Adv. Appl. Prob. **48,** 163–175 (2016)
*doi:*10.1017/apr.2015.11

THE DEGREE PROFILE AND WEIGHT IN APOLLONIAN NETWORKS AND k-TREES

PANPAN ZHANG * ** AND
HOSAM MAHMOUD,* *The George Washington University*

Abstract

We investigate the degree profile and total weight in Apollonian networks. We study the distribution of the degrees of vertices as they age in the evolutionary process. Asymptotically, the (suitably-scaled) degree of a node with a fixed label has a Mittag-Leffler-like limit distribution. The degrees of nodes of later ages have different asymptotic distributions, influenced by the time of their appearance. The very late arrivals have a degenerate distribution. The result is obtained via triangular Pólya urns. Also, via the Bagchi–Pal urn, we show that the number of terminal nodes asymptotically follows a Gaussian law. We prove that the total weight of the network asymptotically follows a Gaussian law, obtained via martingale methods. Similar results carry over to the sister structure of the k-trees, with minor modification in the proof methods, done mutatis mutandis.

Keywords: Random structure; network; random graph; self-similarity; degree profile; phase transition; stochastic recurrence; Pólya urn; martingale

2010 Mathematics Subject Classification: Primary 05082; 90B15
Secondary 60C05

1. Introduction

The ancient Greek mathematician and astronomer Apollonius of Perga (c.262–190 BC) is known for his work on conic sections. His work on circle packing instigated several related dual problems on triangulation (dividing a triangle into smaller ones). In our modern times, graphs based on triangulation (or more generally on enriching simplexes of higher order) are dubbed Apollonian networks, a term perhaps first coined in [1] in 2005. For motivation and important applications, the title of [1] sums it all—'Apollonian networks: simultaneously scale-free, small world, Euclidean, space filling, and with matching graphs'. Each of these phrases is a significant area of modern research by itself. A random structure related to Apollonian networks is the random k-tree. We shall discuss it as an offshoot.

2. Random Apollonian networks

For a general index $k \geq 1$, a *random Apollonian network of index k* is a random graph that grows out of an active k-clique (simplex, the complete graph on k vertices) with vertices labeled 0. At the nth step, an active k-clique is chosen at random (all active cliques being equally likely). We allocate a new vertex, connect it with k edges to the k vertices of the clique, and we then discard the chosen clique; it becomes inactive. In Figure 1 we show the

Received 6 November 2014; revision received 4 April 2015.
* Postal address: Department of Statistics, The George Washington University, 801 22nd St. NW, Washington, DC 20052, USA.
** Email address: pzhang@email.gwu.edu

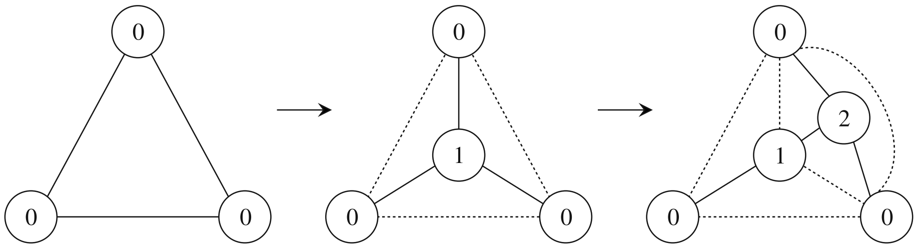

FIGURE 1: The evolution of an Apollonian network of index 3 in two steps.

step-by-step growth of a random Apollonian network of index 3 from an initial triangle in two stages ($n = 2$). In Figure 1 a triangle is active if at least one of its sides is drawn solid, whereas the three sides of an inactive triangle are all dashed.

At the nth step, we add k cliques and deactivate one, a net gain of $(k-1)$ cliques of order k each. Thus, $\tau_n^{(k)}$, the total number of active cliques at time n, is given by

$$\tau_n = \tau_n^{(k)} = (k-1)n + 1.$$

The index k is a fixed parameter of the network, and all the random variables we discuss depend on it. We shall occasionally suppress explicit dependence on k to simplify the notation.

3. Scope and results

Some aspects of Apollonian networks have been investigated in [4]–[6] and [9]. We were inspired by these very recent works to investigate additional properties.

The study of the distribution of degrees in random networks has been a popular topic. Knowing the degree of a node can tell us how popular the node is in a social network, or how much demand there is on it in a routeing network, which can help allocate the appropriate resources.

Exact results about the degree profile in Apollonian networks are in terms of Pochhammer's symbol for the rising factorial and Stirling's numbers of the second kind. Pochhammer's symbol for the rising factorial is $\langle x\rangle_s = x(x+1)\cdots(x+s-1)$, for any $x \in \mathbb{R}$ and any integer $s \geq 0$, with the interpretation that $\langle x\rangle_0 = 1$. The numbers $\left\{ {r \atop i} \right\}$ are Stirling's numbers of the second kind [10].

Let $D_{j,n}^{(k)}$, for $j = 1, \dots, n$, be the degree of the node labeled $j \leq n$ in an Apollonian network of index k and of age n. The random variables $D_{j,n}^{(k)}$, for $j = 1, \dots, n$, describe a profile of degrees in the random graph.

Proposition 1. *The exact moments of* $D_{j,n}^{(k)}$ *are*

$$\begin{aligned}\mathbb{E}[(D_{j,n}^{(k)})^s] = \frac{1}{(k-2)^s}\Bigg((k(k-3))^s + \sum_{r=1}^{s}\binom{s}{r}(k(k-3))^{s-r} \\ \times \left((k-2)^r\left[\left\langle\frac{(k-1)j+1}{k-1}\right\rangle_{n-j}\right]^{-1}\right) \\ \times \sum_{i=1}^{r}(-1)^{r-i}\left\{ {r \atop i} \right\}\left\langle\frac{k}{k-2}\right\rangle_i\left\langle\frac{(k-1)j+1+(k-2)i}{k-1}\right\rangle_{n-j}\Bigg).\end{aligned}$$

https://doi.org/10.1017/apr.2015.11 Published online by Cambridge University Press

In particular, the exact mean of $D^{(k)}_{j,n}$ is

$$\mathbb{E}[D^{(k)}_{j,n}] = \frac{1}{(k-2)}\biggl(k(k-3) + \biggl(k\biggl[\biggl\langle \frac{(k-1)j+1}{k-1} \biggr\rangle_{n-j}\biggr]^{-1}\biggr)\langle j+1\rangle_{n-j}\biggr).$$

A proper normalization gives rise to a limit distribution following from the results in [7] and [19], represented in terms of the standard Riemann's gamma function.

Proposition 2. *As $n \to \infty$, the random variable $D^{(k)}_{j,n}/n^{(k-2)/(k-1)}$ converges in distribution to a random variable D^*_j, which is uniquely characterized by its moments; these moments are*

$$\mathbb{E}[(D^*_j)^s] = \Gamma\biggl(\frac{(k-1)j+1}{k-1}\biggr)\Gamma\biggl(\frac{k}{k-2}+s\biggr) \times \biggl[\Gamma\biggl(\frac{k}{k-2}\biggr)\Gamma\biggl(\frac{(k-1)j+1+(k-2)s}{k-1}\biggr)\biggr]^{-1}.$$

Distributions with moments of the form in Proposition 2 are called gamma-type distributions [15]. The Mittag-Leffler distribution is a special case [14]. When $j = j_n \to \infty$ and $j_n = o(n)$, $\tilde{D}_{j,n}/(n/j_n)^{(k-2)/(k-1)}$ converges in distribution to a random variable with moments $\Gamma(k/(k-2)+s)/\Gamma(k/(k-2))$. These are the moments of a gamma$(k/(k-2), 1)$ random variable.

In various types of random graphs, one considers the number of terminal nodes—nodes that appeared but have yet to participate in further recruiting, as, for instance, leaves in random trees, which were extensively studied in many tree families; see, for example, [13] and [16]. In Apollonian networks, terminal nodes are nodes of minimal degree. Understanding their count is a first step in constructing an alternative degree profile describing the number of nodes of each possible degree in Apollonian networks.

Proposition 3. *Let $T^{(k)}_n$ be the number terminal nodes in an Apollonian network of index k at age n. We then have*

$$\frac{T^{(k)}_n}{n} \to \frac{k-1}{2k-1} \quad \textit{almost surely}$$

and

$$\biggl(T^{(k)}_n - \frac{k-1}{2k-1}n\biggr)n^{-1/2} \xrightarrow{\mathrm{D}} \mathcal{N}\biggl(0, \frac{k(k-1)^2}{(2k-1)^2(3k-1)}\biggr).$$

An important property of an Apollonian network is the 'weight' assigned to a clique. The vertices of a clique have different labels assigned to them. Their sum is the total of the ages assigned at the vertices of the clique and represents a 'weight', and their average represents the mean age of the clique. This is a property at the microscale. A related, and equally important, macro characteristic is the total weight in the entire network (the sum of all the weights in all the cliques), which we call $Y^{(k)}_n$.

Theorem 1. *We have*

$$\frac{Y^{(k)}_n - (k-1)n^2}{n^{3/2}} \xrightarrow{\mathrm{D}} \mathcal{N}\biggl(0, \frac{(k-2)^2(k-1)^3}{k(k+1)(2k-1)}\biggr).$$

https://doi.org/10.1017/apr.2015.11 Published online by Cambridge University Press

4. Proofs

4.1. Degree profile of an Apollonian network

Upon inserting the nth node, we allocate a node labeled n and choose a parent clique for it. The parent (also called the recruiter) may or may not be incident with the node labeled j. If the recruiter is incident with the node labeled j, one edge is added to adjoin the nodes labeled j and n, increasing the degree of the node labeled j by 1. If the recruiter is not incident with the node labeled j, the degree of this node does not change after the nth insertion. This coincides with the dynamics of a Pólya urn scheme on two colors. The idea of embedding an Apollonian network into a Pólya urn was also used in [6].

The urn representing the evolution of the degree of a node of an Apollonian network is triangular, as we discuss next. Suppose that we wait till the node labeled j appears for the first time in the Apollonian network. At this point we color all the active cliques incident with the node labeled j with white and color all the other active cliques with blue, and think of them as the initial number of balls in a white–blue Pólya urn. There are k white cliques in the network (i.e. k white balls in the urn), and $\tau_j^{(k)} - k = (k-1)j + 1 - k$ blue cliques in the network (i.e. $(k-1)j + 1 - k$ blue balls in the urn).

Every time an active white clique is selected to be the recruiter, the degree of the node labeled j is increased by 1. Also, when k new cliques appear, only one of them is to be colored blue (i.e. one blue ball is added to the urn), the rest of the $k-1$ new cliques are to be colored white, and the recruiting clique (white) is deactivated, a net gain of $k-2$ white cliques (i.e. $k-2$ white balls are added to the urn). On the other hand, every time an active blue clique recruits, k new cliques appear, and none of them is incident with the node labeled j. Their appearance in the network does not change the number of cliques incident with the node labeled j. As the recruiting blue clique is rendered inactive, the net gain is $k-1$ blue cliques (i.e. $k-1$ blue balls are added to the urn). The replacement matrix associated with this urn is $\left(\begin{smallmatrix} k-2 & 1 \\ 0 & k-1 \end{smallmatrix}\right)$.

After $n-j$ draws, the urn represents the network at time n. There is a linear relation between $D_{j,n}^{(k)}$ and $W_{j,n-j}^{(k)}$:

$$W_{j,n-j}^{(k)} = (k-2)D_{j,n}^{(k)} - k(k-3) \quad \text{for } n \ge j.$$

We appeal to the recent results in [19] to determine the exact distribution of the number of k-cliques incident with the node labeled j as the network ages, proving Proposition 1.

The asymptotic moments of the scaled random variable D_j^* are obtained by applying Stirling's approximation to the exact moments of $D_{j,n}^{(k)}$, proving Proposition 2.

In particular, as $n \to \infty$, the asymptotic mean and variance of node j (fixed) are given by

$$\mathbb{E}[D_{j,n}^{(k)}] \sim \frac{k\Gamma(j + 1/(k-1))}{(k-2)\Gamma(j+1)} n^{(k-2)/(k-1)}$$

and

$$\operatorname{var}[D_{j,n}^{(k)}] \sim \frac{\Gamma(j+1/(k-1))}{(k-2)^2}\left(\frac{2k(k-1)}{\Gamma(j+(2k-3)/(k-1))} - \frac{k^2\Gamma(j+1/(k-1))}{\Gamma^2(j+1)}\right) n^{2(k-2)/(k-1)}.$$

The same methods are applicable to a choice of $j = j_n$ growing with n. We find two regimes: $j_n = o(n)$ is growing slowly with respect to n (e.g. $j_n = \lceil \ln n + 5 \rceil$), and j_n is of order n (e.g. $j_n = \lfloor \frac{7}{9}n + 4 \rfloor$).

https://doi.org/10.1017/apr.2015.11 Published online by Cambridge University Press

The asymptotic mean of a node with slowly growing label is

$$\mathbb{E}[D_{j,n}^{(k)}] \sim \frac{k}{k-2}\left(\frac{n}{j_n}\right)^{(k-2)/(k-1)}.$$

If $j_n \sim \alpha_n n$, with $0 < \alpha_n \le 1$, we have

$$\mathbb{E}[D_{j,n}^{(k)}] \sim \frac{k}{k-2}(k-3+\alpha_n^{-(k-2)/(k-1)}).$$

Note that, for the very late arrivals with $j_n \sim n$ (e.g. $j_n = n - \lfloor\sqrt[3]{n}\rfloor$), we have $\mathbb{E}[D_{j,n}^{(k)}] \sim k$. A similar analysis shows that these very late arrivals have asymptotic variance converging to 0. Thus, asymptotically these very late arrivals have a degenerate distribution.

4.2. Terminal nodes in random Apollonian networks

Let us color cliques incident with terminal nodes with white and all other cliques with blue. Consider the network at age 1 to be the initial condition of an urn having k white cliques and none blue.

Note that no two terminal nodes ever appear in the same clique. So, a count of white cliques translates directly into a count of terminal nodes (by dividing the former by k). When a white clique recruits (while active), the terminal node in it is no longer terminal. Thus, the other (active) $k-1$ white cliques incident with this (old) terminal node turn into blue. However, new cliques are born to replace the recruiter. The newborn cliques share one new terminal node connected to all k vertices of the old white recruiting clique (now inactive), and all k new cliques created are white; there is no net change in the number of white cliques. On the other hand, if a blue clique recruits; it becomes inactive (loss of one blue clique) and k new active white cliques appear. The corresponding Pólya urn scheme has the replacement matrix $\left(\begin{smallmatrix} 0 & k-1 \\ k & -1 \end{smallmatrix}\right)$.

This is an instance of the well-studied Bagchi–Pal urn [3], and we can conduct a probabilistic analysis based on a known urn theory [2], [3] to prove Proposition 3.

4.3. Weight of a random Apollonian network

As discussed, there are τ_n cliques at the nth step. The ℓth clique is determined by the labels in its vertices, and a k-tuple defined as $(i_{1,n}^{(\ell)}, \dots, i_{k,n}^{(\ell)})$. Note that, at any age $n \ge 0$, for any feasible indices $1 \le j \le k$ and $1 \le \ell \le \tau_n$, the entries $i_{j,n}^{(\ell)}$ are random.

Let us denote the weight of the ℓth clique at age n by $\mathcal{W}_{\ell,n}$. So, it is given by $\mathcal{W}_{\ell,n} = \sum_{j=1}^{k} i_{j,n}^{(\ell)}$. The (overall, global, total) weight of the entire Apollonian network after n steps of evolution, Y_n, is defined as the sum of the weights of all τ_n cliques: $Y_n = \sum_{l=1}^{\tau_n} \mathcal{W}_{\ell,n}$. Henceforth, we call it the *total weight*.

Many of the entries in the tuples defining a clique may be shared with the entries of the k-tuples of other cliques, as many vertices are common between cliques. This introduces strong dependence among the weights assigned, rendering the study of the total weight of the network an interesting challenge.

4.3.1. *A stochastic recurrence for the age and its moments.* Suppose that the ℓth clique is selected as a recruiter at the nth step. A new vertex labeled n will be added, and joined to the k vertices of the recruiter, forming $\binom{k}{k-1} = k$ new cliques in all, and the recruiter will be deactivated. We can formulate a stochastic recurrence:

$$Y_n = Y_{n-1} + \sum_{j=1}^{k}\left(n - \sum_{\substack{s=1\\ s\neq j}}^{k} i_{s,n-1}^{(\ell)}\right) - \sum_{r=1}^{k} i_{r,n-1}^{(\ell)} = Y_{n-1} + (k-2)\mathcal{W}_{\ell,n-1} + kn. \quad (1)$$

https://doi.org/10.1017/apr.2015.11 Published online by Cambridge University Press

In what follows, we use the notation $\mathbb{F}_n$ for the σ-field generated by the first n evolutionary steps.

Proposition 4. *The mean value of Y_n is*

$$\mathbb{E}[Y_n] = n((k-1)n+1).$$

Proof. Toward a recurrence for $\mathbb{E}[Y_n]$, we take the conditional expectation of both sides of the stochastic recurrence (1). We obtain

$$\begin{aligned}\mathbb{E}[Y_n \mid \mathbb{F}_{n-1}] &= Y_{n-1} + \frac{k-2}{\tau_{n-1}} \sum_{\ell=1}^{\tau_{n-1}} \mathcal{W}_{\ell,n-1} + kn \\ &= Y_{n-1} + \frac{k-2}{(k-1)(n-1)+1} Y_{n-1} + kn \\ &= \frac{(k-1)n}{(k-1)(n-1)+1} Y_{n-1} + kn.\end{aligned}$$

Taking another expectation of both sides in the last equation, we obtain a recurrence for the mean of Y_n; namely,

$$\mathbb{E}[Y_n] = \frac{(k-1)n}{(k-1)(n-1)+1}\mathbb{E}[Y_{n-1}] + kn.$$

This recurrence is to be solved under the initial condition $\mathbb{E}[Y_0] = 0$, which yields the average as stated. □

We can go forward forming recurrences and solving them for higher moments, but this gets tiresome very quickly, a manifestation of the combinatorial explosion phenomenon. So, for higher moments, we need a shortcut, such as an asymptotic approach for all moments. We shall take that up in the martingale approach below. However, variance computation remains key to all development. We highlight only the important steps of a lengthy calculation. We start by squaring the basic recurrence (1) and obtain the stochastic recurrence

$$Y_n^2 = (Y_{n-1} + (k-2)\mathcal{W}_{\ell,n-1} + kn)^2.$$

Again we compute the double expectation by first conditioning on $\mathbb{F}_{n-1}$ and then taking a second expectation. In the process, $\mathbb{E}[Y_{n-1}]$ appears on the right-hand side; however, we have already developed this in Proposition 4. So, we use that result, and simplify to obtain the following recursive relationship:

$$\begin{aligned}\mathbb{E}[Y_n^2] = \bigg(1 + \frac{2(k-2)}{(k-1)(n-1)+1}\bigg)\mathbb{E}[Y_{n-1}^2] + 2kn(n-1)((k-1)(n-1)+1) \\ + 2k(k-2)n(n-1) + k^2n^2 + (k-2)^2\mathbb{E}[\mathcal{W}^2_{L_{n-1},n-1}],\end{aligned} \tag{2}$$

where $L_{n-1} \in \{1, 2, \ldots, \tau_{n-1}\}$ is the (uniformly) random index of the recruiting clique chosen at the nth step. It is evident that to proceed we need to assess $\mathbb{E}[\mathcal{W}^2_{L_{n-1},n-1}]$. We are able to calculate it by appealing to the squares of the labels in all the cliques, so we develop that first.

Let us sum the squares of the labels in a clique and define $B_n = B_n^{(k)}$ as the sum of all these sums across the τ_n cliques of the Apollonian network after n steps of evolution.

https://doi.org/10.1017/apr.2015.11 Published online by Cambridge University Press

Lemma 1. *We have*

$$\mathbb{E}[B_n] = \frac{k(k-1)n^3 + (k^2-k+1)n^2 + (k-1)n}{2k-1}.$$

Proof. Conditioning on $\mathbb{F}_{n-1}$ and L_{n-1}, we obtain a conditional relationship

$$B_n \mid \mathbb{F}_{n-1}, L_{n-1} = B_{n-1} + (k-2)\sum_{s=1}^{k}(i_{s,n-1}^{(L_{n-1})})^2 + kn^2,$$

where $i_{s,n-1}^{(L_{n-1})}$, for $s = 1, \dots, k$, are the random labels of the vertices of the recruiting (random) L_{n-1}th clique at time $n-1$. Take the expectation of both sides, with respect to L_{n-1}, to obtain

$$\mathbb{E}[B_n \mid \mathbb{F}_{n-1}] = B_{n-1} + \frac{(k-2)}{\tau_{n-1}}\sum_{\ell=1}^{\tau_{n-1}}\sum_{s=1}^{k}(i_{s,n-1}^{(\ell)})^2 + kn^2 = \frac{(k-1)n}{(k-1)(n-1)+1}B_{n-1} + kn^2.$$

Taking the double expectation and solving the recurrence of $\mathbb{E}[B_n]$, we prove the stated result, completing the proof. □

Lemma 2. *We have*

$$\mathbb{E}[\mathcal{W}_{L_{n-1},n}^2] = \frac{(3k^2-3k+1)}{k(2k-1)}n^2 + O(n).$$

Proof. Define $S_n^2 = \sum_{\ell=1}^{\tau_n}\mathcal{W}_{\ell,n}^2$. By the construction of the Apollonian network, we have a recursive relationship for S_n^2, obtained by conditioning on $\mathbb{F}_{n-1}$ and L_{n-1}:

$$\begin{aligned} S_n^2 \mid \mathbb{F}_{n-1}, L_{n-1} &= S_{n-1}^2 - \mathcal{W}_{L_{n-1},n-1}^2 + (k-1)\sum_{s=1}^{k}(i_{s,n-1}^{(L_{n-1})})^2 \\ &\quad + 2(k-2)\sum_{1\le s<t\le k} i_{s,n-1}^{(L_{n-1})} i_{t,n-1}^{(L_{n-1})} + 2(k-1)n\sum_{s=1}^{k} i_{s,n-1}^{(L_{n-1})} + kn^2. \end{aligned}$$

Take the expectation with respect to L_{n-1} to obtain

$$\begin{aligned} \mathbb{E}[S_n^2 \mid \mathbb{F}_{n-1}] &= S_{n-1}^2 - \frac{S_{n-1}^2}{\tau_{n-1}} + \frac{(k-1)B_{n-1}}{\tau_{n-1}} + \frac{2(k-2)}{\tau_{n-1}}\sum_{\ell=1}^{\tau_{n-1}}\sum_{1\le s<t\le k} i_{s,n-1}^{(\ell)} i_{t,n-1}^{(\ell)} \\ &\quad + \frac{2(k-1)nY_{n-1}}{\tau_{n-1}} + kn^2. \end{aligned}$$

Note that

$$S_n^2 - B_{n-1} = \sum_{\ell=1}^{\tau_{n-1}}\Bigg(\sum_{s=1}^{k} i_{s,n-1}^{(\ell)}\Bigg)^2 - \sum_{\ell=1}^{\tau_{n-1}}\sum_{s=1}^{k}(i_{s,n-1}^{(\ell)})^2 = 2\sum_{\ell=1}^{\tau_{n-1}}\sum_{1\le s<t\le k} i_{s,n-1}^{(\ell)} i_{t,n-1}^{(\ell)}.$$

Simplifying the recurrence yields

$$\mathbb{E}[S_n^2 \mid \mathbb{F}_{n-1}] = \Bigg(1 - \frac{1}{\tau_{n-1}} + \frac{k-2}{\tau_{n-1}}\Bigg)S_{n-1}^2 + \frac{B_{n-1}}{\tau_{n-1}} + \frac{2(k-1)nY_{n-1}}{\tau_{n-1}} + kn^2.$$

https://doi.org/10.1017/apr.2015.11 Published online by Cambridge University Press

Take the expectation, and apply Lemma 1 for the expectation of B_{n-1} and Proposition 4 for the mean total weight after $n-1$ steps. We obtain a recurrence for $\mathbb{E}[S_n^2]$, with asymptotic solution

$$\mathbb{E}[S_n^2] = \frac{(3k^2-3k+1)(k-1)}{k(2k-1)}n^3 + O(n^2).$$

Note that L_{n-1} is selected uniformly at each step, giving

$$\mathbb{E}[\mathcal{W}^2_{L_{n-1},n}] = \frac{1}{\tau_{n-1}}\mathbb{E}[S^2_{n-1}] = \frac{(3k^2-3k+1)}{k(2k-1)}n^2 + O(n). \qquad \square$$

Having obtained all the elements, we can now solve (2) by applying Lemma 2, and we obtain

$$\mathbb{E}[Y_n^2] = (k-1)^2 n^4 + \frac{(k-1)(k^4-2k^3+15k^2-14k+4)}{k(k+1)(2k-1)}n^3 + O(n^2).$$

We are now poised to compute the variance of the total weight from its second moment and the square of its mean. We note a remarkable cancellation of the leading terms (of order n^4), leaving behind a relatively small variance.

Proposition 5. *The variance of Y_n, the age of an Apollonian network of index k at time n, is*

$$\operatorname{var}[Y_n] = \frac{(k-2)^2(k-1)^3}{k(k+1)(2k-1)}n^3 + O(n^2).$$

Some corollaries of the relatively small variance are helpful later in deriving the limit law. We have a good sharp concentration.

Corollary 1. *We have*

$$Y_n = (k-1)n^2 + O_{L_1}(n^{3/2}).$$

(By saying a sequence of random variables Y_n is $O_{L_1}(g(n))$, we mean there exist a positive constant A and a positive integer n_0, such that $\mathbb{E}[|Y_n|] \le A|g(n)|$, for all $n \ge n_0$.)

Proof. From the asymptotics of the mean and variance, as given in Propositions 4 and 5, we have

$$\begin{aligned}\mathbb{E}[(Y_n-(k-1)n^2)^2] &= \mathbb{E}[((Y_n-\mathbb{E}[Y_n]) + (\mathbb{E}[Y_n]-(k-1)n^2))^2] \\ &= \operatorname{var}[Y_n] + (\mathbb{E}[Y_n]-(k-1)n^2)^2 \\ &= O(n^3).\end{aligned}$$

So, by Jensen's inequality,

$$\mathbb{E}[|Y_n-(k-1)n^2|] \le \sqrt{\mathbb{E}[(Y_n-(k-1)n^2)^2]} = O(n^{3/2}),$$

and this proves the corollary. $\square$

4.3.2. *The martingale structure of total weight.* The random variable Y_n is not a martingale. Nonetheless, we can squeeze a martingale out of it by a transformation involving rescaling and relocating the center. We define $M_n = \alpha_n Y_n + \beta_n$, and seek suitable deterministic values α_n and β_n that render M_n a martingale. We are able to determine the coefficients α_n and β_n from fundamental martingale properties:

$$\begin{aligned}\mathbb{E}[M_n \mid \mathbb{F}_{n-1}] &= \alpha_n \mathbb{E}[Y_n \mid \mathbb{F}_{n-1}] + \beta_n \\ &= \alpha_n \biggl(\frac{(k-1)n}{(k-1)(n-1)+1} Y_{n-1} + kn \biggr) + \beta_n \\ &= \alpha_{n-1} Y_{n-1} + \beta_{n-1}.\end{aligned}$$

This is possible if we select α_n and β_n to satisfy the recurrences

$$\alpha_n = \frac{(k-1)(n-1)+1}{(k-1)n} \alpha_{n-1}, \qquad \beta_n = \beta_{n-1} - kn\alpha_n.$$

These recurrences have solutions

$$\alpha_n = \frac{\Gamma(n+1/(k-1))}{\Gamma(1/(k-1))\Gamma(n+1)} \alpha_0, \qquad \beta_n = -\frac{(k-1)n(n+1)\Gamma(n+1/(k-1)+1)}{\Gamma(1/(k-1))\Gamma(n+2)} + \beta_0,$$

for arbitrary choices of α_0 and β_0. For simplicity, we take $\alpha_0 = 1$ and $\beta_0 = 0$.

The scaling and shifting coefficients have simple asymptotics (as $n \to \infty$) obtained from Sterling's approximation to the gamma function:

$$\alpha_n \sim \frac{1}{\Gamma(1/(k-1)) n^{(k-2)/(k-1)}}, \qquad \beta_n \sim -\frac{k-1}{\Gamma(1/(k-1))} n^{k/(k-1)}. \tag{3}$$

4.3.3. *Asymptotic Gaussian law.* We would obtain an asymptotic Gaussian law for the total weight of an Apollonian network, if we verify the conditions of martingale central limit theorem for M_n; see [11, pp. 57–59]. These are the combined conditional Lindeberg's condition and the conditional variance condition. These conditions are concisely expressed in terms of the backward difference operator ∇. Acting on a function h of discrete time, this operator denotes the difference of the function h at two successive time points, i.e. $\nabla h_n = h_n - h_{n-1}$.

The conditional Lindeberg's condition requires that, for some positive sequence ξ_n, and for all $\varepsilon > 0$,

$$U_n := \sum_{j=1}^{n} \mathbb{E}\biggl[\biggl(\frac{\nabla M_j}{\xi_n} \biggr)^2 \mathbf{1}_{\{|\nabla M_j/\xi_n| > \varepsilon\}} \biggm| \mathbb{F}_{j-1} \biggr] \xrightarrow{\mathbb{P}} 0,$$

where '$\xrightarrow{\mathbb{P}}$' denotes convergence in probability. The conditional variance condition requires that, for some constant $H \neq 0$, we have

$$V_n := \sum_{j=1}^{n} \mathbb{E}\biggl[\biggl(\frac{\nabla M_j}{\xi_n} \biggr)^2 \biggm| \mathbb{F}_{j-1} \biggr] \xrightarrow{\mathbb{P}} H^2.$$

When these conditions are satisfied, we obtain

$$\frac{M_n}{\xi_n} \xrightarrow{\mathrm{D}} \mathcal{N}(0, H^2),$$

https://doi.org/10.1017/apr.2015.11 Published online by Cambridge University Press

where the right-hand side is a normally distributed random variable, with variance H. It will turn out that in the case of the total weight of an Apollonian network, the correct scale factor ξ_n is $n^{(k+1)/[2(k-1)]}$.

Before we verify the two conditions, we first establish uniform bounds that are instrumental in checking both of them.

Lemma 3. *The terms* $|\nabla M_j|/n^{1/(k-1)}$ *are absolutely uniformly bounded in* $j = 1, \dots, n$.

Proof. By the construction of the martingale, for each $1 \le j \le n$, we have

$$\begin{aligned}|\nabla M_j| &= |M_j - M_{j-1}|\\ &= |(\alpha_j Y_j + \beta_j) - (\alpha_{j-1} Y_{j-1} + \beta_{j-1})|\\ &\le \alpha_{j-1}\left|\frac{\alpha_j}{\alpha_{j-1}} Y_j - Y_{j-1}\right| + kj\alpha_j.\\ &\le \alpha_{j-1}\left|\left(1 - \frac{k-2}{(k-1)j}\right)Y_j - Y_{j-1}\right| + kj\alpha_{j-1}.\end{aligned}$$

The last inequality is valid because α_j decreases as j increases. The asymptotic properties of α_n (cf. (3)) lead us to conclude that $\alpha_{j-1} \le c_k / j^{(k-2)/(k-1)}$, for some positive c_k, and we obtain

$$\begin{aligned}|\nabla M_j| &\le \frac{c_k}{j^{(k-2)/(k-1)}}\left|\left(1 + O\left(\frac{1}{j}\right)\right)Y_j - Y_{j-1}\right| + \frac{c_k kj}{j^{(k-2)/(k-1)}}\\ &\le \frac{c_k}{j^{(k-2)/(k-1)}}\left(|Y_j - Y_{j-1}| + O\left(\frac{\max_{1\le i\le j} Y_i}{j}\right) + kj\right)\\ &\le \frac{c_k j}{j^{(k-2)/(k-1)}}(k + O(1) + k)\\ &\le h_k n^{1/(k-1)}\end{aligned}$$

for some constant h_k that depends on k only, concluding the proof. □

Lemma 4. *We have*

$$U_n = \sum_{j=1}^{n} \mathbb{E}\left[\left(\frac{\nabla M_j}{n^{(k+1)/2(k-1)}}\right)^2 \mathbf{1}_{\{|\nabla M_j / n^{(k+1)/2(k-1)}| > \varepsilon\}} \,\middle|\, \mathbb{F}_{j-1}\right] \xrightarrow{\mathbb{P}} 0.$$

Proof. By the uniform bound established in Lemma 3, for every $\varepsilon > 0$ there exists $n_0(\varepsilon) > 0$, such that for all $n > n_0(\varepsilon)$, the sets $\{|\nabla M_j / n^{(k+1)/[2(k-1)]}| > \varepsilon\}$ are empty, which implies that the sequence U_n converges almost surely to 0. This almost sure convergence is stronger than the required in-probability convergence. □

Lemma 5. *We have*

$$V_n = \sum_{j=1}^{n} \mathbb{E}\left[\left(\frac{\nabla M_j}{n^{(k+1)/[2(k-1)]}}\right)^2 \,\middle|\, \mathbb{F}_{j-1}\right] \xrightarrow{\mathbb{P}} \frac{(k-2)^2(k-1)^3}{k(k+1)(2k-1)\Gamma^2(1/(k-1))}.$$

Proof. Write

$$V_n = \frac{1}{n^{(k+1)/(k-1)}} \sum_{j=1}^{n} \mathbb{E}[(\nabla(\alpha_j Y_j) + \nabla\beta_j)^2] \mid \mathbb{F}_{j-1}]$$
$$= \frac{1}{n^{(k+1)/(k-1)}} \sum_{j=1}^{n} \mathbb{E}[(\nabla(\alpha_j Y_j))^2 + 2(\nabla(\alpha_j Y_j))\nabla\beta_j + (\nabla\beta_j)^2 \mid \mathbb{F}_{j-1}].$$

Consider the summand in three parts.

(i) The first part is

$$\begin{aligned}
&\mathbb{E}[((\nabla(\alpha_j Y_j))^2 \mid \mathbb{F}_{j-1}] \\
&\quad = \mathbb{E}[(\alpha_j Y_j - \alpha_{j-1} Y_{j-1})^2 \mid \mathbb{F}_{j-1}] \\
&\quad = \alpha_j^2 \mathbb{E}[Y_j^2 \mid \mathbb{F}_{j-1}] + \alpha_{j-1}^2 Y_{j-1}^2 - 2\alpha_j \alpha_{j-1} Y_{j-1} \mathbb{E}[Y_j \mid \mathbb{F}_{j-1}] \\
&\quad = \left(\frac{(k-1)(j-1)+1}{(k-1)j}\alpha_{j-1}\right)^2 \\
&\qquad \times \left[\left(1 + \frac{2(k-2)}{(k-1)(j-1)+1}\right)Y_{j-1}^2 + \left(2kj + \frac{2k(k-2)j}{(k-1)(j-1)+1}\right)Y_{j-1}\right. \\
&\qquad\quad \left. + (k-2)^2 W_{\ell,j-1}^2 + k^2 j^2\right] + \alpha_{j-1}^2 Y_{j-1}^2 - \frac{2((k-1)(j-1)+1)}{(k-1)j}\alpha_{j-1}^2 \\
&\qquad \times \left(\frac{(k-1)j}{(k-1)(j-1)+1} Y_{j-1}^2 + kjY_{j-1}\right).
\end{aligned}$$

(ii) The second part is

$$\begin{aligned}
&\mathbb{E}[2(\nabla(\alpha_j Y_j))\nabla\beta_j \mid \mathbb{F}_{j-1}] \\
&\quad = \mathbb{E}[2(\alpha_j Y_j - \alpha_{j-1} Y_{j-1})(\beta_j - \beta_{j-1}) \mid \mathbb{F}_{j-1}] \\
&\quad = -2kj\alpha_j(\alpha_j \mathbb{E}[Y_j \mid \mathbb{F}_{j-1}] - \alpha_{j-1} Y_{j-1}) \\
&\quad = -2kj \frac{(k-1)(j-1)+1}{(k-1)j} \\
&\qquad \times \alpha_{j-1}\left[\frac{(k-1)(j-1)+1}{(k-1)j}\alpha_{j-1}\left(\frac{(k-1)j}{(k-1)(j-1)+1} Y_{j-1} + kj\right) - \alpha_{j-1} Y_{j-1}\right].
\end{aligned}$$

(iii) The third part is

$$\begin{aligned}
\mathbb{E}[(\nabla\beta_j)^2 \mid \mathbb{F}_{j-1}] &= \mathbb{E}[(\beta_j - \beta_{j-1})^2] \\
&= (-kj\alpha_j)^2 \\
&= k^2 j^2 \left(\frac{(k-1)(j-1)+1}{(k-1)j}\alpha_{j-1}\right)^2.
\end{aligned}$$

Combining these three parts, and using the asymptotic equivalents in Corollary 1 for Y_{j-1}^2 and Y_{j-1}, we see that the summand in V_n can be expressed as

$$\frac{(k-2)^2(k-1)^4}{k(2k-1)(k-1)^2\Gamma^2(1/(k-1))} j^{2/(k-1)} + O_{L_1}(j^{-(k-3)/(k-1)}).$$

https://doi.org/10.1017/apr.2015.11 Published online by Cambridge University Press

Summing these terms for $j = 1, 2, \dots, n$ and letting n go to ∞, we obtain

$$\begin{aligned} V_n &= \frac{1}{n^{(k+1)/(k-1)}} \\ &\quad\times \biggl(\frac{(k-2)^2(k-1)^4}{k(2k-1)(k-1)^2\Gamma^2(1/(k-1))((k+1)/(k-1))} n^{2/(k-1)+1} + O_{L_1}(n^{2/(k-1)}) \biggr) \\ &\to \frac{(k-2)^2(k-1)^3}{k(k+1)(2k-1)\Gamma^2(1/(k-1))} \quad \text{in } \mathrm{L}_1. \end{aligned}$$

This L_1 convergence is stronger than the required in-probability convergence.

Having checked the conditions for the martingale central limit theorem, we conclude that

$$\frac{M_n}{n^{(k+1)/2(k-1)}} \xrightarrow{\mathrm{D}} \mathcal{N}\biggl(0, \frac{(k-2)^2(k-1)^3}{k(k+1)(2k-1)\Gamma^2(1/(k-1))}\biggr).$$

Translating this result into Y_n, we obtain a main result of this investigation. Write the latter as

$$\frac{\alpha_n Y_n^{(k)} + \beta_n}{n^{(k+1)/2(k-1)}} \xrightarrow{\mathrm{D}} \mathcal{N}\biggl(0, \frac{(k-2)^2(k-1)^3}{k(k+1)(2k-1)\Gamma^2(1/(k-1))}\biggr).$$

Using the asymptotics in (3) and making a few adjustments via use of only leading terms, we arrive at

$$\begin{aligned} &\biggl(\frac{Y_n^{(k)}}{\Gamma(1/(k-1))n^{(k-2)/(k-1)}} - \frac{(k-1)n^{k/(k-1)}}{\Gamma(1/(k-1))} \biggr) \Bigl(n^{(k+1)/2(k-1)} \Bigr)^{-1} \\ &\quad \xrightarrow{\mathrm{D}} \mathcal{N}\biggl(0, \frac{(k-2)^2(k-1)^3}{k(k+1)(2k-1)\Gamma^2(1/(k-1))}\biggr), \end{aligned}$$

which is equivalent to the convergence in distribution in Theorem 1, completing its proof. □

5. Random *k*-trees

A structure closely related to Apollonian networks is the k-tree. The k-tree is the same as an Apollonian network in every aspect, except that recruiting cliques are not deactivated; all cliques since the beginning remain active. The k-trees are an old model, introduced in the late 1960s [12]. Interest in the structure has been rekindled in recent times [5], [17]. The 1-tree is an interesting research tool, it is the standard recursive tree and there exists an extensive body of literature. We refer the reader to [18] for an extensive survey.

It is clear that the methods applied to Apollonian networks would work for k-trees (of course, mutatis mutandis) and would produce similar types of result. We summarize these results here, without proof. We shall reuse notation, with tildes. For instance, $\tilde{Y}_n$ denotes the total weight of a k-tree of age n.

By an argument similar to that for Apollonian networks, we create a two-color triangular Pólya urn model in order to study the age profile of the jth node. The replacement matrix is $\bigl(\begin{smallmatrix} k-1 & 1 \\ 0 & k \end{smallmatrix}\bigr)$. An analysis following the steps in Section 4.1 gives us, for fixed j, $0 \le j \le n$, that $\tilde{D}_{j,n}^{(k)}/n^{(k-1)/k}$ converges in distribution to a random variable $\tilde{D}_j^*$ with moments

$$\mathbb{E}[(\tilde{D}_j^*)^s] = \Gamma\biggl(\frac{kj+1}{k}\biggr)\Gamma\biggl(\frac{k}{k-1}+s\biggr)\biggl(\Gamma\biggl(\frac{k}{k-1}\biggr)\Gamma\biggl(\frac{kj+1+(k-1)s}{k}\biggr)\biggr)^{-1}.$$

The urn scheme for $\tilde{T}_n^{(k)}$ is Bernard Friedman's urn. The associated central limit theorem (via martingales) appeared in [8]; of course, it is a special case of the Bagchi–Pal urn [3]. The ball replacement matrix for this urn is $\left(\begin{smallmatrix} 0 & k \\ k & 0 \end{smallmatrix}\right)$, furnishing a Gaussian law

$$\frac{\tilde{T}_n^{(k)} - n/2}{\sqrt{n}} \xrightarrow{\mathrm{D}} \mathcal{N}\left(0, \frac{1}{12}\right).$$

Finally, we obtain the asymptotic Gaussian law for $\tilde{Y}_n^{(k)}$,

$$\frac{\tilde{Y}_n^{(k)} - n^2(k^2/(k+1))}{n^{3/2}} \xrightarrow{\mathrm{D}} \mathcal{N}\left(0, \frac{k^4(k-1)^2}{(2k+1)(k+2)(k+1)^2}\right).$$

References

[1] Andrade, J. S., Jr., Herrmann, H. J., Andrade, R. F. S. and da Silva, L. R. (2005). Apollonian networks: simultaneously scale-free, small world, Euclidean, space filling, and with matching graphs. *Phys. Rev. Lett.* **94,** 018702. (Erratum: **102** (2009), 079901.)

[2] Athreya, K. B. and Karlin, S. (1968). Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Statist.* **39,** 1801–1817.

[3] Bagchi, A. and Pal, A. K. (1985). Asymptotic normality in the generalized Pólya–Eggenberger urn model, with an application to computer data structures. *SIAM J. Algebraic Discrete Methods* **6,** 394–405.

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

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

[6] Ebrahimzadeh, E. *et al.* (2013). On the longest paths and the diameter in random Apollonian networks. *Electron. Notes Discrete Math.* **43,** 355–365.

[7] Flajolet, P., Dumas, P. and Puyhaubert, V. (2006). Some exactly solvable models of urn process theory. in *Fourth Colloquium on Mathematics and Computer Science Algorithms, Trees, Conmbinatorics and Probabilities*, Association of Discrete Mathematics and Theoretical Computer Science, Nancy, pp. 59–118.

[8] Freedman, D. A. (1965). Bernard Friedman's urn. *Ann. Math. Statist.* **36,** 956–970.

[9] Frieze, A. and Tsourakakis, C. E. (2012). On certain properties of random Apollonian networks. In *Algorithms and Models for the Web Graph* (Proc. WAW 2012), Springer, Berlin, pp. 93–112.

[10] Graham, R. L., Knuth, D. E. and Patashnik, O. (1994). *Concrete Mathematics*, 2nd edn. Addison-Wesley, Reading, MA.

[11] Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and Its Application*. Academic Press, New York.

[12] Harary, F. and Palmer, E. M. (1968). On acyclic simplical complexes. *Mathematika* **15,** 115–122.

[13] Janson, S. (2005). Asymptotic degree distribution in random recursive trees. *Random Structures Algorithms* **26,** 69–83.

[14] Janson, S. (2006). Limit theorems for triangular urn schemes. *Prob. Theory Relat. Fields* **134,** 417–452.

[15] Janson, S. (2010). Moments of gamma type and the Brownian supremum process area. *Prob. Surveys* **7,** 1–52, 207–208.

[16] Kuba, M. and Panholzer, A. (2007). On the degree distribution of the nodes in increasing trees. *J. Combin. Theory A* **114,** 597–618.

[17] Panholzer, A. and Seitz, G. (2014). Ancestors and descendants in evolving k-tree models. *Random Structures Algorithms* **44,** 465–489.

[18] Smythe, R. T. and Mahmoud, H. M. (1996). A survey of recursive trees. *Theory Prob. Math. Statist.* **51,** 1–27.

[19] Zhang, P., Chen, C. and Mahmoud, H. (2015). Explicit characterization of moments of balanced triangular Pólya urns by an elementary approach. *Statist. Prob. Lett.* **96,** 149–153.