

Some Node Degree Properties of Series-Parallel Graphs Evolving Under a Stochastic Growth Model

Hosam M. Mahmoud
The George Washington University
Washington, DC 20052, USA

24th International Meeting on Probabilistic,
Combinatorial and Asymptotic Methods for the
Analysis of Algorithms

Menorca, Spain

May 31, 2013

Series-parallel (SP) graphs

Series-parallel (SP) graphs are network models that can represent the flow of, for example, commercial goods from a source to a market.

Consider them directed and the smallest is K_2 .



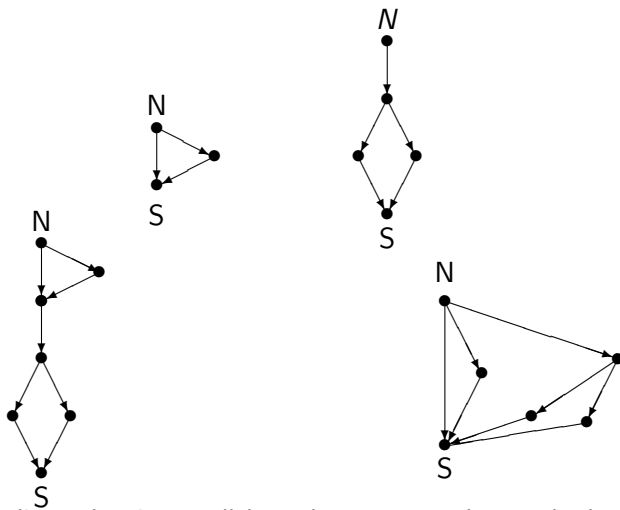


Figure: Two directed series-parallel graphs on top, and a graph obtained by a series composition (bottom left), and another graph obtained by their parallel composition (bottom right).

Models of randomness

To the best of our knowledge, these networks have been studied under two models of randomness:

- ▶ the uniform model, where all SP networks of a certain size are equally likely

Bernasconi, N., Panagiotou, K. and Steger, A. (2008).

Drmotá, M., Giménez, O. and Noy, M. (2010).

Models of randomness

To the best of our knowledge, these networks have been studied under two models of randomness:

- ▶ the uniform model, where all SP networks of a certain size are equally likely

Bernasconi, N., Panagiotou, K. and Steger, A. (2008).

Drmotá, M., Giménez, O. and Noy, M. (2010).

- ▶ The hierarchical lattice model Hambly, B. and Jordan, J. (2004).

Hierarchical lattice model

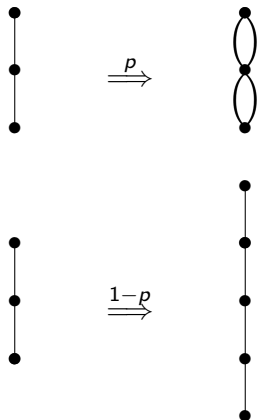


Figure: Hierarchical lattice model, growing too quickly.

An incremental model of randomness

Starting with a directed K_2 (with its sole edge directed from the North to the South Pole), we grow an SP graph in steps. At each step, we choose an edge from the existing graph at random (all edges being equally likely). We subject that edge to either a series extension with probability p , or a parallel doubling with probability $q := 1 - p$.

Edge extension and doubling

► series extension:



\xRightarrow{p}



► Parallel doubling:



$\xRightarrow{1-p}$



Henceforth “random” will always mean this model.

Degrees of certain nodes

We look at certain node degrees in a random SP graph:

- ▶ The degree of a pole. We find its exact distribution, given by a probability formula with alternating signs.
- ▶ For a fixed number s , we study the numbers of nodes of outdegree $1, \dots, s$: asymptotically have a joint multivariate normal distribution.

Pólya urns will systematically provide a working tool.

The degree of a pole

The number of edges coming out of the North Pole is a measure of the volume of trading and the amount of goods that can be shipped out of the source. This number is the North Pole's outdegree (it is also its degree).

Polya urn

Suppose we colored the edges coming out of the North Pole with **white (W)**, and all the other edges with **blue (B)**. We think of the edges as balls in a Pólya urn. Let W_n be the number of white balls in the urn after n edge additions to K_2 . As we start from a directed K_2 , we have $W_0 = 1$, and $B_0 = 0$.

Picking White

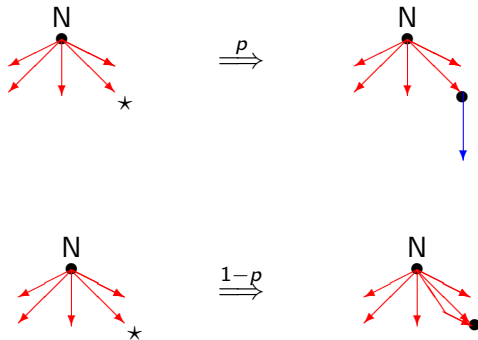


Figure: Picking a white edge (ball).

Picking Blue

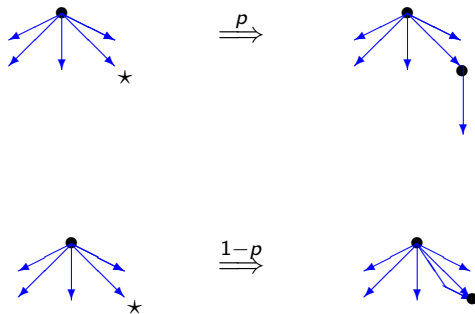


Figure: Picking a blue edge (ball).

Dynamics of the Pólya urn

The dynamics of a two-color Pólya urn scheme are often represented with a replacement matrix, the rows and columns of which are indexed with the two colors, and the entries are the number of balls added. The replacement matrix associated with our urn is

$$\begin{pmatrix} 1 - \text{Ber}(p) & \text{Ber}(p) \\ 0 & 1 \end{pmatrix}$$

Cécile Mailler told us yesterday that triangular arrays have a special behavior

Janson, S. (2005), for arrays with fixed entries.

Analytic tool

It is shown in

Morcrette, B. and Mahmoud, H. (2012), (AofA 2012).

how to get an exact distribution by solving a certain parametric pair of differential equations underlying an urn in $x(t)$ and $y(t)$ for an urn of this type. If $X(t, x(0))$ and $Y(t, y(0))$ are the solution, then $X^{W_0}(t, (x(0)))Y^{B_0}(t, y(0))$ is a history generating function. Specialized to our case, the differential equations are

$$\begin{aligned}x'(t) &= px(t)y(t) + qx^2(t), \\y'(t) &= y^2(t).\end{aligned}$$

Extends work in

Flajolet, P., Dumas, P. and Puyhaubert, V. (2006)

to the case of random entries (via history operators)

Solution

We solve this system under the initial condition $x(0) = u$, and $y(0) = v$, and get

$$x(t) = \frac{uv}{u - uv t - (u - v)(1 - vt)^p},$$
$$y(t) = \frac{v}{1 - vt}.$$

These solutions give rise to the history generating function

$$\sum_{0 \leq w, b, n < \infty} \mathbf{Prob}(W_n = w, B_n = b) u^w v^b z^n$$
$$= \left(\frac{uv}{u - uvz - (u - v)(1 - vz)^p} \right)^{W_0}$$
$$\times \left(\frac{v}{1 - vt} \right)^{B_0}.$$

Solution

Recall that $W_0 = 1$, and $B_0 = 0$. By setting $v = 1$, we get

$$\sum_{n=0}^{\infty} \sum_{w=0}^{\infty} \mathbf{Prob}(W_n = w) u^w z^n = \frac{u}{u - uz - (u - 1)(1 - z)^p}.$$

Theorem

Let W_n be the outdegree (indegree) of the North (South) Pole in a random series-parallel graph. Then, it has the exact probability distribution

$$\mathbf{Prob}(W_n = w) = \sum_{k=0}^{w-1} (-1)^{n+k} \binom{qn - p}{n} \binom{w-1}{k}.$$

Alternating signs

$$\mathbf{Prob}(W_n = w) = \sum_{k=0}^{w-1} (-1)^{n+k} \binom{qk - p}{n} \binom{w-1}{k}.$$

Multivariate analytic combinatorics (Pemantle and Wilson)?

Remark: Probability formulas with alternating signs are remarkable and not always intuitive. There are quite a few of them that appear in similar contexts, such as the classic occupancy problem (a classic work of de Moivre). After all, probability is nonnegative, and somehow cancellations in the formula with alternating signs always occur in a way to produce a nonnegative answer.

Mean

Proposition

Let W_n be the outdegree of the North Pole in a random series-parallel graph. We then have

$$\mathbf{E}[W_n] = \frac{(n+q)(n-1+q)\dots(1+q)}{n!} \sim \frac{1}{\Gamma(q+1)} n^q.$$

Differentiate the probability generating function once with respect to u , and set $u = 1$ to get a generating function of averages

$$\sum_{n=0}^{\infty} \mathbf{E}[W_n] z^n = (1-z)^{p-2}.$$

Extract the n th coefficient.

By symmetry, the South Pole has the same indegree distribution.

Nodes of small degree

The outdegree and indegree of a node in a trading network are indications of the local importance of a trading center to its neighbors. They determine how many neighbors will be affected, if the node becomes dysfunctional.

Indegrees and outdegrees

The indegrees are symmetrical to the outdegrees, for we can imagine the polarity of the graph reversed (the sense of edge orientation leads away from the South Pole), and the indegrees with the old polarity will become outdegrees in the reversed graph. Therefore, it is sufficient to study the outdegrees of the SP graph under the original orientation.

The Pólya urn for Outdegrees

We examine the distribution of the number of nodes of outdegree up to some fixed number, say s :

- ▶ Utilize $s + 1$ colors to code the outdegrees.
- ▶ Color each edge out of a node of outdegree i with color $i = 1, \dots, s$
- ▶ Color $s + 1$ is *special*: we color all the other edges with color $s + 1$; these edges are pointing away from nodes of outdegree $s + 1$ or higher.

The associated Pólya urn

Again, think of the edges as balls in a Pólya urn. This urn evolves in the following way. If at stage n we pick an edge of a nonspecial color i (pointing away from a node of outdegree i), we either extend it (with probability p) into a path of two edges directed away from the North Pole, or double it (with probability q), and a new edge pointing out of the Northern end node is added. In the case of extending the chosen edge, we do not change the outdegree of the Northern end of the edge being extended; we only add a new node of outdegree 1 (a new edge of color 1). In the case of doubling, we change the degree of the Northern end of the edge being doubled—it is increased by 1. Thus, we remove i edges of color i , and add $i + 1$ edges of color $i + 1$. When we pick a special edge, we either increase the outdegree of its northern end, or keep it the same. If the operation is an extension, the number of special edges does not change, but we add one node of outdegree 1 (we add an edge of color 1). If the operation is the doubling of the special edge, the outdegree of the node at the Northern end of the edge goes up by 1 (we add an edge with the special color).

The replacement matrix

The replacement matrix associated with our urn is

$$\begin{pmatrix} 2B - 1 & 2(1 - B) & 0 & \cdots & 0 & 0 \\ B & -2(1 - B) & 3(1 - B) & \cdots & 0 & 0 \\ B & 0 & -3(1 - B) & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ B & 0 & 0 & \cdots & -s(1 - B) & (s + 1)(1 - B) \\ B & 0 & 0 & \cdots & 0 & 1 - B \end{pmatrix}$$

It is a balance urn

Note that the sum across any row of the replacement matrix is 1. Pólya urn schemes satisfying such a condition are called *balanced*.

They enjoy the property that, regardless of the stochastic path followed, τ_n , the total number of balls in the urn after n draws, is deterministic; in our case it is

$$\tau_n = n + 1.$$

Very instrumental

Use of Pólya urn

Let $X_n^{(r)}$ be the number of edges in the SP graph of color r after the random insertion of n edges, and let \mathbf{X}_n be the vector with the $s + 1$ components $X_n^{(1)}, X_n^{(2)}, \dots, X_n^{(s+1)}$. Strong limit laws and asymptotic distributions are known for this type of balanced urn (where all the rows add up to the same constant, which is 1 in our case).

Mean

To deal with the exact mean and covariances, we derive the recurrence equations from the dynamics of the construction. Let \mathcal{F}_n be the sigma field generated by the the first n edge insertions. Let $I_n^{(r)}$ be the indicator of the event that an edge of color r is picked at the n th draw. For color 1, we write the conditional recurrence

$$\begin{aligned}\mathbf{E}[X_n^{(1)} | \mathcal{F}_{n-1}] &= X_{n-1}^{(1)} + \mathbf{E}[(2B - 1)I_n^{(1)} | \mathcal{F}_{n-1}] \\ &\quad + \mathbf{E}[BI_n^{(2)} | \mathcal{F}_{n-1}] \\ &\quad \vdots \\ &\quad + \mathbf{E}[BI_n^{(s+1)} | \mathcal{F}_{n-1}].\end{aligned}$$

Noting the independence of B and \mathcal{F}_{n-1} , we write the latter equation as:

$$\begin{aligned}\mathbf{E}[X_n^{(1)} | \mathcal{F}_{n-1}] &= X_{n-1}^{(1)} + (2p - 1) \mathbf{E}[I_n^{(1)} | \mathcal{F}_{n-1}] \\ &\quad + p \mathbf{E}[I_n^{(2)} | \mathcal{F}_{n-1}] + \cdots + p \mathbf{E}[I_n^{(s+1)} | \mathcal{F}_{n-1}].\end{aligned}$$

Matrix representation

The indicator $I_n^{(r)}$ is a Bernoulli random variable $\text{Ber}(X_{n-1}/\tau_{n-1})$ that conditionally (given \mathcal{F}_{n-1}) has the expectation $X_{n-1}^{(r)}/\tau_{n-1}$. The conditional expectation for the first color then takes the form

$$\begin{aligned}\mathbf{E}[X_n^{(1)} | \mathcal{F}_{n-1}] &= X_{n-1}^{(1)} + (2p - 1) \frac{X_{n-1}^{(1)}}{n} \\ &\quad + p \frac{X_{n-1}^{(2)}}{n} + \dots + p \frac{X_{n-1}^{(s+1)}}{n}.\end{aligned}$$

Note that the coefficients of the random variables come down spanning the entries of the average of the first column of the replacement matrix.

Writing a similar equation for each color, and putting them in matrix form, we get

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

where \mathbf{I} is the $(s+1) \times (s+1)$ identity matrix, and \mathbf{A}^T is the transpose of \mathbf{A} .

Solution via eigenvalues

We can take expectations and write

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

This form can be iterated, and we get

$$\mathbf{E}[\mathbf{X}_n] = \mathbf{R}_n \mathbf{R}_{n-1} \dots \mathbf{R}_1 \mathbf{E}[\mathbf{X}_0]. \quad (1)$$

Observe that the eigenvalues of $\mathbf{E}[\mathbf{A}]$ are

$$\lambda_1 = 1, \quad \text{and} \quad \lambda_r = -(r-1)q, \quad \text{for } r = 2, \dots, s+1.$$

The eigenvalues are real and distinct, with $\lambda_2 = -q < \frac{1}{2} = \frac{1}{2}\lambda_1$. As the eigenvalues are distinct, they give rise to simple Jordan normal forms—the matrix \mathbf{R}_j can be written as where \mathbf{M} is the modal matrix¹ of $\mathbf{E}[\mathbf{A}^T]$, which is invertible, because the eigenvalues are distinct.

Matrix recurrence

The matrix equation can now be simplified to

$$\begin{aligned} \mathbf{E}[\mathbf{X}_n] &= (\mathbf{M}\mathbf{D}_n\mathbf{M}^{-1})(\mathbf{M}\mathbf{D}_{n-1}\mathbf{M}^{-1})\cdots(\mathbf{M}\mathbf{D}_1\mathbf{M}^{-1})\mathbf{E}[\mathbf{X}_0] \\ &= \mathbf{M}\mathbf{D}_n\mathbf{D}_{n-1}\cdots\mathbf{D}_1\mathbf{M}^{-1} \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \end{aligned} \quad (2)$$

Mean

We thus have the exact vector of means:

$$\mathbf{E}[\mathbf{X}_n] = \frac{1}{n!} \mathbf{M} \begin{pmatrix} \Gamma(n+1) & 0 & 0 & \dots & 0 & 0 \\ 0 & \frac{\Gamma(n+1-q)}{\Gamma(1-q)} & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & \frac{\Gamma(n+1-sq)}{\Gamma(1-sq)} \end{pmatrix} \mathbf{M}^{-1} \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

We illustrate this program with the small instance $s = 2$.

Illustration for $s = 2$

Theorem

Let $Y_n^{(r)}$ be the number of nodes of outdegree $r \in \{1, 2\}$ in a random directed series-parallel graph, and let \mathbf{Y}_n be the vector with these two components. We have

$$\begin{aligned}\mathbf{E}[Y_n^{(1)}] &= \frac{\rho(n+1)}{q+1} + \frac{2q\Gamma(n+p)}{(q+1)\Gamma(p)\Gamma(n+1)}, \\ \mathbf{E}[Y_n^{(2)}] &= \frac{\rho q(n+1)}{(2q+1)(q+1)} + \frac{4^p q \Gamma(p - \frac{1}{2}) \Gamma(n+p)}{2\sqrt{\pi}(q+1)\Gamma(-1+2p)\Gamma(n+1)} \\ &\quad - \frac{3q\Gamma(n-1+2p)}{(2q+1)\Gamma(-1+2p)\Gamma(n+1)}.\end{aligned}$$

Bivariate central limit theorem

Also, \mathbf{Y}_n converges in distribution to a bivariate normal vector:

$$\frac{\mathbf{Y}_n - \begin{pmatrix} \frac{p}{q+1} \\ \frac{pq}{(2q+1)(q+1)} \end{pmatrix} n}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(\mathbf{0}, \begin{pmatrix} \frac{2pq(3-p)}{(2-p)^2(3-2p)} & -\frac{2p^2q}{(4-3p)(3-2p)(2-p)^2} \\ -\frac{2p^2q}{(4-3p)(3-2p)(2-p)^2} & \frac{pq(24p^4-157p^3+356p^2-342p+120)}{(5-4p)(4-3p)(3-2p)^2(2-p)^2} \end{pmatrix}\right)$$

Proof Sketch

Note that $Y_n^{(1)} = X_n^{(1)}$, and $Y_n^{(2)} = \frac{1}{2}X_n^{(2)}$. Therefore it suffices to study $Y_n^{(1)} = X_n^{(1)}$ to get the results for $X_n^{(1)}$ and $X_n^{(2)}$.

For exact second moment of $X_n^{(1)}$, we start with a recurrence obtained from the 3×3 replacement matrix:

$$X_n^{(1)} = X_{n-1}^{(1)} + B - (1 - B)\text{Ber}\left(\frac{X_{n-1}^{(1)}}{n}\right). \quad (3)$$

Squaring both sides, we get

$$\begin{aligned} (X_n^{(1)})^2 &= (X_{n-1}^{(1)})^2 + B + (1 - B)\text{Ber}\left(\frac{X_{n-1}^{(1)}}{n}\right) + 2BX_{n-1}^{(1)} \\ &\quad - 2(1 - B)X_n^{(1)}\text{Ber}\left(\frac{X_{n-1}^{(1)}}{n}\right). \end{aligned}$$

Recurrence for second moment of

$$X_n^{(1)}$$

So, the conditional second moment for this color is

$$\mathbf{E}[(X_n^{(1)})^2 | \mathcal{F}_{n-1}] = (X_{n-1}^{(1)})^2 + p + q \frac{X_{n-1}^{(1)}}{n} + 2pX_{n-1}^{(1)} - 2qX_{n-1}^{(1)} \frac{X_{n-1}^{(1)}}{n}.$$

This gives a recurrence for the (unconditional) second moment:

$$\mathbf{E}[(X_n^{(1)})^2] = \left(1 - \frac{2q}{n}\right) \mathbf{E}[(X_{n-1}^{(1)})^2] + \left(2p + \frac{q}{n}\right) \mathbf{E}[X_{n-1}^{(1)}] + p.$$

Plug in $\mathbf{E}[X_n^{(1)}]$, which we have developed. This recurrence, and several other in the sequel, are of the general form

$$a_n = \left(1 - \frac{b}{n}\right) a_{n-1} + h(n), \quad (4)$$

for constant b and known asymptotically linear function $h(n)$, with an asymptotically quadratic solution.

Exact second moment of $X_n^{(1)}$

The solution to the recurrence for $\mathbf{E}[(X_n^{(1)})^2]$ is

$$\begin{aligned} \mathbf{E}[(X_n^{(1)})^2] &= \frac{p(q^2 + 3q + qp + 2qpn + 2 + pn - p)}{(2-p)(1+2q)(1+q)} (n+1) \\ &+ (-p^4 + 2p^2 + 7p^2q^2 - 2qp - 4q^3p \\ &\quad - 5q^2p - 4q^2 + 4q^4 \\ &\quad + 2q^3 - p - 2q + 3p^2q) \Gamma(n+1-2q) \\ &\times ((1+q)(1+2q)(p+2q)(-1+p+2q) \\ &\quad \times \Gamma(n+1)\Gamma(1-2q))^{-1} \\ &+ (2(2p^2n - 2pn + 4qpn + 5qp + 2q^2)) \Gamma(n+p) \\ &\times ((-2+p)(p+2q)(-1+p+2q) \\ &\quad \times \Gamma(n+1)\Gamma(-1+p))^{-1}. \end{aligned}$$

After subtracting off the square of the mean:

$$\mathbf{Var}[X_n^{(1)}] \sim \frac{2pq(3-p)}{(2-p)^2(3-2p)} n.$$

Exact second moment of $X_n^{(2)}$

For the second moment of $X_n^{(2)}$ and the covariance, we only sketch the key steps. We start from a stochastic recurrence (again obtained from the dynamics of the construction):

$$X_n^{(2)} = X_{n-1}^{(2)} + 2(1 - B) \left[\text{Ber} \left(\frac{X_{n-1}^{(1)}}{n} \right) - \text{Ber} \left(\frac{X_{n-1}^{(2)}}{n} \right) \right]. \quad (5)$$

Multiply (3) and (5), and take expectation (handling the Bernoulli random variables via a double expectation).

This gives an exact recurrence for the mixed moment $\mathbf{E}[X_n^{(1)} X_n^{(2)}]$. This recurrence involves $\mathbf{E}[(X_n^{(1)})^2]$, which we already have. Thus, the recurrence is in the form of (4). We solve the recurrence and obtain the exact mixed moment $\mathbf{E}[X_n^{(1)} X_n^{(2)}]$. Extracting leading asymptotics, we get a linear covariance equivalence, as ($n \rightarrow \infty$):

$$\text{Cov}[X_n^{(1)}, X_n^{(2)}] \sim -\frac{4p^2 q}{(4 - 3p)(3 - 2p)(2 - p)^2} n.$$

Finally, square (5), and take expectations. The resulting recurrence has the expectations of $X_n^{(1)}$ and $X_n^{(2)}$, as well as the expectation of their product. We already have all these ingredients in exact form. We plug in the results we have and solve the recurrence (also in the form of (4)) to get $\mathbf{E}[(X_n^{(2)})^2]$. Subtracting off the square of $\mathbf{E}[X_n^{(2)}]$, we get an exact variance. The formula is so huge to be listed, and we only give its linear asymptotic equivalent:

$$\mathbf{Var}[X_n^{(2)}] \sim \frac{4pq(24p^4 - 157p^3 + 356p^2 - 342p + 120)}{(5 - 4p)(4 - 3p)(3 - 2p)^2(2 - p)^2} n. \quad \square$$

Bootstrapping

For higher s , the variances and covariances are significantly more computationally intensive. Nevertheless, the steps are clear. It is a bootstrapped program in the fashion of dynamic programming: Obtain all the results up to color $r - 1$ (in addition to all first moments, obtain all the mixed moments $\mathbf{E}[X^{(i)}X^{(j)}]$, for $i, j = 1, \dots, r - 1$.) Now, write a recurrence for $\mathbf{E}[X_n^{(1)}X_n^{(r)}]$. From known results on urns, we also have the strong law

$$\frac{X_n^{(r)}}{n} \xrightarrow{\text{a.s.}} \frac{r! p q^{r-1}}{(r q + 1)(r - 1) q + 1) \dots (q + 1)}.$$

Thank you