DEGREES AND DISTANCES IN RANDOM AND EVOLVING APOLLONIAN NETWORKS

ISTVÁN KOLOSSVÁRY[1,3], JÚLIA KOMJÁTHY[2,4], AND LAJOS VÁGÓ[1,3]

ABSTRACT. This paper studies Random and Evolving Apollonian networks (RANs and EANs), in d dimension for any $d \geq 2$, i.e. dynamically evolving random d dimensional simplices looked as graphs inside an initial d-dimensional simplex. We determine the limiting degree distribution in RANs and show that it follows a power law tail with exponent $\tau = (2d-1)/(d-1)$. We further show that the degree distribution in EANs converges to the same degree distribution if the simplex-occupation parameter in the nth step of the dynamics is $q_n \to 0$ and $\sum_{n=0}^{\infty} q_n = \infty$. This result gives a rigorous proof for the conjecture of Zhang et al [31] that EANs tend to show similar behavior as RANs once the occupation parameter $q \to 0$. We also determine the asymptotic behavior of shortest paths in RANs and EANs for arbitrary d dimensions. For RANs we show that the shortest path between two uniformly chosen vertices (typical distance), the flooding time of a uniformly picked vertex and the diameter of the graph after n steps all scale as constant times $\log n$. We determine the constants for all three cases and prove a central limit theorem for the typical distances. We prove a similar CLT for typical distances in EANs.

1. INTRODUCTION

The construction of deterministic and random Apollonian networks originates from the problem of Apollonian circle packing: starting with three mutually tangent circles, we inscribe in the interstice formed by the three initial circles the unique circle that is tangent to all of them: this fourth circle is known as the inner Soddy-circle. Iteratively, for each new interstice its inner Soddy-circle is drawn. After infinite steps the result is an Apollonian gasket [13, 23].

An Apollonian network (AN) is the resulting graph if we place a node in the center of each circle and connect two nodes if and only if the corresponding circles are tangent. This model was introduced independently by Andrade et al. [2] and Doye and Massen [19] as a model for networks arising in real-life such as the network of internet cables or links, collaboration network or protein interaction networks. Apollonian networks serve a good model for these networks since they share many similar features that most of them have: a power-law degree distribution, a high clustering coefficient and small distances, usually referred to as the small-world property. Moreover, by construction, Apollonian networks also show high hierarchical structure: a property that is very common in e.g. social networks.

It is straightforward to generalize Apollonian packings to arbitrary d dimensions with mutually tangent d dimensional hyperspheres. Analogously, if each d-hypersphere corresponds to a vertex and vertices are connected by an edge if the corresponding d-hyperspheres are tangent, then we obtain a d-dimensional AN (see [28, 29]).

The network arising by this construction is deterministic. Zhou et al. [32] proposed to randomise the dynamics of the model such that in one step only one interstice is picked uniformly at random and filled with a new circle. This construction in d dimensions yields a d dimensional *random Apollonian network* (RAN) [30]. Using heuristic and rigorous arguments the results in [1, 17, 18, 20, 22, 30, 32] show that RANs have the above mentioned main features of real-life networks.

A different random version of the original Apollonian network was introduced by Zhang et al. [31], called *Evolutionary Apollonian networks* (EAN) where in every step *every* interstice is picked and filled independently of each other with probability q. If an interstice is not filled in a given step, it can be filled in the next step again. We call q the *occupation parameter*. For $q = 1$ we get back the deterministic AN model. It is conjectured in [31] that if $q \to 0$ then EANs show similar features to RANs. To answer this issue, we investigate how the network looks like when $q := q_n$ depends on the time and tends to 0 as the size of the network grows. In this setting, the interesting question is to determine the correct rate for q_n that achieves that EAN shows similar behavior as RAN.

Our contribution. In this paper we give a rigorous proof for the power law degree distribution of RANs in any dimension d. Further, we show that if we let the occupation parameter of EAN to depend on n and choose such that $q_n \to 0$ and $\sum_{n=0}^{\infty} q_n = \infty$, then the limiting degree distribution of EANs is a power law with the same exponent as for RANs. We also determine the clustering coefficient of RANs and EANs. Further, we investigate the asymptotic behavior of shortest paths in random Apollonian networks in arbitrary d dimensions with rigourous methods. We show that the shortest path between two uniformly chosen vertices (typical distance), the flooding time of a uniformly picked vertex and the diameter of the graph after n steps all scale as constant times $\log n$. We determine the constants for all three cases and prove a central limit theorem for the typical distances. We introduce a different approach to describe the structure of the graph that enables us to unfold the hierarchical structure of RANs and EANs and is based on coding used in fractal-related methods.

Now we give the precise definition of the two models.

Random Apollonian networks. A random Apollonian network $\mathrm{RAN}_d(n)$ in d dimensions can be constructed as follows. Initially at step $n = 0$ start from the d-dimensional simplex with an additional vertex in the interior connected to all of the vertices of the simplex. Thus there are initially $d + 1$ d-simplices, that we call *active cliques*. For $n \geq 1$, pick an active clique C_n of $\mathrm{RAN}_d(n - 1)$ *uniformly at random*, insert a node v_n in the interior of C_n and connect v_n with all the vertices of C_n. The newly added vertex v_n forms new cliques with each possible choice of d vertices of C_n. C_n becomes inactive and these newly formed d-simplices become active. The resulting graph is $\mathrm{RAN}_d(n)$. At each step n a $\mathrm{RAN}_d(n)$ has $n + d + 2$ vertices and $nd + d + 1$ active cliques.

There is a natural representation of RANs as evolving triangulations in two dimensions: take a planar embedding of the complete graph on four vertices as in Figure 1 and in each step pick a face of the graph uniformly at random, insert a node and connect it with the vertices of the chosen triangle. The result is a maximal planar graph. Hence, a $(\mathrm{RAN}_2(n))_{n \in \mathbb{N}}$ is equivalent to an increasing family of triangulations by successive addition of faces in the plane, called *stack-triangulations*. Stack-triangultions were investigated in [1] where the authors also considered typical properties under different weighted measures, (e.g. uniformly picked ones having n faces). Under a certain measure stack-triangulations with n faces are an equivalent formulation of $\mathrm{RAN}_2(n)$, see [1] and references therein.
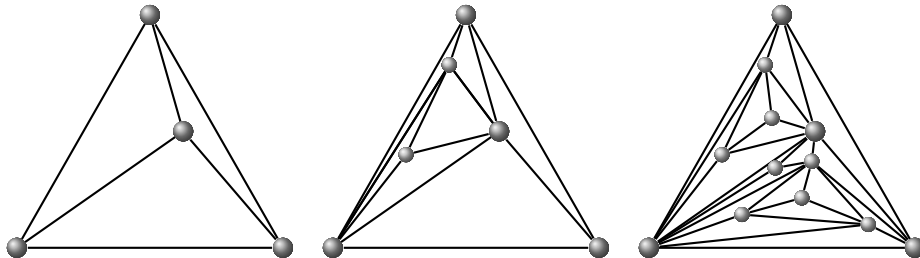
FIGURE 1. A $\mathrm{RAN}_2(n)$ after $n = 0, 2, 8$ steps

Evolutionary Apollonian networks. Given a sequence of occupation parameters $\{q_n\}_{n=1}^{\infty}, 0 \leq q_n \leq 1$, an evolutionary Apollonian network $\mathrm{EAN}_d(n, \{q_n\}) = \mathrm{EAN}_d(n)$ in d dimensions can be constructed iteratively as follows. The initial graph is the same as for a $\mathrm{RAN}_d(0)$ and there are $d+1$ active d-simplices. For $n \geq 1$, pick each active clique of $\mathrm{EAN}_d(n-1)$ independently

of each other with probability q_n. The set of chosen cliques $\mathcal{C}_n$ becomes inactive (the non-picked active cliques stay active) and for every clique $C \in \mathcal{C}_n$ we place a new node $v_n(C)$ in the interior of C that we connect to all vertices of C. This new node $v_n(C)$ together with all possible choices of d vertices from C forms $d+1$ new cliques: these cliques are added to the set of active cliques for every $C \in \mathcal{C}_n$. The resulting graph is $\mathrm{EAN}_d(n)$. The case $q_n \equiv q$ was studied in [31] where it was further suggested that for $q \to 0$ the graph is similar to a $\mathrm{RAN}_d(n)$. We prove their conjecture by showing that EANs obey the same power law exponent as RANs if $q_n \to 0$ and $\sum_{n=0}^{\infty} q_n = \infty$.

Structure of the paper. In Section 2 we state our main results and discuss their relation to other results in the area. Section 3 contains the most important observations about the structure of RANs: we work out an approach of coding the vertices of the graph that enables us to compare the structure of the RAN to a branching process and further, the distance between any two vertices in the graph is given entirely by the coding of these vertices. We also give a short sketch of proofs related to distances in this section. Then we prove rigorously the distance-related theorems in Section 4. Finally in Section 5 we prove the results concerning the degree distributions.

2. MAIN RESULTS

In this section we state our main results and compare it to related literature. First we start with theorems about distances in RANs and EANs, then state our theorems about the limiting degree distributions and clustering coefficient in RANs and EANs, respectively, and finally we compare our results to related literature.

2.1. **Distances in** RAN**s and** EAN**s.** In this subsection we state our results about distances in RANs and EANs. First we define the three main quantities of our investigation: fix n and pick two vertices u and v uniformly at random from $\mathrm{RAN}_d(n)$. Denote by $\mathrm{Hop}_d(n, u, v)$ the hopcount between the vertices u and v, i.e. the number of edges on (one of) the shortest paths between u and v. The flooding time $\mathrm{Flood}_d(n, u)$ is the maximal hopcount from u, while the diameter $\mathrm{Diam}_d(n)$ is the maximal flooding time, formally

$$\mathrm{Flood}_d(n, u) = \max_v \mathrm{Hop}_d(n, u, v) \quad \text{and} \quad \mathrm{Diam}_d(n) = \max_{u,v} \mathrm{Hop}_d(n, u, v).$$

Whenever possiblem d, u and v are suppressed from the notation. We introduce some important notation first. Let

$$Y_d := \sum_{i=1}^{d+1} X_i \tag{2.1}$$

where $(X_i)_{i=1}^{d+1}$ is a collection of independent geometrically distributed random variables with success probability $\frac{i}{d+1}$. We further introduce

$$\mu_d := \mathbb{E}[Y_d] = (d+1)H(d+1), \quad \sigma_d^2 := \mathbb{D}^2[Y_d], \tag{2.2}$$

where $H(d) = \sum_{i=1}^{d} 1/i$. The Large Deviation rate function of Y_d is given by

$$I_d(x) := \sup_{\lambda \in \mathbb{R}} \left\{ \lambda x - \log \left(\mathbb{E}\left[e^{\lambda Y_d} \right] \right) \right\}. \tag{2.3}$$

The next theorem describes the asymptotic behavior of the typical distances in $\mathrm{RAN}_d(n)$.

Theorem 2.1 (Typical distances in RANs)**.** *The hopcount between the vertices corresponding to two uniformly chosen active cliques in a* $\mathrm{RAN}_d(n)$ *satisfies a central limit theorem (CLT) of the form*

$$\frac{\mathrm{Hop}_d(n) - \frac{2}{\mu_d} \frac{d+1}{d} \log n}{\sqrt{2 \frac{\sigma_d^2 + \mu_d}{\mu_d^3} \frac{d+1}{d} \log n}} \xrightarrow{d} Z, \tag{2.4}$$

where μ_d, σ_d^2 as in (2.2) and Z is a standard normal random variable.

Further, the same CLT is satisfied for the distance between two vertices that are picked independently with the size-biased probabilities given by

$$\mathbb{P}(v \text{ is picked} \mid D_v(n) = k) = \frac{(d-1)k - d^2 + d + 2}{dn + d + 1}. \tag{2.5}$$

The next theorem describes the asymptotic behaviour of the flooding time and the diameter:

Theorem 2.2 (Diameter and flooding time in RANs)**.** *Define $\tilde{c}_d$ as the unique solution with $\tilde{c}_d > \frac{d+1}{d}$ to the equation*

$$f_d(c) := c - \frac{d+1}{d} - c \log\left(\frac{d}{d+1}c\right) = -1.$$

Then as $n \to \infty$ with high probability

$$\frac{\mathrm{Diam}_d(n)}{\log n} \xrightarrow{\mathbf{P}} 2\tilde{\alpha}\tilde{\beta}\frac{\tilde{c}_d}{\mu_d},$$
$$\frac{\mathrm{Flood}_d(n)}{\log n} \xrightarrow{\mathbf{P}} \frac{1}{\mu_d}\left(\frac{d+1}{d} + \tilde{\alpha}\tilde{\beta}\tilde{c}_d\right). \tag{2.6}$$

where $(\tilde{\alpha}, \tilde{\beta}) \in (0,1] \times [1, \frac{\mu_d}{d+1}]$ is the solution of the maximization problem with the following constraint:

$$\max\{\alpha\beta : 1 + f(\alpha\tilde{c}_d) - \alpha\beta\frac{\tilde{c}_d}{\mu_d}I_d\left(\frac{\mu_d}{\beta}\right) = 0\}, \tag{2.7}$$

where μ_d, σ_d^2 as in (2.2) and $I_d(x)$ as in (2.3).

Remark 2.3. Observe that the set of (α, β) pairs that satisfy (2.7) is non-empty since for $\alpha = \beta = 1$ by definition $f(\tilde{c}_d) = -1$ and $I_d(\mu_d) = 0$.

Finally, the next theorem describes the asymptotic behavior of the typical distances in $\mathrm{EAN}_d(n)$.

Theorem 2.4 (Typical distances in EANs)**.** *Suppose the sequence of occupation parameters $\{q_n\}$ satisfies $\sum_{n\in\mathbb{N}} q_n = \infty$ and $\sum_{n\in\mathbb{N}} q_n(1 - q_n) = \infty$. Then the hopcount between vertices corresponding to two uniformly chosen active cliques in a $\mathrm{EAN}_d(n)$ satisfies a central limit theorem of the form*

$$\frac{\mathrm{Hop}_d(n) - \frac{2}{\mu_d}\sum_{i=1}^n q_i}{\sqrt{2\frac{\sigma_d^2 + \mu_d}{\mu_d^3}\sum_{i=1}^n q_i(1 - q_i)}} \xrightarrow{d} Z, \tag{2.8}$$

where μ_d, σ_d^2 as in (2.2) and Z is a standard normal random variable.

Further, the same CLT is satisfied for the distance between two vertices that are picked independently with the size-biased probabilities given by

$$\mathbb{P}(v \text{ is picked } |D_v(n) = k||V(n)|) = \frac{(d-1)k - d^2 + d + 2}{d|V(n)| + d + 1}. \tag{2.9}$$

Remark 2.5. Note that in this theorem q_n might or might not tend to 0. The second criterion rules out the case that the $q_n \to 1$ and so the graph becomes essentially deterministic.

2.2. **Degree distribution and clustering coefficient.** First we describe the limiting distribution of RAN_d. Let us denote by $\widetilde{N}_k(n)$ and $\widetilde{p}_k(n)$ the number and the empirical proportion of vertices with degree k at time n respectively, i.e.

$$\widetilde{p}_k(n) := \frac{\widetilde{N}_k(n)}{n+d+2} := \frac{1}{n+d+2}\sum_{i=1}^{n+d+2}\mathbb{1}\{D_i(n) = k\},$$

where $D_i(n)$ stands for the degree of vertex i after the n-th step. Our first theorem describes that this empirical distribution tends to a proper distribution in the ℓ_∞-metric:

Theorem 2.6 (Degree distribution for RANs)**.** *There exists a probability distribution $\{p_k\}_{k=d+1}^\infty$ and a constant c for which*

$$\mathbb{P}\left(\max_k |\widetilde{p}_k(n) - p_k| \geq c\sqrt{\frac{\log n}{n}}\right) = o(1).$$

Further, p_k follows a power law with exponent $(2d-1)/(d-1) \in (2,3]$, more precisely

$$p_k = \frac{d}{2d+1}\frac{\Gamma(k - d + \frac{2}{d-1})}{\Gamma(1 + \frac{2}{d-1})}\frac{\Gamma(2 + \frac{d+2}{d-1})}{\Gamma(k+1-d+\frac{d+2}{d-1})} = k^{-\frac{2d-1}{d-1}}(1 + o(1)), \tag{2.10}$$

where $\Gamma(x)$ is the Gamma function and we used the property that $\Gamma(t+a)/\Gamma(t) = t^a(1+o(1))$.

Our next theorem describes the degree distribution of the graph $\mathrm{EAN}_d(n, \{q_n\})$. Let us denote the set of vertices after n steps by $V(n)$, and write $N_k(n)$ and $p_k(n)$ for the number and the empirical proportion of vertices with degree k at time n, respectively, i.e.

$$p_k(n) := \frac{N_k(n)}{|V(n)|} = \frac{1}{|V(n)|} \sum_{i \in V(n)} \mathbb{1}\{D_i(n) = k\}. \tag{2.11}$$

Theorem 2.7 (Degree distribution for EANs). *Let $d \geq 2$ and $\{q_n\}_{n=0}^\infty$ be probabilities such that*

$$q_n \to 0, \quad \sum_{n=0}^\infty q_n = \infty.$$

Then the degree distribution is tending to the same asymptotic degree distribution $\{p_k\}_{k=d+1}^\infty$ as in the case of RAN_d (2.10), and

$$\mathbb{P}\left(\max_k |p_k(n) - p_k| \geq c\sqrt{\frac{\log |V(n)|}{|V(n)|}}\right) = o(1)$$

for any $c > \sqrt{8}(d+1)^{3/2}$.

The proof of these theorems are given in Sections 5.1 and 5.2. Next we describe the clustering coefficient of RANs and EANs. The clustering coefficient of a node is just the portion of the number of existing edges between its neighbors, compared to the number of all possible edges between those. Here we investigate the clustering coefficient of the whole graph, which is the average of clustering coefficients over the nodes. Since these are direct consequences of the formula for the limiting degree distributions p_k, we state them as corollaries. This corollary is similar to the result in [30, Section 4.2.], but now, that we have established the degree distribution, it has a rigorous proof. This is based on the observation that the clustering coefficient of a vertex with degree k is *deterministic* and equals

$$\frac{d(2k-d-1)}{k(k-1)} \sim \frac{2d}{k}.$$

The reason for this formula is the following: when the degree of a node v increases by one by adding a new vertex w in one of the active cliques v is contained in, then the number of edges between the neighbors of v increases exactly by d, since the newly added vertex w is connected to the d other vertices in the clique. It was observed in simulations and heuristicly proved in [30], that the average clustering coefficient of these networks are converging to a strictly positive constant. Our next corollary determines the exact value of these constants for the two models:

Corollary 2.8 (Clustering coefficient). *The average clustering coefficient of $\mathrm{RAN}_d(n)$ converges to a strictly positive constant as $n \to \infty$, given by*

$$
\begin{aligned}
Cl_d &= \sum_{k=d+1}^\infty \frac{d(2k-d-1)}{k(k-1)} p_k \\
&= \sum_{k=d+1}^\infty \frac{d(2k-d-1)}{k(k-1)} \cdot \frac{d}{2d+1} \frac{\Gamma(k-d+\frac{2}{d-1})}{\Gamma(1+\frac{2}{d-1})} \frac{\Gamma(1+\frac{2d+1}{d-1})}{\Gamma(k-d+\frac{2d+1}{d-1})}.
\end{aligned} \tag{2.12}
$$

Further, the clustering coefficient of $\mathrm{EAN}_d(n, \{q_n\})$ converges to the same value as in (2.12) if $q_n \to 0$ and $\sum_{n \in \mathbb{N}} q_n = \infty$.

2.3. **Related literature.** Several results connected to the asymptotic degree distribution of Apollonian networks are known. It is not hard to see that if a vertex belongs to k active cliques, then the chance of getting a new edge is proportional to k: this argument shows that these models belong to the wide class of *Preferential attachment models* [4, 11, 12]. As a result, some of the classical methods can be adapted to this model.

Using heuristic arguments, Zhang and co-authors [30] obtained that the asymptotic degree exponent should be $\frac{2d-1}{d-1} \in (2, 3]$, which is in good agreement with their simulations. Parallel to writing this paper, we noticed that Frieze and Tsourakakis [22] very recently derived rigorously the exact asymptotic degree distribution of the two-dimensional RAN. Even

though our work is parallel to theirs, the methods are similar: this is coming from the fact that both of the methods used there and in our work are an adaptation of standard methods given in [12, 24]. So, to avoid repetition we decided to only sketch some parts of the proof and include the part that does not follow from their work.

What is entirely new in our paper is that we study the EAN model rigorously. For the degree distribution of EANs only heuristic arguments were known before. Zhang, Rong and Zhou [31] studied the graph series $\mathrm{EAN}_d(n, \{q_n \equiv q\})$. They derived the asymptotic degree exponent using heuristic arguments, and the result fits well with the simulations. They also suggested that as $q \to 0$ the model $\mathrm{EAN}_d(n, \{q_n \equiv q\})$ converges to $\mathrm{RAN}_d(n)$ in some sense. We confirm their claim by deriving the asymptotic degree distribution of $\mathrm{EAN}_d(n, \{q_n\})$ with $\{q_n\}$ such that $q_n \to 0$ and $\sum_{n=0}^{\infty} q_n = \infty$, obtaining the same degree distribution. So the idea of Zhang and co-authors can be made precise in this way.

The statements of Theorem 2.1 are in agreement with previous results. In particular, in [30] the authors estimate the average path length, i.e. the hopcount averaged over all pairs of vertices, and they show that it scales logarithmically with the size of the network.

A refined, but still weaker claim is obtained by Albenque and Marckert [1] concerning the hopcount in two dimensions. They prove that

$$\frac{\mathrm{Hop}(n)}{6/11 \log n} \xrightarrow{\mathbf{P}} 1.$$

The constant $6/11$ is the same as $2(d+1)/(d\mu_d)$ for $d = 2$. They use the previously mentioned notion of stack triangulations to derive the result from a CLT similar to the one in Theorem 2.1, we show an alternative approach using weaker results. The CLT for distances in EANs is new.

Central limit theorems of the form (2.4) for the hopcount have been proven with the addition of exponential or general edge weights for various other random graph models, known usually under the name *first passage percolation*. Janson [25] analysed distances in the complete graphs with i.i.d. exponential edge weights. In a series of papers Bhamidi, van der Hofstad and Hooghiemstra determine typical distances and prove CLT for the hopcount for the Erdős-Rényi random graph [6], the stochastic mean-field model [8], the configuration model with finite variance degrees [5] and quite recently for the configuration model [7] with arbitrary independent and identically distributed edge weights. Inhomogeneous random graphs are handled by Bollobás, Janson and Riordan [10, 27]. Note that in all these models the edges have random weights, while in RANs and EANs all edge weights are 1. The reason for this similarity is hidden in the fact that all these models have an underlying branching process approximation, and the CLT valid for the branching process implies CLT for the hopcount on the graph. The diameter and flooding time of EANs remains a future project.

Further, there are some previous bounds known about the diameter of RANs: Frieze and Tsourakakis [22] establishes the upper bound $2\tilde{c}_2 \log n$ for $\mathrm{RAN}_2(n)$. They use a result of Broutin and Devroye [14] that, combined with the branching process approximation of the structure of RANs we describe in this paper, actually implicitly gives the $2\tilde{c}_d \log n$ upper bound for all d.

Just recently and independently from our work other methods were used to determine the diameter. In [20] Ebrahimzadeh et al. apply the result of [14] in an elaborate way, while Cooper and Frieze in [17] use a more analytical approach solving recurrence relations. We emphasize that the methods in [17, 20] and in the present paper are all qualitatively different. Numerical solution of the maximization problem (2.7) for $d = 2$ yields the optimal $(\tilde{\alpha}, \tilde{\beta})$ pair to be approximately $(0.8639, 1.500)$. The corresponding constant for the diameter is 1.668, which perfectly coincides with the one obtained in [17] and [20]. To the best of our knowledge no result has been proven for the flooding time.

3. STRUCTURE OF RANS AND EANS

3.1. **Tree-structure of RANs and EANs.** The construction method of RANs and EANs enables us to describe a natural way to code the vertices and active cliques of the graph parallel to each other. Let $\Sigma_d := \{1, 2, \ldots, d+1\}$ be the symbols of the alphabet. We give the initial vertices of a $\mathrm{RAN}_d(0)$ the empty word except for the vertex in the middle of the initial d-simplex which gets the symbol $\mathbf{O}$ (root). Then, we label each initial active clique by a different symbol from Σ_d. In step $n = 1$ we assign the newly added vertex u the code $\mathbf{u} = i$

if the clique with label i was chosen. Further, we label the $d + 1$ newly formed d-simplices by ij for all $j \in \Sigma_d$. Similarly for $n \geq 2$, we assign the newly added vertex v the *label of the clique that becomes inactive* and the newly formed active cliques are given the labels $\mathbf{v}j$, $j = 1, \ldots, d + 1$. It is crucial to keep the labeling *consistent* in a geometrical sense, we are going to describe how to do this in Lemma 3.1 below.

Call the length of a code $|\mathbf{u}|$ the generation of the vertex u and denote the deepest common ancestor of u and v by $\mathbf{u} \wedge \mathbf{v}$, thus $|\mathbf{u} \wedge \mathbf{v}| = \min\{k : u_{k+1} \neq v_{k+1}\}$. For codes $\mathbf{u} = u_1 \ldots u_n$ and $\mathbf{v} = v_1 \ldots v_m$ denote the concatenation $u_1 \ldots u_n v_1 \ldots v_m$ by $\mathbf{uv}$. Further, let $u_{(i)}$ denote the position of the last occurrence of the symbol $i \in \Sigma_d$ in $\mathbf{u}$. We introduce the cut-operators $T_i \mathbf{u} := u_1 \ldots u_{(i)-1}$ and $P_i \mathbf{u} = u_{(i)} \ldots u_n$ for all $i \in \Sigma_d$. The following lemma collects the most important combinatorial observations.

Lemma 3.1 (Tree-like properties of the coding)**.** *There exists a way of choosing the coding of the vertices in the following way:*

(a) **consistency:** *The $d + 1$ neighbors of a newly formed vertex u with code $\mathbf{u}$ are $T_i \mathbf{u}$, $i \in \Sigma_d$.*

(b) *The shortest path between any two vertices $\mathbf{u}$ and $\mathbf{v}$ goes through $\mathbf{u} \wedge \mathbf{v}$ or one of its ancestors.*

(c) *For every $\mathbf{u}, \mathbf{v}, \mathbf{w}$ $\mathrm{Hop}(uv, u) \leq \mathrm{Hop}(uvw, u)$.*

(d) *Assume $|\mathbf{u} \wedge \mathbf{v}| < \min\{|\mathbf{u}|, |\mathbf{v}|\}$. Then $\mathrm{Hop}(u, v) \leq \mathrm{Hop}(uw, v)$.*

Part (a) means that if we have a vertex with code $\mathbf{u}$, the first $d+1$ neighbours of the vertex can be determined by cutting off the last pieces of the code of u, up to the last occurence of a given character $i \in \Sigma_d$. This, in two dimensions means the following: suppose at step $n = 0$ the 'left', 'right' and 'bottom' triangles were given the symbols $1, 2$ and 3 respectively. Then later each new node v with code $\mathbf{v}$ in the middle of a triangle gives rise to the new 'left', 'right' and 'bottom' triangles: to these we *have to* assign the labels $\mathbf{v}1, \mathbf{v}2$ and $\mathbf{v}3$ respectively. Part (b) says that if we have two vertices with code $\mathbf{u}$ and $\mathbf{v}$ in the tree, then the shortest path between them must intersect a path from the root to their deepest common ancestor $\mathbf{u} \wedge \mathbf{v}$. Part (c) ensures that the shortest path between a vertex with code $\mathbf{u}$ and one of its descendants $\mathbf{uv}$ cannot go *below* the vertex $\mathbf{uv}$ in the tree. In more detail, this means that even though there are non-tree (=shortcut) edges upwards in the tree along path towards the origins, these shortcut edges do not jump over each other: the shortcut edges from $\mathbf{uvw}$ (an arbitrary descendant of $\mathbf{uv}$) can only go above $\mathbf{uv}$ to vertices to which $\mathbf{uv}$ is also connected to. Finally, part (d) is a corollary of part (b) and (c): it means that if two vertices u, v are not descendants of each other, then the shortest path between them does not go below them in the tree.

Proof. We prove part (a) by induction. We label the initial $d + 1$ active cliques arbitrarily with $i \in \Sigma_d$. The hypotheses clearly holds in this case. Suppose now we already have an active clique with code $\mathbf{u}$, and by induction, we can assume that $\mathbf{u}$ is associated with the clique formed by vertices $(T_1 \mathbf{u}, T_2 \mathbf{u} \ldots, T_{d+1} \mathbf{u})$. When this clique becomes inactive, the vertex $\mathbf{u}$ is added to the graph and the new active cliques are

$$(T_1 \mathbf{u}, \ldots, T_{k-1} \mathbf{u}, T_{k+1} \mathbf{u}, \ldots, T_{d+1} \mathbf{u}, \mathbf{u}) \text{ for all } k \in \Sigma_d. \tag{3.1}$$

Let us denote this clique by $\mathbf{u}k$. We claim that by this choice, the property (a) is maintained, that is, if the vertex $\mathbf{u}k$ will ever be added to the graph, then its neighbours going to be $T_i(\mathbf{u}k), i \in \Sigma_d$. By construction, the neighbours of $\mathbf{u}k$ are exactly the ones in (3.1), and clearly we have $T_k(\mathbf{u}k) = \mathbf{u}$, and $T_i(\mathbf{u}k) = T_i(\mathbf{u})$ for $i \neq k$, so we can write

$$(T_1 \mathbf{u}, \ldots, T_{k-1} \mathbf{u}, T_{k+1} \mathbf{u}, \ldots, T_{d+1} \mathbf{u}, \mathbf{u})$$
$$= (T_1(\mathbf{u}k), \ldots, T_{k-1}(\mathbf{u}k), T_{k+1}(\mathbf{u}k), \ldots, T_{d+1}(\mathbf{u}k), T_k(\mathbf{u}k)).$$

This proves (a).

(b) Note that by part (a), every vertex is connected to $d + 1$ vertices which code length is shorter than $|\mathbf{u}|$, and all these vertices are descendants of each other, i.e. they are in the path from $\mathbf{u}$ to the root. The other vertices $\mathbf{u}$ is connected to are its descendants, i.e. of the form $\mathbf{uw}$ for some $\mathbf{w}$. Hence if we want to build a path from vertex u to v, we must go up in the tree to at least $\mathbf{u} \wedge \mathbf{v}$.

For (c) it is enough to observe that the position of the last occurrence of a symbol in a code can not be earlier than in some prefix of the same code, i.e. $|T_i(\mathbf{uvw})| > |T_i(\mathbf{uv})|$ for all i.

(d) follows from (b) and (c), since $\mathbf{u} \wedge \mathbf{v} = \mathbf{uw} \wedge \mathbf{v}$ and $\mathrm{Hop}(\mathbf{u}, \mathbf{u} \wedge \mathbf{v}) \leq \mathrm{Hop}(\mathbf{uw}, \mathbf{u} \wedge \mathbf{v})$ and also for all ancestors of $\mathbf{u} \wedge \mathbf{v}$. □

The coding in turn gives a natural grouping of the edges. Edges of the initial graph are not given any name. An edge is called a *forward edge* if its endpoints have codes of the form $\mathbf{u}$ and $\mathbf{u}j$ for $j \in \Sigma_d$. All other edges are called *shortcut edges*. So at each step one new forward edge and d shortcut edges are formed. Figure 2 below shows an example in two dimensions. Initially the symbols $1, 2, 3$ were assigned to the left, right, bottom triangles and continued with the coding the same way.
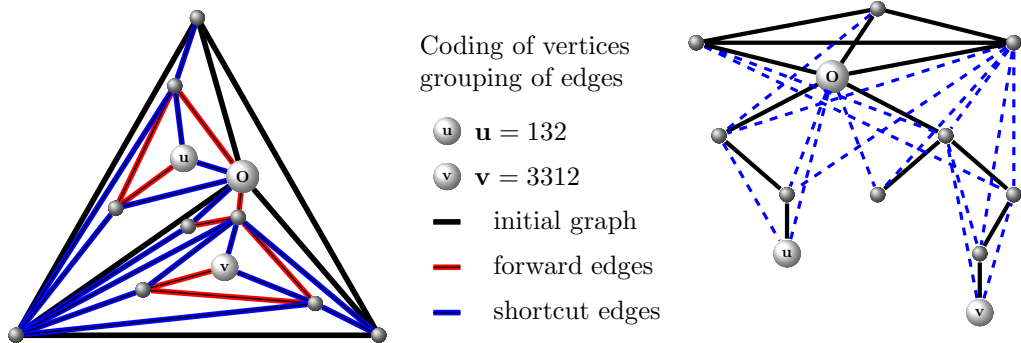

Coding of vertices
grouping of edges

ⓤ $\mathbf{u} = 132$

ⓥ $\mathbf{v} = 3312$

—   initial graph

—   forward edges

—   shortcut edges

FIGURE 2. Tree like structure of a realisation of $\mathrm{RAN}_2(8)$

The grouping of the edges reveals the tree like structure of RANs. Interpreting the initial graph as the root, the forward edges are the *edges of the tree*: along them we can go deeper down in the hierarchy of the graph. The shortcut edges *only run along a tree branch:* between vertices that are in the same line of descent, so we can 'climb up' to the root faster along these edges, see Figure 2.

3.2. **Distances in RANs and EANs: the main idea.** The decomposition above of edges of $\mathrm{RAN}_d(n)$ makes it possible to analyse distances in the graph. Suppose we want to determine the distance between two arbitrary vertices u, v with codes $\mathbf{u}, \mathbf{v}$. Then we can do it as follows:

First, determine the generation of their deepest common ancestor $\mathbf{u} \wedge \mathbf{v}$. Then see, how deep $\mathbf{u}$ and $\mathbf{v}$ are in the tree from $\mathbf{u} \wedge \mathbf{v}$, i.e. determine the length of their code below $\mathbf{u} \wedge \mathbf{v}$. Finally, determine how fast can we reach the deepest common ancestor along the shortcut edges in these two branches, i.e. what is the minimal number of hops we can go up from $\mathbf{u}$ and $\mathbf{v}$ to $\mathbf{u} \wedge \mathbf{v}$?

If we pick u, v uniformly at random, then we have to determine the *typical length of codes* in the tree and the *typical number of shortcut edges* needed to reach the typical common ancestor. If on the other hand we want to analyse the diameter or the flooding time, we have to find a 'long' branch with 'many' shortcut edges. Clearly, one can look at the vertex of *maximal depth* in the tree: but then - by an independence argument about the characters in the code and the length of the code - with high probability the code of the maximal depth vertex in the tree will show *typical behaviour* for the number of shortcut edges. On the other hand, we can calculate how many *slightly shorter branches* are there in the tree. Then, since there are many of them, it is more likely that one of them has a code with more shortcut edges needed than typical. Hence, we study the typical depth and also how many vertices are at larger, atypical depths of a branching process that arises from the forward edges of RANs. The effect of the shortcut edges on the distances is determined using renewal theory (also done in [1]) and large deviation theory. Finally, we optimise parameters such that we achieve the maximal distance by an entropy vs energy argument.

3.3. **Combinatorial analysis of shortcut edges.** Now we investigate the effect of shortcut edges on this tree. By symmetry it is not hard to see that both a uniformly picked individual in the tree has a code where at each position, each character is uniform in Σ_d and further, for two uniformly picked vertices with codes $\mathbf{u}$ and $\mathbf{v}$ the characters in these codes after the position $|\mathbf{u} \wedge \mathbf{v}| + 1$ are again uniformly distributed. Lemma 3.1 part (a) says that the shortcut edges of a vertex $\mathbf{u}$ in the tree are leading exactly to $T_i\mathbf{u}$, the prefixes of $\mathbf{u}$ achieved by deleting every character after the last occurrence of character i in the code. Recall that $P_i\mathbf{u}$ denotes the postfix of $\mathbf{u}$ that starts at the last occurrence of the character $i \in \Sigma_d$ in the code of $\mathbf{u}$. Hence, let us denote the operator that gives the prefix with length $\min_i |T_i\mathbf{u}|$ by $T_{\min}$ and the length of the maximal cut by

$$Y_d^{AN}(\mathbf{u}) := |\mathbf{u}| - |T_{\min}\mathbf{u}| = \max_{i \in \Sigma_d}\{|P_i\mathbf{u}|\}. \tag{3.2}$$

This is the length of the maximal hop we can achieve from the vertex $\mathbf{u}$ upwards in the tree via a shortcut edge.

Consecutively using the operator $T_{\min}$ we can decompose $\mathbf{u}$ into independent blocks, where each block, when reversed, ends at the first appearance of all the characters in Σ_d. We call such a block *full coupon collector block*. E.g. for $\mathbf{u} = 113213323122221131$ this gives $1|132|1332|31222|21131$. Let us denote the total number of blocks needed in this decomposition by

$$N(\mathbf{u}) = \max\{k + 1 : T_{\min}^k\mathbf{u} \neq \varnothing\}. \tag{3.3}$$

Note that this is not the only way to decompose the code in such a way that we always cut only postfixes of the form $P_i\mathbf{u}$: e.g. $1|132|1332|31222211|31$ gives an alternative cut with the same number of blocks.

The following (deterministic) claim establishes that the decomposition along repetitive use of $T_{\min}$ (maximal possible hops) is optimal.

Claim 3.2. *Suppose we have an arbitrary code $\mathbf{u}$ of length n with characters from Σ_d, that we want to decompose into blocks in such a way that* from right to left, *each block ends at the first appearance of some character in that block. Then, the minimal number of blocks needed is given by $N(\mathbf{u})$.*

Proof. Reverse the code to be read from left to right, and call the end of the blocks *barriers*. Denote the position of the i-th barrier by B_i in a decomposition, that does not use full coupon collector blocks all the time, and let $B_i^\star$ stand for the same in the decomposition with full coupon collector blocks. We show that $B_i \leq B_i^\star$, which, combined with the fact that the total number of blocks needed is $\max\{i + 1 : B_i < n\}$, implies the claim. For the first block, clearly $B_1 \leq B_1^\star$ holds. If $B_1 = B_1^\star$, we can cut the first block and compare the next one: let us denote by j the first index that $B_j < B_j^\star$: now the block between B_j and $B_{j+1}^\star$ is *longer* than the block between $B_j^\star, B_{j+1}^\star$, so two things can happen for B_{j+1}: either the block between B_j and $B_{j+1}^\star$ is still an allowed block, and then we can put $B_{j+1} = B_{j+1}^\star$ or it is not, then we have to move the barrier B_{j+1} forward, i.e. $B_{j+1} < B_{j+1}^\star$. From here we can proceed similarly by induction on k for $j + k$. $\square$

For a code of length k with i.i.d. uniform characters from Σ_d let

$$H_k := \max\{i : \sum_{j=1}^{i} Y_d^{(i)} \leq k\}, \tag{3.4}$$

where $Y_d^{(i)}$ are i.i.d copies of Y_d in (2.1), i.e. H_k is the number of consecutive occurrences of full coupon collector blocks in the code.

Lemma 3.3. *Suppose $\mathbf{u}$ is a code of length k with uniform random characters from Σ_d at each position. Then*

$$Y_d^{AN}(\mathbf{u}) \stackrel{d}{=} Y_d \wedge k,$$

$$N(\mathbf{u}) \stackrel{d}{=} H_k.$$

Proof. The last occurrence of any character $i \in \Sigma_d$ in a uniform code is the *first occurrence from backwards* of the same character. Hence, reverse the code of u, and then $|P_i\mathbf{u}|$ is the position of the first occurrence of character i in a uniform sequence of characters, truncated at k, since $|\mathbf{u}| = k$. Maximizing this over all $i \in \Sigma_d$ we get the well-known coupon collector problem, that has distribution Y_d. For the second part, since $N(\mathbf{u})$ cuts down full coupon

collector blocks from the end of the ode of u consecutively, the maximal number of cuts possible is exactly the number of consecutive full coupon collector blocks in the reversed code of $\mathbf{u}$, an i.i.d. code of length k. Since the length of each block has distribution Y_d, and they are independent, the statement follows. $\qquad\square$

Recall μ_d, σ_d^2 from (2.2). From basic renewal theory [21] the following central limit theorem holds

$$\frac{H_k - k/\mu_d}{\sqrt{k\sigma_d^2/\mu_d^3}} \xrightarrow{d} \mathcal{N}(0,1). \tag{3.5}$$

4. Distances in RANs and EANs

Now we make use of the previous section and prove Theorems 2.1, 2.2 and 2.4. Recall that the endpoints of a forward edge $u \to uk$ corresponds to two vertices u and uk ($k \in \Sigma_d$) where the code of u can be obtained by dropping the last symbol of uk. When taking into account only the forward edges of a RAN, there is a natural bijection between the vertices of the RAN and the nodes of a continuous time branching process (CTBP) [3], or a Bellman-Harris process.

Namely, consider a CTBP where the offspring distribution is deterministic: each individual has $d+1$ children and the lifespan of each individual is i.i.d. exponential with mean one. Thus, after birth a node is active for the duration of its lifespan, then splits, becomes inactive and at that instant gives birth to its $d+1$ offspring that become active for their i.i.d. $\mathrm{Exp}(1)$ lifespan.

The bijection between the CTBP at the split times and a RAN_d is the following: the individuals that have already split in the BP are the vertices already present in the RAN, while the active individuals in the BP correspond to the active cliques in the RAN. This holds since in a RAN at every step $d+1$ new active cliques arise in place of the one which becomes inactive. Furthermore, in a RAN an active clique is chosen uniformly at random in each step which is – by the memoryless property of exponential variables – equivalent to the fact that the next individual to split in the BP is a uniformly chosen active individual.

We aim to prove Theorem 2.1 first. Now that the tree-structure of $\mathrm{RAN}_d(n)$ revealed to form a CTBP, we first investigate the distance from the root in this tree. We write G_m for the *generation* of the m-th splitting vertex in the BP, i.e. its graph distance from the root. The next lemma describes the typical size of G_m as well as the *degree of relationship* of two uniformly picked active individuals in the BP.

Lemma 4.1. *Let Z denote a standard normal random variable. Then as $m \to \infty$*

$$\frac{G_m - \frac{d+1}{d}\log m}{\sqrt{\frac{d+1}{d}\log m}} \xrightarrow{d} Z. \tag{4.1}$$

Further, let G_U, G_V denote the generation of two independently and uniformly picked active vertices in the BP after the mth split, and let us write $G_{U \wedge V}$ for the generation of the latest common ancestor of U, V. Then the marginal distribution $G_U \stackrel{d}{=} G_{m+1}$, and

$$\left(\frac{G_U - G_{U \wedge V} - \frac{d+1}{d}\log m}{\sqrt{\frac{d+1}{d}\log m}}, \frac{G_V - G_{U \wedge V} - \frac{d+1}{d}\log m}{\sqrt{\frac{d+1}{d}\log m}} \right) \xrightarrow{d} (Z, Z'), \tag{4.2}$$

where Z, Z' are independent standard normal distributions.

Proof. The lemma follows from the *proof of* the CLT for the *degree of relationship* in Bühler [15, Theorem 3.3]. The distribution of the generation of the m-th chosen vertex G_m is the sum of independent Bernoulli random variables (see [15] or [6]). More precisely,

$$G_m \stackrel{\mathrm{d}}{=} \sum_{i=1}^{m} \mathbb{1}_i, \text{ where } \mathbb{P}\left(\mathbb{1}_i = 1\right) = \frac{d+1}{di+1}, \tag{4.3}$$

The indicator $\mathbb{1}_i$ is 0 or 1 whether in the line of ancestry of the m-th splitting vertex, the individual at the k-th split is surviving from time $k-1$ or newborn. From this identity it is easy to calculate the expectation and variance of G_m:

$$\mathbb{E}\left[G_m\right] = \frac{d+1}{d}\log m, \qquad \mathbb{D}^2\left[G_m\right] = \mathbb{E}\left[G_m\right] + O\left(m^{-1}\right). \tag{4.4}$$

The central limit theorem (4.1) holds for the standardization of G_m since the collection of Bernoulli random variables $\{\mathbb{1}_i\}_{i=1}^m$, $m = 1, 2, \ldots$ satisfy Lindeberg's condition.

For the second statement, $G_U \overset{d}{=} G_{m+1}$ is immediate from the fact that the $m+1$th vertex to split in the process is a uniformly picked vertex among the active vertices.

Next, let us write $\tau_{U \wedge V}$ for the time when the latest common ancestor of U and V splits. Then, conditioned on $\tau_{U \wedge V}$ the following two variables are independent and their joint distribution can be written as

$$(G_U - G_{U \wedge V}, G_V - G_{U \wedge V}) \overset{d}{=} \left(\sum_{i=\tau_{U \wedge V}}^{m+1} \mathbb{1}_i, \sum_{i=\tau_{U \wedge V}}^{m+1} \mathbb{1}'_i \right), \qquad (4.5)$$

where

$$\mathbb{P}\left((\mathbb{1}_i, \mathbb{1}'_i) = (1,0)|\tau_{U \wedge V} < i\right) = \frac{d+1}{di+1}\frac{i-1}{i}$$

$$\mathbb{P}\left((\mathbb{1}_i, \mathbb{1}'_i) = (0,1)|\tau_{U \wedge V} < i\right) = \frac{d+1}{di+1}\frac{i-1}{i}$$

$$\mathbb{P}\left((\mathbb{1}_i, \mathbb{1}'_i) = (1,1), \tau_{U \wedge V} = i|\tau_{U \wedge V} \le i\right) = \frac{(d+1)}{(di+1)i}$$

and conditioned on $\tau_{U \wedge V}$, different indices are independent. To see that $\tau_{U \wedge V}$ has a limiting distribution we can use [15, Lemma 3.3]:

$$\mathbb{P}(\tau_{U \wedge V} \le k) = \prod_{i=k+1}^{m} (1 - \mathbb{P}(\tau_{U \wedge V} = i|\tau_{U \wedge V} \le i)), \qquad (4.6)$$

where the factors on the right hand side are the probabilities that the two ancestral lines do not merge at the i-th split. This is tending to a proper limiting distribution since

$$\sum_{i=1}^{\infty} \mathbb{P}(\tau_{U \wedge V} = i|\tau_{U \wedge V} \le i) = \sum_{i=1}^{\infty} \frac{d+1}{(di+1)di} \le \infty.$$

Hence, $\tau_{U \wedge V}$ has a limiting distribution, and clearly $(m - \tau_{U \wedge V})/m \to 1$. From here, one can show the joint convergence of (4.5) using Lindeberg CLT for linear combinations of $\sum_{i=\tau_{U \wedge V}+1}^{m+1}(\alpha \mathbb{1}_i + \beta \mathbb{1}'_i)$ and get that the two variables in (4.2) tend jointly to a 2 dimensional standard normal variable. $\qquad \square$

Now we are ready to prove Theorem 2.1.

Proof of Theorem 2.1. Pick two uniform active cliques u, v in the graph. We write $\mathbf{u}, \mathbf{v}$ for the codes of these cliques and the corresponding vertices, $G_U = |\mathbf{u}|, G_V = |\mathbf{v}|$ for their generation. As before, we write $\mathbf{u} \wedge \mathbf{v}$ for the latest common ancestor, i.e. their longest common prefix. Let us define the distinct postfixes after $\mathbf{u} \wedge \mathbf{v}$ by

$$\mathbf{u} = (\mathbf{u} \wedge \mathbf{v})\widetilde{\mathbf{u}}, \quad \mathbf{v} = (\mathbf{u} \wedge \mathbf{v})\widetilde{\mathbf{v}}.$$

Note that $|\mathbf{u} \wedge \mathbf{v}| = G_{U \wedge V}$, $|\widetilde{\mathbf{u}}| = G_U - G_{U \wedge V}$, $|\widetilde{\mathbf{v}}| = G_V - G_{U \wedge V}$. By Lemma 3.1 and Claim 3.2 the length of the shortest path between u, v satisfies:

$$\mathrm{dist}(u,v) = N(\widetilde{\mathbf{u}}) + N(\widetilde{\mathbf{v}}),$$

and Lemma 4.1 describes the typical distance between $\mathbf{u}$ and $\mathbf{v}$ along the tree. Note that except the first character, the characters in the codes $\widetilde{\mathbf{u}}$ and $\widetilde{\mathbf{v}}$ are i.i.d. uniform on Σ_d. Hence, by Lemma 3.3 combined with (4.4)

$$\mathbb{E}\left[H_{G_U - G_{U \wedge V}}\right] = \mathbb{E}\left[\mathbb{P}(H_{G_U - G_{U \wedge V}}|G_U - G_{U \wedge V})\right] = \frac{1}{\mu_d}\frac{d+1}{d}\log m(1 + o(1)).$$

To obtain a central limit theorem for $H_{G_U - G_{U \wedge V}}$ observe that

$$\frac{H_{G_U - G_{U \wedge V}} - \frac{1}{\mu_d}\frac{d+1}{d}\log m}{\sqrt{\frac{d+1}{d}\log m\,\sigma_d^2/\mu_d^3}} = \frac{H_{G_U - G_{U \wedge V}} - \frac{1}{\mu_d}(G_U - G_{U \wedge V})}{\sqrt{(G_U - G_{U \wedge V})\sigma_d^2/\mu_d^3}} \cdot \sqrt{\frac{G_U - G_{U \wedge V}}{\frac{d+1}{d}\log m}}$$

$$+ \frac{\frac{1}{\mu_d}(G_U - G_{U \wedge V}) - \frac{1}{\mu_d}\frac{d+1}{d}\log m}{\sqrt{\frac{d+1}{d}\log m\,\sigma_d^2/\mu_d^3}}. \qquad (4.7)$$

The first factor on the right hand side, conditionally on $G_U - G_{U \wedge V}$ with $G_U - G_{U \wedge V} \to \infty$, tends to a standard normal random variable independent of G_U by the renewal CLT in (3.5) and the second factor tends to one in probability by (4.2). By (4.2) again, the second term tends to a $\mathcal{N}(0, \mu_d/\sigma_d^2)$. The two limiting normals are independent, thus

$$\frac{H_{G_U - G_{U \wedge V}} - \frac{1}{\mu_d}\frac{d+1}{d}\log m}{\sqrt{\frac{d+1}{d}\log m\, \sigma_d^2/\mu_d^3}} \xrightarrow{d} \mathcal{N}(0, 1 + \mu_d/\sigma_d^2). \tag{4.8}$$

By conditioning first on $G_{U \wedge V}$, (as in the proof of Lemma 4.1) and using that the characters in the code of $\widetilde{\mathbf{u}}, \widetilde{\mathbf{v}}$ are all i.i.d. uniform in Σ_d, one can similarly show that $(H_{G_U - G_{U \wedge V}}, H_{G_V - G_{U \wedge V}})$ tend jointly to two independent copies of $\mathcal{N}(0, 1 + \mu_d/\sigma_d^2)$ variables. Using now that $\mathrm{Hop}(n) = H_{G_U - G_{U \wedge V}} + H_{G_V - G_{U \wedge V}}$, the first statement of the Theorem 2.1 immediately follows by normalising such that the total variance is 1. The second statement follows by calculating how many active cliques a vertex with degree k is contained: a vertex with degree $d + 1$ is contained in $d + 1$ cliques, and when the degree of a node v increases by 1, then the number of cliques containing v increases by $d - 1$, thus a vertex with degree $k \geq d + 1$ is contained in exactly

$$A_k = k(d - 1) + d^2 - d + 2 \tag{4.9}$$

active cliques. It is also not hard to see that the total number of active clicks after n steps is $A(n) = dn + d + 2$. This means, that picking two vertices x, y according to the size-biased distribution given in (2.5) is equivalent that the uniformly picked active cliques U, V are neighbouring these vertices. The distance between x, y is then between $N(\widetilde{\mathbf{u}}) + N(\widetilde{\mathbf{v}}) - 2, N(\widetilde{\mathbf{u}}) + N(\widetilde{\mathbf{v}})$ since by Lemma 3.1 $x = T_i \mathbf{u}$ for some $i \in \Sigma_i$, hence we can gain at most 1 hop by considering x instead of the clique U and the same holds for y and V. Hence, the CLT for U, V implies a CLT for two vertices picked according to the probabilities in (2.5). $\qquad \square$

Remark 4.2. If we pick two vertices (x,y) of $\mathrm{RAN}_d(n)$ uniformly at random, then we should randomise the index m of G_m to be discrete uniform in $[1, 2, \ldots, m]$. One can easily see that in this case the CLT does not hold anymore. We still have $\frac{\mathrm{dist}(x,y)}{2\frac{d+1}{d}\log m} \xrightarrow{\mathbf{P}} 1$.

Since the proofs are similar, we prove Theorem 2.4 next.

Proof of Theorem 2.4. The proof follows analogous lines to the proof of Theorem 2.1, hence we give only the sketch. The main idea here is that the tree can be viewed as a BP where at step i, each active individual splits with probability q_i or stays active for the next step with probability $1 - q_i$. Hence, Lemma 4.1 can be modified as follows: the generation of a uniformly picked active individual at step m satisfies

$$\widetilde{G}_U \overset{d}{=} \sum_{i=1}^{m} \widetilde{\mathbb{1}}_i,$$

where $\mathbb{P}(\widetilde{\mathbb{1}}_i = 1) = q_i$ and different indices are independent. This corresponds to following the ancestral line of U: in this ancestral line the generation is increased by 1 in the ith step if the individual active at step i split and stays if it did not split. Since splitting happens with probability q_i at step i, the result follows and the CLT for $\widetilde{G}_U$ holds by Lindeberg CLT. Now, for two uniformly picked individuals U, V we have

$$(G_U - G_{U \wedge V}, G_V - G_{U \wedge V}) \overset{d}{=} \left(\sum_{i=\tau_{U \wedge V}}^{m} \widetilde{\mathbb{1}}_i, \sum_{i=\tau_{U \wedge V}}^{m} \widetilde{\mathbb{1}}_i' \right),$$

where different indices are independent and conditioned on $\tau_{U \wedge V}$, $\widetilde{\mathbb{1}}_i, \widetilde{\mathbb{1}}_i'$ are independent indicators with $\mathbb{P}(\widetilde{\mathbb{1}}_i = 1) = \mathbb{P}(\widetilde{\mathbb{1}}_i' = 1) = q_i$. Since the variance $\sum_i q_i(1 - q_i) \to \infty$, the joint CLT follows in a similar manner then for Lemma 4.1 if we can show that $\tau_{U \wedge V}$ has a limiting distribution. For this note that similarly as in (4.6)

$$\mathbb{P}(\tau_{U \wedge V} \leq k) = \prod_{i=k+1}^{m} (1 - \mathbb{P}(\tau_{U \wedge V} = i | \tau_{U \wedge V} \leq i)) \tag{4.10}$$

and the factors on the right hand side express that the two ancestral lines of U, V do not merge yet at step i. Let us write $A(i)$ for the number of active vertices at step i. Then at step i there are $Z_i := \mathrm{Bin}(A(i), q_i)$ many vertices that split, each of them producing $d + 1$

new active vertices, and hence the probability that the two ancestral lines merge at step i, conditioned on $A(i), A(i + 1)$ equals

$$\mathbb{P}\left(\tau_{U \wedge V} = i | \tau_{U \wedge V} \leq i, A(i), A(i + 1)\right) = \frac{Z_i(d + 1)d}{(A(i + 1)(A(i + 1) - 1)}, \qquad (4.11)$$

where $A(i + 1) = A(i) + Z_i(d + 1)$, the new number of active vertices after the ith split. If the sum in $i \in \mathbb{N}$ on the right hand side in the previous display is a.s. finite then (4.10) ensures that $\tau_{U \wedge V}$ has a proper limiting distribution. Hence we aim to show that this is the case whenever the total number of vertices $N(n) \to \infty$. That is,

$$\sum_{i=1}^{\infty} \frac{Z_i(d + 1)d}{A(i + 1)(A(i + 1) - 1)} \leq \infty \qquad \text{a.s. on } \{N(n) \to \infty\}.$$

Since $A(i+1) = N(i+1)d + d + 1$, and $Z_i = N(i+1) - N(i)$, we can approximate the above sum by

$$(d + 1) \sum_{i=1}^{\infty} \frac{d(N(i + 1) - N(i))}{(dN(i + 1))^2}.$$

Now we can interpolate $N(i)$ with a continuous function and then this sum can be approximated by the integral

$$\int_1^{\infty} \frac{N'(x)}{N(x)^2} dx < \infty$$

as long as $N(n) \to \infty$. Note that as long as $\sum_{n \in \mathbb{N}} q_n = \infty$, this is the case by the second Borel-Cantelli lemma: in each step we add at least a new node with probability q_n. The CLT then for the distances follows in the exact same manner as in the proof of Theorem 2.1. □

Next we aim to prove Theorem 2.2, but we need some preliminary lemmas first. Lemma 4.3 shows that the tail of G_m basically behaves as the the tail of the sum of $\log m$ i.i.d. Poi$((d + 1)/d)$ random variables and gives the depth of the longest branch in our BP, while Lemma 4.5 gives a large deviation bound for H_k (recall (3.4)).

Lemma 4.3. *The exact asymptotic tail behaviour of G_m is given by*

$$\lim_{m \to \infty} \frac{\log\left(\mathbb{P}\left(G_m > c \log m\right)\right)}{\log m} = c - \frac{d + 1}{d} - c \log\left(\frac{d}{d + 1}c\right) =: f_d(c). \qquad (4.12)$$

Further, in the branching process tree having m vertices, the deepest branch satisfies

$$\frac{\max_{i \leq m} G_i}{\log m} \xrightarrow{\mathbf{P}} \tilde{c}_d, \qquad (4.13)$$

where $\tilde{c}_d := \{c_d > (d + 1)/d, \ f_d(c_d) = -1\}$ (as defined in Theorem 2.2).

Proof. Let us calculate the following:

$$\frac{1}{\log m} \log \mathbb{E}\left[e^{\theta G_m}\right] = \frac{1}{\log m} \sum_{i=1}^{m} \log\left(1 + \frac{d + 1}{di + 1}(e^\theta - 1)\right).$$

From the series expansion of $\log(1 + x)$ we can see that

$$\lim_{m \to \infty} \frac{1}{\log m} \log \mathbb{E}\left[e^{\theta G_m}\right] = \frac{d + 1}{d}(e^\theta - 1),$$

which is the cumulant generating function of a $\xi = \text{Poi}((d + 1)/d)$ random variable. The rate function of such a random variable is $-f_d(c)$. Hence, by the Gärtner-Ellis theorem for $\xi^{(i)} \sim \xi$ i.i.d.

$$\lim_{m \to \infty} \frac{\log\left(\mathbb{P}\left(G_m > c \log m\right)\right)}{\log m} = \lim_{m \to \infty} \frac{\log\left(\mathbb{P}\left(\sum_{i=1}^{\log m} \xi^{(i)} > c \log m\right)\right)}{\log m} = f_d(c).$$

As for (4.13), since $f_d(\frac{d+1}{d}) = 0$, choosing any $c > \tilde{c}_d$ results in a summable bound on (4.12). Thus by the Borel-Cantelli lemma, for any $c > \tilde{c}_d$ there are only finitely many m such that the event $\{G_m > c \log m\}$ holds, giving the whp upper bound $\tilde{c}_d \log m$ on the depth of the BP. The lower bound can be obtained based on work of Broutin and Devroye [14] since our BP is a special case of so-called random lopsided trees [16, 26]. □

Remark 4.4. The function $f_d(c)$ in (4.12) is strictly negative outside $c = (d+1)/d$. This immediately implies the weaker result $\frac{G_m}{\log m} \xrightarrow{\mathbf{P}} \frac{d+1}{d}$.

Lemma 4.5. *For $1 \leq \beta \leq \mu_d/(d+1)$, the number of consecutive occurrences of full coupon collector blocks in a code of length k satisfies the large deviation*

$$\lim_{k \to \infty} \frac{1}{k} \log \left(\mathbb{P}\left(H_k > \frac{\beta}{\mu_d} k \right) \right) = -\frac{\beta}{\mu_d} I_d\left(\frac{\mu_d}{\beta} \right), \tag{4.14}$$

where the large deviation rate function $I_d(x)$ of Y_d was defined in (2.3).

Proof. Let $Y_d^{(i)}$ be i.i.d. distributed according to Y_d. Since

$$\mathbb{P}\left(H_k > \frac{\beta}{\mu_d} k \right) = \mathbb{P}\left(\sum_{i=1}^{k\beta/\mu_d} Y_d^{(i)} < \left(\frac{\beta}{\mu_d} k \right) \cdot \frac{\mu_d}{\beta} \right),$$

we can apply Cramér's theorem to obtain (4.14). $\qquad\square$

Remark 4.6. The rate function $I_d(x)$ has no explicit form. It can be computed numerically from

$$I_d(x) = \lambda^*(x) \cdot x - \log \mathbb{E}\left[e^{\lambda^*(x)Y_d} \right],$$

where $\lambda^*(x)$ is the unique solution to the equation $\frac{\partial}{\partial \lambda} \log \mathbb{E}\left[e^{\lambda Y_d} \right] = x$ and

$$\log \mathbb{E}\left[e^{\lambda Y_d} \right] = \log d! - d \log(d+1) + (d+1)\lambda - \sum_{i=1}^{d} \log \left(1 - \frac{i}{d+1} e^{\lambda} \right).$$

Now we proceed to the proof of Theorem 2.2.

Proof of Theorem 2.2. First, we have seen at the end of the proof of Theorem 2.1 that switching from vertices to neighbouring active cliques only changes the distances by 2, hence we rather investigate the diameter of the vertices coded by active cliques. Let us denote the set of active cliques at step m by $\mathcal{A}(m)$. We have seen that $|\mathcal{A}(m)| = dm + d + 1$. Our aim is to estimate the number of active cliques u whose codes are relatively long i.e. at least $\alpha \tilde{c}_d \log m$ for $\alpha \in (0, 1]$, and in their code $N(\mathbf{u})$, the number of hops is larger than expected, i.e. at least $n\beta/\mu_d$ for $\beta \in [1, \mu_d/(d+1)]$. Define the indicator variables for each clique $u \in \mathcal{A}(m)$

$$J_u = \mathbb{1}\left[G_u > \alpha \tilde{c}_d \log m, \ H_{G_u} > \frac{\beta}{\mu_d} \alpha \tilde{c}_d \log m \right].$$

Summing over $u \in \mathcal{A}(m)$, the expected number of such cliques is

$$|\mathcal{A}(m)| \mathbb{E}\left[\frac{1}{|\mathcal{A}(m)|} \sum_{u \in \mathcal{A}(m)} J_u \right] = (dm + d + 1)\mathbb{P}\left(G_u > \alpha \tilde{c}_d \log m \right) \mathbb{P}\left[H_{G_u} > \frac{\beta}{\mu_d} \alpha \tilde{c}_d \log m \right],$$

since the length of a code and its coordinates are independent, and the expected value on the left hand side corresponds to a uniformly picked active clique $U \in \mathcal{A}(m)$. By Lemma 4.1 $G_u \stackrel{d}{=} G_{m+1}$, hence (4.12) and (4.14) implies that

$$\mathbb{E}\left[\sum_{u \in \mathcal{A}(m)} J_u \right] = m^{\left(1 + f_d(\alpha \tilde{c}_d) - \alpha \tilde{c}_d \frac{\beta}{\mu_d} I_d\left(\frac{\mu_d}{\beta} \right) \right)(1+o(1))} =: m^{g(\alpha, \beta)(1+o(1))}. \tag{4.15}$$

We wish to choose (α, β) so that $\mathbb{P}\left(\sum_{u \in \mathcal{A}(m)} J_u > 0 \right) > 0$. Thus necessarily $g(\alpha, \beta) \geq 0$. We will need a upper bound on the second moment

$$\mathbb{E}\left[\left(\sum_{u \in \mathcal{A}(m)} J_u \right)^2 \right] = \sum_{u \in \mathcal{A}(m)} \mathbb{E}[J_u] + \sum_{u,v \in \mathcal{A}(m), u \neq v} \mathbb{E}[J_u J_v].$$

Using the proof of Lemma 4.1, we know that the joint distribution of two uniformly picked individuals satisfies that their common ancestor $G_{u \wedge v}$ has a limiting distribution and conditioned on the splitting time $\tau_{u \wedge v}$ of $u \wedge v$, the joint distribution of $G_u - G_{u \wedge v}, G_v - G_{u \wedge v}$ can be described as the sum of indicators. Further, the characters after $u \wedge v$ in both codes are independent and uniform in Σ_d. Hence, it is not hard to see that

$$\sum_{u,v \in \mathcal{A}(m), u \neq v} \mathbb{E}[J_u J_v] \leq \mathbb{E}[\sum_{u \in \mathcal{A}(m)} J_u]^2. \tag{4.16}$$

From a Cauchy-Schwarz inequality followed by (4.16)

$$\mathbb{P}\Big(\sum_{u\in\mathcal{A}(m)} J_u > 0\Big) \geq \frac{\mathbb{E}[\sum_{u\in\mathcal{A}(m)} J_u]^2}{\mathbb{E}\Big[\big(\sum_{u\in\mathcal{A}(m)} J_u\big)^2\Big]} \geq \frac{\mathbb{E}[\sum_{u\in\mathcal{A}(m)} J_u]^2}{\mathbb{E}[\sum_{u\in\mathcal{A}(m)} J_u] + \mathbb{E}[\sum_{u\in\mathcal{A}(m)} J_u]^2} > 0, \quad (4.17)$$

if and only if $g(\alpha,\beta) \geq 0$. From this it is immediate that for the correct order of magnitude of the diameter of the graph we need to pick the largest product $\alpha\beta$ with the property that $g(\alpha,\beta) \geq 0$. Since $g(\alpha,\beta)$ is decreasing both in α and in β, if $g(\alpha,\beta) > 0$ then we can increase one or both of the parameters $\alpha' \geq \alpha, \beta' \geq \beta$ s.t. $\alpha'\beta' > \alpha\beta$ and $g(\alpha',\beta') = 0$. Hence, the diameter can be achieved if we restrict the problem to (α,β) pairs so that the exponent $g(\alpha,\beta)$ equals zero. Thus we arrive at the maximization problem (2.7) in Theorem 2.2:

$$\max_{\alpha,\beta}\Big\{\frac{\tilde{c}_d}{\mu_d}\alpha\beta \ : \ g(\alpha,\beta) = 0, \ (\alpha,\beta) \in (0,1] \times [1, \frac{\mu_d}{d+1}]\Big\}.$$

We show that this maximization problem has a unique solution $(\tilde{\alpha}, \tilde{\beta})$ in Lemma 4.7 below.

Now apply the second moment method in (4.17) with the maximising $(\tilde{\alpha}, \tilde{\beta})$ for a lower bound and any other $(\alpha',\beta') : \alpha'\beta' > \tilde{\alpha}\tilde{\beta}$ and Markov's inequality for an upper bound to finally conclude

$$\frac{\max_{u\in\mathcal{A}(m)} H_{G_u}}{\log m} \xrightarrow{\mathbf{P}} \frac{\tilde{c}_d}{\mu_d}\tilde{\alpha}\cdot\tilde{\beta}. \qquad (4.18)$$

The statement of Theorem 2.2 for the flooding time now follows from the fact that if u is a uniformly picked active clique, then Lemma 4.1 implies that the CLT holds for its generation G_u, and since the characters are uniform in the code of u, similarly as in (4.7), the CLT holds for H_{G_u} as well. Further, since in $\mathrm{Flood}(u,v)$ we maximise the distance over the choice of the other vertex v, clearly whp we can pick v such that the latest common ancestor $u \wedge v$ is the root itself. This combined with the fact that the distance changes only by at most 2 if we consider active cliques instead of vertices in the graph implies

$$\mathrm{Flood}(n) \stackrel{d}{=} H_{G_u} + \max_{v\in\mathcal{A}(n)} H_{G_v}$$

and the statement of the theorem follows from the distributional convergence of $\frac{H_{G_u}}{\log m} \xrightarrow{d} \frac{d+1}{d\mu_d}$ and (4.18).

For the diameter we have

$$\frac{\mathrm{Diam}(n)}{\log m} \stackrel{d}{=} 2\frac{\max_{v\in\mathcal{A}(n)} H_{G_v}}{\log m},$$

since for any $\varepsilon > 0$, whp there are at least two vertices that are not closely related to each other and both satisfy $H_{G_v}/\log m > \frac{\tilde{c}_d}{\mu_d}\tilde{\alpha}\tilde{\beta}(1-\varepsilon)$, but whp there are no vertices that satisfy $H_{G_v}/\log m > \frac{\tilde{c}_d}{\mu_d}\tilde{\alpha}\tilde{\beta}(1+\varepsilon)$. □

We are left to analyse the maximization problem:

Lemma 4.7. *The maximization problem* (2.7) *has a unique solution* $(\tilde{\alpha}, \tilde{\beta}) \in (0,1] \times [1, \frac{\mu_d}{d+1}]$, *and further this solution satisfies*

$$\tilde{\alpha} = \frac{1}{\tilde{c}_d}\frac{d+1}{d}\exp\big\{-I_d'(\mu_d/\tilde{\beta})\big\}, \quad \frac{\tilde{\beta}}{\mu_d}I_d\Big(\frac{\mu_d}{\tilde{\beta}}\Big) = \frac{1 + f_d(\tilde{\alpha}\tilde{c}_d)}{\tilde{\alpha}\tilde{c}_d}.$$

Proof. Define the Lagrange multiplier function $\mathcal{L}(\alpha,\beta,\lambda) := \alpha\beta - \lambda g(\alpha,\beta)$. Necessarily the optimal $(\tilde{\alpha}, \tilde{\beta})$ satisfies $\nabla\mathcal{L}(\tilde{\alpha}, \tilde{\beta}, \tilde{\lambda}) = 0$. The partial derivative $\mathcal{L}(\alpha,\beta,\lambda)_\lambda' = 0$ simply gives the condition $g(\alpha,\beta) = 0$. Further, the optimising $\tilde{\lambda}$ can be expressed from $\mathcal{L}(\alpha,\beta,\lambda)_\alpha' = 0$ and $\mathcal{L}(\alpha,\beta,\lambda)_\beta' = 0$ and satisfies

$$\tilde{\lambda} = \frac{\beta}{\frac{\partial}{\partial\alpha}g(\alpha,\beta)} = \frac{\alpha}{\frac{\partial}{\partial\beta}g(\alpha,\beta)}.$$

After differentiation of $g(\alpha,\beta) = 1 + f_d(\alpha\tilde{c}_d) - \alpha\tilde{c}_d\frac{\beta}{\mu_d}I_d\big(\frac{\mu_d}{\beta}\big)$, rearranging terms and using that $f_d'(x) = -\log\big(\frac{d}{d+1}x\big)$ we obtain the first condition. To check the sufficiency we look at

the bordered Hessian

$$
\begin{bmatrix}
0 & \frac{\partial g}{\partial \alpha} & \frac{\partial g}{\partial \beta} \\
\frac{\partial g}{\partial \alpha} & \frac{\partial^2 \alpha\beta}{\partial \alpha^2} & \frac{\partial^2 \alpha\beta}{\partial \alpha \partial \beta} \\
\frac{\partial g}{\partial \beta} & \frac{\partial^2 \alpha\beta}{\partial \alpha \partial \beta} & \frac{\partial^2 \alpha\beta}{\partial \beta^2}
\end{bmatrix}
=
\begin{bmatrix}
0 & \frac{\partial g}{\partial \alpha} & \frac{\partial g}{\partial \beta} \\
\frac{\partial g}{\partial \alpha} & 0 & 1 \\
\frac{\partial g}{\partial \beta} & 1 & 0
\end{bmatrix}.
$$

Its determinant is $\left(\frac{\partial^2 g(\alpha,\beta)}{\partial \alpha \partial \beta}\right)^2 > 0$, thus the condition is also sufficient. We note that the solution can be approximated by numerical methods. $\qquad\square$

Remark 4.8. We mention here the difficulties in the analysis of the diameter and flooding time of EANs: the main difficulty here is to understand the proper correlation structure of the codes (and shortcut edges) on the vertices of the BP: (a) The corresponding BP tree is fatter than the BP for RAN as soon as $n^{-1} = o(q_n)$. (b) In each step each vertex splits independently of the past with probability q_n. (a) and (b) together imply that even though the marginal distribution of the characters of a uniformly picked clique U is uniform in Σ_d, still it is more likely that the 'neighbouring codes' are also present in the graph and hence codes for which $N(\mathbf{u})$ is large are more likely to appear. Hence we expect that the diameter will have a larger constant in front of $\sum q_i$ than the constant in front of $\log n$ for RAN. (Compare it to the diameter of the deterministic AN: with $q_n \equiv 1$ it is not hard to see that $\mathrm{diam}(\mathrm{AN}_d(n)) = 2n/(d+1)$).

5. Degree distribution of Apollonian networks

In this section we prove the results related to the degree distribution. We start with analysing RANs first.

5.1. **Sketch of Proof of Theorem 2.6.** The proof of Theorem 2.6 determining the degree distribution of RANs consist of two main steps that are described in Lemmas 5.1 and 5.2. The first lemma shows that $N_k(n)$ is getting close to its expectation uniformly in k as $n \to \infty$. The method we describe here is an adaptation of the standard martingale method and similar to that in [9, 12, 24]. Parallel to our work, Frieze and Tsourakakis [22] applied this method to show Lemma 5.1 and 5.2 for two dimension and their proof can be generalized to higher dimensions without any difficulty, hence we only give a sketch of proof here.

Lemma 5.1. *Frieze and Tsourakakis* [22] *Fix $d \geq 2$ and $c_1 > \sqrt{8}(d+1)$. Then*

$$
\mathbb{P}\left(\max_k \left| \widetilde{N}_k(n) - \mathbb{E}\big[\widetilde{N}_k(n)\big] \right| \geq c_1 \sqrt{n \log n}\right) = o(1).
$$

This lemma tells us that $\widetilde{N}_k(n)$ is concentrating around its expected value. From this we immediately get the concentration of $\widetilde{p}_k(n) = \widetilde{N}_k(n)/(n+d+2)$ around its expected value. The second lemma approximates the difference between this expected value and p_k, the stationary distribution.

Lemma 5.2. *There exists a probability distribution $\{p_k\}_{k=d+1}^{\infty}$ for which for any $n \geq 0$ and for any $k \geq d+1$*

$$
\left| \mathbb{E}[\widetilde{N}_k(n)] - p_k(n+d+2) \right| \leq c_2 \sqrt{n \log n}
$$

with some constant c_2. The distribution $\{p_k\}_{k \in \mathbb{N}}$ is determined in (5.12) and it has a power-law asymptotic decay with exponent $\frac{2d-1}{d-1} \in (2,3]$ for $d \geq 2$.

As mentioned above, we do not give the proof of these lemmas here. The methods however are similar then to the ones used in the proof of Lemma 5.3, 5.4 for the EAN below. Given these two lemmas, the proof of Theorem 2.6 follows:

Proof of Theorem 2.6. By triangle inequality Theorem 2.6 immediately follows from Lemmas 5.1 and 5.2 with $c = c_1 + c_2$. $\qquad\square$

5.2. **Proof of Theorem 2.7.** We prove Theorem 2.7 connected to EANs again in two main steps, as in the case of RANs. Recall the definition $p_k(n)$ from (2.11). We denote the additional number of vertices after n steps by $N(n)$, that is, $|V(n)| = N(n)+d+2$. Note that $|V(n)|$ is random, hence here it is better to have a conditional concentration result: Let us denote the sigma algebra generated by $\{N(1), \ldots, N(n)\}$ by $\mathcal{G}_n$. The following lemma tells us that the empirical proportion of degree k vertices is concentrating around its $\mathcal{G}_n$-conditional mean:

Lemma 5.3. *Fix the dimension $d \geq 2$, a constant $c > \sqrt{8}(d+1)^{3/2}$, and a sequence of node arrival probabilities $\{q_n\}_{n=1}^{\infty}$ such that $N(n) \to \infty$ a.s. as $n \to \infty$. Then*

$$\mathbb{P}\left(\max_k \left| p_k(n) - \mathbb{E}\left[N_k(n)|\mathcal{G}_n\right]/|V(n)|\right| \geq c\sqrt{\frac{\log N(n)}{N(n)}} \,\Big| N(n) \right) = o(1). \qquad (5.1)$$

The next lemma tells us that the $\mathcal{G}_n$-conditional mean of the proportion of degree k vertices is tending to p_k given in (2.10):

Lemma 5.4. *Let $d \geq 2$ and let us assume $q_n \to 0$, $N(n) \to \infty$ a.s. Then there exists constants $0 < C_k < \infty$ for which for any $k \geq d+1$ and for any $\delta > 0$*

$$\left|\mathbb{E}\left[N_k(n)|\mathcal{G}_n\right]/|V(n)| - p_k\right| \leq C_k\Big(\sum_{i=1}^{n} \tfrac{N(i+1)-N(i)}{dN(i)} \Big)^{k-d}.$$

The distribution p_k is the same as the asymptotic degree distribution of RAN_d given in (2.10).

Proof of Theorem 2.7. The proof of the theorem is immediate from the triangle inequality and Lemma 5.3, 5.4 once we establish that (a) $N(n) \to \infty$ and (b) the error term in Lemma 5.4 is at most of order $\sqrt{\log N(n)}$. For (a) note that in each step there is always at least one active clique, hence the number of new vertices added in step i is bounded from below by an indicator variable that has success probability q_i. Hence, by the second Borel Cantelli lemma, $N(n) \to \infty$ a.s. as long as $\sum_{n \in \mathbb{N}} q_n = \infty$.

For (b) we need to compare $N(n)$ to the order of the error terms in Lemma 5.4: the number of active cliques $A(n) = dN(n) + d + 1$, hence the number of new vertices after the $n+1$th step satisfies $N(n+1) - N(n) = \mathrm{Bin}(dN(n)+d+1, q_n)$, with $N(1) := \mathrm{Bin}(d+1, q_1)$. Now we can approximate $N(n)$ by a continuous function to estimate the error terms in Lemma 5.4:

$$\sum_{i=1}^{n} \frac{N(i+1)-N(i)}{dN(i)} \leq \frac{C}{d}\int_1^n \frac{N'(x)}{N(x)} \leq C\log N(n).$$

and therefore for any fixed k

$$C_k\Big(\sum_{i=1}^{n} \tfrac{N(i+1)-N(i)}{dN(i)} \Big)^{k-d} \leq \Big(C_k \log N(n)\Big)^{k-d} = o(\sqrt{\log N(n)}),$$

where the constants $C, C_k > 0$ can change along the lines. Hence the statement of Theorem 2.7 follows by combining Lemmas 5.3 and 5.4 by a triangle inequality. □

Remark 5.5 (The order of magnitude of $|V(n)|$)**.** It is elementary to show that the nonnegative martingale

$$M'_n = N(n)\prod_{i=1}^{n-1}(1+dq_i)^{-1} \qquad (5.2)$$

is square integrable if $\sum_{n \in \mathbb{N}} q_n = \infty$ and so there exists a $\xi \geq 0$ random variable

$$\frac{N(n)}{\prod_{i=1}^{n-1}(1+dq_i)} \to \xi \quad a.s.$$

Claim 5.6. *The series p_k given in (2.10) is a probability distribution.*

Proof of Claim 5.6. Clearly $p_k \geq 0$. Combining the formula for p_k in (2.10) with an elementary rewrite of the fraction of the Gamma-functions inside the sum yields

$$\sum_{k=d+1}^{\infty} p_k = \frac{d}{2d+1} \frac{\Gamma\left(1+\frac{2d+1}{d-1}\right)}{\Gamma\left(1+\frac{2}{d-1}\right)} \sum_{k=d+1}^{\infty} \frac{\Gamma\left(k-d+\frac{2}{d-1}\right)}{\Gamma\left(k-d+\frac{2d+1}{d-1}\right)}$$

$$= \frac{d}{2d+1} \frac{\Gamma\left(1+\frac{2d+1}{d-1}\right)}{\Gamma\left(1+\frac{2}{d-1}\right)} \cdot \sum_{k=d+1}^{\infty} \frac{d-1}{d} \left(\frac{\Gamma\left(k-d+\frac{2}{d-1}\right)}{\Gamma\left(k-1-d+\frac{2d+1}{d-1}\right)} - \frac{\Gamma\left(k+1-d+\frac{2}{d-1}\right)}{\Gamma\left(k-d+\frac{2d+1}{d-1}\right)} \right) = 1,$$

since the last sum is telescopic. $\qquad\square$

Now we prove Lemmas 5.3 and 5.4. To do this, the following observations will be useful. By construction, there are $d+1$ active cliques at time $n = 0$. When a new node is born the number of active cliques increases by $d+1-1 = d$, thus at time n there are

$$A(n) = N(n)d + d + 1 \tag{5.3}$$

active cliques given $N(n)$, the number of non-initial nodes at time n. By the same argument as in (4.9), a node with degree $k \geq d+1$ is contained in exactly $A_k = 2 + (k-d)(d-1)$ active cliques.

Proof of Lemma 5.3. We prove the Lemma 5.3 by using the Azuma–Hoeffding inequality in an elaborate way. Let us use the notation $V(n) := \sqrt{N(n)\log N(n)}$, and recall the sigma algebra $\mathcal{G}_n$. We aim to show that there exist a constant $c > 0$ such that

$$\mathbb{P}\left(\max_k \left| N_k(n) - \mathbb{E}\big[N_k(n)|\mathcal{G}_n\big] \right| \geq cV(n) \,\Big|\, \mathcal{G}_n \right) = o(1). \tag{5.4}$$

Taking conditional expectation w.r.t. $N(n)$ in (5.4) immediately gives Lemma 5.3. First note that at time n the maximal degree of any vertex is $N(n) + d - 3$. Thus the left hand side of (5.4) is at most

$$\sum_{k=d+1}^{N(n)+d-3} \mathbb{P}\left(|N_k(n) - \mathbb{E}[N_k(n)|\mathcal{G}_n]| \geq cV(n) \mid \mathcal{G}_n \right).$$

Since there are $N(n) + d - 3$ summables, it is enough to prove that *uniformly in* k with $d+1 \leq k \leq N(n) + d - 3$

$$\mathbb{P}\left(|N_k(n) - \mathbb{E}[N_k(n)|\mathcal{G}_n]| \geq cV(n) \right) = o\left(N(n)^{-1} \right). \tag{5.5}$$

For a fixed time r, let us fix an ordering of the cliques of the graph $\mathrm{EAN}_d(r, \{q_n\})$. Clearly, the number of active cliques $A(r) < (d+1)^r$. To get $\mathrm{EAN}_d(r+1)$ we draw an independent Bernoulli(q_{r+1}) random variable for every clique in $\mathrm{EAN}_d(r, \{q_n\})$. Hence, for $1 \leq r \leq n$ and $0 \leq s \leq A(r)$, it is reasonable to introduce $\mathcal{F}_{r,s}$, the σ-algebra generated by the graph at time $r-1$ and the first s coin flips at time r and $\mathcal{G}_n$. It is straightforward that $\mathcal{G}_n = \mathcal{F}_{1,0} \subseteq \cdots \subseteq \mathcal{F}_{1,d+1} \subseteq \mathcal{F}_{2,0} \subseteq \cdots \subseteq \mathcal{F}_{n,A(n)}$. With this filtration, let us introduce the following Doob-martingale:

$$M_{r,s} = \mathbb{E}\left[N_k(n) \mid \mathcal{F}_{r,s} \right],$$

where k is fixed. Clearly $M_{1,0} = \mathbb{E}\left[N_k(n) \mid \mathcal{G}_n\right]$, and $M_{n,A(n)} = N_k(n)$. Now, we would like to estimate the difference between $M_{r,s}$ and $M_{r,s-1}$. We will see that

$$|M_{r,s} - M_{r,s-1}| \leq 2(d+1) \quad \forall\, r \in \{1, \ldots, n\}, \ \forall\, 1 \leq s < A(r) \leq (d+1)^r. \tag{5.6}$$

From the definition of $M_{r,s}$, we see that the difference is caused by the extra information whether the s-th coin flip raises a new node or not. Let us consider the two different realizations, i.e. in $\mathrm{EAN}(r, s)_a$ there is a new node $v_{r,s}$ at the s-th coin flip and in $\mathrm{EAN}(r, s)_b$ it is not. Note that the number $N(r+1) - N(r)$ of new nodes at time r is *included in the σ algebra* and therefore there must be an s' with $s < s' < A(r)$ that at the s'th coin flip a new node $v_{r,s'}$ will born in $\mathrm{EAN}(r, s)_b$ but not in $\mathrm{EAN}(r, s)_a$. Hence the graphs $\mathrm{EAN}(r+1, 0)_a$ and $\mathrm{EAN}(r+1, 0)_b$ might be coupled in such a way that *the number of nodes are the same* and *every node has the same degree except* for the $d+1$ neighbors of $v_{r,s}$ in $\mathrm{EAN}(r, s)_a$ and the $d+1$ neighbors of $v_{r,s'}$ in $\mathrm{EAN}(r, s)_a$. Since the degree of vertices that are born later than r are not affected by what happens before time r, we can extend this coupling up to time n

such that there are at most $2(d+1)$ nodes with different degrees. Thus, taking expectation with respect to $\mathcal{F}_{r,s-1}$ conserves this difference, which implies (5.6).

We have just proved that $M_{r,s}$ is a martingale with bounded increments. Observe that every new node will create $d+1$ new active cliques and then induce $d+1$ coin flips. Thus there are less than $(d+1)N(n)$ coin flips until time n and so thus $|M_{r,s} - M_{r,s-1}| \neq 0$ only at most $(d+1)N(n)$ times. Hence the Azuma–Hoeffding inequality gives us that

$$\mathbb{P}\left(\left|N_k(n) - \mathbb{E}[N_k(n)|\mathcal{G}_n]\right| \geq a \,\Big|\, \mathcal{G}_n\right) \leq 2\exp\left\{-\frac{a^2}{8N(n)(d+1)^3}\right\}.$$

Now set $a = cV(n)$, $c > \sqrt{8}(d+1)^{3/2}$:

$$\mathbb{P}\left(\left|N_k(n) - \mathbb{E}[N_k(n)|\mathcal{G}_n]\right| \geq cV(n) \mid \mathcal{G}_n\right) \leq 2N(n)^{-\frac{c^2}{24(d+1)^2}} \leq o\left(N(n)^{-1}\right).$$

Note that this bound is *uniform in k*, hence (5.5) and (5.4) follows. □

Proof of Lemma 5.4. We aim to write a recursion for $\mathbb{E}[N_k(n)|\mathcal{G}_n]$. Note that $A(n), N(n+1) - N(n) \in \mathcal{G}_{n+1}$. Let us introduce the nth empirical occupation parameter:

$$\hat{q}_n := \frac{N(n+1) - N(n)}{dN(n) + d + 1} = \frac{1}{A(n)}\sum_{i=1}^{A(n)} \mathbb{1}_{\{\text{the } i\text{th triangle is filled}\}}.$$

Since we assume $N(n) \to \infty$, $\hat{q}_n = q_n(1 + o(1))$ a.s. Given that there are $N(n+1) - N(n)$ successes in $A(n)$ Bernoulli trials, the places of these successful trials are uniformly distributed. Clearly, $N_k(n)$ can change in three different ways in the $n+1$-th step:

(i) A new node can connect to a vertex with degree k. Since a vertex with degree k is contained in A_k active cliques (see (5.3)) the degree stays k from step n to $n+1$ with this happens with $\mathcal{G}_n$-conditonal probability $(1 - \hat{q}_n)^{A_k}$.

(ii) A degree of a vertex can increase to k. A vertex with degree $k-\ell$, $\ell = 1, \ldots, \frac{d-1}{d}k$ is contained in $A_{k-\ell}$ many active cliques, hence this happens with $\mathcal{G}_n$-conditonal probability $\binom{A_{k-\ell}}{\ell}\hat{q}_n^\ell(1 - \hat{q}_n)^{A_{k-\ell}-\ell}$.

(iii) When $k = d+1$, $N_{d+1}(n)$ grows by the number of new nodes $N(n+1) - N(n)$

Hence we can write the following conditional recursion:

$$\mathbb{E}[N_k(n+1)|\mathcal{G}_{n+1}] = \mathbb{E}[N_k(n)|\mathcal{G}_n](1 - \hat{q}_n)^{A_k} + \mathbb{E}[N_{k-1}(n)|\mathcal{G}_n]A_{k-1}\hat{q}_n(1 - \hat{q}_n)^{A_{k-1}-1}$$

$$+ \sum_{\ell=2}^{(d-1)k/d} \mathbb{E}[N_{k-\ell}(n)|\mathcal{G}_n]\binom{A_{k-\ell}}{\ell}\hat{q}_n^\ell(1 - \hat{q}_n)^{A_{k-\ell}-\ell} + \mathbb{1}_{\{k=d+1\}}(N(n+1) - N(n)) \tag{5.7}$$

Now, we first try to find the 'stationary solution' of this recursion in the form $\mathbb{E}[N_k(n)|\mathcal{G}_n] = p_k N(n)$. Then, series expansion in the first term on the right hand side yields that the limiting distribution p_k should satisfy:

$$p_k(N(n+1) - N(n)) = -p_k N(n)A_k\hat{q}_n + p_{k-1}N(n)\hat{q}_n A_{k-1} + O(N(n)\hat{q}_n^2(A_k^2 + A_{k-1}))$$

$$+ \sum_{\ell=2}^{(d-1)k/d} p_{k-\ell}N(n)\hat{q}_n\binom{A_{k-\ell}}{\ell}\hat{q}_n^{\ell-1}(1 + o(1)) + \mathbb{1}_{\{k=d+1\}}(N(n+1) - N(n)) \tag{5.8}$$

Multiply both sides by $(N(n)\hat{q}_n)^{-1}$, and use that $N(n+1) - N(n) = \hat{q}_n A(n)$:

$$p_k\frac{A(n)}{N(n)} = -p_k A_k + p_{k-1}A_{k-1} + \mathbb{1}_{\{k=d+1\}}\frac{A(n)}{N(n)} + O(C_k\hat{q}_n) \tag{5.9}$$

where C_k contains the coefficients of all the smaller order terms. Since $A(n)/N(n) \to d$ as $N(n) \to \infty$, and $\hat{q}_n = q_n(1 + o(1)) \to 0$ as $n \to 0$, the limiting distribution p_k should satisfy

$$p_k(d + A_k) = p_{k-1}A_{k-1} + d\mathbb{1}_{\{k=d+1\}}. \tag{5.10}$$

Using the formula for A_k in (4.9) we equivalently have

$$p_k = p_{k-1}\frac{(k-1)(d-1) - d^2 + d + 2}{k(d-1) - d^2 + 2d + 2} + \mathbb{1}_{\{k=d+1\}}\frac{d}{2d+1}. \tag{5.11}$$

The solution of this recursion is

$$p_k = p_{d+1} \prod_{\ell=d+2}^{k} \frac{\ell - 1 - d + \frac{2}{d-1}}{\ell - d + \frac{d+2}{d-1}} = \frac{d}{2d+1} \frac{\Gamma(k - d + \frac{2}{d-1})}{\Gamma(1 + \frac{2}{d-1})} \frac{\Gamma(2 + \frac{d+2}{d-1})}{\Gamma(k + 1 - d + \frac{d+2}{d-1})}, \quad (5.12)$$

and hence by the properties of Gamma-function we obtain that

$$p_k \sim const \cdot k^{-\frac{2d-1}{d-1}},$$

i.e. the 'stationary solution' has a power-law decay with exponent in $(2,3]$ for $d \geq 2$. Now, the recursion (5.7) initially is not stationary, hence we still need to show that

$$\varepsilon_k(n) := \mathbb{E}[N_k(n)|\mathcal{G}_n] - p_k N(n) \quad (5.13)$$

is tending to zero conditionally on $\mathcal{G}_n$. Using (5.7) and (5.10) it is elementary to check that the following recursion holds for the error terms defined in (5.13):

$$\varepsilon_k(n+1) = \varepsilon_k(n)(1 - \hat{q}_n)^{A_k} + \varepsilon_{k-1}(n)A_{k-1}\hat{q}_n(1 - \hat{q}_n)^{A_{k-1}-1}$$
$$- p_k \left(N(n+1) - N(n)(1 - \hat{q}_n)^{A_k} - (d + A_k)\hat{q}_n N(n) \right)$$
$$- p_{k-1}N(n)\hat{q}_n A_{k-1} \left(1 - (1 - \hat{q}_n)^{A_{k-1}-1} \right) - \mathbb{1}_{\{k=d+1\}}(dN(n) - A(n))\hat{q}_n. \quad (5.14)$$
$$+ \sum_{\ell=2}^{\frac{d-1}{d}k} \mathbb{E}[N_{k-\ell}(n)|\mathcal{G}_n] \binom{A_{k-\ell}}{\ell} \hat{q}_n^\ell (1 - \hat{q}_n)^{A_{k-\ell}-\ell}.$$

Let us denote by $\Delta_k(n) := |\varepsilon_k(n+1) - \varepsilon_k(n)(1 - \hat{q}_n)^{A_k}|$, i.e. the absolute value of the sum of the terms on the right hand side except the first term. Since $(1 - \hat{q}_n)^{A_k} < 1$, we immediately get the upper bound

$$\varepsilon_k(n+1) \leq \varepsilon_k(0) + \sum_{i=1}^{n} \Delta_k(i).$$

With series expansion in the third and fourth term in (5.14) and the identity $dN(n) - A(n) = d + 1$ yields the upper bound

$$\frac{\Delta_k(n)}{N(n)\hat{q}_n} \leq \frac{\varepsilon_{k-1}(n)}{N(n)} A_{k-1} + p_k \left(\frac{d+1}{N(n)} + O(\hat{q}_n^2 A_k^2) \right) + \mathbb{1}_{\{k=d+1\}} \frac{d+1}{N(n)}$$
$$+ p_{k-1}\hat{q}_n A_{k-1}^2 + \sum_{\ell=2}^{\frac{d-1}{d}k} \binom{A_{k-\ell}}{\ell} \hat{q}_n^{\ell-1}, \quad (5.15)$$

where we used that $\mathbb{E}[N_k(n)|\mathcal{G}_n] \leq N(n)$ holds for all n, k. Now clearly we have

$$\Delta_{d+1}(i) = \hat{q}_i(d+1)(1 + p_{d+1}) + o(\hat{q}_i)$$

for any $i \in \mathbb{N}$. Note that $\varepsilon_{d+1}(0) \leq d + 2$ hence for some constants $C_{d+1} < C'_{d+1} < \infty$

$$\varepsilon_{d+1}(n) \leq d + 2 + C_{d+1} \sum_{i=1}^{n-1} \hat{q}_i(1 + o(1)) \leq C'_{d+1} \sum_{i=1}^{n-1} \hat{q}_i.$$

Let us inductively assume that $\varepsilon_{k-1}(i) \leq \varepsilon_{k-1}(0) + C'_{k-1}\left(\sum_{j=1}^{i-1} \hat{q}_j\right)^{k-1-d}$ holds uniformly in i. To carry out the inductive step for k we can sum up (5.15) and use the induction hypothesis to get

$$\varepsilon_k(n) = \varepsilon_k(0) + \sum_{i=1}^{n-1} \hat{q}_i \left(A_{k-1}C'_{k-1}\left(\sum_{j=1}^{i-1} \hat{q}_j\right)^{k-1-d} + c_{k,1} + c_{k,2}\hat{q}_i \right)$$
$$\leq \varepsilon_k(0) + c_{k,3} \sum_{i=1}^{n-1} \hat{q}_i + c_{k,4}\left(\sum_{j \leq i}^{n-1} \hat{q}_i\right)^{k-d} \leq \varepsilon_k(0) + C'_k\left(\sum_{i=1}^{n-1} \hat{q}_i\right)^{k-d}, \quad (5.16)$$

where in the second sum we can use that $\max \hat{q}_i \leq 1$. Hence the induction step is satisfies, so we have just shown that $\varepsilon_k(n) \leq \varepsilon_k(0) + C'_k\left(\sum_{i=1}^{n-1} \hat{q}_i\right)^{k-d}$ holds uniformly in n, which finishes the proof. $\square$

References

[1] M. Albenque and J.-F. Marckert. Some families of increasing planar maps. *Electron. J. Probab.*, 13:no. 56, 1624–1671, 2008.

[2] J. S. Andrade, H. J. Herrmann, R. F. S. Andrade, and L. R. da Silva. Apollonian networks: Simultaneously scale-free, small world, euclidean, space filling, and with matching graphs. *Phys. Rev. Lett.*, 94:018702, 2005.

[3] K. Athreya and P. Ney. *Branching Processes*. Dover Publications, 2004.

[4] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[5] S. Bhamidi, R. van der Hofstad, and G. Hooghiemstra. First passage percolation on random graphs with finite mean degrees. *Ann. Appl. Probab.*, 20(5):1907–1965, 2010.

[6] S. Bhamidi, R. van der Hofstad, and G. Hooghiemstra. First passage percolation on the Erdös-Rényi random graph. *Combinatorics, Probability and Computing*, 20:683–707, 2011.

[7] S. Bhamidi, R. van der Hofstad, and G. Hooghiemstra. Universality for first passage percolation on sparse random graphs. arXiv:1210.6839 [math.PR], 2012.

[8] S. Bhamidi, R. van der Hofstad, and G. Hooghiemstra. Weak disorder asymptotics in the stochastic mean-field model of distance. *Ann. Appl. Probab.*, 22(1):29–69, 2012.

[9] B. Bollobás, C. Borgs, J. Chayes, and O. Riordan. Directed scale-free graphs. In *Proceedings of the fourteenth annual ACM-SIAM symposium on Discrete algorithms, Society for Industrial and Applied Mathematics*, pages 132–139, 2003.

[10] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures and Algorithms*, 31:3–122, 2007.

[11] B. Bollobás and O. Riordan. The diameter of a scale-free random graph. *Combinatorica*, 24:5–34, 2004.

[12] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády. The degree sequence of a scalefree random graph process. *Random Structures Algorithms*, 18:279–290, 2001.

[13] D. Boyd. The sequence of radii of the Apollonian packing. *Mathematics of Computation*, 19:249 – 254, 1982.

[14] N. Broutin and L. Devroye. Large deviations for the weighted height of an extended class of trees. *Algorithmica*, 46:271–297, 2006.

[15] W. Bühler. Generations and degree of relationship in supercritical markov branching processes. *Probability Theory and Related Fields*, 18(2):141–152, 1971.

[16] V. Choi and M. J. Golin. Lopsided trees, i: Analyses. *Algorithmica*, 31:240–290, 2001.

[17] C. Cooper and A. Frieze. The height of random k-trees and related branching processes. arXiv:1309.4342 [math.CO], 2013.

[18] A. Darrasse and M. Soria. Degree distribution of ran structures and boltzmann generation fo trees. In *Conference on Analysis of Algorithms, AofA 07, DMTCS Proceedings*, pages 1–14, 2007.

[19] J. Doye and C. Massen. Self-similar disk packings as model spatial scale-free networks. *Phys. Rev. E*, 71:016128, 2005.

[20] E. Ebrahimzadeh, L. Farczadi, P. Gao, A. Mehrabian, C. M. Sato, N. Wormald, and J. Zung. On the longest paths and the diameter in random Apollonian networks. arXiv:1303.5213 [math.CO], 2013.

[21] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume 1. Wiley, 1968.

[22] A. Frieze and C. E. Tsourakakis. On certain properties of random Apollonian networks. In *Proceedings of the 9th international conference on Algorithms and Models for the Web Graph*, WAW'12, pages 93–112. Springer-Verlag, 2012.

[23] R. L. Graham, J. C. Lagarias, C. L. Mallows, A. R. Wilks, and C. H. Yan. Apollonian circle packings: number theory. *Journal of Number Theory*, 100(1):1 – 45, 2003.

[24] R. v. d. Hofstad. *Random Graphs and Complex Networks*. 2013.

[25] S. Janson. One, two and three times log n/n for paths in a complete graph with random weights. *Combinatorics, Probability and Computing*, 8(4):347–361, 1999.

[26] S. Kapoor and E. M. Reingold. Optimum lopsided binary trees. *J. ACM*, 36:573–590, 1989.

[27] I. Kolossváry and J. Komjáthy. First passage percolation on inhomogeneous random graphs. arXiv:1201.3137 [math.PR], 2013.

[28] Z. Zhang, L. Chen, S. Zhou, L. Fang, J. Guan, and T. Zou. Analytical solution of average path length for Apollonian networks. *Phys. Rev. E*, 77:017102, 2008.

[29] Z. Zhang, F. Comellas, G. Fertin, and L. Rong. High-dimensional Apollonian networks. *Journal of Physics A: Mathematical and General*, 39(8):1811, 2006.

[30] Z. Zhang, L. Rong, and F. Comellas. High-dimensional random Apollonian networks. *Physica A: Statistical Mechanics and its Applications*, 364:610 – 618, 2006.

[31] Z. Zhang, L. Rong, and S. Zhou. Evolving Apollonian networks with small-world scale-free topologies. *Phys. Rev. E*, 74:046105, 2006.

[32] T. Zhou, G. Yan, and B.-H. Wang. Maximal planar networks with large clustering coefficient and power-law degree distribution. *Phys. Rev. E*, 71:046141, 2005.

[1]Budapest University of Technology and Economics, Inter-University Centre for Telecommunications and Informatics 4028 Debrecen, Kassai út 26.

[2]Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands.

E-mail address: `istvanko@math.bme.hu, j.komjathy@tue.nl, vagolala@math.bme.hu`