Local leaders in random networks

Vincent D. Blondel,[1] Jean-Loup Guillaume,[2] Julien M. Hendrickx,[1] Cristobald de Kerchove,[1] and Renaud Lambiotte[1,3]
[1]*Department of Mathematical Engineering, Université catholique de Louvain, 4 avenue Georges Lemaitre, B-1348 Louvain-la-Neuve, Belgium*
[2]*LIP6 (UMR 7606 CNRS), Université Pierre & Marie Curie, 4 place Jussieu, 75005 Paris, France*
[3]*GRAPES, Université de Liège, Sart-Tilman, B-4000 Liège, Belgium*
(Received 27 July 2007; published 13 March 2008)

We consider local leaders in random uncorrelated networks, i.e., nodes whose degree is higher than or equal to the degree of all their neighbors. An analytical expression is found for the probability for a node of degree k to be a local leader. This quantity is shown to exhibit a transition from a situation where high-degree nodes are local leaders to a situation where they are not, when the tail of the degree distribution behaves like the power law $\sim k^{-\gamma_c}$ with $\gamma_c=3$. Theoretical results are verified by computer simulations, and the importance of finite-size effects is discussed.

DOI: 10.1103/PhysRevE.77.036114 PACS number(s): 89.75.Fb, 89.75.Hc, 87.23.Ge, 05.90.+m

I. INTRODUCTION

In the last few years, the study of networks has received increasing attention from the scientific community [1,2] in disciplines as diverse as biology (metabolic and protein interactions), computer and information sciences (the Internet and the World Wide Web), etc. It has been shown that many empirical networks differ from regular lattices by their random structure and by the heterogeneity of the node properties, i.e., nodes inside a given network may exhibit very different topological properties. The best-known case is node degree heterogeneity, which results in fat-tailed degree distributions where many nodes are sparsely connected while a few nodes, or hubs, receive a large number of links [3]. It is now well known that degree heterogeneity [4,5] and especially the presence of hubs are important factors that may radically alter the propagation of *data*, e.g. rumors [6], opinions [7,8], or viruses [9], and may provoke a network's weakness in face of targeted attacks [10,11].

The important role played by hubs in the above processes has therefore motivated a detailed study of the extremal properties of networks. Different contributions [12,13] have focused on the properties of the degree of the leader (the node with the highest degree), in particular on the probability that the leader never changes, and on related leadership statistics [14]. These approaches, based on the theory of extreme statistics [15], have provided an excellent description of the behavior of the global extrema in the network but, surprisingly, the statistics of local extrema have not been considered yet. There are several reasons, though, to focus on *local leaders*, namely, nodes whose degree is larger than or equal to the degree of their neighbors, and on *strict leaders*, namely, nodes whose degree is strictly larger than the degree of their neighbors (see Fig. 1). Such nodes may be viewed as local hubs that trigger the communication between nodes at the local level. Indeed, individuals usually compare their *state* (*e*.g. opinion, wealth, idea, etc.) with the state of their neighbors, thereby suggesting that a local leader might have a dominant role in its own neighborhood, whatever the absolute value of its connectivity. As a rich among the poor, a local leader might therefore have a more dominant role than as a rich among the richest. From a marketing point of view, for instance, the identification of such nodes might be of interest in order to target nodes that play an important role within *circles of friends* [16]. Let us also stress that local leaders form a subset of nodes that might grasp important characteristics of the whole network and could be helpful in order to visualize its internal features.

In this paper, we focus on the properties of local leaders in uncorrelated infinite and finite random networks, i.e., networks where the degrees of neighboring nodes are not correlated [17]. In Sec. II, we derive an analytical formula for the probability P_k for a node of degree k to be a local leader, and show that this probability undergoes a phase transition where the control parameter is the degree distribution itself [18]. When the tail of the distribution decreases faster than a power law $\sim k^{-\gamma_c}$ with $\gamma_c=3$, the probability to be a local leader goes to 1 for large enough values of k. When the tail of the distribution decreases more slowly than $\sim k^{-\gamma_c}$, in contrast, this probability vanishes for large enough degrees. In Sec. III, we validate our theoretical predictions by computer simulations and show how finite-size effects may affect the above transition. In Sec. IV, finally, we conclude and propose generalizations of the concept of local leader.

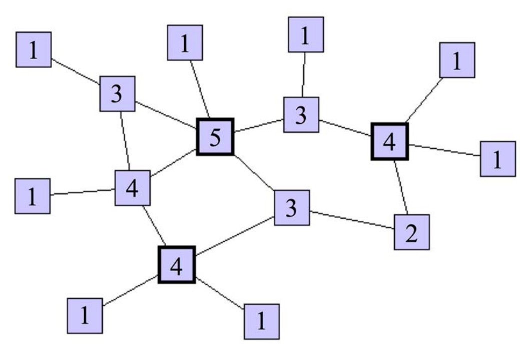

FIG. 1. Sketch of a random network composed of 16 nodes. The network possesses three local leaders; two of them are strict leaders. The numbers inside the vertices represent their degrees.

II. BEING RICH AMONG THE POOR, AND VICE VERSA

Let us consider an undirected random network determined by its degree distribution n_k, i.e., the probability that a randomly chosen node has degree k. By construction, this distribution satisfies the relations

$$\sum_{k=1}^{\infty} n_k = 1, \quad \sum_{k=1}^{\infty} k n_k = z, \tag{1}$$

where $z=2L/N$ is the average degree, N the total number of nodes, and L the total number of links in the network. In the above relations, we have assumed that there are no nodes with degree $k=0$, which is reasonable as such nodes are not part of the network structure.

Let us now evaluate the probability P_k that a node of degree k is a local leader—the case of strict leaders will be briefly discussed at the end of this section. To do so, one first has to look at the probability q_j that a neighbor of the node under consideration has degree j. In a network where the degrees of adjacent nodes are statistically independent, q_j does not depend on the degree k of the local leader, and it is therefore equal to the probability that a randomly chosen link arrives at a node of degree j, so that $q_j = j n_j / z$. One easily verifies that q_j is a probability, i.e., $\Sigma_j q_j = 1$. The probability for this neighbor to have degree $j \le k$ is therefore

$$q_k' = \frac{\sum_{j=1}^{k} j n_j}{z}. \tag{2}$$

By definition, a node with degree k is a local leader if each of its k neighbors has a degree smaller than or equal to k. By using the statistical independence of the degrees of these k neighbors, P_k is found by multiplying Eq. (2) k times,

$$P_k = \left(\frac{\sum_{j=1}^{k} j n_j}{z} \right)^k. \tag{3}$$

In general, P_k is a function of k whose behavior may be evaluated numerically by inserting the degree distribution n_k of the network in Eq. (3) and by performing the summations. In the following, however, we would like to derive general properties of P_k that do not depend on the details of n_k. To do so, let us only focus on the asymptotic behavior of P_k, when k is large, and assume that n_k may be approximated for large enough values of k by a power law $n_k = C k^{-\gamma}$, where C is a normalization constant. The case of pure power laws where $n_k = C k^{-\gamma}$ for all k will be detailed later on.

Let us emphasize that such a tail of the degree distribution is a very general behavior, as it includes scale-free distributions (γ finite), while exponential distributions are recovered in the limit $\gamma \to \infty$. In the following, we focus on general values of γ, with the sole constraint that $\gamma > 2$ so that the average degree is well defined. In that case, $\Sigma_{j=1}^{\infty} j n_j = z$ is a finite number and Eq. (3) may be rewritten as

$$P_k = \left(1 - \frac{\sum_{j=k+1}^{\infty} C j^{-(\gamma-1)}}{z} \right)^k, \tag{4}$$

where we used the fact that $\Sigma_{j=1}^{k} j n_j = \Sigma_{j=1}^{\infty} j n_j - \Sigma_{j=k+1}^{\infty} j n_j$.

For large enough values of k, the summation in (4) may be replaced by an integral so that P_k asymptotically behaves as

$$P_k \approx \left(1 - \frac{C k^{-(\gamma-2)}}{(\gamma-2) z} \right)^k. \tag{5}$$

In order to determine the asymptotic behavior of P_k, it is useful to rewrite Eq. (5) as

$$P_k \approx e^{k \ln[1 - C k^{-(\gamma-2)}/(\gamma-2) z]}, \tag{6}$$

whose dominating term is, when $k^{-(\gamma-2)}$ is sufficiently small,

$$P_k \approx e^{-C k^{-(\gamma-3)}/(\gamma-2) z}. \tag{7}$$

By construction, $\gamma > 2$ and z is positive, so that the asymptotic values of P_k, for large enough values of k, are

$$P_k \underset{k\to\infty}{\longrightarrow} \begin{cases} 1 & \text{for} \quad \gamma > 3, \\ e^{-C/z} & \text{for} \quad \gamma = 3, \\ 0 & \text{for} \quad \gamma < 3. \end{cases} \tag{8}$$

The system therefore undergoes a transition at $\gamma = 3$. If the tail of the degree distribution decreases fast enough, so that $\gamma > 3$, the probability P_k asymptotically goes to 1. Consequently, nodes with a higher degree have a larger probability to be local leaders. When $\gamma < 3$, in contrast, the probability to be a local leader decreases with increasing degree k and asymptotically vanishes, so that, surprisingly, nodes with a larger degree might have a smaller probability to be local leaders.

This result, which may appear intriguing at first sight, can be explained by analyzing the competition between two trends. On one hand, a node with a high degree has a higher probability of having a higher degree than any other particular node, which tends to increase its probability of being a degree leader. On the other hand, a node with a higher degree has more neighbors, which tends to decrease the probability of having a higher degree than all its neighbors [see the exponent k in Eq. (3)]. Depending on the value of γ, the asymptotic behavior is dictated by the first or the second phenomenon, with a transition when $\gamma = 3$, where an equilibrium occurs.

One should also note that the above calculations simplify in term of harmonic functions $H(k,\gamma) \equiv \Sigma_{i=1}^{k} i^{-\gamma}$, when the degree distribution is a pure power law $n_k = C k^{-\gamma}$ for all k, where $C = 1/\Sigma_{k=1}^{\infty} k^{-\gamma} = 1/H(\infty,\gamma)$. Indeed, in that case, the probability to be a local leader P_k reads

$$P_k = \left(\frac{\sum_{j=1}^{k} j^{-(\gamma-1)}}{\sum_{j=1}^{\infty} j^{-(\gamma-1)}} \right)^k = \left(\frac{H(k,\gamma-1)}{H(\infty,\gamma-1)} \right)^k. \quad (9)$$

Using the asymptotics of the harmonic numbers [19]

$$H(k,\gamma-1) = H(\infty,\gamma-1) - \frac{k^{-(\gamma-2)}}{(\gamma-2)}, \quad (10)$$

valid when $\gamma > 2$, it is straightforward to recover the transition (8) where $e^{-C/z}$ is now given by $e^{-1/H(\infty,2)} = e^{-6/\pi^2}$, since $z = H(\infty,\gamma-1)/H(\infty,\gamma)$.

Before going further, let us discuss the case of strict leaders. In that case, the calculations are the same as previously, except that the sums in P_k do not go to k but to $k-1$. However, this difference is vanishingly small for large enough values of k, so that the transition (8) is recovered.

III. SIMULATIONS AND FINITE-SIZE EFFECTS

In this section, we verify the validity of the theoretical predictions (3) and, especially, the existence of the regime $P_k \to 0$ when $\gamma < 3$. One should first stress that the results derived in the previous section are valid for uncorrelated networks composed of an infinite number of nodes. However, whatever the specified degree distribution n_k, a typical realization of the network (in a computer simulation or in a realistic situation) involves only a finite number of nodes. This also implies that the largest degree k_{max} in the network is a finite number. The degree k_{max} of this global leader might be estimated by using tools from the theory of extreme statistics [15], but the main point here is that the global leader is also a local leader. Consequently, the probability for a node of degree k_{max} to be a local leader, when measured in such a system, is $P_{k_{max}} = 1$, in contradiction with the prediction $P_k \to 0$.

In order to highlight this finite-size effect with simulations, it is helpful to consider the truncated power laws defined by

$$n_k = \begin{cases} Dk^{-\gamma} & \text{for} \quad k \le k_{max}, \\ 0 & \text{otherwise}, \end{cases} \quad (11)$$

where the normalization constant D depends on γ and on the cutoff k_{max}, $D = 1/\sum_{k=1}^{k_{max}} k^{-\gamma}$. Such degree distributions offer the possibility to tune the value of the extremal degree k_{max} together with a particularly simple expression for n_k. To generate numerically random uncorrelated networks with the specified degree distribution (11), we proceed as follows [20]. We assign to each node i in a set of N nodes a degree k_i sampled from the probability distribution (11) and impose that $\sum_{i=1}^{N} k_i$ is even. Then the network is constructed by randomly assigning the $L = \sum_{i=1}^{N} k_i/2$ edges while respecting the preassigned degrees k_i. In the simulations, we have considered networks with $N = 10^5$ nodes and averaged the results over 100 realizations of the random process. One should stress that we have considered only truncated distributions such that k_{max} is effectively the maximum degree for each realization of the network, i.e., such that the expected number of nodes with k_{max} verifies $Nn_{k_{max}} \ge 1$. Computer simulations (see Fig. 2) show an excellent agreement with the theoretical prediction (3) and confirm that P_k decreases to values close to 0 when $\gamma < 3$, as predicted by (8), before increasing to 1 due to finite-size effects. When $\gamma > 3$, in contrast, P_k directly increases to 1 (see Fig. 3), as expected. Simulation results are also in perfect agreement with the theoretical prediction (3) in that case.

The above method ensures that the realized network is uncorrelated, even when $\gamma < 3$. Indeed, for small values of k_{max} (i.e., for a maximum degree that scales at most as $N^{1/2}$),

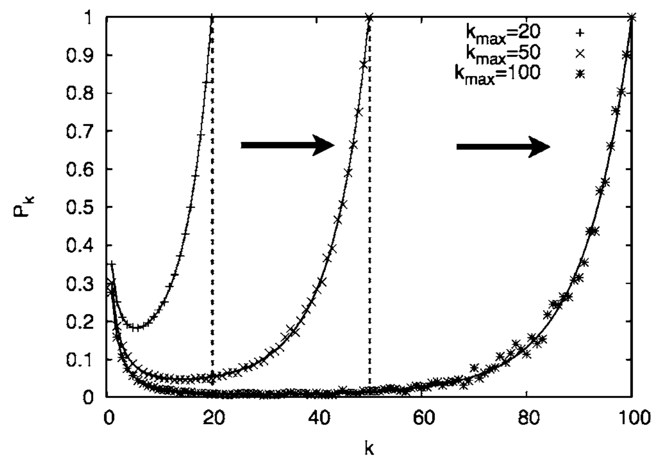

FIG. 2. P_k measured in random networks composed of 10^5 nodes and whose degree distribution is a truncated power law (11) with $\gamma = 2.2$. The results are averaged over 100 realizations. The solid lines are the theoretical prediction (3), evaluated numerically for the degree distributions (11). The value of k where P_k begins to increase toward $P_k = 1$ due to finite-size effects (see main text) is seen to be proportional to k_{max}.

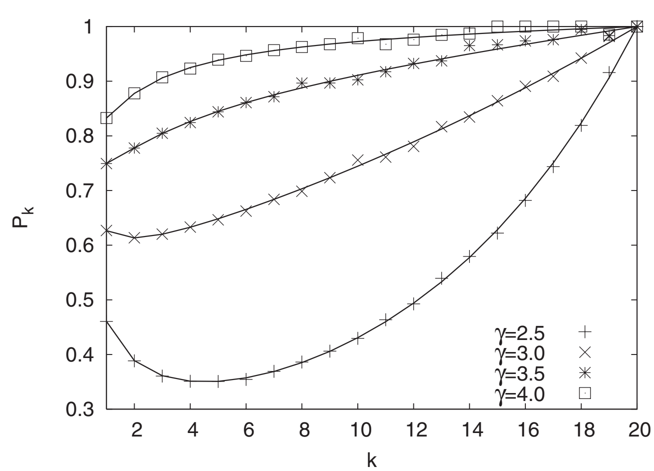

FIG. 3. P_k measured in random networks composed of 10^5 nodes and whose degree distribution is a truncated power law (11) with $\gamma = 2.5$, 3.0, 3.5 and 4.0, respectively. The results are averaged over 100 realizations and $k_{max} = 20$. The solid lines are the theoretical prediction (3), evaluated numerically for the degree distributions (11).

it is known [20–24] that the probability for self-loops or multiple edges to occur is negligible and that the degrees of neighboring nodes are uncorrelated. We have verified this absence of correlations in our networks by measuring the assortativity coefficient [17], which is vanishingly small. For larger values of $k_{\max}$, in contrast, the network densifies and exhibits disassortative correlations. It would have been interesting to look for discrepancies between simulation results and the theoretical prediction (3), thereby highlighting how correlations affect the probability of a node to be a local leader. Unfortunately, for such networks, the maximum degree of the network fluctuates from one realization to another, which implies that the degree distribution is not a continuous function (for large degrees) and that it cannot be approximated by the average degree distribution. Thus the theoretical prediction (3) ceases to be valid in that case. In order to highlight the role of correlations, one therefore needs a method that generates networks with correlations and whose degree distribution does not fluctuate at each realization. A possibility would be to look at growing networks, e.g., the Barabási-Albert model [3,25], where the degree of the nodes would be bounded by some maximum value $k_{\max}$, e.g. a node ceases to receive links if its degree is equal to $k_{\max}$. Such an analysis, however, goes beyond the scope of this paper.

Let us now return to the uncorrelated case when $\gamma<3$. In order to evaluate where finite-size effects become non-negligible, we have focused on the value k_c where P_k is minimum (see Fig. 2) and we have studied the relation between k_c and $k_{\max}$. By inserting the distribution (11) and integrating (3) numerically, one observes that k_c increases linearly with $k_{\max}$, $k_c \approx \alpha k_{\max}$. When $\gamma=2.2$, for instance, one finds $\alpha=0.3189$. This linear dependence has important consequences as it implies that finite-size effects affect only a vanishingly small number of the nodes when $k_{\max}$ is sufficiently large. To show this, let us consider the proportion n_{FS} of nodes affected by the finite-size effects,

$$\begin{aligned} n_{\mathrm{FS}} &= \sum_{k=\alpha k_{\max}}^{k_{\max}} D k^{-\gamma} \\ &\approx \int_{k=\alpha k_{\max}}^{k_{\max}} D k^{-\gamma} dk \\ &= \frac{D}{\gamma-1}(\alpha^{-(\gamma-1)}-1) k_{\max}^{-(\gamma-1)}, \end{aligned} \tag{12}$$

where the summation has been replaced by an integral, as $k_{\max}$ is sufficiently large. The quantity n_{FS} obviously goes to zero when $k_{\max}\to\infty$. Let us note that this limit makes sense only when $N\to\infty$ and that the maximum degree also has to satisfy $k_{\max}<N^{1/2}$ in order to ensure that the network is uncorrelated [20].

Before concluding, let us also derive the behavior of P_k close to $k_{\max}$. In that case, numerical integration shows an exponential decrease in $(k_{\max}-k)$ so that one looks for a solution of the form

$$P_k \approx e^{E(k_{\max}-k)}, \tag{13}$$

where the constant E is found by comparing (13) with

$$P_k = \exp\left[k \ln\left(1 - \sum_{j=k+1}^{k_{\max}} D j^{-(\gamma-1)}/z\right)\right], \tag{14}$$

and by looking at the dominant terms for small values of $k' \equiv k_{\max}-k$. When $k_{\max}$ is sufficiently large, it is straightforward to show that

$$E \approx k_{\max} \ln(1) + k_{\max-1} \ln(1 - D k_{\max}^{-(\gamma-1)}/z) \approx -D k_{\max}^{-(\gamma-2)}/z. \tag{15}$$

This asymptotic behavior has been successfully compared with simulations.

IV. CONCLUSION

In this paper, we have analyzed the statistical properties of local leaders in uncorrelated networks. Such nodes, which may be viewed as local hubs, have a crucial location in a social or information network, as they dominate all their neighbors. Their identification and a better understanding of their properties might therefore be of practical interest. In marketing, for instance, local leaders are good candidates to target in order to maximize a marketing campaign or to minimize the erosion of customers from a company, e.g. to *churn* for mobile operators [26]. We have observed that the probability for a node of degree k to be a local leader undergoes a transition from a rich is rich to a rich is poor situation, which suggests that nodes with a high degree might not be the most influential at the local level. It is interesting to stress that the transition takes place at a realistic value of the power-law exponent $\gamma_c=3$ [27,28], i.e., scale-free distributions usually have an exponent between 2 and 3 [29], and that $\gamma_c=3$ is also the critical value under which the variance diverges. To conclude, one should stress that the local maxima of other node quantities could also give insight into the network structure, e.g., the number of triangles [16]. More general definitions of local leaders could also be considered, e.g. a node of degree k is an α leader if all of its neighbors have degree $k' < k/\alpha$. A generalization of our study to such situations and a comparison with empirical data (where nodes might exhibit degree-degree correlations) could therefore be of interest.

ACKNOWLEDGMENTS

V.B., J.-L.G., J.H., and C.d.K. acknowledge support from the Concerted Research Action (ARC) “Large Graphs and Networks” from the “Direction de la recherche scientifique—Communauté française de Belgique,” from the EU HYCON Network of Excellence (Contract No. FP6-IST-511368), and from the Belgian Program on Interuniversity Attraction Poles initiated by the Belgian Federal Science Policy Office. J.H. is supported by FNRS (Belgian Fund for Scientific Research) R.L. has been supported by European Commission Project No. CREEN FP6-2003-NEST-Path-012864.

[1] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).
[2] A.-L. Barabási, *Linked* (Perseus Publishing, Cambridge, MA, 2002).
[3] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[4] M. Boguñá, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. Lett. **90**, 028701 (2003).
[5] V. Sood and S. Redner, Phys. Rev. Lett. **94**, 178701 (2005).
[6] J. Leskovec, L. A. Adamic, and B. A. Huberman, Proceedings of the 7th ACM Conference on Electronic Commerce (ACM, Ann Arbor, MI, 2006), pp. 228–237, http://doi.acm.org/10.1145/1134707.1134732
[7] S. Galam, Physica A **274**, 132 (1999).
[8] R. Lambiotte, Europhys. Lett. **78**, 68002 (2007).
[9] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. Lett. **85**, 5468 (2000).
[10] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).
[11] R. M. May and A. L. Lloyd, Phys. Rev. E **64**, 066112 (2001).
[12] A.A. Moreira, J.S. Andrade, Jr., and L. A. Nunes Amaral, Phys. Rev. Lett. **89**, 268703 (2002).
[13] P. L. Krapivsky and S. Redner, Phys. Rev. Lett. **89**, 258703 (2002).
[14] T. Łuczak, Random Struct. Algorithms **1**, 287 (1990); P. Erdős and T. Łuczak, *ibid.* **5**, 243 (1994).
[15] J. Galambos, *The Asymptotic Theory of Extreme Order Statistics* (Krieger, Malabar, FL, 1987).
[16] C. de Kerchove, E. Huens, P. Van Dooren, and V. Blondel, *Positive Systems*, Lecture Notes in Control and Information Sciences Vol. 341 (Springer, Berlin, (2006), p. 231.
[17] M. E. J. Newman, Phys. Rev. Lett. **89**, 208701 (2002).
[18] H. E. Stanley, *Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press, Oxford, 1971).
[19] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science* (Addison-Wesley, Reading, MA, 1989).
[20] M. Catanzaro, M. Boguñá, and R. Pastor-Satorras, Phys. Rev. E **71**, 027103 (2005).
[21] J. Park and M. E. J. Newman, Phys. Rev. E **68**, 026112 (2003).
[22] S. Maslov, K. Sneppen, and A. Zaliznyak, Physica A **333**, 529 (2004).
[23] Y. Moreno and A. Vazquez, Eur. Phys. J. B **31**, 265 (2003).
[24] D. Garlaschelli and V. Loffredo, e-print arXiv:cond-mat/0609015.
[25] P. L. Krapivsky and S. Redner, Phys. Rev. E **63**, 066123 (2001).
[26] W.-H. Au, K. C. C. Chan, and X. Yao, IEEE Trans. Evol. Comput. **7**, 532 (2003).
[27] P. L. Krapivsky and S. Redner, Phys. Rev. E **71**, 036118 (2005).
[28] R. Lambiotte and M. Ausloos, Europhys. Lett. **77**, 58002 (2007).
[29] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).