J. Appl. Prob. **53,** 846–856 (2016)
*doi:*10.1017/jpr.2016.44

LONGEST PATHS IN RANDOM APOLLONIAN NETWORKS AND LARGEST *r*-ARY SUBTREES OF RANDOM *d*-ARY RECURSIVE TREES

ANDREA COLLEVECCHIO,* ** *Monash University*
ABBAS MEHRABIAN,*** *University of Waterloo*
NICK WORMALD,* **** *Monash University*

Abstract

Let r and d be positive integers with $r < d$. Consider a random d-ary tree constructed as follows. Start with a single vertex, and in each time-step choose a uniformly random leaf and give it d newly created offspring. Let $\mathcal{T}_{d,t}$ be the tree produced after t steps. We show that there exists a fixed $\delta < 1$ depending on d and r such that almost surely for all large t, every r-ary subtree of $\mathcal{T}_{d,t}$ has less than t^{δ} vertices. The proof involves analysis that also yields a related result. Consider the following iterative construction of a random planar triangulation. Start with a triangle embedded in the plane. In each step, choose a bounded face uniformly at random, add a vertex inside that face and join it to the vertices of the face. In this way, one face is destroyed and three new faces are created. After t steps, we obtain a random triangulated plane graph with $t + 3$ vertices, which is called a random Apollonian network. We prove that there exists a fixed $\delta < 1$, such that eventually every path in this graph has length less than t^{δ}, which verifies a conjecture of Cooper and Frieze (2015).

Keywords: Random Apollonian network; random d-ary recursive tree; longest path; Eggenberger–Pólya urn

2010 Mathematics Subject Classification: Primary 05C80
Secondary 60C05; 05C05

1. Introduction

In this paper we study two important random graph models. The first one is a so-called *random d-ary recursive tree*, defined as follows. Let $d > 1$ be a positive integer. Consider a random d-ary tree evolving as follows. At time 0 it consists of exactly one vertex, ϱ. In the first step ϱ gives birth to d offspring. In each subsequent step we pick, uniformly at random, a vertex with no offspring and connect it with exactly d offspring. At time t this random tree is denoted by $\mathcal{T}_{d,t}$. See [5] for more on random d-ary recursive trees. Let r be a fixed positive integer smaller than d and let S_t denote the size of the largest (possibly nonunique) r-ary subtree of $\mathcal{T}_{d,t}$.

We say that a sequence of events $\{A_k,\ k \in \mathbb{N}\}$ occurs *eventually* (for large k) if there exists an almost surely finite random variable N such that A_k occurs for all $k \geq N$. In this paper all logarithms are natural.

Received 10 February 2015; revision received 9 October 2015.
* Postal address: School of Mathematical Sciences, 9 Rainforest Walk, Clayton, VIC 3800, Australia.
** Email address: andrea.collevecchio@monash.edu
*** Postal address: Department of Combinatorics and Optimization, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1, Canada. Email address: amehrabi@uwaterloo.ca
**** Email address: nick.wormald@monash.edu

Theorem 1. *There exists a fixed $\delta < 1$ such that $S_t < t^\delta$ eventually.*

It turns out that we can take, e.g.

$$\delta = 1 - \frac{d - r}{\mathrm{e} d^{2d} \log(11 d \log d)}$$

in this theorem; see [2, Section 5] for details.

We note at this point that it is possible to analyse quite accurately the size of the subtree obtained by deleting all but the r largest branches at each vertex. However, this does not lead to a proof of the above theorem, since in general this will not be the largest r-ary subtree (see Figure 1 for an example).

The second object we study is a popular random graph model for generating planar graphs with power-law properties, which is defined as follows. Start with a triangle embedded in the plane. At each step, choose a bounded face uniformly at random, add a vertex inside that face and join it to the vertices on the face. In this way, one face is destroyed and three new faces are created. We call this operation *subdividing* the face. After t steps, we have a (random) triangulated plane graph RAN_t with $t + 3$ vertices, $3t + 3$ edges, $2t + 1$ bounded faces, and 1 unbounded face. The random graph RAN_t is called a *random Apollonian network*.

Random Apollonian networks were defined by Zhou *et al.* [13] (see [12] for a generalization to higher dimensions), where it was proved that the diameter of a RAN_t is probabilistically bounded above by a constant times $\log t$. It was shown in [10] and [13] that the RAN_t exhibits a power-law degree distribution for large t. The average distance between two vertices in a typical RAN_t was shown to be $\Theta(\log t)$ by Albenque and Marckert [1], and a central limit theorem was proved by Kolossváry *et al.* [8]. The degree distribution, k largest degrees, k largest eigenvalues (for fixed k), and diameter were studied by Frieze and Tsourakakis [7]. The asymptotic value of the diameter of a typical RAN_t was determined in [6] (and was generalized to higher dimensions in [4] and [8]). We continue this line of research by studying the asymptotic properties of the longest (simple) paths in a RAN_t.

Let $\mathcal{L}_t$ be a random variable denoting the number of vertices in a longest path in a RAN_t. All the limits in this paragraph are as $t \to \infty$. Frieze and Tsourakakis [7] conjectured that there exists a fixed $\delta > 0$ such that $\mathbb{P}(\delta t \le \mathcal{L}_t < t) \to 1$. Ebrahimzadeh *et al.* [6] refuted this conjecture and showed that there exists a fixed $\delta > 0$ such that $\mathbb{P}(\mathcal{L}_t < t/(\log t)^\delta) \to 1$. Cooper and Frieze [3] improved this result by showing that, for every constant $c < \frac{2}{3}$, we have $\mathbb{P}(\mathcal{L}_t \le t \exp(-\log^c t)) \to 1$, and conjectured that there exists a fixed $\delta < 1$ such that $\mathbb{P}(\mathcal{L}_t \le t^\delta) \to 1$. The second main result of this paper is the following, which in particular confirms this conjecture.

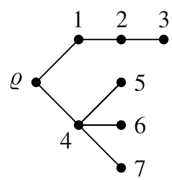

Figure 1: The largest 1-ary subtree of this tree has four vertices; namely, ϱ, 1, 2, and 3. However, if we delete all but the largest branch at each vertex we will obtain a 1-ary subtree with three vertices, such as ϱ, 4, 5. Although this not a d-ary tree for any d, as the vertices do not have the same number of offspring, similar d-ary examples exist.

https://doi.org/10.1017/jpr.2016.44 Published online by Cambridge University Press

Theorem 2. *There exists a fixed $\delta < 1$ such that $\mathcal{L}_t < t^{\delta}$ eventually.*

We can take $\delta = 1 - 4 \times 10^{-8}$, as shown in [2, Section 5].

Regarding lower bounds, it was proved in [6] that $\mathcal{L}_t > (2t+1)^{\log 2/\log 3}$ deterministically, and that $\mathbb{E}[\mathcal{L}_t] = \Omega(t^{0.88})$.

We prove the two main theorems by studying a third object, an infinite tree with weighted vertices, which is introduced in Section 2. Then we prove Theorems 1 and 2 in Sections 3 and 4, respectively. In particular, in the proof of Theorem 2, both infinite 3-ary and 9-ary trees will play a major role.

2. Subtrees of an infinite d-ary tree

We first give an informal description of this section. Fix a positive integer d and consider an infinite rooted d-ary tree, and suppose that a unit amount of sand is injected into the root. The sand then goes down the tree, i.e. the root distributes the sand to its children, and every other vertex does the same (vertices do not keep any sand). The distribution is done randomly: to each vertex ν is associated a random variable $X_\nu \in [0, 1]$, which is the portion of the parent's sand that ν receives. For simplicity, we assume that all vertices use independent copies of a certain random vector for distributing the sand among their children. The 'mass' of a vertex is simply the amount of sand it receives from its parent and distributes to its children. The 'mass' of a finite subtree is just the total mass of its leaves.

Let $r < d$. If each vertex distributes its sand evenly among its children, the mass of every complete r-ary tree of depth n that contains the root would be exactly $(1 - r/d)^n$. The aim of this section is to derive, in the case when the distribution is not too unbalanced, a uniform upper bound of the form $1/\kappa^n$ for the mass of all complete r-ary subtrees of depth n. We would like to argue that in each r-ary subtree, by each level we go down the tree, we lose a $1/\kappa$ fraction of mass on average. We argue this vertex by vertex: for each vertex ν we will define a random variable Υ_ν, which is a lower bound on the portion of mass 'lost' at this vertex by omitting any $d - r$ of its branches. Since vertices behave similarly, all random variables Υ_ν have the same distribution as some random variable Υ. In the main result of this section, Lemma 1, we show that our assertion holds if κ and Υ satisfy a certain 'moment-type' inequality (2).

We now give the formal definitions and arguments. Fix positive integers r and d with $r < d$ and let $\mathcal{T}_{d,\infty}$ be an infinite rooted d-ary tree. Denote the root by ϱ. We denote by $[\nu, \mu]$ the vertices in the path connecting ν to μ, including these two vertices. For a vertex ν, denote its distance from ϱ by $|\nu|$, and its offspring by νi, with $i \in \{1, 2, \dots, d\}$. For $\nu \neq \varrho$, denote by ν^- the parent of ν, i.e. the neighbour μ of ν with $|\mu| = |\nu| - 1$.

To each vertex ν assign a random variable $X_\nu \in (0, 1]$, satisfying the following properties. We have $X_\varrho = 1$. The random variables X_ν, for $\nu \in V(\mathcal{T}_{d,\infty}) \setminus \{\varrho\}$, are identically distributed. Moreover, we assume that the vectors $(X_{\nu 1}, X_{\nu 2}, \dots, X_{\nu d})$ are identically distributed and independent, and that $\sum_{i=1}^{d} X_{\nu i} = 1$. For any vertex ν, define the random variable

$$\Upsilon_\nu = \min\{X_{\nu i_1} + X_{\nu i_2} + \cdots + X_{\nu i_{d-r}} : 1 \le i_1 < i_2 < \cdots < i_{d-r} \le d\}, \tag{1}$$

and let $\Upsilon = \Upsilon_\nu$ for an arbitrary ν.

For each vertex $\nu \in V(\mathcal{T}_{d,\infty})$, define

$$\text{mass}(\nu) = \prod_{\sigma \in [\varrho, \nu]} X_\sigma,$$

and for any set of vertices $A \subset V(\mathcal{T}_{d,\infty})$, let $\text{mass}(A) = \sum_{\nu \in A} \text{mass}(\nu)$.

https://doi.org/10.1017/jpr.2016.44 Published online by Cambridge University Press

Given a nonnegative integer n, consider *level* n of $\mathcal{T}_{d,\infty}$, i.e. the set of vertices at distance n from ϱ. Denote by $\mathcal{G}_{n,r}$ the collection of subsets of at most r^n vertices at level n, with the additional property that they belong to the same r-ary subtree.

The main result of this section is the following.

Lemma 1. *Let λ and κ be positive constants satisfying*

$$d\kappa^{\lambda}\mathbb{E}[(1-\Upsilon)^{\lambda}] < 1. \tag{2}$$

Then eventually, for large n,

$$\max_{C\in\mathcal{G}_{n,r}} \operatorname{mass}(C) \le \kappa^{-n}.$$

Remark 1. This lemma is used twice in this paper with different distributions for Υ and $(X_{\nu 1}, X_{\nu 2}, \ldots, X_{\nu d})$.

To prove this lemma, we first formalize the connection between r-ary subtrees and the random variables Υ_ν via an intermediate quantity which is called the *adjusted mass*. Then we use a standard technique and the independence of the Υ_ν to prove that these variables' average behaviour combine to give the upper bound asserted by the lemma. For each vertex ν, we define its adjusted mass, denoted by $\operatorname{AM}(\nu)$, as

$$\operatorname{AM}(\nu) = \operatorname{mass}(\nu) \prod_{\sigma\in[\varrho,\nu^-]} \frac{1}{1-\Upsilon_\sigma}.$$

Returning to the sand example, this is the amount of mass that is *not* lost on the path from the root to ν, and it is defined this way precisely because of the following lemma. For any $A \subset V(\mathcal{T}_{d,\infty})$, let $\operatorname{AM}(A) = \sum_{\nu\in A}\operatorname{AM}(\nu)$.

Lemma 2. *For every positive integer n and every $C \in \mathcal{G}_{n,r}$, we have* $\operatorname{AM}(C) \le 1$.

Proof. Let $C \in \mathcal{G}_{n,r}$. Define $\operatorname{tree}_r(C)$ to be the r-ary subtree of $\mathcal{T}_{d,\infty}$ whose leaves are precisely the vertices of C. For each vertex ν of $\operatorname{tree}_r(C)$, denote its set of offspring in $\operatorname{tree}_r(C)$ by ν_{off}. Then by the definition of Υ in (1),

$$\operatorname{mass}(\nu_{\text{off}}) \le (1-\Upsilon_\nu)\operatorname{mass}(\nu).$$

Thus, by the definition of $\operatorname{AM}(\cdot)$, we have $\operatorname{AM}(\nu_{\text{off}}) \le \operatorname{AM}(\nu)$. Hence, for any $1 \le k \le n$,

$$\sum_{\substack{\mu\in\operatorname{tree}_r(C)\\ |\mu|=k}} \operatorname{AM}(\mu) \le \sum_{\substack{\nu\in\operatorname{tree}_r(C)\\ |\nu|=k-1}} \operatorname{AM}(\nu).$$

Iterating this, we obtain

$$\operatorname{AM}(C) = \sum_{\nu\in C}\operatorname{AM}(\nu) \le \operatorname{AM}(\varrho) = 1. \qquad \square$$

Proof of Lemma 1. For any positive integer n, define the event

$$\mathcal{C}_n = \bigcap_{\{\nu:\,|\nu|=n\}} \Bigg\{ \prod_{\sigma\in[\varrho,\nu^-]} (1-\Upsilon_\sigma)^{-1} \ge \kappa^n \Bigg\} = \bigcap_{\{\nu:\,|\nu|=n\}} \Big\{ \operatorname{AM}(\nu) \ge \operatorname{mass}(\nu)\kappa^n \Big\}.$$

https://doi.org/10.1017/jpr.2016.44 Published online by Cambridge University Press

We claim that $\mathcal{C}_n$ holds eventually. If this is the case, then for large enough n and any $C \in \mathcal{G}_{n,r}$, we have

$$\text{mass}(C) \le \kappa^{-n}\text{AM}(C) \le \kappa^{-n},$$

where the last inequality is a consequence of Lemma 2.

To complete the proof, we will show that $\mathcal{C}_n$ holds eventually. By the first Borel–Cantelli lemma, it is enough to show that

$$\sum_{n=1}^{\infty} d^n \mathbb{P}\bigg(\prod_{\sigma \in [\varrho, \nu^-]} (1-\Upsilon_\sigma)^{-1} < \kappa^n\bigg) < \infty, \tag{3}$$

where $\nu = \nu(n)$ denotes an arbitrary vertex with $|\nu| = n$. Since the Υ_σ are independent and $\lambda > 0$, the above probability is, by Markov's inequality,

$$\mathbb{P}\bigg(\prod_{\sigma \in [\varrho, \nu^-]} (1-\Upsilon_\sigma)^{\lambda} > \kappa^{-\lambda n}\bigg) \le \mathbb{E}\bigg[\prod_{\sigma \in [\varrho, \nu^-]} (1-\Upsilon_\sigma)^{\lambda}\bigg]\kappa^{\lambda n} = (\mathbb{E}[(1-\Upsilon)^{\lambda}]\kappa^{\lambda})^n.$$

Equation (3) now follows from (2). □

3. Largest r-ary subtrees of random d-ary trees

As the beta and Dirichlet distributions play an important role in this paper, we recall their definitions.

Definition 1. (*Beta and Dirichlet distributions.*) Let $\Gamma(t) = \int_0^\infty x^{t-1}\mathrm{e}^{-x}\,\mathrm{d}x$. For positive numbers α and β, a random variable is said to be distributed as beta(α, β) if it has density

$$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \quad \text{for } x \in (0, 1).$$

The multivariate generalization of the beta distribution is called the Dirichlet distribution. Let $\alpha_1, \alpha_2, \dots, \alpha_n$ be positive numbers. The Dirichlet$(\alpha_1, \alpha_2, \dots, \alpha_n)$ distribution has support on the set

$$\bigg\{(x_1, x_2, \dots, x_n) \colon x_i \ge 0 \text{ for } 1 \le i \le n, \text{ and } \sum_{i=1}^n x_i = 1\bigg\},$$

and its density at point $(x_1, x_2, \dots, x_n)$ is equal to

$$\frac{\Gamma(\sum_{i=1}^n \alpha_i)}{\prod_{i=1}^n \Gamma(\alpha_i)} \prod_{j=1}^n x_j^{\alpha_j - 1}.$$

Note that if the vector $(X_1, X_2, \dots, X_n)$ is distributed as Dirichlet$(\alpha_1, \alpha_2, \dots, \alpha_n)$, then the marginal distribution of X_i is beta$(\alpha_i, \sum_{j \ne i} \alpha_j)$.

Let r and d be fixed positive integers with $r < d$. Let $(B_1, B_2, \dots, B_d)$ be a random vector distributed as Dirichlet$(1/(d-1), 1/(d-1), \dots, 1/(d-1))$, and define the random variable Υ as

$$\Upsilon = \min\{B_{i_1} + B_{i_2} \cdots + B_{i_{d-r}} \colon 1 \le i_1 < i_2 < \cdots < i_{d-r} \le d\}.$$

The main theorem of this section is the following.

https://doi.org/10.1017/jpr.2016.44 Published online by Cambridge University Press

Theorem 3. *Let r and d be fixed positive integers with $r < d$, and let S_t denote the size of the largest r-ary subtree of a random d-ary recursive tree at time t. Let τ, κ, and λ be positive constants satisfying (2) and*

$$\mathrm{e}d \log \tau < (d-1)\tau^{1/(d-1)}, \tag{4}$$

and let $n = \lfloor \log t / \log \tau \rfloor$. There exists a constant K such that eventually (for large t)

$$S_t \le K(r^n + t\kappa^{-n}). \tag{5}$$

Before proving this theorem, we show that it implies Theorem 1.

Proof of Theorem 1. We show that there exist positive constants τ, κ, and λ satisfying (4), (2), and $\kappa > 1$; then we would have $\tau > \mathrm{e}^{d-1} > r$, and the theorem follows from Theorem 3 by choosing any $\delta \in (\max\{1 - \log \kappa / \log \tau, \log r / \log \tau\}, 1)$. As (4) holds for all large enough τ, it suffices to show that there exist $\kappa > 1$ and $\lambda > 0$ satisfying (2). Since $\lim_{\varepsilon \to 0} \mathbb{P}(\Upsilon < \varepsilon) = 0$, we have $\mathbb{E}[(1-\Upsilon)^\lambda] \to 0$ as $\lambda \to \infty$. In particular, there exists $\lambda > 0$ such that $\mathbb{E}[(1-\Upsilon)^\lambda] < 1/d$. Then, we can let

$$\kappa = (d\mathbb{E}[(1-\Upsilon)^\lambda])^{-1/(2\lambda)},$$

which is strictly larger than 1. □

We first provide a high-level idea of the proof of Theorem 3. Observe that a random d-ary recursive tree has $1 + td$ vertices at time t. Think of the generation of a random d-ary recursive tree as a process of storing $1 + td$ balls in the nodes of an infinite d-ary tree in the following manner. Suppose that $1 + td$ balls are injected into the root of our infinite tree. Whenever a vertex receives some balls, one ball is stored in the vertex and will not move any further; the rest, if any, are distributed among its children according to an appropriate variation of a Dirichlet distribution. Once all balls are fixed, the vertices with balls induce a $(1 + td)$-vertex d-ary subtree which, as we show in Lemma 3, corresponds to a random d-ary recursive tree at time t.

We analyse the size of the largest r-ary subtree of this random d-ary recursive tree by approximating this process with the sand distribution process described in the previous section. Note that the number of balls at level n or less is clearly $O(r^n)$. Also, the number of balls at levels greater than n can be bounded by the 'mass' of the corresponding r-ary subtree in the infinite tree, which gives the term $O(t\kappa^{-n})$ in (5). The main issue in building this connection is that the amount of sand can be a real number but the number of balls is always an integer; nevertheless, we can show that these quantities are close enough for our purposes.

We now provide the formal argument. In the rest of this section, $\mathcal{T}_{d,\infty}$ denotes an infinite d-tree rooted at ϱ. We view the random recursive d-ary tree $\mathcal{T}_{d,t}$ as a subtree of $\mathcal{T}_{d,\infty}$. At each time step, we assign a *weight* to each vertex. For each t and each vertex $\nu \in V(\mathcal{T}_{d,\infty})$, if $\nu \notin V(\mathcal{T}_{d,t})$ then let $\mathrm{weight}(\nu, t) = 0$, otherwise, let $\mathrm{weight}(\nu, t)$ denote the number of nonleaf vertices in the subtree of $\mathcal{T}_{d,t}$ rooted at ν.

Lemma 3. *There exist random variables $\{B_\nu\}_{\nu \in V(\mathcal{T}_{d,\infty})}$, with $B_\varrho = 1$, such that for any positive integer t and any $\nu \in V(\mathcal{T}_{d,\infty})$, $\mathrm{weight}(\nu, t)$ is stochastically dominated by a $\mathrm{binomial}(t, \prod_{\sigma \in [\varrho, \nu]} B_\sigma)$ random variable. Moreover, the vectors $(B_{\nu 1}, B_{\nu 2}, \dots, B_{\nu d})$ are independent for $\nu \in V(\mathcal{T}_{d,\infty})$, and are distributed as $\mathrm{Dirichlet}(1/(d-1), 1/(d-1), \dots, 1/(d-1))$.*

Proof. Consider a vertex $\nu \neq \varrho$ and a positive integer t such that $\nu \in V(\mathcal{T}_{d,t})$. Note that at time t, the number of leaves in the branch at v is $(d-1)\mathrm{weight}(\nu, t) + 1$. Hence, given that at

https://doi.org/10.1017/jpr.2016.44 Published online by Cambridge University Press

time $t+1$ the weight of ν^- increases, the probability, conditional on the past, that the weight of ν increases at the same time, is equal to

$$\frac{(d-1)\text{weight}(\nu,t)+1}{(d-1)\text{weight}(\nu^-,t)+1}.$$

Each time a weight increases, its increment is exactly 1. By identifying ν with one colour and its siblings with another colour, the evolution of the numerator of the above expression over time can be modelled using an Eggenberger–Pólya urn, with initial conditions $(1, d-1)$ and reinforcement equal to $d-1$. Moreover, the urns corresponding to distinct vertices are mutually independent.

The limiting distribution describing the Eggenberger–Pólya urn is well known, but we require bounds applying for all t. To this end, we can express the number of times a given colour is chosen by time t in an Eggenberger–Pólya urn as a mixture of binomials with respect to a beta distribution. See, for example, [11, Lemma 1]. In this case, given the initial conditions $(1, d-1)$ and reinforcement $d-1$, the mixture is with respect to beta$(1/(d-1), 1)$. This means that to each vertex ν, we can assign a random variable B_ν distributed as beta$(1/(d-1), 1)$, such that weight(ν, t) conditional on B_ν is binomially distributed with parameters weight$(\nu^-, t)-1$ and B_ν. Set $B_\varrho = 1$ and note that weight$(\varrho, t) = t$. By induction, weight(ν, t), conditional on $\{B_\sigma\}_{\sigma\in[\varrho,\nu]}$, is stochastically smaller than a binomial$(t, \prod_{\sigma\in[\varrho,\nu]} B_\sigma)$.

By the Eggenberger–Pólya urn representation, we also infer that the joint distribution of $(B_{\nu 1}, B_{\nu 2}, \dots, B_{\nu d})$ is Dirichlet$(1/(d-1), 1/(d-1), \dots, 1/(d-1))$ for all ν. See, for example, [11, Lemma 1]. □

Lemma 4. *Let* $B_1, \dots, B_n$ *be independent* beta$(1/(d-1), 1)$ *random variables. For all positive* β*, we have*

$$\mathbb{P}\bigg(\prod_{i=1}^n B_i \le \beta^n\bigg) \le \bigg(\frac{\mathrm{e}\log(1/\beta)\beta^{1/(d-1)}}{d-1}\bigg)^n.$$

Proof. If $\beta \ge \mathrm{e}^{1-d}$ then the right-hand side is at least 1, so we may assume that $0<\beta<\mathrm{e}^{1-d}$. We use Chernoff's technique. Let $\lambda = -1/(d-1) - 1/\log\beta$. We have

$$\mathbb{E}[B_1^\lambda] = \frac{\Gamma(d/(d-1))}{\Gamma(1/(d-1))\Gamma(1)}\int_0^1 x^\lambda x^{-1+1/(d-1)}\,\mathrm{d}x = \frac{1}{(d-1)\lambda+1}.$$

Hence, by Markov's inequality and since the B_i are independent,

$$\mathbb{P}\bigg(\prod_{i=1}^n B_i \le \beta^n\bigg) \le \prod_{i=1}^n \frac{\mathbb{E}[B_1^\lambda]}{\beta^\lambda} = \bigg(\frac{1}{\beta^\lambda((d-1)\lambda+1)}\bigg)^n = \bigg(\frac{\mathrm{e}\log(1/\beta)\beta^{1/(d-1)}}{d-1}\bigg)^n. \qquad \square$$

Proof of Theorem 3. Let $\{B_\nu\}_{\nu\in V(\mathcal{T}_{d,\infty})}$ be as given by Lemma 3. Denote by $\mathcal{G}_{n,r}$ the collection of subsets of the vertices of $\mathcal{T}_{d,\infty}$ at level n, with the property that they belong to the same r-ary subtree. We apply Lemma 1 with mass defined using $X_\sigma = B_\sigma$. Since (2) holds, we conclude that eventually, for large n,

$$\max_{C\in\mathcal{G}_{n,r}} \text{mass}(C) \le \kappa^{-n}. \tag{6}$$

By Lemma 3, weight(ν, t) is stochastically dominated by a binomial$(t, \text{mass}(\nu))$. Chernoff's bound for binomials implies that (see, e.g. [9, Theorem 2.3(b)])

$$\mathbb{P}(\text{weight}(\nu,t) \ge 3t\,\text{mass}(\nu) \mid \text{mass}(\nu) \ge q) \le \mathrm{e}^{-tq}$$

https://doi.org/10.1017/jpr.2016.44 Published online by Cambridge University Press

for every positive q. Since τ satisfies (4), there exists $\tau_1 < \tau$ satisfying

$$\mathrm{e}d \log \tau_1 < (d-1)\tau_1^{1/(d-1)}. \tag{7}$$

Let $\beta = 1/\tau_1$. By Lemma 4, for any vertex ν at level n,

$$\mathbb{P}(\mathrm{mass}(\nu) < \beta^n) \le \bigg(\frac{\mathrm{e}\log(1/\beta)\beta^{1/(d-1)}}{d-1}\bigg)^n.$$

Note that (7) implies that the term in brackets is a constant smaller than $1/d$. We have

$$\begin{aligned}\mathbb{P}\bigg(\bigcup_{\{\mu\colon |\mu|=n\}} \{\mathrm{weight}(\mu,t) \ge 3t\mathrm{mass}(\mu)\}\bigg) &\le d^n \mathbb{P}(\mathrm{weight}(\nu,t) \ge 3t\mathrm{mass}(\nu)) \\ &\le d^n\mathbb{P}(\mathrm{mass}(\nu) < \beta^n) + d^n \mathrm{e}^{-t\beta^n}.\end{aligned}$$

The last expression is summable in n, as $t^{1/n}\beta \ge \tau\beta$ (recall that $n = \lfloor \log t/\log\tau \rfloor$), and $\tau\beta$ is a constant larger than 1. By the first Borel–Cantelli lemma eventually, for large n, we have $\mathrm{weight}(\mu, t) < 3t\mathrm{mass}(\mu)$ for all μ at level n. Combining this and (6) eventually, for large n, we have

$$\max\bigg\{\sum_{\nu\in C} \mathrm{weight}(\nu,t)\colon C \in \mathcal{G}_{n,r}\bigg\} < 3t\kappa^{-n}.$$

Since an r-ary subtree contains at most r^{n+1} vertices at levels n or less, the size of an r-ary subtree is bounded by $r^{n+1} + d\sum_{\nu\in C}\mathrm{weight}(\nu,t)$, where $C \in \mathcal{G}_{n,r}$ are the vertices of the subtree at level n. Thus, eventually, for large t, we have $S_t < r^{n+1} + 3dt\kappa^{-n}$, as required. □

4. Longest paths in random Apollonian networks

We define a tree $\mathcal{T}_{3,t}$, called the $\triangle$-*tree* of a RAN_t, as follows. There is a one-to-one correspondence between the triangles in a RAN_t and the vertices of $\mathcal{T}_{3,t}$. For every triangle $\triangle$ in a RAN_t, we denote its corresponding vertex in $\mathcal{T}_{3,t}$ by $\boldsymbol{v}^{\triangle}$. To construct $\mathcal{T}_{3,t}$, start with a single root vertex ϱ, which corresponds to the initial triangle of a RAN_t. Wherever a triangle $\triangle$ is subdivided into triangles $\triangle_1$, $\triangle_2$, and $\triangle_3$, generate three offspring $\boldsymbol{v}^{\triangle_1}$, $\boldsymbol{v}^{\triangle_2}$, and $\boldsymbol{v}^{\triangle_3}$, for $\boldsymbol{v}^{\triangle}$, and extend the correspondence in the natural manner (see Figure 2 for an illustration). Note that $\mathcal{T}_{3,t}$ is a random 3-ary recursive tree as defined in the introduction and has $3t+1$ vertices and $2t+1$ leaves. The leaves of $\mathcal{T}_{3,t}$ correspond to the bounded faces of a RAN_t. Let $\mathcal{T}_{3,\infty}$ denote an infinite 3-ary tree rooted at ϱ. We view $\mathcal{T}_{3,t}$ as a subtree of $\mathcal{T}_{3,\infty}$. For each vertex $\nu \in V(\mathcal{T}_{3,\infty})$, the *grand-offspring* of ν are its descendants at level $|\nu|+2$. For a triangle $\triangle$ in a RAN_t, $I(\triangle)$ denotes the set of vertices of a RAN_t that are *strictly inside* $\triangle$.

The following lemma introduces the connection with the setup in Section 3. The proof is a simple exercise and can be found in [6, Lemma 3.1].

Lemma 5. *Let $\mathcal{T}_{3,t}$ be the $\triangle$-tree of a RAN_t. Let $\boldsymbol{v}^{\triangle}$ be a vertex of $\mathcal{T}_{3,t}$ with nine grand-offspring $\boldsymbol{v}^{\triangle_1}, \boldsymbol{v}^{\triangle_2}, \dots, \boldsymbol{v}^{\triangle_9}$ in $V(\mathcal{T}_{3,t})$. Then the vertex set of a path in a RAN_t intersects at most eight of the $I(\triangle_i)$s.*

We say that a finite subtree $\mathcal{J}$ of $\mathcal{T}_{3,\infty}$ is *buono* if each vertex of $\mathcal{J}$ has at most eight grand-offspring in $\mathcal{J}$. The motivation for this definition becomes clear in the derivation of Theorem 2 from Theorem 4 below.

https://doi.org/10.1017/jpr.2016.44 Published online by Cambridge University Press

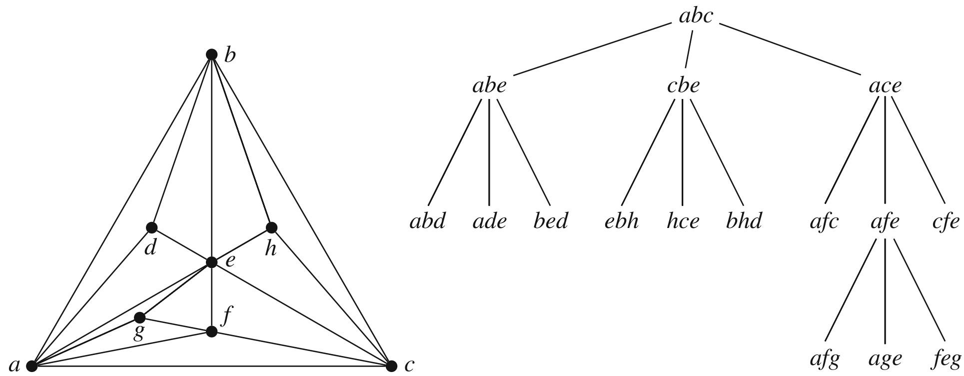

FIGURE 2: A realization of a RAN_5 (*left*) and its corresponding $\triangle$-tree (*right*).

Assume that the four vectors

$$(A_1, A_2, A_3), \qquad (B_{1,1}, B_{1,2}, B_{1,3}), \qquad (B_{2,1}, B_{2,2}, B_{2,3}), \qquad \text{and} \qquad (B_{3,1}, B_{3,2}, B_{3,3})$$

are independent and identically distributed random vectors distributed as $\mathrm{Dirichlet}(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. Define

$$\Upsilon = \min\{A_i B_{i,j} : 1 \le i,\ j \le 3\}.$$

The main theorem of this section is the following.

Theorem 4. *Let τ, κ, λ be positive constants satisfying*

$$3\mathrm{e} \log \tau < 2\sqrt{\tau}, \tag{8}$$

$$9\kappa^{\lambda}\mathbb{E}[(1-\Upsilon)^{\lambda}] < 1, \tag{9}$$

and let n be the largest even integer smaller than $\log t / \log \tau$. Then, there exists a constant K, such that eventually, for large t, the largest buono subtree of $\mathcal{T}_{3,t}$ has at most $K(8^{n/2} + t\kappa^{-n/2})$ vertices.

Remark 2. Note that (8) is precisely (4) of Theorem 3 with $d = 3$, and (9) is precisely (2) of Lemma 1 with $d = 9$. We work mainly with ternary trees in this section. Equation (9) is only used inside the proof of Lemma 6.

We show how this theorem implies Theorem 2.

Proof of Theorem 2. Following the proof of Theorem 1, we can find positive constants τ, κ, and λ satisfying the conditions of Theorem 4, with $\kappa > 1$ and $\tau > 8$. Choose any $\delta \in (\max\{1 - \log(\kappa)/(2\log\tau), \log(8)/2\log\tau\}, 1)$.

Let P be a path in a RAN_t and let $R(P)$ denote the set of vertices $\boldsymbol{v}^{\triangle}$ of $\mathcal{T}_{3,t}$ such that $I(\triangle)$ contains some vertex of P. By Lemma 5, $R(P)$ induces a buono subtree of $\mathcal{T}_{3,t}$. Hence, using Theorem 4 for the second inequality eventually, for large t, we have

$$|V(P)| \le 3 + |R(P)| \le 3 + K(8^{n/2} + t\kappa^{-n/2}) < t^{\delta}. \qquad \square$$

The rest of this section is devoted to the proof of Theorem 4. Note that the setup is exactly that in the previous section with $d = 3$, except we would like to bound the size of a largest *buono*

https://doi.org/10.1017/jpr.2016.44 Published online by Cambridge University Press

subtree and not that of a largest r-*ary* subtree. However, recall that in the proof of Theorem 3 the only fact we used about r-ary subtrees was (6); namely, that $\max_{C\in\mathcal{G}_{n,r}} \text{mass}(C) \le \kappa^{-n}$. So if we could prove a similar inequality for buono subtrees, exactly the same proof works for Theorem 4. Denote by $\mathcal{B}_n$ the collection of subsets of $V(\mathcal{T}_{3,\infty})$ at level n, with the property that they belong to the same buono subtree. Note that each element of $\mathcal{B}_{2n}$ has at most 8^n vertices. The following lemma provides a parallel for (6).

Lemma 6. *Let λ and κ be positive constants satisfying (9). Eventually, for large n,*

$$\max_{C\in\mathcal{B}_n} \text{mass}(C) \le \kappa^{-n/2}.$$

Proof. Note that buono subtrees look very much like 8-ary subtrees of a 9-ary tree if we look at every second level. More precisely, let $\mathcal{T}_{9,\infty}$ be an infinite rooted 9-ary tree obtained from $\mathcal{T}_{3,\infty}$ as follows. The vertices of $\mathcal{T}_{9,\infty}$ are the vertices of $\mathcal{T}_{3,\infty}$ at even levels. A vertex μ is an offspring of ν in $\mathcal{T}_{9,\infty}$ if μ is a grand-offspring of ν in $\mathcal{T}_{3,\infty}$. To each vertex μ assign the random variable $X_\mu = B_\mu B_{\mu^-}$. Buono subtrees of $\mathcal{T}_{3,\infty}$ are translated into 8-ary subtrees of $\mathcal{T}_{9,\infty}$. Note that (9) is precisely (2) of Lemma 1 for $d = 9$. Applying Lemma 1 to $\mathcal{T}_{9,\infty}$ (with $r = 8$ and $d = 9$) concludes the proof. □

Proof of Theorem 4. The proof is almost identical to that of Theorem 3 for $d = 3$, with just two differences, as we want to bound the size of the largest buono subtree instead of that of the largest r-ary subtree. First, we bound the number of vertices at levels n or less by $8^{n/2}$ instead of r^{n+1}. For the remaining vertices, we bound $\max_{C\in\mathcal{B}_n} \text{mass}(C)$ via Lemma 6 instead of bounding $\max_{C\in\mathcal{G}_{n,r}} \text{mass}(C)$ via Lemma 1. □

Acknowledgements

The authors thank the anonymous referee for their beneficial comments which resulted in an improved presentation. A.C. was supported by an ARC Discovery Project (grant number DP140100559) and the STREP project 'Mathemacs'. A.M. was supported by the Vanier Canada Graduate Scholarships program. Most of this work was done while the author was visiting Monash University, Australia. N.W. was supported by an Australian Laureate Fellowships (grant number FL120100125).

References

[1] Albenque, M. and Marckert, J.-F. (2008). Some families of increasing planar maps. *Electron. J. Prob.* **13,** 1624–1671.

[2] Collevecchio, A., Mehrabian, A. and Wormald, N. (2014). Longest paths in Apollonian networks and largest r-ary subtrees of random d-ary recursive trees. Available at https://arxiv.org/abs/1404.2425.

[3] Cooper, C. and Frieze, A. (2015). Long paths in random Apollonian networks. *Internet Math.* **11,** 308–318.

[4] Cooper, C., Frieze, A. and Uehara, R. (2014). The height of random k-trees and related branching processes. *Random Structures Algorithms* **45,** 675–702.

[5] Drmota, M. (2009). *Random Trees: An Interplay Between Combinatorics and Probability*. Springer, Vienna.

[6] Ebrahimzadeh, E. *et al.* (2014). On longest paths and diameter in random Apollonian networks. *Random Structures Algorithms* **45,** 703–725.

[7] Frieze, A. and Tsourakakis, C. E. (2014). Some properties of random Apollonian networks. *Internet Mathematics* **10,** 162–187.

[8] Kolossváry, I., Komjáthy, J. and Vágó, L. (2013). Degrees and distances in random and evolving Apollonian networks. Preprint. Available at https://arxiv.org/abs/1310.3864v1.

[9] McDiarmid, C. (1998). Concentration. In *Probabilistic Methods for Algorithmic Discrete Mathematics* (Algorithms Combin. **16**), Springer, Berlin, pp. 195–248.

[10] MUNGAN, M. (2011). Comment on ‘Apollonian networks: Simultaneously scale-free, small world, Euclidean, space filling, and with matching graphs’. *Phys. Rev. Lett.* **106,** 029802.
[11] PEMANTLE, R. (1988). Phase transition in reinforced random walk and RWRE on trees. *Ann. Prob.* **16,** 1229–1241.
[12] ZHANG, Z., COMELLAS, F., FERTIN, G. AND RONG, L. (2006). High-dimensional Apollonian networks. *J. Phys. A* **39,** 1811–1818.
[13] ZHOU, T., YAN, G. AND WANG, B.-H. (2005). Maximal planar networks with large clustering coefficient and power-law degree distribution. *Phys. Rev. E* **71,** 046141.