

# A Polya Contagion Model for Networks

Mikhail Hayhoe<sup>1</sup>, Fady Alajaji<sup>2</sup>, *Senior Member, IEEE*, and Bahman Gharesifard<sup>3</sup>, *Member, IEEE*

**Abstract**—A network epidemics model based on the classical Polya urn scheme is investigated. Temporal contagion processes are generated on the network nodes using a modified Polya sampling scheme that accounts for spatial infection among neighboring nodes. The stochastic properties and the asymptotic behavior of the resulting network contagion process are analyzed. Unlike the classical Polya process, the network process is noted to be nonstationary in general, although it is shown to be time invariant in its first and some of its second-order statistics and to satisfy martingale convergence properties under certain conditions. Three classical Polya processes, one computational and two analytical, are proposed to statistically approximate the contagion process of each node, showing a good fit for a range of system parameters. Finally, empirical results compare and contrast our model with the well-known discrete time susceptible-infected-susceptible model.

**Index Terms**—Epidemics on networks, martingales, nonstationary stochastic processes, Polya contagion networks.

## I. INTRODUCTION

IN THIS paper, we examine the dynamics and properties of a contagion process, or *epidemic*, on a network. Here, an epidemic can represent a disease [2], a computer virus [3], the spread of an innovation, rumor or idea [4], or the dynamics of competing opinions in a social network [5].

Many different models for the study of infection propagation and curing exist in the literature. Our model, the network Polya contagion process, bears similarities to the well-known susceptible-infected-susceptible (SIS) infection model [6]. In this model, all nodes may initially be healthy or infected. As the epidemic spreads, nodes that are infected can be cured to become healthy, but any healthy node may become infected at any time, regardless of whether they have been cured previously. Epidemics on networks have been intensively studied in recent years; see [7] and references therein and thereafter. The model that we present is an adaptation of the classical Polya contagion process [8]–[10] to a network setting by accounting for spatial

infection between nodes. The classical Polya model has been used to study a variety of epidemics such as the bubonic plague in Peru [11] and the spread of chlamydia in a closed population [12], as well as a wide range of other applications; see [13] for a survey. In this paper, we will examine the stochastic evolution of the network Polya contagion process.

Our model is motivated by the classical Polya contagion process, which evolves by sampling from an urn containing a finite number of red and black balls [8]–[10]. In the network Polya contagion model, each node of the underlying network is equipped with an individual urn; however, instead of sampling from these urns when generating its contagion process, each node has a “super-urn,” created by combining the contents of its own urn with those of its neighbors’ urns. This adaptation captures the concept of spatial infection, since having infected neighbors increases the chance that an individual is infected in the future. This concept of the super-urn sampling mechanism for incorporating spatial interactions was originally introduced in [14] in the context of the image segmentation and labeling problem. We herein adapt the image model of [14] for a network setting and analyze the resulting contagion process affecting each node of the network.

More specifically, we study the time evolution and stochastic properties of the proposed network contagion process. We derive an expression for the temporal  $n$ -fold joint probability distribution of the process. We show that this process, unlike the classical Polya urn process, is, in general, nonstationary and, hence, not exchangeable. For the special case of complete networks, we analytically find the 1-D and 2-D  $(n, 1)$ -step marginal distributions of the contagion process. These results show that even though it is not stationary, the process in this case is nevertheless identically distributed with its latter two marginal distributions being invariant to time shifts. We also establish several martingale properties regarding the network urn compositions, proving that the proportions of red balls in each node’s urn as well as the network average urn proportion converge almost surely to a limit as time grows without bound. We next provide three approximations to the network contagion process by modeling each node’s contagion process via a classical stationary Polya process [10]. In the first one, we approximate each node’s process with the classical Polya process whose correlation parameter is empirically selected so that the Kullback–Leibler divergence measure between its  $n$ -fold joint distribution and that of the original node process is minimized. In the second approximation, we propose an analytical model whose parameters are chosen by matching its first and  $(n, 1)$ -step second-order statistics with those of the original node process, which fits well for large networks. The last approximation uses a classical Polya model with parameters

Manuscript received August 23, 2017; revised August 27, 2017 and August 31, 2017; accepted November 9, 2017. Date of publication December 8, 2017; date of current version December 14, 2018. This work was supported by the Natural Sciences and Engineering Research Council of Canada. This paper was presented in part at the American Control Conference, May 24–26, 2017. Recommended by Associate Editor Ling Shi. (*Corresponding author: Mikhail Hayhoe.*)

M. Hayhoe is with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104 USA (e-mail: mhayhoe@seas.upenn.edu).

F. Alajaji and B. Gharesifard are with the Department of Mathematics and Statistics, Queen’s University, Kingston, ON K7L 3N6, Canada (e-mail: fa@queensu.ca; ghareb@queensu.ca).

Digital Object Identifier 10.1109/TCNS.2017.2781467

chosen analytically that we show fits well for small networks. Finally, simulation results are presented to support the validity of these approximations and to compare our model with the traditional discrete time SIS model, which suggests that the network Polya contagion process captures certain properties of the SIS model, while offering new insights in the case of widespread infection.

The rest of this paper is organized as follows. Section II outlines some preliminary knowledge that will be used throughout this paper. Section III introduces the network contagion process, and Section IV presents its stochastic properties and asymptotic behavior. Section V proposes three approximations for the individual node contagion processes in the network, along with numerical modeling results. Finally, Section VI concludes this paper.

## II. PRELIMINARIES

For a sequence  $v_i = (v_{i,1}, \dots, v_{i,n})$ , we use the notation  $v_{i,s}^t$  with  $1 \leq s < t \leq n$  to denote the vector  $(v_{i,s}, v_{i,s+1}, \dots, v_{i,t})$ . Our technical results rely on notions from stochastic processes, some of which we recall here. Throughout, we assume that the reader is familiar with basic notions of probability theory.

Let  $(\Omega, \mathcal{F}, P)$  be a probability space, and consider the stochastic process  $\{Z_n\}_{n=1}^\infty$ , where each  $Z_n$  is a random variable on  $\Omega$ . We often refer to the indexes of the process as “time” indexes. We recall that the process  $\{Z_n\}_{n=1}^\infty$  is *stationary* if for any  $n \in \mathbb{Z}_{\geq 1}$ , its  $n$ -fold joint probability distribution [i.e., the distribution of  $(Z_1, \dots, Z_n)$ ] is invariant to time shifts. Further,  $\{Z_n\}_{n=1}^\infty$  is *exchangeable* if for any  $n \in \mathbb{Z}_{\geq 1}$ , its  $n$ -fold joint distribution is invariant to permutations of the indexes  $1, \dots, n$ . It directly follows from the definitions that an exchangeable process is stationary. Finally, the process  $\{Z_n\}_{n=1}^\infty$  is called a *martingale* (respectively, *supermartingale*, *submartingale*) with respect to the process  $\{Y_n\}_{n=1}^\infty$  if  $E[|Z_n|] < \infty$  and  $E[Z_{n+1}|Y_n] = Z_n$  almost surely (respectively, less than or equal to, greater than or equal to), for all  $n$ . Precise definitions of all notions, including those of *ergodicity*, can be found in standard texts (e.g., [15], [16]).

We now recall the classical version of the Polya contagion process [8], [10]. Consider an urn with  $R \in \mathbb{Z}_{>0}$  red balls and  $B \in \mathbb{Z}_{>0}$  black balls. We denote the total number of balls by  $T$ , that is,  $T = R + B$ . At each time step, a ball is drawn from the urn. The ball is then returned along with  $\Delta > 0$  balls of the same color. We use an indicator  $Z_n$  to denote the color of ball in the  $n$ th draw

$$Z_n = \begin{cases} 1, & \text{if the } n\text{th draw is red} \\ 0, & \text{if the } n\text{th draw is black.} \end{cases}$$

Let  $U_n$  denote the proportion of red balls in the urn after the  $n$ th draw. Then

$$U_n := \frac{R + \Delta \sum_{t=1}^n Z_t}{T + n\Delta} = \frac{\rho_c + \delta_c \sum_{t=1}^n Z_t}{1 + n\delta_c}$$

where  $\rho_c = \frac{R}{T}$  is the initial proportion of red balls in the urn and  $\delta_c = \frac{\Delta}{T}$  is a correlation parameter. Since we draw balls from this urn at each time step, the conditional probability of drawing

a red ball at time  $n$ , given  $Z^{n-1} = (Z_1, \dots, Z_{n-1})$ , is given by

$$P(Z_n = 1 | Z^{n-1}) = \frac{R + \Delta \sum_{t=1}^{n-1} Z_t}{T + (n-1)\Delta} = U_{n-1}.$$

It can be easily shown that  $\{U_n\}_{n=1}^\infty$  is a martingale [17]. The process  $\{Z_n\}_{n=1}^\infty$ , whose  $n$ -fold joint distribution is denoted by  $Q_{\rho_c, \delta_c}^{(n)}$ , is also exchangeable (hence stationary) and nonergodic with both  $U_n$  and the process sample average  $\frac{1}{n} \sum_{i=1}^n Z_i$  converging almost surely as  $n \rightarrow \infty$  to a random variable governed by the Beta distribution with parameters  $\frac{\rho_c}{\delta_c}$  and  $\frac{1-\rho_c}{\delta_c}$ ; we denote this probability density function (pdf) by  $\text{Beta}(\frac{\rho_c}{\delta_c}, \frac{1-\rho_c}{\delta_c})$  [17], [18]. Finally, the 1-D distribution of the Polya process is  $Q_{\rho_c, \delta_c}^{(1)}(a) = P(Z_n = a) = (\rho_c)^a (1 - \rho_c)^{1-a}$ , for all  $n \in \mathbb{Z}_{\geq 1}$  and  $a \in \{0, 1\}$ . The above classical Polya process  $\{Z_n\}_{n=1}^\infty$  is fully described by its parameters  $\rho_c$  and  $\delta_c$  and, thus, we denote it by  $\text{Polya}(\rho_c, \delta_c)$ .

## III. NETWORK POLYA CONTAGION PROCESS

In this section, we introduce a generalization of the Polya contagion process to networks, where each individual node in the underlying graph that describes the network topology is still equipped with an urn; however, the node’s neighboring structure affects the evolution of its process. This model, hence, captures spatial contagion, since infected neighbors increase the chance of a node being infected in the future.

Consider an undirected graph  $\mathcal{G} = (V, \mathcal{E})$ , where  $V = \{1, \dots, N\}$  is the set of  $N \in \mathbb{Z}_{\geq 1}$  nodes and  $\mathcal{E} \subset V \times V$  is the set of edges. We assume that  $\mathcal{G}$  is connected, that is, there is a path between any two nodes in  $\mathcal{G}$ . We use  $\mathcal{N}_i$  to denote the set of nodes that are neighbors to node  $i$ , that is,  $\mathcal{N}_i = \{v \in V : (i, v) \in \mathcal{E}\}$ , and  $\mathcal{N}_i' = \{i\} \cup \mathcal{N}_i$ . If  $\mathcal{N}_i' = V$  for all  $i \in V$ , the network is called complete; if  $|\mathcal{N}_i| = |\mathcal{N}_j|$  for all  $i, j \in V$ , we call it regular. Each node  $i \in V$  is equipped with an urn, initially with  $R_i \in \mathbb{Z}_{>0}$  red balls and  $B_i \in \mathbb{Z}_{>0}$  black balls (we do not let  $R_i = 0$  or  $B_i = 0$  to avoid any degenerate cases). We let  $T_i = R_i + B_i$  be the total number of balls in the  $i$ th urn,  $i \in \{1, \dots, N\}$ . We use  $Z_{i,n}$  as an indicator for the ball drawn for node  $i$  at time  $n$

$$Z_{i,n} = \begin{cases} 1, & \text{if the } n\text{th draw for node } i \text{ is red} \\ 0, & \text{if the } n\text{th draw for node } i \text{ is black.} \end{cases}$$

However, instead of drawing solely from its own urn, each node draws simultaneously from a “super-urn” created by combining all of the balls in its own urn with the balls in its neighbors’ urns; see Fig. 1. This allows the spatial relationships between nodes to influence their state. This means that  $Z_{i,n}$  is the indicator for a ball drawn from node  $i$ ’s super-urn, and not its individual urn. Hence, the super-urn of node  $i$  initially has  $\bar{R}_i = \sum_{j \in \mathcal{N}_i'} R_j$  red balls,  $\bar{B}_i = \sum_{j \in \mathcal{N}_i'} B_j$  black balls, and  $\bar{T}_i = \sum_{j \in \mathcal{N}_i'} T_j$  balls in total.

We further consider a time-varying version of the classical Polya contagion process, following [19], where at time  $t$  for node  $i \in V$ , after a red ball is drawn, it is returned along with  $\Delta_{r,i}(t)$  red balls to node  $i$ ’s urn, and  $\Delta_{b,i}(t)$  black balls along with the drawn ball are added to node  $i$ ’s urn when a black

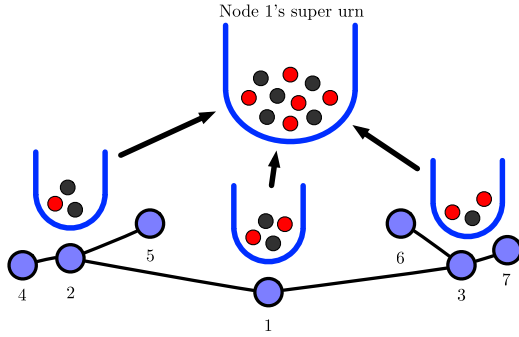


Fig. 1. Illustration of a super-urn in a network.

ball is drawn. When  $\Delta_{r,i}(t) = \Delta_{b,i}(t)$  for all  $t \in \mathbb{Z}_{\geq 1}$ , we write  $\Delta_i(t)$  instead; if the  $\Delta$ 's are not node dependent, we omit the node index. We assume throughout that  $\Delta_{r,i}(t) \geq 0$ ,  $\Delta_{b,i}(t) \geq 0$ , for all  $t \in \mathbb{Z}_{\geq 1}$  and that there exists  $i \in V$  and  $t$  such that  $\Delta_{r,i}(t) + \Delta_{b,i}(t) \neq 0$ ; otherwise, we are simply sampling with replacement.

In the context of epidemics, the red and black balls in an urn, respectively, represent units of ‘‘infection’’ and ‘‘healthiness’’; for example, bacteria and white blood cells. In a super-urn, the bacteria can infect others in the area and the white blood cells contribute to the overall health in the neighborhood of an individual. Drawing red at time  $t$  means the bacteria in the neighborhood were successful in reproduction and so the individual was more infected; otherwise, they were healthier since the white blood cells reproduced. Thus, when  $Z_{i,n} = 1$ , we declare that node  $i$  is *infected* at time  $n$ , and if  $Z_{i,n} = 0$ , then it is *healthy*. We add more units of bacteria once they reproduce, but commonly assume this number  $\Delta_{r,i}(t)$  is the same across all individuals and time because the bacteria do not evolve or become altered. The amount of white blood cells created  $\Delta_{b,i}(t)$  may change since we can give more medicine to certain people to increase their immune response, or vaccinate them so they are better able to fight the disease.

To express the proportion of red balls in the individual urns of the nodes, we define the random vector  $U_n = (U_{1,n}, \dots, U_{N,n})$ , where  $U_{i,n}$  is the proportion of red balls in node  $i$ 's urn after the  $n$ th draw,  $i \in V$ . For node  $i$

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

where the numerator represents the total number of red balls in node  $i$ 's urn after the  $n$ th draw, whereas the denominator is the total number of balls in the same urn. Note that  $U_{i,0} = \frac{R_i}{T_i}$  is the initial proportion of balls in node  $i$ 's urn. For ease of notation, let

$$X_{j,n} = T_j + \sum_{t=1}^n Z_{j,t} \Delta_{r,j}(t) + (1 - Z_{j,t}) \Delta_{b,j}(t). \quad (1)$$

Furthermore, we define the random vector  $S_n = (S_{1,n}, \dots, S_{N,n})$  as the proportion of red balls in the super-urns of the nodes after the  $n$ th draw, so that  $S_{i,n}$  is the proportion of red

balls in node  $i$ 's super-urn after  $n$  draws. Hence, for node  $i$

$$\begin{aligned} S_{i,n} &:= \frac{\bar{R}_i + \sum_{j \in \mathcal{N}'_i} \sum_{t=1}^n Z_{j,t} \Delta_{r,j}(t)}{\sum_{j \in \mathcal{N}'_i} X_{j,n}} \\ &= \frac{\sum_{j \in \mathcal{N}'_i} U_{j,n} X_{j,n}}{\sum_{j \in \mathcal{N}'_i} X_{j,n}}. \end{aligned} \quad (2)$$

Note that  $S_{i,0} = \frac{\bar{R}_i}{T_i}$ .  $S_{i,n}$  is, in fact, a function of the random draw variables of the network, and, in particular, of  $\{Z_j^n\}_{j \in \mathcal{N}'_i}$ , but for ease of notation, when the arguments are clear, we write  $S_{i,n}(Z_1^n, \dots, Z_N^n) = S_{i,n}$ . Then, the conditional probability of drawing a red ball from the super-urn of node  $i$  at time  $n$  given the complete network history, that is, given all past  $n - 1$  draw variables for each node in the network  $\{Z_j^{n-1}\}_{j=1}^N = \{(Z_{1,1}, \dots, Z_{1,n-1}), \dots, (Z_{N,1}, \dots, Z_{N,n-1})\}$ , satisfies

$$\begin{aligned} P(Z_{i,n} = 1 | \{Z_j^{n-1}\}_{j=1}^N) \\ &= \frac{\bar{R}_i + \sum_{j \in \mathcal{N}'_i} \sum_{t=1}^{n-1} Z_{j,t} \Delta_{r,j}(t)}{\sum_{j \in \mathcal{N}'_i} X_{j,n-1}} \\ &= S_{i,n-1}. \end{aligned} \quad (3)$$

That is, the conditional probability of drawing a red ball for node  $i$  at time  $n$  given the entire past  $\{Z_j^{n-1}\}_{j=1}^N$  is the proportion of red balls in its super-urn  $S_{i,n-1}$ . This is, however, analogous to the original Polya case, but instead of relying on the individual proportion of red balls  $U_n$  to describe the conditional probability of drawing red balls, we use the super-urn proportion of red balls since we now draw from there.

**Remark 3.1 (Non-Markovity):** While (3) may appear to suggest some sort of Markovity property, the process is non-Markovian in general. This can easily be seen due to the fact that a draw at time  $n$  requires knowledge of all previous draws for the entire neighborhood.

A main objective throughout the rest of this paper is to study the evolution and stochastic properties of the process defined above. Using the above conditional probability, we can determine the  $n$ -fold joint probability of the entire network  $\mathcal{G}$ : for  $a_i^n \in \{0, 1\}^n$ ,  $i \in \{1, \dots, N\}$ , we have that

$$\begin{aligned} P_{\mathcal{G}}^{(n)}(a_1^n, \dots, a_N^n) \\ &:= P(\{Z_i^n = a_i^n\}_{i=1}^N) \\ &= \prod_{t=1}^n P(\{Z_{i,t} = a_{i,t}\}_{i=1}^N | \{Z_i^{t-1} = a_i^{t-1}\}_{i=1}^N) \\ &= \prod_{t=1}^n \prod_{i=1}^N (S_{i,t-1})^{a_{i,t}} (1 - S_{i,t-1})^{1-a_{i,t}} \end{aligned} \quad (4)$$

where  $S_{i,t} = S_{i,t}(a_1^t, \dots, a_N^t)$  is defined in (3). With the above explicit joint distribution, it is possible to determine the distributions of each node's process. More specifically, using (4), the  $n$ -fold distribution of node  $i$ 's process at time  $t \geq n$  is

$$P_{i,t}^{(n)}(a_{i,t-n+1}, \dots, a_{i,t}) := \sum_{\substack{a_j^{t-n} \in \{0,1\}^{t-n} \\ a_j^t \in \{0,1\}^t : j \neq i}} P_{\mathcal{G}}^{(t)}(a_1^t, \dots, a_N^t).$$

In order to measure the spread of contagion in the network at any given time, we wish to see how likely it is, on average, for a node to be infected at that instant. We, hence, define the *average infection rate* in the network at time  $n$  as the average marginal probability of drawing a red ball

$$\tilde{I}_n := \frac{1}{N} \sum_{i=1}^N P(Z_{i,n} = 1) = \frac{1}{N} \sum_{i=1}^N P_{i,n}^{(1)}(1).$$

Note that  $\tilde{I}_n$  is a function of the network topology  $(V, \mathcal{E})$ , the initial placement of balls  $R_i$  and  $B_i$ , the draw processes  $\{Z_{i,t}\}_{t=1}^n$ , and number of balls added  $\{\Delta_{r,i}(t)\}_{t=1}^n$  and  $\{\Delta_{b,i}(t)\}_{t=1}^n$  for each node  $i \in V$ . Unfortunately for an arbitrary network, the above quantity does not yield an exact analytical formula (except in the simple case of complete networks). As such, in general, it is hard to mathematically analyze the asymptotic behavior of  $\tilde{I}_n$ , which we wish to minimize when attempting to cure an epidemic. Instead, we examine the asymptotic stochastic behavior of a closely related variable given by the average individual proportion of red balls at time  $n$ , namely

$$\tilde{U}_n := \frac{1}{N} \sum_{i=1}^N U_{i,n}$$

which we call the *network susceptibility*. This quantity is related to the conditional probability of drawing a red ball, as seen in (2). Since the individual urn of node  $i$  is in every super-urn in the neighborhood, if  $U_{i,n}$  increases, then  $S_{j,n}$  increases for every  $j \in \mathcal{N}'_i$  and, hence, given the past history, those nodes are more likely to exhibit infected behavior as seen from (3). Note that similar to  $\tilde{I}_n$ ,  $\tilde{U}_n$  is a function of the network variables.

**Remark 3.2 (Finite memory):** It is worth pointing out that a practical adaptation to our model can be considered, where urns have “finite memory” in the sense that the balls added after each draw are only kept in each node’s urn for a finite number of future draws. This model is developed in [18] for the classical Polya process in the context of modeling communication channels, where it is shown that the resulting finite memory contagion process is stationary, Markovian and ergodic. We present the following result that states that in this case the entire state is Markovian and, hence, it is a limited reinforcement model, but we leave an in-depth investigation to a future work. •

**Proposition 3.3 (Finite Memory Markovity):** The entire state of the network Polya contagion process  $\{Z_n\}_{n=1}^\infty$  with finite memory  $M$  is Markovian with memory  $M$ .

**Proof:** By (1) and (3) and the fact that added balls are removed after  $M$  steps, we have for  $n > M$  that

$$\begin{aligned} & P(Z_{i,n} = 1 \mid \{Z_j^{n-1}\}_{j=1}^N) \\ &= \frac{\bar{R}_i + \sum_{j \in \mathcal{N}'_i} \sum_{t=n-M}^{n-1} Z_{j,t} \Delta_{r,j}(t)}{\bar{T}_i + \sum_{j \in \mathcal{N}'_i} \sum_{t=n-M}^{n-1} Z_{j,t} \Delta_{r,j}(t) + (1 - Z_{j,t}) \Delta_{b,j}(t)} \\ &= P\left(Z_{i,n} = 1 \mid \{Z_{j,n-M}^{n-1}\}_{j=1}^N\right). \end{aligned}$$

Using the above result along with conditional independence, for  $(a_1, \dots, a_N) \in \{0, 1\}^N$ , we have for  $n > M$  that

$$\begin{aligned} & P(Z_{1,n} = a_1, \dots, Z_{N,n} = a_N \mid \{Z_j^{n-1}\}_{j=1}^N) \\ &= \prod_{i=1}^N P(Z_{i,n} = 1 \mid \{Z_j^{n-1}\}_{j=1}^N) \\ &= \prod_{i=1}^N P\left(Z_{i,n} = 1 \mid \{Z_{j,n-M}^{n-1}\}_{j=1}^N\right) \\ &= P\left(Z_{1,n} = a_1, \dots, Z_{N,n} = a_N \mid \{Z_{j,n-M}^{n-1}\}_{j=1}^N\right) \end{aligned}$$

and, hence, the entire network process  $\{Z_n\}_{n=1}^\infty$  is Markovian with memory  $M$ . ■

#### IV. STOCHASTIC PROPERTIES

We next examine the stochastic properties of the network contagion process. We assume throughout the beginning of this section that  $\Delta_{r,i}(t) = \Delta_{b,i}(t) = \Delta > 0$ , for all  $i \in V$  and times  $t$ ; that is, the net number of red and black balls added are equal and constant in time for all nodes. In the case of a complete network, the composition of every nodes’ super-urn is identical, since there is only one super-urn that is being drawn from. Thus, for a complete network, the super-urn model is analogous to one urn where multiple draws occur with replacement, which has been recently studied in detail [20]. However, the analysis in [20] is carried out in an aggregate sense, that is, only for the entire urn and not individual processes. Unfortunately, this aggregate approach does not work in a network setting, whereas the super-urn model proposed here is applicable.

##### A. Complete Network Marginal Distributions

We first focus on the special case of complete networks to derive some useful probability distributions; later on, we will obtain other stochastic properties that apply to more general networks.

Given that the network is complete, we focus on one of the nodes, say,  $i \in V$ . For ease of notation, we define  $\bar{T}_j = \sum_{k=1}^N T_k =: \bar{T}$ , and similarly,  $\bar{R}_j =: \bar{R}$ ,  $\bar{B}_j =: \bar{B}$ , for all  $j \in V$ . Defining the events  $A_{n-1} = \{Z_{i,n-1} = a_{n-1}, \dots, Z_{i,1} = a_1\}$  and  $W_{n-1} = \{A_{n-1}, \{Z_{j,1}^{n-1} = b_j^{n-1}\}_{j \neq i}\}$  with  $b_j^{n-1} \in \{0, 1\}^{n-1}$ , and parameters  $\rho = \frac{\bar{R}}{\bar{T}}$  and  $\delta = \frac{N\Delta}{\bar{T}}$ , we can write using (3) under the above assumptions, that

$$\begin{aligned} & P(Z_{i,n} = 1, A_{n-1}) \\ &= \sum_{b_j^{n-1} \in \{0,1\}^{n-1}: j \neq i} P(Z_{i,n} = 1 \mid W_{n-1}) P(W_{n-1}) \\ &= \sum_{b_j^{n-1}: j \neq i} \frac{\bar{R} + \Delta \sum_{t=1}^{n-1} (a_t + \sum_{j \neq i} b_{j,t})}{\bar{T} + \sum_{t=1}^{n-1} \Delta + \sum_{j \neq i} \Delta} P(W_{n-1}) \end{aligned}$$

$$\begin{aligned}
&= \sum_{b_j^{n-1}:j \neq i} \frac{\rho + \frac{\delta}{N} \sum_{t=1}^{n-1} \left( a_t + \sum_{j \neq i} b_{j,t} \right)}{1 + (n-1)\delta} P(W_{n-1}) \\
&= \sum_{b_j^{n-1}:j \neq i} \left[ \rho \frac{P(A_{n-1}, \{Z_j^{n-1} = b_j^{n-1}\}_{j \neq i})}{1 + (n-1)\delta} \right. \\
&\quad + \frac{\delta}{N} \sum_{t=1}^{n-1} \left( a_t \frac{P(A_{n-1}, \{Z_j^{n-1} = b_j^{n-1}\}_{j \neq i})}{1 + (n-1)\delta} \right. \\
&\quad \left. \left. + \sum_{j \neq i} \frac{b_{j,t} P(A_{n-1}, \{Z_j^{n-1} = b_j^{n-1}\}_{j \neq i})}{1 + (n-1)\delta} \right) \right]. \quad (5)
\end{aligned}$$

By examining an arbitrary term  $k \neq i$  in the final sum above, for fixed  $t \in \{1, \dots, n-1\}$ , we can sum out all the other draw variables

$$\begin{aligned}
&\sum_{b_j^{n-1} \in \{0,1\}^{n-1}:j \neq i} b_{k,t} P(A_{n-1}, \{Z_j^{n-1} = b_j^{n-1}\}_{j \neq i}) \\
&= \sum_{b_k^{n-1} \in \{0,1\}^{n-1}} b_{k,t} P(A_{n-1}, Z_k^{n-1} = b_k^{n-1}) \\
&= \sum_{b_{k,t} \in \{0,1\}} b_{k,t} P(A_{n-1}, Z_{k,t} = b_{k,t}) \\
&= P(A_{n-1}, Z_{k,t} = 1). \quad (6)
\end{aligned}$$

Further, by the law of total probability

$$\sum_{b_j^{n-1} \in \{0,1\}^{n-1}:j \neq i} P(A_{n-1}, \{Z_j^{n-1} = b_j^{n-1}\}_{j \neq i}) = P(A_{n-1}). \quad (7)$$

So using (6) and (7), (5) becomes

$$\frac{\rho P(A_{n-1}) + \frac{\delta}{N} \sum_{t=1}^{n-1} \left[ a_t P(A_{n-1}) + \sum_{j \neq i} P(A_{n-1}, Z_{j,t} = 1) \right]}{1 + (n-1)\delta}.$$

Thus, using the law of total probability, we have

$$\begin{aligned}
P(Z_{i,n} = 1) &= \sum_{a^{n-1} \in \{0,1\}^{n-1}} P(Z_{i,n} = 1, A_{n-1}) \\
&= \sum_{a^{n-1}} \frac{\rho P(A_{n-1}) + \frac{\delta}{N} \sum_{t=1}^{n-1} a_t P(A_{n-1})}{1 + (n-1)\delta} \\
&\quad + \frac{\frac{\delta}{N} \sum_{t=1}^{n-1} \sum_{j \neq i} P(A_{n-1}, Z_{j,t} = 1)}{1 + (n-1)\delta} \\
&= \frac{\rho + \frac{\delta}{N} \sum_{t=1}^{n-1} \sum_{j=1}^N P(Z_{j,t} = 1)}{1 + (n-1)\delta}. \quad (8)
\end{aligned}$$

An interesting corollary of this derivation is as follows.

**Lemma 4.1 (Complete network marginal distribution):**

The 1-D marginal distribution of node  $i$ 's contagion draw process  $\{Z_{i,n}\}_{n=1}^{\infty}$  for the  $N$ -node complete network is given by

$$P_{i,n}^{(1)} = P(Z_{i,n} = a) = \rho^a (1 - \rho)^{1-a}$$

where  $i \in V$ ,  $n \geq 1$ , and  $a \in \{0, 1\}$ .

**Proof:** We proceed using a strong induction on  $n \geq 1$ , showing that  $P(Z_{i,n} = 1) = \rho$ , for all nodes  $i \in V$  and all  $n$ . The base case readily holds, since at time  $n = 1$

$$P(Z_{1,1} = 1) = \dots = P(Z_{N,1} = 1) = \frac{\sum_{i=1}^N R_i}{\sum_{i=1}^N T_i} = \rho.$$

Now, assuming that  $P(Z_{j,t} = 1) = \rho$  for all  $j \in V$  and  $t \leq n$  and using (8), we have

$$\begin{aligned}
P(Z_{i,n+1} = 1) &= \frac{\rho + \frac{\delta}{N} \sum_{t=1}^n \sum_{j=1}^N P(Z_{j,t} = 1)}{1 + n\delta} \\
&= \frac{\rho + \sum_{t=1}^n \frac{\delta}{N} N\rho}{1 + n\delta} \\
&= \frac{\rho + \delta \sum_{t=1}^n \rho}{1 + n\delta} = \rho
\end{aligned}$$

which completes the induction argument. The result now follows using the fact that

$$P(Z_{j,n} = 1) + P(Z_{j,n} = 0) = 1 \Rightarrow P(Z_{j,n} = 0) = 1 - \rho$$

for all  $j \in V$  and all  $n$ .  $\blacksquare$

We next show that each node's draw process is not stationary, in general, and, hence, is different from the classical Polya $(\rho_c, \delta_c)$  process.

**Remark 4.2 (Non-stationarity of the network contagion process):** Consider a two-node complete network. Then, using (4), one can obtain (after some simplifications) that

$$\begin{aligned}
P(Z_{1,2} = 1, Z_{1,1} = 1) &= \rho \frac{\rho + (1 + \rho)\frac{\delta}{2}}{1 + \delta} \\
P(Z_{1,3} = 1, Z_{1,2} = 1) \\
&= \sum_{\substack{a_1 \in \{0,1\} \\ b^3 \in \{0,1\}^3}} P(Z_{1,1} = a_1, \{Z_{1,t} = 1\}_{t=2}^3, \{Z_{2,t} = b_t\}_{t=1}^3) \\
&= \rho \frac{4\rho + \delta(2 + 14\rho) + \delta^2(6 + 14\rho) + \delta^3(5 + 3\rho)}{4(1 + \delta)^2(1 + 2\delta)}
\end{aligned}$$

and, hence, the network process is not stationary.  $\bullet$

Since every exchangeable process is necessarily stationary, Remark 4.2 implies that the network Polya process is not exchangeable in general. However, some notions of stationarity remain; in our next result, we will see that there is a consistent relationship between the draws at the first and  $n$ th time steps.

**Lemma 4.3 (Complete network  $(n, 1)$ -step marginal probability):** For the complete network, the 2-D marginal probability that node  $i$ 's draw variables at times  $n$  and 1 are both one is given by

$$P(Z_{i,n} = 1, Z_{i,1} = 1) = \rho \frac{\rho + (1 + (N-1)\rho)\frac{\delta}{N}}{1 + \delta}$$

for  $i \in V$ ,  $n \geq 2$ . Furthermore, for any other node  $k$

$$P(Z_{k,n} = 1, Z_{i,1} = 1) = \rho \frac{\rho + (1 + (N-1)\rho)\frac{\delta}{N}}{1 + \delta}.$$

**Proof:** By Lemma 4.1, we have that  $P(Z_{k,1} = 1) = \rho$  for all  $k \in V$ , so it is enough to show that

$$P(Z_{k,n} = 1 \mid Z_{i,1} = 1) = \frac{\rho + (1 + (N-1)\rho)\frac{\delta}{N}}{1 + \delta} \quad (9)$$

for all  $n$  and nodes  $i$  and  $k$ . Using the law of total probability (3) and after some simplifications, with defining  $W_{n-1} = \{Z_{i,2}^{n-1} = a_2^{n-1}, \{Z_{j,1}^{n-1} = b_{j,1}^{n-1}\}_{j \neq i}\}$  (note that  $W_{n-1}$  is a function of  $\{b_{j,1}^{n-1}\}_{j \neq i}$ , but for simplicity we omit this), we have that

$$\begin{aligned} & P(Z_{i,n} = 1 \mid Z_{i,1} = 1) \\ &= \sum_{\substack{a_2^{n-1} \in \{0,1\}^{n-2} \\ b_{j,1}^{n-1} \in \{0,1\}^{n-1}: j \neq i}} P(Z_{i,n} = 1 \mid Z_{i,1} = 1, W_{n-1}) \\ & \quad \times P(W_{n-1} \mid Z_{i,1} = 1) \\ &= \sum_{a_2^{n-1}, b_{j,1}^{n-1}: j \neq i} \frac{\rho + \frac{\delta}{N}(1 + \sum_{t=2}^{n-1} a_t + \sum_{t=1}^{n-1} \sum_{j \neq i} b_{j,t})}{1 + (n-1)\delta} \\ & \quad \times P(W_{n-1} \mid Z_{i,1} = 1) \\ &= \sum_{a_2^{n-1}, b_{j,1}^{n-1}: j \neq i} \left[ \left( \rho + \frac{\delta}{N} \right) \frac{P(W_{n-1} \mid Z_{i,1} = 1)}{1 + (n-1)\delta} \right. \\ & \quad + \frac{\delta}{N} \sum_{t=2}^{n-1} \frac{a_t P(W_{n-1} \mid Z_{i,1} = 1)}{1 + (n-1)\delta} \\ & \quad \left. + \frac{\delta}{N} \sum_{t=1}^{n-1} \frac{b_{j,t} P(W_{n-1} \mid Z_{i,1} = 1)}{1 + (n-1)\delta} \right]. \end{aligned}$$

Then, after arranging terms and using the law of total probability for

$$\sum_{a_2^{n-1}, b_{j,1}^{n-1}: j \neq i} b_{j,t} P(W_{n-1} \mid Z_{i,1} = 1) = P(Z_{j,t} = 1 \mid Z_{i,1} = 1)$$

we have

$$\begin{aligned} & P(Z_{i,n} = 1 \mid Z_{i,1} = 1) \\ &= \frac{(\rho + \frac{\delta}{N})(1)}{1 + (n-1)\delta} + \frac{\frac{\delta}{N} \sum_{t=2}^{n-1} P(Z_{i,t} = 1 \mid Z_{i,1} = 1)}{1 + (n-1)\delta} \\ & \quad + \frac{\frac{\delta}{N} \sum_{t=1}^{n-1} \sum_{j \neq i} P(Z_{j,t} = 1 \mid Z_{i,1} = 1)}{1 + (n-1)\delta} \\ &= \frac{\rho + \frac{\delta}{N} \sum_{j \neq i} P(Z_{j,1} = 1)}{1 + (n-1)\delta} \\ & \quad + \frac{\frac{\delta}{N} \left[ 1 + \sum_{j=1}^N \sum_{t=2}^{n-1} P(Z_{j,t} = 1 \mid Z_{i,1} = 1) \right]}{1 + (n-1)\delta} \\ &= \frac{\rho(1 + (N-1)\frac{\delta}{N})}{1 + (n-1)\delta} \\ & \quad + \frac{\frac{\delta}{N} \left[ 1 + \sum_{j=1}^N \sum_{t=2}^{n-1} P(Z_{j,t} = 1 \mid Z_{i,1} = 1) \right]}{1 + (n-1)\delta}. \quad (10) \end{aligned}$$

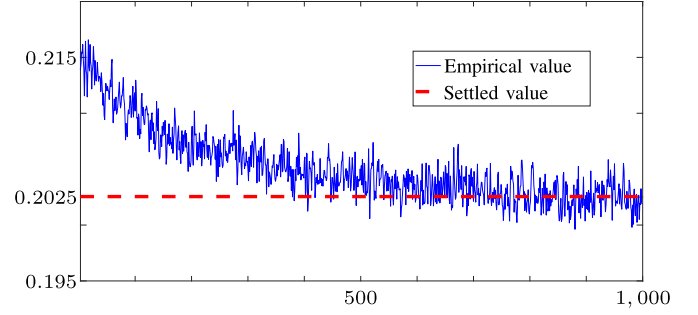


Fig. 2. Simulated values for  $P(Z_{i,n} = 1, Z_{i,n-1} = 1)$  for an arbitrary node  $i$  averaged over 50 000 simulated trials on the network shown in Fig. 3(d). All parameters remained constant throughout all trials; see <http://bit.ly/2tnBix5> for a complete list of them.

It can be similarly shown by symmetry of the complete network that (10) holds for  $P(Z_{k,n} = 1 \mid Z_{i,1} = 1)$  if  $k \neq i$ .

In order to show (9), we proceed using strong induction on  $n \geq 2$ . For the base case, setting  $n = 2$  in (10), we have for any  $i, k \in V$

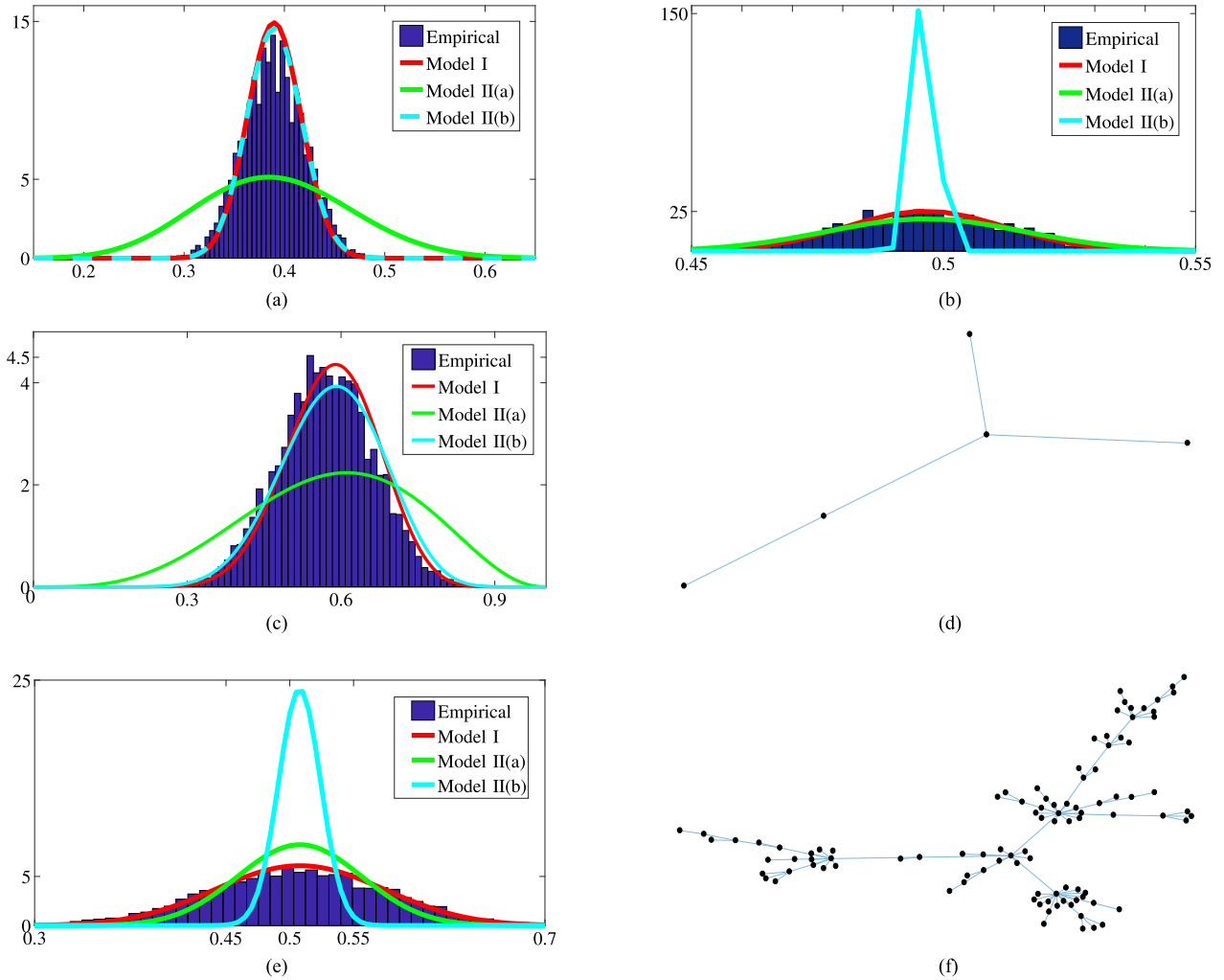
$$\begin{aligned} P(Z_{k,2} = 1 \mid Z_{i,1} = 1) &= \frac{\rho(1 + (N-1)\frac{\delta}{N}) + \frac{\delta}{N}}{1 + \delta} \\ &= \frac{\rho + (1 + (N-1)\rho)\frac{\delta}{N}}{1 + \delta} \end{aligned}$$

as desired. Assume now that  $P(Z_{k,t} = 1 \mid Z_{i,1} = 1)$  is given by (9), for  $2 \leq t \leq n-1$  and any  $i, k \in V$ . Then by (10)

$$\begin{aligned} & P(Z_{k,n} = 1 \mid Z_{i,1} = 1) \\ &= \frac{\rho(1 + (N-1)\frac{\delta}{N})}{1 + (n-1)\delta} \\ & \quad + \frac{\frac{\delta}{N} \left[ 1 + \sum_{j=1}^N \sum_{t=2}^{n-1} P(Z_{j,t} = 1 \mid Z_{i,1} = 1) \right]}{1 + (n-1)\delta} \\ &= \frac{\rho(1 + (N-1)\frac{\delta}{N}) + \frac{\delta}{N} \left[ N(n-2) \frac{\rho + (1 + (N-1)\rho)\frac{\delta}{N}}{1 + \delta} \right]}{1 + (n-1)\delta} \\ &= \frac{1}{1 + (n-1)\delta} \left[ (1 + \delta) \frac{\rho + (1 + (N-1)\rho)\frac{\delta}{N}}{1 + \delta} \right. \\ & \quad \left. + \delta(n-2) \frac{\rho + (1 + (N-1)\rho)\frac{\delta}{N}}{1 + \delta} \right] \\ &= \frac{\rho + (1 + (N-1)\rho)\frac{\delta}{N}}{1 + \delta} \end{aligned}$$

which completes the induction argument.  $\blacksquare$

Although the draw process is not stationary in general, simulated results suggest that it satisfies some asymptotic stationarity properties, in the sense that given sufficient time, the process settles and deviations become very small in magnitude. A representative example is shown in Fig. 2 for the 2-D distribution at times  $n$  and  $n-1$  in the five-node network shown in Fig. 3(d).



**Fig. 3.** Comparison of normalized simulated histograms for the sample average of draws  $\frac{1}{n} \sum_{t=1}^n Z_{i,t}$  and the  $\text{Beta}\left(\frac{\rho_i}{\delta_i}, \frac{1-\rho_i}{\delta_i}\right)$ ,  $\text{Beta}\left(\frac{\rho_i}{\delta_i}, \frac{1-\rho_i}{\delta_i}\right)$ , and  $\text{Beta}\left(\frac{\rho_i}{\delta_i}, \frac{1-\rho_i}{\delta_i}\right)$  pdfs from Models I, II(a), and II(b), respectively, for an arbitrary node  $i$  with  $n = 1000$ , averaged over 5000 simulated trials. Here, the parameters  $\Delta$  as well as  $R_i$  and  $B_i$ , for all  $i \in V$ , were uniformly randomly assigned for each network, and were consistent throughout all trials. See <http://bit.ly/2tnBix5> for a complete list of all parameters used for each network. (a) 10-node complete network histogram. (b) 100-node complete network histogram. (c) 5-node Barabasi-Albert network histogram. (d) 5-node Barabasi-Albert network [23]. (e) 100-node Barabasi-Albert network histogram. (f) 100-node Barabasi-Albert network [23].

## B. Martingale Theorems

We now turn our attention to the martingale properties of the network contagion process, where we do not assume that the network is necessarily complete. Recall that by the martingale convergence theorem [15], [16], if a process  $\{Z_n\}_{n=1}^\infty$  is a martingale (or supermartingale, or submartingale), there exists a random variable  $Z$  such that  $\{Z_n\}_{n=1}^\infty$  converges almost surely to  $Z$  as  $n \rightarrow \infty$ .

**Theorem 4.4 (Individual Urn proportion martingale):** For a network  $\mathcal{G} = (V, \mathcal{E})$ ,  $\Delta_{r,i}(n) = \Delta_{b,i}(n) = \Delta$ , and  $T_i = T$ , for all  $i \in V$  and all  $n$ , the individual proportion of red balls  $\{U_{i,n}\}_{n=1}^\infty$  is a martingale with respect to the draws for the whole network  $\{Z_n\}_{n=1}^\infty = \{(Z_{1,n}, \dots, Z_{N,n})\}_{n=1}^\infty$  if and only if, almost surely

$$\frac{1}{|\mathcal{N}'_i|} \sum_{j \in \mathcal{N}'_i} U_{j,n-1} = U_{i,n-1}.$$

**Proof:** Using the expression for  $U_{i,n}$ , (2) and (3), we have almost surely

$$\begin{aligned} & E[U_{i,n} \mid Z_{n-1}] \\ &= E \left[ \frac{\Delta Z_{i,n} + U_{i,n-1}(T + (n-1)\Delta)}{T + n\Delta} \mid Z_{n-1} \right] \\ &= \frac{U_{i,n-1}(T + (n-1)\Delta)}{T + n\Delta} + \frac{\Delta E[Z_{i,n} \mid Z_{n-1}]}{T + n\Delta} \\ &= U_{i,n-1} \frac{T + (n-1)\Delta}{T + n\Delta} + \frac{\Delta P(Z_{i,n} = 1 \mid Z_{n-1})}{T + n\Delta} \\ &= U_{i,n-1} \left( 1 - \frac{\Delta}{T + n\Delta} \right) + \frac{\Delta \sum_{j \in \mathcal{N}'_i} U_{j,n-1}(T + (n-1)\Delta)}{(T + n\Delta)|\mathcal{N}'_i|(T + (n-1)\Delta)} \\ &= U_{i,n-1} - \frac{\Delta U_{i,n-1}}{T + n\Delta} + \Delta \frac{U_{i,n-1} + \sum_{j \in \mathcal{N}'_i} U_{j,n-1}}{|\mathcal{N}'_i|(T + n\Delta)} \end{aligned}$$

$$\begin{aligned}
&= U_{i,n-1} + \Delta \frac{\left[ \sum_{j \in \mathcal{N}_i} U_{j,n-1} \right] - |\mathcal{N}_i| U_{i,n-1}}{(T + n\Delta)(|\mathcal{N}_i| + 1)} \\
&= U_{i,n-1} + \frac{\Delta \sum_{j \in \mathcal{N}_i} (U_{j,n-1} - U_{i,n-1})}{(T + n\Delta)(|\mathcal{N}_i| + 1)}. \quad (11)
\end{aligned}$$

This implies that  $\{U_{i,n}\}_{n=1}^\infty$  is a martingale with respect to  $\{Z_n\}_{n=1}^\infty$  if and only if

$$\sum_{j \in \mathcal{N}_i} U_{j,n-1} - U_{i,n-1} = 0 \Leftrightarrow \frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} U_{j,n-1} = U_{i,n-1}$$

almost surely.  $\blacksquare$

If the condition in Theorem 4.4 holds, then for any  $i$ , both  $U_{i,n}$  and  $\frac{1}{n} \sum_{t=1}^n Z_{i,t}$  converge almost surely to a limit as  $n \rightarrow \infty$ . However, the condition of Theorem 4.4, barring the trivial single node scenario, is not verifiable. To resolve this issue, we instead examine the evolution of the average proportion of red balls (i.e., the susceptibility) in a regular network.

**Theorem 4.5 (Regular network susceptibility martingale):** For a regular network  $\mathcal{G} = (V, \mathcal{E})$  with  $\Delta_{r,i}(n) = \Delta_{b,i}(n) = \Delta$  and  $T_i = T$  for all nodes  $i \in V$  and times  $n$ , the network susceptibility  $\{\tilde{U}_n\}_{n=1}^\infty$ , where  $\tilde{U}_n = \frac{1}{N} \sum_{i=1}^N U_{i,n}$ , is a martingale with respect to  $\{Z_n\}_{n=1}^\infty$ .

**Proof:** We have, similar to Theorem 4.4, that

$$\begin{aligned}
&E[\tilde{U}_n \mid Z_{n-1}] \\
&= \frac{1}{N} \sum_{i=1}^N E[U_{i,n} \mid Z_{n-1}] \\
&= \frac{1}{N} \sum_{i=1}^N \left[ U_{i,n-1} + \frac{\Delta \sum_{j \in \mathcal{N}_i} (U_{j,n-1} - U_{i,n-1})}{(T + n\Delta)(|\mathcal{N}_i| + 1)} \right] \\
&= \tilde{U}_{n-1} + \sum_{i=1}^N \frac{\Delta \sum_{j \in \mathcal{N}_i} (U_{j,n-1} - U_{i,n-1})}{N(T + n\Delta)(|\mathcal{N}_i| + 1)}.
\end{aligned}$$

Let us examine the second term of the last equality. If this term is zero,  $\{\tilde{U}_n\}_{n=1}^\infty$  is a martingale with respect to  $\{Z_n\}_{n=1}^\infty$ . We now define the adjacency matrix  $[a_{ij}]$  of our network, where the  $(i, j)$ th entry  $a_{ij}$  is 1 if  $(i, j) \in \mathcal{E}$ , and 0 otherwise. Since we assumed that our network was undirected,  $[a_{ij}]$  is symmetric, that is,  $a_{ij} = a_{ji}$  for all  $i, j \in V$ . So

$$\begin{aligned}
&\sum_{i=1}^N \frac{\Delta \sum_{j \in \mathcal{N}_i} (U_{j,n-1} - U_{i,n-1})}{N(T + n\Delta)(|\mathcal{N}_i| + 1)} \\
&= \sum_{i=1}^N \frac{\Delta \sum_{j=1}^N a_{ij} (U_{j,n-1} - U_{i,n-1})}{N(T + n\Delta)(|\mathcal{N}_i| + 1)} \\
&= \frac{\Delta}{N(T + n\Delta)} \sum_{i=1}^N \sum_{j=1}^N \frac{a_{ij} (U_{j,n-1} - U_{i,n-1})}{|\mathcal{N}_i| + 1}.
\end{aligned}$$

Now, we examine the sum of the  $(i, j)$  and  $(j, i)$  components of the double sum, where  $(i, j) \in \mathcal{E}$  (otherwise, both terms are

zero). Recall that  $(i, i) \notin \mathcal{E}, \forall i$ . We have

$$\begin{aligned}
&\frac{a_{ij} (U_{j,n-1} - U_{i,n-1})}{|\mathcal{N}_i| + 1} + \frac{a_{ji} (U_{i,n-1} - U_{j,n-1})}{|\mathcal{N}_j| + 1} \\
&= \frac{a_{ij} (U_{j,n-1} - U_{i,n-1})(|\mathcal{N}_j| + 1)}{(|\mathcal{N}_i| + 1)(|\mathcal{N}_j| + 1)} \\
&\quad + \frac{a_{ji} (U_{i,n-1} - U_{j,n-1})(|\mathcal{N}_i| + 1)}{(|\mathcal{N}_i| + 1)(|\mathcal{N}_j| + 1)} \\
&= \frac{a_{ij} (|\mathcal{N}_j| - |\mathcal{N}_i|)}{(|\mathcal{N}_i| + 1)(|\mathcal{N}_j| + 1)} (U_{j,n-1} - U_{i,n-1}).
\end{aligned}$$

From above, it is clear that this term is zero for all  $i$  and  $j$  by setting  $|\mathcal{N}_j| = |\mathcal{N}_i|$ , that is, in any regular network, and so  $\{\tilde{U}_n\}_{n=1}^\infty$  is a martingale with respect to  $\{Z_n\}_{n=1}^\infty$ .  $\blacksquare$

We next allow the net number of black balls  $\Delta_{b,i}(\cdot)$  to evolve stochastically in time as a function of the past draw history in the network in order to steer  $\{U_{i,n}\}_{n=1}^\infty$  to a limit for every node  $i$ .

**Theorem 4.6 (Individual Urn proportion categories):** In a general network  $\mathcal{G} = (V, \mathcal{E})$  with  $\Delta_{r,i}(n) = \Delta_r$  for all  $n \in \mathbb{Z}_{\geq 1}$  and  $i \in V$ , if we choose  $\{\Delta_{b,i}(n)\}_{n=1}^\infty$  so that

$$\Delta_{b,i}(n) \geq \frac{\Delta_r (1 - U_{i,n-1}) S_{i,n-1}}{U_{i,n-1} (1 - S_{i,n-1})}$$

almost surely for all  $n \in \mathbb{Z}_{\geq 1}$  and  $i \in V$  (respectively, equal to, less than or equal to), then  $\{U_{i,n}\}_{n=1}^\infty$  is a supermartingale (respectively, martingale, submartingale) with respect to  $\{Z_n\}_{n=1}^\infty$ .

**Proof:** We will start with the case of a supermartingale. That is, we wish to show that almost surely for all  $n \in \mathbb{Z}_{\geq 1}$

$$E[U_{i,n} \mid Z_{n-1}] - U_{i,n-1} \leq 0.$$

Define  $\bar{Z}_{i,n} = \sum_{t=1}^n Z_{i,t}$ , and take  $X_{i,n}$  as in (1). Then, we have almost surely

$$\begin{aligned}
&U_{i,n} - U_{i,n-1} \\
&= \frac{R_i + \Delta_r (\bar{Z}_{i,n-1} + Z_{i,n})}{X_{i,n}} - \frac{R_i + \Delta_r \bar{Z}_{i,n-1}}{X_{i,n-1}} \\
&= \frac{\Delta_r Z_{i,n}}{X_{i,n}} - \frac{(R_i + \Delta_r \bar{Z}_{i,n-1})(X_{i,n} - X_{i,n-1})}{X_{i,n-1} X_{i,n}} \\
&= \frac{\Delta_r Z_{i,n}}{X_{i,n}} - \frac{U_{i,n-1} (X_{i,n} - X_{i,n-1})}{X_{i,n}} \\
&= \frac{1}{X_{i,n}} \left[ \Delta_r Z_{i,n} - U_{i,n-1} (\Delta_r Z_{i,n} + \Delta_{b,i}(n) (1 - Z_{i,n})) \right]
\end{aligned}$$

since  $X_{i,n} > 0$  for all  $n \in \mathbb{Z}_{\geq 1}$  almost surely, we can ignore it. Now, since  $U_{i,n-1}$  is almost surely constant given  $Z_{n-1}$

$$E[U_{i,n} \mid Z_{n-1}] - U_{i,n-1} \leq 0 \Rightarrow E[U_{i,n} - U_{i,n-1} \mid Z_{n-1}] \leq 0.$$

That is, we wish to check if, almost surely

$$E[\Delta_r Z_{i,n} (1 - U_{i,n-1}) - \Delta_{b,i}(n) U_{i,n-1} (1 - Z_{i,n}) \mid Z_{n-1}] \leq 0.$$

Now if

$$\Delta_{b,i}(n) \geq \frac{\Delta_r (1 - U_{i,n-1}) S_{i,n-1}}{U_{i,n-1} (1 - S_{i,n-1})}$$



almost surely, we have

$$\begin{aligned}
& E[\Delta_r Z_{i,n}(1 - U_{i,n-1}) - \Delta_{b,i}(n)U_{i,n-1}(1 - Z_{i,n})|Z_{n-1}] \\
& \leq E\left[\Delta_r Z_{i,n}(1 - U_{i,n-1}) - U_{i,n-1}(1 - Z_{i,n})\right. \\
& \quad \left. \times \frac{\Delta_r(1 - U_{i,n-1})S_{i,n-1}}{U_{i,n-1}(1 - S_{i,n-1})} \middle| Z_{n-1}\right] \\
& = \Delta_r(1 - U_{i,n-1})E\left[Z_{i,n} - (1 - Z_{i,n})\frac{S_{i,n-1}}{1 - S_{i,n-1}} \middle| Z_{n-1}\right] \\
& = \Delta_r(1 - U_{i,n-1})\left[S_{i,n-1} - (1 - S_{i,n-1})\frac{S_{i,n-1}}{1 - S_{i,n-1}}\right] \\
& = 0
\end{aligned}$$

where the second to last equality comes from the fact that  $E[Z_{i,n}|Z_{n-1}] = P(Z_{i,n} = 1|Z_{n-1}) = S_{i,n-1}$  almost surely by (3), and that  $S_{i,n-1}$  is almost surely constant given  $Z_{n-1}$ . Thus, as long as  $\Delta_{b,i}(n)$  obeys this bound almost surely for all  $n \in \mathbb{Z}_{\geq 1}$ ,  $\{U_{i,n}\}_{n=1}^{\infty}$  is a supermartingale with respect to  $\{Z_n\}_{n=1}^{\infty}$ . Similarly, if  $\Delta_{b,i}(n)$  is almost surely equal (respectively, less than or equal) to this bound,  $\{U_{i,n}\}_{n=1}^{\infty}$  is a martingale (respectively, submartingale) with respect to  $\{Z_n\}_{n=1}^{\infty}$ . ■

Theorem 4.6 tells us what bounds for  $\{\Delta_{b,i}(n)\}_{n=1}^{\infty}$  must be obeyed almost surely to guarantee that  $\{U_{i,n}\}_{n=1}^{\infty}$  admits an asymptotic limit for all  $i \in V$  in any general network. For instance, this tells us that by choosing  $\Delta_{b,i}(t) = 0$  almost surely for all  $i \in V$  and  $t \in \mathbb{Z}_{\geq 1}$ ,  $\{U_{i,n}\}_{n=1}^{\infty}$  will be a submartingale and will converge to some limiting random variable. While this result is interesting for modeling contagion, it is especially useful in the context of curing.

## V. MODEL APPROXIMATIONS

As previously noted, the dynamics of the network contagion process are complicated, especially when considered on general networks. For this reason, in this section, we develop two useful approximations to this process on a general network that allow us to shed some light on its asymptotic behavior. Throughout this section, unless stated otherwise, we consider general network topologies with  $\Delta_{r,i}(t) = \Delta_{b,i}(t) = \Delta$  for all  $t \in \mathbb{Z}_{\geq 1}$  and  $i \in V$ . However, to match the 1-step and  $(n, 1)$ -step distributions, we make the simplifying assumption that the neighborhood of each node  $i$  can be represented as a complete network, i.e., all of its neighbors are connected to one another, in order to apply Lemmas 4.1 and 4.3.

### A. Approximation: Computational Model

We now introduce our first approximation technique, where we approximate the contagion process of each node in the network with a classical Polya urn process.

**Model I (Computational model):** We approximate the dynamics of any node  $i$ 's contagion process using a classical Polya

process  $\text{Polya}(\rho_c = \rho_i, \delta_c = \hat{\delta}_i)$ , with

$$\rho_i = \frac{\sum_{j \in \mathcal{N}'_i} R_j}{\sum_{j \in \mathcal{N}'_i} T_j}, \quad \text{and} \quad \hat{\delta}_i = \arg \min_{\delta} \frac{1}{n} D\left(P_{i,n}^{(n)} \| Q_{\rho_i, \delta}^{(n)}\right)$$

with

$$Q_{\rho_i, \hat{\delta}_i}^{(n)}(a^n) = \frac{\Gamma\left(\frac{1}{\hat{\delta}_i}\right) \Gamma\left(\frac{\rho_i}{\hat{\delta}_i} + \bar{a}^n\right) \Gamma\left(\frac{1-\rho_i}{\hat{\delta}_i} + n - \bar{a}^n\right)}{\Gamma\left(\frac{1}{\hat{\delta}_i} + n\right) \Gamma\left(\frac{\rho_i}{\hat{\delta}_i}\right) \Gamma\left(\frac{1-\rho_i}{\hat{\delta}_i}\right)}$$

where  $\Gamma(\cdot)$  is the Gamma function,  $a^n = (a_1, \dots, a_n) \in \{0, 1\}^n$ , and  $\bar{a}^n = a_1 + \dots + a_n$ . •

Here  $\rho_c$  is chosen to be the proportion of red balls  $\rho_i$  in the node's super-urn, so that the 1-D distributions of the classical Polya process and the node process  $\{Z_{i,n}\}$  coincide, whereas  $\hat{\delta}_i$  is set by performing a minimization to find the value that best fits  $Q_{\rho_i, \hat{\delta}_i}^{(n)}$  to the distribution of  $\{Z_{i,n}\}_{n=1}^{\infty}$  of node  $i \in V$ . We use a divergence measure, denoted by  $D(\cdot \| \cdot)$ , to observe the quality of the fit.

The explicit derivation of the distribution  $Q_{\rho_i, \hat{\delta}_i}^{(n)}$  can be found in [17] and [21]. This method ensures that the fit of  $Q_{\rho_i, \hat{\delta}_i}^{(n)}$  is as close as possible under the given divergence measure. Since we are measuring the error in using an approximating distribution, we use the Kullback–Leibler divergence [22]; we thus have that

$$\begin{aligned}
\hat{\delta}_i & = \arg \min_{\delta} \frac{1}{n} \sum_{a^n \in \{0,1\}^n} P_{i,n}^{(n)}(a^n) \log \frac{P_{i,n}^{(n)}(a^n)}{Q_{\rho_i, \delta}^{(n)}(a^n)} \\
& = \arg \max_{\delta} \frac{1}{n} \sum_{a^n \in \{0,1\}^n} P_{i,n}^{(n)}(a^n) \log Q_{\rho_i, \delta}^{(n)}(a^n)
\end{aligned}$$

since  $P_{i,n}^{(n)}(a^n) \log P_{i,n}^{(n)}(a^n)$  is independent of  $\hat{\delta}_i$ . The approximating process is stationary and exchangeable, as it is a classical Polya process. We also know (from Section II) that it is nonergodic with its sample average converging almost surely to the  $\text{Beta}\left(\frac{\rho_i}{\hat{\delta}_i}, \frac{1-\rho_i}{\hat{\delta}_i}\right)$  distribution. Calculating an analytic expression for the minimizing  $\hat{\delta}_i$  is not feasible in general, and hence should be performed computationally. However, due to the above minimization, the value of  $\hat{\delta}_i$  is, by definition, the best way to fit a Polya process to the process  $\{Z_{i,n}\}_{n=1}^{\infty}$  for a given  $n$ .

### B. Approximation: Analytical Models

An alternative to Model I is to attempt to find approximations whose parameters can be determined analytically.

**Model II(a) (Large-network analytical model):** For any given node  $i$ , we approximate the dynamics of its process  $\{Z_{i,n}\}_{n=1}^{\infty}$  by using a classical Polya process  $\text{Polya}(\rho_c = \rho_i, \delta_c = \delta'_i)$ , with

$$\rho_i = \frac{\sum_{j \in \mathcal{N}'_i} R_j}{\sum_{j \in \mathcal{N}'_i} T_j}, \quad \text{and} \quad \delta'_i = \frac{\delta_i}{N + (N - 1)\delta_i}$$

where  $\delta_i = \frac{N\Delta}{\sum_{j \in \mathcal{N}'_i} T_j}$ . •

Here the parameters of the classical Polya process are chosen by directly matching its first and  $(n, 1)$ -step second-order

statistics with those of  $\{Z_{i,n}\}_{n=1}^{\infty}$ . This method avoids the computational burden of the previous model by yielding an analytical expression for the correlation parameter of the classical Polya process.

We next prove that under some stationarity and symmetry assumptions, the contagion process running on each node in the network is statistically identical to the classical Polya process of Model II(a).

**Lemma 5.1 (Exact representation):** Suppose that

- 1)  $P(Z_{i,1} = 1 | Z_{j,1}^{n-1} = a^{n-1}) = \rho_i$ , and
- 2)  $P(Z_{i,t} = 1 | Z_{j,1}^{n-1} = a^{n-1}) = P(Z_{k,n} = 1 | Z_{j,1}^{n-1} = a^{n-1})$

for all  $n \geq 1, 2 \leq t < n, i, j, k \in V, a^{n-1} \in \{0, 1\}^{n-1}$ . Then for any node  $i$  in a complete network,  $\{Z_{i,n}\}_{n=1}^{\infty}$  is given exactly by the Polya( $\rho_i, \delta'_i$ ) process.

**Proof:** For any node  $i$ , we wish to show that for all  $n$ , the  $n$ -dimensional distributions of  $\{Z_{i,n}\}_{n=1}^{\infty}$  and the Polya( $\rho_i, \delta'_i$ ) process are identical. It is enough to show that the conditional probability of one event given the whole past is the same, since any joint probability can be written as a product of conditional probabilities. Let us define the events  $A_{n-1} = \{Z_{i,1}^{n-1} = a^{n-1}\}$  and  $B_{n-1} = \{Z_{j,1}^{n-1} = b_{j,1}^{n-1}\}_{j \neq i}$ . Then

$$\begin{aligned}
P_{i|n} &:= P(Z_{i,n} = 1 | A_{n-1}) \\
&= \sum_{b_{j,1}^{n-1} \in \{0,1\}^{n-1}: j \neq i} P(Z_{i,n} = 1 | A_{n-1}, B_{n-1}) P(B_{n-1} | A_{n-1}) \\
&= \sum_{b_{j,1}^{n-1}: j \neq i} \frac{\rho_i + \frac{\delta'_i}{N} \sum_{t=1}^{n-1} (a_t + \sum_{j \neq i} b_{j,t})}{1 + (n-1)\delta'_i} P(B_{n-1} | A_{n-1}) \\
&= \sum_{b_{j,1}^{n-1}: j \neq i} \frac{\rho_i(1 - (N-1)\delta'_i) + \delta'_i \sum_{t=1}^{n-1} (a_t + \sum_{j \neq i} b_{j,t})}{1 + (N(n-2) + 1)\delta'_i} \\
&\quad \times P(B_{n-1} | A_{n-1}) \\
&= \frac{\rho_i(1 - (N-1)\delta'_i) + \delta'_i \sum_{t=1}^{n-1} a_t}{1 + (N(n-2) + 1)\delta'_i} \sum_{b_{j,1}^{n-1}: j \neq i} P(B_{n-1} | A_{n-1}) \\
&+ \frac{\delta'_i}{1 + (N(n-2) + 1)\delta'_i} \sum_{t=1}^{n-1} \sum_{j \neq i} \sum_{b_{j,1}^{n-1}: j \neq i} b_{j,t} P(B_{n-1} | A_{n-1}) \\
&= \frac{(\rho_i(1 - (N-1)\delta'_i) + \delta'_i \sum_{t=1}^{n-1} a_t) \cdot 1}{1 + (N(n-2) + 1)\delta'_i} \\
&+ \frac{\delta'_i \sum_{t=1}^{n-1} \sum_{j \neq i} P(Z_{j,t} = 1 | A_{n-1})}{1 + (N(n-2) + 1)\delta'_i}.
\end{aligned}$$

Then using assumption (1), we have

$$\begin{aligned}
P_{i|n} &= \frac{\rho_i(1 - (N-1)\delta'_i) + \delta'_i \sum_{t=1}^{n-1} a_t + \delta'_i(N-1)\rho_i}{1 + (N(n-2) + 1)\delta'_i} \\
&+ \frac{\delta'_i \sum_{t=2}^{n-1} \sum_{j \neq i} P(Z_{j,t} = 1 | A_{n-1})}{1 + (N(n-2) + 1)\delta'_i} \\
&= \frac{\rho_i + \delta'_i \sum_{t=1}^{n-1} \left[ a_t + \sum_{j \neq i} P(Z_{j,t} = 1 | A_{n-1}) \right]}{1 + (N(n-2) + 1)\delta'_i}
\end{aligned}$$

Now using assumption (2), we get

$$\begin{aligned}
P_{i|n} &= \frac{\rho_i + \delta'_i \left( \sum_{t=1}^{n-1} a_t + \sum_{t=2}^{n-1} \sum_{j \neq i} P_{i|n} \right)}{1 + (N(n-2) + 1)\delta'_i} \\
&= \frac{\rho_i + \delta'_i \left( \sum_{t=1}^{n-1} a_t + (n-2)(N-1)P_{i|n} \right)}{1 + (N(n-2) + 1)\delta'_i}.
\end{aligned}$$

Thus, we have that

$$\begin{aligned}
P_{i|n} &= \frac{\rho_i + \delta'_i \left( \sum_{t=1}^{n-1} a_t + (n-2)(N-1)P_{i|n} \right)}{1 + (N(n-2) + 1)\delta'_i} \\
\Rightarrow P_{i|n} &= \frac{\rho_i + \delta'_i \sum_{t=1}^{n-1} a_t}{1 + (n-1)\delta'_i}
\end{aligned}$$

which is the conditional probability  $P(Z_n = 1 | Z_1^{n-1} = a^{n-1})$  for a Polya( $\rho_i, \delta'_i$ ) process. A similar calculation can be performed for  $P(Z_{i,n} = 0 | Z_1^{n-1} = a^{n-1})$ . ■

Unfortunately in a general network setting assumptions (1) and (2) above do not hold true. However, this result motivates the fact that this analytical approximation is reasonable to use for situations where these assumptions hold within tolerable margins of error; empirical evidence indicates that this occurs for large values of  $N$ , since as  $N$  increases the quality of the fit improves. This approximation, nevertheless, drastically reduces the complexity in analyzing the individual contagion draw processes, as closed-form expressions for the process parameters are available.

**Model II(b) (Small-network analytic model):** Given any node  $i$  in the network with a small to moderate number of nodes, we approximate the dynamics of its contagion process  $\{Z_n\}_{n=1}^{\infty}$  using a Polya( $\rho_i, \delta_i^*$ ) process, where

$$\begin{aligned}
\rho_i &= \frac{\sum_{j \in \mathcal{N}'_i} R_j}{\sum_{j \in \mathcal{N}'_i} T_j}, \quad \text{and} \\
\delta_i^* &= \frac{\delta_i/N}{N + (N-1)\delta_i/N} = \frac{\delta_i}{N^2 + (N-1)\delta_i}
\end{aligned}$$

where  $\delta_i = \frac{N\Delta}{\sum_{j \in \mathcal{N}'_i} T_j}$ . •

The idea behind this model is that we want to remove the dependence on the number of nodes  $N$  from the parameter  $\delta_i = \frac{N\Delta}{T_i}$ , and so we divide each instance of  $\delta_i$  in  $\delta_i^*$  by  $N$ . The idea is that as  $n$  grows, it eventually becomes significantly larger than the relatively small number of nodes  $N$ , and so  $n|\mathcal{N}'_i| \approx n$  for all  $i \in V$ . Hence, we may consider that for a sufficiently large time, we have added  $n\Delta$  balls to the super-urn. Effectively, this means, we are using a correlation parameter of  $\frac{\Delta}{T_i}$  instead of  $\delta = \frac{N\Delta}{T_i}$ . Simulation results confirm that this approximation captures the limit distribution of the original process better than Model II(a) when the number of nodes is small. Fig. 3 displays this relationship. A summary of all models presented in this section, and the scenarios under which they are most suitable, is provided in Table I.

We close this section with numerical demonstrations on the fitness of all models. Fig. 3 shows a representative comparison between the Beta( $\frac{\rho_i}{\delta'_i}, \frac{1-\rho_i}{\delta'_i}$ ) pdf and the simulated histogram of  $\frac{1}{n} \sum_{t=1}^n Z_{i,n}$ , where  $n = 1000$ , for an arbitrary node  $i$  in the

TABLE I  
APPROXIMATION USAGE SCENARIOS

Model	Usage Scenario
I	Exactness valued over analytic simplicity
II(a)	Larger values of $N$ , i.e., large network
II(b)	Small to moderate values of $N$ , i.e., small network

given networks. Recall that the  $\text{Beta}(\frac{\rho_i}{\delta_i}, \frac{1-\rho_i}{\delta_i})$ ,  $\text{Beta}(\frac{\rho_i}{\delta_i}, \frac{1-\rho_i}{\delta_i})$ , and  $\text{Beta}(\frac{\rho_i}{\delta_i}, \frac{1-\rho_i}{\delta_i})$  pdfs are the distributions of the limit random variables to which the sample average of the draw processes of Models I, II(a), and II(b), respectively, converge almost surely, as  $n \rightarrow \infty$  (see Section II). We use complete networks since they satisfy the assumption that all neighborhoods are complete, as well as Barabasi–Albert networks that have been shown to be a good model for real-world social networks [23] and do not satisfy this assumption; however, our results show that the approximations still fit quite well. As expected, Model I provides the best approximation in all scenarios, albeit without an analytic expression for its parameters, which can provide insight into the behavior of the underlying process. Model II(a) fits quite well when the number of nodes in the network is large, as seen in Fig. 3(b) and (e), but fits poorly for a small number of nodes, which is evident in Fig. 3(a) and (c). Model II(b) is the complement of Model II(a) in the sense that it fits very well for a small number of nodes but poorly for a large network. Hence, if analytic expressions for parameters are desired, Models II(a) and II(b) can be used depending on the number of nodes to provide approximations that are marginally worse than the computational exactness of Model I.

### C. Comparison With SIS Model

We now provide a number of empirical results in which we compare our model, with both finite and infinite memory, to the traditional discrete time SIS model [24]. In the SIS model, the parameter  $\delta_{\text{SIS}}$  denotes the probability that a node will recover from infection, and  $\beta_{\text{SIS}}$  is the probability that a node will become infected through contact with a single infected neighbor. The dynamics are described through the probability that any node  $i$  will be infected at time  $t$ ,  $P_i(t)$ , which evolves according to the equation

$$P_i(t+1) = P_i(t)(1 - \delta_{\text{SIS}}) + (1 - P_i(t)) \left( 1 - \prod_{j \in \mathcal{N}_i} (1 - \beta_{\text{SIS}} P_j(t)) \right).$$

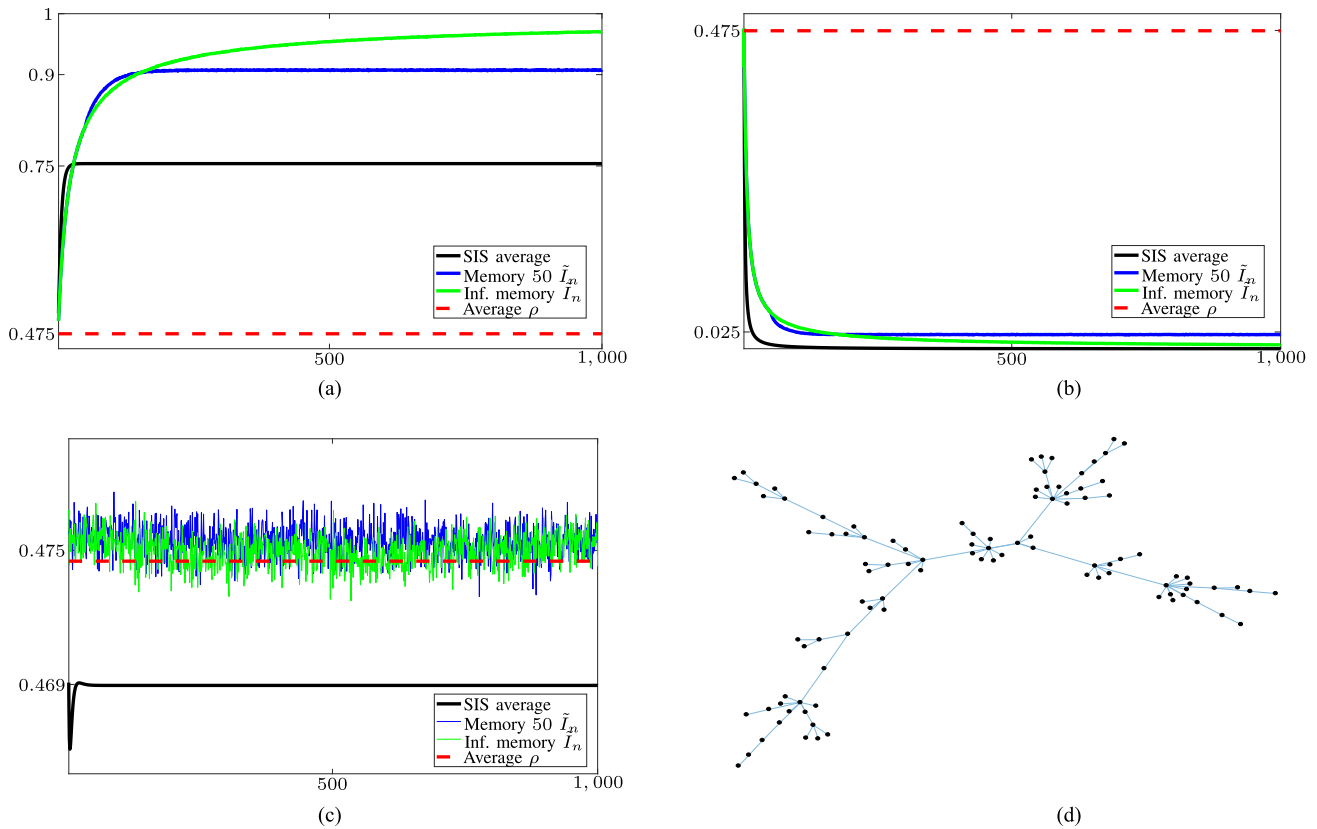
Note in particular that this model exhibits Markovian behavior, since the evolution of the process depends only on the probability of infection from the previous time step. We make the simplifying assumption that  $\delta_{\text{SIS}}$  and  $\beta_{\text{SIS}}$  remain the same for all nodes and throughout time, and hence we will compare it with the network Polya contagion process when  $\Delta_r$  and  $\Delta_b$  are similarly fixed in time and throughout the network.

The concept of an epidemic threshold for the SIS model gives a value through which one may determine whether the epidemic dies, *a priori* using only the system parameters [24]. The threshold condition is directly related to the largest-magnitude eigenvalue  $\lambda_{\text{max}}$  of the adjacency matrix of the underlying graph of the network, and states that if  $\delta_{\text{SIS}} > \beta_{\text{SIS}} \lambda_{\text{max}}$  then the epidemic will be eliminated after some time  $n$ , i.e., eventually  $P_i(t) = 0$  for all  $i$  and all  $t > n$ . Furthermore, it has been shown that this threshold is tight, and indeed if  $\delta_{\text{SIS}} < \beta_{\text{SIS}} \lambda_{\text{max}}$  then some nonzero convergence point exists, called an *endemic state*, and the epidemic will never be eliminated [25].

Fig. 4 compares the behavior of the SIS model and the network Polya contagion process for different selections of these parameters. The initial probabilities of infection  $P_i(0)$  for the SIS model were set to coincide with the initial individual proportions of red balls for the nodes  $\frac{R_i}{T_i}$ . Further, we relate in Fig. 4(a)–(c) the parameters  $\beta_{\text{SIS}}$  and  $\delta_{\text{SIS}}$  to  $\Delta_r$  and  $\Delta_b$ , respectively, using ratios of the largest-magnitude eigenvalue  $\lambda_{\text{max}}$  of the adjacency matrix of the graph shown in Fig. 4(d).

Fig. 4(a) shows a comparison when the SIS model is displaying endemic behavior. We see, here, that after a very short time, the SIS model settles and shortly thereafter the finite memory process settles (albeit to a different value), while for the infinite memory process the individual rates of infection and, hence, the average  $\tilde{I}_n$  continue to increase in time. Since both the SIS model and the finite memory process have limited reinforcement, while the infinite memory process does not, these results are to be expected. Fig. 4(b) displays a comparison where the epidemic threshold is met and the epidemic dies out for the SIS model. Here, we note that  $\tilde{I}_n$  for both the infinite and finite memory processes decreases and approaches zero, albeit not as quickly as the SIS model. Hence, we observe that when the curing parameter  $\Delta_b$  is much larger (in fact, more than five folds larger) than the infection parameter  $\Delta_r$  the epidemic is eliminated, as we expect, and this behavior of the SIS model is captured by the network Polya contagion process. However, the finite memory process does not fully approach zero, since the initial conditions  $R_i$  and  $B_i$  have a much larger influence relative to the infinite memory process. Finally, Fig. 4(c) shows the case where the epidemic does not vanish and the parameters in both models are set to be equal ( $\delta_{\text{SIS}} = \beta_{\text{SIS}}$  and  $\Delta_b = \Delta_r$ ). We observe a similar trend between all models, with the finite and infinite memory processes exhibiting near-identical behavior.

Through these observations, we may conclude that both versions of the network Polya contagion process may apply to the modeling of epidemics, albeit in different applications. The finite memory process exhibits behavior that is more closely related to the SIS model since they are both limited reinforcement processes, and hence it may be best suited to traditional biological diseases. The infinite memory process obeys similar trends, but in the endemic state there are some interesting differences since the effects of the infection continue to spread throughout the population. On the other hand, the SIS model quickly settles and does not change in time. Thus with infinite memory, our process is better suited to modeling opinion dynamics, the spread of ideas, and advertising schemes.



**Fig. 4.** Comparison between discrete time susceptible-infected-susceptible (SIS) model average infection rate  $\frac{1}{N} \sum_{i=1}^N P_i(n)$  and the network Polya contagion process average infection rate  $\tilde{I}_n$ . Simulation results were averaged over 5000 trials, and the initial parameters  $R_i$  and  $B_i$  for all nodes were randomly assigned but consistent throughout all trials for a given case. Here  $\lambda_{\max} \approx 5.05$ ,  $\beta_{\text{SIS}} = 0.15$ , and  $\Delta_r = 2$  for all cases, while  $\delta_{\text{SIS}}$  and  $\Delta_b$  were set according to the ratios given above. (a)  $\frac{\delta_{\text{SIS}}}{\beta_{\text{SIS}}} = \frac{\Delta_b}{\Delta_r} = \frac{\lambda_{\max}}{10}$ . (b)  $\frac{\delta_{\text{SIS}}}{\beta_{\text{SIS}}} = \frac{\Delta_b}{\Delta_r} = 1.01\lambda_{\max}$ . (c)  $\frac{\delta_{\text{SIS}}}{\beta_{\text{SIS}}} = \frac{\Delta_b}{\Delta_r} = 1$  and  $\lambda_{\max} > 1$ . (d) 100-node Barabasi–Albert network [23].

## VI. CONCLUSION

We introduced a network epidemics model based on the classical Polya urn scheme, and we investigated its stochastic properties and asymptotic behavior in detail. We showed that under certain conditions the proportion of red balls in individual urns and the network susceptibility, which are processes used to measure infection, admit limits. Three classical Polya processes were proposed, one computational and two analytical, to statistically approximate the contagion process of each node. Empirical results were presented which show that the approximations are a good fit for a range of system parameters. Our process was also compared empirically with the discrete-time SIS model, showing a similar behavior, particularly in the finite memory mode, while providing different degrees of reinforcement in the endemic state, with the largest reinforcement occurring under the infinite memory mode. Future directions of research include investigations into the curing of these processes, and the further study of the network contagion process with finite memory.

## REFERENCES

- [1] M. Hayhoe, F. Alajaji, and B. Ghahsifard, “A Polya urn-based model for epidemics on networks,” in *Proc. Amer. Control Conf.*, 2017, pp. 358–363.
- [2] L. Kim, M. Abramson, K. Drakopoulos, S. Kolitz, and A. Ozdaglar, “Estimating social network structure and propagation dynamics for an infectious disease,” in *Proc. Int. Conf. Soc. Comput., Behav. Cultural Model., Predict.*, 2014, pp. 85–93.
- [3] M. Garetto, W. Gong, and D. Towsley, “Modeling malware spreading dynamics,” in *Proc. IEEE Int. Conf. Comput. Commun.*, 2003, vol. 3, pp. 1869–1879.
- [4] E. M. Rogers, *Diffusion of Innovations*, 5th ed. New York, NY, USA: Free Press 2003.
- [5] E. Adar and L. A. Adamic, “Tracking information epidemics in blogspace,” in *Proc. IEEE/WIC/ACM Int. Conf. Web Intell.*, 2005, pp. 207–214.
- [6] D. Easley and J. Kleinberg, *Networks, Crowds and Markets: Reasoning about a Highly Connected World*. Cambridge, U.K.: Cambridge Univ. Press, 2010.
- [7] P. V. Mieghem, J. Omic, and R. Kooij, “Virus spread in networks,” *IEEE/ACM Trans. Netw.*, vol. 17, no. 1, pp. 1–14, Feb. 2009.
- [8] F. Eggenberger and G. Polya, “Über die statistik verketteter vorgänge,” *Z. Angew. Math. Mech.*, vol. 3, no. 4, pp. 279–289, 1923.
- [9] G. Polya and F. Eggenberger, “Sur l’interprétation de certaines courbes de fréquences,” *Comptes Rendus C. R.*, vol. 187, pp. 870–872, 1928.
- [10] G. Polya, “Sur quelques points de la théorie des probabilités,” *Annales de l’institut Henri Poincaré*, vol. 1, no. 2, pp. 117–161, 1930.
- [11] A. Rosenblatt, “Sur le concept de contagion de M. G. Polya dans le calcul des probabilités,” *Acad. Nac. de Cien. Exactas, Fis. Nat.*, vol. 3, pp. 186–204, 1940.
- [12] C. F. Martin, L. J. S. Allen, and M. S. Stamp, “An analysis of the transmission of chlamydia in a closed population,” *J. Difference Equations Appl.*, vol. 2, pp. 1–29, 1996.
- [13] R. Pemantle, “A survey of random processes with reinforcement,” *Probab. Surv.*, vol. 4, pp. 1–79, 2007.

- [14] A. Banerjee, P. Burlina, and F. Alajaji, "Image segmentation and labeling using the Polya urn model," *IEEE Trans. Image Process.*, vol. 8, no. 9, pp. 1243–1253, Sep. 1999.
- [15] R. Ash and C. Doléans-Dade, *Probability and Measure Theory*. New York, NY, USA: Academic, 2000.
- [16] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*, 3rd ed. London, U.K.: Oxford Univ. Press, 2001.
- [17] W. Feller, *An Introduction to Probability Theory and its Applications*, vol. 2, 2nd ed. New York, NY, USA: Wiley, 1971.
- [18] F. Alajaji and T. Fuja, "A communication channel modeled on contagion," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 2035–2041, Nov. 1994.
- [19] R. Pemantle, "A time-dependent version of Polya's urn," *J. Theor. Probab.*, vol. 3, no. 4, pp. 627–637, 1990.
- [20] M. Chen and M. Kuba, "On generalized Polya urn models," *J. Appl. Probab.*, vol. 50, no. 4, pp. 1169–1186, 2013.
- [21] N. L. Johnson and S. Kotz, *Urn Models and Their Application*. New York, USA: Wiley, 1977.
- [22] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 2006.
- [23] R. Albert and A. L. Barabási, "Statistical mechanics of complex networks," *Rev. Modern Phys.*, vol. 74, no. 1, pp. 47–97, 2002.
- [24] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, "Epidemic spreading in real networks: An eigenvalue viewpoint," in *Proc. 22nd Int. Symp. Rel. Dist. Syst.*, 2003, pp. 25–34.
- [25] H. J. Ahn and B. Hassibi, "Global dynamics of epidemic spread over complex networks," in *Proc. Conf. Decis. Control.*, 2013, pp. 4579–4585.



**Mikhail Hayhoe** received the B.Sc. degree in mathematics and engineering and the M.A.Sc. degree in applied mathematics in 2015 and 2017, respectively, both from Queens University, Kingston, ON, Canada. He is currently working toward the Ph.D. degree in electrical and systems engineering with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA, USA.

He was a finalist for the Student Best Paper Award at ACC 2017. His research interests include systems and control, optimization, social networks, information theory, big data, and machine learning.



**Fady Alajaji** (S'90–M'94–SM'00) received the B.E. degree with distinction from the American University of Beirut, Beirut, Lebanon, and the M.Sc. and Ph.D. degrees from the University of Maryland, College Park, MD, USA, in 1988, 1990, and 1994, respectively, all in electrical engineering.

He then held a Postdoctoral appointment in 1994 at the Institute for Systems Research, University of Maryland. In 1995, he joined the Department of Mathematics and Statistics, Queen's University, Kingston, ON, Canada, where he is currently a Professor in mathematics and engineering. Since 1997, he has also held a cross appointment with the Department of Electrical and Computer Engineering, Queen's University. In 2013–2014, he worked as acting head with the Department of Mathematics and Statistics, and from 2003 to 2008, he worked as the Chair of the Queen's Mathematics and Engineering Program. His research interests include information theory, joint source-channel coding, error control coding, data compression, and digital communications.

Dr. Alajaji served as Area Editor for Source-Channel Coding and Signal Processing from 2008 to 2015 and as an Editor for Source and Source/Channel Coding from 2003 to 2012 for the IEEE TRANSACTIONS ON COMMUNICATIONS. He served as Organizer and Technical Program Committee member for several international conferences and workshops. He received the Premier's Research Excellence Award from the Province of Ontario in recognition for his research in "the theory and practice of joint source-channel coding in telecommunication systems."



**Bahman Ghahesifard** (M'13) received the B.Sc. degree in mechanical engineering, and the M.Sc. degree in control and dynamics, in 2002 and 2005, respectively, both from Shiraz University, Shiraz, Iran, and the Ph.D. degree in mathematics from Queen's University, Kingston, ON, Canada, in 2009.

He was a Postdoctoral with the Department of Mechanical and Aerospace Engineering, University of California, San Diego, San Diego, CA, USA, 2009–2012 and with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Champaign, IL, USA, from 2012 to 2013. He held a visiting faculty position with the Institute for Systems Theory and Automatic Control, University of Stuttgart, Stuttgart, Germany, in summer 2016. He is currently an Assistant Professor with the Department of Mathematics and Statistics, Queen's University. His research interests include systems and controls, distributed control and optimization, social and economic networks, game theory, geometric control and mechanics, and Riemannian geometry.