*Proceedings of the Edinburgh Mathematical Society* (2018) **61**, 847–868
doi:10.1017/S0013091517000487

# GENERATING INFINITE RANDOM GRAPHS

CSABA BIRÓ AND UDAYAN B. DARJI*
*Department of Mathematics, University of Louisville, Louisville, KY 40292, USA*
(csaba.biro@louisville.edu; ubdarj01@louisville.edu)

(Received 24 May 2017; first published online 22 May 2018)

*Abstract* We define a growing model of random graphs. Given a sequence of non-negative integers $\{d_n\}_{n=0}^{\infty}$ with the property that $d_i \leq i$, we construct a random graph on countably infinitely many vertices $v_0, v_1 \ldots$ by the following process: vertex $v_i$ is connected to a subset of $\{v_0, \ldots, v_{i-1}\}$ of cardinality $d_i$ chosen uniformly at random. We study the resulting probability space. In particular, we give a new characterization of random graphs, and we also give probabilistic methods for constructing infinite random trees.

*Keywords:* Erdős–Rényi graph; random graph; infinite graph; trees; homogeneous structure

2010 *Mathematics subject classification:* Primary 05C63; 05C80
Secondary 05C05; 60C99

## 1. Introduction

Consider the vertex set $\mathbb{N}$. Let $0 < p < 1$ be fixed. For each pair of distinct integers $n, m \in \mathbb{N}$, put an edge between $n$ and $m$ with probability $p$. Let $G$ be the resulting graph on $\mathbb{N}$. A classical 1963 Erdős–Rényi theorem [**10**] states that with probability one, any two such graphs are isomorphic, i.e., there is essentially one random graph on $\mathbb{N}$.

In 1964, Rado [**15**] gave an explicit construction of a graph $R$ which is universal for the collection of all countable graphs. More precisely, he showed that if $G$ and $H$ are any countable graphs and $\phi : G \to H$ a graph homomorphism, then there are embeddings $e_G : G \to R$, $e_H : H \to R$ and a graph homomorphism $\psi : R \to R$ such that $e_H^{-1} \circ \psi \circ e_G = \phi$, i.e., $R$ contains a copy of every countable graph, and every graph homomorphism between countable graphs can be lifted to a graph homomorphism of $R$.

The constructions of Erdős–Rényi and Rado seem very different, but they result in the same graph. The reason for this is that both graphs satisfy the following property: if $(A, B)$ are disjoint, finite sets of vertices, then there are infinitely many vertices $v$ such that there is an edge between $v$ and every element of $A$ and there are no edges between $v$ and any element of $B$. It can be shown by the back and forth method that any two

* Corresponding author.

 

graphs with the above property are isomorphic to each other. A graph with this property is often called the Erdős–Rényi graph, the Rado graph or simply the random graph.

Although the Rado graph is unique and the above definition is very simple, the Rado graph has a rich structure and enjoys attention from mathematicians working in various camps. For example, Cameron [7] gave a number theoretic description of this graph similar to that of the Paley graph. This graph also enjoys attention from model theorists as it is an example of an $\aleph_0$-categorical Fraïssé limit of finite structures. Truss [17] initiated the group theoretic study of the group of automorphisms of the Rado graph. We refer the reader to the survey paper of Cameron [7] for further interesting properties of the random graph along this direction. Connections between percolation theory and the random graphs can be found in the survey paper of van der Hofstad [18]. That the Rado graph can be topologically 2-generated with a great deal of flexibility was shown by the second author and Mitchell [8].

Inspired by the construction of Erdős and Rényi, we introduced a procedure that is flexible enough to generate a large class of infinite graphs, and essentially generalizes the Erdős–Rényi process. We add only finitely many incident vertices to each vertex, determined by a given sequence.

More rigorously, suppose a sequence of integers $\{d_i\}_{i=0}^{\infty}$ is given, with the property $0 \le d_i \le i$ for all $i$. Let $V = \{v_0, v_1, \ldots\}$ be a set of vertices. For $i = 0, 1, \ldots,$ in round $i$, we first choose $A \subseteq \{v_0, \ldots, v_{i-1}\}$ of cardinality $d_i$, with the uniform distribution on the set of subsets of $\{v_0, \ldots, v_{i-1}\}$ that are of size $d_i$. Then, we add edges $v_i u$ for all $u \in A$. The result is a random graph on countably many vertices. We strive to understand the resulting probability space, in particular, we would like to determine the atoms (graphs with positive probability), and cases when there is only one atom with probability 1. In this latter case, we say that the probability space is *concentrated*.

### 1.1. Related literature

#### 1.1.1. *Preferential attachment models*

In the preferential attachment models, the new vertex is adjacent to an earlier vertex or vertices with a probability that depends on the current degree of an existing vertex. One of the first examples of empirical study of this model is by Barabási and Albert [1], and a rigorous mathematical framework was defined by Bollobás and Riordan [2]. These, and subsequent works, study large finite graphs as opposed to the limiting behaviour.

An infinite version of the preferential attachment model (for multigraphs) was studied by Kleinberg and Kleinberg [12]. In their paper, the sequence $d_i$ is constant. Since the preferential attachment model is substantially different from ours, they get very different results, but some of the techniques they use are similar to ours.

#### 1.1.2. *Copying models*

This model was first introduced by Kumar *et al.* [13], and later, a slightly modified and generalized version was defined by Bonato and Janssen [4]. In their construction, besides the sequence $d_i$, an initial finite graph $H$ and a probability $p \in [0, 1]$ are given.

- Let $G_0 = H$.
- To construct $G_i$, add a new vertex $v$ to $G_{i-1}$ and choose its neighbours as follows.
  - Choose a vertex $u \in V(G_{i-1})$ uniformly at random (called the *copy vertex*). Connect $v$ to each neighbour of $u$ with probability $p$.
  - Choose a set of $d_i$ vertices from $V(G_{i-1})$ uniformly at random, and connect $v$ to each vertex in this set.
  - Delete multiple edges if necessary.

Clearly, our process is a special case of this, when $p = 0$.

In [**4**], the authors only study the case when $d_i = \Theta(i^s)$ for some $s \in [0, 1)$. Although we study very similar models, the common special case of our theorems is quite narrow: we imposed the extra condition that $p = 0$, and they imposed strong extra conditions on $d_i$. Nevertheless, it is interesting to note that for the narrow special case when our assumptions coincide, our Theorem 2.3 implies the conclusions of their main theorems (Theorems 2.2 and 2.3 in [**4**]), and more.

#### 1.1.3. *The Janson–Severini process*

Janson and Severini [**11**] introduced a process that also includes ours. Their construction is the following. For all $i = 1, \ldots,$ let $\nu_i$ be a probability distribution on $\{0, 1, \ldots, i\}$. Construct the random graph $G_i$ as follows.

- Let $G_0 = K_1$, the graph on a single vertex.
- Let $D_i$ be a random variable with distribution $\nu_i$, and construct $G_i$ by adding a new vertex to $G_{i-1}$ and connecting it to a uniformly random subset of size $D_i$ of $V(G_{i-1})$.

Of course our model is the special case of theirs when $\nu_i$ is a point mass at $d_i$. In fact, as an application of our theorems, we venture to prove certain limiting behaviour in their general model (which we call the 'double random process') in Corollary 4.4, and Theorem 6.12. However, unlike us, they study the graphons, as limits of their sequence. (Graphons were introduced by Lovász and Szegedy [**14**] and Borgs *et al.* [**5**].) In their main theorem, they determine the limit graphon when $D_n/n \xrightarrow{\text{p}} \nu$ for some probability measure $\nu$ on $[0, 1]$.

## 2. Summary and outline

In § 3 we discuss some minor results. We quickly show how different this model is from the Erdős–Rényi model in that it can easily result in non-concentrated spaces.

The main discussion starts in § 4. The paper contains two major results. In § 4, we prove the first one (stated in this section as Theorem 2.3), which was motivated by the effort of characterizing the sequences that will almost surely result in the Rado graph. We did more than that; we defined a degree of similarity of a graph to the Rado graph, and we can determine from the sequence how similar the resulting graph will be to the Rado graph.

**Definition 2.1.** Let $G$ be a graph and $A, B \subseteq V(G)$. We say that a vertex $v$ is a *witness* for the ordered pair $(A, B)$ if $v$ is adjacent to every vertex in $A$, and $v$ is not adjacent to any vertex in $B$.

**Definition 2.2.** Let $G$ be a graph. For a non-negative integer $k$, we say that $G$ is *$k$-Rado* if every pair of disjoint sets of vertices $(A, B)$ with $|A| \leq k, |B| \leq k$ has infinitely many witnesses.

The number $\mathrm{rado}(G) = \sup\{k : G \text{ is } k\text{-Rado}\}$ is the *radocity* of $G$.

Clearly, every graph is 0-Rado, and if a graph is $k$-Rado, it is also $k'$-Rado for all $k' < k$. Also, by the Erdős–Rényi Theorem, $G$ is isomorphic to the Rado graph if and only if $\mathrm{rado}(G) = \infty$.

We note that the definition of a witness is not new. Clearly, Erdős and Rényi knew about the property, and the same language is used by Spencer in the book [16]. Similar properties for a graph to be $k$-Rado also appeared in the literature. Still, in Spencer's book, the property $A_{r,s}$ is defined as follows: a graph satisfies the property $A_{r,s}$ if every pair of disjoint sets of vertices $(A, B)$ with $|A| = r$, $|B| = s$ has a witness. Note the major difference that $A_{r,s}$ requires only one witness, while $k$-Rado requires infinitely many witnesses, so, e.g., a double ray has $A_{1,1}$, but it is not 1-Rado.

Another similar property is called $n$-e.c. ($n$-existentially closed; see, e.g., [3]). A graph has this property if every pair of disjoint sets of vertices $(A, B)$ with $|A \cup B| = n$ has a witness. So a graph has $n$-e.c. if and only if it has $A_{r,s}$ for all $r + s = n$.

Finally, Winkler used the colourful term *Alice's Restaurant property* for a graph that is $k$-Rado for all $k \geq 0$, in other words, the radocity of the graph is $\infty$. As mentioned above, this happens if and only if the graph is the Rado graph.

In § 4 we will prove the following theorem. It shows that the radocity of the graph is determined by the sequence, not by the random process. In the statement, and throughout this paper, we will use the standard notation $n_{(k)} = n(n-1)\dots(n-k+1)$ with $n_{(0)} = 1$ (even if $n = 0$). In addition, we define $0^0 = 1$ if this power appears as a term of a series.

**Theorem 2.3.** *As before, let $\{d_i\}$ be such that $0 \leq d_i \leq i$. Let*

$$k_1 = \sup\left\{t \in \mathbb{N} : \sum_{n=1}^{\infty} \left(\frac{d_n}{n}\right)^t \left(\frac{n-d_n}{n}\right)^t = \infty\right\},$$

$$k_2 = \sup\left\{t \in \mathbb{N} : \sum_{n=1}^{\infty} \frac{(d_n)_{(t)}(n-d_n)_{(t)}}{(n)_{(2t)}} = \infty\right\}.$$

*Then $k_1 = k_2$, and the process almost surely generates a graph of radocity $k_1$ (and $k_2$).*

As a corollary, we achieve our original motivation.

**Corollary 2.4.** *Let $a_n = \min\{d_n/n, n - d_n/n\}$.*

(i) *If $\sum_{n=1}^{\infty} a_n^k$ diverges for all positive integers $k$, then the process almost surely generates the Rado graph.*

(ii) *If there is a positive integer $k$ for which $\sum_{n=1}^{\infty} a_n^k$ converges, then the process almost surely does not generate the Rado graph.*

This also shows that our result is essentially a generalization of the result of Erdős and Rényi. See § 4 for more details.

### 2.1. Examples

In the following examples, to avoid clutter, we will omit floor and ceiling signs.

- If $d_n = n/2$, then $\mathrm{rado}(G) = \infty$.
- If $0 < c < 1$, and $d_n = cn$, then $\mathrm{rado}(G) = \infty$.
- If $d_n = \sqrt{n}$, then $\mathrm{rado}(G) = 2$.
- If $d_n > 0$ is constant, then $\mathrm{rado}(G) = 1$.
- If $k \geq 1$ integer, and $d_n = n^{(k-1)/k}$, then $\mathrm{rado}(G) = k$.
- If $d_n = \log n$, then $\mathrm{rado}(G) = 1$.

In § 6 we focus on 0–1 sequences. From the discussion above, it is clear that the resulting graphs will almost surely have radocity 0 or 1, but we aim to describe the random graph in more detail.

Recall that a probability space is *concentrated* if there exists a graph $G$ such that the process generates a graph isomorphic to $G$ with probability 1. A graph $G$ is an *atom* of the space if the process generates a graph isomorphic to $G$ with positive probability.

To state a compact theorem, we introduce some elaborate notation to denote certain infinite graphs. Let $T$ be a finite tree. Let $F_T$ be the forest that consists of infinitely many copies of $T$, as components. Let $F_n = \bigcup\{F_T : T \text{ is a tree of size } n\}$. Note that $F_1$ is the countably infinite set with no edges, and $F_2$ is the countably infinite matching.

We will also use the term *ω-tree* for the unique countably infinite tree in which every vertex is of infinite degree.

**Theorem 2.5.** *Suppose $d_n \in \{0, 1\}$ for all $n \in \mathbb{N}$.*

(i) *If $\sum_{i=1}^{\infty} d_i/i = \infty$, then the space is concentrated, the atom is a graph whose components are ω-trees, and the number of components is equal to the number of zeroes in the sequence.*

(ii) *Suppose $\sum_{i=1}^{\infty} d_i/i < \infty$. Let $t_n = \sum_{i=n}^{\infty} d_i/i$, and $k = \min\{\kappa \geq 2 : \sum_l d_l t_{l+1}^{\kappa-2} < \infty\}$. (We set $k = \infty$ if the set in question is empty.) The space has infinitely many atoms, and all of them are of the form $F \cup [\bigcup_{i<k} F_i]$ where $F$ is some finite forest.*

Even though this theorem is not a complete description of the probability space, it describes completely what the atoms are. The distinction of the sequences in part (ii) is extremely subtle, and the proof is very elaborate. Nevertheless, we strived for clarity, and we divided the whole proof into small lemmas, so by the time we are ready to prove the theorem, we can use the machinery that will have been built up.

This theorem is the other major result of the paper, and arguably the more difficult one.

## 3. Non-concentrated spaces

It would perhaps be not completely naïve to think that something similar happens here as in the Erdős–Rényi model. In this section, we demonstrate that this is far from being correct. Therefore, we will show examples of non-concentrated spaces.

The following proposition is actually about a very simple example of *concentration*, but we will use it as a tool to show non-concentration in some other cases.

**Proposition 3.1.** *The sequence* $0, 1, 1, 1, \ldots$ *almost surely generates the* $\omega$*-tree.*

**Proof.** We will prove a more general statement later, see Theorem 2.5. □

**Corollary 3.2.** *Consider a sequence of the form* $d_0, d_1, \ldots, d_k, 1, 1, 1, \ldots$. *Let* $G_1, \ldots, G_l$ *be the set of finite non-isomorphic graphs on* $v_0, \ldots, v_k$ *that can be generated by the process using* $d_1, \ldots, d_k$. *For each* $i$, *let* $G'_i$ *be the graph constructed from* $G_i$ *by attaching an* $\omega$*-tree to every vertex. Then the graphs* $G'_1, \ldots, G'_l$ *are the atoms of the space, with probabilities inherited from the finite part of the process.*

The corollary above shows that it is easy to construct a sequence whose associated probability space is not concentrated, e.g., 0, 1, 2, 1, 2, 1, 1, 1, .... However, these examples are very special in the sense that they are eventually all 0s and 1s, so after that point no more cycles are generated. Nevertheless, the following proposition shows that non-concentrated probability spaces can be found for other kind of sequences.

**Proposition 3.3.** *There exists a sequence* $\{d_i\}$ *with a non-concentrated probability space such that for all positive integers* $N$ *there exists* $n > N$ *such that* $a_n \neq 0$ *and* $a_n \neq 1$.

**Proof.** We will construct a sequence consisting mostly of 1s, but with infinitely many 2s inserted. The sequence starts with $0, 1, 1, 2$. We set $p_0 = 2/3$, and we note that $p_0$ is the probability that the first four vertices include a triangle. Then let $k$ be the least integer such that $k/\binom{k}{2} < 3/4 - p_0$. Set $d_4 = \cdots = d_{k-1} = 1$, and $d_k = 2$. Note that the probability that a triangle is generated by $v_k$ is $p_1 := k/\binom{k}{2}$. In general, after the $l$th 2 in the sequence, let $k$ be a sufficiently large integer for which $d_k$ is not yet defined and

$$\frac{k+l-1}{\binom{k}{2}} < \frac{3}{4} - \sum_{i=0}^{l-1} p_i.$$

Set $d_k = 2$ and set all the elements before $d_k$ that are not yet defined to be 1. Note that the probability that a triangle is generated at $v_k$ equals $p_l = k + l - 1/\binom{k}{2}$. Let $X$ be the random variable that denotes the number of triangles eventually generated in $G$. Owing to the linearity of expectation,

$$\mu = E[X] = \sum_{i=0}^{\infty} p_i.$$

Clearly, from the definition of the sequence, $2/3 \leq \mu \leq 3/4$. That means that

$$\Pr[X = 0] > 0 \quad \text{and} \quad \Pr[X > 0] > 0.$$

The sets $[X = 0]$ and $[X > 0]$ partition the probability space, and neither of them are of measure 0, so the space can not be concentrated. $\square$

## 4. The Rado graph

This section contains the proof of Theorem 2.3, and Corollary 2.4, with additional discussion of some consequences. We will make frequent use of the following basic fact relating infinite products to infinite sums.

**Proposition 4.1.** *Let $\{b_i\}_{i=0}^{\infty}$ be a sequence of real numbers such that $0 < b_i < 1$ and $\{d_i\}$ be a sequence of non-negative integers. Then,*

$$0 < \prod_{i=1}^{\infty}(1-b_i)^{d_i} \iff \sum_{i=1}^{\infty} d_i b_i < \infty.$$

We start with a simple technical lemma.

**Lemma 4.2.** *Fix a non-negative integer $k$. The infinite series*

$$\sum_{n}^{\infty}\left(\frac{d_n}{n}\right)^k\left(\frac{n-d_n}{n}\right)^k \quad \text{and} \quad \sum_{n}^{\infty}\frac{(d_n)_{(k)}(n-d_n)_{(k)}}{(n)_{(2k)}}$$

*either both converge or both diverge.*

**Proof.** If $k \leq 1$, then the statement is trivial. If $k \geq 2$, then partition the terms into three parts: $A = \{i : d_i < k\}$, $B = \{i : n - d_i < k\}$, and $C = \mathbb{N} \setminus (A \cup B)$. It is clear that over the terms indexed by $A$ and $B$, both series converge, so the behaviour is decided by the terms over $C$. For those, we use a generalized limit comparison test and show that the lim inf and lim sup of the ratio of the terms are positive and finite.

To see this last statement, notice that

$$1 \leq \frac{d_n}{d_n}, \frac{d_n}{d_{n-1}}, \ldots, \frac{d_n}{d_n - k + 1} \leq k$$

so

$$1 \leq \frac{(d_n)^k}{(d_n)_{(k)}} \leq k^k.$$

A similar statement can be made about $(n - d_n)^k/(n - d_n)_{(k)}$, so we see that the lim inf of the ratio of the terms is at least 1, and the lim sup is at most $k^{2k}$. $\square$

### 4.1. Proof of Theorem 2.3

Note that $k_1 = k_2$ is a consequence of Lemma 4.2. We will denote this number by $k$, and we will go back and forth between its two equivalent definitions at our convenience.

Now we prove that the graph generated is almost surely $k$-Rado.

The statement is trivial for $k = 0$. Let $A, B$ be two finite disjoint vertex sets with $|A| = |B| = k \geq 1$, and let $N$ be a positive integer. It is sufficient to show that the pair $(A, B)$ has a witness with probability 1 among the vertices $v_N, v_{N+1}, \ldots$.

https://doi.org/10.1017/S0013091517000487 Published online by Cambridge University Press

For a given vertex $v_n$, let $p_n$ be the probability that $v_n$ is a witness for $(A, B)$. Now pick a vertex $v_n$ such that $n > \max\{i : v_i \in A \cup B\}$ and $n \geq N$. Then

$$p_n = \frac{\binom{n-2k}{d_n-k}}{\binom{n}{d_n}} = \frac{(d_n)_{(k)}(n-d_n)_{(k)}}{(n)_{(2k)}}.$$

Note that this holds whether $d_n \geq k$ or $d_n < k$; in the latter case, $p_n = 0$. Hence $\sum_{n=N}^{\infty} p_n$ diverges, and then $\prod_{n=N}^{\infty}(1-p_n) = 0$, which is the probability that the pair $(A, B)$ has no witness beyond (including) $v_N$.

It remains to be proven that if $k < \infty$, then the graph is almost surely not $k+1$-Rado. It suffices to prove that there is a pair $(A, B)$ of finite disjoint vertex sets with $|A| = |B| = k+1$ such that $(A, B)$ almost surely has finitely many witnesses. Indeed, we prove that this is the case for every such pair of vertex sets $(A, B)$. To obtain a contradiction, suppose that this is not true: that is, there are disjoint sets $A, B$ of vertices with $|A| = |B| = k+1$, and the probability that $(A, B)$ has finitely many witnesses is $p < 1$. Let $q_N$ be the probability that $(A, B)$ has no witness beyond (including) $v_N$. We note that

$$q_N \leq p \quad \text{for all } N. \tag{1}$$

On the other hand, similarly, the probability that a given vertex $v_n$ is a witness for $(A, B)$ (if $n$ is large enough) is

$$p_n = \frac{(d_n)_{(k+1)}(n-d_n)_{(k+1)}}{(n)_{(2(k+1))}}.$$

This time, we know that $\sum p_n < \infty$, so $\prod(1-p_n) > 0$. Hence there exists $N$ such that

$$q_N = \prod_{n=N}^{\infty}(1-p_n) > p.$$

But this contradicts (1).

### 4.2. Proof of Corollary 2.4

Suppose that $\sum_{n=1}^{\infty} a_n^k$ diverges for all $k$. Since

$$\sum_{n=1}^{\infty}\left(\frac{d_n}{n}\right)^k\left(\frac{n-d_n}{n}\right)^k \geq \sum_{n=1}^{\infty} a_n^{2k},$$

we get that

$$\sum\left(\frac{d_n}{n}\right)^k\left(\frac{n-d_n}{n}\right)^k$$

diverges for all $k$, and therefore we get that almost surely $\mathrm{rado}(G) = \infty$.

Now suppose that there is a positive integer $k$ for which $\sum_{n=1}^{\infty} a_n^k$ converges. Since

$$a_n^k \geq \left(\frac{d_n}{n}\right)^k \left(\frac{n-d_n}{n}\right)^k \geq \prod_{i=0}^{k-1} \frac{d_n - i}{n} \cdot \frac{n-d_n-i}{n},$$

we have that for large enough $n_0$,

$$\sum_{n=n_0}^{\infty} a_n^k \cdot 2^{2k} \geq \sum_{n=n_0}^{\infty} \left( \prod_{i=0}^{k-1} \frac{d_n - i}{n} \cdot \frac{n-d_n-i}{n} \prod_{i=0}^{2k-1} \frac{n}{n-i} \right) = \sum_{n=n_0}^{\infty} \frac{(d_n)_{(k)}(n-d_n)_{(k)}}{(n)_{(2k)}},$$

and therefore the last sum converges. Thus the graph almost surely has finite radocity.

**Corollary 4.3.** *Let* $a_n = \min\{d_n/n, n - d_n/n\}$*. If* $\limsup a_n > 0$*, then the process almost surely generates the Rado graph.*

**Proof.** Direct consequence of Corollary 2.4. □

The *double random process* is when we choose even the sequence at random, choosing $d_i$ with some distribution from the interval $[0, i]$. Note that Janson and Severini [**11**] study the double random process from a different point of view. The following corollary states that, in some sense, almost all double random processes will result in the Rado graph.

**Corollary 4.4.** *If there exist* $\epsilon > 0$*,* $p_0 > 0$*, and* $M$ *integer such that for* $n > M$*,* $\Pr[\epsilon n \leq d_n \leq (1-\epsilon)n] \geq p_0$*, then the double random process almost surely generates the Rado graph.*

**Proof.** It is easy to see that Corollary 4.3 is almost surely satisfied. □

## 5. Density, sparsity, degrees and stars

The main goal of this section is to analyse how certain 'density' conditions on the sequence will affect the resulting graph. One important result from this section (Theorem 5.3) will also be used in § 6 to analyse zero–one sequences.

For the rest of the section, we will use the notation $s_n = \sum_{i=0}^{n} d_i$, the partial sum of the sequence $\{d_i\}$.

It will be useful to distinguish sequences based on convergence of certain partial sums. When $\sum d_i/i = \infty$, we will refer to this situation as the 'dense' case. The opposite case, when $\sum d_i/i < \infty$, will be called the 'sparse' case. A subcase of the sparse case, when even $\sum s_i d_i/i < \infty$, will be called the 'very sparse case'.

We begin with a simple proposition on binomial coefficients.

**Proposition 5.1.** *Let* $n, d, m \geq 0$ *integers with* $m/n - d \leq 1$*. Then*

$$\left(1 - \frac{m}{n-d}\right)^d \leq \frac{\binom{n-m}{d}}{\binom{n}{d}} \leq \left(1 - \frac{m}{n}\right)^d.$$

https://doi.org/10.1017/S0013091517000487 Published online by Cambridge University Press

**Proof.** We note that

$$\frac{\binom{n-m}{d}}{\binom{n}{d}} = \frac{(n-m)_{(d)}}{(n)_{(d)}} = \prod_{i=0}^{d-1}\left(1-\frac{m}{n-i}\right).$$

Then bound the product by replacing all factors with the largest factor, and then with the smallest factor to obtain the desired inequality. □

**Lemma 5.2.** *Let $v_k$ be a vertex.*

(i) *If $\sum d_i/i = \infty$, then for all $N > k$, $v_k$ almost surely has a neighbour beyond $v_N$.*
(ii) *If $\sum d_i/i < \infty$, then there exists $M$ such that with positive probability $v_k$ has no neighbour beyond $v_M$; furthermore, for all $\epsilon > 0$ there exists an $M' \geq M$ such that $\Pr(v_k$ has a neighbour beyond $v_{M'}) < \epsilon$.*

**Proof.** Let $E_l$ be the event that $v_k$ has no neighbour beyond $v_l$. We will estimate the probability of $E_l$. For any $i > k$, we have

$$\Pr(v_i \not\sim v_k) = \binom{i-1}{d_i} \Big/ \binom{i}{d_i} \quad \text{so } \Pr(E_l) = \prod_{i=l}^{\infty} \binom{i-1}{d_i} \Big/ \binom{i}{d_i}.$$

Using Proposition 5.1, we have

$$\Pr(E_l) \leq \prod_{i=l}^{\infty}\left(1-\frac{1}{i}\right)^{d_i}.$$

If $\sum d_i/i = \infty$, then the product on the right-hand side is zero. Thus, we have that $\Pr(E_l) = 0$ for all $l > i_k$, which proves the first part.

If $\sum d_i/i < \infty$, then there exists an $M$ such that for all $i \geq M$, $d_i/i \leq 1/2$ and $1/(i - d_i) \leq 1$. Then we may use the other part of Proposition 5.1 to get

$$\Pr(E_M) \geq \prod_{i=M}^{\infty}\left(1-\frac{1}{i-d_i}\right)^{d_i}. \tag{2}$$

Also, in this case,

$$\sum_{i=M}^{\infty}\frac{d_i}{i-d_i} = \sum_{i=M}^{\infty}\frac{1}{1-d_i/i}\cdot\frac{d_i}{i} \leq \sum_{i=M}^{\infty} 2\cdot\frac{d_i}{i} < \infty,$$

so $\Pr(E_M) > 0$.

The last statement follows from the fact that the right-hand side in (2) is positive, therefore its tail end converges to 1, so for all $\epsilon > 0$ there exists $M'$ for which $\Pr(E_{M'}) > 1 - \epsilon$. □

**Theorem 5.3 (density and degrees).** *The following statements hold.*

(i) *If $\sum d_i/i = \infty$, then the process almost surely generates a graph in which each vertex is of infinite degree.*

https://doi.org/10.1017/S0013091517000487 Published online by Cambridge University Press

(ii) *If* $\sum d_i/i < \infty$*, then the process almost surely generates a graph in which each vertex is of finite degree.*

**Proof.** Both parts follow from Lemma 5.2. Let $v_k$ be a vertex and let $N > k$ be an integer. In case (i), Lemma 5.2 implies that almost surely $v_k$ has a neighbour beyond $v_N$. This being true for arbitrary $N > k$, we conclude that almost surely $v_k$ has infinitely many neighbours.

In case (ii), the lemma provides that the probability that $v_k$ is of infinite degree is less than $\epsilon$ for all $\epsilon > 0$, and therefore that probability is 0. □

Recall that the bipartite graphs $K_{1,l}$ for $l = 0, 1, 2, \ldots$ are called *stars*. (For convenience, we allow $l = 0$. In this case, $K_{1,l}$ is simply a singleton set.) To emphasize the size of the star, $K_{1,l}$ will often be called an $l$-star. We say that a vertex in a graph is *in a star,* respectively *in an* $l$*-star*, if the connected component of the vertex is a star, respectively an $l$-star.

**Lemma 5.4.** *Suppose* $\sum s_i d_i/i < \infty$*. Then the process almost surely generates a graph* $G$ *which has the property that there exists an* $N_1 = N_1(G)$ *such that for all* $n \geq N_1$ *with* $d_n > 0$*, the vertex* $v_n$ *will attach back to vertices with current degree* 0*. More rigorously, the vertex* $v_n$ *has the property that if* $v_j \sim v_n$*, and* $j < n$*, then* $v_j$ *has no neighbour before* $v_n$*.*

**Proof.** Note that $\sum s_i d_i/i < \infty$ implies $s_n d_n/n \to 0$ as $n \to \infty$, so there exists $N$ such that for all $n > N$, $s_n d_n/n < 1/3$. Consider such an $n$. Below we will compute the probability that at stage $n$, the vertex $v_n$ attaches only to the vertices that are currently of degree 0, i.e., singletons.

Observe that during the process, for every $i$ with $d_i = 0$ one singleton is created, and if $d_i > 0$, then at most $d_i$ singletons are destroyed. One can view this as always creating a singleton and then destroying no more than $2d_i$. So at step $i$, the number of singletons is at least $i - \sum_{j=0}^{i} 2d_j = i - 2s_i$, and then, by Proposition 5.1 and the fact that $d_n/n < 1/3$, the probability that $v_n$ attaches to only singletons is at least

$$\frac{\binom{n-2s_n}{d_n}}{\binom{n}{d_n}} \geq \left(1 - \frac{2s_n}{n - d_n}\right)^{d_n} \geq \left(1 - \frac{2s_n/n}{1 - d_n/n}\right)^{d_n}$$
$$\geq \left(1 - \frac{2s_n/n}{2/3}\right)^{d_n} = \left(1 - \frac{3s_n}{n}\right)^{d_n}.$$

Hence, the probability that this happens to all vertices beyond $N$ is at least

$$\prod_{n=N}^{\infty} \left(1 - \frac{3s_n}{n}\right)^{d_n}.$$

The last product is positive as $\sum s_i d_i/i < \infty$. Hence, we have that for all $\epsilon > 0$ there exists $M$ such that $\prod_{n=M}^{\infty} (1 - (3s_n d_n/n)) > 1 - \epsilon$. That means that with probability greater than $1 - \epsilon$, every vertex beyond $M$ attaches to singletons.

To complete the proof, suppose that the existence of an $N_1$ as in the statement has probability $p < 1$. Choose $\epsilon < 1 - p$. According to the argument above, there exists an

$M$ such that the probability that every vertex beyond $M$ attaches to singletons is greater than $1 - \epsilon > p$, and since $M$ is a suitable choice for $N_1$, this is a contradiction. □

**Theorem 5.5 (very sparse case).** *Suppose $\sum s_i d_i / i < \infty$. Then the process almost surely generates a graph $G$ for which there is $N = N(G)$ such that for all $n > N$, $v_n$ is in a star. Moreover, if $d_n > 0$, then $v_n$ is in a $d_n$-star.*

**Proof.** By Lemma 5.4, the process almost surely generates a graph $G$ for which there is $N = N(G)$ such that for all $n > N$, at stage $n$, either $d_n = 0$ or $v_n$ attaches to $d_n$ many current degree 0 vertices which precede $v_n$. Note that $v_k$, $k > n$, leaves untouched the star generated by $v_n$. Therefore, we obtain that $v_n$ is in a $d_n$-star. If $n > N$ and $d_n = 0$, then the component of $v_n$ in $G$ is either a singleton or a star generated by some $v_m$, $m > n$. □

## 6. Zero-one sequences

As the title suggests, the standing assumption for the section is that $0 \leq d_n \leq 1$ for all $n \geq 0$. We will also assume that there are infinitely many 1s in the sequence, as otherwise we really have a finite sequence and an essentially finite graph (plus isolated vertices), and we get a problem of a very different flavour. It is clear that the number of connected components of the generated graph is equal to the number of 0s in the sequence, and each component is a tree.

### 6.1. Notation

For this section, it will be convenient to introduce some notation to denote certain tuples of indices and sums and products. First, we introduce notation on products and tuples.

We let $\mathbb{N}^{<\mathbb{N}}$ denote the set of all finite strings of $\mathbb{N}$ including the empty string. We define $f : \mathbb{N}^{<\mathbb{N}} \to [0, \infty)$ by

$$f(\sigma) = \prod_{i=1}^{n} \frac{d_{\sigma_i}}{\sigma_i},$$

where $\sigma = (\sigma_1, \ldots, \sigma_n)$. (By convention, $f(\sigma) = 1$ when $\sigma$ is the empty string.) The definition of $f(\sigma)$ depends on the fixed sequence $\{d_n\}$.

Suppose $i, j, l \geq 1$ with $i \leq j$. Then,

$$A_i^l = \{(\sigma_1, \ldots, \sigma_l) \in \mathbb{N}^l \mid \min\{\sigma_1, \ldots, \sigma_l\} = i\},$$

$$B_i^l = \{\sigma \in A_i^l \mid \sigma \text{ is strictly increasing}\},$$

$$B_{i,j}^l = \{(\sigma_1, \ldots, \sigma_l) \in B_i^l \mid \max\{\sigma_1, \ldots, \sigma_l\} \leq j\},$$

$$C_i^l = \{\sigma \in A_i^l \mid \sigma \text{ is injective}\},$$

$$D_i^l = \{i, i+1, \ldots\}^l = \{(\sigma_1, \ldots, \sigma_l) \in \mathbb{N}^l \mid \min\{\sigma_1, \ldots, \sigma_l\} \geq i\}.$$

The following notation is about sums and series:

$$s_{m,n} = \sum_{i=m}^{n} d_i, \quad s_n = s_{0,n} = \sum_{i=0}^{n} d_i, \quad t_{n,m} = \sum_{i=n}^{m} \frac{d_i}{i}, \quad t_n = t_{n,\infty} = \sum_{i=n}^{\infty} \frac{d_i}{i}.$$

### 6.2. The sparse case for zero–one sequences

What we proved in § 5 essentially gives us the behaviour of the probability space in the dense case, when $\sum d_i/i = \infty$, and the very sparse case, when $\sum s_i d_i/i < \infty$. We will summarize these findings (and much more) in Theorem 2.5. This subsection will be devoted entirely to the sparse case. Accordingly, throughout the subsection we will assume that $\sum d_i/i < \infty$. Our findings *will* apply for the very sparse case, giving an alternative proof of the characterization of the space in that case. However, note that the findings of § 5 apply *in the general setting* (not only zero–one), so those theorems still have their importance.

**Proposition 6.1.** $\lim_{i\to\infty} s_i/i = 0$.

**Proof.** Let $\epsilon > 0$ small. Then,

$$\frac{s_n}{n} = \frac{\sum_{i=0}^{n} d_i}{n} \leq \frac{\sum_{i=0}^{\lfloor \epsilon n/2 \rfloor} d_i}{n} + \sum_{i=\lfloor \epsilon n/2 \rfloor+1}^{n} \frac{d_i}{i} \leq \frac{\epsilon}{2} + \sum_{i=\lfloor \epsilon n/2 \rfloor+1}^{\infty} \frac{d_i}{i}.$$

As $\sum d_i/i < \infty$, the second term converges to zero as $n \to \infty$, so eventually it will be less than $\epsilon/2$. □

Let $a(k)_i$ denote the expected number of trees of size $k$ spanned by the vertex set $\{v_0, \ldots, v_i\}$. We will prove a sequence of technical lemmas about the sequences $a(k)$.

**Proposition 6.2.** *For $i \geq 1$ and $k \geq 2$,*

$$a(k)_i = a(k)_{i-1} + \frac{(k-1)a(k-1)_{i-1}}{i} d_i - \frac{k a(k)_{i-1}}{i} d_i.$$

**Proof.** Suppose $d_i = 1$. Consider the process just before we add the edge from $v_i$. Let $p_+$ be the probability that we increase the number of trees of size $k$, and let $p_-$ be the probability that we decrease that number. In either case, the change is $\pm 1$. Let $p_0 = 1 - (p_+ - p_-)$. On the one hand, $p_+ = (k-1)a(k-1)_{i-1}/i$, and $p_- = k a(k)_{i-1}/i$. On the other hand,

$$a(k)_i = p_+(a(k)_{i-1} + 1) + p_-(a(k)_{i-1} - 1) + p_0 a(k)_{i-1} = a(k)_{i-1} + (p_+ - p_-).$$

In the other case, if $d_i = 0$, then $a(k)_i = a(k)_{i-1}$. □

**Lemma 6.3.** *Let $k \geq 2$. Then there exist positive constants $C_1, C_2$ such that, for all $i \geq 1$,*

$$C_1 \sum_{j=k}^{i} \frac{a(k-1)_{j-1}}{j} d_j \leq a(k)_i \leq C_2 \sum_{j=1}^{i} \frac{a(k-1)_{j-1}}{j} d_j.$$

https://doi.org/10.1017/S0013091517000487 Published online by Cambridge University Press

**Proof.** The upper bound is a straightforward consequence of Proposition 6.2. Indeed, $a(k)_i \leq a(k)_{i-1} + (ka(k-1)_{i-1}/i)d_i$, so $a(k)_i \leq k\sum_{j=1}^{i}(a(k-1)_{j-1}/j)d_j$.

For the lower bound, notice that

$$a(k)_i \geq a(k)_{i-1}\left(1-\frac{kd_i}{i}\right)+\frac{a(k-1)_{i-1}}{i}d_i.$$

This implies

$$\begin{aligned}
a(k)_i &\geq \sum_{j=1}^{i}\left[\frac{a(k-1)_{j-1}}{j}d_j\prod_{l=j+1}^{i}\left(1-\frac{kd_l}{l}\right)\right]\\
&\geq \sum_{j=k}^{i}\left[\frac{a(k-1)_{j-1}}{j}d_j\prod_{l=j+1}^{i}\left(1-\frac{kd_l}{l}\right)\right] \geq \sum_{j=k}^{i}\left[\frac{a(k-1)_{j-1}}{j}d_j\prod_{l=k+1}^{\infty}\left(1-\frac{kd_l}{l}\right)\right]\\
&\geq Q\sum_{j=k}^{i}\frac{a(k-1)_{j-1}}{j}d_j,
\end{aligned}$$

where $Q=\prod_{l=k+1}^{\infty}(1-(kd_l/l))$. Note that the second inequality is correct, because for $j=1,\dots,k-2$, we have $a(k-1)_{j-1}=0$ (so the omitted terms are zero), and for $j=k-1$ the omitted term is non-negative. Also note that $\sum d_i/i<\infty$ implies $Q>0$. □

**Lemma 6.4.** *Let $k\geq 2$. Then there exists positive constant $K$ such that for all $i$,*

$$a(k)_i \leq K\sum_{j=1}^{i}d_j t_{j+1}^{k-2}.$$

**Proof.** We proceed by induction on $k$. Let $k=2$. By Lemma 6.3, there exists $C_2$ such that $a(2)_i \leq C_2\sum_{j=1}^{i}(a(1)_{j-1}/j)d_j \leq C_2\sum_{j=1}^{i}d_j$.

Now suppose that $k\geq 3$. By Lemma 6.3 and the induction hypothesis, there exist positive constants $C_2$ and $C$ such that

$$\begin{aligned}
a(k)_i &\leq C_2\sum_{j=1}^{i}\frac{a(k-1)_{j-1}}{j}d_j \leq C\sum_{j=2}^{i}\sum_{l=1}^{j-1}d_l t_{l+1}^{k-3}\frac{d_j}{j} \leq C\sum_{l=1}^{i-1}\sum_{j=l+1}^{i}d_l t_{l+1}^{k-3}\frac{d_j}{j}\\
&\leq C\sum_{l=1}^{i-1}d_l t_{l+1}^{k-3}\sum_{j=l+1}^{i}\frac{d_j}{j} \leq C\sum_{l=1}^{i-1}d_l t_{l+1}^{k-2}.
\end{aligned}$$

□

**Lemma 6.5.** *Let $l\geq 1$. Then,*

$$\sum_{i=1}^{\infty}\sum_{\sigma\in A_i^l}f(\sigma)<\infty.$$

https://doi.org/10.1017/S0013091517000487 Published online by Cambridge University Press

**Proof.** For $l = 1$ the above statement is equivalent to the sparsity condition $\sum d_i/i < \infty$. For $l > 1$, we note that

$$\sum_{\sigma \in A_i^l} f(\sigma) \leq \sum_{j=1}^{l} \sum_{\substack{\sigma \in A_i^l \\ \sigma_j = i}} \frac{d_i}{i} \prod_{\substack{k=1 \\ k \neq j}}^{l} \frac{d_{\sigma_k}}{\sigma_k} = \sum_{j=1}^{l} \frac{d_i}{i} \sum_{\sigma \in D_i^{l-1}} f(\sigma) = \sum_{j=1}^{l} \frac{d_i}{i} t_i^{l-1} = l \frac{d_i}{i} t_i^{l-1}$$

Then,

$$\sum_{i=1}^{\infty} \sum_{\sigma \in A_i^l} f(\sigma) \leq l \sum_{i=1}^{\infty} \frac{d_i}{i} t_i^{l-1}.$$

As $\lim_{i \to \infty} t_i = 0$, we have that $\{t_i\}_{i=1}^{\infty}$ is bounded and hence the desired series converges. □

**Lemma 6.6.** *Suppose that $l \geq 1$. Then,*

$$\sum_{i=1}^{\infty} d_i t_i^l = \infty \implies \sum_{i=1}^{\infty} s_i \sum_{\sigma \in B_i^l} f(\sigma) = \infty.$$

**Proof.** By rearranging and switching the order of summation, we have that

$$\infty = \sum_{i=1}^{\infty} d_i t_i^l = \sum_{i=1}^{\infty} d_i \sum_{j=i}^{\infty} \sum_{\sigma \in A_j^l} f(\sigma) = \sum_{j=1}^{\infty} \sum_{i=1}^{j} d_i \left( \sum_{\sigma \in A_j^l} f(\sigma) \right) = \sum_{j=1}^{\infty} s_j \sum_{\sigma \in A_j^l} f(\sigma). \quad (3)$$

We next observe that if $l = 1$, then $A_i^l = B_i^l$ and the proof is complete. Hence, let us assume that $l \geq 2$. We will next show that

$$\sum_{i=1}^{\infty} s_i \sum_{\sigma \in A_i^l \setminus C_i^l} f(\sigma) < \infty. \quad (4)$$

$$\begin{aligned}
\sum_{i=1}^{\infty} s_i \sum_{\sigma \in A_i^l \setminus C_i^l} f(\sigma) &\leq \sum_{i=1}^{\infty} s_i \sum_{1 \leq j < k \leq l} \sum_{\substack{\sigma \in A_i^l \\ \sigma_j = \sigma_k}} f(\sigma) \\
&= \sum_{i=1}^{\infty} s_i \sum_{1 \leq j < k \leq l} \left( \sum_{\substack{\sigma \in A_i^l \\ \sigma_j = \sigma_k = i}} f(\sigma) + \sum_{m=i+1}^{\infty} \sum_{\substack{\sigma \in A_i^l \\ \sigma_j = \sigma_k = m}} f(\sigma) \right) \\
&\leq \sum_{i=1}^{\infty} s_i \sum_{1 \leq j < k \leq l} \left( \left( \frac{d_i}{i} \right)^2 \sum_{\sigma \in D_i^{l-2}} f(\sigma) + \sum_{m=i+1}^{\infty} \sum_{\substack{1 \leq p \leq l \\ p \notin \{j,k\}}} \sum_{\substack{\sigma \in A_i^l \\ \sigma_j = \sigma_k = m \\ \sigma_p = i}} f(\sigma) \right)
\end{aligned}$$

https://doi.org/10.1017/S0013091517000487 Published online by Cambridge University Press

$$= \sum_{i=1}^{\infty} s_i \sum_{1\le j<k\le l} \left( \left(\frac{d_i}{i}\right)^2 \sum_{\sigma\in D_i^{l-2}} f(\sigma) + \sum_{m=i+1}^{\infty} \sum_{\substack{1\le p\le l\\ p\notin\{j,k\}}} \frac{d_i}{i}\frac{d_m}{m}\frac{d_m}{m} \sum_{\sigma\in D_i^{l-3}} f(\sigma)\right)$$

$$\le \sum_{i=1}^{\infty} s_i \sum_{1\le j<k\le l} \left( \left(\frac{d_i}{i}\right)^2 \sum_{\sigma\in D_i^{l-2}} f(\sigma) + (l-2)\left(\frac{d_i}{i}\right)^2 \sum_{m=i+1}^{\infty} \frac{d_m}{m} \sum_{\sigma\in D_i^{l-3}} f(\sigma)\right)$$

$$\le \sum_{i=1}^{\infty} s_i \sum_{1\le j<k\le l} \left( \left(\frac{d_i}{i}\right)^2 \sum_{\sigma\in D_i^{l-2}} f(\sigma) + (l-2)\left(\frac{d_i}{i}\right)^2 \sum_{\sigma\in D_i^{l-2}} f(\sigma)\right)$$

$$\le (l-1)\binom{l}{2} \sum_{i=1}^{\infty} s_i \left(\frac{d_i}{i}\right)^2 \sum_{\sigma\in D_i^{l-2}} f(\sigma) = (l-1)\binom{l}{2}\sum_{i=1}^{\infty} \frac{s_i}{i}\frac{d_i}{i} t_i^{l-2} < \infty.$$

As before, the last inequality follows as $\{s_i/i\}_{i=1}^{\infty}$ and $\{t_i\}_{i=1}^{\infty}$ are bounded sequences.

Putting (3) and (4) together, we have that

$$\sum_{i=1}^{\infty} s_i \sum_{\sigma\in C_i^l} f(\sigma) = \infty.$$

Noting

$$\sum_{i=1}^{\infty} s_i \sum_{\sigma\in C_i^l} f(\sigma) = l! \sum_{i=1}^{\infty} s_i \sum_{\sigma\in B_i^l} f(\sigma),$$

the proof is complete. □

**Lemma 6.7.** *Let $k \ge 2$. Then, there exist $C$ and $N$ such that for all $i \in \mathbb{N}$*

$$a(2)_i \ge C s_{N,i} \quad \text{and}$$

$$a(k)_i \ge C \sum_{j=3}^{i+3-k} s_{N,j-1} \sum_{\sigma\in B_{j,i}^{k-2}} f(\sigma) \quad \text{for } k \ge 3.$$

**Proof.** Let $N \ge 2$. Since $s_n/n \to 0$, there exists $N$ such that, for all $n > N$, $2s_n/n \le 1/2$,

First consider $k = 2$. From Lemma 6.3, there exists a constant $C$ such that, for all $i \ge 2$,

$$a(2)_i \ge C \sum_{j=2}^{i} \frac{a(1)_{j-1}}{j} d_j.$$

https://doi.org/10.1017/S0013091517000487 Published online by Cambridge University Press

So, for all $i \geq N$,

$$a(2)_i \geq C \sum_{j=N}^{i} \frac{j-2s_{j-1}}{j} d_j \geq C \sum_{j=N}^{i} \frac{j-2s_j}{j} d_j$$
$$= C \sum_{j=N}^{i} \left(1 - \frac{2s_j}{j}\right) d_j \geq \frac{C}{2} \sum_{j=N}^{i} d_j = \frac{C}{2} s_{N,i}.$$

Now assume $k \geq 3$. We will proceed by induction, so we consider $k = 3$ first. There exist $C$ and $C'$ such that, for all $i \geq 3$,

$$a(3)_i \geq C \sum_{j=3}^{i} \frac{a(2)_{j-1}}{j} d_j \geq C' \sum_{j=3}^{i} s_{N,j-1} \frac{d_j}{j} = C' \sum_{j=3}^{i} s_{N,j-1} \sum_{\sigma \in B^1_{j,i}} f(\sigma).$$

Now assume that $k \geq 4$. There exists a $C$ such that, for all $i$,

$$a(k)_i \geq C \sum_{j=k}^{i} \frac{a(k-1)_{j-1}}{j} d_j \geq C \sum_{j=k}^{i} \sum_{l=3}^{j-k+3} s_{N,l-1} \sum_{\sigma \in B^{k-3}_{l,j-1}} f(\sigma) \frac{d_j}{j}$$
$$\geq C \sum_{l=3}^{i-k+3} s_{N,l-1} \sum_{j=l-3+k}^{i} \frac{d_j}{j} \sum_{\sigma \in B^{k-3}_{l,j-1}} f(\sigma) = C \sum_{l=3}^{i-k+3} s_{N,l-1} \sum_{\sigma \in B^{k-2}_{l,i}} f(\sigma). \qquad \square$$

**Lemma 6.8.** *Let $k \geq 2$. Then,*

$$\left[\sum_{j=1}^{\infty} d_j t_j^{k-2} = \infty\right] \implies \left[\lim_{i\to\infty} a(k)_i = \infty\right].$$

**Proof.** The case $k = 2$ follows directly from the previous Lemma 6.7. Hence, assume $k \geq 3$. By Lemma 6.6 and the hypothesis we have that

$$\sum_{j=1}^{\infty} s_j \sum_{\sigma \in B^{k-2}_j} f(\sigma) = \infty.$$

Let $N$ be the constant from Lemma 6.7. Lemma 6.5 and the fact that $0 \leq s_j - s_{N,j-1} \leq N+1$ imply that

$$0 \leq \sum_{j=1}^{\infty} s_j \sum_{\sigma \in B^{k-2}_j} f(\sigma) - \sum_{j=1}^{\infty} s_{N,j-1} \sum_{\sigma \in B^{k-2}_j} f(\sigma) \leq (N+1) \sum_{j=1}^{\infty} \sum_{\sigma \in B^{k-2}_j} f(\sigma)$$
$$\leq (N+1) \sum_{j=1}^{\infty} \sum_{\sigma \in A^{k-2}_j} f(\sigma) < \infty.$$

Hence we have that $\sum_{j=1}^{\infty} s_{N,j-1} \sum_{\sigma \in B^{k-2}_j} f(\sigma) = \infty$.

To complete the proof, we choose $M > 0$. Then there is $i_0$ such that

$$\sum_{j=3}^{i_0+3-k} s_{N,j-1} \sum_{\sigma \in B_j^{k-2}} f(\sigma) > \frac{M+1}{C},$$

where $C$ is the constant from Lemma 6.7. Now for each $1 \le j \le i_0 + 3 - k$, there exists $i_j$ such that

$$\sum_{\sigma \in B_{j,i_j}^{k-2}} f(\sigma) \ge \frac{-1}{C(i_0+3)(s_{i_0+3})} + \sum_{\sigma \in B_j^{k-2}} f(\sigma).$$

Now for all $i \ge \max\{i_0, i_1, \dots i_{i_0+3-k}\}$ we have that

$$\begin{aligned} a(k)_i &\ge C \sum_{j=3}^{i+3-k} s_{N,j-1} \sum_{\sigma \in B_{j,i}^{k-2}} f(\sigma) \ge C \sum_{j=3}^{i_0+3-k} s_{N,j-1} \sum_{\sigma \in B_{j,i_j}^{k-2}} f(\sigma) \\ &\ge C \sum_{j=3}^{i_0+3-k} s_{N,j-1} \left( \frac{-1}{C(i_0+3)(s_{i_0+3})} + \sum_{\sigma \in B_j^{k-2}} f(\sigma) \right) \\ &\ge -1 + C \sum_{j=3}^{i_0+3-k} s_{N,j-1} \sum_{\sigma \in B_j^{k-2}} f(\sigma) > -1 + C \frac{M+1}{C} = M, \end{aligned}$$

completing the proof. □

Recall that a *ray* is a one-way infinite path, i.e., a sequence of vertices $u_1, u_2, \dots$ in a graph such that $u_i \sim u_{i+1}$ for all $i$.

**Theorem 6.9.** *Suppose $\sum d_i/i < \infty$. Then the process almost surely generates a graph with no ray.*

**Proof.** The hypothesis implies that $t_n \to 0$, so there exists a positive integer $N$ such that for all $i \ge N$, $t_i < 1$. Fix $i \ge N$. For $k \ge 2$, let $E_k$ denote the event that the graph will contain a path of length $k$ that starts at the vertex $v_i$ and for every vertex $v_j$ on the path we have $j \ge i$:

$$\begin{aligned} \Pr(E_k) &= \sum_{\sigma \in B_i^{k+1}} \Pr(v_{\sigma_{l+1}} \sim v_{\sigma_l} \text{ for all } 1 \le l \le k) \\ &= \sum_{\sigma \in B_i^{k+1}} \frac{d_{\sigma_2}}{\sigma_2} \cdots \frac{d_{\sigma_{k+1}}}{\sigma_{k+1}} \le \sum_{\sigma \in D_i^k} f(\sigma) = t_i^k. \end{aligned}$$

Hence $\lim_{k\to\infty} \Pr(E_k) = 0$. We conclude that the probability that a ray emanates from the vertex $v_i$ such that every vertex of the ray is beyond $v_i$ is 0. Then almost surely, for all $i \ge N$, such a ray does not exist.

https://doi.org/10.1017/S0013091517000487 Published online by Cambridge University Press

It is easy to see that if the graph had a ray, it would also have a ray $v_{i_0}v_{i_1}\dots$ such that $i_0 \geq N$, and for all $j \in \mathbb{N}$, $i_j \geq i_0$. But we have just seen that the probability of that is 0. □

### 6.3. Proof of Theorem 2.5

Part (i) is an immediate consequence of Theorem 5.3 in the special case of zero-one sequences.

For part (ii), let $m < k$; Lemma 6.8 implies that $a(m)_i \to \infty$, that is, the expected number of components of size $m$ is infinity. But we will prove a stronger statement, namely that almost surely there are infinitely many components of size $m$. We will do this in two steps. First, we will show that almost surely infinitely many components of size $m$ are created. Then, using this fact, we will show that almost surely the final graph has infinitely many components of size $m$.

Let us proceed to show that infinitely many components of size $m$ are created. If $m = 1$, then this follows from the fact that the sequence $\{d_i\}$ contains infinitely many zeros. Let $m \geq 2$. Let $N_j$ be the random variable that counts the number of components of size $m-1$ spanned by vertices $\{v_0,\dots,v_j\}$. Let $E_j$ be the event that a component of size $m$ is created at step $j$ from vertex $v_j$. Then,

$$\begin{aligned}\Pr(E_j) &= \sum_{\ell=0}^{\infty} \Pr(N_{j-1} = \ell)\frac{\ell(m-1)d_j}{j}\\ &= (m-1)\frac{d_j}{j}\sum_{\ell=0}^{\infty} \ell \Pr(N_{j-1} = \ell) = (m-1)\frac{d_j}{j}a(m-1)_{j-1}.\end{aligned}$$

By Lemma 6.3, there exists a constant $C$, such that for all $i \geq 1$,

$$a(m)_i \leq C\sum_{j=1}^{i}\frac{a(m-1)_{j-1}}{j}d_j.$$

Since $a(m)_i \to \infty$, we conclude that

$$\sum_{j=1}^{\infty}\Pr(E_j) = (m-1)\sum_{j=1}^{\infty}\frac{d_j}{j}a(m-1)_{j-1} = \infty.$$

To show that almost surely infinitely many components of size $m$ are created, it suffices to show that for all $i_0 \in \mathbb{N}$, $\Pr(\bigcup_{j=i_0}^{\infty} E_j) = 1$. We will do this using the following theorem, sometimes referred to as the counterpart of the Borel–Cantelli lemma [6].

**Lemma 6.10.** *Let $A_1, A_2, \dots$ be a sequence of events such that $A_i \subseteq A_{i+1}$. Then,*

$$\Pr\left(\bigcup_{j=1}^{\infty} A_j\right) = 1 \iff \sum_{j=1}^{\infty}\Pr(A_{j+1}|A_j^c) = \infty.$$

https://doi.org/10.1017/S0013091517000487 Published online by Cambridge University Press

We apply the above lemma to the situation where $A_j = \bigcup_{i=i_0}^{i_0+j-1} E_i$. We note that for all $j \geq 1$, we have that

$$\Pr(A_{j+1}|A_j^c) \geq \Pr(E_{j+i_0}),$$

implying that

$$\sum_{j=i_0}^{\infty} \Pr(A_{j+1}|A_j^c) = \infty.$$

By Lemma 6.10 and the fact that $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=i_0}^{\infty} E_i$, we have that $\Pr(\bigcup_{j=i_0}^{\infty} E_i) = 1$.

We next show that the final graph almost surely has infinitely many components of size $m$. It suffices to show that for every $i \in \mathbb{N}$ and $\varepsilon > 0$, the probability that the final graph contains a component of size $m$ containing a vertex $v_j$, $j \geq i$, is greater than $1 - \varepsilon$. Indeed, this is the case as it implies that for all $i$, with probability one, a component of size $m$ with a vertex $v_j$, $j \geq i$ will exist. Then, taking intersection over all $i$s, we obtain the desired result with probability one.

If a component of size $m$ is created at vertex $v_j$, the probability that it will not be destroyed is

$$q_j = \prod_{\ell=j+1}^{\infty} \left(1 - \frac{d_\ell m}{\ell}\right).$$

As by hypothesis $\sum d_\ell/\ell = \infty$, we have that $q_j \to 1$ as $j \to \infty$. Since almost surely $E_j$ occurs for infinitely many $j$s, we have that for sufficiently large $j$ we create a component of size $m$ without later destroying it with probability at least $1 - \varepsilon$.

For any $m' \geq k$, Lemma 6.4 shows that the expected number of components of size $m'$ is finite, therefore almost surely there are finitely many components of size $m'$.

There are two things that remain to be proven to finish the proof of the theorem. First, that if $m < k$, and $T$ is a tree with $|T| = m$, then almost surely there are infinitely many components of the graph isomorphic to $T$. From the argument above, we have that almost surely there are infinitely many components of size $m$, and also, if $C$ is a component, $\Pr(C \cong T \mid |C| = m) > 0$, so the statement follows.

The second thing is that almost surely there is no infinite component of the graph. From Theorem 5.3, we know that almost surely each vertex is of finite degree (i.e., the graph is locally finite). Every locally finite connected infinite graph contains a ray (see, e.g., Proposition 8.2.1. in [**9**]). So Theorem 6.9 finishes the proof.

**Corollary 6.11.** *If $\sum_{i=1}^{\infty} d_i s_i / i < \infty$, then the space has infinitely many atoms, and all of them are of the form $F \cup F_1 \cup F_2$, where $F$ is a finite forest.*

**Proof.**

$$\sum_{j=1}^{\infty} d_j t_{j+1} \leq \sum_{j=1}^{\infty} d_j t_j = \sum_{j=1}^{\infty} d_j \sum_{i=j}^{\infty} \frac{d_i}{i} = \sum_{i=1}^{\infty} \frac{d_i}{i} \sum_{j=1}^{i} d_j = \sum_{i=1}^{\infty} \frac{d_i s_i}{i} < \infty,$$

so Theorem 2.5 implies the statement. □

The following theorem is a natural analogue of Corollary 4.4.

**Theorem 6.12.** *Fix $0 < p < 1$, and consider the double random process with $d_i = 1$ with probability $p$, otherwise $d_i = 0$ for $i > 0$. The process almost surely generates infinitely many copies of $\omega$-trees.*

**Proof.** We will prove that the hypotheses of Theorem 2.5(i) are satisfied almost surely. Let $X_n = \sum_{i=1}^{n} d_i/i$. On the one hand, $\mu_n := E[X_n] \geq p \ln n$. On the other hand,

$$\sigma_n^2 := \mathrm{Var}[X_n] = \sum_{i=1}^{n} \mathrm{Var}\left[\frac{d_i}{i}\right] \leq \sum_{i=1}^{n} \frac{p - p^2}{i^2} \leq 2.$$

Fix $M > 0$. Using Chebyshev's inequality, if $\mu_n > M$,

$$\Pr[X_n \leq M] \leq \Pr[|X_n - \mu_n| \geq \mu_n - M] \leq \frac{\sigma_n^2}{(\mu_n - M)^2} \leq \frac{2}{(p \ln n - M)^2} \to 0.$$

Hence, almost surely $X_n \to \infty$. □

**Acknowledgement.** The authors are indebted to the referee for suggesting substantial changes which improved the paper.

## References


1. A.-L. Barabási and R. Albert, Emergence of scaling in random networks, *Science* **286**(5439) (1999), 509–512.
2. B. Bollobás and O. M. Riordan, The diameter of a scale-free random graph, *Combinatorica* **24**(1) (2004), 5–34.
3. A. Bonato, The search for $n$-e.c. graphs, *Contrib. Discrete Math.* **4**(1) (2009), 40–53.
4. A. Bonato and J. C. M. Janssen, Infinite limits and adjacency properties of a generalized copying model, *Internet Math.* **4**(2–3) (2007), 199–223.
5. C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós and K. L. Vesztergombi, Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing, *Adv. Math.* **219**(6) (2008), 1801–1851.
6. F. T. Brusss, A counterpart of the Borel-Cantelli lemma, *J. Appl. Probab.* **17**(4) (1980), 1094–1101.
7. P. J. Cameron, *The random graph revisited*, European congress of mathematics, volume I (Barcelona, 2000), Progress in Mathematics, Volume 201, pp. 267–274 (Birkhäuser, Basel, 2001).
8. U. B. Darji and J. D. Mitchell, Approximation of automorphisms of the rationals and the random graph, *J. Group Theory* **14**(3) (2011), 361–388.
9. R. Diestel, *Graph theory*, 4th edition, Graduate Texts in Mathematics, Volume 173 (Springer, 2010).
10. P. Erdős and A. Rényi, Asymmetric graphs, *Acta Math. Acad. Sci. Hung.* **14** (1963), 295–315.
11. S. Janson and S. Severini, An example of graph limit of growing sequences of random graphs, *J. Comb.* **4**(1) (2013), 67–80.
12. R. D. Kleinberg and J. M. Kleinberg, Isomorphism and embedding problems for infinite limits of scale-free graphs, Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms, ACM, New York, 2005, pp. 277–286 (electronic).
13. R. Kumar, P. Raghavan, S. Rajagopalan, D. Sivakumar, A. Tomkins and E. Upfal, *Stochastic models for the web graph*, 41st Annual Symposium on Foundations of

Computer Science (Redondo Beach, CA, 2000), pp. 57–65 (IEEE Computer Society Press, Los Alamitos, CA, 2000).
14. L. Lovász and B. Szegedy, Limits of dense graph sequences, *J. Combin. Theory Ser. B* **96** (2006), 933–957.
15. R. Rado, Universal graphs and universal functions, *Acta Arith.* **9** (1964), 331–340.
16. J. H. Spencer, *The strange logic of random graphs*, Algorithms and Combinatorics, Volume 22 (Springer, 2001).
17. J. K. Truss, The group of the countable universal graph, *Math. Proc. Cambridge Philos. Soc.* **98**(2) (1985), 213–245.
18. R. van der Hofstad, Percolation and random graphs, in *New perspectives in stochastic geometry* (eds I. S. Molchanov and W. S. Kendall), pp. 173–247 (Oxford University Press, Oxford, 2010).

https://doi.org/10.1017/S0013091517000487 Published online by Cambridge University Press