DIFFUSION-LIMITED AGGREGATION ON A TREE

by

Martin T. Barlow * — U. British Columbia

Robin Pemantle † — U. Wisconsin

Edwin A. Perkins * — U. British Columbia

Summary

We study the following growth model on a regular d-ary tree. Points at distance n adjacent to the existing subtree are added with probabilities proportional to α^{-n}, where $\alpha < 1$ is a positive real parameter. The heights of these clusters are shown to increase linearly with their total size; this complements known results that show the height increases only logarithmically when $\alpha \geq 1$. Results are obtained using stochastic monotonicity and regeneration results which may be of independent interest. Our motivation comes from two other ways in which the model may be viewed: as a problem in first-passage percolation, and as a version of diffusion-limited aggregation (DLA), adjusted so that "fingering" occurs.

1991 AMS Subject Classifications: Primary. 60K40, 60K30, 60K99

Secondary. 60F05, 60F15, 60K35

Key Words: diffusion-limited aggregation, trees, random clusters, first-passage percolation, fingering, regeneration.

* The author's research was supported by an NSERC operating grant.

† The author's research was supported by NSF grants DMS-9300191 and DMS-9353149 and by a Sloan Foundation Fellowship.

1. Introduction

Consider the following dynamical method for growing a subtree of the regular d-ary tree $\mathbb{B}$. Initially the subtree consists of only the root. Vertices are then added one by one from among those neighbouring the current subtree. The choice of which vertices to add is random, with vertices in generation n chosen with probabilities proportional to α^{-n}, where $\alpha > 0$ is a fixed parameter. Let A_n denote the subtree at step n and let $h(A_n) = \max\{|x| : x \in A_n\}$ denote the maximum height of a vertex in A_n. We are interested in the form of the infinite cluster $A_\infty := \cup_{n=0}^{\infty} A_n$, as well as the behaviour of $h(A_n)$ and related quantities as $n \to \infty$. In this paper we treat the case $\alpha < 1$; the case $\alpha \geq 1$ has already been studied. Our main result, contained in Theorem 6.4 and Corollary 6.5, is as follows.

Theorem 1.1 (Strong Law and CLT). Let $\mathbb{B}$ be the regular d-ary tree and $0 < \alpha < 1$. There exist constants $\mu_0(\alpha, d) \in (0,1)$ and $\sigma^2 = \sigma^2(\alpha, d) > 0$ such that

(a) $\lim_{n\to\infty} n^{-1} h(A_n) = \mu_0(\alpha, d)$ a.s.

(b) $n^{-1/2}(h(A_n) - n\mu_0(\alpha, d)) \xrightarrow{w} N(0, \sigma^2)$ as $n \to \infty$.

This model has arisen in a number of different contexts, and, as we will see below, can be interpreted both as diffusion limited aggregation and as first passage percolation on the tree $\mathbb{B}$.

The case $\alpha = 1$ for binary trees arises in binary search algorithms in computer science and has been studied by numerous authors including Pittel (1984) and Devroye (1986). The case $\alpha = 2$ (again for $d = 2$) arises in an entropy estimation procedure of Ziv (1978) and has been studied by Pittel (1985) and (along with $\alpha > 1$) by Aldous and Shields (1988). A discussion of the cases $\alpha = 1$ and $\alpha = 2$ within the general context of random search trees in computer science may be found in Chapters 2 and 6, respectively, of Mahmoud (1992).

If $\alpha \geq 1$ it is easy to see that $A_\infty := \cup_{n=0}^{\infty} A_n$ is almost surely the entire d-ary tree. Simply look at the vacant node x_0 closest to the root, note that ∂A_n has $(d-1)n + 1$ points of which x_0 is the most likely to get filled, and use the fact that $\prod_{k=n}^{\infty}(1 - ((d-1)k+1)^{-1}) = 0$. (This argument also works for any tree $\mathbb{B}$ of bounded degree.) In the case $\alpha \geq 1$ it is therefore natural to look at the growth rates of both $h(A_n)$ and $l(A_n) = \min\{|x| : x \in A_n^c\}$.

Theorem 1.2. Let $\mathbb{B}$ be a binary tree (i.e. $d = 2$).

(a) (Pittel (1984), Devroye (1986)) If $\alpha = 1$, then, writing $\beta_1 = 4.311...$, $\beta_0 = 0.373...$ for the roots of the equation $\frac{1}{2}\beta e^{(1-\beta)/\beta} = 1$,

$$\lim_{n\to\infty} h(A_n)(\log_2 n)^{-1} = \beta_1, \quad \text{and} \quad \lim_{n\to\infty} l(A_n)(\log_2 n)^{-1} = \beta_0 \text{ a.s.}$$

(b) (Aldous and Shields (1988)) If $\alpha > 1$, then

$$\lim_{n\to\infty} h(A_n)(\log_2 n)^{-1} = \lim_{n\to\infty} l(A_n)(\log_2 n)^{-1} = 1 \text{ a.s.}$$

The case $\alpha = 2$ is also in Pittel (1985, Corollary 1). We remark that Aldous and Shields (1988) also study the form of the cluster between $l(A_n)$ and $h(A_n)$, and that while they do not state explicitly the result for $l(A_n)$ and $\alpha > 1$, this is readily derived using their methods. In addition, they mention the growth dynamics for $\alpha < 1$ as an interesting open problem.

Comparing Theorems 1.1 and 1.2 we see that there is a dramatic phase transition at $\alpha = 1$. For $\alpha > 1$ the process $\{A_n\}$ exhibits the same balanced growth as a deterministic procedure in which vertices are added in lexicographic order. Thus the subtree is essentially as short as possible, $(h(A_n) \approx \log_2 n)$ and $l(A_n)/h(A_n) \to 1$. For $\alpha = 1$ the results of Pittel and Devroye show that fluctuations on a logarithmic scale arise, while for $\alpha < 1$ Theorem 1.1(a) implies that the subtree is essentially as long as possible $(h(A_n) \approx n)$. (In this case it is also easy to see that $A_\infty \neq \mathbb{B}$ – see Theorem 2.1 and Remark 2.2 below).

It would be of interest to study this "phase transition" at $\alpha = 1$ more closely. An understanding of the asymptotics of $\mu_0(\alpha, d)$ as $\alpha \uparrow 1$ would be a step in this direction. The bounds in (6.9) show that $\mu_0(\alpha, d) \geq c_1(d) \exp(-\lambda(d)/(1-\alpha))$, (where $\lambda(d) > 0$), and we suspect that

$$0 < \lim_{\alpha \uparrow 1} (1-\alpha) \log \mu_0(\alpha, d)^{-1} < \infty. \tag{1.1}$$

Our original motivation for studying this model was that it is a version of diffusion-limited aggregation (DLA) on a tree. DLA on $\mathbb{Z}^d$ was introduced by Witten and Sander (1981) to model aggregates of a condensing metal vapour, and since then it has been attracted much interest as a model for various physical phenomena (see for example Vicsek (1989)). DLA is a Markov chain taking values in the space of connected finite subsets of $\mathbb{Z}^d$. Given the current configuration, particles diffuse in from infinity according to a random walk conditioned to hit the "boundary" of the current cluster and attach themselves to the first point they hit which is adjacent to the current cluster. Although the process has quite a simple description, there are very few rigorous mathematical results. If A_n is the cluster at step n, $A_0 = \{0\}$ and $h(A_n) = \max\{|x| : x \in A_n\}$, then one hard open problem is to prove the existence, and find the value of

$$\beta_d = \lim_{n\to\infty} \frac{\log h(A_n)}{\log n}.$$

Kesten (1987, 1990) has shown the lim sup of the above ratio is at most $2/(d+1)$. A conjectured value of β_d which agrees quite well with numerical simulations is $(d+1)(d^2+1)^{-1}$ (Lawler (1991, Sec. 2.6)). There is no rigorous lower bound for β_d aside from the trivial $\beta_d \geq d^{-1}$, and hence no rigorous proof of the existence of the "fingering" ($\beta_d > d^{-1}$) which simulations suggest. For some further surveys of DLA from a mathematical perspective see Lawler (1991) and Barlow (1993).

Analyzing DLA on $\mathbb{Z}^d$ is a hard problem, but the same process on a tree is tractable for two reasons:

(i) There is a simple formula for harmonic measure on the boundary of a cluster (see Lemma 1.3 below).

(ii) The absence of loops in the graph means that disjoint parts of the cluster evolve nearly independently.

While the usual heuristic is that a d-ary tree (or *Bethe Lattice*) will exhibit the limiting behaviour for $\mathbb{Z}^d$ as $d \to \infty$, we do not know to what extent our model is relevant to DLA on $\mathbb{Z}^d$. However, it may be interesting to note that, even though the harmonic measure of any cluster is easy to calculate, the proof of Theorem 1.1(a) is still quite long and hard.

To describe more precisely the connection between the model given above and DLA on a tree, we need some more notation. We begin by presenting the notation used to describe an abstract rooted tree with no leaves. In order to be able to move pieces of trees around, we find it convenient to view an arbitrary rooted tree (with countably many vertices) as a subset of the universal rooted tree $\mathbf{T} = \cup_{n=0}^{\infty}\mathbb{N}^n$ where $\mathbb{N}^0 = \{0\}$ and 0 denotes the root of each tree. (Here $\mathbb{N}$ is taken not to include zero.) First, the notation for $\mathbf{T}$ itself is as follows. If $x \in \mathbf{T}$, $|x| = n$ if and only if $x \in \mathbb{N}^n$ and $x|j = (x_1, \ldots, x_j)$ for $j \le |x|$ (we set $x|0 = 0$). If $x, y \in \mathbf{T}$, let $x \oplus y = (x_1, \ldots, x_{|x|}, y_1, \ldots, y_{|y|})$ ($0 \oplus x = x$) and $x \wedge y = (x_1, \ldots, x_j)$ where j is the largest integer such that $x|j = y|j$ ($x \wedge y = 0$ if $j = 0$). We write $x \le y$ if x is an ancestor of y, i.e. $y = x \oplus z$ for some z, and we say x is the *parent* of y if in addition, $|y| = |x| + 1$, and write $x = \mathsf{par}(y)$. Next, we view an arbitrary ordered (the children of each node come with an order) tree as a subtree of $\mathbf{T}$. Given $m : \mathbf{T} \to \mathbb{N}$, inductively define the associated locally finite rooted tree (with no leaves) $\mathbb{B} \subset \mathbf{T}$ by

$$\mathbb{B}(0) = \{0\}, \quad \mathbb{B}(n+1) = \{x \oplus i : x \in \mathbb{B}(n),\ i \in \{1, 2, \ldots, m(x)\}\}, \quad \mathbb{B} = \cup_{n=0}^{\infty}\mathbb{B}(n). \tag{1.2}$$

Note that $m(x)$ is the number of children of x: the condition $m(x) \ge 1$ is equivalent to the assertion that $\mathbb{B}$ has no leaves. If $m(x) = d$ for all x we obtain the regular d-ary tree. The values of $m(x)$ for $x \notin \mathbb{B}$ are of course irrelevant. Intervals $\mathbb{B}(j, k)$, $\mathbb{B}(j, \infty)$, etc. are defined, for example, by $\mathbb{B}(j, k) = \cup_{j \le n \le k}\mathbb{B}(n)$.

A subset A of $\mathbb{B}$ is a *rooted subtree* if $\mathsf{par}(x) \in A$ for all $x \in A$. S (respectively, S_0) denotes the set of all (respectively, all finite) such subtrees. For $A \in S$, the (external) boundary of A is

$$\partial A = \{x \in A^c : \mathsf{par}(x) \in A\},$$

and the height of A is $h(A) = \sup\{|x| : x \in A\}$.

Assume, until we indicate otherwise, that $\mathbb{B}$ is the regular d-ary tree. Fix $\alpha > 0$. Let Q^{μ} denote the law of the random walk $(Y_0, Y_1, \ldots)$ on $\mathbb{B}$ started from initial distribution μ, with transition probabilities given by $p(x, y) = \alpha/(\alpha+d)$ when y is the parent of x, $p(x, y) = 1/(\alpha+d)$ when y is a child of x and x is not the root, and $p(0, y) = 1/d$ when y is a child of 0. Thus $(Y_n, n \in \mathbb{Z}_+)$ is the random walk on $\mathbb{B}$ obtained by assigning conductances α^{-n} to each edge from generation n to generation $n+1$. Write $\tau(A) = \min\{n \in \mathbb{Z}_+ : Y_n \in A\}$ for the hitting time of $A \subset \mathbb{B}$.

The following lemma describes harmonic measure on the boundary of an arbitrary subtree of $\mathbb{B}$.

Lemma 1.3. Let $A \in S_0$ be a non-empty rooted subtree of $\mathbb{B}$ and fix an $N > h(A)$; let $\mu_N = d^{-N}\sum_{x\in\mathbb{B}(N)}\delta_x$ be the uniform measure on $\mathbb{B}(N)$. If $\alpha_0 = \alpha \wedge d$ then

$$Q^{\mu_N}(Y_{\tau(\partial A)} = x|\tau(\partial A) < \infty) = \alpha_0^{-|x|}\Big(\sum_{y\in\partial A}\alpha_0^{-|y|}\Big)^{-1}, \quad x \in \partial A. \tag{1.3}$$

Proof. Note that $|Y_n|$ is a simple reflecting random walk on $\mathbb{Z}_+$ which moves to the right with probability $p = d\alpha^{-1}(1 + d\alpha^{-1})^{-1}$ and to the left with probability $1 - p$. Therefore if $y \in \mathbb{B}(N)$ and $0 \le j \le N$, then

$$Q^{\delta_y}(|Y_n| = j \text{ for some } n \ge 0) = ((1-p)/p) \wedge 1)^{N-j} = (\alpha_0/d)^{N-j}.$$

If $\tau(\partial A) < \infty$, $x \in \partial A$, and $|Y_0| = N > h(A)$, then $Y_{\tau(\partial A)} = x$ if and only if Y_0 is one of the $d^{N-|x|}$ descendants of x in $\mathbb{B}(N)$. Therefore if $x \in \partial A$, then

$$\begin{aligned} Q^{\mu_N}(Y_{\tau(\partial A)} = x, T(\partial A) < \infty) &= \sum_{y \in \mathbb{B}(N)} d^{-N} 1(x \leq y) Q^{\delta_y}(\tau(\partial A) < \infty) \\ &= d^{N-|x|} d^{-N} (\alpha_0/d)^{N-|x|} \\ &= (\alpha_0/d)^N \alpha_0^{-|x|}, \end{aligned}$$

and the result follows. □

As the hitting distribution in (1.3) is independent of N we arrive at the following definition.

Definition. If $A \in S_0$ and $A \neq \emptyset$, then *harmonic measure on ∂A with parameter $\alpha > 0$* is given by

$$H^\alpha_{\partial A}(x) = \alpha^{-|x|} \Big(\sum_{y \in \partial A} \alpha^{-|y|} \Big)^{-1}, \quad x \in \partial A.$$

Strictly speaking, in view of Lemma 1.3, we should restrict to $\alpha \leq d$, but allowing $\alpha > d$ is harmless, and in any case in this work we are only interested in $\alpha \in (0,1)$.

Definition. *DLA on $\mathbb{B}$ with parameter $\alpha > 0$* is the S_0-valued Markov chain $(A_n, n \in \mathbb{N})$ such that $A_1 = \{0\}$ and $A_{n+1} = A_n \cup \{D_{n+1}\}$, where

$$P(D_{n+1} = x | A_n) = H^\alpha_{\partial A_n}(x), \quad x \in \partial A_n. \tag{1.4}$$

It is clear from (1.4) that (for $0 < \alpha \leq d$) the growth model described at the beginning of the Introduction is exactly DLA on $\mathbb{B}$.

For $\alpha < 1$, $H^\alpha_{\partial A_n}$ favours large $|x|$ values in ∂A_n, for $\alpha > 1$ it favours small $|x|$ values and for $\alpha = 1$ we have the uniform law on ∂A_n (the "Eden model"). From the perspective of classical DLA the case $\alpha < 1$ is of greater interest as it is here that we obtain "fingering", i.e. $h(A_n) \gg (\log_2 n)^\beta$ for some $\beta > 1$. This model has been studied in the physics literature – see Vannimenus et al. (1984). As well as some calculations in the case $\alpha = 1$, this paper describes computer simulations which suggested that $h(A_n) \sim c(\alpha)n$ when $d = 2$ and $\alpha < 1$.

A second motivation is that our model is equivalent to *first-passage percolation.* In first-passage percolation, each edge is assigned a positive random variable, called a passage time, and thought of as the time it takes for water (or information, etc.) to pass from one endpoint to the other. To each vertex x one associates a time $T(x)$, which is the minimum over paths from the root to x of the sum of the transit times. This represents the time before x gets wet when the root is a source of water from time 0 onward. The model has been widely studied on $\mathbb{Z}^d$, and some quite nontrivial results have been obtained on trees as well (see Bramson (1978) and Pemantle and Peres (1994) for two examples). On a tree, it is natural to rescale edges so that ones more distant from the root are shorter. For example, the tributaries of a river system behave this way; see also the limit trees of Aldous (1991), whose edge lengths decay exponentially in the distance from the root. While the edge random variables in first passage percolation may, of course, have any distribution on $[0,\infty)$, the simplest case is that of exponentials. In this case the lack of memory property of the exponential distribution implies that the cluster $C(t)$

of vertices which are wet at time t will be a continuous time Markov process on the space of rooted subtrees S.

In Section 2, we prove that DLA on $\mathbb{B}$ arises from first-passage percolation when the passage times are independent exponential random variables with mean α^n (for an edge from $\mathbb{B}(n)$ to $\mathbb{B}(n+1)$). Throughout the paper, we will use these various viewpoints interchangeably; thus we usually refer to the subtree as a cluster or as the DLA, but keep the notion of passage times in the foreground as well, and in fact, most of our analysis takes place in the continuous-time setting of first-passage percolation.

The organization of the paper is as follows. In Section 2 we show that DLA can be embedded in first-passage percolation, and give a number of general results on the model. Theorem 2.1 proves that either $A_\infty = \mathbb{B}$ a.s. or else the cluster A_∞ has unique infinite line of descent (the "backbone"). In the latter case it is possible to decompose A_∞ into the backbone plus a sequence of finite clusters attached to successive points of the backbone. These clusters are i.i.d. given the "percolation times" along the backbone (Theorem 2.5). If $\alpha < 1/d$, a simple law of large numbers argument, based on estimates of cluster sizes in Lemma 2.4, shows that $\liminf_{n\to\infty} n^{-1} h(A_n) \geq c(\alpha, d) > 0$ a.s. (Theorem 2.6). This result exhibits our basic approach while avoiding the technical problems involved when considering the case when α is close to 1. Finally, Theorem 2.8 gives a general strong Markov property.

The hard work is in Sections 3 and 4. Section 4 contains the key estimates on the sizes of the finite clusters, giving bounds in L^1 (Theorem 4.4) and in L^2 (Theorem 4.6). To handle the dependence which arises in these proofs it is necessary along the way to prove stochastic monotonicity results for the conditional distribution of the clusters given the backbone times, and this is done in Section 3 (Lemmas 3.1 and 3.3). Section 5 shows that the percolation times along the backbone form a Markov chain, and that this chain converges exponentially fast to its stationary measure (Theorem 5.2 and Corollary 5.7). This paves the way for a Strong Law of Large Numbers (Theorem 6.1) and Central Limit Theorem (Theorem 6.2) holding for a general class of functionals of the finite clusters. In Section 7 we find a sequence of *regeneration times* for cluster, which allow it to be decomposed into i.i.d. pieces. Combining these results with the theorems in Section 6 we complete the proof of Theorem 1.1. Finally, Section 8 adds some remarks on the asymptotics of the growth dynamics as $\alpha \uparrow 1$, and on a related particle system.

In most of this work we will only be concerned with the regular d-ary tree, and the DLA model described above. However, we may note that most of the results of Sections 2 and 3 hold for a more general model. First, we may consider a general locally finite rooted tree $\mathbb{B}$, defined by (1.2). Secondly, we can fix a function $f : \mathbb{B} \to (0, \infty)$, and take the passage time between $\mathsf{par}(x)$ and x to be exponential with mean $f(x)$. We can then consider a process A_n which has growth probabilities given by

$$P(A_{n+1} = A_n \cup \{x\} | A_n) = f(x)^{-1} \Big(\sum_{y \in \partial A_n} f(y)^{-1} \Big)^{-1}.$$

While these extensions involve no new ideas, describing the tree and the process in this more general setup does require some quite cumbersome notation. Apart from Theorem 2.1, and some simple estimates on cluster size in Lemma 2.4, we will therefore restrict our proofs to the case of a regular tree and $f(x) = \alpha^{|x|}$. An earlier version of this paper, which treats the general case in Sections 2 and 3, is available by anonymous ftp from `ftp.math.ubc.ca` (directory `pub/barlow`.)

A general notational convention is that $c_{i.j}$ denotes a globally defined constant introduced in Section i, whereas c, c', $c(\alpha), \ldots$ may represent different values in different lines. Dependence of $c_{i.j}$ on parameters such as (α, d) will at times be suppressed if there is no ambiguity. The integral of a function φ with respect to a measure μ (or $\mathbb{P}$) is written $\mu(\varphi)$ (or $\mathbb{P}(\varphi)$).

2. The Continuous Time Model for General Trees

Let $\mathbb{B}$ be a general locally finite rooted tree with no leaves and let $f : \mathbb{B} \to (0, \infty)$. Consider the S_0-valued Markov chain $\{A_n\}$ such that $A_1 = \{0\}$ and $A_{n+1} = A_n \cup \{D_{n+1}\}$ where

$$\mathbb{P}(D_{n+1} = x | A_n) = \begin{cases} f(x)^{-1}(\sum_{y \in \partial A_n} f(y)^{-1})^{-1} & \text{for } \; x \in \partial A_n \\ 0 & \text{for } \; x \notin \partial A_n. \end{cases} \tag{2.1}$$

Note that $f(x) = \alpha^{|x|}$ on the regular d-ary tree gives the cluster dynamics of the previous section. Set $A_\infty = \cup_{n=0}^{\infty} A_n$. We now embed $(A_n,\, n \geq 0)$ in a continuous time process.

On some complete $(\Omega, \mathcal{F}, \mathbb{P})$ let $\{U_x : x \in \mathbb{B}\}$ be i.i.d. exponential random variables with mean one, and define $T(x) = \sum_{y \leq x} f(y) U_y$. (We set $T(\mathsf{par}(0)) = 0$). In terms of first passage percolation, if $f(y)U_y$ is the time for liquid to percolate from $\mathsf{par}(y)$ to y and $f(0)U_0$ is the time it takes for 0 to get wet, then $T(x)$ is the time it takes to percolate to x. Let $C(t) = \{x : T(x) \leq t\}$ be the nodes which are wet at time t and define $a(t) : C(t) \to [0, \infty)$ by $a(t)(x) = t - T(x)$. Let $\mathcal{S} = \{(C, a) : C \in S,\; a : C \to [0, \infty)\}$ and let $\mathcal{S}_0$ denote the same set with S_0 in place of S. If Δ is added to $\mathbb{R}$ as a discrete point, define $\Pi_x : \mathcal{S} \to \mathbb{R} \cup \{\Delta\}$ for x in $\mathbb{B}$ by $\Pi_x(C, a) = a(x) 1(x \in C) + \Delta 1(x \notin C)$. Give $\mathcal{S}$ (respectively $\mathcal{S}_0$) the smallest σ-field $\mathcal{F}(\mathcal{S})$ (respectively, $\mathcal{F}(\mathcal{S}_0)$) generated by the maps $\{\Pi_x : x \in \mathbb{B}\}$ (respectively their restrictions to $\mathcal{S}_0$). The process $Y(t) = (C(t), a(t))$ is an $\mathcal{S}$-valued process. $\mathcal{G}_t^0 = \sigma(T(x) \wedge t, x \in \mathbb{B})$ and $\mathcal{G}_t = \mathcal{G}_{t+}^0$ are filtrations on $(\Omega, \mathcal{F})$ and $T(x)$ is a $(\mathcal{G}_t)$ - stopping time for all x in $\mathbb{B}$ because $\{T(x) < t\} = \{T(x) \wedge t < t\} \in \mathcal{G}_t^0$. Clearly Y is $(\mathcal{G}_t)$-adapted because $\Pi_x(Y(t))$ takes on the value $t - T(x)$ on $\{T(x) \leq t\}$ and the value Δ elsewhere.

Let $\mathsf{card}_n = \inf\{t : \#C(t) = n\}$: thus $\{\mathsf{card}_n\}$ is a sequence of a.s. finite $(\mathcal{G}_t)$-stopping times. Clearly $C(\mathsf{card}_{n+1}) = C(\mathsf{card}_n) \cup \{D_{n+1}\}$ where $D_{n+1} \in \partial C(\mathsf{card}_n)$. Using the lack of memory property of the exponential, it is easy to see, as in Section 1 of Aldous and Shields (1988), that the process $A_n = C(\mathsf{card}_n)$ satisfies (2.1). We therefore may, and shall, take $A_n = C(\mathsf{card}_n)$ throughout this work.

Notation. Let $T(\infty) = \lim_{n \to \infty} \mathsf{card}_n \leq \infty$ be the time to percolate to ∞. It is clear that $A_\infty = C(T(\infty)-) := \{x : T(x) < T(\infty)\}$. Set $T(x, y) = T(y) - T(x)$ if $x \leq y$ and $T(x, y) = \infty$ otherwise. Let

$$T(x, \infty) = \lim_{n \to \infty} \min\{T(x, y) : y \geq x, y \in \mathbb{B}(n)\},$$

$$T(x-, \infty) = \lim_{n \to \infty} \min\{T(\mathsf{par}(x), y) : y \geq x, y \in \mathbb{B}(n)\} = f(x) U_x + T(x, \infty).$$

Thus $T(x, \infty)$ is the time to percolate from x to ∞, while $T(x-, \infty)$ is the time to percolate from $\mathsf{par}(x)$ to ∞ through x. If $x \in \mathbb{B}$, let $[x, \infty)$ denote $\{y \in \mathbb{B} : y \geq x\}$ and similarly define (x, ∞), $[0, x]$, etc. For $G \subset \mathbb{B}$ let $\mathcal{F}_G = \sigma(U_x : x \in G)$ and define $\mathcal{F}_n = \sigma(U_x : |x| \leq n)$.

The next result is due to Brennan and Durrett (1986, Sec. 3), but we include a proof because the settings are a little different.

Theorem 2.1. (a) $\mathbb{P}(T(\infty) < \infty) = 0$ or 1.

(b) If $\mathbb{P}(T(\infty) = \infty) = 1$ then $A_\infty = \mathbb{B}$ a.s.

(c) If $\mathbb{P}(T(\infty) < \infty) = 1$ then there is a.s. a unique infinite line of descent in A_∞, i.e., there is a unique sequence $\{\mathsf{spine}_n, n \geq 1\}$ such that

$$(2.2) \qquad \mathsf{spine}_n \in \mathbb{B}(n) \cap A_\infty, \qquad \mathsf{spine}_n = \mathsf{par}(\mathsf{spine}_{n+1}), \quad \text{for all } n \geq 1.$$

Remark 2.2. If $\mathbb{B}$ is a regular d-ary tree and $f(x) = g(|x|)$ then by comparing $T(\infty)$ with the time to percolate to ∞ along a fixed path we see that

$$\sum_{n=0}^{\infty} g(n) < \infty \quad \text{implies} \quad \mathbb{P}(T(\infty) < \infty) = 1.$$

Rather surprisingly, if g is monotone then the converse is also true: $\sum g(n) = \infty$ implies $\mathbb{P}(T(\infty) < \infty) = 0$ – see Pemantle and Peres (1994).

Proof. (a) The event $\{T(\infty) = \infty\} = \bigcap_{x \in \mathbb{B}(n)} \{T(x, \infty) = \infty\}$ is clearly in $\mathcal{T} = \bigcap_{n=1}^{\infty} \mathcal{F}_{\mathbb{B}(n,\infty)}$. The latter is a $0-1$ σ-field by the Kolmogorov $0-1$ theorem.

(b) Clearly $T(\infty) = \infty$ a.s. implies $A_\infty = \{x : T(x) < T(\infty)\} = \mathbb{B}$ a.s.

(c) Fix $j \in \mathbb{N}$. To show that for each j there is a unique choice of spine_j it is enough to show that

$$(2.3) \qquad \text{There exists a unique } X^j \in \mathbb{B}(j) \text{ such that } A_\infty \cap [X^j, \infty) \text{ is infinite,}$$

Assume (2.3) fails. Then there are distinct vertices x_1, x_2 in $\mathbb{B}(j)$ such that with positive probability $A_\infty \cap [x_i, \infty)$ is infinite for both $i = 1, 2$. It follows that, with positive probability,

$$T(\mathsf{par}(x_1)) + T(x_1-, \infty) = T(\infty) = T(\mathsf{par}(x_2)) + T(x_2-, \infty).$$

Therefore conditional on $\mathcal{F}_{j-1}$, $T(x_1-, \infty) - T(x_2-, \infty)$ has an atom at the $\mathcal{F}_{j-1}$-measurable point $T(\mathsf{par}(x_2)) - T(\mathsf{par}(x_1)) \in \mathbb{R}$. On the other hand $T(x_1-, \infty)$ and $T(x_2-, \infty)$ are independent random variables with densities (because $T(x_i-, \infty) = U_{x_i} f(x_i) + T(x_i, \infty)$ with $T(x_i, \infty)$ independent of the exponential random variable U_{x_i}) and are jointly independent of $\mathcal{F}_{j-1}$. Therefore $T(x_1-, \infty) - T(x_2-, \infty)$ has a conditional density given $\mathcal{F}_{j-1}$. This contradiction completes the proof. $\square$

Similar arguments later will require the following quantitative estimate on densities of sums of the variables U_x. The elementary proof is omitted.

Notation. Let $s_p(\alpha) = \prod_{i=1}^{p} (1 - \alpha^i)^{-1}$, $\alpha \in (0, 1)$, $p \in \mathbb{Z}_+ \cup \{\infty\}$, and write $s(\alpha) = s_\infty(\alpha)$.

Lemma 2.3. If $\{U_i, \ i \in \mathbb{Z}_+\}$ are i.i.d. exponential r.v. with mean 1, and $\alpha \in (0, 1)$, then for $p \in \mathbb{Z}_+ \cup \{\infty\}$, $S_p = \sum_{i=0}^{p} \alpha^i U_i$ has a density $u_p(t) \leq s_p(\alpha) e^{-t}$.

Notation. If $x \in \mathbb{B}$, then $\mathbb{B}_x = \{y : x \oplus y \in \mathbb{B}\}$ is the locally finite rooted tree of descendants of x (properly translated) and $\mathbb{B}_x(n) = \{y \in \mathbb{B}_x : |y| = n\}$ is the n^{th} generation of the nodes in $\mathbb{B}_x$. If $x \neq 0$ and $y \in \mathbb{B}_x$, let

$$T^{(x)}(y) = f(x)^{-1} T(\mathsf{par}(x), x \oplus y)$$

be the *rescaled* percolation times for the tree $\mathbb{B}_x$. Let

$$\begin{aligned} T(n) &= \inf\{T(x) : x \in \mathbb{B}(n)\} \quad \text{and} \ , \\ T^{(x)}(n) &= \inf\{T^{(x)}(y) : y \in \mathbb{B}_x(n)\} \end{aligned}$$

be the percolation and rescaled percolation times, respectively, to the n^{th} generation in $\mathbb{B}$ and $\mathbb{B}_x$. Finally, let

$$T^{(x)}(\infty) = T(x-, \infty) f(x)^{-1}$$

be the rescaled time to percolate to infinity in $\mathbb{B}_x$. Note that if $\mathbb{B}$ is a regular d-ary tree and $f(x) = \alpha^{|x|}$, then $\mathbb{B}_x = \mathbb{B}$ for all x, and $\{T^{(x)}(y) : y \in \mathbb{B}\}$ has the same distribution as $\{T(y) : y \in \mathbb{B}\}$.

We now derive upper bounds on the L^1 and L^2 norms of the cluster size at a fixed time, conditioned on being finite. These bounds are crude but their proofs are fairly simple, and the bounds are good enough to enable us to prove that $\liminf n^{-1} h(A_n) > 0$ in the case when $f(x) = \alpha^{|x|}$ and $\alpha \sup_x m(x) < 1$ (c.f. Theorem 2.6 and Remark 2.7). In Sections 3 and 4 we will have to work much harder to obtain better bounds (e.g. Theorem 4.4) which lead to the linear growth of $h(A_n)$ for a regular d-ary tree and all $\alpha \in (0, 1)$.

Lemma 2.4. Let $f(x) = \alpha^{|x|}$ for some $\alpha \in (0, 1)$.

(a) $\mathbb{P}(\#C(t) | T(\infty) > t) \leq c(\alpha) \sum_{n=0}^{\infty} \#\mathbb{B}(n) \alpha^n$ for all $t \geq 0$.

(b) $\mathbb{P}(\#C(t)^2 | T(\infty) > t) \leq c(\alpha) \sum_{z \in \mathbb{B}} \alpha^{|z|} \left(\sum_{m=0}^{\infty} \alpha^m \#\mathbb{B}_z(m) \right)^2$ for all $t \geq 0$.

Proof. (a) If $x \in \mathbb{B}, t \geq 0$ and S_∞ is as in Lemma 2.3, then

$$\begin{aligned} \mathbb{P}(T(x) \leq t < T(\infty)) &\leq \mathbb{P}(\mathbb{P}(t - T(x, \infty) < T(x) \leq t | T(x, \infty))) \\ &\leq s(\alpha) \mathbb{P}(e^{-(t - T(x,\infty))} T(x, \infty)) \qquad \text{(Lemma 2.3)} \\ &\leq s(\alpha) e^{-t} \mathbb{P}(\exp(\alpha^{|x|+1} S_\infty) \alpha^{|x|+1} S_\infty) \\ &\leq (s(\alpha)/(1-\alpha))^2 e^{-t} \alpha^{|x|+1} \qquad \text{(Lemma 2.3).} \end{aligned} \tag{2.4}$$

Using the fact that $\mathbb{P}(T(\infty) > t) \geq \mathbb{P}(U_0 > t) = e^{-t}$, we conclude that

$$\begin{aligned} \mathbb{P}(\#C(t) | T(\infty) > t) &= \sum_{x \in \mathbb{B}} \mathbb{P}(T(x) \leq t < T(\infty)) \mathbb{P}(T(\infty) > t)^{-1} \\ &\leq (s(\alpha)/(1-\alpha))^2 \sum_{n=0}^{\infty} \#\mathbb{B}(n) \alpha^{n+1}. \end{aligned}$$

(b) Let $x_i \in \mathbb{B}$, $i = 1, 2$ be both distinct from $x_1 \wedge x_2$. Let $p = |x_1 \wedge x_2|$, $x_i' = x_i|(p+1)$, and write $x_i = x_i' \oplus y_i$. Then

$$
\begin{aligned}
&\mathbb{P}(T(x_1) \vee T(x_2) \le t < T(\infty)) \\
&\quad \le \mathbb{P}\big(T(x_1 \wedge x_2) + \alpha^{p+1} T^{(x_i')}(y_i) \le t < T(x_1 \wedge x_2) + \alpha^{p+1} T^{(x_i')}(\infty), i = 1, 2\big) \\
&\quad = \mathbb{P}\big(\prod_{i=1}^{2} \mathbb{P}(T^{(x_i')}(y_i) \le (t - T(x_1 \wedge x_2))\alpha^{-p-1} < T^{(x_i')}(\infty) | \mathcal{F}_p)\big) \qquad (2.5) \\
&\quad \le c(\alpha) \mathbb{P}(1(T(x_1 \wedge x_2) < t) \exp(-2\alpha^{-p-1}(t - T(x_1 \wedge x_2)))\alpha^{|x_1| + |x_2| - 2p}) \qquad \text{(by (2.4))} \\
&\quad \le c(\alpha) e^{-t} \alpha^{|x_1| + |x_2| - p}.
\end{aligned}
$$

In the last line we again used Lemma 2.3 with $S_p = T(x_1 \wedge x_2)$. If either x_1 or x_2 equals $x = x_1 \wedge x_2$ the above inequality is clear from (2.4). Decomposing the sum over x_1, x_2 in $\mathbb{B}$ according to the value of $z = x_1 \wedge x_2$, we obtain

$$
\begin{aligned}
\mathbb{P}(\#C(t)^2 1(T(\infty) > t)) &= \sum_{x_1, x_2 \in \mathbb{B}} \mathbb{P}(T(x_1) \vee T(x_2) \le t < T(\infty)) \\
&\le \sum_{z \in \mathbb{B}} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \#\mathbb{B}_z(m) \#\mathbb{B}_z(n) c(\alpha) e^{-t} \alpha^{m+n+|z|} \qquad \text{(by (2.5))} \\
&\le c(\alpha) \mathbb{P}(T(\infty) > t) \Big(\sum_{z \in \mathbb{B}} \alpha^{|z|} \big(\sum_{m=0}^{\infty} \#\mathbb{B}_z(m) \alpha^m \big)^2 \Big).
\end{aligned}
$$

□

These bounds are far from optimal. The very first inequality in the proof of (a) ignores the critical fact that $T(x) < T(\infty)$ only holds for a small proportion of vertices x in $\mathbb{B}(n)$. This leads to the more restrictive conditions on α in Theorem 2.6 below.

In the remainder of the paper we will assume $\mathbb{B}$ is a regular d-ary tree and $f(x) = \alpha^{|x|}$ with $\alpha \in (0, 1)$. By Remark 2.2 we have $T(\infty) < \infty$ a.s. The unique infinite line of descent in A_∞, defined by (2.2), is called $\mathsf{Spine} = \{\mathsf{spine}_n, n \in \mathbb{N}\}$, or the "backbone" of the cluster A_∞. The cluster A_∞ may be partitioned into Spine and a collection of disjoint finite clusters which branch off Spine.

Notation. For $n \in \mathbb{Z}_+$ and $x \in \mathbb{B}(n+1)$, let $\{\mathsf{sib}_j(x) : j < d\}$ denote the siblings of x, i.e., the points in $\{(x|n) \oplus i : i \ne x_{n+1}, i \le d\}$ in increasing order of i. We write $e_{n,j}$ for $\mathsf{sib}_j(\mathsf{spine}_{n+1})$ and for $n \in \mathbb{Z}_+$ and $1 \le j < d$, set

$$\mathsf{Clust}_{n,j} = \{x \in \mathbb{B} : e_{n,j} \oplus x \in A_\infty\}.$$

Thus $\{\mathsf{Clust}_{n,j} : j < d\}$ are the (possibly empty) clusters which branch off the backbone in generation n. Define $a_{n,j} : \mathsf{Clust}_{n,j} \to [0, \infty)$ by

$$a_{n,j}(x) = (T(\infty) - T(e_{n,j} \oplus x))\alpha^{-n-1},$$

and let $Y_{n,j} = (\mathsf{Clust}_{n,j}, a_{n,j})$, which is almost surely in $\mathcal{S}_0$ by Theorem 2.1. Let $W_n = T(\mathsf{spine}_n, \infty)\alpha^{-n-1}$ denote the normalized time to percolate along the backbone from generation n to infinity. Let $\mathbf{W}$ denote the sequence $\{W_n : n = 0, 1, 2, \ldots\}$. For each $t \ge 0$, define a law ν_t on $\mathcal{S}_0$ by

$$\nu_t(\cdot) = \mathbb{P}(Y(t) \in \cdot | T(\infty) > t).$$

Theorem 2.5. Conditional on $(\mathsf{Spine}, \mathbf{W})$, the collection $\{Y_{n,j} : n \in \mathbb{Z}_+, j < d\}$ is independent as n and j vary, and the joint conditional distribution of each $Y_{n,j}$ is given by

$$\mathbb{P}(Y_{n,j} \in \cdot | \mathsf{Spine}, \mathbf{W})(\omega) = \nu_{W_n(\omega)} .$$

The intuitive explanation of this is that the only information passed to the subtree beneath $e_{n,j}$ by conditioning on the backbone is that the time to infinity inside this subtree has to be greater than the time along the backbone.

Proof. Choose $N \in \mathbb{N}$ and $x \in \mathbb{B}(N)$. Consider an event $D = \bigcap_{0 \le n < N, 1 \le j < d} D_{n,j}$, where $D_{n,j}$ is of the form

$$D_{n,j} = \{Y_{n,j} = (b_{n,j}, a) \text{ for some } a \in F_{n,j}\},$$

where $b_{n,j} \in S_0$ and $F_{n,j}$ is the set of nonnegative functions φ on $b_{n,j}$ such that $\varphi(y) \in F^{n,j,y}$ for each y, where $\{F^{n,j,y}\}$ is a specified collection of measurable sets. This class of events generates $\sigma(Y_{n,j} : n \in \mathbb{Z}_+, j < d)$ and is closed under finite intersection, so it suffices to show that

$$\mathbb{P}(D | \mathsf{Spine}, \mathbf{W}) = \prod \nu_{W_n}(D_{n,j}) .$$

If $V_n = U_{\mathsf{spine}_n}$, then clearly $\sigma(\mathbf{W}) = \sigma(V_n : n \in \mathbb{N})$. For $n < N$, let $R_n = T(x|n, x) + T(x, \infty)$ be the time to percolate from $x|n$ to infinity through x. Note that $\mathsf{spine}_N = x$ if and only if for each $0 \le n \le N-1$, the fastest route from $x|n$ to infinity is through x, that is,

$$\{\mathsf{spine}_N = x\} = \bigcap_{0 \le n < N} \bigcap_{1 \le j < d} \{R_n < T(\mathsf{sib}_j(x|(n+1))-, \infty)\} . \tag{2.6}$$

Note also that on $\{\mathsf{spine}_N = x\}$ we have

$$\mathsf{Clust}_{n,j} = \{y' : T^{(\mathsf{sib}_j(x|n+1))}(y') < R_n \alpha^{-n-1}\}$$

and for $y \in \mathsf{Clust}_{n,j}$,

$$\begin{aligned} a_{n,j}(y) &= (T(\infty) - T(\mathsf{sib}_j(x|n+1) \oplus y))\alpha^{-n-1} \\ &= (T(x|n) + R_n - T(\mathsf{sib}_j(x|n+1) \oplus y))\alpha^{-n-1} \\ &= (R_n - T(x|n, \mathsf{sib}_j(x|n+1) \oplus y))\alpha^{-n-1} \\ &= R_n \alpha^{-n-1} - T^{(\mathsf{sib}_j(x|n+1))}(y) . \end{aligned}$$

If B_n are measurable subsets of the positive reals, then using (2.6) and the above, we have

$$\begin{aligned} &\mathbb{P}(V_n \in B_n \text{ for } 1 \le n \le N, \mathsf{spine}_N = x, Y_{n,j} \in D_{n,j} \text{ for } 0 \le n < N, 1 \le j < d) \\ =&\mathbb{P}\Big(1(U_{x|n} \in B_n \text{ for } 1 \le n \le N) \prod_{n=0}^{N-1} \prod_{j=1}^{d-1} \{1(R_n < T(\mathsf{sib}_j(x|n+1)-, \infty)) \\ &\times 1(\{y : T^{(\mathsf{sib}_j(x|n+1))}(y) < R_n \alpha^{-n-1}\} = b_{n,j}) \\ &\times 1(R_n \alpha^{-n-1} - T^{(\mathsf{sib}_j(x|n+1))}(y) \in F^{n,j,y} \text{ for all } y \in b_{n,j})\}\Big) . \end{aligned}$$

Let $\mathcal{G}(x) = \mathcal{F}_{[0,x]} \vee \mathcal{F}_{[x,\infty)}$. Observe that R_n is $\mathcal{G}(x)$-measurable and that $T^{(\mathsf{sib}_j(x|n+1))}(y)$ and $T(\mathsf{sib}_j(x|n+1)-,\infty)$ are $\mathcal{F}_{[\mathsf{sib}_j(x|n+1),\infty)}$- measurable. The collection of σ-fields $\mathcal{F}_{[\mathsf{sib}_j(x|n+1),\infty)}$ for $j < d$ and $0 \le n < N$, together with $\mathcal{G}(x)$, are all mutually independent. Condition the above integrand with respect to $\mathcal{G}(x)$ to see that it equals

$$\begin{aligned}\int 1(U_{x|n}(\omega) \in B_n, 1 \le n \le N) \prod_{n=0}^{N-1} \prod_{j=1}^{d-1} &\Big\{\mathbb{P}\big(R_n(\omega) < T(\mathsf{sib}_j(x|n+1)-,\infty)\big) \\ \times\mathbb{P}\Big(&\{y : T^{(\mathsf{sib}_j(x|n+1))}(y) < R_n(\omega)\alpha^{-n-1}\} = b_{n,j}, \\ &R_n(\omega)\alpha^{-n-1} - T^{(\mathsf{sib}_j(x|n+1))}(y) \in F^{n,j,y} \text{ for all } y \in b_{n,j} \\ &\Big| R_n(\omega)\alpha^{-n-1} < T^{(\mathsf{sib}_j(x|n+1))}(\infty)\Big)\Big\}\, d\mathbb{P}(\omega)\,.\end{aligned}$$

By (2.6) the product of the first factors in the curly braces equals $\mathbb{P}(\mathsf{spine}_N = x|\mathcal{G}(x))(\omega)$. On $\{\mathsf{spine}_N = x\}$ we have $R_n\alpha^{-n-1} = W_n$ for $n < N$, and so the above leads to

$$\begin{aligned}&\mathbb{P}(V_n \in B_n \text{ for } 1 \le n \le N, \mathsf{spine}_N = x, D) \\ &\quad = \int 1(\mathsf{spine}_N = x, U_{\mathsf{spine}_n} \in B_n \text{ for } 1 \le n \le N) \\ &\qquad \times \prod_{n=0}^{N-1} \prod_{j=1}^{d-1} \mathbb{P}\big(\{y : T^{(\mathsf{sib}_j(x|n+1))}(y) < W_n(\omega)\} = b_{n,j}, \\ &\qquad W_n(\omega) - T^{(\mathsf{sib}_j(x|n+1))}(y) \in F^{n,j,y} \text{ for all } y \in b_{n,j} \\ &\qquad |T^{(\mathsf{sib}_j(x|n+1))}(\infty) > W_n(\omega)\big)\, d\mathbb{P}(\omega) \\ &\quad = \int 1(V_n \in B_n \text{ for } 1 \le n \le N, \mathsf{spine}_N = x) \prod_{n=0}^{N-1} \prod_{j=1}^{d-1} \nu_{W_n(\omega)}(D_{n,j})\, d\mathbb{P}(\omega)\,.\end{aligned}$$

We have used the equivalence in law of $\{T^{(x)}(y) : y \in \mathbb{B}\}$ and $\{T(y) : y \in \mathbb{B}\}$ in the last line. $\square$

The above decomposition and the L^1 and L^2 bounds in Lemma 2.4 allow us to use the law of large numbers to establish linear growth of $h(A_n)$ for sufficiently small α. The proof illustrates the basic approach we will take in Section 6 to obtain the result for all $\alpha < 1$.

Theorem 2.6. Assume $\alpha \in (0, d^{-1})$. Then

$$\liminf_{n\to\infty} n^{-1} h(A_n) \ge c_{2.1}(\alpha, d) > 0 \quad \text{a.s.}$$

Proof. If $\sigma(n) = \inf\{k : h(A_k) = n\}$, the result is equivalent to

$$\limsup_{n\to\infty} n^{-1}\sigma(n) \le c < \infty \quad \text{a.s.} \tag{2.7}$$

Decompose $C(T(n))$ into the backbone vertices in $C(T(n))$ (there are at most $n+1$) and the portions of the clusters $\{e_{k,i} \oplus \mathsf{Clust}_{k,i},\ k < n,\ i < d\}$ which are contained in $C(T(n))$. This shows that

$$\sigma(n)/n = \frac{\#C(T(n))}{n} \le \frac{n+1}{n} + n^{-1}\sum_{k=0}^{n-1}\left(\sum_{i=1}^{d-1}\#\mathsf{Clust}_{k,i}\right). \tag{2.8}$$

Conditional on $(\mathbf{W}, \mathsf{Spine})$, $\{\#\mathsf{Clust}_{k,i} : k \ge 0,\ i < d\}$ are independent random variables (Theorem 2.5) such that

$$\begin{aligned}
\mathbb{P}(\#\mathsf{Clust}_{k,i}^2|\mathsf{Spine}, \mathbf{W})(\omega) &= \nu_{W_n(\omega)}((\#C)^2) \\
&\le c(\alpha)\sum_{z\in\mathbb{B}}\alpha^{|z|}(\sum_{m=0}^{\infty}\alpha^m d^m)^2 && \text{(Lemma 2.4)} \\
&\le c(\alpha)(1-\alpha d)^{-3}.
\end{aligned}$$

Therefore $\mu_{k,i}(\omega) = \mathbb{P}(\#\mathsf{Clust}_{k,i}|\mathsf{Spine}, \mathbf{W})(\omega)$ is also uniformly bounded by $c(\alpha, d)$ say, and the strong law of large numbers (applied conditionally) implies that

$$\limsup_{n\to\infty} n^{-1}\sum_{k=0}^{n-1}\sum_{i=1}^{d-1}(\#\mathsf{Clust}_{k,i} - \mu_{k,i}) = 0 \quad \text{a.s.}$$

Use this in (2.8) to see that $\limsup_{n\to\infty}\sigma(n)n^{-1} \le 1 + dc(\alpha, d)$ a.s., thus proving (2.7). ☐

Remark 2.7. In the more general setting of Theorem 2.1 (general $\mathbb{B}$ and f), it is just as easy to decompose A_∞ into Spine (the backbone) and clusters $\{\mathsf{Clust}_{n,j} : j < m(\mathsf{spine}_n), n \in \mathbb{Z}_+\}$ which branch off the backbone in generation n. With only notational changes in the proof, it is then possible to derive an analogue of Theorem 2.5. The lack of scaling means the conditional law of $Y_{n,j}$ will also depend on the tree $\mathbb{B}_{e_{n,j}(\omega)}$ and the appropriately shifted and rescaled version of f. One can then show that Theorem 2.6 remains valid if $\mathbb{B}$ is a rooted tree with no leaves such that $m(x) \le d$ for all x and $f(x) = \alpha^{|x|}$ for some $\alpha \in (0, d^{-1})$. The proof is the same.

Notation. The wide sense past up to x is defined by $\mathcal{E}_x = \mathcal{F}_{(x,\infty)^c}$ and we let $\mathcal{E}_{x-} = \mathcal{F}_{[x,\infty)^c}$.

The following strong Markov property will be used in Section 6. It states that if you stop at a stopping time when $\mathsf{par}(x)$ has been reached, but x has not been reached, then the remaining times to hit vertices from the subtree rooted at x, rescaled, are equal in law to the original system of hitting times. As might be expected, the main difficulty in obtaining this strong Markov property is in getting the statement right.

Theorem 2.8. If σ is an a.s.-finite $(\mathcal{G}_t)$-stopping time and $x \in \mathbb{B}$, then for each measurable $B \in [0,\infty)^{\mathbb{B}}$,

$$\mathbb{P}\Big(((T(x\oplus y) - \sigma)\alpha^{-|x|} : y \in \mathbb{B}) \in B|\mathcal{G}_\sigma \vee \mathcal{E}_{x-}\Big) = \mathbb{P}\big((T(y) : y \in \mathbb{B}) \in B\big) \tag{2.9}$$

almost surely on $\{x \in \partial C(\sigma)\}$.

Proof. Note that $\{x \in \partial C(\sigma)\} = \{T(\mathsf{par}(x)) \le \sigma < T(x)\}$ is in $\mathcal{G}_\sigma$. (Recall that $T(\mathsf{par}(0)) = 0$). Assume first that $\sigma \ge 0$ is constant and consider (2.9) with $\mathcal{G}_t^o$ in place of $\mathcal{G}_t$. Let ν be the exponential law with mean 1 and let ν^F denote product measure on $[0,\infty)^F$. Define $\tilde U_0 = (T(x)-\sigma)\alpha^{-|x|}$ and $\tilde U_y = U_{x\oplus y}$ for $y \in \mathbb{B}\setminus\{0\}$. We claim that

$$\mathbb{P}((\tilde U_y : y \in \mathbb{B}) \in \cdot | \mathcal{G}^o_\sigma \vee \mathcal{E}_{x-}) = \nu^{\mathbb{B}}(\cdot) \text{ a.s. on the event } \{x \in \partial C(\sigma)\}\,. \tag{2.10}$$

Indeed, $\mathcal{G}^o_\sigma \vee \mathcal{E}_{x-} = \sigma(T(y)\wedge\sigma : y \geq x) \vee \mathcal{E}_{x-} \vee \sigma(1(T(x) > \sigma))$, and $T(y)\wedge\sigma = \sigma$ for all $y \geq x$ on $\{x \in \partial C(\sigma)\}$; the latter event is measurable with respect to $\mathcal{E}_{x-} \vee \sigma(1(T(x)>\sigma))$, and so (2.10) is equivalent to

$$\mathbb{P}((\tilde U_y : y \in \mathbb{B}) \in \cdot | \mathcal{E}_{x-} \vee \sigma(1(T(x) > \sigma))) = \nu^{\mathbb{B}}(\cdot) \text{ a.s. on the event } \{x \in \partial C(\sigma)\}\,. \tag{2.11}$$

Let F_1 and F_2 be finite subsets of $\mathbb{B}\setminus\{0\}$ and $[x,\infty)^c$ respectively, let B_0 be a measurable set of nonnegative reals, and let B_j be Borel subsets of $[0,\infty)^{F_j}$ for $j=1,2$. If $G_1 = \{(U_{x\oplus y} : y \in F_1) \in B_1\}$ and $G_2 = \{(U_y : y \in F_2) \in B_2\}$, then

$$\begin{aligned}
\mathbb{P}(\tilde U_0 \in& B_0, G_1, G_2, x \in \partial C(\sigma)) \\
&= \mathbb{P}\big(1((T(x)-\sigma)\alpha^{-|x|} \in B_0, G_2, T(\mathsf{par}(x)) \leq \sigma < T(x))\mathbb{P}(G_1|\mathcal{E}_x)\big) \\
&= \nu^{F_1}(B_1)\mathbb{P}\Big(1(G_2, T(\mathsf{par}(x)) \leq \sigma) \\
&\qquad \times \mathbb{P}(U_x - (\sigma - T(\mathsf{par}(x)))\alpha^{-|x|} \in B_0, U_x > (\sigma - T(\mathsf{par}(x)))\alpha^{-|x|} | \mathcal{E}_{x-})\Big)\,.
\end{aligned}$$

Since $T(\mathsf{par}(x)) \in \mathcal{E}_{x-}$ and U_x is independent of $\mathcal{E}_{x-}$, the lack of memory property of the exponential shows that the conditional expectation term in the last line above is equal to

$$\nu(B_0)\exp[-\alpha^{-|x|}(\sigma - T(\mathsf{par}(x)))] = \nu(B_0)\mathbb{P}(U_x > \alpha^{-|x|}(\sigma - T(\mathsf{par}(x)))|\mathcal{E}_{x-})$$

on the event $\{T(\mathsf{par}(x)) \leq \sigma\}$. Substitute this in the previous equation to conclude that

$$\mathbb{P}(\tilde U_0 \in B_0, (\tilde U_y : y \in F_1) \in B_1, G_2, x \in \partial C(\sigma)) = \nu(B_0)\nu^{F_1}(B_1)\mathbb{P}(G_2, x \in \partial C(\sigma))\,.$$

It is easy to see this implies (2.11) and hence (2.10). Noting that

$$(T(x \oplus y) - \sigma)\alpha^{-|x|} = \sum_{z \in \mathbb{B}, z \leq y} \tilde U_z \alpha^{-|z|}\,,$$

one derives (2.9) for σ constant and $\mathcal{G}^o_\sigma$ in place of $\mathcal{G}_\sigma$. The entire argument generalizes easily to the case where σ is a $(\mathcal{G}^o_t)$-stopping time taking countably many values.

For a general $(\mathcal{G}_t)$-stopping time σ, choose $(\mathcal{G}^o_t)$-stopping times $\sigma_n > \sigma$ so that σ_n is a multiple of 2^{-n} with $\sigma_n \downarrow \sigma$. Note that $x \in \partial C(\sigma)$ implies that $x \in \partial C(\sigma_n)$ for sufficiently large n. Taking limits in the above result, we arrive at (2.9) with $\bigcap_n(\mathcal{G}^o_{\sigma_n} \vee \mathcal{E}_{x-})$ in place of the smaller σ-field $\mathcal{G}_\sigma \vee \mathcal{E}_{x-}$. The result follows. □

Remark 2.9. This result justifies our earlier assertion that the process $A_n = C(\mathsf{card}_n)$ is a Markov chain satisfying (2.1). (Recall that card_n is the time $C(t)$ reaches size n.)

3. Some Stochastic Monotonicity Lemmas

The stochastic monotonicity results derived in this section will play a pivotal role in the proof of the key L^1 and L^2 estimates in Section 4.

There is an obvious isomorphism between $\mathbb{R}^{\mathbb{B}(0,n)}$ and $\mathbb{R}^{\{0\}} \times (\mathbb{R}^{\mathbb{B}(0,n-1)})^{\mathbb{B}(1)}$ which we denote $u \mapsto \bar{u}$ and we extend this isomorphism to functions $\varphi : \mathbb{R}^{\mathbb{B}(0,n)} \to \mathbb{R}$ by defining $\bar{\varphi}(\bar{u}) = \varphi(u)$. We will use the same notation ($u \to \bar{u}$) to denote the isomorphism between $\mathbb{R}^{\mathbb{B}(n)}$ and $\prod_{y\in\mathbb{B}(1)} \mathbb{R}^{\mathbb{B}_y(n-1)}$ and hence define $\bar{\varphi}(\bar{u}) = \varphi(u)$ for $\varphi : \mathbb{R}^{\mathbb{B}(n)} \to \mathbb{R}$. The purpose of this notation is to allow test functions φ to be built recursively, yielding inductive proofs of distributional inequalities.

If $(B, \mathcal{B}, \leq)$ is a partially ordered ordered measurable space and μ, ν are probability laws on $(B, \mathcal{B})$, μ is *stochastically smaller than* ν (write $\mu \prec \nu$) if and only if $\int \psi d\mu \leq \int \psi d\nu$ for every bounded measurable non-decreasing $\psi : B \to \mathbb{R}$. If the law of X is stochastically smaller than that of Y, write $X \prec Y$. If $B = A^T$ for $A \subset \mathbb{R}$ and T a countable set, we always partially order B by $f \leq g$ if and only if $f(t) \leq g(t)$ for all $t \in T$. Partially order $\mathcal{S}$ by $(C, a) \leq (C', a')$ if and only if $C \subset C'$ and $a \leq a'$ on C.

If Γ is a finite subset of $\mathbb{B}$, $t \in \mathbb{R}$, and $n \in \mathbb{Z}_+$, define a probability $\nu(\Gamma, n, t)$ on $(-\infty, 0)^\Gamma$ by

$$\nu(\Gamma, n, t)(\cdot) = \mathbb{P}((t - T(x) : x \in \Gamma) \in \cdot \,|T(n) > t)$$

(for $t \leq 0$ the conditioning is trivial).

Lemma 3.1. $\nu(\mathbb{B}(0,n), n, s) \prec \nu(\mathbb{B}(0,n), n, t)$ whenever $-\infty < s \leq t < \infty$ and $n \in \mathbb{Z}_+$.

Proof. If $n = 0$, $\nu(\mathbb{B}(0), 0, s) = \nu(\mathbb{B}(0), 0, 0)$ for $s \geq 0$ by the lack of memory property of the exponential distribution and is trivially stochastically non-decreasing for $s \leq 0$.

Assume for induction that the result holds for $n-1$. Fix $t \in \mathbb{R}$ and for $z \in \mathbb{B}(1)$, let φ_z be bounded measurable functions on $\mathbb{R}^{\mathbb{B}(0,n-1)}$. Let φ_0 be bounded and measurable on $\mathbb{R}$ and define φ on $\mathbb{R}^{\mathbb{B}(0,n)}$ by $\bar{\varphi}(\bar{u}) = \varphi_0(\bar{u}_0)\prod_{z\in\mathbb{B}(1)} \varphi_z(\bar{u}(z))$. Also let $\varphi^{(\alpha)}(u) = \varphi(\alpha u)$, and set $g(u) = \mathbb{P}(T(n-1) > u)^d$. Use the independence of the vectors $V_z = (T^{(z)}(x) : x \in \mathbb{B}(0, n-1))$ and integrate out the exponential variable U_0 to see that

$$\begin{aligned}
&\mathbb{P}(T(n) > t)\nu(\mathbb{B}(0,n), n, t)(\varphi)\\
&\quad= \int_0^\infty e^{-s}\varphi_0(t-s)\\
&\qquad\times \mathbb{P}\Big(\prod_{z\in\mathbb{B}(1)} 1(s + \alpha T^{(z)}(n-1) > t)\varphi_z(t - s - \alpha T^{(z)}(x) : x \in \mathbb{B}(0, n-1))\Big)\, ds\\
&\quad= \int_0^\infty e^{-s}g((t-s)\alpha^{-1})\varphi_0^{(\alpha)}((t-s)/\alpha)\\
&\qquad\times \prod_{z\in\mathbb{B}(1)} \mathbb{P}(\varphi_z^{(\alpha)}((t-s)\alpha^{-1} - T^{(z)}(x) : x \in \mathbb{B}(0, n-1))|\ T^{(z)}(n-1) > \frac{t-s}{\alpha})\, ds\\
&\quad= \int_0^\infty e^{-s}g((t-s)\alpha^{-1})\Big[\bigotimes_{j=1}^d \nu(\mathbb{B}(0,n-1), n-1, (t-s)\alpha^{-1})\Big](\overline{\varphi^{(\alpha)}}((t-s)\alpha^{-1}, \cdot))\, ds\,.
\end{aligned}$$

The restriction on φ is then easily removed, to yield the above equality for all bounded measurable φ on $\mathbb{R}^{\mathbb{B}(0,n)}$. If

$$\psi(v) = \big[\bigotimes_{j=1}^{d} \nu(\mathbb{B}(0,n-1), n-1, v)\big]\big(\overline{\varphi^{(\alpha)}}(v,\cdot)\big)\,,$$

a change of variables in the above integral leads to

$$\nu(\mathbb{B}(0,n),n,t)(\varphi) = (e^t \mathbb{P}(T(n) > t))^{-1} \alpha \int_{-\infty}^{t/\alpha} e^{\alpha v} g(v)\psi(v)\, dv\,.$$

Take φ to be the constant function 1, hence $\psi \equiv 1$ also, to see that

$$e^t \mathbb{P}(T(n) > t) = \alpha \int_{-\infty}^{t/\alpha} e^{\alpha v} g(v)\, dv\,,$$

and therefore conclude

$$\text{(3.1)} \qquad \nu(\mathbb{B}(0,n),n,t)(\varphi) = \int_{-\infty}^{t/\alpha} e^{\alpha v} g(v)\psi(v)\, dv \Big(\int_{-\infty}^{t/\alpha} e^{\alpha v} g(v)\, dv\Big)^{-1}.$$

If φ is non-decreasing, then ψ is nondecreasing by the induction hypothesis. Since (3.1) expresses $\nu(\mathbb{B}(0,n),n,t)(\varphi)$ as a weighted average of values of ψ with respect to a weighting measure that stochastically increases in t, it follows that $\nu(\mathbb{B}(0,n),n,t)(\varphi)$ is nondecreasing in t and the induction is complete. □

Corollary 3.2. $\nu_s \prec \nu_t$ whenever $0 \le s \le t$.

Proof. Define $\Gamma : [0,\infty) \to \mathbb{R}^{\mathbb{B}}$ by $\Gamma(t)(x) = t - T(x)$ and $\psi : \mathbb{R}^{\mathbb{B}} \to \mathcal{S}$ by

$$\psi(\gamma) = (C,a), \quad C = \{x : \gamma(x) \ge 0\}, \quad a(x) = \gamma(x) \text{ for } x \in C :$$

here $\gamma \in \mathbb{R}^{\mathbb{B}}$. Then ψ is non-decreasing and $Y(t) = \psi(\Gamma(t))$. It therefore suffices to show that $t \to \mathbb{P}(\Gamma(t) \in \cdot \,|T(\infty) > t)$ is stochastically non-decreasing, and for this it suffices to fix $m \in \mathbb{N}$ and show that if $\Gamma_m(t) = \Gamma(t)|_{\mathbb{B}(0,m)}$, then $t \to \mathbb{P}(\Gamma_m(t) \in \cdot\,|T(\infty) > t)$ is stochastically non-decreasing (see Kamae, Krengel and O'Brien (1977, Proposition 2)). By taking limits, one reduces this in turn to proving

$$\text{(3.2)} \qquad t \to \mathbb{P}(\Gamma_m(t) \in \cdot\,|T(n) > t) \text{ is stochastically non-decreasing for all } n \ge m.$$

It suffices to consider (3.2) with $m = n$ since decreasing m only weakens the conclusion. But (3.2) with $m = n$ is precisely the conclusion of Lemma 3.1. □

The inductive arguments from here on require a second set of percolation times, defined analogously to the first but not including the percolation time at the root of each subtree. We apologize for doubling the notation but promise not to do it again.

Notation. Let

$$T_{\mathsf{not}}(n) = \inf\{T(0,x) : x \in \mathbb{B}(n)\} = T(n) - U_0$$
$$T^{(x)}_{\mathsf{not}}(y) = T(x, x \oplus y)\alpha^{-|x|} = T^{(x)}(y) - U_x$$
$$T^{(x)}_{\mathsf{not}}(n) = \inf\{T^{(x)}_{\mathsf{not}}(y) : y \in \mathbb{B}(n)\} = T^{(x)}(n) - U_x\,.$$

In short, the $T^{(x)}$'s include a contribution from U_x while the $T^{(x)}_{\mathsf{not}}$'s do not; times with superscripts are rescaled. It is evident that $\{T^{(x)}_{\mathsf{not}}(y) : y \in \mathbb{B}\}$ is equal in law to $\{T^{(0)}_{\mathsf{not}}(y) : y \in \mathbb{B}\}$. For $|z| \geq 1$, let $\mathsf{First}_n(z)$ denote the a.s. unique vertex in $\mathbb{B}(n)$ such that $T^{(z)}(n) = T^{(z)}(\mathsf{First}_n(z))$, so that $z \oplus \mathsf{First}_n(z)$ is the first descendant of z in generation $|z| + n$ to be reached. (Of course, this is not necessarily the one through which infinity is reached from z). Let First_n be the a.s. unique vertex in $\mathbb{B}(n)$ such that $T(n) = T(\mathsf{First}_n)$. For $z \in \mathbb{B}(n)$ let $\mathbb{P}_z(\cdot) = \mathbb{P}(\cdot\,|\mathsf{First}_n = z)$ and $\{\mu(z,t) : t \geq 0\}$ be a set of regular conditional probabilities on $(-\infty, 0]^{\mathbb{B}(n)}$ for

$$\mathbb{P}_z((T_{\mathsf{not}}(n) - T(0,x) : x \in \mathbb{B}(n)) \in \cdot\,|T_{\mathsf{not}}(n) = t).$$

With the available symmetry, we could have defined $\mu(n,t)$ instead of $\mu(z,t)$, but in this case keeping greater generality also reduces confusion of types.

Lemma 3.3. For any $z \in \mathbb{B} \setminus \{0\}$, there is a version of the set $\{\mu(z,t) : t \geq 0\}$ such that

$$\mu(z,s) \prec \mu(z,t) \text{ for } 0 \leq s \leq t.$$

Question: Is there a useful description of the increasing limit, $\mu(z,\infty)$ of the measures $\mu(z,t)$? How about the measure $\mu(z,0)$? In either case, sending $|z|$ to infinity and rescaling by $\alpha^{-|z|}$ should then yield a locally finite point process on $(-\infty, 0]$.

Proof. We start by establishing a pair of auxiliary results, (3.3) and (3.4), whose proofs could be omitted on a first reading. Let $z = (z_1, z_2, \ldots, z_n) = z_1 \oplus z' \in \mathbb{B}(n)$, $n \geq 1$. The first result is that, conditional on $\mathsf{First}_n = z$ and on the percolation times from the vertex z_1, the vectors $\{T^{(x)}(y) : y \in \mathbb{B}\}$ are i.i.d. as x ranges over the other first generation vertices, and distributed as $\{T(y) : y \in \mathbb{B}\}$ conditioned on $T(n-1) > T_{\mathsf{not}}(n)(\omega)\alpha^{-1}$. More precisely, letting T denote the vector $(T(x) : x \in \mathbb{B})$, we show that

$$\mathbb{P}_z\big((T^{(x)} : |x| = 1, x \neq z_1) \in \cdot\,|\mathcal{F}_{[z_1,\infty)}\big)(\omega) = \bigotimes_{j=1}^{d-1} \mathbb{P}(T \in \cdot\,|T(n-1) > T_{\mathsf{not}}(n)(\omega)\alpha^{-1})\,, \tag{3.3}$$

i.e., the RHS of (3.3) defines a regular conditional probability for the left side.

To prove (3.3), note that

$$\{\mathsf{First}_n = z\} = \{T^{(z_1)}(n-1) = \alpha^{-1}T(0,z)\} \cap \bigcap_{|x|=1, x \neq z} \{T^{(x)}(n-1) > \alpha^{-1}T(0,z)\}$$

almost surely. This shows that if $\varphi_x : \mathbb{R}^{\mathbb{B}} \to \mathbb{R}$ are bounded and measurable for $x \in \mathbb{B}(1)$, then

$$\mathbb{P}\Big(1(\mathsf{First}_n = z) \prod_{|x|=1} \varphi_x(T^{(x)})\Big) = \mathbb{P}\Big(\varphi_{z_1}(T^{(z_1)}) 1(T^{(z_1)}(n-1) = \alpha^{-1}T(0,z))$$
$$\times \prod_{x \in \mathbb{B}(1) \setminus \{z_1\}} 1(T^{(x)}(n-1) > \alpha^{-1}T(0,z))\varphi_x(T^{(x)})\Big)\,.$$

The term in front of the product is $\mathcal{F}_{[z_1,\infty)}$-measurable (since $\mathcal{F}_{[z_1,\infty)}$ is just $\sigma(T^{(z_1)})$) and conditional on $\mathcal{F}_{[z_1,\infty)}$, the vectors $T^{(x)}$ for $x \neq z_1$ are i.i.d. copies of T. Thus

$$\begin{aligned}
&\mathbb{P}_z\Big(\prod_{x\in\mathbb{B}(1)} \varphi_x(T^{(x)})\Big)\mathbb{P}(\mathsf{First}_n = z) \\
&= \int \varphi_{z_1}(T^{(z_1)}(\omega)) \prod_{x\in\mathbb{B}(1)\setminus\{z_1\}} \mathbb{P}(\varphi_x(T)|T(n-1) > \alpha^{-1}T(0,z)(\omega)) \\
&\times \big[1(T^{(z_1)}(n-1)(\omega) = \alpha^{-1}T(0,z)(\omega)) \prod_{x\in\mathbb{B}(1)\setminus\{z_1\}} \mathbb{P}(T^{(x)}(n-1) > \alpha^{-1}T(0,z)(\omega))\big]\, d\mathbb{P}(\omega)\ .
\end{aligned}$$

The term in square brackets equals $\mathbb{P}(\mathsf{First}_n = z|\mathcal{F}_{[z_1,\infty)})(\omega)$ and so we may conclude that

$$\mathbb{P}_z\Big(\prod_{x\in\mathbb{B}(1)} \varphi_x(T^{(x)})\Big) = \int \varphi_{z_1}(T^{(z_1)}) \prod_{x\in\mathbb{B}(1)\setminus\{z_1\}} \mathbb{P}(\varphi_x(T)|T(n-1) > \alpha^{-1}T(0,z)(\omega))\, d\mathbb{P}_z(\omega)\,.$$

From this, (3.3) follows immediately upon noting that $T(0,z) = T_{\mathsf{not}}(n)$ a.s. with respect to $\mathbb{P}_z$.

The second result is that if the joint distribution of the times $T^{(z_1)}_{\mathsf{not}}(\cdot)$ is conditioned on $T^{(z_1)}_{\mathsf{not}}(n-1)$, on U_{z_1} and on $\mathsf{First}_n = z$, then the value of U_{z_1} is irrelevant, and the information $\mathsf{First}_n = z$ may be replaced by the weaker $\mathsf{First}_{n-1}(z_1) = z'$. Formally, if $\mathbb{P}'_{z'}(\cdot)$ denotes $\mathbb{P}(\cdot|\mathsf{First}_{n-1}(z_1) = z')$ and $\varphi : \mathbb{R}^{\mathbb{B}} \to \mathbb{R}$ is bounded and measurable, then, letting $T^{z_1}_{\mathsf{not}}$ denote the vector $(T^{z_1}_{\mathsf{not}}(x) : x \in \mathbb{B})$, we claim

$$\mathbb{P}_z(\varphi(T^{(z_1)}_{\mathsf{not}})|T^{(z_1)}_{\mathsf{not}}(n-1), U_{z_1}) = \mathbb{P}'_{z'}(\varphi(T^{(z_1)}_{\mathsf{not}})|T^{(z_1)}_{\mathsf{not}}(n-1)) \text{ a.s.} \tag{3.4}$$

Note that $\mathbb{P}_z$ is absolutely continuous with respect to $\mathbb{P}'_{z'}$, so that in (3.4) we are asserting that the right side defines a version of the left side.

To prove (3.4), first drop the conditioning on U_{z_1} from the LHS by observing U_{z_1} to be independent of $\sigma(T^{(z_1)}_{\mathsf{not}}(\cdot))$. Next, write

$$\{\mathsf{First}_n = z\} = \{\mathsf{First}_{n-1}(z_1) = z'\} \cap \bigcap_{x\in B(1)\setminus\{z_1\}} \{U_{z_1} + T^{(z_1)}_{\mathsf{not}}(n-1) < T^{(x)}(n-1)\}.$$

The independence of $T^{(z_1)}_{\mathsf{not}}$ and $(U_{z_1}, T^{(x)} : x \in B(1) \setminus \{z_1\})$ shows that

$$\mathbb{P}(\mathsf{First}_n = z|T^{(z_1)}_{\mathsf{not}}) = \mathbb{P}(\mathsf{First}_n = z|1(\mathsf{First}_{n-1}(z_1) = z'), T^{(z_1)}_{\mathsf{not}}(n-1)) \text{ a.s.}$$

If ψ is a bounded, measurable real function, then

$$\begin{aligned}
&\mathbb{P}(\mathsf{First}_n = z)\mathbb{P}_z(\varphi(T^{(z_1)}_{\mathsf{not}})\psi(T^{(z_1)}_{\mathsf{not}}(n-1))) \\
&\quad=\mathbb{P}\big(\varphi(T^{(z_1)}_{\mathsf{not}})\psi(T^{(z_1)}_{\mathsf{not}}(n-1))\mathbb{P}(\mathsf{First}_n = z|1(\mathsf{First}_{n-1}(z_1) = z'), T^{(z_1)}_{\mathsf{not}}(n-1))\big) \\
&\quad=\mathbb{P}\big(\mathbb{P}(\varphi(T^{(z_1)}_{\mathsf{not}})|1(\mathsf{First}_{n-1}(z_1) = z'), T^{(z_1)}_{\mathsf{not}}(n-1))\psi(T^{(z_1)}_{\mathsf{not}}(n-1))1(\mathsf{First}_n = z)\big)\,.
\end{aligned}$$

A separate consequence of the definition of conditional expectation is that almost surely on $\{\mathsf{First}_{n-1}(z_1) = z'\}$ (which contains $\{\mathsf{First}_n = z\}$)

$$\mathbb{P}(\varphi(T_{\mathsf{not}}^{(z_1)})|1(\mathsf{First}_{n-1}(z_1) = z'), T_{\mathsf{not}}^{(z_1)}(n-1)) = \mathbb{P}'_{z'}(\varphi(T_{\mathsf{not}}^{(z_1)})|T_{\mathsf{not}}^{(z_1)}(n-1)) .$$

Combining this with the previous identity shows that

$$\begin{aligned}
&\mathbb{P}_z(\varphi(T_{\mathsf{not}}^{(z_1)})\psi(T_{\mathsf{not}}^{(z_1)}(n-1))) \\
&\quad = \mathbb{P}(\mathbb{P}'_{z'}(\varphi(T_{\mathsf{not}}^{(z_1)})|T_{\mathsf{not}}^{(z_1)}(n-1))\psi(T_{\mathsf{not}}^{(z_1)}(n-1))1(\mathsf{First}_n = z))\mathbb{P}(\mathsf{First}_n = z)^{-1} \\
&\quad = \mathbb{P}_z(\mathbb{P}'_{z'}(\varphi(T_{\mathsf{not}}^{(z_1)})|T_{\mathsf{not}}^{(z_1)}(n-1))\psi(T_{\mathsf{not}}^{(z_1)}(n-1)))
\end{aligned}$$

and (3.4) follows.

The lemma is now proved by induction on n as follows. For $n = 1$, let $Q_z = \delta_0$ and for $x \in \mathbb{B}(1) \setminus \{z\}$, let Q_x be the law of $-\alpha$ times an exponential of mean 1. The lack of memory property for the exponential law shows that $\bigotimes_{x\in\mathbb{B}(1)} Q_x$ is a version of $\mu(z,t)$ for all $t \geq 0$, so the result holds with equality of all the laws.

Assume the result now for $|z| < n$, retaining the above notation. Let $\varphi_x : \mathbb{R}^{\mathbb{B}(n-1)} \to \mathbb{R}$ be bounded and measurable, with $\varphi : \mathbb{R}^{\mathbb{B}(n)} \to \mathbb{R}$ given by $\bar{\varphi}(\bar{u}) = \prod_{x\in\mathbb{B}(1)} \varphi_x(\bar{u}(x))$, i.e., by $\varphi(u(\cdot)) = \prod_{x\in\mathbb{B}(1)} \varphi_x(u(x \oplus y) : y \in \mathbb{B}(n-1))$. Note that if $x \in \mathbb{B}(n-1)$ then almost surely with respect to $\mathbb{P}_z$,

$$\begin{aligned}
T_{\mathsf{not}}(n) &= \alpha(U_{z_1} + T_{\mathsf{not}}^{(z_1)}(n-1)) \\
T_{\mathsf{not}}(n) - \alpha T^{(z_1)}(x) &= \alpha(U_{z_1} + T_{\mathsf{not}}^{(z_1)}(n-1)) - \alpha(U_{z_1} + T_{\mathsf{not}}^{(z_1)}(x)) \\
&= \alpha(T_{\mathsf{not}}^{(z_1)}(n-1) - T_{\mathsf{not}}^{(z_1)}(x)) .
\end{aligned} \tag{3.5}$$

Condition on $\mathcal{F}_{[z_1,\infty)}$ and use (3.3) and (3.5) to see that

$$\begin{aligned}
&\mathbb{P}_z\big(\varphi(T_{\mathsf{not}}(n) - T(0,x) : x \in \mathbb{B}(n))|T_{\mathsf{not}}^{(z_1)}(n-1), U_{z_1}\big)(\omega) \\
&\quad = \mathbb{P}_z\big(\prod_{y\in\mathbb{B}(1)} \varphi_y(T_{\mathsf{not}}(n) - \alpha T^{(y)}(x) : x \in \mathbb{B}(n-1))|T_{\mathsf{not}}^{(z_1)}(n-1), U_{z_1}\big)(\omega) \\
&\quad = \prod_{y\in B(1)\setminus\{z_1\}} \mathbb{P}\big(\varphi_y(\alpha(T_{\mathsf{not}}(n)(\omega)\alpha^{-1} - T(x)) : x \in \mathbb{B}(n-1))|T(n-1) > T_{\mathsf{not}}(n)(\omega)\alpha^{-1}\big) \\
&\qquad \times \mathbb{P}_z\big(\varphi_{z_1}(\alpha(T_{\mathsf{not}}^{(z_1)}(n-1) - T_{\mathsf{not}}^{(z_1)}(x)) : x \in \mathbb{B}(n-1))|T_{\mathsf{not}}^{(z_1)}(n-1), U_{z_1}\big)(\omega) .
\end{aligned}$$

Recall the notation $\varphi^{(\alpha)}(u) = \varphi(\alpha u)$. The previous line now becomes

$$\begin{aligned}
&\mathbb{P}_z\big(\varphi(T_{\mathsf{not}}(n) - T(0,x) : x \in \mathbb{B}(n))|T_{\mathsf{not}}^{(z_1)}(n-1), U_{z_1}\big)(\omega) \\
&\quad = \Big\{ \prod_{y\in\mathbb{B}(1)\setminus\{z_1\}} \big[\nu(\mathbb{B}(n-1), n-1, T_{\mathsf{not}}(n)(\omega)\alpha^{-1})(\varphi_y^{(\alpha)})\big] \Big\} \\
&\qquad \times \mathbb{P}_z\big(\varphi_{z_1}^{(\alpha)}(T_{\mathsf{not}}^{(z_1)}(n-1) - T_{\mathsf{not}}^{(z_1)}(x) : x \in \mathbb{B}(n-1))|T_{\mathsf{not}}^{(z_1)}(n-1), U_{z_1}\big)(\omega) .
\end{aligned}$$

Now use (3.4) to equate the above to

$$\Big\{ \prod_{y\in\mathbb{B}(1)\setminus\{z_1\}} \big[\nu(\mathbb{B}(n-1), n-1, T_{\mathsf{not}}(n)(\omega)\alpha^{-1})(\varphi_y^{(\alpha)})\big]\Big\} \times \mathbb{P}'_{z'}\big(\varphi_{z_1}^{(\alpha)}(T_{\mathsf{not}}^{(z_1)}(n-1) - T_{\mathsf{not}}^{(z_1)}(x) : x \in \mathbb{B}(n-1)) | T_{\mathsf{not}}^{(z_1)}(n-1)\big)(\omega)\,.$$

Use the identity $\mathbb{P}'_{z'}((T_{\mathsf{not}}^{(z_1)}(x) : x \in \mathbb{B}) \in \cdot) = \mathbb{P}_{z'}((T(0,x) : x \in \mathbb{B}) \in \cdot)$ (in words: looking in the subtree from z_1 conditioned on $z_1 \oplus z'$ being the first of its generation reached among the subtree is the same as looking in the whole tree conditioned on z' being the first in its generation) to get

$$\begin{aligned}&\mathbb{P}_z\big(\varphi(T_{\mathsf{not}}(n) - T(0,x) : x \in \mathbb{B}(n)) | T_{\mathsf{not}}^{(z_1)}(n-1), U_{z_1}\big)(\omega)\\ &\quad = \mu(z', T_{\mathsf{not}}^{(z_1)}(n-1)(\omega))(\varphi_{z_1}^{(\alpha)}) \prod_{y\in\mathbb{B}(1)\setminus\{z_1\}} \big[\nu(\mathbb{B}(n-1), n-1, T_{\mathsf{not}}(n)(\omega)\alpha^{-1})(\varphi_y^{(\alpha)})\big]\,.\end{aligned}$$

Condition both sides of the above with respect to $\sigma(T_{\mathsf{not}}(n))$ which is contained in $\sigma(T_{\mathsf{not}}^{(z_1)}, U_{z_1}) \vee \{\mathbb{P}_z\text{-null sets}\}$ and use (3.5) together with the independence of U_{z_1} and $T_{\mathsf{not}}^{(z_1)}(n-1)$ to conclude that

$$\begin{aligned}&\mathbb{P}_z\big(\varphi(T_{\mathsf{not}}(n) - T(0,x) : x \in \mathbb{B}(n)) | T_{\mathsf{not}}(n)\big)(\omega)\\ &= \int_0^{T_{\mathsf{not}}(n)(\omega)/\alpha} \Big[\mu(z', s) \times \bigotimes_{y\in\mathbb{B}(1)\setminus\{z_1\}} \nu(\mathbb{B}(n-1), n-1, T_{\mathsf{not}}(n)(\omega)\alpha^{-1})\Big]\\ &\qquad \times (\overline{\varphi^{(\alpha)}}) \exp(-((T_{\mathsf{not}}(n)(\omega)/\alpha) - s))\mathbb{P}_z(T_{\mathsf{not}}^{(z_1)}(n-1) \in ds)\\ &\qquad \times \Big(\int_0^{T_{\mathsf{not}}(n)(\omega)/\alpha} \exp(-((T_{\mathsf{not}}(n)(\omega)/\alpha) - s))\mathbb{P}_z(T_{\mathsf{not}}^{(z_1)}(n-1) \in ds)\Big)^{-1}\\ &= \int_0^{T_{\mathsf{not}}(n)(\omega)/\alpha} \Big[\mu(z', s) \times \bigotimes_{y\in\mathbb{B}(1)\setminus\{z_1\}} \nu(\mathbb{B}(n-1), n-1, T_{\mathsf{not}}(n)(\omega)\alpha^{-1})\Big]\\ &\qquad \times (\overline{\varphi^{(\alpha)}}) e^s \mathbb{P}_z(T_{\mathsf{not}}^{(z_1)}(n-1) \in ds)\Big(\int_0^{T_{\mathsf{not}}(n)(\omega)/\alpha} e^s \mathbb{P}_z(T_{\mathsf{not}}^{(z_1)}(n-1) \in ds)\Big)^{-1}.\end{aligned}$$

Let $\mu(z,t)(\varphi)$ be defined by the above expression with t in place of $T_{\mathsf{not}}(n)(\omega)$. The above shows that $\{\mu(z,t) : t \geq 0\}$ are regular conditional probabilities for the required conditional distributions. By induction, there is a version of the measures $\mu(z', s)$ that is stochastically nondecreasing in s, while Lemma 3.1 shows that $\bigotimes_{y\in\mathbb{B}(1)\setminus\{z_1\}} \nu(\mathbb{B}(n-1), n-1, t\alpha^{-1})$ is stochastically nondecreasing in t. Thus for each bounded, nondecreasing function, φ on $\mathbb{R}^{\mathbb{B}(n)}$, $\mu(z,t)(\varphi)$ may be written as a weighted average on $[0, t/\alpha]$ of a nondecreasing function:

$$\int_0^{t/\alpha} \psi(s,t)\gamma(ds)\gamma([0,t/\alpha])^{-1}\,,$$

where ψ is nondecreasing in each variable (combining the effects of $\mu(z', s)$ and the measures $\nu(\mathbb{B}(n-1), n-1, t/\alpha)$), and γ is the locally finite measure given by the e^s term times the density of $T_{\mathsf{not}}^{(z_1)}(n-1)$. Such a weighted average is clearly nondecreasing in t. □

Corollary 3.4. If $\theta \geq 0$ and $n \in \mathbb{N}$, there is a nondecreasing version of

$$\mathcal{K}_n(t) = \mathbb{P}\big(\sum_{x \in \mathbb{B}(n)} \exp(-\theta(T(0,x) - T_{\mathsf{not}}(n))) | T_{\mathsf{not}}(n) = t \big).$$

Proof. For $z \in \mathbb{B}(n)$, let

$$\mathcal{K}_{n,z}(t) = \mathbb{P}\big(\sum_{x \in \mathbb{B}(n)} \exp(-\theta(T(0,x) - T_{\mathsf{not}}(n))) | \mathsf{First}_n = z, T_{\mathsf{not}}(n) = t \big).$$

Since the sum is increasing in each $T_{\mathsf{not}}(n) - T(0,x)$, Lemma 3.3 implies that each $\mathcal{K}_{n,z}(t)$ has a nondecreasing version. But by symmetry, $\mathcal{K}_{n,z}$ is independent of z, so is equal to $\mathcal{K}_n$ almost surely. $\square$

The following elementary result is proved by an integration by parts and is stated for future reference.

Lemma 3.5. Assume $p, q : \mathbb{R} \to [0, \infty)$ with p non-decreasing and q non-increasing. Then for any random variable X, $\mathbb{P}(p(X)q(X)) \leq \mathbb{P}(p(X))\ \mathbb{P}(q(X))$. $\square$

4. L^2 Bounds for the Clusters

As we saw in the proof of Theorem 2.6, the key in establishing the linear growth of $h(A_n)$ is a good bound on the size of each cluster off the backbone. These clusters are governed by the laws $(\nu_t, t \geq 0)$ of $C(t)$ conditioned to still be finite at time t (Theorem 2.5). The main results of this section are Theorems 4.4 and 4.6, which give uniform bounds on $\nu_t(\#(C))$ and $\nu_t(\#(C)^2)$ for $t \geq 0$.

Notation. Set $H(t) = \mathbb{P}(T(\infty) > t)$ and $G(t) = e^t H(t)$.

Lemma 4.1. (a) The functions G and H satisfy

$$H(t) = e^{-t}\big(1 + \alpha \int_0^{t/\alpha} e^{\alpha u} H(u)^d du\big), \quad t \geq 0, \tag{4.1}$$

$$G(t) = 1 + \alpha \int_0^{t/\alpha} e^{-u(d-\alpha)} G(u)^d du, \quad t \geq 0. \tag{4.2}$$

(b) $G(t)$ increases to a finite limit $c_{4.1}(\alpha, d)$ as $t \uparrow \infty$ and

$$1 \leq c_{4.1}(\alpha, d) \leq \exp\big(\log(d/(d-\alpha))/(1-\alpha)\big).$$

Proof. Condition on U_0 and use scaling to conclude that

$$H(t) = e^{-t} + \int_0^t e^{-s} H((t-s)/\alpha)^d ds.$$

Setting $u = (t-s)/\alpha$ gives (4.1), and (4.2) is then immediate. Equation (4.2) shows that $G(t)$ increases to a (possibly infinite) limit $c_{4.1}(\alpha, d) \geq 1$ as $t \to \infty$. Lemma 2.3 shows that $H(t) \leq s(\alpha)e^{-t}$ and therefore $H(t) \leq e^{-(t-t_0)^+}$ where $t_0 = \log(s(\alpha))$. Hence from (4.1),

$$\begin{aligned} H(t) &\leq e^{-t} + \alpha e^{-t} \int_0^{t_0} e^{\alpha u} du + 1(t/\alpha > t_0) \alpha e^{-t} \int_{t_0}^{t/\alpha} e^{\alpha u} e^{-d(u-t_0)} du \\ &\leq e^{\alpha t_0 - t} + \alpha(d-\alpha)^{-1} e^{-t+dt_0-(d-\alpha)t_0} = e^{-(t-t_1)}, \end{aligned}$$

where $t_1 = h(t_0) := \alpha t_0 + \log(d/(d-\alpha))$. Iterating this procedure, we obtain $H(t) \leq e^{-(t-t_n)^+}$ for $n \geq 1$, where $t_{n+1} = h(t_n)$ and $\lim_{n\to\infty} t_n = \log(d/(d-\alpha))(1-\alpha)^{-1}$. Let $n \to \infty$ in $G(t) \leq e^{t_n}$ to complete the proof. $\square$

Remark. (4.1) and (4.2) are rather nasty equations since they are non-linear and (worst of all) anticipative. (See Athreya (1985) for some similar equations, arising from the distribution function of the random variable $\sup_{x\in\mathbb{B}} T(x)$.) Solutions to these equations are not unique because $H \equiv 1$ also satisfies (4.1). It is, however, not hard to show that $\mathbb{P}(T(\infty) > t)$ is the unique non-increasing solution H to (4.1) for which $e^t H(t)$ is bounded. Although it seems difficult to get sharp estimates from these equations, in the next section some closely related equations will help us analyze the process $(W_n, n \in \mathbb{Z}_+)$, and in Section 8 we will derive some asymptotic results as $\alpha \uparrow 1$. In particular, the upper bound on $c_{4.1}$ in (b) is by no means optimal – see Remark 8.3.

Notation. If n is a non-negative integer, define a Laplace transform with respect to the real variables $\theta \in [0,1]$ and $\gamma \geq 1$ by

$$r(n,\theta,\gamma) = \mathbb{P}\Big(\sum_{x\in\mathbb{B}(n)} \exp(\theta T_{\mathsf{not}}(n) - \gamma\alpha^{-n-1}(T^{(0)}_{\mathsf{not}}(x) - T_{\mathsf{not}}(n)))\Big),\ n \in \mathbb{Z}_+.$$

Let

$$d_0 = \tfrac{1}{2}d(d-1) - 1.$$

For $\gamma > \alpha$, $n \in \mathbb{N}$ and $0 \leq \delta \leq t$, let

$$S(n,\gamma,\delta,t) = \sum_{x\in\mathbb{B}(n)} 1(t-\delta \leq T(x) \leq t) \exp(-\gamma\alpha^{-n-1}(t - T(n))).$$

To get a feeling for S, set $\delta = t$, and integrate over t to get

$$\int_0^\infty S(n,\gamma,t,t)dt = (\alpha^{n+1}/\gamma) \sum_{x\in\mathbb{B}(n)} \exp(-\gamma\alpha^{-n-1}(T(x) - T(n))).$$

This is $O(\alpha^{n+1})$, provided that not too many times $T(x) - T(n)$ are near 0 on a scale of α^n. The next result gives the critical technical estimate on r, which is then used to show that S is indeed $O(\alpha^n)$ when $\gamma \geq \alpha + 1$. This will in turn lead to Theorems 4.4 and 4.6 on the respective L^1 and L^2 behaviours of the cluster size. Recall that $s(\alpha) = \prod_{k=1}^\infty (1-\alpha^k)^{-1}$.

Lemma 4.2. For all $n \in \mathbb{Z}_+$, $\gamma \geq 1$,

$$r(n, 1, \gamma) \leq c_{4.2}(\alpha, \gamma, d) := d^2(d-1)^{-1}(1-\alpha)s(\alpha)^2 \exp(d_0/(\gamma(1-\alpha))).$$

Proof. The result is trivial if $n = 0$, thus assume $n \geq 1$ and fix $\gamma \geq 1$. For the induction, assume the result for all $n' < n$ and all θ' (what is actually needed is $\theta' = \alpha^j \theta$). Fix $z_1 \in \{1, \dots, d\}$ and set $\tau = \min\{T^{(y)}(n-1) : y \in \mathbb{B}(1) \setminus \{z_1\}\}$. Using the symmetry of the tree when computing $r(n, \theta, \gamma)$, we may sum over only those x with $x_1 = z_1$ and then multiply by d. This leads to

$$\begin{aligned} r(n,\theta,\gamma) &= d \cdot \mathbb{P}\Big(\sum_{x \in \mathbb{B}(n-1)} \exp\big(\theta T_{\mathsf{not}}(n) - \gamma \alpha^{-n-1}(\alpha U_{z_1} + \alpha T^{(z_1)}_{\mathsf{not}}(x) - \alpha T^{(z_1)}(n-1)) \\ &\qquad - \gamma \alpha^{-n-1}(\alpha T^{(z_1)}(n-1) - T_{\mathsf{not}}(n))\big)\Big) \\ &= d \cdot \mathbb{P}\Big(\sum_{x \in \mathbb{B}(n-1)} \exp\big(\theta T_{\mathsf{not}}(n) - \gamma \alpha^{-n}(T^{(z_1)}_{\mathsf{not}}(x) - T^{(z_1)}_{\mathsf{not}}(n-1)) \\ &\qquad - \gamma \alpha^{-n}(T^{(z_1)}(n-1) - \alpha^{-1} T_{\mathsf{not}}(n))\big)\Big). \end{aligned}$$

Divide the above expectation into two terms corresponding to the events $\{T^{(z_1)}(n-1) > \tau\}$ and $\{T^{(z_1)}(n-1) \leq \tau\}$. Use the fact that on the latter event $T^{(z_1)}_{\mathsf{not}}(n-1) = \alpha^{-1} T_{\mathsf{not}}(n)$ while on the former event $\tau = \alpha^{-1} T_{\mathsf{not}}(n)$ to see that

$$r(n, \theta, \gamma) = I_1 + I_2, \qquad \text{where}$$

$$\begin{aligned} I_1 &= d \cdot \mathbb{P}\Big(1(T^{(z_1)}(n-1) > \tau) \exp(\theta \alpha \tau - \gamma \alpha^{-n}(T^{(z_1)}(n-1) - \tau)) \\ &\qquad \times \sum_{x \in \mathbb{B}(n-1)} \exp(-\gamma \alpha^{-n}(T^{(z_1)}_{\mathsf{not}}(x) - T^{(z_1)}_{\mathsf{not}}(n-1)))\Big), \\ I_2 &= d \cdot \mathbb{P}\Big(1(T^{(z_1)}(n-1) \leq \tau) \exp(\theta \alpha T^{(z_1)}(n-1)) \\ &\qquad \times \sum_{x \in \mathbb{B}(n-1)} \exp(-\gamma \alpha^{-n}(T^{(z_1)}_{\mathsf{not}}(x) - T^{(z_1)}_{\mathsf{not}}(n-1)))\Big). \end{aligned} \tag{4.3}$$

The plan is to bound these terms by constant multiples of $r(n-1, \alpha\theta, \gamma)$ and then apply the induction hypothesis. (In fact I_1 will be of smaller order).

Consider I_1 first. The term $1(T^{(z_1)}(n-1) > \tau) \exp(\theta \alpha \tau - \gamma \alpha^{-n}(T^{(z_1)}(n-1) - \tau))$ will make this term relatively small as n becomes large; heuristically, the typical difference between $T(n)$ and $T(x)$ will be of order 1 when $x_1 \neq z_1$, and so the factor of α^{-n} in the exponent makes these terms small. To verify this, fix $z_2 \in B(1) \setminus \{z_1\}$ and set $R = T^{(z_1)}_{\mathsf{not}}(n-1) - T^{(z_2)}_{\mathsf{not}}(n-1)$. Focus on the expression in front of the sum in I_1 and integrate over the pair (U_{z_1}, U_{z_2}) to see that

$$\begin{aligned}
&\mathbb{P}\big(1(T^{(z_1)}(n-1)>\tau)\exp(\theta\alpha\tau-\gamma\alpha^{-n}(T^{(z_1)}(n-1)-\tau))|\mathcal{F}_{(z_1,\infty)}\big)\\
&\quad=(d-1)\mathbb{P}\big(1(U_{z_1}+T^{(z_1)}_{\mathsf{not}}(n-1)>\tau=U_{z_2}+T^{(z_2)}_{\mathsf{not}}(n-1))\\
&\qquad\times\exp(\theta\alpha(U_{z_2}+T^{(z_2)}_{\mathsf{not}}(n-1))-\gamma\alpha^{-n}(U_{z_1}-U_{z_2}+R))|\mathcal{F}_{(z_1,\infty)}\big)\\
&\quad\le(d-1)\mathbb{P}\Big(\Big[\int_0^\infty\int_0^\infty 1(u_2-u_1\le R)\exp(\theta\alpha u_2-\gamma\alpha^{-n}(u_1-u_2))e^{-u_1}\,du_1\,e^{-u_2}\,du_2\Big]\\
&\qquad\times\exp(\theta\alpha T^{(z_2)}_{\mathsf{not}}(n-1)-\gamma\alpha^{-n}R)|\mathcal{F}_{(z_1,\infty)}\Big)\,.
\end{aligned}$$

First integrate u_2 over $(0,u_1+R)$ and then integrate u_1 over $((-R)^+,\infty)$ to bound the above by

$$\begin{aligned}
&(d-1)(\alpha\theta+\gamma\alpha^{-n}-1)^{-1}(2-\alpha\theta)^{-1}\\
&\qquad\qquad\times\mathbb{P}\big(\exp(\theta\alpha T^{(z_2)}_{\mathsf{not}}(n-1)-(2-\alpha\theta)(-R)^++(\alpha\theta-1)R)|\mathcal{F}_{(z_1,\infty)}\big)\\
&\quad=\ (d-1)\alpha^n(\gamma+\alpha^{n+1}\theta-\alpha^n)^{-1}(2-\alpha\theta)^{-1}\exp(\theta\alpha T^{(z_1)}_{\mathsf{not}}(n-1))\\
&\qquad\qquad\times\mathbb{P}\big(\exp(-(2-\alpha\theta)(-R)^+-R)|\mathcal{F}_{(z_1,\infty)}\big)\,.
\end{aligned}$$

Note that $-(2-\alpha\theta)(-R)^+-R\le 0$ for $\alpha,\theta\in[0,1]$, and so this is bounded by

$$(d-1)\alpha^n(\gamma-\alpha^n)^{-1}(2-\alpha\theta)^{-1}\exp(\theta\alpha T^{(z_1)}_{\mathsf{not}}(n-1))\,.$$

Substitute this bound into the expression for I_1 (the summation being $\mathcal{F}_{(z_1,\infty)}$-measurable) to conclude that I_1 is at most

$$\begin{aligned}
&d(d-1)(2-\alpha\theta)^{-1}\alpha^n(\gamma-\alpha^n)^{-1}\\
&\qquad\times P\Big(\sum_{x\in\mathbb{B}(n-1)}\exp(\theta\alpha T^{(z_1)}_{\mathsf{not}}(n-1)-\gamma\alpha^{-n}(T^{(z_1)}_{\mathsf{not}}(x)-T^{(z_1)}_{\mathsf{not}}(n-1)))\Big)\\
&\qquad=(d_0+1)(1-\alpha\theta/2)^{-1}\alpha^n(\gamma-\alpha^n)^{-1}r(n-1,\alpha\theta,\gamma)\,.
\end{aligned}\tag{4.4}$$

Consider now I_2. By symmetry, as θ becomes small the terms in front of the summation in the expression for I_2 should have mean close to 1, and so I_2 should be close to $r(n-1,\alpha\theta,\gamma)$. This is true, but to make the argument rigorous we must control the possible correlations between the summation and the remaining terms in the integrand. This makes use of the monotonicity results from Section 3, and in particular Corollary 3.4.

Let $q(t)=\mathbb{P}(\exp(\theta\alpha U_{z_1})1(U_{z_1}+t\le\tau))$ and note that q is decreasing in t. Use the independence of $\mathcal{F}_{(z_1,\infty)}$ and $\mathcal{E}_{z_1}\supset\sigma(U_{z_1},\tau)$ to see that

$$\begin{aligned}
&\mathbb{P}\big(1(T^{(z_1)}(n-1)\le\tau)\exp(\theta\alpha T^{(z_1)}(n-1))|\mathcal{F}_{(z_1,\infty)}\big)\\
&\qquad=\mathbb{P}\big(1(U_{z_1}+T^{(z_1)}_{\mathsf{not}}(n-1)\le\tau)\exp(\theta\alpha(U_{z_1}+T^{(z_1)}_{\mathsf{not}}(n-1)))|T^{(z_1)}_{\mathsf{not}}(n-1)\big)\\
&\qquad=\exp(\theta\alpha T^{(z_1)}_{\mathsf{not}}(n-1))q(T^{(z_1)}_{\mathsf{not}}(n-1))\,.
\end{aligned}\tag{4.5}$$

The joint independence of τ, U_{z_1} and $T^{(z_1)}_{\mathsf{not}}(n-1)$ shows that

$$\tilde{q}(u) = \mathbb{P}(U_{z_1} + T^{(z_1)}_{\mathsf{not}}(n-1) \leq \tau | U_{z_1} = u)$$

is decreasing in u. The same independence and Lemma 3.5 give

$$\begin{aligned} \mathbb{P}(q(T^{(z_1)}_{\mathsf{not}}(n-1))) &= \mathbb{P}(\exp(\theta\alpha U_{z_1})\tilde{q}(U_{z_1})) \\ &\leq \mathbb{P}(\exp(\theta\alpha U_{z_1}))\mathbb{P}(\tilde{q}(U_{z_1})) = (1-\theta\alpha)^{-1}d^{-1} \end{aligned} \tag{4.6}$$

by symmetry. Using (4.5) in the expression for I_2 we get

$$\begin{aligned} I_2 = d\mathbb{P}\Big(& q(T^{(z_1)}_{\mathsf{not}}(n-1)) \exp(\theta\alpha T^{(z_1)}_{\mathsf{not}}(n-1)) \\ & \times \mathbb{P}\big(\sum_{x \in \mathbb{B}(n-1)} \exp(-\gamma\alpha^{-n}(T^{(z_1)}_{\mathsf{not}}(x) - T^{(z_1)}_{\mathsf{not}}(n-1)))|T^{(z_1)}_{\mathsf{not}}(n-1)\big)\Big). \end{aligned}$$

Corollary 3.4 shows we may assume the conditional expectation is a nondecreasing function of $T^{(z_1)}_{\mathsf{not}}(n-1)$. Recalling that q is nonincreasing, we again use Lemma 3.5 to conclude that

$$\begin{aligned} I_2 &\leq d\mathbb{P}(q(T^{(z_1)}_{\mathsf{not}}(n-1)))\mathbb{P}\big(\exp(\theta\alpha T^{(z_1)}_{\mathsf{not}}(n-1)) \\ &\quad \times \sum_{x \in \mathbb{B}(n-1)} \exp(-\gamma\alpha^{-n}(T^{(z_1)}_{\mathsf{not}}(x) - T^{(z_1)}_{\mathsf{not}}(n-1)))\big) \\ &\leq (1-\theta\alpha)^{-1} r(n-1, \theta\alpha, \gamma), \end{aligned}$$

by (4.6).

Combine the above with (4.3) and (4.4) to see that

$$\begin{aligned} r(n,\theta,\gamma) &\leq (1-\alpha\theta)^{-1}\left(1 + (d_0+1)\alpha^n(\gamma - \alpha^n)^{-1}\right) r(n-1, \alpha\theta, \gamma) \\ &= (1-\alpha\theta)^{-1}\left(1 - \alpha^n\gamma^{-1}\right)^{-1}(1 + d_0\gamma^{-1}\alpha^n) r(n-1, \alpha\theta, \gamma). \end{aligned}$$

We now use induction, recalling that $\gamma \geq 1$, to conclude that

$$\begin{aligned} r(n,1,\gamma) &\leq \Big[\prod_{k=1}^{n-1}(1-\alpha^k)^{-1}\prod_{k=2}^{n}(1-\alpha^k\gamma^{-1})^{-1}\prod_{k=2}^{n}(1+d_0\gamma^{-1}\alpha^k)\Big] r(1, \alpha^{n-1}, \gamma) \\ &\leq s(\alpha)^2(1-\alpha)\exp(d_0\gamma^{-1}(1-\alpha)^{-1}) r(1, \alpha^{n-1}, \gamma)\ . \end{aligned} \tag{4.7}$$

Note that

$$r(1, \alpha^{n-1}, \gamma) \leq d\mathbb{P}(\exp(\alpha^{n-1}T_{\mathsf{not}}(1))) \leq d\mathbb{P}(\exp(\min_{x \in \mathbb{B}(1)} U_x)) = d^2/(d-1),$$

and use this in (4.7) to complete the argument. $\square$

Lemma 4.3. For $\gamma \geq 1+\alpha$, $n \in \mathbb{Z}_+$, $0 \leq \delta \leq t$,

$$\mathbb{P}(S(n,\gamma,\delta,t)) \leq e^{-t} c_{4.2}(\alpha,\gamma-\alpha,d)(\alpha^{n+1} \wedge \delta).$$

Proof. Let γ, δ, t be as above, and let $n \in \mathbb{Z}_+$; for $x \in \mathbb{B}(n)$ we evaluate the corresponding summand in $S(n,\gamma,\delta,t)$. First, integrate out U_0 to see that

$$\begin{aligned}
&\mathbb{P}\big(1(t-\delta \leq T(x) \leq t)\exp(-\gamma\alpha^{-n-1}(t-T(n)))\big)\\
&\quad = \mathbb{P}\Big(1(T^{(0)}_{\mathsf{not}}(x) \leq t)\int_{(t-\delta-T^{(0)}_{\mathsf{not}}(x))^+}^{t-T^{(0)}_{\mathsf{not}}(x)} \exp(-u-\gamma(t-u-T^{(0)}_{\mathsf{not}}(x))\alpha^{-n-1})\,du\\
&\qquad\quad \times \exp(-\gamma(T^{(0)}_{\mathsf{not}}(x)-T_{\mathsf{not}}(n))\alpha^{-n-1})\Big)\ .
\end{aligned}$$

Change variables to $s = t-u-T^{(0)}_{\mathsf{not}}(x)$ and bound the Lebesgue integral from above on the event $\{T^{(0)}_{\mathsf{not}}(x) \leq t\}$ by:

$$\begin{aligned}
&\int_0^{\delta\wedge(t-T^{(0)}_{\mathsf{not}}(x))} \exp(-(\gamma\alpha^{-n-1}-1)s)\,ds\,\exp(-(t-T^{(0)}_{\mathsf{not}}(x)))\\
&\qquad \leq \alpha^{n+1}(\gamma-\alpha^{n+1})^{-1}(1-\exp(-(\gamma\alpha^{-n-1}-1)\delta))\exp(-t+T^{(0)}_{\mathsf{not}}(x))\\
&\qquad \leq e^{-t}\alpha^{n+1}(\gamma-\alpha^{n+1})^{-1}\min\{1,(\gamma\alpha^{-n-1}-1)\delta\}\exp(T^{(0)}_{\mathsf{not}}(x))\\
&\qquad \leq e^{-t}\min\{\alpha^{n+1}(\gamma-\alpha)^{-1},\delta\}\exp(T^{(0)}_{\mathsf{not}}(x))\ .
\end{aligned}$$

Thus

$$\begin{aligned}
&\mathbb{P}\big(1(t-\delta \leq T(x) \leq t)\exp(-\gamma\alpha^{-n-1}(t-T(n)))\big)\\
&\quad \leq e^{-t}\min\{\alpha^{n+1}(\gamma-\alpha)^{-1},\delta\}\mathbb{P}\big(\exp(T_{\mathsf{not}}(n)-(\gamma\alpha^{-n-1}-1)(T^{(0)}_{\mathsf{not}}(x)-T_{\mathsf{not}}(n)))\big)\\
&\quad \leq e^{-t}\min\{\alpha^{n+1}(\gamma-\alpha)^{-1},\delta\}\mathbb{P}\big(\exp(T_{\mathsf{not}}(n)-(\gamma-\alpha)\alpha^{-n-1}(T^{(0)}_{\mathsf{not}}(x)-T_{\mathsf{not}}(n)))\big)\,.
\end{aligned}$$

Summing over x in $\mathbb{B}(n)$ and recalling that $\gamma \geq 1+\alpha$,

$$\begin{aligned}
\mathbb{P}(S(n,\gamma,\delta,t)) &\leq e^{-t}\min(\alpha^{n+1}(\gamma-\alpha)^{-1},\delta)r(n,1,\gamma-\alpha)\\
&\leq e^{-t}c_{4.2}(\alpha,\gamma-\alpha,d)(\alpha^{n+1}\wedge\delta) \quad \text{(by Lemma 4.2).}
\end{aligned}$$

□

Notation. Let $c_{4.3}(\alpha,d) = c_{4.1}(\alpha,d)^d c_{4.2}(\alpha,d-\alpha,d)(1-\alpha)^{-1}$.

In the following result we set $C(0-) = \emptyset$.

Theorem 4.4. (a) For all $0 \leq \delta \leq t$ and n in $\mathbb{Z}_+$

$$\mathbb{P}(\#((C(t)-C((t-\delta)-))\cap\mathbb{B}(n))|T(\infty)>t) \leq c_{4.1}(\alpha,d)^d c_{4.2}(\alpha,d-\alpha,d)(\alpha^{n+1}\wedge\delta).$$

(b) For all $0 \leq \delta \leq t$,

$$\begin{aligned}
&\mathbb{P}(\#(C(t)-C((t-\delta)-))|T(\infty)>t)\\
&\qquad \leq c_{4.1}(\alpha,d)^d c_{4.2}(\alpha,d-\alpha,d)(1-\alpha)^{-1}(\log^+(1/\delta)+1)(\delta\wedge 1)\\
&\qquad \leq c_{4.3}(\alpha,d),
\end{aligned}$$

and so in particular, $\mathbb{P}(\#C(t)|T(\infty)>t) \leq c_{4.3}(\alpha,d)$ for all $t \geq 0$.

(c) $\mathbb{P}(\#(C(t\wedge T(\infty))\cap\mathbb{B}(n))) \leq c_{4.3}(\alpha,d)(1-e^{-t})$ for all $t \geq 0, n \in \mathbb{Z}_+$.

Proof. Clearly $T(\infty) \le T(n) + T(\mathsf{First}_n, \infty)$ and if $z \in \mathbb{B}(n)$ then Lemma 4.1 shows

$$\mathbb{P}(T(z,\infty) > t) = \mathbb{P}(T(\infty) > t\alpha^{-n-1})^d \le c_{4.1}(\alpha,d)^d \exp(-dt\alpha^{-n-1}). \tag{4.8}$$

Therefore,

$$\begin{aligned}
&\mathbb{P}\big(\#((C(t) - C((t-\delta)-)) \cap \mathbb{B}(n))1(T(\infty) > t)\big) \\
&\qquad = \sum_{x \in \mathbb{B}(n)} \mathbb{P}(t - \delta \le T(x) \le t < T(\infty)) \\
&\qquad \le \sum_{x \in \mathbb{B}(n)} \mathbb{P}\big(1(t-\delta \le T(x) \le t)\ \mathbb{P}(T(\mathsf{First}_n, \infty) > t - T(n) | \mathcal{F}_n)\big) \\
&\qquad \le c_{4.1}(\alpha,d)^d \mathbb{P}(S(n,d,\delta,t)) \quad \text{(by (4.8))}.
\end{aligned}$$

Lemma 4.3 therefore shows that

$$\begin{aligned}
&\mathbb{P}\big(\#((C(t) - C((t-\delta)-)) \cap \mathbb{B}(n)) | T(\infty) > t\big) \\
&\qquad \le c_{4.1}(\alpha,d)^d c_{4.2}(\alpha, d-\alpha, d)(\alpha^{n+1} \wedge \delta)(e^t \mathbb{P}(T(\infty) > t))^{-1} \\
&\qquad \le c_{4.1}(\alpha,d)^d c_{4.2}(\alpha, d-\alpha, d)(\alpha^{n+1} \wedge \delta).
\end{aligned} \tag{4.9}$$

This proves (a), and (b) follows upon summing over $n \in \mathbb{Z}_+$.
(c) As in (a) we have

$$\begin{aligned}
\mathbb{P}(\#(C(t \wedge T(\infty)) \cap \mathbb{B}(n))) &\le \sum_{x \in \mathbb{B}(n)} \mathbb{P}(1(T(x) \le t)\mathbb{P}(T(\mathsf{First}_n, \infty) > T(x) - T(n) | \mathcal{F}_n)) \\
&\le \sum_{x \in \mathbb{B}(n)} \mathbb{P}\big(1(T(x) \le t) c_{4.1}(\alpha,d)^d \exp(-d(T(x) - T(n))\alpha^{-n-1})\big) \ \text{(as in (4.8))}.
\end{aligned}$$

Integrate out U_0 to see that the last summation equals

$$\begin{aligned}
&c_{4.1}(\alpha,d)^d \sum_{x \in \mathbb{B}(n)} \mathbb{P}\Big(\exp(-d(T^{(0)}_{\mathsf{not}}(x) - T_{\mathsf{not}}(n))\alpha^{-n-1}) 1(T^{(0)}_{\mathsf{not}}(x) \le t)\big[1 - \exp(-(t - T^{(0)}_{\mathsf{not}}(x)))\big]\Big) \\
&\qquad \le c_{4.1}(\alpha,d)^d (1 - e^{-t}) r(n,0,d) \\
&\qquad \le c_{4.1}(\alpha,d)^d c_{4.2}(\alpha,d,d)(1 - e^{-t}) \quad \text{(by Lemma 4.2)} \\
&\qquad \le c_{4.3}(\alpha,d)(1 - e^{-t}).
\end{aligned}$$

□

Remark 4.5. If $x \in (0,1)$ and $N = [(\log 1/x)/(\log 1/\alpha)]$ then

$$\begin{aligned}
s(\alpha) &\le \Big\{\prod_1^N (1-\alpha)^{-1}(1 + \ldots + \alpha^{k-1})^{-1}\Big\} \exp\Big(\sum_{k=N+1}^{\infty} \alpha^k (1-\alpha^k)^{-1}\Big) \\
&\le (1-\alpha)^{-N} \prod_{k=1}^N (kx)^{-1} \exp(x(1-x)^{-1}(1-\alpha)^{-1}).
\end{aligned}$$

Use Stirling's Formula and optimize over x to see there are constants, $c_{4.4}, c_{4.5} > 0$ such that

$$s(\alpha) \le c_{4.4} \exp(c_{4.5}(1-\alpha)^{-1}) \text{ for all } \alpha \in (0,1). \tag{4.10}$$

Using (4.10) and Lemma 4.1(b) it follows that

$$\begin{aligned} c_{4.3}(\alpha,d) &= c_{4.1}(\alpha,d)^d c_{4.2}(\alpha, d-\alpha, d)(1-\alpha)^{-1} \\ &\le c_{4.4}^2 d^2 (d-1)^{-1} \exp\big((2c_{4.5} + (d_0/(d-1)) + d\log(d/(d-1)))/(1-\alpha)\big) \\ &= c_{4.6}(d)\exp(c_{4.7}(d)/(1-\alpha)). \end{aligned} \tag{4.11}$$

The Central Limit Theorem in Section 6 will require the following L^2 bounds on the conditioned clusters.

Theorem 4.6. There is a $c_{4.8}(\alpha,d) > 0$ such that

$$\mathbb{P}(\#C(t)^2 | T(\infty) > t) \le c_{4.8}(\alpha,d) \qquad \text{for all } t \ge 0.$$

Proof. (In this proof, we suppress the dependence of $c_{4.i}$ on (α,d)). Let $\sum'$ denote summation over x_1, x_2 in $\mathbb{B}$ for which $x_1 \wedge x_2$ is distinct from both x_1 and x_2. Then

$$\begin{aligned} \mathbb{P}(\#C(t)^2 1(T(\infty) > t)) \le 2\mathbb{P}(&\sum_{x_1}\sum_{x_2} 1(x_2 \ge x_1) 1(T(x_2) \le t < T(\infty))) \\ &+ \mathbb{P}(\sum{}' 1(T(x_1) \vee T(x_2) \le t < T(\infty))). \end{aligned} \tag{4.12}$$

If $y \ne 0$ let $\widehat{T}(y) = \inf\{T(x) : |x| = |y|, x \ne y\}$ and let $\widehat{Z}(y)$ be the a.s. unique vertex in $\mathbb{B}(|y|) - \{y\}$ at which this minimum time is attained. Define $\widehat{T}(0) = \infty$ and $\widehat{Z}(0) = 0$. Note that $(\widehat{T}(y), \widehat{Z}(y))$ is $\mathcal{F}_{|y|}$-measurable and $T(\widehat{Z}(y),\infty)$ is $\mathcal{E}_y$-measurable (recall from the beginning of Section 2 that $\mathcal{E}_y$ is the "wide-sense" past up to y). In the second summation in (4.12), write $x_1 = v \oplus i \oplus y_1$ and $x_2 = v \oplus j \oplus y_2$ where $v = x_1 \wedge x_2$ and $i \ne j \in \{1,\dots,d\}$ to get

$$\begin{aligned} &\mathbb{P}(\sum{}' 1(T(x_1) \vee T(x_2) \le t < T(\infty))) \\ &\le \mathbb{P}\Big\{ \sum_v 1(T(v) \le t < \widehat{T}(v) + T(\widehat{Z}(v),\infty)) \\ &\quad \times \mathbb{P}\Big[\sum_{1 \le i}\sum_{\ne j \le d} \Big(\sum_{y_1} 1(T(v) + T^{(v\oplus i)}(y_1)\alpha^{|v|+1} \le t < T(v) + T^{(v\oplus i)}(\infty)\alpha^{|v|+1})\Big) \\ &\quad \times \Big(\sum_{y_2} 1(T(v) + T^{(v\oplus j)}(y_2)\alpha^{|v|+1} \le t < T(v) + T^{(v\oplus j)}(\infty)\alpha^{|v|+1})\Big) \mid \mathcal{E}_v\Big]\Big\}. \end{aligned}$$

$(T^{(v\oplus i)}(\cdot), T^{(v\oplus j)}(\cdot))$, $i \ne j$, are independent and are jointly independent of $\mathcal{E}_v$. The above therefore equals

$$\begin{aligned} &d(d-1)\int \sum_v 1(T(v)(\omega) \le t)\mathbb{P}\big(T(\widehat{Z}(v),\infty) > t - \widehat{T}(v)|\mathcal{F}_{|v|}\big)(\omega) \\ &\qquad \times \big[\mathbb{P}(\#C((t-T(v)(\omega))\alpha^{-|v|-1}) 1(T(\infty) > (t-T(v)(\omega))\alpha^{-|v|-1}))\big]^2 \, d\mathbb{P}(\omega) \\ &\le d(d-1)\int \sum_v 1(T(v)(\omega) \le t)\mathbb{P}(T(0,\infty) > (t-\widehat{T}(v)(\omega))\alpha^{-|v|}) \\ &\qquad \times c_{4.3}^2[\mathbb{P}(T(\infty) > (t-T(v)(\omega))\alpha^{-|v|-1})]^2 \, d\mathbb{P}(\omega)\,. \end{aligned} \tag{4.13}$$

The time $T(0,\infty)$ is equal in law to the minimum of d independent copies of $\alpha T(\infty)$. Lemma 4.1(c) and the fact that $T(|v|) = \min\{\widehat{T}(v), T(v)\}$ therefore show that if $T(v)(\omega) \le t$, then

$$\begin{aligned}\mathbb{P}(T(0,\infty) > (t-\widehat{T}(v)(\omega))\alpha^{-|v|})[\mathbb{P}(T(\infty) > (t-T(v)(\omega))\alpha^{-|v|-1})]^2 \\ \le & c_{4.1}^{d+2}\exp\big\{-d\alpha^{-|v|-1}(t-\widehat{T}(v)(\omega))^+ - 2\alpha^{-|v|-1}(t-T(v)(\omega))\big\} \\ \le & c_{4.1}^{d+2}\exp\big\{-2\alpha^{-|v|-1}(t-T(|v|)(\omega))\big\}\,.\end{aligned}$$

Substitute this into the RHS of (4.13), thus bounding the second term on the RHS of (4.12) by

$$d(d-1)c_{4.3}^2 c_{4.1}^{d+2}\int\sum_v 1(T(v)(\omega)\le t)\exp\big\{-2\alpha^{-|v|-1}(t-T(|v|)(\omega))\big\}d\mathbb{P}(\omega)$$

which is equal to

$$d(d-1)c_{4.3}^2 c_{4.1}^{d+2}\mathbb{P}(\sum_{n=0}^{\infty} S(n,2,t,t))$$

and hence at most $c(\alpha,d)e^{-t}$ by Lemma 4.3.

The first term on the right side of (4.12) equals

$$\begin{aligned}2\mathbb{P}(\sum_x |x|1(T(x)\le t<T(\infty))) &= 2\mathbb{P}(\sum_{n=1}^{\infty} n\#(C(t)\cap\mathbb{B}(n))|T(\infty)>t)\mathbb{P}(T(\infty)>t) \\ &\le c'(\alpha,d)\mathbb{P}(T(\infty)>t) \qquad \text{(by Theorem 4.4(a))}.\end{aligned}$$

Use this and the above bound in (4.12) to conclude

$$\begin{aligned}\mathbb{P}(\#C(t)^2|T(\infty)>t) &\le c'(\alpha,d)+c(\alpha,d)(e^t\mathbb{P}(T(\infty)>t))^{-1} \\ &\le c'(\alpha,d)+c(\alpha,d)\,.\end{aligned}$$

□

5. The Process W

Recall from Section 2 that $W_n = T(\mathsf{spine}_n,\infty)\alpha^{-n-1}$ is the rescaled time to percolate along the backbone from generation n to ∞, and $\mathbf{W} = \{W_n : n\in\mathbb{Z}_+\}$. In this section we will show $\mathbf{W}$ is an ergodic Markov chain which stochastically decreases to its unique invariant measure. Moreover there is an exponentially fast coupling mechanism for the chain.

Notation. For each $x\in\mathbb{B}$, set

$$W(x) = T(x,\infty)\alpha^{-|x|-1} = \alpha^{-1}T_{\mathsf{not}}^{(x)}(\infty)\,.$$

Let F denote the c.d.f. for W_0. Then $W(x)$ is equal in law to W_0 for each $x\in\mathbb{B}$; since W_0 is the minimum of d independent copies of $T(\infty)$, Lemma 4.1 shows that

$$(5.1)\qquad 1-F(t) = \mathbb{P}(W_0>t)\le c_{4.1}(\alpha,d)^d e^{-dt}.$$

Hence we can define a law on $(0,\infty)$, which "tilts" W_0 to the right by α:

$$\Lambda(A) = \mathbb{P}(e^{\alpha W_0} 1(W_0 \in A))/\mathbb{P}(e^{\alpha W_0})\,.$$

Write $\Lambda(t) = \Lambda((0,t))$.

To define the filtration on which the sequence $\mathbf{W}$ is Markov, we introduce two more pieces of notation.

Notation. Let $\bar{\mathcal{F}}_x = \mathcal{E}_x \vee \sigma(W(x))$. Then $x \le y$ implies $\bar{\mathcal{F}}_x \subset \bar{\mathcal{F}}_y$ and if $x \in \mathbb{B}(n)$, then

$$\{\mathsf{spine}_n = x\} = \big\{T(x) + \alpha^{n+1} W(x) < \min\{T(x') + \alpha^{n+1} W(x') : x' \in \mathbb{B}(n) - \{x\}\}\big\} \in \bar{\mathcal{F}}_x.$$

Therefore spine_n is an "$\bar{\mathcal{F}}_x$-stopping point" and we may define a filtration $\{\mathcal{W}_n\}$ by

$$\mathcal{W}_n = \bar{\mathcal{F}}_{\mathsf{spine}_n} = \{A \in \mathcal{F} : A \cap \{\mathsf{spine}_n = x\} \in \bar{\mathcal{F}}_x \ \text{ for all } x \in \mathbb{B}\}.$$

The verification that $\{W_n\}$ is adapted to $\{\mathcal{W}_n\}$ is immediate.

Theorem 5.1. (a) $\mathbf{W} = \{W_n : n \in \mathbb{Z}_+\}$ is a $(\mathcal{W}_n)$-Markov chain such that for Borel sets $B \subseteq (0,\infty)$,

$$\mathbb{P}(W_{n+1} \in B | \mathcal{W}_n) = p(B|W_n) := \Lambda(B|(0, W_n/\alpha))\ .$$

(b) $(W_n, n \in \mathbb{Z}_+)$ is stochastically non-increasing.

Proof. (a) Let $A \in \mathcal{W}_n, x \in \mathbb{B}(n+1)$ and B a Borel subset of $(0,\infty)$. Then

$$\begin{aligned}(5.2)\qquad &\mathbb{P}(W_{n+1} \in B, A, \mathsf{spine}_{n+1} = x)\\ &\quad = \mathbb{P}\Big[1(A, \mathsf{spine}_n = x|n)\mathbb{P}(W(x) \in B,\\ &\qquad\quad U_x + \alpha W(x) < \min_{1\le i\le d-1} U_{\mathsf{sib}_i(x)} + \alpha W(\mathsf{sib}_i(x)) \ \big|\ \bar{\mathcal{F}}_{x|n})\Big].\end{aligned}$$

Both $W(x|n)$ and the event in the conditional probability are $\mathcal{F}_{(x|n,\infty)}$-measurable. Since $\mathcal{F}_{(x|n,\infty)}$ and $\mathcal{E}_{x|n}$ are independent, while $\bar{\mathcal{F}}_{x|n} = \mathcal{E}_{x|n} \vee \sigma(W(x|n))$, it follows that (5.2) equals

$$\begin{aligned}(5.3)\qquad &\mathbb{P}\Big[1(A, \mathsf{spine}_n = x|n)\mathbb{P}(W(x) \in B,\\ &\qquad U_x + \alpha W(x) < \min_{1\le i\le d-1}(U_{\mathsf{sib}_i(x)} + \alpha W(\mathsf{sib}_i(x))) \ \big|\ W(x|n))\Big].\end{aligned}$$

Let $\mathbb{B}(1) = \{z_1, \ldots, z_d\}$, and set

$$\begin{aligned}q(\cdot\,|w) = d\cdot\mathbb{P}(W(z_d) \in \cdot, U_{z_d} + \alpha W(z_d) < &\min_{1\le i\le d-1}(U_{z_i} + \alpha W(z_i))\\ &\big|\ \min_{1\le i\le d}(U_{z_i} + \alpha W(z_i)) = w)\,.\end{aligned}$$

In other words, $q(\cdot\,|w)$ is a regular conditional probability for the right side. The collections $(W(x), U_x, ((W(\mathsf{sib}_i(x)), U_{\mathsf{sib}_i(x)}) : i < d))$ and $(W(z_d), U_{z_d}, ((W(z_i), U_{z_i}) : i < d))$ are equal in law, and

$$W(x|n) = \min\{U_x + \alpha W(x), U_{\mathsf{sib}_i(x)} + \alpha W(\mathsf{sib}_i(x)) : i < d\}.$$

Therefore, using this in (5.3), we deduce that

$$\begin{aligned}\mathbb{P}(W_{n+1} \in B, A, \mathsf{spine}_{n+1} = x) &= d^{-1}\mathbb{P}\big(1(A, \mathsf{spine}_n = x|n) q(B|W(x|n))\big)\\ &= d^{-1}\mathbb{P}\big(1(A, \mathsf{spine}_n = x|n) q(B|W_n)\big).\end{aligned}$$

Sum over $x \in \mathbb{B}(n+1)$ to conclude

$$\mathbb{P}(W_{n+1} \in B \,|\, \mathcal{W}_n) = q(B|W_n) \ \text{ a.s.} \tag{5.4}$$

If $h(y) = \mathbb{P}(\min_{1\le i\le d-1}(U_{z_i} + \alpha W(z_i)) > y)$ then clearly for $A \in \mathcal{B}((0,\infty))$,

$$\begin{aligned} d\mathbb{P}\big(W(z_d) \in B, \min_{1\le i\le d}&(U_{z_i} + \alpha W(z_i)) \in A, U_{z_d} + \alpha W(z_d) < \min_{1\le i\le d-1}(U_{z_i} + \alpha W(z_i))\big)\\ &= d\int\int 1(w \in B, u + \alpha w \in A) h(u + \alpha w) e^{-u} du dF(w)\\ &= d\int\int 1(w \in B, y \in A) 1(w \le y/\alpha) h(y) e^{-y+\alpha w} dF(w) dy\\ &= \int_A \psi(y) \Lambda(B|(0, y/\alpha]) dy,\end{aligned}$$

where $\psi(y) = dh(y)e^{-y}\Lambda(y/\alpha)$. Take $B = (0,\infty)$ to see that ψ is the density of $\min_{i\le d}(U_{z_i} + \alpha W(z_i))$. It follows that $\Lambda(\cdot\,|(0, y/\alpha])$ is a version of $q(\cdot\,|y)$ and (a) is a consequence of (5.4).

(b) We have

$$\begin{aligned}\mathbb{P}(W_1 \ge t) &= d\mathbb{P}(1(W(z_d) \ge t)\mathbb{P}(U_{z_d} + \alpha W(z_d) < \min_{1\le i<d}(U_{z_i} + \alpha W(z_i))|W(z_d)))\\ &\le d\mathbb{P}(W(z_d) \ge t)\mathbb{P}(U_{z_d} + \alpha W(z_d) < \min_{1\le i<d}(U_{z_i} + \alpha W(z_i))) \quad \text{(Lemma 3.5)}\\ &= \mathbb{P}(W(z_d) \ge t) \quad \text{(symmetry)}\\ &= \mathbb{P}(W_0 \ge t).\end{aligned}$$

This proves $W_1 \prec W_0$. The result now follows by induction and the fact that $p(\cdot\,|w)$ is stochastically non-decreasing in w. $\square$

Since W_n decreases stochastically in n, the laws converge weakly, and we shall now show that the limit is nontrivial. The other goals for the remainder of this section are to understand the stationary distribution for $\{W_n\}$ (including existence and convergence to stationarity) and to prove an exponential rate of convergence through a coupling mechanism. The first goal is achieved in the next result.

Theorem 5.2. (a) $\mathbf{W} = \{W_n : n \in \mathbb{Z}_+\}$ converges to a stationary distribution π on $(0,\infty)$.

(b) π is the unique stationary distribution for $\mathbf{W}$ and the stationary chain is ergodic.
(c) There are positive constants $c_{5.1}(p)(p>0)$, and $c_{5.2}$ such that for all $w \geq 0$ and for all $p>0$

$$\begin{aligned}\pi((0,w]) &\leq c_{5.1}(p)w^p \\ \pi([w,\infty)) &\leq c_{5.2}e^{-dw}.\end{aligned} \tag{5.5}$$

We begin the proof with an integral equation that leads to an important bound (5.9) on the left tail of F. Let $F_1(t)=\mathbb{P}(T(\infty)\leq t)$ and $g(x)=1-(1-x)^d$. W_0 is the minimum of d independent copies of $T(\infty)$ and so

$$F(t)=g(F_1(t))\in[F_1(t),dF_1(t)]. \tag{5.6}$$

By conditioning on U_0 we also have

$$F_1(t)=\alpha e^{-t}\int_0^{t/\alpha}e^{\alpha s}F(s)ds\leq F(t/\alpha)(1-e^{-t}). \tag{5.7}$$

Equations (5.6) and (5.7), together with the easy fact that $F_1(t)>0$ for positive t, imply that F and F_1 are infinitely differentiable and satisfy

$$F_1'(t)=-\alpha e^{-t}\int_0^{t/\alpha}e^{\alpha s}F(s)ds+F(t/\alpha),\qquad F'(t)=g'(F_1(t))F_1'(t), \tag{5.8}$$

$$0<F(t)/F(t/\alpha)\leq d(1-e^{-t})\leq td\quad\text{for all } t>0. \tag{5.9}$$

Next, we let β be the midpoint between 1 and α^{-1}, and show that when W_n is small, the probability is at least $1/4$ that $W_{n+1}\geq W_n\beta$.

Lemma 5.3. (a) There is a $t_0>0$ such that $F(t\beta)/F(t/\alpha)<\frac{2}{3}$ and $\Lambda(t\beta)/\Lambda(t/\alpha)<\frac{3}{4}$ for $t\in(0,t_0)$.
(b) For all $0<w\leq 1$, $\Lambda(w)/\Lambda(w/\alpha)\leq dew$.

Proof. Differentiate in (5.8) to see

$$F''(t)=g''(F_1(t))F_1'(t)^2+g'(F_1(t))F_1''(t), \tag{5.10}$$

and as $t\downarrow 0$,

$$\begin{aligned}F_1''(t)&=o(F(t/\alpha))-F(t/\alpha)+F'(t/\alpha)\alpha^{-1}\\&=o(F(t/\alpha))-F(t/\alpha)+g'(F_1(t/\alpha))F_1'(t/\alpha)\alpha^{-1}\\&=o(F(t/\alpha))-F(t/\alpha)+g'(F_1(t/\alpha))\alpha^{-1}(o(F(t/\alpha^2))+F(t/\alpha^2))\quad\text{(from (5.8))}\\&=o(F(t/\alpha^2))-F(t/\alpha)+d\alpha^{-1}F(t/\alpha^2).\end{aligned}$$

Use the latter in (5.10) to conclude that as $t \downarrow 0$

$$\begin{aligned}F''(t) &= o(F(t/\alpha)) + d(o(F(t/\alpha^2)) - F(t/\alpha) + d\alpha^{-1}F(t/\alpha^2)) \quad ((5.8) \text{ again})\\ &= o(F(t/\alpha^2)) + d^2\alpha^{-1}F(t/\alpha^2) - dF(t/\alpha)\\ &> 0 \quad \text{for small } t.\end{aligned}$$

Therefore F is convex near 0 and for $t \in (0, t_0)$ we have

$$\begin{aligned}(F(t/\alpha) - F(t\beta))/F(t/\alpha) \geq \tfrac{1}{2}(F(t/\alpha) - F(t))/F(t/\alpha) &\geq \tfrac{1}{2}(1 - dt) \quad (\text{by } (5.9))\\ &> \tfrac{1}{3} \quad (\text{for } t_0 \text{ small enough}).\end{aligned}$$

For the second inequality in (a) note that for $t \in (0, t_0)$,

$$\Lambda(t\beta)/\Lambda(t/\alpha) \leq F(t\beta)e^t/F(t/\alpha) < \tfrac{3}{4},$$

where we have taken t_0 sufficiently small for the last inequality. This proves (a), and (b) is a trivial consequence of (5.9), just as above. □

Proof of Theorem 5.2. (a), (c). The stochastic monotonicity in Theorem 5.1 shows that (W_n) converges in distribution to a law π on $[0, \infty)$. We must show $\pi(\{0\}) = 0$. If $t_0 > 0$ is as in Lemma 5.3, then for $w \in (0, t_0 \wedge 1]$,

$$\begin{aligned}\mathbb{P}(W_n \leq w) &\leq \mathbb{P}(W_{n+1} \leq w)\\ &= \mathbb{P}(\Lambda((W_n/\alpha) \wedge w)\Lambda(W_n/\alpha)^{-1})\\ &\leq \mathbb{P}(W_n \leq \beta^{-1}w) + \mathbb{P}(1(\beta^{-1}w < W_n \leq w)\Lambda(\beta W_n)\Lambda(W_n/\alpha)^{-1})\\ &\qquad + \mathbb{P}(1(W_n > w)\Lambda(w)\Lambda(w/\alpha)^{-1})\\ &\leq \mathbb{P}(W_n \leq \beta^{-1}w) + \tfrac{3}{4}\mathbb{P}(\beta^{-1}w < W_n \leq w) + dewP(W_n > w),\end{aligned}$$

by Lemma 5.3. Rearrange terms to conclude from the above that

$$\mathbb{P}(\beta^{-1}w < W_n \leq w) \leq 4dew = cw \quad \text{for } w \in (0, t_0 \wedge 1].$$

Iterating the above we find that for $k \geq k_0$,

$$\mathbb{P}(0 < W_n \leq \beta^{-k}) = \sum_{j=k}^{\infty} c\beta^{-j} = c(1 - \beta^{-1})^{-1}\beta^{-k},$$

and therefore

$$\mathbb{P}(0 \leq W_n \leq w) \leq c\beta(1 - \beta^{-1})^{-1}w = c_{5.1}w \qquad \text{for } 0 < w < w_0. \tag{5.11}$$

Let $n \to \infty$ to get the first inequality in (5.5) for $p = 1$, first for w a continuity point of $\pi((0, w])$ in $(0, w_0)$, and then (increase $c_{5.1}$ if necessary) for all $w > 0$. This proves the first part of (c) for $p = 1$ and also shows $\pi(\{0\}) = 0$. Since F is atomless, so is Λ, and $p(\cdot\,|w)$ is weakly continuous in w. Taking limits in $\mathbb{P}(W_{n+1} \in \cdot) = \mathbb{P}(p(\cdot\,|W_n))$, we see that π is a stationary law. The upper bound on $\pi([w,\infty))$ is immediate from (5.1) and $\pi \prec F$ (again increase $c_{5.1}$ if need be).

To prove the upper bound on the left tail of π for general p we consider $p = 2$. The induction argument which will give the result for general p in $\mathbb{N}$ will then be clear. The fact that π is stationary implies that for $w \in (0, t_0)$,

$$\begin{aligned}
\pi\left([0,w]\right) &= \int_0^\infty \frac{\Lambda\left(\frac{x}{\alpha}\wedge w\right)}{\Lambda\left(\frac{x}{\alpha}\right)}\, d\pi(x) \\
&\le \pi\left([0,\beta^{-1}w]\right) + \int 1(\beta^{-1}w < x \le w)\, \frac{\Lambda(w)}{\Lambda(\frac{x}{\alpha})}\, d\pi(x) \\
&\quad + \int 1(w < x \le \frac{w}{\alpha})\, \frac{\Lambda(w)}{\Lambda(\frac{x}{\alpha})}\, d\pi(x) + \int 1(\frac{w}{\alpha} < x)\, \frac{\Lambda(w)}{\Lambda(\frac{x}{\alpha})}\, d\pi(x),
\end{aligned}$$

and therefore,

$$\begin{aligned}
\pi\left((\beta^{-1}w,w]\right) &\le \frac{\Lambda(w)}{\Lambda(w\beta^{-1}\alpha^{-1})}\, \pi\left((\beta^{-1}w,w]\right) + \frac{\Lambda(w)}{\Lambda(\frac{w}{\alpha})}\, \pi\left((w,\frac{w}{\alpha}]\right) + \frac{\Lambda(w)}{\Lambda(\frac{w}{\alpha^2})} \\
&\le \tfrac{3}{4}\, \pi\left((\beta^{-1}w,w]\right) + (dew)c_{5.1}(1)\frac{w}{\alpha} + (de)^2\, \frac{w^2}{\alpha},
\end{aligned}$$

using Lemma 5.3 and the $p = 1$ case. The above implies

$$\pi\left((\beta^{-1}w,w]\right) \le 4\left(dec_{5.1}(1) + (de)^2\right)\alpha^{-1}w^2$$

and the result for $p = 2$ now follows as for $p = 1$ in the above.
(b) (5.8) and (5.9) show that $F_1'(t) \ge e^{-t}F(t/\alpha) > 0$ for all $t > 0$. This and the second part of (5.8) imply $F'(t) > 0$ and hence $\Lambda'(t) > 0$ for all $t > 0$. Therefore $p(\cdot\,|w)$ has a strictly positive continous density on $(0, w/\alpha]$. It follows easily that (W_n) is an indecomposable Markov chain and (b) is a consequence of (a) and Theorem 7.16 of Breiman (1968). $\square$

There is a natural coupling technique for $\mathbf{W}$ which, in addition to refining the above convergence result, will also play an important role in the limit theorems of Section 6. Let $D = \{(w_1,w_2) : w_1 \ge w_2 > 0\}$ and define a Markov kernel on D by (set $0/0 \equiv 0$)

$$\begin{aligned}
\bar{p}(A_1 \times A_2|(w_1,w_2)) &= \int_0^{w_2/\alpha} 1_{A_1\times A_2}(x,x)d\Lambda(x)\Lambda(w_2/\alpha)^{-1}(\Lambda(w_2/\alpha)/\Lambda(w_1/\alpha)) \\
&\quad + \int_{w_2/\alpha}^{w_1/\alpha} 1_{A_1}(x_1)d\Lambda(x_1)(\Lambda(w_1/\alpha) - \Lambda(w_2/\alpha))^{-1} \\
&\quad \times \int_0^{w_2/\alpha} 1_{A_2}(x_2)d\Lambda(x_2)\Lambda(w_2/\alpha)^{-1}(1 - (\Lambda(w_2/\alpha)/\Lambda(w_1/\alpha))).
\end{aligned}$$

If $\bar{p}_i(\cdot\,|(w_1,w_2))$ is the i^{th} marginal of $\bar{p}(\cdot\,|(w_1,w_2))$ $(i = 1,2)$, then

$$
\begin{aligned}
\bar{p}_1(A_1|(w_1,w_2)) &= \int_0^{w_2/\alpha} 1_{A_1}(x)d\Lambda(x)\Lambda(w_1/\alpha)^{-1} + \int_{w_2/\alpha}^{w_1/\alpha} 1_{A_1}(x)d\Lambda(x)\Lambda(w_1/\alpha)^{-1} \\
&= p(A_1|w_1),
\end{aligned}
$$

and similarly one sees that $\bar{p}_2(A_2|(w_1,w_2)) = p(A_2|w_2)$. We extend this coupling of $\mathbf{W}$ to a Markov kernel on $(0,\infty)^2$ by setting

$$
\bar{p}(A_1 \times A_2|(w_1,w_2)) = \bar{p}(A_2 \times A_1|(w_2,w_1)) \quad \text{if } w_2 > w_1.
$$

Then $\bar{p}$ is a Markov kernel on $(0,\infty)^2$ with marginals $\bar{p}_i(\cdot\,|(w_1,w_2)) = p(\cdot\,|w_i)$ and so the induced Markov chain $((W_n^1, W_n^2), n \in \mathbb{Z}_+)$ will be a coupling of (W_n). Clearly $W_0^1 \geq W_0^2$ (respectively, $W_0^1 \leq W_0^2$) implies $W_n^1 \geq W_n^2$ (respectively $W_n^1 \leq W_n^2$) for all $n \geq 0$ a.s., and if $\tau_{\mathsf{couple}} = \min\{n : W_n^1 = W_n^2\}$ then $W_n^1 = W_n^2$ for all $n \geq \tau_{\mathsf{couple}}$ a.s.

If $W_0^1 \geq W_0^2$, the chains will couple at the first time n for which $W_n^1 \leq W_{n-1}^2/\alpha$ and hence it is possible for W_n^2 to jump onto W_n^1. Unfortunately if W_{n-1}^2 is small the probability of $W_n^1 \leq W_{n-1}^2/\alpha$ will be small, and to get a good coupling rate we must bound the time spent by W_n^2 in $(0,\delta]$ for δ small. We start with a stochastic lower bound on the left tail of the Markov kernel of $\mathbf{W}$.

Lemma 5.4. There is a $w_0 \in (0,1]$ and a probability ρ on $(0,\alpha^{-1}]$ such that

$$
p((0,wx]\,|w) \leq \rho((0,x]) \quad \text{for all } x \geq 0 \text{ and } w \in (0,w_0], \tag{5.12}
$$

$$
\int \log x d\rho(x) = m > 0, \tag{5.13}
$$

$$
\rho((0,x]) \leq w_0 dex \leq x \quad \text{for all } x < \alpha^2. \tag{5.14}
$$

Proof. Let $w_1 = (2de)^{-1}$ and define ρ on $(0,w_1] \times (0,1)$ by

$$
\rho(w,x) = \prod_{k=0}^{n-1}(de\alpha^k w) = (dew)^n \alpha^{n(n-1)/2} \quad \text{if} \quad \alpha^n \leq x < \alpha^{n-1}, n \in \mathbb{N}.
$$

Then ρ is non-decreasing in each variable, $\rho \leq 1/2$, $\rho(w,\cdot)$ is right-continuous, and $\rho(w,0+) = 0 \equiv \rho(w,0)$. Let t_0 be as in Lemma 5.3, $0 < w_0 \leq w_1 \wedge t_0$, and define

$$
\rho(x) = \begin{cases} \rho(w_0,x) & \text{if } x < 1, \\ \frac{3}{4} & \text{if } 1 \leq x < \beta, \\ 1 & \text{if } \beta \leq x \leq \alpha^{-1}. \end{cases}
$$

ρ is the distribution function of a law (also denoted by ρ) on $(0,\alpha^{-1}]$. If $(x,w) \in (0,1) \times (0,w_0]$ and $n \in \mathbb{N}$ satisfies $\alpha^n \leq x < \alpha^{n-1}$, then

$$
\begin{aligned}
\Lambda(xw) \leq \Lambda(\alpha^{n-1}w) &\leq \prod_{k=0}^{n-1}(de\alpha^k w)\Lambda(\alpha^{-1}w) \quad (\text{ by Lemma 5.3(b) }) \\
&= \rho(w,x)\Lambda(w/\alpha) \leq \rho(x)\Lambda(w/\alpha).
\end{aligned}
$$

Therefore

$$p((0,wx]|w) = \Lambda(xw)/\Lambda(w/\alpha) \le \rho(x) \quad \text{for } (x,w) \in (0,1) \times (0,w_0].$$

For $(x,w) \in [1,\beta) \times (0,w_0]$ we have, from Lemma 5.3(a),

$$p((0,wx]|w) = \Lambda(xw)/\Lambda(w/\alpha) \le \Lambda(\beta w)/\Lambda(w/\alpha) \le \tfrac{3}{4} = \rho(x),$$

which proves (5.12).

If $\alpha^n \le x < \alpha^{n-1}$ for some $n \ge 3$, then

$$\begin{aligned}\rho(x) = \rho(w_0, x) &= (dew_0)^n \ \alpha^{n(n-1)/2} \\ &\le (dew_0) x^{(n-1)/2} \le (dew_0)x,\end{aligned}$$

which proves (5.14).

Finally,

$$\begin{aligned}\int_0^{1/\alpha} \log x d\rho(x) &\ge \int_0^{\alpha^2} (\log x)(dew_0)dx + (\log\alpha^2)\rho(w_0, 1-) + (\log\beta)/4 \\ &= (dew_0)\int_0^{\alpha^2} (\log x)dx + (\log\alpha^2)dew_0 + (\log\beta)/4,\end{aligned}$$

which is positive if we choose w_0 sufficiently small. □

For $w > 0$, let $\mathbb{P}_w$ be the law of $(W_n, n \in \mathbb{Z}_+)$ starting at $W_0 = w$, and for $(w_1, w_2) \in [0,\infty)$, let $\mathbb{P}_{w_1,w_2}$ be the law of $((W_n^1, W_n^2), n \in \mathbb{Z}_+)$ starting at $(W_0^1, W_0^2) = (w_1, w_2)$. If μ is a law on $(0,\infty)$ or $(0,\infty)^2$, then write $\mathbb{P}_\mu$ for the law of the appropriate chain (W or (W^1, W^2)) with initial distribution μ.

Lemma 5.5. For all $\theta, \varepsilon > 0$, there exist $\delta_{5.1}, c_{5.3}, \lambda_{5.1} > 0$ (depending on (θ,ε)) such that

$$\mathbb{P}_w(\sum_{j=0}^n 1(W_j \le \delta_{5.1}) > \varepsilon n) \le c_{5.3}(w^{-\theta} + 1)e^{-\lambda_{5.1} n} \quad \text{for all } n \in \mathbb{Z}_+, w > 0.$$

Proof. Define a Markov kernel q on $(0,\infty)$ by

$$q((0,y]|w) = \begin{cases} \rho((0,y/w]) & \text{if } w \le w_0 \\ p((0,y]|w_0) & \text{if } w > w_0. \end{cases}$$

Lemma 5.4 implies that $q(\cdot\,|w) \prec p(\cdot\,|w)$ for all $w > 0$. This, the stochastic monotonicity of $p(\cdot\,|w)$, and a standard coupling argument (Kamae et al. (1977, Theorem 2)) show that for each $w > 0$ we may construct Markov chains (W_n) and (X_n) with transition kernels p and q, respectively, on the same probability space such that $X_0 = W_0 = w$ and $W_n \ge X_n$ for all $n \ge 0$. We abuse notation slightly and let $\mathbb{P}_w$ denote the underlying probability. Define a sequence of stopping times $(T_j, j \in \mathbb{Z}_+)$ by

$$\begin{aligned}T_0 &= \min\{k \ge 0 : X_k > w_0\} \\ T_{j+1} &= \min\{k > T_j : X_k > w_0\}.\end{aligned}$$

Let $\{V_j : j \in \mathbb{N}\}$ be i.i.d. random variables, on some $(\Omega, \mathcal{F}, \mathbb{P})$, with law ρ, and set $S_n = \sum_{j=1}^n \log V_j$. Let $T_0'(w) = \min\{k \geq 0 : w \prod_{j=1}^k V_j > w_0\}$, $w > 0$. The definition of q implies that

$$\begin{aligned}\mathbb{P}_w(((\log X_{k\wedge T_0}, k \geq 0), T_0) \in \cdot\,) \\ = \mathbb{P}(((\log w + S_{k\wedge T_0'}, k \geq 0), T_0'(w)) \in \cdot) \text{ for all } w > 0.\end{aligned} \tag{5.15}$$

(5.14) implies $\gamma(\theta) \equiv \mathbb{P}(\exp(-\theta \log V_1)) < \infty$ for $\theta < 1$ and by (5.13) $\gamma'(0) = -\mathbb{P}(\log V_1) < 0$. Choose $0 < \theta < 1$ sufficiently small ($0 < \theta \leq \theta_0$ say) so that $\gamma(\theta) < 1$. Then

$$\begin{aligned}\mathbb{P}_w(T_0 > n) &= \mathbb{P}(\log w + S_k \leq \log w_0 \ \forall k \leq n) \quad (\text{ by } (5.15)) \\ &\leq \mathbb{P}(\exp(-\theta S_n))(w_0/w)^\theta \\ &\leq w^{-\theta}\gamma(\theta)^n \quad (\text{ recall } w_0 \leq 1).\end{aligned} \tag{5.16}$$

The strong Markov property of X now shows that $T_j < \infty$ for all $j \geq 0$ a.s.

If $\delta > 0$ define

$$Y_j = \sum_{k=T_{j-1}+1}^{T_j} 1(X_k \leq \delta) \ , j \in \mathbb{N}; \qquad Y_0 = \sum_{k=0}^{T_0} 1(X_k \leq \delta).$$

For $j \in \mathbb{N}$, the strong Markov property of X implies ($\mathcal{G}_j = \sigma(X_0, \ldots, X_j)$)

$$\begin{aligned}\mathbb{P}_w(Y_j \in A | \mathcal{G}_{T_{j-1}}) &= \mathbb{P}_{X(T_{j-1})}(\mathbb{P}_{X(1)}(Y_0 \in A)) \\ &= \int \mathbb{P}_x(Y_0 \in A) p(dx|w_0) \quad (\text{ since } X_{T_{j-1}} > w_0).\end{aligned} \tag{5.17}$$

Therefore $\{Y_j : j \in \mathbb{N}\}$ are i.i.d. and

$$\begin{aligned}\mathbb{P}_w(Y_j) &= \int \mathbb{P}_x(Y_0) p(dx|w_0) \\ &\leq \int \mathbb{P}\Big(\sum_{k=0}^\infty 1(\log x + S_k < \log \delta)\Big) p(dx|w_0) \quad (\text{ by } (5.15)) \\ &\leq \sum_{k=0}^\infty \int \mathbb{P}(e^{-\theta S_k}) \delta^\theta x^{-\theta} p(dx|w_0) \\ &= (1-\gamma(\theta))^{-1} \int x^{-\theta} p(dx|w_0) \delta^\theta \ = c(\theta)\delta^\theta,\end{aligned} \tag{5.18}$$

where $c(\theta) < \infty$ by (5.12) and (5.14). Use (5.17) and (5.16) to show that for sufficiently small positive β

$$\mathbb{P}_w(e^{\beta Y_j}) = \int \mathbb{P}_x(e^{\beta Y_0}) p(dx|w_0) \leq e^\beta \int \mathbb{P}_x(e^{\beta T_0}) p(dx|w_0) < \infty.$$

(5.18) implies $\frac{d}{d\beta}\mathbb{P}_w(e^{\beta Y_j})|_{\beta=0} \le c(\theta)\delta^\theta$ and so for some $\beta = \beta(\theta) > 0$ we have

$$\mathbb{P}_w(e^{\beta Y_j}) \le 1 + 2c(\theta)\beta\delta^\theta \le \exp(2c(\theta)\beta\delta^\theta). \tag{5.19}$$

Given $\varepsilon > 0$ choose $\delta > 0$ so that $c(\theta)\beta\delta^\theta \le \varepsilon/8$. Then

$$\begin{aligned}
\mathbb{P}_w(\sum_{j=0}^n 1(W_j \le \delta) > \varepsilon n) &\le \mathbb{P}_w(\sum_{j=0}^n 1(X_j \le \delta) > \varepsilon n)\\
&\le \mathbb{P}_w(T_0 + 1 + \sum_{j=1}^n Y_j > \varepsilon n)\\
&\le \mathbb{P}_w(T_0 > \varepsilon n/2 - 1) + \exp(-\beta\varepsilon n/2)\mathbb{P}_w(e^{\beta Y_1})^n\\
&\le w^{-\theta}\gamma(\theta)^{[\varepsilon n/2]-1} + \exp(-\beta\varepsilon n/2 + 2c(\theta)\beta\delta^\theta n)\\
&\qquad (\text{ by } \ (5.16) \text{ and } (5.19))\\
&\le c(w^{-\theta}+1)e^{-\lambda n}
\end{aligned}$$

for some $c = c(\theta, \varepsilon) > 0$ and $\lambda = \lambda(\theta, \varepsilon) > 0$ (by the choice of δ). This proves the result for $0 < \theta \le \theta_0$, and it follows trivially for all $\theta > 0$. ☐

Proposition 5.6. For any $\theta > 0$ there are $c_{4.4}(\theta), \lambda_{5.2}(\theta) > 0$ such that the coupling time τ_{couple} satisfies

$$\mathbb{P}_{w_1,w_2}(\tau_{\mathsf{couple}} > n) \le c_{5.4}((w_1 \wedge w_2)^{-\theta} + 1)e^{-\lambda_{5.2}n} \quad \text{for all } n \in \mathbb{Z}_+,\ (w_1, w_2) \in (0,\infty)^2.$$

In particular, $\tau_{\mathsf{couple}} < \infty$ $\mathbb{P}_\mu-$ a.s. for all laws μ on $(0,\infty)^2$.

Proof. Fix $0 < w_2 \le w_1$, write $\mathbb{P}$ for $\mathbb{P}_{w_1,w_2}$, and let $\mathcal{G}_n = \sigma((W_j^1, W_j^2), j \le n)$. Note that $M(n) = 1(\tau_{\mathsf{couple}} > n)\exp(\sum_{j=0}^{n-1}\Lambda(W_j^2/\alpha))$ is a $(\mathcal{G}_n)$-supermartingale because w.p. 1.

$$\begin{aligned}
\mathbb{P}(M(n+1)|\mathcal{G}_n) &= M(n)\exp(\Lambda(W_n^2/\alpha))\mathbb{P}(\tau_{\mathsf{couple}} > n+1|\mathcal{G}_n)\\
&= M(n)\exp(\Lambda(W_n^2/\alpha))(1 - \Lambda(W_n^2/\alpha)/\Lambda(W_n^1/\alpha)) \ \ (\text{by the definition of } \bar{p})\\
&\le M(n).
\end{aligned}$$

If $\theta > 0$ and $\varepsilon = 1/2$ in Lemma 5.5, then that result gives (for $\delta = \delta_{5.1}$)

$$\begin{aligned}
\mathbb{P}(\tau_{\mathsf{couple}} > n) &\le \mathbb{P}(\tau_{\mathsf{couple}} > n, \sum_{j=0}^{n-1} 1(W_j^2 \le \delta) \le n/2) + \mathbb{P}(\sum_{j=0}^{n-1} 1(W_j^2 \le \delta) > n/2)\\
&\le \mathbb{P}(M(n))\exp(-\Lambda(\delta/\alpha)n/2) + c_{5.3}(w_2^{-\theta}+1)e^{-\lambda_{5.1}n}\\
&\le \exp(\Lambda(w_2/\alpha) - \Lambda(\delta/\alpha)n/2) + c_{5.3}(w_2^{-\theta}+1)e^{-\lambda_{5.1}n}.
\end{aligned}$$

The result follows because $\Lambda(\delta/\alpha) > 0$ (by (5.9)) and $\Lambda(w_2/\alpha) \le 1$. ☐

Corollary 5.7. (a) There are Markov chains $(W_n, n \in \mathbb{Z}_+)$ with law $\mathbb{P}_F$ and $(\overline{W}_n, n \in \mathbb{Z}_+)$ with law $\mathbb{P}_\pi$, defined on the same probability space, such that if

$$\tau_{\mathsf{couple}} = \min\{n \in \mathbb{Z}_+ : W_n = \overline{W}_n\}$$

then:

(i) $W_n = \overline{W}_n$ for all $n \geq \tau_{\mathsf{couple}}$,

(ii) $\mathbb{P}(\tau_{\mathsf{couple}} > n) \leq c_{5.5} e^{-\lambda_{5.3} n}$ for all $n \in \mathbb{Z}_+$ for some $c_{5.5}, \lambda_{5.3} > 0$.

(b) If $W_n = T(X|n,\infty)\alpha^{-n-1}$ and $|\mathcal{L}(W_n) - \pi|$ denotes the total variation distance between the law of W_n and its weak limit π then $|\mathcal{L}(W_n) - \pi| \leq 2c_{5.5} e^{-\lambda_{5.3} n}$, $n \geq 0$.

Proof. (a) Since $F \succ \pi$ by Theorem 5.1(b), there is a law μ on $D = \{(w_1, w_2) : w_1 \geq w_2 > 0\}$ with first and second marginals F and π, respectively. Let $(W_n, \overline{W}_n)$ be a chain with law $\mathbb{P}_\mu$. Then by Proposition 5.6 with $\theta = \frac{1}{2}$

$$\mathbb{P}_\mu(\tau_{\mathsf{couple}} > n) \leq c_{5.4} e^{-\lambda_{5.2}(1/2)n} \int w^{-1/2} + 1 d\pi(w) \leq c_{5.5} e^{-\lambda_{5.3} n} \quad \text{(by (5.5))}.$$

(b) is immediate because (W_n) has law $\mathbb{P}_F$. □

The same argument as in (a) also gives the following corollary.

Corollary 5.8. There exists $c_{5.6} > 0$ such that

$$\mathbb{P}_{\pi\times\pi}(\tau_{\mathsf{couple}} > n) \leq c_{5.6} e^{-\lambda_{5.3} n} \quad \text{for all } n \in \mathbb{Z}_+ . \tag{5.20}$$

6. Limit Theorems

In this section we use the results of the previous sections to prove a Law of Large Numbers (Theorem 6.1) and a Central Limit Theorem (Theorem 6.2), for functionals of the clusters which branch off the backbone. The next section then lists specific limit theorems that follow from these. Recall the setting of Theorem 2.5: the backbone is $\mathsf{spine}_1, \mathsf{spine}_2, \ldots$, the clusters off the backbone, and associated percolation times to infinity are given by $Y_{n,j} = (\mathsf{Clust}_{n,j}, a_{n,j})$, and are conditionally independent with distributions ν_{W_n}. Let $\vec{Y}_n$ denote the vector $(Y_{n,1}, \ldots, Y_{n,d-1}) \in \mathcal{S}_0^{d-1}$.

Notation. If $\varphi : \mathcal{S}_0^{d-1} \to \mathbb{R}$, let

$$\mu_t(\varphi) = \int_{\mathcal{S}_0^{d-1}} \varphi(y_1, \ldots, y_{d-1}) \prod_{j=1}^{d-1} \nu_t(dy_j),$$
$$\bar{\mu}(\varphi) = \int_0^\infty \mu_t(\varphi) d\pi(t).$$

Thus μ_t is the law of $(d-1)$ independent copies of ν_t.

Theorem 6.1. If $\varphi : \mathcal{S}_0^{d-1} \to \mathbb{R}$ is measurable and $\bar{\mu}(\varphi)$ is finite, then

$$\lim_{n\to\infty} n^{-1} \sum_{j=0}^{n-1} \varphi(\vec{Y}_j) = \bar{\mu}(\varphi) \quad \text{a.s..}$$

Theorem 6.2. Suppose that

$$\int_0^\infty \mu_t(\varphi^2)^2 \log^+(\mu_t(\varphi^2))d\pi(t) < \infty. \tag{6.1}$$

Then

$$n^{-1/2} \sum_{j=0}^{n-1} (\varphi(\vec{Y}_j) - \bar{\mu}(\varphi)) \xrightarrow{w} N(0, \sigma_\varphi^2) \quad \text{as } n \to \infty,$$

where

$$\sigma_\varphi^2 = 2\Big(\sum_{j=1}^\infty \mathbb{P}_\pi(\mu_{W_j}(\varphi)\mu_{W_0}(\varphi) - \bar{\mu}(\varphi)^2)\Big) + \bar{\mu}(\varphi^2) - \bar{\mu}(\varphi)^2 < \infty. \tag{6.2}$$

Proof of Theorem 6.1. Let $(W_j, j \in \mathbb{Z}_+)$ and $(\overline{W}_j, j \in \mathbb{Z}_+)$ be the chains with laws $\mathbb{P}_F$ and $\mathbb{P}_\pi$, respectively, which are coupled as in Corollary 5.7 (a). Let $\{U_j : j \in \mathbb{Z}_+\}$ be an independent sequence of i.i.d. random variables which are uniformly distributed on $[0, 1]$. Define $G_t(x) = \mu_t(\{(y_1, \ldots, y_{d-1}) : \varphi(y_1, \ldots, y_{d-1}) \le x\})$ and let $V_j(t) = G_t^{-1}(U_j)$ where $G_t^{-1}(u) = \inf\{x : G_t(x) > u\}$. Therefore $\{V_j(t) : j \in \mathbb{Z}_+\}$ are i.i.d. and have distribution $\mu_t(\varphi(\cdot) \in \cdot)$. Moreover $V_j(t, \omega)$ is jointly measurable because $G_t^{-1}(x)$ is. Theorem 2.5 shows that $(V_j(W_j), j \in \mathbb{Z}_+)$ and $(\varphi(Y^{(j)}), j \in \mathbb{Z}_+)$ are equal in law. By the coupling in Corollary 5.7 it suffices to show

$$\lim_{n\to\infty} n^{-1} \sum_{j=0}^{n-1} V_j(\overline{W}_j) = \bar{\mu}(\varphi) \quad \text{a.s.} \tag{6.3}$$

However, $V_j(\overline{W}_j) = G^{-1}_{\overline{W}_j}(U_j)$ and since $\{(\overline{W}_j, U_j), j \in \mathbb{Z}_+\}$ is clearly stationary and ergodic by Theorem 5.2, $\{V_j(\overline{W}_j), j \in \mathbb{Z}_+\}$ is too. $\bar{\mu}(\varphi)$ is the mean of $V_j(\overline{W}_j)$ and therefore the ergodic theorem implies (6.3). $\square$

Proof of Theorem 6.2. Let $\{(W_j, \overline{W}_j, V_j) : j \in \mathbb{Z}_+\}$ be as in the previous proof. As in the above argument it suffices to show

$$\Phi_n = n^{-1/2} \sum_{j=0}^{n-1} \big(V_j(\overline{W}_j) - \overline{\mu}(\varphi)\big) \xrightarrow{w} N(0, \sigma_\varphi^2).$$

Write $\Phi_n = X_n + Z_n$, where

$$X_n = n^{-1/2} \sum_{j=0}^{n-1} \big(V_j(\overline{W}_j) - \mu_{\overline{W}_j}(\varphi)\big) \;\text{ and }\; Z_n = n^{-1/2} \sum_{j=0}^{n-1} \big(\mu_{\overline{W}_j}(\varphi) - \overline{\mu}(\varphi)\big).$$

We will use the Lindeberg Central Limit Theorem to show that, conditional on $\overline{\mathbf{W}} = (\overline{W}_j)$, $X_n \xrightarrow{w} N(0, \sigma_1^2)$ and then use a Central Limit Theorem for stationary ergodic processes to prove that $Z_n \xrightarrow{w} N(0, \sigma_2^2)$. Introduce

$$s_n^2 = \sum_{j=0}^{n-1} \mathbb{P}\Big(\big(V_j(\overline{W}_j) - \mu_{\overline{W}_j}(\varphi)\big)^2 \Big| \overline{\mathbf{W}} \Big) = \sum_{j=0}^{n-1} \mu_{\overline{W}_j}(\varphi^2) - \mu_{\overline{W}_j}(\varphi)^2.$$

The Ergodic Theorem implies

$$\lim_{n\to\infty} s_n^2 n^{-1} = \overline{\mu}(\varphi^2) - \int \mu_t(\varphi)^2 \, d\pi(t) \equiv \sigma_1^2 \quad \text{a.s.} \tag{6.4}$$

We claim that

$$\lim_{n\to\infty} \mathbb{P}(e^{i\theta X_n} | \overline{\mathbf{W}}) = e^{-\theta^2 \sigma_1^2/2} \qquad \text{for all } \theta \in \mathbb{R}. \tag{6.5}$$

If $\sigma_1^2 = 0$ this is clear from (6.4). Assume $\sigma_1^2 > 0$, let $\varepsilon > 0$ and check the Lindeberg condition, conditional on $\overline{\mathbf{W}}$. For any fixed $K > 0$, use (6.4) and the Ergodic Theorem to conclude

$$\begin{aligned}
&\limsup_{n\to\infty} s_n^{-2} \sum_{j=0}^{n-1} \mathbb{P}\left(\big(V_j(\overline{W}_j) - \mu_{\overline{W}_j}(\varphi)\big)^2 1\big(|V_j(\overline{W}_j) - \mu_{\overline{W}_j}(\varphi)| > \varepsilon s_n\big) |\overline{\mathbf{W}} \right) \\
&\quad \le \lim_{n\to\infty} \sigma_1^{-2} n^{-1} \sum_{j=0}^{n-1} \int \big(\varphi(y) - \mu_{\overline{W}_j}(\varphi)\big)^2 1\big(|\varphi(y) - \mu_{\overline{W}_j}(\varphi)| > K\big) \, d\mu_{\overline{W}_j}(y) \text{ a.s.} \\
&\quad = \sigma_1^{-2} \int \int \big(\varphi(y) - \mu_t(\varphi)\big)^2 1\big(|\varphi(y) - \mu_t(\varphi)| > K\big) \, d\mu_t(y) \, d\pi(t) \text{ a.s.}
\end{aligned}$$

The last expression approaches zero as $K \to \infty$ because $\overline{\mu}(\varphi^2) < \infty$ by (6.1). This gives us the Lindeberg condition with respect to $\mathbb{P}(\cdot|\overline{\mathbf{W}})$ a.s., and (6.5) then follows from the Lindeberg Central Limit Theorem and (6.4).

It is easy to use the exponentially fast coupling given by Corollary 5.8 to see that $\{\overline{W}_j\}$ is strongly mixing with an exponential mixing rate (with the notation of Rio (1995)), $\alpha(n) \le 2c_{5.6} \exp(-\lambda_{5.3} n)$. This means that the same is true of the ergodic process $\{\mu_{\overline{W}_j}(\varphi) - \overline{\mu}(\varphi)\}$, and (6.1) allows us to apply a Central Limit Theorem for strongly mixing stationary processes (Theorem 1 of Rio (1995) and (1.5) of that work) to conclude that

$$\lim_{n\to\infty} \mathbb{P}(e^{i\theta Z_n}) = e^{-\theta^2 \sigma_2^2/2}, \tag{6.6}$$

where

$$\sigma_2^2 = \int \mu_t(\varphi)^2 \, d\pi(t) - \overline{\mu}(\varphi)^2 + 2 \sum_{j=1}^{\infty} \mathbb{P}\big(\mu_{\overline{W}_j}(\varphi)\mu_{\overline{W}_0}(\varphi) - \overline{\mu}(\varphi)^2\big).$$

The required result now follows from (6.5) and (6.6) because $\sigma_\varphi^2 = \sigma_1^2 + \sigma_2^2$. $\square$

Example 6.3. Define $\varphi_1 : \mathcal{S}_0^{d-1} \to [0,\infty)$ by $\varphi_1((C_1,a_1),\dots,(C_{d-1},a_{d-1})) = \sum_{i=1}^{d-1} \#C_i$. Then

$$\begin{aligned}\mu_t(\varphi_1^2) &= (d-1)\int (\#C)^2 d\nu_t + (d-1)(d-2)(\int \#C d\nu_t)^2 \\ &\le (d-1)c_{4.8}(\alpha,d) + (d-1)(d-2)c_{4.3}(\alpha,d)^2 \quad \text{(Theorems 4.4 and 4.6).}\end{aligned}$$

If $C_n = \cup_{i=1}^{d-1}\mathsf{Clust}_{n,i}$, then $\varphi_1(Y^{(n)}) = \#C_n$, and Theorems 6.1 and 6.2 show that

$$\lim_{n\to\infty} n^{-1}\sum_{j=0}^{n-1}\#C_j = \bar\mu \equiv (d-1)\int\int \#C d\nu_t d\pi(t) \quad \text{a.s.,} \tag{6.7}$$

and

$$n^{-1/2}\sum_{j=0}^{n-1}(\#C_j - \bar\mu) \xrightarrow{w} N(0,\sigma_{\varphi_1}^2) \quad \text{as } n\to\infty, \tag{6.8}$$

where a (not very helpful) expression for $\sigma_{\varphi_1}^2 > 0$ may be retrieved from Theorem 6.2. The expression for $\bar\mu$ is more tractable and Theorem 4.4(b) and (4.11) imply that

$$\begin{aligned}0 < \bar\mu = \bar\mu(\alpha,d) &\le (d-1)\sup_t \int \#C d\nu_t \\ &\le (d-1)c_{4.3}(\alpha,d) \\ &\le c_{6.1}(d)\exp(c_{4.7}(d)/(1-\alpha)).\end{aligned} \tag{6.9}$$

We now consider some processes which describe aspects of the growth dynamics of $\{A_n\}$. The size of the cluster at the first time a node of height n is filled is $\mathsf{minsize}(n) = \min\{k : h(A_k) = n\}$. The size of the cluster at the last time a node of height n or less is added is $\mathsf{maxsize}(n) = \max\{k : A_k - A_{k-1} \in \mathbb{B}(0,n)\}$. Let $\mathsf{size}(n) = \#(A_\infty \cap \mathbb{B}(0,n))$ be the number of nodes in A_∞ of height n or less. Then clearly

$$\mathsf{minsize}(n) \le \mathsf{size}(n) \le \mathsf{maxsize}(n). \tag{6.10}$$

Note that each point in $A_\infty \cap \mathbb{B}(0,n)$ is in a cluster $\mathsf{Clust}_{k,i}$ for some $k < n$ or is one of the first $n+1$ vertices along the backbone. If $N(n) = \sum_{k=0}^{n-1}\#(C_k)$, this gives

$$\mathsf{size}(n) \le N(n) + n + 1. \tag{6.11}$$

For the next result $\sigma_{\varphi_1}^2$ is defined in Theorem 6.2 with φ_1 as in Example 6.3, and we define

$$\mathsf{rate} = \bar\mu(\alpha,d) + 1 \ > 1. \tag{6.12}$$

Theorem 6.4. For $\{L(n)\} = \{\mathsf{size}(n)\}, \{\mathsf{maxsize}(n)\}$ or $\{\mathsf{minsize}(n)\}$ the following results hold.
(a) $\lim_{n\to\infty} n^{-1}L(n) = \mathsf{rate}$ a.s.

(b) $\quad n^{-1/2}(L(n) - n\mathsf{rate}) \xrightarrow{w} N(0, \sigma^2_{\varphi_1})$ as $n \to \infty$.

It is not hard to prove this by first using the bounds in Theorem 4.4 to show that $\mathsf{maxsize}(n) - \mathsf{minsize}(n)$ and $N(n)+n+1-\mathsf{size}(n)$ remain bounded in probability as $n \to \infty$, and then applying Example 6.3. The interested reader may find this argument in Lemma 6.5 of the earlier version of this work referred to in the introduction. We will prove Theorem 6.4 in the next Section by showing that the above differences are bounded in L^1 by means of a dynamical decomposition of A_∞ into independent blocks.

Let $\ell(n) = \min\{|x| : x \in A_\infty - A_n\}$ be the height of the shortest vertex which is added after step n. Hence $A_k \cap \mathbb{B}(0, \ell(n) - 1)$ is "frozen" for $k > n$. Clearly $\ell(\mathsf{maxsize}(k) - 1) \le k < \ell(\mathsf{maxsize}(k))$ and so

$$\mathsf{maxsize}(k) = \min\{n : \ell(n) > k\}.$$

It is a simple matter to read off limit theorems for $h(A_n)$ and $\ell(n)$ from the corresponding results for their inverses, $\mathsf{minsize}$ and $\mathsf{maxsize}$, respectively. Let $\mu_0 = \mu_0(\alpha, d) = \mathsf{rate}^{-1}$.

Corollary 6.5. (a) $\lim_{n\to\infty} n^{-1}h(A_n) = \lim_{n\to\infty} n^{-1}\ell(n) = \mu_0$ a.s.

(b)(i) $\quad n^{-1/2}(h(A_n) - n\mu_0) \xrightarrow{w} N(0, \sigma^2_{\varphi_1}\mu_0^3)$ as $n \to \infty$.

(ii) $\quad n^{-1/2}(\ell(n) - n\mu_0) \xrightarrow{w} N(0, \sigma^2_{\varphi_1}\mu_0^3)$ as $n \to \infty$.

Proof. (a) is a trivial consequence of Theorem 6.4 (a) with $L(n) = \mathsf{minsize}(n)$ or $\mathsf{maxsize}(n)$.
(b) Let $H(n) = n^{-1/2}(h(A_n) - n\mu_0)$ and $\Sigma(k) = k^{-1/2}(\mathsf{minsize}(k) - k\mu_0^{-1})$. Fix $x \in \mathbb{R}$ and set $k(n) = [n\mu_0 + x\sqrt{n}] + 1$ where $[z]$ is the greatest integer not exceeding z. If $x_n = (n - k(n)\mu_0^{-1})k(n)^{-1/2}$, then for n large enough so that $k(n) \in \mathbb{N}$,

$$P(H_n \le x) = P(h(A_n) < k(n)) = P(\mathsf{minsize}(k(n)) > n) = P(\Sigma(k(n)) > x_n).$$

Since $\lim x_n = -\mu_0^{-3/2}x$, Theorem 6.4(b) with $L(k) = \sigma(k)$ implies that

$$\lim_{n\to\infty} P(H_n \le x) = P(Z > -\mu_0^{-3/2}x) = P(Z \le \mu_0^{-3/2}x),$$

where Z is a $N(0, \sigma^2_{\varphi_1})$ random variable. (i) follows and a similar argument proves (ii). $\square$

It is easy to translate these a.s. limit theorems as $n \to \infty$ into continuous time results as $t \uparrow T_\infty$. Note that

$$W_{k+1} \le (T_\infty - T^{(k)})\alpha^{-k-1} \le W(\mathsf{First}_k),$$

where First_k is the first vertex in $\mathbb{B}(k)$ which is added to the cluster. Note that $W(\mathsf{First}_k)$ is equal in law to W_0, and $W_{k+1} \succ \pi$ for all k by Theorem 5.1(b). Use the above with the estimates (5.1) and (5.5) (for the left-hand tail of π) and a Borel-Cantelli argument to see that

$$(k-1)^{-1}\alpha^{k-1} \le T_\infty - T^{(k)} \le (\log k)\alpha^k \text{ for large } k \text{ a.s.}$$

By considering $T^{(k)} \le t < T^{(k+1)}$ this gives

(6.13) $\quad h(C_t)^{-1}\alpha^{h(C_t)} < T_\infty - t < \log(h(C_t))\alpha^{h(C_t)}$ for $0 < T_\infty - t$ sufficiently small, a.s.

Theorem 6.6. (a) $\lim_{t\uparrow T_\infty}(h(C(t))(\log_\alpha(T_\infty - t))^{-1} = 1$ a.s.

(b) $\lim_{t\uparrow T_\infty}\#(C(t))(\log_\alpha(T_\infty - t))^{-1} = \mathsf{rate}$ a.s.

Proof. (a) is immediate from (6.13). Corollary 6.5 (a) and a trivial interpolation (recall $A_n = C(\mathsf{card}_n)$) shows that $\lim_{t\uparrow T_\infty} h(C(t))/\#C(t) = \mu_0$ a.s. (b) follows from this and (a). $\square$

Theorem 6.7. (a) $\lim_{n\to\infty} n^{-1}(\log_\alpha(T_\infty - \mathsf{card}_n)) = \mu_0$ a.s.

(b) $n^{-1/2}(\log_\alpha(T_\infty - \mathsf{card}_n) - n\mu_0) \xrightarrow{w} N(0, \sigma^2_{\varphi_1}\mu_0^3)$ as $n \to \infty$.

Proof. For (a) set $t = \mathsf{card}_n$ in Theorem 6.6(b). (b) follows from Corollary 6.5(b)(i) and

$$\lim_{n\to\infty} n^{-1/2}(h(C(\mathsf{card}_n)) - \log_\alpha(T_\infty - \mathsf{card}_n)) = 0 \text{ a.s.} \tag{6.14}$$

(6.14) is an easy consequence of (6.13) and the trivial bound $h(C(\mathsf{card}_n)) \le n$. $\square$

7. A Decomposition of the Infinite Cluster into i.i.d. Blocks.

The drawback of the decompositions of A_∞ into clusters off the backbone (Theorem 2.5) is that it is not a dynamical decomposition. This means that some additional work is needed before the results on the growth rate of $h(A_n)$ such as Corollary 6.5 can be derived from the limit theorems for the clusters $(\mathsf{Clust}_{n,i})$ given in Example 6.3. We now establish a dynamical decomposition of A_∞ into i.i.d. pieces, which will lead to a proof of Theorem 6.4, and will also be used in Section 8 to establish properties of the shape of the cluster "as viewed from the tip".

Let $\mathcal{C}_n = \sigma(Y_{k,i}, W_k : k < n, i < d) \vee \sigma(1(W_n \le \delta_0)) \vee \sigma(\mathsf{spine}_n)$. Here δ_0 is a sufficiently small positive number whose precise value will be prescribed below. Let $\bar{\mathcal{C}}_n = \mathcal{C}_n \vee \sigma(W_n)$. Clearly both $(\mathcal{C}_\cdot)$ and $(\bar{\mathcal{C}}_\cdot)$ are filtrations.

Let $M_n = \min\{a_{n,i}(x) : x \in \mathsf{Clust}_{n,i}, i < d\}$ and recall $C_n = \cup_{i<d}\mathsf{Clust}_{n,i}$. Inductively define regeneration times $\{R_j : j \in \mathbb{Z}_+\}$ by $R_0 = 0$ and

$$\begin{aligned} R_{j+1} = \min\{k > R_j :\ &\min_{R_j \le n < k} M_n\alpha^n > \delta_0\alpha^k, C_n \subset \mathbb{B}(0, k-n-2) \\ &\text{for all } n \in [R_j, k), W_{k-1} > \alpha\delta_0, W_k \le \delta_0\}. \end{aligned}$$

Here $\delta_0 > 0$ will be chosen below. Clearly R_j is a $(\mathcal{C}_n)$-stopping time for all $j \in \mathbb{Z}_+$. The first two and last conditions in the above inductive definition will imply that all the points in $A_\infty \cap \mathbb{B}(0, R_j)$ are added to the cluster before all the points in $A_\infty \cap \mathbb{B}(R_j, \infty)$ and that $\{\mathsf{spine}_{R_j}\}$ is the only point in A_∞ of generation R_j. The next to last condition will ensure the independence of the blocks $B(j)$, defined by

$$B(j) = A_\infty \cap \mathbb{B}(R_j, R_{j+1} - 1).$$

Our goal is to prove the following theorem (for a sufficiently small $\delta_0 > 0$).

Theorem 7.1. The sequence of $(\mathcal{C}_n)$-stopping times, $\{R_j : j \in \mathbb{Z}_+\}$, is a.s. finite and satisfies

(a) $A_\infty \cap \mathbb{B}(R_j) = \{\mathsf{spine}_{R_j}\}$ and $\mathsf{minsize}(R_j) = \mathsf{maxsize}(R_j)$ for all $j \in \mathbb{Z}_+$.

(b) $\mathbb{P}((\#B(j))^2) \le c_{7.1}(\alpha, d)$ for all $j \in \mathbb{Z}_+$.

(c) By (a) we may define $D_j \in S_0$ by

$$B(j) = \{(\mathsf{spine}_{R_j}) \oplus x : x \in D_j\},\ j \in \mathbb{Z}_+.$$

Also let $T(j;x) = T(\mathsf{spine}_{R_j}, (\mathsf{spine}_{R_j}) \oplus x)\alpha^{-R_j}$ for $x \in D_j$. Then $\{(D_j, T(j)) : j \in \mathbb{Z}_+\}$ are independent $\mathcal{S}_0$-valued random vectors and are identically distributed for $j \ge 1$. Moreover for $j \ge 1$,

$$\mathbb{P}((D_j, T(j)) \in \cdot\, |\mathcal{C}_{R_j}) = \int_0^{\delta_0} P((D_0, T^{(0)}_{\mathsf{not}}) \in \cdot\, | W_0 = w) d\Lambda(w)\Lambda([0,\delta_0])^{-1}.$$

Remark. We abuse notation slightly, and extend the definition of $T(j;x)$ to all $x \in \mathbb{B}$. Clearly $T(0;x) = T^{(0)}_{\mathsf{not}}(x)$ for all $x \in \mathbb{B}$.

Lemma 7.2. $\bar{\mathcal{C}}_n \subset \mathcal{W}_n$ for all $n \in \mathbb{Z}_+$.

Proof. By Theorem 5.1 W_n is $\mathcal{W}_n$-measurable, and spine_n is trivially $\mathcal{W}_n$-measurable. Fix $k < n$, $i < d$, and x, x' in $\mathbb{B}$. Then for a Borel set B

$$\{a_{k,i}(x') \in B, \mathsf{spine}_n = x\} = \{W(x|k) - T(x|k, e_i(x|k+1) \oplus x')\alpha^{-k-1} \in B,\ \mathsf{spine}_n = x\} \in \bar{\mathcal{F}}_x.$$

Therefore $Y_{k,i}$ is $\mathcal{W}_n$-measurable. □

Proof of Theorem 7.1(a). Assume $R_j < \infty$ for a fixed $j \in \mathbb{N}$. If $n < R_j, i < d$ and $x \in \mathsf{Clust}_{n,i}$, then

$$\begin{aligned} T(e_i^n \oplus x) = T(\infty) - \alpha^{n+1}a_{n,i}(x) &< T(\infty) - \alpha\delta_0\alpha^{R_j} \\ &\le T(\infty) - \alpha^{R_j+1}W_{R_j} = T(\mathsf{spine}_{R_j}). \end{aligned} \tag{7.1}$$

Note we have used the fact that if $R_j < \infty$ and $j \in \mathbb{N}$, then $M_n\alpha^n > \delta_0\alpha^{R_j}$ for all $n < R_j$ (and not just $R_{j-1} \le n < R_j$). Similarly we have $C_n \subset \mathbb{B}(0, R_j - n - 2)$ for all $n < R_j$ and this clearly implies that all the clusters which break off the backbone before R_j are in $\mathbb{B}(0, R_j - 1)$ and therefore

$$T(R_j) = T(\mathsf{spine}_{R_j}), \quad \mathsf{minsize}(R_j) = \#C(T(\mathsf{spine}_{R_j})). \tag{7.2}$$

By (7.1) no new points are added to the first R_j clusters $(C_0, \dots C_{R_j-1})$ after time $T(\mathsf{spine}_{R_j})$. Hence spine_{R_j} is the last node in $\mathbb{B}(0, R_j)$ added to A_∞. This means

$$\mathsf{maxsize}(R_j) = \#C(T(\mathsf{spine}_{R_j})). \tag{7.3}$$

Clearly $A_\infty \cap \mathbb{B}(R_j) = \{\mathsf{spine}_{R_j}\}$ since spine_{R_j} is the last point added to A_∞ in $\mathbb{B}(0, R_j)$ and the first point added to A_∞ in $\mathbb{B}(R_j)$. (7.2) and (7.3) also give $\mathsf{minsize}(R_j) = \mathsf{maxsize}(R_j)$ and hence (a) is proved once we show $R_j < \infty$ for all $j \in \mathbb{Z}_+$.

The proof of Theorem 7.1(b) requires several preparatory lemmas.

Lemma 7.3. (a) $\nu_t(\min_{x\in C} a(x) \le \delta) \le c_{4.3}(\log^+(1/\delta)+1)(\delta\wedge 1) \equiv g(\delta)$ for all t, $\delta \ge 0$.
(b) There exists $\beta \in (\alpha, 1)$ such that $\nu_t(C \not\subset \mathbb{B}(0, j-1)) \le \beta^{j+1}$ for all $j \in \mathbb{Z}_+$, $t \ge 0$.

Proof. (a) We have

$$
\begin{aligned}
\nu_t(\min_{x\in C} a(x) \le \delta) &\le \mathbb{P}(\max_{x\in C(t)} T(x) \ge t-\delta | T(\infty) > t) \\
&\le \mathbb{P}(\#(C(t) - C((t-\delta)-))|T(\infty) > t) \\
&\le g(\delta) \quad \text{(by Theorem 4.4(b))}.
\end{aligned}
$$

(b) Using Theorem 4.4(a) we have

$$
\begin{aligned}
\nu_t(C \not\subset \mathbb{B}(0, j-1)) &\le \mathbb{P}(\sum_{n\ge j} \#(C(t)\cap \mathbb{B}(n))|T(\infty) > t) \\
&\le \sum_{n\ge j} c_{4.1}^d c_{4.2}\alpha^{n+1} = c_{4.3}\alpha^{j+1},
\end{aligned}
$$

while

$$
\nu_t(C = \emptyset) = \mathbb{P}(U_0 > t|T(\infty) > t) = (e^t\mathbb{P}(T(\infty) > t))^{-1} \ge c_{4.1}^{-1} \text{ (by Lemma 4.1)}.
$$

The result follows trivially from the above two inequalities. $\square$

Lemma 7.4. If $0 < \delta_0 \le \delta_{5.1}(1/2, 1/2)$ ($\delta_{5.1}$ as in Lemma 5.5) there exist $c_{7.2}, c_{7.3} > 0$ such that $N_m = \sum_{k=1}^m 1(W_{k-1} > \alpha\delta_0, W_k \le \delta_0)$ satisfies

$$
\mathbb{P}(\exp(-N_m)|W_0) \le c_{7.2}(W_0^{-1/2}+1)e^{-c_{7.3}m} \qquad \text{for all } m \in \mathbb{N}.
$$

Proof. For δ_0 as above and a fixed $q > 0$ let $M_m = \sum_{k=1}^m 1(W_{k-1} > \alpha\delta_0)(q - 1(W_k \le \delta_0))$. Then

$$
\begin{aligned}
\mathbb{P}(e^{M_m}|\mathcal{W}_{m-1}) &= \exp(M_{m-1} + q1(W_{m-1} > \alpha\delta_0)) \\
&\qquad \times [\exp(-1(W_{m-1} > \alpha\delta_0))\Lambda([0, \delta_0 \wedge W_{m-1}/\alpha)])\Lambda([0, W_{m-1}/\alpha])^{-1} \\
&\qquad + \Lambda((\delta_0, W_{m-1}/\alpha])\Lambda([0, W_{m-1}/\alpha))^{-1}] \quad \text{(Theorem 5.1)} \\
&\le \exp(M_{m-1})[1(W_{m-1} \le \alpha\delta_0) \\
&\qquad + 1(W_{m-1} > \alpha\delta_0)e^q\big(e^{-1}\Lambda([0, \delta_0]) + \Lambda((\delta_0, \infty))\big)\Lambda([0, \infty))^{-1}].
\end{aligned}
$$

As $\Lambda((0, \delta_0]) > 0$ (see (5.9)), we may choose $q = q(\delta_0) > 0$ small enough such that e^{M_m} is a supermartingale. If $N'_m = \sum_{k=1}^m 1(W_k \le \alpha\delta_0)$, then

$$
e^{M_m} = \exp(qm - qN'_m - N_m) \ge 1(N'_m \le m/2)\exp(qm/2 - N_m)
$$

and therefore

$$
\begin{aligned}
\mathbb{P}(\exp(-N_m)|W_0) &\le \mathbb{P}(N'_m > m/2|W_0) + e^{-qm/2}\mathbb{P}(e^{M_m}|W_0) \\
&\le c_{5.3}(W_0^{-1/2}+1)e^{-\lambda_{5.1}m} + e^{-qm/2}.
\end{aligned}
$$

In the last line we used Lemma 5.5 and the supermartingale property of e^{M_m}. $\square$

Lemma 7.5. $\mathbb{P}(W_{R_j} \in B | \mathcal{C}_{R_j}) = \Lambda(B | [0, \delta_0])$ a.s. on $\{R_j < \infty\}$ for all $j \in \mathbb{N}$.

Proof. If $A \in \mathcal{C}_n$ we claim that

$$\mathbb{P}(A|\mathbf{W}) = \mathbb{P}(A|W_0, \ldots, W_{n-1}, 1(W_n \le \delta_0)). \tag{7.4}$$

To see this consider

$$A = \bigcap_{k=0}^{n-1} \bigcap_{i<d} \{Y_{k,i} \in A_{k,i}\} \bigcap \{\mathsf{spine}_n = x\} \cap \{(W_0, \ldots, W_{n-1}) \in B, W_n \in D\}$$

where $D = [0, \delta_0]$ or (δ_0, ∞). (7.4) is then an easy consequence of Theorem 2.5 and the independence of spine_n and $\mathbf{W}$ (which holds by symmetry). (7.4) follows for general $A \in \mathcal{C}_n$ by a monotone class argument.

Assume $A \in \mathcal{C}_n$ is a subset of $\{W_{n-1} > \alpha\delta_0, W_n \le \delta_0\}$. Then

$$\begin{aligned}\mathbb{P}(A, W_n \in B) &= \int \mathbb{P}(A|W_0, \ldots, W_{n-1}, 1(W_n \le \delta_0)) 1(W_n \in B) d\mathbb{P} \quad \text{(by (7.4))} \\ &= \int \mathbb{P}(A|W_{n-1}, 1(W_n \le \delta_0)) \mathbb{P}(W_n \in B|W_{n-1}, 1(W_n \le \delta_0)) d\mathbb{P},\end{aligned}$$

by the Markov property of $\mathbf{W}$. Use the form of the transition kernel found in Theorem 5.1 to see that

$$\begin{aligned}\mathbb{P}(W_n \in B|W_{n-1}, 1(W_n \le \delta_0)) = {} & \Lambda(B|[0, \min(\delta_0, W_{n-1}/\alpha)]) 1(W_n \le \delta_0) \\ & + \Lambda(B|(\delta_0, W_{n-1}/\alpha]) 1(W_n > \delta_0, W_{n-1}/\alpha > \delta_0).\end{aligned}$$

Therefore we have (by our assumption on A)

$$\begin{aligned}\mathbb{P}(A, W_n \in B) &= \int \mathbb{P}(A|W_{n-1}, 1(W_n \le \delta_0)) \Lambda(B|[0, \delta_0]) d\mathbb{P} \\ &= \mathbb{P}(A) \Lambda(B|[0, \delta_0]).\end{aligned} \tag{7.5}$$

Let $j \in \mathbb{N}$, $A \in \mathcal{C}_{R_j}$ and $A_n = A \cap \{R_j = n\} \in \mathcal{C}_n$. Clearly $A_n \subset \{W_{n-1} > \alpha\delta_0, W_n \le \delta_0\}$ and so by (7.5)

$$\begin{aligned}\mathbb{P}(A, R_j < \infty, W_{R_j} \in B) &= \sum_{n=1}^{\infty} \mathbb{P}(A_n, W_n \in B) \\ &= \mathbb{P}(A, R_j < \infty) \Lambda(B|[0, \delta_0]).\end{aligned}$$

□

The next result is an easy consequence of Theorem 2.5.

Lemma 7.6. Let K be a $(\mathcal{C}_n)$-stopping time (possibly infinite). Then conditional on $\mathcal{C}_K \vee \sigma(\mathbf{W})$ and on $\{K < \infty\}$, $\{Y_{K+n,i} : n \in \mathbb{Z}_+, i < d\}$ are independent $\mathcal{S}_0$-valued random vectors such that

$$\mathbb{P}(Y_{K+n,i} \in A | \mathcal{C}_K \vee \sigma(\mathbf{W})) = \nu_{W_{K+n}}(A) \ \text{ a.s. on } \{K < \infty\}.$$

Proof of Theorem 7.1(b). Choose $\theta < \delta_0 \le \delta_{5.1}(\frac{1}{2}, \frac{1}{2}) \wedge 1$ small enough so that

$$p_1(\delta_0) = 1 - \prod_{j=1}^{\infty}(1 - g(\delta_0\alpha^j))^{d-1} + 1 - \prod_{j=1}^{\infty}(1-\beta^j)^{d-1} < 1. \tag{7.6}$$

Here g and β are as in Lemma 7.3. Choose $n_0 \in \mathbb{N}$ sufficiently large so that if $h(\delta_0) = c_{4.3}\delta_0(1 + \log 1/\delta_0 + \log 1/\alpha)$, then (recall $\beta > \alpha$)

$$(d-1)h(\delta_0)j\alpha^j \le \beta^j \ \text{ for } j > n_0 \ \text{ and } \gamma = \beta^{n_0}(p_1(\delta_0) - \beta^{n_0})d(1-\beta)^{-1} + p_1(\delta_0) < 1. \tag{7.7}$$

Fix $j \in \mathbb{Z}_+$ and let $k(0) \equiv R_j$, and $k(0) < k(1) < \ldots < k(n) < \ldots$ denote the successive times for which $W_{k(n)} \le \delta_0$ and $W_{k(n)-1} > \alpha\delta_0$ ($n \in \mathbb{N}$). If $R_j = \infty$ set $k(n) = \infty$ for all n. Each $k(n)$ is a $(\mathcal{C}_n)$-stopping time. Lemma 7.4 and the strong Markov property of W show that each $k(n)$ is finite if R_j is. Let $K(i) = k(in_0)$ and define

$$\begin{aligned}
B_i &= \{\min\{M_n\alpha^n : R_j \le n < K(i)\} \le \delta_0\alpha^{K(i)}\},\\
D_i &= \{C_n \not\subset \mathbb{B}(0, K(i) - n - 2) \text{ for some } n \text{ in } [R_j, K(i))\},\\
A_N &= \cap_{i=1}^N (B_i \cup D_i).
\end{aligned}$$

Note that $R_{j+1} - R_j > m$ and $K(N) \le m + R_j$ together imply A_N (since each $K(i)$ for $i \le N$ must violate one of the defining conditions for R_{j+1}). If

$$N'_m = \sum_{k=1}^m 1(W_{R_j+k} \le \delta_0, W_{R_j+k-1} > \alpha\delta_0),$$

this easily gives for $\varepsilon > 0$

$$\{R_{j+1} - R_j > m\} \subset \{N'_m \le \varepsilon m\} \cup A_{[\varepsilon m/n_0]}. \tag{7.8}$$

If $\omega \in A_N$ and $R_j(\omega) < \infty$, then $\omega \in B_N \cup D_N$, and so for some n in $[R_j, K(N))$, either C_n is not contained in $\mathbb{B}(0, K(N) - n - 2)$ or $M_n\alpha^n \le \delta_0\alpha^{K(N)}$. Choose $i = i(\omega) \in \{1, \ldots, N\}$ such that $K(i-1) \le n < K(i)$ and note that $\omega \in A_{i-1}$ as well ($\omega \in A_j$ for all $j \le N$). This shows that on $\{R_j < \infty\}$,

$$\begin{aligned}
&\mathbb{P}(A_N | \mathcal{C}_{R_j} \vee \sigma(\mathbf{W}))\\
&\quad\le \sum_{i=1}^N \mathbb{P}(A_{i-1} \cap \{\min(M_n\alpha^n : K(i-1) \le n < K(i)) \le \delta_0\alpha^{K(N)}\} | \mathcal{C}_{R_j} \vee \sigma(\mathbf{W}))\\
&\qquad\quad + \mathbb{P}(A_{i-1} \cap \{C_n \not\subset \mathbb{B}(0, K(N) - n - 2) \text{ for some } n \in [K(i-1), K(i))\} | \mathcal{C}_{R_j} \vee \sigma(\mathbf{W}))\\
&\quad= \sum_{i=1}^N \mathbb{P}(1(A_{i-1})\{1 - \prod_{n=K(i-1)}^{K(i)-1} \nu_{W_n}(\min_{x\in C} a(x) > \delta_0\alpha^{K(N)-n})^{d-1} + 1\\
&\qquad\quad - \prod_{n=K(i-1)}^{K(i)-1} \nu_{W_n}(C \subset \mathbb{B}(0, K(N) - n - 2))^{d-1}\} | \mathcal{C}_{R_j} \vee \sigma(\mathbf{W})) \qquad \text{(by Lemma 7.6)}
\end{aligned}$$

$$
\begin{aligned}
&\le \sum_{i=1}^{N} \mathbb{P}(A_{i-1}|\mathcal{C}_{R_j} \vee \sigma(\mathbf{W}))\Big\{1 - \prod_{n=K(i-1)}^{K(i)-1} (1 - g(\delta_0 \alpha^{K(N)-n}))^{d-1} \\
&\qquad\qquad + 1 - \prod_{n=K(i-1)}^{K(i)-1} (1 - \beta^{K(N)-n})^{d-1}\Big\} \qquad \text{(Lemma 7.3)} \\
&\le \mathbb{P}(A_{N-1}|\mathcal{C}_{R_j} \vee \sigma(\mathbf{W}))p_1(\delta_0) \\
&\quad + \sum_{i=1}^{N-1} \mathbb{P}(A_{i-1}|\mathcal{C}_{R_j} \vee \sigma(\mathbf{W}))(d-1)\Big(\sum_{n=K(i-1)}^{K(i)-1} g(\delta_0\alpha^{K(N)-n}) + \beta^{K(N)-n}\Big) \\
&\le \mathbb{P}(A_{N-1}|\mathcal{C}_{R_j} \vee \sigma(\mathbf{W}))p_1(\delta_0) + \sum_{i=1}^{N-1} \mathbb{P}(A_{i-1}|\mathcal{C}_{R_j} \vee \sigma(\mathbf{W})) \sum_{m=n_0(N-i)+1}^{\infty} d\beta^m.
\end{aligned}
$$

In the last line we used the first part of (7.7) and the inequality $K(N) - K(i) \ge (N-i)n_0$. A simple induction argument using the above and the definition of γ in (7.7) gives

(7.9) $$\mathbb{P}(A_N|\mathcal{C}_{R_j} \vee \sigma(\mathbf{W})) \le \gamma^N \quad \text{for all } N \in \mathbb{N} \text{ a.s. on } \{R_j < \infty\}.$$

Use (7.8) with $\varepsilon = c_{7.3}/2$ and the strong Markov property of $\mathbf{W}$ with respect to $(\bar{\mathcal{C}}_n)$ (recall Lemma 7.2 and Theorem 5.1) to see that if $m \in \mathbb{N}$ (recall $\mathbb{P}_w$ is the law of $\mathbf{W}$ starting at w)

$$
\begin{aligned}
&\mathbb{P}(R_{j+1} - R_j > m|\bar{\mathcal{C}}_{R_j}) \\
&\qquad \le \mathbb{P}(N'_m \le \varepsilon m|\bar{\mathcal{C}}_{R_j}) + \mathbb{P}(A_{[\varepsilon m/n_0]}|\bar{\mathcal{C}}_{R_j}) \\
&\qquad \le \exp(\varepsilon m)\mathbb{P}_{W(R_j)}(e^{-N_m}) + \gamma^{-1}(\gamma^{\varepsilon/n_0})^m \qquad \text{(by (7.9))} \\
&\qquad \le c(W(R_j)^{-1/2} + 1)e^{-c'm} \qquad \text{(by Lemma 7.4 and the choice of } \varepsilon\text{)}.
\end{aligned}
$$

Now condition on $\mathcal{C}_{R_j}$, use $F(t) \le dt$ (see (5.9)) if $j = 0$, and this together with Lemma 7.5 if $j \ge 1$, to derive

(7.10) $$\mathbb{P}(R_{j+1} - R_j > m|\mathcal{C}_{R_j}) \le c_{7.4}e^{-c_{7.5}m} \quad \text{for all } j \in \mathbb{Z}_+, m \in \mathbb{N}.$$

Let

$$M'_n = \sum_{i=0}^{n-1} 1(R_j \le i < R_{j+1})(\#C_i - (d-1)\nu_{W_i}(\#C)),\ n \in \mathbb{Z}_+.$$

Lemma 7.6 shows $(M'_n, \mathcal{C}_n \vee \sigma(\mathbf{W}))$ is a martingale and Theorems 2.5 and 4.6, and (7.10) readily show it is L^2-bounded. By (a) each point in $B(j)$ either belongs to a cluster which branched off the backbone at generation $i \in [R_j, R_{j+1})$ or to the backbone itself. Therefore

$$
\begin{aligned}
\mathbb{P}(\#(B(j))^2) &= \mathbb{P}\Big(\big(\sum_i 1(R_j \le i < R_{j+1})(\#(C_i) + 1)\big)^2\Big) \\
&\le 2(\mathbb{P}(M'_\infty)^2) + \mathbb{P}\Big(\big(\sum_i 1(R_j \le i < R_{j+1})(\nu_{W_i}(\#(C))(d-1)+1)\big)^2\Big) \\
&\le 2\mathbb{P}((M'_\infty)^2) + c\mathbb{P}((R_{j+1} - R_j)^2) \qquad \text{(Theorem 4.4)} \\
&\le c_{7.6}
\end{aligned}
$$

by the above and (7.10). This proves (b), and also shows that $R_j < \infty$ for all j a.s.

For (c), one more lemma is required.

Lemma 7.7. For all $j \in \mathbb{N}$ and measurable $A \in [0,\infty)^{\mathbb{B}}$ we have

$$\mathbb{P}((T(j;x), x \in \mathbb{B}) \in A | \mathcal{C}_{R_j}) = \int_0^{\delta_0} \mathbb{P}((T_0(x), x \in \mathbb{B}) \in A | W_0 = w) d\Lambda(w) \Lambda([0,\delta_0])^{-1}.$$

Proof. Let $p((U_x, x \in \mathbb{B} - \{0\}) \in \cdot | w)$ be a regular conditional probability for $(U_x, x \in \mathbb{B} - \{0\})$ given $W_0 = w$. Let $U(j;x) = U((\mathsf{spine}_{R_j}) \oplus x)$ and choose $F \subset \mathbb{B} - \{0\}$ finite and B a measurable subset of $[0,\infty)^F$. We will prove

$$\mathbb{P}((U(j;x), x \in F) \in B | \mathcal{W}_{R_j})(\omega) = p((U_x, x \in F) \in B | W_{R_j}(\omega)) \text{ a.s.} \tag{7.11}$$

Since $T(j;x) = \sum_{0 \neq y \leq x} U(j;y) \alpha^{|y|}$, we can condition on $\mathcal{C}_{R_j} \subset \mathcal{W}_{R_j}$ and use Lemma 7.5 to obtain the desired result. Turning to (7.11), note that for $A \in \mathcal{W}_{R_j}$ and $x_0 \in \mathbb{B}$,

$$\mathbb{P}((U(j;x), x \in F) \in B, A, \mathsf{spine}_{R_j} = x_0) = \mathbb{P}((U_{x_0 \oplus x}, x \in F) \in B, A, \mathsf{spine}_{R_j} = x_0), \tag{7.12}$$

that $A \cap \{\mathsf{spine}_{R_j} = x_0\}$ is in $\bar{\mathcal{F}}_{x_0} = \mathcal{E}_{x_0} \vee \sigma(W(x_0))$, and that $\{(U_{x_0 \oplus x}, x \in F) \in B\} \in \mathcal{F}_{(x_0,\infty)}$. The independence of $\mathcal{E}_{x_0}$ and $\mathcal{F}_{(x_0,\infty)}$ and the inclusion $\sigma(W(x_0)) \subset \mathcal{F}_{(x_0,\infty)}$ therefore shows that (7.12) equals

$$\begin{aligned} &\mathbb{P}(\mathbb{P}((U_{x_0 \oplus x}, x \in F) \in B | W(x_0))\ 1(A, \mathsf{spine}_{R_j} = x_0)) \\ &\qquad = \mathbb{P}(p((U_x, x \in F) \in B | W_{R_j})\ 1(A, \mathsf{spine}_{R_j} = x_0)). \end{aligned}$$

Sum over x_0 to obtain (7.11). □

Since $\mathbf{W}$ and $(Y_{n,i})$ are measurable functions of $T^{(0)}_{\mathsf{not}}(\cdot)$, there is a measurable map $r : [0,\infty)^{\mathbb{B}} \to \mathbb{N}$ such that $R_1 = r(T^{(0)}_{\mathsf{not}})$. It is straightforward to check that $R_{j+1} - R_j = r(T(j;\cdot))$ for all $j \in \mathbb{Z}_+$ (use (a)). Define $\Phi : [0,\infty)^{\mathbb{B}} \to \mathcal{S}_0$ by

$$\begin{aligned} &\Phi(T^{(0)}_{\mathsf{not}}(\cdot)) = (\overline{D_0}(T^{(0)}_{\mathsf{not}}(\cdot)), T^{(0)}_{\mathsf{not}}(\cdot)|_{\overline{D_0}}), \quad \text{where} \\ &\overline{D_0}(T^{(0)}_{\mathsf{not}}(\cdot)) = \{x \in \mathbb{B} : T^{(0)}_{\mathsf{not}}(x) < \lim_{n\to\infty} \inf\{T^{(0)}_{\mathsf{not}}(y) : |y| = n\}, |x| < r(T^{(0)}_{\mathsf{not}})\}. \end{aligned}$$

Then $(D_j, T(j;\cdot)) = \Phi(T(j;\cdot))$ for all $j \in \mathbb{Z}_+$ and so Lemma 7.7 shows that if $A \subset \mathcal{S}_0$ is measurable, then

$$\begin{aligned} \mathbb{P}((D_j, T(j;\cdot)) \in A | \mathcal{C}_{R_j}) &= \int_0^{\delta_0} \mathbb{P}(\Phi(T^{(0)}_{\mathsf{not}}) \in A | W_0 = w) d\Lambda(w) \Lambda([0,\delta_0])^{-1} \\ &= \int_0^{\delta_0} \mathbb{P}((D_0, T^{(0)}_{\mathsf{not}}) \in A | W_0 = w) d\Lambda(w) \Lambda([0,\delta_0])^{-1}. \end{aligned}$$

To complete the proof of Theorem 7.1(c) it suffices to show that $(D_j, T(j;\cdot))$ is $\mathcal{C}_{R_{j+1}}$-measurable for all $j \in \mathbb{Z}_+$. This reduces to showing that for a fixed $x \in \mathbb{B}$,

$$\begin{aligned}&T^{(0)}_{\mathsf{not}}((\mathsf{spine}_{R_j}) \oplus x) 1(T^{(0)}_{\mathsf{not}}((\mathsf{spine}_{R_j}) \oplus x) < \alpha W_0, |x| < R_{j+1} - R_j) \\ &\qquad\qquad\qquad\qquad \text{is } \mathcal{C}_{R_{j+1}}\text{-measurable.}\end{aligned} \tag{7.13}$$

If $n \in \mathbb{Z}_+$ and $x_0 \in \mathbb{B}$ satisfy $n + |x| < |x_0|$ then on $\{\mathsf{spine}_{R_{j+1}} = x_0, R_j = n\}$ we consider the following two cases:

Case 1. $(x_0|n) \oplus x \le x_0$.

Then $(\mathsf{spine}_{R_j}) \oplus x = x_0|(n + |x|)$, $T^{(0)}_{\mathsf{not}}(\mathsf{spine}_{R_j} \oplus x) = \alpha W_0 - \alpha^{n+|x|} W_{n+|x|}$, and so

$$\begin{aligned}\{T^{(0)}_{\mathsf{not}}&((\mathsf{spine}_{R_j}) \oplus x) \in B, T^{(0)}_{\mathsf{not}}((\mathsf{spine}_{R_j}) \oplus x) < \alpha W_0, \mathsf{spine}_{R_{j+1}} = x_0, R_j = n\} \\ &= \{\alpha W_0 - \alpha^{n+|x|} W_{n+|x|} \in B, R_{j+1} = |x_0|, \mathsf{spine}_{|x_0|} = x_0, R_j = n\} \\ &\in \mathcal{C}_{|x_0|} \quad (\text{because } n + |x| < |x_0|).\end{aligned}$$

Case 2. Case 1 fails.

Then there exists $m \in [n, |x_0| - 1) \cap \mathbb{Z}_+$, $i < d$ and x' in $\mathbb{B}$ (depending on (x, x_0)) such that $(\mathsf{spine}_{R_j}) \oplus x = e_i(x_0|m+1) \oplus x'$. Now $T^{(0)}_{\mathsf{not}}(\mathsf{spine}_{R_j}) \oplus x) < \alpha W_0$ if and only if $x' \in C_{m,i}$ in which case $T^{(0)}_{\mathsf{not}}((\mathsf{spine}_{R_j}) \oplus x) = \alpha W_0 - \alpha^{m+1} a_{m,i}(x')$. Therefore

$$\begin{aligned}\{T^{(0)}_{\mathsf{not}}&((\mathsf{spine}_{R_j} \oplus x) \in B, T^{(0)}_{\mathsf{not}}((\mathsf{spine}_{R_j} \oplus x) < \alpha W_0, \mathsf{spine}_{R_{j+1}} = x_0, R_j = n\} \\ &= \{\alpha W_0 - \alpha^{m+1} a_{m,i}(x') \in B, x' \in C_{m,i}, \mathsf{spine}_{|x_0|} = x_0, R_{j+1} = |x_0|, R_j = n\} \\ &\in \mathcal{C}_{|x_0|} \quad (\text{ because } n \le m < |x_0|).\end{aligned}$$

Taking the union over $n < k - |x|$ and $|x_0| = k$ in the above cases we have

$$\begin{aligned}\{T^{(0)}_{\mathsf{not}}((\mathsf{spine}_{R_j}) \oplus x) \in B, T^{(0)}_{\mathsf{not}}&((\mathsf{spine}_{R_j} \oplus x) < \alpha W_0, \\ &\text{for } |x| < R_{j+1} - R_j, R_{j+1} = k\} \in \mathcal{C}_k\end{aligned}$$

and so (7.13) follows. □

We will now use Theorem 7.1 to complete the proof of the main limit theorem, Theorem 6.4. Let $I(n) = [R_{k-1}, R_k) \cap \mathbb{Z}_+$ iff $R_{k-1} \le n < R_k$ and let

$$Z_n = \sum_{j=0}^{\infty} \#C_j 1\big(j \in I(n)\big), n \in \mathbb{Z}_+,$$

be the size of the "regeneration block" spanning generation n. Theorem 7.1(a) implies that

$$\mathsf{maxsize}(n) \le \mathsf{minsize}(n) + Z_n \tag{7.14}$$

and (recall that $N(n)=\sum_{k=0}^{n-1}\#(C_k)$)

$$N(n)+n+1\le \mathsf{minsize}(n)+Z_n. \tag{7.15}$$

Lemma 7.8. The sequence $\{Z_n\}$ is bounded in L^1.

Proof. It follows easily from the exponential estimate (7.10) on the tail of R_k-R_{k-1} and the Renewal Theorem (see (4.16) in Ch. XI of Feller (1971)) that

$$\sup_n \mathbb{P}\left((\#I(n))^q\right)<\infty \text{ for all } q>0. \tag{7.16}$$

Note that Theorems 2.5 and 4.4 together with the definition of $\overline{\mathcal{C}}_j$ imply

$$\mathbb{P}(\#C_j|\overline{\mathcal{C}}_j)\le (d-1)c_{4.3}, \tag{7.17}$$

and that

$$Z_n=\sum_{j=0}^{\infty}\sum_{k=1}^{\infty}\#C_j 1\big(j,n\in[R_{k-1},R_k)\big).$$

Take means in the above to conclude that if $p>1$,

$$\begin{aligned}
\mathbb{P}(Z_n)&\le \sum_{j=0}^{n-1}\sum_{k=1}^{\infty}\mathbb{P}\Big(1\big(j\in[R_{k-1},R_k)\big)\#C_j(R_k-j)^p(n-j)^{-p}\Big)\\
&\quad+\sum_{j=n}^{\infty}\sum_{k=1}^{\infty}\mathbb{P}\Big(1\big(R_{k-1}\le n\le j<R_k\big)\mathbb{P}(\#C_j|\overline{\mathcal{C}}_j)\Big)\\
&\le \sum_{j=0}^{n-1}\mathbb{P}\Big((\#C_j)\big(\#I(j)\big)^p\Big)(n-j)^{-p}+(d-1)c_{4.3}\mathbb{P}\Big(\sum_{j=n}^{\infty}1\big(j\in I(n)\big)\Big)\\
&\le \sum_{j=0}^{n-1}\mathbb{P}\big((\#C_j)^2\big)^{1/2}\mathbb{P}\big(\#I(j)^{2p}\big)^{1/2}(n-j)^{-p}+(d-1)c_{4.3}\mathbb{P}\big(\#I(n)\big).
\end{aligned}$$

Use (7.16) and Theorems 2.5 and 4.6 to see that the final expression above is uniformly bounded in n. □

Proof of Theorem 6.4. For each choice of $L(n)$, writing

$$L(n)=N(n)+n+1+\Delta_n,$$

the inequalities (6.10), (6.11), (7.14) and (7.15) imply that $|\Delta_n|\le Z_n$ and so is bounded in L^1 by the previous Lemma. The Borel-Cantelli Lemma implies $n^{-2}\Delta_{n^2}\to 0$ a.s. and so Example 6.3 gives a.s. convergence along the subsequence $\{n^2\}$. A standard interpolation argument completes the proof of (a). Since $n^{-1/2}\Delta_n\to 0$ in probability, (b) is now immediate from the Central Limit Theorem for $N(n)$ (Example 6.3). □

8. Some Concluding Remarks

While equations (4.1) and (4.2) specify the law of $T(\infty)$ (see the Remark following Lemma 4.1), precise estimates on its distribution seem quite difficult to obtain. It is however possible to derive some asymptotic results as $\alpha \uparrow 1$. Our starting point is the following

Proposition 8.1. (Kingman (1975)). If $\alpha = 1$ then $n^{-1}T(n) \to c_{8.1}(d)$ a.s. as $n \to \infty$, where $c = c_{8.1}(d)$ is the unique root in $(0,1)$ of $dce^{1-c} - 1 = 0$.

As $\alpha \in (0,1)$ will vary in the following, we will use notation such as $T_\alpha(x)$ or $T_\alpha(\infty)$ to denote dependence on α. Bear in mind that the times $T_\alpha(x)$ are all defined on a common probability space as sums of the same variables U_x with different weights.

Theorem 8.2. $(1-\alpha)T_\alpha(\infty) \to c_{8.1}(d)$ as $\alpha \uparrow 1$ a.s. and in L^1.

Proof. Fix $\varepsilon > 0$, and use Proposition 8.1 to choose $K(\varepsilon, \omega)$ such that $n^{-1}T_1(n) > c_{8.1} - \varepsilon$ for $n \geq K(\varepsilon, \omega)$. If $x \in \mathbb{B}$, then

$$\begin{aligned} T_\alpha(0,x) &= \sum_{i=1}^{|x|} \alpha^i (T_1(0, x|i) - T_1(0, x|i-1)) \\ &\geq \sum_{i=K(\varepsilon)}^{|x|} (\alpha^i - \alpha^{i+1}) T_1(0, x|i) + \alpha^{|x|+1} T_1(0,x) \\ &\geq (c_{8.1} - \varepsilon) \sum_{i=K(\varepsilon)}^{|x|} (1-\alpha) i \alpha^i \\ &\geq (c_{8.1} - \varepsilon)(1-\alpha)\alpha^{K(\varepsilon)} \sum_{i=0}^{|x|-K(\varepsilon)} i\alpha^i. \end{aligned}$$

Take the minimum over $|x| = M$ and let $M \to \infty$ to see that

$$\liminf_{\alpha \uparrow 1} (1-\alpha) T_\alpha(\infty) \geq c_{8.1}(d) \quad \text{a.s.} \tag{8.1}$$

For the other direction, note that $T_1(n)/n$ is uniformly L^2-bounded (being stochastically smaller than the average of n i.i.d. exponentials) and so the convergence in Proposition 8.1 holds in L^1 as well. Choose N large enough so that $\mathbb{P}(T_1(N)) \leq (c_{8.1} + \varepsilon)N$. Now choose random vertices x_k inductively N generations apart so that $x_0 = 0$ and x_{k+1} minimizes $T_1(x_k, x_{k+1})$ among all descendants of x_k. The times $T_1(x_k, x_{k+1}) = T_1(x_{k+1}) - T_1(x_k)$ will be i.i.d. Use the crude bound

$$T_\alpha(\infty) \leq \sum_{k=0}^{\infty} \alpha^{kN} T_1(x_k, x_{k+1})$$

together with summation by parts and the Strong Law (as above) to see that $\limsup_{\alpha \uparrow 1} (1-\alpha)T_\alpha(\infty) \leq c_{8.1} + \varepsilon$ almost surely, and that

$$\mathbb{P}[(1-\alpha)T_\alpha(\infty)] \leq \sum_{k=0}^{\infty} (1-\alpha)\alpha^{kN}(c_{8.1} + \varepsilon)N \to c_{8.1} + \varepsilon \tag{8.2}$$

as $\alpha \uparrow 1$ for fixed N. The first conclusion and (8.1) give the required almost sure convergence. Fatou's Lemma and (8.2) then show that the mean value of $(1-\alpha)T_\alpha(\infty)$ approaches $c_{8.1}$ as $\alpha \uparrow 1$. Convergence in L^1 now follows. $\square$

Remark 8.3. Recall from Lemma 4.1 that $e^t\mathbb{P}(T_\alpha(\infty) > t)$ increases to a finite limit $c_{4.1}(\alpha, d)$ as $t \to \infty$. It is possible to show that

$$\lim_{\alpha\uparrow 1}(1-\alpha)\log c_{4.1}(\alpha, d) = c_{8.1}(d). \tag{8.3}$$

The lower bound is an easy consequence of (8.1) and the monotonicity of $e^t\mathbb{P}(T_\alpha(\infty) > t)$ in t. The upper bound is more involved and we will not give a proof. It uses the anticipating equation (4.2). Equation (8.3) shows that the bound on $c_{4.1}$ in Lemma 4.1 is far from optimal. For example if $d = 2$, then $c_{8.1}(d) \approx .23$ and (8.3) implies $c_{4.1}(\alpha, 2) \le \exp((.23+\varepsilon)/(1-\alpha))$ for α close to 1; Lemma 4.1 gives the same kind of bound but with $.23+\varepsilon$ replaced by $\log 2 \approx .69$.

We conclude this paper by mentioning an associated particle system. Set

$$u_t(n) = \#(\partial C_t \cap \mathbb{B}(n)), \quad t \ge 0,\, n \in \mathbb{Z}_+.$$

The "particles" (i.e. sites in $\mathbb{B}$ on the boundary of the cluster C_t) evolve independently: each u-particle at a site $n \in \mathbb{Z}_+$ dies at rate α^n, and is replaced by d particles at $n+1$. The process $U = (u_t(.) : t \ge 0)$ captures the essential features of the DLA processes C_t and A_n: only the labels of the branches are lost. For various limit theorems on the process U in the case $\alpha > 1$ see Aldous and Shields (1988).

To study the evolving cluster for $\alpha < 1$ it is more helpful to consider the following modification of U. Define a random time change σ_t so that C_{σ_t} always adds neighbours to its deepest vertices at a constant rate:

$$\begin{aligned} M_t &= \max\{n : u_t(n) > 0\}, \\ L_t &= \int_0^t \alpha^{-M_s}\,ds, \quad 0 < t < T(\infty), \\ \sigma_t &= \inf\{s : L_s > t\}, \\ V_t(n) &= u_{\sigma_t}(M_{\sigma_t} - n), \quad t \ge 0,\, n \in \mathbb{Z}_+, \end{aligned}$$

where $u_t(n)$ is taken to be zero for $n < 0$. At each time t, V_t is a function on $\mathbb{Z}_+$ counting how many vertices are in ∂C_{σ_t} at each level below the highest one. Note that $L_t < \infty$ if $t < T(\infty)$, and that $\lim_{t\uparrow T(\infty)} L_t = \infty$; thus $\sigma_t < T(\infty)$ for all $t \ge 0$. Straightforward calculations show that the process $\mathbf{V} = (V_t(.) : t \ge 0)$ evolves as follows:

(i) Particles at site n, $n \ge 1$, die at rate α^n and are replaced by d particles at site $n-1$.

(ii) Particles at 0 die at rate 1, and are replaced by d particles at site -1. The whole configuration is then immediately shifted to the right by 1 step.

We call $\mathbf{V}$ the "tip process": it describes the form of the cluster when viewed backwards from the tip. The size of the process $\mathbf{V}$ near 0 arises from the interaction of two effects: first the strongly supercritical branching (at an accelerating rate as particles approach 0), and secondly the right shifts, which move particles away from 0, and so slow down their branching. Since the process $\mathbf{V}$ is a functional of C, it should be possible to deduce many properties of $\mathbf{V}$ from our results on C. However, just as some work was needed to obtain results such as Corollary 6.5 (giving the growth rate of A_n) from the cluster decomposition, so also passing from C to $\mathbf{V}$ is not completely straightforward. A further study of $\mathbf{V}$ may be the subject of a future paper: here we will just give a sketch proof that $\mathbf{V}$ is (in a certain sense) recurrent.

For $f : \mathbb{Z}_+ \to \mathbb{R}$ and $\lambda \in (0,1)$, set $||f||_\lambda = \sum_{n=0}^\infty \lambda^n |f(n)|$.

Theorem 8.4. Let $\lambda \in (0,1)$. Then the process $||V_t||_\lambda$ is recurrent in the sense that there exists $c_{8.2} < \infty$ such that

$$\{t : ||V_t||_\lambda < c_{8.2}\} \quad \text{is unbounded.}$$

Proof. Recall from Section 7 the definition of the regeneration times R_j, and set

$$\xi_j = \#(B(j-1)),$$
$$S_j = \inf\{t \geq 0 : M_{\sigma_t} = R_j\}.$$

It is clear that $S_j < \infty$ for all j and that $\lim_{j\to\infty} S_j = \infty$. Then

$$\begin{aligned} ||V_{S_j}||_\lambda &= \sum_{n=0}^{R_j} \lambda^{R_j - n} u_{\sigma_{S_j}}(n) \\ &= \sum_{i=1}^{j} \sum_{n} 1(R_{i-1} \leq n < R_i) \lambda^{R_j - n} u_{\sigma_{S_j}}(n) + u_{\sigma_{S_j}}(R_j) \leq \sum_{i=1}^{j} \lambda^{j-i} \xi_i + 1. \end{aligned}$$

Here we have used the facts that $R_j \geq R_{j-1} + 1$, and that the cluster $A_n \cap \mathbb{B}(0, R_i)$ is frozen for $n \geq \mathsf{maxsize}(R_i)$. By Theorem 7.1 $\xi_j, j \geq 1$ are i.i.d. with $P(\xi_j^2) < \infty$. Thus (b) follows by comparison with the interval recurrent Markov chain $Z_n = \sum_{i=1}^n \lambda^{n-i} \xi_i$. □

References

Aldous, D.J. (1991). The continuum random tree, I. *Ann. Probab.* **19**, 1 - 28.

Aldous, D.J. and Shields, P. (1980). A diffusion limit for a class of randomly-growing binary trees. *Prob. Th. Rel. Fields* **79**, 509-542.

Athreya, K.B. (1985) Discounted branching random walks. *Adv. Appl. Prob.* **17**, 53-66.

Barlow, M.T. (1993) Fractals and diffusion limited-aggregation. *Bull. Sci. Math.* **117** , 161-169.

Bramson, M. (1978). Minimal displacement of branching random walk. *Zeit. Wahr.* **45**, 89 - 108.

Breiman, Leo (1968). *Probability.* Addison-Wesley, Reading, Mass.

Brennan, M.D. and Durrett, R. (1986). Splitting intervals. *Ann. Probab.* **14**, 1023-1036.

Devroye, L. (1986). A note on the height of binary search trees. *Journal ACM* **33**, 489-498.

Durrett, R. (1991). Probability: theory and examples. Wadsworth, Belmont, CA.

Evans, W., Kenyon, C., Peres, Y. and Schulman, L. (1995). A critical phenomenon in a broadcast process. *Preprint.*

Feller, W. (1971). *An Introduction to Probability Theory and its Applications Volume II, 2nd edition.* John Wiley and Sons, New York.

Kamae, T., Krengel, U. and O'Brien, G.L. (1977). Stochastic inequalities on partially ordered spaces. *Ann. Probab.* **5**, 899-912.

Kesten, H. (1987). How long are the arms in DLA? *J. Phys. A* **20**, L29-L33.

Kesten, H. (1990). Upper bounds for the growth rate of DLA. *Physica A* **168**, 529 - 535.

Kingman, J.F.C. (1975). The first birth problem for an age-dependent branching process. *Ann. Probab.* **3**, 790-801.

Lawler, G.F. (1991). *Intersections of Random Walks.* Birkhäuser, Boston.

Liggett, T.M. (1985). *Interacting Particle Systems.* Springer- Verlag, New York.

Lorden, G. (1970). On excess over the boundary. *Ann. Math. Stat.* **41**, 521-527.

Lyons, R. (1990). Random walks and percolation on a tree. *Ann. Probab.* **18**, 931 - 958.

Mahmoud, H. (1992). *Evolution of Random Search Trees.* John Wiley and Sons, New York.

Newman, C.M. (1980). Normal fluctuations and the FKG inequalities. *Commun. Math. Phys.* **74**, 119-128.

Pemantle, R. and Peres, Y. (1994). Domination between trees and application to an explosion problem. *Ann. Probab.* **22**, 180-194.

Pittel, B. (1984). On growing random binary trees. *J. Math. Anal. Appl.* **103**, 461-480.

Pittel, B. (1985). Asymptotical growth of a class of random trees, *Ann. Probab.* **13**, 414-427.

Rio, E. (1995). The functional law of the iterated logarithm for stationary strongly mixing sequences. *Ann. Probab.* **23**, 1188-1203.

Vannimenus, J., Nickel, B. and Hakim, V. (1984). Models of cluster growth on the Cayley tree. *Phys. Rev. B.* **30**, 391-399.

Vicsek, T. (1989). *Fractal Growth Phenomena.* World Scientific, Singapore.

Witten, T.A. and Sander, L.M. (1981). Diffusion limited aggregation, a kinetic critical approach. *Phys. Rev. Lett.* **47**, 1400-1403.

Ziv, J. (1978). Coding theorems for individual sequences. *IEEE Trans. Inf. Theory* **24**, 405-412.

Address for M. T. Barlow and E. A. Perkins:
Department of Mathematics
University of British Columbia
Vancouver, B.C. V6T 1Z2
Canada

Address for R. Pemantle:
Department of Mathematics
University of Wisconsin
480 Lincoln Drive
Madison, WI 53706 U.S.A.