

Emergent self-organized complex network topology out of stability constraints

Juan I. Perotti,¹ Orlando V. Billoni,¹ Francisco A. Tamarit,¹ Dante R. Chialvo,² and Sergio A. Cannas^{1,*}

¹*Instituto de Física de la Facultad de Matemática,
Astronomía y Física (IFFAMAF-CONICET), Universidad Nacional de Córdoba,
Ciudad Universitaria, 5000 Córdoba, Argentina*

²*Department of Physiology, Northwestern University, Chicago, IL 60611, USA*

Although most networks in nature exhibit complex topology the origins of such complexity remains unclear. We introduce a model of a growing network of interacting agents in which each new agent's membership to the network is determined by the agent's effect on the network's global stability. It is shown that out of this stability constraint, scale free networks emerges in a self organized manner, offering an explanation for the ubiquity of complex topological properties observed in biological networks.

PACS numbers: 89.75.Hc, 89.75.Fb, 89.75.Da

Keywords: Complex networks, stability, evolution

Complex networks of interacting agents are ubiquitous, in a wide range of scales, from the microscopic level of genetic, metabolic and proteins networks to the macroscopic human level of the Internet [1, 2]. All of them exhibit high clustering and relatively short path length compared with random networks. In addition they show scale free topology, namely that the probability of a node to be connected to k other ones display a power law behavior $P(k) \sim k^{-\gamma}$, with exponents $\gamma < 3$, for large values of k [1, 2, 3]. Candidate models for the origin of this topology belong roughly to two classes. The first one is based on the celebrated Barabási-Albert model[4] which key ingredient is the so called *preferential attachment* mechanism, where new nodes have a larger probability to connect with already well connected nodes. The second class of models are based on rewiring algorithms designed to counterbalance the tendency to form hubs with a drift towards shaping random networks[5].

These models have successfully explained the origin of complexity in some networks, but it is recognized that another, equally large, number of cases can not be accounted by either class of models. In particular, growing biological networks involve the coupling of at least two dynamical processes. The first one concerns the addition of new nodes, attached either during a slow evolutionary (i.e., species lifetime) or a relatively faster developmental (i.e., organism life time) process. A second one is the node dynamics which affects and in turn is affected by the growing processes. It is reasonable to expect that different network topology could emerge out of these coupled processes. This Letter is dedicated to discuss a simple model of this problem, showing that complex networks do emerge under general realistic constraints.

Before introducing the model, and to fix ideas, let's dwell on some concrete general examples. First consider a food web, which are constructed through community assembly rules, strongly influenced by the underlying dynamics of species and specific interactions among them[6, 7]. Another example could be neuronal networks,

where the addition of hundreds of thousands new neurons is followed by a dynamical process in which neuronal dynamics and connectivity are interrelated in a way not fully understood. Yet a third example at another scale, could be imagined in the context of social networks, in which novate members can be accepted or rejected based on their individual contribution to a global interest, fitness, performance or profit. In the three examples it is relatively easy to visualize the two processes mentioned above. The consequence of adding a new member with a given connectivity affecting a global in/stability, is represented in these examples by the abundance/lack of food, the neuronal welfare/death or the profits' up/down, respectively. It needs to be noted from the outset that the interest in this letter is not to describe an arbitrary algorithm, but on identifying a dynamical process able to be implemented by natural systems.

Let's consider a system of n interactive agents, whose dynamics is given by a set of differential equations

$$\frac{d\vec{x}}{dt} = \vec{F}(\vec{x}) \quad (1)$$

where \vec{x} is an n -component vector describing the relevant state variables of each agent and \vec{F} is an arbitrary non-linear function. One could imagine that \vec{x} in different systems may represent concentrations of some hormones, or the average density populations in a food web, or the concentration of a chemicals in a biochemical network, or the activity of genes in a gene regulation net, etc. We assume that a given agent i interacts only with a limited set of $k_i < n$ ones; thus F_i depends only on the variables belonging to that set. This defines the interaction network, as was done previously[8].

We will assume that there are two time scales in the dynamics. On the long time scale (much larger than the observation time) the system is subjected to an external flux (migration, mutation, etc.) of new agents that interact with some of the previous ones and can be incorporated into the system or not, so n (and the whole set of

differential equations) can change. On short time scales we assume that n is constant and the dynamics already lead the system to a particular stable stationary state \vec{x}^* defined by $\vec{F}(\vec{x}^*) = 0$ [9]. The stability of that solution is determined by the eigenvalue with maximum real part of the Jacobian matrix $a_{i,j} \equiv \left(\frac{\partial F_i}{\partial x_j} \right)_{\vec{x}^*}$. Therefore a new agent will be incorporated to the network if its inclusion result in a new stable fixed point, that is, if the values of the interaction matrix $a_{i,j}$ are such that the eigenvalue with maximum real part λ of the enlarged Jacobian matrix is negative ($\lambda < 0$). Assuming that isolated agents will reach a stable state by itself after certain characteristic relaxation time, its diagonal elements of the matrix $a_{i,i}$ are negative and given unity value to further simplify the treatment [8]. Their interaction values, (i.e., the non-diagonal matrix elements $a_{i,j}$) will take random values (both positive and negative) taken from some statistical distribution.

These ideas are implemented in a numerical model as follows: At every step the network can either grow or shrink. In each step an attempt is made to add a new node to the existing network, starting from a single agent ($n = 1$). Based on the stability criteria discussed, the attempt can be successful or not. If successful, the agent is accepted, so the existing $n \times n$ matrix grows its size by one column and one row. Otherwise the novate agent will have a probability to be deleted together with some others nodes as further explained below.

More specifically, suppose that we have an already created network with n nodes, such that the $n \times n$ associated interaction matrix $a_{i,j}$ is stable. Then, for the attachment of the $n + 1$ th node we first choose its degree k_{n+1} randomly between 1 and n with equal probability. Then the new agent interaction with the existing network member i is chosen such that non-diagonal matrix elements $(a_{i,n+1}, a_{n+1,i})$ ($i = 1, \dots, n$) are zero with probability $1 - k_{n+1}/n$ and ω with probability k_{n+1}/n , where ω is a real random value uniformly distributed in $[-b, b]$. b determines the interaction range variability and it is one of the two parameters of the model. Then, we calculate numerically λ for the resulting $(n + 1) \times (n + 1)$ matrix. If λ is negative (stable matrix) the new node is accepted. If λ is positive (unstable matrix) it means that the introduction of the new node destabilized the entire system and we will impose that, either the new agent is eliminated or it remains but produces the extinction of a certain number of previous existing agents. In order to further simplify the numerical treatment, we will allow up to $q \leq k_{n+1}$ extinctions, taken from the set of k_{n+1} nodes connected to the new one [10]; q is the other parameter of the model. To choose which nodes are to be eliminated, we first choose one with equal probability in the set of k_{n+1} and remove it. If the resulting $n \times n$ matrix is stable, we start a new trial; otherwise, another node among the remaining $k_{n+1} - 1$ is sorted and removed, repeating

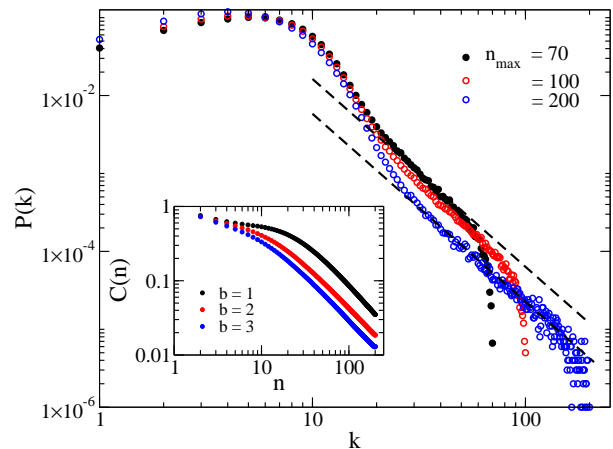


FIG. 1: (Color on-line) Degree distribution $P(k)$ for $b = 2$, $q = 3$ and different values of n_{max} ; the dashed lines correspond to a power law $P(k) \sim k^{-\gamma}$ with $\gamma \approx 2.4$. The inset shows the average connectivity $C(n)$ for different values of b .

the previous procedure. If after q removals the matrix remains unstable, the new node is removed, we return to the original $n \times n$ matrix and start a new trial [11].

First we calculated the average connectivity $C(n)$, defined as the fraction of non-diagonal matrix elements different from zero, averaged over different runs. We found that $C(n) \sim n^{-(1+\epsilon)}$ (inset of Fig. 1) for large values of n , where the exponent ϵ depends on b and q , taking values $0 < \epsilon < 1$. Such behavior is characteristic of food webs [12] and it has been interpreted in terms of self-organized criticality concepts [13]; the present results suggest that this is a general behavior in stability-driven self organized systems.

Next we calculated the degree distribution $P(k)$ of the network with $n = n_{max}$ for different values of b and q . The typical behavior of $P(k)$ is illustrated in Fig. 1 for $b = 2$, $q = 3$ and different values of n_{max} . We see the clear emergence of a power law tail $P(k) \sim k^{-\gamma}$ for large values of n , with a well defined exponent γ , independently of the network size (this figure also shows that the drop in the tail of the distribution is a finite size effect). In Fig. 2 we show $P(k)$ for different values of b at fixed q , where the inset shows the value of the exponent γ as a function of b for different values of q . We see that γ presents a minimum around $b = 2$ for all values of q ; as q increases the exponent decreases and for large enough values of q we obtain a non-trivial value of $\gamma < 3$ for a broad range of values of b .

To exclude the possibility that the observed network topology is trivially associated with a hidden preferential attachment process, we computed the attachment probability $\Pi(k)$, defined as the probability that a new node connects with an already existing node with degree k . Assuming that the average degree $\langle k_i \rangle \ll n$, the attachment probability can be expressed as $\Pi(k) = \sum_i^{n_k} \Pi_i$,

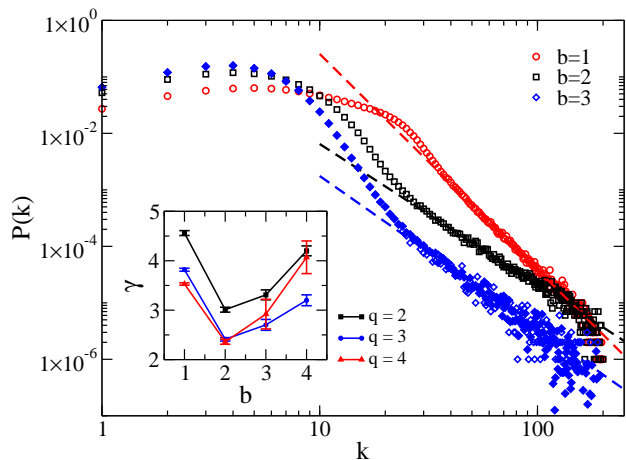


FIG. 2: (Color on-line) Degree distribution $P(k)$ for $q = 3$ different values of b . The inset shows γ as a function of b for different values of q .

where Π_i is the probability that the new node connects to the already existing node i , $n_k \approx n P(k)$ is the number of nodes with degree k and the sum runs over all sites i with degree $k_i = k$. If stability selection would favor some kind of preferential attachment mechanism, (i.e., if new nodes are attached with larger probability to nodes highly connected) we should expect

$$\Pi_i = \frac{k_i}{\sum_{j=1}^n k_j} \approx \frac{k_i}{n(n-1)C(n)} \quad (2)$$

and therefore

$$\Pi(k) \approx \frac{P(k)k}{(n-1)C(n)}. \quad (3)$$

In Fig. 3 the relative attachment probability $\Pi(k)/P(k)$ in the present model for a fixed network size n and different values of b it is compared with the corresponding results for a network of the same size obtained with the Barabási-Albert (BA) [4] algorithm with connectivity $C(n)$. The relative attachment probability shows the expected behavior $\Pi(k)/P(k) \sim k$ for large values of k , consistently with Eq.(3). In the present model the relative attachment probability remains almost constant for a wide range of values of k (including a range of values for which the power law behavior of $P(k)$ has already established), but displays an increasing trend consistent with Eq.(3) for large values of k . In other words, in the present model at variance with the BA model, as the network grows, the assembly mechanisms selected by stability shows a crossover between two regimes: one dominated by preferential attachment and the other not.

Considering that biological systems are probably never in a completely stable situation, we relaxed the condition

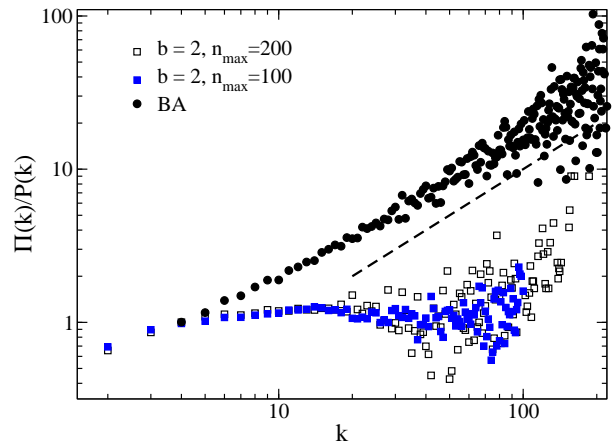


FIG. 3: (Color on-line) Relative attachment probability $\Pi(k)/P(k)$ for different values of $n_{max} = 200$, compared with the corresponding results for a BA network of the same size. The dashed line corresponds to a linear behavior $\Pi(k)/P(k) \sim k$.

of stability $\lambda < 0$ and look at networks growing by allowing λ to take small positive values so that the characteristic time to leave an unstable fixed point $\tau = \lambda^{-1} \gg 1$. By accepting nodes as long as $\lambda < \Delta$ the calculation of $P(k)$ for different values of Δ (positive and negative) showed similar qualitative behavior, with small variations of the γ exponent.

Next we calculated the average path length L between two nodes and the average cluster coefficient Cc for the networks obtained by the present algorithm as a function of the network size n . L is defined as the minimum number of links needed to connect any pair of nodes in the network and Cc is defined as the fraction of connections between topological neighbors of any site[14]. In Fig. 4 we show the typical behavior of $L(n)$ and $Cc(n)$. We see that $Cc(n) \sim n^{-0.75}$ and $L(n) \sim A \ln n + C$. Such scaling behavior is the same one observed in the BA model[1].

As shown in Fig. 4 larger networks becomes less clustered and have longer minimum path on the average. Although in all cases these networks are small world (in the sense that $L \sim \log(n)$), there is an important difference with the networks seen in the Watts-Strogatz scenario[14] in which the addition of shortcuts changes in the same direction both L and Cc . In the present model, L and Cc are inversely related as it can be seen in the single run plotted in panel b of Fig. 4. The data corresponds to values computed every fifty trials, independently if the attempt to add a node was successful or not at that trial. In a sense, this is how a natural network would look like to an observer if one could take snapshots in time. Clearly both quantities fluctuate in opposite directions, as further shown in the inset where the data corresponding to a randomly shuffled network is also plotted for comparison. The behavior of Cc and L is linked with the selection

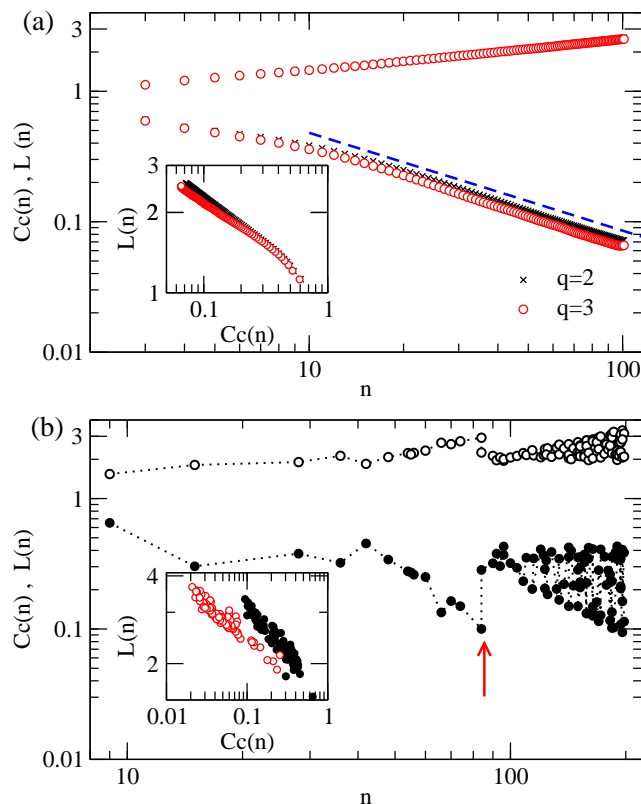


FIG. 4: (Color on-line) (a) Networks' average clustering coefficient $Cc(n)$ and $L(n)$ for $b = 2$ and different values of q as a function of network size. The dashed line is a guide to the eye corresponding to $Cc(n) \sim n^{-0.75}$. Inset shows the same data plotted against each other. (b) $Cc(n)$ and $L(n)$ as a function of network size computed from a single network realization. Data are samples taken every fifty trials, regardless of the resulting stability. Notice how fluctuations increase as the network grows. Inset shows the same data plotted against each other (full circles), in addition to the data computed from a random network with equal size and density of connections (open circles).

dynamics ruling which node is accepted or rejected. The stability constraint favors the nodes with few links, since they modify the matrix $a_{i,j}$ stability much less than new nodes with many links (of course this is reflected in the $P(k)$ density). Thus, most frequently the network grows at the expense of adding nodes with one or few links, producing an increase of L and a decrease of Cc . Most of the times, nodes with many links destabilize the network and are rejected, but when one is finally accepted, a large decrease in L together with an increase in Cc is observed. This sudden change is the signature of a new network hub, as seen in the example denoted with an arrow in Fig. 4b. We also verified that those fluctuations lead to a slow diffusive-like growth of the network size $n(t) \sim t^{1/2}$ (not shown), where the time is measured in number of trials.

The robustness of scale free networks against error

and attack has already been addressed[15] at the level of deleting nodes or links. The present results suggest that it would be interesting to re-study similar properties in terms of finite perturbations affecting the network stability.

Summarizing, the analysis of a very simple model shows that networks with scale free topology can appear as an evolutionary/developmental emergent property, driven by stability selection pressure. This suggests yet another explanation for the ubiquity of scale free topology observed in different networks, especially in natural networks.

Fruitful discussions with P. M. Gleiser are acknowledged. Work supported by CONICET (Argentina), SeCyT, Universidad Nacional de Córdoba (Argentina), FONCyT grant PICT-2005 33305 (Argentina) and NINDS (USA).

* Electronic address: cannas@famaf.unc.edu.ar

- [1] R. Albert and A. -L. Barabási. Reviews of Modern Physics, **74**, 47-97 (2002).
- [2] R. V. Solé, R. Ferrer-Cancho, J. M. Montoya, S. Valverde. Complexity, **8**, 20-33 (2003).
- [3] O. Sporns, D. Chialvo, M. Kaiser, C. Hilgetag. TICS, **63**, 031912 (2004).
- [4] A. -L. Barabási and R. Albert. Science, **286**, 509-512 (1999).
- [5] M. Baiesi and S. S. Manna. Phys. Rev. E, 047103,(2003).
- [6] E. Weiher and P. Keddy. Ecological Assembly Rules, Cambridge U. Press (1999).
- [7] S. L. Pimm, The balance of Nature, The University of Chicago Press (1991).
- [8] R. M. May. Nature. **238**, 413-414 (1972).
- [9] We consider here only stable fixed points of the dynamics which is, at least in ecology, a common assumption[7]. Nevertheless, the proposed mechanism is expected to work as well for more complex attractors (e.g. limit cycles).
- [10] This rule is a bit restrictive, in the sense that one should rather expect an avalanche of extinctions once the system destabilizes. However this minimal network model is already enough to demonstrate scale free topology. An extension allowing extinction avalanches can be considered in the future.
- [11] The process is repeated for $n = n_{max}$ (typically $n_{max} = 200$) and restarted 10^5 times from $n = 1$ to obtain statistics of the networks properties for different values of b and q .
- [12] R. V. Solé and B. Goodwin. Signs of Life (Basic Books, 2000).
- [13] R. V. Solé, D. Alonso, A. McKane. Physica A, **286**, 337-344 (2000).
- [14] D. J. Watts and S. H. Strogatz. Nature, **393**, 440-442 (1998).
- [15] R. Albert, H. Jeong, A. -L. Barabási. Nature, **406** 378-382 (2000).
- [16] M. Aldana and P. Cluzel. Proc. Natl. Acad. Sci. USA, **100**, 8710-8714 (2003).