Reinforcement and Preferential attachment models via Pólya Urns

by

Somya Singh

A thesis submitted to the
Department of Mathematics and Statistics
in conformity with the requirements for
the degree of Doctor of Philosophy

Queen's University
Kingston, Ontario, Canada
August 2023

Abstract

In this thesis, we devise two different types of discrete-time stochastic models using modified Pólya urn schemes. The first set of models concerns interacting contagion networks constructed using two-color (red and black) finite memory Pólya urns in which reinforcing balls are removed M time steps after being added (where M is the "memory" of the urn). The urns interact in the sense that the probability of drawing a red ball (which represents an infectious state or an opinion) for a given urn, not only depends on the ratio of red balls in that urn but also on the ratio of red balls in other urns in a network representing the interconnections, hence accounting for the effect of spatial contagion. The finite memory reinforcement provides a diminishing effect of past draws which represents curing of an infection in an epidemic spread model, or lessening influence of a popular opinion in a social network. We examine the stochastic properties of the underlying Markov draw process and construct a class of dynamical systems to approximate the asymptotic marginal distributions. We also design a consensus achieving connected network of agents via two-color finite memory Pólya urns. The interaction between urns is time-varying and is represented via "super-urns" which combine for each node its own urn with its neighbouring urns. We obtain the consensus value in terms of the network's reinforcement parameters and memory.

In the second part of this thesis, we introduce a novel preferential attachment model using the draw variables of a modified Pólya urn with an expanding number of colors, notably capable of modeling influential opinions (in terms of vertices of high degree) as the graph evolves. Unlike the Barabási-Albert model, the color-coded vertices in conjunction with the time-varying reinforcing parameter in our model allows for the vertices added (born) later in the process to potentially attain a high degree in a way that is not captured by the former. We study the degree count of the vertices in the graphs generated via our model by analyzing the draw vectors of the underlying stochastic process. Furthermore, we compare our model with the Barabási-Albert model via simulations.

Statement of Originality

The following work is my own and I hereby certify that the intellectual content of this thesis is the product of my own work. All references and contributions of other individuals has been cited and sourced appropriately.

Acknowledgements

I deeply thank my advisors, Prof. Fady Alajaji and Prof. Bahman Gharesifard for their guidance, patience, and support. I learned a lot under their supervision. I thank Prof. Yanglei Song and Prof. Felicia Magpantay for their insights and ideas on my research problems. I thank Prof. Philip Paré and Prof. Elaheh Feta for agreeing to be external examiners for my defense. I would also like to thank Prof. Thomas Barthelmé for agreeing to be head delegate for my defence.

My fellow graduate students have been very supportive throughout my PhD journey. A big thanks to all of them: Annika, Sonja, Arpita, Tariq, Henry, Ben, Kaly, Sina, Deepanshu, Sasha, Bora, Neil, Skye, Daniel, Mohammad (sorry if I forgot to add someone to the list!). Apart from my workplace, I am very fortunate to have a very caring and supportive group of friends who were there for me through the thick and thins. Thank you Vasu, Basu, Priyanka, HL, HT, and Imran for always being there for me. Leaving Kingston and not seeing my friends regulary would be difficult for me. Thank you Kaivalya and AJ for regularly talking to me through Google Meet and giving me strength during my thesis writing process.

My deepest gratitude to my entire family for being my pillar of strength and always believing in me. I am grateful to have a very big close-knit family. I have received

loads of love and support from my grandparents (Nanaji, Naniji, Baba, Dadima), aunts and uncles (Ajay fufaji, Anil fufaji, Rekha bua, Usha bua, Mausi, Mausaji, Mama, Mami, Chachu, Chachima) and cousins (Hunny di, Ashu, Shikhar, Aayush, Sakshi, Nikki, Vidit and Ricky).

Finally, I dedicate this thesis and the last part of my acknowledgements to the most important people of my life. Mummy and Daddy have worked really hard and made a lot of sacrifices to make me the person I am today. I learn something new and inspirational on every phone-call with my mom and a big salute to her for keeping me and my family intact after my father's demise.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation and Literature Review

Interacting urn networks are widely used in the field of applied mathematics, biology and computer science to model spread of diseases [34, 35], consensus dynamics [26], image segmentation [12], propagation of computer viruses in connected devices [18], and social networks [58].

The first part of this thesis (Chapters 2 and 3) is concerned with networks of two-color Pólya urns. We are given a network of N urns, represented as a graph with the urns as vertices and interconnections as edges. At time $t = 0$, each urn is composed of some red and some black balls, where different urns can have different initial compositions, but no urn is empty. At each time instant t, a ball is chosen for each urn with probability depending on the composition of the urn itself and of the other urns in the network, and then additional (reinforcing) balls of the color just drawn are added to the urn. Letting $U_{i,t}$ denote the ratio of red balls in urn i at time

t, the draw variable $Z_{i,t}$, denoting the indicator function of a red ball chosen for urn i at time t, is governed by

$$Z_{i,t} = \begin{cases} 1 & \text{w.p.} \quad f(U_{1,t-1}, \cdots, U_{N,t-1}) \\ 0 & \text{w.p.} \quad 1 - f(U_{1,t-1}, \cdots, U_{N,t-1}) \end{cases} \tag{1.1}$$

where "w.p." stands for "with probability" and $f : \mathbb{R}^N \to (0, 1)$ is a real-valued function which accounts for interactions in the network of urns. The stochastic process $\{Z_{i,t}\}_{t=1}^{\infty}$ is commonly known as a reinforcement process generated by an urn model. Although a variety of reinforcement processes are used to develop interacting networks, Pólya urns are the most commonly used urn models (there are a few examples based on other interacting urn networks; e.g., see [38, 52] for interacting Friedman urn networks). In the first part of this thesis, we devise a contagion model through an interacting network of modified Pólya urns and study its stochastic properties. We further develop a consensus achieving network of agents via our urns.

Various other models have been proposed in the literature to portray contagion in networks using interacting Pólya processes. In [34], the concept of "super urn" for a network of Pólya urns is utilized to model spatial contagion, where at every time step and for each urn, a ball is drawn from its "super urn" formed by the collection of the urn's own balls and the ones of its neighbours. In [32, 35], optimal curing and initialization policies were investigated for the network contagion model of [34]. In [49], the authors introduce a symmetric reinforcement scheme for interacting Pólya urn

network with (1.1) given by:

$$Z_{i,t} = \begin{cases} 1 & \text{w.p.} \quad \frac{\alpha}{N} \sum_{i=1}^{N} U_{i,t-1} + (1-\alpha) U_{i,t-1} \\ 0 & \text{w.p.} \quad 1 - \frac{\alpha}{N} \sum_{i=1}^{N} U_{i,t-1} - (1-\alpha) U_{i,t-1} \end{cases}$$

where $\alpha \in [0,1]$ is a fixed number and N is the number of Pólya urns in the network. An example of a more complicated interaction network is given in [16], where a finite connected graph with each vertex equipped with a Pólya urn is considered and at any given time t, only one of the two interacting urns receive balls with probability proportional to a number of balls raised to some fixed power.

An important characteristic of most reinforcement processes generated via urn models is that they are non-Markovian in the sense that the composition of each urn at any given time affects its composition at every time instant thereafter. This property is not realistic when modelling the spread of infection as one should account for the possibility that infection is cured (or that the urn is removed, a possibility that we do not consider here). In this thesis, we consider an interacting Pólya urn network where each urn has a *finite memory*, denoted by $M \geq 1$, in the sense that, at the time instant $t > M$, the reinforcing balls added at time $t - M$ are removed from the urn and hence have no effect on future draws; see Figure 1.1 for an illustration of this finite memory reinforcement. This notion of a finite memory Pólya urn was introduced in [1] (in the context of a single urn modeling error bursts in communication channels) to account for the diminishing effect of past reinforcements on the urn process, which is a realistic assumption when modelling contagion in a population. The resulting network draw variables $\{Z_t\}_{t=1}^{\infty}$ of the Pólya urn with memory M forms a Markov

chain of order M, see [1].

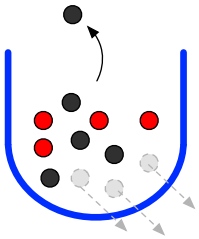


Figure 1.1: Finite memory Pólya urn.

There are other examples in the literature where a Markovian version of the Pólya process is studied. In [4], a rescaled Pólya urn model with randomly fluctuating conditional draw probability is considered. Another Markovian Pólya process is the Pólya-Lundberg process [48], which was recently adapted in [15] to measure the dynamics, among many other models such as [54], of the SARS-CoV-2 pandemic.

The techniques used in the analysis of a finite memory Pólya process are quite different from the ones used for any general random reinforcement process. Standard techniques used for the latter case include the method of moments [28], Martingale methods [31, 49], stochastic approximations [21, 29, 41] and the embedding of reinforcement processes in continuous-time branching processes [8, 9, 37]. A detailed discussion on these methods can be found in [5] and the survey [47].

Another interesting feature of Pólya urns is that their reinforcement schemes represent preferential attachment behavior in the sense that at each time step, we add balls of a particular color with a probability proportional to the number of balls of that color in the urn. This reinforcement of adding balls to the urn proportional to the ratio of balls is commonly known as the "rich gets richer" phenomenon which in the context of randomly growing graphs/networks refers to reinforcements that favor a high degree

vertex to stay influential. Preferential attachment graphs have been widely studied within the areas of statistical mechanics [33, 59], network science [25], probability theory [39, 50] and game theory [53]. One of the most popular models of a preferential attachment graph is the so-called Barabási-Albert model [14], which has since been modified in a variety of ways [6, 42, 61]. Various other models have been devised thereafter to generate preferential attachment graphs; for example in [17] the growth of the random graph is competition based. Given a graph at a certain time step, the new vertex attaches itself to an existing vertex which ends up minimizing a certain cost function. For a vertex, this cost function depends on its centrality and distance from the root ensuring that the vertices with higher degrees have lower cost functions. In [40], a continuous-time equation governing the number of vertices with degree k is formulated to study citation networks. In [27], a randomly growing graph algorithm that combines the features of a geometric random graph and a preferential attachment graph is analysed. In [22], properties of Wikipedia are studied by representing topics as vertices and hyperlinks between them as edges. Several preferential attachment hypergraphs (i.e., graphs in which an edge can join any number of vertices) generating models have also been devised in the literature [10, 30, 36].

Our main objective in Chapter 4 is to introduce a new preferential attachment graph generating algorithm using a modified Pólya urn model. Various versions of Pólya urns have been used to model preferential attachment graphs, for instance, in [24] a generalized Pólya urn process is used to devise a preferential graph generating algorithm (refer to [23, 45] for a detailed description of this generalized Pólya urn process). In [19], a preferential attachment type multi-graph (i.e., a graph that can have more than one edge between a pair of vertices) is constructed using different

variations of the Pólya urn process which was used to study the spread of viruses on the internet in [18]. More elaboration on the similarities between Pólya urns and preferential attachment graphs is given in the survey [47].

The Barabási-Albert model is well known to exhibit a power law distribution as the number of vertices becomes sufficiently large, given by $p(k) \sim k^{-3}$, where $p(k)$ is the probability of randomly selecting a vertex with degree k in the network. Despite the fact that this power law can be used to study various properties of the Barabási-Albert model such as the Hirsch index distribution and the clustering coefficient, see [2, 13, 20] for definitions, the likelihood of vertices gaining new edges is solely determined by their degree. This is not realistic, when modeling scenarios where newly added individuals are accompanied with impactful ideas that can lead to rapid or disruptive influence, regardless of their initially low degree.

Motivated by the above mentioned shortcomings of the Barabási-Albert network, we develop in Chapter 4 a preferential attachment graph generating algorithm in which each vertex is uniquely identified with a color of an expanding color Pólya urn. To address the concern about the degree being the only factor affecting the likelihood of vertices forming new connections in the Barabási-Albert algorithm, we introduce a time-varying reinforcement parameter in our model which enables the vertices to make more or fewer connections depending on their unique color association and the time at which they were introduced in the network, i.e., their birth time.

1.2 Contributions and Organization of the Thesis

The rest of the thesis is organized as follows. In Chapter 2, we design and investigate an interacting network of N two-color finite memory Pólya urns to model the spread of infection (commonly referred to as contagion) where each urn is associated to a node (e.g., "individual") in a general network (e.g., "population") to delineate its "immunity" level. Each Pólya urn in the network contains red and black balls which represent units of "infection" and "healthiness" respectively. The reinforcement of drawing a ball for each urn is mathematically formulated such that a weighted composition of other urns in the network affects the drawing process, hence capturing interaction between urns.

Unlike standard reinforcement processes, we are able to use Markovian properties in our analysis as our model of interacting urns with finite memory M yields an Mth order Markov draw process. However, one drawback of working with a memory-M Markov chain over a network is that the size of its underlying transition probability matrix grows exponentially with both M and the network size. To account for this problem, after having introduced our interacting Pólya urn network and investigated its properties in detail, we formulate a *dynamical system* to tractably approximate its asymptotic behaviour. To obtain this dynamical system, we make the assumption that for any given time $t > M$, the joint probability distribution of draw processes at times $t-1, \cdots, t-M$ for any urn is approximately equal to the product of its marginals. This type of approximation is referred to as a "mean-field approximation" and is commonly used in the literature on compartmental models, such as the well-known susceptible-infectious-susceptible (SIS) model [7, 11, 43, 44, 46, 60]. The key

factor that distinguishes our treatment is the latitude provided by the consideration of memory $M \geq 1$, in contrast to the SIS model which is based on a memory one ($M = 1$) Markov chain. In particular, as our simulations, which are performed for both non-homogeneous and homogeneous networks (of small-to-medium size), verify, the nonlinear dynamical system that we obtain approximates the true (underlying) Markov process with improved accuracy as we increase the memory order M. We also characterize the equilibrium point of this dynamical system, when the nonlinear dynamical system is approximated by its linear part (the latter approximation is exact for the case with memory $M = 1$). More specifically, we show that when $M = 1$, the (linear) dynamical system always admits a unique equilibrium (which can be exactly determined); while for $M > 1$, we note that the linearized dynamical system has a unique equilibrium when its governing (block) matrix has a spectral radius less than unity. In summary, our results provide a novel mathematical framework for the study of epidemics on networks in realistic scenarios where memory is a consideration. The results of Chapter 2 were published in part in [56] and [57].

In Chapter 3, we devise a consensus achieving network using a finite memory Pólya urn network which is somewhat similar to the one studied in Chapter 2, but different in terms of having time-varying interaction dynamics among the urns. Given a connected network $\mathcal{G}$, we equip each agent with an urn, initially consisting of balls of two colors, red and black. The "belief" of each urn (or individual) at a given time instant is the probability that a red ball is chosen in the drawing process. The draw process utilizes the spatial interconnections in $\mathcal{G}$ via the use of "super-urns" which are introduced in [12, 34, 35]. At each time instant and for each urn (or individual), a ball is drawn from its super-urn and returned back to the urn; then reinforcing balls of the color

just drawn are added to the urn for a limited period of M future time instants, where M denotes the memory of the urn. Additionally, and important for our objective of the network reaching consensus, as of time $t = M + 1$, we remove the balls which were present in the urns initially. The significance of removing initial conditions is that the individuals forget about their inherent beliefs as of time $t \geq M + 1$. We analyse our network of finite memory Pólya urns and show that for a memory M, the vector of draw variables forms a time-varying reducible Mth order Markov chain. We study the structure of this Markov process and show that our network achieves consensus. We also present an alternate method to show consensus for a homogeneous connected network. In this method, we obtain a class of linear dynamical systems with time delay which gives the probability of drawing a red ball from the super-urns at any time t. We then obtain the consensus value of this connected homogeneous network by studying the asymptotic properties of these delayed linear dynamical systems. The results of Chapter 3 were published in part in [55].

In Chapter 4, we construct randomly growing undirected graphs using the draw variables of a single modified Pólya urn, with an expanding number of colors. The Pólya urn process is modified in the sense that at each time instant, not only is a ball drawn and returned to the urn along with a number (potentially time-varying) of reinforcing balls of the same color, but another ball of a new color is also added to the urn. This new color corresponds to a new vertex which is added to the graph at this time instant. More specifically, the network is generated by associating each incoming vertex to the new color ball added after each draw and by attaching it to the existing vertex represented by the draw color. The number of colors in the urn grows without bound with the number of draws, and the generated network has a

preferential attachment property as the vertices corresponding to dominant colors (i.e., colors in the urn with a large number of balls) are more likely to attract newly formed vertices as their neighbors. The resulting preferential attachment growing graph is thus constructed via a Pólya urn; this enables us to characterize the degree count of individual vertices in the network through the draw variables of their corresponding colors giving each vertex a unique identity which is absent in the Barabási-Albert model. Indeed, the draw variables of the Pólya urn capture the entire structure of the graph generated and hence it is enough to study the behaviour of these draw variables to understand the properties of the graph. Moreover, we use an extra time-varying parameter to set the number of balls (not necessarily an integer) added to the Pólya urn to reinforce the color of the drawn ball at each time instant. The time-varying nature of this parameter allows us for any given vertex to tweak the likelihood of amplifying or dampening its degree growth depending on the time at which it was introduced in the network; this feature can be used to regulate the dominance (in terms of gaining edges) of high degree vertices over low degree ones in the generated preferential attachment graph. Therefore, unlike the Barabási-Albert algorithm, our model can be used to generate random networks with a variety of degree distributions other than power law distributions.

Finally, conclusions and future directions are presented in Chapter 5.

Chapter 2

Interacting Finite Memory Pólya Networks

2.1 The Model

We consider a network of N finite memory two-color Pólya urns, where each urn can be associated to a node in an arbitrary network. At time $t = 0$, urn i contains R_i red balls and B_i black balls, $i = 1, \ldots, N$. We denote the total number of balls in the ith urn at time $t = 0$ by $T_i = R_i + B_i$, and we assume that there is at least one red and one black ball in each urn at time $t = 0$, i.e., $R_i > 0$ and $B_i > 0$ for all i. We also let $U_{i,t}$ denote the ratio of red balls in urn i at time t, with its initial value (at time $t = 0$) given by $U_{i,0} = R_i/T_i$.

We next define the reinforcement scheme, in the form of *draw variables*, $Z_{i,t}$, associated with urn i at time $t \geq 1$, for our proposed interacting Pólya contagion

network:

$$Z_{i,t} = \begin{cases} 1 & \text{if a red ball is drawn for urn } i \text{ at time } t \\ 0 & \text{if a black ball is drawn for urn } i \text{ at time } t \end{cases} \tag{2.1}$$

where the *process of drawing a ball for urn i* is governed by a particular function of the form (1.1), which is applied *simultaneously* to all urns. We denote the drawing random vector at time t by $Z_t = (Z_{1,t}, Z_{2,t}, \cdots, Z_{N,t})$. If a red ball (respectively, a black ball) is drawn for urn i, we add $\Delta_{r,i}(t)$ red balls (respectively, $\Delta_{b,i}(t)$ black balls) to urn i. This scheme, which we refer to as the *urn scheme*, is often captured by a matrix of the form:

$$\begin{bmatrix} \Delta_{r,i}(t) & 0 \\ 0 & \Delta_{b,i}(t) \end{bmatrix}. \tag{2.2}$$

We assume throughout that $\Delta_{r,i}(t) > 0$ and $\Delta_{b,i}(t) > 0$ for all $i \in \{1, \cdots, N\}$, $t \geq 1$ and that all the urns in the network have same memory (say M), where memory of an urn is defined in Chapter 1 (see also Figure 2.1).

We start with defining an *interaction matrix* S to be an $N \times N$ row-stochastic matrix with non-negative entries, i.e., each row in S sums to one. Entries of the interaction matrix S are denoted by s_{ij}, where $i, j \in \{1, \ldots, N\}$. The interaction matrix S can also be thought of as a weighted adjacency matrix of a directed graph with each vertex equipped with a memory M Pólya urn.

Having defined the interaction matrix, we can now explicitly specify the function f used in the drawing mechanism (2.1). In particular, we set the probability of choosing

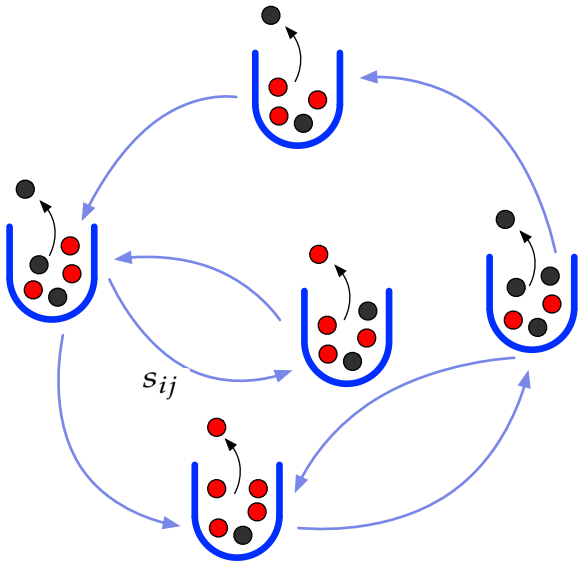


Figure 2.1: The interaction matrix S can also be thought of as a weighted adjacency matrix of a directed graph with each vertex equipped with a Pólya urn.

a red ball from urn i at time t as follows:

$$Z_{i,t} = \begin{cases} 1 & w.p. \quad \sum_{j=1}^{N} s_{ij} U_{j,t-1} \\ 0 & w.p. \quad 1 - \sum_{j=1}^{N} s_{ij} U_{j,t-1}. \end{cases} \tag{2.3}$$

Furthermore, as all draws occur simultaneously as illustrated in Figure 2.1, the draw variables $Z_{i,t}$ and $Z_{i',t}$ are conditionally independent given all past draws in the network, for any $i \neq i'$; hence at any time t,

$$P\left(Z_{1,t}, \ldots, Z_{N,t} | \{Z_{1,k}\}_{k=1}^{t-1}, \ldots, \{Z_{N,k}\}_{k=1}^{t-1}\right) = \prod_{i=1}^{N} P\left(Z_{i,t} | \{Z_{1,k}\}_{k=1}^{t-1}, \ldots, \{Z_{N,k}\}_{k=1}^{t-1}\right). \tag{2.4}$$

Since each urn in the network has memory M, for each time instant $t \geq M+1$, the

ratio of red balls in urn i is given by

$$U_{i,t} = \frac{R_i + \sum_{k=t-M+1}^{t} \Delta_{r,i}(k) Z_{i,k}}{T_i + \sum_{n=t-M+1}^{t} (\Delta_{r,i}(n) Z_{i,n} + \Delta_{b,i}(n)(1 - Z_{i,n}))} \tag{2.5}$$

almost surely.[1]For ease of notation, we define the following (normalized) initial and reinforcement network parameters:

$$\rho_i = \frac{R_i}{T_i}, \qquad \sigma_i = 1 - \frac{R_i}{T_i}, \qquad \delta_{r,i}(t) = \frac{\Delta_{r,i}(t)}{T_i}, \quad \delta_{b,i}(t) = \frac{\Delta_{b,i}(t)}{T_i}. \tag{2.6}$$

Substituting (2.6) in (2.5), we obtain

$$U_{i,t} = \frac{\rho_i + \sum\limits_{k=t-M+1}^{t} \delta_{r,i}(k) Z_{i,k}}{1 + \sum\limits_{n=t-M+1}^{t} (\delta_{r,i}(n) Z_{i,n} + \delta_{b,i}(n)(1 - Z_{i,n}))} \tag{2.7}$$

almost surely. We denote this contagion type network of finite memory Pólya urns with fixed interaction matrix S by IPCN(M, N), where M denotes the memory of urns in the network and N denotes the number of urns in the network. In the next two sections, we discuss the analytic properties of this IPCN(M, N) system.

2.2 Markovian property of IPCN$(\mathbf{M}, \mathbf{N})$

We now establish the Markov property of the network draw process $\{Z_t\}_{t=1}^{\infty}$ in the following proposition:

[1]All identities involving random variables or vectors are (implicitly) understood to hold almost surely.

Proposition 2.1. *For an* IPCN(M, N) *system, the stochastic process given by* $\{Z_t\}_{t=1}^{\infty}$ *is a time-varying Markov chain of order* M*.*

Proof. Let $a_t = (a_{1,t}, \cdots, a_{N,t}) \in \{0,1\}^N$. Using (2.3) and by virtue of the conditional independence of the draw variables $Z_{i,t}$ and $Z_{i',t}$ given all past draws in the network for all $i \neq i'$, we have for $t \geq M$ that

$$\begin{aligned} P(Z_{t+1} =& a_{t+1} | Z_t = a_t, \cdots, Z_1 = a_1) \\ &= \prod_{i=1}^{N} P(Z_{i,t+1} = a_{i,t+1} | Z_t = a_t, \cdots, Z_1 = a_1) \\ &= \prod_{i=1}^{N} \left(a_{i,t+1} \sum_{j=1}^{N} s_{ij} U_{j,t} + (1 - a_{i,t+1})(1 - \sum_{j=1}^{N} s_{ij} U_{j,t}) \right). \end{aligned}$$

As a result, we have that

$$\begin{aligned} &P(Z_{t+1} = a_{t+1} | Z_t = a_t, \cdots, Z_1 = a_1) \\ &= \prod_{i=1}^{N} \left((2a_{i,t+1} - 1) \sum_{j=1}^{N} \frac{s_{ij} \Big(\rho_j + \sum_{k=t-M+1}^{t} \delta_{r,j}(k) a_{j,k} \Big)}{1 + \sum_{n=t-M+1}^{t} (\delta_{r,j}(n) a_{j,n} + \delta_{b,j}(n)(1 - a_{j,n}))} + (1 - a_{i,t+1}) \right) \\ &= P[Z_{t+1} = a_{t+1} | Z_t = a_t, \cdots, Z_{t-M+1} = a_{t-M+1}]. \end{aligned} \tag{2.8}$$

Hence the process $\{Z_t\}_{t=1}^{\infty}$ is a time-varying Mth order Markov chain. □

In order to write the transition probabilities of this Markov chain, we define $\mathbf{W}_t := \{Z_t, Z_{t+1}, \cdots, Z_{t+M-1}\}$. Since $\{Z_t\}_{t=1}^{\infty}$ is an Mth order Markov chain, the order of $\{\mathbf{W}_t\}_{t=1}^{\infty}$ is 1. We use (2.8) to give a general formula for the entries of the transition probability matrix $Q^{(M,N)}$ of the Markov process $\{\mathbf{W}_t\}_{t=1}^{\infty}$, which has 2^{MN}

states. Note that the transition probability, $q_{ab}^{(M,N)}$, of going from state

$$a = ((a_{11}, a_{21}, \cdots, a_{N1}), \cdots, (a_{1M}, a_{2M}, \cdots, a_{NM}))$$

to state

$$b = ((b_{11}, b_{21}, \cdots, b_{N1}), \cdots, (b_{1M}, b_{2M}, \cdots, b_{NM}))$$

in one time step is nonzero if and only if $a_{ij} = b_{i(j-1)}$ for $i \in \{1, \cdots, N\}$ and $j \in \{2, \cdots, M\}$, where a, b are binary NM tuples. If $q_{ab}^{(M,N)}$ is nonzero, it is given by

$$q_{ab}^{(M,N)} := \widetilde{q}_{ab}^{(1)} \widetilde{q}_{ab}^{(2)} \cdots \widetilde{q}_{ab}^{(N)} \tag{2.9}$$

where

$$\widetilde{q}_{ab}^{(d)} = \begin{cases} 1 - \sum_{i=1}^{N} s_{di} \dfrac{(\rho_i + \sum_{k=1}^{M} \delta_{r,i}(k) a_{ik})}{1 + \sum_{n=1}^{M} (\delta_{r,i}(n) a_{in} + \delta_{b,i}(n)(1 - a_{in}))} & \text{if} \quad b_{dM} = 0 \\ \sum_{i=1}^{N} s_{di} \dfrac{(\rho_i + \sum_{k=1}^{M} \delta_{r,i}(k) a_{ik})}{1 + \sum_{n=1}^{M} (\delta_{r,i}(n) a_{in} + \delta_{b,i}(n)(1 - a_{in}))} & \text{if} \quad b_{dM} = 1, \end{cases} \tag{2.10}$$

with $d \in \{1, \cdots, N\}$.

We next show that for time-invariant reinforcement parameters (i.e., $\Delta_{r,i}(t) = \Delta_{r,i}$ and $\Delta_{b,i}(t) = \Delta_{b,i}$ for $t \geq 1$) the Markov chain $\{\mathbf{W}_t\}_{t=1}^{\infty}$ is an irreducible and aperiodic Markov chain.

Lemma 2.2. *For the* IPCN(M, N) *with time-invariant parameters, the transition probability matrix* $Q^{(M,N)}$ *is irreducible and aperiodic.*

Proof. Note that in (2.10), if we set $M = 1$, $\delta_{r,i}(t) = \delta_{r,i}$ and $\delta_{b,i}(t) = \delta_{b,i}$ for $t \geq 1$, then it is possible to go from any state to any state in one time step with a positive transition probability. Hence, the Markov chain is irreducible and aperiodic for memory $M = 1$. For memory $M > 1$ with $\delta_{r,i}(t) = \delta_{r,i}$ and $\delta_{b,i}(t) = \delta_{b,i}$ for $t \geq 1$, to prove irreducibility of the Markov chain, we show that given any two states, it is possible to go from one state to another in finitely many time steps with a positive probability. Let us fix two arbitrary states, $a = ((a_{11}, a_{21}, \cdots, a_{N1}), \cdots, (a_{1M}, a_{2M}, \cdots, a_{NM}))$ and $b = ((b_{11}, b_{21}, \cdots, b_{N1}), \cdots, (b_{1M}, b_{2M}, \cdots, b_{NM}))$. We next construct an M-step path (which occurs with a positive probability) between states a and b.

- Suppose the Markov chain is in state $\mathbf{W}_t = a$ at time t. At time $t + 1$, we go from state a to state,

$$\mathbf{W}_{t+1} = a^{(0)} = \Big(Z_{t+1} = (a_{12}, a_{22}, \cdots, a_{N2}), \\ \cdots, Z_{t+M} = (a_{1M}, a_{2M}, \cdots, a_{NM}), Z_{t+M} = (b_{11}, b_{21}, \cdots, b_{N1})\Big).$$

  Since $a_{ij} = a^{(0)}_{i(j-1)}$ for $i \in \{1, 2, \cdots, N\}$ and $j \in \{2, 3, \cdots, M\}$, the transition probability of going from state a to $a^{(0)}$ is nonzero and can be obtained using (2.9).

- At time $t + 2$ we go from state $a^{(0)}$ to state $a^{(1)}$

$$\mathbf{W}_{t+2} = a^{(1)} = \Big(Z_{t+2} = (a_{13}, \cdots, a_{N3}), \cdots, Z_{t+M+1} = (b_{11}, \cdots, b_{N1}), \\ Z_{t+M+1} = (b_{12}, \cdots, b_{N2})\Big).$$

Following this pattern of adding one N-tuple from state b at each time step, we will reach state b in M time steps. In summary, choosing any initial state, we can reach any other state of the Markov chain in at most M steps. Hence, the Markov chain is irreducible. Also, note that the period of the state with all zeros is one. Since all the states of an irreducible Markov chain have the same period, we obtain that this Markov chain is aperiodic. □

Lemma 2.2 guarantees the existence of a unique stationary distribution for the Markov process $\{\mathbf{W}_t\}_{t=1}^{\infty}$ when the reinforcement parameters are time-invariant. However, since the transition probabilities for this Markov chain (given in (2.10)) are complicated, we cannot analytically determine this unique stationary distribution in general. But we can obtain asymptotic marginal distributions for homogeneous system parameters (i.e., all parameters are identical and time-invariant across all urns): $R_i = R$, $B_i = B$, $\Delta_{r,i}(t) = \Delta_{b,i}(t) = \Delta$ for all $i \in \{1, \cdots, N\}$ and $t \geq 1$. In this (homogeneous) case, we have

$$\rho = \frac{R}{T}, \qquad \sigma = 1 - \frac{R}{T}, \qquad \delta = \frac{\Delta}{T},$$

where $T = R + B$ is the total number of balls in the urns. In the following example and theorem, we illustrate the stochastic properties of the underlying Markov process for a homogeneous IPCN$(1, 2)$ system.

Example 2.3. Given a homogeneous IPCN$(1, 2)$ system with interaction matrix

$$S = \begin{bmatrix} s_{11} & 1 - s_{11} \\ s_{21} & 1 - s_{21} \end{bmatrix}$$

the stationary distribution $\pi = [\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11}]$ for the Markov process $\{(Z_t, Z_{t+1})\}$ (transition probability matrix $Q^{(1,2)}$) is given by

$$\begin{aligned}
\pi_{00} &= \frac{2\sigma^2\delta + \sigma^2 + (1 - s_{11} - s_{21} + 2s_{11}s_{21})\sigma\delta^2}{(1 - s_{11} - s_{21} + 2s_{11}s_{21})\delta^2 + 2\delta + 1} \\
\pi_{01} &= \frac{\rho\sigma(1 + 2\delta)}{(1 - s_{11} - s_{21} + 2s_{11}s_{21})\delta^2 + 2\delta + 1} \\
\pi_{10} &= \frac{\rho\sigma(1 + 2\delta)}{(1 - s_{11} - s_{21} + 2s_{11}s_{21})\delta^2 + 2\delta + 1} \\
\pi_{11} &= \frac{\rho(2\delta - \sigma - 2\sigma\delta + (1 - s_{11} - s_{21} + 2s_{11}s_{21})\delta^2 + 1)}{(1 - s_{11} - s_{21} + 2s_{11}s_{21})\delta^2 + 2\delta + 1}.
\end{aligned}$$

It is easy to see that $\pi = [\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11}]$ satisfies the equation $\pi Q^{(1,2)} = \pi$. We thus have

- $\lim_{t\to\infty} P(Z_{1,t} = 1) = \pi_{10} + \pi_{11} = \rho$, $\lim_{t\to\infty} P(Z_{2,t} = 1) = \pi_{01} + \pi_{11} = \rho$,
- $\lim_{t\to\infty} P(Z_{1,t} = 0) = \pi_{00} + \pi_{01} = \sigma$, $\lim_{t\to\infty} P(Z_{2,t} = 0) = \pi_{00} + \pi_{10} = \sigma$.

Also writing (2.7) for $M = 1$ and taking expectation both sides, we have for $i = 1, 2$ that

$$\lim_{t\to\infty} \mathbb{E}[U_{i,t}] = \frac{\rho + \delta \lim_{t\to\infty} \mathbb{E}[Z_{i,t}]}{1 + \delta} = \frac{\rho + \delta \lim_{t\to\infty} P(Z_{i,t} = 1)}{1 + \delta} = \rho.$$

Hence, irrespective of the used interaction matrix, the asymptotic marginal (one-fold) distributions and urn compositions for the IPCN$(1, 2)$ system are the same as for the single (memory one) Pólya urn studied in [1]; this result is proved in general in Theorem 2.4 below. We, however, next observe that the asymptotic 2-fold draw distributions for the IPCN$(1, 2)$ urns do not match their counterparts for the single Pólya urn process of [1]. Indeed, the 2-fold (joint) distribution vector of the single-urn

(stationary) Pólya Markov chain in [1] is given by

$$\tilde{\pi}^{(2)} = \left[\frac{\sigma(\sigma+\delta)}{1+\delta}, \frac{\rho\sigma}{1+\delta}, \frac{\rho\sigma}{1+\delta}, \frac{\rho(\rho+\delta)}{1+\delta}\right].$$

For the homogeneous IPCN$(1,2)$ system, the joint probability $P(Z_{1,t}=a_1, Z_{1,t+1}=b_1)$ for urn 1 of the homogeneous IPCN$(1,2)$ system is given by

$$\begin{aligned} P(Z_{1,t}=a_1, Z_{1,t+1}=b_1) &= \sum_{a_2,b_2\in\{0,1\}} P(Z_{1,t}=a_1, Z_{1,t+1}=b_1, Z_{2,t}=a_2, Z_{2,t+1}=b_2) \\ &= \sum_{a_2,b_2\in\{0,1\}} P(Z_{1,t+1}=b_1, Z_{2,t+1}=b_2 | Z_{1,t}=a_1, Z_{2,t}=a_2) P(Z_{1,t}=a_1, Z_{2,t}=a_2). \end{aligned}$$

Thus noting the conditional independence of $Z_{1,t+1}$ and $Z_{2,t+1}$ given $(Z_{1,t}, Z_{2,t})$, and using the IPCN$(1,2)$ matrix $Q^{(1,2)}$ along with the fact that

$$\lim_{t\to\infty} P(Z_{1,t}=a_1, Z_{2,t}=a_2) = \pi_{a_1,a_2},$$

we obtain

$$\begin{aligned} \lim_{t\to\infty} P(Z_{1,t}=0, Z_{1,t+1}=0) &= \frac{\sigma(\sigma+\delta)}{1+\delta} - \frac{\pi_{01}(1-s_{11})\delta}{(1+\delta)} \\ \lim_{t\to\infty} P(Z_{1,t}=0, Z_{1,t+1}=1) &= \frac{\sigma\rho}{1+\delta} + \frac{\pi_{01}(1-s_{11})\delta}{(1+\delta)} \\ \lim_{t\to\infty} P(Z_{1,t}=1, Z_{1,t+1}=0) &= \frac{\sigma\rho}{1+\delta} + \frac{\pi_{01}(1-s_{11})\delta}{(1+\delta)} \\ \lim_{t\to\infty} P(Z_{1,t}=1, Z_{1,t+1}=1) &= \frac{\rho(\rho+\delta)}{1+\delta} - \frac{\pi_{01}(1-s_{11})\delta}{(1+\delta)}, \end{aligned}$$

which shows explicitly by how much the asymptotic 2-fold draw distribution for urn 1

deviates from $\tilde{\pi}^{(2)}$. Note that by setting $s_{11} = 1$, the error term $\pi_{01}(1 - s_{11})\delta/(1 + \delta)$ reduces to zero, making the two distributions match, as expected (since when $s_{11} = 1$, urn 1 only interacts with itself). •

Note that it is much harder to derive in closed-form the stationary distribution for the homogeneous IPCN(M, N) system with $M > 1$ and $N > 2$ but we have the following asymptotic marginal probabilities for a homogeneous IPCN(M, N) system.

Theorem 2.4. *For a homogeneous* IPCN(M, N) *system*

$$\lim_{t\to\infty} P(Z_{i,t} = 1) = \rho \tag{2.11}$$

for all urns i in the network.

Proof. Let $\gamma_i = \lim_{t\to\infty} \mathbb{E}[Z_{i,t}]$ for $i \in \{1, 2, \cdots, N\}$. Using (2.10), we obtain

$$\gamma_d = \frac{\rho + \sum_{i=1}^{N} s_{di}(M\gamma_i)\delta}{1 + M\delta}$$

for $d \in \{1, 2, \ldots, N\}$. Let $\mathbf{1}_N = [1, \cdots, 1]^{\mathbf{T}}$ and $\gamma = [\gamma_1, \cdots, \gamma_N]^{\mathbf{T}}$, where $\mathbf{T}$ denotes transposition. Then,

$$(1 + M\delta)\gamma = \rho\mathbf{1}_N + (M\delta)S\gamma$$

which gives

$$(1 + M\delta)(\gamma - \rho\mathbf{1}_N) = (M\delta)S(\gamma - \rho\mathbf{1}_N).$$

Setting $\widetilde{\gamma} := \gamma - \rho\mathbf{1}_N$ in the above equation, we have that

$$S\widetilde{\gamma} = \frac{1 + M\delta}{M\delta}\widetilde{\gamma}.$$

Since the eigenvalues of S have absolute values less than or equal to one (as S is a row-stochastic matrix), we obtain that

$$\widetilde{\gamma} = 0$$

which implies that

$$\gamma_i = \rho \quad \forall \quad i \in \{1, 2 \cdots, N\}.$$

□

Unlike the homogeneous case, we cannot analytically obtain the asymptotic marginal probabilities for general non-homogeneous IPCN(M, N) system. In the next section, we construct a class of time-delayed dynamical systems using *mean-field approximations* to approximate the asymptotic marginal probabilities for the draw variables of urns in IPCN(M, N) system with time-invariant reinforcement parameters.

2.3 Dynamical System Models

Throughout this section, we assume that $\Delta_{r,i}(t) = \Delta_{r,i}$ and $\Delta_{b,i}(t) = \Delta_{b,i}$ for all urns $i \in \{1, \cdots, N\}$ and $t \geq 1$, i.e., the reinforcement parameters are time-invariant. Given an IPCN(M, N) and an urn i, we denote the probability of making a red draw for urn i at time t by

$$P_i(t) := P(Z_{i,t} = 1).$$

2.3.1 Exact Dynamical System for M = 1

We first show that an IPCN$(1, N)$ system can be exactly represented via a linear dynamical system.

Recall from (2.7) and (2.8) with $M = 1$ that the conditional probability of drawing a red ball from urn i at time t for the case of time-invariant reinforcement parameters, given all the draw variables at time $t-1$, is as follows:

$$\begin{aligned} P(Z_{i,t} = 1 | Z_{1,t-1}, Z_{2,t-1}, \cdots, Z_{N,t-1}) &= \sum_{j=1}^{N} \frac{s_{ij}(\rho_j + \delta_{r,j} Z_{j,t-1})}{1 + \delta_{r,j} Z_{j,t-1} + (1 - Z_{j,t-1})\delta_{b,j}} \\ &= \sum_{j=1}^{N} \frac{s_{i,j}(\rho_j + \delta_{r,j})}{1 + \delta_{r,j}} Z_{j,t-1} + \frac{s_{ij}\rho_j}{1 + \delta_{b,j}}(1 - Z_{j,t-1}) \\ &= \sum_{j=1}^{N} [s_{ij}\beta_1^{(j)}(1) Z_{j,t-1} + s_{ij}\beta_1^{(j)}(0)(1 - Z_{j,t-1})] \end{aligned} \tag{2.12}$$

where

$$\beta_1^{(j)}(k) := \frac{\rho_j + k\delta_{r,j}}{1 + k\delta_{r,j} + (1-k)\delta_{b,j}}, \quad j \in \{1, \cdots, N\}, \quad k \in \{0, 1\}.$$

Now taking expectation with respect to $(Z_{1,t-1}, \cdots, Z_{N,t-1})$ on both sides of (2.12), we get

$$P_i(t) = \sum_{j=1}^{N} [\beta_1^{(j)}(1) s_{ij} P_j(t-1) + s_{ij}\beta_1^{(j)}(0)(1 - P_j(t-1))]. \tag{2.13}$$

To this end, defining the vector $P(t)$ as

$$P(t) = [P_1(t), P_2(t), \cdots, P_N(t)]^{\mathbf{T}},$$

we obtain the following dynamical system for the IPCN$(1, N)$ network.

Theorem 2.5. *For the* IPCN$(1, N)$ *system, the infection vector satisfies the equation*

$$P(t) = J_{N,1}P(t-1) + C_{N,1} \tag{2.14}$$

where $J_{N,1} \in \mathbb{R}^{N\times N}$, $C_{N,1} \in \mathbb{R}^{N\times 1}$ *are matrices with respective entries:*

$$[J_{N,1}]_{i\times j} = \frac{s_{ij}(\rho_j + \delta_{r,j})}{(1+\delta_{r,j})} - \frac{s_{ij}\rho_j}{(1+\delta_{b,j})} = s_{ij}(\beta_1^{(j)}(1) - \beta_1^{(j)}(0))$$

$$\textit{and} \quad [C_{N,1}]_{i\times 1} = \sum_{j=1}^{N} \frac{s_{ij}\rho_j}{(1+\delta_{b,j})} = \sum_{j=1}^{N} s_{ij}\beta_1^{(j)}(0).$$

Proof. Follows from (2.13). □

We next examine the equilibrium of this linear dynamical system.

Theorem 2.6. *The linear dynamical system for the* IPCN$(1, N)$ *system given by* (2.14) *has a unique equilibrium point given by* $P^* = (I - J_{N,1})^{-1}C_{N,1}$ *and*

$$\lim_{t\to\infty} P_i(t) = P_i^*$$

for all $i \in \{1, \ldots, N\}$.

Proof. It is enough to show that the spectral radius of the matrix $J_{N,1}$ is less than one;

since the spectral radius is less than, or equal to, the row sum norm of the matrix, it is enough to show that the row sum norm of $J_{N,1}$ is strictly less than 1. Since $0 \leq \rho_j \leq 1$ and $\delta_{r,j}, \delta_{b,j} > 0$ for all $j \in \{1, 2..., N\}$, we have

$$-1 < \frac{(\rho_j + \delta_{r,j})}{(1 + \delta_{r,j})} - \frac{\rho_j}{(1 + \delta_{b,j})} < 1 \tag{2.15}$$

Hence, the sum of absolute values of entries in ith row of the matrix $J_{N,1}$ satisfies

$$\sum_{j=1}^{N} s_{ij} \left| \frac{(\rho_j + \delta_{r,j})}{(1 + \delta_{r,j})} - \frac{\rho_j}{(1 + \delta_{b,j})} \right| < \sum_{j=1}^{N} s_{ij} = 1,$$

which yields the result. □

As an illustration, we find the equilibrium of the linear dynamical system (2.14) for a much simpler IPCN$(1, N)$ system.

Corollary 2.7. *Given an* IPCN$(1, N)$ *system with* $S = I$*,*

$$\lim_{t \to \infty} P_i(t) = \frac{\rho_i(1 + \delta_{r,i})}{1 + \delta_{b,i} + \rho_i(\delta_{r,i} - \delta_{b,i})}. \tag{2.16}$$

Proof. For an IPCN$(1, N)$ system with $S = I$, we have that

$$P(Z_{1,t-1} = a_1, \cdots, Z_{N,t-1} = a_N) = \prod_{j=1}^{N} P(Z_{j,t-1} = a_j).$$

In this case, since $S = I$ and hence the draw variables of urns are independent of each other. The asymptotic value of $P_i(t)$ for $i \in \{1, \cdots, N\}$ is given by the equilibrium

point of the linear dynamical system

$$P(t) = J_{N,1}P(t-1) + C_{N,1}$$

which is given by $P^* \in \mathbb{R}^N$ whose ith component is given by (2.16).

Another way to find this equilibrium point is to write the transition probability matrix for a single urn using (2.8) and solving for stationary distribution to obtain $\lim_{t\to\infty} P_i(t)$. The transition probability matrix for a single non-homogeneous urn i with time-invariant reinforcement parameters is given by

$$Q^{(1,1)} = \begin{bmatrix} \dfrac{\sigma_i + \delta_{b,i}}{1+\delta_{b,i}} & \dfrac{\rho_i}{1+\delta_{b,i}} \\ \dfrac{\sigma_i}{1+\delta_{r,i}} & \dfrac{\rho_i + \delta_{r,i}}{1+\delta_{r,i}} \end{bmatrix}.$$

On solving for the stationary distribution, $[\pi_0, \pi_1]Q^{(1,1)} = [\pi_0, \pi_1]$, we obtain that π_1 indeed equals the right-hand side of (2.16). □

We also illustrate Theorem 2.6 by examining the special homogeneous case. This aligns with the result in Theorem 2.4.

Corollary 2.8. *For a homogeneous* IPCN$(1, N)$ *system, the equilibrium of* (2.14) *is given by* $P^* = \rho\mathbf{1}_N$*, where* $\mathbf{1}_N$ *is vector of ones of size* N*.*

Proof. By Theorem 2.6, the equilibrium P^* is given by

$$P^* = (I - J_{N,1})^{-1}C_{N,1}.$$

Note that the row sums in $(I - J_{N,1})$ are given by $\frac{1}{1+\delta}$, i.e.,

$$(I - J_{N,1})\mathbf{1}_N = \frac{1}{1+\delta}\mathbf{1}_N.$$

Therefore,

$$(I - J_{N,1})^{-1}C_{N,1} = (I - J_{N,1})^{-1}\begin{bmatrix} \frac{\rho}{(1+\delta)} & \frac{\rho}{(1+\delta)} & \cdots & \frac{\rho}{(1+\delta)} \end{bmatrix}^{\mathbf{T}} = \rho\mathbf{1}_N.$$

□

2.3.2 Approximating Dynamical Systems for M > 1

For IPCN(M, N) systems with $M > 1$, we resort to using *mean-field approximations* for the construction of dynamical systems. We assume that for t sufficiently large and $t > M$, for each urn i, $Z_{i,t-1}$,$\cdots$,$Z_{i,t-M}$ are approximately independent of each other; i.e., at any given time instant $t > M$ (where t is sufficiently large), we assume that

$$P(Z_{j,t-1}, Z_{j,t-2}, \cdots, Z_{j,t-M}) \approx \prod_{k=1}^{M} P(Z_{j,t-k}), \tag{2.17}$$

for all $j \in \{1, 2, \cdots, N\}$. For the IPCN$(M, N)$ system, we have from (2.8) that

$$\begin{aligned} &P(Z_{i,t} = 1 | (Z_{1,t-1}, \cdots, Z_{1,t-M}), \cdots, (Z_{N,t-1}, \cdots, Z_{N,t-M})) \\ &= \sum_{j=1}^{N} \frac{s_{ij}(\rho_j + \delta_{r,j}\sum_{k=1}^{M} Z_{j,t-k})}{1 + \sum_{n=1}^{M}(\delta_{r,j}Z_{j,t-n} + \delta_{b,j}(1 - Z_{j,t-n}))}. \end{aligned} \tag{2.18}$$

Now, taking expectation with respect to

$$((Z_{1,t-1},\ldots,Z_{1,t-M}),\ldots,((Z_{N,t-1},\ldots,Z_{N,t-M}))$$

on both sides of (2.18) and using the linearity property of expectation, we obtain

$$P(Z_{i,t}=1)=\sum_{j=1}^{N}\mathbb{E}\left[\frac{s_{ij}(\rho_j+\delta_{r,j}\sum\limits_{k=1}^{M}Z_{j,t-k})}{1+\sum\limits_{n=1}^{M}(\delta_{r,j}Z_{j,t-n}+\delta_{b,j}(1-Z_{j,t-n}))}\right] \tag{2.19}$$
$$=\sum_{j=1}^{N}\sum_{B_M}\frac{s_{ij}(\rho_j+\delta_{r,j}\sum_{k=1}^{M}a_k)}{1+\sum\limits_{n=1}^{M}(\delta_{r,j}a_n+\delta_{b,j}(1-a_n))}P(Z_{j,t-1}=a_1,\cdots Z_{j,t-M}=a_M)$$

where

$$B_M:=\{(a_1,a_2,\cdots a_M)\mid a_k\in\{0,1\}\quad\text{for}\quad k\in\{1,2,\cdots,M\}\}. \tag{2.20}$$

Now we use the *mean-field approximation* (2.17) in (2.19) to obtain the following class of approximating nonlinear dynamical systems

$$P_i(t)\approx\sum_{j=1}^{N}\sum_{B_M}\frac{s_{ij}(\rho_j+\delta_{r,j}\sum_{k=1}^{M}a_k)}{1+\sum\limits_{n=1}^{M}(\delta_{r,j}a_n+\delta_{b,j}(1-a_n))}\prod_{k=1}^{M}P(Z_{j,t-k}=a_k)= \tag{2.21}$$
$$\sum_{j=1}^{N}\sum_{B_M}\frac{s_{ij}(\rho_j+\delta_{r,j}\sum_{k=1}^{M}a_k)}{1+\delta_{r,j}\sum\limits_{n=1}^{M}a_n+\delta_{b,j}(M-\sum\limits_{n=1}^{M}a_n)}\prod_{k=1}^{M}(a_kP_j(t-k)+(1-a_k)(1-P_j(t-k))).$$

For simplicity of notation, we write (2.21) in the following way:

$$P_i(t) \approx \sum_{B_M} \big(\sum_{j=1}^{N} s_{ij} \beta_M^{(j)}(v_M) \big) \prod_{k=1}^{M} (a_k P_j(t-k) + (1-a_k)(1-P_j(t-k))) \tag{2.22}$$

where

$$v_M = \sum_{k=1}^{M} a_k \tag{2.23}$$

with $a_k \in \{0,1\}$ and $k \in \{1,2,\cdots,M\}$, and

$$\beta_M^{(j)}(l) = \frac{\rho_j + l\delta_{r,j}}{1 + l\delta_{r,j} + (M-l)\delta_{b,j}}, \quad j \in \{1,\cdots,N\}, l \in \{0,1,\cdots M\}.$$

We next give a useful rearranged form of (2.22). In particular, even though (2.22) appears to be complicated, after some simplifications, the coefficients of the nonlinear terms follow a binomial pattern. To give an idea of this binomial pattern, we will first present a few examples and then give a proof of the formula for the rearranged form of (2.22).

Example 2.9. We note that for the IPCN$(2,2)$ system, the approximating dynamical system is given by

$$\begin{aligned} P_i(t) \approx & \sum_{j=1}^{2} s_{ij} \beta_2^{(j)}(0) + \sum_{j=1}^{2} \sum_{k=1}^{2} s_{ij} \left(\beta_2^{(j)}(1) - \beta_2^{(j)}(0) \right) P_j(t-k) \\ & + \sum_{j=1}^{2} s_{ij} \left(\beta_2^{(j)}(2) - 2\beta_2^{(j)}(1) + \beta_2^{(j)}(0) \right) \prod_{k=1}^{2} P_j(t-k). \end{aligned}$$

Next, by expansion, we observe that for IPCN$(3,2)$ system, the approximating dynamical system is given by

$$\begin{aligned}
P_i(t) \approx & \sum_{j=1}^{2} s_{ij}\beta_3^{(j)}(0) + \sum_{j=1}^{2}\sum_{k=1}^{3} P_j(t-k)s_{ij}\left(\beta_3^{(j)}(1) - \beta_3^{(j)}(0)\right) \\
& + \sum_{j=1}^{2} P_j(t-1)P_j(t-2)s_{ij}\left(\beta_3^{j}(0) - 2\beta_3^{(j)}(1) + \beta_3^{(j)}(2)\right) \\
& + \sum_{j=1}^{2} P_j(t-2)P_j(t-3)s_{ij}\left(\beta_3^{(j)}(0) - 2\beta_3^{(j)}(1) + \beta_3^{j)}(2)\right) \\
& + \sum_{j=1}^{2} P_j(t-1)P_j(t-3)s_{ij}\left(\beta_3^{(j)}(0) - 2\beta_3^{(j)}(1) + \beta_3^{(j)}(2)\right) \\
& + \sum_{j=1}^{2} P_j(t-1)P_j(t-2)P_j(t-3)s_{ij}\left(3\beta_3^{(j)}(1) - 3\beta_3^{(j)}(2) - \beta_3^{(j)}(0) + \beta_3^{(j)}(3)\right),
\end{aligned}$$

where one can already observe the binomial pattern that we hinted at. •

We will now obtain a rearrangement of (2.22) for a general IPCN(M,N) system.

Theorem 2.10. *For the* IPCN(M,N) *system with* $M>1$*, the approximating dynamical system* (2.22) *can be written as*

$$\begin{aligned}
P_i(t) \approx & \sum_{j=1}^{N} s_{ij}\beta_M^{(j)}(0) \qquad (2.24)\\
& + \sum_{j=1}^{N}\sum_{n=1}^{M}\left[\left(\sum_{k=0}^{n}\left((-1)^{n-k}\binom{n}{k}s_{ij}\beta_M^{(j)}(k)\right)\right)\left(\sum_{\substack{(d_1,\cdots,d_n)\\ \in H_{n,M}}} P_j(t-d_1)\cdots P_j(t-d_n)\right)\right],
\end{aligned}$$

where

$$H_{n,M} := \{(d_1,d_2,\cdots,d_n)\big| d_i \in \{1,\cdots,M\},\ d_i < d_j \ \ for\ \ i<j\ , i,j\in\{1,\cdots,n\}\}.$$

Proof. We show that (2.24) is obtained by a rearrangement of (2.22).

The right-hand side of (2.22) is given by:

$$\sum_{j=1}^{N}\Big[\sum_{B_M} s_{ij}\beta_M^{(j)}(v_M)\prod_{k=1}^{M}\Big(a_kP_j(t-k)+(1-a_k)(1-P_j(t-k))\Big)\Big]. \tag{2.25}$$

The constant term can be extracted from (2.25) by setting $\underline{a}_j = (0, 0, \cdots, 0)$ in B_M for $1 \leq j \leq N$ and is given by

$$\sum_{j=1}^{N}\frac{s_{ij}\rho_j}{1+M\delta_{b,j}}.$$

Now, fixing $j \in \{1, 2, \cdots, N\}$, we expand the term

$$\sum_{B_M} s_{ij}\beta_M^{(j)}(v_M)\prod_{k=1}^{M}(a_kP_j(t-k)+(1-a_k)(1-P_j(t-k))). \tag{2.26}$$

Note that the order of (2.26) is M. In order to get the nth degree term ($1 \leq n \leq M$ in (2.26)), we need to choose n corresponding $P_j(t-k)$'s, where $k \in \{1, 2, \cdots, M\}$ from the product

$$\prod_{k=1}^{M}(a_kP_j(t-k)+(1-a_k)(1-P_j(t-k)))$$

and the rest $M-n$ chosen terms have to be 1. We then look at the coefficient of the chosen nth order term. Note that the coefficients of the chosen $P_j(t-k)$'s are either 1 or -1, depending on the tuple $\underline{a}_j$. Given a tuple $\underline{a}_j$, there are exactly v_M, $P_j(t-k)$'s with coefficients 1 and the rest $n - v_M$ of them have coefficient -1.

Summing over all the possible coefficients of the nth degree term of (2.25) we get

$$\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}s_{ij}\beta_M^{(j)}(k)\sum_{\substack{(d_1,\cdots,d_n)\\ \in H_{n,M}}}P_j(t-d_1)\cdots P_j(t-d_n).$$

Finally, we can obtain the nth degree terms ($1 \leq n \leq M$) for the other $N-1$ urns in exactly the same way as above. □

The analysis of the nonlinear dynamical systems given in (2.24) is clearly more intricate than the one in the case with memory one, where the evaluations were given by a linear dynamical system, namely (2.14). This being said, given that the presence of nonlinearity is due to the product of probabilities, we can use a further approximation by considering the leading linear terms.

Corollary 2.11. *The linear part of the dynamical system* (2.24) *is given by*

$$P_i(t) \approx \sum_{j=1}^{N} s_{ij}\beta_M^{(j)}(0) + \sum_{j=1}^{N}\sum_{k=1}^{M} s_{ij}\left(\beta_M^{(j)}(1) - \beta_M^{(j)}(0)\right)P_j(t-k). \tag{2.27}$$

Proof. Setting $n = 1$ in (2.24) we obtain

$$\begin{aligned}
P_i(t) &\approx \sum_{j=1}^{N} s_{ij}\beta_M^{(j)}(0)\\
&\quad + \sum_{j=1}^{N}\left((-1)\binom{1}{0}s_{ij}\beta_M^{(j)}(0) + (-1)^2\binom{1}{1}s_{ij}\beta_M^{(j)}(k)\right)\Big(\sum_{d\in H_{1,M}}P_j(t-d)\Big)\\
&= \sum_{j=1}^{N} s_{ij}\beta_M^{(j)}(0) + \sum_{j=1}^{N}\sum_{k=1}^{M} s_{ij}(\beta_M^{(j)}(1) - \beta_M^{(j)}(0))P_j(t-k).
\end{aligned}$$

□

Equation (2.27) gives an approximate linear dynamical system for the IPCN(M, N) system. Furthermore, for $M > 1$ network of N urns, we define

$$\widetilde{P}(t) := [P_1(t), \cdots, P_1(t-M), P_2(t), \cdots, P_2(t-M) \cdots, P_N(t), \cdots, P_N(t-M)]^{\mathbf{T}}.$$

Using (2.24) and dropping the nonlinear terms, we can write,

$$\widetilde{P}(t) \approx J_{N,M}\widetilde{P}(t-1) + C_{N,M} \tag{2.28}$$

where, $J_{N,M}$ is a block matrix with N^2 blocks of size $M \times M$.

$$J_{N,M} = \begin{bmatrix} J_{N,M}(1,1) & J_{N,M}(1,2) & \cdots & J_{N,M}(1,N) \\ J_{N,M}(2,1) & J_{N,M}(2,2) & \cdots & J_{N,M}(2,N) \\ \vdots & \ddots & \vdots & \vdots \\ J_{N,M}(N,1) & J_{N,M}(N,2) & \cdots & J_{N,N}(N,N) \end{bmatrix}_{NM \times NM}.$$

Here, the diagonal blocks of matrix, $J_{N,M}(i,i)$ are given by

$$\left[\begin{array}{ccc|c} s_{ii}(\beta_M^{(i)}(1) - \beta_M^{(i)}(0)) & \cdots & s_{ii}(\beta_M^{(i)}(1) - \beta_M^{(i)}(0)) & s_{ii}(\beta_M^{(i)}(1) - \beta_M^{(i)}(0)) \\ \hline \multicolumn{3}{c|}{\mathbf{I}_{(M-1)\times(M-1)}} & \mathbf{0}_{(M-1)\times 1} \end{array}\right]_{M \times M}$$

where $\mathbf{I}_{(M-1)\times(M-1)}$ is the identity matrix of size $M-1$ and $\mathbf{0}_{(M-1)\times 1}$ is the column vector of length $M-1$ with all entries zero. Similarly, the off-diagonal blocks $J_{N,M}(i,j)$

are given by

$$\left[\begin{array}{ccc|c} s_{ij}(\beta_M^{(j)}(1)-\beta_M^{(j)}(0)) & \cdots & s_{ij}(\beta_M^{(j)}(1)-\beta_M^{(j)}(0)) & s_{ij}(\beta_M^{(j)}(1)-\beta_M^{(j)}(0)) \\ \hline \multicolumn{3}{c|}{\mathbf{0}_{(M-1)\times(M-1)}} & \mathbf{0}_{(M-1)\times 1} \end{array}\right]_{M\times M}$$

where, $\mathbf{0}_{(M-1)\times(M-1)}$ is a matrix of size $(M-1)$ with all entries zero. Finally, $C_{N,M}$ is a column matrix with N blocks each of size $1\times M$ given by

$$C_{N,M}(i)=\left[\sum_{j=1}^{N} s_{ij}\beta_M^{(j)}(0) \quad 0 \quad \cdots \quad 0\right]_{M\times 1}^{\mathbf{T}}.$$

The linear dynamical system (2.28) has a unique equilibrium which is given by $(I-J_{N,M})^{-1}C_{N,M}$. Even though we leave the examination of stability properties of the nonlinear dynamical system (2.24) as a future direction, it is worth pointing out that (2.28) asymptotically converges to the unique equilibrium if and only if the spectral radius of $J_{N,M}$ is less than one. The possible dependency of this condition to the interaction matrix and urn properties is also an interesting future direction. A detailed discussion of other future directions is presented in Chapter 5.

2.4 Simulation Results

We provide a set of simulations[2] to illustrate our results. For this purpose, we have considered four different setups which are aimed at demonstrating the impact of memory, as well as initial urn compositions and reinforcement parameters. In particular, for the first two networks with $N=10$ (i.e., Figure 2.2 and Figure 2.3), we

[2]For a complete list of parameters used for generating all figures of this chapter, see the link: https://www.dropbox.com/sh/19py25reaxnfoyn/AABFdBp98J-9Jkd7zzVfTAQ9a?dl=0

use δ_r values that are significantly larger than the δ_b values in Figure 2.2 and δ_b values significantly larger than δ_r values in Figure 2.3. In Figure 2.4, we consider larger size non-homogeneous networks with $N = 100$. We simulate the IPCN(M, N) system for $M = 1, 2, 3$ and their corresponding approximating (nonlinear) dynamical systems given by (2.24). We also simulate the linear approximation (2.27) of the nonlinear dynamical system for each $M = 2, 3$. Recall that for $M = 1$, the linear dynamical system in (2.14) exactly characterizes the underlying Markov draw process. Finally, in Figure 2.5, we simulate a homogeneous IPCN(M, N) system.

Throughout, for the given IPCN(M, N) system, we plot the average empirical sum at time t, which is given by

$$\frac{1}{N}\sum_{i=1}^{N} I_t(i)$$

where

$$I_t(i) = \frac{1}{t}\sum_{n=1}^{t} Z_{i,n}.$$

For each plot, the average empirical sum is computed 100 times and the mean value is plotted against time. For the dynamical systems, we plot the average infection rate at time t, which is given by

$$\frac{1}{N}\sum_{i=1}^{N} P_i(t).$$

We first note from the simulations that for the network with $M = 1$, the linear system in (2.14) matches the empirical sum of the draw process, as expected since in

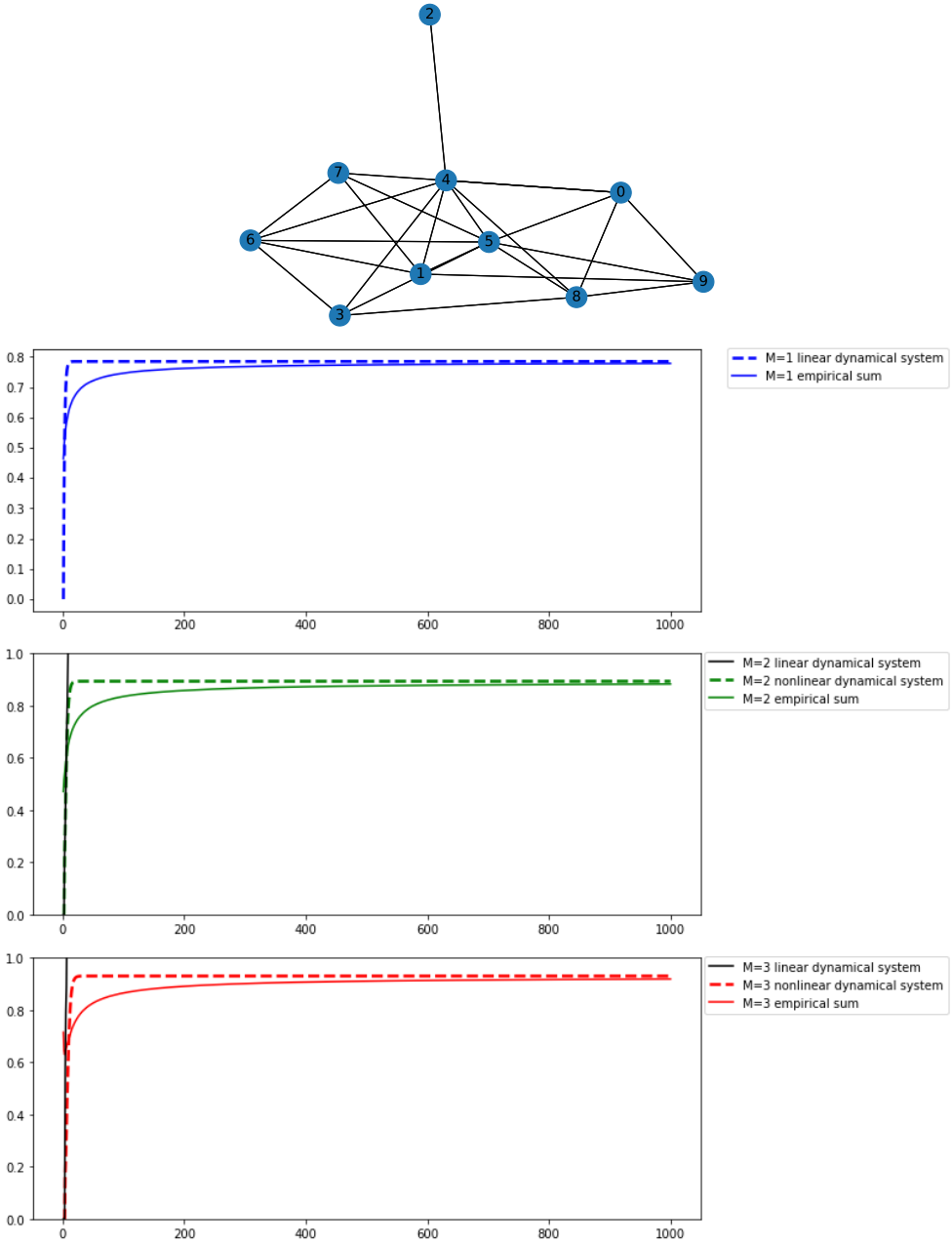


Figure 2.2: Infection rate curves for non-homogeneous IPCN(M, N) systems with $N = 10$ nodes and memory $M = 1, 2, 3$. At $t = 0$, each urn has a total of 25 balls. The number of red balls in each urn at $t = 0$ is chosen randomly between range 5 to 23 so that $\rho's$ lie in the range 0.2 to 0.92. $\Delta_r's$ are chosen randomly between range 60 to 70 and $\Delta_b's$ are randomly chosen between range 20 to 29. For simplicity, we set the initial values $P_i(0), P_i(1), \cdots, P_i(M-1)$ all equal to zero for all urns i in the network.

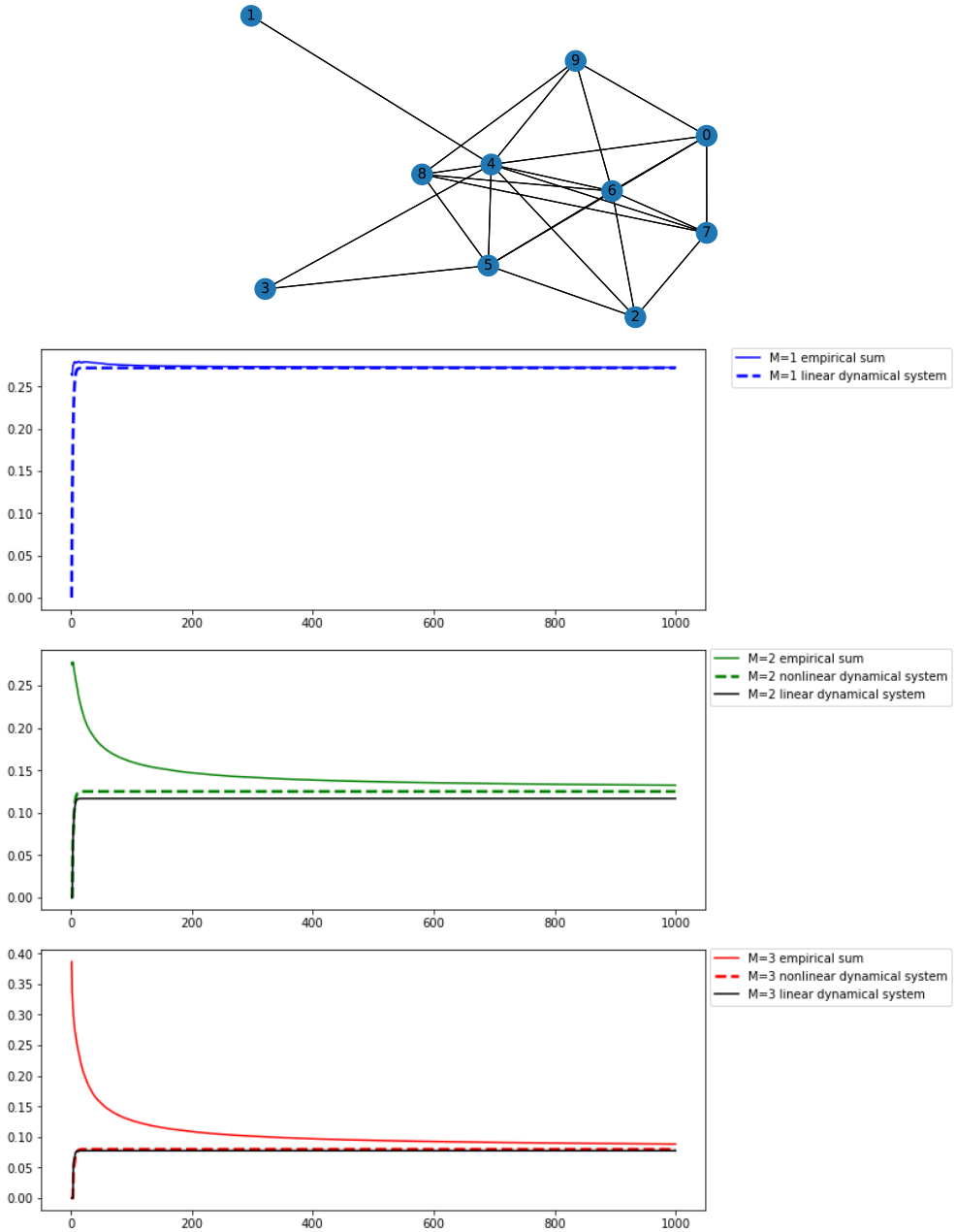


Figure 2.3: Infection rate curves for non-homogeneous IPCN(M, N) systems with $N = 10$ nodes and memory $M = 1, 2, 3$. At $t = 0$, the total number of balls in each urn is 25. The number of red balls in each urn at time $t = 0$ are chosen randomly between the range 2 to 17 so that $\rho's$ lie in the range 0.08 to 0.68. $\Delta'_r s$ are chosen randomly in the range 12 to 30. $\Delta'_b s$ are chosen in the range 61 to 80. For simplicity, we set the initial values $P_i(0), P_i(1), \cdots, P_i(M-1)$ all equal to zero for all urns i.

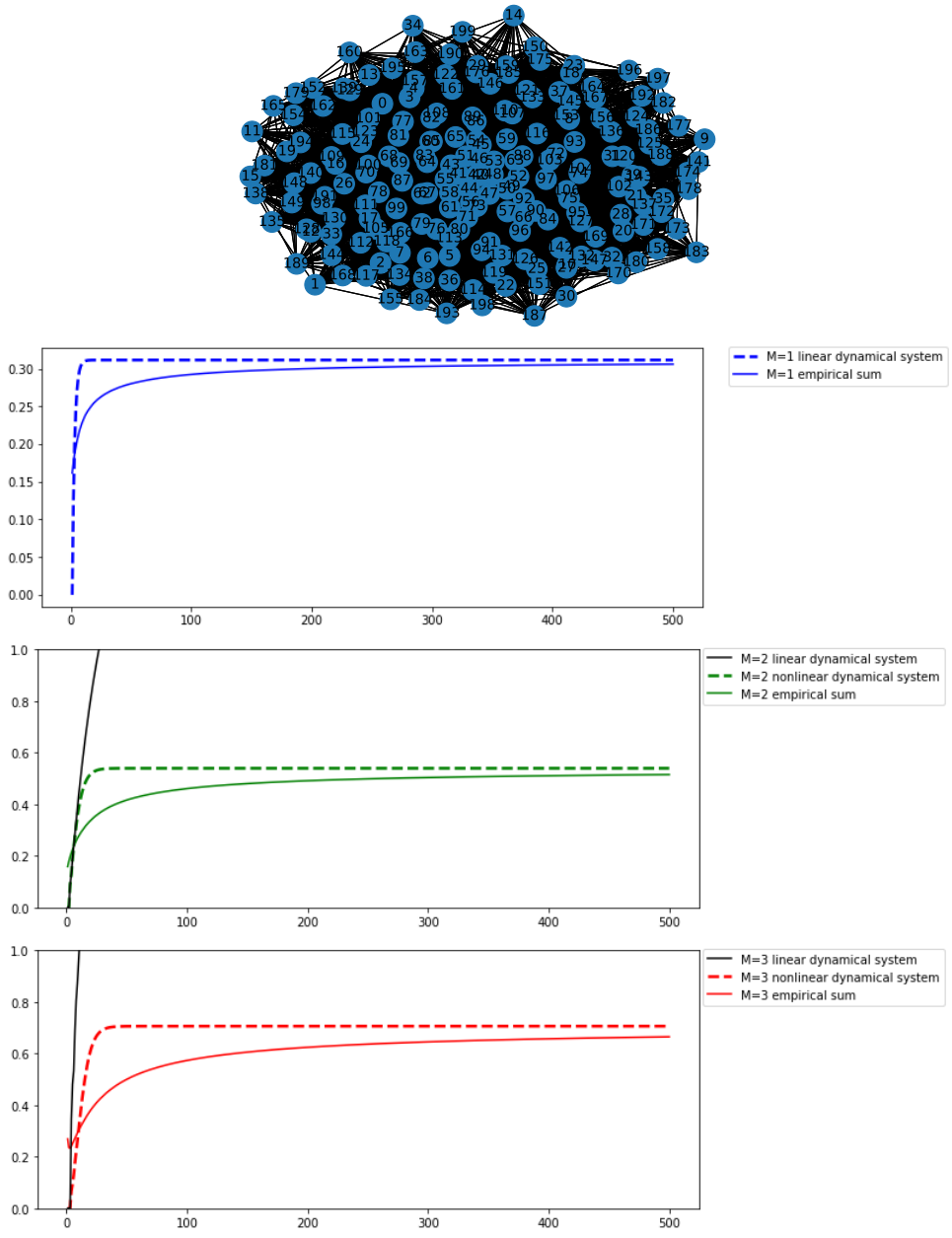


Figure 2.4: Infection rate curves for non-homogeneous IPCN(M, N) Barabási-Albert systems [3] with $N = 100$ nodes and memory $M = 1, 2, 3$. At $t = 0$, the total number of balls in each urn is 25. The number of red balls in each urn at time $t = 0$ are chosen randomly between the range 1 to 10 so that $\rho's$ lie in the range 0.04 to 0.4. $\Delta_r's$ are chosen randomly in the range 40 to 50. $\Delta_b's$ are chosen in the range 15 to 25. For simplicity, we set the initial values $P_i(0), P_i(1), \cdots, P_i(M-1)$ all equal to zero for all urns i.

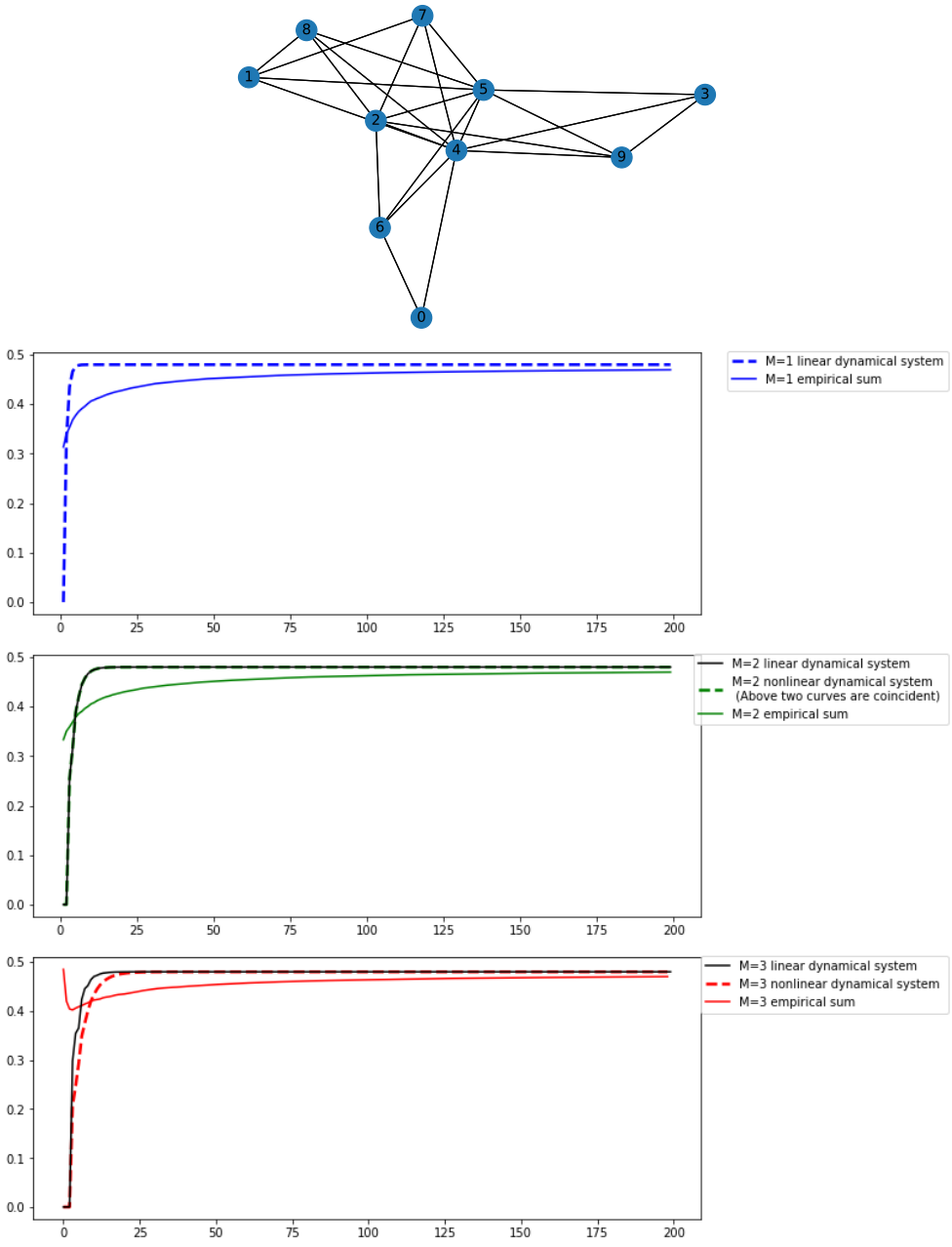


Figure 2.5: Infection rate curves for homogeneous IPCN(M, N) systems with $N = 10$ nodes and memory $M = 1, 2, 3$. We set $\rho = 0.48$, $\delta_r = \delta_b = 0.44$ for all the urns in the network. For simplicity, we set the initial values $P_i(0), P_i(1), \cdots, P_i(M-1)$ all equal to zero for all urns i.

this case the linear system is exact.

We next observe that the nonlinear dynamical system (2.24) is always a good approximation for the IPCN(M, N) system.

Note that in (2.24), the order of the approximating nonlinear dynamical system is equal to the memory of the IPCN(M, N) system and therefore, when we drop nonlinear terms from (2.24) to obtain the linear approximation (2.27), as we expect, the approximation gets worse. For $M > 1$, we can see this worsening of linear approximation in Figure 2.2 and Figure 2.4. However, in some exceptional cases, the linear approximation performs well. An example of this behavior is presented in Figure 2.3, where the linear approximations perform as well as the nonlinear ones. An important aspect of these simulations is that the reinforcement parameters play a major role in determining the asymptotic value of the probability of infection. For example in Figure 2.2, since the δ_r parameters are significantly larger than the δ_b parameters (i.e., infection is much more likely than recovery), the asymptotic value of the plots is higher (i.e., the urns tend towards having a larger composition of red balls). Similarly in Figure 2.3, since the δ_b values are significantly larger than the δ_r values, the asymptotic value of the plots are lower (i.e., the urns tend towards having a larger proportion of black balls). Furthermore, the better performance of the linear system observed in Figure 2.3 relative to Figure 2.2 and Figure 2.4 is attributed to the fact that the constant term in the linear approximation (given by (2.27)) increases when δ_r is increased and decreases when δ_b is increased. Depending on how large δ_r is, the probability of infection as approximated by (2.27) can exceed 1 and hence the linear approximation does not perform well for these cases. Whereas, no matter how large δ_b gets, the probability of infection never gets smaller than 0 and hence the

linear approximation performs comparatively better in this case.

Lastly, we observe from the simulations for the homogeneous IPCN(M, N) system in Figure 2.5 that the empirical sum as well as the linear and nonlinear dynamical approximations converge to ρ irrespective of the memory of the system. This phenomenon is indeed shown in Theorem 2.4 for any homogeneous IPCN(M, N) system.

Chapter 3

A Consensus Model based on Finite Memory Pólya urns

In Chapter 2, we formulated a network of finite memory Pólya urns interacting via a constant interaction matrix. In this chapter, we introduce a finite memory interacting Pólya urn process over a connected network which models consensus dynamics for interacting individuals. More specifically, each urn (individual) in the network is initially equipped with some red and black balls, whose proportions correspond to the individual's opinion (or belief) on a certain color. Unlike the previous chapter, we consider the interactions between urns to be time dependent, for which we use the concept of "super-urns" [12, 34, 35] to define the draw variables. As defined in Chapter 1, for an undirected network of urns, a super-urn of an urn (say urn i) consists of all the balls present in urn i and its neighbouring urns. At each time t, a ball is drawn from the super-urn of each urn i simultaneously and we define the draw

variables as follows:

$$Z_{i,t} = \begin{cases} 1 & \text{if a red ball is drawn from the super urn of urn } i \\ 0 & \text{if a black ball is drawn from the super urn of urn } i. \end{cases} \tag{3.1}$$

We can write (3.1) in terms of coefficients of a time dependent interaction matrix (similar to (2.3) in Chapter 2) as follows:

$$Z_{i,t} = \begin{cases} 1 & \text{w.p.} \quad \sum_{j=1}^{N} s_{ij}(t-1)U_{j,t-1} \\ 0 & \text{w.p.} \quad 1 - \sum_{j=1}^{N} s_{ij}(t-1)U_{j,t-1}, \end{cases} \tag{3.2}$$

such that

$$s_{ij}(t-1) = \frac{a_{ij}X_{j,t-1}}{\sum_{k=1}^{N} a_{ik}X_{k,t-1}},$$

where a_{ij} is the (i,j)th entry of the adjacency matrix of the network and $X_{k,t}$ is the total number of balls (red + black) in urn k at time t. Note that, unlike the IPCN(M,N) system which is established in Chapter 2, the interaction matrix is time-varying here.

We also add an extra reinforcement to the urns of removing all the initial balls (i.e., all balls which were present in the urns at time $t = 0$) at time $t = M$. This extra removal of initial balls signifies that agents completely forget about their initial beliefs at the Mth time step. Taking this extra reinforcement into account, the ratio of red balls in urn i at time t (given by $U_{i,t}$) as formulated in (2.5) in Chapter 2 is given by:

$$U_{i,t} = \frac{R_i + \sum_{k=1}^{t-1} \Delta_{r,i}(k)Z_{i,k}}{T_i + \sum_{n=1}^{t-1} \Delta_{r,i}(n)Z_{i,n} + \sum_{n=1}^{t-1}(1 - \Delta_{b,i}(n))Z_{i,n}} \qquad \text{for} \quad t \leq M, \tag{3.3}$$

$$U_{i,t} = \frac{\sum_{k=t-M}^{t-1} \Delta_{r,i}(k) Z_{i,k}}{\sum_{n=t-M}^{t-1} \Delta_{r,i}(n) Z_{i,n} + \sum_{n=t-M}^{t-1} \Delta_{b,i}(n)(1 - Z_{i,n})} \qquad \text{for} \quad t \geq M+1. \tag{3.4}$$

almost surely. Furthermore, similar to the IPCN(M, N) system in Chapter 2, the draw variables for this model are also conditionally independent given all the past draws in the network. Defining $Z_t = (Z_{1,t}, Z_{2,t}, \cdots, Z_{N,t})$ as the network wide draw tuple, we arrive at the following result.

Lemma 3.1. *The stochastic process $\{Z_t\}_{t=1}^{\infty}$ is a time-varying Mth order Markov chain.*

Proof. Let $a_t = (a_{1,t}, \cdots, a_{N,t}) \in \{0,1\}^N$. Using (3.4) and by virtue of the conditional independence stated in (2.4) in Chapter 2, for $t \geq M+1$ we have that

$$P(Z_{t+1} = a_{t+1} | Z_t = a_t, \cdots, Z_1 = a_1) =$$
$$\prod_{i=1}^{N} \left(\frac{a_{i,t+1} \Big(\sum_{j \in \mathcal{N}_i'} \sum_{k=t-M+1}^{t} \Delta_{r,j}(k) a_{j,k} \Big) + (1 - a_{i,t+1}) \Big(\sum_{j \in \mathcal{N}_i'} \sum_{k=t-M+1}^{t} \Delta_{b,j}(k)(1 - a_{j,k}) \Big)}{\sum_{j \in \mathcal{N}_i'} \sum_{n=t-M+1}^{t} (\Delta_{r,j}(n) a_{j,n} + \Delta_{b,j}(n)(1 - a_{j,n}))} \right) \tag{3.5}$$

where $\mathcal{N}_i'$ is the set of all neighbours of urn i and the urn i itself. As a result, we have that for all $t \geq M+1$,

$$P(Z_{t+1} = a_{t+1} | Z_t = a_t, \cdots, Z_1 = a_1) = P(Z_{t+1} = a_{t+1} | Z_t = a_t, \cdots, Z_{t-M+1} = a_{t-M+1}). \tag{3.6}$$

Hence the process $\{Z_t\}_{t=1}^{\infty}$ is a time-varying Mth order Markov chain. □

3.1 Consensus in General Networks

In this section, we obtain the consensus result for our connected network of finite memory Pólya urns (which we denote by $\mathcal{G}_N$). To begin with, we present the definition of consensus among agents in an interconnected population.

Definition 3.2. *If V represents a group of agents who can access beliefs of a limited number of agents in V, prescribed by a graph $\mathcal{G}$ with vertex set V of size N, and $x_i(t) \in \mathbb{R}$ represents the belief of agent $i \in V$ at time t, then consensus is achieved when*

$$\lim_{t\to\infty} |x_i(t) - x_j(t)| = 0$$

for all $i, j \in V$.

Setting $\mathbf{W}_t := (Z_t, Z_{t+1}, \cdots, Z_{t+M-1})$, Lemma 3.1 states that $\{\mathbf{W}_t\}_{t=1}^{\infty}$ is a Markov chain of order one. Note that for a network of size N and memory M, the Markov chain $\{\mathbf{W}_t\}_{t=1}^{\infty}$ has 2^{MN} states. Before stating the next theorem, we define the following:

Recall that $\mathcal{N}_i'$ is the set of all neighbours of urn i and the urn i itself. We define

$$\mathcal{N}_i^{(\ell)} := \bigcup_{k\in\mathcal{N}_i^{(\ell-1)}} \mathcal{N}_k^{(\ell-1)},$$

where $\ell \geq 1$, and $\mathcal{N}_k^{(0)} := \mathcal{N}_k'$.

Note that in a connected network of urns $\mathcal{G}_N$, for every urn $i \in \{1, 2, \cdots, N\}$, there exists $n \geq 1$ such that

$$\mathcal{N}_i' \cup \mathcal{N}_i^{(1)} \cup \mathcal{N}_i^{(2)} \cup \cdots \cup \mathcal{N}_i^{(n)} = \mathcal{G}_N. \tag{3.7}$$

We now use (3.7) to classify the states of the Markov chain $\{\mathbf{W}_t\}_{t=1}^{\infty}$ as either absorbing or transient.

Theorem 3.3. *The Markov chain $\{\mathbf{W}_t\}_{t=1}^{\infty}$ has two absorbing states which are state $\mathbf{0}$ with all entries zero and state $\mathbf{1}$ with all entries one. The remaining states are transient, i.e., $\{\mathbf{W}_t\}_{t=1}^{\infty}$ is an absorbing Markov chain.*

Proof. We denote a state of the Markov chain $\{\mathbf{W}_t\}_{t=1}^{\infty}$ by the following length-NM tuple

$$a := \big((a_{11}, a_{21}, \cdots, a_{N1}), \cdots, (a_{1M}, a_{2M}, \cdots, a_{NM})\big)$$

where $a \in \{0,1\}^{NM}$. Let $\mathbf{0}$ (resp., $\mathbf{1}$) be the state for which $a_{ij} = 0$ (resp., $a_{ij} = 1$) for all $i \in \{1, \cdots, N\}$ and $j \in \{1, 2, \cdots, M\}$. Using (3.5), we obtain that

$$P(\mathbf{W}_{t+1} = \mathbf{0} \,|\, \mathbf{W}_t = \mathbf{0}) = \prod_{i=1}^{N} P(Z_{i,t+1} = 0 \,|\, \mathbf{W}_t = \mathbf{0}) = 1.$$

Similarly, $P(\mathbf{W}_{t+1} = \mathbf{1} \,|\, \mathbf{W}_t = \mathbf{1}) = 1$. Hence $\mathbf{0}$ and $\mathbf{1}$ are both absorbing states of the Markov chain $\{\mathbf{W}_t\}_{t=1}^{\infty}$. We now show that the remaining states of $\mathbf{W}_t$ are transient. It is enough to show that for any state $b \notin \{\mathbf{0}, \mathbf{1}\}$, there exists a time t_b such that

$$P(\mathbf{W}_{t_b} = \mathbf{0} \,|\, \mathbf{W}_1 = b) > 0. \tag{3.8}$$

To show this, we construct a finite length path from state b to state $\mathbf{0}$ which occurs with positive probability.

- Suppose the Markov chain is in state $\mathbf{W}_t = b$ at time t, with

$$b := \big((b_{11}, b_{21}, \cdots, b_{N1}), \cdots, (b_{1M}, b_{2M}, \cdots, b_{NM})\big).$$

Note that there exists a component $b_{ij} = 0$ for some $i \in \{1, 2, \cdots, N\}$ and $j \in \{1, 2, \cdots, M\}$.

- Let

$$b^{'} := \big((b^{'}_{11}, b^{'}_{21}, \cdots, b^{'}_{N1}), \cdots, (b^{'}_{1M}, b^{'}_{2M}, \cdots, b^{'}_{NM})\big)$$

be a state of the Markov chain $\{\mathbf{W}_t\}_{t=0}^{\infty}$

with $b^{'}_{kj} = b_{kj}$ for all $k \in \{1, 2, \cdots, N\}$ and $j \in \{1, 2, \cdots, M-1\}$. Also, $b^{'}_{kM} = 0$ for all $k \in \mathcal{N}_i^{'}$ and $b^{'}_{kM} = b_{kM}$ for all $k \notin \mathcal{N}_i^{'}$. We will now show that we can go from state b to $b^{'}$ in a single time step, i.e.,

$$P(\mathbf{W}_{t+1} = b^{'} \,|\, \mathbf{W}_t = b) > 0.$$

Note that

$$\begin{aligned} P(\mathbf{W}_{t+1} = b^{'} \,|\, \mathbf{W}_t = b) &= P\big(Z_{t+M} = (b^{'}_{1M}, b^{'}_{2M}, \cdots, b^{'}_{NM}) \,|\, \mathbf{W}_t = b\big) \\ &= \prod_{k=1}^{N} P(Z_{k,t+M} = b^{'}_{k,M} \,|\, \mathbf{W}_t = b). \end{aligned} \tag{3.9}$$

At time $t+M-1$, after making the draws and adding and removing corresponding balls, the super urn of $k \in \mathcal{N}_i^{'}$ contains a black ball because $Z_{i,t+j-1} = b_{i,j} = 0$ for some $j \in \{1, 2, \cdots, M\}$. At time $t + M$, it is possible to draw a black ball from the super urn of k with probability

$$P(Z_{k,t+M} = b^{'}_{KM} = 0 \,|\, \mathbf{W}_t = b) > 0.$$

For the super urn of $k \notin \mathcal{N}_i^{'}$, at time $t+M$, it is possible to draw a ball which was added to the urn k at time $t+M-1$ i.e.,

$$P(Z_{k,t+M} = b_{KM}^{'} = b_{KM} \,|\, \mathbf{W}_t = b) > 0.$$

Hence each term of the product in (3.9) is strictly positive.

- If $b^{'} = \mathbf{0}$, then we are done. Otherwise, at the next time step, i.e., at time $t+M+1$, we draw a black ball from super urns of $j \in \mathcal{N}_i^{(1)}$ (it is possible to draw a black ball from such a super urn because $Z_{k,t+M} = 0$ for all $k \in \mathcal{N}_i^{'}$). For super urns of $j \notin \mathcal{N}_i^{(1)}$, it is possible to draw a ball which was added to the urn j at time $t+M$.

- Repeating the above procedure, by the virtue of (3.7), we will eventually hit (with positive probability) the state $\mathbf{0}$ at some time.

□

Using this structure of the Markov chain $\{\mathbf{W}_t\}_{t=1}^{\infty}$, we now obtain the consensus result for our connected network $\mathcal{G}_N$ of finite memory Pólya urns.

Theorem 3.4. *For a general connected network $\mathcal{G}_N$,*

$$\lim_{t\to\infty} P(Z_{i,t} = 1) = \lim_{t\to\infty} P(Z_{j,t} = 1) \quad \textit{for all} \quad i, j \in \{1, 2, \cdots N\}.$$

Proof. We denote the limiting distributions of the Markov chain $\{\mathbf{W}_t\}_{t=1}^{\infty}$ by Π, where

the entries of Π are denoted by $\pi_{k_1 k_2 \cdots k_M}$ with

$$k_j = (k_{1j}, k_{2j}, \cdots, k_{Nj}) \in \{0,1\}^N \quad \text{for} \quad j \in \{1, 2, \cdots, M\}.$$

The subscript $k_1 k_2 \cdots k_M$ denotes the state of the Markov chain. By Theorem 3.3, the limiting distribution of $\{\mathbf{W}_t\}_{t=1}^{\infty}$ is given by $\Pi = (1-\pi, 0, \cdots, 0, \pi)$, $0 \leq \pi \leq 1$, where first and the last states of the Markov chain are the absorbing states (corresponding to states **0** and **1**, respectively). Since $\{\mathbf{W}_t\}_{t=1}^{\infty}$ is a reducible Markov chain, there is no unique limiting distribution. Also, the limiting distribution can vary depending on $\mathbf{W}_1$, i.e., the initial state of the Markov chain. The marginal limiting distribution for an urn i is given by:

$$\lim_{t\to\infty} P(Z_{i,t} = 1) = \sum_{\substack{k_{ij}=1 \\ j\in\{1,2,\cdots,M\}}} \pi_{k_1 k_2 \cdots k_M} = \pi + 0 = \pi. \tag{3.10}$$

Hence,

$$\lim_{t\to\infty} P(Z_{i,t} = 1) = \pi = \lim_{t\to\infty} P(Z_{j,t} = 1)$$

for $i, j \in \{1, \cdots, N\}$, proving the claim. □

In the proof of Theorem 3.4, we observe that the consensus value is given by π which is the asymptotic belief of each individual in the network $\mathcal{G}_N$.

Even though it is hard to analytically solve for π in terms of the initial state for a general network, we can compute the consensus value, i.e. π, for smaller networks. One such example is presented next.

Example 3.5. Consider a complete network of 3 urns each with memory $M = 1$ and

$\Delta_{r,i}(t) = \Delta_{b,i}(t) = 1$ for $i \in \{1, 2, 3\}$ and $t \geq 1$. For simplicity, we relabel the state space as following:

$$\{000\} \to 0; \{001\} \to 1; \{010\} \to 2; \{100\} \to 3$$

$$\{011\} \to 4; \{101\} \to 5; \{110\} \to 6; \{111\} \to 7$$

We compute the transition probability matrix for the Markov chain $\{\mathbf{W}_t\}_{t=1}^{\infty}$ using (3.5) as follows:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{8}{27} & \frac{4}{27} & \frac{4}{27} & \frac{4}{27} & \frac{2}{27} & \frac{2}{27} & \frac{2}{27} & \frac{1}{27} \\ \frac{8}{27} & \frac{4}{27} & \frac{4}{27} & \frac{4}{27} & \frac{2}{27} & \frac{2}{27} & \frac{2}{27} & \frac{1}{27} \\ \frac{8}{27} & \frac{4}{27} & \frac{4}{27} & \frac{4}{27} & \frac{2}{27} & \frac{2}{27} & \frac{2}{27} & \frac{1}{27} \\ \frac{1}{27} & \frac{2}{27} & \frac{2}{27} & \frac{2}{27} & \frac{4}{27} & \frac{4}{27} & \frac{4}{27} & \frac{8}{27} \\ \frac{1}{27} & \frac{2}{27} & \frac{2}{27} & \frac{2}{27} & \frac{4}{27} & \frac{4}{27} & \frac{4}{27} & \frac{8}{27} \\ \frac{1}{27} & \frac{2}{27} & \frac{2}{27} & \frac{2}{27} & \frac{4}{27} & \frac{4}{27} & \frac{4}{27} & \frac{8}{27} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

For a state i, let $q_{i,0}$ be the probability that the Markov chain starting in state i will eventually end up in state $\{000\}$. Similarly, let $q_{i,7}$ be the probability that the Markov chain starting in state i will eventually end up in state $\{111\}$. Then by symmetry of the Markov chain, we have

$$q_{1,0} = q_{2,0} = q_{3,0} \quad \text{and} \quad q_{4,0} = q_{5,0} = q_{6,0},$$

$$q_{1,7} = q_{2,7} = q_{3,7} \quad \text{and} \quad q_{4,7} = q_{5,7} = q_{6,7}.$$

Now, using the transition probability matrix we obtain the following equations:

$$q_{6,0} = \frac{1}{27} + \frac{6}{27}q_{1,0} + \frac{12}{27}q_{6,0}$$

$$q_{1,0} = \frac{8}{27} + \frac{12}{27}q_{1,0} + \frac{6}{27}q_{6,0}.$$

We further have

$$q_{1,0} + q_{1,7} = 1 \quad \text{and} \quad q_{6,0} + q_{6,7} = 1.$$

Solving above equations yields

$$q_{1,7} = \frac{1}{3} \quad \text{and} \quad q_{6,7} = \frac{2}{3}.$$

which means that if the initial state of this Markov chain is $1, 2$ or 3, then the consensus value is $\frac{1}{3}$, and if the initial state of the Markov chain is $4, 5$ or 6, then the consensus value is $\frac{2}{3}$.

3.2 Consensus in Homogeneous Networks

In this section, we present an alternate approach to show consensus for homogeneous connected networks by constructing a class of linear dynamical systems with time delay. We further derive the exact consensus value obtained in such networks by examining the asymptotic behaviour of these dynamical systems. By homogeneous here, we mean that all reinforcement parameters are identical, i.e., $\Delta_{r,i}(t) = \Delta_{b,i}(t) = \Delta$ for all

$i \in \{1, 2, \cdots, N\}$ and $t \geq 1$. However we allow the initial composition to be different among the urns (i.e., even though the individuals update their beliefs with the same reinforcement parameters, their initial beliefs can be different). Rewriting (3.5) with the homogeneous conditions, we obtain that for $t \geq M + 1$,

$$P(Z_{i,t} = 1 | Z_{t-1}, Z_{t-2}, \cdots, Z_{t-M+1}) = \frac{\sum_{j \in \mathcal{N}_i'} \sum_{n=t-M}^{t-1} Z_{j,n}}{(1 + d_i)M}. \tag{3.11}$$

where d_i is the degree of urn i in the network $\mathcal{G}_N$. Now, taking expectation of both sides with respect to the random variables Z_{t-1}, Z_{t-2}, $\cdots$, Z_{t-M+1} in (3.11), we obtain

$$P(Z_{i,t} = 1) = \frac{\sum_{j \in \mathcal{N}_i'} \sum_{n=t-M}^{t-1} P(Z_{j,n} = 1)}{(1 + d_i)M}. \tag{3.12}$$

We further define $P_i(t) := P(Z_{i,t} = 1)$, to write (3.12) as a discrete-time linear dynamical system given by

$$P_i(t) = \frac{\sum_{j \in \mathcal{N}_i'} \sum_{n=t-M}^{t-1} P_j(n)}{(1 + d_i)M}. \tag{3.13}$$

In (3.13), $P_i(t)$ depends on $P_i(t-1), \cdots, P_i(t-M)$ in a linear fashion, and therefore, in the homogeneous case, we obtain a linear dynamical system with time delay. We next write the dynamical system (3.13) in matrix form. Define

$$P(t) := (P_1(t), \cdots, P_N(t))^T,$$
$$X_{t,M} := (P(t), \cdots, P(t - M + 1))^T.$$

Let

$$B_{N,M} = \frac{1}{M}(I_N + D)^{-1}(I_N + A),$$

where I_N is the identity matrix of size N, D is a diagonal matrix for which ith diagonal entry is d_i, and A is the adjacency matrix of the connected network $\mathcal{G}_N$. We hence have the following linear dynamical system:

$$X_{t,M} = J_{N,M} X_{t-1,M} \tag{3.14}$$

where

$$J_{N,M} = \left[\begin{array}{c|c|c|c|c} B_{N,M} & B_{N,M} & B_{N,M} & \cdots & B_{N,M} \\ \hline \mathbf{I}_N & \mathbf{0}_N & \mathbf{0}_N & \cdots & \mathbf{0}_N \\ \hline \mathbf{0}_N & \mathbf{I}_N & \mathbf{0}_N & \cdots & \mathbf{0}_N \\ \hline \vdots & \vdots & \vdots & \vdots & \vdots \\ \hline \mathbf{0}_N & \mathbf{0}_N & \cdots & \mathbf{I}_N & \mathbf{0}_N \end{array}\right]$$

is a stochastic block matrix of size $NM \times NM$. It has M^2 blocks each of which is a square matrix of size N. In the matrix $J_{N,M}$, $\mathbf{I}_N$ is identity matrix of size N and $\mathbf{0}_N$ is a square matrix of size N with all entries 0.

We next establish the asymptotic behaviour of the linear dynamical system with time delay in (3.14).

Theorem 3.6. *If $\mathcal{G}_N$ is a connected homogeneous network with memory M, then we have that*

$$\lim_{t\to\infty} X_{t,M}(i) = \pi \quad \textit{for} \quad i \in \{1, 2, \cdots, NM\}, \tag{3.15}$$

where

$$\pi = \sum_{j=1}^{M} \frac{(M-j+1)}{M} \sum_{i=1}^{N} v_{1,i} X_{M,M}(i),$$

$X_{t,M}(i)$ *is the* i*th entry of the column vector* $X_{t,M}$*, and*

$$\bar{v} = ((v_{1,1}, \cdots, v_{1,N}), \cdots, (v_{M,1}, \cdots, v_{M,N}))$$

is the l_1*-normalized left eigenvector of the matrix* $J_{N,M}$ *in* (3.14) *corresponding to eigenvalue* $\lambda = 1$*. Moreover,* $(v_{1,1}, \cdots, v_{1,N})$ *is a left eigenvector of* $B_{N,1}$ *also corresponding to eigenvalue* $\lambda = 1$*.*

Proof. Since $\mathcal{G}_N$ is connected, $B_{N,M}$ is a primitive matrix. Since all the entries of $J_{N,M}$ are non-negative and all the blocks in the first row of $J_{N,M}$ are primitive, for some positive integer $k > 0$, all the blocks of $J_{N,M}^k$ will be sum of positive powers of $B_{N,M}$. Since, $B_{N,M}$ is a primitive matrix, there exists $h > k$ such that $J_{N,M}^h$ has all positive entries. Hence $J_{N,M}$ is a primitive matrix. Since $J_{N,M}$ is a stochastic and a primitive matrix, it is a transition probability matrix for an irreducible Markov chain and the normalized left eigenvector of the matrix $J_{N,M}$, which we denote by

$$\bar{v} = ((v_{1,1} \cdots, v_{1,N}), \cdots, (v_{M,1}, \cdots, v_{M,N})),$$

is the unique stationary distribution for this Markov chain. We can write (3.14) as

$$X_{t,M} = J_{N,M}^{t-M} X_{M,M}$$

Taking $t \to \infty$ in the above equation, we obtain

$$\lim_{t\to\infty} X_{t,M} = V_{NM} X_{M,M} \tag{3.16}$$

where the limit in $X_{t,M}$ is taken entry wise and V_{NM} is a square matrix of size NM with each row given by $\bar{v}$ (e.g., see [51] for the asymptotic behaviour of irreducible ergodic Markov chains).

The symmetry in the structure of the block matrix $J_{N,M}$ makes it possible to find a useful relationship between the entries of the left eigenvector of $J_{N,M}$ corresponding to eigenvalue $\lambda = 1$ in terms of the memory parameter M and the matrix B_{N1}.

The equation

$$\bar{v} J_{N,M} = \bar{v}$$

yields the following recursive relations:

$$\begin{aligned}
(v_{1,1}, \cdots, v_{1,N}) B_{N,M} + (v_{2,1}, \cdots, v_{2,N}) &= (v_{1,1}, \cdots, v_{1,N}) \\
(v_{1,1}, \cdots, v_{1,N}) B_{N,M} + (v_{3,1}, \cdots, v_{3,N}) &= (v_{2,1}, \cdots, v_{2,N}) \\
&\vdots \\
(v_{1,1}, \cdots, v_{1,N}) B_{N,M} + (v_{M,1}, \cdots, v_{M,N}) &= (v_{M-1,1}, \cdots, v_{M-1,N})
\end{aligned}$$

which upon solving yields

$$(v_{1,1}, \cdots, v_{1,N}) B_{N,M} = (v_{1,1}, \cdots, v_{1,N})$$

and

$$v_{j,i} = \frac{(M-j+1)}{M} v_{1,i},\ j = 1, \ldots, M, i = 1, \ldots, N. \tag{3.17}$$

We obtain (3.15) by substituting (3.17) in (3.16). □

3.3 Simulation Results

In this section, we present simulations to illustrate the consensus behaviour of our network of urns.[1] In Figure 3.1, we define the empirical sum of urn i at time t as

$$I_t(i) = \frac{1}{t} \sum_{n=1}^{t} Z_{i,n}. \tag{3.18}$$

For each time instant t, the empirical sum $I_t(i)$ for node i (i.e., urn or agent/individual i) is computed 100 times and the arithmetic mean value is plotted against time. We observe in Figure 3.1 that our network exhibits a consensus behaviour with the empirical sum for all the urns eventually reaching the same value (we plot the empirical sums for only 7 urns in Figure 3.1 for better visibility of the curves). In this figure, the values of the Δ_r's and Δ_b's are taken to be in the range 5 to 15.

We indeed remark that in the long run, the empirical beliefs of the urns (agents) about the red colored balls align to a value of about 20%; i.e., the agents eventually gravitate towards favoring the viewpoint represented by the black colored balls.

In Figure 3.2, we plot the trajectory of the delayed linear dynamical systems

[1] For a complete list of parameters used for generating all figures of this chapter, refer to the link: `https://www.dropbox.com/sh/ojvmeo79wbbdv3g/AAA5onqqo0TrCU7I0iteuzRGa?dl=0`

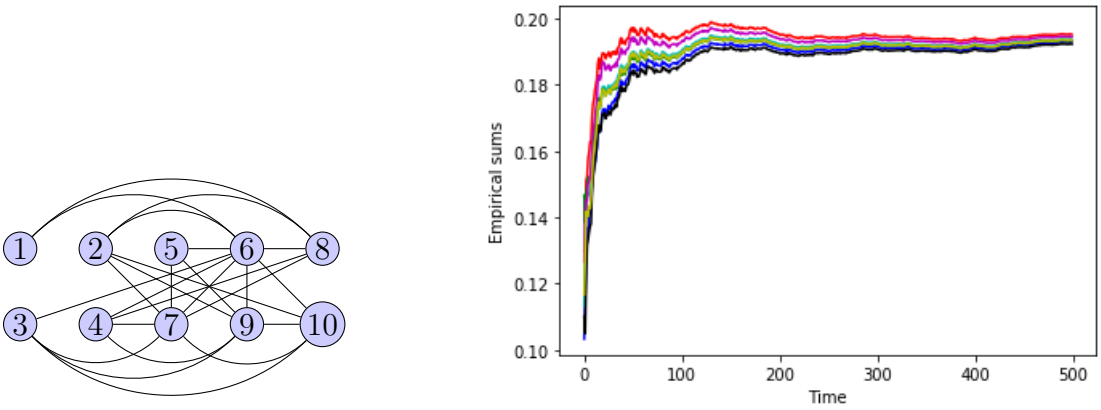


Figure 3.1: Empirical sum for first seven urns in a network with 10 urns with memory $M = 1$. Initial ratio of red balls, Δ_r's and Δ_b's are all taken to be different. We observe that asymptotically the empirical sum of all the network urns approach a consensus value.

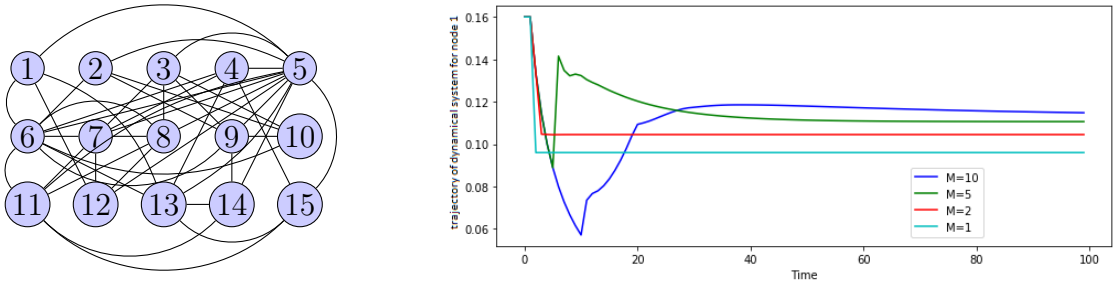


Figure 3.2: A 15-node connected homogeneous network of finite memory Pólya urns. We have $\Delta = 5$ and even through the network is homogeneous, the initial ratio of red balls are different for all the urns.

obtained in (3.13) for different values of the memory parameter M which illustrates the variation of the consensus value with memory for a connected homogeneous network. As seen in Section 3.2, we can verify the consensus value of a homogeneous connected network using Theorem 3.6 by computing the fixed points using (3.15). As an illustration, in the 15-node network of Figure 3.2, we used the following vector of initial ratio of red balls

$$\begin{aligned}\rho &= \Big(U_{1,0}, U_{2,0} \ldots, U_{15,0}\Big) \\ &= (0.16, 0.08, 0.12, 0.04, 0.04, 0.04, 0.24, 0.04, 0.16, 0.12, 0.24, 0.08, 0.16, 0.2, 0.12).\end{aligned}$$

Also for $M = 1$, letting v denote the normalized left eigenvector of the matrix $B_{15,1}$ corresponding to eigenvalue 1, we obtain (via computations carried up to the nearest two digits) that

$$v = (0.04, 0.04, 0.06, 0.06, 0.13, 0.1, 0.07, 0.06, 0.06, 0.04, 0.04, 0.04, 0.07, 0.04, 0.04).$$

Then, using (3.15), the fixed point is given by $\langle \rho, v \rangle = 0.0988$ (where $\langle \cdot, \cdot \rangle$ is the standard inner product). This fixed point is the same as shown by the light blue curve for $M = 1$ in Figure 3.2 (the same behaviour is observed for $M = 2, 5$ and 10); hence, these simulations indeed verify Theorem 3.6.

Chapter 4

A Preferential Attachment Model based on a Pólya Urn

4.1 The Model

We construct a sequence of undirected graphs $\mathcal{G}_t$, where $t \geq 0$ denotes the time index, using a Pólya reinforcement process. We start with $\mathcal{G}_0 = (V_0, \mathcal{E}_0)$, where the initial vertex and the edge set are respectively, $V_0 = \{c_1\}$ and $\mathcal{E}_0 = \{(1, 1)\}$, i.e., a self-loop on vertex 1. At each time step $t \geq 1$, a new vertex enters the graph and forms an edge with an existing vertex. The latter vertex is selected according to the draw variable of a Pólya urn with an expanding number of colors as follows:

- At time $t = 0$, the Pólya urn consists of a single ball of color c_1.
- At each time instant $t \geq 1$, we draw a ball and return it to the urn along with $\Delta_t > 0$ additional (reinforcing) balls of the same color. We also add a ball of a new color c_{t+1}. We then introduce a new vertex to the graph $\mathcal{G}_{t-1}$ (corresponding

to the color c_{t+1}) and connect it with the vertex whose color ball is drawn at time t. This results in the newly formed graph $\mathcal{G}_t$. Note that at time $t = 0$, the urn consist of only one c_1 color ball. Hence, the draw variable at time $t = 1$ is deterministic and corresponds to drawing a c_1 color ball.

At any given time instant t, we define the draw random vector

$$\mathbf{Z}_t := (Z_{1,t}, Z_{2,t}, \cdots, Z_{t,t})$$

of length t, where

$$Z_{j,t} = \begin{cases} 1 & \text{if a } c_j \text{ color ball is drawn at time } t \\ 0 & \text{otherwise} \end{cases} \qquad \text{for } 1 \leq j \leq t. \tag{4.1}$$

The vector $\mathbf{Z}_t$ is a standard unit vector for all time instances $t \geq 1$, and since at time $t = 1$ there is only c_1 color ball present in the urn, $\mathbf{Z}_1 = Z_{1,1} = 1$. We denote the "composition" of the Pólya urn at any given time instant t by the random vector

$$\mathbf{U}_t := (U_{1,t}, U_{2,t}, \cdots, U_{t+1,t}),$$

where

$$U_{j,t} = \frac{\text{Number of balls of color } c_j \text{ in the urn at time } t}{\text{Total number of balls in the urn at time } t}, \tag{4.2}$$

for $1 \leq j \leq t + 1$. In the following lemma, we express the vector $\mathbf{U}_t$ in terms of the draw variables.

Lemma 4.1. *Given $t \geq 0$, $\boldsymbol{U}_t$ is given by*

$$\boldsymbol{U}_t = \frac{1}{1 + t + \sum_{k=1}^{t} \Delta_k} \Big(1 + \sum_{n=1}^{t} \Delta_n Z_{1,n}, 1 + \sum_{n=2}^{t} \Delta_n Z_{2,n}, \cdots, 1 + \Delta_t Z_{t,t}, 1 \Big). \tag{4.3}$$

Proof. To compute the ratio in (4.2), recall that at time $n = 0$, we have one ball in the urn (this ball is of color c_1) and for each time instant $n \geq 1$, we add $\Delta_n + 1$ balls to the urn (Δ_n of the color drawn and 1 of the new color c_{n+1}). Hence the total number of balls in the urn at time t is given by $1 + \sum_{n=1}^{t} (\Delta_n + 1)$.

To determine the number of balls of color c_j in the urn after the tth draw, we note that the first c_j color ball is added to the urn at time $j - 1$. After that, at every time instant n (where $j \leq n \leq t$) at which a c_j color ball is drawn, we add Δ_n balls of c_j color to the urn. Hence, the number of balls of color c_j in the urn at time t is equal to $1 + \sum_{n=j}^{t} \Delta_n Z_{j,n}$. Therefore, the ratio of color c_j balls in the urn at time t in (4.2) is given by:

$$U_{j,t} = \frac{1 + \sum_{n=j}^{t} \Delta_n Z_{j,n}}{1 + t + \sum_{k=1}^{t} \Delta_k} \quad \text{for } 1 \leq j \leq t + 1, \tag{4.4}$$

which yields (4.3). □

Remark 4.2. *As expected, the sum of the components of* $\mathbf{U}_t$ *in* (4.3) *is one for all* $t \geq 0$. *To see this we first note that, for* $t = 0$, $\mathbf{U}_0 = U_{1,0} = 1$. *For any time* $t \geq 1$, *we*

have the following from (4.3)*:*

$$\begin{aligned}\sum_{j=1}^{t+1} U_{j,t} &= \frac{1}{1+t+\sum_{k=1}^{t}\Delta_k}((t+1)+\sum_{n=1}^{t}\Delta_n Z_{1,n}+\sum_{n=2}^{t}\Delta_n Z_{2,n}+\cdots+\Delta_t Z_{t,t})\\ &= \frac{1}{1+t+\sum_{k=1}^{t}\Delta_k}((t+1)+\sum_{i=1}^{t}\sum_{n=i}^{t}\Delta_n Z_{i,t})\\ &= \frac{1}{1+t+\sum_{k=1}^{t}\Delta_k}((t+1)+\sum_{n=1}^{t}\Delta_n\sum_{i=n}^{t} Z_{i,t}) \qquad (4.5)\end{aligned}$$

but since $\mathbf{Z}_t$ *is a standard unit vector for all* $t \geq 1$*, the right-hand side of* (4.5) *simplifies as follows:*

$$\sum_{j=1}^{t+1} U_{j,t} = \frac{1}{1+t+\sum_{k=1}^{t}\Delta_k}((t+1)+\sum_{n=1}^{t}\Delta_n) = 1.$$

An illustration of our model is given in Figure 4.1 where we show a sample path of the random vectors $\mathbf{U}_t$ and $\mathbf{Z}_t$, for $t \geq 4$ and with $\Delta_t = 2$. We further write the conditional probabilities of the draw variables given the past. More specifically, for $1 \leq j \leq t$, using (4.4), we have that

$$\begin{aligned}P(\mathbf{Z}_t = \mathbf{e}_{j,t}\,|\,\mathbf{Z}_{t-1},\cdots,\mathbf{Z}_1) &= P(Z_{j,t}=1\,|\,\mathbf{Z}_{t-1},\cdots,\mathbf{Z}_1)\\ &= P(\text{a } c_j \text{ color ball is drawn at time } t\,|\,\mathbf{Z}_{t-1},\mathbf{Z}_{t-2},\cdots,\mathbf{Z}_1)\\ &= U_{j,t-1} = \frac{1+\sum_{n=j}^{t-1}\Delta_n Z_{j,n}}{1+(t-1)+\sum_{k=1}^{t-1}\Delta_k}, \qquad (4.6)\end{aligned}$$

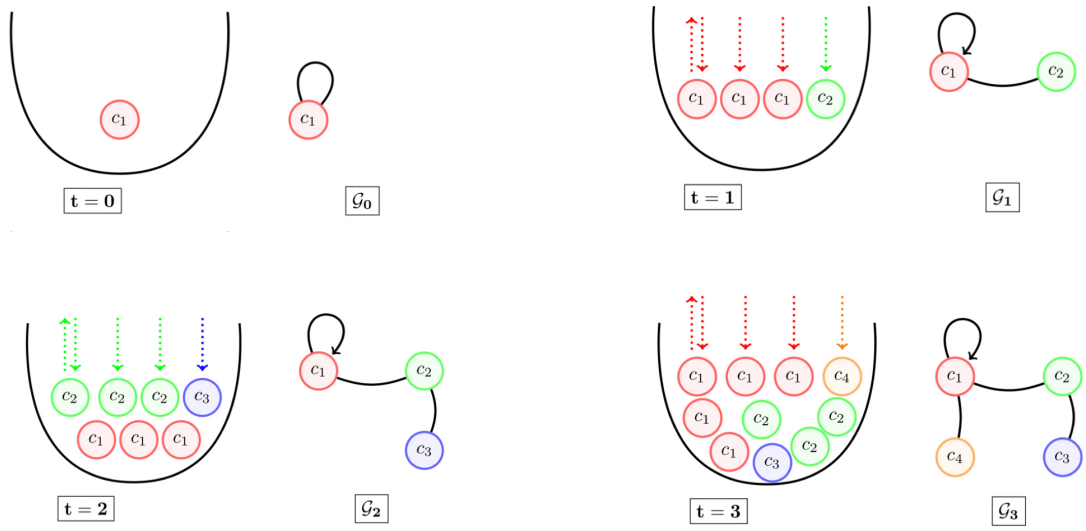


Figure 4.1: We illustrate a sample path for constructing a preferential attachment graph using an expanding color Pólya urn with $\Delta_t = 2$. For $t = 0$, the urn has only one ball of color c_1. This urn corresponds to $\mathcal{G}_0$ and $\mathbf{U}_0 = U_{1,0} = 1$. For $t = 1$, the c_1 color ball is drawn from and returned to the urn (i.e., $\mathbf{Z}_1 = Z_{1,1} = 1$). Two additional c_1 color balls are added to the urn along with a new c_2 color ball and so $\mathbf{U}_1 = (3/4, 1/4)$. For $t = 2$, a c_2 color ball is drawn from and returned to the urn (i.e., $\mathbf{Z}_2 = (0, 1)$). Two additional c_2 color balls are added to the urn along with a new c_3 color ball; hence $\mathbf{U}_2 = (3/7, 3/7, 1/7)$. For $t = 3$, a c_1 color ball is drawn from and returned to the urn (i.e., $\mathbf{Z}_3 = (1, 0, 0)$). Two additional c_1 color balls are added along with a new c_4 color ball; thus $\mathbf{U}_3 = (5/10, 3/10, 1/10, 1/10)$.

where $\mathbf{e}_{j,t}$ represents a standard unit vector of length t whose jth component is 1. Considering the case where $j = t$ in (4.6), we obtain

$$U_{t,t-1} = \frac{1 + \sum_{n=t}^{t-1} \Delta_n Z_{t,n}}{t + \sum_{k=1}^{t-1} \Delta_k} = \frac{1}{t + \sum_{k=1}^{t-1} \Delta_k} = P(Z_{t,t} = 1)$$

and hence

$$P(Z_{t,t} = 1 \,|\, \mathbf{Z}_{t-1}, \mathbf{Z}_{t-2}, \cdots, \mathbf{Z}_1) = P(Z_{t,t} = 1) = \frac{1}{t + \sum_{k=1}^{t-1} \Delta_k}, \tag{4.7}$$

i.e., the conditional probability of drawing a ball of color c_t at time t equals the marginal probability of drawing a ball of color c_t at time t. Similarly, we have that

$$P(Z_{t,t} = 0 \,|\, \mathbf{Z}_{t-1}, \mathbf{Z}_{t-2}, \cdots, \mathbf{Z}_1) = P(Z_{t,t} = 0) = \frac{t - 1 + \sum_{n=1}^{t-1} \Delta_n}{t + \sum_{k=1}^{t-1} \Delta_k}, \tag{4.8}$$

implying that $Z_{t,t}$ is independent of the random vectors $\{\mathbf{Z}_{t-1}, \mathbf{Z}_{t-2}, \cdots, \mathbf{Z}_1\}$. More generally, we obtain the marginal probability for the random variable $Z_{j,t}$ for any $1 \le j \le t$ by taking expectation on both sides in (4.6) with respect to the random vectors $\mathbf{Z}_{t-1}, \mathbf{Z}_{t-2}, \cdots, \mathbf{Z}_1$ as follows:

$$\begin{aligned} P(Z_{j,t} = 1) = \mathbb{E}(U_{j,t-1}) &= \frac{1 + \sum_{n=j}^{t-1} \Delta_n P(Z_{j,n} = 1)}{t + \sum_{k=1}^{t-1} \Delta_k} \\ &= 1 - P(Z_{j,t} = 0), \quad \text{for} \quad 1 \le j \le t. \end{aligned} \tag{4.9}$$

For $j = t$, the formula in (4.9) for $P(Z_{t,t} = 1)$ reduces to (4.7), but for $j < t$, the formula for $P(Z_{j,t} = 1)$ is a recursive function of the marginal probabilities of past

draw variables: $P(Z_{j,1} = 1), \cdots, P(Z_{j,t-1} = 1)$.

We further note that, for graph $\mathcal{G}_t$, the edge between the new vertex to one of the existing vertices in $\mathcal{G}_{t-1}$ is made using the realization of the draw vector $\mathbf{Z}_t$. Using (4.6), we observe that the conditional probability $P(\mathbf{Z}_t = \mathbf{e}_{j,t} | \mathbf{Z}_{t-1}, \cdots, \mathbf{Z}_1)$ can be written in terms of the draw variables $Z_{j,j}, \cdots, Z_{j,t-1}$. Hence, all the spatial information of the graph $\mathcal{G}_t$ is *encoded* in the sequence of random draw vectors $\{\mathbf{Z}_1, \ldots, \mathbf{Z}_{t-1}, \mathbf{Z}_t\}$. We illustrate this property in the following example, where we retrieve the graph $\mathcal{G}_4$ using $\{\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3, \mathbf{Z}_4\}$.

Example 4.3. *Consider the following realizations for the random draw vectors* $\mathbf{Z}_1$, $\mathbf{Z}_2$, $\mathbf{Z}_3$ *and* $\mathbf{Z}_4$ *for all* $t \geq 1$*:*

$$\mathbf{Z}_1 = 1, \qquad \mathbf{Z}_2 = (1, 0),$$
$$\mathbf{Z}_3 = (0, 1, 0), \qquad \mathbf{Z}_4 = (0, 1, 0, 0).$$

By construction, the graph $\mathcal{G}_0$ *consists of only one vertex* c_1 *with a self-loop. Since we start with only one ball of color* c_1 *in the urn, the random variable* $\mathbf{Z}_1 = 1$ *is deterministic and results in an edge drawn between the* c_1 *vertex and the new incoming vertex* c_2*. For* $t = 2$*, since* $\mathbf{Z}_2 = (1, 0)$*, the new incoming vertex* c_3 *is connected to* c_1*. Similarly for* $t = 3$*, the new vertex* c_4 *is connected to* c_2 *and finally, for* $t = 4$*, using the value of* $\mathbf{Z}_4$*, we connect* c_5 *to* c_2*. Hence, the graph* $\mathcal{G}_4$ *is as shown in Figure 4.2.*

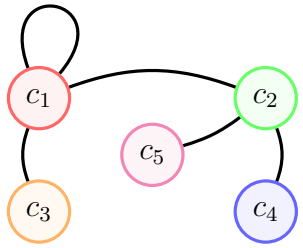


Figure 4.2: An illustration of how the sequence of draw vectors $\{\mathbf{Z}_4 = (0,1,0,0), \mathbf{Z}_3 = (0,1,0), \mathbf{Z}_2 = (1,0), \mathbf{Z}_1 = 1\}$ determines $\mathcal{G}_4$.

4.2 Analysing the degree count of the vertices in $\mathcal{G}_t$

The goal of this section is to establish a formula for the probability distribution of degree count of a fixed vertex in the preferential attachment graph constructed via our modified Pólya process until time t (including time t). We obtain this formula by writing the degree of a fixed vertex at time t in terms of the total number of balls of color corresponding to this vertex drawn until time t. To this end, let random variable $N_{j,t}$ count the number of draws of color c_j from the urn until time t (including time t). Since by construction, for a fixed color c_j, at any time $t \geq j-1$ the degree of vertex c_j at time t, denoted by $d_{j,t}$, is one more than the number of times a c_j color ball is drawn from the urn until time t; the additional one here is due to the fact that at each time instant, the new vertex which is added has degree one. Therefore,

$$d_{j,t} = 1 + N_{j,t} \quad \text{for all } 1 \leq j \leq t+1,$$

where $N_{t+1,t} = 0$. Also, note that the first time a color c_j ball can be drawn is at time j. In the following theorem, we establish an analytical expression for the marginal

probability of random variable $N_{j,t}$.

Theorem 4.4. *Fix $t \geq 1$. For a color c_j, $1 \leq j \leq t$, we have that*

$$P(N_{j,t} = k) = \tag{4.10}$$

$$\begin{cases} \displaystyle\sum_{(i_1,\cdots,i_k)\in\mathcal{A}_{j,k}^{(t)}} \frac{\prod_{a=1}^{k}\left(1+\sum_{b=1}^{a-1}\Delta_{i_b}\right)\prod_{p=j,p\notin\{i_1,\cdots,i_k\}}^{t}\left((p-1)+\sum_{l=1,l\notin\{i_1,\cdots,i_k\}}^{p-1}\Delta_l\right)}{\prod_{n=j-1}^{t-1}\left((n+1)+\sum_{m=1}^{n}\Delta_m\right)} & \text{for } 1 \leq k \leq t-j+1 \\ \dfrac{\prod_{p=j}^{t}\left((p-1)+\sum_{l=1}^{p-1}\Delta_l\right)}{\prod_{n=j-1}^{t-1}\left((n+1)+\sum_{m=1}^{n}\Delta_m\right)} & \text{for } k=0, \end{cases}$$

where

$$\mathcal{A}_{j,k}^{(t)} = \begin{cases} \{(i_1, i_2, \cdots, i_k) \mid 1 = i_1 < i_2 < \cdots < i_k \leq t\} & \text{for } j = 1 \\ \{(i_1, i_2, \cdots, i_k) \mid j \leq i_1 < i_2 < \cdots < i_k \leq t\} & \text{for } 1 < j \leq t. \end{cases} \tag{4.11}$$

Proof. Note that the set $\mathcal{A}_{j,k}^{(t)}$ defined in (4.11) gives all possible ways in which k elements can be chosen from a set of $t-j+1$ consecutive integers. In the context of our model, this set represents all possible k length tuples of time instants such that a color c_j ball is drawn at each of these time instants. For $j = 1$, the first draw at $t = 1$ is deterministic and hence $i_1 = 1$ for $j = 1$ as given in (4.11). For $1 \leq k \leq t-j+1$, we have the following:

$$
\begin{aligned}
&P(N_{j,t} = k) \\
&= \sum_{(i_1,\cdots,i_k)\in\mathcal{A}_{j,k}^{(t)}} P(Z_{j,j} = 0, \cdots, Z_{j,i_1-1} = 0, Z_{j,i_1} = 1, Z_{j,i_1+1} = 0, \cdots Z_{j,i_2-1} = 0, \\
&Z_{j,i_2} = 1, Z_{j,i_2+1} = 0, \cdots, Z_{j,i_3-1} = 0, Z_{j,i_3} = 1, \cdots, Z_{j,i_k-1} = 0, Z_{j,i_k} = 1, \\
&Z_{j,i_k+1} = 0, \cdots, Z_{j,t} = 0) \\
&= \sum_{(i_1,\cdots,i_k)\in\mathcal{A}_{j,k}^{(t)}} [P(Z_{j,t} = 0 | Z_{j,j} = 0, \cdots, Z_{j,i_1-1} = 0, Z_{j,i_1} = 1, Z_{j,i_1+1} = 0, \cdots, \\
&Z_{j,i_2-1} = 0, Z_{j,i_2} = 1, Z_{j,i_2+1} = 0, \cdots, Z_{j,i_3-1} = 0, Z_{j,i_3} = 1, \cdots, Z_{j,i_k-1} = 0, \\
&Z_{j,i_k} = 1, Z_{j,i_k+1} = 0, \cdots, Z_{j,t-1} = 0) P(Z_{j,j} = 0, \cdots, Z_{j,i_1-1} = 0, Z_{j,i_1} = 1, \\
&Z_{j,i_1+1} = 0, \cdots, Z_{j,i_2-1} = 0, Z_{j,i_2} = 1, Z_{j,i_2+1} = 0, Z_{j,i_3-1} = 0, Z_{j,i_3} = 1, \\
&\cdots, Z_{j,i_k-1} = 0, Z_{j,i_k} = 1, Z_{j,i_k+1} = 0, \cdots, Z_{j,t-1} = 0)]. \qquad (4.12)
\end{aligned}
$$

By substituting (4.6) in the conditional probability expressions in (4.12), we obtain the following:

$$
\begin{aligned}
&P(Z_{j,t} = 0 | Z_{j,j} = 0, \cdots, Z_{j,i_1-1} = 0, Z_{j,i_1} = 1, Z_{j,i_1+1} = 0, \cdots, Z_{j,i_2-1} = 0, Z_{j,i_2} = 1, \\
&Z_{j,i_2+1} = 0, \cdots, Z_{j,i_3-1} = 0, Z_{j,i_3} = 1, \cdots, Z_{j,i_k-1} = 0, Z_{j,i_k} = 1, Z_{j,i_k+1} = 0, \\
&\cdots, Z_{j,t-1} = 0) \\
&= 1 - \frac{1 + \sum_{n=j}^{t-1} \Delta_n Z_{j,n}}{t + \sum_{m=1}^{t-1} \Delta_m}
\end{aligned}
$$

$$= 1 - \frac{1 + \sum_{l=1}^{k} \Delta_{i_l}}{t + \sum_{m=1}^{t-1} \Delta_m}$$

$$= \frac{(t-1) + \sum_{l=1, l \notin \{i_1, \cdots, i_k\}}^{t-1} \Delta_l}{t + \sum_{m=1}^{t-1} \Delta_m}. \tag{4.13}$$

Now, substituting the conditional probability expression obtained in (4.13) in (4.12), yields

$$P(N_{j,t} = k) = \sum_{(i_1, \cdots, i_k) \in \mathcal{A}_{j,k}^{(t)}} \frac{((t-1) + \sum_{l=1, l \notin \{i_1, \cdots, i_k\}}^{t-1} \Delta_l)}{t + \sum_{m=1}^{t-1} \Delta_m} P(Z_{j,j} = 0, \cdots, Z_{j,i_1 - 1} = 0,$$

$$Z_{j,i_1} = 1, Z_{j,i_1+1} = 0, \cdots, Z_{j,i_2-1} = 0, Z_{j,i_2} = 1, Z_{j,i_2+1} = 0, \cdots, Z_{j,i_3-1} = 0, Z_{j,i_3} = 1,$$

$$\cdots, Z_{j,i_k-1} = 0, Z_{j,i_k} = 1, Z_{j,i_k+1} = 0, \cdots, Z_{j,t-1} = 0). \tag{4.14}$$

Similar to (4.12) and (4.14), we continue to recursively write the joint probability as a product of conditional and marginal probabilities and substitute the expressions for the conditional probability using (4.13) as follows:

$$P(N_{j,t} = k)$$

$$= \sum_{(i_1, \cdots, i_k) \in \mathcal{A}_{j,k}^{(t)}} \left(\frac{(t-1) + \sum_{l=1, l \notin \{i_1, \cdots, i_k\}}^{t-1} \Delta_l}{t + \sum_{m=1}^{t-1} \Delta_m} \right) \left(\frac{(t-2) + \sum_{l=1, l \notin \{i_1, \cdots, i_k\}}^{t-2} \Delta_l}{t - 1 + \sum_{m=1}^{t-2} \Delta_m} \right) P(Z_{j,j} = 0,$$

$$\cdots, Z_{j,i_1-1} = 0, Z_{j,i_1} = 1, Z_{j,i_1+1} = 0, \cdots, Z_{j,i_2-1} = 0, Z_{j,i_2} = 1, Z_{j,i_2+1} = 0,$$

$$\cdots, Z_{j,i_3-1} = 0, Z_{j,i_3} = 1, \cdots, Z_{j,i_k-1} = 0, Z_{j,i_k} = 1, Z_{j,i_k+1} = 0, \cdots, Z_{j,t-2} = 0)$$

$$\vdots$$

$$= \sum_{(i_1,\cdots,i_k)\in\mathcal{A}_{j,k}^{(t)}} \Bigg[\left(\frac{(t-1)+\sum_{l=1,l\notin\{i_1,\cdots,i_k\}}^{t-1}\Delta_l}{t+\sum_{m=1}^{t-1}\Delta_m} \right) \left(\frac{(t-2)+\sum_{l=1,l\notin\{i_1,\cdots,i_k\}}^{t-2}\Delta_l}{t-1+\sum_{m=1}^{t-2}\Delta_m} \right) \cdots$$

$$\left(\frac{i_k+\sum_{l=1,l\notin\{i_1,\cdots,i_k\}}^{i_k}\Delta_l}{i_k+1+\sum_{m=1}^{i_k}\Delta_m} \right) \left(\frac{1+\sum_{l=1}^{k-1}\Delta_{i_l}}{i_k+\sum_{m=1}^{i_k-1}\Delta_m} \right) \left(\frac{i_k-2+\sum_{l=1,l\notin\{i_1,\cdots,i_k\}}^{i_k-2}\Delta_l}{i_k-1+\sum_{m=1}^{i_k-2}\Delta_m} \right) \cdots$$

$$\left(\frac{i_{k-1}+\sum_{l=1,l\notin\{i_1,\cdots,i_k\}}^{i_{k-1}}\Delta_l}{i_{k-1}+1+\sum_{m=1}^{i_{k-1}}\Delta_m} \right) \left(\frac{1+\sum_{l=1}^{k-2}\Delta_{i_l}}{i_{k-1}+\sum_{m=1}^{i_{k-1}-1}\Delta_m} \right) \left(\frac{i_{k-1}-2+\sum_{l=1,l\notin\{i_1,\cdots,i_k\}}^{i_{k-1}-2}\Delta_l}{i_{k-1}-1+\sum_{m=1}^{i_{k-1}-2}\Delta_m} \right)$$

$$\cdots \left(\frac{1}{i_1+\sum_{m=1}^{i_1-1}\Delta_m} \right) \left(\frac{i_1-2+\sum_{l=1,l\notin\{i_1,\cdots,i_k\}}^{i_1-2}\Delta_l}{i_1-1+\sum_{m=1}^{i_1-2}\Delta_m} \right) \cdots$$

$$\cdots \left(\frac{j+\sum_{l=1,l\notin\{i_1,\cdots,i_k\}}^{j}\Delta_l}{j+1+\sum_{m=1}^{j}\Delta_m} \right) \left(\frac{j-1+\sum_{l=1,l\notin\{i_1,\cdots,i_k\}}^{j-1}\Delta_l}{j+\sum_{m=1}^{j-1}\Delta_m} \right) \Bigg]$$

$$= \sum_{(i_1,\cdots,i_k)\in\mathcal{A}_{j,k}^{(t)}} \Bigg[\frac{\prod_{a=1}^{k}(1+\sum_{b=1}^{a-1}\Delta_{i_b})}{\prod_{n=j-1}^{t-1}((n+1)+\sum_{m=1}^{n}\Delta_m)} \Bigg(\prod_{l_1=j-1}^{i_1-2} (l_1+\sum\nolimits_{l=1,l\notin\{i_1,\cdots,i_k\}}^{l_1}\Delta_l) \Bigg)$$

$$\Bigg(\prod_{l_2=i_1}^{i_2-2} (l_2+\sum\nolimits_{l=1,l\notin\{i_1,\cdots,i_k\}}^{l_2}\Delta_l) \Bigg) \Bigg(\prod_{l_3=i_2}^{i_3-2} (l_3+\sum\nolimits_{l=1,l\notin\{i_1,\cdots,i_k\}}^{l_3}\Delta_l) \Bigg) \cdots$$

$$\cdots \Bigg(\prod_{l_k=i_{k-1}}^{i_k-2} (l_k+\sum\nolimits_{l=1,l\notin\{i_1,\cdots,i_k\}}^{l_k}\Delta_l) \Bigg) \Bigg(\prod_{l_{k+1}=i_k}^{t-1} (l_{k+1}+\sum\nolimits_{l=1,l\notin\{i_1,\cdots,i_k\}}^{l_{k+1}}\Delta_l) \Bigg) \Bigg]$$

$$= \sum_{(i_1,i_2,\cdots,i_k)\in\mathcal{A}_{j,k}^{(t)}} \frac{\prod_{a=1}^{k}(1+\sum_{b=1}^{a-1}\Delta_{i_b}) \prod_{p=j,p\notin\{i_1,\cdots,i_k\}}^{t}((p-1)+\sum_{l=1,l\notin\{i_1,\cdots,i_k\}}^{p-1}\Delta_l)}{\prod_{n=j-1}^{t-1}((n+1)+\sum_{m=1}^{n}\Delta_m)}.$$

Therefore, (4.10) holds for $1 \le k \le t-j+1$. We determine $P(N_{j,t}=0)$ as follows:

$$\begin{aligned} P(N_{j,t}=0) &= P(Z_{j,j}=0, Z_{j,j+1}=0, \cdots, Z_{j,t-1}=0, Z_{j,t}=0) \\ &= P(Z_{j,j}=0) \prod_{n=j+1}^{t} P(Z_{j,n}=0 \,|\, Z_{j,j}=0, Z_{j,j+1}=0, \cdots, Z_{j,n-1}=0). \end{aligned} \tag{4.15}$$

Now, using (4.6) for the conditional probabilities in (4.15) we obtain

$$P(N_{j,t}=0) = \left(1-\frac{1}{j+\sum_{m=1}^{j-1}\Delta_m}\right)\left(1-\frac{1}{j+1+\sum_{m=1}^{j}\Delta_m}\right)\cdots\left(1-\frac{1}{t+\sum_{m=1}^{t-1}\Delta_m}\right)$$
$$= \frac{\prod_{p=j}^{t}((p-1)+\sum_{l=1}^{p-1}\Delta_l)}{\prod_{n=j-1}^{t-1}((n+1)+\sum_{m=1}^{n}\Delta_m)}.$$

Hence (4.10) holds for $k=0$. □

The analytic formula obtained in (4.10) is quite involved when the reinforcement parameter Δ_t is time-varying. For the special case of $\Delta_t=\Delta$ for all $t\geq 1$, Theorem 4.4 reduces to the following corollary.

Corollary 4.5. *Fix $t\geq 1$. For a color c_j, $1\leq j\leq t$ and $\Delta_t=\Delta$ for all $t\geq 1$, the marginal probability for $N_{j,t}$ is given by:*

$$P(N_{j,t}=k) = \tag{4.16}$$

$$\begin{cases} \displaystyle\sum_{(i_1,i_2,\cdots,i_k)\in\mathcal{A}_{j,k}^{(t)}} \frac{\prod_{a=1}^{k}\big(1+(a-1)\Delta\big)\prod_{p=j,p\notin\{i_1,\cdots,i_k\}}^{t}\big((p-1)(\Delta+1)-\Delta\sum_{l=1}^{k}\mathbb{1}(i_l\leq p-1)\big)}{\prod_{n=j-1}^{t-1}\big((\Delta+1)n+1\big)} & \text{for } 1\leq k\leq t-j+1 \\ \displaystyle\frac{\prod_{p=j}^{t}(p-1)(\Delta+1)}{\prod_{n=j-1}^{t-1}\big((\Delta+1)n+1\big)} & \text{for } k=0, \end{cases}$$

where the set $\mathcal{A}_{j,k}^{(t)}$ is defined in (4.11) *and $\mathbb{1}(\mathcal{E})$ is the indicator function of the event $\mathcal{E}$.*

Since the first time instant at which a c_j color ball can be drawn from the Pólya urn is at time j, the total number of draws of a c_j color ball till time t can be at most $t-j+1$. Therefore,

$$\sum_{k=0}^{t-j+1} P(d_{j,t} = k+1) = \sum_{k=0}^{t-j+1} P(N_{j,t} = k) = 1,$$

which implies that $P(N_{j,t} = k)$ is a probability mass function on the support set $\{0, 1, \cdots, t-j+1\}$. We next verify that $P(N_{j,t} = k)$ obtained in (4.16) does indeed sum up to one (over k ranging from zero to $t-j+1$) and is hence a legitimate probability mass function. For simplicity, we focus on the case with $\Delta_t = \Delta$; the proof for the general case follows along similar lines. To this end, we write the set $\mathcal{A}_{j,k}^{(t)}$ as the following disjoint union:

$$\begin{aligned}
\mathcal{A}_{j,k}^{(t)} &= \{(i_1, i_2, \cdots, i_k) \mid j \le i_1 < i_2 < \cdots < i_k \le t\} \\
&= \{(i_1, i_2, \cdots, i_k) \mid j \le i_1 < i_2 < \cdots < i_k \le t-1\} \sqcup \mathcal{B}_{j,k}^{(t)} \\
&= \mathcal{A}_{j,k}^{(t-1)} \sqcup \mathcal{B}_{j,k}^{(t)}, \qquad (4.17)
\end{aligned}$$

where $\mathcal{B}_{j,k}^{(t)} := \{(i_1, \cdots, i_{k-1}, t) \mid j \le i_1 < \cdots < i_{k-1} \le t-1\}$. Note that

$$\mathcal{B}_{j,k}^{(t)} = \{(i_1, \cdots, i_{k-1}, t) \mid (i_1, \cdots, i_{k-1}) \in \mathcal{A}_{j,k-1}^{(t-1)}\}. \qquad (4.18)$$

Theorem 4.6. *Fix $t \ge 1$. For a color c_j, $1 \le j \le t$ and $\Delta_t = \Delta$ for all $t \ge 1$, we*

have that

$$\sum_{k=0}^{t-j+1} P(N_{j,t} = k) = 1. \tag{4.19}$$

Proof. We write the left-hand side of (4.19) using (4.16):

$$\frac{\prod_{p=j}^{t}(p-1)(\Delta+1)}{\prod_{n=j-1}^{t-1}((\Delta+1)n+1)} \quad + \\ \sum_{k=1}^{t-j+1} \sum_{(i_1,\cdots,i_k)\in\mathcal{A}_{j,k}^{(t)}} \frac{\prod_{a=1}^{k}(1+(a-1)\Delta) \prod_{p=j,p\notin\{i_1,\cdots,i_k\}}^{t}((p-1)(\Delta+1)-\Delta\sum_{l=1}^{k}\mathbb{1}(i_l\leq p-1))}{\prod_{n=j-1}^{t-1}((\Delta+1)n+1)} \tag{4.20}$$

Therefore, showing that (4.19) holds is equivalent to showing that:

$$\prod_{p=j}^{t}(p-1)(\Delta+1) \quad + \\ \sum_{k=1}^{t-j+1} \sum_{(i_1,\cdots,i_k)\in\mathcal{A}_{j,k}^{(t)}} \left(\prod_{a=1}^{k}\left(1+(a-1)\Delta\right) \prod_{p=j,p\notin\{i_1,\cdots,i_k\}}^{t}\left((p-1)(\Delta+1)-\Delta\sum_{l=1}^{k}\mathbb{1}(i_l\leq p-1)\right)\right) \\ = \prod_{n=j-1}^{t-1}((\Delta+1)n+1). \tag{4.21}$$

We prove (4.21) by induction on $t-j+1 \geq 1$.

Base Case: $t-j+1=1$ or $t=j$. For this case, the left-hand side of (4.21) is the following:

$$\prod_{p=j}^{j}(p-1)(\Delta+1)$$

$$+\sum_{\mathcal{A}_{j,1}^{(j)}}\Big(\prod_{a=1}^{1}\big(1+(a-1)\Delta\big)\prod_{p=j,p\neq j}^{j}\big((p-1)(\Delta+1)-\Delta\sum_{l=1}^{1}\mathbb{1}(i_l\leq p-1)\big)\Big)$$

$$=(\,j-1)(\Delta+1)+1$$

which, upon simplification and noting that the set $\mathcal{A}_{j,1}^{(j)}=\{j\}$, equals the right-hand side of (4.21) for $t-j+1=1$.

Induction Step: We now show the induction step: assuming that (4.21) is true for $t-j+1=s$, we show that it holds for $t-j+1=s+1$. We thus assume that the following holds:

$$\prod_{p=j}^{j+s-1}((p-1)(\Delta+1))$$
$$+\sum_{k=1}^{s}\sum_{\mathcal{A}_{j,k}^{(j+s-1)}}\Big(\prod_{a=1}^{k}\big(1+(a-1)\Delta\big)\prod_{p=j,p\notin\{i_1,\cdots,i_k\}}^{j+s-1}\big((p-1)(\Delta+1)-\Delta\sum_{l=1}^{k}\mathbb{1}(i_l\leq p-1)\big)\Big)$$
$$=\prod_{n=j-1}^{j+s-2}((\Delta+1)n+1). \tag{4.22}$$

We next show the induction step using (4.22), by starting from the right-hand side:

$$\prod_{n=j-1}^{j+s-1}((\Delta+1)n+1)=((\Delta+1)(s+j-1)+1)\prod_{n=j-1}^{j+s-2}((\Delta+1)n+1)$$
$$\overset{(a)}{=}((\Delta+1)(s+j-1)+1)\prod_{p=j}^{j+s-1}(p-1)(\Delta+1)$$
$$+\sum_{k=1}^{s}((\Delta+1)(s+j-1)+1)\sum_{\mathcal{A}_{j,k}^{(j+s-1)}}\Big(\prod_{a=1}^{k}\big(1+(a-1)\Delta\big)\prod_{p=j,p\notin\{i_1,\cdots,i_k\}}^{j+s-1}\big((p-1)(\Delta+1)$$
$$-\Delta\sum_{l=1}^{k}\mathbb{1}(i_l\leq p-1)\big)\Big)$$

$$
\begin{aligned}
&\overset{(b)}{=} \sum_{k=1}^{s}((\Delta+1)(s+j-1)-\Delta k+\Delta k+1)\sum_{\mathcal{A}_{j,k}^{(j+s-1)}}\Bigg(\prod_{a=1}^{k}\big(1+(a-1)\Delta\big)\\
&\qquad\qquad\prod_{p=j,p\notin\{i_1,\cdots,i_k\}}^{j+s-1}\big((p-1)(\Delta+1)-\Delta\sum_{l=1}^{k}\mathbb{1}(i_l\le p-1)\big)\Bigg)\\
&+\prod_{p=j}^{j+s}(p-1)(\Delta+1) \quad + \quad \prod_{p=j}^{j+s-1}(p-1)(\Delta+1)\\
&\overset{(c)}{=} \prod_{p=j}^{j+s}(p-1)(\Delta+1) \quad + \quad \prod_{p=j}^{j+s-1}(p-1)(\Delta+1)\\
&+\sum_{k=1}^{s}((\Delta+1)(s+j-1)-\Delta k)\sum_{\mathcal{A}_{j,k}^{(j+s-1)}}\Bigg(\prod_{a=1}^{k}\big(1+(a-1)\Delta\big)\\
&\qquad\qquad\prod_{p=j,p\notin\{i_1,\cdots,i_k\}}^{j+s-1}\big((p-1)(\Delta+1)-\Delta\sum_{l=1}^{k}\mathbb{1}(i_l\le p-1)\big)\Bigg)\\
&+\sum_{k=1}^{s}(\Delta k+1)\sum_{\mathcal{A}_{j,k}^{(j+s-1)}}\Bigg(\prod_{a=1}^{k}\big(1+(a-1)\Delta\big)\prod_{p=j,p\notin\{i_1,\cdots,i_k\}}^{j+s-1}\big((p-1)(\Delta+1)\\
&\qquad\qquad-\Delta\sum_{l=1}^{k}\mathbb{1}(i_l\le p-1)\big)\Bigg)\\
&\overset{(d)}{=} \prod_{p=j}^{j+s}(p-1)(\Delta+1) \quad + \quad \prod_{p=j}^{j+s-1}(p-1)(\Delta+1)\\
&+\sum_{k=1}^{s}\sum_{\mathcal{A}_{j,k}^{(j+s-1)}}\Bigg(\prod_{a=1}^{k}\big(1+(a-1)\Delta\big)\prod_{p=j,p\notin\{i_1,\cdots,i_k\}}^{j+s}\big((p-1)(\Delta+1)-\Delta\sum_{l=1}^{k}\mathbb{1}(i_l\le p-1)\big)\Bigg)\\
&+\sum_{k=1}^{s}\sum_{\mathcal{A}_{j,k}^{(j+s-1)}}\Bigg(\prod_{a=1}^{k+1}\big(1+(a-1)\Delta\big)\prod_{p=j,p\notin\{i_1,\cdots,i_k\}}^{j+s-1}\big((p-1)(\Delta+1)-\Delta\sum_{l=1}^{k}\mathbb{1}(i_l\le p-1)\big)\Bigg)\\
&\overset{(e)}{=} \prod_{p=j}^{j+s}(p-1)(\Delta+1) \quad + \quad \prod_{p=j}^{j+s-1}(p-1)(\Delta+1)\\
&+\sum_{k=1}^{s}\sum_{\mathcal{A}_{j,k}^{(j+s-1)}}\Bigg(\prod_{a=1}^{k}\big(1+(a-1)\Delta\big)\prod_{p=j,p\notin\{i_1,\cdots,i_k\}}^{j+s}\big((p-1)(\Delta+1)-\Delta\sum_{l=1}^{k}\mathbb{1}(i_l\le p-1)\big)\Bigg)
\end{aligned}
$$

$$
\begin{aligned}
&+\sum_{k=2}^{s+1}\sum_{\mathcal{A}_{j,k-1}^{(j+s-1)}}\Bigg(\prod_{a=1}^{k}\big(1+(a-1)\Delta\big)\prod_{p=j,p\notin\{i_1,\cdots,i_{k-1}\}}^{j+s-1}\big((p-1)(\Delta+1)-\Delta\sum_{l=1}^{k-1}\mathbb{1}(i_l\le p-1)\big)\Bigg)\\
&\overset{(f)}{=}\prod_{p=j}^{j+s}(p-1)(\Delta+1)\quad+\quad\prod_{p=j}^{j+s-1}(p-1)(\Delta+1)\\
&+\sum_{k=1}^{s}\sum_{\mathcal{A}_{j,k}^{(j+s-1)}}\Bigg(\prod_{a=1}^{k}\big(1+(a-1)\Delta\big)\prod_{p=j,p\notin\{i_1,\cdots,i_k\}}^{j+s}\big((p-1)(\Delta+1)-\Delta\sum_{l=1}^{k}\mathbb{1}(i_l\le p-1)\big)\Bigg)\\
&+\sum_{k=2}^{s+1}\sum_{\mathcal{B}_{j,k}^{(j+s)}}\Bigg(\prod_{a=1}^{k}\big(1+(a-1)\Delta\big)\prod_{p=j,p\notin\{i_1,\cdots,i_{k-1},j+s\}}^{j+s}\big((p-1)(\Delta+1)-\Delta\sum_{l=1}^{k-1}\mathbb{1}(i_l\le p-1)\big)\Bigg)\\
&\overset{(g)}{=}\prod_{p=j}^{j+s}(p-1)(\Delta+1)\quad+\quad\prod_{p=j}^{j+s-1}(p-1)(\Delta+1)\\
&+\sum_{\mathcal{A}_{j,1}^{(j+s-1)}}\Bigg(\prod_{a=1}^{1}\big(1+(a-1)\Delta\big)\prod_{p=j,p\notin\{i_1\}}^{j+s}\big((p-1)(\Delta+1)-\Delta\sum_{l=1}^{1}\mathbb{1}(i_l\le p-1)\big)\Bigg)\\
&+\sum_{k=2}^{s}\sum_{\mathcal{A}_{j,k}^{(j+s)}}\Bigg(\prod_{a=1}^{k}\big(1+(a-1)\Delta\big)\prod_{p=j,p\notin\{i_1,\cdots,i_k\}}^{j+s}\big((p-1)(\Delta+1)-\Delta\sum_{l=1}^{k}\mathbb{1}(i_l\le p-1)\big)\Bigg)\\
&+\prod_{a=1}^{s+1}(1+(a-1)\Delta)\\
&\overset{(h)}{=}\prod_{p=j}^{j+s}(p-1)(\Delta+1)+\sum_{k=1}^{s+1}\sum_{\mathcal{A}_{j,k}^{(j+s)}}\Bigg(\prod_{a=1}^{k}\big(1+(a-1)\Delta\big)\prod_{p=j,p\notin\{i_1,\cdots,i_k\}}^{j+s}\big((p-1)(\Delta+1)\\
&\qquad\qquad-\Delta\sum_{l=1}^{k}\mathbb{1}(i_l\le p-1)\big)\Bigg).
\end{aligned}
$$

In the above set of equations, we obtain (a) by substituting (4.22) in the left-hand side of (a). In (b), we add and subtract Δk to the term $((\Delta+1)(s+j-1)+1)$ and split the summation across the terms $((\Delta+1)(s+j-1)-\Delta k)$ and $(\Delta k+1)$ in (c). In (d), we absorb the terms $((\Delta k+1)(s+j-1)-\Delta k)$ and $(\Delta k+1)$ into the product. We replace k by $k-1$ in the fourth term on the left-hand side of (e). We obtain the

fourth term on the right-hand side of (f) using (4.18). On the right-hand side of (g), the second term can be written as follows:

$$\prod_{p=j}^{j+s-1}(p-1)(\Delta+1)$$
$$= \sum_{\mathcal{B}_{j,1}^{(j+s)}} \Bigg(\prod_{a=1}^{1} \big(1+(a-1)\Delta\big) \prod_{p=j, p\notin\{j+s\}}^{j+s} \big((p-1)(\Delta+1) - \Delta \sum_{l=1}^{0} \mathbb{1}(i_l \leq p-1)\big)\Bigg)$$

which is merged with the third term on the right-hand side of (g) to obtain the $k=1$ term on the right-hand side of (h). Similarly, for the terms for $k=2$ to $k=s$, we merge both of the terms on the right-hand side of (f) using (4.17) to obtain the fourth term on the right-hand side of (g). The last term on the right-hand side of (g) is evaluation of the fourth term on the right-hand side of (f) at $k=s+1$. Finally (h) is obtained by writing all the terms under one summation. Hence the proof follows from induction on $t-j+1$. □

4.3 Simulation Results

In this section, we present a comparative study[1] between our model and the Barabási-Albert model in terms of following three features:

- Structural differences in small-sized graphs;
- Degree distributions of the graphs obtained;
- Expected birth time of vertices with a fixed degree.

[1]For details about the simulations of this chapter, refer to the following link: `https://drive.google.com/drive/folders/1uOmz4B6RQ0hRmuu_CTfb02jJ03B1SyEl?usp=share_link`

We first illustrate the structural differences between the graphs generated by our model and the Barabási-Albert model. In the next set of simulations, we compare the degree distribution of both models by plotting the probability of randomly choosing a k degree vertex versus k (on a $\log - \log$ scale) for a graph generated until a fixed time instant. We give the degree distribution of graphs generated for 5000 time steps (averaged over 250 simulations) via the standard Barabási-Albert model and our model with different choices of the reinforcement parameter Δ_t and discuss the similarities and differences obtained in the degree distributions. In the third set of simulations, we compare both models in terms of vertices expected birth time versus degree, which we define as follows.

Definition 4.7. *Given a random network/graph generated until time t, we define the vertices* ***expected birth time*** *for a fixed degree k, where $1 \leq k \leq t$, as the conditional expectation of all the times when the vertices which have degree k at termination time t were introduced, given the draw vectors until time t. It is denoted by $\overline{b}_t(k)$ and is given by the following expression:*

$$\overline{b}_t(k) = \mathbb{E}\Big[\sum\nolimits_{j=1}^{t} (j-1) \mathbb{1}\big((\sum\nolimits_{n=j}^{t} Z_{j,n} = 1) = k \big) \big| \mathbf{Z}_1, \mathbf{Z}_2, \cdots, \mathbf{Z}_t \Big] \tag{4.23}$$

where $\mathbf{Z}_1, \mathbf{Z}_2, \cdots, \mathbf{Z}_t$ are the draw vectors of the Pólya urn. Note that we write $(j-1)$ in (4.23) because the vertex j is introduced in the network/graph at time $j-1$.

In the experiments, we determine the empirical version of (4.23), which we call the ***average birth time*** *and denote it by $\mathrm{b}_t(k)$ for degree k and termination time t.*

It is given by

$$\mathrm{b}_t(k) = \frac{\sum_{j=1}^{t}(j-1)\mathbb{1}\big((\sum_{n=j}^{t} Z_{j,n}=1)=k\big)}{\sum_{j=1}^{t}\mathbb{1}\big((\sum_{n=j}^{t} Z_{j,n}=1)=k\big)}. \tag{4.24}$$

In Figure 4.3, two 15-vertex networks are depicted, one generated by our model (on the left-hand side) and the other by the Barabási-Albert model (on the right-hand side). We make two observations: First, in contrast with the Barabási-Albert model, in our model all vertices are labelled by distinct colors. This one-to-one correspondence between vertices and colors encodes all the information of the generated graph in the draw vectors of the underlying Pólya urn. Second, the maximum degree achieved is higher in the left-hand side network generated via our model (which achieves a maximum degree of 11 compared to 6 in the right-hand side Barabási-Albert network). This happens along one sample path due to the choice of reinforcement parameter (here $\Delta_t = 5$ for all $t \geq 1$) in our model which allows for an already selected vertex to be chosen with a higher probability than in the case of the Barabási-Albert model where the vertices are chosen proportional to their degree.

In Figures 4.4 and 4.5, we plot the degree distribution (averaged over 250 simulations) of networks generated for four different choices of Δ_t: $1, \ln(t), f(t), g(t)$, where the functions $f(t)$ and $g(t)$ are defined in (4.25). We observe the deviation of the degree distribution of the graphs generated via our model for the above mentioned choices of Δ_t from the degree distribution of the Barabási-Albert network which follows the relation $p(k) \sim k^{-3}$, where $p(k)$ is the probability of randomly choosing a vertex of degree k in the network. As observed in Figure 4.4 (a), the slopes of the degree

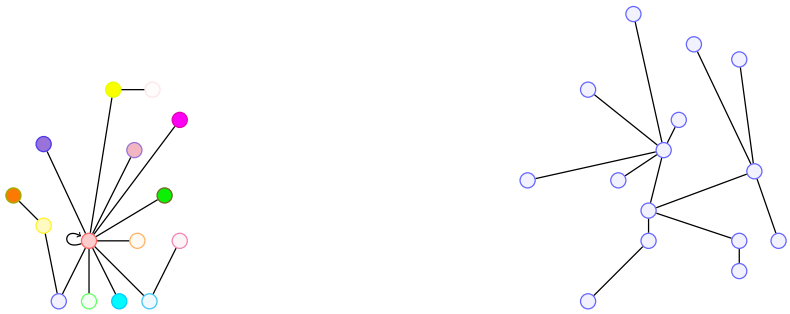
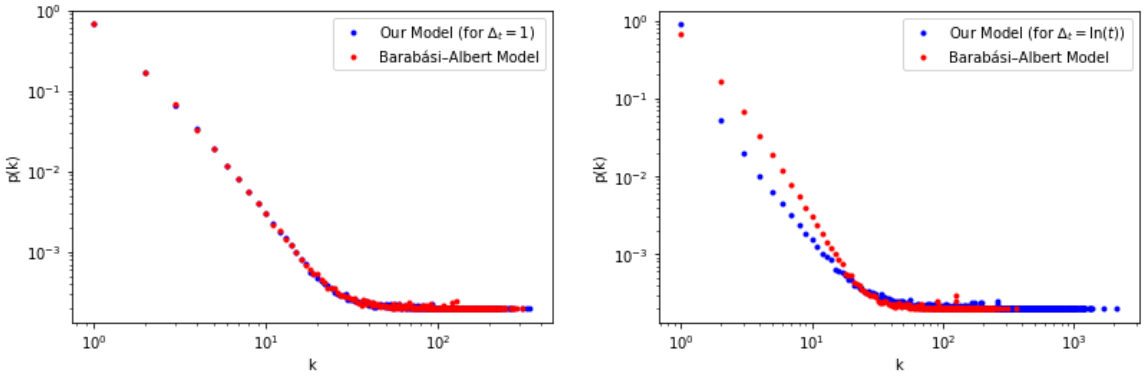


Figure 4.3: On the left-hand side is a 15-vertex network generated via the the draws from a Pólya urn with expanding colors and $\Delta_t = 5$ for all $t \geq 1$ and on the right-hand side is a network with 15 vertices generated via Barabási-Albert model. For our model, unlike the Barabási-Albert model, each vertex is represented by a distinct color which corresponds to a color type of balls in the Pólya urn at that time instant. Furthermore, the extra reinforcement parameter Δ_t in our model provides versatility in the level of preferential attachment. The parameter $\Delta_t = 5$ in our model enables the central vertex of the graph on the left-hand side to obtain a higher degree (11 in this case) as compared to the right-hand side Barabási-Albert network in which the highest degree achieved is 6.

distribution plots for the Barabási-Albert model and our model with $\Delta_t = 1$ are very similar and hence the networks generated via our model for $\Delta_t = 1$ are expected to show a degree distribution corresponding to $p(k) \sim k^{-3}$. The similarity between the Barabási-Albert algorithm and our model with $\Delta_t = 1$ can be represented in the following way:

$$
\begin{aligned}
&P(\text{incoming vertex at time } t \text{ connects to vertex corresponding to color } c_j) \\
&\quad = \text{ratio of } c_j \text{ color balls in the expanding color Pólya urn at time } t-1 \\
&\quad = \frac{\text{degree of vertex corresponding to color } c_j \text{ in graph } \mathcal{G}_{t-1}}{\text{sum of degrees in graph } \mathcal{G}_{t-1}} \\
&\quad = P(\text{incoming vertex connects to the vertex added at time } j-1 \\
&\qquad \text{in a standard Barabási-Albert network}).
\end{aligned}
$$



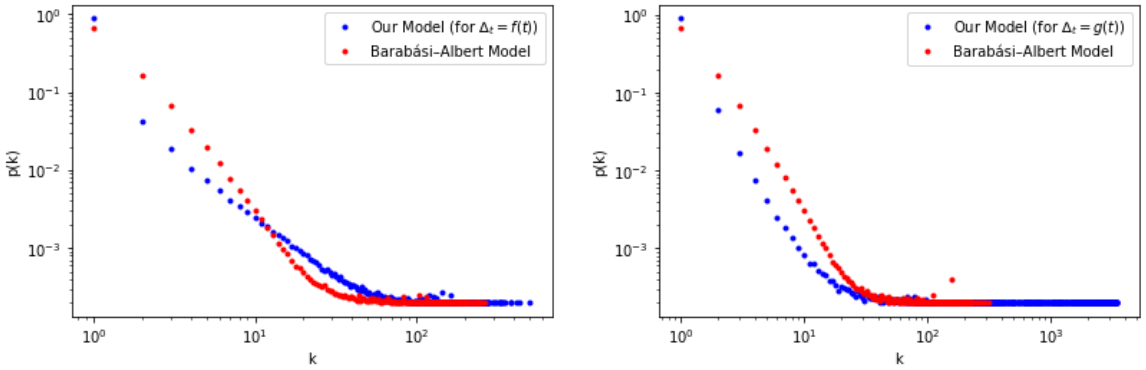
(a) $\Delta_t = 1$ (b) $\Delta_t = \ln(t)$

Figure 4.4: Degree distributions of networks generated until time 5000 (averaged over 250 simulations) for the Barabási-Albert model and our model with $\Delta_t = 1$ and $\Delta_t = \ln(t)$. In **(a)** the degree distributions of both models are nearly identical. While, in **(b)** the degree distributions are quite different.

Hence in the case when $\Delta_t = 1$, the mechanisms of both models for iteratively constructing new vertices and edges are equivalent. However the initialization of our model is different from the Barabási-Albert model. In our model, the initial graph has only one vertex with a self-loop, whereas in the Barabási-Albert model, the initial graph can potentially have more than one vertex equipped with an edge set and no self-loops. Even though the initialization of both models are different, the equivalent procedures for adding new vertices and edges between our model with $\Delta_t = 1$ and the standard Barabási-Albert model ensure that the generated graphs via both models will show similar properties for sufficiently large t. However, it is difficult to analytically solve for the degree distribution and other properties of our model as its reinforcement dynamics is much more involved than the Barabási-Albert model.

In Figure 4.4 (b), we observe that the degree distribution of our model with $\Delta_t = \ln(t)$ significantly differs from the degree distribution of Barabási-Albert model.

The former has a lower probability of obtaining lower degree vertices (degree range $10^0 - 10^1$) as compared to the latter but has a slightly higher probability of gaining moderate degree vertices (degree range $50 - 150$). Additionally, the maximum degree attained in the case of $\Delta_t = \ln(t)$ for our model in Figure 4.4 is much higher ($\sim 10^3$ as compared to only 200 in Barabási-Albert network).



(a) $\Delta_t = f(t)$ (b) $\Delta_t = g(t)$

Figure 4.5: Degree distribution of the Barabási-Albert model and our model generated for two different choices of Δ_t, **(a)** $\Delta_t = f(t)$ and **(b)** $\Delta_t = g(t)$, where $f(t)$ and $g(t)$ are defined in (4.25). Both plots are averaged over 250 simulations, where each simulation is a generation of a 5000-vertex graph.

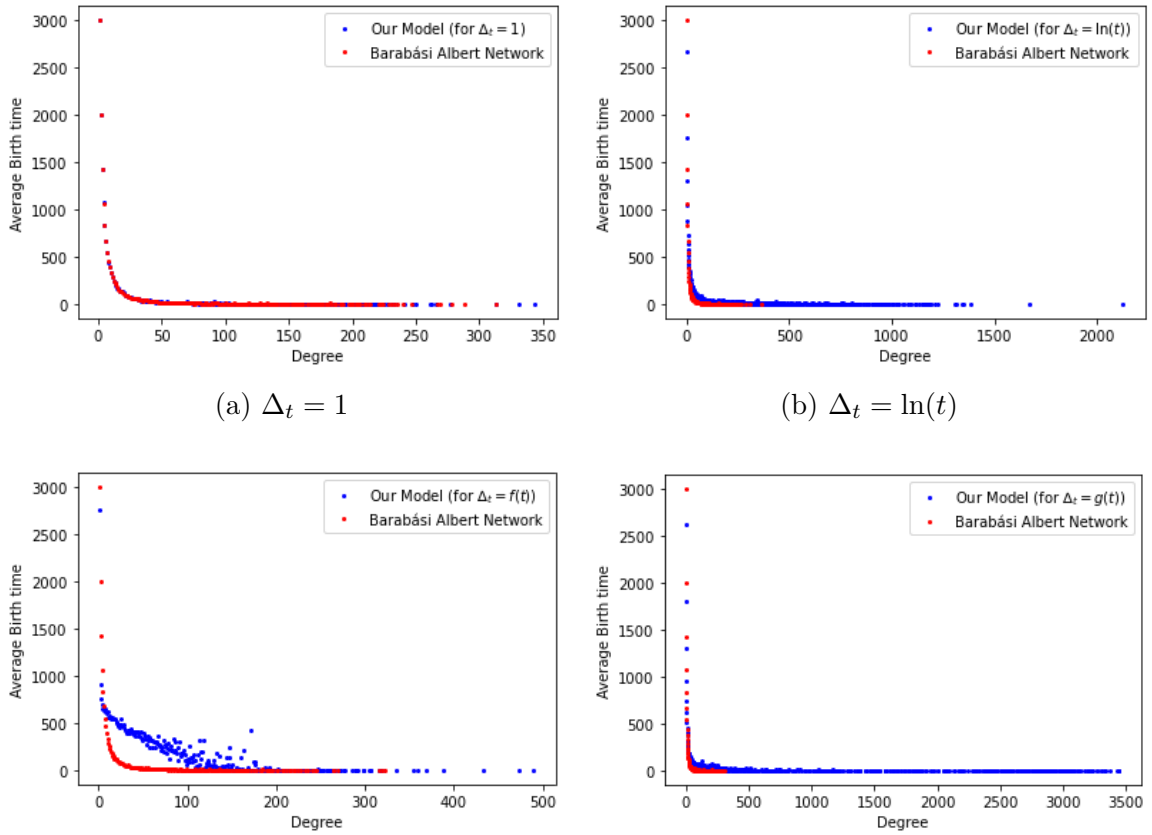
In Figure 4.5, we present two more cases in which the degree distributions differ substantially between our model and the Barabási-Albert model. More specifically, we generate networks via our model for Δ_t being an increasing step function $f(t)$ and

a decreasing continuous function $g(t)$ given by:

$$f(t) = \begin{cases} 1 & \text{for} \quad 0 \leq t < 1000 \\ 10 & \text{for} \quad 1000 \leq t < 2500 \\ 100 & \text{for} \quad 2500 \leq t \leq 5000, \end{cases} \qquad g(t) = \begin{cases} 10 & \text{for} \quad 0 \leq t \leq 1000 \\ \frac{10^4}{t} & \text{for} \quad 1000 \leq t \leq 2000 \\ 5 & \text{for} \quad 2000 \leq t \leq 3000 \\ \frac{15 \times 10^3}{t} & \text{for} \quad 3000 \leq t \leq 4000 \\ 3.75 & \text{for} \quad 4000 \leq t \leq 5000. \end{cases} \tag{4.25}$$

We remark from Figure 4.5 that the maximum degree attained in both networks (generated via $\Delta_t = f(t)$ and $\Delta_t = g(t)$) is higher than in the Barabási-Albert network. Furthermore, there are more moderate degree nodes (with degree $10^1 - 10^2$) in the case of $\Delta_t = f(t)$ and less moderate degree nodes in the case of $\Delta_t = g(t)$ when compared to the Barabási-Albert network.

In the next set of simulations, we compare (4.24) for our network generated with $\Delta_t = 1, \ln(t), f(t)$ and $g(t)$ and the Barabási-Albert network. Figure 4.6(a) demonstrates that in both the Barabási-Albert model and our model for $\Delta_t = 1$ vertices of the same degree are born at similar times. The stark similarities between the Barabási-Albert model and our model for $\Delta_t = 1$ in Figures 4.4 and 4.6 strongly suggest that both models generate networks with very similar structures; however a rigorous analytic study is required to confirm if our model with $\Delta_t = 1$ is stochastically equivalent to the standard Barabási-Albert model. In Figure 4.6(b), we observe that for our model with $\Delta_t = \ln(t)$, the network shows slightly more connectivity in the vertices which are born at similar times as compared to the Barabási-Albert



(a) $\Delta_t = 1$ (b) $\Delta_t = \ln(t)$

(c) $\Delta_t = f(t)$ (d) $\Delta_t = g(t)$

Figure 4.6: Vertices average birth time versus degree for our model using **(a)** $\Delta_t = 1$; **(b)** $\Delta_t = \ln(t)$; **(c)** $\Delta_t = f(t)$ and **(d)** $\Delta_t = g(t)$ (where the functions $f(t)$ and $g(t)$ are given in (4.25)) and for the Barabási-Albert network. All networks are generated for 5000 time steps and the average of 250 such networks is plotted.

network. This effect of same age vertices showing more connectivity when compared to Barabási-Albert networks is much more amplified when our model uses $\Delta(t) = f(t)$ as shown in Figure 4.6(c). Both cases provide a much richer algorithm for generating real-life networks in which the "rich gets richer" phenomenon needs to be dampened as it allows the more recently born vertices to get more connectivity. In contrast, Figure 4.6(d) shows an amplification of the "rich gets richer" phenomenon when compared to the Barabási-Albert model as the first two richest vertices achieve a significantly higher degree (around 3500) compared to all other vertices. The rest of the vertices have very similar connectivity as that of the Barabási-Albert network. The choice $\Delta_t = g(t)$ of the reinforcement parameter in our model provides an algorithm to generate graphs which are spatially similar to the Barabási-Albert network but demonstrate a higher effect of preferential attachment.

Chapter 5

Conclusion

5.1 Thesis Summary

In the first part of this thesis, we devised an interacting network of two-color finite memory Pólya urns. We formulated two types of interaction among the urns; using a fixed (time-invariant) stochastic matrix and through (time-variant) "super-urns" which consist of all the balls of the urn and its neighbours. In particular, in Chapter 2, we presented a network of two-color finite memory Pólya urns interacting via a fixed stochastic interaction matrix. We showed that the underlying Markov process is irreducible and aperiodic when the reinforcement parameters are time-invariant and hence has a unique stationary distribution. We also derived the exact asymptotic marginal infection distribution for the homogeneous case, i.e., when all the urns have identical initial conditions and reinforcement parameters. For the non-homogeneous interacting Pólya contagion network with time-invariant reinforcement parameters, we constructed dynamical systems to evaluate the network's infection propagation.

We showed that when memory $M = 1$, the probability of infection can be exactly represented by a linear dynamical system which has a unique equilibrium point to which the solution asymptotically converges. For memory $M > 1$, we used mean-field approximations to construct an approximating dynamical system which is nonlinear in general; we obtained a linearization of this dynamical system and characterize its equilibrium. We provided simulations comparing the corresponding linear and approximating nonlinear dynamical systems with the original stochastic process. Notably, we demonstrated that the approximating nonlinear dynamical system performs well for all tested values of memory and network size.

In Chapter 3, we demonstrated that a connected network of finite memory Pólya urns interacting through "super-urns" can be used to model opinion dynamics in a social network. Using the properties of the underlying reducible Markov process, we proposed a provably correct consensus dynamics using this model. For the case with homogeneous reinforcement parameters across individuals, we provided a delayed dynamical system that can be used to study the asymptotic properties of this model and determine explicitly the consensus value.

In the second part of this thesis (Chapter 4), we formulated an algorithm for generating preferential attachment graphs using a modified Pólya urn with expanding colors and a time-varying reinforcement parameter Δ_t. The network obtained is similar to the Barabási-Albert network for the case $\Delta_t = 1$ and gains a significant amount of versatility when Δ_t is a time-varying function. We analysed the draw vectors of the underlying stochastic process and derived the probability distribution of a random variable counting the draws of a particular color of this Pólya process. This random variable can be written in terms of the degree of the vertex in the constructed

preferential attachment network corresponding to this color. We provided simulation evidence for the structural similarities between our model and the Barabási-Albert network for $\Delta_t = 1$ and also justified the richness and versatility of our model for general Δ_t.

5.2 Future Work

In this section, we discuss possible directions for future research.

(1) For the IPCN(M, N) system presented in Chapter 2, the underlying Markov process of the M-length draw vectors is irreducible and aperiodic (see Lemma 2.2) when the reinforcement parameters are time-invariant. Even though, it may not be feasible to obtain a closed form formula for the unique stationary distribution for the general non-homogeneous case, one might be able to solve for the stationary distribution in some simpler cases such as the one presented in Example 2.3 for a homogeneous IPCN$(2, 2)$ system. Furthermore, using (2.11) and setting the stationary distribution of the Markov chain $\{\mathbf{W}_t\}_{t=1}^{\infty}$ for a homogeneous IPCN(M, N) as Π, where the entries of Π are denoted by $\pi_{k_1 k_2, \cdots, k_M}$ with $k_j = (k_{1j}, k_{2j}, \cdots, k_{Nj}) \in \{0, 1\}^N$ for $j \in \{1, 2, \cdots, M\}$, we obtain

$$\lim_{t\to\infty} P(Z_{i,t} = 1) = \sum_{\substack{k_{ij}=1 \\ j\in\{1,2,\cdots,M\}}} \pi_{k_1 k_2 \cdots k_M} = \rho. \tag{5.1}$$

One can analyse $\Pi Q^{(M,N)} = \Pi$, where the entries of the transition probability matrix $Q^{(M,N)}$ are given in (2.9), in conjunction with (5.1) to solve for the stationary distribution of $\{\mathbf{W}_t\}_{t=1}^{\infty}$ for some homogeneous cases.

(2) An interesting future direction for the IPCN(M, N) system which is developed in Chapter 2 is to consider time-varying interactions between urns. While we do analyse one form of time-varying interaction matrix devised via the concept of "super-urns" defined in [34] to formulate a consensus achieving interacting network of urns in Chapter 3, additional work is required to study the stochastic properties of a network of urns with general time-varying interactions. Furthermore, it is much more challenging to construct a class of dynamical systems for IPCN(M, N) system with time-varying interactions as carried in Section 2.3 of Chapter 2 using mean-field approximations.

(3) Another useful future direction which we have not explored in this thesis is to set different memories for different Pólya urns in the interacting networks.

(4) A realistic extension for our model in Chapter 4 is to consider removal of edges after some time to account for the fading away of popularity of an idea after some time in a social network. An effective way to add this feature to our model is to equip the Pólya urn (via which the graph is constructed) with a finite memory as in [1]. The removal of balls from the urns can be used to setup removal of edges from the corresponding graph. However, it is problematic to use a time-invariant memory for the urn because the removal of every edge after a fixed amount of time will result in the graph not growing. One way of addressing this concern would be to consider a time-varying memory for the urn, such that the removal of the edges due to finite memory is slower than the addition of edges and hence the resultant graph grows with time. For instance, we can set the memory of the urn as $\lceil \ln(t) \rceil$, i.e., a ball is removed from the urn

$\lceil \ln(t) \rceil$ steps after its addition to the urn.

(5) Finally, the preferential attachment graph of Chapter 4 relies on the draw variables of a Pólya urn with a number of colors that grows without bound as time goes to infinity. Another possible extension of this model is to use an L-color Pólya urn instead to formulate the graph generating algorithm, where L is a *fixed* number of colors. One way in which we can setup this algorithm is as follows: At time $t = 0$, the Pólya urn consists of balls of L colors. Correspondingly, we have an initial graph $\mathcal{G}_0$ with L vertices (each vertex corresponding to a unique color of the urn) and degree of each vertex equals the number of balls of its corresponding color in the urn. At each time instant $t \geq 1$, we draw a ball and return it to the urn along with $\Delta_t > 0$ balls of the same color. We then introduce a vertex to the graph $\mathcal{G}_{t-1}$ (corresponding to the color drawn) and connect it with the existing vertex of the color drawn which has the maximum degree among all the vertices of that color. This results in the newly formed graph $\mathcal{G}_t$.

In contrast to our original model in which each vertex of the graph is uniquely represented by a color in the urn, in this extension, the vertices are grouped into N categories (each represented by a color of the urn). Each category is given an initial influence factor in terms of the initial (at time $t = 0$) number of balls of its color in the urn. The use of a finite number L of colors (unlike the growing number of colors used in Chapter 4) for generating a preferential attachment graph may facilitate the exact analysis of the asymptotic behavior of the graph's degree distribution.

Bibliography

[1] F. Alajaji and T. Fuja. A communication channel modeled on contagion. *IEEE Transactions on Information Theory*, 40(6):2035–2041, 1994.

[2] R. Albert and A. L. Barabási. Topology of evolving networks: Local events and universality. *Physical Review Letters*, 85(24), 2000.

[3] R. Albert and A. L. Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47–97, 2002.

[4] G. Aletti and I. Crimaldi. The rescaled Pólya urn: Local reinforcement and chi-squared goodness-of-fit test. *Advances in Applied Probability*, 54:849–879, 2022.

[5] G. Aletti, I. Crimaldi, and A. Ghiglietti. Synchronization of reinforced stochastic processes with a network based interaction. *The Annals of Applied Probability*, 27(6):3787–3844, 2017.

[6] T. F. A. Alves, G. A. Alves, A. Macedo-Filho, R. S. Ferreira, and F. W. S. Lima. The diffusive epidemic process on Barabasi–Albert networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2021(4):043203, 2021.

[7] H. Andersson and T. Britton. *Stochastic Epidemic Models and their Statistical Analysis*, volume 151. Springer Science & Business Media, 2012.

[8] K. B. Athereya and S. Karlin. Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *The Annals of Mathematical Statistics*, 39(6):1801–1817, 1968.

[9] K. B. Athreya and P. E. Ney. *Branching Processes*. Dover Publications, New York, 2004.

[10] C. Avin, Z. Lotker, Y. Nahum, and D. Peleg. Random preferential attachment hypergraph. In *Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 398–405. ACM, 2019.

[11] R. N. Azizan, C. Thrampoulidis, and B. Hassibi. Improved bounds on the epidemic threshold of exact SIS models on complex networks. In *Proceedings of the IEEE 55th Conference on Decision and Control (CDC)*, pages 3560–3565. IEEE, 2016.

[12] A. Banerjee, P. Burlina, and F. Alajaji. Image segmentation and labeling using the Polya urn model. *IEEE Transactions on Image Processing*, 8(9):1243–1253, 1999.

[13] A. L. Barabási. Scale-free networks: A decade and beyond. *Science*, 325(5939):412–413, 2009.

[14] A. L. Barabási and R. Albert. Emergence of scaling in random networks science. *Science*, 286(5439):509–512, 1999.

[15] N. R. Barraza, G. Pena, and V. Moreno. A non-homogeneous Markov early epidemic growth dynamics model. Application to the SARS-CoV-2 pandemic. *Chaos, Solitons & Fractals*, page 139:110297, 2020.

[16] M. Benaïm, I. Benjamini, J. Chen, and Y. Lima. A generalized Pólya's urn with graph based interactions. *Random Structures & Algorithms*, 46(4):614–634, 2015.

[17] N. Berger, C. Borgs, J. Chayes, R.M. D'Souza, and R.D. Kleinberg. Degree distribution of competition-induced preferential attachment graphs. *Combinatorics, Probability & Computing*, 14(6):697–721, 2005.

[18] N. Berger, C. Borgs, J. Chayes, and A. Saberi. On the spread of viruses on the internet. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms*, volume 5, pages 301–310, 2005.

[19] N. Berger, C. Borgs, J. Chayes, and A. Saberi. Asymptotic behavior and distributional limits of preferential attachment graphs. *Annals of Probability*, 42(1):1–40, 2014.

[20] B. Bollobás and O. Riordan. The diameter of a scale-free random graph. *Combinatorica*, 24(1):5–34, 2004.

[21] V. S. Borkar. *Stochastic Approximation: A Dynamical Systems Viewpoint*. Cambridge University Press, Cambridge, UK, 2008.

[22] A. Capocci, V.D.P. Servedio, F. Colaiori, L.S. Buriol, D. Donato, S. Leonardi, and G. Caldarelli. Preferential attachment in the growth of social networks: The internet encyclopedia Wikipedia. *Phys. Rev. E*, 74(3):036116, 2006.

[23] F. Chung, S. Handjani, and D. Jungreis. Generalization of Pólya's urn problem. *Annals of Combinatorics*, 7:141–153, 2003.

[24] A. Collevecchio, C. Cotar, and M. LiCalzi. On a preferential attachment and generalized Pólya's urn model. *Annals of Applied Probability*, 23(3):1219–1253, 2013.

[25] G. Csányi and B. Szendrői. Structure of a large social network. *Physical Review E*, 69:5, 2004.

[26] A. Fazeli and A. Jadbabaie. On consensus in a correlated model of network formation based on a Pólya urn process. In *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference*, pages 2341–2346, 2011.

[27] A. D. Flaxman, A. M. Frieze, and J. Vera. A geometric preferential attachment model of networks. *Internet Mathematics*, 3(2):187–205, 2006.

[28] D. A. Freedman. Bernard Friedman's urn. *The Annals of Mathematical Statistics*, 36(3):956–970, 1965.

[29] U. Gangopadhyay and K. Maulik. Stochastic approximation with random step sizes and urn models with random replacement matrices. *Annals of Applied Probability*, 29(4):2033–2066, 2019.

[30] F. Giroire, N. Nisse, T. Trolliet, and M. Sulkowska. Preferential attachment hypergraph with high modularity. *Network Science*, 10(4):400–429, 2022.

[31] R. Gouet. Martingale functional central limit theorems for a generalized Pólya urn. *The Annals of Probability*, 21(3):1624–1639, 1993.

[32] G. Harrington, F. Alajaji, and B. Gharesifard. Initialization and curing policies for Pólya contagion networks. *SIAM Journal on Control and Optimization*, 60(2):S170–S195, 2022.

[33] A. K. Hartmann and M. Weigt. Statistical mechanics perspective on the phase transition in vertex covering of finite-connectivity random graphs. *Theoretical Computer Science*, 265(1):199–225, 2001.

[34] M. Hayhoe, F. Alajaji, and B. Gharesifard. A Pólya contagion model for networks. *IEEE Transactions on Control of Network Systems*, 5:1998–2010, 2018.

[35] M. Hayhoe, F. Alajaji, and B. Gharesifard. Curing epidemics on networks using a Pólya contagion model. *IEEE/ACM Transactions on Networking*, 5:2085–2097, 2019.

[36] M. Inoue, T. Pham, and H. Shimodaira. A hypergraph approach for estimating growth mechanisms of complex networks. *IEEE Access*, 10:35012–35025, 2022.

[37] S. Janson. Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Processes and their Applications*, 110(2):177 – 245, 2004.

[38] G. Kaur and N. Sahasrabudhe. Interacting urns on a finite directed graph. *Journal of Applied Probability*, 60(1):166–188, 2023.

[39] W. König. Orthogonal polynomial ensembles in probability theory. *Probability Surveys*, 2:385 – 447, 2005.

[40] P. L. Krapivsky, S. Redner, and F. Leyvraz. Connectivity of growing random networks. *Phys. Rev. Lett.*, pages 4629–4632, 2000.

[41] S. Laurelle and G. Pagés. Randomized urn models revisited using stochastic approximation. *The Annals of Applied Probability*, 23(4):1409–1436, 2013.

[42] M. Lin, G. Wang, and T. Chen. A modified earthquake model based on generalized Barabási–Albert scale-free networks. *Communications in Theoretical Physics*, 46(6):1011, 2006.

[43] W. Mei, S. Mohagheghi, S. Zampieri, and F. Bullo. On the dynamics of deterministic epidemic propagation over networks. *Annual Reviews in Control*, 44:116–128, 2017.

[44] C. Nowzari, V. M. Preciado, and G. J. Pappas. Analysis and control of epidemics: A survey of spreading processes on complex networks. *IEEE Control Systems Magazine*, 36(1):26–46, 2016.

[45] R. I. Oliveira. The onset of dominance in balls-in-bins processes with feedback. *Random Structures Algorithms*, 34:454–477, 2009.

[46] P. E. Paré, C. L. Beck, and A. Nedić. Epidemic processes over time-varying networks. *IEEE Transactions on Control of Network Systems*, 5(3):1322–1334, 2017.

[47] R. Pemantle. A survey of random processes with reinforcement. *Probability Surveys*, 4:1–79, 2007.

[48] D. Pfeifer. Pólya-Lundberg process. *Encyclopedia of Statistical Sciences*, 7, 1986.

[49] P. D. Pra, P. Y. Louis, and I. G. Minelli. Synchronazation via interacting reinforcement. *Journal of Applied Probability*, 51:556–568, 2016.

[50] M. Z. Rácz and A. Sridhar. Correlated randomly growing graphs. *The Annals of Applied Probability*, 32(2):1058 – 1111, 2022.

[51] S. M. Ross. *Introduction to probability models*. Academic Press, 10th edition, 2010.

[52] N. Sahasrabudhe. Synchronization and fluctuation theorems for interacting Friedman urns. *Journal of Applied Probability*, 53:1221–1239, 2016.

[53] F. C. Santos and J. M. Pacheco. Scale-free networks provide a unifying framework for the emergence of cooperation. *Physical Review Letters*, 95:098104, 2005.

[54] B. She, J. Liu, S. Sundaram, and P. E. Paré. On a networked SIS epidemic model with cooperative and antagonistic opinion dynamics. *IEEE Transactions on Control of Network Systems*, 9(3):1154–1165, 2022.

[55] S. Singh, F. Alajaji, and B. Gharesifard. Consensus using a network of finite memory Pólya urns. *IEEE Control Systems Letters*, 6:2780–2785, 2022.

[56] S. Singh, F. Alajaji, and B. Gharesifard. A finite memory interacting Pólya contagion network and its approximating dynamical systems. *SIAM Journal on Control and Optimization*, 60(2):347–369, 2022.

[57] S. Singh, F. Alajaji, and B. Gharesifard. Modeling network contagion via interacting finite memory Pólya urns. In *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, pages 348–353, 2022.

[58] B. Skyrms and R. Pemantle. A dynamic model of social network formation. *Proceedings of the National Academy of Sciences - PNAS*, 97(16):9340–9346, 2000.

[59] C. Tsallis and R. Oliveira. Complex network growth model: Possible isomorphism between nonextensive statistical mechanics and random geometry. *Chaos*, 32(5):053126–053126, 2022.

[60] Y. Wang, D. Chakrabarti, Wang C., and C. Faloutsos. Epidemic spreading in real networks: An eigenvalue viewpoint. In *Proceedings of the 22nd International Symposium on Reliable Distributed Systems*, pages 25–34, 2003.

[61] X. Zhongdong, Z. Guanghui, and B. Wang. Using modified Barabási and Albert model to study the complex logistic network in eco-industrial systems. *International Journal of Production Economics*, 140(1):295–304, 2012.