Extracting topological features from dynamical measures in networks of Kuramoto oscillators

Luce Prignano and Albert Díaz-Guilera

Departament de Física Fonamental, Universitat de Barcelona, 08028 Barcelona, Spain
(Received 28 February 2011; revised manuscript received 15 February 2012; published 29 March 2012)

The Kuramoto model for an ensemble of coupled oscillators provides a paradigmatic example of nonequilibrium transitions between an incoherent and a synchronized state. Here we analyze populations of almost identical oscillators in arbitrary interaction networks. Our aim is to extract topological features of the connectivity pattern from purely dynamical measures based on the fact that in a heterogeneous network the global dynamics is not only affected by the distribution of the natural frequencies but also by the location of the different values. In order to perform a quantitative study we focused on a very simple frequency distribution considering that all the frequencies are equal but one, that of the pacemaker node. We then analyze the dynamical behavior of the system at the transition point and slightly above it as well as very far from the critical point, when it is in a highly incoherent state. The gathered topological information ranges from local features, such as the single-node connectivity, to the hierarchical structure of functional clusters and even to the entire adjacency matrix.

I. INTRODUCTION

Currently, it is widely acknowledged that complex patterns of interaction are as ubiquitous in nature as in society [1]. Nonetheless, further research is required to completely understand how the topology affects the system dynamics [2,3], particularly how global dynamical properties are related to the units dynamics and the interactions between them. A unique answer cannot be provided since complex networks respond differently depending on the dynamical processes that take place within them [4].

One of the most interesting of these macroscopically defined dynamical processes is synchronization, an emerging phenomenon in which populations of interacting units display a common periodic behavior [5,6]. Indeed, understanding the role of connectivity in synchronization has been the subject of intense research in recent years [7]. On the one hand, much work has focused on the generic properties of dynamical systems, mainly looking for necessary and sufficient conditions that would grant that a population of units under a set of simple dynamical rules is able to synchronize [8]. On the other hand, much progress has been made by studying precise models of phase oscillators, one of the most paradigmatic being the model proposed by Kuramoto [9,10], where the interaction between the units is proportional to the sine of the phase difference.

In the present work, we will continue along this line and analyze a population of Kuramoto oscillators with a precise distribution of frequencies. The original work by Kuramoto and many subsequent studies considered that the oscillators, each coupled equally to all the others, had natural frequencies taken from a given distribution. The nonzero width of those distributions made the units follow different trajectories, whereas the interaction term made their phases approach. In fact, depending on the width of the frequency distribution, there is a critical value of the interaction strength above which the units tend to entrain their phases and hence leave the incoherent regime. If the natural frequencies of the oscillators are identical, a unique outcome is possible as the only attractor of the dynamics is a completely synchronized state in which all the oscillators end up in a common phase. And this

occurs for any initial conditions and for any (connected) topology [11].

In systems with regular patterns of connectivity (including all to all) the only complexity comes from the frequency distribution, whereas in more realistic (nonhomogeneous) patterns, not only the frequency values matter but also the precise location as well [3,12].

Here we will focus on a particular frequency distribution, one that is just one step away from the homogeneous case. Such a distribution has identical frequencies for all oscillators except one. This singular oscillator, with a higher frequency than the rest, has received the name *pacemaker*, and its effect in populations of Kuramoto oscillators has been analyzed [13,14]. In Ref. [13], Kori and Mikhailov consider a special case where the pacemaker affects its neighbors but is not affected by them; under these conditions they find numerically that the range of frequencies of the pacemaker for which the system can attain global synchronization depends on the "depth" of the network, where the depth is defined as the maximum distance from the pacemaker to peripheral nodes. Radicchi and Meyer-Ortmanns [14] consider regular structures for which the conditions to synchronize can be analytically computed.

In this paper we use several properties of the heterogeneity induced by the existence of the pacemaker to find useful relations between topology and dynamics. On one hand, by knowing the topology one should be able to infer the dynamical properties of the network. On the other hand, by measuring the dynamics some structural properties can be inferred, and this will be our purpose.

First, we use a similar procedure to the one used in Refs. [13,14], showing that there is a critical value for the frequency of the pacemaker above which the (frequency) synchronized state cannot exist. This is related to the existence of a synchronized solution (also exploited in Ref. [15]) that applies to any subset of oscillators. We find, however, that from a practical point of view the most restrictive condition is usually for the equation of the pacemaker that involves its connectivity, and hence there is a clear relationship between the critical frequency and the pacemaker connectivity that can be used as an experimental measure of the degree.

In order to get more details on the network structure we analyze the system above the critical value where correlations between dynamical evolution of the nodes appear. Such correlations enable us to reveal the hierarchical organization and to recover the network connectivity.

The structure of the paper is as follows. First, in Sec. II, we characterize the coherent state and the transition to the incoherent one by means of a proper definition of the order parameter. Then, in Sec. III, we qualitatively analyze the behavior of the system when it is not in the frequency-locked state. Section IV is devoted to studying the relation between local connectivity and the ability of the system to reach a synchronized (frequency-locked) state. In Sec. V we focus on the system slightly above the transition toward the incoherent state. We show that it is possible to perform some hierarchical analysis concerning the connectivity network. Finally, in Sec. VI we study the system far above the critical point, in a regime characterized by short range correlations, where it becomes easy to identify the nodes directly connected to the pacemaker. Thus the reconstruction of the whole connectivity pattern is accurate and fast.

II. SYNCHRONIZATION AND PHASE TRANSITION

In the original Kuramoto model [9,10], the phases of the oscillators evolve according to the following equation:

$$\dot{\varphi}_i = \omega_i + \sigma \sum_{j=1}^{N} \sin(\varphi_j - \varphi_i), \tag{1}$$

where N is the total number of units of the system, ω_i is the natural frequency of unit i, taken from a distribution, and σ stands for the coupling strength. This case corresponds to a fully connected topology; i.e., each unit interacts with all the other ones. The ability of the system to reach a coherent state, for a given coupling strength, depends only on the width of the distribution of natural frequencies.

Here we want to consider arbitrary connectivity patterns. In this situation, the behavior of the system can no longer be understood in terms of the ratio between the distribution width and the coupling strength only. Where the natural frequencies values are located is also relevant since on a generic interaction network nodes are no longer equivalent.

From now on we are using the two-level hierarchical network of nine nodes represented in Fig. 1 as a benchmark, and, when not otherwise stated, all the figures refer to that connection pattern. This network has been presented in Ref. [16] as a very simple example of the class of deterministic scale-free hierarchical networks proposed by Ravasz and Barabasi in Ref. [17]. We choose this small regular connectivity pattern as a simple paradigmatic example showing general properties of the studied systems since it makes it easy to recognize the role of each node.

Let us rewrite the equation for the evolution of the phases including a symmetric connectivity matrix a_{ij} that takes a value of 1 (0) if nodes i and j are connected (disconnected):

$$\dot{\varphi}_i = \omega_i + \sum_{j=1}^{N} a_{ij} \sin(\varphi_j - \varphi_i), \tag{2}$$
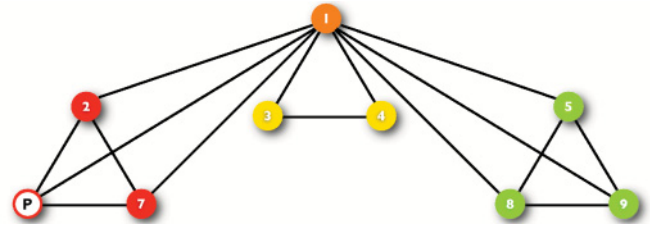

FIG. 1. (Color online) Hierarchic network that will be used as a benchmark. In this particular setting the pacemaker is located on a peripheral node (marked as P) of degree $k_p = 3$. The other nodes are grouped into sets using different colors. The elements of each set are topologically equivalent if we look at the network from the point of view of the pacemaker. Consequently, their dynamical evolution is identical.

where we have rescaled time by setting $\sigma = 1$. Now we consider all the oscillators to have the same natural frequency (0, without loss of generality), except one of them, called the pacemaker, whose frequency is $\omega \neq 0$. It is precisely this extremely simple choice of frequencies that enables us to study the roles played by individual oscillators.

If a stationary state exists, then all the effective frequencies will take constant values, and the following conditions have to be satisfied:

$$\sum_{j=1}^{N} a_{ij} \sin(\varphi_j - \varphi_i) = \Omega_i \quad \forall\, i \neq p, \tag{3}$$

$$\omega + \sum_{j=1}^{N} a_{pj} \sin(\varphi_j - \varphi_p) = \Omega_p, \tag{4}$$

where $\{\Omega_i\}$ are the effective frequencies of the oscillators. Notice that summing up Eqs. (3) and (4) the coupling terms cancel because of the symmetry of the interaction, which results in

$$\sum_{i=1}^{N} \Omega_i = \omega. \tag{5}$$

Looking at Eqs. (3) and (4), it is easy to recognize that there is an interplay between two effects. On the one hand, the width of the frequencies distribution (in our present case this role is played by ω itself) tends to keep the evolution of the oscillators apart since each one follows its natural frequency. On the other hand, the interaction term makes them approach their phases as well as their effective frequencies. Then we conclude that if the pacemaker's natural frequency is small enough, the interaction term dominates, and after a transient time, all effective frequencies Ω_i will be identical:

$$\Omega_i = \omega/N \quad \forall\, i, \tag{6}$$

including the pacemaker. In this case we can say that the system is in a frequency-locked state since all oscillators have the same frequency, although the phases are not equal because there is a coupling term (that of the pacemaker) that cannot vanish.

When increasing the pacemaker frequency ω, some oscillators cannot keep the phase difference, and the frequency-locked state is broken. The left-hand side of Eq. (3) is indeed
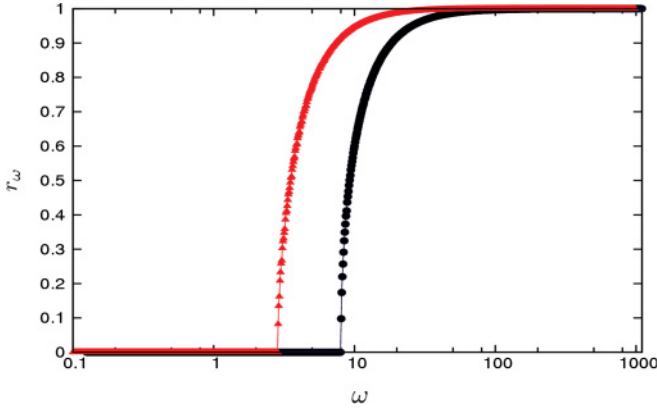
FIG. 2. (Color online) Order parameter (7) as a function of the natural frequency of the pacemaker. Different curves correspond to different settings: circles refer to the pacemaker located on node 1 in Fig. 1 ($k_p = 8$), and triangles refer to the pacemaker on node 2 ($k_p = 3$). The average value $\langle r_\omega \rangle_t$ for $\omega > \omega_p^c$ was calculated on a time window $\Delta t = 100$.

bounded because of the sine terms, whereas the right side increases as the pacemaker frequency is increased. A similar conclusion can be deduced from Eq. (4). Consequently, there will be a transition from a synchronized to an incoherent state. Thus we can define the critical value ω_p^c as the maximum value of the natural frequency of the pacemaker for which the system can attain global synchronization.

Such a transition for a population of phase oscillators is typically characterized by an order parameter R, defined through the equation $Re^{i\Psi} = \sum_j e^{i\varphi_j}$, where Ψ is a global phase (not constant) [18].

In the present work, following [12,19], we adopt another order parameter that is a normalized measure of the effective frequency dispersion (standard deviation):

$$r_\omega = \sqrt{\frac{\sum_{i=1}^{N} [\dot{\varphi}_i/\langle\omega\rangle - 1]^2}{N - 1}}, \qquad (7)$$

where $\langle\omega\rangle$ is the average effective frequency of the oscillators population, a constant quantity always equal to ω/N. According to its definition, r_ω takes values in the interval [0,1] (see Fig. 2). It should be noticed that, since above the critical frequency the system is not able to reach a steady state any longer, calculation of the order parameter (7) requires performing averages over an appropriate time window. Anyway, the value of $\langle\omega\rangle$ does not change because what we found in Eq. (5) is a general result, even for instantaneous values of the effective frequencies.

To find the precise value of the critical frequency we apply the Newton-Raphson method (NR) and check, as a function of the frequency ω, whether the synchronized solution of Eqs. (3) and (4) exists. To simulate the dynamics of the system in the incoherent state ($\omega > \omega_p^c$) we take as initial phases $\{\varphi_i(0)\}$ the stationary values of the differences provided by the NR solution for $\omega = \omega_p^c$. The system of equations (2) is numerically integrated with Euler's method (first order), unless otherwise stated, at fixed time step $\delta t = 10^{-2}$.
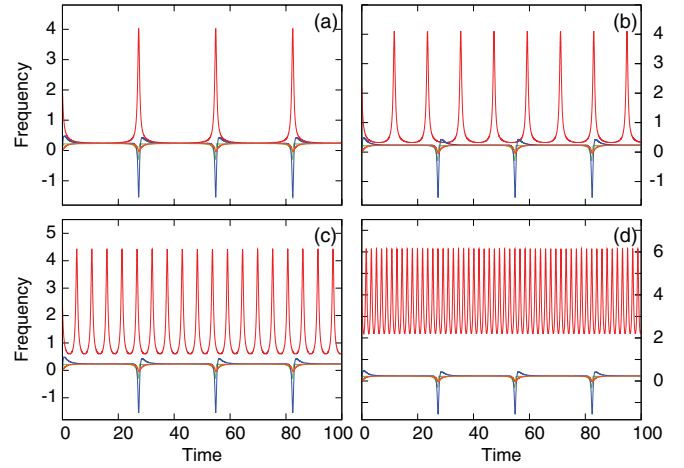

FIG. 3. (Color online) Effective frequencies above the critical point as functions of time. That of the pacemaker (red top curve), in this particular setting located on node 3 in Fig. 1, is on average much larger than the others (lower curves). The plots show a pacemaker natural frequency value that is (a) 1.01, (b) 1.05, (c) 1.2, and (d) 2 times its critical value. Time starts after a transient lag $T_s = 10$.

III. INCOHERENT STATE

Above the critical frequency ω_p^c the system is no longer in a stationary state, and hence the effective frequencies are no longer constant.

Numerical simulations show that, after a transient time, the system enters into a "periodic" state (see Fig. 3). The features of this periodic state are not affected by the initial conditions, and they only depend on the pacemaker frequency and location. It is precisely this fact that enables us to infer topological properties from dynamical measurements.

Figure 4 summarizes what we have learned up to now, shedding light on some interesting details. The time average of the effective frequency of the pacemaker $\langle\dot{\varphi}_p\rangle_t$ and that of one of its neighbor $\langle\dot{\varphi}_j\rangle_t$ are plotted as functions of the pacemaker natural frequency. These quantities are calculated from numerical simulations, taking into account appropriate time windows.

Starting from small values of ω, the picture shows how all the effective frequencies increase together linearly, following the reference line $\Omega_i = \omega/N$ defined by Eq. (6). Then, when ω reaches the critical value ω_p^c, they do separate. Initially, the average effective frequency of the pacemaker goes through a more than linear increasing, while the others start decreasing, keeping their (average) values very close to each other. For even larger values, when $\omega \gg \omega_p^c$, Fig. 4 shows how the average effective frequency $\langle\dot{\varphi}_p\rangle_t$ tends to ω, asymptotically increasing along a new reference line with a slope equal to 1. At the same time, $\langle\dot{\varphi}_i\rangle_t$ for $i \neq p$ goes to zero, as required by the conservation law (5).

IV. CRITICAL FREQUENCY AND LOCAL TOPOLOGY

In this section we explore the relation between the topology of the network and the value of the critical natural frequency of the pacemaker depending on the node where it is located.
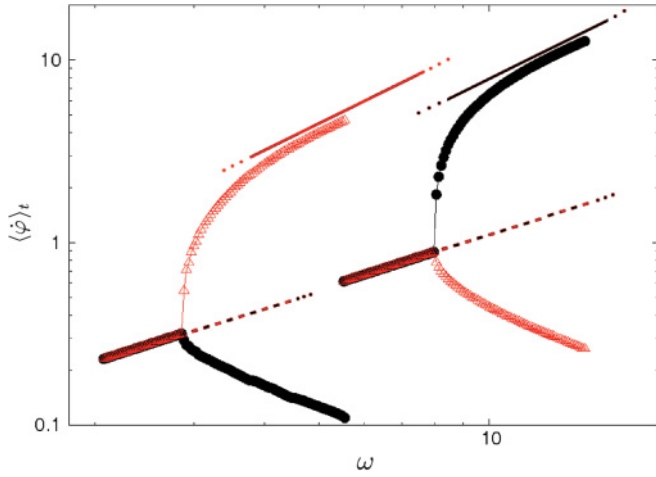
FIG. 4. (Color online) Average effective frequencies as a function of the natural frequency of the pacemaker ω. The behavior of two oscillators is shown: node 1 (black circles) and node 2 (red triangles). On the right side the pacemaker is node 1, and on the left one it is node 2. Initially. the frequencies are synchronized and they increase linearly with a slope of $1/N$ (dashed line), as expected from Eq. (6). Then, when ω reaches the critical value, which is different for different locations of the pacemaker, they grow apart. Far above the critical values, the average frequency of the pacemakers approach asymptotically a new reference line with a slope of 1 (solid line). The time averages were performed on a time window $\Delta t = 100$.

Let us begin by writing the equation for the pacemaker in the synchronized state. As a consequence of Eq. (6), we have

$$\omega + \sum_{j=1}^{N} a_{jp} \sin(\varphi_j - \varphi_p) = \omega/N. \tag{8}$$

This equation links the natural frequency of the pacemaker to the constant values of the phase differences between it and its neighbors, when all the units are oscillating with the same effective frequency. Since the number of non-null terms a_{jp} in the previous expression is given by the number of nodes connected to the pacemaker and $\sin(\varphi_j - \varphi_i) \in [-1,1]$, the degree (or connectivity) of the pacemaker is a bound for the absolute value of the sum in Eq. (8).

Thus there is an upper bound for the critical frequency:

$$\omega_p^c \leqslant k_p \frac{N}{N-1}, \tag{9}$$

where k_p is the degree of the pacemaker. Indeed, any value larger than the right-hand term in inequality (9) is surely unable to satisfy Eq. (8), and hence the system is unable to be frequency synchronized.

Notice that we have obtained this bound by taking into account a single equation, that of the pacemaker. We can write for any oscillator the analog of Eq. (8) as follows:

$$\sum_{j=1}^{N} a_{ji} \sin(\varphi_j - \varphi_i) = \omega/N, \quad \forall \, i \neq p. \tag{10}$$

It is easy to verify that no stricter condition can arise from any of these equations [20]. However, stronger bounds could exist due to the combination of Eq. (8) and some of Eqs. (10).

Let us consider a set of $(n + 1)$ connected nodes, among which the pacemaker is included [21]. Labeling them by an increasing index $i = 1, 2, \ldots, n + 1 = p$ and summing up their equations, we obtain

$$(n+1)\frac{\omega}{N} = \omega + \sum_{i=1}^{n+1} \sum_{j=1}^{N} a_{ij} \sin(\varphi_j - \varphi_i). \tag{11}$$

If two nodes in the considered group are neighbors, their respective interaction terms cancel each other. So the number of remaining terms of the sums in Eq. (11) is given by

$$K_{\text{out}} = \sum_{i=1}^{n+1} k_i - \sum_{i,j=1}^{n+1} a_{ij}, \tag{12}$$

where k_i is the degree of the ith node and K_{out} is equal to the number of links connecting the nodes of the considered set to external ones.

Consequently, Eq. (11) can be rewritten as

$$(n+1)\frac{\omega}{N} = \omega + \sum_{l=1}^{K_{\text{out}}} \sin(\phi_l), \tag{13}$$

where $\phi_l = \varphi_j - \varphi_i$, is the phase difference between two connected nodes i and j which are, respectively, inside and outside the group.

We are now able to write the expression of the upper bound for the critical frequency ω_p^c in a generalized form:

$$\omega_p^c \leqslant K_{\text{out}} \frac{N}{N-(n+1)} = N \frac{K_{\text{out}}}{N_{\text{out}}}, \tag{14}$$

where N_{out} stands for the number of nodes not belonging to the considered set. Eq. (14) reduces to the previous upper bound if one chooses $n = 0$.

In this way we can write a very large number of conditions, that is, the number of the connected sets of nodes that include the pacemaker and whose size ranges from 1 to $N - 1$. Among these, the strongest one is that for which the ratio $K_{\text{out}}/N_{\text{out}}$ takes its minimum value. This is a combinatorial problem, which is in principle very simple but is hard from a computational point of view since the number of conditions grows at least exponentially with the network size.

Minimizing the ratio $K_{\text{out}}/N_{\text{out}}$, we find the strictest condition on ω_c^p that can be expressed in the form of a single equation. No other equation obtained as a linear combination of Eqs. (3) and (4) may provide a stronger bound. This condition is analogous to the necessary condition for global synchronization concerning the surface (here K_{out}) of any subset of nodes derived in Ref. [15] for randomly distributed natural frequencies and generic oscillators. However, these conditions are not sufficient. In our case, it is not certain that the K_{out} remaining sine terms of Eq. (13) are allowed to take their minimal values simultaneously. This kind of problem directly involves the sine function arguments that may not be independent since they are differences between pairs of phases and we are dealing with a system of N coupled equations. It may happen that two or more phases are tied to each other by a certain set of equations of the kind $f_i(\varphi_i, \{\varphi_{i_j}\}) = 0$ (where nodes $\{i_j\}$ are neighbors of node i). Consequently, we cannot minimize the sum of sine terms on a hypercube $[0; 2\pi]^{K_{\text{out}}}$, but

we have to restrict ourselves to a hypersurface of dimension $K_{out} - m$, where m is the number of constraints. A system may experience this kind of difficulty (which we can regard as a kind of angle frustration) only if cycles are present and there is some anisotropy and only when $1 < k_p < N - 1$. Therefore, for a good number of regular connectivity patterns, such as those analyzed in Ref. [14], there is not such a problem, and it is possible to analytically calculate the entire set of values $\{\omega_c^{(p)}\}$, $p = 1, \ldots, N$.

As a simple, analytically solvable network let us consider a Cayley tree with coordination number z, made up of S shells. For each node it is indeed possible to single out a connected "group" such that $K_{out} = 1$, taking in all the nodes on the branch starting from the considered pacemaker. In this way we are minimizing the ratio K_{out}/N_{out} so that we can consider the strictest equation among Eqs. (14). Moreover, since there is no cycle, there are no problems of angle frustration either. Therefore, the obtained expressions give the correct values, not just bounds. In this way we obtain for the critical frequency

$$\omega_c^{(s)} = N \frac{1}{N - \sum_{i=0}^{S-s}(z-1)^i},$$

where s is the shell of the pacemaker.

Even though in real complex networks it is not so easy to calculate $\{\omega_c^p\}$, we have empirically verified that only in a few cases is the critical frequency much smaller than its first upper bound (9). This can be clearly observed in Fig. 5, where we plotted the ratios between the real critical values and the
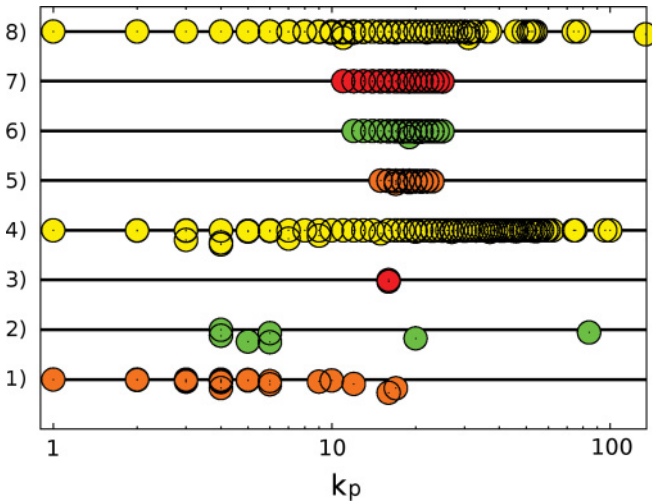

FIG. 5. (Color online) Critical frequency of a pacemaker as a function of its degree for a set of networks. We have divided the critical frequency by the degree and by $N/(N-1)$ such that the bound given by Eq. (9) is 1. We have shifted the data for the different networks, and the horizontal lines are the reference (equal to 1) for each case. From bottom to top the networks are (1) the Zachary club social network [22] used in community detection applications, (2) a hierarchical network of 125 nodes and 3 levels [17], (3) a network of 4 communities of 32 nodes each used as benchmark in community detection algorithms [22] where all the nodes have the same degree, (4) a network of jazz bands [23], (5–7) three networks of three levels of community structure used to relate topological and temporal scales in synchronization [24], and (8) the *Caenorhabditis elegans* neural network [25].

corresponding upper bound for every choice of the pacemaker in several networks.

The accuracy of this estimation enables us to use it in the opposite direction, i.e., to get an estimation of the pacemaker degree from an experimental measure of the critical frequency. We can invert Eq. (9), obtaining

$$k_p \geqslant \omega_p^c \frac{N-1}{N}, \qquad (15)$$

but since the right term is not an integer, the smallest allowed value for k_p is

$$k_p^* = \left[\omega_p^c \frac{N-1}{N} + 1 \right], \qquad (16)$$

where $[x]$ stands for the integer part of x. We can conclude that Eq. (16) gives the correct value of k_p whenever

$$\omega_p^c \in \left[(k_p - 1)\frac{N}{N-1}, k_p \frac{N}{N-1} \right].$$

This fact implies that the estimator (16) for the degree of the pacemaker is very reliable. Indeed, it only fails when the critical frequency is really smaller than its bound (9).

V. SLIGHTLY ABOVE THE CRITICAL POINT

In this and in the next section we translate the rich dynamical information that the system provides in the incoherent state into useful topological information. Here we focus on the behavior of the system slightly above the critical point, while in Sec. VI we will analyze the system when the natural frequency of the pacemaker is many time larger than its critical value.

We are interested in estimating how similar two nodes are from a global topological perspective. For this purpose we need to define an appropriate correlation function, able to relate the dynamical responses of pairs of oscillators.

Looking for the expression of a good correlation function, we get no help from the average values $\langle \dot{\varphi}_i \rangle_t = \int_0^\infty \varphi_i(t)dt$. Indeed, in this regime, all the oscillators, except the pacemaker, have the same average effective frequency. On the contrary, it can be useful to look at the difference between instantaneous values. We measure the frequency of every oscillator at each time, inside a suitable interval. In order to define a correlation, that is, a quantity that has to be non-negative and symmetric with respect to nodes indexes i and j, it is reasonable to start from a power of the absolute value of the difference $|\dot{\varphi}_i^{(p)}(t) - \dot{\varphi}_j^{(p)}(t)|$, where (p) stands for the pacemaker that induces the considered dynamical evolution. Therefore, we propose

$$c_{ij}^p(t) = 1 - \sqrt{\frac{|\dot{\varphi}_i^{(p)}(t) - \dot{\varphi}_j^{(p)}(t)|}{\omega}}.$$

Dividing by ω makes the argument of the root less than 1 because, even if the frequencies may take negative values (see Fig. 3), the condition $|\dot{\varphi}_i(t)| \ll \omega$ always holds.

The period of the effective frequency's oscillation depends on which node is the pacemaker. Then, in order to compute averages in time that are really independent from the considered interval, we have to choose a time window many times larger than the oscillation period. Furthermore, since
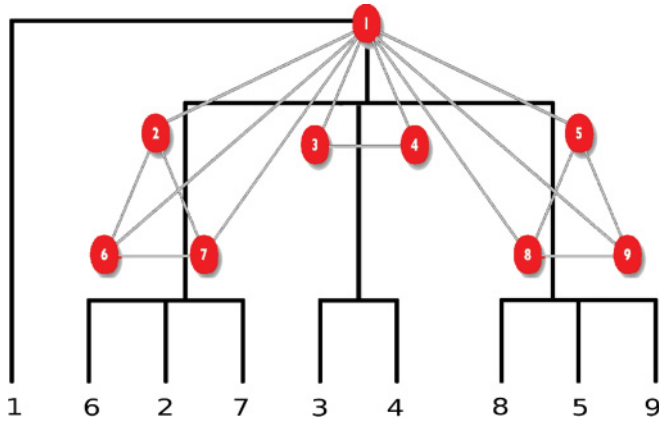
FIG. 6. (Color online) The network of Fig. 1 and its corresponding dendrogram. Correlations are calculated by averaging over a time window $\Delta t = 60$, after a transient lag $T_s = 10$, for $\omega = 1.1\omega_p^c$.

$|\dot{\varphi}_i - \dot{\varphi}_p| \gg |\dot{\varphi}_i - \dot{\varphi}_j|$ for any $i, j \neq p$, we decide to exclude these contributions, taking into account only terms of the kind $c_{i,j}^p$ where $i \neq p$ and $j \neq p$.

Finally, in order to remove the dependence from the index p we have to average all the possible pacemakers. Summarizing in a compact expression, our correlation function can be written as follows:

$$c_{ij} = 1 - \frac{1}{N-2} \sum_{\substack{p=1 \\ p \neq i,j}}^{N} \frac{1}{t_1 - t_0} \int_{t_0}^{t_1} \sqrt{\frac{|\dot{\varphi}_i^{(p)}(t) - \dot{\varphi}_j^{(p)}(t)|}{\omega}} \, dt. \tag{17}$$

A. Hierarchical organization

Once we have obtained the correlation matrix we can proceed to some hierarchical analysis. In the present work we use the standard unweighted pair group method average (UPGMA) [26] algorithm to compute such diagrams. What we find out is a hierarchy of dynamical communities, whose meaning is immediately understandable in the case of small networks, such as our benchmark in Fig. 1 (see Fig. 6). Obviously, this simple network does not need any analysis to obtain its hierarchical organization, but this methodology can be very useful when applied to functional hierarchical network.

As a paradigmatic example, let us consider the corticocortical network of a cat at the macroscopic level. We look at each cortical area as a basic unit, modeling it as a Kuramoto oscillator, finding similar results as in [27,28].

In Fig. 7 we show that, going down along our dendrogram starting from the root, it is possible to recognize two communities clearly separated. Then, the right branch splits into two parts, and the left one undergoes two subsequent bifurcations, so that it is possible to identify three groups of nodes on it. At this level we have five communities. Four of them correspond to well-known physiological subsystems: the frontolimbic (FL), the somatosensory motor (SM), the auditory (A), and the visual (V). The fifth one (HUBS) is composed, except for a single area [29], by superhubs, sometimes considered as a
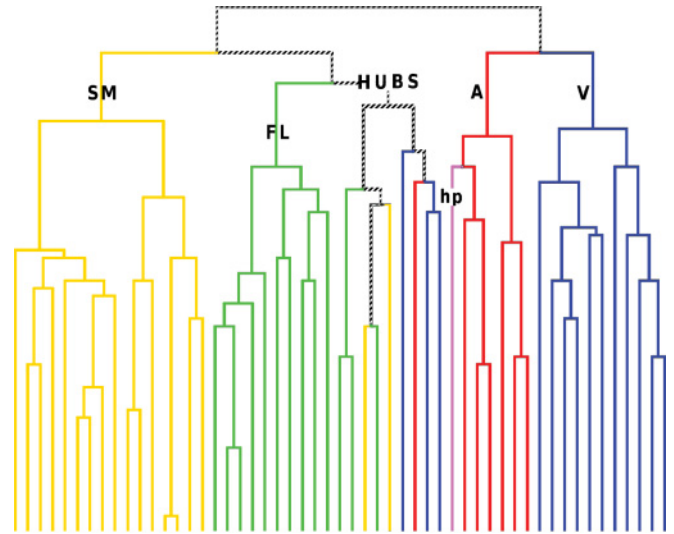

FIG. 7. (Color online) Dendrogram of the cortical brain network of a cat. Different colors correspond to different subsystems: the frontolimbic (FL), the somatosensory motor (SM), the auditory (A), and the visual (V). The rich club is labeled with HUBS, while the branch indicated with the label hp is the area that belongs to the hippocampus, and it is out of place. Correlations are calculated by averaging over a time window $\Delta t = 100$, after a transient lag $T_s = 10$, for $\omega = 1.1\omega_p^c$.

metacommunity (rich club) [27,28]. The most relevant aspect of our hierarchical analysis is that there is no way to recognize this metacommunity if the dendrogram is constructed by means of static methods. It cannot be obtained throughout correlation matrices generated from the adjacency matrix using, for instance, Pearson's coefficient [30] either. Nor can these nodes emerge as a community when the modularity function is maximized. Indeed, maximizing the modularity, we obtain as an optimal partition the same four groups corresponding to the four physiological sub-systems.

In general, complex networks can be organized, and thus analyzed, at different hierarchical levels. For social networks it is very important that a group is tight, so that the multiple connections within the group give rise to the concept of community. On the contrary, in biological networks the most crucial concept is function rather than connectivity per se. Therefore, methods that rely on the connections within groups and maximize modularity will not be enough to identify biological units based primarily on function [31,32]. In this case, our method, which analyzes the dynamical correlation between units, provides a better approach to inferring functional relationships.

One of the known problems of the methods commonly used for detecting community structures in complex networks is the existence of the so-called resolution limit, found by Fortunato and Barthelemy [33]. This issue is related to the impossibility for the methods based on modularity optimization to go beyond a certain resolution, which is related to the community size and to the number of links between communities. The paradigmatic example of the problem is a network formed by "cliques" (small groups of totally connected nodes), which are very sparsely connected. We have checked such structures and found that, dynamically, the correlations are

very strong within the cliques and not among nodes belonging to different modules, showing that our method for detecting the hierarchical organization is not affected by the resolution limit problem.

B. Recovering network topology

Let us now take a step backward and recover something we had previously discarded. In the sum of Eq. (17) we had excluded terms in which one of the indexes was equal to p since they were *heterogeneous*. So $c_{ij} = \frac{1}{N-2} \sum_{p \neq i,j} c_{ij}^p$. However, the set of elements c_{pj}^p, $p = 1, \ldots, N$, also contains information. We may ask ourselves which oscillators are most strongly correlated with the pacemaker and if they share some topological property. The simplest hypothesis is that the set of k_p largest c_{pj}^p identifies the neighbors of the pacemaker. This is reasonable since, even if the pacemaker is very weakly correlated to the rest of the oscillators, coefficients c_{pj}^p are not uniform, and the topological distance is the most immediate quantity that we may suppose this variability is related to. In Sec. IV we showed how to find out an estimator of the degree of each node from the critical frequencies. Thus if we are able to select the possible neighbors, we would be, in principle, able to reconstruct the entire network.

The first problem we face in the attempt to validate this hypothesis is that our list of likely neighbors gives us an asymmetric and weighted adjacency matrix, whose elements are

$$a'_{pj_i} = c_{pj_i}^p \quad \text{for} \quad i = 1, \ldots, k_p^*,$$
$$a'_{pj_i} = 0 \quad \text{for} \quad i = k_p^* + 1, \ldots, N,$$

where k_p^* is the estimator for the degree of the pacemaker given by (16) and $c_{pj_i}^p > c_{pj_l}^p$ whenever $i \leqslant k_p^*$ and $l > k_p^*$.

Moreover $a'_{mn} \neq a'_{nm}$ since, generally speaking, $c_{mn}^m \neq c_{mn}^n$. Therefore we have to remove the weights and symmetrize this matrix. Here we propose an algorithm to perform this task that is at the same time simple and efficient. It consists of four steps.

(1) Symmetrize the matrix in the usual way: $a_{mn}^s = (a'_{mn} + a'_{nm})/2$.

(2) Compute a list of temporary degree $k'_n \geqslant k_n^*$ as the number of non-null elements a_{nm}^s.

(3) Order all the non-zero values a_{nm}^s in a list, from the smaller to the larger.

(4) Check which ones among the corresponding likely links have to be removed, starting from the weakest one.

We proceed as follows: given a pair of nodes m and n whose link is the weakest one, if and only if $k'_m > k_m^*$ and $k'_n > k_n^*$, we remove that link, setting $a_{mn}^s = a_{nm}^s = 0$. In this case both k'_m and k'_n are reduced by one unit. Otherwise, we go to the next link, going on along the entire list until reaching the strongest link.

This method is rooted in the hypothesis, empirically very well verified, that the matrix a_{mn}^s contains all the links of the real network plus a number of *false positive* ones, i.e., that there is no *false negative* link. Thus we need just remove, never add, edges.

Moreover it works properly only if our estimators $\{k_n^*\}$ of the actual degrees $\{k_n\}$ are correct; otherwise, we may make

additional errors. Fortunately, it is a very infrequent problem. The sole hypothesis we make is that the probability for a link to be a "false" one is a monotonously decreasing function of the correlation between the nodes it joins.

Finally, the method does not ensure that in the final estimated network $k'_n = k_n^* \; \forall n$ because it is possible that even if $k'_n > k_n^*$, the nth oscillator has no possible neighbor whose temporary degree is larger than its estimated one. Sometimes this fact may cause new errors; others times it acts as a compensation of the underestimation of the real degrees.

In order to quantify how good a reconstruction is, we introduce the following error definition:

$$\text{err}(\%) = \frac{Fp + Fn}{L} \times 100,$$

where Fp and Fn are, respectively, the number of false positive (spurious) and false negative (missing) links in the reconstructed network and L is the number of edges in the original connectivity pattern. Globally speaking, we can state that our method allows for a reconstruction of an arbitrary connectivity pattern with good precision. Taking into account the networks in Table I, on average, we have $\text{err}(\%) = 6.5$.

Among these networks there are artificial as well as real connectivity patterns. They were selected to be representative of several classes of networks, including hierarchical as well as nonhierarchical, with and without community structure, regular and irregular. For this reason, the average error calculated on this set of benchmarks can be considered to be a good estimator of the accuracy of the proposed reconstruction method when applied on a given unknown connectivity pattern.

TABLE I. Results of the reconstruction on several networks. The columns give the size of the system N, the total number of links in the original network L, the total error in the estimation of the degrees $K_{\text{err}} = \sum_{i=1}^N |k_i - k_i^*|$, the total number of links in the reconstructed network before the removal of exceeding links L', the number of false positive F_p and false negative F_n links in this network, the same for the final reduced network (L'_r, F_p/F_n), and the final total error [err(%)]. From the first row, the networks are our usual benchmark [16], a ring of 6 cliques of 3 nodes [33], a hierarchical network of 25 nodes and 2 levels [17], the Zachary club social network [22], a ring of 16 cliques of 3 nodes [33], the cortical brain network of a cat [34], a hierarchical network of 125 nodes and 3 levels [17], a network of 4 communities of 32 nodes each [22], and 2 networks of 3 levels of community structure [24].

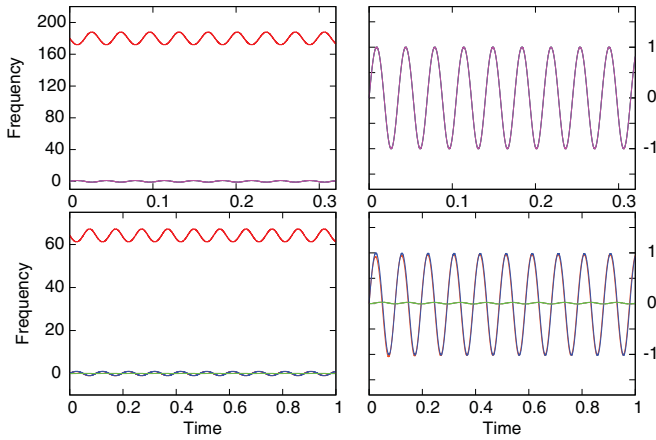| N	L	K_{err}	L'	F_p/F_n	L'_r	F_p/F_n	err(%)
9	15	0	15	0/0	15	0/0	0
18	24	0	24	0/0	24	0/0	0
25	66	0	82	16/0	66	0/0	0
34	78	7	99	27/6	75	7/10	21.8
48	64	0	64	0/0	64	0/0	0
53	391	0	445	53/0	392	5/4	2.3
125	394	33	475	81/0	383	1/12	3.3
128	1024	0	1060	57/21	1026	36/34	6.8
256	2311	0	3223	1040/128	2324	259/246	21.8
256	2301	0	2851	607/57	2312	116/105	9.6

FIG. 8. (Color online) Effective frequencies as function of time far above the critical point ($\omega = 20\omega_p^c$). Plots refer to the same network used for the previous pictures, in the case of two different choices of the pacemaker: node 1 ($k_1 = 8$) above and node 2 ($k_2 = 3$) below. On the left-hand side we plotted the frequency of all the nodes in the network. On the right side the scale has been changed, and the pacemakers are left out. Notice how above, where all the nodes are neighbors of the pacemaker, we may observe a unique curve. On the contrary, below there are two different kinds of oscillations. The largest ones are those of the neighbors of the pacemaker; the others belong to the oscillator not directly connected to it. Time starts after a transient lag $T_s = 1$. The integration time step used is $\delta t = 10^{-4}$.

VI. FAR FROM THE CRITICAL POINT

Far above the critical point the system behaves quite differently. As clearly shown in Fig. 8 (left panels) all units are characterized by effective frequencies that, after a transient time, oscillate around precise values that are equal to their own natural frequency. From this point of view, by increasing the natural frequency of the pacemaker the coupling is less and less important. But, in any case, there are still remnants of the interactions since the amplitudes of the oscillations decay very quickly with the distance from the pacemaker. Indeed, the frequencies of the neighbors of the pacemaker oscillate with an amplitude that is roughly $A_{\text{neigh}} \simeq 2$, while all the other oscillators are almost at rest compared with them. These conditions allow us to recognize the neighbors of a given pacemaker even if we do not know how many there are. Therefore, we may define a simplified correlation function that better suits this situation and that only connects each pacemaker with its neighbors:

$$c_{pi}^F = \frac{\max_t[\dot{\varphi}_i(t)] - \min_t[\dot{\varphi}_i(t)]}{\max_t[\dot{\varphi}_p(t)] - \min_t[\dot{\varphi}_p(t)]} = \frac{A_i}{A_p}. \qquad (18)$$

The above expression is the ratio between two positive terms (amplitudes), and it is equal to 1 for $i = p$.

On any connectivity pattern, the values c_{pi}^F are distributed along a set of stairs whose highest step is easy to identify even if we consider short time windows. The transient time, indeed, is always very short in this regime. We do not need anything else in order to completely reconstruct the entire connection topology.

All we have to do is to compute the values c_{pi}^F for each pacemaker. After finding out the maximum values $\max_{i \neq p} c_{pi}^F$

$\forall p$, we choose an appropriate threshold, say 0.5. A node j will be a neighbor of the pacemaker p if $c_{pj}^F/(\max_{i \neq p} c_{pi}^F) \geqslant 0.5$. Now we are able to construct a connectivity matrix.

Let us notice that in this case there is no need for symmetrization since the adjacency matrix constructed in this way is already symmetric because this method is based on a reliable general property that holds for any connectivity pattern. The use of a threshold is therefore, in principle, unnecessary since all the neighbors have the same amplitude of the frequency oscillation, when the pacemaker natural frequency is above a certain value. But since this value is not known *a priori* and it may be very large if the distribution of the degrees among the neighbors of the pacemaker is very wide, it is useful from an empirical point of view. It is important to stress that, even if we are still in a regime where some degree of *heterogeneity* among the neighbors is conserved, there is no chance to make any errors in the recovered topology. Indeed, the amplitudes of the frequency oscillations of the pacemaker's neighbors is at least one order of magnitude larger than that of any other oscillator (see Figs. 8 and 9). By means of this method all the topologies considered in Table I are properly reconstructed, without errors.

In addition, not all nodes need to be considered as pacemakers. While the method discussed in Sec. V B requires us to perform dynamical measures for every possible location of the pacemaker, for the current description this is not necessary. Indeed, we can look for the neighbors of a number $N' < N$ of pacemakers in order to get all the connections in the considered network. From an experimental point of view, adopting the conceptual framework proposed in Ref. [35],
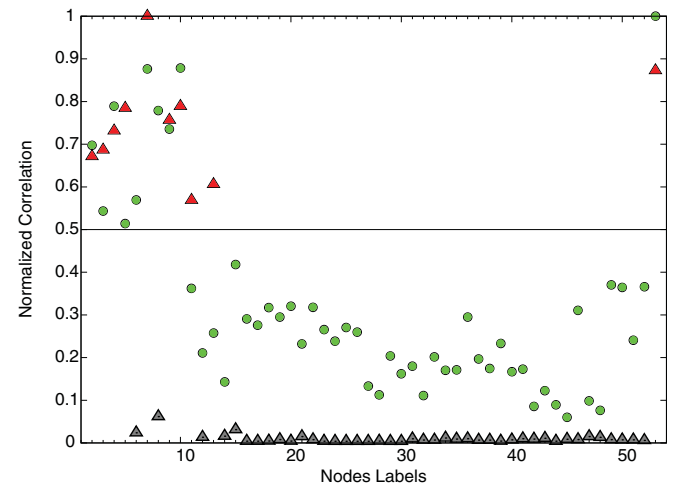

FIG. 9. (Color online) Normalized correlations of the cortical brain network of a cat for a pacemaker on node 1 ($k_p = 10$). The circles are the correlation values calculated through Eq. (17) for $\omega = 1.2\omega_p^c$ ($\Delta t = 100$, $T_s = 20$). The triangles correspond to the correlations given by expression (18) when $\omega = 20\omega_p^c$, calculated on a time window $\Delta t = 1$ and waiting a transient time $T_s = 0.1$. All the values have been divided by the maximum of each set (excluding the autocorrelation). Notice that while in the first case there is an almost continuous spectrum of values, in the second one it is easy to identify a group of points [red triangles] above the line at 0.5 clearly separated from the rest. Those are the ten neighbors of node 1.
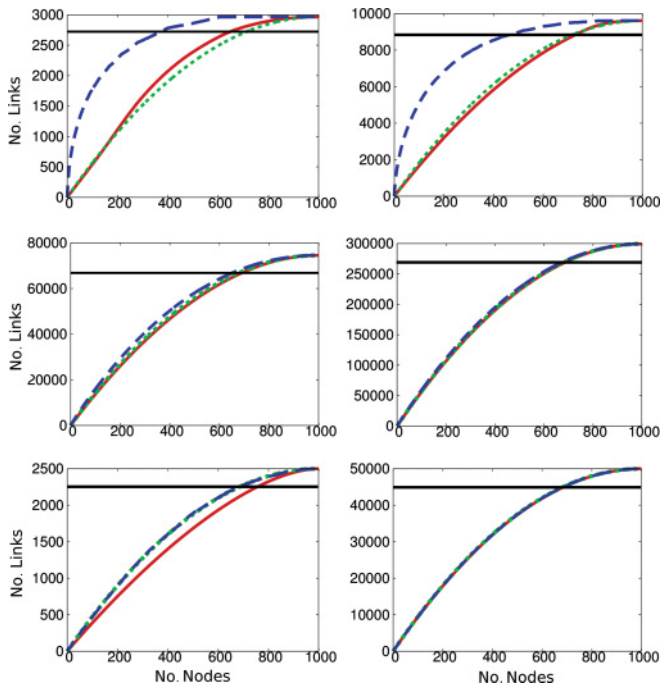
FIG. 10. (Color online) Average number of reconstructed links as a function of the number of nodes we considered as pacemakers (number of trials). From the top to the bottom, the considered networks are a pair of Barabasi-Albert networks, with parameter $k = 3$ (left) and $k = 10$ (right); a pair of Erdos-Reyni graphs with average degree equal to 15 (left) and 60 (right); and a pair of random regular graphs with degree 5 (left) and 100 (right). The size is $N = 1000$ for all plots. Different lines corresponds to different selection algorithms. Blue dashed lines stand for the ordered sequence on the basis of the critical frequency values; the red solid lines are for the random walk; the green dotted lines are for random extractions. Both the random walk and the random extractions are averaged over 1000 samples. The horizontal black line marks 90% of links: notice how in any case we never need more than 70% of the nodes in order to reconstruct 90% of the links, decreasing to 30%–40% in the case of the ordered sequence for scale-free networks. Correlations are computed under the same conditions as those in Fig. 9.

we may consider the choice of a certain pacemaker as the application of a drift on a given unit in a system of identical coupled oscillators. This means that it is possible to solve the problem with less than N experiments.

The criterion for choosing the ordered sequence of nodes on which we locate the pacemaker can vary. We may operate a random extraction or we may start from a randomly chosen node and then move to one of its neighbors along a random walk. Another option, which is much more convenient, especially in the case of scale-free networks, can be adopted if the critical frequencies associated with each oscillator are known. We can order the nodes according to decreasing critical frequency, starting from the highest one. In this way we proceed from larger to smaller (estimated) degrees, taking an important advantage if the degree distribution is not uniform and there are hubs. The hubs, indeed, provide information about a large number of links by means of very few experiments (Fig. 10).

VII. CONCLUSIONS

Systems of nonidentical Kuramoto oscillators have been recently shown to display a degree of synchronization that depends strongly on the topology of the underlying complex network. Here these dynamical properties, particularly by setting different types of correlations between the dynamical evolution of the oscillators, have been used to gather information on the connectivity patterns. Remarkably, this is the case for most experimental situations, where the *a priori* unknown connectivity of a particular network is inferred from purely dynamical measurements.

When the oscillators are identical (all of them having the same natural frequency), any topological configuration has a unique attractor, which is the complete synchronized state, meaning that the oscillators end up in such a state that all effective frequencies and phases are identical. This state does not offer any information about the topology. We perturb this setting by allowing one of the oscillators to have a different natural frequency than the rest. This unit is called the pacemaker of the network. Such a perturbation causes the final state to no longer be phase synchronized. But if the natural frequency of the pacemaker is not very different from the value of the rest of the population, the system still will retain a certain degree of synchronization since the whole system can evolve with the same effective frequency. However, if the frequency difference becomes larger, the system will be unable to find any kind of synchronization. The threshold between the former case and this latter case is a well-defined value, which is strictly dependent on the location of the pacemaker in the network. In this context, we can use the correlations between the effective frequencies of the oscillators in such an incoherent state to reproduce the network connectivity.

Moreover, we show that the dynamical correlations in different situations, whether close to or far from the critical point, provide complementary information on the network.

(1) Working around the critical point, we are able to estimate the degree of each pacemaker merely by its critical frequency.

(2) Slightly above the transition point the hierarchical structure of the whole network (related to functional modules) can be obtained from the correlations between effective frequencies. A further refinement enables us to recover the whole connection network with a good degree of accuracy.

(3) Far above the critical point it is possible to recognize which oscillators are directly connected to an individual pacemaker from a very short measurement of the time evolution of the effective frequencies. In this way we can recover the connectivity pattern, and this method turns out to be much more precise and more efficient than the previous one.

In summary, this paper deals with different approaches relating dynamical properties of individual nodes to the topology of the network. The topological properties inferred from dynamics can be local (the existence of a link between two nodes) as well as global (hierarchical organization of the nodes in the functional network). In particular, for a scale-free network and if the node degrees are known (or have been estimated from the critical frequencies), considering 30% of the possible pacemakers, always selecting the most connected nodes, will be enough to reconstruct approximately 90% of the links.

Other papers have considered the reconstruction of the network from dynamical information. Similar to our proposal with specific targets, Tegner *et al.* [36] analyzed the dynamical response of a gene-regulatory network by changing expression levels of particular genes. On the other hand, di Bernardo *et al.* [37] considered the global effect of different types of perturbations to infer the network topology. This approach has been followed recently also by Gorur Shandilya and Timme [38], who assumed that there is some information about the dynamical evolution of the isolated units and about the coupling. Our method, based on the change of the frequency of a single unit and how it enhances correlations among the nodes, can be more effective in oscillatory systems. In any case, for practical purposes the method chosen will depend on the specific details of the experimental setup, and even a combination of different ones can be the most appropriate.

[1] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[2] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).

[3] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Phys. Rep. **424**, 175 (2006).

[4] A. Barrat, M. Barthelemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, Cambridge, 2008).

[5] A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization* (Cambridge University Press, Cambridge, 2001).

[6] G. V. Osipov, J. Kurths, and C. Zhou, *Synchronization in Oscillatory Networks* (Springer, Berlin, 2007).

[7] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, Phys. Rep. **469**, 93 (2008).

[8] A. T. Winfree, *The Geometry of Biological Time* (Springer, Berlin, 1980).

[9] Y. Kuramoto, in *International Symposium on Mathematical Problems in Theoretical Physics*, Lecture Notes in Physics Vol. 39, edited by H. Araki (Springer, New York, 1975), pp. 420–422.

[10] J. A. Acebrón, L. L. Bonilla, C. J. Pérez-Vicente, F. Ritort, and R. Spigler, Rev. Mod. Phys. **77**, 137 (2005).

[11] There are, however, some limitations to these results, for instance, when the units are placed in regular lattices, such as one-dimensional rings, where other attractors different from the synchronized (equal phases) state may arise [39].

[12] L. Buzna, S. Lozano, and A. Díaz-Guilera, Phys. Rev. E **80**, 066120 (2009).

[13] H. Kori and A. S. Mikhailov, Phys. Rev. Lett. **93**, 254101 (2004).

[14] F. Radicchi and H. Meyer-Ortmanns, Phys. Rev. E **73**, 036218 (2006).

[15] F. Mori, Phys. Rev. Lett. **104**, 108701 (2010).

[16] A. Arenas, A. Díaz-Guilera, and C. J. Pérez-Vicente, Phys. D **224**, 27 (2006).

[17] E. Ravasz and A.-L. Barabási, Phys. Rev. E **67**, 026112 (2003).

[18] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence* (Springer, New York, 1984).

[19] I. Sendiña-Nadal, J. Buldu, I. Leyva, and S. Boccaletti, PLoS ONE **3**, e2644 (2008).

[20] Applying the same argument to the equation of the ith node, we obtain $\omega_p^c \leqslant k_i N$, whose smallest possible value is N, which is the largest possible value for the bound (9).

[21] It is not necessary to take into account the groups that do not include the pacemaker since the bound we obtain for ω_p^c when summing up $n + 1$ equations including the pacemaker or the remaining $N - n + 1$ (not including the pacemaker) is the same.

[22] M. Girvan and M. E. J. Newman, Proc. Natl. Acad. Sci. USA **99**, 7821 (2002).

[23] P. M. Gleiser and D. H. Zanette, Eur. Phys. J. B **53**, 233 (2006).

[24] A. Arenas, A. Díaz-Guilera, and C. J. Pérez-Vicente, Phys. Rev. Lett. **96**, 114102 (2006).

[25] A. Arenas, A. Fernandez, and S. Gomez, Handb. Biol. Networks **10**, 243 (2009).

[26] P. H. A. Sneath and R. R. Sokal, *Numerical Taxonomy* (W. H. Freeman, San Francisco, 1973).

[27] L. Zemanová, C. Zhou, and J. Kurths, Phys. D **224**, 202 (2006).

[28] C. Zhou, L. Zemanová, G. Zamora, C. C. Hilgetag, and J. Kurths, Phys. Rev. Lett. **97**, 238103 (2006).

[29] The cortical area that is not a superhub is a border area that can be seen as a hub only joined with another one very similar to it but is regarded as a superhub in itself anyway.

[30] J. L. Rodgers and W. A. Nicewander, Am. Stat. **42**, 59 (1988).

[31] R. Sharan, I. Ulitsky, and R. Shamir, Mol. Syst. Biol. **3**, 88 (2007).

[32] S. C. Janga, J. J. Díaz-Mejía, and G. Moreno-Hagelsieb, Metab. Eng. **13**, 1 (2011).

[33] S. Fortunato and M. Barthelemy, Proc. Natl. Acad. Sci. USA **104**, 36 (2007).

[34] J. W. Scannell, G. A. P. C. Burns, C. C. Hilgetag, M. A. O'Neil, and M. P. Young, Cereb. Cortex **9**, 277 (1999).

[35] M. Timme, Phys. Rev. Lett. **98**, 224101 (2007).

[36] J. Tegner, M. K. S. Yeung, J. Hasty, and J. Collins, Proc. Natl. Acad. Sci. USA **100**, 5944 (2003).

[37] D. di Bernardo, M. Thompson, T. Gardner, S. Chobot, E. Eastwood, A. Wojtovich, S. Elliott, S. Schaus, and J. Collins, Nat. Biotechnol. **23**, 377 (2005).

[38] S. Gorur Shandilya and M. Timme, New J. Phys. **13**, 013004 (2011).

[39] A. Díaz-Guilera and A. Arenas, in *Bio-Inspired Computing and Communication*, Lecture Notes in Computer Science Vol. 5151, edited by P. Liò, E. Yoneki, J. Crowcroft and D. C. Verma (Springer, Berlin Heidelberg, 2008), pp. 184-191.