Network control principles predict neuron function in the *C. elegans* connectome

Supplementary Information

Contents

I. Dataset and Nomenclature 2

A. *C. elegans* dataset 2

B. *C. elegans* neuronal classes 3

II. Network Control Analyses 4

A. Structural controllability 5

B. Structural controllability of the *C. elegans* connectome 7

C. Self-loops 12

III. Experimental Validations 15

A. Previous experiments 15

B. New experiments 16

IV. Extended Discussions 20

A. Robustness analyses 20

B. Linearised model of neural dynamics 21

C. Detailed control results for individual neurons within larger classes 21

D. Additional analyses 22

References 24

I. DATASET AND NOMENCLATURE

A. *C. elegans* dataset

Caenorhabditis elegans is a model organism, in part thinks to the mapping of its wiring diagram (connectome), which has facilitated decades of investigation into the structure and function of its nervous system [S1–S9]. The hermaphrodite worm has 302 neurons and 96 muscle cells that are identifiable and consistent across individuals. These neurons are connected with chemical synapses (directed) and electric gaps junctions (bidirectional), and some of the neurons are connected to muscles by neuromuscular junctions (directed). In the paper we used the data of *C. elegans*' neuronal network in [S8–S11], consisting of 279 non-pharyngeal neurons and 95 body wall muscles (in addition to the neurons in the pharyngeal system, CANL/R and VC06, which do not make connections with the rest of the network, and the representative muscle for the vulva and anus are excluded). Specifically, the edges (i.e. chemical synapses and gap junctions) between neurons are publicly available in Section "2.1 Connectivity Data" of [S11]; the data of the edges (i.e. neuromuscular junctions) from neurons to muscles are in Section "2.3 Neuron Connections to Sensory Organs and Body Muscles" of [S11]. Note that we did not include any description of the connectivity between muscles in our network. Electrical couplings between neighboring muscle cells act to synchronize their activity and interfere with fine control of individual muscles. Further mechanical limits imposed by muscle stretch receptors and the anatomy of the animal also clearly constrain the potential activity states attainable by neighboring muscle cells. Hence we did not include electrical couplings between adjacent muscle cells in our analysis. Our hypothesis is that any neural nodes which confer greater controllability to individual muscle outputs within these constraints are likely to be critical for normal locomotor behaviour.

Since edge weights do not affect the analysis of *structural controllability* of a network[S12–S16] (see also Sec. II), we considered the connectome to be unweighted, i.e. regardless of the number of synaptic connections between a pair of neurons, we de-

scribe the connectivity with one edge. Therefore, the network has 514 bidirectional edges and 2746 directed edges (2194 chemical synapses and 552 neuromuscular junctions). Extended Data Figure 1 depicts the unweighted neuronal network of the adult worm where the nodes are classified into four categories [S9]: sensory neurons, interneurons, motor neurons and muscles.

B. *C. elegans* neuronal classes

Most of the 279 neurons are present in symmetrical left-right pairs (e.g. ADAL, ADAR), where both members of the pair have similar connectivity to the rest of the network and are expected to have similar function. Neurons are therefore typically ablated in pairs or larger groups (see below) in laser ablation experiments. In line with this we treat each left-right pair of neurons as a single neuronal class and, unless otherwise specified, remove both members of the class both in controllability analyses and in experimental ablations. Note that some neurons do not have a symmetrical counterpart (such as DVA), while other neuronal classes, defined on similar anatomical grounds, contain a larger number of neurons. For example, there are 6 IL1 neurons, with left-right pairs located ventrally, dorsally and laterally: IL1VL, IL1VR, IL1DL, IL1DR, IL1LL, IL1LR. Motor neurons also exist in large classes of 6 or more cells, such as the nine DA cells DA01 to DA09. We followed the classification from WormAtlas [S1, S17], obtaining 103 distinct neuronal classes. We also referred to neuronal classes without the anatomy-specific suffixes "L" (left), "R" (right), "V" (ventral), "D" (dorsal), "L" (lateral) or the neuron-specific numbering. For example, neurons ADAL and ADAR are referred to as neuronal class ADA.

II. NETWORK CONTROL ANALYSES

As shown in the main text, the analyses of each neuron's role in the locomotion of a worm can be mathematically formalised as a network control problem:

$$\begin{cases} \dot{\mathbf{x}}(t) = A\mathbf{x}(t) + B\mathbf{u}(t), \\ \mathbf{y}(t) = C\mathbf{x}(t), \end{cases} \tag{S1}$$

where the vector $\mathbf{x}(t) \in \mathbb{R}^{(N+M)\times 1}$ captures the states of N neurons and M muscles at time t, the vector $\mathbf{y}(t) \in \mathbb{R}^{M\times 1}$ describes the states of M muscles at time t, the vector $\mathbf{u}(t) \in \mathbb{R}^{S\times 1}$ describes the S external control signals at time t, the matrix $A \in \mathbb{R}^{(N+M)\times(N+M)}$ represents the connectome of *C. elegans*, the matrix $B \in \mathbb{R}^{(N+M)\times S}$ represents the receptor neurons on which the S external signals are imposed, and the matrix $C \in \mathbb{R}^{M\times(N+M)}$ represents the target nodes that we aim to control. Note that, while self-loops A_{ii} can affect the full controllability [S18, S19], in SI Sec. II C we took into account the difference of nodal dynamics between neurons and muscles, and proved that under this assumption the nodal dynamics do not affect our method and criterion described below, allowing us to omit self-loops hereafter.

For example, the system of Extended Data Figure 2(a) can be described by Eq. S1 where $N = 2$, $M = 2$, $S = 1$, $A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ * & 0 & 0 & 0 \\ * & 0 & 0 & 0 \\ 0 & 0 & * & 0 \end{bmatrix}$, $B = \begin{bmatrix} * \\ 0 \\ 0 \\ 0 \end{bmatrix}$, and $C = \begin{bmatrix} 0 & * & 0 & 0 \\ 0 & 0 & 0 & * \end{bmatrix}$. A $(*)$ in matrix A indicates the existence of a link, the $(*)$ in B describes the node that is connected by the external signals, and a $(*)$ in C represents a node that links to the outputs.

The statement that system (S1) is controllable means that, with a suitable choice of signals $\mathbf{u}(t)$, the M target nodes can be moved to any desired final states, i.e. $\mathbf{y}(t)$ can reach an arbitrary point in the M-dimensional state space, within a finite time. According

to Kalman's criterion [S20], system (S1) is *controllable* if and only if

$$\text{rank}[CB, CAB, CA^2B, \ldots, CA^{N+M-1}B] = M. \tag{S2}$$

Yet, directly applying Kalman's criterion to assessing the controllability of real networks is often practically infeasible, for two main reasons:
(i) To calculate the rank of the controllability matrix for (S2) we need to know the value of each entry in matrices A, B and C, i.e. not only the wiring diagram but also the weights and signs of each edge which are often not precisely known in real network data;
(ii) The controllability matrix $[CB, CAB, CA^2B, \ldots, CA^{N+M-1}B]$ is extremely ill-conditioned, hence the numerical calculation of its rank is computationally prohibitive and inaccurate for networks with more than about 50 nodes [S21], potentially leading to incorrect conclusions.

To avoid the two shortcomings of Kalman's criterion, in this paper we considered *structural controllability* (also called *generic controllability*) and applied it to analysing muscular control in *C. elegans*.

A. Structural controllability

Structural controllability deals with the controllability of a system (S1) by considering A, B, and C as *structural* matrices, i.e. the entries of these matrices are fixed zeros or independent free parameters. Hence structural controllability depends only on the presence/absence of edges (i.e. the wiring diagram of a network), not requiring the accurate weights or signs of the edges [S12, S13, S22–S25]. The system (S1) is deemed structurally controllable if there exist free parameters in the matrices A, B and C such that the system is controllable in the sense of Kalman's criterion (S2). Note that if the system (S1) is structurally controllable, it is controllable for almost all possible values of the nonzero entries in A, B and C except some pathological cases for which the algebraic variety in parameter space has Lebesgue measure zero [S13, S26]. Therefore, structural controllability

is a generic property of a system, hence it also being called generic controllability.

Take the system in Extended Data Figure 2(a) as an example. The system matrices are as follows:

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ a_{21} & 0 & 0 & 0 \\ a_{31} & 0 & 0 & 0 \\ 0 & 0 & a_{43} & 0 \end{bmatrix}, \; B = \begin{bmatrix} b_1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \text{ and } C = \begin{bmatrix} 0 & c_{12} & 0 & 0 \\ 0 & 0 & 0 & c_{24} \end{bmatrix}.$$

Hence, the controllability matrix

$$K = [CB, CAB, CA^2B, CA^3B] = \begin{bmatrix} 0 & c_{12}a_{21}b_1 & 0 & 0 \\ 0 & 0 & c_{24}a_{43}a_{31}b_1 & 0 \end{bmatrix}.$$

Since the parameters a_{21}, a_{43}, a_{31}, b_1, c_{12}, c_{24} are nonzero, we have $\text{rank}(K) = 2 = M$, i.e. its controllability does not depend on the specific values of these parameters. Thus, the system is structurally controllable. Note that, if the target nodes we aim to control are nodes 2 and 3, instead of 2 and 4, the controllability matrix becomes

$$K = \begin{bmatrix} 0 & c_{12}a_{21}b_1 & 0 & 0 \\ 0 & c_{23}a_{31}b_1 & 0 & 0 \end{bmatrix},$$

hence $\text{rank}(K) = 1 < M$ for any values of the nonzero parameters, i.e. the system becomes uncontrollable. Therefore, the single input on node 1 is able to control nodes 2 and 4, but not able to control nodes 2 and 3.

For the system in Extended Data Figure 2(b) [S13], the matrices A, B and C are as follows:

$$A = \begin{bmatrix} 0 & 0 & 0 \\ a_{21} & 0 & a_{23} \\ a_{31} & a_{32} & 0 \end{bmatrix}, \; B = \begin{bmatrix} b_1 \\ 0 \\ 0 \end{bmatrix}, \text{ and } C = \begin{bmatrix} 0 & c_{12} & 0 \\ 0 & 0 & c_{23} \end{bmatrix}.$$

We have the controllability matrix:

$$K = [CB, CAB, CA^2B] = \begin{bmatrix} 0 & c_{12}a_{21}b_1 & c_{12}a_{23}a_{31}b_1 \\ 0 & c_{23}a_{31}b_1 & c_{23}a_{32}a_{21}b_1 \end{bmatrix}.$$

For almost all parameters we have $\text{rank}(K) = 2 = M$ thus the system is controllable, except for the pathological case $\frac{a_{21}}{a_{31}} = \frac{a_{23}a_{31}}{a_{32}a_{21}}$. Yet, a slight change of any link's weight will rule the pathological case out and make the system controllable. Therefore, the system is *structurally* controllable [S13].

B. Structural controllability of the *C. elegans* connectome

We start by introducing the *linking graph* [S27], an expanded directed network associated with system (S1), which is useful in our analyses of structural controllability of the *C. elegans* connectome. The basic idea of the linking graph is to incorporate time steps in the graphical representation of system (S1) so that it is able to capture the system's dynamics.

The construction of a linking graph is illustrated in Extended Data Figure 2(c). Formally, in a *linking graph* $G_{\text{linking}} = G_{\text{linking}}(A, B, C)$ defined on three structural matrices $\{A, B, C\}$, the node set of G_{linking} is $V_A \cup V_B \cup V_C$, where

$$V_A = \{v_{i,t}^A | i = 1, 2, \ldots, N + M; t = 1, 2, \ldots, N + M\},$$
$$V_B = \{v_{i,t}^B | i = 1, 2, \ldots, S; t = 0, 1, \ldots, N + M - 1\},$$
$$V_C = \{v_i^C | i = 1, 2, \ldots, M\};$$

and the edge set is $E_A \cup E_B \cup E_C$, where

$$E_A = \{v_{j,t}^A \to v_{i,t+1}^A | a_{ij} \neq 0, t = 1, 2, \ldots, N + M - 1\},$$
$$E_B = \{v_{j,t}^B \to v_{i,t+1}^A | b_{ij} \neq 0, t = 0, 1, \ldots, N + M - 1\},$$
$$E_C = \{v_{j,t}^A \to v_i^C | c_{ij} \neq 0, t = N + M\}.$$

The *linking size* of G_{linking} is defined as the maximum number of *disjoint* paths from the node set V_B to the set V_C (paths are disjoint if they do not have any common nodes).

As shown in Extended Data Figure 2(d), the calculation of linking size can be mapped to a multi-source-multi-sink maximum flow problem with the constraint that the capacities of each node in V_A and each edge in E_A are 1. This maximum flow problem can be solved using several existing algorithms [S28], hence we can efficiently calculate the linking size of a system.

If the output matrix $C = \mathbf{I}_{(N+M)}$ (the identity matrix of size $(N + M)$), i.e. the case of aiming to control all the $N + M$ nodes, the number of controllable nodes is equal to the linking size of the linking graph associated with the system [S27]. However, when we aim to control only a fraction of nodes, the number of controllable nodes is *not* equal to the linking size. In fact, according to [S22], we have

$$z \leq z^*, \tag{S3}$$

where z is the actual number of structurally controllable nodes in system (S1), and z^* is the linking size of the associated G_{linking}. This means that the linking size z^* provides only an *upper bound* for the number of structurally controllable nodes in the given target node set. Therefore, the linking size cannot be directly applied for determining the actual number of controllable muscles in a healthy worm or for assessing whether the removal of a neuronal class will decrease the number of controllable muscles.

Given a wiring diagram and the input nodes, determining how many nodes in a target node set are structurally controllable is a challenging problem and has not been fully solved since the 1980s [S22, S29], when the problem was first posed. In the following we offer a *lower bound* of z, which we successfully applied to the controllability analyses of the *C. elegans* connectome.

When the nodes in the target set do not receive the external control signals $\mathbf{u}(t)$ directly, the network can be separated into two subnetworks, G_{O} and G_{M} (see the schematic picture Extended Data Figure 2(e)), where G_{O} contains a subnetwork G_{S}, each node in which receives the external signals directly, and a subnetwork G_{D}, each node in which directly connects to the nodes in the target set. The set of nodes in G_{M}, G_{S} and G_{D} are denoted

by V_{M}, V_{S} and V_{D} respectively. Firstly, taking into account the high density of real neural networks, search for a subnetwork in $(G_{\mathrm{O}} - G_{\mathrm{D}})$ that can be fully structurally controlled by the nodes V_{S}; Secondly, add G_{D}'s nodes (and their associated links) one-by-one into the subnetwork found in the first step, until the new subnetwork cannot be fully controlled by V_{S}. Denote the V_{D}'s subset that are in the final subnetwork G_{F} found in this step as $V_{\mathrm{D}\text{-}}$; Thirdly, search for a maximum matching [S30] from $V_{\mathrm{D}\text{-}}$ to V_{M}, obtaining a node set $U \subseteq V_{\mathrm{M}}$ that are covered by this matching. Note that there may be many configurations of U but its cardinality z_* is unique; Finally, we have the following proposition:

Proposition: The actual number of controllable nodes in the target set V_{M} is larger than or equal to z_*, i.e.

$$z \geq z_*. \tag{S4}$$

Proof: Given a network, the set of input nodes and the set of target nodes, we first show that adding any nodes and links into the network does not decrease the number of structurally controllable nodes in the target set. Recall that each element of the controllability matrix K has a specific graphical meaning. For example, the network of Extended Data Figure 2(b) can be redrawn as Extended Data Figure 2(f) where the virtual nodes in_i and out_i correspond to the ith input and the ith output respectively. The controllability matrix element

$$K_{q,S*l+r} = \sum_{\mathrm{P}\in\mathrm{Path}_{q,r}(l+2)} \pi_{\mathrm{P}}$$

where $\mathrm{Path}_{q,r}(l+2)$ is the set of paths from rth input to qth output with length $l+2$, π_{P} is the product of the weights of all edges in the path P, $q = 1, \ldots, M$, $l = 0, \ldots, N+M-1$ and $r = 1, \ldots, S$). For the case of Extended Data Figure 2(f), there is only one path from in_1 to out_2 that has length 4, hence the element $K_{2,1*(4-2)+1} = K_{2,3} = \pi_{\mathrm{P}} = c_{23}a_{32}a_{21}b_1$, agreeing with the calculation in Sec. II A. If any node and its associated links is added into the network, the cardinality of $\mathrm{Path}_{q,r}(l+2)$ will increase or keep unchanged. This means that the corresponding element in matrix K will be composed of more than or equal to the original number of π_{P}, unable to enhance the linear dependence between the columns of

matrix K (since all elements in matrices A, B and C are free parameters). As rank(K) is defined as the maximum number of linearly independent column vectors of matrix K, we have the conclusion that adding any node and its associated links into a network does not decrease the number of structurally controllable nodes in the target set. Since $G_{\mathrm{F}} \subseteq G_{\mathrm{O}}$, the nodes in $V_{\mathrm{D}\text{-}}$ are certainly structurally controllable in the original network, and the actual number of structurally controllable nodes in V_{D} cannot be smaller than $|V_{\mathrm{D}\text{-}}|$. As $V_{\mathrm{D}\text{-}}$ are structurally controllable, by the definition of structural controllability the signals that the nodes $V_{\mathrm{D}\text{-}}$ receive from the inputs are linearly independent of each other. Thus, the maximum matching [S30] from $V_{\mathrm{D}\text{-}}$ to V_{M} ensures that each node in U also receives linearly independent signals, i.e. all nodes in U are structurally controllable, leading to the inequality S4. □

We applied the two bounds to the *C. elegans* connectome, finding that $z^* = z_* = 89$ for muscular control regarding anterior or posterior gentle touch, i.e. $V_{\mathrm{S}} = \{\mathrm{ALML}, \mathrm{ALMR}, \mathrm{AVM}\}$ or $V_{\mathrm{S}} = \{\mathrm{PLML}, \mathrm{PLMR}\}$ respectively. Hence, the actual number of controllable muscles in the healthy worm is $z = 89$. It is noteworthy that we found that the neuron set V_{D} in the healthy *C. elegans* connectome, i.e. all muscles' neighbours, are independently controllable, meaning that the signals they receive from the receptors are independent of each other. We then repeated the analyses to assess whether the number of controllable muscles decreases after the ablation of each neuronal class. The effect of neuronal ablations on the controllability of motor neurons was assessed in a similar spirit, i.e. when the target nodes are the set of motor neurons that have direct connections to muscles, instead of the set of 95 muscles.

In the main text we used a relevant subnetwork of the *C. elegans* connectome and showed the control mechanism for our new predictions of the involvement of PDB and DD04-06 (but not DD01-03) in worm locomotion. In the following we describe in detail the network control mechanisms for other neuronal classes which we predicted to be involved in worm locomotion.

Assessing AVA class: The subnetwork shown in Extended Data Figure 3(a) com-

prises the predicted neuronal class (green), the affected muscles (purple), these muscles nn (nearest neighbours) (yellow), the anterior gentle touch receptor neurons (blue) and any other neurons lying on shortest paths between input and output (gray). In the healthy worm, using the method described above we proved that the seven neighbouring neurons receive independent signals from ALML/R and AVM. Hence, according to the control principle of Fig. 3d in the main text, all the six muscles are independently controllable. When AVAL/R are ablated, while there is no longer a path between the three receptor neurons and the neurons DA07 and AS08, signals can still pass to all six muscles. Yet, the ablation of AVAL/R forces the signal to the muscles to go through only five neurons, unable to control all six muscles.

Assessing AS class: As shown in Extended Data Figure 3(b), in the healthy worm, the 16 neighbouring neurons (yellow), which we proved to be receiving independent signals from the receptors (blue), are able to independently control all the 14 muscles (purple) according to the control mechanism displayed in Fig. 3d of the main text. If all AS neurons (green) are removed, the 11 remaining neighbouring neurons are unable to independently control the 14 muscles in this subnetwork. Hence AS neuron class is significant in the muscular control of worms.

Assessing DA class: As shown in Extended Data Figure 4(a), in the healthy worm, according to the control mechanism displayed in Fig. 3d of the main text, all the 18 muscles (purple) in this subnetwork are independently controllable. If all DA neurons (green) are removed, the 16 remaining neighbouring neurons (yellow) are unable to independently control these 18 muscles. Hence DA neuron class plays a significant role in the muscular control of worms.

Assessing DB class: As shown in Extended Data Figure 4(b), in the healthy worm, according to the control mechanism displayed in Fig. 3c of the main text, 15 of the 18 muscles (purple) in this subnetwork are independently controllable. If all DB neurons (green) are removed, the 12 remaining neighbouring neurons (yellow) are able to independently control 12 of these 18 muscles, i.e. the number of controllable muscles in this

subnetwork decreases from 15 to 12, suggesting that the ablation of DB neurons limits the flexibility of worm locomotion.

Assessing VA class: As shown in Extended Data Figure 5(a), in the healthy worm, according to the control mechanism displayed in Fig. 3d of the main text, all the 14 muscles (purple) in this subnetwork are independently controllable. If all VA neurons (green) are removed, the 10 remaining neighbouring neurons (yellow) are unable to independently control these 14 muscles. Hence VA neuron class plays a significant role in the muscular control of worms.

Assessing VB class: As shown in Extended Data Figure 5(b), in the healthy worm, according to the control mechanism displayed in Fig. 3c of the main text, 14 of the 15 muscles (purple) in this subnetwork are independently controllable. If all VB neurons (green) are removed, the 10 remaining neighbouring neurons (yellow) are able to independently control 10 of these 15 muscles, i.e. the number of controllable muscles in this subnetwork decreases from 14 to 10, suggesting that the ablation of VB neurons limits the flexibility of worm locomotion.

Assessing VD class: As shown in Extended Data Figure 5(c), in the healthy worm, according to the control mechanism displayed in Fig. 3d of the main text, all the 14 muscles (purple) in this subnetwork are independently controllable. If all VD neurons (green) are removed, the 9 remaining neighbouring neurons (yellow) are unable to independently control these 14 muscles, i.e. the number of controllable muscles decreases from 14 to 9. Hence VD neuronal class plays a significant role in the muscular control of worms.

C. Self-loops

The non-zero diagonal elements A_{ii} in the adjacency matrix A represent the intrinsic dynamics of node i, hence each node has a self-loop. Previous studies showed that self-loops can affect the full controllability of a network [S18, S19]. In [S19], Zhao *et al* found that when all nodes have identical non-zero self-loops, the maximum matching result [S13] for full controllability is still valid.

Here we show that, for the target controllability in *C. elegans* connectome, the assumption of identical self-loops for all nodes is not necessary. We can relax this assumption to the case that the intrinsic dynamics of neurons is different from that of muscles. In other words, for the sake of mathematical modeling, we have made the simplifying assumption that the dynamics of each neuron have similar characteristics (e.g. that they are all non-spiking, have the same active currents, and depend on the same calcium dynamics), and are sufficiently distinct from those of the muscles. In the following, we prove that the method and criterion used in this paper is valid under this assumption.

The adjacency matrix A can be recast to $A = \begin{bmatrix} A_{\alpha\alpha} & A_{\alpha\beta} \\ A_{\beta\alpha} & A_{\beta\beta} \end{bmatrix}$, where sub-matrix $A_{\alpha\alpha}$ represents the adjacency matrix of sub-network G_o in Extended Data Figure 2(a), $A_{\beta\beta}$ is the adjacency matrix of sub-network G_M, $A_{\beta\alpha}$ captures the edges from G_o to G_M, and $A_{\alpha\beta}$ encodes the edges from G_o to G_M. Similarly, we can rewrite $B = \begin{bmatrix} B_\alpha \\ B_\beta \end{bmatrix}$ and $C = \begin{bmatrix} C_\alpha & C_\beta \end{bmatrix}$. For the *C. elegans* connectome considered in this paper, G_o is the sub-network consisting of all neurons and G_M is the sub-network of muscles. Since only neurons in G_o sense external stimuli and output nodes are all in G_M, we have $B_\beta = \mathbf{0}$ and $C_\alpha = \mathbf{0}$. In *C. elegans* connectome there is no edges from G_M to G_o, hence $A_{\alpha\beta} = \mathbf{0}$, and $A_{\beta\beta} = \mathbf{0}$. Therefore, $CA^kB = C_\beta A_{\beta\alpha} A_{\alpha\alpha}^{k-1} B_\alpha$, and the controllability matrix for the connectome without self-loops is

$$\begin{aligned} K &= [CB, CAB, CA^2B, \ldots, CA^{N+M-1}B] \\ &= [C_\beta B_\alpha, C_\beta A_{\beta\alpha} B_\alpha, C_\beta A_{\beta\alpha} A_{\alpha\alpha} B_\alpha, \ldots, C_\beta A_{\beta\alpha} A_{\alpha\alpha}^{N-M-2} B_\alpha] \end{aligned} \tag{S5}$$

In our relaxed assumption, the self-loop of neurons is different from that of muscles, i.e. the adjacency matrix with these self-loops becomes

$$\hat{A} = \begin{bmatrix} A_{\alpha\alpha} + l_1 I_\alpha & \mathbf{0} \\ A_{\beta\alpha} & l_2 I_\beta \end{bmatrix}, \tag{S6}$$

where I_α and I_β are identity matrices and $l_1 \neq 0$, $l_2 \neq 0$, $l_1 \neq l_2$. Denote the controllability matrix for the connectome with the self-loops (Eq. S6) by $\hat{K} = [CB, C\hat{A}B, C\hat{A}^2B, \ldots, C\hat{A}^{N+M-1}B]$, we obtain the following proposition:

Proposition: Under the relaxed assumption (Eq. S6), adding self-loops into the connectome does not change the rank of controllability matrix, i.e. $\mathrm{rank}(\hat{K}) = \mathrm{rank}(K)$.

Proof: Inserting the expression of $\hat{A}$ (i.e. Eq. S6) into $\hat{K}$, we obtain the sub-matrices of $\hat{K}$:

$$\begin{aligned} C\hat{A}^k B &= C_\beta \left(\sum_{i=1}^{k} l_2^{i-1} A_{\beta\alpha} (A_{\alpha\alpha} + l_1 I_\alpha)^{k-i} \right) B_\alpha \\ &= \underbrace{C_\beta A_{\beta\alpha} (A_{\alpha\alpha} + l_1 I_\alpha)^{k-1} B_\alpha}_{\mathrm{ta}(k)} + \underbrace{C_\beta \left(\sum_{i=2}^{k} l_2^{i-1} A_{\beta\alpha} (A_{\alpha\alpha} + l_1 I_\alpha)^{k-i} \right) B_\alpha}_{\mathrm{tb}(k)} \end{aligned} \tag{S7}$$

for $k \geq 2$. For the cases $k = 0$ and $k = 1$, i.e. the first and second sub-matrices in $\hat{K}$, the forms are exactly the same as those in K, i.e. $CB = C_\beta B_\alpha$ and $C\hat{A}B = C_\beta A_{\beta\alpha} B_\alpha$. The sub-matrix $C\hat{A}^k B$ is split into two terms $\mathrm{ta}(k)$ and $\mathrm{tb}(k)$ for the convenience of the following derivation. Based on the expression S7, we have the recursions $\mathrm{tb}(k)=l_2 C\hat{A}^{k-1}B$ for $k \geq 2$.

Because elementary column operations (i.e. interchanging two columns; adding a multiple of one column to another; multiplying a column by a nonzero number and adding the result to another column) do not change the rank of a matrix, we perform these three elementary operations on $\hat{K}$.

(i): Firstly, since $\mathrm{tb}(k)=l_2 C\hat{A}^{k-1}B$, we multiply $C\hat{A}^{k-1}B$ by $-l_2$ and add the result to $C\hat{A}^k B$. After these operations, the controllable matrix $\hat{K}$ becomes $[C_\beta B_\alpha, C_\beta A_{\beta\alpha} B_\alpha, C_\beta A_{\beta\alpha}(A_{\alpha\alpha} + l_1 I_\alpha) B_\alpha, \ldots, C_\beta A_{\beta\alpha}(A_{\alpha\alpha} + l_1 I_\alpha)^{N+M-2} B_\alpha]$.

(ii): As $A_{\alpha\alpha}$ and $l_1 I_\alpha$ commute, $(A_{\alpha\alpha} + l_1 I_\alpha)^k$ can be expanded to an expression containing the terms of $A_{\alpha\alpha}^k$, $A_{\alpha\alpha}^{k-1}$, …, $A_{\alpha\alpha}$, and I_α. Hence, secondly, we multiply $C_\beta A_{\beta\alpha} B_\alpha$ by an appropriate non-zero number and add the result to $C_\beta A_{\beta\alpha}(A_{\alpha\alpha} + l_1 I_\alpha) B_\alpha$. After this operation, the sub-matrix $C_\beta A_{\beta\alpha}(A_{\alpha\alpha} + l_1 I_\alpha) B_\alpha$ contains only the term of $C_\beta A_{\beta\alpha} A_{\alpha\alpha} B_\alpha$. We

perform this operation on higher-order sub-matrices, for example, multiplying $C_\beta A_{\beta\alpha} B_\alpha$ by an appropriate number and multiplying $C_\beta A_{\beta\alpha} A_{\alpha\alpha} B_\alpha$ by another appropriate number, and adding the sum of these two results to the sub-matrix $C_\beta A_{\beta\alpha}(A_{\alpha\alpha}+l_1 I_\alpha)^2 B_\alpha$, then this sub-matrix can be reduced to a matrix that contains only the term of $C_\beta A_{\beta\alpha} A_{\alpha\alpha}^2 B_\alpha$. After repeating this operations to all sub-matrices $C_\beta A_{\beta\alpha}(A_{\alpha\alpha} + l_1 I_\alpha)^k B_\alpha$, the controllability matrix $\hat{K}$ contains only the sub-matrices $C_\beta A_{\beta\alpha} A_{\alpha\alpha}^k B_\alpha$.

Therefore, after the elementary matrix operations (i) and (ii) $\hat{K}$ turns into the same form of K. Operations (i) and (ii) do not change the rank of $\hat{K}$, so rank($\hat{K}$)=rank(K). □

III. EXPERIMENTAL VALIDATIONS

A. Previous experiments

We applied the network control framework described in Sec. II to the adult nematode connectome, predicting twelve neuronal classes whose removal alters the controllability of muscles or motor neurons. Eleven of the twelve neuronal classes have been confirmed by previous laser ablation experiments or indirect evidences (Fig. 1a in the main text).

These include seven motor neuron classes: DA, DB, DD, VA, VB, VD, and AS. Indeed, at the simplest level, locomotion in *C. elegans* results from the alternate contraction and relaxation of the dorsal and ventral body wall muscles which are controlled directly by these seven neuron classes [S31–S34]. Specifically, DA, DB, and DD innervate the *dorsal* body wall muscles and have been previously implicated in locomotion due to their morphology and connectivity profiles [S1] and their physiological activity during locomotion [S35]. Ablation experiments confirmed that DA, DB and DD cells are necessary for backward and forward locomotion and coordinated motion in either direction [S31]. While no ablation experiments exist for the post-embryonically born motor neurons VA, VB and VD, they are the ventral counterparts of the dorsal motor neurons DA, DB and DD from a morphological and connectivity point of view [S36]. Several studies have confirmed their

physiological activity during locomotion and bending [S35, S37–S39] and severe locomotory defects were observed in mutants which lack differentiated VA motor neurons [S40] and in animals with genetically targeted disruption of VA and VC [S41]. Our prediction of AS neurons is again supported by the similarity in their morphology and connection profile to VA and DA neurons, and by the experimentally observed correlation between AS activity and dorsal bending [S37]. Given that our predictions also identify the premotor interneuronal class AVA (Extended Data Figure 3(a)), we find that the relevance to locomotion in the worm for eight of the nine neural classes predicted by the control analysis is supported either by direct laser ablation studies, or by indirect evidence based on genetics or neuronal activity.

B. New experiments

Laser ablation: Laser ablation was carried out as described in [S42]. DD neurons were ablated in L1 animals, using AQ2968 (*ljIS134[Punc-47::GFP-SL2-tagRFP-T]*, a gift from Dr. Victoria Butler) as a fluorescent marker for DDs. Since PDB is born later, it was ablated in L2/early L3 animals, using OH904 (*otIS33[kal-1::gfp]*, [S43]) as a marker. For mock ablations, animals of the same genotype as those to be ablated were mounted and treated with sodium azide in the same way as ablated animals.

Tracking: Worms were maintained at room temperature on Nematode Growth Medium (NGM) plates seeded with OP50 bacteria until the point of tracking. The 3 cm low peptone plates kept at 4°C were dried for approximately 24 hours at room temperature prior to use. On the day of tracking, low peptone plates were seeded with 20 μL of OP50 bacteria and allowed to dry. At least 20 young-adult hermaphrodites per strain were tracked for 15 minutes, spontaneously behaving on food. Worms were moved to their tracking plate using an eyelash hair and allowed to acclimatise for 30 minutes before tracking. To avoid potential bias resulting from heterogeneity in room conditions, recording was randomised across multiple trackers and across multiple days. Ablated animals were matched to mock-ablated controls tracked on the same day. The camera magnifi-

cation was set to between 3.5 and 4.5 μm per pixel at a 640×480 video resolution and computer vision software was used to a control a motorised stage allowing the camera to follow the worm (see [S44]).

Analysis: Videos were segmented and analysed using Worm Tracker software developed in the Schafer lab [S44]. Specifically, each movie frame was segmented using the Otsu threshold [S45], and the worm was taken to be the largest connected component in the resulting image. The curvature of the outline of this connected component was determined, and the two points of highest curvature were taken to be the head and the tail. The skeleton was found by tracing the midline of the outline between these two points. The skeleton was divided into 49 equally spaced points, which were used to define 48 tangent angles. The mean of these angles was then subtracted. A total of 702 features (see [S44] for a complete list) describing various locomotor parameters were computed for each recording. Extracted features were then analysed for statistical differences in GraphPad Prism.

Eigen Projections – Our initial analysis focused on eigen projection features [S44]. The eigen projection features are a measure of worm posture. They are the projections onto the first six eigenworms which together account for 97% of the variance in posture. The eigenworms were computed from 15 N2 videos as previously described in [S46]. Briefly, 48 tangent angles are calculated along the skeleton and rotated to have a mean angle of zero. Principal components analysis is performed on the pooled angle data and we keep the 6 principal components (or eigenworms) that capture the most variance. The first eigenworm roughly corresponds to body curvature. The next two eigenworms are akin to sine and cosine waves encoding the travelling wave during crawling. The fourth eigenworm captures most of the remaining variance at the head and tail. Projected amplitudes are calculated from the posture in each frame.

DD neuron ablations: We first performed the experiments for the individual ablation of DD02 and DD05. We observed several significant differences between DD05 and mock-ablated animals in features related to eigen projection 4 (EP4), which relates to fine

Significant features for mock vs DD4-ablated							
Significant features	p-value	Mock DD4	DD4 ablation	Difference	SE of difference	t ratio	df
Negative Forward EP4	$7.82 * 10^{-4}$	-1.198	-0.961	-0.238	0.065	3.652	38
Absolute Forward EP4	$3.10 * 10^{-3}$	1.178	0.974	0.204	0.065	3.16	38
Positive Forward EP4	0.025	1.134	0.976	0.158	0.068	2.333	38

Significant features for mock vs DD5-ablated							
Significant features	p-value	Mock DD5	DD5 ablation	Difference	SE of difference	t ratio	df
Negative EP4	$7.7 * 10^{-3}$	-1.034	-0.945	-0.087	0.032	2.720	104
Absolute EP4	0.020	1.034	0.966	0.069	0.029	2.359	104
Absolute Forward EP4	0.028	1.065	0.984	0.082	0.037	2.222	103
Negative Forward EP4	0.029	-1.064	-0.973	-0.091	0.041	2.208	103

Significant features for mock vs DD2-ablated							
Significant features	p-value	Mock DD	DD2 ablation	Difference	SE of difference	t ratio	df
None							

Significant features for mock vs DD3-ablated							
Significant features	p-value	Mock DD	DD3 ablation	Difference	SE of difference	t ratio	df
None							

TABLE I: Statistically significant features for individual ablations of DD neurons, comparing mock-ablated controls to specific DD ablations. Statistical test: multiple t-test, significance level $\alpha = 0.05$.

movements at the animal's head and tail. These features were not significantly different between DD02 and mock-ablated animals (Table I). We repeated the DD02 and DD05 ablations in a second replicate, and the EP4 features remained statistically significantly different in the DD05-ablated but not in the DD02-ablated animals. To assess additional DD neurons, we ablated a second neuron predicted to affect control (DD04) and a second neuron predicted not to affect control (DD03) and evaluated EP4 features. In this experiment, DD04-ablated animals were significantly different in several EP4 features from mock-ablated animals, while DD03-ablated animals were not significantly different in any EP4 features. Extended Data Figure 6 shows the EP4 time series plots of sample videos of DD-ablated animals with corresponding mock-ablated controls, along with still images from these videos, highlighting reduced tail movement while moving forward (lower absolute/negative EP4 correlated with ablation.

PDB ablations: For PDB ablation, we observed significant differences in several

Significant features for mock vs PDB-ablated							
Significant features	p-value	Mock PDB	PDB-ablated	Difference	SE of difference	t ratio	df
Absolute EP1	0.034	2.089	2.266	-0.177	0.082	2.158	75
Paused EP1	0.043	0.258	-0.270	0.527	0.255	2.064	75
Negative Paused EP1	0.020	-1.708	-2.058	0.351	0.147	2.384	75
Negative EP1	0.015	-1.987	-2.250	0.263	0.106	2.492	75

TABLE II: Statistically significant features for PDB ablations , comparing mock-ablated controls to PDB ablations. 'EP1' is short for 'Eigen Projection 1'. Statistical test: multiple t-test, significance level $\alpha = 0.05$.

Omega turn D/V bias for mock vs PDB-ablated				
Significant features	p-value	Mock PDB (n)	PDB ablation (n)	Difference
Omega turn ventral bias (by worm)	$3.5 * 10^{-3}$	0.86(30)	0.66 (35)	0.20
Omega turn ventral proportion	$4.7 * 10^{-4}$	0.82 (127)	0.64 (196)	0.18

TABLE III: Ventral vs dorsal omega turn bias in mock and PDB-ablated animals. Statistical tests: for omega turn bias by worm, two-tailed t-test (n indicates worms exhibiting turns); for omega ventral proportion, two-tailed z-test for two proportions (n indicates total turns).

eigenworm 1 features when compared to mock-ablated animals (Table II). We carried out a second set of ablations and these features remained statistically significant. Each of these differences correlates with more highly negative EP1 values, suggesting that large bends (i.e. omega turns) might be less biased to the ventral side than typically observed in normal animals [S47, S48]. To investigate this directly, we compared the frequencies of dorsal and ventral omega turns as scored by the tracking system [S44] in mock- and PDB-ablated animals. We found that indeed the ventral bias of omega turns, considered either worm by worm (i.e. the average of the ventral/total turn proportions measured for each worm) or as a total proportion, was significantly less pronounced in ablated animals (Table III)). Extended Data Figure 6 shows the EP1 time series plots of sample videos of a PDB-ablated animal with corresponding mock-ablated controls, along with still images from these videos, showing a dorsal omega turn in a PDB-ablated animal contrasting with a ventral omega turn typically seen in normal and mock-ablated animals.

IV. EXTENDED DISCUSSIONS

A. Robustness analyses

The *C. elegans* neuronal network is the best mapped connectome so far. Despite the fact that the network we used in this paper is from the most widely accepted, reliable dataset available, it is known that this data contains inaccuracies and uncertainties. We thus test the robustness of our predictions to imperfections in the connectome by calculating the probability that each of the predicted neuron classes remains significant in locomotion after alterations to the network. We consider three types of connection alterations: randomly deleting weak edges, randomly adding or randomly rewiring edges between neurons in the *C. elegans* connectome data.

Weak link deletion: The data in [S11] records the weight, i.e. the number of connections, between a pair of nodes. It is reasonable to assume that, the smaller the weight, the less stereotyped is the edge, prompting us to test the robustness of our predictions if the edges with weights not larger than one are randomly deleted. Extended Data Figure 7(a) shows that the predictions are robust when weak links are deleted. *Random link addition*: We also test the robustness in the scenario that a number of added edges randomly connect neurons. The robustness result for muscle control is shown in Extended Data Figure 7(b), and that for the controllability of motor neurons is shown in Extended Data Figure 7(d).

Random link rewiring: We finally consider the alteration of randomly rewiring the existing edges between neurons in *C. elegans* connectome data. The random rewiring was implemented by double-edge swapping:

Step 1. randomly select two edges, u→v and x→y, where the edges u→y and x→v do not exist in the network;

Step 2. remove the edges u→v and x→y, and create two new edges u→y and x→v;

Step 3. repeat steps 1 & 2 until the desired number of edges have been swapped.

Note that the degree of each neuron is preserved in this process. The robustness result

for muscle control is shown in Extended Data Figure 7(c), and that for the controllability of motor neurons is shown in Extended Data Figure 7(e).

B. Linearised model of neural dynamics

While neural dynamics are inherently nonlinear, our analysis of the control role of neurons is built on the linearised model described by Eq. 2 in the main text. Therefore, we outline its strengths and weaknesses in the following.

(i) The linearised model may be an accurate approximation for the nonlinear behaviours of neural networks in some scenarios. For example, the spontaneous activity pattern of a neural network with Wilson-Cowan nonlinear dynamics can be captured by the linearised model [S49], and both linear and nonlinear models have been successfully applied to predict function from structure of neural networks [S50].

(ii) The controllability of the linearised system is known to determine the local controllability of a full nonlinear system. For example, if a system is locally controllable along a specific trajectory in the state space, then the corresponding nonlinear system is also controllable along the same trajectory [S51]. Indeed, the nonlinear controllability of motifs with nonidentical connections exhibits the same properties as its linear counterpart [S52], and the recent work [S53] shows that linear controllability predictions are consistent with simulations of neuronal networks with Wilson-Cowan nonlinear dynamics.

Nevertheless, future work using more realistic nonlinear models will be an important extension to the current framework.

C. Detailed control results for individual neurons within larger classes

In the main text we showed that, within the class of DD neurons, the individual ablation of DD04, DD05 and DD06 neurons led to a decrease in the number of controllable muscles, while removing DD01, DD02 or DD03 does not affect the muscular controllability. This means that the individual neurons within a neuronal class could play different roles in

muscular control, prompting us to systemically explore the roles of individual neurons of each large class in the muscular control of *C. elegans*.

Note that while the number of controllable muscles is unique, there are different configurations of the set of muscles that are controlled. For example, in Extended Data Figure 8(a), 15 of the 18 muscles are controllable, yet which 15 muscles are controlled is not unique, i.e. there exist different sets containing 15 muscles that are controlled. Therefore, we calculated the probability vector describing the probabilities with which each muscle is controlled, by running the analysis described in Sec. II many times for each ablation. We then compared the probability vector for the healthy connectome and that for the neuron-ablated connectome. The difference between these two vectors tells us the probability that each muscle loses its controllability induced by the ablation of the particular neuron(s).

Extended Data Figure 8(a) shows the probability with which muscles will lose their controllability after the individual DB neurons or the whole DB class is ablated. Each cross indicates a direct connection between a neuron and a muscle cell. One can see that a direct connection does not mean that the ablation of this neuron will affect the controllability of this muscle. Moreover, even if there is not a direct connection from a neuron to a muscle, the ablation of this neuron can also reduce the controllability of this muscle.

Extended Data Figure 8(b) shows the probability with which control is lost over each muscle following the ablation of PDB, a single-neuron class. Extended Data Figures 8(c,d), 9 and 10 show the role of individual neurons within other large classes, including AVA, AS, DA, DD, VA, VB, and VD.

D. Additional analyses

Other sensory neurons as input: In the experiments we approximated gentle touch with freely-moving worms behaving spontaneously on food. In this scenario, the animals necessarily receive further stimuli from their environment. This prompted us to assess

whether our key predictions were valid for gentle touch only or for any sensory input that triggers locomotor response. We repeated the control analysis described in Sec. II.B, yet with other sensory neurons as input, specifically FLP, PVD, and ASH, which, like the touch neurons, are sufficient to induce locomotion. Using these new classes as inputs gave rise to the same set of predictions, i.e. precisely the same neuron classes were theorised to be essential for locomotion. This then suggests that our results capture general locomotory responses to arbitrary sensory stimuli, rather than being specific to gentle touch.

As with any model, we expect that the inclusion of more detailed biological information would further enhance the abilities of our framework. For example, knowledge of the weight and sign of each synapse (excitatory vs inhibitory) would allow us to take into account the energy required for control. With even more physiological information (e.g. dynamical activity patterns) one could also calculate *control time*. These additional ingredients would allow us to investigate not just *if* control is achieved but *how* (via which pathways and mechanisms).

S1. White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. The structure of the nervous system of the nematode Caenorhabditis elegans: the mind of a worm. *Phil. Trans. R. Soc. B* **314**, 1–340 (1986).

S2. Hall, D. H. & Russell, R. L. The posterior nervous system of the nematode caenorhabditis elegans: serial reconstruction of identified neurons and complete pattern of synaptic interactions. *J. Neurosci.* **11**, 1–22 (1991).

S3. Dixon, S. J. & Roy, P. J. Muscle arm development in caenorhabditis elegans. *Development* **132**, 3079–3092 (2005).

S4. Jarrell, T. A. *et al.* The connectome of a decision-making neural network. *Science* **337**, 437–444 (2012).

S5. Bargmann, C. I. & Marder, E. From the connectome to brain function. *Nature Methods* **10**, 483–490 (2013).

S6. Towlson, E. K., Vértes, P. E., Ahnert, S. E., Schafer, W. R. & Bullmore, E. T. The rich club of the c. elegans neuronal connectome. *J. Neurosci.* **33**, 6380–6387 (2013).

S7. Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Rev. Neurosci.* **10**, 186–198 (2009).

S8. Chen, B. L., Hall, D. H. & Chklovskii, D. B. Wiring optimization can relate neuronal structure and function. *Proc. Natl. Acad. Sci. USA* **103**, 4723–4728 (2006).

S9. Varshney, L. R. *et al.* Structural properties of the Caenorhabditis elegans neuronal network. *PLoS Comput. Biol.* **7**, e1001066 (2011).

S10. Chen, B. L.-J. *Neuronal network of C. elegans: from anatomy to behavior.* Ph.D. thesis, Cold Spring Harbor Laboratory (2007).

S11. Neuronal connectivity II. `http://www.wormatlas.org/neuronalwiring.html`. Accessed: 2016-09-05.

S12. Lin, C. T. Structural controllability. *IEEE Transactions on Automatic Control* **19**, 201–208 (1974).

S13. Liu, Y.-Y., Slotine, J.-J. & Barabási, A.-L. Controllability of complex networks. *Nature* **473**, 167–173 (2011).

S14. Jia, T. *et al.* Emergence of bimodality in controlling complex networks. *Nature Commun.* **4** (2013).

S15. Menichetti, G., DallAsta, L. & Bianconi, G. Network controllability is determined by the density of low in-degree and out-degree nodes. *Phys. Rev. Lett.* **113**, 078701 (2014).

S16. Gao, J., Liu, Y.-Y., D'Souza, R. M. & Barabási, A.-L. Target control of complex networks. *Nature Commun.* **5** (2014).

S17. Altun, Z. & Hall, D. Nervous system, general description. In *WormAtlas* (2011). URL doi:10.3908/wormatlas.1.18.

S18. Cowan, N. J., Chastain, E. J., Vilhena, D. A., Freudenberg, J. S. & Bergstrom, C. T. Nodal dynamics, not degree distributions, determine the structural controllability of complex networks. *PloS One* **7**, e38398 (2012).

S19. Zhao, C., Wang, W.-X., Liu, Y.-Y. & Slotine, J.-J. Intrinsic dynamics induce global symmetry in network controllability. *Sci. Rep.* **5**, 8422 (2015).

S20. Kalman, R. E. Mathematical description of linear dynamical systems. *J. Soc. Indus. Appl. Math. Ser. A* **1**, 152–192 (1963).

S21. Gu, S. *et al.* Controllability of structural brain networks. *Nature Commun.* **6** (2015).

S22. Murota, K. & Poljak, S. Note on a graph-theoretic criterion for structural output controllability. *IEEE Transactions on Automatic Control* **35**, 939–942 (1990).

S23. Murota, K. *Systems analysis by graphs and matroids: structural solvability and controllability*, vol. 3 (Springer Science & Business Media, 2012).

S24. Nepusz, T. & Vicsek, T. Controlling edge dynamics in complex networks. *Nature Phys.* **8**, 568–573 (2012).

S25. Ruths, J. & Ruths, D. Control profiles of complex networks. *Science* **343**, 1373–1376 (2014).

S26. Shields, R. W. & Pearson, J. B. Structural controllability of multi-input linear systems. *IEEE Transactions on Automatic Control* **21**, 203–212 (1976).

S27. Poljak, S. On the generic dimension of controllable subspaces. *IEEE Transactions on Automatic Control* **35**, 367–369 (1990).

S28. Ahuja, R. K., Magnanti, T. L. & Orlin, J. B. *Network Flows: Theory, Algorithms, and Applications* (Prentice Hall, 1993).

S29. Dion, J.-M., Commault, C. & Van Der Woude, J. Generic properties and control of linear structured systems: a survey. *Automatica* **39**, 1125–1144 (2003).

S30. Hopcroft, J. E. & Karp, R. M. An $n\hat{5}/2$ algorithm for maximum matchings in bipartite graphs. *SIAM Journal on computing* **2**, 225–231 (1973).

S31. Chalfie, M. *et al.* The neural circuit for touch sensitivity in Caenorhabditis elegans. *J. Neurosci.* **5**, 956–964 (1985).

S32. Zhen, M. & Samuel, A. D. *C. elegans* locomotion: small circuits, complex functions. *Curr. Opin. Neurobiol.* **33**, 117–126 (2015).

S33. Donnelly, J. L. *et al.* Monoaminergic orchestration of motor programs in a complex c. elegans behavior. *PLoS Biol.* **11**, e1001529 (2013).

S34. Haspel, G. & O'Donovan, M. J. A perimotor framework reveals functional segmentation in the motoneuronal network controlling locomotion in Caenorhabditis elegans. *J. Neurosci.* **31**, 14611–14623 (2011).

S35. Haspel, G., O'Donovan, M. J. & Hart, A. C. Motoneurons dedicated to either forward or backward locomotion in the nematode Caenorhabditis elegans. *J. Neurosci.* **30**, 11151–11156 (2010).

S36. Gjorgjieva, J., Biron, D. & Haspel, G. Neurobiology of Caenorhabditis elegans locomotion: where do we stand? *Bioscience* **64**, 476–486 (2014).

S37. Faumont, S. *et al.* An image-free opto-mechanical system for creating virtual environments and imaging neuronal activity in freely moving Caenorhabditis elegans. *PLoS One* **6**, e24666 (2011).

S38. Kawano, T. *et al.* An imbalancing act: gap junctions reduce the backward motor circuit activity to bias c. elegans for forward locomotion. *Neuron* **72**, 572–586 (2011).

S39. Wen, Q. *et al.* Proprioceptive coupling within motor neurons drives c. elegans forward

locomotion. *Neuron* **76**, 750–761 (2012).

S40. White, J., Southgate, E. & Thomson, J. Mutations in the Caenorhabditis elegans unc–4 gene alter the synaptic input to ventral cord motor neurons. *Nature* **355**, 838–841 (1992).

S41. Harbinder, S. *et al.* Genetically targeted cell disruption in Caenorhabditis elegans. *Proc. Natl. Acad. Sci. USA* **94**, 13128–13133 (1997).

S42. Bargmann, C. I. & Avery, L. Laser killing of cells in caenorhabditis elegans. *Methods in Cell Biology* **48**, 225–250 (1995).

S43. Aurelio, O., Hall, D. H. & Hobert, O. Immunoglobulin-domain proteins required for maintenance of ventral nerve cord organization. *Science* **295**, 686–690 (2002).

S44. Yemini, E., Jucikas, T., Grundy, L. J., Brown, A. E. & Schafer, W. R. A database of Caenorhabditis elegans behavioral phenotypes. *Nature Methods* **10**, 877–879 (2013).

S45. Schafer, W. R. & Strimmer, J. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Stat. Appl. Genet. Mol. Biol.* **4** (2005).

S46. Stephens, G. J., Johnson-Kerner, B., Bialek, W. & Ryu, W. S. Dimensionality and dynamics in the behavior of c. elegans. *PLoS Comput. Biol.* **4**, e1000028 (2008).

S47. Gray, J. M., Hill, J. J. & Bargmann, C. I. A circuit for navigation in Caenorhabditis elegans. *Proc. Natl. Acad. Sci. USA* **102**, 3184–3191 (2005).

S48. Huang, K. M., Cosman, P. & Schafer, W. R. Machine vision based detection of omega bends and reversals in C. elegans. *J. Neurosci. Methods* **158**, 323–336 (2006).

S49. Galán, R. F. On how network architecture determines the dominant patterns of spontaneous neural activity. *PLoS One* **3**, e2148 (2008).

S50. Honey, C. *et al.* Predicting human resting-state functional connectivity from structural connectivity. *Proc. Natl. Acad. Sci. USA* **106**, 2035–2040 (2009).

S51. Coron, J.-M. *Control and Nonlinearity* (American Mathematical Society, 2009).

S52. Whalen, A. J., Brennan, S. N., Sauer, T. D. & Schiff, S. J. Observability and controllability of nonlinear networks: The role of symmetry. *Phys. Rev. X* **5**, 011005 (2015).

S53. Muldoon, S. F. *et al.* Stimulation-based control of dynamic brain networks. *PLoS Comput. Biol.* **12**, e1005076 (2016).