



The dynamics of polarized beliefs in networks governed by viral diffusion and media influence

Mohammad Reza Sanatkar^{1,2}

Received: 17 June 2019 / Revised: 12 February 2020 / Accepted: 16 February 2020 / Published online: 2 March 2020
© Springer-Verlag GmbH Austria, part of Springer Nature 2020

Abstract

This paper studies the evolution of polarized beliefs governed by the intertwined dynamics of viral diffusion and media influence in influence networks. It addresses the question of how different forms of influence interact with each other. First, we propose a Markov chain to model the dynamics of individuals as they transition between three belief states (**neutral**, **positive** and **negative**) based on the states of their neighbors. This stochastic system assumes that individuals are influenced via the links of the network or through the global effect of mass media. For exponential and scale-free networks, we approximate this stochastic system by deterministic differential equations and define the homogeneous mean-field system and heterogeneous mean-field system, respectively. Studying stability conditions for these deterministic dynamical systems, we analyze the fraction of neutral, positive and negative individuals in the population. Also, we determine the conditions under which desired dynamical transitions happen for the targeted population. These conditions allow us to predict macroscopic measures of dynamics of adoption in influence networks. Finally, the derived analytical results are validated using simulations of four synthetic networks: Watts–Strogatz, random regular, Barabasi–Albert and small-world forest-fire, as well as five real-world networks: ego-Facebook, Deezer, Livemocha, a Facebook interaction network and Douban. Also, we demonstrate how the proposed model can be leveraged by marketing campaigns for optimal resource allocations between viral marketing and media marketing to minimize the number of final **negative** individuals in different network settings.

Keywords Polarized belief · Diffusion dynamics · Influence networks · Viral marketing · Media marketing

1 Introduction

To design effective marketing strategies that promote brand awareness, the adoption of innovations or the popularity of new products, it is crucial to take into account the influence networks of targeted populations. In influence networks, nodes represent individuals and links describe influence relationships between them. The topology of an influence network specifies the underlying structure of influence relationships between individuals. Marketing strategies can be divided into the two broad categories: viral and media marketing. Viral marketing exploits the structures of influence networks to activate existing influence to target potential adopters, and form global cascades of adopters in influence

networks. In particular, it is designed based on word-of-mouth and encourages individuals to share product information with their social contacts. Therefore, the medium for viral marketing is the links of social networks. On the other hand, media marketing treats all individuals as atomized objects of global media influence without taking into account their social networks. Media marketing is a broadcast mechanism that acts externally on influence networks because all the individuals receive the media influence directly from the same source. TV and newspaper ads are examples of media marketing, while personalized referrals and recommendations are examples of viral marketing. Studying the dynamics of viral diffusion and media influence in networks helps firms to design their marketing campaigns according to the characteristics of diffusion dynamics in the targeted populations.

1.1 Related work

Studying the diffusion of new ideas, beliefs and technologies, collectively called innovation, started about 120 years

✉ Mohammad Reza Sanatkar
mr.sanatkar@gmail.com

¹ Duke University, Durham, NC, USA

² Present Address: San Francisco, CA, USA

ago (Dearing 2008). Marketing campaigns can be seen as external change agents that interact with the dynamics of innovation in influence networks to maximize their primary output measures. The goal of marketing campaigns is to effect maximal change with minimal cost (Dearing 2008). Diffusion of innovations, viral marketing, has been studied extensively (Watts and Dodds 2007; Bakshy et al. 2011; Kempe et al. 2003; Leskovec et al. 2007; Kwak et al. 2010; Cha et al. 2010; Watts and Peretti 2007; Aral and Walker 2012; Cheng et al. 2014; Leskovec et al. 2006; Gomez Rodriguez et al. 2010; Sun et al. 2009; Margaris et al. 2016; Ribeiro 2014; De Bruyn and Lilien 2008; Saxena and Kumar 2019; Gomez Rodriguez and Schölkopf 2012; Du et al. 2013; Aral and Dhillon 2018; Vosoughi et al. 2018; Aral and Eckles 2019; Althoff et al. 2017; Cheng et al. 2018; Lazer et al. 2018; Kizilcec et al. 2018; Fink et al. 2016; Sarkar et al. 2017). Aral and Walker (2011) studied effects of viral product design strategies on creating word-of-mouth dynamics. In particular, they designed and conducted a randomized field experiment testing the effectiveness of active-personalized referrals and passive-broadcast notifications. They found that passive-broadcast viral messaging capabilities induce a higher increase in social contagion compared to active-personalized messaging. Sarkar et al. (2017) analyzed the occurrence of the following two lifecycle events of cascades: (1) the maximum growth period and (2) the period where declining in adoption starts. They studied the impact of network topology on the characteristics of these two periods based on the causality analysis of temporal adoption events (Sanatkar 2016).

Leskovec et al. (2007) studied dynamics of viral marketing in a recommendation network with 4 million people. They found that most recommendation chains do not turn into large cascades, and viral marketing is more effective for expensive products recommended to small and well-connected communities. Also, they explained the propagation of recommendations and cascade sizes in this network by a stochastic model. Cheng et al. (2018) studied the impact of diffusion protocols on cascade growth in online social networks. Analyzing the recurring classes of diffusion protocols, they found two key factors in the construction of such protocols are: (1) how much effort is required to participate in the cascade, where greater needed effort slows down the growth of the cascade, and (2) the social cost of not participating in the cascade, where higher social cost results in an increase in the cascade's adoption likelihood.

In the diffusion of innovation research community, one approach to diffusion maximization is to focus first on determining influential individuals (or influentials). Influentials are opinion leaders who have the credibility to influence a disproportionately large number of individuals. Aral and Walker (2012) showed that in a representative sample of 1.3 million Facebook users, influential individuals are less

susceptible to influence compared to non-influential individuals. Also, they observed that influentials cluster in the network, whereas non-influential individuals do not. Watts and Dodds (2007) studied the role of influentials in marketing and the formation of public opinion using a series of computer simulations of interpersonal influence processes. In particular, they studied the conditions under which local cascades can turn into global cascades and showed that under most conditions, the global cascades are driven by the critical mass of individuals who are easily influenced and not by influentials. More and Lingam (2019) took into account the rate of influence spread in addition to influence maximization in order to select subset of the network nodes as the influentials. Specifically, they proposed an algorithm based on gradient to achieve a balance between influence maximization and the influence spread rate.

Aral and Dhillon (2018) developed a class of empirically motivated influence models based on the network assortativity and the joint distribution of influence and susceptibility to identify more realistic sets of key influentials in order to efficiently disseminate information in social networks. Bakshy et al. (2011) investigated the word-of-mouth dynamics among Twitter users using the Twitter follower graph. They found that users with a large number of followers and users who have been influential in the past are the ones that generate the largest cascades. Also, they concluded that marketers can reliably benefit from word-of-mouth diffusion if large numbers of potential influencers are targeted. The diffusion process is also studied extensively in the context of epidemiology (Sanatkar et al. 2015a, b; Sahneh et al. 2013; Boguá et al. 2003; Pastor-Satorras and Vespignani 2001; Van Mieghem et al. 2009). They analyzed propagation of diseases over networks based on the stages of diseases in hosts where the transitions between these stages are modeled by several dynamics.

A myriad of studies show that the decision process of individuals is affected by mixtures of interpersonal and media influence (Watts and Dodds 2007). Analyzing empirical diffusion patterns over seven different online domains, Goel et al. (2012) concluded that these diffusion patterns motivate models that explicitly take into account media marketing in addition to viral marketing. While most studies of the diffusion process assume person-to-person networks as the only medium for diffusion of innovations, a number of recent studies consider both global influence by external sources and interpersonal networks as mechanisms of diffusion (Myers et al. 2012; Farajtabar et al. 2014; Kleineberg and Boguá 2014; Goel et al. 2012). Goel et al. (2012) showed that popular events regularly grow via combination of media marketing and viral diffusion. Myers et al. (2012) presented a model of information emergence in networks in which information can reach an individual through his neighbors or via the influence of an external source. They

fitted the parameters of their model to a complete one-month trace of the emergence of URL mentions in the Twitter network. They found that only about 71% of the tweets with URL mentions can be explained by viral propagation in the network of Twitter's followers, and the remaining 21% are due to external sources.

Also, Farajtabar et al. (2014) modeled intensities of endogenous and exogenous events in networks of individuals by multivariate Hawkes processes. They derived a time-dependent linear relationship that describes the relationship between the overall network activity and the intensity of exogenous events. Also, they computed the required level of external influence applied to the network to attain a desired activity level using a convex optimization framework. Kleineberg and Boguñá (2014) presented a two-layer model for the evolution of online social networks under viral spreading mechanisms and mass media influence. Based on the empirical validation of their model, they found that viral influence is 4–5 times stronger than mass media influence for the studied online social network.

1.2 Our proposed model

In this paper, we propose a stochastic system to model adoption process of polarized beliefs governed by viral diffusion and media influence at the individual level. The primary difference between our model and other recent studies (Goel et al. 2012; Myers et al. 2012; Farajtabar et al. 2014; Kleineberg and Boguñá 2014) that model both interpersonal and media influence is the following: We consider a third state, the so-called **negative** state, to represent those individuals who hold positions against the innovation in addition to two existing states [**neutral** (susceptible) and **positive** (adoption)] in Goel et al. (2012), Myers et al. (2012), Farajtabar et al. (2014) and Kleineberg and Boguñá (2014). To the best of our knowledge, our proposed stochastic system is the first model that takes into account both interpersonal and media influence for polarized beliefs in influence networks. In many real-world scenarios such as political debates, same-sex marriage, abortion and gun control, in additions to pro and neutral individuals, there exists a third group of individuals that are neither **positive**, adopter, nor **neutral**, susceptible, but against the innovation (Guerra et al. 2013; Yardi and Boyd 2010).

Our proposed model aims at describing such polarized belief dynamics over influence networks. Polarization on different issues becomes more and more widespread in America. In particular, regarding the recent political elections, ordinary people increasingly distrust those from the other political party. Democrats and Republicans consider the other party's members selfish and closed-minded (Iyengar et al. 2019). Strong political perspectives results in conflict and unwillingness to listen to the people from the opposite

political party (Shi et al. 2017). In particular, homogeneous social networks strengthen polarization of partisan preferences, which results in a decrease in tolerance for alternative views in addition to reduction in opportunities for crosscutting political interactions (Lazer et al. 2018). A non-political example of polarization is related to the recent emergence of smart watches. People have had polarized ideas toward smart watches. While many people support using them because of the convenience that comes with receiving notifications on their wrists, many people think that we are already very preoccupied with our smartphones in our daily lives, and we do not need another device. Moreover, Amato et al. (2017) explain the widespread presence of polarization in social issues via modeling the coupling interactions between social influence and decisions of individuals as different layers of a multiplex network.

First, we propose a stochastic system to model the dynamics of polarized beliefs at the individual level. This stochastic system is a Markov chain and is called the individual-based stochastic (IN-STOCH) system. This system is described by a set of individual-based transition events that govern dynamics of polarized belief propagation in influence networks. Then, using a mean-field analysis, we approximate the IN-STOCH system in the large population limit by deterministic differential equations so-called the homogeneous mean-field (HOM-MEAN) and the heterogeneous mean-field (HET-MEAN) systems for exponential and scale-free networks, respectively. The HOM-MEAN system is based on homogeneity and randomness assumptions of influence networks. However, the HET-MEAN system does not assume homogeneity for influence networks and instead assumes that the nodes with different degrees can potentially demonstrate different dynamics. In other words, the HET-MEAN system keeps tracks of the fractions of **neutral**, **negative** and **positive** separately for the nodes with different degrees. The HET-MEAN system has higher complexity than HOM-MEAN system which is a requirement for scale-free networks.

We show that the HOM-MEAN system has at most three equilibrium points and one of these is always unstable. Then, we prove that the stability of the other two depends on the parameters of the model and the average node degrees in the influence network. In particular, one of the two stable equilibria has a zero fraction of **negative** individuals, the so-called **negative-free** equilibrium points. We are interested in **negative-free** equilibrium points based on the chosen measure for the success of marketing campaigns, which is minimizing the number of individuals with **negative** state. In practice, it is extremely optimistic to assume that a marketing campaign can reduce the number of individuals with **negative** state all the way to zero. However, the conditions derived for the stability of the **negative-free** equilibrium points will ensure that the marketing strategies chosen by

marketing campaigns are in the direction of reducing the number of **negative** individuals, while the actual convergence to **negative**-free equilibrium points that guarantees no **negative** individuals at all is asymptotic and might not happen during the finite time-spans of activities of marketing campaigns. Based on the stability of the **negative**-free equilibrium point, we derive conditions for local and global convergence of the fraction of **negative** individuals to zero. Critical values of model parameters corresponding to these conditions characterize different phase transitions of polarized dynamics.

Then, we derive the dynamical system describing the HET-MEAN system that distinguishes between nodes with different degrees. The derived HET-MEAN dynamical system depends on the node degree distribution of the influence networks. We show that this dynamical system has two equilibria. Next, we state that the necessary and sufficient conditions of local exponential stability for the **negative**-free equilibrium point can be investigated by computing the eigenvalues of the Jacobian matrix at the equilibrium point. Also, we derive a sufficient condition for the global stability of the **negative**-free equilibrium point if it is already locally stable. However, since the HET-MEAN system equations depend on the node degree distributions of the influence networks, the local stability condition of the **negative**-free equilibrium point cannot be derived in a closed form without fixing the node degree distribution. As an example, we derive the local exponential stability of the equilibrium point of the HET-MEAN system by fixing the node degree distribution of the influence network to the node degree distribution of the Barabasi–Albert (BA) networks (Barabasi and Albert 1999). In order to compute the stability condition for the BA networks, first, we compute an approximation for the probability of an edge to be connected to a **neutral** node at the equilibrium point using the node degree distribution of the BA networks. Then, we use this approximation to approximate the final fraction of **neutral** nodes at the equilibrium point. Finally, we compute an approximation for the probability of an edge to be connected to **negative** nodes. We employ this probability approximation to derive the local stability condition for the equilibrium point of the BA networks.

Finally, we validate the predictions of final fractions of **neutral**, **positive** and **negative** individuals by the HOM-MEAN and HET-MEAN systems via simulating the IN-STOCH system over the following synthetic networks: random regular, Watts–Strogatz, Barabasi–Albert networks and the small-world forest-fire networks (Drossel and Schwabl 1992) as well as five real-world networks: ego-Facebook (Leskovec and McAuley 2012), Deezer (Rozemberczki et al. 2019), Livemocha (Zafarani and Liu 2009), a Facebook interaction network (Viswanath et al. 2009) and Douban (Zafarani and Liu 2009). Also, we show that how the HOM-MEAN and HET-MEAN systems can be leveraged by marketing

campaigns to divide resources in an optimal way between viral marketing and media marketing according to characteristics of influence networks and targeted populations. The HET-MEAN system can be used for similar use cases by marketing campaign for scale-free influence networks.

The rest of this paper is organized as follows. In Sect. 2, we propose the IN-STOCH system to model the individual-based dynamics of polarized beliefs. Section 3 contains the HOM-MEAN system to approximate the IN-STOCH system for exponential networks at population level. In Sect. 4, we derive the HET-MEAN system to approximate the IN-STOCH system for scale-free networks. Finally, in Sect. 5, we use simulation results to validate our theoretical analysis.

2 Individual-based stochastic system

In this section, we propose a stochastic model, called the individual-based stochastic (IN-STOCH) system, of the adoption process of polarized dynamics at the individual level. This stochastic system assumes that individuals are influenced via the links of influence networks or through the global mass media. In social networks such as Facebook where individuals are either associated (e.g., friends) or unassociated, the influence network is assumed to be an undirected graph. In social networks such as Twitter where individuals follow one another, the influence network is assumed to be a directed graph. In this work, we consider only undirected influence networks. The extension to directed influence networks is straightforward.

In this model, at any given time, each node is in one of the three states: **neutral** (N), **positive** (P) and **negative** (U). Each state represents a different mindset of individuals in the adoption process. An individual in the **neutral** state does not have a strong opinion toward the innovation, or has not received any information about the innovation. On the other hand, **positive** individuals are the ones who have adopted the innovation. For example, they already bought the product and are satisfied with it, or are willing to buy it. The **negative** state represents those individuals with negative opinions toward the innovation or product. For instance, they are who already bought the product but are not happy with it, or have been convinced by some of their friends not to buy it.

Polarized beliefs propagate among individuals based on the following mechanisms (as shown in Fig. 1). A **neutral** individual may become **positive** due to viral adoption or media influence. Viral (or word-of-mouth) adoption is modeled by a stochastic reaction process. A neutral node may virally adopt the innovation with a probability β per each of its **positive** neighbors. On the other hand, media influence is modeled by a stochastic diffusion process. Under this process, a **neutral** individual will become **positive** with probability α because of media influence. It is noted that the media influence probability

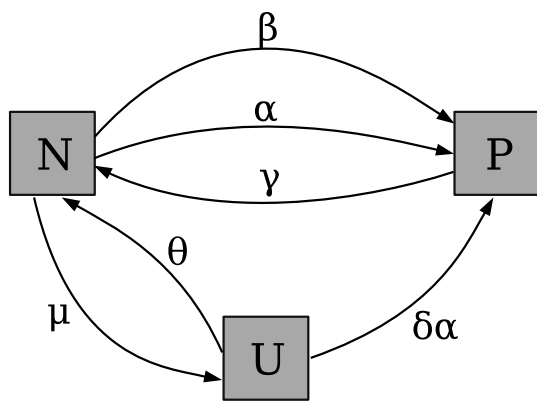


Fig. 1 Flowchart of the individual-based stochastic system. A neutral node may become positive with a probability β per its positive neighbors and because of media influence with a probability α . A positive node may return to the neutral state with a probability γ . A neutral node may become negative with a probability μ per its negative neighbors. A negative node may return to the neutral state with a probability θ and to the positive state with a probability $\delta\alpha$

α is identical for all individuals and independent of the states of their neighbors.

A neutral individual may become negative under a word-of-mouth process with a probability μ per each of its negative neighbors. A negative individual may become neutral with a probability θ . One cause for this transition can be a series of compensating actions taken by marketing campaigns to make negative individuals neutral. Here, we assume that mass media for a negative user is not as effective as mass media for a neutral user. A negative user may become positive with a probability $\delta\alpha$ under media influence, where $\delta \in [0, 1]$. A positive individual may become neutral with a probability γ . One example for this stochastic process can be transitions of those individuals who have bought the product, but, after using it for a while, are not satisfied with it.

In fact, the IN-STOCH system is a Markov chain that governs the state transitions of individuals with respect to the states of their neighbors. For any given node i , let R_i denote the set of its neighbors. Also, $P_{i,n}(t)$, $P_{i,p}(t)$ and $P_{i,u}(t)$ denote the probability of node i being at state neutral, positive and negative, respectively, at time step t . The state of node j at time step t is denoted by $x_j(t)$. Let $R_{i,p}(t) := \sum_{j \in R_i} \mathbb{1}(x_j(t) = P)$ and $R_{i,u}(t) := \sum_{j \in R_i} \mathbb{1}(x_j(t) = U)$, where $\mathbb{1}$ denotes the indicator function. Therefore, the IN-STOCH system is equivalent to the following Markov chain:

$$\begin{aligned}
 P_{i,n}(t + 1) &= \gamma P_{i,p}(t) + \theta P_{i,u}(t) \\
 &\quad + (1 - \alpha - \beta R_{i,p}(t) - \mu R_{i,u}(t)) P_{i,n}(t), \\
 P_{i,p}(t + 1) &= (\alpha + \beta R_{i,p}(t)) P_{i,n}(t) + \delta\alpha P_{i,u}(t) \\
 &\quad + (1 - \gamma) P_{i,p}(t), \\
 P_{i,u}(t + 1) &= \mu R_{i,u}(t) P_{i,n}(t) + (1 - \theta - \delta\alpha) P_{i,u}(t).
 \end{aligned}
 \tag{1}$$

This model considers three belief states and a specific subset of possible transitions. We choose this specific subset based on two criteria: (1) The resultant mean-field model is tractable and the closed-form analytical stability analysis of the mean-field model is feasible and (2) the selected set of transitions allows each individual to transient from any state to another state in at most two time steps. For example, even though based on the chosen transitions, there is no direct transition from positive state to negative state, an individual might perform such a transition via first transition to neutral state from positive state and then to negative state from the neutral state. However, we note that these transitions do not cover all possible adoption scenarios and were chosen based on the adoption scenarios we envision. For example, for the transition from the neutral state to the negative state, we only consider viral adoption and not media influence. It is because our focus is those marketing campaigns that are designed to promote a product and therefore only contribute to media influence transition from the neutral state to the positive state. That being said, there could exist scenarios in which two rival marketing campaigns work against each other where one of them aims at promoting the product and the other one's goal is to degrade the image of the product in public. Here, the proposed model only considers the scenarios where there exists a single marketing campaigns promoting the product. Developing a model that allows the simultaneous existence of two viral marketing campaigns could be a potential next step in this line of research.

The goal of intervening agencies and marketing campaigns is to steer a targeted population to a targeted state. These agencies intervene to alter the dynamics of influence networks and maximize their primary output measures. Many different output measures can be chosen by intervening agencies based on the characteristics of targeted populations and innovations. Some examples of these output measures are driving the overall percentage of adoption to a certain level or making the adoption level homogeneous across users belonging to different communities. In this paper, the asymptotic fraction of individuals with negative state is chosen as the primary output measure. In other words, we design the intervention process realized by a combination of viral diffusion and media influence to minimize (e.g., set to zero) the asymptotic fraction of negative individuals. Our main reason for choosing this specific output measure is that the lack of negative individuals prevents the viral dynamics from drifting toward the negative state.

3 Exponential networks

In this section, we use a mean-field approach to derive a set of deterministic differential equations so-called the homogeneous mean-field (HOM-MEAN) system to approximate the IN-STOCH system at population level for exponential networks. Networks with exponential degree distributions are called exponential networks. An exponential degree distribution peaks at the average node degree $\langle k \rangle$ and decays exponentially fast for node degrees k when $k \gg \langle k \rangle$ or $k \ll \langle k \rangle$. The aim is to predict macroscopic measures of the IN-STOCH system. Then, we study asymptotic behaviors of targeted populations by examining the stability of the equilibrium points of the HOM-MEAN system. We show that this dynamical system has at most three equilibrium points. Based on our chosen output measure for marketing campaigns, the desirable equilibrium is the one whose fraction of **negative** individuals is zero. We derive critical values for dynamical phase transitions of populations by studying local and global stability of this equilibrium point. In particular, two sets of conditions are derived that guarantee the local and global convergence of the fraction of **negative** individuals to zero.

Let $n(t)$, $p(t)$ and $u(t)$ be the fraction of **neutral**, **positive** and **negative** individuals, respectively, at time t so that $n(t) + p(t) + u(t) = 1$. For a variable that depends on time, such as $n(t)$, we use n to denote $n(t)$ and \dot{n} to denote the time derivative $n'(t)$. Thus, the HOM-MEAN dynamical system is described with the following coupled differential equations in continuous time:

$$\begin{aligned} \dot{n} &= -k\beta pn - \alpha n + \gamma p - k\mu nu + \theta u, \\ \dot{p} &= k\beta pn + \alpha n - \gamma p + \delta \alpha u, \\ \dot{u} &= k\mu nu - \theta u - \delta \alpha u. \end{aligned} \tag{2}$$

The HOM-MEAN system is derived assuming the independence between states of individuals. Also, it assumes that the fluctuations in exponential degree distributions around average node degrees can be neglected. This assumption is consistent with the empirical literature on influentials (Brock and Durlauf 2001). That is, node degree distributions of influence networks display a relatively little variation around their averages. For influence networks with non-local connectivity, we expect that the derived analytical results at population level based on the HOM-MEAN system will approximate the behavior of targeted population (Pastor-Satorras and Vespignani 2001). Using $n + p + u = 1$ reduces (2) to the two-dimensional system:

$$\begin{aligned} \dot{n} &= k\beta n^2 - (k\beta + \alpha + \gamma)n + (k\beta - k\mu)nu \\ &\quad + (\theta - \gamma)u + \gamma, \\ \dot{u} &= k\mu nu - (\theta + \delta \alpha)u. \end{aligned} \tag{3}$$

3.1 Equilibrium points

To study the steady-state behavior of the HOM-MEAN system, we start by finding its equilibrium points. We want to choose a set of values for parameters that ensure the stability of the equilibrium points with no **negative** individuals. The equilibrium points of the dynamical system are found by imposing the stationary conditions. That is, $\dot{n} = 0$ and $\dot{u} = 0$. Let (\bar{n}, \bar{u}) be an equilibrium point of this dynamical system. From

$$\dot{u} = k\mu \bar{n} \bar{u} - (\theta + \delta \alpha) \bar{u} = 0, \tag{4}$$

we see that either $\bar{u} = 0$ or $\bar{n} = (\theta + \delta \alpha)/(k\mu)$.

Replacing $\bar{u} = 0$ in

$$\begin{aligned} \dot{n} &= k\beta \bar{n}^2 - (k\beta + \alpha + \gamma) \bar{n} \\ &\quad + (k\beta - k\mu) \bar{n} \bar{u} + (\theta - \gamma) \bar{u} + \gamma = 0, \end{aligned} \tag{5}$$

we have

$$k\beta \bar{n}^2 - (k\beta + \alpha + \gamma) \bar{n} + \gamma = 0. \tag{6}$$

Lemma 1 *The dynamical system in Eq. (3) has at least one equilibrium point with no **negative** individuals if and only if*

$$\alpha + \gamma \leq k\beta + \sqrt{(k\beta + \alpha + \gamma)^2 - 4k\beta\gamma}. \tag{7}$$

Proof The candidates for the equilibrium points with no **negative** individuals are the solutions of (6). Equation (6) has a real solution if and only if $(k\beta + \alpha + \gamma)^2 - 4k\beta\gamma \geq 0$. Let $\Delta := (k\beta + \alpha + \gamma)^2 - 4k\beta\gamma$. Now, we show that Δ is always nonnegative. We have

$$(k\beta - \gamma)^2 + \alpha^2 + \gamma^2 + 2k\beta\alpha + 2\alpha\gamma \geq 0. \tag{8}$$

By expanding and rearranging, we can show that

$$\Delta = k^2\beta^2 + \alpha^2 + \gamma^2 - 2k\beta\gamma + 2k\beta\alpha + 2\alpha\gamma. \tag{9}$$

Therefore, (6) has at least one real solution and at maximum two solutions that can be written as

$$\begin{aligned} \bar{n}_1 &= \frac{k\beta + \alpha + \gamma + \sqrt{\Delta}}{2k\beta}, \\ \bar{n}_2 &= \frac{k\beta + \alpha + \gamma - \sqrt{\Delta}}{2k\beta}. \end{aligned} \tag{10}$$

Both \bar{n}_1 and \bar{n}_2 are always nonnegative because $\sqrt{\Delta} \leq k\beta + \alpha + \gamma$. $(\bar{n}_1, 0)$ and $(\bar{n}_2, 0)$ are equilibrium points if \bar{n}_1 and \bar{n}_2 , respectively, are less than or equal to one. It is because the state space of the HOM-MEAN system is $[0, 1] \times [0, 1]$. Now, we show that if (7) is satisfied, \bar{n}_2 will be less than or equal to 1. Because $\bar{n}_2 \leq \bar{n}_1$, if $\bar{n}_2 \leq 1$, then

the HOM-MEAN system has at least one equilibrium point with no **negative** individuals. $\bar{n}_2 \leq 1$ if and only if

$$\alpha + \gamma \leq k\beta + \sqrt{\Delta}. \tag{11}$$

□

The HOM-MEAN system may have another equilibrium point with nonzero fraction of **negative** individuals. Replacing $\bar{n} = (\theta + \delta\alpha)/(k\mu)$ in

$$\begin{aligned} \dot{n} &= k\beta\bar{n}^2 - (k\beta + \alpha + \gamma)\bar{n} \\ &+ (k\beta - k\mu)\bar{n}\bar{u} + (\theta - \gamma)\bar{u} + \gamma = 0, \end{aligned} \tag{12}$$

we have

$$\bar{u} = \frac{-k\mu\gamma + (\delta\alpha + \theta)(k\beta + \alpha + \gamma) - \frac{\beta}{\mu}(\theta + \delta\alpha)^2}{k\beta(\theta + \delta\alpha) - k\mu(\gamma + \delta\alpha)}. \tag{13}$$

3.2 Stability of equilibrium points

In this section, we study the stability of the equilibrium points of the HOM-MEAN system. In particular, we show that, of the two possible equilibrium points whose fraction of **negative** individuals is zero, only $(\bar{n}_2, 0)$ can be locally stable, and $(\bar{n}_1, 0)$ is always unstable. Also, conditions for global asymptotic stability of $(\bar{n}_2, 0)$ are derived.

An equilibrium point of a dynamical system is locally exponentially stable if and only if the real parts of all eigenvalues of its Jacobian matrix, computed at the equilibrium point, are negative, and it is unstable if at least one of the eigenvalue of the Jacobian matrix, computed at the equilibrium point, has a positive real part (Khalil 2000). Let $A(n^*, u^*)$ be the Jacobian matrix of the HOM-MEAN system in (3) at point (n^*, u^*) , where $A_{11}(n^*, u^*) = \frac{\partial n}{\partial n}|_{(n^*, u^*)}$, $A_{12}(n^*, u^*) = \frac{\partial n}{\partial u}|_{(n^*, u^*)}$, $A_{21}(n^*, u^*) = \frac{\partial u}{\partial n}|_{(n^*, u^*)}$ and $A_{22}(n^*, u^*) = \frac{\partial u}{\partial u}|_{(n^*, u^*)}$. Hence, we have

$$\begin{aligned} A_{11}(n^*, u^*) &= 2k\beta n^* + (k\beta - k\mu)u^* - (k\beta + \alpha + \gamma), \\ A_{12}(n^*, u^*) &= (k\beta - k\mu)n^* + \theta - \gamma, \\ A_{21}(n^*, u^*) &= k\mu u^*, \\ A_{22}(n^*, u^*) &= k\mu n^* - \theta - \delta\alpha. \end{aligned} \tag{14}$$

Computing the matrix A at $(\bar{n}_1, 0)$, we have

$$\begin{aligned} A_{11} &= \sqrt{\Delta}, \\ A_{12} &= \frac{\beta - \mu}{2\beta} (k\beta + \alpha + \gamma + \sqrt{\Delta}) + \theta - \gamma, \\ A_{21} &= 0, \\ A_{22} &= \frac{\mu}{2\beta} (k\beta + \alpha + \gamma + \sqrt{\Delta}) - \theta - \delta\alpha. \end{aligned} \tag{15}$$

The Jacobian matrix computed at $(\bar{n}_1, 0)$ is an upper triangular matrix. Also, it is a real matrix because $\Delta \geq 0$ as we show in Lemma 1. Therefore, its two eigenvalues are real and equal to A_{11} and A_{22} . We can conclude that $(\bar{n}_1, 0)$ is always unstable because A_{11} is always positive. The Jacobian matrix at $(\bar{n}_2, 0)$ can be computed as:

$$\begin{aligned} A_{11} &= -\sqrt{\Delta}, \\ A_{12} &= \frac{\beta - \mu}{2\beta} (k\beta + \alpha + \gamma - \sqrt{\Delta}) + \theta - \gamma, \\ A_{21} &= 0, \\ A_{22} &= \frac{\mu}{2\beta} (k\beta + \alpha + \gamma - \sqrt{\Delta}) - \theta - \delta\alpha. \end{aligned} \tag{16}$$

The two eigenvalues of the Jacobian matrix are real and equal to A_{11} and A_{22} because the Jacobian matrix is upper triangular and Δ is always nonnegative. The local stability of $(\bar{n}_2, 0)$ is only determined by A_{22} because A_{11} is always negative. In the following theorem, we derive conditions under which $(\bar{n}_2, 0)$ is locally exponentially stable.

Theorem 1 *If (7) in Lemma 1 is satisfied, then $(\bar{n}_2, 0)$ is locally exponentially stable if and only if*

$$\begin{aligned} \text{(a)} \quad &\mu < \frac{2\beta(\theta + \delta\alpha)}{k\beta + \alpha + \gamma}, \text{ or} \\ \text{(b)} \quad &\mu > \frac{2\beta(\theta + \delta\alpha)}{k\beta + \alpha + \gamma} \text{ and} \\ &\gamma < \frac{1}{\mu - \frac{1}{k}(\delta\alpha + \theta)} \left((\delta\alpha + \theta)(\beta + \frac{1}{k}\alpha) - \frac{\beta}{k\mu}(\theta + \delta\alpha)^2 \right). \end{aligned} \tag{17}$$

Proof If (7) in Lemma 1 is satisfied, then $(\bar{n}_2, 0)$ is locally exponentially stable if and only if A_{22} is negative. That is,

$$\frac{\mu}{2\beta} (k\beta + \alpha + \gamma - \sqrt{\Delta}) - \theta - \delta\alpha < 0. \tag{18}$$

The inequality in (18) holds if and only if

$$-\sqrt{\Delta} < \frac{2\beta}{\mu}(\theta + \delta\alpha) - (k\beta + \alpha + \gamma). \tag{19}$$

If the RHS of (19) is positive, the inequality in (19) holds because its LHS is always negative. The RHS of (19) is positive if and only if

$$\mu < \frac{2\beta(\theta + \delta\alpha)}{k\beta + \alpha + \gamma}, \tag{20}$$

which yields constraint (a). However, if the RHS of (19) is negative, then the inequality in (19) is satisfied if and only if

$$\Delta > \left(\frac{2\beta}{\mu}(\theta + \delta\alpha) - (k\beta + \alpha + \gamma) \right)^2. \tag{21}$$

The inequality in (21) is equivalent to

$$k\mu\gamma < \delta\alpha(k\beta + \alpha + \gamma) + \theta(k\beta + \alpha + \gamma) - \frac{\beta}{\mu}(\theta + \delta\alpha)^2, \quad (22)$$

which yields constraint (b). □

The constraints in Theorem 1 show two different regimes that both guarantee convergence of the fraction of **negative** individuals to zero. In constraint (a), the condition for this convergence is characterized by μ . We call the regime corresponding to this constraint, regime 1. On the other hand, in constraint (b), γ characterizes the convergence to zero for the fraction of **negative** individuals. The regime corresponding to this constraint is called regime 2. We define **negative-free** thresholds with respect to the conditions in Theorem 1 to represent the desired phase transitions that guarantee the convergence to zero for the fraction of **negative** individuals. Let $\eta_1 := \frac{2\beta(\theta + \delta\alpha)}{k\beta + \alpha + \gamma}$ denote the **negative-free** threshold for regime 1. Also, let $\eta_2 := \frac{1}{\mu - \frac{1}{k}(\delta\alpha + \theta)} \left((\delta\alpha + \theta)(\beta + \frac{1}{k}\alpha) - \frac{\beta}{k\mu}(\theta + \delta\alpha)^2 \right)$ denote the **negative-free** threshold for regime 2. In regime 1, $(\bar{n}_2, 0)$ is exponentially stable if and only if $\mu < \eta_1$. In regime 2, $(\bar{n}_2, 0)$ is exponentially stable if and only if $\mu > \eta_1$ and $\gamma < \eta_2$.

HOM-MEAN systems with larger **negative-free** thresholds are more resistant to processes that push individuals toward the **negative** state. Therefore, intervening agencies should design their intervention process to make **negative-free** thresholds as large as possible. In regime 1, increasing the average node degree decreases the **negative-free** threshold, η_1 , so that decreasing the chance of stability of $(\bar{n}_2, 0)$. In other words, increasing the average node degree helps the word-of-mouth process that causes individuals become **negative**. In regime 2, condition (b) in Theorem 1 reduces to

$$\gamma < \frac{\beta}{\mu}(\delta\alpha + \theta), \quad (23)$$

as $k \rightarrow \infty$. Equation (23) shows the significance of both viral diffusion and media influence for highly connected influence networks. As $k \rightarrow \infty$,

$$\eta_2 \rightarrow \frac{\beta}{\mu}(\delta\alpha + \theta). \quad (24)$$

Theorem 1 provides us with the necessary and sufficient conditions of the local stability of $(\bar{n}_2, 0)$, but not its global stability. In the following theorem, we derive the conditions that ensure the equilibrium point with a nonzero fraction of **negative** individuals,

$$\begin{aligned} \bar{n} &= (\theta + \delta\alpha)/(k\mu) \\ \bar{u} &= \frac{-k\mu\gamma + (\delta\alpha + \theta)(k\beta + \alpha + \gamma) - \frac{\beta}{\mu}(\theta + \delta\alpha)^2}{k\beta(\theta + \delta\alpha) - k\mu(\gamma + \delta\alpha)}, \end{aligned} \quad (25)$$

does not lie in the state space of the HOM-MEAN system.

Theorem 2 *If one of the conditions in Theorem 1 is satisfied, then the $(\bar{n}_2, 0)$ point is the only stable equilibrium point of the HOM-MEAN system if one of the following conditions is satisfied:*

- (a) $\mu < \frac{1}{k}(\theta + \delta\alpha)$,
- (b) $\mu < \mu_0$ and $\gamma > \gamma_0$,
- (c) $\mu > \mu_0$ and $\gamma < \gamma_0$,
- (d) $\mu < \mu_0$, $\gamma < \gamma_0$ and $(\delta\alpha + \theta)(\alpha + \gamma) + k\mu\delta\alpha > \frac{\beta}{\mu}(\theta + \delta\alpha)^2$,
- (e) $\mu > \mu_0$, $\gamma > \gamma_0$ and $(\delta\alpha + \theta)(\alpha + \gamma) + k\mu\delta\alpha < \frac{\beta}{\mu}(\theta + \delta\alpha)^2$,

where $\mu_0 = \beta(\theta + \delta\alpha)/(\gamma + \delta\alpha)$ and $\gamma_0 = \left((\delta\alpha + \theta)(\beta + \frac{1}{k}\alpha) - \frac{\beta}{k\mu}(\theta + \delta\alpha)^2 \right) / \left(\mu - \frac{1}{k}(\delta\alpha + \theta) \right)$.

Proof Given $(\bar{n}_1, 0)$ is always an unstable equilibrium point, we derive the conditions that guarantee the equilibrium point with a nonzero fraction of **negative** individuals in (25) does not lie in the state space, $[0, 1] \times [0, 1]$. If $(\theta + \delta\alpha)/(k\mu) > 1$, then this equilibrium does not lie in the state space. Condition (a) is equivalent to $(\theta + \delta\alpha)/(k\mu) > 1$. Conditions (b) and (c) ensure that \bar{u} in (25) is negative. Let N and D denote the numerator and denominator of \bar{u} in (25), respectively. \bar{u} in (25) is negative if and only if either $N > 0$ and $D < 0$, or $N < 0$ and $D > 0$. The inequalities $N < 0$ and $D > 0$ yield condition (b), and $N > 0$ and $D < 0$ lead to condition (c). If \bar{u} in (25) is positive, $N > 0$ and $D > 0$, or $N < 0$ and $D < 0$, \bar{u} being greater than one ensures that this equilibrium point does not lie in the state space. The inequalities $N > 0$ and $D > 0$, and \bar{u} being greater than 1 yield condition (d), and if $N < 0$ and $D < 0$, they yield condition (e). □

If $(\bar{n}_2, 0)$ is locally stable but not globally stable, then its region of attraction quantifies how robust this equilibrium point is with respect different initial states. The region of attraction for $(\bar{n}_2, 0)$ is the subset of the state space which trajectories initiating from, asymptotically converge to $(\bar{n}_2, 0)$. Determining the region of attraction is important since tells us how far the initial state can be from $(\bar{n}_2, 0)$ and still converge to $(\bar{n}_2, 0)$. In the following theorem, we estimate the region of attraction for $(\bar{n}_2, 0)$.

Theorem 3 *If $(\bar{n}_2, 0)$ is locally exponentially stable, trajectories converge to this equilibrium point if their initial states lie in the set R_D defined as follows:*

$$R_D = [0, 1]^2 \cap \hat{R}_D, \tag{27}$$

where

$$\hat{R}_D = \{(n, u) | p_1(n - \bar{n}_2)^2 + 2p_2(n - \bar{n}_2)u + p_4u^2 < \eta\}, \tag{28}$$

where

$$\begin{aligned} p_1 &= -\frac{1}{2A_{11}}, \\ xp_2 &= \frac{A_{12}}{2A_{11}(A_{11} + A_{22})}, \text{ and} \\ p_4 &= \frac{-A_{12}^2}{2A_{11}A_{22}(A_{11} + A_{22})}. \end{aligned} \tag{29}$$

A_{11} , A_{12} , and A_{22} are derived in (16). Also, η is a positive number less than $\lambda_{\min}(P)r^2$ where $r = 1/(f + \frac{1}{2}\sqrt{g^2 + h^2})$ and $P = \begin{bmatrix} p_1 & p_2 \\ p_2 & p_4 \end{bmatrix}$. And, $f = 2p_1k\beta$, $g = 2(p_1(k\beta - k\mu) + p_2(k\beta + k\mu))$, $h = 2(p_2(k\beta - k\mu) + p_4k\mu)$.

Proof See ‘‘Appendix A.’’ □

4 Scale-free networks

In this section, we use a mean-field approach to approximate the IN-STOCH system for scale-free networks by a dynamical system called the heterogeneous mean-field (HET-MEAN) system that is represented by a set of deterministic differential equations. Unlike exponential networks, the degree distributions of scale-free networks have strong fluctuations around their average node degrees. This characteristic of scale-free networks motivates a different mean-field analysis from the one for exponential networks, that distinguishes between the dynamics of the nodes with different node degrees. In particular, the HET-MEAN system separately keeps tracks of the fraction of neutral, positive and negative nodes with different degrees k .

Because of the dependency of the HET-MEAN system equations on the node degree distributions of influence networks, it is not possible to derive closed-form formulas to predict the final fraction of neutral, positive and negative individuals without fixing the node degree distribution. In other words, stability analysis of equilibrium points of the HET-MEAN system and derivation of closed-form prediction formulas cannot be done without choosing a specific node degree distribution. To demonstrate such a use case, in this section, after developing results for a general form scale-free network, we focus on the BA networks which are special type of scale-free networks. Fixing the node degree distribution to be the node degree distribution of the BA networks, we derive closed-form approximation for its

corresponding negative-free equilibrium point and negative-free conditions.

Let $n_k(t)$, $p_k(t)$ and $u_k(t)$ be the fraction of neutral, positive and negative individuals with node degree k at time t . Therefore, $n_k(t) + p_k(t) + u_k(t) = 1$. The HET-MEAN system assumes the independence between states of individuals and is described with the following coupled differential equations in continuous time:

$$\begin{aligned} \dot{n}_k &= -k\beta\Phi(p)n_k - \alpha n_k + \gamma p_k - k\mu\Phi(u)n_k + \theta u_k, \\ \dot{p}_k &= k\beta\Phi(p)n_k + \alpha n_k - \gamma p_k + \delta \alpha u_k, \\ \dot{u}_k &= k\mu\Phi(u)n_k - \theta u_k - \delta \alpha u_k, \end{aligned} \tag{30}$$

where k spans over all the distinct node degrees in the given influence network; $\Phi(p)$ and $\Phi(u)$ denote the probabilities that any given edge is connected to a positive and a negative node, respectively. Therefore, $\Phi(p) + \Phi(u) + \Phi(n) = 1$, where $\Phi(n)$ is the probability that any given edge is connected to a neutral node. Let $P(k)$ denote the node degree distribution of the influence network. Therefore, using $P(k)$, one can compute $\Phi(p)$, $\Phi(u)$ and $\Phi(n)$ given p_k , u_k and n_k are known, as follows:

$$\Phi(x) = \frac{1}{\langle k \rangle} \sum kP(k)x_k, \tag{31}$$

where x can be either p , n or u and $\langle k \rangle$ denotes the average node degree computed. Using $n_k + p_k + u_k = 1$, we reduce (30) to the following system:

$$\begin{aligned} \dot{n}_k &= k(\beta - \mu)\Phi(u)n_k + k\beta\Phi(n)n_k - (k\beta + \alpha + \gamma)n_k \\ &\quad + (\theta - \gamma)u_k + \gamma, \\ \dot{u}_k &= k\mu\Phi(u)n_k - (\theta + \delta \alpha)u_k. \end{aligned} \tag{32}$$

The variables of the reduced mean-field dynamical system in (32) are $\{n_k\}$ and $\{u_k\}$. The dimension of this dynamical system state space is $2K$, where K is the number of distinct node degrees.

4.1 Equilibrium points

In this section, first, we show that the HET-MEAN system has only one negative-free equilibrium point candidate. Then, we prove that this negative-free equilibrium point candidate is always a valid equilibrium point for the HET-MEAN system. In the following lemma, using the stationary equation of the HET-MEAN system, we show that there exists only one negative-free equilibrium point candidate for this dynamical system.

Lemma 2 *The only negative-free equilibrium point candidate of the HET-MEAN system is the following:*

$$\left\{ \left(\bar{u}_k = 0, \bar{n}_k = \frac{\gamma}{k\beta(1 - \bar{\Phi}(n)) + \alpha + \gamma} \right) \right\}. \tag{33}$$

Proof The stationary conditions for the HET-MEAN dynamical system are $\dot{n}_k = 0$ and $\dot{u}_k = 0$. Imposing these stationary conditions, we find the equilibrium points of the HET-MEAN system. Let $\{(\bar{u}_k, \bar{n}_k)\}$ be an equilibrium point of this dynamical system. From $\dot{u}_k = 0$, we see that either $\bar{u}_k = 0$ or

$$\bar{n}_k = \frac{\theta + \delta\alpha}{k\mu\bar{\Phi}(u)} \bar{u}_k, \tag{34}$$

where $\bar{\Phi}(u)$ and $\bar{\Phi}(n)$ denote the values of $\Phi(u)$ and $\Phi(n)$ that satisfy stationary conditions, respectively. Here, we focus on the equilibrium point that satisfies $\bar{u}_k = 0$ for every k . It is clear that this equilibrium point candidate is **negative-free** since for every k , the fraction of **negative** nodes is equal to 0. This equilibrium point being **negative-free**, for its corresponding $\bar{\Phi}(u)$, we can write

$$\bar{\Phi}(u) = 0. \tag{35}$$

Next, to fully characterize this **negative-free** equilibrium point candidate, we compute its \bar{n}_k components. Replacing $u_k = 0$ in $n_k = 0$, we have

$$\bar{n}_k = \frac{\gamma}{k\beta(1 - \bar{\Phi}(n)) + \alpha + \gamma}. \tag{36}$$

Therefore, the **negative-free** equilibrium point candidate can be written as follows:

$$\left\{ \bar{u}_k = 0, \bar{n}_k = \frac{\gamma}{k\beta(1 - \bar{\Phi}(n)) + \alpha + \gamma} \right\}. \tag{37}$$

□

In the following Lemma, we show that the **negative-free** equilibrium point candidate computed in Lemma 2 is always an equilibrium point for the HET-MEAN system. This result is significant since it shows that there is always a possibility for marketing campaigns to design their marketing strategies for scale-free influence networks to drive the dynamics of networks to converge to a **negative-free** state.

Lemma 3 *The negative-free equilibrium point candidate for the HET-MEAN system computed in Lemma 2 is always an equilibrium point for the HET-MEAN system.*

Proof The **negative-free** equilibrium point candidate (37) will be an equilibrium point if and only if it lies in the state space of the dynamical system. That is $\forall k$,

$$0 \leq \bar{n}_k \leq 1. \tag{38}$$

Equation (38) is equivalent to $0 \leq \bar{\Phi}(n) \leq 1$. Now, we show that $0 \leq \bar{\Phi}(n) \leq 1$ always holds. We have

$$\begin{aligned} \bar{\Phi}(n) &= \frac{1}{\langle k \rangle} \sum k \bar{n}_k P(k) \\ &= \frac{1}{\langle k \rangle} \sum \frac{\gamma k P(k)}{k\beta(1 - \bar{\Phi}(n)) + \alpha + \gamma}. \end{aligned} \tag{39}$$

Let $g(\bar{\Phi}(n)) := \sum \gamma k P(k) / (k\beta(1 - \bar{\Phi}(n)) + \alpha + \gamma) / \langle k \rangle$. Therefore, $0 \leq \bar{\Phi}(n) \leq 1$ if and only if the equation $\bar{\Phi}(n) = g(\bar{\Phi}(n))$ has a solution for $\bar{\Phi}(n)$ in $[0, 1]$. We have

$$\begin{aligned} g(0) &= \frac{1}{\langle k \rangle} \sum \frac{k P(k)}{1 + \frac{1}{c}(k\beta + \alpha)}, \\ g(1) &= \frac{\gamma}{\gamma + \alpha}. \end{aligned} \tag{40}$$

We have $0 \leq g(0) \leq 1$ and $0 \leq g(1) \leq 1$. Therefore, $\bar{\Phi}(n) = g(\bar{\Phi}(n))$ has a solution in $[0, 1]$. □

4.2 Stability of negative-free equilibrium point

We examine the stability of the **negative-free** equilibrium point (37) by first computing the partial derivatives of the HET-MEAN system (32) at this equilibrium point. Then, we use these partial derivatives to form the $2K \times 2K$ Jacobian matrix. Finally, we study the eigenvalues of this matrix to analyze the stability of the **negative-free** equilibrium point. In the following, we compute the partial derivatives of the HET-MEAN system at the **negative-free** equilibrium point. The partial derivatives $\frac{\partial \dot{n}_k}{\partial n_k}$ can be written as follows:

$$\frac{\partial \dot{n}_k}{\partial n_k} = k\beta\bar{\Phi}(n) + \frac{k^2}{\langle k \rangle} \beta \bar{n}_k P(k) - (k\beta + \alpha + \gamma). \tag{41}$$

For $\frac{\partial \dot{n}_k}{\partial n_j}, \forall j \neq k$, we can write

$$\frac{\partial \dot{n}_k}{\partial n_j} = \frac{kj}{\langle k \rangle} \beta \bar{n}_k P(j). \tag{42}$$

Computing $\frac{\partial \dot{n}_k}{\partial u_k}$, we have

$$\frac{\partial \dot{n}_k}{\partial u_k} = \frac{k^2}{\langle k \rangle} (\beta - \mu) \bar{n}_k P(k) + \theta - \gamma. \tag{43}$$

$\frac{\partial \dot{n}_k}{\partial u_j}, \forall j \neq k$, can be derived as follows:

$$\frac{\partial \dot{n}_k}{\partial u_j} = \frac{kj}{\langle k \rangle} (\beta - \mu) \bar{n}_k P(j). \tag{44}$$

We have $\frac{\partial u_k}{\partial n_j} = 0, \forall j$. Computing $\frac{\partial \dot{u}_k}{\partial u_k}$, we can write

$$\frac{\partial \dot{u}_k}{\partial u_k} = \frac{k^2}{\langle k \rangle} \mu \bar{n}_k P(k) - (\theta + \delta\alpha). \tag{45}$$

Finally, computing $\frac{\partial \dot{u}_k}{\partial u_j}, \forall j \neq k$, we have

$$\frac{\partial \dot{u}_k}{\partial u_j} = \frac{k_j}{\langle k \rangle} \mu \bar{n}_k P(j). \tag{46}$$

Using the above-derived partial derivatives, we can form the Jacobian matrix at the **negative-free** equilibrium point. We form the two following submatrices of the Jacobian matrix to simplify the stability analysis of the equilibrium point. Let $J_{\bar{n}}$ denote the $K \times K$ submatrix formed by $\partial \dot{n}_k / \partial n_j$ and $\partial \dot{n}_k / \partial n_j$; and $J_{\bar{u}}$ denote the $K \times K$ submatrix formed by $\partial \dot{u}_k / \partial u_k$ and $\partial \dot{u}_k / \partial u_j$. In the following theorem, we rely on these two $K \times K$ matrices to derive necessary and sufficient conditions for locally exponentially stability of the **negative-free** equilibrium point.

Theorem 4 *The negative-free equilibrium point (37) is locally exponentially stable if and only if the real parts of all eigenvalues of $J_{\bar{n}}$ and $J_{\bar{u}}$ are negative.*

Proof The equilibrium point is locally exponentially stable if and only if the real parts of all eigenvalues of the Jacobian matrix, computed at this equilibrium point, are negative. Because $\partial \dot{u}_k / \partial n_j = 0$ at the equilibrium point in (37), the eigenvalues of the Jacobian matrix are equal to the union of the eigenvalues of $J_{\bar{n}}$ and $J_{\bar{u}}$. Therefore, this equilibrium point is locally exponentially stable if and only if the real parts of all eigenvalues of $J_{\bar{u}}$ and $J_{\bar{n}}$ are negative. \square

In the following theorem, we provide a sufficient condition for the global stability of the **negative-free** equilibrium point of the HET-MEAN system.

Theorem 5 *The negative-free equilibrium point (37) is globally exponentially stable if the condition of Theorem 4 is satisfied and $\mu < \frac{1}{k}(\theta + \delta\alpha)$.*

Proof If the condition of Theorem 4 is satisfied, then the **negative-free** equilibrium point is locally exponentially stable, and in order to prove it is globally exponentially stable, One approach is to ensure that the HET-MEAN system does not have any other equilibrium point. The HET-MEAN system, in addition to the **negative-free** equilibrium point, has only one more equilibrium point candidate. This equilibrium point candidate has nonzero final fraction of **negative** individuals. In order to compute this equilibrium point, imposing the stationary condition $\dot{u}_k = 0$, we have $\bar{n}_k = \bar{u}_k(\theta + \delta\alpha)/(k\mu\bar{\Phi}(u))$. Replacing \bar{n}_k by $\bar{u}_k(\theta + \delta\alpha)/(k\mu\bar{\Phi}(u))$ in $n = 0$, we have

$$\bar{u}_k = \frac{\gamma k \mu \bar{\Phi}(u)}{I + (\gamma - \theta) k \mu \bar{\Phi}(u)}, \tag{47}$$

where $I = (\theta + \delta\alpha)(k(\mu - \beta)\bar{\Phi}(u) + k\beta(1 - \bar{\Phi}(n)) + \alpha + \gamma)$. Based on $\bar{n}_k = \bar{u}_k(\theta + \delta\alpha)/(k\mu\bar{\Phi}(u))$, \bar{n}_k has a linear relationship with \bar{u}_k with the coefficient $C = (\theta + \delta\alpha)/(k\mu\bar{\Phi}(u))$. Using (31), it can be shown that $\bar{\Phi}(n) = C\bar{\Phi}(u)$. Therefore, we have

$$\bar{\Phi}(n) = \frac{\theta + \delta\alpha}{k\mu\bar{\Phi}(u)} \bar{\Phi}(u) = \frac{\theta + \delta\alpha}{k\mu}. \tag{48}$$

We use the above closed-form formula derived for $\bar{\Phi}(n)$ to force this equilibrium candidate to be outside of the state space of the HET-MEAN system. This idea is realized by forcing $\bar{\Phi}(n) > 1$ which translates into the following condition:

$$\mu < \frac{1}{k}(\theta + \delta\alpha). \tag{49}$$

\square

4.3 Brabasi–Albert networks

In this section, we apply our theoretical results for the HET-MEAN system to Barabasi–Albert (BA) networks, which are examples of scale-free networks (Barabasi and Albert 1999). The underlying idea behind the BA networks is that individuals tend to make connections with individuals who are more well-known and who have a larger number of existing connections. This is an intuitive phenomenon that is observed on social networks. To construct a BA network, we start from a small number m_0 of disconnected nodes; every new node forms m edges to existing nodes such that the probability of a new edge to be connected to an old node is proportional to the degree of old node (Barabasi and Albert 1999). The constructed network has the following degree distribution

$$P(k) = \begin{cases} 2m^2 k^{-3}, & \text{if } k \geq m \\ 0, & \text{otherwise} \end{cases}, \tag{50}$$

with average node degree $\langle k \rangle = 2m$ (Barabasi and Albert 1999).

For the equilibrium point with no **negative** individuals in (37), we have $\bar{n}_k = \gamma/(k\beta(1 - \bar{\Phi}(n)) + \alpha + \gamma)$. In order to compute the steady-state fraction of **neutral** nodes $\bar{n} = \sum P(k)\bar{n}_k$, first, we need to compute $\bar{\Phi}(n)$. We can calculate $\bar{\Phi}(n)$ using a self-consistency equation that is derived from the fact that \bar{n}_k is also a function of $\bar{\Phi}(n)$. We approximate the sum in (51) by an integral in order to derive the self-consistency equation. Therefore, we have

$$\begin{aligned} \bar{\Phi}(n) &= \frac{1}{\langle k \rangle} \sum_{k=m}^{\infty} kP(k)\bar{n}_k \\ &\approx \frac{1}{2m} \int_m^{\infty} \frac{2m^2k^{-2}\gamma}{k\beta(1-\bar{\Phi}(n)) + \alpha + \gamma} dk. \end{aligned} \tag{51}$$

Computing the integral in (51), we derive the following self-consistency equation for $\bar{\Phi}(n)$:

$$\begin{aligned} \ln \left(1 + \frac{\alpha + \gamma}{m\beta(1 - \bar{\Phi}(n))} \right) \\ = \frac{\alpha + \gamma}{m\beta\gamma} \left(\gamma - \alpha \frac{\bar{\Phi}(n)}{1 - \bar{\Phi}(n)} \right), \end{aligned} \tag{52}$$

where $\bar{\Phi}(n)$ can be found numerically. Also, using the integral approximation, we have

$$\begin{aligned} \bar{n} &= \sum_{k=m}^{\infty} P(k)\bar{n}_k \approx \int_m^{\infty} 2m^2k^{-3} \frac{\gamma}{k\beta(1 - \bar{\Phi}(n)) + \alpha + \gamma} dk \\ &= \frac{2m^2\gamma\beta^2(1 - \bar{\Phi}(n))^2}{(\alpha + \gamma)^3} \ln \left(1 + \frac{\alpha + \gamma}{m\beta(1 - \bar{\Phi}(n))} \right) \\ &\quad + 2m^2\gamma \left(\frac{\beta(1 - \bar{\Phi}(n))}{(\alpha + \gamma)^2m} - \frac{1}{2(\alpha + \gamma)m^2} \right). \end{aligned} \tag{53}$$

For the other equilibrium point corresponding to $\bar{n}_k = \bar{u}_k(\theta + \delta\alpha)/(k\mu\bar{\Phi}(u))$, first we calculate $\bar{\Phi}(u)$ through a self-consistency equation using the fact that \bar{u}_k is a function of $\bar{\Phi}(u)$. Employing the integral approximation, we can write

$$\begin{aligned} \bar{\Phi}(u) &= \frac{1}{\langle k \rangle} \sum_{k=m}^{\infty} kP(k)\bar{u}_k \\ &\approx \frac{1}{2m} \int_m^{\infty} 2m^2k^{-2}\bar{u}_k dk \\ &= \bar{\Phi}(u) \frac{msf}{g} \ln \left(1 + \frac{g}{m(f\bar{\Phi}(u) + h)} \right), \end{aligned} \tag{54}$$

where $s = \gamma\mu^2/f$; $f = \mu((\mu - \beta)(\theta + \delta\alpha) + (\gamma - \theta))$; $g = (\theta + \delta\alpha)(\beta(\delta\alpha - \theta) + \mu(\alpha + \gamma))$; $h = \beta\mu(\theta + \delta\alpha)$. Using the self-consistency equation in (54), we can derive explicitly $\bar{\Phi}(u)$ as follows:

$$\bar{\Phi}(u) = \frac{g}{mf} \left(e^{\frac{g}{m\gamma\mu^2}} - 1 \right)^{-1} - \frac{h}{f}. \tag{55}$$

Finally, using the integral approximation, we can derive the fraction of negative individuals as

$$\bar{u} = \sum_{k=m}^{\infty} P(k)\bar{u}_k \approx \int_m^{\infty} 2m^2k^{-3}\bar{u}_k, \tag{56}$$

where \bar{u}_k is derived in (47). Computing the integral in (56), we have

$$\bar{u} = H \left(1 - \frac{m}{g} (f\bar{\Phi}(u) + h) \ln \left(1 + \frac{g}{(f\bar{\Phi}(u) + h)m} \right) \right), \tag{57}$$

where $H = 2msf\bar{\Phi}(u)/g$.

However, Theorem 4 can be used to determine the stability of the equilibrium point with no negative individuals in (37) for the BA networks; here, we derive alternative criteria to examine the stability of this equilibrium point. Given $\bar{\Phi}(u)$ is the probability that any given edge is connected to a negative individual, having no negative node in steady state is equivalent to $\bar{\Phi}(u) \leq 0$.

Lemma 4 For the BA networks, asymptotic fraction of negative individuals will be 0 if

$$e^{\frac{g}{mc\mu^2}} - 1 \geq \frac{1}{m} \left(\frac{1}{\mu}(\delta\alpha - \theta) + \frac{1}{\beta}(\alpha + c) \right). \tag{58}$$

Proof The inequality in (58) is equivalent to $\bar{\Phi}(u) \leq 0$. □

The main advantage of Lemma 4 over Theorem 4 to check whether the fraction of negative individuals is equal 0 for the BA networks, is its lower computational cost. It is because for Theorem 4, we need to compute the eigenvalues of two $K \times K$ matrices.

5 Simulation results

In this section, we compare the analytical approximations by the HOM-MEAN and HET-MEAN systems with the simulation results of the IN-STOCH system on the following synthetic networks: random regular, Watts–Strogatz (WS), BA networks the small-world forest-fire networks (Drossel and Schwabl 1992) as well as five real-world networks: ego-Facebook (Leskovec and Mcauley 2012), Deezer (Rozemberczki et al. 2019), Livemocha (Zafarani and Liu (2009), a Facebook interaction network (Viswanath et al. 2009) and Douban (Zafarani and Liu 2009). In particular, we compare the analytical approximations for the steady-state fraction of negative individuals with the results of simulating the IN-STOCH system over the influence networks. Also, using an example of exponential networks and a case study over the Deezer network, we demonstrate how potentially the proposed model can be leveraged by marketing campaigns to design their resource allocations between viral marketing and media marketing to minimize the number of final negative individuals in different network settings.

For all the simulation results of the IN-STOCH system in this section, initial fractions of neutral, positive and negative nodes are randomly drawn from the probability simplex for each network realization. Because of the lack of access

to real-world polarities, we rely on the following method to make the initial states of the individuals realistic by choosing their initial polarities based on their neighbors. For each positive and negative state, we randomly choose 100 nodes as the root nodes for the positive and negative clusters. Based on simulation results, we found that the number of root nodes does not impact the final steady state across different networks. Then, we propagate positive and negative states in these clusters using the Breadth First Search algorithm till the number of positive and negative nodes reaches the drawn initial fractions of positive and negative from the probability simplex.

5.1 Exponential networks

In this section, we focus on WS and random regular networks that lie in the category of exponential networks. For these networks, the analytical approximations are based on the results of Theorems 1 and 2 in combination with the derived equilibrium point with a nonzero fraction of negative individuals in (25). If either Theorem 1 or Theorem 2 is satisfied, then the analytical approximation for steady-state fraction of negative individuals is equal to 0; otherwise, it is equal to (25). In the first experiment, simulations are performed using randomly generated WS networks with $N = 10000$ nodes. WS networks are used as reference point for homogeneous networks that model small-world networks. WS networks are shown to be good candidates to model real-world networks in the context of social networks (Watts and Strogatz 1998). Small-world networks are networks with short average path lengths and high clustering. To construct a WS network, we start from a ring of N nodes. Choosing an even number k , each node is connected with its $k/2$ nearest clockwise neighbors and $k/2$ nearest counter clockwise neighbors. Then, with probability p , every link connected to a clockwise neighbor is rewired to a randomly chosen node. As a result of this process, a WS network with the average node degree k is created. One of the important characteristics of these network is their small fluctuations in connectivity due to their exponential degree distributions.

We iterate the rules of the IN-STOCH system for 2000 steps. Three different ensembles of WS networks are studied with average node degrees of $k = 12, 16$ and 20 and with probability of rewiring $p = 0.2$. In Fig. 2, both analytical and simulation results of final fraction of negative individuals as a function of μ/β in the WS networks are presented. The final fraction of negative individuals in the steady state is averaged over ten different realizations of WS networks for each case. The analytical approximations of Theorems 1 and 2, based on the HOM-MEAN system, are in total agreement with the simulation results of the IN-STOCH system; they correctly predict when final fractions of negative individuals are 0. When neither Theorem 1

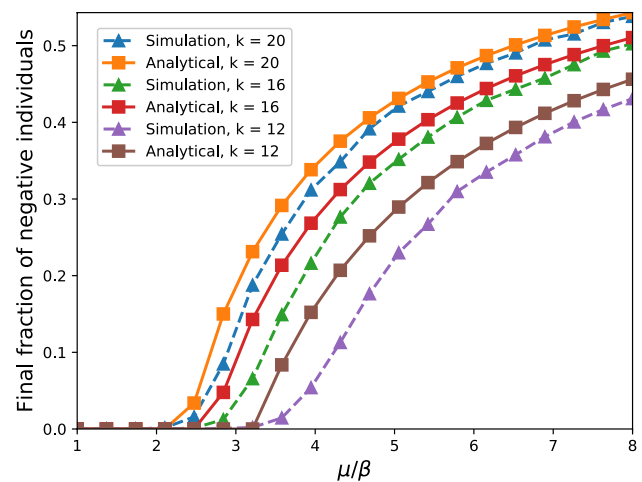


Fig. 2 Final fraction of negative individuals for WS networks as a function of μ/β . Solid lines are theoretical predictions. Dashed lines are simulation results. Parameter values: $\theta = 0.01, \beta = 0.001, \alpha = 0.01, \delta = 0.4$ and $\gamma = 0.01$

nor Theorem 2 holds, the analytical approximation for the steady-state fraction of negative individuals by (25) is in good agreement with the simulation results with the average absolute difference of 0.033, 0.021 and 0.015, respectively, for the average node degree of 12, 16 and 20. This agreement becomes stronger by increasing the average node degree. Also, Theorem 1 predicts that regime 1 is active for all the three cases, which is confirmed by simulation results. The negative-free threshold decreases by increasing the average node degree. It shows that for the given scenario, increasing the average node degree helps the word-of-mouth process by negative individuals more than the word-of-mouth process by positive individuals.

In the second experiment, simulations are performed in random regular networks with $N = 10000$ nodes. Random regular networks are similar to regular networks in the sense that all nodes have identical node degrees. However, neighbors of each node are chosen randomly. Iterating the rules of the IN-STOCH system is done over three ensembles of random regular networks with node degrees of $k = 12, 16$ and 20 for 2000 steps. In Fig. 3, we plot the analytical approximations and the simulation results of final fractions of negative individuals versus μ/β . The plotted simulation results for each node degree are resultant of averaging the simulation results of ten different realizations of random regular networks. Similar to WS networks, the simulation results are in total agreement with the analytical approximations of Theorems 1 and 2. And, when neither of these two theorems hold, the steady-state fractions of negative individuals predicted by (25) are in good agreement with the simulation results with the average absolute difference of 0.020, 0.013 and 0.007, respectively, for node degrees 12, 16 and 20. Following the same trend as WS networks, this

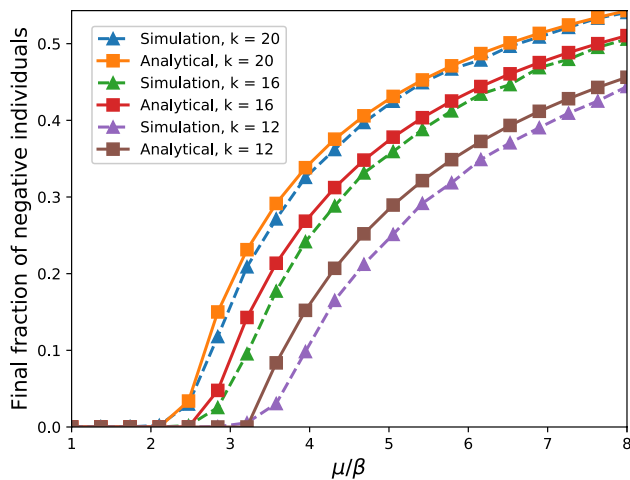


Fig. 3 Final fraction of **negative** individuals for random regular networks as a function of μ/β . Solid lines are theoretical predictions. Dashed lines are simulation results. Parameter values: $\theta = 0.01$, $\beta = 0.001$, $\alpha = 0.01$, $\delta = 0.4$ and $\gamma = 0.01$

agreement becomes stronger by increasing the node degree. However, for all three values of node degrees, (25) predicts more accurately the steady-state fraction of **negative** individuals for the random regular networks compared to the WS networks. This is because the HOM-MEAN system assumes that all nodes have identical node degrees, which is true for random regular networks, but not for WS networks. Similar to WS networks, the **negative-free** threshold decreases by increasing the average node degree.

5.2 Scale-free networks

In this section, we validate the analytical results for the scale-free networks based on the HET-MEAN system using simulation of the proposed stochastic adoption model over a number of synthetic and real-world networks that demonstrate the characteristics of scale-free networks. We start with BA networks. The accuracy of predictions of the HET-MEAN system is studied based on simulations of the IN-STOCH system over the BA networks with $N = 1000$ nodes. Three different ensembles of BA networks with the parameters of $m = 8, 12$ and 16 are randomly generated. For the all generated BA networks, m_0 is equal to m . For all cases, the simulation total time steps are equal to 5000. Figure 4 plots the approximations of steady-state fractions of **negative** individuals by the HET-MEAN system and the simulation results versus μ/β . The analytical approximations are computed as follows: If condition (58) in Lemma 4 is satisfied, then the final fraction of **negative** individuals is equal to 0; otherwise, (57) is used to compute the final fraction of **negative** individuals. The plotted simulation results for each value of m are resultant of averaging the simulation results of ten different realizations of the BA networks. We observe

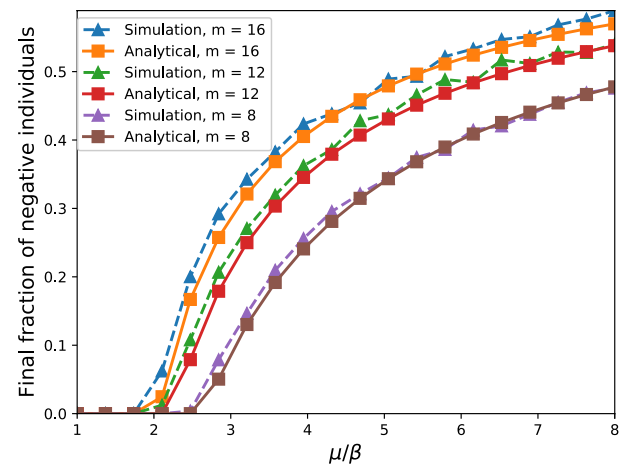


Fig. 4 Final fraction of **negative** individuals for BA networks as a function of μ/β . Solid lines are theoretical predictions. Dashed lines are simulation results. Parameter values: $\theta = 0.01$, $\delta = 0.4$ and $\gamma = 0.01$

the analytical approximations in a good agreement with the simulation results with the average absolute difference of 0.006, 0.011 and 0.0132, respectively, for m parameter of 8, 12 and 16.

In order to predict the final fraction of **negative** individuals for real-world networks, one of the following three approaches can be taken. (1) Approximation of their empirical degree distributions with one of the discussed degree distributions in this paper and directly employ their mean-field predictions. For example, if the empirical node degree distribution of a real-world network is concentrated around its average node degree, then the analytical results of the HOM-MEAN system can be used directly. (2) If the empirical node degree distribution is not concentrated around the average node degree, one can approximate the empirical node degree distribution with a closed-form mathematical formula using a curve fitting method and plug the curve fitted distribution into the framework developed for scale-free networks in this paper and follow the same steps as we did for the BA networks to derive analytical closed-form prediction results. (3) Another possible approach when the empirical node degree distribution is not concentrated around its average node degree is to use the computed empirical node degree distribution to numerically solve the HET-MEAN system equations (32) in order to predict the final fraction of **negative**, **positive** and **neutral** nodes. The main gain of numerically solving the HET-MEAN system compared to the directly simulation of networks' dynamics using the stochastic IN-STOCH system is the significance reduction in the computational cost. It is because the dimension of the stochastic IN-STOCH system is equal to the number of the nodes in the network, whereas the dimension of the HET-MEAN system is equal to the number of distinct node

degrees in the network. To illustrate this point, we consider ego-Facebook network.

Ego-Facebook network (Leskovec and Mcauley 2012) is an undirected network consisting of 4039 nodes and 88234 edges with average node degree 43.7 and standard deviation 52.4; it was generated by using a Facebook app based on friends lists. Figure 5 shows the node degree histogram of this network. While the number of nodes is 4039, the dominant values of the histogram is limited to only about 200 node degrees. This shows the computational gain that can be accomplished with using the HET-MEAN system instead of the IN-STOCH system. In general, it is expected that this computational gain becomes more significant as networks become larger since the gap between the number of nodes and the dominant values of the node degree histograms increases. In order to numerically solve the HET-MEAN system (32) for ego-facebook network, first, we compute its empirical degree distribution. Then, we simulate the dynamical system in (32) till its state evolution reaches the steady state. Finally, we report the fraction of negative individuals at the steady state. Also, we simulate the IN-STOCH system over this network for 10000 steps by iterating the rules of the IN-STOCH system. Figure 6 depicts the results of numerical solving of the HET-MEAN system and the simulation of the IN-STOCH system for the ego-Facebook network with ten different random initializations. It shows a good agreement between the predictions of these two systems for the final fraction of negative individuals with the average absolute difference of 0.028.

Then, we compare the predictions by the HET-MEAN (32) with the simulation result of the IN-STOCH system for the Deezer network (Rozemberczki et al. 2019). This dataset represents friendship relationships among users of the music streaming service Deezer during November 2019. In particular, we use the part of this dataset with the individuals

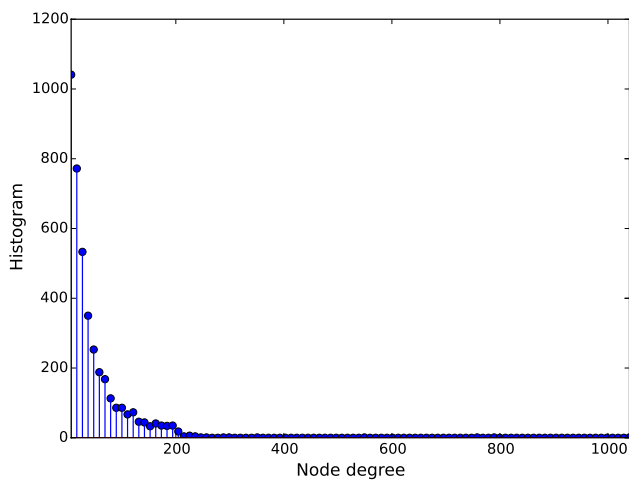


Fig. 5 Histogram of node degrees of the ego-Facebook network

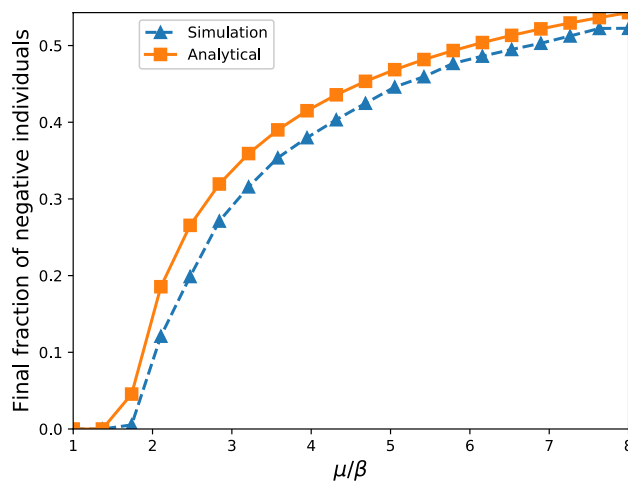


Fig. 6 Final fraction of negative individuals in the ego-Facebook network as a function of μ/β . The solid line is the theoretical predictions. The dashed line is the simulation results. Parameter values: $\theta = 0.01$, $\beta = 0.001$, $\alpha = 0.01$, $\delta = 0.4$ and $\gamma = 0.01$

from Hungary. The network represented by this part of dataset consists of 54573 nodes and 498202 edges. We numerically solve Equation (32) using the computed node degree histogram of the Deezer network till the convergence of the HET-MEAN dynamical system. Also, the IN-STOCH system is simulated over this network for 2000 steps and ten different random initializations. Figure 7 shows the simulation results of the final fraction of negative individuals for IN-STOCH system as well as the numerical solutions of the HET-MEAN system for the Deezer network. A good agreement between analytical and simulation results with the average absolute difference of 0.0195 is observed.

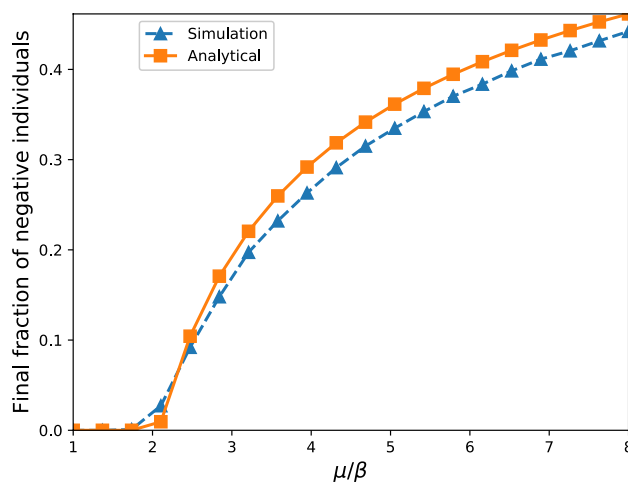


Fig. 7 Final fraction of negative individuals in the Deezer network as a function of μ/β . The solid line is the theoretical predictions. The dashed line is the simulation results. Parameter values: $\theta = 0.01$, $\beta = 0.001$, $\alpha = 0.01$, $\delta = 0.4$ and $\gamma = 0.01$

Next, we examine the predictions of the HET-MEAN system for a real-world network called Livemocha (Zafarani and Liu 2009). This network represents friendship relationships among users of the online language learning community Livemocha and is formed by 104103 nodes and 2193083 undirected edges. Figure 8 compares the numerical solution of the HET-MEAN system and the simulation results of the IN-STOCH system after 5000 steps averaged over ten different random initializations for the final fraction of **negative** individuals. There exists a good agreement between the predictions of the HET-MEAN system and the results of the IN-STOCH system with the average absolute difference of 0.008.

Furthermore, the prediction of the HET-MEAN system is compared with the simulation result of the IN-STOCH system over a Facebook interaction network (Viswanath et al. 2009). This network represents 817035 logged activities among 63731 users of New Orleans regional Facebook network over a period of two years, where the activities among users are denoted by undirected edges. Figure 9 compares the numerical solution of the HET-MEAN system and the simulation results of the IN-STOCH system after 5000 steps averaged over ten different random initializations for the final fraction of **negative** individuals. There exists a good agreement between the predictions of the HET-MEAN system and the results of the IN-STOCH system with the average absolute difference of 0.023.

The last real-world network that we use to study the accuracy of the predictions of the HET-MEAN system is the Douban network (Zafarani and Liu 2009). This undirected network consists of 154,908 nodes and 327,162 edges and represents the friendship network of the Chinese recommendation Web site Douban for books, movies and music.

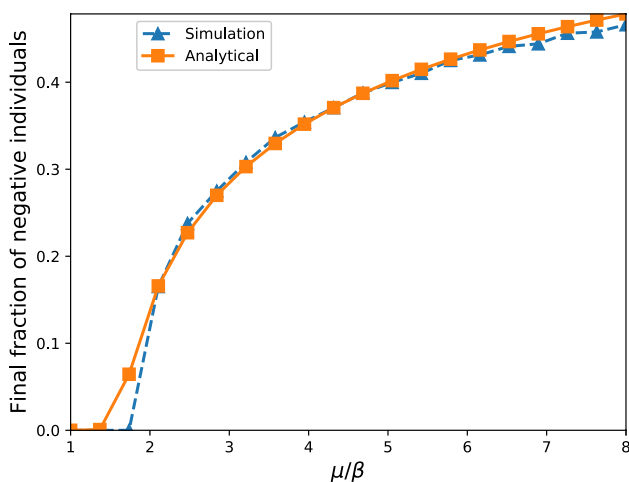


Fig. 8 Final fraction of **negative** individuals in the Livemocha network as a function of μ/β . The solid line is the theoretical predictions. The dashed line is the simulation results. Parameter values: $\theta = 0.01$, $\beta = 0.001$, $\alpha = 0.01$, $\delta = 0.4$ and $\gamma = 0.01$

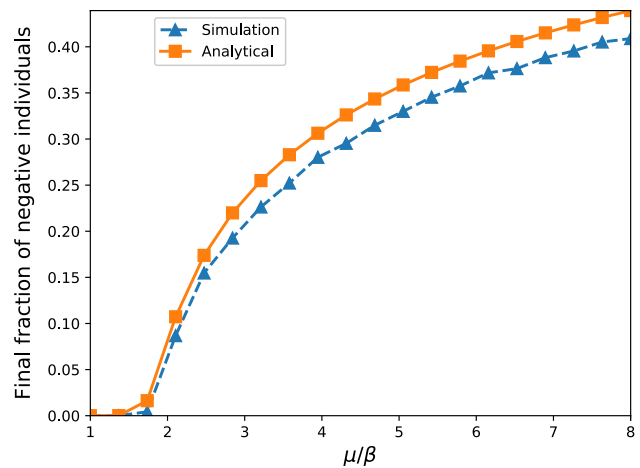


Fig. 9 Final fraction of **negative** individuals in a Facebook interaction network as a function of μ/β . The solid line is the theoretical predictions. The dashed line is the simulation results. Parameter values: $\theta = 0.01$, $\beta = 0.001$, $\alpha = 0.01$, $\delta = 0.4$ and $\gamma = 0.01$

Figure 10 shows the comparison of the HET-MEAN system predicted final fraction of **negative** individuals with their counterparts obtained via the simulation of the IN-STOCH system for 5000 steps averaged over ten different random initializations. The average absolute difference between the predictions of the HET-MEAN system and the results of the IN-STOCH system is equal to 0.005, which demonstrates a good agreement between them.

Next, we examine the proposed IN-STOCH system and the mean-field system equations (32) on the small-world forest-fire networks (Drossel and Schwabl 1992). The small-world forest-fire networks are constructed over two-dimensional lattice graphs. At each given time, each vertex of such

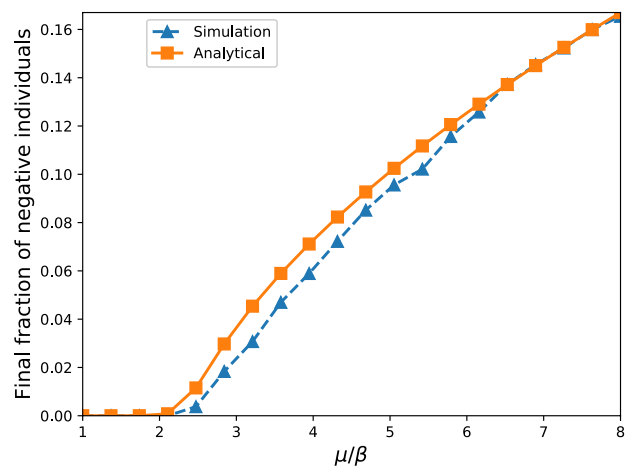


Fig. 10 Final fraction of **negative** individuals in the Douban network as a function of μ/β . The solid line is the theoretical predictions. The dashed line is the simulation results. Parameter values: $\theta = 0.01$, $\beta = 0.001$, $\alpha = 0.01$, $\delta = 0.4$ and $\gamma = 0.01$

lattice graphs is occupied by either a tree or a burning tree or it is empty. The following dynamical system modeling fires in forests governs the final patterns of trees: (1) Tree grows with the probability p at an empty vertex. (2) A vertex with burning tree becomes an empty vertex and all of its neighboring vertices' trees switch to burning trees. (3) A tree becomes a burning tree with the lightning probability of f . Following (Drossel and Schwabl 1992), each vertex has at most one long-range connection (shortcut) with the probability q to a randomly chosen vertex.

Here, we realize a small-world forest-fire network over 100×100 two-dimensional lattice with $p = 0.05$, $f = 0.000001$ and $q = 0.5$. For initialization of the lattice, we randomly choose 0.55 fraction of vertices to be occupied by trees and the rest to be empty. We iterate over the forest-fire dynamical model for 100 steps assuming the local neighborhood of each node consists of its 8 nearest neighbors. Then, the IN-STOCH system is simulated over this network by iterating of the update rules of the IN-STOCH system, for ten different random initializations of the IN-STOCH system. Figure 11 shows the average final fraction of **negative** individuals versus values of μ/β for these ten simulations of the IN-STOCH system as well as the analytical predictions computed by numerically solving the HET-MEAN system in (32). Similar to the case studies of the real-world networks, we rely on numerically solving the HET-MEAN system in (32) instead of using closed-form analytical predictions of the final fraction of **negative** individuals to demonstrate that numerically solving the HET-MEAN system can be applied to synthetic networks as well. Figure 11 illustrates a relatively good agreement between the simulation results and analytical predictions with the average absolute difference of 0.032.

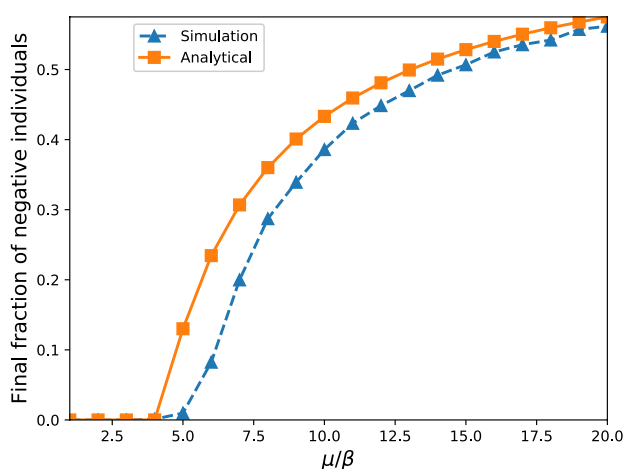


Fig. 11 Final fraction of **negative** individuals in a forest-fire network as a function of μ/β . The solid line is the theoretical predictions. The dashed line is the simulation results. Parameter values: $\theta = 0.01$, $\beta = 0.001$, $\alpha = 0.01$, $\delta = 0.4$ and $\gamma = 0.01$

5.3 Viral marketing versus media marketing

In this section, the HOM-MEAN and HET-MEAN systems are used to compare the impact of viral diffusion with media influence for exponential influence and scale-free networks, respectively. In our framework, a marketing scenario is fully characterized by the values of γ , θ , δ , μ and the topology of influence network. The fraction β/α represents the relative ratio investment of marketing campaigns in viral marketing compared to media marketing. The set $\{(\beta, \alpha) \in [0, 1] \times [0, 1]\}$ spans the space of all different intervening strategies which are formed by simultaneous viral and media marketing. Moreover, Ω denotes the set of those intervening strategies that result in final zero **negative** individuals, and Ψ denotes the set of those intervening strategies that result in final nonzero **negative** individuals. Therefore, the sets Ω and Ψ form a partition for the mentioned set of all intervening strategies.

Figure 12 depicts Ω and Ψ corresponding to two different scenarios for the HOM-MEAN system where the x -axis is β and the y -axis is α . These two scenarios have identical transition probabilities and only differ in the average node degrees. Blue circles represent intervening strategies in Ω , and intervening strategies in Ψ are represented by red circles. For scenario A, the average node degree is 40, and the average node degree for scenario B is 4. The significant observation is that in scenario B with the average node degree 4, non-strong viral marketing ($\beta < 0.2$) could be compensated by strong media marketing ($\alpha > 0.4$) such that the final fraction of **negative** individuals becomes zero. However, in scenario A with the average node degree 40, no matter how strong is media marketing, when viral marketing is weak ($\beta < 0.2$), the final fraction of **negative** individuals becomes nonzero. Therefore, we conclude that for scenario A, the focus of the marketing campaigns must be viral marketing. On the other hand, we observe that for scenario B, if media marketing is strong enough ($\alpha > 0.4$), then no matter how strong the viral marketing is, the final fraction of **negative** individuals becomes zero. Therefore, the focus of marketing campaigns for scenario B should be media marketing.

Figure 13 shows Ω and Ψ computed by numerically solving the HET-MEAN system (32) corresponding to the Deezer network for two different marketing scenarios. The x -axes denote β which is representative of strength of viral marketing and the y -axes denote α which is representative of strength of media marketing. The two marketing scenarios only differ in the value of δ which impacts the effectiveness of media marketing. In particular, greater δ results in higher likelihood of transition of individuals from **negative** state to **positive** state because of media marketing. For scenario A with δ value of 0.1, we observe that viral marketing dominates media marketing. It is because no matter how strong is media marketing, whether the final

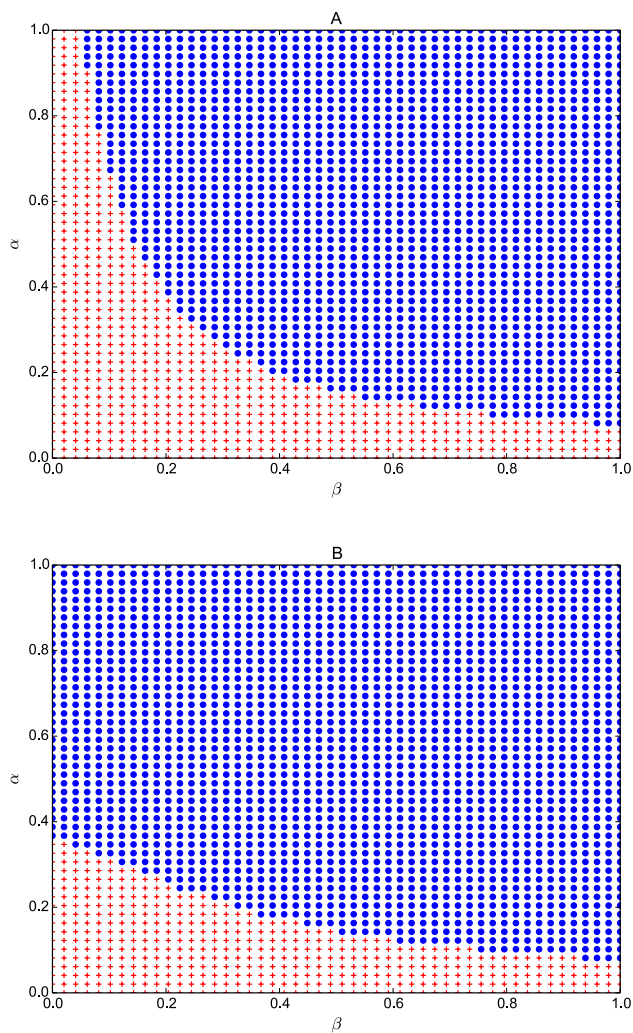


Fig. 12 Comparison between viral diffusion and media influence over the WS network. Blue circles represent intervening strategies in Ω (zero **negative** individuals), and intervening strategies in Ψ (nonzero **negative** individuals) are represented by red circles. x -axis is viral marketing β and y -axis is media marketing α . Parameter values: $\gamma = 0.5$, $\theta = 0.0001$, $\delta = 0.05$ and $\mu = 0.008$. top **a** $k = 40$ bottom **B** $k = 4$

fraction of **negative** individuals becomes zero or not, only depends on the strength of viral marketing. In particular, strong enough viral marketing ($\beta > 0.6$) guarantees final zero fraction of **negative** individuals. Therefore, for this marketing scenario, the focus of marketing campaign must be viral marketing and not media marketing. On the other hand, for scenario B with δ value of 0.5, the role of media marketing becomes more significant. In particular, weak viral marketing ($\beta < 0.6$) can be compensated with strong media marketing ($\alpha > 0.48$) while such a strategy was not achievable for scenario A.

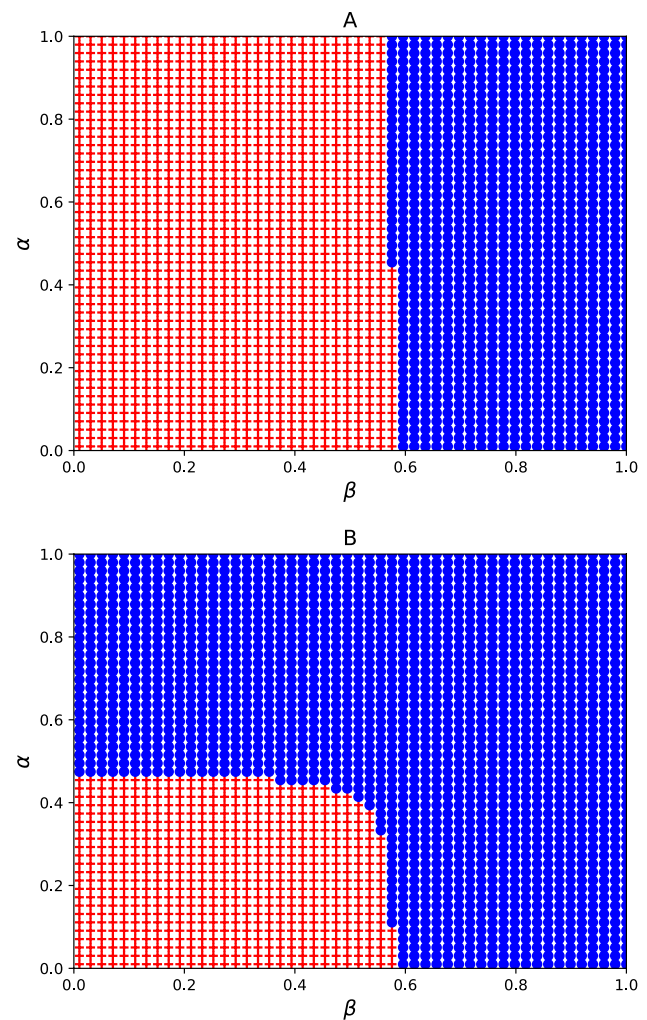


Fig. 13 Comparison between viral diffusion and media influence over the Deezer network. Blue circles represent intervening strategies in Ω (zero **negative** individuals), and intervening strategies in Ψ (nonzero **negative** individuals) are represented by red circles. x -axis is viral marketing β and y -axis is media marketing α . Parameter values: $\gamma = 0.001$, $\theta = 0.5$ and $\mu = 0.001$. top **a** $\delta = 0.1$ bottom **b** $\delta = 0.5$

6 Conclusion

In this paper, we propose the IN-STOCH system to model the adoption of polarized beliefs in influence networks, governed by both viral diffusion and media influence. We study the intertwined dynamics of these two different forms of influence and how they interact over influence networks. Using a mean-field approach, we derive the HOM-MEAN system to approximate the IN-STOCH system for exponential networks. Based on the stability of the equilibrium points of the HOM-MEAN system, we derive the conditions to guarantee the convergence of the fraction of **negative** individuals to 0. Also, we derive the HET-MEAN system to approximate the IN-STOCH system for scale-free networks and show that the HET-MEAN system always has exactly

one equilibrium point with no **negative** individuals. In particular, we apply our results for the HET-MEAN system to the BA networks and compute a simplified condition for the stability of the equilibrium point with no **negative** individuals for the BA networks. Finally, we validate our theoretical approximations by simulation of the IN-STOCH system over the following synthetic networks: WS, random regular, BA and a forest-fire network, as well as the following five real-world networks: ego-Facebook, Deezer, Livemocha, a Facebook interaction network and Douban.

Simulation results show that the HOM-MEAN system can accurately predict the convergence to zero for the fraction of **negative** individuals for WS and random regular networks. For cases where the final fractions of **negative** individuals are not 0, the analytical approximation given by the HOM-MEAN system is in good agreement with simulation results for WS and random regular networks. This agreement also improves when the average node degrees increase. In general, the analytical approximations show a higher accuracy for random regular networks compared to WS networks since random regular networks satisfy the underlying assumption of the HOM-MEAN system, all nodes have the same number of neighbors. Relying on our closed-form analytical formulas for the HOM-MEAN system, one can predict the final fraction of **negative** individuals with no computational cost for exponential networks no matter how large they are.

We approximate the dynamics of the IN-STOCH system for scale-free networks using the HET-MEAN system based on a mean-field approach. It is shown that the HET-MEAN system has always a single **negative**-free equilibrium point. Then, we state necessary and sufficient conditions for local stability of this equilibrium point as well as a sufficient condition for its global stability. Next, for Barabasi–Albert (BA) networks, we derive a sufficient condition to guarantee that the final fraction of **negative** individuals converges to 0.

The analytical approximations for scale-free networks by the HET-MEAN system are validated by simulations of the IN-STOCH system over different synthetic and real-world networks that demonstrate the characteristics of scale-free networks. In particular, we analyzed our analytical results over five real-world networks: ego-Facebook, Deezer, Livemocha, a Facebook interaction network and Douban as well as two synthetic networks: Barabasi–Albert (BA) and forest-fire. For BA network, it is shown that the analytical results obtained by studying the equilibrium point of the HET-MEAN system are in a good agreement with the simulation results. In particular, the HET-MEAN system can accurately predict the cases where the steady-state fractions of **negative** individuals are equal to 0. For the real-world networks and the forest-fire network, we show that numerically solving the HET-MEAN system using their node degree histograms can accurately predict the final

fraction of **negative** individuals with significant reduction in computational cost compared to directly simulation of the IN-STOCH system over these networks.

Finally, we rely on our analytical results for the HOM-MEAN and HET-MEAN systems to demonstrate how marketing campaigns can leverage the proposed framework to design their marketing strategies. In particular, using the HOM-MEAN system, we show that for two exponential networks with different degrees of connectivity, marketing campaigns need to shift their resources between viral marketing and media marketing to minimize the final number of individuals with **negative** state. Also, the HET-MEAN system is used to compare the effectiveness of viral marketing and media marketing for the Deezer network corresponding to two different marketing scenarios. We show that for the marketing scenario with smaller δ , the focus of marketing campaigns must be viral marketing and for the marketing scenario with greater value of δ , media marketing can be used to compensate for weak viral marketing.

Acknowledgements The contents of this paper were partially presented in Sanatkar (2016) with the following differences. This paper includes a new theorem that provides a sufficient condition for the global stability of the **negative**-free equilibrium point of the HET-MEAN system. Also, regarding the simulation results, different from Sanatkar (2016), initial fractions of **neutral**, **positive** and **negative** nodes are randomly drawn from the probability simplex for each network realization, and in order to make the initial states of the individuals realistic, their initial polarities are designed to depend on their neighbors using the Breadth First Search algorithm. Simulation results are extended for five new networks: the Deezer network, Livemocha network, a Facebook interaction network, the Douban network and the small-world forest-fire network. Moreover, for the study of the optimal resource allocation between viral marketing and media marketing, we present new results for the Deezer network.

Appendix

Proof of Theorem 3 If $(\bar{n}_2, 0)$ is locally exponentially stable, the Jacobian matrix A in (16), computed at $(\bar{n}_2, 0)$, is negative definite. We define \hat{n} and \hat{u} as follows:

$$\begin{aligned} \hat{n} &= n - \bar{n}_2, \\ \hat{u} &= u. \end{aligned} \tag{59}$$

Replacing n and u with \hat{n} and \hat{u} in (3), we can write

$$\begin{aligned} \dot{\hat{n}} &= k\beta\hat{n}^2 + (2k\beta n^* - (k\beta + \alpha + \gamma))\hat{n} + (k\beta - k\mu)\hat{n}\hat{u} \\ &\quad + ((k\beta - k\mu)n^* + \theta - \gamma)\hat{n}, \\ \dot{\hat{u}} &= k\mu\hat{n}\hat{u} + (k\mu n^* - \theta - \delta\alpha)\hat{u}. \end{aligned} \tag{60}$$

The origin of the transformed dynamical system in (60) is corresponding to $(\bar{n}_2, 0)$ for the HOM-MEAN system in (3). We use Lyapunov functions to estimate the region of attraction for the origin of the transformed dynamical system. A

Lyapunov function can be found by solving the Lyapunov equation

$$PA + A^T P = -Q, \tag{61}$$

where Q is a positive definite matrix. Taking $Q = I$, the unique solution is the positive definite matrix

$$P = \begin{bmatrix} p_1 & p_2 \\ p_2 & p_4 \end{bmatrix}, \tag{62}$$

where

$$\begin{aligned} p_1 &= -\frac{1}{2A_{11}}, \\ p_2 &= \frac{A_{12}}{2A_{11}(A_{11} + A_{22})}, \\ p_4 &= \frac{-A_{12}^2}{2A_{11}A_{22}(A_{11} + A_{22})}. \end{aligned} \tag{63}$$

The quadratic function $V(\hat{n}, \hat{u}) = p_1 \hat{n}^2 + 2p_2 \hat{n}\hat{u} + p_4 \hat{u}^2$ is a Lyapunov function for the transformed dynamical system in a certain neighborhood of the origin (Khalil 2000). To find the region of attraction of the origin, we need to determine a domain D around the origin where V is positive definite and V' is negative definite. Since P is positive definite, D can be determined by studying negative definiteness of V' around the origin. We can write

$$\begin{aligned} V'(\hat{n}, \hat{u}) &= -\hat{n}^2 + (2p_4 \hat{u} + 2p_2 \hat{n})k\mu \hat{n}\hat{u} \\ &\quad - \hat{u}^2 + (2p_1 \hat{n} + 2p_2 \hat{u})(k\beta \hat{n}^2 + (k\beta - k\mu)\hat{n}\hat{u}) \\ &= -\hat{n}^2 - \hat{u}^2 + f\hat{n}^3 + g\hat{n}^2 \hat{u} + h\hat{n}\hat{u}^2, \end{aligned} \tag{64}$$

where $f = 2p_1 k\beta$, $g = 2(p_1(k\beta - k\mu) + p_2(k\beta + k\mu))$ and $h = 2(p_2(k\beta - k\mu) + p_4 k\mu)$. Let $X := [\hat{n} \ \hat{u}]^T$. Hence, it can be written

$$\begin{aligned} V' &= -\|X\|^2 + f\hat{n}^3 + \hat{n}\hat{u}(g\hat{n} + h\hat{u}) \\ &= -\|X\|^2 + f\hat{n}^3 + \hat{n}\hat{u}[g \ h]X \\ &\leq -\|X\|^2 + f\|X\|^3 + \frac{1}{2}\|X\|^2 \sqrt{g^2 + h^2} \|X\| \\ &\leq -\|X\|^2 + (f + \frac{1}{2}\sqrt{g^2 + h^2})\|X\|^3. \end{aligned} \tag{65}$$

Therefore, V' is negative if $\|X\| < 1/(f + \frac{1}{2}\sqrt{g^2 + h^2})$. Let $r := 1/(f + \frac{1}{2}\sqrt{g^2 + h^2})$. And, we have

$$\eta < \min_{\|x\|=r} x^T P x = \lambda_{\min}(P)r^2 \tag{66}$$

Therefore, we can write

$$R_D = [0, 1]^2 \cap \hat{R}_D, \tag{67}$$

where

$$\hat{R}_D = \{(n, u) | p_1(n - n^*)^2 + 2p_2(n - n^*)u + p_4 u^2 < \eta\}. \tag{68}$$

□

References

Althoff T, Jindal P, Leskovec J (2017) Online actions with offline impact: how online social networks influence online and offline user behavior. In: Proceedings of the tenth ACM international conference on web search and data mining, Association for Computing Machinery, New York, NY, USA, pp 537–546. <https://doi.org/10.1145/3018661.3018672>

Amato R, Díaz-Guilera A, Kleineberg K-K (2017) Interplay between social influence and competitive strategical games in multiplex networks. *Sci Rep* 7:1

Aral S, Dhillon PS (2018) Social influence maximization under empirical influence models. *Nat Hum Behav* 2:375–382

Aral S, Walker D (2011) Creating social contagion through viral product design: a randomized trial of peer influence in networks. *Manag Sci* 57(9):1623–1639

Aral S, Walker D (2012) Identifying influential and susceptible members of social networks. *Science* 337(6092):337–341

Bakshy E, Hofman JM, Mason WA, Watts DJ (2011) Everyone’s an influencer: quantifying influence on twitter. In: Proceedings of the fourth ACM international conference on web search and data mining, WSDM ’11, New York, NY, USA, 2011. ACM, pp 65–74

Barabasi A-L, Albert R (1999) Emergence of scaling in random networks. *Science* 286(5439):509–512

Boguá M, Pastor-Satorras R, Vespignani A (2003) Epidemic spreading in complex networks with degree correlations. In: Romualdo P-S, Miguel R, Albert D-G (eds) *Statistical mechanics of complex networks*, vol 625. Lecture notes in physics. Berlin, Springer, pp 127–147

Brock WA, Durlauf SN (2001) Discrete choice with social interactions. *Rev Econ Stud* 68(2):235–260

Cha M, Haddadi H, Benevenuto F, Gummadi KP (2010) Measuring user influence in twitter: The million follower fallacy. In: ICWSM ’10: proceedings of international AAAI conference on weblogs and social

Cheng J, Adamic L, Dow PA, Kleinberg JM, Leskovec J (2014) Can cascades be predicted? In: Proceedings of the 23rd international conference on world wide web, WWW ’14, New York, NY, USA. ACM, pp 925–936

Cheng J, Kleinberg JM, Leskovec J, Liben-Nowell D, State B, Subbian K, Adamic LA (2018) Do diffusion protocols govern cascade growth? In: ICWSM

De Bruyn A, Lilien GL (2008) A multi-stage model of word-of-mouth influence through viral marketing. *Int J Res Mark* 25(3):151–163

Dearing JW (2008) Evolution of diffusion and dissemination theory. *J Public Health Manag Pract JPHMP* 14(2):99–108

Drossel B, Schwabl F (1992) Self-organized criticality in a forest-fire model. *Phys A Stat Mech Appl* 191(1):47–50

Du N, Song L, Gomez-Rodriguez M, Zha H (2013) Scalable influence estimation in continuous-time diffusion networks. In: Burges CJC, Bottou L, Welling M, Ghahramani Z, Weinberger KQ (eds) *Advances in neural information processing systems* 26. Curran Associates, Inc, Red Hook, pp 3147–3155

Farajtabar M, Du N, Gomez Rodriguez M, Valera I, Zha H, Song L (2014) Shaping social activity by incentivizing users. In: Burges CJC et al (eds) *Advances in neural information processing systems* 27. Curran Associates Inc, Red Hook

- Fink C, Schmidt A, Barash V, Cameron C, Macy M (2016) Complex contagions and the diffusion of popular Twitter hashtags in Nigeria. *Soc Netw Anal Min* 6:12
- Goel S, Watts DJ, Goldstein DG (2012) The structure of online diffusion networks. In: Proceedings of the 13th ACM conference on electronic commerce, EC '12, New York, NY, USA, 2012. ACM, pp 623–638. <https://doi.org/10.1145/2229012.2229058>
- Gomez RM, Leskovec J, Krause A (2010) Inferring networks of diffusion and influence. In: Proceedings of the 16th ACM SIGKDD international conference on knowledge discovery and data mining, KDD '10, New York, NY, USA. ACM, pp 1019–1028
- Gomez RM, Schölkopf B (2012) Influence maximization in continuous time diffusion networks. In: Proceedings of the 29th international conference on machine learning (ICML 2012), New York, NY, USA, July 2012. Omnipress, pp 313–320
- Guerra PHC, Meira Jr W, Cardie C, Kleinberg R (2013) A measure of polarization on social media networks based on community boundaries. In: Emre K, Nicole BE, Bernie H, Paul R, Ian S (eds) ICWSM. The AAAI Press
- Iyengar S, Lelkes Y, Levendusky M, Malhotra N, Westwood SJ (2019) The origins and consequences of affective polarization in the United States. *Annu Rev Polit Sci* 22(1):129–146
- Kempe D, Kleinberg J, Tardos É (2003) Maximizing the spread of influence through a social network. In: Proceedings of the ninth ACM SIGKDD international conference on knowledge discovery and data mining, KDD '03, New York, NY, USA, 2003. ACM, pp 137–146
- Khalil HK (2000) Nonlinear systems, third edn. Prentice-Hall, New York
- Kizilcec RF, Bakshy E, Eckles D, Burke M (2018) Social influence and reciprocity in online gift giving. In: Proceedings of the 2018 CHI conference on human factors in computing systems, CHI '18, New York, NY, USA, 2018. ACM, pp 126:1–126:11
- Kleineberg K-K, Boguñá M (2014) Evolution of the digital society reveals balance between viral and mass media influence. *Phys Rev X* 4:031046
- Kwak H, Lee C, Park H, Moon S (2010) What is twitter, a social network or a news media? In: Proceedings of the 19th international conference on world wide web, WWW '10, New York, NY, USA, 2010. ACM, pp 591–600
- Lazer DMJ, Baum MA, Benkler Y, Berinsky AJ, Greenhill KM, Menczer F, Metzger MJ, Nyhan B, Pennycook G, Rothschild D, Schudson M, Sloman SA, Sunstein CR, Thorson EA, Watts DJ, Zittrain JL (2018) The science of fake news. *Science* 359(6380):1094–1096
- Leskovec J, McAuley JJ (2012) Learning to discover social circles in ego networks. In: Pereira F, Burges CJC, Bottou L, Weinberger KQ (eds) Advances in neural information processing systems 25. Curran Associates, Inc, Red Hook, pp 539–547
- Leskovec J, Adamic LA, Huberman BA (2007) The dynamics of viral marketing. *ACM Trans Web I*(1):5
- Leskovec J, Singh A, Kleinberg J (2006) Patterns of influence in a recommendation network. In: Proceedings of the 10th Pacific-Asia conference on advances in knowledge discovery and data mining, PAKDD'06, Berlin, Heidelberg. Springer-Verlag, pp 380–389
- Margaris D, Vassilakis C, Georgiadis P (2016) Recommendation information diffusion in social networks considering user influence and semantics. *Soc Netw Anal Min* 6:12
- More JS, Lingam C (2019) A gradient-based methodology for optimizing time for influence diffusion in social networks. *Soc Netw Anal Min* 9:1–10
- Myers SA, Zhu C, Leskovec J (2012) Information diffusion and external influence in networks. In: Proceedings of the 18th ACM SIGKDD international conference on knowledge discovery and data mining, Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/2339530.2339540>
- Pastor-Satorras R, Vespignani A (2001) Epidemic dynamics and endemic states in complex networks. *Phys Rev E* 63:066117
- Ribeiro B (2014) Modeling and predicting the growth and death of membership-based websites. In: Proceedings of the 23rd international conference on world wide web, WWW '14, Republic and Canton of Geneva, Switzerland. International World Wide Web Conferences Steering Committee, pp 653–664
- Rozemberczki B, Davies R, Sarkar R, Sutton C (2019) Gemsec: graph embedding with self clustering. In: Proceedings of the 2019 IEEE/ACM international conference on advances in social networks analysis and mining 2019. ACM, pp 65–72
- Sahneh FD, Scoglio CM, Miegheem PV (2013) Generalized epidemic mean-field model for spreading processes over multilayer complex networks. *IEEE ACM Trans Netw* 21(5):1609–1620
- Sanatkar MR (2016) The dynamics of polarized beliefs in networks governed by viral diffusion and media influence. Ph.D. thesis
- Sanatkar MR, White WN, Natarajan B, Scoglio CM, Garrett KA (2015a) Epidemic threshold of an SIS model in dynamic switching networks. *IEEE Trans Syst Man Cybern Syst PP*(99):1–1
- Sanatkar MR, Scoglio C, Natarajan B, Isard SA, Garrett KA (2015b) History, epidemic evolution, and model burn-in for a network of annual invasion: soybean rust. *Phytopathology* 105(7):947–955
- Sarkar S, Guo R, Shakarian P (2017) Understanding and forecasting lifecycle events in information cascades. *Soc Netw Anal Min* 7:55
- Saxena B, Kumar P (2019) A node activity and connectivity-based model for influence maximization in social networks. *Soc Netw Anal Min* 9:12
- Shi F, Teplitskiy M, Duede E, Evans J (2017) The wisdom of polarized crowds. *Nat Hum Behav* 3:329–336
- Sinan A, Dean E (2019) Protecting elections from social media manipulation. *Science* 365:858–861
- Sun E, Rosenn I, Marlow C, Lento T (2009) Gesundheit! modeling contagion through Facebook news feed. In: International AAAI conference on weblogs and social media
- Van Mieghem P, Omic J, Kooij R (2009) Virus spread in networks. *IEEE ACM Trans Netw* 17(1):1–14
- Viswanath B, Mislove A, Cha M, Gummadi KP (2009) On the evolution of user interaction in Facebook. In: Proceedings of the workshop on online social networks, pp 37–42
- Vosoughi S, Roy D, Aral S (2018) The spread of true and false news online. *Science* 359:1146–1151
- Watts DJ, Dodds PS (2007) Influentials, networks, and public opinion formation. *J Consum Res* 34(4):441–458
- Watts DJ, Peretti J (2007) Viral marketing for the real world. Harvard Business Review, Boston
- Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. *Nature* 393(6684):409–10
- Yardi S, Boyd D (2010) Dynamic debates: an analysis of group polarization over time on twitter. *Bull Sci Technol Soc* 30(5):316–327
- Zafarani R, Liu H (2009) Social computing data repository at ASU. Tempe, AZ: Arizona State University, School of Computing, Informatics and Decision Systems Engineering. <http://socialcomputing.asu.edu>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.