

# Evolving networks and the development of neural systems

Samuel Johnson, J. Marro and Joaquín J. Torres

Departamento de Electromagnetismo y Física de la Materia and Instituto Carlos I de Física Teórica y Computacional, Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain.

E-mail: samuel@onsager.ugr.es, jmarro@ugr.es and jtorres@onsager.ugr.es

**Abstract.** It is now generally assumed that the heterogeneity of most networks in nature probably arises via preferential attachment of some sort. However, the origin of various other topological features, such as degree-degree correlations and related characteristics, is often not clear and attributed to specific functional requirements. We show how it is possible to analyse a very general scenario in which nodes gain or lose edges according to any (e.g., nonlinear) functions of local and/or global degree information. Applying our method to two rather different examples of brain development – synaptic pruning in humans and the neural network of the worm *C. Elegans* – we find that simple biologically motivated assumptions lead to very good agreement with experimental data. In particular, many nontrivial topological features of the worm’s brain arise naturally at a critical point.

PACS numbers: 64.60.aq, 89.75.Fb, 87.85.dm, 05.40.-a

Submitted to: *J. Stat. Mech.* (accepted, 2010)

## 1. Introduction

The conceptual simplicity of a *network* – a set of nodes, some pairs of which connected by edges – often suffices to capture the essence of cooperation in complex systems. Ever since Barabási and Albert presented their evolving network model [1], in which linear preferential attachment leads asymptotically to a scale-free degree distribution (the degree,  $k$ , of a node being its number of neighbouring nodes), there have been many variations or refinements to the original scenario [2, 3, 4, 5, 6, 7] (for a review, see [8]). In [9], we showed how topological phase transitions and scale-free solutions could emerge in the case of nonlinear rewiring in fixed-size networks. Now we extend our scope to more general and realistic situations, considering the evolution of networks making only minimal assumptions about the attachment/detachment rules. In fact, all we assume is that these probabilities factorize into two parts: a local term which depends on node degree, and a global term, which is a function of the mean degree of the network.

Our motivation can be found in the mechanisms behind many real-world networks, but we focus, for the sake of illustration, on the development of biological neural networks, where nodes represent neurons and edges play the part of synaptic interaction [10, 11, 12]. Experimental neuroscience has shown that enhanced electric activity induces synaptic growth and dendritic arborization [13, 14, 15, 16]. Since the activity of a neuron depends on the net current received from its neighbours, which is higher the more neighbours it has, we can see node degree as a proxy for this activity – accounting for the local term alluded to above. On the other hand, synaptic growth and death also depend on concentrations of various molecules, which can diffuse through large areas of tissue and therefore cannot in general be considered local. A feature of brain development in many animals is *synaptic pruning* – the large reduction in synaptic density undergone throughout infancy. Chechik *et al.* [17, 18] have shown that via an elimination of less needed synapses, this can reduce the energy consumed by the brain (which in a human at rest can account for a quarter of total energy used) while maintaining near optimal memory performance. Going on this, we will take the mean degree of the network – or mean synaptic density – to reflect total energy consumption, hence the global terms in our attachment/detachment rules.

An alternative approach would be to consider some kind of model neurons explicitly and couple the probabilities of synaptic growth and death to neuronal dynamic variables, such as local and global fields. In a Hopfield network, for example, the expected value of the field (total incoming current) at node  $i$  is proportional to  $i$ 's degree [19], the total current (energy consumption) in the network therefore being proportional to the mean degree; qualitatively, these observations are likely to hold also in more realistic situations [20], although relations need not be linear. Co-evolving networks of this sort are currently attracting attention, with dynamics such as Prisoner's Dilemma [21], Voter Model [22] or Random Walkers [23]. Although we consider this line of work particularly interesting, for generality and analytical tractability we opt here

to use only topological information for the attachment/detachment rules, although our results can be applied to any situation in which the dynamical states of the elements at the nodes can be functionally related to degrees<sup>‡</sup>.

Following a brief general analysis, we show how appropriate choices of functions induce the system to evolve towards heterogeneous (sometimes scale-free) networks while undergoing synaptic pruning in quantitative agreement with experiments. At the same time, degree-degree correlations emerge naturally, thus making the resulting networks *disassortative* – as tends to be the case for most biological networks – and leading to realistic small-world parameters.

## 2. Basic considerations

Consider a simple undirected network with  $N$  nodes defined by the adjacency matrix  $\hat{a}$ , the element  $\hat{a}_{ij}$  representing the existence or otherwise of an edge between nodes  $i$  and  $j$ . Each node can be characterized by its degree,  $k_i = \sum_j \hat{a}_{ij}$ . Initially, the degrees follow some distribution  $p(k, t = 0)$  with mean  $\kappa(t)$ . We wish to study the evolution of networks in which nodes can gain or lose edges according to stochastic rules which only take into account local and global information on degrees. So as to implement this in the most general way, we will assume that at every time step, each node has a probability of gaining a new edge,  $P_i^{\text{gain}}$ , to a random node; and a probability of losing a randomly chosen edge,  $P_i^{\text{lose}}$ . We assume these factorize as  $P_i^{\text{gain}} = u(\kappa)\pi(k_i)$  and  $P_i^{\text{lose}} = d(\kappa)\sigma(k_i)$ , where  $u$ ,  $d$ ,  $\pi$  and  $\sigma$  can be arbitrary functions, but impose nothing else other than normalization.

For each edge that is withdrawn from the network, two nodes decrease in degree:  $i$ , chosen according to  $\sigma(k_i)$ , and  $j$ , a random neighbour of  $i$ 's; so there is an added effective probability of loss  $k_j/(\kappa N)$ . Similarly, for each edge placed in the network, not only  $l$  chosen according to  $\pi(k_l)$  increases its degree; a random node  $m$  will also gain, with the consequent effective probability  $N^{-1}$  (though see§). Let us introduce the notation  $\tilde{\pi}(k) \equiv \pi(k) + N^{-1}$  and  $\tilde{\sigma}(k) \equiv \sigma(k) + k/(\kappa N)$ . Network evolution can now be seen as a *one step process* [24] with transition rates  $u(\kappa)\tilde{\pi}(k)$  and  $d(\kappa)\tilde{\sigma}(k)$ . The expected value for the increment in a given  $p(k, t)$  at each time step (which we equate with a temporal derivative) defines a master equation for the degree distribution [9]:

$$\begin{aligned} \frac{dp(k, t)}{dt} &= u(\kappa)\tilde{\pi}(k-1)p(k-1) + d(\kappa)\tilde{\sigma}(k+1)p(k+1) \\ &\quad - [u(\kappa)\tilde{\pi}(k) + d(\kappa)\tilde{\sigma}(k)]p(k, t). \end{aligned} \quad (1)$$

Assuming now that  $p(k, t)$  evolves towards a stationary distribution,  $p_{\text{st}}(k)$ , then this must necessarily satisfy detailed balance since it is a one step process [24]; i.e., the

‡ For instance, the stationary distribution of walkers used for edge dynamics in [23] is actually obtained purely from topological information, although it can only be written in terms of local degrees for undirected networks.

§ We are ignoring the small corrections that arise because  $j \neq i$  and  $l \neq m$ , which in any case would disappear if self-connections were allowed.

flux of probability from  $k$  to  $k + 1$  must equal the flux from  $k + 1$  to  $k$ , for all  $k$  [25]. This condition (sufficient for (1) to be zero) can be written as

$$\frac{\partial p_{\text{st}}(k)}{\partial k} = \left[ \frac{u(\kappa_{\text{st}})}{d(\kappa_{\text{st}})} \frac{\tilde{\pi}(k)}{\tilde{\sigma}(k+1)} - 1 \right] p_{\text{st}}(k), \quad (2)$$

where we have substituted a difference for a partial derivative and  $\kappa_{\text{st}} \equiv \sum_k k p_{\text{st}}(k)$ . Setting  $\pi$  and  $\sigma$  so as to be normalized to one (i.e.,  $\sum_k p(k)\pi(k) = \sum_k p(k)\sigma(k) = 1, \forall t$ ), which is equivalent to saying that at each time step exactly  $u(\kappa)$  nodes are chosen to gain edges and  $d(\kappa)$  to lose them, then in the stationary state we will have  $u(\kappa_{\text{st}}) = d(\kappa_{\text{st}})$  since the total number of edges will be conserved. From (2) we can see that  $p_{\text{st}}(k)$  will have an extremum at some value  $k_e$  if it satisfies  $\tilde{\pi}(k_e) = \tilde{\sigma}(k_e+1)$ .  $k_e$  will be a maximum (minimum) if the numerator in (2) is smaller (greater) than the denominator for  $k > k_e$ , and viceversa for  $k < k_e$ . Assuming, for example, that there is one and only one such  $k_e$ , then a maximum implies a relatively homogeneous distribution, while a minimum means  $p_{\text{st}}(k)$  will be split in two, and therefore highly heterogeneous. More intuitively, if for nodes with large enough  $k$  there is a higher probability of gaining edges than of losing them, the degrees of these nodes will grow indefinitely, leading to heterogeneity. If, on the other hand, highly connected nodes always lose more edges than they gain, we will obtain quite homogeneous networks. From this reasoning we can see that a particularly interesting case (which turns out to be critical) is that in which  $\pi(k)$  and  $\sigma(k)$  are such that

$$\tilde{\pi}(k) = \tilde{\sigma}(k) \equiv v(k), \quad \forall k. \quad (3)$$

According to (2), condition (3) means that for large  $k$ ,  $\partial p_{\text{st}}(k)/\partial k \rightarrow 0$ , and  $p_{\text{st}}(k)$  flattens out – as for example a power-law does.

The standard Fokker-Planck approximation for the one step process defined by (1) is [24]:

$$\begin{aligned} \frac{\partial p(k, t)}{\partial t} &= \frac{1}{2} \frac{\partial^2}{\partial k^2} \{ [d(\kappa) \tilde{\sigma}(k) + u(\kappa) \tilde{\pi}(k)] p(k, t) \} \\ &+ \frac{\partial}{\partial k} \{ [d(\kappa) \tilde{\sigma}(k) - u(\kappa) \tilde{\pi}(k)] p(k, t) \}. \end{aligned} \quad (4)$$

For transition rates which meet condition (3), (4) can be written as:

$$\begin{aligned} \frac{\partial p(k, t)}{\partial t} &= \frac{1}{2} [d(\kappa) + u(\kappa)] \frac{\partial^2}{\partial k^2} [v(k) p(k, t)] \\ &+ [d(\kappa) - u(\kappa)] \frac{\partial}{\partial k} [v(k) p(k, t)]. \end{aligned} \quad (5)$$

Ignoring boundary conditions, the stationary solution must satisfy, on the one hand,  $v(k)p_{\text{st}}(k) = Ak + B$ , so that the diffusion is stationary, and, on the other,  $u(\kappa_{\text{st}}) = d(\kappa_{\text{st}})$ , to cancel out the drift. For this situation to be reachable from any initial condition,  $u(\kappa)$  and  $d(\kappa)$  must be monotonous functions, decreasing and increasing respectively.

### 3. Synaptic pruning

As a simple example, we will first consider global probabilities which have the linear forms:

$$u[\kappa(t)] = \frac{n}{N} \left( 1 - \frac{\kappa(t)}{\kappa_{\max}} \right) \quad \text{and} \quad d[\kappa(t)] = \frac{n}{N} \frac{\kappa(t)}{\kappa_{\max}}, \quad (6)$$

where  $n$  is the expected value of the number of additions and deletions of edges per time step, and  $\kappa_{\max}$  is the maximum value the mean degree can have. This choice describes a situation in which the higher the density of synapses, the less likely new ones are to sprout and the more likely existing ones are to atrophy – a situation that might arise, for instance, in the presence of a finite quantity of nutrients. Again taking into account that  $\pi$  and  $\sigma$  are normalized to one, summing over  $P_i^{\text{gain}} - P_i^{\text{lose}}$  we find that the increment in  $\kappa(t)$  is  $d\kappa(t)/dt = 2\{u[\kappa(t)] - d[\kappa(t)]\} = 2(n/N)[1 - 2\kappa(t)/\kappa_{\max}]$  (independently of the local probabilities). Therefore, the mean degree will increase or decrease exponentially with time, from  $\kappa(0)$  to  $\frac{1}{2}\kappa_{\max}$ . Assuming that the initial condition is, say,  $\kappa(0) = \kappa_{\max}$ , and expressing the solution in terms of the *mean synaptic density* – i.e.,  $\rho(t) \equiv \kappa(t)N/(2V)$ , with  $V$  the total volume considered – we have

$$\rho(t) = \rho_f \left( 1 + e^{-t/\tau_p} \right), \quad (7)$$

where we have defined  $\rho_f \equiv \rho(t \rightarrow \infty)$  and the time constant for pruning is  $\tau_p = \rho_f N/n$ . This equation was fitted in figure 1 to experimental data on layers 1 and 2 of the human auditory cortex|| obtained during autopsies by Huttenlocher and Dabholkar [26].

It seems reasonable to assume that the initial overgrowth of synapses is due to the transient existence of some kind of growth factors. If we account for these by including a nonlinear, time-dependent term  $g(t) \equiv a \exp(-t/\tau_g)$  in the probability of growth, i.e.,  $u[\kappa(t), t] = (n/N)[1 - \kappa(t)/\kappa_{\max} + g(t)]$ , leaving  $d[\kappa(t)]$  as before, we find that  $\rho(t)$  becomes

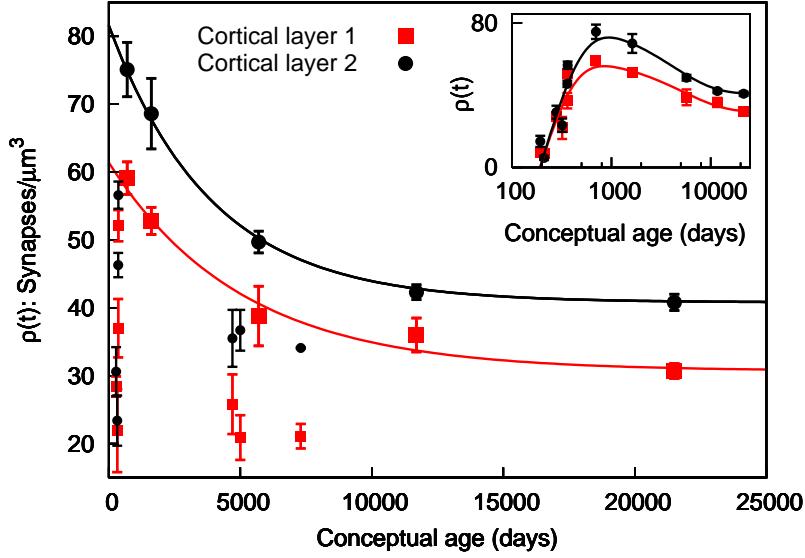
$$\rho(t) = \rho_f \left[ 1 + e^{-t/\tau_p} - \left( 1 + e^{-t_0/\tau_p} \right) e^{-\frac{t-t_0}{\tau_g}} \right], \quad (8)$$

where  $t_0$  is the time at which synapses begin to form ( $t = 0$  corresponds to the moment of conception) and  $\tau_g$  is the time constant related to growth. The inset in figure 1 shows the best fit to the auditory cortex data. Since the contour conditions  $\rho_f$  and (for (8))  $t_0$  are simply taken as the value of the last data point and the time of the first one, in each case, the time constants  $\tau_p$  and  $\tau_g$  are the only parameters needed for the fit.

### 4. Phase transitions

The drift-like evolution of the mean degree we have just illustrated with the example of synaptic pruning is independent of the local probabilities  $\pi(k)$  and  $\sigma(k)$ . The effect of these is rather in the diffusive behaviour which can lead, as mentioned, either to

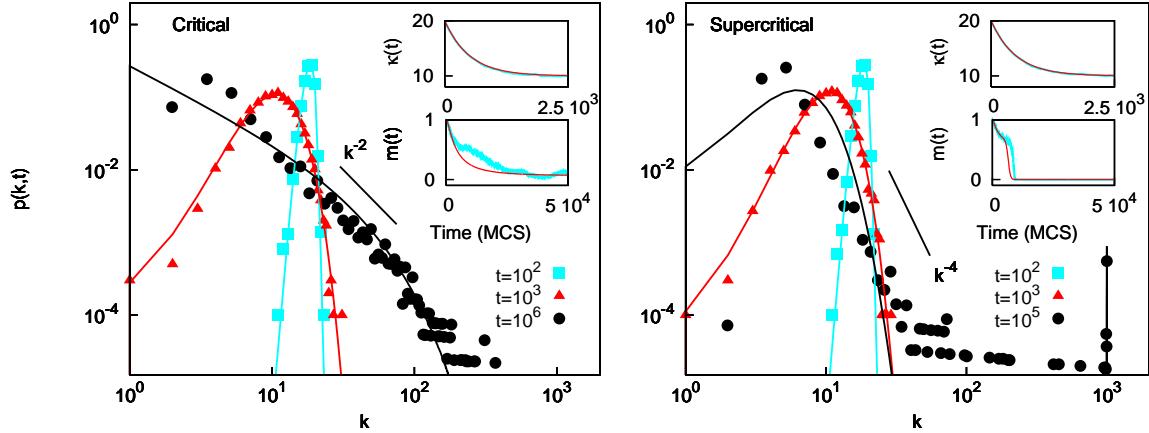
|| Data points for three particular days (smaller symbols) are omitted from the fit, since we believe these must be from subjects with inherently lower synaptic density.



**Figure 1.** Synaptic densities in layers 1 (red squares) and 2 (black circles) of the human auditory cortex against time from conception. Data from [26], obtained by directly counting synapses in tissues from autopsies. Lines follow best fits to (7), where the parameters were: for layer 1,  $\tau_p = 5041$  days; and for layer 2,  $\tau_p = 3898$  days (for  $\rho_f$  we have used the last data points: 30.7 and 40.8 synapses/ $\mu\text{m}^3$ , for layers 1 and 2 respectively). Data pertaining to the first year and to days 4700, 5000 7300, shown with smaller symbols, where omitted from the fit. Assuming the existence of transient growth factors, we can include the data points for the first year in the fit by using (8). This is done in the inset (where time is displayed logarithmically). The best fits were: for layer 1,  $\tau_g = 151.0$  and  $\tau_p = 5221$ ; and for layer 2,  $\tau_g = 191.1$  and  $\tau_p = 4184$ , all in days (we have approximated  $t_0$  to the time of the first data points, 192 days).

homogeneous or to heterogeneous final states. A useful bounded order parameter to characterize these phases is therefore  $m \equiv \exp(-\sigma^2/\kappa^2)$ , where  $\sigma^2 = \langle k^2 \rangle - \kappa^2$  is the variance of the degree distribution ( $\langle \cdot \rangle \equiv N^{-1} \sum_i \langle \cdot \rangle$  represents an average over nodes). We will use  $m_{\text{st}} \equiv \lim_{t \rightarrow \infty} m(t)$  to distinguish between the different phases, since  $m_{\text{st}} = 1$  for a regular network and  $m_{\text{st}} \rightarrow 0$  for one following a highly heterogeneous distribution. Although there are particular choices of probabilities which lead to (5), these are not the only critical cases, since the transition from homogeneous to heterogeneous stationary states can come about also with functions which never meet condition (3). Rather, this is a classic topological phase transition, the nature of which depends on the choice of functions [27, 28, 29].

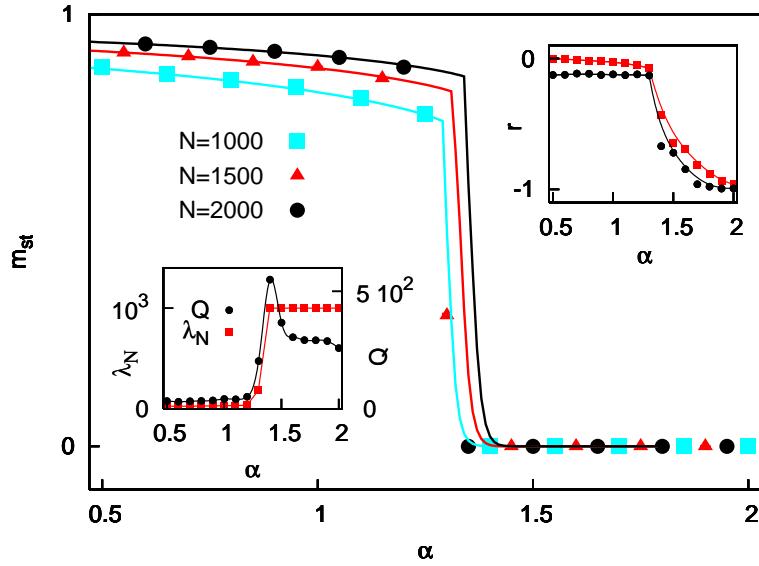
Evolution of the degree distribution is shown in figure 2 for critical and supercritical choices for the probabilities, as given by MC simulations (starting from regular random graphs) and contrasted with theory. (The subcritical regime is not shown since the stationary state has a distribution similar to the ones at  $t = 10^3$  MCS in the other regimes.) The disparity between the theory and the simulations for the final distributions is due to the build up of certain correlations not taken into account in our



**Figure 2.** Evolution of the degree distributions of networks beginning as regular random graphs with  $\kappa(0) = 20$  in the critical (top) and supercritical (bottom) regimes. Local probabilities are  $\sigma(k) = k/(\langle k \rangle N)$  in both cases, and  $\pi(k) = 2\sigma(k) - N^{-1}$  and  $\pi(k) = k^{3/2}/(\langle k^{3/2} \rangle N)$  for the critical and supercritical ones, respectively. Global probabilities as in (6), with  $n = 10$  and  $\kappa_{\max} = 20$ . Symbols in the main panels correspond to  $p(k, t)$  at different times as obtained from MC simulations. Lines result from numerical integration of (1). Insets show typical time series of  $\kappa$  and  $m$ . Light blue lines are from MC simulations and red lines are theoretical, given by (7) and (1), respectively.  $N = 1000$ .

analysis. This is because the existence of some very highly connected nodes reduces the probability of there being very low degree nodes. In particular, if there are, say,  $x$  nodes connected to the rest of the network, then a natural cutoff,  $k_{\min} = x$ , emerges. Note that this occurs only when we restrict ourselves to simple networks, i.e., with only one edge allowed for each pair of nodes. This topological phase transition is shown in figure 3, where  $m_{\text{St}}$  is plotted against parameter  $\alpha$  for global probabilities as in (6) and local ones  $\pi(k) \sim k^\alpha$  and  $\sigma(k) \sim k$ . This situation corresponds to one in which edges are eliminated randomly while nodes have a power-law probability of sprouting new ones (note that power-laws are good descriptions of a variety of monotonous response functions, yet only require one parameter). Although, to our knowledge, there are not yet enough empirical data to ascertain what degree distribution the structural topology of the human brain follows, it is worth noting that its functional topology, at the level of brain areas, has been found to be scale-free with an exponent very close to 2 [30].

In general, most other measures can be expected to undergo a transition along with its variance. For instance, highly heterogeneous networks (such as scale-free ones) exhibit the small-world property, characterized by a high *clustering coefficient*,  $C \gg \langle k \rangle / N$ , and a low *mean minimum path*,  $l \sim \ln(N)$  [31]. A particularly interesting topological feature of a network is its *synchronizability* – i.e., given a set of oscillators placed at the nodes and coupled via the edges, how wide a range of coupling strengths will result in them all becoming synchronized. Barahona and Pecora showed analytically that, for linear oscillators, a network is more synchronizable the lower the relation



**Figure 3.** Phase transitions in  $m_{st}$  for  $\pi(k) \sim k^\alpha$  and  $\sigma(k) \sim k$ , and  $u(\kappa)$  and  $d(\kappa)$  as in (6).  $N = 1000$  (blue squares), 1500 (red triangles) and 2000 (black circles);  $\kappa(0) = \kappa_{\max} = 2n = N/50$ . Corresponding lines are from numerical integration of (1). The bottom left inset shows values of the highest eigenvalue of the Laplacian matrix (red squares) and of  $Q = \lambda_N/\lambda_2$  (black circles), a measure of unsynchronizability;  $N = 1000$ . The top right inset shows transitions for the same parameters in the final values of Pearson's correlation coefficient  $r$  (see section 5), both for only one edge allowed per pair of nodes (red squares) and without this restriction (black circles).

$Q = \lambda_N/\lambda_2$  – where  $\lambda_N$  and  $\lambda_2$  are the highest and lowest non-zero eigenvalues of the Laplacian matrix ( $\hat{\Lambda}_{ij} \equiv \delta_{ij}k_i - \hat{a}_{ij}$ ), respectively [32]. The bottom left inset in figure 3 displays the values of  $Q$  and  $\lambda_N$  obtained for the different stationary states. There is a peak in  $Q$  at the critical point. It has been argued that this tendency of heterogeneous topologies to be particularly unsynchronizable poses a paradox given the wide prevalence of scale-free networks in nature, a problem that has been deftly got around by considering appropriate weighting schemes for the edges [33, 34] (see also ¶, and [36] for a review). However, there is no generic reason why high synchronizability should always be desirable. In fact, it has recently been shown that heterogeneity can improve the dynamical performance of model neural networks precisely because the fixed points are easily destabilised [37] (as well as conferring robustness to thermal fluctuations and improving storage capacity [19]). This makes intuitive sense, since, presumably, one would not usually want all the neurons in one's brain to be doing exactly the same thing. Therefore, this point of maximum *unsynchronizability* at the phase transition may be a particularly advantageous one.

On the whole, we find that three classes of network – homogeneous, scale-free (at the critical point) and ones composed of starlike structures, with a great many small-degree

¶ Using pacemaker nodes, scale-free networks have also been shown to emerge via rules which maximize synchrony [35].

nodes connected to a few hubs – can emerge for any kind of attachment/detachment rules. It follows that a network subject to some sort of optimising mechanism, such as Natural Selection for the case of living systems, could thus evolve towards whichever topology best suits its requirements by tuning these microscopic actions.

## 5. Correlations

Most real networks have been found to exhibit degree-degree correlations, also known as *mixing* by degree [38, 39]. They can thus be classified as *assortative*, when the degree of a typical node is positively correlated with that of its neighbours, or *disassortative*, when the correlation is negative. This property has important implications for network characteristics such as connectedness and robustness [40, 41]. A useful measure of this phenomenon is Pearson’s correlation coefficient applied to the edges [39, 41, 8]:  $r = ([k_l k'_l] - [k_l]^2)/([k_l^2] - [k_l]^2)$ , where  $k_l$  and  $k'_l$  are the degrees of each of the two nodes pertaining to edge  $l$ , and  $[\cdot] \equiv (\langle k \rangle N)^{-1} \sum_l (\cdot)$  represents an average over edges;  $|r| \leq 1$ . Writing  $\sum_l (\cdot) = \sum_{ij} \hat{a}_{ij} (\cdot)$ ,  $r$  can be expressed in terms of averages over nodes:

$$r = \frac{\langle k \rangle \langle k^2 k_{nn}(k) \rangle - \langle k^2 \rangle^2}{\langle k \rangle \langle k^3 \rangle - \langle k^2 \rangle^2}, \quad (9)$$

where  $k_{nn}(k)$  is the mean nearest-neighbour-degree function; i.e., if  $k_{nn,i} \equiv k_i^{-1} \sum_j \hat{a}_{ij} k_j$  is the mean degree of the neighbours of node  $i$ ,  $k_{nn}(k)$  is its average over all nodes such that  $k_i = k$ . Whereas most social networks are assortative ( $r > 0$ ) – due, probably, to mechanisms such as homophily [39] – almost all other networks, whether biological, technological or information-related, seem to be generically disassortative. The top right inset in figure 3 displays the stationary value of  $r$  obtained in the same networks as in the main panel and lower inset. It turns out that the heterogeneous regime is disassortative, the more so the larger  $\alpha$ . (Note that a completely homogeneous network cannot have degree-degree correlations, since all degrees are the same.) It is known that the restriction of having at most one edge per pair of nodes induces disassortativity [42, 43]. However, in our case this is not the sole origin of the correlations, as can also be seen in the same inset of figure 3, where we have plotted  $r$  for networks in which we have lifted the restriction and allowed any number of edges per pair of nodes. In fact, when multiple edges are allowed, the correlations are slightly stronger.

To understand how these correlations come about, consider a pair of nodes  $(i, j)$ , which, at a given moment, can either be occupied by an edge or unoccupied. We will call the expected times of permanence for occupied and unoccupied states  $\tau_{ij}^O$  and  $\tau_{ij}^U$ , respectively. After sufficient evolution time (so that occupancy becomes independent of the initial state<sup>+</sup>), the expected value of the corresponding element of the adjacency matrix,  $E(\hat{a}_{ij}) \equiv \hat{\epsilon}_{ij}$ , will be

$$\hat{\epsilon}_{ij} = \frac{\tau_{ij}^O}{\tau_{ij}^O + \tau_{ij}^U}.$$

<sup>+</sup> Note that this will always happen eventually since the process is ergodic.

If  $p_{ij}^+$  ( $p_{ij}^-$ ) is the probability that  $(i, j)$  will become occupied (unoccupied) given that it is unoccupied (occupied), then  $\tau_{ij}^0 \sim 1/p_{ij}^-$  and  $\tau_{ij}^u \sim 1/p_{ij}^+$ , yielding

$$\hat{\epsilon}_{ij} = \left( 1 + \frac{p_{ij}^-}{p_{ij}^+} \right)^{-1}.$$

Taking into account the probability that each node has of gaining or losing an edge, we obtain\*:  $p_{ij}^+ = u(\langle k \rangle)N^{-1}[\pi(k_i) + \pi(k_j)]$  and  $p_{ij}^- = d(\langle k \rangle)[\sigma(k_i)/k_i + \sigma(k_j)/k_j]$ . Then, assuming that the network is sparse enough that  $p_{ij}^- \gg p_{ij}^+$  (since the number of edges is much smaller than the number of pairs), and particularising for power-law local probabilities  $\pi(k) \sim k^\alpha$  and  $\sigma(k) \sim k^\beta$ , the expected occupancy of the pair is

$$\hat{\epsilon}_{ij} \simeq \frac{p_{ij}^+}{p_{ij}^-} = \frac{u(\langle k \rangle)}{d(\langle k \rangle)} \frac{\langle k^\beta \rangle}{\langle k^\alpha \rangle N} \left( \frac{k_i^\alpha + k_j^\alpha}{k_i^{\beta-1} + k_j^{\beta-1}} \right).$$

Considering the stationary state, when  $u(\langle k \rangle) = d(\langle k \rangle)$ , and for the case of random deletion of edges,  $\beta = 1$  (so that the only nonlinearity is due to  $\alpha$ ), the previous expression reduces to

$$\hat{\epsilon}_{ij} \simeq \frac{\langle k \rangle}{2\langle k^\alpha \rangle N} (k_i^\alpha + k_j^\alpha). \quad (10)$$

(Note that this matrix is not consistent term by term, since  $\sum_j \hat{\epsilon}_{ij} \neq k_i$ , although it is globally consistent:  $\sum_{ij} \hat{\epsilon}_{ij} = \langle k \rangle N$ .) The nearest-neighbour-degree function is now

$$k_{nn}(k_i) = \frac{1}{k_i} \sum_j \hat{\epsilon}_{ij} k_j = \frac{\langle k \rangle}{2\langle k^\alpha \rangle} (\langle k \rangle k_i^{\alpha-1} + \langle k^{\alpha+1} \rangle k_i^{-1})$$

(a decreasing function for any  $\alpha$ ), with the result that Pearson's coefficient becomes

$$r = \frac{1}{\langle k^\alpha \rangle} \left( \frac{\langle k \rangle^3 \langle k^{\alpha+1} \rangle - \langle k^2 \rangle^2 \langle k^\alpha \rangle}{\langle k \rangle \langle k^3 \rangle - \langle k^2 \rangle^2} \right). \quad (11)$$

More generally, one can understand the emergence of these correlations in the following way. For the network to become heterogeneous, we must have  $\pi(k) + N^{-1} \geq \sigma(k) + k/(\langle k \rangle N)$  for large enough  $k$ , so that highly connected nodes do not lose more edges than they can acquire (see section 2). This implies that  $\pi(k)$  must be increasing and approximately linear or superlinear. The expected value of the degree of a node  $i$ , chosen according to  $\pi(k_i)$ , is then  $E(k_i) = N^{-1} \sum_k \pi(k)k \gtrsim \langle k^2 \rangle / \langle k \rangle$ , while that of its new, randomly chosen neighbour,  $j$ , is only  $E(k_j) = \langle k \rangle$ . This induces disassortative correlations which can never be compensated by the breaking of edges between nodes whose expected degree values are  $N^{-1} \sum_k \sigma(k)k$  and  $\langle k^2 \rangle / \langle k \rangle$  if  $\sigma(k)$  is an increasing function. It thus ensues that a scenario such as the one analysed in this paper will never lead to assortative networks except for some cases in which  $\sigma(k)$  is a decreasing function – meaning that less connected nodes should be more likely to lose edges.

\* Again, we are ignoring corrections due to the fact that  $i$  is necessarily different from  $j$ .

Assortativity could, however, arise if there were some bias also on the node chosen to be  $i$ 's neighbour, e.g. on the postsynaptic neuron – which is precisely what happens in most social networks, where individuals do not generally choose their friends, partners, etc. randomly. Although there seem to be other reasons for the ubiquity of disassortative networks in nature [44], it is possible that the generality of the scenario studied here may also play a part.

We can use the expected value matrix  $\hat{\epsilon}$  to estimate other magnitudes. For example, the clustering coefficient, as defined by Watts and Strogatz [31], is an average over nodes of  $C_i$ , with  $C_i$  the proportion of  $i$ 's neighbours which are connected to each other; so its expected value is  $E(C_i) = \hat{\epsilon}_{jl}$  conditioned to  $j$  and  $l$  being neighbours of  $i$ 's. This means that, on average, we can make the approximation that  $k_j = k_l = \langle k_{nn} \rangle = \langle k \rangle [\langle k \rangle \langle k^{\alpha-1} \rangle + \langle k^{\alpha+1} \rangle \langle k^{-1} \rangle] / (2\langle k^\alpha \rangle)$ . Substituting this value in (10), and taking into account that one edge of  $j$ 's and one of  $l$ 's are taken up by  $i$ , we have

$$C \simeq \frac{\langle k \rangle}{\langle k^\alpha \rangle N} (\langle k_{nn} \rangle - 1)^\alpha. \quad (12)$$

For a rough estimate of the mean minimum path (the minimum path between two nodes being the smallest number of edges one has to follow to get from one to the other), we can proceed as in [45]. For a given node, let us define the number of nearest neighbours,  $z_1$ , next-nearest neighbours,  $z_2$ , and in general  $m$ th neighbours,  $z_m$ . Using the relation  $z_m = z_1 (z_2/z_1)^{m-1}$ , and assuming that the network is connected and can be obtained in  $l$  steps, this yields

$$1 + \sum_1^l z_m = N. \quad (13)$$

On average,  $z_1 = \langle k \rangle$  and  $z_2 = \langle k \rangle [(1 - C)\langle k_{nn} \rangle - 1]$  (since for each second nearest neighbour, one edge goes to the reference node and a proportion  $C$  to mutual neighbours). Now, if  $N \gg z_1$  and  $z_2 \gg z_1$ , (13) leads to

$$l \simeq 1 + \frac{\ln(N/\langle k \rangle)}{\ln[(1 - C)\langle k_{nn} \rangle - 1]}. \quad (14)$$

## 6. The *C. Elegans* neural network

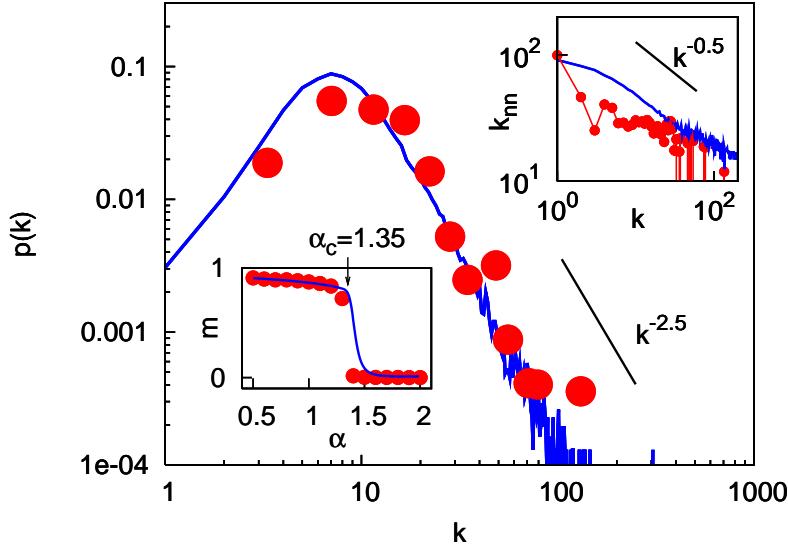
There exists a biological neural network which has been entirely mapped (although not, to the best of our knowledge, at different stages of development) – that of the much-investigated worm *C. Elegans* [46, 31]. With a view to testing whether such a network could arise via simple stochastic rules of the kind we are here considering, we ran simulations for the same number of nodes,  $N = 307$ , and (stationary) mean degree,  $\langle k \rangle = 14.0$  (in the simple, undirected representation of the network). Using the global probabilities given by (6) and local ones  $\pi(k) \sim k^\alpha$  and  $\sigma(k) \sim k$  (as in figure 3), we obtain a surprising result. Precisely at the critical point,  $\alpha = \alpha_c \simeq 1.35$ , there are some remarkable similarities between the biological network and the ones produced by the model.

Figure 4 displays the degree distributions, both for the empirical network and for the average (stationary) simulated network corresponding to the critical point, while the top inset shows the mean-nearest-neighbour degree function  $k_{nn}(k)$  for the same networks. Both  $p(k)$  and  $k_{nn}(k)$  of the simulated networks can be seen to be very similar to those measured in the biological one. Furthermore, as is displayed in table 1, the clustering coefficient obtained in simulation is almost the same as the empirical one. The mean minimum path is similar though slightly smaller in simulation, probably due to the worm’s brain having modules related to functions [47]. Finally, Pearson’s coefficient is also in fairly good agreement, although the simulated networks are actually a bit more disassortative. It should, however, be stressed that the simulation results are for averages over 100 runs, while the biological system is equivalent to a single run; given the small number of neurons, statistical fluctuations can be fairly large, so one should refrain from attributing too much importance to the precise values obtained – at least until we can average over 100 worms. Table 1 also shows the values of  $C$ ,  $l$  and  $r$  both as estimated from the theory laid out in section 5, and for the equivalent network in the *configuration model* [39] – generally taken as the null model for heterogeneous networks, where the probability of an edge existing between nodes  $i$  and  $j$  is  $k_i k_j / (\langle k \rangle N)$ . It is clear that whereas the configuration-model predictions deviate substantially from the magnitudes measured in the *C. Elegans* neural network, the growth process we are here considering accounts for them quite well. It is interesting that it should be at the

**Table 1.** Values of small-world parameters  $C$  and  $l$ , and Pearson’s correlation coefficient  $r$ , as measured in the neural network of the worm *C. Elegans* [46], and as obtained from simulations in the stationary state ( $t = 10^5$  steps) for an equivalent network at the critical point when edges are removed randomly – i.e., for  $\alpha = 1.35$  and  $\beta = 1$ .  $N = 307$ ,  $\kappa_{\text{st}} = 14.0$ ; averages over 100 runs and global probabilities as in (6). Theoretical estimates correspond to (12), (14) and (11) applied to the networks generated by the same simulations. The last column lists the respective *configuration model* values:  $C$  and  $l$  are obtained theoretically as in [39], while  $r$ , from MC simulations as in [43], is the value expected due to the absence of multiple edges. (See also figure 4.)

	Experiment	Simulation	Theory	Config.
$C$	0.28	0.28	0.23	0.15
$l$	2.46	2.19	1.86	1.96
$r$	-0.163	-0.207	-0.305	-0.101

critical point that a structural topology so similar to the empirical one emerges, since it seems that the brain’s functional topology may also be related to a critical point [49, 48].



**Figure 4.** Degree distribution (binned) of the *C. Elegans* neural network (circles) [46] and that obtained with MC simulations (line) in the stationary state ( $t = 10^5$  steps) for an equivalent network in which edges are removed randomly ( $\beta = 1$ ) at the critical point ( $\alpha = 1.35$ ).  $N = 307$ ,  $\kappa_{\text{st}} = 14.0$ , averages over 100 runs. Global probabilities as in (6). The slope is for  $k^{-5/2}$ . Top right inset: mean-neighbour-degree function  $k_{\text{nn}}(k)$  as measured in the same empirical network (circles) and as given by the same simulations (line) as in the main panel. The slope is for  $k^{-1/2}$ . Bottom left inset:  $m_{\text{st}}$  of equivalent network for a range of  $\alpha$ , both from simulations (circles) and as obtained with (1). (See also table 1.)

## 7. Discussion

With this work we have attempted, on the one hand, to extend our understanding of evolving networks so that any choice of transition probabilities dependent on local and/or global degrees can be treated analytically, thereby obtaining some model-independent results; and on the other, to illustrate how such a framework can be applied to realistic biological scenarios. For the latter, we have used two examples relating to rather different nervous systems:

- i) synaptic pruning in humans, for which the use of nonlinear global probabilities reproduces the initial increase and subsequent depletion in synaptic density in good accord with experiments – to the extent that nonmonotonic data points spanning a lifetime can be very well fitted with only two parameters; and
- ii) the structure of the *C. Elegans* neural network, for which it turns out that by only considering the numbers of nodes and edges, and imposing random deletion of edges and power-law probability of growth, the critical point leads to networks exhibiting many of the worm’s nontrivial features – such as the degree distribution, small-world parameters, and even level of disassortativity.

These examples indicate that it is not farfetched to contemplate how many

structural features of the brain or other networks – and not just the degree distributions – could arise by simple stochastic rules like the ones considered; although, undoubtedly, other ingredients such as natural modularity [47], a metric [50] or functional requirements [12] can also be expected to play a role in many instances. We hope, therefore, that the framework laid out here – in which for simplicity we have assumed the network to be undirected and to have a fixed size, although generalizations are straightforward – may prove useful for interpreting data from a variety of fields. It would be particularly interesting to try to locate and quantify the biological mechanisms assumed to be behind this kind of network dynamics.

## Acknowledgements

We are very grateful to Dante R. Chialvo and to Alex Arenas for critical readings of the manuscript and fruitful suggestions, as well as to Víctor M. Eguíluz, Miguel A. Muñoz, Pedro Garrido, Jorge F. Mejias and Sebastiano de Franciscis for useful discussions. This work was supported by Junta de Andalucía project FQM-01505 and by Spanish MEC-FEDER project FIS2009-08451.

## References

- [1] Barabási A-L and Albert R 1999 Emergence of scaling in random networks *Science* **286** 509
- [2] Albert R and Barabási A-L 2000 Topology of evolving networks: local events and universality *Phys. Rev. Lett.* **85** 5234
- [3] Bianconi G and Barabási A-L 2000 Bose-Einstein condensation in complex networks *Phys. Rev. Lett.* **86** 5632
- [4] Krapivsky P L, Redner S and Leyvraz F 2000 Connectivity of growing random networks *Phys. Rev. Lett.* **85** 4629
- [5] Bianconi G and Barabási A-L 2001 Competition and multiscaling in evolving networks *EPL* **54** 436
- [6] Park K and Lai Y-C 2005 Self-organized scale-free networks *Phys. Rev. E* **72** 026131
- [7] Ree S 2007 Generation of scale-free networks using a simple preferential-rewiring dynamics *Physica A* **376** 692698
- [8] Boccaletti S, Latora V, Moreno Y, Chavez M and Hwang D-U 2006 Complex networks: Structure and dynamics *Phys. Rep.* **424** 175–308
- [9] Johnson S, Torres J J and Marro J 2009 Nonlinear preferential rewiring in fixed-size networks as a diffusion process *Phys. Rev. E* **79** 050104(R)
- [10] Amit D J 1989 *Modeling Brain Function* (Cambridge: Cambridge Univ. Press)
- [11] Torres J J and Varona P 2010 *Modeling Biological Neural Networks* in *Handbook of Natural Computing* ed G Rozenberg *et al* (Berlin: Springer Verlag) *in press*
- [12] Sporns O, Chialvo D R, Kaiser M and Hilgetag C C 2004 Organization, development and function of complex brain networks *Trends Cogn. Sci.* **8** 418–25
- [13] Lee K S, Schottler F, Oliver M and Lynch G 1980 Brief bursts of high-frequency stimulation produce two types of structural change in rat hippocampus *J. Neurophysiol.* **44** 247
- [14] Frank E 1997 Synapse elimination: For nerves it's all or nothing *Science* **275** 324
- [15] Klintsova A Y and Greenough W T 1999 Synaptic plasticity in cortical systems *Current Opinion in Neurobiology* **9** 203

- [16] De Roo M, Klauser P, Mendez P, Poglia L and Muller D 2008 Activity-dependent PSD formation and stabilization of newly formed spines in hippocampal slice cultures *Cerebral Cortex* **18** 151
- [17] Chechik G, Meilijson I and Ruppin E 1998 Synaptic pruning in development: A computational account *Neural Comp.* **10** 1759
- [18] Chechik G, Meilijson I and Ruppin E 1999 Neuronal regulation: A mechanism for synaptic pruning during brain maturation *Neural Comp.* **11** 2061
- [19] Torres J J, Muñoz M A, Marro J and Garrido P L 2004 Influence of topology on a neural network performance *Neurocomputing* **58–60**, 229
- [20] P J Magistretti 2009 Low-cost travel in neurons *Science* **325** 1349–51
- [21] Poncela J, Gómez-Gardeñes J, Floría L M, Sánchez A and Moreno Y 2008 Complex cooperative networks from evolutionary preferential attachment *PLoS ONE* **3**(6): e2449
- [22] Vazquez F, Eguíluz V M and San Miguel M 2008 Generic absorbing transition in coevolution dynamics *Phys. Rev. Lett.* **100** 108702
- [23] Antiqueira L, Rodrigues F A and L da F Costa 2009 Modeling connectivity in terms of network activity *J. Stat. Mech.* L09005
- [24] van Kampen N G 1992 Stochastic Processes in Physics and Chemistry (Amsterdam: North-Holland)
- [25] Marro J and R Dickman 1999 Nonequilibrium Phase Transitions in Lattice Models (Cambridge: Cambridge Univ. Press)
- [26] Huttenlocher P R and Dabholkar A S 1997 Regional differences in synaptogenesis in human cerebral cortex *Journal of Comparative Neurology* **387** 167
- [27] Park J and Newman M E J 2004 Statistical mechanics of networks *Phys. Rev. E* **70** 066117
- [28] Burda Z, Jurkiewicz J and Krzywicki A 2004 Statistical mechanics of random graphs *Physica A* **344** 56–61
- [29] Derényi I, Farkas I, Palla G, Vicsek T 2004 Topological phase transitions of random networks *Physica A* **344** 583–90
- [30] Eguíluz V M, Chialvo D R, Cecchi G A, Baliki M and Apkarian A V 2005 Scale-free brain functional networks *Phys. Rev. Lett.* **94** 018102
- [31] Watts D J and Strogatz S H 1998 Collective dynamics of small-world networks *Nature* **393** 440
- [32] Barahona M and Pecora L M 2002 Synchronization in small-world systems *Phys. Rev. Lett.* **89** 054101
- [33] Motter A E, Zhou C and Kurths J 2005 Network synchronization, diffusion, and the paradox of heterogeneity *Phys. Rev. E* **71** 016116
- [34] Chavez M, Hwang D-U, Amann A, Hentschel H G E and Boccaletti S 2005 Synchronization is enhanced in weighted complex networks *Phys. Rev. Lett.* **94** 218701
- [35] Sendiña-Nadal I, Buldú J M, Leyva I and Boccaletti S 2008 Phase locking induces scale-free topologies in networks of coupled oscillators *PLoS ONE* **3**: e2644
- [36] Arenas A, Díaz-Guilera A, Kurths J, Moreno Y and Zhou C 2008 Synchronization in complex networks, *Phys. Rep.* **469** 93
- [37] Johnson S, Marro J and Torres J J 2008 Functional optimization in complex excitable networks *EPL* **83** 46006
- [38] Pastor-Satorras R, Vázquez A and Vespignani A 2001 Dynamical and correlation properties of the Internet *Phys. Rev. Lett.* **87** 258701
- [39] Newman M E J 2003 The structure and function of complex networks *SIAM Reviews* **45** 167
- [40] Newman M E J 2002 Mixing patterns in networks *Phys. Rev. Lett.* **89** 208701
- [41] Newman M E J 2003 Assortative mixing in networks *Phys. Rev. E* **67** 026126
- [42] Park J and Newman M E J 2003 The origin of degree correlations in the Internet and other networks *Phys. Rev. E* **68** 026112
- [43] Maslov S, Sneppen K and Zaliznyak A 2004 Pattern detection in complex networks: correlation profile of the Internet *Physica A* **333** 529–40
- [44] Johnson S, Torres J J, Marro J and Muñoz M A 2010 The entropic origin of disassortativity in

complex netwokrs submitted.

- [45] Albert R and Barabási A-L 2002 Statistical mechanics of complex networks *Rev. Mod. Phys.* **74**
- [46] White J G, Southgate E, Thompson J N and Brenner S 1986 The structure of the nervous system of the nematode *Caenorhabditis Elegans* *Phil. Trans. R. Soc. London* **314** 1  
Data at: <http://cdg.columbia.edu/cdg/datasets>
- [47] Arenas A, Fernández A and Gómez S 2008 A complex network approach to the determination of functional groups in the neural system of *C. Elegans* *Bio-Inspired Computing and Communication, Lect. Notes Comp. Sci.* **5151** 9-18
- [48] Chialvo D R, Balenzuela P and Fraiman D 2008 The Brain: What is critical about it?, *Collective Dynamics: Topics on Competition and Cooperation in the Biosciences, AIP* **1028** 28
- [49] Chialvo D R 2004 Critical brain networks *Physica A* **340** 756-65
- [50] Kaiser M 2004 Spatial growth of real-world networks *Phys. Rev. E* **69** 036103