Temporal profiles of avalanches on networks

James P Gleeson

MACSI, Department of Mathematics and Statistics, University of Limerick, Ireland;
james.gleeson@ul.ie

Rick Durrett

Department of Mathematics, Duke University, Durham, NC, USA

19 Dec 2016

Abstract

An avalanche or cascade occurs when one event causes one or more subsequent events, which in turn may cause further events in a chain reaction. Avalanching dynamics are studied in many disciplines, with a recent focus on average avalanche shapes, i.e., the temporal profiles that characterize the growth and decay of avalanches of fixed duration. At the critical point of the dynamics the average avalanche shapes for different durations can be rescaled so that they collapse onto a single universal curve. We apply Markov branching process theory to derive a simple equation governing the average avalanche shape for cascade dynamics on networks. Analysis of the equation at criticality demonstrates that nonsymmetric average avalanche shapes (as observed in some experiments) occur for certain combinations of dynamics and network topology; specifically, on networks with heavy-tailed degree distributions. We give examples using numerical simulations of models of information spreading, neural dynamics, and threshold models of behaviour adoption.

1 Introduction

The dynamics of avalanches or cascades are studied in many disciplines. Examples include the spreading of disease (or information) from human to human [1, 2], avalanches of neuron firings in the brain [3], and the "crackling noise" exhibited by earthquakes and magnetic materials [4]. Of particular interest are cases with dynamics poised at a critical point, where universal scalings of avalanches are observed. The most commonly studied feature of such systems is the distribution of avalanche sizes, which has a power-law scaling at the critical point. The observation of heavy-tailed distributions of avalanche sizes has therefore been used to indicate whether a system is critical. However, power-law distributions can also arise from mechanisms other than criticality [5, 6], so recently attention has focussed more upon the temporal aspects of avalanches, which also exhibit universal characteristics at criticality.

The average avalanche shape is determined by averaging the temporal profiles of all avalanches that have a fixed duration T (see Fig. 1). At criticality, the average avalanche shape is a universal function of the rescaled time t/T, meaning that the average avalanche shapes for different durations can be rescaled to collapse onto a single curve [4]. This feature has recently been used as a sensitive

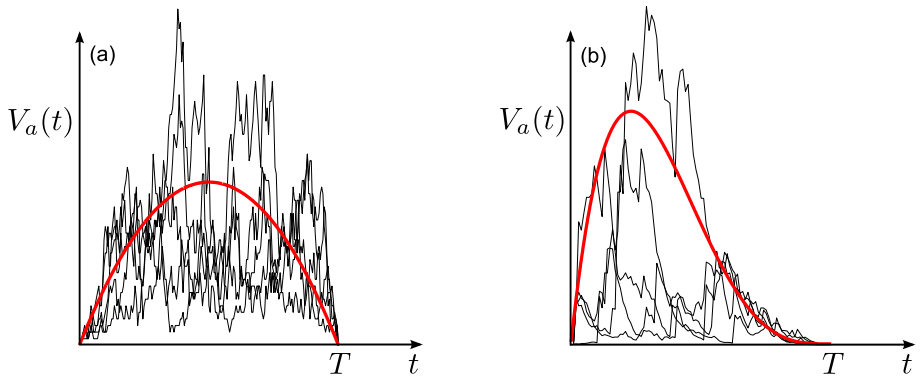

Figure 1: Examples of (a) symmetric and (b) nonsymmetric average avalanche shapes. In each panel, the black curves show five examples of individual avalanches that all have duration T. The average avalanche shape for the duration T (red curve) is found by averaging the temporal profiles of all such avalanches.

test for criticality in a range of dynamics, from the Barkhausen effect in ferromagnetic materials [7] to neural avalanches [3, 8] and electroencephalography (EEG) recordings from hypoxic neonatal cortex [9]. While average avalanche shapes are typically parabolic (and symmetric) functions of time like the example in Fig. 1(a), nonsymmetric (left-skewed) avalanche shapes resembling Fig. 1(b) have also been observed in experiments. For example, early observations of nonsymmetric avalanche shapes in experiments on Barkhausen noise [4] raised doubts about whether the theoretical model used in [10,11] was in the correct universality class. Although this discrepancy between theory and experiment was later resolved by a more detailed theory for avalanche propagation [12,13], several instances of nonsymmetric avalanche shapes (e.g., the neural avalanches in Fig. S4 of [3]) still lack explanation. Despite some progress in modelling avalanche profiles using random walks [9,14] and self-organized criticality models [15–18], the factors that cause nonsymmetric average avalanche shapes remain poorly understood.

The characteristics of avalanches that occur on networks depend on both the network connectivity and the node-to-node dynamics of the cascade [19]. Cascading models have been applied, for example, to power-grid blackouts [20], epidemic outbreaks [21], and to the propagation of memes (pieces of digital information) through online social networks [22, 23]. The avalanche-size distribution at criticality is known to depend non-trivially on the degree distribution of the underlying network [24]; in this paper we focus instead on the temporal profile of cascades, i.e., the average avalanche shape, and how it is affected by the network degree distribution. Using a mathematical derivation of the average avalanche shape for Markovian dynamics, we demonstrate that—as in other universality-breaking examples [25]—networks with heavy-tailed degree distributions can give rise to qualitatively different results from those found on networks with finite-variance degrees. However, the dynamics of the avalanching process are also important: we show that in fact it is the interaction between the dynamics and the network topology that determines whether average avalanche shapes are symmetric or not.

2 Dynamics and networks

To define the *average avalanche shape* we consider the set S_T of all avalanches that are of duration T (meaning that the avalanche has terminated at a time T after its first event, with no further events occurring at any time $t > T$). Each avalanche a in the set S_T is described by a function

$V_a(t)$, which is the number of events that occur at time t in that avalanche (so $V_a(t) = 0$ for $t > T$). The avalanche shape for the duration T is defined as the average of the functions $V_a(t)$, taken over all the avalanches a in set S_T.

We consider networks that are defined by their degree distributions, but are otherwise maximally random ("configuration-model" networks [26]). For undirected networks, the degree distribution p_k is the probability that a randomly-chosen node has degree (number of neighbours) equal to k. For directed networks, the joint degree distribution p_{jk} is the probability that a random node has in-degree j and out-degree k. We denote the mean degree by z (so $z = \sum_k k p_k$ for undirected networks and $z = \sum_{j,k} k p_{jk} = \sum_{j,k} j p_{jk}$ for directed networks). Such configuration-model networks are locally tree-like, which facilitates the use of the branching-process approximations that we employ. We assume that the networks consist of a single connected component and that they are large enough to permit us to use infinite-size approximations.

Our focus is on discrete-state dynamics, where each node of the network can be in one of a set of discrete states at each moment in time; transitions between states may occur continuously in time, or only at discrete time steps. Cascades occur when nodes successively switch to one specific state, which we will call the "active" state; we will generically refer to all other states as "inactive". Once a node is activated (i.e., once it transitions to the active state), it affects its neighbouring nodes by increasing the probability that they will also become activated at a later time. We focus on unidirectional dynamics, meaning that in the case where a node activates some of its neighbouring nodes and then subsequently becomes inactive, the neighbouring nodes cannot directly re-activate it. One important class of such dynamics includes cases where an activated node cannot subsequently return to the inactive state, and so cannot be re-activated (this class is called "monotonic dynamics" in [27]). Another class takes place on tree-like directed networks, which have negligible numbers of loops, so that activation of node i can affect its out-neighbours, but there exists no path for the out-neighbours to subsequently affect the state of node i (even if node i returns to the inactive state).

Each cascade is assumed to be initiated by a randomly-chosen single node, called the "seed" node, which is activated at the beginning of the process while all other nodes are inactive; subsequent to the activation of the seed node, the cascade of activation of nodes proceeds according to the rules of the model under consideration. Some examples of such models are as follows.

- Threshold models, on an undirected network: Each node i possesses a positive threshold R_i that is assigned randomly from a distribution. When an inactive node i of degree k_i is chosen for updating it considers the number m_i of its neighbours who are active, and makes a decision according to the rules of the specific model. In the Watts threshold model [28], for example, node i becomes active if the fraction m_i/k_i is greater than, or equal to, the node's threshold R_i, i.e., the node activates if the fraction of its neighbours who are active is sufficently large, otherwise the node remains inactive. An alternative type of threshold model is that of Centola and Macy [29], wherein node i actives if the total number of active neighbours (rather than the fraction of such neighbours) is large enough: $m_i \geq R_i$ (see also the discussion of threshold models and their relation to coordination games in [27]).

- The neuronal dynamics model of [3]: Each node in a weighted, directed network represents a neuron. The weight ϕ_{ij} on the edge connecting node i to node j is chosen at random from a uniform distribution on the interval $(0, \phi_{\max})$, where $\phi_{\max}$ is a tunable parameter [30, 31]. (In [3] the values of ϕ_{ij} are instead inferred from experiments on neural networks). Using the same model parameters as [3], the neurons are modelled by binary-state elements as a very simple approximation of integrate-and-fire dynamics: whenever neuron i fires (becomes active), it causes neuron j to become active (in the next discrete time step) with probability

ϕ_{ij}. After a neuron fires, it is returned to the inactive state in the next time step. In [3], exogenous input noise is added to the system to ensure continuous neural activity. Since we focus purely on the avalanche dynamics, we instead randomly select a node to be the seed node of the cascade and activate it in the first time step, and record the ensuing avalanche of activations.

- The model of [32] for meme propagation on a directed social network (like Twitter): Each node (of N) in an directed network represents a user of the social network. The out-degree k_i of a node i is the number of its "followers" in the network: these are the users that receive the "tweets" (or distinct pieces of digital information, generically called "memes") sent by node i. Each user also retains a memory of the last meme received from the nodes it follows via a "screen" that is overwritten when a new meme is received. (More realistic models that incorporate longer memory are described in [33,34]). In each time step (with $\Delta t = 1/N$) one node is chosen at random and with probability μ the node "innovates" by creating a new meme, placing it on its screen, and tweeting this meme to all its followers (where the new meme overwrites any existing memes on their screens). Alternatively (with probability $1-\mu$), the chosen node "retweets" the meme that is currently on its screen. A newly-innovated meme can therefore experience an avalanche of popularity as it is retweeted multiple times before it eventually is forgotten by all users in the network, at which time the avalanche terminates. The analyses of [32] and [34] show that the avalanche dynamics of the memes are critical in the limit $\mu \to 0$ and subcritical for $\mu > 0$. Subsequent work [34] showed that the distribution of avalanche sizes for empirical data on hashtags can be fitted using a value of the innovation probability μ of 0.033.

Other examples of unidirectional dynamics to which our results apply include the zero-temperature random-field Ising model [4] and the susceptible-infected-recovered (SIR) disease-spread model on networks [21].

3 Analytical approach

3.1 The offspring distribution

The central quantity in our branching process analysis is called the *offspring distribution*, denoted by q_k ($k = 0, 1, 2, \ldots$). Roughly speaking, this distribution gives the likelihood that if a cascade of activation reaches a node (i.e., if one of the node's neighbours becomes active), that the node will activate and expose k other neighbouring nodes to potential activation (see Appendix A for details of the branching process approximation). For an undirected network, we define q_k as

$$\text{(undirected)} \quad q_k = \frac{k+1}{z} p_{k+1} v_{k+1} \quad \text{for } k \geq 1, \tag{1}$$

noting that the probability of reaching a node of degree $k+1$ by travelling along a random edge is $\frac{k+1}{z} p_{k+1}$, and if this edge spreads the activation to the node it has k remaining inactive neighbours. The quantity v_k is the probability that a node i of degree k is *vulnerable* [28], meaning that the activation of a single neighbouring node (at time t_1) will lead to the activation of node i at some time $t > t_1$, assuming that no other neighbour of node i becomes active by time t.

For a directed network, q_k is defined as

$$\text{(directed)} \quad q_k = \sum_j \frac{j}{z} p_{jk} v_{jk} \quad \text{for } k \geq 1, \tag{2}$$

where the factor $\frac{j}{z}p_{jk}$ represents the probability of reaching a node of in-degree j and out-degree k by travelling along a random edge of the network. The vulnerability v_{jk} is the probability that the activation of a single in-neighbour (at time t_1) of a node i (of in-degree j and out-degree k) will lead to the activation of node i at some time $t > t_1$, assuming no other in-neighbour of node i becomes active by time t. In both the directed and undirected cases, the probability of having zero offspring is given by $q_0 = 1 - \sum_{k\geq 1} q_k$.

We now give examples of how the vulnerabilities v_k and v_{jk} are calculated for the models introduced in Sec. 2.

- Threshold models on an undirected network: The vulnerability v_{k+1} is the probability that a node i, of degree $k+1$, activates when exactly one of its neighbours is active, i.e., when $m_i = 1$. According to the rules of the Watts threshold model, for example, node i will become active if its threshold R_i is less than, or equal to, $1/(k+1)$, which is the fraction of its neighbours that are active when $m_i = 1$. The probability that the node's threshold is less than this value is given by

$$\text{(Watts)} \quad v_{k+1} = F\left(\frac{1}{k+1}\right), \tag{3}$$

  where F is the cumulative distribution function of the thresholds. In the Centola-Macy threshold mode, a node with one active neighbour will activate if its threshold is less than or equal to 1, and so the vulnerability in this case is

$$\text{(Centola-Macy)} \quad v_{k+1} = F(1). \tag{4}$$

- In the neuronal dynamics model of [3] a node is vulnerable if it becomes active when one of its in-neighbours fires. According to the rules of the model, this occurs with a probability that depends on the edge between the two nodes, but is chosen from a uniform distribution on $(0, \phi_{\max})$, independently of the degrees of the nodes. The probability that a random node with in- and out-degrees j and k is vulnerable is therefore the average of this uniform distribution, i.e., $v_{jk} = \phi_{\max}/2$.

- The model of [32] for meme propagation on a directed social network: We focus on a chosen meme and assume that this meme has been tweeted by an in-neighbour of node i (i.e., by one of the users followed by user i). The probability that node i will subsequently retweet this meme (before user i's memory is overwritten by other tweets it receives) is calculated in Sec. IVA of [34] as

$$v_{jk} = \frac{1-\mu}{j+\mu}, \tag{5}$$

  where μ is the innovation probability.

Inserting these v_k and v_{jk} functions into Eqs. (1) and (2), respectively, defines the offspring distribution for each of the models.

The *branching number* defines whether the dynamical process on a given network is subcritical, critical, or supercritical. The branching number is the mean of the offspring distribution, i.e., the expected number of "children" per "parent":

$$\xi = \sum_k k\, q_k, \tag{6}$$

and the value $\xi = 1$ is the critical value, separating the subcritical case ($\xi < 1$) from the supercritical case ($\xi > 1$). In the critical case, power-law distributions of avalanche sizes are observed [32, 34] but in this paper we focus on the temporal profiles of the avalanches.

3.2 Average avalanche shape

The detailed derivation of our results from the theory of Markov branching processes is given in Appendix B. For models with continuous-time updating a particularly simple result is found: the average avalanche shape $A(t)$ for avalanches of duration T can be expressed as

$$A(t) = \frac{Q(T-t)\left[f'\left(Q(T)\right) - f'\left(Q(T-t)\right)\right]}{f\left(Q(T-t)\right) - Q(T-t)}, \tag{7}$$

where $f(x)$ is the generating function for the offspring distribution,

$$f(s) = \sum_k q_k s^k \tag{8}$$

and $Q(t)$ is the fraction of avalanches that are extinct by time t, which is given by the solution of the ordinary differential equation

$$\frac{dQ}{dt} = f(Q) - Q, \text{ with } Q(0) = 0. \tag{9}$$

Thus the average avalanche shape for a given offspring distribution and duration T can be calculated by solving only one ordinary differential equation, Eq. (9), and then using the solution function $Q(t)$ in the explicit formula of Eq. (7). We also show (in Appendix E) that the average avalanche shape for discrete-time updating can be found in a similar fashion, but the resulting expression is less amenable to analysis than the continuous-time case. However, qualitatively similar results are found for both continuous-time and discrete-time updating (see Fig. S2), so we focus mainly on the solution given by Eqs. (7) and (9).

In Figure 2 we show the avalanche shapes that are given by Eq. (7) in the case where the offspring distribution q_k is a Poisson distribution with mean ξ. In the critical case ($\xi = 1$) we see from Figs. 2(a) and 2(b) that a rescaling of time and of avalanche height causes the avalanche shapes for different durations to collapse onto a single symmetric curve: this scaling collapse (although not the shape of the curve) is predicted by the universality arguments of [4]. Note that a case where Eq. (9) is exactly solvable (binary fission) is examined in Appendix C and is shown to give a parabolic avalanche shape at criticality. However, for subcritical ($\xi < 1$) avalanches the profiles are non-parabolic (Fig. 2(c)) and do not collapse onto a universal curve (Fig. 2(d)). For supercritical ($\xi > 1$) avalanches (where we only consider those avalanches which terminate at a finite time T: a non-zero fraction of avalanches also exist that never terminate), very similar shapes are observed to the subcritical case (Figs. 2(e) and 2(f)).

Next, we consider the case where the offspring distribution has a power-law tail:

$$q_k \sim C\, k^{-\gamma} \quad \text{as } k \to \infty, \tag{10}$$

with exponent γ in the range $2 < \gamma < 3$ so that the second moment of the offspring distribution is infinite. (Here, and throughout the paper, we use C to denote a constant prefactor in an asymptotic scaling relation). Such cases are of practical interest because scale-free degree distributions are well-known [25,26,35] to strongly affects dynamics on networks, and (as we discuss in Sec. 3.4) a scale-free degree distribution can, for certain dynamics, lead to offspring distributions such as Eq. (10). Note that the second moment of the offspring distribution is related to the second derivative of the generating function of Eq. (8) evaluated at $s = 1$, so this case corresponds to $f''(1) = \infty$.

Using offspring distributions of the form (10) in Eqs. (7) and (9) gives the avalanche shapes shown in Figure 3, where we focus on the critical case for three different values of the power-law

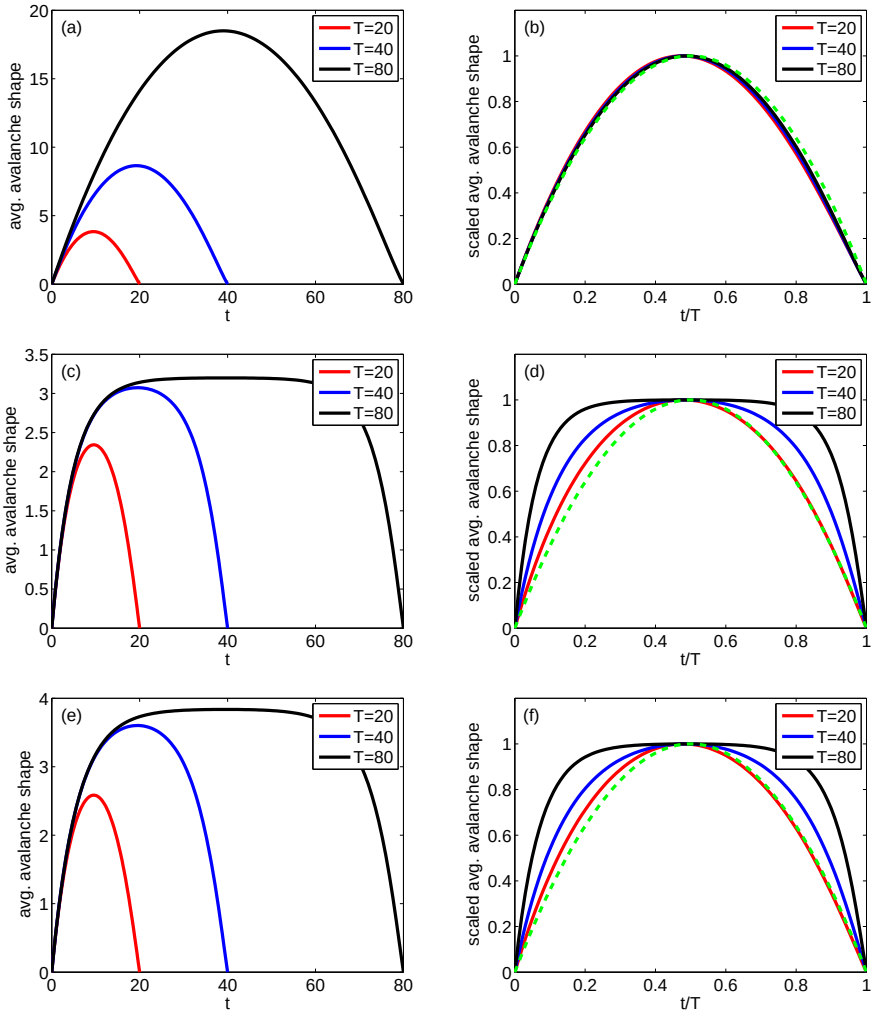

Figure 2: (Left column: panels (a), (c) and (e)): Average avalanche shapes from Eq. (7), for Poisson offspring distribution q_k, with mean ξ. Each panel in the right column (panels (b), (d) and (e)) shows the same function as in the panel to its left, but plotted (as, for example, in [7]) in rescaled time t/T and rescaled vertically to have a maximum value of 1. Panels (a) and (b): critical case with $\xi = 1$; panels (c) and (d): subcritical case, $\xi = 0.8$; panels (e) and (f): supercritical case, $\xi = 1.2$. The dashed green curve in panels (b), (d), and (f) is the parabola $4\frac{t}{T}\left(1 - \frac{t}{T}\right)$.

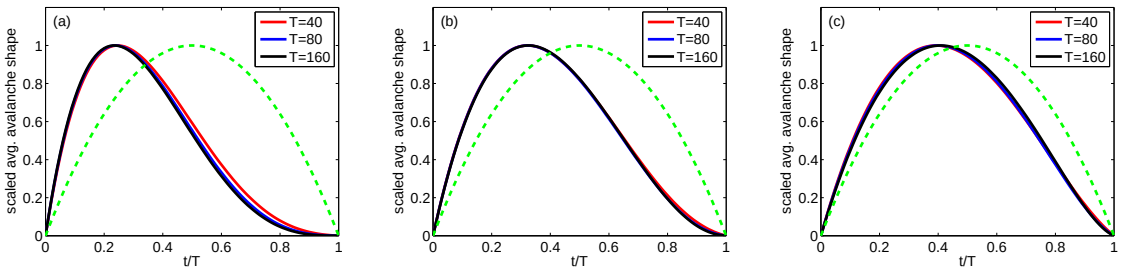

Figure 3: Rescaled average avalanche shapes for power-law offspring distribution: $q_k = C\,k^{-\gamma}$ for $k \geq 1$, with constant C chosen to give the critical case $\xi = 1$. Power-law exponents are: (a) $\gamma = 2.3$, (b) $\gamma = 2.5$, (c) $\gamma = 2.9$; the dashed green curve is the same parabola as in Fig. 2.

exponent γ. Clearly, the avalanche shapes are nonsymmetric, with a leftward skew. (This contrasts with the right-skewed avalanche shapes found from random-walk models with long memory [14]). A detailed asymptotic analysis of the governing equation (Appendix F) enables us to conclude that, as $T \to \infty$ (and $T - t \to \infty$), the average avalanche shape scales as

$$A(t) \sim \begin{cases} C\frac{t}{T}\,(T-t) & \text{if } f''(1) \text{ is finite,} \\ \\ C\frac{t}{T}\,(T-t)^{\frac{1}{\gamma-2}} & \text{if } q_k \propto k^{-\gamma} \text{ as } k \to \infty, \text{ with } 2 < \gamma < 3, \end{cases} \tag{11}$$

where the constant prefactor C is independent of T. Note that parabolic avalanche shapes (with peak at $t = T/2$) are seen in the large-T limit whenever the offspring distribution q_k has finite second moment, but the shape is nonsymmetric (with peak at $t = (\gamma-2)T/(\gamma-1) < T/2$) for power-law distributions with γ between 2 and 3.

True power-law tails are never seen in real systems, due to finite-size effects. If the offspring distribution instead has a truncated power-law form, with an exponential cutff for $k \gg \kappa$:

$$q_k \sim C\,k^{-\gamma} e^{-\frac{k}{\kappa}} \quad \text{as } k \to \infty, \tag{12}$$

and with $2 < \gamma < 3$, then the avalanche shapes for all durations T do not collapse onto a single curve, even at criticality. As the asymptotic analysis of Appendix F reveals, the shape for large T is determined by the survival function $1 - Q(t)$: this is the fraction of avalanches that remain alive at time t. For the offspring distribution of Eq. (12), the survival function $1 - Q(t)$ scales as $t^{-\frac{1}{\gamma-2}}$ for early times, but as t^{-1} for later times; the crossover from one regime to the other is determined by the exponential cutoff κ in the offspring distribution (see Fig. 4(a)). Therefore, it is possible to observe nonsymmetric shapes for avalanches with relatively short durations T, but the longer-duration avalanches (the $T \to \infty$ limit) revert to the parabolic shape typical of offspring distributions with finite second moment, see Fig. 4(b).

3.3 Other characteristic temporal shapes

The Markov branching process approach that we use to derive the average avalanche shape in Eq. (7) can also be applied to calculate other temporal characteristics of avalanches. In Appendix B, for example, we derive a formula for the variance of the avalanche shape (i.e., the variance of the set of functions $\{V_a(t) \,|\, a \in S_T\}$). The overall shape of the standard deviation is found to be similar to

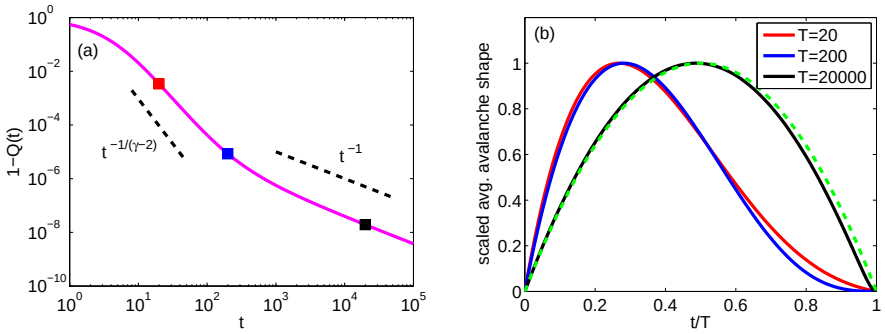

Figure 4: Survival function $1 - Q(t)$ (panel (a)) and rescaled avalanche shape (panel (b)) for offspring distributions with truncated power-law form: $q_k \propto k^{-\gamma} e^{-k/\kappa}$ with $\gamma = 2.3$, $\kappa = 10^6$. The constant of proportionality is determined by the criticality condition $\xi = 1$. The coloured squares in panel (a) mark the durations $T = 20$, $T = 200$, $T = 2 \times 10^4$ of the avalanches whose average shapes are plotted in panel (b). Note the $T = 20$ and $T = 200$ cases have the nonsymmetric profiles typical of power-law q_k, but the $T = 2 \times 10^4$ case closely matches the parabolic profile expected for offspring distributions with finite second moment.

the average avalanche shape: The asymptotic result for the critical case (in the limit $T \to \infty$ and $T - t \to \infty$, see Appendix F) is that the coefficient of variation at time t for avalanches of duration T is

$$\mathrm{CV}(t) = \frac{\sqrt{\text{variance}}}{A(t)} \sim \begin{cases} \frac{1}{\sqrt{2}} & \text{if } f''(1) \text{ is finite,} \\ \sqrt{\frac{3-\gamma+(\gamma-2)\frac{t}{T}}{(\gamma-1)\frac{t}{T}}} & \text{if } q_k \propto k^{-\gamma} \text{ as } k \to \infty, \text{ with } 2 < \gamma < 3. \end{cases} \tag{13}$$

Note that the coefficient of variation is constant (independent of t and T) for the case where parabolic avalanche shapes occur, so the standard deviation also has a parabolic profile. However, in the power-law case with $\gamma < 3$, the shape of the standard deviation is even more skewed than that of the average avalanche shape profile: the coefficient of variation limits to a constant as $t \to T$, but it diverges like $1/\sqrt{t}$ as $t \to 0$.

The universality of the average avalanche shape has been used in analyzing experimental data; specifically, the collapse of shapes for avalanches of different durations can help identify whether the dynamics is critical or not [3,4,7,9]. However, one drawback of the average avalanche shape is that it requires an accurate assessment of the time T at which each avalanche terminates. Pinpointing such termination times can be difficult in empirical data, especially for avalanches of information on social networks, many of which exhibit very long lifespans [34, 36, 37]. Another characteristic temporal shape that we can calculate analytically is the average shape of all avalanches that have not terminated by a given observation time T: such a characteristic may prove easier to calculate for empirical data than the standard avalanche shape. In Appendix D we show that the average non-terminating avalanche shapes at various observation times T collapse onto a single curve when the dynamics are critical. The universal curve is again found to have a parabolic form (but with peak at $t = T$) if $f''(1)$ is finite, and to have a skewed (non-parabolic) shape if the offspring distribution is power-law.

As we demonstrate with our numerical simulations below, an even simpler temporal profile can provide a very sensitive measure of whether an avalanching system is critical or not. The average

number of events observed at a time t after an avalanche begins is given by the average of $V_a(t)$ over the entire set of avalanches, regardless of the duration of the avalanche. (Note avalanches that terminated at a time T with $T < t$ contribute zero to the measure at time t). In Appendix D we show that the average number of such events at time t is an exponentially decaying function of t if the dynamics is subcritical, an exponentially growing function of t for supercritical dynamics, and is a constant (independent of t) for the critical case. This temporal characteristic is particularly useful in understanding the numerical simulation results of Sec. 4 below.

3.4 Power-law offspring distributions from network dynamics

In Secs. 3.2 and 3.3 we have shown that nonsymmetric average avalanche shapes occur within Markovian branching processes when the offspring distribution q_k has a power-law tail with exponent γ between 2 and 3. Here we examine how such offspring distributions might arise from unidirectional dynamics on undirected and directed networks, using the relationships given by Eqs. (1) and (2).

In the case of an undirected network, we suppose that the vulnerability v_k depends on the node degree k as

$$v_k \sim C\, k^{-\nu} \text{ as } k \to \infty. \tag{14}$$

Then, if the degree distribution of the network has a power-law tail:

$$p_k \sim C\, k^{-\alpha} \text{ as } k \to \infty, \tag{15}$$

the large-k asymptotics of q_k is given by Eq. (1) as

$$q_k \sim C\, k^{1-\alpha-\nu} \tag{16}$$

and so we write

$$\text{(undirected)} \quad \gamma = -1 + \alpha + \nu \tag{17}$$

for the power-law exponent of the offspring distribution. Note that in general the value of γ will be different from the power-law exponent α of the network's degree distribution.

The case of a directed network is complicated by the existence of the joint distribution p_{jk} of in-degree j and out-degree k. If we assume the simplest case of nodes having independent in- and out-degree, then the joint distribution factorizes: $p_{jk} = p_j^{\text{in}}\, p_k$, and the large-$k$ behaviour of the vulnerability can be specified by the weighted sum over in-degrees as

$$\sum_j \frac{j}{z}\, p_j^{\text{in}}\, v_{jk} \sim C\, k^{-\nu}. \tag{18}$$

Assuming a power-law out-degree distribution, as in Eq. (15), Eq. (2) yields

$$q_k \sim C\, k^{-\alpha-\nu} \text{ as } k \to \infty, \tag{19}$$

so that the power-law exponent of the offspring distribution is

$$\text{(directed)} \quad \gamma = \alpha + \nu. \tag{20}$$

4 Numerical simulations

In this section we report on numerical simulations of unidirectional dynamics on networks, to assess the applicability of the branching process theory developed in Sec. 3. We run numerical simulations of the example dynamics listed in Sec. 2 on synthetic (configuration-model) and real-world networks, recording average avalanche shapes (and other temporal characteristics) for comparison with our theory. The branching process paradigm is only an approximation for dynamics on networks, as its assumptions are invalidated by the existence of loops in the network and by the finite size of the networks [26]. We demonstrate that while such features indeed impact upon the agreement with theoretical results, the main qualitative feature of critical dynamics that we have identified—the appearance of nonsymmetric avalanches shapes when the offspring distributions of Eqs. (1) or (2) have sufficiently heavy tails—is indeed observable in numerical simulations of cascade dynamics on networks. Matlab/Octave code for the simulations is available for download from [38].

4.1 Meme propagation on a directed network

Figure 5 shows results from simulations of the meme propagation model of [32]. Memes are tweeted from user to user according to the rules described in Sec. 2; the popularity of each meme is tracked in the simulations, and the number of events in the avalanche profile of a meme is the number of times it is tweeted within a time interval. The updating within the model proceeds in time steps of $\Delta t = 1/N$, so for large N (we use networks with $N = 10^5$ nodes) the discrete-time dynamics approximates continuous-time updating. We record the average avalanche shape determined by all memes whose avalanches terminate at time T (using a bin of duration 0.5, so $T = 9$, for example, includes all avalanches that terminate at times in the range $[8.5, 9]$).

Two network structures are compared (see Table 1): the panels in the left column of Fig. 5 (panels (a), (c) and (e)) present results for a network with scale-free out-degree distribution $p_k \propto k^{-\alpha}$ with $\alpha = 2.5$, while the panels in the right column are for a network where every node has exactly $z = 10$ followers (note the mean degree of the two networks are approximately equal). In both cases the in-degree distribution is Poisson, and in-degrees and out-degrees of nodes are independent. We set the innovation parameter μ to zero, so we expect the dynamics to be critical (from Eqs. (2), (5) and (6)). Moreover, using the vulnerability from Eq. (5) in Eq. (18) gives a ν value of 0 for this model, so Eq. (20) predicts nonsymmetric avalanche shapes for values of α, the tail exponent of the network's out-degree distribution, between 2 and 3; note we use $\alpha = 2.5$ in the left column of Fig. 5.

According to the theory of Sec. 3, the rescaled average avalanche shape curves for different (and sufficiently large) avalanche durations T should collapse onto a single curve. We see good agreement with this prediction in Fig. 5(a): note the distinctively nonsymmetric shape of the collapsed curve. Although the average avalanche shapes for the z-regular out-degree network in Fig. 5(b) are not fully converged by $T = 9$, they are evidently approaching the parabolic profile expected for the case where the offspring distribution has finite second moment. Panels (c) and (d) contrast the average non-terminating avalanche shapes that are found on the two networks. On the power-law network the shape is clearly non-parabolic, while the match to the asymptotic expression of Eq. (66) is very good for the case of finite $f''(1)$. On both types of network the error bars are smaller than in panels (a) and (b), reflecting the fact that the set of avalanches that have not terminated by time T is much larger than the set of avalanches that terminate at T, so the average shape is better estimated using the larger data set. Panels (e) and (f) show the average number of events per unit time, as discussed at the end of Sec. 3.3. The criticality of the process is reflected in the fact that the average number of events (tweets) per unit time is constant over the timescale of the simulations.

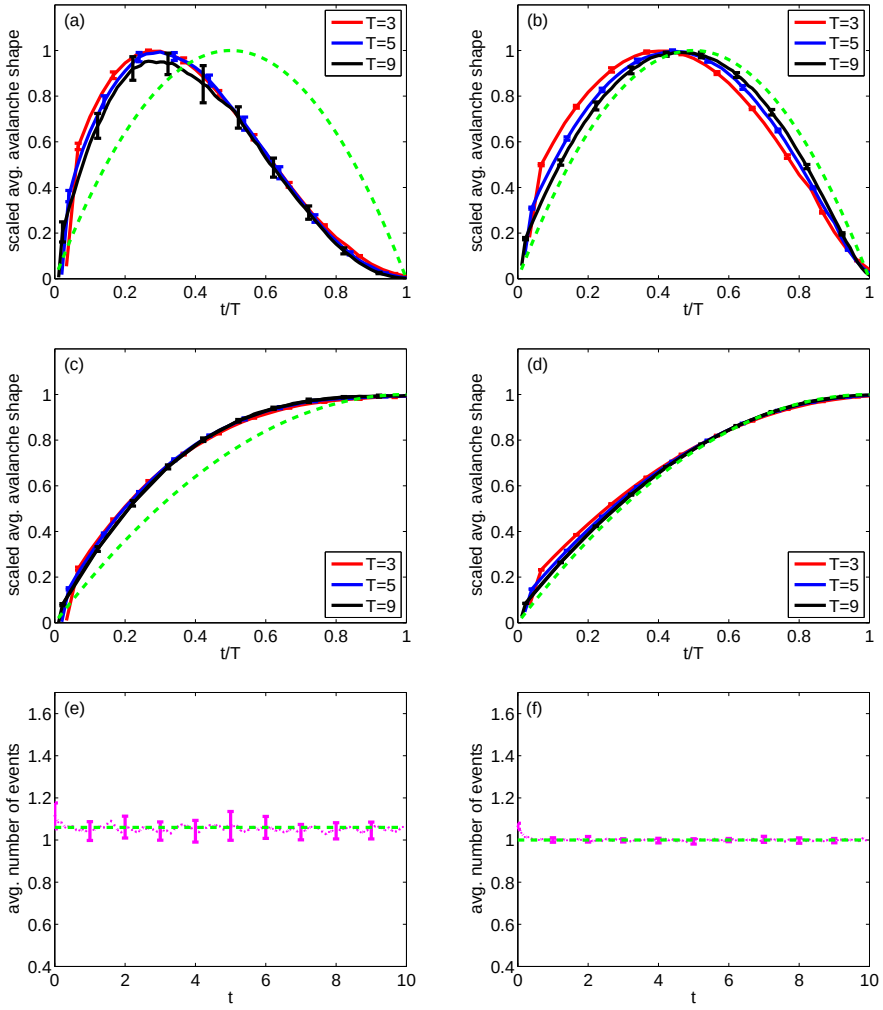

Figure 5: Numerical simulation results for critical meme-popularity avalanches within the model of [32]. The panels in the left column show results from a network with a power-law out-degree distribution, with exponent $\alpha = 2.5$. The results in the right column are from a network where every node has exactly $z = 10$ out-neighbours. See Table 1 for further details. Panels (a) and (b): rescaled average avalanche shapes. Panels (c) and (d): rescaled average non-terminating avalanche shapes. Panels (e) and (f): average number of tweets of the meme per unit time.

Fig.	Dynamics	Directed network	ν	Criticality condition	Avalanches n_A	Panels (a), (c), (e)			Panels (b), (d), (f)		Replicas n_R
						$p_k \propto k^{-\alpha}$	z	N	$p_k = \delta_{k,z}$	N	
5	Meme propagation	Yes	0	$\mu = 0$	10^6	$\alpha = 2.5$	10.6	10^5	$z = 10$	10^5	12
6	Neuronal dynamics	Yes	0	$\phi_{\max} = \frac{2}{z}$	10^7	$\alpha = 2.5$	10.6	10^5	$z = 10$	10^5	24
7	Threshold model	No	0	$\theta_{\max} = \frac{\langle k^2 - k \rangle}{z}$	10^6	$\alpha = 3.3$	2.9	10^6	$z = 3$	10^5	24
8	Meme propagation	Yes	0	$\mu = 0$	1.4×10^6	$z = 21.75$		$N = 81,306$			6

Table 1: Details of the numerical simulations that produce the results shown in Figs. 5–8. The parameter ν is determined by the vulnerability from Eq. (14) or Eq. (18). The criticality condition is $\xi = 1$, determined from Eq. (6) and Eq. (1) or Eq. (2). In each experiment, n_A individual avalanches are simulated to calculate the average avalanche shape and other measures. Results in the left column of each Figure (panels (a), (c) and (e)) are for networks with power-law p_k, with N nodes and mean degree z, while those in the right column of each Figure (panels (b), (d) and (f)) are for regular random networks. For each avalanche the seed node is changed and the order of node updates (for Figs. 5 and 8) or the dynamical parameters (the edge weights ϕ_{ij} for Fig. 6 or the node thresholds R_i for Fig. 7) are randomized. In order to quantify the robustness of the results, the complete experiment that results in the average avalanche shape is repeated a total of n_R times and the error bars in the Figures denote the standard deviation of the measures over the set of replica experiments.

4.2 Neuronal dynamics model on a directed network

Figure 6 shows results obtained from simulations of the neuronal dynamics model of [3], where the parameter $\phi_{\max}$ is tuned so as to poise the dynamics near to criticality. As in Fig. 5, the left column of results is for a network with power-law out-degree distribution of exponent $\alpha = 2.5$, while those in the right column are for a z-regular network. All neurons are synchronously updated in each discrete time step and we record the number of activated neurons at each step as the "events" of the avalanche; the avalanche terminates when no new neurons are activated.

Although the details of this discrete-time case differ from the continuous-time case of Fig. 5, the results again qualitatively agree with theory: here the vulnerability $v_{jk} = \phi_{\max}/2$ is independent of node degree k, and so $\nu = 0$ in Eq. (18), giving $\gamma = \alpha = 2.5$ from Eq. (20) for the power-law network. As expected, we see nonsymmetric average avalanche shapes and non-parabolic average non-terminating avalanche shapes in panels (a) and (c), while the corresponding results on the finite-second-moment network (panels (b) and (d)) agree closely with the parabolic asymptotic shapes of Eqs. (11) and (66).

4.3 Threshold model on an undirected network

In Figure 7 we consider threshold dynamics on an undirected network. Specifically, we show results for a Centola-Macy threshold model, with a uniform distribution of thresholds on the interval $(0, \theta_{\max})$, where the parameter $\theta_{\max}$ is tuned to place the dynamics close to criticality. The vulnerability v_k of Eq. (4) is a constant in this case, so $\nu = 0$ in Eq. (14). As a result, Eq. (17) gives $\gamma = -1 + \alpha$ as the exponent of the offspring distribution. We therefore expect nonsymmetric average avalanche profiles when the power-law exponent α of the network degree distribution lies between 3 and 4; note we use $\alpha = 3.3$ in the left column of Fig. 7. It is noteworthy that the network degree distribution has finite variance in this case: it is the interaction between the network topology and the Centola-Macy dynamics that leads to nonsymmetric avalanche shapes. (In contrast, for the Watts threshold model with uniformly distributed thresholds, Eq. (3) gives $\nu = 1$, hence $\gamma = \alpha$, and nonsymmetric profiles appear only for networks with degree exponents α in the range $2 < \alpha < 3$, i.e., for degree distributions with infinite variance). The updating is synchronous, i.e., all nodes are updated at each discrete time step, and the number of avalanche events at each time step is the number of nodes that are newly activated.

In this case, we can clearly see one of the limitations of the branching process approximation: the seed node for each cascade is chosen uniformly at random from all the nodes, but subsequently activated nodes in the cascade are reached with a probability proportional to their degree, as in Eq. (1). Therefore the branching process picture does not correctly capture the first step of the cascade and this discrepancy can be seen in the early-time shape of the avalanches, particularly in panels (a) and (b) of Fig. 7. The theory could be extended to deal with this issue (as in [39] for example), but the effects on the average profiles diminishes as longer-duration avalanches are considered.

A more serious limitation of the theory's accuracy is presented by Fig. 7(e), where we see the deviation of the average number of events away from the constant value that indicates critical dynamics (see the discussion in Sec. 3.3). In fact, the finite size of the network and the heavy-tailed degree distribution mean that the activated nodes are quickly (within about 10 time steps) replacing inactive nodes throughout a significant fraction of the network, so that some of the "new" branches emanating from an activated node are in fact connected to previously-activated nodes, contrary to the branching process assumption of independence. As a result, the spreading efficiency decreases over time, and the branching process—despite initially being at criticality—becomes subcritical.

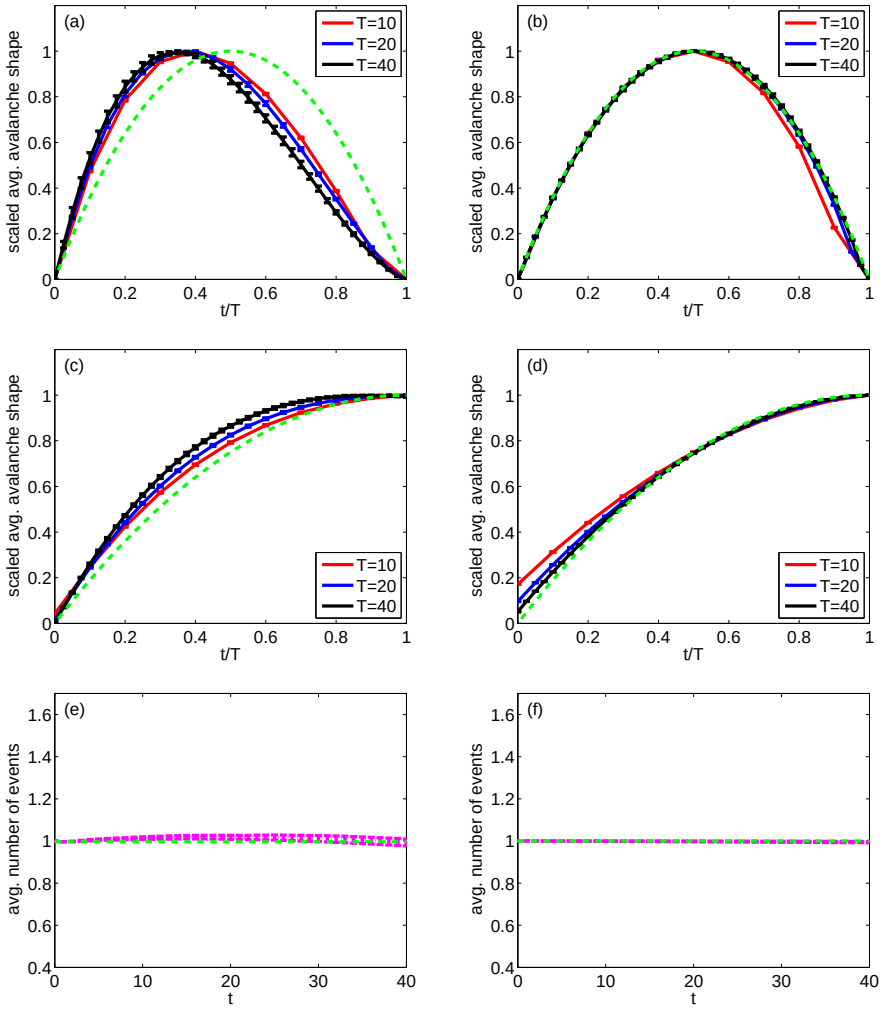

Figure 6: Results of numerical simulation of the neuronal dynamics model of [3] at criticality. As in Fig. 5, the results in the left column are for a network with power-law out-degree distribution ($\alpha = 2.5$), while those in the right column are for a network with z-regular out-degrees, see Table 1. Panels (a) and (b): rescaled average avalanche shapes. Panels (c) and (d): rescaled average non-terminating avalanche shapes. Panels (e) and (f): average number of firing neurons per discrete time step.

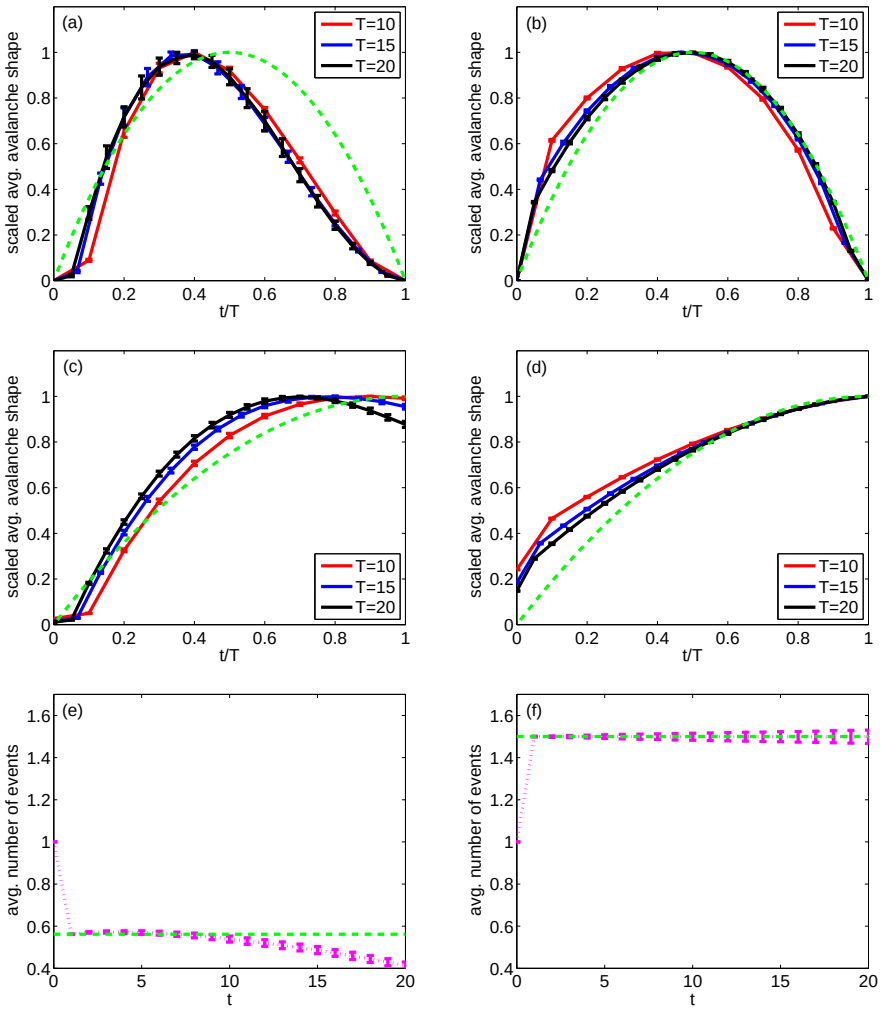

Figure 7: Results of numerical simulation of a Centola-Macy threshold model [29] at criticality. Left column results are for a scale-free network ($\alpha = 3.3$); right column results are for a z-regular random network, see Table 1. Panels (a) and (b): rescaled average avalanche shapes. Panels (c) and (d): rescaled average non-terminating avalanche shapes. Panels (e) and (f): average number of newly-activated nodes per discrete time step.

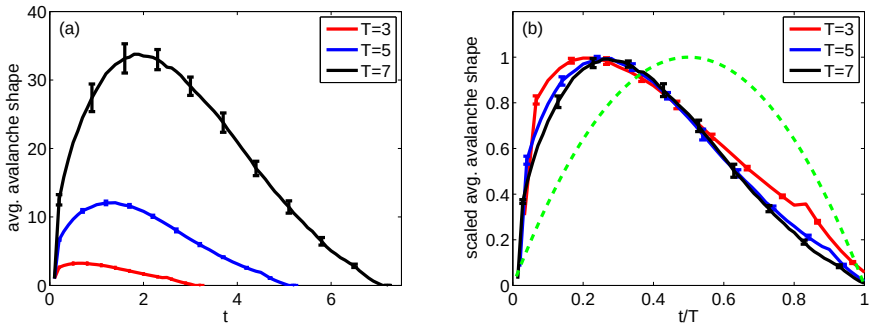

Figure 8: Numerical simulation results for critical meme-popularity avalanches [32] on the Twitter network of [40,41], see Table 1. Although this network is not tree-like, the qualitative predictions of our theory still hold: the average avalanche shapes (panel (a)) show a good collapse under rescaling (panel (b)), with a distinctly nonsymmetric shape.

Nevertheless, as Fig. 7(a) shows, the average avalanche shapes still collapse quite well, and clearly are nonsymmetric; the average shape of non-terminating avalanches (in panel (c)) is a more sensitive indicator of the loss of criticality due to the finite-size of the network.

4.4 Meme propagation on a directed real-world network

All networks used in the simulations above were configuration-model networks. However, real-world networks are known to differ significantly from configuration-model networks with the same degree distribution [26]. For example, degree-degree correlations and closed triangles of nodes ("clustering") are common in many social networks, but are relatively rare in the corresponding configuration-model networks. In our final example, we therefore use the sample of the Twitter network that is made available for download from [40] by the authors of [41]. Unlike a configuration-model directed network, this dataset contains a high proportion of reciprocated links (i.e., instances where node i follows node j and node j also follows node i); in fact 48% of all links in the network are reciprocated. This feature means that the network is not tree-like, and the branching process assumptions are not true. Nevertheless, Fig. 8 shows that results from numerical simulations of the critical meme propagation model of [32] are in very good qualitative agreement with the predictions of our theory: a good collapse of the average avalanche shapes for different durations T is found (Fig. 8(b)), with a clear left skew that is consistent with the heavy-tailed distribution of number of Twitter followers found in empirical data [34, 42]. As in other examples where tree-based theory is more accurate than expected on real-world networks [43], this result demonstrates that the predictions of the theory are quite robust to violations of the assumptions used for the mathematical derivation.

5 Conclusions

In this paper we have examined the link between Markov branching processes and unidirectional cascade dynamics on networks, with a focus on the temporal profile of avalanches. Our main result is Eq. (7), which gives the average avalanche shape for avalanches of duration T, and requires only the solution of a single ordinary differential equation (Eq. (9)). The input to this equation is the offspring distribution of the branching process, which is determined from the network structure and the dynamical system of interest by Eq. (1) or Eq. (2). In our analysis of the avalanche shape given

by Eq. (7), we have demonstrated that nonsymmetric avalanche shapes can arise at criticality when the offspring distribution has a power-law tail. Using numerical simulations of threshold models, neuronal dynamics, and online information-sharing, we show that nonsymmetric shapes can occur for common models running on networks with power-law degree distributions. However, it is important to note that a heavy-tailed degree distribution is not sufficient to guarantee nonsymmetric avalanche shapes: It is the interplay between the cascade dynamics and the network topology, as can be seen from the formulas in Eqs. (17) and (20) for the exponent of the offspring distribution, that determines the symmetry of the average avalanche shape. The results of numerical simulations verify the qualitative predictions of the branching process theory, despite finite-size effects and other violations of the assumptions of the theory.

In addition, our theoretical approach enables us to identify other characteristic temporal shape functions (e.g., the average non-terminating avalanche shape of Eq. (53)) that may prove useful when experimentalists seek to identify critical behaviour from the temporal signatures in a finite data set. Given the relevance of branching process descriptions to cascades in a range of fields (e.g., neuroscience [44], social networks [34], crackling noise [4], etc.), it is hoped that these insights may find many applications. We anticipate several possible directions for extensions of the methodology introduced here; notably, removing the Markovian assumption to apply a similar analysis for non-Markovian cascades [34, 39], and extending the theory to multilayer networks [45, 46].

Acknowledgements

Helpful discussions with Gareth Baxter, Sergey Dorogovtsev, Ali Faqeeh, Peter Fennell, Kristina Lerman, Kevin O'Sullivan, and Mason Porter are gratefully acknowledged. This work was supported by Science Foundation Ireland (grant numbers 15/SPP/E3125 and 11/PI/1026). We acknowledge the SFI/HEA Irish Centre for High-End Computing (ICHEC) for the provision of computational facilities.

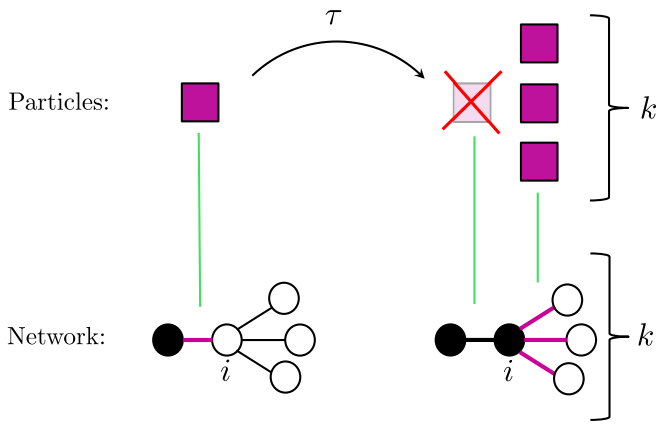

Figure S1: The green lines demonstrate the analogy between the particles of a continuous-time branching process (top), and the active-inactive deges (coloured magenta) of a cascade process on a network (bottom) as described in Appendix A.

A Linking branching processes to discrete-state dynamics on networks

In this Appendix we demonstrate how branching processes can be used to describe unidirectional dynamics on configuration-model networks. The paradigm of continuous-time branching processes describes *particles* who each live for a lifetime τ, where the value of τ is drawn from a prescribed distribution. In the Markov case, the distribution of τ values is exponential (and we choose the timescales so that the mean of the distribution is 1); in the Galton-Watson process time is discrete, so every particle has a lifetime of exactly 1 time unit. (The framework we describe here can also be used to link non-Markovian dynamics with the appropriate (Bellman-Harris) branching processes [47]; however, in this paper we restrict our attention to the Markovian continuous-time, and discrete-time, cases). At the end of its lifetime, the particle dies; simultaneously, a number k of identical "children" particles appear, where k is a random integer from a distribution q_k (the offspring distribution). Each avalanche of the branching process is initiated at time 0 by a single particle, and terminates at time T when there are no particles remaining alive; the temporal evolution of the cascade process (i.e., the avalanche shape) can be described in terms of the number of particles alive at time t, for times t with $0 < t < T$.

Figure S1 shows how unidirectional dynamics on a network can be described in terms of a branching process. We identify a particle of the branching process with an "active-inactive (AI) edge" of the network, i.e., an edge that links an active node to an inactive node (see the magenta edges in the network segments of Fig. S1) [48]. An edge labelled as an AI edge remains in this state for a time τ; then, with probability v_{k+1} (for a node of degree $k+1$ in an undirected network) the node i becomes active. Since the edge of interest now joins two active nodes, it is no longer an AI edge and is considered to be a dead particle. However, node i becoming active has created k new AI edges, one for each link from i to its inactive neighbours.

Thus, we identify the particles of a branching process with the AI edges of the network. The offspring distribution of the branching process is determined from the network structure and dynamics by considering that the probability of reaching a node of degree $k+1$ (such as node i) by

travelling along a random edge of the network is $\frac{k+1}{z}p_{k+1}$, and the probability that its single active neighbour will activate node i is given by v_{k+1}. Hence, the offspring distribution for an undirected network is given by Eq. (1). Similar arguments show that the corresponding offspring distribution for a directed network is given by Eq. (2).

B Derivation of average avalanche shape

In our theoretical approach, we model the avalanche dynamics using a continuous-time branching process [47]. Each "particle" survives for a random lifetime τ (drawn from the exponential distribution with mean 1) and at the end of its lifetime, is replaced by k "children", where the number of offspring k is $0, 1, 2, \ldots$ with probability q_k. The mean number ξ of children per parent is given by Eq. (6); the critical case $\xi = 1$ is of particular interest, but our derivation is also valid for other cases.

Following the notation of Chapter III of [47], we define $Z(t)$ as a one-dimensional continuous time Markov branching process. The (time-homogeneous) transition probabilities are defined as

$$P_{1j}(t) = \text{Prob}\left\{Z(\tau + t) = j | Z(\tau) = 1\right\} \tag{21}$$

and the corresponding generating function is

$$F(s,t) = \sum_k P_{1k}(t) s^k. \tag{22}$$

Note that the probability of extinction by time t is

$$Q(t) \equiv \text{Prob}\left\{Z(t) = 0 | Z(0) = 1\right\} = F(0,t). \tag{23}$$

We consider an avalanche as a process starting with a single particle ($Z(0) = 1$), with lifetimes of particles that are exponentially distributed with parameter 1. Upon death, a particle leaves k offspring with probability q_k, $k = 0, 1, 2, \ldots$; the generating function for the offspring distribution is defined in Eq. (8), with $\xi = f'(1)$.

For a fixed duration $T > 0$, we define the *avalanche path probability* $\pi_n(t)$ to be the probability that n particles are alive at time t ($0 \le t \le T$), conditioned on there being exactly one particle alive at time 0 and exactly one particle alive at time T:

$$\pi_n(t) = \text{Prob}\left\{Z(t) = n | Z(0) = 1 \text{ and } Z(T) = 1\right\}. \tag{24}$$

Because of the Markov property, $\pi_n(t)$ is proportional to the product of the probabilities of having a path that (i) has n particles alive at time t and (ii) goes from n particles at time t to one particle at time T, so

$$\pi_n(t) \propto P_{1n}(t) P_{n1}(T - t). \tag{25}$$

We can write

$$P_{n1}(T - t) = n \left[P_{10}(T - t)\right]^{n-1} P_{11}(T - t) \tag{26}$$

to reflect the fact that the n particles active at time t can be reduced to a single particle at time T only if $n - 1$ of them have no descendants alive at time T, with the remaining one (which can be chosen in n ways) having one descendant at time T. The correctly normalized avalanche path probability distribution is therefore

$$\pi_n(t) = \frac{P_{1n}(t)\, n \left[P_{10}(T - t)\right]^{n-1} P_{11}(T - t)}{\sum_n P_{1n}(t)\, n \left[P_{10}(T - t)\right]^{n-1} P_{11}(T - t)}. \tag{27}$$

The denominator of (27) can be written as

$$F'\left(P_{10}(T-t),t\right)P_{11}(T-t), \tag{28}$$

where, for brevity, the prime denotes differentiation with respect to the generating function variable s (i.e., the first argument of $F(s,t)$). Noting that $P_{10}(t) = F(0,t) = Q(t)$ and $P_{11}(t) = F'(0,t)$, the denominator of (27) can be further simplified to

$$F'\left(F(0,T-t),t\right)F'(0,T-t) = F'(0,T), \tag{29}$$

where we have used the Markov property [47]

$$F(s,T) = F\left(F(s,T-t),t\right). \tag{30}$$

Thus, the avalanche path probability distribution is

$$\pi_n(t) = \frac{P_{1n}(t)\,n\left[P_{10}(T-t)\right]^{n-1}P_{11}(T-t)}{F'(0,T)}, \tag{31}$$

and the generating function for this distribution is

$$\begin{aligned} G(s,t) &= \sum_n \pi_n(t)s^n \\ &= \frac{1}{F'(0,T)}\sum_n P_{1n}(t)\,n\left[P_{10}(T-t)\right]^{n-1}P_{11}(T-t)s^n \\ &= \frac{sF'\left(sP_{10}(T-t),t\right)P_{11}(T-t)}{F'(0,T)}. \end{aligned} \tag{32}$$

The average avalanche shape is determined by the expected number of particles alive at time t, where the expectation is over the set of avalanche paths conditioned as in Eq. (24). We calculate this using the generating function $G(s,t)$ as follows:

$$\begin{aligned} G'(1,t) &= \frac{1}{F'(0,T)}\left[F'\left(P_{10}(T-t),t\right)P_{11}(T-t) + F''\left(P_{10}(T-t),t\right)P_{10}(T-t)P_{11}(T-t)\right] \\ &= 1 + \frac{F(0,T-t)F'(0,T-t)F''\left(F(0,T-t),t\right)}{F'(0,T)}. \end{aligned} \tag{33}$$

Since we are considering Markov branching processes, we can use the Kolmogorov forward and backward equations [47] for the generating function $F(s,t)$:

$$\begin{aligned} \frac{\partial}{\partial t}F(s,t) &= \left(f(s)-s\right)F'(s,t) \\ \frac{\partial}{\partial t}F(s,t) &= f\left(F(s,t)\right) - F(s,t), \end{aligned} \tag{34}$$

and by eliminating $\partial F/\partial t$ from this pair of equations we write $F'(s,t)$ in terms of $F(s,t)$:

$$F'(s,t) = \frac{f\left(F(s,t)\right) - F(s,t)}{f(s)-s} \quad \text{for } s \neq 1. \tag{35}$$

Differentiating with respect to s and substituting F' using (35) yields

$$F''(s,t) = \frac{\left[f\left(F(s,t)\right) - F(s,t)\right]\left[f'\left(F(s,t)\right) - f'(s)\right]}{\left[f(s)-s\right]^2}, \tag{36}$$

and evaluating at $s = F(0, T-t)$ and using (30), we have

$$F''\left(F(0,T-t),t\right) = \frac{\left[f\left(F(0,T)\right) - F(0,T)\right]\left[f'\left(F(0,T)\right) - f'\left(F(0,T-t)\right)\right]}{\left[f\left(F(0,T-t)\right) - F(0,T-t)\right]^2}. \tag{37}$$

Using (35) and (37) in (33) gives

$$\begin{aligned} G'(1,t) &= 1 + \frac{F(0,T-t)\left[f'\left(F(0,T)\right) - f'\left(F(0,T-t)\right)\right]}{f\left(F(0,T-t)\right) - F(0,T-t)} \\ &= 1 + \frac{Q(T-t)\left[f'\left(Q(T)\right) - f'\left(Q(T-t)\right)\right]}{f\left(Q(T-t)\right) - Q(T-t)}, \end{aligned} \tag{38}$$

where $Q(t) = F(0,t)$ solves the ordinary differential equation obtained from the second equation of (34) at $s = 0$:

$$\frac{dQ}{dt} = f(Q) - Q, \qquad \text{with } Q(0) = 0. \tag{39}$$

We have therefore reduced the problem of calculating the average avalanche shape to the solution of a single ordinary differential equation. Given an offspring distribution and a duration T, we can easily solve Eq. (39) using standard numerical methods for $t = 0$ to T, then substitute $Q(t)$ into Eq. (38) to obtain the average avalanche shape

$$A(t) \equiv G'(1,t) - 1, \tag{40}$$

as given in Eq. (7).

To calculate the variance of the avalanche shape, we again use the generating function $G(s,t)$ to write

$$\mathrm{var}\left(n(t)\right) = \left\langle n^2 \right\rangle - \langle n \rangle^2 = G''\left(1,t\right) + G'\left(1,t\right) - \left(G'(1,t)\right)^2. \tag{41}$$

Differentiating Eq. (32) with respect to s, then setting s equal to 1 (and using the identities $P_{10}(T-t) \equiv F(0,T-t)$ and $P_{11}(T-t) = F'(0,T-t)$) yields

$$G''(1,t) = \frac{F'(0,T-t)F(0,T-t)}{F'(0,T)}\left\{2F''\left(F(0,T-t),t\right) + F'''\left(F(0,T-t),t\right)F(0,T-t)\right\}. \tag{42}$$

Using Eq. (35) on the first term, and Eq. (36) and differentiation of Eq. (35) on the terms in braces leads (after some algebra) to the result

$$\begin{aligned} \mathrm{var}\left(n(t)\right) = &\frac{Q(T-t)}{\left[f(Q(T-t) - Q(T-t)\right]^2} \times \\ &\times \left\{ f(Q(T-t))\left[f'(Q(T) - f'(Q(T-t) - Q(T-t)f''(Q(T-t)\right]\right. \\ &\quad + Q(T-t)\left[f'(Q(T-t))^2 - f'(Q(T))f'(Q(T-t))\right. \\ &\quad \left.\left. + f''(Q(T))\left[f(Q(T)) - Q(T)\right] + Q(T-t)f''\left(Q(T-t)\right)\right]\right\}. \end{aligned} \tag{43}$$

C Exactly solvable critical case

We consider binary fission, with $p_0 = p_2 = 1/2$ being the only non-zero offspring probabilities. Hence the offspring generating function is

$$f(s) = \frac{1}{2} + \frac{1}{2}s^2, \tag{44}$$

and the ordinary differential equation (39) can be solved exactly for $Q(t)$:

$$Q(t) = \frac{t}{2+t}. \tag{45}$$

Using these functions in Eq. (38) gives the exact avalanche shape

$$A(t) = G'(1,t) - 1 = \frac{t}{2+T}(T-t) = \frac{T^2}{2+T}\frac{t}{T}\left(1 - \frac{t}{T}\right), \tag{46}$$

which is a parabolic function of t with maximum at $t = T/2$. The corresponding variance of the avalanche shape is found from Eq. (43) to be

$$\frac{T^4}{2(2+T)^2}\left(\frac{t}{T}\right)^2\left(1 - \frac{t}{T}\right)^2, \tag{47}$$

and so the coefficient of variation for this case is $\mathrm{CV} = 1/\sqrt{2}$, independent of t and T.

D Derivation of non-terminating avalanche shape

We consider here the average shape of all avalanches that have not terminated by the observation time T. We use the notation introduced in Appendix B, but here we define (for a fixed observation time $T > 0$) the avalanche path probability $\tilde{\pi}_n(t)$ to be the probability that n particles are alive at time t ($0 \le t \le T$), conditioned on there being exactly one particle alive at time 0 and *at least one* particle alive at time T:

$$\tilde{\pi}_n(t) = \mathrm{Prob}\left\{Z(t) = n | Z(0) = 1 \text{ and } Z(T) > 0\right\}. \tag{48}$$

Using the Markov property, we have

$$\tilde{\pi}_n(t) \propto P_{1n}(t)\left[1 - P_{n0}(T-t)\right] = P_{1n}(t)\left[1 - P_{10}(T-t)^n\right]. \tag{49}$$

The correctly normalized avalanche path probability distribution is therefore

$$\begin{aligned}\tilde{\pi}_n(t) &= \frac{P_{1n}(t)\ \left[1 - P_{10}(T-t)^n\right]}{\sum_n P_{1n}(t)\ \left[1 - P_{10}(T-t)^n\right]} \\ &= \frac{P_{1n}(t)\ \left[1 - P_{10}(T-t)^n\right]}{1 - Q(T)},\end{aligned} \tag{50}$$

where we have used the Markov property Eq. (30) in the denominator to write

$$\sum_m P_{1m}(t) P_{10}(T-t)^m = F\left(P_{10}(T-t), t\right) = F\left(F(0, T-t), t\right) = F(0,T) = Q(T). \tag{51}$$

The average non-terminating avalanche shape is then given by

$$\begin{aligned}A_{NT}(t) &= \sum_n n\tilde{\pi}_n(t) \\ &= \sum_n \frac{n\, P_{1n}(t)\ \left[1 - Q(T-t)^n\right]}{1 - Q(T)} \\ &= \frac{F'(1,t) - Q(T-t)F'\left(Q(T-t), t\right)}{1 - Q(T)}.\end{aligned} \tag{52}$$

Using the relation given by Eq. (35) for the second term in the numerator we get

$$A_{NT}(t) = \frac{F'(1,t) - Q(T-t)\frac{f(Q(T))-Q(T)}{f(Q(T-t))-Q(T-t)}}{1-Q(T)}, \tag{53}$$

so it remains only to find an expression for $F'(1,t)$.

Recall that $F'(1,t) = \sum_n nP_{1n}(t)$, which is the expected number of particles alive at time t. This quantity therefore measures the average number of events observed at time t across an ensemble comprised of *all* avalanches (with no conditioning on avalanche duration); as noted in Section 3.3, it provides a simple measure of whether an observed ensemble of avalanches is from a subcritical, critical, or supercritical system. Differentiating the backward Kolmogorov equation (Eq. (34)) with respect to s and setting s equal to 1 gives the ordinary differential equation satisfied by $F'(1,t)$:

$$\frac{d}{dt}F'(1,t) = \left[f'(1) - 1\right] F'(1,t), \tag{54}$$

which has solution

$$F'(1,t) = \exp\left[(\xi - 1)t\right] \tag{55}$$

(recall $\xi = f'(1)$ is the branching number). Therefore, critical systems (those with $\xi = 1$) have a constant expected number of events per unit time, while in sub- (respectively, super-) critical systems, the number of events decays (resp., grows) exponentially with the age t of the avalanches.

Substituting the expression (55) into Eq. (53) gives a formula for the average non-terminating avalanche shape in terms of the extinction fraction $Q(t)$; the asymptotic analysis of this shape is given in Appendix F.

E Discrete-time branching processes

Here we consider discrete-time (Galton-Watson) branching processes instead of the continuous-time process used in Appendices B–D. The arguments leading to Eq. (33) remain valid in the discrete-time case, but we cannot now use the Kolmogorov equations to further simplify Eq. (33). However, the Galton-Watson process has the governing equation [47]

$$F(s,t+1) = f\left(F(s,t)\right), \tag{56}$$

from which we can derive the relations

$$F'(s,t+1) = f'\left(F(s,t)\right) F'(s,t) \tag{57}$$

and

$$F''(s,t+1) = f'\left(F(s,t)\right) F''(s,t) + f''\left(F(s,t)\right) \left[F'(s,t)\right]^2, \tag{58}$$

with initial conditions

$$F(s,0) = s, \quad F'(s,0) = 1, \quad F''(s,0) = 0. \tag{59}$$

For any chosen values of the variable s and duration T, we can iterate Eqs. (56)–(58) to find the values of $F(s,t)$, $F'(s,t)$ and $F''(s,t)$ at any (integer) time value t between 0 and T; evaluating at the appropriate value of s then enables us to calculate the average avalanche shape $A(t) = G'(1,t) - 1$ from Eq. (33).

Figure S2 demonstrates that the discrete-time and continuous-time (from Eq. (7)) formulations give very similar results for the average avalanche shapes; accordingly, we concentrate in the main text on the continuous-time approach, for which analytical insights (e.g., asymptotic behaviour) are easier to obtain.

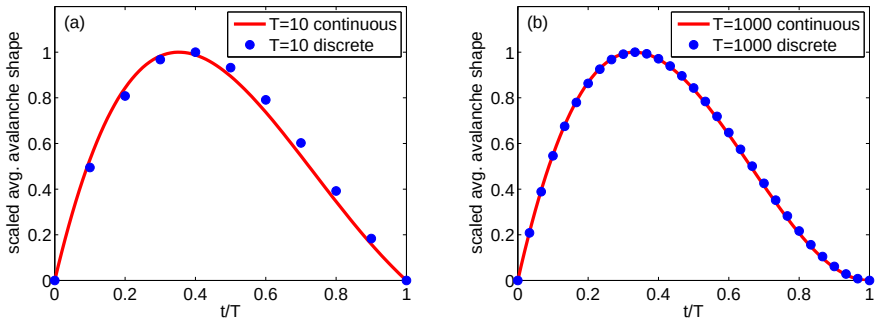

Figure S2: Average avalanche shapes at criticality ($\xi = 1$) from the continuous-time approach of Eq. (7) (curves) and the discrete-time approach of Eqs. (56)–(58) (symbols). The offspring distribution q_k is power-law, with exponent $\gamma = 2.5$. The results of both approaches are very similar even for relatively short-duration avalanches ($T = 10$ in panel (a)), and are almost indistinguishable for long-lived avalanches ($T = 1000$ in panel (b)).

F Large-T asymptotics at criticality

To consider the asymptotic behaviour of the avalanche shape for a general offspring distribution, we begin with the large-time asymptotics of $Q(t)$. Equation (39) can be rewritten as

$$\int \frac{dQ}{f(Q) - Q} = \int dt, \tag{60}$$

and the large-t asymptotic behaviour corresponds to Q values near 1. Expanding $f(Q)$ near $Q = 1$, and assuming that $q_k \sim Dk^{-\gamma}$ as $k \to \infty$ with $2 < \gamma < 3$ yields [49]

$$f(Q) \sim 1 - f'(1)(1 - Q) + D\Gamma(1 - \gamma)(1 - Q)^{\gamma - 1} \quad \text{as } Q \to 1^-. \tag{61}$$

If the offspring distribution has a finite second moment then Eq. (61) is replaced by

$$f(Q) \sim 1 - f'(1)(1 - Q) + \frac{f''(1)}{2}(1 - Q)^2 \quad \text{as } Q \to 1^-. \tag{62}$$

In the critical case $f'(1) = 1$ and substituting from Eq. (61) into Eq. (60) gives

$$\int \frac{dQ}{D\Gamma(1 - \gamma)(1 - Q)^{\gamma - 1}} = \int dt$$

which integrates to

$$1 - Q \sim C_1 t^{-\frac{1}{\gamma - 2}} \quad \text{as } t \to \infty, \tag{63}$$

where $C_1 = [D(\gamma - 2)\Gamma(1 - \gamma)]^{-\frac{1}{\gamma - 2}}$. (If $f''(1)$ is finite, then (63) still holds with γ replaced by 3 and C_1 replaced by $2/f''(1)$.)

Differentiating (61) (or (62)) and substituting for $Q(t)$ from (63) leads to

$$f'(Q(t)) \sim 1 - C_2\, t^{-1} \quad \text{as } t \to \infty \tag{64}$$

and hence (38) yields the asymptotic avalanche shape, in the limit $T \to \infty$ and $T - t \to \infty$, as

$$A(t) \sim C_3\, T^{\frac{1}{\gamma - 2}} \frac{t}{T} \left(1 - \frac{t}{T}\right)^{\frac{1}{\gamma - 2}}, \tag{65}$$

where C_2 and $C_3 = (\gamma - 1)/C_1$ are independent of t and T, and this formula holds also for finite $f''(1)$ if γ is replaced by 3. We note that the asymptotic avalanche shape is parabolic for $\gamma = 3$ (i.e., if $f''(1)$ is finite), but its peak is at $t/T = (\gamma - 2)/(\gamma - 1) < 1/2$ for γ values between 2 and 3.

Similarly, the asymptotic behaviour of the average non-terminating avalanche shape at criticality is determined from Eq. (53) to be (in the limit $T \to \infty$ and $T - t \to \infty$):

$$A_{NT}(t) \sim \begin{cases} C_4\, t\left(2 - \frac{t}{T}\right) & \text{if } f''(1) \text{ is finite,} \\ \\ C_4\, T^{\frac{1}{\gamma-2}} \left(1 - \left(1 - \frac{t}{T}\right)^{\frac{\gamma-1}{\gamma-2}}\right) & \text{if } q_k \sim k^{-\gamma} \text{ as } k \to \infty, \text{ with } 2 < \gamma < 3, \end{cases} \tag{66}$$

where C_4 is independent of t and T.

The asymptotic behaviour of the variance of the avalanche shape is found by analyzing Eq. (43) in the same way, yielding

$$\mathrm{var}\,(n(t)) \sim T^{\frac{2}{\gamma-2}} \frac{\gamma - 1}{C_1^2} \frac{t}{T} \left(1 - \frac{t}{T}\right)^{\frac{2}{\gamma-2}} \left(3 - \gamma + \frac{t}{T}(\gamma - 2)\right) \tag{67}$$

(with $\gamma = 3$ again recovering the finite-$f''(1)$ case). Using this formula and the asymptotics of the mean avalanche shape from Eq. (65) gives the result for the coefficient of variation of the avalanche shape in Eq. (13) of the main text.

References

[1] Oscar A Pinto and Miguel A Muñoz. Quasi-neutral theory of epidemic outbreaks. *PloS One*, 6(7):e21946, 2011.

[2] Javier Borge-Holthoefer, Raquel A Baños, Sandra González-Bailón, and Yamir Moreno. Cascading behaviour in complex socio-technical networks. *Journal of Complex Networks*, 1(1):3–24, 2013.

[3] Nir Friedman, Shinya Ito, Braden AW Brinkman, Masanori Shimono, RE Lee DeVille, Karin A Dahmen, John M Beggs, and Thomas C Butler. Universal critical dynamics in high resolution neuronal avalanche data. *Physical Review Letters*, 108(20):208102, 2012.

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

[5] Michael PH Stumpf and Mason A Porter. Critical truths about power laws. *Science*, 335(6069):665–666, 2012.

[6] Mark EJ Newman. Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 46(5):323–351, 2005.

[7] Stefanos Papanikolaou, Felipe Bohn, Rubem Luis Sommer, Gianfranco Durin, Stefano Zapperi, and James P Sethna. Universality beyond power laws and the average avalanche shape. *Nature Physics*, 7(4):316–320, 2011.

[8] John M Beggs and Nicholas Timme. Being critical of criticality in the brain. *Frontiers in Physiology*, 3:163, 2012.

[9] James A Roberts, Kartik K Iyer, Simon Finnigan, Sampsa Vanhatalo, and Michael Breakspear. Scale-free bursting in human cortex following hypoxia at birth. *The Journal of Neuroscience*, 34(19):6557–6572, 2014.

[10] Amit P Mehta, Andrea C Mills, Karin A Dahmen, and James P Sethna. Universal pulse shape scaling function and exponents: Critical test for avalanche models applied to Barkhausen noise. *Physical Review E*, 65(4):046139, 2002.

[11] Matthew C Kuntz and James P Sethna. Noise in disordered systems: The power spectrum and dynamic exponents in avalanche models. *Physical Review B*, 62(17):11699, 2000.

[12] Stefano Zapperi, Claudio Castellano, Francesca Colaiori, and Gianfranco Durin. Signature of effective mass in crackling-noise asymmetry. *Nature Physics*, 1(1):46–49, 2005.

[13] Francesca Colaiori. Exactly solvable model of avalanches dynamics for Barkhausen crackling noise. *Advances in Physics*, 57(4):287–359, 2008.

[14] Andrea Baldassarri, Francesca Colaiori, and Claudio Castellano. Average shape of a fluctuation: Universality in excursions of stochastic processes. *Physical Review Letters*, 90(6):060601, 2003.

[15] Lasse Laurson, Mikko J Alava, and Stefano Zapperi. Power spectra of self-organized critical sandpiles. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(11):L11001, 2005.

[16] Matthias Rybarsch and Stefan Bornholdt. Avalanches in self-organized critical neural networks: A minimal model for the neural SOC universality class. *PloS One*, 9(4):e93090, 2014.

[17] Paolo Massobrio, Valentina Pasquale, and Sergio Martinoia. Self-organized criticality in cortical assemblies occurs in concurrent scale-free and small-world networks. *Scientific Reports*, 5:10578, 2015.

[18] Janina Hesse and Thilo Gross. Self-organized criticality as a fundamental property of neural systems. *Frontiers in Systems Neuroscience*, 8:166, 2014.

[19] Daniel B Larremore, Marshall Y Carpenter, Edward Ott, and Juan G Restrepo. Statistical properties of avalanches in networks. *Physical Review E*, 85(6):066131, 2012.

[20] Ian Dobson. Estimating the propagation and extent of cascading line outages from utility data with a branching process. *IEEE Transactions on Power Systems*, 27(4):2146–2155, 2012.

[21] Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani. Epidemic processes in complex networks. *Reviews of Modern Physics*, 87(3):925, 2015.

[22] Justin Cheng, Lada Adamic, P Alex Dow, Jon Michael Kleinberg, and Jure Leskovec. Can cascades be predicted? In *Proceedings of the 23rd International Conference on World Wide Web*, pages 925–936. ACM, 2014.

[23] Sharad Goel, Ashton Anderson, Jake Hofman, and Duncan J Watts. The structural virality of online diffusion. *Management Science*, 62(1):180–196, 2015.

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

[25] Filippo Radicchi and Claudio Castellano. Breaking of the site-bond percolation universality in networks. *Nature Communications*, 6, 2015.

[26] Mark Newman. *Networks: an introduction*. Oxford University Press, 2010.

[27] Mason A Porter and James P Gleeson. *Dynamical systems on networks: A tutorial*. Springer, 2016.

[28] Duncan J Watts. A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences of the United States of America*, 99(9):5766–5771, 2002.

[29] Damon Centola and Michael Macy. Complex contagions and the weakness of long ties. *American Journal of Sociology*, 113(3):702–734, 2007.

[30] RE Lee DeVille and Charles S Peskin. Synchrony and asynchrony in a fully stochastic neural network. *Bulletin of Mathematical Biology*, 70(6):1608–1633, 2008.

[31] RE Lee DeVille, Charles S Peskin, and Joel H Spencer. Dynamics of stochastic neuronal networks and the connections to random graph theory. *Mathematical Modelling of Natural Phenomena*, 5(2):26–66, 2010.

[32] James P Gleeson, Jonathan A Ward, Kevin P O'Sullivan, and William T Lee. Competition-induced criticality in a model of meme popularity. *Physical Review Letters*, 112(4):048701, 2014.

[33] Lillian Weng, Alessandro Flammini, Alessandro Vespignani, and Fillipo Menczer. Competition among memes in a world with limited attention. *Scientific Reports*, 2:335, 2012.

[34] James P Gleeson, Kevin P O'Sullivan, Raquel A Baños, and Yamir Moreno. Effects of network structure, competition and memory time on social spreading phenomena. *Physical Review X*, 6(2):021019, 2016.

[35] Richard Durrett. *Random graph dynamics*. Cambridge University Press, 2007.

[36] Kristina Lerman, Rumi Ghosh, and Tawan Surachawala. Social contagion: An empirical study of information spread on Digg and Twitter follower graphs. *arXiv:1202.3162*, 2012.

[37] Kristina Lerman. Information is not a virus, and other consequences of human cognitive limits. *Future Internet*, 8(2):21, 2016.

[38] Matlab/Octave simulation codes and network datasets for the avalanche examples used in this paper. `http://www.ul.ie/gleesonj/avalanches`.

[39] José Luis Iribarren and Esteban Moro. Branching dynamics of viral information spreading. *Physical Review E*, 84(4):046116, 2011.

[40] SNAP Network datasts. `http://snap.stanford.edu/data/egonets-Twitter.html`.

[41] Julian J McAuley and Jure Leskovec. Learning to discover social circles in ego networks. In *NIPS*, volume 2012, pages 548–56, 2012.

[42] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. What is Twitter, a social network or a news media? In *Proceedings of the 19th International Conference on World Wide Web*, pages 591–600. ACM, 2010.

[43] Sergey Melnik, Adam Hackett, Mason A Porter, Peter J Mucha, and James P Gleeson. The unreasonable effectiveness of tree-based theory for networks with clustering. *Physical Review E*, 83(3):036112, 2011.

[44] Woodrow L Shew, Hongdian Yang, Shan Yu, Rajarshi Roy, and Dietmar Plenz. Information capacity and transmission are maximized in balanced cortical networks with neuronal avalanches. *The Journal of Neuroscience*, 31(1):55–63, 2011.

[45] Mikko Kivelä, Alex Arenas, Marc Barthelemy, James P Gleeson, Yamir Moreno, and Mason A Porter. Multilayer networks. *Journal of Complex Networks*, 2(3):203–271, 2014.

[46] Stefano Boccaletti, Ginestra Bianconi, Regino Criado, Charo I Del Genio, Jesús Gómez-Gardeñes, Miguel Romance, Irene Sendiña-Nadal, Zhen Wang, and Massimiliano Zanin. The structure and dynamics of multilayer networks. *Physics Reports*, 544(1):1–122, 2014.

[47] Krishna B Athreya and Peter E Ney. *Branching Processes*. Springer Science & Business Media, 2012.

[48] John K McSweeney. Single-seed cascades on clustered networks. *arXiv:1510.00056*, 2015.

[49] Herbert S Wilf. *generatingfunctionology*. Elsevier, 2013.