Entangled networks, super-homogeneity and optimal network topology

Luca Donetti,[1] Pablo I. Hurtado,[2] and Miguel A. Muñoz[1]

[1] *Departamento de Electromagnetismo y Física de la Materia and Instituto Carlos I de Física Teórica y Comput. Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain*

[2] *Department of Physics, Boston University, Boston, Massachusetts 02215, USA*

(Dated: February 10, 2005)

A new family of graphs, *entangled networks*, with optimal properties in many respects, is introduced. By definition, their topology is such that optimizes synchronizability for generic dynamical processes. These networks are shown to have an extremely homogeneous structure: degree, node-distance, betweenness, and loop distributions are all very narrow. They are characterized as well by a very interwoven (entangled) structure with short average distances, large loops, and no well-defined community-structure (poor modularity). More importantly, we show that this family of nets exhibits an excellent performance with respect to other connectivity or flow properties such as robustness against errors and attacks, minimal first-passage time of random walks, good searchability, efficient communication, etc. These remarkable features convert entangled networks in a powerful and useful concept, optimal or almost-optimal in many senses, and with plenty of potential applications in network design, computer science, or neuroscience.

PACS numbers: 89.75.Hc,05.45.Xt,87.18.Sn

The ubiquitous presence of networks in Nature and social sciences is one of the main findings in the study of complex systems. The topology of such networks has been profusely studied [1] and some basic architectures have been discovered. The *scale-free* one, characterized by a power-law connectivity-distribution, is probably the most widely studied and celebrated, while other examples are small-world, hierarchical, Apollonian, static networks, etc [1]. Right after the first topological studies, the interest has shifted to the analysis of functional or dynamical aspects of processes occurring on networks, the evolution of the network topology, and the interplay between these last two dynamical features. Indeed, this "network perspective" has become a new paradigmatic way to look at complex systems. One particular issue that has attracted much interest because of its conceptual relevance and practical implications is the study of the *synchronizability* of individual dynamical processes occurring at the vertices of a given network. How does synchronizability depend upon network topology? This problem is much more general than it seems at first sight, as it is directly related to the question of how difficult it is to transmit information across the net or how difficult is for the sites to "talk" to each other. For example, a recently addressed important task is to determine the most efficient topology for communication networks both with and without traffic congestion [2]. Other problems as the minimization of first-passage times of random walkers on networks, the optimal topology in social networks to reach consensus, or the performance optimization of Hopfield neural-networks [3, 4] are also similar in essence. Hence, the issue of synchronizability is linked to many specific problems in different disciplines as computer science, biology, sociology, etc. [2, 5]. Some aspects of these problems have been already tackled; a key contribution is due to Barahona and Pecora (BP) [5] who established a criterion based on spectral techniques to determine the stability of synchronized states on networks.

The criterion is as follows. Consider a dynamical process $\dot{x}_i = F(x_i) - \sigma \sum_j L_{ij} H(x_j)$, where x_i with $i \in 1, 2, ..., N$ are dynamical variables, F and H are the evolution and the coupling functions respectively, σ is a constant, and L_{ij} is the Laplacian matrix, defined by $L_{ii} = k_i$ (the connectivity degree of node i), $L_{ij} = -1$ if nodes i and j are connected, and $L_{ij} = 0$ otherwise. A standard linear stability analysis can be performed by i) expanding around a synchronized state $x_1 = x_2 = \ldots = x_N = x^s$ with x^s solution of $\dot{x^s} = F(x^s)$, ii) diagonalizing L to find its N eigenvalues $0 = \lambda_1 < \lambda_2 \leq \ldots \leq \lambda_N$, and iii) writing equations for the normal modes y_i of perturbations $\dot{y}_i = [F'(x^s) - \sigma\lambda_i H'(x^s)]\, y_i$ which have all the same form but different effective coupling $\alpha = \sigma\lambda_i$. BP observed that the maximum Lyapunov exponent is in general negative only within a bounded interval $[\alpha_A, \alpha_B]$, and a decreasing (increasing) function below (above) (see fig. 1 in [5]). Requiring all effective couplings to lie within such an interval, $\alpha_A < \sigma\lambda_2 \leq \ldots \leq \sigma\lambda_N < \alpha_B$, it is straightforward to conclude that a synchronized state is linearly stable on a network if and only if $\frac{\lambda_N}{\lambda_2} < \frac{\alpha_B}{\alpha_A}$. Notice that the left hand side depends only on the network topology while the right hand side depends exclusively on the dynamics (through F and G, and x^s). Moreover, the interval $(\alpha_A,\ \alpha_B)$ in which the synchronized state is stable is larger for smaller eigenratios λ_N/λ_2, whence one concludes that *a network exhibits better synchronizability if the ratio $Q = \lambda_N/\lambda_2$ is as small as possible*, independently of the dynamics.

This letter is devoted (i) to build-up networks with a fixed number of nodes N and average connectivity $\langle k \rangle$, exhibiting a degree of synchronizability as high as pos-

sible (i.e. minimizing Q), (ii) to explore the topological features converting them into highly synchronizable, and (iii) to study their connection to networks optimizing other flow or connectivity properties relevant in neurocomputing, computer science, or graph theory.

First, we overview how Q behaves in some well-known topologies. For networks with the small-world property [1] Q is smaller than for deterministic graphs or purely random networks [5]. This was attributed to the existence of short characteristic paths between sites. However, Nishikawa *et al.* in a study of other small-world networks concluded that Q decreases as some heterogeneity measures decrease, even if the average distance increases [6]. Also, Hong et al. concluded that Q decreases whenever the betweenness heterogeneity decreases [7]. In order to extend and systematize these results and construct optimal synchronizable networks, and in the absence of a better strategy, we define a numerical algorithm able to minimize Q and search for such optimal nets.

Our optimization algorithm is a modified simulated annealing [8] initialized with a random network with N nodes and an average connectivity-degree $\langle k\rangle$. At each step the number of rewiring trials is randomly extracted from an exponential distribution. Attempted rewirings are (i) rejected if the updated network is disconnected, and otherwise (ii) accepted if $\delta Q = Q_{final} - Q_{initial} < 0$, or (iii) accepted with probability [9] $p = \min\left(1, [1-(1-q)\delta Q/T]^{1/(1-q)}\right)$ (where T is a temperature-like parameter) if $\delta Q \geq 0$. In the $q \to 1$ limit the usual Metropolis algorithm is recovered [9], while we choose $q = -3$ as it gives the fastest convergence (though results do not depend on this). The first N rewirings are performed at $T = \infty$, and they are used to calculate a new T such that the largest δQ among the first N ones would be accepted with large probability; in particular, we take $T = (1-q)\cdot(\delta Q)_{max}$. T is kept fixed for $100N$ rewiring trials or $10N$ accepted ones, whichever occurs first. Then, T is decreased by 10% and the process iterated until there is no change during 5 successive temperature steps, assuming that a (relative) minimum of Q has been found. Most of these details can be changed without affecting significatively the final results, while the main drawback of the algorithm is that the calculation of eigenvalues is slow.

The network found by different runs of the algorithm is unique (in most of the cases) as long as N is small enough ($N \lesssim 30$), while they are slightly different if N is larger. This indicates that the eigenvalue-ratio absolute minimum is not always found, and that the evolving network can remain trapped in some "metastable" state. Nevertheless, the final values of Q are very similar from run to run as shown in fig. 1. This fact makes us confident that a reasonably good and robust approximation to the optimal topology is obtained in general, though, strictly speaking, we cannot guarantee that the optimal solution has been actually found. To gain some insight

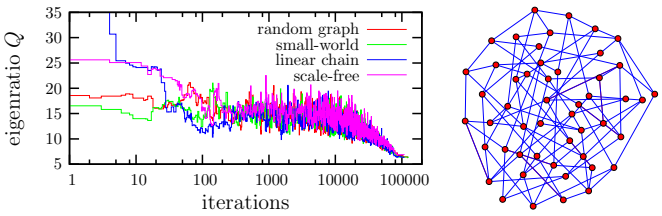

FIG. 1: Eigenvalue ratio, Q as a function of the number of algorithmic iterations. Starting from different initial conditions, with $N = 50$, and $\langle k\rangle = 4$, the algorithm converges to networks, as the depicted one, with very similar values of Q.

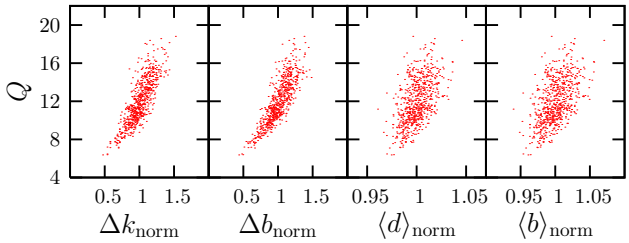

FIG. 2: Relation between the ratio Q and (i) node-connectivity standard deviation, (ii) betweenness standard deviation, (iii) average node-distance, and (iv) average betweenness. The subscript "norm" stands for normalization with respect to the respective mean-values, centering all the measured quantities around 1.

into the topological traits favoring a small Q, we measure different quantities during the algorithmic evolution and plot them versus the changing eigenratio. It turns out (as shown in fig. 2) that there is a strong correlation between the tendency of Q to decrease and an increase in the homogeneity (lowering variances) of the degree, average-distance and betweenness distributions. In a nutshell, the more synchronizable the network the more homogeneous it is. Also, the average distance and betweenness tend to diminish with Q, though these quantities are much less sensitive that their corresponding standard deviations (fig. 2). The emerging narrow betweenness distribution is in sharp contrast with that of networks with a strong community structure [10]. Indeed, a well known method to detect communities consists in removing progressively links with the largest betweenness [10]. The method leads to sound results whenever the betweenness is broadly distributed. Hence, well-defined communities do not exist in the emerging optimal networks.

Further inspection of these networks reveals another significant trait: the absence of short loops. This can be quantified by the *girth* (length of the shortest loop) or more accurately by the average length, $\langle \ell \rangle$, of the shortest loop passing through each node. Indeed, for small values of N and k, it is possible to identify the resulting optimized networks, as they have been studied in the mathematical literature: some of them are *cage graphs*. Let us recall that a (k, g)-cage graph is a k-regular graph (i.e. with a delta-peaked connectivity distribution) of

girth g having the minimum possible number of nodes. For $k = 3$ and $N = 10, 14$, and 24, respectively, the optimal nets found by the algorithm are cage-graphs with girth 5, 6, and 7 (called *Petersen*, *Heawood* and *McGee* graphs) respectively (see fig. 3 and [11]). For other values of N cage graphs do not exist but, in all cases, networks with very narrow shortest-loop distributions, with large mean values, are the optimal ones.

In general, we call the emerging structures **entangled networks**: all sites are very much alike (super-homogeneity) and the links form a very intrincated or interwoven structure (no community-structure, poor modularity, and large shortest-loops). Every single site is close to any other one (short average distance) owing not to the existence of intermediate highly connected hubs, as in scale-free nets [1], but as the result of a very "democratic" or entangled structure in which properties such as site-to-site distance, betweenness, and minimum-loop-size are very homogeneously distributed (see figs. 1, 3).

We have tried to use our (so far, partial) understanding of the entangled-topology to generate them more efficiently. For example, the constraint of homogeneity in the degree distribution can easily be implemented by starting up the simulation with regular graphs (or almost regular graphs) and performing changes respecting such a property (by randomly selecting pairs of links and exchanging their endpoints). A much faster convergence to optimal nets is obtained in this way. Other topological constraints are not so easy to implement. We have performed simulations using target functions different from Q in the optimization algorithm. Functions as the average distance, average betweenness, or homogeneity measures (such as the distance variance or the betweenness variance), or $\langle \ell \rangle$ are not sufficient: they need to be optimized simultaneously, in some proper way, to obtain reasonable outputs. We have tried different combinations of these quantities. The best convergence and results are obtained for the following combination of the betweenness, b, the betweenness variance, Δb and $\langle \ell \rangle$: $U = \frac{((\Delta b)^2 + \langle b \rangle^2)}{N} - \langle \ell \rangle$. The optimization of U is much faster than the minimization of the eigenratio as U is faster to compute than Q. For small networks the final result is as good as the one of the original method but, unfortunately, when N increases results worsen, though the computational time is always relatively small. This failure means that a full topological understanding of (large) entangled networks has not been reached yet.

In order to put our findings into context, we discuss some connections between our networks and known results and concepts in graph-theory. General considerations show that $\lambda_N \in [k, 2k]$ for regular graphs [12, 13]. As the variability of λ_N is very limited, optimizing Q is almost equivalent in most cases to maximizing λ_2 (also called *spectral gap*), as we have verified numerically. It is also known that for any family of regular graphs, G_m (m is the family index), in which the size N goes to infinity for large m the inequality $\lambda_2 \leq k - 2\sqrt{k-1}$ holds asymptotically, providing an upper bound for the spectral gap. Finally, it can be shown that for any family G_m in which the girth goes to infinity for large m, $\lambda_2 \geq k - 2\sqrt{k-1}$ asymptotically, meaning that the optimal gap value is obtained whenever the girth diverges for large N [12, 13]. Moreover, this remains true if the number of loops of finite length is not extensive. These results are in accordance with our observation of large girths and large $\langle \ell \rangle$ for entangled networks (even if they are not at the large-N limit). Another link with graph theory is provided by the concept of *expanders*. These are highly connected sparse graphs, with applications which include the design of super-efficient communication networks among many others [13, 14], and are defined as follows [12, 13]. Given a subset S of nodes in a graph G, its "edge boundary" is the set of links between nodes in S and nodes in its complement. The "expansion parameter" h of a graph G is the minimum ratio of the edge-boundary of a set and the set itself. A sequence of regular graphs G_m is a *family of expanders* if its size N diverges for large m and h is always larger than a given positive constant. This means that the boundary of any subset is always a non-vanishing fraction of the subset itself. Note that a large value of h corresponds to a very intrincated (entangled) network, where it is not possible to isolate subsets with a small boundary (or, in other words, where communities are poorly defined). Also, the expansion property is strictly related to the spectral gap [13]: $\frac{\lambda_2}{2} \leq h(G) \leq \sqrt{2k\lambda_2}$, meaning that (families of) entangled networks are expanders. *Ramanujan graphs* [13] are defined as k-regular graphs of size N with $\lambda_2 \geq k - 2\sqrt{k-1}$. Hence, these graphs are optimal expanders [13]. A family of entangled networks, will typically be a Ramanujan one (as λ_2 tends to be maximized) and, therefore, a (close to optimal) family of expanders. The explicit construction of expanders and Ramanujan graphs is a currently active field in graph-theory [13, 14], and it could serve as a starting point for explicit entangled-network design.

Finally, we discuss some properties of entangled networks as related to other optimization or flow problems:

i) In a recent paper [15], the optimization of network robustness against random and/or intentional removal of nodes has been studied. It was concluded that for generalized random graphs in the limit $N \to \infty$ the most robust topology (maximizing the percolation threshold) is characterized by a degree distribution with no more than 3 distinct node connectivities; i.e. with a rather homogeneous degree-distribution. To study the possible connection with our super-homogeneous entangled networks, let us recall that the initial topology we have considered (i.e. k-regular graphs) is already the optimal solution for robustness-optimization against errors and attacks in random networks [15]. A natural question to ask is whether further Q-optimization has some effect on the

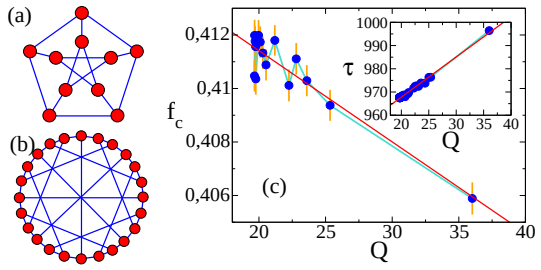

FIG. 3: Cage graphs for $k = 3$ and (a) $g = 5$ (Petersen) and (b) $g = 7$ (McGee). Right panel: Percolation threshold (main) and average first-passage time (inset) as a function of the eigenratio Q, as obtained during the optimization of a network with 500 nodes and $\langle k \rangle = 3$. The initial network corresponds to a 3-regular graph with $N = 500$.

network robustness. The answer is yes, as shown in fig. 3 where the percolation threshold for random or intentional attacks (which for regular graphs coincide), f_c, is plotted versus Q for a particular Q-optimization run. This further improvement of the robustness is possible because entangled networks include correlations, absent in random graphs [15], which favor the resilience to attacks. This tendency is maintained for increasing N, confirming that entangled networks are also extremely efficient from the robustness point of view (this remains true for *reliability* against link removal [16]).

ii) The problem of optimal topologies for local search with congestion has been tackled recently [2]. It has been shown that whenever the density of information packets traveling through the network is above a given threshold (congestion), the optimal network topology for packet flow towards its target node is a highly homogeneous, isotropic configuration, where all the nodes have essentially the same degree, betweenness, etc [2]. We encounter the same super-homogeneity, characteristic of optimally synchronizable graphs, revealing that entangled networks are also optimal (or perform extremely well) for packet flow and local searches with congestion.

iii) A typical measure of the network performance for flow properties is the average first-passage time, τ, of random walks. It is defined as the average time it takes to a random walker to arrive for the first time to a given node starting from another one. For a k-regular graph, τ can be expressed in terms of the Laplacian eigenvalues as $\tau \propto \sum \lambda_n^{-1}$, where the sum runs over all non-zero eigenvalues [17]. The largest contribution comes from $1/\lambda_2$, therefore minimizing Q guarantees a small τ (see inset in fig. 3), providing more evidence that entangled nets exhibit a very good performance for flow problems.

iv) Recently Kim concluded that neural networks with lower clustering coefficient exhibit much better performance than others [4]. Entangled nets have a very low clustering coefficient as only large loops exist and, therefore, they are natural and excellent candidates to have a good performance and large capacity.

All these features suggest that entangled networks, defined here as networks which optimize synchronizability, are also extremely good with respect to many different highly desirable properties in networks. This allows us to state the following *conjecture*: Given N and an average number of links per site k, there exists a network topology (that of entangled nets) with (almost) optimal properties in a global sense, characterized by homogeneous degree, betweenness, and distance distributions, large girths, large average shortest loops, no community-structure, and small diameters. A more precise topological characterization of entangled graphs, as well as the definition of an algorithmic procedure to build them up (similar to those existing for expanders and Ramanujan graphs), remain open and challenging problems with a huge amount of potential applications. Finding clear-cut examples of this topology in the real world, beyond artificial networks, is also a fundamental task to further gauge the role played by this family of networks in Nature.

We acknowledge useful discussions with D. Cassi, and P. L. Krapivsky, as well as financial support from the spanish MCyT (FEDER) under project BFM2001-2841, the MECD, and the EU COSIN-project-IST2001-33555.

[1] S. H. Strogatz, Nature **410**, 268 (2001). A. L. Barabási, Rev. Mod. Phys. **74**, 47 (2002). M. E. J. Newman, SIAM Review **45**, 167 (2003).
[2] R. Guimera et al. Phys. Rev. Lett. **89**, 248701 (2002).
[3] J. J. Torres et al. Neurocomputing, **58-60**, 229 (2004).
[4] B. J. Kim, Phys. Rev. E **69**, 045101(R) (2004).
[5] M. Barahona and L. M. Pecora, Phys. Rev. Lett. **89**, 054101 (2002). L. M. Pecora and T. L. Carroll, Phys. Rev. Lett. **80**, 2109 (1998).
[6] T. Nishikawa et al. Phys. Rev. Lett. **91**, 014101 (2003)
[7] H. Hong et al. Phys. Rev. E. **69**, 067105 (2004).
[8] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Science **220** 671 (1983).
[9] T. J. P. Penna, Phys. Rev. E **51**, R1 (1995).
[10] M. Girvan, M. E. J. Newman, Proc. Natl. Acad. Sci. USA **99**, 7821-7826 (2002).
[11] See http://mathworld.wolfram.com/CageGraph.html
[12] B. Bollobás, *Extremal Graph Theory* Academic Press, New York. 1978. W. Tutte, *Graph Theory As I Have Known It*, Oxford U. Press, New York, (1998). F. Chung, *Spectral Graph Theory*, Number 92 in CBMS Region Conference Series in Mathematics. Am. Math. Soc., 1997.
[13] P. Sarnak, Notices of the AMS, **51**, 762 (2004). G. Davidoff, P. Sarnak and A. Valette, *Elementary Number Theory, Group Theory and Ramanujan Graphs*, London Math. Soc. Students Texts, 55, Cambridge (2003).
[14] O. Reingold, S. Vadhan, and A. Wigderson, Ann. of Math. **155**, 157 (2002).
[15] A.X.C. Valente, A. Sarkar, and H.A. Stone, Phys. Rev. Lett. **92**, 118702 (2004).
[16] W. Myrvold, in *Proceedings of the 8th Int. Conf. on Graph Theory, Combinatorics*, Vol. II, pag. 650 (1998).
[17] L. Lovász, in *Combinatorics, Paul Erdős is Eighty* (V2), Keszthely, Hungary (1993).