Avalanches in an excitable network

Reza Rastegar* Alexander Roitershtein †

April 1, 2019

Abstract

We study propagation of avalanches in a certain excitable network. The model is a particular case of the one introduced in [23], and is mathematically equivalent to an endemic variation of the Reed-Frost epidemic model introduced in [27]. Two types of heuristic approximation are frequently used for models of this type in applications, a branching process for avalanches of a small size at the beginning of the process and a deterministic dynamical system once the avalanche spreads to a significant fraction of a large network. In this paper we prove several results concerning the exact relation between the avalanche model and these limits, including rates of convergence and rigorous bounds for common characteristics of the model.

MSC2010: Primary 60J10, 60J85, secondary 92D25, 90B15, 60K40.
Keywords: avalanches in networks, excitable networks, cascading failure, criticality, branching processes, dynamic graphs.

1 Introduction

We study a discrete-time Markov model of the propagation of avalanches in a large network. Avalanches here is a general term referring to a cascading spread of a node's feature in a network of linked objects. The exact nature of the feature is immaterial for our purposes, it may have either positive or negative effect on particular aspects of the network performance. We generally refer to network's nodes possessing the feature as excited. Initially, a set of nodes becomes excited as a result of an external simulation. Once an avalanche is triggered, excited nodes can transmit this feature to currently non-excited ones, creating cascades (generations) of excited nodes evolving in discrete time.

Examples of avalanches in networks that have been studied in applications include epidemics, outages in a power grid, information spread in a human network, cascades of firing neurons in cortex, viruses in a computer network, forest fire etc [18, 23]. The model that we consider in this paper belongs to the class of chain-binomial Markov models [6, 37]. Two types of approximation are frequently used for models of this type, the first one is a branching process approximation for cascades of a small size at the beginning of the process

*Occidental Petroleum Corporation, Houston, TX 77046, USA; e-mail: reza_rastegar2@oxy.com
†Dept. of Statistics, Texas A&M University, College Station, TX 77843; e-mail: alexander@stat.tamu.edu

[2, 7, 12, 15, 18, 29, 32], and the second one is an approximation by a deterministic dynamical system once the avalanche spreads to a significant fraction of a large network [5, 10, 22, 27].

The approximations link the asymptotic behavior of immensely complex stochastic processes on a large network to relatively well understood mathematical objects, allowing to gain a qualitative insight into statistical proprieties of the network model. Each of the two approximation processes is in a rigorous sense a limit of the original model in a certain regime. The interpretation of the link between the original model and the approximations is not trivial because some essential features are not preserved when the limit is taken. For instance, the branching approximation for the model studied in this paper is in essence a linearization eliminating the dependence between the nodes (cf. [23]), it is monotone in all basic parameters while the original model is not. Likewise, while the Markov chain describing the evolution of the avalanche magnitude (its transition kernel is specified in (2) below) converges to zero with probability one for any set of parameters, there is a regime in which the approximating dynamical system converges to its non-zero global stable point (see Section 6 below). Usually, the relation between network models and their approximations is studied using either heuristic arguments or numeric computations.

In this paper we focus on a rigorous analytical comparison between the avalanche model and the above mentioned limits, including rates of convergence and rigorous bounds for common characteristics of the model. Loosely speaking, some of our results can be viewed as a second order correction to the branching approximation. In order to make the comparison between the avalanche model and its branching approximation we use coupling constructions based on canonical schemes of stochastic coupling of binomial and Poisson random variables (see Sections 3 and 5 below). Most of our results are new, some complement the results obtained in [23] through heuristic perturbation arguments.

Typically, cascading models exhibit a phase transition between a subcritical regime characterized by a short duration and small size of the avalanche and a supercritical one characterized by long lasting avalanches that eventually affect a non-zero fraction of the network before disappear. It is often argued that regimes near the criticality exhibit the most rich and advantageous for the network performance behavior. See, for instance, recent surveys [13, 30] and references therein. The phenomenon of criticality is of a special interest also because of a universal nature of the phenomenon as well as mathematical challenges in its study. For an interesting discussion of the relation between branching processes and self-organized criticality see [20, 38]. The model that we investigate in this paper is lacking a trivial monotonicity in parameters, but nevertheless exhibits the phase transition with a distinct critical set of parameters.

We proceed with a formal definition of the avalanche model considered in this paper. Fix an integer $n \geq 3$ and denote $V_n = \{1, \ldots, n\}$. The set V_n models the nodes of a network with n nodes. Let $\Omega_n := 2^{V_n}$ be the space of subsets of V_n, and consider the following *avalanche process* $(A_k)_{k \in \mathbb{Z}_+}$ on Ω_n. Here and henceforth $\mathbb{Z}_+$ denotes the set of non-negative integers. Assume that the initial state $A_0 \in \Omega_n \backslash \{\emptyset, V_n\}$ is neither an empty set nor the whole network. Let $p \in (0, 1)$ be given, and $q = 1 - p$. Formally, the sequence A_k is a discrete-time Markov chain in the state space Ω_n with transition kernel given by

$$P(A_{k+1} = B \,|\, A_k = A) = \begin{cases} (1 - q^{|A|})^{|B|} \cdot (q^{|A|})^{n-|A|-|B|} & \text{if } B \subset A^c \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

where A and B are arbitrary elements of Ω_n, A^c denotes the complement of the set A in V_n, and $|A|$ denotes the cardinality of A.

We refer to A_k and A_k^c as, respectively, the *excited* and *resting* states at time k. The interpretation is that an excited state turns into rested in the next instance of time, but before that can excite any of the resting states, each one with probability p. Thus the (conditional, given the excited nodes A_k) probability that a node $x \in A_k^c$ will *not* become excited in the next iteration is equal to $q^{|A_k|}$. We further assume that the excitement mechanisms of different resting nodes are independent each of other at any given instant of time, and hence the product on the right-hand side of (1).

The model is a particular case of the avalanche model introduced in [23], where the "excitant probability" is p, uniformly across all links in the network. Formally, the model described by (1) coincides with the model of the spread of an endemic infection introduced in [27]. In contrast to [27], we concentrate in this paper on the case when p is $O(1/n)$ rather than $O(1)$ for large n. More realistic versions of excitable networks are considered, for instance, in [21] and [24, 25]. We remark that our proof methods can be partially extended and applied to more complex networks, cf. [35].

Let $\{G_k(n,p) : k \in \mathbb{Z}_+\}$ be an i. i. d. sequence of Erdős-Rényi graphs with percolation parameter p, sharing V_n as the common vertex set. An equivalent, dynamic graph viewpoint on the avalanche model is that every node excited at time k excites with certainty all its resting neighbors in the random graph $G_k(n,p)$. Though this observation is not used anywhere in the paper, it immediately provides some heuristic insights into the behavior of the avalanche model when n is large. For instance, if $p = c/n$ then with overwhelming probability, when $c \in (0,1)$ all connected components of the Erdős-Rényi graph are of order $O(\log n)$, while if $c > 1$ there is a connected component of the size $O(n)$ [9]. Heuristically, an implication for the avalanche model that one may expect, is that the avalanche starting on a single node has a little chance to spread over a non-zero fraction of the network in the former case whereas the probability of such an avalanche to eventually reach the size of order $O(n)$ is non-zero in the latter regime. This heuristic observation is formally confirmed in Section 4. More generally, the dynamic graph viewpoint might 1) serve as an indication of the existence of a phase transition in terms of the asymptotic behavior of the model on the scale $p = c/n$ at $c = 1$, and 2) be perceived as a fundamental reason behind the phase transition.

In this paper we focus on the Markov chain $(X_t)_{t\in\mathbb{Z}_+}$, where $X_t = |A_t|$. Transition kernel of this Markov chain is as follows:

$$P(X_{k+1} = j | X_k = i) = \binom{n-i}{j} (1-q^i)^j (q^i)^{n-i-j}, \qquad i = 0,1,\ldots,n, \quad j \le n-i, \tag{2}$$

with the convention that $\binom{0}{0} = 1$ and $\binom{0}{j} = 0$ for all $i \in \mathbb{N}$.

Let T and S denote, respectively, the duration and the total size of the avalanche:

$$T = \inf\{k \in \mathbb{Z}_+ : X_k = 0\} \qquad \text{and} \qquad S = \sum_{k=0}^{T} X_k, \tag{3}$$

with the usual convention that $\inf \emptyset = +\infty$. Remark that zero is the unique absorbing state and $P(T < \infty) = 1$ regardless of the initial state X_0. Both the quantities are arguably among the most important general characteristics of avalanches in a complex network. In Section 5 we study the distribution function of avalanche duration T. The asymptotic behavior of S for large values of n is the content of Theorem 3.4 in Section 3.

The organization of the paper is as follows. In Section 2 we obtain first estimates for the expected value of several fundamental characteristics of the avalanches, such as the current size X_k, current heterogeneity $X_k(n-X_k)$, and the total size S. Some of these basic estimates are subsequently refined and improved. In Section 3 we introduce the key technical tool of our study, a branching process to which the avalanche chain converges weakly as n tends to infinity, provided that the initial value X_0 is kept fixed and the "intensity factor" pn converges to a positive limit. We use the branching approximation to study the total size of the avalanche in the subcritical regime. In particular, Theorem 3.4 gives the rate of convergence to the limiting value as n goes to infinity. In Section 4 we address the question whether an avalanche that started on a few initially excited nodes can propagate to a non-zero fraction of the network (asymptotically, when n is large). In particular, Theorem 4.2 provides lower and upper bounds with a qualitatively matching asymptotic behavior for this probability for a given network size n. In Section 5 we study duration of the avalanche using branching approximation and its natural coupling with the avalanche chain. In Section 6 we are concerned with an approximation of the avalanche chain by a deterministic dynamical system. While the branching approximation is adequate as long as $X_k \ll n$, the deterministic approximation is suitable when $X_k/n = O(1)$.

2 Basic estimates for a subcritical network

The aim of this section is to give bounds on the expected number of excited nodes at a given time, total size of the avalanche, and the probability that a given node is excited at a fixed time $k \in \mathbb{N}$. Throughout the section we consider a single network, that is we assume that $n \in \mathbb{N}$ and $p \in (0,1)$ in (2) are fixed. Most of the results here are of an auxiliary nature, but some, in particular Propositions 2.4, 2.5, and 2.8, appear to be of independent interest. Propositions 2.4 is concerned with the evolution of a measure of heterogeneity for the network, Proposition 2.5 gives a uniform on V_n upper bound on the probability that a given node $x \in V_n$ is excited at time k, and Proposition 2.8 gives tight lower and upper bounds for the expected total size of the avalanche in the subcritical regime. For future convenience, we formulate our results in terms of arbitrary bounds $c > 0$ and $d > 0$ that satisfy the following condition:

Condition 2.1. $0 < d \le pn \le c$.

The next series of propositions is formulated for an arbitrary $c > 0$, even though the results are primarily useful in the case when $c \in (0,1)$. As we will see in the next section, this case exactly corresponds to a subcritical regime of the avalanche model, which turns out to be p satisfying the condition $pn < 1$.

Let $(\mathcal{F}_k)_{k\in\mathbb{Z}_+}$ be the natural filtration for the sequence $(X_k)_{k\in\mathbb{Z}_+}$, that is $\mathcal{F}_k$ is the σ-algebra generated by $X_0, \ldots, X_k$. It follows from (2) that with probability one,

$$E(X_{k+1} \,|\, \mathcal{F}_k) \le (n - X_k)\Big[1 - \Big(1 - \frac{c}{n}\Big)^{X_k}\Big] \le cX_k \frac{n - X_k}{n} \le cX_k. \tag{4}$$

Notice that the formula remains formally true when $X_k = 0$. Thus, we have:

Proposition 2.2. *Let Condition 2.1 hold. Then the sequence* $(X_k c^{-k})_{k\in\mathbb{Z}_+}$ *is a supermartingale with respect to the filtration* $(\mathcal{F}_k)_{k\in\mathbb{Z}_+}$.

This straightforward estimate will be improved for the entire range of the network parameters (np is less, greater, or equal to 1) in Section 6 below.

For $k \in \mathbb{Z}_+$, let

$$H_k = X_k(n - X_k). \tag{5}$$

The following is immediate from (4).

Corollary 2.3. *Let Condition 2.1 hold. Then* $E(X_k) \le \frac{c^k}{n}E(H_0) \le \frac{c^k}{n}E(X_0)$ *for all* $k \in \mathbb{N}$.

For $k \in \mathbb{Z}_+$ and $x \in V_n$, let

$$\mathcal{E}_k(x) = \begin{cases} 1 & \text{if } x \in A_k \\ 0 & \text{if } x \notin A_k \end{cases} \tag{6}$$

be the indicator of the event that a given node $x \in V_n$ is excited at time $k \in \mathbb{Z}_+$. Remark that $H_k = X_k(n-X_k)$ can be interpreted as a (non-normalized) measure of the heterogeneity of the network because

$$H_k = X_k(n - X_k) = \Big(\sum_{x\in V_n} \mathcal{E}_k(x)\Big) \cdot \Big(\sum_{x\in V_n} (1 - \mathcal{E}_k(x))\Big) = \sum_{x,y\in V_n} \mathcal{E}_k(x)(1 - \mathcal{E}_k(y)),$$

and thus $\frac{2H_k}{n(n-1)}$ is the probability that two nodes randomly chosen at time k are not in the same state. It turns out (see Section 6 in this paper) that in the subcritical case $np \in (0, 1)$ (and in fact also in the critical case $np = 1$), for large values of n, loosely speaking, the number of excited nodes decreases to zero almost monotonically. This observation motivates the following result, which in particular implies that the expected heterogeneity decreases to zero monotonically in the subcritical regime (see also Corollary 6.5 below).

Proposition 2.4. *Let Condition 2.1 hold. Then the sequence* $(H_k c^{-k})_{k\in\mathbb{Z}_+}$ *is a supermartingale with respect to the filtration* $(\mathcal{F}_k)_{k\in\mathbb{Z}_+}$.

Proof. It follows from (2) that

$$E(X_{k+1}^2|X_k) \quad = \quad (n - X_k)(1 - q^{X_k})q^{X_k} + (n - X_k)^2(1 - q^{X_k})^2. \tag{7}$$

Notice that the formula remains true when $X_k = 0$. It follows that

$$\begin{aligned} &E[X_{k+1}(n - X_{k+1})|X_k] = nE(X_{k+1}|X_k) - E(X_{k+1}^2|X_k) \\ &\quad = n(n - X_k)(1 - q^{X_k}) - (n - X_k)(1 - q^{X_k})q^{X_k} - (n - X_k)^2(1 - q^{X_k})^2 \\ &\quad = (n - X_k)(1 - q^{X_k})[n - q^{X_k} - (n - X_k)(1 - q^{X_k})] \\ &\quad = (n - X_k)(1 - q^{X_k})[X_k(1 - q^{X_k}) + (n - 1)q^{X_k}] \\ &\quad \le (n - 1)(n - X_k)(1 - q^{X_k}) \le n(n - X_k)\Big(1 - \Big(1 - \frac{c}{n}\Big)^{X_k}\Big) \\ &\quad \le cX_k(n - X_k), \end{aligned} \tag{8}$$

where in the first inequality we used the fact that $X_k \le n - 1$, and hence

$$X_k(1 - q^{X_k}) + (n - 1)q^{X_k} \le n - 1.$$

The proof of the proposition is complete. □

Recall $\mathcal{E}_k(x)$ from (6). Let

$$\xi_k(x) = P(\mathcal{E}_k(x) = 1), \qquad x \in V_n \tag{9}$$

and

$$\eta_k = P(X_k > 0) = P(T > k), \qquad k \in \mathbb{Z}_+,$$

where T is the duration of the avalanche defined in (3). We have the following:

Proposition 2.5. *Let Condition 2.1 hold. Then $\xi_k(x) \le \frac{c^k}{n^2} E(H_0)$ for all $k \in \mathbb{N}$ and $x \in V_n$.*

The proof of the proposition is given below in this section, after the statement of a corollary. The claim is trivial if the distribution of X_0 is invariant with respect to permutation of nodes (see the proof below), but takes slightly more effort to establish when the symmetry is broken and the nodes cannot be treated as stochastically identical.

We now proceed with the corollary. By virtue of (2),

$$P(X_{k+1} = 0 \,|\, X_k) \ge q^{\frac{n^2}{4}}$$

uniformly on X_k, and hence

$$\eta_k = P(X_k > 0) < (1 - q^{\frac{n^2}{4}})^k \qquad \forall\, k \in \mathbb{N}. \tag{10}$$

The following result is a refinement of this naive estimate for the subcritical regime. By Chebyshev's inequality,

$$\eta_k = P(X_k \ge 1) \le E(X_k) = E\Big(\sum_{x \in V_n} \mathcal{E}_k(x)\Big) = \sum_{x \in V_n} \xi_k(x).$$

This yields

Corollary 2.6. *Let Condition 2.1 hold. Then $\eta_k \le \frac{c^k}{n} E(H_0)$ for all $k \in \mathbb{N}$.*

Remark that the result in the corollary will be further improved in Theorem 5.1 below using a comparison to a branching process and known estimates for the extinction time of the latter. In particular, it turns out that if $np = 1$ in (2), then

$$\eta_k \le 1 - (\frac{k}{k+2})^{i_0} \le \frac{2i_0}{k+2} \qquad \forall\, k \in \mathbb{N},$$

and the bound is asymptotically tight (see the lower bound in Theorem 5.1 and also Corollary 5.5).

Proof of Proposition 2.5. We say that the distribution of A_0 is *exchangeable* if it is independent of a particular labeling of the network's nodes, that if $P(B \subset A_0) = P(\sigma B \subset A_0)$ for any permutation (one-to-one and onto relabeling) $\sigma : V_n \to V_n$ and a cluster of nodes $B \subset V_n$.

Fix any $x \in V_n$. First, observe that if the distribution of A_0 is exchangeable, then

$$\xi_k(x) = E(\mathcal{E}_k(x)) = \frac{1}{n} E\Big(\sum_{x \in V_n} \mathcal{E}_k(x)\Big) = \frac{1}{n} E(X_k) \leq \frac{c^k}{n} E(H_0)$$

by virtue of Proposition 2.2.

Consider now an auxiliary avalanche process where A_0 is chosen uniformly over subsets of V_n of a given size m. We will denote the conditional law of the process $P(\,\cdot\,|A_0)$ by $P_{m,x}$ when $x \in A_0$ and by $P_{m,\overline{x}}$ otherwise. Then, in view of the result for exchangeable A_0 and Proposition 2.5,

$$\begin{aligned}
\frac{m(n-m)c^k}{n^2} &= \frac{c^k}{n^2} E(H_0) \geq P(\mathcal{E}_k(x) = 1) \\
&= \frac{\binom{n-1}{m}}{\binom{n}{m}} P_{m,\overline{x}}(\mathcal{E}_k(x) = 1) + \frac{\binom{n-1}{m-1}}{\binom{n}{m}} P_{m,x}(\mathcal{E}_k(x) = 1) \\
&\geq \frac{\binom{n-1}{m}}{\binom{n}{m}} P_{m,\overline{x}}(\mathcal{E}_k(x) = 1) = \frac{n-m}{n} P_{m,\overline{x}}(\mathcal{E}_k(x) = 1),
\end{aligned}$$

which implies

$$P_{m,\overline{x}}(\mathcal{E}_k(x) = 1) \leq \frac{mc^k}{n}. \tag{11}$$

Therefore,

$$\begin{aligned}
P_{j,x}(\mathcal{E}_{k+1}(x) = 1) &= \sum_{m=0}^{n-j} \binom{n-j}{m} (1-q^j)^m q^{j(n-j-m)} P_{m,\overline{x}}(\mathcal{E}_k(x) = 1) \\
&\leq c^k \sum_{m=0}^{n-j} \binom{n-j}{m} (1-q^j)^m q^{j(n-j-m)} \frac{m}{n} \\
&= \frac{c^k}{n} E_{j,x}(X_1) = \frac{c^k (n-j)(1-q^j)}{n} \leq \frac{c^{k+1} j(n-j)}{n^2}.
\end{aligned} \tag{12}$$

In particular, we have established that

$$P_{j,x}(\mathcal{E}_k(x) = 1) \leq \frac{c^{k-1}(n-j)(1-q^j)}{n} \leq c^k \frac{j(n-j)}{n^2} \leq \frac{jc^k}{n}. \tag{13}$$

Turning now to $P_{j,\overline{x}}(\mathcal{E}_{k+1}(x)$, write

$$\begin{aligned}
P_{j,\overline{x}}(\mathcal{E}_{k+1}(x) = 1) &= \sum_{m=1}^{n-j-1} \binom{n-j-1}{m} (1-q^j)^m q^{j(n-j-m-1)} P_{m,\overline{x}}(\mathcal{E}_k(x) = 1) \\
&\quad + \sum_{m=1}^{n-j-1} \binom{n-j-1}{m-1} (1-q^j)^{m-1} q^{j(n-j-m)} P_{m,x}(\mathcal{E}_k(x) = 1).
\end{aligned}$$

Using (13) along with (11) , we obtain

$$
\begin{aligned}
P_{j,\bar{x}}(\mathcal{E}_{k+1}(x)) &\le \frac{c^k}{n}\sum_{m=1}^{n-j-1}\binom{n-j-1}{m}(1-q^j)^m q^{j(n-j-m-1)}m \\
&\quad + \frac{c^k}{n}\sum_{m=1}^{n-j-1}\binom{n-j-1}{m-1}(1-q^j)^{m-1}q^{j(n-j-m)}m \\
&= \frac{c^k}{n}E_{j,\bar{x}}(X_1) = \frac{c^k(n-j)(1-q^j)}{n} \le \frac{c^{k+1}j(n-j)}{n^2}. \qquad (14)
\end{aligned}
$$

The claim follows now from (12) and (14). □

We next investigate the total number of excited nodes in a subcritical regime. We will use the following lower bound for $1-q^i$.

Lemma 2.7. *Under Condition 2.1,* $1-q^i \ge \frac{di}{n} - \frac{c^2i^2}{2n^2}$ *for all* $i \in \mathbb{Z}_+$.

Proof. The lemma is trivial for $i = 0, 1$. For $i \ge 2$, using the Lagrange form of the second order remainder in Taylor's series for $f(p) = (1-p)^i$ around zero,

$$
(1-p)^i = 1 - ip + \frac{i(i-1)}{2}p^2(1-p_*)^{i-2} \le 1 - ip + \frac{i(i-1)}{2}p^2 \le 1 - \frac{di}{n} + \frac{c^2i^2}{2n^2}
$$

for some $p_* \in (0,p)$. □

Recall S from (3). We have:

Proposition 2.8. *Suppose that Condition 2.1 holds with* $c \in (0,1)$. *Then,*

$$
\frac{E(X_0)}{1-d} - \frac{3E(X_0^2)}{n(1-c)^3} \le E(S) \le \frac{E(X_0)}{1-c}. \qquad (15)
$$

Proof. To prove the proposition, we will first obtain a suitable lower bound for $E(X_k)$. It follows from Lemma 2.7 that

$$
\begin{aligned}
E(X_{k+1}|X_k) &\ge (n-X_k)\Big(\frac{dX_k}{n} - \frac{c^2X_k^2}{2n^2}\Big) = dX_k - \frac{c^2X_k^2}{2n} - \frac{dX_k^2}{n} + \frac{c^2X_k^3}{2n^2} \\
&\ge dX_k - \frac{3cX_k^2}{2n}, \qquad (16)
\end{aligned}
$$

where we used the fact that $d < c$ and $c^2 < c$. Using the identity in (7), we obtain

$$
\begin{aligned}
E(X_k^2) &= E[(n-X_{k-1})(1-q^{X_{k-1}})q^{X_{k-1}} + n^2(1-q^{X_{k-1}})^2] \\
&\le E[n(1-q^{X_{k-1}}) + n^2(1-q^{X_{k-1}})^2] \\
&\le E\Big[n\cdot\frac{cX_{k-1}}{n} + n^2\frac{c^2X_{k-1}^2}{n^2}\Big] = E(cX_{k-1} + c^2X_{k-1}^2). \qquad (17)
\end{aligned}
$$

Iterating,

$$\begin{aligned} E(X_k^2) &\le E(cX_{k-1}+c^2X_{k-1}^2) \le E(cX_{k-1}+c^3X_{k-2}+c^4X_{k-2}^2) \\ &\le c^{2k}E(X_0^2)+\sum_{j=1}^{k} c^{2j-1}E(X_{k-j}) \le c^{2k}E(X_0^2)+\sum_{j=1}^{k} c^{k+j-1}E(X_0) \\ &\le c^{2k}E(X_0^2)+\frac{c^k}{1-c}E(X_0). \end{aligned}$$

Therefore,

$$E(X_{k+1}) \ge dE(X_k)-\frac{3c}{2n}E\Big[c^{2k}E(X_0^2)+\frac{c^k}{1-c}E(X_0)\Big].$$

Iterating again, we obtain

$$\begin{aligned} E(X_{k+1}) &\ge d^{k+1}E(X_0)-\frac{3c}{2n}\sum_{j=0}^{k}\Big[c^{2(k-j)}d^jE(X_0^2)+\frac{c^{k-j}d^j}{1-c}E(X_0)\Big] \\ &\ge d^{k+1}E(X_0)-\frac{3c}{2n}\sum_{j=0}^{k}\Big[c^{2k+j}E(X_0^2)+\frac{c^k}{1-c}E(X_0)\Big] \\ &\ge d^{t+1}E(X_0)-\frac{3c^{k+1}}{2n(1-c)}[c^kE(X_0^2)+(k+1)E(X_0)]. \end{aligned}$$

Using this bound along with the upper bound in Corollary 2.3, we obtain that

$$d^kE(X_0)-\frac{3c^k}{2n(1-c)}[c^kE(X_0^2)+kE(X_0)] \le E(X_k) \le c^kE(X_0).$$

Therefore, summing over all indexes from zero to k,

$$\frac{E(X_0)}{1-d}-\frac{3}{2n(1-c)}\Big[\frac{1}{1-c^2}E(X_0^2)+\frac{c}{(1-c)^2}E(X_0)\Big]. \le E\Big(\sum_{k=0}^{\infty}X_k\Big) \le \frac{E(X_0)}{1-c}$$

Taking in account that $c<1$, $1-c^2<(1-c)^2$, and $E(X_0)\le E(X_0^2)$, we obtain the lower bound in the form given in the statement of the proposition. □

We remark that though the constant $\frac{3E(X_0^2)}{(1-c)^3}$ in front of $1/n$ in the correction term at the left-hand side of (15) is not optimal, the lower bound captures correctly the dependence of this term on n. The latter result is formally stated in part (iii) of Theorem 3.4 below.

3 Poisson approximation and the size of the avalanche

The primary goal of this section is to study the asymptotic behavior of the total size of the avalanche for a certain ensemble of comparable avalanche models. The underlying family of models is introduced in equation (18) and Assumption 3.1 below, and the main result of this section is stated in Theorem 3.4. The secondary purpose of this section is to introduce a branching process approximation which will be used throughout the rest of the paper.

In the rest of the paper, along with a single network X, we will often consider a family of Markov chains $X^{(n)} = (X_k^{(n)})_{k\in\mathbb{Z}_+}$, each governed by a transition kernel of the same type as in (2), namely

$$P_n(i,j) := P(X_{k+1}^{(n)} = j | X_k^{(n)} = i) = \binom{n-i}{j}(1-q_n^i)^j (q_n^i)^{n-i-j}, \tag{18}$$

for some $q_n \in (0,1)$ and all $i = 0, 1, \dots, n,\ j \le n-i$. In this definition we maintain the convention that $\binom{0}{0} = 1$ and $\binom{0}{j} = 0$ for all $i \in \mathbb{N}$ in (18). Therefore, all Markov chains in this collection eventually absorb at zero. Typically we will impose the following comparability assumption on the family of avalanche models under consideration:

Assumption 3.1.

(i) There exists $\lambda > 0$ such that $\lim_{n\to\infty} np_n = \lambda$, where $p_n = 1 - q_n$.

(ii) All $X^{(n)}$ have the same initial state, namely $X_0^{(n)} = i_0$ for some $i_0 \in \mathbb{N}$ and all $n \in \mathbb{N}$.

Some of our asymptotic estimates will be stated in terms of arbitrary numbers $c > 0$, $d > 0$, and $i_0 \in \mathbb{N}$ that satisfying the following condition. This condition is an analogue of Condition 2.1 for a family of networks that satisfies Assumption 3.1.

Assumption 3.2. *Let $p_n \in (0,1)$, $n \in \mathbb{N}$, be given and consider a family of avalanche models $\{X^{(n)} : n \in \mathbb{N}\}$ with transition kernels defined in (18). Assume that part (ii) of Assumption 3.1 is in force and, furthermore, there exist constants $c > 0$, $d \in (0, c)$, and $n_0 \in \mathbb{N}$ such that*

(i) $np_n \in [c, d]$ for all $n \ge n_0$.

(ii) If part (i) of Assumption 3.1 holds and $\lambda > 1$, then $d > 1$.

(iii) If part (i) of Assumption 3.1 holds and $\lambda < 1$, then $c < 1$.

It follows from (2) that under Assumption 3.1, for any $i \in \mathbb{N}$, $j \in \mathbb{Z}_+$, and $k \in \mathbb{N}$,

$$\lim_{n\to\infty} P(X_{k+1}^{(n)} = j | X_k^{(n)} = i) = e^{-\lambda i}\frac{(\lambda i)^j}{j!}. \tag{19}$$

Let $Z^{(\lambda)} = (Z_k^{(\lambda)})_{k\in\mathbb{Z}_+}$ be a Markov chain on $\mathbb{Z}_+$ with absorption state at zero, Poisson transition kernel

$$P(Z_{k+1}^{(\lambda)} = j | Z_k^{(\lambda)} = i) = e^{-\lambda i}\frac{(\lambda i)^j}{j!}, \qquad i \in \mathbb{N},\ j \in \mathbb{Z}_+,$$

and the same initial state $Z_0^{(\lambda)} = i_0$, the same as for all $X^{(n)}$. We can assume without loss of generality that $Z^{(\lambda)}$ is a Galton-Watson branching process with a Poisson offspring distribution, namely

$$Z_{k+1}^{(\lambda)} = \sum_{j=1}^{Z_k^{(\lambda)}} Y_{k,j}^{(\lambda)} \tag{20}$$

for a collection of independent Poisson random variables $Y = \{Y_{k,j}^{(\mu)} : k \in \mathbb{Z}_+, j \in \mathbb{N}, \mu > 0\}$ such that for all $i \in \mathbb{Z}_+$,

$$P(Y_{k,j}^{(\mu)} = i) = e^{-\mu} \frac{\mu^i}{i!}.$$

The sum in the right-hand side of (20) is assumed to be zero if $Z_k^{(\lambda)} = 0$, that is $Z_k^{(\lambda)}$ is formally defined for all $k \in \mathbb{Z}_+$.

The convergence in (19) implies the weak convergence of the sequence of Markov processes $X^{(n)}$ to the branching process $Z^{(\lambda)}$ as $n \to \infty$ [19]. To illustrate the functionality of the branching approximation, Fig. (1) and (2) below provide plots of $E(T)$ as a function of the initial state i_0 for $n = 100$ and $n = 1000$, in each case for four values of the parameter $c = np$ concentrated around the theoretical phase transition value $c = 1$ suggested by the branching approximation. Note that $E(T) < \infty$ by virtue of (10).

Let Q be $(n-1) \times (n-1)$ matrix with entries $Q(i, j) = P(X_{k+1} = j \,|\, X_k = i)$. To evaluate the expectation we use the following standard Markov chain matrix calculation:

$$E(T) = \sum_{m=0}^{\infty} P(T > m) = \sum_{m=0}^{\infty} Q^m e = (I - Q)^{-1} e,$$

where $e \in \mathbb{R}^{n-1}$ is an $(n-1)$-vector with all entries equal to one and I is the $(n-1)$-dimensional unit matrix. To compute the inverse matrix in the above expression we used the packages "numpy" and "decimal" on Python 3.5 with the computation precision set to 400 decimal points.

An intuitive reason for the uniformly (on i_0) large values of $E(T)$ and high persistence of the avalanche in the supercrtical regime, when n is large, is that the stochastic path of the Markov chain X is well approximated by a trajectory of a deterministic dynamical system that is locally Lipschitz, and consequently is quickly attracted to its unique (non-zero) global stable point (see Section 6 below for details). Heuristically, it appears that the Markov chain spends most of its time before the absorbtion being "trapped" in a neighborhood of the stable point. The numerical simulations show that the phase transition in the avalanche model doesn't occur at exactly $c = 1$ for either $n = 100$ or 1000. While the phase transition is fairly smooth for $n = 100$, it is considerably more sharp and conspicuous for $n = 1000$. Overall, one can conclude that the branching approximation gives a useful qualitative insight into the existence of an asymptotic phase transition in the avalanche model.

In what follows we will exploit the following explicit monotone coupling of $X^{(n)}$ with a branching process. For future convenience, we state the result in terms of a family of avalanche models rather than a single network. At the base of the construction is a standard coupling between a binomial $B(n, p)$ and a Poisson$(-n \log(1-p))$ random variables.

Proposition 3.3. *Let Assumption 3.1 hold. Then for every $n \geq n_0$ there exists a Markov chain $(X_k^{(c,n)}, Z_k^{(c,n)})_{k \in \mathbb{Z}_+}$ on $\mathbb{Z}_+^2$ such that the following holds true:*

(i) $(X_k^{(c,n)})_{k \in \mathbb{Z}_+}$ is distributed the same as $(X_k^{(n)})_{k \in \mathbb{Z}_+}$.

(ii) $(Z_k^{(c,n)})_{k \in \mathbb{Z}_+}$ is distributed the same as $(Z_k^{(c)})_{k \in \mathbb{Z}_+}$.

(iii) With probability one, $X_0^{(c,n)} = Z_0^{(c,n)}$ and $X_k^{(c,n)} \leq Z_k^{(c,n)}$ for all $k \in \mathbb{N}$.

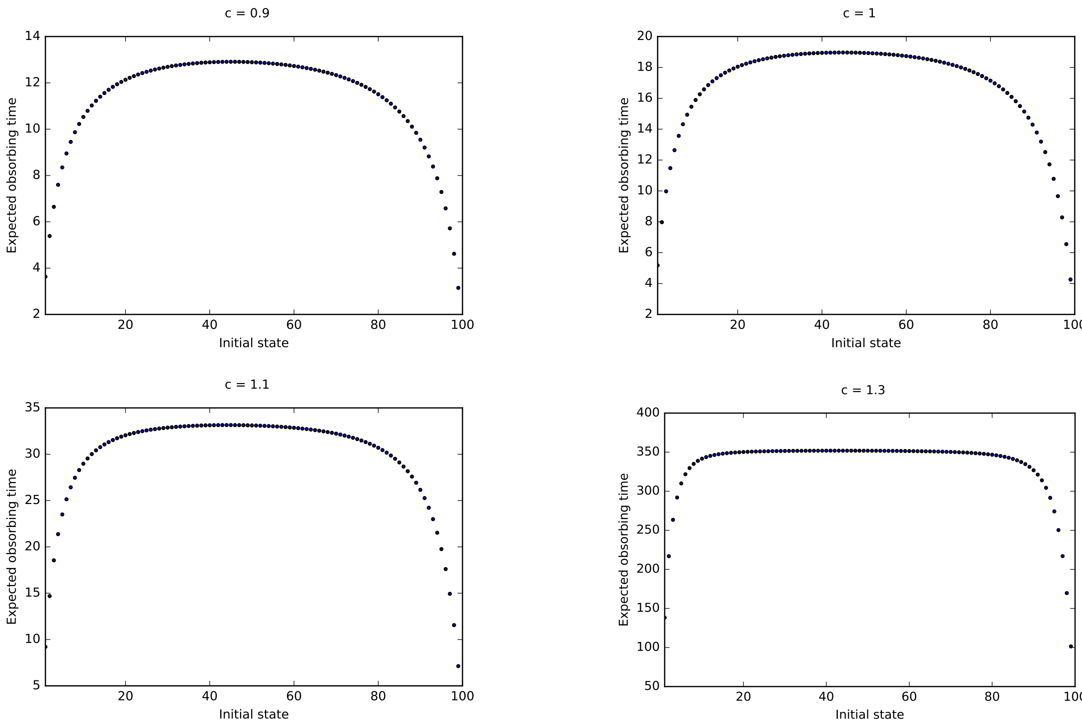

Figure 1: Plot of the function $f(i_0) = E(T \mid X_0 = i_0)$ for $n = 100$ and several values of the parameter $c = np$ ranging from $c = 0.9$ to $c = 1.3$.

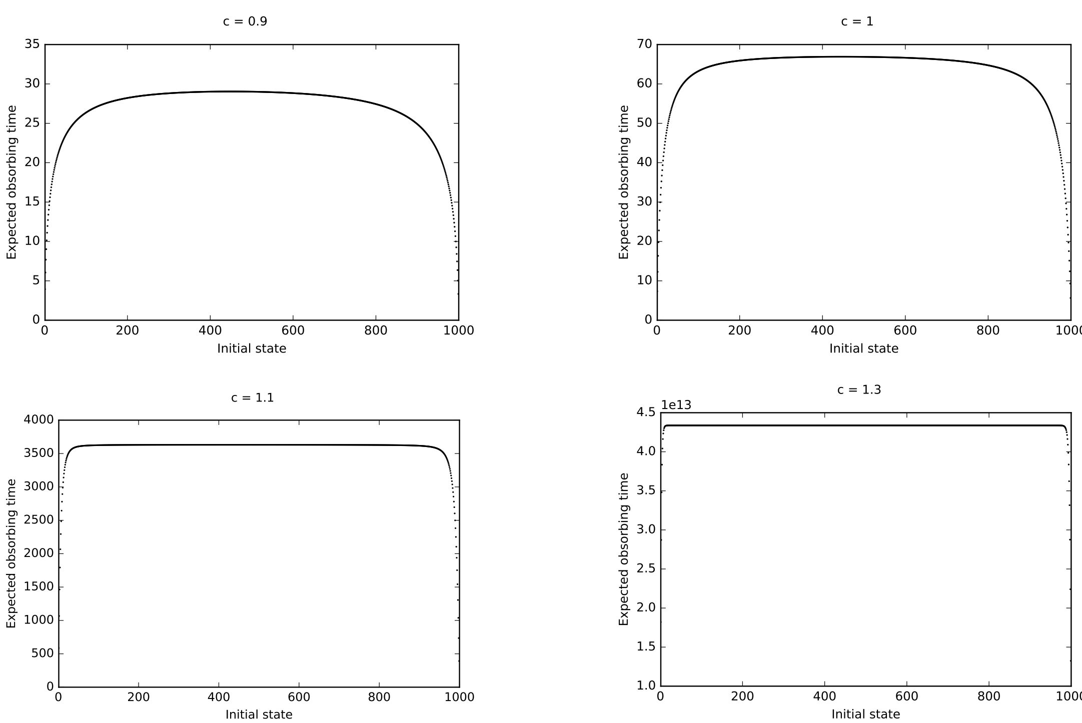

Figure 2: Plot of the function $f(i_0) = E(T \mid X_0 = i_0)$ for $n = 1000$ and several values of the parameter $c = np$ ranging from $c = 0.9$ to $c =$ ‘1.3.

Proof of Proposition 3.3. We have

$$-\log q_n = -\log(1-p_n) \le -\log\Big(1-\frac{c}{n}\Big), \qquad \forall\ n \ge n_0. \tag{21}$$

It is easy to check that for any $n \ge n_0$ and $i \in \mathbb{N}$,

$$-\log q_n \le \frac{c}{n-1} \le \frac{c}{n-i}. \tag{22}$$

Indeed, if $f(x) = \frac{c}{x-1} + \log\Big(1-\frac{c}{x}\Big)$, then $\lim_{x\to+\infty} f(x) = 0$ and for any $x > 1$,

$$f'(x) = -\frac{c}{(x-1)^2} + \frac{c}{x^2-cx} \le -\frac{c}{(x-1)^2} + \frac{c}{x^2-x} = \frac{c}{x-1}\Big[-\frac{1}{x-1}+\frac{1}{x}\Big] < 0,$$

implying that $f(x) > 0$ for $x > 1$. Let $\{Y_{n,k,j}^{(c,x,z)} : n \in \mathbb{N}, k \in \mathbb{Z}_+, j \in \mathbb{N}, x \in \mathbb{Z}_+, z \in \mathbb{Z}_+, x < n\}$ be a collection of independent Poisson random variables such that

$$P(Y_{n,k,j}^{(c,x,z)} = i) = e^{-\frac{cz}{n-x}}\frac{\big(\frac{cz}{n-x}\big)^i}{i!}, \qquad i \in \mathbb{Z}_+.$$

Further, let $U = \{U_{n,k,j}^{(c,x)} : n \in \mathbb{N}, k \in \mathbb{Z}_+, j \in \mathbb{N}, x \in \mathbb{Z}_+, x < n\}$ be a collection of independent Bernoulli variables which is independent of the family of Poisson variables Y and such that

$$P(U_{n,k,j}^{(c,x)} = 1) = \frac{1-q_n^x}{1-e^{-\frac{cx}{n-x}}} \qquad \text{and} \qquad P(U_{n,k,j}^{(c,n)} = 0) = \frac{q_n^x - e^{-\frac{cx}{n-x}}}{1-e^{-\frac{cx}{n-x}}}.$$

Finally, set

$$B_{n,k,j}^{(c,x)} = U_{n,k,j}^{(c,n)} \mathbf{1}_{\{Y_{k,j}^{(c,x,x)}>0\}}$$

and define a Markov chain of integer triples $(X_k^{(c,n)}, Q_k^{(c,n)}, Z_k^{(c,n)})_{k\in\mathbb{Z}_+}$ through the initial condition $Z_0^{(c,n)} = Q_0^{(c,n)} = X_k^{(c,n)} = i_0$ and the recursion

$$\begin{cases} Z_{k+1}^{(c,n)} &= \sum_{j=1}^{n-x}(Y_{n,k,j}^{(c,x,x)} + Y_{n,k,j}^{(c,x,z-x)}) \\ \\ Q_{k+1}^{(c,n)} &= \sum_{j=1}^{n-x} Y_{n,k,j}^{(c,x,x)} \qquad\qquad\qquad\qquad \text{if} \quad X_k^{(c,n)} = x,\ Z_k^{(c,n)} = z. \\ \\ X_{k+1}^{(c,n)} &= \sum_{j=1}^{n-x} B_{n,k,j}^{(c,x)} \end{cases} \tag{23}$$

By induction, $P(X_k^{(c,n)} \le Q_k^{(c,n)} \le Z_k^{(c,n)}) = 1$ for all $k \in \mathbb{N}$. Furthermore, by our construction $(X_k^{(c,n)})_{k\in\mathbb{Z}_+}$ is distributed the same as $(X_k^{(n)})_{k\in\mathbb{Z}_+}$ while $(Z_k^{(c,n)})_{k\in\mathbb{Z}_+}$ is distributed the same as the branching process $(Z_k^{(c)})_{k\in\mathbb{Z}_+}$. □

Our next result concerns the total size of the avalanche, namely the total number of excited sites created by the avalanche during its entire life span. Let

$$S_n = \sum_{k=0}^{\infty} X_k^{(n)}.$$

Note that $P(S_n < \infty) = 1$ since $X^{(n)}$ is an irreducible Markov chain with a unique absorbing state at zero. The following theorem complements the bounds provided by Proposition 2.8 for a single network with a fixed $n \in \mathbb{N}$. The theorem relates asymptotic characteristics of $X^{(n)}$ to their counterparts for the limiting branching process $Z^{(\lambda)}$.

Theorem 3.4. *Let Assumption 3.1 hold. Then*

(i)

$$\lim_{n\to\infty} E(S_n) = \begin{cases} \frac{i_0}{1-\lambda} & \text{if } \lambda \in (0,1) \\ +\infty & \text{if } \lambda \geq 1. \end{cases}$$

(ii) If $\lambda \leq 1$, S_n converges in distribution as $n \to \infty$ to the Borel-Tanner distribution with parameters i_0 and λ, that is

$$\lim_{n\to\infty} P(S_n = j) = \frac{i_0}{j}\frac{(\lambda j)^{j-i_o}}{(j-i_0)!}e^{-\lambda j}, \qquad j \geq i_0. \tag{24}$$

(iii) If $\lambda > 1$, then for any $m \geq i_0$,

$$\lim_{n\to\infty} P(S_n > m) = 1 - \alpha_\lambda + \sum_{j=m+1}^{\infty} \frac{i_0}{j}\frac{(\lambda j)^{j-i_o}}{(j-i_0)!}e^{-\lambda j},$$

where α_λ is the extinction probability of the branching process $Z^{(\lambda)}$, that is the unique in $(0,1)$ root of the fixed point equation $\alpha_\lambda = e^{-(1-\alpha_\lambda)\lambda}$.

(iv) If $\lambda < 1$,

$$\lim_{n\to\infty} n\Big[\frac{i_0}{1-\lambda} - E(S_n)\Big] = \frac{3i_0\lambda^2}{2(1-\lambda)^2(1+\lambda)} + \frac{i_0^2(2\lambda-\lambda^2)}{2(1-\lambda)^2(1+\lambda)}.$$

For $\mu > 0$, let $S^{(\mu)} = \sum_{k=0}^{\infty} Z_k^{(\mu)}$. By the Otter-Dwass theorem, the limiting distribution in (24) is the distribution of $S^{(\lambda)}$ [16]. Similarly to other models using an approximation by the Poisson branching process, the distribution tails of the total size of the underlying population (avalanche in our case) at the critical regime obey a power law. Indeed, (24) and Stirling's formula implies that when $\lambda = 1$, for large values of j and n, $P(S_n = j)$ is well-approximated by $\frac{i_0}{j}\frac{(\lambda j)^{j-i_o}}{(j-i_0)!}e^{-\lambda j} \sim i_0\sqrt{\frac{1}{2\pi}}j^{-3/2}$.

Proof of Theorem 3.4.
(i) If $\lambda < 1$ the result in (i) follows from Proposition 2.8. If $\lambda \geq 1$, then a version of Fatou's lemma for weakly convergent sequences implies that for any $J \in \mathbb{N}$,

$$\liminf_{n\to\infty} E\Big(\sum_{k=0}^{J} X_k^{(n)}\Big) \geq E\Big(\sum_{k=0}^{J} Z_k^{(\lambda)}\Big) = i_0 \sum_{k=0}^{J} \lambda^k,$$

and the result follows by taking J to infinity. We remark in passing that estimates similar to (15) show that in fact $\lim_{n\to\infty} E\Big(\sum_{k=0}^{J} X_k^{(n)}\Big) = i_0 \sum_{k=0}^{J} \lambda^k$.

(ii) Let $c > 0$ and $d > 0$ be as in Condition 3.2. Assume first that $\lambda \in (0,1)$. A simple argument can be given in order to prove the result in this case. To prove the convergence of S_n to $S^{(\lambda)}$ we will consider exponential generating functions $E(e^{-\alpha S_n})$ and $E(e^{-\alpha S^{(c)}})$, $\alpha > 0$, $c > 0$, and use the inequality $e^{-\alpha x} - e^{-\alpha y} \leq \alpha(y-x)$ which is true for any $y > 0$ and $x \in (0,y)$. It follows from (23) that

$$0 \leq E(e^{-\alpha S_n}) - E(e^{-\alpha S^{(c)}}) \leq \alpha E(S^{(c)} - S_n).$$

By Proposition 2.8, for any $n \geq n_0$,

$$\begin{aligned} 0 &\leq \limsup_{n\to\infty} [E(e^{-\alpha S_n}) - E(e^{-\alpha S^{(c)}})] \\ &\leq \limsup_{n\to\infty} [E(e^{-\alpha S_n}) - E(e^{-\alpha S^{(c)}})] \leq \alpha\Big(\frac{i_0}{1-c} - \frac{i_0}{1-d}\Big), \end{aligned}$$

which yields the result since the parameters c and d can be chosen arbitrarily close to λ.

Assume now that $\lambda = 1$. Without loss of generality we can assume that $c > 1$. Let $A_c = \{\lim_{k\to\infty} Z_k^{(c)} = 0\}$ be the event of extinction for the branching process $(Z_k^{(c)})_{k\in\mathbb{Z}_+}$. It follows from Proposition 3.3 that for any integer $m \geq i_0$,

$$P(S_n > m) \leq P(S^{(c)} > m; A_c) + P(\overline{A}_c) = \sum_{j>m} \frac{i_0}{j} \frac{(cj)^{j-i_o}}{(j-i_0)!} e^{-cj} + P(\overline{A}_c), \tag{25}$$

where $\overline{A}_c$ is the complement event $A_c = \{\lim_{k\to\infty} Z_k^{(c)} = +\infty\}$, and the second identity is an instance of the Otter-Dwass theorem for supercritical branching process, see Theorem 1 in [16]. By letting first n got to infinity and then c approach $\lambda = 1$, we obtain

$$\limsup_{n\to\infty} P(S_n > m) \leq \sum_{j>m} \frac{i_0}{j} \frac{(j)^{j-i_o}}{(j-i_0)!} e^{-j}. \tag{26}$$

On the other hand, Fatou's lemma implies that for any $J \in \mathbb{N}$,

$$\liminf_{n\to\infty} P(S_n > m) \;\; \geq \;\; \liminf_{n\to\infty} P\Big(\sum_{k=0}^{J} X_k^{(n)} > m\Big) \geq P\Big(\sum_{k=0}^{J} Z_k^{(1)} > m\Big). \tag{27}$$

Letting first n and then J go to infinity, we obtain that

$$\liminf_{n\to\infty} P(S_n > m) \geq P(S^{(1)} > m) = \sum_{j>m} \frac{i_0}{j} \frac{(j)^{j-i_o}}{(j-i_0)!} e^{-j}. \tag{28}$$

Combining this estimate with (26) completes the proof of part (ii) for $\lambda = 1$.

(iii) The proof is similar to that of part (ii) for $\lambda = 1$. More precisely, (25) and (27) with $Z^{(1)}$ replaced by $Z^{(\lambda)}$ remain correct for any $m > i_0$, $c > \lambda$, and $J \in \mathbb{N}$. By letting first n go to infinity and then c approach λ in (25), we obtain the following counterpart of (26):

$$\limsup_{n\to\infty} P(S_n > m) \leq \sum_{j>m} \frac{i_0}{j} \frac{(j\lambda)^{j-i_o}}{(j-i_0)!} e^{-j\lambda} + 1 - \alpha_\lambda. \tag{29}$$

By letting first n and then J go to infinity in (28), we obtain from the Otter-Dwass theorem that

$$\liminf_{n\to\infty} P(S_n > m) \geq P(S^{(\lambda)} > m) = \sum_{j>m} \frac{i_0}{j} \frac{(j\lambda)^{j-i_o}}{(j-i_0)!} e^{-j\lambda} + 1 - \alpha_\lambda.$$

Combining this estimate with (29) completes the proof of part (iii) of the theorem.

(iv) Using the Lagrange form of the second order remainder in Taylor's series for the function $f(p) = (1-p)^i$ around zero, we obtain that for all $n \in \mathbb{N}$ and $i \in \mathbb{N}$,

$$q_n^i = (1-p_n)^i = 1 - ip_n + \frac{i(i-1)}{2} p_n^2 (1-\beta_{n,i})^{i-2}$$

for some $\beta_{n,i} \in (0, p_n)$. Therefore,

$$\begin{aligned} E(X_{k+1}^{(n)}) &= E[(n - X_k^{(n)})(1 - q_n^{X_k^{(n)}})] \\ &= np_n E(X_k^{(n)}) + \frac{np_n^2}{2} E[X_k^{(n)}(X_k^{(n)} - 1)] \\ &\quad - p_n E[(X_k^{(n)})^2] + \frac{p_n^2}{2} E[(X_k^{(n)})^2 (X_k^{(n)} - 1)(1 - \beta_{n,X_k^{(n)}})^{X_k^{(n)}-2}]. \end{aligned} \tag{30}$$

It follows from the coupling construction given by Proposition 3.3 that

$$\sum_{k=0}^{\infty} E[(X_k^{(n)})^3] < \sum_{k=0}^{\infty} E[(Z_k^{(c)})^3] < \infty$$

for any $c \in (\lambda, 1)$ and $n \geq n_0$. Hence, summing up both the sides of (30) from $k = 1$ to infinity, we obtain

$$\begin{aligned} E(S_n) &= \frac{i_0}{1-np_n} + \frac{np_n^2}{2(1-np_n)} E\Big[\sum_{k=0}^{\infty} X_k^{(n)}(X_k^{(n)} - 1)\Big] \\ &\quad - \frac{p_n}{1-np_n} E\Big[\sum_{k=0}^{\infty} (X_k^{(n)})^2\Big] + o(1/n). \end{aligned}$$

Therefore, by the dominated convergence theorem (here we use again Proposition 3.3 which shows that $X_k^{(n)}$ is stochastically dominated by $Z_k^{(c)}$),

$$\begin{aligned} \lim_{n\to\infty} n\Big[\frac{i_0}{1-c} - E(S_n)\Big] &= -\frac{\lambda^2}{2(1-\lambda)} E\Big[\sum_{k=0}^{\infty} (Z_k^{(\lambda)})^2 - \frac{i_0}{1-\lambda}\Big] + \frac{\lambda}{1-\lambda} E\Big[\sum_{k=0}^{\infty} (Z_k^{(\lambda)})^2\Big] \\ &= \frac{i_0\lambda^2}{2(1-\lambda)^2} + \frac{2\lambda - \lambda^2}{2(1-\lambda)} E\Big[\sum_{k=0}^{\infty} (Z_k^{(\lambda)})^2\Big]. \end{aligned} \tag{31}$$

Furthermore,

$$E\Big[\sum_{k=0}^{\infty} (Z_k^{(\lambda)})^2\Big] = i_0^2 + E\Big[\sum_{k=1}^{\infty} [\lambda Z_{k-1}^{(\lambda)} + \lambda^2 (Z_{k-1}^{(\lambda)})^2]\Big],$$

which implies that (the sum is finite, for instance, because it is dominated by a finite second moment of the Borel-Tanner distribution of $S^{(\lambda)}$)

$$E\Big[\sum_{k=0}^{\infty}(Z_k^{(\lambda)})^2\Big] = \frac{\lambda i_0}{(1-\lambda)(1-\lambda^2)} + \frac{i_0^2}{1-\lambda^2}.$$

Substituting this identity into (31) yields the result in part (iv) of the theorem. □

4 Spread to a non-zero fraction of the network

In this section we are concerned with the question whether an avalanche initiated by just a few excited nodes has a substantial potential to spread to a large fraction of the network. The results for the supercritical regime are stated in Theorems 4.2 and 4.3, whereas the critical and subcritical regimes are addressed in Theorem 4.4.

First we will consider a single network with given parameters n and p. For an arbitrary real number $J > 0$ we define

$$h_J(i) = P\Big(\max_{k\in\mathbb{Z}_+} X_k \geq J\Big|X_0 = i\Big) = P(X_{T_J} \geq J|X_0 = i), \tag{32}$$

where

$$T_J = \min\{k \in \mathbb{N} : X_k = 0 \text{ or } X_k \geq J\}. \tag{33}$$

We begin with the supercrtical regime, namely the case when $np > 1$. Consequently, without loss of generality we can assume that $d > 1$ in Condition 2.1. Let $\varepsilon_d > 0$ be a positive constant such that

$$\frac{1-e^{-d\varepsilon}}{\varepsilon}(1-\varepsilon) > 1 \qquad \forall\, \varepsilon \in (0, \varepsilon_d]. \tag{34}$$

Note that such ε_d exists because $\lim_{\varepsilon\to 0}\frac{1-e^{-d\varepsilon}}{\varepsilon}(1-\varepsilon) = d > 1$. Further, for $\mu \in (0,\infty)\backslash\{1\}$, let $\alpha_\mu \neq 1$ denote the unique in $(0,\infty)\backslash\{1\}$ solution of the fixed point equation

$$\alpha_\mu = e^{-(1-\alpha_\mu)\mu}, \qquad \alpha_\mu \neq 1. \tag{35}$$

We remark that $\alpha_\mu < \mu\alpha_\mu < 1$ if $\mu > 1$, and $\alpha_\mu > \mu\alpha_\mu > 1$ if $\mu \in (0,1)$. This is true because (35) is equivalent to $\mu\alpha_\mu e^{-\mu\alpha_\mu} = \mu e^{-\mu}$, and the function $f(x) = xe^{-x}$ has a unique local maximum at $x = 1$.

Observe that the right-hand side of (35) is $E\Big(\alpha_\mu^{Z_{k+1}^{(\mu)}}\Big|Z_k^{(\mu)} = 1\Big)$, and hence $\Big(\alpha_\mu^{Z_k^{(\mu)}}\Big)_{k\in\mathbb{Z}_+}$ is a martingale with respect to its natural filtration. If $\mu > 1$, then α_μ is the extinction probability of the supercritical Poisson branching process $Z^{(\mu)}$. Furthermore, if $\mu < 1$ then $1/\alpha_\mu$ is the extinction probability of the dual supercritical process $Z^{(\mu\alpha_\mu)}$.

The main technical result of this section is the following proposition.

Proposition 4.1. *Suppose that Condition 2.1 is satisfied with $d > 1$. Let ε_d be a constant that satisfies condition* (34)*. Then* *(a) There is a constant $\rho = \rho(\varepsilon_d) \in (0,1)$ such that $E(\rho^{X_{k+1}}|X_k = i) \le \rho^i$ for all $i \in [1, n\varepsilon_d)$ and $k \in \mathbb{Z}_+$. Furthermore, $\rho(\varepsilon_d)$ can be chosen in such a way that*

$$\lim_{\varepsilon_d \to 0} \rho(\varepsilon_d) = \alpha_d. \tag{36}$$

(b) $E(\alpha_c^{X_{k+1}}|X_k = i) \ge \alpha_c^i$ for all integers $i \in [1, n)$ and $k \in \mathbb{Z}_+$.

Proof.
(a) First, we choose a real constant $\gamma > 0$ in such a way that

$$q^i \le \left(1 - \frac{d}{n}\right)^i \le 1 - \frac{\gamma i}{n}$$

for any integer $i \in [1, n\varepsilon_d)$. Since

$$q^i \le \left(1 - \frac{d}{n}\right)^i = \left[\left(1 - \frac{d}{n}\right)^n\right]^{\frac{i}{n}} \le e^{-\frac{di}{n}},$$

it suffices to find $\gamma > 0$ such that

$$e^{-dx} \le 1 - \gamma x$$

for any $x \in [0, \varepsilon_d)$. To this end, for a fixed $\gamma > 0$ let $f_\gamma(x) = 1 - \gamma x - e^{-dx}$. Then $f_\gamma(0) = 0$ and

$$f'_\gamma(x) = -\gamma + de^{-dx} \qquad \text{and} \qquad f''_\gamma(x) = -d^2 e^{-dx} < 0.$$

Thus $f_\gamma(x) > 0$ for any $x \in (0, \varepsilon_d)$ provided that

$$f'_\gamma(0) = -\gamma + d > 0 \qquad \text{and} \qquad f_\gamma(\varepsilon_d) = 1 - \gamma\varepsilon_d - e^{-d\varepsilon_d} \ge 0.$$

Since $\frac{1-e^{-d\varepsilon_d}}{\varepsilon_d} < d$, we can put

$$\gamma = \frac{1 - e^{-d\varepsilon_d}}{\varepsilon_d}. \tag{37}$$

Then for any $\rho \in (0,1)$ and integer $i \in [1, n\varepsilon_d)$,

$$\begin{aligned} E(\rho^{X_{k+1}}|X_k = i) &= (\rho(1-q^i) + 1 \cdot q^i)^{n-i} = (\rho + (1-\rho) \cdot q^i)^{n-i} \\ &\le \left(\rho + (1-\rho) \cdot \left(1 - \frac{\gamma i}{n}\right)\right)^{n-i} = \left(1 - (1-\rho)\frac{\gamma i}{n}\right)^{n-i} \\ &\le e^{-(1-\rho)\gamma(1-\varepsilon_d)i}. \end{aligned} \tag{38}$$

Thus, we can set ρ to be the unique solution of the fixed point equation

$$\rho = e^{-(1-\rho)\gamma(1-\varepsilon_d)}. \tag{39}$$

Note that $\rho = \alpha_\mu$ with $\mu = \gamma(1-\varepsilon_d) = \frac{1-e^{-d\varepsilon_d}}{\varepsilon_d}(1-\varepsilon_d)$. The limit result in (36) follows immediately from (39) and (37).

(b) By Proposition 3.3 the process $Z^{(c)}$ stochastically dominates X. Therefore, taking in account that $\alpha_c < 1$. we obtain

$$E(\alpha_c^{X_{k+1}}|X_k = i) \ge E(\alpha_c^{Z_{k+1}^{(c)}}|Z_k^{(c)} = i) = \alpha_c^i.$$

The proof of the proposition is complete. □

Doob's optional stopping theorem implies that for any $\varepsilon \in (0, \varepsilon_d)$ and $i \in [1, n\varepsilon)$,

$$\rho^i \ge E(\rho^{X_{T_{n\varepsilon}}}|X_0 = i) \ge h_{n\varepsilon}(i)\rho^n + (1 - h_{n\varepsilon}(i))$$

and

$$\alpha_c^i \le E(\alpha_c^{X_{T_{n\varepsilon}}}|X_0 = i) \le h_{n\varepsilon}(i)\alpha_c^{n\varepsilon} + (1 - h_{n\varepsilon}(i)).$$

This yields the following result:

Theorem 4.2. *Suppose that Condition 2.1 is satisfied with $d > 1$. Then*

$$\frac{1-\rho^i}{1-\rho^n} \le h_{n\varepsilon}(i) \le \frac{1-\alpha_c^i}{1-\alpha_c^{n\varepsilon}},$$

for all $\varepsilon \in (0, \varepsilon_d)$ and $i \in [1, n\varepsilon)$.

We next consider the asymptotic behavior of the avalanche model when the network size approaches infinity. Similarly to (32) and (33), for any $J > 0$ and $n \in \mathbb{N}$ we define

$$h_J^{(n)}(i) = P\Big(\max_{k\in\mathbb{Z}_+} X_k^{(n)} \ge J \Big| X_0^{(n)} = i\Big) = P(X_{T_J^{(n)}}^{(n)} \ge J | X_0^{(n)} = i), \tag{40}$$

where

$$T_J^{(n)} = \min\{k \in \mathbb{N} : X_k^{(n)} = 0 \text{ or } X_k^{(n)} \ge J\}. \tag{41}$$

Suppose that Assumption 3.1 (and consequently Condition 3.2) hold with $\lambda > 1$. Theorem 4.2 then implies that for any constants $i \in \mathbb{N}$, $\varepsilon \in (0, \varepsilon_d)$, and a function $\psi : \mathbb{N} \to \mathbb{N}$ such that

$$\lim_{n\to\infty} \psi(n) = +\infty \qquad \text{and} \qquad \limsup_{n\to\infty} \frac{\psi(n)}{n} = 0, \tag{42}$$

we have

$$1 - \rho^i \le \liminf_{n\to\infty} h_{\psi(n)}^{(n)}(i) \le \limsup_{n\to\infty} h_{\psi(n)}^{(n)}(i) \le 1 - \alpha_c^i,$$

where $\rho = \rho(\varepsilon)$, as defined in the statement of part (a) of Proposition 4.1. Because of the second condition in (42) we can chose the constant $\varepsilon > 0$ to be as small as we wish. Therefore, by virtue of (36),

$$1 - \alpha_d^i \le \liminf_{n\to\infty} h_{\psi(n)}^{(n)}(i) \le \limsup_{n\to\infty} h_{\psi(n)}^{(n)}(i) \le 1 - \alpha_c^i.$$

Since the constants c and d can be chosen arbitrarily close to λ, we arrive to

$$\lim_{n\to\infty} h_{\psi(n)}^{(n)}(i) = 1 - \alpha_\lambda^i. \tag{43}$$

It turns out that condition (42) can be relaxed as follows:

Theorem 4.3. *Suppose that Assumption 3.1 is satisfied with $\lambda > 1$. Then* (43) *holds for any integer $i \in \mathbb{N}$ and a function $\psi : \mathbb{N} \to \mathbb{N}$ such that $\psi(n) < n$ and $\lim_{n\to\infty} \psi(n) = +\infty$.*

A similar result for a frequency-dependent Wright-Fisher model has been obtained in [12]. The proof of the theorem is very similar to that of Theorem 3.8 in [12], and therefore is omitted. We remark that the proof requires a uniform in n upper bound estimate on $P(X_{k+1}^{(n)} = 0|X_k^{(n)} = m)$ for given $m \in \mathbb{N}$. One can use, for instance, the following bound:

$$P(X_{k+1}^{(n)} = 0|X_k^{(n)} = m) = q^{m(n-m)} \geq \Big(1 - \frac{c}{n}\Big)^{m(n-m)} \geq e^{-2cm}$$

for all n large enough.

We now turn to the study of the maximal number of excited sites in the subcritical and critical regimes.

Theorem 4.4. *Let Assumption 3.1 hold.*

(a) If $\lambda < 1$, then for any integers $i \in \mathbb{N}$ and $m > i$,

$$\limsup_{n\to\infty} h_m^{(n)}(i) \leq \frac{\alpha_\lambda^i - 1}{\alpha_\lambda^m - 1}. \tag{44}$$

(b) If $\lambda = 1$, then for any integers $i \in \mathbb{N}$ and $m > i$,

$$\limsup_{n\to\infty} h_m^{(n)}(i) \leq \frac{i}{m}. \tag{45}$$

Remark 4.5. *Let $M^{(\lambda)} = \max_{k\in\mathbb{Z}_+} Z_k^{(\lambda)}$ and $M_n = \max_{k\in\mathbb{Z}_+} X_k^{(n)}$. The estimates on the right-hand side of* (44) *and* (45) *are classical upper bounds for $P(M^{(\lambda)} \geq m|Z_0^{(\lambda)} = i)$ [26]. Note that the estimates are not trivial in the sense that, in general, $X^{(n)}$ is not dominated by the limiting branching process $Z^{(\lambda)}$ because it is possible that $np_n > \lambda$. However, since both $M^{(\lambda)}$ and M_n are a-priori finite with probability one,* (19) *implies that M_n converges to $M^{(\lambda)}$ in distribution as $n \to \infty$. Hence one can expect that the bounds are meaningful for the avalanche model when n is large. The bound in* (45) *is known to be asymptotically accurate as $m \to \infty$, namely $\lim_{m\to\infty} P(M^{(1)} \geq m|Z_0^{(1)} = i) = \frac{i}{m}$ [26]. For the subcritical process, Theorem $\hat{2}$ in [31] suggests that the correct order of $P(M^{(\lambda)} \geq m|Z_0^{(\lambda)} = i)$ as $m \to \infty$ is $m^{-1}\alpha_\lambda^{-m}$, up to a constant that depends on i.*

Before we prove Theorem 4.4 we state a direct consequence for our model of some well-known results for branching processes in the critical and subcritical regimes. The first part is an implication of a result in [26] mentioned in Remark 4.5, the second part can be derived from a result of [3], and the third one follows from Theorem $\hat{2}$ in [31], all three are based on a comparison to branching process invoking Proposition 3.3. We continue to use the notation for maxima of a branching process introduced in the above remark.

Proposition 4.6. *There exists a sequence of positive constants $\{\varepsilon_m > 0 : m \in \mathbb{N}\}$ such that* $\lim_{m\to\infty} \varepsilon_m = 0$ *and the following holds true:*

(a) If $np = 1$ in (2)*, then*

(i) $P(\max_{k\in\mathbb{Z}_+} X_k > m) \le \frac{i_0}{m}(1+\varepsilon_m)$ *for all integer* $m \in (i_0, n)$.

(ii) $E(\max_{0\le j\le k} X_j) \le \log k(1+\varepsilon_k)$ *for all* $k \in \mathbb{N}$.

(b) If $c := np < 1$ *in* (2)*, then there exists a constant* $B = B(c, i_0) > 0$ *that depends on* c *and* i_0 *only (but not on* n *and* p*) such that* $P(\max_{k\in\mathbb{Z}_+} X_k > m) \le \frac{B(c,i_0)}{m\alpha_c^m}(1+\varepsilon_m)$ *for all integer* $m \in (i_0, n)$.

We remark that we chose the sequence ε_m in the statement of the proposition to be the same in all three cases exclusively for a notational convenience.

We now return to Theorem 4.4.

Proof of Theorem 4.4.
(a) Let $n_0 \in \mathbb{N}$ and $d \in (0, \lambda)$ satisfy Condition 3.2. Fix any $\varepsilon > 0$ and, similarly to (37), let

$$\gamma = \gamma(\varepsilon) = \frac{1 - e^{-d\varepsilon}}{\varepsilon}. \tag{46}$$

Notice that $\gamma < d < 1$. Similarly, to (38), for any $n \ge n_0$, $\rho > 1$ and integer $i \in [1, n\varepsilon)$,

$$E(\rho^{X_{k+1}^{(n)}} | X_k^{(n)} = i) \le \Big(1 + (\rho - 1)\frac{\gamma i}{n}\Big)^{n-i} \le e^{(\rho-1)\gamma i}.$$

Recall (35) and set $\rho = \alpha_\gamma$. Thus $E(\rho^{X_{k+1}^{(n)}} | X_k^{(n)} = i) \le \rho^i$ for any $n \ge n_0$ and $\in [1, n\varepsilon)$. Doob's optional theorem implies that

$$\rho^i \ge E\Big(\rho^{X_{T_m^{(n)}}^{(n)}} \Big| X_0^{(n)} = i\Big) \ge h_m^{(n)}(i)\rho^m + (1 - h_m^{(n)}(i)).$$

Therefore,

$$h_m^{(n)}(i) \le \frac{\rho^i - 1}{\rho^m - 1} = \frac{\alpha_\gamma^i - 1}{\alpha_\gamma^m - 1} \qquad \forall\ n \ge n_0.$$

By taking ε to zero in (46) we obtain

$$\limsup_{n\to\infty} h_m^{(n)}(i) \le \frac{\alpha_d^i - 1}{\alpha_d^m - 1}. \tag{47}$$

Since in this argument d is an arbitrary number in $(0, \lambda)$, the result of part (a) follows by taking d to λ in the above inequality.

(b) Observe that (47) is still true for $\lambda = 1$ and any $d \in (0, 1)$. Furthermore, $\lim_{d\to 1} \alpha_d = 1$, and hence, by the L'Hospital rule,

$$\limsup_{n\to\infty} h_m^{(n)}(i) \le \lim_{\alpha\to 1} \frac{\alpha^i - 1}{\alpha^m - 1} = \frac{i}{m}.$$

The proof of the proposition is complete. □

5 Duration of the avalanche

The goal of this section is to evaluate distribution tails of the avalanche's duration. The expected value of the duration is discussed in some detail elsewhere [27, 34] using a mixture of numerical and computational approaches. The main result here is Theorem 5.1 which provides estimates for a single network. Some consequences for a network ensemble satisfying Assumption 3.1 are drawn in Corollaries 5.5 and 5.6 at the end of the section. The basic idea of the proofs is to compare the avalanche model to a branching process during a time frame which on one hand is large enough for an asymptotic pattern to emerge and, on the other hand, is sufficiently small so that with an asymptotically overwhelming probability, the paths of the avalanche Markov chain X and a coupled branching process wouldn't diverge until its end. The proofs rely on asymptotically tight estimates for the branching process given in [1].

For $c \in (0,1)$ let

$$s(s) = \frac{2-c}{c} \qquad \text{and} \qquad r(c) = \frac{ce^{-c}}{e^{-c}-(1-c)}. \tag{48}$$

Recall T from (3) and α_c from (35).

Theorem 5.1. *Consider an avalanche process $(X_k)_{k\in\mathbb{Z}_+}$ with $X_0 = i_0$. Let $c = np$.*

(i) If $c > 1$, then there exist constants $\theta > 0$ and $K > 0$ such that for any pair of constants $x > 0$ and $m \in \mathbb{N}$ which satisfies the condition $xc^m < n$,

$$\Big(\frac{\alpha_c s_1(1-c^m\alpha_c^m)}{s_1 - c^m\alpha_c^m}\Big)^{i_0} \le P(T \le m) \le \Big(\frac{\alpha_c r_1(1-c^m\alpha_c^m)}{r_1 - c^m\alpha_c^m}\Big)^{i_0} + \frac{3c^{\frac{3(m+1)}{2}}mx^{3/2}}{2n} + K^{i_0}e^{-\theta x},$$

where $s_1 = s(\alpha_c c)$ and $r_1 = r(\alpha_c c)$.

(ii) If $c < 1$, then for any pair of constants $m, J \in \mathbb{N}$ such that $m < n$ and $i_0 < J < n$,

$$\Big(\frac{s_2(1-c^m)}{s_2 - c^m}\Big)^{i_0} \le P(T \le m) \le \Big(\frac{r_2(1-c^m)}{r_2 - c^m}\Big)^{i_0} + \frac{3cmJ^2}{2n} + \frac{\alpha_c^{i_0}-1}{\alpha_c^J - 1},$$

where $s_2 = s(c)$ and $r_2 = r(c)$.

(iii) If $c = 1$, then for any $m \in \mathbb{N}$ such that $m < n$,

$$\Big(\frac{m}{m+2}\Big)^{i_0} \le P(T \le m) \le \Big(\frac{m}{m+e-1}\Big)^{i_0} + \frac{3}{2}\Big(\frac{3mi_0^2}{n}\Big)^{1/3}.$$

The above bounds for the distribution function of T are originated in their counterparts for the limiting branching process, see Lemma 5.3 below. The latter estimates are borrowed from [1]. We remark that in the supercritical regime $c > 1$, a similar result for a frequency-dependent Wright-Fisher model has been proved in [12] (see Theorem 3.9 there). By taking n to infinity in the conclusions of the theorem, one can obtain tight asymptotic bounds for the avalanche model under Assumption 3.1, see Corollaries 5.5 and 5.6 below for details.

Proof of Theorem 5.1. Let $d_{TV}(X,Y)$ denote the total variation distance between the distributions of the random variables X and Y. That is, if X and Y are both non-negative and integer-valued, $d_{TV}(X,Y) = \frac{1}{2}\sum_{n=0}^{\infty}|P(X=n) - P(Y=n)|$. By the coupling inequality, $d_{TV}(X,Y) \le P(X \ne Y)$. Furthermore, there exists a maximal coupling, that is a random pair $(\widetilde{X}, \widetilde{Y})$ such that $\widetilde{X}$ is distributed the same as X, $\widetilde{Y}$ is distributed the same as Y, and $P(\widetilde{X} \ne \widetilde{Y}) = d_{TV}(X,Y)$ [36].

We will use the following inequalities. For the first claim see, for instance, Theorem 4 and subsequent Remark 1.1.4 in [11], and for the second one Theorem 1.C(i) in [8].

Lemma 5.2.

(i) Let $X = BIN(n,p)$ have the binomial distribution with parameters $n \in \mathbb{N}$, $p \in (0,1)$ and Y have the Poisson distribution with parameter $c = np$. Then $d_{TV}(X,Y) \le \frac{p}{2} \cdot \min\{1, c\}$.

(ii) Let X and Y be two Poisson random variables with parameters $\mu > 0$ and $c > \mu$, respectively. Then $d_{TV}(X,Y) \le \min\{1, \frac{1}{\sqrt{c}}\} \cdot |\mu - c|$.

Using the above results, we can construct a coupling of the Markov chain $(X_k)_{k\in\mathbb{Z}_+}$ and the limiting branching process $(Z_k)_{k\in\mathbb{Z}_+}$ as follows. The resulting process $(\widetilde{X}_k, \widetilde{Z}_k)_{k\in\mathbb{Z}_+}$ is a Markov chain. Suppose that the random pairs $(\widetilde{X}_t, \widetilde{Z}_t)$ have been sampled for all $t \le k$ and that $\widetilde{X}_t = \widetilde{Z}_t$ for all $t \le k$. Let i be the common value of $\widetilde{Z}_k$ and $\widetilde{X}_k$. Sample the next pair $(\widetilde{X}_{k+1}, \widetilde{Z}_{k+1})$ using the maximal coupling for X_{k+1} under the conditional distribution $P(X_{k+1} \in \cdot | X_k = i)$ and $Z_{k+1}^{(c)}$ under the conditional distribution $P(Z_{k+1}^{(c)} \in \cdot | Z_k^{(c)} = i)$. After the random time $\tau := \min\{k \in \mathbb{N} : \widetilde{X}_k \ne \widetilde{Z}_k\}$, sample $(\widetilde{Z}_t)_{t\ge\tau}$ and $(\widetilde{X}_t)_{t\ge\tau}$ independently. Using Lemma 5.2 and at first approximating $\widetilde{X}_{k+1}$ by a Poisson random variable with parameter $(n-i)(1-q^i)$, we obtain that for $c > 1$,

$$\begin{aligned} P(\widetilde{X}_{k+1} \ne \widetilde{Z}_{k+1} | \widetilde{X}_k = \widetilde{Z}_k = i) &\le \frac{1}{2}(1-q^i) + \frac{1}{\sqrt{ci}}|(n-i)(1-q^i) - ci| \\ &\le \frac{ci}{2n} + \frac{1}{\sqrt{ci}}[ci - (n-i)(1-q^i)]. \end{aligned}$$

Using the bound in Lemma 2.7 we conclude that in the case when $c > 1$,

$$P(\widetilde{X}_{k+1} \ne \widetilde{Z}_{k+1} | \widetilde{X}_k = \widetilde{Z}_k = i) \le \frac{ci}{2n} + \frac{c^{1/2}i^{3/2}}{2n} + \frac{c^{3/2}i^{3/2}}{2n} \le \frac{3c^{3/2}i^{3/2}}{2n}. \tag{49}$$

Similarly, when $c \le 1$, without making an assumption on whether $ci \le 1$ or not,

$$P(\widetilde{X}_{k+1} \ne \widetilde{Z}_{k+1} | \widetilde{X}_k = \widetilde{Z}_k = i) \le \frac{ci}{2n} + \frac{ci^2}{2n} + \frac{c^2i^2}{2n} \le \frac{3ci^2}{2n}. \tag{50}$$

Recall

$$\tau = \inf\{k \in \mathbb{N} : \widetilde{X}_k \ne \widetilde{Z}_k\}$$

and let

$$\sigma = \inf\{k \in \mathbb{N} : \widetilde{Z}_k = 0\}.$$

To evaluate the distribution function of T we will use the following inequalities:

$$\begin{aligned} P(T \leq m) &\leq P(T \leq m, \tau > m) + P(\tau \leq m) \\ &\leq P(\sigma \leq m) + P(\tau \leq m) \end{aligned} \tag{51}$$

and

$$P(T \leq m) \geq P(\sigma \leq m). \tag{52}$$

The latter inequality holds true because the avalanche process X is stochastically dominated by the branching process $Z^{(c)}$ by virtue of Proposition 3.3.

The following lemma summarizes results of [1] regarding the distribution (subdistribution if the process is supercritical) function $P(\sigma \leq m)$. Specific bounds for the extinction time of a Poisson distribution are derived in Theorem 2 of [1].

Recall α_c from (35) and $s(c), r(c)$ from (48).

Lemma 5.3 ([1]).

(i) If $c > 1$, *then for any* $m \in \mathbb{N}$, $(\frac{\alpha_c s_1(1-c^m\alpha_c^m)}{s_1 - c^m\alpha_c^m})^{i_0} \leq P(\sigma \leq m) \leq (\frac{\alpha_c r_1(1-c^m\alpha_c^m)}{r_1 - c^m\alpha_c^m})^{i_0}$, *where* $s_1 = s(\alpha_c c)$ *and* $r_1 = r(\alpha_c c)$.

(ii) If $c < 1$, *then for any* $m \in \mathbb{N}$, $(\frac{s_2(1-c^m)}{s_2-c^m})^{i_0} \leq P(\sigma \leq m) \leq (\frac{r_2(1-c^m)}{r_2-c^m})^{i_0}$, *where* $s_2 = s(c)$ *and* $r_2 = r(c)$.

(iii) If $c = 1$, *then for any* $m \in \mathbb{N}$, $(\frac{m}{m+2})^{i_0} \leq P(\sigma \leq m) \leq (\frac{m}{m+e-1})^{i_0}$.

We next estimate $P(\tau \leq m)$. For $k \in \mathbb{N}$ let $W_k = \max_{0 \leq i \leq k} \widetilde{Z}_i$. By the Markov property of $(\widetilde{X}_k, \widetilde{Z}_k)_{k \in \mathbb{Z}_+}$, for any $J \in \mathbb{N}$ we have:

$$\begin{aligned} P(\tau \leq m) &\leq P(\tau \leq m \text{ and } W_m < J) + P(W_m \geq J) \\ &\leq P\Big(\bigcup_{k=1}^{m} \{\widetilde{X}_{k-1} = \widetilde{Z}_{k-1} < J, \, \widetilde{X}_k \neq \widetilde{Z}_k\}\Big) + P(W_m \geq J) \\ &\leq m \cdot P(\widetilde{X}_k \neq \widetilde{Z}_k \,|\, \widetilde{X}_{k-1} = \widetilde{Z}_{k-1} < J) + P(W_m \geq J). \end{aligned} \tag{53}$$

The first part of the next lemma is an improved version of Lemma 5.6 in [12].

Lemma 5.4.

(i) Suppose that $c > 1$. *Then there exist constants* $\theta > 0$ *and* $K > 0$ *such that for any* $x > 0$ *and* $m \in \mathbb{N}$ *we have* $P(W_m \geq xc^m) \leq K^{i_0} e^{-\theta x}$.

(ii) If $c < 1$, *then for any* $m \in \mathbb{N}$ *and integer* $J \geq i_0$, *we have* $P(W_m \geq J) \leq \frac{\alpha_c^{i_0} - 1}{\alpha_c^J - 1}$.

(iii) If $c = 1$, *then for any* $m \in \mathbb{N}$ *and integer* $J \geq i_0$, *we have* $P(W_m \geq J) \leq \frac{i_0}{J}$.

Proof of Lemma 5.4.
(i) Let $U_k := c^{-k}\widetilde{Z}_k$. Then $(U_k)_{k \in \mathbb{Z}_+}$ is a martingale with respect to its natural filtration. For any $\theta > 0$, $f(x) = e^{\theta x}$ is a convex function and hence the sequence $e^{\theta U_k}$, $k \in \mathbb{Z}_+$, form a submartingale. Hence, by Doob's maximal inequality,

$$P(W_m \geq xc^m) \;\; \leq \;\; P(\max_{0 \leq k \leq m} e^{\theta U_k} \geq e^{\theta x}) \leq e^{-\theta x} E(e^{\theta U_m}).$$

The result now follows from Theorem 4 in [4] which states that $\sup_{m\in\mathbb{N}} E(e^{\theta U_m}) < +\infty$ for some $\theta > 0$.

(ii) and (iii) This is a direct implication of Theorem 2 in [26]. $\square$

The claims in the theorem follow now by combining the bounds in (49)–(53) with the estimates in Lemmas 5.3 and 5.4. In the case when $c = 1$, the optimization problem over the optimal choice of the parameter J in the upper bound for $P(T \le m)$ can be solved explicitly, and we choose $J = (\frac{ni_0^2}{3m})^{1/3}$. $\square$

We will next consider a family of avalanches under Assumption 3.1. For $J > 0$ let

$$T^{(n)} = \min\{k \in \mathbb{N} : X_k^{(n)} = 0\}. \tag{54}$$

The following result shows that the bounds in Lemma 5.3 hold asymptotically for the avalanche model.

Corollary 5.5. *Let Assumption 3.1 hold.*

(i) If $\lambda > 1$, then for any $m \in \mathbb{N}$,

$$\begin{aligned}\Big(\frac{\alpha_\lambda s_3(1-\lambda^m\alpha_\lambda^m)}{s_3-\lambda^m\alpha_\lambda^m}\Big)^{i_0} &\le \liminf_{n\to\infty} P(T^{(n)} \le m) \\ &\le \limsup_{n\to\infty} P(T^{(n)} \le m) \le \Big(\frac{\alpha_\lambda r_3(1-\lambda^m\alpha_\lambda^m)}{r_3-\lambda^m\alpha_\lambda^m}\Big)^{i_0},\end{aligned}$$

where $s_3 = s(\alpha_\lambda\lambda)$ and $r_3 = r(\alpha_\lambda\lambda)$.

(ii) If $\lambda < 1$, then for any $m \in \mathbb{N}$,

$$\Big(\frac{s_4(1-\lambda^m)}{s_4-\lambda^m}\Big)^{i_0} \le \liminf_{n\to\infty} P(T^{(n)} \le m) \le \limsup_{n\to\infty} P(T^{(n)} \le m) \le \Big(\frac{r_4(1-\lambda^m)}{r_4-\lambda^m}\Big)^{i_0},$$

where $s_4 = s(\lambda)$ and $r_4 = r(\lambda)$.

(iii) If $\lambda = 1$, then for any $m \in \mathbb{N}$,

$$\Big(\frac{m}{m+2}\Big)^{i_0} \le \liminf_{n\to\infty} P(T^{(n)} \le m) \le \limsup_{n\to\infty} P(T^{(n)} \le m) \le \Big(\frac{m}{m+e-1}\Big)^{i_0}.$$

Proof. Informally speaking, the obvious strategy for the proof is to take n to infinity in the conclusions of Theorem 5.1. Let $\alpha_1 = r(1) = s(1) = 1$. Observe that for any $\lambda > 0$, $\lim_{\mu\to\lambda} \alpha_\mu = \alpha_\lambda$, and for any $\lambda \in [0,1]$,

$$\lim_{\mu\to\lambda} s(\mu) = s(\lambda), \quad \lim_{\mu\to\lambda} r(\mu) = r(\lambda),$$

where in the case $\lambda = 1$ the latter limits are understood as left limits. Furthermore, the limits in Lemma 5.3 are continuous functions of c. This is evident for $c \ne 1$, and for $c = 1$ by the L'Hospital rule we have

$$\lim_{c\to 1^-} \frac{s(c)(1-c^m)}{s(c)-c^m} = \lim_{c\to 1^-} \frac{-(1-c^m)-m(2-c)c^{m-1}}{-1-(m+1)c^m} = \frac{m}{m+2}$$

and

$$
\begin{aligned}
\lim_{c\to 1^-} \frac{r(c)(1-c^m)}{r(c)-c^m} &= \lim_{c\to 1^-} \frac{(e^{-c}-ce^{-c})(1-c^m)-me^{-c}c^m}{e^{-c}-ce^{-c}-mc^{m-1}(e^{-c}-1+c)-c^m(-e^{-c}+1)} \\
&= \frac{me^{-1}}{me^{-1}-e^{-1}+1} = \frac{m}{m+e-1},
\end{aligned}
$$

as desired.

For $k \in \mathbb{N}$ and $\mu > 0$, let $W_k^{(\mu)} = \max_{0\le i\le k} Z_i^{(\mu)}$. In comparison with the proof of Theorem 5.1, the only nuance here is that before letting n go to infinity the term $P(W_m^{(np_n)} \ge J)$ in (53), which a priori is not uniform in n, should be replaced with $P(W_m^{(c)} \ge J_n)$, where c is the constant introduced in Condition 3.2 and J_n is a suitable sequence which in particular satisfies $\lim_{n\to\infty} J_n = +\infty$. [1] □

Another interesting consequence of Theorem 5.1 is the following result which shows that on the right scale, the asymptotic behavior of the avalanche model coincides with that of the branching process (see, for instance, compact introduction section in [28] for the underlying branching process estimates).

Corollary 5.6. *Let Assumption 3.1 hold.*

(i) Suppose that $\lambda > 1$. *Let* $(\beta_n)_{n\in\mathbb{N}}$ *be a sequence of positive constants such that*

$$
\lim_{n\to\infty} \beta_n \log n = +\infty \qquad \text{and} \qquad \limsup_{n\to\infty} \beta_n < \frac{2}{3\log\lambda}.
$$

Let $m_n = \beta_n \log n$. *Then* $\lim_{n\to\infty} P(T^{(n)} > m_n) = 1 - \alpha_\lambda^{i_0}$.

(ii) Suppose that $\lambda > 1$. *Let* $(\beta_n)_{n\in\mathbb{N}}$ *be a sequence of positive constants such that*

$$
\lim_{n\to\infty} \frac{\beta_n(\log n)^3}{n} = 0 \qquad \text{and} \qquad \liminf_{n\to\infty} \beta_n > -\frac{1}{3\log\lambda}.
$$

Let $m_n = \beta_n \log n$. *Then* $\lim_{n\to\infty} \frac{1}{m_n} \log P(T^{(n)} > m_n) = \log\lambda$.

(iii) Assume $\lambda = 1$. *Let* $(m_n)_{n\in\mathbb{N}}$ *be a sequence of positive integers such that* $\lim_{n\to\infty} m_n = +\infty$ *and* $\lim_{n\to\infty} \frac{m_n^4}{n} = 0$. *Then* $\lim_{n\to\infty} m_n P(T^{(n)} > m_n) = 2i_0$.

Proof. In each of the three cases ($\lambda >, <, = 1$) the sequence β_n is chosen in such a way that the term corresponding to $P(W_n \ge m)$ in the key estimate (53) and consequently in the conclusions of Theorem 5.1 is dominated by the bounds for $P(\sigma \le m)$ which are supplied by Lemma 5.3. As the result, the term vanishes as n goes to infinity, and the asymptotic behavior of $P(T^{(n)} > m_n)$ coincides with its counterpart for branching processes. The only subtlety is in the proof of part (iii), where the asymptotic behavior exhibited by the lower and upper bounds in Lemma 5.3 do not match, namely $\lim_{n\to\infty} m\{1-(\frac{m}{m+2})^{i_0}\} = 2i_0$ while $\lim_{n\to\infty} m\{1-(\frac{m}{m+e-1})^{i_0}\} = (e-1)i_0$. However, Kolmogorov's estimate (see, for instance, Theorem C in [28]) ensures that $\lim_{m\to\infty} P(\sigma > m) = 2i_0$. In view of (53) this yields the desired result in the case $\lambda = 1$. □

[1] Technically, this part of the proof makes the claim a corollary to a slight modification of Theorem 5.1 rather than to the theorem itself.

6 Deterministic approximation

The branching approximation reflects the dynamics of the process suitably when the latter is subcritical and the size of avalanche decays exponentially fast, similarly to the branching approximation process. In contrast, in the supercritical regime, namely when Condition 2.1 is satisfied with $c > 1$ or Assumption 3.1 holds with $\lambda > 1$, the results in Section 2 tend to provide a little or no information on the asymptotic behavior of the model, and even a partial improvement of this situation is clearly desirable. However, even in the supercritical case, one should expect that with $\varepsilon \ll 1$, the branching approximation is adequate for $\sim \varepsilon \log n$ of first steps while the ratio $\frac{X_k}{n}$ remains small, cf. Section 6.3.1 of [15]. On the other hand, once the ratio becomes of order one, one can expect that the dynamics of the model will be almost deterministic and follow a "mean-field" equation due to the law of large numbers. The section is devoted to the study of this deterministic approximation of the model. The relevance of this regime to a supercrtical avalanche model is formally elucidated by the results in Section 4.

The rest of the section is divided into three subsections. Section 6.1 discusses an heuristic computation based on a complete decoupling of a week dependence between the network nodes, that shades an additional light on the nature of the deterministic approximation. In Section 6.2 the deterministic dynamical system is formally introduced and studied. Interestingly enough, the underlying dynamics resembles but is richer than that of the logistic equation. This was already observed in [14, 27]. Section 6.3 addresses the question for how long the trajectories of an avalanche Markov chain and the corresponding deterministic dynamical system remain close each to other, provided they began at the same point.

6.1 An extra motivation and a precise comparison result

Our motivation in this section is partially coming from the following heuristic computation, which is an adaptation to our framework of the one proposed in [17] for a model of avalanches in a Boolean network. Consider an avalanche Markov chain $(X_k)_{k\in\mathbb{Z}_+}$ with transition kernel introduced in (2), and recall $\mathcal{E}_k$ from (6) and $\xi_k(x)$ from (9). If the random variables $\{\mathcal{E}_k(x) : x \in V_n\}$ were independent and identically distributed for some $k \in \mathbb{Z}_+$, we could relate their common value ξ_k to the common value ξ_{k+1} of $\xi_{k+1}(x)$, $x \in V_n$, by means of the following recursion [17]:

$$\begin{aligned}\xi_{k+1} &= (1-\xi_k)\sum_{i=0}^{n-1}\binom{n-1}{i}\xi_k^i(1-\xi_k)^{n-1-i}(1-q^i)\\ &= (1-\xi_k)[1-(q\xi_k+1-\xi_k)^n] = (1-\xi_k)[1-(1-p\xi_k)^n].\end{aligned} \tag{55}$$

Consequently, the following inequality would hold:

$$\xi_{k+1} \geq (1-\xi_k)(1-e^{-pn\xi_k}) = (1-\xi_k)(1-e^{-c\xi_k}), \tag{56}$$

where we denote $c = np$. We remark that similar decoupling arguments are often used in a physics literature to justify a mean-field approximation in a complex locally tree-like network (hence a fairly weak dependence between the nodes), see, for instance, [23, 24, 25, 32, 33].

Of course, $\mathcal{E}_k(x)$ are not independent and in general are not identically distributed (the latter depends on whether the distribution of X_0 is exchangeable or not, cf. Proposition 2.5). For a different, but somewhat related model of avalanches, it is argued in [17] (see the comment [4] in the References section there) that the above i. i. d. assumption "is true for large networks with well-behaved degree distributions, but excludes networks with hubs which output to a substantial fraction of the nodes." It is of interest to note that, in agreement with this general heuristic principle, a suitable modification of (55) and, consequently, of (56) serve in a certain rigorous sense as a good approximation for the avalanche model. This is accomplished in (64) below.

We will now proceed with a rigorous modification of the above heuristic calculation. For $\alpha > 0$ let

$$g_\alpha(x) = (1-x)(1-e^{-\alpha x}), \qquad x \in [0,1],$$

and

$$\chi_\alpha = \max_{x\in(0,1)} g_\alpha(x), \quad \text{and} \quad \nu_\alpha = \operatorname*{argmax}_{x\in(0,1)} g_\alpha(x).$$

Note that ν_α is uniquely defined for all $\alpha > 0$. Further, if $\alpha > 1$, let ζ_α be the unique on $(0,1)$ solution to the fixed point equation

$$g_\alpha(\zeta_\alpha) = \zeta_\alpha, \qquad \zeta_\alpha \in (0,1).$$

Remark that $\zeta_\alpha = g_\alpha(\zeta_\alpha) < 1-\zeta_\alpha$ implies that $\zeta_\alpha < 1/2$. The following lemma summarizes other basic properties of the function g_α that we are going to use (see also Theorem 6.6 in Section 6.2 below). The proof of the lemma is omitted, the dependence of the graph of g_α on the parameter α as well as basic properties of g_α are illustrated in Fig. 3 below. For more details see, for instance, [14] where the function g_α is studied systematically in a similar context.

Lemma 6.1.

(i) *$g_\alpha(x)$ is increasing on $(0,\nu_\alpha)$ and decreasing on $(\nu_\alpha,1)$.*

(ii) *If $\alpha \le 1$, then $g_\alpha(x) < x$ for all $x \in (0,1)$.*

(iii) *If $\alpha > 1$, then $g_\alpha(x) > x$ on $(0,\zeta_\alpha)$ and $g_\alpha(x) < x$ on $(\zeta_\alpha,1)$.*

(iv) *$g_\alpha(x) < \alpha x$ for all $\alpha > 0$ and $x \in (0,1)$.*

(v) *There exists a transitional value $\alpha_{\rm tr} > 1$ such that $\nu_\alpha > \zeta_\alpha$ if and only if $\alpha < \alpha_{\rm tr}$.*

Numerical simulations indicate that $\alpha_{\rm tr} \approx 2.46742$. The role of the threshold in the dynamics of the model is discussed in more detail in [14] and [34].

For an avalanche model with transition kernel (2) and fixed network size $n \in \mathbb{N}$, let

$$\varphi_k = \frac{1}{n}E(X_k), \qquad k \in \mathbb{Z}_+. \tag{57}$$

Note that if the distribution of X_0 is exchangeable (i. e. invariant with respect to permutation of the network nodes), then $\varphi_k = \xi_k(x)$, where ξ_k is defined in (9), for all $k \in \mathbb{Z}_+$ and $x \in V_n$. By the bounded convergence theorem, for any fixed $n \in \mathbb{N}$, $\lim_{k\to\infty} \varphi_k = 0$ regardless of the choice of parameter $p > 0$. We have:

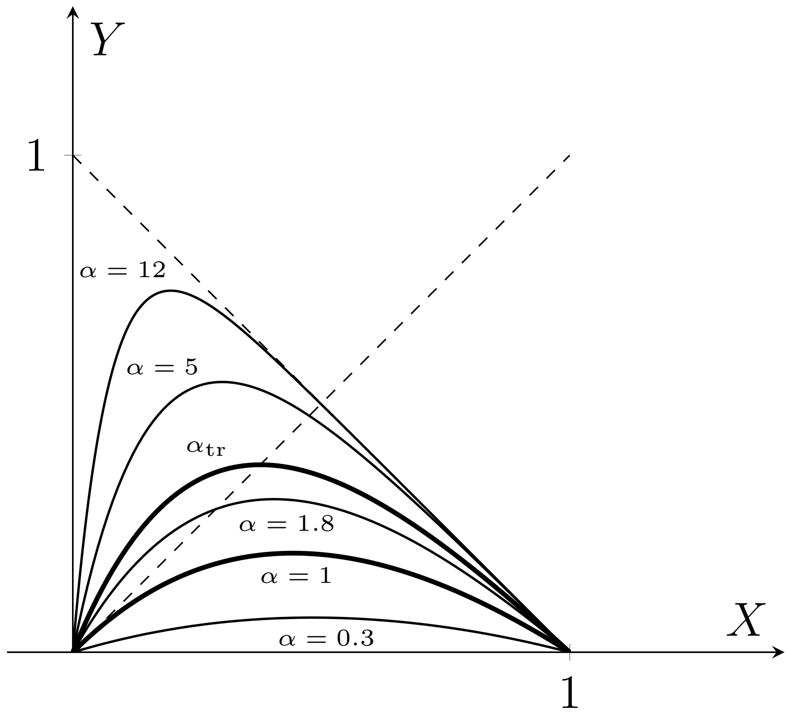

Figure 3: Graph of the function $g_\alpha(x) = (1-x)(1-e^{-\alpha x})$ for several values of the parameter α, including the critical branching value $\alpha = 1$, for which $g_1'(0) = 1$, and $\alpha = 2.46742$ which is a close approximation to the transitional value α_{tr}.

Theorem 6.2. *Let $(X_k)_{k\in\mathbb{Z}_+}$ be an avalanche model with transition kernel* (2)*. Denote*

$$\alpha = -n\log(1-p). \tag{58}$$

The following holds true:

(i) Let $(\phi_k)_{k\in\mathbb{Z}_+}$ be defined recursively by setting $\phi_0 = \frac{1}{n}E(X_0)$ and

$$\phi_{k+1} = 1 - e^{-\alpha\phi_k}.$$

Then $\varphi_k \le \phi_k$ for all $k \in \mathbb{Z}_+$.

(ii) $\sup_{k\in\mathbb{Z}_+}\varphi_k \le \max\{\phi_0, \chi_\alpha\}$.

(iii) Let $(\psi_k)_{k\in\mathbb{Z}_+}$ be defined recursively by setting $\psi_0 = \frac{1}{n}E(X_0)$ and

$$\psi_{k+1} = g_\alpha(\psi_k).$$

If $\alpha \le \alpha_{\text{tr}}$ and

$$\varphi_0 \le \zeta_\alpha,$$

then

$$\varphi_k \le \psi_k \qquad \forall\, k \in \mathbb{Z}_+,$$

and, moreover,

$$\sup_{k\in\mathbb{Z}_+} \varphi_k \le \zeta_\alpha \le \chi_\alpha.$$

Note that in view of the fact that $1-e^{-\alpha\phi_k} < \alpha\phi_k$ and the result in part (iv) of Lemma 6.1, the theorem improves the result in the basic Corollary 2.3 even in the subcritical case $\alpha < 1$. The relevance of the phase transition at α_{tr} to the dynamics of the avalanche Markov chain is further discussed in Section 6.2 below, at the paragraph following Corollary 6.3.

Proof of Theorem 6.2.

(i) It follows from (2), (58), and Jensen's inequality which is applied in the last step to the concave on $(0,1)$ function $g(x) = (1-x)(1-e^{-\alpha x})$ that

$$\begin{aligned} \varphi_{k+1}(x) &= E\Big[\Big(1-\frac{X_k}{n}\Big)(1-q^{X_k})\Big] = E\Big[\Big(1-\frac{X_k}{n}\Big)\Big(1-(1-p)^{X_k}\Big)\Big] \\ &= E\Big[\Big(1-\frac{X_k}{n}\Big)\Big(1-\Big\{(1-p)^p\Big\}^{\frac{X_k}{p}}\Big)\Big] \\ &= E\Big[\Big(1-\frac{X_k}{n}\Big)\Big(1-e^{-\frac{\alpha X_k}{n}}\Big)\Big] \le (1-\varphi_k)(1-e^{-\alpha\varphi_k}). \end{aligned} \tag{59}$$

In particular, $\varphi_{k+1} \le 1-e^{-\alpha\varphi_k}$. Since $f(x) = 1-e^{-\alpha x}$ is a monotone increasing function, an induction argument shows that $\varphi_k \le \phi_k$ for all $k \in \mathbb{Z}_+$.

(ii) The claim is immediate from (59).

(iii) Observe that g_α is monotone increasing on $(0,\zeta_\alpha)$ for $\alpha \le \alpha_{\text{tr}}$. Therefore, it follows from (59) by using induction on k, that $\varphi_k \le \psi_k \le \zeta_\alpha$ for all $k \in \mathbb{Z}_+$. Moreover, if $\alpha \le 1$ then $\psi(x) < x$ on $(0,\zeta_\alpha)$, and hence $\psi_{k+1} \le \psi_k$ for all $k \in \mathbb{Z}_+$, while if $\alpha \in (1,\alpha_{\text{tr}}]$ then $\psi(x) > x$ on $(0,\zeta_\alpha)$, and hence $\psi_{k+1} \ge \psi_k$ for all $k \in \mathbb{Z}_+$. Since g_α has either one fixed point at zero (for $\alpha \le 1$) or two at zero and ζ_α (for $\alpha > 1$), this implies that $\lim_{k\to\infty}\varphi_k = 0$ for $\alpha \le 1$ while $\sup_{k\in\mathbb{Z}_+}\varphi_k \le \sup_{k\in\mathbb{Z}_+}\psi_k \le \lim_{k\to\infty}\varphi_k = \zeta_\alpha$ in the case that $\alpha \in (1,\alpha_{\text{tr}}]$. □

6.2 Deterministic approximation for large size networks

We turn now to a study of an ensemble of avalanche models which satisfies Assumption 3.1. First we discuss a direct implication of the results in Theorem 6.2 for the asymptotic behavior of the expected fraction of excited nodes in a network of the ensemble. The main results of this section are stated afterwards in Theorems 6.4 (regarding asymptotic behavior of $\frac{1}{n}X_k^{(n)}$ for large n) and Corollary 6.5 (a consequence of the results in Theorem 6.4 for the heterogeneity of a network in the ensemble).

First, observe that continuity of the results in Theorem 6.2 in the parameters α and the initial data φ_0 together with the monotonicity of g_α on the interval $(0,\zeta_\alpha)$ lead to the following corollary to the theorem.

Corollary 6.3. *Let Assumption 3.1 hold, and suppose that the following limit exists and belongs to* $(0,1)$:

$$\varepsilon_0 = \lim_{n\to\infty} \frac{1}{n} E(X_0^{(n)}).$$

For $n \in \mathbb{N}$ *and* $k \in \mathbb{Z}_+$ *let*

$$\varphi_{n,k} = \frac{1}{n} E(X_k^{(n)}). \tag{60}$$

be the counterpart of φ_k *introduced for a single network in* (57).

Then the following holds true:

(i) Let $(\phi_k)_{k\in\mathbb{Z}_+}$ *be defined recursively by setting* $\phi_0 = \varepsilon_0$ *and*

$$\phi_{k+1} = 1 - e^{-\lambda\phi_k}.$$

Then $\limsup_{n\to\infty} \varphi_{n,k} \le \phi_k$ *for all* $k \in \mathbb{Z}_+$.

(ii) $\limsup_{n\to\infty} \varphi_{n,k} \le \chi_\lambda$ *for all* $k \in \mathbb{N}$.

(iii) Recall $\alpha_{\rm tr}$ *from Lemma 6.1. Let* $(\psi_k)_{k\in\mathbb{Z}_+}$ *be defined recursively by setting*

$$\psi_0 = \varepsilon_0 \qquad \text{and} \qquad \psi_{k+1} = g_\lambda(\psi_k). \tag{61}$$

If $\lambda < \alpha_{\rm tr}$ *and* $\varepsilon_0 < \zeta_\lambda$, *then*

$$\limsup_{n\to\infty} \varphi_{n,k} \le \psi_k \qquad \text{and} \qquad \limsup_{n\to\infty} \varphi_{n,k} \le \zeta_\lambda \le \chi_\lambda$$

for all $k \in \mathbb{Z}_+$.

Let

$$x_{n,k} = \frac{X_k^{(n)}}{n}, \qquad n \in \mathbb{N},\ k \in \mathbb{Z}_+. \tag{62}$$

Under Assumption 3.1, define the asymptotic branching factor as a function $b : (0,1) \to \mathbb{R}$ by setting

$$b(x) = \lim_{n\to\infty} \frac{1}{x} E(x_{n,k+1} \,|\, x_{n,k} = x) = \frac{g_\lambda(x)}{x}, \qquad x \in (0,1).$$

It turns out (see [14] for details) that the behavior of the sequence $b(\psi_k)$ for $\alpha \in (1, \alpha_{\rm tr})$ and $\alpha > \alpha_{\rm tr}$ differ qualitatively, that is a secondary phase transition in the avalanche model occurs at the transitional value $\alpha_{\rm tr}$. For instance, if $\alpha \in (1, \alpha_{\rm tr})$ and ψ_0 is sufficiently small, then $b(\psi_k)$ increases monotonically to one as $k \to \infty$. In contrast, that's not necessarily true when $\alpha > \alpha_{\rm tr}$. For example, in the case $\alpha > \alpha_{\rm tr}$, if for some $m \in \mathbb{N}$, $\psi_0 < \nu_\alpha$ and $\psi_m = \nu_m$, then the sequence ψ_k increases monotonically at the first m steps and then converging to one by oscillating consequently between values which are larger and smaller than one [14, p. 81]. Note that by choosing m sufficiently large, we can place ψ_0 as close to zero (the only fixed point of the equation $x = g_\alpha(x)$ other than ζ_α) as we wish.

Let $\stackrel{P}{\to}$ denote convergence in probability as the size of the network n goes to infinity. The following theorem is an adaptation to our setup of Theorems 1 and 3 in [10] (see also [22] for earlier similar results).

Theorem 6.4. *Let Assumption 3.1 hold. Recall $x_{n,k}$ from* (62) *and suppose that*

$$x_{n,0} \overset{P}{\to} \psi_0$$

for some constant $\psi_0 \in (0,1)$. Then the following holds true:

(a) $x_{n,k} \overset{P}{\to} \psi_k$ *for all $k \in \mathbb{Z}_+$, where $\psi_{k+1} = g_\lambda(\psi_k)$.*

(b) Let

$$y_{n,k} = \sqrt{n}(x_{n,k} - \psi_k), \qquad n \in \mathbb{N},\, k \in \mathbb{Z}_+$$

and set

$$v(x) := g_\lambda(x)e^{-\lambda x} = (1-x)e^{-\lambda x}(1-e^{-\lambda x}), \qquad x \in [0,1].$$

Suppose that in addition to (62)*, $y_{n,0}$ converges weakly, as n goes to infinity, to some (possibly random) Y_0. Then the sequence $y^{(n)} := (y_{n,k})_{k\in\mathbb{Z}_+}$ converges in distribution, as n goes to infinity, to a time-inhomogeneous Gaussian $AR(1)$ sequence $(Y_k)_{k\in\mathbb{Z}_+}$ defined by*

$$Y_{k+1} = g'_\lambda(\psi_k)Y_k + e_k, \tag{63}$$

where e_k, $k \in \mathbb{Z}_+$, are independent Gaussian variables, each e_k distributed as $N(0, v(\psi_k))$.

Proof. Let $\mathfrak{Z} := \{\mathfrak{Z}_{k,i}^{(n,j)} : n,i,j \in \mathbb{N}, k \in \mathbb{Z}_+\}$ be a collection of independent Bernoulli variables with

$$P(\mathfrak{Z}_{k,i}^{(n,j)} = 1) = 1 - q_n^j \qquad \text{and} \qquad P(\mathfrak{Z}_{k,i}^{(n,j)} = 1) = q_n^j.$$

Thus, without loss of generality, we can assume that

$$X_{k+1}^{(n)} = \sum_{i=1}^{n-X_k^{(n)}} \mathfrak{Z}_{k,i}^{(n,X_k^{(n)})}.$$

Let $\mathfrak{Y} := \{\mathfrak{Y}_{k,i}^{(n,j)} : n,i,j \in \mathbb{N}, k \in \mathbb{Z}_+\}$ be another collection of independent Bernoulli variables defined on the same probability space, and such that

$$P(\mathfrak{Y}_{k,i}^{(n,j)} = 1) = 1 - e^{-\frac{\lambda j}{n}} \qquad \text{and} \qquad P(\mathfrak{Y}_{k,i}^{(n,j)} = 1) = e^{-\frac{\lambda j}{n}}.$$

For $n \in \mathbb{N}$, let $\alpha_n = np_n$. By using the maximal coupling for two Bernoulli variables, we can and will assume that the pairs $(\mathfrak{Z}_{k,i}^{(n,j)}, \mathfrak{Y}_{k,i}^{(n,j)})$ are independent $\{0,1\}^2$-random variables, and

$$P(\mathfrak{Z}_{k,i} \neq \mathfrak{Y}_{k,i}) = \left| e^{-\frac{\lambda j}{n}} - \left(1 - \frac{\alpha_n}{n}\right)^j \right|.$$

For $n \in \mathbb{N}$, define a new sequence $\widetilde{X}^{(n)} = (\widetilde{X}_k^{(n)})_{k\in\mathbb{Z}_+}$ by setting $\widetilde{X}_0^{(n)} = \psi_0$ and

$$\widetilde{X}_{k+1}^{(n)} = \sum_{i=1}^{n-\widetilde{X}_k^{(n)}} \mathfrak{Y}_{k,i}^{(n,\widetilde{X}_k^{(n)})}.$$

Theorems 1 and 3 in [10] ensures that the results in the theorem, both the LLN and CLT, hold if we replace $X^{(n)}$ with $\widetilde{X}^{(n)}$. Thus, in order to prove the theorem, it suffices to show that $(X_k^{(n)} - \widetilde{X}_k^{(n)}) \overset{P}{\to} 0$ for all $k \in \mathbb{Z}_+$ (see, for instance, Remark (i) in [10, p. 60] and/or the last paragraph in the proof of Theorem 3 there, which both assert that the main result of [19] goes through to a non-homogeneous chain setting, and therefore the weak convergence in question is implied by the convergence of transition kernels). To this end, observe that

$$\begin{aligned} E\Big(\frac{1}{n}\Big|X_{k+1}^{(n)} - \widetilde{X}_{k+1}^{(n)}\Big|\Big) &\le \frac{1}{n}E\Big(\sum_{i=1}^{n-X_k^{(n)}} \Big|\mathfrak{Z}_{k,i}^{(n,X_k^{(n)})} - \mathfrak{Y}_{k,i}^{(n,X_k^{(n)})}\Big|\Big) \le E\Big(\Big|\mathfrak{Z}_{k,1}^{(n,X_k^{(n)})} - \mathfrak{Y}_{k,1}^{(n,X_k^{(n)})}\Big|\Big) \\ &= P\Big(\mathfrak{Z}_{k,1}^{(n,X_k^{(n)})} \ne \mathfrak{Y}_{k,1}^{(n,X_k^{(n)})}\Big) = E\Big(\Big|e^{-\frac{\lambda X_k^{(n)}}{n}} - \Big(1-\frac{\alpha_n}{n}\Big)^{X_k^{(n)}}\Big|\Big), \end{aligned}$$

and hence the claim can be proved by induction, using the bounded convergence theorem. □

Part (a) of the above result and the bounded convergence theorem imply that

$$\lim_{n\to\infty} \varphi_{n,k} = \psi_k \qquad \forall\, k \in \mathbb{Z}_+, \tag{64}$$

where $\varphi_{n,k}$ is defined in (57). Note that this limit identity is a reminiscent of the heuristic (55) and (56) in our framework.

It is not hard to prove (cf. Remark (v) on p. 61 of [10], see also [22]) that if $\psi_0 \overset{P}{\to} \zeta_\lambda$ in the statement of Theorem 6.4 and, in addition, $y_{n,0}$ converges weakly, as n goes to infinity, to some Y_0, then the linear recursion (63) can be replaced with

$$Y_{k+1} = g_\lambda'(\zeta_\lambda)Y_k + \widetilde{e}_k,$$

where $\widetilde{e}_k$, $k \in \mathbb{Z}_+$, are i. i. d. Gaussian variables, each $\widetilde{e}_k$ distributed as $N(0, v(\zeta_\lambda))$. One then can show (see the proof of Theorem 6.7 below) that $|g_\lambda'(\zeta_\lambda)| < 1$, and hence Markov chain Y_k has a stationary distribution, see [10, p. 61] for more details.

Recall H_k from (5) and define a normalized heterogeneity $h_{n,k}$ by

$$h_{n,k} = \frac{H_{n,k}}{2n(n-1)}, \qquad n \in \mathbb{N}, k \in \mathbb{Z}_+.$$

Notice that if the distribution of $X_0^{(n)}$ is exchangeable, then $x_{n,k}$ is a probability that two nodes in the network generating $X_k^{(n)}$ randomly chosen at time k have different types. The following is immediate from Theorem 6.4.

Corollary 6.5. *Under the conditions of Theorem 6.4, the following holds true:*

(a) For $x \in [0,1]$ let $r(x) = \frac{1}{2}x(1-x)$. Then $h_{n,k} \overset{P}{\to} r(\psi_k)$ for all $k \in \mathbb{Z}_+$.

(b) Let

$$\widetilde{y}_{n,k} = \sqrt{n}(h_{n,k} - r(\psi_k)), \qquad n \in \mathbb{N},\ k \in \mathbb{Z}_+.$$

Suppose that in addition to (62), $y_{n,0} = \sqrt{n}(x_{n,0} - \psi_0)$ converges weakly, as n goes to infinity, to some (possibly random) Y_0. Then the sequence $\widetilde{y}^{(n)} := (\widetilde{y}_{n,k})_{k\in\mathbb{Z}_+}$ converges in distribution, as n goes to infinity, to a time-inhomogeneous Gaussian $AR(1)$ sequence $(\widetilde{Y}_k)_{k\in\mathbb{Z}_+}$, where $\widetilde{Y}_k = \frac{1}{2}(1-2\psi_k)Y_k$ and Y_k is defined in (63).

Proof. The convergence in probability of $h_{n,k}$ follows from the continuous mapping theorem and the result in part (a) of Theorem 6.4. To prove the weak convergence of $\widetilde{y}_{n,k}$, write

$$\widetilde{y}_{n,k} = \sqrt{n}(r(x_{n,k}) - r(\psi_k)) + \frac{\sqrt{n}}{2} X_{n,k}(n - X_{n,k})\Big(\frac{1}{n(n-1)} - \frac{1}{n^2}\Big)$$

and observe that

$$\Big|\frac{\sqrt{n}}{2} X_{n,k}(n - X_{n,k})\Big(\frac{1}{n(n-1)} - \frac{1}{n^2}\Big)\Big| \leq \frac{\sqrt{n}}{2(n-1)} \longrightarrow_{n\to\infty} 0.$$

Furthermore, by the mean value theorem,

$$\sqrt{n}(r(x_{n,k}) - r(\psi_k)) = r'(x^*_{n,k})\sqrt{n}(x_{n,k} - \psi_k) = (1 - 2x_{n,k})\sqrt{n}(x_{n,k} - \psi_k)$$

for some $x^*_{n,k}$ between $x_{n,k}$ and ψ_k. In view of the result in part (b) of Theorem 6.4, the claim in part (b) of this theorem follows now by another application of the continuous mapping theorem. □

The discrete-time dynamical system ψ_k is studied in details in [14]. In particular, the following result is proved there (Theorem 1 in [14]):

Theorem 6.6 ([14])**.** *Let $\lambda > 0$ and $\varepsilon_0 \in (0, 1)$ be given, and the sequence $(\psi_k)_{k\in\mathbb{Z}_+}$ is defined as in* (61)*. Then the following holds true:*

(i) If $\lambda \in (0, 1]$, then $\lim_{k\to\infty} \psi_k = 0$ for any $\psi_0 \in [0, 1]$.

(ii) If $\lambda > 1$, then for all $\psi_0 \in [0, 1]$, $\lim_{k\to\infty} \psi_k = \zeta_\lambda$, where $\zeta_\lambda \in (0, 1/2)$ is the unique positive solution to the fixed point equation $g_\lambda(x) = x$.

Together with the results in Section 4 and Theorem 6.4, this theorem implies that when n is large, with high probability, a supercritical Markov chain $X^{(n)}$ will be eventually trapped for a long time in a neighborhood of the global stable point ζ_λ of the map g_λ. The next section is devoted to the proof of a certain qualitative form of this informal observation.

6.3 Comparison of the stochastic and deterministic trajectories

The main results of this section are stated in Theorem 6.7 (supercritical case) and Theorem 6.9 (critical and subcritical case).

Theorem 6.4 suggests that when both n and the first generation $X^{(n)}_0$ in the avalanche process are substantially large, the trajectory of the deterministic sequence ψ_k can serve as a good approximation to the path of the Markov chain $X^{(n)}_k$. Note however that, at least for a supercritical process, two trajectories cannot in principle stay close each to other forever since while the latter converges to zero with probability one, the former tends to a non-zero limit by virtue of Theorem 6.6.

The following theorem offers some insight into the duration of the time when the deterministic and the stochastic paths stay fairly close each to other, before they become significantly separated each from another at the first time. The theorem is a suitable modification of some results in [5]. In words, the theorem asserts that if Assumption 3.1 holds with $\lambda > 1$ and the scaled initial population $x_{n,0}$ is close enough to the stable point of the map g_λ and n is large, the trajectory of the avalanche model will stay close to the deterministic sequence ψ_k for a time which is exponentially large in the network size n. For reader's convenience we give a short detailed proof of the theorem which generally follows the line of argument in [5] but is different in several details.

Theorem 6.7. *Suppose that Assumption 3.1 is satisfied with $\lambda > 1$. For $\delta > 0$, let*

$$\tau_n(\delta) = \inf\{k \in \mathbb{Z}_+ : |x_{n,k} - \psi_k| \geq \delta\}, \tag{65}$$

where the sequence ψ_k is defined in (61).

There exist an interval $(a, b) \subset (0, 1)$ including ζ_λ and constants $\gamma > 0$, $\delta_0 \in (0, 1)$, and $n_0 \in \mathbb{N}$ such that if $x_{n,0} \in (a + \delta_0, b - \delta_0)$, and $n > n_0$, then for any $m \in \mathbb{N}$ and $\delta \in (0, \delta_0)$ we have

$$P(\tau_n(\delta) > m) \geq (1 - 2e^{-\gamma\delta^2 n})^m \geq 1 - 2me^{-\gamma\delta^2 n}.$$

Furthermore, the constants a, b, δ_0 and γ depend on the sequence of parameters $(p_n)_{n\in\mathbb{N}}$ through λ only (this is not necessarily true for n_0, which in general is not exclusively determined by the value of λ).

Proof. Pick first $b \in (0, 1)$ and then $a \in (0, 1)$ in such that a manner that

$$0 < a < \min\{\nu_\lambda, \zeta_\lambda, g_\lambda(b)\} < \max\{\nu_\lambda, \zeta_\lambda\} < b < 1.$$

Let $I = (a, b)$ and $h = \min\{g_\lambda(a), g_\lambda(b)\}$. Then

$$a < h, \qquad \nu_\lambda < b, \qquad g(I) \subset (h, \nu_\lambda).$$

Thus, if we set

$$\varepsilon = \min\{b - \zeta_\lambda, h - a, (b - a)/2\},$$

we get

$$g(a, b) \subset (a + \varepsilon, b - \varepsilon) \qquad \text{and} \qquad \zeta_\lambda \in (a + \varepsilon, b - \varepsilon). \tag{66}$$

The latter assertion is true because $g_\lambda(\zeta_\lambda) = \zeta_\lambda$ and the point ζ_λ belongs to the interval (a, b) which is mapped into $(a + \varepsilon, b - \varepsilon)$ by g_λ.

Next, observe that for any $x \in (0, 1)$,

$$g'_\lambda(x) = -1 + (1 + \alpha - \alpha x)e^{-\alpha x} > -1,$$

and

$$g''_\lambda(x) < 0 \text{ (and hence, } g'_\lambda \text{ is decreasing)}, \qquad g'_\lambda(1) = -1 + e^{-\alpha} < 0.$$

Furthermore,

$$g'_\lambda(\zeta_\lambda) < 1$$

because $g'_\lambda(\zeta_\lambda) < 0$ when $\zeta_\lambda > \nu_\lambda$, and when $\zeta_\lambda < \nu_\lambda$ the graph of g_λ intersects the line $y = x$ at ζ_λ going upward from the left to the right, and hence the slope is less than one at the point of intersection.

It follows that we can choose $a, b \in (0,1)$ in such a way that (66) holds true for some $\varepsilon > 0$, and

$$\text{There exists } \varrho \in (0,1) \text{ such that } |g'_\lambda(x)| < \varrho \text{ on } (a,1). \tag{67}$$

From now on assume that the constants $a, b \in (0,1)$ and $\varepsilon > 0$ satisfy (66) and (67). Pick any $\delta \in (0, \varepsilon)$, and assume that for some $n \in \mathbb{N}$ and $k \in \mathbb{Z}_+$, $X_k^{(n)}$ satisfies the following two conditions:

$$\begin{array}{ll} 1. & x_{n,k} \in (a+\delta, b-\delta) \\ 2. & |x_{n,k} - \psi_k| < \delta. \end{array} \tag{68}$$

It follows from (2) that for a given $X_k^{(n)} < bn$,

$$\begin{aligned} P\Big(\Big|x_{n,k+1} - (1-x_{n,k})\Big(1 - q_n^{X_k^{(n)}}\Big)\Big| \geq \frac{(1-\varrho)\delta}{2} \,\Big|\, X_k^{(n)}\Big) &\leq 2e^{-\frac{\delta^2 n(1-x_{n,k})}{2(1-\varrho)^2}} \\ &\leq 2e^{-\frac{\delta^2 n(1-b)}{2(1-\varrho)^2}}, \end{aligned} \tag{69}$$

where in the first step we applied Hoeffding's inequality for binomial variables. Furthermore, if n is large enough, then under the condition (68), we have

$$\Big|g_\lambda(x_{n,k}) - (1-x_{n,k})\Big(1 - q_n^{X_k^{(n)}}\Big)\Big| \leq \frac{(1-\varrho)\delta}{2}$$

and

$$|g_\lambda(x_{n,k}) - g_\lambda(\psi_k)| \leq \varrho\delta,$$

which together imply

$$\Big|g_\lambda(\psi_k) - (1-x_{n,k})\Big(1 - q_n^{X_k^{(n)}}\Big)\Big| \leq \frac{(1-\varrho)\delta}{2} + \varrho\delta \leq \frac{(1+\varrho)\delta}{2}.$$

Combining the last inequality with (69), we obtain that

$$P\Big(|x_{n,k+1} - \psi_{k+1}| \geq \delta \,\Big|\, X_k^{(n)}\Big) \leq 2e^{-\frac{\delta^2 n(1-b)}{2(1-\varrho)^2}}$$

for any $X_{n,k}$ that satisfies condition (68). Taking in account (66), we arrive to the following conclusion:

Lemma 6.8. *If $X_k^{(n)}$ satisfies condition (68), then (conditionally on $X_k^{(n)}$) $X_{k+1}^{(n)}$ satisfies the same condition with probability larger than $2e^{-\frac{\delta^2 n(1-b)}{2(1-\varrho)^2}}$, uniformly on $X_{k+1}^{(n)}$.*

Let $A_{n,m,\delta}$ be the event that (68) is satisfied for $k = 0, 1, \ldots, m$, and recall $\tau_n(\delta)$, from (65). By the Markov property, the lemma implies that under the conditions of the theorem, if (68) is satisfied for $k = 0$, then

$$P(\tau_n(\delta) > m) \geq P(A_{n,m,\delta}) \geq \Big(1 - 2e^{-\frac{\delta^2 n(1-b)}{2(1-\varrho)^2}}\Big)^m \geq 1 - 2me^{-\frac{\delta^2 n(1-b)}{2(1-\varrho)^2}},$$

completing the proof of the theorem. □

The counterpart of Theorem 6.4 for $\lambda \leq 1$ is easier to prove since the behavior of the derivative g'_λ on $[0, 1]$ is "more friendly" in this case, in that $|g\prime\lambda(x)| < 1$ and $g_\lambda(x) < x$ for all $x \in (0, 1)$. For $\varepsilon \in (0, 1)$ let

$$\varsigma_n(\varepsilon) = \inf\{k \in \mathbb{Z}_+ : x_{n,k} < \varepsilon\}. \tag{70}$$

Recall $\tau_n(\delta)$ from (65). Let $\lfloor x \rfloor = \max\{k \in \mathbb{Z} : k \leq x\}$ denote the integer part of the real number x. We have:

Theorem 6.9. *Let Assumption 3.1 hold with $\lambda \leq 1$, and suppose that $x_{n,0} = \psi_0$ for some $\psi_0 \in (0, 1)$ and all $n \in \mathbb{N}$. Pick any $\delta \in (0, 1)$ such that*

$$\delta < \psi_0 - \psi_1,$$

where ψ_k are defined in (61). Let $a \in (0, 1)$ be the minimal root of the equation $g_\lambda(a) = \delta$, and define

$$\varrho = \begin{cases} \max\{\lambda, |g'_\lambda(\psi_0)|\} & \text{if } \lambda < 1, \\ \max\{g'_\lambda(a), |g'_\lambda(\psi_0)|\} & \text{if } \lambda = 1. \end{cases} \tag{71}$$

Then the following holds true.

(i) Let $m_0 = \lfloor \log \frac{\psi_0}{\delta} \rfloor + 1$. Then for any $n \in \mathbb{N}$,

$$P(\varsigma_n(\delta) > m_0) \geq \Big(1 - 2e^{-\frac{\delta^2 n(1-\psi_0)}{2(1-\varrho)^2}}\Big)^{m_0} \geq 1 - 2m_0 \exp\Big\{-\frac{\delta^2 n(1-\psi_0)}{2(1-\varrho)^2}\Big\}.$$

(ii) If $\lambda \in (0, 1)$, then for any $n, m \in \mathbb{N}$,

$$P(\tau_n(\delta) > m) \geq \Big(1 - 2e^{-\frac{\delta^2 n(1-\psi_0)}{2(1-\varrho)^2}}\Big)^m \geq 1 - 2m \exp\Big\{-\frac{\delta^2 n(1-\psi_0)}{2(1-\varrho)^2}\Big\}.$$

Proof. Observe that (67) holds with ϱ introduced in (71). Furthermore, ψ_k is a decreasing sequence since $g_\lambda(x) < x$ for $x \in (0, 1)$ when $\lambda \leq 1$. The rest of the proof is similar to that of Theorem 6.7, namely the induction argument based on (68) and Lemma 6.8 carries over. Note that we can replace $1 - b$ by $1 - \psi_0 = 1 - \varepsilon_0$ in the conclusions of the lemma because the sequence ψ_k is monotone decreasing under the conditions of the theorem. □

References

[1] A. Agresti, *Bounds on the extinction time distribution of a branching process*, Adv. in Appl. Probab. **6** (1974), 322–335.

[2] H. Andersson, *Limit theorems for a random graph epidemic model*, Ann. Appl. Probab. **8** (1998), 1331–1349.

[3] K. B. Athreya, *On the maximum sequence in a critical branching processes*, Ann. Probab. **16** (1988), 502–507.

[4] K. B. Athreya, *Large deviation rates for branching processes–I. Single type case*, Ann. Appl. Probab. **4** (1994), 779–790.

[5] O. Aydogmus, *On extinction time of a generalized endemic chain-binomial model*, Math. Biosci. **279** (2016), 38–42.

[6] N. T. J. Bailey, *The Mathematical Theory of Infectious Diseases and Its Applications* (2nd edn.), Hafner Press/Macmillan Publishing Co., New York, 1975.

[7] A. D. Barbour, *Couplings for locally branching epidemic processes*, J. Appl. Probab. **51** (2014), 43–56.

[8] A. D. Barbour, L. Holst, and S. Janson, *Poisson Approximation (Oxford Studies in Probability, Vol. 2)*. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1992.

[9] B. Bollobás, *Random Graphs* (2nd ed.), Cambridge University Press, 2001.

[10] F. M. Buckley and P. K. Pollett, *Limit theorems for discrete-time metapopulation models*, Probab. Surv. **7** (2010), 53–83.

[11] S. Chatterjee, P. Diaconis, and E. Meckes, *Exchangeable pairs and Poisson approximation*, Probab. Surv. **2** (2005), 64–106.

[12] T. Chumley, O. Aydogmus, A. Matzavinos, and A. Roitershtein, *Moran-type bounds for the fixation probability in a frequency-dependent Wright-Fisher model*, J. Math. Biol. **76** (2018), 1–35.

[13] L. Cocchi, L. L. Gollo, A. Zalesky, and M. Breakspear, *Criticality in the brain: A synthesis of neurobiology, models and cognition*, Prog. Nerobiol. **158** (2017), 132–152.

[14] K. L. Cooke, D. F. Calef, and E. V. Level, *Stability or chaos in discrete epidemic models*, In: V. Lakshmikantham (Ed.), *Nonlinear Systems and Applications*, Academic Press, New York, 1977, pp. 73–93.

[15] R. Durrett, *Probability Models for DNA Sequence Evolution*, 2nd ed., Springer Series in Probability and its Applications, Springer, New York, 2010.

[16] M. Dwass, *The total progeny in a branching process and a related random walk*, J. Appl. Probab. **6** (1969), 682–686.

[17] J. P. Gleeson, *Mean size of avalanches on directed random networks with arbitrary degree distributions*, Phys. Rev. E **77** (2008), 057101.

[18] J. P. Gleeson and R. Durrett, *Temporal profiles of avalanches on networks*, Nat. Commun. **8** (2017), 1227.

[19] A. F. Karr, *Weak convergence of a sequence of Markov chains*, Z. Wahrsch. Verw. Gebiete **33** (1975), 41–48.

[20] C. T. Kello, *Critical branching neural networks*, Psychol. Rev. **120** (2013), 230–254.

[21] O. Kinouchi and M. Copelli, *Optimal dynamical range of excitable networks at criticality*, Nat. Phys. **2** (2006), 348.

[22] F. C. Klebaner and O. Nerman, *Autoregressive approximation in branching processes with a threshold*, Stochastic Process. Appl. **51** (1994), 1–7.

[23] D. B. Larremore, M. Y. Carpenter, E. Ott, and J. G. Restrepo, *Statistical properties of avalanches in networks*, Phys. Rev. E **85** (2012), 066131.

[24] D. B. Larremore, W. L. Shew, and J. G. Restrepo, *Predicting criticality and dynamic range in complex networks: effects of topology*, Phys. Rev. Lett. **106** (2011), 058101.

[25] D. B. Larremore, W. L. Shew, E. Ott, and J. G. Restrepo, *Effects of network topology, transmission delays, and refractoriness on the response of coupled excitable systems to a stochastic stimulus*, Chaos **21** (2011), 025117.

[26] T. Lindvall, *On the maximum of a branching process*, Scand. J. Statist. **3** (1976), 209–214.

[27] I. M. Longini Jr., *A chain binomial model of endemicity*, Math. Biosci. **50** (1980), 85–93.

[28] R. Lyons, R. Pemantle, and Y. Peres, *Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes*, Ann. Probab. **23** (1995), 1125–1138.

[29] S. A. Moosavi, A. Montakhab, and A. Valizadeh, *Refractory period in network models of excitable nodes: self-sustaining stable dynamics, extended scaling region and oscillatory behavior*, Scientific Reports **7** (2017), 7107.

[30] M. A. Mūnoz, *Colloquium: Criticality and dynamical scaling in living systems*, Rev. Mod. Phys **90** (2018), 031001.

[31] O. Nerman, *On the maximal generation size of a non-critical Galton-Watson process*, Scand. J. Statist. **4** (1977), 131–135.

[32] P. Rämö, J. Kesseli, and O. Yli-Harja, *Perturbation avalanches and criticality in gene regulatory networks*, J. Theoret. Biol. **242** (2006), 164–170.

[33] P. Rämö, S. Kauffman, J. Kesseli, and O. Yli-Harja, *Measures for information propagation in Boolean networks*, Phys. D **227** (2007), 100–104.

[34] R. Rastegar and A. Roitershtein, *Duration of avalanches in an excitable network*, in preparation.

[35] R. Rastegar and A. Roitershtein, *Approximation schemes for avalanches in a complex network*, work in progress.

[36] H. Thorisson, *Coupling, Stationarity, and Regeneration*, Springer, New-York, 2000.

[37] C. H. Weiß and P. K. Pollett, *Chain binomial models and binomial autoregressive processes*, Biometrics **68** (2012), 815–824.

[38] S. Zapperi, K. B. Lauritsen, and H. E. Stanley, *Self-organized branching processes: mean-field theory for avalanches*, Phys. Rev. Lett. **75** (1995), 4071.