

Virality detection and control strategies in rumor models

Eva Rifa ^{*}

*Eurecat, Centre Tecnològic de Catalunya, Barcelona, Spain
and Facultat de Física, Universitat de Barcelona (UB), Barcelona, Spain*

Julian Vicens 

Eurecat, Centre Tecnològic de Catalunya, Barcelona, Spain

Emanuele Cozzo 

*Universitat de Barcelona Institute of Complex Systems (UBICS), Universitat de Barcelona, Barcelona, Spain;
Departament de Física de la Matèria Condensada, Universitat de Barcelona, Barcelona, Spain;
and Communication Networks & Social Change research group (CNSC), Universitat Oberta de Catalunya, Barcelona, Spain*



(Received 7 October 2025; accepted 18 February 2026; published 24 March 2026)

We study the dynamics and intervention strategies of a rumor using the modified Maki-Thompson model. A key challenge in social networks is distinguishing between natural increases in transmissibility and artificial injections of rumor spreaders, such as through broadcast events or astroturfing. Using stochastic simulations, we compare two scenarios: one with organic growth in transmissibility and another with externally injected spreaders. Although both lead to high autocorrelation, only the organic growth produces oscillatory patterns in autocorrelation at multiple lags, an effect we can analytically explain using the N -intertwined mean-field approximation. This distinction offers a practical tool to identify the origin of rumor virality and also infer its transmissibility. Our approach is validated analytically and tested on real-world data from Twitter during the announcement of the Higgs boson discovery. In addition to detection, we also explore control strategies. We show that the average lifetime of a rumor can be manipulated through targeted interventions: placing spreaders at specific locations in the network. Depending on their placement, these interventions can either extend or shorten the lifespan of the rumor.

DOI: [10.1103/2bh8-p6rx](https://doi.org/10.1103/2bh8-p6rx)

I. INTRODUCTION

In the digital era, social networks and platforms have become central elements in the dissemination of information, influencing the dynamics of public discourse and communication. It has long been recognized that as long as dominant social media platforms are owned by monopolistic companies whose primary goal is profit maximization, they cannot serve as spaces where rumor propagation is truly spontaneous. That is why virality, understood as the widespread reproduction of a message, is not a neutral or uniform process. Furthermore, these platforms have enabled mechanisms of persuasion in the battle over public opinion, including disinformation campaigns and the emergence of covert manipulation tactics such as astroturfing [1].

Astroturfing refers to the practice of simulating the appearance of a grassroots movement through the mass mobilization

of fake or automated accounts [1]. While this practice predates the internet, its current scale, amplified by platforms such as X (formerly Twitter), makes it cheaper, more global, and more difficult to detect [2]. It has been used in a variety of contexts, including aggressive marketing strategies and electoral manipulation. The Internet Research Agency's campaigns during the 2016 U.S. presidential elections [3] are a well-known example, but recent studies have documented similar operations in South Korea [4] and Thailand [5], among others [6]. Since ground truth is rarely available [4], researchers suggest focusing on identifying coordinated behavior rather than specific content [7]. Astroturfing must be driven by the interests of economic or political power. When this is not the case, we can more generally speak of an external impulse to the system. For example, the diffusion through traditional media (television, radio, streaming) can trigger similar dynamics without manipulative intent [8,9]. These dynamics allow us to delve deeper into the study of viral phenomena and develop strategies to manage the propagation of information online.

One of the most extensively studied models for rumor spreading is the Maki-Thompson (MT) model [10], which, similar to epidemiological models, describes individuals transitioning between states of ignorance, spreading, and stifling depending on their interactions. For many years, it was believed that this model would always have a nonzero proba-

^{*}Contact author: eva.rifa@eurecat.org

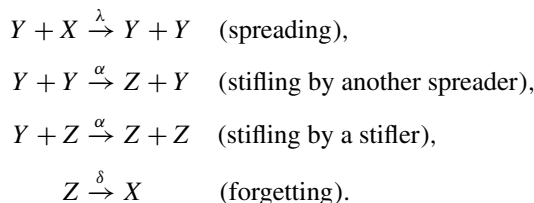
bility of reaching a nonvanishing fraction of the population [11] and, therefore, that it did not exhibit a phase transition. However, a recent study by Ferraz de Arruda *et al.* [12] demonstrated the existence of a critical phase transition. This finding, which arises from accounting for node-to-node correlations neglected in mean-field approaches, challenged long-standing assumptions and opened directions for research in rumor dynamics.

The existence of a critical transition in the rumor model has important implications. In particular, it enables the study of early warning signals (EWS), statistical indicators that detect when systems are approaching a tipping point and experiencing critical slowing down (CSD). EWS have been successfully applied in fields such as epidemiology [13], ecology and climate science [14], while, as far as we know, they have not yet been explored in the context of rumor dynamics. Moreover, while in recent years intervention strategies and forecasting have become standard tools in epidemic modeling [15], only a few studies have taken a similar approach to investigate the spread of rumors [16,17], which do not account for the system's critical dynamics.

The current context, together with theoretical advances in rumor spreading modeling, leads us to ask the following questions: Can we distinguish whether the rumor virality is driven by an external impulse (such as a broadcast or astroturfing campaign) or by an internal increase in transmissibility? And what are the optimal strategies to control the reach of a rumor? Answering these questions is not only theoretically relevant but also brings significant practical implications. Distinguishing between the organic growth of rumors and coordinated external behavior is crucial for identifying disinformation campaigns and preserving the integrity of online discourse. Furthermore, by identifying the optimal way for controlling the rumor spread, it can serve as a basis for designing robust communication strategies to amplify beneficial information and mitigate harmful content on digital platforms. This article is organized as follows: Sec. II presents the model, Sec. III shows the results, and Sec. IV provides the discussion.

II. MODEL DEFINITION

We consider the modified version of the MT rumor spreading model [12], where each individual (node i) can be in one of three states: ignorant ($X_i = 1$), spreader ($Y_i = 1$), or stifter ($Z_i = 1$), where $X_i + Y_i + Z_i = 1$. These are modeled as Bernoulli random variables taking values in $\{0, 1\}$. Transitions between states occur as independent Poisson processes on an undirected network defined by adjacency matrix $A = (A_{ij})$. The transition rules are as follows:



The difference with an epidemiological SIRS (Susceptible-Infected-Recovered-Susceptible) model resides in the removal of spreaders. While in the latter the transition is spontaneous,

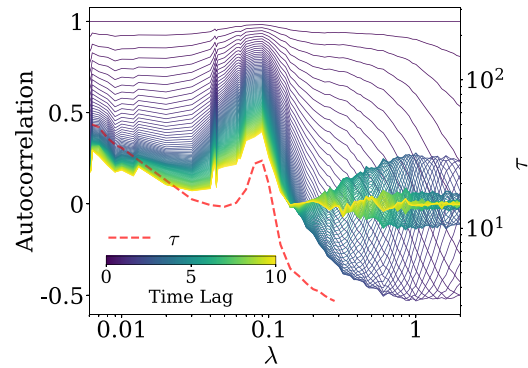


FIG. 1. The autocorrelation as a function of the spreading rate (λ) at multiple lags of the density of spreaders starting from a single spreader. The red dashed line indicates the average lifetime of finite realizations (τ). The autocorrelation is computed from simulations using the quasistationary method. The parameters used are $\delta = 0.89$, $\alpha = 0.5$ for a network of size $N = 1000$ and $\langle k \rangle = 10$.

in the former it depends on the contacts. With respect to the original MT model [10], in which $\delta = 0$ and stiflers never become susceptible to be spreaders again, in the modified MT model, $\delta > 0$ introduces a spontaneous transition from stifter to ignorant. Thus, while the original model has an infinite number of absorbing states, namely, any state with $Y_i = 0 \forall i$, the modified model has a unique (rumor-free) absorbing state ($X_i = 1 \forall i$).

As showed in Ref. [12], a phase transition exists both in the MT original and modified model. In the original MT model, the critical point separates a regime in which the number of stiflers when the process reaches one of the infinite absorbing states does not scale with the system size, from one in which it diverges with the system size. In the modified model, the critical point separates a subcritical regime in which the process ends in the absorbing state, from a regime in which it stays in an active state. The lifetime of a rumor diverges at the critical point. However, while in epidemiological models the lifetime decays exponentially in the subcritical regime, in considered rumor models has a nonmonotonic power-law behavior (see Fig. 1).

The system evolves as a continuous-time Markov chain on the constrained state space $\{0, 1\}^{3N}$, with transition rates determined by pairwise interactions. Thus, it admits in principle an exact analysis. However, it becomes intractable for large N , so we consider a first-order approximation by taking expectations of the indicator variables:

$$\begin{aligned} \frac{dx_i}{dt} &= \delta z_i - \lambda \sum_{k=1}^N A_{ki} x_i y_k, \\ \frac{dy_i}{dt} &= \lambda \sum_{k=1}^N A_{ki} x_i y_k - \alpha \sum_{k=1}^N A_{ki} y_i (y_k + z_k), \\ \frac{dz_i}{dt} &= -\delta z_i + \alpha \sum_{k=1}^N A_{ki} y_i (y_k + z_k), \end{aligned} \quad (1)$$

where $x_i(t) = \mathbb{E}[X_i(t)]$, $y_i(t) = \mathbb{E}[Y_i(t)]$, and $z_i(t) = \mathbb{E}[Z_i(t)]$. This approximation captures the average behavior of the system while accounting for local network structure via the adjacency matrix.

III. RESULTS

First, we study EWS in the rumor model (Sec. III A 1) and whether they allow us to distinguish organic growth from external injections (Sec. III A 2). Second, we derive these results analytically (Secs. III A 3 and III A 4), and third, we present a case study (Sec. III A 5). Finally, we show how targeted interventions can lengthen or shorten the rumor's lifetime (Sec. III B).

A. Analysis of rumor dynamics through early warning signals

1. Autocorrelation patterns of the rumor model

We investigate whether EWS can provide insights into the dynamics of rumor spreading. One of the most commonly used statistical indicators for detecting critical transitions is the autocorrelation [13,15], which measures the correlation of a signal with a delayed version of itself. As a system approaches a critical point, its recovery rate from perturbations typically slows down, leading to an increase in autocorrelation. This phenomenon, known as CSD, makes autocorrelation a powerful tool for anticipating shifts in critical complex systems.

We begin by analyzing the autocorrelation of the density of spreaders at different lags (the time separation between two observations) across varying spreading rates to explore the different regimes of the model. Additionally, we examine the lifetime (τ) [18] of the rumor as a marker to identify the critical point (λ_c) (see Appendix A 3 for details). As expected, the autocorrelation increases near the λ_c , consistent with critical slowing down (see Fig. 1).

However, an important subtlety arises in the subcritical regime, where the system evolves very slowly due to the rarity of spreading events. In this regime, autocorrelation remains artificially high, not because the system is approaching a transition, but because it is barely changing. This slow behavior leads to high autocorrelation values that can be misleading if interpreted in isolation. This finding suggests that, in rumor models, autocorrelation should be interpreted in the context of dynamical regime of the system.

2. Distinguishing organic growth from an external impulse

We now focus on realistic rumor spreading dynamics and investigate whether EWS can distinguish different types of virality, understood as distinct pathways through which a rumor can reach large fraction of the population. In particular, we examine whether EWS are able to discriminate between externally induced virality and that arising endogenously from the system itself.

To address this question, we designed two scenarios that lead to a similar density of spreaders, but that correspond to different routes toward a viral regime. The first scenario consists of a density pulse that mimics a sudden intrusion of new spreaders interacting with the rumor, a process associated with astroturfing strategies or, more generally, external shocks. The second scenario corresponds to an abrupt increase in the spreading rate, causing the system to transition from a subcritical to a supercritical regime, thus giving rise to endogenous virality emerging from the internal dynamics of the system. In both experiments, we employ two early warning

indicators to track changes in the system's dynamics. One of them is the autocorrelation, which we compute from the density of spreaders, capturing the similarity of the system's state across successive time intervals. The other is the mutual information (MI) between node states, which quantifies the amount of information obtained about one node by observing the other.

To simulate the first scenario, corresponding to an organic growth of the interest in the rumor, we instantaneously tune the transmissibility parameter from the subcritical to the supercritical regime. The second scenario represents an external impulse and is implemented through an artificial injection of spreaders. In this case, we consider four variants of the injection mechanism, accounting for two independent factors. On the one hand, we distinguish between the activity of the injected spreaders: either permanent spreaders, which remain continuously active, or transient spreaders, which behave like normal agents and follow the same rumor spreading rules. On the other hand, we consider the spatial placement of the injected spreaders, distinguishing between random injection and localized injection near existing spreader nodes. A schematic representation of the scenarios and the corresponding indicators is shown in Fig. 2, while further methodological details are provided in Appendix A.

For the transmissibility increase scenario, we observe that the mutual information exhibits a sharp increase at the time of the perturbation and remains high thereafter. Similarly, the autocorrelation displays a peak at the perturbation, followed by a relaxation to a value higher than that observed prior to the transition. In the case of artificially injected spreaders, the perturbation induces an abrupt change in the time series of spreaders density, which is also reflected in the early warning indicators across all injection variants. In particular, the mutual information shows a sudden increase at the perturbation both for the transmissibility increase and for the injection of permanent spreaders.

The autocorrelation, however, exhibits more heterogeneous behavior depending on the injection mechanism. When permanent spreaders are introduced, a pronounced peak is observed at the perturbation, after which the autocorrelation stabilizes at a higher or lower value depending on whether the injection is random or localized. In contrast, when the injected spreaders behave as normal agents, the autocorrelation takes longer to stabilize, and in the case of localized injection, it displays a drop rather than a peak.

The key challenge is therefore to distinguish, based solely on the post-perturbation time series of spreaders, whether the observed change originates from an internal dynamical transition or from an external influence. Our results indicate that this distinction remains ambiguous, as the temporal evolution of the autocorrelation in the two scenarios exhibits qualitatively similar features.

To resolve this ambiguity, we examine the autocorrelation across lags [Fig. 3(a)]. In the case of a permanent increase in transmissibility that leaves the system in a metastable steady state, correlations decay as damped oscillations; an external injection, by contrast, produces an exponential decay. For the increase in transmissibility, we can also see oscillations when we measure only the time series for the new spreaders density [Fig. 3(b)], counted as the new infected nodes during

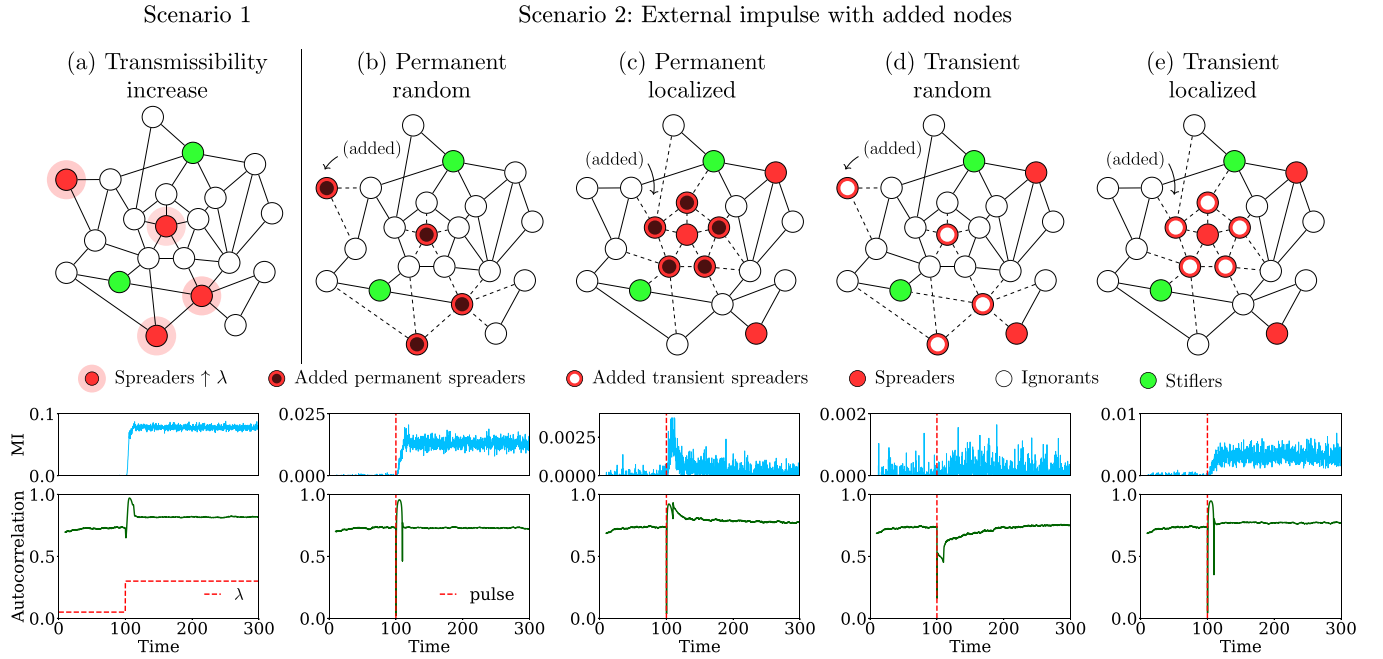


FIG. 2. Schematic representation of two scenarios: (1) sudden increase in the transmissibility rate (λ) and (2) external impulse with fixed spreading rate. Top panels show network representations: Solid edges represent existing connections, while dashed edges indicate externally added nodes. In scenario (2), four variants are simulated: (b) a pulse of permanent spreaders added at a random locations; (c) a localized pulse of permanent spreaders added adjacent to an existing spreader; (d) a pulse of transient spreaders (converted to common population) added at random locations; and (e) a localized pulse of transient spreaders added adjacent to an existing spreader. Bottom panels show the mutual information (MI) and autocorrelation of the spreaders density. Results are averaged over 1000 realizations using a rolling window of 50 time steps. Parameters are $\alpha = 0.5$, $\delta = 0.89$, $N = 1000$, $\langle k \rangle = 10$, and the external perturbation is of 50 new spreaders. See Appendix A for simulation details.

each time step. This finding is very useful when working with real data, where it is often difficult to define when a spreader becomes a stifter. Therefore, it will be much easier for us to measure the density of new spreaders. Furthermore, oscillations are also visible for different network structures (see Appendix B).

3. The NIMFA autocorrelation in the steady state

In this section, we developed an analytical approximation for the autocorrelation in the steady state using the N -intertwined mean-field approximation (NIMFA) [19–21]

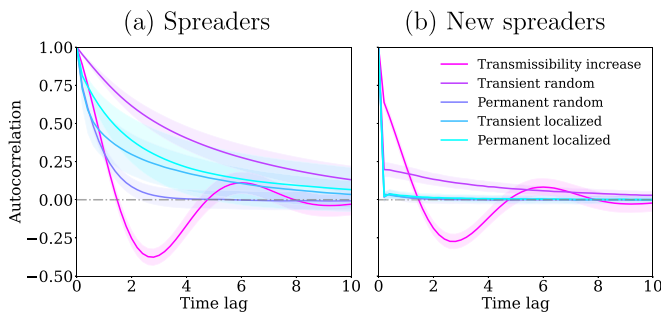


FIG. 3. Autocorrelation of (a) spreaders and (b) new spreaders density as a function of lag after the perturbation for the different scenarios of Fig. 2.

that allows us to write effective rates ($\tilde{\lambda}_\infty$, $\tilde{\alpha}_\infty$) for each node transition.

$$\tilde{\lambda}_{i\infty} \equiv \lambda \sum_{k=1}^N A_{ki} y_{i\infty}, \quad (2)$$

$$\tilde{\alpha}_{i\infty} \equiv \alpha \sum_{k=1}^N A_{ki} (y_{k\infty} + z_{k\infty}). \quad (3)$$

The autocorrelation of the spreading state (V) of a node j between time s and t is

$$R_j(s, t) \triangleq \frac{E[V_j(s)V_j(t)] - E[V_j(s)]E[V_j(t)]}{\sqrt{\text{Var}[V_j(s)]\text{Var}[V_j(t)]}}. \quad (4)$$

In a first-order approximation, we assume that the probabilities are independent for Bernoulli random variables $E[V_i(t)V_j(t)] = E[V_i(t)]E[V_j(t)]$. Upon further derivation (see Appendix C), the analytical expression for the autocorrelation is obtained:

$$R_{i\infty}(h) = \frac{1}{1 - v_{i\infty}} \left[\frac{1}{\tilde{\alpha}\tilde{\lambda} + \tilde{\alpha}\delta + \tilde{\lambda}\delta} \left[\delta\tilde{\lambda} + \frac{\tilde{\alpha}}{2}(\tilde{\lambda} + \delta)(e^{\mu_2 h} + e^{\mu_3 h}) + \frac{\tilde{\alpha}(\tilde{\alpha}\tilde{\lambda} - \tilde{\lambda}^2 + \tilde{\alpha}\delta - \delta^2)}{2w}(e^{\mu_2 h} - e^{\mu_3 h}) \right] - v_{i\infty} \right], \quad (5)$$

where

$$\mu_{2,3} = -\frac{1}{2}(\delta + \tilde{\lambda} + \tilde{\alpha}) \pm w, \quad (6)$$

$$w = \frac{1}{2}\sqrt{(\delta + \tilde{\lambda} + \tilde{\alpha})^2 - 4(\tilde{\lambda}\delta + \tilde{\alpha}\delta + \tilde{\lambda}\tilde{\alpha})}, \quad (7)$$

which corresponds to a damped oscillator with angular frequency w when the eigenvalues of the transition rate matrix become complex.

4. Conditions for the emergence of oscillations

We observe in simulations that oscillations emerge only when the system reaches a metastable steady state, which occurs for sufficiently large spreading rate λ . This behavior contrasts with the trivial absorbing state, where the rumor does not spread through the network and both y and z vanish ($y, z = 0$). Analytically, we have shown that oscillations can only occur when the frequency w [given by Eq. (7)] becomes complex, that is, when the argument of the square root in the expression is negative. For a random regular network with average degree k , this condition when $y = 0$ simplifies to

$$\delta^2 + (\alpha k z_\infty)^2 - 2\alpha k z_\infty \delta > 0. \quad (8)$$

From this expression, we see that when $y = 0$, w is always real. Thus, oscillations are analytically ruled out in the absorbing state, and can only arise in the regime where $y > 0$ at the metastable state where the rumor spreading persists. Similar to the SIS (Susceptible-Infected-Susceptible) model in epidemiology, the system exhibits a phase transition at the critical spreading rate λ_c , below which it remains in the absorbing (rumor-free) state and above which it reaches an active metastable regime. This provides strong evidence that the appearance of oscillations is only possible above the critical threshold λ_c .

To explore this regime analytically, we study the fixed points of the mean-field equations [Eq. (1)] in the steady state. Fixed points correspond to the long-term behavior of the system when the time derivatives vanish. In addition to the absorbing state fixed point mentioned above, the system admits a nontrivial (metastable) fixed point. For a random regular network, where all nodes are assumed to be statistically identical ($x_i = x$, $y_i = y$, and $z_i = z$) and using the identity, $\sum_{i=1}^N A_{ki} = k$ along with the normalization condition $x + y + z = 1$, we obtain the following steady-state solution:

$$\begin{aligned} x_\infty &= \frac{\alpha}{\alpha + \lambda}, \\ y_\infty &= \frac{\lambda \delta}{\alpha \lambda k + \alpha \delta + \lambda \delta}, \\ z_\infty &= \frac{\alpha \lambda^2 k}{(\alpha + \lambda)(\alpha \lambda k + \alpha \delta + \lambda \delta)}. \end{aligned} \quad (9)$$

This nontrivial fixed point characterizes the metastable regime in which a nonzero fraction of spreader and stiflers coexist, and from which oscillatory behavior may emerge depending on the parameter values.

We now compare the analytical results with simulations. In Fig. 4 (a), we show the autocorrelation of the density of spreaders obtained from simulations. In Fig. 4 (b), we show the NIMFA autocorrelation for a random regular network.

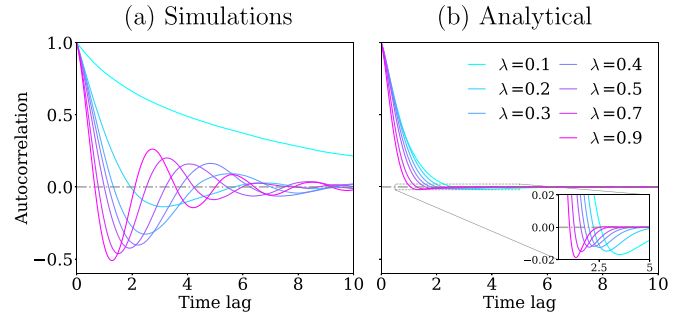


FIG. 4. Autocorrelation as a function of the lags for different spreading rates. Panel (a) shows results from simulations, oscillating at $\lambda > \lambda_c$. Panel (b) shows the NIMFA autocorrelation as a function of the lag at different rates. Parameters: $N = 1000$, $\langle k \rangle = 10$, $\alpha = 0.5$, and $\delta = 1$.

Although the analytical results do not visibly display oscillations as clearly as in the simulations, we can still compute the oscillation period from the model equations and compare it to the period of the simulations. As shown in Fig. 5, the analytically derived period agrees well with that observed in simulations. It is important to note that this agreement holds only for specific parameter values; in this case, $\alpha = 0.5$ and $\delta = 1$.

We can approximate the lower and upper bounds for the spreading rate λ that allow oscillations by substituting the steady-state values from Eq. (9) into Eq. (7) using the parameters $\delta = 1$, $\alpha = 0.5$, and $k = 10$. Under these conditions, oscillations are present in the range $0.068\,0854 < \lambda < 6.031\,05$. The lower bound is compatible with an approximation of the critical point λ_c for the same parameter set (see Fig. 6). The upper bound, however, does not align with simulations, where oscillations continue to persist beyond this value.

We also aim to understand what happens when we vary the parameters δ and α . Since the steady-state values y_∞ and z_∞ inherently depend on these parameters, we will analyze the limiting behavior of Eq. (7) derived from the NIMFA autocorrelation as δ and α vary. If y_∞ and z_∞ are well defined

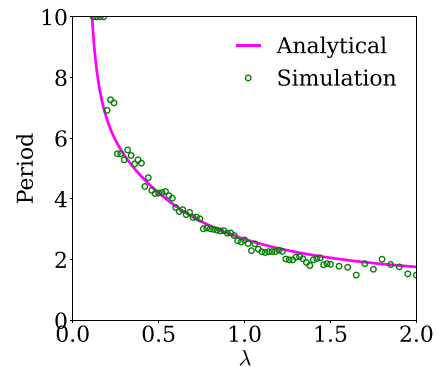


FIG. 5. Period of the autocorrelation as a function of lags at different rates (λ). The analytical approximation (pink line) is computed directly from the equations. The green circles are for the simulations computed with Lomb-Scargle periodogram [22,23]. Parameters: $\langle k \rangle = 10$, $N = 1000$, $\alpha = 0.5$, and $\delta = 1$.

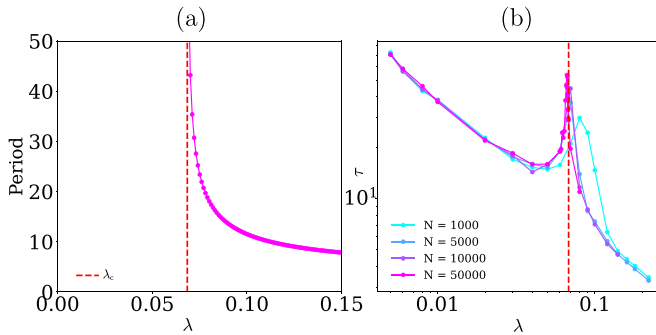


FIG. 6. (a) Analytical approximation of the period of the autocorrelation as a function of the spreading rate (λ). The red dashed line indicates the lower bound of the existence of oscillations. (b) Lifetime, the red dashed line (lower bound of oscillatory behavior) is compatible with an approximation of λ_c . The parameters are $N = 1000$, $k = 10$, $\alpha = 0.5$, and $\delta = 1$.

in the limit $\delta = 0$, then Eq. (7) yields purely real eigenvalues, as the square root reduces to a perfect square. As δ increases, the equation similarly shows that the eigenvalues can remain real if δ is sufficiently large, eliminating oscillatory behavior.

To study this behavior quantitatively, we substitute the steady-state values y_∞ and z_∞ , obtained from the fixed point analysis Eq. (9) into Eq. (7). In particular, the NIMFA predicts that the critical transmissibility λ_c increases with the forgetting rate δ , indicating that faster forgetting makes spreading mode difficult. However, simulations reveal the opposite: Increasing δ lowers λ_c , although the period shrinks and oscillations become noisier. This discrepancy arises because in the real dynamics, stiflers accumulate around spreaders and inhibit the transmission of the rumor; a large δ removes these nodes more quickly and thus facilitates spreading. The NIMFA does not account for such nearest-neighbor correlations and it cannot capture this mechanism.

We also analyze the effect on varying the parameter α , which governs how quickly spreaders become stiflers. Larger values of α lead to faster stifling and consequently, higher values of λ_c are required for the rumor to persist. In this case, both the mean-field predictions and simulations show qualitatively consistent trends (λ_c increases with α), although the numerical values differ. A more detailed analysis can be found in Appendix C 2.

Overall, both simulations and mean-field analysis show that oscillations arise in the active metastable regime, which appears above the critical spreading rate λ_c . By examining the steady-state solutions of the mean-field equations, we determine the conditions for oscillatory behavior and identify the parameter ranges that support it. Although the NIMFA autocorrelation correctly predicts the critical point, it deviates from simulations by failing to account for local interaction effects.

5. Case study

High-impact events in online platforms often give rise to viral dynamics through multiple and coexisting mechanisms. In particular, bursts of attention may result from external amplification, such as mass media exposure, or from changes in the internal dynamics that enhance transmissibility. In the

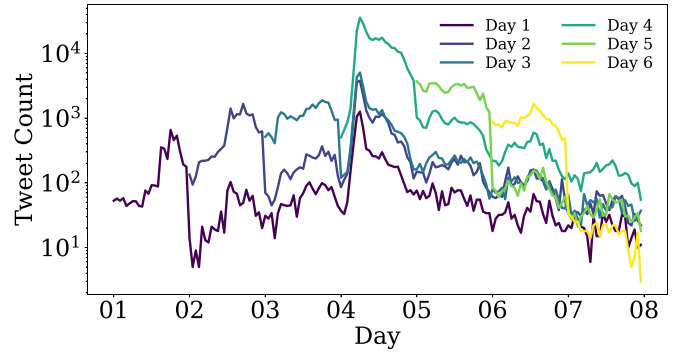


FIG. 7. Tweet count time series for each group of daily new users from the Higgs boson discovery Twitter dataset.

context of our framework, these mechanisms correspond to distinct types of virality: exogenous virality driven by external shocks and endogenous virality emerging from the system's internal dynamics. As in the synthetic scenarios considered above, real-world data typically combine both effects, making it challenging to disentangle them based solely on activity volume. Here, we apply our method to a real-world dataset consisting of Twitter activity surrounding the announcement of the Higgs boson discovery, collected between July 1 and July 7, 2012 [24]. This dataset provides a valuable opportunity to study the emergence of viral behavior during a high-impact event characterized by an external shock.

Our goal is to determine whether the internal dynamics of rumor transmission changed following the official announcement of the discovery on July 4th. Specifically, we aim to infer whether the observed increase in activity reflects a genuine increase in transmissibility among users, indicative of endogenous virality, or whether it was driven by an influx of new participants attracted by widespread media coverage, corresponding to exogenous virality. In other words, we aim to distinguish between an intrinsic change in the rumor spread and an increase in volume caused by external user intrusion.

We observe that users who joined the conversation in the early days continued tweeting throughout the week, and that the announcement on July 4th led to both a sharp increase in activity and a large influx of new users (see Fig. 7).

To analyze the data within the methodological framework introduced above, we segmented the new spreaders time series using a breakpoints detection method [25], allowing us to isolate the system's behavior before and after the announcement. For each segment, we calculated the autocorrelation at different lags (Fig. 8). Notably, after the peak in activity, the oscillation period in the autocorrelation signal shortens, a pattern which according to our analysis is compatible with an increase in transmissibility. This suggests that, beyond the external amplification from media coverage, the internal spread of the rumor on Twitter became more intense.

From an applied perspective, this approach provides a form of nowcasting, enabling the inference of real-time shifts in the system's internal state from limited data. In practice, observing only half of an oscillation period is sufficient to estimate changes in transmissibility. Remarkably, our results align with those obtained from more computationally intensive Monte

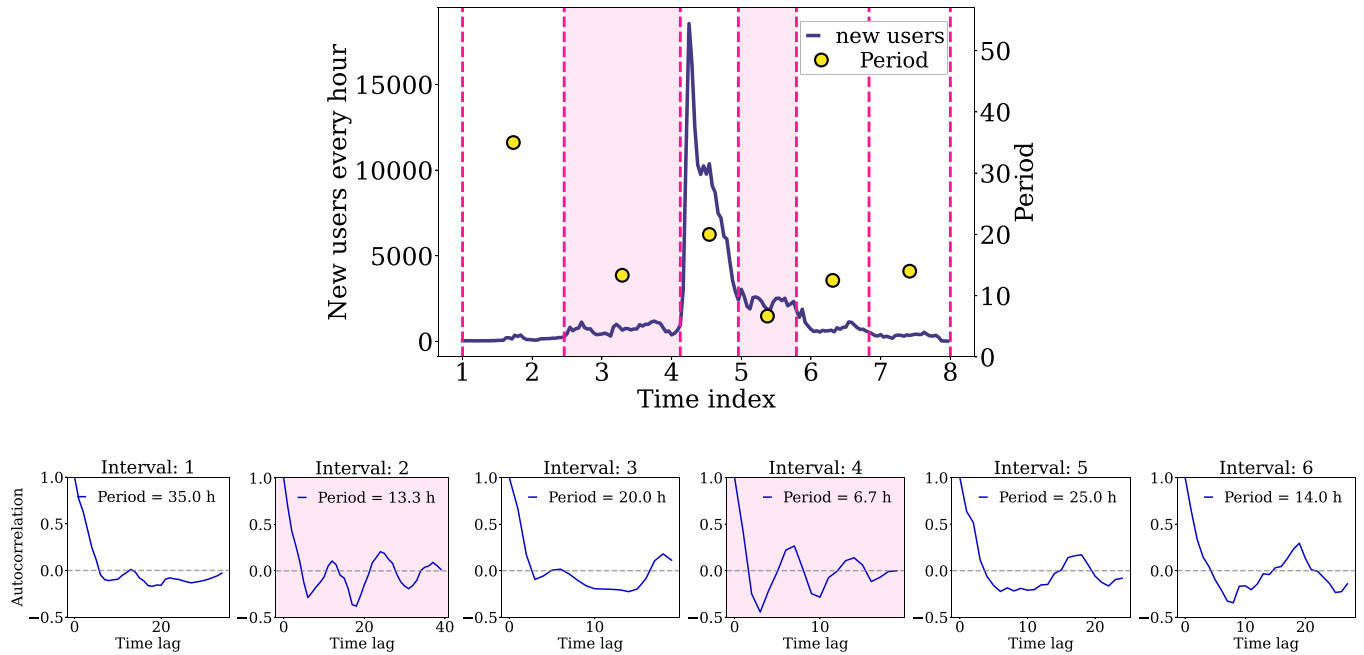


FIG. 8. Analysis of time series of the Twitter activity surrounding the Higgs boson discovery. The main plot shows the time series of new spreaders every hour, having a peak on the fourth day, containing also the statistical breakpoints. The smaller plots are the autocorrelation computed from the time series for each time series breakpoints segment. Then, the period of the autocorrelation (right axis) oscillations is plotted in circles at the main plot. The analyzed segments before and after the peak are highlighted in red.

Carlo simulations on the same dataset [26], but our approach is significantly simpler and more efficient to implement.

B. Control strategies

In this final section, we explore control strategies of rumor propagation. In practical settings, such as mitigating disinformation or prolonging beneficial awareness campaigns, being able to control how long a rumor persists in a network is highly valuable.

Based on the subcritical regime results from Ref. [12], we investigate whether the targeted placement of spreaders can modify the lifetime of a rumor, defined as the average time it takes for a finite simulation starting from a single spreader to reach the absorbing (rumor-free) state.

1. Extending lifetime

First, we consider a simple intervention: Adding a new spreader at the beginning of the simulation, placed at random distance from the seed, as shown in Fig. 9(a). This minimal change leads to a substantial increase in the average lifetime of the rumor. In the subcritical regime, where spreading is scarce and localized, each spreader initiates an independent cascade, and the system only reaches the absorbing state once both cascades have terminated. Therefore, the total lifetime can be approximated as the maximum time of two independent exponentials:

$$T_{\text{independent spreaders}} \approx \frac{3}{2} T_{\text{single spreader}}. \quad (10)$$

This behavior is supported by the analytical lifetime approximation developed in Ref. [12], which considers the minimal set of transitions required to reach the absorbing

state. For a single initial spreader at node i , the shortest extinction sequence is as follows::

- (1) Node i spreads the rumor to a neighbor j , $\langle \tau_1 \rangle$.

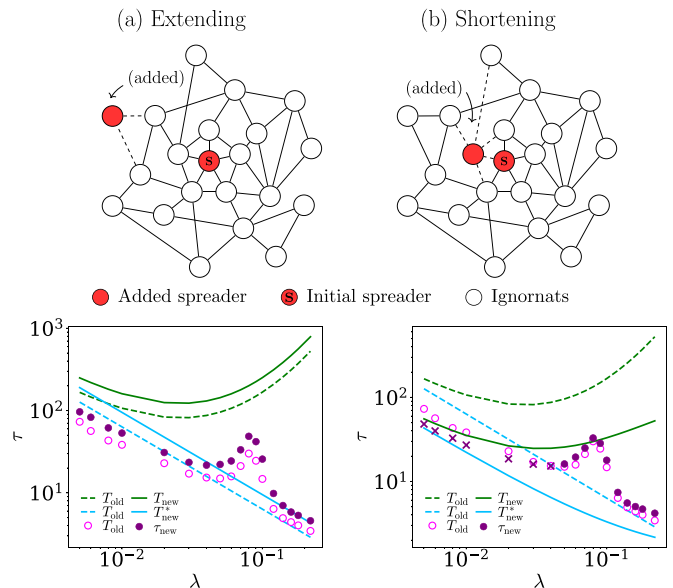


FIG. 9. Network representation (top) and lifetime results (bottom) for the two intervention scenarios: (a) addition of an external spreader far from the initial spreader and (b) addition of an external spreader close to the initial spreader. The initial spreader is denoted by s . Filled (empty) circles indicate the lifetime with (without) intervention. Solid (dashed) lines show the corresponding analytical approximations. Crosses mark cases where intervention reduces the lifetime. Parameters: $\alpha = 0.5$, $\delta = 1$, $N = 1000$, and $\langle k \rangle = 10$.

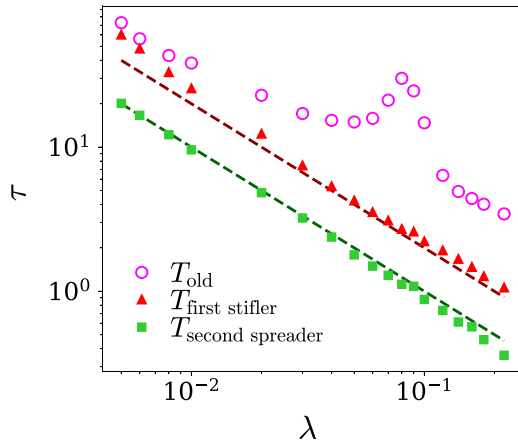


FIG. 10. Average temporal observables of the process starting from a single spreader. Circles indicate the lifetime. Triangles denote the average time of appearance of the first stifier, and squares the average time of appearance of the second spreader. Dashed lines represent the corresponding analytical approximations. Parameters: $\alpha = 0.5$, $\delta = 1$, $N = 1000$, and $\langle k \rangle = 10$.

- (2) One of the two nodes (i or j) becomes a stifier, $\langle \tau_2 \rangle$.
- (3) The second node becomes a stifier, $\langle \tau_3 \rangle$.
- (4) A stifier forgets the rumor and becomes ignorant, $\langle \tau_4 \rangle$.
- (5) The remaining node also forgets the rumor, $\langle \tau_5 \rangle$.

Since transitions are governed by the continuous-time Markov chain, the total lifetime can be approximated by counting the number of times the process fails to reach the absorbing state plus the time it succeeds [12]:

$$T = \sum_{i=1}^{\infty} i \langle \tau_1 \rangle (1 - P_{\text{abs}})^i + P_{\text{abs}} \langle \tau_{1 \rightarrow 5} \rangle \approx \langle T_{\text{abs}} \rangle. \quad (11)$$

As explained in Ref. [12], in the subcritical regime the time that dominates is that of spreading events. Knowing this, we can know with precision the time it will take for the second spreader and the first stifier to appear, as seen in Fig. 10. In contrast with the previous results, we note that this simple analytical approximation of the phenomenon fits perfectly with the simulations. Although we have not designed any intervention for this case, we believe that it may be useful for further research. On average, the time of the appearance is

$$\langle \tau_{\text{second spreader}} \rangle = \frac{1}{k\lambda} \quad \text{and} \quad \langle \tau_{\text{first stifier}} \rangle = \frac{2}{k\lambda}. \quad (12)$$

2. Shortening lifetime

While extending the lifetime of a rumor may be desirable in cases like public awareness campaigns, there are many contexts (such as combating disinformation) where the goal is the opposite: to shorten the rumor's duration and minimize its impact. We now explore whether the lifetime of a rumor can be reduced through targeted interventions.

To this end, we introduce a second spreader in close proximity to the seed node. Specifically, we place the new spreader in a neighboring node before the first spreading event occurs. In the subcritical regime, where spreading is already inefficient, this small change has a significant effect. Because the two spreaders are so close, they are likely to interact early.

This increases the likelihood of early extinction, effectively bypassing the long transient dynamics characteristic of the subcritical regime conditions.

By placing two spreaders close together, we effectively increase the chance that steps (2) and (3) from the events sequence occur earlier, leading to early extinction of the rumor. Simulations confirm that this simple intervention can significantly shorten the expected lifetime in the subcritical regime [see Fig. 9(b)]. Analytically, we approximate it by adding a probability that the two spreaders interact and on average the process finishes earlier. The full derivation can be found in Appendix D. The time to reach the absorbing state for $\lambda \ll 1$ and $\lambda \ll \alpha, \delta$ regime can be approximated as

$$T^* = \frac{1}{2(\langle k \rangle - 1)\lambda} \times \frac{1 - \left(\frac{1}{2\langle k \rangle} + \frac{\alpha}{\alpha + \delta} \right)}{\left(\frac{1}{2\langle k \rangle} + \frac{\alpha}{\alpha + \delta} \right)^2} + \left(\frac{1}{2\langle k \rangle} + \frac{\alpha}{\alpha + \delta} \right) \times \left(\frac{1}{\alpha + \delta} + \frac{1}{2\alpha} + \frac{3}{2\delta} \right). \quad (13)$$

These results highlight how proximity-based interventions can be strategically employed to modify rumor lifetime, based on a lightweight and scalable method.

IV. DISCUSSION

This study advances our understanding of rumor spreading in complex networks and provides practical tools that are applicable to real-world social systems. We show that early warning signals, such as autocorrelation, can be used not only to detect the onset of viral behavior in rumors but also to infer shifts in the underlying transmission dynamics. Building on this, we propose strategies to either extend or shorten the lifetime of a rumor, enabling more effective amplification or suppression depending on the application.

We explored whether a sudden burst in activity caused by external events, such as media broadcasts or astroturfing campaigns, can be distinguished from genuine increases in internal transmissibility. By analyzing the temporal evolution of the lag-1 autocorrelation, computed as a rolling average over time, we find that different spreading mechanisms leave distinct statistical signatures. External spreaders introduced near active nodes lead to an increase in autocorrelation, potentially mimicking the effect of higher transmissibility, and the temporal pattern depends on whether these spreaders remain permanently active or follow the same dynamics as others. In the latter case, the autocorrelation decays more slowly after the perturbation.

These results are especially of interest from a practical perspective. In fact, previously proposed astroturfing detection methods rely on the analysis of interaction networks, linguistic content, and user metadata [4,27], being computationally and data intensive, our results allow early detection methods based on the statistical analysis of the new spreader time series only.

While temporal lag-1 autocorrelation alone may be insufficient to distinguish these scenarios, analyzing the autocorrelation as a function of the lag reveals a crucial signal: Only internally driven growth produces oscillatory patterns across lags. These oscillations arise when the system enters a

metastable regime above the critical transmissibility threshold λ_c . Using the NIMFA, we derived an analytical expression for this behavior as a damped oscillator when the eigenvalues of the infinitesimal generator matrix are complex. Although oscillations are strongly damped in the analytical model, the predicted period aligns closely with simulations, allowing analytical estimation of the critical point and establishing autocorrelation as a robust indicator for identifying spreading mechanisms.

The emergence of oscillations connects naturally to nonequilibrium statistical mechanics. It is well established that oscillations cannot occur in equilibrium steady-state systems. Oscillations require the system to be driven away from equilibrium, which in stochastic terms corresponds to a nonzero net flux (a measure of nonequilibrium strength) that breaks the detailed balance. For systems with three or more states, theory predicts that at a critical flux value a phase transition can occur, producing oscillatory overshoots [28,29]. From a Markov chain perspective, this transition corresponds to the generator matrix acquiring complex conjugate eigenvalues with negative real parts.

Returning to our model, we observe notable discrepancies between the NIMFA predictions and simulations when varying parameters such as the forgetting rate δ . While the NIMFA suggests that higher δ hinders spreading, simulations reveal the opposite. This mismatch stems from the NIMFA's neglect of local correlations, such as stifler clustering around spreaders, which can block further transmission. A higher forgetting rate reduces these local obstructions, indirectly facilitating spread.

To validate our approach, we applied it to Twitter activity surrounding the July 2012 announcement of the Higgs boson discovery. Our aim was to determine whether the postannouncement spike was driven solely by media amplification or also by an increase in internal transmissibility. Measuring autocorrelation at multiple lags before and after the peak, we found a higher frequency of oscillations afterward, consistent with an internal transmissibility increase. This result demonstrates the method's applicability to real-world data.

We further demonstrate that targeted interventions can substantially alter a rumor's duration. These interventions involve introducing external spreaders at specific network locations. A rumor lasts longer when the new spreader is placed far from existing ones, as spatial separation leads to the formation of independent spreading processes that interfere less with each other. Conversely, placing a spreader near an existing one shortens the lifetime in the subcritical regime. This accelerates the transitions from spreader to stifler states, where the rumor is slow and reaches few nodes. As part of this framework, we also estimate the timing of key events, such as the appearance of a second spreader or a stifler. Interestingly, extending a rumor's lifetime is generally easier than shortening it, an insight that aligns with intuitive expectations. These results extend and complement previous ones [16] by elucidating the role of the modified expected lifetime of the rumor when new spreaders are introduced in the populations and their locations.

In summary, we show that autocorrelation analysis across multiple lags can reliably distinguish internally driven spreading from externally introduced activity. Oscillatory patterns are unique to endogenous growth and can be predicted analytically,

with the mean-field approximation accurately capturing both the critical threshold and the oscillation period for certain parameter ranges, even though its performance declines when local correlations become dominant. Also, we demonstrated that the lifetime of a rumor can be controlled through targeted placement of spreaders within a network.

These findings contribute not only to the theoretical modeling of rumor dynamics but also have broader implications for online communication environments. By identifying how rumor volume and visibility can be artificially amplified, our findings help explain how “false volume” can be created in online platforms. In doing so, we highlight mechanisms that are often employed by powerful actors, and facilitated by large technology platforms, to manipulate online discourse using fake accounts and coordinated activity. This challenges the notion of social networks as neutral spaces for information exchange and underscores the importance of developing tools to detect and mitigate such manipulation.

ACKNOWLEDGMENTS

This work was financially supported by the Catalan Government through the funding grant ACCIÓ-Eurecat (Project TRAÇA2023—CLINT). E.R. is a fellow of Eurecat's “Vicente Lopez” Ph.D. grant program, E.C. acknowledges support from the Spanish Grant No. PID2021-128005NB-C22, funded by MCIN/AEL 10.13039/501100011033, and from Generalitat de Catalunya under Project No. 2021-SGR-00856.

E.R., J.V., and E.C. conceptualized the study, E.R. performed data analyses and wrote the original draft, and all authors critically discussed the results, revised the paper, and approved the final manuscript.

DATA AVAILABILITY

The data that support the findings of this article are openly available at [30]. Additional simulation data and extended analysis scripts are available from the authors upon reasonable request.

APPENDIX A: METHODS

1. Networks

Synthetic networks were generated in Python using NetworkX's [31] functions: `random_regular_graph(d,n)` for the random regular networks [32], `powerlaw_cluster_graph(n,m,p)` for the clustered scale-free [33], and `erdos_renyi_graph(n,p_er)` for the Erdős-Rényi graph [34].

2. Monte Carlo simulations

We use an optimized Gillespie algorithm [35] to simulate the exact stochastic trajectories of the continuous-time Markov process [36].

In order to collect significant statistics in the subcritical regime where the system would rapidly reach the absorbing state and to approximate the steady-state behavior, we apply the quasistationary method [37]. With it, we prevent the

system from being absorbed by reintroducing a saved active configuration when the absorption state is reached.

3. Lifetime

To compute the lifetime of finite realizations, each run is initialized with a single randomly chosen spreader and the rest of the population as ignorants. The simulation proceeds until the system reaches an absorbing state [18].

4. Early warning signal indicators

a. Mutual information

The MI quantifies the amount of information that one random variable contains about another, measuring the statistical dependence between them and capturing linear and non-linear relationships. It is defined as [14]

$$M_{X,Y} = \sum_{i,j} p_{X,Y}(i,j) \log \left(\frac{p_{X,Y}(i,j)}{p_X(i)p_Y(j)} \right), \quad (\text{A1})$$

where $p_{X,Y}(i,j)$ is the joint probability distribution of variables X and Y , and $p_X(i)$ and $p_Y(j)$ are their corresponding marginal distributions. We have used `mutual_info_score` from Python library `scikit-learn` [38]. Specifically, we have computed the mutual information over symbolic time series on a rolling time window of 50 steps picking up two nodes at a random distance. Then, we have computed the average of 1000 realizations for each time step.

b. Autocorrelation

Autocorrelation measures the correlation of a signal with a delayed copy of itself. To compute it, we have used function `statsmodels.tsa.stattools.acf` from Python library `statsmodels` [39]. Similar to the mutual information, we have computed it with a rolling window of 50 steps, from the spreaders density time series, then averaged over 1000 realizations.

APPENDIX B: AUTOCORRELATION FOR DIFFERENT NETWORK STRUCTURES

In this Appendix, we analyze the autocorrelation of the density of spreaders when the transmissibility rate λ is high, across different network structures. We generated a clustered power-law network (degree exponent ≈ 3) whose maximum degree was 196 and whose average degree was 9.932, and an Erdős-Rényi random graph with a maximum degree of 21 and an average degree of 9.972. In Fig. 11, we plotted the autocorrelation as a function of lag and we observe clear oscillations.

APPENDIX C: THE NIMFA AUTOCORRELATION IN THE STEADY STATE

1. Analytical derivation

We derive the autocorrelation function for the spreading state of a node in the steady state of the Maki-Thompson modified model using the NIMFA [20,21].

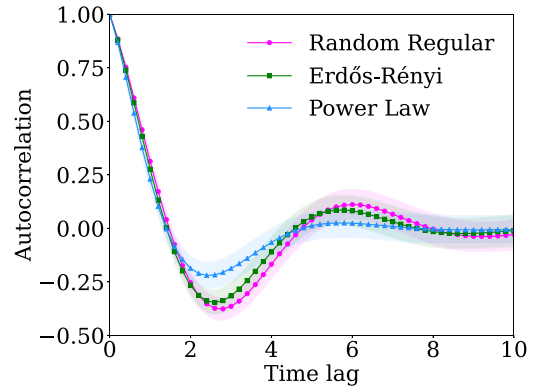


FIG. 11. Autocorrelation of spreaders density as a function of lag for a high spreading rate $\lambda = 0.3$ for an Erdős-Rényi and a power-law network ($\gamma \approx 3$) with $\langle k \rangle \approx 10$. The period of the oscillations is similar, the Erdős-Rényi has a similar behavior with the random regular.

The autocorrelation of the spreading state $V_j(t) \in \{0, 1\}$ of a node j between times s and t is defined as

$$R_j(s, t) \triangleq \frac{E[V_j(s)V_j(t)] - E[V_j(s)]E[V_j(t)]}{\sqrt{\text{Var}[V_j(s)]\text{Var}[V_j(t)]}}. \quad (\text{C1})$$

In the steady state ($t \rightarrow \infty$), we denote the time difference as lag h and write $R_{j\infty}(h) \equiv R_j(t, t+h)$.

Using the NIMFA, we can define effective transmission and stifling rates. For node i , the steady-state effective rates are

$$\tilde{\lambda}_i \equiv \lambda \sum_{k=1}^N A_{ki} y_{i\infty} \quad \text{and} \quad \tilde{\alpha}_i \equiv \alpha \sum_{k=1}^N A_{ki} (y_{k\infty} + z_{k\infty}). \quad (\text{C2})$$

These rates define the mean-field evolution of the spreading probability $v_i(t)$, approximated as

$$\frac{dv_i(t)}{dt} = \left[1 - v_i(t) - \frac{1}{\delta} v_i(t) \tilde{\alpha} \right] \tilde{\lambda} - v_i(t) \tilde{\alpha}. \quad (\text{C3})$$

The nontrivial steady-state spreading probability is

$$v_{i\infty} = \frac{\tilde{\lambda}}{\tilde{\lambda} \left(1 + \frac{\tilde{\alpha}}{\delta} \right) + \tilde{\alpha}}. \quad (\text{C4})$$

Additionally, one can bound this by assuming a maximum neighborhood degree d_i :

$$0 \leq v_{i\infty} \leq \frac{\lambda}{\lambda \left(1 + \frac{d_i}{\delta} \right) + \alpha}. \quad (\text{C5})$$

We now construct the infinitesimal generator Q_i for the Markov process governing the state transitions of node i :

$$Q_{i\infty} = \begin{pmatrix} -\tilde{\lambda}_i & \tilde{\lambda}_i & 0 \\ 0 & -\tilde{\alpha}_i & \tilde{\alpha}_i \\ \delta & 0 & -\delta \end{pmatrix}, \quad (\text{C6})$$

this defines the transitions between ignorant, spreader, and stifier states. The transition matrix at lag h is

$$P_{i\infty}(h) = e^{Q_{i\infty}h} = U_i e^{\Lambda_i} U_i^{-1}, \quad (\text{C7})$$

where U_i and Λ_i are the eigenvector matrix and the diagonal eigenvalue matrix of $Q_{i\infty}h$, respectively. The eigenvalues of $Q_{i\infty}$ are

$$\mu_{2,3} = -\frac{1}{2}(\delta + \tilde{\lambda} + \tilde{\alpha}) \pm \frac{1}{2}\sqrt{(\delta + \tilde{\lambda} + \tilde{\alpha})^2 - 4(\tilde{\lambda}\delta + \tilde{\alpha}\delta + \tilde{\lambda}\tilde{\alpha})}. \quad (C8)$$

Let us define

$$a = -\frac{1}{2}(\delta + \tilde{\lambda} + \tilde{\alpha}), \quad w = \frac{1}{2}\sqrt{(\delta + \tilde{\lambda} + \tilde{\alpha})^2 - 4(\tilde{\lambda}\delta + \tilde{\alpha}\delta + \tilde{\lambda}\tilde{\alpha})}. \quad (C9)$$

We focus on the transition probability of the remaining spreader

$$P_{i\infty}(h)_{2,2} = \Pr[V_{i\infty}(t+h) = 1 | V_{i\infty}(t) = 1]. \quad (C10)$$

After simplifying, we get the expression

$$P_{i\infty}(h)_{2,2} = \frac{1}{\tilde{\alpha}\tilde{\lambda} + \tilde{\alpha}\delta + \tilde{\lambda}\delta} \left[\delta\tilde{\lambda} + \frac{\tilde{\alpha}}{2}(\tilde{\lambda} + \delta)(e^{\mu_2 h} + e^{\mu_3 h}) + \frac{\tilde{\alpha}(\tilde{\alpha}\tilde{\lambda} - \tilde{\lambda}^2 + \tilde{\alpha}\delta - \delta^2)}{2w}(e^{\mu_2 h} - e^{\mu_3 h}) \right]. \quad (C11)$$

Since in the steady state $E[V_{i\infty}(t)] = v_{i\infty}$, the autocorrelation is

$$R_{i\infty}(h) = \frac{P_{i\infty}(h)_{2,2} - v_{i\infty}}{1 - v_{i\infty}}. \quad (C12)$$

Substituting $P_{i\infty}(h)_{2,2}$, we obtain the expression for the autocorrelation

$$R_{i\infty}(h) = \frac{1}{1 - v_{i\infty}} \left[\frac{1}{\tilde{\alpha}\tilde{\lambda} + \tilde{\alpha}\delta + \tilde{\lambda}\delta} \left[\delta\tilde{\lambda} + \frac{\tilde{\alpha}}{2}(\tilde{\lambda} + \delta)(e^{\mu_2 h} + e^{\mu_3 h}) + \frac{\tilde{\alpha}(\tilde{\alpha}\tilde{\lambda} - \tilde{\lambda}^2 + \tilde{\alpha}\delta - \delta^2)}{2w}(e^{\mu_2 h} - e^{\mu_3 h}) \right] - v_{i\infty} \right]. \quad (C13)$$

When the eigenvalues ($\mu_{2,3}$) are complex, we get the closed-form expression

$$R_i = \frac{1}{1 - v_{i\infty}} \left[\frac{1}{\tilde{\alpha}\tilde{\lambda} + \tilde{\alpha}\delta + \tilde{\lambda}\delta} [\delta\tilde{\lambda} + Ae^{ah}(\cos wh - \phi)] - v_{i\infty} \right],$$

where

$$A = \sqrt{(\tilde{\alpha}(\tilde{\lambda} + \delta))^2 + \left(\frac{\tilde{\alpha}(\tilde{\alpha}\tilde{\lambda} - \tilde{\lambda}^2 + \tilde{\alpha}\delta - \delta^2)}{w} \right)^2}, \quad (C14)$$

$$\phi = \arctan \left(\frac{\tilde{\alpha}(\tilde{\alpha}\tilde{\lambda} - \tilde{\lambda}^2 + \tilde{\alpha}\delta - \delta^2)}{w\tilde{\alpha}(\tilde{\lambda} + \delta)} \right), \quad (C15)$$

which corresponds to a damped oscillator.

2. Conditions for the emergence of oscillations

The autocorrelation shows oscillations when the eigenvalues of the infinitesimal generator are complex. We can get a full expression if we substitute the steady-state values from Eq. (9) into Eq. (7):

$$\omega^2 = \left(\delta + \frac{\lambda^2 k \delta}{\alpha \lambda k + \alpha \delta + \lambda \delta} + \frac{\alpha k \lambda}{\alpha + \lambda} \right)^2 - 4 \left(\frac{\lambda^2 k \delta}{\alpha \lambda k + \alpha \delta + \lambda \delta} \delta + \frac{\alpha k \lambda}{\alpha + \lambda} \delta + \frac{\alpha k \lambda}{\alpha + \lambda} \frac{\lambda^2 k \delta}{\alpha \lambda k + \alpha \delta + \lambda \delta} \right). \quad (C16)$$

The critical transmissibility λ_c can be found with the following expression:

$$\delta^2 + \left(\frac{\lambda^2 k \delta}{\alpha \lambda k + \alpha \delta + \lambda \delta} \right)^2 + \left(\frac{\alpha k \lambda}{\alpha + \lambda} \right)^2 = -2 \left(\frac{\lambda^2 k \delta}{\alpha \lambda k + \alpha \delta + \lambda \delta} \delta + \frac{\alpha k \lambda}{\alpha + \lambda} \delta + \frac{\alpha k \lambda}{\alpha + \lambda} \frac{\lambda^2 k \delta}{\alpha \lambda k + \alpha \delta + \lambda \delta} \right). \quad (C17)$$

Since we cannot get a closed expression for λ_c , we compute the limits from when Eq. (7) is complex. For the case $\lambda \ll \delta, \alpha$:

$$w^2 \underset{\lambda \ll \delta, \alpha}{\approx} \delta - 2k\lambda\delta + \lambda^2 \left(k^2 - 2\frac{k}{\alpha}\delta \right) + \mathcal{O}(\lambda^3) < 0, \quad (C18)$$

the minimum value for λ at second order is

$$\lambda^* = \frac{2k\delta - \sqrt{\frac{8k\delta^3}{\alpha}}}{2(k^2 - \frac{2k\delta}{\alpha})}. \quad (C19)$$

From this, we can see that λ_c will grow as δ becomes larger. In the same way, λ_c will grow as α becomes larger. We

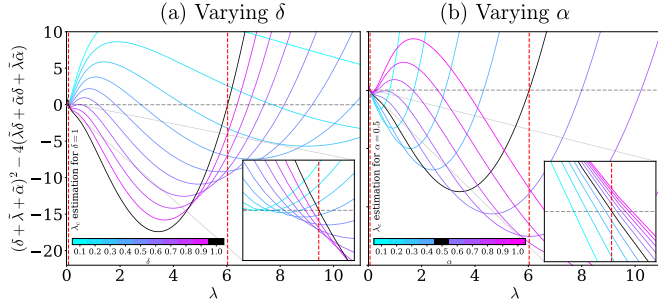


FIG. 12. (a) Comparison of the condition for the oscillations, varying δ from 0.1 to 1. When the expression is negative, the eigenvalue is complex and the oscillations appear. In black, the results with $\delta = 1$. (b) Comparison of the condition for the oscillations, varying α from 0.1 to 1. When the expression is negative, the eigenvalue is complex, and the oscillations appear. In black, the results with $\alpha = 0.5$.

can observe this in Fig. 12, where we plot the condition that makes the oscillations appear as a function of λ , for different δ [Fig. 12(a)] and for different α [Fig. 12(b)]. The oscillations start to appear when the function is negative, so λ_c is the first time the function crosses 0. However, if we compare these results with the simulations, we observe important differences. We know where is λ_c because it is when the lifetime diverges. In Fig. 13(a), we plotted the lifetime as a function of rate λ for different values of δ . We see that the value of λ_c decreases as the parameter δ increases. This is in contradiction with the NIMFA results. For parameter α [Fig. 13(b)], we see that the result qualitatively matches the behavior for the second-order approximation: As α increases, λ_c increases. Finally, we plot the period of the oscillations for different parameter values α and δ . In Fig. 14(a), we plotted the oscillations period of the simulations for different values of δ . We see that as δ increases, the oscillations period decreases, and therefore, it is compatible that oscillations start to appear earlier for larger δ , as we can know by looking at the lifetime. In Fig. 14(b), we can see that the period of the oscillations decreases as the parameter α decreases, making also compatible with the lifetime results.

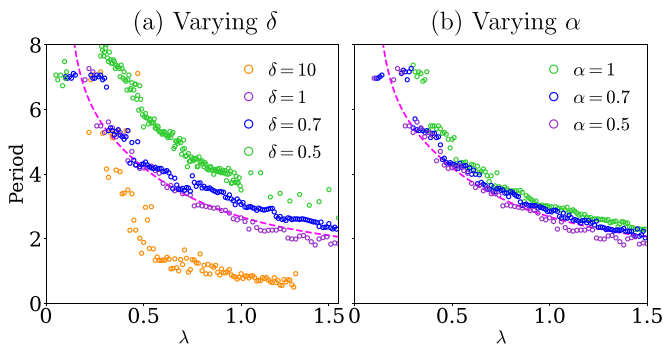


FIG. 13. Average lifetime as a function of rates λ : (a) for different values of $\delta \in \{0.5, 0.7, 1\}$, with $\alpha = 0.5$, and (b) for different values of $\alpha \in \{0.5, 0.7, 1\}$, with $\delta = 1$. The horizontal lines indicate the approximate values of λ_c , determined from the divergence point of the lifetime.

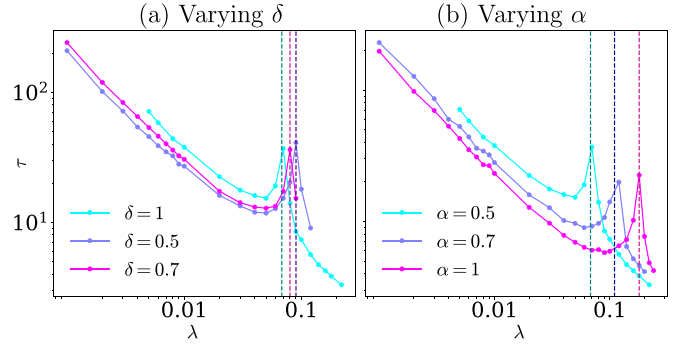


FIG. 14. Period of the autocorrelation as a function of rates λ : (a) for different values of $\delta \in \{0.5, 0.7, 1, 10\}$, with $\alpha = 0.5$, and (b) for different values of $\alpha \in \{0.5, 0.7, 1\}$, with $\delta = 1$. The pink dashed line shows the analytical result for $\delta = 1$, $\alpha = 0.5$, which matches the simulations. For other combinations of parameters, the analytical approximation does not coincide with simulations.

APPENDIX D: APPROXIMATION OF THE LIFETIME FOR THE INTERVENTION

The lifetime can be approximated as follows [12]: Assuming that node i is a spreader, the shortest sequence of events that will drive the process to the absorbing state consists of the following: (1) Node i spreads the information to node j , (2) either i or j turns into stifter, (3) the other node turns into a stifter, (4) either i or j turns into ignorant, and (5) the other node turns into an ignorant. Since our process is a continuous Markov chain, only one event occurs at a time; the respective probabilities P and expected times $\langle \tau \rangle$ of these events are

$$P_1 = 1, \quad \langle \tau_1 \rangle = \frac{1}{\langle k \rangle \lambda}, \quad (D1)$$

$$P_2 = \frac{2\alpha}{2\alpha + 2\lambda(\langle k \rangle - 1)}, \quad \langle \tau_2 \rangle = \frac{1}{2\alpha + 2\lambda(\langle k \rangle - 1)}, \quad (D2)$$

$$P_3 = \frac{\alpha}{\alpha + \lambda(\langle k \rangle - 1) + \delta}, \quad \langle \tau_3 \rangle = \frac{1}{\alpha + \lambda(\langle k \rangle - 1) + \delta}, \quad (D3)$$

$$P_4 = P_5 = 1, \quad 2\langle \tau_4 \rangle = \langle \tau_5 \rangle = \frac{1}{\delta}. \quad (D4)$$

The probability and the expected time for reaching the absorbing state through this chain are

$$P_{\text{abs}} = \frac{\alpha^2}{(\alpha + (\langle k \rangle - 1)\lambda)(\alpha + \delta + (\langle k \rangle - 1)\lambda)}, \quad (D5)$$

$$\langle \tau_{1 \rightarrow 5} \rangle = \frac{1}{\langle k \rangle \lambda} + \frac{3}{2\delta} + \frac{3}{2\alpha + 2\lambda(\langle k \rangle - 1)}. \quad (D6)$$

In the $\lambda \ll \delta$ and $\lambda \ll \alpha$ regime, $\langle \tau_1 \rangle$ is the largest time:

$$T = \sum_{i=1}^{\infty} i \langle \tau_i \rangle (1 - P_{\text{abs}})^i + P_{\text{abs}} \langle \tau_{1 \rightarrow 5} \rangle \approx \langle T_{\text{abs}} \rangle. \quad (D7)$$

For the case the lifetime is reduced (see Sec. III B 2) in the subcritical regime, the equations are

$$T = \langle \tau_1 \rangle \frac{1 - P_{\text{abs}}}{(P_{\text{abs}})^2} + P_{\text{dir}} \langle \tau_{2 \rightarrow 4} \rangle + P_{\text{ind}} \langle \tau_{1 \rightarrow 4} \rangle, \quad (D8)$$

where

$$P_{\text{abs}} = P_{\text{dir}} + P_{\text{ind}}, \quad (\text{D9})$$

$$P_{\text{ind}} = \frac{\alpha^2}{(\alpha + (\langle k \rangle - 2)\lambda)(\alpha + \delta + (\langle k \rangle - 2)\lambda)}, \quad (\text{D10})$$

$$P_{\text{dir}} = \frac{1}{2(\langle k \rangle - 1)} \frac{\alpha}{\lambda + \alpha}, \quad (\text{D11})$$

$$\langle \tau_1 \rangle = \frac{1}{(\langle k \rangle - 1)\lambda}, \quad (\text{D12})$$

$$\langle \tau_{2 \rightarrow 4} \rangle = \frac{1}{\alpha + \delta + (\langle k \rangle - 2)\lambda} + \frac{1}{2\alpha + 2(\langle k \rangle - 2)\lambda} + \frac{3}{2\delta}. \quad (\text{D13})$$

Finally,

$$\begin{aligned} T = & \frac{1}{2(\langle k \rangle - 1)\lambda} \times \frac{1 - \frac{1}{2(\langle k \rangle - 1)} \frac{\alpha}{\lambda + \alpha} - \frac{\alpha^2}{(\alpha + (\langle k \rangle - 2)\lambda)(\alpha + \delta + (\langle k \rangle - 2)\lambda)}}{\left(\frac{1}{2(\langle k \rangle - 1)} \frac{\alpha}{\lambda + \alpha} + \frac{\alpha^2}{(\alpha + (\langle k \rangle - 2)\lambda)(\alpha + \delta + (\langle k \rangle - 2)\lambda)} \right)^2} + \left(\frac{1}{2\langle k \rangle} \frac{\alpha}{\lambda + \alpha} \right) \\ & \times \left(\frac{1}{\alpha + \delta + (\langle k \rangle - 2)\lambda} + \frac{1}{2\alpha + 2(\langle k \rangle - 2)\lambda} + \frac{3}{2\delta} \right) + \left(\frac{\alpha^2}{(\alpha + (\langle k \rangle - 2)\lambda)(\alpha + \delta + (\langle k \rangle - 2)\lambda)} \right) \\ & \times \left(\frac{1}{(\langle k \rangle - 1)\lambda} + \frac{1}{\alpha + \delta + (\langle k \rangle - 2)\lambda} + \frac{1}{2\alpha + 2(\langle k \rangle - 2)\lambda} + \frac{3}{2\delta} \right), \end{aligned} \quad (\text{D14})$$

and for the $\lambda \ll 1$ and $\lambda \ll \alpha, \delta$ regime,

$$T^* = \frac{1}{2(\langle k \rangle - 1)\lambda} \times \frac{1 - \left(\frac{1}{2\langle k \rangle} + \frac{\alpha}{\alpha + \delta} \right)}{\left(\frac{1}{2\langle k \rangle} + \frac{\alpha}{\alpha + \delta} \right)^2} + \left(\frac{1}{2\langle k \rangle} + \frac{\alpha}{\alpha + \delta} \right) \times \left(\frac{1}{\alpha + \delta} + \frac{1}{2\alpha} + \frac{3}{2\delta} \right). \quad (\text{D15})$$

-
- [1] Proyecto UNA, *La Viralidad Del Mal* (Editorial Descontrol, Barcelona, 2024).
- [2] S. Arce-García, E. Said-Hung, and D. Mottareale-Calvanese, Types of Astroturfing campaigns with disinformation and polarized content during the pandemic in Spain, *icono14* **21**, (2023).
- [3] C. A. Bail, B. Guay, E. Maloney, A. Combs, D. S. Hillegus, F. Merhout, D. Freelon, and A. Volfovsky, Assessing the Russian internet research agency's impact on the political attitudes and behaviors of American Twitter users in late 2017, *Proc. Natl. Acad. Sci. USA* **117**, 243 (2020).
- [4] F. B. Keller, D. Schoch, S. Stier, and J. Yang, Political astroturfing on Twitter: How to coordinate a disinformation campaign, *Polit. Commun.* **37**, 256 (2020).
- [5] C. Sirikupt, Drowning out dissent: The Thai military's quest to fabricate popular support on Twitter, *Int. J. Press./Polit.* **31**, 207 (2026).
- [6] X. Wang, J. Li, E. Srivatsavaya, and S. Rajtmajer, Evidence of inter-state coordination amongst state-backed information operations, *Sci. Rep.* **13**, 7716 (2023).
- [7] D. Schoch, F. B. Keller, S. Stier, and J. Yang, Coordination patterns reveal online political astroturfing across the world, *Sci. Rep.* **12**, 4572 (2022).
- [8] S. Gelper, M. J. Lovett, and R. Peres, The effect of second screening on repeat viewing: Insights from large-scale mobile diary data, *J. Acad. Mark. Sci.* **53**, 907 (2025).
- [9] W. Brannon and D. Roy, The speed of news in Twitter (X) versus radio, *Sci. Rep.* **14**, 11939 (2024).
- [10] D. P. Maki and M. Thompson, *Mathematical Models and Applications, with Emphasis on the Social, Life, and Management Sciences* (Prentice-Hall, Englewood Cliffs, NJ, 1973).
- [11] A. Barrat, M. Barthélemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, 2008).
- [12] G. Ferraz de Arruda, L. G. S. Jeub, A. S. Mata, F. A. Rodrigues, and Y. Moreno, From subcritical behavior to a correlation-induced transition in rumor models, *Nat. Commun.* **13**, 3049.
- [13] S. M. O'Regan and J. M. Drake, Theory of early warning signals of disease emergence and leading indicators of elimination, *Theor. Ecol.* **6**, 333 (2013).
- [14] S. V. George, S. Kachhara, and G. Ambika, Early warning signals for critical transitions in complex systems, *Phys. Scr.* **98**, 072002 (2023).
- [15] W. Ullon and E. Forgoston, Controlling epidemic extinction using early warning signals, *Int. J. Dyn. Control* **11**, 851 (2023).
- [16] S. Belen, C. Y. Kaya, and C. E. M. Pearce, Impulsive control of rumours with two broadcasts, *ANZIAM J.* **46**, 379 (2005).
- [17] C. Pearce, Y. Kaya, and S. Belen, Impulsive control of a sequence of rumour processes, in *Continuous Optimiza-*

- tion: *Current Trends and Modern Applications*, edited by V. Jeyakumar and A. Rubinov (Springer US, Boston, MA, 2005), pp. 387–407.
- [18] M. Boguñá, C. Castellano, and R. Pastor-Satorras, Nature of the epidemic threshold for the susceptible-infected-susceptible dynamics in networks, *Phys. Rev. Lett.* **111**, 068701 (2013).
- [19] P. Van Mieghem, J. Omic, and R. Kooij, Virus spread in networks, *IEEE/ACM Trans. Networking* **17**, 1 (2009).
- [20] P. Van Mieghem, *Performance Analysis of Complex Networks and Systems* (Cambridge University Press, Cambridge, 2014).
- [21] Q. Liu and P. Van Mieghem, Autocorrelation of the susceptible-infected-susceptible process on networks, *Phys. Rev. E* **97**, 062309 (2018).
- [22] N. R. Lomb, Least-squares frequency analysis of unequally spaced data, *Astrophys. Space Sci.* **39**, 447 (1976).
- [23] J. D. Scargle, Studies in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data, *Astrophys. J.* **263**, 835 (1982).
- [24] M. De Domenico, A. Lima, P. Mougél, and M. Musolesi, The anatomy of a scientific Rumor, *Sci. Rep.* **3**, 2980 (2013).
- [25] C. Truong, L. Oudre, and N. Vayatis, Selective review of offline change point detection methods, *Signal Process.* **167**, 107299 (2020).
- [26] G. Ferraz de Arruda, F. Aparecido Rodrigues, P. Martín Rodríguez, E. Cozzo, and Y. Moreno, A general Markov chain approach for disease and rumour spreading in complex networks, *J. Complex Networks* **6**, 215 (2018).
- [27] J. Ratkiewicz, M. Conover, M. Meiss, B. Gonçalves, S. Patil, A. Flammini, and F. Menczer, Truthy: Mapping the spread of astroturf in microblog streams, in *Proceedings of the 20th International Conference Companion on World Wide Web, WWW'11* (Association for Computing Machinery, New York, NY, USA, 2011), pp. 249–252.
- [28] C. Jia, M. Qian, and D. Jiang, Overshoot in biological systems modelled by Markov chains: A non-equilibrium dynamic phenomenon, *IET Syst. Biol.* **8**, 138 (2014).
- [29] C. Jia and M. Qian, Nonequilibrium enhances adaptation efficiency of stochastic biochemical systems, *PLoS One* **11**, 1 (2016).
- [30] E. Rifà, ruscs: Rumor Spreading Complex Networks Simulations, Zenodo (2026), <https://doi.org/10.5281/zenodo.18713674>.
- [31] A. A. Hagberg, D. A. Schult, and P. J. Swart, Exploring Network Structure, Dynamics, and Function using NetworkX, in *Proceedings of the 7th Python in Science Conference*, edited by G. Varoquaux, T. Vaught, and J. Millman (Pasadena, CA, USA, 2008), pp. 11–15.
- [32] A. Steger and N. C. Wormald, Generating random regular graphs quickly, *Comb. Probab. Comput.* **8**, 377 (1999).
- [33] P. Holme and B. J. Kim, Growing scale-free networks with tunable clustering, *Phys. Rev. E* **65**, 026107 (2002).
- [34] P. Erdős and A. Rényi, On random graphs, *Publ. Math.* **6**, 290 (1959).
- [35] N. Gutjahr, P. Hövel, and A. Viol, Controlling extended criticality via modular connectivity, *J. Phys.: Complexity* **2**, 035023 (2021).
- [36] D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, *J. Phys. Chem.* **81**, 2340 (1977).
- [37] M. M. de Oliveira and R. Dickman, How to simulate the quasistationary state, *Phys. Rev. E* **71**, 016129 (2005).
- [38] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, Scikit-learn: Machine learning in Python, *J. Mach. Learn. Res.* **12**, 2825 (2011).
- [39] S. Seabold and J. Perktold, Statsmodels: Econometric and statistical modeling with Python, in *Proceedings of the 9th Python in Science Conference*, edited by S. van der Walt and J. Millman (Austin, TX, 2010), pp. 92–96.