Packing arborescences in random digraphs

Carlos Hoppen[*]   Roberto F. Parente[†]   Cristiane M. Sato[‡]

April 16, 2018

Abstract

We study the problem of packing arborescences in the random digraph $\mathcal{D}(n, p)$, where each possible arc is included uniformly at random with probability $p = p(n)$. Let $\lambda(\mathcal{D}(n, p))$ denote the largest integer $\lambda \geq 0$ such that, for all $0 \leq \ell \leq \lambda$, we have $\sum_{i=0}^{\ell-1}(\ell - i)|\{v : d^{in}(v) = i\}| \leq \ell$. We show that the maximum number of arc-disjoint arborescences in $\mathcal{D}(n, p)$ is $\lambda(\mathcal{D}(n, p))$ a.a.s. We also give tight estimates for $\lambda(\mathcal{D}(n, p))$ depending on the range of p.

1 Introduction and main result

Many important problems in discrete mathematics deal with packing structures with some desired property into a larger structure, and their goal is typically to find as many disjoint structures with the desired property as possible. Several classical results in combinatorial optimization fit into this general framework. For instance, the maximum matching problem can be seen as packing vertex-disjoint edges. We also highlight Tutte's [21] and Nash-Williams's [20] results on packing spanning trees, as well as Menger's [19] and Mader's [18] results on packing paths. See the book by Cornuéjols [6] for many more examples.

Given the extensive literature on this topic, it is only natural that there is a great number of packing results in extremal combinatorics and random structures. For instance, the problem of packing Hamiltonian cycles in random structures has been studied quite intensively since the 1980s (see [3, 4, 5, 11, 15, 16, 17]). In the particular case of digraphs, some significant results have been obtained recently (see [7, 8, 9]).

Recently, Gao, Pérez-Giménez and the third author [12, 13] obtained results concerning packing spanning trees in random graphs. As usual, given a function $p \colon \mathbb{N} \to [0, 1]$ and a positive integer n, we let $\mathcal{G}(n, p)$ be the random graph on $[n] = \{1, \ldots, n\}$ such that each edge is included independently with probability p. Moreover, given a sequence of probability spaces $(\Omega_i, \mathcal{F}_i, \Pr_i)_{i \in \mathbb{N}}$, we say that a sequence of events $(A_i)_{i \in \mathbb{N}}$ holds *asymptotically almost surely* (a.a.s. for short) if $\Pr_n(A_n) \to 1$ as $n \to \infty$.

Theorem 1.1 (Pu–Pérez-Giménez–Sato[1] [12, 13])**.** *For $p = p(n) \in [0, 1]$, the maximum number of edge-disjoint spanning trees in $\mathcal{G}(n, p)$ is a.a.s.*

$$\min \left\{ \delta(\mathcal{G}(n, p)), \left\lfloor m(\mathcal{G}(n, p))/(n - 1) \right\rfloor \right\}.$$

It is easy to see that $\delta(\mathcal{G}(n, p))$ and $\left\lfloor m(\mathcal{G}(n, p))/(n - 1) \right\rfloor$ are upper bounds for the number of edge-disjoint spanning trees since every spanning tree has at least one edge incident to every vertex and has exactly $n - 1$ edges. The following classical result proved by Tutte and Nash-Williams is the main tool in [12, 13] to prove that the maximum is achieved by one of these two parameters.

Theorem 1.2 (Tutte [21] and Nash-Williams [20])**.** *Given a graph $G = (V, E)$ and an integer $k \geq 0$, G contains k edge-disjoint spanning trees if and only if, for every partition of V with ℓ parts, the number of edges with ends in different parts is at least $k(\ell - 1)$.*

It is quite natural that this result (which is actually a min-max relation) can be successfully used for random graphs since the partition condition is essentially an expansion condition and random graphs are well known to have nice expansion properties.

Our main result is an analogue of Theorem 1.1 for digraphs. A *digraph $D = (V, A)$* is given by its finite set V of *vertices* and its set $A \subset \{(u, v) \in V^2 \colon u \neq v\}$ of *arcs*. We say that an arc (u, v) *leaves u* and *enters v*, or, alternatively, that it *points at v*. The *underlying graph* of a digraph $D = (V, A)$ is the graph (actually multigraph) obtained by ignoring orientations on arcs. Our result deals with packing arborescences, which are an analogue of spanning trees in digraphs. Indeed, an *arborescence* of a digraph is a spanning sub-digraph such that its underlying graph is a rooted tree and each vertex except the root has in-degree 1 and the root has in-degree zero. Roughly speaking, an arborescence is a spanning tree with the arcs "pointing away" from the root. Let $\mathcal{D}(n, p)$ denote the random digraph on $[n] = \{1, \ldots, n\}$ such that each arc is included independently at random with probability p. Let $\tau(\mathcal{D}(n, p))$ denote the maximum number of arc-disjoint arborescences in $\mathcal{D}(n, p)$. For every digraph D and $v \in D$, let the *in-degree $d_D^{in}(v)$* of v be the number of arcs entering v in D, while the *out-degree $d_D^{out}(v)$* of v is the number of arcs leaving v in D. Our main result may be stated as follows.

Theorem 1.3. *For every $p = p(n) \in [0, 1]$, the maximum number of arc-disjoint arborescences in $\mathcal{D}(n, p)$ a.a.s. satisfies*

$$\tau(\mathcal{D}(n, p)) = \lambda(\mathcal{D}(n, p)), \tag{1}$$

where $\lambda(\mathcal{D}(n, p))$ is the maximum integer $\lambda \geq 0$ such that, for all $0 \leq \ell \leq \lambda$,

$$\sum_{i=0}^{\ell-1} (\ell - i) |\{v \colon d_{\mathcal{D}(n,p)}^{in}(v) = i\}| \leq \ell. \tag{2}$$

Moreover,

(a) *if $p = (\log(n) - h(n))/(n - 1)$ with $h(n) = \omega(1)$, then $\lambda(\mathcal{D}(n, p)) = 0$ a.a.s.;*

(b) *if $p = (\log(n) + h(n))/(n - 1)$ with $h(n) = O(\log \log n)$, then $\lambda(\mathcal{D}(n, p)) \in \{\delta^{in}, \delta^{in} + 1\}$ a.a.s.;*

---

[1]The result stated here is not the strongest result obtained in [12, 13].

(c) if $p = (\log(n) + h(n))/(n-1)$ with $h(n) = o(\log n)$ and $h(n) = \Omega(\log \log n)$, then $\lambda(\mathcal{D}(n,p)) \sim \delta^{in}$ a.a.s.

One interesting feature of our result is that $\tau(\mathcal{D}(n,p))$ has a very strong relation with the number of vertices with low degrees. This differs from the graph case in the following sense. Theorem 1.1 tells us that, for random graphs, the obstacles to pack spanning trees are quite simple: either we do not have enough edges to get more spanning trees or we exhausted the edges incident with a vertex. Our result shows that for random digraphs the obstacles to pack arborescences are more intricate while still arising from natural constraints. This is due to the fact that the root of an arborescence plays a special role, which does not happen for undirected graphs. In our case, the reason why $\lambda(\mathcal{D}(n,p))$ is an upper bound for $\tau(\mathcal{D}(n,p))$ is that, in order to pack ℓ arborescences, every vertex of $\mathcal{D}(n,p)$ whose in-degree is $\ell - i$ must be the root of at least i arborescences since its in-degree would be exhausted. Quite interestingly, our condition does not involve the out-degrees.

Similarly to the undirected case, the core of our proof relies on a result on combinatorial optimization, which was proved by Frank [10] and is an analogue of Theorem 1.2 for digraphs. Instead of dealing with partitions, Frank's result imposes conditions on subpartitions. A *subpartition* of a set S is a collection of pairwise disjoint non-empty subsets of S. Note that, unlike a partition, a subpartition does not need to include every element of S. For every digraph $D = (V, A)$ and $S \subseteq V$, let $d_D^{in}(S)$ denote the number of arcs entering S (from $V \setminus S$). For future reference, let also $d_D^{out}(S)$ be the number of arcs leaving S (to $V \setminus S$) in D.

Theorem 1.4 (Frank [10]). *Let $D = (V, A)$ be a digraph and let $k \geq 0$ be an integer. Then D contains k arc-disjoint arborescences if, and only if, for every subpartition $\mathcal{P}$ of V, we have*

$$\sum_{U \in \mathcal{P}} d_D^{in}(U) \geq k(|\mathcal{P}| - 1). \tag{3}$$

One of the difficulties of working with subpartitions instead of partitions is that some vertices may be not included in any part and the arcs entering such vertices do not contribute to the summation in (3), which is something that did not occur in the graph case.

In terms of previous results about arborescences in random digraphs, Bal, Bennett, Cooper, Frieze, and Prałat [2] have proved that in the random digraph process (where the arcs are added one-by-one), the digraph contains an arborescence a.a.s. in the step where there is single a vertex with in-degree zero[2].

Organization of the paper. This paper is organized as follows. In Section 2, we introduce the main definitions and notation used in the paper. In Section 3, we present the main properties of $\mathcal{D}(n,p)$ that are used: in Section 3.2 we study properties of the degrees in $\mathcal{D}(n,p)$; in Section 3.3 we show the relation between $\lambda(\mathcal{D}(n,p))$ and the minimum in-degree; and in Section 3.4 we prove a few basic expansion properties of $\mathcal{D}(n,p)$. Finally, in Section 4 we combine the results from Section 3 with the result by Frank (Theorem 1.4) to complete the proof of Theorem 1.3.

---

[2]This is part of their main result about rainbow arborescences.

2 Definitions and notation

In this section, we define the main concepts used in this paper. We will repeat a few definitions already presented in the introduction so that the reader can easily find any of them.

Definition 2.1. *(Random digraph $\mathcal{D}(n,p)$) Given a function $p = p(n)\colon \mathbb{N} \to [0,1]$, let $\mathcal{D}(n,p)$ denote the random digraph with vertex set $[n] = \{1, 2, \ldots, n\}$ such that each of the $n(n-1)$ arcs is included independently at random with probability p.*

Definition 2.2. *(Neighbourhoods and degrees) Given a digraph $D = (V, A)$ and $v \in V$, we define the* in-neighbourhood *of v, denoted by $\Gamma_D^{in}(v)$, as the set $\{u \in V\colon (u, v) \in A\}$. Similarly, we define the* out-neighbourhood *of v, denoted by $\Gamma_D^{out}(v)$, as the set $\{u \in V\colon (v, u) \in A\}$. Moreover, we define the* in-degree *of v as $d_D^{in}(v) = |\Gamma_D^{in}(v)|$ and the* out-degree *of v as $d_D^{out}(v) = |\Gamma_D^{out}(v)|$. That is, $d_D^{in}(v)$ is the number of arcs "entering" v and $d_D^{out}(v)$ is the number of arcs "leaving" v. Let $\delta^{in}(D) = \min_{v \in V} d_D^{in}(v)$ and $\delta^{out}(D) = \min_{v \in V} d_D^{out}(v)$.*

Definition 2.3. *(Cuts) Given a digraph $D = (V, A)$ and disjoint sets $S, S' \subseteq V$, we define $A_D(S, S')$ as the set of arcs $(u, v) \in A$ such that $u \in S$ and $v \in S'$.*

Definition 2.4. *(Induced digraphs) Given a digraph $D = (V, A)$ and $S \subseteq V$, we define $D[S]$ as the digraph with vertex set S with edge set $A_D[S] = \{(u, v) \in A\colon u \in S,\ v \in S\}$.*

Definition 2.5. *(Arborescences) An* arborescence *of a digraph $D = (V, A)$ is a digraph $T = (V, A_T)$ where $A_T \subseteq A$ such that the underlying graph of T is tree and each vertex except the root has in-degree 1 and the root has in-degree zero. Let $\tau(D)$ denote the maximum number of arc-disjoint arborescences in D.*

Definition 2.6. *Given a digraph $D = (V, A)$, let $\lambda(D)$ denote the maximum integer $\lambda \geq 0$ such that, for all $0 \leq \ell \leq \lambda$,*

$$\sum_{i=0}^{\ell-1} (\ell - i)|\{v\colon d_D^{in}(v) = i\}| \leq \ell. \tag{4}$$

We use $d^{in}(v)$ to denote $d_{\mathcal{D}(n,p)}^{in}(v)$ for ease of notation. Similarly, $d^{out}(v) = d_{\mathcal{D}(n,p)}^{out}(v)$, $\delta^{in} = \delta^{in}(\mathcal{D}(n,p))$, $\delta^{out} = \delta^{out}(\mathcal{D}(n,p))$, $\tau = \tau(\mathcal{D}(n,p))$ and $\lambda = \lambda(\mathcal{D}(n,p))$, and so on.

In all results in this paper, except stated otherwise, the probability space is the one defined by $\mathcal{D}(n,p)$ and the asymptotics are for n going to infinity. We use standard asymptotic notation, which may be found in [14, Section 1.2].

In many proofs, we will use the well known subsubsequence principle, which states that, if x is a constant and (x_n) is a real sequence whose subsequences always have a subsubsequence converging to x, then $x_n \to x$.

3 Properties of the random digraph $\mathcal{D}(n,p)$

In this section, we study the behaviour of the degrees in $\mathcal{D}(n,p)$. We also prove some simple properties about cuts in $\mathcal{D}(n,p)$. In Section 3.1, we state two basic results on binomial random variables that are used throughout the paper. For basic probabilistic results (such as Markov's and Chebyshev's inequality), we refer the reader to Alon and Spencer [1].

3.1 Properties of binomial random variables

In this section, we state two results on binomial random variables.

Theorem 3.1 (Chernoff's bounds [14]). *Let $X_1, \ldots, X_n$ denote n independent Bernoulli variables. Let $X = \sum_{i=1}^{n} X_i$ and let $\mu = \mathbb{E}[X]$. Then, for any $0 < \tau < 1$,*

$$\Pr(X \geq (1+\tau)\mu) \leq e^{-\tau^2 \mu/3}, \tag{5}$$

$$\Pr(X \leq (1-\tau)\mu) \leq e^{-\tau^2 \mu/2}. \tag{6}$$

Lemma 3.2 (Lemma 16 [12, 13]). *For every constant $\eta > 0$ there exist positive constants C_1 and C_2 such that the following holds for any function $0 \leq p = p(n) \leq 1/\sqrt{n}$ and every integer $0 < k \leq (1-\eta)np$. Let $X \sim \mathrm{Bin}(n, p)$. Then,*

$$\Pr(X \leq k) = C \left(\frac{e^{-pn}}{\sqrt{k}} \right) \left(\frac{epn}{k} \right)^k, \quad \text{with } C_1 \leq C \leq C_2. \tag{7}$$

3.2 Degrees in $\mathcal{D}(n, p)$

In this section, we present some results on the minimum in-degree and out-degree in $\mathcal{D}(n, p)$. We also prove some properties of vertices with low degree.

The following lemma is an application of Lemma 3.2 to the in-degrees and out-degrees of $\mathcal{D}(n, p)$.

Lemma 3.3. *Let $0 < \eta < 1$ be a constant. There exist constants $C_1 > 0$ and $C_2 > 0$ such that, for any function $\alpha = \alpha(n) \in (0, 1-\eta]$ and any function p satisfying $0.9 \log n/(n-1) \leq p \leq 1/\sqrt{n}$, the following holds[3]:*

(i) There exists $C = C(n) \in [C_1, C_2]$ such that, for every $v \in [n]$,

$$\Pr\left(d^{in/out}(v) \leq \alpha p(n-1)\right) = \frac{C}{\sqrt{\alpha p(n-1)}} \exp\left(-p(n-1)\left(1 - \alpha \log\left(\frac{e}{\alpha}\right)\right)\right).$$

(ii) $\Pr\left(\delta^{in/out} \leq \alpha p(n-1)\right) \leq \dfrac{C_2}{\sqrt{\alpha p(n-1)}} \exp\left(\log n - p(n-1)\left(1 - \alpha \log\left(e/\alpha\right)\right)\right);$

(iii) $\Pr\left(\delta^{in/out} > \alpha p(n-1)\right) \leq \dfrac{\sqrt{\alpha p(n-1)}}{C_1} \exp\left(p(n-1)\left(1 - \alpha \log\left(e/\alpha\right)\right) - \log n\right).$

(iv) $\Pr\left(\exists v \in [n] \text{ s.t. } \min\{d^{in}(v), d^{out}(v)\} \leq \alpha p(n-1)\right)$

$$\leq \frac{(C_2)^2}{\alpha p(n-1)} \exp\left(\log n - 2p(n-1)\left(1 - \alpha \log\left(\frac{e}{\alpha}\right)\right)\right).$$

Proof. The proof of (i)–(iii) is basically the same as the proof of [12, 13, Lemma 18]. We include it here for the sake of completeness.

---

[3]Here we use $d^{in/out}(v)$ to denote either $d^{in}(v)$ or $d^{out}(v)$.

For every $v \in V$, the in-degree $d^{in}(v)$ of v has distribution $\text{Bin}(n-1, p)$. Thus, by Lemma 3.2, there exist C_1 and C_2 (depending only on η) and a constant $C = C(n) \in [C_1, C_2]$ such that

$$\Pr\left(d^{in}(v) \le \alpha p(n-1)\right) = C\left(\frac{e^{-p(n-1)}}{\sqrt{\alpha p(n-1)}}\right)\left(\frac{e}{\alpha}\right)^{\alpha p(n-1)}$$

$$= \frac{C}{\sqrt{\alpha p(n-1)}} \exp\left(-p(n-1)\left(1 - \alpha \log\left(\frac{e}{\alpha}\right)\right)\right). \quad (8)$$

This proves (i). Let Y denote the number of vertices $v \in [n]$ such that $d^{in}(v) \le \alpha p(n-1)$. Then, by (8), we conclude that

$$\mathbb{E}(Y) \ge \frac{C_1}{\sqrt{\alpha p(n-1)}} \exp\left(\log n - p(n-1)\left(1 - \alpha \log\left(\frac{e}{\alpha}\right)\right)\right)$$

and

$$\mathbb{E}(Y) \le \frac{C_2}{\sqrt{\alpha p(n-1)}} \exp\left(\log n - p(n-1)\left(1 - \alpha \log\left(\frac{e}{\alpha}\right)\right)\right). \quad (9)$$

Thus, (ii) follows by applying Markov's inequality. Since the in-degrees of distinct vertices are independent random variables, Y is a binomial random variable with probability p' given by (8). Thus, $\text{Var}(Y) = np'(1-p') \le \mathbb{E}Y$ and so, by Chebyshev's inequality, we obtain

$$\Pr(Y = 0) \le \frac{\text{Var}(Y)}{(\mathbb{E}Y)^2} \le \frac{1}{\mathbb{E}Y} \le \frac{\sqrt{\alpha p(n-1)}}{C_1} \exp\left(-\log n + p(n-1)\left(1 - \alpha \log\left(\frac{e}{\alpha}\right)\right)\right).$$

Proving (i), (ii) and (iii) for d^{out} and δ^{out} is analogous.

For any $v \in [n]$, (iv) follows trivially from (9) (and its analogue for $d^{out}(v)$) and the fact that $d^{in}(v)$ and $d^{out}(v)$ are independent random variables. ∎

Corollary 3.4. *Let $\alpha \in (0, 1)$ and let $h(n)$ be a function such that $h(n) = o(\log n)$. Let $p = p(n) = (\log(n) + h(n))/(n-1)$. Then there is $C > 0$ such that $\delta^{in} < \alpha \log n$ with probability at least $1 - n^{-C}$.*

Proof. It suffices to prove the result for $h(n) \ge 0$. Let $\alpha \in (0, 1)$ be a constant. Note that $\beta := 1 - \alpha \log(e/\alpha)$ is a constant less than 1. Lemma 3.3(iii) leads to

$$\Pr(\delta^{in} > \alpha p(n-1)) \le \frac{\sqrt{\alpha p(n-1)}}{C_1} \exp\left(\beta(\log n + h(n)) - \log n\right) = o(n^{(\beta-1)/2}).$$

∎

Lemma 3.5. *Let $p = p(n) = (\log(n) + h(n))/(n-1)$ be such that $h(n) \le C' \log\log n$, where C' is a positive constant. Then, for any constant $C > C'$, $\delta^{in} \le C$ a.a.s.*

Proof. The expected number of vertices with in-degree C in $\mathcal{D}(n, p)$ is

$$n\binom{n-1}{C}p^C(1-p)^{n-C-1} \ge \exp\left(\log n + C\log\log n - C\log C - (n-1)p + o(1)\right)$$

$$\ge \exp\left(-h(n) + C\log\log n - C\log C + o(1)\right) = \omega(1).$$

As in the proof of Lemma 3.3(iii), we may apply Chebyshev's inequality to conclude that the minimum in-degree is at most C a.a.s. ∎

It turns out that the following real function appears often in our paper:

$$F(x) = 1 - x \log \frac{e}{x} = 1 - x + x \log x. \tag{10}$$

This is a continuous and strictly decreasing function for $x \in (0, 1)$. Moreover, $F(1) = 0$ and $\lim_{x \to 0^+} F(x) = 1$. In particular, for every $\phi > 1$, there exists a single $\alpha \in (0, 1)$ such that $F(\alpha) = 1/\phi$.

Corollary 3.6. *Let $\phi > 1$ be a constant. Let $\alpha \in (0, 1)$ be such that $F(\alpha) = 1 - \alpha + \alpha \log \alpha = 1/\phi$. For $p = p(n) \sim \phi \log(n)/(n - 1)$, we have $\delta^{in} \sim \alpha p(n - 1)$ and $\delta^{out} \sim \alpha p(n - 1)$ a.a.s.*

Proof. Let $\gamma = \gamma(n)$ be such that $\gamma(n) \log n = p(n - 1)$. Thus, $\gamma \sim \phi$. Let $\varepsilon \in (0, \min\{\alpha, 1 - \alpha\})$ be a constant.

We have $\lim_{n \to \infty} \gamma(n) F(\alpha + \varepsilon) < F(\alpha) \lim_{n \to \infty} \gamma(n) = 1$ and then by Lemma 3.3(iii), we obtain $\delta^{in} \leq (\alpha + \varepsilon)p(n - 1)$ with probability going to 1. On the other hand, since $\lim_{n \to \infty}(\gamma(n) F(\alpha - \varepsilon)) > 1$, by Lemma 3.3(ii), we have $\delta^{in} \geq (\alpha - \varepsilon)p(n - 1)$ a.a.s. The proof for δ^{out} is similar. ∎

Definition 3.7. *We say that a vertex $v \in [n]$ is ε-in-light if $d^{in}(v) \leq \delta^{in} + \varepsilon np$ and ε-out-light if $d^{out}(v) \leq \delta^{out} + \varepsilon np$.*

Lemma 3.8. *For every constant $\varphi \geq 0.9$, there exists $\varepsilon > 0$ such that the following holds for sufficiently large n. For any $0.9 \log n/(n - 1) \leq p \leq \varphi \log n/(n - 1)$, with probability at least $1 - n^{-0.18}$, there is no pair (u, v) of ε-in-light vertices such that $uv \in A$ or $\Gamma^{in}(v) \cap \Gamma^{in}(u) \neq \varnothing$ and there is no pair (u, v) of ε-out-light vertices such that $uv \in A$ or $\Gamma^{out}(v) \cap \Gamma^{out}(u) \neq \varnothing$.*

Proof. We claim that it is possible to choose α and ε so that, for large n, we have

$$\Pr(\delta^{in} > \alpha(n - 1)p) \leq n^{-0.19} \text{ and } \Pr(d^{in}(v) \leq (\alpha + \varepsilon)p(n - 1)) \leq n^{-0.7}, \ \forall v \in V. \tag{11}$$

Assuming that this claim holds, and conditioning on the event that $\delta^{in} \leq \alpha(n - 1)p$, let S be the set of vertices $v \in [n]$ such that $d^{in}(v) \leq (\alpha + \varepsilon)p(n - 1)$. Note that S contains all ε-in-light vertices. Then, for any vertices $u, v \in V$, by our choice of α and ε, we have

$$\Pr(uv \in A, u \in S \text{ and } v \in S) = p \Pr(v \in S | uv \in A) \Pr(u \in S | uv \in A)$$
$$\leq (1 + o(1))pn^{-1.4}.$$

Since we have at most $n(n - 1)$ choices for (u, v) and $p \leq \varphi \log n/(n - 1)$, the expected number of pairs of adjacent vertices in S is $(1 + o(1))\varphi n^{-0.4} \log n$. Thus, the probability that there are adjacent ε-in-light vertices is at most $n^{-0.19} + (1 + o(1))\varphi n^{-0.4} \log n \leq \frac{1}{4} n^{-0.18}$ for sufficiently large n.

For any vertices $u, v, z \in V$, by our choice of α and ε, we have

$$\Pr(zu, zv \in A, u \in S \text{ and } v \in S) = p^2 \Pr(v \in S | zv \in A) \Pr(u \in S | zu \in A)$$
$$\leq (1 + o(1))p^2 n^{-1.4}.$$

Hence, since we have at most $n(n - 1)(n - 2)$ choices for (u, v, z) and $p \leq \varphi \log n/(n - 1)$, the expected number of pairs of adjacent ε-in-light vertices is $(1 + o(1))\varphi n^{-0.4} \log^2 n$ and the result follows as above. Note that the same argument applies for the out-degree.

Finally we show how to choose α and ε to obtain the desired bounds for the probabilities in (11). To this end, let $F(x) = 1 - x + x \log x$ be the function defined in (10), and let β be the constant such that $\beta \sim (n-1)p/\log n$. By hypothesis, $0.9 \leq \beta \leq \varphi$. Choose α so that $\beta F(\alpha) = 0.81$. Then the RHS of Lemma 3.3(iii) is at most $O(1) \exp(-0.2 \log n + (1/2) \log \log n) \leq n^{-0.19}$. We can then choose $\varepsilon > 0$ so that $\beta F(\alpha + \varepsilon) = 0.71$ and the RHS of Lemma 3.3(i) becomes $O(1) \exp(-\beta \log n F(\alpha + \varepsilon) - (1/2) \log \log n) \leq \exp(-0.7 \log n - (1/2) \log \log n) \leq n^{-0.7}$. ∎

The next result follows immediately from Chernoff's inequality (Theorem 3.1).

Lemma 3.9. *Let $\psi = \psi(n) = \omega(1)$ and $\phi = 1/\sqrt{\log n}$. Let $p = p(n)$ be such that $(n-1)p = \psi \log n$ a.a.s. for all $v \in [v]$ we have $(1-\phi)p(n-1) \leq d^{in}(v) \leq (1+\phi)p(n-1)$ and $(1-\phi)p(n-1) \leq d^{out}(v) \leq (1+\phi)p(n-1)$.*

For a function $p \colon \mathbb{N} \to [0,1]$ and integers $n, k \in \mathbb{N}$, we consider the random variable $Y_k = Y_k(\mathcal{D}(n,p))$ that counts the number of vertices of in-degree k in $\mathcal{D}(n,p)$.

Definition 3.10. *Let $\delta^* = \delta^*(\mathcal{D}(n,p))$ denote the minimum integer $k \geq 0$ such that $\mathbb{E}[Y_k] \geq 1$ in $\mathcal{D}(n,p)$.*

The following result shows the relation between δ^* and δ^{in}.

Lemma 3.11. *Let $p = p(n) \sim \log(n)/(n-1)$. Then $\delta^{in} \in \{\delta^* - 1, \delta^*, \delta^* + 1\}$ a.a.s. Moreover, we have that, if $\mathbb{E}[Y_{\delta^*-1}] = o(1)$, then $\delta^{in} \in \{\delta^*, \delta^*+1\}$ and, if $\mathbb{E}[Y_{\delta^*-1}] = \Omega(1)$, then $\delta^{in} \in \{\delta^* - 1, \delta^*\}$.*

Proof. It is straightforward to show that there exists $k^*(n) = o(\log n)$ such that $\mathbb{E}[Y_{k^*(n)}] = \omega(1)$. Thus, $\delta^* = o(\log n)$. For every $k = o(\log n)$, we have

$$\mathbb{E}[Y_{k-1}] / \mathbb{E}[Y_k] \sim k/((n-1)p) \tag{12}$$

First assume that $\mathbb{E}[Y_{\delta^*-1}] = o(1)$. By Markov's inequality and by (12), we have $\delta^{in} \geq \delta^*$ a.a.s.

Again by (12), we obtain $\mathbb{E}[Y_{\delta^*+1}] \sim np/\delta^* \cdot \mathbb{E}[Y_{\delta^*}] = \omega(1)$. Hence, by Chebychev's inequality (using the independence of the in-degrees), we have $\Pr(Y_{\delta^*+1} = 0) = o(1)$, so that $\delta^{in} \leq \delta^* + 1$ a.a.s. The proof for the case $\mathbb{E}[Y_{\delta^*-1}] = \Omega(1)$ is similar. ∎

3.3 Estimating $\lambda(\mathcal{D}(n,p))$

In this section, we give tight estimates for $\lambda = \lambda(\mathcal{D}(n,p))$ depending on the range of p.

The easiest case is $p = \omega(\log n/n)$, when our estimate follows immediately from Lemma 3.9 and the fact that λ is between the minimum and the maximum in-degree.

Corollary 3.12. *If $\phi = \phi(n) = \omega(1)$ is a function and $p = p(n) \sim \phi \log(n)/(n-1)$, then $\lambda(n) \sim p(n-1)$ a.a.s.*

Next we consider other ranges of p.

Lemma 3.13. *Let $\phi > 1$ be a constant. If $p = p(n) \sim \phi \log(n)/(n-1)$, then $\lambda \sim \delta^{in}$.*

Proof. By Corollary 3.6, for $\alpha \in (0, 1)$ such that $F(\alpha) = 1/\phi$, we have $\delta^{in} \sim \alpha p(n-1)$ a.a.s. Given $\varepsilon > 0$, Lemma 3.3(i) ensures that there is $0 < \beta < 1$ (depending on α and ε) such that, for any vertex v, we have

$$\Pr(d^{in}(v) \leq \alpha(1+\varepsilon)p(n-1)) = \Theta\left(n^{-\beta}\right).$$

Since the in-degrees of distinct vertices are independent, we may apply Chernoff's inequality (Theorem 3.1) to the binomial random variable counting the number of vertices whose in-degree is at most $\alpha(1+\varepsilon)p(n-1)$ to conclude that there are $\Theta(n^{1-\beta})$ such vertices a.a.s. Since $n^{1-\beta} = \omega(\alpha(1+\varepsilon)p(n-1))$, this implies that $\lambda \leq \alpha(1+\varepsilon)p(n-1) \sim (1+\varepsilon)\delta^{in}$ a.a.s. Since $\lambda \geq \delta^{in}$ holds trivially, our result follows. ∎

In the next cases, we use the following simple fact.

Claim 3.14. *Let D be a digraph and $k \geq 0$ an integer. If $Y_k > k + 1$, then $\lambda(D) \leq k$.*

Lemma 3.15. *Let $p = p(n) = (\log(n) + h(n))/(n-1)$ be such that $h(n) = O(\log \log n)$. Then $\lambda(\mathcal{D}(n,p)) \in \{\delta^{in}, \delta^{in} + 1\}$ a.a.s.*

Proof. Consider δ^* from Definition 3.10. First assume that $\mathbb{E}[Y_{\delta^*-1}] = o(1)$. We shall prove that $\lambda \leq \delta^* + 1$ a.a.s., which leads to the desired conclusion because $\lambda \geq \delta^{in}$ and Lemma 3.11 ensures that $\delta^{in} \in \{\delta^*, \delta^*+1\}$ a.a.s.. The definition of δ^* implies that $\mathbb{E}[Y_\delta^*] \geq 1$, and Lemma 3.5 implies that there exists $C > 0$ such that $\delta^{in} \leq C$ a.a.s., and hence $\delta^* = O(1)$. Moreover, $\mathbb{E}[Y_{\delta^*+1}] \sim np/\delta^* \cdot \mathbb{E}[Y_{\delta^*}]$, so that $\mathbb{E}[Y_{\delta^*+1}] \geq np/\delta^*(1 + o(1)) = \omega(1)$. By Chernoff's inequality (Theorem 3.1), $\Pr(Y_{\delta^*+1} \leq \mathbb{E}[Y_{\delta^*+1}]/2) \leq \exp(-A \log n))$, where $A > 0$. Thus, $Y_{\delta^*+1} > \delta^* + 2$ a.a.s. and so $\lambda \leq \delta^* + 1$ a.a.s. by Claim 3.14.

The proof is similar for the case $\mathbb{E}[Y_{\delta^*-1}] = \Omega(1)$, where we prove that $\lambda \leq \delta^*$ a.a.s. The result then follows by the subsubsequence principle. ∎

Lemma 3.16. *Let $p = p(n) = (\log(n) + h(n))/(n-1)$ be such that $h(n) = o(\log n)$ and $h(n) = \omega(\log \log n)$. Then $\lambda(\mathcal{D}(n,p)) \sim \delta^{in}$ a.a.s.*

Proof. The proof is very similar to the proof for Lemma 3.15. Let $\varepsilon > 0$ be a constant and fix $T = \lfloor(1+\varepsilon)\delta^{in}\rfloor$. By Lemma 3.3(ii), we have $T > \delta^{in}$ a.a.s. We will address only the case where $\mathbb{E}[Y_{\delta^*-1}] = o(1)$. By Lemma 3.11, we have $\delta^{in} \in \{\delta^*, \delta^* + 1\}$ a.a.s. To show that $\lambda \leq T$, we prove that $Y_T > T + 1$ a.a.s. By Corollary 3.4, we have that $\delta^{in} = o(\log n)$ and so $T = o(\log n)$ as well. Then, for $k = \delta^{in}$, we have

$$\frac{\mathbb{E}[Y_T]}{\mathbb{E}[Y_k]} \sim \left(\frac{np}{k}\right)^{T-k} = \omega(T). \tag{13}$$

Since $\mathbb{E}[Y_k] = \Omega(1)$, we have that $Y_T = \omega(T)$ a.a.s. by Chernoff's inequality (Theorem 3.1), which implies that $\lambda \leq T$ a.a.s. ∎

3.4 Expansion properties

In this section, we investigate properties of the cuts of $\mathcal{D}(n,p)$. We start by proving a simple result about the number of arcs going from a "large" set to another "large" set of vertices.

Lemma 3.17. *Let $f(n) \to \infty$ and let ζ be a positive constant. There exists a positive constant C such that, for $p = p(n) \in [f(n)/n, 1]$ and large n, the probability that there exist disjoint sets $S, S' \subseteq [n]$ with size at least ζn such that $|A(S', S)| < \zeta^2 n^2 p / 2$ is at most n^{-C}.*

Proof. Let $S, S' \subseteq [n]$ be disjoint sets with size at least ζn. Then $|A(S, S')|$ has distribution $\text{Bin}(|S'||S|, p)$. Thus, $\mathbb{E}(|A(S', S)|) = |S||S'|p$ and by Chernoff's inequality (Theorem 3.1), we have

$$\Pr\left(|A(S', S)| \leq \frac{\zeta^2 n^2 p}{2} \right) \leq \exp\left(-\frac{\zeta^2 n^2 p}{8} \right). \tag{14}$$

By the union bound, the probability that there exist disjoint sets $S, S' \subseteq [n]$ with size at least ζn such that $|A(S', S)| < \zeta^2 n^2 p / 2$ is at most

$$\sum_{s, s' \geq \zeta n} \binom{n}{s} \binom{n}{s'} \exp\left(-\frac{\zeta^2 n^2 p}{4} \right) \leq 4^n \exp\left(-\frac{\zeta^2 n^2 p}{4} \right) < \exp\left(2n - \frac{\zeta^2 n^2 p}{4} \right),$$

and the result follows since $np \geq f(n) \to \infty$. ∎

Next we prove a lemma about the number of induced arcs in sets that are "not too large". Later we will use this lemma to argue that many arcs must leave such sets.

Lemma 3.18. *Consider a function $f = f(n) \to \infty$, and let ϕ be positive constant. There exists a positive constant ζ such that, for $p = p(n) \in [f/n, 1]$ and large n, the probability that there exists $S \subseteq [n]$ with size $|S| \leq \zeta n$ such that $|A[S]| > \phi n |S| p$ is at most $e^{-f^2/2}$.*

Proof. Let $\zeta > 0$ be sufficiently small so that $e\zeta/\phi \leq e^{-1/\phi^2}$. Let $S \subseteq [n]$ be a set with size $s \leq \zeta n$. If $s \leq \phi np$, then $|A[S]| \leq s(s-1) \leq \phi nps$. So assume $s \geq \phi np$. Then the probability that $|A[S]| > \lceil \phi nps \rceil$ is at most $\binom{s(s-1)}{\lceil \phi nps \rceil} p^{\lceil \phi nps \rceil}$. Thus, the expected number of sets S with size at most ζn with $|A[S]| > \phi pn|S|$ is at most

$$\sum_{\phi np \leq s \leq \zeta n} \binom{n}{s} \binom{s^2}{\lceil \phi nsp \rceil} p^{\lceil \phi nsp \rceil} \leq \sum_{\phi np \leq s \leq \zeta n} \left(\frac{ne}{s} \right)^s \left(\frac{se}{\phi n} \right)^{\lceil \phi nsp \rceil}$$

$$= \sum_{\phi np \leq s \leq \zeta n} \left(\left(\frac{e^2}{\phi} \right) \left(\frac{se}{\phi n} \right)^{\frac{\lceil \phi nsp \rceil}{s} - 1} \right)^s$$

$$\leq \sum_{\phi np \leq s \leq \zeta n} \left(\left(\frac{e^2}{\phi} \right) \left(\frac{\zeta e}{\phi} \right)^{\frac{\lceil \phi nsp \rceil}{s} - 1} \right)^s \quad \text{since } s \leq \zeta n \text{ and } np \geq f \to \infty$$

$$\leq \sum_{\phi np \leq s \leq \zeta n} \left(\left(\frac{e^2}{\phi} \right) \left(\frac{\zeta e}{\phi} \right)^{\phi np - 1} \right)^s \quad \text{since } \zeta e/\phi < 1$$

$$\leq \sum_{\phi np \leq s \leq \zeta n} \left(\beta e^{-\frac{np}{\phi}} \right)^s, \quad \text{for } \beta = e/\zeta \text{ since } \zeta e/\phi < e^{-1/\phi^2}$$

$$\leq 2 (\beta e^{-\frac{np}{\phi}})^{\phi np} \leq e^{-(np)^2/2} \leq e^{-f^2/2},$$

for n sufficiently large. The result then follows by Markov's inequality. ∎

In the next lemma, we compare $d^{in}(S)$ and δ^{in} in the range where $p = (1+\Omega(1)) \log n/(n-1)$ for S of size from 2 to $n-2$.

Lemma 3.19. *Let ψ be a positive constant. There is a positive constant C such that, for $p = p(n) \in [(1 + \psi) \log n/(n - 1), 1]$ and large n, the probability that there exists a set $S \subset [n]$ with $2 \le |S| \le n - 2$ such that $d^{in}(S) < 1.5 \delta^{in}$ is at most n^{-C}.*

Proof. Let $\gamma(n)$ be such that $p(n - 1) = \gamma(n) \log n$. First assume that $\gamma(n) \ge 100$. Let $0 < \varepsilon_1 < 5/2 - \sqrt{6}$, $\tau = 2(1+\varepsilon_1)/10$ and $\varepsilon_2 > 0$ be such that $2(1-\tau)(1-\varepsilon_1) \ge 1.5(1+\varepsilon_2)$.

Let $S \subseteq [n]$ and let $\bar{S} = [n] \backslash S$. Then $d^{in}(S) = d^{out}(\bar{S}) = |A(\bar{S}, S)|$ which is distributed as $\text{Bin}(s\bar{s}, p)$ where $s = |S|$ and $\bar{s} = |\bar{S}|$. Then, by Chernoff's inequality (Theorem 3.1) and the union bound, the probability that there exists a set $S \subset [n]$ with $2 \le |S| \le n - 2$ such that $d^{in}(S) < (1 - \tau)s\bar{s}p$ is at most

$$\sum_{2 \le s \le n-2} \binom{n}{s} \exp\left(-\frac{\tau^2 s\bar{s}p}{2}\right) \le 2 \sum_{2 \le s \le n/2} \binom{n}{s} \exp\left(-\frac{\tau^2 s\bar{s}p}{2}\right)$$

$$\le 2 \sum_{2 \le s \le n/2} \exp\left(s + s \log n - s \log s - \frac{\tau^2 |S||\bar{S}|p}{2}\right)$$

$$\le 2 \sum_{2 \le s \le n/2} \exp\left(s \log n \left(1 + \frac{1 - \log s}{\log n} - \frac{\gamma\tau^2}{4}\right)\right) \quad (15)$$

$$\le 2 \sum_{2 \le s \le n/2} \exp\left(-\varepsilon_1 s \log n\right) \le 4n^{-2\varepsilon_1}. \quad (16)$$

To obtain (15), we use $|S||\bar{S}| = s(n - s) \ge sn/2$, and to obtain (16), we use $\gamma\tau^2/4 \ge (1 + \varepsilon_1)^2$. Note that, if $|S| \le \varepsilon_1 n$ or $|\bar{S}| \le \varepsilon_1 n$, by our choice of τ, ε_1 and ε_2,

$$(1 - \tau)|S||\bar{S}|p \ge (1 - \tau)2(1 - \varepsilon_1)np \ge 1.5(1 + \varepsilon_2)np. \quad (17)$$

If $|S| \ge \varepsilon_1 n$ and $|\bar{S}| \ge \varepsilon_1 n$, then for sufficiently large n,

$$(1 - \tau)|S||\bar{S}|p \ge (1 - \tau)\varepsilon_1^2 n^2 p > 1.5(1 + \varepsilon_2)np. \quad (18)$$

It is easy to see that $\delta^{in} \le (1 + \varepsilon_2)np$ with very large probability. Indeed, by Chernoff's inequality (Theorem 3.1), the total number of arcs in the random digraph satisfies $|A| \ge (1 + \varepsilon_2)n(n - 1)p$ with probability at most $\exp(-\varepsilon_2^2 n(n - 1)p/2) \le n^{-C}$ for any $C > 0$.

Now assume that $1 + \psi \le \gamma(n) \le 100$. By Corollary 3.6, there exist $\alpha_1(n)$, $\alpha_2(n)$ and a constant $x_1 > 0$ such that

$$x_1 < \alpha_1(n) < \alpha_2(n) < 1, \quad \alpha_2(n) - \alpha_1(n) \le x_1/7$$

and

$$\alpha_1 pn \le \delta^{in} \le \alpha_2 pn \quad \text{and} \quad \alpha_1 pn \le \delta^{out} \le \alpha_2 pn.$$

Let $\zeta > 0$ be given by Lemma 3.18 applied to $\phi = x_1/7$ and $f(n) = \gamma(n) \log n$. Let also C_2 be given by Lemma 3.17 for ζ and $f(n)$.

Fix $S \subseteq [n]$. If $|S|, |\bar{S}| \geq \zeta n$, by Lemma 3.17, $d^{in}(S) = |A(\bar{S}, S)| \geq (\zeta^2/2)n^2 p \geq 1.5\alpha_2 np \geq 1.5\delta^{in}$ with probability at least $1 - n^{-C_2}$ for sufficiently large n. If $2 \leq |S| < \zeta n$, by Lemma 3.18, with probability at least $1 - e^{-(\gamma \log n)^2/2}$, for sufficiently large n,

$$d^{in}(S) = \sum_{v \in S} d^{in}(v) - |A[S]| \geq \delta^{in}|S| - \phi np|S|$$

$$\geq \delta^{in}|S| - \alpha_1 np|S|/7 \geq 6\delta^{in}|S|/7 \geq 1.5\delta^{in}.$$

If $2 \leq |\bar{S}| \leq \zeta n$, then

$$d^{out}(\bar{S}) = \sum_{v \in \bar{S}} d^{out}(v) - |A[\bar{S}]| \geq \delta^{out}|\bar{S}| - \phi np|\bar{S}| \tag{19}$$

$$> |\bar{S}|np(\alpha_1 - \phi) \geq 1.5np\alpha_2 \geq 1.5\delta^{in} \tag{20}$$

since $2(\alpha_1 - \phi) > 1.5(\alpha_1 + \phi) \geq 1.5\alpha_2$, and the result follows since $d^{in}(S) = d^{out}(\bar{S})$. Observe that it suffices to fix $C < C_2$ to get the desired result. ∎

The next lemma will be useful when we apply Theorem 1.4 to subpartitions with a very large class, namely subpartitions where one part contains a $(1 - \varepsilon)$-fraction of the vertex set.

Lemma 3.20. *There exist positive constants ϕ and ψ such that the following holds. For any function $g = g(n)$ such that $0 \leq g(n) = o(\log n)$, there exist positive constants $\varepsilon > 0$ and $C > 0$ such that, for any $p = p(n) \in [(\log(n) - g(n))/(n - 1), (\log(n) + g(n))/(n - 1)]$ and for large n, with probability at least $1 - n^{-C}$, there is no partition (X, Y, Z) of $[n]$ of the vertex set of $\mathcal{D}(n, p)$ satisfying the following conditions:*

(i) $|X| \geq (1 - \varepsilon)n$,

(ii) $|Y| \geq |Z|$ or $|Z| \leq \phi p(n - 1)$,

(iii) $d^{in}(X) + d^{in}(Y) \leq \psi p(n - 1)|Y|$.

Proof. Let $\phi \leq 3/400$ and $\psi < 3/20$. Let $\zeta > 0$ be obtained by Lemma 3.18 with $f(n) = \log n - g(n)$ and ϕ. Let $\varepsilon \leq \zeta$ and let (X, Y, Z) be a partition of $[n]$ satisfying conditions (i) and (ii) of the lemma. By Lemma 3.3(iv) with $\alpha = 0.09$, we have, for n suficiently large, with probability at least $1 - n^{-0.2}$, that $d^{in}(v) + d^{out}(v) \geq 0.18(n - 1)p$ for all $v \in Y$. Lemma 3.18, we have that $|A[Y]| \leq \phi(n - 1)p|Y|$ and $|A[Y \cup Z]| \leq \phi(n - 1)p(|Y| + |Z|)$ with probability at least $1 - e^{-\log^2 n/4}$. Thus,

$$d^{in}(X) + d^{in}(Y) = \sum_{v \in Y} \left(d^{in}(v) + d^{out}(v)\right) - 2|A[Y]| - |A(Y, Z)| + |A(Z, X)|.$$

$$\geq 0.18(n - 1)p|Y| - 2\phi(n - 1)p|Y| - \min\{|A[Y \cup Z]|, |Y||Z|\}$$

$$\geq 0.18(n - 1)p|Y| - 2\phi(n - 1)p|Y| - 2\phi(n - 1)p|Y|$$

$$\geq 0.15(n - 1)p|Y|,$$

as required. ∎

In the next two lemmas, we bound $d^{in}(S)$ in the range $p \sim \log n/(n - 1)$.

Lemma 3.21. *Let $g = g(n)$ be a function such that $0 \leq g(n) = o(\log n)$. There exist positive constants $\eta > 0$ and $C > 0$ with the following properties. For all functions $(\log n - g(n))/(n-1) \leq p = p(n) \leq (\log n + g(n))/(n-1)$, with probability at least $1 - n^{-C}$, there are at least two vertices with in-degree zero or there is no $S \subseteq [n]$ with size $2 \leq |S| \leq \eta n$ such that $d^{in}(S) < \max\{\delta^{in} + 1, 2\delta^{in}\}$.*

Proof. Let $p = p(n)$ be a function as in the statement of the lemma and let $\varepsilon > 0$ and C_1 be obtained through Lemma 3.8 with $\varphi = 1.1$. By Corollary 3.4 we have $\delta^{in} < (\varepsilon/16) \log n$ with probability at least $1 - n^{C_0}$ for some constant C_0. Let $\eta = \zeta > 0$ given by Lemma 3.18 applied to $f(n) = \log n - g(n)$ and $\phi \leq \varepsilon/16$. Assume that the random digraph has at most one vertex with in-degree zero and fix $S \subseteq [n]$ with size $2 \leq |S| \leq \eta n$. By Lemma 3.18, with probability at least $1 - e^{-\log^2 n/4}$, $|A[S]| \leq \varepsilon n |S| p/16$. Let S_ℓ denote the ε-in-light vertices in S and let $S_h = S \setminus S_\ell$.

First assume that all vertices in S are ε-in-light. By Lemma 3.8, with probability at least $1 - n^{-C_1}$, no pair of ε-in-light vertices are adjacent, and thus $d^{in}(S) \geq |S|\delta^{in} \geq 2\delta^{in} \geq \delta^{in} + 1$ if $\delta^{in} > 0$. If $\delta^{in} = 0$, because there is a single vertex with $d^{in}(v) = 0$, we have $d^{in}(S) \geq |S| - 1 \geq 1 = \max\{2\delta^{in}, \delta^{in} + 1\}$. Next suppose that there is at least one vertex u in S that is not ε-in-light. Note that $d^{in}(u) \geq \delta^{in} + \varepsilon n p$, which implies that, for $|S| < \varepsilon n p/2$, we have $d^{in}(S) \geq |A(\bar{S}, u)| \geq d^{in}(u) - |S| + 1 \geq \delta^{in} + \varepsilon n p/2 \geq 2\delta^{in} + 1$ for large n. So we can assume that $|S| \geq \varepsilon n p/2$.

If $|S_h| \geq |S|/8$, then $d^{in}(S) \geq \sum_{v \in S_h} d^{in}(v) - |A[S]| \geq |S|\varepsilon n p/8 - \varepsilon n p|S|/16 \geq \varepsilon(n-1)p/8 \geq 2\delta^{in} + 1$. So assume that $|S_h| \leq |S|/8$. Thus, $|S_\ell| \geq 7|S|/8$. Then $d^{in}(S) \geq \delta^{in}|S_\ell| - |S_h|$ since no pair of ε-in-light vertices have a common in-neighbour by Lemma 3.8. If $\delta^{in} > 0$, then $d^{in}(S) \geq |S_\ell| - |S_h| \geq 3\varepsilon(n-1)p/8 \geq 2\delta^{in} + 1$. So assume that $\delta^{in} = 0$. Since there is a single vertex with in-degree zero, we have $d^{in}(S) \geq |S_\ell| - 1 - |S_h| \geq \varepsilon(n-1)p/4 \geq 2\delta^{in} + 1$. ∎

Lemma 3.22. *Let $\phi > 0$ be a constant and $g = g(n)$ be a function such that $0 \leq g(n) = o(\log n)$. There exist positive constants $\eta > 0$ and $C > 0$ such that the following holds for all functions $(\log n - g(n))/(n-1) \leq p = p(n) \leq (\log n + g(n))/(n-1)$. With probability at least $1 - n^{-C}$, there exist two vertices with in-degree zero or there is no $S \subseteq [n]$ with size $\phi \log n \leq |S| \leq \eta n$ such that $d^{out}(S) < 2\delta^{in} + 1$.*

Proof. Let $\phi > 0$ be constant. Let $p = (\log(n) + h(n))/(n-1)$ such that $|h(n)| \leq g(n)$ and let $\varepsilon > 0$ and C_1 be obtained through Lemma 3.8 with $\varphi = 1.1$. We may assume that $\phi < \varepsilon/16$. By Corollary 3.4 we have $\delta^{in} < (\phi/16) \log n$ with probability at least $1 - n^{C_0}$ for some constant C_0. Let $\eta = \zeta > 0$ from Lemma 3.18 applied to ϕ and let $\psi \leq \phi/2$.

If $h(n) < -\log(\psi \log n)$, at least two vertices have in-degree zero a.a.s. by Chernoff's inequality (Theorem 3.1), so we consider the case $h(n) \geq -\log(\psi \log n)$.

Assume that at most one vertex has in-degree zero and fix $S \subseteq [n]$ with size $\phi \log n \leq |S| \leq \eta n$. By Lemma 3.18, $|A[S]| \leq \varepsilon n |S| p/16$ with probability at least $1 - e^{-\log^2 n/4}$. Let S_ℓ denote the ε-out-light vertices in S and let $S_h = S \setminus S_\ell$. If $|S_h| \geq |S|/8$, then $d^{out}(S) \geq \sum_{v \in S_h} d^{out}(v) - |A[S]| \geq \varepsilon n p|S|/8 - \varepsilon n p|S|/16 \geq 2\delta^{in} + 1$. So assume that $|S_h| \leq |S|/8$, and thus $|S_\ell| \geq 7|S|/8$. Then $d^{out}(S) \geq \delta^{out}|S_\ell| - |S_h|$ since no pair of ε-out-light vertices have a common out-neighbour. If $\delta^{out} > 0$, then $d^{out}(S) \geq |S_\ell| - |S_h| \geq (3/8)\varepsilon(n-1)p \geq 2\delta^{in} + 1$. We may assume that $\delta^{out} = 0$. We first estimate the number of vertices with out-degree zero. The expected number of vertices with out-degree zero

13

is $n(1-p)^{n-1} = \exp(-h(n) + o(1))$. It is obvious that the smaller is $h(n)$, the higher is the number of vertices with out-degree zero. If $h(n) = -\log(\psi \log n)$, then the expected number of vertices with out-degree zero is equal to $\psi \log n(1 + o(1))$. Thus, by Chernoff's inequality (Theorem 3.1), there exists $C > 0$ such that the probability that there are more than $5\phi \log n/8$ vertices with out-degree zero is at most $\exp(-C \log n)$. Then, $d^{out}(S) \geq |S_\ell| - (5\phi/8) \log n - |S|/8 \geq 3|S|/4 - 5\phi \log n/8 \geq 2\delta^{in} + 1$ for $h(n) \geq -\log(\psi \log n)$. $\blacksquare$

4   Proof of Theorem 1.3

In the proof of Theorem 1.3, we consider four different probability regimens, which are described in the following result.

Lemma 4.1. *Let $p = p(n) \in [0, 1]$. If*

  (i) *$p \leq (\log(n) - g(n))/(n - 1)$, for a function $g(n) = \Omega(\log n)$; or*

  (ii) *$(\log(n) - g(n))/(n - 1) \leq p \leq (\log(n) + g(n))/(n - 1)$, for a function $g(n) = o(\log n)$; or*

  (iii) *$p \sim (1 + \psi) \log(n)/(n - 1)$, for a constant $\psi > 0$; or*

  (iv) *$p = g(n) \log(n)/(n - 1)$, for a function $g(n) = \omega(1)$,*

then $\tau(\mathcal{D}(n, p)) = \lambda(\mathcal{D}(n, p))$ a.a.s.

   We will now show how Theorem 1.3 follows from the lemma above. Let $p = p(n) \in [0, 1]$. Let A_n be the event that $\tau(\mathcal{D}(n, p)) = \lambda(\mathcal{D}(n, p))$ and let $\bar{A}_n$ be its complement. We will show that $\lim_{n \to \infty} \Pr(\bar{A}_n) = 0$. We will use the subsubsequence principle. To this end, let $(n_i)_{i \in \mathbb{N}}$ be an arbitrary increasing sequence where $n_i \in \mathbb{N}$ for all $i \in \mathbb{N}$. It suffices to show that there is a subsequence $(m_i)_{i \in \mathbb{N}}$ of $(n_i)_{i \in \mathbb{N}}$ such that $\lim_{i \to \infty} \Pr(\bar{A}_{m_i}) = 0$.

   Let $h(n) = (n - 1)p/\log n$. Note that $h(n) \geq 0$. Thus, there exists a subsequence $(m_i)_{i \in \mathbb{N}}$ of $(n_i)_{i \in \mathbb{N}}$ such that $\lim_{i \to \infty} h(m_i) = c$, where c is a nonnegative constant or ∞. If $c < 1$, we have $\lim_{i \to \infty} \Pr(\bar{A}_{m_i}) = 0$ by Lemma 4.1(i). If $c = 1$, we have $\lim_{i \to \infty} \Pr(\bar{A}_{m_i}) = 0$ by Lemma 4.1(ii). If $1 < c < \infty$, we have $\lim_{i \to \infty} \Pr(\bar{A}_{m_i}) = 0$ by Lemma 4.1(iii). If $c = \infty$, we have $\lim_{i \to \infty} \Pr(\bar{A}_{m_i}) = 0$ by Lemma 4.1(iv). Thus, we may apply the subsubsequence principle and conclude that $\Pr(\bar{A}_n) = o(1)$.

Proof of Lemma 4.1. Case (i): Suppose that $p \leq (\log(n) - g(n))/(n - 1)$, for a function $g(n) = \omega(1)$. If we have two vertices with in-degree 0, then $\tau = \lambda = 0$. The expected number of vertices with in-degree 0 is $n(1 - p)^{n-1} \geq \exp(g(n) + o(1)) \to \infty$. So, by Chernoff's inequality (Theorem 3.1), there are at least two vertices with in-degree 0 a.a.s.

   For the remaining cases, let $\mathcal{P} = (V_0, \ldots, V_t)$ be a partition of $[n]$. We need to show that $\sum_{i=1}^{t} d^{in}(V_i) \geq \lambda(t - 1)$, since this is equivalent to proving that $\tau(\mathcal{D}(n, p)) \geq \lambda(\mathcal{D}(n, p))$ by Theorem 1.4.

   Case (ii): Fix $g(n)$ given in this case. Let ϕ and ψ of Lemma 3.20, and consider $\varepsilon > 0$ and $C > 0$ given by it. Fix $\eta_1 > 0$ and C_1 as in Lemma 3.21 and let $\eta_2 > 0$ and C_2 be obtained by applying Lemma 3.22 to ϕ. Let $\alpha = \min\{\varepsilon, \eta_1, \eta_2\}$. We may assume that the number of vertices with in-degree zero is at most one, otherwise $\tau = \lambda = 0$ trivially.

Suppose first that there exists $j > 0$ such that $|V_j| \geq (1 - \alpha)n$. Let $B = \bigcup_{i>0,\ i\neq j} V_i$. Note that the result is trivial if $B = \emptyset$, as $t - 1 = 0$ in this case, thus suppose $|B| > 0$. If $|V_0| \leq \phi np$ or $|V_0| \leq |B|$, then $d^{in}(V_j) + d^{in}(B) = \Omega(p(n-1)|B|)$ with probability $1 - n^{-C}$ by Lemma 3.20. Since $\lambda = O(\delta^{in})$ almost surely by Lemmas 3.15 and 3.16 and $\delta^{in} = o(\log n)$ by Corollary 3.4, we have $d^{in}(V_j) + d^{in}(B) = \omega(\lambda|B|)$, so that

$$\sum_{i=1}^{t} d^{in}(V_i) \geq d^{in}(V_j) + d^{in}(B) \geq \lambda|B| \geq \lambda(t-1).$$

So assume that $|V_0| > \phi np$ and $|V_0| > |B| \geq 1$. Let $I = \{i > 0 : |V_i| = 1,\ d^{in}(V_i) \leq \lambda\}$. Observe that $d^{in}(V_j) = d^{out}(V_0 \cup B)$ and by Lemma 3.22 we have $d^{in}(V_j) = d^{out}(V_0 \cup B) \geq 2\delta^{in} + 1$, which is at least λ a.a.s. (see Lemmas 3.15 and 3.16).

Moreover, for every $V_i \in \mathcal{P} \setminus I$ for $i \neq j$, by Lemma 3.21, $d^{in}(V_i) \geq \max\{\delta^{in} + 1, 2\delta^{in}\}$, which is at least λ a.a.s. Let $V(I) = \bigcup_{i \in I} V_i$. Then we have

$$\sum_{i=1}^{t} d^{in}(V_i) \geq \lambda(t - |I|) + \sum_{i \in I} d^{in}(V_i) = t\lambda - \sum_{v \in V(I)} (\lambda - d^{in}(v)) \geq t\lambda - \lambda,$$

by the definition of λ.

Next suppose that $|V_i| \leq (1 - \alpha)n$ for all $i > 0$. If $2 \leq |V_i| \leq \alpha n$, then $d^{in}(V_i) \geq \max\{\delta^{in} + 1, 2\delta^{in}\}$ by Lemma 3.21 and $d^{in}(V_i) \geq \lambda$ a.a.s. by Lemmas 3.15 and 3.16. If $|V_i| \geq \alpha n$ for some i, since we have $|\overline{V}_i| \geq \alpha n$ and so, by Lemma 3.17 with $\zeta = \alpha$, $d^{in}(V_i) = |A(\overline{V}_i, V_i)| \geq \alpha^2 n^2 p / 2 \geq \lambda$. Thus, a.a.s.

$$\sum_{i=1}^{t} d^{in}(V_i) \geq \lambda(t - |I|) + \sum_{i \in I} d^{in}(V_i) \geq t\lambda - \lambda,$$

again by the definition of λ.

Case (iii): By Lemma 3.19, every set $S \subseteq [n]$ of size in $[2, n-2]$ has $d^{in}(S) \geq 1.5\delta^{in}$ a.a.s. By Lemma 3.13 we have $\lambda \sim \delta^{in}$ a.a.s., which implies that $d^{in}(S) \geq \lambda$. Let $I = \{i > 0 : |V_i| = 1,\ d^{in}(V_i) \leq \lambda - 1\}$. If $|V_i| \leq n - 2$ for all i, then a.a.s. $d^{in}(V_i) \geq \lambda$ for all $i \notin I$ and so

$$\sum_{i=1}^{t} d^{in}(V_i) \geq \lambda(t - |I|) + \sum_{i \in I} d^{in}(V_i) \geq t\lambda - \lambda,$$

by the definition of λ.

Now suppose, without loss of generality, that $|V_1| = n - 1$ (the case $|V_1| = n$ is trivial). Then there is a single vertex $v \notin V_1$. We can assume that $t = 2$ since the case $t = 1$ is trivial. Thus, $V_2 = \{v\}$ and $d^{in}(V_1) + d^{in}(V_2) = d^{in}(v) + d^{out}(v) \geq \delta^{in} + \delta^{out}$. By Corollary 3.6 we have $\delta^{in} + \delta^{out} \geq 1.5\delta^{in}$ a.a.s. By Lemma 3.13, we conclude that $d^{in}(V_1) + d^{in}(V_2) \geq 1.5\delta^{in} \geq \lambda$ a.a.s.

Case (iv): We may proceed as in the previous case, since, for every $S \subseteq [n]$ of size in $[2, n-2]$, we again have $\delta^{in} \sim \lambda$ and $d^{in}(S) \geq 1.5\delta^{in} \geq \lambda$ a.a.s. (since $\delta^{in} \sim (n-1)p$ a.a.s. by Lemma 3.9 and $\lambda \sim (n-1)p$ a.a.s. by Corollary 3.12). This leads to the desired result with the above arguments if $|V_i| \leq n - 2$ for every i. Otherwise, we use Lemma 3.9 to show that $\delta^{in} + \delta^{out} \geq 1.5\delta^{in}$ a.a.s., and we may again repeat the analysis of the previous case. ∎

References

[1] Noga Alon and Joel H. Spencer, *The probabilistic method*, second ed., Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience [John Wiley & Sons], New York, 2000, With an appendix on the life and work of Paul Erdős.

[2] Deepak Bal, Patrick Bennett, Colin Cooper, Alan Frieze, and Paweł Prałat, *Rainbow arborescence in random digraphs*, Journal of Graph Theory **83** (2016), no. 3, 251–265.

[3] Sonny Ben-Shimon, Michael Krivelevich, and Benny Sudakov, *On the resilience of hamiltonicity and optimal packing of hamilton cycles in random graphs*, SIAM Journal on Discrete Mathematics **25** (2011), no. 3, 1176–1193.

[4] Béla Bollobás, *The evolution of random graphs*, Trans. Amer. Math. Soc. **286** (1984), no. 1, 257–274.

[5] Béla Bollobás and Alan M. Frieze, *On matchings and Hamiltonian cycles in random graphs*, Random graphs '83 (Poznań, 1983), North-Holland Math. Stud., vol. 118, North-Holland, Amsterdam, 1985, pp. 23–46.

[6] Gérard Cornuéjols, *Combinatorial optimization: Packing and covering*, CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics, 2001.

[7] Asaf Ferber, Matthew Kwan, and Benny Sudakov, *Counting Hamilton cycles in sparse random directed graphs*, ArXiv e-prints (2017).

[8] Asaf Ferber and Eoin Long, *Packing and counting arbitrary hamilton cycles in random digraphs*, Random Structures Algorithms, to appear.

[9] Asaf Ferber, Rajko Nenadov, Andreas Noever, Ueli Peter, and Nemanja Škorić, *Robust hamiltonicity of random directed graphs*, Journal of Combinatorial Theory, Series B **126** (2017), 1 – 23.

[10] András Frank, *Covering branchings*, Acta Sci. Math. (Szeged) **41** (1979), no. 1-2, 77–81.

[11] Alan Frieze and Michael Krivelevich, *On two Hamilton cycle problems in random graphs*, Israel J. Math. **166** (2008), 221–234.

[12] Pu Gao, Xavier Pérez-Giménez, and Cristiane M. Sato, *Arboricity and spanning-tree packing in random graphs*, Random Structures Algorithms, to appear.

[13] _____, *Arboricity and spanning-tree packing in random graphs with an application to load balancing*, ArXiv e-prints (2013).

[14] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński, *Random graphs*, Wiley-Interscience, New York, 2000.

[15] Fiachra Knox, Daniela Kühn, and Deryk Osthus, *Edge-disjoint Hamilton cycles in random graphs*, Random Structures Algorithms **46** (2015), no. 3, 397–445.

[16] Michael Krivelevich and Wojciech Samotij, *Optimal packings of Hamilton cycles in sparse random graphs*, SIAM J. Discrete Math. **26** (2012), no. 3, 964–982.

[17] Daniela Kühn and Deryk Osthus, *Hamilton decompositions of regular expanders: applications*, J. Combin. Theory Ser. B **104** (2014), 1–27.

[18] Wolfgang Mader, *A reduction method for edge-connectivity in graphs*, Ann. Discrete Math. **3** (1978), 145–164, Advances in graph theory (Cambridge Combinatorial Conf., Trinity College, Cambridge, 1977).

[19] Karl Menger, *Zur allgemeinen kurventheorie*, Fundamenta Mathematicae **10** (1927), no. 1, 96–115 (ger).

[20] Crispin St. John Alvah Nash-Williams, *Edge-disjoint spanning trees of finite graphs*, J. London Math. Soc. **36** (1961), 445–450.

[21] William Thomas Tutte, *On the problem of decomposing a graph into n connected factors*, J. London Math. Soc. **36** (1961), 221–230.