Diffusion approximation of a network model of meme popularity

Kleber A. Oliveira†, Samuel Unicomb, and James P. Gleeson
MACSI, Department of Mathematics and Statistics, University of Limerick, Limerick, Ireland

Models of meme propagation on social networks, in which memes compete for limited user attention, can successfully reproduce the heavy-tailed popularity distributions observed in online settings. While system-wide popularity distributions have been derived analytically, the dynamics of individual meme trajectories have thus far evaded description. To address this, we formulate the diffusion of a given meme as a one-dimensional stochastic process, whose fluctuations result from aggregating local network dynamics using classic and generalised central limit theorems, with the latter based on stable distribution theory. Ultimately, our approach decouples competing trajectories of meme popularities, allowing them to be simulated independently, and thus parallelised, and expressed in terms of Fokker-Planck equations.

I. INTRODUCTION

The dynamics of information diffusion in networked systems have been investigated extensively in recent years [1]. Broadly interpreted as transmissible pieces of information [2], memes have been collected and analysed in a number of online empirical studies [3–6]. Of particular interest is meme popularity, which corresponds to the number of times that a given piece of information content has been shared by a social media user to their network neighbours, or followers. Empirically, distributions of meme popularity are characterised by their high variance, and indeed heavy tails. That is, while most memes die out soon after being introduced in favour of competing content, a small number are heavily shared, both through viral cascades and the broadcasting effect of nodes with high numbers of followers [7].

The distributions of meme popularity observed in empirical settings have been recovered in simple models, through a mechanism known as competition-induced criticality, or CIC [8]. In these models, memes are viewed as being in competition for the finite resource that is user attention. The constraint of finite attention has become key in many models of social network dynamics [9–12]. In models of CIC, this constraint is typically represented by a screen attributed to each user, on which a single meme can be held; extensions to this model can allow a list of memes [13]. Memes are transmitted when users broadcast to their followers the contents of their own screen, overwriting any content that was previously on the screens of the followers. Typical of critical phenomena on complex networks [14], dynamic quantities here can be described in terms of branching processes.

This approach to describe the phenomenon is also known as a tree-like approximation, in the sense it assumes cascade growth outcomes typically neglect the presence of cycles in the network structure – a useful assumption to a very reasonable degree [15]. Additionally, the CIC models can be understood as neutral models [16, 17], meaning no meme has any inherent characteristic to help them grow more popular than others; in other words, the popularity dynamics are solely controlled by macroparameters such as the underlying network structural characteristics and the rate new memes are introduced in the system, while trajectories of individual memes are highly stochastic.

Although branching process techniques have been successful in deriving system-wide distributions of meme popularity, they are poorly suited to describing the time evolution of the popularity of any single meme. In simulations of the CIC dynamics, the popularity of a given meme evolves according to a discrete-time jump process, where the number of screens that it occupies increases and decreases due to competition with other memes for user attention. We examine an approximation that considers the fraction of screens occupied by memes to be a continuous variable.

Viewing meme popularity as a continuous, stochastic quantity suggests connection to related models. For example, Di Santo et al. [18] encode the statistics of several avalanche types through the formalism of stochastic equations. Notarmuzi et al. [19] also produce a diffusion process description for the quality-biased competition model (which is inspired by the CIC model) that provides novel analytical understanding of the mechanisms prescribed by agent-based simulations. In this paper, we carefully examine the discrete-to-continuous limit to produce novel descriptions of meme dynamics making use of the classic or generalised central limit theorems according to certain cases, and examine the accuracy and parameter dependence of the resulting approximations.

The remainder of this paper is organised as follows. In Section II, we begin by describing the competition-induced criticality model, as well as the simulation scheme used in numerical experiments. Then in Section II A, we identify distinct cases of local fluctuations that can be aggregated by means of central limit theorems. In Sections II B and II C, we distinguish between the case of finite- and infinite-variance degree distributions, making use of classic and generalised central limit theorems to derive stochastic, finite-difference equations modelling popularity dynamics. In Section III, we validate our analysis, by systematically comparing the results of our one-dimensional model to those obtained through computational simulations of the CIC model on various networks.
II. THE DIFFUSION APPROXIMATION

In this section we describe a diffusion approximation of the meme popularity model of Ref. [8]. Following that work, we assume a directed network of size \( N \), with edge direction representing the direction of the flow of information in a social network. As such, the out-degree \( k \) of a node is the number of followers, for instance, in an online social network like Twitter. We assume this to follow a distribution \( p_k \). The average degree is \( \langle k \rangle \), and its variance is \( \sigma^2 \). For simplicity, we assume that \( \langle k \rangle \) is integer and that the in-degree distribution is \( \langle k \rangle \)-regular, and wired according to the configuration model, meaning the network is maximally random up to in- and out-degree. Each node represents a screen on which a single meme can be displayed. Each meme has a discrete type, or colour, and is copied from the screen of a given node to its followers when the node “tweets.”

In each discrete time step, the system is updated by having one node is selected uniformly at random to tweet. With probability \( \mu \), the active node innovates and tweets an original meme, overwriting the previously held meme, as well as those of its out-neighbours. Otherwise, i.e., with probability \( 1 - \mu \), the node “retweets” the meme currently on its screen and so overwrites the screens of its out-neighbours with the retweeted meme. We define the elementary time unit as \( dt = 1/N \) between successive updates, such that a macroscopic time step of 1 is when \( N \) updates have occurred across the system. We initialise the system by assuming every screen to be blank, and assume that if a blank screen is chosen for update, it innovates a new meme with probability 1. This model was studied in Ref. [8] using a branching process approximation for the distribution of meme popularity (accumulated number of retweets). In this paper, we focus on characterising the stochastic trajectories of individual memes, which necessitates a discrete-to-continuous approximation to study the fraction of screens occupied by a meme at a given time.

A. Fluctuation cases

We focus on a particular meme, which we shall refer to as the red meme, that occupies a fraction \( s \) of all screens in the system. The evolution of this meme is a stochastic function of the network topology \( p_k \), the innovation rate \( \mu \), as well as the current density \( s \) of the meme itself. The temporal evolution of \( s \) can be modelled by identifying four cases, or outcomes, of an update step. These are indexed by \( j = 1, 2, 3 \) and 4, as outlined in Table I. In cases 1 and 2, the selected node innovates original content, and in cases 3 and 4 it simply retweets. In cases 1 and 3, the selected node contains the red meme, and in cases 2 and 4, some other meme. Note that cases 1, 2 and 4 can only diminish the number of red memes on screens across the system, by overwriting with some other meme. In contrast, case 3 can only increase the number of red memes, by overwriting the screens of its followers with the red meme. The sign of the contribution and the probability \( \pi_j \) of case \( j \) arising are given in columns 4 and 5 of Table I. Further, upon the assumption that red memes are distributed with density \( s \) randomly amongst all nodes, we can model the fluctuation in \( s \) at a given step, in each case \( j \). The probability generating function, or pgf, corresponding to the out-degree distribution \( p_k \) is defined to be \( f(x) = \sum_k p_k x^k \). Accordingly, the pgf of the number of neighbours that do or do not contain meme \( s \) is given by composition of \( f \) with the pgf of the Bernoulli distribution: \( 1 - s + s x \) or \( s + (1 - s) x \), respectively. By shifting and scaling by the probabilities \( \pi_j \), we define in the final column of Table I the pgfs \( g_j(x) \) for the integer change in the absolute number of nodes containing the red meme due to each case \( j \). The goal of a diffusion approximation is to aggregate a large number of fluctuations resulting from microscopic dynamics to estimate an effective change in the variable of interest over a larger time interval. This is achieved using central limit theorems, which deal with sums of random variables. In this way, we approximate the temporal evolution of \( s \), using our knowledge of the fluctuations in \( s \) due to individual updates, as described by the distributions \( g_j(x) \). We group together \( n \) successive update events, and estimate the resulting change to the red meme, \( \Delta s \). The time step in the diffusion approximation is \( \Delta t = n dt = n/N \). In choosing a value for \( n \), we weigh two considerations. The first is that \( n \) must be sufficiently large for the sum of fluctuations to be well-approximated using the central limit theorem. The second is that \( n \) must not be so large that the value of \( s \) varies appreciably over the interval.

| case \( j \) | innovates | red meme selected | contribution | \( \pi_j \) | \( g_j(x) \) |
|------------|-----------|------------------|-------------|---------|---------|
| 1          | yes       | yes              | -           | \( \mu s \) | \( 1 - \pi_1 + \pi_2 f(1 - s + sx) \) |
| 2          | yes       | no               | -           | \( \mu(1 - s) \) | \( 1 - \pi_2 + \pi_2 f(1 - s + sx) \) |
| 3          | no        | yes              | +           | \( (1 - \mu)s \) | \( 1 - \pi_3 + \pi_3 f(s + (1 - s)x) \) |
| 4          | no        | no               | -           | \( (1 - \mu)(1 - s) \) | \( 1 - \pi_4 + \pi_4 f(1 - s + sx) \) |
as the model parameters. In both cases, the scaled and shifted sum of \( n \) independent copies of \( C_j \) converges in distribution to a random variable \( Z \), such that

\[
C_{j1} + C_{j2} + C_{j3} + \ldots + C_{jn} \xrightarrow{d} \frac{Z + b_{j,n}}{a_{j,n}}
\]

for constants \( a_{j,n} \) and \( b_{j,n} \), as \( n \to \infty \). In the classic central limit theorem, \( Z \sim \mathcal{N}(0,1) \), the standard normal distribution, and in the generalised central limit theorem, \( Z \sim S(\alpha, \beta) \), with \( S \) the standardised stable distribution, and \( \alpha \) and \( \beta \) are shape parameters to be described in Sec. II C as well as Appendix A. As such, to approximate \( \Delta s \), it remains only to relate the scale and shift constants, \( a_{j,n} \) and \( b_{j,n} \) for each case \( j \) to the variable \( s \), as well as the model parameters. The relevant results from the classic and generalised central limit theorems are given in Sec. II B below.

Following the arguments above, the time evolution of \( s \) can be found by repeatedly sampling random variables \( Z \), and updating the constants \( a_{j,n} \) and \( b_{j,n} \) according to the central limit theorem at each step. Each step corresponds to \( n \) updates, where the system time advances by \( \Delta t \). This process can be expressed as a finite-difference regime, where

\[
\Delta s = g(s)\Delta t + h(s)\Delta t^{1/2}\frac{Z}{\sqrt{n}}, \quad \text{with} \quad Z \sim \mathcal{N}(0,1) \tag{2}
\]

and

\[
\Delta s = g(s)\Delta t + h(s)\Delta t^{1/\alpha}Z, \quad \text{with} \quad Z \sim S(\alpha, \beta) \tag{3}
\]

when \( p_k \) has finite- and infinite-variance, respectively. The drift and diffusion, \( g(s) \) and \( h(s) \), result from combining the terms \( a_{j,n} \) and \( b_{j,n} \), and are therefore functions of \( s \), the innovation rate \( \mu \), and the underlying network topology, as described below. We integrate Eqs. (2) and (3) using discrete time steps (cf. the Euler-Maruyama scheme), producing trajectories of \( s \) over time.

We briefly describe the initial condition on \( s \) for Eqs. (2) and (3). As described above, every meme enters the system through an innovation tweet, so it initially occupies the screen of the innovating node, as well as its \( k \) followers. The initial value of \( s \) is therefore \((1 + k)/N\), where \( k \) is a random integer with distribution \( p_k \). However, we find that setting \( s(0) = (1 + k)/N \) removes the dependence on \( k \) and yields indistinguishable results, so we choose the latter in this work. Finally, we note that for finite \( \Delta t \), fluctuations can cause the value of the density \( s \) to decrease below zero, which is unphysical. As such, we treat \( s = 0 \) as an absorbing boundary condition, corresponding to the red meme going extinct.

## B. Finite-variance degree distribution

When the out-degree distribution \( p_k \) of the underlying network has finite variance \( \sigma^2 \), fluctuations in the density \( s \) of a given meme have finite variance. This is true on the whole, as well as for fluctuations of each type \( j \). In each case, the sum of these fluctuations is approximated by \( (Z_j + b_{j,n})/a_{j,n} \), as per the classic central limit theorem, with the noise term \( Z_j \) drawn from the standard normal distribution, \( \mathcal{N}(0,1) \). In order to compute \( a_{j,n} \) and \( b_{j,n} \), we require the mean and variance of the fluctuations of each type \( C_j \). Both are straightforward to compute, as detailed in Table I. Then, \( a_{j,n} \) is given by

\[
a_{j,n} = \frac{1}{\sigma_j} n^{-\frac{1}{2}} \tag{4}
\]

and

\[
b_{j,n} = n\langle C_j \rangle a_{j,n}, \tag{5}
\]

where \( \sigma_j^2 = \langle C_j^2 \rangle - \langle C_j \rangle^2 \) is the variance of \( C_j \), and \( \langle C_j \rangle \) is its mean. We can then simplify the signed sum of the four \( (Z_j + b_{j,n})/a_{j,n} \) terms by grouping their deterministic and stochastic components, \( b_{j,n}/a_{j,n} \) and \( Z_j/a_{j,n} \), respectively. The former is simply \( n\langle C_j \rangle \), and the latter can be simplified using rules for scaling and summing normal random variables. In this way, we produce the drift and diffusion terms in Eq. (2), respectively

\[
g(s) = -\langle C_1 \rangle - \langle C_2 \rangle + \langle C_3 \rangle - \langle C_4 \rangle, \tag{6}
\]

and

\[
h(s) = (\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2)^{\frac{1}{2}} N^{-\frac{1}{2}}. \tag{7}
\]

Recall that \( \Delta t = n dt = n/N \), where \( n \) is the number of update steps aggregated, and \( N \) is system size. The \( s \) dependency of \( h(s) \) and \( g(s) \) arises through \( \langle C_j \rangle \) and \( \sigma_j^2 \), respectively, as per Table I. The noise term \( Z \) associated with \( h(s) \) in Eq. (2) is drawn from \( \mathcal{N}(0,1) \). The accuracy of this approximation can be seen in Fig. I(a) and (b), and will be discussed below.

## C. Infinite-variance degree distribution

In the case where the out-degree distribution \( p_k \) has infinite variance, fluctuations in the density \( s \) of a given
TABLE II. Properties of the generating functions \( g_j(x) \) required by the central limit theorems. The mean and variance of \( C_j \) are used in the classical central limit theorem, and the tail parameters \( c_j \) in the generalised central limit theorem. The probability \( \pi_j \) case \( j \) arising is given in the previous table. Note \( c \) is the tail parameter of the underlying degree distribution.

| \( j \) | \( (C_j) \) | \( (C_j^2) - \langle C_j \rangle^2 \) | \( c_j \) |
|---|---|---|---|
| 1 | \( \pi_1(s(k) + 1) \) | \( \pi_1(s^2(\sigma^2 - \langle k \rangle^2) + 2\langle k \rangle s + \langle C_1 \rangle - \langle C_1 \rangle^2) \) | \( c_{1.1} s^{-1} \) |
| 2 | \( \pi_2 s(k) \) | \( \pi_2 s^2(\sigma^2 - \langle k \rangle^2) + \langle C_2 \rangle - \langle C_2 \rangle^2 \) | \( c_{2.1} s^{-1} \) |
| 3 | \( \pi_3(1 - s)/k \) | \( \pi_3(1 - s)^2(\sigma^2 - \langle k \rangle^2) + \langle C_3 \rangle - \langle C_3 \rangle^2 \) | \( c_{3.1}(1 - s)^{-1} \) |
| 4 | \( \pi_4 s(k) \) | \( \pi_4 s^2(\sigma^2 - \langle k \rangle^2) + \langle C_4 \rangle - \langle C_4 \rangle^2 \) | \( c_{4.1} s^{-1} \) |

meme have infinite variance. This is true on the whole, as well as for fluctuations \( C_j \) of each type \( j \). As such, the classic central limit theorem no longer applies. However, for some choices of \( p_k \), namely, those whose tail probabilities are asymptotically power-laws, with exponent \( \gamma \) in the range \( 2 < \gamma < 3 \), fluctuations of each type sum to \( (Z_j + b_j)/a_{j,n} \), where \( Z_j \) is an \( \alpha \)-stable random variable. This is described by the generalised central limit theorem \cite{20}, with the variable \( Z_j \) drawn from the standardised stable distribution \( S(\alpha, \beta) \), and \( a_{j,n} \) and \( b_{j,n} \) are norming constants (see Appendix A for information regarding the parameterisation and sampling of stable distributions).

The first shape parameter \( \alpha \) is the stability parameter, and is related to the power-law decay of the variables being summed. To determine the value of \( \alpha \) for fluctuations \( C_j \) of each of the four types, we consider first the underlying degree distribution. In this work, we use power-law degree distributions \( p_k \sim k^{-\gamma} \), where \( 2 < \gamma < 3 \). As such, the sum of degrees themselves are asymptotically \( \alpha \)-stable, with \( \alpha = \gamma - 1 \). Since the change in the density of a given meme is closely related to the degree \( k \) of the tweeting node, fluctuations also have tail probabilities decaying with power-law exponent \( \gamma \). As a consequence, fluctuations \( C_j \) of each type sum to \( (Z_j + b_{j,n})/a_{j,n} \), asymptotically, with \( Z_j \) being \( \alpha \)-stable. This can be seen \cite{22} by examining the generating functions of these distributions, \( g_j(x) \), in Table I.

The second shape parameter \( \beta \) is related to skewness, and is determined by the relative weights of the upper and lower tail probabilities of the variable being summed \cite{20}. The tail weights of fluctuations \( C_j \) are related to the tail of the underlying degree distribution \( p_k \). In particular, degrees have upper tail weight \( c \) given by the relation \( k^{\alpha}(1 - P_k) \rightarrow c \) as \( k \rightarrow \infty \). Here \( P_k \) is the cumulative distribution function of the degree distribution. In general for stable distributions, the lower tail weight is defined similarly, but here is zero since degrees are non-negative. In other words, \( p_k \) is totally skewed to the right. The tail weights \( c_j \) of the fluctuations \( C_j \) are related to \( c \), but moderated by the density \( s \), as noted in the final column of Table I. Along with the stability parameter \( \alpha \), the tail weights \( c_j \) determine the scaling term \( a_{j,n} \) in the generalised central limit theorem, describing the sum of fluctuations of type \( j \). For distributions of this type \cite{20}, the scaling constant \( a_{j,n} \) in the generalised central limit theorem is given by

\[
a_{j,n} = \left( \frac{2\Gamma(\alpha) \sin(\pi\alpha/2)}{\pi c_j} \right)^{1/\alpha} n^{\frac{1}{\alpha}},
\]

where \( \Gamma \) is the gamma function, and \( n \) as before is the number of terms in the sum, as per Eq. (1), and the generalised central limit theorem. In the interval \( 1 < \alpha < 2 \) that is of interest, the value of \( b_{j,n} \) is the same as in the finite-variance case, given by Eq. (5). In principle, computing \( a_{j,n} \) and \( b_{j,n} \), and sampling \( \alpha \)-stable variables (see Appendix A) allows us to approximate the change in \( s \) after \( n \) node updates using Eq. (8).

As in the finite-variance case, we can simplify the signed sum of the four \( (Z_j + b_{j,n})/a_{j,n} \) terms by grouping together their deterministic and stochastic parts. First, we consider the terms \( b_{j,n}/a_{j,n} \), and then use the relation \( g(s) \), which is the same here as in Eq. (1). Second, the signed sum over \( Z_j/a_{j,n} \) can be computed using a basis property of stable laws, namely, that the sum of \( \alpha \)-stable random variables is also \( \alpha \)-stable. In this way, the diffusion term can be shown to be

\[
h(s) = \left( \frac{\pi(c_1 + c_2 + c_3 + c_4)}{2\Gamma(\alpha) \sin(\pi\alpha/2)} \right)^{1/\alpha} N^{-(\alpha-1)/\alpha},
\]

with the dependence of \( h(s) \) on \( s \) arising through the tail weights \( c_j \), as per Table II. The random term \( Z \) associated with \( h(s) \) in Eq. (9) is drawn from the standardised stable distribution with shape parameters \( \alpha \) and \( \beta = (c_1 - c_2 + c_3 - c_4)/(c_1 + c_2 + c_3 + c_4) \). Here we note that while the distributions of \( C_j \) are totally skewed to the right, their signed sum is not. That is, since \( j = 1 \), \( 2 \) and \( 4 \) contribute negatively, meaning the lower tail is no longer zero. The \( j = 3 \) case contributes positively, giving the upper tail. As such, the value of \( \beta \) for the signed sum of the terms \( Z_j/a_{j,n} \) is given by the relative weights of the upper and lower tails.

Finally, it is useful to point out some differences in Eqs. (2) and (3), particularly in regards to the factor \( \Delta t = n dt = \eta/n \) in the diffusion term. In the finite-variance case noise \( Z \) scales as \( \Delta t^{1/2} \), and in the infinite-variance case scales as \( \Delta t^{1/\alpha} \). When \( \alpha = 2 \), stable theory recovers the normal case. Further, it may be noteworthy that \( Z \) is independent of \( s \) in Eq. (2), but dependent on \( s \) in Eq. (3), via the skewness parameter \( \beta \).
D. Popularity

In the dynamics of the model introduced in Ref. [8], we associate with a meme not only a time-dependent screen density \( s \), but also a popularity, \( y \), a non-negative integer indicating the number of times a given meme has been retweeted since it first appeared. Because popularity is observable in online social networks like Twitter, where it corresponds to retweets, it is easier to compare with data than screen abundance. As such, we model the evolution of \( y \) along with \( s \). The is straightforward to do, since \( y \) is simply incremented by one when a selected node contains the red meme and retweets rather than innovates. This describes case \( j = 3 \), which occurs with probability \( \pi_3 \), as per Table I. We model the fluctuations in \( y \) over \( n \) node selections as \( n \) samples from the Bernoulli distribution with probability \( \pi_3 \), resulting in the equation

\[
\Delta y = g(s) \Delta t + h(s) \Delta t Z, \quad \text{with} \quad Z \sim N(0, 1), \quad (10)
\]

where \( g(s) = \pi_3 N \), the expected value of the Bernoulli distribution, and \( h^2(s) = \pi_3 (1 - \pi_3) N \), the standard deviation of the same distribution. Finally, note that in simulation there is a boundary condition, where \( y \) must remain constant if \( s \) goes to zero. Note also that \( y \) in Eq. (10) is continuous due to the noise term, so we round \( y \) to the nearest integer when reporting it below.
III. RESULTS

Here we discuss a series of computational experiments in which we validate the diffusion approximation through comparison to network-based simulation of the CIC model. We begin by surveying a range of model parameters in Section III A to verify that the diffusion approximation is effective in a variety of settings. Then in Section III B we use the Fokker-Planck equation to derive an expression of the time evolution of the moments of the density $s$ of a given meme. Finally in Section III C we examine the effect of the time step $\Delta t = nd$ on the accuracy of the diffusion approximation.

A. Accuracy of diffusion approximation

In Fig. 1 we use the diffusion approximation to recover a number of experimental results observed in network-based simulation. In Fig. 1(a), we measure the distribution of popularities $y$, as a function of meme age $a$, which is the time elapsed between when a meme was introduced to the system through innovation, and observation time $t$. The follower distribution $p_k$ is 10-regular, and the innovation rate $\mu$ is zero, except initially when blank screens are selected. We observe that at all the ages shown, $a = 0.1, 1$ and $10$, the tail of complementary cumulative distribution of popularities $P_y$ decays exponentially. For larger ages, popularities follow a power law with exponent $-1/2$. This is typical of avalanche size distributions at a dynamical critical point, or of outbreak sizes in a neutral model [17], and recovers the results of Ref. [8]. These observations are reproduced almost exactly by the diffusion approximation, see Fig. 1(b), particularly when $a$ is large. When $a = 0.1$, there appears to be a very slight difference in the rate of the exponential cutoff. The diffusion approximation uses $n = 10^3$.

In Fig. 1(c), we repeat this experiment almost identically, the only difference being that the follower distribution $p_k$ is power-law, rather than regular. The power-law exponent is $\gamma = 2.5$, and we set a minimum degree cutoff of $k = 4$. Here we recover a well-known result, explained in Ref. [8], that popularity $P_y$ decays as a power-law with exponent $-2/3$, for large ages $a$. For early ages, a power law decay is also apparent, but this appears to be a reflection of the out-degree distribution. That is, when a meme first enters the system, it appears on the screen of the innovating node, as well as its $k$ followers. Since the follower distribution has exponent $-5/2$, this appears as $-3/2$ in the ccdf picture. Once again, these observations are recovered with high accuracy by the diffusion approximation in Fig. 1(d), especially for larger $a$.

Finally, in Fig. 1(e) we examine the effect of allowing innovation with probability $\mu = 0.1$ and 0.3. We restrict ourselves to the popularity distribution for memes of age $a = 10$. Otherwise, experiments are identical to Fig. 1(c), which is reproduced here as the $\mu = 0$ case for comparison. When $\mu = 0.1$, the distribution of popularities $y$ is made more narrow. For larger values, $\mu = 0.3$, the distribution once again tends to a power law with exponent $-3/2$. As before, this appears to be a sign that there are relatively few memes to survive to this age, due to the increased competition for screen space. As such, most memes die out soon after being born, meaning their popularity was proportional to the number of followers $k$ that they were initially broadcast to. This is supported by the fact that this appears to be a limiting distribution for $\mu > 0.3$, not shown. These results are supported by the diffusion approximation, Fig. 1(f).

B. Fokker-Planck equation and its moments

In this section we use the Fokker-Planck equation to derive the mean and variance of meme density $s$, as the fraction of screens occupied by a given meme, as a function of the time since it was introduced. It is this that we refer to by time $t$ from now on, rather than total system time. We restrict ourselves to the case where the follower distribution $p_k$ has finite-variance. The Fokker-Planck equation [23] describes the time evolution of $p(s, t)$,
the probability density function of screen fraction $s$ as a function of time $t$. Beginning with a Langevin equation describing a single trajectory, such as Eq. (2), the Fokker-Planck equation can be written

$$\frac{\partial}{\partial t} p(s, t) = -\frac{\partial}{\partial s} [g(s)p(s, t)] + \frac{\partial^2}{\partial s^2} [D(s)p(s, t)],$$

(11)

where $g(s)$ and $D(s) = h^2(s)/2$ are the drift and diffusion terms of the single particle equation, defined in Eqs. (3) and (7), respectively. We assume that initially, $p(s, 0) = \delta(1 + \langle k \rangle)/N$, where $\delta$ is the Dirac delta, since all individual trajectories start at this value. As it stands, the Fokker-Planck equation cannot be solved explicitly for our system, but can of course be integrated numerically using standard techniques.

It is possible to obtain an analytic estimate for $\langle s \rangle$ and $\sigma_s$, without solving the Fokker-Planck equation for $p(s, t)$ directly. To this end, observe that at time $t$, the $m$-th moment of $s$ is given by

$$\langle s^m \rangle = \int_0^1 s^m p(s, t) ds.$$  

(12)

As such, by multiplying both sides of Eq. (11) by $s^m$, and integrating over all $s$, we obtain an ordinary differential equation describing the time evolution of the $m$th moment $\langle s^m \rangle$. In general, this too cannot be solved in closed form. However, by including only the first-order terms in $s$ in the drift and diffusion expressions $g(s)$ and $h(s)$, respectively, the equation becomes tractable. In particular, when $m = 1$ and $2$, the resulting equation can be solved for $\langle s \rangle$ and $\langle s^2 \rangle$ using standard techniques, which in turn yield an expression for $\sigma_s$. Concretely, this procedure yields

$$\langle s \rangle = \kappa_1 e^{-\mu(\langle k \rangle + 1)t}$$

(13)

and

$$\sigma_s^2 = \kappa_2 \left( e^{-\mu(\langle k \rangle + 1)t} - e^{-2\mu(\langle k \rangle + 1)t} \right),$$

(14)

where $\kappa_1$ and $\kappa_2$ are constant functions of network topology, through $\langle k \rangle$ and $\sigma$, and the innovation probability $\mu$. The first constant, $\kappa_1 = (\langle k \rangle + 1)/N$, is simply the initial value of $s$, while the second, $\kappa_2 = (\mu(\langle k \rangle + \langle (k)^2 + \sigma^2 \rangle)/(1 - \mu))/\mu N^2$, indicates that $\sigma_s$ is an increasing function of $\langle k \rangle$ and $\sigma$. The effectiveness of this approximation is evident in Fig. 1(a), where it is clear that the first-order approximation of $s$ is appropriate. In this figure, solid curves result from network-based simulation of CIC. For each value of $\mu$, we simulate the network dynamics for $100N$ update steps, and track the screen count $s$ for each meme, as a function of the time since it was introduced to the system. There appears to be only a very small systematic error in $\sigma_s$, with the theoretical moments overestimating the value observed in network simulation. It is likely that the linearisation of the diffusion term causes this. No such error is apparent in Fig. 1(a).

C. Dependence on $n$, or time step $\Delta t$

In this section we examine how the parameter $n$ affects the quality of the diffusion approximation. This parameter is defined in Eq. (1), and is the number of update events that are aggregated to produce a single time step of the diffusion approximation. It appears in the finite-difference implementation, Eqs. (2) and (3), as $\Delta t = n dt = n/N$. Our choice of $n$ in practice is determined by two competing effects. On one hand, the larger the value of $n$, the better the approximation of the sum of the fluctuations $C_j$ as a normal distribution, as per the central limit theorem. This is especially true when $s$ is small, such as at early times. On the other hand, the smaller the value of $n$, the better the finite-difference solution. It is generally true that finite-difference methods are more accurate the smaller the value of $\Delta t = n dt$, as we assume that the variable in question, here $s$, is constant over the integration time step. As such, $n$ must be large enough for the central limit theorem to apply, but small enough that $s$ does not vary appreciably over that interval.

We verify the effect of increasing $n$ in Fig. 3(a) and (b), where solid curves show the result of the diffusion approximation. Experimentally, we use the same settings as in Fig. 1(b) for memes at ages $a = 10$. We set $n = 10^3$ in that figure, which is reproduced as a baseline result here because it matches well to the numerical results of Fig. 1(a). The popularity curves are visibly distorted as we increase $n$ from $10^3$ up to $10^7$, both when the follower distribution $p_k$ is regular and power-law. It is difficult to know a priori which value of $n$ to choose, in this work we have used a value of $n = 10^3$, unless stated otherwise.

IV. DISCUSSION

In this work we have approximated a model of meme popularity as a one-dimensional stochastic process, and shown it to be effective in recovering a range of critical spreading phenomena on networks. Depending on whether the degree distribution has finite or infinite variance, dynamic fluctuations can be approximated by Gaussian and $\alpha$-stable noise, respectively. The effectiveness of this approximation is revealed, shedding light on the mechanisms that ultimately drive dynamical outcomes. In particular, our results show that in the finite-variance case, only the first two moments $\langle k \rangle$ and $\sigma$ of the follower distribution $p_k$ play a role. In the infinite variance case, only the mean $\langle k \rangle$, the power-law exponent $\alpha$ and its tail weight $\omega$ are important. As such, any approximation using the full out-degree distribution $p_k$ is wasteful by comparison. In addition, the diffusion approximation shows that memes can be modelled independently of one another. That is, each realisation of the Langevin equations, modelling the spread of a single meme, is independent of all other realisations. A clear advantage of this is that trajectories can be computed in
FIG. 3. Complementary cumulative distribution of meme popularity $y$, using the diffusion approximation with various time step sizes $\Delta t = n \, dt = n/N$. The follower distribution $p_k$ is 10-random regular, (a), and power-law with lower cutoff at $k = 4$, exponent 2.5, and resulting tail weight $c = 6.629$, (b). For each value of $n$, $10^7$ trajectories were computed, either until memes reached an age of $a = 10$, or until they died out. Curves at $n = 10^4$ in (a) and (b) correspond to the $a = 10$ curves in Fig. 1(b) and (d), respectively.

parallel, something that is impossible in network-based simulations of the dynamics. This gain in efficiency is especially noteworthy in the limit of large networks, where memory is obviously a constraint when simulating networks themselves, in contrast to the diffusion approximation. In a similar vein, the diffusion approximation allows us to sample as many, or as few trajectories as we like. While a large number of memes may emerge in a large network, their distribution may be revealed by a much smaller sample of memes using the diffusion approximation. Small networks may present the opposite problem, in which case the diffusion approximation allows us to extrapolate by sampling additional trajectories.

Another advantage of the diffusion approximation is that it naturally gives rise to a suite of analytical tools. In particular, diffusion processes have a corresponding Fokker-Planck equation, which we have exploited to derive expressions for the temporal evolution of the moments of meme abundance as an ensemble. We used this to calculate the mean and variance of meme abundance as a function of time and demonstrated that a small-s linearisation gives accurate closed-form solutions. This new perspective paves the way for future work, given the ease with which slight modifications can be made to allow the study of different meme spreading dynamics, such as when memes attract user attention heterogeneously (i.e., some are more likely to spread than others), sometimes referred to as fitness models [24].

One limitation of our work is that we have assumed that the in-degree, meaning the number of individuals that a give node follows, is constant across the network. Allowing this distribution to be arbitrary leads to small, but noticeable disagreements in our results, such as the accuracy of the Fokker-Planck approximation of the moments of $s$. Extending our analytical results to account for this is a potential direction for future work.

ACKNOWLEDGEMENTS

This work was partly supported by Science Foundation Ireland under grant numbers 16/IA/4470 (all authors), 12/RC/2289 P2 (J.G.) and 16/RC/3918 (J.G.).

† Corresponding author: Kleber.Oliveira@ul.ie

Appendix A: Sampling stable variables

In this appendix we describe the numerical procedure for simulating $\alpha$-stable variables as is necessary for the diffusion term in Eq. 3. Following the formulation of the Chamber-Mallows-Stuck algorithm from Ref. [20], we let $\Theta$ and $W$ be independent random variables with $\Theta$ uniformly distributed on $(-\pi/2, \pi/2)$ and $W$ exponentially distributed with mean 1. We focus on the interval $1 < \alpha < 2$. For the skewness parameter $\beta$ in Eq. 3, we have $-1 \leq \beta \leq 1$, determined by the tail parameters of the variables being summed, as described in the main text. We define $\theta = \tan^{-1}(\beta \tan(\pi \alpha/2))/\alpha$. Then, the random variable $Z$ given by

$$Z \sim \frac{\sin(\theta + \Theta)}{(\cos \alpha \theta \cos \Theta)^{1/\alpha}} \left(\frac{\cos((\alpha - 1)\theta)}{W}\right)^{1-\alpha/\alpha}$$  \quad (A1)

follows a $S(\alpha, \beta)$ distribution [20]. In this paper we have used $\alpha = 1.5$ throughout. For clarity, recall that $\alpha$ here is related to the exponent of the power-law degree distribution $\gamma$ by $\alpha = \gamma - 1$, and $\beta$ is determined by the relative weights of the tail probabilities $c_j$, specifically $\beta = (-c_1 - c_2 + c_3 - c_4)/(c_1 + c_2 + c_3 + c_4)$, which in turn are defined in Table 1. In the same vein, normal variables can be obtained starting with uniform random variables using the Box-Muller transform. Note that two parameterisations are outlined in Ref. [20], and flagged notationally by 0 or 1. Throughout this work we adhere to the latter, which is written $S(\alpha, \beta; 1)$ in Ref. [20].
1] Z.-K. Zhang, C. Liu, X.-X. Zhan, X. Lu, C.-X. Zhang, and Y.-C. Zhang, Dynamics of information diffusion and its applications on complex networks, Phys. Rep. 651, 1 (2016).

2] E. Ferrara, M. JafariAsbagh, O. Varol, V. Qazvinian, F. Menczer, and A. Flammini, Clustering memes in social media, in Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ACM, New York, 2013) p. 548–555.

3] N. O. Hodas and K. Lerman, The simple rules of social contagion, Sci. Rep. 4, 4343 (2014).

4] J. Cheng, L. Adamic, P. A. Dow, J. M. Kleinberg, and J. Leskovec, Can cascades be predicted?, in Proceedings of the 23rd International Conference on World Wide Web (ACM, New York, 2014) p. 925–936.

5] J. M. Miotto, H. Kantz, and E. G. Altmann, Stochastic dynamics and the predictability of big hits in online videos, Phys. Rev. E 95, 032311 (2017).

6] D. Notarmuzi, C. Castellano, A. Flammini, D. Mazzilli, and F. Radicchi, Universality, criticality and complexity of information propagation in social media, Nat. Commun. 13, 1308 (2022).

7] S. Goel, A. Anderson, J. Hofman, and D. J. Watts, The structural virality of online diffusion, Manage. Sci. 62, 180 (2016).

8] J. P. Gleeson, J. A. Ward, K. P. O’Sullivan, and W. T. Lee, Competition-induced criticality in a model of meme popularity, Phys. Rev. Lett. 112, 048701 (2014).

9] L. Weng, A. Flammini, A. Vespignani, and F. Menczer, Competition among memes in a world with limited attention, Sci. Rep. 2, 335 (2012).

10] G. L. Ciampaglia, A. Flammini, and F. Menczer, The production of information in the attention economy, Sci. Rep. 5, 9452 (2015).

11] P. Lorenz-Spreen, B. M. Monsted, P. Hövel, and S. Lehmann, Accelerating dynamics of collective attention, Nat. Commun. 10, 1759 (2019).

12] K. Lerman, Information is not a virus, and other consequences of human cognitive limits, Future Internet 8, 10.3390/fi8020021 (2016).

13] J. P. Gleeson, K. P. O’Sullivan, R. A. Baños, and Y. Moreno, Effects of network structure, competition and memory time on social spreading phenomena, Phys. Rev. X 6, 021019 (2016).

14] S. N. Dorogovtsev, A. V. Goltsev, and J. F. Mendes, Critical phenomena in complex networks, Reviews of Modern Physics 80, 1275 (2008).

15] S. Chandra, E. Ott, and M. Girvan, Critical network cascades with re-excitable nodes: Why treelike approximations usually work, when they break down, and how to correct them, Phys. Rev. E 101, 062304 (2020).

16] R. A. Bentley, S. Carrignon, D. J. Ruck, S. Valverde, and M. J. O’Brien, Neutral models are a tool, not a syndrome, Nat. Hum. Behav. 5, 807 (2021).

17] O. A. Pinto and M. A. Muñoz, Quasi-neutral theory of epidemic outbreaks, PLOS ONE 6, 1 (2011).

18] S. di Santo, P. Villegas, R. Burioni, and M. A. Muñoz, Simple unified view of branching process statistics: Random walks in balanced logarithmic potentials, Phys. Rev. E 95, 032115 (2017).

19] D. Notarmuzi and C. Castellano, Analytical study of quality-biased competition dynamics for memes in social media, EPL 122, 28002 (2018).

20] J. P. Nolan, [Univariate Stable Distributions], 1st ed., Springer Series in Operations Research and Financial Engineering (Springer International Publishing, 2020).

21] D. J. Higham., An algorithmic introduction to numerical simulation of stochastic differential equations, SIAM Rev. 43, 525 (2001).

22] A. Faqeeh, S. Osat, F. Radicchi, and J. P. Gleeson, Emergence of power laws in noncritical neuronal systems, Phys. Rev. E 100, 010401 (2019).

23] D. Sornette, [Critical Phenomena in Natural Sciences], 2nd ed., Springer Series in Synergetics (Springer-Verlag Berlin Heidelberg, 2006).

24] D. F. Oliveira and K. S. Chan, The effects of trust and influence on the spreading of low and high quality information, Physica A 525, 657 (2019).