From Market Saturation to Social Reinforcement: Understanding the Impact of Non-Linearity in Information Diffusion Models*

Tobias Friedrich
Hasso Plattner Institute,
University of Potsdam
Potsdam, Germany
tobias.friedrich@hpi.de

Andreas Göbel
Hasso Plattner Institute,
University of Potsdam
Potsdam, Germany
andreas.goebel@hpi.de

Nicolas Klodt
Hasso Plattner Institute,
University of Potsdam
Potsdam, Germany
nicolas.klodt@hpi.de

Martin S. Krejca
Laboratoire d'Informatique (LIX),
CNRS, École Polytechnique, Institut
Polytechnique de Paris
Paris, France
martin.krejca@polytechnique.edu

Marcus Pappik
Hasso Plattner Institute,
University of Potsdam
Potsdam, Germany
marcus.pappik@hpi.de

ABSTRACT

Diffusion of information in networks is at the core of many problems in AI. Common examples include the spread of ideas and rumors as well as marketing campaigns. Typically, information diffuses at a non-linear rate, for example, if markets become saturated or if users of social networks reinforce each other's opinions. Despite these characteristics, this area has seen little research, compared to the vast amount of results for linear models, which exhibit less complex dynamics. Especially, when considering the possibility of re-infection, no fully rigorous guarantees exist so far.

We address this shortcoming by studying a very general non-linear diffusion model that captures saturation as well as reinforcement. More precisely, we consider a variant of the SIS model in which vertices get infected at a rate that scales polynomially in the number of their infected neighbors, weighted by an infection coefficient λ. We give the first fully rigorous results for thresholds of λ at which the expected survival time becomes super-polynomial. For cliques we show that when the infection rate scales sub-linearly, the threshold only shifts by a poly-logarithmic factor, compared to the standard SIS model. In contrast, super-linear scaling changes the process considerably and shifts the threshold by a polynomial term. For stars, sub-linear and super-linear scaling behave similar and both shift the threshold by a polynomial factor. Our bounds are almost tight, as they are only apart by at most a poly-logarithmic factor from the lower thresholds, at which the expected survival time is logarithmic.

KEYWORDS

information diffusion; epidemic models; non-linear infection; survival time

*for the full version see [6]

Proc.of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand. © 2024 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

ACM Reference Format:
Tobias Friedrich, Andreas Göbel, Nicolas Klodt, Martin S. Krejca, and Marcus Pappik. 2024. From Market Saturation to Social Reinforcement: Understanding the Impact of Non-Linearity in Information Diffusion Models. In *Proc.of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024*, IFAAMAS, 9 pages.

1 INTRODUCTION

Information diffusion processes on graphs are widely studied in the area of AI [11, 12, 15, 25, 28–30] and in other domains, modeling various graph processes, such as spread of infections [16, 24] and computer viruses [2, 3], social influence and the spread of ideas [14], and viral marketing campaigns [1].

Commonly, information diffusion processes are modeled as epidemiological models over networks (see [24] for an extensive survey). In these models, each vertex of the host network is in a state, such as *susceptible* or *infected*, and transitions between these states at variable rates that depend on the states of all vertices in the network. A very prominently studied epidemiological model is the *SIS model*. In this model, each susceptible vertex becomes infected by each of its infected neighbors independently with a system-wide infection rate of $\lambda \in \mathbb{R}_{>0}$, and each infected vertex turns susceptible independently with a normalized rate of 1.

The SIS-model includes the possibility for vertices to re-infect after recovering from the infection. That accounts, for example, for infections that do not grant immunity [23] or for bloggers that can post the same message multiple times [17]. Therefore, it is possible for the infection to stay active for a very long time by infecting the same vertices over and over again. The quantity that measures how long the process contains infected vertices is known as the *survival time* and marks an important property for networks in the SIS model.

Due to its relevance, the survival time of the standard SIS model has been studied extensively for decades both empirically [24] as well as mathematically rigorously, for the latter on infinite [e.g., 8, 18, 22] and on finite graphs [e.g., 2, 3, 7]. Combined, these results show for a large variety of different finite graph classes a sharp transition, with respect to the infection rate λ, from a survival time

Table 1: Our threshold results for the infection coefficient λ in the modified SIS process with infection rate $\lambda I^{1+\alpha}$, where I is the number of infected neighbors of a vertex. The table gives regimes for λ depending on the host graph structure, the number of vertices n and the infection exponent α, in which the expected survival time is logarithmic or super-polynomial in n respectively. The case $\alpha = 0$ corresponds to the known results for the standard SIS model.

	$T \in O(\log(n))$	$T \geq 2^{n^{\Omega(1)}}$
Clique		
$\alpha < 0$	$\lambda \in O(n^{-1})$	$\lambda \in \omega(n^{-1}\log(n)^{-\alpha})$
$\alpha > 0$	$\lambda \in o(n^{-1-\alpha})$	$\lambda \in \omega(n^{-1-\alpha}\log(n)^{\alpha})$
$\alpha = 0$	$\lambda \leq n^{-1}$ [7]	$\lambda \geq (1+\varepsilon)n^{-1}$ [7]
Star		
	$\lambda \in O\left(n^{-\frac{1}{2}-\frac{\alpha}{2(2+\alpha)}}\right)$	$\lambda \in \omega\left(n^{-\frac{1}{2}-\frac{\alpha}{2(2+\alpha)}}\log(n)^{\frac{4}{1+\alpha}}\right)$

logarithmic in the graph size to one that is super-polynomial. The regime of λ where this change occurs is known as the *epidemic threshold*. This regime is mostly independent of the starting configuration of the process as long as at least one vertex starts infected.

In the majority of the epidemiology models studied, vertices get infected at a rate that scales linearly with the number of their infected neighbors. However, experiments have shown that there are processes that cannot be modeled with such simple assumptions [10, 21]. For example, when behavior is spread over social networks, there is a social reinforcement effect that leads to much higher adoption rates when the number of infected neighbors is high [4]. Similar effects can be observed for biological contagions [27]. In some cases the opposite effect happens, for example, during market saturation, when trying to convince customers to buy a product [15].

As non-linear infection rates are prominent in natural processes, it is important to study how they alter the insights that have been gained for linear processes. To this end, we consider an altered version of the standard SIS model, parameterized by an *infection exponent* $\alpha \in \mathbb{R}_{>-1}$. In this version, susceptible vertices do not get infected by their neighbors independently but, instead, a susceptible vertex with i infected neighbors is infected with a rate of $\lambda i^{1+\alpha}$. Note that we call λ the *infection coefficient*, and for $\alpha = 0$, it coincides with the infection rate of the standard SIS model.

This model as well as generalizations and variants thereof have already been theoretically studied [9, 19, 26, 31]. These works use mean-field theory, simplifying the original process by making approximation assumptions. In these works the host graph is assumed to be a clique in order to derive differential equations that model the dynamics of the process. This results in a threshold for the infection coefficient at which the simplified process changes from having the all-susceptible state as unique global equilibrium to having two equilibria from which one is not the all-susceptible state and is globally stable. This threshold gives an estimate for the epidemic threshold on cliques. To the best of our knowledge, no fully rigorous results on the epidemic threshold exist so far.

Our Contribution. We analyze the epidemic threshold of a SIS variant with non-linear infection rates in a rigorous mathematical manner. To the best of our knowledge, these are the first results that study the process in the non-linear setting without any simplifications. We prove both upper and lower bounds for the epidemic threshold for cliques (Corollaries 7 and 10, assuming one initially infected vertex) and for stars (Corollary 16, assuming the center to be initially infected). Further, our results encompass settings with sub-linear as well as super-linear infection rates in case of a constant infection exponent α. In all cases, our upper and lower bounds of the epidemic threshold are almost tight, i.e. different by at most poly-logarithmic factors. Our results are summarized in Table 1. Note that in this setting the survival time increases monotonically when adding extra vertices and edges, hence our lower bounds carry over to graphs with large clique or star subgraphs.

For cliques of size n, we see a clear difference between the sub-linear and the super-linear setting. For sub-linear infection rates (Corollary 7, $\alpha \in (-1, 0)$), the epidemic threshold remains similar to the linear setting [7, Section V.C] and increases from $1/n$ to $\omega(\log(n)^{-\alpha}/n)$.

For the super-linear setting for cliques of size n (see Corollary 10), the infection exponent α of the model also has an impact on the power of n. The infection survives already, in expectation, a time exponential in n once $\lambda \in \omega(n^{-1-\alpha}\log(n)^{\alpha})$, decreasing the threshold from the linear setting by a factor in the order of $\log(n)/n^{\alpha}$. However, the survival time has a very high variance, in the sense that the process survives exponentially long with a probability at most super-polynomial in $-n$ (see Lemma 8). This long survival time occurs once the process hits a critical mass of infected vertices. In contrast, the process in the linear model survives exponentially long, already with a probability linear in n^{-1}.

For stars with n leaves, we provide a unified bound for the sub- and super-linear setting (see Corollary 16). In both cases, the epidemic threshold deviates roughly by a factor of $n^{-\alpha/(2(2+\alpha))}$ from the linear threshold of $n^{-1/2}$ [7, Theorems 5.1 and 5.2]. This shows that the effect of changing the infection rate is also well pronounced when deviating from the linear setting.

Overall, on stars, changing the infection rate has a strong impact on the epidemic threshold in all settings. On cliques, the impact is strong when the infection rate scales super-linearly, but the effect is far less prominent in a sub-linear scaling.

2 PRELIMINARIES

We consider a variation of the SIS model in which the rate at which vertices get infected scales non-linearly in the number of its infected neighbors. The process is defined as follows.

Let $G = (V, E)$ be a finite, undirected graph with vertex set V and edge set E. Further let $\lambda \in \mathbb{R}_{>0}$ and $\alpha \in \mathbb{R}_{>-1}$. A contact process C with infection coefficient λ and infection exponent α on G is a continuous-time Markov process over partitions of V into susceptible vertices and infected vertices. The transition of the process is decided by Poisson point processes on the vertices, which we call Poisson clocks. Each infected vertex has a Poisson clock with rate 1 which when it triggers *heals* it and moves it to the susceptible set. Each susceptible vertex has a Poisson clock with a variable rate. For every vertex $v \in V$ and time $t \in \mathbb{R}_{\geq 0}$, let $N_{t,v}$ be the number of infected neighbors of v. Each susceptible vertex v has a Poisson clock that infects it at rate $\lambda N_{t,v}^{1+\alpha}$. Note that we restrict α to be larger than -1 as otherwise the clock rate is either not defined or

strictly positive for susceptible vertices with no infected neighbors. We also assume α to be constant in the number of vertices.

We aim to calculate the *survival time* of the contact process, the first point in time at which the process reaches the only absorbing state, which is the the state in which all vertices are susceptible.

In our proofs we sometimes consider the discrete version of that process in which one step is exactly one trigger of a clock. That means that in one step exactly one vertex heals or one vertex gets infected. We use the fact that while at least one vertex is infected all clocks together trigger at a rate of at least 1 and at most polynomial to transfer the bounds from the discrete version to the continuous process.

We use stochastic domination to transfer results from one random variable to another. We say that a random variable $(X_t)_{t\in\mathbb{R}}$ *dominates* another random variable $(Y_t)_{t\in\mathbb{R}}$ if and only if there exists a coupling $(X'_t, Y'_t)_{t\in\mathbb{R}}$ in a way such that for all $t \in \mathbb{R}_{\geq 0}$ we have $X'_t \geq Y'_t$. For example for two contact processes C and C' that only differ in the fact that C has a higher infection coefficient, the number of infected vertices in C dominates the number of infected vertices in C' as the processes can be coupled in a way such that all healing clocks trigger at the same time and infections in C' imply infections in C at the same time. The domination then directly implies that the survival time of C dominates the survival time of C', so the survival time in our model increases monotonically with the infection coefficient.

When we say that some event happens asymptotically almost surely (a.a.s.) that means that for increasing number of vertices in the considered graph the event happens with a probability of $1 - o(1)$.

Some of the processes that we analyze are very similar to the well-known gambler's ruin problem, as they increase and decrease by one with certain probabilities until they reach a limit in either direction. We consider the following version of the gambler's ruin problem.

Theorem 1 (Gambler's ruin [5, page 345]). *Let $(P_t)_{t\in\mathbb{N}}$ be the amount of money that a player has in a gambler's ruin game that has a probability of $p \neq 1/2$ for them to win in each step. Let $q = 1 - p$. The game ends at time T when the player either reaches the lower bound l or the upper bound u of money. Then*

(1) $\Pr[P_T = l] = \frac{1-(p/q)^{u-P_0}}{1-(p/q)^{u-l}}$;

(2) $\Pr[P_T = u] = \frac{1-(q/p)^{P_0-l}}{1-(q/p)^{u-l}}$.

Wald's equation helps us calculate the survival time by partitioning the process into phases and then bounding the number and length of those phases.

Theorem 2 (Wald's equation [20, page 346]). *Let $X_1, X_2, \ldots$ be nonnegative, independent, identically distributed random variables with distribution X. Let T be a stopping time for this sequence. If T and X have bounded expectation, then*

$$\mathbb{E}\left[\sum_{i=1}^{T} X_i\right] = \mathbb{E}[T] \cdot \mathbb{E}[X].$$

The following theorem bounds the expected value of the maximum of n exponentially distributed random variables which helps as bounding the time until all vertices heal at least once.

Theorem 3 ([20, page 33]). *Let $n \in \mathbb{N}_{>0}$, and let $\{X_i\}_{i\in[n]}$ be independent random variables that are each exponentially distributed with parameter $\lambda \in \mathbb{R}_{>0}$. Let $m = \max_{i\in[n]} X_i$, and let H_n be the n-th harmonic number. Then*

$$\mathbb{E}[m] = \frac{H_n}{\lambda} < \frac{1 + \ln(n+1)}{\lambda}.$$

We use Chernoff bounds to bound the value of binomially distributed random variables.

Theorem 4 ([20, Theorem 4.4, Theorem 4.5]). *Let $X_1, \ldots, X_n$ be independent Poisson trials, $X = \sum_{i=1}^{n} X_i$, $\mu = \mathbb{E}[X]$ and $\delta \in (0, 1)$. Then*

(1) $\Pr[X \geq (1+\delta)\mu] \leq e^{-\mu\delta^2/3}$,

(2) $\Pr[X \leq (1-\delta)\mu] \leq e^{-\mu\delta^2/2}$.

3 THRESHOLDS ON CLIQUES

In the standard SIS process on a clique there exists a number of infected vertices at which vertices heal and get infected at the same rate, called the equilibrium point. For the contact process with nonlinear infection rate, depending on whether the scaling is sub- or super-linear, this equilibrium is attracting or repelling, respectively. A sub-linear scaling leads to an attracting equilibrium, which yields a threshold close to $1/n$ (see Corollary 7). A super-linear scaling leads to a repelling equilibrium which makes it very unlikely to reach it. Hence, the infection always dies out fast a.a.s. as shown in Lemma 8. However, there still is a threshold above which the expected survival time is super-polynomial as if the infection crosses the equilibrium, the survival time becomes extremely large. The threshold is roughly $n^{-1-\alpha}$ (see Corollary 10) which is different to the sub-linear case.

3.1 Sub-Linear Scaling

When the infection rate scales sub-linearly, there is an equilibrium that is attracting. Also, when being in a state with a number of infected vertices that is a constant factor away from the equilibrium, it is already twice as likely to go towards the equilibrium than going away from it in each step. That means that the survival time is exponential in the equilibrium value. Hence, we get a threshold at the point where this exponential becomes super-polynomial in n. We first show the lower bound on the expected survival time.

Theorem 5. *Let G be a clique with $n \in \mathbb{N}_{>0}$ vertices. Further, let C be a contact process with infection coefficient $\lambda \in \mathbb{R}_{>0}$ and infection exponent $\alpha \in (-1, 0)$ on G that starts with exactly one infected vertex. Let T be the survival time of C. If $(\lambda n)^{-1/\alpha} \leq n/2$ and $(\lambda n/4)^{-1/\alpha} \geq 2$, then $\mathbb{E}[T] \in \Omega\left(2^{(\lambda n)^{-1/\alpha}}/n\right)$.*

Proof. We show that while there are at most $(\lambda n/4)^{-1/\alpha}$ infected vertices, the probability of infecting a new vertex is at least twice as high as the probability to heal one in the next step. Therefore the process dominates a gambler's ruin instance with a biased coin of probability 2/3, which has an expected exponential time to reach its lower bound.

Let $c \in \mathbb{R}_{\geq 1}$. Consider a state with $I = (\lambda n/c)^{-1/\alpha}$ infected vertices. From this state, vertices heal at a rate of I and because $(\lambda n)^{-1/\alpha} \leq n/2$, new vertices get infected at a rate of

$$\lambda I^{1+\alpha}(n-I) \geq \lambda I^{1+\alpha} n/2 = \lambda I (\lambda n/c)^{-\alpha/\alpha} n/2 = \frac{c}{2} I.$$

Note that for $c \geq 4$, the rate at which new vertices get infected is at least twice as high as the rate at which vertices heal. Hence, while there are at most $I = (\lambda n/4)^{-1/\alpha}$ infected vertices, the discrete version of the contact process dominates a gambler's ruin instance A with a biased coin with probability 2/3.

For all $i \in \mathbb{N}, i < I$, let p_i be the probability that starting with i infected vertices, the infection dies out before reaching I infected vertices. We get using Theorem 1

$$p_1 \leq \frac{1-2^{I-1}}{1-2^I} = \frac{2^{I-1}-1}{2^I-1} \leq \frac{1}{2}$$

and

$$p_{I-1} \leq \frac{1-2^1}{1-2^I} = \frac{1}{2^I-1}.$$

Starting with one infected vertex, the probability to reach I infected vertices before the infection dies out is $1 - p_1 \geq 1/2$. From that point on, the number of times that the infection reaches I infected vertices again from below dominates a geometric random variable X with success probability $p_{I-1} \leq \frac{1}{2^I-1}$. Note that between each of those times, a vertex has to heal, which happens at a rate of at most n. That means that the expected time between two of those events is at least $1/n$. Together with Wald's Equation (Theorem 2), that gives us $\mathbb{E}[T] \geq 1/2 \cdot \mathbb{E}[X]/n$. Plugging in the expected value of X concludes the proof. □

We get a very similar result for the upper bound by upper bounding the probability to increase the number of infected vertices in the next step instead of lower bounding it. The detailed proof can be found in the full version of the paper [6].

Theorem 6. *Let G be a clique with $n \in \mathbb{N}_{>0}$ vertices. Further, let C be a contact process with infection coefficient $\lambda \in \mathbb{R}_{>0}$ and infection exponent $\alpha \in (-1, 0)$ on G that starts with exactly one infected vertex. Let T be the survival time of C. If $\lambda n \geq 2$, then $\mathbb{E}[T] \in \mathrm{O}\left((\lambda n)^{(\lambda n)^{-1/\alpha}}\right)$.*

Using those two results, we can pinpoint the threshold relatively precisely.

Corollary 7. *Let G be a clique with $n \in \mathbb{N}_{>0}$ vertices. Further, let C be a contact process with infection coefficient $\lambda \in \mathbb{R}_{>0}$ and infection exponent $\alpha \in (-1, 0)$ on G that starts with exactly one infected vertex. Let T be the survival time of C.*

(1) If $\lambda \in \omega(\log(n)^{-\alpha}/n)$ then $\mathbb{E}[T]$ is super-polynomial in n.
(2) If $\lambda \in \mathrm{O}(1/n)$ then $\mathbb{E}[T]$ is constant in n.

3.2 Super-Linear Scaling

When the infection rate scales super-linearly, the equilibrium point becomes repellent. When being a constant factor away from that equilibrium, there is a constant drift away from the equilibrium. Therefore, the process is very unlikely to reach this equilibrium, but when it does, the infection might infect all vertices and stay above the equilibrium for a very long time. The expected survival time is therefore decided by the product of the probability of reaching the equilibrium and the expected time the process stays above the equilibrium.

We bound the probability of reaching the equilibrium first. Then we calculate the expected survival time after reaching the equilibrium. Putting those two results together gives us the lower and upper bound for the expected survival time. Note that for the upper bound on the survival time we do not need the second lemma as in the regime we consider the equilibrium value is above the number of vertices and can therefore never be reached.

The following lemma bounds the probability to reach the equilibrium from both sides.

Lemma 8. *Let G be a clique with $n \in \mathbb{N}_{>0}$ vertices. Further, let C be a contact process with infection coefficient $\lambda \in \mathbb{R}_{>0}$ and infection exponent $\alpha \in \mathbb{R}_{>0}$ on G that starts with exactly one infected vertex. Let p be the probability that the infection reaches a state with $(\lambda n/2)^{-1/\alpha}$ infected vertices. If $1 \leq (\lambda n/2)^{-1/\alpha} \leq n/2$, then*

$$2^{1-(2\lambda n)^{-1/\alpha}} \geq p \geq (\lambda n/2)^{(\lambda n/2)^{-1/\alpha}}.$$

Proof. We bound the process below $(\lambda n/2)^{-1/\alpha}$ infected vertices with one gambler's ruin instance each for the upper and lower bound by upper and lower bounding the probability to infect a new vertex in the next step.

Let $c \in \mathbb{R}_{\geq 1/2}$. Consider a state with $I = (c\lambda n)^{-1/\alpha}$ infected vertices. From this state, vertices heal at a rate of I and because $(\lambda n/2)^{-1/\alpha} \leq n/2$, new vertices get infected at a rate of

$$\begin{aligned} \lambda I^{1+\alpha}(n-I) &\leq \lambda I^{1+\alpha} n \\ &= \lambda I (c\lambda n)^{-\alpha/\alpha} n \\ &= I/c \end{aligned} \tag{1}$$

and

$$\begin{aligned} \lambda I^{1+\alpha}(n-I) &\geq \lambda I^{1+\alpha} n/2 \\ &= \lambda I (c\lambda n)^{-\alpha/\alpha} n/2 \\ &= \frac{1}{2c} I. \end{aligned} \tag{2}$$

Note that these bounds are exactly the same bounds we got for negative α. However, now that α is positive, a higher c actually decreases I instead of increasing it. That means that for I below $(\lambda n)^{-1/\alpha}$, it is more likely to heal vertices than to infect new ones.

Let p_1 be the probability that the infection reaches a state with $(2\lambda n)^{-1/\alpha}$ infected vertices. As $(2\lambda n)^{-1/\alpha} \leq (\lambda n/2)^{-1/\alpha}$ it holds $p_1 \geq p$. Choosing $c = 2$ in Equation (1) implies that below $(2\lambda n)^{-1/\alpha}$ infected vertices, the probability to infect a new vertex in the next step is at most half as high as the probability to heal one. Therefore the number of infected vertices in the discrete version of C dominates a gambler's ruin instance A_1 with a biased coin of probability 1/3 which gives us using Theorem 1 that

$$p_1 \leq \frac{1-2^1}{1-2^{(2\lambda n)^{-1/\alpha}}} = \frac{1}{2^{(2\lambda n)^{-1/\alpha}}-1} \leq 2^{1-(2\lambda n)^{-1/\alpha}}.$$

For the lower bound on p note that the probability to infect a vertex in the next step is minimized when the number of infected vertices is 1. That corresponds to $c = (\lambda n)^{-1}$ in Equation (2), which gives us that as long as the infection did not die out, the probability to infect a vertex in the next step is at least $\frac{\lambda n/2}{\lambda n/2+1}$. Therefore the number of infected vertices of the discrete version of C is dominated by a gambler's ruin instance A_2 with a biased coin of probability $\frac{\lambda n/2}{\lambda n/2+1}$ which gives us

$$p \geq \frac{1-(\lambda n/2)^{-1}}{1-(\lambda n/2)^{-(\lambda n/2)^{-1/\alpha}}} = \frac{(\lambda n/2)^{-1}-1}{(\lambda n/2)^{-(\lambda n/2)^{-1/\alpha}}-1} \geq (\lambda n/2)^{(\lambda n/2)^{-1/\alpha}}. \qquad \square$$

The following lemma lower bounds the time we stay above the equilibrium. Its proof is very similar to Theorem 5 and uses that while we are a constant factor above the equilibrium, the probability to infect a new vertex is at least twice as high as the probability to heal one in the next step. That gives us an exponential survival time. The detailed proof can be found in the full version of the paper [6].

Lemma 9. *Let G be a clique with $n \in \mathbb{N}_{>0}$ vertices. Further, let C be a contact process with infection coefficient $\lambda \in \mathbb{R}_{>0}$ and infection exponent $\alpha \in \mathbb{R}_{>0}$ on G that starts with $(\lambda n/2)^{-1/\alpha}$ infected vertices. Let T be the survival time of C. If $(\lambda n/2)^{-1/\alpha} \in o(n)$, then $\mathbb{E}[T] \geq 2^{\Theta(n)}$.*

Putting those two results together gives us the following threshold.

Corollary 10. *Let G be a clique with $n \in \mathbb{N}_{>0}$ vertices. Further, let C be a contact process with infection coefficient $\lambda \in \mathbb{R}_{>0}$ and infection exponent $\alpha \in \mathbb{R}_{>0}$ on G that starts with exactly one infected vertex. Let T be the survival time of C.*

(1) If $\lambda \in \omega\left(n^{-1-\alpha}\log(n)^{\alpha}\right)$ then $\mathbb{E}[T]$ is exponential in n.
(2) If $\lambda \in o\left(n^{-1-\alpha}\right)$ then $\mathbb{E}[T]$ is constant in n.

Proof. First consider $\lambda \in \omega\left(n^{-1-\alpha}\log(n)^{\alpha}\right)$. Note that this implies that $\lambda n^{-1/\alpha} \in o(n/\log(n))$. Now consider all λ such that $\lambda n \geq 1$. For all larger λ we get the same results using the fact that the expected survival time scales monotonically with λ. Now the conditions for Lemma 8 and Lemma 9 are fulfilled. Let E_0 be the event that the infection reaches a state with $(\lambda n/2)^{-1/\alpha}$ infected vertices. By Lemma 8 it holds that

$$\Pr[E_0] \geq (\lambda n/2)^{(\lambda n/2)^{-1/\alpha}} \geq (\lambda n/2)^{o(n/\log(n))} \geq 2^{-o(n)}.$$

By Lemma 9 it holds that $\mathbb{E}[T \mid E_0] \geq 2^{\Theta(n)}$. Together we get

$$\mathbb{E}[T] \geq \Pr[E_0]\cdot\mathbb{E}[T \mid E_0] \geq 2^{-o(n)}\cdot 2^{\Theta(n)} \geq 2^{\Theta(n)}.$$

Now consider $\lambda \in o\left(n^{-1-\alpha}\right)$. Let $c \in \mathbb{R}_{>0}$. Consider a state with $I = (c\lambda n)^{-1/\alpha}$ infected vertices. From this state, vertices heal at a rate of I and new vertices get infected at a rate of

$$\lambda I^{1+\alpha}(n-I) \leq \lambda I^{1+\alpha} n = \lambda I (c\lambda n)^{-\alpha/\alpha} n = I/c.$$

As $\lambda \in o\left(n^{-1-\alpha}\right)$, it holds that $(\lambda n)^{-1/\alpha} \in \omega(n)$. Therefore for each reachable state it holds that $c \geq 2$, which means that the probability to heal a vertex is always at least twice as high as the probability to infect one. Therefore the process is dominated by a biased gambler's ruin instance with probability of $1/3$ to increase by 1. The gambler's ruin instance has a constant expected time to reach its lower bound and because triggers always happen at a rate of at least 1, that also gives us a constant upper bound for $\mathbb{E}[T]$. $\square$

4 THRESHOLDS ON STARS

For the star with $n \in \mathbb{N}$ leaves we start by simplifying the process by only considering the number of infected leafs I_t at step t and whether the center is infected or not. We then get the following transition rates

when the center is infected:

$I_{t+1} = I_t + 1$	at rate $\lambda(n-I_t)$,
$I_{t+1} = I_t - 1$	at rate I_t,
the center heals	at rate 1.

when the center is healthy:

$I_{t+1} = I_t + 1$	at rate 0,
$I_{t+1} = I_t - 1$	at rate I_t,
the center gets infected	at rate $\lambda I_t^{1+\alpha}$.

This time our analysis works for both positive and negative α. We first define $\beta = \lambda^2 n(\lambda n)^{\alpha}$ because this value dictates the survival time of the process. We now start with the upper bound which is derived similarly to the normal setting.

Lemma 11. *Let G be a star with $n \in \mathbb{N}_{>0}$ leaves. Further, let C be a contact process with infection coefficient $\lambda \in (0,1)$ and infection exponent $\alpha \in \mathbb{R}_{>-1}$ on G that starts with infected center and susceptible leaves. Let T be the survival time of C. If $\lambda \in o\left(n^{-\alpha/(1+\alpha)}\right)$ and $\lambda \in \omega\left(n^{-1}\right)$, then $\mathbb{E}[T] \leq 2^{\Theta(\beta)}\log(n)$.*

Proof. We first show that the process stays at fewer than $3\lambda n$ infected leaves for most of the time even if the center is permanently infected. We then lower bound the probability for the infection to die out in a single center-healthy phase when it starts with at most

$3\lambda n$ infected leaves. Bounding the expected length of center-healthy and center-infected phases concludes the proof.

Consider the modified process C' that behaves like C with the exception that it ignores all of the healing triggers of the center. The number of infected leaves in C' dominates the number of infected leaves in C. While there are I' infected leaves in C', leaves heal at a rate of I' and new leaves get infected at a rate of $\lambda(n - I') \leq \lambda n$. Therefore, while $I' \geq 2\lambda n$, leaves heal at least twice as fast as new leaves get infected. That means that between $2\lambda n$ and $3\lambda n$ infected leaves, the discrete version of the process is dominated by a gambler's ruin instance with a biased coin with success probability $1/3$. As this instance takes in expectation an exponential time in λn to reach its upper limit and only a linear time to drop back to its lower bound, the time that C' spends above $3\lambda n$ infected leaves is negligible compared to the entire survival time T. As C' dominates C, the same holds for C. Therefore, in the following we assume that the number of infected leaves stays below $3\lambda n$ without reducing the expected survival time too much with this assumption.

Consider a state with at most $3\lambda n$ infected leaves and healthy center. Let E be the event that the infection dies out before infecting the center. In a state with I infected leaves, leaves heal at a rate of I and the center gets infected at rate $\lambda I^{1+\alpha}$. In order for E to happen, all of the leaves have to heal before the center gets infected. It holds for all $x \in [0, 1.59]$ that $e^{-x} \leq 1 - x/2$. Note that $\lambda \in o\left(n^{-\alpha/(1+\alpha)}\right)$ implies $\lambda(\lambda n)^\alpha \in o(1)$. Hence, for sufficiently large n, we can use the previous inequality with $x/2 = \lambda(\lambda n)^\alpha$. With that, we bound the probability of E by

$$\begin{aligned}
\Pr[E] &\geq \prod_{i=1}^{3\lambda n} \frac{i}{i + \lambda i^{1+\alpha}} = \prod_{i=1}^{3\lambda n} \frac{1}{1 + \lambda i^\alpha} \\
&= \prod_{i=1}^{3\lambda n} \left(1 - \frac{\lambda i^\alpha}{1 + \lambda i^\alpha}\right) \geq \prod_{i=1}^{3\lambda n} \left(1 - \lambda i^\alpha\right) \\
&\geq \prod_{i=1}^{3\lambda n} e^{-2\lambda i^\alpha} = e^{-2\lambda \sum_{i=1}^{3\lambda n} i^\alpha} \\
&= e^{-\Theta\left(\lambda(\lambda n)^{1+\alpha}\right)} \\
&= e^{-\Theta(\beta)}. \qquad (3)
\end{aligned}$$

The second to last step uses the Euler-Maclaurin Formula (see [13]).

Let S be the number of center-healthy phases of C that start with at most $3\lambda n$ infected leaves before the infection dies out. By Equation (3), S is dominated by a geometric random variable with parameter $e^{-\Theta(\beta)}$. Each center-infected phase lasts in expectation for 1 time unit as it ends when the center heals which happens at rate 1. By Theorem 3, each center-healthy phase lasts at most $\log(n)$ time units in expectation as it ends the latest when the last leaf heals which is determined by the maximum of n exponential random variables with mean 1. To bound T, we also need to consider the time T' spent above $3\lambda n$ infected leaves, but as we argued before, that time is much smaller than the rest of T. We get

$$\begin{aligned}
\mathbb{E}[T] &\leq \mathbb{E}[S] \cdot (1 + \log(n)) + T' \\
&\leq e^{\Theta(\beta)} \cdot (1 + \log(n)) + T' \\
&\leq 2^{\Theta(\beta)} \log(n). \qquad \square
\end{aligned}$$

For the lower bound, the idea is to look at the process while the number of infected vertices is in between $\lambda n/8$ and $\lambda n/4$ and splitting this range into $\sqrt{\beta}$ many equally sized blocks. We then use the center-healthy phases and center-infected phases as the steps of our process. Now in a center-healthy phase, the probability of decreasing the number of infected vertices by more than a block is exponentially small in $\sqrt{\beta}$. In a center-healthy phase, the probability of not increasing by a block before healing the center is sub constant. We now build a gambler's ruin game on the $\sqrt{\beta}$ many blocks as states by taking as a step a full cycle between two center infections. The probability of decreasing by more than one block is so small that we ignore it. The probability of increasing by a block in total is much higher than the probability of decreasing by one (much more than twice as much). So our process dominates a biased gambler's ruin on $\sqrt{\beta}$ many states with a biased coin of probability $2/3$. Therefore, the time it takes to die out is with high probability exponential in $\sqrt{\beta}$.

We first bound the probability of healing too many vertices in a center-healthy phase.

Lemma 12. *Let G be a star with $n \in \mathbb{N}_{>0}$ leaves. Further, let C be a contact process with infection coefficient $\lambda \in \mathbb{R}_{>0}$ and infection exponent $\alpha \in \mathbb{R}_{>-1}$ on G. Let $x, y \in \mathbb{N}$. Further, let p_y^x be the probability to drop in a center-healthy phase from x infected vertices to at most y. Let $z = y$ if α is positive and $z = x$ otherwise. If $\lambda z^\alpha \leq 1$, then $p_y^x \leq e^{-(x-y)\frac{\lambda z^\alpha}{2}}$.*

Proof. While there are $I \in \mathbb{N}$ infected leaves, leaves heal at a rate of I and the center gets infected at a rate of $\lambda I^{1+\alpha}$. Hence, in the next step a leaf heals with a probability of $\frac{I}{I+\lambda I^{1+\alpha}} = \frac{1}{1+\lambda I^\alpha}$. Note that this probability is monotonically decreasing or increasing in I when α is positive or negative respectively. For our choice of z, that means that in the interval between y and x, this probability is maximized at z. Now in order to drop from x infected leaves to y in a center-healthy phase, in each state in between a leaf has to heal. Using $\lambda z^\alpha \leq 1$, we bound that probability by

$$\begin{aligned}
p_y^x &= \prod_{i=y+1}^{x} \frac{1}{1 + \lambda i^\alpha} \leq \left(\frac{1}{1 + \lambda z^\alpha}\right)^{x-y} \\
&= \left(1 - \frac{\lambda z^\alpha}{1 + \lambda z^\alpha}\right)^{x-y} \leq \left(1 - \frac{\lambda z^\alpha}{2}\right)^{x-y} \\
&\leq e^{-(x-y)\frac{\lambda z^\alpha}{2}}. \qquad \square
\end{aligned}$$

Now we lower bound the probability of infecting enough vertices in a center-infected phase.

Lemma 13. *Let G be a star with $n \in \mathbb{N}_{>0}$ leaves and let $z \in \mathbb{R}_{>0}$. Further, let C be a contact process with infection coefficient $\lambda \in (0, 1)$ and infection exponent $\alpha \in \mathbb{R}_{>-1}$ on G that starts with infected center and at most $\lambda n/4 - \frac{\lambda n}{4z}$ infected leaves. Let $\lambda n \in \omega(1)$. Then if $z \in o(\lambda n)$, the probability of the event E that the number of infected leaves increases by at least $\frac{\lambda n}{4z}$ before the center heals is at least $e^{-\Theta(1/z)}$.*

Proof. We first show that it is likely that there are at least $\frac{\lambda n}{z}$ steps that do not heal the center and to then lower bound the probability that enough of these steps infect new leaves.

While the center is infected, each leaf either heals at a rate of 1 or gets infected at a rate of λ depending on whether it is infected or not. Thus, each leaf changes its state at a rate of at least λ. The center heals at rate one. That means that each step has a probability of at most $\frac{1}{\lambda n}$ to heal the center. Therefore, the number of steps S before the center heals dominates a geometric random variable X with parameter $\frac{1}{\lambda n}$. We get

$$\Pr\left[S \geq \frac{\lambda n}{z}\right] \geq \Pr\left[X \geq \frac{\lambda n}{z}\right] \geq \left(1 - \frac{1}{\lambda n}\right)^{\frac{\lambda n}{z}} \geq e^{-\frac{1}{\lambda n}\frac{\lambda n}{z}} \geq e^{-\Theta(\frac{1}{z})}.$$

Now when there are at least $\frac{\lambda n}{z}$ steps before the center heals, then if at least 5/8-th of those infect new vertices that implies E. While there are at most $\lambda n/4$ infected leaves, leaves heal at a rate of at most $\lambda n/4$ and new vertices get infected at a rate of at least $\lambda(n - \lambda n/4) \geq \lambda n/2$. Hence, each step has a probability of at least 2/3 to infect a leaf. That means that the number of the steps out of the first S steps that infect a new leaf dominates a binomial random variable $B \sim \mathrm{Bin}(S, 2/3)$. We get using Chernoff bounds (Theorem 4)

$$\begin{aligned}\Pr\left[E \,\middle|\, S \geq \frac{\lambda n}{z}\right] &\geq \Pr\left[B \geq \frac{5}{8}S \,\middle|\, S \geq \frac{\lambda n}{z}\right] \\ &\geq \Pr\left[B \geq (1 - \frac{1}{16})\mathbb{E}[B] \,\middle|\, S \geq \frac{\lambda n}{z}\right] \\ &\geq 1 - e^{-\frac{1}{2\cdot 16^2}\cdot\frac{\lambda n}{z}}.\end{aligned}$$

This gives us a lower bound for the probability of E of

$$\begin{aligned}\Pr[E] &\geq \Pr\left[S \geq \frac{\lambda n}{z}\right] \cdot \Pr\left[E \,\middle|\, S \geq \frac{\lambda n}{z}\right] \\ &\geq e^{-\Theta(\frac{1}{z})} \cdot (1 - e^{-\frac{1}{2\cdot 16^2}\cdot\frac{\lambda n}{z}}) \\ &\geq e^{-\Theta(\frac{1}{z})}.\end{aligned}$$

□

Putting those results together gives us the following lower bound.

Theorem 14. *Let G be a star with $n \in \mathbb{N}_{>0}$ leaves. Further, let C be a contact process with infection coefficient $\lambda \in (0,1)$ and infection exponent $\alpha \in \mathbb{R}_{>-1}$ on G that starts with infected center and at least $\lambda n/4$ infected leaves. Let T be the survival time of C. If $\lambda(\lambda n)^\alpha \leq 1$ and $\beta \in \omega(1)$, then a.a.s. it holds $T \geq 2^{\Theta\left(\sqrt{\beta}\right)}$.*

Proof. We look at the process while the number of infected vertices is in between $\lambda n/8$ and $\lambda n/4$ and split this interval into $\sqrt{\beta}$ many equally sized blocks. We then only consider in which of those blocks the number of infected vertices is each time the center gets infected. We show that the resulting process with high probability dominates a gambler's ruin instance with a biased coin of probability 2/3 which gives us the desired bound.

We define the process X that is coupled to C as follows. X transitions to different values at exactly the times at which the center gets infected in C. It takes as values the number of infected leaves in C divided by $\frac{\lambda n}{8\sqrt{\beta}}$ rounded down to an integer. To bound the survival time T, we consider the process until one of two events happens: Either X decreases by more than 1 in a step or it reaches $\sqrt{\beta}$. We only consider C below $\lambda n/4$ infected vertices. Every time when a center gets infected while more leaves are infected we ignore it.

In order for X to reduce by more than 1 in a step, either the center-infected phase or the center-healthy phase has to reduce the number of infected leaves by at least $\frac{\lambda n}{16\sqrt{\beta}}$. While below $\lambda n/4$ infected leaves, center-infected phases heal vertices at least twice as fast as they heal leaves which makes it exponentially unlikely to reduce the number of infected leaves by too much. For the center-healthy phases, by Lemma 12 the probability to reduce the number of infected leaves in a center-healthy phase by $\frac{\lambda n}{16\sqrt{\beta}}$ while the number of infected leaves is between $\lambda n/8$ and $\lambda n/4$ is at most $e^{-\Theta\left(\frac{\lambda n}{\sqrt{\beta}}\cdot\lambda(\lambda n)^\alpha\right)} = e^{-\Theta\left(\sqrt{\beta}\right)}$. That means that the time until that event happens is at least geometrically distributed with probability $e^{-\Theta\left(\sqrt{\beta}\right)}$ which means that it takes a.a.s. $e^{\Theta\left(\sqrt{\beta}\right)}$ until then.

Now assume that X never reduces by more than 1 in a step. By Lemma 13, the probability that X increases by at least 1 in a step is at least $e^{-\Theta\left(1/\sqrt{\beta}\right)}$ which is at least 2/3 for sufficiently large n as $\beta \in \omega(1)$. That means that X dominates a gambler's ruin instance with a biased coin of probability 2/3 in the range between $\sqrt{\beta}$ and $2\sqrt{\beta}$. This gambler's ruin instance has a.a.s. a time of $2^{\Theta\left(\sqrt{\beta}\right)}$ until it reaches its lower bound.

Now the infection cannot die out before X reaches $\sqrt{\beta}$. We argued that this does not happen before either X reduces by more than 1 in a step or the gambler's ruin instance reaches $\sqrt{\beta}$. As shown before, both of those events take a.a.s. $2^{\Theta\left(\sqrt{\beta}\right)}$ time. As each phase considered in X needs the center to heal and then infect which takes at least 1 time unit in expectation, this gives us the desired bound for T. □

We now show that starting with only the center infected, we reach a state with $\lambda n/4$ infected leaves with high probability.

Lemma 15. *Let G be a star with $n \in \mathbb{N}_{>0}$ leaves. Further, let C be a contact process with infection coefficient $\lambda \in (0,1)$ and infection exponent $\alpha \in \mathbb{R}_{>-1}$ on G that starts with infected center and susceptible leaves. If $\lambda(\lambda n)^\alpha \leq 1$ and $\beta \in \omega(\log(n)^4)$, C reaches a state with $\lambda n/4$ infected leaves asymptotically almost surely.*

Proof. We show that asymptotically almost surely we reach a state from which we need a super-constant amount of center-healthy phases for the infection to die out. We then show that each center-infected phase has a constant probability to reach $\lambda n/4$ infected leaves. That makes it very unlikely not to reach $\lambda n/4$ infected leaves before the infection dies out.

By Lemma 13, the probability that the process reaches a state with at least $\frac{\lambda n}{\beta^{\frac{1}{2(1+\alpha)}}}$ infected leaves is at least $e^{-\Theta\left(1/\beta^{\frac{1}{2(1+\alpha)}}\right)}$. Now consider phases of the process that start and end with the center

infecting or the infection dying out. By Lemma 12 between $\frac{\lambda n}{2\beta^{\frac{1}{2(1+\alpha)}}}$ and $\frac{\lambda n}{\beta^{\frac{1}{2(1+\alpha)}}}$ infected leaves, the probability p to heal more than $\frac{\lambda n}{\beta^{\frac{1}{2(1+\alpha)}} \cdot \sqrt[4]{\beta}}$ is at most

$$p \leq e^{-\frac{\lambda n}{\beta^{\frac{1}{2(1+\alpha)}} \cdot \sqrt[4]{\beta}} \cdot \lambda \left(\frac{\lambda n}{2\beta^{\frac{1}{(1+\alpha)}}}\right)^{\alpha}} \leq e^{-\frac{\beta}{\beta^{\frac{1+\alpha}{2(1+\alpha)}} \sqrt[4]{\beta}}} \leq e^{-\sqrt[4]{\beta}}.$$

As $\beta \in \omega\left(\log(n)^4\right)$, that means that asymptotically almost surely, there are at least $\sqrt[4]{\beta}$ many of those phases before the infection dies out. By Lemma 13, each of those phases has a constant probability to infect $\lambda n/4$ many leaves. Therefore we reach a state with $\lambda n/4$ infected leaves before the infection dies out asymptotically almost surely. □

Plugging in the values for λ and combining the previous lemmas gives us the following bounds

Corollary 16. *Let G be a star with $n \in \mathbb{N}_{>0}$ leaves. Further, let C be a contact process with infection coefficient $\lambda \in \mathbb{R}_{>0}$ and infection exponent $\alpha \in \mathbb{R}_{>-1}$ on G that starts with infected center and susceptible leaves. Let T be the survival time of C.*

(1) If $\lambda \in \omega\left(n^{-1/2-\frac{\alpha}{2(2+\alpha)}} \log(n)^{4/(2+\alpha)}\right)$ then T is a.a.s. super-polynomial in n.

(2) If $\lambda \in \mathrm{O}\left(n^{-1/2-\frac{\alpha}{2(2+\alpha)}}\right)$ then $\mathbb{E}[T]$ is at most logarithmic in n.

Proof. First consider $\lambda \in \omega\left(n^{-1/2-\frac{\alpha}{2(2+\alpha)}} \log(n)^{4/(2+\alpha)}\right)$. We show the survival time lower bound for $\lambda \in \mathrm{o}\left(n^{-\alpha/(1+\alpha)}\right)$ and $\lambda \in \mathrm{o}(1)$ as it also holds for all larger λ then because of the monotonicity of the survival time. Note that such a λ always exists as $-1/2 - \frac{\alpha}{2(2+\alpha)} < -\alpha/(1+\alpha)$ and $-1/2 - \frac{\alpha}{2(2+\alpha)} < 0$ for all $\alpha \in \mathbb{R} \setminus [-2,-1]$. Now $\lambda \in \mathrm{o}\left(n^{-\alpha/(1+\alpha)}\right)$ and $\lambda \in \mathrm{o}(1)$ imply both $\lambda \leq 1$ and $\lambda(\lambda n)^{\alpha} \leq 1$ for sufficiently large n. The fact that $\lambda \in \omega\left(n^{-1/2-\frac{\alpha}{2(2+\alpha)}} \log(n)^{4/(2+\alpha)}\right)$ implies that $\beta \in \omega\left(\log(n)^4\right)$. Hence, both Theorem 14 and Lemma 15 are applicable.

By Lemma 15, C reaches a state with $\lambda n/4$ a.a.s and by Theorem 14 it then survives a.a.s. for a time of at least $2^{\Theta\left(\sqrt{\beta}\right)}$. As $\beta \in \omega\left(\log(n)^4\right)$, this time is super-polynomial in n which concludes the proof for the first case.

Now consider $\lambda \in \mathrm{O}\left(n^{-1/2-\frac{\alpha}{2(2+\alpha)}}\right)$. We show the survival time upper bound for $\lambda \in \Theta\left(n^{-1/2-\frac{\alpha}{2(2+\alpha)}}\right)$ and it follows for smaller λ because of the linearity of the survival time. Note that for $\alpha > -2$ it holds $-1/2 - \frac{\alpha}{2(2+\alpha)} > -1$ and hence $\lambda \in \omega\left(n^{-1}\right)$. Also for all $\alpha \in \mathbb{R} \setminus [-2,-1]$ it holds $-1/2 - \frac{\alpha}{2(2+\alpha)} < -\alpha/(1+\alpha)$ and $-1/2 - \frac{\alpha}{2(2+\alpha)} < 0$ which imply $\lambda \in \mathrm{o}\left(n^{-\alpha/(1+\alpha)}\right)$ and $\lambda < 1$ for sufficiently large n. Hence Lemma 11 is applicable which gives us that $\mathbb{E}[T] \leq 2^{\Theta(\beta)} \log(n)$. Noting that for $\lambda \in \Theta\left(n^{-1/2-\frac{\alpha}{2(2+\alpha)}}\right)$, $\beta \in \Theta(1)$, concludes the proof. □

5 CONCLUSION AND OUTLOOK

We conducted the first fully rigorous analysis of the SIS model with non-linear infection rates in the number of infected neighbors. Our results provide almost-tight lower and upper bounds on the epidemic threshold for both star and clique graphs. Those are important graph classes because they provide insights for the two extremes of very little connectivity and full connectivity, respectively. Moreover, since we consider the SIS process, survival time lower bounds also carry over to graphs that contain those classes as sub graphs. Hence, our lower bounds carry over to graphs with very high degree vertices.

On cliques with sub-linear scaling, the process behaves very similar to the normal SIS process and the threshold only shifts by a poly-logarithmic factor. When the infection scales super-linearly, the process completely changes as at the threshold, there exists an equilibrium point that is repellent. That means that the process dies out very fast asymptotically almost surely but if it reaches the equilibrium, it survives so long that the overall expected survival time is still exponential. The threshold in this setting is by a polynomial factor smaller than in the sub-linear case. On stars, we show that sub-linear and super-linear scaling do not differ that much and the threshold changes by a polynomial factor that depends on the scaling of the infection.

As a next step it is very interesting to see to which extend these results carry over to larger graph classes or even real-world graphs. A starting point for that could be running experiments on such graphs. For further theoretical results, we believe that expansion is the most promising property to analyze. That would give results for a lot of graph classes like Erdős–Rényi graph. High expansion results in a very good bound on the number of edges between susceptible and infected vertices. In the standard SIS model, this translates to a bound on the rate at which vertices get infected. However, in our model, this is not immediate, as the rate at which vertices get infected also depends on how their edges are distributed. It makes a difference whether all susceptible vertices have roughly the same number of infected neighbors or whether some of them have many more than others. This makes our model harder to analyze than the standard model.

Another direction would be to look at more advanced infection models like the SIRS model, which adds temporary immunity to the process, and adjust those in the same way to non-linear infection rates. That gives a better understanding into under which conditions the adjusted scaling makes a difference and how much of a difference it is.

The adjusted infection function can model infections that scale polynomially with the number of infected neighbors. It would be interesting to also look into other functions like for example threshold functions that stay at 0 until a specific threshold is reached and only then start increasing.

ACKNOWLEDGMENTS

The authors thank George Skretas for helpful discussions and his guidance on the literature of influence maximization.

Andreas Göbel was funded by the project PAGES (project No. 467516565) of the German Research Foundation (DFG).

Marcus Pappik was funded by the HPI Research School on Data Science and Engineering.

REFERENCES

[1] Nitin Agarwal and Huan Liu. 2008. Blogosphere: Research Issues, Applications, and Tools. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*.

[2] Noam Berger, Christian Borgs, Jennifer T. Chayes, and Amin Saberi. 2005. On the spread of viruses on the internet. In *Symposium on Discrete Algorithms (SODA)*. 301–310. https://doi.org/10.5555/1070432.1070475

[3] Christian Borgs, Jennifer Chayes, Ayalvadi Ganesh, and Amin Saberi. 2010. How to distribute antidote to control epidemics. *Random Structures & Algorithms* 37, 2 (2010), 204–222. https://doi.org/10.1002/rsa.20315

[4] Damon Centola. 2010. The spread of behavior in an online social network experiment. *Science* 329, 5996 (2010), 1194–1197. https://doi.org/10.1126/science.1185231

[5] William Feller. 1968. *An Introduction to Probability Theory and its Applications* (3 ed.). Vol. 1. John Wiley & Sons.

[6] Tobias Friedrich, Andreas Göbel, Nicolas Klodt, Martin S. Krejca, and Marcus Pappik. 2024. From Market Saturation to Social Reinforcement: Understanding the Impact of Non-Linearity in Information Diffusion Models. arXiv:2401.10818 [math.PR]

[7] Ayalvadi Ganesh, Laurent Massoulié, and Don Towsley. 2005. The effect of network topology on the spread of epidemics. In *International Conference on Computer Communications (INFOCOM)*. 1455–1466. https://doi.org/10.1109/INFCOM.2005.1498374

[8] T. E. Harris. 1974. Contact interactions on a lattice. *The Annals of Probability* 2, 6 (1974), 969 – 988. https://doi.org/10.1214/aop/1176996493

[9] Herbert W Hethcote and Pauline van den Driessche. 1991. Some epidemiological models with nonlinear incidence. *Journal of Mathematical Biology* 29, 3 (1991), 271–287.

[10] Nathan O Hodas and Kristina Lerman. 2014. The simple rules of social contagion. *Scientific reports* 4, 1 (2014).

[11] Keisuke Ikeda, Takeshi Sakaki, Fujio Toriumi, and Satoshi Kurihara. 2016. An examination of a novel information diffusion model: considering of twitter user and twitter system features. In *AAMAS*. Springer, 180–191.

[12] Julie Jiang, Xiang Ren, and Emilio Ferrara. 2023. Retweet-BERT: Political Leaning Detection Using Language Features and Information Diffusion on Social Networks. In *Proceedings of the International AAAI Conference on Web and Social Media*. 459–469.

[13] Victor G Kac and Pokman Cheung. 2002. *Quantum calculus*. Springer.

[14] David Kempe, Jon Kleinberg, and Éva Tardos. 2003. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*. 137–146.

[15] Fang Kong, Jize Xie, Baoxiang Wang, Tao Yao, and Shuai Li. 2023. Online Influence Maximization under Decreasing Cascade Model. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*. 2197–2204.

[16] Jure Leskovec, Andreas Krause, Carlos Guestrin, Christos Faloutsos, Jeanne VanBriesen, and Natalie Glance. 2007. Cost-effective outbreak detection in networks. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*. 420–429.

[17] Jure Leskovec, Mary McGlohon, Christos Faloutsos, Natalie Glance, and Matthew Hurst. 2007. Patterns of cascading behavior in large blog graphs. In *Proceedings of the 2007 SIAM international conference on data mining*. SIAM, 551–556.

[18] Thomas M. Liggett. 1996. Multiple transition points for the contact process on the binary tree. *The Annals of Probability* 24, 4 (1996), 1675–1710. https://doi.org/10.1214/aop/1041903202

[19] Wei-min Liu, Simon A Levin, and Yoh Iwasa. 1986. Influence of nonlinear incidence rates upon the behavior of SIRS epidemiological models. *Journal of mathematical biology* 23 (1986), 187–204.

[20] Michael Mitzenmacher and Eli Upfal. 2017. *Probability and Computing: Randomization and Probabilistic Techniques in Algorithms and Data Analysis* (2 ed.). Cambridge university press.

[21] Bjarke Mønsted, Piotr Sapieżyński, Emilio Ferrara, and Sune Lehmann. 2017. Evidence of complex contagion of information in social media: An experiment using Twitter bots. *PloS one* 12, 9 (2017).

[22] Danny Nam, Oanh Nguyen, and Allan Sly. 2022. Critical value asymptotics for the contact process on random graphs. *Trans. Amer. Math. Soc.* 375, 12 (2022). https://doi.org/10.1090/tran/8399

[23] Mark EJ Newman. 2003. The structure and function of complex networks. *SIAM review* 45, 2 (2003), 167–256.

[24] Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani. 2015. Epidemic processes in complex networks. *Reviews of Modern Physics* 87, 3 (2015), 925–979. https://doi.org/10.1103/RevModPhys.87.925

[25] Abdul Razaque, Syed Rizvi, Meer Jaro Khan, Muder Almiani, and Amer Al Rahayfeh. 2022. State-of-art review of information diffusion models and their impact on social network vulnerabilities. *Journal of King Saud University - Computer and Information Sciences* 34, 1 (2022), 1275–1294.

[26] Guillaume St-Onge, Iacopo Iacopini, Vito Latora, Alain Barrat, Giovanni Petri, Antoine Allard, and Laurent Hébert-Dufresne. 2021. Influential groups for seeding and sustaining nonlinear contagion in heterogeneous hypergraphs. *Communications Physics* 5, 25 (2021). https://doi.org/10.1038/s42005-021-00788-w

[27] Guillaume St-Onge, Hanlin Sun, Antoine Allard, Laurent Hébert-Dufresne, and Ginestra Bianconi. 2021. Universal nonlinear infection kernel from heterogeneous exposure on higher-order networks. *Physical Review Letters* 127, 158301 (2021), 158301–1–158301–7. https://doi.org/10.1103/PhysRevLett.127.158301

[28] Yangke Sun, Bogdan Cautis, and Silviu Maniu. 2023. Social Influence-Maximizing Group Recommendation. In *Proceedings of the International AAAI Conference on Web and Social Media*. 820–831.

[29] Bryan Wilder, Laura Onasch-Vera, Juliana Hudson, Jose Luna, Nicole Wilson, Robin Petering, Darlene Woo, Milind Tambe, and Eric Rice. 2018. End-to-End Influence Maximization in the Field.. In *AAMAS*, Vol. 18. 1414–1422.

[30] Bo Yan, Kexiu Song, Jiamou Liu, Fanku Meng, Yiping Liu, and Hongyi Su. 2019. On the Maximization of Influence Over an Unknown Social Network.. In *AAMAS*, Vol. 19. 13–17.

[31] Yihui Zhang, Shaoting Tang, Sen Pei, Shu Yan, Shijin Jiang, and Zhiming Zheng. 2015. Health behavior spreading with similar diminishing returns effect. *Physica A: Statistical Mechanics and its Applications* 425 (2015), 18–26.