



A LETTERS JOURNAL EXPLORING  
THE FRONTIERS OF PHYSICS

OFFPRINT

**Modeling self-sustained activity cascades  
in socio-technical networks**

P. PIEDRAHITA, J. BORGE-HOLTHOEFER, Y. MORENO  
and A. ARENAS

EPL, **104** (2013) 48004

Please visit the new website  
[www.epljournal.org](http://www.epljournal.org)



A LETTERS JOURNAL EXPLORING  
THE FRONTIERS OF PHYSICS

## AN INVITATION TO SUBMIT YOUR WORK

[www.epljournal.org](http://www.epljournal.org)

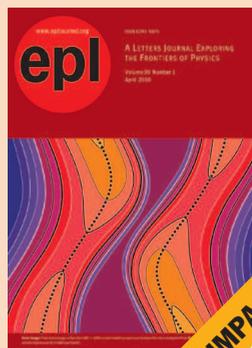
### **The Editorial Board invites you to submit your letters to EPL**

EPL is a leading international journal publishing original, high-quality Letters in all areas of physics, ranging from condensed matter topics and interdisciplinary research to astrophysics, geophysics, plasma and fusion sciences, including those with application potential.

The high profile of the journal combined with the excellent scientific quality of the articles continue to ensure EPL is an essential resource for its worldwide audience. EPL offers authors global visibility and a great opportunity to share their work with others across the whole of the physics community.

### **Run by active scientists, for scientists**

EPL is reviewed by scientists for scientists, to serve and support the international scientific community. The Editorial Board is a team of active research scientists with an expert understanding of the needs of both authors and researchers.



**IMPACT FACTOR**  
**2.753\***  
\* As ranked by ISI 2010

[www.epljournal.org](http://www.epljournal.org)

**IMPACT FACTOR**

**2.753\***

\* As listed in the ISI® 2010 Science  
Citation Index Journal Citation Reports

**OVER**

**500 000**

full text downloads in 2010

**30 DAYS**

average receipt to online  
publication in 2010

**16 961**

citations in 2010  
37% increase from 2007

*"We've had a very positive experience with EPL, and not only on this occasion. The fact that one can identify an appropriate editor, and the editor is an active scientist in the field, makes a huge difference."*

**Dr. Ivar Martin**

Los Alamos National Laboratory,  
USA

**Six good reasons to publish with EPL**

We want to work with you to help gain recognition for your high-quality work through worldwide visibility and high citations.

- 1 Quality** – The 40+ Co-Editors, who are experts in their fields, oversee the entire peer-review process, from selection of the referees to making all final acceptance decisions
- 2 Impact Factor** – The 2010 Impact Factor is 2.753; your work will be in the right place to be cited by your peers
- 3 Speed of processing** – We aim to provide you with a quick and efficient service; the median time from acceptance to online publication is 30 days
- 4 High visibility** – All articles are free to read for 30 days from online publication date
- 5 International reach** – Over 2,000 institutions have access to EPL, enabling your work to be read by your peers in 100 countries
- 6 Open Access** – Articles are offered open access for a one-off author payment

Details on preparing, submitting and tracking the progress of your manuscript from submission to acceptance are available on the EPL submission website [www.epletters.net](http://www.epletters.net).

If you would like further information about our author service or EPL in general, please visit [www.epljournal.org](http://www.epljournal.org) or e-mail us at [info@epljournal.org](mailto:info@epljournal.org).

**EPL is published in partnership with:**



European Physical Society



Società Italiana  
di Fisica



EDP Sciences

**IOP Publishing**

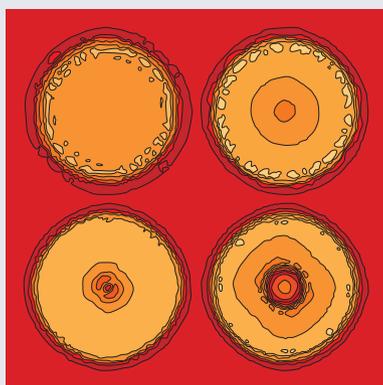
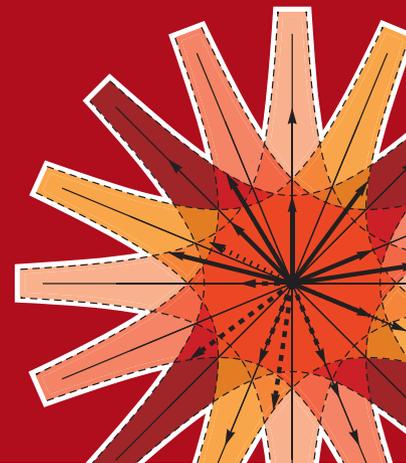
IOP Publishing



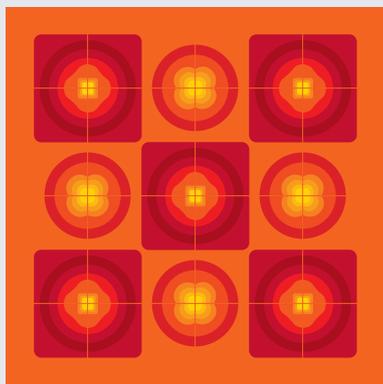
A LETTERS JOURNAL  
EXPLORING THE FRONTIERS  
OF PHYSICS

**EPL Compilation Index**

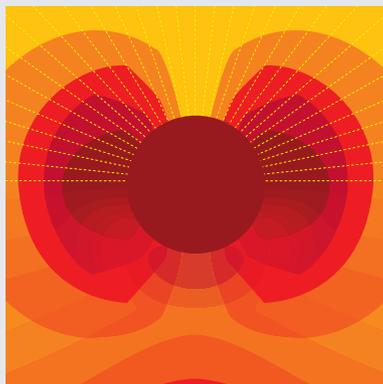
[www.epljournal.org](http://www.epljournal.org)



Biaxial strain on lens-shaped quantum rings of different inner radii, adapted from **Zhang et al** 2008 *EPL* **83** 67004.



Artistic impression of electrostatic particle-particle interactions in dielectrophoresis, adapted from **N Aubry and P Singh** 2006 *EPL* **74** 623.



Artistic impression of velocity and normal stress profiles around a sphere that moves through a polymer solution, adapted from **R Tuinier, J K G Dhont and T-H Fan** 2006 *EPL* **75** 929.

Visit the EPL website to read the latest articles published in cutting-edge fields of research from across the whole of physics.

Each compilation is led by its own Co-Editor, who is a leading scientist in that field, and who is responsible for overseeing the review process, selecting referees and making publication decisions for every manuscript.

- Graphene
- Liquid Crystals
- High Transition Temperature Superconductors
- Quantum Information Processing & Communication
- Biological & Soft Matter Physics
- Atomic, Molecular & Optical Physics
- Bose-Einstein Condensates & Ultracold Gases
- Metamaterials, Nanostructures & Magnetic Materials
- Mathematical Methods
- Physics of Gases, Plasmas & Electric Fields
- High Energy Nuclear Physics

If you are working on research in any of these areas, the Co-Editors would be delighted to receive your submission. Articles should be submitted via the automated manuscript system at [www.epletters.net](http://www.epletters.net)

If you would like further information about our author service or EPL in general, please visit [www.epljournal.org](http://www.epljournal.org) or e-mail us at [info@epljournal.org](mailto:info@epljournal.org)



**IOP Publishing**

**Image:** Ornamental multiplication of space-time figures of temperature transformation rules (adapted from T. S. Bíró and P. Ván 2010 *EPL* **89** 30001; artistic impression by Frédérique Swist).

# Modeling self-sustained activity cascades in socio-technical networks

P. PIEDRAHITA<sup>1</sup>, J. BORGE-HOLTHOEFER<sup>1</sup>, Y. MORENO<sup>1,2</sup> and A. ARENAS<sup>3,4</sup>

<sup>1</sup> *Instituto de Biocomputación y Física de Sistemas Complejos (BIFI), Universidad de Zaragoza  
Mariano Esquillor s/n, 50018 Zaragoza, Spain*

<sup>2</sup> *Departamento de Física Teórica, Universidad de Zaragoza - 50009 Zaragoza, Spain*

<sup>3</sup> *Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili - 43007 Tarragona, Spain*

<sup>4</sup> *IPHES, Institut Català de Paleoecologia Humana i Evolució Social - C/Escorxador s/n, 43003 Tarragona, Spain*

received 24 July 2013; accepted in final form 14 November 2013  
published online 10 December 2013

PACS 89.65.-s – Social and economic systems

PACS 89.75.Fb – Structures and organization in complex systems

PACS 89.75.Hc – Networks and genealogical trees

**Abstract** – The ability to understand and eventually predict the emergence of information and activation cascades in social networks is core to complex socio-technical systems research. However, the complexity of social interactions makes this a challenging enterprise. Previous works on cascade models assume that the emergence of this collective phenomenon is related to the activity observed in the local neighborhood of individuals, but do not consider what determines the willingness to spread information in a time-varying process. Here we present a mechanistic model that accounts for the temporal evolution of the individual state in a simplified setup. We model the activity of the individuals as a complex network of interacting integrate-and-fire oscillators. The model reproduces the statistical characteristics of the cascades in real systems, and provides a framework to study the time evolution of cascades in a state-dependent activity scenario.

Copyright © EPLA, 2013

The proliferation of social networking tools —and the massive amounts of data associated to them— has evidenced that modeling social phenomena demands a complex, dynamic perspective. Physical approaches to social modeling are contributing to this transition from the traditional paradigm (scarce data and/or purely analytical models) towards a data-driven new discipline [1–4]. This shift is also changing the way in which we can analyze social contagion and its most interesting consequence: the emergence of information cascades in the Information and Communication Technologies (ICT) environment. Theoretical approaches, like epidemic and rumor dynamics [5–7], reduce these events to physically plausible mechanisms. These idealizations deliver analytically tractable models, but they attain only a qualitative resemblance to empirical results [8], for instance regarding cascade size distributions.

The vast majority of models to this end —including the threshold model, overviewed in the next section— are based on a dynamical process that determines individuals' activity (transmission of information), and this activity is propagated according to certain rules usually

based on the idea of social reinforcement, *i.e.* the more active neighbors an individual has, the larger his probability to become also active, and thus to contribute to the transmission of information. Yet, the challenge of having mechanistic models that include more essential factors, like the self-induced (intrinsic, spontaneous) propensity of individuals to transmit information, still remain open —though some contributions emerge in this fast-growing field [9,10].

Furthermore, the availability of massive amounts of microblogging data logs, like Twitter, places scholars in the position to scrutinize the patterns of real activity and model them. These patterns indicate that avalanche phenomena are not isolated events. Instead, users engaged in a certain topic repeatedly participate, affecting each other and giving rise to an heterogeneous collection of cascades emerging over time, which cannot be modeled independently of each other.

Accordingly, we propose a new framework that extends the classical threshold model to accommodate the temporal evolution of interdependent cascading events. Our proposal conceals also, in an idealized manner, other desirable

ingredients, such as genuine complex contagion or self-induced motivation, to participate.

**Time-constrained activity cascades.** – Before addressing the model itself, we need to understand what is meant to be modeled. There is a general consensus around the concept of cascade, which—in the ICT environment—can be outlined in the following way: the basic criterion to include a node  $i$  in the cascade where  $j$  belongs to is to guarantee that i)  $i$  and  $j$  became neighbors at  $t_1$  (the notion of “friend” must be understood broadly here); ii)  $i$  received a piece of information from  $j$ , who had previously sent it out, at time  $t_2$ ; and finally iii) the node  $i$  sends out a piece of information at time  $t_3$ . Typically, no strict time restriction exists besides the fact that  $t_1 < t_2 < t_3$ : the emphasis is generally placed on whether the *same* content is flowing [8]. This content-based view is useful when considering very specific pieces of information (e-mail chain letters [11] or URL forwarding [12,13], for instance), but renders a scenario in which the only possible transition for a node (user) is from inactive (susceptible) to active (infected).

Such vision oversees the fact that online platforms allow users to share contents, but also (more often than not) to spread behavior. Indeed, discussion over a topic typically happens not by mere information retransmission, but by iterated activity (variable units of information expressed in online text, evolving over time) [14] which influences (motivates) other users to join. A richer, time-constrained representation can then be obtained if conditions i) to iii) above are accepted, except that for  $i$  to be included in an avalanche started at  $j$ , the piece of information being transmitted may or may not be the same, and  $t_3 - t_2 \leq \Delta\tau$ , where  $\tau$  is an arbitrary (typically small) time lapse. In other words, the condition for  $i$  and  $j$  to be included in the same cascade is to exhibit temporally correlated, quasi-synchronized activity [8,15,16]. With this slight modification, not one but multiple cascading events can be measured from an activity data set (for instance, a collection of time-stamped tweets), and a single user may participate many times in the same cascade, in different times.

Empirical (real) cascades hereafter refer to such time-constrained representation.

**The threshold model.** – Along the lines of content-based cascades, the reputed threshold model [17] (and its networked version [18]) mimics social dynamics, where the pressure to engage a behavior increases as more friends adopt that same behavior. Briefly, the networked threshold model assigns a fixed threshold  $\tau$ , drawn from a distribution  $0 \leq g(\tau) \leq 1$ , to each node (individual) in a complex network of size  $N$  and an arbitrary degree distribution  $p_k$ . Each node is marked as *inactive* except an initial seeding fraction of active nodes, typically  $\Phi_0 = 1/N$ . Denoting  $a_i$  the number of active neighbors, a node  $i$  with degree  $k_i$  updates its state becoming active whenever the fraction of active neighbors  $a_i/k_i > \tau_i$ . The simulation of this mechanistic process evolves following this rule until an

equilibrium is reached, *i.e.*, no more updates occur. Given this setup, the *cascade condition* in degree-uncorrelated networks can be derived from the growth of the initial fraction of active nodes, who on their turn might induce the one-step-to-activation (vulnerable) nodes. Therefore, large cascades can only occur if the average cluster size of vulnerable nodes diverges. This condition is met at [18,19]

$$F = \sum_k k(k-1)\rho_k p_k = \langle k \rangle, \quad (1)$$

where  $\rho_k$  is the density of nodes with degree  $k$  close to their activation threshold,  $p_k$  is the fraction of nodes of degree  $k$  and  $\langle k \rangle$  is the average degree [18].

For  $F < \langle k \rangle$  all the clusters of vulnerable nodes are small, and the initial seed can not spread beyond isolated groups of early adopters; on the contrary, if  $F > \langle k \rangle$  then small fraction of disseminators may unleash—with finite probability—large cascades. More recently, the cascade condition has been analytically determined for different initial conditions [19] as well as for modular and correlated networks [20,21], while placing the threshold model in the more general context of critical phenomena and percolation theory [22].

As mentioned, the model has a limited scope since it can account only for one-shot events, for instance the diffusion of a single rumor or the adoption of an innovation. Also, this framework leaves no room for spontaneous initiative: even low-threshold nodes—those with higher propensity to participate in a cascade—will not be recruited unless their neighbors act upon them. Empirical evidence suggests, instead, that once an agent becomes active that behavior will be sustained, and reinforced, over time [23]. This creates a form of enduring activation that will be affected and affect other agents over time in a recursive way. Indeed, events evolve in time—and so do the cascades elicited therein [16], as a consequence of dynamical changes in the states of agents as dynamics progress. Cascades are then events that brew over time in a system that holds some memory of past interactions. Moreover, the propensity to be active in the propagation of information sometimes depends on other factors than raw social influence, *e.g.*, mood, personal implication, opinion, etc.

**Integrate-and-fire model: analytical approach.** – In this paper, we present a model with self-sustained activity, where system-wide events emerge as microscopical conditions become increasingly correlated. We capitalize on the classical integrate-and-fire oscillator (IFO) model by Mirollo and Strogatz [24]. In this model, each node in a network of size  $N$  is characterized by a voltage-like state  $m \in [0, 1]$  of an oscillator, which monotonically increases with phase  $\phi$  until it reaches  $m = 1$ , and then it fires or *activates* (emits information to its coupled neighbors), and immediately deactivates (resets its state to  $m = 0$ ). The pulsatile dynamics, makes that each time a node becomes active, the state of its  $k$  neighbors is increased by  $\varepsilon$ . More precisely,  $m$  is uniformly distributed at  $t = 0$  and evolves

such that

$$m = f(\phi) = \frac{1}{w} \ln(1 + [e^w - 1]\phi) \quad (2)$$

parametrized by  $w > 0$  to guarantee that  $f$  is concave down. Whenever  $m_i = 1$  then *instantaneously*  $m_j = \min(1, m_j + \varepsilon)$ , if the edge  $(i, j)$  exists. Thus, a node may reach activation either by itself (spontaneously, as its phase comes to an end) or because of its neighbors' action (which may pull it up to action).

Integrate-and-fire models have been extremely useful to assess the bursty behavior and the emergence of cascades in neuronal systems represented in lattices [25] and complex networks [26,27]. We propose to model social systems as a complex network of IFOs representing the time evolving activation of individuals. Within this framework, spontaneous propensity —activation regardless exogenous factors— is guaranteed; while contagion is genuinely complex, *i.e.* the number of necessary external influences (if any) to show activity varies in time. Remarkably, activity is purely periodical only if the oscillators are either isolated (disconnected), dynamically uncoupled ( $\varepsilon = 0$ ) or the dynamics have reached full (irreversible) synchronization. These three scenarios are irrelevant in terms of social modeling. On the other hand, some traces of periodicity in users' activity in socio-technical platforms —like Twitter— have been widely studied at the aggregate level (see, for instance, [28]), and they exist also at the individual level.

The model comprises two free parameters,  $w$  and  $\varepsilon$ , which are closely related. Dissipation  $w$  may be interpreted as the willingness or *intrinsic propensity* of agents to participate in a certain diffusion event: the larger  $w$ , the shorter it takes for a node to enter the tip-over interval  $1 - \varepsilon < m < 1$ . Conversely,  $\varepsilon$  quantifies the amount of influence an agent exerts onto its neighbors when it shows some activity. Larger  $\varepsilon$ 's will be more consequential for agents, forcing them more rapidly into the tip-over region. Both quantities affect the level of *motivation*  $m$  of a given agent. Note that  $\varepsilon$  in the current framework evokes  $\tau$  in the classical threshold model, in the sense that both determine the width of the tip-over region. Finally, the phase is translated into time steps, and then prescribed as  $\phi(t) = t$ .

To gain some analytical insight, we use eq. (1) to derive the cascade condition in this new framework. Note that now the distribution of activity is governed by  $\rho(t) = 1 - \int_0^{1-\varepsilon} g(m, t) dm$  where  $g(m, t)$  corresponds to the states' probability distribution at a certain time  $t$ . For an initial uniform distribution of motivation  $m$  and a fixed  $\varepsilon$ , the condition for the emergence of cascades reads at time  $t = 0$

$$\varepsilon \sum_k k(k-1)p_k = \langle k \rangle \quad (3)$$

(see inset (a) in fig. 1). And in general for any time

$$\rho(t) \sum_k k(k-1)p_k = \langle k \rangle, \quad (4)$$

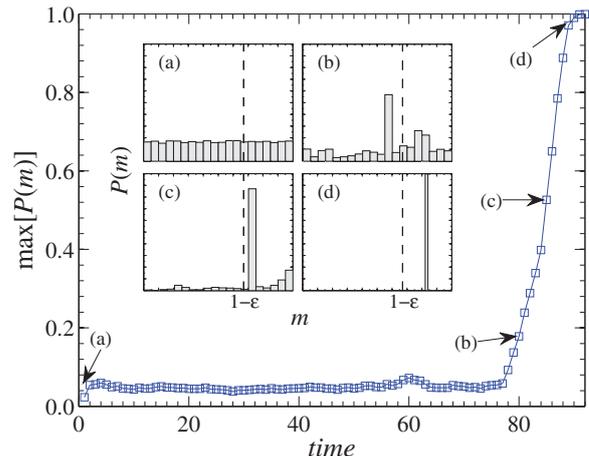


Fig. 1: (Color online) Insets (a)–(d): motivation  $m$  probability distributions of four different representative times along the synchronization window. Each snapshot depicts the  $m$ -state histogram of the  $N$  oscillators. The dynamics begins with a random uniform distribution of  $m$ -states (inset (a)) and it progressively narrows during the transition to synchrony (inset (d)). Main: largest fraction of synchronized nodes across time. The path to synchronization evolves steadily at a low level, and eventually suffers an abrupt transition.

which implies that the cascade condition depends on time in our proposed framework. Clearly, in this scenario  $\rho(t)$  is not a function of the node degree  $k$ , as opposed to  $\rho_k$  in Watts' proposal.

As the dynamics evolve in time, the states of the nodes progressively correlate and, consequently, the distribution of states changes dramatically. The evolution of the states distribution is depicted in fig. 1. The initially uniform distribution  $g(m, 0)$  (inset (a)) evolves towards a Dirac  $\delta$  function (inset (d)) as the network approaches global synchronization, *i.e.* global cascade. We have not been able to find a closed analytical expression for the consecutive composition of the function  $g(m, t)$  after an arbitrary number of time steps to reveal the evolution of  $\rho(t)$ , nonetheless it can be solved numerically. Equation (4) reduces to

$$\rho(t)(\langle k^2 \rangle - \langle k \rangle) = \langle k \rangle. \quad (5)$$

The cascade condition is thus

$$\frac{\rho(t)}{1 + \rho(t)} = \frac{\langle k \rangle}{\langle k^2 \rangle}, \quad (6)$$

that exactly corresponds to the bond percolation critical point on uncorrelated networks [29–31]. For the case of random Poisson networks  $\langle k^2 \rangle \sim \langle k \rangle^2$ , then

$$\frac{\rho(t)}{1 + \rho(t)} = \frac{1}{\langle k \rangle}. \quad (7)$$

It is worth highlighting that eqs. (6) and (7) represent an advance in our understanding of non-linear, pulse-coupled dynamics, regardless of our (social) interpretation.

**Integrate-and-fire model: numerical results.** –

We can now explore the cascade condition in the  $(\varepsilon, \langle k \rangle)$  phase diagram in fig. 2, and compare the analytical predictions with results from extensive numerical simulations. Since the time to full synchronization (global cascade) is different for each  $(\varepsilon, \langle k \rangle)$ , we introduce *cycles*. One cycle is complete whenever every node in the network has fired at least once. In this way we bring different time scales to a common, coarse-grained temporal ground, allowing for comparison. The regions where cascades size  $S_c$  above a prescribed threshold  $\xi$  ( $0.25N$ , in fig. 2) are possible are color-coded for each cycle  $0, 25, 75, 100, \dots$ , and black is used in regions where cascades do not reach  $\xi$  (labeled as N.C., “no cascades”, in fig. 2). Note that if cascades are possible for a cycle  $c$ , they will be possible also for any  $c' \geq c$ . This figure renders an interesting scenario: on the one hand, it suggests the existence of critical  $\varepsilon_c$  values below which the cascade condition is systematically frustrated (black area in the phase diagram). On the other, it establishes how many cycles it takes for a particular  $(\varepsilon, \langle k \rangle)$  pair to attain macroscopical cascades (full synchronization) —which becomes an attractor thereafter, for undirected connected networks. Given the cumulative dynamics of the current framework, in contrast with Watts’ model, the region in which global cascades are possible grows with  $\langle k \rangle$ .

Turning to the social sphere, these results open the door to predicting how long it takes for a given topology, and a certain level of inter-personal influence, to achieve system-wide events. Furthermore, the existence of a limiting  $\varepsilon_c$  determines whether such events can happen at all.

Additionally, the predictions resulting from eq. (6) are represented as dashed lines in fig. 2. For the sake of clarity, we only include predictions for  $c = 0$  (dashed black),  $c = 25$  (dashed gray) and  $c = 150$  (dashed white). Projections from this equation run close to numerical results in both homogeneous (fig. 2(a)) and inhomogeneous networks (fig. 2(b)) with degree distribution  $p(k) = k^{-\gamma}$ , although some deviations exist. Noteworthy, eq. (6) clearly overestimates the existence of macroscopic cascades in the case of scale-free networks at  $c = 0$ . Indeed,  $\rho(0) = \varepsilon$  does not yet incorporate the inherent dynamical heterogeneity of a scale-free topology, thus eq. (6) is a better predictor as the dynamics loose memory of the hardwired initial conditions. In the general case  $c > 0$ , deviations are due to the fact that the analytical approach in the current work is not developed beyond first order. Second order corrections to this dynamics (including dynamical correlations) should be incorporated to the analysis in a similar way to that in [21], however it is beyond the scope of the current presentation.

According to Mirolo and Strogatz [24], synchronicity emerges more rapidly when  $w$  or  $\varepsilon$  is large; then the time taken to synchronize, *i.e.* to observe global cascades, is inversely proportional to the product  $w\varepsilon$ . In our simulations in the next Section, we use this cooperative effect between coupling and willingness to fix  $\varepsilon$ , which is set to a set of

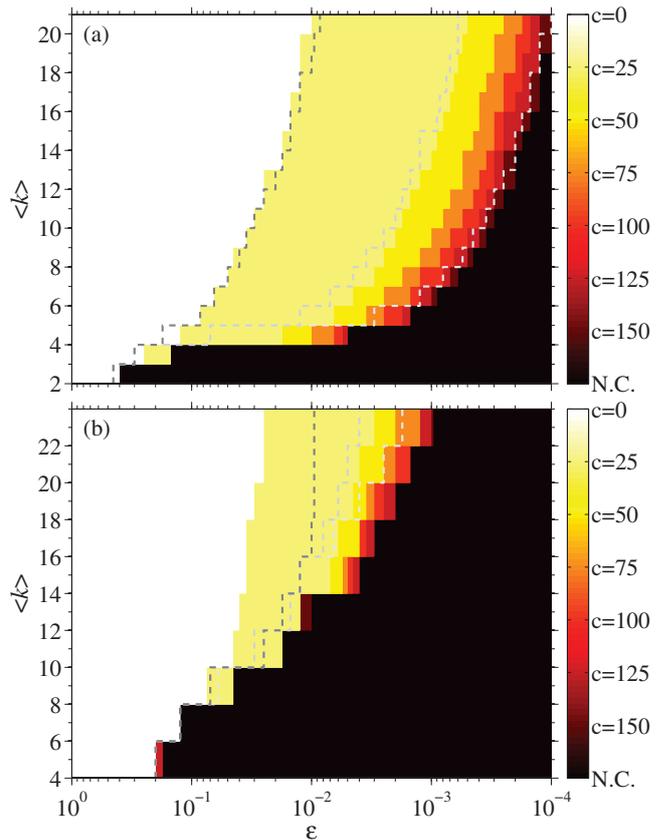


Fig. 2: (Color online)  $(\varepsilon, \langle k \rangle)$  cascade diagram for different cycles (coded by color), with fixed  $w = 3$ . Vertical axis and each dashed line define a confined region in which global cascades might occur according to eq. (6) and for a specific cycle (here we show only the expected zones for  $c = 0$  —dashed white— and  $c = 150$  —dashed gray). Results are obtained for synthetic Erdős-Rényi (a) and scale-free with  $\gamma = 3$  (b) uncorrelated networks of size  $N = 10^4$ . A cascade is considered “macroscopical” if the synchronized cluster  $S_c \geq 0.25N$ . Color codes indicate the existence of at least one cascade  $S > S_c$  in numerical simulations; analytical predictions are averaged over 200 networks with random initial conditions. Note that the cascade condition in (a) often underestimates the actual cascade regions because it does not take into account second-order interactions; the same applies in the lower panel (b), except for  $c = 0$  where the analytical prediction overestimates the results because the inclusion of the hub into the cascade is improbable starting from a uniform distribution.

values (slightly above or below)  $\varepsilon \simeq \varepsilon_c$ , and empirically estimate  $w$  to attain a good matching between observed cascade distributions and our synthetic results.

**Application to real data.** – To illustrate the explanatory power of the dynamical threshold model, we use data from [www.twitter.com](http://www.twitter.com). They comprise a set of  $\sim 0.5$  million Spanish messages publicly exchanged through this platform from the 25th of April to the 25th of May, 2011 [32]. In this period a sequence of civil protests and demonstrations took place, including camping events

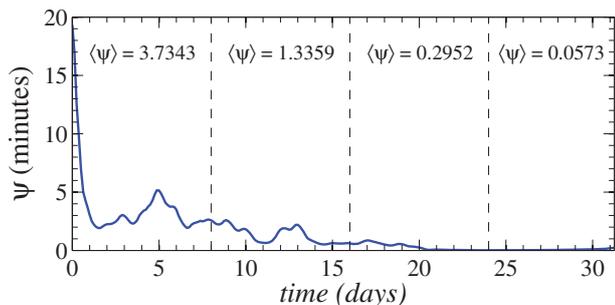


Fig. 3: (Color online) Average inter-event times  $\Psi$  of the whole collection of data. To measure them, an overlapping sliding window scheme has been used (windows span 1 day, the offset between windows is 12 hours). To estimate  $w$  for 2 different periods of the protests, we take their corresponding time slices and compute inter-event times grand averages  $\langle \Psi \rangle$ . Thus, in fig. 4 a  $w \sim 10^{-1}$  will be used for the first period and  $w \sim 20$  for the last one. The scaling  $w \sim 1/\Psi$  is merely an heuristic estimation, and some fine-tune is necessary to achieve a satisfactory matching.

in the main squares of several cities beginning on the 15th of May and growing in the following days. Notably, a pulse-based model suits well with the affordances of this social network, in which any emitted message is instantly broadcasted to the author’s immediate neighborhood —its set of *followers*. For the whole sample, we queried for the list of followers for each of the emitting users, discarding those who did not show outgoing activity during the period under consideration. The set of users  $N = 87569$  plus their following relations constitute the topological support (directed network) for the dynamical process running on top of it. The average number of followers of this network is  $\langle k \rangle = 69$  and its degree distribution scales like  $p(k) \sim k^{-1.5}$ . Note that, unlike other substrates, friendship networks exhibit a high level of reciprocity. In particular,  $r = 0.45$  (as defined in [33]) for this particular case, which implies that many links can effectively be regarded as undirected. This is why the theory —developed for undirected networks— is a reasonable approach for this particular case.

On top of the described network, we measure the empirical *time-constrained activity cascade* (or simply “cascade”) size distribution for different periods, as explained previously. To test the proposed model, we run the dynamics on the same topology for some given parameters  $(w, \varepsilon)$ . Since (as mentioned)  $\varepsilon$  is fixed to the network’s particular  $\varepsilon_c \approx 10^{-3}$ , we only need to determine which values of  $w$  can adequately represent the evolving nature of events. Figure 3 represents the average inter-event times  $\Psi$  in the data for each 32 days of activity around the 15 M movement. Grand averages  $\langle \Psi \rangle$  are computed for 8-day windows, roughly corresponding to the periods for which we offer model fittings (see below). As expected,  $\Psi$  drops abruptly in the first days —when the movement is brewing— and smoothly decreases afterwards, until

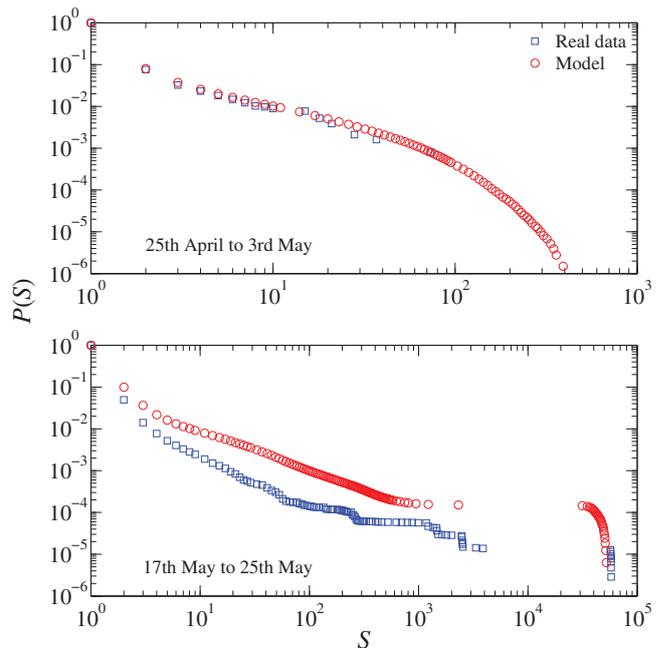


Fig. 4: (Color online) Cascade size cumulative distributions  $P(S)$  of real data (blue squares) and the model counterpart (red circles). We have considered two time windows which significantly differ: first eight days (top) for which we have set  $w = 0.1$ ; last eight days (bottom) for which we have  $w = 30.0$ . Note that  $\varepsilon_c \approx 10^{-3}$ . The model performs well in both periods, the relative error of the slope in the linear region is  $< 1\%$  (see slope values on the figure). Real-data distributions are measured as described in the main text, see also [15,16].

$\Psi \approx 0$ . Intuitively, the willingness of an individual to participate in the protest is proportional to real-world excitation level, *i.e.* it scales like the inverse of  $\Psi$ , and thus we envisage that  $w \sim 1/\Psi$ .

In fig. 4 we compare empirical *vs.* synthetic cascade size distributions for different periods of the protests: the “slow-growth” phase (25th April to 3rd May; blue squares in the upper panel) for the first days, when the protest is limited to some online activists; and the “explosive” phase (19th to 25th May; blue squares in the lower panel), which comprehends the most active interval —the reaction to the Spanish government ban on demonstrations around local elections on the 22nd May. The proposed dynamics is run on the same topology for different  $w$  values, with remarkable success (red circles), though the bottom panel does not show so good of an agreement as the top one. In particular, a small range around  $w \sim 1/\langle \Psi \rangle$  values was tested in the simulations, seeking a minimization of the relative error of the slope in the linear region of  $P(S)$ , *i.e.* the cascade size distribution.

Moreover, as seen in fig. 4, our model is able to capture different regimes. Admittedly, the gap observed in the data (fig. 4, bottom panel) is reminiscent of a supercritical regime, in which one either has small-size events or system-wide cascades. The change between regimes,

namely, from the subcritical one represented in fig. 4, top panel, and the super-critical phase needs the activity of the system to be long-lived, as this is a transition that takes place at different time windows. Note that this is a quite relevant feature of the model here introduced, since existing threshold models do not allow that individuals engage in more than one cascade (recall that once one individual is active remains so forever) and therefore cannot lead to the same kind of temporal transition observed in the real data.

**Conclusions.** – Summarizing, we have proposed a time-dependent continuous self-sustained model of social activity. The model can be analyzed in the context of previous cascade models, and encompasses novel phenomenology as the time dependence of the critical value of the emergence of cascades. We interpret it under a social perspective, where collective behavior is seen as an evolving phenomenon resulting from inter-personal influence, contagion and memory —while, collaterally, it sheds new light to the behavior of pulse-coupled oscillator dynamics. In a general perspective, our modeling framework offers an alternative approach to the analysis of interdependent decision making and social influence. It complements threshold models and complex contagion taking into account time dynamics and recursive activation, and also splits motivation into two components: intrinsic propensity and strength of social influence. We also anticipate that the exploration of the whole parametric space would lead to new insights about the effects of social influence and interdependence in social collective phenomena.

\* \* \*

We gratefully acknowledge SANDRA GONZÁLEZ-BAILÓN for interesting discussions and comments on the draft. This work has been partially supported by MINECO through Grant FIS2011-25167 and FIS2012-38266; Comunidad de Aragón (Spain) through a grant to the group FENOL; Generalitat de Catalunya through the grant SGR2009-838; and by the EC FET-Proactive Project PLEXMATH (grant 317614). AA also acknowledges partial financial support from the ICREA Academia and the James S. McDonnell Foundation.

## REFERENCES

- [1] WATTS D. J., *Annu. Rev. Sociol.*, **30** (2004) 243.
- [2] LAZER D., PENTLAND A. S., ADAMIC L., ARAL S., BARABÁSI A. L., BREWER D., CHRISTAKIS N., CONTRACTOR N., FOWLER J. and GUTMANN M. *et al.*, *Science*, **323** (2009) 721.
- [3] CONTE R., GILBERT N., CIOFFI-REVILLA C., DEFFUANT G., KERTESZ J., LORETO V., MOAT S., NADAL J.-P., SANCHEZ A., NOWAK A. *et al.*, *Eur. Phys. J. ST*, **214** (2012) 325.
- [4] GILES J., *Nature*, **488** (2012) 448.
- [5] RAPOPORT A., *Bull. Math. Biophys.*, **15** (1953) 523.
- [6] DALEY D. J. and KENDALL D. G., *Nature*, **204** (1964) 1118.
- [7] GOFFMAN W., *Nature*, **204** (1964) 225.
- [8] BORGE-HOLTHOEFER J., BAÑOS R. A., GONZÁLEZ-BAILÓN S. and MORENO Y., *J. Complex Netw.*, **1** (2013) 3, doi: 10.1093/cmnet/cnt006.
- [9] GLEESON J. P., CELLAI D., ONNELA J.-P., PORTER M. A. and REED-TSOCHAS F., arXiv:1305.7440 (2013).
- [10] HOGG TAD, LERMAN KRISTINA and SMITH LAURA M., *Hum. J.*, **2** (2013) 25, arXiv:1308.2705 (2013).
- [11] LIBEN-NOWELL D. and KLEINBERG J., *Proc. Natl. Acad. Sci. U.S.A.*, **105** (2008) 4633.
- [12] BAKSHY E., HOFMAN J., MASON W. and WATTS D., *Everyone’s an influencer: quantifying influence on twitter*, in *Proceedings of the 4th ACM International Conference on Web Search and Data Mining (ACM)* 2011, pp. 65–74.
- [13] LERMAN K., GHOSH R. and SURACHAWALA T., arXiv:1202.3162 (2012).
- [14] LESKOVEC J., BACKSTROM L. and KLEINBERG J., *Meme-tracking and the dynamics of the news cycle*, in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (ACM)* 2009, pp. 497–506.
- [15] GONZÁLEZ-BAILÓN S., BORGE-HOLTHOEFER J., RIVERO A. and MORENO Y., *Sci. Rep.*, **1** (2011) 197.
- [16] BORGE-HOLTHOEFER J., RIVERO A. and MORENO Y., *Phys. Rev. E*, **85** (2012) 066123.
- [17] GRANOVETTER M., *Am. J. Sociol.* (1978) 1420.
- [18] WATTS D., *Proc. Natl. Acad. Sci. U.S.A.*, **99** (2002) 5766.
- [19] GLEESON J. and CAHALANE D., *Phys. Rev. E*, **75** (2007) 056103.
- [20] GALSTYAN A. and COHEN P., *Phys. Rev. E*, **75** (2007) 036109.
- [21] GLEESON J., *Phys. Rev. E*, **77** (2008) 046117.
- [22] HACKETT A., MELNIK S. and GLEESON J., *Phys. Rev. E*, **83** (2011) 056107.
- [23] BORGE-HOLTHOEFER J., RIVERO A., GARCÍA I., CAUHÉ E., FERRER A., FERRER D., FRANCO S. D., IÑIGUEZ D., PÉREZ M., RUIZ G. *et al.*, *PLoS ONE*, **6** (2011) e23883.
- [24] MIROLLO R. and STROGATZ S., *SIAM J. Appl. Math.*, **50** (1990) 1645.
- [25] CORRAL A., PÉREZ-VICENTE, DÍAZ-GUILERA and ARENAS A., *Phys. Rev. Lett.*, **74** (1995) 118.
- [26] TIMME M., WOLF F. and GEISEL T., *Phys. Rev. Lett.*, **89** (2002) 258701.
- [27] ROXIN A., RIECKE H. and SOLLA S. A., *Phys. Rev. Lett.*, **92** (2004) 198101.
- [28] GOLDBERGER S. A. and MACY M. W., *Science*, **333** (2011) 1878.
- [29] MOLLOY M. and REED B., *Random Struct. Algorithms*, **6** (1995) 161.
- [30] MOLLOY M. and REED B., *Combin. Probab. Comput.*, **7** (1998) 295.
- [31] DOROGOVTSSEV S., GOLTSEV A. and MENDES J., *Rev. Mod. Phys.*, **80** (2008) 1275.
- [32] <http://cosnet.bifi.es/research-lines/online-social-systems/15m-dataset>.
- [33] GARLASCHELLI D. and LOFFREDO M. I., *Phys. Rev. Lett.*, **93** (2004) 268701.