Chapter 1

An optimization approach to the structure of the neuronal layout of *C. elegans*

Alex Arenas, Alberto Fernández and Sergio Gómez
Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, E-43007 Tarragona, Spain

1.1 Introduction

Between 1899 and 1904, in the masterpiece *Textura del sistema nervioso del hombre y de los vertebrados* (published in separated folded sheets during these years), S. Ramón y Cajal established the linchpin of modern neuroscience [1]. Among its capital contributions, he stated the *law of maximum economy in space, time and inter-connective matter*, that explicitly hypothesizes about an optimization of the structure and function of the nervous systems during evolution, reflected in an economical principle for informational driving processes in neuronal circuitries. This fascinating elucidation of the complex structure of nervous systems, has been however very difficult to quantify with real data. The topological mapping of each one of the neurons of a vertebrate's brain is still out of nowadays technical possibilities. However, it exists an invertebrate organism for which the complete neuronal layout is known, the nematode *C. elegans*, see Fig. 1.1. The current computational capabilities and the disposal of such a connectivity data set allow us to explore the conjecture of S. Ramón y Cajal about the "wiring economy principle".

In this chapter we will review a recent optimization approach to the wiring connectivity in *C. elegans*, discussing the possible outcomes of the optimization process, its dependence on the optimization parameters, and its validation with the actual neuronal layout data. We will follow the

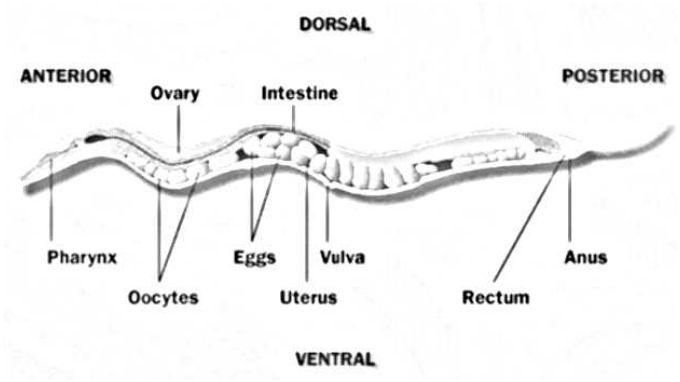

Fig. 1.1 Sketch of *C. elegans* anatomy.

main procedure described in the work by Chen *et al.* [2–4]. The results show that the current approach to optimization of neuronal layouts is still not conclusive, and then the "wiring economy principle" remains unproved.

1.2 The dataset

The nematode *Caenorhabditis Elegans* has become in biology the experimental organism par excellence to understand the mechanisms underlying a whole animal's behavior, at the molecular and cellular levels [5, 6]. It has been extensively studied to understand particular biological phenomena, with the expectation that discoveries made in this organism will provide insight into the workings of other organisms. It is one of the usually called model organisms. In particular, model organisms are widely used to explore potential causes and treatments for human disease when human experimentation would be unfeasible or unethical. This approximation is supported by the fact of common descent of all living organisms, and the conservation of metabolic and developmental pathways and genetic material over the course of evolution. In fact, this goal was a primary motivation behind the development of *C. elegans* as an experimental organism 40 years ago. Yet it has proven surprisingly difficult to obtain a mechanistic understanding of how the *C. elegans* nervous system generates behavior, despite the existence of a "wiring diagram" that contains a degree of information about neural connectivity unparalleled in any organism. Studying model organisms can be informative, but generalizations should be carefully considered.

The structural anatomy of *C. elegans* is basically that of a cylinder around 1 millimeter in length and 0.1 millimeter in diameter, see Fig. 1.1. In the following, we will use the common hypothesis of study of a one dimensional entity. We are interested in its neuronal system, in particular in the position along the body of the different neurons and its interconnections. The current work uses the public data found in [7]. The construction of this data set started with the work by Albertson *et al.*, and White *et al.* [8, 9], and has been contributed by many authors since then, in the multimedia project *Wormatlas* [7]. The particular wiring diagram we use was revised and completed by Chen *et al.* [3] using other valuable sources [11, 12]. The wiring information we have used is structured in four parts: connectivity data between neurons, neuron description, neuron connections to sensory organs and body muscles, and neuronal lineage. The architecture of the nervous system of *C. elegans* shows a bilaterally symmetric body plan. With a few exceptions, neurons in *C. elegans* have a simple uni- or bipolar morphology that is typical for invertebrates. Synapses between neurons are usually formed *en passant* and each cell has multiple presynaptic regions dispersed along the length of the axon.

The neuronal network connectivity of the *C. elegans* can be represented as a weighted adjacency matrix of 279 nonpharyngeal neurons, out of a total of 302 neurons (pharyngeal neurons are not considered in this work because they are not reported in the above mentioned database). The abstraction at this point consists in to assume that the nervous system of the *C. elegans* can be modeled as a network, where nodes represent the center of the cell bodies, and the links represent synapses, see Fig. 1.2. The order and nomenclature of the neurons in the matrix follows that of [12], for a detailed biological record of the dataset see [7]. The position of neurons has been defined in the data set as follows: i) neuron location is considered at the center of the cell body projected onto the anterior–posterior axis of the worm, ii) a neuron is assumed to make a single connection to a given sensory organ, iii) the position of each muscle is defined as the midpoint between anterior and posterior extremities of the sarcomere region, and iv) there is a lack of data specifying the location of individual synapses in the worm.

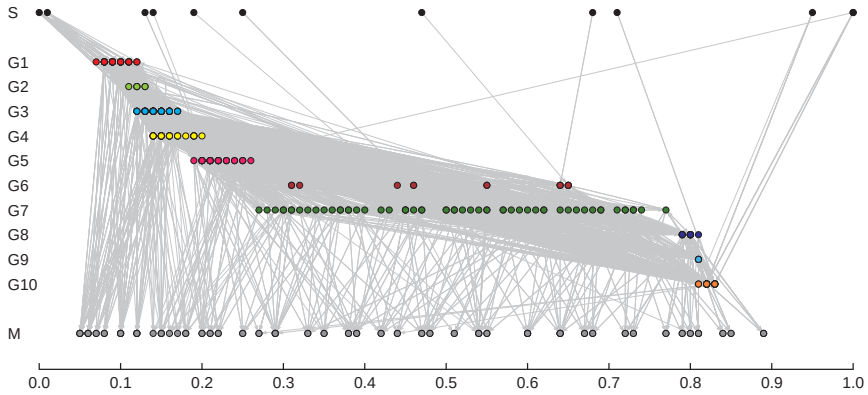

Fig. 1.2 *C. elegans* layout of sensorial (S), motor (M) and non-pharyngeal neuronal cells, and its connectivity. Neurons in the same ganglia are given the same vertical offset for clarity: G1) anterior ganglion, G2) dorsal ganglion, G3) lateral ganglion, G4) ventral ganglion, G5) retrovesicular ganglion, G6) posterolateral ganglion, G7) ventral cord neuron group, G8) pre-anal ganglion, G9) dorsorectal ganglion, G10) lumbar ganglion. The bottom ruler shows the longitudinal assigned coordinates, with values 0.0 and 1.0 for the head and tail of the worm respectively.

1.3 Optimization model problem

Assuming that the neuronal wiring of *C. elegans* has been optimized by natural evolution, it seems plausible to formulate an optimization model problem whose results will reproduce a neuronal layout in agreement with the real data, and support then the hypothesis of the wiring economy principle. The model must be robust to the variation of the parameters, and statistically significant. Here we follow the general formulation presented in [2, 3]. The optimization problem, is stated as a cost function that must be minimized. The cost is separated in two different contributions: a connection cost between neurons, and a connection cost between neurons and sensorial organs or muscles. This difference arises from the different connectivity patterns and also because its mathematical convenience, note that the position of sensors and muscles is taken as input data, constraining then the optimization problem. We will work in the scope of the dedicated-wire model (following the terminology in [3]) which simply means that every synapse has its own wire. At difference, the shared-wire model introduced also in [3] considers a neuron as a wire (segment) with multiple synapses; we do not include its analysis here because it does not provide betters results.

Mathematically, the general formulation of wiring cost is:

$$C^{\text{tot}} = C^{\text{int}} + C^{\text{ext}} \tag{1.1}$$

where

$$C^{\text{int}} = \frac{1}{2\alpha_A} \sum_i \sum_j A_{ij}(x_i - x_j)^2 \tag{1.2}$$

$$C^{\text{ext}} = \frac{1}{\alpha_S} \sum_i \sum_k S_{ik}(x_i - s_k)^2 + \frac{1}{\alpha_M} \sum_i \sum_r M_{ir}(x_i - m_r)^2 \tag{1.3}$$

and

- $\mathbf{A} = (A_{ij})$ neuron-neuron connectivity matrix $(N_A \times N_A)$
- $\mathbf{S} = (S_{ik})$ neuron-sensor connectivity matrix $(N_A \times N_S)$
- $\mathbf{M} = (M_{ir})$ neuron-motor connectivity matrix $(N_A \times N_M)$
- $\mathbf{x} = (x_1, \ldots, x_{N_A})^T$ neurons positions vector
- $\mathbf{s} = (s_1, \ldots, s_{N_S})^T$ sensors positions vector
- $\mathbf{m} = (m_1, \ldots, m_{N_M})^T$ motors positions vector
- α_A, α_S, and α_M are parameters to be determined by normalization constraints

Everything will be treated as input data except the position of the neurons $\mathbf{x}$. This quadratic expression of the cost is just a convention, other exponents have been considered in [3] with no improvement on the optimization results. Moreover, it is mathematically convenient because in this form the system is analytically solvable. The optimization of the total cost (1.1) is obtained by imposing

$$\frac{\partial C}{\partial x_a} = 0\,, \quad a = 1, \ldots, N_A \tag{1.4}$$

They yield a system of linear equations whose solution, in matrix form, reads

$$\mathbf{x} = \mathbf{Q}^{-1} \left(\frac{1}{\alpha_S} \mathbf{S}\mathbf{s} + \frac{1}{\alpha_M} \mathbf{M}\mathbf{m} \right) \tag{1.5}$$

where

$$Q_{ij} = \left(\frac{1}{\alpha_A} \sum_\ell \tilde{A}_{i\ell} + \frac{1}{\alpha_S} \sum_k S_{ik} + \frac{1}{\alpha_M} \sum_r M_{ir} \right) \delta_{ij} - \frac{1}{\alpha_A} \tilde{A}_{ij} \tag{1.6}$$

being δ_{ij} the elements of the identity matrix, and

$$\tilde{\mathbf{A}} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^T) \tag{1.7}$$

Note that the symmetrization of matrix $\mathbf{A}$ is not an initial constraint but a consequence of the optimization process.

It is convenient to define also the following constants for the determination of the α parameters:

$$\tau_A = \frac{1}{2N_A} \sum_i \sum_j A_{ij} \tag{1.8}$$

$$\tau_S = \frac{1}{2N_A} \sum_i \sum_k S_{ik} \tag{1.9}$$

$$\tau_M = \frac{1}{2N_A} \sum_i \sum_r M_{ir} \tag{1.10}$$

The actual neuronal layout of *C. elegans* is represented in Fig. 1.2. In the plot we have separated neurons at different areas and also differentiated them from sensor neurons and muscles. The first interesting observation is that the connectivity is clearly biased towards the head of the animal, where more sensorial connections are established. The muscular connectivity is also more dense in the anterior part of the body although not as dense than for the sensors. Realizing that our constraints are the positions of sensor neurons and muscles, we expect a significant scatter of the predictions in the region between the middle and the posterior part of the body. Moreover, the scatter will result in an under-prediction of the position of neurons, the predicted positions will be biased towards the anterior part.

1.4 Results and discussion

Eq. 1.5 gives the position of the neurons in the abstracted optimization model. To measure the success of this method, the mean absolute difference between the actual ($\mathbf{x}$) and predicted ($\mathbf{x}'$) neuron positions is used:

$$E = \frac{1}{N_A} \sum_i |x_i - x'_i| \tag{1.11}$$

We have prepared several experiments in the scope of the current optimization problem, to check the reliability of the current approach to get support on the hypothesis of wiring optimization in the neuronal layout of *C. elegans*. Here we expose the set up and results of each experiment. Finally a summary is presented in Table 1.1.

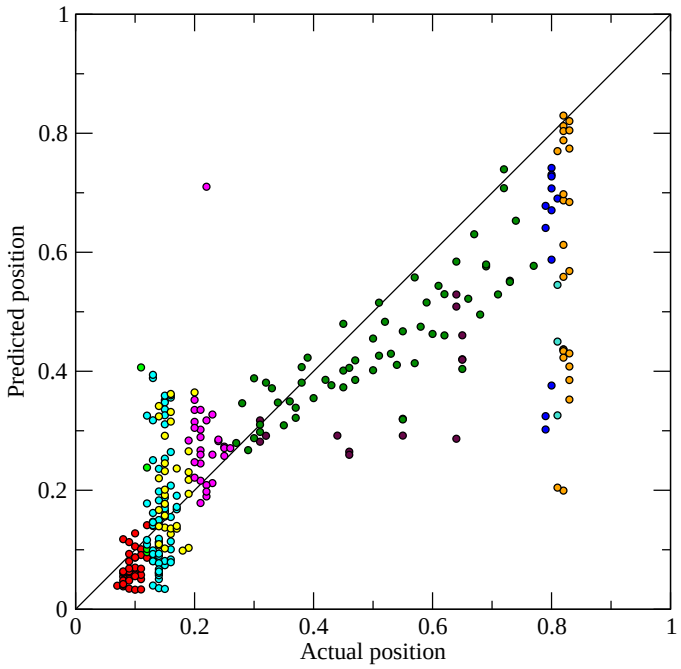

Fig. 1.3 Experiment 1. Predicted versus real neuronal layout of *C. elegans*, setup detailed in text. Colors correspond to the different ganglia shown in Fig. 1.2.

1.4.1 *Experiment 1*

First, we reproduce the results presented in [3] for the dedicated-wire model. In this case, the matrices **A**, **S** and **M** correspond to the weighted connectivity matrices of the real data. The parameters α_A, and α_M are set to normalize the neuron-neuron, and neuron-muscle interaction by the average number of synapses per neurite, 29.44 (or equivalently 58.88 synapses per neuron divided by two neurites per neuron). In the referenced work α_S is set to 1, with no apparent reason, so we did it. The results are depicted in Fig. 1.3, different colors corresponds to the different ganglia shown in Fig. 1.2. The error E in this approximation is 9.69% [1]. The authors contrasted this result with the positioning of neurons uniformly at random along the body of the worm, which raises an $E \approx 34\%$, moreover they computed that the probability of obtaining by chance the results of the

[1] In the original paper the authors find an error of 9.71%, we attribute this difference to the use of different data, since we are using the last update of the data in [7].

optimization is of order 10^{-68}. They also compare the optimization results against the null hypothesis that more related neurons are positioned closer to each other. To this end they use instead of $\mathbf{A}$ the "relatedness" matrix, which is a matrix connecting neurons by number of jumps in the lineage tree, substituting non existent connections in this matrix by a uniform repulsive force [2]. We will reproduce this null hypothesis in Experiment 6. The authors find an error in this case of $E = 26.1\%$ far bigger than the optimization result, supporting then the optimization process as a meaningful description of the relationship between neuronal arrangement and connectivity in the worm.

1.4.2 *Experiment 2*

We noticed that the results presented in [3] can be quantitatively improved by a separate normalization of the neuron-neuron interaction, and the neuron-muscle interaction. Given that the adjacency matrix $\mathbf{A}$, and $\mathbf{M}$ are distinguishable, it makes sense to compute the normalization $\alpha_A = \tau_A$, and $\alpha_M = \tau_M$. In doing this slight change in the parameters, the error reduces to $E = 8.75\%$ which is noticeable in this scenario (nearly a 10% relative improvement with respect to the result in Experiment 1). However, qualitatively the results do not seem to change, see Fig. 1.4.

1.4.3 *Experiment 3*

If the modification in the normalization is a key factor in the optimization process, one expects this to hold also for the normalization of the neuron-sensor connections. Mathematically, there is no handicap for this consideration because it simply implies $\alpha_S = \tau_S$, and it provides consistency with the prescription above. We used this normalization, plus the one in Experiment 2, and obtained the unsatisfactory result of an increasing in the error to $E = 10.15\%$, see Fig. 1.5. This surprising effect raises doubts on the approach, because this sensibility to the normalization parameters was unexpected.

[2] This last prescription seems to us unnecessary because the relatedness matrix even with the zero values for non existent connections is non singular.

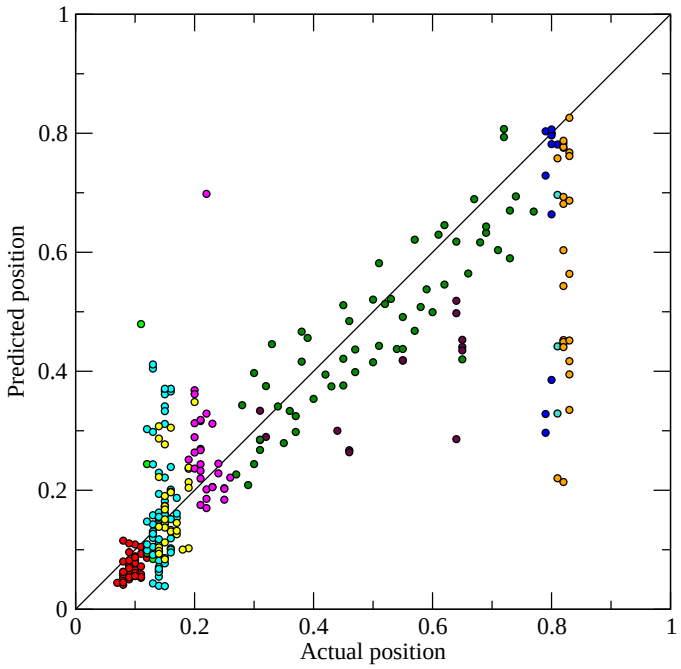

Fig. 1.4 Experiment 2. Predicted versus real neuronal layout of *C. elegans*, setup detailed in text. Colors correspond to the different ganglia shown in Fig. 1.2.

1.4.4 *Experiment 4*

After the last observation, we decided to change the input data considering that the neuron-neuron connectivity matrix is an unweighted matrix. This seems a more realistic approach, since multiple synapses between neurons are not built upon multiple wires but only two neurites with multiple *en passant* synapses. We filtered $\mathbf{A}$ assigning 1 if there is a connection between the neurons, disregarding its number, and 0 otherwise. The matrices $\mathbf{M}$ and $\mathbf{S}$ were remained unchanged. Using the normalization exposed in Experiment 2, $\alpha_A = \tau_A$, and $\alpha_M = \tau_M$, the error now is reduced to $E = 8.33\%$, see Fig. 1.6. Now the results are more striking because they point out not only that the normalization factor notably affects the results, but also that the consideration of the weights (number of synapses) between neurons makes no real difference in the outcome of the optimization process.

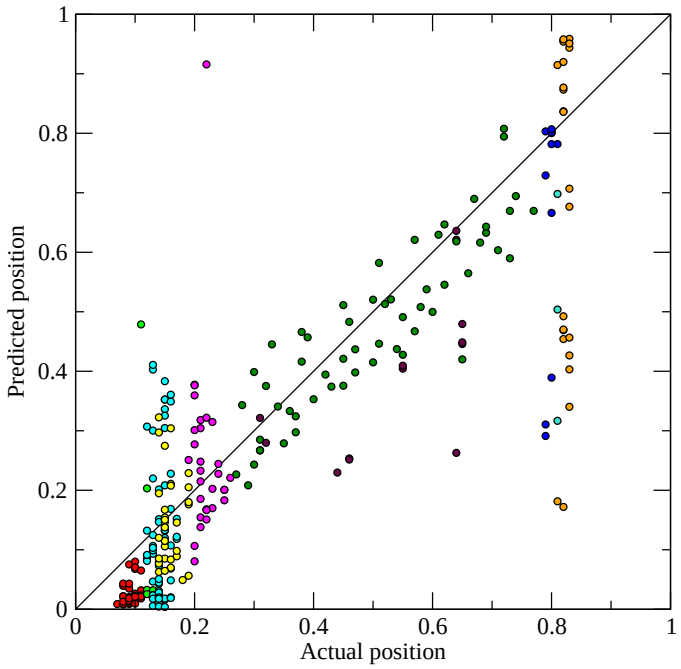

Fig. 1.5 Experiment 3. Predicted versus real neuronal layout of *C. elegans*, setup detailed in text. Colors correspond to the different ganglia shown in Fig. 1.2.

1.4.5 *Experiment 5*

Equivalently to Experiment 3, we introduce the normalization of the neuron-sensor connections $\alpha_S = \tau_S$ in the set up exposed in Experiment 4, i.e. unweighted connectivity matrix. Again the results of the optimization process are worst, with an error $E = 9.62\%$, see Fig. 1.7. Definitely the inclusion of the normalization of the neuron-sensor connections provide worst results in terms of global error. The main reason is that the normalization of neuron-sensor connections enhances even more the connectivity towards the anterior part of the animal. However, we think that its elimination is not mathematically consistent, and then it must be preserved in the same way we preserve the normalization of the other type of connections.

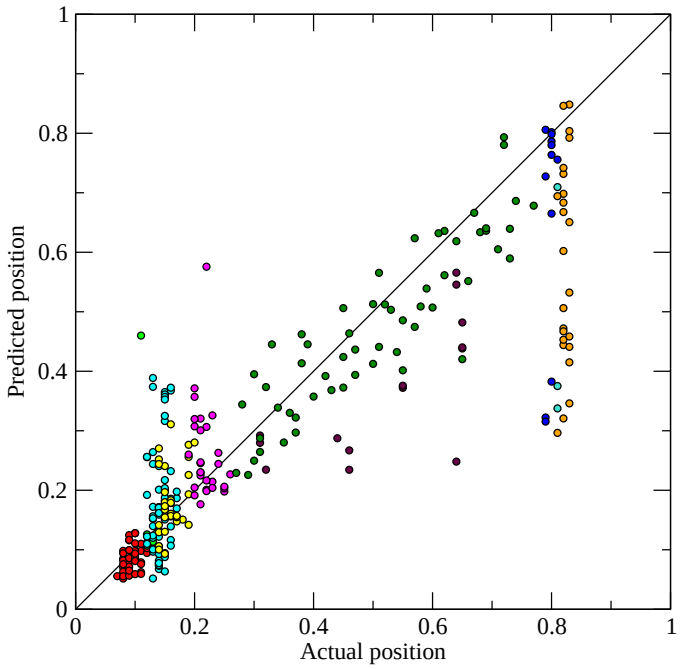

Fig. 1.6 Experiment 4. Predicted versus real neuronal layout of *C. elegans*, setup detailed in text. Colors correspond to the different ganglia shown in Fig. 1.2.

1.4.6 *Experiment 6*

At the light of the previous results, we think that the contribution of the actual neuronal connectivity in the global error of the optimization process is very low. Following the idea in [3], we used the relatedness of neurons, which is a measure of distance in the lineage tree (not linearly correlated with the connectivity), instead of the adjacency matrix. The authors in [3] did something similar and provided an error significantly larger than the error considering the actual connectivity. We simply substituted the adjacency matrix by the relatedness matrix, and normalized according to the previous Eq. 1.8. The error in this case is $E = 11.76\%$, slightly larger than the error in Experiment 1 but absolutely comparable, moreover the scatter of the neuron positions is qualitatively equivalent to that in Experiments 1 to 5, see Fig. 1.8.

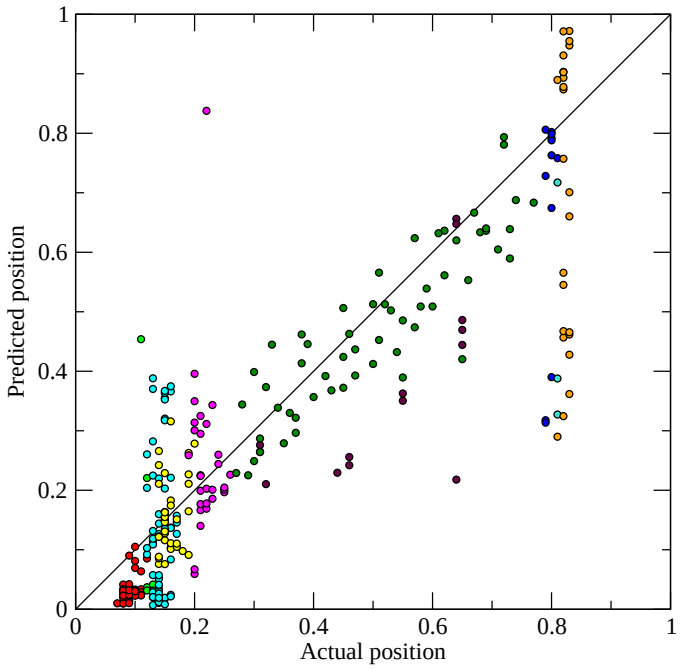

Fig. 1.7 Experiment 5. Predicted versus real neuronal layout of *C. elegans*, setup detailed in text. Colors correspond to the different ganglia shown in Fig. 1.2.

1.4.7 *Experiment 7*

Finally, we want to show the results of a really out-minded experiment where the true adjacency matrix is replaced by an all-to-all connectivity between neurons. The result obtained in this particular case, raises an error $E = 11.60\%$ which is comparable to the results so far obtained using the true connectivity matrix, see Fig. 1.9. This experiment raises serious doubts about the validity of the optimization procedure proposed so far, as a way to test the wiring economy principle in neural networks.

1.4.8 *Summary*

The summary of the experiments results are presented in Table 1.1. These show that the outcome of the optimization problems are basically driven by the information we provide *a priori*, i.e. the position of the sensor and muscle neurons, which determine the main anchorage of neurons' connections,

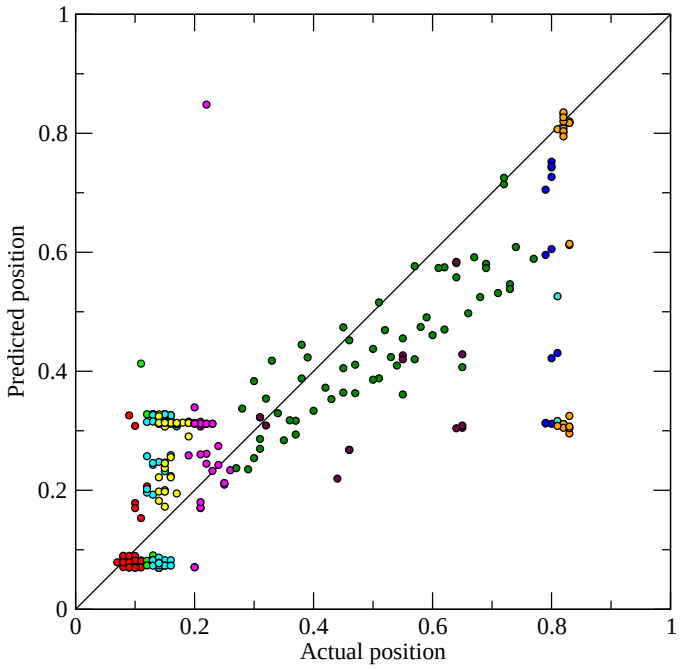

Fig. 1.8 Experiment 6. Predicted versus real neuronal layout of *C. elegans*, setup detailed in text. Colors correspond to the different ganglia shown in Fig. 1.2.

with very low impact of the neuron-neuron connectivity in the process. We also present a persuasive and forceful comparison of the resulting layouts of the experiments at Fig. 1.10. Note that the kernel of the prediction success is in the ventral cord neuron group (G7) in the central area of the worm, which is not surprising given the biological nature of these neurons, see [10]. The ventral cord is spatially coherent; neurons, particularly interneurons running in the cord, maintain their positions relative to their nearest neighbors in spite of local distortions produced by intrusions of cell bodies, which means that any optimization procedure where the motor neurons are fixed will provide, given the motor-neuron connectivity, a good balance in the positioning of neurons of G7. This is pretty uninformative about the actual position of the rest of neurons but eventually gives a bound for the error that dominates the out-coming results. In our opinion, these findings arise important concerns about the method to support the wiring economy principle in neuronal networks, and encourage the scientific community to find

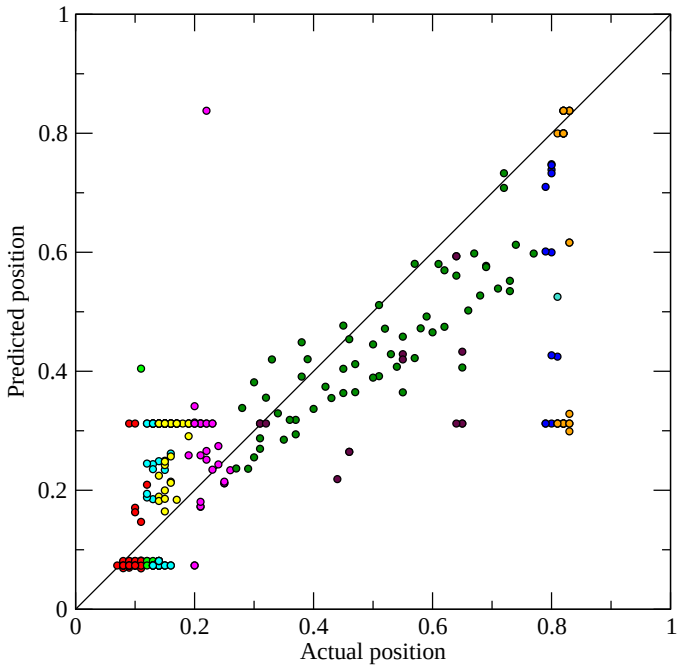

Fig. 1.9 Experiment 7. Predicted versus real neuronal layout of *C. elegans*, setup detailed in text. Colors correspond to the different ganglia shown in Fig. 1.2.

new approaches to reveal the mechanisms governing the biological topology of neuronal networks.

Table 1.1 Summary of the different experiments and results

Exp.	Parameters	E
1	$A_{ij} >= 0$ (see Ref. 3) $\alpha_A = \alpha_M = \tau_A + \tau_M = 29.44$ $\alpha_S = 1$	9.69%
2	$A_{ij} >= 0$ $\alpha_A = \tau_A = 26.10$ $\alpha_M = \tau_M = 3.34$ $\alpha_S = 1$	8.75%
3	$A_{ij} >= 0$ $\alpha_A = \tau_A = 26.10$ $\alpha_M = \tau_M = 3.34$ $\alpha_S = \tau_S = 0.15$	10.15%
4	$A_{ij} \in \{0, 1\}$ $\alpha_A = \tau_A = 8.20$ $\alpha_M = \tau_M = 3.34$ $\alpha_S = 1$	8.33%
5	$A_{ij} \in \{0, 1\}$ $\alpha_A = \tau_A = 8.20$ $\alpha_M = \tau_M = 3.34$ $\alpha_S = \tau_S = 0.15$	9.62%
6	A_{ij} replaced by relatedness R_{ij} (lineage distance) $\alpha_A = \tau_A = 2255.97$ $\alpha_M = \tau_M = 3.34$ $\alpha_S = \tau_S = 0.15$	11.76%
7	$A_{ij} = 1\,,\ \forall i, j$ $\alpha_A = \tau_A = 139.60$ $\alpha_M = \tau_M = 3.34$ $\alpha_S = \tau_S = 0.15$	11.60%

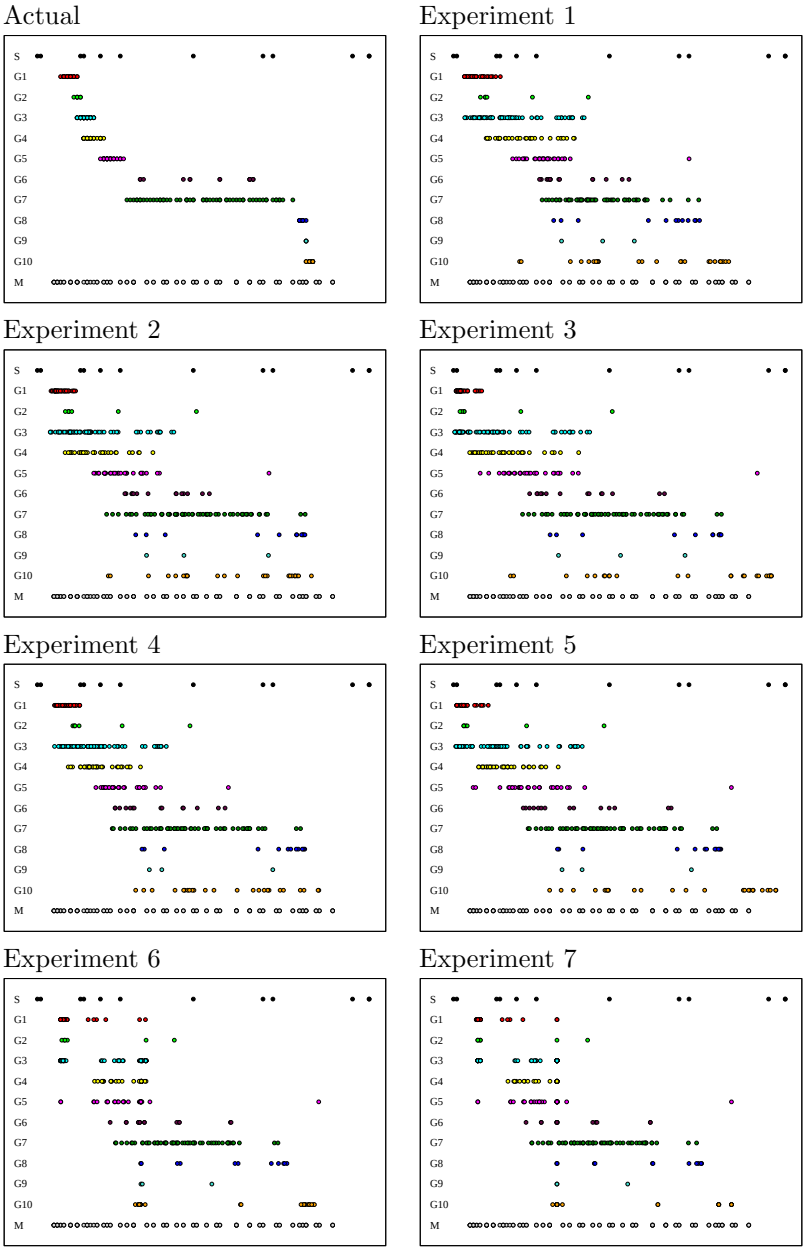

Fig. 1.10 Comparison of the *C. elegans* real neuronal layout and predictions of the experiments described in the text.

Bibliography

[1] Ramón y Cajal, S. (1899). Textura del Sistema Nervioso del Hombre y de los Vertebrados. Translation: Texture of the Nervous System of Man and the Vertebrates. Vol. 1. 1899, New-York: Springer (1999) 631.

[2] Chklovskii, D. B. (2004). Exact Solution for the Optimal Neuronal Layout Problem, *Neural Computation* **16**(10), pp. 2067–2078.

[3] Chen, B. L., Hall, D. H. and Chklovskii, D. B. (2006). Wiring optimization can relate neuronal structure and function, *Proc. Natl. Acad. Sci. USA* **103**, pp. 4723–4728.

[4] Chen, B. L. and Chklovskii, D. B. (2006). Placement and routing optimization in the brain, *Proc. Int. Symposium on Physical Design (ISPD'06)* (San José, California, USA), pp. 136–141.

[5] Wood, W.B., ed. (1998). The nematode Caenorhabditis elegans, Cold Spring Harbor Laboratory: Cold Spring Harbor, New York. 25.

[6] Schafer, W.R. (2005). Deciphering the Neural and Molecular Mechanisms of C. elegans Behavior, *Current Biology* **15**, pp. R723–R729.

[7] http://www.wormatlas.org/handbook/nshandbook.htm/nswiring.htm

[8] Albertson, D.G. and Thomson, J.N. (1976). The Pharynx of Caenorhabditis elegans, *Phil. Trans. R. Soc. London, series B* **275**, pp. 299–325.

[9] White, J.G., Southgate, E., Thompson, J.N., Brenner, S. (1986). The structure of the nervous system of the nematode caenorhabditis elegans, *Phil. Trans. Royal Soc. London. Series B* **314**, pp. 1–340.

[10] White, J.G., Southgate, E., Thompson, J.N., Brenner, S. (1976). The structure of the ventral nerve cord of the nematode caenorhabditis elegans, *Phil. Trans. Royal Soc. London. Series B* **275**, pp. 327–348.

[11] Durbin, R.M. (1987). Studies on the Development and Organisation of the Nervous System of Caenorhabditis elegans. PhD Thesis, University of Cambridge.

[12] Achacoso, T.B., Yamamoto, W.S. (1992). AY's Neuroanatomy of C. elegans for Computation. CRC Press, Boca Raton.