Modeling Self-Developing Biological Neural Networks

Hugues Berry, Olivier Temam

Modeling Self-Developing Biological Neural Networks

Hugues Berry *, Olivier Temam

Alchemy group, INRIA Futurs, Parc Club Orsay Université, ZAC des vignes, 4 rue Jacques Monod, 91893 Orsay Cedex France

Abstract

Recent progress in chips-neuron interface suggests real biological neurons as long-term alternatives to silicon transistors. The first step to designing such computing systems is to build an abstract model of self-assembled biological neural networks, much like computer architects manipulate abstract models of transistors. In this article, we propose a model of the *structure* of biological neural networks. Our model reproduces most of the graph properties exhibited by *Caenorhabditis elegans*, including its *small-world* structure and allows generating surrogate networks with realistic biological structure, as would be needed for complex information processing/computing tasks.

Key words: Biological neural networks, Network structure, Growth model, Graph properties, Computing devices

1. Introduction

Carbon nanotubes look like a promising alternative technology to silicon chips because the manufacturing process, possibly based upon self-assembly, will be much cheaper than current CMOS processes [9]. On the other hand, these individual components may turn out to be much slower than current transistors, exhibit lots of manufacturing defects, and may be difficult to assemble into complex and irregular structures like today's custom processors. Current research are focused on building increasingly large structures of carbon nanotubes and understanding how they can be transformed into computing devices [14].
However, carbon nanotubes, though the most promising and short-term, is not the only possible alternative to silicon chips. Other emerging technologies, even if they are less familiar to chip designers, should be explored as well. In this article, we focus our attention on *biological neurons*.

* Corresponding author.
Email addresses: hugues.berry@inria.fr (Hugues Berry), olivier.temam@inria.fr (Olivier Temam).

They share some properties with carbon nanotubes: they have a low design cost, but they will provide even slower components, a significant percentage of these components will be similarly faulty, and it will be hard to assemble them into complex, irregular *pre-determined* structures. On the other hand, they have a significant asset over carbon nanotubes: we already know it is possible to self-assemble them into very large structures capable of complex information processing tasks.

While proposing computing structures based on biological neurons may seem preposterous at first sight, G. Zeck and P. Fromherz [18,56] at the Max Planck Institute for Biochemistry in Martinsried, Germany, have recently demonstrated they can interface standard silicon chips with biological neurons, pass electrical signals back and forth through one or several biological neurons, much like we intend to do with carbon nanotubes, i.e., hybrid carbon nanotubes/standard CMOS chips [21]. Moreover, based on this research work, Infineon (one of the main European chip manufacturers) has recently announced it is investigating a prototype of a chip (called "NeuroChip" that can interconnect a grid of transistors with a network of biological neurons [29], based on Fromherz's research work. So, while we will not claim this research direction should be mainstream, it is certainly worth exploring.

Now, computing machines, such as current processor architectures, are designed using a very abstract model of the physical properties of transistors and circuits. Typically, what processor architects really use (e.g., at Intel or other chip manufacturers) is how many logic gates can be traversed in a single clock cycle, and how many logic gates can be laid out on a single chip. They do not deal with the complex physics occurring at the transistor level but rely upon a very abstract and simplified model of the undergoing physical phenomena. Similarly, if we want to start thinking about computing systems built upon biological neurons, we must come up with sufficiently abstract models of biological networks of neurons that will enable the design of large systems without dealing with the individual behavior of biological neurons.

The vast literature on artificial neural networks provides little indications on the *structures* of biological neural networks [24]. To understand what kind of computing systems can be built upon biological neurons, we must first understand the kind of *structures* into which biological neurons can self-assemble. Consequently, we have turned to biology for that issue, and the current article is a joint work between computer science and biology research groups. We start with the biological neural network of *Caenorhabditis elegans*, which has been described in great details in [2,53]. Based on this work, Oshio *et al.* [40] have recently built a database which describes this biological neural network and facilitates its manipulation. Using this map as an oracle, we define a model of network growth in real space and provide empirical evidence that the characteristics of networks built upon this model and the above mentioned biological network closely match. Since this model describes the network *growth* using simple local rules, it can be used to represent much larger networks, as would be needed for computing systems. In other words, it allows the generation of surrogate networks with structures comparable to that of *C. elegans* .

There are many studies on biological neural networks, but they mostly focus on the identification of *regular* biological networks with clear structures, such as the basic circuits of the neocortex [15,16], and seldom account for the seemingly irregular structure of the vast majority of biological neural networks. We provide a *statistical* description of these apparently unstructured biological networks, that can be used as a building block for computing systems studies. Future work will focus on analyzing the evolution and learning properties of neural networks with such structures.

In Section 2, we present the biological neural network of *C. elegans* and study its properties. In Section 3, we build a network model with similar properties, provide empirical evidence that

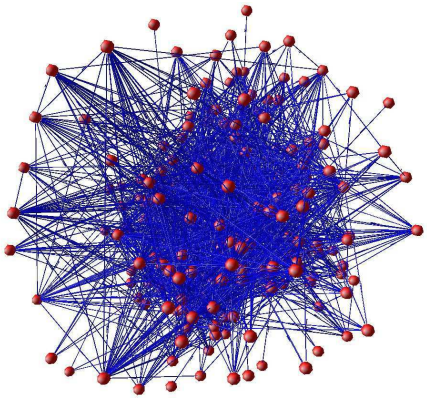

Fig. 1. Visual illustration of the neural network of *C. elegans* . Neurons are displayed as nodes and connections between them are symbolized as links. Spatial positions are arbitrary.

it closely emulates the neural network of *C. elegans* and provide a detailed comparison of the model and its biological counterpart.

2. A biological neural network

C. elegans is a millimetric worm with a simple network of 302 neurons. All the connections between its neurons have been mapped [2,53] and are believed to be relatively well conserved between individual worms. To construct a graph model of this system, we used the electronic database recently published by Oshio *et al*. [40]. A part of this system, comprising 20 neurons and referred to as the "pharyngeal system" is dedicated to control rhythmic contractions of a muscular pump that sucks food into the worm body [2]. This system is almost totally disconnected from the rest of the network. Following Morita *et al.* [37], we neglected here the pharyngeal system and only deal with the remaining 282 neurons. We then further neglected those neurons for which no connection had been described, as well as the connections to non specified cells. At the end, the network thus consisted of 265 neurons. Unlike Morita *et al.* [37], we treated each link as directed, i.e., we differentiated links from neuron i to neuron j and links from j to i; however, we collapsed multiple identical links into a single one. Furthermore, *C. elegans* neural network displays a great number of gap junctions. These are electrical synapses (as opposed to chemical ones) that provide electrically conductive links between two neurons. Contrarily to chemical synapses, these electrical couplings are bidirectional (i.e. the gap junction conductance depends on the voltage *difference* between the two neurons). Here, we treated gap junctions as pairs of links with opposite directions. Overall, we obtain 2335 unique links (or 10234 connections if we allow redundant links with the same orientation between two neurons).

Figure 1 shows a visual illustration of the corresponding neural network. A visual inspection of this figure, especially the peripheral nodes, [1] indicates that the network is rather heterogeneous: strongly connected nodes coexist with sparsely connected ones. We further tried to estimate the nature of the probability distribution of the connectivity (or graph degree), as it plays a fundamental role in characterizing the network type. The probability distribution of the connectivity in *C.*

[1] On paper, the core of the network structure is barely visible, but on a screen, it can be inspected through zooming and 3D manipulations; however the peripheral structure is the same as the core structure.

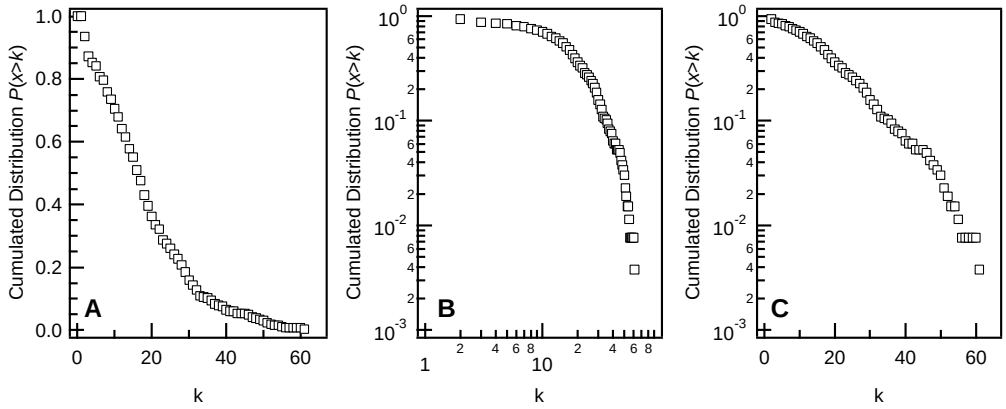

Fig. 2. Cumulated distributions of the connectivity, $P(x > k)$, where k is the total connectivity (i.e. the number of links to and from each node), for the *C. elegans* neural network shown Fig 1. The same data are presented as a linear-linear (A), a log-log (B) and log-linear plot (C).

elegans neural network has been controversial. A first study claimed the distribution was compatible with a power-law (graphs determined by power-law distributions are also called "scale-free" graphs) [5]. Not long after, this result was contradicted by an article from H.E. Stanley's team that studied outgoing and incoming connectivity separately (and ignored gap junctions) and showed that both distributions were exponential, thus excluding scale-free properties [3]. Finally, Morita *et al.* put forward correlations among incoming, outgoing and gap junctions to explain that the total degree (incoming + outgoing + gap junctions) was neither exponential nor displayed a clear power law decrease [37]. Figure 2 presents the distribution of the total connectivity for the *C. elegans* neural network displayed Fig 1. The center panel is a replot of the left one, in log-log coordinates. A power law decrease would yield a straight line in this representation, which is clearly not the case. Further, the right panel is another replot of the same data, in log-linear coordinates. Here, a straight line would indicate an exponential decrease. Clearly, the curve is closer to an exponential decay than a power-law one. However, the agreement is far from perfect. Thus, our results confirm that connectivity distribution for *C. elegans* neural network is not scale-free, but rather vaguely exponential.
We will see in the next section that several network characteristics are necessary to emulate this network structure; more importantly, we will extract simple local rules governing the network growth, enabling the development of potentially large but realistic biological networks using the same rules.

3. A model of biological neural networks

3.1. *The model*

Small-World graphs and neural networks. The global behavior of most large systems emerges from local interactions between their numerous components. At an abstract level, these systems can often be viewed as graphs, with each link representing the interaction between two components. Such graph theory approaches have proven successful in understanding the global proper-

ties of several complex systems originating from highly disparate fields, from the biological to social and technological domain. Hence the same (or similar) reasoning can be applied to analyze cell metabolism [30], the citation of scientific articles [39], software architecture [49], the Internet [5] or electronic circuits [27]. A common feature of all these networks is that their physical structure reflects their assembly and evolution, so that their global features can be understood on the basis of a small set of simple local rules that control their growth. The most common statistical structures resulting from these local rules are the so-called *small-world* and *scale-free* networks. In the broadest sense, the "small-world" phenomenon relates to sparse networks that display a low mean shortest path (or more precisely one that scales as the logarithm of the network size). However, in recent literature, the term "small-world network" is usually employed in the sense of Watts & Strogatz's model [52], where "small-world" networks are networks with both a low mean shortest path (i.e. displaying the "small-world" phenomenon) *and* a clustering index that is much higher than in comparable purely random networks (formal definitions of these network structure observables are given below). Actually, this definition encompasses non-trivial networks in which both the local (high clustering index) and global (short mean shortest path) properties are very efficient. Another important frequently encountered property in real-world networks is the scale-free structure, that is defined by a connectivity probability density function that decreases as a power law instead of an exponential decay. At a much coarser grain, graph theory methods have recently been applied to networks of cortical areas [48,17], i.e., not networks of neurons but networks of neuron *areas*, with the prospect of understanding the network functions. Since we target the characterization of networks of biological *neurons*, we study the neural network of the millimetric worm *C. elegans* at the level of individual *neurons*, and attempt to derive a network growth model that closely emulates it.

Most complex (sparse) networks can be categorized into four families [6]: purely random networks, regular networks, small-world networks and scale-free networks (note however that these categories are not exclusive, as many scale-free networks are also small-world ones). In *purely random networks* (also known as Erdős-Rényi graphs), two nodes i and j are connected with a predefined probability independently of all others. These graphs are characterized by short paths between two nodes (denoted λ) and a low clustering (denoted $\langle C \rangle$). On the opposite, *regular* graphs (where each node has the same connectivity) are characterized by a high clustering and usually display a large average shortest path. Between these two extremes, *small-world* networks (in the sense of [52]) display both small average shortest paths and a high degree of clustering. For most small-world networks, $P(k)$, i.e., the probability density function of the connectivity k, decreases very quickly (exponentially) beyond the most probable value of k, which thus sets the connectivity scale. However, in some graphs (such as the Internet), $P(k)$ decreases as a power-law of k ($P(k) \propto k^{-\gamma}$), i.e., in a much slower way [42]. In this case, nodes with a very high connectivity (hubs) can also be present with a significant probability so that the connectivity does not display a clear scale, hence the term "scale-free" networks.

We now formally introduce the parameters of a network model. Besides the number of nodes N and number of links K, the structural characteristics of complex networks are mainly quantified by their link density ρ, average connectivity $\langle k \rangle$, connectivity distribution $P(k)$, average shortest path λ and average clustering coefficient $\langle C \rangle$ [47]. The network density ρ is the density of links out of the $N(N-1)$ possible directed links [2] (recall multiple links between two nodes are considered a unique link and self-connections are forbidden)

[2] Each node can have at most $N-1$ outgoing links, so the maximum number of links is $N(N-1)$.

$$\rho = K/\left(N^2 - N\right) \tag{1}$$

The connectivity (or *degree*) k_i of node number i is the number of links coming from or directed to node i. $P(k)$ is the probability density function of the k_i's and $\langle k \rangle$ their average over all the nodes in the network. Let $d(i,j)$ be the shortest path (in number of neurons) between neuron i and j, then λ is its average over the network

$$\lambda = 1/\left(N^2 - N\right) \sum_{i,j} d(i,j) \tag{2}$$

The clustering coefficient of a node i with k_i (incoming plus outgoing) connections is defined by

$$C_i = E_i/\left(k_i^2 - k_i\right) \tag{3}$$

where E_i is the number of connections among the k_i neighbors of node i, excluding the connections between a neighbor and node i itself. The average clustering coefficient $\langle C \rangle$ is the average of the C_i's over all nodes and expresses the probability that two nodes connected to a third one are also connected together (degree of cliquishness).

We also quantify the average level of asymmetry between incoming and outgoing links. To this aim, we calculate for each node an asymmetry index

$$\alpha_i = \frac{|k_i^{out} - k_i^{in}|}{k_i} \tag{4}$$

where k_i^{out} and k_i^{in} are, respectively, the number of links leading out of (out-degree) or into (in-degree) node i (i.e. $k_i^{out} + k_i^{in} = k_i$). The average value of the α_is over the network, $\langle \alpha \rangle$, expresses the tendency of the nodes to have unbalanced out-degree and in-degree values. Its value is 0 when all nodes in the network have as many incoming links as outgoing ones, and 1 when nodes have exclusively incoming or outgoing links.

Further information about the network structure can be obtained by inspecting the k-dependence of $C(k)$, the average clustering coefficient restricted to nodes of connectivity k

$$C(k) = \frac{\sum_i \delta_{k_i k} C_i}{\sum_i \delta_{k_i k}} \tag{5}$$

with Kronecker's Delta

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

In many real-world networks, such as actor networks or the World Wide Web, $C(k)$ decreases as a power law of k [46], indicating that high degree nodes are more likely to have a poorly interconnected neighborhood. This property is a strong indicator of the hierarchical organization of these scale-free networks.

Finally, we quantified the correlation among the connectivities of two connected nodes by the conditional probability $P(k'|k)$ that a link with connectivity k is linked to a node with connectivity k' [41]. To this aim, we compute for each node i the average connectivity $k_{nn,i}$ of i's neighbors [7] (a neighbor being a node having at least one link leading out of or into i)

$$k_{nn,i} = \frac{1}{n_i} \sum_i a_{ij} k_j \tag{6}$$

where n_i is the number of i's neighbors and $a_{ij} = 1$ if j is a neighbor of i, 0 else. The average value of k_{nn} restricted to nodes of connectivity k is then

Network	ρ	$\langle k \rangle$	λ	$\langle C \rangle$	$\langle \alpha \rangle$
C. elegans	0.033	17.62	3.19	0.173	0.358
random	0.033	17.62	2.79	0.0334	0.192
model	0.033	17.58	3.23	0.181	0.421

Table 1
Structural characteristics of the neural network of *C. elegans* shown in Figure 1, a comparable random (Erdős-Rényi) network and the network obtained with the proposed growth model. Data for the random and model networks are averages over 100 network realizations. See text for definition of the listed properties. The parameters for the model network are $L_x = L_y = 15$, $L_z = 300$ (neuron size units), $P_{new} = 0.00130$, $\xi = 10$, $N = 265$

$$k_{nn}(k) = \frac{\sum_i \delta_{k_i k} k_{nn,i}}{\sum_i \delta_{k_i k}} \tag{7}$$

The evolution of $k_{nn}(k)$ with k is an important indicator of mixing properties in the network (mixing by node degree). If k_{nn} increases with k, highly connected nodes are more likely to be connected to highly connected nodes. This property, known as "assortative mixing", is often found in social networks, for example [7], while the opposite ("disassortative mixing") is a property of the Internet, for instance [41].

The main structural characteristics of the *C. elegans* neural network are indicated in Table 1. Compared to a random network with the same density, this neural network has a similar average shortest path but the clustering has increased almost fivefold. This means that, in the *C. elegans* neural network, one neuron can reach any other neuron in only three connections on average. This is a clear sign of small-world properties. Considering the network is treated here as a directed graph, these results are coherent with previously published estimates [52,13]. Another characteristic of *C. elegans* network, that can be seen from Table 1, is the slight asymmetry of its node connectivity. Indeed, compared to a random network, the value of the asymmetry index $\langle \alpha \rangle$ for the network of *C. elegans* is almost twice its value in a comparable random network. Hence, the neurons of *C. elegans* are more likely to present unbalanced in- and out-degrees than if the connection directions were random.

In biological neural networks, distance matters. Most network growth models do not consider the physical *distance* between two nodes (neurons) [34]. For instance, most scale-free networks are obtained through a preferential attachment rule which postulates that new nodes are linked to the already most connected nodes [5]. Not only this development rule implies some global control mechanism (i.e., a node must somehow know which are the most connected nodes) which seems unlikely in the case of a neuronal system, but it also implies that long connections are just as likely as shorter ones. Similar arguments can be opposed to the Watts-Strogatz rewiring algorithm that generates small-world networks through addition of long-range connections to a pre-existing regular circular network [52]. An improved variation of the Watts-Strogatz algorithm restricts rewiring to a local spatial neighborhood around each node [13] thus implicitly introducing the distance factor. However, these last two models are not network *growth* models and it is unlikely that they can be modify to yield growth models. They thus do not provide a biologic realistic metaphor. In opposition to these models, we address in the present work the specific case of biological neuron network *growth* in real three dimensional space.

Several observations support the key notion of physical distance. Long distance connections are expensive in biological neural networks because they imply large volumes of metabolically active tissue to be maintained and long transmission delays [12]. Such long-distance connections are

thus much less likely than short-distance ones, just like long-distance links between two nodes in Internet or in airport transportation systems are more unlikely than short ones. Indeed, connecting two nodes on the Internet and maintaining the resulting physical line comes together with a cost proportional to the line's length, which favors shorter links [55]. Likewise, fuel cost, flight length limitations or geo-political reasons [23] are possible explanations for the increased probability of small-distance connections in airplane networks [19]. Moreover, the total wiring length cost seems to be a crucial factor of cortical circuit development [10,11]. The network structure itself depends on the wiring length. For instance, small-world properties have been shown to emerge naturally upon minimization of the Euclidean distance between nodes [35]. Furthermore, Kaiser et al. [33,32] have recently shown that the network structure during growth in a metric space is influenced by neuron density (number of neurons per unit volume) when growth occurs in a spatially constrained domain.

A network growth model for *C. elegans* **.** We now propose a network growth model in a three-dimensional space. Neurons are abstracted as cubical volumes of unit size. The position of each neuron on the cubic lattice is defined by the integer coordinates (i, j, k) of its center of mass and spans over the volume comprised between $(i - 1/2, j - 1/2, k - 1/2)$ and $(i + 1/2, j + 1/2, k + 1/2)$. The lattice dimensions are L_x, L_y and L_z, defining a volume of $L_x \times L_y \times L_z$ unit sizes. We start by placing a unique neuron at the center of the three-dimensional lattice. Each step of the growth algorithm then consists of six elementary substeps:

(i) choose a neuron n at random among the neurons already connected in the network (*origin* neuron). Let (i, j, k) be the spatial coordinates of n on the lattice.

(ii) Then choose a *destination* site (i', j', k') at distance d with probability

$$P(d) = 1/\xi \exp(-d/\xi) \tag{8}$$

where d is the Euclidean distance between (i, j, k) and (i', j', k'), and ξ is a parameter that sets the rate of $P(d)$ decay with d. If the chosen destination site is located outside the lattice borders, go back to substep 1. Thanks to the exponential distance distribution, the probability to create a connection of a given length (at a certain distance) decreases rapidly with the length, which accounts for the high cost of long wires. Note that, in biological neural networks, new connections are established through cell outgrowths (neurites) from existing neurons; these outgrowths are guided by gradients of chemical concentration which similarly decay rapidly with distance. Exponentially decaying-distance probability is classically used in the modeling of networks for which long-distance wirings are very costly [55,23]. Furthermore, recent evaluations of the probability distribution of the internode length in the *C. elegans* network have clearly evidenced such an exponential decay (albeit with several characteristic length scales) [1]. The $1/\xi$ prefactor is added as a normalization term so that the distribution function is unity for infinite distances. Note however that in practice, because distances greater than the lattice volume are rejected, the possible distances are indeed constrained by the lattice size.

(iii) If a neuron n' of the network already exists at the destination site (i', j', k'), a connection is created between n and n'.

(iv) If there is no neuron at the destination site, a new neuron n' is placed at the destination site and a connection is created between n and n' with probability P_{new}; the value of P_{new} is discussed below.

(v) If a connection has been created during one of the two preceding steps, its direction ($n \to n'$ or $n' \to n$) is chosen with probability $P_{n \to n'} = 1 - P_{n' \to n} = k_n^{out} / \left(k_n^{out} + k_n^{in}\right)$. This process accounts for the property that, *C. elegans* neurons exhibit slightly unbalanced in-

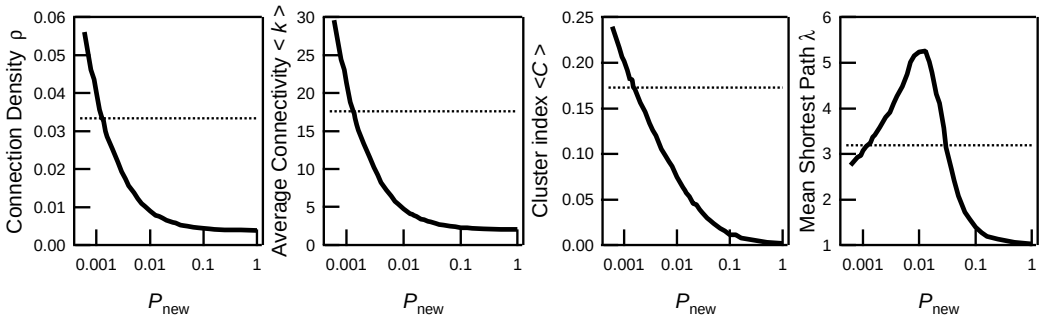

Fig. 3. Structural characteristics of the networks obtained with the proposed model as a function of the probability P_{new} that a new neuron connects to the network. Values are averages over 100 realizations. The dotted line indicates the corresponding value found for the network of *C. elegans* . Other parameters were $L_x = L_y = 15$, $L_z = 300$ (neuron size units), $\xi = 10$ and $N = 265$

and out-degrees (see above).

(vi) go back to substep 1.

This algorithm iterates until the network contains a prescribed number of neurons N. Because we ultimately want to compare the model results with *C. elegans* we set $N = 265$ in this study. For most of the presented results, the size of the lattice volume in which the network growth is restricted was determined from *C. elegans* body dimensions. The body length of the adult varies between 1.00 [4] and 1.30 mm [36], and its body volume is close to 4.0 nL [25]. Assuming the body is a $L_z = 1.20\,\mathrm{mm}$-long square cylinder hence yields $L_x = L_y \approx 60\,\mu\mathrm{m}$. The diameter of a neuron cell body (soma) varies according to the estimations between 3.8 [22] and $5\,\mu\mathrm{m}$ [54]. Using a $4\,\mu$m-diameter, we obtain $L_x = L_y = 15$ and $L_z = 300$ neuron (soma) size units.

3.2. *Results*

New neurons are unlikely to be created in already cluttered areas. The probability P_{new} that a new neuron is integrated in the preexisting network is the most important parameter in the model. As seen in Figure 3, connection density, average connectivity and clustering index increase as P_{new} decreases while the evolution of the average shortest path λ is biphasic, with a maximum at $P_{new} \approx 0.01$. Thus, decreasing P_{new} below 0.01 yields networks with increasingly strong small-world properties together with increasingly high average connectivity.

Interestingly, the results of Figure 3 show that the studied structural properties of the networks obtained with our algorithm match that of *C. elegans* neural network for $P_{new} \approx 1.3 \times 10^{-3}$. The main structural characteristics of the networks obtained using this algorithm with $P_{new} = 1.30 \times 10^{-3}$ are listed in Table 1, and can be compared to the values obtained for *C. elegans* . Clearly, the values obtained with the model are in very good agreement with those observed in the real network. Even the node asymmetry of the model networks is fairly close (within 15%) to that observed in *C. elegans* . Further comparisons between the real network and our model can be found in Figure 4. The connectivity distribution of the real network (Fig. 4A) is clearly well approximated by the model with $P_{new} = 1.30 \times 10^{-3}$. The clustering index for nodes of connectivity k, $C(k)$ is displayed in Fig. 4B. For the *C. elegans* network, albeit strong fluctuations

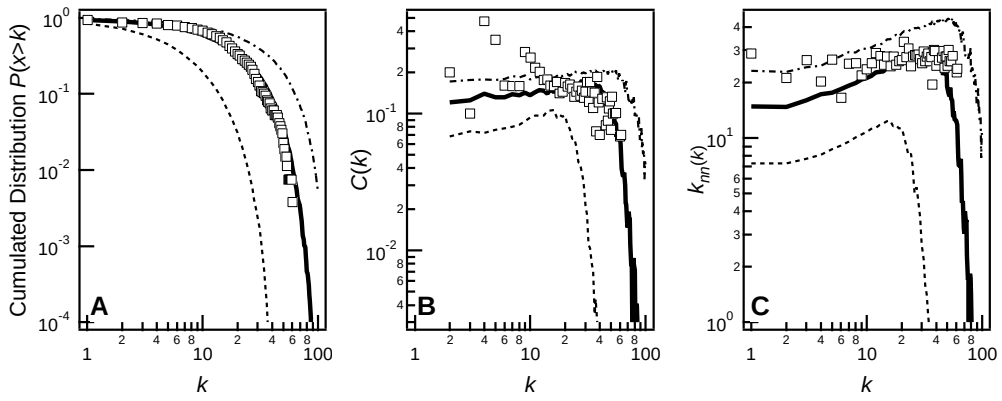

Fig. 4. Comparisons of the structure of *C. elegans* network (white squares) with those obtained using the proposed algorithm with the probability of newcomer insertion $P_{new} = 0.00130$ (thick full lines). For comparison, the results obtained with the model using $P_{new} = 0.00600$ (dashed line) or 0.00060 (dashed-dotted line) are also displayed. (A) Cumulated probability distribution of the connectivity k; (B) k-dependence of the average clustering coefficient restricted to nodes of connectivity k, $C(k)$ (Eq.5); (C) k-dependence of the average neighbor connectivity of nodes with degree k, $k_{nn}(k)$ (Eq.7). Other parameters for the model results were $L_x = L_y = 15$, $L_z = 300$ (neuron size units), $\xi = 10$ and $N = 265$. Results for the model are averages over 100 realizations.

are observed at low k values ($k < 10$), $C(k)$ is largely constant up to $k \approx 40$ and displays a rapid decay after this cutoff value. Hence, in this network, the clustering index of most nodes does not strongly depend on their connectivity with the exception of highly connected ones that tend to have slightly less interconnected neighborhoods. All these properties are satisfactorily captured by our model, including a close to independent behavior for $k < 40$, and a cutoff decay for highly connected nodes. Finally Fig. 4C shows the average connectivity $k_{nn}(k)$ of the neighbors of a node as a function of the node connectivity k. A clear-cut behavior of *C. elegans* network is the independence of $k_{nn}(k)$ with respect to k. This indicates a perfect mixing by node degree: the network is neither assortative nor disassortative. The behavior of our model related to this point is partly satisfactory. The evolution of $k_{nn}(k)$ for the real network is globally well rendered by the model (at least for $k > 10$), albeit the behavior of the model for the less connected nodes is more assortative than observed in *C. elegans* . However, it must be noted that, even for the low values of k, the increase in the model results remains moderate (from $k = 1$ to $k = 10$, $k_{nn}(k)$ increases from ≈ 15 to ≈ 21.)

Taken together, these results indicate that the present model (with $P_{new} = 1.30 \times 10^{-3}$ and the other parameters given in Fig.4) yields networks that are statistically very similar to that of *C. elegans* . It thus provides a simple way to generate network surrogates for the structure of *C. elegans* neural networks and can be used to study the dynamical properties of these kinds of structures (at least for those properties that do not critically depend on assortativity at low connectivities). A crucial parameter for the model is the probability P_{new} that a neuron of the network uncovers a newcomer and connects it. Interestingly, realistic structural values are obtained when this probability is rather low ($P_{new} \approx 1.3 \times 10^{-3}$). A biological interpretation of this value is that natural neural networks would be very reluctant to admit new neurons in the network (as only 1 trials out of ≈ 800 would be statistically successful). In fact, this probability (which is actually a probability by time step, i.e. a probability rate) encompasses both the rate

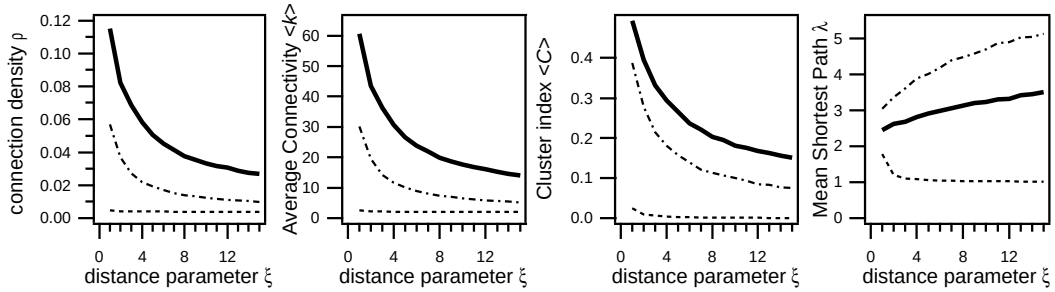

Fig. 5. Structural characteristics of the networks obtained with the proposed model as a function of the relative distance parameter ξ (see Eq. 8). The value of newcomer neuron integration probability P_{new} was 0.0013 (full thick line), 0.0060 (dashed-dotted line) or 1.00 (dashed line). Results are averages over 100 network realizations. Other parameters were $L_x = L_y = 15$, $L_z = 300$ (neuron size units) and $N = 265$.

at which neurons belonging to the network discover newcomers that are not already part of it, and the rate at which, once uncovered, these newcomer neurons are incorporated in the network. Interestingly, recent results in neurobiology suggests that this rate is indeed low in real biological neural networks. For instance, the time necessary for neurons to develop from isolated cells to a fully connected mature network *in vitro* is very long, of the order of a month [50]. Likewise, in adult mice hippocampus, newborn neuron incorporation is a complex process that needs more than a month before newcomers are fully incorporated in the network and reach a mature morphology [51,20]. Finally, it has recently been suggested that the lack of neural turnover and/or replacement of injured neurons in several parts of the adult brain is not due to the absence of potentially competent cell, but, more probably, to a strong reluctance of the neurons to accept newcomers into an already established neural network [45]. In light of these findings, our results suggest that this strong reluctance could be an important factor for the structure of the networks. For instance a somewhat trivial effect concerns the high average connectivity observed in biological neural networks: if new neurons can hardly emerge in already cluttered areas, connections are mostly drawn among existing neurons, hence the connectivity increases.
In the rest of the paper, we turn to the study of the model itself and show how the structure of the obtained networks depends on the model parameters.

Influence of neuron density. We first present results obtained when the local neuron density changes. To this aim, we varied the parameter ξ of the connection distance probability, Eq.8. When ξ decreases, the probability that two neurons are connected by long-distance connections decreases, so that local neuron density increases. Figure 5 presents the results obtained for various ξ values and three values for P_{new} (the other parameters are set to the same values as Fig. 4). Setting $P_{new} = 1$, i.e. a new neuron is certain to be created in a currently empty location, our algorithm is closely related to the two-dimensional model recently proposed by Kaiser and Hilgetag [33,32]. In their paper, the obtained networks progressively acquire small-world properties when neuron density increases. In our case, with $P_{new} = 1$, the obtained networks are mainly insensible to local density and remain random networks with very few connections. When P_{new} decreases, however, increasing local neuron densities has a marked effect on the networks. On the one hand, connection density (and average connectivity) increases with increasing local density. On the other hand, this increase in connectivity comes together with increased clustering and decreasing mean-shortest path. In other words, when the probability P_{new} is low, increasing

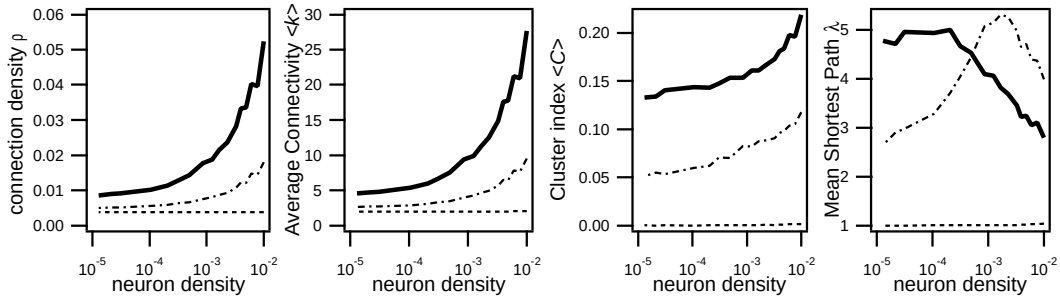

Fig. 6. Structural characteristics of the networks obtained with the proposed model as a function of the global average neuron density $\frac{N}{L_x \times L_y \times L_z}$ with $P_{new} = 0.0013$ (full thick line), 0.0060 (dashed-dotted line) or 1.00 (dashed line). The different density values are obtained through variations of the domain volume ($L_x \times L_y \times L_z$), preserving the same aspect ratio as in previous figures (i.e. $L_x/L_z = 20$) and the square section (i.e. $L_x = L_y$). The results are average values obtained over 100 network realizations. Other parameters were $\xi = 10$, $N = 265$

local neuron density yields networks with small-world properties. Note however that the "intensity" of the network small-world property slightly fades when local neuron density increases (i.e. ξ decreases). Indeed, the relative values of the clustering index and the mean shortest path (i.e. relative to comparable random networks) changes monotonously from 6.0 to 4.5, and from 1.15 to 1.30, respectively, when ξ is decreased from 12 to 1.

Another approach to increase neuron density consists in increasing the global (average) density. To this aim, we varied the volume of the lattice domain in which the network growth is restricted, while keeping constant its aspect ratio (L_z/L_x) and preserving the square section ($L_x = L_y$). We then define the (global) neuron density as the average number of neurons by available volume, $\frac{N}{(L_x \times L_y \times L_z)}$. Taken together, the corresponding results (Fig. 6) are similar to those obtained through increase of the local neuron density, except that the mean shortest path may present saturating or biphasic behaviors. But here again, the neuron density has no influence on the networks obtained with $P_{new} = 1$, that remains extremely sparsely connected.

Shape might also matter. The overall shape of the domain can be an important determinant of local neuron density as well. In the following, we present results obtained keeping the lattice volume constant while varying one of the domain length ("baguette-like" lattices with square sections). Figure 7 presents the evolution of the network properties when the aspect ratio (or length-to-width ratio, L_z/L_x) is varied. The overall volume of the lattice domain is kept constant to within $< 0.5\%$ of its value in Figures 3- 5 (i.e. $15 \times 15 \times 300$ in neuron soma size units). When $P_{new} = 1$, the structural characteristics do once again not depend on the aspect ratio. However, with lower P_{new} values, all quantities (except λ) tend to increase as the domain anisotropy increases. Most notably, for low P_{new} values (such as 0.0013), the average connectivity in highly anisotropic domains is almost sevenfold its value when the network is grown in an isotropic cubic domain.

Hence, the decrease of the lattice cross section acts in synergy with low P_{new} values to further increase local neuron density. Figure 8 illustrates this effect with two examples of model networks grown in a moderately anisotropic domain ($L_Z/L_x = 20$). With $P_{new} = 1$ (Fig. 8, Left), the network is sparsely connected and poorly clustered. In strong contrast, the network obtained

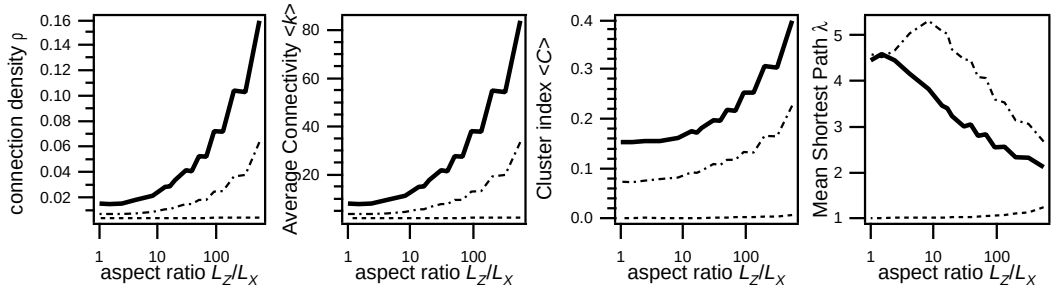

Fig. 7. Structural characteristics of the networks obtained in "baguette-like" space domains with square sections ($L_x = L_y$). The length-to-width ratio L_z/L_x is varied at constant domain volume (i.e. constant to within $< 0.5\%$ of $15 \times 15 \times 300$). The probability to connect to a new neuron $P_{new} = 0.0013$ (full thick line), 0.0060 (dashed-dotted line) or 1.00 (dashed line). Shown are averages over 100 realizations of the network. Other parameters are $L_x = L_y$ (square section), $\xi = 10$, $N = 265$.

with $P_{new} = 0.00130$, (Fig. 8, Right) is densely connected and highly clustered. Occasionally, as in the network displayed Fig. 8 (right), remote neurons can form a local group of clustered neurons, far from the main core area. These structures remind of the organization of the neurons into ganglions in *C. elegans* neural network. These results suggest that the anisotropy of the space domain inside which network growth is restricted, could be an important factor to determine network connectivity and clustering.

This result could be of importance because many biological neural networks are restricted to grow into highly anisotropic domains. This is of course the case of the worm *C. elegans* . But these conclusions could also be of interest for the organization of neocortical minicolumns, that are chiefly cylinders of $30-50$ μm diameter and $3-6$ mm length and contain of the order of 100 neurons [38,31]. The neurons inside a given column are generated during the middle of gestation, some distance away from the cortex, in the same underlying ventricle zone [43]. They then migrate from there to their cortical area following radial bundles of glial fascicles. Because of this scaffolding by the glial cells, the migrating neurons are spatially confined to a radially elongated domain all along the fetal cerebral wall. This spatial restriction appears to be a major determinant of the columnar cytoarchitecture in the neocortical minicolumn [44]. Furthermore, the formation of these minicolumns seems an intrinsically kinetic phenomenon, as neurons destined to deeper cortical layers are generated earliest and are bypassed by those of the more superficial layers, arriving later in the developing column [43,44]. Hence, minicolumn formation is a network growth process whose structure is mainly imposed by a highly anisotropic restricting space domain. We thus think that our model, endowed with adequate modifications (e.g. to account for laminar organization) could be used to study network structuration in such cases.

4. Discussion

A major characteristic of our model is that it is spatially embedded, i.e. the neurons explicitly live in a three-dimensional Euclidean space. In spite of the possible importance of this physical embedding, most of the works about the topological properties of complex networks have ignored it. In most studies, the only studied distance is the graph distance, i.e. the path length (in number of links) between two nodes. However, several recent works have been devoted to the study

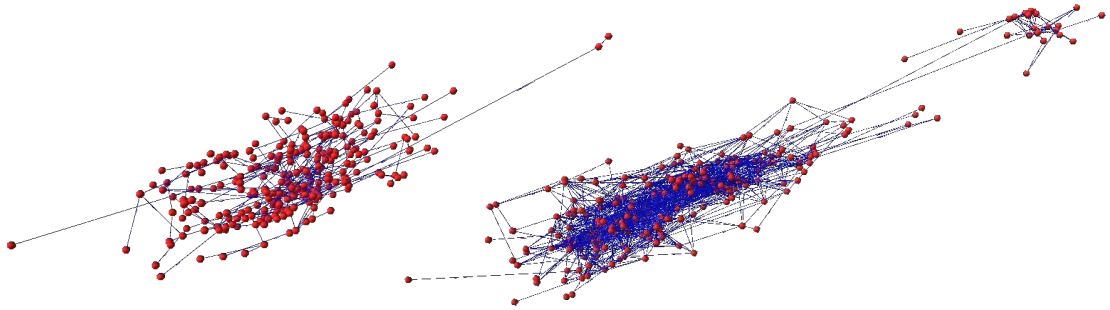

Fig. 8. Two examples of model networks grown in moderately anisotropic domains ($L_z/L_x = 20$) and $P_{new} = 1.00$ (Left) or $P_{new} = 0.00130$ (Right). Other parameters are $L_x = L_y = 15$ (square section), $\xi = 10$ and $N = 265$. Note that the neuron radius is not to scale.

of networks in Euclidean physical space (reviewed in [8], for instance). To account for the tendency of real-world networks to present more short-range than large-distance connections, some of these works have studied the structures generated when the average physical distance between the nodes is minimized. This process is known to yield a rich variety of structures, including small-world and scale-free ones, when the distance to be minimized is the graph distance [28]. Some models have compared the networks obtained when the optimization pressure is put on the graph distance or on the Euclidean distance or intermediate situations between both [35,19]. In a first report, Mathias and Gopal studied a network of nodes arranged on a two dimensional ring. When only the graph distance is taken into account, the obtained networks are random star-like ones with long-distance links and a few hubs, while optimizing the Euclidean distance between nodes only rather leads to regular networks. Small-world networks are found in the intermediate regime, where the optimization pressure is partly directed towards Euclidean distance and partly towards graph distance minimization [35]. More recently, Gastner and Newman obtained comparable results when the nodes are randomly positioned on a square [19]. Furthermore, similar theoretical considerations about syntactical networks in sentences has led to the hypothesis that part of language organization can be attributed to minimization (or at least constraining) of the Euclidean (physical) distance between linked words [26].

In contrast with these models, our model does not necessitate to minimize the global distance between nodes. The small-world structure emerges as a result of the interplay between the low probability of newcomer incorporation (P_{new}) and the rapid decay of the connection probability when the Euclidean distances between two neurons increases. This property can actually also be found in Kaiser and Hilgetag's model [33,32]. This is a spatial network growth model in which, at each time step, a newcomer tries to connect to each neuron of the pre-existing network with a probability that decays exponentially fast with the Euclidean distance between the two neurons. This behavior is reminiscent of the famous "preferential attachment" model of Barabási and Albert for generating scale-free networks, where newcomer nodes attach to preexisting ones with a probability proportional to the connectivity of the destination node [5]. Varying the model parameters, Kaiser and Hilgetag's model yields different structures, including scale-free and small-world networks. However, the parameter range yielding small-world networks in this model is extremely narrow. Furthermore the obtained clustering indices and connectivities are still far

from those observed in *C. elegans* . The principal difference between the present model and the model proposed by Kaiser & Hilgetag is that in our case, the connection probability between neurons that are already parts of the network remains high, at least for densely populated areas. Hence, once incorporated, a newcomer neuron can extensively continue to increase the number of its connections with the rest of the network. This allows to obtain much higher connectivity and clustering.

As shown by the results presented above, using appropriate parameters, our model yields networks with structures that are statistically very similar to that of *C. elegans* . It thus provides a simple way to generate network surrogates for the structure of *C. elegans* neural networks and can be used to study the properties of neuron dynamics on these network structures. Hence, a first direction for future work will consist in the study of structure-dynamics relationships in these kinds of structures. Most noticeably, as the model defines the network *growth* properties, it can be used to generate large-size neural networks, that will be needed for solving computing tasks.

A further step will consist in improving the model accuracy/realism by integrating known but abstract characteristics of the behavior of individual neurons. Through this combined model, we will investigate the application of such biological neural networks to computing tasks, assuming the experimental setups described in [18]. In this perspective, our aim is to obtain a sufficiently abstract model of biological networks of neurons that will enable the design of large systems without dealing with the individual behavior of biological neurons. We think that the availability of such abstract models will be a crucial chokepoint that will have to be overcome if we want to build computing systems using real biological neurons.

5. Acknowledgement

This research was supported by the French National Research Agency (ANR, project number "JC05_63935").

References

[1] Y.Y. Ahn, B. J. Kim, and H.G. Jeong. Wiring cost in the organization of a biological network, 2005. arXiv.org:q-bio/0505009.

[2] D.G. Albertson and J.N. Thomson. The pharynx of Caenorhabditis elegans. Phil. Trans. R. Soc. London B, 275:229–325, 1976.

[3] L.A.N. Amaral, A. Scala, M. Barthélémy, and H.E. Stanley. Classes of small-world networks. Proc. Natl. Acad. Sci. USA, 10:11149–11152, 2000.

[4] L. Avery and B.B. Shtonda. Food transport in the *C. elegans* pharynx. J. Exp. Biol., 206:2441–2457, 2003.

[5] A.L. Barabási and R. Albert. Emergence of scaling in random networks. Science, 286:509–512, 1999.

[6] A.L. Barabási and Z.N. Oltvai. Network biology: understanding the cell's functional organization. Nature Rev. Genetics, 5:101–113, 2004.

[7] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani. The architecture of complex weighted networks. Proc. Natl. Acad. Sci. USA, 101:3747–3752, 2004.

[8] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang. Complex networks: structure and dynamics. Phys. Rep., 424:175–308, 2006.

[9] Y. Chen, D. A. A. Ohlberg, X. Li, D. R. Stewart, R. S. Williams, J. O. Jeppesen, K. A. Nielsen, J. F. Stoddart, D. L. Olynick, and E. Anderson. Nanoscale molecular-switch devices fabricated by imprint lithography. Appl. Phys. Lett., 82:1610–1612, 2003.

[10] C. Cherniak. Component placement optimization in the brain. J. Neurosci., 14:2418–2427, 1994.

[11] C. Cherniak. Neural component placement. Trends Neurosc., 18:522–527, 1995.

[12] D.B. Chklovskii, T. Schikorski, and C.F. Stevens. Wiring optimization in cortical circuits. Neuron, 34:341–347, 2002.

[13] J. W. Clark and A. T. Eggebrecht. The small world of the nobel nematode Caenorhabditis elegans. In J. da Providencia and F. B. Malik, editors, Condensed Matter Theories, volume 18. Nova Science Publishers, Hauppauge, N.Y., 2003. *to appear*.

[14] A. Dehon. Nanowire-based programmable architectures. J. Emerg. Technol. Comput. Syst., 1:109–162, 2005.

[15] R. Douglas and K. Martin. Neocortex. In G. M. Shepherd, editor, The synaptic organization of the brain, pages 459–509. Oxford University Press, Oxford, 1998.

[16] R.J. Douglas and K.A.C. Martin. Neuronal circuits of the neocortex. Annu. Rev. Neurosci., 27:419–451, 2004.

[17] V.M. Eguiluz, D.R. Chialvo, G. Cecchi, M. Baliki, and A.V. Apkarian. Scale-free brain functional networks. Phys. Rev. Lett., 94:018102, 2005.

[18] P. Fromherz. Neuroelectronic interfacing: Semiconductor chips with ion channels, nerve cells, and brain. In Rainer Waser, editor, Nanoelectronics and informaiton technology, pages 781–810. Wiley–VCH, Berlin, 2003.

[19] M.T. Gastner and M.E.J. Newman. The spatial structure of networks. Eur. Phys. J. B, 49:2471–252, 2006.

[20] S. Ge, E.L.K. Goh, Y. Kitabatake, G. Ming, and H. Song. Gaba regulates synaptic integration of newly generated neurons in the adult brain. Nature, 439:589–593, 2006.

[21] S. C. Goldstein and M. Budiu. NanoFabrics: Spatial computing using molecular electronics. In Proceedings of the 28th International Symposium on Computer Architecture 2001, 2001.

[22] M.B. Goodman, D.H. Hall, L. Avery, and S.R. Lockery. Active currents regulate sensitivity and dynamic range in *C. elegans* neurons. Neuron, 20:763–772, 1998.

[23] R. Guimerà and L.A.N. Amaral. Modeling the world-wide airport network. Eur. Phys. J. B, 38:381–385, 2004.

[24] S. Haykin. Neural Networks. Prentice Hall Intl, London, UK, 2nd edition, 1999.

[25] T. Hirose, Y. Nakano, Y. Nagamatsu, T. Misumi, H. Ohta, and Y. Ohshima. Cyclic GMP-dependent protein kinase EGL-4 controls body size and lifespan in *C. elegans*. Development, 130:1089–1099, 2003.

[26] R. Ferrer i Cancho. Euclidean distance between syntactically linked words. Phys. Rev. E, 70:056135, 2004.

[27] R. Ferrer i Cancho, C. Janssen, and R.V. Solé. The topology of technology graphs: Small world patterns in electronic circuits. Phys. Rev. E., 64:32767, 2001.

[28] R. Ferrer i Cancho and R.V. Solé. Optimization in complex networks. In Statistical physics of complex networks, Lecture notes in physics, pages 114–125. Springer, Berlin, Germany, 2003.

[29] Infineon. Neurochip with integrated electronics in research phase. http://www.infineon.com/cgi/ecrm.dll/jsp/showfrontend.do? lang=EN&channel_oid=-11398, 2003.

[30] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, and A.L. Barabási. The large-scale organization of metabolic networks. Nature, 407:651–654, 2000.

[31] E.G. Jones. Microcolumns in the cerebral cortex. Proc. Natl. Acad. Sci. USA, 97:5019–5021, 2000.

[32] M. Kaiser and C.C. Hilgetag. Modelling the development of cortical systems networks. Neurocomputing, 58–60:297–302, 2004.

[33] M. Kaiser and C.C. Hilgetag. Spatial growth of real-world networks. Phys. Rev. E, 69:036103, 2004.

[34] M. Kimura, K. Saito, and N. Ueda. Modeling of growing networks with directional attachment and communities. Neural Networks, 17:975–988, 2004.

[35] N. Mathias and V. Gopal. Small worlds: how and why. Phys. Rev. E, 63:021117, 2001.

[36] K. Morita, A.J. Flemming, Y. Sugihara, M. Mochii, Y. Suzuki, S. Yoshida, W.B. Wood, Y. Kohara, A.M. Leroi, and N. Ueno. A *Caenorhabditis elegans* TGF-β, DBL-1, controls the expression of LON-1, a PR-related protein, that regulates polyploidization and body length. EMBO J., 21:1063–1073, 2002.

[37] S. Morita, K.-I. Oshio, Y. Osana, Y. Funabashi, K. Oka, and K. Kawamura. Geometrical structure of the neuronal network of Caenorhabditis elegans. Physica A, 298:553–561, 2001.

[38] V. B. Mountcastle. The columnar organization of the neocortex. Brain, 120:701–722, 1997.

[39] M.E.J. Newman. The structure of scientific collaboration networks. Proc. Natl. Acad. Sci. USA, 98:404–109, 2001.

[40] K. Oshio, Y. Iwasaki, S. Morita, Y. Osana, S. Gomi, E. Akiyama, K. Omata, K. Oka, and K. Kawamura. Database of synaptic connectivity of c. elegans for computation. Technical Report 3, CCeP, Keio Future, Keio University, Japan, 2003. The connectivity database is available online at http://ims.dse.ibaraki.ac.jp/research/database_en.html.

[41] R. Pastor-Satorras, A. Vázquez, and A. Vespignani. Dynamical and correlation properties of the internet. Phys. Rev. Lett., 87:258701, 2001.

[42] R. Pastor-Satorras and A. Vespignani. Evolution and Structure of the Internet. Cambridge University Press, Cambridge, UK, 2004.

[43] P. Rakic. Neurobiology of Neocortex, chapter Intrinsic and extrinsic determinants of neocortical parcellation: A radial unit model, pages 5–27. John Wiley and Sons, Ltd, 1988.

[44] P. Rakic. A small step for the cell, a giant leap for mankind: a hypothesis of neocortical expansion during evolution. Trends Neurosci., 18:383–388, 2000.

[45] P. Rakic. Immigration denied. Nature, 427:685–686, 2004.

[46] E. Ravasz and A.-L. Barabási. Hierarchical organization in complex networks. Phys. Rev. E, 67:026112, 2003.

[47] O. Sporns. Graph theory methods for the analysis of neural connectivity patterns. In R. Kötter, editor, Neuroscience Databases: A Practical Guide, pages 171–186. Klüwer, Boston, MA, 2002.

[48] O. Sporns, D.R. Chialvo, M. Kaiser, and C.C. Hilgetag. Organisation, development and function of complex brain networks. Trends Cog. Sci., 8:418–425, 2004.

[49] S. Valverde, R. Ferrer i Cancho, and R.V. Solé. Scale-free networks from optimal design. Europhys. Lett., 60:512–517, 2002.

[50] J. van Pelt, I. Vajda, P.S. Wolters, M.A. Corner, and G.J.A. Ramakers. Dynamics and plasticity in developing neuronal networks in vitro. Progress Brain Res., 147:173–187, 2005.

[51] H. van Praag, A.F. Schinder, B.R. Christie, B. Toni, T.D. Palmer, and F.H. Gage. Functional neurogenesis in the adult hippocampus. Nature, 415:1030–1034, 2002.

[52] D.J. Watts and S.H. Strogatz. Collective dynamics of 'small-world' networks. Nature, 393:440–442, 1998.

[53] J.G. White, E. Southgate, J.N. Thomson, and S. Brenner. The structure of the nervous system of the nematode Caenorhabditis elegans. Phil. Trans. R. Soc. London B, 314:1–340, 1986.

[54] S.R. Wicks, C.J. Roehrig, and C.H. Rankin. A dynamics network simulation of the nematode tap withdrawal circuit: predictions concerning synaptic function using behavioral criteria. J. Neurosci., 16:4017–4031, 1996.

[55] S. H. Yook, H. Jeong, and A. L. Barabási. Modeling the internet's large-scale topology. Proc. Natl. Acad. Sci. USA, 99:13382–13386, 2002.

[56] G. Zeck and P. Fromherz. Noninvasive neuroelectronic interfacing with synaptically connected snail neurons immobilized on a semiconductor chip. Proc. Natl. Acad. Sci. USA, 98:10457–10462, 2001.