Probability in the Engineering and Informational Sciences, **28**, 2014, 565–572.
doi:10.1017/S0269964814000138

SOME PROPERTIES OF BINARY SERIES-PARALLEL GRAPHS

Hosam M. Mahmoud
Department of Statistics
The George Washington University
Washington, DC 20052
USA
E-mail: hosam@gwu.edu

We introduce an incremental growth model for directed *binary* series-parallel (SP) graphs. The vertices of a directed binary SP graphs can only have outdgrees 1 or 2. We show that the number of vertices of outdegree 1 have a normal distribution (so, necessarily, the vertices of outdegree 2 have a normal distribution, too). Furthermore, we study the average length of a random walk between the poles of the graph. The asymptotic equivalent of the latter property includes the golden ratio. Pólya urns will systematically provide a coding method to initiate the studies.

1. INTRODUCTION

Series-parallel (SP) graphs are used to model flow in a bipolar network. For example, the network can represent routes between a source (producer) and a sink (market), and the SP graph models the flow of commercial goods from the producer to the market.

There are a few definitions of families of SP graphs. We use the common notation K_n for the complete graph on n vertices. In one definition, an SP graph is an undirected connected graph that does not contain K_4 as a minor. This definition does not distinguish two vertices as a source and a sink. One other popular recursive definition constructs SP graphs from smaller members of the family. In this variant, the smallest SP graph is K_2. The two vertices are called *poles* (one *North* and one *South*). Larger SP graphs are obtained 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 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.

Several discussions in the sequel will be connected to the the order and size of a graph. The *order* of a graph is the number of its vertices, and the *size* is the number of its edges. Models of randomness that have been proposed are the uniform model, where all SP networks of a certain size are equally likely [2,4], the hierarchical lattice model [5], where at each stage of the growth every edge is either serialized or parallelized, and the incremental model introduced in [7], in which one edge at a time is serialized or parallelized.

© Cambridge University Press 2014 0269-9648/14 $25.00

https://doi.org/10.1017/S0269964814000138 Published online by Cambridge University Press

Applications often impose restrictions on the growth of random graphs. There can be a restriction on the outdegrees of a vertex. For instance, each town along the trading routes from the producer to the market may have a limited budget to build at most two roads toward the market. Our aim is to study the effect of such a restriction on SP networks, to put in respective the performance of such restricted random models in comparison with unrestricted random SP networks. We introduce a stochastic model for the growth of directed binary SP graphs, and study several of their properties.

The *binary directed series-parallel graph* (henceforth binary SP graph) grows as follows. The smallest binary SP graph is the directed K_2, which we draw vertically and take the convention that the top vertex is the North Pole, which is the source, and the bottom vertex is the South Pole, which is the sink; the sense of the directed edge is from the North Pole to the South Pole. At each step of the evolution, a directed edge is chosen at random, say uv, directed from vertex u to vertex v. If uv is the only edge out of u (i.e., the outdegree of u is 1), we double or *parallelize* the edge, otherwise (u has outdegree 2), we extend or *serialize* it. To parallelize uv, we create a new edge, also directed from u to v, raising the outdegree of u from 1 to 2. To serialize uv, we remove that edge, and replace it with two directed edges ux (directed from u to x), and xv (directed from x to v). Thus, a new vertex, x, appears. Note that orientation in the plane is part of the definition of a binary SP graph: a vertex with outdegree 2 has two children, distinguished as one left child and one right child. Figure 1 shows a few small binary SP graphs and their probabilities (written above the North Pole of each graph). In Figure 1, the two binary SP graphs of size 3 illustrate the plane embedding — they are two different binary SP graphs, although they are isomorphic graphs in the graph-theoretic sense.

2. PÓLYA URNS

In this paper, certain Pólya urn schemes appear as a tool for the formulation of properties binary SP graphs. We say a brief word about these schemes and establish notation. A k-color *Pólya urn scheme* is an urn containing balls of up to k different colors, subject to rules of evolution in discrete time. After each unit of time, a ball is drawn uniformly at random (all balls being equally likely) from the urn, and the color of the ball is observed. The ball is then placed back in the urn, and depending on its color, a certain number of balls of each color is added. If the ball drawn is of color i, we add $A_{i,j}$ balls of color j, for $j = 1, \ldots, k$. (Interpret "add" to mean take out balls, if $A_{i,j}$ is negative.) The dynamics of the evolution of the urn can thus be represented by the *replacement matrix*

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

3. THE ORDER OF A RANDOM BINARY SP GRAPH

In the binary SP growth model, we add an edge at each step, starting from a directed K_2. After n steps we have $n+1$ edges in the graph. What is the order of the graph at this point?

Let us color the edges of the binary graph according to the following scheme. Let edges emanating out of a vertex of outdegree 2 be colored with white, and edges emanating out

https://doi.org/10.1017/S0269964814000138 Published online by Cambridge University Press

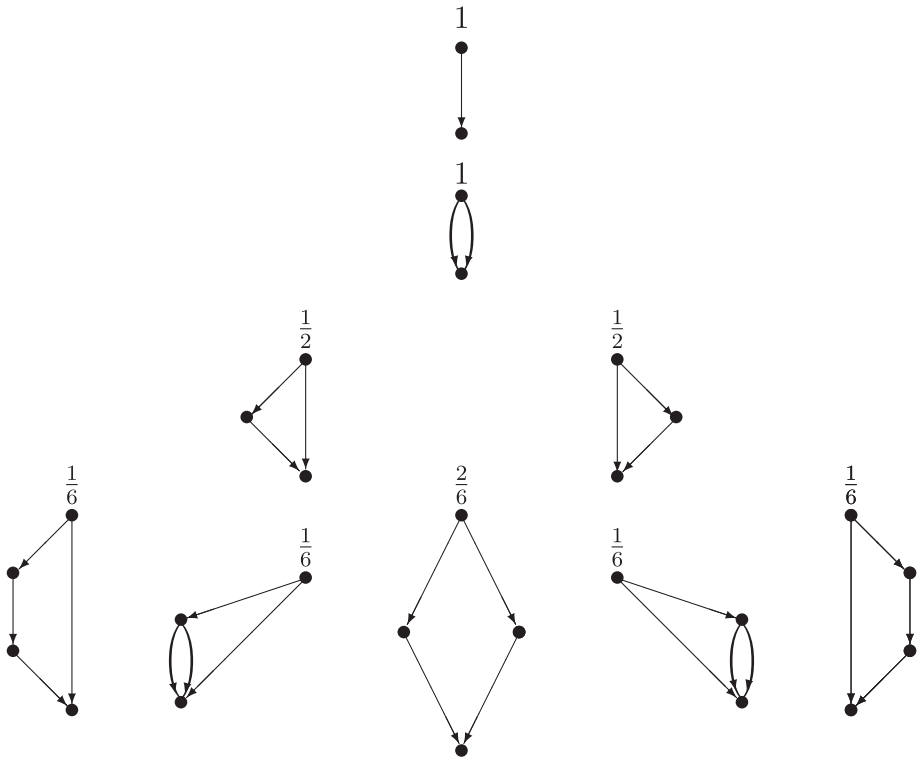


FIGURE 1. Some small binary SP graphs and their probabilities. In each of these graphs, the top vertex is the source and the bottom vertex is the sink.

of a vertex of outdegree 1 be colored with blue. After n edge insertions, let the number of white and blue balls be W_n and B_n, respectively.

There is a Pólya urn underlying the growth of the graph. Think of the edges of the binary SP graph as balls in an urn. At step n, if the randomly chosen edge (ball) is white, we serialize it into two, the one closer to the source is white and the one further from the source is blue, a net gain of one blue ball. If the randomly chosen edge is blue, it gets parallelized into two white edges: we remove one blue edge and add two white.

The replacement matrix associated with our urn is

$$\begin{pmatrix} 0 & 1 \\ 2 & -1 \end{pmatrix}; \tag{1}$$

Pólya urns of this type have already been analyzed. They appear in the analysis of random binary search trees (see [3]).

THEOREM 1: *Let X_n be the order of a random binary SP graph of size n. This order follows a central limit tendency:*

$$\frac{X_n - \frac{2}{3}n}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{2}{45}\right).$$

PROOF: In [3], each leaf of a binary search tree is associated with two white balls, and each non-leaf (so-called internal) node is associate with one blue. The overall dynamics of the

https://doi.org/10.1017/S0269964814000138 Published online by Cambridge University Press

binary search tree coincide with a Pólya urn scheme with the replacement matrix (1). The count of the leaves, L_n, satisfies

$$\frac{L_n - \frac{1}{3}n}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{2}{45}\right).$$

In our binary SP graph context, $2L_n$ has the same distribution as W_n. In the binary SP, the white edges appear in pairs, and each pair emanates out of one common vertex of outdegree 2, whereas each blue edges emanates out of a vertex of outdegree 1. Thus, we have $X_n = \frac{1}{2}W_n + B_n + 1$ (the 1 accounts for the South Pole), that is, $X_n = \frac{1}{2}W_n + (n - W_n) + 1 = n - \frac{1}{2}W_n + 1$, and the theorem follows. ■

4. TYPICAL GRAPH LENGTH

We next investigate the average length of a random walk from the North Pole to the South Pole (i.e., the average length of a randomly constructed typical path directed from the North Pole to the South Pole). Start a walk from the North Pole toward the South Pole by constructing a path that follows edges coming out of vertices of outdegree 1, and randomizes the choice for edges coming out of vertices of outdegree 2: at a vertex of outdegree 2, choose to come down the left or the right edge with equal probability. Continue in this fashion, till the South Pole is reached. We call such a path a *random walk* from the North Pole to the South Pole on the binary SP graph.

By symmetry, the length of a random walk on the binary SP graph is distributed like the left silhouette, which we define as follows. Start at the North Pole. At each vertex down the path, if the vertex you are at has outdegree 1, follow the sole edge emanating out of it. If the vertex has outdegree 2, follow the left edge emanating out of it. Continue in this fashion, till the South Pole is reached. The path so constructed is the *left silhoutte* of the binary SP graph.

To compute the average length of a random walk on the binary SP graph (also the average length of the left silhouette), we code the edges of the binary SP graph by three colors: Color 1 is for edges on the left silhouette that come out of vertices of outdogree 1, Color 2 is for edges on the left silhouette that come out of vertices of outdogree 2, and Color 3 is for edges not on the left silhouette. Since the number of colors here is small, we can name them: Colors 1, 2 and 3 can be taken to be white (W), blue (B) and red (R). Figure 2 illustrates this color coding.

Let W_n, B_n and R_n be respectively the number of white, blue and red edges in the graph after n edge insertions. As we start from a directed K_2, we have $W_0 = 1$, $B_0 = R_0 = 0$. The evolution occurs as follows. At step n, we choose an edge at random. If the edge is white, it is on the left silhouette, and the vertex it is emanating from has outdegree 1. In this case, we parallelize the edge: that white edge disappears and is replaced by two parallel edges, one blue on the left silhouette, and one red not on the left silhouette. If the chosen edge is blue, we serialize it, the pair replacing it is a short path of length 2 on the left silhouette, the edge of that path closer to the North Pole emanates from a vertex with outdegree 2, thus it is blue; the edge closer to the South Pole is white (a net gain of one white edge). If the chosen edge is red, it is not on the left silhouette: we serialize or parallelize it according to the growth rules (according to whether it is coming out of a vertex of outdegree 1 or 2). Whether we serialize or parallelize this edge, the contribution is another red edge also not on the left silhouette (a net gain of one red edge). The replacement matrix associated with

https://doi.org/10.1017/S0269964814000138 Published online by Cambridge University Press

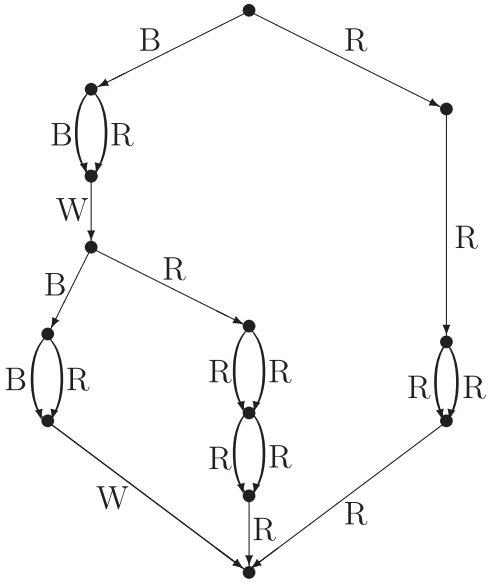


FIGURE 2. The color coding of a binary SP graph: edges on the silhouette that emanate out of a vertex with outdegree 1 are white (W), edges on the silhouette that emanate out of a vertex with outdegree 2 are blue (B), and edges not on the silhouette are red (R).

this Pólya urn scheme is

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

The eigenvalues of this matrix are $\lambda_1 = 1, \lambda_2 = \phi - 1$, and $\lambda_3 = -\phi$, where ϕ is the golden ratio $\frac{1}{2}(\sqrt{5}+1)$. So, $\lambda_1 = 1$ is the principal eigenvalue, with a corresponding left eigenvector

$$\mathbf{x}_1 = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

The other two left eigenvectors, corresponding to λ_2 and λ_3, respectively, are

$$\mathbf{x}_2 = \begin{pmatrix} \phi - 2 \\ 1 - \phi \\ 1 \end{pmatrix}, \quad \mathbf{x}_3 = \begin{pmatrix} -\phi - 1 \\ \phi \\ 1 \end{pmatrix}.$$

According to standard Pólya urn theory, we have

$$\frac{W_n}{n} \xrightarrow{a.s.} \lambda_1 v_1 = 0;$$

see [1] or [6]. This only tells us that the scale n is too high to produce a nontrivial limit. In search for the right scaling, we resort to an exact calculation of the average number of edges of each color.

https://doi.org/10.1017/S0269964814000138 Published online by Cambridge University Press

THEOREM 2: *Let S_n be the length (in terms of number of edges) of the left silhouette in a random binary SP after n edge insertions. As $n \to \infty$, we have the asymptotic relation*

$$\mathbf{E}[S_n] = \frac{\phi}{\sqrt{5}\,\Gamma(\phi)}\, n^{\phi-1} + O\big(n^{\phi-2}\big).$$

PROOF: After n edge insertions, let I_n^X be the indicator of picking color $X \in \{W, B, R\}$ in the nth step. The number of edges of the three colors satisfy the simultaneous recurrences:

$$W_n = W_{n-1} - I_n^W + I_n^B, \tag{2}$$

$$B_n = B_{n-1} + I_n^W, \tag{3}$$

$$R_n = R_{n-1} + I_n^W + I_n^R.$$

Let $\mathcal{F}_n$ be the sigma field generated by the first n edge insertions. Conditioning on $\mathcal{F}_{n-1}$, we have

$$\begin{aligned}
\mathbf{E}\big[W_n \,|\, \mathcal{F}_{n-1}\big] &= W_{n-1} - \mathbf{E}\big[I_n^W \,|\, \mathcal{F}_{n-1}\big] + \mathbf{E}\big[I_n^B \,|\, \mathcal{F}_{n-1}\big],\\
\mathbf{E}\big[B_n \,|\, \mathcal{F}_{n-1}\big] &= B_{n-1} + \mathbf{E}\big[I_n^W \,|\, \mathcal{F}_{n-1}\big],\\
\mathbf{E}\big[R_n \,|\, \mathcal{F}_{n-1}\big] &= R_{n-1} + \mathbf{E}\big[I_n^W \,|\, \mathcal{F}_{n-1}\big] + \mathbf{E}\big[I_n^R \,|\, \mathcal{F}_{n-1}\big].
\end{aligned}$$

These conditional expectations can be written as

$$\begin{aligned}
\mathbf{E}\big[W_n \,|\, \mathcal{F}_{n-1}\big] &= W_{n-1} - \frac{W_{n-1}}{n} + \frac{B_{n-1}}{n},\\
\mathbf{E}\big[B_n \,|\, \mathcal{F}_{n-1}\big] &= B_{n-1} + \frac{W_{n-1}}{n},\\
\mathbf{E}\big[R_n \,|\, \mathcal{F}_{n-1}\big] &= R_{n-1} + \frac{W_{n-1}}{n} + \frac{R_{n-1}}{n}.
\end{aligned}$$

Let $\mathbf{X}_n$ be the vector with the three components W_n, B_n and R_n. Taking (unconditional) averages, the three simultaneous recurrences can be written in matrix form:

$$\mathbf{E}[\mathbf{X}_n] = \mathbf{E}[\mathbf{X}_{n-1}] + \frac{1}{n}\mathbf{A}^T\mathbf{E}[\mathbf{X}_{n-1}] = \left(\mathbf{I} + \frac{1}{n}\mathbf{A}^T\right)\mathbf{E}[\mathbf{X}_{n-1}] := \mathbf{Q}_n\mathbf{E}[\mathbf{X}_{n-1}],$$

where $\mathbf{A}^T$ is the transpose of $\mathbf{A}$, and $\mathbf{I}$ is the 3×3 identity matrix. As the three eigenvalues are distinct, they give rise to simple Jordan normal forms — the matrix $\mathbf{Q}_j$ can be written as

$$\mathbf{Q}_j = \mathbf{M}\begin{pmatrix} 1+\frac{\lambda_1}{j} & 0 & 0\\ 0 & 1+\frac{\lambda_2}{j} & 0\\ 0 & 0 & 1+\frac{\lambda_3}{j}\end{pmatrix}\mathbf{M}^{-1},$$

where

$$\mathbf{M} = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \mathbf{x}_3] = \begin{pmatrix} 0 & \phi-2 & -\phi-1\\ 0 & 1-\phi & \phi\\ 1 & 1 & 1\end{pmatrix}.$$

Let

$$\mathbf{D}_j = \begin{pmatrix} 1+\frac{\lambda_1}{j} & 0 & 0\\ 0 & 1+\frac{\lambda_2}{j} & 0\\ 0 & 0 & 1+\frac{\lambda_3}{j}\end{pmatrix},$$

https://doi.org/10.1017/S0269964814000138 Published online by Cambridge University Press

and proceed with

$$\mathbf{E}[\mathbf{X}_n] = \left(\mathbf{M}\,\mathbf{D}_n\,\mathbf{M}^{-1}\right)\left(\mathbf{M}\,\mathbf{D}_{n-1}\,\mathbf{M}^{-1}\right)\cdots\left(\mathbf{M}\,\mathbf{D}_1\,\mathbf{M}^{-1}\right)\mathbf{E}[\mathbf{X}_0]$$

$$= \mathbf{M}\,\mathbf{D}_n\mathbf{D}_{n-1}\cdots\mathbf{D}_1\;\mathbf{M}^{-1}\begin{pmatrix}1\\0\\0\end{pmatrix}.$$

We thus have the exact vector of means:

$$\mathbf{E}[\mathbf{X}_n] = \frac{1}{n!}\,\mathbf{M}\begin{pmatrix}\dfrac{\Gamma(n+\lambda_1)}{\Gamma(1+\lambda_1)} & 0 & 0\\ 0 & \dfrac{\Gamma(n+\lambda_2)}{\Gamma(1+\lambda_2)} & 0\\ 0 & 0 & \dfrac{\Gamma(n+\lambda_3)}{\Gamma(1+\lambda_3)}\end{pmatrix}\mathbf{M}^{-1}\begin{pmatrix}1\\0\\0\end{pmatrix}.$$

Multiplying out, we obtain (with the help of symbolic manipulation system, like MAPLE)

$$\mathbf{E}[W_n] = \frac{\pi(\phi-1)\,\Gamma(n+\phi) + \phi\,\sin(\pi\phi)\,\Gamma^2(\phi)\,\Gamma(n+1-\phi)}{\sqrt{5}\,\pi\,\Gamma(\phi)\,\Gamma(n+1)},$$

$$\mathbf{E}[B_n] = \frac{\pi\,\Gamma(n+\phi) - \sin(\pi\phi)\,\Gamma^2(\phi)\Gamma(n+1-\phi)}{\sqrt{5}\,\pi\,\Gamma(\phi)\,\Gamma(n+1)}.$$

Both white and blue edges appear only on the left silhouette, and the length of that path is $W_n + B_n$. Hence, we have

$$\mathbf{E}[S_n] = \mathbf{E}[W_n] + \mathbf{E}[B_n] = \frac{\pi\phi\,\Gamma(n+\phi) + (\phi-1)\,\sin(\pi\phi)\,\Gamma^2(\phi)\,\Gamma(n+1-\phi)}{\sqrt{5}\,\pi\,\Gamma(\phi)\,\Gamma(n+1)}.$$

The asymptotic formula given is a consequence of Stirling's approximation of the ratio of Gamma functions. ■

Note that

$$\lim_{n\to\infty}\frac{\mathbf{E}[B_n]}{\mathbf{E}[W_n]} = \phi.$$

COROLLARY 1: *Let $\tilde{S}_n$ be the length of a random walk from the North Pole to the South Pole on a random binary SP graph of size n. As $n \to \infty$, we have*

$$\mathbf{E}[\tilde{S}_n] = \frac{\phi}{\sqrt{5}\,\Gamma(\phi)}\,n^{\phi-1} + O\left(n^{\phi-2}\right).$$

5. CONCLUDING REMARKS

We conducted a probabilistic analysis of some properties of random binary SP graphs, which are models for commercial networks for transferring goods between a producer and a market. The vertices are restricted by budgetary considerations to have at most two edges emanating out of them. The analysis shows that about 1/3 of the vertices of a large binary SP graph are forced to consume their entire budget (i.e., have two children), and

https://doi.org/10.1017/S0269964814000138 Published online by Cambridge University Press

this random number is concentrated around its average. The routes between the producer and the market are of average length of the order $n^{\phi-1} = n^{1/\phi} \approx n^{0.618}$, a sublinear length, so the average length does not degenerate into large orders of magnitude that do not favor the use of the model.

How typical is this average pole-to-pole path length in random binary SP graphs? This issue could be best answered, if we determine the order of the variance. An exact variance computation is daunting, and we leave this for future research. For the moment, we have only the order; we give the reader a sense of how it can be obtained. We can square each of the two Eqs. (2) and (3), then take expectations. The term $\mathbf{E}[W_{n-1}B_{n-1}]$ will appear in both. We can then take the product of (2) and (3), to get a third recurrence in $\mathbf{E}[W_nB_n]$ itself. We can proceed by considering the three equations for $\mathbf{E}[W_n^2]$, $\mathbf{E}[B_n^2]$ and $\mathbf{E}[W_nB_n]$ as a simultaneous recurrence system. This system contains $\mathbf{E}[W_n]$ and $\mathbf{E}[B_n]$, which have complicated exact values. From there we can make progress by considering asymptotic equivalents, instead of exact values, or by dropping terms toward tight inequalities. From upper and lower bounds, the variance turns out to be large, of the order n^{ϕ}, and there can be large fluctuations about the mean path length, rendering it not the most typical value. The asymptotic variance here, and for an entire class of similar Pólya urn models, remains as an open problem, worthy of vigorous pursuit.

Acknowledgment

The author thanks a referee for a thoughtful report. By following the recommendation of that report the exposition was significantly improved.

References

1. Athreya, K. & Karlin, S. (1968). Embedding of urn schemes into continuous time Markov branching process and related limit theorems. *Ann. Math. Stat.* 39: 1801–1817.
2. Bernasconi, N., Panagiotou, K. & Steger, A. (2008). On the degree sequences of random outerplanar and series-parallel graphs. *Lect. Notes Comput. Sci.* 5171: 303–316.
3. Devroye, L. (1991). Limit laws for local counters in random binary search trees. *Random Struct. Algorithms* 2: 303–316.
4. Drmota, M., Giménez, O. & Noy, M. (2010). Vertices of given degree in series-parallel graphs. *Random Struct. Algorithms* 36: 273–314.
5. Hambly, B. & Jordan, J. (2004). A random hierarchical lattice: the series-parallel graph and its properties. *Adv. Appl. Probab.* 36: 824–838.
6. Janson, S. (2004). Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stoch. Process. Appl.* 110: 177–245.
7. Mahmoud, H. (2013). Some node degree properties of series-parallel graphs evolving under a stochastic growth model. *Probab. Eng. Inf. Sci.* 27: 297–307.

https://doi.org/10.1017/S0269964814000138 Published online by Cambridge University Press