Local properties of a random mapping model

Jennie C. Hansen[*] and Jerzy Jaworski[‡]

July 10, 2007

Abstract

In this paper we investigate the 'local' properties of a random mapping model, $T_n^{\hat{D}}$, which maps the set $\{1, 2, ..., n\}$ into itself. The random mapping $T_n^{\hat{D}}$ was introduced in a companion paper [**?**] is constructed using a collection of exchangeable random variables $\hat{D}_1, ..., \hat{D}_n$ which satisfy $\sum_{i=1}^{n} \hat{D}_i = n$. In the random digraph, $G_n^{\hat{D}}$, which represents the mapping $T_n^{\hat{D}}$, the in-degree sequence for the vertices is given by the variables $\hat{D}_1, \hat{D}_2, ..., \hat{D}_n$, and, in some sense, $G_n^{\hat{D}}$ can be viewed as an analogue of the general independent degree models from random graph theory. By local properties we mean the distributions of random mapping characteristics related to a given vertex v of $G_n^{\hat{D}}$ - for example, the numbers of predecessors and successors of v in $G_n^{\hat{D}}$. We show that the distribution of several variables associated with the local structure of $G_n^{\hat{D}}$ can be expressed in terms of expectations of simple functions of $\hat{D}_1, \hat{D}_2, ..., \hat{D}_n$. We also consider two special examples of $T_n^{\hat{D}}$ which correspond to random mappings with preferential and anti-preferential attachment, respectively, and determine, for these examples, exact and asymptotic distributions for the local structure variables considered in this paper.

[*]Actuarial Mathematics and Statistics, Heriot–Watt University,Edinburgh EH14 4AS, UK. E-mail address: J.Hansen@ma.hw.ac.uk

[†]Faculty of Mathematics and Computer Science, Adam Mickiewicz University, Umultowska 87, 61-614 Poznań, Poland. E-mail address: jaworski@amu.edu.pl

[‡]J.Jaworski acknowledges the generous support by the Marie Curie Intra-European Fellowship No. 501863 (RANDIGRAPH) within the 6th European Community Framework Programme.

1 Introduction

The study of random mapping models was initiated independently by several authors in the 1950s (see companion paper [?] and the references therein) and the properties of these models have received much attention in the literature. In particular, these models have been useful as models for epidemic processes, and have natural applications in cryptology (see, for example, [?, ?, ?, ?, ?, ?, ?, ?, ?]). To date, the most widely studied models have been special cases of a general model denoted by $T_{\mathbf{p}(n)}$, which can be defined as follows: Let $[n]$ denote the set of integers $\{1, 2, ..., n\}$ and let $\mathcal{M}_n$ denote the set of all mappings from $[n]$ into $[n]$. For each $n \geq 1$, let $\mathbf{p}(n) = \{p_{ij}(n) : 1 \leq i, j \leq n\}$ be an array such that $p_{ij}(n) \geq 0$ for $1 \leq i, j \leq n$ and $\sum_{j=1}^{n} p_{ij}(n) = 1$ for every $1 \leq i \leq n$, and let $X_1^n, X_2^n, ..., X_n^n$ be independent random variables such that $\Pr\{X_i^n = j\} = p_{ij}(n)$ for all $1 \leq i, j \leq n$. Then the random mapping $T_{\mathbf{p}(n)} : [n] \to [n]$ is defined (in terms of the variables $X_1^n, X_2^n, ..., X_n^n$) by

$$T_{\mathbf{p}(n)}(i) = j \quad \text{iff} \quad X_i^n = j \tag{1.1}$$

for all $1 \leq i, j \leq n$. It follows from (**??**) that the distribution of $T_{\mathbf{p}(n)}$ is given by

$$\Pr\{T_{\mathbf{p}(n)} = f\} = \prod_{i=1}^{n} p_{if(i)}(n) \tag{1.2}$$

for each $f \in \mathcal{M}_n$. Any mapping $f \in \mathcal{M}_n$ can be represented as a directed graph $G(f)$ on a set of vertices labelled $1, 2, ..., n$, such that there is a directed edge from vertex i to vertex j in $G(f)$ if and only if $f(i) = j$. So $G_{\mathbf{p}(n)} \equiv G(T_{\mathbf{p}(n)})$ is a random directed graph on a set of vertices labelled $1, 2, ..., n$ which represents the action of the random mapping $T_{\mathbf{p}(n)}$ on $[n]$. We note that since each vertex in $G_{\mathbf{p}(n)}$ has out-degree 1, the components of $G_{\mathbf{p}(n)}$ consist of directed cycles with directed trees attached. Also, it follows from the definition of $T_{\mathbf{p}(n)}$ that the variables $X_1^n, X_2^n, \ldots X_n^n$ can be interpreted as the independent 'choices' of the vertices $1, 2, \ldots, n$ in the random digraph $G_{\mathbf{p}(n)}$ (see, in addition, Mutafchiev [?] and Jaworski [?]).

The example of $T_{\mathbf{p}(n)}$ which is best understood is the uniform random mapping, $T_n \equiv T_{\mathbf{p}(n)}$, where $p_{ij}(n) = \frac{1}{n}$ for all $1 \leq i, j \leq n$ (see, for example, the monograph by Kolchin [?]). In particular, in the context of applications the asymptotic distributions of variables such as the number of predecessors and the number of successors of a vertex in G_n have received much attention (see [?, ?, ?, ?, ?, ?, ?]). In another direction, Berg, Jaworski, and Mutafchiev (see [**?**, **?**, **?**, **?**]) have investigated the structure of $G_{\mathbf{p}(n)}$ when $\mathbf{p}(n)$ is given by $p_{ii}(n) = q$ for some $0 \leq q \leq 1$ and all $1 \leq i \leq n$, and $p_{ij}(n) = \frac{1-q}{n-1}$ for all $1 \leq i, j \leq n$ such that $i \neq j$. Finally, Aldous, Miermont, and Pitman (see [**?**] and [**?**]) have recently investigated the asymptotic structure of $G_{\mathbf{p}(n)}$, where $\mathbf{p}(n)$ is given by $p_{ij}(n) = p_j(n) > 0$ for all $1 \leq i, j \leq n$, by using an ingenious coding of the mapping $T_{\mathbf{p}(n)}$ as a stochastic process on the interval $[0, 1]$. Their results are closely related to earlier work on the relationship between random mappings and random forests (see Pitman [**?**] and references therein).

The common feature in all the models discussed above is that each vertex in $G_{\mathbf{p}(n)}$ 'chooses' the vertex that it is mapped to independently of the 'choices' made by all other vertices. In this paper we consider the properties a new random mapping model, $T_n^{\hat{D}}$, which was introduced in a companion paper [**?**]. In the model $T_n^{\hat{D}}$ the vertex 'choices' are not necessarily independent. The model is constructed by first specifying the in-degrees $\hat{D}_1, \hat{D}_2, ..., \hat{D}_n$ of the vertices labelled $1, 2, ..., n$, and then selecting a random mapping uniformly from all mappings with the given in-degree sequence $\hat{D}_1, \hat{D}_2, ..., \hat{D}_n$. After defining the model in Section 2, we show in Section 3 that the distributions of many important random mapping statistics for $T_n^{\hat{D}}$ can be computed as expectations of functions of the (random) in-degree sequence $\hat{D}_1, \hat{D}_2, ..., \hat{D}_n$. In Section 4 we apply these results to two special examples - the preferential and anti-preferential attachment models - which turn out to be equivalent to special cases of $T_n^{\hat{D}}$.

2 The model

In order to define the model $T_n^{\hat{D}}$, we adopt the following notation. For $n \geq 1$, 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 digraph $G(f)$ which represents the mapping f, and define $\vec{d}(f) \equiv (d_1(f), ..., d_n(f))$. Also, given a vector $\vec{d} \equiv (d_1, d_2, ..., d_n)$ of non-negative 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 $\vec{d}$.

To define $T_n^{\hat{D}}$, we start with a collection of non-negative integer-valued exchangeable random variables $\hat{D}_1, \hat{D}_2, ..., \hat{D}_n$ such that $\sum_{i=1}^n \hat{D}_i = n$. Given the event $\{\hat{D}_i = d_i, i = 1, 2, ..., n\}$ (with $\Pr\{\hat{D}_i = d_i, i = 1, 2, ..., n\} > 0$), we define the conditional distribution of $T_n^{\hat{D}}$ by

$$\Pr\{T_n^{\hat{D}} = f \mid \hat{D}_i = d_i, i = 1, 2, ..., n\} = \begin{cases} \frac{\prod_{i=1}^n d_i!}{n!} & \text{if} \quad d_i(f) = d_i, i = 1, 2, ..., n \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

In other words, *given* $(\hat{D}_1, \hat{D}_2, ..., \hat{D}_n) = (d_1, d_2, ..., d_n) = \vec{d}$, $T_n^{\hat{D}}$ is uniformly distributed over $\mathcal{M}_n(\vec{d})$. It follows from (**??**) that for any $f \in \mathcal{M}_n$,

$$\Pr\{T_n^{\hat{D}} = f\} = \frac{\prod_{i=1}^n (d_i(f))!}{n!} \Pr\big\{\hat{D}_i = d_i(f), 1 \leq i \leq n\big\}. \tag{2.2}$$

Given the random mapping $T_n^{\hat{D}}$, let $G_n^{\hat{D}} \equiv G(T_n^{\hat{D}})$ denote the random digraph on n labelled vertices which represents $T_n^{\hat{D}}$. We note that it follows from the exchangeability of the variables $\hat{D}_1, \hat{D}_2, ..., \hat{D}_n$ and (**??**) that, for any permutation $\sigma : [n] \to [n]$, we have

$$\sigma \circ T_n^{\hat{D}} \overset{d}{\sim} T_n^{\hat{D}} \circ \sigma \overset{d}{\sim} T_n^{\hat{D}}.$$

In other words, the distribution of the corresponding digraph $G_n^{\hat{D}}$ is invariant under re-labelling of the vertices of the graph.

An important class of examples can constructed as follows. Suppose that $D_1, D_2, \ldots, D_n$ are i.i.d. non-negative integer-valued random variables with $\Pr\{\sum_{i=1}^n D_i = n\} > 0$, and let $\hat{D}_1, \hat{D}_2, ..., \hat{D}_n$ be a sequence of random variables with joint distribution is given by

$$\Pr\big\{\hat{D}_i = d_i, 1 \leq i \leq n\big\} = \Pr\Big\{D_i = d_i, 1 \leq i \leq n \,\Big|\, \sum_{i=1}^n D_i = n\Big\}.$$

Clearly, the variables $\hat{D}_1, \hat{D}_2, ..., \hat{D}_n$ are exchangeable with $\sum_{i=1}^n \hat{D}_i = n$, so we can use $\hat{D}_1, \hat{D}_2, ..., \hat{D}_n$ to construct $T_n^{\hat{D}}$ and $G_n^{\hat{D}}$. We note that it is easy to check that if $D_1, D_2, ..., D_n$ are i.i.d. Poisson variables, then $\hat{D}_1, \hat{D}_2, ..., \hat{D}_n$ have a multinomial distribution with parameters n and $(1/n, 1/n, ..., 1/n)$ and the corresponding random mapping $T_n^{\hat{D}}$ is just the usual uniform random mapping. There are interesting interpretations of $T_n^{\hat{D}}$ in the cases where the underlying i.i.d. variables $D_1, D_2, ..., D_n$ have **(i)** a generalised negative binomial distribution, and **(ii)** a binomial $Bin(m, p)$ distribution. In particular, case (i) corresponds to a random mapping with 'preferential attachment' and case (ii) corresponds to a random mapping with 'anti-preferential attachment'.

In this paper we consider some local properties of the random digraph $G_n^{\hat{D}}$ which represents the random mapping $T_n^{\hat{D}}$. By local properties we mean the distributions of random mapping characteristics related to a given vertex v - for example, the numbers of predecessors and successors of v in $G_n^{\hat{D}}$. To investigate such variables, we introduce some further notation and definitions. For any $f \in \mathcal{M}_n$ and any positive integer m, let $f^{(m)}$ denote the m^{th} iterate of f, and for every $i \in [n]$, define $f^{(0)}(i) \equiv i$. We say that $i \in [n]$ is a *cyclic* vertex of f if for some $m > 0$, $f^{(m)}(i) = i$. In particular, if i is a cyclic vertex of f then vertex i lies on a cycle in the digraph $G(f)$. We also note that every component of $G(f)$ consists of a directed cycle with trees, directed towards the cycle, attached to it. For any $f \in \mathcal{M}_n$, let $\mathcal{L}(f)$ denote the set of cyclic vertices in the component of $G(f)$ which contains the vertex 1.Define $\ell(f) = |\mathcal{L}(f)|$ and define $h(f)$, the height of vertex 1 in $G(f)$, by

$$h(f) = \min\{k \geq 0 : f^{(k)}(1) \in \mathcal{L}(f)\}.$$

Next, let

$$\mathcal{P}(f) \equiv \{j \in [n] : f^{(k)}(j) = 1 \text{ for some } k \geq 0\},$$

denote the predecessors of vertex 1 under f, and let

$$\mathcal{S}(f) \equiv \{j \in [n] : f^{(k)}(1) = j \text{ for some } k \geq 0\},$$

denote the successors of vertex 1 under f. We define $p(f) = |\mathcal{P}(f)|$ and $s(f) = |\mathcal{S}(f)|$. In this paper we are interested in the local properties of $G_n^{\hat{D}}$ which are described by the random variables $\ell_n^{\hat{D}} \equiv \ell(T_n^{\hat{D}})$, $h_n^{\hat{D}} \equiv h(T_n^{\hat{D}})$, $p_n^{\hat{D}} \equiv p(T_n^{\hat{D}}$ and $s_n^{\hat{D}} \equiv s(T_n^{\hat{D}})$. We mention here that the distributions for the number of cyclic vertices, the number of components, and the size of a typical component of $G_n^{\hat{D}}$ have been determined in a companion paper [**?**].

3 Results

In this section we derive general formulas for the distributions of the variables $\ell_n^{\hat{D}}$, $h_n^{\hat{D}}$, $p_n^{\hat{D}}$ and $s_n^{\hat{D}}$ described above. The joint distribution of $h_n^{\hat{D}}$ and $\ell_n^{\hat{D}}$ is given by the following theorem and the distributions for $\ell_n^{\hat{D}}$, $h_n^{\hat{D}}$, and $s_n^{\hat{D}}$are obtained as corollaries.

Theorem 1. *For $0 \le x, y \le n-1$,*

$$\Pr\{h_n^{\hat{D}} = x, \ell_n^{\hat{D}} = y+1\} = \frac{1}{n}\mathbf{E}\bigg(\hat{D}_1(\hat{D}_1 - 1)\hat{D}_2 \ldots \hat{D}_{x+y}\bigg) \cdot \mathbf{I}_{\{x \neq 0, x+y+1 \le n\}}$$

$$+\frac{1}{n}\mathbf{E}\bigg(\hat{D}_1\hat{D}_2 \ldots \hat{D}_{y+1}\bigg) \cdot \mathbf{I}_{\{x=0\}}.$$

Proof. First suppose that $1 \le x \le n-1$ and $0 \le y \le n-x-1$. Let $k = x+y$, then we have

$$\Pr\{h_n^{\hat{D}} = x, \ell_n^{\hat{D}} = y+1\} = \sum_{\substack{C \subseteq [n] \setminus \{1\} \\ s.t. |C| = k}} \Pr\{\mathcal{S}_n^{\hat{D}} = C \cup \{1\}, h_n^{\hat{D}} = x, \ell_n^{\hat{D}} = y+1\}.$$

$$= \binom{n-1}{k} \Pr\{\mathcal{S}_n^{\hat{D}} = C' \cup \{1\}, h_n^{\hat{D}} = x, \ell_n^{\hat{D}} = y+1\} \tag{3.1}$$

where $C' = \{2, 3, ..., k+1\}$ and $\mathcal{S}_n^{\hat{D}} \equiv \mathcal{S}(T_n^{\hat{D}})$. We note that the second equality above holds since the distribution of $G_n^{\hat{D}}$ is invariant under re-labelling its vertices. Next, let $S(k)$ denote the set of all permutations, σ, of $\{2, 3, ..., k+1\}$, and for any $\sigma \in S(k)$, let

$$\mathcal{A}_\sigma = \big\{(T_n^{\hat{D}})^{(m)}(1) = \sigma(m+1) \text{ for } 1 \le m \le k \text{ and } (T_n^{\hat{D}})^{(k+1)}(1) = \sigma(x+1)\big\}. \tag{3.2}$$

Then we have

$$\Pr\{\mathcal{S}_n^{\hat{D}} = C' \cup \{1\}, h_n^{\hat{D}} = x, \ell_n^{\hat{D}} = y+1\} = \sum_{\sigma \in S(k)} \Pr\{\mathcal{A}_\sigma\} = k!\Pr\{\mathcal{A}_{id}\}. \tag{3.3}$$

Again, the second equality follows from the invariance of the distribution of $G_n^{\hat{D}}$ under the re-labelling of its vertices. To determine $\Pr\{\mathcal{A}_{id}\}$ we write

$$\Pr\{\mathcal{A}_{id}\} = \sum_{\substack{d_i \ge 0 \\ s.t. \sum d_i = n}} \Pr\left\{\mathcal{A}_{id} \middle| \hat{D}_i = d_i, 1 \le i \le n\right\} \Pr\{\hat{D}_i = d_i, 1 \le i \le n\} \tag{3.4}$$

We note that

$$\Pr\left\{\mathcal{A}_{id} \middle| \hat{D}_i = d_i, 1 \le i \le n\right\} \neq 0$$

if and only if $d_i \geq 1$ for $2 \leq i \leq k+1$ and, in addition, $d_{x+1} \geq 2$. In this case, it follows from (**??**) and straightforward counting arguments, that

$$\Pr\left\{\mathcal{A}_{id}\middle|\hat{D}_i = d_i, 1 \leq i \leq n\right\} =$$

$$= \frac{(n-k-1)!}{d_1!(d_2-1)!\cdot\cdot(d_{x+1}-2)!\cdot\cdot(d_{k+1}-1)!d_{k+2}!\cdot\cdot d_n!} \times \frac{d_1!\cdots d_n!}{n!}$$

$$= \frac{d_2 d_3 \cdots d_{x+1}(d_{x+1}-1)\cdots d_{k+1}}{n(n-1)\cdots(n-k)}. \qquad (3.5)$$

Observe that formula (**??**) is still valid when $d_i = 0$ for some $2 \leq i \leq k+1$ or $d_{x+1} = 1$. So it follows from (**??**) and (**??**) that

$$\Pr\{\mathcal{A}_{id}\} = \frac{E(\hat{D}_2\cdot\cdot\hat{D}_{x+1}(\hat{D}_{x+1}-1)\cdot\cdot\hat{D}_{k+1})}{n(n-1)\cdots(n-k)}$$

$$= \frac{E(\hat{D}_1(\hat{D}_1-1)\hat{D}_2\cdot\cdot\hat{D}_k)}{n(n-1)\cdots(n-k)}. \qquad (3.6)$$

The last equality holds by the exchangeability of $\hat{D}_1, \hat{D}_2, ..., \hat{D}_n$. So, combining (**??**), (**??**), and (**??**), we obtain

$$\Pr\{h_n^{\hat{D}} = x, \ell_n^{\hat{D}} = y+1\} = \frac{1}{n}\mathbf{E}\Big(\hat{D}_1(\hat{D}_1-1)\hat{D}_2...\hat{D}_{x+y}\Big) \qquad (3.7)$$

since $k = x+y$.

Next, suppose that $x = 0$ and $1 \leq y \leq n-1$. Then, as above, we have

$$\Pr\{h_n^{\hat{D}} = 0, \ell_n^{\hat{D}} = y+1\} = \sum_{\substack{C\subseteq[n]\setminus\{1\} \\ s.t. |C|=y}} \Pr\{\mathcal{S}_n^{\hat{D}} = C\cup\{1\}, h_n^{\hat{D}} = 0, \ell_n^{\hat{D}} = y+1\}$$

$$= \binom{n-1}{y}\Pr\{\mathcal{S}_n^{\hat{D}} = C'\cup\{1\}, h_n^{\hat{D}} = 0, \ell_n^{\hat{D}} = y+1\} \qquad (3.8)$$

where $C' = \{2, 3, ..., y+1\}$. Again, let $S(y)$ denote the set of all permutations, σ, of C', and for any $\sigma \in S(y)$, let

$$\mathcal{B}_\sigma = \big\{(T_n^{\hat{D}})^{(m)}(1) = \sigma(m+1) \text{ for } 1 \leq m \leq y \text{ and } (T_n^{\hat{D}})^{(y+1)}(1) = 1\big\}. \quad (3.9)$$

Then, as above, we have

$$\Pr\{\mathcal{S}_n^{\hat{D}} = C' \cup \{1\}, h_n^{\hat{D}} = 0, \ell_n^{\hat{D}} = y+1\} = \sum_{\sigma \in S(y)} \Pr\{\mathcal{B}_\sigma\} = k!\Pr\{\mathcal{B}_{id}\}. \quad (3.10)$$

We write

$$\Pr\{\mathcal{B}_{id}\} = \sum_{\substack{d_i \geq 0 \\ s.t. \sum d_i = n}} \Pr\left\{\mathcal{B}_{id} \middle| \hat{D}_i = d_i, 1 \leq i \leq n\right\} \Pr\{\hat{D}_i = d_i, 1 \leq i \leq n\} \quad (3.11)$$

and we note that

$$\Pr\left\{\mathcal{B}_{id} \middle| \hat{D}_i = d_i, 1 \leq i \leq n\right\} \neq 0$$

if and only if $d_i \geq 1$ for $1 \leq i \leq y+1$. In this case, it follows from (**??**) and counting that

$$\Pr\left\{\mathcal{B}_{id} \middle| \hat{D}_i = d_i, 1 \leq i \leq n\right\} =$$

$$= \frac{(n-y-1)!}{(d_1-1)!(d_2-1)! \cdot \cdot (d_{y+1}-1)! d_{y+2}! \cdot \cdot d_n!} \times \frac{d_1! \cdots d_n!}{n!}$$

$$= \frac{d_1 d_2 \cdot \cdot d_{y+1}}{n(n-1)\cdots(n-y)}. \quad (3.12)$$

Observe that formula (**??**) is still valid when $d_i = 0$ for some $1 \leq i \leq y+1$. So, combining (**??**), (**??**), and (**??**), we obtain

$$\Pr\{h_n^{\hat{D}} = 0, \ell_n^{\hat{D}} = y+1\} = \frac{1}{n}\mathbf{E}\left(\hat{D}_1 \hat{D}_2 \ldots \hat{D}_{y+1}\right). \quad (3.13)$$

Finally, for $x = y = 0$, we have

$$\Pr\{h_n^{\hat{D}} = 0, \ell_n^{\hat{D}} = 1\} = \Pr\{s_n^{\hat{D}} = 1\} = \Pr\{T_n^{\hat{D}}(1) = 1\} =$$

$$= \sum_{\substack{d_i \geq 0 \\ s.t. \sum d_i = n}} \Pr\left\{T_n^{\hat{D}}(1) = 1 \middle| \hat{D}_i = d_i, 1 \leq i \leq n\right\} \Pr\{\hat{D}_i = d_i, 1 \leq i \leq n\} \quad (3.14)$$

We note that $\Pr\left\{T_n^{\hat{D}}(1) = 1 \middle| \hat{D}_i = d_i, 1 \leq i \leq n\right\} \neq 0$ if and only if $d_1 \geq 1$, and in this case we have

$$\Pr\left\{T_n^{\hat{D}}(1) = 1 \middle| \hat{D}_i = d_i, 1 \leq i \leq n\right\} = \frac{(n-1)!}{(d_1-1)!(d_2)!\cdots d_n!} \times \frac{d_1!\cdots d_n!}{n!} = \frac{d_1}{n}. \tag{3.15}$$

Equation (**??**) remains valid when $d_1 = 0$. So it follows from (**??**) and (**??**) that

$$\Pr\{h_n^{\hat{D}} = 0, \ell_n^{\hat{D}} = 1\} = \frac{E(\hat{D}_1)}{n}$$

and (**??**) holds in the case $y = 0$. □

Corollary 1. *For* $0 \leq k \leq n-1$,

$$\Pr\left\{s_n^{\hat{D}} = k+1\right\} =$$

$$\frac{k}{n}\mathbf{E}\left(\hat{D}_1(\hat{D}_1-1)\hat{D}_2 ... \hat{D}_k\right) + \frac{1}{n}\mathbf{E}\left(\hat{D}_1\hat{D}_2\hat{D}_3 ... \hat{D}_{k+1}\right).$$

Proof. The corollary follows from Theorem 1 and the observation that for $0 \leq k \leq n-1$

$$\Pr\left\{s_n^{\hat{D}} = k+1\right\} = \sum_{x=0}^{k} \Pr\{h_n^{\hat{D}} = x, \ell_n^{\hat{D}} = k-x+1\}.$$

□

We also immediately obtain from Theorem 1:

Corollary 2. *For* $1 \leq x \leq n-1$,

$$\Pr\{h_n^{\hat{D}} = x\} = \frac{1}{n}\sum_{y=0}^{n-x-1} \mathbf{E}\left(\hat{D}_1(\hat{D}_1-1)\hat{D}_2 ... \hat{D}_{x+y}\right)$$

and

$$\Pr\{h_n^{\hat{D}} = 0\} = \frac{1}{n}\sum_{y=0}^{n-1} \mathbf{E}\left(\hat{D}_1\hat{D}_2 ... \hat{D}_{y+1}\right).$$

Corollary 3. *For $0 \leq y \leq n-1$*

$$\Pr\{\ell_n^{\hat{D}} = y+1\} = \frac{1}{n}\sum_{x=1}^{n-y-1} \mathbf{E}\Big(\hat{D}_1(\hat{D}_1 - 1)\hat{D}_2 ... \hat{D}_{x+y}\Big) + \frac{1}{n}\mathbf{E}\Big(\hat{D}_1\hat{D}_2 ... \hat{D}_{y+1}\Big)$$

where the sum above is interpreted as 0 if $y = n-1$.

Next, we consider the distribution of $p_n^{\hat{D}}$, the number of predecessors of vertex 1 in $G_n^{\hat{D}}$. In order to determine this distribution we need to count the number of directed trees, rooted at 1, with a specified degree sequence. To state the required tree counting lemma, we adopt some notation. First, for any $k \geq 2$, let $\mathcal{T}_k$ denote the set of all labelled trees on the vertices $1, 2, ..., k$ such that each tree $t \in \mathcal{T}_k$ is rooted at vertex 1 and the edges of t are oriented so that the (shortest) path from any vertex v to the root 1 is directed towards 1. For any $t \in \mathcal{T}_k$ and any vertex $v \in t$, we let $d_v(t)$ denote the in-degree of v in t. We note that we must have $d_1(t) \geq 1$ and $\sum_{i=1}^k d_i(t) = k-1$. Finally, for any non-negative integers $d_1, d_2, ..., d_k$, such that $d_1 \geq 1$ and $\sum_{i=1}^k d_i = k-1$, let

$$\mathcal{T}_k(d_1, d_2, ..., d_k) = \{t \in \mathcal{T}_k : d_i(t) = d_i, 1 \leq i \leq k\}$$

and let $t_k(d_1, d_2, ..., d_k) = |\mathcal{T}_k(d_1, d_2, ..., d_k)|$. Then we have the following result:

Lemma 1. *Suppose that $k \geq 2$ and $d_1, d_2, ..., d_k$ are non-negative integers such that $d_1 \geq 1$ and $\sum_{i=1}^k d_i = k-1$, then*

$$t_k(d_1, d_2, ..., d_k) = \frac{(k-2)!}{(d_1 - 1)! d_2! \cdots d_k!}. \tag{3.16}$$

Proof. Suppose that $k \geq 2$ and $d_1, d_2, ..., d_k$ are non-negative integers such that $d_1 \geq 1$ and $\sum_{i=1}^k d_i = k-1$ and let $\mathcal{S}(d_1, d_2, ..., d_k)$ denote the set of all sequences of length $k-1$ such that for any $s \in \mathcal{S}(d_1, d_2, ..., d_k)$ and for any $1 \leq i \leq k$, the integer i appears d_i times in s and the sequence s ends with the integer 1. To prove equation (**??**) we use a Prufer tree code [**?**] to construct a bijection between $\mathcal{S}(d_1, d_2, ..., d_k)$ and $\mathcal{T}_k(d_1, d_2, ..., d_k)$. Specifically, the Prufer encoding constructs a sequence $s \in \mathcal{S}(d_1, d_2, ..., d_k)$ for every $t \in \mathcal{T}_k(d_1, d_2, ..., d_k)$ as follows:

1. Suppose that t is a labelled, directed tree rooted at 1. Suppose that v is the smallest vertex in t such that in-degree $d_v(t) = 0$ and suppose that $v \rightarrow w$ is a directed edge in t. Delete vertex v from t and add w to the sequence s.

2. If t is the trivial tree consisting only of vertex 1, STOP. Otherwise, go to Step 1 and repeat.

The Prufer decoding constructs a tree $t \in \mathcal{T}_k(d_1, d_2, ..., d_k)$ for every $s \in \mathcal{S}(d_1, d_2, ..., d_k)$ as follows:

1. Start with k labelled isolated vertices, and let $K = (1, 2, ..., k)$ denote the ordered list of numbers $1, 2, ..., k$.
2. Suppose that i is the smallest number in list K which does not appear in sequence s and suppose that j is the first number in the sequence s. Add the directed edge $i \to j$ to the graph and remove i from the list K and j from the sequence s.
3. If $K = \{1\}$, STOP. Otherwise, go to Step 2 and repeat.

The Prufer coding and encoding are inverse operations (see [**?**]), so

$$|\mathcal{T}_k(d_1, d_2, ..., d_k)| = |\mathcal{S}(d_1, d_2, ..., d_k)| = \frac{(k-2)!}{(d_1-1)!d_2! \cdots d_k!}$$

as desired. □

Given Lemma 1, we can prove

Theorem 2. *For* $0 \le k \le n-1$,

$$\Pr\Big\{p_n^{\hat{D}} = k+1\Big\} = \frac{n-k}{n(k+1)}\Pr\Big\{\sum_{i=1}^{k+1} \hat{D}_i = k\Big\} + \frac{1}{n}\Pr\Big\{\sum_{i=1}^{k+1} \hat{D}_i = k+1\Big\}.$$

Proof. First, suppose that $1 \le k \le n-1$. Then we have

$$\Pr\{p_n^{\hat{D}} = k+1\} = \sum_{\substack{C \subseteq [n] \setminus \{1\} \\ s.t. |C| = k}} \Pr\{\mathcal{P}_n^{\hat{D}} = C \cup \{1\}\} = \binom{n-1}{k} \Pr\{\mathcal{P}_n^{\hat{D}} = C' \cup \{1\}\} \tag{3.17}$$

where $C' = \{2, 3, ..., k+1\}$ and $\mathcal{P}_n^{\hat{D}} \equiv \mathcal{P}(T_n^{\hat{D}})$. Next, observe that

$$\Pr\{\mathcal{P}_n^{\hat{D}} = C' \cup \{1\}\} =$$

$$= \sum_{\substack{d_i \ge 0 \\ \sum d_i = n}} \Pr\{\mathcal{P}_n^{\hat{D}} = C' \cup \{1\} \big| \hat{D}_i = d_i, 1 \le i \le n\} \Pr\{\hat{D}_i = d_i, 1 \le i \le n\}$$

$$= \sum_{\substack{d_i \geq 0 \\ \sum d_i = n}} \sum_{\ell=1}^{n} \Pr\big\{\mathcal{P}_n^{\hat{D}} = C' \cup \{1\}, T_n^{\hat{D}}(1) = \ell \big| \hat{D}_i = d_i, 1 \leq i \leq n\big\} \Pr\{\hat{D}_i = d_i, 1 \leq i \leq n\}. \tag{3.18}$$

Now suppose $\vec{d} = (d_1, d_2, ..., d_n)$ such that $\sum_{i=1}^{n} d_i = n$, and for $1 \leq \ell \leq n$, let

$$\mathcal{M}_n(\vec{d}, C' \cup \{1\}, \ell) = \{f \in \mathcal{M}_n : \vec{d}(f) = \vec{d}, \mathcal{P}(f) = C', f(1) = \ell\}. \tag{3.19}$$

Then it follows from (**??**) that for $1 \leq \ell \leq n$,

$$\Pr\big\{\mathcal{P}_n^{\hat{D}} = C' \cup \{1\}, T_n^{\hat{D}}(1) = \ell \big| \hat{D}_i = d_i, 1 \leq i \leq n\big\} = \frac{|\mathcal{M}_n(\vec{d}, C', \ell)|}{n!(\prod_{i=1}^{n} d_i!)^{-1}}. \tag{3.20}$$

So to determine $\Pr\{\mathcal{P}_n^{\hat{D}} = C' \cup \{1\}, T_n^{\hat{D}}(1) = \ell | \hat{D}_i = d_i, 1 \leq i \leq n\}$ we need to count the set $\mathcal{M}_n(\vec{d}, C', \ell)$. There are three cases to consider.

First, suppose that $1 < \ell \leq k+1$. If $\mathcal{P}_n^{\hat{D}} = C' \cup \{1\}$ and $T_n^{\hat{D}}(1) = \ell$, then 1 is a cyclic vertex of $T_n^{\hat{D}}$ and the vertex set of the connected component in $G_n^{\hat{D}}$ which contains 1 is just $\{1, 2, ..., k+1\}$. Hence, we have $\Pr\{\mathcal{P}_n^{\hat{D}} = C' \cup \{1\}, T_n^{\hat{D}} = \ell | \hat{D}_i = d_i, 1 \leq i \leq n\} \neq 0$ if and only if $\sum_{i=1}^{k+1} d_i = k+1$, $d_\ell \geq 1$, and $d_1 \geq 1$. So, suppose that $\sum_{i=1}^{k+1} d_i = k+1$, $d_\ell \geq 1$, $d_1 \geq 1$, and $f \in \mathcal{M}_n(\vec{d}, C', \ell)$. If the directed edge from 1 to ℓ in $G(f)$ is deleted, we obtain a directed tree on the vertices $1, 2, ..., k+1$ with in-degree sequence $d_1, ..., d_\ell - 1, ..., d_{k+1}$ and root at vertex 1, and a random mapping graph on the vertices $k+2, ..., n$ with in-degree sequence $d_{k+1}, ..., d_n$. It follows that to count $\mathcal{M}_n(\vec{d}, C', \ell)$ it is enough to count certain trees and certain random mapping graphs, i.e.

$$|\mathcal{M}_n(\vec{d}, C', \ell)| = t_{k+1}(d_1, ..., d_\ell - 1, ..., d_{k+1}) \times \frac{(n-k-1)!}{d_{k+2}! \cdots d_n!}. \tag{3.21}$$

Substituting (**??**) into (**??**) and appealing to Lemma 1, we obtain in this case

$$\begin{aligned} &\Pr\{\mathcal{P}_n^{\hat{D}} = C' \cup \{1\}, T_n^{\hat{D}}(1) = \ell | \hat{D}_i = d_i, 1 \leq i \leq n\} \\ &\quad = \frac{(k-1)!(n-k-1)!}{n!} d_1 d_\ell \mathbf{I}_{\{\sum_{i=1}^{k+1} d_i = k+1\}}. \end{aligned} \tag{3.22}$$

We note that equation (**??**) remains valid when $d_1 = 0$ or $d_\ell = 0$, and so it holds for *all* degree sequences $d_1, d_2, ..., d_n$ such that $\sum_{i=1}^{n} d_i = n$.

Similarly, in the case when $\ell = 1$, we have $\Pr\{\mathcal{P}_n^{\hat{D}} = C' \cup \{1\}, T_n^{\hat{D}}(1) = 1 | \hat{D}_i = d_i, 1 \le i \le n\} \neq 0$ if and only if $\sum_{i=1}^{k+1} d_i = k+1$, and $d_1 \ge 2$. Now provided $\sum_{i=1}^{k+1} d_i = k+1$ and $d_1 \ge 2$, we have by the same argument as given above

$$|\mathcal{M}_n(\vec{d}, C', 1)| = t_{k+1}(d_1 - 1, d_2, ..., d_{k+1}) \times \frac{(n-k-1)!}{d_{k+2}! \cdots d_n!}. \tag{3.23}$$

Again, substituting (**??**) into (**??**) and appealing to Lemma 1, we obtain

$$\Pr\{\mathcal{P}_n^{\hat{D}} = C' \cup \{1\}, T_n^{\hat{D}}(1) = 1 \big| \hat{D}_i = d_i, 1 \le i \le n\}$$

$$= \frac{(k-1)!(n-k-1)!}{n!} d_1(d_1 - 1)\mathbf{I}_{\{\sum_{i=1}^{k+1} d_i = k+1\}}. \tag{3.24}$$

Again, we note that equation (**??**) remains valid when $d_1 = 0$ or 1, and so it holds for *all* degree sequences $d_1, d_2, ..., d_n$ such that $\sum_{i=1}^n d_i = n$.

Finally, suppose that $\ell > k+1$. If $\mathcal{P}_n^{\hat{D}} = C' \cup \{1\}$ and $T_n^{\hat{D}}(1) = \ell$, then 1 is not a cyclic vertex of $T_n^{\hat{D}}$ and the graph induced by $T_n^{\hat{D}}$ on the vertex set $C' \cup \{1\}$ is a directed tree, rooted at 1. Hence, we have $\Pr\{\mathcal{P}_n^{\hat{D}} = C', T_n^{\hat{D}} = \ell | \hat{D}_i = d_i, 1 \le i \le n\} \neq 0$ if and only if $\sum_{i=1}^{k+1} d_i = k$, $d_\ell \ge 1$, and $d_1 \ge 1$. So suppose that $\sum_{i=1}^{k+1} d_i = k$, $d_\ell \ge 1$, $d_1 \ge 1$, and $f \in \mathcal{M}_n(\vec{d}, C', \ell)$. If the directed edge from 1 to ℓ in $G(f)$ is deleted, we obtain a directed tree on the vertices $1, 2, ..., k+1$ with in-degree sequence $d_1, d_2, ..., d_{k+1}$ and root at vertex 1, and a random mapping graph on the vertices $k+2, ..., n$ with in-degree sequence $d_{k+1}, ..., d_\ell - 1, ..., d_n$. So, in this case, we have

$$|\mathcal{M}_n(\vec{d}, C', \ell)| = t_{k+1}(d_1, d_2, ..., d_{k+1}) \times \frac{(n-k-1)!}{d_{k+2}! \cdot \cdot (d_\ell - 1)! \cdot \cdot d_n!}. \tag{3.25}$$

Substituting (**??**) into (**??**) and appealing to Lemma 1, we obtain

$$\Pr\{\mathcal{P}_n^{\hat{D}} = C' \cup \{1\}, T_n^{\hat{D}}(1) = \ell \big| \hat{D}_i = d_i, 1 \le i \le n\}$$

$$= \frac{(k-1)!(n-k-1)!}{n!} d_1 d_\ell \mathbf{I}_{\{\sum_{i=1}^{k+1} d_i = k\}}. \tag{3.26}$$

We note that equation (**??**) remains valid when $d_1 = 0$ or $d_\ell = 0$, and so it holds for *all* degree sequences $d_1, d_2, ..., d_n$ such that $\sum_{i=1}^n d_i = n$.

It follows from (**??**), (**??**), (**??**), (**??**), and the identities

$$\left(d_1(d_1-1)+\sum_{\ell=2}^{k+1} d_1 d_\ell\right)\mathbf{I}_{\{\sum_{i=1}^{k+1} d_i=k+1\}} = k d_1 \mathbf{I}_{\{\sum_{i=1}^{k+1} d_i=k+1\}}$$

and

$$\sum_{\ell=k+2}^{n} d_1 d_\ell \mathbf{I}_{\{\sum_{i=1}^{k+1} d_i=k\}} = (n-k) d_1 \mathbf{I}_{\{\sum_{i=1}^{k+1} d_i=k\}},$$

that

$$\Pr\{\mathcal{P}_n^{\hat{D}} = C' \cup \{1\}\} =$$

$$= \sum_{\substack{d_i\geq 0\\ \sum d_i=n}} \sum_{\ell=1}^{k+1} \Pr\{\mathcal{P}_n^{\hat{D}} = C'\cup\{1\}, T_n^{\hat{D}}(1)=\ell | \hat{D}_i = d_i, 1\leq i\leq n\}\Pr\{\hat{D}_i = d_i, 1\leq i\leq n\}$$

$$+ \sum_{\substack{d_i\geq 0\\ \sum d_i=n}} \sum_{\ell=k+2}^{n} \Pr\{\mathcal{P}_n^{\hat{D}} = C'\cup\{1\}, T_n^{\hat{D}}(1)=\ell | \hat{D}_i = d_i, 1\leq i\leq n\}\Pr\{\hat{D}_i = d_i, 1\leq i\leq n\}$$

$$= \frac{k!(n-k-1)!}{n!} \sum_{\substack{d_i\geq 0\\ \sum d_i=n}} d_1 \mathbf{I}_{\{\sum_{i=1}^{k+1} d_i=k+1\}} \Pr\{\hat{D}_i = d_i, 1\leq i\leq n\}$$

$$+ \frac{(k-1)!(n-k)!}{n!} \sum_{\substack{d_i\geq 0\\ \sum d_i=n}} d_1 \mathbf{I}_{\{\sum_{i=1}^{k+1} d_i=k\}} \Pr\{\hat{D}_i = d_i, 1\leq i\leq n\}$$

$$= \frac{k!(n-k-1)!}{n!} E\left(\hat{D}_1 \middle| \sum_{i=1}^{k+1} \hat{D}_i = k+1\right) \Pr\left\{\sum_{i=1}^{k+1} \hat{D}_i = k+1\right\}$$

$$+ \frac{(k-1)!(n-k)!}{n!} E\left(\hat{D}_1 \middle| \sum_{i=1}^{k+1} \hat{D}_i = k\right) \Pr\left\{\sum_{i=1}^{k+1} \hat{D}_i = k\right\}.$$

$$= \frac{k!(n-k-1)!}{n!} \Pr\left\{\sum_{i=1}^{k+1} \hat{D}_i = k+1\right\} + \frac{k!(n-k)!}{n!(k+1)} \Pr\left\{\sum_{i=1}^{k+1} \hat{D}_i = k\right\}. \quad (3.27)$$

The last equality follows since, by the exchangeability of $\hat{D}_1, ..., \hat{D}_{k+1}$, $E(\hat{D}_1 | \sum_{i=1}^{k+1} \hat{D}_i = k+1) = 1$ and $E(\hat{D}_1 | \sum_{i=1}^{k+1} \hat{D}_i = k) = \frac{k}{k+1}$. So, in the case $1 \leq k \leq n-1$, the result follows from (**??**) and (**??**).

Finally, in the case $k = 0$ we have

$$\begin{aligned}
\Pr\{p_n^{\hat{D}} = 1\} &= \Pr\{\hat{D}_1 = 0\} + \Pr\{\hat{D}_1 = 1, T_n^{\hat{D}}(1) = 1\} \\
&= \Pr\{\hat{D}_1 = 0\} + \Pr\{T_n^{\hat{D}}(1) = 1 | \hat{D}_1 = 1\}\Pr\{\hat{D}_1 = 1\} \\
&= \Pr\{\hat{D}_1 = 0\} + \frac{1}{n}\Pr\{\hat{D}_1 = 1\}
\end{aligned}$$

as required. □

4 Examples

In this section we consider the preferential and anti-preferential attachment models which are defined below. In a companion paper [?] it was shown that these models are equivalent to special cases of $T_n^{\hat{D}}$. Using this equivalence, we apply the results obtained in Section 3 to investigate local properties of these models.

A Preferential Attachment Model

We begin by defining $T_n^\rho : [n] \to [n]$, a random mapping with 'preferential attachment', where $\rho > 0$ is a fixed parameter. For $1 \leq k \leq n$, we define $T_n^\rho(k) = \xi_k^{(\rho,n)}$ where $\xi_1^{(\rho,n)}, \xi_2^{(\rho,n)}, ..., \xi_n^{(\rho,n)}$ is a sequence of random variables whose distributions depend on the evolution of an urn scheme. The distribution of each $\xi_k^{(\rho,n)}$ is determined by a (random) n-tuple of non-negative weights $\vec{a}(k) = (a_1(k), a_2(k), ..., a_n(k))$ where, for $1 \leq j \leq n$, $a_j(k)$ is the 'weight' of the j^{th} urn at the *start* of the k^{th} round of the urn scheme. Specifically, given $\vec{a}(k) = \vec{a} = (a_1, ..., a_n)$, we define

$$\Pr\left\{\xi_k^{(\rho,n)} = j \middle| \vec{a}(k) = \vec{a}\right\} = \frac{a_j}{\sum_{i=1}^n a_i}.$$

The random weight vectors $\vec{a}(1), \vec{a}(2), ..., \vec{a}(n)$ associated with the urn scheme are determined recursively. For $k = 1$, we set $a_1(1) = a_2(1) = \cdots = a_n(1) = \rho > 0$. For $k > 1$, $\vec{a}(k)$ depends on both $\vec{a}(k-1)$ and the value of $\xi_{k-1}^{(\rho,n)}$ as follows: Given that $\xi_{k-1}^{(\rho,n)} = j$, we set $a_j(k) = a_j(k-1) + 1$ and for all other $i \neq j$, we set $a_i(k) = a_i(k-1)$ (i.e. if $\xi_{k-1}^{(\rho,n)} = j$ then a 'ball' with weight 1 is added to the j^{th} urn).

The random mapping T_n^ρ as defined above is a preferential attachment model in the following sense. Since, for $1 \leq k \leq n$, we have $T_n^\rho(k) = \xi_k^{(\rho,n)}$, and since the (conditional) distribution of $\xi_k^{(\rho,n)}$ depends on the state of the urn scheme at the start of round k, it is clear that vertex k is more likely to be mapped to vertex j if the weight $a_j(k)$ is (relatively) large, i.e. if several of the vertices $1, 2, ..., k-1$ have already been mapped to vertex j . The distribution of T_n^ρ is given by Theorem 3 below and was obtained in [**?**].

Theorem 3. *Suppose that $D_1^\rho, D_2^\rho, ...$ are i.i.d. random variables with a generalized negative binomial distribution given by*

$$\Pr\{D_1^\rho = k\} = \frac{\Gamma(k+\rho)}{k!\Gamma(\rho)} \left(\frac{\rho}{1+\rho}\right)^\rho \left(\frac{1}{1+\rho}\right)^k \quad \textit{for} \quad k = 0, 1, ...,$$

where $\rho > 0$ is a fixed parameter.
For $n \geq 1$, let $\hat{D}(\rho, n) = (\hat{D}_{1,n}^\rho, \hat{D}_{2,n}^\rho, ..., \hat{D}_{n,n}^\rho)$ be a sequence of variables with joint distribution given by

$$\Pr\{\hat{D}_{i,n}^\rho = d_i, 1 \leq i \leq n\} = \Pr\left\{D_i^\rho = d_i, 1 \leq i \leq n \,\middle|\, \sum_{i=1}^n D_i^\rho = n\right\}.$$

Then for every $n \geq 1$, the random mappings $T_n^\rho : [n] \to [n]$ and $T_n^{\hat{D}(\rho,n)} : [n] \to [n]$ have the same distribution.

Since $T_n^\rho \overset{d}{\sim} T_n^{\hat{D}(\rho,n)}$, it follows that the random digraphs $G_n^\rho \equiv G(T_n^\rho)$ and $G_n^{\hat{D}(\rho,n)}$ have the same distribution. So we can investigate the local properties of G_n^ρ by applying Theorem 1 (and its corollaries) and Theorem 2 to $G_n^{\hat{D}(\rho,n)}$. In the calculations that follow we adopt the following notation: if $f(s) = \sum_{k=0}^\infty a_k s^k$, then $[s^n]f(s) = a_n$, the coefficient of s^n in the power series expansion of $f(s)$. We also use the fact that the the probability generating function for D_1^ρ is given by

$$E(s^{D_1^\rho}) = \left(\frac{\rho}{1+\rho-s}\right)^\rho. \tag{4.1}$$

We begin by noting that for any integer $1 \leq x \leq n$ we have

$$E\Big(\hat{D}_{1,n}^\rho \hat{D}_{2,n}^\rho \cdots \hat{D}_{x,n}^\rho\Big) = E\Big(D_1^\rho D_2^\rho \cdots D_x^\rho \,\Big|\, \sum_{i=1}^n D_i^\rho = n\Big)$$

$$= \frac{[s^n]E\Big(D_1^\rho s^{D_1^\rho} D_2^\rho s^{D_2^\rho} \cdots D_x^\rho s^{D_x^\rho} s^{D_{x+1}^\rho} \cdots s^{D_n^\rho}\Big)}{[s^n]E\Big(s^{D_1^\rho} \cdots s^{D_n^\rho}\Big)}$$

$$= \frac{[s^n]\Big(E\big(D_1^\rho s^{D_1^\rho}\big)\Big)^x \Big(E\big(s^{D_1^\rho}\big)\Big)^{n-x}}{[s^n]\Big(E\big(s^{D_1^\rho}\big)\Big)^n}. \tag{4.2}$$

The last equality holds since the variables $D_1^\rho, D_2^\rho, ..., D_n^\rho$ are independent and identically distributed. Now since

$$E\Big(D_1^\rho s^{D_1^\rho}\Big) = s\left(\frac{\rho}{1+\rho-s}\right)^{\rho+1}, \tag{4.3}$$

it follows from (**??**)-(**??**) and routine calculations that

$$E\Big(\hat{D}_{1,n}^\rho \hat{D}_{2,n}^\rho \cdots \hat{D}_{x,n}^\rho\Big) = \frac{\rho^x (n)_x}{(n\rho + x - 1)_x} \tag{4.4}$$

where $(n)_x \equiv (n)(n-1)\cdots(n-x+1)$. Similarly, for $1 \le x \le n-1$, we have

$$E\Big(\hat{D}_{1,n}^\rho(\hat{D}_{1,n}^\rho - 1)\hat{D}_{2,n}^\rho \cdots \hat{D}_{x,n}^\rho\Big) = E\Big(D_1^\rho(D_1^\rho - 1)D_2^\rho \cdots D_x^\rho \,\Big|\, \sum_{i=1}^n D_i^\rho = n\Big)$$

$$= \frac{[s^n]E\Big(D_1^\rho(D_1^\rho - 1)s^{D_1^\rho} D_2^\rho s^{D_2^\rho} \cdots D_x^\rho s^{D_x^\rho} s^{D_{x+1}^\rho} \cdots s^{D_n^\rho}\Big)}{[s^n]E\Big(s^{D_1^\rho} \cdots s^{D_n^\rho}\Big)}$$

$$= \frac{[s^n]E\Big(D_1^\rho(D_1^\rho - 1)s^{D_1^\rho}\Big)\Big(E\big(D_1^\rho s^{D_1^\rho}\big)\Big)^{x-1}\Big(E\big(s^{D_1^\rho}\big)\Big)^{n-x}}{[s^n]\Big(E\big(s^{D_1^\rho}\big)\Big)^n}. \tag{4.5}$$

Since

$$E\Big(D_1^\rho(D_1^\rho - 1)s^{D_1^\rho}\Big) = \left(\frac{1+\rho}{\rho}\right) s^2 \left(\frac{\rho}{1+\rho-s}\right)^{\rho+2}, \tag{4.6}$$

it follows from (**??**) and routine calculations that

$$E\Big(\hat{D}_{1,n}^\rho(\hat{D}_{1,n}^\rho - 1)\hat{D}_{2,n}^\rho \cdots \hat{D}_{x,n}^\rho\Big) = \rho^x (1+\rho)\frac{(n)_{x+1}}{(n\rho + x)_{x+1}}. \tag{4.7}$$

It now follows from Theorem 3, Corollary 2, (**??**), and (**??**) that for $1 \leq k \leq n-1$ and $\rho > 0$

$$\Pr\{h(T_n^\rho) = k\} = \Pr\{h_n^{\hat{D}(\rho,n)} = k\} = \frac{1}{n} \sum_{y=0}^{n-k-1} E(\hat{D}^\rho_{1,n}(\hat{D}^\rho_{1,n} - 1)\hat{D}^\rho{}_{2,n} \cdots \hat{D}^\rho_{k+y,n})$$

$$= \frac{1}{n} \sum_{x=k}^{n-1} \frac{\rho^x (1+\rho)(n)_{x+1}}{(n\rho + x)_{x+1}} \tag{4.8}$$

and

$$\Pr\{h(T_n^\rho) = 0\} = \frac{1}{n} \sum_{x=1}^{n} \frac{\rho^x (n)_x}{(n\rho + x - 1)_x}. \tag{4.9}$$

It is straightforward to check that for fixed $\rho > 0$ and under the assumption that $k = \lfloor x\sqrt{n} \rfloor$ for some fixed $0 < x < \infty$,

$$\Pr\{h(T_n^\rho) = k\} \sim \frac{1}{\sqrt{n}} \int_x^\infty \left(\frac{1+\rho}{\rho}\right) \exp\left(\frac{-(1+\rho)u^2}{2\rho}\right) du. \tag{4.10}$$

Likewise, for $0 \leq k \leq n-1$, it follows from Theorem 3, Corollary 3, (**??**), and (**??**) that

$$\Pr\{\ell(T_n^\rho) = k+1\} =$$

$$\frac{1}{n} \sum_{x=1}^{n-k-1} E(\hat{D}^\rho_{1,n}(\hat{D}^\rho_{1,n} - 1)\hat{D}^\rho{}_{2,n} \cdots \hat{D}^\rho_{k+x,n}) + \frac{1}{n} E(\hat{D}^\rho_{1,n}\hat{D}^\rho_{2,n} \cdots \hat{D}^\rho_{k+1,n})$$

$$= \frac{1}{n} \sum_{y=k+1}^{n-1} \frac{\rho^y (1+\rho)(n)_{y+1}}{(n\rho + y)_{y+1}} + \frac{1}{n} \frac{\rho^{k+1}(n)_{k+1}}{(n\rho + k)_{k+1}}. \tag{4.11}$$

Again, for fixed $\rho > 0$ and under the assumption that $k = \lfloor y\sqrt{n} \rfloor$ for some fixed $0 < y < \infty$, we obtain

$$\Pr\{\ell(T_n^\rho) = k+1\} \sim \frac{1}{\sqrt{n}} \int_y^\infty \left(\frac{1+\rho}{\rho}\right) \exp\left(\frac{-(1+\rho)u^2}{2\rho}\right) du. \tag{4.12}$$

So it follows from (**??**) and (**??**) that the variables $\frac{h(T_n^\rho)}{\sqrt{n}}$ and $\frac{\ell(T_n^\rho)}{\sqrt{n}}$ converge in distribution to a variable Y_α with density

$$f_\alpha(y) = \int_y^\infty \alpha e^{-\alpha u^2/2} du \quad \text{for } y \in (0, \infty) \tag{4.13}$$

and parameter $\alpha = \frac{\rho+1}{\rho}$.

Next, for $0 \le k \le n-1$, we obtain from Theorem 3, Corollary 1, (**??**), and (**??**)

$$\begin{aligned}
\Pr\{s(T_n^\rho) = k+1\} &= \frac{k}{n} E(\hat{D}_{1,n}^\rho(\hat{D}_{1,n}^\rho - 1)\hat{D}_{2,n}^\rho \cdots \hat{D}_{k,n}^\rho) \\
&\quad + \frac{1}{n} E(\hat{D}_{1,n}^\rho \cdots \hat{D}_{k+1,n}^\rho) \\
&= (k(1+\rho)+\rho)\frac{(n-k)\rho^k (n)_k}{n(n\rho+k)_{k+1}} \qquad (4.14) \\
&= \frac{\rho^k (n-1)_k}{(n\rho+k-1)_k} - \frac{\rho^{k+1}(n-1)_{k+1}}{(n\rho+k)_{k+1}}.
\end{aligned}$$

For fixed $\rho > 0$ and under the assumption that $k = \lfloor x\sqrt{n} \rfloor$ for some fixed $0 < x < \infty$, we obtain by the usual asymptotic calculations

$$\Pr\{s(T_n^\rho) = k+1\} \sim \frac{1}{\sqrt{n}} \frac{(\rho+1)x}{\rho} \exp\left(-\left(\frac{\rho+1}{\rho}\right)\frac{x^2}{2}\right).$$

Finally, for $0 \le k \le n-1$, it follows from Theorem 2 and Theorem 3 that

$$\Pr\{p(T_n^\rho) = k+1\} = \frac{n-k}{n(k+1)} \Pr\left\{\sum_{i=1}^{k+1} \hat{D}_{i,n}^\rho = k\right\} + \frac{1}{n} \Pr\left\{\sum_{i=1}^{k+1} \hat{D}_{i,n}^\rho = k+1\right\}$$

$$= \frac{n-k}{n(k+1)} \Pr\left\{\sum_{i=1}^{k+1} D_i^\rho = k \,\middle|\, \sum_{i=1}^{n} D_i^\rho = n\right\} + \frac{1}{n} \Pr\left\{\sum_{i=1}^{k+1} D_i^\rho = k+1 \,\middle|\, \sum_{i=1}^{n} D_i^\rho = n\right\}.$$

It now follows by routine generating function calculations, that for
$0 \le k \le n-2$

$$\begin{aligned}
&\Pr\{p(T_n^\rho) = k+1\} \qquad (4.15) \\
&= \frac{\rho n + n - 1}{n} \binom{n}{k+1} \frac{\Gamma(\rho(k+1)+k)\Gamma\big((\rho+1)(n-k-1)\big)\Gamma(\rho n)}{\Gamma(\rho(k+1))\Gamma(\rho(n-k-1))\Gamma(\rho n + n)} \\
&= \frac{1}{n}\binom{n-1}{k} \frac{\rho}{\rho+1} \frac{((\rho+1)(k+1)-1)_k((\rho+1)(n-k-1))_{n-k-1}}{((\rho+1)(k+1)-1)((\rho+1)n-2)_{n-2}},
\end{aligned}$$

while

$$\Pr\{p(T_n^\rho) = n\} = \frac{1}{n}.$$

One can note that the distribution above is a variant of a special case of the *quasi-hypergeometric distribution I* (see (2.122) in [**?**]). We also obtain, by the usual asymptotic calculations, that for fixed $\rho > 0$ and $k = 0, 1, \ldots,$

$$\begin{aligned} &\lim_{n\to\infty} \Pr\{p(T_n^{\rho}) = k+1\} \qquad (4.16) \\ &= \frac{1}{\rho^k}\binom{(\rho+1)(k+1)}{k}\frac{1}{k+1}\left(\frac{(\rho+1)(k+1)-k}{(\rho+1)(k+1)-1}\right)\left(\frac{\rho}{\rho+1}\right)^{(\rho+1)(k+1)} \\ &= \frac{\rho}{(\rho+1)k+\rho}\binom{(\rho+1)k+\rho}{k}\left(\frac{1}{\rho+1}\right)^{k}\left(\frac{\rho}{\rho+1}\right)^{(\rho+1)k+\rho-k}, \end{aligned}$$

which is a special case of the *generalized binomial distribution* (see (2.121) in [**?**]).

An Anti-Preferential Attachment Model

In this section we consider $T_n^m : [n] \to [n]$, a random mapping with 'anti-preferential attachment', where $m \geq 1$ is a fixed interger parameter. For $1 \leq k \leq n$, we define $T_n^m(k) = \eta_k^{(m,n)}$ where, as in the definition of T_n^{ρ}, the variables $\eta_1^{(m,n)}, \eta_2^{(m,n)}, .., \eta_n^{(m,n)}$ depend on the evolution of an urn scheme. The distribution of each variable $\eta_k^{(m,n)}$ is determined by a (random) n-tuple of non-negative weights $\vec{b}(k) = (b_1(k), b_2(k), \ldots, b_n(k))$ where, for $1 \leq j \leq n$, $b_j(k)$ is the number of balls in the j^{th} urn at the start of the k^{th} round of the urn scheme. Specifically, given $\vec{b}(k) = \vec{b} = (b_1, \ldots, b_n)$, we define

$$\Pr\left\{\eta_k^{(m,n)} = j \middle| \vec{b}(k) = \vec{b}\right\} = \frac{b_j}{\sum_{i=1}^{n} b_i}$$

The random weight vectors $\vec{b}(1), \vec{b}(2), \ldots, \vec{b}(n)$ associated with the urn scheme are determined recursively. For $k = 1$, we set $b_1(1) = b_2(1) = \cdots = b_n(1) = m$. For $k > 1$, $\vec{b}(k)$ depends on both $\vec{b}(k-1)$ and the value of $\eta_{k-1}^{(m,n)}$ as follows: Given that $\eta_{k-1}^{(m,n)} = j$, we set $b_j(k) = b_j(k-1) - 1$ and for all other $i \neq j$, we set $b_i(k) = b_i(k-1)$ (i.e. if $\eta_{k-1}^{(m,n)} = j$ then a ball is removed from the j^{th} urn).

The random mapping T_n^m as defined above is an anti-preferential attachment model in the following sense. Since, for $1 \leq k \leq n$, we have $T_n^m(k) = \eta_k^{(m,n)}$, and since the (conditional) distribution of $\eta_k^{(m,n)}$ depends on the state of the urn scheme at the start of round k, it is clear that vertex k is less likely to ‘ choose’ vertex j if the weight $b_j(k)$ is (relatively) small, i.e. if several of the vertices $1, 2, ..., k-1$ have already been mapped to vertex j. It is also clear from the definition of T_n^m that the in-degree of any vertex in the random digraph $G_n^m \equiv G(T_n^m)$ is at most m and in the case $m = 1$, T_n^1 is a (uniform) random permutation. The distribution of T_n^m is given by Theorem 4 below and was obtained in [**?**].

Theorem 4. *Suppose that $D_1^m, D_2^m, ...$ are i.i.d. $Bin(m, p)$ variables where $m \geq 1$ is a fixed integer parameter.*
Let $\hat{D}(m,n) = (\hat{D}_{1,n}^m, \hat{D}_{2,n}^m, ..., \hat{D}_{n,n}^m)$ be a sequence of variables with joint distribution given by

$$\Pr\{\hat{D}_{i,n}^m = d_i, 1 \leq i \leq n\} = \Pr\left\{ D_i^m = d_i, 1 \leq i \leq n \,\middle|\, \sum_{i=1}^n D_i^m = n \right\}.$$

Then the random mappings T_n^m and $T_n^{\hat{D}(m,n)}$ have the same distribution.

Since $T_n^m \overset{d}{\sim} T_n^{\hat{D}(m,n)}$, it follows that the random digraphs $G_n^m \equiv G(T_n^m)$ and $G_n^{\hat{D}(n,m)}$ have the same distribution. So, as in the case of the preferential attachment model, we can investigate the local properties of G_n^m by considering the local properties of $G_n^{\hat{D}(n,m)}$.

We begin by noting that for any integer $1 \leq x \leq n-1$ we have

$$E(\hat{D}_{1,n}^m(\hat{D}_{1,n}^m - 1)\hat{D}_{2,n}^m \cdots \hat{D}_{x,n}^m) = E\left(D_{1,n}^m(D_{1,n}^m - 1)D_{2,n}^m \cdots D_{x,n}^m \,\middle|\, \sum_{i=1}^n D_{i,n}^m = n \right)$$

$$
\begin{aligned}
&= \sum_{\vec{d}\ s.t.\ \sum_{i=1}^{n} d_i = n} (d_1-1)d_1 d_2\cdots d_x \times \frac{\binom{m}{d_1}\cdots\binom{m}{d_n}}{\binom{nm}{n}} \\
&= \sum_{t=x+1}^{\min(n,xm)} \sum_{\substack{\vec{d}\ s.t.\ \sum_{i=1}^{x} d_i = t \\ and\ \ \sum_{i=1}^{n} d_i = n}} (d_1-1)d_1 d_2\cdots d_x \times \frac{\binom{m}{d_1}\cdots\binom{m}{d_n}}{\binom{nm}{n}} \\
&= \sum_{t=x+1}^{\min(n,xm)} \sum_{\vec{d}\ s.t.\ \sum_{i=1}^{x} d_i = t} (d_1-1)d_1 d_2\cdots d_x \times \frac{\binom{m}{d_1}\cdots\binom{m}{d_x}\binom{nm-xm}{n-t}}{\binom{nm}{n}} \\
&= m^x(m-1) \sum_{t=x+1}^{\min(n,xm)} \sum_{\vec{d}\ s.t.\ \sum_{i=1}^{x} d_i = t} \frac{\binom{m-2}{d_1-2}\binom{m-1}{d_2-1}\cdots\binom{m-1}{d_x-1}\binom{nm-xm}{n-t}}{\binom{nm}{n}} \\
&= m^x(m-1) \sum_{t=x+1}^{\min(n,xm)} \frac{\binom{xm-x-1}{t-x-1}\binom{nm-xm}{n-t}}{\binom{nm}{n}} = m^x(m-1)\frac{\binom{nm-x-1}{n-x-1}}{\binom{nm}{n}} \\
&= m^x(m-1)\frac{(n)_{x+1}}{(nm)_{x+1}}\,. \qquad (4.17)
\end{aligned}
$$

In the summations above the sum is always taken over those degree sequences for which the binomial coefficients are defined. We also adopt the formal convention that $\binom{0}{0}=1$. By similar calculations we also obtain for $1 \le x \le n$ (and $m \ge 2$)

$$
E(\hat{D}^m_{1,n}\hat{D}^m_{2,n}\cdots\hat{D}^m_{x,n}) = m^x\frac{\binom{nm-x}{n-x}}{\binom{nm}{n}} = m^x\frac{(n)_x}{(nm)_x}. \qquad (4.18)
$$

It follows from Theorem 4, Corollary 2, (**??**), and (**??**), that for $1 \le k \le n-1$ and $m \ge 2$ we have

$$
\begin{aligned}
\Pr\{h(T^m_n)=k\} &= \frac{1}{n}\sum_{t=k}^{n-1} E(\hat{D}^m_{1,n}(\hat{D}^m_{1,n}-1)\hat{D}^m_{2,n}\cdots\hat{D}^m_{t,n}) \\
&= \frac{1}{n}\sum_{t=k}^{n-1}\frac{m^t(m-1)\binom{nm-t-1}{n-t-1}}{\binom{nm}{n}}, \qquad (4.19)
\end{aligned}
$$

and

$$
\Pr\{h(T^m_n)=0\} = \frac{1}{n}\sum_{t=1}^{n} E(\hat{D}^m_{1,n}\hat{D}^m_{2,n}\cdots\hat{D}^m_{t,n}) = \frac{1}{n}\sum_{t=1}^{n} m^t\frac{\binom{nm-t}{n-t}}{\binom{nm}{n}}. \qquad (4.20)
$$

It is routine to check that for fixed $m \geq 2$ and under the assumption that $k = \lfloor x\sqrt{n} \rfloor$ for some fixed $0 < x < \infty$, we obtain

$$\Pr\{h(T_n^m) = k\} \sim \frac{1}{\sqrt{n}} \int_x^\infty \left(\frac{m-1}{m}\right) \exp\left(-\left(\frac{m-1}{m}\right)\frac{u^2}{2}\right) du. \quad (4.21)$$

Likewise, it follows from Theorem 4, Corollary 1, (**??**), and (**??**), that for $0 \leq k \leq n-1$ and $m \geq 2$

$$\begin{aligned} \Pr\{\ell(T_n^m) = k+1\} &= \frac{1}{n} \sum_{t=k+1}^{n-1} E(\hat{D}_{1,n}^m(\hat{D}_{1,n}^m - 1)\hat{D}_{2,n}^m \cdots \hat{D}_{t,n}^m) + \frac{1}{n} E(\hat{D}_{1,n}^m \hat{D}_{2,n}^m \cdots \hat{D}_{k+1,n}^m) \\ &= \frac{1}{n} \sum_{t=k+1}^{n-1} \frac{m^t(m-1)\binom{nm-t-1}{n-t-1}}{\binom{nm}{n}} + \frac{1}{n} m^{k+1} \frac{\binom{nm-t}{n-t}}{\binom{nm}{n}}. \quad (4.22) \end{aligned}$$

Again, for fixed $m \geq 2$ and under the assumption that $k = \lfloor y\sqrt{n} \rfloor$ for some fixed $0 < y < \infty$, we obtain

$$\Pr\{\ell(T_n^m) = k+1\} = \frac{1}{\sqrt{n}} \int_y^\infty \left(\frac{m-1}{m}\right) \exp\left(-\left(\frac{m-1}{m}\right)\frac{u^2}{2}\right) du. \quad (4.23)$$

It follows form (**??**) and (**??**) that, as in the case of the preferential model, the variables $\frac{h(T_n^m)}{\sqrt{n}}$ and $\frac{\ell(T_n^m)}{\sqrt{n}}$ converge in distribution to Y_α but with $\alpha = \frac{m-1}{m}$ in this case.
Next, for $0 \leq k \leq n-1$ and $m \geq 2$, we have

$$\begin{aligned} \Pr\{s(T_n^m) = k+1\} &= \frac{k}{n} E(\hat{D}_{1,n}^m(\hat{D}_{1,n}^m - 1)\hat{D}_{2,n}^m \cdots \hat{D}_{k,n}^m) + \frac{1}{n} E(\hat{D}_{1,n}^m \cdots \hat{D}_{k+1,n}^m) \\ &= \frac{m^k}{n}(m(k+1) - k)\frac{\binom{nm-k-1}{n-k-1}}{\binom{nm}{n}} \\ &= \frac{m^k}{n}(k(m-1) + m)\frac{(n)_{k+1}}{(nm)_{k+1}} \\ &= m^k \frac{(n-1)_k}{(nm)_k} - m^{k+1}\frac{(n-1)_{k+1}}{(nm)_{k+1}}. \quad (4.24) \end{aligned}$$

It is straightforward to check that for fixed $m \geq 2$ and under the assumption that $k = \lfloor x\sqrt{n} \rfloor$ for some fixed $0 < x < \infty$, we have

$$\Pr\{s(T_n^m) = k+1\} \sim \frac{1}{\sqrt{n}} \frac{(m-1)x}{m} \exp\left(-\left(\frac{m-1}{m}\right)\frac{x^2}{2}\right).$$

Finally, it follows from Theorem 2 and Theorem 4 that for $0 \leq k \leq n-1$ and $m \geq 2$

$$\Pr\{p(T_n^m) = k+1\} = \tag{4.25}$$

$$\begin{aligned}
&= \frac{n-k}{n(k+1)} \Pr\left\{\sum_{i=1}^{k+1} \hat{D}_{i,n}^m = k\right\} + \frac{1}{n} \Pr\left\{\sum_{i=1}^{k+1} \hat{D}_{i,n}^m = k+1\right\} \\
&= \frac{n-k}{n(k+1)} \Pr\left\{\sum_{i=1}^{k+1} D_i^m = k \,\middle|\, \sum_{i=1}^{n} D_i^m = n\right\} + \frac{1}{n} \Pr\left\{\sum_{i=1}^{k+1} D_i^m = k+1 \,\middle|\, \sum_{i=1}^{n} D_i^m = n\right\} \\
&= \frac{n-k}{n(k+1)} \frac{\binom{m(k+1)}{k}\binom{m(n-k-1)}{n-k}}{\binom{mn}{n}} + \frac{1}{n} \frac{\binom{m(k+1)}{k+1}\binom{m(n-k-1)}{n-k-1}}{\binom{mn}{n}} \\
&= \frac{(mn-n+1)\binom{m(k+1)}{k+1}\binom{m(n-k-1)}{n-k-1}}{n(mk+m-k)\binom{mn}{n}} \\
&= \binom{n-1}{k} \frac{m}{m(k+1)} \frac{(m(k+1))_k (m(n-k-1))_{n-k-1}}{(mn)_{n-1}}.
\end{aligned}$$

Again we should mention that the distribution above is a special case of the *quasi-hypergeometric distribution I* (see (2.122) in [?]). Straightforward asymptotic calculations establish that for $m \geq 2$ and $k \in \{0, 1, 2, ...\}$,

$$\lim_{n\to\infty} \Pr\{p(T_n^m) = k+1\} = \frac{1}{(m-1)^k} \binom{mk+m}{k} \frac{1}{k+1} \left(1 - \frac{1}{m}\right)^{mk+m}.$$

We note that the asymptotic distribution for $p(T_n^m)$ is the Consul distribution (see [?] p.98) with parameters m and $\theta = \frac{1}{m}$.

5 Final remarks

One of the main advantages of the random mapping model $T_n^{\hat{D}}$ is that we have a calculus for this model which allows us to determine the distributions of several variables associated with the structure of $G_n^{\hat{D}}$ in terms of expectations of simple functions of $\hat{D}_1, \hat{D}_2, ..., \hat{D}_n$. As we have seen above, in the special case where the variables $\hat{D}_1, \hat{D}_2, ..., \hat{D}_n$ have the same distribution as a collection of i.i.d. variables $D_1, D_2, ..., D_n$ *conditioned* on $\sum_{i=1}^n D_i = n$, it is straightforward to use this calculus to obtain exact and asymptotic distributions for various important random variables associated with the local structure of $G_n^{\hat{D}}$. The calculus for $T_n^{\hat{D}}$ also illustrates the fundamental importance of the distribution of the underlying degree sequence $\hat{D}_1, ..., \hat{D}_n$ to the structure of the random mapping digraph. This suggests that in various modelling applications the key to fitting a random mapping model is to fit the joint distribution of the vertex in-degree data.

As a example of model fitting, we mention the work of Arney and Bender on random mappings with constraints on coalescence [**?**]. Their work was motivated, in part, by the analysis of shift register data. In order to model a random shift register they put a uniform measure on $\mathcal{M}_n^{\{0,1,2\}}$, the set of all mappings $f : [n] \to [n]$ such that, for every $1 \le i \le n$, $|f^{-1}(i)|$, the number of pre-images of i under f, equals $0, 1$, or 2. So, if $f \in \mathcal{M}_n^{\{0,1,2\}}$, then every vertex in $G_n(f)$ has in-degree equal $0, 1$, or 2. Arney and Bender observed that in some respects their model does not fit the shift register data. In particular, their model predicts $0.293n$ vertices with in-degree 0 whereas the average number of vertices with in-degree 0 in a random shift register is $n/4$. Their method depends on the asymptotic analysis of combinatorial generating functions, so to adjust their model they introduce an extra weighting into their generating functions in order to get better agreement between their model and the data. With some effort they are able to extract asymptotic distributions and mean values for certain statistics of interest. In contrast, by using the model $T_n^{\hat{D}}$, we can more easily and naturally capture the local structure of the shift register data. Specifically, suppose that $\hat{D}_1, \hat{D}_2, ..., \hat{D}_n$ have the same distribution as n independent $Bin(2, \frac{1}{2})$ variables, $D_1, D_2, ..., D_n$, conditioned on $\sum_{i=1}^n D_i = n$. Then $\Pr\{\hat{D}_1 = 0\} = \frac{1}{4}(1 + \frac{1}{2n-2})^{-1}$ and the expected number of vertices with in-degree 0 in $T_n^{\hat{D}}$ is asymptotic to $\frac{n}{4}$. Using this approach we can obtain *exact* distributions as well as the asymptotic results obtained by Arney and Bender. We are also able to obtain exact and asymptotic distributions for variables such as the number of predecessors of a vertex, which were not obtained by Arney and Bender.

References

[1] Aldous, D., Miermont, G. and Pitman, J. (2004). Brownian Bridge Asymptotics for Random $p-$mappings . *Electonic J. Probab.* **9**, 37-56.

[2] ALDOUS, D., MIERMONT, G. AND PITMAN, J. (2005). Weak Convergence of Random $p-$mappings and the Exploration Process of of Inhomogeneous Continuum Random Trees.. *Probab. Th. Related Fields.* **133**, 1-17.

[3] ARNEY, J. AND BENDER, E. A. (1982). Random mappings with constraints on coalescence and number of origins. *Pacific J. Math.* **103** 269–294.

[4] ARRATIA, R. AND TAVARÉ, S. (1992). Limit theorems for combinatorial structures via discrete process approximations. *Random Structures and Algorithms* **3,** 321–345.

[5] ARRATIA, R., STARK, D. AND TAVARÉ, S. (1995). Total variation asymptotics for Poisson process approximations of logarithmic combinatorial assemblies. *Ann. Probab.* **23**, 1347–1388.

[6] BERG, S. (1981). On snowball sampling, random mappings and related problems, *J. Appl. Prob.* **18**, 283–290.

[7] BERG, S. (1983) Random contact processes, snowball sampling and factorial series distributions, *J. Appl. Prob.* **20**, 31–46.

[8] BERG S. AND JAWORSKI J., Probability distributions related to the local structure of a random mapping, in: Frieze A. and Łuczak T. (eds): *Random Graphs, Volume 2.* John Wiley & Sons (1989), 1–21.

[9] BURTIN, Y. D. (1980). On simple formula for random mappings and its applications. *J. Appl. Prob.* **17**, 403–414.

[10] FLAJOLET, P., ODLYZKO, A. M. (1990) Random mapping statistics, *Lecture Notes in Computer Science* **434**, Advances in Cryptology - Eurocrypt Proceedings. 329–354.

[11] HANSEN, J. C. AND JAWORSKI, J. (2007). A new random mapping model. To appear in *Random Structures and Algorithms.*

[12] HANSEN, J. C. AND JAWORSKI, J. (2007). A random mapping with preferential attachment. *In preparation.*

[13] JAWORSKI, J. (1990). Random mappings with independent choices of the images. in:*Random Graphs*, Vol.1, Wiley, 89–101.

[14] Jaworski, J. (1998). Predecessors in a random mapping. *Random Structures and Algorithms* **13**, *No.3 and No.4*, 501–519.

[15] Jaworski, J. (1999). Epidemic processes on digraphs of random mappings. *J. Appl. Prob.* **36**, 780–798.

[16] Johnson, N. L., Kotz, S., and Kemp, A. W. (1992). *Univariate Discrete Distributions*, 2nd edition, JohnWiley & Sons, New York.

[17] Kolchin, V. F. (1986). *Random Mappings*, Optimization Software Inc., New York.

[18] Mutafchiev, L. (1982) A limit distribution related to random mappings and its application to an epidemic process, *Serdica* **8**, 197–203.

[19] Mutafchiev, L. (1984) On some stochastic problems of discrete mathematics, in: *Proceedings of XIII Spring Conference of the Union of Bulgarian Mathematicians*, 57–80.

[20] Pitman, J.. (2001) Random mappings, forest, and subsets associated with the Abel-Cayley-Hurwitz multinomial expansions, *Seminaire Lotharingien de Combinatoire* **46**:45 pp.

[21] Pittel, B. (1983) On distributions related to transitive closures of random finite mappings, *Ann. Probab.* **11**, 428–441.

[22] Prufer, H. (1918) Neuer Beweis eines Satzes uber Permutationen, *Archiv fur Mathematik und Physik.* **27**, 142-144.

[23] Quisquater, J.-J. and Delescaille, J-P. (1990) How easy is collision search? Application to DES *Lecture Notes in Computer Science* **434**, Advances in Cryptology - Eurocrypt Proceedings. 429–434.

[24] van Oorschot, P. C.. and Wiener, M. J.. (1994) Parallel collision search with applications to hash functions and discrete logarithms. *Proc. of the 2nd ACM Conference on Computer and Communications Security* 210–218.