DEPTH-FIRST SEARCH PERFORMANCE IN A RANDOM DIGRAPH WITH GEOMETRIC DEGREE DISTRIBUTION

PHILIPPE JACQUET AND SVANTE JANSON

Abstract. We present an analysis of the depth-first search algorithm in a random digraph model with geometric outdegree distribution. We give also some extensions to general outdegree distributions. This problem posed by Donald Knuth in his next to appear volume of *The Art of Computer Programming* gives interesting insight in one of the most elegant and efficient algorithm for graph analysis due to Tarjan.

1. Introduction

The motivation of this paper is a new section in Donald Knuth's *The Art of Computer Programming* [5], which is dedicated to Depth-First Search (DFS) in a digraph. Briefly, the DFS starts with an arbitrary vertex, and explores the arcs from that vertex one by one. When an arc is found leading to a vertex that has not been seen before, the DFS explores the arcs from it in the same way, in a recursive fashion, before returning to the next arc from its parent. This eventually yields a tree containing all descendants of the the first vertex (which is the root of the tree). If there still are some unseen vertices, the DFS starts again with one of them and finds a new tree, and so on until all vertices are found. We refer to [5] for details as well as for historical notes. (See also S1–S2 in Section 4.) Note that the digraphs in [5] and here are multi-digraphs, where loops and multiple arcs are allowed. (Although in our random model they are few and usually not important.) The DFS algorithm generates a spanning forest (the *depth-first forest*) in the digraph, with all arcs in the forest directed away from the roots. Our main purpose is to study the distribution of the depth of vertices in the depth-first forest, starting with a random digraph G.

Furthermore, the DFS algorithm in [5] classifies the arcs in the digraph into the following five types, see Figure 1 for examples:

- *loops*;
- *tree arcs*, the arcs in the resulting depth-first forest;
- *back arcs*, the arcs which point to an ancestor of the current vertex in the current tree;
- *forward arcs*, the arcs which point to an already discovered descendant of the current vertex in the current tree;
- *cross arcs*, all other arcs (these point to an already discovered vertex which is neither a descendant nor an ancestor of the current vertex, and might be in another tree).

We will study the numbers of arcs of different types. (See further the exercises in [5].)

The random digraph model that we consider has n vertices and a given outdegree distribution $\mathbf{P}$. The outdegrees (number of outgoing arcs) of the n vertices are independent random numbers with this distribution. The endpoint of each arc is uniformly selected at random among the n vertices, independently of all other arcs. (Therefore, an arc can loop back to the starting vertex, and multiple arcs can occur.) We consider asymptotics as $n \to \infty$ for a fixed outdegree distribution.

Date: 28 March 2022.

Key words and phrases. Combinatorics, Depth-First Search, Random Digraphs.

Supported by the Knut and Alice Wallenberg Foundation.

We thank Donald Knuth for posing us questions and conjectures that led to the present paper.

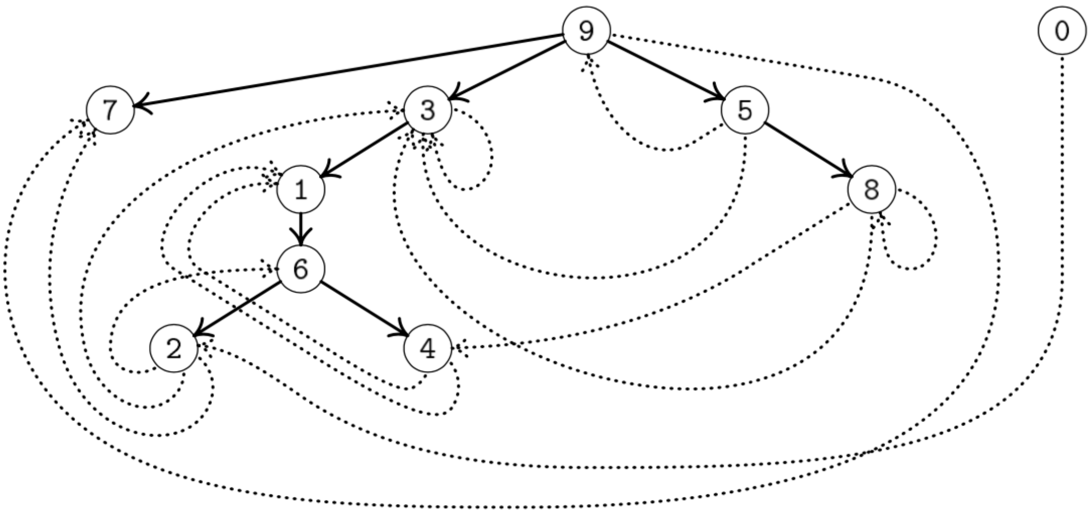

FIGURE 1. Example of a depth-first forest (jungle) from [5], by courtesy of Donald Knuth. Tree arcs are solid (*e.g.* ⑨→ ③). For example, ③--→ ③ is a loop, ②--→ ③ is a back arc, ⑨--→ ⑦ is a forward arc, ⑧--→ ④ and ⓪--→ ② are cross arcs.

We will focus on the case of a geometric outdegree distribution; the lack-of-memory property in this case leads to interesting features and a simpler analysis. We describe in Section 4 how the results can be extended to general outdegree distributions using a variation of the method. The paper will study the following outdegree distributions in the following order:

- a geometric distribution;
- a shifted geometric distribution (starting from integer 1 instead of zero);
- a general distribution.

Remark 1. Related results are proved by [3] for DFS in an undirected Erdős–Rényi graph $G(n, \lambda/n)$; see also [2]. The main result of [3] shows convergence of the depth profile in the depth-first forest to a certain deterministic limit. Although this is for a different random graph model, DFS on $G(n, \lambda/n)$ is the same as DFS on the Erdős–Rényi digraph $D(n, \lambda/n)$, which is essentially the same as the digraph studied in the present paper with outdegree distribution $\mathrm{Po}(\lambda)$. Hence the result in [3] is essentially the special case $\mathbf{P} = \mathrm{Po}(\lambda)$ of our result for the depths. The proofs are quite different. DFS in the random digraph $D(n, p)$ has also been considered previously, for example in the proof of [6, Theorem 3].

This is an extended abstract of [4], which contains further results and complete proofs not included here; in particular we treat there general outdegree distributions in detail.

1.1. **Some notation.** We denote the given outdegree distribution by $\mathbf{P}$. We let η, often with subscripts, denote random variables with this distribution. In particular, we denote the outdegree of vertex v by $\eta(v)$. Recall that our standing assumption is that these outdegrees are i.i.d. (independent and identically distributed) with $\eta_v \sim \mathbf{P}$. We let v_t denote the t-th vertex found by the DFS, and simplify notation by letting $\eta_t := \eta(v_t)$ be its outdegree. It follows from the construction of the DFS that also the random variables η_t, $t = 1, \dots, n$ are i.i.d. with distribution $\mathbf{P}$; this fundamental property will be used repeatedly without further mention.

The mean outdegree, *i.e.*, the expectation $\mathbb{E}\,\eta$ of $\mathbf{P}$, is denoted by λ. We assume throughout, for technical reasons, that the second moment $\mathbb{E}\,\eta^2 < \infty$.

As usual, w.h.p. means *with high probability*, *i.e.*, with probability $1 - o(1)$. We use $\overset{\mathrm{p}}{\longrightarrow}$ for convergence in probability, and $\overset{\mathrm{d}}{\longrightarrow}$ for convergence in distribution of random variables.

Moreover, let (a_n) be a sequence of positive numbers, and X_n a sequence of random variables. We write $X_n = o_{\mathrm{p}}(a_n)$ if, as $n \to \infty$, $X_n/a_n \overset{\mathrm{p}}{\longrightarrow} 0$, *i.e.*, if for every $\varepsilon > 0$, we have $\mathbb{P}(|X_n| > \varepsilon a_n) \to 0$. Furthermore, $X_n = o_{L^2}(a_n)$ means $\mathbb{E}\big[|X_n/a_n|^2\big] \to 0$, and $X_n = O_{L^2}(a_n)$ means $\mathbb{E}\big[|X_n/a_n|^2\big] = O(1)$. Note that $X_n = o_{L^2}(a_n)$ implies $X_n = o_{\mathrm{p}}(a_n)$, and that $X_n = O_{L^2}(a_n)$ implies $X_n = o_{L^2}(\omega_n a_n)$ and thus $X_n = o_{\mathrm{p}}(\omega_n a_n)$, for any sequence $\omega(n) \to \infty$. Note also that $X_n = o_{L^2}(a_n)$ implies $\mathbb{E}\, X_n = o(a_n)$, and similarly for $O_{L^2}(a_n)$; thus error terms of this type implies immediately estimates for expectations and second moments. In particular, for the most common case below, $X_n = O_{L^2}(n^{1/2})$ is equivalent to $\mathbb{E}\, X_n = O(a_n)$ and $\operatorname{Var} X_n = O(n)$.

We define $\rho_0(x)$, for $x \geqslant 0$, as the largest solution in $[0, 1)$ to

$$1 - \rho_0 = e^{-x\rho_0}. \tag{1}$$

As is well known, $\rho_0(x)$ is the survival probability of a Galton–Watson process with $\mathrm{Po}(x)$ offspring distribution. We have $\rho_0(x) = 0$ for $x \leqslant 1$ and $0 < \rho_0(x) < 1$ for $x > 1$.

All logarithms are natural.

2. Depth analysis with geometric outdegree distribution

In this section we assume that the outdegree distribution is geometric $\mathrm{Ge}(1-p)$ for some fixed $0 < p < 1$, and thus has mean

$$\lambda := \mathbb{E}\, \eta = \frac{p}{1-p}. \tag{2}$$

When doing the DFS on a random digraph of the type studied in this paper, we generally reveal the outdegree of a vertex as soon as we find it. (See S1–S2 in Section 4.) However, for a geometric outdegree distribution, because of its lack-of-memory property, we do not have to immediately reveal the outdegree when we find a new vertex v. Instead, we only check whether there is at least one outgoing arc (probability p), and if so, we find its endpoint and explore this endpoint if it has not already been visited; eventually, we return to v, and then we check whether there is another outgoing arc (again probability p, by the lack-of-memory property), and so on. This will yield the important Markov property in the construction in the next subsection.

In the following, by a *future arc* from some vertex, we mean an arc that at the current time has not yet been seen by the DFS.

2.1. **Depth Markov chain.**

Our aim is to track the evolution of the search depth as a function of the number t of discovered vertices. Let v_t be the t-th vertex discovered by the DFS ($t = 1, \dots, n$), and let $d(t)$ be the depth of v_t in the resulting depth-first forest, *i.e.*, the number of tree edges that connect the root of the current tree to v_t. The first found vertex v_1 is a root, and thus $d(1) = 0$.

The quantity $d(t)$ follows a Markov chain with transitions ($1 \leqslant t < n$):

(i) $d(t + 1) = d(t) + 1$.
This happens if, for some $k \geqslant 1$, v_t has at least k outgoing arcs, the first $k - 1$ arcs lead to vertices already visited, and the kth arc leads to a new vertex (which then becomes v_{t+1}). The probability of this is

$$\sum_{k=1}^{\infty} p^k \Big(\frac{t}{n}\Big)^{k-1} \Big(1 - \frac{t}{n}\Big) = \frac{(1 - t/n)p}{1 - pt/n}. \tag{3}$$

(ii) $d(t + 1) = d(t)$, assuming $d(t) > 0$.
This holds if all arcs from v_t lead to already visited vertices, *i.e.*, (i) does not happen, and furthermore, the parent of v_t has at least one future arc leading to an unvisited vertex. These two events are independent. Moreover, by the lack-of-memory property,

the number of future arcs from the parent of v_t has the same distribution $\mathrm{Ge}(1-p)$. Hence, the probability that one of these future arcs leads to an unvisited vertex equals the probability in (3). The probability of (ii) is thus

$$\Bigl(1-\frac{(1-t/n)p}{1-pt/n}\Bigr)\frac{(1-t/n)p}{1-pt/n}. \tag{4}$$

(iii) $d(t+1)=d(t)-\ell$, assuming $d(t)>\ell\geqslant 1$.
This happens if all arcs from v_t lead to already visited vertices, and so do all future arcs from the ℓ nearest ancestors of v_t, while the $(\ell+1)$th ancestor has at least one future arc leading to an unvisited vertex. The argument in (ii) generalizes and shows that this has probability

$$\Bigl(1-\frac{(1-t/n)p}{1-pt/n}\Bigr)^{\ell+1}\frac{(1-t/n)p}{1-pt/n}. \tag{5}$$

(iv) $d(t+1)=d(t)-\ell$, assuming $d(t)=\ell\geqslant 0$.
By the same argument as in (ii) and (iii), except that the $(\ell+1)$th ancestor does not exist and we ignore it, we obtain the probability

$$\Bigl(1-\frac{(1-t/n)p}{1-pt/n}\Bigr)^{\ell+1}. \tag{6}$$

Note that (iv) is the case when $d(t+1)=0$ and thus v_{t+1} is the root of a new tree in the depth-first forest.

We can summarize (i)–(iv) in the formula

$$d(t+1)=\bigl(d(t)+1-\xi_t\bigr)^+, \tag{7}$$

where $x^+:=\max\{x,0\}$, and ξ_t is a random variable, independent of the history, with the distribution

$$\mathbb{P}(\xi_t=k)=(1-\pi_t)^k\pi_t,\quad k\geqslant 0,\qquad\text{with}\qquad \pi_t:=\frac{(1-t/n)p}{1-pt/n}=1-\frac{1-p}{1-pt/n}. \tag{8}$$

In other words, ξ_t has the geometric distribution $\mathrm{Ge}(\pi_t)$. Define

$$\widetilde{d}(t):=\sum_{i=1}^{t-1}(1-\xi_i), \tag{9}$$

and note that (9) is a sum of independent random variables. Then (7) and induction yield

$$d(t)=\widetilde{d}(t)-\min_{1\leqslant j\leqslant t}\widetilde{d}(j),\qquad 1\leqslant t\leqslant n. \tag{10}$$

Remark 2. Similar formulas have been used for other, related, problems with random graphs and trees, where trees have been coded as walks, see for example [1, Section 1.3]. Note that in our case, unlike e.g. [1], $\widetilde{d}(t)$ may have negative jumps of arbitrary size.

Remark 3. We can also express these relations using generating functions. Let $p(t,z)$ be the probability generating function $\mathbb{E}\,z^{\xi_t}$ of ξ_t, *i.e.*,

$$p(t,z):=\frac{(1-t/n)p}{1-pt/n}\sum_{\ell\geqslant 0}\left(1-\frac{(1-t/n)p}{1-pt/n}\right)^{\ell}z^{\ell}=\frac{(1-t/n)p}{1-pt/n-(1-p)z}, \tag{11}$$

and let $f(t,z):=E[z^{d(t)}]$. We then have the identity, equivalent to (7),

$$f(t+1,z)=\mathcal{N}\bigl[R(t,z)f(t,z)\bigr] \tag{12}$$

where $R(t,z):=p(t,1/z)z$ and $\mathcal{N}$ is the operator on power series in $z^{\pm1}$:

$$\mathcal{N}g(z)=\Pi^+g(z)+(\Pi^-g)(1) \tag{13}$$

where Π^+ is the operator which removes the strictly negative powers of z and Π^- is the operator which removes the nonnegative powers of z. Thus we have, since $f(1,z)=1$,

$$f(t+1,z)=\mathcal{N}R(t,z)\mathcal{N}R(t-1,z)\mathcal{N}\cdots\mathcal{N}R(1,z). \tag{14}$$

At least some of the results below can be derived using this formalism, but we will not employ it in the present paper.

2.2. **Main result for depth analysis.** Note first that (9) implies that the expectation of $\widetilde{d}(t)$ is

$$\mathbb{E}\big[\widetilde{d}(t)\big]=\sum_{i=1}^{t-1}(1-\mathbb{E}\,\xi_i)=\sum_{i=1}^{t-1}\Big(1-\frac{1-\pi_i}{\pi_i}\Big)=\sum_{i=1}^{t-1}\Big(1-\frac{1-p}{p(1-i/n)}\Big). \tag{15}$$

Let $\theta:=t/n$. We fix a $\theta^*<1$ and obtain that, uniformly for $\theta\leqslant\theta^*$, recalling (2),

$$\mathbb{E}\big[\widetilde{d}(t)\big]=\int_0^t\Big(1-\frac{1}{\lambda(1-x/n)}\Big)\,\mathrm{d}x+O(1)=n\widetilde{\ell}(\theta)+O(1), \tag{16}$$

where

$$\widetilde{\ell}(\theta):=\int_0^\theta\Big(1-\frac{1}{\lambda(1-x)}\Big)\,\mathrm{d}x=\theta+\frac{1}{\lambda}\log(1-\theta). \tag{17}$$

Note that the derivative $\widetilde{\ell}'(\theta)=1-\lambda^{-1}/(1-\theta)$ is (strictly) decreasing on $(0,1)$, *i.e.*, $\widetilde{\ell}$ is concave. Moreover, if $\lambda>1$ (*i.e.*, $p>\frac12$) which we call the *supercritical* case, then $\widetilde{\ell}'(0)>0$, and (17) shows that $\widetilde{\ell}(\theta)$ is positive and increasing for $\theta<\theta_0:=1-\lambda^{-1}=(2p-1)/p$. After the maximum at θ_0, $\widetilde{\ell}(\theta)$ decreases and tends to $-\infty$ as $\theta\nearrow 1$. Hence, there exists a $0<\theta_1<1$ such that $\widetilde{\ell}(\theta_1)=0$; we then have $\widetilde{\ell}(\theta)>0$ for $0<\theta<\theta_1$ and $\widetilde{\ell}(\theta)<0$ for $\theta>\theta_1$. We will see that in this case the depth-first forest w.h.p. contains a giant tree, of order and height both linear in n, while all other trees are small.

On the other hand, if $\lambda\leqslant 1$ (*i.e.*, $p\leqslant\frac12$) (the *subcritical* and *critical* cases), then $\widetilde{\ell}'(0)\leqslant 0$ and $\widetilde{\ell}(\theta)$ is negative and decreasing for all $\theta\in(0,1)$. In this case, we define $\theta_0:=\theta_1:=0$ and note that the properties just stated for $\widetilde{\ell}$ still hold (rather trivially). We will see that in this case w.h.p. all trees in the depth-first forest are small.

Note that in all cases, θ_1 is the largest solution in $[0,1)$ to

$$\log(1-\theta_1)=-\lambda\theta_1. \tag{18}$$

Remark 4. The equation (18) may also be written $1-\theta_1=\exp(-\lambda\theta_1)$, which shows that $\theta_1=\rho_0(\lambda)$, the survival probability of a Galton–Watson process with $\mathrm{Po}(\lambda)$ offspring distribution defined in (1).

We define $\widetilde{\ell}^+(\theta):=[\widetilde{\ell}(\theta)]^+$. Thus, by (17) and the comments above,

$$\widetilde{\ell}^+(\theta)=\begin{cases}\theta+\lambda^{-1}\log(1-\theta), & 0\leqslant\theta\leqslant\theta_1,\\ 0, & \theta_1\leqslant\theta\leqslant 1.\end{cases} \tag{19}$$

We can now state one of our main results. Proofs are given in the next subsection.

Theorem 5. *We have*

$$\max_{1\leqslant t\leqslant n}\big|d(t)-n\widetilde{\ell}^+(t/n)\big|=O_{L^2}(n^{1/2}). \tag{20}$$

Corollary 6. *The height* Υ *of the depth-first forest is*

$$\Upsilon:=\max_{1\leqslant t\leqslant n}d(t)=\upsilon n+O_{L^2}(n^{1/2}), \tag{21}$$

where

$$\upsilon=\upsilon(p):=\widetilde{\ell}^+(\theta_0)=\begin{cases}0, & 0<\lambda\leqslant 1,\\ 1-\lambda^{-1}-\lambda^{-1}\log\lambda, & \lambda>1\end{cases} \tag{22}$$

Moreover, we can show that the height Υ is asymptotically normally distributed. Details are given in the full paper [4].

Corollary 7. *The average depth* $\overline{d}$ *in the depth-first forest is*

$$\overline{d} := \frac{1}{n}\sum_{t=1}^{n} d(t) = \alpha n + O_{L^2}(n^{1/2}), \tag{23}$$

where

$$\alpha = \alpha(p) := \frac{1}{2}\theta_1^2 - \frac{1}{\lambda}\Big((1-\theta_1)\log(1-\theta_1) + \theta_1\Big) = \frac{\lambda-1}{\lambda}\theta_1 - \frac{1}{2}\theta_1^2. \tag{24}$$

We have $\alpha = 0$ *if and only if* $\lambda \geqslant 1$, i.e., $p \leqslant 1/2$.

Remark 8. When $p > \frac{1}{2}$, the height is thus linear in n, unlike many other types of random trees. This might imply a rather slow performance of algorithms that operate on the depth-first forest.

2.3. **Proofs.**

Proof of Theorem 5. Since (9) is a sum of independent random variables, $\widetilde{d}(t) - \mathbb{E}\,\widetilde{d}(t)$ $(t = 1, \dots, n)$ is a martingale, and Doob's inequality yields, for all $T \leqslant n$,

$$\mathbb{E}\big[\max_{t\leqslant T} |\widetilde{d}(t) - \mathbb{E}\,\widetilde{d}(t)|^2\big] \leqslant 4\,\mathbb{E}\big[|\widetilde{d}(T) - \mathbb{E}\,\widetilde{d}(T)|^2\big] = 4\sum_{i=1}^{T-1} \operatorname{Var}(\xi_i). \tag{25}$$

Fix $\theta^* < 1$, and assume, as we may, that $\theta^* > \theta_1$. Let $T^* := \lfloor n\theta^* \rfloor$, and consider first $t \leqslant T^*$. For $i < T^*$, we have $\operatorname{Var}\xi_i = O(1)$, and thus, for $T = T^*$, the sum in (25) is $O(T^*) = O(n)$. Consequently, (25) yields

$$\max_{t\leqslant T^*} \big|\widetilde{d}(t) - \mathbb{E}\,\widetilde{d}(t)\big| = O_{L^2}(n^{1/2}). \tag{26}$$

Hence, by (16),

$$M^* := \max_{t\leqslant T^*} \big|\widetilde{d}(t) - n\widetilde{\ell}(t/n)\big| = O_{L^2}(n^{1/2}). \tag{27}$$

For $t \leqslant T^*$, the definition of M^* in (27) implies

$$\Big|\min_{1\leqslant j\leqslant t} \widetilde{d}(j) - n \min_{1\leqslant j\leqslant t} \widetilde{\ell}(j/n)\Big| \leqslant M^*. \tag{28}$$

Moreover, for $t/n \leqslant \theta_1$, we have $\min_{1\leqslant j\leqslant t} \widetilde{\ell}(j/n) = O(1/n)$, while for $t/n \geqslant \theta_1$, we have $\min_{1\leqslant j\leqslant t} \widetilde{\ell}(j/n) = \widetilde{\ell}(t/n)$. Hence, for all $t \leqslant T^*$,

$$\min_{1\leqslant j\leqslant t} \widetilde{\ell}(j/n) = \widetilde{\ell}(t/n) - \widetilde{\ell}^+(t/n) + O(1/n), \tag{29}$$

and thus, by (28),

$$\Big|\min_{1\leqslant j\leqslant t} \widetilde{d}(j) - n\widetilde{\ell}(t/n) + n\widetilde{\ell}^+(t/n)\Big| \leqslant M^* + O(1). \tag{30}$$

Finally, by (10), (27) and (30),

$$\big|d(t) - n\widetilde{\ell}^+(t/n)\big| \leqslant 2M^* + O(1). \tag{31}$$

This holds uniformly for $t \leqslant T^*$, and thus, by (27),

$$\max_{1\leqslant t\leqslant T^*} \big|d(t) - n\widetilde{\ell}^+(t/n)\big| = O_{L^2}(n^{1/2}). \tag{32}$$

It remains to consider $T^* < t \leqslant n$. Then the argument above does not quite work, because $\pi_t \searrow 0$ and thus $\operatorname{Var}\xi_t \nearrow \infty$ as $t \nearrow n$. We therefore modify ξ_t. We define $\widehat{\pi}_t := \max\{\pi_t, \pi_{T^*}\}$; thus $\widehat{\pi}_t = \pi_t$ for $t \leqslant T^*$ and $\widehat{\pi}_t > \pi_t$ for $t > T^*$. We may thus define independent random variables $\widehat{\xi}_t$ such that $\widehat{\xi}_t \sim \operatorname{Ge}(\widehat{\pi}_t)$ and $\widehat{\xi}_t \leqslant \xi_t$ for all $t < n$. (Thus,

$\widehat{\xi}_t = \xi_t$ for $t \leqslant T^*$.) The argument above works for the modified variables for all $t \leqslant n$. Since the modification can only increase $d(t)$, it follows that

$$\max_{T^*<t\leqslant n} d(t) = O_{L^2}(n^{1/2}), \tag{33}$$

which completes the proof since $\widetilde{\ell}(t/n) = 0$ for $t > T^*$. We omit the details. □

Proof of Corollary 6. Immediate from Theorem 5 and (17), since we have $\max_t \widetilde{\ell}^+(t/n) = \max_\theta \widetilde{\ell}^+(\theta) + O(1/n)$ and $\max_\theta \widetilde{\ell}^+(\theta) = \widetilde{\ell}^+(\theta_0) = \widetilde{\ell}(\theta_0)$. □

Proof of Corollary 7. By Theorem 5,

$$\frac{1}{n}\sum_{t=1}^{n} d(t) = \sum_{t=1}^{n} \widetilde{\ell}^+(t/n) + O_{L^2}\big(n^{1/2}\big) = n\alpha + O_{L^2}\big(n^{1/2}\big), \tag{34}$$

where

$$\begin{aligned}\alpha := \int_0^1 \widetilde{\ell}^+(x)\,\mathrm{d}x &= \int_0^{\theta_1} \widetilde{\ell}(x)\,\mathrm{d}x = \int_0^{\theta_1} \Big(x + \lambda^{-1}\log(1-x)\Big)\,\mathrm{d}x \\ &= \frac{1}{2}\theta_1^2 - \lambda^{-1}\Big((1-\theta_1)\log(1-\theta_1) + \theta_1\Big),\end{aligned} \tag{35}$$

which yields (24), using (18). □

2.4. **The trees in the forest.**

Theorem 9. *Let N be the number of trees in the depth-first forest. Then*

$$N = \psi n + O_{L^2}(n^{1/2}), \tag{36}$$

where

$$\psi = \psi(p) := 1 - \theta_1 - \frac{\lambda}{2}(1-\theta_1)^2. \tag{37}$$

Proof. Let $J_t := \mathbf{1}\{d(t) = 0\}$, the indicator that vertex t is a root and thus starts a new tree. Thus $N = \sum_1^n J_t$.

If $\theta_1 > 0$ (*i.e.*, $\lambda > 1$), then Theorem 5 shows that w.h.p. $d(t) > 0$ in the interval $(1, n\theta_1)$, except possibly close to the endpoints. Thus the DFS will find one giant tree of order $\approx \theta_1 n$, possibly preceded by a few small trees, and, as we will see later, followed by many small trees. To obtain a precise estimate, we note that there exists a constant $c > 0$ such that $\widetilde{\ell}(\theta) \geqslant \min\{c\theta, c(\theta_1 - \theta)\}$ for $\theta \in [0, \theta_1]$. Hence, if $t \leqslant n\theta_1$ and $d(t) = 0$, then $\widetilde{d}(t) \leqslant d(t) = 0$ by (10) and, recalling (27),

$$M^* \geqslant n\widetilde{\ell}(t/n) \geqslant c\min\{t, n\theta_1 - t\}. \tag{38}$$

Consequently, $d(t) = 0$ with $t \leqslant n\theta_1$ implies $t \in [1, c^{-1}M^*] \cup [n\theta_1 - c^{-1}M^*, n\theta_1]$. The number of such t is thus $O(M^* + 1) = O_{L^2}(n^{1/2})$, using (27).

Let $T_1 := \lceil n\theta_1 \rceil$. We have just shown that (the case $\theta_1 = 0$ is trivial)

$$\sum_{t=1}^{T_1-1} J_t = O_{L^2}(n^{1/2}). \tag{39}$$

It remains to consider $t \geqslant T_1$. Let

$$\mu_t := \mathbb{E}\,\xi_t = \frac{1-\pi_t}{\pi_t} = \frac{1-p}{p(1-t/n)} = \frac{1}{\lambda(1-t/n)}. \tag{40}$$

For any integer $k \geqslant 0$, the conditional distribution of $\xi_t - k$ given $\xi_t \geqslant k$ equals the distribution of ξ_t. Hence,

$$\mathbb{E}\big[(\xi_t - k)^+\big] = \mathbb{E}\big[\xi_t - k \mid \xi_t \geqslant k\big]\,\mathbb{P}(\xi_t \geqslant k) = \mu_t\,\mathbb{P}(\xi_t - k \geqslant 0). \tag{41}$$

We use again the stochastic recursion (7). Let $\mathcal{F}_t$ be the σ-field generated by $\xi_1,\dots,\xi_{t-1}$. Then $d(t)$ is $\mathcal{F}_t$-measurable, while ξ_t is independent of $\mathcal{F}_t$. Hence, (7) and (41) yield

$$\begin{aligned}\mathbb{E}\bigl[d(t+1)\mid\mathcal{F}_t\bigr]&=\mathbb{E}\bigl[d(t)+1-\xi_t\mid\mathcal{F}_t\bigr]+\mathbb{E}\bigl[(\xi_t-1-d(t))^+\mid\mathcal{F}_t\bigr]\\&=d(t)+1-\mu_t+\mu_t\,\mathbb{P}\bigl[\xi_t-1-d(t)\geqslant 0\mid\mathcal{F}_t\bigr]\\&=d(t)+1-\mu_t+\mu_t\,\mathbb{P}\bigl[d(t+1)=0\mid\mathcal{F}_t\bigr]\\&=d(t)+1-\mu_t+\mu_t\,\mathbb{E}\bigl[J_{t+1}\mid\mathcal{F}_t\bigr].\end{aligned}\tag{42}$$

We write $\Delta d(t):=d(t+1)-d(t)$ and $\overline{J}_t:=1-J_t$. Then (42) yields

$$\mathbb{E}\bigl[\Delta d(t)-1+\mu_t\overline{J}_{t+1}\mid\mathcal{F}_t\bigr]=0.\tag{43}$$

Define

$$\mathcal{M}_t:=\sum_{i=1}^{t-1}\mu_i^{-1}\bigl(\Delta d(i)-1+\mu_i\overline{J}_{i+1}\bigr)=\sum_{i=1}^{t-1}\Bigl(\mu_i^{-1}\Delta d(i)-\mu_i^{-1}+\overline{J}_{i+1}\Bigr).\tag{44}$$

Then $\mathcal{M}_t$ is $\mathcal{F}_t$-measurable, and (43) shows that $\mathcal{M}_t$ is a martingale. We have, with $\Delta\mathcal{M}_t:=\mathcal{M}_{t+1}-\mathcal{M}_t$, using (7),

$$|\Delta\mathcal{M}_t|\leqslant\mu_t^{-1}\bigl|d(t+1)-d(t)-1\bigr|+\overline{J}_{t+1}\leqslant\mu_t^{-1}\xi_t+1,\tag{45}$$

and thus, since $\pi_t\leqslant p<1$ for all t by (8),

$$\mathbb{E}|\Delta\mathcal{M}_t|^2\leqslant 2\mu_t^{-2}\,\mathbb{E}\,\xi_t^2+2=2\Bigl(\frac{\pi_t}{1-\pi_t}\Bigr)^2\frac{1-\pi_t+(1-\pi_t)^2}{\pi_t^2}+2=O(1).\tag{46}$$

Hence, uniformly for all $T\leqslant n$,

$$\mathbb{E}\,\mathcal{M}_T^2=\sum_{t=1}^{T-1}\mathbb{E}\,|\Delta\mathcal{M}_t|^2=O(T)=O(n).\tag{47}$$

The definition (44) yields

$$\mathcal{M}_n-\mathcal{M}_{T_1}=\sum_{t=T_1}^{n-1}\mu_t^{-1}\Delta d(t)-\sum_{t=T_1}^{n-1}\mu_t^{-1}+\sum_{t=T_1}^{n-1}\overline{J}_{t+1}.\tag{48}$$

By a summation by parts, and interpreting $\mu_n^{-1}:=0$,

$$\sum_{t=T_1}^{n-1}\mu_t^{-1}\Delta d(t)=\sum_{t=T_1+1}^{n}\bigl(\mu_{t-1}^{-1}-\mu_t^{-1}\bigr)d(t)-\mu_{T_1}^{-1}d(T_1).\tag{49}$$

As t increases, μ_t increases by (40), and thus $\mu_{t-1}^{-1}-\mu_t^{-1}>0$. Hence, (49) implies

$$\begin{aligned}\Bigl|\sum_{t=T_1}^{n-1}\mu_t^{-1}\Delta d(t)\Bigr|&\leqslant\sum_{t=T_1+1}^{n}\bigl(\mu_{t-1}^{-1}-\mu_t^{-1}\bigr)\sup_{i>T_1}|d(t)|+\mu_{T_1}^{-1}|d(T_1)|\leqslant 2\mu_{T_1}^{-1}\sup_{i\geqslant T_1}|d(t)|\\&=O_{L^2}(n^{1/2})\end{aligned}\tag{50}$$

by (20), since $\widetilde{\ell}^+(t/n)=0$ for $t\geqslant T_1\geqslant n\theta_1$. Furthermore, (47) shows that $\mathcal{M}_n,\mathcal{M}_{T_1}=O_{L^2}(n^{1/2})$. Hence, (48) yields

$$\sum_{t=T_1+1}^{n}J_t=n-T_1-\sum_{t=T_1+1}^{n}\overline{J}_t=n-T_1-\sum_{t=T_1}^{n-1}\mu_t^{-1}+O_{L^2}(n^{1/2})=n\psi+O_{L^2}(n^{1/2}),\tag{51}$$

where

$$\psi:=1-\theta_1-\int_{\theta_1}^{1}\lambda(1-x)\,\mathrm{d}x=1-\theta_1-\frac{\lambda}{2}(1-\theta_1)^2.\tag{52}$$

The result follows by (51) and (39). □

The argument in the proof of Theorem 9 shows also the following; we omit the details.

Theorem 10. *If $\lambda > 1$ ($p > \frac{1}{2}$), then the largest tree $\mathbf{T}_1$ in the depth-first forest has order $|\mathbf{T}_1| = \theta_1 n + O_{L^2}(n^{1/2})$.*

By different methods (for $\lambda = 1$ adapted from [1]), we can prove the following complements, which show the same behaviour as the largest component in a random Erdős–Rényi graph.

Theorem 11. (i) *If $\lambda = 1$ ($p = \frac{1}{2}$), then $|\mathbf{T}_1|/n^{2/3}$ converges in distribution to a positive random variable.*
(ii) *If $\lambda < 1$ ($p < \frac{1}{2}$), then, for some constants $c, C > 0$, w.h.p. $c \log n \leqslant |\mathbf{T}_1| \leqslant C \log n$.*

The limit distribution in (i) is of the type found in [1] for components in random graphs.

2.5. **Types of arcs.** Recall from the introduction the classification of the arcs in the digraph G. Since we assume that the outdegrees are $\mathrm{Ge}(1-p)$ and independent, the total number of arcs, M say, has a negative binomial distribution with mean λn, and, by a weak version of the law of large numbers,

$$M = \lambda n + O_{L^2}(n^{1/2}). \tag{53}$$

In the following theorem, we give the asymptotics of the number of arcs of each type.

Theorem 12. *Let L, T, B, F and C be the numbers of loops, tree arcs, back arcs, forward arcs, and cross arcs in the random digraph. Then*

$$L = O_{L^2}(1), \tag{54}$$

$$T = \tau n + O_{L^2}(n^{1/2}), \tag{55}$$

$$B = \beta n + O_{L^2}(n^{1/2}), \tag{56}$$

$$F = \varphi n + O_{L^2}(n^{1/2}), \tag{57}$$

$$C = \chi n + O_{L^2}(n^{1/2}), \tag{58}$$

where

$$\tau := \chi := 1 - \psi = \theta_1 + \frac{\lambda}{2}(1-\theta_1)^2, \tag{59}$$

$$\beta := \varphi := \lambda\alpha = (\lambda - 1)\theta_1 - \frac{\lambda}{2}\theta_1^2. \tag{60}$$

The equalitites $\tau = \chi$ and $\beta = \varphi$ mean asymptotic equality of the corresponding expectations of numbers of arcs. In fact, there are exact equalities.

Theorem 13. *For any n, $\mathbb{E}\, T = \mathbb{E}\, C$ and*

$$\mathbb{E}\, B = \mathbb{E}\, F = \lambda\, \mathbb{E}\, \overline{d} = \beta n + O(n^{1/2}). \tag{61}$$

Remark 14. Knuth [5] conjectures, based on exact formulas for small n, that, much more strongly, B and F have the same distribution for every n. (Note that T and C do not have the same distribution; we have $T \leqslant n - 1$, while C may take arbitrarily large values.)

Partial proof of Theorems 12 and 13. L: A simple argument with generating functions shows that the number of loops at a given vertex v is $\mathrm{Ge}(1 - p/(n - np + p))$; these numbers are independent, and thus $L \sim \mathrm{NegBin}\big(n, 1 - p/(n - np + p)\big)$ with $\mathbb{E}\, L = p/(1-p) = \lambda = O(1)$ and $\mathrm{Var}(L) = p(1 - p + p/n)/(1-p)^2 = O(1)$ [5]. Moreover, it is easily seen that asymptotically, L has a Poisson distribution, $L \xrightarrow{\mathrm{d}} \mathrm{Po}(\lambda)$.

T: This follows immediately from Theorem 9, since $T = n - N$.

B, F: Let v, w be two distinct vertices. If the DFS finds w as a descendant of v, then there will later be $\mathrm{Ge}(1-p)$ arcs from w, and each has probability $1/n$ of being a back arc to v. Similarly, there will be $\mathrm{Ge}(1-p)$ future arcs from v, and each has probability $1/n$ of being a forward arc to w. Hence, if I_{vw} is the indicator that w is a descendant of v, and B_{vw} $[F_{vw}]$ is the number of back [forward] arcs vw, then

$$\mathbb{E}\,B_{wv} = \mathbb{E}\,F_{vw} = \frac{\lambda}{n}\,\mathbb{E}\,I_{vw}. \tag{62}$$

Summing over all pairs of distinct v and w, we obtain

$$\mathbb{E}\,B = \mathbb{E}\,F = \frac{\lambda}{n}\,\mathbb{E}\sum_{w}\sum_{v\neq w} I_{vw} = \frac{\lambda}{n}\,\mathbb{E}\sum_{w} d(w) = \lambda\,\mathbb{E}\,\overline{d}, \tag{63}$$

and (61) follows by Corollary 7.

The proofs of (56)–(58) and the equality $\mathbb{E}\,T = \mathbb{E}\,C$ are given in the full paper [4]. □

3. Depth, trees and arc analysis in the shifted geometric outdegree distribution

In this section, the outdegree distribution is $\mathrm{Ge}_1(1-p) = 1 + \mathrm{Ge}(1-p)$. Thus we now have the mean

$$\lambda = \frac{1}{1-p}.$$

As in Section 2, the depth $d(t)$ is a Markov chain given by (7), but the distribution of ξ_t is now different. The probability (3) is replaced by $(1-t/n)/(1-pt/n)$, but the number of future arcs from an ancestor is still $\mathrm{Ge}(1-p)$, and, with $\theta := t/n$,

$$\mathbb{P}\big(\xi_t = k\big) = \begin{cases} \overline{\pi}_t := \frac{1-\theta}{1-p\theta}, & k = 0, \\ (1-\overline{\pi}_t)(1-\pi_t)^{k-1}\pi_t, & k \geqslant 1, \end{cases} \tag{64}$$

where $\pi_t = p\overline{\pi}_t$ is as in (8). The probability generating function of ξ_t is, instead of (11),

$$p(t,z) = \overline{\pi}_t + (1-\overline{\pi}_t)\frac{\pi_t z}{1-(1-\pi_t)z} = (1-\theta)\frac{1-(1-p)z}{1-p\theta-(1-p)z}. \tag{65}$$

The rest of the analysis does not change, and the results in Theorems 5–12 still hold, but we get different values for many of the constants.

We now have $\mathbb{E}\,\xi_t = \frac{(1-p)\theta}{p(1-\theta)}$ and instead of (16) we have $\mathbb{E}\,\widetilde{d}(t) = n\widetilde{\ell}(\theta) + O(1)$ where now $\widetilde{\ell}(\theta)$ takes the new value

$$\widetilde{\ell}(\theta) := \frac{1}{p}\theta + \frac{1-p}{p}\log(1-\theta). \tag{66}$$

Note that $\widetilde{\ell}(\theta_1) = 0$ still gives (18), now with $\lambda = 1/(1-p)$, and that $\lambda > 1$ for every p. Differentiating (66) shows that the maximum point $\theta_0 = p > 0$.

Figure 2 shows $\widetilde{\ell}(\theta)$ for both geometric distributions.

Straightforward calculations yield

$$\upsilon := \widetilde{\ell}(p) = 1 + \frac{1-p}{p}\log(1-p), \tag{67}$$

$$\alpha := \frac{1}{p}\left(\frac{\theta_1^2}{2} - \frac{1}{\lambda}((1-\theta_1)\log(1-\theta_1) - \frac{1}{\lambda}\theta_1\right) = \theta_1 - \frac{\theta_1^2}{2p}, \tag{68}$$

$$\psi := 1 - \theta_1 - \frac{\lambda}{2}(1-\theta_1)^2. \tag{69}$$

Note that the value for ψ is the same as in (37). This is no coincidence; we show by the method in Section 4 that this holds for all offspring distributions with the same mean λ.

Now the expected numbers of back and forward arcs differ since $\mathbb{E}\,B = \lambda\,\mathbb{E}\,\overline{d} \sim \lambda\alpha n$ and $\mathbb{E}\,F = (\lambda-1)\,\mathbb{E}\,\overline{d} \sim (\lambda-1)\alpha n$ because the average number of future arcs at a vertex after a descendant have been created is $\lambda - 1$. Thus the equality $\beta = \varphi$ and the equality of the expected number of back and forward arcs in Theorems 12 and 13 was an artefact of the geometric degree distribution.

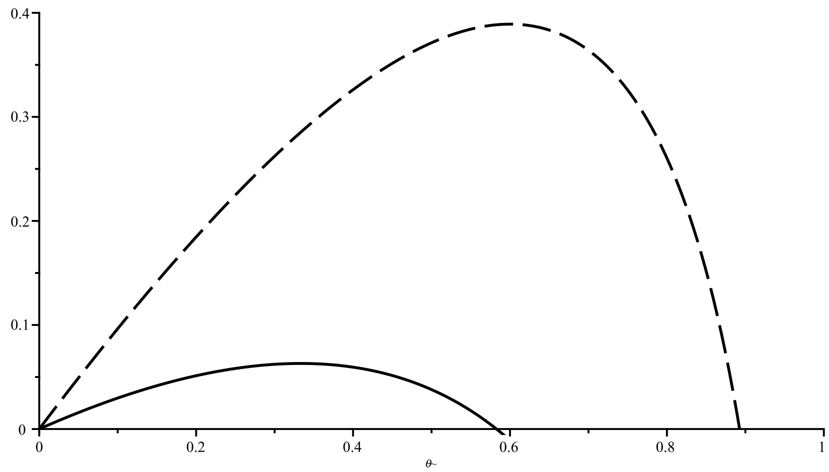

FIGURE 2. $\widetilde{\ell}(\theta)$, the asymptotic search depth, for geometric distribution (solid) and shifted geometric distribution (dashed) with $p = 0.6$.

The estimates (54)–(58) in Theorem 12 hold, with the constants now given by

$$\tau := 1 - \psi = \theta_1 + \frac{\lambda}{2}(1-\theta_1)^2, \tag{70}$$

$$\beta := \lambda\alpha = \lambda\theta_1 - \frac{\lambda}{2p}\theta_1^2 = \lambda\theta_1 - \frac{\lambda^2}{2(\lambda-1)}\theta_1^2, \tag{71}$$

$$\varphi := (\lambda-1)\alpha = (\lambda-1)\theta_1 - \frac{\lambda}{2}\theta_1^2 = \frac{\lambda}{2} - \tau, \tag{72}$$

$$\chi := \frac{\lambda}{2} - \beta = \frac{\lambda}{2}(1-\theta_1)^2 + \frac{\lambda}{2(\lambda-1)}\theta_1^2. \tag{73}$$

Note that τ and φ are as in Theorem 12, while β and χ are different. In particular, $\beta \neq \varphi$ as noted above. Similarly, $\chi \neq \tau$, and thus the equality of the expected numbers of tree arcs and cross arcs in Theorem 13 also was an effect of the geometric distribution.

4. Stack index analysis and forest size for a general outdegree distribution

In this section, we consider a general outdegree distribution $\mathbf{P}$, with mean λ and finite variance. When the outdegree distribution is general, the depth does not longer follow an easy Markov chain, since we should keep track of the number of children seen so far at each level of the branch of the tree toward the current vertex.

Instead we get back a Markov chain if we replace the depth by the stack index $I(t)$. The DFS can be regarded as keeping a stack of unexplored arcs, for which we have seen the start vertex but not the end. The stack evolves as follows:

S1. If the stack is empty, pick a new vertex v that has not been seen before (if there is no such vertex, we have finished). Otherwise, pop the last arc from the stack, and reveal its endpoint v (which is uniformly random over all vertices). If v already is seen, repeat.

S2. (v is now a new vertex) Reveal the outdegree m of v and add to the stack m new arcs from v, with unspecified endpoints. GOTO S1.

Let again v_t be the tth vertex seen by the DFS, and let $I(t)$ be the size of the stack when $v(t)$ is found (but before we add the arcs from v_t). Also let η_t be the outdegree of v_t. Then

$I(1) = 0$ and, in analogy with (7),

$$I(t+1) = \bigl(I(t) + \eta_t - 1 - \xi_t\bigr)^+, \qquad 1 \leqslant t < n, \tag{74}$$

where ξ_t is the number of arcs leading to already seen vertices before we find a new one; we have $\mathbb{P}(\xi = k) = (1 - \frac{t}{n})(\frac{t}{n})^k$ and thus $\xi_t \sim \mathrm{Ge}(1 - t/n)$.

In analogy with (9), we define also

$$\widetilde{I}(t) := \sum_{i=1}^{t-1} (\eta_i - 1 - \xi_i) = \sum_{i=1}^{t-1} \zeta_i, \tag{75}$$

where we define $\zeta_t := \eta_t - \xi_t - 1$. Then, as in (10),

$$I(t) = \widetilde{I}(t) - \min_{1 \leqslant j \leqslant t} \widetilde{I}(j). \tag{76}$$

Note that

$$\mathbb{E}\,\zeta_t = \mathbb{E}\,\eta_t - \mathbb{E}\,\xi_t - 1 = \lambda - \frac{t/n}{1 - t/n} - 1 = \lambda - \frac{1}{1 - t/n}. \tag{77}$$

Hence, uniformly in $t/n \leqslant \theta^* $ for any fixed $\theta^* < 1$,

$$\mathbb{E}\,\widetilde{I}(t) = \sum_{i=1}^{t-1} \mathbb{E}\,\zeta_i = (t-1)\lambda - \sum_{i=1}^{t-1} \frac{1}{1 - t/n} = n\tilde{\iota}(t/n) + O(1), \tag{78}$$

where

$$\tilde{\iota}(\theta) := \int_0^\theta \Bigl(\lambda - \frac{1}{1-\tau}\Bigr)\,\mathrm{d}\tau = \lambda\theta + \log(1-\theta). \tag{79}$$

Let

$$\tilde{\iota}^+(\theta) := [\tilde{\iota}(\theta)]^+ = \begin{cases} \lambda\theta + \log(1-\theta), & 0 \leqslant \theta \leqslant \theta_1, \\ 0, & \theta_1 \leqslant \theta \leqslant 1, \end{cases} \tag{80}$$

where again θ_1 is the largest root in $[0, 1)$ of (18), now with $\lambda = \mathbb{E}\,\eta_1$, the mean of $\mathbf{P}$. The proof of Theorem 5 applies with very minor differences, and yields:

Theorem 15. *Suppose that the outdegree distribution has finite variance. Then*

$$\max_{1 \leqslant t \leqslant n} \bigl| I(t) - n\tilde{\iota}^+(t/n) \bigr| = O_{L^2}(n^{1/2}). \tag{81}$$

Moreover, v_{t+1} is a root if and only if $I(t) + \zeta_t = I(t) + \eta_1 - 1 - \xi_t < 0$, cf. (74). The arguments in the proof of Theorem 9 apply with minor differences, and show:

Theorem 16. *Theorems 9 and 10 hold for any outdegree distribution with finite variance, with* $\psi := 1 - \theta_1 - \frac{\lambda}{2}(1 - \theta_1)^2$.

Figure 3 shows the parameter ψ as a function of the average degree λ.

Moreover, we have:

Theorem 17. *Theorem 11*(i) *holds for any outdegree distribution with mean* $\lambda = 1$ *and positive, finite variance* $\mathrm{Var}\,\eta$.

Theorem 11(ii) *holds for any outdegree distribution with mean* $0 < \lambda < 1$ *and some finite exponential moment* $\mathbb{E}\,e^{x_0\eta} < \infty$ *for some* $x_0 > 0$.

Remark 18. Theorem 11(i) does not hold for the non-random $\eta = 1$, *i.e.*, when the digraph is a random mapping; in this case the largest trees have sizes of orders $n^{1/2}$, not $n^{2/3}$. (See also [5, Exercise 36 (preliminary number)].)

Moreover, we can use the stack index to find the depth of the nodes, which leads to

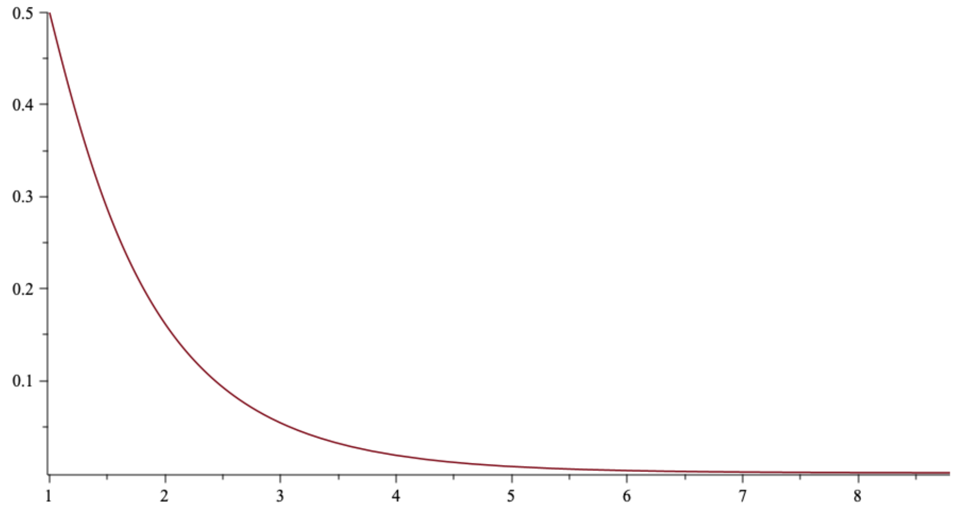

FIGURE 3. ψ, as function of λ.

Theorem 19. *Let $\rho(\theta)$ be the survival probability of a Galton–Watson process with offspring distribution obtained from* $\mathbf{P}$ *by thinning, killing each child with probability θ. Let*

$$\widetilde{\ell}^{+}(\theta) := \begin{cases} \widetilde{\ell}(\theta) = \int_0^\theta \rho(x)\,\mathrm{d}x, & 0 \leqslant \theta \leqslant \theta_0, \\ \widetilde{\ell}(\breve{\theta}), \ \textit{where } \breve{\theta} \in (0,\theta_0) \textit{ and } \widetilde{\iota}(\breve{\theta}) = \widetilde{\iota}(\theta), & \theta_0 < \theta < \theta_1, \\ 0, & \theta_1 \leqslant \theta \leqslant 1. \end{cases} \tag{82}$$

Then, we have

$$\max_{1 \leqslant t \leqslant n} \left| d(t) - n\widetilde{\ell}^{+}(t/n) \right| = o_{L^2}(n), \tag{83}$$

For the geometric and shifted geometric distributions in Section 2 and 3, we see from (17), (66) and (79) that $\widetilde{\ell}(\theta)$ and $\widetilde{\iota}(\theta)$ are proportional, and thus the expected stack size $\mathbb{E}\,\widetilde{I}(t)$ and depth $\mathbb{E}\,\widetilde{d}(t)$ asymptotically are proportional by a fixed factor independent of t for $t \in (0,\theta_1)$. However, although $\widetilde{\ell}(\theta)$ and $\widetilde{\iota}(\theta)$ always have the same root θ_1, they are in general not proportional; in fact, assuming $\lambda > 1$, this happens only when the outdegree distribution is geometric, or a geometric distribution with $\mathbb{P}(\eta_i = 0)$ modified.

Using the stack index $I(t)$, we can extend many of the results from Section 2 to arbitrary outdegree distributions with finite variance, but not all. Thus some results for geometric distributions remain conjectures in general. In particular, we mention the following. (See the comment after Corollary 6.)

Conjecture 20. *We conjecture that as in the geometric case, the height Υ is asymptotically normally distributed for any supercritical outdegree distribution with finite variance.*

REFERENCES

[1] David Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.* **25** (1997), no. 2, 812–854.

[2] Sahar Diskin & Michael Krivelevich. On the performance of the Depth First Search algorithm in supercritical random graphs. Preprint, 2021. arXiv:2111.07345

[3] Nathanaël Enriquez, Gabriel Faraud & Laurent Ménard. Limiting shape of the depth first search tree in an Erdős-Rényi graph. *Random Structures Algorithms* **56** (2020), no. 2, 501–516.

[4] Svante Janson and Philippe Jacquet, Depth-First Search performance in random digraphs. Preprint, 2022.

[5] Donald E. Knuth, *The Art of Computer Programming*, Section 7.4.1.2 (Preliminary draft, 13 February 2022). http://cs.stanford.edu/~knuth/fasc12a.ps.gz
[6] Michael Krivelevich & Benny Sudakov. The phase transition in random graphs: a simple proof. *Random Structures Algorithms* **43** (2013), no. 2, 131–138.

Inria Saclay Ile de France, France
Email address: philippe.jacquet@inria.fr

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden
Email address: svante.janson@math.uu.se
URL: http://www2.math.uu.se/∼svante/