

Optimally wired subnetwork determines neuroanatomy of *Caenorhabditis elegans*

Alfonso Pérez-Escudero* and Gonzalo G. de Polavieja†

Neural Processing Laboratory, Instituto "Nicolás Cabrera" de Física de Materiales, and Department of Theoretical Physics, Universidad Autónoma de Madrid, 28049 Madrid, Spain

Edited by Solomon H. Snyder, Johns Hopkins University School of Medicine, Baltimore, MD, and approved September 7, 2007 (received for review April 5, 2007)

Wiring cost minimization has successfully explained many structures of nervous systems. However, in the nematode *Caenorhabditis elegans*, for which anatomical data are most detailed, wiring economy is thought to play only a partial role and alone has failed to account for the grouping of neurons into ganglia [Chen BL, Hall DH, Chklovskii DB (2006) *Proc Natl Acad Sci USA* 103:4723–4728; Kaiser M, Hilgetag CC (2006) *PLoS Comput Biol* 2:e95; Ahn Y-Y, Jeong H, Kim BJ (2006) *Physica A* 367:531–537]. Here, we test the hypothesis that optimally wired subnetworks can exist within nonoptimal networks, thus allowing wiring economy to give an improved prediction of spatial structure. We show in *C. elegans* that the small subnetwork of wires connecting sensory and motor neurons with sensors and muscles, comprising only 15% of connections, is close to optimal and alone predicts the main features of the spatial segregation of neurons into ganglia and encephalization. Moreover, a method to dissect networks into optimal and nonoptimal components is shown to find a large near-optimal subnetwork of 84% of neurons with a very low position error of 5.4%, and that explains clustering of neurons into ganglia and encephalization to fine detail. In general, we expect realistic networks not to be globally optimal in wire cost. We thus propose the strategy of using near-optimal subnetworks to understand neuroanatomical structure.

encephalization | ganglia | optimization | wiring economy | anatomy

Why do neurons organize spatially in ganglia, strata, brain areas, and maps? Spatial patterns might have several origins. Evolutionary and developmental constraints underlie spatial organization in nervous systems (1). Neuromodulation can be more effective when neurons form spatial groups (2). Processing can require the use of different wire lengths and a spatial organization of delays (3). Wiring economy (4) can also explain spatial patterns and has the advantage of being quantitatively testable. This principle states that the morphology of nervous systems is such that the cost of interconnecting its parts is the minimum possible given that the conditions for the proper functioning of the system are met. It is simple to see that wiring economy can result in the organization of the nervous system into spatial structures. Imagine, for example, the simple case of two types of neurons, say, *R* or *L* neurons responding to the excitation of right and left eyes in a human, respectively. If *R* neurons are heavily connected among themselves and the same for the *L* neurons, whereas the *R*–*L* connections are few, it can save wire to have the *R* neurons and *L* neurons segregated into two different spatial groups, because mixing them would increase the amount of wire used.

The wiring economy principle has been applied to many structures in nervous systems. It has been used to explain large brain structures, such as the existence of brain areas (5), neocortical folding (6), retinotopic maps (7, 8), ocular dominance patterns (5, 9), orientation maps (5, 10, 11), segregation of gray and white matter (12–14), and the arrangement of areas in vertebrate cortex and ganglia in *C. elegans* (15–18). Wiring economy has also been

used to explain the structure of neurons, including their dimensions (19–21) and branching angles (22, 23).

The most detailed application of wiring economy has been to the entire nervous system of *Caenorhabditis elegans* (24). These authors updated the wiring diagram of *C. elegans* (25, 26) and used methods developed previously for circuit board design (27, 28) to predict neuronal layout (18). Predicted positions differed from actual ones on average <10% the length of the animal. These results showed that wiring economy is at work in *C. elegans*, but the layout was found to be nonoptimal (29, 30). The actual spatial organization in distinct ganglia could not be obtained from wiring economy applied to the complete network unless extra factors related to axonal guidance and command neuron function were added to the theoretical prediction (24).

Here, we use wiring economy alone to predict the clustering of neurons in ganglia in *C. elegans*. The article is organized as follows. First, we show that the optimization of the subnetwork of connections from sensory and motor neurons to sensors and muscles, respectively, is responsible for the main features of the clustering in ganglia. Second, we show that wiring economy further applied to the connections among sensory and motor neurons lowers the mean error of predicted neuron positions and improves details of the prediction of clustering. Third, we propose a method to dissect networks into optimal and nonoptimal components. Application to the experimental configuration of neurons in *C. elegans* finds an optimal subnetwork of 84% of neurons that explains the neuroanatomy of the animal down to fine details and using only wiring economy.

Results

Wiring Optimization of Complete Network Does Not Predict Actual Clustering. The *C. elegans* network is formed by 279 neurons, excluding pharyngeal and unconnected neurons, organized in clusters known as ganglia (Fig. 1A; see ref. 25). There are 10 ganglia known as anterior (G1 in Fig. 1A), dorsal (G2), lateral (G3), ventral (G4), retrovesicular (G5), posterolateral (G6), ventral cord (G7), preanal (G8), dorsorectal (G9), and lumbar (G10). Ganglia G1–G5 are clustered in the head and G8–G10 in the tail. The head also has a high concentration of synapses in a region known as nerve ring. There are 73 sensory neurons (neurons connected to sensors and other neurons), 113 motor neurons (neurons connected to muscles and other neurons), 13 neurons that are simultaneously sensory and

Author contributions: A.P.-E. and G.G.d.P. designed research, performed research, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

*Present address: Instituto de Óptica "Daza de Valdés," Consejo Superior de Investigaciones Científicas, Serrano 121, 28006 Madrid, Spain.

†To whom correspondence should be addressed. E-mail: gonzalo.polavieja@uam.es.

This article contains supporting information online at www.pnas.org/cgi/content/full/0703183104/DC1.

© 2007 by The National Academy of Sciences of the USA

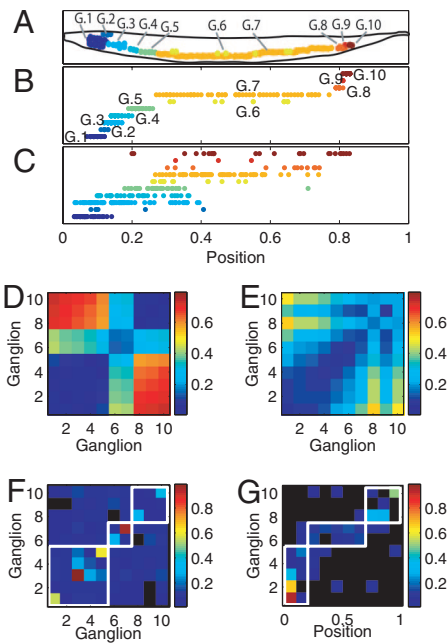


Fig. 1. Optimization of the complete *C. elegans* network fails to predict actual clustering. (A) Soma positions in the actual nematode are organized in spatial clusters known as ganglia. (B) Soma positions in the actual nematode. Somas in different ganglia are offset for clarity. (C) Soma position obtained by minimizing the wiring cost of the complete network. Neurons in different ganglia in the actual network are offset for clarity. (D) Average distance between somas belonging to the same ganglion (diagonal elements) and different ganglia (nondiagonal elements) for the actual nematode. (E) Same as D but obtained for soma positions minimizing the total wiring cost. (F) Average number of connections per neuron between ganglia. Diagonal squares, connections between neurons belonging to the same ganglion; nondiagonal squares, connections between neurons belonging to different ganglia. White boxes enclose regions of the graph corresponding to ganglia overlapping in the one dimension considered. Elements with a value of exactly 0 are colored in black. (G) Average number of connections per neuron between ganglia and organs that lie in each of the 10 bins we divided the length of the nematode. White boxes enclose regions corresponding to connections between ganglia and organs located at the same position in the body as the ganglion. Elements with a value of exactly 0 are colored in black. C and E were obtained with $\alpha = \beta = 1/29.3$.

motor, and 80 interneurons (neurons connected only to other neurons).

We adopt the “dedicated wire” cost model of (24) for the *C. elegans* network. Because the length of this nervous system is more than 10 times its diameter, the model considers a single dimension for simplicity. The network is modeled as a set of nodes representing the cell bodies connected by wires, which represent axons and dendrites. Although in *C. elegans* each axon makes more than one synapse, the model in principle assumes that each synapse has a dedicated wire. The total wiring cost W is written as the sum of three costs, one for connections among neurons, a second one for connections between neurons and sensors, and a third one for connections between neurons and muscles

$$W = \frac{1}{2} \alpha \sum_{i,j=1}^N A_{ij} |x_i - x_j|^\xi + \sum_{i,k=1}^{N,S} B_{ik} |x_i - s_k|^\xi + \beta \sum_{i,l=1}^{N,M} C_{il} |x_i - m_l|^\xi, \quad [1]$$

with A_{ij} , B_{ik} , and C_{il} the number of synapses between neurons i and j , neuron i and sensor k , and neuron i and muscle l , respectively. N , S , and M are the number of neurons, sensors, and muscles,

respectively. x_i is the 1D projection of the position of neuron i , with s_k and m_l analogously for the fixed positions of sensor k and muscle l , respectively. All positions have values from 0 to 1 (where 0 is the head and 1 is the tail). The $1/2$ factor multiplying A_{ij} takes into account the fact that the distance between each pair of neurons is summed twice (note that A is a symmetric matrix). α and β are normalization constants to take into account differences in average cost of the sensory neuron, motor neuron, and interneuron groups. These coefficients can account, for example, for part of the difference between the structure of the “dedicated wire” model and the real network. In the actual network, each neurite that connects two neurons or a neuron and a muscle holds on average 29.3 synapses. Because the model assumes a wire per synapse, the extra cost can be discounted by dividing the wire cost of neuron-to-neuron and neuron-to-muscle connections by 29.3, that is, by making $1/\alpha = 1/\beta = 29.3$ (24). On the other hand, actual connections between neurons and sensors do have a dedicated neurite for each synapse, so, in this case, the model is already in good agreement with reality. Wire cost may be due, for example, to wire volume (5, 18, 22), signal delay and attenuation (31, 32), metabolic costs (33, 34), noise and developmental costs (35). Its functional form is unknown but it is clear that it should increase with wire length. Here, it is written as a power ξ of the wire length. Because wiring cost minimization is analytical for the quadratic case, $\xi = 2$, results have always been obtained for this exponent except where indicated otherwise. Numerical tests suggest that these results are robust with respect to a change in the precise form of the cost (see *Materials and Methods*). The optimal neuron positions minimizing the total cost W in Eq. 1 for $\xi = 2$ can be obtained analytically as follows [see refs. 18, 27, and 28 and supporting information (SI) Text]:

$$\vec{x} = Q^{-1}[B\vec{s} + \beta C\vec{m}] \quad [2a]$$

$$Q_{ip} = \delta_{ij} \left(\alpha \sum_{j=1}^N A_{ij} + \sum_{k=1}^S B_{ik} + \beta \sum_{l=1}^M C_{il} \right) - \alpha A_{ip}. \quad [2b]$$

The optimal positions of the 279 nonpharyngeal neurons calculated from Eq. 2 and using the connectivity matrices and sensor and muscle positions of the real nematode (see *Materials and Methods*) compare well with the actual positions of the somas with a low average error of $e_p = 9.71\%$, much lower than the average error of random positions, $e_{p,rand} = 34.6\%$ (24). However, the correct clustering is not predicted by wire cost minimization of the full network (Fig. 1 B–E). Neurons within each ganglion are closer to each other than to neurons in other ganglia, although in the 1D projection some ganglia overlap. The first five ganglia, located in the head of the nematode, are compact and overlap in the 1D projection. The same happens with the last three, which are located in the tail. The other two (ganglia 6 and 7) are much more elongated and located at the mid-body region. Fig. 1B shows the positions of the neurons along the actual nematode, grouped by ganglia, and offset for clarity. Fig. 1C shows the positions of the predicted positions of somas obtained from wire cost minimization applied to the complete network. As in Fig. 1B, neurons are grouped by ganglia in the actual animal and offset for clarity. Neurons belonging to the same ganglion in the actual network are now dispersed. Fig. 1D shows for the actual nematode the mean distance between neurons in the same ganglion (diagonal squares), and between neurons that lie in different ganglia (off-diagonal squares) (see *Materials and Methods*). The clustering structure seen in the actual nematode, Fig. 1D, is however very different to the one predicted by wire cost minimization applied to the complete network (Fig. 1E; see also Fig. 1C for an intuitive picture). To quantify how far the predicted clustering structure is from the actual one, we subtracted the clustering graphs in Fig. 1D and E and calculated the mean resulting value as the clustering error, $e_c = \sum_{i,j}^G |a_{ij}(\text{actual}) - a_{ij}(\text{predicted})|/G^2$, where

$G = 10$ is the number of ganglia, and a_{ij} (actual) and a_{ij} (predicted) are the elements of the actual and predicted clustering graphs, respectively. The clustering error e_c has values between 0 (perfect prediction) and 1 (animal length), and we multiply it by 100 to give values as percentage of animal length, similarly to the neuron position error. Its value for the optimization of the complete network is $e_c = 16.4\%$ the length of the animal, a value closer to the case of random positions, $e_{c,rand} = 21.8\%$, than to a perfect prediction. In the following, we always use the notation e_c for clustering error as defined above and e_p for average position error.

Although the actual clustering in *C. elegans* is not consistent with wiring cost minimization of the complete network, we wondered whether it could be consistent with wiring economy of subnetworks. If clustering is due to wiring economy, there could be in principle the following two simple scenarios for its origin. Wiring economy would predict clustering when there are more connections within a group of neurons than between neurons belonging to different groups. In this case, neurons belonging to the same group or ganglion would cluster to minimize the cost of intra-ganglion wires. Fig. 1F shows the number of connections per neuron between the neurons of the same ganglion (diagonal elements) and between neurons of different ganglia (off-diagonal elements). Elements representing exactly 0 connections per neuron are colored in black. Clustering would naturally arise in the optimal layout if diagonal elements were much stronger than nondiagonal ones. Furthermore there should be much higher connectivity between ganglia that are close to each other (e.g., ganglia 1 to 5) than between distant ganglia (e.g., ganglia 3 and 10). White lines in Fig. 1F limit these regions of the graph for which high connectivity between ganglia would be consistent with the actual layout. For example, elements within the big white box in the lower-left corner represent connectivity among the first five ganglia, which are very close to each other in the actual nematode. The other two white boxes have a similar interpretation. Elements outside the white boxes represent connectivity between distant ganglia. Although elements with highest connectivity values lie within the white boxes, elements outside them are also important with each ganglion connected to almost all ganglia, so a simple picture for clustering does not emerge.

An alternative explanation for the clustering of neurons in ganglia would consist in a high interconnectivity between neurons and sensors and muscles. Fig. 1G shows the number of connections per neuron between each ganglion and the organs, sensors and muscles, located within small segments of the nematode. In this case, the connectivity is much more consistent with the actual layout of the ganglia: ganglia 1 to 5 are strongly connected to a small region at the head of the nematode, ganglia 6 and 7 are connected to a large region in the middle of the body, and ganglia 8 to 10 are connected to a small region at the tail. Vertical white lines in Fig. 1G limit these regions of connectivity, and horizontal white lines separate the three groups of ganglia. It is clear that each ganglion is strongly connected only to one region of the animal, and very weakly connected, or not connected at all (black elements in Fig. 1G), to organs in other regions. This connectivity pattern suggests a simple picture for clustering based on wiring economy and a high interconnectivity between sensory and motor neurons and spatial patches of organs.

Wiring Optimization of Subnetwork of Connections from Sensory and Motor Neurons to Organs Predicts Main Features of Clustering. We therefore tested the prediction of cost minimization of wires connecting sensory and motor neurons with sensors and muscles, respectively. The associated cost for this case can be obtained from the total cost in Eq. 1 making $\alpha = 0$, that is, eliminating from the optimization the costs of neuron-to-neuron wires,

$$W = \sum_{i,k=1}^{\hat{N},S} B_{ik}|x_i - S_k|^\xi + \beta \sum_{i,l=1}^{\hat{N},M} C_{il}|x_i - m_l|^\xi, \quad [3]$$

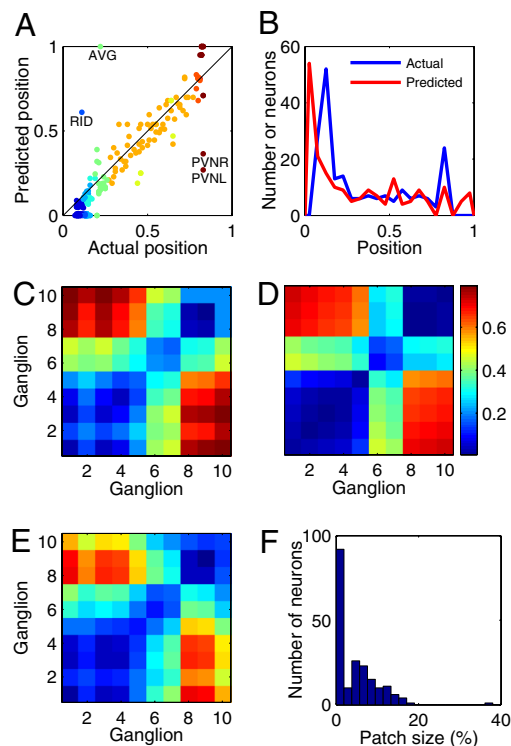


Fig. 2. Optimization of neuron-to-organ connections predicts main features of clustering. (A) Optimal positions of sensory and motor neurons, obtained by minimizing the cost of wiring them to organs, versus their actual positions. Colors distinguish different ganglia. (B) Density sensory and motor neurons along the nematode for actual and optimal positions. (C) Average distance between neurons belonging to the same ganglion (diagonal elements) or different ganglia (nondiagonal elements) for optimized sensory and motor neuron positions. (D) Same as C but for the actual nematode. (E) Same as C but optimizing the complete network. (F) Histogram of sizes of the minipatches of skin to which each neuron connects.

with $\hat{N} = 199$ the total number of sensory and motor neurons. Minimization of this cost function for cost exponent $\xi = 2$ gives explicitly the optimal position for sensory and motor neurons of the form (see *SI Text*)

$$x_i = \frac{\sum_{k=1}^S B_{ik} S_k + \beta \sum_{l=1}^M C_{il} m_l}{\sum_{k=1}^S B_{ik} + \beta \sum_{l=1}^M C_{il}}. \quad [4]$$

Therefore, wiring economy in this case reduces to a simple prediction by which each sensory and motor neuron should be located at the center of mass of the sensors and muscles it connects to. Despite its simplicity, the center of mass formula (Eq. 4) predicts to a very good approximation the neuronal layout of the 199 sensory and motor neurons of *C. elegans* (mean position error $e_p = 9.08\%$, $r = 0.923$ for $\beta = 1$) (Fig. 2A). Also, the density of the neurons along the animal is similar to the actual one, predicting encephalization as seen in the real animal although with brain ganglia shifted to a more anterior position (Fig. 2B). More importantly, the main features of the clustering pattern in ganglia are correctly predicted. The predicted distances between neurons from the same ganglion and between neurons from different ganglia (Fig. 2C, clustering error $e_c = 6.99\%$), compare well with the actual clustering pattern of sensory and motor neurons (Fig. 2D).

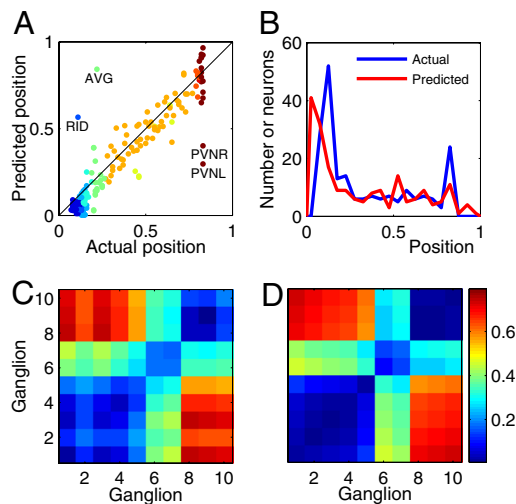


Fig. 3. Connections among sensory and motor neurons improve predicted clusters. (A) Optimal positions of sensory and motor neurons, obtained by minimizing their connections to organs and among themselves, versus their actual positions. Colors distinguish different ganglia. (B) Neuron density along the nematode, for actual and optimal positions. (C) Average distance between neurons belonging to the same ganglion (diagonal elements) or different ganglia (nondiagonal elements) for optimized sensory and motor neuron positions. (D) Same as C but for the actual nematode.

The predictions obtained using the center of mass formula in Eq. 4 are extremely robust. Varying the parameters β and the cost power ξ have an effect on average position error of 1% and on clustering error $<0.3\%$ because of the fact that sensors and muscles are connected to small skin patches (Fig. 2F; see also *SI Text*). It is interesting to compare the results of this simple center-of-mass calculation with the predictions for sensory and motor neurons when optimizing the complete network, Eq. 2. The difference between the two calculations is that the complete network one also

includes all neuron-to-neuron connections. For the complete network calculation we found that average position error e_p and clustering error e_c are lowest for $\alpha = 0.05$ and $\beta = 1.5$. However, even for this best case the clustering is worse than that obtained using the simple center-of-mass calculation (Fig. 2E, $e_c = 7.26\%$). For other values like $\alpha = \beta = 1/29.3$ chosen in reference (24) the clustering error is even higher (SI Fig. 6A, $e_c = 12.3\%$). We thus conclude that the net effect of the neuron-to-neuron connections in the optimization worsens the predictions. The rest of the paper is dedicated to prove that predictions can however be improved further by including part of the neuron-to-neuron connections. In this way we will be able to predict neuron positions for more than just the 199 sensory and motor neurons and explain neuroanatomy to finer details.

Optimal Connections Among Sensory and Motor Neurons Improve Predicted Clusters. We started by adding the connections among sensory neurons, motor neurons and between sensory and motor neurons. The effect of these neuron to neuron connections is obtained from wiring economy of the subnetwork of sensory and motor neurons using the total cost W in Eq. 1 but restricting the sums to sensory and motor neurons. We found an improvement over the center of mass formula in Eq. 4 for neuronal positions, clustering structure and neuron statistics along the animal for a large parameter range, $\alpha \leq 0.1$ and $\beta \geq 0.06$. The best results for neuronal positions and clustering structure were found for $\alpha = 0.05$ and $\beta = 1.5$ (in the remaining, we always use these parameter values). Neuronal layout (Fig. 3A, $e_p = 7.71\%$, $r = 0.93$) and clustering in ganglia (Fig. 3B and C; $e_c = 5.23\%$) are closer to actual ones (Fig. 3D).

Dissection into Optimal and Nonoptimal Subnetworks. To proceed further, we needed a method to dissect the complete network into optimal and nonoptimal subnetworks in the sense of Eq. 2. The number of subnetworks in a system of N neurons is 2^N , an astronomically large number for analysis. Instead, we have used a method of analysis that uses of the order of N calculations. In the following, we illustrate the core of our method using the toy

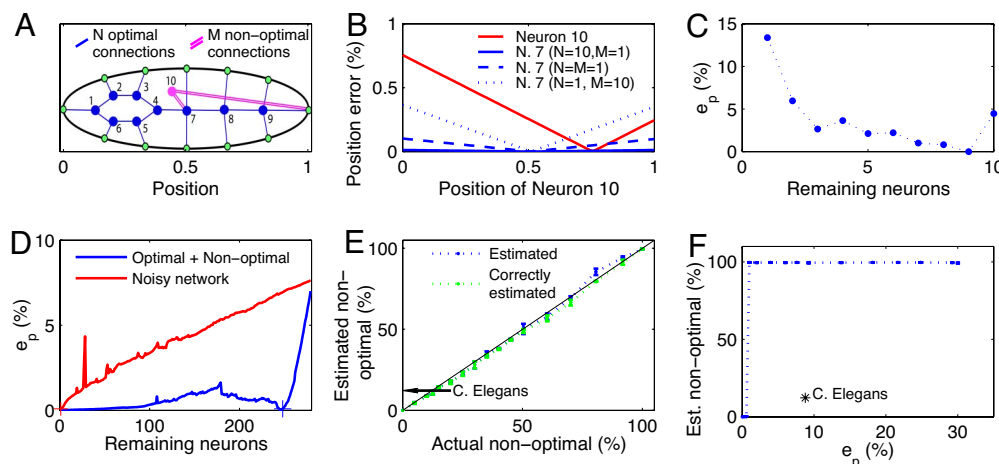


Fig. 4. Method to dissect networks into optimal and nonoptimal subnetworks. (A) Toy network configuration. Neurons 1–9 (blue) are optimal, and neuron 10 (pink) is located at random. Blue and pink links represent N and M connections each, respectively. (B) Position error for nonoptimal neuron 10 (red) and optimal neuron 7 (blue) as a function of the position of the nonoptimal neuron 10 (rest of neurons have vanishing error). Step 1 in the dissection method classifies neurons by their position error: neuron 10 is classified as the worst one when the red line is above the blue line ($N > M$ and neuron 10 located sufficiently far from its center-of-mass position), then neuron 7, and then the rest. (C) Step 2 in the dissection method calculates the average position error e_p in networks of decreasing size (from right to left) eliminating worst located neurons in the order determined in step 1 of the method. Minimum of the error located at size 9. (D) Same as C but for a network configuration with the *C. elegans* connectivity and with optimal and nonoptimal subnetworks (blue) and for a noisy network with neurons located at positions obtained adding to optimal positions noise following a uniform distribution of $\text{std} = 8\%$ (red). Plus signs indicate the point separating optimal and nonoptimal subnetworks. (E) Estimated size of the nonoptimal subnetwork (blue) and percentage of correctly estimated nonoptimal neurons (green). (F) Dissection method does not find a separation between optimal and nonoptimal components in noisy networks built without optimal and nonoptimal components.

connectivity illustrated in Fig. 4A, with 9 optimal neurons in blue and the 10th neuron in pink in a random position, also connected to sensors or muscles (green circles). Blue and pink links indicate N and M connections, respectively. The first step of the method consists in classifying the neurons by their position error measured for each neuron as the difference between its actual position and its locally optimal position, calculated as the position of its center-of-mass that in general reads as

$$x_i^{CM} = \frac{\alpha \sum_{j \neq i} A_{ij} x_j^{\text{actual}} + \sum_k B_{ik} S_k + \beta \sum_l C_{il} m_l}{\alpha \sum_{j \neq i} A_{ij} + \sum_k B_{ik} + \beta \sum_l C_{il}}, \quad [5]$$

with x_j^{actual} the actual positions of the rest of the neurons ($\alpha = \beta = 1$ in toy model). For the example in Fig. 4A, all of the optimally placed neurons have vanishing position errors, except neuron 7 that is connected also to the nonoptimal neuron 10. Although neuron 7 is optimal, its estimated position error depends linearly on the position of neuron 10 with a slope given by $M/(M + 4N)$. As long as the nonoptimal neuron 10 is located sufficiently far from its center-of-mass position and the optimal links outnumber nonoptimal ones, $N > M$, neuron 7 is then correctly classified as having less position error than neuron 10, (Fig. 4B, continuous and dashed blue line). The second step of the method consists in performing a full optimization using Eq. 2 with the complete network first and then further optimization calculations eliminating one by one the neurons in order of decreasing position error as determined by step 1, and for each of these subnetworks we calculate the average position error e_p . In our example this procedure corresponds to an optimization calculation with all neurons first, then we eliminate neuron 10, followed by neuron 7 and then any of the rest, and calculate for each case the average position error, Fig. 4C. The lowest average position error is found for neurons 1–9, that is, the method finds the optimal subnetwork. Two lessons can be extracted from this simple example. First, this simple method works better the lower the number of nonoptimal links. Second, it is clear that step 1 can be improved by using an iterative classification. Once neuron 10 is classified as the worst one, the remaining 9 are reclassified. From these 9 neurons the worst neuron is selected and put as second worst in the list. This reclassification process is continued until the last neuron. In this way optimal neurons linked to nonoptimal ones, like neuron 7, are not necessary classified as worse than the rest by the dissection method. In the rest of the paper we use this iterative step 1 unless stated otherwise.

We then applied the dissection method to artificial network configurations with the same connectivity as the *C. elegans* network to test its efficacy in separating optimal and nonoptimal subnetworks. We built network configurations with a given percentage of randomly selected nonoptimal neurons with positions taken from a uniform distribution of width given by the length of the animal. The rest of the neurons are optimal, that is, obtained from an optimization calculation including only these neurons. We applied to these network configurations the iterative step 1 to classify the neurons by position error and step 2 to calculate the average position error e_p in subnetworks of decreasing size by gradually eliminating the worst neurons in the list of step 1 (Fig. 4D, blue curve). The average position error e_p goes down to a value very close to zero (as seen from right to left) at a network size that separates the optimal and nonoptimal subnetworks. The dissection method was found to have an excellent performance in the estimation of the size of the nonoptimal network for nonoptimal subnetworks up to a size of 50% and deteriorates only slightly for larger nonoptimal networks (Fig. 4E, blue). The percentage of neurons correctly classified as nonoptimal (truly nonoptimal neurons in the estimated nonoptimal subnetwork) is also excellent with little misclassification until a size of 50% of the total network and a slight deterioration for increasing

size (Fig. 4E, green). Similar results are found for neurons in nonoptimal networks located randomly following a Gaussian instead of a uniform distribution (SI Fig. 7B). Crucial for the success of the method is the iterative step 1, without which the deterioration of the estimation already starts at a size of nonoptimal subnetwork of 25% (SI Fig. 8B).

Networks formed by a near-optimal subnetwork with a given low noise added to optimal positions and a nonoptimal subnetwork with larger noise are also dissected to a good approximation when the nonoptimal subnetwork has a size $< 50\%$ of the total network (SI Fig. 9B). The distinct feature of the average position error e_p in step 2 for this case is not a minimum but an abrupt change in the slope when reducing the size of the network. This point of abrupt slope change separates well the optimal from the nonoptimal subnetworks (SI Fig. 9A). We have implemented a robust numerical algorithm to find this separation point automatically (SI Text). Noisy networks, that is, networks in which neurons are randomly displaced from its optimal location but do not have optimal and nonoptimal subnetworks, also have lower average position error e_p when eliminating the worst located neurons. Interestingly, however, these noisy networks never show, whatever the noise level, a special point that could be used to separate optimal from nonoptimal subnetworks like a minimum or a change in slope (Fig. 4D, red line for uniform noise and SI Fig. 9A, red line for Gaussian noise). As a consequence, our algorithm classifies noisy networks either as fully optimal when the noise is very low or as fully nonoptimal for higher noise (Fig. 4F for uniform noise and SI Fig. 9C for Gaussian noise).

Dissection of the actual *C. elegans* network results in a nonoptimal subnetwork of 12% of neurons (34 of 279) with 20 interneurons, 8 sensory neurons and 6 motor neurons. The average position error e_p is lowered from 8.8% for the complete network to 5.4% for the near-optimal subnetwork (Fig. 5A, see SI Table 1 for a list of the nonoptimal neurons). More importantly, the clustering error e_c is lowered very significantly from 14% to 2.1% (Fig. 5B) so the clustering diagram of the optimal subnetwork (Fig. 5B, Large Inset) is now very similar to the experimental one down to fine details like the clustering in ganglia within the brain (Fig. 1D). Encephalization, for example, is explained by using wiring economy for the optimal subnetwork (SI Fig. 10). By using the optimal subnetwork and a

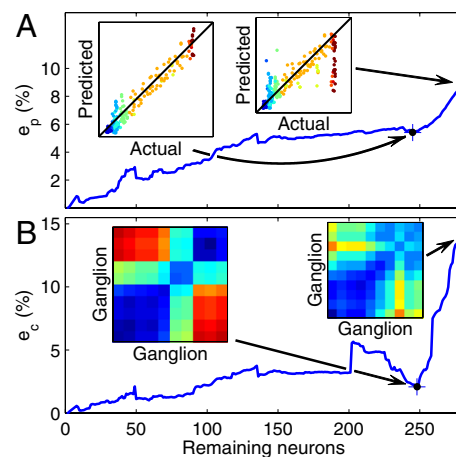


Fig. 5. Dissection of the actual *C. elegans* configuration into near-optimal and nonoptimal subnetworks. (A) Average position error e_p in networks of decreasing size (from right to left) eliminating worst located neurons in the order determined in step 1 of the method. Separation of optimal and nonoptimal subnetworks was found automatically at the point with a plus sign corresponding to a large slope change. Insets: Predicted versus actual position for the near-optimal subnetwork and for the complete network. (B) Same as A but for the clustering error. (Insets) Clustering diagram for near-optimal subnetwork (compare with actual one in Fig. 1D) and for total network.

simple model for synapse location at the middle point between somas, predicted location of synapses is found to be consistent with the actual nerve ring (SI Fig. 11). Note that the error of the near-optimal subnetwork in *C. elegans* does not go to zero, consistent with a network in which both the near-optimal and nonoptimal subnetworks have noise of different variances (SI Fig. 9D).

Wiring economy applied to the near-optimal subnetwork in *C. elegans* accounts for the clustering of neurons into ganglia as shown by the low clustering error of $e_c = 2.1\%$. The origin of this low value is 2-fold. First, the dissection method has eliminated badly located nonoptimal neurons thus reducing average position error. Second, in the optimization of the complete network, the nonoptimal subnetwork is not passive but has a negative effect on the near-optimal subnetwork. The clustering error of the near-optimal subnetwork in the complete network optimization is found to be 3% higher than alone. To further understand this effect, we have searched for the nonoptimal neurons responsible for increased position errors in the near-optimal neurons when both are forming the complete network. We removed one neuron at a time and calculated the change in position error in each of the near-optimal neurons (SI Fig. 12A). Neuron labels in the figure are located such that most damaging nonoptimal neurons are at the bottom and most affected near-optimal neurons to the left. The most damaging nonoptimal neurons are the interneurons AVA, DVA, PVC and AKL. The near-optimal neurons most affected by these nonoptimal neurons are PHC and PHB, but many more are affected and this effect is in part responsible for the low clustering error when eliminating the nonoptimal subnetwork. Nonoptimal neurons also affect the position error of other nonoptimal neurons, but the effects are smaller than on near-optimal neurons and in this case there is no clear core of most affecting neurons (SI Fig. 12B). Only DVA appears again as one of the more damaging neurons.

Discussion

The *C. elegans* nervous system was found to contain a subnetwork of 84% of neurons wired almost optimally. This subnetwork explains the origin of the structure of the whole system, in particular the segregation of neurons into ganglia and encephalization. A small subset of the connections of this large near-optimal subnetwork formed by the links of sensory and motor neurons to sensors and muscles is responsible for the main features of the neuroanatomy. Connections among neurons are responsible for further neuroanatomical detail, like the formation of ganglia within the head ganglia and the location of this head ganglia within the body. An analysis of nonoptimal neurons revealed that interneurons

AVA, DVA, PVC, and AKL have a negative effect on the position of near-optimal neurons in a complete network optimization. This result is consistent with the idea that these interneurons might be spatially constrained in the actual nematode. Interestingly, DVA is a stretch receptor neuron expressing TRP-4, the *elegans* homologue of the mechanosensitive TPRN channel (36). AVA and PVC express UNC-8, a DEG/EnaC family member homologue to subunits of a mechanically gated ion channel (37). Internal body sensors like those implicated in proprioception could be added to the theory in a form identical to external sensors when experimental data becomes available.

Materials and Methods

Data. Network connectivity and actual positions of somas, sensors, and muscles were taken from the revised data published in ref. 24, available at www.wormatlas.org.

Construction of Interganglia Distance Graphs. Interganglia distance graphs, or clustering graphs, as those in Fig. 1D and E, were built in the following way. Neurons were grouped by the ganglion they belong to in the real nematode. Then, distances between every neuron of ganglion i and every neuron of ganglion j are computed (or, if $i = j$, between every two neurons of the same ganglion). The average of these distances is the ij -th element of the interganglia distance graph. Note that these matrices are symmetrical by construction.

Construction of Interganglia Connectivity Graphs. Square (ij) of Fig. 1F represents the number of connections per neuron between ganglia i and j . It was computed by dividing the total number of connections between ganglia i and j by the sum of the number of neurons of the two ganglia. When $i = j$, it is computed as the total number of connections between neurons of the same ganglion divided by the number of neurons that form the ganglion. For Fig. 1G, the length of the nematode was divided into 10 bins of equal length. Square (ij) of this figure represents the number of connections between ganglion i and the organs that fall into the j -th bin, divided by the number of neurons in ganglion i .

We thank Brian Burton, Vicente Pérez-Díaz, and Adela Escudero-Berrián for critical comments on the manuscript. This work was supported by the Spanish Ministry of Science and Education (MEC) (G.G.d.P.) and Comunidad de Madrid (Biociencia program) (G.G.d.P.). A.P.-E. acknowledges fellowships from Consejo Superior de Investigaciones Científicas, MEC, and Instituto de Salud Carlos III.

1. Striedter GF (2004) *Principles of Brain Evolution* (Sinauer, Sunderland, MA).
2. Nusbaum MP, Blitz DM, Swensen AM, Wood D, Marder E (2001) *Trends Neurosci* 24:146–154.
3. Konishi M (1999) in *Elsevier's Encyclopedia of Neuroscience*, eds Adelman G, Smith BH (Elsevier, Amsterdam), pp 1906–1908.
4. Ramón y Cajal S (1899) *Textura del Sistema Nervioso del Hombre y de los Vertebrados* (Nicolás Moya, Madrid) (Spanish); trans Pasik P, Pasik T (1999) *Texture of the Nervous System of Man and the Vertebrates* (Springer, New York).
5. Mitchison G (1991) *Proc R Soc London B Biol Sci* 245:151–158.
6. Van Essen DC (1997) *Nature* 385:313–318.
7. Cowey A (1979) *Q J Exp Psychol* 31:1–17.
8. Nelson ME, Bower JM (1990) *Trends Neurosci* 13:403–408.
9. Chklovskii DB, Koulakov A (2000) *Physica A* 284:318–334.
10. Durbin R, Mitchison G (1990) *Nature* 343:644–647.
11. Koulakov AA, Chklovskii DB (2001) *Neuron* 29:519–527.
12. Murre JM, Sturdy DP (1995) *Biol Cybern* 73:529–545.
13. Ruppin E, Schwartz EL, Yeshurun Y (1993) *Biol Cybern* 70:89–94.
14. Wen Q, Chklovskii DB (2005) *PLoS Comput Biol* 1:e78.
15. Cherniak C (1994) *J Neurosci* 18:522–527.
16. Cherniak C (1995) *Trends Neurosci* 14:2418–2427.
17. Klyachko VA, Stevens CF (2003) *Proc Natl Acad Sci USA* 100:7937–7941.
18. Chklovskii DB (2004) *Neural Comput* 16:2067–2078.
19. Cherniak C, Changizi M, Won Kang D (1999) *Phys Rev E* 59:6001–6009.
20. Chklovskii DB (2000) *J Neurophysiol* 83:2113–2119.
21. Chklovskii DB, Stepanyants A (2003) *BMC Neuroscience* 4:18.
22. Cherniak C (1992) *Biol Cybern* 66:503–510.
23. Shefi O, Harel A, Chklovskii DB, Ben-Jacob E, Ayali A (2003) *Neurocomputing* 58–60:487–495.
24. Chen BL, Hall DH, Chklovskii DB (2006) *Proc Natl Acad Sci USA* 103:4723–4728.
25. White JG, Southgate E, Thomson JN, Brenner S (1986) *Philos Trans R Soc London Ser B* 314:1–340.
26. Hall DH, Russell RL (1991) *J Neurosci* 11:1–22.
27. Hall K (1970) *Management Sci* 17:219–229.
28. Kleinhaus J, Sigl G, Johannes F, Antreich K (1991) *IEEE Trans CAD* 10:356–365.
29. Ahn, Y.-Y, Jeong H, Kim BJ (2006) *Physica A* 367:531–537.
30. Kaiser M, Hilgetag CC (2006) *PLoS Comput Biol* 2:e95.
31. Rushton WA (1951) *J Physiol (London)* 115:101–122.
32. Rall W, Burke RE, Holmes WR, Jack JJ, Redman SJ, Segev I (1992) *Physiol Rev* 72:5159–5186.
33. Laughlin SB, de Ruyter van Steveninck R, Anderson JC (1998) *Nat Neurosci* 1:36–41.
34. Attwell D, Laughlin SB (2001) *J Cereb Blood Flow Metab* 21:1133–1145.
35. Tessier-Lavigne M, Goodman CS (1996) *Science* 274:1123–1133.
36. Li W, Feng Z, Sternberg PW, Xu XZ (2006) *Nature* 440:684–687.
37. Tavernarakis N, Shreffler W, Wang S, Driscoll M (1997) *Neuron* 18:107–119.