

## EPL, **97** (2012) 48004 doi: 10.1209/0295-5075/97/48004

## Epidemic phase transition of the SIS type in networks

P. VAN MIEGHEM<sup>(a)</sup>

Delft University of Technology - Mekelwey 4, NL-2628 CD Delft, The Netherlands, EU

received 9 August 2011; accepted in final form 11 January 2012 published online 15 February 2012

PACS 89.20.-a – Interdisciplinary applications of physics PACS 89.75.Hc – Networks and genealogical trees

Abstract – By making only one approximation of a mean-field type, we determine the nature of the SIS type of epidemic phase transition in any network: the steady-state fraction of infected nodes  $y_{\infty}$  is linear in  $(\tau_c^{-1} - \tau^{-1})$  for effective infection rates  $\tau \downarrow \tau_c$ , the derivative of  $y_{\infty}$  at the epidemic threshold  $\tau_c = \frac{1}{\lambda_1}$  is exactly computed and depends on the largest eigenvalue  $\lambda_1$  of the adjacency matrix and on the first- and third-order moments of the corresponding eigenvector. Since coupled oscillators in a network synchonize at a coupling strength proportional to  $\frac{1}{\lambda_1}$ , a similar characterization of the phase transition is suggested. The behavior of  $y_{\infty}$  around  $\tau_c$  was the missing part in the general steady-state theory of a SIS-type epidemic on a network.

Copyright © EPLA, 2012

**Introduction.** – The increasing threats from cybercrime and the expected outbreak of a new lethal virus in our increasing human population justify studies on virus spread in networks. In particular, the role of the network in the virus contamination process, which is still insufficiently well understood, is the scope of this paper. A remarkable property of the susceptible-infected-susceptible (SIS) virus model [1,2] is the appearance of a phase transition when the effective infection rate  $\tau = \frac{\beta}{\delta}$  approaches the epidemic threshold  $\tau_c = \frac{1}{\lambda_1}$ , where  $\lambda_1$  is the largest eigenvalue of the adjacency matrix A, also called the spectral radius<sup>1</sup>. Below the epidemic threshold  $\tau_c$ , the network is virus free in the steady-state, while above  $\tau_c$ , there is always a fraction of nodes that remains infected.

Our analysis, based on the N-intertwined SIS virus spread model [4], specifies, for the first time, the precise behavior of the phase transition for values of the effective infection rate at  $\tau = \tau_c + \epsilon$ , where  $\epsilon > 0$  is arbitrarily small. In earlier work (see ref. [5]), only the position of the epidemic threshold, *i.e.*  $\tau_c = \frac{1}{\lambda_1}$ , has been determined and most analyses were only able to describe the region  $\tau \leq \tau_c$ . Our method illustrates the power of an algebraic and spectral analysis of graphs (see, *e.g.*, [3]). The first and third moments of the normalized eigenvector  $x_1$  belonging to  $\lambda_1$  appear in the main result, Theorem 1 below. In the Kuramoto model of coupled oscillators in a network, which is as basic [6] as the SIS virus model, the onset of synchronization [7] occurs at a coupling strength that is also inversely proportional to  $\lambda_1$ . Since these two dynamic processes on a network possess the same  $\frac{1}{\lambda_1}$  type of critical threshold between two phases, we speculate that first and third moments of  $x_1$  also determine the synchronized state as estimated below.

## The fraction $y_{\infty}$ of infected nodes in the steady-state. –

The N-intertwined SIS model in brief. A network is represented by an undirected graph G(N, L) with N nodes and L links. The network topology is described by a symmetric adjacency matrix A, in which the element  $a_{ij} = a_{ji} = 1$  if there is a link between nodes *i* and *j*, otherwise  $a_{ij} = 0$ . The state of a node *i* is specified by a Bernoulli random variable  $X_i \in \{0, 1\}$ :  $X_i = 0$  for a healthy node and  $X_i = 1$  for an infected node. A node *i* at time *t* can be in one of the two states: *infected*, with probability  $v_i(t) = \Pr[X_i(t) = 1]$  or healthy, with probability  $1 - v_i(t)$ . We assume that the curing process per node i is a Poisson process with rate  $\delta$ , and that the infection rate per link is a Poisson process with rate  $\beta$ . This is the general description of the simplest type of a SIS virus spread model in a network. This SIS model can be expressed exactly in terms of a continuous-time Markov model with  $2^N$  states as shown in [4]. Unfortunately, the exponentially increasing state space with N prevents the determination of the set of  $\{v_i(t)\}_{1 \leq i \leq N}$  in realistic networks, which has triggered a spur of research to find good approximate solutions.

<sup>&</sup>lt;sup>(a)</sup>Present address: Faculty of Electrical Engineering, Mathematics and Computer Science - P.O. Box 5031, 2600 GA Delft, The Netherlands, EU; E-mail: P.F.A.VanMieghem@tudelft.nl.

<sup>&</sup>lt;sup>1</sup>In a graph G with N nodes and L links, the spectral radius is bounded [3] by  $\frac{2L}{N}\sqrt{1+\frac{\operatorname{Var}[D]}{E[D]}} \leq \lambda_1 \leq d_{\max}$ , where D is the degree of an arbitrary node and  $d_{\max}$  is the maximum degree in G. Equality is attained for regular graphs.

In contrast to all published SIS type of models, the N-intertwined model, proposed and investigated in depth in [4] and reviewed in [8], only makes *one* approximation in the exact SIS model and is applicable to *all* graphs. In particular, the actual infection rate of node i due its infected neighbors,  $\beta \sum_{j=1}^{N} a_{ij} 1_{\{X_j(t)=1\}}$ , is replaced by its mean<sup>2</sup>  $\beta \sum_{j=1}^{N} a_{ij} v_j(t)$ . This "mean-field" type of approximation transforms the exact set of  $2^N$  linear differential equations into a set of N non-linear differential equations (see (1) and (2) below) and introduces both the epidemic phase transition and the "meta-stable state"<sup>3</sup>. The steady-state in the non-linear N-intertwined virus spread model refers to the meta-stable state, which is reached exponentially rapidly and which reflects real epidemics more closely. Simulations in [4] indicate that the accuracy of the N-intertwined model (compared to the exact SIS Markov model) increases with N. In [10], we show that the mean-field approximation upper bounds the exact  $\Pr[X_i(t) = 1]$  (which is useful to guarantee epidemic safety bounds in real networks) and that it implies that the random variables  $X_i$  and  $X_i$  are implicitly assumed to be independent. Since this basic assumption is increasingly good for large N, we expect that the deductions from the N-intertwined model are asymptotically (for  $N \to \infty$ ) almost exact for real-world networks. The homogenous N-intertwined model, where the infection and curing rate is the same for each link and node in the network, has been extended to a heterogeneous setting in [5].

The governing differential equation in the N-intertwined model for a node i is

$$\frac{\mathrm{d}v_i(t)}{\mathrm{d}t} = \beta(1 - v_i(t)) \sum_{j=1}^N a_{ij} v_j(t) - \delta v_i(t).$$
(1)

In words, the time derivative of the infection probability of a node *i* consists of two competing processes: 1) while healthy with probability  $(1 - v_i(t))$ , all infected neighbors, an event with probability  $\sum_{j=1}^{N} a_{ij}v_j(t)$ , try to infect the node *i* with rate  $\beta$  and 2) while infected with probability  $v_i(t)$ , the node *i* is cured at rate  $\delta$ . Defining the vector  $V(t) = [v_1(t)v_2(t)\cdots v_N(t)]^T$ , the matrix representation based on (1) becomes

$$\frac{\mathrm{d}V(t)}{\mathrm{d}t} = (\beta A - \delta I) V(t) - \beta \operatorname{diag}(v_i(t)) A V(t), \quad (2)$$

where  $\operatorname{diag}(v_i(t))$  is the diagonal matrix with elements  $v_1(t), v_2(t), \ldots, v_N(t)$ . The mean-field model of Kephart and White [11], that is identitical to the "small degree variance" model of Pastor-Satorras and Vespignani [12], is found from the set (1) by assuming that each node

experiences a same infection. That uniform-homogeneouslocal assumption reduces the set of N equations to one and replaces the actual number of neighbors  $\sum_{j=1}^{N} a_{ij}$ of node *i*, by an average *r*, resulting in  $\frac{dv(t)}{dt} = \beta r(1 - v(t))v(t) - \delta v(t)$ . When the variance of the degree distribution is large, Pastor-Satorras and Vespignani [12] propose a heterogeneous variant,  $\frac{dv(t|k)}{dt} = \beta k(1 - v(t|k))\Theta\{v(t)\} - \delta v(t|k)$ , where v(t|k) is the probability of infection of a node with degree *k* and  $\Theta\{v(t)\}$  is the probability that a link points to an infected node. The latter  $\Theta\{v(t)\}$  is a function of v(t|k) and solved self-consistently in [12]. Unfortunately, the relation between this "heterogeneous" variant and the *N*-intertwined model is less obvious.

In the sequel, we focus on the steady-state, where  $v_{i\infty} = \lim_{t\to\infty} v_i(t)$  and  $\lim_{t\to\infty} \frac{dv_i(t)}{dt} = 0$ . From (1), we obtain

$$v_{i\infty} = \frac{\beta \sum_{j=1}^{N} a_{ij} v_{j\infty}}{\beta \sum_{j=1}^{N} a_{ij} v_{j\infty} + \delta} = 1 - \frac{1}{1 + \tau \sum_{j=1}^{N} a_{ij} v_{j\infty}}.$$
 (3)

Beside the trivial solution  $v_{i\infty} = 0$ , (3) illustrates that there is another positive solution reflecting the meta-stable state in which we are interested here. For regular graphs, where each node has degree d, symmetry in the steadystate implies that  $v_{i\infty} = v_{\infty}$  for all nodes i and it follows from (3) with the definition of the degree  $d_i = \sum_{j=1}^{N} a_{ij}$ that

$$v_{\infty;\text{regular}} = y_{\infty;\text{regular}}(\tau) = 1 - \frac{1}{\tau d}$$
 (4)

where  $y_{\infty} = \frac{1}{N} \sum_{i=1}^{N} v_{i\infty}$  is the fraction of infected nodes in the steady-state.

General relations. After left-multiplication of the steady-state version of (2) by the vector

$$V_{\infty}^{T} \operatorname{diag}(v_{i\infty}^{k-1}) = \begin{bmatrix} v_{1\infty}^{k} & v_{2\infty}^{k} & \cdots & v_{N\infty}^{k} \end{bmatrix}$$

which we denote by  $(V_{\infty}^k)^T$ , we obtain the scalar equation

$$\left(V_{\infty}^{k}\right)^{T}V_{\infty} = \sum_{j=1}^{N} v_{j\infty}^{k+1} = \tau\left(\left(V_{\infty}^{k}\right)^{T}AV_{\infty} - \left(V_{\infty}^{k+1}\right)^{T}AV_{\infty}\right).$$
(5)

For k = 0 in (5), and introducing the all one vector  $u = \lim_{k \to 0} V_{\infty}^k$ , we arrive at the fundamental relation for  $y_{\infty} \in [0, 1]$  in terms of the vector  $V_{\infty}$ :

$$Ny_{\infty} = u^T V_{\infty} = \tau \left( u - V_{\infty} \right)^T A V_{\infty}.$$
 (6)

For k = 1 in (5), the norm  $||V_{\infty}||_2^2 = V_{\infty}^T V_{\infty} = \sum_{j=1}^N v_{j\infty}^2$ obeys

$$V_{\infty}^{T}V_{\infty} = \tau \left( V_{\infty}^{T}AV_{\infty} - V_{\infty}^{T} \operatorname{diag}(v_{i\infty})AV_{\infty} \right).$$
(7)

<sup>&</sup>lt;sup>2</sup>The expectation of the indicator function  $1_X$ , where X is a random variable, equals  $E[1_X] = \Pr[X]$ .

<sup>&</sup>lt;sup>3</sup>In the exact SIS model, the steady-state is the healthy state, which is the only absorbing state in the Markov process. However, in networks of realistic size N, this steady-state is only reached after an unrealistically long time [9].

An eigenvector approach. Since the eigenvectors  $x_1, x_2, \ldots, x_N$  belonging to the eigenvalues  $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_N$  of the adjacency matrix A span the N-dimensional vector space, we can write the steady-state infection probability vector  $V_{\infty}(\tau)$  as a linear combination of the eigenvectors of A,

$$V_{\infty}(\tau) = \sum_{k=1}^{N} \gamma_k(\tau) x_k, \qquad (8)$$

where the coefficient  $\gamma_k(\tau) = x_k^T V_{\infty}(\tau)$  is the scalar product of  $V_{\infty}(\tau)$  and the eigenvector  $x_k$  and where the eigenvector  $x_k$  obeys the normalization  $x_k^T x_k = 1$ . Physically, (8) maps the dynamics  $V_{\infty}(\tau)$  of the process onto the eigenstructure of the network, where  $\gamma_k(\tau)$  determines the importance of the process in a certain eigendirection of the graph. The definition  $y_{\infty}(\tau) = \frac{1}{N} u^T V_{\infty}(\tau)$  shows that

$$y_{\infty}(\tau) = \frac{1}{N} \sum_{k=1}^{N} \gamma_k(\tau) u^T x_k.$$
(9)

Substitution of (8) into (6) yields

$$y_{\infty}(\tau) = \frac{\tau}{N} \sum_{k=1}^{N} \lambda_k \gamma_k(\tau) \left( u^T x_k - \gamma_k(\tau) \right).$$
(10)

For irregular graphs, generally,  $\gamma_m(\tau) = x_m^T V_{\infty}(\tau) \neq 0$  for m > 1 and all eigenvalues and eigenvectors in (10) play a role. Moreover,  $\gamma_m(\tau)$  can be negative, as well as  $\lambda_m$ , while  $\sum_{k=1}^N \lambda_k = 0$  (see [3], p. 30). The larger the spectral gap  $\lambda_1 - \lambda_2$  and the smaller  $|\lambda_N|$ , the more  $y_{\infty}$  is determined by the dominant k = 1 term in (10), and the more its viral behavior approaches that of a regular graph. Graphs with large spectral gap possess strong topological robustness [3], in the sense that it is difficult to tear that network apart.

Just above the epidemic threshold. – The Perron-Frobenius Theorem [3] states that all components of  $x_1$ are non-negative (and strictly positive when the graph G is connected). Moreover, all components of  $V_{\infty}(\tau)$  are positive as well for  $\tau > \tau_c$  so that  $\gamma_1(\tau) > 0$  and  $\gamma_1(\tau) >$  $\gamma_m(\tau)$  for all m > 1. The fact that the epidemic threshold occurs at  $\tau = \tau_c = \frac{1}{\lambda_1}$  has been proved in several papers, see, *e.g.*, [4,13]. Here, we recall the fundamental lemma for the *N*-intertwined SIS model, proved in [4].

**Lemma 1.** There exists a value  $\tau_c = \frac{1}{\lambda_1} > 0$  and for  $\tau < \tau_c$ , there is only the trivial steady-state solution  $V_{\infty} = 0$ . Beside the  $V_{\infty} = 0$  solution, there is a second, non-zero solution for all  $\tau > \tau_c$ . For  $\tau = \tau_c + \varepsilon$ , it holds that  $V_{\infty} = \alpha x_1$ , where  $\varepsilon, \alpha > 0$  are arbitrarily small constants and where  $x_1$  is the eigenvector belonging to the largest eigenvalue  $\lambda_1$  of the adjacency matrix A.

Lemma 1 shows that, for all graphs,  $V_{\infty} = \alpha x_1 + \xi y$ , where y is a vector orthogonal to  $x_1$ ,  $\alpha$  tends to zero as  $\tau \downarrow \tau_c$ , while  $\xi$  tends faster to zero in that limit than  $\alpha$ . The following theorem is our fundamental result:

**Theorem 1.** For any graph, the steady-state fraction of infected nodes  $y_{\infty}$  obeys

$$y_{\infty}(\tau) = \frac{1}{N} \frac{u^T x_1}{\lambda_1 \sum_{j=1}^N (x_1)_j^3} \left(\tau_c^{-1} - \tau^{-1}\right) + O\left(\tau_c^{-1} - \tau^{-1}\right)^2$$
(11)

when  $\tau$  approaches the epidemic threshold  $\tau_c$  from above.

**Proof:** The proof of Lemma 1 suggests when  $\tau \downarrow \tau_c$  that  $\alpha = \alpha_0(\tau_c^{-1} - \tau^{-1})^q + o((\tau_c^{-1} - \tau^{-1})^q)$  and  $\xi = \xi_0(\tau_c^{-1} - \tau^{-1})^{\kappa} + o((\tau_c^{-1} - \tau^{-1})^{\kappa})$ , where  $\kappa > q > 0$ . Comparing to the definition (8) in the limit that  $\tau$  approaches the epidemic threshold  $\tau_c$  from above, we have that  $\gamma_1(\tau) = c_1(\tau_c^{-1} - \tau^{-1})^q + o((\tau_c^{-1} - \tau^{-1})^q)$  and  $\gamma_k(\tau) = c_k(\tau_c^{-1} - \tau^{-1})^{\kappa} + o((\tau_c^{-1} - \tau^{-1})^{\kappa})$ , where  $c_1 > 0$  because all components of  $V_\infty$  must be non-negative.

When combining both expressions (9) and (10) for  $y_{\infty}$ , we obtain

$$\sum_{k=1}^{N} \left(\lambda_k - \tau^{-1}\right) \gamma_k(\tau) u^T x_k = \sum_{k=1}^{N} \lambda_k \gamma_k^2(\tau).$$
(12)

Substitution of (8) into (7) yields

$$V_{\infty}^{T}V_{\infty} = \sum_{m=1}^{N} \sum_{k=1}^{N} \gamma_{k}(\tau)\gamma_{m}(\tau)x_{m}^{T}x_{k}$$
$$= \tau \left(\sum_{m=1}^{N} \sum_{k=1}^{N} \lambda_{k}\gamma_{k}(\tau)\gamma_{m}(\tau)x_{m}^{T}x_{k}\right)$$
$$-\sum_{k=1}^{N} \sum_{q=1}^{N} \sum_{k=1}^{N} \lambda_{k}\gamma_{q}(\tau)\gamma_{m}(\tau)\gamma_{k}(\tau)x_{m}^{T}\operatorname{diag}((x_{q})_{i})x_{k}\right).$$

Invoking orthogonality of eigenvectors,  $x_m^T x_k = \delta_{km}$ , yields

$$\sum_{k=1}^{N} (\lambda_{k} - \tau^{-1}) \gamma_{k}^{2}(\tau) = \sum_{k=1}^{N} \sum_{q=1}^{N} \sum_{k=1}^{N} \lambda_{k} \gamma_{q}(\tau) \gamma_{m}(\tau) \gamma_{k}(\tau) \times \sum_{j=1}^{N} (x_{m})_{j} (x_{q})_{j} (x_{k})_{j}.$$
 (13)

When  $\tau \downarrow \tau_c$ , (12) reveals that the left-hand side is of the order

$$c_{1}u^{T}x_{1}\left(\tau_{c}^{-1}-\tau^{-1}\right)^{q+1} + \sum_{k=2}^{N}\left(\lambda_{k}-\lambda_{1}\right)c_{k}u^{T}x_{k}\left(\tau_{c}^{-1}-\tau^{-1}\right)^{\kappa} + o\left(\left(\tau_{c}^{-1}-\tau^{-1}\right)^{\min(q+1,\kappa)}\right),$$

whereas, due to  $\kappa > q$ , the right-hand side is of the order

$$\lambda_1 c_1^2 \left(\tau_c^{-1} - \tau^{-1}\right)^{2q} + \sum_{k=2}^N \lambda_k c_k^2 \left(\tau_c^{-1} - \tau^{-1}\right)^{2\kappa} + o\left(\left(\tau_c^{-1} - \tau^{-1}\right)^{\min(2\kappa, 2q)}\right) = \lambda_1 c_1^2 \left(\tau_c^{-1} - \tau^{-1}\right)^{2q} + o\left(\left(\tau_c^{-1} - \tau^{-1}\right)^{2q}\right).$$

7 7

Suppose that a)  $\kappa < q+1$ , then equating corresponding powers in  $\tau_c^{-1} - \tau^{-1}$  at both left- and right-hand side yields that  $\kappa = 2q$ . The other alternative b) is that  $\kappa \ge q+1$ , which leads to q+1=2q or q=1 and  $\kappa \ge 2$ .

Similarly, the left-hand side of (13) is of the order

$$\sum_{k=1}^{N} (\lambda_k - \tau^{-1}) \gamma_k^2(\tau) = c_1^2 (\tau_c^{-1} - \tau^{-1})^{1+2q} + \sum_{k=2}^{N} (\lambda_k - \tau^{-1}) c_k^2 (\tau_c^{-1} - \tau^{-1})^{2\kappa} + o\left((\tau_c^{-1} - \tau^{-1})^{\min(2q+1,2\kappa)}\right),$$

while the right-hand side of (13) is of the order

$$\sum_{k=1}^{N} \sum_{q=1}^{N} \sum_{k=1}^{N} \lambda_k \gamma_q(\tau) \gamma_m(\tau) \gamma_k(\tau) \sum_{j=1}^{N} (x_m)_j (x_q)_j (x_k)_j = c_1^3 (\tau_c^{-1} - \tau^{-1})^{3q} \lambda_1 \sum_{j=1}^{N} (x_1)_j^3 + O\left((\tau_c^{-1} - \tau^{-1})^{2q+\kappa}\right).$$

Equating corresponding powers in  $\tau_c^{-1} - \tau^{-1}$  again leads to two cases for the exponents: c) if  $2\kappa < 2q + 1$ , then  $2\kappa = 3q$  and the second-order term shows that 1 + 2q = $2q + \kappa$  or  $\kappa = 1$ , so that  $q = \frac{2}{3}$ , while d) if  $2\kappa \ge 2q + 1$ , then 1 + 2q = 3q, thus q = 1 and  $2\kappa \ge 3$ . But the secondorder term shows that  $2\kappa = 2q + \kappa$ , or  $\kappa = 2q = 2$ . Case c) is inconsistent with either case a) and b), while case d) is consistent with case b). Thus, we conclude that q = 1 and  $\kappa = 2$ . Equating corresponding powers in  $\tau_c^{-1} - \tau^{-1}$  yields that

$$c_1 = \left(\lambda_1 \sum_{j=1}^N (x_1)_j^3\right)^{-1}$$

After substituting the order terms into the fraction of infected nodes (9), we arrive at (11).  $\Box$ 

After transforming  $s = \frac{1}{\tau}$ , the slope at the epidemic threshold follows from (11) as

$$\frac{\mathrm{d}y_{\infty}(s)}{\mathrm{d}s}\bigg|_{s=\lambda_{1}} = -\frac{1}{N} \frac{u^{T}x_{1}}{\lambda_{1}\sum_{j=1}^{N}(x_{1})_{j}^{3}} = -\frac{1}{\lambda_{1}N} \frac{\sum_{j=1}^{N}(x_{1})_{j}}{\sum_{j=1}^{N}(x_{1})_{j}^{3}}.$$
(14)

For example, for  $\tau \downarrow \tau_c$  or  $s \uparrow \lambda_1$ , the steady-state fraction of infected nodes  $y_{\infty}$  in the complete bipartite graph  $K_{nm}$ has the slope

$$\left.\frac{\mathrm{d}y_{\infty}\left(s\right)}{\mathrm{d}s}\right|_{s=\lambda_{1}} = -\frac{1}{N}\frac{u^{T}x_{1}}{\lambda_{1}\sum_{j=1}^{N}\left(x_{1}\right)_{j}^{3}} = -\frac{2}{N}$$

which agrees with the exact formula [4]

$$y_{\infty}(s) = \frac{(mn - s^2)}{N} \left\{ \frac{1}{s+m} + \frac{1}{s+n} \right\},$$
 (15)

while, for regular graphs, where  $\lambda_1 = d$  and  $\frac{u^T x_1}{\sum_{j=1}^N (x_1)_j^3} = N$ , (11) without higher-order terms coincides with (4).

The appearance of the three first moments of the largest eigenvector in Theorem 1 is intriguing. In [3], we show that  $1 \leq u^T x_1 \leq \sqrt{N}$  and that the upper bound is reached in regular graphs. Invoking the Hölder inequality ([3], p. 309) using the normalization  $x_1^T x_1 = 1$  shows that

$$N\left(\sum_{j=1}^{N} (x_1)_j^3\right)^2 \ge 1.$$

An upper bound follows from

$$\sum_{j=1}^{N} (x_1)_j^3 \leqslant \max_{1 \leqslant j \leqslant N} (x_1)_j \sum_{j=1}^{N} (x_1)_j^2 = \max_{1 \leqslant j \leqslant N} (x_1)_j \leqslant 1.$$

For a connected graph, the Perron-Frobenius theorem implies that  $\max_{1 \leq j \leq N} (x_1)_j < 1$ , because all vector components are positive and  $x_1^T x_1 = 1$ . Hence, we find that

$$\frac{1}{\sqrt{N}} \leqslant \sum_{j=1}^{N} (x_1)_j^3 \leqslant \max_{1 \leqslant j \leqslant N} (x_1)_j.$$
 (16)

Since equality (at both sides) in (16) is reached for the regular graph, where  $(x_1)_i = \frac{1}{\sqrt{N}}$ , these bounds are the best possible among all graphs. With these bounds, the slope in (14) at the epidemic threshold lies in between

$$\frac{1}{\lambda_1 N \max_{1 \leqslant j \leqslant N} (x_1)_j} \leqslant - \left. \frac{\mathrm{d} y_\infty(s)}{\mathrm{d} s} \right|_{s=\lambda_1} \leqslant \frac{1}{\lambda_1},$$

where the upper bound is again reached for regular graphs. The more irregular a graph is, the smaller the slope, which agrees with simulations in [14]. Further, the inequality shows that slope decreases at least as fast as  $\tau_c = \frac{1}{\lambda_1}$  when the network size N increases. Since most real-world networks possess an irregular degree distribution close to a power law, both  $\tau_c$  and the slope at the epidemic threshold  $\tau_c$  are small so that simulations in large networks might suggest that  $y_{\infty}(\tau) \simeq \xi(\tau - \tau_c)^{\gamma}$  with  $\gamma > 1$ .

**Discussion.** – Critical parameters at a phase transition [15] often follow power laws. Applied to the virus spread problem, the general form can be written as  $y_{\infty}(\tau) = \xi(\tau - \tau_c)^{\gamma}$ , where  $\xi, \gamma$  and  $\tau_c$  reflect physical properties of the virus spread process and of the graph. Here, the critical exponent equals  $\gamma = 1$  and  $\xi = -\lambda_1^2 \frac{\mathrm{d}y_{\infty}}{\mathrm{d}s}\Big|_{s\uparrow\lambda_1}$ 

specified in (14). In contrast to thermodynamics or Gibbs measures, such as the magnetization, the phase transition in virus spread on networks is entirely different in nature: there are no infinite slopes, nor jumps. Restrepo *et al.* [7] approximately derived the behavior of the order parameter r (instead of the fraction of synchronized oscillators) around the critical threshold  $g_c = \frac{g_0}{\lambda_1}$  as a function of the coupling strength g as

$$r^{2}(g) \simeq r_{0} \frac{\lambda_{1}^{2} \left(\sum_{j=1}^{N} (x_{1})_{j}\right)^{2}}{N \left(d_{\mathrm{av}}\right)^{2} \sum_{j=1}^{N} (x_{1})_{j}^{4}} \left(\frac{g}{g_{c}} - 1\right) \left(\frac{g}{g_{c}}\right)^{-3},$$

where  $g_0$  and  $r_0$  are constants and  $d_{av}$  is the average degree. Heuristically combining the inequalities

 $(\sum_{j=1}^N (x_1)_j^4)^{-1} \leqslant (\sum_{j=1}^N (x_1)_j^3)^{-2}$  and  $\frac{\lambda_1}{d_{\rm av}} \geqslant 1$  may suggest REFER that

$$r^{2}(g) \simeq \frac{r_{0}}{N} \left( \frac{\sum_{j=1}^{N} (x_{1})_{j}}{\sum_{j=1}^{N} (x_{1})_{j}^{3}} \right)^{2} \left( \frac{g}{g_{c}} - 1 \right) \left( \frac{g}{g_{c}} \right)^{-3},$$

where only first- and third-order moments appear as in Theorem 1.

Expression (11) illustrates that, among all graphs with N nodes and L links, the regular graph with degree  $d = \frac{2L}{N}$  has the largest epidemic threshold  $\tau_c = \frac{1}{d}$ , but also the largest (in absolute value) derivative  $\frac{dy_{\infty}}{ds}\Big|_{s\uparrow\lambda_1}$  because of equality in the general relation  $\lambda_1 \ge \frac{2L}{N}$  and in the bounds in (16). This means that a higher effective infection rate  $\tau$  is needed to cause a non-zero steady-state fraction of nodes in the regular graph to be permanently infected, but that, slightly above that critical rate  $\tau_c$ , a higher relative fraction of nodes is infected than in other graphs. In other words, the change in virus conductivity [14] at  $\tau = \tau_c + \epsilon$  is highest in regular graphs.

The behavior around  $\tau = \tau_c$  was the missing part in the general steady-state theory, because an accurate continued fraction expansion around  $\tau \to \infty$  exists [4], whose convergence is, unfortunately, the worst for values of  $\tau$ just above and at  $\tau_c$ . In other words, a few convergents in the continued fraction are insufficient to approximate the behavior of  $y_{\infty}(\tau)$  around  $\tau_c$  accurately, and just the onset of permanent infection in a network is of practical importance because heavy infection  $\tau \to \infty$  hardly occurs due to our knowledge of several immunization strategies [16]. Concavity of  $y_{\infty}(\tau)$  for  $\tau \ge \tau_c$ , proved in [5] bridges the gap between the regime around  $\tau = \tau_c$  and  $\tau \to \infty$ . Finally, as in [17], we propose as future work an accurate comparison<sup>4</sup> between the N-intertwined virus model and the exact SIS Markov chain to assess the relative error of the mean-field approximation around  $\tau = \tau_c$ .

## \* \* \*

This research was supported by Next Generation Infrastructures (Bsik) and the EU FP7 project ResumeNet (project No. 224619).

REFERENCES

- CASTELLANO C. and PASTOR-SATORRAS R., *Phys. Rev.* Lett., **105** (2010) 218701.
- [2] BARRAT A., BARTELEMY M. and VESPIGNANI A., Dynamical Processes on Complex Networks (Cambridge University Press, Cambridge, UK) 2008.
- [3] VAN MIEGHEM P., Graph Spectra for Complex Networks (Cambridge University Press, Cambridge, UK) 2011.
- [4] VAN MIEGHEM P., OMIC J. and KOOIJ R. E., *IEEE/ACM Trans. Netw.*, **17** (2009) 1.
- [5] VAN MIEGHEM P. and OMIC J., In-homogeneous virus spread in networks, Delft University of Technology, report 2008081 (www.nas.ewi.tudelft.nl/people/Piet/ TUDelftReports) 2008.
- [6] STROGATZ S. H, Physica D, 143 (2000) 1.
- [7] RESTREPO J. G., OTT E. and HUNT BRIAN R., Phys. Rev. E, 71 (2005) 036151.
- [8] VAN MIEGHEM P., Computing, 93 (2011) 147.
- [9] GANESH A., MASSOULIÉ L. and TOWSLEY D., The effect of network topology on the spread of epidemics, in IEEE INFOCOM05, Vol. 2 (IEEE) 2005, pp. 1455–1466.
- [10] OMIC J. and VAN MIEGHEM P., Epidemic spreading in networks - variance of the number of infected nodes, Delft University of Technology, report 20090707 (www.nas. ewi.tudelft.nl/people/Piet/TUDelftReports) 2009.
- [11] KEPHART J. O. and WHITE S. R., Direct-graph epidemiological models of computer viruses, in Proceedings of the 1991 IEEE Computer Society Symposium on Research in Security and Privacy (1991) pp. 343–359.
- [12] PASTOR-SATORRAS R. and VESPIGNANI A., Phys. Rev. E, 63 (2001) 066117.
- [13] CHAKRABARTI D., WANG Y., WANG C., LESKOVEC J. and FALOUTSOS C., ACM Trans. Inf. Syst. Secur. (TISSEC), 10 (2008) 1.
- [14] YOUSSEF M., KOOIJ R. E. and SCOGLIO C., Viral conductance: Quantifying the robustness of networks with respect to spread of epidemics, to be published in J. Comput. Sci., doi:10.1016/j.jocs.2011.03.001 (2011).
- [15] DOROGOVTSEV S. N. and GOLTSEV A. V., Rev. Mod. Phys., 80 (2008) 1275.
- [16] CHEN Y., PAUL G., HAVLIN S., LILJEROS F. and STAN-LEY H. E., Phys. Rev. Lett., 101 (2008) 058701.
- [17] CASTELLANO C. and PASTOR-SATORRAS R., Phys. Rev. Lett., 96 (2006) 038701.

 $<sup>^4\,{\</sup>rm The}$  last section in [4] gives a comparison for small networks for which the exact Markov chain with  $2^N$  states can still be evaluated.