The graph structure of a deterministic automaton chosen at random

Xing Shi Cai, Luc Devroye

School of Computer Science, McGill University of Montreal, Canada,

xingshi.cai@mail.mcgill.ca

lucdevroye@gmail.com

April 11, 2016

An n-state deterministic finite automaton over a k-letter alphabet can be seen as a digraph with n vertices which all have k labeled out-arcs. Grusho [20] proved that whp in a random k-out digraph there is a strongly connected component of linear size, i.e., a giant, and derived a central limit theorem. We show that whp the part outside the giant contains at most a few short cycles and mostly consists of tree-like structures, and present a new proof of Grusho's theorem. Among other things, we pinpoint the phase transition for strong connectivity.

Keywords: random digraphs; deterministic finite automaton

1 Introduction

1.1 The model and the history

The deterministic finite automaton (DFA) is widely used in computational complexity theory. Formally, a DFA is a 5-tuple $(Q, \Sigma, \delta, q_0, F)$, where Q is a finite set called the set of states, Σ is a finite set called the alphabet, $\delta : Q \times \Sigma \to Q$ is the transition function, $q_0 \in Q$ is the start state, and $F \subseteq Q$ is the set of accept states. If q_0 and F are ignored, a DFA with n states and a k-alphabet can be seen as a digraph with vertices $[n] \equiv \{1 \dots, n\}$ in which each vertex has k out-arcs labeled by $1, \dots, k$ (a *k-out digraph*). Note that such a digraph can have self-loops and multi-arcs. For a basic introduction to DFA and its applications, see [37].

Let $\mathcal{D}_{n,k}$ denote a digraph chosen uniformly at random from all k-out digraphs of n vertices. Equivalently $\mathcal{D}_{n,k}$ is a random k-out digraph of n vertices with the endpoints of its kn arcs chosen independently and uniformly at random.

When $k = 1$, $\mathcal{D}_{n,k}$ is equivalent to a uniform random mapping from $[n]$ to itself, which has been well studied by Kolchin [27], Flajolet and Odlyzko [18], and Aldous and Pitman [2]. In $\mathcal{D}_{n,1}$, the largest strongly connected component (SCC) has expected size $\Theta(\sqrt{n})$, and so does the size of the longest cycle. However, as shown later, for $k \geq 2$, the largest SCC has expected size $\Theta(n)$.

From now on we assume that $k \geq 2$. Let $\mathcal{S}_v$ (the *spectrum* of v) be the set of vertices in $\mathcal{D}_{n,k}$ that are reachable from vertex v, including v itself. In 1973 Grusho [20] first proved that $(|\mathcal{S}_1| - \nu_k n)/\sigma_k \sqrt{n}$ converges in distribution to a standard normal, where ν_k and σ_k are explicitly defined constants.

Given a set of vertices $\mathcal{S} \subseteq [n]$, call $\mathcal{S}$ *closed* if there are no arcs that start from vertices in $\mathcal{S}$ and end at vertices in $\mathcal{S}^c \equiv [n] \setminus \mathcal{S}$. Let $\mathcal{G}_n$ be the set of vertices in the largest closed SCC in $\mathcal{D}_{n,k}$. (If the largest closed SCC is not unique, let $\mathcal{G}_n$ be the vertex set of the largest closed SCC that contains the smallest vertex-label.) We call $\mathcal{G}_n$ the giant. Grusho also proved that $|\mathcal{G}_n|$ has the same limit distribution as $|\mathcal{S}_1|$ by showing that with high probability (whp) $\mathcal{G}_n$ is reachable from all vertices and that $|\mathcal{S}_1| - |\mathcal{G}_n| = o_p(\sqrt{n})$ (see [22] for the notation). His proof relies on a result by Sevast'yanov [35] which approximates the exploration of $\mathcal{S}_1$ with a Gaussian process.

In 2012 Carayol and Nicaud [10] proved a local limit theorem for $|\mathcal{S}_1|$ by analyzing the limit behavior of the probability that $|\mathcal{S}_1| = s$ for an s close to $\nu_k n$. Their proof depends on a theorem by Korshunov [28] which says that conditioned on every vertex having in-degree at least one, the probability that $\mathcal{S}_1 = [n]$ tends to some constant. Carayol and Nicaud derived a simple and explicit formula of this constant from their theorem. (The same formula is also proved by Lebensztayn [29] with a more analytic approach using Lagrange series.)

Lately the simple random walk (SRW) on $\mathcal{D}_{n,k}$ has gained some attention for its applications in machine learning. Addario-Berry, Balle, and Perarnau [1] studied the stationary distribution of the SRW by analyzing the distances in $\mathcal{D}_{n,k}$. They proved that the diameter and the typical distance, rescaled by $\log n$, converge in probability to explicit constants. Angluin and Chen [3] studied the rate of the convergence to the stationary distribution of the SRW. They also suggested an algorithm for learning a uniformly random DFA under Kearns' statistical query model [26].

1.2 Our results and a sketch of proof

A digraph can be uniquely decomposed into SCCs which form a directed acyclic graph (DAG) through a process called condensation that contracts every SCC into a single vertex while keeping all the arcs between SCCs [5]. The condensation DAG of $\mathcal{D}_{n,k}$ is denoted by $\mathcal{D}^{\mathrm{A}}_{n,k}$.

Let $\mathcal{G}^c_n \equiv [n] \setminus \mathcal{G}_n$, i.e., $\mathcal{G}^c_n$ is the set of vertices that are outside the giant. The structure of $\mathcal{D}^{\mathrm{A}}_{n,k}$ depends on $\mathcal{D}_{n,k}[\mathcal{G}^c_n]$, the digraph induced by $\mathcal{G}^c_n$. Our analysis shows that in $\mathcal{D}_{n,k}[\mathcal{G}^c_n]$ the total number of cycles and the number of cycles of a fixed length both converge to Poisson distributions with constant means. So the number of cycles and the length of the longest cycle are both $O_p(1)$ (see [22]). Furthermore, these cycles are vertex-disjoint whp. Therefore, almost every vertex in $\mathcal{G}^c_n$ is a SCC itself and $\mathcal{D}^{\mathrm{A}}_{n,k}$ is very much like $\mathcal{D}_{n,k}$ with the giant contracted into a single vertex.

The *d-core* of an undirected graph is the maximum induced subgraph in which all vertices have degree at least d. Similarly the *d-in-core* of a digraph can be defined as the maximum induced sub-digraph in which all vertices have in-degree at least d. Let $\mathcal{O}_n$ denote the set of vertices in the one-in-core of $\mathcal{D}_{n,k}$. Note that $\mathcal{G}_n \subseteq \mathcal{O}_n$ since a SCC induces a sub-digraph with each vertex having in-degree at least one. Also note that cycles cannot exist outside $\mathcal{O}_n$, for otherwise they contradict the maximality of $\mathcal{O}_n$. Now assume that every vertex can reach $\mathcal{G}_n$, which happens whp by Grusho [20]. Then $\mathcal{D}_{n,k}$ can be divided into three layers: the center is $\mathcal{G}_n$; then comes $\mathcal{O}_n \setminus \mathcal{G}_n$, which consists of cycles outside $\mathcal{G}_n$ and paths from these cycles to $\mathcal{G}_n$; the outermost is $\mathcal{O}_n^c \equiv [n] \setminus \mathcal{O}_n$, which is acyclic.

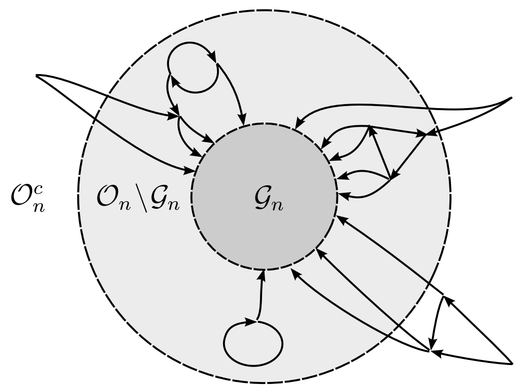

Figure 1: Three layers of $\mathcal{D}_{n,k}$: the giant $\mathcal{G}_n$; the one-in-core $\mathcal{O}_n$; and the whole graph.

Since there cannot be many vertices in cycles outside the giant, the middle layer $\mathcal{O}_n \setminus \mathcal{G}_n$ must be very "thin". Thus if we can prove $(|\mathcal{O}_n| - \nu_k n)/\sqrt{n}$ converges to a normal distribution, then we can also prove it for $|\mathcal{G}_n|$. The event $|\mathcal{O}_n| = s$ happens if and only if there is a set of vertices $\mathcal{S}$ with $|\mathcal{S}| = s$ such that: (a) $\mathcal{D}_{n,k}[\mathcal{S}]$, the sub-digraph induced by $\mathcal{S}$, has minimum in-degree one (*surjective*) and there are no arcs going from $\mathcal{S}$ to $\mathcal{S}^c$ (*closed*), which we refer to as $\mathcal{S}$ being a *k-surjection* (since $\mathcal{D}_{n,k}[\mathcal{S}]$ is equivalent to a surjective function from $[ks]$ to $[s]$); (b) $\mathcal{D}_{n,k}[\mathcal{S}^c]$ is acyclic. The probability of (a) can be computed by counting the number of surjective functions. And we are able to show that the probability of (b) converges to a constant. Note that for a fixed set $\mathcal{S}$ (a) and (b) are independent because they depend on the endpoints of two disjoint sets of arcs. Thus we can get the limit of $\mathbb{P}\{\mathcal{O}_n = \mathcal{S}\}$. Since the one-in-core of a digraph is unique, $\mathbb{P}\{|\mathcal{O}_n| = s\} = \sum_{\mathcal{S}\subseteq[n]:|\mathcal{S}|=s} \mathbb{P}\{\mathcal{O}_n = \mathcal{S}\}$. Thus we can finish the proof by computing the characteristic function of $(|\mathcal{O}_n| - \nu_k n)/\sqrt{n}$.

Note that although our formula for $\mathbb{P}\{|\mathcal{O}_n| = s\}$ is inspired by and resembles Carayol and Nicaud's formula for $\mathbb{P}\{|\mathcal{S}_1| = s\}$, we actually prove the result from scratch without relying on previous work. Since we are able to derive explicit expressions of all the constants in our formula, the computation of the characteristic function becomes quite simple. Furthermore, to our knowledge this is the first self-contained proof. Thus in

Section 2 we prove:

Theorem 1 (Central limit law)**.** *Let $\mathcal{Z}$ denote a standard normal random variable. Then as $n \to \infty$,*

$$\frac{|\mathcal{O}_n| - \nu_k n}{\sigma_k \sqrt{n}} \xrightarrow{d} \mathcal{Z}, \qquad \frac{|\mathcal{G}_n| - \nu_k n}{\sigma_k \sqrt{n}} \xrightarrow{d} \mathcal{Z}, \qquad \frac{\max_{v \in [n]} |\mathcal{S}_v| - \nu_k n}{\sigma_k \sqrt{n}} \xrightarrow{d} \mathcal{Z},$$

where ν_k and σ_k are constants defined by

$$\nu_k \equiv \frac{\tau_k}{k}, \qquad \sigma_k^2 \equiv \frac{\tau_k}{k e^{\tau_k}(1 - k e^{-\tau_k})},$$

and τ_k is the unique positive solution of $1 - \tau_k/k - e^{-\tau_k} = 0$.

Remark. Equivalently, ν_k is the unique positive solution of $1 - \nu_k = e^{-k\nu_k}$ and

$$\sigma_k^2 = \frac{\nu_k(1 - \nu_k)}{1 - k(1 - \nu_k)}.$$

Let $G(n, m)$ be a Erdős–Rényi random graph, i.e., a graph chosen uniformly at random from all graphs with n vertices and m edges [16]. It is well-known that for $k > 1$, $|\mathcal{C}^n_{\max}|$—the size of the largest component in $G(n, m = nk/2)$—is $(\nu_k + o(1))n$ whp. Moreover, $(|\mathcal{C}^n_{\max}| - \nu_k n)/\sqrt{n}$ also converges in distribution to a normal random variable with variance σ_k^2 (see, e.g., Durrett [14]). Intuitively, this is because the in-degree of a vertex in $\mathcal{D}_{n,k}$ has asymptotically a Poisson distribution of mean k. Thus a backward exploration process from vertex in $\mathcal{D}_{n,k}$ is approximately a Galton-Watson process with survival probability ν_k, as is the exploration process starting from a vertex in $G(n, m = nk/2)$.

Section 3 studies the part of $\mathcal{D}_{n,k}$ outside the giant, which determines the structure of $\mathcal{D}^{\mathrm{A}}_{n,k}$ and supports the proof of Theorem 1. Our results are summarized in two theorems, where all our logarithms are natural:

Theorem 2 (Cycles outside the giant)**.** *We have:*

(a) Let L_n be the length of the longest cycle in $\mathcal{D}_{n,k}[\mathcal{G}_n^c]$. Then $L_n = O_p(1)$.

(b) Let C_n be the number of cycles in $\mathcal{D}_{n,k}[\mathcal{G}_n^c]$. Then

$$C_n \xrightarrow{d} \mathrm{Poi}\left(\log \frac{1}{1 - k e^{-\tau_k}}\right),$$

where $\mathrm{Poi}(x)$ denotes the Poisson distribution with mean x.

(c) Let $C_{n,\ell}$ be the number of cycles of length ℓ in $\mathcal{D}_{n,k}[\mathcal{G}_n^c]$. Then for all fixed $\ell \geq 1$,

$$C_{n,\ell} \xrightarrow{d} \mathrm{Poi}\left(\frac{(k e^{-\tau_k})^\ell}{\ell}\right).$$

Theorem 3 (Spectra outside the giant)**.** *Let $\mathcal{S}'_v \equiv \mathcal{S}_v \cap \mathcal{G}_n^c$, i.e., $\mathcal{S}'_v$ is the spectrum of v in $\mathcal{D}_{n,k}[\mathcal{G}_n^c]$. Let $\mathrm{dist}(v,u)$ be the distance from v to u, i.e., the length of the shortest directed path from v to u. Then*

(a) $\mathbb{P}\left\{\cup_{v \in \mathcal{G}_n^c}[\mathrm{arc}(\mathcal{D}_{n,k}[\mathcal{S}'_v]) - |\mathcal{S}'_v| \geq 1]\right\} = o(1)$*, where* $\mathrm{arc}(\cdot)$ *denotes the number of arcs. In other words, whp every spectrum in $\mathcal{D}_{n,k}[\mathcal{G}_n^c]$ is a tree or a tree plus an extra arc.*

(b) Let $S_n \equiv \max_{v \in \mathcal{G}_n^c} |\mathcal{S}'_v|$. Let $\lambda_k \equiv (k - \tau_k)\left(\frac{\tau_k}{k-1}\right)^{k-1}$. Then

$$\frac{S_n}{\log n} \xrightarrow{p} \frac{1}{\log(1/\lambda_k)}.$$

(c) Let $W_n \equiv \max_{v \in \mathcal{G}_n^c} \min_{u \in \mathcal{G}_n} \mathrm{dist}(v,u)$, i.e., the maximum distance to $\mathcal{G}_n$. Then

$$\frac{W_n}{\log_k \log n} \xrightarrow{p} 1.$$

(d) Let M_n be the length of the longest path in $\mathcal{D}_{n,k}[\mathcal{G}_n^c]$. Then

$$\frac{M_n}{\log n} \xrightarrow{p} \frac{1}{\log(e^{\tau_k}/k)}.$$

(e) Let $D_n \equiv \max_{v \in \mathcal{G}_n^c} \max_{u \in \mathcal{S}'_v} \mathrm{dist}(v,u)$. Then

$$\frac{D_n}{\log n} \xrightarrow{p} \frac{1}{\log(e^{\tau_k}/k)}.$$

The rest of the paper gives some other results regarding this model. Section 4 shows that $\mathcal{D}_{n,k}$ exhibits a phase transition for strong connectivity. Section 5 extends some of our results to simple k-out digraphs. Section 6 suggests some extensions of this model.

Remark. Lemma 9 shows that $|\mathcal{O}_n| - |\mathcal{G}_n| = O_p(1)$. The intuition is that a digraph with minimal in-degree and out-degree at least one is likely to have a large SCC. This phenomenon is also observed in $D(n,p)$, which is a random digraph of n vertices with each possible arc existing independently with probability p. Pittel and Poole [33, thm. 1.3] showed that in $D(n,p)$ the $(1,1)$-core—the maximal induced sub-digraph in which each vertex has in-degree and out-degree at least one—differs from the largest SCC in size by at most $O((\log n)^8)$, whp. This intuition is also used for studying the asymptotic counts of strongly connected digraphs (see Pérez-Giménez and Wormald [34], Pittel [32]).

2 The size of the one-in-core

2.1 The law of large numbers for the one-in-core

To prove Theorem 1, we first need to narrow the range of $|\mathcal{O}_n|$ to close to $\nu_k n$.

Theorem 4 (Law of large numbers)**.** *For all fixed* $\delta \in (0, 1/2)$*,*

$$\mathbb{P}\left\{|\mathcal{O}_n| \notin \mathcal{I}_n\right\} \leq \frac{1 + o(1)}{n},$$

where $\mathcal{I}_n \equiv [\nu_k n - n^{1/2+\delta}, \nu_k n + n^{1/2+\delta}]$*.*

Thus $|\mathcal{O}_n|/n \xrightarrow{p} \nu_k$, which gives the theorem its name.

Let K_s be the number of k-surjections of size s in $\mathcal{D}_{n,k}$. Then it suffices to show that $\mathbb{P}\left\{\sum_{s \notin \mathcal{I}_n} K_s \geq 1\right\} \leq (1 + o(1))/n$. As argued in the introduction, for a set of vertices $\mathcal{S}$ to be the one-in-core, it must also be a k-surjection, i.e., every vertex in $\mathcal{D}_{n,k}[\mathcal{S}]$, the sub-digraph induced by $\mathcal{S}$, must have minimum in-degree one ($\mathcal{S}$ is *surjective*), and there are no arcs going from $\mathcal{S}$ to $\mathcal{S}^c$ ($\mathcal{S}$ is *closed*). Thus

$$\mathbb{P}\left\{\mathcal{S} \text{ is a } k\text{-surjection}\right\} = \mathbb{P}\left\{\mathcal{S} \text{ is surjective} \mid \mathcal{S} \text{ is closed}\right\} \mathbb{P}\left\{\mathcal{S} \text{ is closed}\right\}.$$

Computing the limit of the two factors shows that:

Lemma 1. *We have*

$$\mathbb{P}\left\{\sum_{s \notin \mathcal{I}_n} K_s \geq 1\right\} \leq \frac{1 + o(1)}{n}.$$

And for $s \in \mathcal{I}_n$

$$\mathbb{E}K_s \sim \frac{1}{\sqrt{2\pi(1 - ke^{-\tau_k})n}}\, g\left(\frac{s}{n}\right) \quad \left[f\left(\frac{s}{n}\right)\right]^n,$$

where

$$g(x) \equiv \frac{1}{\sqrt{x(1-x)}}, \qquad f(x) \equiv \left[\frac{x^{k-1}\gamma_k}{(1-x)^{(1-x)/x}}\right]^x,$$

and $\gamma_k \equiv \left(\frac{k}{e^{\tau_k}}\right)^k (e^{\tau_k} - 1)$*.*

Theorem 4 follows immediately. The proof of Lemma 1 is postponed to the appendix. (The two functions $f(x)$ and $g(x)$ are also studied by Carayol and Nicaud [10].)

2.2 The central limit law of the one-in-core

In this section we prove the part of Theorem 1 about $|\mathcal{O}_n|$. The rest of the theorem appears as corollaries in Section 3. Let $\partial\mathcal{O}_n = |\mathcal{O}_n| - \nu_k n$. Then $\partial\mathcal{O}_n$ takes values in $[n] - \nu_k n \equiv \{s : \nu_k n + s \in [n]\}$. As Theorem 4 shows, whp $\partial\mathcal{O}_n \leq n^{1/2+\delta}$ for all fixed $\delta \in (0, 1/2)$. Thus it suffices to consider only the probability that $\partial\mathcal{O}_n$ takes value in the set

$$\mathcal{J}_n \equiv ([n] - \nu_k n) \cap \left[-n^{1/2+\delta}, n^{1/2+\delta}\right],$$

for some fixed $\delta \in (0, 1/2)$. Thus the characteristic function of $\partial\mathcal{O}_n/\sqrt{n}$ is

$$\begin{aligned}\phi_n(t) &= \sum_{s \in ([n]-\nu_k n)\setminus\mathcal{J}_n} e^{its/\sqrt{n}}\mathbb{P}\left\{\partial\mathcal{O}_n = s\right\} + \sum_{s \in \mathcal{J}_n} e^{its/\sqrt{n}}\mathbb{P}\left\{\partial\mathcal{O}_n = s\right\} \\ &= o(1) + \sum_{s \in \mathcal{J}_n} e^{its/\sqrt{n}}\mathbb{P}\left\{\partial\mathcal{O}_n = s\right\}.\end{aligned}$$

Let $\mathcal{S}$ be a set of vertices with $|\mathcal{S}| = \nu_k n + s$ for some $s \in \mathcal{J}_n$. Recall that $\mathcal{O}_n = \mathcal{S}$ if and only if $\mathcal{S}$ is a k-surjection and $\mathcal{D}_{n,k}[\mathcal{S}^c]$ is acyclic, two events that are independent. By Theorem 5 in Section 3.2, $\mathbb{P}\{\mathcal{D}_{n,k}[\mathcal{S}^c] \text{ is acyclic}\} \sim 1 - ke^{-\tau_k}$. Also recall that K_x counts the number of k-surjections of size x. It follows from Lemma 1 that

$$\begin{aligned}\mathbb{P}\{\partial\mathcal{O}_n = s\} &= \sum_{\mathcal{S}\subseteq[n]:|\mathcal{S}|=\nu_k n+s} \mathbb{P}\{\mathcal{O}_n = \mathcal{S}\} \\ &= \sum_{\mathcal{S}\subseteq[n]:|\mathcal{S}|=\nu_k n+s} \mathbb{P}\{\mathcal{S} \text{ is a } k\text{-surjection}\} \times \mathbb{P}\{\mathcal{D}_{n,k}[\mathcal{S}^c] \text{ is acyclic}\} \\ &\sim (1-ke^{-\tau_k})\mathbb{E}K_{\nu_k n+s} \\ &= \sqrt{\frac{1-ke^{-\tau_k}}{2\pi}}\ \frac{1}{\sqrt{n}}\, g\left(\nu_k + \frac{s}{n}\right)\ \left[f\left(\nu_k+\frac{s}{n}\right)\right]^n,\end{aligned}$$

where K_x, $f(x)$ and $g(x)$ are defined as in the previous subsection.

If $s \in \mathcal{J}_n$, then Lemma A6 in the appendix shows that

$$g\left(\nu_k + \frac{s}{n}\right) = \left(1 + O\left(\frac{|s|}{n}\right)\right)\frac{1}{\sigma_k\sqrt{1-ke^{-\tau_k}}},$$

and

$$f\left(\nu_k + \frac{s}{n}\right) = \exp\left\{-\frac{s^2}{2\sigma_k^2 n^2}\right\} + O\left(\frac{|s|^3}{n^3}\right).$$

Therefore, choosing δ small enough, e.g., $\delta = 1/9$, we have

$$\begin{aligned}\sum_{s\in\mathcal{J}_n} e^{its/\sqrt{n}}\mathbb{P}\{\partial\mathcal{O}_n = s\} &\sim \frac{1}{\sqrt{2\pi\sigma_k^2}}\frac{1}{\sqrt{n}}\sum_{s\in\mathcal{J}_n} e^{its/\sqrt{n}}\exp\left\{-\frac{s^2}{2\sigma_k n}\right\} \\ &= o(1) + \frac{1}{\sqrt{2\pi\sigma_k^2}}\int_{-n^\delta}^{n^\delta} e^{itx}\exp\left\{-\frac{x^2}{2\sigma_k^2}\right\}\,\mathrm{d}x \\ &= o(1) + \frac{1}{\sqrt{2\pi\sigma_k^2}}\int_{-\infty}^{\infty} e^{itx}\exp\left\{-\frac{x^2}{2\sigma_k^2}\right\}\,\mathrm{d}x \\ &= o(1) + \exp\left(\frac{\sigma_k^2 t^2}{2}\right).\end{aligned}$$

Thus the characteristic function of $\partial\mathcal{O}_n/\sqrt{n}$ converges to $\exp(\sigma_k^2 t^2/2)$, the characteristic function of $\sigma_k\mathcal{Z}$. It follows from the central limit theorem that $\partial\mathcal{O}_n/\sqrt{n}$ converges to $\sigma_k\mathcal{Z}$ in distribution. Note that using the estimates of this section, we actually have a local limit theorem for $|\mathcal{O}_n|$.

3 The structure of the directed acyclic graph

3.1 De-randomizing the giant

Since a SCC induces a sub-digraph in which each vertex has in-degree at least one, a closed SCC is also a k-surjection. Lemma 1 implies that whp all k-surjections are of sizes in $\mathcal{I}_n \equiv [\nu_k n - n^{1/2+\delta}, \nu_k n + n^{1/2+\delta}]$. When this happens, as $\nu_k > 1/2$ (Lemma A1), there exists one and only one closed SCC and it is $\mathcal{G}_n$. And if $\mathcal{G}_n$ is the only closed SCC, then every vertex must be able to reach it. This can be summarized as:

Lemma 2. *Whp $|\mathcal{G}_n| \in \mathcal{I}_n$ and $\mathcal{G}_n$ is reachable from all vertices.*

Since $e^{-\tau_k} \equiv 1 - \tau_k/k \equiv 1 - \nu_k$, the above lemma implies that whp $||\mathcal{G}_n^c| - e^{-\tau_k} n| \leq n^{1/2+\delta}$. Thus the structure of $\mathcal{D}_{n,k}[\mathcal{G}_n^c]$, the sub-digraph induced by $\mathcal{G}_n^c \equiv [n] \setminus \mathcal{G}_n$, should be close to that of a sub-digraph induced by a fixed set of vertices whose size is close to $e^{-\tau_k} n$. Formally, we have:

Lemma 3. *Let f_n be a sequence of integer-valued functions on a sequence of digraphs. Let X be an integer-valued random variable. If there exists a sequence $\varepsilon_n \to 0$ such that*

$$\sup_{\mathcal{V}_n \subseteq [n] : |\mathcal{V}_n| \in \mathcal{I}_n} \| f_n(\mathcal{D}_{n,k}[\mathcal{V}_n^c]), X \|_{\text{TV}} \leq \varepsilon_n,$$

where $\mathcal{V}_n^c \equiv [n] \setminus \mathcal{V}_n$ and $\| \cdot, \cdot \|_{\text{TV}}$ denotes the total variation distance, then

$$f_n(\mathcal{D}_{n,k}[\mathcal{G}_n^c]) \xrightarrow{d} X.$$

Proof. Define the event $E_n = [|\mathcal{G}_n| \in \mathcal{I}_n]$. Let m be an integer, let $\mathcal{V}_n \subseteq [n]$ be a fixed set of vertices with $|\mathcal{V}_n| \in \mathcal{I}_n$. Recall that since $\nu_k > 1/2$, $|\mathcal{V}_n| > n/2$ for large n. Thus the event $[\mathcal{G}_n = \mathcal{V}_n]$ depends only on the induced sub-digraph $\mathcal{D}_{n,k}[\mathcal{V}_n]$, which is independent of $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$. Therefore the two events $[\mathcal{G}_n = \mathcal{V}_n]$ and $[f_n(\mathcal{D}_{n,k}[\mathcal{V}_n^c]) = m]$ are independent. Using this observation and Lemma 2, we have

$$\begin{aligned}
&\mathbb{P}\{f_n(\mathcal{D}_{n,k}[\mathcal{G}_n^c]) = m\} \\
&= \mathbb{P}\{[f_n(\mathcal{D}_{n,k}[\mathcal{G}_n^c]) = m] \cap E_n^c\} + \mathbb{P}\{[f_n(\mathcal{D}_{n,k}[\mathcal{G}_n^c]) = m] \cap E_n\} \\
&= o(1) + \sum_{\mathcal{V}_n \subseteq [n] : |\mathcal{V}_n| \in \mathcal{I}_n} \mathbb{P}\{f_n(\mathcal{D}_{n,k}[\mathcal{V}_n^c]) = m \mid \mathcal{G}_n = \mathcal{V}_n\} \mathbb{P}\{\mathcal{G}_n = \mathcal{V}_n\} \\
&\leq o(1) + \sum_{\mathcal{V}_n \subseteq [n] : |\mathcal{V}_n| \in \mathcal{I}_n} (\mathbb{P}\{X = m\} + \varepsilon_n) \mathbb{P}\{\mathcal{G}_n = \mathcal{V}_n\} \\
&\leq o(1) + \mathbb{P}\{X = m\}.
\end{aligned}$$

Similarly we have $\mathbb{P}\{f_n(\mathcal{D}_{n,k}[\mathcal{G}_n^c]) = m\} \geq \mathbb{P}\{X = m\} + o(1)$. Since this applies to all integers m, $f_n(\mathcal{D}_{n,k}[\mathcal{G}_n^c]) \xrightarrow{d} X$. □

Corollary 1. *Let $\mathcal{E}_n$ be a sequence of sets of digraphs. If there exists a sequence $\varepsilon_n \to 0$ such that*

$$\sup_{\mathcal{V}_n \subseteq [n] : |\mathcal{V}_n| \in \mathcal{I}_n} \mathbb{P}\{\mathcal{D}_{n,k}[\mathcal{V}_n^c] \notin \mathcal{E}_n\} \leq \varepsilon_n,$$

then whp $\mathcal{D}_{n,k}[\mathcal{G}_n^c] \in \mathcal{E}_n$.

Proof. This follows from the previous lemma by taking $X \equiv 1$ and f_n to be the indicator function that a digraph is in $\mathcal{E}_n$. $\square$

The rest of this section proves Theorem 2 and Theorem 3. But instead of working on $\mathcal{G}_n^c$ directly, we prove similar theorems on fixed sets of vertices, and then apply the above lemma or its corollary to get the final result.

3.2 Cycles outside the giant

In this subsection, we show the following:

Theorem 5. *Let $\omega_n \to \infty$ be an arbitrary sequence. There exists a sequence $\varepsilon_n = o(1)$ such that for all fixed sets of vertices $\mathcal{V}_n \subseteq [n]$ with $|\mathcal{V}_n| \in \mathcal{I}_n$, we have:*

(a) Let L_n^* be the length of the longest cycle in $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$. Then $\mathbb{P}\{L_n^* > \omega_n\} \le \varepsilon_n$.

(b) The probability that $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$ contains vertex-intersecting cycles is at most ε_n.

(c) Let $C_{n,\ell}^*$ be the number of cycles of length ℓ in $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$. Let $X_\ell = \mathrm{Poi}((ke^{-\tau_k})^\ell/\ell)$. Then for all fixed ℓ, $\left\|C_{n,\ell}^*, X_\ell\right\|_{\mathrm{TV}} \le \varepsilon_n$.

(d) Let C_n^* be the number of cycles in $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$. Let $X = \mathrm{Poi}(\log \frac{1}{1-ke^{-\tau_k}})$. Then $\|C_n^*, X\|_{\mathrm{TV}} \le \varepsilon_n$. As a result, $|\mathbb{P}\{C_n^* = 0\} - 1/(1-ke^{-\tau_k})| \le 2\varepsilon_n$.

Theorem 2 follows from the above theorem and Lemma 3. Our proof is inspired by Cooper and Frieze's work on the directed configuration model [12]. Note that the Cooper-Frieze model is different from that studied by us. In their model, both in-degrees and out-degrees are predetermined, whereas we require all out-degrees to be k but the in-degrees are random.

The intuition behind Theorem 5 is that when two cycles share vertices, they contain fewer vertices than arcs. So if we fix the "shape" of a pair of such cycles, the number of ways to label them times the probability that they both exist is $o(1)$. Thus whp cycles in $\mathcal{V}_n^c$ are vertex-disjoint and the total number of cycles has a distribution close to a sum of independent indicator random variables.

In the following proof, instead of finding the exact ε_n, we derive implicit $o(1)$ upper bounds for probabilities and total variation distances which only requires that $|\mathcal{V}_n| \in \mathcal{I}_n$.

Lemma 4. *Let $\overline{C_n^*} \equiv \sum_{1\le\ell\le\omega_n} C_{n,\ell}^*$. Then $\mathbb{P}\left\{C_n^* \ne \overline{C_n^*}\right\} = o(1)$.*

Proof. Define $(x)_\ell \equiv x(x-1)\cdots(x-\ell+1)$. Then the number of all possible cycles of length ℓ is $(|\mathcal{V}_n^c|)_\ell k^\ell/\ell$. (Note that we are also considering the labels on arcs, which makes the counting easier.) And the probability that such a cycle exists is $n^{-\ell}$. Recalling that $|\mathcal{V}_n^c| \in [e^{-\tau_k}n - n^{1/2+\delta}, e^{-\tau_k}n + n^{1/2+\delta}]$, we have

$$\mathbb{E}\left[C_{n,\ell}^*\right] = \frac{1}{\ell}(|\mathcal{V}_n^c|)_\ell k^\ell \left(\frac{1}{n}\right)^\ell \le \left(ke^{-\tau_k}\left(1 + O\left(n^{-1/2+\delta}\right)\right)\right)^\ell. \tag{1}$$

Since $ke^{-\tau_k} \equiv k - \tau_k < 1$ (Lemma A1), there exists a constant $c_1 < 1$ such that the above is less than c_1^ℓ for n large enough. Since $C_n^* \neq \overline{C_n^*}$ if and only if $\sum_{\ell > \omega_n} C_{n,\ell}^* \geq 1$,

$$\mathbb{P}\left\{C_n^* \neq \overline{C_n^*}\right\} = \mathbb{P}\left\{\sum_{\ell>\omega_n} C_{n,\ell}^* \geq 1\right\} \leq \mathbb{E}\left[\sum_{\ell>\omega_n} C_{n,\ell}^*\right] \leq O\left(c_1^{\omega_n}\right) = o(1). \qquad \square$$

Since $L_n^* > \omega_n$ if and only if $\overline{C_n^*} \neq C_n^*$, part (a) of Theorem 5 follows. From now on let $\omega_n = \log\log n$. We show that:

Lemma 5. *Let X and X_ℓ be as in Theorem 5. Then $\left\|\mathrm{Poi}(\mathbb{E}\overline{C_n^*}), X\right\|_{\mathrm{TV}} = o(1)$. And for all $\ell \leq \omega_n$, $\left\|\mathrm{Poi}(\mathbb{E}C_{n,\ell}^*), X_\ell\right\|_{\mathrm{TV}} = o(1)$.*

Proof. For all $\ell \leq \omega_n$, by (1) we have

$$\mathbb{E}C_{n,\ell}^* = \frac{1}{\ell}\left(e^{-\tau_k} n + O\left(n^{1/2+\delta}\right)\right)_\ell k^\ell \left(\frac{1}{n}\right)^\ell = \frac{(ke^{-\tau_k})^\ell}{\ell}(1 + O(\ell n^{-1/2+\delta})).$$

Thus

$$\mathbb{E}\overline{C_n^*} = \sum_{1\leq\ell\leq\omega_n} \mathbb{E}\left[C_{n,\ell}^*\right] = \log\left(\frac{1}{1-ke^{-\tau_k}}\right) + O\left(\omega_n n^{-1/2+\delta}\right).$$

Therefore $\mathbb{E}\overline{C_n^*} \to \mathbb{E}X$ and $\mathbb{E}C_{n,\ell}^* \to \mathbb{E}X_\ell$, which implies the lemma. $\square$

Proof of Theorem 5. By the two previous lemmas, it suffices to show that

$$\left\|\overline{C_n^*}, \mathrm{Poi}(\mathbb{E}\overline{C_n^*})\right\|_{\mathrm{TV}} = o(1), \quad \left\|C_{n,\ell}^*, \mathrm{Poi}(\mathbb{E}C_{n,\ell}^*)\right\|_{\mathrm{TV}} = o(1) \quad \text{for all fixed } \ell.$$

We prove this by using a theorem of Arratia et al. [4]. (A similar result is proved by Barbour et al. [6]). The method is known as the Chen-Stein method because it was first developed by Chen [11] who applied Stein's theory [38] on probability metrics to Poisson distributions.

Let $\mathcal{C}$ be the space of all possible cycles of length at most ω_n in $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$. For $\alpha \in \mathcal{C}$, let $\mathcal{B}_\alpha \subseteq \mathcal{C}$ be the set of cycles that are vertex-intersecting with α. Let $\mathbb{1}_\alpha$ be the indicator that a cycle α appears in $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$. Define

$$b_1 \equiv \sum_{\alpha\in\mathcal{C}} \sum_{\beta\in\mathcal{B}_\alpha} \mathbb{E}\mathbb{1}_\alpha \mathbb{E}\mathbb{1}_\beta, \qquad b_2 \equiv \sum_{\alpha\in\mathcal{C}} \sum_{\beta\in\mathcal{B}_\alpha:\beta\neq\alpha} \mathbb{E}\left[\mathbb{1}_\alpha\mathbb{1}_\beta\right], \qquad b_3 \equiv \sum_{\alpha\in\mathcal{C}} s_\alpha,$$

where

$$s_\alpha = \mathbb{E}\left|\mathbb{E}\left[\mathbb{1}_\alpha | \sigma\left(\mathbb{1}_\beta : \beta \in \mathcal{C}\setminus\mathcal{B}_\alpha\right)\right] - \mathbb{E}\mathbb{1}_\alpha\right|,$$

and $\sigma(\cdot)$ denotes the sigma algebra generated by $(\cdot)$. Theorem 1 of Arratia et al. [4] states that

$$\left\|\overline{C_n^*},\ \mathrm{Poi}(\mathbb{E}\overline{C_n^*})\right\|_{\mathrm{TV}} \leq 2(b_1 + b_2 + b_3).$$

If $\beta \in \mathcal{C}\setminus\mathcal{B}_\alpha$, then α and β are vertex-disjoint. Thus $\mathbb{1}_\alpha$ and $\mathbb{1}_\beta$ are independent and $s_\alpha = 0$ for all $\alpha \in \mathcal{C}$, i.e., $b_3 = 0$. It suffices to show that b_1 and b_2 are $o(1)$.

Let $|\alpha|$ denote the length of a cycle α. Fix $\ell_1 \leq \omega_n$ and $\ell_2 \leq \omega_n$. There are at most $|\mathcal{V}_n^c|^{\ell_1} k^{\ell_1}$ cycles of length ℓ_1. For $|\alpha| = \ell_1$, there are at most $\ell_1 |\mathcal{V}_n^c|^{\ell_2 - 1} k^{\ell_2}$ cycles of length ℓ_2 that share at least one vertex with α. Since $(|\mathcal{V}_n^c|)^\ell = (1+o(1))(e^{-\tau_k} n)^\ell$ for $\ell \leq \omega_n$,

$$\begin{aligned}\sum_{\alpha \in \mathcal{C}:|\alpha|=\ell_1} \sum_{\beta \in \mathcal{B}_\alpha:|\beta|=\ell_2} \mathbb{E}\mathbb{1}_\alpha \mathbb{E}\mathbb{1}_\beta &\leq (1+o(1)) \left[(e^{-\tau_k} n)^{\ell_1} k^{\ell_1}\right] \left[\ell_1 (e^{-\tau_k} n)^{\ell_2 - 1} k^{\ell_2}\right] \left(\frac{1}{n}\right)^{\ell_1+\ell_2} \\ &= (1+o(1)) \frac{1}{e^{-\tau_k} n} \left[\ell_1 (e^{-\tau_k} k)^{\ell_1}\right] \left[(e^{-\tau_k} k)^{\ell_2}\right].\end{aligned}$$

Therefore

$$\begin{aligned} b_1 &= \sum_{1\leq \ell_1 \leq \omega_n} \sum_{1 \leq \ell_2 \leq \omega_n} \sum_{\alpha \in \mathcal{C}:|\alpha|=\ell_1} \sum_{\beta \in \mathcal{B}_\alpha:|\beta|=\ell_2} \mathbb{E}\mathbb{1}_\alpha \mathbb{E}\mathbb{1}_\beta \\ &\leq (1+o(1)) \frac{1}{e^{-\tau_k} n} \sum_{\ell_1 \geq 1} \sum_{\ell_2 \geq 1} \left[\ell_1 (k e^{-\tau_k})^{\ell_1}\right] \left[(k e^{-\tau_k})^{\ell_2}\right] \\ &\leq (1+o(1)) \frac{1}{e^{-\tau_k} n} \left[\sum_{\ell_1 \geq 1} \ell_1 (k e^{-\tau_k})^{\ell_1}\right] \left[\sum_{\ell_2 \geq 1} (k e^{-\tau_k})^{\ell_2}\right]\end{aligned}$$

which is $O(1/n)$ since both sums converge.

To compute b_2, we upper bound the number of pairs of vertex-intersecting cycles that could possibly appear in $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$ at the same time. Let α and β be such a pair. Let $V(\alpha), A(\alpha), V(\beta), A(\beta)$ be the vertex set and (labeled) arc set of α and β respectively. Let $\alpha \cup \beta$ be the digraph of vertex set $V = V(\alpha) \cup V(\beta)$ and arc set $A = A(\alpha) \cup B(\beta)$. Assume that $|V| = s$ and $|A| = s + t$. Note that as α and β share at least one vertex, $t \geq 1$. Since $V \subset [n]$, we can relabel the s vertices in $\alpha \cup \beta$ with $[s]$ such that the order of the vertex labels is maintained. The result is a digraph with vertex set $[s]$ and $s+t$ arcs labeled with $[k]$. There are at most $(s^2)^{s+t} k^{s+t}$ such digraphs, since there are at most s^2 choices of endpoints and k choices of labels for each of the $s+t$ arcs. Each digraph of this type corresponds to at most $\binom{|\mathcal{V}_n^c|}{s} \leq |\mathcal{V}_n^c|^s$ pairs of cycles like α and β. Thus there are at most $|\mathcal{V}_n^c|^s (s^2)^{s+t} k^{s+t}$ such pairs. Summing over s and t, we have

$$\begin{aligned} b_2 &\leq \sum_{1 \leq s \leq 2\omega_n} \sum_{1 \leq t \leq 2\omega_n} |\mathcal{V}_n^c|^s (s^2)^{s+t} k^{s+t} \mathbb{E}\left[\mathbb{1}_\alpha \mathbb{1}_\beta\right] \\ &\leq \sum_{1 \leq s \leq 2\omega_n} \sum_{1 \leq t \leq 2\omega_n} \left(e^{-\tau_k} n + n^{1/2+\delta}\right)^s (2\omega_n)^{2\times 4\omega_n} k^{s+t} \frac{1}{n^{s+t}} \\ &\leq (2\omega_n)^{8\omega_n} \sum_{1 \leq s \leq 2\omega_n} \sum_{1 \leq t \leq 2\omega_n} \frac{\left(n + e^{\tau_k} n^{1/2+\delta}\right)^s}{n^s} (k e^{-\tau_k})^s \frac{k^t}{n^t} \qquad (2) \\ &\leq O\left(\frac{1}{n}\right) (2\omega_n k)^{8\omega_n} \sum_{1 \leq s \leq 2\omega_n} \sum_{1 \leq t \leq 2\omega_n} (1 + e^{\tau_k} n^{-1/2+\delta})^{2\omega_n} \qquad (k e^{-\tau_k} < 1/2) \\ &\leq O\left(\frac{1}{n}\right) (2\omega_n k)^{8\omega_n} (2\omega_n)^2 \left(1 + O\left(n^{-1/2+\delta} \omega_n\right)\right) \to 0,\end{aligned}$$

where the last step we use that $\omega_n = \log\log n$.

Thus part (d) of Theorem 5 for C_n^* is proved. We can prove part (c) for $C_{n,\ell}^*$ using the same method by limiting $\mathcal{C}$ to contain only cycles of a fixed length ℓ. Note that the above inequality shows that the probability that there exist vertex-intersecting cycles in $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$ is $o(1)$, thus part (b) is also proved. □

The method used above can be easily adapted to prove similar results for undirected cycles, like the following lemma which is needed in the study of spectra in $\mathcal{D}_{n,k}[\mathcal{G}_n^c]$:

Lemma 6. *Let $\psi_n \to \infty$ be an arbitrary sequence. There exists a sequence $\varepsilon_n = o(1)$ such that for all fixed sets of vertices $\mathcal{V}_n$ with $|\mathcal{V}_n| \in \mathcal{I}_n$, we have:*

(a) The probability that $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$ contains an undirected cycle of length greater than ψ_n is at most ε_n.

(b) The probability that $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$ contains vertex-intersecting undirected cycles is at most ε_n.

Proof. Let U_ℓ be the number of undirected cycles of length ℓ in $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$. Then

$$\mathbb{E}\left[U_\ell\right] \leq \frac{1}{\ell}(|\mathcal{V}_n^c|)^\ell (2k)^\ell \frac{1}{n^\ell} \leq \left(2ke^{-\tau_k}(1+n^{-1/2+\delta})\right)^\ell,$$

where the 2 comes from the fact that each edge in an undirected cycle has two possible directions. Since $2ke^{-\tau_k} = 2(k-\tau_k) < 1$ (Lemma A1), with exact the same argument of Lemma 4, we can show that $\mathbb{E}\left[\sum_{\ell>\psi_n} U_\ell\right] = o(1)$ for all $\psi_n \to \infty$. Thus (a) is proved.

Now choose $\psi_n = \log\log n$. Again we can show that whp there are no vertex-intersecting undirected cycles of length at most ψ_n by repeating the computation of b_2 in the proof of Theorem 5 with $ke^{-\tau_k}$ replaced by $2ke^{-\tau_k}$ in (2). □

3.3 Spectra outside the giant

In this section, we prove Theorem 3 (spectra outside the giant). Instead of working on $\mathcal{G}_n^c$ directly, we again prove similar results on a fixed set of vertices and then apply Lemma 3 to finish the proof.

3.3.1 The tree-like structure of some spectra

We prove part (a) of Theorem 3. Let $\mathcal{V}_n \subseteq [n]$ with $|\mathcal{V}_n| \in \mathcal{I}_n \equiv [\nu_k n - n^{1/2+\delta}, \nu_k n + n^{1/2+\delta}]$ be a fixed set of vertices. For $v \in \mathcal{V}_n^c \equiv [n] \setminus \mathcal{V}_n$, let $\mathcal{S}_v^*$ be the spectrum of v in $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$, the sub-digraph induced by $\mathcal{V}_n^c$. The following lemma shows that whp every spectrum in $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$ induces a sub-digraph that is a tree or a tree plus one extra arc:

Lemma 7. *We have*

$$\sup_{\mathcal{V}_n \subseteq [n]: |\mathcal{V}_n| \in \mathcal{I}_n} \mathbb{P}\left\{\cup_{v \in \mathcal{V}_n^c}[\mathrm{arc}(\mathcal{D}_{n,k}[\mathcal{S}_v^*]) - |\mathcal{S}_v^*| \geq 1]\right\} = o(1),$$

where $\mathrm{arc}(\cdot)$ *denotes the number of arcs.*

Proof. For $v \in \mathcal{V}_n^c$, if $\mathrm{arc}(\mathcal{D}_{n,k}[\mathcal{S}_v^*]) \geq |\mathcal{S}_v^*| + 1$, then $\mathcal{D}_{n,k}[\mathcal{S}_v^*]$ must contain at least two undirected cycles. By Lemma 6, whp all undirected cycles in $\mathcal{D}_{n,k}[\mathcal{S}_v^*]$ are vertex-disjoint. Therefore, if $\mathcal{D}_{n,k}[\mathcal{S}_v^*]$ contains two undirected cycles, then whp they are vertex-disjoint and connected by an undirected path.

Let $X_{r,s,t}$ be the number of pairs of undirected cycles of length r and s respectively that are connected by an undirected path of length t. In such a structure the number of arcs is $r+s+t$ while the number of vertices is $r+s+t-1$. Since $|\mathcal{V}_n| \in \mathcal{I}_n$, we have $|\mathcal{V}_n^c| = n - |\mathcal{V}_n| \in \mathcal{I}_n^c \equiv [e^{-\tau_k} n - n^{1/2+\delta}, e^{-\tau_k} n + n^{1/2+\delta}]$. Thus

$$\mathbb{E}X_{r,s,t} \leq (|\mathcal{V}_n^c|)^{r+s+t-1} (2k)^{r+s+t} \left(\frac{1}{n}\right)^{r+s+t} \leq O\left(\frac{1}{n}\right) \left(2ke^{-\tau_k} + \frac{2k}{n^{1/2-\delta}}\right)^{r+s+t}.$$

Summing over all possible r, s and t shows that

$$\begin{aligned}\sum_{1\leq r\leq n} \sum_{1\leq s\leq n} \sum_{1\leq t\leq n} \mathbb{E}X_{r,s,t} &\leq O\left(\frac{1}{n}\right) \sum_{1\leq r} \sum_{1\leq s} \sum_{1\leq t} \left(2ke^{-\tau_k} + \frac{2k}{n^{1/2-\delta}}\right)^{r+s+t} \\ &\leq O\left(\frac{1}{n}\right) \left(\sum_{1\leq i} \left(2ke^{-\tau_k} + \frac{2k}{n^{1/2-\delta}}\right)^{i}\right)^{3},\end{aligned}$$

which is $o(1)$ since the sum in the brackets converges. □

3.3.2 The maximum size of spectra

This section proves part (b) of Theorem 3 (the sizes of spectra outside the giant).

Lemma 8. *Let $\varepsilon > 0$ be a constant. Then*

$$\sup_{\mathcal{V}_n \subseteq [n]: |\mathcal{V}_n| \in \mathcal{I}_n} \mathbb{P}\left\{ \left| \frac{\max_{v\in\mathcal{V}_n^c} |\mathcal{S}_v^*|}{\log n} - \frac{1}{\log(1/\lambda_k)} \right| > \varepsilon \right\} = o(1),$$

where $\lambda_k \equiv (k-\tau_k)\left(\frac{\tau_k}{k-1}\right)^{k-1}$.

The exploration of $\mathcal{D}_{n,k}[\mathcal{S}_v^*]$ can be coupled with a colouring process. Initially, colour all vertices in $\mathcal{V}_n$ green, all vertices in $\mathcal{V}_n^c$ yellow, and all arcs white. Then:

(i) Colour the vertex v black, and colour the k arcs that start from v red. (Red arcs start from vertices in $\mathcal{S}_v^*$ but their endpoints are not determined yet.)

(ii) Pick an arbitrary red arc. Choose its endpoint uniformly at random from all the n vertices. Colour this arc with the colour of its chosen endpoint vertex. (So a yellow arc goes to a vertex that is not already in $\mathcal{S}_v^*$, a black arc goes to a vertex that is already in $\mathcal{S}_v^*$.) If the chosen vertex is yellow, colour this vertex black and colour all its arcs red.

(iii) If there are no red arcs left, terminate. Otherwise go to the previous step.

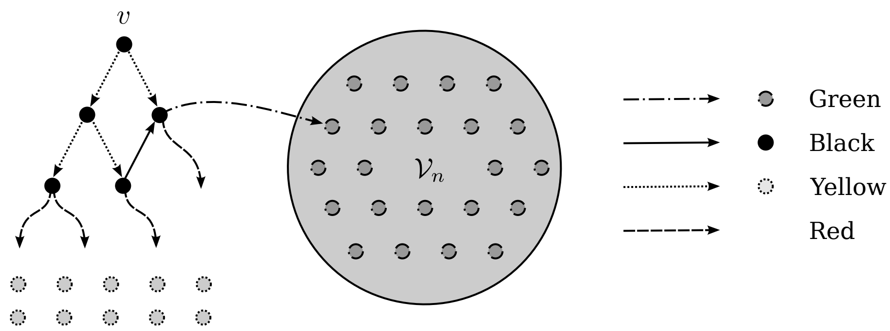

Figure 2: The colouring process.

In the end, $\mathcal{S}_v^*$ consists of all black vertices, and arcs that start from vertices in $\mathcal{S}_v^*$ have one of three colors: green arcs go to $\mathcal{V}_n$; yellow arcs form a spanning tree of $\mathcal{D}_{n,k}[\mathcal{S}_v^*]$ rooted at v; black arcs connect vertices in $\mathcal{S}_v^*$ but they are not part of the yellow spanning tree, so they are in cycles in $\mathcal{D}_{n,k}[\mathcal{S}_v^*]$. Figure 2 depicts the colouring process.

We use random variables R_t and Y_t to track the number of red arcs and yellow vertices after the t-th red arc is colored. Thus $R_0 = k$ and $Y_0 = |\mathcal{V}_n^c| - 1$. When a red arc is colored, if a yellow vertex is chosen as its endpoint, then the number of red arcs increases by $(k-1)$ and the number of yellow vertices decreases by one. Otherwise the number of red arcs decreases by one and the number of yellow vertices remains unchanged. Thus for $t \geq 1$,

$$R_t = R_{t-1} + k\xi_t - 1 = k \sum_{i=1}^{t} \xi_i - (t-k), \quad \text{and} \quad Y_t = Y_{t-1} - \xi_t = |\mathcal{V}_n^c| - 1 - \sum_{i=1}^{t} \xi_i,$$

where ξ_t are independent Bernoulli Y_t/n (the probability that a yellow vertex is chosen). Let $T \equiv \min\{t : R_t \leq 0\}$. Then $|\mathcal{S}_v^*| = T/k$, since T is the total number arcs that have been colored and $|\mathcal{S}_v^*|$ is the total number of vertices that have been colored.

Let $(\overline{\xi_t})_{t\geq 1}$, be i.i.d. Bernoulli $(e^{-\tau_k} + n^{-1/2+\delta})$. Since $Y_t/n \leq |\mathcal{V}_n^c|/n \leq e^{-\tau_k} + n^{-1/2+\delta}$, we have $\overline{\xi_t} \succeq \xi_t$, where $\succeq$ denotes stochastically greater than (see [36]). Therefore there exists a coupling such that $\overline{\xi_t} \geq \xi_t$ for all t almost surely. Let $\overline{T}_t \equiv \min\{t : k\sum_{i=1}^{t} \overline{\xi_i} - (t-k) \leq 0\}$. Then $\overline{T} \geq T$ almost surely. (The random variable $\underline{T}$ is called the total progeny of a Galton-Watson process with offspring distribution $\overline{\xi_1}$. For an introduction to Galton-Watson processes see [13]). It is well know that if $\mathbb{E}\overline{\xi_1} < 1$, which is true in this case, then $\mathbb{E}\overline{T} = k/(1 - \mathbb{E}\overline{\xi_1}) = O(1)$. Thus $\mathbb{E}T = O(1)$.

Proof of the upper bound. Let $\omega_n = \lfloor (1+\varepsilon)\log n / \log(1/\lambda_k) \rfloor + 1$. Since $\overline{T} \geq T$,

$$\mathbb{P}\left\{T \geq k\omega_n\right\} \leq \mathbb{P}\left\{\overline{T} \geq k\omega_n\right\} \leq \mathbb{P}\left\{\frac{\sum_{i=1}^{k\omega_n} \overline{\xi_i}}{k\omega_n} \geq \frac{1}{k_n}\right\}$$

where $k_n = k\omega_n/(\omega_n - 1)$. Hoeffding [21] showed that

$$\mathbb{P}\left\{\frac{\mathrm{Bin}(m,p)}{m} \geq p + x\right\} \leq \left\{\left(\frac{p}{p+x}\right)^{p+x} \left(\frac{1-p}{1-p-x}\right)^{1-p-x}\right\}^m.$$

where $\mathrm{Bin}(m,p)$ denotes a binomial (m,p) random variable. Recalling that $\mathbb{E}\overline{\xi_1} = e^{-\tau_k} + n^{-1/2+\delta} \equiv 1 - \tau_k/k + n^{-1/2+\delta}$ and $\lambda_k \equiv (k-\tau_k)\left(\frac{\tau_k}{k-1}\right)^{k-1}$, it follows from Hoeffding's inequality that $\mathbb{P}\{T \geq k\omega_n\}$ is at most

$$\begin{aligned}
\left[\left(\frac{\mathbb{E}\overline{\xi_1}}{1/k_n}\right)\left(\frac{1-\mathbb{E}\overline{\xi_1}}{1-1/k_n}\right)^{k_n-1}\right]^{\omega_n} &= \left[(k-\tau_k)\left(\frac{\tau_k}{k-1}\right)^{k-1} + O(n^{-1/2+\delta})\right]^{\omega_n+O(1)} \\
&= O(\lambda_k^{\omega_n})\left(1+O\left(n^{-1/2+\delta}\right)\right)^{\omega_n} \\
&= O\left(n^{-(1+\varepsilon)}\right). \qquad (3)
\end{aligned}$$

Since $k|\mathcal{S}_v^*| = T$, by the union bound

$$\mathbb{P}\left\{\cup_{v\in\mathcal{V}_n^c}|\mathcal{S}_v^*| \geq \omega_n\right\} \leq n\mathbb{P}\left\{\overline{T} \geq k\omega_n\right\} = O\left(n^{-\varepsilon}\right). \qquad \square$$

Proof of the lower bound. Let $\psi_n \equiv \lceil (1-\varepsilon)\log n/\log(1/\lambda_k)\rceil$. To show that whp there exists a $v \in \mathcal{V}_n^c$ such that $|\mathcal{S}_v^*| \geq \psi_n$, pick an arbitrary yellow vertex and run the colouring process. If at least ψ_n vertices are colored black (success) in the process then terminate. Otherwise (failure) pick another yellow vertex and repeat the colouring process until one trial succeeds. If the colouring process is repeated for at most $t_n \equiv \lfloor n/(\log n)^3 \rfloor$ times, then at most $a_n \equiv t_n\psi_n = O(n/(\log n)^2)$ vertices are colored black in the end. Therefore, the probability that the number of red arcs increases after colouring one red arc is at least $(|\mathcal{V}_n^c| - a_n)/n$.

Let $(\underline{\xi_i})_{i\geq 1}$ be i.i.d. Bernoulli $(|\mathcal{V}_n^c| - a_n - \psi_n)/n$. Let $\underline{T} = \min\{t : k\sum_{i=1}^t \underline{\xi_i} - (t-k) \leq 0\}$. Then in each of the first t_n iterations, the probability of a success is at least $\mathbb{P}\{\underline{T} \geq k\psi_n\} \geq \mathbb{P}\{\underline{T} = k\psi_n\}$. (For a detailed proof, see van der Hofstad's discussion of the Erdős–Rényi model [39, chap. 4.2.2].) By the hitting-time theorem of Galton-Watson processes [41],

$$\mathbb{P}\{\underline{T} = k\psi_n\} = \frac{1}{\psi_n}\mathbb{P}\left\{k\sum_{i=1}^{k\psi_n}\underline{\xi_i} = k(\psi_n - 1)\right\}.$$

Since $\sum_{i=1}^{k\psi_n}\underline{\xi_i}$ is a binomial random variable, the above equals

$$\frac{1}{\psi_n}\binom{k\psi_n}{\psi_n - 1}\left(\frac{|\mathcal{V}_n^c| - a_n - \psi_n}{n}\right)^{\psi_n - 1}\left(1 - \frac{|\mathcal{V}_n^c| - a_n - \psi_n}{n}\right)^{k\psi_n - (\psi_n - 1)} \equiv b_n.$$

By Stirling's approximation [17, pp. 407]

$$\binom{k\psi_n}{\psi_n - 1} = \Theta(1)\binom{k\psi_n}{\psi_n} = \frac{1}{\Theta\left(\sqrt{\psi_n}\right)}\left[\frac{k}{(1-1/k)^{k-1}}\right]^{\psi_n}.$$

Recalling that $a_n \equiv O\left(n/(\log n)^2\right)$ and $\psi_n \equiv \lceil (1-\varepsilon)\log n/\log(1/\lambda_k)\rceil$, we have, in view of $|\mathcal{V}_n^c| = e^{-\tau_k}n + O\left(n^{1/2+\delta}\right)$,

$$\left(\frac{|\mathcal{V}_n^c| - a_n - \psi_n}{n}\right)^{\psi_n - 1} = \left(e^{-\tau_k} - O\left(\frac{1}{(\log n)^2}\right)\right)^{\psi_n - 1} = \Theta\left(e^{-\tau_k\psi_n}\right),$$

and

$$\begin{aligned}\left(1-\frac{|\mathcal{V}_n^c|-a_n-\psi_n}{n}\right)^{k\psi_n-(\psi_n-1)} &= \left(1-e^{-\tau_k}+O\left(\frac{1}{(\log n)^2}\right)\right)^{k\psi_n-(\psi_n-1)} \\ &= \Theta\left(\left(\frac{\tau_k}{k}\right)^{(k-1)\psi_n}\right).\end{aligned}$$

Recall that $e^{-\tau_k} \equiv 1-\tau_k/k$. Therefore

$$\lambda_k \equiv (k-\tau_k)\left(\frac{\tau_k}{k-1}\right)^{k-1} = ke^{-\tau_k}\left(\frac{\tau_k}{k-1}\right)^{k-1} = \frac{k}{(1-1/k)^{k-1}}e^{-\tau_k}\left(\frac{\tau_k}{k}\right)^{k-1}.$$

Putting everything together, we have

$$b_n = \Theta\left(\frac{1}{\psi_n}\frac{1}{\sqrt{\psi_n}}\left[\frac{k}{(1-1/k)^{k-1}}e^{-\tau_k}\left(\frac{\tau_k}{k}\right)^{k-1}\right]^{\psi_n}\right) = \Theta\left(\frac{\lambda_k^{\psi_n}}{\psi_n^{3/2}}\right) = \Theta\left(\frac{n^{-1+\varepsilon}}{\psi_n^{3/2}}\right).$$

So the probability that all the first $t_n \equiv \lfloor n/(\log n)^3 \rfloor$ trials fail is at most

$$(1-b_n)^{t_n} \le \exp\{-b_n t_n\} = \exp\left\{\Theta\left(-\frac{n^{\varepsilon}}{(\log n)^{9/2}}\right)\right\} = o(1). \qquad \square$$

By Lemma 2, whp $\mathcal{G}_n$ is reachable from all vertices. When this happens, $\mathcal{O}_n \setminus \mathcal{G}_n$ consists of vertices either on cycles in $\mathcal{D}_{n,k}[\mathcal{G}_n^c]$ or on paths from these cycles to $\mathcal{G}_n$. Since the number of such cycles and the length of the longest one of them are both $O_p(1)$, Lemma 8 implies that $|\mathcal{O}_n| - |\mathcal{G}_n| = O_p(\log n)$. Thus

$$\frac{|\mathcal{G}_n|-\nu_k n}{\sqrt{n}} = \frac{|\mathcal{O}_n|-\nu_k n}{\sqrt{n}} - O_p\left(\frac{\log n}{\sqrt{n}}\right) \xrightarrow{d} \mathcal{Z},$$

which is the second part of Theorem 1.

In fact we can show that $|\mathcal{O}_n| - |\mathcal{G}_n| = O_p(1)$. This seems to be obvious since in $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$ the expected size of a spectrum is $O(1)$ and the number of cycles is $O_p(1)$. However, it is not trivial because $\mathbb{1}_{[v \text{ is on a cycle}]}$ and $|\mathcal{S}_v^*|$ are not independent. For a proof using Cayley's formula, see Lemma 9 in the next section (Section 3.3.3).

We can also use Lemma 8 to show that

$$\frac{\max_{v\in[n]} |\mathcal{S}_v| - |\mathcal{G}_n|}{\log n} \xrightarrow{p} \frac{1}{\log(1/\lambda_k)},$$

which finishes the last part of Theorem 1, i.e., $(\max_{v\in[n]} |\mathcal{S}_v| - \nu_k n)/\sigma_k\sqrt{n} \xrightarrow{d} \mathcal{Z}$. Let A_n be the event that every vertex can reach $\mathcal{G}_n$. Assuming A_n happens, $\mathcal{G}_n \subseteq \mathcal{S}_v$ for all $v \in [n]$. Thus for all $\varepsilon > 0$,

$$\begin{aligned}&\mathbb{P}\left\{\left|\frac{\max_{v\in[n]} |\mathcal{S}_v| - |\mathcal{G}_n|}{\log n} - \frac{1}{\log(1/\lambda_k)}\right| > \varepsilon\right\} \\ &\le \mathbb{P}\left\{\left[\left|\frac{\max_{v\in[n]} |\mathcal{S}_v'|}{\log n} - \frac{1}{\log(1/\lambda_k)}\right| > \varepsilon\right] \cap A_n\right\} + \mathbb{P}\{A_n^c\} = o(1).\end{aligned}$$

Since $|\mathcal{S}_1| \le \max_{v\in[n]} |\mathcal{S}_v|$ and whp $|\mathcal{S}_1| \ge |\mathcal{G}_n|$, we also recover Grusho's central limit law of $|\mathcal{S}_1|$.

3.3.3 The size of the middle layer

Lemma 9 and Corollary 1 imply that $|\mathcal{O}_n| - |\mathcal{G}_n| = O_p(1)$.

Lemma 9. *Let $\omega_n \to \infty$ be an arbitrary sequence of nonnegative numbers. Then*

$$\sup_{\mathcal{V}_n \subseteq [n]:|\mathcal{V}_n| \in \mathcal{I}_n} \mathbb{P}\left\{ \sum_{v \in \mathcal{C}(\mathcal{V}_n^c)} |\mathcal{S}_v^*| \geq \omega_n \right\} = o(1),$$

where $\mathcal{C}(\mathcal{V}_n^c)$ denotes the set of vertices on cycles in $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$, and $\mathcal{S}_v^*$ is the spectrum of v in $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$, the sub-digraph induced by $\mathcal{V}_n^c$.

Proof. By Theorem 5 and Lemma 7, in $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$ whp: (a) there are at most $\sqrt{\omega_n}$ vertices on cycles, i.e., $|\mathcal{C}(\mathcal{V}_n^c)| \leq \sqrt{\omega_n}$; (b) every $\mathcal{S}_v^*$ induces either a tree or a tree plus one extra arc; (c) $\max_{v \in \mathcal{G}_n^c} |\mathcal{S}_v^*| = O(\log n)$. Now assume all these events happen. If $\sum_{v \in \mathcal{C}(\mathcal{V}_n^c)} |\mathcal{S}_v^*| \geq \omega_n$, then (a) implies there is at least one vertex $u \in \mathcal{C}(\mathcal{V}_n^c)$ with $|\mathcal{S}_u^*| \geq \sqrt{\omega_n}$. By (b), $\mathcal{S}_u^*$ induces a sub-digraph that consists of exactly one cycle and isolated trees with their roots on this cycle. If $|\mathcal{S}_u^*| = \ell$, we call the induced sub-digraph an ℓ-eye. Note that by (c) there are no ℓ-eyes with $\ell > (\log n)^2$.

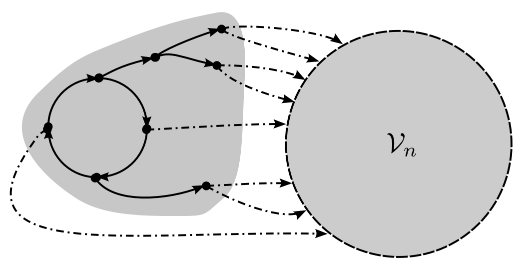

Figure 3: The leftmost shaded part of this figure is an ℓ-eye.

Let $\mathcal{S} \subseteq \mathcal{V}_n^c$ with $|\mathcal{S}| = \ell$ be a set of vertices. If $\mathcal{S}$ induces an ℓ-eye $\mathcal{D}_e$, then there are ℓ arcs that start and end at specific vertices in $\mathcal{S}$ decided by $\mathcal{D}_e$, which happens with probability $(1/n)^\ell$. If $\mathcal{S} = \mathcal{S}_u^*$ for some vertex $u \in \mathcal{S}$, call $\mathcal{S}$ a *partial spectrum*. For $\mathcal{S}$ to be a partial spectrum, the other $(k-1)\ell$ arcs that start from $\mathcal{S}$ must end at $\mathcal{V}_n$, which happens with probability $(|\mathcal{V}_n|/n)^{(k-1)\ell}$. So the probability that $\mathcal{S}$ induces a fixed $\mathcal{D}_e$ and $\mathcal{S}$ is a partial spectrum is $(1/n)^\ell(|\mathcal{V}_n|/n)^{(k-1)\ell}$.

By Cayley's formula [7], there are $\ell^{\ell-1}$ ways that $\mathcal{S}$ can form a rooted tree. In such a tree, there are at most ℓ^2 ways to add an extra arc to make it an ℓ-eye. In a vertex-labeled ℓ-eye, there are at most k^ℓ ways to label the arcs. So the number of ℓ-eyes can be induced by $\mathcal{S}$ is less than $\ell^{\ell-1}\ell^2 k^\ell$. And there are $\binom{|\mathcal{V}_n^c|}{\ell}$ ways to choose $\mathcal{S}$.

Let X_ℓ be the number of ℓ-eyes induced by partial spectra. Recall that $\nu_k \equiv \tau_k/k = 1 - e^{-\tau_k}$. Thus $|\mathcal{V}_n| \in \mathcal{I}_n \equiv [\nu_k n - n^{1/2+\delta}, \nu_k n + n^{1/2+\delta}]$ implies that $|\mathcal{V}_n^c| \leq e^{-\tau_k} n + n^{1/2+\delta}$.

So for $\ell \le (\log n)^2$, by the above arguments,

$$\begin{aligned}
\mathbb{E}X_\ell &\le \binom{|\mathcal{V}_n^c|}{\ell} \ell^{\ell-1}\ell^2 k^\ell \left(\frac{1}{n}\right)^\ell \left(\frac{|\mathcal{V}_n|}{n}\right)^{(k-1)\ell} \\
&\le \frac{(e^{-\tau_k} n + n^{1/2+\delta})^\ell}{(\ell/e)^\ell} \ell^{\ell+1} k^\ell \left(\frac{1}{n}\right)^\ell \left(\frac{\tau_k}{k} + n^{-1/2+\delta}\right)^{(k-1)\ell} \\
&= \left[e\left(e^{-\tau_k} + n^{-1/2+\delta}\right) k \left(\frac{\tau_k}{k} + n^{-1/2+\delta}\right)^{k-1}\right]^\ell \ell \\
&= \left(1 + O\left(\ell n^{-1/2+\delta}\right)\right) \left(k e^{1-\tau_k} \left(\frac{\tau_k}{k}\right)^{k-1}\right)^\ell \ell \\
&\equiv \left(1 + O\left(\ell n^{-1/2+\delta}\right)\right) \rho_k^\ell \ell.
\end{aligned}$$

By Lemma A1, $\rho_k < 1$. Since $\sqrt{\omega_n} \to \infty$,

$$\sum_{\sqrt{\omega_n} \le \ell \le (\log n)^2} \mathbb{E}X_\ell \le \left[1 + O\left(\frac{(\log n)^2}{n^{1/2-\delta}}\right)\right] \sum_{\sqrt{\omega_n} \le \ell}^{\infty} \ell(\rho_k)^\ell = o(1).$$

Thus whp there are no ℓ-eyes induced by partial spectra with $\ell \in [\sqrt{\omega_n}, (\log n)^2]$. □

3.3.4 The distance to the giant

This subsection proves part (c) of Theorem 3.

Lemma 10. *For all $\varepsilon > 0$,*

$$\sup_{\mathcal{V}_n \subseteq [n] : |\mathcal{V}_n| \in \mathcal{I}_n} \mathbb{P}\left\{\left|\frac{\max_{v \in \mathcal{V}_n^c} W_v^*}{\log_k \log n} - 1\right| > \varepsilon\right\} = o(1),$$

where $W_v^* \equiv \min_{u \in \mathcal{V}_n} \operatorname{dist}(v, u)$, i.e., W_v^* is the length of the shortest path from v to $\mathcal{V}_n$.

Let $v \in \mathcal{V}_n^c$ be a vertex. If $W_v^* > 1$, then all neighbors of v are in $\mathcal{V}_n^c$, and most likely there are k of them. So $\mathbb{P}\{W_v^* > 1\} \approx (|\mathcal{V}_n^c|/n)^k \approx e^{-\tau_k k}$. If $W_v^* > 2$, then the neighbors of v's neighbors are all in $\mathcal{V}_n^c$, and most likely there are k^2 of them. So $\mathbb{P}\{W_v^* > 2\} \approx (|\mathcal{V}_n^c|/n)^{k+k^2} \approx e^{-\tau_k(k+k^2)}$. Repeating this argument shows that $\mathbb{P}\{W_v^* > x\} \approx \exp\{-\tau_k(k + k^2 \dots k^x)\} = e^{-\tau_k \Theta(k^x)}$, which is $o(1/n)$ when $x \ge (1 + \varepsilon) \log_k \log n$.

To make the above intuition rigorous, the colouring process defined in the previous subsection needs to be slightly modified. Let v be the vertex where the process has started. When choosing a red arc to colour, instead of choosing one arbitrarily from all red arcs, choose one arbitrarily from those that are closest to v. Thus at the end, the yellow arcs consist of not just a spanning tree but a breadth-first-search (bfs) spanning tree of $\mathcal{D}_{n,k}[\mathcal{S}_v^*]$. If $\mathcal{V}_n$ (the set of green vertices) is contracted into a single green vertex, then the green arcs together with yellow arcs form a DAG. Let $\mathcal{T}_v$ denote this DAG. Then W_v^* is the length of the shortest path from v to the green vertex contracted from $\mathcal{V}_n$. Figure 4 shows an example of $\mathcal{T}_v$.

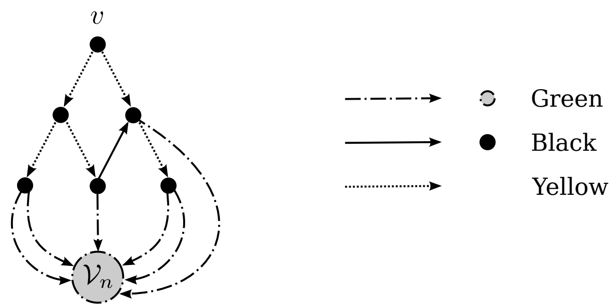

Figure 4: An example of $\mathcal{T}_v$.

Proof. Let $\omega_n = \lfloor (1+\varepsilon)\log_k \log n \rfloor$. Call the arcs whose endpoints are at distance i to v the i-th layer of $\mathcal{T}_v$. The event $W_v^* > \omega_n$ implies that the first ω_n layers of arcs in $\mathcal{T}_v$ are all yellow arcs and thus they form a tree of height ω_n. By Lemma 7, whp there are no $v \in \mathcal{V}_n^c$ such that $\mathcal{D}_{n,k}[\mathcal{S}_v^*]$ contains more than one black arc. Thus whp in every $\mathcal{T}_v$ all internal (non-leaf) vertices except at most one have out degree k. Let A_n denote this event. Assuming A_n happens, $W_v^* > \omega_n$ implies that there are at least $\Theta(k^{\omega_n}) = \Theta(\log n)^{1+\varepsilon}$ yellow arcs in the first ω_n layers of $\mathcal{T}_v$. Thus in the colouring process, the first $\Theta(\log n)^{1+\varepsilon}$ arcs choose their endpoints in $\mathcal{V}_n^c$. The probability that this happens is at most $(|\mathcal{V}_n^c|/n)^{\Theta(\log n)^{1+\varepsilon}}$. Since $|\mathcal{V}_n| \in \mathcal{I}_n$, $|\mathcal{V}_n^c| = n - |\mathcal{V}_n| \le e^{-\tau_k} n + n^{1/2+\delta}$. Then by the union bound,

$$
\begin{aligned}
\mathbb{P}\left\{\cup_{v\in\mathcal{V}_n^c}[W_v^* > \omega_n]\right\} &\le \sum_{v\in\mathcal{V}_n^c} \mathbb{P}\left\{[W_v^* > \omega_n]\cap A_n\right\} + \mathbb{P}\left\{A_n^c\right\} \\
&\le n(|\mathcal{V}_n^c|/n)^{\Theta(\log n)^{1+\varepsilon}} + o(1) \\
&\le n(e^{-\tau_k} + n^{-1/2+\delta})^{\Theta(\log n)^{1+\varepsilon}} + o(1) = o(1).
\end{aligned}
$$

Thus whp $\max_{v\in\mathcal{V}_n^c} W_v^* \le \omega_n$.

Let $\psi_n = \lceil (1-\varepsilon)\log_k \log n \rceil$. To show that whp there is a vertex v with $W_v^* \ge \psi_n$, run the colouring process starting from an arbitrary yellow vertex v until either an arc is colored black or green (failure), or the first $\psi_n - 1$ layers of $\mathcal{T}_v$ are colored yellow (success). So to succeed, the first $\psi_n - 1$ layers of $\mathcal{T}_v$ form a full k-ary tree, i.e., the first $k + k^2 + \cdots + k^{\psi_n - 1} = \Theta(k^{\psi_n}) = \Theta(\log n)^{1-\varepsilon}$ arcs must be colored yellow. If the process fails, we pick another yellow vertex and try again until one trial succeeds. Since the colouring process stops before colouring the ψ_n layer of $\mathcal{T}_v$, each trial colors at most $\Theta(k^{\psi_n}) = \Theta(\log n)^{1-\varepsilon}$ vertices black. If the process is tried at most $\lceil n/(\log n)^2 \rceil$ times, then at most $b_n \equiv \lceil n/(\log n)^2 \rceil O(\log n)^{1-\varepsilon} = O(n/(\log n)^{1+\varepsilon})$ vertices are colored black. Therefore, each arc has probability at least $(|\mathcal{V}_n^c| - b_n)/n$ to be colored yellow during the first $\lceil n/(\log n)^2 \rceil$ trials. Since $|\mathcal{V}_n| \in \mathcal{I}_n$, $|\mathcal{V}_n^c| = n - |\mathcal{V}_n| \ge e^{-\tau_k} n - n^{1/2+\delta}$. Thus the probability to succeed in one trial is at least

$$
\left(\frac{|\mathcal{V}_n^c| - b_n}{n}\right)^{O(\log n)^{1-\varepsilon}} \ge \left[e^{-\tau_k} - O\left(\frac{1}{(\log n)^{1+\varepsilon}}\right)\right]^{O(\log n)^{1-\varepsilon}} = e^{-O(\log n)^{1-\varepsilon}}.
$$

Therefore, the probability that the first $\lceil n/(\log n)^2 \rceil$ trials fail is at most

$$\left(1 - e^{-O(\log n)^{1-\varepsilon}}\right)^{\lceil n/(\log n)^2 \rceil} \le \exp\left\{-e^{-O(\log n)^{1-\varepsilon}} \frac{n}{(\log n)^2}\right\} = o(1).$$

Thus whp $\max_{v \in \mathcal{V}_n^c} W_v^* \ge \psi_n$. □

3.3.5 The longest path outside the giant

This subsection proves (d) and (e) of Theorem 3.

Lemma 11. *For all $\varepsilon > 0$, we have:*

$$\sup_{\mathcal{V}_n \subseteq [n] : |\mathcal{V}_n| \in \mathcal{I}_n} \mathbb{P}\left\{\left|\frac{m(\mathcal{V}_n^c)}{\log n} - \frac{1}{\log(e^{\tau_k}/k)}\right| > \varepsilon\right\} = o(1),$$

where $m(\mathcal{V}_n^c)$ denotes the length of the longest path in $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$; and

$$\sup_{\mathcal{V}_n \subseteq [n] : |\mathcal{V}_n| \in \mathcal{I}_n} \mathbb{P}\left\{\left|\frac{d(\mathcal{V}_n^c)}{\log n} - \frac{1}{\log(e^{\tau_k}/k)}\right| > \varepsilon\right\} = o(1).$$

where $d(\mathcal{V}_n^c)$ denotes the maximal distance between two connected vertices in $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$.

Since $m(\mathcal{V}_n^c) \ge d(\mathcal{V}_n^c)$, it suffices to prove the upper bound for $m(\mathcal{V}_n^c)$ and the lower bound for $d(\mathcal{V}_n^c)$.

Proof of the upper bound. Let $\omega_n = (1+\varepsilon)\log n/\log(e^{\tau_k}/k)$. Let X_ℓ be the number of labeled paths of length ℓ in $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$. There are less than $|\mathcal{V}_n^c|^{\ell+1}k^\ell$ possible such paths. Each of them exists with probability $(1/n)^\ell$. Recall that $|\mathcal{V}_n| \in \mathcal{I}_n$ implies $|\mathcal{V}_n^c| \le e^{-\tau_k}n + n^{1/2+\delta}$. Thus

$$\mathbb{E}X_\ell \le |\mathcal{V}_n^c|^{\ell+1}k^\ell \left(\frac{1}{n}\right)^\ell \le \left(e^{-\tau_k}n + n^{1/2+\delta}\right)\left(ke^{-\tau_k} + kn^{-1/2+\delta}\right)^\ell.$$

Since $ke^{-\tau_k} < 1$ (Lemma A1), for n large enough,

$$\sum_{\omega_n < \ell < |\mathcal{V}_n^c|} \mathbb{E}X_\ell \le n \sum_{\omega_n < \ell} (ke^{-\tau_k} + kn^{-1/2+\delta})^\ell = O\left(n\left(ke^{-\tau_k}\right)^{\omega_n}\right) = O\left(n^{-\varepsilon}\right).$$

Thus $\mathbb{P}\left\{m(\mathcal{V}_n^c) > \omega_n\right\} = O\left(n^{-\varepsilon}\right)$. □

Proof of the lower bound. Let $\psi_n \equiv \lceil (1-\varepsilon)\log n/\log(1/ke^{-\tau_k}) \rceil$. To show there are two vertices at distance within $[\psi_n, \infty)$, pick an arbitrary yellow vertex v and run the colouring process until either a vertex at distance ψ_n from v has been colored (success), or $\lceil (\log n)^2 \rceil$ vertices have been colored (failure), or the process terminates because all vertices that are reachable from v in $\mathcal{D}_{n,k}[\mathcal{V}_n^c]$ has been discovered (failure). If the process fails, we pick another yellow vertex and try again until one trial succeeds.

If at most $t_n \equiv \lfloor n/(\log n)^4 \rfloor$ trials are made, then at most $\lceil (\log n)^2 \rceil t_n = O\left(n/(\log n)^2\right)$ vertices are colored. So in the first t_n trials, when an arc is colored, the probability that it is colored yellow is at least $\mu_n \equiv (|\mathcal{V}_n^c| - O\left(n/(\log n)^2\right))/n = e^{-\tau_k} - O\left(1/(\log n)^2\right)$. Let $(Z_m)_{m\geq 0}$ be a Galton-Watson process with offspring distribution $\mathrm{Bin}(k, \mu_n)$ and $Z_0 = 1$. In other words, $Z_{m+1} = \sum_{j=1}^{Z_m} X_{m,j}$, where $(X_{m,j})_{m\geq 0, j\geq 1}$ are i.i.d. $\mathrm{Bin}(k, \mu_n)$. Then the probability that one trial succeeds is at least $\mathbb{P}\{Z_{\psi_n} > 0\}$ minus the probability that in a trial $\lceil (\log n)^2 \rceil$ vertices have been colored, which is $O\left(n^{-1-\varepsilon}\right)$ by (3) in Lemma 8.

Let $\varphi_m(y) = \mathbb{E} y^{Z_m}$, i.e., $\varphi_m(y)$ is the probability generating function of Z_m. Thus $\mathbb{P}\{Z_m = 0\} = \varphi_m(0)$. Since $ke^{-\tau_k} < 1/2$ (Lemma A1), for n large enough $k\mu_n < 1/2$. So we can apply Lemma A7 in the appendix to show that

$$\varphi_m(0) \leq 1 - (k\mu_n)^m + \left(1 - \frac{1}{2^m}\right)(k\mu_n)^{m+1} < 1 - \frac{1}{2}(k\mu_n)^m, \quad \text{for all } m \geq 0.$$

Recalling that $\psi_n \equiv \lceil (1-\varepsilon)\log n / \log(1/ke^{-\tau_k}) \rceil$,

$$\mathbb{P}\{Z_{\psi_n} > 0\} = 1 - \varphi_{\psi_n}(0) > \frac{1}{2}\left(ke^{-\tau_k} - O\left(\frac{1}{(\log n)^2}\right)\right)^{\psi_n} = \Omega\big(n^{-1+\varepsilon}\big).$$

So the probability that one trial succeeds is $\Omega(n^{-1+\varepsilon}) - O\left(n^{-1-\varepsilon}\right) = \Omega(n^{-1+\varepsilon})$. (The $O\left(n^{-1-\varepsilon}\right)$ term is the probability that one trial colors too many vertices.) Thus the probability that the first $t_n \equiv \lfloor n/(\log n)^4 \rfloor$ trials fail is at most

$$\big(1 - \Omega\big(n^{-1+\varepsilon}\big)\big)^{t_n} \leq \exp\left\{-\Omega\left(\frac{1}{n^{1-\varepsilon}}\left\lfloor \frac{n}{(\log n)^4} \right\rfloor\right)\right\} = \exp\left\{-\Omega\left(\frac{n^{\varepsilon}}{(\log n)^4}\right)\right\} = o(1).$$

Therefore whp $d(\mathcal{V}_n^c) \geq \psi_n$. □

4 Phase transition in strong connectivity

Now instead of assuming that k is fixed, let $k \to \infty$ as $n \to \infty$. Let K be a fixed integer. We can construct $\mathcal{D}_{n,k}$ by first generating $\mathcal{D}_{n,K}$ and then adding arcs with labels in $\{K+1, \ldots, k\}$ into it. By Lemma 2, for all $\varepsilon > 0$, there exists a K depending only on ε such that whp in $\mathcal{D}_{n,K}$ the largest closed SCC has size at least $(1-\varepsilon)n$ and is reachable from all vertices. Since adding arcs can only increase the size of this SCC, whp $\mathcal{D}_{n,k}$ has a SCC of size at least $(1-\varepsilon)n$ that is reachable from all vertices.

In fact, if k increases fast enough, then whp $\mathcal{D}_{n,k}$ is strongly connected. More precisely, $\mathcal{D}_{n,k}$ exhibits a phase transition for strong connectivity similar to the analogous event in the Erdős–Rényi model [15].

Theorem 6. *If $k - \log n \to -\infty$, then whp $\mathcal{D}_{n,k}$ is not strongly connected. If $k - \log n \to \infty$, then whp $\mathcal{D}_{n,k}$ is strongly connected.*

If there is a vertex with in-degree zero, then obviously the digraph is not strongly connected. Thus the following lemma proves the lower bound in Theorem 6.

Lemma 12. *If $k - \log n \to -\infty$, whp $\mathcal{D}_{n,k}$ contains a vertex of in-degree zero.*

Given a set of vertices $\mathcal{S}$, if there are no arcs that start from $\mathcal{S}^c \equiv [n] \setminus \mathcal{S}$ and end at $\mathcal{S}$, then call $\mathcal{S}$ a *non-leaf.* If $\mathcal{D}_{n,k}$ is not strongly connected, then there must exist a non-leaf set of vertices $\mathcal{S}$ with $|\mathcal{S}| < n$. Thus the following lemma implies the upper bound in Theorem 6.

Lemma 13. *If $k - \log n \to +\infty$, whp there does not exist a non-leaf set of vertices $\mathcal{S}$ with $|\mathcal{S}| < n$.*

We omit the proofs of the above two lemmas as they use standard first and second moment methods.

5 The simple digraph model, the number of self-loops and multiple arcs

A *simple* digraph is one in which there are no self-loops and there is no more than one arc from one vertex to another. Let $\mathcal{D}^*_{n,k}$ denote a simple k-out digraph with n vertices chosen uniformly at random from all such digraphs. $\mathcal{D}^*_{n,k}$ can be viewed as $\mathcal{D}_{n,k}$ restricted to the event that $\mathcal{D}_{n,k}$ is simple. This section proves the following theorem:

Theorem 7. *The probability that $\mathcal{D}_{n,k}$ is simple converges to $e^{-k-\binom{k}{2}}$ as $n \to \infty$.*

Theorem 7 can be proved directly as follows. Let $\mathbb{1}_v$ be the indicator that the k arcs starting from vertex v do not end at v and do not end at the same vertex. Then

$$\mathbb{P}\left\{\mathbb{1}_v = 1\right\} = \frac{(n-1)(n-2)\cdots(n-k)}{n^k} = 1 - \frac{k(k+1)}{2n} + O\left(\frac{1}{n^2}\right).$$

Since $\mathcal{D}_{n,k}$ is simple if and only if $\cap_{v=1}^n [\mathbb{1}_v = 1]$ happens, we have

$$\begin{aligned}\mathbb{P}\left\{\mathcal{D}_{n,k} \text{ is simple}\right\} &= \mathbb{P}\left\{\cap_{v=1}^n \left[\mathbb{1}_v = 1\right]\right\} = \prod_{v=1}^n \mathbb{P}\left\{\mathbb{1}_v = 1\right\} \\ &= \left(1 - \frac{k(k+1)}{2n} + O\left(\frac{1}{n^2}\right)\right)^n \to e^{-k(k+1)/2} = e^{-k-\binom{k}{2}}.\end{aligned}$$

However, we can say more about self-loops and multiple arcs between vertices. Let $\mathcal{I} \equiv [n] \times [k]$. For $(v,i) \in \mathcal{I}$, define the random variable $\mathbb{1}_{v,i}$ to be the indicator that the arc with label i starting from vertex v forms a self-loop. Let $\mathcal{J} \equiv \{(v,i,j) \in [n] \times [k] \times [k] : i < j\}$. For $(v,i,j) \in \mathcal{J}$, define the random variable $\mathbb{1}_{v,i,j}$ to be the indicator that the two arcs starting from vertex v with labels i and j both end at the same vertex. Let $S_n = \sum_{\alpha \in \mathcal{I}} \mathbb{1}_\alpha$ and $M_n = \sum_{\alpha \in \mathcal{J}} \mathbb{1}_\alpha$. Then $[S_n = 0] \cap [M_n = 0]$ if and only if $\mathcal{D}_{n,k}$ is simple.

Lemma 14. *Let S and M be two independent Poisson random variables of means k and $\binom{k}{2}$ respectively. Then $(S_n, M_n) \xrightarrow{d} (S, M)$ as $n \to \infty$. In fact,*

$$\|(S_n, M_n), (S, M)\|_{\mathrm{TV}} = O\left(\frac{1}{n}\right).$$

Indeed the lemma implies that as $n \to \infty$,

$$\mathbb{P}\{\mathcal{D}_{n,k} \text{ is simple}\} = \mathbb{P}\{S_n = M_n = 0\} \to \mathbb{P}\{S = 0\}\,\mathbb{P}\{M = 0\} = e^{-k}e^{-\binom{k}{2}}.$$

Remark. Bollobás [9] proved a theorem similar to Lemma 14 for the configuration model (see also Bollobás [8, sec. 2.4]). Many authors have extended this result under various conditions, see, e.g., McKay [30], McKay and Wormald [31], Janson and Luczak [25], Janson [23]. Our proof uses Stein's method, which may also be applied to self-loops and multiple edges in the configuration model to get proofs shorter than previous ones.

Proof of Lemma 14. We use the Chen-Stein method [11]. Since the probability that an arc forms a self-loop is $1/n$,

$$\mathbb{E}S_n = \sum_{(v,i)\in\mathcal{I}} \mathbb{E}\mathbb{1}_{v,i} = kn\frac{1}{n} = k.$$

Thus $\mathbb{E}S = k = \mathbb{E}S_n$. Since the probability that two arcs with the same start point have the same endpoint is also $1/n$,

$$\mathbb{E}M_n = \sum_{v\in[n]} \sum_{1\le i<j\le k} \mathbb{E}\mathbb{1}_{v,i,j} = n\binom{k}{2}\frac{1}{n} = \frac{k(k-1)}{2}.$$

Thus $\mathbb{E}M = k(k-1)/2 = \mathbb{E}M_n$.

For $\alpha \in \mathcal{I} \cup \mathcal{J}$, let

$$\mathcal{B}_\alpha = \{\beta \in \mathcal{I} \cup \mathcal{J} : \mathbb{1}_\beta \text{ and } \mathbb{1}_\alpha \text{ are dependent}\}.$$

(Note that $\mathbb{1}_\alpha \in \mathcal{B}_\alpha$.) Define

$$b_1 \equiv \sum_{\alpha\in\mathcal{I}\cup\mathcal{J}} \sum_{\beta\in\mathcal{B}_\alpha} \mathbb{E}[\mathbb{1}_\alpha]\,\mathbb{E}[\mathbb{1}_\beta], \quad b_2 \equiv \sum_{\alpha\in\mathcal{I}\cup\mathcal{J}} \sum_{\beta\in\mathcal{B}_\alpha:\alpha\ne\beta} \mathbb{E}[\mathbb{1}_\alpha\mathbb{1}_\beta], \quad b_3 \equiv \sum_{\alpha\in\mathcal{I}\cup\mathcal{J}} s_\alpha,$$

where

$$s_\alpha = \mathbb{E}\left|\mathbb{E}[\mathbb{1}_\alpha \,|\, \sigma(\mathbb{1}_\beta : \beta \in [\mathcal{I}\cup\mathcal{J}] \setminus \mathcal{B}_\alpha)] - \mathbb{E}\mathbb{1}_\alpha\right|.$$

By [11, thm. 2], if $b_1 + b_2 + b_3 \to 0$, then $(S_n, M_n) \xrightarrow{d} (S, M)$. Since $\mathbb{1}_\alpha$ is independent of the random variables $\mathbb{1}_\beta$ with $\beta \in [\mathcal{I}\cup\mathcal{J}] \setminus \mathcal{B}_\alpha$, we have $s_\alpha = 0$ and thus $b_3 = 0$.

For $(v, i) \in \mathcal{I}$, $\mathbb{1}_{v,i}$ depends on the random variables $\mathbb{1}_{v,r,s}$ with $1 \le r < s \le k$ and $i \in \{r, s\}$, of which there are $k-1$. Thus $|\mathcal{B}_{v,i}| = 1 + (k-1) = k < 2k$. For $(v, i, j) \in \mathcal{J}$, $\mathbb{1}_{v,i,j}$ depends on $\mathbb{1}_{v,i}$ and $\mathbb{1}_{v,j}$. It also depends on the random variables

$\mathbb{1}_{v,r,s}$ with $1 \le r < s \le k$ and $\{r,s\} \cap \{i,j\} \ne \emptyset$, of which there are $2(k-1)-1 = 2k-3$. Thus $|\mathcal{B}_{v,i,j}| = 2 + 2k - 3 < 2k$. So for all $\alpha \in \mathcal{I} \cup \mathcal{J}$, $|\mathcal{B}_\alpha| < 2k$. Therefore

$$\begin{aligned} b_1 &= \sum_{\alpha \in \mathcal{I}} \sum_{\beta \in \mathcal{B}_\alpha} \mathbb{E}[\mathbb{1}_\alpha] \mathbb{E}[\mathbb{1}_\beta] + \sum_{\alpha \in \mathcal{J}} \sum_{\beta \in \mathcal{B}_\alpha} \mathbb{E}[\mathbb{1}_\alpha] \mathbb{E}[\mathbb{1}_\beta] \\ &< nk \times 2k \times \frac{1}{n} \times \frac{1}{n} + n\binom{k}{2} \times 2k \times \frac{1}{n} \times \frac{1}{n} = O\left(\frac{1}{n}\right). \end{aligned}$$

Consider $(v,i) \in \mathcal{I}$. If $\beta \in \mathcal{B}_{v,i} \cap \mathcal{I}$, then $\beta = (v,i)$. If $\beta \in \mathcal{B}_{v,i} \cap \mathcal{J}$, then $\beta = (v,r,s)$ for some (r,s) with $i \in \{r,s\}$. Then $\mathbb{1}_{v,i}\mathbb{1}_{v,r,s} = 1$ if and only if the two arcs starting from vertex v labeled r and s respectively both end at v. Thus $\mathbb{E}[\mathbb{1}_{v,i}\mathbb{1}_{v,r,s}] = 1/n^2$. Therefore

$$b_{2,\mathcal{I}} \equiv \sum_{\alpha \in \mathcal{I}} \sum_{\beta \in \mathcal{B}_\alpha : \beta \ne \alpha} \mathbb{E}[\mathbb{1}_\alpha \mathbb{1}_\beta] = \sum_{\alpha \in \mathcal{I}} \sum_{\beta \in \mathcal{B}_\alpha \cap \mathcal{J}} \mathbb{E}[\mathbb{1}_\alpha \mathbb{1}_\beta] < nk \times 2k \times \frac{1}{n^2} = O\left(\frac{1}{n}\right).$$

Consider $(v,r,s) \in \mathcal{J}$. If $(v,i) \in \mathcal{B}_{v,r,s}$, then $(v,r,s) \in \mathcal{B}_{v,i}$. Thus by the above argument $\mathbb{E}[\mathbb{1}_{v,r,s}\mathbb{1}_{v,i}] = 1/n^2$. If $(v,i,j) \in \mathcal{B}_{v,r,s}$ and $(i,j) \ne (r,s)$, then $|\{r,s\} \cup \{i,j\}| = 3$. So $\mathbb{1}_{v,r,s}\mathbb{1}_{v,i,j} = 1$ iff the three arcs starting from vertex v with labels in $\{r,s\} \cup \{i,j\}$ all end at the same vertex. Thus $\mathbb{E}[\mathbb{1}_{v,r,s}\mathbb{1}_{v,i,j}] = 1/n^2$. Therefore

$$b_{2,\mathcal{J}} \equiv \sum_{\alpha \in \mathcal{J}} \sum_{\beta \in \mathcal{B}_\alpha : \beta \ne \alpha} \mathbb{E}[\mathbb{1}_\alpha \mathbb{1}_\beta] < n\binom{k}{2} \times 2k \times \frac{1}{n^2} = O\left(\frac{1}{n}\right).$$

Thus $b_2 \equiv b_{2,\mathcal{I}} + b_{2,\mathcal{J}} = O(1/n)$. □

Corollary 2. *Let $\mathcal{E}$ be a set of digraphs. If $\mathcal{D}_{n,k} \in \mathcal{E}$ whp, then $\mathcal{D}^*_{n,k} \in \mathcal{E}$ whp.*

Proof. We have

$$\mathbb{P}\left\{\mathcal{D}^*_{n,k} \notin \mathcal{E}\right\} = \mathbb{P}\{\mathcal{D}_{n,k} \notin \mathcal{E} \,|\, \mathcal{D}_{n,k} \text{ is simple}\} \le \frac{\mathbb{P}\{\mathcal{D}_{n,k} \notin \mathcal{E}_n\}}{\mathbb{P}\{\mathcal{D}_{n,k} \text{ is simple}\}} \to 0. \qquad \square$$

This corollary implies that all previous results in the form of "whp $\mathcal{D}_{n,k} \ldots$" can be automatic translated into "whp $\mathcal{D}^*_{n,k} \ldots$". For example, the statement of Theorem 3 with $\mathcal{D}_{n,k}$ replaced by $\mathcal{D}^*_{n,k}$ is still true.

Corollary 3. *Let $\mathcal{D}^{**}_{n,k}$ be a digraph chosen uniformly at random from all simple and arc-unlabeled k-out digraphs with n vertices. If whp $\mathcal{D}_{n,k}$ has property $\boldsymbol{P}$ where $\boldsymbol{P}$ does not depend on arc-labels, then whp $\mathcal{D}^{**}_{n,k}$ has property $\boldsymbol{P}$.*

Proof. Note that: (a) for each digraph in the space of $\mathcal{D}^{**}_{n,k}$, there $(k!)^n$ ways to arc-label it to get $(k!)^n$ different digraphs in the space of $\mathcal{D}^*_{n,k}$; (b) no two different arc-unlabeled digraphs can be turned into the same digraph by arc-labeling. So there exists a $(k!)^n$-to-one surjective mapping from the space of $\mathcal{D}^*_{n,k}$ to the space of $\mathcal{D}^{**}_{n,k}$. Thus $\mathcal{D}^{**}_{n,k}$ can be viewed as $\mathcal{D}^*_{n,k}$ with arc labels removed. Since **P** does not depend on arc-labels, it follows from Corollary 2 that whp $\mathcal{D}^{**}_{n,k}$ has property **P**. □

6 Extensions

The typical distance H_n of $\mathcal{D}_{n,k}$ is the distance between two vertices v_1 and v_2 chosen uniformly at random. If v_1 cannot reach v_2, then $H_n = \infty$. Addario-Berry et al. [1] proved that conditioned on $H_n < \infty$, $H_n/\log_k n \xrightarrow{p} 1$. We have found an alternative proof for this result using the path counting technique invented by van der Hofstad [40, chap. 3.5]. The proof can be found in a longer version of this paper at [`http://arxiv.org/abs/1504.06238`].

Addario-Berry et al. [1] also proved that the diameter of the giant component divided by $\log n$ converges in probability to $1/\log(k)+1/\log(1/\lambda_k)$. Recall that the longest path outside the giant divided by $\log n$ converges in probability to $1/\log(1/\lambda_k)$. This seems to be a strong indication that it might be possible to derive a new proof for the diameter of the giant.

Recall that $\mathcal{D}^*_{n,k}$ is a simple k-out digraph with n vertices chosen uniformly at random from all such digraphs. Section 5 proved that if whp $\mathcal{D}_{n,k}$ has property $\mathbf{P}$, then whp $\mathcal{D}^*_{n,k}$ has property $\mathbf{P}$. But results like Theorem 1, the central limit law of the one-in-core, cannot be transferred to $\mathcal{D}^*_{n,k}$ automatically. We believe that it might be possible to achieve get the same result for $\mathcal{D}^*_{n,k}$ following the line of Janson and Luczak's treatment of the configuration model [24].

A natural generalization of $\mathcal{D}_{n,k}$ is to have a deterministic out-degree sequence, as in the directed configuration model, instead of requiring each vertex to have out-degree exactly k. With some constraints on the out-degree sequence, most of our results should hold for this generalized model. Furthermore, we could let each vertex choose its out-degree independently at random from an out-degree distribution. Again by adding some restrictions on the out-degree distribution, most of our results should still hold.

The problem of generating a uniform random surjective function with fixed domain size is an open problem. Theorem 1 implies a simple algorithm for choosing a $[km] \to [m]$ surjective function uniformly at random. Let $n = \lceil m/\nu_k \rceil$. Then we generate a $\mathcal{D}_{n,k}$. If $|\mathcal{O}_n| = m$, i.e., if the one-in-core in $\mathcal{D}_{n,k}$ contains m vertices, then $\mathcal{D}_{n,k}[\mathcal{O}_n]$ is equivalent to a uniform random sample of a $[km] \to [m]$ surjective function. Otherwise we try again until $|\mathcal{O}_n| = m$. Theorem 1 shows that $\mathbb{P}\{|\mathcal{O}_n| = m\} = \Theta(1/\sqrt{m})$. Thus the expected number of $\mathcal{D}_{n,k}$ needed to be generated is $\Theta(\sqrt{m})$. Since generating a $\mathcal{D}_{n,k}$ takes $\Theta(m)$ time, the expected running time of the whole algorithm is $\Theta(m^{3/2})$. But we believe that $\Theta(m)$ should be achievable.

Acknowledgment

The authors thank Laura Eslava, Hamed Hatami, Guillem Perarnau, Bruce Reed, Henning Sulzbach and Yelena Yuditsky for valuable comments on this work, and Denis Thérien for pointing out the importance of the model.

References

[1] L. Addario-Berry, B. Balle, and G. Perarnau. On the diameter and the stationary distribution of random r-out digraphs. 2015.

[2] D. J. Aldous and J. Pitman. Brownian bridge asymptotics for random mappings. *Random Structures and Algorithms*, 5(4):487–512, 1994.

[3] D. Angluin and D. Chen. Random walks on random uni-regular graphs. 2015.

[4] R. Arratia, L. Goldstein, and L. Gordon. Two moments suffice for poisson approximations: The chen-stein method. *The Annals of Probability*, 17(1):9–25, 1989.

[5] J. Bang-Jensen and G. Z. Gutin. *Digraphs: Theory, Algorithms and Applications*. Springer Monographs in Mathematics. Springer, 2009.

[6] A. D. Barbour, L. Holst, and S. Janson. *Poisson Approximation*. Oxford University Press, 1993.

[7] N. L. Biggs, E. K. Lloyd, and R. J. Wilson. *Graph Theory 1736–1936*. Oxford University Press, 1976.

[8] B. Bollobás. *Random Graphs*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2nd edition, 2001.

[9] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311–316, 1980.

[10] A. Carayol and C. Nicaud. Distribution of the number of accessible states in a random deterministic automaton. In *29th International Symposium on Theoretical Aspects of Computer Science*, volume 14 of *Leibniz International Proceedings in Informatics*, 194–205. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2012.

[11] L. H. Y. Chen. Poisson approximation for dependent trials. *The Annals of Probability*, 3(3):534–545, 1975.

[12] C. Cooper and A. Frieze. The size of the largest strongly connected component of a random digraph with a given degree sequence. *Combinatorics, Probability and Computing*, 13:319–337, 2004.

[13] R. Durrett. *Probability: Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 4th edition, 2010.

[14] R. Durrett. *Random Graph Dynamics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2007.

[15] P. Erdős and A. Rényi. On random graphs i. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959.

[16] P. Erdős and A. Rényi. On the evolution of random graphs. *Publication of the Mathematical Institute of the Hungarian Academy of Sciences*, 5:17–61, 1960.

[17] P. Flajolet and R. Sedgewick. *Analytic Combinatorics.* Cambridge University Press, 2009.

[18] P. Flajolet and A. Odlyzko. Random mapping statistics. In *Advances in Cryptology–EUROCRYPT '89*, volume 434 of *Lecture Notes in Computer Science*, 329–354. Springer, 1990.

[19] I. J. Good. An asymptotic formula for the differences of the powers at zero. *The Annals of Mathematical Statistics*, 32(1):249–256, 1961.

[20] A. A. Grusho. Limit distributions of certain characteristics of random automaton graphs. *Mathematical Notes of the Academy of Sciences of the USSR*, 14(1):633–637, 1973.

[21] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):1330, 1963.

[22] S. Janson, T. Łuczak, and A. Rucinski. *Random Graphs.* Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, 2011.

[23] S. Janson. The probability that a random multigraph is simple. II. *Journal of Applied Probability*, 51A:123–137, 12 2014.

[24] S. Janson and M. J. Luczak. Asymptotic normality of the k-core in random graphs. *The Annals of Applied Probability*, 18(3):1085–1137, 2008.

[25] S. Janson and M. J. Luczak. A new approach to the giant component problem. *Random Structures and Algorithms*, 34(2):197–216, 2009.

[26] M. Kearns. Efficient noise–tolerant learning from statistical queries. *Journal of the ACM*, 45(6):983–1006, 1998.

[27] V. F. Kolchin. *Random Mappings.* Translations series in mathematics and engineering. Optimization Software, Inc., Publications Division, 1986.

[28] A. Korshunov. The number of initially connected automata. *Problemy Kibernetiki*, 34:5–82, 1978.

[29] E. Lebensztayn. On the asymptotic enumeration of accessible automata. *Discrete Mathematics and Theoretical Computer Science*, 12(3):75–79, 2010.

[30] B. D. McKay. Asymptotics for symmetric 0-1 matrices with prescribed row sums. *Ars Combinatoria*, 19:15–25, 1985.

[31] B. D. McKay and N. C. Wormald. Asymptotic enumeration by degree sequence of graphs with degrees $o(n^{1/2})$. *Combinatorica*, 11(4):369–382, 1991.

[32] B. Pittel. Counting strongly-connected, moderately sparse directed graphs. *Random Structures and Algorithms*, 43(1):49–79, 2013.

[33] B. Pittel and D. Poole. Asymptotic distribution of the numbers of vertices and arcs of the giant strong component in sparse random digraphs. *Random Structures and Algorithms*, 2015. To appear.

[34] X. Pérez-Giménez and N. Wormald. Asymptotic enumeration of strongly connected digraphs by vertices and edges. *Random Structures and Algorithms*, 43(1):80–114, 2013.

[35] B. A. Sevast'yanov. Convergence to gaussian and poisson processes of the distribution of the number of empty cells in the classical problem of pellets. *Teoriya Veroyatnostei i ee Primeneniya*, 12(1):144–153, 1967.

[36] M. Shaked and J. Shanthikumar. *Stochastic Orders*. Springer Series in Statistics. Springer, 2007.

[37] M. Sipser. *Introduction to the Theory of Computation*. Cengage Learning, 2012.

[38] C. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the 6th Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Probability Theory*, 583–602. University of California Press, 1972.

[39] R. van der Hofstad. *Random Graphs and Complex Networks*, volume 1. 2014. URL `http://www.win.tue.nl/~rhofstad/NotesRGCN.pdf`.

[40] R. van der Hofstad. *Random Graphs and Complex Networks*, volume 2. 2014. URL `http://www.win.tue.nl/~rhofstad/NotesRGCNII.pdf`.

[41] R. van der Hofstad and M. Keane. An elementary proof of the hitting time theorem. *The American Mathematical Monthly*, 115(8):753–756, 2008.

Appendix

1. Inequalities for constants

Lemma A1. *Assume that* $k \geq 2$.

(a) There exists exactly one $\tau_k > 0$ *such that* $1 - \tau_k/k - e^{-\tau_k} = 0$*;*

(b) $0 < k - \tau_k < 1/2$*;*

(c) $1/2 < 1 - \frac{1}{2k} < \nu_k \equiv \tau_k/k < 1$*;*

(d) $\lambda_k \equiv (k - \tau_k) \left(\frac{\tau_k}{k-1}\right)^{k-1} < \lambda'_k \equiv (k - \tau_k) e^{1-k+\tau_k} < 1$*;*

(e) $\gamma_k \equiv \left(\frac{k}{e\tau_k}\right)^k (e^{\tau_k} - 1) < 1$*;*

(f) $\rho_k \equiv k e^{1-\tau_k} \left(\frac{\tau_k}{k}\right)^{k-1} < 1$*;*

(g) $\lambda_k = \Theta(ke^{-k})$ *as* $k \to \infty$*.*

Proof. Let $\eta(x) = 1 - x/k - e^{-x}$. Since $\eta''(x) = -e^{-x} < 0$, $\eta(x)$ is strictly concave. Since $\eta(k - 1/2) > 0$, and $\eta(k) < 0$, $\eta(x) = 0$ must have exactly one positive solution and this solution must be in $(k - 1/2, k)$. Thus (a) and (b) are proved. (c) follows since $\tau_k/k > 1 - 1/k \geq 1/2$. For (d) note that $\lambda_k < {\lambda_k}'$ as $1 - x < e^{-x}$ for all $x \neq 0$. For $\lambda'_k < 1$ note that

$$\log {\lambda_k}' = \log(k - \tau_k) + 1 - (k - \tau_k) = \log\left[1 - (1 - (k - \tau_k))\right] + 1 - (k - \tau_k) < 0,$$

since $\log(1 - x) < -x$ for all $x \in (0, 1)$.

For (e), first use $\tau_k/k \equiv 1 - e^{-\tau_k}$ to get

$$\gamma_k = \frac{1}{e^k(1 - e^{-\tau_k})^k} e^{\tau_k}(1 - e^{-\tau_k}) = e^{\tau_k - k}(1 - e^{-\tau_k})^{1-k}.$$

Then use $ke^{-\tau_k} \equiv k - \tau_k$ to get

$$\begin{aligned}
\log \gamma_k &= \tau_k - k + (1 - k)\log(1 - e^{-\tau_k}) \\
&= (\tau_k - k) + \log(1 - e^{-\tau_k}) - k\log(1 - e^{-\tau_k}) \\
&< (\tau_k - k) - e^{-\tau_k} + k(e^{-\tau_k} + e^{-2\tau_k}) \\
&= (\tau_k - k) + (k - \tau_k) + e^{-\tau_k}(k - \tau_k - 1) < 0,
\end{aligned}$$

since $-x > \log(1 - x) > -x - x^2$ for all $x \in (0, 1/2)$ and $e^{-\tau_k} = 1 - \nu_k \in (0, 1/2)$.

For (f), use $\tau_k < k$ from (a) to get

$$\tau_k \equiv k(1 - e^{-\tau_k}) < k(1 - e^{-k}). \tag{4}$$

Therefore,

$$\frac{\tau_k}{k} \equiv 1 - e^{-\tau_k} < 1 - \exp\left\{-k\left(1 - e^{-k}\right)\right\}.$$

Again by (a), $\tau_k > k - 1/2$. Thus

$$\tau_k \equiv k(1 - e^{-\tau_k}) > k(1 - e^{-k+\frac{1}{2}}). \tag{5}$$

Therefore,

$$ke^{-\tau_k} < k\exp\left\{-k\left(1 - e^{-k+\frac{1}{2}}\right)\right\}.$$

The above bounds imply that

$$\rho_k \equiv ke^{1-\tau_k}\left(\frac{\tau_k}{k}\right)^{k-1} < k\exp\left\{1 - k\left(1 - e^{-k+\frac{1}{2}}\right)\right\}\left(1 - \exp\left\{-k\left(1 - e^{-k}\right)\right\}\right)^{k-1}.$$

Using this bound, numeric computations show that $\rho_2 < 0.945651$. When $k \geq 3$, the above upper bound is less than

$$k \exp\left\{1 - k\left(1 - e^{-\frac{5}{2}}\right)\right\},$$

which takes its maximal value at $k = 3$ for $k \in [3, \infty)$. This maximal value is about 0.52. Thus $\rho_k < 1$ for all $k \geq 2$.

By (4) and (5), $k - \tau_k = ke^{-k+O(1)}$ and $\tau_k/k = 1 - e^{-k+O(1)}$ as $k \to \infty$. Therefore

$$\begin{aligned}\lambda_k &\equiv (k - \tau_k)\left(\frac{\tau_k}{k-1}\right)^{k-1} \\ &= (k - \tau_k)\left(\frac{\tau_k}{k}\right)^{k-1}\left(\frac{k}{k-1}\right)^{k-1} \\ &= ke^{-k+O(1)}\left(1 - e^{-k+O(1)}\right)^{k-1} e(1 + o(1)) = ke^{-k+O(1)}.\end{aligned}$$

Thus (g) is proved. □

2. The sizes of k-surjections

In this section we prove Lemma 1. Recall that K_s is the number of k-surjections of size s in $\mathcal{D}_{n,k}$. We first deal the case that s is small:

Lemma A2. *$\mathbb{P}\{K_1 \geq 1\} \leq 1/n^{k-1} \leq 1/n$.*

Proof. A single vertex is a k-surjection if and only if all its k arcs are self-loops. Thus

$$\mathbb{P}\{K_1 \geq 1\} \leq \sum_{v \in [n]} \mathbb{P}\{v \text{ has only self-loops}\} = n\left(\frac{1}{n}\right)^k \leq \frac{1}{n^{k-1}} \leq \frac{1}{n}.$$ □

Lemma A3. *$\mathbb{P}\left\{\sum_{2 \leq s \leq an} K_s \geq 1\right\} = o(1/n)$, for all fixed $a \in \left(0, e^{-1/(k-1)}\right)$.*

Proof. We can choose $\varepsilon \in (0,1)$ such that $2(k-1)(1-\varepsilon) > 1$ since $k \geq 2$. Let

$J = \{2, \ldots, \lfloor an \rfloor\}$. Then

$$
\begin{aligned}
\mathbb{P}\left\{\sum_{s \in J} K_s \geq 1\right\} &\leq \sum_{s \in J} \sum_{\mathcal{S} \subseteq [n]:|\mathcal{S}|=s} \mathbb{P}\{\mathcal{S} \text{ is closed}\} \\
&= \sum_{s \in J} \binom{n}{s} \left(\frac{s}{n}\right)^{ks} \\
&\leq \sum_{s \in J} \left(\frac{en}{s}\right)^{s} \left(\frac{s}{n}\right)^{ks} \qquad \text{(Stirling's approximation)} \\
&= \sum_{2 \leq s \leq n^{\varepsilon}} \left[e \left(\frac{s}{n}\right)^{k-1}\right]^{s} + \sum_{n^{\varepsilon} < s < an} \left[e \left(\frac{s}{n}\right)^{k-1}\right]^{s} \\
&\leq \left[e \left(\frac{n^{\varepsilon}}{n}\right)^{k-1}\right]^{2} \sum_{2 \leq s+2} \left[e \left(\frac{n^{\varepsilon}}{n}\right)^{k-1}\right]^{s} + \sum_{n^{\varepsilon} < s} \left(e \times a^{k-1}\right)^{s} \\
&= O\left(n^{-2(k-1)(1-\varepsilon)}\right) + O\left(\left(ea^{k-1}\right)^{n^{\varepsilon}}\right),
\end{aligned}
$$

where both terms are $o(1/n)$ due to our choice of ε and a. □

When s is large, we need to take into account the probability that $\mathcal{S}$ is surjective. Let $\left\{ {x \atop y} \right\}$ denote Stirling's number of the second kind, i.e., the number of ways to put x balls into y unordered bins such that there are no empty bins [17, pp. 64]. Then

$$
\mathbb{P}\{\mathcal{S} \text{ is surjective} \mid \mathcal{S} \text{ is closed}\} = \frac{\left\{ {ks \atop s} \right\} s!}{s^{ks}},
$$

where the numerator is the number of ways to choose endpoints for the ks arcs in $\mathcal{S}$ so that minimum in-degree is one, and the denominator is the total number of ways to choose endpoints for ks arcs in $\mathcal{S}$. Thus

$$
\begin{aligned}
\mathbb{P}\{\mathcal{S} \text{ is a } k\text{-surjection}\} &= \mathbb{P}\{\mathcal{S} \text{ is surjective} \mid \mathcal{S} \text{ is closed}\} \, \mathbb{P}\{\mathcal{S} \text{ is closed}\} \\
&= \frac{\left\{ {ks \atop s} \right\} s!}{s^{ks}} \left(\frac{s}{n}\right)^{ks} = \frac{\left\{ {ks \atop s} \right\} s!}{n^{ks}}.
\end{aligned}
$$

Good [19] established an asymptotic estimation of Stirling's numbers of the second kind

$$
\left\{ {ks \atop s} \right\} \sim \frac{(ks)!}{s!} \frac{(e^{\tau_k} - 1)^s}{{\tau_k}^{ks} \sqrt{2\pi ks(1 - ke^{-k})}}.
$$

Applying this and Stirling's approximation for factorials, we have

$$
\begin{aligned}
\mathbb{P}\{\mathcal{S} \text{ is a } k\text{-surjection}\} &\sim \frac{(ks)!}{s!} \frac{(e^{\tau_k} - 1)^s}{{\tau_k}^{ks} \sqrt{2\pi ks(1 - ke^{-k})}} \frac{s!}{n^{ks}} \\
&\sim \frac{1}{\sqrt{1 - ke^{-\tau_k}}} \left[\left(\frac{s}{n}\right)^{k} \gamma_k\right]^{s}, \qquad (6)
\end{aligned}
$$

where $\gamma_k \equiv (k/e\tau_k)^k \left(e^{\tau_k} - 1\right) < 1$ (see Lemma A1).

Lemma A4. *There exists a constant* $b \in (\nu_k, 1)$ *such that* $\mathbb{P}\left\{\sum_{bn \le s \le n} K_s \ge 1\right\} = o(1/n)$.

Proof. Let $b > \nu_k$ be a constant decided later. If $|\mathcal{S}| = s \in [bn, n]$, then by (6)

$$\mathbb{P}\{\mathcal{S} \text{ is a } k\text{-surjection}\} = O\left(\left[\left(\frac{s}{n}\right)^k \gamma_k\right]^s\right) \le O(\gamma_k^s) \le O\left(\gamma_k^{bn}\right).$$

Since $b > \nu_k > 1/2$ (Lemma A1),

$$\binom{n}{s} \le \binom{n}{bn} = O\left(\frac{1}{\sqrt{n}}\left[\frac{1}{b^b(1-b)^{1-b}}\right]^n\right).$$

Therefore

$$\mathbb{P}\{K_s \ge 1\} \le \binom{n}{s}\mathbb{P}\{\mathcal{S} \text{ is a } k\text{-surjection}\} \le O\left(\left[\frac{\gamma_k^b}{b^b(1-b)^{1-b}}\right]^n\right).$$

Since the quantity in the square brackets goes to $\gamma_k < 1$ as $b \to 1$, we can pick a b close enough to one such that $\mathbb{P}\left\{\sum_{bn \le s \le n} K_s \ge 1\right\} = o(1/n)$. □

Let $a \in (0, \nu_k)$ and $b \in (\nu_k, 1)$ be two constants such that the upper bounds in Lemma A3 and A4 hold. If $|\mathcal{S}| = xn$ with $x \in (a, b)$ and xn integer-valued, then by (6) and Stirling's approximation

$$\begin{aligned}
\mathbb{E}K_{xn} &= \binom{n}{xn}\mathbb{P}\{\mathcal{S} \text{ is a } k\text{-surjection}\} \\
&\sim \frac{1}{\sqrt{2\pi x(1-x)n}}\left[\frac{1}{(x)^x(1-x)^{1-x}}\right]^n \frac{1}{\sqrt{1-ke^{-\tau_k}}}\left(x^k\gamma_k\right)^{xn} \\
&= \frac{1}{\sqrt{2\pi(1-ke^{-\tau_k})n}}g(x)\left[f(x)\right]^n \qquad (7)
\end{aligned}$$

where

$$g(x) \equiv \frac{1}{\sqrt{x(1-x)}}, \qquad f(x) \equiv \left[\frac{x^{k-1}\gamma_k}{(1-x)^{(1-x)/x}}\right]^x.$$

Lemma A5. *For all fixed* $a \in (0, \nu_k)$, $b \in (\nu_k, 1)$ *and* $\delta \in (0, 1/2)$, $\mathbb{P}\left\{\sum_{s \in J} K_s \ge 1\right\} = o(1/n)$, *where* $J = [an, \nu_k n - n^{\frac{1}{2}+\delta}] \cup [\nu_k n + n^{\frac{1}{2}+\delta}, bn]$.

Proof. Let $h(x) \equiv \log f(x)$. Lemma A6 shows that as $x \to \nu_k$,

$$h(x) = -\frac{(x-\nu_k)^2}{2\sigma_k^2} + O\left(|x-\nu_k|^3\right),$$

and that $h(x)$ is strictly increasing on (a, ν_k) and strictly decreasing on (ν_k, b). It follows from $|s/n - \nu_k| > n^{-1/2+\delta}$ that $h(s/n) \le -n^{2\delta-1}/2\sigma_k^2 + O\left(n^{3\delta-3/2}\right)$. As for $g(x)$, it is bounded on (a,b). Thus by (7) and Markov's inequality

$$\begin{aligned}\log(n^2\mathbb{P}\{K_s \ge 1\}) &\le \log(n^2\mathbb{E}K_s)\\ &= \log\left(n^2 O\left(n^{-1/2}\right) f\left(\frac{s}{n}\right)^n\right)\\ &= O\left(\log n\right) + nh\left(\frac{s}{n}\right)\\ &\le O\left(\log n\right) - \frac{n^{2\delta}}{2\sigma_k^2} + O\left(n^{3\delta-1/2}\right),\end{aligned}$$

which goes to $-\infty$. In other words, $\mathbb{P}\{K_s \ge 1\} = o\left(1/n^2\right)$. So $\mathbb{P}\left\{\sum_{s\in J} K_s \ge 1\right\} = o\left(1/n\right)$. □

Lemma 1 follows immediately from Lemma A2, A3, A4, and A5.

3. Special functions

Lemma A6. *Let $f(x)$, $g(x)$ and $h(x)$ be defined as in the previous subsection. Let ν_k, τ_k and σ_k be as in Lemma A1. Then*

(a) As $x \to \nu_k$, $g\left(x\right) = g(\nu_k) + O\left(|x-\nu_k|\right) = \left(1 + O\left(|x-\nu_k|\right)\right)/(\sigma_k\sqrt{1-ke^{-\tau_k}})$.

(b) $h(x)$ and $f(x)$ are strictly increasing on $(1-\frac{1}{k}, \nu_k)$ and strictly decreasing on $(\nu_k, 1)$.

(c) As $x \to \nu_k$,

$$h(x) = h(\nu_k) + O(|x-\nu_k|^3) = -\frac{(x-\nu_k)^2}{2\sigma_k^2} + O(|x-\nu_k|^3),$$

which implies that

$$f(x) = e^{h(x)} = \exp\left\{-\frac{(x-\nu_k)^2}{2\sigma_k^2}\right\} + O(|x-\nu_k|^3).$$

Proof. For (a), recall that $\sigma_k^2 \equiv \tau_k/(ke^{\tau_k}(1-ke^{-\tau_k}))$. Thus $\sigma_k^2(1-ke^{-\tau_k}) = \nu_k(1-\nu_k)$. Then $g(\nu_k) = 1/\sqrt{\nu_k(1-\nu_k)} = 1/\sigma_k\sqrt{1-ke^{-\tau_k}}$. Since $g'(x)$ is bounded around ν_k, by Taylor's theorem,

$$g(x) = g(\nu_k) + O(|x-\nu_k|) = \left(1 + O\left(|x-\nu_k|\right)\right)\frac{1}{\sigma_k\sqrt{1-ke^{-\tau_k}}}, \qquad \text{as } x \to \nu_k.$$

Let $r(x) = \log\left(f(x)^{1/x}\right) = h(x)/x$. Using $\tau_k/k \equiv 1 - e^{-\tau_k} \equiv \nu_k$ shows that

$$\gamma_k = \left(\frac{1}{e\nu_k}\right)^k e^{\tau_k}\nu_k = {\nu_k}^{-k+1}e^{-k+\tau_k} = {\nu_k}^{-k+1}(e^{-\tau_k})^{(k-\tau_k)/\tau_k} = {\nu_k}^{-k+1}(1-\nu_k)^{(1-\nu_k)/\nu_k}.$$

Then $r(\nu_k) = \log\left({\nu_k}^{k-1}(1-\nu_k)^{(\nu_k-1)/\nu_k}\gamma_k\right) = \log(1) = 0$,

$$r'(x) = \frac{k}{x} + \frac{1}{x^2}\log(1-x), \qquad \text{and} \qquad r''(x) = -\frac{k}{x^2} - \frac{2\log(1-x)}{x^3} - \frac{1}{x^2(1-x)}.$$

Therefore $r'(\nu_k) = 0$ and $r''(\nu_k) = -1/(\nu_k \sigma_k^2)$.

Since $h(x) = xr(x)$,

$$h'(x) = r(x) + xr'(x), \qquad h''(x) = 2r'(x) + xr''(x) = \frac{k}{x} - \frac{1}{x(1-x)}.$$

Thus $h(\nu_k) = 0$, $h'(\nu_k) = 0$ and $h''(\nu_k) = -1/\sigma_k^2$. Also recalling that $1 - \frac{1}{k} < 1 - \frac{1}{2k} < \nu_k < 1$ (Lemma A1), $h(x)$ is strictly concave on $(1-\frac{1}{k}, 1)$, reaching maximum at ν_k. Thus (b) is proved. The two asymptotic equations in (c) follow from Taylor's theorem. □

4. Probability generating functions of Galton-Watson processes

Lemma A7. *Let $\mu \in (0, \frac{1}{2k})$ be a constant where $k \geq 2$. Let $(Z_m)_{m\geq 0}$ be a Galton-Watson process with $Z_0 \equiv 1$ and offspring distribution* Bin(k, μ). *Let $\varphi_m(y) \equiv \mathbb{E}y^{Z_m}$. Then*

$$\varphi_m(0) \leq 1 - (k\mu)^m + \left(1 - \frac{1}{2^m}\right)(k\mu)^{m+1}.$$

Proof. We use induction. Let $c_m = 1 - 1/2^m$. For $m = 1$,

$$\varphi_1(y) = \mathbb{E}y^{Z_1} = (1 - \mu(1-y))^k.$$

Since $\mu > 0$ and $k \geq 2$, by Taylor's theorem,

$$\varphi_1(0) = (1-\mu)^k \leq 1 - k\mu + \frac{(k\mu)^2}{2} = 1 - k\mu + c_1(k\mu)^2.$$

It is well known that for $m > 1$, $\varphi_m(y) = \varphi_1(\varphi_{m-1}(y))$ (see [13]). Assuming the lemma holds for m, then

$$\begin{aligned}
\varphi_{m+1}(0) &= \varphi_1(\varphi_m(0)) = \left(1 - \mu\left(1 - \varphi_m(0)\right)\right)^k \\
&\leq \left(1 - \mu\left((k\mu)^m - c_m(k\mu)^{m+1}\right)\right)^k \\
&\leq 1 - k\mu\left((k\mu)^m - c_m(k\mu)^{m+1}\right) + \frac{k^2}{2}\mu^2\left((k\mu)^m - c_m(k\mu)^{m+1}\right)^2 \\
&= 1 - (k\mu)^{m+1} + c_m(k\mu)^{m+2} + \frac{(k\mu)^m}{2}(1 - c_m k\mu)^2(k\mu)^{m+2} \\
&\leq 1 - (k\mu)^{m+1} + c_{m+1}(k\mu)^{m+2},
\end{aligned}$$

since $k\mu < 1/2$ and $c_{m+1} = c_m + 1/2^{m+1}$. □