Neurocomputing 427 (2021) 238–261

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Multilayer network analysis of *C. elegans*: Looking into the locomotory circuitry

Thomas Maertens [a], Eckehard Schöll [a,b], Jorge Ruiz [a,b], Philipp Hövel [a,b,c,*,1]

[a] Institut für Theoretische Physik, Technische Universität Berlin, Hardenbergstraße 36, 10623 Berlin, Germany
[b] Bernstein Center for Computational Neuroscience Berlin, Humboldt-Universität zu Berlin, Berlin, Germany
[c] School of Mathematical Sciences, University College Cork, Western Road, Cork T12 XF64, Ireland

ARTICLE INFO

Article history:
Received 15 June 2020
Revised 18 September 2020
Accepted 10 November 2020
Available online 27 November 2020
Communicated by Zidong Wang

Keywords:
Connectome of *C. elegans*
Multilayer networks
Neuronal dynamics
Central pattern generators
Motion behavior
Harmonic waves
Synchronization
Feedback control
Hindmarsh-Rose model

ABSTRACT

We investigate how locomotory behavior is generated in the brain focusing on the paradigmatic connectome of nematode *Caenorhabditis elegans* (*C. elegans*) and on neuronal and muscular activity patterns that control forward locomotion. We map the neuronal network of the worm as a multilayer network that takes into account various neurotransmitters and neuropeptides. Using logistic regression analysis, we predict the neurons of the locomotory subnetwork. Combining Hindmarsh-Rose equations for neuronal activity with a leaky integrator model for muscular activity, we study the dynamics within this subnetwork and predict the forward locomotion of the worm using a harmonic wave model. The application of time-delayed feedback control reveals synchronization patterns that contribute to a coordinated locomotion of *C. elegans*. Analyzing the synchronicity when the activity of certain neurons is silenced informs us about their significance for a coordinated locomotory behavior. Since the information processing is the same in humans and *C. elegans*, the study of the locomotory circuitry provides new insights for understanding how the brain generates motion behavior.

1. Introduction

The human brain is a complex neuronal network with hundreds of billions of neurons arranged in a multitude of interconnected circuits. A major goal of neuroscience research is to understand how mental processes and behavior are controlled by the brain [30, Chapter 1]. In order to understand how behavior is generated, it is essential to understand the structure and the function of the individual circuits, along with the underlying patterns of neuronal activity [10,13]. Therefore, it is necessary to identify the neurons and their connections within each circuit. In view of the complexity of the brain, simpler models need to be considered.

The nematode *C. elegans* is an important animal model for almost all areas of experimental biology. The tiny creature provides a complete description of a development lineage, a nervous system, and a genome [26]. Although the genome of the worm is very simple, it has remarkable similarities to the human genome [44]. The neurons of *C. elegans* and humans are almost identical [36]. The nerve ring located at the head region of the worm is an equivalent counterpart to the human central nervous system [15]. Compared to the complexity of the brain, the nervous system of the adult *C. elegans* hermaphrodite can be treated more easily since it consists of only 302 neurons fixed. It serves as a prototype for investigations of neuronal networks because all synaptic connections between the neurons have been completely mapped by electron microscopy with reasonable accuracy at the cellular level [1,5,13,16,78,80].

Locomotion is the most important behavior of *C. elegans* and is made possible through 95 body wall muscles [2]. For example, the worm must be able to move in order to search for food, conspecifics, or improved conditions [26] but also to react to certain environmental influences, such as a gentle body touch. The *C. elegans* nervous system includes three main types of neurons: sensory neurons, interneurons, and motor neurons. In case of a body touch, information processing starts with sensory neurons that transmit information to interneurons. The latter in turn stimulate motor neurons that stimulate muscle cells so that the worm reacts

* Corresponding author.
E-mail addresses: thomas.maertens@gmx.net (T. Maertens), schoell@physik.tu-berlin.de (E. Schöll), jorge.ruiz@bccn-berlin.de (J. Ruiz), philipp.hoevel@ucc.ie (P. Hövel).
URLs: http://www.itp.tu-berlin.de/schoell (E. Schöll), http://publish.ucc.ie/researchprofiles/D019/philipphoevel (P. Hövel).
[1] ORCID: 0000-0002-1370-4272.

https://doi.org/10.1016/j.neucom.2020.11.015
0925-2312/

with locomotion [19, Section II]. In general, information processing of external influences in humans does not function differently even if the human nervous system is much more elaborate. For example, humans can also perceive touches via receptors in their skin [54] and react to them with movements of corresponding body parts. Therefore, studying the locomotory subnetwork of *C. elegans* could provide new insights for understanding how the brain generates motion behavior.

In this study, we map the somatic nervous system of *C. elegans* as a multilayer network including neurons and body wall muscles. For a better understanding of the properties and processes within the network, the underlying neurotransmitters and neuropeptides in chemical synapses are taken into account. These represent different modes of interactions between the nodes and define the multilayer network [5]. This makes it possible to apply logistic regression models to the network in order to identify the locomotory subnetwork that comprises all neurons involved in locomotion behavior (cf. Section 2). Subsequently, we focus the patterns of neuronal and muscular activity for a circuit of the locomotory subnetwork that initiates forward and backward locomotion in response to a touch stimulus on the head or tail. In order to describe the locomotion of the worm, we implement a harmonic wave model which is based on muscular activity. The peculiarity of our model is that the body posture of *C. elegans* can be predicted by the activity of each muscle pair consisting of a dorsal and a ventral muscle (cf. Section 3). This allows us to identify synchronization patterns that contribute to a coordinated locomotion behavior of *C. elegans* using time-delayed feedback control and to perform synchronization analyses from which conclusions can be drawn about the significance of neurons within the locomotory circuitry (cf. Section 4).

2. Multilayer network analysis

In this section, we introduce a multilayer network of *C. elegans* and based on this we predict the locomotory subnetwork using logistic regression.

2.1. Construction of the multilayer network

In order to create the network, we make use of several datasets (see Table 1). For the neurons, the information of class affiliation and neuron type is taken into account (ID 1). Neuronal connectivity describes how neurons are connected to each other, either through chemical synapses or gap junctions (ID 2). The neurons in the network are extended by 95 body wall muscle cells as these are essential for analysis of locomotion. Motor neurons act on muscle cells through neuromuscular junctions (ID 3). To better understand the interactions of chemical synapses, they are separated into various neurotransmitters and neuropeptides. This requires transmitter and receptor information for the individual neurons and information about which transmitters and receptors couple together (ID 4). By considering different neurotransmitters and neuropeptides, the network receives many extra layers that provide a deeper knowledge of the *C. elegans* connectome and biological neuronal networks in general.

Table 1
Datasets with network specific information. The neuron and connectivity data (ID 1 – ID 3) can be found on WormAtlas [82]. The transmitter and receptor data (ID 4) are partly provided by [5] and have been supplemented by own work.

ID	Dataset	Content
1	Neurons	Name, class, and neuron type
2	Neuronal connectivity	Synapse type with number of synapses between presynaptic and postsynaptic partners
3	Neuromuscular connectivity	Number of synapses between neurons and body wall muscle cells
4	Transmitters and receptors	Neurotransmitter, neuropeptide, and neuroreceptor information for the neurons

The mapped network consists of 279 neurons and 95 muscle cells (Fig. 1a). The neurons can be categorized as sensory neurons, interneurons, and motor neuron. Although the neurons can be multifunctional, only one type is highlighted in the figure (see also Table A.2). In total, there are 3,538 distinct directed connections between the nodes formed by electrical, chemical, and neuromuscular synapses (see also Table A.1). Note that the number of synapses is recorded for all connections in the network but is not taken into account here. In order to assign the number of synapses between neuron pairs with multiple neurotransmitters and neuropeptides, we assume that these are co-released [20].

Since the chemical links are represented by corresponding transmitter types, the multilayer network is defined by all these modes of interactions which are depicted in Fig. 1b-h: Acetylcholine (ACh), glutamate (Glu), *gamma*-aminobutyric acid (GABA), monoamine (MA), peptide, electrical, and neuromuscular connections, respectively. ACh covers 33.4% of the 3,538 connections in the network and forms the largest layer. It is widely spread among all neurons and does not prefer a specific neuron type. Glu uses 20.5% of the connections. In comparison with the ACh layer, significantly fewer motor neurons are involved. GABA represents the smallest network layer and is released on 3.8% of the connections. Most of them are established by interneurons and motor neurons. Thereafter, the MA transmitters follow with 5.8%. These connections are independent of the neuron type, but most of the motor neurons are predominantly postsynaptic. In the group of MAs, dopamine accounts for about two thirds (cf. Fig. A.1b). The peptide layer utilizes 11.8% of the connections in the network, which primarily exist between interneurons and sensory neurons. About 31% of the peptides are peptide transmitters and about 69% are cotransmitters (cf. Fig. A.1c, and d). Electrical transmission accounts for 29.1% of the connections and is therefore the second largest layer after ACh. 15.5% of the connections employ neuromuscular junctions, which connect motor neurons and body wall muscles. Background information on the different layers is summarized in Appendix B, and the data preparation process is detailed in Appendix A.

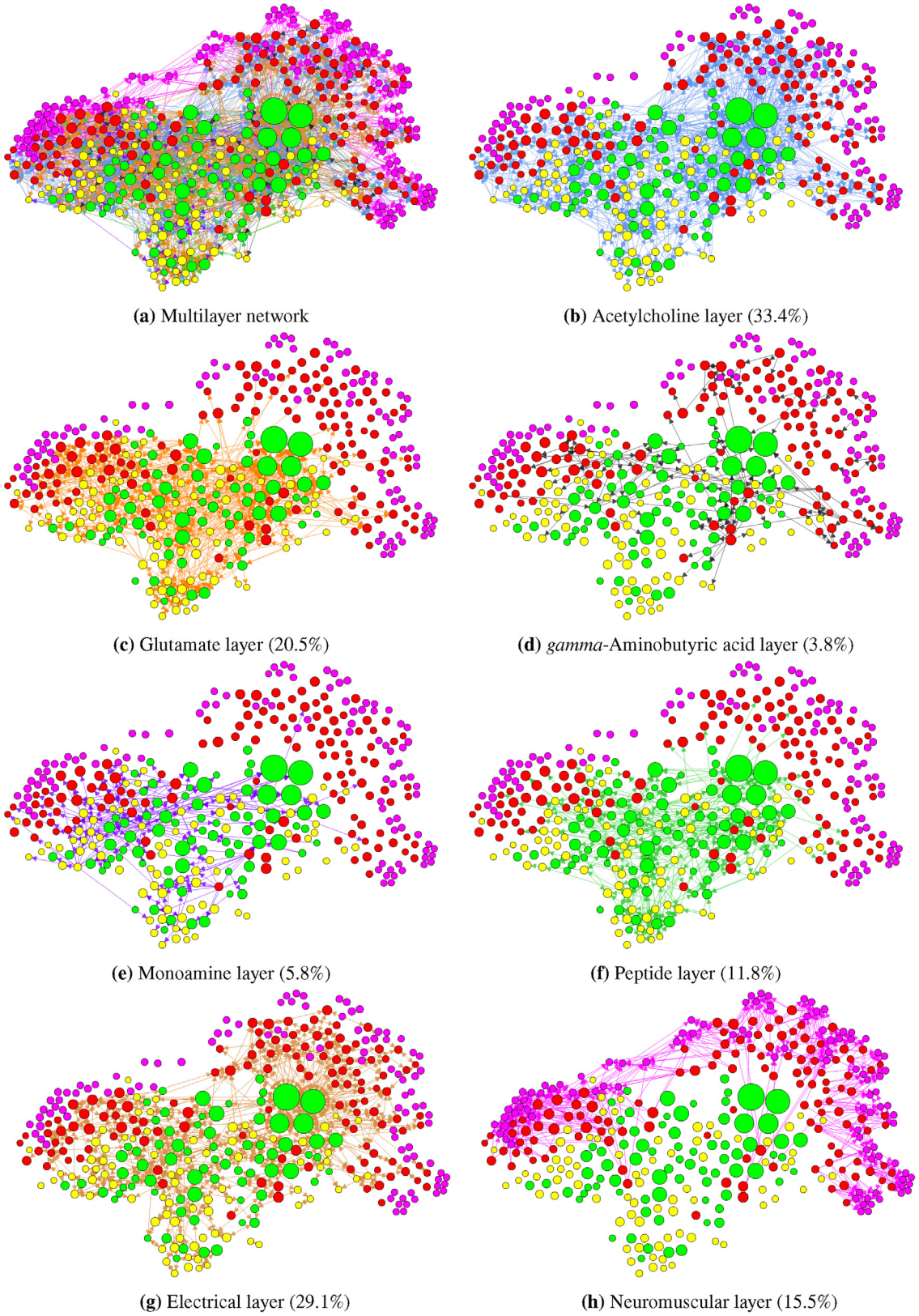
The transmitters utilized in the network allow for logistic regression analysis to predict the neurons involved in locomotion behavior, as detailed in the following.

2.2. Logistic regression analysis

Binary model – Logistic regression is a classification algorithm which can be used to predict a binary response variable $Y_i \in \{0,1\}$ based on a set of independent explanatory variables $x_1,\dots,x_m$. Since the expectation value of the response variable $\mathbb{E}(Y_i)$ lies in the interval $[0,1]$, the binary model applies the sigmoid function as response function

$$\sigma_i(z_i(\boldsymbol{c})) = \frac{1}{1+e^{-z_i(\boldsymbol{c})}}, \quad i = 1,2,\dots,n, \tag{1}$$

where i is the observation index, $z_i(\boldsymbol{c}) = c_0 + c_1x_{1i} + c_2x_{2i} + \dots + c_mx_{mi}$ is the linear combination of m explanatory variables $x_{1i}, x_{2i}, \dots, x_{mi}$ with $m+1$ parameters $\boldsymbol{c} = (c_0, c_1, c_2, \dots, c_m)$. The outcome for observation i can be interpreted as the probability that the response variable Y_i is equal to one, $\sigma_i(z_i(\boldsymbol{c})) = \mathbb{E}(Y_i) = \mathbb{P}(Y_i = 1)$. The goal of logistic regression analysis is to find parameters that best fit empirical observed data in Eq. (1). This can be achieved using the maximum likelihood estimation [59, Chapter 4].



(a) Multilayer network

(b) Acetylcholine layer (33.4%)

(c) Glutamate layer (20.5%)

(d) *gamma*-Aminobutyric acid layer (3.8%)

(e) Monoamine layer (5.8%)

(f) Peptide layer (11.8%)

(g) Electrical layer (29.1%)

(h) Neuromuscular layer (15.5%)

Fig. 1. The *C. elegans* multilayer network. Nodes (illustrated as circles) represent 80 sensory neurons (yellow), 76 interneurons (green), 123 motor neurons (red), and 95 body wall muscles (magenta). A larger node size indicates a higher total node degree. In total, 3,538 distinct directed connections exist between the nodes. The different connection types are illustrated as colored arrows. Panel (a) is an overlay of all other panels (b)-(h). The placement algorithm for the neurons is FORCE ATLAS available in the program GEPHI.

Discriminative power – The discriminative power P measures the ability of a logistic regression model or potential explanatory variables to distinguish correctly between realizations (0 or 1) of the response variable y_i. We define the power as

$$P = \frac{1}{N_{Y=1}N_{Y=0}} \sum_{l=1}^{N_{Y=1}} \sum_{m=1}^{N_{Y=0}} \Psi\left(x_{l,Y=1}, x_{m,Y=0}\right),$$
$$\Psi\left(x_{l,Y=1}, x_{m,Y=0}\right) = \begin{cases} 1, & \text{if } x_{l,Y=1} > x_{m,Y=0} \\ 0, & \text{if } x_{l,Y=1} = x_{m,Y=0} \\ -1, & \text{if } x_{l,Y=1} < x_{m,Y=0} \end{cases}, \tag{2}$$

where $x_{l,Y=1}$ and $x_{m,Y=0}$ are realizations of the empirical distributions $X_{Y=1}$ and $X_{Y=0}$ with numbers of observations $N_{Y=1}$ and $N_{Y=0}$.

The power can be derived from the receiver operating characteristic (ROC) and the cumulative accuracy profile (CAP). Both ROC and CAP are important concepts to visualize the discriminative power or separation ability of a model. The information contained in a ROC or CAP curve can be aggregated into a single number, the area under the ROC curve (AUROC) or the accuracy ratio (AR). The AR can be interpreted as a simplified representation of AUROC since $\text{AR} = 2 \cdot \text{AUROC} - 1$ and is also known as Gini coefficient or power statistics. Note that the Mann–Whitney statistics can be introduced as a quantity equivalent to AUROC [21, Chapter 13], which allows us to calculate the discriminative power as introduced in Eq. (2). From a statistical point of view, the power P can be interpreted as the probability to uncover a difference when there really is one. The advantage of defining the power using the Mann-Whitney statistics is that 95% confidence intervals can be calculated which better account for the uncertainty associated with small sample sizes (see Appendix C for more details about logistic regression and discriminative power).

2.3. Data basis for logistic regression

Next, we introduce the data basis to which we apply the logistic regression tools in order to predict the neurons of the locomotory subnetwork. In compiling the data, it must be ensured that these are related to the locomotion behavior of the worm. We achieve this by considering the shortest paths between sensory neurons and muscle cells. In the development of regression models, we restrict ourselves to 5 sensory neurons through which *C. elegans* perceives gentle touches on the head and tail. The worm reacts on this by moving forward or backward. The shortest paths are illustrated in Fig. 2. In general, the transfer of information in the network begins with sensory neurons and proceeds via interneurons to motor neurons. The latter in turn can activate muscle cells. The particularity of *C. elegans* is that the stimulation and relaxation of muscle cells is effected by two types of motor neurons: excitatory and inhibitory. Moreover, the inhibitory cells are generally activated by excitatory cells which are included in the shortest paths. However, Fig. 2 has only an illustrative purpose and is not complete. Many of the motor neurons in the third group would also appear in the fourth group because the motor neurons themselves are interconnected. All locomotion-relevant interneurons and motor neurons are contained in the shortest paths.

The shortest paths are classified as paths going completely through the locomotory subnetwork and others. This is done first on the level of the neurons (results are taken from [83] and depicted in Fig. 6) and then on the level of the shortest paths:

$$LS_p = \begin{cases} 1, & \text{if neuron } p \text{ belongs to one of the classes AVD (2),} \\ & \text{AVA(2), PVC(2), AVB(2), VA(12), DA(9), AS(11),} \\ & \text{VB(11), DB(7), DD(6), VD(13), or PDB(1)} \\ 0, & \text{otherwise} \end{cases},$$
$$y_i = \begin{cases} 1, & \text{if the shortest path } sp_i \text{ between sensory neuron and} \\ & \text{muscle cell only involves neurons with } LS_p = 1 \\ 0, & \text{otherwise} \end{cases}.$$

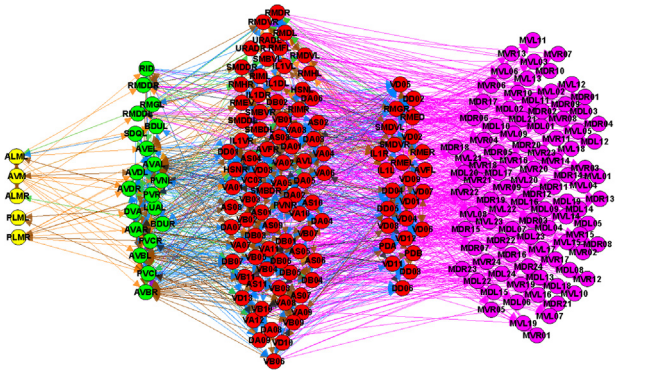


Fig. 2. The *C. elegans* information flow from touch sensory neurons to muscle cells. The information flows along the shortest paths from the sensory neurons (ALML, ALMR, AVM, PLML, PLMR) to specific interneurons for processing and then to specific motor neurons to complete an action. The latter can be directly connected to the muscle cells, or they can previously activate a second layer of motor neurons. The color code is as in Fig. 1.

The variable LS_p contains the information of being part of the locomotory subnetwork for the neurons $neuron_p$ with index $p = 1, 2, \ldots, 279$, and y_i holds the same information for the shortest paths sp_i with index $i = 1, 2, \ldots, n$. While LS_p plays a role in finding key factors for a logistic regression model in the univariate factor analysis, y_i is used in the multivariate optimization where factors are combined and parameters are fitted. The numbers of the shortest paths are presented in Table 2. Logistic regression models are developed on the dataset with 5 touch sensory neurons. In total, there are 12,431 pathways of which 8,450 pass through the locomotory subnetwork. The validation of the developed models is based on the dataset that includes all sensory neurons. In this case, 72,912 of a total of 156,888 pathways exploit the locomotory subnetwork.

2.4. Model development

After the data basis has been set up, the next step is to search for factors that have a high explanatory contribution regarding the locomotory subnetwork.

Univariate power analysis – Since the question of connectivity is of central importance in a network, we calculate the power values (2) for different degree distributions regarding the different neuron types in the shortest paths (see exemplary Fig. 2). There are a total of 20 interneurons, 8 of them belong to the locomotory subnetwork $LS_p = 1$, and 12 do not $LS_p = 0$. To analyze the power of

Table 2
Logistic regression datasets. The datasets represent the shortest paths between (a) the touch sensory neurons (ALML, ALMR, AVM, PLML, PLMR) and (b) all sensory neurons and the body wall muscles where the second neuron is an interneuron and the third and fourth (if existing) are both motor neurons. Logistic regression models are developed on (a) and tested on (b).

Shortest path length	Number of observations for $Y_i = 1$	for $Y_i = 0$	in total
(a) Touch sensory neurons			
3	1,691	540	2,231
4	6,759	3,441	10,200
Total	**8,450**	**3,981**	**12,431**
(b) All sensory neurons			
3	14,784	13,439	28,223
4	58,128	70,537	128,665
Total	**72,912**	**83,976**	**156,888**

these 20 neurons, the in- and outdegree within the multilayer network (Fig. 1) must be determined first. In Fig. 3a, the power P is visualized for different transmitter types. For the peptides, it is also relevant whether they are released as classical transmitters or as cotransmitters (cf. Fig. A.1d). In the lower part of the figure, the highest observed values among the 8 neurons of the locomotory subnetwork $LS_p = 1$ are also given for better understanding.

The indegree of the interneurons yields many factors with extremely good power. The values for ACh, Glu, electrical transmission, and peptide cotransmission are greater or equal to 75% suggesting that most neurons of the locomotory subnetwork $LS_p = 1$ can be distinguished from other neurons $LS_p = 0$. The positive power values indicate that most neurons of the locomotory subnetwork should have higher indegree values than other neurons. Glu has the best power with 95.8%, whereby the highest observed indegree among the neurons of the locomotory subnetwork is 24. The indegree of the electrical connections is also an excellent factor to separate the neurons. The power value of 75% also signifies that there is a difference between the neurons, but this time, the uncertainty is a little higher which is shown by the broader confidence interval. On the other side, the power values of GABA, MA, and peptide transmission are not unsatisfactory. From a probabilistic point of view, however, it is very uncertain to state a difference. The confidence intervals indicate that the true power value could be lower than 5%, which means that the null hypothesis "There is no difference" should not be rejected (on a significance level of 5%). The reason for this uncertainty is the low sample size of 20 neurons. Note that it is sufficient to postulate a difference as long as power values are greater than 60%.

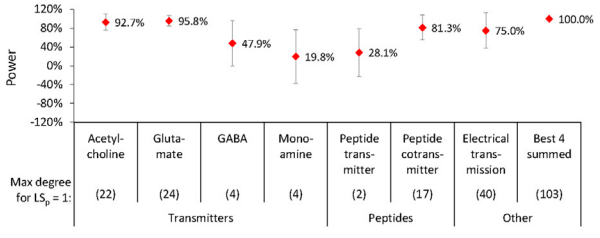
In addition, the indegree values of the interneurons can be combined for different transmission types. The summation of the values for ACh, Glu, electrical transmission, and peptide cotransmission lead to a power of 100% meaning that the neurons of the locomotory subnetwork can be perfectly separated from other neurons. The locomotory interneurons are generally characterized by a high outdegree since they convey signals from various sensory neurons to a large number of motor neurons. On the other hand, locomotion is the most important behavior of *C. elegans*. The worms must be able to respond to environmental changes by adapting their locomotion behavior. Therefore, the interneurons of the locomotory subnetwork have also a high indegree due to different sensory inputs that include many different transmitters. This is already indicated in Fig. 1. The desired 8 interneurons have the highest total degree among all neurons. The outdegree power values of the interneurons are presented in Fig. C.3.

Next, we consider the first layer of motor neurons (excitatory) in the shortest paths (see exemplary Fig. 2). It is assumed that the 8 interneurons of the locomotory subnetwork are known, and only those motor neurons are selected that are postsynaptic to them. This limits the number of motor neurons to 62 of which 54 belong to the locomotory subnetwork. The power values are depicted in Fig. 3b. There are two factors with negative power values lower than −70%, Glu and peptide transmission. In this case, the minus sign indicates that the motor neurons of the locomotory subnetwork $LS_p = 1$ have more lower indegree values than other neurons $LS_p = 0$. The indegree for peptide transmission has the best power value with absolute 87.5%. The motor neurons of the locomotory subnetwork can be distinguished from most other neurons because they do not have peptide transmitters. Note that the power cannot be further increased by simple combination of factors. All tested combinations result in a lower absolute power than for the peptide transmitters. For example, the combination with Glu decreases the power to absolute 80.3%. The outdegree power values can be seen in Fig. C.4a.

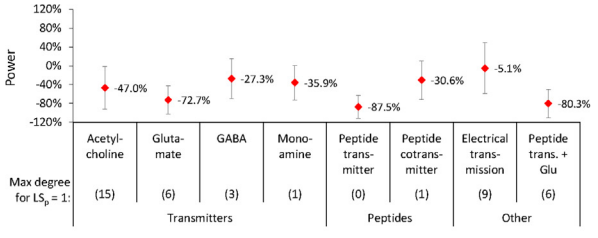
Apart from connectivity, it can also be sufficient to consider only the existence of specific transmitters for the neurons. In this case, no degree distributions need to be calculated, and a better qualitative understanding of the neurons in the network can eventually be gained (see Fig. C.4b and Appendix C.3). The in- and outdegree power values of the second layer of motor neurons (inhibitory) are provided in Fig. C.5. The results are comparable with those of the excitatory motor neurons, but they do not contribute to the multivariate optimization, which we will investigate next.

Multivariate optimization – In the multivariate analysis, the factors discussed in the previous section are combined using logistic regression (1) to further increase their power (2). The basis are the shortest paths between touch sensory neurons and muscle cells (cf. Table 2a), which include all locomotion-relevant interneurons and motor neurons. For predicting the locomotory subnetwork, we have shortlisted three models. The factors and parameters used including the resulting negative twofold log-likelihood are provided in Table 3.

All models require only two factors. Models 1 and 3 consider the combined in- and outdegree of the interneurons as the first factor. The best factor from the indegree analysis is summed up with the best factor from the outdegree analysis (for details see Figs. 3a, C.3). The inclusion of the outdegree is not absolutely necessary because the indegree power in the univariate analysis is already 100%, but it makes the prediction of the logistic regression models a little more robust. On the other hand, Model 2 considers only the indegree of the interneurons as the first factor. Models 1 and 2 utilize as the second factor the peptide transmitter indegree of the first layer motor neurons. Model 3 takes into account a qualitative factor that aims only for the existence of certain transmitters and peptides. First, the existence of the ACh transmitter is binary coded (1 or 0). Then, the same is done for NLP and PDF peptides, which are subtracted from it (for details see Fig. C.4b). The peptide transmitter indegree is the only factor with a negative coefficient because motor neurons of the locomotory subnetwork do not have such connections, which is indicated by negative power values in the univariate analysis. For the other factors applies that the neurons of the locomotory subnetwork can be recognized by having more higher values than other neurons. The negative twofold log-likelihood indicates that Model 1 should have the highest power followed by Model 2 and Model 3.



(a) Interneurons



(b) First layer motor neurons

Fig. 3. Indegree power values.

Table 3
Logistic regression models to predict neurons involved in locomotion behavior of *C. elegans*. The table provides the estimated coefficients for the explanatory variables, the estimated constant, and the resulting negative twofold log-likelihood for the regression models. The latter is given by the statistical program SPSS. Under the null hypothesis "The model has a perfect fit", the fit is the better the lower its value. "Best 4 summed" includes the indegree for ACh, Glu, electrical transmission, and peptide cotransmission.

Model	Interneuron		First layer motor neuron		Constant	−2·Log- Likelihood
	Indegree (Best 4 summed)	Outdegree (ACh + electrical)	Indegree (Peptide transm.)	Quality factor (ACh - NLP - PDF)		
1		0.164	−25.3		−10.3	3,478.8
2	0.212		−25.1		−6.5	3,714.9
3		0.157		9.7	−19.5	3,908.7

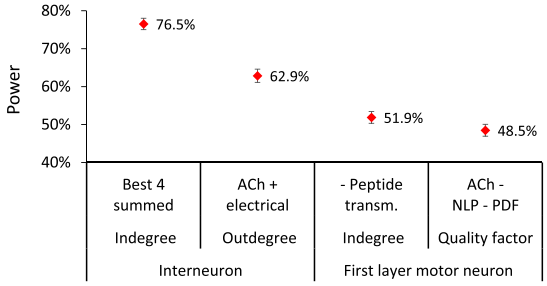


Fig. 4. Individual power values for the chosen regression factors on the development dataset.

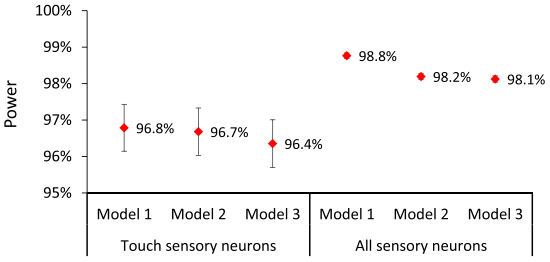


Fig. 5. Power values of the logistic regression models.

Considering the shortest paths, the power values of the selected factors are compared in Fig. 4. The interneuron indegree only achieves a power value of 76.5% meaning that the information concerning the interneurons is no longer unique. For example, many paths can lead through a particular interneuron of the locomotory subnetwork, but in dependence of the postsynaptic motor neurons, each path is classified as path through the locomotory subnetwork $y_i = 1$ or not $y_i = 0$. Therefore, the same interneuron can be part of both, which makes the information redundant. For this reason, the power cannot be 100% as it is the case in the univariate analysis where a distinct group of neurons is considered. The idea behind this procedure is as follows: If the neurons of the distinct groups can be separated by certain factors, then these factors should also be able to distinguish the shortest paths in combination. The other factors in Fig. 4 have a lower power than the interneuron idegree, but their separation ability is still more than acceptable. The quality factor of the first layer motor neurons has the lowest value with $P = 48.5\%$.

Finally, Fig. 5 illustrates the resulting power for the fitted regression models on the development (left) and the validation dataset (right). The models show power values between 96% and 97% for the touch sensory neurons. If the models are applied on the shortest paths with all sensory neurons, slightly better values between 98% and 99% are obtained. Note that there are 90 neurons with sensory function (instead of 5), and 93 interneurons are postsynaptic to them (instead of 20), but only 8 belong to the locomotory subnetwork $LS_p = 1$ as assumed. The interneurons in

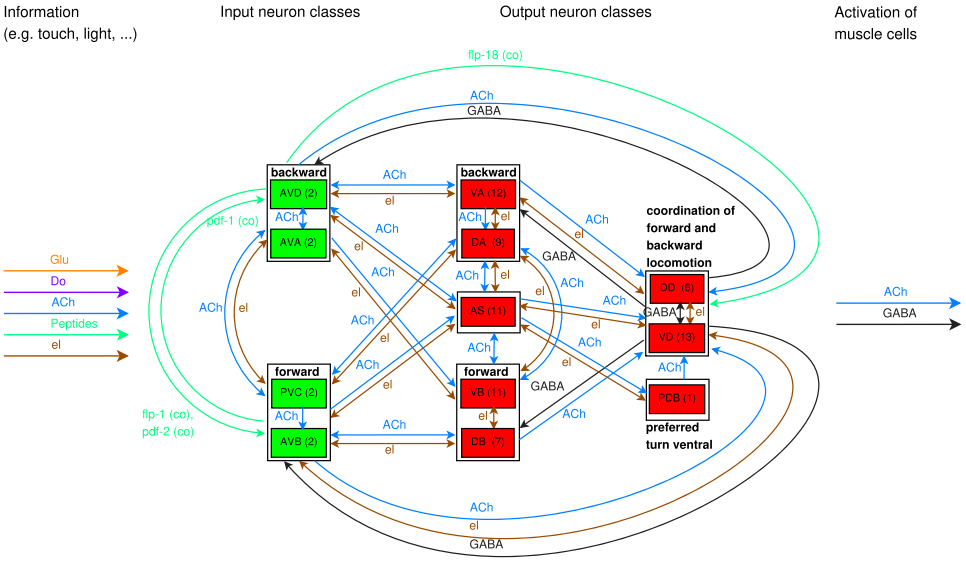


Fig. 6. The *C. elegans* main motor program for forward and backward locomotion (locomotory circuitry). The command interneurons (green) initiate the forward and backward locomotion by primarily activating the first layer motor neurons. These stimulate muscle contraction via the transmitter ACh. In addition, they activate the second layer motor neurons, the DD and VD neurons, which relax muscle contraction via the transmitter GABA. The numbers in brackets indicate the number of neurons per neuron class.

turn are presynaptic to 106 motor neurons (instead of 62) where 54 belong to the locomotory subnetwork. Using the classification scheme in Appendix C.2, the achieved power values are outstanding. They indicate that the shortest paths can be well separated by the logistic regression models. One interesting question is what predictions can be made on this basis, which we will investigate more closely.

It remains to be mentioned that none of the regression models employs factors for the second layer motor neurons. The reason is that the power on the development dataset could not be significantly increased by using these factors, which is probably due to high correlation effects between the first layer and second layer motor neurons. The positive aspect is that the wanted inhibitory motor neurons in the second layer are almost all only postsynaptic to the wanted excitatory motor neurons in the first layer. Therefore, it is sufficient to predict the interneurons and the excitatory motor neurons. However, the power can be slightly increased by taking more factors into account, but the power gain does not justify such complications.

2.5. Predicting the locomotory subnetwork

In this section, the performance of Models 1 and 2 is examined in more detail.

Prediction accuracy with touch sensory neurons – The output of the logistic regression models on the shortest paths is classified by the threshold 0.5. Values greater than or equal to the threshold indicate paths through the locomotory subnetwork, and values smaller than the threshold indicate other paths. In practice, the threshold should be adjusted so that the predicted proportion of paths through the locomotory subnetwork corresponds to the actual observed proportion. Since Models 1 and 2 have an outstanding discriminative power and the proportion of $y_i = 1$ changes therefore only marginally when predicting the development dataset (cf. Table 2a), this is not necessary. Table 4 provides the performance of the models for the touch sensory neurons. The results for all sensory neurons can be found in Table C.6.

The prediction of Model 1 is very close to a perfect result. In total, the model correctly predicts 99.2% of all shortest paths, and none of the paths through the locomotory subnetwork $y_i = 1$ are misclassified. On the other hand, some paths that do not completely pass through the locomotory subnetwork $y_i = 0$ are incorrectly classified as paths through the locomotory subnetwork. However, the proportion of incorrectly classified paths is only 2.6%. The prediction accuracy of Model 2 is slightly lower with a total performance of 96.9% because there is a larger misclassification of paths of 9.7% that do not belong to the locomotory subnetwork. The reason for this is that Model 2 does not take into account the outdegree of the interneurons. Nevertheless, both models deliver an excellent performance suggesting that they should be appropriate for use on the larger dataset with all sensory neurons (cf. Table 2b).

Model predictions with all sensory neurons – We now consider the shortest paths that are classified as paths through the locomotory subnetwork by Models 1 and 2 with threshold 0.5 and analyze the underlying neurons. Both model predictions include all neurons of the locomotory subnetwork $LS_p = 1$. Besides, Model 1 (2) predicts 8 (29) further neurons, which are presented in Table C.7. Note that for all of these neurons except for three it can be assumed on the basis of literature research and connectivity analysis that they are involved in locomotion behavior of *C. elegans* (cf. Section C.4). For example, the SMD and RMD motor neurons are connected with muscle cells in the head and neck and contribute to multiple navigation behaviors. Therefore, the locomotory subnetwork can be divided into different circuits. While the starting neurons $LS_p = 1$ belong to a circuit that mainly initiates forward and backward locomotion (main motor program) of *C. elegans* [19, Section I], the SMD and RMD neurons can be assigned to a circuit for navigation [28]. Model 2 is capable of capturing both circuits since its first factor is little less strict. Without considering the outdegree, a few more interneurons are allowed that are presynaptic to additional potential motor neurons. Altogether, the results can be regarded as a success since in the end we obtain more neurons involved in locomotion behavior than originally assumed. However, in order to make a more precise statement regarding the power or performance, a reparametrization of the regression models including the classification threshold should be done.

After predicting the neurons of the locomotory subnetwork using logistic regression, the next step is to examine their underlying dynamics.

3. Simulation of dynamics

In this section, we investigate the patterns of neuromuscular activity that occur in response to a touch stimulus on the tail. Since *C. elegans* can only react with forward locomotion, we restrict ourselves to the main motor program that initiates forward and backward locomotion to certain stimuli. Therefore, the results by [83] ($LS_p = 1$) including the 5 touch sensory neurons serve as neuronal basis.

3.1. Locomotory circuitry

Fig. 6 depicts the utilized circuit for forward and backward locomotion. The neurons are summarized in classes, such as AS and PDB, and components of classes, such as the forward locomotion component of interneurons. The neuron classes including their neuron members have been defined by White et al. [80]. Most neuron classes in *C. elegans* are composed of a pair of two bilaterally symmetric neurons, which can be observed across the L/R axis. This also applies to the interneurons. For example, the neuron AVBL (AVBR) is located on the left (right) side of the worm. The motor neurons are distributed along the midline of the worm. The first letter indicates connections to either dorsal or ventral muscle cells. The second letter gives the type of the neurons (A-, B-, and D-type neurons). The exception are the AS neurons, which are connected to dorsal muscle cells. The motor neuron classes contain an asymmetry in neuron numbers. The neurons are numbered from head to tail. For example, AS11 lies in the tail of the worm.

Table 4
Model performance on the development dataset (touch sensory neurons).

Locomotory subnetwork	Prediction Model 1			Prediction Model 2		
	correct	incorrect	correct in %	correct	incorrect	correct in %
yes	8,450	0	100.0	8,450	0	100.0
no	3,876	105	97.4	3,594	387	90.3
Total	**12,326**	**105**	**99.2**	**12,044**	**387**	**96.9**

If the worm experiences a gentle touch on its tail, the sensory neurons PLML, PLMR will register this and provide input to the forward interneurons PVC via Glu and electrical transmission but also to the backward interneurons AVD and AVA only via Glu. The forward component primarily activates the B-type motor neurons VB and DB via ACh and electrical transmission. Subsequently, the VB (DB) neurons stimulate ventral (dorsal) muscle cells via ACh and, at the same time, the second layer motor neurons DD (VD) via ACh and electrical transmission, which are connected to muscle cells on the opposite side. The D-type motor neurons DD (VD) have an inhibitory effect and relax muscle cells via GABA so that the simultaneous contraction of ventral and dorsal muscles is prevented. Therefore, these neurons enable the coordination of the movement of the worm. The explanation for the backward locomotion is analog [19, Section I]. The AS neurons are not included in many models and studies of locomotion of *C. elegans*. These are only connected to dorsal muscle cells and therefore support the DA and DB neurons to cause dorsal muscle contraction. The PDB neuron probably causes a ventral bias when performing large wavelength body bends that occur during turning [83]. It was included for the sake of completeness but should not play a decisive role for the initiation of forward locomotion. Furthermore, some peptides acting as ACh cotransmitters can be identified in Fig. 6. These are neglected in the following. Note that the layered structure shown in the figure is also in good agreement with a community analysis of dynamical correlations [58].

During locomotion, the worms generate rhythmic body undulations. This raises the important question of how these are related to neuromuscular dynamics.

3.2. Modeling neuromuscular activity

The neuron dynamics can be described in terms of the three-dimensional Hindmarsh-Rose system [35]

$$\begin{aligned}\dot{p}_i(t) &= q_i(t) - ap_i(t)^3 + bp_i(t)^2 - n_i(t) + I_{\text{ext},i} - g_{\text{Glu}}(p_i(t) - V_{\text{exc}}) \\ &\quad \times \sum_{k=1}^{N} C_{\text{Glu},ki} S(p_k(t)) - g_{\text{ACh}}(p_i(t) - V_{\text{exc}}) \sum_{k=1}^{N} C_{\text{ACh},ki} S(p_k(t)) \\ &\quad - g_{\text{GABA}}(p_i(t) - V_{\text{inh}}) \sum_{k=1}^{N} C_{\text{GABA},ki} S(p_k(t)) + g_{\text{el}} \sum_{k=1}^{N} L_{ik} p_k(t), \quad \text{(d)} \\ \dot{q}_i(t) &= c - dp_i(t)^2 - q_i(t), \quad \text{(e)} \\ \dot{n}_i(t) &= r[s(p_i(t) - p_0) - n_i(t)], \quad \text{(f)}\end{aligned}$$

where $p_i(t), i = 1, \ldots, N$, denotes the membrane potential of the i-th neuron at time t, $q_i(t)$ represents the fast current, either Na^+ or K^+, and $n_i(t)$ the slow current, for example, Ca^{2+}. The parameter r is the time scale separation between fast and slow variables.

The $N \times N$ coupling matrices $\mathbf{C}_{\text{Glu}}, \mathbf{C}_{\text{ACh}}$, and $\mathbf{C}_{\text{GABA}}$ contain as weights the number of chemical synapses between pairs of neurons for the transmitters Glu, ACh, and GABA, respectively. Since the chemical coupling is nonlinear, the membrane potential $p_k(t)$ is transformed by the sigmoid function $S(p) = 1/[1 + \exp(-\lambda(p - \theta))]$, which acts as a continuous mechanism for the activation and deactivation of chemical synapses. The reversal potentials for excitatory and inhibitory chemical synapses are denoted by V_{exc} and V_{inh}, respectively.

The connectivity of the electrical synapses is described in terms of the Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{G}$ where the coupling matrix $\mathbf{D}$ and the diagonal degree matrix $\mathbf{G}$ comprise as weights the number of electrical synapses between each neuron pair and the total number of electrical synapses per neuron, respectively. Since the electrical coupling is linear, the membrane potential $p_k(t)$ must not be transformed.

Throughout this paper, the system parameters are set as follows: $a = 1, b = 3, c = 1, d = 5, r = 0.005, s = 4, p_0 = -1.6, V_{\text{exc}} = 2, V_{\text{inh}} = -1.5, I_{\text{ext},i} = 9$ for the neurons PLML/R else $0, \lambda = 10$, and $\theta = -0.25$. For the coupling strengths, we use the values: $g_{\text{el}} = 0.20, g_{\text{Glu}} = 0.10, g_{\text{ACh}} = 0.10$, and $g_{\text{GABA}} = 0.15$.

Note that the Hindmarsh-Rose model is based on the global behavior of the neuron and does not aim to model in detail the microscopic biochemical processes. Therefore, it is simpler than Hodgkin–Huxley [37] or Morris-Lecar [51] models which attempt to reproduce the electrophysiological process of biological neurons [72] based on conductance models with many governing equations and coefficients. Instead, the Hindmarsh-Rose model uses dimensionless variables (and parameters) and must be scaled to become biologically applicable, but it describes well the neuronal behavior observed in experimental biology. The sigmoid function $S(p)$ can be parametrized differently for each transmitter type. For the sake of simplicity, we adopt the parameters λ and θ for all types. The inhibitory effect of the transmitter GABA is achieved by the parameter $V_{inh} = -1.5$, which takes into account that inhibitory synapses can also be excitatory. More important is to find a parametrization for the coupling strengths. Electrical synapses have properties that are different from chemical ones. For instance, they conduct nerve impulses extraordinarily fast (almost instantaneous). Communication between neurons is possible without delay, which is not typical for chemical synapses. Therefore, electrical synapses are most frequently observed in neuronal circuits that require the fastest possible response [38]. In terms of the neuronal dynamics, shorter signal propagation times imply stronger couplings. Consequently, we set the electrical coupling as the strongest. The explanation for the chemical coupling strengths is based on the locomotory circuitry (Fig. 6). The inhibitory transmitter GABA is employed by the DD and VD neurons. In contrast to the excitatory motor neurons, they activate muscle cells on the opposite side in order to prevent simultaneous contraction of dorsal and ventral muscles. For compensating the longer signaling pathways, the coupling of GABA is assumed to be slightly stronger than ACh.

The muscle dynamics is based on the neuronal activity and can be modeled as leaky integrators [7,41]

$$\dot{A}_l(t) = \frac{1}{\eta}\left[\sum_{s=1}^{N}\left(g_{\text{ACh}} E_{\text{ACh},sl} p_s(t) - g_{\text{GABA}} E_{\text{GABA},sl} p_s(t)\right) - A_l(t) + const\right] \quad (4)$$

with a characteristic time scale of $\eta = 100$ ms, which roughly agrees with response times of obliquely striated muscles [49]. The activation state of the muscle cells at time t is represented by the dimensionless variable $A_l(t), l = 1, \ldots, M$, where M denotes the number of muscle cells. The $N \times M$ coupling matrices $\mathbf{E}_{\text{ACh}}$ and $\mathbf{E}_{\text{GABA}}$ contain as weights the number of neuromuscular synapses between neurons and muscle cells for the transmitters ACh and GABA, respectively. The coupling strengths g_{ACh} and g_{GABA} are as specified above. The cosmetic parameter $const = 1$ ensures that the muscle activation is more positive than negative.

Note that also electrical synapses exist between body wall muscles. However, the electrical coupling between muscles is assumed to be too weak and can therefore be neglected [8].

3.3. Modeling forward locomotion

The thrust for the forward locomotion of *C. elegans* is provided by 95 body wall muscles located in 4 quadrants along the midline [2]. For simplicity, we limit our mathematical description to the dorsal and ventral quadrant on the right-hand side and assume that the muscles are arranged in rows and numbered $1, 2, \ldots, 24$ from head to tail, with each dorsal muscle having a ventral partner. This takes into account that the worm crawls on its side and gen-

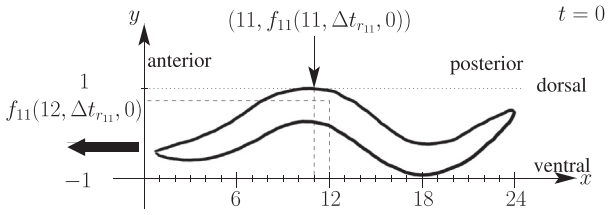
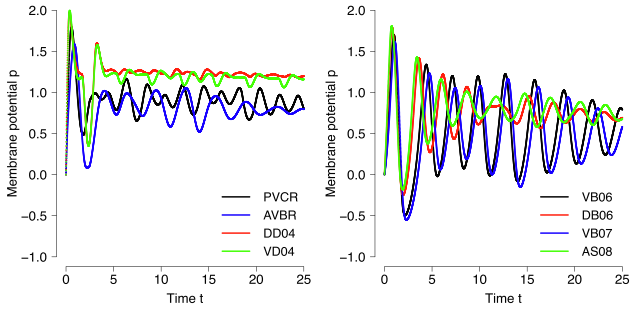


Fig. 7. Schematic diagram of coordinates. At time $t = 0$, the muscle pairs $m = 1, \ldots, 24$ lie along the x-axis. The worm moves forward in negative x-direction. The harmonic wave $f_m(x), x \in [1, 24]$ [cf. Eq. (5)], corresponds to the elongation of the worm in y-direction depending on the position x of the m-th muscle pair. Note that at time $t = 0$, the function f_m has either a minimum or a maximum at position $x = m$, as exemplified for f_{11}, which is adjusted by the phase offset ϕ_m (cf. Table D.10).



(a) Isopotential neurons **(b)** Oscillatory motor neurons

Fig. 8. Neuronal activity patterns. Two types of activity patterns can be identified among the neurons of the locomotory circuitry: (a) isopotentials and (b) local oscillations.

erates sinusoidal undulations. Then, the locomotion of the worm can be described as harmonic wave. Since the undulations spread from the head to posterior, we model the elongation (orthogonal to the direction of motion) of each muscle pair $m = 1, 2, \ldots, 24$ consisting of a dorsal and a ventral muscle as harmonic wave

$$\begin{aligned} f_m(x, \Delta t_{r_m}, t) &= B \sin\left[\gamma(x, \Delta t_{r_m}, t)\right] \\ &= B \sin\Big[\tfrac{2\pi}{\lambda} x - \bar{\omega}(\Delta t_{r_m}) t + \phi_m \\ &\quad + \sum_{s_m=1}^{r_m \geqslant 1} \left[\bar{\omega}(\Delta t_{s_m}) - \bar{\omega}(\Delta t_{s_m-1})\right] t_{s_m}\Big] \end{aligned} \tag{5}$$

where x is the position of the m-th muscle pair, t is the time, B is the amplitude, λ the wavelength, $\overline{\omega}$ the mean angular frequency between extreme body bends, ϕ_m the phase offset, and $\Delta\overline{\omega}$ the change of the mean angular frequency. The coordinates are chosen such that $x \in [1, 24]$ for each harmonic wave $f_m(x)$ based on the m-th muscle pair indicating the body curvature of the worm at time t (see Fig. 7). The index $r_m = 0, 1, 2, \ldots, N_m$ refers to extreme values in the time series of subtracted muscle activity $A_m^{\text{ventral}}(t) - A_m^{\text{dorsal}}(t)$ (cf. Fig. 10).

The first term $2\pi x/\lambda$ in the sine function defines the position of the harmonic wave f_m at time $t = 0$ without phase shift. The wavelength is estimated to be $\lambda = 18$ and corresponds to 75% of the length of the worm. The second term describes the propagation of the harmonic wave in positive x-direction so that the worm moves forward in negative x-direction. The mean angular frequency results from the temporal difference between two subsequent minimum and maximum values r_m and $r_m + 1$ (or vice versa) leading to $\overline{\omega}(\Delta t_{r_m}) = \pi/\Delta t_{r_m}$ with $\Delta t_{r_m} = t_{r_m+1} - t_{r_m}$ and $t_{r_m} \leqslant t < t_{r_m+1}$. The third term represents the initial phase ϕ_m, which is calculated on the basis of linear functions (see Table D.10). With every subsequent extreme value, phase jumps occur because the mean angular frequency changes. These jumps are compensated by the fourth term. The corresponding change of the mean angular frequency $\overline{\omega}(\Delta t_{s_m}) - \overline{\omega}(\Delta t_{s_m-1})$ multiplied by time t_{s_m} results in the required phase correction, which is cumulative as expressed by the summation. Finally, the amplitude of the harmonic waves is set to $B = 1$. The forward locomotion of the worm is described by the body wave which results as an average value over all harmonic waves $f_m(x), x \in [1, 24]$, based on the m-th muscle pair.

3.4. Simulation results

In order to integrate the dynamical Eqs. (3) and (4), the statistical software R is utilized. The simulation time is 25 time units with timesteps of 0.005. The wave model (5) is applied as a downstream process on the simulated time series for muscular activity with timesteps of 0.05.

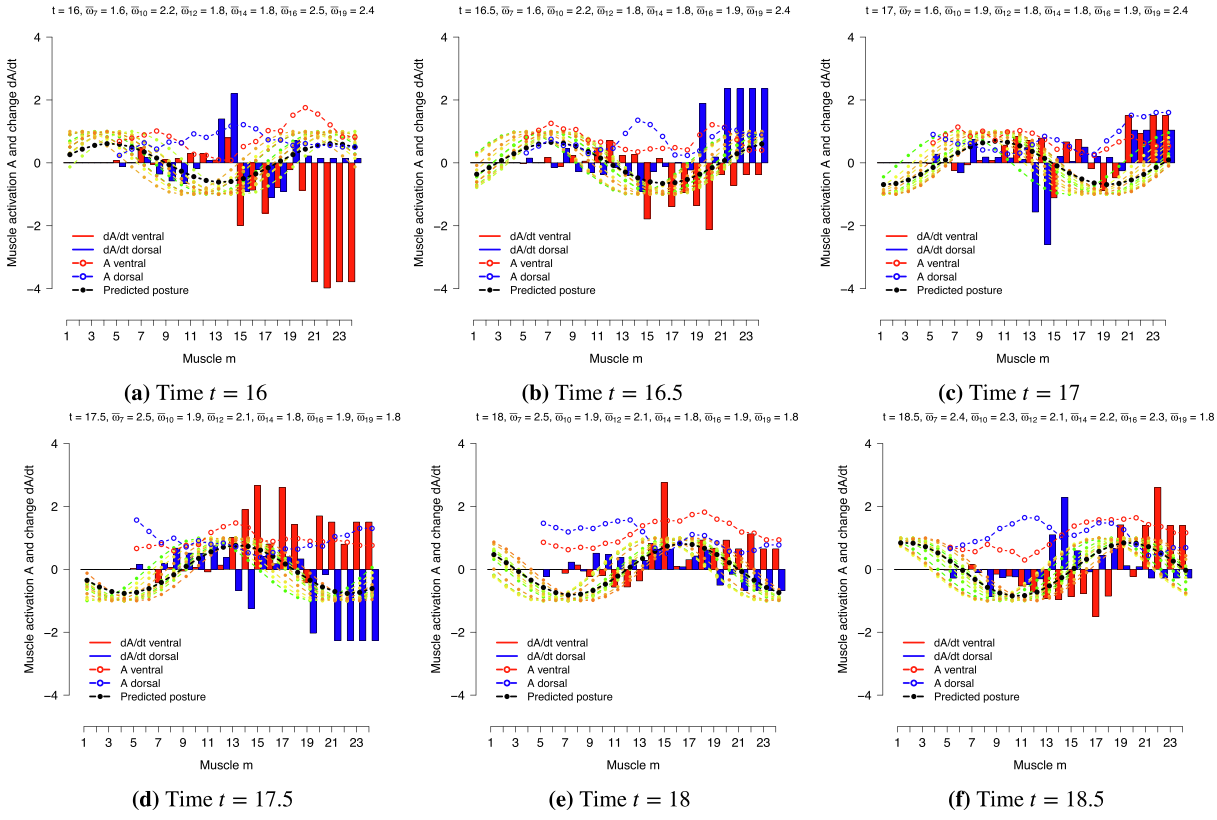
Neuronal activity – Many neurons in *C. elegans*, especially those involved in locomotion, do not fire classic action potentials [79]. This, however, does not exclude signal processing in *C. elegans* [27]. Therefore, the time series represent the behavior of the membrane potential below action potential threshold, and there is a huge operating range where different patterns (regenerative events) can occur [47].

The neurons of the locomotory circuitry are essentially either isopotentials or local oscillators (see Fig. 8). The neurons PVCR and AVBR are representative for the most interneurons and show a rather constant activity (very small oscillations) after considering a transient time of about 5 time units. Ablation experiments proved that these are active and non-oscillating during forward locomotion [26].

The neurons DB06, VB06, VB07, and AS08 are representative for most first layer motor neurons (cf. Fig. 6) and exhibit harmonic oscillations with non-constant amplitude over time. The latter is not subject of this study. Comparing different neurons within a class, their oscillations appear to be slightly shifted to each other. In this context, the neurons DB06, VB06, and VB07 are shown in Fig. 8b because they reproduce the experimentally observed activity of comparable neurons [24, Fig. 5]. The simulated membrane potential of DB06 is in anti-phase to those of VB06 and VB07, whereas the membrane potential of VB07 precedes that of VB06. Moreover, the motor neurons function as local oscillators to generate alternating bends of the *C. elegans* body. The intrinsic oscillations are amplified by the isopotential activity of the interneurons (Fig. 8a) in order to drive forward and backward movements [79]. The same must apply to the AS neurons because they exhibit a similar oscillatory behavior, as indicated by AS08 in Fig. 8b.

The neurons DD04 and VD04 represent the behavior of all DD and most VD neurons and show a very constant activity over time. Within the classes DD and VD, the neurons are connected by many electrical synapses, which could be a reason for this behavior because they must be able to quickly transmit signals from excitatory motor neurons to opposite muscle cells in order to coordinate locomotion. In other words, they contribute to an anti-phase activation of ventral and dorsal muscles, which is necessary for alternating body bends (cf. Fig. 10b). Similar to the shifts of membrane potential oscillations for excitatory motor neurons (Fig. 8b), different levels of membrane potential can be detected within the classes DD and VD or among interneurons (Fig. 8a).

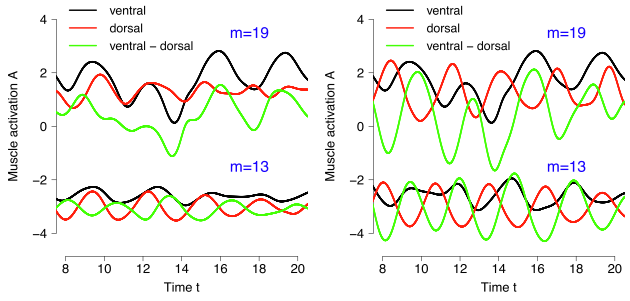
A central pattern generator (CPG) is a component of a neuronal network that is capable of generating rhythmic oscillations to drive motor behaviors such as locomotion. Pattern generation may result



(a) Time $t = 16$ **(b)** Time $t = 16.5$ **(c)** Time $t = 17$

(d) Time $t = 17.5$ **(e)** Time $t = 18$ **(f)** Time $t = 18.5$

Fig. 9. Forward locomotion of *C. elegans* with time-delayed feedback control. The posture of *C. elegans* can be described as harmonic wave and results from ventral (red curve) and dorsal (blue curve) muscle activity, which is simulated with Eqs. (3) and (6) [parameters of feedback control are provided in Tables D.8-D.9] and plotted over the muscle index m. Based on this, the harmonic waves $f_m(x), x \in [1, 24], m = 5, 7, 8, \cdots, 24$, are calculated [cf. Eq. (5)] and averaged (black curve). Besides the averaged wave, 12 individual waves with index $m = 5, 7, 9 - 14, 16, 17, 19, 20$ are shown in green, orange, and yellowish color, which can be interpreted as its fluctuations. On top of the diagrams, chosen mean angular frequencies for the harmonic wave model are given. In addition to the muscle activation, its temporal change is in.dicated by bars.



(a) Uncontrolled **(b)** With feedback control

Fig. 10. Effect of time-delayed feedback control on muscular activity for two muscle pairs. Panel (a) and (b) correspond to the uncontrolled [Eqs. (3) and (4)] and controlled case [Eqs. (3) and (6) – additional parameters are provided in Tables D.8-D.9]. The black and red curve refer to ventral and dorsal muscle activation. The green curve is their difference, which indicates the body curvature at $x = m$. Note that the time series for the muscle pair $m = 13$ do not reflect the actual muscle activation but are shifted downwards from the original level.

from interactions between neurons (network-based rhythmicity), or through interactions among currents in individual neurons (pacemaker oscillator neurons) [22,26]. The simulation results in Fig. 8b prove that the main motor program of *C. elegans* is a neuronal circuit with intrinsic oscillatory activities. Therefore, it fulfils the prerequisites of a central pattern generator. Furthermore, we conclude that the coupling of neurons within the circuit (Fig. 6) is responsible for the oscillatory behavior. For this reason, the CPG is an intrinsic property of the multilayer network of *C. elegans* which generates the sinus rhythm for locomotion by the interaction of different neuron types.

Muscular activity – The neuronal harmonic oscillations generated in the locomotory circuitry are transferred to body wall muscles for locomotion. For a coordinated locomotion of *C. elegans*, an anti-phase muscular activity is required between dorsal and ventral muscle cells. However, this is not the case for most muscle pairs in our simulation. The harmonic waves (5) – adjusted on the basis of muscular activity – diverge significantly. Later, we will see how the problem can be solved to a large extent by using time-delayed feedback control. However, we already take the results into account below.

Fig. 9 captures the dynamics of the muscles at different times. The dorsal (ventral) muscle activation A is represented by the dotted blue (red) curve and its temporal change $\dot{A}$ by the blue (red) bars. Since the locomotory circuitry (Fig. 6) mainly covers the forward and backward locomotion of *C. elegans*, some of the muscles in the head and neck are missing. This concerns the dorsal and ventral muscles $1 - 4$ and 6. For the latter, the mean activation of the muscles 5 and 7 is calculated. Note that the displayed muscle activation A does not reflect the actual simulated values. For better visibility, the time series are first normalized to values between 0 and 2 in the time interval $[8, 20]$ and then smoothed with a simple second order moving average in order to avoid a serrated appearance. However, the modeling of forward locomotion is independent of this and done beforehand. The black curve results from the average over all 19 harmonic waves $f_m(x), x \in [1, 24], m = 5, 7, 8, \cdots, 24$, calculated with Eq. (5) and constitutes the final posture of the

worm. In addition to the averaged body wave, 12 individual harmonic waves are shown in the graphics in green, orange, and yellowish color, which can be interpreted as its fluctuation. For some of them, the mean angular frequencies are given at the top of the diagrams indicating a coordinated locomotion if they are close together. The body wave propagates to the right in time implying that the worm moves forward to the left. Note that the simulated ventral and dorsal muscle activation only partially corresponds with the predicted body curvature of *C. elegans*. Possible reasons are discussed in Section 5.2.

After studying the neuronal and muscular activity patterns that drive forward locomotion, we investigate synchronization patterns and significant neurons in the following that contribute to coordinated movements of *C. elegans*.

4. Synchronicity of motion

Based on the simulations of muscular activity (4), the harmonic waves (5) are not well synchronized. This corresponds to an uncoordinated locomotion of *C. elegans*. In order to enhance the synchronicity of harmonic waves, we consider a feedback control scheme. Subsequently, synchronicity is quantified over relevant simulation time using the Kuramoto order parameter as introduced in later Section 4.2. This parameter then forms the reference for further simulations in which certain neurons of the locomotory circuitry are silenced. As a consequence, conclusions can be drawn about the importance of neurons for a coordinated locomotion behavior.

4.1. Time-delayed feedback control

In Fig. 8b, neuronal oscillations are shifted to each other. This implies the existence of different time delays between the motor neurons. For this reason, these must also exist in muscular activity. The neuronal oscillations should be transferred to muscle cells via neuromuscular synapses in such a way that there is an anti-phase activation between ventral and dorsal muscles. Only then, a smooth and coordinated locomotion of *C. elegans* is possible. However, the neuromuscular connections are not perfect but partly based on estimates. This concerns in particular the given numbers of synapses [82]. Therefore, anti-phase behavior is not observed and a coordinated locomotion of the worm is not given. However, the coordination of locomotion can be improved with time-delayed feedback control, which is a general powerful control method in nonlinear systems [61,71,70]. Such delayed feedback mechanisms are often present in neuronal systems due to intrinsic propagation and processing delays.

For this purpose, the muscle dynamics is remodeled as

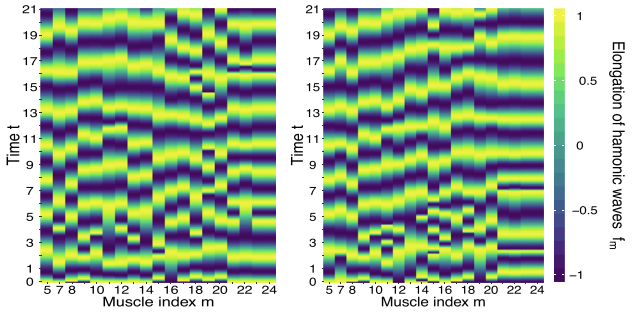
$$\dot{A}_l(t) = H(A_l(t), p_s(t)) - \frac{K_l}{\eta}\left[g_{\mathrm{ACh}}\sum_{s=1}^{N} E_{\mathrm{ACh},sl}(p_s(t)-p_s(t-\tau_s)) - g_{\mathrm{GABA}}\sum_{s=1}^{N} E_{\mathrm{GABA},sl}(p_s(t)-p_s(t-\tau_s))\right], \tag{6}$$

where function H corresponds to the right side of Eq. (4). The second term represents the control force that entails two new components: (i) the time delay τ_s for the N neurons and (ii) the feedback strength K_l for the M muscles. One peculiarity is that the control itself is delayed. It takes place after the integration of the dynamical system and is based on the comparison of synchronization between different harmonic waves.

Without feedback control, the muscle activation A is simulated with Eqs. (3) and (4). Based on the resulting time series, the harmonic waves $f_m(x), x \in [1,24], m = 5,7,8,\cdots,24$, are fitted using Eq. (5). The observation of harmonic waves over time shows that less than one third propagates approximately synchronously over a longer time interval. These are the waves with the muscle index $m = 5,9,17,20$ which serve as a reference for the others. When feedback control is applied, Eq. (4) is replaced by Eq. (6). The other harmonic waves with index $m = 7,8,10-16,18,19,21-24$ are then gradually calibrated to the reference. For each harmonic wave to be calibrated, all combinations of time delay and feedback strength resulting from the intervals [0.25,5] and [0.25,3], respectively, in steps of 0.25 are simulated. The feedback strength K_l is initially activated for all muscle cells that are related to the corresponding muscle index m of the considered harmonic wave, and the time delay τ_s is activated for all motor neurons which connect to these. Afterwards, the propagation of the time-delayed wave is compared with the reference over time. Whether or not the synchronization between them is suitable is assessed visually, but optimally it should look as in Fig. 9. There, one can see that the individual waves in green, orange, and yellowish color propagate nearly synchronized to the right over time. However, the parametrization is not straightforward because there are many interdependencies between the neuromuscular connections. When sufficient synchronicity is not observed, changes are made for time delays τ_s or feedback strengths K_l with respect to the motor neurons or muscle cells involved. For example, only motor neurons are considered whose neuromuscular connections overlap as little as possible with others. Likewise, only dorsal or ventral muscles can be included. Changing the parameters of feedback control causes the harmonic waves to propagate more or less synchronously over time. Depending on the parametrization, the wave to be calibrated may temporarily run ahead or behind the reference, or may overtake it. However, there is no precise rule for this, but it varies from case to case. During the calibration process, it has been found that the harmonic waves can be reasonably synchronized within the time interval [8,20]. In total, 35 motor neurons and 51 muscle cells are affected by the control force. The estimated parameters are provided in Tables D.8–D.9.

Fig. 10 illustrates the effect of time-delayed feedback control on the activation of muscles with index $m = 13$ and $m = 19$. The dorsal and ventral muscle activity exhibit most of the time a rather in-phase behavior. By activating the control, the oscillations of the dorsal and ventral muscles are more regular and almost in anti-phase.

In addition, Fig. 11 visualizes the propagation of 19 harmonic



(a) Uncontrolled **(b)** With feedback control

Fig. 11. Synchronization of harmonic waves with time-delayed feedback control. The elongation $f_m(x = m)$ [cf. Eq. (5)] is plotted over time t at position $x = m$ for each muscle pair with index $m = 5,7,8,\cdots,24$ in the absence [Eqs. (3) and (4)] and presence [Eqs. (3) and (6) – additional parameters are provided in Tables D.8-D.9] of time-delayed feedback control in panels (a) and (b), respectively. Without feedback control, no pattern of locomotion can be detected. With feedback control, the amplitude generated in the anterior body propagates almost linearly in time to posterior muscles, which makes it a phase-shifted synchronization. The more apparent the linear behavior, the better is the coordination of forward locomotion. For the muscles in the tail, the waves spread rather constantly in time. This is due to the fact that the same muscular connections are assumed for them [82].

waves in the absence and presence of feedback control. For each muscle pair with index $m = 5, 7, 8, \cdots, 24$, the elongation $f_m(x = m)$ [cf. Eq. (5)] is plotted over time up to $t = 21$ at position $x = m$. Without feedback control, no synchronization pattern can be identified. With feedback control, the amplitude generated in the anterior body indicated by yellow or dark blue color propagates almost linearly in time to posterior muscles in the considered time interval $[8, 20]$. Since the anterior waves show a similar behavior with a slight time lag, this is a phase-shifted synchronization. As the coordinated locomotion behavior of *C. elegans* becomes more probable, the more clearly the linearity can be seen in the synchronization plot, resembling a travelling wave. This is not the case for the muscles in the tail of the worm by design of the network. Actual high-power electron microscopes are not able to cover the neuromuscular connections in this region [16]. Therefore, the corresponding connections between motor neurons and muscle cells in the connectivity data are estimated and identical for muscles in the tail [82]. Fig. 9 illustrates this clearly. The temporal changes of muscle activation $\dot{A}$ indicated by bars display an almost identical behavior for the posterior muscles.

4.2. Kuramoto order parameter

In order to measure the synchronization between harmonic waves, we define the Kuramoto order parameter R as

$$R(t) = \frac{1}{W}\left|\sum_{m=5,7-24} \exp\left[i\gamma(x = 0, \Delta t_{r_m}, t)\right]\right|. \tag{7}$$

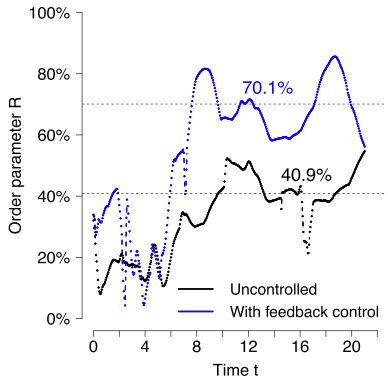


Fig. 12. Measuring synchronicity of harmonic waves during forward locomotion of *C. elegans*. The black and blue curve correspond to the Kuramoto order parameter given by Eq. (7) without [Eqs. (3) and (4)] and with feedback control [Eqs. (3) and (6) – additional parameters are provided in Tables D.8-D.9]. The dotted lines mark the average values in the time interval $[8, 20]$.

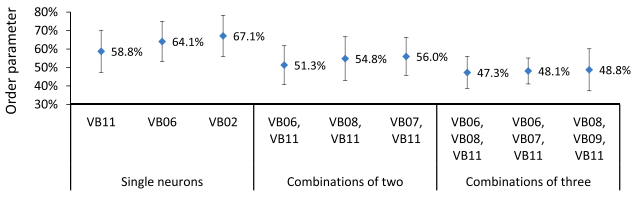
The order parameter represents the phase coherence of $W = 19$ waves. As explained earlier, the muscle pairs with index $m = 1 - 4, 6$ are not considered. The phase $\gamma(x, \Delta t_{r_m}, t)$ is equal to the argument of the sine function in Eq. (5). Since the phase differences between the harmonic waves are the same for each position on the x-axis, the position is fixed and can be set to any desired value. Here, we use $x = 0$ without loss of generality. $R = 1$ corresponds to full in-phase synchronization and is denoted by 100%, and $R = 0$ corresponds to complete desynchronization.

Based on the simulations of the harmonic waves, the order parameter (7) is calculated in time steps of 0.05 and considered in the time interval $[8, 20]$. The results are graphically displayed in Fig. 12. The synchronicity of the 19 harmonic waves expressed in the order parameter fluctuates over time around its average value. With time-delayed feedback control, the time-averaged order parameter is about 70.1%. This corresponds to an improvement of 29.2% over the uncontrolled system. Even if the time-averaged order parameter of 40.9% for the uncontrolled system may indicate that there is some partial synchronization between the harmonic waves, it is still at a level that does not facilitate a coordinated forward locomotion of *C. elegans* (cf. Fig. 11). The time-averaged order parameter of the controlled system serves as reference for simulations where neuronal activity is silenced.

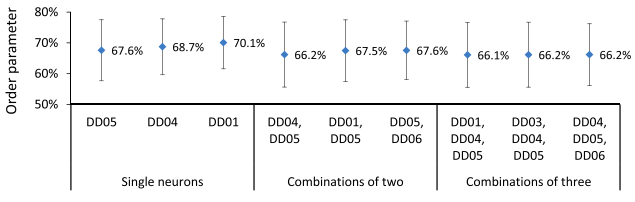
4.3. Silencing of neuronal activity

The starting point is the nearly synchronous propagation of harmonic waves in the presence of feedback control. If the activity of neurons is silenced, the synchronicity of locomotion of *C. elegans* is disturbed. This disturbance is considered in the time interval $[8, 20]$ and can be quantified by the time-averaged order parameter $\langle R(t)\rangle_t$. Note that the time interval $[0, 8]$ displays transient effects and is therefore disregarded. The neurons that cause the strongest decline of the time-averaged order parameter are the most significant. All neurons within relevant classes are successively silenced individually, in combinations of two, and combinations of three. Our silencing strategy is to hold the membrane potential at the constant value zero throughout the simulation. For the different selections, the top three results with the lowest time-averaged order parameter ± its standard deviation are considered.

Fig. 13 displays the results for the classes VB and DD. While silenced activity of VB neurons reduces the order parameter most, it causes the smallest changes for DD neurons. If neurons within the VB class are silenced individually, VB11, VB06, and VB02 have the strongest impact on the coordinated locomotion of *C. elegans*. In combination with these, the neurons VB07-VB09 become also relevant. The combined silencing of VB06, VB08, and VB11 leads to a significant decrease of the order parameter by more than 10% compared to the individual silencing. Therefore, the locomotion of the worm is highly uncoordinated. Within the DD class, the individual silencing of DD05, DD04, and DD01 shows a small to minor impact on the order parameter, which can only be marginally enhanced by the combined silencing of neurons. For exam-



(a) Silencing activity of VB neurons



(b) Silencing activity of DD neurons

Fig. 13. Most significant VB and DD motor neurons for synchronicity of harmonic waves. Without silencing of neuronal activity, the reference value of the time-averaged order parameter is 70.1%.

Table 5
Most significant motor neurons for a coordinated locomotion of *C. elegans* identified by silencing the neuronal activity of singles, doubles, and triples (see Figs. 13 and D.6–8).

Class	Neurons
AS (11)	AS01-AS03, AS05, AS06
DA (9)	DA01-DA05, DA07
DB (7)	DB01-DB04, DB06, DB07
DD (6)	DD01, DD03-DD06
VA (12)	VA04, VA08, VA10, VA12
VB (11)	VB02, VB06-VB09, VB11
VD (13)	VD02, VD09, VD10, VD12, VD13

ple, silencing DD05 individually leads to an order parameter of 67.6%, but silencing DD05 together with DD01 and DD04 only results in a reduction by 1.5%. Although the effect may seem very weak, the individual silencing of neuronal activity regarding DD04 or DD05 is already sufficient to impair the *C. elegans* locomotion [83]. In this case, the combined silencing of neuronal activity hardly provides additional information. The results for the other classes can be found in Figs. D.6-D.8. Table 5 summarizes the identified neurons in the top three single, double, and triple selection.

5. Discussion & conclusions

We have modeled the somatic nervous system of *C. elegans* as a multilayer network whose nodes comprise both neurons and muscle cells. The different layers are defined by different modes of interactions between the nodes and include important neurotransmitters and neuropeptides. This physiological approach allows for a better understanding of the network. On the one hand, it enables to conduct a logistic regression analysis in order to predict neurons involved in locomotion behavior. On the other hand, it allows to study the dynamics of complex circuits like the locomotory circuitry (Fig. 6), which initializes forward and backward locomotion of *C. elegans*.

5.1. Regression models predict the locomotory subnetwork with outstanding results

Logistic regression analysis can be operationalized in the multilayer network by considering the shortest paths between touch sensory neurons and body wall muscles. The starting point is the identification of key factors for the neurons in the shortest paths. This is based on the measurements of the discriminative power. Subsequently, logistic regression models can be built on them. As a result, we have explored various models with only two factors (cf. Table 3) that predict the shortest paths with high accuracy. The power values of the models lie between 96% and 99% (cf. Fig. 5) and indicate an outstanding separation ability. While the best model correctly predicts 99.2% of all the shortest paths on the development dataset, the second best model achieves about 96.9% (cf. Table 4). Both models correctly classify the shortest paths that only consist of neurons involved in locomotion behavior. A misclassification of less than 10% only occurs in pathways that also involve other neurons. In contrast to the development dataset, the shortest paths on the test dataset are not limited to 5 sensory neurons but consider all sensory neurons in the network. Based on the test dataset, one model predicts 29 additional neurons involved in locomotion behavior and therefore extends the initial selection. For 26 of them, it can be confirmed by literature research and connectivity analysis that these are indeed involved in locomotion behavior (cf. Tables C.6-C.7). For this reason, the models are well suited to predict the locomotory subnetwork. Moreover, the achieved results suggest that this method can also be used to predict other subnetworks. Since different subnetworks are characterized by their structure along with the underlying neuron and transmitter types, there is a good chance to identify them using logistic regression. This applies in particular to *C. elegans* since all neurons and their connectivity are known. Even if the information about neurons and connectivity is incomplete, logistic regression analysis is not necessarily useless as long as there is a tendency in the data. The procedure proposed in this study can be applied to identify circuits in the brain that control the movement of certain parts of the body since the information processing is the same in humans and *C. elegans*. In general, it could be very useful for studying larger biological neuronal networks than *C. elegans*. However, the application of this method is limited by the fact that sufficient information about the neurons must already be gathered. In conclusion, logistic regression analysis can be seen as a reasonable complement to optogenetic and ablation experiments.

5.2. Harmonic wave model describes forward locomotion

Concerning the dynamics, we have shown that the forward locomotion of *C. elegans* can be described using a harmonic wave model. The particularity of our model is that a harmonic wave is adjusted for each muscle pair consisting of a ventral and a dorsal muscle. This seems to be reasonable since the worms generate harmonic oscillations that spread from the head to posterior. The average value over the harmonic waves then represents the body curvature of *C. elegans* during locomotion.

Rhythmicity of locomotion is generated by neuronal oscillations – The neuronal basis is the locomotory circuitry (Fig. 6) in which the rhythmicity of alternating body bends is generated in order to drive forward locomotion. We have captured the main dynamical behaviors using the Hindmarsh-Rose equations. Our simulation results (cf. Fig. 8b) suggest that the sinus rhythm is network-based and results from couplings within the locomotory circuitry and interactions between different neuron types. Therefore, the circuit can be referred to as a central pattern generator. Since it initiates both forward and backward locomotion, oscillatory behavior can be observed in all classes of the first layer motor neurons. In the simulations, these oscillations are very robust against input of touch sensory neurons. With respect to a touch stimulus on the head or tail, only tiny phase shifts of them can be observed. This raises the question whether these tiny changes can already explain triggering forward and backward locomotion. Importantly, the expected patterns of neuronal activity are correctly reflected by the Hindmarsh-Rose system although it is basically a model for describing action potential behavior. Since no external current is applied to the neurons of the locomotory circuitry, the simulated dynamics shows pure coupling effects. Note that a current is only considered for the sensory neurons PLML/R, which perceive gentle touches on the tail. The set parameter value causes a fast oscillatory behavior rather than spiking behavior. However, this does not play a decisive role as they can only marginally influence the behavior of the other neurons due to the chosen coupling strengths. A higher current was applied because it had a slightly positive impact on the alignment of the harmonic waves. As a potential outlook, it would be worthwhile to examine the locomotory circuitry in more detail in order to understand exactly how the coupling must be designed to generate rhythmic oscillations. This could also play an important role in the brain. For example, the brainstem contains centers that control the rhythm for heartbeat and respiration. Since both are coupled together, the heartbeat can synchronize with the respiratory rhythm [55]. On the other side, there are neuromechanical models in the literature that can realistically reproduce the forward loco-

motion of *C. elegans* physical body [7,41]. However, these are based on a simplified neuronal circuit disregarding the asymmetry in neuron numbers. Together with other studies, those models suggest that proprioception is closely associated with the generation of body rhythms, but this does not exclude the coexistence with central pattern generators [22,23,26]. Our findings support the latter. However, we cannot rule out that even smaller units of the locomotory circuitry (Fig. 6) are capable of independent rhythm generation. For example, the study by [23, Fig. 9] suggests a multi-oscillator model for forward locomotion based on optogenetic and ablation experiments where two units of the first layer motor neurons generate rhythmic alternating body bends of *C. elegans*. Pre-motor interneurons activate or suppress this circuit. The propagation of these bends along the body is achieved by stretch receptor feedback. Since proprioceptive coupling mechanisms are not included in our model, we use time-delayed feedback control in order to coordinate the movements of the worm.

Anti-phase muscular activity is largely prevented – The generated neuronal harmonic oscillations are subsequently transferred to body wall muscles via neuromuscular synapses. Regarding the muscular activity, many problems occur that prevent a plausible description of locomotion behavior. The biggest problem is that the simulated activation of opposite dorsal and ventral muscles is usually not at a comparable level and does not show an anti-phase behavior, as illustrated in Fig. 10b. In addition, there is only little correlation between activities of the different muscle pairs along the midline of the worm. One reason for this is probably that the locomotory circuitry may still have missing parts. Logistic regression analysis has revealed that 26 additional neurons could be involved in locomotion behavior. This is especially true for the SMD and RMD motor neurons, which drive dorsoventral undulations in the head and neck of *C. elegans* and propagate them posteriorly through stretch-receptor feedback. Therefore, proprioceptive mechanisms also contribute to the coordination of locomotion [41]. Moreover, extrasynaptic neurotransmission between head motor neurons may also play an important role in ensuring optimal efficiency of forward locomotion [73,43]. Apart from that, the SMD and RMD motor neurons are involved in multiple navigation behaviors and can therefore be assigned to a circuit for navigation [28]. We have excluded navigation in our analyses because we wanted to examine first whether rhythmicity is generated in locomotory circuitry. Nevertheless, it cannot be excluded that a smooth coordinated locomotion of *C. elegans* at least depends on both circuits. The next step would be to analyze the dynamics if they are coupled together. Another important reason is that the neuromuscular connectivity data is far from perfect. The neuron-to-muscle connections in the tail of the worm are unknown, so the same connections are assumed for the latter muscles. Furthermore, the underlying number of synapses is estimated as an average value for many connections [82]. However, this does not properly account for the asymmetric structure of the motor neurons in the locomotory circuitry [77]. Besides, we have assumed that dorsal and ventral muscles lie in a row so that each dorsal muscle faces a ventral muscle. In fact, the muscle pairs behind the neck of the worm are not directly opposite to each other but slightly staggered along the main body axis [2]. As a consequence, the simulated muscle dynamics does not seem plausible across all muscles. For this reason, we have tried to improve the quality of the time series with scaling and smoothing efforts for visualization (Fig. 9).

Time-delayed feedback control enhances synchronicity of harmonic waves – In terms of the harmonic wave model on top of muscular activity, the aforementioned factors prevent a synchronous propagation of adjusted harmonic waves. In order to enhance synchronicity, we make use of time-delayed feedback control which is effective in the time interval $[8, 20]$. The overall synchronicity can be quantified with the time-averaged Kuramoto order parameter. As a result, we were able to increase the synchronicity by 29.2% to 70.1% (cf. Fig. 12). Although the initial value of 40.9% for the uncontrolled system may indicate some partial synchronizations between the harmonic waves, these do not contribute to a coordinated locomotion behavior of *C. elegans*. The effect of time-delayed feedback control can clearly be seen in the synchronization plots (Fig. 11). With the exception of muscle pairs in the tail of the worm, all adjusted harmonic waves propagate nearly synchronized but phase-shifted in time. This means that undulations generated in the anterior body of *C. elegans* spread linearly in time towards the tail, which describes a coordinated forward locomotion. Without feedback control, such a pattern cannot be detected. Therefore, the harmonic wave model in combination with time-delayed feedback control solves many problems faced by insufficient quality of connectivity data or considering only parts of the locomotory subnetwork. The approach also seems to be reasonable since the investigated neuronal activity supports the existence of different time delays for motor neurons. However, the downside of the model is that the parametrization in terms of complex circuits can prove to be very difficult.

5.3. Silencing of neurons reveals significance for coordinated locomotion behavior

Finally, the modeling of forward locomotion enables us to perform synchronicity analyses from which conclusions can be drawn about the significance of neurons within the locomotory circuitry. The synchronization of harmonic waves expressed in the time-averaged order parameter can be determined for different simulations where certain neurons are silenced. This in turn interferes with the coordination of forward locomotion of *C. elegans*. The stronger the disturbance, the more important are the silenced neurons. In all motor neuron classes, we have silenced neurons individually, in combinations of two, and combinations of three by holding their membrane potential at the constant value zero throughout the simulation. Note that this strategy still includes coupling effects with other neurons, but it should not weaken the relevance of the analysis. Since the neurons of the locomotory circuitry are essentially isopotential with the ability to produce regenerative responses [47], the constant value of zero may be due to a very high membrane resistance. A similar effect could be achieved using optogenetics where neuronal activity is manipulated with light in order to study the impact on behavior [17,40]. For specifying the most significant neurons, we consider the top three results for each selection and highlight all neurons included. Drawing a hard line regarding the order parameter does not capture important neurons properly because the measure can take on different levels within the different classes. In general, silenced VB and VA motor neurons have the strongest influence on the coordination of locomotion. These generate oscillatory behavior for ventral muscles. The DB and DA motor neurons are responsible for dorsal muscles. Since they share this role with the AS motor neurons, this may be one reason why the effect is slightly lower. In comparison, silencing the activity of DD and VD motor neurons leads to the smallest decrease of the order parameter, which changes only marginally between the different selections of singles, doubles, and triples. This could be due to the fact that these neurons generally exhibit a rather constant behavior (cf. Fig. 8a). Nevertheless, the individual silencing of their neuronal activity can already be sufficient to impair the locomotion of *C. elegans* [83]. Moreover, we have examined the interneurons in the locomotory circuitry in the same way. As a result, all interneuron classes are important for a coordinated locomotion behavior. Beyond that, similar statements can be made by applying network control principles to the *C. elegans* connectivity data [83]. Since our investiga-

tions are not only based on neuronal and muscular connectivity but also include the underlying dynamics of neurons and muscle cells, our findings should provide a good indication of significant neurons in the locomotory circuitry (cf. Figs. 13 and D. 6–8). However, it must be pointed out that these are reserved for the implemented parametrizations of the utilized models. The better the harmonic waves can be synchronized in advance, the more meaningful such results will be.

5.4. Conclusions

In summary, we have shown how a specific circuit of locomotion can be identified in the multilayer network of *C. elegans* using logistic regression. Furthermore, we have introduced various dynamical models and physical methods in order to understand the underlying patterns of neuronal and muscular activity during forward locomotion. This provides a good basis for explaining further circuits and understanding how these can be coupled with each other.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by Deutsche Forschungsgemeinschaft (DFG) in the framework of Collaborative Research Center 910. PH acknowledges further support by DFG under Grant No. HO4695/3–1. JR acknowledges support by the German Academic Exchange Service (DAAD) and by the National Agency for Research and Development (ANID): Scholarship Program DAAD/BECAS Chile, 2016 (57221134).

Appendix A. Data description and data preparation

A.1. Neuronal and muscular connectivity

Neuronal connectivity – The nervous system of the adult *C. elegans* hermaphrodite contains 302 neurons and can be divided into two nearly completely isolated systems: a large somatic nervous system (282 neurons) and a small pharyngeal nervous system (20 neurons). The wiring diagram of the somatic nervous system can be found on WormAtlas [82] and was provided by [78].

In the dataset, the numbers of electrical and chemical synapses are specified for each neuron pair. The electrical connections are labeled "EJ". For chemical synapses, the type of synapse must be taken into account:

- **Monoadic**: Send ("S") – Neuron 1 is presynaptic to Neuron 2.
- **Polyadic**: Send poly ("Sp") – Neuron 1 is presynaptic to more than one postsynaptic partner. Neuron 2 is just one of these postsynaptic neurons.

Nearly two thirds of the synapses are polyadic, but not all polyadic synapses have been faithfully marked as such. In this study, the type of synapse is not of interest. To get rid of the synapse type, the number of synapses must be aggregated (summed) over identical neuron pairs in the selection "Send" plus "Send poly". In total, the wiring diagram consists of 2,194 unidirectional chemical connections and 514 bidirectional electrical connections. Since the nervous system is modeled as a directed network, the number of electrical connections doubles to 1,028.

Table A.1
Data basis connectivity data. The somatic nervous system of *C. elegans* is modeled as a directed network whose nodes represent neurons and body wall muscles connected by chemical synapses, gap junctions, and neuromuscular junctions. The multilayer network in Fig. 1 has a total of 3,538 distinct connections because chemical and electrical connections overlap.

Connection type	Number of connections	Number of synapses
chemical	2,194	6,394
electrical	1,028	1,774
muscular	548	1,791
Total	**3,770**	**9,959**

The numbers of chemical and electrical synapses (gap junctions) are 6,394 and 1,774 in total (cf. Table A.1). Note that the connections are made by 279 neurons. The neurons CANL, CANR, and VC06 are excluded since they do not have connections with other neurons.

Neuromuscular connectivity – If the worm is cut from above (dorsally) along the midline and splayed out laterally, four muscle quadrants can be identified. From left to right, these are the quadrants: dorsal left, ventral left, ventral right, and dorsal right. Each quadrant contains 24 muscle cells with the exception of the ventral left quadrant which contains 23 cells. This results in a total of 95 muscle cells. Within each quadrant, the muscles lie in double rows and are numbered from head to tail. The first 16 muscles across the four quadrants belong to the head, the next 16 to the neck, and the rest of them to the body. While the muscles lie side by side in the head, they are slightly staggered along the main body axis in the remaining regions ([2], [50, Section III]). Neuron-to-muscle connections are also available on WormAtlas [82] and based on the works of [18,78,80]. It is assumed that body wall muscles are activated by motor neurons. In total, 552 motor neurons make connections to body wall muscles. The connections of the neurons CEPVL, CEPVR, ADEL, and AVKR are not considered because they do not function as motor neurons. The total number of neuromuscular synapses is 1,791 (cf. Table A.1). Note that a large proportion of neuromuscular synapses is based on average values and that the neuron-to-muscle connections in the tail of the worm are only estimated [82].

A.2. Neuron functions

In *C. elegans*, three main types of neurons can be distinguished: sensory neurons, interneurons, and motor neurons. This information can be extracted from [53]. Note that it is also available on WormAtlas [82] but must be looked up individually. In general, neurons can have multiple functions and combine the properties of all three main types. The dataset contains up to two functions whereby the first specified function is assumed to be the main

Table A.2
Neuron functions in *C. elegans*. The somatic nervous system contains 282 neurons in total, which are classified into sensory neurons, interneurons, and motor neurons.

Neurons that function as	Number	also function as	Number
Sensory neurons	82	-	73
		Interneurons	3
		Motor neurons	6
Interneurons	89	-	78
		Motor neurons	7
		Sensory neurons	4
Motor neurons	111	-	83
		Interneurons	24
		Sensory neurons	4
		Total	**282**

function. There are a total of 82 sensory neurons, 89 interneurons, and 111 motor neurons in the somatic nervous system of *C. elegans*. Six of the sensory neurons and seven of the interneurons also operate as motor neurons resulting in a total of 124 neurons with the function of a motor neuron (cf. Table A.2). We changed the function of the neurons PVDL and PVDR from interneuron to sensory neuron. For the neurons RIML and RIMR, the function as interneuron is additionally assigned to the function as motor neuron [82].

A.3. Neurotransmitter and neuroreceptor data

By considering actual transmitter and receptor data collected for the individual neurons, the chemical layer in the network can be divided into several sublayers. Besides the classical transmitters acetylcholine (ACh), glutamate (Glu), and *gamma*-aminobutyric acid (GABA), we consider different monoamine transmitters (MA) and neuropeptides. Table A.3 provides all transmitters and receptors utilized in this study.

The data for the MAs can be found by Bentley et al. [5]. Hypothetical dopamine receptors, such as *dop-5* and *dop-6*, are also taken into account to increase the number of connections in the network. In total, there are 20 neurons with MA transmitters and 232 neurons with MA receptors in network. The data for the classical transmitters ACh, Glu, and GABA have been collected by ourselves. There are a total of 218 neurons with classical transmitters and 210 neurons with matching receptors. Another important class of neurochemicals are neuropeptides, which are also provided by [5]. In total, there are 164 neurons with neuropeptides and 185 neurons with corresponding receptors. In this study, we use the transmitter and receptor data including the subset of peptides to map the chemical connections with respect to the underlying transmitter types.

Note that for MAs and neuropeptides large extrasynaptic signalling networks exist in *C. elegans*. Extrasynaptic connections are generated by the diffusion of neurotransmitters and neuropeptides to adjacent (extracellular) synapses. Such connections occur mainly outside the synaptic connectome and are referred to as wireless connections. It is well established that they play an important role in brain function [5]. The inclusion of extrasynaptic neurotransmission would add many more layers to the network, which is not done.

A.4. Mapping of chemical connections

Only $1,529$ connections (about 70%) of the $2,194$ chemical connections can be covered with the utilized transmitter and receptor

Table A.3
Neurotransmitter and neuroreceptor data. For each transmitter type, the corresponding receptors are specified.

Class	Family	Type	Matching receptors	Source
Transmitters	Monoamines	Dopamine	dop-1, dop-2, dop-3, dop-4, dop-5, dop-6, lgc-53	Bentley et al. (2016)
		Octopamine	octr-1, ser-3, ser-6	
		Serotonin	mod-1, ser-1, ser-4, ser-5, ser-7	
		Tyramine	lgc-55, ser-2, tyra-2, tyra-3	
	Classical transmitters	Acetylcholine	acc-1, acc-2, acc-4, acr-12, acr-14, acr-15, acr-16, acr-18, acr-2, acr-23, acr-5, deg-3, des-2, gar-1, gar-2, gar-3, lev-8, lgc-12, lgc-27, lgc-46, unc-29, unc-63	Jorge Ruiz (2017)
		gamma-Aminobutyric acid	exp-1, gab-1, gbb-1, ggr-1, ggr-2, lgc-35, lgc-37, lgc-38	
		Glutamate	avr-15, glc-3, glr-1, glr-2, glr-3, glr-4, glr-5, glr-6, glr-8, mgl-1, mgl-3, nmr-1, nmr-2	
Peptides	FLPs*	flp-1	npr-11, npr-4	Bentley et al. (2016)
		flp-4	npr-4	
		flp-5	npr-11	
		flp-10	egl-6	
		flp-13	frpr-4	
		flp-15	npr-3	
		flp-17	egl-6	
		flp-18	npr-1, npr-11, npr-4, npr-5	
		flp-21	npr-1, npr-11, npr-2, npr-5	
		flp-24	npr-17	
	NLPs*	nlp-1	npr-11	
		nlp-12	ckr-2	
	NTCs*	ntc-1	ntr-1	
	PDFs*	pdf-1	pdfr-1	
		pdf-2	pdfr-1	

* FLP – FMRFamide-like peptide.
NLP – Neuropeptide-like protein.
NTC – Nematocin (Oxytocin/Vasopressin-related peptide).
PDF – Pigment-dispersing factor.

data. For the remaining 665 connections (about 30%), either the information about the transmitter or the receptor is missing, or the transmitters and receptors are not compatible. To have a more complete view of the network of *C. elegans*, the missing transmitter types are estimated for these connections. Furthermore, it is assumed that the estimated transmitters are valid for all neurons of the same class [82].

If the transmitter information for one neuron is missing, the least common set of transmitters is searched for all postsynaptic receptors in order to establish at least one connection with all unassigned postsynaptic neurons. If several options remain afterwards, not all transmitter types are assigned but preferred ones are chosen. Sometimes a decision can be made with the help of WormAtlas [82] by studying the transmitter and receptor information for the individual neurons. If still no result can be found, the transmitter with the highest probability of occurrence is utilized (compare exemplary with Fig. A.1). Two examples are given below:

- **Example 1**: Neuron A has no transmitter information, neuron X has the ACh receptor *acc-1* as well as the Glu receptor *glr-1*, neuron Y has the dopamine receptor *dop-1*, and neuron Z has the ACh receptor *acc-2*. The least common set of transmitters for connecting with the neurons X, Y and Z is given by ACh and dopamine, which are assigned to Neuron A.
- **Example 2**: If neuron Z from the first example has also a Glu receptor like *glr-2*, no unique decision can be made about the transmitters ACh and Glu. If no solution can be found with the help of WormAtlas, frequency distributions are considered. In this case, the transmitter ACh has the higher probability of occurrence and is assigned to Neuron A along with dopamine.

The same procedure is used when the receptor information is missing for specific neurons. This time, the least common set of transmitters among all presynaptic neurons is desired. For the underlying transmitters, no specific receptors can be determined. It can only be indicated that for certain transmitters there must be at least one matching receptor.

Finally, the last approach is also applied in the case that the information for transmitters and receptors does not correspond. In general, neurons have significantly fewer transmitters than receptors. In Table A.3, it can be seen that there is a huge variety for the latter. In a number of neurons, many receptors do not play a role in the formation of synaptic connections since they can only couple with transmitters that are not released. However, these could be important for extrasynaptic connections, but this is not considered. In the case of synaptic transmission, it is more likely that the existence of receptors for specific transmitters can be indicated than vice versa. Therefore, transmitters are considered and receptors are assigned. Our approach for estimating transmitter types for chemical connections does not claim to be the best solution, but it can be carried out without much effort, and the results are plausible. The findings are provided in Tables A.4 and A.5.

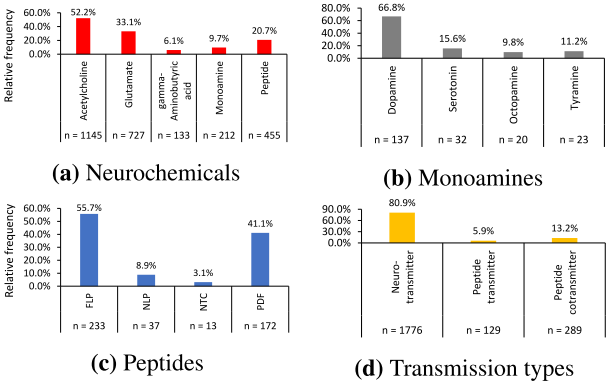


Fig. A.1. Frequency distributions for the chemical connections. (a) Relative frequency of different neurochemicals on a total of 2,194 chemical connections (overlap is 21.8%). (b) Relative frequency of the monoamine transmitters on a total of 205 monoamine connections (overlap is 3.4%). (c) Relative frequency of the peptide families on a total of 418 peptide connections (overlap is 8.9%). (d) Relative frequency of neurotransmission in general.

Table A.4
Estimated transmitters for chemical connections.

Neuron class	Predicted transmitters	Neuron class	Predicted transmitters
AVH (2)	ACh, flp-21, flp-24, Glu, ntc-1	PVM (1)	dopamine, GABA, Glu
AVJ (2)	ACh, Glu	PVW (2)	Glu, ntc-1, serotonin
AWA (2)	dopamine, flp-21, Glu, tyramine	RID (1)	ACh
		RIP (2)	ACh

Table A.5
Estimated transmitter receptors for chemical connections.

Neuron class	Predicted receptors (unspecified) for	Neuron class	Predicted receptors (unspecified) for
ADA (2)	Ach, Glu, octopamine, pdf-1	OLL (2)	Ach, GABA, Glu
ADE (2)	Glu, nlp-1	OLQ (4)	dopamine
ADF (2)	ACh	PDA (1)	Ach
ADL (2)	Ach, dopamine	PDB (1)	ACh
AFD (2)	Ach, Glu	PDE (2)	Ach, flp-1, GABA, Glu
AIA (2)	Ach	PHC (2)	GABA, Glu
AIB (2)	Ach, flp-21, pdf-1	PLN (2)	ACh
AIM (2)	Ach, Glu, pdf-1	PVC (2)	GABA
AIN (2)	Glu	PVM (1)	dopamine, flp-1
AIZ (2)	Glu	PVN (2)	ACh, flp-10, GABA, Glu
ALA (1)	dopamine, Glu	PVP (2)	Ach, GABA, Glu
ALN (2)	Glu, pdf-1	PVR (1)	ACh, dopamine, Glu
AQR (1)	ACh	PVT (1)	Glu
AS (11)	Glu	PVW (2)	ACh, GABA
ASE (2)	Ach	RIB (2)	flp-21, GABA
ASG (2)	Ach, Glu	RIC (2)	Ach
ASH (2)	Ach	RID (1)	Glu, pdf-1
ASJ (2)	ACh, GABA, Glu	RIF (2)	Ach
ASK (2)	Ach, GABA, Glu	RIG (2)	dopamine, pdf-1
AUA (2)	Ach	RIH (1)	Glu, pdf-1
AVA (2)	dopamine, nlp-1, octopamine	RIM (2)	flp-1, GABA, octopamine
AVB (2)	flp-1	RIP (2)	ACh, dopamine, Glu
AVD (2)	Ach	RIR (1)	Ach, pdf-1
AVE (2)	dopamine, flp-1, GABA	RIS (1)	ACh
AVF (2)	GABA, Glu	RIV (2)	ACh, dopamine, Glu, octopamine
AVG (1)	GABA		
AVH (2)	Ach, nlp-1	RMD (6)	flp-1, GABA, pdf-1
AVJ (2)	ACh, dopamine, GABA, nlp-1, pdf-1	RMF (2)	ACh, flp-1, Glu, octopamine, pdf-1
AVK (2)	ACh, dopamine, GABA, octopamine, pdf-1	RMG (2)	Ach
		RMH (2)	Ach, dopamine, Glu, pdf-1
AVL (1)	Ach, GABA, Glu, pdf-1		
AVM (1)	flp-1	SAA (4)	dopamine, flp-1, nlp-1
AWA (2)	Ach, Glu	SAB (3)	Ach, GABA
AWB (2)	ACh	SDQ (2)	Glu
AWC (2)	Ach, flp-21	SIA (4)	flp-1, Glu
BAG (2)	Ach	SIB (4)	pdf-1
BDU (2)	ACh	SMB (4)	dopamine, flp-1, Glu, octopamine, pdf-1
CEP (4)	GABA, Glu, pdf-1		
DA (9)	Glu	SMD (4)	flp-1, octopamine
DD (6)	Glu	URA (4)	Ach, nlp-1
DVA (1)	flp-1	URB (2)	Ach, dopamine
DVC (1)	GABA, pdf-1	URX (2)	Glu
IL1 (6)	dopamine, Glu	URY (4)	ACh, GABA
IL2 (6)	Glu	VA (12)	Glu
LUA (2)	Ach	VD (13)	GABA, Glu

Using the estimated transmitters and receptors, all chemical connections in the network are mapped with at least one type of transmitter. The result is illustrated in Fig. A.1. The chemical layer consists of 2,194 neuron connections. About 52% of the connections are covered by the transmitter ACh, which represents the largest sublayer. Then follows Glu with about 33% and the group of the MAs with almost 10% (cf. Fig. A.1a). In the group of the MAs, the dopamine transmitter has the largest proportion with about 67% (cf. Fig. A.1b). The dopamine sublayer is slightly larger than the GABA sublayer, which occupies about 6% of the chemical connections. Although the peptides have a larger proportion at around 21%, only about 31% of them act as peptide transmitters (cf. Fig. A.1d). The sublayer of peptide transmitters is comparable to the sublayer of GABA. The remaining 69% of the peptides are cotransmitters which always occur together with neurotransmitters. Among all pepdites, the FLP and PDF family are the most frequently encountered with about 56% and 41% (cf. Fig. A.1c).

Appendix B. Functions of layers in the network

This appendix provides details on the different layers depicted in Fig. 1b-h that give rise to the overall network shown in Fig. 1a.

Acetylcholine layer – ACh covers 33.4% of the 3,538 connections in the network and forms the largest layer. It is widely spread among all neurons and does not prefer a specific neuron type. In humans, ACh acts at skeletal neuromuscular junctions, at neuromuscular junctions between the vagus nerve and cardiac muscle fibers, and at a variety of locations within the central nervous system. While the function of ACh at neuromuscular junctions is well known, its role in the central nervous system is not well understood [60, Chapter 6]. In *C. elegans*, ACh is involved in many behaviors like locomotion [81,32], egg-laying [4], feeding [62], and defecation [76].

Glutamate layer – Glu uses 20.5% of the connections in the network and is therefore the third largest layer. In comparison with the ACh layer, significantly fewer motor neurons are involved. In the human brain, Glu is the most important transmitter for normal brain function. It is estimated that more than half of all brain synapses release this substance [60]. In *C. elegans*, it contributes to foraging behavior [34], long-term memory [65], and spontaneous switches from forward to backward movement referred to as reversals [85,9].

***gamma*-Aminobutyric acid layer** – GABA represents the smallest layer in the network and is released on 3.8% of the connections. Most of them are established by interneurons and motor neurons. About one third of the synapses in the brain use GABA as their neurotransmitter, which is most frequently found in interneurons of local circuits. In contrast to ACh and Glu, GABA has an inhibitory effect [60, Chapter 6]. In *C. elegans*, it can also act as excitatory transmitter which depends on the neuroreceptor. As an inhibitory transmitter, GABA regulates the head movements while foraging [80] or relaxes muscle cells during locomotion [48].

Monoamine layer – MA transmitters have a proportion of 5.8% of the network connections. These are independent of the neuron type, but most of the motor neurons are predominantly postsynaptic. In the group of MAs, dopamine accounts for about two thirds (cf. Fig. A.1b). In humans, MAs regulate many brain functions and are also present in the peripheral nervous system. They are entangled in a wide range of behaviors that range from central homeostatic functions to cognitive phenomena, such as attention. The transmitters play an important role in the brain because defects in the MA function can lead to psychiatric disorders. The dopamine transmitter, for instance, is necessary for the coordination of body movements. A degeneration of particular dopaminergic neurons can lead to characteristic motor dysfunction as it is the case with parkinson's disease [60, Chapter 6]. In *C. elegans*, the MAs affect a variety of behaviors including egg-laying, pharyngeal pumping, locomotion, and learning. A good overview is given by [12]. The dopamine transmitter, for example, is responsible for the modulation of locomotion behavior and for learning. The modulation of locomotion behavior enables the worms to react on environmental changes [68] and to search efficiently for new food sources [34]. Learning allows the worms to change their behavior based on previous experience. For example, the animals react to a non-localized mechanical stimulus such as plate tapping by either moving backwards or increasing their forward locomotion rate. Repeated tapping on the plate causes the worms to become habituated to the stimulus, and they exhibit a reduced frequency of reversals [66].

Peptide layer – Peptides are utilized in 11.8% of the network connections, which are primarily established by interneurons and sensory neurons. Most of the peptides originate from the FLP and PDF family. About 31% of the peptides are peptide transmitters and about 69% are cotransmitters (cf. Fig. A.1c and d). The number of peptides in humans is estimated to be over 1,000, over 100 have already been identified. Peptide transmission is involved in the perception of pain, modulation of emotions, and regulation of complex reactions to stress. Peptide cotransmission enhances or dampens synaptic activity and can influence many functions such as food intake, metabolism, social behavior, learning, and memory ([67], [60, Chapter 6]). In *C. elegans*, peptides affect many behaviors including locomotion, dauer formation, egg-laying, sleep, learning, social behavior, mechano-, and chemosensation. Numerous peptides of the FLP family are involved in feeding behavior [45]. On the other side, peptides can fulfill a unique functions as *nlp-22* is a regulator of *C. elegans* sleep-like state (lethargus) during a larval transition stage [52]. For this reason, neuropeptides should be considered as the third layer of information flow in neuronal communication next to chemical and electrical transmission. This study only uses a small subset of peptides found in *C. elegans*, and many of their functions are still unknown. The actual number of neuropeptides in the worm exceeds 300 [64,14]. Like monoamines, neuropeptides have a small wired network but a large wireless network [5]. This raises the opportunity that neuropeptides could be involved in all *C. elegans* behaviors.

Electrical layer – Electrical transmission accounts for 29.1% of the network connections and is therefore the second largest layer behind ACh. Electrical synapses are found in all nervous systems. They enable a direct, passive flow of electrical current from one neuron to another. In contrast to chemical synapses, the current can flow in both directions (bidirectional) and the transmission is extraordinarily fast (virtually instantaneous). Communication is possible without delay, which is not typical for chemical synapses. For this reason, they are found in places where quick actions are necessary and have the general purpose of coordinating and synchronizing network activity among neuron groups. For example, certain neurons in the brainstem are synchronized by electrical synapses to produce rhythmic breathing. The same applies to populations of interneurons in the cerebral cortex, thalamus, and cerebellum [60, Chapter 5]. In *C. elegans*, electrical transmission plays an important role in locomotion behavior and development [31,74].

Neuromuscular layer – Neuromuscular junctions are comparable with chemical synapses and have a proportion of 15.5% of the network connections. They are made from motor neurons to body wall muscles. In *C. elegans*, the transmitters ACh and GABA can be assigned to it. While ACh stimulates muscle contraction, GABA relaxes muscles cells [63,42]. In humans, it is well known that ACh is released by spinal motor neurons and results in contraction of skeletal muscles [60, Chapter 5].

Appendix C. Logistic regression analysis

Logistic regression analysis is a statistical procedure used for classification problems. An application-oriented introduction to the methods of logistic regression can be found in [59, Chapter 4]. In this study, we introduce the binary model.

C.1. Binary logistic regression model

The binary model considers a response variable (random number) which is denoted as Y and accepts the two values 0 and 1. The expectation value $\mathbb{E}(Y)$ lies in the interval $[0,1]$. By applying a response function

$$F(z),\ z\in\mathbb{R}, \text{with values in the interval} [0,1],$$

the linear combination $z = c_0 + c_1x_1 + c_2x_2 + \ldots + c_mx_m$ of m explanatory variables $x_1, x_2, \ldots, x_m$ is restricted to the same interval whereby c_0 is a constant and $c_1, c_2, \ldots, c_m$ represent coefficients. This results in the following estimate

$$\mathbb{E}(Y) = \mathbb{P}(Y = 1) = F(z).$$

The binary model possesses as response function the logistic function (also known as sigmoid function)

$$F(t) = \frac{1}{1+\exp(-t)}, \quad t\in\mathbb{R}. \tag{C.1}$$

This function has a nonlinear s-shaped curve that asymptotically approaches the values 0 and 1 from above and below, and its values can be interpreted as the probability that the response variable equals one.

For the observation i, the values of the response variable Y (0 or 1) and the explanatory variables $x_1, x_2, \ldots, x_m$ can be specified:

$$Y_i, x_{1i}, x_{2i}, \ldots, x_{mi}, \quad i = 1, 2, \ldots, n.$$

The linear regression term for observation i is

$$z_i = c_0 + c_1x_{1i} + c_2x_{2i} + \ldots + c_mx_{mi}, \quad i = 1, 2, \ldots, n.$$

The random variables $Y_1, Y_2, \ldots, Y_n$ are assumed to be independent. The $m+1$ dimensional vector of the unknown parameters is designated as $\boldsymbol{c} = [c_0, c_1, c_2, \ldots, c_m]^T$. The probability for the occurrence of $Y_i = 1$ is abbreviated with

$$\sigma_i(\boldsymbol{c}) = \mathbb{P}(Y_i = 1).$$

Then, according to the above estimate, the binary logistic regression model reads as [59, Chapter 4]

$$\sigma_i(\boldsymbol{c}) = F(z_i(\boldsymbol{c})) = \frac{1}{1+\exp(-z_i(\boldsymbol{c}))}, \quad i = 1, 2, \ldots, n. \tag{C.2}$$

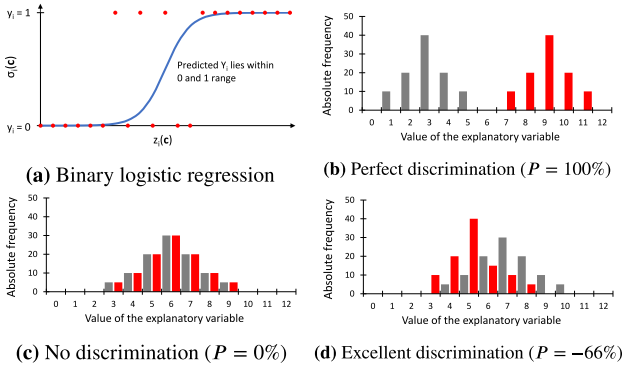
Since logistic regression predicts probabilities and not just classes, the unknown parameters c_j can be estimated using the maximum likelihood method. This method maximizes the probability of the parameters for the given data. The response probability p_i depending on realizations (0 or 1) of the response variable y_i satisfies

$$p_i(y_i) = \begin{cases} \frac{1}{1+\exp(-z_i(\boldsymbol{c}))}, & \text{if } y_i = 1 \\ 1 - \frac{1}{1+\exp(-z_i(\boldsymbol{c}))}, & \text{if } y_i = 0 \end{cases},$$

which can be summarized as

$$p_i(y_i) = \left(\frac{1}{1+\exp(-z_i(\boldsymbol{c}))}\right)^{y_i}\left(1 - \frac{1}{1+\exp(-z_i(\boldsymbol{c}))}\right)^{1-y_i}.$$

For all i observations together, the probability theorem for independent events can be used to construct the likelihood function which needs to be maximized:



(a) Binary logistic regression **(b)** Perfect discrimination ($P = 100\%$)

(c) No discrimination ($P = 0\%$) **(d)** Excellent discrimination ($P = -66\%$)

Fig. C.2. Logistic regression model and example distributions with corresponding discriminative power. (a) Binary logistic regression. The logistic function σ_i is illustrated as blue line. Realizations of the response variable y_i are displayed as red circles. (b)-(d) Example distributions of an explanatory variable with corresponding discriminative power. The empirical distribution $X_{Y=1}$ ($X_{Y=0}$) is indicated by red (grey) bars.

$$L(\boldsymbol{c}) = \prod_{i=1}^{n}\left(\frac{1}{1+\exp(-z_i(\boldsymbol{c}))}\right)^{y_i}\cdot\left(1 - \frac{1}{1+\exp(-z_i(\boldsymbol{c}))}\right)^{1-y_i} \overset{!}{=} \max.$$

Taking the natural logarithm yields the log-likelihood function

$$\begin{aligned}\mathscr{L}(\boldsymbol{c}) = &\sum_{i=1}^{n}\left[y_i \ln\left(\frac{1}{1+\exp(-z_i(\boldsymbol{c}))}\right)\right] \\ &+\left[(1-y_i)\ln\left(1 - \frac{1}{1+\exp(-z_i(\boldsymbol{c}))}\right)\right],\end{aligned} \tag{C.3}$$

which has identical extreme values but is easier to calculate. In many program packages, the maximization is performed by the Newton–Raphson algorithm where the zero point is approximated by iteration [75, Section 3.3].

To evaluate the fitted model, the log-likelihood is often multiplied by -2. The negative twofold log-likelihood is approximately χ^2 - distributed with $n-(m+1)-1$ degrees of freedom where n is the number of observations and $m+1$ the number of parameters. The expression $-2\mathscr{L}$ is denoted as deviance and is comparable to the residual sum of squares of the linear regression. With a perfect fit, the deviance is zero. Under the null hypothesis "The model has a perfect fit", the fit is the better the lower the value $-2\mathscr{L}$ [75, Section 3.5]. With the fitted model, the values 0 or 1 can be assigned to the response variable Y in dependence on a threshold value. The binary logistic regression model is depicted in Fig. C.2a.

The requirement for creating a logistic regression model is that the explanatory variables should not be highly correlated with each other because this can cause estimation problems [6].

C.2. Power analysis to identify key factors

The quality of a regression model is determined by its ability to distinguish correctly between realizations (0 or 1) of the response variable y_i. To figure out which explanatory variables x_k have a high explanatory contribution to the model (C.2), the power P is introduced. This measure lies in the interval $[-1,1]$ and can be defined as

$$\begin{aligned}P &= \frac{1}{N_{Y=1}N_{Y=0}}\sum_{l=1}^{N_{Y=1}}\sum_{m=1}^{N_{Y=0}} S\left(x_{l,Y=1}, x_{m,Y=0}\right), \\ S\left(x_{l,Y=1}, x_{m,Y=0}\right) &= \begin{cases} 1, & \text{if } x_{l,Y=1} > x_{m,Y=0} \\ 0, & \text{if } x_{l,Y=1} = x_{m,Y=0} \\ -1, & \text{if } x_{l,Y=1} < x_{m,Y=0} \end{cases},\end{aligned} \tag{C.4}$$

where $x_{l,Y=1}$ and $x_{m,Y=0}$ are realizations of the empirical distributions $X_{Y=1}$ and $X_{Y=0}$ with numbers of observations $N_{Y=1}$ and $N_{Y=0}$.

The power can be derived from the receiver operating characteristic (ROC) and the cumulative accuracy profile (CAP). Both ROC and CAP are important concepts to visualize the discriminative power (separation ability) of a model. They are applied in many scientific disciplines, for instance, in biology, information technology, and engineering sciences. The concepts convey the same message but present it in different ways. The information contained in a ROC or CAP curve can be aggregated into a single number, the area under the ROC curve (AUROC) or the accuracy ratio (AR). The AR can be interpreted as a simplified representation of AUROC since

$$AR = 2 \cdot AUROC - 1 \tag{C.5}$$

and is also known as Gini coefficient or power statistics. Since the Mann–Whitney statistics can be introduced as an equivalent to the area under the ROC curve [21, Chapter 13], we obtain the power in (C.4) using (C.5). Eq. (C.4) indicates that P can be expressed in terms of probabilities

$$\mathbb{E}(P) = \mathbb{P}(X_{Y=1} > X_{Y=0}) - \mathbb{P}(X_{Y=1} < X_{Y=0}), \tag{C.6}$$

which allows for an intuitive interpretation.

If all observed values of the distribution $X_{Y=1}$ are larger than those of the distribution $X_{Y=0}$, the power P is equal to 100%. The realizations (0 or 1) of the response variable y_i can be perfectly separated (cf. Fig. C.2b). The same applies if both distributions switch the sides so that the power is now −100%. If both distributions overlap completely, the power is zero, and the values of the response variable cannot be separated with logistic regression (cf. Fig. C.2c). As a third example, a case is shown where the distributions overlap slightly resulting in a power of −66%. The minus sign indicates that the distribution $X_{Y=1}$ has more observations with lower values than the distribution $X_{Y=0}$. The absolute power value of 66% indicates an excellent regression model (cf. Fig. C.2d).

The power of a logistic regression model can be classified as follows [39, Section 5.2.4]:

- If $P = 0\%$, the model cannot separate between observed values (0 or 1) of the response variable y_i (predicted probabilities are pure random, like a coin flip).
- If $0\% \leq |P| < 40\%$, the model is considered to have a poor separation ability.
- If $40\% \leq |P| < 60\%$, the model is considered to have an acceptable separation ability.
- If $60\% \leq |P| < 80\%$, the model is considered to have an excellent separation ability.
- If $|P| \geqslant 80\%$, the model is considered to have an outstanding separation ability.

It should be noted that in practice it is extremely unusual to observe absolute power values greater than 80%. Since the power P measures the discriminative power of potential explanatory variables, it is well suited to identify key factors for the regression model. In this case, it is extremely unusual to observe absolute power values greater than 40%. At the lower end, factors with $|P| > 5\%$ (on a significance level of 5%) can also be important for the regression model. Those factors can significantly increase the power of the regression model in combination with other uncorrelated factors. Please note that factors with a negative power value do not restrict the application of logistic regression. For the regression model, they simply mean negative coefficients. However, it should be checked whether a negative relationship to the values of the response variable is plausible.

From a statistical point of view, the power P can be interpreted as the probability to uncover a difference when there really is one. On a significance level of 5%, a power value of $|P| > 5\%$ indicates that the null hypothesis "There is no difference" must be rejected. To get a feeling for the accuracy of the calculated power (C.4), it is typical to specify a confidence interval. An efficient method to compute the 95% confidence intervals based on the Mann–Whitney statistic is provided by [21, Chapter 13]. This method becomes increasingly important for small sample sizes with a relatively small number of observations for $y_i = 1$ or $y_i = 0$. The uncertainty is considered more appropriately by significantly broader confidence intervals compared with other program packages.

C.3. Complementary univariate power analyses for developing logistic regression models

Interneuron connectivity – The results for considering the outdegree of the interneurons are depicted in Fig. C.3. For electrical connections, the power value of 75% is the same as in the indegree analysis (Fig. 3a) since they are bidirectional. Only ACh exhibits a higher power with 83.3%. The combined oudegree of both transmission types increases the power by 11.5% to 94.8%, which is not far from a perfect model. For the other transmission types, no statement can be made due to low values or high uncertainty of the power.

First layer motor neuron connectivity – The outdegrees of the first layer motor neurons reveal one factor with excellent power of −75%. The motor neurons of the locomotory subnetwork $LS_p = 1$ have no outgoing connections with peptide cotransmission. In combination with GABA transmission, the power can be increased to absolut 79.9% (see Fig. C.4a).

Qualitative view on the first layer motor neurons – Instead of focusing on the connectivity of neurons in the network, it is also possible to target just the presence of specific transmitters. For example, the presence of certain transmitters among neurons can be binary coded. In the simplest case, 1 stands for yes and 0 for nNo. This is done for ACh along with peptides of the families NLP and PDF. The results are plotted in Fig. C.4b. The factor "PDF peptide or not" has already a good single power of −62.5%. The neurons of the locomotory subnetwork do not have PDF peptides, which can be seen in the distribution. The same is true for NLP peptides. If NLPs and PDFs are subtracted from ACh, the overlap of neurons of the locomotory subnetwork $LS_p = 1$ and other neurons $LS_p = 0$ can be minimized resulting in a power of 81.9%. This power value lies between the absolute values of the respective best factor from the indegree and outdegree analysis with 87.5% and 79.9% (see Figs. 3b and C.4a). Note that no qualitative consideration has been made for the interneurons since they can be perfectly distinguished by their connectivity. However, this could also work fine for them.

Second layer motor neuron connectivity – Last, the second layer motor neurons are investigated (see exemplary Fig. 3b). It is assumed that the interneurons and the first layer motor neurons of the locomotory subnetwork are known. Only the motor neurons are considered that are postsynaptic to these of the first layer. In total, there are 66 motor neurons of which 60 belong to the locomotory subnetwork. New in this group are the motor neurons of the classes DD and VD with the exception of DD01, VD03, VD10,

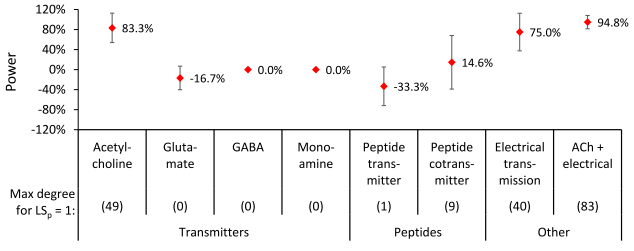
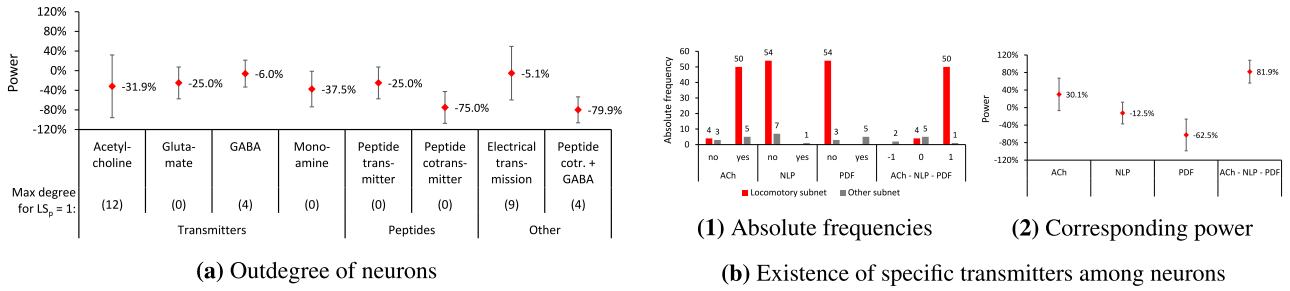


Fig. C.3. Outdegree power values for the interneurons.

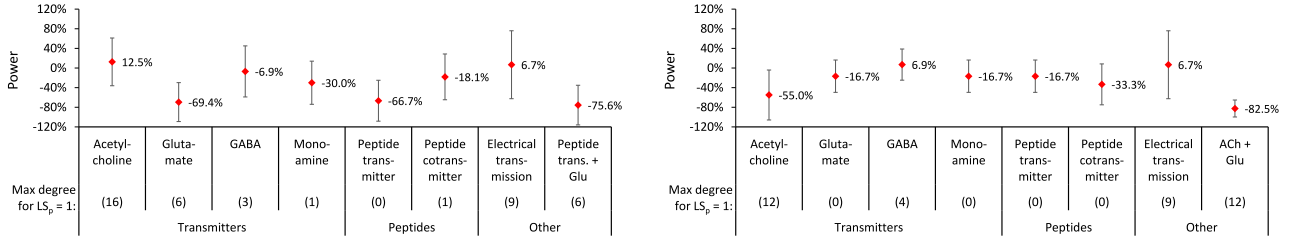


(a) Outdegree of neurons

(b) Existence of specific transmitters among neurons

(1) Absolute frequencies **(2)** Corresponding power

Fig. C.4. Power values for the first layer motor neurons.



(a) Indegree

(b) Outdegree

Fig. C.5. Power values for the second layer motor neurons.

and VD13. In addition, many motor neurons of the first layer are also present here. Note that this is not visible in Fig. 2 since it is only for illustrative purposes. For example, motor neurons of the class VB are all directly connected to muscles, but they have also postsynaptic partners in their class. The postsynaptic partners occur therefore also in the second layer of motor neurons in the shortest paths. Fig. C.5 illustrates the resulting power values for (a) indegree and (b) outdegree distributions. For the indegree, most of the second layer motor neurons can be excellently separated because they do not have incoming connections with peptide transmitters. In this case, the absolute power value of 66.7% is 20.8% lower than that for the first layer motor neurons (see Fig. 3b). On the other hand, Glu is the best factor for logistic regression with −69.4% power. If both factors are combined, the power rises to absolute 75.6%. For the outdegree, no clear conclusion can be drawn at first glance for the individual factors due to high uncertainties. If the outdegree of ACh and Glu is combined, a power of −82.5% can be achieved indicating an outstanding separation ability.

C.4. Complementary information for predicting the locomotory subnetwork

Prediction accuracy with all sensory neurons – The prediction accuracy of the logistic regression models on the test dataset with all sensory neurons is detailed in Table C.6. In total, Model 1 (2) predicts 99.2% (90.1%) of the shortest path correctly, and none of the paths through the locomotory subnetwork $y_i = 1$ are misclassified. For paths that do not completely pass through the locomotory subnetwork $y_i = 0$, there is a misclassification of 1.5% (18.5%). The latter value is significantly higher because Model 2 does not take into account the outdegree of the interneurons. However, this is not a problem but an interesting finding because it predicts additional neurons for the locomotory subnetwork.

New findings based on model predictions with all sensory neurons – Using the predicted shortest paths $y_i = 1$, Model 1 (2) indicates 8 (29) further neurons for the locomotory subnetwork, which are listed in Table C.7. For all of these neurons except AVL, RMFL, and V03, it can be assumed on the basis of literature research and connectivity analysis that they affect the locomotion behavior of *C. elegans*, which is discussed in the following.

The interneurons are considered first. The neuron class AVE together with the classes AVA and AVD initiate the backward locomotion of *C. elegans* ([19, Section I], [84]). The neuron class AIB plays a role in turns. The classes AIB and AIZ promote turns whereas the classes AIY and AIA inhibit turns [25]. The RIG neurons are involved in reversal behavior, and they are likely responsible for increased reversal rates [11]. The AVJ and AVF neurons are suggested to function generally as modulators of the command interneurons that promote forward movement. Therefore, they are involved in the coordination of locomotion. Especially in periods of active egg-laying, they coordinate egg-laying and locomotion [33,69].

Next, the motor neurons are considered. The AVL neuron is involved in the defecation. It controls together with the interneuron DVB the expulsion muscle contraction in the defecation motor program [3, Section III]. Since this neuron has no apparent influence on the locomotion behavior of the worm, it is not considered to be part of the locomotory subnetwork. The function of the PDA neuron is currently unknown, but it could be a counterpart to the

Table C.6
Model performance on the test dataset (all sensory neurons).

Locomotory subnetwork	Prediction Model 1			Prediction Model 2		
	correct	incorrect	correct in %	correct	incorrect	correct in %
yes	72,912	0	100.0	72,912	0	100.0
no	82,730	1,246	98.5	68,407	15,569	81.5
Total	**155,642**	**1,246**	**99.2**	**141,319**	**15,569**	**90.1**

Table C.7
Additional predicted neurons for the locomotory subnetwork of *C. elegans*. The assignment of the neuron type is arbitrary. The neurons are indicated as motor neurons if there are direct connections to muscle cells. Many of the listed neurons are multifunctional. The involvement of the neurons AVL, RMFL, and V03 in locomotion behavior is either unknown or cannot clearly be indicated.

Neuron Type	Model 2 additionally predicts		Same for	Suggested
	neuron class	which neurons	Model 1	function
Interneuron	AVE (2)			- Drive backward movement
	AIB (2)			- Promoting turns
	RIG (1/2)	RIGL		- Involved in reversal behavior
	AVJ (1/2)	AVJL		- Coordination of movement
Motor neuron	AVF (2)			- Coordination of movement
	AVL (1)		yes	- Involved in defecation process
	PDA (1)		yes	- Promoting turns
	RID (1)		yes	- Maintain forward movement
	PVN (1/2)	PVNR	yes	- Coordination of movement
	RIM (1/2)	RIML	yes	- Maintain backward movement
	RIV (2)			- Promoting turns
	RMD (6)			- Promoting turns
	RME (2/4)	RMEL, RMEV		- Promoting turns
	RMF (1/2)	RMFL	yes	
	SMD (4)		only SMDVR	- Promoting turns
	VC (1/6)	VC03	yes	- Inhibition of egg-laying

PDB neuron. The latter generally ensures that ventral turns are preferred by the worm [83]. The PDA neuron is connected to the dorsal muscle MDL21, which lies opposite to the muscles connected with the PDB neuron. The PDB is not only connected with one but with two ventral muscles. This could be the reason why ventral turns are preferred. The RID neuron modulates the motor state of *C. elegans* to maintain forward locomotion [46]. The function of the PVN neurons in the locomotion of the worm is currently not investigated. These are presynaptic to command interneurons of the classes AVA, AVB, AVD, AVE, PVC, and to coordinating neurons of the classes AVJ and AVF, as well as to some motor neurons of the classes VD and DD. Therefore, they could contribute to the coordination of the locomotion. For the neurons of the DD and VD class, it is well established that they coordinate forward and backward movements [19, Section I]. The RIM neurons play a role in the backward locomotion. For example, photostimulation of RIM causes the worm to reverse [29]. As a reaction to an anterior touch, head movements are suppressed and the backward movement is maintained [57]. On the other hand, RIM is suggested to inhibit the initiation of reversals during locomotion [56]. The RIV neurons are involved in the local foraging behavior consisting of reversals and deep omega-shaped turns. The ablation of the neurons reduces the frequency of omega turns [28]. The neurons of the classes RME, SMD, and RMD are connected with head and neck muscles. The SMD and RMD motor neurons drive dorsoventral undulations [41] and are needed for multiple navigation behaviors [28], such as the local foraging behavior. The RME neurons are linked with the head bending amplitude during forward locomotion. The undulatory activity of RME is regulated by SMD neurons via extrasynaptic neurotransmission, which ensures optimal efficiency of forward locomotion [73]. The extrasynaptic neuromodulation between RIM and RME neurons allows deep head bending during omega turns and plays a role in the escape behavior of *C. elegans* [43]. Regarding the RMF neurons, no conclusions can be drawn without a detailed analysis. Therefore, RMFL is not counted to the locomotory subnetwork. The role of the VC motor neurons in locomotion is also unclear. VC04 and VC05 innervate the vulva muscles and are involved in egg-laying [26]. The first three VC neurons innervate motor neurons of the VD and DD class and could therefore contribute to the coordination of locomotion. This is not contradictory to the fact that the release of acetylcholine from these VC neurons inhibits egg-laying behavior. The neuron VC06 has no connections. It is hard to say whether the neurons VC01-VC03 are part of the locomotory subnetwork or belong to the egg-laying program. Therefore, VC03 is not counted.

In summary, it can be assumed that probably 26 of the neurons in Table C.7 are involved in locomotion behavior of *C. elegans* except AVL, RMFL, and V03. Model 1 only predicts 5 of them because it only allows little margin regarding the interneurons. Model 1 includes the indegree and outdegree of the interneurons and Model 2 only the indegree.

Appendix D. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.neucom.2020.11.015.

References

[1] D.G. Albertson, J.N. Thomson, The pharynx of Caenorhabditis elegans, Phil. Trans. R. Soc. London, Ser. B, Biol. Sci. 275 (1976) 299–325, https://doi.org/10.1098/rstb.1976.0085.
[2] Z.F. Altun, D.H. Hall, Muscle system, somatic muscle, WormAtlas (2009), https://doi.org/10.3908/wormatlas.1.7.
[3] Avery, L., Thomas, J.H., 1997. Feeding and Defecation. In: D.L. Riddle, T. Blumenthal, B.J. Meyer, and J.R. Priess, editors. C. elegans II. 2nd edition. Cold Spring Harbor. Cold Spring Harbor Laboratory Press. URL:https://www.ncbi.nlm.nih.gov/books/NBK20138..
[4] I.A. Bany, M.Q. Dong, M.R. Koelle, Genetic and cellular basis for acetylcholine inhibition of Caenorhabditis elegans egg-laying behavior, J. Neurosci. 23 (2003) 8060–8069, https://doi.org/10.1523/JNEUROSCI.23-22-08060.2003.
[5] B. Bentley, R. Branicky, C.L. Barnes, Y.L. Chew, E. Yemini, E.T. Bullmore, P.E. Vértes, W.R. Schafer, The multilayer connectome of Caenorhabditis elegans, PLoS Comput. Biol. 12 (2016), https://doi.org/10.1371/journal.pcbi.1005283 e1005283.
[6] V. Bewick, L. Cheek, J. Ball, Statistics review 14: Logistic regression, Crit. Care 9 (2005) 112–118, https://doi.org/10.1186/cc3045.

[7] J.H. Boyle, S. Berri, N. Cohen, Gait modulation in C. elegans: An integrated neuromechanical model, Front. Comput. Neurosci. 6 (2012), https://doi.org/10.3389/fncom.2012.00010.
[8] J.H. Boyle, N. Cohen, Caenorhabditis elegans body wall muscles are simple actuators, Biosystems 94 (2008) 170–181, https://doi.org/10.1016/j.biosystems.2008.05.025.
[9] P.J. Brockie, J.E. Mellem, T. Hills, D.M. Madsen, A.V. Maricq, The C. elegans glutamate receptor subunit nmr-1 is required for slow nmda-activated currents that regulate reversal frequency during locomotion, Neuron 31 (2001) 617–630, https://doi.org/10.1016/S0896-6273(01)00394-4.
[10] J.L. Cazemier, F. Clascá, P.H.E. Tiesinga, Connectomic analysis of brain networks: Novel techniques and future directions, Front. Neuroanatomy 10 (2016), https://doi.org/10.3389/fnana.2016.00110.
[11] M.Y. Chao, J. Larkins-Ford, T.M. Tucey, A.C. Hart, lin-12 notch functions in the adult nervous system of C. elegans, BMC Neuroscience 6 (2005) 45, https://doi.org/10.1186/1471-2202-6-45.
[12] Chase, D.L., Koelle, M.R., 2007. Biogenic amine neurotransmitters in C. elegans. WormBook: the online review of C. elegans biology, 1–15. doi: 10.1895/wormbook.1.132.1..
[13] Chen, B.L., 2007. Neuronal Network of C. elegans: from Anatomy to Behavior. Ph.d. thesis. The Watson School of Biological Sciences. URL:https://www.wormatlas.org/images/BethChenThesis.pdf..
[14] Chew, Y.L., Grundy, L.J., Brown, A.E.X., Beets, I., Schafer, W.R., 2018. Neuropeptides encoded by nlp-49 modulate locomotion, arousal and egg-laying behaviours in Caenorhabditis elegans via the receptor seb-3. Philosophical Transactions of the Royal Society of London. Series B, Biological sciences 373. doi: 10.1098/rstb.2017.0368..
[15] H. Chiu, A. Alqadah, C.F. Chuang, C. Chang, C. elegans as a genetic model to identify novel cellular and molecular mechanisms underlying nervous system regeneration, Cell Adhesion Migration 5 (2011) 387–394, https://doi.org/10.4161/cam.5.5.17985.
[16] S.J. Cook, T.A. Jarrell, C.A. Brittin, Y. Wang, A.E. Bloniarz, M.A. Yakovlev, K.C.Q. Nguyen, L.T.H. Tang, E.A. Bayer, J.S. Duerr, H.E. Bülow, O. Hobert, D.H. Hall, S.W. Emmons, Whole-animal connectomes of both Caenorhabditis elegans sexes, Nature 571 (2019) 63–71, https://doi.org/10.1038/s41586-019-1352-7.
[17] K. Deisseroth, Optogenetics: 10 years of microbial opsins in neuroscience, Nat. Neurosci. 18 (2015) 1213–1225, https://doi.org/10.1038/nn.4091.
[18] S.J. Dixon, P.J. Roy, Muscle arm development in Caenorhabditis elegans, Development 132 (2005) 3079–3092, https://doi.org/10.1242/dev.01883.
[19] Driscoll, M., Kaplan, J., 1997. Mechanotransduction. In: D.L. Riddle, T. Blumenthal, B.J. Meyer, and J.R. Priess, editors. C. elegans II. 2nd edition. Cold Spring Harbor. Cold Spring Harbor Laboratory Press. URL:https://www.ncbi.nlm.nih.gov/books/NBK20177..
[20] S. El Mestikawy, A. Wallén-Mackenzie, G.M. Fortin, L. Descarries, L.E. Trudeau, From glutamate co-release to vesicular synergy: vesicular glutamate transporters, Nat. Rev. Neurosci. 12 (2011) 204–216, https://doi.org/10.1038/nrn2969.
[21] Engelmann, B., Rauhmeier, R., 2011. The Basel II Risk Parameters. Springer Berlin Heidelberg, Berlin, Heidelberg. doi: 10.1007/978-3-642-16114-8..
[22] Fouad, A.D., 2018. Analysis Of Rhythm Generation In The Caenorhabditis Elegans Motor Circuit. Ph.D. thesis. URL:https://repository.upenn.edu/edissertations/2903..
[23] Fouad, A.D., Teng, S., Mark, J.R., Liu, A., Alvarez-Illera, P., Ji, H., Du, A., Bhirgoo, P. D., Cornblath, E., Guan, S.A., Fang-Yen, C., 2018. Distributed rhythm generators underlie Caenorhabditis elegans forward locomotion. eLife 7. doi: 10.7554/eLife.29913..
[24] Gao, S., Guan, S.A., Fouad, A.D., Meng, J., Kawano, T., Huang, Y.C., Li, Y., Alcaire, S., Hung, W., Lu, Y., Qi, Y.B., Jin, Y., Alkema, M., Fang-Yen, C., Zhen, M., 2018. Excitatory motor neurons are local oscillators for backward locomotion. eLife 7. doi: 10.7554/eLife.29915..
[25] P.A. Garrity, M.B. Goodman, A.D. Samuel, P. Sengupta, Running hot and cold: behavioral strategies, neural circuits, and the molecular machinery for thermotaxis in C. elegans and drosophila, Genes Dev. 24 (2010) 2365–2382, https://doi.org/10.1101/gad.1953710.
[26] J. Gjorgjieva, D. Biron, G. Haspel, Neurobiology of Caenorhabditis elegans locomotion: Where do we stand?, Bioscience 64 (2014) 476–486, https://doi.org/10.1093/biosci/biu058.
[27] M.B. Goodman, D.H. Hall, L. Avery, S.R. Lockery, Active currents regulate sensitivity and dynamic range in C. elegans neurons, Neuron 20 (1998) 763–772, https://doi.org/10.1016/S0896-6273(00)81014-4.
[28] J.M. Gray, J.J. Hill, C.I. Bargmann, A circuit for navigation in Caenorhabditis elegans, Proc. National Acad. Sci. USA 102 (2005) 3184–3191, https://doi.org/10.1073/pnas.0409009101.
[29] Z.V. Guo, A.C. Hart, S. Ramanathan, Optical interrogation of neural circuits in Caenorhabditis elegans, Nat. Methods 6 (2009) 891–896, https://doi.org/10.1038/nmeth.1397.
[30] H. Haken, Principles of Brain Functioning: A Synergetic Approach to Brain Activity, Behavior and Cognition, in: volume 67 of Springer Series in Synergetics, Springer, Berlin and Heidelberg, 1996, https://doi.org/10.1007/978-3-642-79570-1.
[31] D.H. Hall, Gap junctions in C. elegans: Their roles in behavior and development, Devel. Neurobiol. 77 (2017) 587–596, https://doi.org/10.1002/dneu.22408.
[32] S. Hallam, E. Singer, D. Waring, Y. Jin, The C. elegans neurod homolog cnd-1 functions in multiple aspects of motor neuron fate specification, Development 127 (2000) 4239–4252.
[33] L.A. Hardaker, E. Singer, R. Kerr, G. Zhou, W.R. Schafer, Serotonin modulates locomotory behavior and coordinates egg-laying and movement in Caenorhabditis elegans, J. Neurobiol. 49 (2001) 303–313, https://doi.org/10.1002/neu.10014.
[34] T. Hills, P.J. Brockie, A.V. Maricq, Dopamine and glutamate control area-restricted search behavior in Caenorhabditis elegans, J. Neurosci. 24 (2004) 1217–1225, https://doi.org/10.1523/JNEUROSCI.1569-03.2004.
[35] J. Hizanidis, N.E. Kouvaris, G. Zamora-López, Z.L. Gorka, A. Díaz-Guilera, C.G. Antonopoulos, Chimera-like states in modular neural networks, Sci. Rep. 6 (2016) 19845, https://doi.org/10.1038/srep19845.
[36] Hobert, O., 2013. The neuronal genome of Caenorhabditis elegans. WormBook: the online review of C. elegans biology, 1–106doi: 10.1895/wormbook.1.161.1..
[37] A.L. Hodgkin, A.F. Huxley, A quantitative description of membrane current and its application to conduction and excitation in nerve, J. Physiol. 117 (1952) 500–544, https://doi.org/10.1113/jphysiol.1952.sp004764.
[38] S.G. Hormuzdi, M.A. Filippov, G. Mitropoulou, H. Monyer, R. Bruzzone, Electrical synapses: a dynamic signaling system that shapes the activity of neuronal networks, Biochimica et Biophysica Acta – Biomembranes 1662 (2004) 113–137, https://doi.org/10.1016/j.bbamem.2003.10.023.
[39] D.W. Hosmer, S. Lemeshow, Applied Logistic Regression, John Wiley & Sons Inc, Hoboken, NJ, USA, 2000, https://doi.org/10.1002/0471722146.
[40] S.J. Husson, A. Gottschalk, A.M. Leifer, Optogenetic manipulation of neural activity in C. elegans: from synapse to circuits and behaviour, Biol. Cell 105 (2013) 235–250, https://doi.org/10.1111/boc.201200069.
[41] E.J. Izquierdo, R.D. Beer, From head to tail: a neuromechanical model of forward locomotion in Caenorhabditis elegans, Phil. Trans. R. Soc. London, Series B, Biol. Sci. 373 (2018), https://doi.org/10.1098/rstb.2017.0374.
[42] Jorgensen, E.M., 2005. Gaba. WormBook: the online review of C. elegans biology, 1–13. doi: 10.1895/wormbook.1.14.1..
[43] Y. Kagawa-Nagamura, K. Gengyo-Ando, M. Ohkura, J. Nakai, Role of tyramine in calcium dynamics of gabaergic neurons and escape behavior in Caenorhabditis elegans, Zoological Letters 4 (2018), https://doi.org/10.1186/s40851-018-0103-1.
[44] C.H. Lai, C.Y. Chou, L.Y. Ch'ang, C.S. Liu, W. Lin, Identification of novel human genes evolutionarily conserved in Caenorhabditis elegans by comparative proteomics, Genome Res. 10 (2000) 703–713, https://doi.org/10.1101/gr.10.5.703.
[45] Li, C., Kim, K., 2008. Neuropeptides. WormBook: the online review of C. elegans biology, 1–36. doi: 10.1895/wormbook.1.142.1..
[46] Lim, M.A., Chitturi, J., Laskova, V., Meng, J., Findeis, D., Wiekenberg, A., Mulcahy, B., Luo, L., Li, Y., Lu, Y., Hung, W., Qu, Y., Ho, C.Y., Holmyard, D., Ji, N., McWhirter, R., Samuel, A.D.T., Miller, D.M., Schnabel, R., Calarco, J.A., Zhen, M., 2016. Neuroendocrine modulation sustains the C. elegans forward motor state. eLife 5. doi: 10.7554/eLife.19887..
[47] S.R. Lockery, M.B. Goodman, The quest for action potentials in C. elegans neurons hits a plateau, Nat. Neurosci. 12 (2009) 377–378, https://doi.org/10.1038/nn0409-377.
[48] S.L. McIntire, E. Jorgensen, H.R. Horvitz, Genes required for gaba function in Caenorhabditis elegans, Nature 364 (1993) 334–337, https://doi.org/10.1038/364334a0.
[49] B. Milligan, N. Curtin, Q. Bone, Contractile properties of obliquely striated muscle from the mantle of squid (alloteuthis subulata) and cuttlefish (sepia officinalis), J. Exp. Biol. 200 (1997) 2425–2436.
[50] Moerman, D.G., Fire, A., 1997. Muscle: Structure, Function, and Development. In: D.L. Riddle, T. Blumenthal, B.J. Meyer, and J.R. Priess, editors. C. elegans II. 2nd edition. Cold Spring Harbor. Cold Spring Harbor Laboratory Press. URL: https://www.ncbi.nlm.nih.gov/books/NBK20130..
[51] C. Morris, H. Lecar, Voltage oscillations in the barnacle giant muscle fiber, Biophys. J . 35 (1981) 193–213, https://doi.org/10.1016/S0006-3495(81)84782-0.
[52] M.D. Nelson, N.F. Trojanowski, J.B. George-Raizen, C.J. Smith, C.C. Yu, C. Fang-Yen, D.M. Raizen, The neuropeptide nlp-22 regulates a sleep-like state in Caenorhabditis elegans, Nature Commun. 4 (2013) 2846, https://doi.org/10.1038/ncomms3846.
[53] K. Oshio, Y. Iwasaki, S. Morita, Y. Osana, S. Gomi, E. Akiyama, K. Omata, K. Oka, K. Kawamura, Database of Synaptic Connectivity of C. elegans for Computation. Technical report of ccep, keio future, no.3, Keio University, 2003, URL:http://ims.dse.ibaraki.ac.jp/ccep.
[54] D.M. Owens, E.A. Lumpkin, Diversification and specialization of touch receptors in skin, Cold Spring Harbor Perspectives Med. 4 (2014), https://doi.org/10.1101/cshperspect.a013656.
[55] S. Perry, N.A. Khovanova, I.A. Khovanov, Control of heart rate through guided high-rate breathing, Sci. Rep. 9 (2019) 1545, https://doi.org/10.1038/s41598-018-38058-5.
[56] B.J. Piggott, J. Liu, Z. Feng, S.A. Wescott, X.Z.S. Xu, The neural circuits and synaptic mechanisms underlying motor initiation in C. elegans, Cell 147 (2011) 922–933, https://doi.org/10.1016/j.cell.2011.08.053.
[57] J.K. Pirri, M.J. Alkema, The neuroethology of C. elegans escape, Curr. Opin. Neurobiol. 22 (2012) 187–193, https://doi.org/10.1016/j.conb.2011.12.007.
[58] A. Pournaki, L. Merfort, J. Ruiz, N.E. Kouvaris, P. Hövel, J. Hizanidis, Synchronization patterns in modular neuronal networks: a case study of C. elegans, Front. Appl. Math. Stat. 5 (2019) 52, https://doi.org/10.3389/fams.2019.00052.
[59] H. Pruscha, Statistisches Methodenbuch, Springer, Berlin Heidelberg, Berlin, Heidelberg, 2006, https://doi.org/10.1007/3-540-29305-1.

[60] Purves, D., Augustine, George J., Fitzpatrick, David, Hall, W.C., LaMantia, A.S., Mooney, R., White, L.E. (Eds.), 2018. Neuroscience. Sixth edition ed., Oxford University Press Sinauer Associates is an imprint of Oxford Universitiy Press, New York and Oxford..
[61] K. Pyragas, Continuous control of chaos by self-controlling feedback, Phys. Lett. A 170 (1992) 421–428, https://doi.org/10.1016/0375-9601(92)90745-8.
[62] D.M. Raizen, R.Y. Lee, L. Avery, Interacting genes required for pharyngeal excitation by motor neuron mc in Caenorhabditis elegans, Genetics 141 (1995) 1365–1382.
[63] Rand, J.B., 2007. Acetylcholine. WormBook: the online review of C. elegans biology, 1–21. doi: 10.1895/wormbook.1.131.1..
[64] N. Ringstad, Neuromodulation: The fevered mind of the worm, Curr. Biol. 27 (2017) R315–R317, https://doi.org/10.1016/j.cub.2017.03.005.
[65] J.K. Rose, K.R. Kaun, S.H. Chen, C.H. Rankin, Glr-1, a non-nmda glutamate receptor homolog, is critical for long-term memory in Caenorhabditis elegans, J. Neurosci. 23 (2003) 9595–9599, https://doi.org/10.1523/JNEUROSCI.23-29-09595.2003.
[66] J.K. Rose, C.H. Rankin, Analyses of habituation in Caenorhabditis elegans, Learning & Memory 8 (2001) 63–69, https://doi.org/10.1101/lm.37801.
[67] A.F. Russo, Overview of neuropeptides: Awakening the senses?, Headache 57 (Suppl 2) (2017) 37–46, https://doiorg/10.1111/head.13084.
[68] E.R. Sawin, R. Ranganathan, H. Horvitz, C. elegans locomotory rate is modulated by the environment through a dopaminergic pathway and by experience through a serotonergic pathway, Neuron 26 (2000) 619–631, https://doi.org/10.1016/S0896-6273(00)81199-X.
[69] W.R. Schafer, Deciphering the neural and molecular mechanisms of C. elegans behavior, Curr. Biol. 15 (2005) R723–9, https://doi.org/10.1016/j.cub.2005.08.020.
[70] E. Schöll, S.H.L. Klapp, P. Hövel, Control of Self-Organizing Nonlinear Systems, Springer, Berlin, 2016, https://doi.org/10.1007/978-3-319-28028-8.
[71] E. Schöll, H.G. Schuster, Handbook of Chaos Control, Wiley, Weinheim, 2008, https://doi.org/10.1002/9783527622313.
[72] A.C. Scott, The electrophysics of a nerve fiber, Rev. Mod. Phys. 47 (1975) 487–533, https://doi.org/10.1103/RevModPhys.47.487.
[73] Shen, Y., Wen, Q., Liu, H., Zhong, C., Qin, Y., Harris, G., Kawano, T., Wu, M., Xu, T., Samuel, A.D., Zhang, Y., 2016. An extrasynaptic gabaergic signal modulates a pattern of forward movement in Caenorhabditis elegans. eLife 5. doi: 10.7554/eLife.14197..
[74] K.T. Simonsen, D.G. Moerman, C.C. Naus, Gap junctions in C. elegans, Front. Physiol. 5 (2014) 40, https://doi.org/10.3389/fphys.2014.00040.
[75] P. Stein, M. Pavetic, M. Noack, Multivariate Analyseverfahren. Lecture notes, Universität Duisburg-Essen, 2015, URL:https://www.uni-due.de/imperia/md/content/soziologie/stein_11_2015_multivariate.pdf.
[76] J.H. Thomas, Genetic analysis of defecation in Caenorhabditis elegans, Genetics 124 (1990) 855–872.
[77] Tolstenkov, O., van der Auwera, P., Steuer Costa, W., Bazhanova, O., Gemeinhardt, T.M., Bergs, A.C., Gottschalk, A., 2018. Functionally asymmetric motor neurons contribute to coordinating locomotion of Caenorhabditis elegans. eLife 7. doi: 10.7554/eLife.34997..
[78] L.R. Varshney, B.L. Chen, E. Paniagua, D.H. Hall, D.B. Chklovskii, Structural properties of the Caenorhabditis elegans neuronal network, PLoS Comput. Biol. 7 (2011), https://doi.org/10.1371/journal.pcbi.1001066 e1001066.
[79] Q. Wen, S. Gao, M. Zhen, Caenorhabditis elegans excitatory ventral cord motor neurons derive rhythm for body undulation, Phil. Trans. R. Soc. London, Ser. B, Biol. Sci. 373 (2018), https://doi.org/10.1098/rstb.2017.0370.
[80] J.G. White, E. Southgate, J.N. Thomson, S. Brenner, The structure of the nervous system of the nematode Caenorhabditis elegans, Phil. Trans. R. Soc. London, Ser. B, Biol. Sci. 314 (1986) 1–340, https://doi.org/10.1098/rstb.1986.0056.
[81] A.R. Winnier, J.Y. Meir, J.M. Ross, N. Tavernarakis, M. Driscoll, T. Ishihara, I. Katsura, D.M. Miller, Unc-4/unc-37-dependent repression of motor neuron-specific genes controls synaptic choice in Caenorhabditis elegans, Genes Dev. 13 (1999) 2774–2786, https://doi.org/10.1101/gad.13.21.2774.
[82] WormAtlas [Internet], 2020. URL:http://www.wormatlas.org..
[83] G. Yan, P.E. Vértes, E.K. Towlson, Y.L. Chew, D.S. Walker, W.R. Schafer, A.L. Barabási, Network control principles predict neuron function in the Caenorhabditis elegans connectome, Nature 550 (2017) 519–523, https://doi.org/10.1038/nature24056.
[84] M. Zhen, A.D.T. Samuel, C. elegans locomotion: small circuits, complex functions, Curr. Opin. Neurobiol. 33 (2015) 117–126, https://doi.org/10.1016/j.conb.2015.03.009.
[85] Y. Zheng, P.J. Brockie, J.E. Mellem, D.M. Madsen, A.V. Maricq, Neuronal control of locomotion in C. elegans is modified by a dominant mutation in the glr-1 ionotropic glutamate receptor, Neuron 24 (1999) 347–361, https://doi.org/10.1016/S0896-6273(00)80849-1.

Thomas Maertens, born in 1982, holds a master's degree in theoretical physics from Technical University of Berlin (2020). He previously achieved his diploma in business administration from HTW Berlin (2010) and has worked for 4 years as a data analyst in the field of credit rating. His research interests include computational neuroscience and artificial neural networks. The focus lies on modeling the neural network of the worm *C. elegans*.

Eckehard Schöll, born in 1951, is a Professor of Theoretical Physics at the Technical University of Berlin and a Principal Investigator of the Bernstein Center of Computational Neuroscience Berlin. He studied physics at the University of Tübingen (Diplom 1976), and holds PhD degrees in mathematics from the University of Southampton (UK, 1978) and in physics from RWTH Aachen (Germany, 1981), and an Honorary Doctorate from Saratov State University (Russia, 2017). He is an expert in the field of nonlinear dynamics and head of the group Nonlinear Dynamics and Control. His work pertains to research in the fields of mathematics and physics, particularly semiconductor physics, neurodynamics, complex systems and networks, and bifurcation theory. His latest research is also related to topics in biology and the social sciences, e.g. simulation of the dynamics in socioeconomic or neuronal networks. He is one of the forerunners into the research of chimera states.

Jorge Ruiz, born in 1983, is a Doctoral candidate at the Technical University of Berlin and currently based at the Bernstein Center of Computational Neuroscience Berlin. He obtained his MSc in Physics degree at the Pontifical Catholic University of Chile researching in quantum field theory, and his BSc in Physics degree at the same university studying physical theories in different dimensions. His current research has been devoted mainly to design algorithms for different problems using diverse datasets to test their quality, covering topics such as complex networks, dynamical systems, and lately a new interest in multi-agent systems.

Philipp Hövel, born in 1980, is a mathematician and physicist. He received his Diplom in Physics (2004), Mathematics (2006), Dr. rer. nat. (2009), and Habilitation (2017) from Technical University of Berlin. He is currently working as lecturer in the School of Mathematical Sciences at University College Cork, Ireland. His research mission is to lift the boundaries between data-oriented science, theoretical approaches, and numerical simulations addressing interdisciplinary questions based on an overlap of nonlinear dynamics, network science, and control theory. His areas of mathematical expertise include complex systems, bifurcation theory, delay differential equations, and complex networks. His interdisciplinary research concept has led to better insight and fundamental understanding of synchronization processes and other dynamical phenomena. In addition, the combination with empirical data sets will allow for investigations of real-world relevance in areas such as neuroscience, epidemiology, and beyond.