

# Co-Virality of Competing Content over OSNs?

Khushboo Agarwal and Veeraruna Kavitha

{agarwal.khushboo, vkavitha}@iitb.ac.in, IEOR, IIT Bombay, Mumbai 400076, India

**Abstract**—The marked increase in advertisements over online social networks (OSNs) necessitates the study of content propagation. We analyse the viral markets with content providers competing for the propagation of similar posts over OSNs. Towards this, we required a new variant of the branching process (BP), which we named as “Branching process with attack”; the entities upon wake up attempt to attack and acquire the opposite population; furthermore, each entity produces its offsprings as is usually considered in BPs. In addition to providing expressions for the growth rates of individual posts, dichotomy etc., we explore the co-existence/co-virality of posts; can the competing content spread and explode (number of unread/live copies of both posts grow significantly with time) simultaneously over the network? We prove that either one or both populations/posts get extinct or the populations settle to a unique co-existence equilibrium and derive the corresponding asymptotic ratios of the two populations/posts. Our analysis applies to large population networks focusing on mass behaviour, rather than micro details. Our study provides insights into two crucial design aspects, the number of seed users and the quality of the post. We performed Monte-Carlo simulations on synthetic and ego-twitter dataset by SNAP to support our findings.

**Index Terms**—Viral Competing Markets, Branching process, Attack

## I. INTRODUCTION

Social media is universally prevalent today; the shared content is shared (again), liked or dis-liked by the users and thus may get viral over the network. This gives the content providers (CPs) an opportunity to share the product information with customers and cheaply amplify their growth; a strategy well known as “Viral Marketing” (e.g., [6], [7], [13]). Variety of content propagates simultaneously through such platforms, and enjoy the resultant benefits; however, they face strict competition due to competing contents.

In [6], a timeline structure holding the content is considered to study the competing content dissemination over online social networks (OSNs). Here the competition was due to the placement of competing content at various levels of the timeline, but the paper did not consider the aspect that a user would choose one among the competing contents. When a user, already shared with one content (say content A), is also shared with another content (say content B), then both the posts are available on its timeline. If the posts are in direct competition, user may choose the latter post. Thus one can say that ‘content A is attacked and acquired by content B’. We study such systems and refer them as *viral competing markets*. Each CP has to decide on two factors: (i) the initial number

of users (called seed users) for whom it may have to pay to share its content; (ii) the quality of its post. Some of the key questions that we explore are:

- 1) At what growth rate does each content propagate?
- 2) With what probability does a given content get extinct?
- 3) Can the posts of competitors co-exist over the network?
- 4) How should a CP trade-off between investing on seed users and designing an attractive post?

Some of these questions were investigated previously (e.g., [2], [5], [12], [14]), but the presence of competition poses new challenges and provides a new outlook towards viral markets.

*Our approach and contributions:* There are several approaches for studying content propagation over the network. A set of literature focuses on micro details, like degree and particularity (basically, connections) of the users in the network for their designs and uses tools like random graphs (e.g., [5], [14]) or epidemiology based models (e.g. [2], [12]) for the study. However, these models can not capture some important aspects related to content propagation like virality, i.e., the explosion of (the number of shares of) content over the network. Furthermore, a post can witness a huge surge in the shares in a short duration, which can only be captured by continuous-time BPs. *We consider large population networks, where all users are typically the same (in stochastic sense), and focus on macro-details related to content propagation.* Branching processes are a great choice for such a study (e.g., [6], [7], [13]). It provides analysis for growth patterns, extinction probability and other measures, and captures phenomenon which are specific to content propagation and virality. Some existing research considers competing contents over the network using techniques other than BPs [2], [5], [12]; however, those models are inadequate for studying *co-virality of the posts*, i.e., concurrent explosion (the number of copies of both the posts exploding simultaneously over the network) of the posts. We aim to understand these phenomenon.

There is a vast literature that studies variants of BPs (e.g., [1], [4], [8]). The prey-predator type BP [4] is the only BP that can capture some aspects of competition. Here, one population (predator) attacks, while the other population (prey) progresses like a standard BP. They do not consider ‘double-sided attack’, ‘acquisition’ and ‘size based attacks’; but, such aspects are required to capture ‘any post attacking and acquiring any other post’. Further they consider discrete-time models, while the users visit the OSN asynchronously resulting in continuous-time evolution. Thus, the existing BPs are insufficient to mimic critical aspects of viral competing markets.

We introduce a new variant of Multi-type branching process, named as *Branching process with attack* (BPA). In standard

The work of first author is partially supported by Prime Minister’s Research Fellowship (PMRF), India.

BP, upon wake up, each entity produces a random number of its own offsprings before dying; in our model, it also attacks and captures the opposite population. We study the continuous-time population-size dependent and Markovian variant. The unread copies of each post are modelled as one population.

We contribute towards both branching process and viral marketing literature. In addition to studying growth rates, dichotomy etc., we provide a distinct result in contrast to “winner takes it all” (studied in [5], [12] and others). We prove using *stochastic approximation techniques* that either one or both posts get extinct, or the limit proportions converge to a unique fixed ratio which is a function of mean offsprings and attack capacity. The last possibility is a rare phenomenon and occurs when mean number of offsprings and/or initial population is high (as in large OSNs). But, on ego-twitter dataset co-existence is observed even when started with moderate seed users in significant number of sample runs.

By one of our results, any CP could capture the market if it invests more on seed users than on designing an attractive post, even if it has low market share/credibility. One such recent example is *the attack over WhatsApp by Signal*. The tweet ‘Use Signal’ by Elon Musk (with 42.6 million followers) provided equivalent number of seed users. This led to massive surges in downloads (nearly 1 million downloads a day) for Signal, even in the presence of market giants like WhatsApp.

## II. VIRAL COMPETING MARKETS

We consider content propagation over online social networks by two content providers (CPs) competing for similar kind of product/service. The main goal of each CP is to reach out to more audience through viral marketing; a certain fraction of the users receiving the post may provide business to the CP. Each CP initially posts its content to few selected users and these are referred to as seed users. Whenever a receiver visits the application (referred to as OSN, Online social network) over which the post has been shared, it views the post. We refer the time instance at which the user views the post as its “wake-up” time. After viewing, the user shares it with some or all of its contacts/friends. The number of shares depend upon the extent to which the user appreciates the post. Our main aim is to analyze *the propagation of such competing content over OSNs*.

### A. Competing content to Branching process:

As is usually done in viral marketing literature (e.g., [6], [7], [13]), we are modeling this propagation using an appropriate branching process (BP). We detail the dynamics of the problem, along side, describing the BP based modeling details. Towards the end, we would observe that the branching process modeling the propagation of competing content is very different from the models considered in the literature.

- We refer the two CPs respectively by  $x$  and  $y$  type CPs. If the user receives only one kind of post (say of  $x$ -CP), at the time it views the post/it wakes-up, we refer it as  $x$ -post. In this case, the user would obviously share (if at all) only  $x$ -post to its friends. Once viewed, the user would not read the post again, we then say that the number of unread  $x$ -users/posts

have reduced by one, and are increased by the number of shares; this is exactly like a death in an appropriate BP, after producing random number of offsprings (shares here).

- If the user receives both the posts at its wake-up, (due to competing nature of the two posts) it may chose one among the two to share; we refer such a post/user as  $x$  or  $y$ -post/user respectively depending upon its choice. It is also possible that it shares to none (zero offsprings). Initially the user must have been shared (say)  $x$ -post, we then refer it as  $x$ -user. But when another friend forwards the  $y$ -post before the user wakes-up, then the user gets converted to  $y$ -user if it prefers  $y$ -post over the  $x$ -post at it’s wake-up. *It is this conversion, which we refer as attack (by  $y$ -user)*, that makes the underlying branching process very different from the ones studied in the literature. As before the corresponding unread posts are reduced by one, once the user views (and shares) them.

- If the user is shared with multiple copies of the same post, then the user would consider only the latest share.

The posts propagate from one user to another in the above described manner. As mentioned before, there is a possibility of the rival post reaching the user before it views the first received post. If the user prefers to share the rival post, then one can say that ‘the rival post attacked and acquired the first post’; we name such a process as *branching process with attack (BPA)*. We analyse these novel BPs, along with viral competing markets.

### B. Modelling details

Let  $c_0^x, c_0^y$  be the number of seed users of  $x, y$  posts chosen by the respective CPs. We assume that any user wakes-up after an exponentially distributed time with parameter  $\lambda$  (i.e, we consider Markovian BPs); *this makes the model mathematically tractable and is a reasonable approximation to model scenario like ours* (see [6], [13]). Let the number of friends ( $F$ ) of a typical user be random which is independently and identically distributed (IID) across the users in the network;  $m_f := E[F]$ . Let  $Z_t := (C_t^x, C_t^y)$  be the number of current live/unread copies of respective posts at time  $t$ .

Upon wake up, a user shares the post (of say  $x$ -type) to its friends with probability  $\eta_x$  depending upon the quality<sup>1</sup> of the post. The  $x$ -CP achieves desired  $\eta_x$  by appropriate allocation of resources towards designing an attractive post. In all, we assume that a user with  $x$ -post ( $y$ -post) shares to  $Bin(F, \eta_x)$  (resp.  $Bin(F, \eta_y)$ ) of its friends, where  $Bin(\cdot, \cdot)$  is a Binomial random variable. Among these friends to whom the post is to be shared by  $x$ -user, there will be 3 fractions:

- a fraction of users who had not received any post yet, these would correspond to new live copies of  $x$ -post,  $\xi_x$ ;
- another fraction that already received the  $y$ -post, these would correspond to the attacked population,  $\xi_{xy}$ ; and,
- a third fraction who already received the  $x$ -post and will not be interested in another copy of it.

We assume that for any post type (say  $x$ -post) and any time  $t$ ,  $\gamma(C_t^x)$  fraction of users will have  $x$ -post, among any given

<sup>1</sup>it reflects the enticing nature of the post and the urge that the users feel to like/forward the post; could depend on its design, irresistible offers etc.

subset of the network; further, we assume  $\gamma(\cdot)$  is an increasing function, which converges to  $\gamma^*$ . This modeling is reasonable for huge networks, as the fraction of users with a given post is incomparable with the network size.

In all, when a  $x$ -user forwards, the number of shares to a subset of its friends without any ( $x$  or  $y$ ) post is given by,

$$\xi_x(Z_t) := \text{Bin}(F, \eta_x(1 - \gamma(C_t^x) - \gamma(C_t^y))),$$

while the number of shares to that subset of its friends with  $y$ -post,  $\xi_{xy}(C_t^y) := \text{Bin}(F, \eta_x \gamma(C_t^y))$ . Among the latter fraction (that received both the posts), only  $p_{xy}$  sub-fraction prefers the new ( $x$ ) post; this preference of the user changes user's type to  $x$ -type and this is *exactly equivalent to an attack over  $y$ -posts by  $x$ -post*; the number of (successfully) attacked  $y$ -posts are:

$$\zeta_{xy}(C_t^y) := \text{Bin}(\min\{\xi_{xy}(C_t^y), C_t^y\}, p_{xy}). \quad (1)$$

Note that *the successfully attacked  $y$ -posts are acquired by the  $x$ -type*. Let  $z = (c^x, c^y)$  be a realisation of the random vector  $Z = (C^x, C^y)$ . The expected values of the *population size dependent attacks and offsprings* and their limits are defined by (where  $z \rightarrow \infty$  denotes  $c^x \rightarrow \infty$  and  $c^y \rightarrow \infty$ ):

$$\begin{aligned} m_x(z) &:= E[\xi_x(z)] = m_f \eta_x (1 - \gamma(c^x) - \gamma(c^y)), & (2) \\ m_x^* &:= \lim_{z \rightarrow \infty} m_x(z) = m_f \eta_x (1 - 2\gamma^*) \\ m_{xy}(c^y) &:= E[\zeta_{xy}(c^y)] = E[\min\{\text{Bin}(F, \eta_x \gamma(c^y)), c^y\}] p_{xy}, \text{ and,} \\ m_{xy}^* &:= \lim_{c^y \rightarrow \infty} m_{xy}(c^y) = m_f \eta_x \gamma^* p_{xy}. \end{aligned}$$

When a user with  $y$ -post shares, we have analogous dynamics. *The network structure is usually the same towards any post; hence  $\gamma(\cdot)$  is same for both. It is  $\eta_x, \eta_y$  (depends upon the design of posts by CPs),  $p_{xy}, p_{yx}$  (depends upon the CPs and network, details later) and  $c_0^x, c_0^y$  (seed users chosen by CPs) that distinguishes the course of propagation of the two posts.*

Observe that the offsprings and ‘attacks’,  $\{\xi_x(Z_t)\}$  and  $\{\xi_{xy}(C_t^y)\}$ , are IID across those  $x$ -wake up epochs for which  $Z_t = z$  for a fixed  $z$ ; similar is the case with  $\{\xi_y(Z_t)\}$  and  $\{\xi_{yx}(C_t^x)\}$ . We further assume *super critical* conditions and finite first moments:

**A.1** For each  $i, j \in \{x, y\}$  with  $i \neq j$  and any  $z$ :

$$1 < m_i(z), m_i^* < \infty, \text{ and } m_{ij}(c^j), m_{ij}^* < \infty.$$

Under super-critical conditions, both the posts can get viral in the absence of competition, i.e., when they progress independently (see [1]). Further, we assume that  $0 < P(\xi_i(z) = 0) < 1$ , for each  $i \in \{x, y\}$  and any  $z$ , thus ensuring a possibility of zero as well as higher number of shares.

### C. Dynamics:

We analyse the system by studying the number of copies of the posts at transition epochs, i.e., at time instances at which a user wakes-up. We begin with some notations. Let us denote the event that  $x$ -user wakes up as  $x \uparrow$  and similarly define  $y \uparrow$ . Let  $\tau_n$  represent the  $n^{\text{th}}$  transition epoch<sup>2</sup>. Let  $C_n^x := C_{\tau_n}^x = \lim_{t \downarrow \tau_n} C_t^x$  be the number of  $x$ -posts immediately after  $\tau_n$ . Similarly define  $C_n^y$ . Note that the *time taken by the first user to wake-up*,  $(\tau_{n+1} - \tau_n)$ , after the  $n^{\text{th}}$  transition epoch

<sup>2</sup>If both posts get extinct at  $n^{\text{th}}$  epoch, we set  $\tau_k := \tau_n$  for all  $k \geq n$ .

is exponentially distributed with parameter  $\lambda S_n$ , with the sum  $S_n := C_n^x + C_n^y$  representing the total number of unread copies on OSN corresponding to the two posts. One can summarize the system (and BPA) using the following description of the events at transition epochs. When a  $x$ -user wakes up ( $x \uparrow$ ), it shares to  $\xi_x$  number of its friends, attacks  $\zeta_{xy}$  number of  $y$ -type and dies<sup>3</sup>, i.e., (see (1)),

$$\begin{aligned} C_n^x &= C_{n-1}^x + \tilde{\xi}_n^x(Z_{n-1}) + \zeta_{xy,n}(C_{n-1}^y), \\ C_n^y &= C_{n-1}^y - \zeta_{xy,n}(C_{n-1}^y), \quad \tilde{\xi}_n^x(Z_{n-1}) := \xi_n^x(Z_{n-1}) - 1. \end{aligned} \quad (3)$$

We have similar transitions when  $y \uparrow$ , with  $\tilde{\xi}_n^y(Z_{n-1}) := \xi_n^y(Z_{n-1}) - 1$ . Observe that the (successfully) attacked posts are acquired by the attacking type.

**Some more modeling aspects:** In social networks typically the *posts are stored in timelines* (e.g., as in [6]), the inverse stacks on the user's application interface. When a post is shared to the user, this post sits on the top of the timeline and all other posts shift down by one position. Thus, the competing posts have the potential to attack each other; when a user has both the posts it can prefer the newer post (one on higher levels) to a bigger extent (and/or the interest may depend upon the source of the post). *If the user always prefers the newer post*, this can be modelled by considering  $p_{xy} = p_{yx} = 1$ . On the other hand, if one of the CPs (say  $x$ ) is more influential than the other, we can model this by  $p_{xy} > p_{yx}$ ; *in this case, the user may prefer the  $x$ -post even when it is placed at a lower level*. As already mentioned, one (say  $x$ -CP) designing a more attractive post can be captured by  $\eta_x > \eta_y$ .

In short, each CP can control the propagation of its post by appropriate choice of  $c_0^x/c_0^y$  (seed users) and  $\eta_x/\eta_y$  (attractiveness of the post), for given influence factors  $p_{xy}$  and  $p_{yx}$ ; we investigate the significance of these parameters.

**Independent Propagation:** If one models simultaneous propagation of the competing content without considering attack, we will have the two posts propagating independent of each other ( $p_{xy} = p_{yx} = 0$ ). We get,  $m_x = m_f \eta_x$  and  $m_y = m_f \eta_y$ ; basically *the existence of any post (represented by  $\gamma(\cdot)$ ) does not alter the prospects of any other post*. This can be modelled using two independent single type BPs, which we refer briefly as BPNAs (no attack). *We compare the conclusions drawn by such independent processes with those obtained using our BPA model below, to illustrate the drawbacks of the formal models to study viral competing markets*. We provide numerical comparison using synthetic and ego-twitter dataset (in section IV).

## III. ANALYSIS

### A. Disparity in the number of seed users

Consider a scenario with huge disparity in seed users of the two CPs. We immediately have (proof in Appendix):

**Theorem 1: [Seed users]** Assume  $m_{xy}(1) > 0$  and A.1. For any  $\epsilon > 0$  and  $c_0^y$ , there exists a  $\bar{c}^x < \infty$  such that

$$P(C_n^y \xrightarrow{n \rightarrow \infty} 0 | Z_0 = (c_0^x, c_0^y)) > 1 - \epsilon \text{ for all } c_0^x \geq \bar{c}^x.$$

<sup>3</sup>the number of unread copies of  $x$ -post decreases by 1.

The result is true even when  $x$  and  $y$  are interchanged. ■  
Thus *the CP with higher seed users eventually captures the market with high probability*. This is true irrespective of the influence factors ( $p_{xy}, p_{yx}$ ) and the relative attractiveness of the posts ( $\eta_x, \eta_y$ ). In fact, all it requires is that the CP with higher seed users (say  $x$ -CP) can attack, i.e.,  $p_{xy} > 0$ . In other words, *even the post of the market giant ( $p_{yx} \gg p_{xy} > 0$ ) can get extinct with high probability, if the smaller CP uses exorbitantly large number of seed users ( $c_0^x \gg c_0^y$ )*.

On the contrary, if one models viral competing markets using *BPNA*s, by independence, *the number of seed users of any CP has no influence on the propagation of the post of the other CP; this leads to misleading conclusions* (section IV).

From Theorem 1, we can also conclude that if a CP (say  $x$ -type) with some influence ( $p_{xy} > 0$ , i.e., at least few users with both the posts on their timeline prefer this post) invests more in sharing its post to large number of seed users, instead of investing on designing a better post (i.e.,  $\eta_x$ ), then irrespective of the the quality of the other post ( $y$ -type),  $x$ -CP is always in a better position. In fact, if the  $y$ -CP shares its post to lesser number of seed users, then there is a high possibility that  $x$ -post reaches more users and  $x$ -CP can capture the market. Thus, *it is more crucial to procure higher seed users*, which is not aptly indicated by *BPNA* based models.

### B. Limit proportions

Transient analysis (study of growth patterns, limit proportions etc.) is an important aspect for Markov chains that are predominantly transient, like BPs under super-critical regime. It is a common practice to scale the process appropriately that enables convergence to a finite limit, to understand the otherwise transient, exploding process. We consider a very different type of scaling ( $\Theta_n$  defined below) and adopt a new approach using *stochastic approximation* (SA) techniques (e.g. [10]) to derive (time) limit of the proportion  $\beta_t := C_t^x/S_t$  of the two posts. *These limits shed light on the fractional occupancy of the competing posts over the network after passage of sufficiently long time and co-virality of the competing posts*.

Let  $\S_n$  represent the sample mean formed by the sequence of offsprings/shares plus the initial number of posts (see (3)):

$$\S_n = \frac{1}{n} \left( \sum_{k=1}^n \tilde{\xi}_k + c_0^x + c_0^y \right), \quad \tilde{\xi}_k := \tilde{\xi}_k^x H_k + \tilde{\xi}_k^y H_k^c, \quad (4)$$

where  $H_k = 1 - H_k^c$  is the indicator that  $x \uparrow$  at  $k^{th}$ -epoch; *for simpler notations, we omit the dependence on  $Z_n, C_n^x, C_n^y$  at places*. Observe that the total population ( $\nu_e$  extinction epoch),

$$S_n = n \S_n 1_{\{n < \nu_e\}}, \quad \nu_e := \inf\{n : S_n = 0\}. \quad (5)$$

Further, also observe (same is the case for  $C_n^y$ ):

$$C_n^x \leq S_n \leq n|\S_n| \text{ for all } n. \quad (6)$$

If the process was population independent, for example if  $\gamma(\cdot) \equiv \gamma$  and  $\eta_x = \eta_y$ , by law of large numbers,  $\S_n \rightarrow m_f \eta_x - 1$  a.s., while  $\psi_n := S_n/n \rightarrow m_f \eta_x - 1$  only in survival sample paths. For general case, one can construct upper bounding processes which exhibit similar property. These

observations form the main basis for analysing the limits of proportions  $\psi_n$  and  $\theta_n := C_n^x/n$ . Observe  $\beta_n = \theta_n/\psi_n$  and let  $\Theta_n := [\psi_n, \theta_n, t_n]$ ,  $\Theta_0 := [s_0, c_0^x, 0]$ , with  $t_n := \sum_{k=1}^n 1/k$ .

The evolution of  $\Theta_n$  can be captured by a 3-dimensional stochastic approximation based scheme as below (see (3)):

$$\begin{aligned} \psi_n &= \psi_{n-1} + \epsilon_n \left( \tilde{\xi}_n - \psi_{n-1} \right) K_n, \\ \theta_n &= \theta_{n-1} + \epsilon_n \left[ H_n \left( \tilde{\xi}_n^x + \zeta_{xy,n} I_n \right) - H_n^c \zeta_{yx,n} I_n - \theta_{n-1} \right] J_n, \\ t_n &= t_{n-1} + \epsilon_n, \text{ with } \epsilon_n := 1/n, \quad I_n := 1_{\{\theta_{n-1} < \psi_{n-1}\}}, \\ J_n &:= 1_{\{\theta_{n-1} > 0\}}, \text{ and } K_n = 1_{\{\psi_{n-1} > 0\}}. \end{aligned} \quad (7)$$

Observe that  $C_n^x = \eta(t_n)\theta_n$ ,  $C_n^y = (\psi_n - \theta_n)\eta(t_n)$ , where  $\eta(t) := \max\{n : t_n \leq t\}$ . The ODE that can approximate (7) is given by (see [10], and Appendix for more details):

$$\begin{aligned} \dot{\Theta} &= \bar{g}(\Theta) \text{ with } \bar{g}_t(\Theta) = 1, \\ \bar{g}_\psi(\Theta) &= \left[ \frac{\theta}{\psi} m_x(z) + \left( 1 - \frac{\theta}{\psi} \right) m_y(z) - 1 - \psi \right] 1_{\{\psi > 0\}}, \\ \bar{g}_\theta(\Theta) &= \left[ \frac{\theta}{\psi} \left( m_x(z) - 1 + m_{xy}(c^y) 1_{\{\theta < \psi\}} \right) \right. \\ &\quad \left. - \left( 1 - \frac{\theta}{\psi} \right) m_{yx}(c^x) 1_{\{\theta < \psi\}} - \theta \right] 1_{\{\theta > 0\}}, \text{ where,} \\ z(t) &:= (c^x(t), c^y(t)), \quad c^x(t) := \theta(t)\eta(t), \quad c^y(t) := (\psi - \theta)(t)\eta(t). \end{aligned} \quad (8)$$

We will prove that the ODE indeed approximates (7) and derive further results using [10, Theorem 2.2, pp. 131] in Theorem 2 (given below); *the ODE approximation is derived for the general case as in the hypothesis, however the analysis of the ODE (Lemma 1 in Appendix) is valid only for the special population-size independent case ( $\eta_x = \eta_y$  and  $\gamma(\cdot) \equiv \gamma$ , where  $\gamma < \infty$ )*. For general case, the Theorem is only a conjecture and we are working towards completing the proof.

Basically we provide the required justifications, identify the attractors and the domain of attraction of the ODE and finally derive the following result (proof in Appendix) under the additional assumptions:

**A.2** We assume finite second moments, i.e., for each  $z$  and  $i, j \in \{x, y\}, i \neq j$ :  $E(\xi_i^2(z)) < \infty$ ,  $E(\xi_{ij}^2(c^j)) < \infty$ .

**A.3** For some finite  $\bar{c}^y, \bar{c}^x, \kappa_{ij}$ , for  $i, j \in \{x, y\}, i \neq j$ ,

$$\kappa_{ij} \min\{c^j, \bar{c}^j\} \leq m_{ij}(c^j) \leq \bar{c}^j \kappa_{ij} = m_{ij}^*,$$

i.e.,  $m_{xy}(\cdot), m_{yx}(\cdot)$  are bounded by piece-wise linear functions.

**Theorem 2: [Limit Proportions]** Assume A.1-A.3. The sequence  $(\psi_n, \theta_n)$  converges a.s. to one of the following limits:

- (i)  $(0, 0)$ , i.e., both population types get extinct,
- (ii)  $(m_y^* - 1, 0)$ , i.e., only  $y$ -population survives,
- (iii)  $(m_x^* - 1, m_x^* - 1)$ , i.e., only  $x$ -population survives, or
- (iv)  $(\psi^\circ, \theta^\circ)$ , with  $\theta^\circ = \psi^\circ \beta^\circ$  and  $\psi^\circ = \beta^\circ m_x^* + (1 - \beta^\circ) m_y^* - 1$ , i.e., both populations co-exist, where

$$\beta^\circ = \begin{cases} \frac{m_{yx}^*}{m_{xy}^* + m_{yx}^*} & \text{when } m_x^* = m_y^*, \text{ and else,} \\ \frac{1}{2} + \frac{m_{xy}^* + m_{yx}^* - \sqrt{(m_x^* - m_y^* + m_{xy}^* - m_{yx}^*)^2 + 4m_{xy}^* m_{yx}^*}}{2(m_x^* - m_y^*)} & \end{cases} \quad (9)$$

**Remarks:** • One can easily prove that the probability of the events (i)-(iii) in the above are non-zero. *The probability of*

co-virality, that of the event (iv) is a more interesting question which is considered in the next subsection for finite time.

- After simple algebra using (9), one can show that  $\beta^\circ > 1/2$  if and only if  $m_x + (m_{xy}^* - m_{yx}^*) < m_y + (m_{yx}^* - m_{xy}^*)$ , i.e., if and only if the combined reproduction capacity ( $m_x^*$ ) and net attack capacity ( $m_{xy}^* - m_{yx}^*$ ) of  $x$ -type is smaller than that of  $y$ -post. Thus surprisingly a ‘weaker post’ occupies a higher proportion at limit. But, when viewed from a different perspective, for the two posts to co-exist, there should be some sort of balance and that is possible only if *the proportion of the post with higher potential is smaller than that with the lower potential*.

### C. Finite-time Co-existence

Having investigated the possible values of limit proportions, we proceed to explore if at all the two posts can co-exist over the network with non-zero probability. Towards this, we restrict our attention to finite time analysis, and asymptotic analysis is for future. We have (proof in Appendix):

**Theorem 3: [Finite-time Co-existence]** Define the two types of  $(\beta)$  neighbourhoods of  $\beta^\circ$  of (9) respectively by,

$$\mathcal{N}_\epsilon := \{z : \beta \in [\beta^\circ - \epsilon, \beta^\circ + \epsilon]\} \text{ and } \mathcal{N}_{\epsilon, \Delta} := \mathcal{N}_\epsilon \cap \{s = \Delta\}.$$

Define exit time for any set  $A$ ,  $\Gamma_A := \inf_t \{Z_t \in A^c\}$  (infimum of empty set is  $\infty$ ). Then, for any  $\tilde{\epsilon} > \epsilon > 0$  and any  $T < \infty$ :  $\inf_{z \in \mathcal{N}_{\epsilon, \Delta}} P_z(\Gamma_{\mathcal{N}_\epsilon} \leq T) \rightarrow 0$  as  $\Delta \rightarrow \infty$ . ■

If the two competing posts start/reach in a ratio close to  $\beta^\circ$  and in sufficiently large numbers, then their proportions remain close to  $\beta^\circ$  in any finite time, with high probability. *Thus, we can have/continue co-existence or simultaneous propagation of the posts on the network for any finite time.* We verify the same through simulations (see Table IV in section IV).

**Co-virality of competing posts:** In the sample paths where  $\psi_n \rightarrow 0$ , both posts get extinct. Consider the survival sample paths (i.e.,  $\psi_n = S_n/n \not\rightarrow 0$ ) in Theorem 2, then  $S_n \rightarrow \infty$ . This leads either to: (a) high disparity in the number of two posts (as suggested by Theorem 2.ii-iii); or (b) the limit proportion converging to  $\beta^\circ$  (by Theorem 2.iv). In the first case, by Theorem 1, the post with lower number of unread copies gets extinct (and not just fraction  $C_n^x/n \rightarrow 0$ ) and hence co-virality is not possible. In the latter scenario, both the posts grow exponentially large (see Theorem 4 in next subsection) over the network while remaining in a proportion close to  $\beta^\circ$ . *This is the co-virality event. However, at present, we do not have an estimate of the probability of such an event over infinite time.* Nevertheless, Theorem 3 confirms continuation of co-virality event over finite-time horizons with high probability, if and when it reaches a neighborhood of  $\beta^\circ$ .

### D. Growth Patterns and Extinction probability

In this subsection we analyze the growth patterns of the evolution of the competing posts. We utilize the well-known results of the BPs to derive the same. Without loss of generality consider  $\eta_x \geq \eta_y$ . Define the following upper and lower rates:

$$\underline{\alpha} := \lambda(m_y^* - 1) \text{ and } \bar{\alpha} := \lambda(m_x(0, 0) - 1). \quad (10)$$

We have the following results (proof in Appendix):

**Theorem 4: [Growth and extinction]** Assume A.1-A.2 and  $\eta_x \geq \eta_y$ . Then:

(i) the growth rate of  $\{S_n\}$  process is upper bounded by  $\bar{\alpha}$  and lower bounded by  $\underline{\alpha}$  as below:

$$P(\limsup_n S_n e^{-\bar{\alpha}\tau_n} < \infty) = 1, \quad P(\liminf_n S_n e^{-\underline{\alpha}\tau_n} > 0) > 0.$$

(ii) the result is also true for  $x$  as well as  $y$ -post propagation processes, i.e., for example,

$$P(\limsup_n C_n^x e^{-\bar{\alpha}\tau_n} < \infty) = 1 \text{ and } P(\liminf_n C_n^x e^{-\underline{\alpha}\tau_n} > 0) > 0$$

(iii) the extinction probability  $q := \sqrt{P_{z_0=(1,1)}(S_n \rightarrow 0)}$  is bounded between,  $\underline{q} \geq q \geq \bar{q}$ , with

$$\underline{q} := \sqrt{P_{(1,1)}(\underline{S}_n \rightarrow 0)}, \quad \bar{q} := \sqrt{P_{(1,1)}(\bar{S}_n \rightarrow 0)},$$

where  $\underline{q}, \bar{q}$  respectively satisfy  $\underline{f}(s) = s$  and  $\bar{f}(s) = s$ , with  $\underline{f}$  and  $\bar{f}$  being PGFs of random variables  $\tilde{\xi}_y^* := \lim_{z \rightarrow \infty} \tilde{\xi}_y^*(z)$  a.s. and  $\tilde{\xi}^x(0, 0)$  respectively. ■

The significance of such theorems is well understood in BP literature, and we explain the same in our context:

(a) Let  $\bar{W}_x := \limsup_n C_n^x e^{-\bar{\alpha}\tau_n}$  and  $\underline{W}_x := \liminf_n C_n^x e^{-\underline{\alpha}\tau_n}$ . Then for large enough  $N$ , with an appropriate  $\epsilon > 0$ , we have:

$$(\underline{W}_x - \epsilon)e^{\underline{\alpha}\tau_n} < C_n^x < (\bar{W}_x + \epsilon)e^{\bar{\alpha}\tau_n} \text{ for all } n \geq N. \quad (11)$$

Observe that  $\tau_n \rightarrow \infty$  a.s. on survival sample paths<sup>4</sup>. Thus from (11), we have exponential growth of number of unread copies of any post, with a rate that is upper and lower bounded respectively by  $\bar{\alpha}$  and  $\underline{\alpha}$ .

(b) the probability of extinction of both the posts, i.e., the probability that  $S_n$  (number of total posts) becomes zero for some  $n < \infty$  is given by part (iii) of the Theorem.

**Population-size independent case:** Consider a special case with  $\eta_x = \eta_y$ ,  $\gamma(\cdot) \equiv \gamma$  and any values of  $p_{xy}, p_{yx}$  (need not be equal). Here, the upper and lower bounding processes in the proof of the theorem (in Appendix) coincide, as from (10)  $\alpha = \underline{\alpha} = \bar{\alpha} = \lambda(m_f \eta_x (1 - 2\gamma) - 1)$ .

Also, clearly from (3),  $\{S_n\}_n$  corresponds to a standard BP (the dynamics of individual propagation is complicated, however the sum population evolves exactly like a single type BP) and hence exhibits dichotomy (see [1]): either it has asymptotic (exponential) growth when it survives ( $\bar{W}_s = \underline{W}_s > 0$ ) or it gets extinct ( $\bar{W}_s = \underline{W}_s = 0$ ).

From Theorem 4.iii, the extinction probabilities are also equal,  $q = \bar{q} = \underline{q}$ . This is the probability that both the posts are wiped out from the network in a finite time. Further consider symmetric case:  $c_0^x = c_0^y$  and  $p_{xy} = p_{yx}$ . Then by symmetry, the  $x$ -post survives with probability more than half that of  $S_n$ , i.e.,  $(1 - q^{2c_0^x})/2$  (more is possible due to co-existence); in standard BPs the descendants of each seed user evolve independently (see [1]). However, observe that the probability of survival of  $x$ -post without attack (in BPNA model) equals  $(1 - q^{c_0^x})$ , which would be substantially larger, as  $(1 - q^{c_0^x}) > (1 - q^{2c_0^x})/2$ .

From Theorem 4.ii, the growth rate of individual posts with attack equals  $\alpha$ , which equals that in BPNA only when  $\gamma = 0$ .

<sup>4</sup> $\tau_n$  is lower bounded by maximum among  $n$  exponential random variables

#### IV. NUMERICAL EXPERIMENTS

In this section, we perform extensive Monte-Carlo (MC) simulations on both synthetic data and ego-network dataset of Twitter provided by SNAP [11] to draw further insights into viral competing markets and to validate our theoretical results. Towards this, we simulate our process till a pre-defined system time  $T$  and define the respective extinction probability of total number of live posts, individual posts, and the probability of co-existence as:  $q_s^T := P(S_t = 0 \text{ for some } t \leq T)$ ,  $q_x^T := P(C_t^x = 0 \text{ for some } t \leq T)$ ,  $p_{co}^T := P(C_t^x > 0, C_t^y > 0 \text{ for all } t \leq T)$ . Similarly define  $q_y^T$ .

##### A. Observations using synthetic data

We consider population-size independent case, i.e., consider that  $\gamma(\cdot)$  is a constant, while population-size dependent case is considered with ego-twitter dataset. We performed MC simulations till  $T = 10^6$  for 3200 instances with parameters:  $\lambda = 0.0002$ ;  $F \stackrel{d}{\sim} \text{Poisson}(4)$ ;  $\xi_x, \xi_y \stackrel{d}{\sim} \text{Bin}(F, 0.2667)$ ;  $\xi_{xy}, \xi_{yx} \stackrel{d}{\sim} \text{Bin}(F, 0.053)$ ;  $p_{xy}, p_{yx} = 0.3$  in Table I, i.e., with  $m_x = m_y = 1.067$  and  $m_{xy}^* = m_{yx}^* = 0.064$ . We observe that the  $y$ -post gets extinct with increasing probabilities as  $c_0^x$  increases (as in Theorem 1). For the case without attack (BPNA) as  $T \rightarrow \infty$ ,  $q_s^T, q_x^T, q_y^T \rightarrow q^{c_0^x + c_0^y}, q^{c_0^x}, q^{c_0^y}$  and  $p_{co}^T \rightarrow (1 - q^{c_0^x})(1 - q^{c_0^y})$ , where  $q$  is defined in Theorem 4; the estimates of the last four columns well match.

TABLE I: Both posts have equal potential:  $c_0^y = 2$

| $c_0^x$ | With attack and acquisition |         |         |            | Without attack and acquisition |         |         |            |
|---------|-----------------------------|---------|---------|------------|--------------------------------|---------|---------|------------|
|         | $q_s^T$                     | $q_x^T$ | $q_y^T$ | $p_{co}^T$ | $q_s^T$                        | $q_x^T$ | $q_y^T$ | $p_{co}^T$ |
| 2       | 0.589                       | 0.792   | 0.797   | 0          | 0.592                          | 0.768   | 0.775   | 0.049      |
| 4       | 0.451                       | 0.609   | 0.842   | 0          | 0.453                          | 0.590   | 0.757   | 0.106      |
| 10      | 0.204                       | 0.287   | 0.917   | 0          | 0.203                          | 0.261   | 0.773   | 0.169      |
| 16      | 0.094                       | 0.123   | 0.970   | 0          | 0.092                          | 0.119   | 0.785   | 0.188      |
| 30      | 0.016                       | 0.020   | 0.996   | 0          | 0.013                          | 0.018   | 0.766   | 0.228      |

From columns 2 and 6, the extinction probabilities of the total posts is the same with/without attack (see subsection III-D with  $\eta_x = \eta_y$ ), however that of the individual posts are very different in the two models; this shows that *ignoring the existence of competing content can be misleading*. Observe from  $p_{co}^T$  of column 5 that co-existence is not possible with competition, which is mis-judged by BPNA models, column 9.

In Table II,  $x$ -post is more attractive ( $\xi_x \stackrel{d}{\sim} \text{Bin}(F, 0.3325)$ ) and attacks more ( $\xi_{xy} \stackrel{d}{\sim} \text{Bin}(F, 0.0667)$ ), other parameters as in Table I. Clearly  $x$ -post has more potential to grow larger, attack  $y$ -post and we capture the same in Table II where  $q_s^T$  and  $q_x^T$  have reduced significantly in comparison to Table I; while  $y$ -post gets extinct with higher probabilities. This difference is unnoticed by BPNA models (see  $q_y^T$  without attack of row 1 in the two tables).

TABLE II:  $x$ -post has higher potential:  $c_0^y = 2$

| $c_0^x$ | With attack and acquisition |         |         |            | Without attack and acquisition |         |         |            |
|---------|-----------------------------|---------|---------|------------|--------------------------------|---------|---------|------------|
|         | $q_s^T$                     | $q_x^T$ | $q_y^T$ | $p_{co}^T$ | $q_s^T$                        | $q_x^T$ | $q_y^T$ | $p_{co}^T$ |
| 2       | 0.218                       | 0.396   | 0.822   | 0          | 0.235                          | 0.305   | 0.742   | 0.188      |
| 6       | 0.022                       | 0.041   | 0.981   | 0          | 0.021                          | 0.028   | 0.729   | 0.265      |
| 10      | 0.002                       | 0.005   | 0.998   | 0          | 0.001                          | 0.002   | 0.700   | 0.298      |

In Figure 1 we fix  $c_0^x = 2$ , while rest of the parameters are as in Table II. Even with,  $m_x + m_{xy}^* > m_y + m_{yx}^*$ , i.e., *even*

when the combined reproduction and attack power of  $x$ -type is bigger than that of  $y$ -type, the  $x$ -post can get extinct with large probabilities if seed users of  $y$ -post are large enough (for example  $c_0^y \geq 10$  in the figure), thus re-affirming Theorem 1.

TABLE III: Co-existence of populations

| $10^{-5} c_0^x$ | $c_0^x/c_0^y$ | $m_x$  | $m_y$ | $10^{-5} N$ | $c_N^x$    | $\beta^o$ | $c_N^x/s_N$ |
|-----------------|---------------|--------|-------|-------------|------------|-----------|-------------|
| 1.01            | 1.005         | 2.9998 | 3.0   | 25974       | 2614884763 | 0.501     | 0.503       |
| 10              | 1             | 2      | 2     | 13141       | 659107240  | 0.5       | 0.501       |
| 10              | 0.618         | 3.0    | 2.98  | 10070       | 765240476  | 0.382     | 0.382       |
| 10              | 0.618         | 2.92   | 2.90  | 10534       | 766418907  | 0.382     | 0.381       |

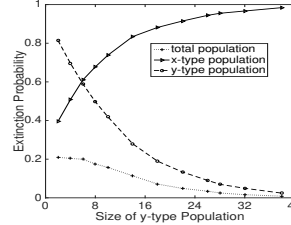


TABLE IV: Approaching co-existence!

| $m_x, m_y, m_c^*$ | $c_0^x, c_0^y$ | $\beta^o$ | $\%_{in}$ |
|-------------------|----------------|-----------|-----------|
| (300, 300, 10)    | (2500, 3001)   | 0.5       | 11.2      |
| (300, 300, 10)    | (3000, 3001)   | 0.5       | 10.4      |
| (300, 300, 10)    | (3000, 3601)   | 0.5       | 10.1      |
| (300, 300, 30)    | (3000, 3601)   | 0.5       | 5.1       |
| (300, 290, 20)    | (3000, 4611)   | 0.44      | 5.3       |

Fig. 1: High disparity in seed users

Next, we reinforce co-existence results of Theorems 2 and 3 with  $m_{xy}^* = m_{yx}^* = 0.02$  in Table III to also support our conjecture (Theorem 2 with  $m_x > m_y$ ). We see that the limit fractions estimated through MC simulations well match the theoretical ones (9). Note here that  $(c_0^x, c_0^y)$  are very large.

Recall that in Theorem 3, we showed that if posts start close to  $\beta^o$  with large number of seed users, then for any finite time, the fraction remains close to  $\beta^o$ . In Table IV, we corroborate this result by taking  $\beta_0$  near  $\beta^o$  and tabulating the fraction of sample paths (average is over 1000 sample paths) in which the ratio converges even closer (within 5%) to  $\beta^o$ , after  $10^4$  transitions. Here,  $m_{xy}^* = m_{yx}^* = m_c^*$ . It is clear from Table IV that in considerable number of sample paths, the fraction of posts converge within 0.05-band around  $\beta^o$ , even when the process started outside the band.

##### B. Validation through ego-twitter dataset by SNAP

We study the post propagation over twitter dataset provided by SNAP [11]. We could extract 37, 255 users and 1, 048, 576 directed connections using this data. We are yet to utilize the information of circles provided by SNAP. The data had insufficient cross connections and  $\eta_x, \eta_y$  are naturally small parameters, which resulted in very few instances of the post being forwarded to a user that already had a post. Hence we introduced a parameter  $\varphi$ ; while forwarding shares to  $\eta_x$  or  $\eta_y$  fraction of users we consider  $\varphi$  sub-fraction specifically among those friends that already have the post. We also introduced on average two extra friends per user to increase the cross connections. We then ran the usual MC based simulations, picked seed users and friends randomly from the dataset, generated exponential clocks and the relevant user wake-ups/visits, forwarding of the post randomly as explained in section II-B. We also *considered independent propagation of the two posts by setting  $\varphi = 0$ ,  $p_{xy} = p_{yx} = 0$ , which we refer as BPNA simulations*; this simulation neglects the fact that the users predominantly pick one among the two competing posts, when they receive both. Our aim is to show

the drastic differences in the two types of the propagations, alongside demonstrating the validity of our theoretical results. The estimates are after averaging for 1000 sample paths.

We consider a first example in Table V with equal number of seed users. We set  $\eta_x, \eta_y = 0.05$ ,  $\lambda = 5$ ,  $(T, \varphi) = (2.1, 0.6)$ ,  $p_{xy}, p_{yx} = 1$ . We observe that the probability of co-existence (last column), as well as existence (column 3,4) are quite small even in spite of starting with seed users as large as 8. This shows the significance of attack inherently present in the simultaneous propagation of competing posts. Had it been independent propagation, these probabilities would have been much higher (theoretically equal  $(1 - q^{c_0^x})$ , see section III-D). We also run independent BPNA simulations for  $c_0^x = c_0^y = 8$  and observed that  $1 - q_x^T = 1 - q_y^T = 0.9$  and  $p_{co}^T = 0.48$ .

TABLE V: With equal seeds

| $c_0^x$ | $c_0^y$ | $1 - q_x^T$ | $1 - q_y^T$ | $p_{co}^T$ |
|---------|---------|-------------|-------------|------------|
| 4       | 4       | 0.613       | 0.631       | 0.127      |
| 5       | 5       | 0.672       | 0.658       | 0.186      |
| 8       | 8       | 0.748       | 0.754       | 0.282      |

TABLE VI: BPA versus BPNA

|         |         | BPA     |         | BPNA    |         |
|---------|---------|---------|---------|---------|---------|
| $c_0^x$ | $c_0^y$ | $q_x^T$ | $q_y^T$ | $q_x^T$ | $q_y^T$ |
| 3       | 20      | 0.751   | 0.020   | 0.138   | 0       |
| 4       | 20      | 0.682   | 0.018   | 0.044   | 0       |
| 5       | 30      | 0.66    | 0.004   | 0.028   | 0       |

In the second example of Table VI, we consider disparity in the number of seed users. Here, we set  $\eta_x = \eta_y = 0.2$ ,  $\lambda = 5$ ,  $T = 0.6$ . We also tabulate results with BPNA simulations in the last 2 columns. We see a huge disparity in the two sets of extinction probabilities; for example, when competition is considered, even with 5 seed users  $x$ -post gets extinct with more than 0.6 probability while with independent runs the extinction is only 0.028.

## V. CONCLUSIONS

Online social platforms are usually flooded with variety of content; some of which are competing with each other. The content is stored in an ordered manner as in timelines of Facebook, Twitter etc. Thus, a user has a wide variety of posts to read, which provides them multiple (commercial) choices, however, the competing contents are always at a risk of losing their chances. When a (advertising/commercial) post in a user's timeline gets shifted down by a newer competing post, the new post snatches away the opportunities of (attacks and acquires) the old post depending upon the popularity and/or the freshness of the two contents.

We propose a new variant of continuous-time population-size dependent branching process, namely Branching process with Attack, that mimics this 'attack' and 'acquire' phenomenon of the competing content. This new variant poses new questions along with the old set of questions (related to growth rates, number of shares and extinction probabilities etc.), that of *co-existence/co-explosion* of the competing populations. In the context of viral competing markets, this new question translates to *co-virality* possibilities, i.e., simultaneous spread and explosion of the competing content over the network.

Our work has two-fold contributions towards the literature: branching processes, as well as, the viral competing markets. We adopted a *novel approach of using stochastic approximation techniques with branching processes* to derive the time-asymptotic limit proportions for individual populations.

*In particular, in regard to viral competing markets, we showed that either one or both the posts get extinct. If the posts co-exist, they converge to a fixed (unique) ratio.* Interestingly, this ratio is smaller for the post which is more attractive and has higher net attack capacity.

Any CP can capture the market if it manages to procure large number of seed users in comparison to its opponent, even when the other CP is a market giant. In other words, seed users play a critical role. In contrast, when the two CPs start with comparable number of seed users, there is a possibility of co-virality. We also discuss the growth rates and extinction probabilities of individual posts.

We validated our results through Monte-Carlo simulations. Using real traces, we showed that simple BPs (when attack and acquisition is ignored) are insufficient to model viral competing markets and give erroneous conclusions; for example, on twitter-dataset, models that neglect attack-acquisition aspect underestimate the extinction probabilities up to 93.55%.

The structure of timelines have different types of influences on post-propagation (as is recently observed in [6], some posts may be lost before the user's visit to its timeline). It would be interesting to study this aspect along with competing posts. Further, using the performance results of this paper, one can consider relevant game-theoretic and optimization problems.

## VI. APPENDIX

**Proof of Theorem 1<sup>5</sup>:** We are considering conditional probability, given that  $C_0^x = c_0^x$  and  $C_0^y = c_0^y$  are the respective initial sizes. Let  $\tau$  be the time epoch before one among the  $c_0^y$  number of particles wakes up; observe that multiple  $x$ -type particles might have woken up during this time. Then,  $\tau \geq \tau_e$ , where  $\tau_e$  is exponentially distributed random variable with parameter  $\lambda c_0^y$  and equals the minimum wake-up time among  $c_0^y$  number of  $y$ -particles. Note that some of them could have been attacked/acquired in the meanwhile, and hence  $\tau_e$  is only a lower bound. Let  $N_{\#}$  be the number of transitions (i.e., wake-ups) of  $x$ -type before a  $y$ -type particle wakes up; then clearly  $N_{\#} \geq \text{Bin}(c_0^x, 1 - e^{-\lambda\tau})$  a.s.; the lower bound is obtained by considering only those among  $c_0^x$  particles that woke-up. Let  $\widetilde{C}^y$  denote the size of the  $y$ -type population at the next  $y$ -transition. Let  $P_0(\cdot)$  represent the conditional probability  $P(\cdot | C_0^x = c_0^x, C_0^y = c_0^y)$  and similarly let  $E_0[\cdot]$  represent the conditional expectation  $E[\cdot | C_0^x = c_0^x, C_0^y = c_0^y]$ . Now with  $\zeta := \zeta_{xy}(1)$  (recall  $\zeta_{xy}(1) \leq \zeta_{xy}(c^y)$  a.s. for any  $c^y \geq 1$ ) we have the following:

$$\begin{aligned} P_0(C^y \rightarrow 0) &\geq P_0(\widetilde{C}^y = 0) = P_0\left(c_0^y < \sum_{i=1}^{N_{\#}} \zeta_{xy,i}\right) \\ &\geq P_0\left(c_0^y < \sum_{i=1}^{c_0^x} \zeta_i 1_{\{\tau_{\uparrow}\}}\right) = E_0\left[P_0\left(c_0^y < \sum_{i=1}^{c_0^x} \zeta_i 1_{\{\tau_{\uparrow}\}} \mid \tau\right)\right], \end{aligned}$$

where  $1_{\{\tau_{\uparrow}\}}$  is indicator of the event that the  $x$ -particle under consideration has woken-up before  $\tau$ .

Consider a fixed  $\bar{\tau}$  such that (possible as  $c_0^y$  is fixed)

$$P_0(\tau \geq \bar{\tau}) > P_0(\tau_e \geq \bar{\tau}) = e^{-\lambda c_0^y \bar{\tau}} > \sqrt{1 - \epsilon}. \quad (12)$$

<sup>5</sup>This proof is loosely inspired by the proofs provided in [4].

Let  $\Phi$  be the complementary CDF of standard normal random variable and define the following for the above choice of  $\bar{\tau}$ :

$$\bar{Z}_{c_0^x} := \frac{\sum_{i=1}^{c_0^x} \zeta_i 1_{\{x \uparrow\}} - c_0^x m_{xy}(1)(1 - e^{-\lambda \bar{\tau}})}{\text{Var}(\zeta_i)}.$$

Let  $\bar{z}_{c_0^x} := \frac{c_0^y - c_0^x m_{xy}(1)(1 - e^{-\lambda \bar{\tau}})}{\text{Var}(\zeta_i)}$ . Observe that  $\{\zeta_i 1_{\{x \uparrow\}}\}_i$  are IID random variables and hence by Central Limit Theorem and Portmanteau Theorem for  $\varepsilon := 1 - \sqrt{1 - \epsilon}$ , there exists a  $\bar{c}_0^x < \infty$  such that<sup>6</sup> for all  $c_0^x \geq \bar{c}_0^x$ :

$$P(\bar{Z}_{c_0^x} > \bar{z}_{c_0^x}) \geq P_0(\bar{Z}_{c_0^x} > \bar{z}_{c_0^x}) \geq \Phi(z_{c_0^x}) - \varepsilon/2.$$

If required, choose  $\bar{c}_0^x$  further large such that

$$P(\bar{Z}_{c_0^x} > \bar{z}_{c_0^x}) \geq 1 - \varepsilon \text{ for all } c_0^x \geq \bar{c}_0^x, \quad (13)$$

and this is possible because  $\Phi(a) \rightarrow 1$  when  $a \rightarrow -\infty$  (hence  $\Phi(\bar{z}_{c_0^x}) \rightarrow 1$  as  $\bar{c}_0^x \rightarrow \infty$ ). Using the bounds (12), (13) and conditioning on  $\tau$ , we have for all  $c_0^x \geq \bar{c}_0^x$ :

$$\begin{aligned} P_0(C^y \rightarrow 0) &\geq P_0(C^y \rightarrow 0; \tau \geq \bar{\tau}) \\ &\geq P_0(E[\bar{Z}_{c_0^x} > \bar{z}_{c_0^x} | \tau; \tau \geq \bar{\tau}]) \geq P_0(\bar{Z}_{c_0^x} > \bar{z}_{c_0^x}) P_0(\tau \geq \bar{\tau}) \\ &= P(\bar{Z}_{c_0^x} > \bar{z}_{c_0^x}) P_0(\tau \geq \bar{\tau}) \geq (1 - \varepsilon). \end{aligned} \quad \blacksquare$$

**Proof of Theorem 2:** Define  $L_n := [L_n^\psi, L_n^\theta, L_n^t]^T$ , where,

$$\begin{aligned} L_n^\psi &= \left( \tilde{\xi}_n - \psi_{n-1} \right) K_n, \quad L_n^t = 1 \text{ and} \\ L_n^\theta &= \left[ H_n \left( \tilde{\xi}_n^x + \zeta_{xy,n} I_n \right) - H_n^c \tilde{\zeta}_{yx,n} I_n - \theta_{n-1} \right] J_n. \end{aligned} \quad (14)$$

Define  $\mathcal{F}_n := \sigma\{C_k^x, C_k^y, \tau_k : 1 \leq k \leq n\}$  for all  $n$ . Observe that  $E[L_n | \mathcal{F}_{n-1}] = \bar{g}(\Theta_n)$  (right hand side of ODE (8)).

Now, we will prove the result using [10, Theorem 2.2, pp. 131], as  $\bar{g}(\cdot)$  is only measurable. To this end, observe from (4)

that  $\tilde{\xi}_k(z) \leq \tilde{\xi}_k(0, 0)$  a.s. Say  $\tilde{\xi}_k^y(0, 0) \stackrel{d}{\leq} \tilde{\xi}_k^x(0, 0)$ . Define a new process  $\tilde{\xi}_n$ , which evolves like  $\xi_n$ , but with  $\tilde{\xi}_k(z)$  in (4) replaced by IID terms  $\tilde{\xi}_k^x(0, 0)$ . By appropriate coupling arguments, one can dominate  $\xi_n$  by  $\tilde{\xi}_n$  almost surely and then:

$$C_n^x \leq S_n \leq n |\tilde{\xi}_n| \leq n |\tilde{\xi}_n| \text{ for all } n. \quad (15)$$

By law of large numbers,  $\tilde{\xi}_n \rightarrow m_x(0, 0) - 1$  a.s.

For using said result in [10], we first need to prove (a.s.) equicontinuity of sequence  $\Theta^n(t) := \Theta_n + \sum_{i=n}^{m(t_n+t)-1} \epsilon_i L_i$ , with  $m(t)$  of [10] equal to  $\eta(t)$ . This proof goes through exactly as in the proof of [10, Theorem 2.1, pp. 127] because of the following reasons: the random vector  $L_n$  is comprised of  $\theta_n, \psi_n$  and IID random variables and by (15), it suffices to show that  $\sup_n E|\tilde{\xi}_n|^2 < \infty$ , which is trivially true under A.2; further, we exactly have  $E[L_n | \mathcal{F}_{n-1}] = \bar{g}(\Theta_n)$  (here  $\beta_n$  in [10, Assumption A.2.2] is 0), as well the projection term is zero. Further,  $\{\Theta^n(0)\}_n$  is bounded a.s. by strong law of large numbers as applied to  $\{\tilde{\xi}_n\}_n$ .

In Lemma 1, we identify the attractors<sup>7</sup> of (8), with  $\theta^o, \beta^o$  as in (9) for the special case with  $\eta_x = \eta_y = \eta$  and  $\gamma(\cdot) \equiv \gamma$ ;

<sup>6</sup>Because the Gaussian measure is measure of a continuous random variable, the set under consideration is Gaussian-continuity set (see [3]).

<sup>7</sup>A set  $A$  is said to be Asymptotically stable in the sense of Liapunov, if there exist a neighbourhood (called domain of attraction,  $D(A)$ ) starting in which the ODE trajectory converges to  $A$  as time progresses (e.g., [10]).

the proof of Lemma 1 is only a conjecture for general case. Proof is now completed sample-path wise.

First consider the sample-paths in which  $\psi_n \rightarrow 0$ . Then clearly,  $(\psi_n, \theta_n) \rightarrow (0, 0)$ . Observe in complementary sample paths,  $\psi_n \rightarrow m - 1$  a.s., where  $m := m_f \eta(1 - 2\gamma)$ .

Consider the sample-paths in which  $\{\theta_n\}$  sequence does not exit neighbourhood  $\mathcal{N}_\epsilon(\theta^o) := \{\theta : |\theta - \theta^o| < \epsilon\}$  infinitely often (i.o.), for every  $\epsilon > 0$ ; that is<sup>8</sup>,  $\theta_n \rightarrow \theta^o$  and so  $(\psi_n, \theta_n) \rightarrow (m - 1, \theta^o)$ .

Now consider the remaining sample paths, then there exists<sup>9</sup> at least one  $\epsilon > 0$  such that  $\{\psi_n, \theta_n, t_n\}$  sequence visits the compact  $S_\epsilon$  of Lemma 1 i.o. By [10, Theorem 2.2, pp. 131] as applied to these sample paths, the sequence converges to the attractor  $A$  of Lemma 1.  $\blacksquare$

**Lemma 1: [Attractors]** For ODE (8), the set  $A := \{(m - 1, 0), (m - 1, m - 1)\} \times \{\infty\}$  is locally asymptotically stable in the sense of Liapunov. For any  $\epsilon > 0$ , the set<sup>10</sup>

$$S_\epsilon = \left\{ (\psi, \theta) : \psi \in \bar{\mathcal{N}}_\epsilon(m - 1), \frac{\theta}{\psi} \in [0, 1] - \mathcal{N}_\epsilon(\beta^o) \right\} \times [T_0, \infty],$$

is compact and is in the domain of attraction of  $A$ , when  $T_0$  is such that  $\eta(T_0)\beta^o(m - 1 - \epsilon) \geq \max\{\bar{c}^y \kappa_{xy} / \kappa_{yx}, \bar{c}^x \kappa_{yx} / \kappa_{xy}\}$ .

Further, the equilibrium point  $(m - 1, \theta^o, \infty)$  is not stable.  $\blacksquare$

**Proof :** For  $t$  component, we consider the extended positive real line including  $\infty$ , with an appropriate metric; proofs in [10] would go through even for this. With this, the required compactness is true. The  $\psi$ -component of the ODE (8) has the following solution:

$$\psi(t) = \begin{cases} e^{-t}(\psi(0) - m + 1) + m - 1, & \text{when } \psi(0) > 0, \\ \psi(0), & \psi(0) \leq 0. \end{cases}$$

Thus  $(m - 1)$  is asymptotically stable with  $(0, \infty)$  as domain of attraction. For  $\theta$  component, one needs to substitute solution  $\psi(t)$  in its ODE ( $\bar{g}_\theta$  of (8)) to analyze. By considering  $\psi^* = m - 1$ , the equilibrium points<sup>11</sup> for the ODE corresponding to  $\theta$  component are 0,  $m - 1$  or  $\theta^o$ .

To test the stability of the above equilibrium points we consider the ODE representing the ratio  $\beta = \theta/\psi$  (derived using (8) and with  $I_{\theta/\psi} := 1_{\{0 < \theta < \psi\}}$ ):

$$\dot{\beta} = \frac{I_{\theta/\psi}}{\psi} \left( \beta m_{xy}(c^y(t)) - (1 - \beta) m_{yx}(c^x(t)) \right).$$

Consider any  $\psi \in \bar{\mathcal{N}}_\epsilon(m - 1)$ ,  $t \geq T_0$  and  $\beta > \beta^o$ , then by A.3 we have:

$$\begin{aligned} \dot{\beta} &\geq \frac{I_{\theta/\psi}}{\psi} \left\{ \frac{\theta}{\psi} \kappa_{xy} \min\{\bar{c}^y, (\psi - \theta)\eta(t)\} - \left(1 - \frac{\theta}{\psi}\right) \kappa_{yx} \bar{c}^x \right\} \\ &= \frac{I_{\theta/\psi}}{\psi} \min \left\{ \left( \bar{c}^y \kappa_{xy} + \bar{c}^x \kappa_{yx} \right) \frac{\theta}{\psi} - \bar{c}^x \kappa_{yx}, \right. \\ &\quad \left. \frac{\theta}{\psi} (\psi - \theta)\eta(t) \kappa_{xy} - \left(1 - \frac{\theta}{\psi}\right) \bar{c}^x \kappa_{yx} \right\} > 0. \end{aligned}$$

<sup>8</sup>From Lemma 1,  $(m - 1, \theta^o, \infty)$  is only an equilibrium point and not an attractor; nonetheless the actual dynamics can still converge to it.

<sup>9</sup>If  $\{\theta_n\}$  sequence exits neighbourhood  $\mathcal{N}_\epsilon(\theta^o)$  i.o., and say  $\{\psi_n\}$  entered  $\bar{\mathcal{N}}_\epsilon(m - 1)$ , then choose  $\epsilon < \frac{(1 - \beta^o)\bar{\epsilon}}{m - 1 + \bar{\epsilon}}$ .

<sup>10</sup>Define  $\bar{\mathcal{N}}_\epsilon(\theta^o) := \{\theta : |\theta - \theta^o| \leq \epsilon\}$ .

<sup>11</sup>In this context, the point  $\bar{\theta}$  is an equilibrium point if  $\bar{g}_\theta(\psi^*, \bar{\theta}, \infty) = 0$ .



In the above, the first term is positive by A.3 and definition of  $\beta^L$  and the second one is positive by choice of  $T_0$ , as after dividing the second term by  $(\psi - \theta)/\psi$  we get:

$$\frac{\theta}{\psi} \psi \eta(t) \kappa_{xy} - \bar{c}^x \kappa_{yx} > \beta^o (m - 1 - \epsilon) \eta(T_0) \kappa_{xy} - \bar{c}^x \kappa_{yx} > 0.$$

Thus the derivative of ratio ( $\beta$ ) is positive, and hence the ratio increases and  $\beta_t \rightarrow 1$ , when initialized with  $\beta_0 > \beta^o$ . Similarly, when initialized with  $\beta_0 < \beta^o$ , the derivative  $\dot{\beta} < 0$  throughout and  $\beta_t \rightarrow 0$ .

Using the above arguments one can also conclude that from any neighbourhood of  $(m-1, \theta^o, \infty)$ , there exist points starting from which the ODE converges either to  $(m-1, m-1, \infty)$  or  $(m-1, 0, \infty)$  and hence the last part. ■

**Proof of Theorem 3:** We are given  $z \in \mathcal{N}_{\epsilon, \Delta}$ , i.e.,  $s = \Delta := k\delta$  for some  $k, \delta > 0$ . Now observe that (with  $\beta_i^c = 1 - \beta_i$ ):

$$\beta_n = \frac{\beta_0 \delta + \sum_{i=1}^n \frac{U_i}{k}}{\delta + \sum_{i=1}^n \frac{G_i}{k}}, \quad (U_i, G_i) := \begin{cases} (-\zeta_{yx,i}, \tilde{\xi}_i^y), & \text{w.p. } \beta_{i-1}^c \\ (\tilde{\xi}_i^x + \zeta_{xy,i}, \tilde{\xi}_i^x) & \text{w.p. } \beta_{i-1}. \end{cases}$$

Observe that  $\Gamma_{\mathcal{N}_\epsilon} = \min\{\Gamma_+^{(k)}, \Gamma_-^{(k)}\}$ , where  $\Gamma_+^{(k)} = \min\{n : \beta_n > \beta^o + \tilde{\epsilon}\}$  and with  $s = k\delta$ :

$$\Gamma_-^{(k)} := \Gamma^{(k)} := \min \left\{ n : \frac{\beta_0 \delta + \sum_{i=1}^n \frac{U_i}{k}}{\delta + \sum_{i=1}^n \frac{G_i}{k}} < \beta^o - \tilde{\epsilon} \right\}.$$

Suffices to analyze one of  $\Gamma_+^{(k)}, \Gamma_-^{(k)}$ , then proof follows for the other in the same way. We first analyse the case where the number of offsprings and attacks are finite. In particular, assume that for each  $i, j \in \{x, y\}$  and  $i \neq j$ , we have  $\zeta_{ij}, |\tilde{\xi}^i|$  are upper bounded by  $\bar{\zeta}_{ij}, \bar{\xi}^i < \infty$  a.s. respectively. We consider a special sample path  $\omega^l$  where  $y \uparrow$  continuously, produces  $\xi^y$  offsprings and attacks  $\bar{\zeta}_{yx}$  individuals of  $x$ -type. Observe that such a sample path would take the least time to exit  $\mathcal{N}_\epsilon^*$  and clearly is given by:

$$\begin{aligned} \underline{\Gamma}^{(k)} &:= \min \left\{ n : \frac{\beta_0 \delta - n \frac{\bar{\zeta}_{yx}}{k}}{\delta + n \frac{\bar{\xi}^y}{k}} < \beta^o - \tilde{\epsilon} \right\} \\ &= \min \left\{ n : n > \frac{\delta(\tilde{\epsilon} + \beta_0 - \beta^o)k}{\bar{\xi}^y(\beta^o - \tilde{\epsilon}) + \bar{\zeta}_{yx}} \right\} = \left\lceil \frac{\delta(\tilde{\epsilon} + \beta_0 - \beta^o)k}{\bar{\xi}^y(\beta^o - \tilde{\epsilon}) + \bar{\zeta}_{yx}} \right\rceil. \end{aligned}$$

Thus  $\underline{\Gamma}_k$  is increasing in  $k$  and  $\underline{\Gamma}^{(k)} \rightarrow \infty$  as  $k \rightarrow \infty$ . Now, by construction of special sample path  $\omega^l$ ,  $\Gamma^{(k)} > \underline{\Gamma}^{(k)}$  a.s., and thus,  $\Gamma^{(k)} \rightarrow \infty$  as  $k \rightarrow \infty$  a.s. and this completes the proof for bounded attacks and offsprings.

For general case, define  $E_k := \{\underline{\Gamma}^{(k)} \leq T\}$ ,  $\bar{E}$  as complement of

$$E := \{\cap_{n=1}^T (\tilde{\xi}_n^y < \bar{\xi}^y)\} \cap \{\cap_{n=1}^T (\zeta_{yx,n} < \bar{\zeta}_{yx})\}$$

We have,  $\inf_{z \in \mathcal{N}_{\epsilon, \Delta}} P_z(E_k) = \inf_{z \in \mathcal{N}_{\epsilon, \Delta}} \left( P_z(E_k, E) + P_z(E_k, E^c) \right)$ .

The second term (on right side) is upper bounded by  $T[P_z(\tilde{\xi}_n^y > \bar{\xi}^y) + P_z(\zeta_{yx,n} > \bar{\zeta}_{yx})]$ . For any  $\kappa > 0$ , this term is upper bound by  $\kappa/2$  for an appropriate choice of  $\bar{\xi}^y$  and  $\bar{\zeta}_{yx}$  (finite first moments). For this choice of  $\bar{\xi}^y$  and  $\bar{\zeta}_{yx}$ , choose  $k$  large enough such that first term is also less than  $\kappa/2$ . ■

**Proof of Theorem 4:** From the details of sub-section II-B,  $\tilde{\xi}_y^* := \lim_{z \rightarrow \infty} \tilde{\xi}^y(z)$  exists a.s. We now describe two single type population-size independent BPs that evolve at their own transition (wake-up) epochs as below (like in (3)):

$$\bar{S}_n = \left( \bar{S}_{n-1} + \tilde{\xi}_n^x(0, 0) \right) 1_{\bar{\nu}_e \leq n} \quad \text{and} \quad \underline{S}_n = \left( \underline{S}_{n-1} + \tilde{\xi}_{y,n}^* \right) 1_{\underline{\nu}_e \leq n},$$

where  $\bar{\nu}_e$  and  $\underline{\nu}_e$  are their respective extinction epochs as in (4)-(5). Clearly the offsprings/shares that update the original  $\{S_t\}$  process can be lower/upper bounded (almost surely) by those in the above equation. Thus from (4), the evolution of  $S_n$  is given and upper/lower bounded almost surely for all  $n$  by (three processes are updated at their own transition epochs)

$$S_n = S_{n-1} + \tilde{\xi}_{n-1}(Z_n), \quad \text{and} \quad \underline{S}_n \leq S_n \leq \bar{S}_n. \quad (16)$$

Further more, clearly, the wake-up epochs will be faster in the upper bounding systems (i.e.,  $\underline{\tau}_n \geq \tau_n \geq \bar{\tau}_n$  a.s.) and hence:

$$S_n e^{-\bar{\alpha} \tau_n} \leq \bar{S}_n e^{-\bar{\alpha} \bar{\tau}_n} \quad \text{and} \quad S_n e^{-\alpha \tau_n} \geq \underline{S}_n e^{-\alpha \underline{\tau}_n} \quad \text{a.s.}$$

From the standard theory of BPs,  $\{S_n e^{-\alpha \tau_n}\}$  as well as  $\{\bar{S}_n e^{-\bar{\alpha} \bar{\tau}_n}\}$  converge a.s., as well as the limits are non-zero with positive probability [1]. This proves (i). For part (ii), for example, observe that

$$C_n^x e^{-\bar{\alpha} \tau_n} \leq S_n e^{-\bar{\alpha} \tau_n},$$

and the first claim of part (i) follows. Further, consider the sample paths, where  $y \uparrow$  repeatedly and die. Then,  $x$ -type population survive independently, like a single type branching process, at rate  $\lambda(m_f \eta_x(1 - \gamma^*) - 1) > \underline{\alpha}$ , and hence the other claim of part (ii) is done. Part (iii) follows in a similar way. ■

## REFERENCES

- [1] Athreya, Krishna B. "Branching process." Encyclopedia of Environmentrics 1 (2006).
- [2] Beutel, Alex, et al. "Interacting viruses in networks: can both survive?." Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. 2012.
- [3] Billingsley, Patrick. Convergence of probability measures. John Wiley & Sons, 2013.
- [4] Coffey, John, and Wolfgang J. Bühler. "The Galton-Watson predator-prey process." Journal of applied probability (1991): 9-16.
- [5] Deijfen, Maria, and Remco van der Hofstad. "The winner takes it all." The Annals of Applied Probability 26.4 (2016): 2419-2453.
- [6] Dhouchak, Ranbir, Veeraruna Kavitha, and Eitan Altman. "A viral timeline branching process to study a social network." 2017 29th International Teletraffic Congress (ITC 29). Vol. 3. IEEE, 2017.
- [7] Iribarren, José Luis, and Esteban Moro. "Branching dynamics of viral information spreading." Physical Review E 84.4 (2011): 046116.
- [8] Klebaner, Fima Chaim. "On population-size-dependent branching processes." Advances in Applied Probability 16.1 (1984): 30-55.
- [9] Klebaner, Fima C. "Geometric growth in near-supercritical population size dependent multitype Galton-Watson processes." The Annals of Probability 17.4 (1989): 1466-1477.
- [10] Kushner, Harold, and G. George Yin. Stochastic approximation and recursive algorithms and applications. Vol. 35. Springer Science & Business Media, 2003.
- [11] Leskovec, Jure, and Andrej Krevl. "SNAP Datasets." SNAP Datasets: Stanford Large Network Dataset Collection. <http://snap.stanford.edu/data> (2014).
- [12] Prakash, B. Aditya, et al. "Winner takes all: competing viruses or ideas on fair-play networks." Proceedings of the 21st international conference on World Wide Web. 2012.
- [13] Van der Lans, Ralf, et al. "A viral branching model for predicting the spread of electronic word of mouth." Marketing Science 29.2 (2010).
- [14] Zhou, Jingya, et al. "Cost-efficient viral marketing in online social networks." World Wide Web 22.6 (2019): 2355-2378.