

PREDECESSORS AND SUCCESSORS IN RANDOM MAPPINGS WITH EXCHANGEABLE IN-DEGREES

JENNIE C. HANSEN,* *Heriot-Watt University*

JERZY JAWORSKI,** *Adam Mickiewicz University*

Abstract

In this paper we characterise the distributions of the number of predecessors and of the number of successors of a given set of vertices, A , in the random mapping model, $T_n^{\hat{D}}$ (see Hansen and Jaworski (2008)), with exchangeable in-degree sequence $(\hat{D}_1, \hat{D}_2, \dots, \hat{D}_n)$. We show that the exact formulae for these distributions and their expected values can be given in terms of the distributions of simple functions of the in-degree variables $\hat{D}_1, \hat{D}_2, \dots, \hat{D}_n$. As an application of these results, we consider two special examples of $T_n^{\hat{D}}$ which correspond to random mappings with preferential and anti-preferential attachment, and determine the exact distributions for the number of predecessors and the number of successors in these cases. We also characterise, for these two special examples, the asymptotic behaviour of the expected numbers of predecessors and successors and interpret these results in terms of the threshold behaviour of epidemic processes on random mapping graphs. The families of discrete distributions obtained in this paper are also of independent interest.

Keywords: Random mapping; exchangeable in-degrees; preferential attachment; anti-preferential attachment; urn scheme; epidemic process

2010 Mathematics Subject Classification: Primary 60C05

Secondary 05C80

1. Introduction

In this paper we investigate the distributions of the numbers of predecessors and successors of vertices from a given set in random mappings with exchangeable in-degrees. In order to describe such mappings, we begin with a general definition of a random mapping model. For positive integer n , let $[n]$ denote the set of integers $\{1, 2, \dots, n\}$ and let \mathcal{M}_n denote the set of all functions from $[n]$ into $[n]$. A random mapping $T: [n] \rightarrow [n]$ is a random element of the space of mappings \mathcal{M}_n . We note that any $f \in \mathcal{M}_n$ can be represented as a directed graph, $G(f)$, on vertices $1, 2, \dots, n$ such that there is a directed edge from vertex i to j if and only if $f(i) = j$. So, if T is a random mapping on $[n]$ then $G(T)$ is a random directed graph on n labelled vertices. Since each vertex in $G(T)$ has out-degree 1, the components of the random digraph $G(T)$ consist of directed cycles with directed trees attached to the cycles.

Various random mapping models have been studied since the 1950s and the structure of their corresponding digraphs have received much attention in the literature; see, for example, [5], [6], [15], [19], [23], [24], [27], [29], and the references therein. In particular, these models have been considered as models for epidemic processes and they have a natural application in

Received 20 October 2011; revision received 11 December 2012.

* Postal address: Actuarial Mathematics and Statistics Department and The Maxwell Institute for Mathematical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK. Email address: j.hansen@hw.ac.uk

** Postal address: Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Umultowska 87, 61-614 Poznań, Poland. Email address: jaworski@amu.edu.pl

the analysis of cryptographic systems (e.g. DES), in applications of Pollard’s algorithm, and in random number generation. In the context of applications, two statistics of particular interest for a random mapping T are the number of predecessors and the number of successors in the digraph $G(T)$ of a set of vertices $A \subseteq [n]$. In the case of the uniform mapping $T_n: [n] \rightarrow [n]$, where $\mathbb{P}\{T_n = f\} = 1/n^n$ for each $f \in \mathcal{M}_n$, the exact and asymptotic distributions of the number of predecessors and the number of successors of vertices from a given set in $G_n \equiv G(T_n)$ have been extensively investigated (see [3], [4], [6], [8], [21], [22], and [25]). In another direction, these variables have also been investigated for the evolutionary model, $T_{n,q}: [n] \rightarrow [n]$, which is a generalization of the uniform model. The model $T_{n,q}$ depends on an additional parameter $0 \leq q \leq 1$ and is defined so that, for $1 \leq i \leq n$, $\mathbb{P}\{T_{n,q}(i) = i\} = q$ and, for $1 \leq i, j \leq n$ such that $i \neq j$, $\mathbb{P}\{T_{n,q}(i) = j\} = (1 - q)/(n - 1)$. Clearly, for $T_{n,q}$, the distributions of the numbers of predecessors and successors will depend also on the parameter $0 \leq q \leq 1$ and this dependence has been characterised in [16] and [17].

Both the uniform model T_n and the evolutionary model $T_{n,q}$ described above are examples of random mappings with independent vertex choices. In this paper we consider, instead, the structure of random mappings with exchangeable in-degrees where vertex choices are not necessarily independent. This class of random mappings was introduced in [11] (see also [10] and [12]) and can be defined as follows. Suppose that $f \in \mathcal{M}_n$. Then, for $1 \leq i \leq n$, we let $d_i(f)$ denote the in-degree of vertex i in the functional digraph $G(f)$ which represents the mapping f , and define $\vec{d}(f) \equiv (d_1(f), \dots, d_n(f))$. Also, given a vector $\vec{d} \equiv (d_1, d_2, \dots, d_n)$ of nonnegative integers such that $\sum_{i=1}^n d_i = n$, define

$$\mathcal{M}_n(\vec{d}) \equiv \{f \in \mathcal{M}_n : \vec{d}(f) = \vec{d}\}$$

to be the set of all mappings $f \in \mathcal{M}_n$ with in-degree sequence \vec{d} and note that the size of the set $\mathcal{M}_n(\vec{d})$ is given by

$$|\mathcal{M}_n(\vec{d})| = \frac{n!}{\prod_{i=1}^n d_i!}.$$

Now suppose that $\hat{D}_1, \hat{D}_2, \dots, \hat{D}_n$ is a collection of nonnegative integer-valued exchangeable random variables such that $\hat{D}_1 + \hat{D}_2 + \dots + \hat{D}_n = n$. Then, given the event $\{\hat{D}_i = d_i, i \in [n]\}$ (with $\mathbb{P}\{\hat{D}_i = d_i, i = 1, 2, \dots, n\} > 0$), we define the conditional distribution of the random mapping $T_n^{\hat{D}}$ with exchangeable degree sequence $(\hat{D}_1, \hat{D}_2, \dots, \hat{D}_n)$ to be the uniform distribution over $\mathcal{M}_n(\vec{d})$. So, the distribution of $T_n^{\hat{D}}$ is given by

$$\mathbb{P}\{T_n^{\hat{D}} = f\} = \frac{\prod_{i=1}^n (d_i(f))!}{n!} \mathbb{P}\{\hat{D}_i = d_i(f), 1 \leq i \leq n\}$$

for any $f \in \mathcal{M}_n$.

One of the most important and attractive features of the random mapping $T_n^{\hat{D}}$, as defined above, is that many distributions of statistics related to the structure of $G_n^{\hat{D}} \equiv G(T_n^{\hat{D}})$ (e.g. the number of components in $G_n^{\hat{D}}$, the size of a typical component in $G_n^{\hat{D}}$, etc.) can be expressed in terms of a calculus based on the joint distribution of the variables $(\hat{D}_1, \dots, \hat{D}_n)$. In addition, this calculus turns out to be straightforward to use for a large class of random mappings with exchangeable in-degrees which can be constructed as follows. Suppose that D_1, D_2, \dots, D_n are independent and identically distributed (i.i.d.) nonnegative integer-valued random variables with $\mathbb{P}\{\sum_{i=1}^n D_i = n\} > 0$, and let $\hat{D}_1, \hat{D}_2, \dots, \hat{D}_n$ be a sequence of random variables with

joint distribution given by

$$\mathbb{P}\{\hat{D}_i = d_i, 1 \leq i \leq n\} = \mathbb{P}\left\{D_i = d_i, 1 \leq i \leq n \mid \sum_{i=1}^n D_i = n\right\}.$$

Clearly, the variables $\hat{D}_1, \hat{D}_2, \dots, \hat{D}_n$ are exchangeable with $\sum_{i=1}^n \hat{D}_i = n$, so we can use $\hat{D}_1, \hat{D}_2, \dots, \hat{D}_n$ to construct $T_n^{\hat{D}}$ and $G_n^{\hat{D}}$. In the special case when the underlying i.i.d. variables D_1, D_2, \dots, D_n have a generalised negative binomial distribution we obtain a random mapping with ‘preferential attachment’, and when they have a binomial $\text{Bin}(m, p)$ distribution, we obtain a random mapping with ‘anti-preferential attachment’. For both of these natural and interesting examples, the calculus developed for random mappings with exchangeable in-degrees has been used to obtain the exact and asymptotic distributions for various structure variables associated with these models. We also note that when the underlying i.i.d. variables D_1, D_2, \dots, D_n have a Poisson distribution, then $T_n^{\hat{D}}$ corresponds to the uniform random mapping model. So, random mappings with exchangeable in-degrees can also be viewed as a generalisation, in a different direction, of the classical uniform model T_n .

In this paper we show that there is also a calculus for the distributions of the number of predecessors and the number of successors of a set of vertices $A \subseteq [n]$ in $G_n^{\hat{D}} \equiv G(T_n^{\hat{D}})$, where $T_n^{\hat{D}}$ is a random mapping with exchangeable in-degrees. In Section 2 we show exactly how these distributions depend on the joint distribution of the exchangeable in-degree variables $\hat{D}_1, \hat{D}_2, \dots, \hat{D}_n$ for $T_n^{\hat{D}}$. In Section 3 we apply this calculus to obtain exact formulae for the distributions of the number of predecessors and the number of successors in the special examples of mappings with preferential and anti-preferential attachment, respectively. We also investigate the asymptotic behaviour of the expected value of these distributions. We also note that the discrete distributions obtained in Section 3 are of general independent interest. Finally, in Section 4 we discuss the application of our results to the characterisation of the threshold behaviour of epidemic processes on random mapping digraphs and we suggest directions for further research.

2. Main results

For any $f \in \mathcal{M}_n$, we note that every component of $G(f)$ consists of a directed cycle with trees, directed towards the cycle, attached to it. Recall that if there exists an oriented path from i to j in $G(f)$ then j is said to be a *successor* of i while i is said to be a *predecessor* of j . More formally, for any $f \in \mathcal{M}_n$ and any positive integer ℓ , let $f^{(\ell)}$ denote the ℓ th iterate of f , and, for every $i \in [n]$, define $f^{(0)}(i) \equiv i$. Let

$$\mathcal{S}_f(j) \equiv \{i \in [n] : f^{(\ell)}(j) = i \text{ for some } \ell \geq 0\}$$

denote the successors of vertex j under f , and let

$$\mathcal{P}_f(i) \equiv \{j \in [n] : f^{(\ell)}(j) = i \text{ for some } \ell \geq 0\}$$

denote the predecessors of vertex i under f . Moreover, let

$$\mathcal{S}_f(A) = \bigcup_{j \in A} \mathcal{S}_f(j), \quad \mathcal{P}_f(B) = \bigcup_{i \in B} \mathcal{P}_f(i),$$

i.e. with respect to the mapping f , $\mathcal{S}_f(A)$ is the set of all successors of the vertices of $A \subset [n]$, and $\mathcal{P}_f(B)$ is the set of all predecessors of the vertices of $B \subset [n]$.

In this paper we are interested in the properties of $G_n^{\hat{D}}$ which are described by the random variables

$$s_n^{\hat{D}}(A) = |\mathcal{S}_{T_n^{\hat{D}}}(A)|, \quad p_n^{\hat{D}}(B) = |\mathcal{P}_{T_n^{\hat{D}}}(B)|.$$

We note that it follows from the definition of $T_n^{\hat{D}}$ and the exchangeability of the variables $\hat{D}_1, \hat{D}_2, \dots, \hat{D}_n$ that the distribution of $G_n^{\hat{D}}$ is invariant under relabelling of the vertices of $G_n^{\hat{D}}$. In particular, for any sets $A, B \subset [n]$, we have

$$s_n^{\hat{D}}([a]) \stackrel{D}{\sim} s_n^{\hat{D}}(A), \quad p_n^{\hat{D}}([b]) \stackrel{D}{\sim} p_n^{\hat{D}}(B), \tag{2.1}$$

where $a = |A|$ and $b = |B|$. Throughout this paper, for any $1 \leq a \leq n$, $s_n^{\hat{D}}[a]$ will denote the number of successors in $G_n^{\hat{D}}$ of a given a -element subset of the vertex set of $G_n^{\hat{D}}$. Similarly, for any $1 \leq b \leq n$, $p_n^{\hat{D}}[b]$ will denote the number of predecessors in $G_n^{\hat{D}}$ of a given b -element subset of the vertex set of $G_n^{\hat{D}}$. We note from (2.1) that the distributions of $s_n^{\hat{D}}[a]$ and $p_n^{\hat{D}}[b]$ depend only on the parameters a, b , and n and on the distribution of the variables $\hat{D}_1, \hat{D}_2, \dots, \hat{D}_n$.

In this section we obtain exact formulae for the distributions of the variables $s_n^{\hat{D}}[a]$ and $p_n^{\hat{D}}[b]$ defined above. We begin by adopting some additional notation. Given $1 \leq b \leq n$, we define the ‘cut’ digraph $G_{n,b}^{\hat{D}}$ by deleting, for each $1 \leq i \leq b$, the edge from i to $T_n^{\hat{D}}(i)$ in $G_n^{\hat{D}}$, and we let $T_{n,b}^{\hat{D}}$ denote the mapping from $[b+1, n] \equiv \{b+1, b+2, \dots, n\}$ to $[n]$ which corresponds to the directed edges of $G_{n,b}^{\hat{D}}$. For $1 \leq i \leq n$, we also introduce the ‘cut’ variable $\tilde{D}_i(n, b) \equiv d_i(G_{n,b}^{\hat{D}})$, the in-degree of vertex i in the ‘cut’ graph $G_{n,b}^{\hat{D}}$. Finally, let $\mathcal{M}_{n,b}$ denote the set of all mappings $h: [b+1, n] \rightarrow [n]$, and, for any nonnegative integers $\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_n$ such that $\sum_{i=1}^n \tilde{d}_i = n - b$, let

$$\mathcal{M}_{n,b}(\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_n) = \{h \in \mathcal{M}_{n,b} : d_i(G(h)) = \tilde{d}_i, 1 \leq i \leq n\}.$$

The conditional distribution of $T_{n,b}^{\hat{D}}$ is given by the following lemma.

Lemma 2.1. *Suppose that $\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_n$ are nonnegative integers such that $\sum_{i=1}^n \tilde{d}_i = n - b$. Then, for every $h \in \mathcal{M}_{n,b}(\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_n)$,*

$$\mathbb{P}\{T_{n,b}^{\hat{D}} = h \mid \tilde{D}_i(n, b) = \tilde{d}_i, 1 \leq i \leq n\} = \left(\binom{n-b}{\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_n} \right)^{-1},$$

that is, given $\{\tilde{D}_i(n, b) = \tilde{d}_i, 1 \leq i \leq n\}$, the distribution of $T_{n,b}^{\hat{D}}$ is uniform over $\mathcal{M}_{n,b}(\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_n)$.

Proof. Suppose that $\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_n$ are nonnegative integers such that $\sum_{i=1}^n \tilde{d}_i = n - b$, and that d_1, d_2, \dots, d_n are nonnegative integers such that $\sum_{i=1}^n d_i = n$ and $d_i \geq \tilde{d}_i$ for $1 \leq i \leq n$. Then, for every $h \in \mathcal{M}_{n,b}(\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_n)$, we have

$$\begin{aligned} &\mathbb{P}\{T_{n,b}^{\hat{D}} = h \mid \hat{D}_i = d_i, \tilde{D}_i(n, b) = \tilde{d}_i, 1 \leq i \leq n\} \\ &= \frac{\mathbb{P}\{T_{n,b}^{\hat{D}} = h \mid \hat{D}_i = d_i, 1 \leq i \leq n\}}{\mathbb{P}\{\tilde{D}_i(n, b) = \tilde{d}_i, 1 \leq i \leq n \mid \hat{D}_i = d_i, 1 \leq i \leq n\}} \\ &= \frac{b!}{(d_1 - \tilde{d}_1)! \cdots (d_n - \tilde{d}_n)!} \left[\left(\frac{(n-b)!}{\tilde{d}_1! \cdots \tilde{d}_n!} \right) \left(\frac{b!}{(d_1 - \tilde{d}_1)! \cdots (d_n - \tilde{d}_n)!} \right) \right]^{-1} \\ &= \left(\binom{n-b}{\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_n} \right)^{-1} \tag{2.2} \end{aligned}$$

since, given $\hat{D}_i = d_i$ for $1 \leq i \leq n$, the distribution of $T_n^{\hat{D}}$ is uniform over $\mathcal{M}_n(d_1, d_2, \dots, d_n)$. It is also clear that

$$\sum_{\substack{\vec{d} \text{ such that } \sum d_i = n \\ d_i \geq \tilde{d}_i}} \mathbb{P}\{\hat{D}_i = d_i, 1 \leq i \leq n \mid \tilde{D}_i(n, b) = \tilde{d}_i, 1 \leq i \leq n\} = 1. \tag{2.3}$$

The result now follows from (2.2), (2.3), and the total probability theorem.

Next, we give an exact formula for the distribution of $p_n^{\hat{D}}[b]$ in terms of the ‘cut’ variables $\tilde{D}_1(n, b), \tilde{D}_2(n, b), \dots, \tilde{D}_n(n, b)$.

Proposition 2.1. *Suppose that $1 \leq b \leq n$, and let $p_n^{\hat{D}}[b]$ denote the number of predecessors of a given b -element subset of the vertex set in $G_n^{\hat{D}}$. Then, for $0 \leq t \leq n - b$,*

$$\mathbb{P}\{p_n^{\hat{D}}[b] = b + t\} = \frac{b}{b + t} \mathbb{P}\left\{\sum_{i=1}^{b+t} \tilde{D}_i(n, b) = t\right\}.$$

Proof. It follows from (2.1) that we can assume throughout the proof that the given b -element subset of vertices is $[b] = \{1, 2, \dots, b\}$. Now consider the ‘cut’ digraph $G_{n,b}^{\hat{D}}$. Since $G_{n,b}^{\hat{D}}$ is obtained by deleting edges from $1 \leq i \leq b$ to $T_n^{\hat{D}}(i)$ in $G_n^{\hat{D}}$, it follows that $G_{n,b}^{\hat{D}}$ is, in general, a disjoint union of $\mathcal{F}(T_{n,b}^{\hat{D}})$, a directed forest of trees rooted at the vertices labelled $1, 2, \dots, b$, and of components which consist of directed cycles with directed trees attached. It follows from the definition of $G_{n,b}^{\hat{D}}$ that $\mathcal{P}_{T_n^{\hat{D}}}([b])$ is the vertex set of $\mathcal{F}(T_{n,b}^{\hat{D}})$, and so

$$p_n^{\hat{D}}[b] = |\mathcal{F}(T_{n,b}^{\hat{D}})|. \tag{2.4}$$

Now suppose that $t = 0$. It follows from (2.4) that

$$\begin{aligned} \mathbb{P}\{p_n^{\hat{D}}[b] = b\} &= \mathbb{P}\{|\mathcal{F}(T_{n,b}^{\hat{D}})| = b\} \\ &= \mathbb{P}\{\tilde{D}_i(n, b) = 0 \text{ for } 1 \leq i \leq b\} \\ &= \mathbb{P}\left\{\sum_{i=1}^b \tilde{D}_i(n, b) = 0\right\}. \end{aligned}$$

So the result holds for $t = 0$.

Next, suppose that $1 \leq t \leq n - b$. Since the distribution of $G_n^{\hat{D}}$ is invariant under relabelling of the vertices, we have

$$\mathbb{P}\{p_n^{\hat{D}}[b] = b + t\} = \binom{n - b}{t} \mathbb{P}\{\mathcal{P}_{T_n^{\hat{D}}}([b]) = [b + t]\}. \tag{2.5}$$

Let $\mathcal{A}(b, t, n, \hat{D})$ denote the event that $\mathcal{F}(T_{n,b}^{\hat{D}})$ is a directed forest on $[b + t]$ which is rooted on $[b]$. Then we have

$$\begin{aligned} &\mathbb{P}\{\mathcal{P}_{T_n^{\hat{D}}}([b]) = [b + t]\} \\ &= \mathbb{P}\{\mathcal{A}(b, t, n, \hat{D})\} \\ &= \mathbb{P}\left\{\mathcal{A}(b, t, n, \hat{D}) \mid \sum_{i=1}^{b+t} \tilde{D}_i(n, b) = t\right\} \mathbb{P}\left\{\sum_{i=1}^{b+t} \tilde{D}_i(n, b) = t\right\}. \end{aligned} \tag{2.6}$$

We use the following lemma to compute the right-hand side of (2.6).

Lemma 2.2. *Suppose that $1 \leq b \leq n$ and $1 \leq t \leq n - b$, and that $\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_{b+t}$ are nonnegative integers such that $\sum_{i=1}^b \tilde{d}_i \geq 1$ and $\sum_{i=1}^{b+t} \tilde{d}_i = t$. Then the number of directed forests on $[b + t]$, rooted at $[b]$, with in-degree sequence given by $\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_{b+t}$, is equal to*

$$\frac{(t - 1)!}{\tilde{d}_1! \tilde{d}_2! \cdots \tilde{d}_{b+t}!} \sum_{i=1}^b \tilde{d}_i,$$

where we interpret $0! = 1$.

Proof. Suppose that $1 \leq b \leq n$ and $1 \leq t \leq n - b$, and that $\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_{b+t}$ are nonnegative integers such that $\sum_{i=1}^b \tilde{d}_i \geq 1$ and $\sum_{i=1}^{b+t} \tilde{d}_i = t$. Let $\mathcal{T}_{b,t}(\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_{b+t})$ denote the set of directed forests on $1, 2, \dots, b+t$, rooted at $1, 2, \dots, b$ with in-degree sequence $\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_{b+t}$, and let $\mathcal{S}(\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_{b+t})$ denote the set of sequences of length t such that each $1 \leq i \leq b + t$ appears \tilde{d}_i times in the sequence and the last term in the sequence is some $i \in [b]$. Using a straightforward adaptation of the Prüfer tree code (see [26] and also [10]), we obtain a bijection between $\mathcal{T}_{b,t}(\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_{b+t})$ and $\mathcal{S}(\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_{b+t})$. The lemma now follows since

$$\mathcal{S}(\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_{b+t}) = \frac{(t - 1)!}{\tilde{d}_1! \tilde{d}_2! \cdots \tilde{d}_{b+t}!} \sum_{i=1}^b \tilde{d}_i.$$

We note that Lemma 2.2 can also be proved by appealing to Moon’s formula (see [20]) for the number of trees with a given degree sequence.

To complete the proof of Proposition 2.1, suppose that $\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_n$ are nonnegative integers such that

- (C1) $\sum_{i=1}^b \tilde{d}_i \geq 1,$
- (C2) $\sum_{i=1}^{b+t} \tilde{d}_i = t,$
- (C3) $\sum_{i=b+t+1}^n \tilde{d}_i = n - b - t.$

Then it follows from Lemma 2.1 and Lemma 2.2 that

$$\begin{aligned} & \mathbb{P}\{\mathcal{A}(b, t, n, \hat{D}) \mid \hat{D}_i(n, b) = \tilde{d}_i, 1 \leq i \leq n\} \\ &= \frac{(t - 1)! \sum_{i=1}^b \tilde{d}_i (n - b - t)!}{\tilde{d}_1! \tilde{d}_2! \cdots \tilde{d}_{b+t}! \tilde{d}_{b+t+1}! \cdots \tilde{d}_n!} \bigg/ \frac{(n - b)!}{\tilde{d}_1! \tilde{d}_2! \cdots \tilde{d}_n!} \\ &= \binom{n - b}{t}^{-1} \frac{1}{t} \sum_{i=1}^b \tilde{d}_i. \end{aligned} \tag{2.7}$$

Finally, let $\tilde{d}(b, t, n)$ denote the set of vectors $(\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_n)$ that satisfy conditions (C1)–(C3). Then it follows from (2.7) that

$$\begin{aligned} & \mathbb{P}\left\{\mathcal{A}(b, t, n, \hat{D}) \mid \sum_{i=1}^{b+t} \tilde{D}_i(n, b) = t\right\} \\ &= \sum_{(\tilde{d}_1, \tilde{d}_2, \dots, \tilde{d}_n) \in \tilde{d}(b, t, n)} \binom{n - b}{t}^{-1} \left(\frac{1}{t} \sum_{i=1}^b \tilde{d}_i\right) \\ & \quad \times \mathbb{P}\left\{\tilde{D}_i(n, b) = \tilde{d}_i, 1 \leq i \leq n \mid \sum_{i=1}^{b+t} \tilde{D}_i(n, b) = t\right\} \end{aligned}$$

$$\begin{aligned}
 &= \binom{n-b}{t}^{-1} \frac{1}{t} \mathbb{E} \left(\sum_{i=1}^b \tilde{D}_i(n, b) \mid \sum_{i=1}^{b+t} \tilde{D}_i(n, b) = t \right) \\
 &= \binom{n-b}{t}^{-1} \frac{b}{b+t}.
 \end{aligned} \tag{2.8}$$

The last equation in (2.8) follows from the exchangeability of the variables $\tilde{D}_1(n, b), \tilde{D}_2(n, b), \dots, \tilde{D}_n(n, b)$. The result now follows from (2.5), (2.6), and (2.8).

Theorem 2.1. *Let $p_n^{\hat{D}}[b]$ denote the number of predecessors of a given b -element subset of the vertex set in $G_n^{\hat{D}}$. Then, for $0 \leq t \leq n - b$,*

$$\mathbb{P} \{ p_n^{\hat{D}}[b] = b + t \} = \binom{n-b}{t} \frac{b}{b+t} \sum_{j=0}^b \frac{\binom{b}{j}}{\binom{n}{t+j}} \mathbb{P} \left\{ \sum_{i=1}^{b+t} \hat{D}_i = j + t \right\}.$$

Moreover,

$$\mathbb{E}(p_n^{\hat{D}}[b]) = b + b \sum_{k=1}^{n-b} \frac{(n-b)_k}{(n)_k} \mathbb{E}(\hat{D}_1 \hat{D}_2 \dots \hat{D}_k).$$

Proof. As in the proof of Proposition 2.1, we can assume without loss of generality that the given subset of vertices is $[b] = \{1, 2, \dots, b\}$. From Proposition 2.1, we have

$$\begin{aligned}
 &\mathbb{P} \{ p_n^{\hat{D}}[b] = b + t \} \\
 &= \frac{b}{b+t} \mathbb{P} \left\{ \sum_{i=1}^{b+t} \tilde{D}_i(n, b) = t \right\} \\
 &= \frac{b}{b+t} \sum_{j=0}^b \mathbb{P} \left\{ \sum_{i=1}^{b+t} \tilde{D}_i(n, b) = t \mid \sum_{i=1}^{b+t} \hat{D}_i = j + t \right\} \mathbb{P} \left\{ \sum_{i=1}^{b+t} \hat{D}_i = j + t \right\}.
 \end{aligned} \tag{2.9}$$

To compute the conditional probability in (2.9), it is convenient to introduce the following urn process. Start with an urn that contains \hat{D}_i balls labelled i for $1 \leq i \leq n$. Select balls sequentially from the urn such that at each step a ball is selected uniformly and at random from the balls still in the urn. This process generates a random mapping from $[n]$ to $[n]$ which is defined by mapping $k \in [n]$ to the label of the ball removed from the urn at the k th step. It is straightforward to check that this random mapping and $T_n^{\hat{D}}$ have the same distribution. It is also clear that we can define the ‘cut’ variables $\tilde{D}_i(n, b)$ in terms of this urn scheme. Specifically, for $1 \leq i \leq n$, $\tilde{D}_i(n, b)$ equals the number of balls labelled i that are left in the urn after the first b have been removed. So, given that $\sum_{i=1}^{b+t} \hat{D}_i = j + t$, we obtain $\sum_{i=1}^{b+t} \tilde{D}_i(n, b) = t$ if and only if in the first b draws from the urn, we choose exactly $b - j$ balls with labels that are greater than $b + t$ (and j balls with labels that are not greater than $b + t$). It follows that

$$\mathbb{P} \left\{ \sum_{i=1}^{b+t} \tilde{D}_i(n, b) = t \mid \sum_{i=1}^{b+t} \hat{D}_i = j + t \right\} = \frac{\binom{j+t}{j} \binom{n-t-j}{b-j}}{\binom{n}{b}} = \binom{n-b}{t} \frac{\binom{b}{j}}{\binom{n}{j+t}}.$$

The first result now follows by substitution into (2.9).

To obtain the formula for the expectation of the number of predecessors $p_n^{\hat{D}}[b]$, we define, for $i = b + 1, \dots, n$, $I_i[b]$ to be the indicator variable corresponding to the event

$\mathcal{A}_i[b] = \bigcup_{j=1}^b \mathcal{A}_{i,j}[b]$, where $\mathcal{A}_{i,j}[b]$ is the event that there is a path from i to j in $G_n^{\hat{D}}$ for which all internal vertices are in $\{b + 1, \dots, n\}$. It is clear that, for $b + 1 \leq i \leq n$, the events $\mathcal{A}_{i,1}[b], \mathcal{A}_{i,2}[b], \dots, \mathcal{A}_{i,b}[b]$ are disjoint, so

$$\begin{aligned} \mathbb{E}(p_n^{\hat{D}}[b]) &= \mathbb{E}\left(b + \sum_{i=b+1}^n I_i[b]\right) \\ &= b + \sum_{i=b+1}^n \sum_{j=1}^b \mathbb{P}\{\mathcal{A}_{i,j}[b]\} \\ &= b + b(n - b)\mathbb{P}\{\mathcal{A}_{b+1,1}[b]\}. \end{aligned} \tag{2.10}$$

Now observe that

$$\mathbb{P}\{\mathcal{A}_{b+1,1}[b]\} = \sum_{k=1}^{n-b} \mathbb{P}\{\mathcal{A}_{b+1,1}(k, [b])\}, \tag{2.11}$$

where $\mathcal{A}_{b+1,1}(k, [b])$ denotes the event that there is a path with k directed edges from $b + 1$ to 1 in $G_n^{\hat{D}}$ for which all internal vertices are in $\{b + 1, \dots, n\}$. Finally, we note that, using the same approach as in [10], we have

$$\mathbb{P}\{\mathcal{A}_{b+1,1}(k, [b])\} = \binom{n - b - 1}{k - 1} (k - 1)! \frac{1}{(n)_k} \mathbb{E}(\hat{D}_1 \hat{D}_2 \cdots \hat{D}_k). \tag{2.12}$$

The result now follows from (2.10)–(2.12).

To obtain the distribution of $s_n^{\hat{D}}[a]$, we exploit the following ‘duality’ (see [16]) between the successors and predecessors of $T_n^{\hat{D}}$. For any $A \subseteq [n]$ and $B \subseteq [n] \setminus A$, with $|A| = a$ and $|B| = b$,

$$B \subseteq [n] \setminus \mathcal{S}_{T_n^{\hat{D}}}(A) \quad \text{if and only if} \quad A \subseteq [n] \setminus \mathcal{P}_{T_n^{\hat{D}}}(B), \tag{2.13}$$

since each of the inclusions in (2.13) is equivalent to the statement that no vertex in B is ‘reachable’ from A , and, therefore,

$$\mathbb{P}\{B \subseteq [n] \setminus \mathcal{S}_{T_n^{\hat{D}}}(A)\} = \mathbb{P}\{A \subseteq [n] \setminus \mathcal{P}_{T_n^{\hat{D}}}(B)\}. \tag{2.14}$$

We also note that, since the distribution of $G_n^{\hat{D}}$ is invariant under relabelling of the vertices, we have

$$\mathbb{P}\{B \subseteq [n] \setminus \mathcal{S}_{T_n^{\hat{D}}}(A)\} = \mathbb{P}\{B' \subseteq [n] \setminus \mathcal{S}_{T_n^{\hat{D}}}(A)\} \tag{2.15}$$

for any other $B' \subseteq [n] \setminus A$ such that $|B'| = b$. It follows from (2.1) and (2.15) that, since the random variable $n - s_n^{\hat{D}}(A)$ can be expressed as a sum of indicators of events that a given vertex is not a successor of A , we have

$$(n - a)_b \mathbb{P}\{B \subseteq [n] \setminus \mathcal{S}_{T_n^{\hat{D}}}(A)\} = \mathbb{E}_b(n - s_n^{\hat{D}}(A)) = \mathbb{E}_b(n - s_n^{\hat{D}}[a]), \tag{2.16}$$

where $\mathbb{E}_t(X)$ denotes the t th factorial moment of the random variable X . By a similar argument, we also have

$$(n - b)_a \mathbb{P}\{A \subseteq [n] \setminus \mathcal{P}_{T_n^{\hat{D}}}(B)\} = \mathbb{E}_a(n - p_n^{\hat{D}}(B)) = \mathbb{E}_a(n - p_n^{\hat{D}}[b]). \tag{2.17}$$

Equations (2.14), (2.16), and (2.17) give us the following useful relation between the number of successors of the vertices $A \subseteq [n]$, $|A| = a$, and the number of predecessors of the vertices

$B \subseteq [n] \setminus A$, $|B| = b$, in terms of factorial moments:

$$\frac{\mathbb{E}_a(n - p_n^{\hat{D}}[b])}{(n - b)_a} = \frac{\mathbb{E}_b(n - s_n^{\hat{D}}[a])}{(n - a)_b}. \tag{2.18}$$

Using Theorem 2.1 and (2.16)–(2.18), we obtain the exact formulae for the distribution and expected value of the number of successors of the vertices labelled $1, 2, \dots, a$ in $G_n^{\hat{D}}$.

Theorem 2.2. *Let $s_n^{\hat{D}}[a]$ denote the number of successors of a given a -element subset of the vertex set in $G_n^{\hat{D}}$. Then, for $0 \leq s \leq n - a$,*

$$\begin{aligned} \mathbb{P}\{s_n^{\hat{D}}[a] = a + s\} &= \binom{n - a}{s} \sum_{j=0}^s \binom{s}{j} (-1)^j \sum_{t=0}^{s-j} \binom{s - j}{t} \frac{n + j - a - s}{n + j - a - s + t} \\ &\quad \times \sum_{i=0}^{n+j-a-s} \frac{\binom{n+j-a-s}{i}}{\binom{n}{t+i}} \mathbb{P}\left\{ \sum_{k=1}^{n+j-a-s+t} \hat{D}_k = i + t \right\}. \end{aligned}$$

Moreover,

$$\begin{aligned} \mathbb{E}(s_n^{\hat{D}}[a]) &= n - \sum_{j=0}^{n-1-a} \frac{(n - a)_{j+1}}{(n)_{j+1}} \frac{n - j}{j + 1} \mathbb{P}\left\{ \sum_{k=1}^{j+1} \hat{D}_k = j \right\} \\ &\quad - \sum_{j=0}^{n-1-a} \frac{(n - a)_{j+1}}{(n)_{j+1}} \mathbb{P}\left\{ \sum_{k=1}^{j+1} \hat{D}_k = j + 1 \right\}. \end{aligned}$$

Proof. Again, it follows from (2.1) that we can assume throughout the proof that the given a -element subset of vertices is $[a] = \{1, 2, \dots, a\}$. The result now follows from inclusion–exclusion, (2.14), and (2.16):

$$\begin{aligned} \mathbb{P}\{s_n^{\hat{D}}[a] = a + s\} &= \mathbb{P}\{n - s_n^{\hat{D}}[a] = n - a - s\} \\ &= \sum_{b=n-a-s}^{n-a} \binom{b}{n - a - s} (-1)^{b-n+a+s} \frac{\mathbb{E}_b(n - s_n^{\hat{D}}[a])}{b!} \\ &= \sum_{b=n-a-s}^{n-a} \binom{n - a}{b} \binom{b}{n - a - s} (-1)^{b-n+a+s} \mathbb{P}\{[a] \subseteq [n] \setminus \mathcal{P}_{T_n^{\hat{D}}}([a + 1, a + b])\} \\ &= \sum_{b=n-a-s}^{n-a} \binom{n - a}{b} \binom{b}{n - a - s} (-1)^{b-n+a+s} \\ &\quad \times \sum_{t=0}^{n-a-b} \binom{n - a - b}{t} \mathbb{P}\{\mathcal{P}_{T_n^{\hat{D}}}([a + 1, a + b]) = [a + 1, a + b + t]\} \\ &= \sum_{b=n-a-s}^{n-a} \binom{n - a}{b} \binom{b}{n - a - s} (-1)^{b-n+a+s} \\ &\quad \times \sum_{t=0}^{n-a-b} \binom{n - a - b}{t} \binom{n - b}{t}^{-1} \mathbb{P}\{p_n^{\hat{D}}[b] = b + t\} \end{aligned} \tag{2.19}$$

$$\begin{aligned}
 &= \binom{n-a}{s} \sum_{j=0}^s \binom{s}{j} (-1)^j \sum_{t=0}^{s-j} \binom{s-j}{t} \frac{n+j-a-s}{n+j-a-s+t} \\
 &\quad \times \sum_{i=0}^{n+j-a-s} \frac{\binom{n+j-a-s}{i}}{\binom{n}{t+i}} \mathbb{P} \left\{ \sum_{k=1}^{n+j-a-s+t} \hat{D}_k = i+t \right\}.
 \end{aligned}$$

The formula for the expected value of $s_n^{\hat{D}}[a]$ follows immediately from Theorem 2.1 since by (2.18) we have

$$\mathbb{E}(s_n^{\hat{D}}[a]) = n - \sum_{j=0}^{n-1-a} \frac{(n-a)_{j+1}}{(n-1)_j} \mathbb{P}\{p_n^{\hat{D}}[1] = j+1\}. \tag{2.20}$$

3. Examples

In this section we consider the application of our main results to the digraph structure of random mappings with preferential attachment and anti-preferential attachment.

3.1. Random mappings with preferential attachment

The definition of random mappings with preferential attachment is based upon the following sequential urn scheme. Start with n urns, numbered 1 to n and each containing a ball with weight ρ , where $\rho > 0$ is fixed. Urns are sequentially selected and balls are added to the urns as follows. At each stage, for $1 \leq i \leq n$, the probability that urn i is selected is proportional to the weight of the balls in urn i . If urn i is selected at some stage then a ball of weight 1 is added to urn i before the next urn selection is made. We define the random mapping $T_n^\rho : [n] \rightarrow [n]$ based on the first n selections in the scheme defined above. Specifically, we define

$$T_n^\rho(i) = j$$

for $1 \leq i, j \leq n$ if on the i th selection, urn j is chosen.

In order to determine the distributions of the numbers of predecessors and successors of a set of vertices in $G_n^\rho \equiv G(T_n^\rho)$, we use the fact (see [11]) that T_n^ρ has the same distribution as a random mapping $T_n^{\hat{D}(\rho, n)}$, with exchangeable in-degree sequence $\hat{D}(\rho, n) = (\hat{D}_{1,n}^\rho, \hat{D}_{2,n}^\rho, \dots, \hat{D}_{n,n}^\rho)$ defined as follows. Suppose that $D_1^\rho, D_2^\rho, \dots$ are i.i.d. random variables with a generalized negative binomial distribution given by

$$\mathbb{P}\{D_1^\rho = k\} = \frac{\rho^{(k)}}{k!} \left(\frac{\rho}{1+\rho}\right)^\rho \left(\frac{1}{1+\rho}\right)^k \quad \text{for } k = 0, 1, \dots, \tag{3.1}$$

where ρ is a positive parameter and $x^{(j)} = x(x+1)\dots(x+j-1)$. Then, for $n \geq 1$, let $\hat{D}(\rho, n) = (\hat{D}_{1,n}^\rho, \hat{D}_{2,n}^\rho, \dots, \hat{D}_{n,n}^\rho)$ be a sequence of variables with joint distribution given by

$$\mathbb{P}\{\hat{D}_{i,n}^\rho = d_i, 1 \leq i \leq n\} = \mathbb{P}\left\{D_i^\rho = d_i, 1 \leq i \leq n \mid \sum_{i=1}^n D_i^\rho = n\right\}. \tag{3.2}$$

We note that, since $T_n^\rho \stackrel{D}{\sim} T_n^{\hat{D}(\rho, n)}$, we also have

$$s_n^\rho[a] \stackrel{D}{\sim} s_n^{\hat{D}(\rho, n)}[a], \quad p_n^\rho[b] \stackrel{D}{\sim} p_n^{\hat{D}(\rho, n)}[b],$$

where $s_n^\rho[a]$ and $p_n^\rho[b]$ denote the number of successors and the number of predecessors, respectively, of given a -element and b -element vertex subsets in G_n^ρ . Thus, we can determine the distributions of $s_n^\rho[a]$ and $p_n^\rho[b]$ by appealing to Theorems 2.1 and 2.2 applied to $G_n^{\hat{D}(\rho, n)}$.

Theorem 3.1. Suppose that $1 \leq b \leq n$, and let $p_n^\rho[b]$ denote the number of predecessors of the vertices from a given b -element subset of the vertex set in G_n^ρ . Then, for $0 \leq t \leq n - b$,

$$\mathbb{P}\{p_n^\rho[b] = b + t\} = \binom{n-b}{t} \frac{b}{b+t} \frac{(\rho(b+t))^{(t)} (\rho(n-b-t))^{(n-b-t)}}{(\rho n)^{(n-b)}}.$$

Moreover,

$$\mathbb{E}(p_n^\rho[b]) = b \sum_{k=0}^{n-b} \rho^k \frac{(n-b)_k}{(\rho n)^{(k)}}.$$

Proof. Suppose that $1 \leq b \leq n$ are fixed integers and that $\rho > 0$ is fixed. Since $T_n^\rho \stackrel{D}{\sim} T_n^{\hat{D}(\rho, n)}$, where the distribution of $\hat{D}(\rho, n)$ is given by (3.2), we have, for $0 \leq t \leq n - b$,

$$\mathbb{P}\{p_n^\rho[b] = b + t\} = \mathbb{P}\{p_n^{\hat{D}(\rho, n)}[b] = b + t\}. \tag{3.3}$$

Next we note that it follows from the definition of the i.i.d. variables $D_1^\rho, D_2^\rho, \dots$ that, for any integers $k \geq 0$ and $m \geq 1$,

$$\mathbb{P}\left\{\sum_{i=1}^m D_i^\rho = k\right\} = \frac{(\rho m)^{(k)}}{k!} \left(\frac{\rho}{1+\rho}\right)^{\rho m} \left(\frac{1}{1+\rho}\right)^k. \tag{3.4}$$

So, it follows from (3.2) and (3.4) that

$$\begin{aligned} & \mathbb{P}\left\{\sum_{i=1}^{b+t} \hat{D}_{i,n}^\rho = j + t\right\} \\ &= \frac{\mathbb{P}\{\sum_{i=1}^{b+t} D_i^\rho = j + t\} \mathbb{P}\{\sum_{i=b+t+1}^n D_i^\rho = n - j - t\}}{\mathbb{P}\{\sum_{i=1}^n D_i^\rho = n\}} \\ &= \binom{n}{j+t} \frac{(\rho(b+t))^{(j+t)} (\rho(n-b-t))^{(n-j-t)}}{(\rho n)^{(n)}} \\ &= \binom{n}{j+t} \frac{(\rho(b+t))^{(t)} (\rho(n-b-t))^{(n-b-t)}}{(\rho n)^{(n)}} \\ & \quad \times (t + \rho(b+t))^{(j)} (n-b-t + \rho(n-b-t))^{(b-j)}. \end{aligned} \tag{3.5}$$

The thesis now follows from (3.3), (3.5), Theorem 2.1, and the identity

$$\sum_{j=0}^b \binom{b}{j} x^{(j)} y^{(b-j)} = (x + y)^{(b)}$$

(which follows from the fact that rising factorials are Sheffer sequences of binomial type; see [28]). The formula for the expectation of $p_n^\rho[b]$ follows from Theorem 2.1 since (see [10])

$$\mathbb{E}(\hat{D}_{1,n}^\rho \hat{D}_{2,n}^\rho \cdots \hat{D}_{k,n}^\rho) = \rho^k \frac{(n)_k}{(\rho n)^{(k)}}.$$

We note that the distribution above is closely related to quasi-hypergeometric distribution I (see [7], [18], and the references therein). One can easily check that the results for the moments obtained for this distribution formally coincide with the results for the number of predecessors, but the parameters in our model are outside the range of parameters usually given for these distributions.

In order to describe the asymptotic behaviour of $\mathbb{E}(p_n^\rho[b])$ as $n \rightarrow \infty$, we introduce a function $f: \mathbb{R}_+^2 \rightarrow \mathbb{R}$ which is defined by

$$f(c, \alpha) = c \exp\left(\frac{\alpha c^2}{2}\right) \int_{\alpha c}^\infty \exp\left(\frac{-u^2}{2\alpha}\right) du. \tag{3.6}$$

We note that it is an exercise in calculus to show that, for fixed $\alpha > 0$, $f(c, \alpha)$ increases to 1 as $c \rightarrow \infty$. Using this fact and straightforward asymptotic calculations, we obtain the following result from Theorem 3.1.

Corollary 3.1. *Suppose that $\rho > 0$ is fixed.*

(i) *If $b = o(\sqrt{n})$ then*

$$\mathbb{E}(p_n^\rho[b]) \sim b \sqrt{\frac{\rho\pi n}{2(1+\rho)}}.$$

(ii) *If $c > 0$ is fixed and $b = \lfloor c\sqrt{n} \rfloor$, then*

$$\mathbb{E}(p_n^\rho[b]) \sim f\left(c, \frac{\rho}{1+\rho}\right)n,$$

where f is as defined by (3.6).

Theorem 3.2. *Let $s_n^\rho[a]$ denote the number of successors of the vertices from a given a -element subset of the vertex set in G_n^ρ . Then, for $0 \leq s \leq n - a$,*

$$\mathbb{P}\{s_n^\rho[a] = a + s\} = \binom{n-a}{s} \frac{(a+s)(\rho+1)-1}{(\rho n)^{(a+s)}} \sum_{j=0}^s \binom{s}{j} (-1)^j (\rho(a+s-j))^{(a+s-1)}.$$

Moreover,

$$\mathbb{E}(s_n^\rho[a]) = n - \sum_{j=1}^{n-a} \binom{n-a}{j} \frac{(\rho j)^{(j-1)}(\rho(n-j))^{(n-j)}}{(\rho n)^{(n-1)}}.$$

Proof. The distribution for $s_n^\rho[a]$ is obtained from (2.19) and Theorem 3.1 via some manipulation of the summations. For completeness, the main steps of these manipulations are included for the reader:

$$\begin{aligned} \mathbb{P}\{s_n^\rho[a] = a + s\} &= \binom{n-a}{s} \sum_{b=n-a-s}^{n-a} \binom{s}{n-a-b} (-1)^{b-n+a+s} \\ &\quad \times \sum_{t=0}^{n-a-b} \binom{n-a-b}{t} \frac{b}{b+t} \frac{(\rho(b+t))^{(t)}(\rho(n-b-t))^{(n-b-t)}}{(\rho n)^{(n-b)}} \\ &= \binom{n-a}{s} \sum_{j=0}^s \binom{s}{j} (-1)^j \sum_{t=0}^{s-j} \binom{s-j}{t} \frac{n+j-a-s}{n-a-t} \\ &\quad \times \frac{(\rho(n-a-t))^{(s-j-t)}(\rho(a+t))^{(a+t)}}{(\rho n)^{(a+s-j)}} \\ &= \binom{n-a}{s} \sum_{t=0}^s \binom{s}{t} (\rho(a+t))^{(a+t)} \\ &\quad \times \sum_{j=0}^{s-t} \binom{s-t}{j} (-1)^{s-t-j} \frac{n+j-a-s}{n-a-t} \frac{(\rho(n-a-t))^{(j)}}{(\rho n)^{(a+t+j)}} \end{aligned}$$

$$\begin{aligned}
 &= \binom{n-a}{s} \sum_{t=0}^s \binom{s}{t} \frac{(-1)^{s-t} (\rho(a+t))^{(a+t)}}{(\rho n)^{(a+t)}} \\
 &\quad \times \left[\sum_{j=0}^{s-t} \binom{s-t}{j} (-1)^j \frac{(\rho(n-a-t))^{(j)}}{(\rho n + a + t)^{(j)}} \right. \\
 &\quad \left. + \frac{\rho(s-t)}{\rho n + a + t} \sum_{j=1}^{s-t} \binom{s-t-1}{j-1} (-1)^{j-1} \frac{(\rho(n-a-t)+1)^{(j-1)}}{(\rho n + a + t + 1)^{(j-1)}} \right] \\
 &= \binom{n-a}{s} \sum_{t=0}^s \binom{s}{t} \frac{(-1)^{s-t} (\rho(a+t))^{(a+t)}}{(\rho n)^{(a+t)}} \\
 &\quad \times \left[\sum_{j=0}^{s-t} \binom{s-t}{j} (-1)^j \frac{(\rho(n-a-t))^{(j)}}{(\rho n + a + t)^{(j)}} \right. \\
 &\quad \left. + \frac{\rho(s-t)}{\rho n + a + t} \sum_{j=0}^{s-t-1} \binom{s-t-1}{j} (-1)^j \frac{(\rho(n-a-t)+1)^{(j)}}{(\rho n + a + t + 1)^{(j)}} \right].
 \end{aligned}$$

Now, using the identity

$$\sum_{j=0}^N \binom{N}{j} (-1)^j \frac{x^{(j)}}{(x+y)^{(j)}} = \frac{y^{(N)}}{(x+y)^{(N)}},$$

which follows from Gauss’s hypergeometric theorem (see [1], and also Frisch’s identity in [9]), we obtain

$$\sum_{j=0}^{s-t} \binom{s-t}{j} (-1)^j \frac{(\rho(n-a-t))^{(j)}}{(\rho n + a + t)^{(j)}} = \frac{((1+\rho)(a+t))^{(s-t)}}{(\rho n + a + t)^{(s-t)}}.$$

Similarly,

$$\sum_{j=0}^{s-t-1} \binom{s-t-1}{j} (-1)^j \frac{(\rho(n-a-t)+1)^{(j)}}{(\rho n + a + t + 1)^{(j)}} = \frac{((1+\rho)(a+t))^{(s-t-1)}}{(\rho n + a + t + 1)^{(s-t-1)}}.$$

Hence,

$$\begin{aligned}
 &\mathbb{P}\{s_n^\rho[a] = a + s\} \\
 &= \binom{n-a}{s} \sum_{t=0}^s \binom{s}{t} \frac{(-1)^{s-t} (\rho(a+t))^{(a+t)}}{(\rho n)^{(a+t)}} \\
 &\quad \times \frac{((1+\rho)(a+t))^{(s-t-1)}}{(\rho n + a + t)^{(s-t)}} [s + a + \rho(a+t) - 1 + \rho(s-t)] \\
 &= \binom{n-a}{s} \frac{(1+\rho)(s+a) - 1}{(\rho n)^{(a+t)}} \sum_{j=0}^s \binom{s}{j} (-1)^j (\rho(a+s-j))^{(a+s-1)}.
 \end{aligned}$$

Finally, the formula for the expectation of $s_n^\rho[a]$ follows from (2.20) and Theorem 3.1 for $b = 1$.

Corollary 3.2. *Suppose that $a = o(n)$ and $\rho > 0$ is fixed. Then*

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E}(s_n^\rho[a])}{n} = 0.$$

Proof. Suppose that $\rho > 0$ is fixed and $a = o(n)$. Then we obtain, from (2.20),

$$\mathbb{E}(s_n^\rho[a]) = n - (n - a) \sum_{j=0}^{n-a-1} \frac{(n - a - 1)_j}{(n - 1)_j} \mathbb{P}\{p_n^\rho[1] = j + 1\}. \tag{3.7}$$

It also follows from Theorem 3.1 that, for fixed $\rho > 0$ and any fixed integer $j > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P}\{p_n^\rho[1] = j + 1\} = \frac{\rho}{(\rho + 1)j + \rho} \binom{(\rho + 1)(j + 1)}{j} \left(\frac{1}{1 + \rho}\right)^j \left(\frac{\rho}{\rho + 1}\right)^{(\rho + 1)j + \rho - j}.$$

So, for any $\varepsilon > 0$, there is some integer $\ell(\varepsilon) > 0$ such that, for all large n ,

$$\mathbb{P}\{p_n^\rho[1] \leq \ell(\varepsilon)\} \geq 1 - \varepsilon, \tag{3.8}$$

and it follows from (3.7) and (3.8) that

$$\begin{aligned} \mathbb{E}(s_n^\rho[a]) &\leq n - (n - a) \frac{(n - a - 1)_{\ell(\varepsilon)}}{(n - 1)_{\ell(\varepsilon)}} \mathbb{P}\{p_n^\rho[1] \leq \ell(\varepsilon)\} \\ &\leq n - (n - a) \frac{(n - a - 1)_{\ell(\varepsilon)}}{(n - 1)_{\ell(\varepsilon)}} (1 - \varepsilon). \end{aligned} \tag{3.9}$$

The result follows since we obtain, from (3.9),

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E}(s_n^\rho[a])}{n} \leq \varepsilon.$$

3.2. Random mappings with anti-preferential attachment

In this section we consider the digraph structure of random mappings with anti-preferential attachment. Random mappings with anti-preferential attachment were introduced in [11] and can be described as follows. Start with n urns, numbered 1 to n , each containing $m \geq 1$ balls, where m is a fixed integer. Balls are removed one at a time and at random from the urns in such a way that the probability that a ball is removed from urn j in a given draw is equal to the number of balls in urn j before the draw divided by the total number of balls still in the urns before the draw. Then, for $1 \leq j \leq n$, we define $T_n^m(i) = j$ if a ball is removed from the j th urn on the i th draw.

As in the case of random mappings with preferential attachment, it is known (see [11]) that T_n^m has the same distribution as a random mapping $T_n^{D(m,n)}$, with exchangeable in-degree sequence $\hat{D}(m, n) = (\hat{D}_{1,n}^m, \hat{D}_{2,n}^m, \dots, \hat{D}_{n,n}^m)$ defined as follows. Suppose that D_1^m, D_2^m, \dots are i.i.d. random variables with a binomial distribution $\text{Bin}(m, p)$. Then, for $n \geq 1$, let $\hat{D}(m, n) = (\hat{D}_{1,n}^m, \hat{D}_{2,n}^m, \dots, \hat{D}_{n,n}^m)$ be a sequence of variables with joint distribution given by

$$\mathbb{P}\{\hat{D}_{i,n}^m = d_i, 1 \leq i \leq n\} = \mathbb{P}\left\{D_i^m = d_i, 1 \leq i \leq n \mid \sum_{i=1}^n D_i^m = n\right\}. \tag{3.10}$$

Again, since $T_n^m \stackrel{D}{\sim} T_n^{\hat{D}(m,n)}$, we also have

$$s_n^m[a] \stackrel{D}{\sim} s_n^{\hat{D}(m,n)}[a], \quad p_n^m[b] \stackrel{D}{\sim} p_n^{\hat{D}(m,n)}[b],$$

where $s_n^m[a]$ and $p_n^m[b]$ denote the number of successors and the number of predecessors, respectively, of given a -element and b -element vertex subsets in $G_n^m \equiv G(T_n^m)$. Thus, as in the case of preferential mappings, we can determine the distributions of $s_n^m[a]$ and $p_n^m[b]$ by appealing to Theorems 2.1 and 2.2 applied to $G_n^{\hat{D}(m,n)}$.

Theorem 3.3. *Let $p_n^m[b]$ denote the number of predecessors of the vertices from a given b -element subset of the vertex set in G_n^m . Then, for $0 \leq t \leq n - b$,*

$$\mathbb{P}\{p_n^m[b] = b + t\} = \binom{n - b}{t} \frac{b}{b + t} \frac{(m(b + t))_t (m(n - b - t))_{n - b - t}}{(mn)_{n - b}}.$$

Moreover,

$$\mathbb{E}(p_n^m[b]) = b \sum_{k=0}^{n - b} m^k \frac{(n - b)_k}{(nm)_k}.$$

Proof. Suppose that $1 \leq b \leq n$ and $m \geq 1$ are fixed integers. Since $T_n^m \stackrel{D}{\sim} T_n^{\hat{D}(m,n)}$, where the distribution of $\hat{D}(m, n)$ is given by (3.10), we have, for $0 \leq t \leq n - b$,

$$\mathbb{P}\{p_n^m[b] = b + t\} = \mathbb{P}\{p_n^{\hat{D}(m,n)}[b] = b + t\}. \tag{3.11}$$

Next, we note that

$$\begin{aligned} & \mathbb{P}\left\{\sum_{i=1}^{b+t} \hat{D}_{i,n}^m = j + t\right\} \\ &= \frac{\mathbb{P}\{\sum_{i=1}^{b+t} D_i^m = j + t\} \mathbb{P}\{\sum_{i=b+t+1}^n D_i^m = n - j - t\}}{\mathbb{P}\{\sum_{i=1}^n D_i^m = n\}} \\ &= \frac{\binom{(b+t)m}{j+t} p^{j+t} (1 - p)^{(b+t)m - j - t} \binom{(n - b - t)m}{n - j - t} p^{n - j - t} (1 - p)^{(n - b - t)m - n + j + t}}{\binom{nm}{n} p^n (1 - p)^{nm - n}} \\ &= \binom{n}{j + t} \frac{((b + t)m)_{j+t} ((n - b - t)m)_{n - j - t}}{(nm)_n} \\ &= \binom{n}{j + t} \frac{((b + t)m)_t ((n - b - t)m)_{n - b - t}}{(nm)_n} \\ & \quad \times ((b + t)m - t)_j ((n - b - t)m - n + b + t)_{b - j}. \end{aligned}$$

Equation (3.11), Theorem 2.1, and Vandermonde’s identity,

$$\sum_{j=0}^b \binom{b}{j} (x)_j (y)_{b - j} = (x + y)_b$$

(which follows from the fact that the falling factorials are also Sheffer sequences of binomial type), lead immediately to the thesis. The formula for the expectation of $p_n^m[b]$ follows from Theorem 2.1 since (see [10])

$$\mathbb{E}(\hat{D}_{1,n}^m \hat{D}_{2,n}^m \cdots \hat{D}_{k,n}^m) = m^k \frac{(n)_k}{(nm)_k}.$$

We note that the distribution above is closely related to quasi-Pólya distribution I (see [7], [18], and the references therein). One can easily check that the results for the moments obtained for this distribution formally coincide with the results for the number of predecessors, but the parameters in our model are outside the range of parameters usually given for these distributions.

Using straightforward asymptotic calculations, we obtain the following result from Theorem 3.3.

Corollary 3.3. *Suppose that $m \geq 2$ is a fixed integer.*

(i) *If $b = o(\sqrt{n})$ then*

$$\mathbb{E}(p_n^m[b]) \sim b \sqrt{\frac{mn\pi}{2(m-1)}}.$$

(ii) *If $c > 0$ is fixed and $b = \lfloor c\sqrt{n} \rfloor$, then*

$$\mathbb{E}(p_n^m[b]) \sim f\left(c, \frac{m}{m-1}\right)n,$$

where f is as defined by (3.6).

Theorem 3.4. *Let $s_n^m[a]$ denote the number of successors of the vertices from a given a -element subset of the vertex set in G_n^m . Then, for $0 \leq s \leq n - a$,*

$$\mathbb{P}\{s_n^m[a] = a + s\} = \binom{n-a}{s} \sum_{j=0}^s \binom{s}{j} (-1)^j \frac{(m-1)(a+s) + j}{(mn)_{a+s}} (m(a+s-j))_{a+s-1}.$$

Moreover,

$$\mathbb{E}(s_n^m[a]) = n - \sum_{j=1}^{n-a} \binom{n-a}{j} \frac{(mj)_{j-1} (m(n-j))_{n-j}}{(mn)_{n-1}}.$$

Proof. First, we note that by (2.19) we have

$$\begin{aligned} \mathbb{P}\{s_n^m[a] = a + s\} &= \sum_{b=n-a-s}^{n-a} \binom{n-a}{b} \binom{b}{n-a-s} (-1)^{b-n+a+s} \\ &\quad \times \sum_{t=0}^{n-a-b} \binom{n-a-b}{t} \binom{n-b}{t}^{-1} \mathbb{P}\{p_n^m[b] = b + t\}. \end{aligned}$$

Next, applying Theorem 3.3 and manipulating the summations above as in the proof of Theorem 3.2, we obtain

$$\begin{aligned} \mathbb{P}\{s_n^m[a] = a + s\} &= \binom{n-a}{s} \sum_{t=0}^s \binom{s}{t} \frac{(-1)^{s-t} (m(a+t))_{a+t}}{(mn)_{a+t}} \\ &\quad \times \left[\sum_{j=0}^{s-t} \binom{s-t}{j} (-1)^j \frac{(m(n-a-t))_j}{(mn-a-t)_j} \right. \\ &\quad \left. + \frac{m(s-t)}{mn-a-t} \sum_{j=0}^{s-t-1} \binom{s-t-1}{j} (-1)^j \frac{(m(n-a-t)-1)_j}{(mn-a-t-1)_j} \right]. \end{aligned}$$

Now, using the identity

$$\sum_{j=0}^N \binom{N}{j} (-1)^j \frac{(x)_j}{(y+x)_j} = \frac{(y)_N}{(x+y)_N},$$

which also follows from Gauss’s hypergeometric theorem (see [1] and also [9]), we obtain

$$\sum_{j=0}^{s-t} \binom{s-t}{j} (-1)^j \frac{(m(n-a-t))_j}{(mn-a-t)_j} = \frac{((m-1)(a+t))_{s-t}}{(mn-a-t)_{s-t}}.$$

Similarly,

$$\begin{aligned} & \frac{m(s-t)}{mn-a-t} \sum_{j=0}^{s-t-1} \binom{s-t-1}{j} (-1)^j \frac{(m(n-a-t)-1)_j}{(mn-a-t-1)_j} \\ &= \frac{m(s-t)}{mn-a-t} \frac{((m-1)(a+t))_{s-t-1}}{(mn-a-t-1)_{s-t-1}} \\ &= m(s-t) \frac{((m-1)(a+t))_{s-t-1}}{(mn-a-t)_{s-t}}. \end{aligned}$$

Hence,

$$\begin{aligned} & \mathbb{P}\{s_n^m[a] = a+s\} \\ &= \binom{n-a}{s} \sum_{t=0}^s \binom{s}{t} \frac{(-1)^{s-t} (m(a+t))_{a+t}}{(mn)_{a+t}} \\ & \quad \times \frac{((m-1)(a+t))_{s-t-1}}{(mn-a-t)_{s-t}} [(m-1)(a+t) - s + t + 1 + m(s-t)] \\ &= \binom{n-a}{s} \frac{(m-1)(a+s) + 1}{(mn)_{a+s}} \sum_{j=0}^s \binom{s}{j} (-1)^j (m(a+s-j))_{a+s-1}. \end{aligned}$$

Finally, the formula for the expectation of $s_n^m[a]$ follows from (2.20) and Theorem 3.3 for $b = 1$.

Corollary 3.4. *Suppose that $a = o(n)$ and $m > 1$ is a fixed positive integer. Then*

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E}(s_n^m[a])}{n} = 0.$$

Proof. The proof of the result is analogous to the proof of Corollary 3.2 and follows from the fact that, in this case, for fixed $j \geq 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P}\{p_n^m[1] = j+1\} = \frac{1}{(m-1)^j} \binom{mj+m}{m} \frac{1}{j+1} \left(1 - \frac{1}{m}\right)^{mj+m}.$$

4. Final remarks and discussion

In this paper we have shown that, for a random mapping, $T_n^{\hat{D}}$, with exchangeable in-degrees, the exact distributions and expected values of the numbers of successors and of predecessors of a set of vertices A in $G_n^{\hat{D}}$ can be expressed in terms of the joint distribution of the in-degree variables $\hat{D}_1, \hat{D}_2, \dots, \hat{D}_n$. Since the uniform random mapping T_n is a special case of the

model $T_n^{\hat{D}}$, our results immediately generalise the known results for T_n . As an application of these results, we have also obtained the distributions of these variables for the important special examples, T_n^ρ and T_n^m , of random mappings with preferential and anti-preferential attachment, respectively. We note that these results generalise earlier work in [10] on the numbers of successors and predecessors of a single vertex in T_n^ρ and T_n^m , respectively. However, different techniques were required in this paper and the results obtained above, along with the results in [10], lead to interesting identities when we consider the predecessors and successors of a single vertex. In addition, we have characterised for T_n^ρ and T_n^m the asymptotic behaviour of the expected values of the numbers of predecessors and successors. We note that, for these models, one can also consider the asymptotic distributions of these variables, but these calculations are outside the scope of this paper.

In this paper we also characterised the asymptotic behaviour of the expected value for the numbers of predecessors and of successors in a random mapping because this behaviour is closely related to the threshold behaviour of epidemic processes on random mapping digraphs. Specifically, suppose that $\tilde{T}_n : [n] \rightarrow [n]$ is a random mapping which is represented by the random digraph $G(\tilde{T}_n)$ and suppose that some a -element subset A of the vertices $[n]$ are initially infected with a contagious disease and the disease spreads to other elements of $[n]$ along arcs of the random digraph $G(\tilde{T}_n)$. Three types of epidemic process can be considered depending upon the way in which the disease spreads. If it spreads only forward, that is, in the direction of orientation of the arcs, it is called a *direct epidemic process* (DEP). If it spreads only backward, it is called an *inverse epidemic process* (IEP). If it spreads in both directions, the process is termed a *two-sided epidemic process* (TEP). Clearly, the total numbers of elements which are eventually infected in DEP, IEP, and TEP on a digraph $G(\tilde{T}_n)$ are given as the cardinalities of sets of all successors, all predecessors of elements from A , and the set of all vertices in the connected components to which elements from A belong, respectively. Following Gertsbakh [9], a function $h_{IEP} = h_{IEP}(\tilde{T}_n)$ is called the *threshold* for IEP on $G(\tilde{T}_n)$ representing a random mapping \tilde{T}_n if, for a fixed γ , $0 < \gamma < 1$,

$$\lim_{n \rightarrow \infty} \mathbb{P}\{p_{\tilde{T}_n}[h] > \gamma n\} = \begin{cases} 0 & \text{if } h = o(h_{IEP}), \\ 1 & \text{if } h_{IEP} = o(h). \end{cases}$$

Similarly, threshold functions h_{DEP} and h_{TEP} can be defined for DEP and TEP, respectively. We note that, for uniform random mappings, the exact and asymptotic distributions for the number of predecessors of elements from a given set of vertices for the uniform models were given by Burtin [6]. The respective results for successors were given by Berg [3], [4] and Pittel [25], who also gave the results related to TEP. The solution to the main problem stated by Gertsbakh, namely, to find the thresholds for DEP, IEP, and TEP, follows immediately from these results (some results concerning the thresholds were given independently by Mutafchiev [21], [22]). In the case of random mappings with preferential or anti-preferential attachment, we can conclude from the asymptotic results for the expected value of the numbers of successors and predecessors that, when the parameters ρ and m are fixed, we obtain, for both G_n^ρ and G_n^m ,

- $h_{DEP} = n$, i.e. there is no proper threshold function for DEP on G_n^ρ and G_n^m ,
- $h_{IEP} = \sqrt{n}$ for both G_n^ρ and G_n^m .

It is also known (see [11]) that, for both G_n^ρ and G_n^m , with ρ and m fixed, the size of the component that contains vertex 1 is $O(n)$. So, for both G_n^ρ and G_n^m , we have

- $h_{TEP} = 1$, i.e. there is no proper threshold function for TEP on G_n^ρ and G_n^m .

In light of the results obtained in this paper, it would be interesting to investigate in more detail the asymptotic behaviour of the distributions of the numbers of predecessors and of successors in the random digraphs G_n^ρ and G_n^m . In particular, we would like to study how the limiting distributions depend on the parameters ρ and m and on the size of the set of ‘infected’ vertices as $n \rightarrow \infty$. For example, complementary results (see [12]) on the component structure of G_n^ρ suggest that the limiting distributions for $p_n^\rho[b]$ and $s_n^\rho[a]$ will depend on whether $\rho n \rightarrow \infty$ or $\rho n \rightarrow \beta > 0$ or $\rho n \rightarrow \beta = 0$ as $n \rightarrow \infty$ as well as on the size of a and b as a function of n . In the case of random mappings with anti-preferential attachment, there are qualitative changes in the component structure in models which lie in a continuum between T_n^2 and T_n^1 (see [13]) and we would expect that the epidemic process threshold behaviour in this continuum will differ from the threshold behaviour in G_n^m when $m \geq 2$.

Finally, in another direction, we note that the epidemic process results for uniform random mappings were generalized for epidemic processes associated with the evolutionary model mentioned in the introduction. The results obtained for the evolutionary model are interesting for many reasons, e.g. they provide a full answer to questions stated by Islam [14] (see also [2]) and they also lead to a very nice relation between Abel sums and noncentral Stirling numbers as well as to families of interesting discrete distributions. Since our results for mappings with preferential and anti-preferential attachments display a striking formal similarity to results for the uniform model, where powers are replaced by rising or falling factorials, respectively, it is likely that an investigation of an evolutionary model related to the ‘cut’ variables introduced in Section 2 would lead to even more general families of discrete distributions. We also note that it may be possible to obtain exact and asymptotic results for TEP on G_n^ρ and G_n^m , with the help of results given in this paper, although we cannot directly use the relation which was the key to obtain the corresponding results for mappings with independent choices of images (see [17]).

Acknowledgements

The authors are grateful to an anonymous referee for suggestions which improved the exposition of the paper and for pointing out an alternative proof of Lemma 2.2. J. Jaworski acknowledges the support of the Marie Curie Intra-European Fellowship No. 236845 (RANDOMAPP) within the 7th European Community Framework Programme and by the National Science Centre under grant number DEC-2011/01/B/ST1/03943.

References

- [1] BAILEY, W. N. (1935). *Generalised Hypergeometric Series*. Cambridge University Press.
- [2] BALL, F., MOLLISON, D. AND SCALIA-TOMBA, G. (1997). Epidemics with two levels of mixing. *Ann. Appl. Prob.* **7**, 46–89.
- [3] BERG, S. (1981). On snowball sampling, random mappings and related problems. *J. Appl. Prob.* **18**, 283–290.
- [4] BERG, S. (1983). Random contact processes, snowball sampling and factorial series distributions. *J. Appl. Prob.* **20**, 31–46.
- [5] BERG, S. AND MUTAFCHIEV, L. (1990). Random mappings with an attracting center: Lagrangian distributions and a regression function. *J. Appl. Prob.* **27**, 622–636.
- [6] BURTIN, Y. D. (1980). On simple formula for random mappings and its applications. *J. Appl. Prob.* **17**, 403–414.
- [7] CONSUL, P. C. AND FAMOYE, F. (2006). Quasi-probability models. In *Lagrangian Probability Distributions*, Birkhäuser, Boston, pp. 69–91.
- [8] GERTSBAKH, I. B. (1977). Epidemic processes on a random graph: some preliminary results. *J. Appl. Prob.* **14**, 427–438.
- [9] GOULD, H. W. (1972). *Combinatorial Identities: A Standardized Set of Tables Listing 500 Binomial Coefficient Summations*. Henry W. Gould, Morgantown.
- [10] HANSEN, J. C. AND JAWORSKI, J. (2008). Local properties of random mappings with exchangeable in-degrees. *Adv. Appl. Prob.* **40**, 183–205.

- [11] HANSEN, J. C. AND JAWORSKI, J. (2008). Random mappings with exchangeable in-degrees. *Random Structures Algorithms* **33**, 105–126.
- [12] HANSEN, J. C. AND JAWORSKI, J. (2009). A random mapping with preferential attachment. *Random Structures Algorithms* **34**, 87–111.
- [13] HANSEN, J. C. AND JAWORSKI, J. (2013). Structural transition in random mappings. Submitted.
- [14] ISLAM, M. N., O'SHAUGHNESSY, C. D. AND SMITH, B. (1996). A random graph model for the final-size distribution of household infections. *Statist. Medicine* **15**, 837–843.
- [15] JAWORSKI, J. (1990). Random mappings with independent choices of the images. In *Random Graphs '87*, eds M. Karoński, J. Jaworski and A. Ruciński, John Wiley, Chichester, pp. 89–101.
- [16] JAWORSKI, J. (1998). Predecessors in a random mapping. *Random Structures Algorithms* **13**, 501–519.
- [17] JAWORSKI, J. (1999). Epidemic processes on digraphs of random mappings. *J. Appl. Prob.* **36**, 780–798.
- [18] JOHNSON, N. L., KOTZ, S. AND KEMP, A. W. (1992). *Univariate Discrete Distributions*, 2nd edn. John Wiley, New York.
- [19] KOLCHIN, V. F. (1986). *Random Mappings*. Optimization Software, New York.
- [20] MOON, J. W. (1970). *Counting Labelled Trees* (Canad. Math. Monog. 1). Canadian Mathematical Congress, Montreal.
- [21] MUTAFCHIEV, L. (1981). Epidemic processes on random graphs and their threshold function. *Serdica* **7**, 153–159.
- [22] MUTAFCHIEV, L. (1982). A limit distribution related to random mappings and its application to an epidemic process. *Serdica* **8**, 197–203.
- [23] MUTAFCHIEV, L. (1984). On some stochastic problems of discrete mathematics. In *Mathematics and Mathematical Education* (Sunny Beach, 1984), pp. 57–80.
- [24] PITMAN, J. (2001). Random mappings, forest, and subsets associated with Abel-Cayley-Hurwitz multinomial expansions. *Seminaire Lothar. Combin.* **46**, 45pp.
- [25] PITTEL, B. (1983). On distributions related to transitive closures of random finite mappings. *Ann. Prob.* **11**, 428–441.
- [26] PRÜFER, H. (1918). Neuer beweis eines satzes über permutationen. *Arch. Math. Phys.* **27**, 142–144.
- [27] ROSS, S. M. (1981). A random graph. *J. Appl. Prob.* **18**, 309–315.
- [28] ROTA, G. C. *et al.* (1975). *Finite Operator Calculus*. Academic Press.
- [29] STEPANOV, V. E. (1971). Random mappings with a single attracting center. *Theory Prob. Appl.* **16**, 155–162.