arXiv:1106.1940v1 [cs.SI] 10 Jun 2011

THE DEGREE SEQUENCE OF RANDOM APOLLONIAN NETWORKS

CHARALAMPOS E. TSOURAKAKIS

Abstract. We analyze the asymptotic behavior of the degree sequence of Random Apollonian Networks [11]. For previous weaker results see [10, 11].

1. Introduction

1.1. **Results.** Random Apollonian Networks (RANs) is a popular model of planar graphs with power law properties [11], see also Section 1.3. In this note we analyze the degree sequence of RANs. For earlier, weaker results see [10, 11]. Our main result is Theorem 1.

Theorem 1. *Let $Z_k(t)$ denote the number of vertices of degree k at time t, $k \geq 3$. For any $k \geq 3$ there exists a constant b_k (depending on k) such that for t sufficiently large*

$$|\mathbb{E}\left[Z_k(t)\right] - b_k t| \leq K, \text{ where } K = 3.6.$$

Furthermore let $\lambda > 0$. For any $k \geq 3$

$$\mathbf{Pr}\left[|Z_k(t) - \mathbb{E}\left[Z_k(t)\right]| \geq \lambda\right] \leq 2e^{-\frac{\lambda^2}{72t}}.$$

1.2. **Related Work.** Bollobás, Riordan, Spencer and Tusnády [3] proved rigorously the power law distribution of the Barabási-Albert model [2]. Random Apollonian Networks were introduced in [11]. Their degree sequence was analyzed inaccurately in [11] (see comment in [10]) and subsequently in [10] using physicist's methodology. Cooper & Uehara [4] and Gao [8] analyzed the degree distribution of random k trees, a closely related model to RANs. In RANs –in contrast to random k trees– the random k clique chosen at each step has never previously been selected. For example in the two dimensional case any chosen triangular face is being subdivided into three new triangular faces by connecting the incoming vertex to the vertices of the boundary. Darrasse and Soria analyzed the degree distribution of random Apollonian network structures in [6].

1.3. **Model.** For convenience, we summarize the RAN model here, see also [11]. A RAN is generated by starting with a triangular face and doing the following until the network reaches the desired size: pick a triangular face uniformly at random, insert a vertex inside the sampled face and connect it to the vertices of the boundary.

1.4. **Prerequisites.** In Section 2 we invoke the following lemma.

Lemma 1 (Lemma 3.1, [5]). *Suppose that a sequence $\{a_t\}$ satisfies the recurrence*

$$a_{t+1} = (1 - \frac{b_t}{t+t_1})a_t + c_t$$

Key words and phrases. Degree Distribution, Random Apollonian Networks.
Supported by NSF grant ccf1013110.

for $t \geq t_0$*. Furthermore suppose* $\lim_{t\to+\infty} b_t = b > 0$ *and* $\lim_{t\to+\infty} c_t = c$*. Then* $\lim_{t\to+\infty} \frac{a_t}{t}$ *exists and*

$$\lim_{t\to+\infty} \frac{a_t}{t} = \frac{c}{1+b}.$$

We also use in Section 2 the Azuma-Hoeffding inequality 2, see [1, 9].

Lemma 2 (Azuma-Hoeffding inequality)**.** *Let* $(X_t)_{t=0}^n$ *be a martingale sequence with* $|X_{t+1}-X_t| \leq c$ *for* $t = 0, \ldots, n-1$*. Also, let* $\lambda > 0$*. Then:*

$$\mathbf{Pr}\left[|X_n - X_0| \geq \lambda\right] \leq 2\exp\Big(-\frac{\lambda^2}{2c^2n}\Big)$$

2. Proof of Theorem 1

For simplicity, let $N_k(t) = \mathbb{E}\left[Z_k(t)\right]$, $k \geq 3$. Also, let $d_v(t)$ denote the degree of vertex v at time t and let $\mathbf{1}(d_v(t) = k)$ be an indicator variable which equals 1 if $d_v(t) = k$, otherwise 0. Then, for any $k \geq 3$ we can express the expected number $N_k(t)$ of vertices of degree k as a sum of expectations of indicator variables:

$$N_k(t) = \sum_v \mathbb{E}\left[\mathbf{1}(d_v(t) = k)\right]. \tag{1}$$

We distinguish two cases in the following.

• <u>Case 1 $k = 3$:</u>

Observe that a vertex of degree 3 is created only by an insertion of a new vertex. The expectation $N_3(t)$ satisfies the following recurrence[1]

$$N_3(t+1) = N_3(t) + 1 - \frac{3N_3(t)}{2t+1}. \tag{2}$$

The basis for Recurrence (2) is $N_3(1) = 4$. We prove the following lemma which shows that $\lim_{t\to+\infty} \frac{N_3(t)}{t} = \frac{2}{5}$.

Lemma 3. $N_3(t)$ *satisfies the following inequality:*

$$|N_3(t) - \frac{2}{5}t| \leq K, \text{ where } K = 3.6 \tag{3}$$

Proof. We use induction. Assume that $N_3(t) = \frac{2}{5}t + e_3(t)$. We wish to prove that for all t, $|e_3(t)| \leq 3.6$. The result trivially holds for $t = 1$. Assume the result holds for some t. We shall show it holds for $t + 1$.

[1]The three initial vertices participate in one less face than their degree. However, this leaves the asymptotic analysis unchanged.

$$N_3(t+1) = N_3(t) + 1 - \frac{3N_3(t)}{2t+1} \Rightarrow$$
$$e_3(t+1) = e_3(t) + \frac{3}{5} - \frac{6t + 15e_3(t)}{10t+5} = e_3(t) + \frac{3}{5(2t+1)} - \frac{3e_3(t)}{2t+1} \Rightarrow$$
$$|e_3(t+1)| \leq K(1 - \frac{3}{2t+1}) + \frac{3}{5(2t+1)} \leq K$$

Hence by induction, Inequality (3) holds for all $t \geq 1$.

□

• Case 2 $k \geq 4$:

For $k \geq 4$ the following Equation holds for each indicator variable $\mathbf{1}(d_v(t) = k)$:

$$\mathbb{E}\left[\mathbf{1}(d_v(t+1) = k)\right] = \mathbb{E}\left[\mathbf{1}(d_v(t) = k)\right](1 - \frac{k}{2t+1}) + \mathbb{E}\left[\mathbf{1}(d_v(t) = k-1)\right]\frac{k-1}{2t+1} \tag{4}$$

Therefore, substituting in Equation (1) the expression from Equation 4 we obtain for $k \geq 4$

$$N_k(t+1) = N_k(t)(1 - \frac{k}{2t+1}) + N_{k-1}(t)\frac{k-1}{2t+1}. \tag{5}$$

Now, we use induction to show that $\lim_{t\to+\infty} \frac{N_k(t)}{t}$ for $k \geq 4$ exists.

Lemma 4. *For $k \geq 3$ $\lim_{t\to+\infty} \frac{N_k(t)}{t}$ exists. Let $b_k = \lim_{t\to+\infty} \frac{N_k(t)}{t}$. Then, $b_3 = \frac{2}{5}, b_4 = \frac{1}{5}, b_5 = \frac{4}{35}$ and for $k \geq 6$ $b_k = \frac{24}{k(k+1)(k+2)}$. Furthermore, for all $k \geq 3$*

$$|N_k(t) - b_k t| \leq K, \text{ where } K = 3.6. \tag{6}$$

Proof. For $k = 3$ the result holds by Lemma 3. We use induction. Rewrite Recurrence (5) as: $N_k(t+1) = (1 - \frac{b_t}{t+t_1})N_k(t) + c_t$ where $b_t = k/2$, $t_1 = 1/2$, $c_t = N_{k-1}(t)\frac{k-1}{2t+1}$. Clearly, $\lim_{t\to+\infty} b_t = k/2 > 0$ and by the inductive hypothesis $\lim_{t\to+\infty} c_t = \lim_{t\to+\infty} b_{k-1}t\frac{k-1}{2t+1} = b_{k-1}(k-1)/2$. Hence, by invoking Lemma 1 we obtain

$$b_k = \lim_{t\to+\infty} \frac{N_k(t)}{t} = \frac{(k-1)b_{k-1}/2}{1 + k/2} = b_{k-1}\frac{k-1}{k+2}.$$

Therefore $b_3 = \frac{2}{5}$, $b_4 = \frac{1}{5}$, $b_5 = \frac{4}{35}$ for any $k \geq 6$, $b_k = \frac{24}{k(k+1)(k+2)}$ which shows a power law degree distribution with exponent 3.

Finally consider the proof of Inequality (6). The case $k = 3$ was proved in the Lemma 3. Assume the result holds for some $k \geq 3$, i.e., $|e_k(t)| \leq K$ where $K = 3.6$. We will show it holds for $k+1$ too. Let $e_k(t) = N_k(t) - b_k t$. Substituting in Recurrence (5) and using the fact that $b_{k-1}(k-1) = b_k(k+2)$ we obtain the following:

$$e_k(t+1) = e_k(t) + \frac{k-1}{2t+1}e_{k-1}(t) - \frac{k}{2t+1}e_k(t) \Rightarrow$$
$$|e_k(t+1)| \le |(1-\frac{k}{2t+1})e_k(t)| + |\frac{k-1}{2t+1}e_{k-1}(t)| \le K(1-\frac{1}{2t+1}) \le K$$

Hence by induction, Inequality (6) holds for all $k \ge 3$. □

It's worth pointing out that Lemma 4 agrees with [7] where it was shown that the maximum degree is $\Theta(\sqrt{t})$. Finally, the Lemma 5 proves the concentration of $Z_k(t)$ around its expected value for $k \ge 3$. This lemma applies the Azuma-Hoeffding inequality 2 and completes the proof of Theorem 1.

Lemma 5. *Let $\lambda > 0$. For any $k \ge 3$*

$$\mathbf{Pr}\left[|Z_k(t) - \mathbb{E}\left[Z_k(t)\right]| \ge \lambda\right] \le 2e^{-\frac{\lambda^2}{72t}}. \tag{7}$$

Proof. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space induced by the construction of a Random Apollonian Network (see Section 1.3) after t insertions. Fix k ($k \ge 3$) and let $(X_i)_{i\in\{0,1,\ldots,t\}}$ be the martingale sequence defined by $X_i = \mathbb{E}\left[Z_k(t)|\mathcal{F}_i\right]$, where $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_i$ is the σ-algebra generated by the RAN process after i steps. Notice $X_0 = \mathbb{E}\left[Z_k(t)|\{\emptyset,\Omega\}\right] = N_k(t)$, $X_t = Z_k(t)$. We show that $|X_{i+1} - X_i| \le 6$ for $i = 0,..,t-1$. Let $P_j = (Y_1,\ldots,Y_{j-1},Y_j)$, $P'_j = (Y_1,\ldots,Y_{j-1},Y'_j)$ be two sequences of face choices differing only at time j. Also, let $\bar{P}, \bar{P}'$ continue from P_j, P'_j until t. We call the faces Y_j, Y'_j special with respect to $\bar{P}, \bar{P}'$. We define a measure preserving map $\bar{P} \mapsto \bar{P}'$ in the following way: for every choice of a non-special face in process $\bar{P}$ at time l we make the same face choice in $\bar{P}'$ at time l. For every choice of a face inside the special face Y_j in process $\bar{P}$ we make an isomorphic (w.r.t., e.g., clockwise order and depth) choice of a face inside the special face Y'_j in process $\bar{P}'$. Since the number of vertices of degree k can change by at most 6, i.e., the (at most) 6 vertices involved in the two faces Y_j, Y'_j the following holds:

$$|\mathbb{E}\left[Z_k(t)|P\right] - \mathbb{E}\left[Z_k(t)|P'\right]| \le 6.$$

Furthermore, this holds for any P_j, P'_j. We deduce that X_{i-1} is a weighted mean of values, whose pairwise differences are all at most 6. Thus, the distance of the mean X_{i-1} is at most 6 from each of these values. Hence, for any one step refinement $|X_{i+1} - X_i| \le 6\ \forall i \in \{0,\ldots,t-1\}$. By applying the Azuma-Hoeffding inequality as stated in Lemma 2 we obtain

$$\mathbf{Pr}\left[|Z_k(t) - \mathbb{E}\left[Z_k(t)\right]| \ge \lambda\right] \le 2e^{-\frac{\lambda^2}{72t}}. \tag{8}$$

□

A corollary immediately obtained by the previous lemma by setting $\lambda = \sqrt{t\log t}$ is the following:

Corollary 1. $\mathbf{Pr}\left[|Z_k(t) - \mathbb{E}\left[Z_k(t)\right]| \ge \sqrt{t\log t}\right] = o(1)$.

Acknowledgements

The author would like to thank Deepak Bal and Alan Frieze for helpful comments on the manuscript.

References

[1] Azuma, K.: *Weighted sums of certain dependent variables* Tohoku Math J 3, pp. 357-367, 1967

[2] Barabási, A., Albert, R.: *Emergence of Scaling in Random Networks* Science 286, pp. 509-512, 1999

[3] Bollobás, B., Riordan, O., Spencer, J., Tusnády, G.: *The Degree Sequence of a Scale Free Random Graph Process* Random Struct. Algorithms, 18(3), pp. 279-290, 2001

[4] Cooper, C., Uehara, R.: *Scale Free Properties of random k-trees* Mathematics in Computer Science, 3(4), pp. 489–496, 2010

[5] Chung Graham, F., Lu, L.: *Complex Graphs and Networks* American Mathematical Society, No. 107 (2006)

[6] Darrasse, A., Soria, M.: *Degree distribution of random Apollonian network structures and Boltzmann sampling* 2007 Conference on Analysis of Algorithms, AofA 07, DMTCS Proceedings.

[7] Frieze, A., Tsourakakis, C.E.: *High Degree Vertices, Eigenvalues and Diameter of Random Apollonian Networks* Available at Arxiv `http://arxiv.org/abs/1104.5259`

[8] Gao, Y.: *The degree distribution of random k-trees* Theoretical Computer Science, 410(8-10), 2009.

[9] Hoeffding, W.: *Probability inequalities for sumes of bounded random variables* J Amer Statist Assoc 58, pp. 13-30, 1963

[10] Wu, Z.-X., Xu, X.-J., Wang, Y.-H.: *Comment on "Maximal planar networks with large clustering coefficient and power-law degree distribution"* Physical Review, E 73, 058101 (2006)

[11] Zhou, T., Yan, G., Wang, B.H: *Maximal planar networks with large clustering coefficient and power-law degree distribution* Physical Review, E 71, 046141 (2005)

Department of Mathematical Sciences, Carnegie Mellon University, 5000 Forbes Av., 15213, Pittsburgh, PA, U.S.A

E-mail address: `ctsourak@math.cmu.edu`