Degree profile of hierarchical lattice networks

Yarong Feng[1] **Hosam Mahmoud**[2]
Ludger Rüschendorf[3]

September 1, 2015

Abstract. We study the degree profile of random hierarchical lattice networks. At every step, each edge is either serialized (with probability p) or parallelized (with probability $1-p$). We establish an asymptotic Gaussian law for the number of nodes of outdegree 1, and show how to extend the derivations to encompass asymptotic limit laws for higher outdegrees. The asymptotic joint distribution of the number of nodes of outdegree 1 and 2 is shown to be bivariate normal. No phase transition with p is detected in these asymptotic laws.

For the limit laws, we use ideas from the contraction method. The recursive equations which we get involves coefficients and toll terms depending on the recursion variable and thus are not in the standard form of the contraction method. Yet, an adaptation of the contraction method goes through, showing that the method has promise for a wider range of random structures and algorithms.

AMS subject classifications. Primary: 05C82, 90B15, 05C21; Secondary: 60C05, 05C80, 60F05, 68Q87, 65Q30.

Keywords: Network, random graph, flow in graph, random structure, phase transition, recurrence, contraction method, metric, Wasserstein distance.

1 Introduction

Today we witness proliferation of all kinds of networks (social, hardwired, roadmaps, big organizations, etc.). There is a need to propose and analyze

[1]Department of Statistics, The George Washington University, Washington, D.C. 20052, U.S.A. E-mail: yfeng14@gwu.edu

[2]Department of Statistics, The George Washington University, Washington, D.C. 20052, U.S.A. E-mail: hosam@gwu.edu

[3]Department of Mathematical Stochastics, University of Freiburg, Eckerstraße 1, D-79104 Freiburg, Germany. E-mail: ruschen@stochastik.uni-freiburg.de

associated models. The class of series-parallel (SP) graphs is of particular relevance to bipolar networks, where there is a flow (of commodities like, say, commercial merchandise) from a source (producer) to a sink (market).

At the core of several SP definitions is the notion of a complete graph. We use the common notation K_n for the complete graph on n vertices. There are a few definitions of families of SP graphs. In one definition, an SP graph is an undirected connected graph that does not contain K_4 as a minor. By this definition, typical SP graphs start out from very small complete graphs, and as they grow, they forbid the existence of large complete graphs. Note that this definition does not distinguish two vertices as a source and a sink. One other popular recursive definition builds SP graphs inductively from smaller members of the family. In this variation, the smallest SP graph is K_2. The two vertices are called *Poles* (one *North* and one *South*). Larger SP graphs are built from smaller ones by one of two compositions: a *series composition*, which identifies the South Pole of a graph with the North Pole of the other, or a *parallel composition*, which identifies the two North Poles together, and the two South Poles together.

For the network flow application we have in mind, we take the second definition of SP graphs. Moreover, we think of SP graphs as directed, with orientation assigned to the edges to allow the flow to move from the North Pole to the South Pole. The *size* of an SP graph is the number of edges in it.

Several models of randomness have been proposed for SP graphs. They include the uniform model, where all SP networks of a certain size are equally likely [1, 4], the hierarchical lattice model [7], where at each stage of the growth every edge is either serialized or parallelized, and the incremental model introduced in [9], in which one randomly chosen edge at a time is serialized or parallelized. The reference [10] presents a variation with a binary degree restriction on the nodes of the SP graph.

Our aim in this article is to study the profile of node degrees in the hierarchical lattice model introduced in [7]. In this model, at each discrete time step, *every* directed edge in the graph experiences an evolution. At a stage, an edge is *serialized* with probability p, or *parallelized* with probability $q := 1 - p$. The graph evolves in the following manner. Suppose uv is a directed edge from vertex u to vertex v that exists at time step $n - 1$. To serialize uv (with probability p), we replace the edge with two directed edges ux (directed from u to x), and xv (directed from x to v). Thus, a new vertex, x, appears. To parallelize uv (with probability q), we create a new edge, also directed from u to v. We shall call p the *index* of the hierarchical lattice. It is also natural to think of the number of steps of evolution as the *age* of the network.

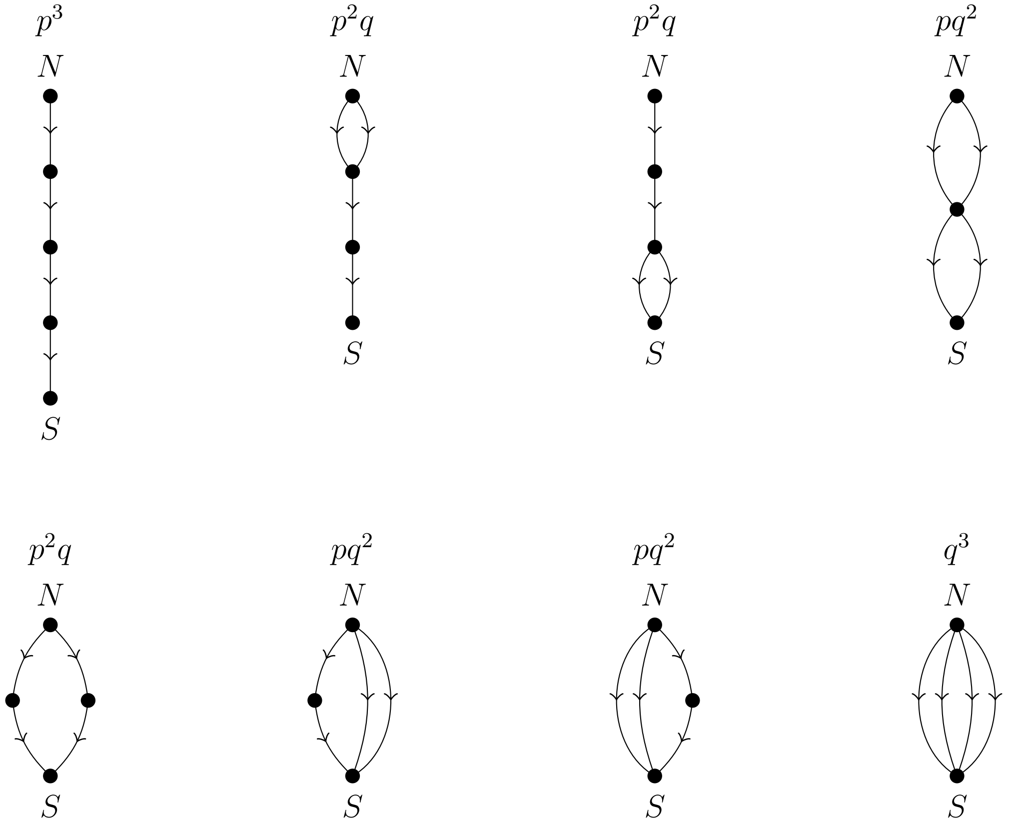

Figure 1: The eight hierarchical lattice graphs of size 4, and their probabilities.

Figure 1 shows all eight hierarchical lattice SP graphs that can arise in two steps of evolution and their probabilities. Note that orientation in the plane is part of the definition of a hierarchical lattice SP graph (network): A vertex with outdegree k has k children, distinguished as leftmost, second from the left, third from the left, and so on till the kth from the left (i.e., the rightmost). For instance, the second and third SP graphs in the second row of Figure 1 are isomorphic graphs, but are considered as different hierarchical lattice networks.

The hierarchical lattice graph model is quite interesting, as it may depend critically on p. If p is too small, most of the operations of edge evolution tend to parallelize, and one would expect a short fat graph extending between N and S. At the other end of the spectrum, if p is too large, most of the operations of edge evolution tend to serialize, and one would expect a long scrawny graph between N and S. Are there interesting phase transitions in between? The authors of [7] report that the effective resistance across the graphs, first-passage percolation on the graphs and the Cheeger constant of the graphs all exhibit a drastic change of behavior at $p = \frac{1}{2}$. An instance of these phase transitions is manifested in the effective resistance, R_n, between the poles. It converges to 0 almost surely, for $p < \frac{1}{2}$, and diverges to ∞ almost surely, for $p > \frac{1}{2}$. Whereas at $p = \frac{1}{2}$, $\frac{1}{n} \ln R_n$ converges to 0, but the scale $\frac{1}{n}$ is not enough to bring $\ln R_n$ down to 0, when $p > \frac{1}{2}$. By contrast, some phenomena, such as the order (number of nodes) *do not* exhibit a phase change.

2 Scope

We study the distribution of the number of nodes of a given outdegree by the contraction method. We shall see Gaussian limits adding to the list of properties that do not exhibit phase transition. A motivation for studying the outdegrees of nodes is that they are an important factor in the functionality of the network. The failure of a node with high outdegree can paralyze large components in the network and disrupt the flow along many paths.

The structure of the rest of the manuscript is in sections. Section 3 is on the notation used throughout. In Section 4, we set up a hierarchical recurrence system among the number of nodes of various outdegrees. At the bottom of this inductive construction is the number of nodes of outdegree 1 (the smallest outdegree that can appear in the network), and we need to analyze this separately. This is taken up thoroughly in Section 5, which is divided into two subsections dealing with two main themes: concentration laws in Subsection 5.1 and Gaussian laws in Subsection 5.2. The method used for the asymptotic distribution in Subsection 5.2 is an adaptation of the contraction method. The same methods are extended in Section 6 to nodes of higher outdegrees. In Section 7, we conclude with some remarks. Some technical details are relegated to two appendices.

3 Notation

We use the notation $X_n^{(k)}$ for the number of nodes of outdegree k in a hierarchical lattice graph at age n. In what follows, $\text{Bin}(m, p)$ denotes a binomial random variable that counts the number of successes in m independent identically distributed experiments, with rate of success p per experiment. The notation Uniform $(0, 1)$ will be used for a random variable that is uniformly distributed on the interval (0,1). Also, $\mathcal{N}(0, \sigma^2)$ will stand for a centered normally distributed random variable with variance σ^2. The notation $\mathbb{I}_{\mathcal{E}}$ will be used for the indicator of the event $\mathcal{E}$, that is, a function that assumes the value 1, if $\mathcal{E}$ occurs, and assumes the value 0, otherwise.

We shall have need for the compounded random variable $\text{Bin}(S, p)$, where S is random (the distribution of such random variable is sometimes called a hierarchical model or a mixture model, and the associated measure is sometimes called a random measure). Such a random variable is generated by first obtaining a value for S from a prior distribution, then using this value as the number of experiments in the binomial distribution. These hierarchical models are core in Bayesian statistics.

In particular, we shall need the compounded binomial random variables $\text{Bin}(X_n^{(k)}, p)$. Technically speaking, such a binomial random variable should be represented as a sum of $X_n^{(k)}$ independent identically distributed random variables (indicators), where each indicator is described by a Uniform $(0, 1)$ random variable U, and for each experiment the U is independent of all else. The uniform variable is defined on the same probability space as all the variables of the network.We shall succinctly write $\text{Bin}(X_n^{(k)}, p)$ to actually mean $\sum_{i=1}^{X_n^{(k)}} \mathbb{I}_{\{U_i<p\}}$, for U_i's being $X_n^{(k)}$ independent identically distributed Uniform (0,1) random variables.We shall alternate in our choice of notation—the notation $\text{Bin}(X_n^{(k)}, p)$ is more compact than the summation form, and will be used when adequate for simpler computations, like those for mean values. However, the need arises in variance and higher moments computations for a notation that captures subtle dependencies, and in these situations we use the full summation form.We can take the space on which sequences of random variables and their limits are defined to be Skorohod's.

From probability theory we use standard convergence notation: The symbols $\xrightarrow{\text{a.s.}}$, $\xrightarrow{\text{P}}$, and $\xrightarrow{\text{D}}$ are respectively for convergence almost surely, in probability and in distribution, whereas $\overset{D}{=}$ is for exact equality in distribution.

4 A hierarchical recurrence system for the node degrees

We develop in this section a system of stochastic recurrence equations for $X_n^{(k)}$, for $k = 1, 2, \dots$. It will turn out that $X_n^{(k)}$ is related to $X_{n-1}^{(j)}$, for all $j \le k$. Thus, we are creating a hierarchical recurrence system in which $X_n^{(k)}$ depends on the number of nodes of the same outdegree or smaller, and the entire history of the evolution. At the basis of this inductive system, we need a representation for $X_n^{(1)}$. We set up this hierarchical recurrence system in two technical lemmata.

Lemma 1. *Let $X_n^{(1)}$ be the number of nodes of outdegree 1 in a hierarchical lattice network with index p at age n. We then have*

$$X_n^{(1)} \overset{D}{=} 2\mathrm{Bin}^*\big(X_{n-1}^{(1)}, p\big) + \mathrm{Bin}^{**}\big(2^{n-1} - X_{n-1}^{(1)}, p\big), \qquad (1)$$

with boundary conditions $X_0^{(1)} = 1$. Here $\mathrm{Bin}^*\big(X_{n-1}^{(1)}, p\big)$ and $\mathrm{Bin}^{**}\big(2^{n-1} - X_{n-1}^{(1)}, p\big)$ are conditionally independent given $X_{n-1}^{(1)}$.

Proof. Let us use the notation $\mathcal{E}_n^{(\ge 2)}$ to mean the number of edges emanating out of nodes of outdegree at least 2 when the network is at age n. Note that the number of edges doubles after each step of evolution—there are 2^n edges, when the network is at age n. Thus, we have

$$\mathcal{E}_n^{(\ge 2)} = 2^n - X_n^{(1)}. \qquad (2)$$

At age n, a node of outdegree 1 could appear in two ways. The first way is when the edge out of a node of outdegree 1 in the previous step is serialized, it adds a new edge with a northern node of outdegree 1. Since the original node of outdegree 1 is preserved as well, there are $2\mathrm{Bin}^*(X_{n-1}^{(1)}, p)$ contributions to nodes of outdegree 1. The second way is when an edge out of a node of outdegree higher than 2 is serialized, it also adds a new node of outdegree 1. We have $\mathrm{Bin}^{**}(\mathcal{E}_n^{(\ge 2)}, p)$ contributions in this way. Using (2), we get the desired result. □

Lemma 2. *Let $X_n^{(k)}$ be the number of nodes of outdegree $k \ge 2$ in a hierarchical lattice network with index p at age n. We then have*

$$X_n^{(k)} \overset{D}{=} \mathrm{Bin}_k\big(X_{n-1}^{(k)}, p^k\big) + \sum_{i=\lceil \frac{k}{2} \rceil}^{k-1} \mathrm{Bin}_i\left(X_{n-1}^{(i)}, \binom{i}{k-i} p^{2i-k} q^{k-i}\right),$$

with boundary conditions $X_n^{(k)} = 0$. *Here, for* $i \neq j$, *the random variables* $\text{Bin}_i\big(X_{n-1}^{(i)}, p\big)$ *and* $\text{Bin}_j\big(X_{n-1}^{(j)}, p\big)$ *are conditionally independent given* $X_{n-1}^{(i)}$ *and* $X_{n-1}^{(j)}$.

Proof. At age n, a node of outdegree k could appear in two ways. The first way is when every edge out of a node of outdgree k in the previous step is serialized (an event that occurs with probability p^k); the original node of outdegree k is preserved. There are a total of $\text{Bin}_k\big(X_{n-1}^{(k)}, p^k\big)$ contributions.

There can be contributions to $X_n^{(k)}$ from edges out of nodes of outdegree $i < k$, too. If a proper number of such edges is parallelized, they can increase the outdegree at their common northern vertex to k. More precisely, suppose we have a node of outdegre i at age $n - 1$. None, some or all of these edges can be parallelized. If $i < \lceil \frac{k}{2} \rceil$, no matter how the edges out of this node evolve, they will not induce a change in the number of nodes of outdgree k—even if all of them get parallelized, the outdegree of the northern vertex of these edges doubles to $2i < k$. Only edges out of nodes of outdegree $i = \lceil \frac{k}{2} \rceil, \dots, k - 1$ can evolve in proper combinations to increase the number of nodes of outdegree k at age n. If a node has outdegree i, for $\lceil \frac{k}{2} \rceil \le i \le k-1$, $2i - k$ among these edges can be serialized and $k - i$ can be parallelized. The occurrence of this event increases the number of nodes of outdegree k by 1. The $k - i$ edges to be parallelized can be chosen in $\binom{i}{k-i}$ ways. □

5 Probabilistic analysis of the number of nodes of outdegree 1

The hierarchical system of stochastic recurrences in Lemmata 1 and 2 is instrumental in conducting a probabilistic analysis, giving averages and beyond. We start from the bottom of the hierarchy, i.e., from nodes of outdegree 1.

Proposition 1. *The average and variance of* $X_n^{(1)}$, *the number of nodes of outdegree 1 in a hierarchical lattice network with index* p *at age* n, *are given by*

$$\mathbb{E}[X_n^{(1)}] = \frac{p}{2-p}\, 2^n - \frac{p^{n+1}}{2-p} + p^n \sim \frac{p}{2-p}\, 2^n,$$

and

$$\begin{aligned}\mathbb{V}\mathrm{ar}[X_n^{(1)}] &= \frac{2pq(p+1)2^n}{(2-p)(2-p^2)} + \frac{6p^nq}{(2-p)} - \frac{2(2p+3)qp^{2n}}{(2-p^2)} \\ &\sim \frac{2pq(p+1)}{(2-p)(2-p^2)}\, 2^n.\end{aligned}$$

Proof. Conditioning the stochastic recurrence in Lemma 1 on $X_{n-1}^{(1)}$, then taking a double expectation, we obtain

$$\mathbb{E}\big[X_n^{(1)}\big] = p\,\mathbb{E}\big[X_{n-1}^{(1)}\big] + 2^{n-1}p.$$

This recurrence is amenable to direct iteration, and with the boundary condition in Lemma 1, the stated result follows.

The variance follows from a similar recursive formulation, so we only outline it. Square both sides of the stochastic recurrence in Lemma 1, condition on $X_{n-1}^{(1)}$, and take expectation, to get

$$\begin{aligned}\mathbb{E}\big[(X_n^{(1)})^2 \,|\, X_{n-1}^{(1)}\big] &= 4\,\mathbb{E}\Big[\big(\mathrm{Bin}^*\big(X_{n-1}^{(1)}, p\big)\big)^2 \,|\, X_{n-1}^{(1)}\Big] \\ &\quad + \mathbb{E}\Big[\big(\mathrm{Bin}^{**}\big(2^{n-1} - X_{n-1}^{(1)}, p\big)\big)^2 \,|\, X_{n-1}^{(1)}\Big] \\ &\quad + 4\,\mathbb{E}\big[\mathrm{Bin}^*\big(X_{n-1}^{(1)}, p\big) \,|\, X_{n-1}^{(1)}\big] \\ &\qquad \times \mathbb{E}\big[\mathrm{Bin}^{**}\big(2^{n-1} - X_{n-1}^{(1)}, p\big) \,|\, X_{n-1}^{(1)}\big].\end{aligned}$$

Note that in the cross-product term we use conditional independence. Using well known facts about the binomial random variable, then taking a double expectation, we get a recurrence for the unconditional expectation $\mathbb{E}\big[(X_n^{(1)})^2\big]$, the solution of which (under the boundary condition given in Lemma 1) is obtained. The variance follows by subtracting off the square of established mean. □

5.1 Concentration laws

The relatively small variance gives us concentration laws.

Theorem 1. *Let $X_n^{(1)}$ be the number of nodes of outdegree 1 in a hierarchical lattice network at age n. Then, we have*

$$\frac{X_n^{(1)}}{2^n} \xrightarrow{\text{a.s.}} \frac{p}{2-p} =: c_p^{(1)}, \tag{3}$$

and further,

$$\mathbb{E}\big|X_n^{(1)} - 2^n c_p^{(1)}\big| = O\big(2^{n/2}\big).$$

Proof. By Chebyshev's inequality, for any fixed $\varepsilon > 0$, we write

$$\mathbb{P}\Big(\big|X_n^{(1)} - \mathbb{E}\big[X_n^{(1)}\big]\big| > \varepsilon\Big) \le \frac{\mathbb{V}\mathrm{ar}[X_n^{(1)}]}{\varepsilon^2}.$$

Replace ε by $\varepsilon\mathbb{E}\big[X_n^{(1)}\big]$ to get

$$\begin{aligned}
\mathbb{P}\left(\left|\frac{X_n^{(1)}}{\mathbb{E}\big[X_n^{(1)}\big]} - 1\right| > \varepsilon\right) \quad &\le \quad \frac{\dfrac{2pq(p+1)2^n}{(2-p)(2-p^2)} + \dfrac{6p^n q}{(2-p)} - \dfrac{2(2p+3)qp^{2n}}{(2-p^2)}}{\varepsilon^2\Big(\dfrac{2^n p}{2-p} - \dfrac{p^{n+1}}{p-2} + p^n\Big)^2} \\
&= \quad O\Big(\frac{1}{2^n}\Big).
\end{aligned}$$

This fast rate of decline in the probabilities renders the series of the probabilities summable:

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\left|\frac{X_n^{(1)}}{\mathbb{E}\big[X_n^{(1)}\big]} - 1\right| \ge \varepsilon\right) < \infty.$$

According to the Borel-Cantelli Lemma, we have

$$\frac{X_n^{(1)}}{\mathbb{E}[X_n^{(1)}]} \xrightarrow{\text{a.s.}} 1.$$

Applying Proposition 1, we get the first result.

From the asymptotics of the mean and variance, as given in Proposition 1, we have

$$\begin{aligned}
\mathbb{E}\big[\big(X_n^{(1)} - 2^n c_p^{(1)}\big)^2\big] &= \mathbb{E}\big[\big(\big(X_n^{(1)} - \mathbb{E}\big[X_n^{(1)}\big]\big) + \big(\mathbb{E}\big[X_n^{(1)}\big] - 2^n c_p^{(1)}\big)\big)^2\big] \\
&= \mathbb{V}\mathrm{ar}\big[X_n^{(1)}\big] + \big(\mathbb{E}\big[X_n^{(1)}\big] - 2^n c_p^{(1)}\big)^2 \\
&= O(2^n).
\end{aligned}$$

By Jensen's inequality we have

$$\mathbb{E}|X_n^{(1)} - 2^n c_p^{(1)}| \le \sqrt{\mathbb{E}\big[\big(X_n^{(1)} - 2^n c_p^{(1)}\big)^2\big]} = O\big(2^{n/2}\big).$$

□

5.2 An asymptotic Gaussian distribution

Higher moments are hard to compute by direct recurrence. Each one is more complex and is given by longer formulæ than the one before. We need a shortcut for the computation of all asymptotic moments, i.e., the asymptotic distribution. A tool suitable for this task is the contraction method. For basics of the method and several useful variations see [3, 11, 12, 13, 14, 15].

Consider random variables obtained by suitable normalization (shifting and scaling) of the variables in Lemma 1. It is natural to normalize by exact centering and use the exact standard deviation as a scale factor. Thanks to Slutsky's theorem [8]; Pages 146–147, we get the same results, if we use an asymptotic equivalent of the mean to shift, and the correct order of the standard deviation to scale; let us introduce the normalized random variable

$$Y_n^{(1)} := \frac{X_n^{(1)} - 2^n c_p^{(1)}}{\sqrt{2^n}}.$$

We normalize the representation in Lemma 1 by writing it in the form

$$\begin{aligned}
Y_n^{(1)} &= \frac{X_n^{(1)} - 2^n c_p^{(1)}}{\sqrt{2^n}} \\
&\stackrel{\mathcal{D}}{=} 2\,\frac{\mathrm{Bin}^*\big(X_{n-1}^{(1)}, p\big) - pX_{n-1}^{(1)}}{\sqrt{pqX_{n-1}^{(1)}}} \times \sqrt{\frac{pqX_{n-1}^{(1)}}{2^n}} \\
&\qquad + \frac{\mathrm{Bin}^{**}\big(2^{n-1} - X_{n-1}^{(1)}, p\big) - p\big(2^{n-1} - X_{n-1}^{(1)}\big)}{\sqrt{pq\big(2^{n-1} - X_{n-1}^{(1)}\big)}} \\
&\qquad\qquad \times \sqrt{\frac{pq\big(2^{n-1} - X_{n-1}^{(1)}\big)}{2^n}} \\
&\qquad + \frac{2pX_{n-1}^{(1)} + p\big(2^{n-1} - X_{n-1}^{(1)}\big) - 2^n c_p^{(1)}}{\sqrt{2^n}} \\
&\stackrel{\mathcal{D}}{=}: 2Z_nR_n + \tilde{Z}_n\tilde{R}_n + \frac{pY_{n-1}^{(1)}}{\sqrt{2}},
\end{aligned} \tag{4}$$

where

$$Z_n = \frac{\mathrm{Bin}^*\big(X_{n-1}^{(1)}, p\big) - pX_{n-1}^{(1)}}{\sqrt{pqX_{n-1}^{(1)}}},$$

$$\tilde{Z}_n = \frac{\mathrm{Bin}^{**}\big(2^{n-1} - X_{n-1}^{(1)}, p\big) - p\big(2^{n-1} - X_{n-1}^{(1)}\big)}{\sqrt{pq\big(2^{n-1} - X_{n-1}^{(1)}\big)}},$$

$$R_n = \sqrt{\frac{pqX_{n-1}^{(1)}}{2^n}}, \qquad \tilde{R}_n = \sqrt{\frac{pq\big(2^{n-1} - X_{n-1}^{(1)}\big)}{2^n}}.$$

Note that Z_n and $\tilde{Z}_n$ are conditionally independent given $X_{n-1}^{(1)}$.

The distributional equation (4) does not fit in the usual framework considered in the literature on the contraction method, as it is not in the form of iterative subproblems and a toll function that does not depend on any of the parts. If we consider $Y_{n-1}^{(1)}$ to be the iterative part, the remaining variables are a toll dependent on $Y_{n-1}^{(1)}$. Nonetheless, the equation has the general spirit of the distributional equations in the contraction method for recursive algorithms, and can be handled by similar ideas, as it is asymptotically of the usual form of the contraction method.

We shall prove that $Y_n^{(1)}$ converges to a normal limit $Y^{(1)}$ in distribution. We shall do this in four steps:

Step 1: We *guess* the distributional equation satisfied by the limit $Y^{(1)}$.
Step 2: We prove that the distributional limiting equation satisfied by $Y^{(1)}$ has a *unique* solution.
Step 3: We *prove* that $Y_n^{(1)}$ converges to $Y^{(1)}$.
Step 4: The unique solution is shown to be the *normal distribution.*

Step 1

By Theorem 1, we have the convergence relations

$$R_n \xrightarrow{\text{a.s.}} \sqrt{\frac{pqc_p^{(1)}}{2}} =: R, \qquad \tilde{R}_n \xrightarrow{\text{a.s.}} \sqrt{\frac{pq(1 - c_p^{(1)})}{2}} =: \tilde{R}.$$

We also have the two normalized binomial random variables converging to normal limits:

$$Z_n = \frac{\text{Bin}^*(X_{n-1}^{(1)}, p) - pX_{n-1}^{(1)}}{\sqrt{pqX_{n-1}^{(1)}}} \xrightarrow{\text{D}} Z,$$

$$\tilde{Z}_n = \frac{\text{Bin}^{**}(2^{n-1} - X_{n-1}^{(1)}, p) - p(2^{n-1} - X_{n-1})}{\sqrt{pq(2^{n-1} - X_{n-1}^{(1)})}} \xrightarrow{\text{D}} \tilde{Z},$$

where Z and $\tilde{Z}$ are standard normal random variates. It is plausible to surmise that representation (4) induces the following limiting distributional equation

$$Y^{(1)} \overset{\mathcal{D}}{=} \frac{p}{\sqrt{2}} Y^{(1)} + 2ZR + \tilde{Z}\tilde{R}, \tag{5}$$

where Z, $\tilde{Z}$, and $Y^{(1)}$ are mutually indepdendent. The presumed independence comes from the fact that Z_n and $\tilde{Z}_n$ are conditionally independent. Moreover, one sees, by the central limit theorem and using that $X_{n-1}^{(1)} \xrightarrow{\text{a.s.}} \infty$: that

$$\begin{aligned}
\mathbb{P}\big(Z_n \le z, \tilde{Z}_n \le \tilde{z} \,|\, Y_{n-1}^{(1)}\big) &= \mathbb{P}\big(Z_n \le z, \tilde{Z}_n \le \tilde{z} \,|\, X_{n-1}^{(1)}\big) \\
&= \mathbb{P}\big(Z_n \le z \,|\, X_{n-1}^{(1)}\big)\, \mathbb{P}\big(\tilde{Z}_n \le \tilde{z} \,|\, X_{n-1}^{(1)}\big) \\
&\xrightarrow{\text{a.s.}} \mathbb{P}\big(Z \le z\big)\, \mathbb{P}\big(\tilde{Z} \le \tilde{z}\big) \\
&= \mathbb{P}\big(Z \le z, \tilde{Z} \le \tilde{z} \,|\, Y^{(1)}\big).
\end{aligned}$$

Consequently, we would have $(Z_n, \tilde{Z}_n, Y_n^{(1)}) \xrightarrow{\text{D}} (Z, \tilde{Z}, Y^{(1)})$, if we established $Y_n^{(1)} \xrightarrow{\text{D}} Y^{(1)}$.

The methodology in the steps that follow is based on establishing convergence w.r.t. the Wasserstein distance. The *Wasserstein distance of order p* (also called the minimal ℓ_p distance) between two distribution functions F and G is defined by

$$\ell_p(F, G) = \inf ||W - Z||_p,$$

where the infimum is taken over all random variables W and Z having the respective distribution functions F and G (with $||\,.\,||_p$ being the usual $\mathcal{L}_p$ norm). If F_n is a sequence of distribution functions, it is known [2] that convergence in the first-order Wasserstein distance implies, and in fact is equivalent to, weak convergence, as well as convergence of the first moment.

Step 2

In what follows we denote $p/\sqrt{2}$ by h and let $\mathcal{F}_1$ be the space of all distribution functions with finite first moment. View the right-hand side of (5) as a mapping from $\mathcal{F}_1$ into itself.

Let Y and Y' be two random variables with distribution functions $F_Y \in \mathcal{F}_1$, and $F_{Y'} \in \mathcal{F}_1$, such that (Y, Y') is independent of $((Z, R), (\tilde{Z}, \tilde{R}))$. Define the transformation $\mathcal{T} : \mathcal{F}_1 \to \mathcal{F}_1$ by $\mathcal{T}(F_Y)$ being the law of $hY + 2ZR + \tilde{Z}\tilde{R} =: hY + S$. We then have

$$\ell_1\big(\mathcal{T}(F_Y), \mathcal{T}(F_{Y'})\big) \le \mathbb{E}\big[\big|\big(hY + S\big) - \big(hY' + S\big)\big|\big] = h\,\mathbb{E}\big[|Y - Y'|\big].$$

Taking the infimum over (Y, Y') that is a coupling of $(F_Y, F_{Y'})$, we find that

$$\ell_1\big(\mathcal{T}(F_Y), \mathcal{T}(F_{Y'})\big) \le h\,\ell_1(F_Y, F_{Y'}) < \ell_1(F_Y, F_{Y'}).$$

Thus, the mapping $\mathcal{T}$ is contracting. As the mapping is a contraction in a complete metric space, there is a unique fixed-point in $\mathcal{F}_1$ satisfying (5), by Banach's fixed-point theorem [5].

An alternative proof of the uniqueness of the solution to the limiting distributional equation (5) can be found in [16], Theorem 1.5 and 1.6.

Step 3

We need some work to formally justify the limit equation guessed in Step 1. This is done by "coupling" the random variables on the same probability space, and showing that the first-order Wasserstein distance between the distributions of $Y_n^{(1)}$ and $Y^{(1)}$ converges to 0. Let F_n be the distribution function of $Y_n^{(1)}$, and F be the distribution function of $Y^{(1)}$.

Lemma 3. *(The coupling lemma). There exists a coupling (Y_n, Y) of (F_n, F), such that $b_n^{(1)} := \ell_1(F_n, F) = \mathbb{E}|Y_n - Y| \to 0$. In particular, this implies*

$$Y_n^{(1)} \xrightarrow{\text{D}} Y^{(1)}.$$

Proof. We shall show that the first-order Wasserstein distance between F_n and F converges to 0. Let (Y_n, Y) be an optimal coupling of (F_n, F), for all $n \ge 0$. Define

$$\hat{X}_n^{(1)} = \sqrt{2^n}\,Y_n + 2^n c_p^{(1)}.$$

Then, $\hat{X}_n^{(1)}$ is defined on the space of the coupling and has the distribution of $X_n^{(1)}$, and Y_n is a normalized binomial in the form

$$Y_n = \frac{\hat{X}_n^{(1)} - 2^n c_p^{(1)}}{\sqrt{2^n}}.$$

We can then define the associated random variables

$$\hat{Z}_n = \frac{\mathrm{Bin}^*\big(\hat{X}_{n-1}^{(1)}, p\big) - p\hat{X}_{n-1}^{(1)}}{\sqrt{pq\hat{X}_{n-1}^{(1)}}},$$

$$\hat{\tilde{Z}}_n = \frac{\mathrm{Bin}^{**}\big(2^{n-1} - \hat{X}_{n-1}^{(1)}, p\big) - p\big(2^{n-1} - \hat{X}_{n-1}^{(1)}\big)}{\sqrt{pq\big(2^{n-1} - \hat{X}_{n-1}^{(1)}\big)}},$$

$$\hat{R}_n = \sqrt{\frac{pq\hat{X}_{n-1}^{(1)}}{2^n}}, \qquad \hat{\tilde{R}}_n = \sqrt{\frac{pq\big(2^{n-1} - \hat{X}_{n-1}^{(1)}\big)}{2^n}}.$$

We call the limits of these variables $\hat{Z}$, $\hat{\tilde{Z}}$, $\hat{R}$, and $\hat{\tilde{R}}$ respectively. As shown in Step 1, $(\hat{Z}_n, \hat{\tilde{Z}}_n) \xrightarrow{\mathrm{D}} (\hat{Z}, \hat{\tilde{Z}})$, where $\hat{Z}$ and $\hat{\tilde{Z}}$ are independent standard normal random variates. Furthermore, $\hat{R}_n \xrightarrow{\mathrm{D}} \hat{R}$ and $\hat{\tilde{R}}_n \xrightarrow{\mathrm{D}} \hat{\tilde{R}}$ (the hatted limits are the same as R and $\tilde{R}$, which are actually constants and these convergence relations take place in probability, as well).

By this construction, $hY_{n-1} + 2\hat{Z}_n\hat{R}_n + \hat{\tilde{Z}}_n\hat{\tilde{R}}_n$ has the same distribution as Y_n, and $hY + 2\hat{Z}\hat{R} + \hat{\tilde{Z}}\hat{\tilde{R}}$ has the same distribution as Y. Thus, they are a coupling of (F_n, F) and as a consequence we get:

$$b_n^{(1)} = \ell_1(F_n, F) \le \mathbb{E}\big|(hY_{n-1} + 2\hat{Z}_n\hat{R}_n + \hat{\tilde{Z}}_n\hat{\tilde{R}}_n) - (hY + 2\hat{Z}\hat{R} + \hat{\tilde{Z}}\hat{\tilde{R}})\big|.$$

We shall show that $b_n^{(1)} \to 0$; subsequently, we have $Y_n \xrightarrow{\mathrm{D}} Y$.

Recall that $h = p/\sqrt{2}$. Note that $0 < h < 1$. By the triangle inequality, we get

$$\begin{aligned}
b_n^{(1)} &\le h\mathbb{E}|Y_{n-1} - Y| + 2\,\mathbb{E}|\hat{Z}_n\hat{R}_n - \hat{Z}\hat{R}| + \mathbb{E}|\hat{\tilde{Z}}_n\hat{\tilde{R}}_n - \hat{\tilde{Z}}\hat{\tilde{R}}| \\
&= h\,\mathbb{E}|Y_{n-1} - Y| + 2\,\mathbb{E}|\hat{Z}_n\hat{R}_n + \hat{R}_n\hat{Z} - \hat{R}_n\hat{Z} - \hat{Z}\hat{R}| \\
&\qquad + \mathbb{E}|\hat{\tilde{Z}}_n\hat{\tilde{R}}_n + \hat{\tilde{R}}_n\hat{\tilde{Z}} - \hat{\tilde{R}}_n\hat{\tilde{Z}} - \hat{\tilde{Z}}\hat{\tilde{R}}| \\
&\le h\,b_{n-1}^{(1)} + 2\,\mathbb{E}|\hat{R}_n(\hat{Z}_n - \hat{Z})| + 2\,\mathbb{E}|\hat{Z}(\hat{R}_n - \hat{R})| + \mathbb{E}|\hat{\tilde{R}}_n(\hat{\tilde{Z}}_n - \hat{\tilde{Z}})| \\
&\qquad + \mathbb{E}|\hat{\tilde{Z}}(\hat{\tilde{R}}_n - \hat{\tilde{R}})|.
\end{aligned}$$

We can bound all the four nonrecursive terms by the Cauchy-Schwarz inequality in the following way:

$$\begin{aligned}
\mathbb{E}|\hat{Z}(\hat{R}_n - \hat{R})| &\le \sqrt{\mathbb{E}[\hat{Z}^2]\,\mathbb{E}[(\hat{R}_n - \hat{R})^2]} \\
&= \sqrt{\mathbb{E}[(\hat{R}_n - \hat{R})^2]} \\
&= \left(\mathbb{E}\left[\left(\sqrt{\frac{pq\hat{X}^{(1)}_{n-1}}{2^n}} - \sqrt{\frac{pqc^{(1)}_p}{2}}\right)^2\right]\right)^{1/2} \\
&= \left(\mathbb{E}\left[\left(\frac{pq\hat{X}^{(1)}_{n-1}}{2^n} + \frac{pqc^{(1)}_p}{2} - 2pq\sqrt{\frac{c^{(1)}_p\hat{X}^{(1)}_{n-1}}{2^{n+1}}}\right)\right]\right)^{1/2}.
\end{aligned}$$

Recall that $\hat{X}^{(1)}_n/2^n \le 1$. By the continuous mapping theorem, dominated convergence and Theorem 1, the limit of the latter is

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}|\hat{Z}(\hat{R}_n - \hat{R})| &\le \left(\mathbb{E}\left[\lim_{n\to\infty}\left(\frac{pq\hat{X}^{(1)}_{n-1}}{2^n} + \frac{pqc^{(1)}_p}{2} - 2pq\sqrt{\frac{c^{(1)}_p\hat{X}^{(1)}_{n-1}}{2^{n+1}}}\right)\right]\right)^{1/2} \\
&= \left(\mathbb{E}\left[\left(\frac{pqc^{(1)}_p}{2} + \frac{pqc^{(1)}_p}{2} - 2pq\sqrt{\frac{c^{(1)}_p \times c^{(1)}_p}{4}}\right)\right]\right)^{1/2} \\
&= 0,
\end{aligned}$$

i.e. $\lim_{n\to\infty} \mathbb{E}|\hat{Z}(\hat{R}_n - \hat{R})| = 0$. We also have

$$\lim_{n\to\infty} \mathbb{E}|\hat{\hat{Z}}(\hat{\hat{R}}_n - \hat{\hat{R}})| = 0.$$

We similarly have

$$\mathbb{E}|\hat{R}_n(\hat{Z}_n - \hat{Z})| \le \mathbb{E}|\hat{Z}_n - \hat{Z}| = o(1),$$

$$\mathbb{E}|\hat{\hat{R}}_n(\hat{\hat{Z}}_n - \hat{\hat{Z}})| \le \mathbb{E}|\hat{\hat{Z}}_n - \hat{\hat{Z}}| = o(1).$$

Combining the bounds, we see that

$$b^{(1)}_n \le hb^{(1)}_{n-1} + o(1).$$

It follows that, for any fixed $\varepsilon > 0$, there exists a positive index $n_0(\varepsilon)$, such that

$$b_n^{(1)} \le h b_{n-1}^{(1)} + \varepsilon, \qquad \text{for } n \ge n_0(\varepsilon).$$

For large n, we can proceed iteratively:

$$\begin{aligned} b_n^{(1)} &\le \varepsilon + h b_{n-1}^{(1)} \\ &\le \varepsilon + \varepsilon h + h^2 b_{n-2}^{(1)} \\ &\;\vdots \\ &\le \varepsilon(1 + h + h^2 + \cdots + h^{n-n_0-1}) + h^{n-n_0} b_{n_0}^{(1)} \\ &\le \varepsilon(1 + h + h^2 + \cdots) + h^{n-n_0} b_{n_0}^{(1)} \\ &= \frac{\varepsilon}{1-h} + h^{n-n_0} b_{n_0}^{(1)} \\ &\to \frac{\varepsilon}{1-h}, \qquad \text{as } n \to \infty. \end{aligned}$$

As $\varepsilon > 0$ is arbitrary to start with, we have $b_n^{(1)} \to 0$. This convergence asserts that $Y_n \xrightarrow{\mathrm{D}} Y$. □

Step 4

Lastly, we want to determine the unique solution to (5). This leads us to a main result of this investigation.

Theorem 2. *Let $X_n^{(1)}$ be the number of nodes of outdegree 1 in a hierarchical lattice network with index p at age n. Then, we have*

$$\frac{X_n^{(1)} - \frac{p}{2-p} 2^n}{2^{n/2}} \xrightarrow{\mathrm{D}} \mathcal{N}\Big(0, \frac{2pq(p+1)}{(2-p)(2-p^2)}\Big).$$

Proof. The normally distributed random variate $\mathcal{N}(0, \sigma^2)$ solves (5), for an appropriate choice of the variance σ^2. It is not hard to check that

$$\mathcal{N}(0,\sigma^2) \;\overset{D}{=}\; \frac{p}{\sqrt{2}} \mathcal{N}_1(0,\sigma^2) + 2\,\mathcal{N}_2(0,1)\sqrt{\frac{pqc_p^{(1)}}{2}} + \mathcal{N}_3(0,1)\sqrt{\frac{pq(1-c_p^{(1)})}{2}},$$

where the three normally distributed random variates on the right-hand side are independent, is a solution of (5), if

$$\mathcal{N}(0,\sigma^2) \;\overset{D}{=}\; \mathcal{N}\Big(0, \frac{1}{2}p^2\sigma^2 + 2pqc_p^{(1)} + \frac{1}{2}pq(1-c_p^{(1)})\Big),$$

that is, if σ^2 is a solution to the algebraic equation

$$\sigma^2 = \frac{1}{2}p^2\sigma^2 + 2pqc_p^{(1)} + \frac{1}{2}pq(1 - c_p^{(1)}).$$

The solution to the latter is

$$\sigma^2 = \frac{2pq(p+1)}{(2-p)(2-p^2)}. \tag{6}$$

Indeed, $\mathcal{N}\Big(0, \frac{2pq(p+1)}{(2-p)(2-p^2)}\Big)$ is a solution to (5); Step 2 guarantees that it is a unique solution, and Step 3 establishes the convergence. □

6 Probabilistic analysis of the number of nodes of higher outdegree

We first use Lemma 2 to inductively produce the expectations of the number of nodes of higher degrees. It is easy to show from a simple recurrence that the order of the graph (number of vertices in it) is $O(2^n)$, with asymptotic $2^n p$ average. This is also noted in [7]. So, $\mathbb{E}[X_n^{(k)}]$, for each fixed k, is of the asymptotic form $c_k 2^n + o(2^n)$, as $n \to \infty$ (some of the c_k's might be 0). But then, we have $\sum_{k=1}^{\infty} c_k = p$. Taking expectations of the recurrence (2), then scaling by 2^n and passing to the limits (guaranteed to exit), as $n \to \infty$, we find a recurrence relating the coefficients:

$$c_k = \frac{1}{2}pc_k + \frac{1}{2}\sum_{i=\lceil \frac{k}{2}\rceil}^{k-1} \binom{i}{k-i} p^{2i-k} q^{k-i} c_i.$$

The following table gives the first few values of c_k, for the unbiased case $p = q$; each value is approximated to four decimal places. Note that nodes of the first seven smallest degrees constitute a proportion of more than 97% of the nodes in the network.

k	1	2	3	4	5	6	7
c_k	0.3333	0.0952	0.0254	0.0172	0.0070	0.0054	0.0035

In principle, the hierarchical system of recurrences in Lemma 2 can be used to inductively develop limit distributional equations for the (normalized) number of nodes of higher degrees. We analyze $X_n^{(2)}$ as an illustration. First, we obtain the following stochastic recurrence for $X_n^{(2)}$ by letting $k = 2$ in Lemma 2:

$$X_n^{(2)} \overset{D}{=} \mathrm{Bin}_2\big(X_{n-1}^{(2)}, p^2\big) + \mathrm{Bin}_1\big(X_{n-1}^{(1)}, q\big). \tag{7}$$

Using the same approach as that in the proof of Proposition 1, we find the first moment of $X_n^{(2)}$, which is

$$\begin{aligned}\mathbb{E}\big[X_n^{(2)}\big] &= \frac{pq}{(2-p)(2-p^2)}2^n + \frac{p^2+p-2}{p(2-p^2)}p^{2n} + \frac{2q}{p(2-p)}p^n \qquad (8)\\ &\sim \frac{pq}{(2-p)(2-p^2)}2^n.\end{aligned}$$

We square (7) toward second moment calculation of $X_n^{(2)}$. Upon taking averages, we get the term $\mathbb{E}\big[X_{n-1}^{(1)}X_{n-1}^{(2)}\big]$. Therefore, we need to develop $\mathbb{E}\big[X_n^{(1)}X_n^{(2)}\big]$ first. The two variables in the expectation are dependent, hence the expectation cannot be obtained by taking the product of $\mathbb{E}\big[X_n^{(1)}\big]$ and $\mathbb{E}\big[X_n^{(2)}\big]$. Our strategy is to develop a recurrence for $\mathbb{E}\big[X_n^{(1)}X_n^{(2)}\big]$ and solve it.

To clearly determine the dependency between the binomial random variables in (1) and (7), we need to first refine our recurrence for $X_n^{(1)}$. In (2), $2^n - X_n^{(1)}$ is used to represent the total number of edges out of nodes of outdegree at least 2. Separating out edges emanating out of nodes of outdegree 2, we get an alternative representation for (2), namely

$$2^n - X_n^{(1)} = 2X_n^{(2)} + \big(2^n - X_n^{(1)} - 2X_n^{(2)}\big),$$

where $2X_n^{(2)}$ is the total number of edges out of nodes of outdegree 2, and $2^n - X_n^{(1)} - 2X_n^{(2)} =: \mathcal{E}_n^{(\geq 3)}$ is the total number of edges out of nodes of outdegree at least 3. Therefore, (1) can be rewritten as

$$X_n^{(1)} = 2\mathrm{Bin}^*\big(X_{n-1}^{(1)}, p\big) + \mathrm{Bin}'\big(2X_{n-1}^{(2)}, p\big) + \mathrm{Bin}''\big(\mathcal{E}_n^{(\geq 3)}, p\big). \tag{9}$$

Here $\mathrm{Bin}^*\big(X_{n-1}^{(1)}, p\big)$, $\mathrm{Bin}'\big(2X_{n-1}^{(2)}, p\big)$ and $\mathrm{Bin}''\big(\mathcal{E}_n^{(\geq 3)}, p\big)$ are conditionally independent (given $X_{n-1}^{(1)}$ and $X_{n-1}^{(2)}$).

Multiplying (7) by (9) yields a recurrence for $X_n^{(1)}X_n^{(2)}$,

$$\begin{aligned} X_n^{(1)}X_n^{(2)} \stackrel{D}{=}\; & 2\mathrm{Bin}^*\big(X_{n-1}^{(1)}, p\big)\,\mathrm{Bin}_2\big(X_{n-1}^{(2)}, p^2\big) + 2\mathrm{Bin}^*\big(X_{n-1}^{(1)}, p\big)\,\mathrm{Bin}_1\big(X_{n-1}^{(1)}, q\big) \\ & + \mathrm{Bin}^{'}\big(2X_{n-1}^{(2)}, p\big)\,\mathrm{Bin}_2\big(X_{n-1}^{(2)}, p^2\big) + \mathrm{Bin}^{'}\big(2X_{n-1}^{(2)}\big)\,\mathrm{Bin}_1\big(X_{n-1}^{(1)}, q\big) \\ & + \mathrm{Bin}^{''}\big(\mathcal{E}_n^{(\geq 3)}, p\big)\,\mathrm{Bin}_2\big(X_{n-1}^{(2)}, p^2\big) \\ & + \mathrm{Bin}^{''}\big(\mathcal{E}_n^{(\geq 3)}, p\big)\,\mathrm{Bin}_1\big(X_{n-1}^{(1)}, q\big). \end{aligned} \tag{10}$$

Taking double expectations on both sides, we get a recurrence for $\mathbb{E}\big[X_n^{(1)}X_n^{(2)}\big]$. For simplicity, we shall denote the averages of the six terms on the right hand side in (10) by (a), (b), (c), (d), (e) and (f), respectively, and find their expectations one by one.

Note that the two binomial random variables in each of (a), (d), (e) and (f) are conditionally independent, therefore their expectations can be calculated in a manner like what we did for $\mathbb{E}\big[\big(X_n^{(1)}\big)^2\big]$. Also, observe that (b) can be rewritten as $2\mathrm{Bin}^*\big(X_{n-1}^{(1)}, p\big)\Big(X_{n-1}^{(1)} - \mathrm{Bin}^*\big(X_{n-1}^{(1)}, p\big)\Big)$, the expectation of which can be calculated similarly.

Unfortunately, the above approach could not be applied to get the expectation of (c), because the two Binomial random variables in (c) are conditionally dependent. Here the need arises for use of the more refined representation of the binomial random variable as a sum of indicators, as alluded to in Section 3. We write

$$\mathrm{Bin}^{'}\big(2X_{n-1}^{(2)}, p\big) = \sum_{i=1}^{X_{n-1}^{(2)}} \Big(\mathbb{I}_{\{U_{L_i}<p\}} + \mathbb{I}_{\{U_{R_i}<p\}}\Big),$$

$$\mathrm{Bin}_2\big(X_{n-1}^{(2)}, p^2\big) = \sum_{j=1}^{X_{n-1}^{(2)}} \mathbb{I}_{\{U_{L_j}<p\}}\mathbb{I}_{\{U_{R_j}<p\}}.$$

Here, the subscript $L_i(R_i)$ represents the left (right) edge out of the ith node of outdegree 2, respectively, and the product $\mathbb{I}_{\{U_{L_i}<p\}}\mathbb{I}_{\{U_{R_i}<p\}}$ indicates the event that the left and right edges out of the ith node of outdegree 2 are being serialized simultaneously. By construction, $\mathbb{I}_{\{U_{L_i}<p\}}$ is independent of $\mathbb{I}_{\{U_{L_j}<p\}}$, for any $i \neq j$, and $\mathbb{I}_{\{U_{L_i}<p\}}$ is independent of $\mathbb{I}_{\{U_{R_j}<p\}}$, for any i and j.

Using this new representation, (c) can be rewritten as

$$\begin{aligned}
&\mathrm{Bin}'\big(2X_{n-1}^{(2)},p\big)\,\mathrm{Bin}_2\big(X_{n-1}^{(2)},p^2\big)\\
&\quad=\sum_{i=1}^{X_{n-1}^{(2)}}\Big(\mathbb{I}_{\{U_{L_i}<p\}}+\mathbb{I}_{\{U_{R_i}<p\}}\Big)\sum_{j=1}^{X_{n-1}^{(2)}}\Big(\mathbb{I}_{\{U_{L_j}<p\}}\mathbb{I}_{\{U_{R_j}<p\}}\Big)\\
&\quad=\sum_{i=1}^{X_{n-1}^{(2)}}\sum_{j=1}^{X_{n-1}^{(2)}}\mathbb{I}_{\{U_{L_i}<p\}}\mathbb{I}_{\{U_{L_j}<p\}}\mathbb{I}_{\{U_{R_j}<p\}}+\sum_{i=1}^{X_{n-1}^{(2)}}\sum_{j=1}^{X_{n-1}^{(2)}}\mathbb{I}_{\{U_{R_i}<p\}}\mathbb{I}_{\{U_{L_j}<p\}}\mathbb{I}_{\{U_{R_j}<p\}}.
\end{aligned}$$

The expectations of these two terms can be calculated easily by conditioning on $X_{n-1}^{(2)}$ and using independence of the indicators. For instance, for $i\neq j$, we have

$$\mathbb{E}\big[\mathbb{I}_{\{U_{R_i}<p\}}\mathbb{I}_{\{U_{L_j}<p\}}\mathbb{I}_{\{U_{R_j}<p\}}\big]=\mathbb{E}\big[\mathbb{I}_{\{U_{R_i}<p\}}\big]\mathbb{E}\big[\mathbb{I}_{\{U_{L_j}<p\}}\big]\mathbb{E}\big[\mathbb{I}_{\{U_{R_j}<p\}}\big]=p^3,$$

whereas, for $i=j$, we have

$$\mathbb{E}\big[\mathbb{I}_{\{U_{R_i}<p\}}\mathbb{I}_{\{U_{L_i}<p\}}\mathbb{I}_{\{U_{R_i}<p\}}\big]=\mathbb{E}\big[\mathbb{I}_{\{U_{R_i}<p\}}\mathbb{I}_{\{U_{L_i}<p\}}\big]=p^2.$$

Therefore, we can break up the sums into components where $i=j$, and components where $i\neq j$, and use the expectations as discussed. So, now we obtain a recurrence for $\mathbb{E}\big[X_n^{(1)}X_n^{(2)}\big]$, namely

$$\begin{aligned}
\mathbb{E}\big[X_n^{(1)}X_n^{(2)}\big]&=p^3\,\mathbb{E}\big[X_{n-1}^{(1)}X_{n-1}^{(2)}\big]+pq\,\mathbb{E}\big[(X_{n-1}^{(1)})^2\big]\\
&\qquad+(p^32^{n-1}+2p^2q)\,\mathbb{E}\big[X_{n-1}^{(2)}\big]+(pq2^{n-1}-2pq)\,\mathbb{E}\big[X_{n-1}^{(1)}\big].
\end{aligned}$$

This linear recurrence can be explicitly solved, as all the nonrecursive parts have been established in Proposition 1 and (8). With $\mathbb{E}\big[X_n^{(1)}X_n^{(2)}\big]$ in our possession, we are able to get the exact variance of $X_n^{(2)}$, and the exact covariance between $X_n^{(1)}$ and $X_n^{(2)}$. As the formulæ are lengthy and less relevant to future analysis, we will only report their leading terms here, and relegate the exact expressions to Appendix A. Asymptotically, the variance of $X_n^{(2)}$, and the covariance between $X_n^{(1)}$ and $X_n^{(2)}$ are

$$\begin{aligned}
\mathbb{V}\mathrm{ar}\big[X_n^{(2)}\big]&\sim\frac{p(4-4p-2p^2-2p^3+10p^4-13p^5+10p^6-2p^7-p^8)}{(2-p)(2-p^2)(2-p^3)(2-p^4)}\,2^n\\
&=:\alpha_{22}2^n, \qquad (11)\\
\mathbb{C}\mathrm{ov}\big[X_n^{(1)},X_n^{(2)}\big]&\sim-\frac{2p^2(1-2p+2p^2-p^3)}{(2-p)(2-p^2)(2-p^3)}\,2^n=:\alpha_{12}2^n. \qquad (12)
\end{aligned}$$

Note that the coefficient $-2p^2(1-2p+2p^2-p^3)/((2-p)(2-p^2)(2-p^3))$ is always negative for $p \in (0,1)$; the increase of nodes of outdegree 1 occurs at the expense of nodes of outdegree 2, and vice versa.

In a way similar to the proof of Theorem 1, we get concentration laws for $X_n^{(2)}$:

$$\frac{X_n^{(2)}}{2^n} \xrightarrow{\text{a.s.}} \frac{pq}{(2-p)(2-p^2)} =: c_p^{(2)}, \tag{13}$$
$$\mathbb{E}|X_n^{(2)} - c_p^{(2)} 2^n| = O(2^{\frac{n}{2}}).$$

Having established these concentration results, we are now ready to get the asymptotic joint distribution of $\left(X_n^{(1)}, X_n^{(2)}\right)$ by the contraction method, adapting notation and the four steps mentioned above to be in the multivariate setting. The steps are very similar to those in the proof of the asymptotic distribution of $X_n^{(1)}$. Toward a bivariate form of Step 1 in succinct notation, we rewrite (4) as

$$Y_n^{(1)} = hY_{n-1}^{(1)} + S_n, \tag{14}$$

with $S_n = 2Z_nR_n + \tilde{Z}_n\tilde{R}_n$, and normalize $X_n^{(2)}$ as

$$\begin{aligned} Y_n^{(2)} &:= \frac{X_n^{(2)} - c_p^{(2)} 2^n}{\sqrt{2^n}} \\ &\stackrel{D}{=} W_n^{(2)} V_n^{(2)} + W_n^{(1)} V_n^{(1)} + \frac{p^2}{\sqrt{2}} Y_{n-1}^{(2)} + \frac{q}{\sqrt{2}} Y_{n-1}^{(1)} \\ &= \frac{p^2}{\sqrt{2}} Y_{n-1}^{(2)} + \frac{q}{\sqrt{2}} Y_{n-1}^{(1)} + S'_n, \end{aligned} \tag{15}$$

where $S'_n := W_n^{(2)} V_n^{(2)} + W_n^{(1)} V_n^{(1)}$,

$$W_n^{(2)} := \frac{\mathrm{Bin}_2\left(X_{n-1}^{(2)}, p^2\right) - p^2 X_{n-1}^{(2)}}{\sqrt{p^2(1-p^2)X_{n-1}^{(2)}}}, \quad W_n^{(1)} := \frac{\mathrm{Bin}_1\left(X_{n-1}^{(1)}, q\right) - qX_{n-1}^{(1)}}{\sqrt{pqX_{n-1}^{(1)}}},$$

$$V_n^{(2)} := \sqrt{\frac{p^2(1-p^2)}{2} \times \frac{X_{n-1}^{(2)}}{2^{n-1}}}, \quad V_n^{(1)} := \sqrt{\frac{pq}{2} \times \frac{X_{n-1}^{(1)}}{2^{n-1}}}.$$

From (14) (15) we have a bivariate distributional recursion:

$$\mathbf{Y_n} := \begin{pmatrix} Y_n^{(1)} \\ Y_n^{(2)} \end{pmatrix} \stackrel{\mathbf{D}}{=} \begin{pmatrix} h & 0 \\ \frac{q}{\sqrt{2}} & \sqrt{2}h^2 \end{pmatrix} \begin{pmatrix} Y_{n-1}^{(1)} \\ Y_{n-1}^{(2)} \end{pmatrix} + \begin{pmatrix} S_n \\ S'_n \end{pmatrix} =: \mathbf{AY_{n-1}} + \mathbf{B_n}. \tag{16}$$

Note that $W_n^{(1)}$ and $W_n^{(2)}$ are conditionally independent given $X_{n-1}^{(1)}$ and $X_{n-1}^{(2)}$ and both are asymptotically normal. Denote their limits by $W^{(1)}, W^{(2)}$ respectively. Also note that

$$V_n^{(1)} \xrightarrow{\text{a.s.}} \sqrt{\frac{pqc_p^{(1)}}{2}} =: V^{(1)}, \quad V_n^{(2)} \xrightarrow{\text{a.s.}} \sqrt{\frac{p^2(1-p^2)c_p^{(2)}}{2}} =: V^{(2)}.$$

An application of the multivariate central limit theorem shows that $\mathbf{B_n} \xrightarrow{\mathbf{D}} \mathbf{B}$, where

$$\mathbf{B} := \begin{pmatrix} S \\ S' \end{pmatrix} = \begin{pmatrix} 2ZR + \tilde{Z}\tilde{R} \\ W^{(2)}V^{(2)} + W^{(1)}V^{(1)} \end{pmatrix}$$

is a normal vector with

$$\mathbb{V}\text{ar}[S] = \frac{pq(1+3c_p^{(1)})}{2}, \ \mathbb{V}\text{ar}[S'] = \frac{p^2(1-p^2)c_p^{(2)} + pqc_p^{(1)}}{2},$$

and

$$\mathbb{C}\text{ov}[S, S'] = \frac{p^3q^2 - p^2q(2-p^2)}{(2-p)(2-p^2)}.$$

A thorough explanation and the calculation of $\mathbb{C}\text{ov}[S, S']$ can be found in Appendix B.

Under the assumption that $\mathbf{Y_n} \xrightarrow{\mathbf{D}} \mathbf{Y}$ we therefore get from (16) the limiting equation:

$$\mathbf{Y} = \mathbf{AY} + \mathbf{B}, \tag{17}$$

with $\mathbf{Y}$ independent of $\mathbf{B}$. $\mathbf{Y}$ has a distribution to be determined from analysis of the fixed point equation (17) in further steps.

We next argue Step 2. We extend the notions and notations to bivariate cases, and let $\mathbf{Y}$ and $\tilde{\mathbf{Y}}$ be two bivariate random vectors with distribution functions $F_{\mathbf{Y}}$ and $F_{\tilde{\mathbf{Y}}}$ in $\mathcal{F}_1$, which now means the space of bivariate distributions with finite first moment. We can go through Step 2 for the bivariate transformation

$$\mathcal{T}^{(2)} : \mathcal{F}_1 \times \mathcal{F}_1 \to \mathcal{F}_1 \times \mathcal{F}_1,$$

defined by $\mathcal{T}^{(2)}(F_{\mathbf{Y}})$ being the law of $\mathbf{AY} + \mathbf{B}$.

The desired contracting property of $\mathcal{T}^{(2)}$ follows from the calculation

$$\begin{aligned}\ell_1\Big(\mathcal{T}^{(2)}(\tilde{\mathbf{Y}}),\mathcal{T}^{(2)}(\mathbf{Y})\Big) &\le \mathbb{E}\big|(h\tilde{Y}^{(1)}+S)-(hY^{(1)}+S)\big| \\ &\quad + \mathbb{E}\Big|\Big(\sqrt{2}\,h^2\tilde{Y}^{(2)}+\frac{q}{\sqrt{2}}\tilde{Y}^{(1)}+S'\Big) \\ &\quad - \Big(\sqrt{2}\,h^2Y^{(2)}+\frac{q}{\sqrt{2}}Y^{(1)}+S'\Big)\Big| \\ &= h\mathbb{E}\big|\tilde{Y}^{(1)}-Y^{(1)}\big|+\sqrt{2}\,h^2\mathbb{E}\big|\tilde{Y}^{(2)}-Y^{(2)}\big| \\ &\quad + \frac{q}{\sqrt{2}}\,\mathbb{E}\big|\tilde{Y}^{(1)}-Y^{(1)}\big| \\ &\le \max\Big(\sqrt{2}\,h^2, h+\frac{q}{\sqrt{2}}\Big)\,\ell_1(F_{\tilde{\mathbf{Y}}},F_{\mathbf{Y}}).\end{aligned}$$

Note that $\max\Big(\sqrt{2}\,h^2, h+\frac{q}{\sqrt{2}}\Big)<1$, thus the transformation $\mathcal{T}^{(2)}$ is contracting. Uniqueness follows just as in the case of $k=1$.

We now briefly sketch the distance computations required in Step 3. We work with "hatted" counterparts defined on the space of an optimal coupling $(\mathbf{Y}_n,\mathbf{Y})$.

We compute the Wasserstein distance

$$\begin{aligned}b_n^{(2)} := \ell_1(F_{\mathbf{Y_n}},F_{\mathbf{Y}}) &\le \mathbb{E}\big|(hY_{n-1}^{(1)}+\hat{S}_n)-(hY^{(1)}+\hat{S})\big| \\ &\quad + \mathbb{E}\big|(2h^2Y_{n-1}^{(2)}+\frac{q}{\sqrt{2}}Y_{n-1}^{(1)}+\hat{S}'_n) \\ &\qquad - (2h^2Y^{(2)}+\frac{q}{\sqrt{2}}Y^{(1)}+\hat{S}')\big| \\ &\le 2h^2\mathbb{E}|Y_{n-1}^{(2)}-Y^{(2)}|+\Big(h+\frac{q}{\sqrt{2}}\Big)\mathbb{E}|Y_{n-1}^{(1)}-Y^{(1)}| \\ &\quad + \mathbb{E}|\hat{S}_n-\hat{S}|+\mathbb{E}|\hat{S}'_n-\hat{S}'| \\ &\le 2h^2b_{n-1}^{(2)}+\Big(h+\frac{q}{\sqrt{2}}-2h^2\Big)b_n^{(1)} \\ &\quad + \mathbb{E}|\hat{S}_n-\hat{S}|+\mathbb{E}|\hat{S}'_n-\hat{S}'|.\end{aligned}$$

In our discussion of the case $k=1$, we have shown that both $b_n^{(1)}$ and $\mathbb{E}|\hat{S}_n-\hat{S}|$ converge to 0. By very similar considerations (triangle inequalities, Cauchy-Schwarz inequalities, etc.), we can also show that $\mathbb{E}|\hat{S}'_n-\hat{S}'|$ converges to zero. This puts the latter distance calculation in the form

$$b_n^{(2)} \le 2h^2b_{n-1}^{(2)}+a_n,$$

where a_n is $o(1)$. We saw in the case $k = 1$ that an inequality of this type was sufficient to prove that $b_n^{(1)} \to 0$. Retracing these steps on the last inequality, we can now easily produce a proof that $b_n^{(2)} \to 0$, as $n \to \infty$, completing Step 3.

The next theorem essentially checks Step 4. We use $\mathcal{N}_{\mathbf{2}}(\mathbf{0}, \mathbf{\Sigma})$ to mean a bivariate normal distribution with mean $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and covariance matrix $\mathbf{\Sigma}$. In the following, the symbols $c_p^{(1)}$, $c_p^{(2)}$, are those in (3) and (13), and the symbols α_{12} and α_{22} are defined in (12) and (11). For consistency of bivariate notation and æsthetics, we rename σ^2 (which appeared in (6)) as α_{11}.

Theorem 3. *Let $X_n^{(1)}$ and $X_n^{(2)}$ be the number of nodes of outdegrees 1 and 2 in a hierarchical lattice network with index p at age n. Then, we have*

$$\frac{\begin{pmatrix} X_n^{(1)} \\ X_n^{(2)} \end{pmatrix} - 2^n \begin{pmatrix} c_p^{(1)} \\ c_p^{(2)} \end{pmatrix}}{2^{n/2}} \xrightarrow{\text{D}} \mathcal{N}_{\mathbf{2}}\left(\mathbf{0}, \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{12} & \alpha_{22} \end{pmatrix}\right).$$

Proof. Let $\mathbf{A} = \begin{pmatrix} h & 0 \\ \frac{q}{\sqrt{2}} & \sqrt{2}h^2 \end{pmatrix}$. We shall show that the bivariate normal distribution solves (17), with covariance matrix satisfying

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{12} & \alpha_{22} \end{pmatrix} = \mathbf{A} \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{12} & \alpha_{22} \end{pmatrix} \mathbf{A}^T + \begin{pmatrix} \mathbb{V}\text{ar}[S] & \mathbb{C}\text{ov}[S, S'] \\ \mathbb{C}\text{ov}[S, S'] & \mathbb{V}\text{ar}[S'] \end{pmatrix}.$$

Solving for $\alpha_{11}, \alpha_{12}, a_{22}$, we get

$$\alpha_{11} = \frac{2pq(p+1)}{(2-p)(2-p^2)},$$

$$\alpha_{12} = -\frac{2p^2(1-2p+2p^2-p^3)}{(2-p)(2-p^2)(2-p^3)},$$

and

$$\alpha_{22} = \frac{p(4-4p-2p^2-2p^3+10p^4-13p^5+10p^6-2p^7-p^8)}{(2-p)(2-p^2)(2-p^3)(2-p^4)}.$$

As a consequence of the previous steps we get the result.

□

7 Concluding remarks

We investigated $X_n^{(k)}$, the number of nodes of outdegree k in a random hierarchical lattice network at age n. We determined the exact first two moments of $X_n^{(1)}$. The structure offers quite a bit of challenge when we get to exact higher moments, requiring very intensive combinatorial computation. As a general paradigm, most second moments in this investigation are obtained with the help of a symbolic computation system (we used Maple ©). We also obtained an asymptotic Gaussian law for $X_n^{(1)}$. Owing to a recursive stochastic recurrence system relating $X_n^{(k)}$ to the number of nodes of smaller outdegrees, we can in principle go forward and develop similar limit laws for $X_n^{(k)}$, for $k \geq 1$. As an illustration, we showed how to extend the derivations for $X_n^{(1)}$ to get a bivariate Gaussian law for $(X_n^{(1)}, X_n^{(2)})$. The study needed a variation of the contraction method including toll terms depending on the recursion variable. Our success with the particular instance of hierarchical lattice network leads us to expect tweaked forms of the contraction method using some variant of the four steps we outlined in our approach, will succeed and be important tools for random structures that tend to double their size quickly (say in one step). These structures are in the vogue [6, 17]. It remains to be seen if a general methodology and broad theorems can be developed for this and similar types of recursions allowing to conclude limiting results for further interesting functions of the network.

Acknowledgment

We are indebted to Professors Ralph Neininger and Uwe Rösler for an encouraging round of discussion.

References

[1] Bernasconi, N., Panagiotou, K. and Steger, A. (2008). On the degree sequences of random outerplanar and series-parallel graphs. *Lecture Notes in Computer Science*, **5171**, 303–316.

[2] Bickel, P. and Freedman D. (1981). Some asymptotic theory for the bootstrap. *The Annals of Statistics*, **9**, 1196–1217.

[3] Cramer, M. and Rüschendorf, L. (1996). Convergence of a branching type recursion. *Annales de l'Iinstitut Henri Poincaré, Probabilités et Statistiques*, **32**, 725–741.

[4] Drmota, M., Giménez, O. and Noy, M. (2010). Vertices of given degree in series-parallel graphs. *Random Structures & Algorithms*, **36**, 273–314.

[5] Farmakis, I., Moskowitz, M. (2013) . *Fixed Point Theorems and Their Applications*. World Scientific, New Jersey.

[6] Feng, Q. (2014). Random fast growth models for treelike networks. *Private Communication*.

[7] Hambly, B. and Jordan, J. (2004). A random hierarchical lattice: the series-parallel graph and its properties. *Advances in Applied Probability*, **36**, 824–838.

[8] Karr, A. (1993). *Probability*. Springer, New York.

[9] Mahmoud, H. (2013). Some node degree properties of series-parallel graphs evolving under a stochastic growth model. *Probability in the Engineering and Informational Sciences*, **27**, 297–307.

[10] Mahmoud, H. (2014). Some properties of binary series-parallel graphs. *Probability in the Engineering and Informational Sciences* **28**, 565–572.

[11] Neininger, R. and Rüschendorf, L. (2004). A general limit theorem for recursive algorithms and combinatorial structures. *The Annals of Applied Probability*, **14**, 378–418.

[12] Neininger, R. (2001). On a multivariate contraction method for random recursive structures with applications to Quicksort. *Random Structures and Algorithms*, **19**, 498–524.

[13] Rachev, S. and Rüschendorf, L. (1995). Probability metrics and recursive algorithms. *Advances in Applied Probability*, **27**, 770–799.

[14] Rösler, U. (1991). A limit theorem for "Quicksort." *RAIRO Inform. Théor. Appl.*, **25**, 85–100.

[15] Rösler, U. and Rüschendorf, L. (2001). The contraction method for recursive algorithms. *Algorithmica*, **29**, 3–33.

[16] Vervaat, W. (1979). On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Advances in Applied Probability*, **11**, 750–783.

[17] Zhang, Z., Qi, Y., Zhou, S., Gao, S. and Guan, J. (2010). Explicit determination of mean firrst-passage time for random walks on deterministic uniform recursive trees. *Physical Review*, **E, 81**, 016114.

Appendix A

The exact covariance between $X_n^{(1)}$ and $X_n^{2)}$ is

$$
\begin{aligned}
\mathbb{C}\mathrm{ov}\big[X_n^{(1)} X_n^{(2)}\big] = {} & \frac{6-6p}{(p+1)(2-p)}\, p^n + \frac{2p^2(p^3-2p^2+2p-1)}{(2-p)(2-p^2)(2-p^3)}\, 2^n \\
& - \frac{(p^2+p-2)}{(2-p)(2-p^2)}\, (2p)^{2n} \\
& + \frac{(8p^5+2p^4-10p^3-16p^2-16p+32)}{p(p+1)(2-p)(2-p^2)(2-p^3)}\, p^{3n} \\
& + \frac{4p^3-6p^2-2p+4}{(2-p)^2(2-p^2)} (2p)^n \\
& - \frac{(2p^4-12p^3+2p^2+40p-32)}{p(2-p)^2(2-p^2)}\, p^{2n} \\
& - \frac{p^2(1-p)}{(2-p)^2(2-p^2)}\, 4^n \\
& \qquad - \Big(p^n - \frac{p^{n+1}}{2-p} + \frac{p}{2-p}\, 2^n\Big) \\
& \qquad\quad \times \Big(\frac{2-2p}{p(2-p)}\, p^n + \frac{p^2+p-2}{p(2-p^2)}\, p^{2^n} \\
& \qquad\qquad + \frac{p(1-p)}{(2-p)(2-p^2)}\, 2^n\Big).
\end{aligned}
$$

The exact variance of $X_n^{(2)}$ is

$$\mathbb{V}\mathrm{ar}\big[X_n^{(2)}\big] = \frac{p^4+10p^3+9p^2-14p-6}{p^2(p+1)(2-p^2)}\,p^{2n} + \frac{-2p^4-10p^3+8p^2-2p+6}{p(p^2+p+1)(p+1)(2-p)}\,p^n + \frac{-4p^4-12p^3-4p^2+20}{p^2(p+1)(2-p^3)}\,p^{3n} + \frac{p^7+4p^6+7p^5+9p^4+2p^3-3p^2-6p-14}{p^2(p+1)(p^2+p+1)(2-p^4)}\,p^{4^n} + \frac{p(4-4p-2p^2-2p^3+10p^4-13p^5+10p^6-2p^7-p^8)}{(2-p)(2-p^2)(2-p^3)(2-p^4)}\,2^n.$$

Appendix B

We show that $\begin{pmatrix} S_n \\ S'_n \end{pmatrix} \xrightarrow{\mathrm{D}} \mathcal{N}_{\mathbf{2}}(\mathbf{0}, \mathbf{\Sigma})$ with $\mathbf{\Sigma}$ specified below. Note that $\begin{pmatrix} S_n \\ S'_n \end{pmatrix}$ can be written in the following way:

$$\begin{pmatrix} S_n \\ S'_n \end{pmatrix} = \begin{pmatrix} 2Z_nR_n + \tilde{Z}_n\tilde{R}_n \\ W_n^{(1)}V_n^{(1)} + W_n^{(2)}V_n^{(2)} \end{pmatrix} = \begin{pmatrix} 2Z_nR_n \\ W_n^{(1)}V_n^{(1)} \end{pmatrix} + \begin{pmatrix} \tilde{Z}_n\tilde{R}_n \\ W_n^{(2)}V_n^{(2)} \end{pmatrix} = \begin{pmatrix} 2R_n & 0 \\ 0 & V_n^{(1)} \end{pmatrix}\begin{pmatrix} Z_n \\ W_n^{(1)} \end{pmatrix} + \begin{pmatrix} \tilde{R}_n & 0 \\ 0 & V_n^{(2)} \end{pmatrix}\begin{pmatrix} \tilde{Z}_n \\ W_n^{(2)} \end{pmatrix}.$$

Recall that

$$\tilde{Z}_n = \frac{\mathrm{Bin}(2^{n-1} - X_{n-1}^{(1)}, p) - p(2^{n-1} - X_{n-1}^{(1)})}{\sqrt{pq(2^{n-1} - X_{n-1}^{(1)})}},$$

where $2^{n-1} - X_{n-1}^{(1)} = 2X_{n-1}^{(2)} + \mathcal{E}_n^{(\geq 3)}$. Let

$$\tilde{Z}_n^{(1)} = \frac{\mathrm{Bin}(X_{n-1}^{(2)}, p) - pX_{n-1}^{(2)}}{\sqrt{pq(2^{n-1} - X_{n-1}^{(1)})}},$$

$$\tilde{Z}_n^{(2)} = \frac{\mathrm{Bin}(X_{n-1}^{(2)} + \mathcal{E}_n^{(\geq 3)}, p) - p(X_{n-1}^{(2)} + X_{n-1}^{(\geq 3)})}{\sqrt{pq(2^{n-1} - X_{n-1}^{(1)})}}.$$

We can further write $\begin{pmatrix} S_n \\ S'_n \end{pmatrix}$ as

$$\begin{pmatrix} 2R_n & 0 \\ 0 & V_n^{(1)} \end{pmatrix} \begin{pmatrix} Z_n \\ W_n^{(1)} \end{pmatrix} + \begin{pmatrix} \tilde{R}_n & 0 \\ 0 & V_n^{(2)} \end{pmatrix} \begin{pmatrix} \tilde{Z}_n^{(1)} \\ W_n^{(2)} \end{pmatrix} + \begin{pmatrix} \tilde{R}_n & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{Z}_n^{(2)} \\ 0 \end{pmatrix},$$

where $\begin{pmatrix} Z_n \\ W_n^{(1)} \end{pmatrix}$, $\begin{pmatrix} \tilde{Z}_n^{(1)} \\ W_n^{(2)} \end{pmatrix}$ and $\begin{pmatrix} \tilde{Z}_n^{(2)} \\ 0 \end{pmatrix}$ are mutually independent, and each of them can be written as the sum of independent, zero-mean random vectors. For example, we can write the first vector as

$$\begin{aligned}
\begin{pmatrix} Z_n \\ W_n^{(1)} \end{pmatrix} &= \frac{1}{\sqrt{pqX_{n-1}^{(1)}}} \begin{pmatrix} \mathrm{Bin}(X_{n-1}^{(1)}, p) - pX_{n-1}^{(1)} \\ \mathrm{Bin}(X_{n-1}^{(1)}, q) - qX_{n-1}^{(1)} \end{pmatrix} \\
&= \frac{1}{\sqrt{pqX_{n-1}^{(1)}}} \sum_{i=1}^{X_{n-1}^{(1)}} \begin{pmatrix} \mathbb{I}_{\{U_i < p\}} - p \\ \mathbb{I}_{\{U_i < q\}} - q \end{pmatrix},
\end{aligned}$$

where $\begin{pmatrix} \mathbb{I}_{\{U_i < p\}} - p \\ \mathbb{I}_{\{U_i < q\}} - q \end{pmatrix}$ and $\begin{pmatrix} \mathbb{I}_{\{U_j < p\}} - p \\ \mathbb{I}_{\{U_j < q\}} - q \end{pmatrix}$ are independent and have mean 0, for $i \neq j$. Note also that they are (conditionally) identically distributed. Therefore, the multivariate central limit theorem guarantees that $\begin{pmatrix} Z_n \\ W_n^{(1)} \end{pmatrix}$ converges in distribution to a mean-zero bivariate normal random vector. Similar arguments can show that $\begin{pmatrix} \tilde{Z}_n^{(1)} \\ W_n^{(2)} \end{pmatrix}$ and $\begin{pmatrix} \tilde{Z}_n^{(2)} \\ 0 \end{pmatrix}$ both converge in distribution to bivariate normal random vectors in the limit. Note that the three limit random vectors are mutually independent. Since

$$\begin{pmatrix} 2R_n & 0 \\ 0 & V_n^{(1)} \end{pmatrix} \xrightarrow{\text{a.s.}} \begin{pmatrix} 2R & 0 \\ 0 & V^{(1)} \end{pmatrix}, \begin{pmatrix} \tilde{R}_n & 0 \\ 0 & V_n^{(2)} \end{pmatrix} \xrightarrow{\text{a.s.}} \begin{pmatrix} \tilde{R} & 0 \\ 0 & V^{(2)} \end{pmatrix},$$

and

$$\begin{pmatrix} \tilde{R}_n & 0 \\ 0 & 0 \end{pmatrix} \xrightarrow{\text{a.s.}} \begin{pmatrix} \tilde{R} & 0 \\ 0 & 0 \end{pmatrix},$$

an application of Slutsky's theorem proves that the limit of $\begin{pmatrix} S_n \\ S'_n \end{pmatrix}$ is a linear combination of three zero-mean, independent bivariate normal random vectors, hence is itself a zero-mean bivariate normal random vector.

We next take up $\mathbb{C}\text{ov}[S, S']$, i.e. $\mathbf{\Sigma}$. Note that $\mathbb{C}\text{ov}[S, S'] = \mathbb{C}\text{ov}[2ZR + \tilde{Z}\tilde{R}, W^{(2)}V^{(2)} + W^{(1)}V^{(1)}]$. We can find it by taking the limit of $\mathbb{C}\text{ov}[2Z_nR_n + \tilde{Z}_n\tilde{R}_n, W_n^{(2)}V_n^{(2)} + W_n^{(1)}V_n^{(1)}]$, which needs careful calculation due to the intricate dependency structure between $(Z_n, \tilde{Z}_n)$ and $(W_n^{(1)}, W_n^{(2)})$.

Recall that Z_n and $\tilde{Z}_n$, $W_n^{(1)}$ and $W_n^{(2)}$ are conditionally independent. Furthermore, Z_n and $W_n^{(2)}$, $\tilde{Z}_n$ and $W_n^{(1)}$ are conditionally independent as well, because they depend on different sets of edges in the network. For example, Z_n depends on edges coming out of nodes of outdegree 1, while $W_n^{(2)}$ is related to edges out of nodes of outdegree 2. On the other hand, Z_n and $W_n^{(2)}$, $\tilde{Z}_n$ and $W_n^{(2)}$ are dependent. For example, $\tilde{Z}_n$ depends on edges out of nodes of outdegree greater than or equal to 2, which includes outdegree 2, thus is related to $W_n^{(2)}$.

In view of the above analysis, we have

$$\begin{aligned}
&\mathbb{C}\text{ov}[2Z_nR_n + \tilde{Z}_n\tilde{R}_n, W_n^{(2)}V_n^{(2)} + W_n^{(1)}V_n^{(1)}] \\
&\qquad = \mathbb{C}\text{ov}[2Z_nR_n, W_n^{(2)}V_n^{(2)}] + \mathbb{C}\text{ov}[2Z_nR_n, W_n^{(1)}V_n^{(1)}] \\
&\qquad\quad + \mathbb{C}\text{ov}[\tilde{Z}_n\tilde{R}_n, W_n^{(2)}V_n^{(2)}] + \mathbb{C}\text{ov}[\tilde{Z}_n\tilde{R}_n, W_n^{(1)}V_n^{(1)}] \\
&\qquad = \mathbb{C}\text{ov}[2Z_nR_n, W_n^{(1)}V_n^{(1)}] + \mathbb{C}\text{ov}[\tilde{Z}_n\tilde{R}_n, W_n^{(2)}V_n^{(2)}],
\end{aligned}$$

where both terms can be found in the same way as $\mathbb{C}\text{ov}[X_n^{(1)}, X_n^{(2)}]$. We have

$$
\begin{aligned}
\mathbb{C}\mathrm{ov}[2Z_nR_n, W_n^{(1)}V_n^{(1)}] &= \mathbb{E}[2Z_nR_nW_n^{(1)}V_n^{(1)}] - \mathbb{E}[2Z_nR_n]\,\mathbb{E}[W_n^{(1)}V_n^{(1)}] \\
&= \mathbb{E}[2Z_nR_nW_n^{(1)}V_n^{(1)}] \\
&= \mathbb{E}\Big[2\,\frac{\mathrm{Bin}(X_{n-1}^{(1)}, p) - pX_{n-1}^{(1)}}{\sqrt{2^n}} \times \frac{\mathrm{Bin}(X_{n-1}^{(1)}, q) - qX_{n-1}^{(1)}}{\sqrt{2^n}}\Big] \\
&= \mathbb{E}\left[-2\,\frac{\big(pX_{n-1}^{(1)} - \mathrm{Bin}(X_{n-1}^{(1)}, p)\big)^2}{2^n}\right] \\
&= -\frac{\mathbb{E}[\mathbb{E}[\big(pX_{n-1}^{(1)} - \mathrm{Bin}(X_{n-1}^{(1)}, p)\big)^2 | X_{n-1}^{(1)}]]}{2^{n-1}} \\
&= -\frac{pq\mathbb{E}[X_{n-1}^{(1)}]}{2^{n-1}} \\
&= -\frac{p^2q}{2-p},
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{C}\mathrm{ov}[\tilde{Z}_n\tilde{R}_n, W_n^{(2)}V_n^{(2)}] &= \mathbb{E}[\tilde{Z}_n\tilde{R}_nW_n^{(2)}V_n^{(2)}] - \mathbb{E}[\tilde{Z}_n\tilde{R}_n]\,\mathbb{E}[W_n^{(2)}V_n^{(2)}] \\
&= \mathbb{E}[\tilde{Z}_n\tilde{R}_nW_n^{(2)}V_n^{(2)}] \\
&= \mathbb{E}\Bigg[\frac{\mathrm{Bin}(2^{n-1} - X_{n-1}^{(1)}, p) - p(2^{n-1} - X_{n-1}^{(1)})}{\sqrt{2^n}} \\
&\qquad \times \frac{\mathrm{Bin}(X_{n-1}^{(2)}, p^2) - p^2X_{n-1}^{(2)}}{\sqrt{2^n}}\Bigg] \\
&= \frac{1}{2^n}\Bigg(\mathbb{E}\Big[\mathrm{Bin}(2^{n-1} - X_{n-1}^{(1)}, p)]\,\mathrm{Bin}(X_{n-1}^{(2)}, p^2)\Big] \\
&\qquad - \mathbb{E}\Big[p(2^{n-1} - X_{n-1}^{(1)})\,\mathrm{Bin}(X_{n-1}^{(2)}, p^2)\Big] \\
&\qquad - \mathbb{E}\Big[p^2X_{n-1}^{(2)}\,\mathrm{Bin}(2^{n-1} - X_{n-1}^{(1)}, p)\Big] \\
&\qquad + \mathbb{E}\Big[p(2^{n-1} - X_{n-1}^{(1)})\,p^2X_{n-1}^{(2)}\Big]\Bigg).
\end{aligned}
$$

The last three terms can be calculated by taking double expectation. The first one needs some work. Note that $\mathrm{Bin}(2^{n-1} - X_{n-1}^{(1)}, p) = \mathrm{Bin}(2X_{n-1}^{(2)}, p) +$

$\mathrm{Bin}(\mathcal{E}_n^{(\geq 3)}, p)$, and $\mathrm{Bin}(\mathcal{E}_n^{(\geq 3)}, p)$ is conditionally independent of $\mathrm{Bin}(X_{n-1}^{(2)}, p^2)$. Therefore, we have

$$\begin{aligned}
&\mathbb{E}\Big[\mathrm{Bin}(2^{n-1} - X_{n-1}^{(1)}, p)]\,\mathrm{Bin}(X_{n-1}^{(2)}, p^2)\Big] \\
&\quad = \mathbb{E}\Big[\mathrm{Bin}(2X_{n-1}^{(2)}, p)\,\mathrm{Bin}(X_{n-1}^{(2)}, p^2)\Big] + \mathbb{E}\Big[\mathrm{Bin}(\mathcal{E}_n^{(\geq 3)}, p)\,\mathrm{Bin}(X_{n-1}^{(2)}, p^2)\Big],
\end{aligned}$$

where the first expectation has been calculated in Section (6) to be $\mathbb{E}[2p^3(X_{n-1}^{(2)})^2 + 2p^2 q X_{n-1}^{(2)}]$.

Combining all the terms yields

$$\mathbb{C}\mathrm{ov}[\tilde{Z}_n \tilde{R}_n, W_n^{(2)} V_n^{(2)}] = \frac{p^2 q \mathbb{E}[X_{n-1}^{(2)}]}{2^{n-1}} \sim \frac{p^3 q^2}{(2-p)(2-p^2)}.$$

Upon taking limits, we have

$$\mathbb{C}\mathrm{ov}[S, S'] = \frac{p^3 q^2 - p^2 q(2-p^2)}{(2-p)(2-p^2)}.$$