



HAL
open science

Viral Marketing Branching Processes

Ranbir Dhouchak, Veeraruna Kavitha, Eitan Altman

► **To cite this version:**

Ranbir Dhouchak, Veeraruna Kavitha, Eitan Altman. Viral Marketing Branching Processes. Computer Communications, 2023, 198, pp.140-156. 10.1016/j.comcom.2022.11.015 . hal-03858831

HAL Id: hal-03858831

<https://inria.hal.science/hal-03858831v1>

Submitted on 17 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Viral Marketing Branching Processes

Ranbir Dhouchak, Veeraruna Kavitha
IEOR, IIT Bombay, India

Eitan Altman
INRIA, Sophia-Antipolis, France
University of Avignon, France

Abstract

¹ We consider the inherent timeline structure of the appearance of content in online social networks (OSNs) while studying content propagation. We model the propagation of a post/content of interest by an appropriate multi-type branching process. The branching process allows one to predict the emergence of global macro properties (e.g., the spread of a post in the network) from the laws and parameters that determine local interactions. The local interactions largely depend upon the timeline (an inverse stack capable of holding many posts and one dedicated to each user) structure and the number of friends (i.e., connections) of users, etc. We explore the use of multi-type branching processes to analyze the viral properties of the post, e.g., to derive the expected number of shares, the probability of virality of the content, etc.

In OSNs, the new posts push down the existing contents in timelines, which can greatly influence content propagation; our analysis considers this influence. We find that one leads to draw incorrect conclusions when the timeline (TL) structure is ignored: a) for instance, even less attractive posts are shown to get viral; b) ignoring TL structure also indicates erroneous growth rates. More importantly, one cannot capture some interesting paradigm shifts/phase transitions; for example, virality chances are not monotone with network activity parameter, as shown by analysis including TL influence.

In the last part, we integrate the online auctions into our viral marketing model. We study the optimization problem considering real-time bidding. We again compared the study with and without considering the TL structure for varying activity levels of the network. We find that the analysis without TL structure fails to capture the relevant phase transitions, thereby making the study incomplete.

Keywords: Online Social Network, Viral Marketing, Branching Processes

1. Introduction

The advent of the Internet has transformed the advertising industry in various ways. With the constant year-on-year growth of the number of users, the global userbase of the Internet passed the 3.5 billion, and 4.9 billion mark respectively in 2017 and 2021, constituting nearly 50% and 63% of the earth's population [23]. This has made the Internet a powerful tool for organizations to interact with users and advertise their prod-

¹This work is sponsored by DST-INRIA project LION.

ucts/services in a personalized manner. In particular, Online Social Networks (OSNs) such as Facebook, Twitter, YouTube, etc., play an instrumental role in the overall digital advertising of the products/services of various organizations. Users on these OSNs exchange volumes of information/data in the form of images, blogs, texts, videos, etc. Due to immense activities of the users in OSNs, the marketing/advertising companies promote their commercial content by leveraging the strengths of these OSNs.

In *viral marketing*, the content providers or advertisers create content that are appealing to the users (for example by giving offers, discounts, or by advertising in an attractive manner); some example strands of such work can be found in [29, 15, 26]; When users find the service/product good enough, they involuntarily spread a word about it, triggering word-of-mouth. Users share the content with their friends, and the information is thus spread through OSNs. In the abstract sense, information spreads like a virus from one person to another, and hence called viral marketing. However, the content propagation has additional complexities which must be incorporated in the model to accurately investigate the process/phenomenon. And we study the same in this paper.

1.1. Motivation and scope of research: Timeline structure

OSNs store volumes of information consumed by the users. At the user level, these segments of information (called posts) are arranged based on their chronological order for the display to users [25]. In other words, posts appear at different levels (each level holds one post) based on their newness on each user’s page in an OSN, for instance, News Feed on Facebook. We call this reverse chronological appearance of the posts a ‘timeline’ (TL), one dedicated to each user. This order of storing contents on timelines (TLs) and the related dynamics have a great influence on content propagation. *However, no attention is paid to the TL structure of the posts/contents appearing on a user’s page in viral marketing literature.* We study the content propagation phenomenon over OSNs, considering the inherent TL structure.

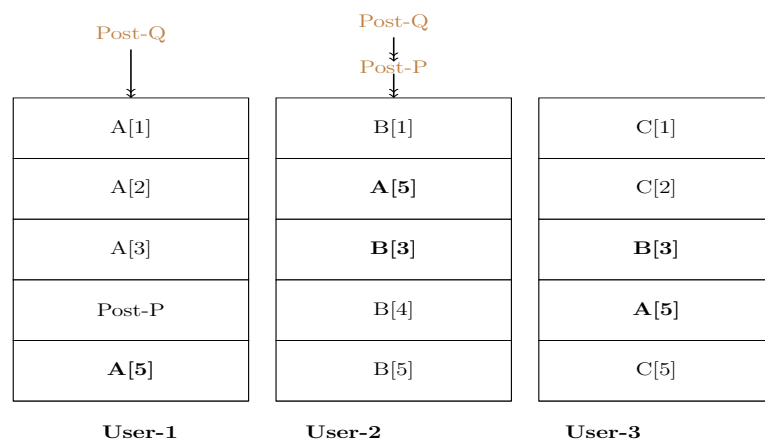


Figure 1: TL structure : User-1 has posts A[1], A[2], A[3], Post-P and A[5] on its TL, at the instance when Post-Q is being forwarded to it; after the forward, this TL will have Post-Q followed by A[1], A[2], A[3] and Post-P. User-2 has two forwards Post-P followed by Post-Q, because of which B[4] and B[5] are removed from its TL. User-3 has no forward hence its TL is unchanged.

A typical example of TL structure (for three users) is shown in Figure 1. Referring to the Figure 1, the natural question to ask is: how many users have a particular post

of interest? The next immediate question is: at what level does that post reside (i.e., the position)? For instance, all the three users have the post ‘A[5]’ on their TLs, but at different levels. It is clear that the posts positioned on the top of the TLs receive more attention/visibility compared to the ones at lower levels. Further, the arrival of new contents keeps shifting/pushing down the existing contents of a TL (as shown in Figure 1). Consequently, a particular content of interest may reach lower levels before the user visits its TL, and may miss the user’s attention. Technically, a user can scroll through an indefinite number of posts at various levels. However, it is known that users’ attention is limited to the first few levels [24, 18]. We consider this aspect while analyzing viral-content propagation. We observe (theoretically as well as numerically) that this aspect makes a huge difference in the conclusions related to such a study. In addition, the TL structure also influences content propagation because of various other aspects: a) multiple posts with the same user (the attention gets divided); and b) decreasing interest towards reading contents at lower levels of TLs, etc. We discover that without taking these key elements into consideration, one may draw incorrect conclusions.

When the content of competing content providers circulates through the same social network and at the same point of time, many more factors influence content propagation. These influences are more complicated with TL structure; for example, a user might neglect the content of a low influential content provider when it (simultaneously) has a competing-content. Alternatively one might be interested in forwarding both contents, or the user might be interested only in forwarding the post that appears first. It might be possible that the content of a content provider gets viral, but not that of the others, etc. These aspects require the study of decomposable branching processes and the same along with the propagation of the competing content is considered in [8, Part-2].

Content propagation over OSNs follows a number of models based on factors such as empirical evidences, the structure of an OSN, etc. There has been an extensive literature on the content propagation over the OSNs, and an important approach for modeling the dynamics of content propagation has been branching processes (see [29, 31, 12, 11], etc). Authors in [15] studied information diffusion in the real viral marketing campaigns (involving 31000 individuals) and showed that the branching processes (BPs) explain the dynamics of information diffusion. The BPs are adequate to incorporate the characteristics of content diffusion (e.g., phase transition-epidemic threshold) and provide explicit expressions for many important performance measures. As an example, authors in [29] provided a discrete time branching model to predict the spread of a campaign. Using the theory of BPs, they estimated the performance of the campaign via various measures such as the number of forwarded e-mails, the number of viral e-mails, etc., as a function of system parameters. Other studies (for example, [26, 4]), reinforce that the BP can well fit the content propagation trajectories collected from real data. Authors in [13] showed that posts can get viral either due to broadcast type of forwarding or due to the branching type of forwarding, using a large empirical study of viral posts; they further showed that the viral cascades can also result from a combination of the two varieties. We hence use BPs, in particular multi-type BPs, to model content propagation influenced by TLs.

Our approach and contributions

In a BP, a parent produces identically and independently distributed (IID) offsprings. When one models content propagation over an OSN as a BP, any parent should produce a number of offsprings that are identically distributed as that of the other parents and are independent of the offspring produced previously. Moreover, parents keep producing

offsprings even when the population explodes. This is possible only when the OSN has an infinite population. The OSNs such as Facebook or Twitter have a huge number of users, it is reasonable to assume that the OSNs have an infinite population (unbounded number of users). Further, when users have identically distributed number of friends, then the BPs can model the content propagation over OSNs. This simplifies modeling and analysis. We use Multi-Type Branching Processes (MTBPs) (for example, [20, 21]) to model the influence of TL structure on content propagation in OSNs. We further extract the realistic features of content sharing in a typical OSN and incorporate them appropriately in our model. The BPs can mimic most of the phenomenon that influences content propagation. For example, one can model the effects of multiple posts being forwarded to the same friend, multiple forwards of the same post, etc.

A post on a higher level on TL has a better chance of being read by the user. Posts of appealing nature (e.g., with irresistible offers) have a great chance of being in circulation (we quantify this chance using a post quality factor). Posts of similar nature appearing at lower levels of TL have smaller chances of appreciation, etc. To study all these factors, one needs to differentiate the TLs that have the ‘post’ at different levels, and this is possible only through multi-type BPs. The following are some elements of our approach:

- Using the well-known results of multi-type branching processes (MTBPs), we obtain closed-form expressions for some performance measures which provide insights for the performance of the campaign, e.g., visibility of the content, virality², etc. For the other measures, we either have approximate (time asymptotic) closed-form expressions or simple fixed point equations whose solution provides the required measures.
- We study the influence of various network (structural) parameters on content propagation. We also study the effects of posts sliding down the TLs due to network activity. As the mean number of friends (network activity) increases, one can expect content to spread more rapidly (monotonous behavior). Contrary to that, we discover non-monotonous behavior in the virality of content. This phenomenon is fundamentally due to TL structure.

The paper also includes an important and novel application of our content-propagation analysis. In the last part, we integrate online auctions into the viral marketing model. We study an optimization problem considering real-time bidding. We compare the study considering the TL structure to that neglecting the TL structure, for varying levels of network activity. Our observations are similar: drastically different optimizers and the analysis without TL structure fails to capture the relevant phase transitions.

Some initial results of this work appeared in [6]. The major new inclusions in this journal version are: a) the proofs were omitted in [6], which are included in the current paper; b) we include a theoretical comparison between the system that considers the effects of TLs and the model that neglects the same to draw important conclusions about the deficiency of the latter model; c) the elaborate application on viral marketing, real-time bidding, and optimal budget allocation is completely new; and finally, d) we provide a validation of our mathematical model using simulations on Twitter data set [33].

²A post is said to be viral if a large number of posts are already shared and if the number of users with posts is exploding with time, i.e., it grows exponentially fast with time.

1.2. Related work

The huge growth in the activity on the Internet has generated a wide interest in understanding content propagation on the Internet. Previously, peer-to-peer (P2P) networks, as in [10, 31], have played an important role in the communication of the content. While of late there has been a lot of interest in content propagation over OSNs, see [29, 15, 2, 4]. The P2P networks pull the required information from their peers, while in viral marketing, the information is pushed for marketing purposes. In viral marketing, one needs to keep pushing information by passing it on to seed users or initial nodes to keep the flow going on. We consider content propagation over OSNs like Facebook, Twitter, etc., where the information (called post/message) is again pushed, but involuntarily. Here the post/content is forwarded to few initial users, and the post gets viral based on the interest generated among the users and the extensive sharing.

There is a vast literature that studies the propagation of content over OSNs. Many models discretize the time and study content propagation across the discrete time slots, for example as in [2]. As argued in [5, 29, 15] and references therein, a continuous time version (events occur at continuously distributed random time instances) is a better model and we consider the same. In the majority of the works which primarily use graph-theoretic models, the information is spread at maximum to one user at any message forward event, see [2, 5, 4]. However, when a user visits an OSN (e.g., Facebook, Twitter, WhatsApp, Telegram, etc.), it reads multiple unread posts one after the other and forwards each of them to some/all of its friends, for example as in [11, 9]; it may not at all forward some of the posts or may rarely forward a post. In [9], authors discuss and analyze this type of content propagation, influenced by the structure/capacity of the timeline and where the dynamics are continued by similar forwarding/sharing of posts; they consider dynamics of post-propagation over a single timeline used by many (more relevant for public timelines), while we consider one timeline per person.

Authors in [29, 15], etc., study viral marketing problems, where the marketing message is pushed continuously via emails, banner advertisements, search engines, etc. This scenario allows multiple forwards of the same post, and is analysed using BPs. However, they do not consider the influence of other posts using the same medium, and the other effects of TLs. As already mentioned, these aspects majorly influence the analysis. In [12] the authors analyze content propagation using discrete time branching processes; they discuss several interesting aspects including offspring distribution (given the degree distribution), life time distribution, average depth of the trees etc. This paper focuses on aspects like the number of generations survived, but (as mentioned by authors themselves) does not consider continuous time aspects like: a) when the different members of a generation received the post; b) when they forwarded the post; and c) how the unread copies of the post evolve with time. They also do not consider timeline structure.

Branching processes have been used in analysing various types of networks, such as, polling systems [17] which have been used to model local area networks and P2P networks [30], etc. We use branching processes not only to study the time evolution of the contents of interest (extinction and viral growth) but also to provide a spatio-temporal description of the process. We model the evolution of the number of timelines that have a given content at a given level of the timeline (e.g., top of the timeline).

2. System description

We consider a giant OSN like Facebook, Twitter, VKontakte, etc., with a large number of users. In this paper, we track content of interest corresponding to one specific content provider. In the second part of this work in [8], we consider the contents of multiple (competing) content providers³. Users use these networks to connect to other users to share photos, news, events/activities taking place around them, commercial content, etc. We briefly refer to these pieces of information as a post. Recall that these posts appear at different levels (on the screen) based on their newness, for instance, News Feed on Facebook. When a user visits⁴ the OSN, they read the posts on their timeline and share a post, upon finding it appealing/useful, with some of their friends (users connected to them). In this sharing process, the post appears on the top level of the timelines of those friends with whom the post is shared. This brings about a change in the appearance of contents on the timelines of recipients of the post. Basically, the existing contents of these TLs shift one level down each (see Figure 1). And a user can share as many posts as it wants. The number of shares of a particular post by a particular user depends upon: a) the distribution of its number of friends; and b) the extent to which the user liked the post. And extensive sharing of the post amongst the users potentially makes the post viral. It is evident that the sharing of a post depends on how engaging the *content provider* (CP) designs its post. There are some more aspects which influence the content propagation. For example, users may become reluctant to read/share the contents on the lower levels of their TLs. When they see multiple posts of similar nature, they may appreciate few posts while the remaining ones receive reduced attention. We study all those aspects and the dynamics created by the actions (e.g., like, share, etc) of the users, which have a major impact on the propagation of the commercial content.

Continuous time branching processes

The continuous time branching processes (CTBPs) are often good candidates for modeling viral marketing models. We describe these processes briefly as follows (see [20] for more details). Let $X(0)$ be the number of initial particles in a CTBP. Each of these particles stays alive for an exponentially distributed time with parameter, say λ and then dies. The ‘death’ times of the particles are independent of the others. Hence, the first death occurs after exponentially distributed time with parameter $X(0)\lambda$. Upon its death, it produces a random number (say ζ) of offspring, which join the existing population. The number of particles immediately after the death of the first particle changes to $X(0) - 1 + \zeta$. The ‘death’ times are again exponentially distributed (by the memoryless property), and the process continues. It is well-known that (under certain assumptions) the BPs have a certain dichotomy: a) either the population gets extinct (death of last individual TL), or b) the population grows exponentially fast with time; and c) there is no third way (see [20]).

2.1. Dynamics of content propagation and branching process

The content propagation in a typical OSN is as follows. Let us say we are interested in the propagation of post- \mathbf{P} when the process starts with $X(0)$ number of seed TLs.

³The extension of this work is submitted separately for details see Part-2 in [8].

⁴The users ‘visit’ OSNs at random intervals of time and in each ‘visit’ it browses some/all new posts.

We track the post- \mathbf{P} for the first N levels of TLs. It is important to note that $X(0)$ posts remain unread until one of the corresponding users visits its TL (and possibly forwards/rejects the post). We call the number of such TLs as the *number of unread TLs* (NU-TLs). If a user, among $X(0)$, visiting its TL finds post- \mathbf{P} attractive, it reads the post and may share the same with a random number of its friends. And post- \mathbf{P} would be placed on the top level of the recipient TLs. As shown in Figure 2, the recipient TL has post- \mathbf{P} on the top, and the remaining posts shift down one level each. If some more

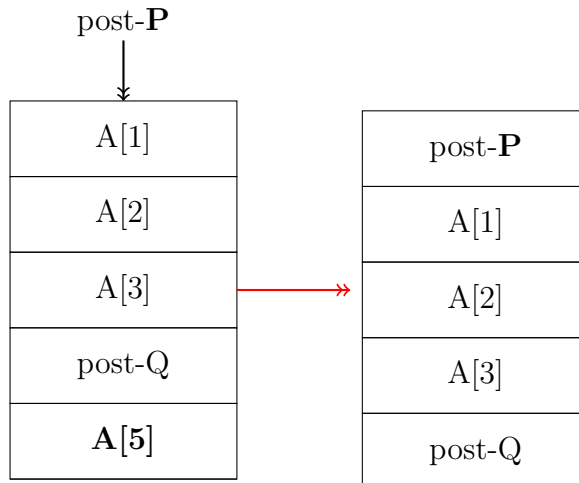


Figure 2: Shifting of the contents on a TL

posts are shared again with some of these recipient TLs, the contents further shift down the corresponding TLs. For instance, when one more post is shared after the post- \mathbf{P} with the same shared user, the post- \mathbf{P} resides on the second level of the corresponding TL.

We first argue that the continuous time version of the branching process fits the content propagation better than the discrete counterpart. In a CTBP, any one of the existing particles ‘dies’ after exponentially distributed time while in a discrete-time version all the particles of a generation ‘die’ together. When the number of copies of CP-post grows fast (i.e., when the post is viral), the time period between two subsequent changes decreases rapidly as time progresses. This is also well captured by CTBP, which mimics the content dynamics better. Also, the continuous time models are often considered a better choice for modeling information diffusion in social networks-related literature (see, e.g. [16, 32]).

As the underlying OSN is huge, one can say that the visit times of users (as defined in footnote 3) are virtually independent of each other. We assume memory-less visit times, i.e., the users visit their TLs at intervals that are exponentially distributed as in a CTBP. The sharing process generates a random number, say ζ , of new TLs holding post- \mathbf{P} . If the user does not read or share the post after visiting its TL, then $\zeta = 0$. If sharing process is independent and identical across the users, the new TLs ζ resemble IID offsprings in a CTBP, and the effective NU-TLs with post- \mathbf{P} appear like the particles of a CTBP. When one of the users of these NU-TLs (including the new ones) visits its TL and starts sharing the post- \mathbf{P} (as before), then the content propagation dynamics again resemble a CTBP.

However, the CTBP described above does not capture some aspects related to post-propagation process. Post- \mathbf{P} can disappear from some of the TLs, before the corresponding user’s visits. To be precise, the post- \mathbf{P} would disappear from a TL with $(N - l + 1)$ or more shares, if initially post- \mathbf{P} were at level l . For example, the post ‘ $\mathbf{A}[5]$ ’ is lost

by the arrival of post- \mathbf{P} in Figure 2. In all, the propagation of content in an OSN is influenced by two factors: a) the evolution of TLs with post- \mathbf{P} when some other posts are shared with them (contents on the TL shift down); and b) the sharing dynamics of post- \mathbf{P} between different TLs.

If we consider a CTBP with a single type of population, all the particles will have the same death rate and offspring distribution (for example, [20]). However, the disappearance of post- \mathbf{P} from a TL depends upon the level at which the post is available. Further, we will see that many more aspects of the dynamics depend upon the level at which the post- \mathbf{P} resides. Thus clearly, the single type CTBP is not sufficient, and we require a multi-type continuous time branching process (MTBP). An MTBP describes the population dynamics in the scenarios with a finite number of population-types. All the particles belonging to one type have the same death rate and offspring distribution; however, these parameters could be different across different types. To model the rich behaviour of the propagation dynamics, we will require (details in later sections) a particle of a certain type to produce offspring of other types. This modeling feature is readily available with MTBPs. We will show that the propagation dynamics can be well modeled by an appropriate MTBP, where for any $l \leq N$, all the TLs with post- \mathbf{P} in level l form one type of population. Additionally branching processes have an interesting feature: it suffices to study the evolution of the population with one initial/seed particle; to be more specific, the analysis starting with multiple seeds can be derived using the analysis with one seed particle (details are in later sections).

Assumptions: We track the post of the CP and study the time evolution of the post over TLs till first N levels. We assume a TL with posts of the CPs is not written⁵ with the post of the same CP again. In a huge social network, it reasonable to assume that the probability of the same post being shared again with any user is very small.

In [1], we consider a new variant of a BP that extends the study of content propagation without the above assumption. The new variant is called saturated total-population-size dependent BP, which caters to the fact that the users will not be interested in a post, when they receive the same post the second time, and so on.

3. Single content provider model

We consider a single CP and refer to its post as the CP-post. The TLs containing CP-post may have it at any level from one to N . These TLs also contain the other posts, and the movement of these posts can also affect the propagation of the CP-post. And our focus would be on CP-post. We say a user is of type l , if its TL contains the CP-post on level l and the top $l - 1$ levels do not contain the CP-post. Let $X_l(t)$ represent the number of unread TLs (NU-TLs) of type l at time t . We study the time evolution of $\mathbf{X}(t) := \{X_1(t), X_2(t), \dots, X_N(t)\}$. We will show below that the N -valued vector process $\mathbf{X}(t)$ is an MTBP under suitable conditions.

3.1. Modeling details

Birth-death process via shift and share transitions: To model the content propagation process by an appropriate BP, one needs to specify the ‘death’ of an existing

⁵We say a TL is written when a friend of it shares a post which changes its content.

parent (a TL with ‘unread’ CP-post in our case) and the distribution of its offsprings. A user of type l is said to ‘die’ either when its TL is written by another user or when the user itself wakes up (visits its TL) and shares the post with some of its friends. In the former event, exactly one user of type $(l + 1)$ (if $l < N$) is ‘born’ while the latter event gives birth to a random number of offspring of type 1 or 2 or $\dots N$ depending upon how many posts the user shares along with the CP-post. For example, if the user shares two posts say ‘message-X’ and ‘CP-post’ one after the other with its friends; then, CP-post will be on the top of those recipient TLs/users, while ‘message -X’ is on the second level of the same TLs. In general, if $l - 1$ ($l \leq N$) posts are shared with the same user after the CP-post, then the CP-post becomes available on the l -th level. We have new users with CP-post at l -th level of their TL, and thus type l offsprings are produced. Assume that offsprings of type l are produced with probability ρ_l , with $\rho_1 > 0$, $\rho_l \geq 0$ and $\sum_l \rho_l = 1$.

In general, users have a lethargy to view/read all the posts. We represent this via a level-based reading probability, r_l , which represents the probability that a typical user reads the post on level l . It is reasonable to assume $r_1 \geq r_2 \dots \geq r_N \geq 0$. We have two types of transitions that modify the MTBP, which we call *shift and share transitions*. In the share transition, a user first reads the CP-post and based on the interest generated, it shares CP-post with a random number of friends. Figure 3 below describes the share transition.

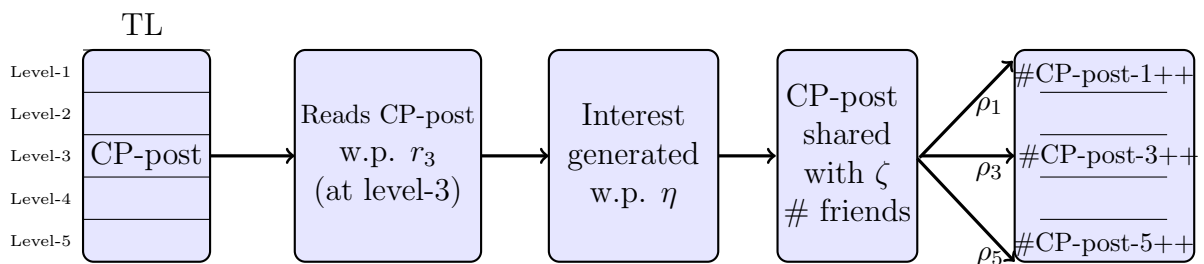


Figure 3: Share transition. Here, w.p. - ‘with probability’, # - number, and ++ - increase.

In the shift transition, the TL of the user with the CP-post is written by other users, and the position of CP-post shifts down.

CP-post propagation dynamics: Let \mathcal{G}_1 represent the subset of users with CP-post at some level, while \mathcal{G}_2 includes the others. The number of users in \mathcal{G}_1 at time t equals,

$$X(t) := \sum_{l \leq N} X_l(t). \quad (1)$$

We assume the OSN and hence \mathcal{G}_2 has infinitely many users, and this remains the same irrespective of the size of \mathcal{G}_1 , which is finite at any finite time. Thus, the transitions between \mathcal{G}_2 and \mathcal{G}_1 are more significant, and one can neglect the transitions within \mathcal{G}_1 . It is obvious that we are not interested in transitions within \mathcal{G}_2 (users without CP-post). We thus model the action of these groups in the following consolidated manner:

- **Share transition:** Each user from \mathcal{G}_1 wakes up after $\exp(\nu)$ time (exponentially distributed with parameter ν) to visit its TL and writes to a random (IID) number of users of \mathcal{G}_2 (refer to Figure 2).
- **Shift transition:** The TL of one of the users of \mathcal{G}_1 is written by one of the users of \mathcal{G}_2 (basically some one from \mathcal{G}_2 forwards a message and one among the recipients

is from \mathcal{G}_1), and the time intervals between two successive writes are exponentially distributed with parameter λ (refer to Figure 3).

The state of the network, $\mathbf{X}(t)$, changes when the first of the above-mentioned events occurs. At time t , we have $X(t)$ (see equation (1)) number of users in \mathcal{G}_1 and thus (first) one of them wakes up according to exponential distribution with parameter $X(t)\nu$. Similarly, the first TL/user of the group \mathcal{G}_1 is written with a post after exponential time with parameter $X(t)\lambda$. Thus, the state $\mathbf{X}(t)$, changes after exponential time with parameter $X(t)\lambda + X(t)\nu$. Thus, the rate of transitions at any time is proportional to $X(t)$, the NU-TLs at that time, and hence, the rate of transitions increase sharply as time progresses, if the post gets viral. Considering all the modeling aspects, the IID offspring vector generated by one l -type user are as below (w.p. means with probability):

$$\xi_l = \begin{cases} \mathbf{e}_{l+1}\mathbb{1}_{l < N} & \text{w.p. } \theta := \frac{\lambda}{\lambda + \nu} \text{ and} \\ \zeta \mathbf{e}_i & \text{w.p. } (1 - \theta)r_l \rho_i \quad \forall i \leq N \\ \mathbf{0} & \text{w.p. } (1 - \theta)(1 - r_l). \end{cases} \quad (2)$$

where \mathbf{e}_l represents standard unit vector of size N with one in the l -th position, $\mathbf{0} = (0, \dots, 0)$, $\mathbb{1}_A$ represents the indicator, ζ is the random number of friends to whom the post is shared and r_l is the probability with which the user reads/views a post on level l . The first row in (2) is due to shift transitions and the second row is due to a share, while the last row is because the user did not read the post at level l . Figure 4 demonstrates the transitions.

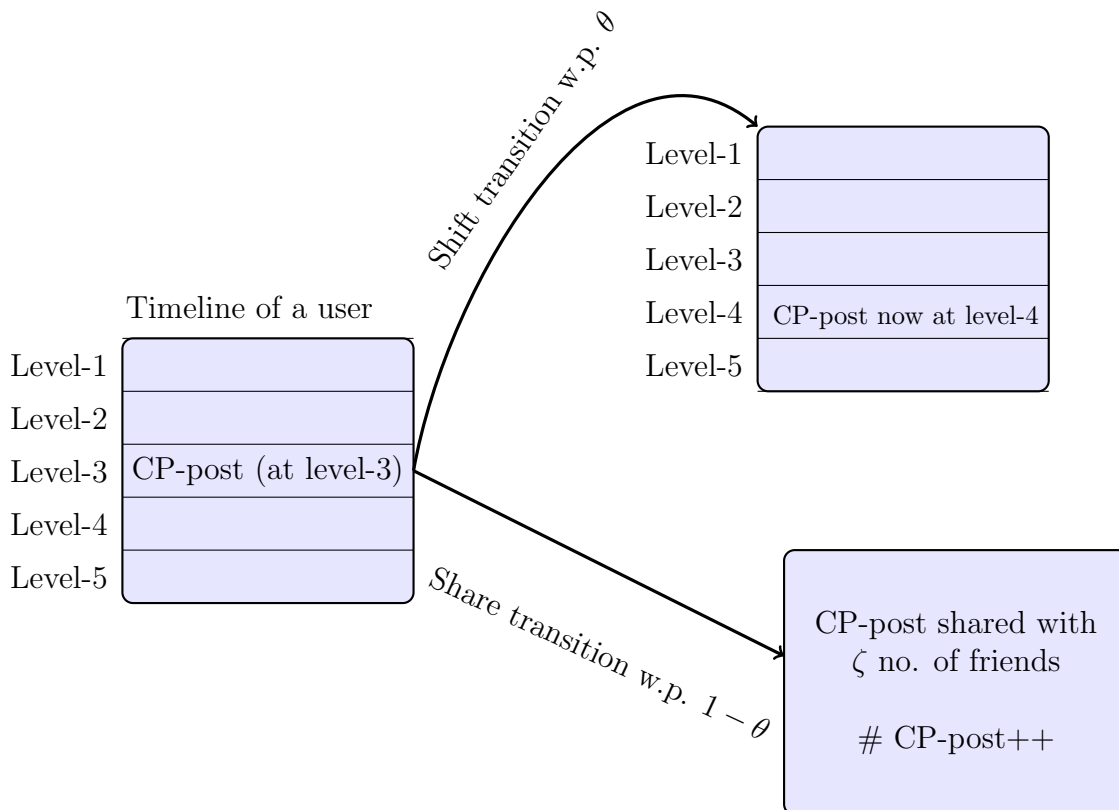


Figure 4: Propagation of CP-post: transitions

Recall that users (offsprings) of type i are produced with probability ρ_i during the share transitions. From equation (2), the offspring distribution is identical at all time instances t , ζ can be assumed independent across users, and hence ξ_l are IID offsprings

from any type l user. Further, all the transitions occur after memoryless exponential times, and hence $\mathbf{X}(t)$ is an MTBP with N -types (e.g. [21]).

PGFs and post quality factor: Let $f_F(s, \beta)$ be the probability generating function (PGF) of the number of friends, \mathbb{F} , of a typical user, parametrized by β . For example, $f_F(s, \beta) = \exp(\beta(s - 1))$ is for Poisson distributed \mathbb{F} , while, $f_F(s, \beta) = (1 - \beta)/(1 - \beta s)$ is for geometric \mathbb{F} . Let $m = f'_F(1, \beta)$ (derivative w.r.t. s) represent the corresponding mean. A user shares the post with some/all of its friends (ζ of equation (2)) based on how engaging the post is. Let the post quality factor η quantify the extent of the CP-post engagement on a (continuous) scale of 0 to 1 where $\eta = 0$ means the worst and $\eta = 1$ is the best quality. We assume that the mean of the number of shares is proportional to this quality factor. In other words, $m(\eta) = \tilde{m}\eta$ represents the post quality dependent mean of the random shares for some $\tilde{m} > 0$. Let $f(s, \eta, \beta)$ represent the PGF of ζ , the number of shares. When \mathbb{F} is Poisson distributed, the above PGF and the expected value $E[\zeta]$ respectively equal:

$$f(s, \eta, \beta) = f_F(s, \eta\beta) = \exp(\beta\eta(s - 1)) \text{ for any } s \text{ and } m(\eta) = \eta\beta.$$

When \mathbb{F} is Geometric distributed, one may assume the post quality dependent parameter

$$\beta_\eta = (1 - \beta)/(1 - \beta + \beta\eta), \text{ which ensures } m(\eta) = \eta\beta.$$

And then the PGF of ζ is given by $f(s, \eta, \beta) = f_F(s, \beta_\eta) = (1 - \beta_\eta)/(1 - \beta_\eta s)$. One can derive such PGFs for other distributions of \mathbb{F} . *Interestingly enough, we find that most of the analysis does not depend upon the distribution of \mathbb{F} but only on its expected value.*

Let $\mathbf{s} := (s_1, \dots, s_N)$ and $\bar{\mathbf{f}}(\mathbf{s}, \eta) := \sum_{i=1}^N f(s_i, \eta, \beta)\rho_i$ (β is dropped in $\bar{\mathbf{f}}(\mathbf{s}, \eta)$ for simpler notation). The post quality factor dependent PGF, of the offspring distribution of the overall BP, is given by (see equation (2)):

$$\boxed{h_l(\mathbf{s}) = \theta (s_{l+1}\mathbb{1}_{l < N} + \mathbb{1}_{l=N}) + (1 - \theta)r_l\bar{\mathbf{f}}(\mathbf{s}, \eta) + (1 - \theta)(1 - r_l)}. \quad (3)$$

The first term is due to shift transitions by which the type changes to $l + 1$ (or dies when $l = N$), the second term is due to new offsprings, while the last term is due to the possibility that the user does not read the post at level l (see equation (2)).

3.2. Generator matrix

The key ingredient required for the analysis of any MTBP is its generator matrix. We begin with the generator for MTBP that represents the evolution of unread TLs with CP-post. We refer to this process briefly as TL-CTBP, timeline continuous-time branching process. The generator matrix, A , is given by $A = (a_{lk})_{N \times N}$, where,

$$a_{lk} = a_l \left(\partial h_l(\mathbf{s}) / \partial s_k \Big|_{\mathbf{s}=\mathbf{1}} - \mathbb{1}_{\{l=k\}} \right),$$

and a_l represents the transition rate of a type- l particle (see [21] for details). For our case, from previous discussions $a_l = \lambda + \nu$ for all l . Further, using equation (3), the matrix A for our single CP case is given by (with $c := (1 - \theta)m\eta$, $c_l = c\rho_l$)

$$A = (\lambda + \nu) \begin{bmatrix} c_1 r_1 - 1 & c_2 r_1 + \theta & \cdots & c_{N-1} r_1 & c_N r_1 \\ c_1 r_2 & c_2 r_2 - 1 & \cdots & c_{N-1} r_2 & c_N r_2 \\ & \vdots & & & \\ c_1 r_{N-1} & c_2 r_{N-1} & \cdots & c_{N-1} r_{N-1} - 1 & c_N r_{N-1} + \theta \\ c_1 r_N & c_2 r_N & \cdots & c_{N-1} r_N & c_N r_N - 1 \end{bmatrix}. \quad (4)$$

For example, the second element of the first row of the above matrix is obtained by differentiating $h_1(\cdot)$ of (3) with s_2 , which equals (when $N > 2$)

$$\theta + (1 - \theta)r_1 \left. \frac{\partial \bar{\mathbf{f}}(\mathbf{s}, \eta)}{\partial s_2} \right|_{\mathbf{s}=\mathbf{1}} = \theta + (1 - \theta)r_1 \left. \frac{\partial f(s_2, \eta, \beta)}{\partial s_2} \right|_{\mathbf{s}=\mathbf{1}} \rho_2$$

and then substituting $\mathbf{s} = \mathbf{1}$; also observe that $\partial f(s_2, \eta, \beta)/\partial s_2|_{\mathbf{s}=\mathbf{1}} = m\eta$. The largest eigenvalue and the corresponding eigenvectors of the above generator matrix are instrumental in obtaining the analysis of TL-CTBP [21] and the following lemma establishes important properties about the same. We also prove that the resulting TL-CTBP is positive regular⁶, which is an important property that establishes the simultaneous survival/extinction of all types of TLs.

Lemma 1. *i) When $0 < \theta < 1$, the matrix e^{At} (A in (4)) for any t is positive regular.*
ii) Let α be the maximal real eigenvalue of the generator matrix A and let inner product $\mathbf{r} \cdot \mathbf{c} := \sum_{i=1}^N r_i c_i$. Then $\alpha \in (\mathbf{r} \cdot \mathbf{c} - 1, \mathbf{r} \cdot \mathbf{c} - 1 + \theta)(\lambda + \nu)$. When the reading probabilities have special form $r_l = d_1 d_2^l$ (for some $0 \leq d_1, d_2 \leq 1$), then

$$\alpha \rightarrow (\mathbf{r} \cdot \mathbf{c} - 1 + \theta d_2)(\lambda + \nu) \text{ as } N \rightarrow \infty.$$

iii) The left and right eigenvectors \mathbf{u}, \mathbf{v} corresponding to α satisfy the following equations $c_1 \mathbf{r} \cdot \mathbf{u} = \sigma u_1$ and $c_1 \mathbf{r} \cdot \mathbf{v} = \sigma v_N$ where $\sigma := \alpha/(\lambda + \nu) + 1$. We have

$$u_l = \sum_{i=0}^{l-1} \frac{\rho_{l-i}}{\rho_1} \left(\frac{\theta}{\sigma}\right)^i u_1, \quad 2 \leq l \leq N \quad \text{and} \quad v_l = \sum_{i=0}^{N-l} \left(\frac{\theta}{\sigma}\right)^i \frac{r_{l+i}}{r_N} v_N, \quad 1 \leq l \leq N-1.$$

Proof: *The proof is given in Appendix. ■*

At the finest details, we now developed a full-fledged MTBP that models the content propagation. The multitype continuous time branching processes (MTBPs) are well studied in the literature (for example, [20, 21]). The analysis of MTBP largely depends upon its generator matrix. Lemma 1 describes the characteristics of the generator matrix specific to our model. It results in positive regularity of TL-CTBP, i.e., the generator matrix A (4) is positive regular. The largest eigenvalue α of A characterizes the growth rate of NU-TLs. We later see that left and right eigenvectors, \mathbf{u} and \mathbf{v} (corresponding to α) characterize the visibility of the CP-post. Using the characterizations of Lemma 1 and the rich theory of MTBPs, we derive various performance measures specific to this content propagation.

Before we proceed, we would like to elaborate on the special form considered for reading probabilities in Lemma 1. In this special case, we assume that the reading probabilities decrease with the level of TL in a geometric manner ($r_l = d_1 d_2^l$ for any l). This special form aids us in deriving more insights into the problem and we believe this is a reasonable assumption.

The CP would be interested in many related performance measures as a function of the post quality factor and we consider the same in the next section.

⁶A matrix B is called positive regular (irreducible) if there exists an n such that the matrix B^n has all strict positive entries. A BP is positive regular when its mean matrix is positive regular. With A as generator, the positive regularity is guaranteed if e^A is positive regular (e.g. [21]).

3.3. Performance analysis

If the CP invests sufficiently in preparing the content/post and ensures good quality, the post can get viral. To be more precise, when it designs a post such that η is above the virality threshold (see (7)), then it has positive probability of not getting extinct or getting viral. The relevant central questions related to a BP which are also relevant to our content propagation process include: a) what is the extinction probability, i.e., the probability with which the entire population gets extinct?; b) what is the rate at which the population grows?; c) what is the total progeny? etc. We apply the well-known results to address the above questions in our context and derive some performance measures. We employ fixed point techniques to obtain the other performance measures. We begin with the probability of extinction.

We conclude this sub-section by making an important observation: the overall evolution of the post depends on the number of initial/seed TLs with CP-post, however, it is sufficient to consider the analysis with one seed TL. This is because of the following properties of the BPs (e.g., [20]): i) the growth rate does not change with the number of seeds; ii) the extinction probability with multiple seeds can be derived using that for single seed user. For example, for single type process, the extinction probability with s -seed users equals, $q_s = q_1^s$, where q_1 is the same probability with 1 seed user.

3.3.1. Extinction probabilities

Depending upon the context of the problem, for instance, an awareness campaign, the CP may be interested in knowing the chances of dissemination of its information to a large population, i.e., the chance of virality of its post. This probability can be obtained directly using the extinction probability of the corresponding MTBP, as explained below. The CP-post is said to be extinct when it disappears completely off the OSN, i.e., none of the N -length TLs contain the CP-post eventually (as time progresses). Let q_l be the probability with which the process gets extinct when TL-CTBP starts with one TL of type l ,

$$q_l := P(\mathbf{X}(t) = \mathbf{0} \text{ for some } t > 0 | \mathbf{X}(0) = \mathbf{e}_l).$$

Let $\mathbf{q} := \{q_1, q_2, \dots, q_N\}$ represent the vector of extinction probabilities.

Under positive regularity conditions of Lemma 1.(i), when a BP is not extinct, the population grows exponentially fast to infinity (see [28, 21], etc). This fact is established for our TL-CTBP in Theorem 1, provided in the later subsections. *Thus, we have a dichotomy: either the post gets viral at an exponential rate, or otherwise, it dies completely (gets extinct).* And hence the extinction probability equals one minus the probability of virality. We now have the following with notations as defined in Lemma 1.

Lemma 2. *Assume $0 < \theta < 1$ and $E[\mathbb{F} \log \mathbb{F}] < \infty$ with $\mathbb{F} \log(\mathbb{F}) := 0$ when $\mathbb{F} = 0$. Then clearly $E[\zeta \log \zeta] < \infty$ for any post quality factor η . Hence we have the following:*

- (i) *If $\alpha \leq 0$, extinction occurs w.p.1, i.e., $\mathbf{q} = \mathbf{1} = (1, \dots, 1)$;*
- (ii) *If $\alpha > 0$, then⁷ $\mathbf{q} < \mathbf{1}$, i.e., the post gets viral with positive probability irrespective of the type of the seed TL. In this case the extinction probability vector \mathbf{q} is the unique solution of the equation, $h(\mathbf{s}) = \mathbf{s}$, and lies in the interior of $[0, 1]^N$.*

⁷Vector $\mathbf{q} < \mathbf{s}$ if $q_i < s_i$ for all components i .

Proof It follows from [21, Theorems 1-2] and by Lemma 1. ■

It is easy to verify that the hypotheses of this lemma are easily satisfied by many distributions. For example, Poisson, Geometric, etc., satisfy $E[\mathbb{F} \log \mathbb{F}] < \infty$.

By Lemma 2.(ii) the extinction probabilities are obtained by solving $h(\mathbf{s}) = \mathbf{s}$. The extinction probability can be obtained by conditioning on events and is given as below when the process starts a type- l TL:

$$q_l = \theta \left(q_{l+1} \mathbb{1}_{\{l < N\}} + \mathbb{1}_{\{l=N\}} \right) + (1 - \theta) r_l \bar{\mathbf{f}}(\mathbf{q}, \eta) + (1 - \theta)(1 - r_l). \quad (5)$$

One can simplify the above set of equations starting from $l = N$ and then via backward induction to prove that the below set of equations solve (5):

$$q_{N-l} = (q_N - 1) \sum_{i=0}^l \theta^{l-i} \frac{r_{N-i}}{r_N} + 1 \text{ for any } 1 \leq l < N. \quad (6)$$

The solution of the above provides the extinction probabilities.

Virality Threshold: By Lemma 2.(ii) the CP-post *gets viral, i.e., the TL-CTBP survives and explodes with non-zero probability*, when $\alpha > 0$. When N is sufficiently large, by Lemma 1.(ii) and Lemma 2.(ii),

$$\alpha \approx (m\eta(1 - \theta)\boldsymbol{\rho} \cdot \mathbf{r} - 1 + \theta d_2)(\lambda + \nu) = (m\eta\boldsymbol{\rho} \cdot \mathbf{r} - 1)\nu - (1 - d_2)\lambda. \quad (7)$$

It is well-known that the BPs survive with positive probability if the largest eigenvalue of the generator matrix, A , is positive (supercritical process). We have a similar statement for large N , i.e., such TL-CTBPs can survive when $m\eta(1 - \theta)\boldsymbol{\rho} \cdot \mathbf{r} > 1 - \theta d_2$ (see equation(7)), for a BP pitted against the shifting process. The virality threshold, denoted by $\bar{\eta}$, is defined in terms of network parameters and is given by

$$\bar{\eta} > \frac{1 - \theta d_2}{m(1 - \theta)\boldsymbol{\rho} \cdot \mathbf{r}}. \quad (8)$$

Thus, the virality chances are influenced by post quality η , shift factor $(1 - \theta)$, by the types of posts produced as given by $\boldsymbol{\rho}$, the mean number of friends m and the reading probabilities \mathbf{r} . In effect, the virality chances are influenced by factor $(1 - \theta)\eta\boldsymbol{\rho} \cdot \mathbf{r}$.

No-TL Case: What if all the effects of the TLs were neglected?

The majority of the works as in [15, 29], consider a study of content propagation without considering TL structure, and as mentioned before, such a study is incomplete in networks where the TLs significantly influence the propagation. This is more so in the networks where the activity is significant and the TL of any user is usually written with a lot of new content before the next visit of the user to the network. We would like to compare our conclusions with the case when the effects of TLs are neglected in such networks. Without considering the TL structure, there will be:

1. No notion of post residing at various levels, i.e., all posts reside at one level only, and so $N = 1$; consequently its reading probability is one ($r_1 = 1$), and further, $\rho_1 = 1, \rho_i = 0 \forall i > 1$.
2. No notion of shifting effect, consequently $\theta = 0$ which is equivalent to saying $\lambda = 0$.

The remaining modeling details of the content propagation are the same as before. In view of this, it is evident that the content propagates according to a single type continuous time Markov branching process. Thus, the analysis of this case boils down to a special case of the TL model (with $\lambda = 0, N = 1, \rho_1 = 1, r_l = 1 \forall l$). For this special case, from equation (7), the rate of growth say α_{No-TL} is given by:

$$\alpha_{No-TL} \approx (m\eta - 1)\nu.$$

Observe that the post gets viral when $m\eta > 1$, as is well understood in branching and viral marketing literature [29]. However, as mentioned before this neglects the key aspects of content propagation—effects of TLs. It is accompanied by an erroneous conclusion that the virality chances are influenced by m and η only. While in reality there is additional influence, which is summarized by factor $(1 - \theta)\eta\rho \cdot \mathbf{r}$ (see equation (8)).

In No-TL case, the extinction probability is obtained by solving $q = f(q, \eta)$ (substituting the parameter values in equation (5)). It again becomes evident that the effect of post-residing at various levels is disappeared. The extinction probability is the same, whether it is started with one CP-post on level 1 or level 9. This is again a wrong interpretation and the solutions of the equation (6)/(5) provide the correct extinction probabilities when one considers the influence of TLs as in this work.

Influence of the Network Connectivity on Extinction Probability: When the mean number of friends $m = E[F]$ is small, the majority of users in the network are connected to few friends, and one can imagine a network that is sparsely connected. On the other hand, when $m = E[F]$ is large, a sizable number of users have a large number of friends and we probably have a better-connected network. We study the impact of network connectivity on extinction probability by studying its dependency on m . When mean m increases, the network becomes more active as sharing different posts becomes more pronounced. The TLs are flooded with different posts rapidly, so do the TLs containing post-**P**, and one might anticipate an increase in its virality chances. However, these TLs also receive the other posts rapidly, resulting in rapid shifts to their contents. Thus, with an increase in m , the λ increases, and so does θ . We observe an interesting phenomenon in Figure 5, with respect to the virality chances $1 - q_\rho$ with $q_\rho := \sum_l q_l \rho_l$, when λ is set proportional to mean m . To begin with, the virality chances $1 - q_\rho$ improve (q_ρ decreases) with mean m , as anticipated. However, if one increases m further, we notice an increase in q_ρ . Basically, increased m implies more shares of post-**P** to new users *but it also implies post-**P** is missed more often*. This phenomenon is mainly observed because of timeline structure: when TL structure is neglected, any user will view all the posts with equal interest irrespective of their levels. And consequently, one would not have noticed the effect of m on extinction probability (as in No-TL case). There seems to be *an optimal number of mean friends, which is best suited for post propagation*.

3.3.2. Time evolution of the NU-TLs

The number of unread timelines (NU-TLs), at various time instances, may serve as an indicator of the reach of the CP-post. The reach of CP-post is another yardstick of the campaign's effectiveness. Here, we obtain the time evolution of NU-TLs. We have the following theorem which is instrumental in obtaining the expected number of NU-TLs in the viral scenario:

Theorem 1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an appropriate probability space and let $\{\mathcal{F}_t\}$ be the natural filtration for TL-CTBP $\mathbf{X}(\cdot)$, i.e., for each t , \mathcal{F}_t is the σ -algebra generated by $\{\mathbf{X}(\mathbf{t}'); \mathbf{t}' \leq t\}$.*

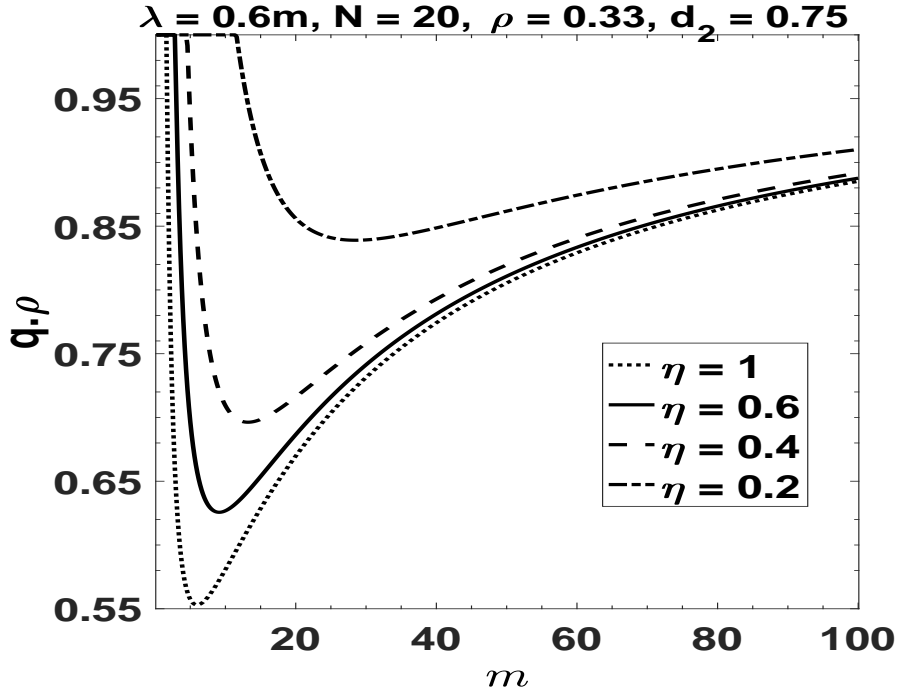


Figure 5: Extinction vs $m = E[\mathbb{F}]$, where $q_\rho = \mathbf{q} \cdot \rho$. Little benefit in producing high quality content in OSNs with high mean degree!

\mathbf{t} }. The process $\{\mathbf{v} \cdot \mathbf{X}(\mathbf{t}) e^{-\alpha \mathbf{t}}; \mathbf{t} \geq \mathbf{0}\}$, with \mathbf{v}, α as in Lemma 1, is a non negative martingale (with natural filtration) and,

$$\lim_{t \rightarrow \infty} \mathbf{X}(\mathbf{t}, \omega) e^{-\alpha \mathbf{t}} = \mathbf{W}(\omega) \mathbf{u} \text{ for almost all } \omega, \quad (9)$$

where W is a non negative random variable (depends upon l and other factors) that satisfies⁸: $P_l(W = 0) = q_l$, $E_l[W] = v_l$ for each l , with $\mathbf{u} \cdot \mathbf{v} = 1$.

Proof: Under the assumptions of Lemma 2, the TL-CTBP satisfies the hypotheses of Theorem 1 of [21]. ■

The CP-post gets extinct on the sample paths with $W = 0$ in equation (9) (see [21] for details, also observe that $P_l(W = 0) = q_l$ from the above theorem). It gets viral in the complementary paths, i.e., when $W > 0$, as is also evident from the limit⁹ given by equation (9).

On the viral paths, we have two important measures: 1) the growth rate α , and 2) the visibility of the post. The growth rate characterizes the rate at which the post spreads through the OSN. From (9), the TLs grow exponentially fast with time at the rate α (given by equation (7)), i.e., according to $e^{\alpha t}$. And the other measure, *the visibility of the post* can be determined by the number of *potential* users that can read the post and

⁸ We use E_l and P_l to represent the conditional expectation and probability respectively when TL-CTBP starts with one l -type TL.

⁹Note that for large t , the NU-TLs $\mathbf{X}(\mathbf{t}, \omega) \approx \mathbf{W}(\omega) \mathbf{u} e^{\alpha t}$, which grows exponentially fast when $W(\omega) > 0$.

thereby get influenced to buy the product/service. Recall that users attention is limited to the first few number of levels. Clearly, the visibility of the post depends on the level at which it resides on the NU-TLs. The more the number of TLs having post on higher levels, the more the visibility. The number of potential users viewing the post on the level l is approximately $r_l u_l e^{\alpha t}$ (large t) where u_i is the i -th component of vector \mathbf{u} . We define the visibility of the post at level say l as the *fraction* of NU-TLs holding the post at level l after a long time t which is given as $u_l / \sum_i u_i$.

We also obtain the time evolution of the *expected value of NU-TLs*. This result is obtained as a corollary of Theorem 1.

Corollary 1. *When $\alpha > 0$ and starting with one type- i seed TL, $\sum_{l=1}^N E_i [X_l(t)] = e^{\alpha t} v_i \sum_l u_l$. Further, when $r_i = d_1 d_2^i$, $\rho_i = \tilde{\rho}^i$ (with $\sum_i \rho_i = 1$ and $0 < \rho \leq 1$) for all i , we have*

$$\sum_{l=1}^N E [X_l(t)] = \varrho e^{\alpha t} d_2^{i-1} \text{ with } \varrho := (1 - d_2 \rho) \left(\frac{1}{1 - \rho} - \frac{\theta}{\rho} \frac{1}{\sigma - \theta} \right) \frac{(\sigma - \theta d_2)(\sigma \rho - \theta)}{(\sigma - \theta)(\rho - \theta)}. \quad (10)$$

Proof: Using the fact that $v \cdot X(t) e^{-\alpha t}$ is a martingale and $\mathbf{u} \cdot \mathbf{v} = \mathbf{1}$, one can write the following

$$\begin{aligned} E [\mathbf{v} \cdot \mathbf{X}(t) e^{-\alpha t}] &= E [v \cdot X(0) e^{-\alpha \times 0}] = v_i; \quad \mathbf{u} \cdot \mathbf{E} [\mathbf{v} \cdot \mathbf{X}(t) e^{-\alpha t}] = \mathbf{u} \mathbf{v}_i \quad (\text{as } \mathbf{X}_i(0) = \mathbf{1}) \\ \mathbf{u} \cdot \mathbf{E} [\mathbf{v} \cdot \mathbf{X}(t) e^{-\alpha t}] &= \mathbf{u} \cdot \mathbf{v} \mathbf{E} [\mathbf{X}(t) e^{-\alpha t}] = \mathbf{E} [\mathbf{X}(t) e^{-\alpha t}] = \mathbf{u} \mathbf{v}_i \quad \because \mathbf{u} \cdot \mathbf{v} = \mathbf{1}. \end{aligned}$$

Thus $E [\mathbf{X}(t)] = \mathbf{u} \mathbf{v}_i e^{\alpha t}$. Further, by taking the sum of individual component of the expected value of the random vector

$$\sum_{l=1}^N E [X_l(t)] = E \left[\sum_{l=1}^N X_l(t) \right] = e^{\alpha t} v_i \sum_l u_l, \quad \text{because } N \text{ is finite.}$$

Substituting the value of $v_i \sum_l u_l$ from equation (34) in Appendix, we get the desired result

$$\sum_{l=1}^N E [X_l(t)] = e^{\alpha t} d_2^{i-1} (1 - d_2 \rho) \left(\frac{1}{1 - \rho} - \frac{\theta}{\rho} \frac{1}{\sigma - \theta} \right) \frac{(\sigma - \theta d_2)(\sigma \rho - \theta)}{(\sigma - \theta)(\rho - \theta)} = \varrho e^{\alpha t} d_2^{i-1}. \quad \blacksquare$$

We observe that the rate of increase of the expected NU-TLs is again α , however the constant is determined by d_2 (the factor by which the reading probability decreases with level), decreases geometrically with level i of the parent TL and ρ (a factor that determines the number of posts simultaneously forwarded).

3.3.3. Time evolution of the number of shares

We derive another important performance measure, the expected number of shares of the post, before a given time t . This measure gives the total spread of the post, i.e., the total number of shares a post gets in the given time-frame (e.g., the number of shares in Facebook). It is basically the total number of distinct TLs (i.e., users) that received a copy of the post before time t . *It is important to observe here that ‘number of shares’*

is different from the well-known ‘total progeny’¹⁰ of the underlying BP. The ‘number of shares’ is due to offspring generated by share transition only, while the ‘total progeny’ is due to both ‘share’ as well as ‘shift’ transition offspring. We discuss the number of shares in viral ($\mathbf{q} < \mathbf{1}$) and non viral (sure extinction) scenario.

Viral scenarios: We employ probability generation based technique to obtain the time evolution of number of shares. Let $Y(t)$ be the accumulated number of shares till time t and let $Y = \lim_{t \rightarrow \infty} Y(t)$ (can also be infinity) be the eventual number of shares. The following Lemma captures the time evolution of number of shares.

Lemma 3. *Let $\mathbf{y}(t) := [y_1(t) \cdots, y_N(t)]$ with $y_l(t) := E_l[Y(t)] = E[Y(t) | \mathbf{X}(0) = \mathbf{e}_l]$, the expected number of shares till time t when started with one l -type TL for each l . If $\alpha > 0$, we have*

$$\mathbf{y}(t) = e^{At} \left(\mathbf{1} + (\lambda + \nu) \mathbf{A}^{-1} \mathbf{k} \right) - (\lambda + \nu) \mathbf{A}^{-1} \mathbf{k} \text{ where } \mathbf{k} = [\mathbf{1} - \theta, \mathbf{1} - \theta, \dots, \mathbf{1} - \theta, \mathbf{1}]^T. \quad (11)$$

Proof: *The proof is given in Appendix.* ■

Thus, the expected number of shares grow exponentially fast with time for viral scenarios. Further, the growth rate α (see eqn. (7)) is the same as that for the unread posts. From (38) of Appendix, for large t , the expected shares when started with one type- l particle is:

$$y_l(t) \approx e_{l,0} e^{\alpha t} \text{ with } e_{l,0} = v_l \sum_{i=1}^N u_i \left(1 + \frac{\nu}{\alpha} \right) + v_l \frac{\lambda}{\alpha} u_N. \quad (12)$$

Non viral scenarios: When population gets extinct with probability one, the expected number of total shares is finite. One can directly obtain the expected number of shares by conditioning on the first transition event as follows:

$$y_l := E_l[Y] = \theta y_{l+1} \mathbb{1}_{\{l < N\}} + (1 - \theta) r_l (m\eta + m\eta \mathbf{y} \cdot \boldsymbol{\rho}); \quad \text{for all } l \leq N. \quad (13)$$

On recursively simplifying the above system of equations backward, we obtain the following for any $l \leq N$:

$$y_l = (1 - \theta) m\eta (1 + \mathbf{y} \cdot \boldsymbol{\rho}) \sum_{i=0}^{N-l} \theta^{N-l-i} r_{N-i}. \quad (14)$$

Summing the above over l after multiplying with ρ_l , we obtain:

$$\mathbf{y} \cdot \boldsymbol{\rho} = \sum_{l=1}^N \rho_l y_l = (1 - \theta) m\eta (1 + \mathbf{y} \cdot \boldsymbol{\rho}) \sum_l \rho_l \sum_{i=0}^{N-l} \theta^{N-l-i} r_{N-i}.$$

Thus, the FP equation for $\mathbf{y} \cdot \boldsymbol{\rho}$ is linear and hence we have a unique FP solution for $\mathbf{y} \cdot \boldsymbol{\rho}$ whenever

$$(1 - \theta) m\eta \sum_l \rho_l \sum_{i=0}^{N-l} \theta^{N-l-i} r_{N-i} < 1.$$

¹⁰The total number of offspring produced so far, by the BP.

If $(1 - \theta)m\eta\mathbf{r}\cdot\boldsymbol{\rho} - 1 + \theta = \mathbf{r}\cdot\mathbf{c} - 1 + \theta < 0$, from Lemma 1.(ii) $\alpha < 0$ and the process would be extinct w.p. one. In this scenario:

$$\begin{aligned} (1 - \theta)m\eta \sum_l \rho_l \sum_{i=0}^{N-l} \theta^{N-l-i} r_{N-i} &\leq (1 - \theta)m\eta \sum_l \rho_l r_l \sum_{i=0}^{N-l} \theta^{N-l-i} \\ &= (1 - \theta)m\eta \sum_l \rho_l r_l \frac{1 - \theta^{N-l-1}}{1 - \theta} = m\eta\mathbf{r}\cdot\boldsymbol{\rho} < 1 \end{aligned}$$

because $r_1 \geq r_2 \cdots \geq r_N$. We can similarly show using the limit of the eigenvalue α of Lemma 1, that when the process is extinct w.p. one, the above condition is always satisfied asymptotically. To be more precise the condition is satisfied for all N bigger than a threshold \bar{N} , whenever the process is extinct w.p. one.

We thus have the following unique FP for $\mathbf{y}\cdot\boldsymbol{\rho}$ under the conditions discussed above:

$$\mathbf{y}\cdot\boldsymbol{\rho} = \frac{(1 - \theta)m\eta \sum_l \rho_l \sum_{i=0}^{N-l} \theta^{N-l-i} r_{N-i}}{1 - (1 - \theta)m\eta \sum_l \rho_l \sum_{i=0}^{N-l} \theta^{N-l-i} r_{N-i}}. \quad (15)$$

One can substitute the above in equation (14) to obtain y_l for all l :

$$y_l = \frac{(1 - \theta)m\eta \sum_{i=0}^{N-l} \theta^{N-l-i} r_{N-i}}{1 - (1 - \theta)m\eta \sum_l \rho_l \sum_{i=0}^{N-l} \theta^{N-l-i} r_{N-i}}. \quad (16)$$

Also, it is easy to verify that the FP is unique, by uniqueness of the FP solutions for $\mathbf{y}\cdot\boldsymbol{\rho}$. Note that in No-TL case, the number of shares is computed using

$$y = E[Y] = m\eta(1 + y) \implies y = \frac{m\eta}{1 - m\eta} \quad (17)$$

which is again inaccurate.

Special case: Say $r_i = d_1 d_2^i$, $\rho_i = \tilde{\rho} \rho^i$ (with $\sum_i \rho_i = 1$ and $0 < \rho \leq 1$) for all i , one can easily simplify the above. We have the following

$$\begin{aligned} \sum_l \rho_l \sum_{i=0}^{N-l} \theta^{N-l-i} r_{N-i} &= d_1 \tilde{\rho} \sum_l \rho^l \sum_{i=0}^{N-l} \theta^{N-l-i} d_2^{N-i} = d_1 \tilde{\rho} \sum_l \rho^l d_2^l \sum_{i=0}^{N-l} \theta^{N-l-i} d_2^{N-l-i} \\ &= d_1 \tilde{\rho} \sum_l \rho^l d_2^l \sum_{i=0}^{N-l} \theta^i d_2^i = d_1 \tilde{\rho} \sum_l \rho^l d_2^l \frac{(1 - (\theta d_2)^{N-l+1})}{1 - \theta d_2} \\ &= \frac{d_1 \tilde{\rho}}{1 - \theta d_2} \left(\rho d_2 \sum_{l=0}^{N-1} \rho^l d_2^l - d_2 (\theta d_2)^N \rho \sum_{l=0}^{N-1} \rho^l \theta^{-l} \right) = \frac{d_1 \tilde{\rho}}{1 - \theta d_2} \left(\rho d_2 \frac{1 - (\rho d_2)^N}{1 - \rho d_2} - (d_2)^{N+1} \rho \frac{\theta^N - \rho^N}{\theta(\theta - \rho)} \right). \end{aligned}$$

Substituting this in equation (15) and under the limit $N \rightarrow \infty$, we obtain the following compact expression (where $\tilde{\rho} = (1 - \rho)/\rho$ now in the limit):

$$\mathbf{y}\cdot\boldsymbol{\rho} \approx \frac{O_{mean}}{1 - O_{mean}} \text{ with } O_{mean} := (1 - \theta)m\eta \frac{(1 - \rho)d_1 d_2}{(1 - \theta d_2)(1 - \rho d_2)}. \quad (18)$$

3.4. Validation of the number of shares

We validate our theoretical expression for the expected number of shares by Monte Carlo simulation based on a real dataset - *Stanford Large Network Dataset Collection*

(SNAP) dataset as provided in ego-Facebook, Social Networks section [33]. The dataset consists of friends' list of 4039 Facebook users and undirected connections among them. The sum of the number of friends of all these 4039 users (undirected connections) stands at 88234. To judiciously validate the theoretical finding, we add new users to the existing dataset as it has insufficient users originally. Basically, we split the friends of the nodes that have higher degree of connections into multiple sets. We then created new users and made undirected connections by randomly choosing the nodes from each of the above-mentioned sets. (We now have a total of 20109 users.)

We emulate the content propagation on the above dataset as follows. TL of each user has five levels ($N = 5$). The starting type-1 seed TL reads the CP-post with probability r_1 , shares it with a random number of friends from its friends' list (as in the dataset) while influenced by the post quality factor η . We incorporated all the other details, e.g., shifting, the lifetime of a TL, etc., into the simulation. The number of shares in each sample path (realization) at some fixed time points is noted down. We then compute the average number of shares generated in 8000 such sample paths (and at the fixed time instances), i.e., the time evolution of the expected number of shares. In Figure 6, the time evolution of the expected number of shares (theoretical as well as MC based) is plotted.

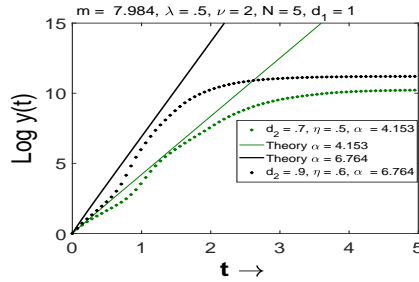


Figure 6: Time evolution of the expected shares: Theory vs Simulation

As the number of users are finite (dataset), the trajectory of $\log y_1(t)$ begins to saturate as time elapses in Monte Carlo simulation. While theoretically, the expected number of shares continues to grow indefinitely. From Figure 6, the theoretical trajectory of $\log y_1(t)$ well matches to that of the simulation based trajectory till saturation.

4. Viral marketing and real time bidding

The performance measures obtained in the previous sections can be useful in many advertisement or campaign related objectives such as brand awareness, search engine optimization, maximizing the number of clicks to a post (advertisement, abbreviated as, ad), etc. In this section, we will study online auctioning for advertisements in viral marketing using the performance measures as obtained in the previous sections.

The publishers of OSNs sell the advertisement inventory to various CPs via auction mechanism commonly known as *real time bidding* [3]. For example, Facebook auctions billions of advertisement space inventory every day, and the ads of the winners are served. Real-time bidding enables the CPs to automatically submit their bids in real time, and the advertisement of the highest worth (based on bid amount and its performance) is thus served. By virtue of auctioning, a natural competition occurs among the CPs for winning auctions. A CP has to win the auction to get sufficient number of seed (initial)

timelines. The virality/sharing of the post further depends upon the quality of the advertisement/post (recall the post quality factor η). On summarizing, the CP has to invest in two aspects: a) the bid amount to win the auction, and b) the amount spent on the design of the post (η). Recall that designing of a post could include providing authentic information about your services/products, or providing quality content, or giving offers, etc. Inappropriately tailored post can make users lose interest in the post, and thereby reducing the virality chances.

CPs typically have wide-ranging objectives while advertising on OSNs. For example, a CP may be interested in enhancing the brand awareness of its products. Brand awareness plays a central role in users' decision making for a purchase. Such an objective is achieved if the brand promotional post gets viral. Recall, we say a post gets viral if it spreads on a massive scale via its sharing among the users. Given that a post gets viral, a CP may be interested in knowing how fast the post spreads, i.e., the rate of virality. Other objectives, a CP may be interested in, include: maximizing the number of clicks on its post, improving its reputation, increasing its presence in the marketplace, etc.

In previous sections, we derived some of these performance measures. For example, we obtained the time evolution of the number of shares and NU-TLs which characterize the rate of virality. We also obtained the expression for the probability of virality. On the other hand, in non-viral (sure extinction) scenarios, we computed the expected number of total shares before extinction. We provided explicit expressions for some of the performance measures as a function of controllable parameters while others are represented as the solutions of appropriate FP (fixed point) equations. *One can use these measures to study a relevant optimization problem taking auctions into account.* In particular, and without loss of generality, we take the expected shares/NU-TLs as an indicative of the performance of CP's posts.

4.1. Optimal budget allocation

In the single CP model¹¹, the CP (indirectly) competes with other CPs only for advertisement inventory space (i.e., for winning initial seeds). This is because the other CPs are advertising unrelated content. We consider the details related to winning auctions, and then the resultant rewards derived by single CP. The CP has to first win the auction, and then its post will propagate via the shares as discussed before. Recall that these shares generate revenue to the CP. Therefore, it becomes important for the (concerned) CP to know the bid distribution of the other CPs. In particular, we need the highest bid of the agents participating in the auction. Authors in [3] show that the maximum bid value follows the log-normal distribution with parameters mean μ_b and variance σ_b^2 . Let \mathcal{M} denote the distribution of the maximum bid values M_b , then, it follows from [3]:

$$\log M_b \sim N(\mu_b, \sigma_b^2) \quad \text{where } N(\cdot, \cdot) \text{ is Gaussian distribution.} \quad (19)$$

As mentioned before, we consider the expected value of NU-TLs as one among several choices of performance measures to study the optimization problem. More specifically, we consider the sum of the expected number of users with CP-post at various levels, $\sum_{l=1}^N E[X_l(t)]$ for some large t , as an indicator of CP's revenue. Note that when the

¹¹When we consider the study of competing content in [8], the CPs further compete over relative visibility of their own content (details in [8]).

characteristics of underlying social network (e.g., sparsely connected) are such that the probability of extinction is *one*, i.e., in non-viral scenario, the CP gets *zero* reward as the NU-TLs become zero after some time. Whereas in the viral scenarios (i.e., $\mathbf{q} < \mathbf{1}$), we have $\bar{\eta} \leq \eta \leq 1$ (see (8)) and the CP gets $\sum_{l=1}^N E[X_l(t)]$ provided that it wins the auction.

Authors in [22] state that the winner of the auction is decided based on the bid amount and the corresponding quality of the post/advertisement collectively. In other words, the CP wins the auction when the bid amount x and η collectively exceeds M_b , the maximum bid by other CPs, i.e., $x\eta > M_b$. Thus, the probability of winning the bid is $P(M_b < x\eta)$, which is the cumulative density function (CDF) of log-normal distribution. Given that the CP wins the auction, its content is placed at the top-level of one TL, i.e., we begin content propagation with one seed TL of type-1. By the time the seed user visits its TL, the post might have shifted down or might disappear completely from the TL.

In all, the CP invests: i) x amount for bidding so as to win the auction, and ii) $\kappa_1\eta$ amount for preparing the post, where $\kappa_1 > 0$. Let us say the CP wants to maximize its utility, denoted by $\mathbf{C}(x, \eta)$ where

$$\mathbf{C}(x, \eta) = \begin{cases} \left(\log E\left(\sum_l X_l(t)\right) - \kappa_2(x + \kappa_1\eta) \right) P(M_b < x\eta), & \text{if } \bar{\eta} \leq \eta \leq 1 \\ 0, & \text{else,} \end{cases}$$

where the weightage κ_2 captures trade-off between the reward $\log E\left(\sum_l X_l(t)\right)$ and the overall cost $x + \kappa_1\eta$.

The close-form expression of CDF of log-normal distribution (here erf is the error function),

$$P(M_b < x\eta) = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{\log x\eta - \mu_b}{\sqrt{2}\sigma_b}\right) = \frac{1}{2} + \operatorname{sign}(f(x\eta)) \frac{1}{\sqrt{\pi}} \int_0^{|f(x\eta)|} e^{-z^2} dz,$$

where $\operatorname{sign}(a) = +1$ if $a \geq 0$ and -1 otherwise; and $f(x\eta) = \frac{\log x\eta - \mu_b}{\sqrt{2}\sigma_b}$. Using Corollary 1, we rewrite it as

$$\mathbf{C}(x, \eta) = \begin{cases} \left(\log \left(e^{\alpha t v_i} \sum_l u_l \right) - \kappa_2(x + \kappa_1\eta) \right) \left(\frac{1}{2} + \operatorname{sign}(f(x\eta)) \frac{1}{\sqrt{\pi}} \int_0^{|f(x\eta)|} e^{-z^2} dz \right), & \text{if } \bar{\eta} \leq \eta \leq 1 \\ 0, & \text{else.} \end{cases}$$

Thus, the optimization problem is stated as:

$$O1 : \max_{x, \eta} \mathbf{C}(x, \eta) \quad \text{s.t. } x \geq 0 \text{ and } 0 \leq \eta \leq 1. \quad (20)$$

In some scenarios, the CP is constrained by limited budget, say \bar{B} . Then, the question is how to allocate/divide the same into bid amount x and η -related cost, such that the revenue is maximized. For such cases, we consider constrained optimization (revenue maximization) problem under the budget constraint $\mathcal{B}(x, \eta) = x + \kappa_1\eta \leq \bar{B}$. This leads to the formulation of a *variant of the above stated optimization problem*:

$$O2 : \max_{x, \eta} \log \left(E \left[\sum_l X_l(t) \right] \right) P(M_b \leq x\eta) \quad \text{s.t. } \bar{\eta} \leq \eta \leq 1, x \geq 0, x + \kappa_1\eta \leq \bar{B}. \quad (21)$$

Optimizers of the above problem give the best allocation of the available budget subject to the following factors: i) winning the auction, and ii) maintaining post quality such that overall spending does not exceed \bar{B} .

Proposition 1. Any pair of optimizers, (x^*, η^*) , of $O2$ satisfy $x^* + \kappa_1 \eta^* = \bar{B}$.

Proof The proof is given in Appendix. ■

Due to the complex nature of the underlying objective functions, it is hard to analyse both of the optimization problems analytically. In particular, we are interested in obtaining the optimizers and study their variations with different system parameters in both the optimization problems. Let us say C^* and C_{con}^* be optimal objective values of $O1$ and $O2$. We compare and contrast the optimizers and objective values of $O1$ and $O2$ in the plots of Figures 7 and 8, where the sub-figures have expansion of only x^* and η^* curves.

Figures 7 and 8 depict the CP's spending, i) on the bid amount for winning the auction, x ; and ii) on the post quality factor η in order to maximize its utility. When the CP has a limited budget, i.e., as in optimization problem $O2$, we see in Figure 7 that x^* increases whereas η^* decreases as the mean number of friends increases (m). Eventually, both settle at their respective constant values, i.e., $x^* \approx 2.08$, $\eta^* \approx 0.69$. This pattern is attributed to two factors: i) the cost factor for η , i.e., κ_1 is comparable to \bar{B} ; and ii) the increasing mean number of friends accounts for the steady decrease of η to 0.69. In other words, as m increases, the post can get viral with smaller η (see equation (8)), and hence, the CP tends to proportionally invest more in winning bid. This kind of trend is seen only when x and η are taken together as in budget constraint. While in optimization problem $O1$, due to the absence of budget constraint, we immediately see in Figure 8 that x^* increases unrestricted and η^* also increases to its maximum value one as the network activity increases (measure by m). Hence, we see higher optimal objective values attained in $O1$ compared to those of $O2$. Basically, the CP can utilize the increasing connectivity of network (i.e., m increases) by investing more in x and η . Note that the trend is different for the extinction probability as in Figure 5. There we saw that as the connectivity of the network increases, the virality chances decrease. However, as seen now, the virality chances reduce but the expected shares still improve.

While in $O2$, the CP steadily increases allocation in x , and hence, proportionally invests lesser in η . And both eventually settle to the constant values.

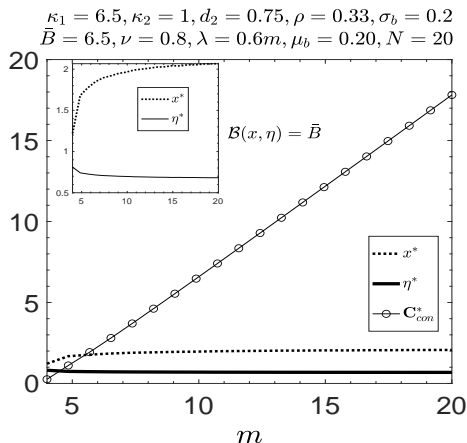


Figure 7: $O2$ with TL & $\lambda \propto m$

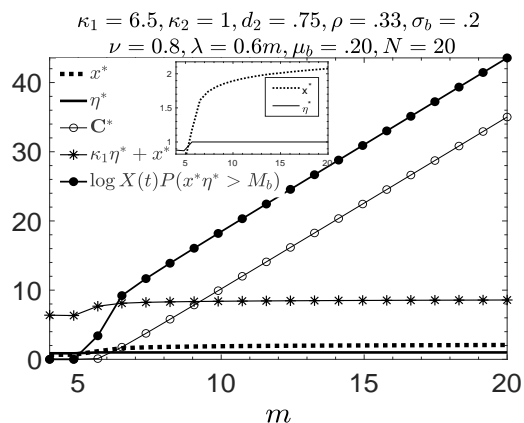


Figure 8: $O1$ with TL & $\lambda \propto m$.

When mean number of friends increases, it is natural that the CPs would bid more as it would be easier for a content to get viral (recall $\alpha \propto m$). In other words, an increase in m implies an increase in μ_b ($\mu_b \propto m$). When the mean of bid distribution increases, it gets difficult to win the auction in constrained problem $O2$ (see Figure 9). Consequently, the CP has to invest more in winning the auction, which comes at the cost of reducing the post quality η (x can not increase unrestricted due to the budget constraint). Further, as explained earlier the increasing mean number accounts for the steady decrease in η^* , and thereby, x^* increases and both of them converge to the fixed values as can be seen in Figure 9. Note that the objective value C_{con}^* , in this case, decreases after $m \approx 5$ because the allocation to x is considerably higher than that seen in Figure 7. Again in Figure 10, without the budget constraint (in $O1$), we do not see the trend. The optimal value increases with an increase in m , as x^* can take unrestricted values.

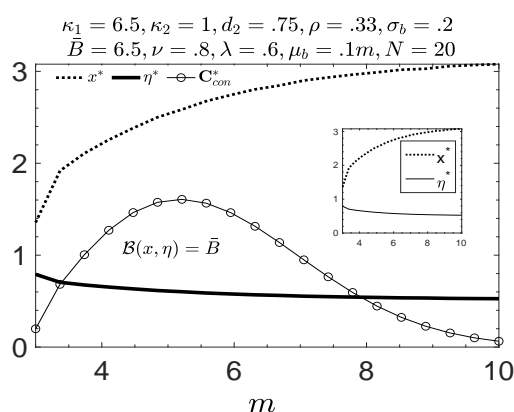


Figure 9: $O2$ with TL & $\mu_b \propto m$

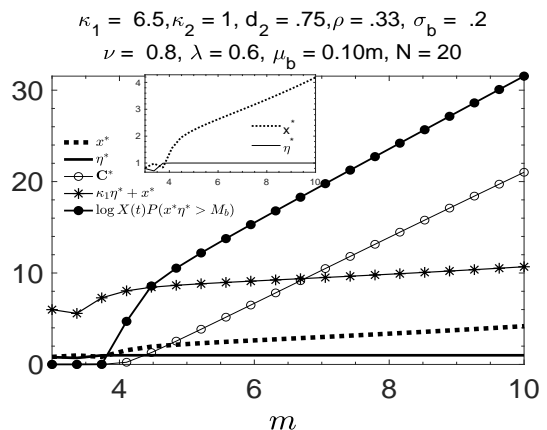


Figure 10: $O1$ with TL & $\mu_b \propto m$

Impact of timeline structure on optimizers: Earlier we studied the impact of TL structure on the post propagation. We now see through Figures 11 and 12 how neglecting the TL structure influences the optimizers. In No-TL case, the optimal values in both versions $O1$ and $O2$ are higher than that of their respective timeline scenarios, overestimating the realistic optimal value. Also, the realizable optimal objective value (TL case) further gets compromised when it is accompanied by adopting No-TL case optimizers. They may be sub-optimal for TL scenario for, e.g., in the context of $O2$ problem with No-TL case concluding $x^* \approx 2.27$, $\eta^* \approx 0.64$ (see Figure 12) to be optimizers is fallacious (the actual optimizers in the TL case as in Figures 7 and 8 are $x^* \approx 2.08$, $\eta^* \approx 0.69$). Thus, ignoring TL can cause a CP to make sub-optimal decisions and may indicate

false trends.

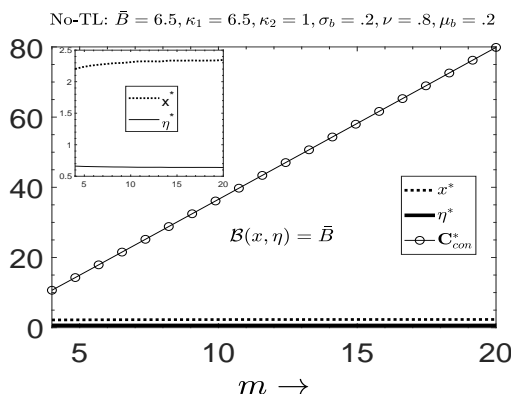


Figure 11: Problem *O2* without TL

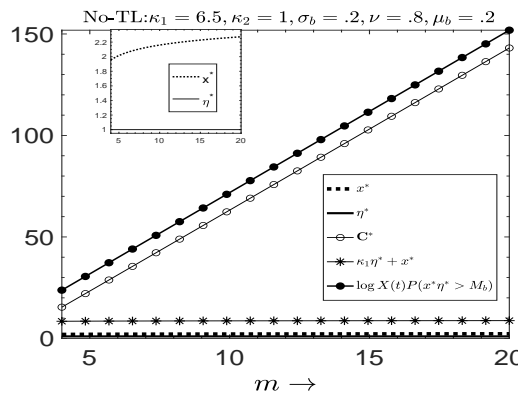


Figure 12: Problem *O1* without TL

Conclusions

We studied the propagation of a post of interest over an OSN with large number of users. We modeled the propagation of the post, considering the timeline (TL) structure, by an appropriate multi-type branching process. As is well known, we found that the underlying branching process exhibits a certain dichotomy: either the post gets extinct or goes viral. We obtained various performance measures such as the time evolution of the number of unread posts, the expected number of shares, the probability of virality, etc. We showed that the expected number of shares grow at the same rate as the number of unread posts. We compared our findings with the results that one would obtain without considering TL structure. We observed that the branching model without considering the TL structure leads to erroneous conclusions, for the systems that are significantly influenced by the TL structure. For instance, we found that a study without TLs shows that even less attractive posts can get viral. It also indicates erroneous growth rates. More importantly, we also observe that without TL effects, one cannot capture some interesting paradigm shifts/phase transitions in certain behavioral patterns. For example, as the network becomes more active, one anticipates that it is more beneficial to engage in the network. The studies which do not incorporate these effects of TL lead to this erroneous conclusion; and argue that the virality chances increase monotonically as the mean number of friends increases (m). We demonstrated that virality chances do not increase monotonically with the number of friends. After a certain value of m , it decreases for some intermittently active networks (medium m values). To be more specific, for some range of parameters, less active networks are preferable to more active networks.

Lastly, we integrated online auctions into our viral marketing model. We studied the optimization problem considering the online auctions. We again compared the study with and without considering TL structure for varying activity levels of the network. We observed that the analysis without considering TL structure fails to capture phase transitions, thereby making the overall study incomplete. Our study provides a framework using which, one can estimate important performance measures related to content propagation over online social networks, which further can be used in solving relevant optimization/game theoretic problems.

[1] Agarwal, K. and Veeraruna Kavitha, "Saturated total-population dependent branch-

- ing process and viral markets”, CDC 2022 and available in arXiv preprint, arXiv:2203.16239.
- [2] Chen, Wei, Yajun Wang, and Siyu Yang. “Efficient influence maximization in social networks”, Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2009.
 - [3] Cui, Ying, et al. “Bid landscape forecasting in online ad exchange marketplace”, Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2011.
 - [4] Doerr, Benjamin, Mahmoud Fouz, and Tobias Friedrich. “Why rumors spread so quickly in social networks”, Communications of the ACM, 2012.
 - [5] Du, MFB Nan, Yingyu Liang, and L. Song. “Continuous-time influence maximization for multiple items”, CoRR, abs/1312.2164, 2013.
 - [6] Dhouchak, Ranbir, Veeraruna Kavitha, and Eitan Altman. “A viral timeline branching process to study a social network”, 29th International Teletraffic Congress (ITC 29), 2017.
 - [7] Dhouchak, Ranbir, and Veeraruna Kavitha. “Decomposable Branching Processes and ViralMarketing”, arXiv preprint arXiv:1907.00160 2019.
 - [8] Dhouchak, Ranbir, Veeraruna Kavitha, and Eitan Altman. “A Viral Timeline Branching Process to Study a Social Network, arXiv preprint arXiv:1705.09828
 - [9] Eitan Altman, et al. “Competition over timeline in social networks”, IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2013.
 - [10] Eitan Altman, Philippe Nain, Adam Shwartz, Yuedong Xu. “Predicting the Impact of Measures Against P2P Networks: Transient Behaviour and Phase Transition”, IEEE Transactions on Networking (ToN), 2013.
 - [11] Gleeson, James P., et al. “Competition-induced criticality in a model of meme popularity”, Physical review letters 112.4, 2014.
 - [12] Gleeson, J. P., Onaga, T., Fennell, P., Cotter, J., Burke, R., & OSullivan, D. J. “Branching process descriptions of information cascades on Twitter”, Journal of Complex Networks, 2020.
 - [13] Goel, S., Anderson, A., Hofman, J., & Watts, D. J. The structural virality of online diffusion. Management Science, 2016.
 - [14] S. Hautphenne. “Extinction probabilities of supercritical decomposable branching processes”, Journal of Applied Probability, 2012.
 - [15] Iribarren, Jose Luis, and Esteban Moro. “Branching dynamics of viral information spreading”, Physical Review E 84.4, 2011.
 - [16] Iribarren, Jos Luis, and Esteban Moro. “Information diffusion epidemics in social networks”, arXiv preprint arXiv:0706.0641 2007.
 - [17] J.A.C. Resing, “Polling systems and multitype branching processes”, Queueing Systems, December 1993.

- [18] Lerman, Kristina. “Information is not a virus, and other consequences of human cognitive limits”, *Future Internet* 8.2, 2016.
- [19] H. Kesten, and BP. Stigum. “Limit theorems for decomposable multi-dimensional Galton-Watson processes”, *Journal of Mathematical Analysis and Applications*, 1967.
- [20] Krishna B Athreya and Peter E Ney. “Branching processes”, Springer Science & Business Media, 2012.
- [21] Krishna Balasundaram Athreya. “Some results on multitype continuous time Markov branching processes”, *The Annals of Mathematical Statistics*, 1968.
- [22] Mahdian, Mohammad, and Kerem Tomak. “Pay-per-action model for online advertising”, *Proceedings of the 1st international workshop on Data mining and audience intelligence for advertising*, ACM, 2007.
- [23] Meeker, Mary, and Liang Wu. “Internet trends 2018.” or “Measuring digital development: Facts and figures 2021” available at <https://www.itu.int/en/ITU-D/Statistics/Pages/facts/default.aspx>.
- [24] Nielsen, Jakob. “Scrolling and attention”, Nielsen Norman Group, 2010.
- [25] Piantino, S., Case, R., Funiak, S., Gibson, D. K., Huang, J., Mack, R. D., ... & Young, S. (2014). U.S. Patent No. 8,726,142. Washington, DC: U.S. Patent and Trademark Office.
- [26] Stewart, David B., Michael T. Ewing, and Dineli R. Mather. “A conceptual framework for viral marketing”, Australian and New Zealand Marketing Academy (ANZMAC) Conference, 2009.
- [27] Sundaram, Rangarajan K, “A first course in optimization theory”, Cambridge university press, 1996.
- [28] Theodore E Harris. “The theory of branching processes”. Courier Corporation, 2002.
- [29] Van der Lans, Ralf, et al. “A viral branching model for predicting the spread of electronic word of mouth”, *Marketing Science* 29, 2010.
- [30] Xiangying Yang and Gustavo de Veciana. “Service Capacity of Peer-to-Peer Networks”, *IEEE Infocom*, 2004.
- [31] X. Yang and G.D. Veciana, “Service Capacity of Peer to Peer Networks”, *Proc. of IEEE Infocom*, 2004.
- [32] Zadeh, Amir Hassan, and Ramesh Sharda. “Modeling brand post popularity dynamics in online social networks”, *Decision Support Systems* 65, 2014.
- [33] <https://snap.stanford.edu/data/>

Proofs are given in the Appendix on the next page.

Appendix

Proof of Lemma 1: i) The matrix e^{At} for any $t > 0$ is positive regular iff e^A is [21], because $A + I$ has only non-negative entries. Thus it is sufficient to prove e^A is positive regular. Without loss of generality we can drop the multiplier $\lambda + \nu$. Then the matrix A can be written in the following way $A = A_1 + A_2$, where

$$A_1 = \begin{bmatrix} c_1 r_1 & c_2 r_1 + \theta & \cdot & c_{N-1} r_1 & c_N r_1 \\ c_1 r_2 & c_2 r_2 & \cdot & c_{N-1} r_2 & c_N r_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ c_1 r_{N-1} & c_2 r_{N-1} & \cdot & c_{N-1} r_{N-1} & c_N r_{N-1} + \theta \\ c_1 r_N & c_2 r_N & \cdot & c_{N-1} r_N & c_N r_N \end{bmatrix}$$

and $A_2 = -I$, where I is the identity matrix. Thus, $e^A = e^{A_1} e^{A_2} = e^{-1} e^{A_1}$ since the matrices commute. For any i , one can express

$$e^{A_i} = I + A_i + \frac{A_i^2}{2!} + \frac{A_i^3}{3!} + \dots \quad (22)$$

Also $e^{A_2} = e^{-1} I$ commutes with e^{A_1} . A matrix is positive regular if there exists an n such that A^n has all positive entries. If $c_l > 0$ and $r_l > 0$ for all l , then A_1 is trivially positive regular and hence e^A is also positive regular.

Consider a general case, where some of the constants can be zero, in particular consider the case with $c_l = 0 \forall l > 1$ and $c_1 > 0$. For this case:

$$A_1 = \begin{bmatrix} c_1 r_1 & \theta & 0 & 0 & \cdot & 0 & 0 & 0 & 0 \\ c_1 r_2 & 0 & \theta & 0 & \cdot & 0 & 0 & 0 & 0 \\ c_1 r_3 & 0 & 0 & \theta & \cdot & 0 & 0 & 0 & 0 \\ c_1 r_4 & 0 & 0 & 0 & \cdot & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ c_1 r_{N-3} & 0 & 0 & 0 & \cdot & 0 & \theta & 0 & 0 \\ c_1 r_{N-2} & 0 & 0 & 0 & \cdot & 0 & 0 & \theta & 0 \\ c_1 r_{N-1} & 0 & 0 & 0 & \cdot & 0 & 0 & 0 & \theta \\ c_1 r_N & 0 & 0 & 0 & \cdot & 0 & 0 & 0 & 0 \end{bmatrix}$$

Then it is clear that

$$A_1^2 = \begin{bmatrix} c_1^2 r_1^2 + \theta c_1 r_2 & \theta c_1 r_1 & \theta^2 & \cdot & 0 & 0 \\ c_1^2 r_1 r_2 + \theta c_1 r_3 & \theta c_1 r_2 & 0 & \cdot & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ c_1^2 r_1 r_{N-2} + \theta c_1 r_{N-1} & \theta c_1 r_{N-2} & 0 & \cdot & 0 & \theta^2 \\ c_1^2 r_1 r_{N-1} + \theta c_1 r_N & \theta c_1 r_{N-1} & 0 & \cdot & 0 & 0 \\ c_1^2 r_1 r_N & \theta c_1 r_N & 0 & \cdot & 0 & 0 \end{bmatrix} \quad (23)$$

The third power $A_1^3 = A_1^2 A_1$ will have first three columns positive because the first two columns in A_1^2 is have strict positive terms and the first 2×3 sub matrix of A_1

$$\begin{bmatrix} c_1 r_1 & \theta & 0 \\ c_1 r_2 & 0 & \theta \end{bmatrix}$$

has at least one positive entry in every column (and all other terms of A_1 are non-negative). Continuing this way one can verify that A_1^N has all positive entries by induction. Basically once A_1^n has first n columns with only positive entries, because the first $n \times (n+1)$ sub-matrix of A_1 has atleast one positive entry in every column, the matrix $A_1^{n+1} = (A_1^n) \times A_1$ will have its first $n+1$ columns with only positive entries. Further A_1^n has only non negative entries for any $n \in \mathbb{N}$. From (22) it is direct that e^{A_1} is positive regular and so is $e^{-1}e^{A_1}$.

For the general case, when only some of $\{c_l\}$ are non-zero since terms are non-negative, the positive regularity follows from the above case and expansion (22). The result is true as long as $c_1 > 0$, i.e., as long as $\rho_1 > 0$.

Proof of parts (ii)-(iii): We proved that e^A is positive regular. By Frobenius-Perron theory of positive regular matrices: a) there exists an eigenvalue, call it e^α , of the matrix e^A whose algebraic and geometric multiplicities are one and which dominates all the other eigenvalues in the absolute sense. In fact, α would be a real eigenvalue of matrix A , and it dominates the real components of all other eigenvalues of the matrix A ; b) there exists a left eigenvector \mathbf{u} and a right eigenvector \mathbf{v} , both with all positive components, corresponding to α . Fix one such set of left and right eigenvectors \mathbf{u}, \mathbf{v} .

Note that the eigenvectors of matrices A and e^A are the same. Any left eigenvector of α , in particular \mathbf{u} , satisfies $\mathbf{u}A = \alpha\mathbf{u}$ and hence we get the following system of equations relating \mathbf{u} and α

$$\begin{aligned} (\lambda + \nu)c_1\mathbf{r}\cdot\mathbf{u} - (\lambda + \nu)u_1 &= \alpha u_1 \text{ or in other words } c_1\mathbf{r}\cdot\mathbf{u} = \frac{\alpha + \lambda + \nu}{\lambda + \nu}u_1, \text{ and similarly} \\ c_l\mathbf{r}\cdot\mathbf{u} + \theta u_{l-1} &= \frac{\alpha + \lambda + \nu}{\lambda + \nu}u_l, \quad l \geq 2. \end{aligned} \quad (24)$$

Simplifying the above we obtain the following relation among various components of left eigenvector \mathbf{u} : for any $l \leq N$

$$u_l = \sum_{i=0}^{l-1} \frac{\rho_{l-i}}{\rho_1} \left(\frac{\theta}{\sigma}\right)^i u_1; \quad \sum_{i=1}^N u_i = \sum_{l=1}^N \frac{\rho_l}{\rho_1} \sum_{i=0}^{N-l} \left(\frac{\theta}{\sigma}\right)^i u_1 \text{ where } \sigma := \frac{\alpha + \lambda + \nu}{\lambda + \nu}. \quad (25)$$

Following exactly the same procedure, we obtain the relation among various components of right eigenvector \mathbf{v} which are

$$v_l = \sum_{i=0}^{N-l} \frac{r_{l+i}}{r_N} \left(\frac{\theta}{\sigma}\right)^i v_N \quad \forall l = 1, 2, \dots, N-1. \quad (26)$$

This completes the proof of part (iii).

Fix \mathbf{u}, \mathbf{v} as before, and consider the following linear function of σ' :

$$P(\sigma') := (\mathbf{r}\cdot\mathbf{c})\mathbf{r}\cdot\mathbf{u} + \theta \sum_{i=1}^{N-1} r_{i+1}u_i - \sigma'\mathbf{r}\cdot\mathbf{u} \quad (27)$$

where $\mathbf{r}\cdot\mathbf{c} := \sum_{i=1}^N r_i c_i$ etc. Multiplying either side of the equation (24) with r_l and then summing over l we notice that σ is a zero of $P(\cdot)$. In other words, eigenvalue $\alpha = (\sigma^* - 1)(\lambda + \nu)$, where σ^* is a zero of $P(\cdot)$. Because $u_i > 0$ for all l , $\mathbf{r}\cdot\mathbf{u} > 0$ and similarly $\mathbf{r}\cdot\mathbf{c} > 0$. Thus σ is the only zero of $P(\cdot)$. It is clear that

$$P(\mathbf{r}\cdot\mathbf{c}) = \theta \sum_{i=1}^{N-1} r_{i+1}u_i > 0.$$

Since r_i s are monotonic, i.e., because $r_1 \geq r_2 \geq \dots \geq r_N$,

$$P(\mathbf{r} \cdot \mathbf{c} + \theta) = \theta \sum_{i=1}^{N-1} r_{i+1} u_i - \theta \mathbf{r} \cdot \mathbf{u} < 0.$$

Thus, the only zero of $P(\cdot)$ lies in the open interval $(\mathbf{r} \cdot \mathbf{c}, \mathbf{r} \cdot \mathbf{c} + \theta)$. Thus $\alpha \in (\mathbf{r} \cdot \mathbf{c} - 1, \mathbf{r} \cdot \mathbf{c} + \theta - 1)(\lambda + \nu)$.

Consider the special case with $r_l = d_1 d_2^l$, where d_1 and $d_2 \leq 1$ are constants, then clearly the only root of equation (27) σ equals

$$\sigma = \mathbf{r} \cdot \mathbf{c} + \theta d_2 \frac{\sum_{i=1}^{N-1} r_i u_i}{\mathbf{r} \cdot \mathbf{u}} = \mathbf{r} \cdot \mathbf{c} + \theta d_2 \left(1 - \frac{r_N u_N}{\mathbf{r} \cdot \mathbf{u}} \right).$$

Now we study the convergence of σ as $N \rightarrow \infty$. It is obvious that the eigenvectors/eigenvalues corresponding to different N would be different. We would normalize them by choosing the eigenvector \mathbf{u} with $u_1 = 1$ for any N . With such a choice, it is clear from (25) that u_N remains bounded even when we let $N \rightarrow \infty$. Thus as $N \rightarrow \infty$

$$\sigma = \mathbf{r} \cdot \mathbf{c} + \theta d_2 \left(1 - \frac{r_N u_N}{\mathbf{r} \cdot \mathbf{u}} \right) \rightarrow \mathbf{r} \cdot \mathbf{c} + \theta d_2 \quad \text{as } N \rightarrow \infty \quad \because (r_N \rightarrow 0).$$

Thus, as the number of TL levels increase the largest eigenvalue, α of matrix A converges to $(\mathbf{r} \cdot \mathbf{c} + \theta d_2 - 1)(\lambda + \nu)$. ■

Computation of $v_l \sum_l u_l$: Referring to Theorem 1, the left and right eigenvectors of the matrix A are

$$u_l = \sum_{i=0}^{l-1} \frac{\rho^{l-i}}{\rho_1} \left(\frac{\theta}{\sigma} \right)^i u_1, \quad v_l = \sum_{i=0}^{N-l} \frac{r_{l+i}}{r_N} \left(\frac{\theta}{\sigma} \right)^i v_N; \quad l \geq 2 \quad \text{and} \quad \sigma = \frac{\alpha}{(\lambda + \nu)} + 1.$$

When $\rho_i = \tilde{\rho} \rho^i$, $r_i = d_1 d_2^i$ with $0 < d_1, d_2, \rho < 1$. On substituting these values, we obtain

$$\begin{aligned} u_l &= \frac{1 - \left(\frac{\theta}{\sigma \rho} \right)^l}{1 - \frac{\theta}{\sigma \rho}} \rho^{l-1} u_1 = \frac{\rho^{l-1} - \frac{1}{\rho} \left(\frac{\theta}{\sigma} \right)^l}{1 - \frac{\theta}{\sigma \rho}} u_1, \\ \sum_{l=1}^N u_l &= \frac{u_1}{\left(1 - \frac{\theta}{\sigma \rho} \right)} \left(\frac{1 - \rho^N}{1 - \rho} - \frac{\theta}{\sigma \times \rho} \frac{1 - \left(\frac{\theta}{\sigma} \right)^N}{1 - \frac{\theta}{\sigma}} \right), \quad \text{and,} \\ v_l &= \frac{1 - \left(\frac{\theta d_2}{\sigma} \right)^{N-l+1}}{1 - \frac{\theta d_2}{\sigma}} d_2^{l-N} v_N = \frac{\frac{d_2^l}{d_2^N} - \left(\frac{\theta}{\sigma} \right)^{N+1} d_2 \left(\frac{\sigma}{\theta} \right)^l}{1 - \frac{\theta d_2}{\sigma}} v_N. \end{aligned}$$

Thus,

$$\begin{aligned} v_l \sum_l u_l &= \frac{\frac{d_2^l}{d_2^N} - \left(\frac{\theta}{\sigma} \right)^{N+1} d_2 \left(\frac{\sigma}{\theta} \right)^l}{1 - \frac{\theta d_2}{\sigma}} v_N \frac{u_1}{\left(1 - \frac{\theta}{\sigma \rho} \right)} \left(\frac{1 - \rho^N}{1 - \rho} - \frac{\theta}{\sigma \times \rho} \frac{1 - \left(\frac{\theta}{\sigma} \right)^N}{1 - \frac{\theta}{\sigma}} \right) \\ &= \frac{\frac{d_2^l}{d_2^N} - \left(\frac{\theta}{\sigma} \right)^{N+1} d_2 \left(\frac{\sigma}{\theta} \right)^l}{\left(1 - \frac{\theta d_2}{\sigma} \right) \left(1 - \frac{\theta}{\sigma} \right)} \left(\frac{1 - \rho^N}{1 - \rho} - \frac{\theta}{\sigma \times \rho} \frac{1 - \left(\frac{\theta}{\sigma} \right)^N}{1 - \frac{\theta}{\sigma}} \right) u_1 v_N. \end{aligned} \quad (28)$$

We require the value of $u_1 v_N$ towards obtaining $v_l \sum_l u_l$. For this, we will use the fact that $\mathbf{u} \cdot \mathbf{v} = 1$. So,

$$\sum_{l=1}^N u_l v_l = \sum_{l=1}^N \frac{\rho^{l-1} - \frac{1}{\rho} \left(\frac{\theta}{\sigma}\right)^l}{1 - \frac{\theta}{\sigma\rho}} \frac{\frac{d_2^l}{d_2^N} - \left(\frac{\theta}{\sigma}\right)^{N+1} d_2 \left(\frac{\sigma}{\theta}\right)^l}{1 - \frac{\theta d_2}{\sigma}} u_1 v_N = 1.$$

Observe that

$$\begin{aligned} & \sum_{l=1}^N \left(\rho^{l-1} - \frac{1}{\rho} \left(\frac{\theta}{\sigma}\right)^l \right) \times \left(\frac{d_2^l}{d_2^N} - \left(\frac{\theta}{\sigma}\right)^{N+1} d_2 \left(\frac{\sigma}{\theta}\right)^l \right) \\ &= \frac{d_2}{d_2^N} \sum_{l=1}^N (d_2 \rho)^{l-1} - \frac{1}{\rho d_2^N} \sum_{l=1}^N \left(\frac{\theta d_2}{\sigma}\right)^l - \sum_{l=1}^N \left(\frac{\theta}{\sigma}\right)^{N+1} \frac{d_2}{\rho} \left(\frac{\sigma \rho}{\theta}\right)^l + \sum_{l=1}^N \frac{d_2}{\rho} \left(\frac{\theta}{\sigma}\right)^{N+1} \\ &= \frac{d_2}{d_2^N} \frac{1 - (d_2 \rho)^N}{1 - d_2 \rho} - \frac{1}{\rho d_2^N} \frac{\theta d_2}{\sigma} \frac{1 - \left(\frac{\theta d_2}{\sigma}\right)^N}{1 - \frac{\theta d_2}{\sigma}} - \frac{d_2}{\rho} \left(\frac{\theta}{\sigma}\right)^{N+1} \frac{\sigma \rho}{\theta} \frac{1 - \left(\frac{\sigma \rho}{\theta}\right)^N}{1 - \frac{\sigma \rho}{\theta}} + \frac{d_2}{\rho} \left(\frac{\theta}{\sigma}\right)^{N+1} N. \end{aligned}$$

Substituting the above into $\sum_{l=1}^N u_l v_l = 1$ equation, we have:

$$\frac{u_1 v_N}{\left(1 - \frac{\theta}{\sigma\rho}\right) \left(1 - \frac{\theta d_2}{\sigma}\right)} \left(\frac{d_2}{d_2^N} \frac{1 - (d_2 \rho)^N}{1 - d_2 \rho} - \frac{1}{\rho d_2^N} \frac{\theta d_2}{\sigma} \frac{1 - \left(\frac{\theta d_2}{\sigma}\right)^N}{1 - \frac{\theta d_2}{\sigma}} - \frac{d_2}{\rho} \left(\frac{\theta}{\sigma}\right)^{N+1} \frac{\sigma \rho}{\theta} \frac{1 - \left(\frac{\sigma \rho}{\theta}\right)^N}{1 - \frac{\sigma \rho}{\theta}} + \frac{d_2}{\rho} \left(\frac{\theta}{\sigma}\right)^{N+1} N \right) = 1,$$

$$\text{or, } u_1 v_N = \frac{\left(1 - \frac{\theta}{\sigma\rho}\right) \left(1 - \frac{\theta d_2}{\sigma}\right)}{\left(\frac{d_2}{d_2^N} \frac{1 - (d_2 \rho)^N}{1 - d_2 \rho} - \frac{1}{\rho d_2^N} \frac{\theta d_2}{\sigma} \frac{1 - \left(\frac{\theta d_2}{\sigma}\right)^N}{1 - \frac{\theta d_2}{\sigma}} - \frac{d_2}{\rho} \left(\frac{\theta}{\sigma}\right)^{N+1} \frac{\sigma \rho}{\theta} \frac{1 - \left(\frac{\sigma \rho}{\theta}\right)^N}{1 - \frac{\sigma \rho}{\theta}} + \frac{d_2}{\rho} \left(\frac{\theta}{\sigma}\right)^{N+1} N \right)}. \quad (29)$$

Now substituting the value of $u_1 v_N$ in equation (28)

$$\begin{aligned} v_l \sum_l u_l &= \frac{\frac{d_2^l}{d_2^N} - \left(\frac{\theta}{\sigma}\right)^{N+1} d_2 \left(\frac{\sigma}{\theta}\right)^l}{\left(1 - \frac{\theta d_2}{\sigma}\right) \left(1 - \frac{\theta}{\sigma}\right)} \left(\frac{1 - \rho^N}{1 - \rho} - \frac{\theta}{\sigma \times \rho} \frac{1 - \left(\frac{\theta}{\sigma}\right)^N}{1 - \frac{\theta}{\sigma}} \right) \\ &\quad \times \frac{\left(1 - \frac{\theta}{\sigma\rho}\right) \left(1 - \frac{\theta d_2}{\sigma}\right)}{\left(\frac{d_2}{d_2^N} \frac{1 - (d_2 \rho)^N}{1 - d_2 \rho} - \frac{1}{\rho d_2^N} \frac{\theta d_2}{\sigma} \frac{1 - \left(\frac{\theta d_2}{\sigma}\right)^N}{1 - \frac{\theta d_2}{\sigma}} - \frac{d_2}{\rho} \left(\frac{\theta}{\sigma}\right)^{N+1} \frac{\sigma \rho}{\theta} \frac{1 - \left(\frac{\sigma \rho}{\theta}\right)^N}{1 - \frac{\sigma \rho}{\theta}} + \frac{d_2}{\rho} \left(\frac{\theta}{\sigma}\right)^{N+1} N \right)} \\ &= \frac{\frac{d_2^l}{d_2^N} - \left(\frac{\theta}{\sigma}\right)^{N+1} d_2 \left(\frac{\sigma}{\theta}\right)^l}{\left(1 - \frac{\theta}{\sigma}\right)} \left(\frac{1 - \rho^N}{1 - \rho} - \frac{\theta}{\sigma \rho} \frac{1 - \left(\frac{\theta}{\sigma}\right)^N}{1 - \frac{\theta}{\sigma}} \right) \\ &\quad \times \frac{\left(1 - \frac{\theta}{\sigma\rho}\right)}{\left(\frac{d_2}{d_2^N} \frac{1 - (d_2 \rho)^N}{1 - d_2 \rho} - \frac{1}{\rho d_2^N} \frac{\theta d_2}{\sigma} \frac{1 - \left(\frac{\theta d_2}{\sigma}\right)^N}{1 - \frac{\theta d_2}{\sigma}} - \frac{d_2}{\rho} \left(\frac{\theta}{\sigma}\right)^{N+1} \frac{\sigma \rho}{\theta} \frac{1 - \left(\frac{\sigma \rho}{\theta}\right)^N}{1 - \frac{\sigma \rho}{\theta}} + \frac{d_2}{\rho} \left(\frac{\theta}{\sigma}\right)^{N+1} N \right)} \\ &= \frac{d_2^l - \left(\frac{\theta d_2}{\sigma}\right)^{N+1} \left(\frac{\sigma}{\theta}\right)^l}{\left(1 - \frac{\theta}{\sigma}\right)} \left(\frac{1 - \rho^N}{1 - \rho} - \frac{\theta}{\sigma \rho} \frac{1 - \left(\frac{\theta}{\sigma}\right)^N}{1 - \frac{\theta}{\sigma}} \right) \\ &\quad \times \frac{\left(1 - \frac{\theta}{\sigma\rho}\right)}{\left(\frac{d_2}{d_2^N} \frac{1 - (d_2 \rho)^N}{1 - d_2 \rho} - \frac{\theta d_2}{\sigma \rho} \frac{1 - \left(\frac{\theta d_2}{\sigma}\right)^N}{1 - \frac{\theta d_2}{\sigma}} - \frac{1}{\rho} \left(\frac{\theta d_2}{\sigma}\right)^{N+1} \frac{\sigma \rho}{\theta} \frac{1 - \left(\frac{\sigma \rho}{\theta}\right)^N}{1 - \frac{\sigma \rho}{\theta}} + \frac{1}{\rho} \left(\frac{\theta d_2}{\sigma}\right)^{N+1} N \right)}. \end{aligned}$$

On simplifying and using property of limits:

$$\begin{aligned} \lim_{N \rightarrow \infty} v_l \sum_{l=1}^N u_l &= \lim_{N \rightarrow \infty} \frac{d_2^l - \left(\frac{\theta d_2}{\sigma}\right)^{N+1} \left(\frac{\sigma}{\theta}\right)^l}{\left(1 - \frac{\theta}{\sigma}\right)} \left(\frac{1 - \rho^N}{1 - \rho} - \frac{\theta}{\sigma \rho} \frac{1 - \left(\frac{\theta}{\sigma}\right)^N}{1 - \frac{\theta}{\sigma}} \right) \\ &\times \lim_{N \rightarrow \infty} \frac{\left(1 - \frac{\theta}{\sigma \rho}\right)}{\left(d_2 \frac{1 - (d_2 \rho)^N}{1 - d_2 \rho} - \frac{\theta d_2}{\sigma \rho} \frac{1 - \left(\frac{\theta d_2}{\sigma}\right)^N}{1 - \frac{\theta d_2}{\sigma}} - d_2 \frac{\left(\frac{\theta d_2}{\sigma}\right)^N - (d_2 \rho)^N}{1 - \frac{\sigma \rho}{\theta}} + \frac{1}{\rho} \left(\frac{\theta d_2}{\sigma}\right)^{N+1} N \right)}. \end{aligned} \quad (30)$$

Note that in viral scenario $\alpha > 0$ and $\sigma > 1$ and hence $\theta/\sigma < 1$, so

$$\lim_{N \rightarrow \infty} \left(\frac{\theta}{\sigma}\right)^{N+1} = 0, \quad \lim_{N \rightarrow \infty} \left(\frac{\theta d_2}{\sigma}\right)^{N+1} = 0, \quad \lim_{N \rightarrow \infty} N \left(\frac{\theta d_2}{\sigma}\right)^{N+1} = 0 \quad \because \quad d_2 < 1. \quad (31)$$

In what follows, for any fixed l , we have

$$\lim_{N \rightarrow \infty} \frac{d_2^l - \left(\frac{\theta d_2}{\sigma}\right)^{N+1} \left(\frac{\sigma}{\theta}\right)^l}{\left(1 - \frac{\theta}{\sigma}\right)} \left(\frac{1 - \rho^N}{1 - \rho} - \frac{\theta}{\sigma \rho} \frac{1 - \left(\frac{\theta}{\sigma}\right)^N}{1 - \frac{\theta}{\sigma}} \right) = \frac{d_2^l}{\left(1 - \frac{\theta}{\sigma}\right)} \left(\frac{1}{1 - \rho} - \frac{\theta}{\sigma \rho} \frac{1}{1 - \frac{\theta}{\sigma}} \right) \quad (32)$$

$$\lim_{N \rightarrow \infty} \frac{\left(1 - \frac{\theta}{\sigma \rho}\right)}{\left(d_2 \frac{1 - (d_2 \rho)^N}{1 - d_2 \rho} - \frac{\theta d_2}{\sigma \rho} \frac{1 - \left(\frac{\theta d_2}{\sigma}\right)^N}{1 - \frac{\theta d_2}{\sigma}} - d_2 \frac{\left(\frac{\theta d_2}{\sigma}\right)^N - (d_2 \rho)^N}{1 - \frac{\sigma \rho}{\theta}} + \frac{1}{\rho} \left(\frac{\theta d_2}{\sigma}\right)^{N+1} N \right)} = \frac{\left(1 - \frac{\theta}{\sigma \rho}\right)}{\left(\frac{d_2}{1 - d_2 \rho} - \frac{\theta d_2}{\sigma \rho} \frac{1}{1 - \frac{\theta d_2}{\sigma}} \right)} \quad (33)$$

Substituting these limits (equations (32) and (33)) in equation (30), we get

$$\begin{aligned} \lim_{N \rightarrow \infty} v_l \sum_{l=1}^N u_l &= \frac{d_2^l}{\left(1 - \frac{\theta}{\sigma}\right)} \left(\frac{1}{1 - \rho} - \frac{\theta}{\sigma \rho} \frac{1}{1 - \frac{\theta}{\sigma}} \right) \frac{\left(1 - \frac{\theta}{\sigma \rho}\right)}{\left(\frac{d_2}{1 - d_2 \rho} - \frac{\theta d_2}{\sigma \rho} \frac{1}{1 - \frac{\theta d_2}{\sigma}} \right)} \\ &= \frac{d_2^{l-1}}{\left(1 - \frac{\theta}{\sigma}\right)} \left(\frac{1}{1 - \rho} - \frac{\theta}{\rho \sigma - \theta} \right) \frac{\left(1 - \frac{\theta}{\sigma \rho}\right)}{\left(\frac{1}{1 - d_2 \rho} - \frac{\theta}{\rho \sigma - \theta d_2} \right)} \end{aligned}$$

Thus, we have

$$v_l \sum_l u_l = d_2^{l-1} (1 - d_2 \rho) \left(\frac{1}{1 - \rho} - \frac{\theta}{\rho \sigma - \theta} \right) \frac{(\sigma - \theta d_2)(\sigma \rho - \theta)}{(\sigma - \theta)(\rho - \theta)}. \quad (34)$$

■

Proof of Lemma 3 : Let $\mathbf{j}_x = \{\mathbf{j}_{x_1}, \mathbf{j}_{x_2}, \dots, \mathbf{j}_{x_N}\}$ be the number of TLs of type $1, 2, \dots, N$ respectively, and y be the total number of shares. It is easy to observe that $y \geq \sum_i j_{x_i}$. We write it in short form as $y \geq \mathbf{j}_x$. Define $\mathbf{s}_x^{\mathbf{j}_x} := \prod_i s_{x_i}^{j_{x_i}}$. Then the PGF of TL-CTBP can be written as

$$F_1(\mathbf{s}, t) = \sum_{\mathbf{j}_x=0}^{\infty} \sum_{y \geq \mathbf{j}_x}^{\infty} P_{(\mathbf{e}_{1,1}) \rightarrow (\mathbf{j}_x, y)}(t) \mathbf{s}_x^{\mathbf{j}_x} s_y^y, \quad \text{and,} \quad \frac{\delta F_1(\mathbf{s}, t)}{\delta t} = \sum_{\mathbf{j}_x=0}^{\infty} \sum_{y \geq \mathbf{j}_x}^{\infty} P'_{(\mathbf{e}_{1,1}) \rightarrow (\mathbf{j}_x, y)}(t) \mathbf{s}_x^{\mathbf{j}_x} s_y^y.$$

The above is obtained by conditioning on the events of the first transition; the populations generated by two parents evolve independently of each other and the procedure is similar

to the standard procedure used in these kind of computations (for example, [20]). Let $\xi = (\xi_1, \xi_2, \dots, \xi_N)$ represent the offsprings produced by one parent and let $\bar{\xi} := \sum_i \xi_i$.

By backward equation, we have $P'_{1k}(t) = \sum_j q_{1j} P_{jk}(t)$; in our case it is

$$\begin{aligned} \frac{\delta F_1(\mathbf{s}, t)}{\delta t} &= (\lambda + \nu) \left((1 - \theta) r_1 \sum_{\xi} \sum_{\mathbf{j}_x=0}^{\infty} \sum_{y \geq \mathbf{j}_x}^{\infty} P_1(\xi) P_{(\xi, \bar{\xi}+1) \rightarrow (\mathbf{j}_x, y)}(t) \mathbf{s}_x^{\mathbf{j}_x} s_y^y + \theta F_2(\mathbf{s}, t) \right. \\ &\quad \left. - \sum_{\mathbf{j}_x=0}^{\infty} \sum_{y \geq \mathbf{j}_x}^{\infty} P_{(\mathbf{e}_1, 1) \rightarrow (\mathbf{j}_x, y)}(t) \mathbf{s}_x^{\mathbf{j}_x} s_y^y + (1 - \theta)(1 - r_1) s_y \right) \\ \frac{\delta F_1(\mathbf{s}, t)}{\delta t} &= (\lambda + \nu) \left((1 - \theta) r_1 \sum_{\xi} P_1(\xi) \Pi_{i=1}^N \left(\sum_{\mathbf{j}_x=0}^{\infty} \sum_{y \geq \mathbf{j}_x}^{\infty} P_{(\mathbf{e}_i, 1) \rightarrow (\mathbf{j}_x, y)}(t) \mathbf{s}_x^{\mathbf{j}_x} s_y^y \right)^{\xi_i} s_y \right. \\ &\quad \left. + \theta F_2(\mathbf{s}, t) - F_1(\mathbf{s}, t) + (1 - \theta)(1 - r_1) s_y \right) \\ \frac{\delta F_1(\mathbf{s}, t)}{\delta t} &= (\lambda + \nu) \left((1 - \theta) r_1 s_y f_1(\mathbf{F}(\mathbf{s}, t)) + \theta F_2(\mathbf{s}, t) - F_1(\mathbf{s}, t) + (1 - \theta)(1 - r_1) s_y \right) \end{aligned}$$

where $\mathbf{F}(\mathbf{s}, t) := \{F_1(\mathbf{s}, t), F_2(\mathbf{s}, t), \dots, F_N(\mathbf{s}, t)\}$. Similarly we can write for any l

$$\begin{aligned} \frac{\delta F_l(\mathbf{s}, t)}{\delta t} &= (\lambda + \nu) \left((1 - \theta) r_l s_y f_l(\mathbf{F}(\mathbf{s}, t)) + \theta \left(\mathbb{1}_{l < N} F_{l+1}(\mathbf{s}, t) + s_y \mathbb{1}_{l=N} \right) \right. \\ &\quad \left. - F_l(\mathbf{s}, t) + (1 - \theta)(1 - r_l) s_y \right). \end{aligned}$$

Let $\dot{y}_l(t) = \frac{\delta^2 F_l(\mathbf{s}, t)}{\delta t \delta s_y} \Big|_{\mathbf{s}=1} \forall l = \{1, 2, \dots, N\}$ represent the time derivative of number shares till time t when started with a type l progenitor. We have the following expression

$$\begin{aligned} \dot{y}_1(t) &= (\lambda + \nu) \left((1 - \theta) r_1 f_1(1) + (1 - \theta) r_1 \sum_{i=1}^N \frac{\delta f_1(\mathbf{F}(\mathbf{s}, t))}{\delta F_i(\mathbf{s}, t)} \frac{\delta F_i(\mathbf{s}, t)}{\delta s_y} \Big|_{\mathbf{s}=1} + (1 - \theta)(1 - r_1) \right. \\ &\quad \left. + \theta \frac{\delta F_2(\mathbf{s}, t)}{\delta s_y} \Big|_{\mathbf{s}=1} - \frac{\delta F_1(\mathbf{s}, t)}{s_y} \Big|_{\mathbf{s}=1} \right) \\ &= (\lambda + \nu) \left((1 - \theta) r_1 + (1 - \theta)(1 - r_1) + (1 - \theta) r_1 m \eta \sum_{i=1}^N \rho_i y_i(t) + \theta y_2(t) - y_1(t) \right) \\ &= (\lambda + \nu) \left(1 - \theta + r_1 \sum_{i=1}^N c_i y_i(t) + \theta y_2(t) - y_1(t) \right). \end{aligned}$$

Similarly, we can write the above for any l

$$\dot{y}_l(t) = (\lambda + \nu) \left(1 - \theta + r_l \sum_{i=1}^N c_i y_i(t) + \theta y_{l+1}(t) \mathbb{1}_{l < N} - y_l(t) + \theta \mathbb{1}_{l=N} \right). \quad (35)$$

The above can be written in matrix form as

$$\frac{1}{\lambda + \nu} \begin{bmatrix} \dot{y}_1(t) \\ \dot{y}_2(t) \\ \vdots \\ \dot{y}_{N-1}(t) \\ \dot{y}_N(t) \end{bmatrix} = \begin{bmatrix} c_1 r_1 - 1 & c_2 r_1 + \theta & \cdots & c_{N-1} r_1 & c_N r_1 \\ c_1 r_2 & c_2 r_2 - 1 & \cdots & c_{N-1} r_2 & c_N r_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ c_1 r_{N-1} & c_2 r_{N-1} & \cdots & c_{N-1} r_{N-1} - 1 & c_N r_{N-1} + \theta \\ c_1 r_N & c_2 r_N & \cdots & c_{N-1} r_N & c_N r_N - 1 \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_{N-1}(t) \\ y_N(t) \end{bmatrix} + \begin{bmatrix} 1 - \theta \\ 1 - \theta \\ \vdots \\ 1 - \theta \\ 1 \end{bmatrix}.$$

Solving the above set of equations, we obtain:

$$\begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_{N-1}(t) \\ y_N(t) \end{bmatrix} = e^{At} \begin{bmatrix} y_1(0) \\ y_2(0) \\ \vdots \\ y_{N-1}(0) \\ y_N(0) \end{bmatrix} + e^{At} \int_0^t e^{-As} (\lambda + \nu) \begin{bmatrix} 1 - \theta \\ 1 - \theta \\ \vdots \\ 1 - \theta \\ 1 \end{bmatrix} ds \quad (36)$$

$$= e^{At} \begin{bmatrix} y_1(0) \\ y_2(0) \\ \vdots \\ y_{N-1}(0) \\ y_N(0) \end{bmatrix} + e^{At} A^{-1} (I - e^{-At}) (\lambda + \nu) \begin{bmatrix} 1 - \theta \\ 1 - \theta \\ \vdots \\ 1 - \theta \\ 1 \end{bmatrix} \quad (37)$$

With $\mathbf{y}(t) := \{y_1(t), y_2(t), \dots, y_N(t)\}$, we can represent the above as:

$$\mathbf{y}(t) := \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_{N-1}(t) \\ y_N(t) \end{bmatrix} = e^{At} \left(\mathbf{1} + (\lambda + \nu) A^{-1} \mathbf{k} \right) - (\lambda + \nu) A^{-1} \mathbf{k}$$

where $\mathbf{k} = [1 - \theta, 1 - \theta, \dots, 1 - \theta, 1]^T$.

$$\mathbf{y}(t) = e^{At} \left(\mathbf{1} + (\lambda + \nu) A^{-1} \mathbf{k} \right) - (\lambda + \nu) A^{-1} \mathbf{k}$$

From [21, equation (45)], e^{At} can be approximated for large t . By which, we can write

$$\begin{aligned} \mathbf{y}(t) &\approx e^{\alpha t} \mathbf{v} \mathbf{u}' \left(\mathbf{1} + (\lambda + \nu) A^{-1} \mathbf{k} \right) - (\lambda + \nu) A^{-1} \mathbf{k} \quad (38) \\ &\approx e^{\alpha t} \left(\mathbf{v} \sum_{i=1}^N u_i + \frac{\lambda + \nu}{\alpha} \mathbf{v} \mathbf{u}' \mathbf{k} \right) - (\lambda + \nu) A^{-1} \mathbf{k} \\ &\approx \mathbf{v} e^{\alpha t} \left(\sum_i u_i \left(1 + \frac{\lambda + \nu}{\alpha} (1 - \theta) \right) + \frac{\lambda + \nu}{\alpha} u_N \right) - (\lambda + \nu) A^{-1} \mathbf{k} \\ &\approx \mathbf{v} e^{\alpha t} \sum_i u_i \left(1 + \frac{1 - \theta}{\mathbf{r.c} - 1 + \theta d_2} \right) - (\lambda + \nu) A^{-1} \mathbf{k}. \quad \blacksquare \end{aligned}$$

Proof of Proposition 1: We will prove that at optimality the budget constraint is tight, i.e., $x + \kappa_1 \eta = \bar{B}$ using the Lagrangian relaxation method. To do so, we first change the inequality budget constraint, $x + \kappa_1 \eta \leq \bar{B}$, to equality constraint as follows. Let s^2 (ensuring it to be ≥ 0) be slack variable such that $x + \kappa_1 \eta + s^2 = \bar{B}$. Similarly, we have

$$\eta - s_1^2 = \bar{\eta}, \quad \eta + s_2^2 = 1, \quad x - s_3^2 = 0 \quad \text{for the constraints } \eta \geq \bar{\eta}, \quad \eta \leq 1 \quad \text{respectively}$$

where s_1^2, s_2^2, s_3^2 are the slack/surplus variables. The Lagrangian function $L(x, \eta, \Lambda)$ with Lagrangian multiplier $\Lambda, \Lambda_1, \Lambda_2, \Lambda_3$ is given as

$$\max_{x, \eta} \underbrace{\log E \left(\sum_l X_l(t) \right) P(\text{Bid} \leq x \eta) - \Lambda (\bar{B} - x - \kappa_1 \eta - s^2) - \Lambda_1 (\bar{\eta} - \eta + s_1^2) - \Lambda_2 (1 - \eta - s_2^2) - \Lambda_3 (x - s_3^2)}_{L(x, \eta, \Lambda, \Lambda_1, \Lambda_2, \Lambda_3)}.$$

The critical points of $\log E\left(\sum_l X_l(t)\right)P(Bid \leq x\eta)$ with the given constraint, say

$$\mathcal{G} := -x - \kappa_1\eta - s^2, \quad \mathcal{G}_1 := \bar{\eta} - \eta + s_1^2, \quad \mathcal{G}_2 := 1 - \eta - s_2^2, \quad \mathcal{G}_3 := x - s_3^2$$

are obtained by solving the following system of simultaneous equations^{12 13}

$$\begin{aligned} \frac{\partial \log E\left(\sum_l X_l(t)\right)P(M_b < x\eta)}{\partial x} &= \Lambda \frac{\partial \mathcal{G}}{\partial x} + \Lambda_1 \frac{\partial \mathcal{G}_1}{\partial x} + \Lambda_2 \frac{\partial \mathcal{G}_2}{\partial x} + \Lambda_3 \frac{\partial \mathcal{G}_3}{\partial x} \\ \implies \log E\left(\sum_l X_l(t)\right) \frac{e^{-f(x\eta)^2}}{\sqrt{2\pi\sigma_b x}} &= -\Lambda + \Lambda_3, \end{aligned} \quad (39)$$

$$\begin{aligned} \frac{\partial \log E\left(\sum_l X_l(t)\right)P(M_b \leq x\eta)}{\partial \eta} &= \Lambda \frac{\partial \mathcal{G}}{\partial \eta} + \Lambda_1 \frac{\partial \mathcal{G}_1}{\partial \eta} + \Lambda_2 \frac{\partial \mathcal{G}_2}{\partial \eta} + \Lambda_3 \frac{\partial \mathcal{G}_3}{\partial \eta} \\ \implies \log E\left(\sum_l X_l(t)\right) \frac{e^{-f(x\eta)^2}}{\sqrt{2\pi\sigma_b \eta}} + P(M_b \leq x\eta) \frac{\partial \log E\left(\sum_l X_l(t)\right)}{\partial \eta} &= -\kappa_1\Lambda - \Lambda_1 - \Lambda_2 \end{aligned} \quad (40)$$

and, $\frac{\partial \log E\left(\sum_l X_l(t)\right)P(M_b \leq x\eta)}{\partial s} = \Lambda \frac{\partial \mathcal{G}}{\partial s} + \Lambda_1 \frac{\partial \mathcal{G}_1}{\partial s} + \Lambda_2 \frac{\partial \mathcal{G}_2}{\partial s} + \Lambda_3 \frac{\partial \mathcal{G}_3}{\partial s} \implies 0 = 2\Lambda s.$ (41)

And so, $x + \kappa_1\eta + s^2 = \bar{B}$. We now compute the gradient w.r.t. to s_1, s_2, s_3 :¹⁴

$$\frac{\partial \log E\left(\sum_l X_l(t)\right)P(M_b \leq x\eta)}{\partial s_1} = \Lambda \frac{\partial \mathcal{G}}{\partial s_1} + \Lambda_1 \frac{\partial \mathcal{G}_1}{\partial s_1} + \Lambda_2 \frac{\partial \mathcal{G}_2}{\partial s_1} + \Lambda_3 \frac{\partial \mathcal{G}_3}{\partial s_1} \implies 0 = -2s_1\Lambda_1 \quad (42)$$

$$\frac{\partial \log E\left(\sum_l X_l(t)\right)P(M_b \leq x\eta)}{\partial s_2} = \Lambda \frac{\partial \mathcal{G}}{\partial s_2} + \Lambda_1 \frac{\partial \mathcal{G}_1}{\partial s_2} + \Lambda_2 \frac{\partial \mathcal{G}_2}{\partial s_2} + \Lambda_3 \frac{\partial \mathcal{G}_3}{\partial s_2} \implies 0 = 2s_2\Lambda_2 \quad (43)$$

$$\frac{\partial \log E\left(\sum_l X_l(t)\right)P(M_b \leq x\eta)}{\partial s_3} = \Lambda \frac{\partial \mathcal{G}}{\partial s_3} + \Lambda_1 \frac{\partial \mathcal{G}_1}{\partial s_3} + \Lambda_2 \frac{\partial \mathcal{G}_2}{\partial s_3} + \Lambda_3 \frac{\partial \mathcal{G}_3}{\partial s_3} \implies 0 = -2s_3\Lambda_3 \quad (44)$$

$$\text{and } \bar{\eta} - \eta + s_1^2 = 0, \quad 1 - \eta - s_2^2 = 0, \quad x - s_3^2 = 0. \quad (45)$$

Referring to equation (44), we have either $s_3 = 0$ or $\Lambda_3 = 0$. If $s_3 = 0$, then $x = 0$ (see equation (45)); which is clearly not an optimal solution (zero objective value) as the objective can be improve when $x > 0$. In particular, we do not need to compute $\Lambda_1, \Lambda_2, s_1, s_2, s_3$ for this proof. We only need to prove that $s = 0$. For this, observe that equation (41) gives that either $\Lambda = 0$ or $s = 0$. However, $\Lambda \neq 0$ because $\log E\left(\sum_l X_l(t)\right) \frac{e^{-f(x\eta)^2}}{\sqrt{2\pi\sigma_b x}}$ is positive (recall $\Lambda_3 = 0$). Therefore, we must have $s = 0$, which consequently brings out the tightness of budget constraint $x + \kappa_1\eta = \bar{B}$. Hence proved. \blacksquare

¹²see http://users.wpi.edu/~pwdavis/Courses/MA1024B10/1024_Lagrange_multipliers.pdf

¹³and http://www.math.harvard.edu/archive/21a_spring_09/PDF/11-08-Lagrange-Multipliers.pdf

¹⁴Note that we mainly require equations (41) and (44) for this proof.