Diffusion in Social Networks: Effects of Monophilic Contagion, Friendship Paradox and Reactive Networks

Buddhika Nettasinghe, *Student Member, IEEE,* Vikram Krishnamurthy, *Fellow, IEEE,* and Kristina Lerman

Abstract—We consider SIS contagion processes over networks where, a classical assumption is that individuals' decisions to adopt a contagion are based on their immediate neighbors. However, recent literature shows that some attributes are more correlated between two-hop neighbors, a concept referred to as *monophily*. This motivates us to explore monophilic contagion, the case where a contagion (e.g. a product, disease) is adopted by considering two-hop neighbors instead of immediate neighbors (e.g. you ask your friend about the new iPhone and she recommends you the opinion of one of her friends). We show that the phenomenon called *friendship paradox* makes it easier for the monophilic contagion to spread widely. We also consider the case where the underlying network stochastically evolves in response to the state of the contagion (e.g. depending on the severity of a flu virus, people restrict their interactions with others to avoid getting infected) and show that the dynamics of such a process can be approximated by a differential equation whose trajectory satisfies an algebraic constraint restricting it to a manifold. Our results shed light on how graph theoretic consequences affect contagions and, provide simple deterministic models to approximate the collective dynamics of contagions over stochastic graph processes.

Index Terms—Friendship Paradox, Diffusion, SIS Model, Social Networks, Monophily, Reactive Network, Random Graphs.

---

1 INTRODUCTION

THE Susceptible-Infected-Susceptible (SIS) model [1] is one of the widely used models in the study of diffusion processes e.g. spread of contagious diseases, information etc. This paper focuses on a discrete time version of the SIS model on a network that involves two steps (reviewed in detail in Sec. 2) at each time instant:

Step 1 - An individual (called an agent) m is sampled uniformly from the population.

Step 2 - Agent m (from step 1) observes her neighbors (defined by the underlying network) and decides to choose one of the two possible states: infected or susceptible.

In this context, the aim of this paper is to answer the following questions:

Aim 1: How does sampling agent m from a non-uniform distribution change dynamics of the SIS model?

Aim 2: Assume that the diffusion process is based on *monophilic contagion* where agent m observes two-hop neighbors (friends of friends) in step 2. How does this change the dynamics of the SIS model?

Aim 3: Assume that the underlying network is a *reactive network* i.e. the network evolves as a Markov chain whose transition probability matrix is parameterized by the diffusion state. How can one obtain a tractable model of the collective dynamics of the diffusion and graph process?

- *B. Nettasinghe and V. Krishnamurthy are with Cornell Tech, 2 West Loop Rd, New York, NY 10044, USA and, School of Electrical and Computer Engineering, Cornell University, NY, USA.*
  E-mail: {dwn26, vikramk}@cornell.edu
- *K. Lerman is with Information Sciences Institute, University of Southern California, 4676 Admiralty Way, Marina del Rey, CA 90292, USA.*
  E-mail: lerman@isi.edu.

Next, we discuss the motivation behind the above three aims and how they relate to the literature in detail.

1.1 Motivation

The impact of network structure on diffusion processes has been studied extensively in literature. Notable works include [1], [2], [3], [4], [5], [6]. Three key assumptions made in most of these works are:

i. Each node is equally likely to update her state at each time instant

ii. The decision of an individual to adopt a contagion or not is affected only by the immediate neighbors of a network fully defined by its degree distribution

iii. The underlying social network remains same throughout the diffusion process.

Motivation for Aim 1: Contradictory to the assumption i, the probability that an individual updates her state (e.g. opinion) depends on her social connectivity (i.e. degree in the social network) according to recent findings [7]. Motivated by such findings, our Aim 1 (in Sec. 3) studies the modified SIS model where the state of a random friend (sampled proportional to degree) evolves at each time instant. This modification to the standard SIS model reflects the fact that high degree nodes evolve more often in real-world social networks as found in [7]. The main result of Sec. 3 shows that this modification results in different dynamics (compared to the standard SIS model).

Motivation for Aim 2: It has been shown in several recent works that the individuals' attributes and decisions in real world social networks are affected by and/or correlated with two-hop neighbors (i.e. friends of friends). For example, [8] recently showed that certain attributes of individuals

might be more similar to friends of friends (referred to as "the company you are kept in") than to the attributes of their friends (referred to as "the company you keep"). This phenomenon is referred to as *monophily*. This should be contrasted to *homophily* [9] where attributes of individuals are similar to their friends. Further, [10] also discusses how social influence extends beyond immediate neighbors and considers a model where strategic agents (on a social network) are affected by both immediate as well as two-hop neighbors. Motivated by the works in [8], [10], our second aim (in Sec. 4) explores the case where the diffusion process is based on *monophilic contagion* i.e. agents take their friends of friends (two-hop neighbors) into account (in contrast to assumption ii) to decide whether to adopt the contagion[1]. Further, Sec. 4 also explores how assortativity (which is a property that is not captured by the degree distribution of the network) and the graph theoretic phenomenon called *friendship paradox* affects the diffusion process.

Motivation for Aim 3: We note that most real world networks are of random nature and evolve rapidly during the diffusion process. More generally, the underlying network may evolve in a manner that depends on the state of the diffusion process as well e.g. depending on the state of a spreading disease (fraction of infected individuals for example), people might restrict their interactions with others and thus, changing the structure of the underlying contact network [11]. Hence, modeling a network as a deterministic graph does not capture this (diffusion state dependent) evolution of real world networks. Motivated by this, in our third aim (in Sec. 5), the underlying network is modeled as a Markovian graph process whose transition probabilities at each time instant depend on the state of the diffusion process.

1.2 Related Work

Apart from the literature discussed in Sec. 1.1 that serve as the main motivation for our work, several other works also pursued related research directions on friendship paradox, diffusion processes and evolving networks.

Related to the second aim of this paper, [12] studied how different methods of perceiving the number of friends of friends (e.g. mean, median) affect the time taken to reach consensus in networks (with different assortativities) as a result of the friendship paradox. Authors experimentally showed that these different perception models result in different times for the individuals in the social network to reach consensus. This is related to our second aim (monophilic contagion in Sec. 4) in that our aim is also to compare diffusions driven by two different perception methods: observing friends vs observing friends of friends. [13] experimentally (using Twitter and mobile phone data) discovered that friendship paradox is less pronounced among most frequently contacted friends and, used a modified Susceptible-Infected model to numerically illustrate how this will reduce time taken to reach consensus. Further,

[14] also experimentally explored the connection between diffusion processes and friendship paradox and proposed a viral marketing approach which seeds random neighbors of random nodes instead of random nodes. The main contrast between these works and our second research aim is that our work focuses on the collective effects of two-hop (monophilic) contagion, assortativity and friendship paradox on diffusion processes whereas previous work focused on aspects related one-hop friends. The effects of degree correlations on diffusion processes have been explored in several works. Excellent reviews can be found in [15], [16]. Further, the connection between degree correlations and effects of friendship paradox has been studied in [12], [17], [18]. The main difference between these previous works and our results on assortativity (Sec. 4.2) relies on the fact that we restrict our attention to the particular case of monophilic contagion which has not been explored literature previously. Further, we also aim to illustrate how the effect of friendship paradox on monophilic contagion is amplified or reduced by the assortativity.

Related to the third aim this paper, several works proposed and analyzed evolving graph models: [19] studied susceptible-infected- susceptible (SIS) model on an adaptive network where susceptible individuals are allowed to rewire their connections with infected individuals and show that rewiring leads to different properties compared to classical SIS model, [20], [21] studied variants of adaptive susceptible-infected-susceptible (ASIS) model where susceptible individuals are allowed to temporarily cut edges connecting them to infected nodes in order to prevent the spread of the infection, [22] analyzed the stability of epidemic processes over time-varying networks and provides sufficient conditions for convergence, [23] studied a SIS process over a static contact network where the nodes have partial information about the epidemic state and react by limiting their interactions with their neighbors when they believe the epidemic is currently prevalent, [24] considered a diffusion process on a network whose evolution functionally depends on the set of initial seeds of the diffusion process. In contrast to these prior works, the model we discuss in Sec. 5 is more general in the sense that the underlying network is treated as a Markov process with a (functional) dependency on the state of the contagion. Further, our aim is to model monophilic contagion on a reactive network in a manner that is amenable to data-driven statistical inference methods.

1.3 Main Results and Organization

Main results of this paper are as follows:

1) Random friends (i.e. nodes sampled with probability proportional to degree) being sampled in step 1 of the SIS model results in different dynamics compared to the classical SIS model. However, critical thresholds on the parameters of the SIS model (which determine if the contagion will eventually die away or not) are not affected by this change.

2) The critical threshold corresponding to monophilic contagion is smaller than the critical threshold of non-monophilic contagion as a result of friendship paradox. This makes monophilic contagion easier to spread

---

[1]It should be noted that the concept of monophily presented in [8] does not give a causal interpretation but only the correlation between attributes of two-hop neighbors. What we consider is *monophilic contagion* (motivated by monophily): the contagion caused by the influence of two hop neighbors.

compared to non-monophilic contagion. Further, network disassortativity (negative degree-degree correlation) amplifies the effect of friendship paradox on monophilic contagion.

3) The collective dynamics of an SIS process (based on monophilic contagion) on a reactive network can be modeled by an ordinary differential equation (ODE) with an algebraic constraint. This provides a simple deterministic model of the stochastic dynamics of a complex system (an SIS process on a random graph, both evolving on the same time scale).

Organization: Section 2 reviews the mean-field dynamics model of an SIS process and the concept of friendship paradox. Sec. 3 explores the effect of non-uniformly sampled m in step 1 of SIS-model and states the first main result. Sec. 4 derives the critical thresholds for monophilic contagion, states the second main result and discusses some of its generalizations. Finally, Sec. 5 explores the collective dynamics of the SIS process on a reactive network and states the third main result.

2 PRELIMINARIES: APPROXIMATION OF SIS MODEL AND FRIENDSHIP PARADOX

In this section, the basic SIS diffusion process and how it can be modeled using deterministic mean-field dynamics is reviewed briefly. Then, the friendship paradox, some of its recent generalizations and applications are discussed. These results related to mean-field dynamics and friendship paradox are used in subsequent sections to obtain the main results.

2.1 Discrete time SIS Model

Consider a social network represented by an undirected graph $G = (V, E)$ where $V = \{1, 2, \ldots, M\}$ denotes the set of nodes. At each discrete time instant n, a node $v \in V$ of the network can take the state $s_n^{(v)} \in \{0, 1\}$ where, 0 denotes the susceptible state and 1 denotes the infected state. The degree $d(v) \in \{1, \ldots, D\}$ of a node $v \in V$ is the number of nodes connected to v and, $M(k)$ denotes the total number of nodes with degree k. Then, the degree distribution $P(k) = \frac{M(k)}{M}$ is the probability that a randomly selected node has degree k. Further, we also define the population state $\bar{x}_n(k)$ as the fraction of nodes with degree k that are infected (state 1) at time n i.e.

$$\bar{x}_n(k) = \frac{1}{M(k)} \sum_{v \in V} \mathbb{1}_{\{d(v) = k, \, s_n^{(v)} = 1\}}, \quad k = 1, \ldots, D. \quad (1)$$

For this setting, we adopt the SIS model used in [1], [25] which is as follows briefly.

Discrete Time SIS Model: At each discrete time instant n,

Step 1: A node $m \in V$ is chosen with uniform probability $p^X(m) = 1/M$ where, M is the number of nodes in the graph.

Step 2: The state $s_n^{(m)} \in \{0, 1\}$ of the sampled node m (in Step 1) evolves to $s_{n+1}^{(m)} \in \{0, 1\}$ with transition probabilities that depend on the degree of m, number of infected neighbors of m, population state of the network $\bar{x}_n$

Note that the above model is a Markov chain with a state space consisting of 2^M states (since each of the M nodes can be either infected or susceptible at any time instant). Due to this exponentially large state space, the discrete time SIS model is not mathematically tractable. However, we are interested only in the fraction of the infected nodes (as opposed to the exact state out of the 2^M states) and therefore, it is sufficient to focus on the dynamics of the population state $\bar{x}_n$ defined in (1) instead of the exact state of the infection.

2.2 Mean-Field Dynamics Model for SIS Diffusion Processes

Mean-Field dynamics refers to a simplified model of a (stochastic) system where the stochastic dynamics are replaced by deterministic dynamics. Most of the research related to mean-field dynamics are based on the seminal work of Kurtz [26] on population dynamics models. In this section, we first discuss how mean-field dynamics can be used as a deterministic model of a SIS diffusion process over an undirected network. Since an SIS diffusion over a social network is a Markov process whose state space grows exponentially with the number of nodes, mean-field dynamics offers a deterministic model that is analytically tractable [1], [2], [3], [25]. The following result from [25] shows how mean-field dynamics model closely approximates the stochastic dynamics of the true population state $\bar{x}_n$.

Theorem 1 (Mean-Field Dynamics). *1) The population state defined in (1) evolves according to the following stochastic difference equation driven by martingale difference process:*

$$\bar{x}_{n+1}(k) = \bar{x}_n(k) + \frac{1}{M}[P_{01}(k, \bar{x}_n) - P_{10}(k, \bar{x}_n)] + \zeta_n \quad (2)$$

where,
$$P_{01}(k, \bar{x}_n) = (1 - \bar{x}_n(k)) \times$$
$$\mathbb{P}(s_{n+1}^m = 1 | s_n^m = 0, d(m) = k, \bar{x}_n) \quad (3)$$

$$P_{10}(k, \bar{x}_n) = \bar{x}_n(k) \mathbb{P}(s_{n+1}^m = 0 | s_n^m = 1, d(m) = k, \bar{x}_n). \quad (4)$$

are the scaled transition probabilities of the states and, ζ_n is a martingale difference process with $||\zeta_n||_2 \leq \frac{\Gamma}{M}$ for some positive constant Γ.

2) Consider the mean-field dynamics process associated with the population state:

$$x_{n+1}(k) = x_n(k) + \frac{1}{M}(P_{01}(k, x_n) - P_{10}(k, x_n)) \quad (5)$$

where, $P_{01}(k, x_n)$ and $P_{10}(k, x_n)$ are as defined in (3), (4) and $x_0 = \bar{x}_0$. Then, for a time horizon of T points, the deviation between the mean-field dynamics (5) and the actual population state $\bar{x}_n$ of the SIS model satisfies

$$\mathbb{P}\{\max_{0 \leq n \leq T} ||x_n - \bar{x}_n||_\infty \geq \epsilon\} \leq C_1 \exp(-C_2 \epsilon^2 M) \quad (6)$$

for some positive constants C_1, C_2 providing $T = O(M)$.

First part of Theorem 1 is the martingale representation of a Markov chain (which is the population state $\bar{x}_n$).

Note from (2) that the dynamics of the population state $\bar{x}_n$ resemble a stochastic approximation recursion (new state is the old state plus a noisy term). Hence, the trajectory of the population state $\bar{x}_n$ should converge (weakly) to the deterministic trajectory given by the ODE corresponding to the mean-field dynamics in (5) as the size of the network M goes to infinity i.e. the step size of the stochastic approximation algorithm goes to zero (for details, see [27], [28]). Second part of the theorem provides an exponential bound on the deviation of the mean-field dynamics approximation from the actual population state for a finite length of the sample path. In the subsequent sections of this paper, the mean-field approximation (5) is utilized to study the effects of various sampling methods and friendship paradox on the SIS model of information diffusion.

2.3 Friendship Paradox

Friendship paradox refers to a graph theoretic consequence that was introduced in 1991 by Scott. L. Feld in [29]. We briefly review of the main results related to friendship paradox in this subsection. Feld's original statement of the friendship paradox is "on average, the number of friends of a random friend is always greater than or equal to the number of friends of a random individual". Here, a random friend refers to a random end node Y of a uniformly chosen edge (a pair of friends). This statement is formally stated in Theorem 2. Further, Theorem 3 (based on [30]) states that a similar result holds when the degrees of a random node X and random friend Z of a random node X are compared as well.

Theorem 2 (Friendship Paradox - Version 1 [29]). *Let $G = (V, E)$ be an undirected graph, X be a node chosen uniformly from V and, Y be a uniformly chosen node from a uniformly chosen edge $e \in E$. Then,*

$$\mathbb{E}\{d(Y)\} \geq \mathbb{E}\{d(X)\}, \qquad (7)$$

where, $d(X)$ denotes the degree of X.

Theorem 3 (Friendship Paradox - Version 2 [30]). *Let $G = (V, E)$ be an undirected graph, X be a node chosen uniformly from V and, Z be a uniformly chosen neighbor of a uniformly chosen node from V. Then,*

$$d(Z) \geq_{fosd} d(X) \qquad (8)$$

where, $\geq_{fosd}$ denotes the first order stochastic dominance[2].

The intuition behind Theorem 2 and Theorem 3 stems from the fact that individuals with a large number of friends (high degree nodes) appear as the friends of a large number of individuals. Hence, these high degree nodes can contribute to an increase in the average number of friends of friends. On the other hand, individuals with smaller number of friends appear as friends of a smaller number of individuals. Hence, they do not cause a significant change in the average number of friends of friends.

[2]A discrete random variable X (with a cumulative distribution function F_X) first order stochastically dominates a discrete random variable Y (with a cumulative distribution function F_Y), denoted $X \geq_{fosd} Y$ if, $F_X(n) \leq F_Y(n)$, for all n. Further, first order stochastic dominance implies larger mean.

Friendship paradox, which in essence is a sampling bias observed in undirected social networks has gained attention as a useful tool for estimation and detection problems in social networks. For example, [31] proposes to utilize friendship paradox as a sampling method for reduced variance estimation of a heavy-tailed degree distribution, [32], [33], [34] explore how the friendship paradox can be used for detecting a contagious outbreak quickly, [14], [35], [36], [37], [38] utilizes friendship paradox for maximizing influence in a social network, [39] proposes friendship paradox based algorithms for efficiently polling a social network (e.g. to forecast an election) in a social network, [40] studies how the friendship paradox in a game theoretic setting can systematically bias the individual perceptions. Further, [13], [17], [18], [41], [42], [43], [44], [45], [46] present and analyze further generalizations of the classical friendship paradox.

3 EFFECT OF THE SAMPLING DISTRIBUTION IN THE STEP 1 OF THE SIS MODEL

Recall from Sec. 2.3 that we distinguished between three sampling methods for a network $G = (V, E)$: a random node X, a random friend Y and, a random friend Z of a random node. Further, recall that in the discrete-time SIS model explained in Sec. 2.1, the node m whose state evolves is sampled uniformly from V i.e. $m \stackrel{d}{=} X$. This section studies the effect of random friends (Y or Z) evolving at each time instant instead of random nodes (X) i.e. the cases where $m \stackrel{d}{=} Y$ and $m \stackrel{d}{=} Z$. Following is our main result in this section:

Theorem 4. *Consider the discrete time SIS model on a graph $G = (V, E)$ presented in Sec. 2.1.*

1) *If the node m is a uniformly chosen node Y from a uniformly chosen edge $e \in E$ (i.e. node m with degree $d(m)$ is chosen with probability $p^Y(m) = \frac{d(m)}{\sum_{v \in V} d(v)}$), then the stochastic dynamics of the SIS model can be approximated by,*

$$x_{n+1}(k) = x_n(k) + \frac{1}{M}\frac{k}{\bar{k}}\big(P_{01}(k, x_n) - P_{10}(k, x_n)\big), \quad (9)$$

where $\bar{k}$ is the average degree of the graph $G = (V, E)$.

2) *If the node m is a random neighbor Z of a random node X, then the stochastic dynamics of the SIS model can be approximated by,*

$$x_{n+1}(k) = x_n(k) + \frac{1}{M}\left(\sum_{k'} \frac{P(k)}{P(k')}P(k|k')\right) \times$$
$$\big(P_{01}(k, x_n) - P_{10}(k, x_n)\big), \quad (10)$$

where $\bar{k}$ is the average degree of the graph $G = (V, E)$, P is the degree distribution and $P(k|k')$ is the probability that a random neighbor of a degree k' node is of degree k. Further, if the network is a degree-uncorrelated network i.e. $P(k|k')$ does not depend on k', then (10) will be the same as (9).

Proof. Note that the population state $\{\bar{x}_n\}_{n \geq 0}$ is a Markov chain with a state space of the size $\Pi_{d=1}^{M}(M(d) + 1)$. Let P^{pop} denote the transition probability matrix of this Markov chain and e_i denote the $\Pi_{d=1}^{M}(M(d) + 1)$ dimensional column vector with 1 in the ith position and zeros in all

other positions. Then, the Martingale representation of this Markov chain is,

$$\rho_{n+1} = (P^{pop})'\rho_n + \eta_n \tag{11}$$

where, ρ_n are states taking values in the space $\{e_1, \ldots, e_{\Pi_{d=1}^M (M(d)+1)}\}$, η_n is martingale difference noise. Then, by multiplying with the state level matrix, we get

$$\bar{x}_{n+1}(k) = \mathbb{E}\{\bar{x}_n(k)|\bar{x}_n\} + \gamma_n \tag{12}$$

where, γ_n is the product of martingale difference noise η_n and state level matrix. Then,

$$\bar{x}_{n+1}(k) = \mathbb{E}\{\bar{x}_n(k)|\bar{x}_n\} + \gamma_n \tag{13}$$

$$= \mathbb{P}(s_{n+1}^{(m)} = 1, s_n^{(m)} = 0, d(m) = k|\bar{x}_n) \times (\bar{x}_n(k) + \frac{1}{M(k)}) +$$

$$\mathbb{P}(s_{n+1}^{(m)} = 0, s_n^{(m)} = 1, d(m) = k|\bar{x}_n) \times (\bar{x}_n(k) - \frac{1}{M(k)}) +$$

$$(1 - \mathbb{P}(s_{n+1}^{(m)} = 1, s_n^{(m)} = 0, d(m) = k|\bar{x}_n) -$$

$$\mathbb{P}(s_{n+1}^{(m)} = 0, s_n^{(m)} = 1, d(m) = k|\bar{x}_n))(\bar{x}_n(k)) + \gamma_n \tag{14}$$

Let,

$$A = \mathbb{P}(s_{n+1}^{(m)} = 1, s_n^{(m)} = 0, d(m) = k|\bar{x}_n)$$
$$B = \mathbb{P}(s_{n+1}^{(m)} = 0, s_n^{(m)} = 1, d(m) = k|\bar{x}_n).$$

Then, we get

$$\bar{x}_{n+1}(k) = \bar{x}_n(k) + \frac{1}{M(k)}(A - B) + \gamma_n. \tag{15}$$

Then, the first and second parts of the Theorem 4 follow by decomposing the joint distributions of A, B with respective to the degree distributions of a random friend Y and a random friend Z of a random node X respectively.
□

Theorem 4 shows that, if the node m sampled in the step 1 of the SIS model (explained in Sec. 2.1), is chosen to be a random friend or a random friend of a random node, then different elements $x_n(k)$ of the mean-field approximation evolves at rates that are different to the classical case where $m \stackrel{d}{=} X$. This result allows us to model the dynamics of the population state in the more involved case where, frequency of the evolution of an individual is proportional his/her degree (part 1 - e.g. high degree nodes change opinions more frequently due to higher exposure) and also depends on the degree correlation (part 2 - e.g. nodes being connected to other similar/different degree nodes changes the frequency of changing the opinion).

Remark 1 (Invariance of the critical thresholds to the sampling distribution in step 1). The stationary condition for the mean-field dynamics is obtained by setting $x_{n+1}(k) - x_n(k) = 0$ for all $k \geq 1$. Comparing (5) with (9) and (10), it can be seen that this condition yields the same expression $P_{01}(k, x_n) - P_{10}(k, x_n) = 0$, for all three sampling methods (random node - X, random end of a random link Y and, a random neighbor Z of a random node). Hence, the critical thresholds of the SIS model are invariant to the distribution from which the node m is sampled in step 1. This leads us to Sec. 4 where, modifications to the step 2 of the SIS model are analyzed in terms of the critical thresholds.

4 CRITICAL THRESHOLDS FOR UNBIASED-DEGREE NETWORKS

In Sec. 3 of this paper, we focused on the step 1 of the SIS model and, showed that different sampling methods for selecting the node m result in different mean-field dynamics with the same stationary conditions. In contrast, the focus of this section is on the step 2 of the SIS model and, how changes to this step would result in different stationary conditions and critical thresholds.

4.1 Critical Thresholds for Monophilic and Non-Monophilic Contagions

Recall the SIS model reviewed in Sec. 2.1 again. We limit our attention to the case of *unbiased-degree* networks and viral adoption rules discussed in [47].

Unbiased-degree network: In an unbiased-degree network, neighbors of agent m sampled in the step 1 of the SIS model are $d(m)$ (degree of agent m) number of uniformly sampled agents (similar in distribution to the random variable X) from the network. Therefore, in an unbiased-degree network, any agent is equally likely to be a neighbor of the sampled (in the step 1 of the SIS model) agent m.

Viral adoption rules[3]: If the sampled agent m (in the step 1 of the SIS model) is an infected agent, she becomes susceptible with a constant probability δ. If the sampled agent m (in the step 1 of the SIS model) is a susceptible (state 0) agent, she samples $d(m)$ (degree of m) number of other agents $X_1, X_2, \ldots, X_{d(m)}$ (neighbors of m in the unbiased-degree network) from the network and, adopts the contagion based on one of the following rules:

Case 1 - Non-monophilic adoption rule: For each sampled neighbor X_i, m observes the state of X_i. Hence, agent m observes the states of $d(m)$ number of random nodes. Let a_m^X denote the number of infected agents among $X_1, \ldots, X_{d(m)}$. Then, the susceptible agent m becomes infected with probability $\nu\frac{a_m^X}{D}$ where, $0 \leq \nu \leq 1$ is a constant and D is the largest degree of the network.

Case 2 - Monophilic adoption rule: For each sampled neighbor X_i, m observes the state of a random friend $Z_i \in \mathcal{N}(X_i)$ of that neighbor. Hence, agent m observes the states of $d(m)$ number of random friends $Z_1, \ldots, Z_{d(m)}$ of random nodes $X_1, \ldots, X_{d(m)}$. Let a_m^Z be the number of infected agents among $Z_1, \ldots, Z_{d(m)}$. Then, the susceptible agent m becomes infected with probability $\nu\frac{a_m^Z}{D}$ where, $0 \leq \nu \leq 1$ is a constant and D is the largest degree of the network.

In order to compare the effects of non-monophilic and monophilic adoption rules, we look at the conditions on the model parameters for which, each rule leads to a positive fraction of infected nodes starting from a small fraction of infected nodes i.e. a positive stationary solution to the mean-field dynamics (5). Our main result is the following:

---

[3]The two rules (monophilic and non-monophilic) are called viral adoption rules as they consider the total number of infected nodes (denoted by a_m^X and a_m^Z in case 1 and case 2 respectively) in the sample in contrast to the persuasive adoption rules that consider the fraction of infected nodes in the sample [3].

Theorem 5. *Consider the SIS model described in Sec. 2.1. Define the effective spreading rate as $\lambda = \frac{\nu}{\delta}$ and let X be a random node and Z be a random friend of X.*

1) *Under the non-monophilic adoption rule (Case 1), the mean-field dynamics equation (5) takes the form,*

$$x_{n+1}(k) = x_n(k) + \frac{1}{M}\left((1 - x_n(k))\frac{\nu k \theta_n^X}{D} - x_n(k)\delta\right) \tag{16}$$

where,

$$\theta_n^X = \sum_k P(k)x_n(k) \tag{17}$$

is the probability that a randomly chosen node X at time n is infected. Further, there exists a positive stationary solution to the mean field dynamics (16) for case 1 if and only if

$$\lambda > \frac{D}{\mathbb{E}\{d(X)\}} = \lambda_X^* \tag{18}$$

2) *Under the monophilic adoption rule (Case 2), the mean-field dynamics equation (5) takes the form,*

$$x_{n+1}(k) = x_n(k) + \frac{1}{M}\left((1 - x_n(k))\frac{\nu k \theta_n^Z}{D} - x_n(k)\delta\right) \tag{19}$$

where,

$$\theta_n^Z = \sum_k \left(\sum_{k'} P(k')P(k|k')\right)x_n(k) \tag{20}$$

is the probability that a randomly chosen friend Z of a randomly chosen node X at time n is infected[4]. Further, there exists a positive stationary solution to the mean field dynamics (19) if and only if

$$\lambda > \frac{D}{\mathbb{E}\{d(Z)\}} = \lambda_Z^* \tag{21}$$

Proof. **Part 1: Non-monophilic adoption rule:** The proof of the first part is inspired by [3], [47] that consider the unbiased degree networks with non-monophilic adoption rules with continuous-time evolutions (as opposed to the discrete time case considered here). The main purpose of the first part is to provide a comparison of the non-monophilic adoption rule with the monophilic adoption rule (part 2). Consider the mean-field dynamics given in 5. The probability of a susceptible agent agent m (with degree $d(m) = k$) sampled at time instant n for the unbiased degree network adopting the contagion can be derived as follows:

$$\mathbb{P}(s_{n+1}^m = 1 | s_n^m = 0, d(m) = k, \bar{x}_n) = \tag{22}$$

$$\sum_{a=1}^k \frac{\nu a}{D}\binom{k}{a}(\theta_n^X)^a(1 - \theta_n^X)^{(k-a)} = \frac{\nu k \theta_n^X}{D} \tag{23}$$

where,

$$\theta_n^X = \sum_k P(k)x_n(k) \tag{24}$$

[4]We use $P(k|k')$ to denote the conditional probability that a node with degree k' is connected to a node with degree k. More specifically $P(k|k') = \frac{e(k,k')}{q(k)}$ where $e(k, k')$ is the joint degree distribution of the network and $q(k)$ is the marginal distribution that gives the probability of random end (denoted by random variable Y in Theorem 2) of random link having degree k. We also use σ_q to denote the variance of $q(k)$ in subsequent sections.

is the probability that a randomly chosen node X is infected at time instant n. Eq. (22) is based on the following argument. The neighbors of m (in the case of non-monophilic adoption rule) are are randomly sampled nodes (X) and therefore, the number a of infected neighbors (out of k total) follows a binomial distribution with parameter θ_n^X. Since the probability of being infected when a susceptible node m has a (out of k) infected neighbors is $\frac{\nu a}{D}$, the probability of a degree k susceptible node becoming infected at time m is the expectation of $\frac{\nu a}{D}$ with respect to the binomial distribution as calculated in (22). Further, the probability that an infected node (independent of the degree) becomes susceptible is δ as assumed in the viral adoption rule i.e.

$$\mathbb{P}(s_{n+1}^m = 0 | s_n^m = 1, d(m) = k, \bar{x}_n) = \delta. \tag{25}$$

Then, substituting (25), (22) in to the mean-field dynamics 5 yields the mean-field dynamics approximation (16) for non-monophilic adoption rules.

In order to obtain the critical thresholds for non-monophilic case, consider the stationary condition of 16, $x_{n+1}(k) - x_n(k) = 0$, which yields the stationary population state $x(k)$ to follow,

$$x(k) = \frac{\frac{\nu k \theta^X}{D}}{\frac{\nu k \theta^X}{D} + \delta} = \frac{\lambda k \theta^X}{\lambda k \theta^X + D} \tag{26}$$

where,

$$\theta^X = \sum_k P(k)x(k) = \rho \tag{27}$$

is the probability that a randomly chosen node X is infected during stationary state (defined in 41). Then, by substituting (26) in (27), we get

$$\rho = \sum_k P(k)\frac{\lambda k \rho}{\lambda k \rho + D} = H_{\lambda,P}^X(\rho). \tag{28}$$

The diffusion prevails (without dying away) when (28) has a positive solution. Further, (28) is an increasing, concave function with $H_{\lambda,P}^X(0) = 0$ and $H_{\lambda,P}^X(1) < 1$.

Hence, in order for the (28) to have a positive solution,

$$\left.\frac{dH_{\lambda,P}^X(\rho)}{d\rho}\right|_{\rho=0} > 1, \tag{29}$$

which then yields,

$$\lambda > \frac{D}{\sum_k kP(k)} = \frac{D}{\mathbb{E}d(X)} \tag{30}$$

Part 2: Monophilic Adoption Rule: Consider the mean-field dynamics given in 5. By following steps similar to the part 1 of the proof, The probability of a susceptible agent agent m (with degree $d(m) = k$) sampled at time instant n for the unbiased degree network adopting the contagion is:

$$\mathbb{P}(s_{n+1}^m = 1 | s_n^m = 0, d(m) = k, \bar{x}_n) = \frac{\nu k \theta_n^Z}{D} \tag{31}$$

where,

$$\theta_n^Z = \sum_k \left(\sum_{k'} P(k')P(k|k')\right)x_n(k) \tag{32}$$

is the probability that a randomly chosen friend Z of a randomly chosen node X is infected at time instant n. To

understand how the expression for θ_n^Z is derived, recall that $P(k|k')$ is the probability that a randomly chosen friend of a node with degree k' is degree k. Then, taking expectation of $P(k|k')$ with respect to probability of sampling a node with degree k' yields $\sum_{k'} P(k')P(k|k')$ as the probability of a random friend Z (of a random node X) having a degree k.

Further, the probability that an infected node (independent of the degree) becomes susceptible is δ as assumed in the viral adoption rule i.e.

$$\mathbb{P}(s_{n+1}^m = 0 | s_n^m = 1, d(m) = k, \bar{x}_n) = \delta. \quad (33)$$

Then, substituting (33), (31) in to the mean-field dynamics 5 yields the mean-field dynamics approximation (19) for non-monophilic adoption rules.

In order to obtain the critical thresholds for non-monophilic case, consider the stationary condition of 19, $x_{n+1}(k) - x_n(k) = 0$, which yields the stationary population state $x(k)$ to follow,

$$x(k) = \frac{\frac{\nu k \theta^Z}{D}}{\frac{\nu k \theta^Z}{D} + \delta} = \frac{\lambda k \theta^Z}{\lambda k \theta^Z + D} \quad (34)$$

where,

$$\theta^Z = \sum_k \left(\sum_{k'} P(k')P(k|k') \right) x(k) \quad (35)$$

is the probability that a random friend Z of a randomly node X is infected during stationary state (defined in 41). Then, by substituting (34) in (35), we get

$$\theta^Z = \sum_k \left(\sum_{k'} P(k')P(k|k') \right) \frac{\lambda k \theta^Z}{\lambda k \theta^Z + D} = H_{\lambda,P}^Z(\theta^Z). \quad (36)$$

The diffusion prevails (without dying away) when (36) has a positive solution. Further, (36) is an increasing, concave function of θ^Z with $H_{\lambda,P}^Z(0) = 0$ and $H_{\lambda,P}^Z(1) < 1$.

Hence, in order for the (36) to have a positive solution,

$$\left. \frac{dH_{\lambda,P}^Z(\theta^Z)}{d\theta^Z} \right|_{\theta^Z = 0} > 1, \quad (37)$$

which then yields,

$$\lambda > \frac{D}{\sum_k k(\sum_{k'} P(k')P(k|k'))} = \frac{D}{\mathbb{E}d(Z)}. \quad (38)$$

□

The infection spreading under the monophilic adoption rule (Case 2 of Theorem 5) can also be thought of as representing the network by the square graph (corresponding to the square of the adjacency matrix of the original network). Proceeding that way would also yield the same critical threshold as in the Case 2 of Theorem 5. Theorem 5 allows us to analyze the effects of friendship paradox and degree-assortativity on the contagion process as discussed in the next subsection.

4.2 Effects of Friendship Paradox and Degree Correlation on the Monophilic Contagion

Theorem 5 showed that the critical thresholds of the mean-filed dynamics equation (5) for the two adoption rules (non-monophilic and monophilic contagion) are different. Following is an immediate corollary of Theorem 5 which gives the ordering of these critical thresholds using the friendship paradox stated in Theorem 2.

Corollary 6. *The critical thresholds λ_X^*, λ_Z^* in (18), (21) for the cases of non-monophilic (case 1) and monophilic (case 2) adoption rules satisfy*

$$\lambda_Z^* \leq \lambda_X^*. \quad (39)$$

Corollary 6 shows that in the case of monophilic adoption rule, it is easier (smaller effective spreading rate) for the infection to spread to a positive fraction of the agents as a result of the friendship paradox. Hence, observing random friends of random neighbors for adopting a contagion makes it easier for the contagion to spread instead of dying away (in unbiased-degree networks). This shows how friendship paradox can affect the spreading of a contagion over a network.

Remark 2. *If we interpret an individual's second-hop connections as weak-ties, then Theorem 5 and Corollary 6 can be interpreted as results showing the importance of weak-ties in contagions (in the context of a SIS model and an unbiased-degree network). See the seminal works in [48], [49] for the importance and definitions of weak-ties in the sociology context.*

The ordering $\lambda_Z^* \leq \lambda_X^*$ of the critical thresholds in Corollary 6 holds irrespective of any other network property. However, the magnitude of the difference of the critical thresholds $\lambda_X^* - \lambda_Z^*$ depends on the neighbor-degree correlation (assortativity) coefficient defined as,

$$r_{kk} = \frac{1}{\sigma_q^2} \sum_{k,k'} kk' \Big(e(k,k') - q(k)q(k') \Big) \quad (40)$$

using the notation in Footnote 4. To intuitively understand this, consider a star graph that has a negative assortativity coefficient (as all low degree nodes are connected to the only high degree node). Therefore, a randomly chosen node X from the star graph has a much smaller expected degree $\mathbb{E}\{d(X)\}$ than the expected degree $\mathbb{E}\{d(Z)\}$ of a random friend Z of the random node X compared to the case where the network has a positive assortativity coefficient.

In order to experimentally illustrate the effects of the assortativity on the critical threshold, we start with a random power-law graph (with power-law coefficient $\alpha = 2.4$) $G_{initial} = (V_{intial}, E_{intial})$ with $|V_{intial}| = 5000$ generated using the configuration model [50]. More specifically, we generate k half-edges for each node with $k \sim ck^{-2.4}$ (where c is the normalizing constant) and then, connect each half-edge to the another randomly selected half-edge avoiding self loops. This model will result in a power-law degree distribution i.e. $P(k) = ck^{-\alpha}$. Then, three networks with different assortativity coefficients r_{kk} (-0.2, 0.0 and 0.2) were generated by rewiring the edges of $G_{initial}$ according the edge rewiring procedure [51]. In this method, two random links $(v_1, v_2), (u_1, u_2)$ are chosen at each iteration and they are replaced with new edges $(v_1, u_1), (v_2, u_2)$ if it increases
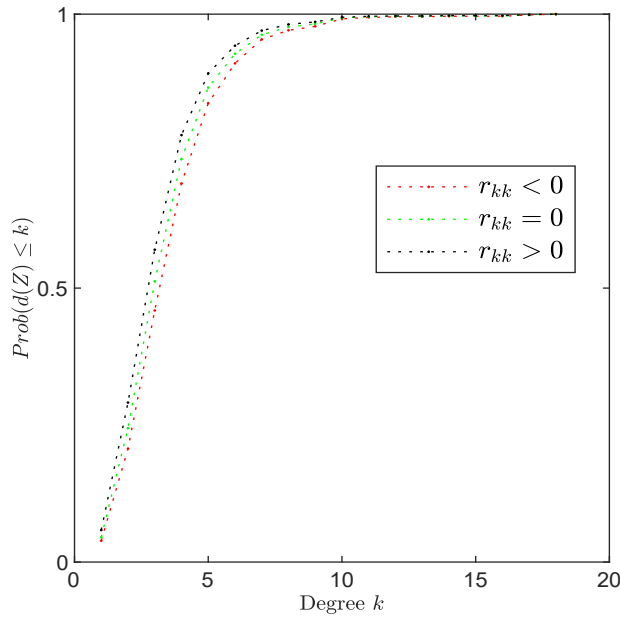
(a) CDFs of the of the degree $d(Z)$ of a random friend Z of a random node for three networks with same degree distribution but different assortativity r_{kk} values. Note that the CDFs are point-wise increasing with r_{kk} showing that $\mathbb{E}\{d(Z)\}$ decreases with r_{kk}.
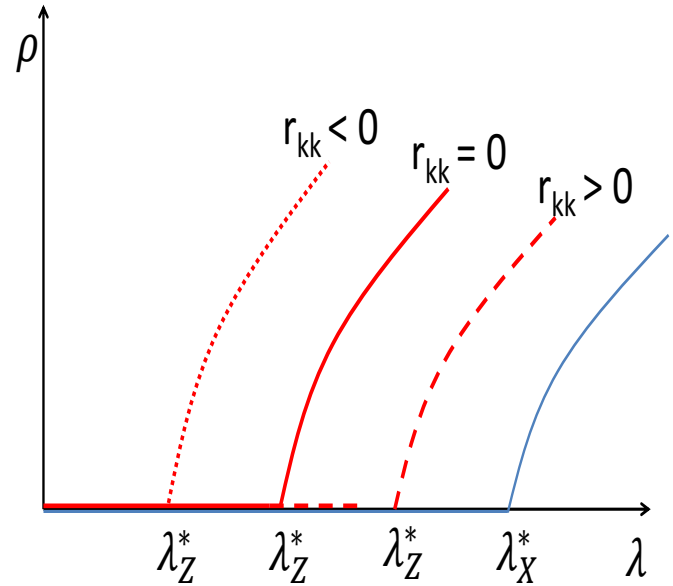
(b) Variation of the stationary fraction ρ of infected nodes with the effective spreading rate λ for the case 1 (blue) and case 2 (red), illustrating the ordering of the critical thresholds of cases 1,2 and the effect of assortativity.

Fig. 1: Comparison of non-monophilic and monophilic adoption rules and the effect of assortativity on the critical thresholds of the monophilic adoption rule.

(respectively, decreases) the value of the assortativity coefficient r_{kk}. The process is repeated until the desired value of the assortativity coefficient r_{kk} is achieved. The CDFs of the distribution of degree $d(Z)$ of a random friend Z of a random node for the three networks with the three different assortativies r_{kk} are shown in Fig. 1(a). It can be seen that the CDFs are point wise increasing with the r_{kk} showing that CDFs with smaller assortativity dominates the CDFs with larger assortativities. This stochastic dominance corresponds to larger expected degree $\mathbb{E}\{d(Z)\}$ and smaller critical threshold $\lambda_Z^* = \frac{D}{\mathbb{E}d(Z)}$ for networks with smaller assortativities.

Further, the critical threshold λ_Z^* for the monophilic contagion (case 2 of Theorem 5) is related to the critical threshold presented in [52] which is the inverse of the largest eigenvalue of the adjacency matrix in the following sense. In Case 2 of Theorem 5, the critical threshold becomes $\lambda_Z^{uc} = \frac{D\mathbb{E}\{d(X)\}}{\mathbb{E}\{d^2(X)\}}$ when the network has no assortativity ($r_{kk} = 0$). According to [53], this is inversely proportional to the largest eigenvalue of the adjacency matrix. Hence, in the special case where the network has no assortativity, the critical threshold λ_Z^* becomes proportional to the inverse of the largest eigenvalue of the adjacency matrix.

Next, we consider the stationary fraction of the infected nodes

$$\rho = \sum_k P(k)x(k) \tag{41}$$

where $P(k)$ is the degree distribution and $x(k), k = 1, \ldots, D$ are the stationary states of the mean-field dynamics in (5). Fig. 1b illustrates how the stationary fraction of the infected nodes ρ would vary with the effective spreading rate λ for non-monophilic (blue curve) and monophilic conatgion (red curves), showing the difference between the two cases and the effect of assortativity.

4.3 Modeling the existence of both monophilic and non-monophilic contagion

The main results related to the monophilic contagion (Theorem 5 and Corollary 6) can be easily modified to the setting where both friends and friends of friends affect the decisions of individuals. To achieve this, one can assume that at each time instant, the node m sampled in the first step of the SIS model (Sec 2.1) makes a decision based on friends of friends with probability $q \in [0, 1]$ and makes the decision based on immediate friends with probability $1 - q$. Then, it can be shown (by following steps similar to Theorem 5) that the critical threshold corresponding to this case is:

$$\lambda_q^* = \frac{D}{q\mathbb{E}\{d(Z)\} + (1-q)\mathbb{E}\{d(X)\}} \tag{42}$$

where D is the largest degree of the network, X is a random node and Z is a random friend of a random node. Hence, $q = 1$ corresponds to the monophilic contagion with critical threshold $\lambda_Z^* = \frac{D}{\mathbb{E}\{d(Z)\}}$ and $q = 0$ corresponds to the non-monophilic contagion with critical threshold $\lambda_X^* = \frac{D}{\mathbb{E}\{d(X)\}}$.

Then, it follows from friendship paradox (Theorem 2 and Theorem 3) that

$$\lambda_Z^* \leq \lambda_q^* \leq \lambda_X^*, \quad \forall q \in [0,1]. \qquad (43)$$

Tuning parameter q can be used to extend the monophilic contagion model to a wide array of practical settings where friends as well as friends of friends affect an individual's decision to update her state. For example, in Facebook, immediate friends of individuals directly affect their opinions/actions. However, friends of friends also affect the opinions/actions of individuals to a certain extent as their activity is often visible through activities (such as "sharing" and "liking") of immediate friends. With the continuous tuning parameter $q \in [0,1]$, such real world scenarios can also be analyzed using the monophilic contagion model. A second hypothetical example is the situation where a person seeks her friend's opinion about which product to buy (e.g. Apple or Samsung). Then, with probability q the friend might recommend the opinion of one her friends and with probability $1 - q$ provide her own opinion.

5 COLLECTIVE DYNAMICS OF SIS PROCESSES AND REACTIVE NETWORKS UNDER MONOPHILIC CONTAGION

So far in Sec. 3 and Sec. 4, the underlying social network on which the contagion spreads was treated as a deterministic graph and, the mean-field dynamics (5) was used to approximate the SIS-model. In contrast, this section explores the more general case where the underlying social network also randomly evolves at each time step n (of the SIS-model) in a manner that depends on the population state $\bar{x}_n$. Our aim is to obtain a tractable model that represents the collective dynamics of the SIS-model and the evolving graph process. As explained in Sec. 1.1 with examples, the motivation for this problem comes from the real world networks that evolves depending on the state of diffusions on them. Before stating our main result, we first define a reactive network and discuss the main assumptions.

Definition 1 (Reactive Network). *A reactive network is a Markovian graph process* $\{G_n\}_{n \geq 0}$ *with a state space* $\mathcal{G} = \{\mathcal{G}_1, \ldots, \mathcal{G}_N\}$ *consisting of N graphs and transition probabilities parameterized by the population state* $\bar{x}_n$ *i.e.* $G_{n+1} \sim P_{\bar{x}_n}(\cdot | G_n)$.

In Definition 1, the parameterization of the transition probabilities by the population state $\bar{x}_n$ represents the dependency of the graph process on the current state of the SIS information diffusion process. In other words, the population state $\bar{x}_n$ at time instant n determines the probability distribution with which the graph evolves from G_n to G_{n+1}. The name *reactive network* is derived from this functional dependency of the graph evolution on the population state.

We make the following two assumptions on the reactive network $\{G_n\}_{n \geq 0}$ (Definition 1).

Assumption 1. *Each graph* $\mathcal{G}_i \in \mathcal{G}$, $i = 1, \ldots, N$ *has the same degree distribution* $P(k)$ *but different conditional degree distributions* $P_{\mathcal{G}_1}(k | k'), \ldots, P_{\mathcal{G}_N}(k | k')$.

Assumption 2. *The transition probability matrix* $P_{\bar{x}_n}$ *of the reactive network* $\{G_n\}_{n \geq 0}$ *(Definition 1) is irreducible and aperiodic with a unique stationary distribution* $\pi_{\bar{x}_n}$ *for all values of the population state* $\bar{x}_n$.

The Assumption 1 imposes the constraint that each graph in the state space has the same degree distribution but different conditional distributions. Hence, the state space $\mathcal{G}$ can contain graphs with different higher order properties but the same degree distribution. Under this assumption, the number of nodes $M(k)$ with degree k will remain the same at each time instant n and hence, the new population state $\bar{x}_{n+1}(k)$ at time instant $n + 1$ can still be expressed as the old population state $\bar{x}_n(k)$ plus an update term as in Theorem 1. In other words, Assumption 1 allows the use of mean-field dynamics (Theorem 1) to model the dynamics of the population state of a SIS model on a reactive network. The second assumption is standard in Markov chain models and it ensures the convergence to a unique stationary distribution.

Before proceeding to state our main result in this section, we first briefly discuss the motivation for the context of the result and how it relates to the earlier sections that focused on monophilic contagion. We consider an SIS process under monophilic contagion spreading on a reactive network (Definition 1) that satisfies the Assumptions 1 and 2. The motivation for considering this setting is two fold:

1) Firstly, Sec. 4.2 showed that monophilic contagion is affected by the assortativity of the network. Considering monophilic contagion on a reactive network under Assumptions 1 and 2 allows us to model and study the more general case where assortativity of the network evolves depending on the state of the diffusion.

2) Secondly, considering a diffusion process on an evolving network is more realistic compared to models based on a deterministic network as discussed in detail in Sec. 1.1. Our main result provides a model (that can also be generalized to settings beyond monophilic contagion) which allows contagion dependent state evolution.

In this context, our main result is the following.

Theorem 7 (Collective Dynamics of SIS-model and Reactive Network). *Consider a reactive network* $\{G_n\}_{n \geq 0}$ *(Definition 1) with state space* $\mathcal{G}$ *and transition probabilities* $P_{\bar{x}_n}(\cdot | G_n)$ *(parameterized by the population state* $\bar{x}_n$*) satisfying the Assumptions 1 and 2. Let the kth element of the vector* $H(x_n, G_n)$ *be*

$$H_k(x_n, G_n) = (1 - x_n(k))\frac{\nu k \theta_n^Z}{D} - x_n(k)\delta \quad where, \quad (44)$$

$$\theta_n^Z = \sum_k \left(\sum_{k'} P(k') P_{G_n}(k | k') \right) x_n(k). \qquad (45)$$

Further, assume that $H(x, \mathcal{G}_i)$ *is Lipschitz continuous in x for all* $\mathcal{G}_i \in \mathcal{G}$. *Then, as the number of nodes M tends to infinity, the sequence of the population state vectors* $\{\bar{x}_n\}_{n \geq 0}$ *generated by the SIS model under monophilic contagion over the reactive network converges weakly to the trajectory of the deterministic differential equation*

$$\frac{dx}{dt} = \mathbb{E}_{G \sim \pi_x}\{H(x, G)\} \quad (ODE) \qquad (46)$$

$$P_x' \pi_x = \pi_x. \quad (algebraic\ constraint) \qquad (47)$$

Proof. The following result from [28] will be used to establish the weak convergence of the sequence of population states $\{\bar{x}_n\}_{n \geq 0}$ in Theorem 7.

Consider the stochastic approximation recursion,

$$\bar{x}_{n+1} = \bar{x}_n + \epsilon \mathcal{H}(\bar{x}_n, G_n), \quad n = 0, 1, \dots \quad (48)$$

where $\epsilon > 0$, $\{G_n\}$ is a $\mathcal{G}$ valued random process and, $\bar{x}_n \in \mathbb{R}^M$ is the output of recursion at time $n = 0, 1, \dots$. Further, let

$$\bar{x}^\epsilon(t) = \bar{x}_n \text{ for } t \in [n\epsilon, n\epsilon + \epsilon], \quad n = 0, 1, \dots, \quad (49)$$

which is a piecewise constant interpolation of $\{\bar{x}_n\}$. In this setting, the following result holds.

Theorem 8. *Consider the stochastic approximation algorithm (48). Assume*
SA1: $\mathcal{H}(x, G)$ *is uniformly bounded for all* $x \in \mathbb{R}^M$ *and* $G \in \mathcal{G}$.
SA2: *For any* $l \geq 0$, *there exists* $h(x)$ *such that*

$$\frac{1}{N} \sum_{n=l}^{N+l-1} \mathbb{E}_l \{\mathcal{H}(x, G_n)\} \to h(x) \text{ as } N \to \infty. \quad (50)$$

where, $\mathbb{E}_l\{\cdot\}$ denotes expectation with respect to the sigma algebra generated by $\{G_n : n < l\}$.
SA3: *The ordinary differential equation (ODE)*

$$\frac{dx(t)}{dt} = h(x(t)), \quad x(0) = \bar{x}_0 \quad (51)$$

has a unique solution for every initial condition.

Then, the interpolated estimates $\theta^\epsilon(t)$ defined in (49) satisfies

$$\lim_{\epsilon \to 0} \mathbb{P}\Big(\sup_{0 \leq t \leq T} |\bar{x}^\epsilon(t) - x(t)| \geq \eta \Big) = 0 \text{ for all } T > 0, \eta > 0 \quad (52)$$

where, (t) is the solution of the ODE (51).

Next, we will use Theorem 8 to show how the dynamics of the population state can be approximated by and ODE with an algebraic constraint in the case of a reactive network.

By Part 2 of Theorem 1, the stochastic dynamics of the state $\bar{x}_n$ can be replaced by their mean-field dynamics x_n as follows:

$$x_{n+1} = x_n + \frac{1}{M} H(x_n, G_n) \quad (53)$$

where $H(x_n, G_n)$ is as defined in Theorem 7. Note that (53) resembles (48).

SA1 condition - Each element $H_k(x, G)$ of $H(x, G)$ (for any x, G in the domain) is a difference of two values(each in the interval $[0, 1]$). Hence, SA1 condition holds.

SA2 condition - As a result of the law of large numbers of the Markovian graph process $\{G_n\}$, SA 2 holds with

$$h(x) = \mathbb{E}_{G \sim \pi_x} \{H(x, G)\} \quad (54)$$

where, π_x is the unique stationary distribution satisfying $P'_x \pi_x = \pi_x$.

SA3 condition - Lipschitz continuity of $h(x)$ is a sufficient condition for the existence of a unique solution for a nonlinear ODE. Hence, SA3 condition holds.

Therefore, the result follows from Theorem 8. $\square$

Theorem 7 states that the dynamics of the population state of the SIS diffusion (under monophilic contagion) on a reactive network can be approximated by an ODE (46) with an algebraic constraint (47). The core idea behind this result (and the proof that leads to it) can also be understood as follows in order to gain some intuition. Due to the Assumption 1, the mean-field dynamics

$$x_{n+1} = x_n + \frac{1}{M} H(x_n, G_n) \quad (55)$$

can be used to model the evolution of the population state of the SIS process over network despite the fact that it is evolving. Then, as the number of nodes M becomes large (i.e. the scaling factor $\frac{1}{M}$ goes to zero), the sequence $\{x_n\}_{n \geq 0}$ evolves on a slow time scale compared to the reactive network $\{G_n\}_{n \geq 0}$. In other words, it will be a system where $\{x_n\}_{n \geq 0}$ evolves on a slow time scale (due to the large M) and $\{G_n\}_{n \geq 0}$ evolves on a fast time scale. Stochastic averaging theory results (used in the proof) for such two time scale problems state that, the fast dynamics of the reactive network $\{G_n\}_{n \geq 0}$ can be approximated by their average on the slow time scale of the population state $\{x_n\}_{n \geq 0}$. In other words, the $\frac{1}{M} H(x_n, G_n)$ can be replaced by its $\mathbb{E}_{G \sim \pi_x} \{H(x, G)\}$ with respect to the stationary distribution π_x of the Markov chain and thus yielding the ODE (46). The algebraic constraint follows from the fact that the π_x is the eigenvector with unit eigenvalue of the parameterized transition probability matrix P_x of the reactive network.

From a statistical modeling perspective, Theorem 7 provides a useful means of approximating the complex dynamics of two inter-dependent stochastic processes (diffusion process and the stochastic graph process) by an ODE (46) whose trajectory $x(t)$ at each time instant $t > 0$ is constrained by the algebraic condition (47). Having an algebraic constraint restricts the number of possible sample paths of the population state vector $\{\bar{x}_n\}_{n \geq 0}$. More specifically, the D-dimensional vector $x(t)$ (at each time instant $t > 0$) should satisfy the equation (47). Hence, from a statistical inference/filtering perspective, this makes estimation/prediction of the population state easier. For example, the algebraic condition can be used in Bayesian filtering algorithms (such as the one proposed in [25]) to estimate the population state with a better accuracy.

6 CONCLUSION

This paper explored the SIS diffusion processes over social networks using a discrete-time model where, a randomly sampled node (at each time instant) faces the decision of updating her state (infected or susceptible) based on the states of her friends. The mean-field dynamics was adopted to model the dynamics due to the exponentially large state space of the contagion process. It was shown that distribution with which the updating node is chosen lead to different mean-field dynamics, but they induce the same critical threshold on model parameters that decides whether the contagion will spread or die out. Further, it was shown that monophilic contagion (taking a decision by observing friends of friends) make it easier (compared to standard non-monophilic contagion where the decision is taken by observing friends) for a contagion to spread

instead dying out. The reason for this contrast between monophilic and non-monophilic contagions was shown to be the friendship paradox whose effect is further amplified by the network disassortativity (negative neighbor degree correlation). Finally, the case where underlying network is a reactive network that randomly evolves depending on the state of the contagion was studied. It was shown that the complex collective dynamics of the two (functionally) dependent stochastic processes (SIS process and the random graph process) can be approximated by a deterministic ODE whose trajectory satisfies an algebraic constraint. Our main results shed light on how graph theoretic and sociological concepts such as friendship paradox and weak-ties affect diffusion processes over social networks. Further, they provide simple deterministic models for complex collective dynamics of contagions over stochastic graph processes.

REFERENCES

[1] D. López-Pintado, "Diffusion in complex social networks," *Games and Economic Behavior*, vol. 62, no. 2, pp. 573–590, 2008.

[2] M. O. Jackson and B. W. Rogers, "Relating network structure to diffusion properties through stochastic dominance," *The BE Journal of Theoretical Economics*, vol. 7, no. 1, 2007.

[3] D. López-Pintado, "Influence networks," *Games and Economic Behavior*, vol. 75, no. 2, pp. 776–787, 2012.

[4] R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in scale-free networks," *Physical review letters*, vol. 86, no. 14, p. 3200, 2001.

[5] M. O. Jackson and L. Yariv, "Diffusion of behavior and equilibrium properties in network games," *American Economic Review*, vol. 97, no. 2, pp. 92–98, 2007.

[6] D. J. Watts, "A simple model of global cascades on random networks," *Proceedings of the National Academy of Sciences*, vol. 99, no. 9, pp. 5766–5771, 2002.

[7] N. O. Hodas and K. Lerman, "How visibility and divided attention constrain social contagion," in *Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom)*. IEEE, 2012, pp. 249–257.

[8] K. M. Altenburger and J. Ugander, "Monophily in social networks introduces similarity among friends-of-friends," *Nature Human Behaviour*, vol. 2, no. 4, p. 284, 2018.

[9] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks," *Annual review of sociology*, vol. 27, no. 1, pp. 415–444, 2001.

[10] B. Fotouhi, N. Momeni, B. Allen, and M. A. Nowak, "Conjoining uncooperative societies facilitates evolution of cooperation," *Nature Human Behaviour*, vol. 2, no. 7, pp. 492–499, 2018. [Online]. Available: https://doi.org/10.1038/s41562-018-0368-6

[11] J. E. Aledort, N. Lurie, J. Wasserman, and S. A. Bozzette, "Non-pharmaceutical public health interventions for pandemic influenza: an evaluation of the evidence base," *BMC public health*, vol. 7, no. 1, p. 208, 2007.

[12] E. Lee, S. Lee, Y.-H. Eom, P. Holme, and H.-H. Jo, "Impact of perception models on friendship paradox and opinion formation," *arXiv preprint arXiv:1808.04170*, 2018.

[13] J. P. Bagrow, C. M. Danforth, and L. Mitchell, "Which friends are more popular than you?: Contact strength and the friendship paradox in social networks," in *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*. ACM, 2017, pp. 103–108.

[14] S. Lattanzi and Y. Singer, "The power of random neighbors in social networks," in *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*. ACM, 2015, pp. 77–86.

[15] M. Boguná, R. Pastor-Satorras, and A. Vespignani, "Epidemic spreading in complex networks with degree correlations," *arXiv preprint cond-mat/0301149*, 2003.

[16] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, "Epidemic processes in complex networks," *Reviews of modern physics*, vol. 87, no. 3, p. 925, 2015.

[17] Y.-H. Eom and H.-H. Jo, "Generalized friendship paradox in complex networks: The case of scientific collaboration," *Scientific Reports*, vol. 4, Apr. 2014.

[18] K. Lerman, X. Yan, and X.-Z. Wu, "The "majority illusion" in social networks," *PloS one*, vol. 11, no. 2, p. e0147617, 2016.

[19] T. Gross, C. J. D. DLima, and B. Blasius, "Epidemic dynamics on an adaptive network," *Physical review letters*, vol. 96, no. 20, p. 208701, 2006.

[20] M. Ogura and V. M. Preciado, "Epidemic processes over adaptive state-dependent networks," *Physical Review E*, vol. 93, no. 6, p. 062316, 2016.

[21] ——, "Optimal containment of epidemics in temporal and adaptive networks," in *Temporal Network Epidemiology*. Springer, 2017, pp. 241–266.

[22] P. E. Paré, C. L. Beck, and A. Nedić, "Epidemic processes over time-varying networks," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 3, pp. 1322–1334, 2018.

[23] K. Paarporn, C. Eksin, J. S. Weitz, and J. S. Shamma, "Networked SIS epidemics with awareness," *IEEE Transactions on Computational Social Systems*, vol. 4, no. 3, pp. 93–103, 2017.

[24] B. Nettasinghe and V. Krishnamurthy, "Influence maximization over Markovian graphs: A stochastic optimization approach," *IEEE Transactions on Signal and Information Processing over Networks*, 2018.

[25] V. Krishnamurthy, S. Bhatt, and T. Pedersen, "Tracking infection diffusion in social networks: Filtering algorithms and threshold bounds," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 3, no. 2, pp. 298–315, 2017.

[26] T. G. Kurtz, *Approximation of population processes*. SIAM, 1981, vol. 36.

[27] V. Krishnamurthy, *Partially Observed Markov Decision Processes*. Cambridge University Press, 2016.

[28] H. J. Kushner and G. Yin, *Stochastic approximation and recursive algorithms and applications*, ser. Applications of mathematics. New York: Springer, 2003, no. 35.

[29] S. L. Feld, "Why your friends have more friends than you do," *American Journal of Sociology*, vol. 96, no. 6, pp. 1464–1477, 1991.

[30] Y. Cao and S. M. Ross, "The friendship paradox." *Mathematical Scientist*, vol. 41, no. 1, 2016.

[31] Y.-H. Eom and H.-H. Jo, "Tail-scope: Using friends to estimate heavy tails of degree distributions in large-scale complex networks," *Scientific reports*, vol. 5, p. 09752, 2015.

[32] N. A. Christakis and J. H. Fowler, "Social network sensors for early detection of contagious outbreaks," *PloS one*, vol. 5, no. 9, p. e12948, 2010.

[33] M. Garcia-Herranz, E. Moro, M. Cebrian, N. A. Christakis, and J. H. Fowler, "Using friends as sensors to detect global-scale contagious outbreaks," *PloS one*, vol. 9, no. 4, p. e92413, 2014.

[34] L. Sun, K. W. Axhausen, D.-H. Lee, and M. Cebrian, "Efficient detection of contagious outbreaks in massive metropolitan encounter networks," *Scientific reports*, vol. 4, p. 5099, 2014.

[35] L. Seeman and Y. Singer, "Adaptive seeding in social networks," in *Foundations of Computer Science (FOCS), 2013 IEEE 54th Annual Symposium on*. IEEE, 2013, pp. 459–468.

[36] T. Horel and Y. Singer, "Scalable methods for adaptively seeding a social network," in *Proceedings of the 24th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 2015, pp. 441–451.

[37] D. A. Kim, A. R. Hwong, D. Stafford, D. A. Hughes, A. J. O'Malley, J. H. Fowler, and N. A. Christakis, "Social network targeting to maximise population behaviour change: a cluster randomised controlled trial," *The Lancet*, vol. 386, no. 9989, pp. 145–153, 2015.

[38] V. Kumar, D. Krackhardt, and S. Feld, "Network interventions based on inversity: Leveraging the friendship paradox in unknown network structures," Yale University, Tech. Rep., 2018.

[39] B. Nettasinghe and V. Krishnamurthy, ""What do your friends think?": Efficient polling methods for networks using friendship paradox," *arXiv preprint arXiv:1802.06505*, 2018.

[40] M. O. Jackson, "The friendship paradox and systematic biases in perceptions and social norms," *arXiv preprint arXiv:1605.04470*, 2016.

[41] N. O. Hodas, F. Kooti, and K. Lerman, "Friendship paradox redux: Your friends are more interesting than you," in *Seventh International AAAI Conference on Weblogs and Social Media*, 2013.

[42] D. J. Higham, "Centrality-friendship paradoxes: When our friends are more important than us," *arXiv preprint arXiv:1807.01496*, 2018.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSE.2019.2909015, IEEE Transactions on Network Science and Engineering

12

[43] J. B. Kramer, J. Cutler, and A. Radcliffe, "The multistep friendship paradox," *The American Mathematical Monthly*, vol. 123, no. 9, pp. 900–908, 2016.

[44] J. Bollen, B. Gonçalves, I. van de Leemput, and G. Ruan, "The happiness paradox: your friends are happier than you," *EPJ Data Science*, vol. 6, no. 1, p. 4, 2017.

[45] B. Fotouhi, N. Momeni, and M. G. Rabbat, "Generalized friendship paradox: An analytical approach," in *International Conference on Social Informatics*. Springer, 2014, pp. 339–352.

[46] S. Pal, F. Yu, Y. Novick, A. Swami, and A. Bar-Noy, "Quantifying the strength of the friendship paradox," in *International Workshop on Complex Networks and their Applications*. Springer, 2018, pp. 460–472.

[47] D. López-Pintado, "An overview of diffusion in complex networks," in *Complex Networks and Dynamics*. Springer, 2016, pp. 27–48.

[48] A. Rapoport, "Spread of information through a population with socio-structural bias: I. assumption of transitivity," *The bulletin of mathematical biophysics*, vol. 15, no. 4, pp. 523–533, 1953.

[49] M. S. Granovetter, "The strength of weak ties," *American Journal of Sociology*, vol. 78, no. 6, pp. 1360–1380, 1973.

[50] M. Molloy and B. Reed, "A critical point for random graphs with a given degree sequence," *Random structures & algorithms*, vol. 6, no. 2-3, pp. 161–180, 1995.

[51] M. E. Newman, "Assortative mixing in networks," *Physical review letters*, vol. 89, no. 20, p. 208701, 2002.

[52] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, "Epidemic spreading in real networks: An eigenvalue viewpoint," in *22nd International Symposium on Reliable Distributed Systems, 2003. Proceedings*. IEEE, 2003, pp. 25–34.

[53] J. G. Restrepo, E. Ott, and B. R. Hunt, "Approximating the largest eigenvalue of network adjacency matrices," *Physical Review E*, vol. 76, no. 5, p. 056119, 2007.

PLACE PHOTO HERE

Kristina Lerman is a Principal Scientist at the University of Southern California Information Sciences Institute and holds a joint appointment as a Research Associate Professor in the USC Computer Science Department. Trained as a physicist, she now applies network analysis and machine learning to problems in computational social science, including crowdsourcing, social network and social media analysis. Her recent work on modeling and understanding cognitive biases in social networks has been covered by the Washington Post, Wall Street Journal, and MIT Tech Review.

PLACE PHOTO HERE

Buddhika Nettasinghe received the M.A.Sc. degree from the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, Canada in 2016. He is currently a Ph.D. student with the School of Electrical and Computer Engineering, Cornell University and Cornell Tech, New York, NY, USA. His current research interests include statistical inference and learning, network science, computational social science and complex systems.

PLACE PHOTO HERE

Vikram Krishnamurthy (F'05) received the Ph.D. degree from the Australian National University, Canberra, ACT, Austrailia, in 1992. He is currently a Professor in the School of Electrical and Computer Engineering, Cornell University and Cornell Tech, New York, NY, USA. From 2002 to 2016, he was the Canada Research Chair Professor in statistical signal processing at the University of British Columbia, Vancouver, BC, Canada. His current research interests include statistical signal processing and stochastic control with applications in social networks. He served as a Distinguished Lecturer for the IEEE signal processing society and Editor-in-Chief of IEEE Journal Selected Topics in Signal Processing. He received an honorary doctorate from KTH (Royal Institute of Technology), Stockholm, Sweden in 2013.