Supplemental Material : "Classes of critical avalanche dynamics in complex networks."

Filippo Radicchi,[1] Claudio Castellano,[2] Alessandro
Flammini,[1] Miguel A. Muñoz,[3] and Daniele Notarmuzi[1]

[1]*Center for Complex Networks and Systems Research,
Luddy School of Informatics, Computing, and Engineering,
Indiana University, Bloomington, Indiana 47408, USA*[∗]
[2]*Istituto dei Sistemi Complessi (ISC-CNR), Via dei Taurini 19, I-00185 Roma, Italy*
[3]*Departamento de Electromagnetismo y Física de la Materia e Instituto Carlos I de Física Teórica
y Computacional. Facultad de Ciencias. Universidad de Granada. E-18071, Granada, Spain*

∗ filiradi@indiana.edu

I. THE BRANCHING PROCESS

The standard Branching Process (BP), also called Galton-Watson process, [7, 10, 14, 31] is a stochastic (Markov) process that describes the evolution of the non-negative integer $X(t)$ at discrete times $t = \{0, 1, 2, ...\}$ as

$$X(t+1) = \sum_{i=1,...X(t)} x_i(t) .$$

The initial condition is $X(0) = 1$; $x_i(t)$, $\forall i, t \in \mathbb{N}$ are independent and identically distributed random variables taking integer values. These random variables are extracted from some arbitrary probability distribution function $P(x)$, with $x \in \mathbb{N}$. This process can be interpreted as a contagion process in a directed tree, where, starting from one infected seed node, the epidemic propagates to a random number of nearest neighbors, branching out from the seed node. A single instance of BP is, in the standard jargon, named avalanche. The integer stochastic variable X can reach at any time step – with some probability – the absorbing-state value $X = 0$, from which it cannot possibly escape. The time T when the system reaches the absorbing state $X = 0$ is the duration of the avalanche. The total number S of infected individuals before reaching the absorbing state $X = 0$ is instead the size of the avalanche

$$S = \sum_{t=0,1,...T} X(t). \tag{1}$$

The BP is critical if the average value of $P(x)$ is equal to unity, so that for each infected node there is, on average, one infection, so that there is no intrinsic tendency for the overall number of infected nodes at a given time to either increase or decrease. In the interpretation of BP as a spreading process on a directed tree, if for simplicity we assume that each node has k^{out} out-going connections, the critical regime is obtained by setting the spreading probability along individual edges equal to $1/k^{out}$. At criticality, size and duration probability distribution functions of avalanches can be written as

$$\begin{aligned} P(S) &\sim S^{-\tau} \mathcal{G}_{\mathcal{S}}(S/S_C) \\ F(T) &\sim T^{-\alpha} \mathcal{G}_{\mathcal{T}}(T/T_C) \end{aligned}, \tag{2}$$

with $\tau = 3/2$ and $\alpha = 2$, where $\mathcal{G}_{\mathcal{S}}(S/S_C)$ and $\mathcal{G}_{\mathcal{T}}(T/T_C)$ are cut-off (scaling) functions, and the cut-off scales, S_C and T_C, depend only on the system size (in the case it is finite) [12]. Moreover, the average avalanche size scales with its duration as $\langle S \rangle \sim T^\chi$, where the exponent χ needs to obey the general scaling relation [1, 26]

$$\chi = \frac{\alpha - 1}{\tau - 1} , \tag{3}$$

and thus $\chi = 2$. A particularly simple proof of these results for the case in which the underlying tree is homogeneous with $k^{out} = 2$ can be found in Ref. [6]. A more systematic derivation –for different types of underlying regular or random tree topologies– can be obtained within the generating-function formalism [22, 24, 32]; for instance, already back in 1949, Otter computed the solution for the case of a Poissonian distribution of branches per node [19].

The exponent values $\tau = 3/2$ and $\alpha = 2$ are extremely universal and robust; they emerge in many different types of propagation processes such as directed percolation, the contact process, the voter model, susceptible-infected-susceptible model, susceptible-infected-recovered model, and many others, as long as the underlying pattern of connections is either a high-dimensional lattice or a sufficiently homogeneous network [11, 17, 18]. In Ref. [6], this super-universality is explained using an effective Langevin equation, called the demographic random walker.

Although MF critical exponents appear almost everywhere, some papers [8, 9, 25] pointed out that there is a simple way to break this extremely robust type of scaling: it suffices to take a distribution $P(x)$ of possible offspring values with divergent second moment, i.e., a branching number distributed as a power law with exponent $2 < \gamma < 3$. Specifically, if

$$P(x) = \frac{1}{Z(\gamma - 1)} \begin{cases} x^{-\gamma} & \text{, for } x \geq 1 \\ Z(\gamma - 1) - \sum_{x=1}^{\infty} x^{-\gamma} & \text{, for } x = 0 \end{cases}, \tag{4}$$

where $Z(\cdot)$ is the Riemann Zeta function, then it can be shown [8, 9, 25] that critical exponents τ and α become functions of γ as

$$
\begin{aligned}
\tau &= \gamma/(\gamma - 1) \\
\alpha &= (\gamma - 1)/(\gamma - 2).
\end{aligned}
\tag{5}
$$

The expressions in Eq. (5) converge to the standard values $3/2$ and 2, respectively, when $\gamma \to 3$, i.e., when the second moment of the offspring distribution becomes finite. Critical exponents exhibit anomalous values for smaller values of γ; further, logarithmic corrections to scaling are observed at the marginal value $\gamma = 3$ [9]. Observe also that the anomalous-exponent values are larger than the standard ones, implying that the probability of observing large avalanches is smaller in the anomalous case, despite of the probability of having nodes with very large number of offsprings. Intuitively, this observation stems from the fact that as large numbers of offspring are possible, this implies that –in order to keep the average value equal to unity– the probability to have zero offspring must be much larger than in the standard case, causing then avalanches to die earlier and to be smaller on average. The existence of anomalous values for the critical exponents of avalanches can be also understood in terms of the effective Langevin-equation approach [29]. Such an approach includes a standard Gaussian noise, stemming from the central limit theorem for the addition of stochastic variables with finite variance; however, if the variance is not finite, the Gaussian noise needs to be replaced by a Levy-stable distribution, leading to different results [29].

II.   NETWORKS

We consider unweighted graphs of size N and topology specified by the adjacency matrix A. The generic element $A_{ij} = 1$ if a directed connection $i \to j$ exists, whereas $A_{ij} = 0$, otherwise. We consider networks with no self-loops, so that $A_{ii} = 0$ for all i. The in-degree $k_i^{in} = \sum_j A_{ji}$ of node i is defined as the total number of connections pointing to node i. Similarly, the out-degree $k_i^{out} = \sum_j A_{ij}$ of node i is defined as the total number of connections departing from node i.

Undirected networks are described in the same exact way as above, with the only difference that the adjacency matrix A is such that $A_{ij} = A_{ji}$, for all pairs of nodes i and j. Clearly, no distinction between in-coming and out-going connections is possible in the case of undirected networks. We just indicate with $k_i = \sum_j A_{ij}$ the degree of node i.

In our analysis, we consider the following types of networks.

A.   Synthetic directed networks

We generate instances of a model that has been considered in the context of avalanche dynamics by Gleeson et al. [8]. In the construction of the model, one imposes the out-degree for every single node in the network. This information is contained in the imposed out-degree sequence $\{k_1^{out}, k_2^{out}, \ldots, k_N^{out}\}$, with average value equal to $\langle k^{out} \rangle$. For every node i, k_i^{out} directed connections are created by choosing k_i^{out} nodes uniformly at random among all nodes in the network. The resulting in-degree distribution $P(k^{in})$ is a Poisson distribution with average equal to $\langle k^{in} \rangle = \langle k^{out} \rangle$. In our analysis, we generate out-degree sequences composed of random variates extracted from the power-law distribution $P(k^{out}) \sim [k^{out}]^{-\gamma}$ for $k^{out} \in [4, k_{max}^{out}]$ and $P(k^{out}) = 0$, otherwise. In our analysis, we use $\gamma = 2.1$ and two distinct max values for the out-degree, i.e., $k_{max}^{out} = N - 1$ and $k_{max}^{out} = \sqrt{N}$. For both k_{max}^{out} values, we generate a single instance of the network model for each of the following network sizes: $N = 10^3$, $N = 10^4$, $N = 10^5$, $N = 10^6$, and $N = 10^7$. We use those single instances in all our analyses. Notice $k_{max}^{out} = N - 1$ is a hard cutoff for the degree distribution, but in each realization of the degree sequence the maximum out-degree is actually of the order of $N^{1/(\gamma - 1)} \ll N$.

B.   Synthetic undirected networks

We create synthetic undirected networks using the configuration model [16]. The model is very similar to the one described in the previous section. The model requires that a degree sequence $\{k_1, k_2, \ldots, k_N\}$ is imposed. Connections are randomly created by selecting pairs of stubs. In our implementation of

Network	Type	N	M	k_{max}^{out}	ν_1	ω_N	ω_{N-1}	θ_{2M}	Ref.	Url
AS Caida	Undirected	26,475	53,381	2,628	0.020	69.64	51.13	59.41	[13]	url
Twitter 15M	Directed	82,253	5,890,705	30,825	0.248	384.82	181.90	655.51	[3]	url
Youtube	Directed	509,332	4,272,379	27,286	0.012	208.41	161.85	184.93	[15]	url

Table I. Summary table of the real-world networks considered in our analysis. From left to right, we report: name of the network, type of network, number of nodes N, number of edges M, maximal value of the outdegree k_{max}^{out}, second smallest eigenvalue ν_1 of the graph Laplacian, largest ω_N and second largest ω_{N-1} eigenvalues of the adjacency matrix, largest eigenvalue θ_{2M} of the non-backtracking matrix, reference to the paper where the dataset has been first considered, url of the website where we downloaded the dataset. The numerical values appearing in the table refer to the largest strongly connected component of the graphs.

the model, we discard eventual multiple connections and self-loops appearing in the graph; this means that the imposed degree sequence may not be exactly reproduced in the resulting instance of the model. For simplicity, we restrict our attention to cases where effective values of node degrees do not differ much from those initially imposed in the creation of the model instances. Specifically, we consider the case where the imposed degrees are random variates taken from the power-law distribution $P(k) \sim k^{-\gamma}$ for $k \in [4, \sqrt{N}]$ and $P(k) = 0$, otherwise. Setting the maximal value of the degree to $\sqrt{N}$ leads to a negligible fraction of discarded multiple connections and self-loops during the generation of individual instances of the model [4]. In our analysis, we use $\gamma = 2.1$ to generate a single instance of the network model for each of the following network sizes: $N = 10^3$, $N = 10^4$, $N = 10^5$, and $N = 10^7$. We use those single instances in all our analyses.

C. Real networks

We consider three real-world networks: (a) an undirected graph representing a snapshot of the Internet at the Autonomous system level [13]; (b) the directed Twitter network of the Spanish 15M movement [3]; (c) a directed graph representing a portion of the Youtube social network [15]. A summary of the topological features that regard the largest strongly connected component of the various networks is provided in Table I. The degree distributions of the networks are displayed in Figure 1. We find that the outdegree distribution of all networks is compatible with $P(k^{out}) \sim [k^{out}]^{-\gamma}$ with $\gamma = 2.1$ for all networks.
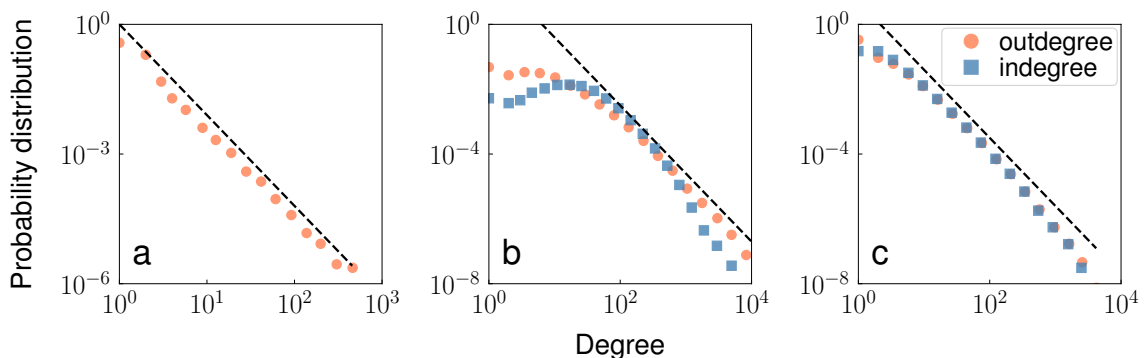

Figure 1. Degree distribution of the three real-world networks considered in our analysis: (a) a snapshot of the Internet at the Autonomous system level [13]; (b) the Twitter network of the Spanish 15M movement [3]; (c) a portion of the Youtube social network [15]. For the two directed networks, we plot the distributions of the out- and indegrees. In the various panels, the black dashed lines indicate power-law decay with exponent $\gamma = 2.1$.

III. MODELS FOR AVALANCHE DYNAMICS

We consider different models originating avalanche dynamics in networks. Regardless of the specific model considered, we focus our attention on the size S and duration T of the avalanches generated by individual randomly chosen nodes. In the following, we describe the various models and their implementation in our numerical simulations. For every model, we provide definitions of avalanches, and avalanche duration and size.

Together with this paper, we release source code with the implementation of all the processes described below. The code is available at http://homes.sice.indiana.edu/filiradi/resources.html.

A. Competition-Induced-Criticality (CIC) model

We consider the model introduced by Gleeson *et al.* to describe meme dynamics in networks [8]. The model runs on a directed and unweighted network with adjacency matrix A and total number of nodes N. The following description and implementation of the model are identical on both directed and undirected networks.

The state $\sigma_i(t)$ of node i at time t is a discrete-valued variable corresponding to the arbitrary index of a meme. At each elementary step of the dynamical process, one node i is selected at random. With probability μ, the state of node i changes to a new value not already present in the network. This step corresponds to the creation of a new meme by node i. Then, regardless of the fact that a new meme was created or not by node i, the meme of node i is shared to all its neighbors. Sharing means that we set $\sigma_j(t + dt) = \sigma_i(t)$, for all j such that $A_{ij} = 1$. The state of all other nodes not involved in the elementary dynamical step remains unchanged. A reasonably good definition for the time increment dt is based on the assumption that the dynamics is given by a standard Poisson process where the rate of sharing events per unit of time equals $1/N$, so that dt is a random variate extracted from the exponential distribution $P(dt) \sim \exp[-dt/N]$. Alternatively, one can assume that exactly one sharing event happens at every regular interval of time $dt = 1/N$.

Two basic meme-based measures are of interest in the model: 1) the popularity S of the meme, corresponding to the total number of times a meme is actually shared; 2) the duration or lifetime T of the meme, corresponding to the total amount of time that a meme is present in the network before it is forgotten. Both these quantities are considered for memes originated from a single randomly chosen node in the network, and they are interpreted in terms of BP observables: S represents the size of the avalanche generated by the meme in the network; T is instead its duration. Please note that on a finite network and for $\mu = 0$, after a sufficiently long time, a single meme will occupy the states of all nodes in the network for ever. This event corresponds to an avalanche of infinite size and duration. Infinite avalanches are excluded from our analysis.

An *ad litteram* implementation of the above model is computationally expensive. The reason is that one needs to account for the simultaneous dynamics of a multitude of competing memes. However, given that the focus is on the properties of avalanches associated with individual memes, one can greatly speed-up computations required for the simulation of the process by looking at the dynamics of one meme at a time. Assuming that we are observing a single meme only, the state of the generic node i is a *de facto* a binary variable: $\sigma_i(t) = 1$ means that node i is occupied by the observed meme at time t; $\sigma_i(t) = 0$ indicates that node i is not occupied by the observed meme at time t. On the basis of the configuration of the system at time t and the topology of the network, we can divide the nodes in three disjoint sets: the set $\mathcal{C}(t)$ composed of nodes in state 1; the set $\mathcal{F}(t)$ composed of nodes in state 0, but having at least one connection pointing to a node in the set $\mathcal{C}(t)$; the set $\mathcal{O}(t)$ composed of all other nodes. As nodes in the set $\mathcal{O}(t)$ do not contribute to dynamics of the observed meme directly, their actual evolution doesn't need to be explicitly simulated by the algorithm. Only three elementary events that lead to the change in the total number of nodes having state equal to 1 can happen:

1. A randomly selected node in the set $\mathcal{C}(t)$ shares the meme with its neighbors. This event happens with probability $(1-\mu)|\mathcal{C}(t)|/(|\mathcal{C}(t)| + |\mathcal{F}(t)|)$, where we indicated with $|\mathcal{C}(t)|$ and $|\mathcal{F}(t)|$ the size of the sets $\mathcal{C}(t)$ and $\mathcal{F}(t)$, respectively. This event corresponds to the sharing of the observed meme, thus popularity increases as $S \to S + 1$.

2. A randomly selected node in the set $\mathcal{C}(t)$ creates a new meme, thus its state changes to 0; the new meme is shared with the neighbors of the selected node. Please note that some of the neighbors of

the selected node can be part of the set $\mathcal{C}(t)$. This event happens with probability $\mu \, |\mathcal{C}(t)|/(|\mathcal{C}(t)| + |\mathcal{F}(t)|)$.

3. A randomly selected node in the set $\mathcal{F}(t)$ shares a meme to other nodes in the set $\mathcal{C}(t)$. This event happens with probability $|\mathcal{F}(t)|/(|\mathcal{C}(t)| + |\mathcal{F}(t)|)$.

After one of these events happens, time increases as $t \to t + dt$. Please note that the new sets $\mathcal{C}(t + dt)$ and $\mathcal{F}(t + dt)$ can be updated with local operations only. The simulation starts by building the initial sets $\mathcal{C}(0)$ and $\mathcal{F}(0)$ depending on the initial configuration $\sigma_i(t = 0)$ for all i. Our initial conditions are such that for all nodes i we have $\sigma_i(t = 0) = 0$. Then, one random node i is selected for the creation of the observed meme, i.e., $\sigma_i(t = 0) = 1$; the creation of the meme is followed by a first sharing event to all neighbors of the initial node, so that $\sigma_j(t = 0) = 1$ for all j such that $A_{ij} = 1$. Also, the initial value of popularity is $S = 1$. The simulation then proceeds until $\mathcal{C}(t = T) = \emptyset$. As mentioned earlier, there are potentially two slightly different ways of defining the time unit in the model. Assuming that the dynamics is described by a Poisson process, the increment of time dt considered at each of the above steps is a random variable extracted from the exponential distribution $P(dt) \propto \exp\left[-dt/\tau\right]$, with $\tau = |\mathcal{C}(t)| + |\mathcal{F}(t)|$. Alternatively, one can consider a time increment $dt = R/N$, where R is a random variable extracted from the geometric distribution with probability of success equal to $(|\mathcal{C}(t)| + |\mathcal{F}(t)|)/N$. This second scenario corresponds to the case in which exactly one sharing event is performed at every regular interval of time $1/N$.

B.   Voter model (VOT)

This model is very similar to the CIC model described above. The only difference regards how sharing is performed. The description and implementation of the model on undirected networks is identical to the one provided below for directed networks.

A generic node i can be found at time t in two possible states, either $\sigma_i(t) = 1$ or $\sigma_i(t) = 0$. Each elementary step of the dynamics consists in choosing a node i at random; then, one in-coming connection of node i is chosen at random, with node j found on the other side of the edge. Sharing happens following the direction of the connection, so that $\sigma_i(t + dt) = \sigma_j(t)$. The state of all other nodes remains unchanged. Also in the voter model, one can define the increment of time dt in the two ways, as explained above for the CIC model.

For simplicity, we use an implementation of the model very similar to the one described for the CIC model. We still rely on the three sets $\mathcal{C}(t)$, $\mathcal{F}(t)$, and $\mathcal{O}(t)$, defined exactly as in the CIC model. For the voter model, possible events that may lead to a change of the nodes in state 1 are:

1. A node i in the set $\mathcal{C}(t)$ is randomly selected. One random neighbor j is selected too. This event happens with probability $|\mathcal{C}(t)|/(|\mathcal{C}(t)| + |\mathcal{F}(t)|)$. The state of node i evolves as $\sigma_i(t + dt) = \sigma_j(t)$. If the state of node j is $\sigma_j(t) = 1$, then popularity increases as $S \to S + 1$.

2. A node i in the set $\mathcal{F}(t)$ is randomly selected. One random neighbor j is selected too. The state of node i evolves as $\sigma_i(t + dt) = \sigma_j(t)$. If the state of node j is $\sigma_j(t) = 1$, then popularity increases as $S \to S + 1$. This event happens with probability $|\mathcal{F}(t)|/(|\mathcal{C}(t)| + |\mathcal{F}(t)|)$.

After one of these events happens, time increases as $t \to t + dt$. Please note that the new sets $\mathcal{C}(t + dt)$ and $\mathcal{F}(t + dt)$ can be obtained by updating $\mathcal{C}(t)$ and $\mathcal{F}(t)$ with local operations only. The simulation starts by building the initial sets $\mathcal{C}(0)$ and $\mathcal{F}(0)$. Our initial conditions are such that one random node i is selected. The state of node i is set as $\sigma_i(t = 0) = 1$; all other nodes $j \neq i$ are instead in the state $\sigma_j(t = 0) = 0$. Popularity is initially set as $S = 1$. The simulation then proceeds until $\mathcal{C}(t = T) = \emptyset$. Increments of time dt can be defined exactly as in the case of the CIC model.

C.   Invasion process (IP)

The invasion process is a sort of inverse version of the voter model. Still, node i can influence node j only if a directed connection $i \to j$ exists in the network. The description and implementation of the models on undirected networks is identical to one provided below for directed networks.

A generic node i can be found at time t in two possible states, either $\sigma_i(t) = 1$ or $\sigma_i(t) = 0$. Each elementary step of the dynamics consists in choosing a node i at random; then, one out-going connection of node i is chosen at random, with node j found at the end of the edge. Sharing goes in the same direction as the one indicated by the edge, so that $\sigma_j(t + dt) = \sigma_i(t)$. The state of all other nodes remains unchanged. Also here, one can define the increment of time dt as in the CIC model.

For simplicity, we use an implementation of the model very similar to the one described for the CIC and the voter models. We still rely on the three sets $\mathcal{C}(t)$, $\mathcal{F}(t)$, and $\mathcal{O}(t)$. The set $\mathcal{F}(t)$ is now composed of all nodes in state 0 with at least one in-coming connection from nodes in the set $\mathcal{C}(t)$. There are only two possible events that may lead to a change in the number of the nodes in state 1. These are:

1. A node i in the set $\mathcal{C}(t)$ is randomly selected. One random neighbor j is selected too. This event happens with probability $|\mathcal{C}(t)|/(|\mathcal{C}(t)| + |\mathcal{F}(t)|)$. The state of node j evolves as $\sigma_j(t + dt) = \sigma_i(t)$. Popularity increases as $S \to S + 1$.

2. A node i in the set $\mathcal{F}(t)$ is randomly selected. One random neighbor j is selected too. The state of node j evolves as $\sigma_j(t + dt) = \sigma_i(t)$. This event happens with probability $|\mathcal{F}(t)|/(|\mathcal{C}(t)| + |\mathcal{F}(t)|)$.

After one of these events happens, time increases as $t \to t + dt$. Please note that the new sets $\mathcal{C}(t + dt)$ and $\mathcal{F}(t + dt)$ can be obtained by updating $\mathcal{C}(t)$ and $\mathcal{F}(t)$ with local operations only. The simulation starts by building the initial sets $\mathcal{C}(0)$ and $\mathcal{F}(0)$. Our initial conditions are such that one random node i is selected. The state of node i is set as $\sigma_i(t = 0) = 1$; all other nodes $j \neq i$ are instead in the state $\sigma_j(t = 0) = 0$. Popularity is initially set as $S = 1$. The simulation then proceeds until $\mathcal{C}(t = T) = \emptyset$. Increments of time dt are defined exactly as in the case of the CIC and the voter model.

D. Link dynamics (LD)

We consider also link dynamics, a model with characteristics similar to those of the voter model and the invasion process. The description and implementation of the model on undirected networks is identical to one provided below for directed networks.

A generic node i can be found at time t in two possible states, either $\sigma_i(t) = 1$ or $\sigma_i(t) = 0$. Each elementary step of the dynamics consists in choosing an edge at random, say $i \to j$. Then the state of node j evolves as $\sigma_j(t + dt) = \sigma_i(t)$. The state of all other nodes remains unchanged. Also here, one can define the increment of time dt as in the CIC model.

For simplicity, we use an implementation of the model very similar to those considered for the voter model and the invasion process. We rely on the three sets $\mathcal{C}(t)$, $\mathcal{F}(t)$, and $\mathcal{O}(t)$. In this case, the sets do not contain nodes but edges. In particular, the set $\mathcal{C}(t)$ is the set of all edges having both nodes at the ends of the edge in state 1; the set $\mathcal{F}(t)$ is the set of all edges having one node at the end of the edge in state 1 and the other in state 0; $\mathcal{O}(t)$ is the set of all other edges, where both nodes at the ends of the edge are in state 0. There are only two possible events that contribute to the evolution of the avalanche. These are:

1. An edge $\mathbf{e} = i \to j$ is chosen at random from the set $\mathcal{C}(t)$. This happens with probability $|\mathcal{C}(t)|/(|\mathcal{C}(t)| + |\mathcal{F}(t)|)$. Popularity increases as $S \to S + 1$.

2. An edge $\mathbf{e} = i \to j$ is chosen at random from the set $\mathcal{F}(t)$. This happens with probability $|\mathcal{F}(t)|/(|\mathcal{C}(t)| + |\mathcal{F}(t)|)$. The state of node j evolves as $\sigma_j(t + dt) = \sigma_i(t)$. If the state of node i is $\sigma_i(t) = 1$, then popularity increases as $S \to S + 1$.

After one of these events happens, time increases as $t \to t + dt$. Please note that the new sets $\mathcal{C}(t + dt)$ and $\mathcal{F}(t + dt)$ can be obtained by updating $\mathcal{C}(t)$ and $\mathcal{F}(t)$ with local operations only. The simulation starts by building the initial sets $\mathcal{C}(0)$ and $\mathcal{F}(0)$. Our initial conditions are such that one random node i is selected. The state of node i is set as $\sigma_i(t = 0) = 1$; all other nodes $j \neq i$ are instead in the state $\sigma_j(t = 0) = 0$. Popularity is initially set as $S = 1$. The simulation then proceeds until $\mathcal{C}(t = T) = \emptyset$. Increments of time dt are defined exactly as in the case of the CIC and the voter model.

E. Susceptible-Infected-Susceptible (SIS) model

In the case of the SIS model, our implementation coincides with the following description. The state of node i at time t can assume two values: $\sigma_i(t) = 1$ meaning that the node is infected; $\sigma_i(t) = 0$ meaning

that the node is susceptible. At each time t, we consider two sets: $\mathcal{C}(t)$ is the set of all nodes in state 1; $\mathcal{E}(t)$ is the set of all directed edges $\mathbf{e} = i \to j$ such that $i \in \mathcal{C}(t)$ and $j \notin \mathcal{C}(t)$. At time t, one the two events may happen:

1. With probability $\lambda|\mathcal{E}(t)|/[\lambda|\mathcal{E}(t)| + |\mathcal{C}(t)|]$, a random edge $\mathbf{e} = i \to j$ is selected from the set $\mathcal{E}(t)$. Node j changes its state as $\sigma_j(t + dt) = 1$, and popularity increases as $S \to S + 1$.

2. With probability $|\mathcal{C}(t)|/[\lambda|\mathcal{E}(t)| + |\mathcal{C}(t)|]$, a random node i is selected from the set $\mathcal{C}(t)$. Node i changes its state, i.e., $\sigma_i(t + dt) = 0$.

After one of these two events happened, the sets $\mathcal{C}(t)$ and $\mathcal{E}(t)$ are updated with local operations only. The initial configuration considered in our simulations consists in choosing a random node i and set $\sigma_i(t = 0) = 1$; for all $j \neq i$ instead we have $\sigma_j(t = 0) = 0$. The size of the avalanche is initially set as $S = 1$. Every simulation proceeds until $\mathcal{C}(t = T) = \emptyset$, with T duration of the avalanche. At each stage of the algorithm, time increases by the increment dt. This is chosen to be a random variable extracted from the exponential distribution $P(dt) \propto \exp\left[-dt/(\lambda|\mathcal{E}(t)| + |\mathcal{C}(t)|)\right]$. This consists in assuming that infections happen at rate λ per unit time, and one infected node becomes susceptible per unit of time.

F. Susceptible-Infected-Recovered (SIR) model

The description, and implementation, of the SIR model is very similar to the one of the SIS model. The main difference is that the state of node i at time t can assume three values: $\sigma_i(t) = 1$ meaning that the node is infected; $\sigma_i(t) = 0$ meaning that the node is susceptible; $\sigma_i(t) = 2$ meaning that the node is recovered. At each time t, we consider two sets: $\mathcal{C}(t)$ is the set of all nodes in state 1; $\mathcal{E}(t)$ is the set of all directed edges $\mathbf{e} = i \to j$ such that $i \in \mathcal{C}(t)$ and $j \notin \mathcal{C}(t)$ and $\sigma_j(t) = 0$. At time t, one the two events may happen:

1. With probability $\lambda|\mathcal{E}(t)|/[\lambda|\mathcal{E}(t)| + |\mathcal{C}(t)|]$, a random edge $\mathbf{e} = i \to j$ is selected from the set $\mathcal{E}(t)$. Node j changes its state as $\sigma_j(t + dt) = 1$, and popularity increases as $S \to S + 1$.

2. With probability $|\mathcal{C}(t)|/[\lambda|\mathcal{E}(t)| + |\mathcal{C}(t)|]$, a random node i is selected from the set $\mathcal{C}(t)$. Node i recovers, i.e., $\sigma_i(t + dt) = 2$.

After one of these two events happened, the sets $\mathcal{C}(t)$ and $\mathcal{E}(t)$ are updated with local operations only. The initial configuration considered in our simulations consists in choosing a random node i and set $\sigma_i(t = 0) = 1$; for all $j \neq i$ instead we have $\sigma_j(t = 0) = 0$. The size of the avalanche is initially set as $S = 1$. Every simulation proceeds until $\mathcal{C}(t = T) = \emptyset$, with T duration of the avalanche. At each stage of the algorithm, time increases by the increment dt. This is chosen to be a random variable extracted from the exponential distribution $P(dt) \propto \exp\left[-dt/(\lambda|\mathcal{E}(t)| + |\mathcal{C}(t)|)\right]$. This consists in assuming that infections happen at rate λ per unit time, and one infected node on average recovers per unit of time.

G. Contact process (CP)

In the case of the contact process, our implementation of the model is an *ad litteram* adaptation of the following description. The state of node i at time t can assume two values: $\sigma_i(t) = 1$ meaning that the node is infected; $\sigma_i(t) = 0$ meaning that the node is susceptible. At time of t, a node i is extracted at random from the set of infected nodes $\mathcal{C}(t)$. Then, one the two following events happens:

1. With probability $\lambda/(1 + \lambda)$, a random connection of node i is selected. If j is the index of the node at the end of the selected connection, then the state of that node becomes $\sigma_j(t + dt) = \sigma_i(t)$. If $\sigma_j(t) = 0$, then popularity increases as $S \to S + 1$.

2. With probability $1/(1 + \lambda)$, the node i changes its state, i.e., $\sigma_i(t + dt) = 0$.

After one of these two events happened, the set $\mathcal{C}(t)$ is updated with local operations only. The initial configuration considered in our simulations consists in choosing a random node i and set $\sigma_i(t = 0) = 1$; for all $j \neq i$ instead we have $\sigma_j(t = 0) = 0$. The size of the avalanche is initially set as $S = 1$. Every simulation proceeds until $\mathcal{C}(t = T) = \emptyset$, with T duration of the avalanche. The increment of time dt is chosen as a random variable from the exponential distribution $P(dt) \propto \exp\left[-dt/|\mathcal{C}(t)|\right]$. Alternatively, one can assume that $dt = R/N$, with R random variate from a geometric distribution with success probability $|\mathcal{C}(t)|/N$.

IV. RESULTS OF NUMERICAL SIMULATIONS

We performed numerical simulations of all the processes described above on different networks topologies. For each network and dynamical model, we always measured the size S and the duration T of avalanches seeded by a single randomly chosen node. Dynamical processes are set as close as possible to their critical regime. Specifically, for the models VOT, IP and LD, our implementation is such that models are exactly at their critical point. In CIC, we set the probability $\mu = 0$ to be in the critical regime. For CP, we set the rate $\lambda = 1$ to set the model in its critical regime. In SIS we approximated the critical regime by setting $\lambda = 1/\omega_N$, with ω_N largest eigenvalue of the adjacency matrix of the graph. For SIS, we set $\lambda = 1/\theta_{2M}$, where θ_{2M} is the largest eigenvalue of the non-backtracking matrix of the graph.

A. Synthetic networks

We performed numerical simulations of the various models described above on networks created using the directed and undirected version of the configuration models, described in Secs. II A and II B, respectively. For each network and model, we compute the distributions $P(S)$ and $P(T)$ of avalanche size S and duration T, respectively, as well as the relation $\langle S \rangle$ *vs.* T. Numerical results are obtained by simulating 10^6 avalanches; each avalanche is seeded by a single randomly chosen node. We considered different network sizes N, and compare with the expected scaling in the standard and anomalous BPs. Results are reported in Figs. 2, 5, 8, 11, 14, 17, 20 for the directed configuration model, and in Figs. 4, 7, 10, 13, 16, 19, 22 for the undirected configuration model. We summarize our results in Table II.

Avalanche model	Network model	Universality Class	Figure
CIC	Directed CM, $k_{max}^{out} = N - 1$	Anomalous	2
CIC	Directed CM, $k_{max}^{out} = \sqrt{N}$	Standard	3
CIC	Undirected CM	Standard	4
VOT	Directed CM, $k_{max}^{out} = N - 1$	Anomalous	5
VOT	Directed CM, $k_{max}^{out} = \sqrt{N}$	Standard	6
VOT	Undirected CM	Standard	7
IP	Directed CM, $k_{max}^{out} = N - 1$	Standard	8
IP	Directed CM, $k_{max}^{out} = \sqrt{N}$	Standard	9
IP	Undirected CM	Standard	10
LD	Directed CM, $k_{max}^{out} = N - 1$	Anomalous	11
LD	Directed CM, $k_{max}^{out} = \sqrt{N}$	Standard	12
LD	Undirected CM	Standard	13
SIS	Directed CM, $k_{max}^{out} = N - 1$	Anomalous	14
SIS	Directed CM, $k_{max}^{out} = \sqrt{N}$	Standard	15
SIS	Undirected CM	Standard	16
SIR	Directed CM, $k_{max}^{out} = N - 1$	Anomalous	17
SIR	Directed CM, $k_{max}^{out} = \sqrt{N}$	Standard	18
SIR	Undirected CM	Standard	19
CP	Directed CM, $k_{max}^{out} = N - 1$	Standard	20
CP	Directed CM, $k_{max}^{out} = \sqrt{N}$	Standard	21
CP	Undirected CM	Standard	22

Table II. Summary table for avalanche statistics on synthetic networks models. From left to right we report the following information: acronym of the avalanche model, name of the network model, universality class observed for the combination avalanche/network models, number of the figure where the results of the numerical simulations are visualized. Universality class is determined on the basis of visual inspection of the numerical results.
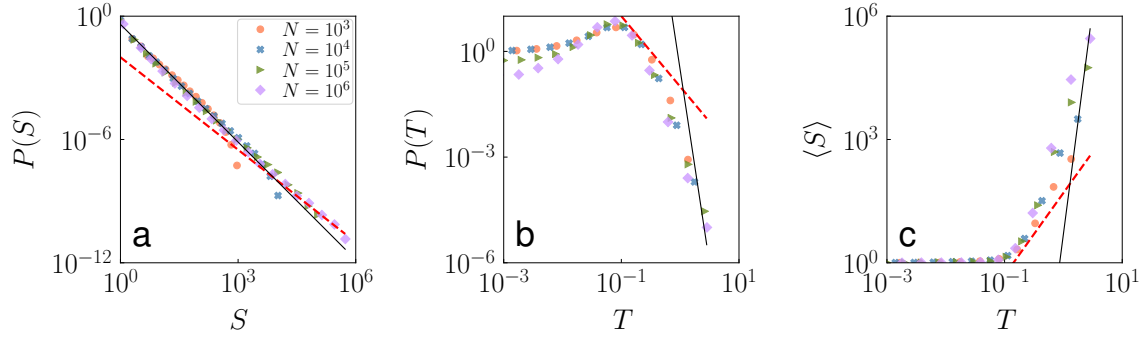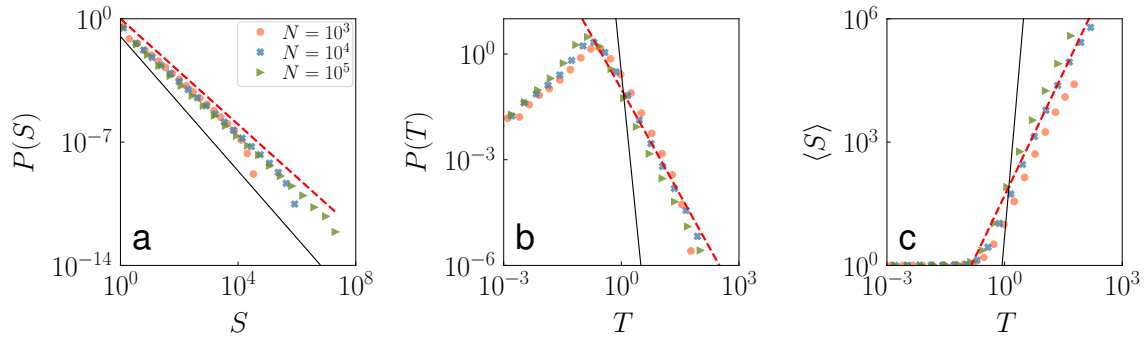
9
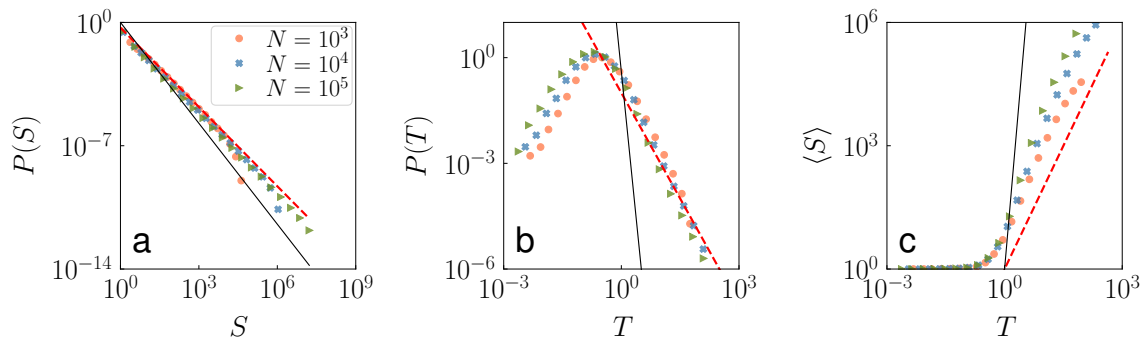
Figure 2. CIC on directed synthetic networks. Networks are constructed with the CM as described in section II A. Here $k_{max}^{out} = N - 1$. We then simulate 10^6 avalanches seeded by a single randomly chosen node. In the various panels, the red dashed line indicates standard BP exponents, whereas the black full line serves as a reference for anomalous BP exponents. (a) Probability distribution of the number of spreading events S. (b) Probability distribution of cascade duration T. (c) Average number of spreading events $\langle S \rangle$ as a function of the avalanche duration.

Figure 3. CIC on directed synthetic networks. Networks are constructed with the CM as described in section II A. Here $k_{max}^{out} = \sqrt{N}$. We then simulate 10^6 avalanches seeded by a single randomly chosen node. In the various panels, the red dashed line indicates standard BP exponents, whereas the black full line serves as a reference for anomalous BP exponents. (a) Probability distribution of the number of spreading events S. (b) Probability distribution of cascade duration T. (c) Average number of spreading events $\langle S \rangle$ as a function of the avalanche duration.

Figure 4. CIC on undirected synthetic networks. Networks are constructed with the UCM as described in section II B. We then simulate 10^6 avalanches seeded by a single randomly chosen node. In the various panels, the red dashed lines indicate standard BP exponents, whereas the black full lines serve as a reference for anomalous BP exponents. (a) Probability distribution of the number of spreading events S. (b) Probability distribution of cascade duration T. (c) Average number of spreading events $\langle S \rangle$ as a function of the avalanche duration.
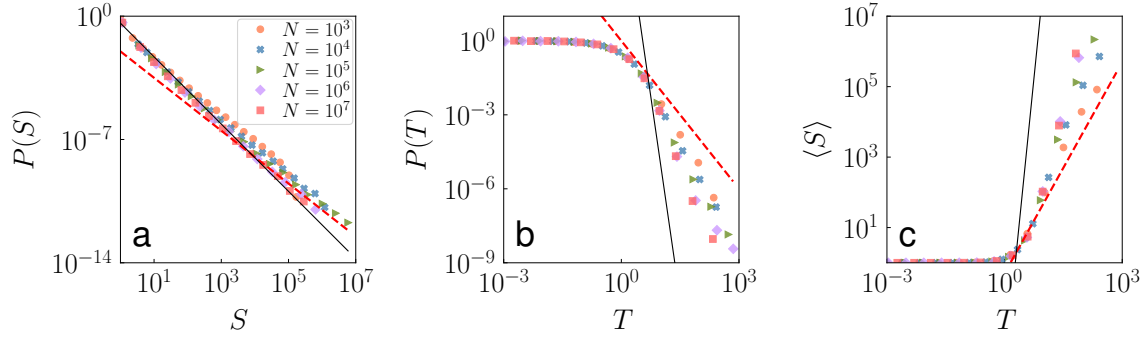
Figure 5. VOT on directed synthetic networks. The description of the various panels is the same as in Fig. 2.
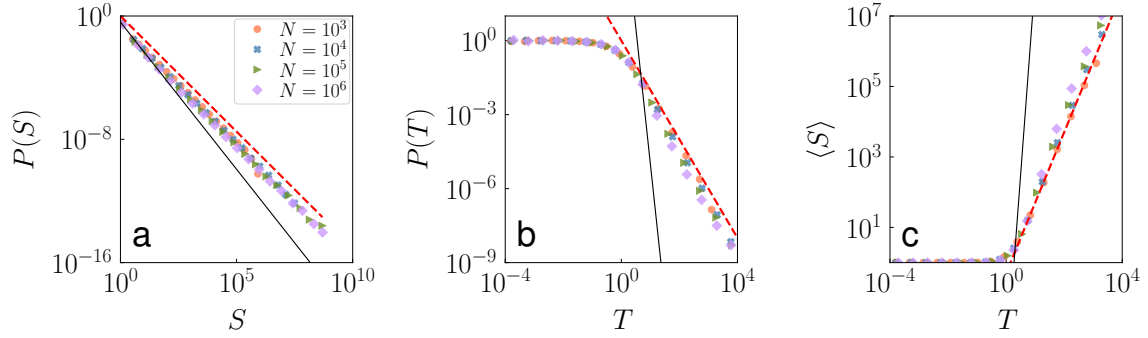

Figure 6. VOT on directed synthetic networks. The description of the various panels is the same as in Fig. 3.
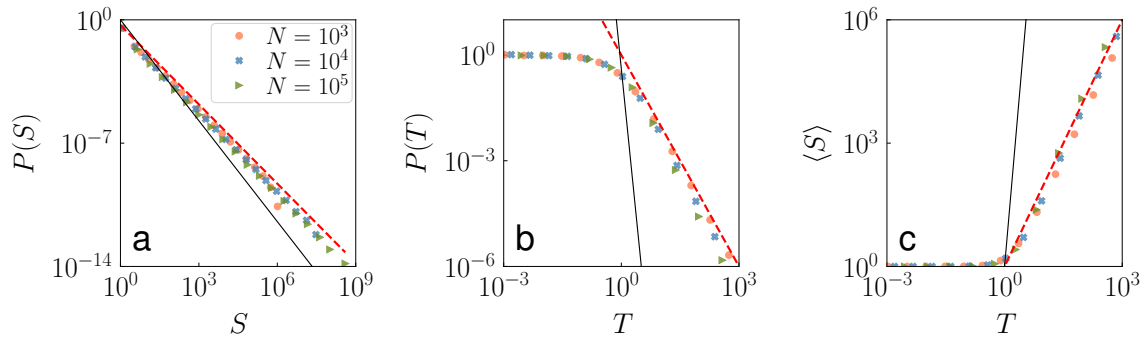

Figure 7. VOT on undirected synthetic networks. The description of the various panels is the same as in Fig. 4.
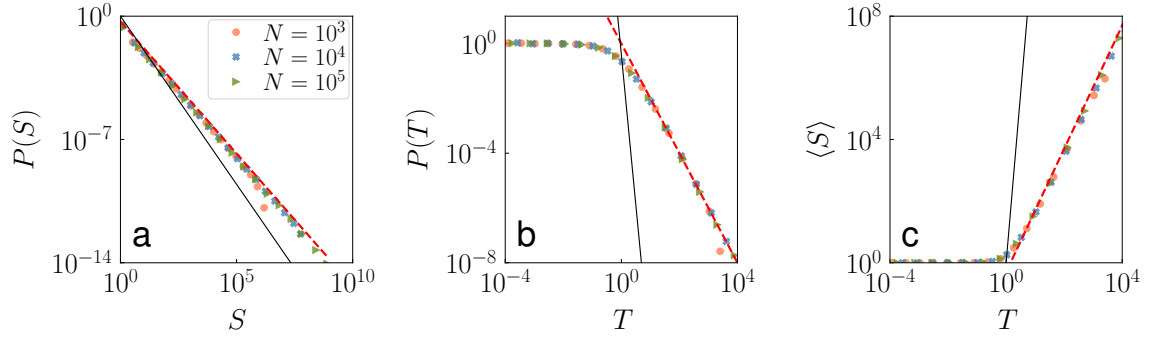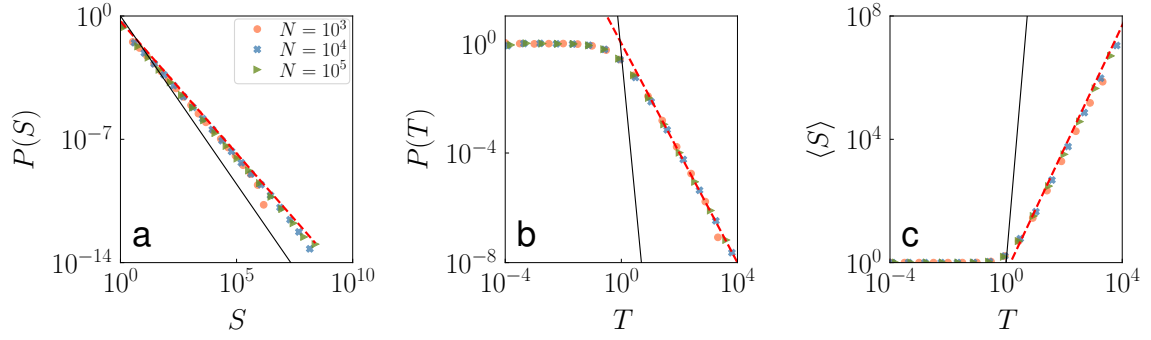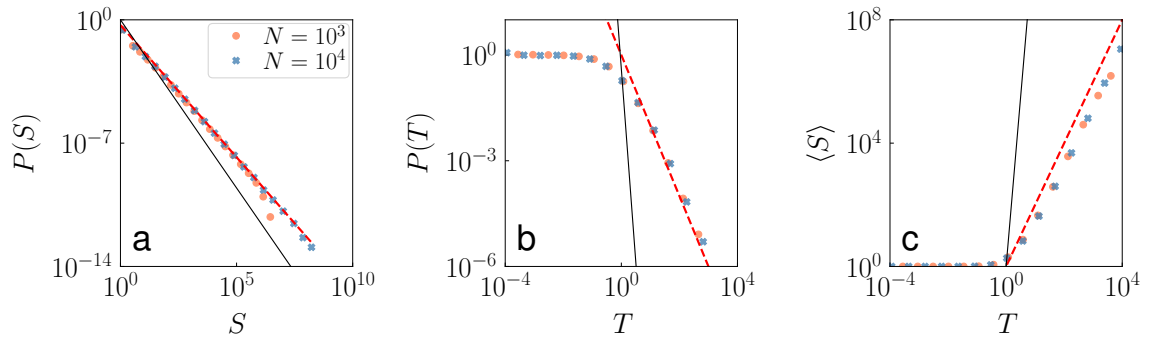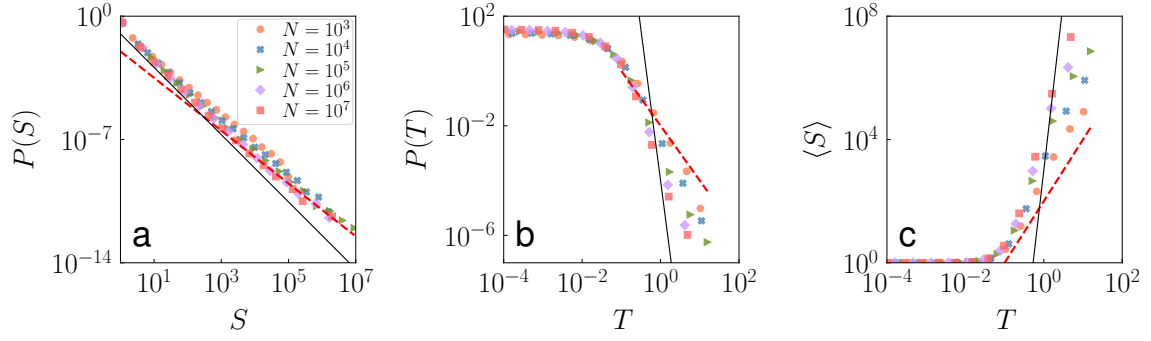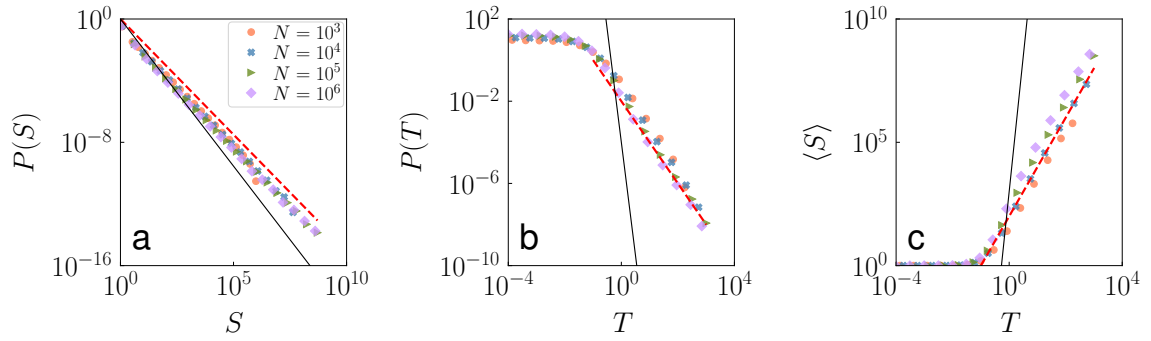
Figure 8. IP on directed synthetic networks. The description of the various panels is the same as in Fig. 2.

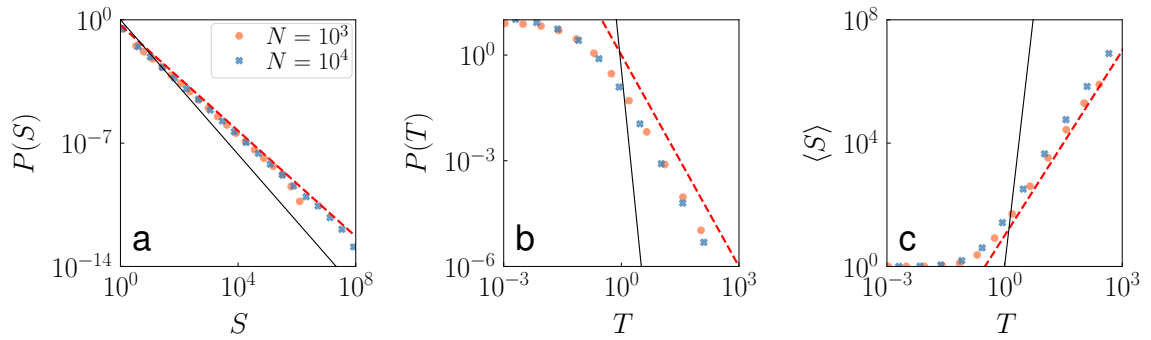Figure 9. IP on directed synthetic networks. The description of the various panels is the same as in Fig. 3.

Figure 10. IP on undirected synthetic networks. The description of the various panels is the same as in Fig. 4.
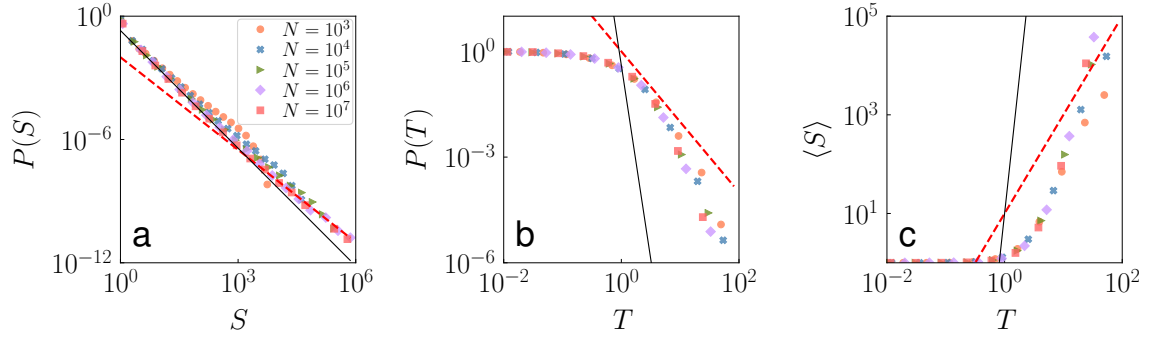
Figure 11. LD on directed synthetic networks. The description of the various panels is the same as in Fig. 2.

Figure 12. LD on directed synthetic networks. The description of the various panels is the same as in Fig. 3.

Figure 13. LD on undirected synthetic networks. The description of the various panels is the same as in Fig. 4.

Figure 14. SIS on directed synthetic networks. The description of the various panels is the same as in Fig. 2.
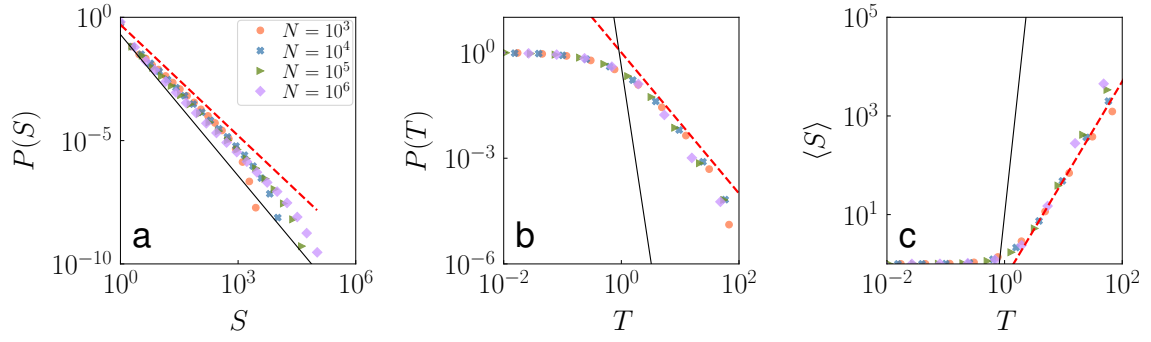

Figure 15. SIS on directed synthetic networks. The description of the various panels is the same as in Fig. 3.
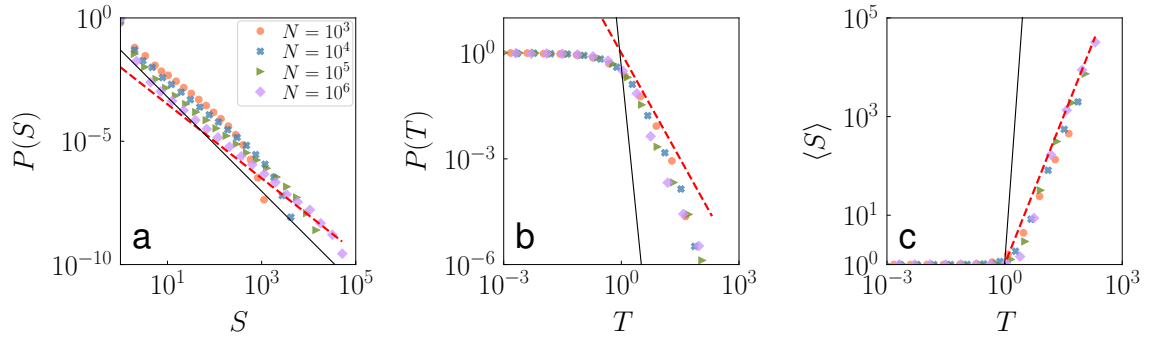

Figure 16. SIS on undirected synthetic networks. The description of the various panels is the same as in Fig. 4.
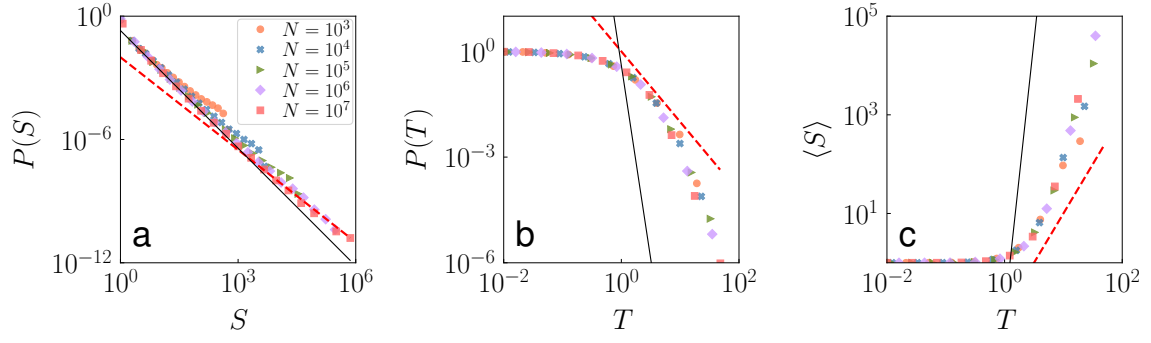
14

Figure 17. SIR on directed synthetic networks. The description of the various panels is the same as in Fig. 2.
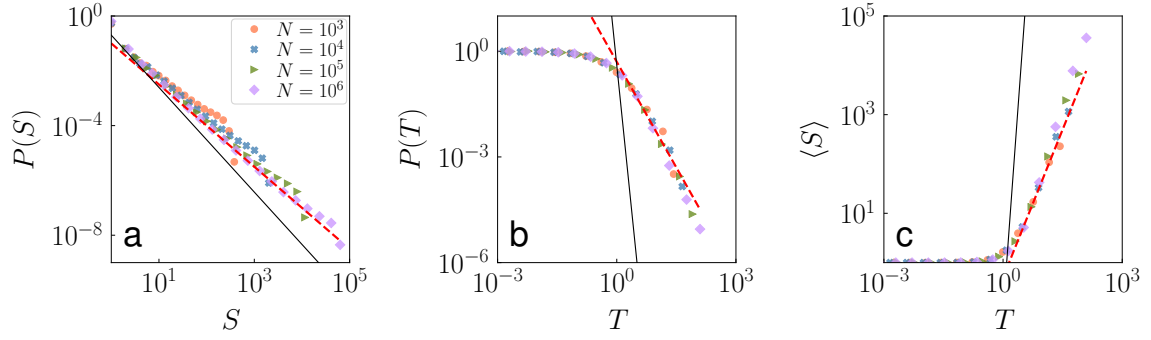

Figure 18. SIR on directed synthetic networks. The description of the various panels is the same as in Fig. 3.
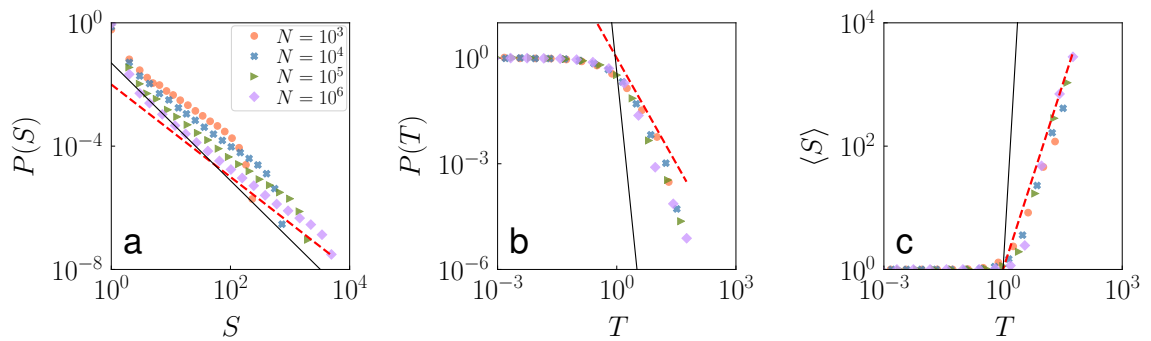

Figure 19. SIR on undirected synthetic networks. The description of the various panels is the same as in Fig. 4.
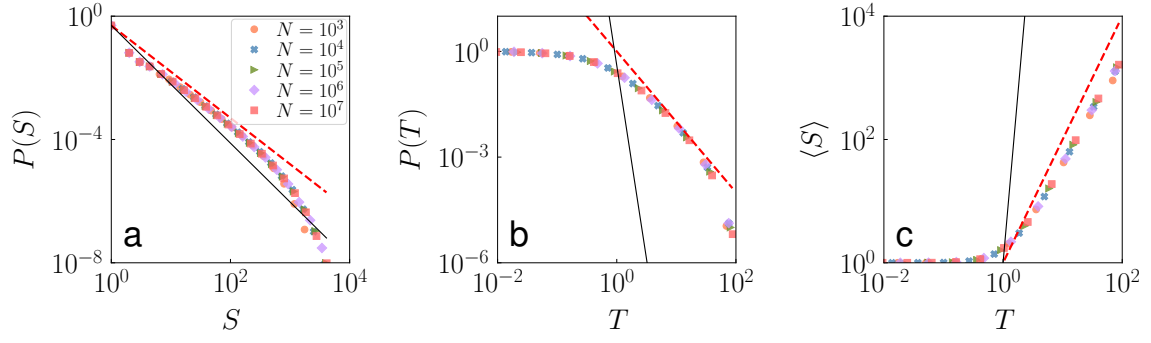
Figure 20. CP on directed synthetic networks. The description of the various panels is the same as in Fig. 2.
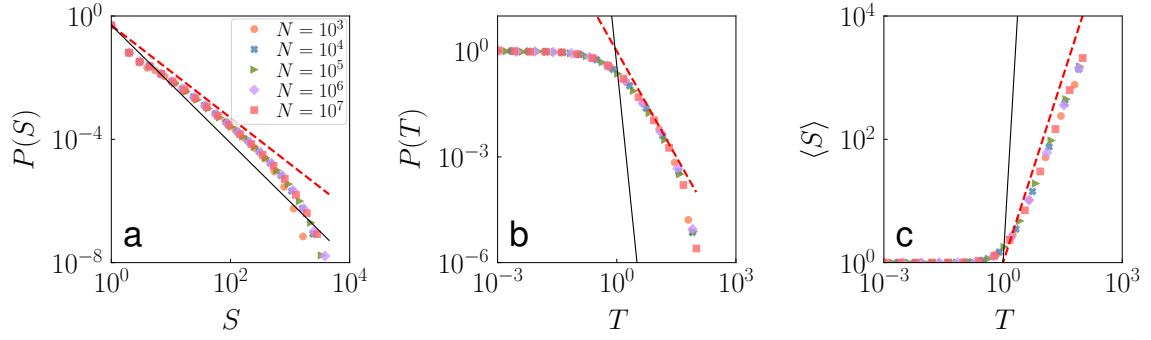

Figure 21. CP on directed synthetic networks. The description of the various panels is the same as in Fig. 3.
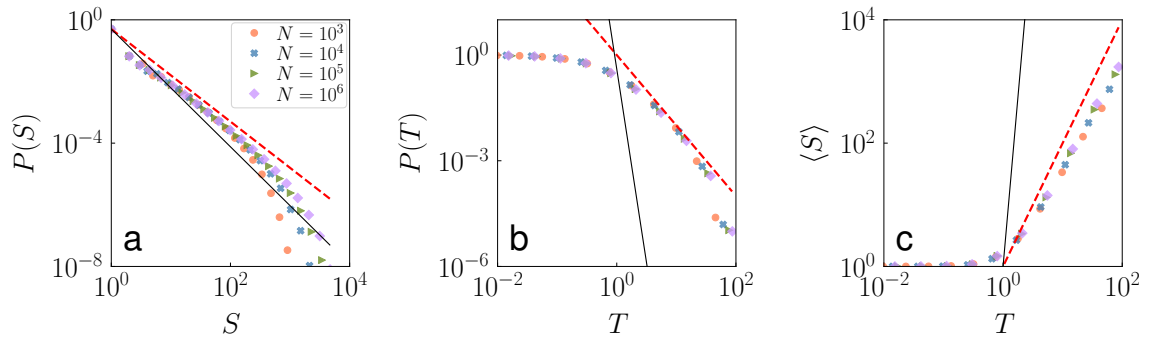

Figure 22. CP on undirected synthetic networks. The description of the various panels is the same as in Fig. 4.

B. Test of robustness

To properly assess and interpret the results obtained for avalanche dynamical models on synthetic networks, we considered two additional tests.

First, we quantified the role of finite-size effect in the determination of critical exponents for the pure branching process. Specifically, we performed simulations of a branching process where the the probability of creating x new branches is given by

$$P(x) = \begin{cases} 1 - \left(\sum_{x=x_{min}}^{x_{max}} x^{-\gamma}\right)\left(\sum_{x=x_{min}}^{x_{max}} x^{-\gamma+1}\right)^{-1} & \text{, if } x = 0 \\ x^{-\gamma}\left(\sum_{x=x_{min}}^{x_{max}} x^{-\gamma}\right)\left(\sum_{x=x_{min}}^{x_{max}} x^{-\gamma+1}\right)^{-1} & \text{, if } x_{min} \le x \le x_{max} \\ 0 & \text{, otherwise} \end{cases} \quad (6)$$

The former probability reduces to the one of Eq. (4) for $x_{min} = 1$ and $x_{max} = \infty$. In our tests, we set $x_{min} = 4$, and $x_{max} = N$, with N assuming the same values as those of the networks considered in our numerical simulations. Results are displayed in Fig. 23 for $\gamma = 2.1$ and in Fig. 24 for $\gamma = 2.5$. We note that the scalings for $P(T)$ and $\langle S \rangle$ vs. T predicted by theory are not reproduced for $\gamma = 2.1$. This fact is in line with what observed in results reported above for avalanches in finite-size networks. In the case of BP with $x_{max} = \infty$, the issue may be due to other types of finite-size effects, including the lack of proper divergence for the moments of random variates generated in computer algorithms [23]. Scaling exponents for $P(T)$ and $\langle S \rangle$ vs. T closer to those theoretically predicted are obtained for $\gamma = 2.5$. In this case, however, the distinction between anomalous vs. standard exponents for $P(S)$ becomes almost unnoticeable.
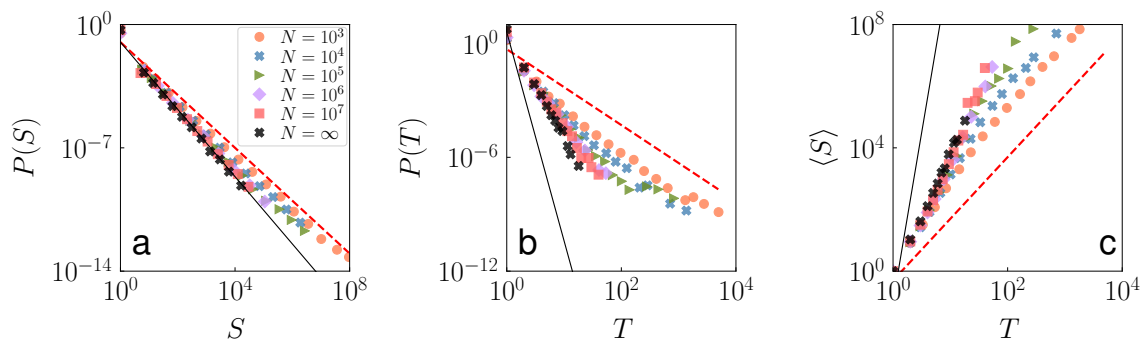
Figure 23. Simulations of the branching process with branching probability defined as in Eq. (6). We set $x_{min} = 4$, and consider different values $x_{max} = N$. Here, the power-law exponent is $\gamma = 2.1$.

Figure 24. Same as in Fig. 23, but for $\gamma = 2.5$.

Second, we considered some of the dynamical avalanche models on synthetic directed networks generated with $k_{min} = 4$, $k_{max} = N - 1$, and $\gamma = 2.5$. Results are provided in Fig. 25 for CIC, and in 26

17

for VOT. For this value of γ it is impossible to distinguish anomalous from normal scaling of the size distribution, but instead one can see evidence of anomalous scaling in the dependence of $P(T)$ and $\langle S \rangle$ on T.

Figure 25. CIC on directed synthetic networks. Same as in Fig. 2, but for networks with degree exponent $\gamma = 2.5$.
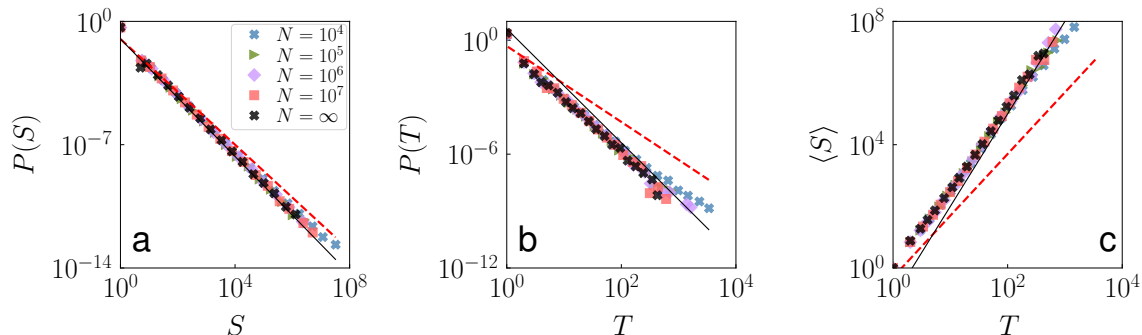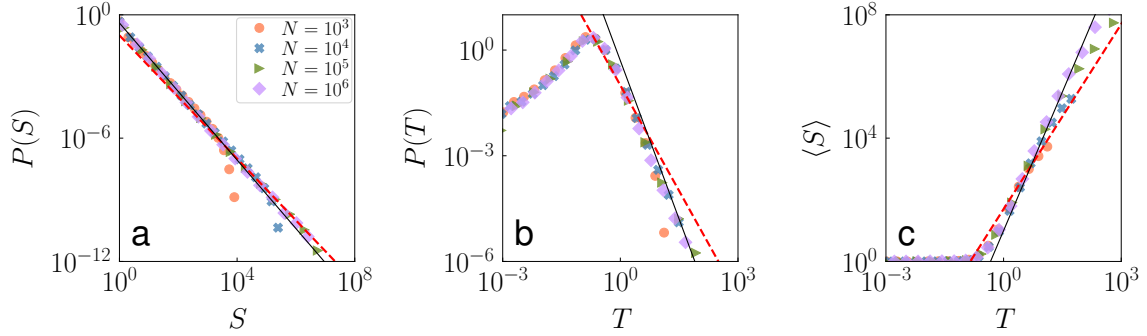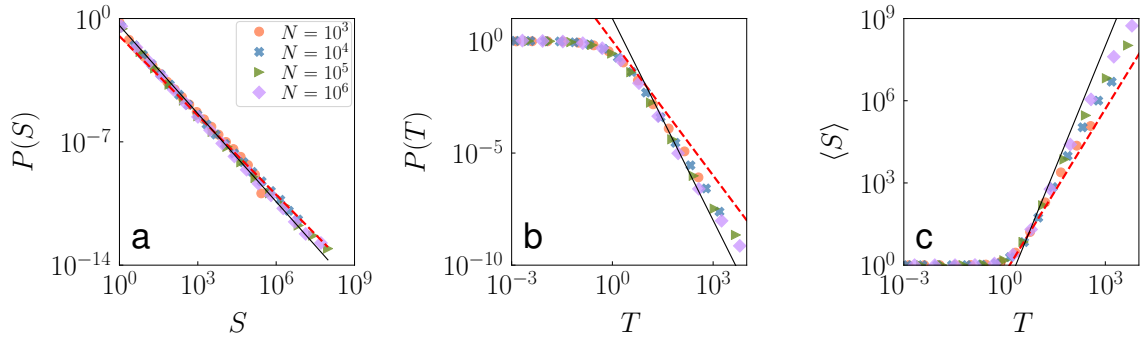

Figure 26. VOT on directed synthetic networks. Same as in Fig. 5, but for networks with degree exponent $\gamma = 2.5$.

C. Real-world networks

We performed numerical simulations of avalanche dynamics running on top of the real-world topologies of Table I for a selection of dynamical models. Results are reported in the main manuscript and Figures 27 and 28. Numerical results are obtained by simulating 10^6 avalanches, each seeded by a single randomly chosen node.

V. LANGEVIN EQUATIONS FOR AVALANCHE DYNAMICS IN COMPLEX NETWORKS

We derive here Langevin equations describing the dynamics of all models considered above, except for SIR, which is excluded being the only model where nodes are not described by binary variables. In spite of this fact, numerical results for the SIR model are consistent with those obtained for the other models. We assume that dynamics is taking place on a network with adjacency matrix A and N nodes. The section is organized as follows. First, we describe the general procedure to derive Langevin equations for a model in which nodes can assume only two states. Then, we explicitly illustrate the procedure for the CIC and SIS models. Properties of the other models can be derived by mapping them to the case of either CIC or SIS.

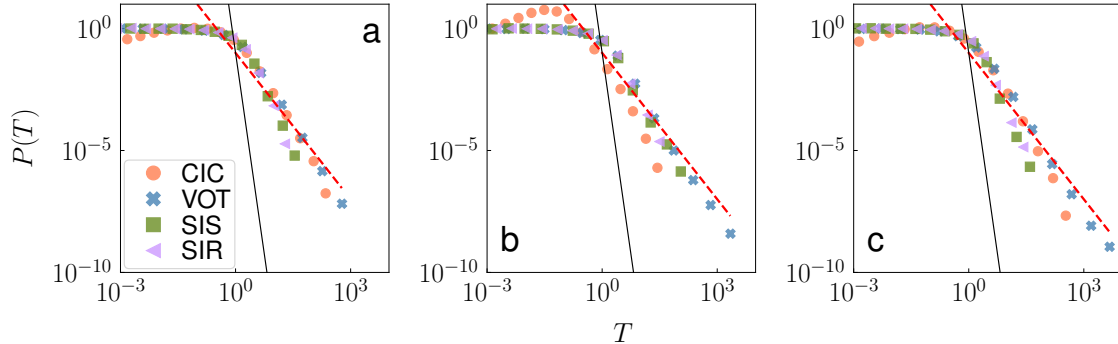Figure 27. Avalanches in real-world networks. We consider the following networks: (a) undirected graph representing a snapshot of the Internet at the Autonomous system level [13]; (b) directed Twitter network of the Spanish 15M movement [3]; (c) directed graph representing a portion of the Youtube social network [15]. We measure the distribution $P(T)$ of avalanche duration T for some dynamical models. Different symbols and colors refer to different avalanche dynamical models. The red dashed line represent standard BP critical exponents, while the full black line indicates the power-law decay expected for the anomalous BP. Note that the out-degree distributions of these networks are all well modeled by power laws with exponent $\gamma = 2.1$.
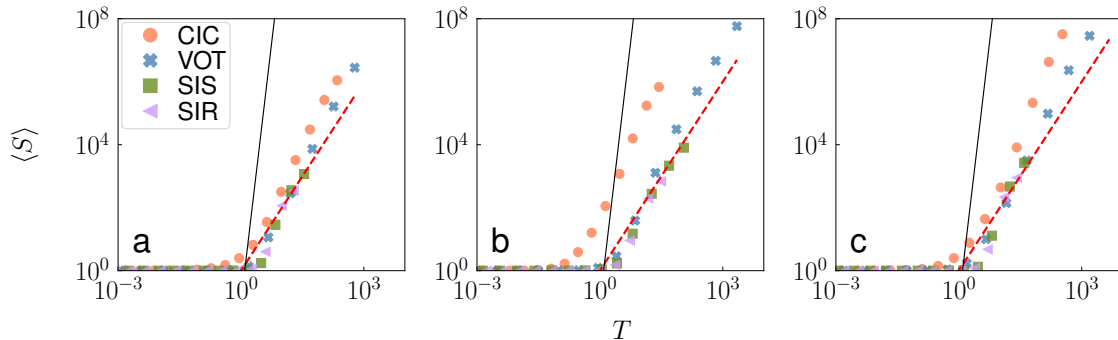
Figure 28. Same as in Figure 27, but for the relation between the average value of the size of the avalanche $\langle S \rangle$ and its duration. The red dashed line represent standard BP critical exponents, while the full black line indicates the power-law decay expected for the anomalous BP. Note that the out-degree distributions of these networks are all well modeled by power laws with exponent $\gamma = 2.1$.

A. Summary of the approach

We follow a procedure similar to one described in Ref. [2]. The state of the node i at time t is a random time-dependent variable $\sigma_i(t)$. The variable can assume two possible values, namely $\sigma_i(t) = 0, 1$. Its evolution can be described by the equation

$$\sigma_i(t + dt) = \sigma_i(t)\,\zeta_i(dt) + (1 - \sigma_i(t))\,\eta_i(dt)\,. \tag{7}$$

The variables $\zeta(dt)$ and $\eta(dt)$ are dichotomous random variables taking values

$$\zeta_i(dt) = \begin{cases} 0 & \text{, with probability } P_{\zeta_i}(0) \\ 1 & \text{, with probability } P_{\zeta_i}(1) = 1 - P_{\zeta_i}(0) \end{cases} \tag{8}$$

and

$$\eta_i(dt) = \begin{cases} 1 & \text{, with probability } P_{\eta_i}(1) \\ 0 & \text{, with probability } P_{\eta_i}(0) = 1 - P_{\eta_i}(1) \end{cases}. \tag{9}$$

19

In the notation above, $P_{\zeta_i}(0)$ is the probability to observe the transition $\sigma_i(t) = 1 \to \sigma_i(t + dt) = 0$; $P_{\eta_i}(1)$ is the probability to observe the transition $\sigma_i(t) = 0 \to \sigma_i(t + dt) = 1$. The values of the probabilities P_{ζ_i} and P_{η_i} depend on the details of the stochastic dynamical model under consideration.

Here, we will focus on the global (network-wide) average of the variables $\sigma_i(t)$, i.e.,

$$\rho(t) = \frac{1}{N} \sum_i \sigma_i(t) . \tag{10}$$

Neglecting correlations and assuming finite variance of the σ_i, we can invoke the central limit theorem, and assume that $\rho(t)$ obeys a normal distribution. Its dynamics can be therefore approximated by the Langevin equation

$$\frac{d\rho(t)}{dt} = \Psi[\rho(t)] + \sqrt{D[\rho(t)]}\, \xi(t) . \tag{11}$$

In the equation above, $\xi(t)$ is a stochastic variable extracted from a standard normal distribution. The drift term is defined as

$$\Psi[\rho(t)] = \lim_{dt \to 0} \frac{\langle \Delta\rho(t)|\vec{\sigma}(t)\rangle}{dt} = \frac{1}{N} \sum_i \Psi_i[\vec{\sigma}(t)] ,$$

whereas the elements that compose the diffusion term are

$$D[\rho(t)] = \lim_{dt \to 0} \frac{\langle [\Delta\rho(t)]^2|\vec{\sigma}(t)\rangle}{dt} = \frac{1}{N^2} \sum_i D_i[\vec{\sigma}(t)] ,$$

with $\vec{\sigma}(t) = [\sigma_1(t), \sigma_2(t), \ldots, \sigma_N(t)]^T$ a vector that encodes the state of all node variables at time t and $\Delta\rho(t) := \rho(t + dt) - \rho(t)$. The functions $\Psi_i[\vec{\sigma}(t)]$ and $D_i[\vec{\sigma}(t)]$ are defined explicitly below.

Because of the boolean nature of the variables ζ_i and η_i, their average values coincide with the probability of these variables taking value 1, i.e. $\langle \zeta_i \rangle = P_{\zeta_i}(1)$ and $\langle \eta_i \rangle = P_{\eta_i}(1)$. For the models we have considered here, we can always write these probabilities in terms of a gain and a loss function, χ_i and ℓ_i, respectively, as

$$\begin{aligned} P_{\zeta_i}(1) &= 1 - dt\, \ell_i \\ P_{\eta_i}(1) &= dt\, \chi_i \end{aligned} . \tag{12}$$

Each model will be entirely specified by the choice of χ_i and ℓ_i. However, the general expressions (12) allow to explicitly compute the general form of the Langevin equation (11).

The drift term is given by

$$\langle \Delta\rho(t)|\vec{\sigma}(t)\rangle = \frac{1}{N} \sum_i [\sigma_i\langle\zeta_i\rangle + (1 - \sigma_i)\langle\eta_i\rangle - \sigma_i] = \frac{dt}{N} \sum_i [(1 - \sigma_i)\chi_i - \sigma_i\ell_i] ,$$

so that

$$\Psi_i[\vec{\sigma}(t)] = (1 - \sigma_i)\chi_i - \sigma_i\ell_i .$$

To compute the diffusion term, note that

$$\langle [\Delta\rho(t)]^2|\vec{\sigma}(t)\rangle = \langle [\rho(t + dt)]^2|\vec{\sigma}(t)\rangle - \langle \rho(t + dt)|\vec{\sigma}\rangle^2 ,$$

thanks to the properties of conditional variance. Both these terms can be written as a diagonal sum over a single node and an off diagonal sum over all distinct couples (i, j). By using $\langle \zeta_i\zeta_j \rangle = \langle\zeta_i\rangle\langle\zeta_j\rangle$ and similar relations for η_i, the off diagonal terms equal. By also noting that $\sigma_i^2 = \sigma_i$, $(1 - \sigma_i)^2 = (1 - \sigma_i)$ and $\sigma_i(1 - \sigma_i) = 0$, the difference among the diagonal terms can be written as

$$\langle [\Delta\rho(t)]^2|\vec{\sigma}(t)\rangle = \frac{1}{N^2} \sum_i \sigma_i \left(\langle\zeta_i^2\rangle - \langle\zeta_i\rangle^2\right) + (1 - \sigma_i)\left(\langle\eta_i^2\rangle - \langle\eta_i\rangle^2\right) .$$

Finally, note that the second moment of ζ_i and η_i equals the first and that, thanks to the expressions (12), $\langle \eta_i \rangle^2 = \mathcal{O}(dt^2)$ and $\langle \zeta_i \rangle^2 = 1 - 2\,dt\,\ell_i + \mathcal{O}(dt^2)$. It is then immediate to see that

$$D[\rho] = \frac{1}{N^2} \sum_i [\sigma_i \ell_i + (1 - \sigma_i)\chi_i] .$$

In summary then, the dynamics of the density is described by the following set of equations:

$$
\begin{cases}
d\rho(t)/dt = \Psi[\rho(t)] + \sqrt{D[\rho(t)]}\xi(t) \\[2mm]
\Psi[\rho(t)] = \frac{1}{N} \sum_i \Psi_i[\vec{\sigma}(t)] = \frac{1}{N} \sum_i [(1 - \sigma_i)\chi_i - \sigma_i \ell_i] \\[2mm]
D[\rho(t)] = \frac{1}{N^2} \sum_i D_i[\vec{\sigma}(t)] = \frac{1}{N^2} \sum_i [(1 - \sigma_i)\chi_i + \sigma_i \ell_i]
\end{cases}
\qquad . \qquad (13)
$$

Please note that we didn't write explicitly the time dependence of the various variables to keep the notation as simple as possible.

Whereas Eqs. (13) can be used to describe the approximate behaviour of the system at arbitrary time t, we will mostly focus our attention on the long-term behavior of the dynamics. To understand the long-term dynamics of the system, we will look at the ensemble average value (over an infinite number of realizations of the dynamics) of the variables $s_i(t) := \langle \sigma_i(t) \rangle$ only. We refer to this approximation as the individual-based mean-field approximation (IBMFA) [21]: taking the expectation value of Eq. (7) we obtain an equation involving $\langle \sigma_i(t)\chi_i(t) \rangle$ and analogously for the other term on the rhs of the equation. The IBMFA consists into assuming that $\langle \sigma_i(t)\chi_i(t) \rangle = \langle \sigma_i(t) \rangle \langle \chi_i(t) \rangle$. The IBMFA dynamical equations read as

$$\frac{ds_i(t)}{dt} = [1 - s_i(t)]g_i(t) - s_i(t)l_i(t) . \qquad (14)$$

In the equation above, $g_i(t) := \langle \chi_i(t) \rangle$ and $l_i(t) := \langle \ell_i(t) \rangle$ are the average values (over an infinite number of simulations of the model) of the gain and loss terms defined earlier. Please note that s_i, g_i and l_i are deterministic variables, and they differentiate from their stochastic counterparts σ_i, χ_i and ℓ_i, respectively.

In the following, we will use the IBMFA for two main purposes: (i) establishing the amount of time necessary before the system enters in its long-term dynamical regime; (ii) determining the properties of the system in the long-term dynamical. We will then use this knowledge into the Langevin formalism to derive statistical properties that describe only avalanches that are long (large) enough to reach the long-term dynamical regime of the system.

B.   CIC model

We consider first the case of the CIC model. According to the rules of CIC dynamics, the probabilities P_{ζ_i} and P_{η_i} are

$$P_{\zeta_i}(0) = dt \left[\mu + \mu \sum_j A_{ji} + (1 - \mu) \sum_j A_{ji}(1 - \sigma_j) \right] = dt \left[\mu(k_i^{in} + 1) + (1 - \mu) \sum_j A_{ji}(1 - \sigma_j) \right] , \quad (15)$$

and

$$P_{\eta_i}(1) = dt\,(1 - \mu) \sum_j A_{ji}\sigma_j , \qquad (16)$$

Please note that we didn't write explicitly the time dependence of the variables σs to keep the notation as simple as possible. In the following, we will continue to use this light notation; explicit time dependence of variables will appear only in some cases to give emphasis to such a dependence. The first term in Eq. (15) accounts for the probability that node i creates a new meme; the second term accounts for the

probability that a neighbor of node i creates and shares a new meme; the last term is the probability of a neighbor of node i to share the meme already in its memory, being this meme different from the one under observation. The r.h.s. of Eq. (16) accounts for the fact that a neighbor of node i shares the meme under observation. The above equations take the form (12) for

$$\chi_i = (1 - \mu) \sum_j A_{ji} \sigma_j \tag{17}$$

and

$$\ell_i = \mu(k_i^{in} + 1) + (1 - \mu) \sum_j A_{ji}(1 - \sigma_j) . \tag{18}$$

1. Individual-based mean-field approximation

Insights regarding the physical behavior of the system can be obtained by working under the IBMFA, and write Eq. (14) as

$$\frac{ds_i}{dt} = -\mu(1 + k_i^{in}) \, s_i + (1 - \mu) \sum_j A_{ji} \left[s_j(1 - s_i) - s_i(1 - s_j) \right] . \tag{19}$$

This equation arises from the fact that the average values of Eqs. (17) and (18) are

$$g_i := \langle \chi_i \rangle = (1 - \mu) \sum_j A_{ji} s_j .$$

and

$$l_i := \langle \ell_i \rangle = \mu(k_i^{in} + 1) + (1 - \mu) \sum_j A_{ji}(1 - s_j) ,$$

respectively. We note that the summation in the r.h.s. of Eq. (19) can be manipulated as

$$\sum_j A_{ji}[s_j(1 - s_i) - s_i(1 - s_j)] = \sum_j A_{ji}(s_j - s_i) = \sum_j (A_{ji} - k_i^{in} \delta_{ji}) s_j = - \sum_j L_{ji} \, s_j .$$

In the equation above, we made use of the Kronecker delta function, $\delta_{ji} = 1$ if $j = i$, and $\delta_{ji} = 0$, otherwise. Further, we indicated with $L = K^{in} - A$ the Laplacian operator of the graph, with K^{in} diagonal matrix containing the in-degree of the nodes [5, 30]. We note that Eq. (19) can be rewritten as

$$\frac{ds_i}{dt} = -\mu(1 + k_i^{in}) s_i - (1 - \mu) \sum_j L_{ji} \, s_j , \tag{20}$$

or in matrix-vector format as

$$\frac{d\vec{s}}{dt} = -\mu(I + K^{in}) \, \vec{s} - (1 - \mu) \, L^T \, \vec{s} . \tag{21}$$

Here I is the identity matrix and L^T is the transpose of the graph Laplacian. In essence, under the mean-field approximation, the dynamics of a meme in the CIC model is described by a simple diffusion equation with a dissipative term. Let us now consider the system at the critical point $\mu = 0$.
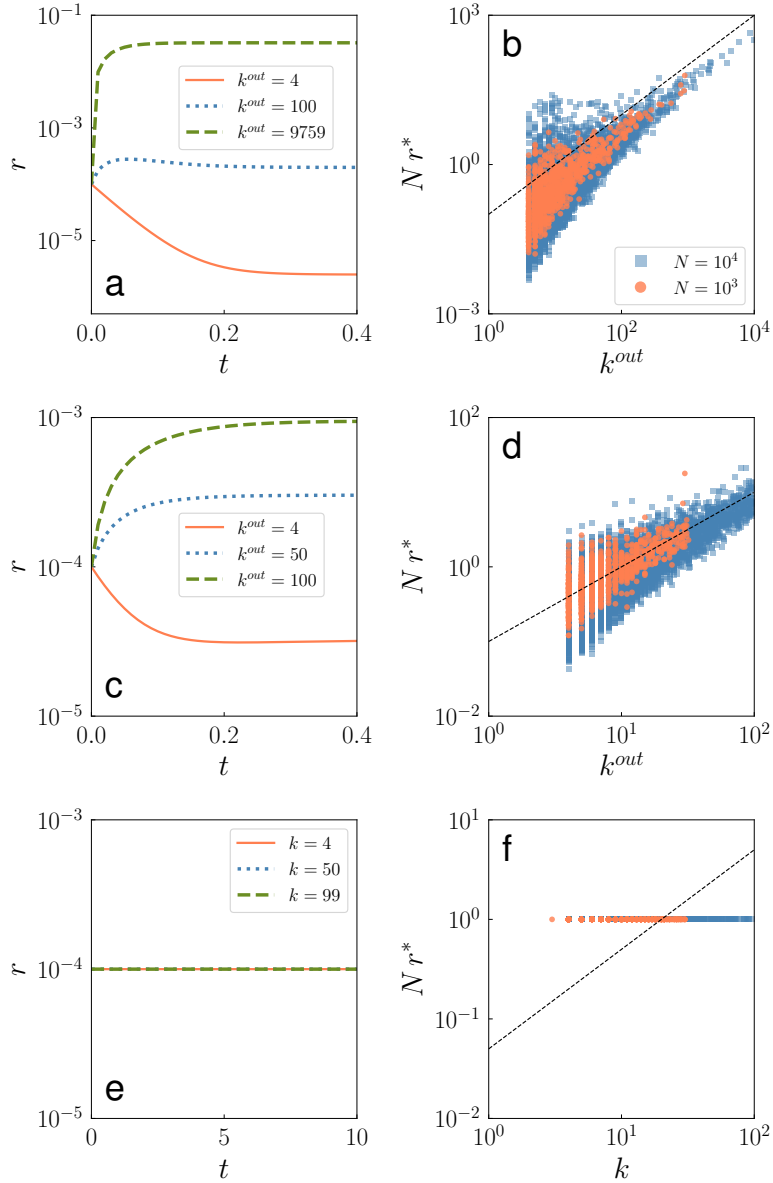
Figure 29. We consider the CIC model for $\mu = 0$ on directed networks constructed according to the model described in section II A. We consider initial conditions where $s_i(t = 0) = 0$ for all i, except for a single seed node j such that $s_j(t = 0) = 1$. (a) We construct a network with $N = 10^4$ and $k_{max}^{out} = N - 1$. Then, we numerically integrate Eq. (21) and plot $r(t) = 1/N \sum_i s_i(t)$ as a function of t. The three different lines correspond to different initial conditions where the selected seed nodes have radically different out-degree values k^{out}. (b) Dependence of the asymptotic value r^* on the out-degree of the initial seed k^{out}. We consider two network sizes, $N = 10^3$ and $N = 10^4$. The dashed line is proportional to k^{out}. (c) and (d) Same as in panels a and b, respectively, but for networks generated according to the directed configuration model with $k_{max}^{out} = \sqrt{N}$. (e) and (f) Same as in panels a and b, respectively, but for networks generated according to the undirected configuration model with $k_{max} = \sqrt{N}$. In all cases the exponent of the out-degree distribution is $\gamma = 2.1$.

a. Undirected networks

If the network is undirected the matrix L is symmetric and the solution of Eq. (21) can be written using the eigendecomposition of L as

$$\vec{s}(t) = \sum_n [\vec{v}_n \cdot \vec{s}(0)] \, e^{-\nu_n t} \, \vec{v}_n . \tag{22}$$

23

In the above equation, $\vec{v}_n$ is the n-th eigenvector of the matrix L corresponding to the n-th eigenvalue ν_n and $\vec{v}_n \cdot \vec{s}(0)$ is the scalar product between the vectors $\vec{v}_n$ and the initial condition $\vec{s}(0)$. We remind that the eigenvalues of L are such that $0 = \nu_1 \leq \nu_2 \leq \cdots \leq \nu_N$, where the multiplicity of the null eigenvalue represents the number of connected components in the network [30]. For simplicity and without lack of generality, we assume that the entire network is composed of a single connected component. From Eq. (22) $\vec{s}$ converges exponentially fast, after a timescale $1/\nu_2$, to the stationary value

$$\vec{s}^* = [(\vec{v}_1)^T \cdot \vec{s}(0)] \, \vec{v}_1 \, , \tag{23}$$

Since the eigenvector $\vec{v}_1$ has all components equal $\vec{v}_1 = \frac{1}{\sqrt{N}} (1, \ldots, 1)^T$, this means that all s_i converge to the same stationary value

$$s_i = r^* = \|[(\vec{v}_1)^T \cdot \vec{s}(0)] \cdot \vec{v}_1\| = N^{-1}. \tag{24}$$

The form of the $\vec{v}_1$ implies that r^* does not depend on which node is the initial seed.

Note also that the sum over i of Eq. (20) at the critical point $\mu = 0$ equals 0 in the case of undirected network, pointing out that the density is a conserved quantity in CIC model under the IBMFA. All these results are numerically confirmed in Figure 29(e-f).

   b. *Directed networks*

If instead the network is directed, the eigenfunction expansion generally doesn't hold. Assuming that the network is composed of a single strongly connected component, then the asymptotic value of each s_i has been computed in Ref. [30] as

$$s_i = [(\vec{v}_1^{(l)})^T \cdot \vec{s}(t=0)] \, (\vec{v}_1^{(r)})_i \, , \tag{25}$$

where now $\vec{v}_1^{(l)}$ and $\vec{v}_1^{(r)}$ are the first left and right eigenvectors of L^T, respectively. In particular, $\vec{v}_1^{(r)}$ is again proportional to $\vec{1}$, so that all variables s tend to a common value given by $r^* = \|[(\vec{v}_1^{(l)})^T \cdot \vec{s}(0)] \cdot \vec{v}_1^{(r)}\|$. In this case this asymptotic value depends on the initial seed. Figure 29 shows that for the random directed networks considered in our work the value of r^* is proportional to the initial seed's out-degree k^{out}.

2. *Langevin approach*

For $\mu = 0$, we know, from the previous description of the individual-based mean-field theory, that if t is large enough, then $s_i = r^*$. Recall that each variable s_i is defined as the average of σ_i over an infinite number of realizations of the process. Consequently, s_i is a deterministic realization-independent variable. To preserve the stochastic nature of the variables σ_i and, at the same time, exploit the lesson we learn from the IBMFA, we make a further approximation: instead of considering the full stochastic nature of the σ_i as prescribed by Eq. (7), we approximate these variables, in the long-term limit, as Bernoulli random variables with a success probability $P(\sigma_i = 1) = r^*$. This probability does not depend on i, in analogy with the IBMFA description, nor it depends on time because of the nature of this approximation (long-term limit).

Finally, we perform what is usually called adiabatic approximation [2]: as the IBMFA suggests, we assume that the microscopic degrees of freedom have reached a stationary state in which they weakly fluctuate around their average values, allowing to replace them with their averages and preserving the stochastic nature of the macroscopic variable only, which is the density ρ in our case. In the present framework, the average does not need to be taken over multiple realizations: it can be thought as a temporal average over each realization, under the constraint that the temporal average is taken entirely in the stationary limit of large times. Under the previous approximation of Bernoulli variables, we set $\sigma_i = \rho^*$. This choice of the notation aims to remind to the reader that the variable ρ^* is inspired by the deterministic variable r^*, but is defined in the context of the stochastic Langevin description.

Replacing σ_i with ρ^* the drift term in Eq. (11) is automatically zero. The i-th component of the diffusion term appearing in Eq. (13) reads as

$$D_i[\vec{\sigma}] = \sum_j A_{ji} \left(\sigma_i(1 - \sigma_j) + \sigma_j(1 - \sigma_i) \right) = \sum_j A_{ji} \left(\rho^*(1 - \rho^*) + \rho^*(1 - \rho^*) \right) = 2 \, k_i^{in} \, \rho^*(1 - \rho^*) \, .$$

The diffusion term is then

$$\frac{1}{N^2} \sum_i D_i[\vec{\sigma}] = \frac{2\,\rho^*(1-\rho^*)\sum_i k_i^{in}}{N^2} = 2\,\rho^*(1-\rho^*)\,\frac{\langle k \rangle}{N} \;,$$

where $\langle k \rangle$ is the average degree of the network. Under the adiabatic approximation of the (approximated) Bernoulli variables, the Langevin equation takes the form

$$\frac{d\rho}{dt} = \xi \sqrt{\rho\,(1-\rho)} \sqrt{\frac{2\langle k \rangle}{N}} \;. \tag{26}$$

Eq. (26) is identical to the equations considered in Ref [6], thus allowing us to state that duration and size of avalanches obey power-law distributions characterized by exponents typical of the mean-field branching process.

C.   Link Dynamics

Link Dynamics can be treated similarly. According to the rules of the dynamics, the probabilities P_{ζ_i} and P_{η_i} are

$$P_{\zeta_i}(0) = \frac{dt}{\langle k \rangle} \sum_j A_{ji}(1-\sigma_j)$$

and

$$P_{\eta_i}(1) = \frac{dt}{\langle k \rangle} \sum_j A_{ji}\sigma_j \;.$$

The factor $dt/\langle k \rangle$ is the rate at which each link is selected [28]. The above equations take the form (12) for

$$\ell_i = \frac{1}{\langle k \rangle} \sum_j A_{ji}(1-\sigma_j) \;,$$

and

$$\chi_i = \frac{1}{\langle k \rangle} \sum_j A_{ji}\sigma_j \;.$$

Equations (17) and (18) are identical to the above expressions for $\mu = 0$ apart from a factor $\langle k \rangle^{-1}$. Rescaling time by this factor, results for the CIC model apply.

D.   Voter Model

Also the Voter model can be treated similarly. According to the rules of the dynamics, the probabilities P_{ζ_i} and P_{η_i} are

$$P_{\zeta_i}(0) = \frac{dt}{k_i^{in}} \sum_j A_{ji}(1-\sigma_j)$$

and

$$P_{\eta_i}(1) = \frac{dt}{k_i^{in}} \sum_j A_{ji}\sigma_j \ .$$

In this case the term $1/k_i^{in}$ is the probability of picking one among the k_i^{in} links pointing to node i. These equations correspond to the choice

$$\ell_i = \frac{1}{k_i^{in}} \sum_j A_{ji}(1 - \sigma_j) \ ,$$

and

$$\chi_i = \frac{1}{k_i^{in}} \sum_j A_{ji}\sigma_j \ .$$

If we apply the IBMFA, we obtain

$$\frac{d\vec{s}(t)}{dt} = -\mathcal{L}^T \vec{s}(t) \ , \tag{27}$$

where $\mathcal{L}$ is the normalized graph laplacian, i.e.,

$$\mathcal{L} = K_{in}^{-1} L \ .$$

The kernel of this matrix is spanned by the same vectors that span the kernel of the graph laplacian L [5]. Thus, each s_i will converge exponentially fast to a constant value. All considerations valid for the CIC model can be extended to the Voter model as well.

E. SIS model

We consider here the SIS model with recovery rate equal to one, and epidemic rate equal to λ. We can write

$$P_{\zeta_i}(0) = dt$$

and

$$P_{\eta_i}(1) = dt\,\lambda \sum_j A_{ji}\sigma_j \ ,$$

implying

$$\chi_i = \lambda \sum_j A_{ji}\sigma_j \tag{28}$$
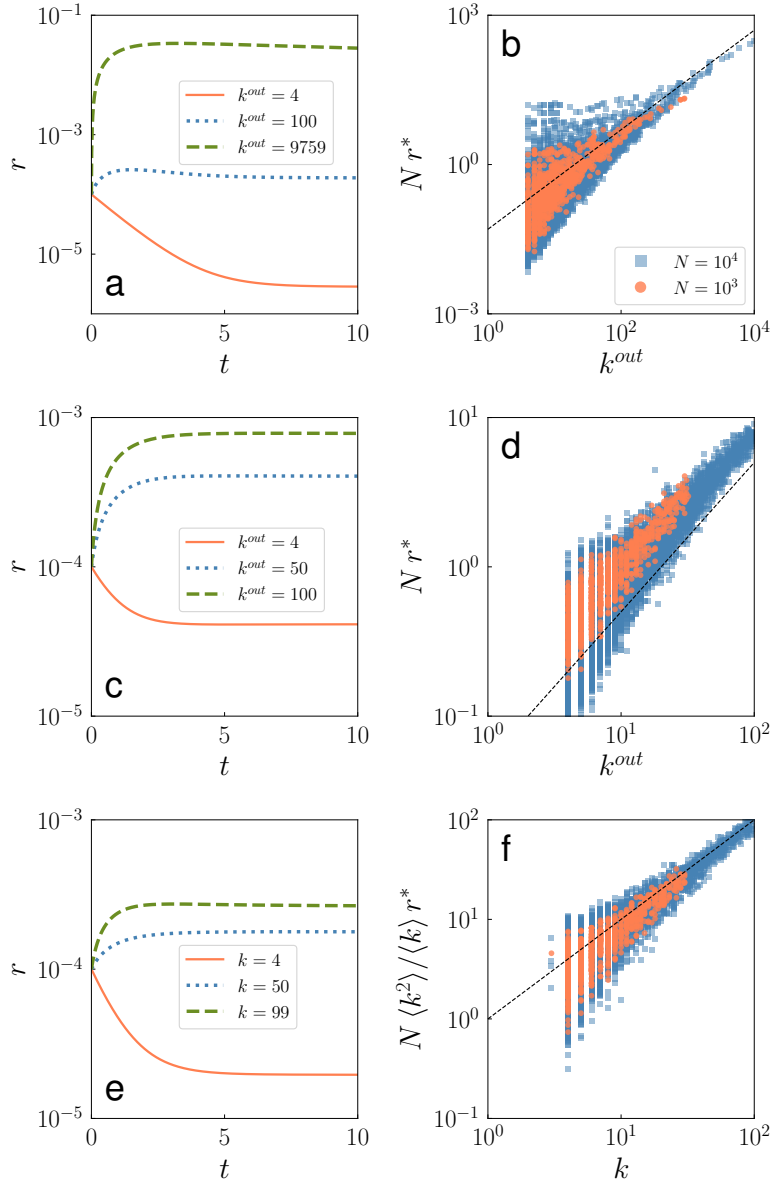
and

$$\ell_i = 1 \ . \tag{29}$$

Figure 30. Same as in Figure 29, but for the SIS model. Results are obtained by numerically integrating Eqs. (30). In panels a and b, we consider instances of the directed configuration model with $k_{max}^{out} = N - 1$. In panels c and d, we consider instances of the directed configuration model with $k_{max}^{out} = \sqrt{N}$. In panels e and f instead, we consider instances of the undirected configuration model.

1. Individual-based mean-field approach

As for the case of the CIC model, insights regarding the physical behavior of the system can be obtained by working under the individual-based mean-field approach, and write Eq. (14) as

$$\frac{ds_i(t)}{dt} = -s_i(t) + \lambda[1 - s_i(t)] \sum_j A_{ji} s_j(t).$$ (30)

The equation above follows immediately from the fact that $g_i(t) := \langle \chi_i(t) \rangle = \lambda \sum_j A_{ji} s_j(t)$ and $l_i(t) := \langle \ell_i(t) \rangle = 1$.

27

As the phase transition of the SIS model is continuous, in the vicinity of the critical point the whole density of infected individuals is small. So will be each s_i. We can thus linearize the previous equation to obtain

$$\frac{ds_i}{dt} = -s_i + \lambda \sum_j A_{ji} s_j \ . \tag{31}$$

The entire system of equations for every node can be therefore written in the compact matrix-vector form as

$$\frac{d\vec{s}}{dt} = (\lambda A^T - I)\vec{s} \ . \tag{32}$$

The endemic state will start to emerge when the largest eigenvalue of the operator $\lambda A^T - I$ equals one so that the critical value of the epidemic rate is given by

$$\lambda_c = \frac{1}{\omega_N} \ , \tag{33}$$

with ω_N the largest eigenvalue of the matrix A [1].

a. *Undirected networks*

We can repeat an argument similar to the one used for the CIC model, expanding the solution as

$$\vec{s}(t) = \sum_n \left(\vec{w}_n \cdot \vec{s}(0) \right) e^{-t/\tau_n} \vec{w}_n \ , \tag{34}$$

where $\vec{w}_n$ is the n-th eigenvector of A with associated eigenvalue ω_n, $\vec{s}(0)$ is the initial condition. The coefficient associated with the n-th eigenvector decays exponentially fast to zero, over a typical time scale equal to $\tau_n = \omega_N/(\omega_N - \omega_n)$. The relaxation time to the stationary state is thus due to the exponential decay to zero of the second eigenmode, whose decay time is $t^* = \tau_{N-1} = \omega_N/[\omega_N - \omega_{N-1}]$. For $t \gg t^*$ the components of the vector $\vec{s}$ are hence proportional to those of the principal eigenvector of the matrix A

$$s_i = s^* w_{N,i} \ , \tag{35}$$

where $s^* = \vec{w}_N \cdot \vec{s}(0)$.

We take as initial condition $s_i(0) = \delta_{i,i_{seed}}$, where i_{seed} is the node where the avalanche is seeded. Hence $s^* = w_{N,i_{seed}}$. For $\gamma < 5/2$ [20] the components of the principal eigenvector are

$$w_{N,i} = \frac{k_i}{[N\langle k^2 \rangle]^{1/2}}. \tag{36}$$

As a consequence, in the stationary state the variables s_i reach the asymptotic value,

$$s_i = w_{N,i_{seed}} w_{N,i} = \frac{k_i k_{i_{seed}}}{N\langle k^2 \rangle}. \tag{37}$$

At odds with the case of CIC dynamics, in the stationary state the variables s_i do not reach the same asymptotic value, being s_i dependent on k_i and thus, in general, $s^* \neq r^*$. For the same reason the value of r^* depends on the initial seed, and in particular it is proportional to k_{seed}. In Fig. 30 (e,f) we plot $r^* = 1/N \sum_i s_i$: according to Eq. (37) this quantity is equal to $k_{seed}\langle k \rangle / N\langle k^2 \rangle$. The figure confirms this prediction.

———

[1] We restrict ourselves to the case $\gamma < 5/2$. For $\gamma > 5/2$ the SIS epidemic transition is due to a completely different mechanism and Eq. (33) is not a good estimate of the critical value [21].

b.   Directed networks

As in the previous case, we set $\lambda_c = 1/\omega_N$ and note that the stationary solution of Eq. (32) is again proportional to the principal (right) eigenvector of A^T, $\vec{w}_N^{(r)}$. This can be simply understood in a discrete-time description, where the stationary solution emerges as a consequence of the repeated application of the matrix $\lambda A^T - I$.

We show in Fig. 30 (upper panels) that also in the directed case the constant r^* is proportional to k_{out} of the initial seed, but with a different scaling with the system size N. As Eq. (36) states for the undirected case, it is safe to presume that the stationary value of s_i again depends on i, i.e. that the vector $\vec{w}_N^{(r)}$ is not constant.

2.   Langevin approach

As in the CIC case, the IBMFA tells us that, in the long time limit, the microscopic variables can be well approximated by Bernoulli variables with probability of success $P(\sigma_i = 1) = w_{N,i}^{(r)} r^*$. Performing, again as in the CIC case, the adiabatic approximation, the stochastic variables σ_i can be thought as weakly fluctuating and thus be approximated as

$$\sigma_i \simeq \rho^* \, w_{N,i}^{(r)} \, .$$

The drift terms $\Psi_i[\vec{\sigma}] = 0$. Neglecting quadratic terms, we can write the third equation in (11) as

$$D_i[\vec{\sigma}] = \sigma_i + \lambda(1 - \sigma_i) \sum_j A_{ji}\sigma_j = \sigma_i + \lambda \sum_j A_{ji}\sigma_j = 2\sigma_i = 2\rho^* \, w_{N,i}^{(r)} \, .$$

We can thus write

$$D[\vec{\sigma}] = \frac{1}{N^2} \sum_i D_i[\vec{\sigma}] = \frac{2\rho^* \sum_i w_{N,i}^{(r)}}{N^2} = \frac{2}{N} \, \rho^* \, \langle w_N^{(r)} \rangle \, ,$$

where $\langle w_N^{(r)} \rangle$ is the (network-wide) average value of the components of the eigenvector $\vec{w}_N^{(r)}$. We can then write a Langevin Eq. (11) as

$$\frac{d\rho}{dt} = \xi \, \sqrt{\rho} \, \sqrt{\frac{2 \, \langle w_N^{(r)} \rangle}{N}} \, . \tag{38}$$

Still, this equation is of the same form as those considered in Ref [6], thus allowing us to state that duration and size of avalanches obey power-law distributions characterized by exponents typical of the mean-field branching process.

In summary, for SIS dynamics these results point out that, on undirected networks, MF exponents should be observed for long times, after a preasymptotic regime lasting for a time $t^* = \omega_N/[\omega_N - \omega_{N-1}]$. For scale-free networks with $\gamma < 5/2$ the spectral gap between the largest and the second-largest eigenvalue of the adjacency matrix diverges as the size grows [27]. As a consequence in the thermodynamic limit t^* tends to 1 and only the MF regime can be observed, in agreement with the tendency observed in Fig. (16).

F.   Contact process

According to the rules of the dynamics, the probabilities P_{ζ_i} and P_{η_i} are

$$P_{\zeta_i}(0) = dt$$

and

$$P_{\eta_i}(1) = \lambda\, dt \sum_j \frac{A_{ji}}{k_j^{out}}\, \sigma_j \ .$$

The out-degree dependence is suppressed by the factor $1/k_j^{out}$. As a consequence, CP cannot display anomalous behavior even in networks with broad out-degree distributions.

G. Invasion process

According to the rules of the dynamics, the probabilities P_{ζ_i} and P_{η_i} are

$$P_{\zeta_i}(0) = dt \sum_j \frac{A_{ji}}{k_j^{out}}\, (1 - \sigma_j)$$

and

$$P_{\eta_i}(1) = dt \sum_j \frac{A_{ji}}{k_j^{out}}\, \sigma_j \ .$$

The same considerations as of CP apply here as well.

[1] Andrea Baldassarri, Francesca Colaiori, and Claudio Castellano. Average shape of a fluctuation: Universality in excursions of stochastic processes. *Phys. Rev. Lett.*, 90(6):060601, 2003.

[2] Marián Boguñá, Claudio Castellano, and Romualdo Pastor-Satorras. Langevin approach for the dynamics of the contact process on annealed scale-free networks. *Physical Review E*, 79(3):036110, 2009.

[3] Javier Borge-Holthoefer, Alejandro Rivero, Iñigo García, Elisa Cauhé, Alfredo Ferrer, Darío Ferrer, David Francos, David Iñiguez, María Pilar Pérez, Gonzalo Ruiz, et al. Structural and dynamical patterns on online social networks: the spanish may 15th movement as a case study. *PloS one*, 6(8):e23883, 2011.

[4] Michele Catanzaro, Marián Boguná, and Romualdo Pastor-Satorras. Generation of uncorrelated random scale-free networks. *Physical Review E*, 71(2):027103, 2005.

[5] John S Caughman and JJP Veerman. Kernels of directed graph laplacians. *the electronic journal of combinatorics*, 13(1):39, 2006.

[6] Serena di Santo, Pablo Villegas, Raffaella Burioni, and Miguel A Muñoz. Simple unified view of branching process statistics: Random walks in balanced logarithmic potentials. *Physical Review E*, 95(3):032115, 2017.

[7] William Feller. Two singular diffusion problems. *Ann. Math.*, pages 173–182, 1951.

[8] JP Gleeson, JA Ward, KP O'Sullivan, and WT Lee. Competition-induced criticality in a model of meme popularity. *Physical Review Letters*, 112(4):048701, 2014.

[9] K-I Goh, D-S Lee, B Kahng, and D Kim. Sandpile on scale-free networks. *Physical review letters*, 91(14):148701, 2003.

[10] Theodore E Harris. *The theory of branching processes*. Dover, New York, 1989.

[11] Hans Karl Janssen. Survival and percolation probabilities in the field theory of growth models. *J. Phys. Cond. Matter*, 17(20):S1973, 2005.

[12] Leo P Kadanoff, Sidney R Nagel, Lei Wu, and Su-min Zhou. Scaling and universality in avalanches. *Phys. Rev. A*, 39(12):6524, 1989.

[13] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. Graphs over time: densification laws, shrinking diameters and possible explanations. In *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, pages 177–187. ACM, 2005.

[14] T.M. Liggett. *Interacting Particle Systems*. Classics in Mathematics. Springer, New York, 2004.

[15] Alan Mislove, Massimiliano Marcon, Krishna P. Gummadi, Peter Druschel, and Bobby Bhattacharjee. Measurement and Analysis of Online Social Networks. In *Proceedings of the 5th ACM/Usenix Internet Measurement Conference (IMC'07)*, San Diego, CA, October 2007.

[16] Michael Molloy and Bruce Reed. A critical point for random graphs with a given degree sequence. *Random structures & algorithms*, 6(2-3):161–180, 1995.

[17] Miguel A Muñoz, Ronald Dickman, Alessandro Vespignani, and Stefano Zapperi. Avalanche and spreading exponents in systems with absorbing states. *Phys. Rev. E*, 59(5):6175, 1999.

[18] Miguel A Muñoz, G Grinstein, and Yuhai Tu. Survival probability and field theory in systems with absorbing states. *Phys. Rev. E*, 56(5):5101, 1997.

[19] Richard Otter. The multiplicative process. *Ann. Math. Stat.*, 20:206–224, 1949.

[20] Romualdo Pastor-Satorras and Claudio Castellano. Distinct types of eigenvector localization in networks. *Sci. Rep.*, 6:18847, jan 2016.

[21] Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani. Epidemic process in compex networks. *Reviews of modern physics*, 87(3):925, 2015.

[22] Michael Plischke and Birger Bergersen. *Equilibrium statistical physics*. World Scientific, Singapore, 2006.

[23] Filippo Radicchi. Underestimating extreme events in power-law behavior due to machine-dependent cutoffs. *Physical Review E*, 90(5):050801, 2014.

[24] Sidney Redner. *A guide to first-passage processes*. Cambridge University Press, Cambridge, 2001.

[25] A Saichev, A Helmstetter, and D Sornette. Power-law distributions of offspring and generation numbers in branching models of earthquake triggering. *Pure and applied geophysics*, 162(6-7):1113–1134, 2005.

[26] James P Sethna, Karin A Dahmen, and Christopher R Myers. Crackling noise. *Nature*, 410(6825):242–250, 2001.

[27] Diogo H Silva, Silvio C Ferreira, Wesley Cota, Romualdo Pastor-Satorras, and Claudio Castellano. Spectral properties and the accuracy of mean-field approaches for epidemics on correlated networks. *arXiv preprint arXiv:1907.02144*, 2019.

[28] Vishal Sood, Tibor Antal, and Sidney Redner. Voter models on heterogeneous networks. *Physical Review E*, 77(4):041121, 2008.

[29] D Sornette. *Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization and Disorder: Concepts and Tools (Springer Series in Synergetics)*. Springer, 2006.

[30] JJP Veerman and E Kummel. Diffusion and consensus on weakly connected directed graphs. *arXiv preprint arXiv:1807.09846*, 2018.

[31] Henry William Watson and Francis Galton. On the probability of the extinction of families. *J. Roy. Anthropol. Inst.*, 4:138–144, 1875.

[32] Stefano Zapperi, Kent Bækgaard Lauritsen, and H Eugene Stanley. Self-organized branching processes: mean-field theory for avalanches. *Phys. Rev. Lett.*, 75(22):4071, 1995.