Volume 5 Issue 6 ISSN 2631-4444

Degrees in random m-ary hooking networks

Kiran R. Bhutani[1], Ravi Kalpathy[1], and Hosam Mahmoud[2]

[1]Department of Mathematics, The Catholic University of America, Washington, D.C. 20064, U.S.A.

[2]Department of Statistics, The George Washington University, Washington, D.C. 20052, U.S.A.

The theme in this paper is a composition of random graphs and Pólya urns. The random graphs are generated through a small structure called the seed. Via Pólya urns, we study the asymptotic degree structure in a random m-ary hooking network and identify strong laws. We further upgrade the result to second-order asymptotics in the form of multivariate Gaussian limit laws. We give a few concrete examples and explore some properties with a full representation of the Gaussian limit in each case. The asymptotic covariance matrix associated with the Pólya urn is obtained by a new method that originated in this paper and is reported in [25].

1 Introduction

Many random structures grow by adding small components. For example, networks grow by linking new small structures, and urns grow by adding balls, etc. In this manuscript, we deal with a certain flavor of network growth via "hooking" components of fixed structure (all additions are graphs that have the same form). The basic repetitive form is called a *seed*, which has a particular node in it called the *hook*.

Initially, the network is just the seed. At every discrete point in time, a copy of the seed is brought to the current network and adjoined to it by fusing its hook into a node (vertex) chosen from the network. After a long period of growth, we have a large network. Figure 1 shows the step-by-step growth of a hooking network grown from a triangular seed in four steps. We have not specified how a node is chosen in the network. It could be specified deterministically by an external agent, such as an AI optimizer, or it could be done according to a probability distribution on the nodes of the existing network. In the latter case, the growth of the network is a stochastic process. In this light, Figure 1 would be only one realization of the process.

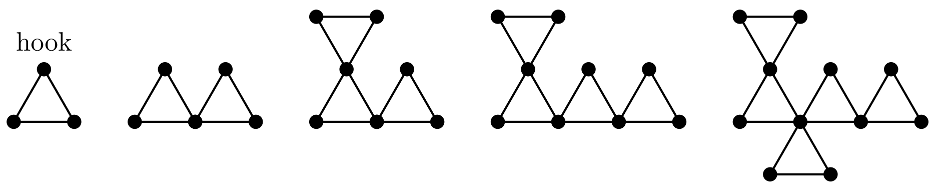

Figure 1: A hooking network grown in four time steps from a triangle (seed).

Alternative nomenclature appears in the jargon. For instance, adjoining the seed to the network is called by some hooking. The word "fusing" also appears frequently. The chosen receptacle node in the graph to host the hook of the newly added seed is referred to as a "latch." The hooking (fusing) operation is achieved by identifying the hook and the chosen latch. The act of adjoining a seed at a step may then be considered as "latching" or "recruiting" by the latch.

Kiran R. Bhutani: bhutani@cua.edu, 0000-0003-4823-0887
Ravi Kalpathy: kalpathy@cua.edu, 0000-0001-5874-2427
Hosam Mahmoud: hosam@gwu.edu, 0000-0003-0962-9406

Accepted in Compositionality on 2023-06-02. Click on the title to verify.

There has been recently a flurry of research papers on hooking networks (though some do not call it that) [3–7, 9–13, 15, 23, 24, 27]; see [29] for a comprehensive background.

Applications can be found in organic chemistry, social networks and many other areas. For instance, in polymer chemistry, monomers link together into a covalently bonded network of macromolecules to form polymers [4, 5, 7]. Social networks and the like grow by adding communities in which one member of a new community is a friend of one member of an existing social network, where the degree of a node in the network is a reflection of the popularity of the network subscriber it represents.

With the degrees of nodes in a network being of prime importance, our purpose in this paper is to study the degrees in a certain kind of network (with a limit on the number of hookings at every node). As the degrees in the network and the balls in a certain kind of Pólya urns have the same generative mechanism, we rely on the urn composition to produce theorems for the network structure considered in this manuscript.

We note that the study of the asymptotic degree structure in hooking networks has already been conducted (under different assumptions) in [9–11, 23]. The authors of these papers did not impose a limit on the number of hookings at a particular place.

Some applications put a cap on the number of times a component can be added at one place. A popular example is the binary search tree, which arises in computer science as an efficient data structure that expedites the fast retrieval of data [8, 21]. The binary imposition comes from the physical implementation of the nodes in computer memory. A node is comprised of space for fields of information with distinguishable pointers to the left and right subtrees containing nodes of a similar structure. These pointers occupy different locations in the computer memory.

A few models of hooking networks [4, 10–12] have been introduced for applications with no limit on the number of hookings that can be made at a particular node or edge. In this paper, we turn our attention to m-ary hooking networks, where at most m (finite) copies of a seed can join a node in the network. Figures 2–3 will aid the reader visualize this construct.

Via Pólya urns, we study the asymptotic degree structure in these m-ary networks. We identify strong laws for the number of nodes of various degrees in the network. We further upgrade the result to second-order asymptotics in the form of multivariate Gaussian limit laws. We compute the covariances of the Pólya urn by a new method that originated in this paper and is reported in [25].

1.1 The m-ary hooking network

An m-ary network grows as follows: We start with a connected *seed* graph G_0 with vertex set of size τ_0. A vertex in the seed is designated as a *hook*. Each copy of the seed that joins the network subsequently uses its own copy of that vertex for hooking. At step n, a copy of the seed is hooked into the graph $G_{n-1} = (V_{n-1}, \mathcal{E}_{n-1})$ that exists at time $n-1$ to produce the graph $G_n = (V_n, E_n)$. So, τ_0 is the size of V_0, and $\tau_n = (\tau_0 - 1)n + \tau_0$. For instantiation, see Figure 1 in which the seed is a triangle, and $\tau_0 = 3, \tau_1 = 5, \tau_2 = 7, \tau_3 = 9$, and $\tau_4 = 11$.

We think of the graph G_n as a network of *age* n. The hooking is accomplished by fusing together the hook of a new copy of the seed and a *latch* (vertex) randomly chosen from the network at age $n-1$. A vertex in the network can qualify as a latch at most $m \geq 1$ times.

We consider a *uniform* probability model that equally likely selects any of the insertion positions in the network as a latch. Initially, there are m insertion positions associated with each vertex in G_0. Every hooking takes away an insertion position from a vertex. After a node recruits $r \leq m$ times, there remain $m - r$ insertion positions. When all m positions are taken at a vertex (i.e., the vertex has recruited m times), the vertex is saturated and is no longer active in recruiting.

A visual device helps us discern the insertion positions. We represent each insertion position at a node as a virtual node connected by an edge to the vertex. The graph carrying the virtual nodes is called the *extended network*. Initially, there are m virtual nodes attached to each vertex of the seed, which is the graph G_0. Figure 2 shows a seed to be used in building a binary network ($m = 2$) and its extension. The nodes of the seed are shown as bullets and the virtual nodes in the extended seed are shown as squares. The colors in the square nodes encode the degree and history of the nodes they are attached to. Initially every node has two square nodes (insertion positions) attached to it and colors representing their degrees at the start. For instance, initially each node of degree 3 carries two white (color 1) square nodes, and each node of degree 7 carries two blue (color 2) square nodes. These colors will change over time according to a scheme that is explained later.

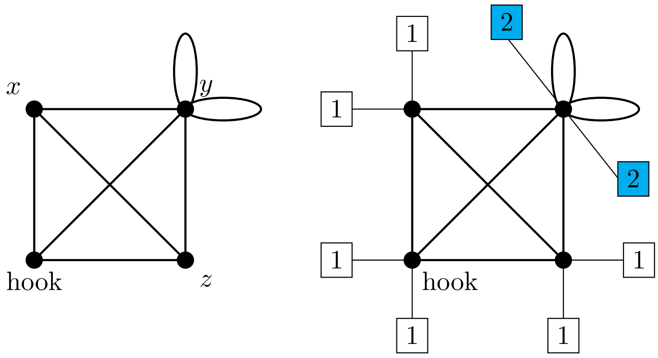

Figure 2: A seed (left) for a binary network and its extension (right).

Figure 3 shows a binary network grown from the seed in Figure 2 in four steps. The four hookings occur at the nodes labeled x (twice), y (once), and hook (once) in Figure 2. Note in Figure 3 the appearance of additional colors (to be explained later). Note also that node x no longer carries virtual nodes, as it has used all its recruiting chances, having recruited twice. There is no change at the node labeled z, but now each of the hook and y carries only one virtual node, having only recruited once (with one extra chance left at recruiting).

The seed in this particular binary hooking network has nodes of degree 3 and 7. The hook is of degree 3. Hooking the seed into nodes of degree 3 increases the degree to 6, which can ultimately increase to 9; hooking the seed into nodes of degree 7 increases the degree to 10, which can ultimately increase to 13. The six degrees 3, 7, 6, 9, 10, 13 are the only admissible degrees. A coloring scheme needs six tracking colors, one color for each admissible degree.

1.2 Organization

This section is continued in Subsection 1.3, which sets up the notation used throughout. The sequel is organized in sections. We investigate the distribution of degrees in the network G_n, for large n. It is demonstrated that asymptotically the admissible degrees have joint multivariate Gaussian laws. A paramount device for this investigation is the Pólya urn [22]. In Section 2, we say a word about Pólya urns and mention the specific results on which our investigation falls back. In Section 3, we associate a specific Pólya urn with m-ary networks.

Applying urn theory requires extensive linear algebra computation for large matrices. We take up most of the computation to find eigenvalues and eigenvectors in Section 3. The main results, presented in Section 4, are strong laws for the degrees, and asymptotic Gaussian laws, for any $m \ge 1$. In Section 5, we give some specific instances in four subsections. The examples illustrate the usual cases and properties that appear in degenerate cases. The lengthy details for the construction of the covariance matrix in the binary example in Section 5 are relegated to an appendix.

1.3 Notation

We use the notation $[r]$ to denote the set $\{1, 2, \dots, r\}$. For a real number x, we denote the falling factorial $x(x-1)\dots(x-r+1)$ by $(x)_r$; the usual interpretation of $(x)_0$ is 1.

Suppose the distinct degrees in the seed are $d_1 < d_2 < \cdots < d_k$, for some $k \ge 1$, and there are n_ℓ vertices of degree d_ℓ in the seed, for $\ell = 1, \dots, k$. Throughout, the hook is considered to be of degree $d_i =: h$, and we assume there are $n_j \ge 1$ nodes of degree d_j in the seed. Hence, the degrees that appear in the m-ary network are $d_j + sh$, for $j \in [k]$ and for $s = 0, 1, \dots, m$. Note that the numbers $d_j + sh$ need not be distinct. We present an example of this in Subsection 5.3.

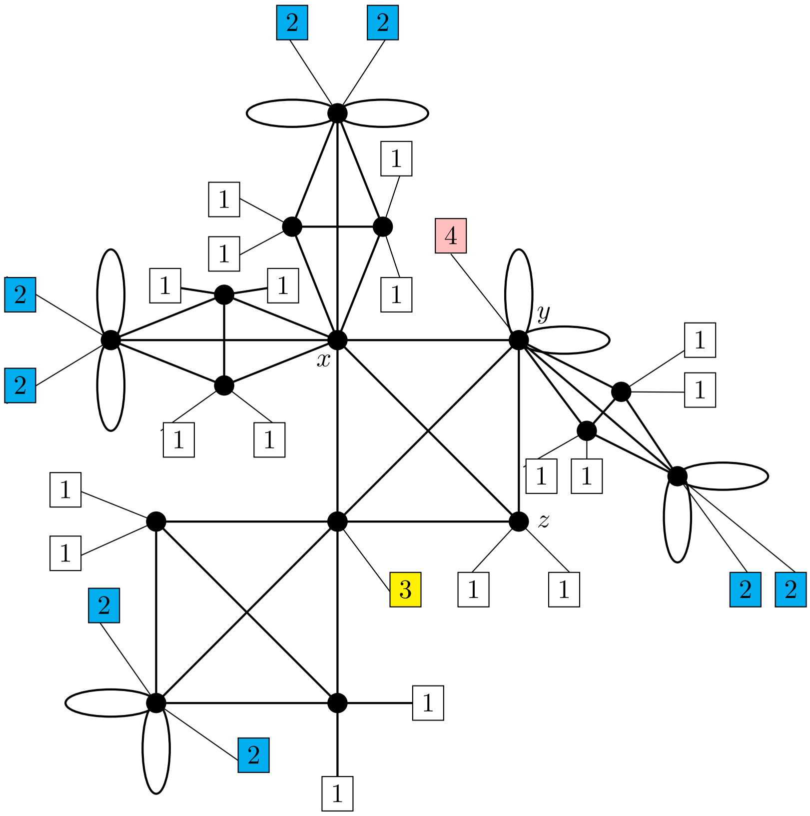

Figure 3: An extended binary network at age 4 (after four hookings). Colors 1, 2, 3, and 4 correspond to active nodes of degrees 3, 7, 6, and 10, respectively.

We print matrices and vectors in boldface, often subscripted in a way to reveal the dimension or the position in a block matrix. The distinction is clear from the context. We use $\mathbf{0}_j$ to denote either a vector of j zeros or a $j \times j$ matrix of zeros. The intent will be determined by the context. We use $\mathbf{I}_j$ for the $j \times j$ identity matrix. For a matrix $\mathbf{C}$, we denote the transpose by $\mathbf{C}^T$; the notation $|\mathbf{C}|$ is its determinant. We refer to a diagonal matrix that has the numbers $b_1, \dots, b_s$ on the diagonal as $\mathbf{Diag}(b_1, \dots, b_s)$. For a function $g(n)$, the notation $\mathbf{O}(g(n))$ and $\mathbf{o}(g(n))$ respectively represent a matrix in which all the entries are $O(g(n))$ and $o(g(n))$ in the usual sense.

Acknowledgements

The authors are thankful to Professor Alexander Boris Levin for a valuable discussion on the linear algebra. We also thank two anonymous reviewers for their insightful comments and helpful suggestions. The referees' sincere and generous effort improved the exposition.

2 Pólya urns

A multicolor Pólya urn scheme is initially nonempty. Suppose up to c colors, numbered $1, 2, \dots, c$, can appear over the course of time. At each time step, a ball is drawn at random from the urn and its color is noted. If the color of the ball withdrawn is i, we put it back in the urn and add $a_{i,j}$ balls of color $j \in [c]$, and the drawing is continued. These dynamics are captured in a $c \times c$ replacement matrix:

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,c} \\ a_{2,1} & a_{2,2} & \dots & a_{2,c} \\ \vdots & \vdots & \ddots & \vdots \\ a_{c,1} & a_{c,2} & \dots & a_{c,c} \end{pmatrix}.$$

In a general setting, the entries of the replacement matrix can be negative (which means removing balls) or even random. All the urns that appear in this paper have deterministic entries. We restrict all further presentations to fixed replacement matrices. Let $X_{n,i}$ be the number of balls of color $i \in [c]$ in the urn after n draws, and let $\mathbf{X}_n$ be the random vector with these components. The urn scheme is said to be *balanced*, if the number of balls added at each step is constant, say $\theta \ge 0$, (that is, $\sum_{j=1}^{c} a_{i,j} = \theta$, for all $i \in [c]$). The parameter θ is called the *balance factor*. The scheme is called *tenable* if it is always possible to draw balls and execute the rules. In a tenable scheme, the urn never becomes empty and the scheme never calls for taking out balls of a color, when there is not a sufficient number of balls of that color present in the urn. We shall focus only on balanced tenable urn schemes with deterministic replacement matrices, as that is all we need in our investigation of m-ary hooking networks.

Suppose the c eigenvalues of $\mathbf{A}^T$ are $\lambda_1, \dots, \lambda_c$ and are labeled according to decreasing order of their real parts. That is, we arrange the eigenvalues so that

$$\Re\,\lambda_1 \ge \Re\,\lambda_2 \ge \cdots \ge \Re\,\lambda_c.$$

The eigenvalue λ_1 is called the *principal eigenvalue* and the corresponding eigenvector $\mathbf{v}_1$ is called the *principal eigenvector*. In the context of urns, the scale of $\mathbf{v}_1$ is chosen so that $||\mathbf{v}_1||_1$ is 1. For a tenable balanced urn scheme with a deterministic replacement matrix, the principal eigenvalue is the balance factor by Perron–Frobenius theorem, i.e., $\lambda_1 = \theta$; see Chapter 8 in [17].

The number of balls in the scheme follows a strong law [1]:

$$\frac{1}{n}\mathbf{X}_n \xrightarrow{a.s.} \lambda_1 \mathbf{v}_1. \tag{1}$$

Smythe [28] and Janson [18] present a theory for classes of generalized urn models, wherein one finds joint Gaussian laws.

Under mild regularity conditions, strong laws and Gaussian limit laws for properly normalized versions of $\mathbf{X}_n$ are available in terms of the eigenvalues and eigenvectors of $\mathbf{A}^T$. In practice, the urn schemes associated with the growth of random structures are often tenable, growing, balanced and the corresponding replacement matrices have finite entries. Further, most applications involve schemes in which no color is redundant, in the sense that from the starting condition every color appears infinitely often, even if it is not present at the start. For this restricted class of urn schemes, the normality conditions are:

(1) $\lambda_1 > 0$ is real and simple.

(2) $\frac{\Re(\lambda_2)}{\lambda_1}$ is at most $\frac{1}{2}$.

These conditions are certainly met in the urn associated with an m-ary network.

The theory in [18, 28] covers urn schemes with random entries under some rather general regularity conditions. These conditions are significantly simplified in the case of tenable balanced deterministic replacement matrices. For the latter class, asymptotically there is an underlying joint multivariate normal distribution, if λ_1 is real positive and simple (of multiplicity 1) and all the components of $\mathbf{v}_1$ are positive with $\Re\,\lambda_2 < \frac{1}{2}\Re\,\lambda_1$. In such a case, we have a multivariate Gaussian law:

$$\frac{1}{\sqrt{n}}(X_n - \lambda_1 \mathbf{v}_1) \xrightarrow{\mathcal{D}} \mathcal{N}(\mathbf{0}_c, \mathbf{\Sigma}_c), \tag{2}$$

where the limit is a multivariate normal vector with mean vector $\mathbf{0}_c$, and some $c \times c$ covariance matrix $\mathbf{\Sigma_c}$.

The method in [18] for the computation of the covariance matrix associated with a Pólya urn is quite elaborate. More recently, several authors revisited these issues from an existential point of view [19, 26] and from a computational point of view [25].

The general theory in [18] covers matrices with random entries. According to [19], in the case of small-index urns (with $\lambda_2 < \frac{1}{2}\lambda_1$), $\frac{1}{n}\mathbb{C}\mathrm{ov}[\mathbf{X}_n]$ converges to the limit matrix $c \times c$, and [25] specifies the limiting matrix $\mathbf{\Sigma}_c$ as the one that solves a linear matricial equation. Specialized to the case of deterministic Pólya urns, $\mathbf{\Sigma}_c$ is the solution to the equation

$$\lambda_1 \mathbf{\Sigma}_c = \mathbf{A}^T \mathbf{\Sigma}_c + \mathbf{\Sigma}_c \mathbf{A} + \lambda_1 \mathbf{A}^T \big(\mathbf{Diag}(x_1, \dots, x_c) - \mathbf{v}_1 \mathbf{v}_1^T\big)\mathbf{A}, \tag{3}$$

where the diagonal matrix has the c components of the principal eigenvector $\mathbf{v}_1 = (x_1, \dots, x_c)^T$ on its diagonal.

3 The Pólya urn underlying m-ary hooking networks

We map the virtual nodes (corresponding to actively recruiting vertices) onto a Pólya urn on mk colors. This number of colors corresponds to the evolution of k different initial degrees, and each node of any particular initial degree can recruit m times. Ultimately, this mapping leads us to strong laws for active and inactive nodes in the network.

Recall that k is the number of distinct degrees in the seed, h is the degree of the hook, and n_ℓ is the number of vertices of degree d_ℓ in the seed, for $\ell = 1, \dots, k$. For $j \in [k]$ and $s = 0, \dots, m-1$, we associate colors $j+sk$ respectively with *active* nodes that are originally of degree d_j, and experience latching s times.

As the node degrees increase by hooking, they lose insertion positions. The fewer insertion positions left (if any) change colors.

Think of the virtual nodes as balls in a Pólya urn. The urn evolves as follows: When a virtual node of color $j \neq i$ is chosen as the insertion position at time n, an extended seed is hooked to the latch it emanates from. (Recall that the hook is of degree d_i.) We determine the replacement rules of the urn by distinguishing several cases according to the color of the drawn ball. Consider latching at a node of degree d_j, for $j \in [k]$. For $\ell \in [k]$, $\ell \neq j$ and $\ell \neq i$, the hooking adds mn_ℓ virtual nodes of color ℓ to the extended network G_n (mn_ℓ balls of color ℓ to the urn), resulting in adding n_ℓ actual nodes of degree d_ℓ each to G_n. The dynamics impose two exceptions at $\ell = j \neq i$ and at $\ell = i$. At $\ell = j$, the hooking adds only $mn_j - m$ virtual nodes of color j (balls of color j in the urn) as m virtual nodes of color j from the graph G_{n-1} are lost in G_n. Also, the incoming hook loses its m virtual nodes, so we add only $mn_i - m$ balls of color i.

The rules for the case $j = i$ are similar. We add mn_ℓ nodes of color $\ell \neq i$. The hooking occurs at a node of the hook degree d_i, with m virtual nodes lost from G_{n-1} and m more from the incoming seed. So, we only add $mn_i - 2m$ virtual balls of color i to the urn.

In the binary example of Figures 2–3, the seed has the degrees 3 and 7; the associated external nodes are encoded with colors 1 and 2, respectively. Nodes of degree 3 can progress to be of degree 6, and nodes of degree 7 can progress to be of degree 10; the corresponding external nodes are of colors 3 and 4, respectively. Then again, nodes of degree 6 can progress to be of degree 9, and nodes of degree 10 can progress to be of degree 13, meeting the quota of at most two hookings, so, they carry no external nodes, and no colors are associated with them. The scheme needs four colors in all.

Remark 3.1. *It is possible for some seed structures to have the numbers $d_j + rk$ and $d_{j'} + sk$ being equal. An example of this arises from a seed of a binary network that has the two degrees $d_1 = 2$ and $d_2 = 4$, with $k = 2$, with a hook of degree $d_1 = 2$. In this instance, $d_1 + 2 = 4$ and $d_2 = 4$ (with $j = 1, j' = 2, r = 1$ and $s = 0$). There are two types of nodes of degree 4, depending on their recruiting history. Nodes of degree 2 grow to become of degree 4 and 6, and nodes of degree 4 grow to become of degree 6 and 8. The urn gives a fine ramification of nodes of a degree like 4, distinguishing their virtual nodes (by colors) as virtual nodes of color 2 (initially attached to nodes of degree 4), and virtual nodes of color 3 (attached to nodes initially of color 1 and recruited once). The total number of nodes of degree 4, regardless of their history, is a combination of both counts.*

The replacement matrix is an $mk \times mk$ matrix represented in m^2 blocks. That is, we have $\mathbf{A} = [\mathbf{H}_{i,j}]_{1 \le i,j \le m}$, and each block $\mathbf{H}_{i,j}$ is of size $k \times k$. The foregoing discussion specifies the top left block:

$$\mathbf{H}_k = \mathbf{H}_{1,1}$$
$$= m \begin{pmatrix} n_1 - 1 & n_2 & \dots & n_{i-1} & n_i - 1 & n_{i+1} & \dots & n_k \\ n_1 & n_2 - 1 & \dots & n_{i-1} & n_i - 1 & n_{i+1} & \dots & n_k \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \\ n_1 & n_2 & \dots & n_{i-1} - 1 & n_i - 1 & n_{i+1} & \dots & n_k \\ n_1 & n_2 & \dots & n_{i-1} & n_i - 2 & n_{i+1} & \dots & n_k \\ n_1 & n_2 & \dots & n_{i-1} & n_i - 1 & n_{i+1} - 1 & \dots & n_k \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \\ n_1 & n_2 & \dots & n_{i-1} & n_i - 1 & n_{i+1} & \dots & n_k - 1 \end{pmatrix}.$$

When a virtual node of color $j \in [k]$ is taken for the next hooking, we replace m virtual nodes of that color with $m-1$ virtual nodes of color $j+k$, and make no other changes. Hence, we have the blocks

$$\mathbf{H}_{1,2} = (m-1)\mathbf{I}_k, \qquad \mathbf{H}_{1,r} = \mathbf{0}_k, \qquad \text{for } r = 3, \dots, m.$$

If a virtual node of color $j = r + sk > k$, for $r \in [k]$, $s \in [m-1]$ is the insertion position, we add mn_j virtual nodes of color j, for $j \neq i$ (they come with the hooked seed), and add only $mn_i - m$ virtual nodes of color i, as the extended seed loses its m virtual node upon hooking. So, the blocks $\mathbf{H}_{2,1}, \mathbf{H}_{3,1}, \dots, \mathbf{H}_{m,1}$ are all the same, and all equal to a matrix which we call $\mathbf{H}'_k$. The matrix $\mathbf{H}'_k$ is identical to $\mathbf{H}_k$, except that its diagonal entries have an extra m. That is, we have $\mathbf{H}'_k = \mathbf{H}_k + m\mathbf{I}_k$.

If a ball of color j, for $j = 1, \dots, k$ is drawn, we adjust the ball count of such a color according to the block $\mathbf{H}_{1,1}$. In addition, we increase the number of balls of color $j+k$ by $m-1$ as the remaining virtual nodes that are siblings of the one taken now have a history of recruiting once. In other words, the block $\mathbf{H}_{1,k}$ is set to $(m-1)\mathbf{I}_k$.

After recruiting s times, for $1 \le s \le m-1$, a node of original degree d_j has $m-s$ external nodes attached to it. These balls receive the color $j+sk$. When a ball of this color is drawn from the urn, the node it belongs to has now recruited $s+1$ times. We have one more insertion position taken, leaving only $m-s-1$ insertion positions (virtual nodes) attached to the owning node. The remaining insertion positions change colors to reflect the fact that the node they belong to has recruited $(s+1)$ times. The action in the urn is to put the drawn ball back, take out $m-s$ balls of color $j+sk$ and add to the urn $m-s-1$ balls of color $j+(s+1)k$. We have explained how the blocks $\mathbf{H}_{i,j}$, for $i = 2, \dots m-1$, $j = 2, \dots m$ are formed.

When a node recruits its mth seed, it becomes inactive. So, a node recruiting its last seed, gets lost. This is reflected as a -1 in the replacement matrix. Thus, the bottom-right $k \times k$ block is the negative of a $k \times k$ identity matrix, any other block that is not named explicitly in the preceding explanation is set to $\mathbf{0}_k$.

The full replacement matrix is

$$\mathbf{A} = \begin{array}{|c|c|c|c|c|c|c|}
\hline
\mathbf{H}_k & (m-1)\mathbf{I}_k & \mathbf{0}_k & \mathbf{0}_k & \cdots & \mathbf{0}_k & \mathbf{0}_k \\
\hline
\mathbf{H}'_k & -(m-1)\mathbf{I}_k & (m-2)\mathbf{I}_k & \mathbf{0}_k & \cdots & \mathbf{0}_k & \mathbf{0}_k \\
\hline
\mathbf{H}'_k & \mathbf{0}_k & -(m-2)\mathbf{I}_k & (m-3)\mathbf{I}_k & \cdots & \mathbf{0}_k & \mathbf{0}_k \\
\hline
\mathbf{H}'_k & \mathbf{0}_k & \mathbf{0}_k & -(m-3)\mathbf{I}_k & \cdots & \mathbf{0}_k & \mathbf{0}_k \\
\hline
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\hline
\mathbf{H}'_k & \mathbf{0}_k & \mathbf{0}_k & \mathbf{0}_k & \cdots & 2\mathbf{I}_k & \mathbf{0}_k \\
\hline
\mathbf{H}'_k & \mathbf{0}_k & \mathbf{0}_k & \mathbf{0}_k & \cdots & -2\mathbf{I}_k & \mathbf{I}_k \\
\hline
\mathbf{H}'_k & \mathbf{0}_k & \mathbf{0}_k & \mathbf{0}_k & \cdots & \mathbf{0}_k & \textbf{-}\mathbf{I}_k \\
\hline
\end{array}.$$

Note that an urn scheme with such a replacement matrix is balanced. Hence, (3) applies for the calculation of covariances.

Example 3.1. *Consider a ternary network ($m = 3$) built from the seed in Figure 2. The associated replacement matrix is*

$$\begin{pmatrix}
3 & 3 & 2 & 0 & 0 & 0 \\
6 & 0 & 0 & 2 & 0 & 0 \\
6 & 3 & -2 & 0 & 1 & 0 \\
6 & 3 & 0 & -2 & 0 & 1 \\
6 & 3 & 0 & 0 & -1 & 0 \\
6 & 3 & 0 & 0 & 0 & -1
\end{pmatrix}.$$

3.1 Eigenvalues

To find the eigenvalues of $\mathbf{A}^T$ (which are the same as those of $\mathbf{A}$), we solve for the roots of the characteristic polynomial $|\mathbf{A} - \lambda \mathbf{I}_{mk}| = 0$. Some block operations help us get through. We determine a rectangular block in $\mathbf{A}$ by specifying the position of its top left cell and bottom right cell—the block (i, j)–(k, ℓ) is comprised of all the cells (p, q), with $i \leq p \leq k$ and $j \leq q \leq \ell$. For example, $\mathbf{A}$ itself is the block $(1, 1)$–(mk, mk). Multiply each entry of the bottom $m \times mk$ block by -1 and add this block to all the blocks above it. More precisely, multiply each entry of the block $((m-1)k + 1, 1)$–(mk, mk) by -1, then add the ith row of that block to row $i + rk$, for $r = 0, \ldots, (m-1)$. This produces the matrix

$\mathbf{H}_k - \mathbf{H}'_k - \lambda\mathbf{I}_k$	$(m-1)\mathbf{I}_k$	$\mathbf{0}_k$	$\cdots$	$\mathbf{0}_k$	$\mathbf{I}_k + \lambda\mathbf{I}_k$
$\mathbf{0}_k$	$-(m-1)\mathbf{I}_k - \lambda\mathbf{I}_k$	$(m-2)\mathbf{I}_k$	$\cdots$	$\mathbf{0}_k$	$\mathbf{I}_k + \lambda\mathbf{I}_k$
$\mathbf{0}_k$	$\mathbf{0}_k$	$-(m-2)\mathbf{I}_k - \lambda\mathbf{I}_k$	$\cdots$	$\mathbf{0}_k$	$\mathbf{I}_k + \lambda\mathbf{I}_k$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$\mathbf{0}_k$	$\mathbf{0}_k$	$\mathbf{0}_k$	$\cdots$	$2\mathbf{I}_k$	$\mathbf{I}_k + \lambda\mathbf{I}_k$
$\mathbf{0}_k$	$\mathbf{0}_k$	$\mathbf{0}_k$	$\cdots$	$-2\mathbf{I}_k - \lambda\mathbf{I}_k$	$2\mathbf{I}_k + \lambda\mathbf{I}_k$
$\mathbf{H}'_k$	$\mathbf{0}_k$	$\mathbf{0}_k$	$\cdots$	$\mathbf{0}_k$	$-\mathbf{I}_k - \lambda\mathbf{I}_k$

,

which has the same characteristic polynomial as $\mathbf{A}$. Recall the relation $\mathbf{H}_k - \mathbf{H}'_k = -m\mathbf{I}_k$, and so the top left block of the latter matrix is a diagonal matrix as well. Now add the sum of the first $m - 1$ column-blocks to the mth to get the matrix

$-m\mathbf{I}_k - \lambda\mathbf{I}_k$	$(m-1)\mathbf{I}_k$	$\mathbf{0}_k$	$\cdots$	$\mathbf{0}_k$	$\mathbf{0}_k$
$\mathbf{0}'_k$	$-(m-1)\mathbf{I}_k - \lambda\mathbf{I}_k$	$(m-2)\mathbf{I}_k$	$\cdots$	$\mathbf{0}_k$	$\mathbf{0}_k$
$\mathbf{0}_k$	$\mathbf{0}_k$	$-(m-2)\mathbf{I}_k - \lambda\mathbf{I}_k$	$\cdots$	$\mathbf{0}_k$	$\mathbf{0}_k$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\mathbf{0}_k$	$\mathbf{0}_k$	$\mathbf{0}_k$	$\cdots$	$-2\mathbf{I}_k - \lambda\mathbf{I}_k$	$\mathbf{0}_k$
$\mathbf{H}'_k$	$\mathbf{0}_k$	$\mathbf{0}_k$	$\cdots$	$\mathbf{0}_k$	$\mathbf{H}'_k - \mathbf{I}_k - \lambda\mathbf{I}_k$

,

which has the same characteristic polynomial as $\mathbf{A}$. Expanding the determinant by the mth column of blocks, we get the characteristic polynomial

$$|\mathbf{H}'_k - \lambda\mathbf{I}_k - \mathbf{I}_k|\,(-m-\lambda)^k(-(m-1)-\lambda)^k \cdots (-2-\lambda)^k = 0.$$

So, $-r$, for $r = 2, 3, \ldots, m$, are eigenvalues, each with multiplicity k.

Additional eigenvalues come from the equation

$$|\mathbf{H}'_k - \lambda\mathbf{I}_k - \mathbf{I}_k| = 0.$$

These eigenvalues are determined by a method similar to the one used for the block matrices. In two steps we get a diagonal matrix, first by multiplying the bottom row by -1 and adding it to all the blocks above it, then adding the first $k - 1$ columns to the last:

$$|\mathbf{H}'_k - \lambda\mathbf{I}_k - \mathbf{I}_k|$$

$$= \begin{vmatrix} -1-\lambda & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \lambda+1 \\ 0 & -1-\lambda & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \lambda+1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & -1-\lambda & \lambda+1 \\ mn_1 & mn_2 & mn_3 & \ldots & mn_{i-1} & mn_i - m & mn_{i+1} & \ldots & mn_{k-1} & mn_k - 1 - \lambda \end{vmatrix}$$

$$= \begin{vmatrix} -1-\lambda & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -1-\lambda & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & -1-\lambda & 0 \\ mn_1 & mn_2 & mn_3 & \ldots & mn_{i-1} & mn_i - m & mn_{i+1} & \ldots & mn_{k-1} & \sum_{j=1}^{k} mn_j - m - 1 - \lambda \end{vmatrix}.$$

Expanding the determinant by the last column, we get the polynomial

$$(-\lambda - 1)^{k-1}\Big(m\sum_{j=1}^{k} n_j - m - 1 - \lambda\Big) = 0, \tag{4}$$

so the additional eigenvalues are -1 (with multiplicity $k-1$) and $m\sum_{j=1}^{k} n_j - m - 1 = m\tau_0 - m - 1$ (with multiplicity 1). To summarize, the eigenvalues of $\mathbf{A}^T$ are the same as those of $\mathbf{A}$ which are shown in Table 1.

Table 1: The eigenvalues and their multiplicities.

Eigenvalue	multiplicity
$\lambda_1 = m\tau_0 - m - 1$	1
$\lambda_2, \dots, \lambda_k = -1$	$k-1$
$\lambda_{k+1}, \dots, \lambda_{2k} = -2$	k
$\lambda_{2k+1}, \dots, \lambda_{3k} = -3$	k
$\vdots$	$\vdots$
$\lambda_{(m-1)k+1}, \dots, \lambda_{mk} = -m$	k

3.2 The principal eigenvector

Let $\mathbf{v}_1 = (x_1, \dots, x_{mk})^T$. To deal with the matrices at the level of blocks, let us consider $\mathbf{v}_1$ as m vectors, $\mathbf{y}_1, \dots, \mathbf{y}_m$, (each of k components) stacked atop of each other. That is, $\mathbf{y}_{s+1}$ is the segment $(x_{1+sk}, \dots, x_{k+sk})^T$, for $s = 0, \dots, m-1$. We now solve

$$\left[\begin{array}{|c|c|c|c|c|c|}
\hline
\mathbf{H}_k^T & (\mathbf{H}'_k)^T & (\mathbf{H}'_k)^T & (\mathbf{H}'_k)^T & \cdots & (\mathbf{H}'_k)^T \\
\hline
(m-1)\mathbf{I}_k & -(m-1)\mathbf{I}_k & \mathbf{0}_k & \mathbf{0}_k & \cdots & \mathbf{0}_k \\
\hline
\mathbf{0}_k & (m-2)\mathbf{I}_k & -(m-2)\mathbf{I}_k & \mathbf{0}_k & \cdots & \mathbf{0}_k \\
\hline
\mathbf{0}_k & \mathbf{0}_k & (m-3)\mathbf{I}_k & \mathbf{0}_k & \cdots & \mathbf{0}_k \\
\hline
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\hline
\mathbf{0}_k & \mathbf{0}_k & \mathbf{0}_k & \mathbf{0}_k & \dots & \text{-}\mathbf{I}_k \\
\hline
\end{array}\right]
\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_m \end{pmatrix} = \lambda_1 \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_m \end{pmatrix}.$$

For $r = 2, \dots, m$, the rth row of blocks gives us the equation

$$(m - r + 1)\mathbf{y}_{r-1} - (m - r + 1)\mathbf{y}_r = \lambda_1 \mathbf{y}_r,$$

which is the same as

$$\mathbf{y}_r = \frac{m - r + 1}{m - r + 1 + \lambda_1}\mathbf{y}_{r-1}.$$

Thus, recursively, we get all the segments in terms of $\mathbf{y}_1$. Namely, we have

$$\mathbf{y}_r = \frac{(m-1)_{r-1}}{(m - 1 + \lambda_1)_{r-1}}\mathbf{y}_1 = \frac{(m-1)_{r-1}}{(m\tau_0 - 2)_{r-1}}\mathbf{y}_1. \tag{5}$$

We can then get $\mathbf{y}_1$ from the first row of blocks as the solution to

$$\mathbf{H}_k^T\mathbf{y}_1 + (\mathbf{H}'_k)^T\mathbf{y}_2 + \cdots + (\mathbf{H}'_k)^T\mathbf{y}_m = \lambda_1\mathbf{y}_1.$$

The ith row (corresponding to an active raw hook that has not yet recruited) of this matricial equation reads

$$\begin{aligned}(mn_i - m)x_1 + (mn_i - m)x_2 + \cdots + (mn_i - m)x_{i-1} + (mn_i - 2m)x_i \\ + (mn_i - m)x_{i+1} + \cdots + (mn_i - m)x_{mk} = \lambda_1 x_i.\end{aligned}$$

Using the fact that $||\mathbf{v}_1||_1$ is normalized to 1, we rearrange to get

$$m(n_i - 1)\sum_{j=1}^{mk} x_j - mx_i = m(n_i - 1) - mx_i = \lambda_1 x_i.$$

Plugging in the value of the principal eigenvalue, we obtain

$$x_i = \frac{m(n_i - 1)}{m\tau_0 - 1}.$$

We next turn to the jth row, for $j \in [k]$, and $j \neq i$ to find

$$mn_j x_1 + \cdots + mn_j x_{j-1} + m(n_j - 1)x_j + mn_j x_{j+1} + \cdots + mn_j x_{mk} = \lambda_1 x_j.$$

By manipulation similar to the case of x_i, we conclude that

$$x_j = \frac{mn_j}{m\tau_0 - 1}.$$

We have determined the segment $\mathbf{y}_1$. It is

$$\mathbf{y}_1 = \frac{m}{m\tau_0 - 1} \begin{pmatrix} n_1 \\ n_2 \\ \vdots \\ n_{i-1} \\ n_i - 1 \\ n_{i+1} \\ \vdots \\ n_k \end{pmatrix}. \tag{6}$$

4 Strong laws and joint Gaussian distributions

Let D_{n,d_j+sh} be the number of nodes that are originally of degree d_j, and experience latching s times in an m-ary network at age n, for $j \in [k]$, $s = 0, 1, \dots, m$. The urn scheme is associated only with the active nodes. An active node of degree $d_j + sh$ has $m - s$ virtual nodes attached to it. Therefore, we have $D_{n,d_j+sh} = X_{n,j+sk}/(m - s)$, with $0 \le s < m$, where $X_{n,j+sk}$ is the number of balls of color $j + sk$ in the urn after n draws. Inactive nodes cannot recruit any more and do not have corresponding balls in the urn. Some extra work is needed to determine a strong law for inactive nodes.

As we shall see in later examples, the matrix $\mathbf{A}^T$ may not be invertible.* When $(\mathbf{A}^T)^{-1}$ exists, we call the network *invertible*.

Theorem 4.1. *Let D_{n,d_j+sh} be the number of nodes that are originally of degree d_j, and experience latching s times in an invertible m-ary network at age n, for $j \in [k]$, $s = 0, 1, \dots, m$, and let $\mathbf{D}_n$ be the vector with these components. Suppose the principal eigenvector of the urn associated with the network is $\mathbf{v}_1 = (x_1, \dots, x_{mk})^T$. The degree vector follows the strong law*

$$\frac{1}{n}\mathbf{D}_n = \frac{1}{n} \begin{pmatrix} D_{n,d_1} \\ D_{n,d_2} \\ \vdots \\ D_{n,d_k} \\ D_{n,d_1+h} \\ D_{n,d_2+h} \\ \vdots \\ D_{n,d_k+mh} \end{pmatrix} \xrightarrow{a.s.} \mathbf{D}^* := \alpha_m \begin{pmatrix} \frac{(m-1)_0}{(m\tau_0-2)_0}\,\mathbf{y}_1 \\ \frac{(m-1)_1}{(m\tau_0-2)_1}\,\mathbf{y}_1 \\ \vdots \\ \frac{(m-1)_{m-1}}{(m\tau_0-2)_{m-1}}\,\mathbf{y}_1 \\ \frac{(m-1)_{m-1}}{\alpha_m(m\tau_0-2)_{m-1}}\,\mathbf{y}_1 \end{pmatrix},$$

where $\alpha_m = m\tau_0 - m - 1$, and $\mathbf{y}_1$ is given in (6).

*Invertability excludes only the case $\lambda_1 = 0$, which means $m(\tau_0 - 1) = 1$. This can only happen, if $m = 1$ with $\tau_0 = 2$.

Proof. Let $X_{n,r}$ be the number of virtual nodes of color $r \in [mk]$ in the extended network G_n at time n (balls of color r in the urn after n draws), and denote by $\mathbf{X}_n$ the vector with these mk components.

By (1), we have

$$\frac{1}{n}\mathbf{X}_n \xrightarrow{a.s.} \lambda_1 \mathbf{v}_1 = (m\tau_0 - m - 1)\mathbf{v}_1.$$

We determine the principal eigenvector $\mathbf{v}_1$ by the recursive algorithm in (5), with (6) at the basis of the induction.

The vector of color counts relates to the degrees in the following way. For each active node in the graph of degree $d_j + sh$, for $j \in [k]$ and $s = 0, \dots, m-1$, there are $m - s$ virtual nodes of color $j + sk$ attached to it. So, D_{n,d_j+sh} is $X_{n,j+sk}/(m-s)$, specifying the first mk components of $\mathbf{D}_n$, as in the statement of the theorem.

For the inactive nodes, of degrees $d_1 + mh, \dots, d_k + mh$, their counts relate to the principal eigenvector in the following way. Let $Y_{n,j}$ be the number of times a ball of color $j \in [mk]$ is drawn from the urn by time n, and let $\mathbf{Y}_n$ be the vector with these components. A pick from color $r \in [k]$ (with $r \neq i$) contributes mn_j balls of color j in the urn, except at $r = j$, or $r = i$, where the contribution is reduced by m, owing to the hooking. This gives us the relation

$$X_{n,j} = \sum_{\substack{r=1 \\ r\neq j, r\neq i}}^{k} mn_j Y_{n,r} + (mn_i - m)Y_{n,i} + (mn_j - m)Y_{n,j} + X_{0,j}.$$

Arguing similarly, when $j = i$, we obtain

$$X_{n,i} = \sum_{\substack{r=1 \\ r\neq i}}^{k} mn_j Y_{n,r} + (mn_i - 2m)Y_{n,i} + X_{0,i}.$$

Collected in matrix notation, these relations are

$$\mathbf{X}_n = \mathbf{A}^T \mathbf{Y}_n + \mathbf{X}_0. \tag{7}$$

When the inverse of $\mathbf{A}^T$ exists, we can invert the relation into

$$\mathbf{Y}_n = (\mathbf{A}^T)^{-1}(\mathbf{X}_n - \mathbf{X}_0), \tag{8}$$

and a strong law for $\mathbf{Y}_n$ ensues:

$$\frac{1}{n}\mathbf{Y}_n = \frac{1}{n}(\mathbf{A}^T)^{-1}(\mathbf{X}_n - \mathbf{X}_0) \xrightarrow{a.s.} (\mathbf{A}^T)^{-1}(\lambda_1 \mathbf{v}_1) = (\mathbf{A}^T)^{-1}(\mathbf{A}^T \mathbf{v}_1) = \mathbf{v}_1.$$

Note that $Y_{n,j+mk} = D_{n,d_j+mh}$. Therefore, the last k components of $\mathbf{D}_n$ are the last k components of the principle eigenvector:

$$\frac{1}{n}\begin{pmatrix} D_{d_1+mh} \\ D_{d_2+mh} \\ \vdots \\ D_{d_k+mh} \end{pmatrix} \xrightarrow{a.s.} \begin{pmatrix} x_{(m-1)k+1} \\ x_{(m-1)k+2} \\ \vdots \\ x_{mk} \end{pmatrix}.$$

□

Remark 4.1. *Theorem 4.1 requires the invertibility of* $\mathbf{A}^T$*. If* $\mathbf{A}^T$ *is not invertible, the associated urn gives us only strong laws for the degrees of active nodes and determining the number of inactive nodes requires some other reasoning.*

Example 4.1. *An example of noninvertible networks in Remark 4.1 is a degenerate unary network* ($m = 1$) *grown out of a path of length 1. This is a degenerate network that grows as a path and the only active nodes in it are the two vertices at the two ends of the path. In this instance, we have* $\mathbf{A} = [0]$*, and* $\mathbf{A}^T$ *has no inverse. The urn argument gives us the strong law* $D_{n,1}/n \xrightarrow{a.s.} 0$*. Here, we have the relation* $D_{n,2} = nY_{n,1}$*, and* $D_{n,2}/n \to 1$*.*

The urn scheme associated with an m-ary network is a generalized scheme in the sense discussed in Section 2. In particular, we have $\Re\lambda_2 < \frac{1}{2}\Re\lambda_1$. So, when all the components of the principal eigenvector are positive, the central limit theorem in (2) applies to the active nodes. When one of the components of $\mathbf{v}_1$ is 0, we call the network *degenerate*. In the present work, a sufficient condition for nondegeneracy is $n_i > 1$. Note the distinction between degeneracy and invertibility.

Theorem 4.2. *Let D_{n,d_j+sh} be the number of nodes that are originally of degree d_j, and experience latching s times in a nondegenerate m-ary network at age n, for $j \in [k]$, $s = 0, 1, \ldots, m$. Let $\mathbf{D}_n$ be the vector with these components, and $\mathbf{D}^*$ be the almost-sure limit in Theorem 4.1. As $n \to \infty$, we have*

$$\frac{1}{\sqrt{n}}(\mathbf{D}_n - n\mathbf{D}^*) \xrightarrow{\mathcal{D}} \mathcal{N}(\mathbf{0}_{mk}, \mathbf{\Sigma}_{mk}),$$

where $\mathbf{\Sigma}_{mk}$ is an $mk \times mk$ covariance matrix.

Proof. We have proved that a central limit theorem applies to the active nodes. By (8), the components of $\mathbf{Y}_n$ are linear combinations of the components of $\mathbf{X}_n$. It follows that $\mathbf{Y}_n$ also asymptotically follows a multivariate Gaussian law. In particular, the vector of the numbers of inactive nodes (the last k components of $\mathbf{Y}_n$) follows a multivariate Gaussian law and the mk components of $\mathbf{D}_n$ are together asymptotically Gaussian. □

5 Specific instances with small m

The calculations outlined in the previous sections are tractable for small m. We consider the unary and binary cases.

5.1 Unary hooking networks

Consider the case $m = 1$. As a corollary to Theorem 4.1, we get a strong law for the degrees in a random unary hooking network.

Corollary 5.1. *Let D_{n,d_j+sh} be the number of nodes that are originally of degree d_j, and experience latching s times in a nondegenerate unary network at age n, for $j \in [k]$ and $s = 0, 1$. Then, we have*

$$\frac{1}{n}\begin{pmatrix} D_{n,d_1} \\ D_{n,d_2} \\ \vdots \\ D_{n,d_k} \\ D_{n,d_1+h} \\ D_{n,d_2+h} \\ \vdots \\ D_{n,d_k+h} \end{pmatrix} \xrightarrow{a.s.} \frac{1}{\tau_0 - 1}\begin{pmatrix} (\tau_0 - 2)n_1 \\ \vdots \\ (\tau_0 - 2)n_{i-1} \\ (\tau_0 - 2)(n_i - 1) \\ (\tau_0 - 2)n_{i+1} \\ \vdots \\ (\tau_0 - 2)n_k \\ n_1 \\ \vdots \\ n_{i-1} \\ n_i - 1 \\ n_{i+1} \\ \vdots \\ n_k \end{pmatrix}.$$

Note the degenerate case $n_i = 1$. In this case, the ith component of the limiting vector is 0. So, G_0 has one node of degree d_i (the degree of the hook in the seed), and sooner or later it will progress by fusing to become an inactive node of degree $2d_i$.

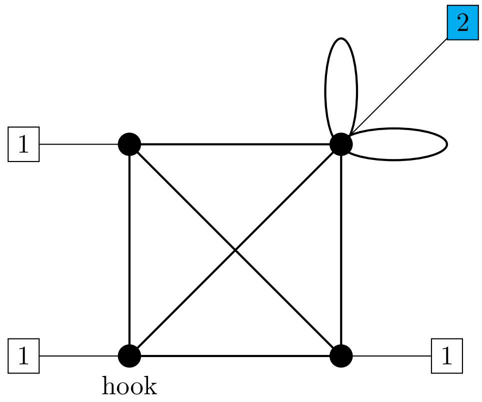

Figure 4: An extended seed for a unary network.

5.2 A concrete nondegenerate unary case

Let the seed be the one shown in Figure 4. The figure shows the extended seed, with virtual nodes at the insertion positions colored white (color 1) and blue (color 2).

In this example, we have $k = 2$, $d_1 = 3$, $d_2 = 7$, and $\tau_0 = 4$. There are $n_1 = 3$ nodes of degree 3 and only $n_2 = 1$ node of degree 7. The admissible degrees are $3, 6, 7, 10$. Nodes of degree $d_3 = 6$ and $d_4 = 10$ are inactive—they do not recruit after they appear. The only active nodes are of degrees 3 and 7. We need only two colors, one (white, color 1) to correspond to nodes of degree 3 and another (blue, color 2) to correspond to nodes of degree 7. The replacement matrix is

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 2 & 0 \end{pmatrix}.$$

This is a nondegenerate invertible unary case, as $(\mathbf{A}^T)^{-1}$ exists. By Corollary 5.1, we have

$$\frac{1}{n}\begin{pmatrix} D_{n,3} \\ D_{n,7} \\ D_{n,6} \\ D_{n,10} \end{pmatrix} \xrightarrow{a.s.} \frac{1}{3}\begin{pmatrix} 4 \\ 2 \\ 2 \\ 1 \end{pmatrix}.$$

In this instance, the replacement matrix is small and amenable to straightforward variance calculation. In fact, this type of urn is known in the literature as a Bagchi–Pal urn, first analyzed in [2]. That paper gives a simple formula for the variance of the number of white balls. According to [2], $\mathbb{V}\text{ar}[D_{n,3}] = \mathbb{V}\text{ar}[X_{n,1}] = 1/9$. We add two balls after each drawing, and we have

$$X_{n,1} + X_{n,2} = D_{n,3} + D_{n,7} = 2n + 4.$$

The linear dependence tells us that $\mathbb{V}\text{ar}[D_{n,3}] = \mathbb{V}\text{ar}[D_{n,7}]$ and $\mathbb{C}\text{ov}[D_{n,3}, D_{n,7}] = -\mathbb{V}\text{ar}[X_{n,1}] = -1/9$. Further, by (7), we have

$$\begin{aligned} D_{n,3} &= Y_{n,1} + 2Y_{n,2} + 3 = D_{n,6} + 2D_{n,10} + 3, \\ D_{n,7} &= Y_{n,1} + 1 = D_{n,6} + 1. \end{aligned}$$

With these relations, we complete the covariance matrix for the central limit theorem:

$$\frac{1}{\sqrt{n}}\left(\begin{pmatrix} D_{n,3} \\ D_{n,7} \\ D_{n,6} \\ D_{n,10} \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 4 \\ 2 \\ 2 \\ 1 \end{pmatrix}\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{9}\begin{pmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1, & -1 & -1 & 1 \end{pmatrix}\right).$$

5.3 A case with degeneracy

Consider the extended seed shown in Figure 5, from which we build a unary network. In this example, we have $k = 3$ distinct degrees in the seed, which are $d_1 = 1, d_2 = 2$, and $d_3 = 3$. There are $n_1 = 1$ node of degree 1, $n_2 = 2$ nodes of degree 2, and $n_3 = 1$ node of degree 3. We need three colors for the active nodes, with a ball of color j corresponding to an active node of degree j, for $j = 1, 2, 3$.

The replacement matrix is

$$\mathbf{A} = \begin{pmatrix} -1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 2 & 0 \end{pmatrix}.$$

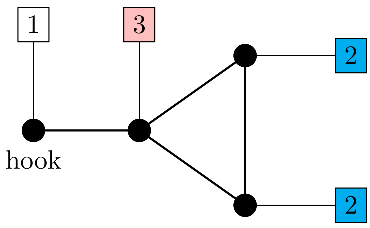

Figure 5: A seed leading to a degenerate unary network.

By (1), we obtain

$$\frac{1}{n}\begin{pmatrix} X_{n,1} \\ X_{n,2} \\ X_{n,3} \end{pmatrix} \xrightarrow{a.s.} \frac{1}{3}\begin{pmatrix} 0 \\ 4 \\ 2 \end{pmatrix}.$$

Let us distinguish nodes of degree 2 as active, with the count $\hat{D}_{n,2}$, and inactive with count $\tilde{D}_{n,2}$. Likewise, we distinguish nodes of degree 3 as active, with the count $\hat{D}_{n,3}$, and inactive with count $\tilde{D}_{n,3}$. All nodes of degree 4 are inactive. The actual number of nodes of degree 2 is $D_{n,2} = \hat{D}_{n,2} + \tilde{D}_{n,2}$, and the actual number of degree 3 is $D_{n,3} = \hat{D}_{n,3} + \tilde{D}_{n,3}$.

The network in this example has $\tau_0 = 4$. By Corollary 5.1, we have the strong convergence

$$\frac{1}{n}\begin{pmatrix} D_{n,1} \\ \hat{D}_{n,2} \\ \hat{D}_{n,3} \\ \tilde{D}_{n,2} \\ \tilde{D}_{n,3} \\ D_{n,4} \end{pmatrix} \xrightarrow{a.s.} \frac{1}{3}\begin{pmatrix} 0 \\ 4 \\ 2 \\ 0 \\ 2 \\ 1 \end{pmatrix},$$

from which we find the limiting vector of degrees:

$$\frac{1}{n}\begin{pmatrix} D_{n,1} \\ D_{n,2} \\ D_{n,3} \\ D_{n,4} \end{pmatrix} \xrightarrow{a.s.} \frac{1}{3}\begin{pmatrix} 0 \\ 4 \\ 4 \\ 1 \end{pmatrix}.$$

As one of the components of $\mathbf{v}_1$ is 0, the central limit theorem in Smythe [28] does not apply. However, by an alternative formulation in Janson (Theorem 3.22 in [18]) we still get a central limit theorem.

For the inactive nodes (of degrees 2, 3 and 4), we use the construction in (8). Specialized to the unary network at hand, the limiting covariance of the vector on the left of (8) follows:

$$\frac{1}{n}\mathbb{C}\mathrm{ov}\left[\begin{pmatrix}Y_{n,1}\\Y_{n,2}\\Y_{n,3}\end{pmatrix}\right] \to (\mathbf{A}^T)^{-1}\lim_{n\to\infty}\left(\frac{1}{n}\mathbb{C}\mathrm{ov}[\mathbf{X}_n]\right)\left((\mathbf{A}^T)^{-1}\right)^T$$
$$= \frac{1}{9}\begin{pmatrix}0 & 0 & 0\\0 & 1 & -1\\0 & -1 & 1\end{pmatrix}.$$

For notational convenience, we refer to the elements of the matrix on the right as $y_{r,s}$, for $r,s \in \{1,2,3\}$.

The three components of the vector on the left-hand side of (8) are $Y_{n,1} = \tilde{D}_{n,2}$, $Y_{n,2} = \tilde{D}_{n,3}$ and $Y_{n,3} = \tilde{D}_{n,4}$. So, the last matrix is the bottom right 3×3 block in the full 6×6 limiting covariance matrix among all the nodes (active and inactive). Let us call that latter 6×6 matrix $\mathbf{G}$, and refer to its elements by $g_{r,s}$, for $r,s \in [6]$.

As we can only have at most one inactive node of degree 1, and at most one inactive node of degree 2, we have $\frac{1}{n}D_{n,1} \xrightarrow{a.s.} 0$, and $\frac{1}{n}\tilde{D}_{n,2} \xrightarrow{a.s.} 0$. So,

$$\frac{1}{n}\mathbb{V}\mathrm{ar}[D_{n,1}] \to 0, \qquad \frac{1}{n}\mathbb{V}\mathrm{ar}[\tilde{D}_{n,2}] \to 0.$$

Also any limiting covariances involving these degrees are 0. Reading off (7) component-wise, we get three separate equations:

$$X_{n,1} = -Y_{n,1} + 1, \tag{9}$$
$$X_{n,2} = 2Y_{n,1} + Y_{n,2} + 2Y_{n,3} + 2, \tag{10}$$
$$X_{n,3} = Y_{n,1} + Y_{n,2} + 1. \tag{11}$$

From these three equations, we can get the top left 3×3 block of the full 6×6 limiting covariance. For instance, by taking the variance of (10), and scaling by n^{-1}, we get

$$\frac{1}{n}\mathbb{V}\mathrm{ar}[X_{n,2}] \to 4y_{1,1} + y_{2,2} + 4y_{3,3} + 4y_{1,2} + 8y_{1,3} + 4y_{2,3} = \frac{1}{9};$$

recall that $y_{1,1} = 0$, $y_{1,2} = 0$, and $y_{1,3} = 0$.

Adding any two of the equations (9)–(11), gives us one covariance in the top left 3×3 block of the full 6×6 matrix $\mathbf{G}$.

Reorganize (10) in the form

$$X_{n,2} - Y_{n,2} = 2Y_{n,1} + 2Y_{n,3} + 2,$$

and take the variance, to get

$$\mathbb{V}\mathrm{ar}[X_{n,2}] + \mathbb{V}\mathrm{ar}[Y_{n,2}] - 2\,\mathbb{C}\mathrm{ov}[X_{n,2}, Y_{n,2}]$$
$$= 4\,\mathbb{V}\mathrm{ar}[Y_{n,1}] + 4\,\mathbb{V}\mathrm{ar}[Y_{n,3}] + 8\,\mathbb{C}\mathrm{ov}[Y_{n,1}, Y_{n,3}].$$

Taking the limit at a scale of n^{-1}, we get

$$\frac{1}{n}\mathbb{C}\mathrm{ov}[\hat{D}_{n,2}, \tilde{D}_{n,3}] \to g_{2,5} = -\frac{1}{2}(4y_{1,1} + 4y_{3,3} + 8y_{1,3} - y_{2,2} - g_{2,2}) = -\frac{1}{9};$$

recall that $y_{1,1} = 0$ and $y_{1,3} = 0$.

Reorganize (10) in a form separating $2Y_{n,3}$ on the left, a similar computation gives $\mathbb{C}\mathrm{ov}[\hat{D}_{n,2}, D_{n,4}]$. Taking the limit at a scale of n^{-1}, we get

$$\frac{1}{n}\mathbb{C}\mathrm{ov}[\hat{D}_{n,2}, D_{n,4}] \to g_{2,5} = -\frac{1}{4}(4y_{1,1} + y_{2,2} + 4y_{1,2} - 4y_{3,3} - g_{2,2}) = \frac{1}{9};$$

recall that $y_{1,1} = 0$ and $y_{1,2} = 0$.

The trickiest elements of $\mathbf{G}$ are $g_{3,5} = g_{5,3}$ and $g_{3,6} = g_{6,3}$. No individual equation among (9)–(11) has sufficient information to determine these elements, but a combination of (10) and (11) gives us what we want. Add these two equations, with $2Y_{n,2}$ written on the left of (10), scale by n^{-1} and take the limits, to get

$$g_{2,2} + g_{3,3} + 2g_{2,3} + 4y_{2,2} - 4g_{2,6} - \lim_{n\to\infty} \frac{4}{n}\mathbb{V}\text{ar}[\hat{D}_{n,2}, \tilde{D}_{n,3}] = 9y_{1,1} + 4y_{3,3} + 12y_{1,3};$$

recall that $y_{1,1} = 0$ and $y_{1,3} = 0$. The remaining limit is $g_{3,5} = -\frac{1}{9}$. A similar manipulation, with $2Y_{n,3}$ on the left, gives us $g_{3,6} = \frac{1}{9}$. The complete covariance matrix is

$$\mathbf{G} = \lim_{n\to\infty} \frac{1}{n}\mathbb{C}\text{ov}\left[\begin{pmatrix} D_{n,1} \\ \hat{D}_{n,2} \\ \hat{D}_{n,3} \\ \tilde{D}_{n,2} \\ \tilde{D}_{n,3} \\ D_{n,4} \end{pmatrix}\right] = \frac{1}{9}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & -1 & 1 \\ 0 & -1 & 1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 & 1 & -1 \\ 0 & 1 & 1 & 0 & -1 & 1 \end{pmatrix}.$$

The graph theory point of view does not distinguish node degrees by history. A network analyst may wish to make decisions based on the plain degrees. For instance, for the unary network grown from the seed in Figure 5, in a social context all nodes of degree 3 may mean three friends. The analyst is only interested in the limiting value of $\frac{1}{n}\mathbb{C}\text{ov}[\mathbf{D}_n]$. The entries of this matrix are related to $\mathbf{G}$. For example $D_{n,2} = \hat{D}_{n,2} + \tilde{D}_{n,2}$, giving us the limiting covariance relation

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{V}\text{ar}[D_{n,2}] = g_{2,2} + g_{4,4} + g_{2,4} = \frac{1}{9}.$$

Completing the computations for all the entries of $\lim_{n\to\infty} \frac{1}{n}\mathbb{C}\text{ov}[\mathbf{D}_n]$, we get the central limit theorem

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}}\left(\begin{pmatrix} D_{n,1} \\ D_{n,2} \\ D_{n,3} \\ D_{n,4} \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 0 \\ 4 \\ 4 \\ 1 \end{pmatrix}\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{9}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & -2 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}\right).$$

Remark 5.1. *Curiously,* $\frac{1}{n}\mathbb{V}\text{ar}[D_{n,3}] \to 0$*. A moment of reflection leads us to see that active and inactive nodes of degree 3 appear together in every step, once the initial hook becomes inactive and the node of degree 3 in the seed copy latched into it also recruits. That is, after the latter event,* $D_{n,3} = \hat{D}_{n,3} + \tilde{D}_{n,3}$ *behaves deterministically.*

5.4 A binary case

Consider the extended seed shown in Figure 2. It is the same seed we considered in the first example on unary networks but now we build a random binary network out of it. In this binary example, we have $k = 2$ distinct degrees in the seed, which are $d_1 = 3$, and $d_2 = 7$. There are $n_1 = 3$ nodes of degree 3 and $n_2 = 1$ node of degree 7. We need four colors for the active nodes, with color $1, 2, 3, 4$ respectively corresponding to the active nodes of degrees $3, 7, 6, 10$.

The replacement matrix is

$$\mathbf{A} = \begin{pmatrix} 2 & 2 & 1 & 0 \\ 4 & 0 & 0 & 1 \\ 4 & 2 & -1 & 0 \\ 4 & 2 & 0 & -1 \end{pmatrix}.$$

According to Theorem 4.1, we get

$$\frac{1}{n}\mathbf{D}^*_n = \frac{1}{n}\begin{pmatrix} D_{n,3} \\ D_{n,7} \\ D_{n,6} \\ D_{n,10} \\ D_{n,9} \\ D_{n,13} \end{pmatrix} \xrightarrow{a.s.} \frac{1}{21}\begin{pmatrix} 60 \\ 30 \\ 10 \\ 5 \\ 2 \\ 1 \end{pmatrix} =: \mathbf{D}^*.$$

The limiting covariance matrix is computed in the appendix. The corresponding central limit theorem is

$$\frac{1}{\sqrt{n}}\left(\begin{pmatrix} D_{n,3} \\ D_{n,7} \\ D_{n,6} \\ D_{n,10} \\ D_{n,9} \\ D_{n,13} \end{pmatrix} - \frac{1}{21}\begin{pmatrix} 60 \\ 30 \\ 10 \\ 5 \\ 2 \\ 1 \end{pmatrix}\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \frac{5}{882}\begin{pmatrix} 24 & -16 & -27 & 11 & 3 & 5 \\ -16 & 20 & 11 & -19 & 5 & -1 \\ -27 & 11 & 40 & -8 & -13 & -3 \\ 11 & -19 & -8 & 24 & -3 & -5 \\ 3 & 5 & -13 & -3 & 10 & -2 \\ 5 & -1 & -3 & -5 & -2 & 6 \end{pmatrix}\right).$$

6 Concluding remarks

We studied random m-ary networks that grow from a seed, and each node in the network has m hooking positions. The degrees in the network evolve over time. Via a connection to the composition of certain Pólya urns, we are able to extract theorems for the degrees in the m-ary network. Namely a strong law (Theorem 4.1) and a multivariate central limit theorem (Theorem 4.2) were developed for m-ary networks.

Appendix

We illustrate the procedure in (3) on the binary network in Section 5.4. The limit $\mathbf{\Sigma}_c$ is 4×4, with the elements q_{ij}, for $i, j \in [4]$.

The urn associated with this network has $\lambda_1 = \theta = 5$, and the principal eigenvector is $(x_1, x_2, x_3, x_4)^T = \frac{1}{21}(12, 6, 2, 1)^T$. We wish to solve the matricial equation

$$\mathbf{A}^T\mathbf{\Sigma}_4 + \mathbf{\Sigma}_4\mathbf{A} + \lambda_1\mathbf{A}^T\mathbf{Diag}(x_1, x_2, x_3, x_4)\mathbf{A} - \lambda_1\mathbf{A}^T\mathbf{v}_1\mathbf{v}_1^T\mathbf{A} - \lambda_1\mathbf{\Sigma}_4 = \mathbf{0}_4.$$

We extract elements from the left-hand side and equate them to 0. It is enough to extract the 10 elements on and above the diagonal, since $\mathbf{\Sigma}_4 = [q_{i,j}]_{1\le i,j\le 4}$ is symmetric. Going through this extraction, we get

$$\begin{aligned}
-q_{1,1} + 8q_{1,2} + 8q_{1,3} + 8q_{1,4} + \frac{240}{49} &= 0, \\
2q_{1,1} + 2q_{1,3} + 2q_{1,4} - 3q_{1,2} + 4q_{2,2} + 4q_{2,3} + 4q_{2,4} - \frac{160}{49} &= 0, \\
q_{1,1} - 4q_{1,3} + 4q_{2,3} + 4q_{3,3} + 4q_{3,4} - \frac{440}{147} &= 0, \\
q_{1,2} - 4q_{1,4} + 4q_{2,4} + 4q_{3,4} + 4q_{4,4} + \frac{200}{147} &= 0, \\
4q_{1,2} + 4q_{2,3} + 4q_{2,4} - 5q_{2,2} + \frac{200}{49} &= 0, \\
q_{1,2} - 6q_{2,3} + 2q_{1,3} + 2q_{3,3} + 2q_{3,4} + \frac{200}{147} &= 0, \\
q_{2,2} - 6q_{2,4} + 2q_{1,4} + 2q_{3,4} + 2q_{4,4} - \frac{320}{147} &= 0, \\
2q_{1,3} - 7q_{3,3} + \frac{970}{441} &= 0, \\
q_{2,3} - 7q_{3,4} + q_{1,4} - \frac{250}{441} &= 0, \\
2q_{2,4} - 7q_{4,4} + \frac{610}{441} &= 0.
\end{aligned}$$

This is a standard system of linear equations, with the solution

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{C}\mathrm{ov}\left[\begin{pmatrix} X_{n,1} \\ X_{n,2} \\ X_{n,3} \\ X_{n,4} \end{pmatrix}\right] = \frac{5}{441}\begin{pmatrix} 48 & -32 & -27 & 11 \\ -32 & 40 & 11 & -19 \\ -27 & 11 & 20 & -4 \\ 11 & -19 & -4 & 12 \end{pmatrix}.$$

We remind the reader that colors $1, 2, 3, 4$ correspond to degrees $3, 7, 6, 10$ respectively and that $Y_{n,3}$ and $Y_{n,4}$ correspond to $D_{n,9}$ and $D_{n,13}$. Additionally, with

$$D_{n,3} = \frac{1}{2} X_{n,1}, \qquad D_{n,7} = \frac{1}{2} X_{n,2}, \qquad D_{n,6} = X_{n,3}, \qquad D_{n,10} = X_{n,4},$$

we have the variances and covariances in the top left 4×4 block:

$$\begin{aligned}
\frac{1}{n}\mathbb{V}\mathrm{ar}[D_{n,3}] &= \frac{1}{4n}\mathbb{V}\mathrm{ar}[X_{n,1}] \to \frac{20}{147}, \qquad \frac{1}{n}\mathbb{V}\mathrm{ar}[D_{n,7}] = \frac{1}{4n}\mathbb{V}\mathrm{ar}[X_{n,2}] \to \frac{50}{441}, \\
\frac{1}{n}\mathbb{V}\mathrm{ar}[D_{n,6}] &= \frac{1}{n}\mathbb{V}\mathrm{ar}[X_{n,3}] \to \frac{100}{441}, \qquad \frac{1}{n}\mathbb{V}\mathrm{ar}[D_{n,10}] = \frac{1}{n}\mathbb{V}\mathrm{ar}[X_{n,4}] \to \frac{20}{147}, \\
\frac{1}{n}\mathbb{C}\mathrm{ov}[D_{n,3}, D_{n,7}] &= \frac{1}{4n}\mathbb{C}\mathrm{ov}[X_{n,1}, X_{n,2}] \to -\frac{40}{441}, \\
\frac{1}{n}\mathbb{C}\mathrm{ov}[D_{n,3}, D_{n,6}] &= \frac{1}{2n}\mathbb{C}\mathrm{ov}[X_{n,1}, X_{n,3}] \to -\frac{15}{98}, \\
\frac{1}{n}\mathbb{C}\mathrm{ov}[D_{n,3}, D_{n,10}] &= \frac{1}{2n}\mathbb{C}\mathrm{ov}[X_{n,1}, X_{n,4}] \to \frac{55}{882}, \\
\frac{1}{n}\mathbb{C}\mathrm{ov}[D_{n,7}, D_{n,6}] &= \frac{1}{2n}\mathbb{C}\mathrm{ov}[X_{n,2}, X_{n,3}] \to \frac{55}{882}, \\
\frac{1}{n}\mathbb{C}\mathrm{ov}[D_{n,7}, D_{n,10}] &= \frac{1}{2n}\mathbb{C}\mathrm{ov}[X_{n,2}, X_{n,4}] \to -\frac{95}{882}, \\
\frac{1}{n}\mathbb{C}\mathrm{ov}[D_{n,6}, D_{n,10}] &= \frac{1}{n}\mathbb{C}\mathrm{ov}[X_{n,3}, X_{n,4}] \to -\frac{20}{441}.
\end{aligned}$$

By the symmetry of a covariance matrix, we can complete the top left 4×4 block of the full 6×6 limiting covariance matrix among all the degrees.

For the inactive nodes (of degrees 9 and 13), we use the construction in (8). Specialized to the binary network at hand, we have

$$\mathbf{Y}_n = (\mathbf{A}^T)^{-1} \left(\begin{pmatrix} X_{n,1} \\ X_{n,2} \\ X_{n,3} \\ X_{n,4} \end{pmatrix} - \begin{pmatrix} 6 \\ 2 \\ 0 \\ 0 \end{pmatrix} \right).$$

The limiting covariance of the vector on the left, at the scale of n^{-1}, follows:

$$\begin{aligned}
\frac{1}{n}\mathbb{C}\mathrm{ov}\left[\begin{pmatrix} Y_{n,1} \\ Y_{n,2} \\ Y_{n,3} \\ Y_{n,4} \end{pmatrix}\right] &\to (\mathbf{A}^T)^{-1} \lim_{n\to\infty} \left(\frac{1}{n}\mathbb{C}\mathrm{ov}[\mathbf{X}_n]\right) \left((\mathbf{A}^T)^{-1}\right)^T \\
&= \frac{5}{882} \begin{pmatrix} 24 & -16 & -3 & -5 \\ -16 & 20 & -5 & 1 \\ -3 & -5 & 10 & -2 \\ -5 & 1 & -2 & 6 \end{pmatrix}.
\end{aligned}$$

The bottom two components of the vector on the left-hand side are $Y_{n,3} = D_{n,9}$ and $Y_{n,4} = D_{n,13}$. So, the bottom 2×2 block in the last matrix are the bottom right 2×2 block of the full 6×6 limiting covariance matrix among all the degrees.

We determine the rest of the full 6×6 matrix from relation (7), which in this instance becomes

$$\begin{pmatrix} X_{n,1} \\ X_{n,2} \\ X_{n,3} \\ X_{n,4} \end{pmatrix} = \begin{pmatrix} 2Y_{n,1} + 4Y_{n,2} + 4Y_{n,3} + 4Y_{n,4} + 6 \\ 2Y_{n,1} + 2Y_{n,3} + 2Y_{n,4} + 2 \\ Y_{n,1} - Y_{n,3} \\ Y_{n,2} - Y_{n,4} \end{pmatrix}. \tag{12}$$

Extract the third component and write $Y_{n,3}$ on the left-hand side. Taking the variance, we get

$$\mathbb{V}\mathrm{ar}[X_{n,3}] + \mathbb{V}\mathrm{ar}[Y_{n,3}] + 2\,\mathbb{C}\mathrm{ov}[D_{n,6}, D_{n,9}] = \mathbb{V}\mathrm{ar}[Y_{n,1}].$$

After scaling, we get

$$\begin{aligned}\frac{1}{n}\mathrm{Cov}[D_{n,6}, D_{n,9}] &= \frac{1}{2n}\mathbb{V}\mathrm{ar}[Y_{n,1}] - \frac{1}{2n}\mathbb{V}\mathrm{ar}[Y_{n,3}] - \frac{1}{2n}\mathbb{V}\mathrm{ar}[X_{n,3}]\\ &\to \frac{1}{2}\Big(\frac{20}{147} - \frac{25}{441} - \frac{100}{441}\Big)\\ &= -\frac{65}{882}.\end{aligned}$$

The fourth component in (12) can be handled similarly to give the limit $\lim_{n\to\infty}\frac{1}{n}\mathrm{Cov}[D_{n,10}, D_{n,13}] = -\frac{25}{882}$. Another equation comes from the top component in (12). Upon reorganization, we write

$$X_{n,1} - 4Y_{n,4} = 2Y_{n,1} + 4Y_{n,2} + 4Y_{n,3} + 6.$$

Taking the variance and using $X_{n,1} = 2D_{n,3}$, $Y_{n,4} = D_{n,13}$, we can solve for $\lim_{n\to\infty}\frac{1}{n}\mathrm{Cov}[D_{n,3}, D_{n,13}] = \frac{25}{882}$.

With a different reorganization of the top components in (12), bringing $4Y_{n,3} = 4D_{n,9}$ to the left-hand side, we obtain the limit $\lim_{n\to\infty}\frac{1}{n}\mathrm{Cov}[D_{n,3}, D_{n,9}] = \frac{5}{294}$.

Coming from the second equation from the top in (12), the relation

$$X_{n,2} = 2Y_{n,1} + 2Y_{n,1} + 2Y_{n,1} + 2$$

can be handled in a similar manner in two steps, once with $2Y_{n,3}$ on the left, and a second time with $2Y_{n,4}$ on the left, to respectively produce the limits

$$\lim_{n\to\infty}\frac{1}{n}\mathrm{Cov}[D_{n,7}, D_{n,9}] = \frac{25}{882}, \qquad \lim_{n\to\infty}\frac{1}{n}\mathrm{Cov}[D_{n,7}, D_{n,13}] = -\frac{5}{882}.$$

No single equation taken from (12) gives us the remaining elements. We combine the third and the fourth in the form

$$X_{n,3} + X_{n,4} = Y_{n,1} + Y_{n,2} - Y_{n,3} - Y_{n,4}.$$

With $Y_{n,3}$ written on the left, we get $\lim_{n\to\infty}\frac{1}{n}\mathrm{Cov}[D_{n,9}, D_{n,10}] = -\frac{5}{294}$. The calculation needs the limiting value $\lim_{n\to\infty}\frac{1}{n}\mathrm{Cov}[D_{n,6}, D_{n,9}]$, and this has already been determined. A reorganization with $Y_{n,4}$ written on the left, we get $\lim_{n\to\infty}\frac{1}{n}\mathrm{Cov}[D_{n,6}, D_{n,13}] = -\frac{5}{294}$. The calculation needs the limiting value $\lim_{n\to\infty}\frac{1}{n}\mathrm{Cov}[D_{n,10}, D_{n,13}]$, and this has already been determined.

References

[1] Athreya, K. and Karlin, S. (1968). Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *The Annals of Mathematical Statistics* **39**, 1801–1817. https://doi.org/10.1214/aoms/1177698013.

[2] Bagchi, A. and Pal, A. (1985). Asymptotic normality in the generalized Pólya-Eggenberger urn model with applications to computer data structures. *SIAM Journal on Algebraic and Discrete Methods* **6**, 394–405. https://doi.org/10.1137/0606041.

[3] Bahrani, M. and Lumbroso, J. (2019). Split-decomposition trees with prime nodes: Enumeration and random generation of cactus graphs. *2018 Proceedings of the Fifteenth Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, 143–157. https://doi.org/10.1137/1.9781611975062.13.

[4] Bhutani, K., Kalpathy, R. and Mahmoud, H. (2021). Average measures in polymer graphs. *International Journal of Computer Mathematics: Computer Systems Theory* **6:1**, 37–53. https://doi.org/10.1080/23799927.2020.1860134.

[5] Bhutani, K., Kalpathy, R. and Mahmoud, H. (2022). Random networks grown by fusing edges via urns. *Network Science*, 1–14. https://doi.org/10.1017/nws.2022.30.

[6] Bhutani, K., Kalpathy, R. and Mahmoud, H. (2023). Random multi-hooking networks. *Probability in the Engineering and Informational Sciences*, 1–15. https://doi.org/10.1017/s0269964822000523.

[7] Bhutani, K., Kalpathy, R., Mahmoud, H. and Ofonedu, A. (2023+). Some empirical and theoretical attributes of random multi-hooking networks (under review).

[8] Brown, G. and Shubert, B. (1984). On random binary trees. *Mathematics of Operations Research* **9**, 43–65. https://doi.org/10.1287/moor.9.1.43.

[9] Chen, C. and Mahmoud, H. (2016). Degrees in random self-similar bipolar networks. *Journal of Applied Probability* **53**, 434–447. https://doi.org/10.1017/jpr.2016.11.

[10] Desmarais, C. and Holmgren, C. (2019). Degree distributions of generalized hooking networks. *2019 Proceedings of the Sixteenth Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, 103–110. https://doi.org/10.1137/1.9781611975505.11.

[11] Desmarais, C. and Holmgren, C. (2020). Normal limit laws for vertex degrees in randomly grown hooking networks and bipolar networks. *The Electronic Journal of Combinatorics* **27(2)**, P2.45. https://doi.org/10.37236/9139.

[12] Desmarais, C. and Mahmoud, H. (2021). Depths in hooking networks. *Probability in the Engineering and Informational Sciences*, 1–9. https://doi.org/10.1017/s0269964821000164.

[13] Drmota, M., Gittenberger, B. and Panholzer, A. (2008). The degree distribution of thickened trees. *DMTCS Proceedings, Fifth Colloquium on Mathematics and Computer Science* **AI**, 149–162. https://doi.org/10.46298/dmtcs.3561.

[14] Freedman, D. (1965). Bernard Friedman's urn. *The Annals of Mathematical Statistics* **36**, 956–970. https://doi.org/10.1214/aoms/1177700068.

[15] Gopaladesikan, M., Mahmoud, H. and Ward, M. (2014). Building random trees from blocks. *Probability in the Engineering and Informational Sciences* **28**, 67–81. https://doi.org/10.1017/s0269964813000338.

[16] Holmgren, C., Janson, S. and Sileikis, M. (2017). Multivariate normal limit laws for the numbers of fringe subtrees in m-ary search trees and preferential attachment trees. *The Electronic Journal of Combinatorics* **24**, p2.51. https://doi.org/10.37236/6374.

[17] Horn, R. and Johnson, C. (1985). *Matrix Analysis.* Cambridge University Press, Cambridge, UK.

[18] Janson, S. (2004). Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Processes and Their Applications* **110**, 177–245. https://doi.org/10.1016/j.spa.2003.12.002.

[19] Janson, S. (2020). Mean and variance of balanced Pólya urns. *Advances in Applied Probability* **52** , 1224–1248. https://doi.org/10.1017/apr.2020.38.

[20] Kalpathy, R. and Mahmoud, H. (2016). Degree profile of m-ary search trees: A vehicle for data structure compression. *Probability in the Engineering and Informational Sciences* **30**, 133–123. https://doi.org/10.1017/s0269964815000303.

[21] Knuth, D. (1998). *The Art of Computer Programming, Vol. 3: Sorting and Searching*, 2nd Ed. Addison-Wesley, Boston, Massachusetts.

[22] Mahmoud, H. (2008). *Pólya Urn Models.* Chapman-Hall, Orlando, Florida.

[23] Mahmoud, H. (2019). Local and global degree profiles of randomly grown self-similar hooking networks under uniform and preferential attachment. *Advances in Applied Mathematics* **111**, 101930. https://doi.org/10.1016/j.aam.2019.07.006.

[24] Mahmoud, H. (2019). A spectrum of series-parallel graphs with multiple edge evolution. *Probability in the Engineering and Informational Sciences* **33**, 487–499. https://doi.org/10.1017/s0269964818000505.

[25] Mahmoud, H. (2021). Covariances in Pólya urn schemes. *Probability in the Engineering and Informational Sciences* **37**, 60–71. https://doi.org/10.1017/s0269964821000450.

[26] Pouyanne, N. (2008). An algebraic approach to Pólya processes. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* **44**, 293–323. https://doi.org/10.1214/07-aihp130.

[27] Resnick, S. and Samorodnitsky, G. (2016). Asymptotic normality of degree counts in a preferential attachment model. *Advances in Applied Probability* **48**, 283–299. https://doi.org/10.1017/apr.2016.56.

[28] Smythe, R. (1996). Central limit theorems for urn models. *Stochastic Processes and Their Applications* **65**, 115–137. https://doi.org/10.1016/s0304-4149(96)00094-4.

[29] van der Hofstad, R. (2016). *Random Graphs and Complex Networks, Vol. 1.* Cambridge University Press, Cambridge, UK.