Identification of network modules by optimization of ratio association

L. Angelini
*TIRES-Center of Innovative Technologies for Signal Detection and Processing, University of Bari,
70126 Bari, Italy; Dipartimento Interateneo di Fisica, 70126 Bari, Italy; and Istituto Nazionale di Fisica
Nucleare, Sezione di Bari, 70126 Bari, Italy*

S. Boccaletti
*CNR-Istituto dei Sistemi Complessi, Via Madonna del Piano, 10, 50019 Sesto Fiorentino (FI), Italy
and Italian Embassy in Tel Aviv, Trade Tower, 25 Hamered Street, Tel Aviv, 68125 Israel*

D. Marinazzo, M. Pellicoro, and S. Stramaglia
*TIRES-Center of Innovative Technologies for Signal Detection and Processing, University of Bari,
70126 Bari, Italy; Dipartimento Interateneo di Fisica, 70126 Bari, Italy; and Istituto Nazionale di Fisica
Nucleare, Sezione di Bari, 70126 Bari, Italy*

We introduce a novel method for identifying the modular structures of a network based on the maximization of an objective function: the ratio association. This cost function arises when the communities detection problem is described in the probabilistic autoencoder frame. An analogy with kernel k-means methods allows us to develop an efficient optimization algorithm, based on the deterministic annealing scheme. The performance of the proposed method is shown on real data sets and on simulated networks. © *2007 American Institute of Physics*. [DOI: 10.1063/1.2732162]

The structure of a complex network may be described by identifying the modules of which it is composed. The concept of the module is qualitative: nodes are more connected within their modules than with the rest of the network, and its quantification is still a subject of debate. Modularity, a quantity related to the correlation between the probability of having an edge joining two sites and the fact that the sites belong to the same modules, has been widely accepted as a measure for module identification. Here we provide a new description of this important problem. We analyze the use of a novel objective function, the ratio association, measuring the coherency between modules. Ratio association emerges in the probabilistic autoencoder frame, performing a lossy compression of the network's structures. An analogy to kernel k-means allows the development of an efficient algorithm for the optimization of ratio association. The power of the proposed technique is assessed by showing the structures found by ratio optimization on real datasets and on simulated networks. The likelihood of the probabilistic autoencoder model may be used to select the optimal number of modules.

I. INTRODUCTION

A hierarchical structure of modules characterizes the topology of most real-world network systems.[1] In social networks, for instance, these modules are densely connected groups of individuals belonging to social communities. *Modules* (also called *community structures*) are defined as tightly connected subgraphs of a network, i.e., subsets of nodes *within* which the density of links is very high, while *between* which connections are much sparser. These tight-knit modules constitute units that separately (and in parallel) contribute to the collective functioning of the network. For instance,

the presence of subgroups in *biological* and *technological* networks is at the basis of their functioning. Hence the issue of detecting and characterizing module structures in networks received a considerable amount of attention.

Rigorously, the identification of the hierarchy of modules of a network is equivalent to the *graph partitioning* problem in computer science, which is known to be a nondeterministic polynomial time-complete problem.[2] A series of efficient heuristic methods has been proposed over the years to cope with this problem. These include methods based on *spectral analysis*,[3] or *hierarchical clustering methods* developed in the context of social networks analysis.[4] Among the different techniques, we recall the modularity identification based on the statistical properties of a system of spins,[5] and hierarchical clustering techniques exploiting the central concept of modularity.[6,7] The modularity Q is a measure of the correlation between the probability of having an edge joining two sites and the fact that the sites belong to the same modules (see Ref. 7 for the mathematical definition of Q). Methods directly based on the optimization of Q have been proposed,[8,9] while recently a spectral technique has been introduced[10] exploiting the information of the *modularity matrix*, which, for a given graph, has the property of being dense even when the adjacency matrix is sparse.

Furthermore, another recent stream of research has been initiated by the relevant observation that topological hierarchies are associated with dynamical time scales in the transient of a synchronization process.[11] Such an observation inspired the introduction of a fast technique able to detect and identify the modules of a complex network from the cluster desynchronization scenario of phase oscillators.[12]

In this paper, we introduce a new technique for modules identification, which efficiently optimizes an objective function called *ratio association*. Precisely, once the number of

modules n_c is fixed, the optimization process leads to a fast (in a time scaling linearly with the number of nodes N in the network) detection of the corresponding modules. The efficiency of the algorithm is due to an equivalence with the kernel k-means,[13] which we exploit in the deterministic annealing frame.[14] We show that the optimization of ratio association may be motivated in the probabilistic autoencoder frame, a paradigm that has been used to derive cost functions for data clustering;[15] the same cost function has been used in Ref. [16] for classification of time-series data. In order to select the number of modules n_c, the quality of the solution is to be assessed. This can be achieved in two ways, on the basis of the modularity of the solution, or according to the likelihood of the probabilistic autoencoder model.

The paper is organized as follows. In the next section, we describe our method, while in Sec. III some applications are shown. Section IV summarizes our conclusions.

II. THE METHOD

Given a set of data vectors $\{\mathbf{x}_i\}_{i=1}^N$, with $\mathbf{x}_i \in \mathbb{R}^n$, the goal of the kernel k-means is to find a q-way disjoint partition[17] $\{\pi_c\}_{c=1}^q$ of the data (where π_c represents the cth cluster) such that the following objective function is minimized:

$$\mathcal{I}(\{\pi_c\}_{c=1}^q) = \sum_{c=1}^q \sum_{\mathbf{x}_i \in \pi_c} \|\Phi(\mathbf{x}_i) - \mathbf{m}_c\|^2, \qquad (1)$$

where

$$\mathbf{m}_c = \frac{\Sigma_{\mathbf{x}_i \in \pi_c} \Phi(\mathbf{x}_i)}{|\pi_c|}. \qquad (2)$$

Here, $|\pi_c|$ is the cardinality of the subset π_c, and Φ is a function mapping the x vectors onto a (generally) higher-dimensional space (if Φ is the identity function, the above equations recover the standard definition of k-means).

Expanding the distance term $\|\Phi(\mathbf{x}_i) - \mathbf{m}_c\|^2$ in the objective function, one obtains

$$\Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_i) - \frac{2\Sigma_{\mathbf{x}_j \in \pi_c} \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)}{|\pi_c|}$$
$$+ \frac{\Sigma_{\mathbf{x}_j \in \pi_c} \Sigma_{\mathbf{x}_\ell \in \pi_c} \Phi(\mathbf{x}_j) \cdot \Phi(\mathbf{x}_\ell)}{|\pi_c|^2}. \qquad (3)$$

Notice that in Eq. (3), all computations involving data points are in the form of inner products. As a result, one can use the *kernel trick*: if one can compute the dot product $K_{ij} = \Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x}_j)$ efficiently, then one is able to compute distances between points in this mapped space without having to explicitly know the mapping of $\mathbf{x}_i$ and $\mathbf{x}_j$ onto $\Phi(\mathbf{x}_i)$ and $\Phi(\mathbf{x}_j)$, respectively. It is known that *any* positive semidefinite matrix K can be thought of as a kernel matrix.[18] Using the kernel matrix, Eq. (1) can be rewritten as

$$\mathcal{I}(\{\pi_c\}_{c=1}^q) = \sum_{c=1}^q \sum_{i \in \pi_c} \left(K_{ii} - \frac{2\Sigma_{j \in \pi_c} K_{ij}}{|\pi_c|} + \frac{\Sigma_{j \in \pi_c} \Sigma_{\ell \in \pi_c} K_{j\ell}}{|\pi_c|^2} \right).$$
$$(4)$$

Suppose that the graph $G = (V, A)$ is given, where V is the set of N vertices and A is the adjacency matrix [the elements A_{ij} are one (zero) whenever an edge is (is not) present between vertices i and j]. If $\mathcal{A}$ and $\mathcal{B}$ are two disjoint subsets of V, we furthermore define links$(\mathcal{A}, \mathcal{B}) = \Sigma_{i \in \mathcal{A}, j \in \mathcal{B}} A_{ij}$.

The idea is to fix the number of modules n_c into which we want to efficiently partition the original graph, and to look for the n_c-way disjoint partition of V ($\{\pi_c\}_{c=1}^{n_c}$), which maximizes the following objective function, called *ratio association*:[19]

$$\mathcal{R}(\{\pi_c\}_{c=1}^{n_c}) = \sum_{c=1}^{n_c} \frac{\text{links}(\pi_c, \pi_c)}{|\pi_c|}. \qquad (5)$$

Let us now associate to the given graph a $N \times N$ kernel matrix as follows:

$$K = \sigma I + A, \qquad (6)$$

where I is the identity matrix and σ is a real number chosen to be sufficiently large so that K comes out to be positive definite. Now, given a n_c-way disjoint partition $\{\pi_c\}_{c=1}^{n_c}$ of the graph, the corresponding value of the ratio association and the objective function of kernel k-means are related as follows:

$$\mathcal{I}(\{\pi_c\}_{c=1}^{n_c}) = (N - n_c)\sigma - \mathcal{R}(\{\pi_c\}_{c=1}^{n_c}). \qquad (7)$$

An important point follows: $\mathcal{I}$ attains its minimum in correspondence with the same partition providing the maximum of $\mathcal{R}$, independently of σ, as was shown in Ref. [20] when considering the standard iterations of k-means. Therefore, the kernel k-means minimization may be straightforwardly used to find the n_c optimal clustering of the graph, by simply maximizing the ratio association. The ratio association may be derived in the probabilistic autoencoder frame as described in the Appendix.

Hence, we can use graph clustering to discover modules structures. As we deal here with modular structures maximizing the ratio association, in the following we will handle the optimization problem by deterministic annealing.

Let ρ_{ic} be the probability that vertex i belongs to the cth module. We write

$$\rho_{ic} = \frac{e^{-\beta \xi_{ic}}}{\sum_{c'=1}^{n_c} e^{-\beta \xi_{ic'}}}, \qquad (8)$$

where, according to (4),

$$\xi_{ic} = K_{ii} - \frac{2\Sigma_{j=1}^N K_{ij}\rho_{jc}}{\Sigma_{m=1}^N \rho_{mc}} + \frac{\Sigma_{j,\ell=1}^N K_{j\ell}\rho_{jc}\rho_{\ell c}}{(\Sigma_{m=1}^N \rho_{mc})^2}, \qquad (9)$$

and K is given by (6).

Starting from a random configuration of $\{\rho\}$ and $\{\xi\}$, Eqs. (8) and (9) are solved iteratively while exponentially increasing β. At large β, $\rho_i(c)$ are all zero except for one element providing the module to whom the vertex i has to be assigned.

Notice that the annealing procedure leads to a final partition of vertices that still has a tiny dependence on the starting configuration, hence the algorithm is to be run several times, selecting the partition leading to the lowest value of $\mathcal{I}$.
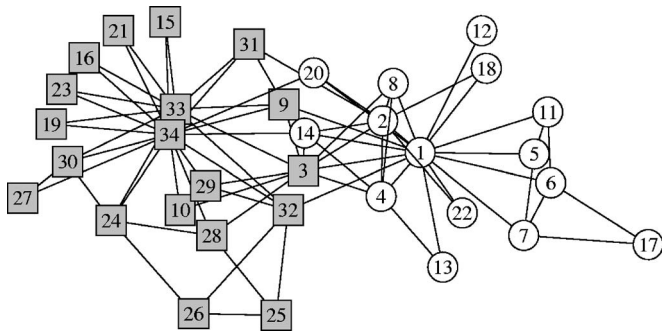
FIG. 1. The Zachary karate club network. The two modules identified by the proposed algorithm are colored in gray and white, respectively. Squares and circles indicate the two real communities described by Zachary (Ref. 21). Notice that our technique fully reveals the true subdivision.
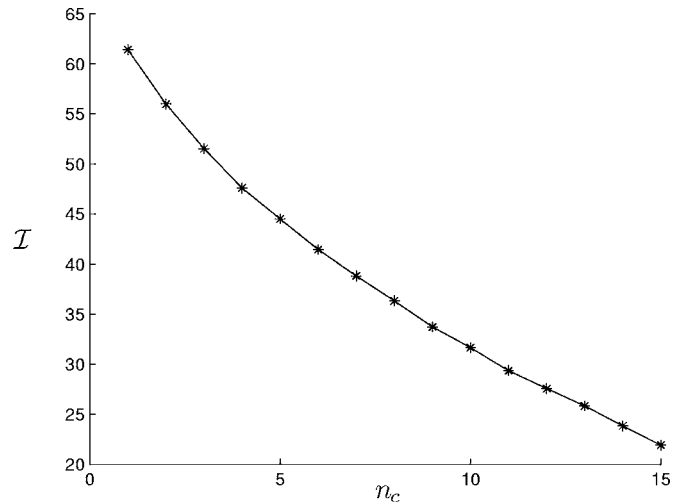

FIG. 2. The objective function $\mathcal{I}$ for the karate club modular structure found by the proposed algorithm vs the number of communities n_C. Here $\sigma=2$, but results are stable against variations of σ.

As is typical of deterministic annealing approaches, the complexity of the algorithm is $\mathcal{O}(n_c N \langle z \rangle)$, where $\langle z \rangle$ is the average number of edges per vertex. Note, however, that in the proposed method, the number of modules is to be specified in advance. Therefore, for a full hierarchical description of the original network (i.e., when one wants also to determine the optimal n_c), the algorithm has to be run by varying n_c between its minimum (2) and its maximum (N) value, leading to an overall complexity $\mathcal{O}(N^3 \langle z \rangle)$.

The choice of σ deserves few comments. As said before, to enforce the positive definiteness of K, and thus to establish the connection to kernel k-means, σ must be sufficiently large. However, since varying σ does not change the global optimum, one may choose σ in the most convenient way from the computational point of view, even though K will not be ensured to be positive definite.

III. APPLICATIONS

Let us first discuss the application of the proposed method to the well-known Zachary karate club network,[21] shown in Fig. 1. We point out that the output of the algorithm is independent of $\sigma \in [0\,20]$ (notice that K, in this case, is positive for $\sigma > 5$).

When selecting $n_c=2$ (i.e., when trying to split the network in two modules), Fig. 1 shows that one fully recovers the true subdivision of the data set, with $\mathcal{R}=8.0139$ and modularity $Q=0.3715$.

When repeating the analysis at varying n_c, Fig. 2 reports the value of $\mathcal{I}$, corresponding to the solution, as a function of n_c: it is a strictly decreasing function. In Fig. 3, we plot the modularity Q of the solution versus n_C: the maximum is $Q=0.420$ and corresponds to a partition of the graph into four modules, in perfect agreement with the outcome of other techniques previously tested on the Zachary karate club network. Finally, Fig. 4 reports the ratio association $\mathcal{R}$ versus n_C, making evident the validity of Eq. (7).

The selection of n_c may also be done on the basis of the average log-likelihood of the autoencoder (see the Appendix). In Fig. 5, we plot $L-L_0$ versus n_c, where L is the average log-likelihood of the data set, while L_0 is the same quantity evaluated on a network with the same number of nodes

and links but with links randomly assigned to pairs of nodes.[22] According to the criterion of the largest $\Delta L = L - L_0$, both $n_c=2$ and 4 are suitable partitions.

To evaluate the performance of the proposed technique, we generate a set of random graphs featuring a well defined modular structure. Precisely, all graphs are generated with $N=128$ nodes and $K=1024$ edges. The nodes are distributed into four modules, containing 32 nodes each. Pairs of nodes belonging to the same module (to different modules) are linked with probability p_{in} (p_{out}). p_{out} is taken so that the average number z_{out} of edges a node forms with members of other communities can be controlled (in our trials z_{out} has been varied between 0 and 10). p_{in} is chosen so as to maintain a constant total average node degree $\langle k \rangle = 16$. Notice that, as z_{out} increases, the modular structure of the network becomes weaker and harder to identify. As the *real* modular structure is directly imposed by the generation process here, the performance of the identification method can be assessed
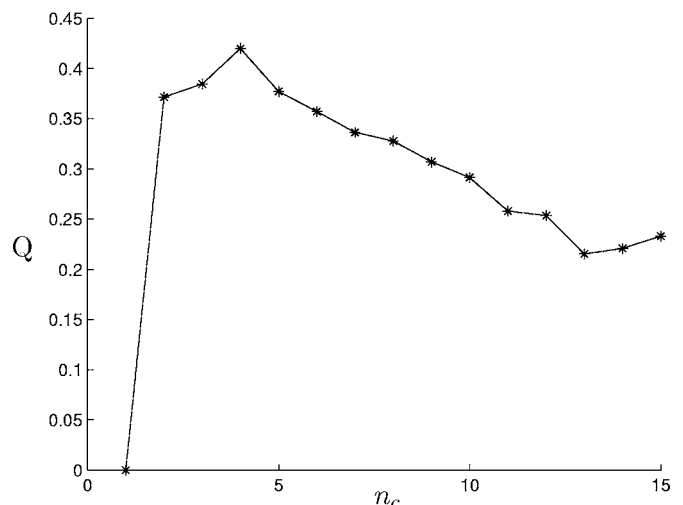

FIG. 3. The modularity Q for the karate club modular structure found by the proposed algorithm vs n_c. As in Fig. 2, we use here $\sigma=2$, but results are stable against variations of σ.
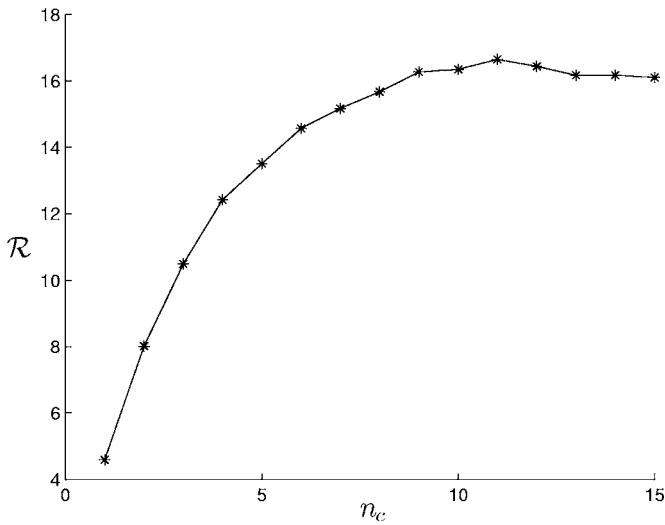
FIG. 4. The ratio association $\mathcal{R}$ for the karate club modular structure found by the proposed algorithm vs n_c. Same stipulations on σ as in the captions of Figs. 2 and 3.

FIG. 6. The fraction p of correctly classified nodes is plotted vs z_{out}, the average number of edges a node forms with members of other modules, for the proposed algorithm (stars), for the Girvan and Newman method (Ref. 7) (empty squares) and for the method introduced by Duch and Arenas (Ref. 9) (empty circles). Each point refers to an ensemble average over 100 different network realizations. Same stipulations for $\sigma=2$ as in the captions of Figs. 2–4.

by monitoring the fraction p of correctly classified nodes versus z_{out}. In Fig. 6, we report a comparative analysis of p versus z_{out} for the proposed algorithm, for the Girvan and Newman method,[7] and for the method introduced by Duch and Arenas.[9] The result is that the accuracy attained by our method comes out to be slightly better than that of Ref. 9. Note that while applying our algorithm to these networks, we have selected n_c by maximizing $L-L_0$ (similar results are obtained maximizing the modularity, see Fig. 7).

While dealing with medium-large real networks, finding the number of communities is a computationally heavy task. An efficient strategy to cope this problem consists of repeated division into two communities, a suboptimal strategy to optimize the ratio association. We use our algorithm first to divide the network into two parts, then divide those parts, and so forth: the overall complexity is $\mathcal{O}(N^2\langle z\rangle)$ in this case. Note that after dividing the network in two, we delete the edges falling between the two parts and then apply the algorithm to each subgraph. This is correct, for ratio association
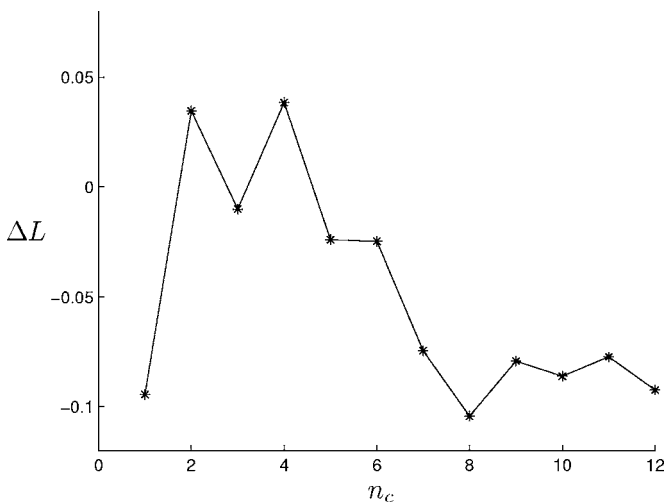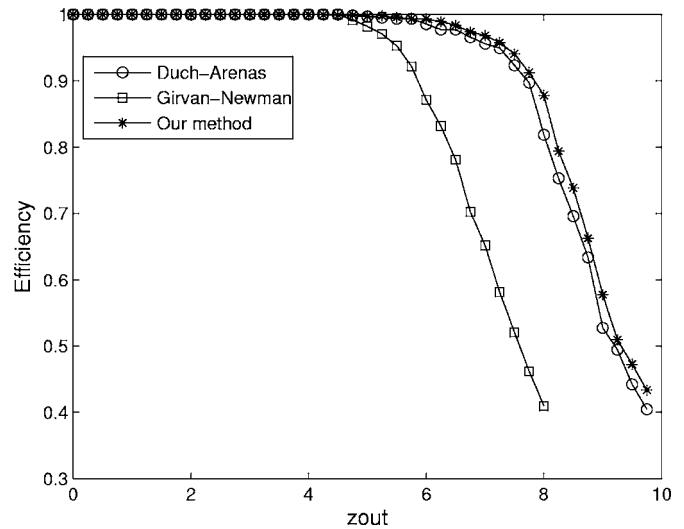
optimization, as the degree of nodes does not appear in the definition of ratio association. Moreover, we adopt the prescription that, at any stage, we leave undivided communities with size smaller than a threshold S_{th}. To show the application of this strategy, we consider here the Pretty Good Privacy (PGP) web of trust social network,[25] containing 10680 nodes. The maximum modularity of this dataset, obtained by the method,[26] is 0.7329 and it is achieved at $n_c=80$. Using the method described in Ref. 9, the maximum modularity is 0.8459 and corresponds to $n_c=365$. In Fig. 8, we depict $L-L_0$, as found by our algorithm, versus n_c. In accordance with the method,[9] we find that the optimal number of communities, maximizing ΔL, is $n_c=360$ (local maxima are found also at $n_c=100$ and 190). Having thus determined the

FIG. 5. $\Delta L = L - L_0$ (see the text) is plotted vs n_c for the Zachary network.
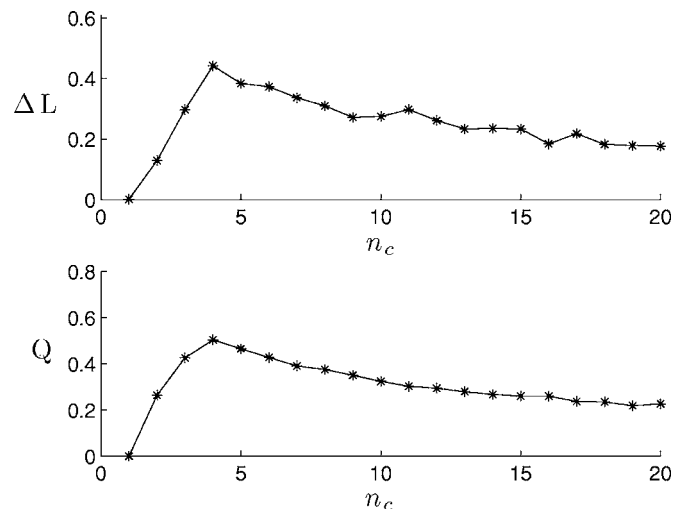

FIG. 7. Top: $\Delta L = L - L_0$ (see the text) is plotted vs n_c for a randomly generated network with $z_{out}=4$. Bottom: the modularity of the modular structure found by the proposed algorithm on the same network.
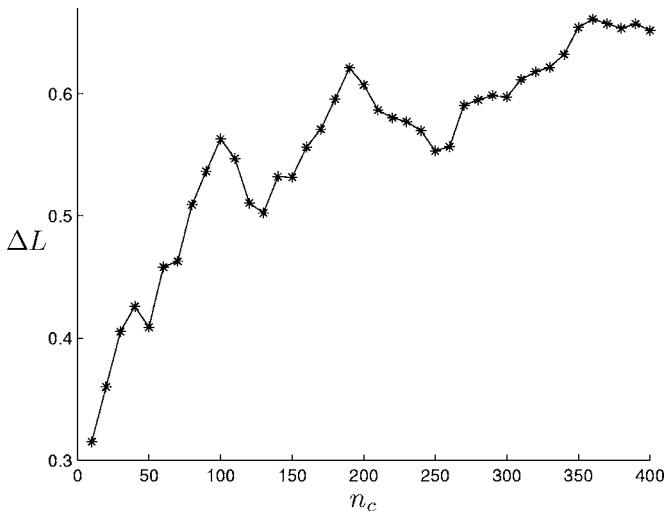
FIG. 8. $\Delta L = L - L_0$ is plotted vs n_c for the PGP network. The curve is obtained using repeated division in two parts, and $S_{th} = 16$.

FIG. 10. The scaling of the CPU time is reported as a function of the number of vertices N at fixed n_c.

optimal number of communities, we subsequently run our algorithm with $n_c = 360$ so as to find all the 360 communities in one step. To describe the properties of the communities structure we obtain, we sort the communities according to their sizes and we plot, in Fig. 9, the size $S(i)$ of the ith community. We note that there are eight communities with size greater than 100, while communities with size smaller than 20 exhibit the functional relation $S(i) \sim i^{0.61}$: this scaling law may be connected with the scale-free structures discovered in the PGP network.[25]

We finally report in Fig. 10 the CPU time needed to complete a given partition as a function of the number of vertices. The curve confirms that, for a given n_c, the computational demand scales linearly with the network size.

IV. CONCLUSIONS

In conclusion, we introduced a novel method for identifying the modular structures of a network based on the maxi-
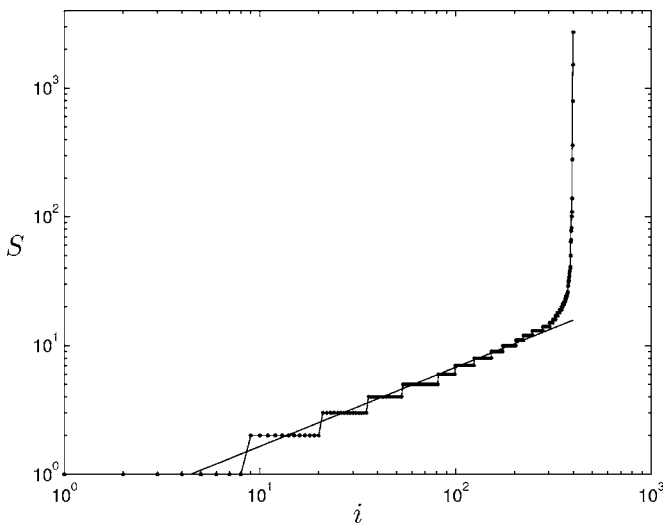

FIG. 9. The cardinality $S(i)$ of the ith community (after sorting in ascending order) is plotted vs i. The linear fit, in the left part of the curve, is $S \sim i^{0.61}$.
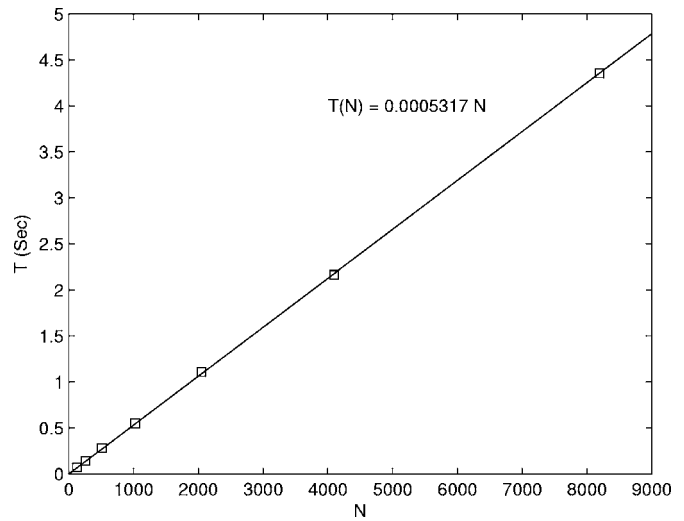
mization of an objective function: the ratio association. This objective function emerges in the frame of probabilistic autoencoders, thus providing a new description of the communities detection problem, in terms of a lossy compression of the structures, as well as a new selection strategy for the number of modules by means of the log-likelihood of the model.

ACKNOWLEDGMENTS

APPENDIX: PROBABILISTIC AUTOENCODER FRAME

In this appendix, we show that ratio association may be derived in the Probabilistic Autoencoder Framework. We briefly discuss autoencoders described by one-stage folded Markov chains.[15] Let us consider a point x, in a data space, sampled with probability distribution $P_0(x)$; a code index $\alpha \in \{1, \ldots, q\}$ is assigned to x according to conditional probabilities $P(\alpha|x)$. A reconstructed version of the input, x', is then obtained by use of the Bayesian decoder:

$$P(x'|\alpha) = \frac{P(\alpha|x')P_0(x')}{P(\alpha)}. \tag{A1}$$

The joint distribution of x, x', and α, describing this encoding-decoding process, is

$$P(x, x', \alpha) = P_0(x)P(\alpha|x)P(x'|\alpha); \tag{A2}$$

owing to (A1), the joint distribution reads

$$P(x, x', \alpha) = \frac{P_0(x)P_0(x')P(\alpha|x)P(\alpha|x')}{P(\alpha)}. \tag{A3}$$

The conditional probabilities $\{P(\alpha|x)\}$ are the free parameters that must be adjusted to force the autoencoder to emulate the identity map on the data space.

Let $s(x, x')$ be a measure of the similarity between input and output; the average similarity is then given by

$$S = \sum_{\gamma=1}^{q} \int dx \int dx' \frac{P_0(x) P_0(x') P(\alpha|x) P(\alpha|x')}{P(\alpha)} s(x, x'). \quad \text{(A4)}$$

A good autoencoder is obviously characterized by a high value of S. Given a set of data vectors $\{\mathbf{x}_i\}_{i=1}^{N}$, partitioning these points in q modules corresponds, in this frame, to design an autoencoder, with q code indexes, acting on data space. Choosing the encoder to be deterministic leads to

$$P(\alpha|x) = \delta_{\alpha\gamma(x)}, \quad \text{(A5)}$$

$\gamma(x) \in \{1, \ldots, q\}$ being the code index associated with x. The estimate for the average similarity, based on the dataset at hand, is given by

$$\hat{S} = \frac{1}{N} \sum_{\alpha=1}^{q} \frac{\sum_{i,j=1}^{N} \delta_{\alpha\gamma_i} \delta_{\alpha\gamma_j} s_{ij}}{\sum_{k=1}^{N} \delta_{\alpha\gamma_k}}. \quad \text{(A6)}$$

If the similarity matrix s_{ij} is identified with the kernel matrix K, we obtain

$$N\hat{S} = n_c \sigma + \mathcal{R}.$$

Therefore, maximization of the ratio association is equivalent to designing the most effective autoencoder, the effectiveness being measured by the average similarity.

Now we consider the average log-likelihood[23] of data $\{\mathbf{x}_i\}_{i=1}^{N}$,

$$\sum_{\gamma=1}^{q} \int dx P(x, \gamma) \log P(x|\gamma). \quad \text{(A7)}$$

We may easily obtain an estimate of this quantity, which measures how good the autoencoder frame is to model the dataset. Using kernel density estimation,[24] we easily obtain

$$L = \frac{1}{N} \sum_{i=1}^{N} \log\left(\frac{\text{links}(\{i\}, \pi_{\gamma(i)})}{|\pi_{\gamma(i)}|} \right); \quad \text{(A8)}$$

the numerator, in the formula above, is the number of links from node i to nodes in the same module as i, whereas the denominator is the cardinality of the module to which i belongs.

[1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Phys. Rep. **424**, 175 (2006).
[2] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness* (W. H. Freeman & Company, San Francisco, 1979).
[3] A. Pothen, H. D. Simon, and K. P. Liou, SIAM J. Matrix Anal. Appl. **11**, 430 (1990).
[4] S. Wasserman and K. Faust, *Social Networks Analysis* (Cambridge University Press, Cambridge, 1994).
[5] J. Reichardt and S. Bornholdt, Phys. Rev. Lett. **93**, 218701 (2004).
[6] M. E. J. Newman, Eur. Phys. J. B **38**, 321 (2004).
[7] M. Girvan and M. Newman, Proc. Natl. Acad. Sci. U.S.A. **99**, 7821 (2002).
[8] R. Guimera and L. A. N. Amaral, Nature **433**, 895 (2005).
[9] J. Duch and A. Arenas, Phys. Rev. E **72**, 027104 (2005).
[10] M. E. J. Newman, Proc. Natl. Acad. Sci. U.S.A. **103**, 8577 (2006).
[11] A. Arenas, A. Díaz-Guilera, and C. J. Pérez-Vicente, Phys. Rev. Lett. **96**, 114102 (2006).
[12] S. Boccaletti, M. Ivanchenko, V. Latora, A. Pluchino, and A. Rapisarda (unpublished).
[13] M. Girolami, IEEE Trans. Neural Netw. **13**, 780 (2002).
[14] K. Rose, E. Gurewitz, and G. Fox, Pattern Recogn. Lett. **11**, 589 (1990).
[15] L. Angelini, L. Nitti, M. Pellicoro, and S. Stramaglia, Phys. Lett. A **285**, 279 (2001).
[16] T. Schreiber and A. Schmitz, Phys. Rev. Lett. **79**, 1476 (1997).
[17] By q-way disjoint partition it is meant a set of q disjoint data subsets, whose union is the whole data.
[18] J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis* (Cambridge University Press, Cambridge, 2004).
[19] J. Shi and J. Malik, IEEE Trans. Pattern Anal. Mach. Intell. **22**, 888 (2000).
[20] I. S. Dhillon, Y. Guan, and B. Kulis, *A Unified View of Kernel k-means, Spectral Clustering and Graph Cuts*, University of Texas at Austin, Dept. of Computer Sciences, UTCS Technical Report TR-04-25 (2005).
[21] W. W. Zachary, J. Anthropol. Res. **33**, 452 (1977).
[22] We randomly generate 100 such random networks and take L_0 as the average; however, we find that L_0 varies less than 1% over the realizations.
[23] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, International Student Edition (McGraw-Hill, New York, 1985).
[24] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification* (Wiley, New York, 2001).
[25] X. Guardiola, R. Guimera, A. Arenas, A. Diaz-Guilera, and L. A. N. Amaral, eprint cond-mat/0206240.
[26] M. E. J. Newman, Phys. Rev. E **69**, 066133 (2004).