How Complex Contagions Spread Quickly in Preferential Attachment Models and Other Time-Evolving Networks

Roozbeh Ebrahimi, Jie Gao, Golnaz Ghasemiesfeh, and Grant Schoenbeck,

Abstract—The k-*complex contagion* model is a social contagion model which describes the diffusion of behaviors in networks where the successful adoption of a behavior requires influence from multiple contacts. It has been argued that complex contagions better model behavioral changes such as adoption of new beliefs, fashion trends or expensive technology innovations. A contagion in this model starts from a set of initially infected seeds and progresses in rounds. In any round any node with at least $k > 1$ infected neighbors becomes infected.
Previous work on k-complex contagions was focused on networks with uniform degree distributions. However, many real-world network topologies have non-uniform degree distribution and evolve over time. We analyze the spreading rate of a k-complex contagion in a general family of time-evolving networks which includes the preferential attachment (PA) model. We prove that if the initial seeds are chosen as the k earliest nodes in a network of this family, a k-complex contagion covers the entire network of n nodes in $O(\log n)$ rounds with high probability (w.h.p). We prove that the choice of the seeds is crucial: in the PA model, even if a much larger number of seeds are chosen *uniformly randomly*, the contagion stops prematurely w.h.p. Although the earliest nodes in a PA model are likely to have high degrees, it is actually the evolutionary graph structure of such models that facilitates fast spreading of complex contagions. The general family of time-evolving graphs with this property even contains networks without a power law degree distribution.
Finally, we prove that when a k-complex contagion starts from an arbitrary set of initial seeds on a general graph, determining if the number of infected vertices is above a given threshold is $\mathbf{P}$-complete. Thus, one cannot hope to categorize all the settings in which k-complex contagions percolate in a graph.

Index Terms—Social Networks, Complex Contagion, The Preferential Attachment Model, Time-Evolving Networks, Branching Processes, Stochastic Coupling

1 Introduction

Social behavior is undoubtedly one of the defining characteristics of us as a species. Social acts are influenced by the behavior of others while at the same time influencing them. Understanding the dynamics of influence and modeling it in social networks is thus a key step in comprehending the emergence of new behaviors in societies. Similar to rumors or viruses, behavior changes manifest contagion like properties while spreading in a social network. Some of these contagions are beneficial (e.g., adopting a healthy lifestyle) or profitable (e.g., viral marketing), while some others are destructive and undesirable (e.g., teenage smoking or alcohol abuse). To effectively promote desirable contagions and discourage undesirable ones, the first step is to understand how these contagions spread in networks and what important parameters facilitate fast spreading.

In this paper, we focus on contagions that are *complex* [1], contagions that require social reaffirmation from multiple neighbors, as opposed to *simple* ones, which can spread through a single contact, e.g., viruses or rumors. It has been argued in sociology literature that when agents' actions and behavioral changes are involved, complex contagions represent most of the realistic settings – making an important distinction between the *acquisition* of information and the decision to *act* on the information. While it takes only a single tie for people to hear about a new belief or fashion, "it is when they see people they know getting involved, that they become most susceptible to recruitment", Centola and Macy [1]. Many examples of complex contagions have been reported in social studies, e.g., buying pricey new technologies, decision to migrate, etc. [2], [3]. Studies of online social networks have also confirmed complex contagions. A study on Facebook showed that having two or more friends already on Facebook who are not connected to each other substantially increases the

- *R. Ebrahimi, PhD, currently works as an engineer at Google Inc. CA, 94043. E-mail: rebrahimi@google.com.*
- *J. Gao, Associate Professor at the Department of Computer Science, Stony Brook University, NY 11794. E-mail: jgao@cs.stonybrook.edu.*
- *G. Ghasemiesfeh, PhD student at the Department of Computer Science, Stony Brook University, NY 11794. E-mail: gghasemiesfe@cs.stonybrook.edu.*
- *G.Schoenbeck, Assistant Professor at the Department of Computer Science, University of Michigan, MI 48109. E-mail: schoenb@umich.edu*

likelihood of joining Facebook [4]. A study on Twitter showed a similar phenomenon, that repeated exposures to a topic plays an important role in the diffusion of hashtags [5].

Simple contagions have been extensively studied [6]. Simple contagions can spread fast in social networks because these networks typically have the small world property. In contrast, fast spreading of complex contagions appears to be much more delicate and difficult. Previous studies [1], [7], [8] show that for a number of small world models, in which simple contagions are super fast, complex contagions are *exponentially* slower. These results use a simple yet elegant model called the k-complex contagion – starting from a set of initially infected seeds, any node with at least k infected neighbors gets infected ($k = \Theta(1)$). Remarkably, these results are the only known rigorous analyses on complex contagions, despite the crucial importance of them in modeling a wide range of social behaviors. The difficulty of formal analysis arises in two aspects. First, the required multiple infections means that subsequent exposures do not always have diminishing returns which turns out to be mathematically challenging. For example, it violates submodularity and subadditivity, on which many analyses depend. Second, the superadditive character of complex contagions means that they are integrally related to community structure, as they spread better in dense regions of a network [3]. Indeed, the analysis in small world models shows that the social network structure is crucial in enabling successful complex contagions [7], [8].

In this paper we provide the first rigorous study of how complex contagions spread in time-evolving graphs. A particularly interesting case of this family is the preferential attachment (PA) model with power law degree distribution. None of the models we studied in [7], [8] had power law degree distribution, which limits how the mathematical results can be used to understand contagions in the real world. Our results in this paper fill this gap.

In a graph with power law degree distribution, the number of nodes with degree d is proportional to $1/d^{\gamma}$, where $\gamma = \Theta(1) > 0$. One of the most studied generative models with a power law degree distribution is the PA model. Price in [9] attributed the appearance of power law degree distributions to the mechanism of "cumulative advantage", known as preferential attachment, phrased by Barabasi *et al.* [10]. The PA model considers an evolving network in which newcomers link to nodes already in the network. In general, a new node chooses its neighbor using the preferential attachment rule with probability $p \in [0, 1]$ and choose a neighbor uniformly randomly with probability $1 - p$. Under the preferential attachment rule, a potential neighbor node is selected with a probability proportional to its current degree. Thus, as the network evolves, high degree nodes have an advantage in attracting even more links.

In a time-evolving network with a power-law distribution, not all the nodes are homogeneous. This mirrors reality in that people may be different in how connected they are, and hence have different positions in a social network. A number of previous works acknowledge such differences and compute the 'network' value of a user, e.g., the expected profit from sales to the customers who may influence others, and so on [11]. This heterogeneity allows us to study the effect of selecting the initial seeds, an aspect of complex contagions not examined in previous theoretical work.

Our results. The main result of this paper is to show that complex contagions spreads in polylogarithmic number of rounds in a general family of time-evolving networks that includes the preferential attachment models [12], [10], [13]. We prove that if the initial seeds are the k oldest nodes in a network of this family, a k-complex contagion covers the entire network of n nodes in $O(\log n)$ steps.

For the PA model, when the probability of creating edges using the preferential attachment rule, p, is in $[0, 1)$ (ref. to Def. 1) we conjecture that the diameter is $\Theta(\log n)$ with high probability, and so our result is tight up to a constant factor[1]. This means that, if the initial seeds are properly chosen, the speed of simple and complex contagions differ only by a constant factor. When $p = 1$, it is known [16], [15] that the diameter is $\Theta(\log n/ \log\log n)$, and so in this setting complex contagions are at most a $\log\log n$ factor slower than simple contagions.

We also show that the choice of the initial seeds is crucial: there exists a polynomial threshold $f(n)$ such that if $o(f(n))$ initial seeds are chosen uniformly at random in the PA model, the contagion almost surely does *not* spread. This is in stark contrast to the case were we only need to infect the k–a constant–oldest nodes. Second, we show that if $\Omega(f(n)\log n)$ initial seeds are randomly infected, then, with high probability, the k oldest nodes become infected and by the above results, the whole graph gets infected in $O(\log n)$ rounds. This signifies not only the importance of the choice of initial seeds, but also the delicacy of the diffusion in a complex contagion.

The oldest nodes in a PA model are likely to have high degrees. However, we remark that it is actually not the power law degree distribution per se that facilitates the spread of complex contagions, but rather the evolutionary graph structure of such models. Indeed, the time-evolving family also includes heavily concentrated degree distributions with the largest degree being only $O(\log n)$, e.g. the PA model with $p = 0$.

While one might hope to categorize all the settings in which complex contagions spread, we show that this is unlikely. We prove that given a graph, a list of initially infected nodes, and a threshold, it is $\mathbf{P}$-complete to decide if the number of infected nodes surpasses the threshold or not.

1. Dommers *et al.* [14] show that, if the exponent of the power-law distribution is ≥ 3, then the PA model has a diameter of $\Theta(\log n)$. Berger *et al.* [15] prove that if $p \in [0, 1)$ in Def. 1, then the exponent of the power-law distribution is ≥ 3. However, while Berger *et al.* use the same PA model as Def. 1, the model in Dommers *et al.* is slightly different. It is beyond the scope of this paper to extend the results of Dommers *et al.* to this setting, but we know of no barriers to doing so.

Thus, in some sense, the best one can do (in the worst-case) is to simulate the contagion.

The main thrust of this paper is theoretical but we believe that the analysis is fundamental in how we understand contagions in the real world. A contagion in reality may involve a lot of factors that are not captured by our model: the variations in personal influence, tie strength, and resistance to social influence, to name a few. Characterizing these factors under a specific contagion scenario using real data sets is an research direction that we are working on simultaneously but is beyond the scope of this paper.

Organization of this paper. In Section 2, we explore some of the related work. Section 3 provides formal definitions and models. We prove our main result on the polylogarithmic speed of k-complex contagions in Sections 5, 6, 7 and 8. However, we first provide an overview of the challenges and techniques for this result in Section 4. Section 9 provides simulation results that match the polylogarithmic speed bound that we prove theoretically. Section 10 addresses the random choice of initial seeds for a k-complex contagion in the PA graph. In Section 11 we present a theorem showing that computing the extent of k-complex contagions in general graphs is $\mathbf{P}$-complete.

2 Related Work

Diffusion of information/viruses has been an active research topic in different areas of science. For a full review, please refer to [7]. We describe the most relevant results here.

Our model of complex contagions belongs to the general family of *threshold models* in the study of diffusions. In the *threshold model*, each node has a threshold on the number of infected neighbors needed to become infected [17] – in a k-complex contagion, all nodes have the same threshold k. The threshold model is motivated by coordination games in which a user benefits from adopting the same behavior as her neighbors, which tend to have threshold strategies in equilibrium. Most studies focus on the stable states, and structural properties that prevent complete adoption of the advanced technology or the better behaviors [18]. Montanari *et al.* [19] is one of the few that relate the steady state convergence speed of the game to the network structure.

Diffusion of simple contagions in PA models has been extensively studied [15], [20], [21]. On the empirical side, there are many studies of diffusion in networks [22], [5]. Most of the studies related to ours examine influence on Twitter. Bakshy *et al.* [23] observed that users who have been influential in the past and have a large in-degree would generate the largest cascades.

When the initial seeds are chosen uniformly randomly, the diffusion is termed as the *bootstrap percolation* [24], [25]. Previous work on bootstrap percolation have examined the Erdos-Renyi graph [26], the random regular graph [27], the configuration model [28] with power-law exponent $\gamma > 3$, and the Chung-Lu model with power-law exponent $2 < \gamma < 3$ [29]. In all these results, when the graphs have constant average degree (similar to our setting), when the size of initial seed set is sublinear, the process will not cover a significant part of the graph. Coja-Oghlan *et al.* in [30] studied the problem of finding a "contagious set", a set whose infection would lead to the infection of the entire graph, in d-regular expander graphs in the bootstrap percolation setting. Under different expansion assumptions, they give upperbounds on the size of contagious sets in the 2-complex contagion diffusion model.

3 Preliminaries

Preferential Attachment Model: There are different definitions of the preferential attachment model, in which the difference lies in the subtle ways that the edges are created. We first explain the *independent model* [12].

Definition 1. *We define the* ***independent preferential attachment model****,* $\mathrm{PA}_{p,m}(n)$*: We start with a complete graph on* $m+1$ *nodes. At each subsequent time step* $t = m+2, \cdots, n$ *a node* v *arrives and adds* m *edges to the existing vertices in the network. Denote the graph containing the first* $n-1$ *nodes as* G_{n-1}*. For each new vertex, we choose* $w_1, w_2, \cdots, w_m$ *vertices, possibly with repetitions from the existing vertices in the graph. Specifically, nodes* $w_1, w_2, \cdots, w_m$ *are chosen independently of each other conditioned on the past. For each* i*, with probability* p *(*$0 \leq p \leq 1$*),* w_i *is selected from the set of vertices of* G_{n-1} *with probability proportional to the vertices' degree in* G_{n-1}*; and with probability* $1-p$*,* w_i *is selected uniformly at random. Then we draw edges between the new vertex and the* w_i*'s. Repeated* w_i*'s cause multiple edges. Note that* $\sum_{v \in V(G_n)} deg(v) = 2mn$*.*

There are two other variations of the PA model. In the *conditional* model [13], a new edge is chosen conditioned on it being different from the other edges already built; in the *sequential* model [31], the m edges of the new node v are built sequentially such that the i-th edge of v is chosen preferentially assuming the previous $i-1$ edges of v have been included in the graph and their degrees are counted. In the following we mainly focus our discussion on the *sequential* PA model. Our analyses will also apply to the other two variations with slight modifications.

k-complex contagion: We define a k-complex contagion process in an undirected graph, where $k = O(1)$. We assume that we are given a graph G, which might have been generated by an evolving process. But in the contagion process, the topology of G is fixed.

Definition 2. *Given a graph* G*, a* ***k-complex contagion*** $\mathrm{CC}(G, k, \mathcal{I})$ *is a contagion that initially infects vertices of* $\mathcal{I}$*, initial seeds, and spreads over graph* G*. The contagion proceeds in rounds. At each round, each vertex with at least* k *infected neighbors becomes infected. The vertices of* $\mathcal{I}$ *are infected in round* 0*. We are interested in the minimum number of rounds for all the nodes to be infected.*

4 Challenges and Proof Overview

We prove that when initial seeds are chosen as the oldest k nodes, k-complex contagions in a family of time-evolving

networks infect every node in $O(\log n)$ rounds. This family includes all the variants of the PA graph. We provide a proof overview before diving into the technical details.

Let D be a graph created according to the PA model (Def. 1). First, let us sketch a proof for $k = 1$, i.e. that with high probability D has diameter $O(\log n)$. Then we show where the analogous proof runs into trouble for $k > 1$. This will motivate the machinery that we develop. Label the vertices $1, 2, 3, \ldots$ according to their arrival order. We sketch a proof that the distance from an arbitrary node v to vertex 1 is $O(\log n)$ w.h.p. and the result follows from a union bound. Consider the following procedure: **a)** Start at v; **b)** Follow the edge out of v whose end point, u has the lowest label; **c)** If the label of u is 1, stop. Otherwise, repeat the procedure for node u. We claim that this procedure ends in $O(\log n)$ steps w.h.p. Consider that at some point, the process is at vertex u. Consider the induced subgraph on the vertices $\{1, 2, \ldots, u\}$. If we have no prior knowledge, then it is easy to show that the lowest labelled neighbor of u will be, in expectation, at most αu for some $\alpha < 1$. The result follows from standard concentration arguments. However, the process has knowledge of the graph when a vertex u is processed. Namely, it knows the neighbors of all the vertices it has previously processed! Fortunately, it is not too hard to show that if all these endpoints have indices greater than u, then the marginal distribution of edges on the induced subgraph of vertices $\{1, 2, \ldots, u\}$ remains unchanged.

Things go awry when we let $k = 2$. First, we need better concentration to be able to handle many nodes at the same time. With $k = 1$, if we get unlucky and the first few steps did not move backward much from v, we are still doing at least as well as when we started. However, when $k = 2$ and the first ℓ steps did not move backward much, we have 2^ℓ vertices to process which is a problem when $\ell = \Omega(1)$.

One solution is to partition the graph into stages. Let stage 0 contain the first k vertices, while stage i contains the vertices labeled between $k(1+\epsilon)^{i-1}$ and $k(1+\epsilon)^i$. Thus, each stage will have a $(1+\epsilon)$ fraction more vertices than the last. The probability that a vertex in stage i does not connect to k vertices in previous stages can be upper bounded by a constant that depends on k and ϵ and thus can be made arbitrarily small. We can show that it takes at most an (expected) constant number of steps to get from one stage to the previous stages. While this is sufficient for the proof to work when $k = 1$, it is not enough when $k \geq 2$. The reason is that only knowing the expectation does not give a tight enough bound when we process many vertices. We need to bound the maximum rather than just the average.

To solve this problem, we model the above process as a ***labeled branching process***, introduced in Section 5. A branching process is a Markov process modeling a population where individuals in generation i produce some number of individuals in generation $i + 1$ according to a probability distribution. In a ***labeled branching process***, each individual has a label, and the probability distribution of producing an offspring is dependent on the labels of the parent/offspring. We intend to couple the random process that creates D with a labeled branching process B. The labels in B are proxies of the stages of nodes in D. After the coupling, the height of D is bounded by the extinction time of B. We use a potential function argument to study the extinction time of the labeled branching process. We show that with high probability, the population becomes extinct in $O(\log n)$ generations. The coupling argument must make correspondence between the nodes/edges in D and nodes/branches in B and thus relies on showing that the marginal probabilities of creating edges in B and in D match. The edges of D are created in the arrival order of a PA graph (Def. 1). However, B reveals nodes/edges from last to first. That is, the root branches (edges) are the first edges to be revealed in a branching process and the root corresponds to the node labeled n. Therefore, the coupling argument should follow a revealing process that processes nodes in the reverse arrival order of the PA graph.

Unfortunately, at this point, more subtle problems arise. As k is increased to 2, new dependencies appear. Say we are processing the 100-th arriving node, which has neighbors with arriving orders 33 and 50. Then when we go to process node 50, we have information about node 33, namely that it connects to node 100. In general, we are processing a node u, but the process has already revealed many outgoing edges from nodes $\{w\}_{w>u}$ to a node $s_{s<u}$, then the outgoing edges of u are more likely to be connected to s in the PA graph conditioned on the information revealed so far. In contrast, in the arrival order of the PA graph, at the time u created its edges, s might not have had a high degree and thus the edges of u would not be likely to be connected to s. This ruins the above approach. To rectify things, we need to be careful about the order in which the edges are revealed. Instead of revealing the neighbors of a particular vertex we query if individual edges (e.g. (u, v)) exist in the graph. By the end, we have queried all the edges, but we do so in a carefully chosen order. We do not "process v" any more. Instead this ordering processes two edge points at a time. However, when we process an edge (u, v) we are able to relate the probability that this edge exists to a probability that it is created in a more natural revelation ordering (similar to the definition of PA).

We formally define a ***revealing process*** that generates a graph according to an arbitrary ordering of edges (Sec. 6). We also describe the above staging schema for such *revealing processes*, which we call the ***staging property***. In Section 7, we show that under a specific ordering of edges, ***backward-forward (BF) ordering***, *revealing processes with staging property* can be coupled with a labeled branching process B. In Section 8, we show that under the BF ordering there exists a revealing process with staging property for the PA graph. Using the results of Sections 5, 6, 7, 8, we can overcome the aforementioned proof difficulties and upperbound the speed of k-complex contagions starting from the first arriving nodes as initial seeds.

5 Labeled Branching Processes

In this section, we describe one of our main tools in analyzing the speed of complex contagions on time-evolving graphs. We define a ***labeled branching process*** and analyze its extinction time.

Definition 3. *For constants m and $0 < \alpha \leq 1$, we call a branching process a $\boldsymbol{B(m, x, \alpha)}$-**labeled branching process**, if*

1) *It starts with one node (root) labeled x at depth 0 (where x is a positive integer);*
2) *at each subsequent depth, every i-labeled node (where $i \neq 0$) produces m children. Each child has label $i-1$ with probability α and label i with probability $1-\alpha$;*
3) *0-labeled nodes produce no children.*

The following lemma bounds the extinction time of a labeled branching process by $O(\log n)$, when there are $x = O(\log n)$ labels and α is a constant satisfying $\alpha > 1-1/m$.

Lemma 1. *If $\alpha > 1-1/m$, and $x = c_1 \log n$ for a constant c_1, then the probability that $B(m, x, \alpha)$ has not died out after depth $t = c_2 \log(n)$ is at most $n^{-(c_3+1)}$, where c_3 is a constant, $c_2 = (c_3 + 1 + c_1/\log_{md}(e))/\log(1/\delta)$, $d = m\alpha/(1-m(1-\alpha))$, and $\delta = m(1-\alpha)+1/m-(1-\alpha)$.*

Proof: We refer to a node as an $\boldsymbol{(i-1)}$-***labeled origin*** if it is $(i-1)$-labeled but its parents are not. Let d be the expected number of $(i-1)$-labeled origin descendants of an i-labeled node v. First note that d does not depend on i. Clearly, any $(i-1)$-labeled children of v are $(i-1)$-labeled origins, and any i-labeled children of v will produce in expectation d descendants that are $(i-1)$-labeled origins. This gives us the equation that $d = m\alpha + m(1-\alpha)d$. Assuming that $\alpha > 1 - 1/m$, we can solve $d = m\alpha/(1 - m(1-\alpha))$. Then by independence, the expected number of 0-labeled leaves of the root of the branching process is d^x.

We define a potential function $\phi(t)$ on the branching process B at time t. Let $N_t(j)$ be the number of j-labeled nodes of B at depth t. Note that $N_0(x) = 1$, and $N_0(j) = 0$ for $0 \leq j \leq x-1$. Let

$$\phi(t) = \sum_{j=1}^{x} N_t(j)(md)^j.$$

We can verify that $\phi(0)$ is a polynomial in n, because $\phi(0) = (md)^x = (md)^{c_1 \log n} = n^{c_1/\log_{md}(e)}$.

Next, we show that this potential function decreases exponentially with the time. Let $\mathcal{F}_t = \sigma(\mathcal{N}_t(0), \ldots, \mathcal{N}_t(x))$ be the sigma algebra generated by the random variables $\{\mathcal{N}_t(0), \ldots, \mathcal{N}_t(x)\}$.

Claim 2. $\mathbb{E}[\phi(t+1)|\mathcal{F}(t)] \leq \delta\phi(t)$, *where* $\delta = m(1-\alpha) + 1/m - (1-\alpha)$.

Proof: At level t, a node v of label i contributes $(md)^i$ to $\phi(t)$ for depth t. Node v's contribution to $\phi(t+1)$ at depth $t+1$ is at most $m(\alpha(md)^{i-1} + (1-\alpha)(md)^i)$ in expectation. We factor $(md)^i$ out, insert the value for d from above and simplify to get δ. Notice that as long as $\alpha > 1 - 1/m$ we have that $\delta < 1$. □

We prove by induction that $\mathbb{E}[\phi(t)] < \delta^t\phi(0)$.

Base case $t = 1$. $\mathcal{F}(0)$ is the trivial sigma-algebra so we drop it in our notation. $\mathbb{E}[\phi(1)] \leq \delta\phi(0)$.

Inductive case. Assume that $\mathbb{E}[\phi(z)] \leq \delta^z\phi(0)$. We have

$$\begin{aligned}\mathbb{E}[\phi(z+1)] &= \mathbb{E}[\mathbb{E}[\phi(z+1)|\mathcal{F}_z]] \\ &\leq \mathbb{E}[\delta\phi(z)] \quad \text{by Claim 2,} \\ &\leq \delta^{z+1}\phi(0) \quad \text{by induction hypothesis.}\end{aligned}$$

Let $c_2 = (c_3 + 1 + c_1/\log_{md}(e))/\log(1/\delta)$. Then $\mathbb{E}[\phi(c_2 \log n)] = \delta^{c_2 \log n}\phi(0) < n^{-(c_3+1)}$. If a node at time $t = c_2 \log n$ existed it would contribute at least $(md)^1 \geq 1$ to ϕ. Thus, by Markov's inequality, we conclude that the probability that there are any nodes on the level t is at most $n^{-(c_3+1)}$. □

Our notion of *labeled branching process* is closely related to the notion of *multitype Galton-Watson branching processes* in the Markov process literature [32]. Although the extinction time of multitype processes have been studied before [32], this literature has not explored the extinction time when the number of types in the process is not a constant. In our setting however, the number of types (labels) is $\Omega(1)$ and Lemma 1 can be generalized to any number of labels bigger than $\log n$ with slight modification. In this sense, Lemma 1 might be useful in its own right in multitype Galton-Watson branching processes theory.

6 Revealing Processes and the Staging Property

In this section we first define the notion of a ***revealing process***. A revealing process takes a graph generation process over time and changes the order in which the random edges of the graph are *revealed/created*. We then introduce the notion of ***revealing processes that possess a staging property***. A graph that is generated by a *revealing process with staging property* can be divided into a number of ordered *stages* such that the probability that an outgoing edge of a vertex from stage i does NOT land in preceding stages is bounded. Under some specific ordering of edges, *revealing processes with staging property* bear similarities to the notion of the *labeled branching process* that we analyzed in Sec. 5. We use this similarity and couple the two random processes. We finally introduce the aforementioned specific ordering on all the possible edge incidences, which we call the ***backward-forward ordering***. Let $\mathcal{G}$ be a distribution of graphs that is defined by a graph generation process over time.

Definition 4. *We will say that distribution $\mathcal{G}$ $\boldsymbol{m}$-**generates** a graph over time if: i) The process $\mathcal{G}$ starts with a complete graph on $m+1$ nodes at time 0. At each time step at most one vertex arrives. The i-th arriving node is labeled index i. ii) Each arriving vertex v has at least m edges to previously added vertices[2]. For each edge $v \to u$, $u < v$.*

While the edges of a graph H generated by $\mathcal{G}$ are created in a specific ordering, it is possible to ***reveal*** the edges of H in other arbitrary orderings. We rigorously define such

2. These edges are possibly generated in a randomized way.

a ***revealing process*** through a *filtration process with an adapted probability measure*.

Definition 5. *Let n be the size of the graph to be genrated. Let $u < v \in [n]$ and $j \in [m]$. Each triple (u, v, j) is called an* ***edge incidence triple****.*

An *edge incidence triple* (u, v, j) corresponds to the j-th edge that could be (potentially) issued by node v linking to u in the (randomly) generated graph. Let T be the number of edge incidence triples in an m-generated graph of size n.

Definition 6. *A total* ***ordering*** $\mathcal{O}$ *on the edge incidence triples is a function $\mathcal{O} : \mathbb{N} \to \mathbb{N}^3$ such that there is a natural bijection between $\mathcal{O}([T])$ and the set of edge incident triples. We treat $\mathcal{O}$ as being 0-indexed.*

Definition 7. *We define a* ***revealing process*** *through a family of σ-algebras (filtration).*

1) *Let $\mathbf{\Omega} = \{(u, v, j) | (u, v, j)$ is an edge incidence triple$\}$ be the set of all possible outcomes and $\boldsymbol{\mathcal{F}_{\mathcal{O},0} = \{\emptyset, \Omega\}}$ be the trivial σ-algebra. For each i, let* **event** $\boldsymbol{\mathcal{E}_{\mathcal{O},i}}$ *to be the yes/no answer to the* outcome $\mathcal{O}[i]$.
2) *For each $i \in [T]$, we take $\boldsymbol{\mathcal{F}_{\mathcal{O},i}}$ to be the* natural *filtration: $\boldsymbol{\mathcal{F}_{\mathcal{O},i} = \sigma(\mathcal{E}_{\mathcal{O},0}, \ldots, \mathcal{E}_{\mathcal{O},i-1})}$. Since $\mathcal{F}_{\mathcal{O},i} \subseteq \mathcal{F}_{\mathcal{O},i+1}$, $\boldsymbol{(\Omega, \mathcal{F}_{\mathcal{O},t})}$ forms a* ***filtration****.*
3) *Let $\boldsymbol{\{Q_{\mathcal{O},t}\}}$* ***be a family of probability measures adapted*** *to the filtration $(\Omega, \mathcal{F}_{\mathcal{O},t})$, meaning $Q_{\mathcal{O},i}$ is* a probability measure *with respect to $\mathcal{F}_{\mathcal{O},i}$ for each i.*

Then, $(\Omega, \mathcal{F}_{\mathcal{O},t}, Q_{\mathcal{O},t})$ is a ***revealing process*** *which generates a graph G with n vertices.*

From Def. 7 a *revealing process* is also a graph generation process, but a process where the size of the graph, the arriving order of the nodes and their labels is already determined from the beginning (Def. 4). We intentionally call it a *revealing process* because we are interested in defining $\{Q_{\mathcal{O},t}\}$ such that $(\Omega, \mathcal{F}_{\mathcal{O},t}, Q_{\mathcal{O},t})$ is compatible with some distribution of graphs $\mathcal{G}$, that is, it generates graphs with the same probability distribution as $\mathcal{G}$.

Definition 8. *Let G be a graph generated by a* revealing process *$(\Omega, \mathcal{F}_{\mathcal{O},t}, Q_{\mathcal{O},t})$. We say that $(\Omega, \mathcal{F}_{\mathcal{O},t}, Q_{\mathcal{O},t})$ has the* $\boldsymbol{(r, m, \alpha)}$***-staging property*** *if there exists an ordering on the vertices of G and an ordered partition $S_0, S_1, \ldots, S_r$ of the nodes into $r + 1$ stages (the nodes in stage i are ordered before those of $i + 1$) such that:*

i) $|S_0| < \log(n)$;
ii) Each vertex has m edges to nodes prior in the ordering;
iii) Assume that node v is in stage i. Let W be the set of nodes in stage i that precede v (they have smaller indices). Let $\boldsymbol{R_{\mathcal{O},(W,v,j)}}$ ***be the probability of the jth outgoing edge of v*** **landing** ***in the set W****, i.e.:*

$$R_{\mathcal{O},(W,v,j)} = \sum_{u \in W} Q_{\mathcal{O},h}(\mathcal{E}_{\mathcal{O},h}) \quad \text{where } h = \mathcal{O}^{-1}[(u, v, j)].$$

Then it must be that $R_{\mathcal{O},(W,v,j)} \leq 1 - \alpha$.

A graph G generated by $(\Omega, \mathcal{F}_{\mathcal{O},t}, P_{\mathcal{O},t})$ with staging property is said to be $\boldsymbol{(r, m, \alpha)}$***-staged****.*

The ***backward-forward ordering*** sorts the incidence triples by the decreasing order of the edge-receiving vertices, and for nodes with the same landing vertices sorts them by the increasing order of the edge-issuing vertices. We are interested in the BF ordering, because $(\Omega, \mathcal{F}_{BF,t}, Q_{BF,t})$ with (r, m, α)-staging property have similar stochastic characteristics to $B(m, r, \alpha)$-labeled branching processes.

Definition 9. *A* ***backward-forward (BF) ordering*** *on triples is as follows: $(u_1, v_1, j_1) < (u_2, v_2, j_2)$ if a) $u_1 > u_2$ or; b) if $u_1 = u_2$, and $v_1 < v_2$ or; c) if $u_1 = u_2$, $v_1 = v_2$, and $j_1 < j_2$.*

7 Speed of Complex Contagions in Staged Graphs Generated by a BF-ordered Revealing Process

In this section, we use the similarities between a $(\Omega, \mathcal{F}_{BF,t}, Q_{BF,t})$ with (r, m, α)-staging property and a $B(m, r, \alpha)$-labeled branching process to form a coupling argument between them. Such a coupling allows us to apply Lem. 1 and bound the length of the longest path from a node u to node 1 in a graph generated by the former process by the depth of a corresponding labeled branching process. We additionally show that the speed of k-complex contagions in an *m-generated network* starting with the oldest nodes as the initial seeds is bounded by the length of the longest path to the initial seeds in a graph.

Let G be a graph generated by a *revealing process* $(\Omega, \mathcal{F}_{BF,t}, Q_{BF,t})$ *with $(O(\log n), k, \alpha)$-staging property* (where $\alpha > 1 - 1/k$). The following theorem states that starting from the oldest nodes, a k-complex contagion on G has polylogarithmic speed with high probability. It is noteworthy to observe that the same scenario also happens for k-complex contagions on graphs generated by a *revealing process* $(\Omega, \mathcal{F}_{BF,t}, Q_{BF,t})$ *with $(O(\log n), m, \alpha)$-staging property* where $m \geq k$. We assume both k, m are constants.

Theorem 3. *Let G be a graph generated by a* revealing process $(\Omega, \mathcal{F}_{BF,t}, Q_{BF,t})$ with (x, k, α)-staging property *where $\alpha > 1 - 1/k$, and $x = O(\log n)$. Let $\mathcal{I}$ be the set of k first arrived vertices in G according to Def. 4. A k-complex contagion $\mathrm{CC}(G, k, \mathcal{I})$ will infect the entire graph with probability $1 - 1/n^{c_3}$ in less than $c_2 \log n$ number of rounds where $c_2 = (c_3 + 1 + x/(\log n \log_{kd}(e)))/\log(1/\delta) + 1$, $d = k\alpha/(1 - k(1 - \alpha))$, and $\delta = k(1 - \alpha) + 1/k - (1 - \alpha)$.*

Proof: Consider a directed subgraph of G, in which we only keep the k edges from each vertex pointing to the smaller labeled vertices. We say u follows v if there is a directed edge from u to v. Node u becomes infected in the next round if it follows k infected neighbors. By removing extra edges and making the propagation directed we only make the contagion spread slower. Thus, we get an upper bound on the speed.

We prove by induction that the time it takes to infect a vertex v is no greater than the length of the *longest* path

from v to the vertices in $\mathcal{I}$ in this directed graph. The first k vertices have longest paths of length 0 to $\mathcal{I}$ and are infected at time 0. Assume the hypothesis for nodes with path length ℓ. Let $\ell+1$ be the length of the longest path from a vertex u to $\mathcal{I}$. Then the k out-neighbors of u have paths of length at most ℓ to the first k vertices. By induction, they are infected at time ℓ, and so is u at time $\ell+1$.

Pick an arbitrary node u. We will show that u is infected in time $O(\log n)$ with probability $1-1/n^{c_3+1}$. Then taking a union bound on all nodes, we will have our result. Note that if u is in stage 0, then it will be infected in time $\log n$ because stage 0 has only $\log n$ nodes and the path back to the original k vertices makes progress at each step, and thus takes time at most $\log n$. Next, we let u be in stage $i > 0$. We will bound the time t it takes all paths starting at u to get back to stage 0, and this will bound the time to infect u by $t + \log(n)$. Next, we only need to show that $t \leq (c_2 - 1)\log n$ with probability at least $1 - n^{-(c_3+1)}$.

Coupling the longest path with the branching process. We will create a coupling so that the longest path from u to stage 0 is bounded by the time it takes an appropriate labeled branching process to terminate. Let $B(y)$ denote a $B(k, i, \alpha)$-labeled branching process rooted at node y (ref. to Def. 3). We consider the branching process $B(\hat{u})$ that is rooted at node $\hat{u}$ labeled i. Node $\hat{u}$ corresponds to the node u in G and because u is in stage i, $\hat{u}$ is also labeled i. We use the same letter to show correspondence between the branching process and the graph nodes, while node letters in $B(u)$ will carry the $\hat{}$ hat. G is generated by $(\Omega, \mathcal{F}_{BF,t}, Q_{BF,t})$, which means that the random choices corresponding to edge incidence triples are revealed according to the BF ordering.

We couple the j-th branch of $\hat{u}$ to the j-th neighbor of u in G. If the j-th neighbor of u is NOT in stage i, then we couple this to the j-th branch of $\hat{u}$ so that its label is $i-1$. This coupling is truthful to the marginal probabilities:

a) The probability that the j-th edge of u is in stage i, $R_{BF,(W,v,j)}$, is at most $1-\alpha$ according to the staging property (Def. 8);
b) and the probability that $\hat{u}$ has a branch of label i is $1-\alpha$.

Consider a fixed node v in G; we explain how we find the corresponding node $\hat{v}$ in the branching process. We wait until all the oriented edge triples (v, w, k) have been revealed by $(\Omega, \mathcal{F}_{BF,t}, Q_{BF,t})$. When all these triples have been revealed, we know if v has:

1) No corresponding parent in the branching process tree;
2) Exactly one corresponding parent $\hat{p}$ in the branching process tree;
3) More than one parent in the branching process tree.

We treat these cases as follows:

1) We don't couple the probabilities;
2) We correspond the child of $\hat{p}$ with v and name it $\hat{v}$. We couple the events as we described above;
3) We know which parent is deeper in branching process, we couple with this branch and ignore the rest.

The detailed coupling procedure maintains the invariant that the label of $\hat{v}$ is always $\geq$ the stage of the corresponding node v in G. Lemma 1 states that the $B(k, x, \epsilon)$-labeled branching process $B(\hat{u})$ dies out after $(c_2 - 1)\log n$ levels with probability at least $1 - n^{(-c_3+1)}$. Hence, the length of the longest path from u to initial nodes is also $\leq (c_2 - 1)\log n$ with probability at least $1 - n^{-(c_3+1)}$. □

8 Staging Property in PA Graphs

In this section, we first show that there exists a family of probability measures $\{P_{BF,t}\}$ adapted to the filtration process $(\Omega, \mathcal{F}_{BF,t})$ such that the revealing process $(\Omega, \mathcal{F}_{BF,t}, P_{BF,t})$ is *compatible* with the *sequential preferential attachment model*, $\mathrm{PA}_{p,m}(n)$, which means that $(\Omega, \mathcal{F}_{BF,t}, P_{BF,t})$ generates graphs with the same probability distribution as $\mathrm{PA}_{p,m}(n)$. As a stand-alone result, this lemma states that by knowing the size of the network, it is possible to generate a PA graph backwards. We then prove that $(\Omega, \mathcal{F}_{BF,t}, P_{BF,t})$ satisfies a $(\log n, m, 2/3)$-staging property. By combining these two results with Theorem 3, we get that a k-complex contagion with initial seeds as the oldest nodes in the $\mathrm{PA}_{p,m}(n)$ infects all the nodes in $O(\log n)$ rounds with high probability.

The ***Arrival-Time (AT) ordering*** on edge incidence triples is a sequential ordering of all possible edges that corresponds to the order that they are built in the graph generation process $\mathcal{G}$. That is, a node that arrives earlier will have its edges placed earlier. For the edges placed by the same node v, we sort them according to the inverse arriving order of their tails.

Definition 10. *An **arrival-time (AT) ordering** on triples is as follows:* $(u_1, v_1, j_1) < (u_2, v_2, j_2)$ *if a)* $v_1 < v_2$ *or; b) if* $v_1 = v_2$ *and* $j_1 < j_2$ *or; c) if* $v_1 = v_2$ *and* $j_1 = j_2$ *and* $u_1 > u_2$.

Before defining the aforementioned family of probability measures $\{P_{BF,t}\}$, we give a set of useful definitions.

Definition 11. *Let* $\mathcal{O}[i] = (u, v, j)$. *For each* $u \in [n]$*:*

- $\mathbf{rev}_{\mathcal{O},i}(v, j) = \{w | \{w\} \in \mathcal{F}_{\mathcal{O},i}$ ***and*** $\exists\, x, w = (x, v, j)\}$ *as the set of* edge incidence triple outcomes *corresponding to the* j*-th outgoing edge of node* v *that the filtration* $(\Omega, \mathcal{F}_{\mathcal{O},t})$ *has **revealed** thus far.*
- $\mathbf{exam}_{\mathcal{O},i}(v, j) = \{x \,|\{w\} \in \mathcal{F}_{\mathcal{O},i}$ ***and*** $w = (x, v, j)\}$ *as the set of vertices we have* examined *thus far to determine the receiving end of the* j*-th outgoing edge of vertex* v.
- *Let* $\mathbf{1}_{\mathcal{E}_{\mathcal{O},i}}(w)$ *be the **indicator function** for an outcome* $w \in \Omega$*, i.e.:*

$$\mathbf{1}_{\mathcal{E}_{\mathcal{O},i}}(w) = \begin{cases} 1 & \text{if } w \in \mathcal{E}_{\mathcal{O},i} \\ 0 & \text{if } w \notin \mathcal{E}_{\mathcal{O},i} \end{cases}$$

Since each event $\mathcal{E}_{\mathcal{O},i}$ *is concerned with exactly one outcome in* Ω*, we can drop the* (w) *from our notation and use* $1_{\mathcal{E}_{\mathcal{O},i}}$.

- *We define the in-degree of a node in $\mathcal{F}_{\mathcal{O},i}$ based on the edges originated from a range of vertices $[x_1, x_2]$ as*

$$\mathbf{ideg}_{\boldsymbol{\mathcal{O}},\boldsymbol{i}}(\boldsymbol{u}, [\boldsymbol{x_1}, \boldsymbol{x_2}]) = \sum_{0\le r\le i,\ \exists\ y,z\ s.t.\ \mathcal{O}[r]=(u,y,z),\ \ x_1\le y\le x_2,\ 0\le z\le m} 1_{\mathcal{E}_{\mathcal{O},r}}$$

We define a family of *probability measures* $\{P_{\mathcal{O},t}\}$ *adapted* to the filtration $(\Omega, \mathcal{F}_{\mathcal{O},t})$ that mimics the *preferential attachment* rule (Def. 1)[3]. However, this does NOT guarantee that $(\Omega, \mathcal{F}_{\mathcal{O},t}, P_{\mathcal{O},t})$ is compatible with $\mathrm{PA}_{p,m}(n)$ for an arbitrary ordering.

Definition 12. *Let $\mathcal{O}[i] = (u, v, j)$. The **probability measure $\boldsymbol{P_{\mathcal{O},i}}$** is as follows.*

***1)** If $\sum_{t=0}^{i}\sum_{\forall x,\ w=(x,v,j)} 1_{\mathcal{E}_{\mathcal{O},t}}(w) > 0$ then the j-th outgoing edge of v has already been created and thus*

$$\boldsymbol{P_{\mathcal{O},i}(\mathcal{E}_{\mathcal{O},i}) = 0,}$$

***2)** Otherwise, the edge (u, v, j) is created* preferentially *with probability $\boldsymbol{q_{\mathcal{O},i}(u, v, j)}$ and* uniformly *with probability $\boldsymbol{1 - q_{\mathcal{O},i}(u, v, j)}$.*

$$\boldsymbol{P_{\mathcal{O},i}(\mathcal{E}_{\mathcal{O},i})} = q_{\mathcal{O},i}(u,v,j)\frac{m + \mathrm{ideg}_{\mathcal{O},i}(u,[u,v])}{2m(v-1)+2(j-1)} + \frac{(1-q_{\mathcal{O},i}(u,v,j))}{v-1-|\mathrm{exam}_{\mathcal{O},i}(v,j)|}$$

We define $q_{\mathcal{O},i}(u, v, j)$ to be an updated version of p in Def. 1 using Bayes' rule. This is necessary because the probability that the edge (u, v, j) is chosen preferentially *changes based on the information in $\mathcal{F}_{\mathcal{O},i}$ that the jth edge of v was not in* $\mathrm{rev}_{\mathcal{O},i}(v, j)$.

$$\boldsymbol{q_{\mathcal{O},i}(u,v,j)} = \frac{ps_{\mathcal{O},i}(u,v,j)}{ps_{\mathcal{O},i}(u,v,j) + (1-p)\frac{v-1-|\mathrm{exam}_{\mathcal{O},i}(v,j)|}{v-1}}$$

$$\boldsymbol{s_{\mathcal{O},i}(u,v,j)} = 1 - \sum_{x\in\mathrm{exam}_{\mathcal{O},i}(v,j)} \frac{m + \mathrm{ideg}_{\mathcal{O},i}(x,[x,v])}{2m(v-1)+2(j-1)}$$

Note that the above family of probability measures is only *preferential* in a *pseudo post-hoc* sense in that it assumes that the total sum of degrees of nodes before v excluding the out-going edges of v is $m(v-1)$ at all times during the filtration. Also, note that a graph G generated by $(\Omega, \mathcal{F}_{AT,t}, P_{AT,t})$ is a sequential PA graph, $\mathrm{PA}_{p,m}(n)$. We now argue that $(\Omega, \mathcal{F}_{BF,t}, P_{BF,t})$ is compatible with $(\Omega, \mathcal{F}_{AT,t}, P_{AT,t})$.

Lemma 4. *Fix any m-generated graph H of size n. Let $P_{AT}(H)$ and $P_{BF}(H)$ denote the probability that H is generated by $(\Omega, \mathcal{F}_{AT,t}, P_{AT,t})$ and $(\Omega, \mathcal{F}_{BF,t}, P_{BF,t})$ respectively. Then $P_{AT}(H) = P_{BF}(H)$.*

Proof: We can write the generation probability of H as the conjunction of the edge incidences (occurring/not occurring). There is a unique set of *fixed* edge incidence events $\mathcal{E}_{AT,0}, \mathcal{E}_{AT,1}, \ldots, \mathcal{E}_{AT,T}$ that exactly correspond to the occurrence of the graph H[4]. Recall that $\mathcal{F}_{AT,i}$ is the set of events $\{\mathcal{E}_{AT,0}, \mathcal{E}_{AT,1}, \ldots, \mathcal{E}_{AT,i-1}\}$. So, we have

$$P_{AT}(H) = P_{AT,0}(\mathcal{E}_{AT,0}|\mathcal{F}_{AT,0}) \times \ldots P_{AT,T}(\mathcal{E}_{AT,T}|\mathcal{F}_{AT,T}). \quad (1)$$

$P_{BF}(H)$ is also concerned with the same unique set of edge incidence events albeit in a different order $BF[0], BF[1], \ldots, BF[T]$:

$$P_{BF}(H) = P_{BF,0}(\mathcal{E}_{BF,0}|\mathcal{F}_{BF,0}) \times \ldots P_{BF,T}(\mathcal{E}_{BF,T}|\mathcal{F}_{BF,T}). \quad (2)$$

We prove by induction on the index of the BF ordering that if $AT[i_1] = BF[i_2] = (u, v, j)$, then

$$P_{BF,i_2}(\mathcal{E}_{BF,i_2}|\mathcal{F}_{BF,i_2}) = P_{AT,i_1}(\mathcal{E}_{AT,i_1}|\mathcal{F}_{AT,i_1})^{5}. \quad (3)$$

The claim of the lemma will follow as $P_{BF}(H)$ and $P_{AT}(H)$ will be the multiplication of equal probabilities in Equations 1 and 2.

We first prove $\mathrm{exam}_{AT,i_1}(v, j) = \mathrm{exam}_{BF,i_2}(v, j)$.

$\mathrm{exam}_{AT,i_1}(v, j) = \{u + 1, u + 2, \ldots, v - 1\}$ By Def. 10

$\mathrm{exam}_{BF,i_2}(v, j) = \{u + 1, u + 2, \ldots, v - 1\}$ By Def. 9

$\Rightarrow \mathrm{exam}_{AT,i_1}(v, j) = \mathrm{exam}_{BF,i_2}(v, j)$

Base Case: $i_2 = 0$**.** We consider the first edge incidence triple revealed in the BF ordering, $BF[0] = (n-1, n, 1)$. Since $\mathrm{exam}_{BF,0}(n, 1) = \emptyset$, and $\mathrm{ideg}_{BF,1}(n-1, [n-1, n]) = 0$ we have that:

$$s_{BF,0}(n-1,n,1) = 1$$
$$q_{BF,0}(n-1,n,1) = p$$
$$P_{BF,0}(\mathcal{E}_0) = p\frac{m}{2m(n-1)} + (1-p)\frac{1}{n-1}.$$

There exists an index h such that $AT[h] = (n-1, n, 1)$. By Def. 10, $AT^{-1}[(f, n, 1)] > AT^{-1}[(n-1, n, 1)] = h$ for $f < n-1$. Hence, $\mathrm{exam}_{AT,h}(n, 1) = \emptyset$. Again, by Def. 10, h is the smallest index such that the AT ordering looks at the incoming edges to node $n-1$, hence $\mathrm{ideg}_{AT,h}(n-1, [n-1, n]) = 0$. Therefore, we have:

$$s_{AT,h}(n-1,n,1) = 1$$
$$q_{AT,h}(n-1,n,1) = p$$
$$P_{AT,h}(\mathcal{E}_h) = p\frac{m}{2m(n-1)} + (1-p)\frac{1}{n-1} = P_{BF,0}(\mathcal{E}_0).$$

Inductive Case. We run induction on the index d of the edge incidence events in the backward forward order. Suppose the d-th edge incident event in the backward forward order is the d'-th event in the arriving order. That is, $BF[d] = (u, v, j) = AT[d']$. By induction hypothesis, for any index e with $e < d$, let e' be the index of the same event in the arriving order, i.e., $BF[e] = AT[e']$. Then

$$P_{BF,e}(\mathcal{E}_{BF,e}|\mathcal{F}_{BF,e}) = P_{AT,e'}(\mathcal{E}_{AT,e'}|\mathcal{F}_{AT,e'}).$$

We have shown that $\mathrm{exam}_{BF,d}(v, j) = \mathrm{exam}_{AT,d'}(v, j)$. We now exhibit that for $\forall x \in \mathrm{exam}_{BF,d}(v, j)$, $\mathrm{ideg}_{BF,d}(x, [x, v]) = \mathrm{ideg}_{AT,d'}(x, [x, v])$.

First, by closely looking at Def. 9, we see that $\forall x \in \mathrm{exam}_{BF,d}(v, j) = \{u + 1, u + 2, \ldots, v - 1\}$, we have

$$\begin{aligned}&\mathrm{ideg}_{BF,d}(x,[x,v])|\mathcal{F}_{BF,d}\\ &= \sum_{0\le r\le d,\ \exists\ y,z\ \text{s.t.}\ BF[r]=(x,y,z),\ x\le y\le v,\ 0\le z\le m} 1_{\mathcal{E}_{BF,r}}\\ &= \sum_{z=1}^{j-1}\sum_{y=v} 1_{\mathcal{E}_{BF,t}} + \sum_{z=1}^{m}\sum_{y=x}^{v-1} 1_{\mathcal{E}_{BF,t}},\\ &\qquad \text{where } t = BF^{-1}[(x,y,z)].\end{aligned}$$

3. This means that $P_{\mathcal{O},i}$ is *a probability measure* with respect to $\mathcal{F}_{\mathcal{O},i}$ for each i.

4. T is the total number of edge incidence triples

5. Note that $\mathcal{E}_{BF,i_2} = \mathcal{E}_{AT,i_1}$ are the same edge incidence events but we denote them differently for convenience.

Also, from Def. 10, we have that $\forall x \in \text{exam}_{AT,d'}(v,j)$:

$$\begin{aligned}
&\text{ideg}_{AT,d'}(x,[x,v])|\mathcal{F}_{AT,d'} \\
&\quad= \sum_{0\le r\le d',\ \exists\ y,z \text{ s.t. } AT[r]=(x,y,z),\ x\le y\le v,\ 0\le z\le m} 1_{\mathcal{E}_{AT,r}} \\
&\quad= \sum_{z=1}^{j-1}\sum_{y=v} 1_{\mathcal{E}_{AT,t}} + \sum_{z=1}^{m}\sum_{y=x}^{v-1} 1_{\mathcal{E}_{AT,t}}, \\
&\qquad \text{where } t = AT^{-1}[(x,y,z)].
\end{aligned}$$

Note that by Def. 9 and Def. 10 all indices t, t' above are smaller than d and d' respectively. Since the edge incidence events we are conditioning on are fixed in H, both $\text{ideg}_{AT,d'}(x,[x,v])|\mathcal{F}_{AT,d'}$ and $\text{ideg}_{BF,d}(x,[x,v])|\mathcal{F}_{BF,d}$ compute to a number and equal each other. Hence we have:

$$\begin{aligned}
&\forall\ x \in \text{exam}_{BF,d}(v,j), \qquad (4)\\
&\text{ideg}_{BF,d}(x,[x,v])|\mathcal{F}_{BF,d} = \text{ideg}_{AT,d'}(x,[x,v])|\mathcal{F}_{AT,d'}.
\end{aligned}$$

By Definition 12, we get that

$$s_{BF,d}(u,v,j)|\mathcal{F}_{BF,d} = s_{AT,d'}(u,v,j)|\mathcal{F}_{AT,d'} \qquad (5)$$
$$q_{BF,d}(u,v,j)|\mathcal{F}_{BF,d} = q_{AT,d'}(u,v,j)|\mathcal{F}_{AT,d'}. \qquad (6)$$

Finally, we have that

$$\begin{aligned}
&\text{ideg}_{BF,d}(u,[u,v])|\mathcal{F}_{BF,d} \\
&\quad= \sum_{0\le r\le d,\ \exists\ y,z \text{ s.t. } BF[r]=(u,y,z),\ u\le y\le v,\ 0\le z\le m} 1_{\mathcal{E}_{BF,r}} \\
&\quad= \sum_{z=1}^{j-1}\sum_{y=v} 1_{\mathcal{E}_{BF,t}} + \sum_{z=1}^{m}\sum_{y=u}^{v-1} 1_{\mathcal{E}_{BF,t}}, \\
&\qquad \text{where } t = BF^{-1}[(u,y,z)], \\
&\text{ideg}_{AT,d'}(u,[u,v])|\mathcal{F}_{AT,d'} \\
&\quad= \sum_{0\le r\le d,\ \exists\ y,z \text{ s.t. } AT[r]=(u,y,z),\ u\le y\le v,\ 0\le z\le m} 1_{\mathcal{E}_{AT,r}} \\
&\quad= \sum_{z=1}^{j-1}\sum_{y=v} 1_{\mathcal{E}_{AT,t}} + \sum_{z=1}^{m}\sum_{y=u}^{v-1} 1_{\mathcal{E}_{AT,t}}, \\
&\qquad \text{where } t = AT^{-1}[(u,y,z)].
\end{aligned}$$

Again, because edge incidence events we are conditioning on are fixed in H, both these terms compute to a number and equal each other. By Definition 12, $P_{BF,d}(\mathcal{E}_{BF,d}|\mathcal{F}_{BF,d})$ and $P_{AT,d'}(\mathcal{E}_{AT,d'}|\mathcal{F}_{AT,d'})$ are equal in every component and thus are equal to each other. This completes the induction proof. □

Lemma 5. *The $(\Omega, \mathcal{F}_{BF,t}, P_{BF,t})$ revealing process satisfies the $(\log n, m, 2/3)$-staging property.*

Proof: We define the stages as follows. Stage S_0 contains the first 2 nodes and for each i, $S_i = \{v_s | (3/2)^i < s \le (3/2)^{i+1}\}$. Let W be the set of nodes in stage i that arrived before v. Let $\mathcal{H}_{BF}$ be the space of all graphs generated by the $(\Omega, \mathcal{F}_{BF,t}, P_{BF,t})$ process. By Lemma 4 we have that $\mathcal{H}_{BF} = \mathcal{H}_{AT}$. Given a fixed graph H, let $1(v^j \to W)$ indicate whether the j-th outgoing edge of v in H lands in the set W. We can rewrite $R_{BF,(W,v,j)}$ as

$$\begin{aligned}
R_{BF,(W,v,j)} &= \sum_{u\in W} Q_{BF,h}(\mathcal{E}_{BF,h}) \text{ where } h = BF^{-1}[(u,v,j)] \\
&= \sum_{H\in\mathcal{H}_{BF}} P_{BF}(H) 1(v^j \to W) \\
&= \sum_{H\in\mathcal{H}_{AT}} P_{AT}(H) 1(v^j \to W) \quad \text{by Lemma 4} \\
&= R_{AT,(W,v,j)}. \qquad (7)
\end{aligned}$$

We bound $R_{AT,(W,v,j)}$ by directly using Def. 1. In the case that the edge of v was chosen uniformly, the probability of choosing an edge in stage S_{i-1} or smaller is greater than $2/3$. In the case that the edge was chosen preferentially, we know that the total sum of nodes before u is $2m(v-1)$, and the sum of degrees for the nodes in stage $i-1$ or smaller is at least $2m(3/2)^i$. Since $v < (3/2)^{i+1}$, then the probability that the preferentially selected neighbor is among the first $i-1$ stages is bigger than $2/3$. Hence $R_{BF,(W,v,j)} = R_{AT,(W,v,j)} < 1/3$. □

Theorem 6 is a direct corollary of Thm. 3 and Lem. 5.

Theorem 6. *Let $\mathcal{I}$ be the set of first k arrived vertices in the $\text{PA}_{p,m}(n)$ graph and let $k \le m = O(1)$, and $0 \le p \le 1$. A k-complex contagion $\text{CC}(G,k,\mathcal{I})$ infects all of $\text{PA}_{p,m}(n)$ in $O(\log n)$ rounds w.h.p.*

Remark 1. *It is noteworthy that the family of graphs $\text{PA}_{p,m}(n)$ does not always generate a power-law graph. In fact, the $\text{PA}_{0,m}(n)$ model generates a heavily concentrated degree distribution with the largest degree being $O(\log n)$. We emphasize that our results about the spread of complex contagions hold for all the members of this family regardless of them having a power-law distribution or not*[6].

9 Simulations

We provide simulation results that show that our theoretical analysis of the polylogarithmic speed of k-complex contagions in PA networks (Thm. 6) also matches a simulated contagion. Figure 1 shows that the number of rounds that it takes a k-complex contagion to infect $\text{PA}_{0.7,3}(n)$, a PA network of size n and parameters $p = 0.7$ and $m = 3$, lies closely with a $O(\log n)$ function.

10 Bootstrap Percolation in the PA Model

In this section, we focus on bootstrap percolation in the independent PA model (Def. 1)[7]. In other terms, we analyze complex contagions when the initial seeds are chosen uniformly at random. First, we show that there exists a polynomial threshold $f(n)$ such that if $o(f(n))$ initial seeds are chosen uniformly at random, the contagion almost surely

6. The $G(n,p)$ graph also has a heavily concentrated distribution with largest degree being $O(\log n)$. However, unlike $\text{PA}_{0,m}(n)$, deterministic choice of a constant number of initial seeds in the $G(n,p)$ would not cause complex contagions to spread [26].

7. However, the same result (Theorem 7) applies to the other two variations of the PA model (Definition 1) with slight modification.

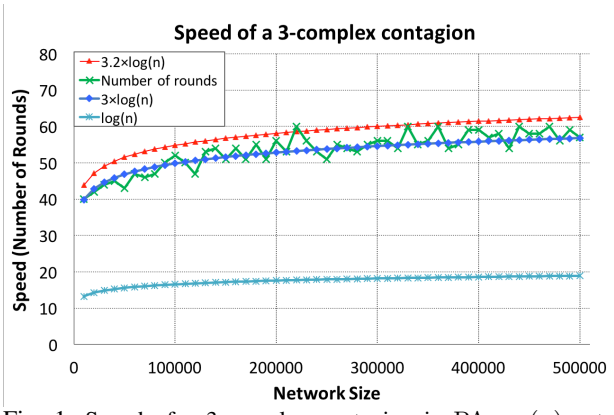

Fig. 1: Speed of a 3-complex contagion in $\mathrm{PA}_{0.7,3}(n)$ networks of different sizes

does *not* spread. Second, we show that if $\Omega(f(n)\log n)$ initial seeds are infected, the whole graph gets infected with high probability in $O(\log n)$ rounds. This shows that the first few nodes in the arriving order of the network are critical in their roles of enabling a complex contagion.

10.1 No New Infections

The following theorem shows that if we select initial seeds randomly in a PA graph, then until the size of these initial seeds is a polynomial in the size of the graph, the contagion almost surely does not spread to any other node.

Theorem 7. *Let $0 < p \leq 1$ and consider the $\mathrm{PA}_{p,m}(n)$ graph. A k-complex contagion $\mathrm{CC}(\mathrm{PA}_{p,m}(n), k, S)$ would not spread to other nodes with probability $1 - o(1)$, if we choose S as follows.*

1) *If $k \geq 2/p$, $S = \{o\left(n^{1-p/2}\right)$ random initial seeds$\}$;*
2) *If $k < 2/p$, $S = \{o\left(n^{1-1/k}\right)$ random initial seeds$\}$.*

Proof: Assume that the network edges are undirected and let $s = |S|$. Let X be the number of infected nodes in the first round. X is the number of nodes that have at least k neighbors in S. We show that the expectation of X is $o(1)$. By Markov's inequality, the number of infected nodes will be zero with probability $1 - o(1)$. Let d_i and $\nu_i(S)$ denote the degree of the i-th node, and the number of neighbors of node i in set S respectively. Then $E[X]$ can be written as:

$$E[X] = \sum_{i=1}^{n} \text{Prob}\left\{\nu_i(S) \geq k\right\}$$
$$= \sum_{i=1}^{n}\sum_{x=k}^{mn} \text{Prob}\left\{\nu_i(S) \geq k | d_i = x\right\} \text{Prob}\left\{d_i = x\right\}.$$

In the proof of Lemma 10, we show that

$$\text{Prob}\left\{\nu_i(S) \geq k | d_i = x\right\} \leq \min\left\{\left(\frac{xs}{n}\right)^k \left(\frac{1}{1 - xs/n}\right), 1\right\}.$$

Let $W = \text{Prob}\left\{\nu_i(S) \geq k | d_i = x\right\}$ and $E[N_x]$ be the expected number of nodes of degree x in the PA graph of n vertices,

$$E[X] \leq \sum_{i=1}^{n}\sum_{x=k}^{mn} W \cdot \text{Prob}\left\{d_i = x\right\} \leq \sum_{x=k}^{mn} W \cdot E[N_x].$$

Thus, a critical step in the proof is to upper bound $E[N_x]$. We utilize the *master equation* method [33] to do this computation. However, instead of directly solving the recurrence as is done for the case of $p = 1$ for the *sequential* PA model in [34] and for the *conditioned* PA model in [35], we upper bound it for all values of $0 < p \leq 1$ in Lemmas. 8 and 9.

Let $N_t(x)$ denotes the number of nodes with degree x in the graph of t vertices and denote by $n_t(x) = E[N_t(x)]$. The following recurrence holds for the $\mathrm{PA}_{p,m}$ model:

$$E[N_{t+1}(x)|N_t(x)] = \left(1 - \frac{a_x}{t}\right) n_t(x) + \frac{a_{x-1}}{t} n_t(x-1) + c_x.$$

in which a_x and c_x are non-negative values that depend on the specific model and $a_{x+1} \geq a_x$.

In the $\mathrm{PA}_{p,m}$ model, each node issues m edges to existing nodes. With probability p, each edge connects to a node with preferential attachment rule and with probability $1-p$, an edge connects to a uniformly random chosen node.

$$a_x = \frac{px}{2} + m(1-p), \quad c_x = \delta_{(m+1)x} = \begin{cases} 1 & x = m+1 \\ 0 & x \neq m \end{cases}.$$

We ignore the possibility of more than one edge being attached to one vertex and the self-loops. The proofs for the following four lemmas are delayed until Subsection 10.3.

Lemma 8. *Let N_x be the number of nodes of degree x in the $\mathrm{PA}_{\mathrm{p,m}}(\mathrm{n})$ model. For all $0 < p \leq 1$, we have that $E[N_n(x)] \leq mn\eta_x$, where $\eta_x = \frac{a_{x-1}}{1+a_x}\eta_{x-1} + \frac{c_x}{1+a_x}$.*

Lemma 9. *In the preferential attachment $\mathrm{PA}_{\mathrm{p,m}}(\mathrm{n})$ model, we have $\eta_x = \Theta(x^{-(1+2/p)})$ for all $0 < p \leq 1$.*

Lemma 10. *Let S be chosen as stated in Theorem 7, X be the number of infected nodes in the first round of the complex contagion $\mathrm{CC}(\mathrm{PA}_{p,m}(n), k, S)$, and $0 < p \leq 1$. We have that,*

$$E[X] = O\left(\frac{s^k}{n^{k-1}} \sum_{x=k}^{n/2s} x^{k-1-2/p} + n \sum_{x=n/2s+1}^{mn} 1/x^{1+2/p}\right).$$

Lemma 11. *Let S be chosen as stated in Theorem 7, X be the number of infected nodes in the first round of the complex contagion $\mathrm{CC}(\mathrm{PA}_{p,m}(n), k, S)$, and $0 < p \leq 1$. We have that $E[X] = o(1)$.*

Applying Markov inequality on Lem. 11 proves Thm. 7. □

10.2 Oldies But Goodies

We use the expected degree of early nodes in the PA model to show that they become infected with high probability once enough random seeds are infected at round 0. Once all the first k nodes in the graph are infected, the k-complex contagion will spread to the rest of the graph quickly. This emphasizes the role of early nodes in the PA model.

A computation of the expected degree of nodes for $p = 1$ and $m = 1$ is presented in [36]. We follow their approach and prove the expected degree for all values of $0 < p \leq 1, m \geq 2$ in the following Lemma (proved in Subsection 10.3). We will work with the independent model here, but the other two variations are similar.

Lemma 12. *Let $d_t(s)$ denote the degree of node s in a $\mathrm{PA}_{p,m}$ at time t, and $0 < p \leq 1$. We have $E[d_n(s)] = \Theta\left((n/s)^{p/2}\right)$.*

Theorem 13. *Let $0 < p \leq 1$. If we choose $\mathcal{I}$ as $\Omega\left(n^{1-p/2}\log n\right)$ random initial seeds, then a k-complex contagion $\mathrm{CC}(\mathrm{PA}_{p,m}(n), k, \mathcal{I})$ will spread to all the nodes with high probability in $O(\log n)$ rounds.*

Proof: We focus on the first k arrived nodes in the PA graph. By Lem. 12, each of the first k nodes have expected degree of at least $m^2(1-p)\left(\frac{n}{k}\right)^{p/2}$. We focus on node v_k, the node that arrived at time k, from now on. If we infect $\Omega\left(n^{1-p/2}\log n\right)$ nodes, v_k would have $\Omega(\log n)$ infected neighbors in expectation in round 0. This would mean that w.h.p. v_k would have $\geq k$ infected neighbors in round 0. This means that all the first k nodes will be infected with high probability in round 1. Once the first k nodes are infected, Thm. 6 can be applied to show that the speed of contagion is $O(\log n)$ with high probability. □

10.3 Proofs of Computational Lemmas

Proof of Lemma 8: We prove the claim by induction on t, the number of nodes in the graph. In the base case $N_0(k) = 0$ for all x, so the claim is trivially true. Suppose that the claim is true for t, i.e., $n_t(x) \leq mt\eta_x$. And $\eta_{x-1} = (1+a_x)\eta_x/a_{x-1} - c_x/a_{x-1}$. By the recurrence we have

$$\begin{aligned} n_{t+1}(x) &\leq \left(1-\tfrac{a_x}{t}\right) mt\eta_x + \tfrac{a_{x-1}}{t} mt\eta_{x-1} + c_x \\ &\leq \left(1-\tfrac{a_x}{t}\right) mt\eta_x \\ &\quad + a_{x-1}m\left((1+a_x)\eta_x/a_{x-1} - c_x/a_{x-1}\right) + c_x \\ &\leq m(t+1)\eta_x. \end{aligned}$$

which proves the statement. □

Proof of Lemma 9: The statement for $p = 1$ is proved in [35]. We follow a similar strategy to prove it for all the values of $0 < p \leq 1$. From the recursive definition of η_x, we can write:

$$\eta_x = \sum_{j=1}^{x} \frac{c_j}{1+a_j} \prod_{i=j+1}^{x} \frac{a_{i-1}}{1+a_i}.$$

However, $c_j = 0$ for all $j > m+2$. Hence for $x \geq m+2$ we can write:

$$\begin{aligned} \eta_x &= \eta_{m+2} \prod_{j=m+3}^{x} \frac{a_{j-1}}{1+a_j} \\ &= \eta_{m+2} \prod_{j=m+3}^{x} \frac{p(j-1)/2 + m(1-p)}{1+pj/2+m(1-p)}. \end{aligned}$$

Define $\alpha_p = \frac{m-mp-p/2}{p/2}$ and $\beta_p = \frac{m-mp+1}{p/2}$ and notice that for $p < 1$, $-1 < \alpha_p < \beta_p$. We have:

$$\begin{aligned} \log(x) &= \log(\eta_{m+2}) + \\ &\sum_{j=m+3}^{x} \log\left(\frac{p(j-1)}{2} + m(1-p)\right) \\ &\quad - \sum_{j=m+3}^{x} \log\left(1 + \frac{pj}{2} + m(1-p)\right) \\ &= \log(\eta_{m+2}) + \sum_{j=m+3}^{x} \log\left(1+\frac{\alpha_p}{j}\right) - \log\left(1+\frac{\beta_p}{j}\right). \end{aligned}$$

$f(x) = \log(1+x)$ is a continuous function. So by the *mean value theorem* we have:

$$\forall j, \quad \exists \psi_j \quad \alpha_j/j < \psi_j < \beta_j/j, \quad f'(\psi_j) = \frac{f(\beta_j) - f(\alpha_j)}{\beta_j - \alpha_j}.$$

Hence we get:

$$\begin{aligned} \log(x) &= \log(\eta_{m+2}) + \sum_{j=m+3}^{x} \left(\frac{\beta_j - \alpha_j}{j}\right) \frac{1}{1+\psi_j} \\ &= \log(\eta_{m+2}) - \frac{2+p}{p} \sum_{j=m+3}^{x} \frac{1}{j(1+\psi_j)}. \end{aligned}$$

Furthermore, we have that

$$\sum_{j=m+3}^{x} \frac{1}{j+\beta_p} \leq \sum_{j=m+3}^{x} \frac{1}{j(1+\psi_j)} \leq \sum_{j=m+3}^{x} \frac{1}{j+\alpha_p};$$

which means that $\eta_x = \Theta\left(x^{-(1+2/p)}\right)$. □

Proof of Lemma 10: Let d_i and $\nu_i(S)$ denote the degree of the i-th node, and the number of neighbors of node i in set S respectively. We have:

$$\begin{aligned} E[X] &= \sum_{i=1}^{n} \text{Prob}\left\{\nu_i(S) \geq k\right\} \\ &= \sum_{i=1}^{n} \sum_{x=k}^{mn} \text{Prob}\left\{\nu_i(S) \geq k | d_i = x\right\} \text{Prob}\left\{d_i = x\right\}. \end{aligned}$$

We can rewrite $\text{Prob}\left\{\nu_i(S) \geq k | d_i = x\right\}$ as:

$$\begin{aligned} &\text{Prob}\left\{\nu_i(S) \geq k | d_i = x\right\} \\ &= \min\left\{\sum_{j=k}^{x} \text{Prob}\left\{\nu_i(S) = j | d_i = x\right\}, 1\right\} \\ &\leq \min\left\{\sum_{j=k}^{x} x^j \left(\frac{s}{n}\right)^j, 1\right\} \\ &\leq \min\left\{\left(\frac{xs}{n}\right)^k \left(\frac{1}{1-xs/n}\right), 1\right\} \qquad \text{if } \frac{xs}{n} < 1. \end{aligned}$$

We claim that if $\frac{xs}{n} < 1/2$, then $\left(\frac{xs}{n}\right)^k \left(\frac{1}{1-xs/n}\right) < 1$ because $\left(\frac{1}{1-xs/n}\right) < 2$, and $k \geq 2$. We upper bound $E[X]$:

$$\begin{aligned} E[X] &= \sum_{i=1}^{n} \sum_{x=k}^{mn} \text{Prob}\left\{\nu_i(S) \geq k | d_i = x\right\} \text{Prob}\left\{d_i = x\right\} \\ &\leq \sum_{x=k}^{mn} \min\left\{\left(\frac{xs}{n}\right)^k \left(\frac{1}{1-xs/n}\right), 1\right\} \sum_{i=1}^{n} \text{Prob}\left\{d_i = x\right\} \\ &\leq \sum_{x=k}^{mn} \min\left\{\left(\frac{xs}{n}\right)^k \left(\frac{1}{1-xs/n}\right), 1\right\} E[N_x]. \end{aligned}$$

Now we cut off the summation at $xs/n = 1/2$. Although this cut-off is not sharp, since we are bounding the expectation from above it is ok.

$$\begin{aligned} E[X] &\leq \sum_{x=k}^{mn} \min\left\{\left(\frac{xs}{n}\right)^k \left(\frac{1}{1-xs/n}\right), 1\right\} E[N_x] \\ &\leq \sum_{x=k}^{n/2s} \left(\frac{xs}{n}\right)^k \left(\frac{1}{1-xs/n}\right) E[N_x] + \sum_{x=n/2s+1}^{mn} E[N_x] \\ &\leq \sum_{x=k}^{n/2s} 2\left(\frac{xs}{n}\right)^k E[N_x] \\ &\quad + \sum_{x=n/2s+1}^{mn} E[N_x] \quad \text{since } \left(\frac{1}{1-xs/n}\right) < 2. \end{aligned}$$

We now can use Lemmas 8, 9:

$$\begin{aligned}E[X] &\le \sum_{x=k}^{n/2s} 2\left(\frac{xs}{n}\right)^k mn\eta_x + \sum_{x=n/2s+1}^{mn} mn\eta_x \quad \text{(Lemma 8)},\\ &\le \sum_{x=k}^{n/2s} 2\left(\frac{xs}{n}\right)^k mn\,\Theta\left(x^{-(1+2/p)}\right)\\ &\quad + \sum_{x=n/2s+1}^{mn} mn\,\Theta\left(x^{-(1+2/p)}\right) \qquad \text{(Lemma 9)},\\ &= O\left(\frac{s^k}{n^{k-1}}\sum_{x=k}^{n/2s} x^{k-1-2/p} + n\sum_{x=n/2s+1}^{mn} x^{-1-2/p}\right).\end{aligned}$$

Proof of Lemma 11: We do case analysis on $E[X]$ based on Lem. 10:

Case 1: $k > 2/p$**.**

$E[X] = O\left(\frac{s^k}{n^{k-1}}\left(\frac{n}{2s}\right)^{k-2/p} + n\left(\frac{n}{2s}+1\right)^{-2/p}\right)$,

which solves to $E[X] = O\left(\frac{s^{2/p}}{n^{2/p-1}}\right)$. If $s = o\left(n^{1-p/2}\right)$, we get $E[X] = o(1)$.

Case 2: $k = 2/p$**.**

$E[X] = O\left(\frac{s^k}{n^{k-1}}\log(n/2s) + n\left(\frac{n}{2s}+1\right)^{-2/p}\right)$,

which solves to $E[X] = O\left(\frac{s^{2/p}\log(n/2s)}{n^{2/p-1}}\right)$. If $s = o\left(n^{1-p/2}(\log n)^{-p/2}\right)$, we get $E[X] = o(1)$. However if $n^{1-p/2}(\log n)^{-p/2} < s = o\left(n^{1-p/2}\right)$, the $\log(n/2s)$ term in $E[X]$ would be $O(1)$ and $E[X] = o(1)$.

Case 3: $k < 2/p$**.**

We have $E[X] = O\left(\frac{s^k}{n^{k-1}} + \frac{s^{2/p}}{n^{2/p-1}}\right)$ that solves to $E[X] = O\left(\frac{s^k}{n^{k-1}}\right)$. If $s = o\left(n^{1-1/k}\right)$, we get that $E[X] = o(1)$ again. □

Proof of Lemma 12: We write a recursive relation based on the edge probabilities.

$$E[d_t(s)|d_{t-1}(s)] = d_{t-1}(s) + pm\frac{d_{t-1}(s)}{2m(t-1)} + (1-p)m\frac{1}{t-1}$$

$$E[d_t(s)] = \frac{2t-2+p}{2t-2}E[d_{t-1}(s)] + \frac{m(1-p)}{t-1}.$$

Starting with $d_s(s) = m$, we get:

$$\begin{aligned}E[d_n(s)] &= \sum_{j=s}^{n}\frac{m(1-p)}{j-1}\prod_{i=j+1}^{n}\frac{t-1+p/2}{t-1}\\ &= \sum_{j=s}^{n}\frac{m(1-p)}{j-1}d_s(s)\prod_{i=j+1}^{n}\frac{t-1+p/2}{t-1}\\ &= \sum_{j=s}^{n}\frac{m^2(1-p)}{j-1}\frac{\Gamma(n-1+p/2)}{\Gamma(n-1)}\frac{\Gamma(j)}{\Gamma(j+p/2)}\\ &= m^2(1-p)\sum_{j=s}^{n}\frac{1}{j-1}\left(\frac{n}{j}\right)^{p/2}\left(1+O\left(\frac{1}{j}\right)\right)\\ &\qquad \text{using Stirling's formula for } \Gamma(.);\\ &= \Theta\left(m^2(1-p)n^{p/2}\sum_{j=s}^{n}\frac{1}{j^{p/2}(j-1)}\right)\\ &= \Theta\left(m^2(1-p)\left(\frac{n}{s}\right)^{p/2}\right). \quad \square\end{aligned}$$

11 Complexity of Computing The Extent of Complex Contagions

We prove that it is $\mathbf{P}$-complete to decide if a k-complex contagion completely infects a graph or stops at a small fraction of its nodes.

Theorem 14. *For any integer $k \ge 2$, given a triple (G, S, M) where G is an undirected graph, S is a subset of vertices, and M is an integer, it is $\mathbf{P}$-complete to determine if the size of the resulting k-complex contagion on G when the vertices of S are initially infected is at least M. Let n be the number of vertices in G. In fact for any $0 < \epsilon < 1$, the promise problem of deciding if the size of the k-complex contagion is n or at most n^ϵ, is promise $\mathbf{P}$-complete.*

The reduction comes from the ***MonotoneCircuitValue*** problem in circuit complexity.

Definition 13. *In the **MonotoneCircuitValue** (MCV) problem we are given a circuit C with `0`, `1`, `AND`, and `OR` gates and one gate g_* designated as output. We insist that C is layered, that is we can partition the gates into levels $\{0, 1\} = L_0, L_1, \ldots, L_{\ell-1}, L_\ell = \{g_*\}$ such that wires always connect gates at levels i and $i+1$ for some $0 \le i \le \ell - 1$. $C \in$ MCV if the circuit is a properly encoded, layered, monotone circuit and evaluates to `1`. Otherwise $C \notin$ MCV.*

Theorem 15 ([37])**.** *The MCV problem is $\mathbf{P}$-complete.*

Proof of Theorem 14: These problems are in $\mathbf{P}$ or promise-$\mathbf{P}$ because an algorithm can simply simulate the contagion and count the number of infected nodes. To show the hardness result the idea is to reduce from MCV. Given such a circuit C we create a graph as follows:

Fix ϵ, k. Given a circuit C with m gates we create the triple (G, S, M) as follows: Let $M = (3k^3m)^{1/\epsilon}$. We next create the vertices of G:

- For each gate g_a of C , we create k vertices $G_a = \{g_a^i\}_{0\le i<k}$.
- For each wire w_{ab} of C connecting gate g_a to gate g_b, create k^2 vertices $W_{ab} = \{w_{ab}^{i,j}\}_{0\le i,j<k}$.
- Create M additional vertices $T = \{t_i\}_{0\le i<M}$.

Next, we create the edges:

- Consider a non constant gate g_c of C with input gates g_a and g_b (assume an arbitrary ordering over the input gates).
  - Add the k^3 edges to connect all vertices in G_x to all vertices in W_{xc} for $x \in \{a, b\}$.
  - If g_c in an `OR` gate, connect $w_{x,c}^{i,j}$ to g_c^i for $0 \le i, j < k$ for $x \in \{a, b\}$.
  - If g_c in an `AND` gate, connect $w_{a,c}^{i,j}$ to g_c^i for $0 \le i < k$ and $0 \le j < \lceil k/2 \rceil$.
  - If g_c in an `AND` gate, connect $w_{b,c}^{i,j}$ to g_c^i for $0 \le i < k$ and $0 \le j < \lfloor k/2 \rfloor$.
- Add the k^2 edges between G_* and t_i for $0 \le i < M$.
- For all the vertices $v \in G \setminus T$, add k edges between v and k vertices of T. But, each vertex of M can

only be used once. Let $R = 3k^2m$. Because every gate has at most 2 in-wires, and each gate/wire has at most k^2 corresponding nodes, R is an upper bound on the number of vertices not in T. Therefore, $M = (3k^3m)^{1/\epsilon} > 3k^3m = kR$ is big enough to satisfy the use-once constraint on the vertices of T.

Let $S = G_1$, the vertices corresponding to the constant `1` gate. It is easy to verify that (G, S, M) can be constructed in logspace[8]. Now, we will show that T is infected if and only if C evaluates to 1. The proof will follow from the following lemma:

Lemma 16. *Consider a $0 \leq t \leq 2\ell$. If t is even, we claim that the only newly infected nodes at time t correspond to gates at level $t/2$ in C which evaluate to `1`. We claim that for an odd t, the only newly infected nodes at time t correspond to the wires w_{ab} connecting gates at level $(t-1)/2$ and $(t+1)/2$ where the gate at level $(t-1)/2$ evaluates to `1`.*

Using the lemma, at time 2ℓ the only nodes that can possibly become infected are those corresponding to the output gate. If they do become infected, then at time $2t+1$ all the nodes of T will become infected. Ultimately, at time $2t+2$ all the graph will become infected.

Notice that each node in T only has one edge outside the nodes of output gate G_*. Therefore, if at time 2ℓ the output gate does not become infected, then at that step no additional nodes become infected and the contagion is over. We prove the lemma first:

Proof: The proof proceeds by induction. At time $t = 0$ this is true, because the only nodes at level 0 are constant gates, and the only constant gates that evaluate to `1` is the `1` gate. By construction $G_1 = S$ and so these vertices are initially infected at time $t = 0$. Assuming that the statement is true up to time $t < 2\ell$, we show that the statement is true at time $t+1$.

The case where t is even: At time $t+1$ any node that becomes infected must be connected to a node that was infected at time t. By the inductive hypothesis, the only nodes that become infected at time t are those that correspond to gates at level $t/2$. By construction, these nodes are connected to nodes corresponding to wires connecting gates at level $t/2-1$ and level $t/2$ as well as nodes corresponding to wires connecting gates at level $t/2$ and level $t/2+1$.

The nodes W_{ab} that correspond to wires w_{ab} connecting a gate g_a at level $t/2-1$ and a gate g_b level $t/2$ are, by construction, attached to the nodes G_a and the nodes G_b. The nodes of W_{ab} can only be infected at time $t+1$ if they were not already infected at time t. By the inductive hypothesis, the nodes of W_{ab} are not infected at time t if and only if g_a evaluates to `0` in which case, again by the inductive hypothesis, the nodes of G_a are also not infected at time t. However, if the nodes corresponding to G_a are not infected at time t, then the nodes in W_{ab} will not be infected at time $t+1$ as, by construction, each node in W_{ab} has only one neighbor outside of G_a and $k \geq 2$.

The case where t is odd: By the inductive hypothesis, the only nodes that become infected at time t are those that correspond to wires that connect nodes in level $(t-1)/2$ and level $(t+1)/2$. By construction, these nodes are connected to nodes corresponding to gates at level $(t-1)/2$ and level $(t+1)/2$. By the inductive hypothesis, all the neighbors that these newly infected nodes' wires connect to at level $(t-1)/2$ are already infected. Let's consider then the nodes corresponding to gates at level $(t+1)/2$.

If g_c is an `OR` gate with inputs g_a and g_b, then, by construction, each node in G_c is attached to k nodes in W_{ac} and k nodes in W_{bc}. Thus, if either g_a or g_b evaluate to 1, then, by the inductive hypothesis, either the nodes in W_{ac} or the nodes in W_{bc} will be infected at time t and thus at time $t+1$ the nodes in G_c will become infected. On the other hand, if neither g_a or g_b evaluate to `1`, then, by the inductive hypothesis, neither the nodes in W_{ac} or the nodes in W_{bc} will be infected at time t. By construction, any other neighbors of nodes in G_c correspond to wires connecting gates at level $(t-1)/2$ and $(t+1)/2$. By the inductive hypothesis, these gates are not infected at time t. Thus, the nodes of G_c will not be infected at time $t+1$.

If g_c is an `AND` gate with inputs g_a and g_b, then, by construction, each node in G_c is attached to $\lceil k/2 \rceil$ nodes in W_{ac} and $\lfloor k/2 \rfloor$ nodes in W_{bc}. Thus, if both g_a or g_b evaluate to 1, then, by the inductive hypothesis, the nodes in W_{ac} and the nodes in W_{bc} will be infected at time t and thus at time $t+1$ the nodes in G_c will become infected. On the other hand, if either g_a or g_b evaluate to `0`, then, by the inductive hypothesis, either the nodes in W_{ac} or the nodes in W_{bc} will be not infected at time t. By construction, any other neighbors of nodes in G_c correspond to wires connecting gates at level $(t-1)/2$ and $(t+1)/2$. By the inductive hypothesis, these gates are not infected at time t. Thus, the nodes of G_c will not be infected at time $t+1$. □

The reduction is complete because:

- If $C \in \mathsf{MCV}$, then T becomes infected and at least M nodes (and all the nodes in the graph) are infected.
- If $C \notin \mathsf{MCV}$, then T does not become infected. Remember that $R = 3k^2m$ is an upper bound on the number of vertices not in T and thus an upper bound on the number of nodes that become infected. But $R = 3k^2m < M^\epsilon = 3k^3m < n^\epsilon$. Thus, fewer than n^ϵ nodes are infected.

8. Note that multiplication, powering, and division are known to be in logspace [38]. However, these results are not needed if we simply compute a number $M > (3k^3m)^{1/\epsilon}$

Acknowledgment

G. Ghasemiesfeh, R. Ebrahimi and J. Gao would like to acknowledge support through NSF DMS-1418255, CCF-1535900, CNS-1618391 and AFOSR FA9550-14-1-0193. G. Schoenebeck would like to acknowledge support through NSF CCF-1535912, CAREER CCF-1452915, a Facebook Faculty Award and a Google Faculty Award.

References

[1] D. Centola and M. Macy, "Complex Contagions and the Weakness of Long Ties," *American Journal of Sociology*, vol. 113, no. 3, pp. 702–734, 2007.

[2] J. S. Coleman, E. Katz, and H. Menzel, *Medical Innovation: A Diffusion Study*. Bobbs-Merrill Co, 1966.

[3] D. Centola, "The spread of behavior in an online social network experiment," *Science*, vol. 329, no. 5996, p. 1194, 2010.

[4] J. Ugander, L. Backstrom, C. Marlow, and J. Kleinberg, "Structural diversity in social contagion," *Proc. National Academy of Sciences*, vol. 109, no. 16, pp. 5962–5966, 2012.

[5] D. M. Romero, B. Meeder, and J. Kleinberg, "Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on Twitter," in *Proc. of the 20th conference on World Wide Web*, 2011, pp. 695–704.

[6] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, 2010.

[7] G. Ghasemiesfeh, R. Ebrahimi, and J. Gao, "Complex contagion and the weakness of long ties in social networks: revisited," in *Proc. of the 14th ACM conference on Electronic Commerce*, 2013, pp. 507–524.

[8] R. Ebrahimi, J. Gao, G. Ghasemiesfeh, and G. Schoenebeck, "Complex contagions in Kleinberg's small world model," in *Proc. of the 2015 Conference on Innovations in Theoretical Computer Science*, pp. 63–72.

[9] D. D. S. Price, "A general theory of bibliometric and other cumulative advantage processes," *Journal of the American Society for Information Science*, pp. 292–306, 1976.

[10] A. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, 1999.

[11] P. Domingos and M. Richardson, "Mining the network value of customers," in *Proc. of the 7th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2001, pp. 57–66.

[12] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, pp. 47–97, 2002.

[13] N. Berger, C. Borgs, J. T. Chayes, and A. Saberi, "Asymptotic behavior and distributional limits of preferential attachment graphs," *Annals of Applied Probability*, 2010.

[14] S. Dommers, R. van der Hofstad, and G. Hooghiemstra, "Diameters in preferential attachment models," *Journal of Statistical Physics*, vol. 139, pp. 72–107, 2010.

[15] N. Berger, C. Borgs, J. T. Chayes, and A. Saberi, "On the spread of viruses on the Internet," in *Proc. of the 16th ACM-SIAM symposium on Discrete algorithms*, 2005, pp. 301–310.

[16] B. Bollobas, "Mathematical results on scale-free random graphs," in *Handbook of Graphs and Networks*. Wiley, 2003, pp. 1–37.

[17] M. Granovetter, "Threshold models of collective behavior," *The American Journal of Sociology*, vol. 83, no. 6, pp. 1420–1443, 1978.

[18] S. Morris, "Contagion," *Review of Economic Studies*, vol. 67, pp. 57–78, 2000.

[19] A. Montanari and A. Saberi, "Convergence to equilibrium in local interaction games," *SIGecom Exch.*, vol. 8, no. 1, pp. 11:1–11:4, 2009.

[20] B. Doerr, M. Fouz, and T. Friedrich, "Social networks spread rumors in sublogarithmic time," in *Proc. of the 43rd ACM symposium on Theory of computing*, 2011, pp. 21–30.

[21] F. Chierichetti, S. Lattanzi, and A. Panconesi, "Rumor spreading in social networks," *Theoretical Computer Science*, vol. 412, no. 24, pp. 2602 – 2610, 2011.

[22] E. Bakshy, B. Karrer, and L. A. Adamic, "Social influence and the diffusion of user-created content," in *Proc. of the 10th ACM Conference on Electronic Commerce*, 2009, pp. 325–334.

[23] E. Bakshy, J. M. Hofman, W. A. Mason, and D. J. Watts, "Everyone's an influencer: quantifying influence on twitter," in *Proc. of the 4th ACM conference on Web Search and Data Mining*, 2011, pp. 65–74.

[24] J. Chalupa, P. L. Leath, and G. R. Reich, "Bootstrap percolation on a bethe lattice," *Journal of Physics C: Solid State Physics*, vol. 12, no. 1, p. L31, 1979.

[25] J. Adler, "Bootstrap percolation," *Physica A: Statistical and Theoretical Physics*, vol. 171, no. 3, pp. 453–470, 1991.

[26] S. Janson, T. Luczak, T. Turova, and T. Vallier, "Bootstrap percolation on the random graph $G_{n,p}$," *Annals of Applied Probability*, vol. 22, no. 5, pp. 1989–2047, 2012.

[27] J. Balogh and B. Pittel, "Bootstrap percolation on the random regular graph," *Random Structures and Algorithms*, vol. 30, pp. 257–286, 2007.

[28] H. Amini, "Bootstrap percolation and diffusion in random graphs with given vertex degrees." *Electr. J. Comb.*, vol. 17, no. 1, 2010.

[29] H. Amini and N. Fountoulakis, "What I tell you three times is true: bootstrap percolation in small worlds," in *Proc. of the 8th conference on Internet and Network Economics*, 2012, pp. 462–474.

[30] A. Coja-Oghlan, U. Feige, M. Krivelevich, and D. Reichman, "Contagious sets in expanders," in *Proc. of the 26th ACM-SIAM Symposium on Discrete Algorithms*, 2015, pp. 1953–1987.

[31] B. Bollobás and O. Riordan, "The diameter of a scale-free random graph," *Combinatorica*, vol. 24, pp. 5–34, 2004.

[32] T. E. Harris, *The theory of branching processes*, ser. Die Grundlehren der Mathematischen Wissenschaften, Bd. 119. Berlin: Springer-Verlag, 1963.

[33] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, "Structure of growing networks with preferential linking." *Phys. Rev. Lett.*, vol. 85, pp. 4633–4636, 2000.

[34] R. Durrett, *Random Graph Dynamics (Cambridge Series in Statistical and Probabilistic Mathematics)*. New York, NY, USA: Cambridge University Press, 2006.

[35] O. Hagberg and C. Wiuf, "Convergence properties of the degree distribution of some growing network models," *Bulletin of Mathematical Biology*, vol. 68, no. 6, pp. 1275–1291, 2006.

[36] B. Bollobás, O. Riordan, J. Spencer, and G. E. Tusnady, "The degree sequence of a scale-free random graph process." *Random Struct. Algorithms*, vol. 18, no. 3, pp. 279–290, 2001.

[37] L. M. Goldschlager, "The monotone and planar circuit value problems are log space complete for p," *SIGACT News*, vol. 9, no. 2, pp. 25–29, 1977.

[38] W. Hesse, E. Allender, and D. A. M. Barrington, "Uniform constant-depth threshold circuits for division and iterated multiplication," *Journal of Computer and System Sciences*, vol. 65, no. 4, pp. 695 – 716, 2002.

Roozbeh Ebrahimi is a software engineer at Google Inc. He finished his PhD in Computer Science at Stony Brook University. He received his B.Sc. in Mathematics from Sharif University of Technology.

Jie Gao is an associate professor of Computer Science at Stony Brook University. She received BS from the special class for the gifted young program at University of Science and Technology of China in 1999 and Ph.D in computer science from Computer Science department, Stanford University in 2004. She received NSF CAREER Award in 2006, and multiple awards at Stony Brook CS department on Excellence in Research. She is currently serving on the editorial board of ACM Transactions on Sensor Networks and Journal of Discrete Algorithms.

Golnaz Ghasemiesfeh is pursuing a PhD in Computer Science at Stony Brook University. She received her M.Sc. in Computer Science from University of Toronto and her B.Sc. in Computer Science from Sharif University of Technology.

Grant Schoenbeck is an assistant professor at the University of Michigan in the Computer Science and Engineering division. His research is supported by the NSF, Facebook, and Google including an NSF CAREER award. Before coming to the University of Michigan in 2012, he was a Postdoctoral Research Fellow at Princeton. Grant received his PhD at UC Berkeley, studied theology at Oxford University, and received his BA in mathematics and computer science from Harvard.