The distribution of inhibitory neurons in the *C. elegans* connectome facilitates self-optimization of coordinated neural activity

Alejandro Morales  
Institute of Economic Research  
National Autonomous University of Mexico  
Mexico City, Mexico  
alejandroe@ciencias.unam.mx

Tom Froese  
Embodied Cognitive Science Unit  
Okinawa Institute of Science and Technology Graduate University  
Okinawa, Japan  
tom.froese@oist.jp

**Abstract**—The nervous system of the nematode soil worm *Caenorhabditis elegans* exhibits remarkable complexity despite the worm’s small size. A general challenge is to better understand the relationship between neural organization and neural activity at the system level, including the functional roles of inhibitory connections. Here we implemented an abstract simulation model of the *C. elegans* connectome that approximates the neurotransmitter identity of each neuron, and we explored the functional role of these physiological differences for neural activity. In particular, we created a Hopfield neural network in which all of the worm’s neurons characterized by inhibitory neurotransmitters are assigned inhibitory outgoing connections. Then, we created a control condition in which the same number of inhibitory connections are arbitrarily distributed across the network. A comparison of these two conditions revealed that the biological distribution of inhibitory connections facilitates the self-optimization of coordinated neural activity compared with an arbitrary distribution of inhibitory connections.

**Index Terms**—artificial neural networks, self-organization, Hopfield networks, artificial life, complex adaptive systems, self-modeling, *C. elegans*, computational neuroscience

I. INTRODUCTION

The brain has been called the most complex object in the universe. It is therefore understandable that we are still lacking a unified theory of how the brain works. The various mechanisms underlying its capacity to generate ordered patterns of neural activity in a spontaneous manner are still being unravelled. Unlike artificially designed systems, like the digital computer, the brain is organized in a hierarchical manner and neural activity is not governed by a central controller. This inherent complexity poses a challenge for rapid progress for brain science, but at least theoretical neuroscience can benefit from general insights coming from recent advances in the theory of complex adaptive systems.

A relatively unknown but promising approach is the study of self-optimization using an algorithm first proposed by Watson et al. [1] on the basis of the Hopfield neural network, and which was subsequently developed in various directions. The core idea is that a Hopfield network can be made to form an associative memory of its own state attractors, namely by repeatedly letting it minimize neural constraints posed by the network architecture and reinforcing those attractors on which it converges. The result is that the network starts to generalize over its local attractors and thereby enhances its capacity to converge on more optimal attractors. In contrast to much other contemporary work in training artificial neural networks, this self-optimization process does not rely on any prior knowledge of the target domain; it is based on a simple form of unsupervised learning (e.g. Hebbian learning). Previous work has successfully applied this algorithm to various kinds of neural network architectures; for example, in networks with symmetric weights [2], [3], networks with continuous activation functions [4], [5], and spiking neural networks [6].

However, in all previous models of self-optimization the neural networks were relatively small (e.g. around 100 neurons) and had a highly artificial topology (e.g. arbitrary initialization of connection weights, arbitrary distribution of negative connections, or perfectly modular arrangement of connections). This leaves it unclear whether the process would successfully scale up to the complexity of biological neural network architectures. In other words, if one of the aims of this particular model of a complex adaptive system is to provide inspiration for theoretical neuroscience, then it would be more meaningful to test this algorithm with a biologically plausible neural network architecture.

Accordingly, as a first step in this direction, we proposed that the *C. elegans* connectome is a suitable starting point [7]. *Caenorhabditis elegans* is a one-millimeter-long soil worm of the nematode family. It is the first animal from which the complete mapping of synaptic connections has been obtained and it is a reference animal model in biology [8], [9]. The neural network dynamics of *C. elegans* have been modeled in several ways [10], and the artificial life research community has also shown an interest in the study of *C. elegans* [11]–[15]. However, it has focused on small circuits that are embodied in a sensorimotor loop, without taking into account the whole connectome.
We took an alternative approach and decided to study the whole connectome as a complex adaptive system. We recently demonstrated that the self-optimization algorithm can be applied to the C. elegans connectome when it is viewed as a directed multigraph [7]. Some previous work on self-optimization employed networks without negative connections, so as a first step we explored the effect of including negative connections in the connectome. In that work we compared two conditions: (1) The self-optimization algorithm was run with only excitatory (positive) connections, and (2) it was run with 30% inhibitory (negative) connections that were arbitrarily assigned at the beginning of the algorithm. We found that the C. elegans connectome consistently showed a tendency to optimize its own connectivity such that neural coordination become more prominent, but the presence of negative connections tended to increase the difficulty of neural coordination.

Given that this difficulty may result from the arbitrariness of the distribution of the negative connections, we next explored self-optimization in the C. elegans connectome by separating it into hierarchically organized functional clusters that are known from the literature [16], [17]. In that work, self-optimization was tested in two scenarios: (1) In each cluster independently from the rest of the connectome, and (2) in each cluster but embedded within the context of the whole connectome. Inhibitory connections were introduced arbitrarily in (1) and distributed in inter-cluster connections in (2). Again, the C. elegans connectome consistently showed a tendency to optimize its own connectivity in both cases, but there was a significant improvement of self-optimization in (2). This finding revealed that it was not necessarily the presence of inhibitory connections per se that makes self-optimization less effective, but rather the precise distribution of the inhibitory connections and how they carve up the overall connectome into distinct clusters.

In this paper, we continued exploring the conditions for self-optimization in the C. elegans connectome by including more biological details in our model. In particular, we took into account the biological distribution of inhibitory connections as specified by the neurotransmitter identity of each of the neurons in the connectome [18], [19] (scenario A), and we compared this with a control condition in which the same number of inhibitory connections are distributed arbitrarily across the network (scenario B). Based on our previous work, we hypothesized that the biological distribution provides a more effective starting point for the self-optimization of connectivity, consistently leading to improved neural coordination. We were able to confirm this hypothesis.

But we also found self-optimization to be generally less effective, at least compared to previous variations of the model in which we had initialized the connectome as a fully interconnected network by padding its biologically existing connections with initially non-functional (i.e. zero-weighted) connections that are biologically non-existent. We speculate about the reasons for this difference at the end of the paper.

Fig. 1. Visualization of the C. elegans connectome in a circular layout with a bundling procedure. Black represents excitatory connections; red represents inhibitory connections. Scenario A: Connectome with 26% inhibitory connections distributed in accordance with each neuron’s neurotransmitter identity. Scenario B: Connectome with 26% inhibitory connections but with an arbitrary distribution. By visual inspection we can observe a greater number of aggregated connections in scenario A, especially more clustering of excitatory connections.

II. METHODS
A. The connectome
We followed the same approach as in our previous work [7], [17] regarding the connectome’s translation to a directed multigraph, with neurons as nodes and connections as edges. This approach enabled us to run the self-optimization algorithm on
a model of the connectome published by Jarrell et al. [20]. That connectome database contains hermaphrodite and male neural information such as connection direction, type of connection (synapse or gap junction), and the number of connections between neurons. We only considered hermaphroditic information because *C. elegans* males, which occur at a low frequency in laboratory strains [21], differ slightly in the number, identity, and synaptic wiring in their nervous system [22].

*C. elegans* consists of only 959 cells, of which 302 are neurons belonging to the nervous system. We took into account 280 neurons, with 5609 connections. We did not consider the 20 pharyngeal neurons because they are treated as belonging to an independent neural system [23]. We also removed the neurons *CANR* and *CANL* because they do not have obvious connections and are classified as end organs [24].

We modeled chemical synapses as single-directed links between neurons (for example, $A \rightarrow B$ indicates that neuron $A$ is presynaptic to neuron $B$, and $B$ is postsynaptic to $A$). Gap junctions are modeled as double-linked neurons (if two neurons, $C$ and $D$, have a gap junction between them, then there are two links: $C \rightarrow D$ and $D \rightarrow C$). The number of connections between neurons was translated into the weight of each edge, normalized in the interval (0, 1). Both edges representing a gap junction were assigned the same weight. Weight values vary between 1 and 81 before normalization. There were only 15 edges whose weight values exceeded 44, and we decided to treat them as outliers and clip them to 1 during normalization. This was done to prevent the majority of connections to become comparatively too small in weight. The effect of clipping these outliers is to increase the scope of state-space explorations of the resulting network, allowing it to visit a more diverse set of attractors. Future work could develop a more elegant way of translating the connectome information into edge weights.

### B. Inhibitory neurons

We categorized all outgoing connections as either excitatory or inhibitory based on a neuron’s neurotransmitter identity [18], [19]. In other words, when a neuron’s identity is determined to be inhibitory, we made the signs of the weights of its outgoing connections negative. According to [18], *C. elegans* neurons can be partitioned into one of four groups based on its neurotransmitter identity, with the following counts:

(a) 147 excitatory neurons  
(b) 26 inhibitory neurons  
(c) 83 simultaneously excitatory and inhibitory neurons  
(d) 24 neurons not characterized (their neurotransmitter identities are unclear)

Excitatory neurons are mostly dopaminergic neurons, while inhibitory neurons are GABA-aminergic. Neurons in group (c) have inhibitory and excitatory neurotransmitters simultaneously (glutamate, serotonin, octopamine, tyramine) [25], [26]. Molecular analysis of *C. elegans* neuron signaling also includes neuropeptide function, receptors and related genes [9], [19], but we are not taking into account these molecules in our current model. Regarding neuron function, most excitatory and inhibitory *C. elegans* neurons are motor neurons. Neurons that have excitatory and inhibitory connections simultaneously are usually mechanosensory neurons and interneurons [24].

We assigned positive and negative outgoing connections to the neurons in group (a) and (b), respectively. Then, for groups (c) and (d) we arbitrarily selected 50% of the neurons as having all of their outgoing connections negative. Under these conditions, we obtained around 26% inhibitory connections in the whole network: a percentage similar to the one used in our previous studies [7], [17] and close to the theoretical approximation in literature [27]–[30].

In Fig. 1, we depict the resulting Hopfield network structure in a circular layout, in which black and red edges indicate excitatory and inhibitory connections, respectively. We applied a bundling algorithm to avoid visual clutter, and to give an indication of possible clustering of connections.

### C. The connectome as a complete graph

In previous work we had added zero-weighted connections to the Hopfield neural network of the connectome in order to turn it into a complete directed graph [7], [17]. In this way, Hebbian learning could include extra functional connections in the network by changing these weights from zero to non-zero values, and we had found that this facilitates the removal of constraints between interactions during the self-optimization process. In the current work, we first tested the restricted network based on the biological connectome of 5,609 edges in a first set of simulation runs. Then we conducted a second set of simulation runs in which we expanded the network into a complete graph by adding extra zero-weighted edges, which resulted in a network with 79,089 edges.

The extra edges do not have anatomical equivalents in the adult connectome, but we could consider that young worms have an expanded connectome that gets reduced in size during maturation due to pruning of neuronal synapses and axons [22], [31]. Future work could explore what happens if we approximate development by slowly pruning the complete graph toward the biological connectome during self-optimization.

### D. Model dynamics

Following the Hopfield network formalism, neurons have binary activation states ($-1, 1$). Asynchronous state updates are based on the following equation:

$$s_i(t+1) = \theta \left( \sum_{j} N \left( \sum_{k} w_{ijk} s_j(t) \right) \right)$$

where $s_i$ is the state of neuron $i$, $w_{ijk}$ is the connection weight between neuron $i$ and neuron $j$ with index $k$ (more than one connection with the same direction could arise between $i$ and $j$), and $\theta$ is a Heaviside threshold function that returns $-1$ and 1 for negative and positive arguments, respectively. In a Hopfield network, a node $i$ satisfies a constraint in its interaction with node $j$ with connection index $k$ if $s_i s_j w_{ijk} > 0$. System
energy \( E \) represents the total constraint satisfaction achieved by the network:

\[
E = -\sum_{i,j,k} w_{ijk}^{O}(t)s_i(t)s_j(t) \tag{2}
\]

where \( w_{ijk}^{O} \) is the original weight configuration of \( w_{ijk} \) defined at the beginning of the algorithm. In other words, if the weights of the connections between two neurons are such that they result in excitation (inhibition), the contribution towards \( E \) will be negative (positive) when the states of the two neurons coincide (are different). When the self-optimization process is ongoing, Hebbian learning is applied to all system connections (i.e., change in weight, \( \Delta w_{ij} = \delta s_is_j, \delta > 0 \)), but technically these weight changes are stored separately in another weight configuration. Keeping track of the accumulating weight changes separately in this manner allows us to plug in the state configurations produced by the learning network into \( w_{ijk}^{O} \) to test how well these states satisfy the constraints posed by the original connectome. We fixed the learning rate \( (\delta) \) to 0.0001 for all simulation runs.

The self-optimization algorithm consists of repeating the following sequence of steps:

1) Arbitrary assignment of states for the neurons (reset).
2) Asynchronous updates of the neurons, typically for sufficient iterations such that activity stabilizes (convergence).
3) Application of Hebbian learning.

Each repetition of steps 1) and 2) is called a reset-convergence cycle. The number of neuron updates performed during step 2) is set to 18,000. We observed this to be an adequate quantity to reach stability at the end of each reset-convergence cycle. For purposes of comparison, we also measured the performance of the reset-convergence cycles before and after self-optimization (i.e., in the absence of Hebbian learning).

### III. RESULTS

We first explored the reset-convergence dynamics of the Hopfield network of the connectome without Hebbian learning. Fig. 2 shows network energy after successive neuron state updates starting from arbitrary initial states. Energy tends to decrease, but the network does not necessarily reach a fixed-point attractor because the presence of inhibitory synapses makes neural coordination more difficult, as we had already previously observed \[7, 17\]. We noted a tendency for the biological distribution of inhibitory connections to facilitate convergence on more coordinated states.

We then investigated this network’s capacity for self-optimization more systematically. First, we performed 1000 reset-convergence cycles without Hebbian learning. Then, we applied self-optimization using 1000 reset-convergence cycles. Finally, we applied to the network 1000 reset-convergence cycles without Hebbian learning using configuration obtained by self-optimization to show its stability. Cycles 1001-2000 follow steps 1, 2, and 3 of the self-optimization algorithm. Cycles 1-1000 and 2001-3000 only follow steps 1 and 2 because they are not learning phases. Results are averaged over 10 different experiments with a different initial random number seed. We performed this set of simulation runs using the network with the biological distribution of inhibitory connections (scenario A), and repeated the simulation runs using the network with an arbitrary distribution of inhibitory connections (scenario B).

The results are shown in Fig. 3, which reveals consistently improved neural coordination in scenario A compared to scenario B. In addition, we find that Hebbian learning manages to improve the performance of both networks, thus demonstrating that self-optimization can be applied to them. However, the resulting improvement in neural coordination following self-optimization is not as pronounced as in our previous work. Therefore, we also tested these networks as complete graphs, which indeed significantly improves self-optimization, as shown in Fig. 4.

### IV. DISCUSSION

We have tested self-optimization in a Hopfield network based on the \textit{C. elegans} connectome. Here we suggest some possible implications of our work for theoretical neuroscience.
Fig. 3. Self-optimization of neural activity in two scenarios. The plot shows a measure of neural state coordination, i.e. the total energy of a Hopfield neural network, after convergence from random initial states, across three phases for ease of comparison: 1) before learning (1–1,000), 2) during the self-optimization process based on Hebbian learning (1,001–2,000), and 3) after learning (2,001–3,000). Scenario A (red): The neural network is initialized with inhibitory connections in accordance with the connectome’s inhibitory neurons. This means that 26% of connections are initialized with negative weights. Scenario B (blue): Control condition in which the network is initialized with 26% negative connections, but their location is distributed arbitrarily across the network. Each scenario was tested with 10 independent runs. Both scenarios show evidence of self-optimization in terms of spontaneous convergence on lower energy values following application of Hebbian learning. However, when inhibitory connections are distributed in accordance with the connectome’s inhibitory neurons, the network tends to consistently find lower energy values, which indicates that more neuron states are coordinated.

Fig. 4. Self-optimization of coordinated neural activity in two scenarios involving excess connections. The plot shows a measure of neural state coordination, i.e. the total energy of a Hopfield neural network, after convergence from random initial states, across three phases for ease of comparison: 1) before learning (1–1,000), 2) during the self-optimization process based on Hebbian learning (1,001–2,000), and 3) after learning (2,001–3,000). Scenario A (red): The neural network is initialized with inhibitory connections in accordance with the connectome’s inhibitory neurons, as in Fig. 3, but this time the network is also padded with excess connections of zero weight that turn it into a potentially fully connected network. Scenario B (blue): Control condition in which the network is initialized with 26% negative connections with an arbitrary distribution across the network, as in Fig. 3, but this time the network is also padded with excess connections of zero weight that turn it into a potentially fully connected network. Each scenario was tested with 10 independent runs. Both scenarios show evidence of self-optimization in terms of spontaneous convergence on lower energy values following application of Hebbian learning. However, when inhibitory connections are distributed in accordance with the connectome’s inhibitory neurons, the network tends to consistently find lower energy values, which indicates that more neuron states are coordinated. The inclusion of excess connections made self-optimization significantly more effective.

Of course, we acknowledge that our model is highly abstract, and that the link with actual brain function is not direct; so we must be cautious. Nevertheless, there are some interesting considerations that can be made. That the network starts with a large amount of different initial attractors can be interpreted in terms of the nervous system being initially pre-configured with a large number of different cell assembly configurations, which then get optimized according to lifetime demands placed on them [32]. The reset-convergence cycle could be interpreted in terms of biological rhythms that modulate neural activity, similar to sleep-wake cycles [6].

We showed that by assigning inhibitory connections according to neuron identity, the network tends to converge on lower energy values from arbitrary initial states. This implies that the distribution of inhibitory neurons in the connectome plays a role in facilitating neural coordination. We also demonstrated that when we repeat such reset-convergence cycles in the presence of accumulating weight changes due to Hebbian learning...
there is a consistent further decrease in energy values. In other words, self-optimization additionally facilitates coordination among neurons. However, we do not find that the improvement generated by self-optimization is significantly facilitated by a realistic distribution of inhibitory connections; the relative decrease in energy values is comparable across the scenarios based on biological and arbitrary distributions.

Adding extra zero-weighted connections to the initial networks notably improves the performance of self-optimization, which then tends to focus the state space of the networks to the most coordinated state configurations. This difference in performance deserves more attention in future work, and the biological equivalent of the excess connections remains to be determined. We mentioned the possibility that an expanded network could correspond to the connectome of an immature worm before the pruning of synapses and axons. Another possibility is that these connections could model potential worm before the pruning of synapses and axons. Another network could correspond to the connectome of an immature network architectures [34].

This improvement of biological realism of our simulation model could go hand in hand with the implementation of more realistic neural dynamics, for instance by using continuous-time, continuous-state neurons like in [5]. Finally, we would draw on the literature on chemical signaling in the worm’s nervous system [33], and combine it with relevant artificial neural network architectures [34].

REFERENCES
[1] R. A. Watson, R. Mills, and C. L. Buckley, “Global adaptation in networks of selfish components: Emergent associative memory at the system scale,” Artificial Life, vol. 17, no. 3, pp. 147–166, 2011.
[2] R. A. Watson, C. L. Buckely, and R. Mills, “Optimization in “self-modeling” complex adaptive systems,” Complexity, vol. 16, no. 5, pp. 17–26, 2011.
[3] R. A. Watson, R. Mills, and C. L. Buckley, “Transformations in the scale of behavior and the global optimization of constraints in adaptive networks,” Adaptive Behavior, vol. 19, no. 4, pp. 227–249, 2011.
[4] M. Zarco and T. Froese, “Self-modeling in Hopfield neural networks with continuous activation function,” Procedia Computer Science, vol. 123, pp. 573–578, 2018.
[5] ———, “Self-optimization in continuous-time recurrent neural networks,” Frontiers in Robotics and AI, vol. 5, p. 96, 2018.
[6] A. Woodward, T. Froese, and T. Ikekagi, “Neural coordination can be enhanced by occasional interruption of normal firing patterns: A self-modeling complex neural network model,” Neural Networks, vol. 62, pp. 39 – 46, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0893608014002068
[7] A. Morales and T. Froese, “Self-optimization in a hopfield neural network based on the C. elegans connectome,” in The 2019 Conference on Artificial Life, H. Fellerman, J. Barcardt, A. Goñi-Moreno, and R. Fuchsln, Eds., MIT Press, 2019, pp. 448–453.
[8] J. G. White, E. Southgate, J. N. Thomson, and S. Brenner, “The structure of the nervous system of the nematode Caenorhabditis elegans,” Philosophical Transactions of the Royal Society, Biological Sciences, vol. 279, no. 938, pp. 299–325, 1976.
[9] L. R. Girard et al., “Wormbook: the online review of Caenorhabditis elegans biology,” Nucleic Acids Research, vol. 35, pp. D472–D475, 2006, doi:10.1093/nar/gkl894.
[33] C. Li, L. S. Nelson, K. Kim, A. Nathoo, and A. C. Hart, “Neuropeptide gene families in the nematode Caenorhabditis elegans,” *Annals of the New York Academy of Sciences*, vol. 897, no. 1, pp. 239–252, 1999.

[34] P. Husbands, A. Philippides, P. Vargas, C. L. Buckley, P. Fine, E. Di Paolo, and M. O’Shea, “Spatial, temporal, and modulatory factors affecting GasNet evolvability in a visually guided robotics task,” *Complexity*, vol. 16, no. 2, pp. 35–44, 2010.

[35] E. Izquierdo and T. Bührmann, “Analysis of a dynamical recurrent neural network evolved for two qualitatively different tasks: Walking and chemotaxis,” in *ALIFE*, 2008, pp. 257–264.