Adv. Appl. Prob. **48,** 865–902 (2016)
*doi:*10.1017/apr.2016.32
© *Applied Probability Trust* 2016

DEGREES AND DISTANCES IN RANDOM AND EVOLVING APOLLONIAN NETWORKS

ISTVÁN KOLOSSVÁRY,* ** *Budapest University of Technology and Economics and Inter-University Centre for Telecommunications and Informatics*
JÚLIA KOMJÁTHY,*** *Budapest University of Technology and Economics, Inter-University Centre for Telecommunications and Informatics and Eindhoven University of Technology*
LAJOS VÁGÓ,* **** *Budapest University of Technology and Economics and Inter-University Centre for Telecommunications and Informatics*

Abstract

In this paper we study random Apollonian networks (RANs) and evolving Apollonian networks (EANs), in d dimensions for any $d \geq 2$, i.e. dynamically evolving random d-dimensional simplices, looked at 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 tends to 0 but is not summable in n. This result gives a rigorous proof for the conjecture of Zhang *et al.* (2006) that EANs tend to exhibit similar behaviour as RANs once the occupation parameter tends to 0. We also determine the asymptotic behaviour of the shortest paths in RANs and EANs for any $d \geq 2$. For RANs we show that the shortest path between two vertices chosen u.a.r. (typical distance), the flooding time of a vertex chosen uniformly at random, and the diameter of the graph after n steps all scale as a constant multiplied by $\log n$. We determine the constants for all three cases and prove a central limit theorem for the typical distances. We prove a similar central limit theorem for typical distances in EANs.

Keywords: Random graph; random network; typical distance; diameter; hopcount; degree distribution

2010 Mathematics Subject Classification: Primary 05C80
Secondary 05C82; 05C12; 90B15; 60J80

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 [14], [25].

Received 6 November 2014; revision received 21 August 2015.
* Postal address: Budapest University of Technology and Economics, Inter-University Centre for Telecommunications and Informatics, 4028 Debrecen, Kassai út 26, Hungary.
** Email address: istvanko@math.bme.hu
*** Postal address: Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands. Email address: j.komjathy@tue.nl
**** Email address: vagolala@math.bme.hu

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

An Apollonian network (AN) is the resulting graph if we place a vertex in the centre of each circle and connect two vertices if and only if the corresponding circles are tangent. This model was introduced independently by Andrade *et al.* [2] and Doye and Massen [21] as a model for real-life networks such as the network of internet cables or links, collaboration networks or protein interaction networks. Apollonian networks can serve as a model for these networks since their main characteristic properties can be observed in the examples above: a power-law degree distribution, a high clustering coefficient, and small distances, usually referred to as the small-world property. Moreover, by construction, ANs also exhibit a hierarchical structure: a property that is very commonly observed in, e.g. social networks.

It is straightforward to generalise Apollonian packings to arbitrary $d \geq 2$ 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 [31] and [34]).

The network arising by this construction is deterministic. Zhou *et al.* [35] proposed to randomise the dynamics of the model such that in one step only one interstice is picked uniformly at random (u.a.r.) and filled with a new circle. This construction yields a d-dimensional *random Apollonian network* (RAN) [32]. Using heuristic and rigorous arguments the results in [1], [18], [19], [22], [24], [32], and [35] show that RANs have the above mentioned main features of real-life networks.

A different random version of the original AN was introduced by Zhang *et al.* [33], called *evolutionary Apollonian networks* (EANs) 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. We call q the *occupation parameter*. For $q = 1$ we get back the deterministic AN model. It was conjectured in [33] that an EAN with parameter q, as $q \to 0$, should show a similar topological behaviour to RANs. To make this statement rigorous, instead of looking at a sequence of evolving EANs with decreasing parameters, we slightly modify the model and investigate the asymptotic behaviour of a single EAN when q might differ in each step of the dynamics. That is, we consider a series $\{q_n\}_{n=1}^{\infty}$ of occupation parameters so that q_n applies for step n of the dynamics, and assume that q_n tends to 0. In this setting, the interesting question is how to determine the correct rate for q_n that achieves the observation that an EAN exhibits similar behaviour as a RAN when the parameter tends to 0.

Definition of RANs. A random Apollonian network in d dimensions ($\mathrm{RAN}_d(n)$) can be constructed as follows. The graph at step $n = 0$ consists of $d + 2$ vertices, embedded in $\mathbb{R}^d$ in such a way that $d + 1$ of them form a d-dimensional simplex, and the $(d + 2)$th vertex is located in the interior of this simplex, connected to all of the vertices of the simplex. This vertex in the interior forms $d + 1$ d-simplices with the other vertices: initially we set the status of these d-simplices as 'active', and call them *active cliques*. For $n \geq 1$, pick an active clique C_n of $\mathrm{RAN}_d(n - 1)$ u.a.r., insert a new vertex 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. Set the status of the clique C_n as 'inactive', and the status of the newly formed d-simplices as '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 u.a.r., insert a vertex and connect it with the vertices of the chosen triangle (face). 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

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

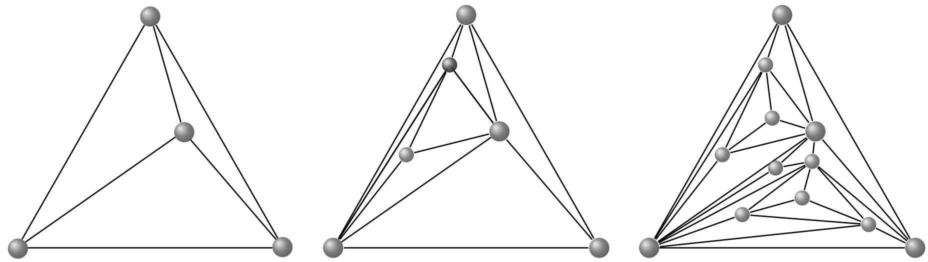

FIGURE 1: A $\mathrm{RAN}_2(n)$ after $n = 0,\ 2,\ 8$ steps.

stack-triangulations. Stack-triangulations were investigated in [1] where the authors also considered typical properties under different weighted measures, e.g. ones picked u.a.r. having n faces. Under a certain measure stack-triangulations with n faces are an equivalent formulation of $\mathrm{RAN}_2(n)$; see [1] and the references therein.

Definition of EANs. Given a sequence of occupation parameters $\{q_n\}_{n=1}^{\infty}, 0 \le q_n \le 1$, an evolutionary Apollonian network in d dimensions ($\mathrm{EAN}_d(n, \{q_n\}) = \mathrm{EAN}_d(n)$) 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 cliques. For $n \ge 1$, choose 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 nonpicked active cliques stay active) and for every clique $C \in \mathcal{C}_n$ we place a new vertex $v_n(C)$ in the interior of C that we connect to all vertices of C. This new vertex $v_n(C)$ 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 an $\mathrm{EAN}_d(n)$. The $q_n \equiv q$ case was studied in [33] 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 $\sum_{n=0}^{\infty} q_n = \infty$.

Remark 1. Note that both in the RAN and EAN models, there is a one-to-one correspondence between cliques and vertices/future vertices: vertex v corresponds to the clique C that became inactive when v was placed in the interior of the d-simplex corresponding to C. In this respect, we call vertices that are already present in the graph *inactive vertices*, and we refer to active cliques as *active vertices*: this notation means that these vertices are not yet present in the graph, but might become present in the next step of the dynamics.

Structure of the paper. In Section 2 we state our main results and discuss their relation to other results in the field. Section 3 contains the most important combinatorial 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 provide 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

We begin with the necessary notation used to state our results.

Fix n and consider two 'active' or 'inactive' vertices u and v from $\mathrm{RAN}_d(n)$ or $\mathrm{EAN}_d(n)$. Denote by $\mathrm{hop}_d(n, u, v)$ the hopcount between the vertices u and v, i.e. the number of edges

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

on (one of) the shortest paths between u and v. The flooding time $\text{flood}_d(n, u)$ is the maximal hopcount from u, while the diameter $\text{diam}_d(n)$ is the maximal flooding time, formally

$$\text{flood}_d(n, u) = \max_v \text{hop}_d(n, u, v) \quad \text{and} \quad \text{diam}_d(n) = \max_{u,v} \text{hop}_d(n, u, v).$$

Whenever possible d, u, and v are suppressed from the notation.

We define $D_v(n)$ as the degree of vertex v after the nth step. Denote by $\widetilde{N}_k(n)$ and $\widetilde{p}_k(n)$ the number and the empirical proportion of inactive vertices with degree k at time n, respectively, in a $\text{RAN}_d(n)$, i.e.

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

where $\mathbf{1}$ is the indicator function.

Analogously, for the graph $\text{EAN}_d(n, \{q_n\})$ we use the notations $N_k(n)$ and $p_k(n)$ defined by

$$p_k(n) := \frac{N_k(n)}{N(n)} = \frac{1}{N(n)} \sum_{i \in V(n)} \mathbf{1}_{\{D_i(n)=k\}}, \tag{2}$$

where $V(n)$ denotes the set of vertices after n steps and $N(n) = |V(n)|$.

Let $(X_i)_{i=1}^{d+1}$ be a collection of independent geometrically distributed random variables with success probability $i/(d+1)$ for X_i. Define the sum

$$Y_d := \sum_{i=1}^{d+1} X_i, \tag{3}$$

where Y_d is commonly referred to as a *full coupon collector block* in a coupon collector problem with $d+1$ coupons. Denote the expectation and variance of Y_d by

$$\mu_d := \mathbb{E}[Y_d] = (d+1)H(d+1), \qquad \sigma_d^2 := \mathbb{D}^2[Y_d], \tag{4}$$

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

$$I_d(x) := \sup_{\lambda \in \mathbb{R}} \{\lambda x - \log(\mathbb{E}[\exp\{\lambda Y_d\}])\}. \tag{5}$$

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

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

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

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

The following is needed for the flooding time and diameter. Consider the function

$$f_d(c) := c - \frac{d+1}{d} - c \log\biggl(\frac{d}{d+1} c\biggr). \tag{6}$$

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

Note that $-f_d(c)$ is the rate function of a $\mathrm{Poi}((d+1)/d)$ random variable. Thus, for $c > (d+1)/d$ the equation $f_d(c) = -1$ has a unique solution which we denote by $\widetilde{c}_d$. Finally, we introduce

$$g(\alpha, \beta) := 1 + f_d(\alpha \tilde{c}_d) - \alpha\beta \frac{\tilde{c}_d}{\mu_d} I_d\left(\frac{\mu_d}{\beta}\right). \tag{7}$$

We say that a sequence of events $\mathcal{E}_n$ happens *with high probability* (w.h.p.) if $\lim_n \mathbb{P}(\mathcal{E}_n) = 1$. Note that 'with high probability' is the same as 'asymptotically almost surely'. We further define for an event A and a σ-algebra $\mathcal{F}$ the conditional probability $\mathbb{P}(A \mid \mathcal{F}) = \mathbb{E}[\mathbf{1}_A \mid \mathcal{F}]$, where $\mathbf{1}_A$ is the indicator of the event A, i.e. it takes value 1 if A holds and 0 otherwise. We will sometimes replace $\mathcal{F}$ by a list of random variables, in this case we drop the σ-algebra notation and list the random variables in the conditioning, and this means conditional on the σ-algebra generated by this list of random variables.

2.1. Distances in RANs and EANs

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

Theorem 1. (Typical distances in RANs.) *With high probability, the hopcount between two active vertices chosen u.a.r. in a* $\mathrm{RAN}_d(n)$ *satisfies a central limit theorem (CLT) of the form*

$$\left(\mathrm{hop}_d(n) - \frac{2}{\mu_d}\frac{d+1}{d}\log n\right)\left(2\frac{\sigma_d^2+\mu_d}{\mu_d^3}\frac{d+1}{d}\log n\right)^{-1/2} \xrightarrow{\mathrm{D}} Z, \tag{8}$$

where μ_d, σ_d^2 as in (4), Z is a standard normal random variable, and '$\xrightarrow{\mathrm{D}}$' denotes convergence in distribution.

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

$$\mathbb{P}(v \text{ is chosen } \mid D_v(n)) = \frac{(d-1)D_v(n) - d^2 + d + 2}{dn + d + 1}. \tag{9}$$

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

Theorem 2. (Diameter and flooding time in RANs.) *Let u denote either an active vertex chosen u.a.r. or an inactive vertex chosen according to the size-biased distribution given in (9). Then as $n \to \infty$, w.h.p.*

$$\frac{\mathrm{diam}_d(n)}{\log n} \xrightarrow{\mathbb{P}} 2\tilde{\alpha}\tilde{\beta}\frac{\tilde{c}_d}{\mu_d}, \qquad \frac{\mathrm{flood}_d(n,u)}{\log n} \xrightarrow{\mathbb{P}} \frac{1}{\mu_d}\left(\frac{d+1}{d} + \tilde{\alpha}\tilde{\beta}\tilde{c}_d\right),$$

where '$\xrightarrow{\mathbb{P}}$' denotes convergence in probability and $(\tilde{\alpha}, \tilde{\beta}) \in (0,1] \times [1, \mu_d/(d+1)]$ is the optimal solution of the maximization problem with the following constraint:

$$\max\{\alpha\beta \colon g(\alpha, \beta) = 0\}. \tag{10}$$

Remark 2. Observe that the set of (α, β) pairs that satisfy the constraint in (10) is nonempty since, for $\alpha = \beta = 1$, by definition $f(\tilde{c}_d) = -1$ and $I_d(\mu_d) = 0$. The fact that the pair $(\tilde{\alpha}, \tilde{\beta})$ is unique is proved in Lemma 5. The maximization problem is also equivalent to first defining the $g(\alpha(x), \beta(x)) := \sup_{\alpha,\beta}\{g(\alpha, \beta) \colon \alpha\beta = x\}$ and then choosing the unique x with $g(\alpha(x), \beta(x)) = 0$, where the existence and uniqueness of such an x follows from the fact that $g(\alpha(x), \beta(x))$ is strictly monotone decreases in x and is continuous. This is shown in Claim 4.

We conclude with the asymptotic behaviour of the typical distances in $\mathrm{EAN}_d(n)$.

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

Theorem 3. (Typical distances in EANs.) *Assume that 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 w.h.p., the hopcount between two active vertices chosen u.a.r. in an* $\mathrm{EAN}_d(n)$ *satisfies a CLT of the form*

$$\biggl(\mathrm{hop}_d(n) - \frac{2}{\mu_d}\sum_{i=1}^{n} q_i\biggr)\biggl(2\frac{\sigma_d^2+\mu_d}{\mu_d^3}\sum_{i=1}^{n} q_i(1-q_i)\biggr)^{-1/2} \xrightarrow{\mathrm{D}} Z,$$

where μ_d, σ_d^2 *as in (4) and* Z *is a standard normal random variable.*

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

$$\mathbb{P}(v \text{ is chosen} \mid D_v(n), N(n)) = \frac{(d-1)D_v(n) - d^2 + d + 2}{d(N(n)-d-2)+d+1}. \tag{11}$$

Remark 3. Note that in this theorem q_n might or might not tend to 0. The second criterion rules out the case when $q_n \to 1$ so fast that the graph becomes essentially deterministic. Further, the statements of Theorems 1 and 3 also hold if one of the vertices is an active vertex chosen uniformly at random and the other vertex is inactive chosen according to the distribution given in (9) and (11), respectively.

2.2. Degree distribution and clustering coefficient

Our first result states that for a $\mathrm{RAN}_d(n)$ the empirical distribution $\widetilde{p}_k(n)$ tends to a proper distribution in the ℓ_∞-metric.

Theorem 4. (Degree distribution for RANs.) *For all* $d \geq 2$ *there exist a probability distribution* $\{p_k\}_{k=d+1}^{\infty}$ *and a constant* c *for which*

$$\lim_{n\to\infty} \mathbb{P}\biggl(\max_k |\widetilde{p}_k(n) - p_k| \geq c\sqrt{\frac{\log n}{n}}\biggr) = 0.$$

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

$$\begin{aligned} p_k &= \frac{d}{2d+1}\frac{\Gamma(k-d+2/(d-1))}{\Gamma(1+2/(d-1))}\frac{\Gamma(2+(d+2)/(d-1))}{\Gamma(k+1-d+(d+2)/(d-1))} \\ &= C(d)k^{-(2d-1)/(d-1)}(1+o_k(1)), \end{aligned} \tag{12}$$

where $o_k(1)$ *denotes a quantity that tends to 0 as* $k \to \infty$, $C(d)$ *is a constant that depends on* d, *and* $\Gamma(x)$ *is the gamma function.*

Remark 4. To obtain the asymptotic behaviour of p_k above we use the property that

$$\frac{\Gamma(t+a)}{\Gamma(t)} = t^a(1+o(1)).$$

For the theorem that describes the degree distribution of the graph $\mathrm{EAN}_d(n, \{q_n\})$, we need the following additional analytic assumption on the sequence $\{q_n\}_{n\in\mathbb{N}}$.

Assumption 1. *Assume, as before, that* $q_n \to 0$ *and* $\sum_{i=1}^{\infty} q_n \to \infty$. *We assume further that there exist constants* c_1 *and* C_1 *(that depend only on the sequence* $\{q_n\}_{n=1}^{\infty}$*) such that*

$$c_1 \leq \frac{\sum_{i=1}^{n} q_i^2 \prod_{j=1}^{i}(1+dq_j)}{q_n \prod_{j=1}^{n}(1+dq_j)} \leq C_1, \tag{13}$$

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

and for every $\varepsilon > 0$*,*

$$\sum_{n=1}^{\infty} \exp\left\{-\varepsilon q_n \exp\left\{d \sum_{j=0}^{n} q_j\right\}\right\} < \infty. \tag{14}$$

Theorem 5. (Degree distribution for EANs.) *Let* $d \geq 2$ *and* $\{q_n\}_{n=0}^{\infty}$ *be probabilities such that Assumption 1 is satisfied. Then the degree distribution tends to the same asymptotic degree distribution* $\{p_k\}_{k=d+1}^{\infty}$ *as in the case of* $\mathrm{RAN}_d(n)$ *given in (12). More precisely, there exists a constant* $C_0 > 0$ *and a random variable* $\eta < \infty$ *such that for any* $k \in \mathbb{N}$

$$\lim_{n\to\infty} \mathbb{P}\Big(\max_{i<k} |p_i(n) - p_i| \geq C_0 k!\, \eta^k q_n\Big) = 0.$$

In particular, the degree distribution converges pointwise for all $i < k = k(n)$ *for any* $k(n)$ *that satisfies* $C_0 k!\, \eta^k q_n \to 0$ *as* $n \to \infty$*.*

In the following lemma we present classes of sequences $\{q_n\}_{n\in\mathbb{N}}$ that satisfy Assumption 1.

Lemma 1. (Regularly varying sequences.) *Let* $L(x)$ *denote a slowly varying function at* ∞*, i.e. for every* $c > 0$*,* $\lim_{x\to\infty} L(cx)/L(x) = 1$*. Then Assumption 1 is satisfied in the following cases:*

(i) $q_n = L(n)/n^{\alpha}$ *for some* $\alpha \in (0, 1)$*;*

(ii) $q_n = (b + o(n^{-\delta}))/n$ *with* $b > 1/d$ *for some arbitrary* $\delta > 0$*.*

Remark 5. Clearly, Lemma 1(ii) does not cover all possible regularly varying functions with $\alpha = 1$ for which Assumption 1 holds: one can check that the assumption holds for other cases as well, e.g. $q_n = b + b'/\log n$ with any $b > 1/d$ and arbitrary $b' \in \mathbb{R}$ works. On the other hand, some cases where $q_n = L(n)/n$ and $L(n) \to \infty$ fail, e.g. $q_n = b/(n \log n)$ does not satisfy (13).

The proof of these theorems and Lemma 1 are given in Sections 5.1 and 5.2. Next we describe the clustering coefficient of RANs and EANs. The clustering coefficient of a vertex is the proportion of the number of existing edges between its neighbours, compared to the number of all possible edges between them. Here we investigate the clustering coefficient of the whole graph, which is the average of clustering coefficients over the vertices. 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 [32, 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* both in RANs and EANs and equals

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

The explanation for this formula is as follows. When the degree of vertex v increases by 1 by adding a new vertex w in one of the active cliques containing v, then the number of edges between the neighbours of v increases exactly by d, since the newly added vertex w is connected to the other d vertices in the clique. It was observed in simulations and heuristicly proved in [32], that the average clustering coefficient of these networks converges to a strictly positive constant. Our next corollary determines the exact value of these constants for the two models.

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

Corollary 1. (Clustering coefficient.) *The average clustering coefficient of the* $\mathrm{RAN}_d(n)$ *converges to a strictly positive constant as* $n \to \infty$*, given by*

$$\begin{aligned} Cl_d &= \sum_{k=d+1}^{\infty} \frac{d(2k-d-1)}{k(k-1)} p_k \\ &= \sum_{k=d+1}^{\infty} \frac{d(2k-d-1)}{k(k-1)} \frac{d}{2d+1} \frac{\Gamma(k-d+2/(d-1))}{\Gamma(1+2/(d-1))} \frac{\Gamma(1+(2d+1)/(d-1))}{\Gamma(k-d+(2d+1)/(d-1))}. \end{aligned} \tag{15}$$

Further, the clustering coefficient of the $\mathrm{EAN}_d(n, \{q_n\})$ *converges to the same value as in (15) if* $q_n \to 0$ *and* $\sum_{n\in\mathbb{N}} q_n = \infty$.

2.3. Related literature

Several results related to the asymptotic degree distribution of ANs are known. It is not difficult to see that if a vertex belongs to k active cliques, then the chance for that vertex to get a new edge is proportional to k: this argument shows that these models belong to the wide class of *preferential attachment models* [4], [10], [13]. As a result, some of the classical methods can be adapted to this model.

Using heuristic arguments, Zhang *et al.* [32] obtained that the asymptotic degree exponent should be $(2d-1)/(d-1) \in (2, 3]$, which is in good agreement with their simulations. Parallel to writing this paper, we noted that Frieze and Tsourakakis [24] very recently derived rigorously the exact asymptotic degree distribution of the two-dimensional RAN. Even though our work is independent of 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 [13] and [30]. So, to avoid repetition we decided to only sketch some parts of the proof and include the part that does not overlap with their work, i.e. the degree distribution of EANs.

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 *et al.* [33] studied the graph series $\mathrm{EAN}_d(n)$ with (fixed) occupation parameter q. They derived the asymptotic degree exponent using heuristic arguments, and the result is in good agreement with the simulations. They also suggested that as $q \to 0$ the model $\mathrm{EAN}_d(n)$ converges to a $\mathrm{RAN}_d(n)$ in some sense. We confirm their claim by deriving the asymptotic degree distribution of the $\mathrm{EAN}_d(n, \{q_n\})$ 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 his co-authors can be made precise in this way.

The statements of Theorem 1 are in agreement with previous results. In particular, in [32] the authors estimated the average path length, i.e. the hopcount averaged over all pairs of vertices, and they showed that it scales logarithmically with the size of the network.

A more refined claim was obtained by Albenque and Marckert [1] concerning the hopcount in two dimensions. They proved that

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

The constant $\frac{6}{11}$ 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 1. We present an alternative approach using weaker results. The CLT for distances in EANs is novel.

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

Central limit theorems of the form (8) 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 [26] analysed distances in the complete graphs with independent and identically distributed (i.i.d.) exponential edge weights. In a series of papers Bhamidi, van der Hofstad, and Hooghiemstra determined typical distances and proved the CLT for the hopcount for the Erdős–Rényi random graph [7], the stochastic mean-field model [5], the configuration model with finite variance degrees [6], and quite recently for the configuration model [8] with arbitrary i.i.d. edge weights from a continuous distribution. Inhomogeneous random graphs were handled by Bollobás *et al.* [11], and in [29]. 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 a CLT for the hopcount on the graph. The diameter and flooding time of EANs remains a future project.

Further, in the literature some bounds are known about the diameter of RANs: Frieze and Tsourakakis [24] established the upper bound $2\tilde{c}_2 \log n$ for $\mathrm{RAN}_2(n)$. They used a result of Broutin and Devroye [15] 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 [22] the authors applied the result of [15] in an elaborate way, while Cooper and Frieze in [18] used a more analytical approach solving recurrence relations. We emphasise that the methods in [18, 22], and in this present paper are all qualitatively different. Numerical solution of the maximization problem (10) 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 $2\widetilde{c_2}/\mu_2 \times 0.8639 \times 1.5 = 1.668$, which perfectly coincides with the one obtained in [18] and [22]. To the best of the authors' knowledge no result has been proven for the flooding time.

3. Structure of RANs and EANs

3.1. Tree-like 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 $\Sigma_d := \{1, 2, \dots, d+1\}$ 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 $\boldsymbol{O}$ (root). Then, we code each initial active clique by a different symbol from Σ_d. In step $n = 1$ we assign the newly added vertex u the code $\boldsymbol{u} = i$ if the clique with code i was chosen. Further, we code the $d+1$ newly formed d-simplices by ij for all $j \in \Sigma_d$ (here, ij means concatenation). Similarly for $n \geq 2$, we assign the newly added vertex v the *code of the clique that becomes inactive*, denoted by $\boldsymbol{v}$, and the newly formed active cliques are given the codes $\boldsymbol{v}j$, $j = 1, \dots, d+1$ (with concatenation again). It is crucial to keep the coding *consistent* in a geometrical sense. We describe how to do this in Lemma 2 below.

Thus, each vertex in the graph has a code that is a concatenation of symbols from Σ_d. For a vertex u we write $\boldsymbol{u} = u_1 u_2 \dots u_\ell$ for its code for some $\ell \in \mathbb{N}$, and we call the length of a code $|\boldsymbol{u}| = \ell$ the *generation* of the vertex u. For two vertices u and v with codes $\boldsymbol{u} = u_1 u_2 \cdots u_n$ and $\boldsymbol{v} = v_1 v_2 \cdots v_m$, respectively, we say that u is an *ancestor* of v if $n < m$ and $u_1 u_2 \cdots u_n = v_1 v_2 \cdots v_n$. We denote the latest common ancestor of u and v by $u \wedge v$ and its code

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

by $\boldsymbol{u} \wedge \boldsymbol{v}$; thus, $|\boldsymbol{u} \wedge \boldsymbol{v}| = \min\{k : u_{k+1} \neq v_{k+1}\}$. For codes $\boldsymbol{u} = u_1 \dots u_n$ and $\boldsymbol{v} = v_1 \dots v_m$, we denote the concatenation $u_1 \dots u_n v_1 \dots v_m$ by $\boldsymbol{uv}$ and the corresponding vertex by uv. Further, let $u_{(i)}$ denote the *last occurrence* of the symbol $i \in \Sigma_d$ in $\boldsymbol{u}$. We introduce the cut-operators $T_i\boldsymbol{u} := u_1 \dots u_{(i)-1}$ and $P_i\boldsymbol{u} = u_{(i)} \dots u_n$ for all $i \in \Sigma_d$.

Remark 6. Note that there is a one-to-one correspondence between the codes of length at most n and vertices of a rooted $(d+1)$-ary tree of depth n. As a result, we use the codes $\boldsymbol{u}$ to denote vertices as well, i.e. we identify vertices in a RAN or EAN with their codes and sometimes refer to $\boldsymbol{u}$ as a vertex. In this respect, the concept 'u is an ancestor of v' precisely means the 'usual' notion of being an ancestor: the unique path from v to the root in the $(d+1)$-ary tree passes through u.

Apart from these 'tree' edges, RANs and EANs have other edges as well. However, we will see below that these extra edges always go upwards (or downwards) on a branch of a tree, hence the crucial tree-like properties of the structure are conserved. We collect the most important combinatorial observations in the following lemma.

Lemma 2. (Tree-like properties of the coding.) *There exists a way of choosing the coding of the vertices of a RAN or EAN so that the following hold.*

(i) *Consistency. The $d+1$ neighbours of a newly formed vertex with code $\boldsymbol{u}$ have codes $T_i\boldsymbol{u}$, $i \in \Sigma_d$. Further, for any edge with endpoints $\boldsymbol{u}$ and $\boldsymbol{v}$ either '$\boldsymbol{u}$ is an ancestor of $\boldsymbol{v}$' or vice versa.*

(ii) *Any shortest path between two vertices with codes $\boldsymbol{u}$ and $\boldsymbol{v}$ must intersect the path from the root to the vertex with code $\boldsymbol{u} \wedge \boldsymbol{v}$.*

(iii) *For any two vertices with codes $\boldsymbol{u}$ and $\boldsymbol{v}$,* $\operatorname{hop}(\boldsymbol{u}, \boldsymbol{v}) = \operatorname{hop}(\boldsymbol{u}, \boldsymbol{u} \wedge \boldsymbol{v}) + \operatorname{hop}(\boldsymbol{v}, \boldsymbol{u} \wedge \boldsymbol{v})$.

Before stating the proof, let us interpret Lemma 2. Lemma 2(i) means that edges are only present between vertices along the same ancestral line. In particular, the first $d+1$ neighbours of a newly added vertex with code $\boldsymbol{u}$ can be determined by removing the last pieces of the code of $\boldsymbol{u}$, up to the last occurrence of a given symbol $i \in \Sigma_d$.

The coding 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 $\boldsymbol{u}$ and $\boldsymbol{u}j$ for $j \in \Sigma_d$. All other edges are called *shortcut edges*. So in a RAN at each step one new forward edge and d shortcut edges are formed.

In Figure 2 we show an example in two dimensions. Suppose at step $n = 0$ the 'left', 'right', and 'bottom' triangles are given the symbols $1, 2$, and 3, respectively. Then later each new vertex u with code $\boldsymbol{u}$ in the middle of a triangle gives rise to the new 'left', 'right', and 'bottom' triangles: to these we *have to* assign the codes $\boldsymbol{v}1$, $\boldsymbol{v}2$, and $\boldsymbol{v}3$, respectively.

In Figure 2(a) a planar embedding of the graph is shown, while in Figure 2(b) the tree-like structure of the same graph becomes more apparent. 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.

Lemma 2(ii) is a consequence of Lemma 2(i). It says that if we have two vertices with code $\boldsymbol{u}$ and $\boldsymbol{v}$ in the tree, then any shortest path between them must intersect a path from the root to their latest common ancestor $\boldsymbol{u} \wedge \boldsymbol{v}$. Finally, Lemma 2(iii) says that the distance between any two vertices can be obtained in essentially the same manner as one would do for a tree: it ensures that a shortest path between a vertex with code $\boldsymbol{u}$ and one of its descendants $\boldsymbol{uw}$ cannot

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

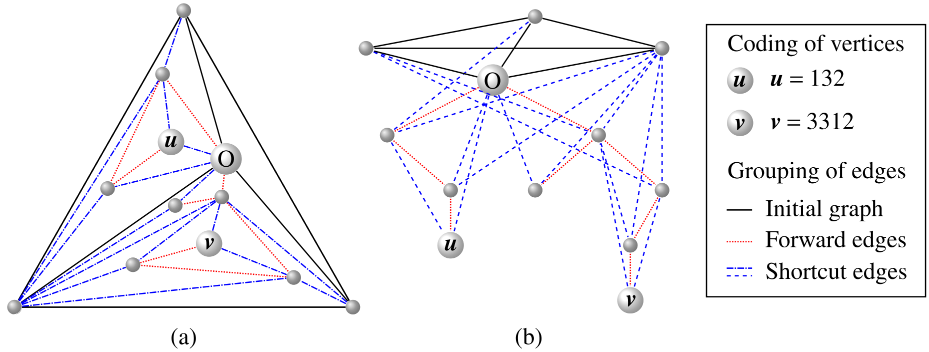

FIGURE 2: Tree-like structure of a realisation of $\mathrm{RAN}_2(8)$. (a) Shows a planar embedding of the graph, and (b) the tree-like structure of the same graph.

go *below* the vertex $\boldsymbol{uw}$ in the tree, and also that if two vertices u and v are not descendants of each other, then a shortest path between them does not go below them in the tree.

Proof of Lemma 2. (i) We use proof by induction. We label the initial $d+1$ active cliques arbitrarily with $i \in \Sigma_d$. The hypotheses clearly holds in this case. In this proof, to identify a clique, we list the code of its vertices, e.g. $C = (\boldsymbol{v}^{(1)}, \dots, \boldsymbol{v}^{(d+1)})$ denotes a clique that is formed by vertices with codes $\boldsymbol{v}^{(1)}, \dots, \boldsymbol{v}^{(d+1)}$. Recall that each clique also uniquely corresponds to the vertex that is placed in the interior of the corresponding d-simplex when the clique becomes inactive. We say that a clique C has code $\boldsymbol{u}$ if this vertex has code $\boldsymbol{u}$. In this notation, we need to find a consistent coding, i.e. one where the clique $(\boldsymbol{v}^{(1)}, \dots, \boldsymbol{v}^{(d+1)})$ gets a code $\boldsymbol{u}$ such that each $\boldsymbol{v}^{(j)}$ can be obtained by removing the code $\boldsymbol{u}$ at the last occurrence of a symbol in Σ_d.

Now suppose that we already have an active clique with code $\boldsymbol{u}$, and, by induction, we can assume that $\boldsymbol{u}$ is associated with the clique formed by vertices $(T_1\boldsymbol{u}, T_2\boldsymbol{u}, \dots, T_{d+1}\boldsymbol{u})$. When this clique becomes inactive, the vertex $\boldsymbol{u}$ is added to the graph and the $d+1$ new active cliques are

$$(T_1\boldsymbol{u}, \dots, T_{k-1}\boldsymbol{u}, T_{k+1}\boldsymbol{u}, \dots, T_{d+1}\boldsymbol{u}, \boldsymbol{u}) \quad \text{for all } k \in \Sigma_d. \tag{16}$$

Let us denote this clique by $\boldsymbol{u}k$ (where $\boldsymbol{u}k$ means concatenating the symbol k to the end of the code $\boldsymbol{u}$). We claim that by this choice, (i) is maintained, i.e. if the vertex $\boldsymbol{u}k$ is ever to be added to the graph, then its neighbours will be exactly $T_i(\boldsymbol{u}k)$, $i \in \Sigma_d$. By construction, the neighbours of $\boldsymbol{u}k$ are exactly the ones in (16), and clearly we have $T_k(\boldsymbol{u}k) = \boldsymbol{u}$, and $T_i(\boldsymbol{u}k) = T_i(\boldsymbol{u})$ for $i \neq k$, so we can write

$$\begin{aligned}(T_1\boldsymbol{u}, \dots, T_{k-1}\boldsymbol{u}, T_{k+1}\boldsymbol{u}, \dots, T_{d+1}\boldsymbol{u}, \boldsymbol{u}) \\ = (T_1(\boldsymbol{u}k), \dots, T_{k-1}(\boldsymbol{u}k), T_{k+1}(\boldsymbol{u}k), \dots, T_{d+1}(\boldsymbol{u}k), T_k(\boldsymbol{u}k)).\end{aligned}$$

This proves (i).

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

(iii) First, we need to observe that the position of the last occurrence of a symbol in a code cannot be earlier than that in some prefix of the same code, i.e. $|T_i(\boldsymbol{uxy})| > |T_i(\boldsymbol{ux})|$ for all i.

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

This implies that a shortest path between any two vertices $\boldsymbol{u}$ and $\boldsymbol{v}$ can be decomposed as $\boldsymbol{u} \to \cdots \to \boldsymbol{y} \to \cdots \to \boldsymbol{v}$, where $\boldsymbol{y}$ is either $\boldsymbol{u} \wedge \boldsymbol{v}$ or one if its ancestors, $\mathcal{P}_u := \boldsymbol{u} \to \cdots \to \boldsymbol{y}$ is a directed path in the graph where each vertex is an ancestor of the previous one on the path, and $\mathcal{P}_v = \boldsymbol{y} \to \cdots \to \boldsymbol{v}$ is a directed path where each vertex is a descendant of the previous one on the path.

Secondly, we need to observe that it is enough to go up to $\boldsymbol{u} \wedge \boldsymbol{v}$ both from $\boldsymbol{u}$ and $\boldsymbol{v}$: suppose that a shortest path is of the form $\boldsymbol{u} \to \cdots \to (\boldsymbol{u} \wedge \boldsymbol{v})\boldsymbol{x} \to \boldsymbol{y} \to (\boldsymbol{u} \wedge \boldsymbol{v})\boldsymbol{z} \to \cdots \to \boldsymbol{v}$, where $\boldsymbol{y}$ is an ancestor of $\boldsymbol{u} \wedge \boldsymbol{v}$, and $\boldsymbol{x}$ and $\boldsymbol{z}$ are such codes that there is a shortcut edge between $(\boldsymbol{u} \wedge \boldsymbol{v})\boldsymbol{x}$ and $\boldsymbol{y}$ and between $(\boldsymbol{u} \wedge \boldsymbol{v})\boldsymbol{z}$ and $\boldsymbol{y}$. Then, from (i) it is obvious that there is also an edge between $(\boldsymbol{u} \wedge \boldsymbol{v})\boldsymbol{x}, \boldsymbol{u} \wedge \boldsymbol{v}$ and between $(\boldsymbol{u} \wedge \boldsymbol{v})\boldsymbol{z}, \boldsymbol{u} \wedge \boldsymbol{v}$. This means that we can change $\boldsymbol{y}$ to $\boldsymbol{u} \wedge \boldsymbol{v}$ in this path to obtain a path of equal length. □

3.2. Distances in RANs and EANs: the main idea

With the help of the grouping of the edges as above, we can determine the distance between two arbitrary vertices u and v with codes $\boldsymbol{u}$ and $\boldsymbol{v}$ as follows.

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

If we pick u and v u.a.r. 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 symbols in the code and the length of the code—w.h.p. 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* there are 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 greater, atypical depths of a branching process that arises from the forward edges of the RANs. The effect of the shortcut edges on the distances is determined using renewal theory (also carried out in [1]) and large deviation theory. Finally, we optimise parameters such that we achieve the maximal distance by an entropy versus energy argument.

3.3. Combinatorial analysis of shortcut edges

Now we investigate the effect of shortcut edges on this tree. Lemma 2(i) says that the shortcut edges of a vertex $\boldsymbol{u}$ in the tree lead exactly to $T_i\boldsymbol{u}$, the prefixes of $\boldsymbol{u}$ achieved by removing the code after the last occurrence of symbol i in the code. Recall that $P_i\boldsymbol{u} = u_{(i)} \dots u_n$ denotes the postfix of $\boldsymbol{u}$ that starts with the last occurrence of the symbol $i \in \Sigma_d$ in the code of $\boldsymbol{u}$, while $T_i\boldsymbol{u} = u_1 \dots u_{(i)-1}$.

Let us denote the operator that gives the prefix with length $\min_i |T_i\boldsymbol{u}|$ by $T_{\min}$ and, further, denote the length of the maximal cut by

$$Y_d^{\mathrm{AN}}(\boldsymbol{u}) := |\boldsymbol{u}| - |T_{\min}\boldsymbol{u}| = \max_{i \in \Sigma_d}\{|P_i\boldsymbol{u}|\}. \tag{17}$$

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

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

Consecutively using the operator $T_{\min}$ we can decompose $\boldsymbol{u}$ into independent blocks, where each block, when reversed, ends at the first position when all the symbols in Σ_d have appeared. We call such a block a *full coupon collector block*. For example, for $\boldsymbol{u} = 1132133231222221131$ this gives 1|132|1332|31222|21131. Let us denote the total number of blocks needed in this decomposition by

$$N(\boldsymbol{u}) = \max\{k+1\colon (T_{\min})^k \boldsymbol{u} \neq \varnothing\}. \tag{18}$$

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 $P_i\boldsymbol{u}$. For example, 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 $T_{\min}$ (longest possible hops) is optimal.

Claim 1. Suppose that we have an arbitrary code $\boldsymbol{u}$ of length n with symbol 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 symbol in that block. Then the minimal number of blocks needed is given by $N(\boldsymbol{u})$.

Proof. Consider two different decompositions of $\boldsymbol{u}$ into blocks: in the first decomposition use the operator $T_{\min}$ consecutively, while in the second one we suppose that at least one block is not a full coupon collector block. Without loss of generality, we may assume that this is the first block from the end of the code $\boldsymbol{u}$. The endpoint of the first hop in the first decomposition is $T_{\min}\boldsymbol{u} := T_{i^*}\boldsymbol{u}$, while in the second decomposition the endpoint is $T_j\boldsymbol{u}$ for some $j \neq i^*$, with $|T_j\boldsymbol{u}| > |T_{\min}\boldsymbol{u}|$. Hence, there is a $\boldsymbol{w}$ such that $T_j\boldsymbol{u} = (T_{\min}\boldsymbol{u})\boldsymbol{w}$. Conclude from Lemma 2(iii) that $\operatorname{hop}(T_{\min}\boldsymbol{u}, \varnothing) \leq \operatorname{hop}(T_j\boldsymbol{u}, \varnothing)$; thus, the number of blocks in the second decomposition cannot be smaller than $N(\boldsymbol{u})$. □

Note that $Y_d^{\mathrm{AN}}(\cdot)$ and $N(\cdot)$ are deterministic operators when applied to a fixed code $\boldsymbol{u}$. Next we state the distributional properties of $Y_d^{\mathrm{AN}}(\boldsymbol{u})$ and $N(\boldsymbol{u})$ when $\boldsymbol{u}$ is the code of a uniformly chosen active clique. The reason for the need of this is that both in the evolution of the RAN and the EAN, once a clique with code $\boldsymbol{u}$ becomes inactive and is replaced with the vertex with code $\boldsymbol{u}$, the $d+1$ new cliques that become active are exactly the direct descendants (children) of the vertex $\boldsymbol{u}$ in the d-ary tree. At each step in the evolution of the RAN, the clique to become inactive is chosen u.a.r. among the active cliques, and in the EAN an independent coin flip with success probability q_n (not depending on the code itself) determines for each active clique if it becomes inactive or stays active for the next step.

As a result of these dynamics, by symmetry it is not difficult to see that an active vertex (active clique) chosen u.a.r. in both the RAN and the EAN has a code where, conditioned on the length of the code, the symbol at each position is chosen u.a.r. in Σ_d. Further, for two vertices chosen u.a.r. with codes $\boldsymbol{u}$ and $\boldsymbol{v}$ the symbols in these codes after the position $|\boldsymbol{u} \wedge \boldsymbol{v}| + 1$ are independent and uniformly distributed in Σ_d.

For every $k \geq 1$, define the random variable

$$H_k := \max\biggl\{\ell\colon \sum_{j=1}^{\ell} Y_d^{(\ell)} \leq k\biggr\}, \tag{19}$$

where $Y_d^{(\ell)}$ are i.i.d. copies of Y_d in (3).

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

Lemma 3. *Suppose that $\boldsymbol{u}$ is a code of length k with symbols chosen u.a.r. from Σ_d at each position. Then*

$$Y_d^{\mathrm{AN}}(\boldsymbol{u}) \stackrel{\mathrm{D}}{=} \min\{Y_d, k\}, \qquad N(\boldsymbol{u}) \stackrel{\mathrm{D}}{=} H_k + 1.$$

Proof. The last occurrence of any symbol $i \in \Sigma_d$ in a uniform code is the *first occurrence from backwards* of the same symbol. Hence, reverse the code of $\boldsymbol{u}$, and then $|P_i\boldsymbol{u}|$ is the position of the first occurrence of symbol i in a uniform sequence of symbols of length k, since $|\boldsymbol{u}| = k$. Clearly, $|P_i\boldsymbol{u}| = k$ if the symbol i does not occur in $\boldsymbol{u}$. As a result, $|P_i\boldsymbol{u}|$ has a geometric distribution with parameter $1/(d+1)$ truncated at k. Maximizing this over all $i \in \Sigma_d$ we obtain the well-known coupon collector problem that has distribution Y_d, truncated again at k. For the second expression, since $N(\boldsymbol{u})$ cuts down full coupon collector blocks from the end of the code of $\boldsymbol{u}$ consecutively, the maximal number of cuts possible is exactly the number of consecutive full coupon collector blocks in the reversed code of $\boldsymbol{u}$, an i.i.d. code of length k. Since the length of each block has distribution $\min\{Y_d, k\}$, and they are independent, the statement follows by observing that the last, nonfull block of the reversed code corresponds to the $+1$ in the statement. □

Recall μ_d, σ_d^2 from (4). From basic renewal theory [23] the following CLT holds as $k \to \infty$:

$$\frac{H_k - k/\mu_d}{\sqrt{k\sigma_d^2/\mu_d^3}} \xrightarrow{\mathrm{D}} \mathcal{N}(0, 1). \tag{20}$$

4. Distances in RANs and EANs

In light of the main idea of the proof in Section 3.2, we begin with the analysis of the tree created by the forward edges of the graph.

4.1. A continuous-time branching process

There is a natural embedding of the evolution of the RAN into the evolution 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 (equivalently, a vertex) has $d+1$ children and the lifespan of each individual is i.i.d. exponential with mean 1. Thus, after birth a vertex 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. Exp(1) lifespan. The process starts with a single individual that is called the *root* and which dies immediately at $t = 0$ giving birth to its $d+1$ children.

The bijection between the CTBP at the split times and a RAN_d is as follows. The individuals that have already split in the CTBP are the vertices already present (inactive vertices) in the RAN_d, while the active (alive) individuals in the CTBP correspond to the active vertices (active cliques) in the RAN_d. This holds since at every step of a RAN_d, $d+1$ new active cliques arise in place of the one that becomes inactive. Further, in a RAN_d an active clique is chosen u.a.r. in each step which is—by the memoryless property of exponential variables—equivalent to the fact that the next individual to split in the CTBP is an active individual chosen u.a.r.

We write $G_U(m)$ for the *generation* of a uniformly chosen active individual in the CTBP after m individuals have split, i.e. its graph distance from the root. In the next two propositions we describe the growth of our CTBP in terms of the typical size of $G_U(m)$ as well as the *degree of relationship* of two active individuals chosen u.a.r. in Proposition 1 and the maximal size of $G_U(m)$ together with its tail behaviour in Proposition 2.

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

Proposition 1. *Let Z denote a standard normal random variable. Then, as $m \to \infty$,*

$$\Big(G_U(m) - \frac{d+1}{d}\log m\Big)\Big(\frac{d+1}{d}\log m\Big)^{-1/2} \xrightarrow{\mathrm{D}} Z. \tag{21}$$

Further, let G_U and G_V denote the generations of two active vertices chosen u.a.r. in the CTBP after the mth split, and write $G_{U\wedge V}$ for the generation of the latest common ancestor of U and V. Then the marginal distribution $G_U \overset{\mathrm{D}}{=} G_U(m)$, and

$$\bigg(\frac{G_U - G_{U\wedge V} - ((d+1)/d)\log m}{\sqrt{((d+1)/d)\log m}}, \frac{G_V - G_{U\wedge V} - ((d+1)/d)\log m}{\sqrt{((d+1)/d)\log m}}\bigg) \xrightarrow{\mathrm{D}} (Z, Z'), \tag{22}$$

where Z, Z' are independent standard normal distributions.

The proposition is an application of [16, Chapter 4.2, Theorem 2.5] to the CTBP studied here with deterministic offspring distribution ($d+1$ children). Before the proof, we need a lemma, that originates from [16, Theorem 3.3], and the first part can also be found, e.g. in [7]. First some notation. Write D_i, S_i for the number of children of the ith splitting vertex and the number of active individuals after the ith split in a CTBP, and for an event A and random variable X write

$$\mathbb{P}_m(A) := \mathbb{P}(A \mid D_i, S_i,\ i = 1, \dots, m), \qquad \mathbb{E}_m[X] := \mathbb{E}[X \mid D_i, S_i,\ i = 1, \dots, m].$$

Claim 2. (i) The generation $G_U(m)$ of an active individual U chosen u.a.r. after the mth split in a CTBP satisfies the following indicator representation:

$$G_U(m) \overset{\mathrm{D}}{=} \sum_{i=1}^{m} \mathbf{1}_i, \tag{23}$$

where $\mathbb{E}_m[\mathbf{1}_i] = D_i/S_i$, and the indicators are independent, conditioned on the sequence $D_i, S_i, i = 1, \dots, m$.

(ii) Let us denote (U, V) a pair of individuals chosen u.a.r. after the mth split. Let us further assume that the latest common ancestor $U \wedge V$ of U and V reproduced at the $\tau_{U\wedge V}$th split. Then, conditioned on $\tau_{U\wedge V}$, the following two variables are independent and their joint distribution can be written as

$$(G_U - G_{U\wedge V}, G_V - G_{U\wedge V}) \overset{\mathrm{D}}{=} \bigg(\sum_{i=\tau_{U\wedge V}}^{m} \mathbf{1}_i, \sum_{i=\tau_{U\wedge V}}^{m} \mathbf{1}'_i\bigg), \tag{24}$$

where

$$\mathbb{P}_m((\mathbf{1}_i, \mathbf{1}'_i) = (1, 0) \mid \tau_{U\wedge V} < i) = \frac{D_i}{S_i}\frac{S_i - D_i}{S_i - 1}, \tag{25a}$$

$$\mathbb{P}_m((\mathbf{1}_i, \mathbf{1}'_i) = (0, 1) \mid \tau_{U\wedge V} < i) = \frac{D_i}{S_i}\frac{S_i - D_i}{S_i - 1}, \tag{25b}$$

$$\mathbb{P}_m((\mathbf{1}_i, \mathbf{1}'_i) = (1, 1), \tau_{U\wedge V} = i \mid \tau_{U\wedge V} \le i) = \frac{D_i(D_i - 1)}{S_i(S_i - 1)} \tag{25c}$$

and, conditioned on $\tau_{U\wedge V}$, different indices are independent.

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

Proof. (i) A proof using the ancestral line can be found in [16, Section 3.A] (see also Section 2.A for clearer explanations), but a proof based on induction can also be worked out. Here we give the core idea of the proof of Bühler. The ancestral line of an individual in a CTBP is the unique path from the individual to the root. For the time interval between the ith and $(i+1)$th split we can allocate a unique individual on the ancestral line that was active in this time interval. For the following observations, we condition on $D_i, S_i, i = 1, \dots, m$. Then $G_m = \mathbf{1}_1 + \mathbf{1}_2 + \cdots + \mathbf{1}_m$, where the indicators $\mathbf{1}_i$ are conditionally independent and $\mathbf{1}_i = 1$ if and only if the ancestor that was alive in the time interval between the ith and $(i+1)$th split was newborn (born at the ith split). Recall that the individual that splits at the ith step is chosen u.a.r., as well as U is also chosen u.a.r. among the S_m many active individuals after the mth split. Since in the interval between the ith and $(i+1)$th split there were exactly D_i many individuals newborn, and S_i many alive, and the ancestor of U is equally likely to be any of them, this yields the probability $\mathbb{P}(\mathbf{1}_i = 1 \mid D_i, S_i) = D_i/S_i$.

(ii) The proof follows from [16, Section 3.B] in a similar manner: after time $\tau_{U\wedge V}$, we write $G_U - G_{U\wedge V}$ as sums of indicators, where $\mathbf{1}_i$ is 1 if and only if the individual alive between the ith and $(i+1)$th on the ancestral line of U is newborn (born at the ith split). We do the same for $G_V - G_{U\wedge V}$ using $\mathbf{1}'_i$s. Conditional on $D_i, S_i, i = 1, \dots, m$, the pairs $(\mathbf{1}_i, \mathbf{1}'_i)$ become independent and their joint distribution is the one given in (25a)–(25c), since, at each step, each pair of active individuals is equally likely to be the ancestors of U and V, and the ancestral lines merge precisely when the ancestors of U and V are two children of the vertex that splits at step i, giving the last line of (25a)–(25c). □

Proof of Proposition 1. The proposition follows from Claim 2. More precisely, we note that in our case $D_i = d+1$ and $S_i = di+1$ are deterministic; hence,

$$G_U(m) \stackrel{\mathrm{D}}{=} \sum_{i=1}^{m} \mathbf{1}_i, \tag{26}$$

where $\mathbb{P}(\mathbf{1}_i = 1) = (d+1)/(di+1)$. From this identity the expectation and variance of G_U follows:

$$\mathbb{E}[G_U(m)] = \frac{d+1}{d}\log m + O(1), \qquad \mathbb{D}^2[G_U(m)] = \mathbb{E}[G_U] + O(m^{-1}). \tag{27}$$

The CLT (21) holds for the standardization of $G_U(m)$ since the collection of Bernoulli random variables $\{\mathbf{1}_i\}_{i=1}^m$, $m = 1, 2, \dots$ satisfies Lindeberg's condition.

For the second statement, $G_U \stackrel{\mathrm{D}}{=} G_U(m)$ is obvious by noting that the marginal of a uniformly chosen pair of vertices is a uniformly chosen vertex. Next, note that the event $(\mathbf{1}_i, \mathbf{1}'_i) = (1,1)$ means that the ancestral lines of U and V merge at the ith split. To see that $\tau_{U\wedge V}$ has a limiting distribution, we use

$$\mathbb{P}(\tau_{U\wedge V} \le k) = \prod_{i=k+1}^{m} (1 - \mathbb{P}(\tau_{U\wedge V} = i \mid \tau_{U\wedge V} \le i)), \tag{28}$$

where the factors on the right-hand side are the probabilities that the two ancestral lines do not merge at the ith split. This tends to a proper limiting distribution since by (25a)–(25c)

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

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

Hence, $\tau_{U\wedge V}$ has a limiting distribution, and clearly $(\log m - \log \tau_{U\wedge V})/\log m \to 1$. Note that this also means that $G_{U\wedge V}$ also has a limiting distribution, independent of m, since $G_{U\wedge V}$ is the generation of the individual that splits at the $\tau_{U\wedge V}$th split.

From here, one can show the joint convergence of (24) using the Lindeberg CLT for linear combinations of $\sum_{i=\tau_{U\wedge V}+1}^{m}(\alpha\,\mathbf{1}_i + \beta\,\mathbf{1}'_i)$ and find that the two variables in (22) tend jointly to a two-dimensional standard normal variable. □

Recall the definition of the function $f_d(c)$ from (6) and the constant $\tilde c_d$ that satisfies $\tilde c_d > (d+1)/d$, $f_d(\tilde c_d) = -1$. We will need the next proposition in the proof of Theorem 2.

Proposition 2. *The exact asymptotic tail behaviour of $G_U(m)$ is given by*

$$\lim_{m\to\infty} \frac{\log(\mathbb{P}(G_U(m) > c\log m))}{\log m} = f_d(c). \tag{29}$$

Further, after m splits the deepest branch in the CTBP satisfies

$$\frac{\max_{i\le m} G_U(i)}{\log m} \xrightarrow{\mathbb{P}} \tilde c_d. \tag{30}$$

Proof. Let $\Lambda_m(\theta) := \log \mathbb{E}[\exp\{\theta G_U(m)/\log m\}]$. Using (26), elementary calculation yields that

$$\Lambda_m(\theta \log m) = \sum_{i=1}^{m} \log\biggl(1 + \frac{d+1}{di+1}(\exp\{\theta\} - 1)\biggr).$$

Hence, from the series expansion of $\log(1+x)$, we see that

$$\lim_{m\to\infty} \frac{1}{\log m}\Lambda_m(\theta \log m) = \frac{d+1}{d}(\exp\{\theta\} - 1),$$

which is the cumulant generating function of a $\xi = \mathrm{Poi}((d+1)/d)$ random variable. The rate function of such a random variable is $-f_d(c)$. Hence, the conditions of the Gärtner–Ellis theorem [20, Section 2.3] are satisfied, which implies (29).

Our CTBP is a special case of so-called random lopsided trees [17], [27]. The maximal depth of such trees was studied by Broutin and Devroye [15] in a more general framework. Thus, (30) is just an application of [15, Theorem 5 and Remark afterwards] with our notation. This completes the proof. □

Remark 7. To see that $\tilde c_d$ should be the correct constant in (30) we can argue that, from (29), it follows that the sum $\sum_m \mathbb{P}(G_U(m) > c\log m) < \infty$ for any $c > \tilde c_d$. Thus, by the Borel–Cantelli lemma, for any such c there are only finitely many m such that the event $\{G_U(m) > c\log m\}$ holds, giving the w.h.p. upper bound $\tilde c_d \log m$ on the depth of the CTBP.

4.2. Proofs of Theorems 1 and 3

Proof of Theorem 1. Pick a pair of active vertices $\boldsymbol{u}, \boldsymbol{v}$ u.a.r. from a $\mathrm{RAN}_d(n)$. We write $|\boldsymbol{u}|, |\boldsymbol{v}|$ for their generation. As before, write $\boldsymbol{u}\wedge\boldsymbol{v}$ for their latest common ancestor, i.e. the longest common prefix of their codes. Define the distinct postfixes $\widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{v}}$ after $\boldsymbol{u}\wedge\boldsymbol{v}$ by

$$\boldsymbol{u} =: (\boldsymbol{u}\wedge\boldsymbol{v})\widetilde{\boldsymbol{u}}, \qquad \boldsymbol{v} =: (\boldsymbol{u}\wedge\boldsymbol{v})\widetilde{\boldsymbol{v}}.$$

By Lemma 2(iii) and Claim 1 the length of the shortest paths between $\boldsymbol{u}, \boldsymbol{v}$ satisfies

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

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

and Proposition 1 describes the typical distance between $\boldsymbol{u}$ and $\boldsymbol{v}$ along the tree (i.e. using only forward edges and no shortcut edges). Since $\boldsymbol{u}$ and $\boldsymbol{v}$ were chosen uniformly at random among the *active vertices* after n splits, we can write $|\boldsymbol{u}| \stackrel{\mathrm{D}}{=} G_U(n)$, $|\boldsymbol{v}| \stackrel{\mathrm{D}}{=} G_V(n)$, where U and V denote two uniformly chosen alive individuals in the CTBP in Section 4.1. By the same reasoning (and dropping the dependence of n for shorter notation),

$$|\boldsymbol{u} \wedge \boldsymbol{v}| \stackrel{\mathrm{D}}{=} G_{U\wedge V}, \qquad |\widetilde{\boldsymbol{u}}| \stackrel{\mathrm{D}}{=} G_U - G_{U\wedge V}, \qquad |\widetilde{\boldsymbol{v}}| \stackrel{\mathrm{D}}{=} G_V - G_{U\wedge V}. \tag{31}$$

Further, by symmetry of the process, the symbols in the codes $\widetilde{\boldsymbol{u}}$ and $\widetilde{\boldsymbol{v}}$ are i.i.d. uniform on Σ_d (after the first symbol in both of the codes, which has to be different by the definition of $\boldsymbol{u} \wedge \boldsymbol{v}$). Hence, by Lemma 3,

$$\operatorname{dist}(\boldsymbol{u}, \boldsymbol{v}) \stackrel{\mathrm{D}}{=} H_{G_U - G_{U\wedge V}} + H_{G_V - G_{U\wedge V}}. \tag{32}$$

Using (27) and the fact from the proof of Claim 2 that $G_{U\wedge V}$ has a limiting distribution, we calculate

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

To obtain a CLT for $H_{G_U - G_{U\wedge V}}$, observe that

$$\begin{aligned} &\frac{H_{G_U - G_{U\wedge V}} - (1/\mu_d)((d+1)/d)\log m}{\sqrt{((d+1)/d)\log m\sigma_d^2/\mu_d^3}} \\ &\quad = \frac{H_{G_U - G_{U\wedge V}} - (1/\mu_d)(G_U - G_{U\wedge V})}{\sqrt{(G_U - G_{U\wedge V})\sigma_d^2/\mu_d^3}} \sqrt{\frac{G_U - G_{U\wedge V}}{((d+1)/d)\log m}} \\ &\qquad + \frac{(1/\mu_d)(G_U - G_{U\wedge V}) - (1/\mu_d)((d+1)/d)\log m}{\sqrt{((d+1)/d)\log m\sigma_d^2/\mu_d^3}}. \end{aligned} \tag{33}$$

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 (20) and the second factor tends to 1 in probability by (22). By (22) again, the second term tends to a $\mathcal{N}(0, \mu_d/\sigma_d^2)$. Since the length of the codes $\widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{v}}$ are independent of the symbols in these codes, $H_{G_U - G_{U\wedge V}} \mid G_U - G_{U\wedge V}$ is independent of $G_U - G_{U\wedge V}$. As a result, the two limiting normals arising from the two summands on the right-hand side of (33) are also independent; thus,

$$\bigg(H_{G_U - G_{U\wedge V}} - \frac{1}{\mu_d}\frac{d+1}{d}\log m\bigg)\bigg(\sqrt{\frac{d+1}{d}\log m\sigma_d^2/\mu_d^3}\bigg)^{-1/2} \xrightarrow{\mathrm{D}} \mathcal{N}\bigg(0, 1 + \frac{\mu_d}{\sigma_d^2}\bigg). \tag{34}$$

By conditioning first on $G_{U\wedge V}$ (as in the proof of Proposition 1) and using the fact that the symbols in the code of $\widetilde{\boldsymbol{u}}, \widetilde{\boldsymbol{v}}$ are all i.i.d. uniform in Σ_d, one can 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. By (32) it follows that $\operatorname{hop}(n) = H_{G_U - G_{U\wedge V}} + H_{G_V - G_{U\wedge V}}$, and the first statement of the Theorem 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 in: a vertex with degree $d+1$ is contained in $d+1$ cliques, and when the degree of a vertex 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

$$Q_k = 2 + (k-d)(d-1) \tag{35}$$

active cliques. This means that the inactive vertex v is connected to exactly Q_k many active vertices with an edge. It is clear that the total number of active vertices after n steps is $A(n) = dn + d + 1$. This implies that choosing two inactive vertices x, y according to the size-biased distribution given in (9) is equivalent to choosing two active cliques U, V chosen u.a.r. that are neighbouring these vertices. The distance between x, y is then between $N(\widetilde{\boldsymbol{u}}) + N(\widetilde{\boldsymbol{v}}) - 2$, $N(\widetilde{\boldsymbol{u}}) + N(\widetilde{\boldsymbol{v}})$ since, by Lemma 2, $x = T_i \boldsymbol{u}$ for some $i \in \Sigma_i$; hence, we can gain at most one 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 (9). □

Remark 8. Denote the generation of the mth splitting vertex by $\widehat{G}_m$. Since at each split in the CTBP exactly one new inactive vertex is created; namely, a uniformly chosen active vertex becomes inactive, we have $\widehat{G}_m \overset{\mathrm{D}}{=} G_U(m-1)$. Hence, if we choose an *inactive vertex* of a $\mathrm{RAN}_d(n)$ *uniformly at random* then its distance from the root has distribution $G_U(X)$, where X is a random variable uniform in the set $\{0, 1, \dots, n-1\}$, with $G_U(0) = 1$. With a similar argument to the one in Claim 2, one can obtain that the latest common ancestor of two inactive vertices chosen u.a.r. also has a limiting distribution, and if $\widehat{\mathrm{hop}}(n)$ denotes the distance between them, one can obtain $\widehat{\mathrm{hop}}(n)/(2((d+1)/d)\log n) \overset{\mathbb{P}}{\to} 1$. But it is also not difficult to see that the CLT does not hold anymore (since it does not hold for $G_U(X)$ for X uniform in $\{0, 1, \dots, n-1\}$).

Proof of Theorem 3. The proof follows analogous lines to the proof of Theorem 1, hence, we only give the sketch. The main idea here is that the tree can be viewed as a CTBP 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, Proposition 1 can be modified as follows. The generation of an active individual picked u.a.r. after the mth split satisfies

$$\widetilde{G}_U(m) \overset{\mathrm{D}}{=} \sum_{i=1}^{m} \widetilde{\mathbf{1}}_i,$$

where $\widetilde{\mathbf{1}}_i = 1$ if and only if the individual on the ancestral line of U is newborn at the ith step. Note that in this case, the indicators are independent even without conditioning, and $\mathbb{P}(\widetilde{\mathbf{1}}_i = 1) = q_i$, since at each step *each* individual splits with the same probability, independently of each other. Since splitting happens with probability q_i at step i, the CLT for $\widetilde{G}_U(m)$ holds by the Lindeberg CLT. Now, for two individuals U and V picked u.a.r., with $G_{U\wedge V}, \tau_{U\wedge V}$ as in Proposition 1, we have

$$(G_U - G_{U\wedge V}, G_V - G_{U\wedge V}) \overset{\mathrm{D}}{=} \Bigg(\sum_{i=\tau_{U\wedge V}}^{m} \widetilde{\mathbf{1}}_i, \sum_{i=\tau_{U\wedge V}}^{m} \widetilde{\mathbf{1}}'_i \Bigg),$$

where different indices are independent and conditioned on $\tau_{U\wedge V}$, $\widetilde{\mathbf{1}}_i, \widetilde{\mathbf{1}}'_i$ are independent indicators with $\mathbb{P}(\widetilde{\mathbf{1}}_i = 1) = \mathbb{P}(\widetilde{\mathbf{1}}'_i = 1) = q_i$. Since the variance $\sum_i q_i(1-q_i) \to \infty$, the joint CLT follows in a similar manner as for Proposition 1 if we can show that $\tau_{U\wedge V}$ has a limiting distribution. For this, note that, similarly as in (28),

$$\mathbb{P}(\tau_{U\wedge V} \leq k) = \prod_{i=k+1}^{m} (1 - \mathbb{P}(\tau_{U\wedge V} = i \mid \tau_{U\wedge V} \leq i)), \tag{36}$$

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

and the factors on the right-hand side express that the two ancestral lines of U and V do not merge yet at step i. Write A_i for the number of active vertices at step i. Then at step i there are $Z_i := \text{bin}(A_i, q_i)$ 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} is

$$\mathbb{P}(\tau_{U\wedge V} = i \mid \tau_{U\wedge V} \le i, A_i, A_{i+1}) = \frac{Z_i(d+1)d}{A_{i+1}(A_{i+1}-1)}, \tag{37}$$

where $A_{i+1} = A_i + dZ_i$, the new number of active vertices after the ith split. The right-hand side of (37) is obtained by observing that if U and V are chosen uniformly at random, each pair of individuals at step i, $A_{i+1}(A_{i+1}-1)/2$ in total, is equally likely to be the ancestors of them, and there are $Z_i(d+1)d/2$ many pairs that make the ancestral lines merge. If the sum in $i \in \mathbb{N}$ on the right-hand side of (37) is almost surely (a.s.) finite then (36) 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 inactive vertices $N(n) \to \infty$, i.e.

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

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 is a.s. finite if and only if the integral

$$\int_1^{\infty} \frac{N'(x)}{N(x)^2}\,\mathrm{d}x$$

is a.s. finite. Note that this is the case if and only if $N(n) \to \infty$. Further, as long as $\sum_{n\in\mathbb{N}} q_n = \infty$, $N(n) \to \infty$ holds a.s. by the second Borel–Cantelli lemma: in each step we add at least a new vertex with probability q_n. Then the CLT for the distances follows in the exact same manner as in the proof of Theorem 1. □

4.3. Proof of Theorem 2

We need some preliminary statements before the proof. Recall from (5) the definition of the large deviation rate function $I_d(x)$ of Y_d and also H_k as the number of consecutive occurrences of full coupon collector blocks in a code of length k from (19).

Lemma 4. *For* $1 \le \beta \le \mu_d/(d+1)$, H_k *satisfies the large deviation*

$$\lim_{k\to\infty} \frac{1}{k} \log\biggl(\mathbb{P}\biggl(H_k > \frac{\beta}{\mu_d}k\biggr)\biggr) = -\frac{\beta}{\mu_d} I_d\biggl(\frac{\mu_d}{\beta}\biggr). \tag{38}$$

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

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

we can apply Cramér's theorem [20, Section 2.2] to obtain (38). □

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

At the end of the proof of Theorem 1 we see that switching from inactive vertices to neighbouring active vertices/cliques only changes the distances by at most 2; hence, it is preferable to investigate the diameter of the graph by active vertices. Denote the set of active vertices at step n by $\mathcal{A}_n$. We index $\mathcal{A}_n$ by vertices $\boldsymbol{u}$ and denote one picked u.a.r. by U. We have seen that $|\mathcal{A}_n| = dn + d + 1$. Our aim is to estimate the expected number of $\boldsymbol{u} \in \mathcal{A}_n$ with distance at least $x\tilde{c}_d \log n/\mu_d$ from the root for some $x \geq 1$.

Recall the definition of the function $g(\alpha, \beta)$ from (7), and define, for an $x \geq 1$,

$$(\alpha(x), \beta(x)) := \arg\sup_{\alpha,\beta}\{g(\alpha, \beta) \colon \alpha\beta = x\}. \tag{39}$$

Claim 3. For any $x \geq 1$, define the indicator variables for each vertex $\boldsymbol{u} \in \mathcal{A}_n$ as

$$J_{\boldsymbol{u}}(x) := \mathbf{1}_{\{N(\boldsymbol{u}) > x\tilde{c}_d \log n/\mu_d\}}.$$

Then, with $(\alpha(x), \beta(x))$ as in (39),

$$\lim_{n\to\infty} \frac{1}{\log n} \log \mathbb{E}\bigg[\sum_{\boldsymbol{u}\in\mathcal{A}_n} J_{\boldsymbol{u}}(x)\bigg] = g(\alpha(x), \beta(x)). \tag{40}$$

Proof. Note that

$$|\mathcal{A}_n|\mathbb{E}\bigg[\frac{1}{|\mathcal{A}_n|}\sum_{\boldsymbol{u}\in\mathcal{A}_n} J_{\boldsymbol{u}}(x)\bigg] = (dn + d + 1)\mathbb{P}\bigg(H_{G_U(n)} \geq \frac{x\tilde{c}_d \log n}{\mu_d}\bigg), \tag{41}$$

where we use Lemma 3 for the distributional identity $N(\boldsymbol{u}) \stackrel{\mathrm{D}}{=} H_{G_U(n)}$ for a uniformly chosen $\boldsymbol{u} \in \mathcal{A}_n$ (see also the argument above (31)). Then

$$\begin{aligned}\mathbb{P}\bigg(H_{G_U(n)} \geq \frac{x}{\mu_d}\tilde{c}_d \log n\bigg)& \\ &\geq \mathbb{P}(G_U(n) > \alpha(x)\tilde{c}_d \log n)\mathbb{P}\bigg(H_{\alpha(x)\tilde{c}_d \log n} > \frac{\beta(x)}{\mu_d}\alpha(x)\tilde{c}_d \log n\bigg),\end{aligned}$$

where we use the fact that $x = \alpha(x)\beta(x)$ and the symbols in a uniformly chosen $\boldsymbol{u}$ are i.i.d. uniform in Σ_d, and H_k is increasing in k. Finally, multiplying by $dn + d + 1$, taking the logarithm, and dividing by $\log n$, applying (29) and (38) (with $k = \alpha(x)c_d \log n$ and only dividing by $\log n$ instead of $\alpha(x)c_d \log n$), we arrive at

$$\begin{aligned}\lim_{n\to\infty} \frac{\log \mathbb{E}[\sum_{\boldsymbol{u}\in\mathcal{A}_n} J_{\boldsymbol{u}}(x)]}{\log n} &\geq 1 + f_d(\alpha(x)\tilde{c}_d) - \alpha(x)\tilde{c}_d \frac{\beta(x)}{\mu_d} I_d\bigg(\frac{\mu_d}{\beta(x)}\bigg) \\ &= g(\alpha(x), \beta(x)).\end{aligned}$$

For the upper bound, fix a small $\varepsilon > 0$, and set

$$\begin{aligned}i^\star(\varepsilon) &:= \max\bigg\{i \colon (\alpha(x) - (i+1)\varepsilon)\tilde{c}_d \geq \frac{(d+1)}{d}\bigg\}, \\ i_\star(\varepsilon) &:= \min\{i \colon (\alpha(x) - i\varepsilon)\tilde{c}_d \geq x\tilde{c}_d\}.\end{aligned}$$

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

Then we can decompose the event $\{H_{G_U(n)} \geq x\tilde{c}_d \log n/\mu_d\}$ according to which ε-length interval $G_U(n)/(\tilde{c}_d \log n)$ falls into, and use the monotonicity of H_k in k to obtain

$$\begin{aligned}
&\mathbb{P}\bigg(H_{G_U(n)} \geq \frac{x}{\mu_d}\tilde{c}_d \log n\bigg) \\
&\quad \leq \mathbb{P}\bigg(H_{(d+1)\log n/d} \geq \frac{x\tilde{c}_d \log n}{\mu_d}\bigg) + \mathbb{P}(G_U(n) > x\tilde{c}_d \log n) \\
&\qquad + \sum_{i=i_\star(\varepsilon)}^{i^\star(\varepsilon)} \mathbb{P}\bigg(H_{(\alpha(x)-i\varepsilon)\tilde{c}_d \log n} \geq \frac{x\tilde{c}_d \log n}{\mu_d}\bigg) \\
&\qquad \times \mathbb{P}((\alpha(x)-(i+1)\varepsilon)\tilde{c}_d \log n < G_U(n) < (\alpha(x)-i\varepsilon)\tilde{c}_d \log n). \qquad (42)
\end{aligned}$$

For the first term on the right-hand side of (42), using (38), we obtain

$$\lim_{n\to\infty} \frac{\log(\mathbb{P}(H_{(d+1)\log n/d} \geq x\tilde{c}_d \log n/\mu_d))}{\log n} = -\frac{x\tilde{c}_d}{\mu_d} I_d\bigg(\frac{(d+1)\mu_d}{x\tilde{c}_d d}\bigg). \qquad (43)$$

For the ith summand in the third term on the right-hand side of (42), we use an upper bound by dropping the upper restriction on $G_U(n)$, and using (38) again and also (29), we obtain

$$\begin{aligned}
&\log\bigg(\mathbb{P}\bigg(H_{(\alpha(x)-i\varepsilon)\tilde{c}_d \log n} \geq \frac{x\tilde{c}_d \log n}{\mu_d}\bigg)\mathbb{P}((\alpha(x)-(i+1)\varepsilon)\tilde{c}_d \log n < G_U(n))\bigg) \\
&\quad \leq \log n\bigg(-\frac{x\tilde{c}_d}{\mu_d} I_d\bigg(\frac{\mu_d(\alpha - i\varepsilon)}{x}\bigg) + f_d((\alpha(x)-(i+1)\varepsilon)\tilde{c}_d)\bigg)(1+o(1)), \qquad (44)
\end{aligned}$$

where the $(1+o(1))$ disappears when dividing by $\log m$ and taking the limit as $n \to \infty$. The second term on the right-hand side of (42) can be treated similarly, except that there is no part coming from the large deviation principle of the $H_\cdot$. This is not surprising since this is the point where the length of the code becomes so large that a typical number of shortcut edges already exceeds $x\tilde{c}_d \log n/\mu_d$.

To finish the upper bound, note that setting $i = i^\star(\varepsilon) + 1$, the right-hand side of (44) exactly yields the right-hand side of (43), while setting $i = i_\star(\varepsilon)$ yields the second term, since in this case the rate function $I_d(\cdot)$ vanishes. Further, note that the terms in (42) are *additive*. This implies that when taking the logarithm and dividing by $\log n$, the largest term will dominate and determine the leading exponent. As a result,

$$\begin{aligned}
&\lim_{n\to\infty} \frac{\mathbb{P}(H_{G_U}(n) \geq x\tilde{c}_d \log n/\mu_d)}{\log n} \\
&\quad \leq \max_{i\in[i_\star(\varepsilon),\, i^\star(\varepsilon)+1]} -\frac{x\tilde{c}_d}{\mu_d} I_d\bigg(\frac{\mu_d(\alpha(x)-i\varepsilon)}{x}\bigg) + f_d((\alpha(x)-(i+1)\varepsilon)\tilde{c}_d).
\end{aligned}$$

To finish, let $\varepsilon \to 0$, and note that the ith term on the right-hand side is $g(z, x/z) - 1$ for some $z \in \mathbb{R}$. Since the maximum of this expression is taken at $z = \alpha(x)$, the proof is complete. □

Claim 4. (*Monotonicity of* $g(\alpha(x), \beta(x))$ *in* x.) The function $g(\alpha(x), \beta(x))$ is continuous and strictly monotone decreasing for $x > (d+1)/(d\tilde{c}_d)$.

Proof. Recall that

$$g(\alpha, \beta) := 1 + f_d(\alpha\tilde{c}_d) - \alpha\beta\frac{\tilde{c}_d}{\mu_d} I_d\bigg(\frac{\mu_d}{\beta}\bigg).$$

The continuity follows from the fact that $g(\alpha, \beta)$ is differentiable. For the monotonicity, consider $x_1 > x_2 > 1$. We have to show that the maximum of the function $g(\alpha, \beta)$ on the hyperbola $\beta = x_1/\alpha$ is smaller than that on $\beta = x_2/\alpha$. Let $g_1 := g(\alpha(x_1), \beta(x_1))$ and $g_2 := g(\alpha(x_2), \beta(x_2))$. Note that $f_d(\alpha\tilde{c}_d) < 0$ and monotone decreasing in α as long as $\alpha > (d+1)/(d\tilde{c}_d)$, while the second term $-\alpha\beta I_d(\mu_d/\beta) < 0$ and monotone decreasing in β as long as $\beta > 1$. Since $x_1 > (d+1)/(d\tilde{c}_d)$, at least one of the inequalities $\alpha(x_1) > (d+1)/(d\tilde{c}_d)$ and $\beta(x_1) > 1$ must hold.

First, suppose that $\alpha(x_1) > (d+1)/(d\tilde{c}_d)$ holds. Then, if $x_2/\beta(x_1) > (d+1)/(d\tilde{c}_d)$ then clearly $g_2 = g(\alpha(x_1), x_1/\alpha(x_1)) < g(x_2/\beta(x_1), \beta(x_1)) \le g_1$ and we are done. If, on the other hand, $x_2/\beta(x_1) < (d+1)/(d\tilde{c}_d)$ then consider the point on the hyperbola $(d+1)/(d\tilde{c}_d), x_2/(d+1)/(d\tilde{c}_d)$. Since we decreased both coordinates,

$$g_1 < g\biggl(\frac{d+1}{d\tilde{c}_d}, \frac{x_2/(d+1)}{d\tilde{c}_d}\biggr)$$

holds as long as $x_2/(d+1)/(d\tilde{c}_d) > 1$. This must hold since otherwise the whole hyperbola $\beta = x_2/\alpha$ would be in the region $\{\alpha < d+1/(d\tilde{c}_d)\} \cup \{\beta < 1\}$, which would mean that $x_2 < (d+1)/(d\tilde{c}_d)$ which contradicts our original assumption.

If $\beta(x_1) > 1$ then the argument is similar by first decreasing β to $x_2/\alpha(x_1)$ or to 1 (whichever is larger), and in case we have to decrease it to 1 then we further decrease $\alpha(x_1)$ to x_2 and again using the fact that $x_2 > 1$ implies that $x_2 > (d+1)/d\tilde{c}_d$. □

Proof of Theorem 2. First, we choose the largest possible x in $J_u(x)$ so that

$$\lim_{n\to\infty} \mathbb{P}\biggl(\sum_{u\in\mathcal{A}_n} J_u(x) > 0\biggr) > 0,$$

i.e. there is at least one active clique that has distance $x\tilde{c}_d \log n/\mu_d$ from the root. Note that if x is such that $g(\alpha(x), \beta(x)) < 0$, then by Claim 3 and Markov's inequality, we have

$$\mathbb{P}\biggl(\sum_{u\in\mathcal{A}_m} J_u(x) > 0\biggr) \le \mathbb{E}\biggl[\sum_{u\in\mathcal{A}_n} J_u(x)\biggr] = n^{g(\alpha(x),\beta(x))(1+o(1))} \to 0. \tag{45}$$

Thus, necessarily x has to have $g(\alpha(x), \beta(x)) \ge 0$. Next we work out the lower bound. For this, we need a upper bound on the second moment

$$\mathbb{E}\biggl[\biggl(\sum_{u\in\mathcal{A}_n} J_u(x)\biggr)^2\biggr] = \mathbb{E}\biggl[\sum_{u\in\mathcal{A}_n} J_u(x)\biggr] + \sum_{u,v\in\mathcal{A}_n,\, u\neq v} \mathbb{E}[J_u(x)J_v(x)].$$

Note that the second term equals

$$(dn+d+1)^2\mathbb{P}\biggl(H_{G_U} > \frac{x\tilde{c}_d \log n}{\mu_d}, H_{G_V} > \frac{x\tilde{c}_d \log n}{\mu_d}\biggr),$$

where H_{G_U} and H_{G_V} is the minimal number of hops needed to reach the root from two active vertices chosen independently and u.a.r. As before, write $U \wedge V$ for the latest common ancestor of U and V. Then we can write

$$\begin{aligned}\mathbb{P}\biggl(H_{G_U} > \frac{x\tilde{c}_d \log n}{\mu_d}, H_{G_V} > \frac{x\tilde{c}_d \log n}{\mu_d}\biggr) \\ = \mathbb{P}\biggl(H_{G_{U\wedge V}} + H_{G_{U-G_{U\wedge V}}} > \frac{x\tilde{c}_d \log n}{\mu_d}, H_{G_{U\wedge V}} + H_{G_{V-G_{U\wedge V}}} > \frac{x\tilde{c}_d \log n}{\mu_d}\biggr).\end{aligned}$$

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

Pick any function $\omega(n) \to \infty$ that also satisfies $\omega(n) = o(\log n)$ (e.g. $\omega(n) = \log\log n$ will do), then we can bound the right-hand side from above as follows:

$$\mathbb{P}(H_{G_{U\wedge V}} > \omega(n)) + \mathbb{P}\bigg(H_{G_U - G_{U\wedge V}} > \frac{x}{\mu_d}\tilde{c}_d \log n - \omega(n),\\ H_{G_V - G_{U\wedge V}} > \frac{x}{\mu_d}\tilde{c}_d \log n - \omega(n)\bigg). \tag{46}$$

Using the proof of Claim 2, we know that the joint distribution of two active individuals U, V picked u.a.r. satisfies the fact that their common ancestor $G_{U\wedge V}$ has a limiting distribution. Hence, for any $\omega(n) \to \infty$,

$$\mathbb{P}(H_{G_{U\wedge V}} > \omega(n)) \to 0. \tag{47}$$

Further, conditioned on the splitting time $\tau_{u\wedge v}$ of $U \wedge V$, we can describe the joint distribution of $G_U - G_{U\wedge V}, G_V - G_{U\wedge V}$ as the sum of indicators, see (24). Further, the two sums are asymptotically independent, and also the symbols in the code $\boldsymbol{u}$, $\boldsymbol{v}$ of U and V are independent and uniform in Σ_d after $\boldsymbol{u} \wedge \boldsymbol{v}$, the code of $U \wedge V$. Hence, choosing a large enough n, the $\omega(n)$ term becomes negligible and we obtain

$$\begin{aligned}
&\mathbb{P}\bigg(H_{G_U - G_{U\wedge V}} > \frac{x\tilde{c}_d \log n}{\mu_d} - \omega(n), H_{G_V - G_{U\wedge V}} > \frac{x\tilde{c}_d \log n}{\mu_d} - \omega(n)\bigg)\\
&\quad = \mathbb{P}\bigg(H_{G_U - G_{U\wedge V}} > \frac{x\tilde{c}_d \log n}{\mu_d}\bigg)\mathbb{P}\bigg(H_{G_V - G_{U\wedge V}} > \frac{x\tilde{c}_d \log n}{\mu_d}\bigg)(1 + o(1))\\
&\quad = \mathbb{P}\bigg(H_{G_U} > \frac{x\tilde{c}_d \log n}{\mu_d}\bigg)^2 (1 + o(1)),
\end{aligned} \tag{48}$$

where the $o(1) \to 0$ as $n \to \infty$, and in the last equation we used again the fact that $G_{U\wedge V}$ has a limiting distribution. Combining (47) with (48) to bound (46), we arrive at

$$\begin{aligned}
&\mathbb{E}\bigg[\bigg(\sum_{u\in\mathcal{A}_n} J_u(x)\bigg)^2\bigg]\\
&\quad = \mathbb{E}\bigg[\sum_{u\in\mathcal{A}_n} J_u(x)\bigg] + (nd + d + 1)^2 \bigg(\mathbb{P}\bigg(H_{G_U} > \frac{x\tilde{c} \log n}{\mu_d}\bigg)^2 (1 + o(1)) + o(1)\bigg)\\
&\quad = \mathbb{E}\bigg[\sum_{u\in\mathcal{A}_n} J_u(x)\bigg] + \mathbb{E}\bigg[\sum_{u\in\mathcal{A}_n} J_u(x)\bigg]^2 (1 + o(1)).
\end{aligned} \tag{49}$$

From a Cauchy–Schwarz inequality followed by (49), we obtain

$$\begin{aligned}
\mathbb{P}\bigg(\sum_{u\in\mathcal{A}_n} J_u(x) > 0\bigg) &\geq \frac{\mathbb{E}[\sum_{u\in\mathcal{A}_n} J_u(x)]^2}{\mathbb{E}[(\sum_{u\in\mathcal{A}_n} J_u(x))^2]}\\
&\geq \frac{\mathbb{E}[\sum_{u\in\mathcal{A}_n} J_u(x)]^2}{\mathbb{E}[\sum_{u\in\mathcal{A}_n} J_u(x)] + \mathbb{E}[\sum_{u\in\mathcal{A}_n} J_u(x)]^2 (1 + o(1))},
\end{aligned} \tag{50}$$

and the right-hand side is strictly positive in the limit as $n \to \infty$ if and only if $g(\alpha, \beta) \geq 0$ (using Claim 3 again for each term on the right-hand side). From this and the monotonicity

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

of $g(\alpha(x), \beta(x))$ in x (see Claim 4), it is immediate that the largest diameter can be achieved when picking $x := \tilde{x}$ so that $g(\alpha(\tilde{x}), \beta(\tilde{x})) = 0$. Apply (45) with $x = \tilde{x}(1 + \varepsilon)$ and (50) with $x := \tilde{x}(1 - \varepsilon)$ to finally conclude that

$$\frac{\max_{u \in \mathcal{A}_n} H_{G_U(n)}}{\log n} \xrightarrow{\mathbb{P}} \frac{\tilde{c}_d}{\mu_d} \tilde{x} \quad \text{as } n \to \infty. \tag{51}$$

The statement of Theorem 2 for the flooding time now follows from the fact that if U is an active clique picked u.a.r. after the nth step of the evolution of the RAN then

$$\mathrm{flood}(n) \stackrel{\mathrm{D}}{=} H_{G_U - G_{U \wedge V}} + \max_{v \in \mathcal{A}(n)} H_{G_V - G_{U \wedge V}}.$$

Now, the proof of Theorem 1 (or Proposition 1) implies that the CLT holds for generation $G_U - G_{U \wedge V}$, and since the symbols are uniform in the code of U, similarly as in (33), the CLT holds for $H_{G_U(n)}$ as well. Further, since in flood(u, v) we maximise the distance over the choice of the other vertex V, clearly w.h.p. 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, and the statement of the theorem follows from, the distributional convergence of

$$\frac{H_{G_U(n)}}{\log n} \xrightarrow{\mathrm{D}} \frac{(d+1)d}{\mu_d}$$

and (51). For the diameter, we have

$$\frac{\mathrm{diam}(n)}{\log n} \stackrel{\mathrm{D}}{=} 2 \frac{\max_{u \in \mathcal{A}(n)} H_{G_u}}{\log n},$$

since for any $\varepsilon > 0$, w.h.p. there are at least two vertices that are not closely related to each other and both satisfy

$$\frac{H_{G_u}}{\log n} > \left(\frac{\widetilde{c}_d}{\mu_d}\right) \tilde{x}(1 - \varepsilon),$$

but w.h.p. there are no vertices that satisfy $H_{G_v}/\log n > (\widetilde{c}_d/\mu_d)\tilde{x}(1 + \varepsilon)$. □

We are left to analyse the maximization problem. First of all, it is elementary to see (e.g. using Claim 4 or elementary two-dimensional calculus) that solving (39) and then choosing $\tilde{x}$ so that $g(\alpha(\tilde{x}), \beta(\tilde{x})) = 0$ is equivalent to the maximization problem in (10). However, two-dimensional techniques provide a better understanding of the solution $\tilde{x} = \alpha(\tilde{x})\beta(\tilde{x})$. In short we write $\alpha(\tilde{x}) := \tilde{\alpha}$, $\beta(\tilde{x}) := \tilde{\beta}$.

Lemma 5. *The maximization problem (10) has a unique solution*

$$(\tilde{\alpha}, \tilde{\beta}) \in (0, 1] \times \left[1, \frac{\mu_d}{d+1}\right],$$

and, further, this solution satisfies

$$\tilde{\alpha} = \frac{1}{\tilde{c}_d} \frac{d+1}{d} \exp\left\{-I_d'\left(\frac{\mu_d}{\tilde{\beta}}\right)\right\}, \qquad \frac{\tilde{\beta}}{\mu_d} I_d\left(\frac{\mu_d}{\tilde{\beta}}\right) = \frac{1 + f_d(\tilde{\alpha}\tilde{c}_d)}{\tilde{\alpha}\tilde{c}_d}.$$

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

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 $\nabla\mathcal{L}(\tilde{\alpha}, \tilde{\beta}, \widetilde{\lambda}) = 0$. The partial derivative $\mathcal{L}(\alpha, \beta, \lambda)'_\lambda = 0$ simply gives the condition $g(\alpha, \beta) = 0$. Further, the optimising $\widetilde{\lambda}$ can be expressed from $\mathcal{L}(\alpha, \beta, \lambda)'_\alpha = 0$ and $\mathcal{L}(\alpha, \beta, \lambda)'_\beta = 0$ and satisfies

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

After differentiation of

$$g(\alpha, \beta) = 1 + f_d(\alpha\tilde{c}_d) - \alpha\tilde{c}_d \frac{\beta}{\mu_d} I_d\left(\frac{\mu_d}{\beta}\right),$$

rearranging terms, and using the fact that $f'_d(x) = -\log((d/(d+1))x)$, we obtain the first condition. To check the sufficiency we look at the bordered Hessian

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

Its determinant is $(\partial^2 g(\alpha, \beta)/\partial\alpha\partial\beta)^2 > 0$; thus, the condition is also sufficient. We note that the solution can be approximated by numerical methods. □

Remark 9. 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 branching process.

(i) The corresponding branching process tree is fatter than the branching process for the RAN as soon as $n^{-1} = o(q_n)$.

(ii) In each step each vertex splits independently of the past with probability q_n.

Both (i) and (ii) together imply that even though we do understand that the marginal distribution of the symbols of a clique U picked u.a.r. 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 $N(\boldsymbol{u})$ 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 the RAN. (Compare it to the diameter of the deterministic AN: with $q_n \equiv 1$ it is not difficult 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 an elementary claim.

Claim 5. The series p_k given in (12) is a probability distribution.

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

Proof. Clearly $p_k \geq 0$. Combining the formula for p_k in (12) with an elementary rewrite of the fraction of the gamma functions inside the sum yields

$$\begin{aligned}\sum_{k=d+1}^{\infty} p_k &= \frac{d}{2d+1}\frac{\Gamma(1+(2d+1)/(d-1))}{\Gamma(1+2/(d-1))}\sum_{k=d+1}^{\infty}\frac{\Gamma(k-d+2/(d-1))}{\Gamma(k-d+(2d+1)/(d-1))}\\ &= \frac{d}{2d+1}\frac{\Gamma(1+(2d+1)/(d-1))}{\Gamma(1+2/(d-1))}\\ &\quad\times\sum_{k=d+1}^{\infty}\frac{d-1}{d}\biggl(\frac{\Gamma(k-d+2/(d-1))}{\Gamma(k-1-d+(2d+1)/(d-1))}\\ &\qquad\qquad -\frac{\Gamma(k+1-d+2/(d-1))}{\Gamma(k-d+(2d+1)/(d-1))}\biggr)\\ &= 1,\end{aligned}$$

since the last sum is telescopic. □

We proceed by analysing RANs first.

5.1. Sketch of the proof of Theorem 4

The proof of Theorem 4 determining the degree distribution of RANs consists of two main steps that are described in Lemmas 6 and 7. Recall the definition of $\widetilde{N}_k(n)$ and $N_k(n)$ from (1) and (2). In Lemma 6 we see that $\widetilde{N}_k(n)$ converges to its expectation uniformly in k as $n \to \infty$. The method we describe here is an adaptation of the standard martingale method and is similar to that in [12], [13], and [30]. Parallel to our work, Frieze and Tsourakakis [24] applied this method to show Lemmas 6 and 7 for two dimensions and their proof can be generalised to higher dimensions without any difficulty; hence, we give only a sketch of proof here.

Lemma 6. (Frieze and Tsourakakis [24].) *Fix $d \geq 2$ and $c_1 > \sqrt{8}(d+1)$. Then*

$$\lim_{n\to\infty}\mathbb{P}\Bigl(\max_k |\widetilde{N}_k(n) - \mathbb{E}[\widetilde{N}_k(n)]| \geq c_1\sqrt{n\log n}\Bigr) = 0.$$

This lemma tells us that $\widetilde{N}_k(n)$ concentrates around its expected value. From this we immediately obtain the concentration of $\widetilde{p}_k(n) = \widetilde{N}_k(n)/(n+d+2)$ around its expected value. Lemma 7 approximates the difference between this expected value and p_k, the stationary distribution.

Lemma 7. (Frieze and Tsourakakis [24].) *There exists a probability distribution $\{p_k\}_{k=d+1}^{\infty}$ for which for any $n \geq 0$ and for any $k \geq d+1$,*

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

with some constant c_2. The distribution $\{p_k\}_{k\in\mathbb{N}}$ is determined in (63) and it has a power-law asymptotic decay with exponent $(2d-1)/(d-1) \in (2,3]$ for $d \geq 2$.

As mentioned above, we do not provide the proof of these lemmas here. The methods, however, are similar to the ones used in the proof of Lemmas 8 and 9 for the EANs below. Given Lemmas 8 and 9, the proof of Theorem 4 is immediate.

Proof of Theorem 4. By the triangle inequality Theorem 4 follows from Lemmas 6 and 7 with $c = c_1 + c_2$. □

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

5.2. Proof of Theorem 5

We prove Theorem 5 about the degree distribution of EANs again in two main steps, as in the case of the RANs. Recall the definition $p_k(n)$ from (2). We denote the total minus the initial number of inactive vertices after n steps by $N^*(n)$, i.e. $N^*(n) = N(n) - d - 2$. Note that $N^*(n)$ is random (and does not necessarily concentrate); hence, any statement about the degree distribution is more accurate when stated conditional on $N^*(n)$ (rather than in terms of its mean, say). Denote the σ-algebra generated by $\{N^*(1), \dots, N^*(n)\}$ by $\mathcal{G}_n$. In the following lemma we see that the empirical proportion of degree k inactive vertices concentrates around its $\mathcal{G}_n$-conditional mean.

Lemma 8. *Fix the dimension $d \geq 2$, a constant $c^\star > \sqrt{24}(d+1)$, and a sequence of vertex arrival probabilities $\{q_n\}_{n=1}^{\infty}$ such that $N^*(n) \to \infty$ a.s. as $n \to \infty$. Then there exists $\varepsilon = \varepsilon(c^\star) > 0$ such that w.h.p.*

$$\mathbb{P}\biggl(\max_k \biggl| p_k(n) - \frac{\mathbb{E}[N_k(n) \mid \mathcal{G}_n]}{N^*(n)} \biggr| \geq c^\star \sqrt{\frac{\log N^*(n)}{N^*(n)}} \biggm| N^*(n) \biggr) = o(N^*(n)^{-\varepsilon(c^\star)}).$$

In the next lemma we see that the $\mathcal{G}_n$-conditional mean of the proportion of degree k inactive vertices tends to p_k (defined in (12)).

Lemma 9. *Let $d \geq 2$ and assume that $\{q_n\}_{n=1}^{\infty}$ is a sequence that satisfies Assumption 1. Then for any $k \geq d+1$ there exists a random variable $\eta_k < \infty$ such that w.h.p.*

$$\biggl| \frac{\mathbb{E}[N_k(n) \mid \mathcal{G}_n]}{N^*(n)} - p_k \biggr| \leq \eta_k q_n (1 + o_n(1)),$$

where the distribution $\{p_k\}_{k \in \mathbb{N}}$ is the same as the asymptotic degree distribution of the RAN_d *given in (12). Further, the random sequence $\{\eta_k\}$ can be chosen to be nondecreasing with $\eta_k \leq C_0 k! \eta^k$ for a constant C_0 and a random variable $\eta < \infty$.*

The random variable η will be explicitly defined at the end of the proof of Lemma 9. We need one additional statement to be able to prove Theorem 5.

Claim 6. (*The order of magnitude of $N^*(n)$.*) There exists a random variable $\xi \geq d+1$ such that

$$\lim_{n \to \infty} N^*(n) \prod_{i=0}^{n-1} (1 + dq_i)^{-1} \to \xi \quad \text{a.s.,} \tag{52}$$

where $q_0 := 1$.

Proof. When a new vertex is added to the graph, the number of active cliques increases by $d + 1 - 1 = d$; thus, at time n there are

$$A(n) = dN^*(n) + d + 1 \tag{53}$$

active cliques given $N^*(n)$. Since each clique that is active at step n turns into an inactive clique (inactive vertex) with probability q_n at step n, the number of new inactive vertices after the $(n+1)$th step satisfies $N^*(n+1) - N^*(n) = \mathrm{bin}(dN^*(n) + d + 1, q_n)$ with $N^*(1) := \mathrm{bin}(d+1, q_1)$. As a result, $\mathbb{E}[N^*(n+1) \mid N^*(n)] = N^*(n)(1 + dq_n) + (d+1)q_n$ with $q_0 := 1$,

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

and it is elementary to show that the process

$$M'_n = N^*(n)\prod_{i=0}^{n-1}(1+dq_i)^{-1} - (d+1)\sum_{i=0}^{n-1}\biggl(q_i\prod_{j=0}^{i}(1+dq_j)^{-1}\biggr)$$

is a square-integrable martingale if $\sum_{n\in\mathbb{N}} q_n = \infty$ and so there exists a random variable $M'_\infty \geq 0$ such that $M'_n \to M'_\infty$ a.s. Equation (52) follows with $\xi := M'_\infty + d + 1 + K$ if we can show that $\sum_{i=1}^{\infty}(q_i\prod_{j=0}^{i}(1+dq_j)^{-1})$ converges to a constant $K \geq 0$. For this, since $1/(1+\mathrm{d}x) \leq \exp\{-x\}$ as long as $x \leq 1$,

$$\sum_{i=1}^{\infty}\biggl(q_i\prod_{j=0}^{i}(1+dq_j)^{-1}\biggr) \leq \sum_{i=1}^{\infty} q_i\exp\biggl\{-\sum_{j=1}^{i}q_j\biggr\}.$$

The right-hand side converges since $\sum_{n=1}^{\infty} q_n \to \infty$. □

Proof of Theorem 5. Recall $c^\star$ and $\varepsilon(c^\star)$ from Lemma 8, and η_k from Lemma 9. Claim 6 implies that $N^*(n) = \Theta(\prod_{i=0}^{n-1}(1+dq_i))$ a.s.; hence, for each fixed k there exists a random integer $n_0(k)$ such that for all $n > n_0(k)$ we have $c^\star\sqrt{\log N^*(n)/N^*(n)} < \eta_k q_n$. Since the sequence η_i is nondecreasing, by the triangle inequality and the union bound, for all $n > n_0(k)$,

$$\mathbb{P}\Bigl(\max_{i\leq k}|p_i(n) - p_i| \geq 3\eta_k q_n\Bigr) \leq \mathbb{P}\biggl(\max_{i\leq k}\biggl|p_i(n) - \frac{\mathbb{E}[N_i(n) \mid \mathcal{G}_n]}{N^*(n)}\biggr| \geq \eta_k q_n\biggr) + \mathbb{P}\biggl(\max_{i\leq k}\biggl|\frac{\mathbb{E}[N_i(n) \mid \mathcal{G}_n]}{N^*(n)} - p_i\biggr| \geq 2\eta_k q_n\biggr).$$

By Lemmas 8 and 9 both terms on the right-hand side tend to 0; hence, the statement of Theorem 5 follows. □

Now we prove Lemmas 8 and 9. To do this, the following observations will be useful. In a similar way as we derived (53) one can show that an inactive vertex with degree $k \geq d+1$ is contained in exactly

$$Q_k = 2 + (k-d)(d-1) \tag{54}$$

active cliques. □

Proof of Lemma 8. We prove Lemma 8 using the Azuma–Hoeffding inequality in an elaborate way. Use the notation $K(n) := \sqrt{N^*(n)\log N^*(n)}$, and recall the σ-algebra $\mathcal{G}_n$. We aim to show that there exists a constant $c > 0$ such that

$$\mathbb{P}\Bigl[\max_k |N_k(n) - \mathbb{E}[N_k(n) \mid \mathcal{G}_n]| \geq cK(n) \Bigm| \mathcal{G}_n\Bigr] = o(N^*(n)^{-\varepsilon}). \tag{55}$$

Taking conditional expectation with respect to $N^*(n)$ of both sides immediately gives Lemma 8. First note that at step n the maximal degree of any inactive vertex is $N^*(n)+d-3$. Thus, the left-hand side of (55) is at most

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

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

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

$$\mathbb{P}(|N_k(n) - \mathbb{E}[N_k(n) \mid \mathcal{G}_n]| \ge c^\star K(n)) = o(N^*(n)^{-(1+\varepsilon)}). \tag{56}$$

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

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

where k is fixed. Clearly $M_{1,0} = \mathbb{E}[N_k(n) \mid \mathcal{G}_n]$ 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

$$|M_{r,s} - M_{r,s-1}| \le 2(d+1) \text{ for all } r \in \{1, \dots, n\}, \text{ for all } 1 \le s < A(r) \le (d+1)^r. \tag{57}$$

From the definition of $M_{r,s}$, we see that the difference is caused by the extra information whether the sth coin flip raises a new vertex or not. Let us consider the two different realizations, i.e. in the $\mathrm{EAN}(r,s)_a$ a new vertex $v_{r,s}$ is added to the graph at the sth coin flip but not in the $\mathrm{EAN}(r,s)_b$. Note that the number $N^*(r+1) - N^*(r)$ of new vertices 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 vertex $v_{r,s'}$ will be added in the $\mathrm{EAN}(r,s)_b$ but not in the $\mathrm{EAN}(r,s)_a$. 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 inactive vertices are the same* and *every vertex has the same degree except* for the $d+1$ neighbours of $v_{r,s}$ in the $\mathrm{EAN}(r,s)_a$ and the $d+1$ neighbours of $v_{r,s'}$ in the $\mathrm{EAN}(r,s)_b$. Since the degree of vertices that were added 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 $2(d+1)$ inactive vertices with different degrees. Thus, taking expectation with respect to $\mathcal{F}_{r,s-1}$ conserves this difference, which implies (57). (This argument is somewhat similar to [30, Lemma 8.5], except here a somewhat different conditioning is needed.)

We have just proved that the martingale $M_{r,s}$ has bounded increments. Observe that every new vertex will create $d+1$ new active cliques and induce $d+1$ coin flips. Thus, there are at most $(d+1)N^*(n)$ coin flips until time n and so $|M_{r,s} - M_{r,s-1}| \ne 0$ only at most $(d+1)N^*(n)$ times (and the number of nonzero coin flips is measurable with respect to $\mathcal{G}_n$). Hence, using the Azuma–Hoeffding inequality, we have, for any $a > 0$,

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

(Note that both side are random variables and the statement holds a.s., also $\mathcal{G}_n$ contains $N^*(n)$.) Now set $a = c^\star K(n)$, $c^\star > \sqrt{24}(d+1)$ (and, therefore, $c^\star > \sqrt{24}(d+1)(1+\varepsilon)^{1/2}$ for some $\varepsilon > 0$) to obtain

$$\mathbb{P}(|N_k(n) - \mathbb{E}[N_k(n) \mid \mathcal{G}_n]| \ge c^\star K(n) \mid \mathcal{G}_n) \le 2N^*(n)^{-(c^\star)^2/24(d+1)^2} \le o(N^*(n)^{-(1+\varepsilon)}).$$

Note that this bound is *uniform in* k; hence, (56) and (55) follow. □

For the proof of Lemma 9, we use the following proposition.

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

Proposition 3. *Let us introduce the nth* empirical occupation parameter

$$\hat{q}_n := \frac{N^*(n+1) - N^*(n)}{dN^*(n) + d + 1} = \frac{1}{A(n)} \sum_{i=1}^{A(n)} \mathbf{1}_{\{the\ ith\ active\ vertex\ at\ step\ n\ becomes\ inactive\}}.$$

Then, $\hat{q}_n/q_n \to 1$ a.s. as long as (14) of Assumption 1 is satisfied.

Proof. Indeed, $\hat{q}_n$ is the empirical occupation parameter of a $\text{bin}(A(n), q_n)$ distribution, so it is reasonable to assume that if n is large enough, $\hat{q}_n$ will tend to the true parameter, q_n. Introduce the event $\mathcal{E}_n(\varepsilon) := \{|\hat{q}_n/q_n - 1| \geq \varepsilon\}$. We can prove $\hat{q}_n/q_n \to 1$ a.s. if we can show that for every $\varepsilon > 0$, $\mathcal{E}_n(\varepsilon)$ happens only for finitely many n. For this, we use a Chernoff bound conditional on $A(n)$ (see, e.g. [30, Theorem 2.21]):

$$\begin{aligned}\mathbb{P}(\mathcal{E}_n(\varepsilon) \mid A(n)) &= \mathbb{P}(|\text{bin}(A(n), q_n) - A(n)q_n| > \varepsilon q_n A(n) \mid A(n)) \\ &\leq \exp\left\{-\frac{A(n)q_n\varepsilon^2}{2}\right\}. \end{aligned} \tag{58}$$

Note that by Claim 6, $A(n)$ is of order $\exp\{d\sum_{i=1}^{n} q_i\}$ a.s.; hence, the right-hand side is summable for any $\varepsilon > 0$ if and only if (14) of Assumption 1 is satisfied. Hence, by the Borel–Cantelli lemma, in these cases we obtain $\hat{q}_n/q_n \to 1$ a.s. □

Proof of Lemma 9. We aim to write a recursion for $\mathbb{E}[N_k(n) \mid \mathcal{G}_n]$. Note that $A(n+1)$, the number of active vertices (by (53)) and the number of new inactive vertices at step n, $N^*(n+1) - N^*(n)$ are both measurable with respect to $\mathcal{G}_{n+1}$.

Given that there are $N^*(n+1) - N^*(n)$ successes in $A(n)$ Bernoulli trials, the places of these successful trials are distributed u.a.r. To compute $\mathbb{E}[N_k(n+1) \mid \mathcal{G}_{n+1}]$ in terms of $\mathbb{E}[N_k(n) \mid \mathcal{G}_n]$ we have to take into account the following three events that result in a vertex of degree k at $n+1$.

(i) A degree of an inactive vertex can increase to k. An inactive vertex with degree $k - \ell$, $\ell = 1, \dots, ((d-1)/d)k$ is contained in $Q_{k-\ell}$ many active cliques, hence, conditioned on $N^*(n+1) - N^*(n)$, $A(n) \in \mathcal{G}_n$, the indicators which are 1 form a uniform subset of all $A(n)$ active cliques. Then, the probability that out of $Q_{k-\ell}$ many cliques exactly ℓ become inactive and the rest do not, is given by the hypergeometric distribution

$$\begin{aligned}&\binom{Q_{k-\ell}}{\ell} \prod_{i=0}^{\ell-1} \frac{N^*(n+1) - N^*(n) - i}{A(n) - i} \prod_{j=0}^{Q_{k-\ell}-\ell-1} \frac{A(n) - (N^*(n+1) - N^*(n)) - j}{A(n) - \ell - j} \\ &\quad \leq \binom{Q_{k-\ell}}{\ell} \hat{q}_n^{\ell} (1 - \hat{q}_n)^{Q_{k-\ell}-\ell} (1 + o_n(1)),\end{aligned} \tag{59}$$

as long as $A(n) \to \infty$, which holds again by Claim 6 and $\sum_{n=1}^{\infty} q_n \to \infty$. Further, the inequality can be replaced by an equality for $\ell = 1$, since in these cases $N^*(n+1) - N^*(n) - i$ cannot become 0 for any i in the first product of (59) and $A(n) \to \infty$. For $\ell \geq 2$, 0 serves as a lower bound on the left-hand side.

(ii) An inactive vertex with degree k at n can conserve its degree. Since an inactive vertex with degree k is contained in Q_k active cliques (see (54)) the degree stays k from step n to $n+1$ with $\mathcal{G}_n$-conditional probability that is given by setting $\ell = 0$ in (59), which is $(1 - \hat{q}_n)^{Q_k}(1 + o_n(1))$.

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

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

Note that the factor $(1 + o_n(1))$ in (59) also depends on k and ℓ. However, one can show that for all $k \geq d + 1$ and all $\ell = 1, \ldots, ((d - 1)/d)k$ this term is bounded from above by

$$\psi_k^+(n) := \left(1 + \frac{2Q_k}{A(n)}\right)^{Q_k},$$

and for $\ell = 0$ it is bounded above by 1. Hence, we can write the following conditional recursion as an upper bound:

$$\begin{aligned} &\mathbb{E}[N_k(n+1) \mid \mathcal{G}_{n+1}] \\ &\quad \leq \mathbb{E}[N_k(n) \mid \mathcal{G}_n](1 - \hat{q}_n)^{Q_k} \\ &\qquad + \psi_k^+(n)\bigg(\mathbb{E}[N_{k-1}(n) \mid \mathcal{G}_n]Q_{k-1}\hat{q}_n(1 - \hat{q}_n)^{Q_{k-1}-1} \\ &\qquad\qquad + \sum_{\ell=2}^{(d-1)k/d} \mathbb{E}[N_{k-\ell}(n) \mid \mathcal{G}_n]\binom{Q_{k-\ell}}{\ell}\hat{q}_n^{\ell}(1 - \hat{q}_n)^{Q_{k-\ell}-\ell}\bigg) \\ &\qquad + \mathbf{1}_{\{k=d+1\}}(N^*(n+1) - N^*(n)). \end{aligned} \tag{60}$$

A similar lower bound can also be given as

$$\begin{aligned} \mathbb{E}[N_k(n+1) \mid \mathcal{G}_{n+1}] \geq \psi_k^-(n)&\mathbb{E}[N_k(n) \mid \mathcal{G}_n](1 - \hat{q}_n)^{Q_k} \\ &+ \mathbb{E}[N_{k-1}(n) \mid \mathcal{G}_n]Q_{k-1}\hat{q}_n(1 - \hat{q}_n)^{Q_{k-1}-1} \\ &+ \mathbf{1}_{\{k=d+1\}}(N^*(n+1) - N^*(n)), \end{aligned} \tag{61}$$

where $\psi_k^-(n) = (1 - Q_k/(A(n) - (N^*(n+1) - N^*(n))))^{Q_k}$. Now, we first find the 'stationary solution' of this recursion in the form $\mathbb{E}[N_k(n) \mid \mathcal{G}_n] = p_k N^*(n)$. Recall that $\hat{q}_n \to 0$, so a series expansion in the first term on the right-hand side yields that the limiting distribution p_k should satisfy

$$\begin{aligned} &p_k(N^*(n+1) - N^*(n)) \\ &\quad \leq -p_k N^*(n)Q_k\hat{q}_n + p_{k-1}N^*(n)\hat{q}_n Q_{k-1} + O(N^*(n)\hat{q}_n^2(Q_k^2 + Q_{k-1})) \\ &\qquad + \sum_{\ell=2}^{(d-1)k/d} p_{k-\ell}N^*(n)\hat{q}_n\binom{Q_{k-\ell}}{\ell}\hat{q}_n^{\ell-1} + \mathbf{1}_{\{k=d+1\}}(N^*(n+1) - N^*(n)). \end{aligned}$$

Multiplying both sides by $(N^*(n)\hat{q}_n)^{-1}$, and using the fact that $N^*(n+1) - N^*(n) = \hat{q}_n A(n)$, we have

$$p_k\frac{A(n)}{N^*(n)} = -p_k Q_k + p_{k-1}Q_{k-1} + \mathbf{1}_{\{k=d+1\}}\frac{A(n)}{N^*(n)} + O(C_k\hat{q}_n),$$

where $C_k < k^2d^2$ by (54), and it estimates the smaller-order terms. Note that the inequality became an equality since similar analysis on the lower bound in (61) yields the same right-hand side. By Proposition 3, we can write $\hat{q}_n = q_n(1 + o(1)) \to 0$ as $n \to 0$. Combining this with $\lim_{n\to\infty} A(n)/N^*(n) = d$ (by (53)), we find that the limiting distribution p_k should satisfy

$$p_k(d + Q_k) = p_{k-1}Q_{k-1} + d\,\mathbf{1}_{\{k=d+1\}}. \tag{62}$$

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

Using the formula for Q_k in (35), we equivalently have

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

The solution of this recursion is

$$\begin{aligned} p_k &= p_{d+1}\prod_{\ell=d+2}^{k}\frac{\ell-1-d+2/(d-1)}{\ell-d+(d+2)/(d-1)} \\ &= \frac{d}{2d+1}\frac{\Gamma(k-d+2/(d-1))}{\Gamma(1+2/(d-1))}\frac{\Gamma(2+(d+2)/(d-1))}{\Gamma(k+1-d+(d+2)/(d-1))}, \end{aligned} \tag{63}$$

and, hence, by the properties of the gamma function, we obtain

$$p_k \sim \text{constant } k^{-(2d-1)/(d-1)},$$

i.e. the 'stationary solution' has a power-law decay with exponent in $(2, 3]$ for $d \geq 2$.

Next we analyse the convergence of $\mathbb{E}[N_k(n) \mid \mathcal{G}_n]/N^*(n)$ to p_k. For this, we need to show that

$$\varepsilon_k(n) := \mathbb{E}[N_k(n) \mid \mathcal{G}_n] - p_k N^*(n) \tag{64}$$

is of the order claimed in Lemma 9 (as $n \to \infty$), conditionally on $\mathcal{G}_n$. Using (60) and (62) it is elementary to check that the following recursion holds for the error terms defined in (64):

$$\begin{aligned} \varepsilon_k(n+1) \leq{}& \varepsilon_k(n)(1-\hat{q}_n)^{Q_k} + \varepsilon_{k-1}(n)Q_{k-1}\hat{q}_n(1-\hat{q}_n)^{Q_{k-1}-1}\psi_k^+(n) \\ &- p_k(N^*(n+1) - N^*(n)(1-\hat{q}_n)^{Q_k} - (d+Q_k)\hat{q}_n N^*(n)) \\ &- p_{k-1}N^*(n)Q_{k-1}\hat{q}_n(1-(1-\hat{q}_n)^{Q_{k-1}-1}\psi_k^+(n)) \\ &+ \mathbf{1}_{\{k=d+1\}}(A(n) - dN^*(n))\hat{q}_n \\ &+ \psi_k^+(n)\sum_{\ell=2}^{((d-1)/d)k}\mathbb{E}[N_{k-\ell}(n) \mid \mathcal{G}_n]\binom{Q_{k-\ell}}{\ell}\hat{q}_n^{\ell}(1-\hat{q}_n)^{Q_{k-\ell}-\ell}. \end{aligned} \tag{65}$$

Denote by $\Delta_k^+(n)$ the absolute value of the sum of all terms but the first one on the right-hand side of (65). Analogously, we denote by $\Delta_k^-(n)$ the absolute value of the sum of all terms but the first and last one. To estimate $\varepsilon_k(n)$ we use induction in k. Suppose that for all ℓ, n such that $\ell \leq k-1$, we have

$$|\varepsilon_\ell(n)| \leq \eta_\ell N^*(n)\hat{q}_n \tag{66}$$

with random variables $\eta_\ell < \infty$ that will be defined at the end of the proof. The induction clearly starts for $\ell = d$ since $\varepsilon_d(n) = 0$ for all n. To advance the induction, observe that

$$\varepsilon_k(n+1) \leq |\varepsilon_k(n)| + \Delta_k^+(n).$$

This inequality holds since when $\varepsilon_k(n) \leq 0$, by (65) we have $\varepsilon_k(n+1) \leq \Delta_k^+(n)$. Otherwise, when $\varepsilon_k(n+1) \leq \varepsilon_k(n)$, the inequality is immediate. There remains the $\varepsilon_k(n+1) > \varepsilon_k(n) \geq 0$ case. In this case the inequality follows from $(1-\hat{q}_n)^{Q_k}\psi_k^+(n) \leq 1$ for a large enough n, since $\hat{q}_n \gg 2Q_k/A(n)$. Similarly, it is easy to see that the lower bound $\varepsilon_k(n+1) \geq -|\varepsilon_k(n)| - \Delta_k^-(n)$ holds. Therefore,

$$|\varepsilon_k(n+1)| \leq |\varepsilon_k(0)| + \sum_{i=1}^{n}\Delta_k(i), \tag{67}$$

where $\Delta_k(i) = \max\{\Delta_k^+(i), \Delta_k^-(i)\}$.

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

With a series expansion in the second, third, and fourth terms in (65), and using the fact that $N^*(n+1) - N^*(n) = \hat{q}_n A(n)$ and the identity $A(n) - dN^*(n) = d+1$ yields the upper bound for

$$\Delta_k^+(n) \le \bigg|\hat{q}_n\bigg(\varepsilon_{k-1}(n)Q_{k-1}\psi_k^+(n) + (d+1)(\mathbf{1}_{\{k=d+1\}} - p_k) + \frac{p_{k-1}Q_{k-1}2Q_k^2}{d}\bigg) + \hat{q}_n^2 N^*(n)\bigg(p_k Q_k^2 + p_{k-1}Q_{k-1}^2 + O(\hat{q}_n) + \sum_{\ell=2}^{((d-1)/d)k} \bar{p}_{k-\ell}(n)\binom{Q_{k-\ell}}{\ell}\hat{q}_n^{\ell-2}(1+o_n(1))\bigg)\bigg|, \tag{68}$$

where we used $\mathbb{E}[N_k(n) \mid \mathcal{G}_n]/N^*(n) = \bar{p}_{k-\ell}(n)(1+o_n(1))$, with the definition $\bar{p}_k(n) := \mathbb{E}[N_k(n) \mid \mathcal{G}_n]/(N^*(n)+d+2)$, that sums up to 1 in k.

A similar inequality holds for $\Delta_k^-(n)$ (without the last sum on the right-hand side of (68)), so we arrive at

$$\Delta_k(n) \le \psi_k^+(n)(c_{k,1}\hat{q}_n|\varepsilon_{k-1}(n)| + c_{k,2}\hat{q}_n + c_{k,3}\hat{q}_n^2 N^*(n)), \tag{69}$$

with $c_{k,1} := Q_{k-1},\ c_{k,2} := (p_k + \mathbf{1}_{\{k=d+1\}})(d+1) + p_{k-1}Q_{k-1}2Q_k^2/d$, and $c_{k,3} := p_k Q_k^2 + p_{k-1}Q_{k-1}^2 + 2^{Q_k}$.

The next inequality is an easy corollary of Assumption 1 and Claim 6. There exist strictly positive random variables ξ_1 and ξ_2 such that

$$\xi_1 \le \frac{\sum_{i=1}^{n-1} N^*(i)\hat{q}_i^2}{\hat{q}_n N^*(n)} \le \xi_2 \le 17\xi^2. \tag{70}$$

holds a.s. Indeed, from Claim 6 it follows that there exists a random constant K such that for all $i > K$, $N^*(i)/\prod_{j=0}^{i}(1+dq_j) \in (\xi/2, 2\xi)$ and $\hat{q}_i/q_i \in (\frac{1}{2}, 2)$ (the latter holds a.s. by Assumption 1 and the argument below (58)). As a result

$$\frac{\sum_{i=K}^{n-1} N^*(i)\hat{q}_i^2}{\sum_{i=K}^{n-1} q_i^2 \prod_{j=1}^{i}(1+dq_j)} \in \left[\frac{\xi}{4}, 4\xi\right], \qquad \frac{N^*(n)\hat{q}_n}{q_n \prod_{j=1}^{n}(1+dq_j)} \in \left[\frac{\xi}{4}, 4\xi\right] \tag{71}$$

hold for all $n \ge K$. The statement of (70) then follows by noting that the sum of the terms indexed by j between 1 and K is a constant (we obtained the $17\xi^2$ as an upper bound on ξ_2 by possibly taking n even larger.)

The estimate in (70) is strong enough to complete the induction step of the proof of Lemma 9. The induction hypothesis in (66) together with (67) and (69) yields

$$\begin{aligned}\varepsilon_k(n) &\le |\varepsilon_k(0)| + \sum_{i=1}^{n-1}\psi_k^+(i)(c_{k,1}\eta_{k-1}\hat{q}_i^2 N^*(i) + c_{k,2}\hat{q}_i + c_{k,3}\hat{q}_i^2 N^*(i)) \\ &\le |\varepsilon_k(0)| + C + \sum_{i=1}^{n-1} 2(c_{k,1}\eta_{k-1}\hat{q}_i^2 N^*(i) + c_{k,2}\hat{q}_i + c_{k,3}\hat{q}_i^2 N^*(i)),\end{aligned}$$

where we used the fact that $\psi_k^+(i) \le 2$ if n is sufficiently large. Next, we can apply the upper bound in (70) on the sum of the first and third terms to obtain, for n large enough,

$$\varepsilon_k(n) \le (2c_{k,1}\eta_{k-1} + c_{k,3})\xi_2\hat{q}_n N^*(n) + 2c_{k,2}\sum_{i=1}^{n-1}\hat{q}_i + |\varepsilon_k(0)|.$$

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

We can advance the induction by noting that $|\varepsilon_k(0)| \le d+2$, while $\sum_{i=1}^{n-1} \hat{q}_i = o(N^*(n)q_n)$ follows from the second statement in (71). Further, by (70) and the definitions of $c_{k,i}$, we see that η_k can be chosen to satisfy the recursion $\eta_k = \xi_2(2Q_{k-1}\eta_{k-1} + 2^{k(d-1)+1} + 2)$. Using $Q_k < k(d-1)$ and $p_k \sim \text{constant}\, k^{(2d-1)/(d-1)}$ (and also using the fact that we can chose $\eta_d = 0$), we obtain $\eta_k \le k!\,\eta^k$ for $\eta := 2(d-1)\xi_2$ and $k > k_0$, where k_0 is deterministic. Therefore, $\eta_k \le C_0 k!\,\eta^k$ holds for all k for some C_0 large enough. □

In the proof of Lemma 1, we will repeatedly use the following theorem, see [28] and [9, Proposition 1.5.8].

Theorem 6. (Karamata's theorem, direct half.) *Let $L(x)$ be a slowly varying function at ∞ and let $\beta > 0$. Then for any fixed $x_0 > 0$,*

$$\lim_{x\to\infty} \frac{1}{x^\beta L(x)} \int_{x_0}^{x} \frac{L(t)}{t^{1-\beta}}\,\mathrm{d}t = \lim_{x\to\infty} \frac{1}{x^{-\beta} L(x)} \int_{x}^{\infty} \frac{L(t)}{t^{\beta+1}}\,\mathrm{d}t = \frac{1}{\beta}.$$

Further, the function $\widetilde{L}(x) := \int_{x_0}^{x} (L(t)/t)\,\mathrm{d}t$ is slowly varying at ∞ and

$$\lim_{x\to\infty} \frac{1}{L(x)} \int_{x_0}^{x} \frac{L(t)}{t}\,\mathrm{d}t = \infty.$$

Proof of Lemma 1. (i) We start with the $\alpha \ne 1$ case by showing that (13) holds. In this proof, C denotes a generic constant with a value that might change even along lines, but only depends on the sequence $\{q_n\}_{n\in\mathbb{N}}$ and nothing else. Further, the Landau symbol Θ has its usual meaning, i.e. for two sequences a_n and b_n, we say that $a_n = \Theta(b_n)$ if there exist constants $0 < c$ and $C < \infty$ such that for all n, $cb_n < a_n < Cb_n$. Using the inequalities $x - x^2/2 \le \log(1+x) \le x + x^2/2$, we obtain

$$\prod_{j=1}^{i}(1+dq_j) = \exp\biggl\{\sum_{j=1}^{i}\log(1+dq_j)\biggr\} = \exp\biggl\{d\sum_{j=1}^{i} q_j(1+o_i(1))\biggr\}, \tag{72}$$

where we use the fact that Theorem 6 gives $\sum_{j=1}^{i} q_j = \Theta(i^{1-\alpha}L(i))$, while $\sum_{j=1}^{i} q_j^2 = \Theta(i^{1-2\alpha}L^2(i))$ if $\alpha \in (0, \frac{1}{2})$, slowly varying if $\alpha = \frac{1}{2}$ and summable if $\alpha \in (\frac{1}{2}, 1)$, and, hence, the sum of the second-order error terms can be substituted into the $o_i(1)$ term in (72). This estimate implies that

$$q_i^2 \prod_{j=1}^{i}(1+dq_j) = \frac{L(i)^2}{i^{2\alpha}} \exp\{\Theta(i^{1-\alpha}L(i))\} \to \infty,$$

and, hence, the sum of the first $\lfloor n/2 \rfloor$ terms add only at most a constant factor to the total sum

$$\sum_{i=1}^{n-1} q_i^2 \prod_{j=1}^{i}(1+dq_j) \le C \sum_{i=\lceil n/2\rceil}^{n-1} q_i^2 \prod_{j=1}^{i}(1+dq_j) \le Cq_n \sum_{i=\lceil n/2\rceil}^{n-1} q_i \prod_{j=1}^{i}(1+dq_j),$$

where, to obtain the second inequality we use the fact that for all $j \in \{\lceil n/2\rceil, \dots n\}$, we have $q_j = L(j)/j^\alpha \le cL(n)/n^\alpha = Cq_n$ for some $C > 0$, since $L(\cdot)$ is slowly varying (this statement follows from a usual Potter's bound; see [9, Theorem 1.5.6]). It remains to study the following sum:

$$\sum_{i=\lceil n/2\rceil}^{n-1} q_i \prod_{j=1}^{i}(1+dq_j) = \sum_{i=n/2}^{n-1} q_i \exp\biggl\{d\sum_{j=1}^{i} q_j(1+o_i(1))\biggr\}.$$

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

Let $q(x)$ be a continuous function so that $q(n) = q_n$, and approximate the sum with the integral, we see that the right-hand side is at most

$$C \int_{n/2}^{n} \exp\left\{ d \int_{1}^{x} q(y)\,\mathrm{d}y (1 + o_x(1)) \right\} q(x)\,\mathrm{d}x. \tag{73}$$

From Theorem 6 and the argument below (72) it follows that the term hidden in $o_x(1)$ is $\Theta(x^{-\alpha} L(x))$ when $\alpha \neq \frac{1}{2}$ and $\widetilde{L}(x)/x^{1/2}$ for some other slowly varying function $\widetilde{L}(x)$ when $\alpha = \frac{1}{2}$. Hence, it is not difficult to see (by, e.g. estimating with two different constant factors in the exponent) that the $(1 + o_x(1))$ factor can be neglected and it follows that the integral in (73) can be bounded from above by

$$C \exp\left\{ d \int_{1}^{n} q(y)\,\mathrm{d}y \right\} \leq C \prod_{i=1}^{n} (1 + dq_i),$$

where in the last step we moved back from the integral to the sum again, similarly as before. (Note that we used the fact that $\sum_{i=1}^{n-1} q_i \to \infty$ as well, which follows trivially when $\alpha \in (0, 1)$). The lower bound can be determined in a similar manner. For the second statement, (14), a simple application of Theorem 6 is sufficient,

$$\sum_{n=1}^{\infty} \exp\left\{ -\varepsilon q_n \exp\left\{ d \sum_{j=0}^{n} q_j \right\} \right\} = \sum_{n=1}^{\infty} \exp\left\{ -\varepsilon \frac{L(n)}{n^{\alpha}} \exp\{ d L(n) n^{1-\alpha} \} \right\},$$

which is summable for every $\varepsilon > 0$ since $1 - \alpha > 0$.

(ii) We show the upper bound in (13). The lower bound can be determined in a similar manner. The same argument to that used around (72) can be repeated and since $\sum_{j=1}^{\infty} b^2/j^2 < C$, we have

$$q_i^2 \prod_{j=1}^{i} (1 + dq_j) \leq \frac{(b + o(i^{-\delta}))^2}{i^2} \exp\{ db \log i + C \} \leq C i^{db-2}, \tag{74}$$

where we also use the fact that the error term $\sum_{j=1}^{i} o(n^{-1-\delta}) < C$ for some constant $C > 0$ depends on the sequence $\{q_n\}_{n \in \mathbb{N}}$ only. Finally, summing (74) in i, the first i_0 terms contribute as a constant, and the rest can be expressed as

$$\sum_{i=1}^{n-1} q_i^2 \prod_{j=1}^{i} (1 + dq_j) \leq C \sum_{i=i_0}^{n-1} i^{db-2} \leq C n^{db-1}, \tag{75}$$

since $db > 1$. On the other hand, the denominator in (13) can be estimated from below as

$$q_n \prod_{j=1}^{n} (1 + dq_j) \geq \frac{b - o(n^{-\delta})}{n} \exp\left\{ d \sum_{j=1}^{n} \frac{b - o(j^{-\varepsilon})}{j} - C \right\} \geq C n^{db-1}. \tag{76}$$

Combining the estimates in (75) and (76) yields the upper bound in (13). The lower bound can be shown analogously and is left to the reader. As for (14), we can use the lower bound (76) to obtain the following upper bound:

$$\exp\left\{ -\varepsilon q_n \exp\left\{ d \sum_{j=1}^{n} q_j \right\} \right\} \leq \exp\{ -\varepsilon C n^{db-1} \},$$

which is summable since $db - 1 > 0$. □

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

Acknowledgements

This work was partially supported by the TÁMOP-4.2.2.C-11/1 /KONV-2012-0001** project. The project has been supported by the European Union and co-financed by the European Social Fund. This research was supported in part by the grant KTIA-OTKA # CNK 77778, funded by the Hungarian National Development Agency (NFÜ) from a source provided by KTIA.

The work of JK is partially financed by the STAR cluster, and also part of the research programme Veni (project number 639.031.447), which are (partly) financed by The Netherlands Organisation for Scientific Research (NWO).

We thank the anonymous referee for the useful suggestions that greatly improved the presentation of the paper.

References

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

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

[3] Athreya, K. B. and Ney, P. E. (2004). *Branching Processes*. Dover, Mineola, NY.

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

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

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

[7] Bhamidi, S., van der Hofstad, R. and Hooghiemstra, G. (2011). First passage percolation on the Erdős–Rényi random graph. *Combin. Prob. Comput.* **20,** 683–707.

[8] Bhamidi, S., van der Hofstad, R. and Hooghiemstra, G. (2012). Universality for first passage percolation on sparse random graphs. Preprint. Available at http://arxiv.org/abs/1210.6839.

[9] Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1989). *Regular Variation* (Encyclopedia Math. Appl. **27**). Cambridge University Press.

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

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

[12] Bollobás, B., Borgs, C., Chayes, J. and Riordan, O. (2003). Directed scale-free graphs. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, ACM, New York, pp. 132–139.

[13] Bollobás, B., Riordan, O., Spencer, J. and Tusnády, G. (2001). The degree sequence of a scale-free random graph process. *Random Structures Algorithms* **18,** 279–290.

[14] Boyd, D. W. (1982). The sequence of radii of the Apollonian packing. *Math. Comput.* **39,** 249–254.

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

[16] Bühler, W. J. (1971). Generations and degree of relationship in supercritical Markov branching processes. *Z. Wahrscheinlichkeitsth.* **18,** 141–152.

[17] Choi, V. and Golin, M. J. (2001). Lopsided trees. I. Analyses. *Algorithmica* **31,** 240–290.

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

[19] Darrasse, A. and Soria, M. (2007). Degree distribution of random Apollonian network structures and Boltzmann sampling. In *2007 Conference on Analysis of Algorithms, AofA 07*, Association of Discrete Mathematics and Theoretical Computer Science, Nancy, pp. 313–324.

[20] Dembo, A. and Zeitouni, O. (2010). *Large Deviations Techniques and Applications* (Stoch. Modelling Appl. Prob. **38**), 2nd edn. Springer, Berlin.

[21] Doye, J. P. K. and Massen, C. P. (2005). Self-similar disk packings as model spatial scale-free networks. *Phys. Rev. E* **71,** 016128.

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

[23] Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd edn. John Wiley, New York.

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press

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

[25] Graham, R. L. *et al.* (2003). Apollonian circle packings: number theory. *J. Number Theory* **100,** 1–45.

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

[27] Kapoor, S. and Reingold, E. M. (1989). Optimum lopsided binary trees. *J. Assoc. Comput. Mach.* **36,** 573–590.

[28] Karamata, J. (1930). Über die Hardy–Littlewoodschen Umkehrungen des Abelschen Stetigkeitssatzes. *Math. Z.* **32,** 319–320.

[29] Kolossváry, I. and Komjáthy, J. (2015). First passage percolation on inhomogeneous random graphs. *Adv. Appl. Prob.* **47,** 589–610.

[30] Van der Hofstad, R. (2016). Random graphs and complex networks. To appear in *Cambridge Series in Statistical and probabilistic Mathematics,* Vol. 1, Cambridge University Press, Cambridge.

[31] Zhang, Z. *et al.* (2008). Analytical solution of average path length for Apollonian networks. *Phys. Rev. E* **77,** 017102.

[32] Zhang, Z., Rong, L. and Comellas, F. (2006). High-dimensional random Apollonian networks. *Physica A* **364,** 610–618.

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

[34] Zhang, Z., Comellas, F., Fertin, G. and Rong, L. (2006). High-dimensional Apollonian networks. *J. Phys. A* **39,** 1811–1818.

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

https://doi.org/10.1017/apr.2016.32 Published online by Cambridge University Press