CONNECTOME-CONSTRAINED LATENT VARIABLE MODELS OF WHOLE-BRAIN NEURAL ACTIVITY

Anonymous authors
Paper under double-blind review

ABSTRACT

Brain-wide measurements of activity and anatomical connectivity of the *C. elegans* nervous system in principle allow for the development of detailed mechanistic computational models. However, there are several challenges. We often do not have direct experimental access to important modeling details such as single-neuron dynamics and the signs and strengths of the synaptic connectivity. Further, neural activity can only be measured in a subset of neurons, often indirectly via calcium imaging, and significant trial-to-trial variability has been observed. To overcome these challenges, we introduce a connectome-constrained latent variable model (CC-LVM) of the unobserved voltage dynamics of the entire *C. elegans* nervous system and the observed calcium signals. We used the framework of variational autoencoders to fit parameters of the mechanistic simulation constituting the generative model of the LVM to calcium imaging observations. A variational approximate posterior distribution over latent voltage traces for all neurons is efficiently inferred using an inference network, and constrained by a prior distribution given by the biophysical simulation of neural dynamics. When applied to a recent dataset, we find that connectomic constraints enable our LVM to predict the activity of neurons whose activity were withheld. We explored models with different degrees of biophysical detail, and found that models with conductance-based synapses provide markedly better predictions than current-based synapses for this system.

1 INTRODUCTION

The anatomical connectivity of the entire *C. elegans* nervous system, including both chemical and electrical synapses, has been known for several decades (White et al., 1986; Varshney et al., 2011; Witvliet et al., 2020). However, well-calibrated and predictive connectome-constrained mechanistic simulations of this nervous system have yet to be demonstrated (Kim et al., 2019; Warrington et al., 2019; Usuyama et al., 2012; Izquierdo & Beer, 2013; Gleeson et al., 2018). This is because currently available experimental data are insufficient to completely constrain computational models. First, the single-neuron and synapse dynamics are generally unknown. Second, the connectome does not directly inform the signs and strengths of individual synapses. Third, the response properties of sensory neurons are incompletely known. Further, it is unclear what level of biophysical detail is necessary to reproduce the essential computational function of the *C. elegans* nervous system.

We use recently available whole brain calcium imaging data to constrain the missing parameters in a connectome-constrained, biophysically detailed simulation of the *C. elegans* nervous system. We start with a simplified non-spiking passive point-neuron model of the voltage dynamics of individual neurons in the circuit. We model inputs to the neurons from electrical synapses, and model nonlinear chemical synapses with either current-based or conductance-based biophysics. The challenge of fitting such a model to data is two-fold. First, the voltage dynamics of the neurons are not directly observed, but rather indirectly measured via their slow calcium dynamics. Second, there is significant trial-to-trial variability in neural activity, suggesting strong initial state dependence in the neural responses even to the same sensory stimulus (Gordus et al., 2015). These issues can both be addressed by treating the collective voltage signals of all neurons in the nervous system as an unobserved latent variable whose dynamics are determined by the simplified connectome-constrained biophysical simulation with unknown neuronal and synaptic parameters.

Our connectome-constrained latent variable model (CC-LVM) of voltage dynamics of the entire *C. elegans* nervous system is a large-scale latent variable model with a very high-dimensional latent space consisting of voltage dynamics of 300 neurons over 5 minutes of time. The generative model for these latent variables are described by stochastic differential equations modeling the nonlinear dynamics of the network activity, and a novel differential equation describing how the calcium signals signals are generated from the voltage dynamics of individual non-spiking neurons. We developed a variational autoencoder based framework for inferring the unobserved voltage dynamics from the observed calcium dynamics of only a subset of the neurons in the nervous system. An inference network enables efficient variational inference of the latent variables, and fitting the unknown neuronal and synaptic parameters of our connectome and biophysics based generative model.

The compact nervous system of *C. elegans* makes it an ideal platform for systems neuroscience. The *C. elegans* nervous system consists of 300 neurons, divided into 118 distinct classes (often bilaterally symmetric pairs) (Sulston & Horvitz, 1977; Kimble & Hirsh, 1979; Sulston, 1983). These neurons can largely be categorized as sensory neurons, interneurons, and motor neurons. The majority of these neurons, about 200, are concentrated in the head, forming the brain of the worm. The synaptic connectivity of *C. elegans* has been mapped with electron microscopy, providing researchers with a complete connectome (White et al., 1986; Witvliet et al., 2020).

We apply the CC-LVM to whole-brain calcium imaging data which captures 170 of the 300 neurons in multiple worms as they respond to chemosensory stimuli (Yemini et al., 2019). In principle, an accurate model of the nervous system constrained by incomplete activity measurements but complete connectivity measurements can enable accurate predictions of activity in neurons which were not recorded. We tested this hypothesis by using the CC-LVM to predict the activity of neurons which were measured, but whose activity was withheld during model training. We also used the model to predict the activity of entire worms, by holding out single trials during training. The CC-LVM predicted activity of withheld neurons and withheld worms significantly better than models unconstrained by the connectome, demonstrating the utility of the connectome even when little is known about the signs and strengths of individual connections. Further, we found that models with conductance-based synapses provide superior predictions to models with current-based synapses. CC-LVMs thus provide a new tool for connectome and activity constrained modeling of neural circuits and for discovering the appropriate level of detail of biophysical model for a given system.

1.1 Prior Work

Previous work has been focused on creating network simulations based on anatomical connectome data, with unknown single neuron synaptic biophysics (Liu et al., 2018; Gleeson et al., 2018; Warrington et al., 2019; Olivares et al., 2021; Kim et al., 2019). These network models provided a holistic view of the entire nervous system and were validated by comparing simulated locomotion with movements of live animals (Izquierdo & Beer, 2016). However, they were not fit against prerecorded calcium fluorescence data, so their simulated neuronal activities are unconfirmed. Latent variable models (LVM) capitalize on recent advances in machine learning to uncover structure from data generated by increasingly advanced neural interfacing technologies (Pei et al., 2021). This framework has been used in other contexts to infer neuronal voltage dynamics from high-dimensional calcium fluorescence recordings (Pandarinath et al., 2018; Aitchison et al., 2017; Speiser et al., 2017). Despite their utility, these models do not account for connectomic data which can greatly improve their predictive power. Our work employs both LVMs and connectomic constraints to create a model that is informed by both neural population dynamics and network structure.

Recently, Bayesian models have been widely applied in neuronal datasets to infer the spiking activity, neural dynamics, and connectivity (Warrington et al., 2019; Aitchison et al., 2017; Speiser et al., 2017). In particular, Warrington et al. (2019). use sequential Monte Carlo (SMC) to impute the intracellular voltage potentials of *C. elegans* neurons from 49 recorded calcium traces (Kato et al., 2015). Their work combined neuronal, body and calcium observation simulators in order to model locomotion and is the first attempt to use the full anatomical connectome to infer unobserved neural voltage dynamics (Boyle et al., 2012; Marblestone, 2014). Their simulator produces a series of exemplar neuronal voltage traces but they do not evaluate the imputed traces of unmeasured neurons against real data so the accuracy of their inferences is unknown. Furthermore, they explore how different methods for approximate Bayesian inference affect parameter estimation on simulated data and show that their method is best. Our model uses a variational auto-encoder (VAE) to perform

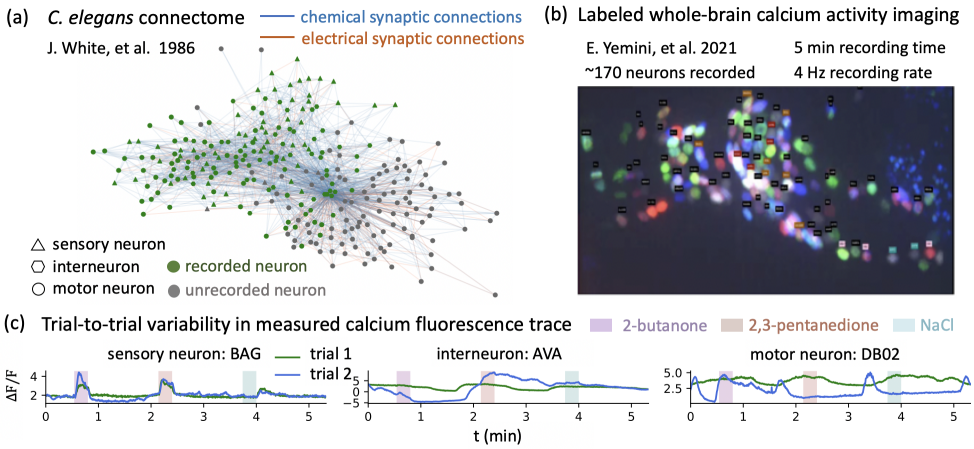

Figure 1: **Anatomical and functional whole-brain datasets available in *C. elegans*. a.** The connectome of *C.elegans* nervous system, with chemical connections in blue and electrical connections in orange (White et al., 1986). Sensory, inter, and motor neurons labeled with different shapes, and neurons recorded in the functional imaging dataset (Yemini et al., 2019) are colored green. **b.** Labeled whole-brain calcium imaging of 170 neurons in the head of the worm while a panel of chemosensory stimuli were presented Yemini et al. (2019). Deterministic multicolor labels allow for the identification of all neurons. **c.** Examples of fluorescence activity traces of three measured neurons, with stimulus delivery marked by the colored bars. Note the significant trial-to-trial variability in neural activity.

inference instead of SMC which generates a different type of distribution and is more computationally efficient because it allows direct sampling from the approximate posterior. Additionally, the CC-LVM is trained on significantly more neurons, and we validate our model by evaluating its ability to predict the activity of neurons held out from the training data. Finally, rather than study how various inference methods can affect parameter estimation in a toy model, we search a space of generative models to optimize voltage predictions by comparing them to real calcium data.

2 Connectome-constrained latent variable model

We constructed a connectome-constrained latent variable model (CC-LVM) of the *C. elegans* nervous system, where each node in our network represents a specific neuron in the animal. The activity of each neuron was modeled with a latent variable analogous to voltage. The dynamics of each neuron in the network is represented with a stochastic non-spiking passive leaky integrator equation with learned time constants and resting membrane potentials (Dayan et al., 2003). The neurons were coupled by both chemical and electrical synapses with learned weights (Dayan et al., 2003). We allowed these weights to be non-zero only where the connectome indicates the existence of synapses.

Given a set of learned parameters (weights, time constants, and resting membrane potentials), the dynamics of the network defines a prior distribution over neural activity trajectories. Importantly, the stochastic nature of the dynamics allows for deviations from perfectly deterministic dynamics. This allows the model to accommodate the observed variability in single-neuron dynamics. This variability has several potential sources, including the unmeasured initial states of the neurons and our incomplete knowledge of the sensory inputs driving the nervous system.

A latent variable model of this scale with a nonlinear generative model defined by the stochastic dynamics is difficult to fit. To address this challenge, we used the probabilistic inference framework of variational auto-encoders (VAE) (Rezende et al., 2014) to train a black-box voltage inference network to predict a posterior distribution over neural activity trajectories. We then used this inference network to train the parameters of the CC-LVM. The resulting LVM has a biologically realistic generative model of the neural dynamics of the *C. elegans* nervous system, and a black-box temporal convolutional inference network which, given sensory stimulus and calcium imaging measurements, predicts a factorized Gaussian distribution over the voltages of all the neurons in the network. Several variants of the LVM were developed: we tested models in which the synaptic connections were modeled as either *current-based* or *conductance-based* (Dayan et al., 2003), and evaluated different levels of connectome constraint. The LVM was optimized with the ELBO (evidence lower bound) objective, which includes both a reconstruction loss to maximize the data likelihood and a Kullback-Leibler (KL) divergence to minimize the distance between prior and posterior distributions.

2.1 NETWORK MODEL WITH PASSIVE POINT-NEURON VOLTAGE DYNAMICS

Neurons in the *C. elegans* nervous system are largely non-spiking (Bargmann, 1998), so we model the voltage dynamics for these neurons as a passive point neurons with a single electrical compartment. Let $\boldsymbol{v} \in \mathbb{R}^N$ denote the voltages of the N neurons. The voltage $v_i(t)$ for each post-synaptic neuron i at the time t was calculated using a first-order leaky integrator equation given by

$$\tau_i \dot{v}_i(t) + v_i(t) = s_i^c(t) + s_i^e(t) + v_i^{rest} + h_i(t), \tag{1}$$

where τ_i is the voltage time constant, h_i is the chemosensory input to the sensory neurons, s_i^c is the chemical synaptic input, s_i^e is the electrical synaptic input, v_i^{rest} is the resting neuron voltage.

We studied two variations of the model, the *current-based model* and the *conductance-based model*, which differ in their formulations of chemical synaptic input s_i^c. Since neurons in the *C. elegans* nervous system are largely non-spiking (Bargmann, 1998), we model the chemical synapses as having graded release of neurotransmitter, rather than the all-or-none quantal release seen in spiking neurons. In both models, we model the amount of neurotransmitter released as proportional to the pre-synaptic voltage v_j, following softplus activation $g(\cdot)$ which sets a minimum voltage below which there is no synaptic release, leading to $W_{ji}^c g(v_j(t))$. In our current-based model, synaptic input s_i^c to a post-synaptic neuron i is directly proportional pre-synaptic neurotransmitter concentration:

$$s_i^c(t) = \sum_j^N W_{ji}^c g(v_j(t)), \tag{2}$$

where W_{ji}^c represents the chemical synaptic weight between pre-synaptic neuron j and post-synaptic neuron i. W_{ji}^c can be either positive or negative depending on whether the synaptic connection is excitatory or inhibitory. If the connection between neurons j and i is nonexistent, W_{ji}^c is set to zero. In the conductance-based model, we model the synaptic current entering the post-synaptic neuron, with more biophysical detail, as

$$s_i^c(t) = \sum_j^N (E_{ji} - v_i(t)) W_{ji}^c g(v_j(t)). \tag{3}$$

Here, the pre-synaptic neurotransmitter concentration $W_{ji}^c g(v_j(t))$ is more accurately modeled as proportional to the conductance at the post-synaptic terminal. The post-synaptic current is then given by the product of the synaptic conductance and the difference between the post-synaptic voltage $v_i(t)$ and the synaptic reversal potential E_{ji}.

In contrast to the current-based synapse whose input is independent of the post-synaptic voltage $v_i(t)$, the conductance-based synapse model more accurately also has a dependence on the post-synaptic voltage. In addition, this model decouples the sign of the synapse from the strength. The reversal potential of a synapses E_{ji} dictates whether a synapse is excitatory or inhibitory. A large and positive E_{ji} corresponds to an excitatory synapse causing depolarization of the post-synaptic neuron. And an inhibitory synapse will have a negative E_{ji} causing hyperpolarization. In this model, we can now independently train the sign of a synapse and its non-negative strength W_{ji}^c, which is not easily possible with current-based synapses. In both the current-based and conductance-based models, the following equation was used to represent electrical synaptic inputs:

$$s_i^e(t) = \sum_j^N W_{ji}^e (v_j(t) - v_i(t)), \tag{4}$$

where W_{ji}^e is restricted to be non-negative and $v_j - v_i$ is the potential difference between presynaptic and postsynaptic neurons. We restrict $W_{ji}^e = W_{ij}^e$. To directly compare the outputs of the LVM to neural activity measurements, our model must generate calcium signals from the voltage traces. We model the calcium concentration $[\mathrm{Ca}]_i$ of each neuron i as a first-order leaky integrator, driven by voltage-gated calcium channels with the same nonlinear current-voltage (I-V) function $g(v_i)$:

$$\tau_{[\mathrm{Ca}]} [\dot{\mathrm{Ca}}]_i(t) + [\mathrm{Ca}]_i(t) = g(v_i(t)), \tag{5}$$

where $\tau_{[\mathrm{Ca}]}$ is a time constant shared across all neurons. We then map calcium concentration $[\mathrm{Ca}]$ into the measured calcium fluorescence signals $\boldsymbol{f}$ via an affine transform with scalar α^f and bias β^f,

$$f_i(t) = \alpha^f [\mathrm{Ca}]_i(t) + \beta^f + \sigma^f \epsilon_i^f(t), \tag{6}$$

with measurement noise represented by a noise amplitude σ_f and a noise term $\epsilon_i^f(t) \sim \mathcal{N}(0, 1)$.

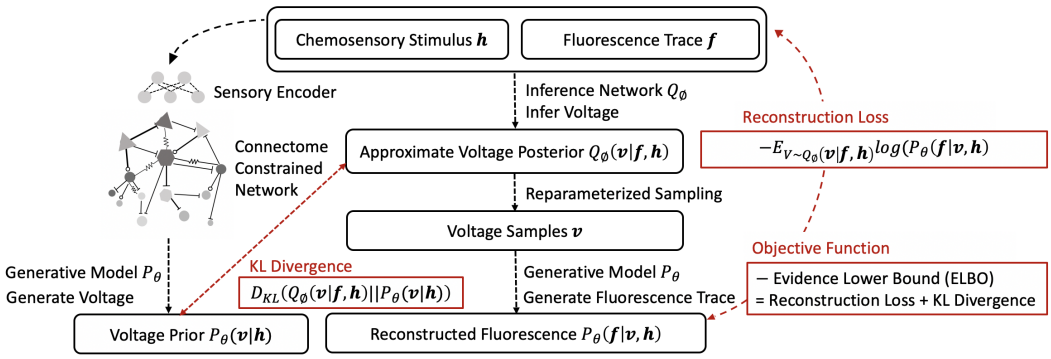

Figure 2: **Schematic of the connectome-constrained latent variable model (CC-LVM).** Sensory inputs are fed through a biophyiscally realistic connectome-constrained network to generate the voltage prior $P_\theta(\boldsymbol{v}|\boldsymbol{h})$. To train the model, we first used an inference network to infer from the data the approximate voltage posterior $Q_\phi(\boldsymbol{v}|\boldsymbol{f})$. Fluorescence traces are then reconstructed via a generative model $P_\theta(\boldsymbol{f}|\boldsymbol{v},\boldsymbol{h})$. We then compared the reconstructed fluorescence traces to the measured traces, and optimized the model with evidence lower bound (ELBO), with a combination of reconstruction loss and KL divergence.

2.2 Variational Inference

Note that voltage dynamics in equation 1 is a deterministic equation. We provide further flexibility to the model in order to model the significant trial-to-trial variability we observed in Figure 1c. Such variability could be caused by different initial state of neurons, or by unknown, unmeasured inputs from environment. In order to address this, we allow for the voltage dynamics to deviate from the deterministic dynamics, and introduce a Gaussian prior over these deviations. This results in the following stochastic version of the voltage dynamics,

$$\tau_i \dot{v}_i(t) + v_i(t) = s_i^c(t) + s_i^e(t) + v_i^{rest} + h_i(t) + \sigma_i^v \epsilon_i^v(t), \tag{7}$$

σ_i^v is the noise amplitude, and the standard normal noise term $\epsilon_i^v \sim \mathcal{N}(0,1)$. With this approach, we re-formulated the neuronal voltage as a latent variable $\boldsymbol{v}$ with the generative model given by 7 and must perform Bayesian inference to estimate their value. Exact Bayesian inference is intractable for this high-dimensional nonlinear stochastic dynamical system. For this reason, we developed a computationally efficient strategy for inferring an approximate posterior distribution $P(\boldsymbol{v}|\boldsymbol{f})$ over the latent variable $\boldsymbol{v}$, given the measured fluorescence data $\boldsymbol{f}$ using variational Bayes in the framework of the variational autoencoder (Kingma & Welling, 2014), see derivation in Appendix A.1.

2.3 Inference Network

The inference network $Q_\phi(\boldsymbol{v}|\boldsymbol{f},\boldsymbol{h})$ generates the approximate posterior distribution of the latent variable $\boldsymbol{v}$ given the measured fluorescence data $\boldsymbol{f}$ and the sensory inputs $\boldsymbol{h}$. It consists of multiple 1D convolutional layers that transform the input data into a Gaussian distributions $v_i(t) \sim \mathcal{N}(\mu_i(t), \sigma_i(t))$ for each neuron $i = 1, ..., N$ at every time point $t = 0, ..., D$. The approximate posterior distribution $Q_\phi(\boldsymbol{v}|\boldsymbol{f},\boldsymbol{h})$ is factorized as

$$Q_\phi(\boldsymbol{v}|\boldsymbol{f},\boldsymbol{h}) = \prod_{t=0,\ldots,D} Q_\phi\left(\boldsymbol{v}(t)|\boldsymbol{f}(0),..,\boldsymbol{f}(D),\boldsymbol{h}(0),..,\boldsymbol{h}(D)\right), \tag{8}$$

D is the length of the time window of data fed into the inference network, and ϕ is its parameters.

2.4 Generative Model

The generative model outputs the calcium fluorescence trace $\boldsymbol{f}$ and neuronal voltage $\boldsymbol{v}$ given sensory inputs $\boldsymbol{h}$. It can be formulated as

$$P_\theta(\boldsymbol{f},\boldsymbol{v}|\boldsymbol{h}) = P_\theta(\boldsymbol{f}|\boldsymbol{v},\boldsymbol{h}) P_\theta(\boldsymbol{v}|\boldsymbol{h}), \tag{9}$$

where θ contains parameters of the generative model. $P_\theta(\boldsymbol{v}|\boldsymbol{h})$ is a biophysically realistic connectome-constrained network with passive point-neuron dynamics. It outputs a prior distribution over voltage $\boldsymbol{v}$ given sensory input $\boldsymbol{h}$. A sensory encoder maps $\boldsymbol{h}$ via a multilayer perceptron (MLP) to sensory neurons. The stimuli is represented as a set of binary vectors $h_k(t)$, which are 1 given the presence

model		current-based			conductance-based				
weight	**parameter**	**learned dense**	**learned sparse**	**connectome sparsity**	**learned dense**	**learned sparse**	**connectome sparsity**	**connectome count**[1]	**connectome count**[2]
W^c_{ji}	range	$(-\infty, \infty)$	$(-\infty, \infty)$	$(-\infty, \infty)$	$[0, \infty)$	$[0, \infty)$	$[0, \infty)$	$[0, \infty)$	$[0, \infty)$
	T^c_{ji}	1	1	$\mathbf{1}[C^c_{ji} > 0]$	1	1	$\mathbf{1}[C^c_{ji} > 0]$	$\mathbf{1}[C^c_{ji} > 0]$	$\mathbf{1}[C^c_{ji} > 0]$
	M^c_{ji}	✓	✓	✓	✓	✓	✓	C^c_{ji}	C^c_{ji}
	α^c	0.01	0.01	0.01	0.01	0.01	0.01	✓	✓
W^e_{ji}	range	$[0, \infty)$	$[0, \infty)$	$[0, \infty)$	$[0, \infty)$	$[0, \infty)$	$[0, \infty)$	$[0, \infty)$	$[0, \infty)$
	T^e_{ji}	1	1	$\mathbf{1}[C^e_{ji} > 0]$	1	1	$\mathbf{1}[C^e_{ji} > 0]$	$\mathbf{1}[C^e_{ji} > 0]$	$\mathbf{1}[C^e_{ji} > 0]$
	M^e_{ji}	✓	✓	✓	✓	✓	✓	C^e_{ji}	C^e_{ji}
	α^e	0.01	0.01	0.01	0.01	0.01	0.01	✓	✓

Table 1: **Definitions and constraints for the LVM variants.** We evaluated a total of 8 model variants. Across these variants, we explored two methods of modeling chemical synapses: current-based and conductance-based (Sec. 2.1). We also tested several different levels of connectome constraint. (Sec. 2.6). For both the chemical synapse W^c_{ji} and electrical synapses W^e_{ji}, the weight matrix $W_{ji} = \alpha T_{ji} M_{ji}$, where T_{ji} is the sparsity matrix, M_{ji} is the matrix of connection magnitudes, and α is a global scalar. Trainable parameters are indicated by a ✓. We compared the connectome-constrained versions to two unconstrained learned models, a learned dense model in which all neurons are connected, and a learned sparse model with a L_2 regularization with connectome synapse count $\boldsymbol{C}$ using $|||\boldsymbol{M}| - |\boldsymbol{C}|||$.

of stimulus k at time t and 0 otherwise. $P_\theta(\boldsymbol{f}|\boldsymbol{v}, \boldsymbol{h})$ is another realistic model that maps the voltage $\boldsymbol{v}$ into reconstructed fluorescence $\boldsymbol{f}$, given sensory inputs $\boldsymbol{h}$. The voltages $\boldsymbol{v}$ are sampled from the approximate posterior distributions generated by the inference network. Details of both distributions are described in Appendix A.2.

2.5 Objective function for generative model and inference network

The variational objective of our latent variable model uses ELBO, which is defined as

$$\mathcal{L} = \mathbb{E}_{Q(\boldsymbol{v}|\boldsymbol{f},\boldsymbol{h})}[\log P_\theta(\boldsymbol{f}, \boldsymbol{v}|\boldsymbol{h}) - \log Q_\phi(\boldsymbol{v}|\boldsymbol{f}, \boldsymbol{h})]. \tag{10}$$

The reconstructed calcium fluorescence $\boldsymbol{f}$ is then fit to the measured fluorescence data by maximizing the ELBO for individual data points. The objective function is a combination of reconstruction loss between the reconstructed and measured fluorescence traces and the KL divergence between voltage prior and posterior (details in Appendix A.3, A.4):

$$\mathcal{L} = -\mathbb{E}_{\boldsymbol{v}\sim Q(\boldsymbol{v}|\boldsymbol{f},\boldsymbol{h})}\left[\log\left(P_\theta\left(\boldsymbol{f}|\boldsymbol{v}, \boldsymbol{h}\right)\right)\right] + D_{\mathrm{KL}}\left(Q_\phi(\boldsymbol{v}|\boldsymbol{f}, \boldsymbol{h}) \| P_\theta(\boldsymbol{v}|\boldsymbol{h})\right), \tag{11}$$

2.6 Connectome constraint for model weights

To build the model using chemical and electrical synapse count, $\boldsymbol{C}^c$ and $\boldsymbol{C}^e$, as a connectome constraint, we factorized the chemical synaptic weight W^c_{ji} as $W^c_{ji} = \alpha^c T^c_{ji} M^c_{ji}$, where α^c represents a global scaling factor which scales synapse counts to per-synapse currents or conductances depending on the model class. T^c_{ji} is a binary value which indicates the connectivity, and M^c_{ji} reflects the magnitude of connection strength. The electrical synaptic weights matrix W^e_{ji} is similarly defined as $W^e_{ji} = \alpha^e T^e_{ji} M^e_{ji}$. We imposed a connectome constraint on both the current-based and conductance-based models, and compared their performance to unconstrained models (Table 1, Appendix A.5). The anatomical connectome can be represented mathematically as two $N \times N$ matrices: one containing the chemical synapse counts ($\boldsymbol{C}^c$), and one containing the sizes of electrical synapses ($\boldsymbol{C}^e$). We designed the *connectome sparsity* constraint by fixing the sparsity of our model to that of the connectome. For chemical connections, the model sparsity $T^c_{ji} = \mathbf{1}[C^c_{ji} > 0]$, and likewise for electrical connections, $T^e_{ji} = \mathbf{1}[C^e_{ji} > 0]$. For both chemical and electrical connections, the magnitudes M^c_{ji} and M^e_{ji} are trainable.

We applied the connectome sparsity constraint to both current-based and conductance-based versions of the LVM. In the conductance-based model, we were able[1] to apply the even stronger *connectome count* constraint: directly assigning the synapse counts to the chemical connection magnitudes ($M^c_{ji} = C^c_{ji}$) and the synapse sizes to the eletrical connection magnitudes ($M^e_{ji} = C^e_{ji}$). Under this constraint, both the sparsity and magnitude are fixed, and only the global scaling factors α^c and α^e

[1]Since in the current-based model, the range of W^c_{ji} is $(-\infty, \infty)$, and because the signs of the synaptic connections are unknown, we could not directly apply the *connectome count* constraint to it.

are trainable. Additionally, we evaluated another variant of this constraint, *connectome count*[2]. This constraint uses parameters of *connectome count*[1] as initialization parameters for optimization, then the model allows M^c_{ji} to deviate from C^c_{ji} during re-training. We compared our connectome-constrained models to two unconstrained models. The *learned dense* model used fully connected sparsity matrices $\boldsymbol{T}^c$ and $\boldsymbol{T}^e$, with the magnitudes $\boldsymbol{M}^c$ and $\boldsymbol{M}^e$ as trainable parameters. The *learned sparse* model used L_2 regularization on both magnitude matrices: $|||\boldsymbol{M}^c| - |\boldsymbol{C}^c|||$ and $|||\boldsymbol{M}^e| - |\boldsymbol{C}^e|||$, so that the fraction of nonzero elements in $\boldsymbol{M}^c$ is equal to that of $\boldsymbol{C}^c$, and likewise for $\boldsymbol{M}^e$ and $\boldsymbol{C}^e$.

3 EXPERIMENTS

Here, we discuss the details of the connectome and labeled-whole brain imaging dataset we used. We also introduce two evaluations we conducted to study the predictive performance of CC-LVMs.

3.1 DATASET

We applied the CC-LVM to a calcium imaging dataset in which immobilized, pan-neuronally labeled *C.elegans* were presented with a panel of chemosensory stimuli (2-butanone, 2,3-pentanedione, and NaCl) (Yemini et al., 2019). The activity of 170 neurons in the head of the animal were measured across 21 individuals. Each recording lasts for 5 min, with an acquisition rate of 4 Hz. In this dataset, every neuron in the worm brain could be identified *in vivo* using a deterministic multicolor landmark strain. Thus, it was the first experiment in which whole brain dynamics could be mapped onto the anatomical connectome at the resolution of single neurons. The connectome constraints we applied utilized the anatomical connectivity data from White et al. (1986). The complete hermaphrodite *C.elegans* was reconstructed using electron microscopy, and the connections between its 300 neurons were mapped. The connectome contains 3464 chemical synapses and 1031 electrical synapses.

3.2 NEURON HOLDOUT EVALUATION

We hypothesized that the CC-LVM which incorporates the complete connectome but has access only to the measured activity of a subset of the neurons may nevertheless be able to accurately predict the activity of unmeasured neurons. We tested this hypothesis by performing neuron holdout evaluations, withholding a single pair of measured neurons from the model during training and then comparing the model predictions to the ground truth calcium fluorescence measurements. We repeated this evaluation for all 107 measured neuron pairs. For every neuron pair, we performed the neuron holdout evaluation for both current and conductance models with different levels of connectome constraints. For these experiments, we calculated the *correlation coefficient* (details in Appendix A.3) between the predicted and measured calcium fluorescence magnitudes to quantify the predictive performance (Figure 3a and Table 2, more predicted traces in Appendix A.7). For each evaluation, we trained 4 models with random initialization, and reported both the mean value and standard error (SE) for each model. We found that *connectome sparsity* and *connectome count* constrained models were better at predicting the activity of held out neurons than unconstrained models.

3.3 WORM HOLDOUT EVALUATION

Another evaluation we performed was to hold out the data from a handful of individual worms, train the model on the remaining worms, and predict the activity of the neurons of the held out individuals. For each of the 8 model variants, we trained on 15 worms, and tested the model on 6 withheld worms. In each case, we trained 4 models with different random initialization. For the worm holdout traces, we calculated the *correlation coefficient* between the predicted and measured calcium fluorescence. We also reported *ELBO*, *reconstruction loss* and *KL divergence* (Appendix A.3) on tested worms to represent the fitting performance of each model. We reported the mean and standard error (SE) from models trained with 4 different random initializations. We found that models using the *connectome count* constraints performed better at predicting calcium traces from held out worms, Table 2.

3.4 COMPARING THE DIFFERENT MODEL VARIANTS

The gold standard for comparing generative models in machine learning is to compare the model evidence $P(\boldsymbol{f})$ on a test dataset. However, this is intractable for many generative models, and so the

model			current-based			conductance-based				
holdout	**metric**		**learned dense**	**learned sparse**	**connectome sparsity**	**learned dense**	**learned sparse**	**connectome sparsity**	**connectome count**[1]	**connectome count**[2]
neuron	*corr* ↑	mean	-0.104	-0.001	0.087	0.081	0.259	0.262	0.318	**0.319**
		SE	±0.013	±0.017	±0.021	±0.012	±0.016	±0.016	±0.016	±**0.016**
worm	*corr* ↑	mean	0.465	0.346	0.440	0.457	0.464	0.463	**0.474**	0.468
		SE	±0.006	±0.007	±0.006	±0.006	±0.006	±0.006	±**0.006**	±0.006
	ELBO↑	mean	-0.772	-1.124	-0.803	-0.849	-0.911	-0.898	-0.854	**-0.697**
		SE	±0.040	±0.032	±0.036	±0.033	±0.032	±0.032	±0.036	±**0.066**
	recon ↓	mean	0.745	0.933	0.754	0.771	0.834	0.0821	0.744	**0.636**
		SE	±0.036	±0.030	±0.035	±0.033	±0.032	±0.031	±0.035	±**0.065**
	KLD ↓	mean	**0.028**	0.192	0.049	0.078	0.077	0.077	0.111	0.061
		SE	±**0.001**	±0.003	±0.002	±0.003	±0.004	±0.003	±0.006	±0.002

Table 2: **Holdout evaluation results.** We performed neuron holdout evaluations on each of the 107 measured neuron pairs in the dataset, and report the average *correlation coefficient* between the predicted and measured fluorescence traces (Sec. 3.2). We also performed worm holdout evaluations, witholding 6 of the 21 individuals from the training (Sec. 3.3). For the worm holdouts, we report several metrics: *correlation coefficient*, *ELBO*, *reconstruction loss*, and *KL divergence*. Overall, the conductance-based model under *connectome count* constraint produced the best predictions among the 8 models.

evidence lower bound (*ELBO*) is used for model fitting and sometimes also for model comparison. However, the *ELBO* is not a tight lower bound, and so is not guaranteed to provide the same ordering over models as the model evidence $P(\boldsymbol{f})$. In this paper, exploiting the interpretable nature of our mechanistic generative model, we developed a stricter criterion for comparing models based on the neuron holdout evaluation. Neuron holdout predictive performance is biologically meaningful, as it evaluates the ability of a model to make testable predictions about neurons which were not directly measured. We found that overall, the conductance-based CC-LVMs under *connectome count* constraint make the best neuron holdout predictions as quantified by the *correlation coefficient*. We found that model accuracies at neuron holdout prediction are not very tightly correlated with their test *ELBO*. While *connectome count*[2] still achieves the best performance, the *learned dense* model achieves the second best performance, despite performing worst in the neuron holdout evaluation. This suggests that the *learned dense* model has potentially discovered an equally accurate but mechanistically incorrect way to model the calcium imaging data.

Stronger connectome constraint improves the neuron holdout predictions. We found that our CC-LVM with the *connectome count* constraint produced significantly better predictions than CC-LVM with the *connectome sparsity* constraint, and both connectome constrained models outperformed unconstrained LVMs with the same dynamics (Table 2). This suggests that the neuron activity predictions are improved by stronger connectome constraint. Note that *connectome count*[2] achieves slightly better performance than *connectome count*[1], as its parameters get initialized from *connectome count*[1], then are allowed to deviate during re-training. We also explored whether CC-LVMs achieve better performance due solely to the smaller number of trainable parameters. Comparing the connectome-constrained models to the *learned sparse* LVM with sparsity regularization, we found that the learned sparse model performed worse. This indicates that the connectome constraint improves predictability due to its topology, not due to the number of trainable parameters.

Perhaps relatedly, we found superior predictions for interneurons and motor neurons, compared to sensory neurons. While the inputs to motor neurons and especially interneurons are largely contained in the connectome, the inputs to sensory neurons are not as well constrained by the connectome since they also respond to external stimuli. Instead, we trained a blackbox model to predict the tunings of sensory neurons to experimentally presented odors, but this likely does not fully characterize all the sensory inputs to the worm. This result additionally suggests the utility of the connectome, for neurons where the connectome provides information about the dominant sources of input.

Conductance-based models outperform current-based models. We found that conductance-based models achieve better performance in neuron holdout evaluations, despite needing to estimate the reversal potential $\boldsymbol{E}$ in addition to the non-negative strength of a chemical synapse $\boldsymbol{W}^c$. This could be for several reasons. The biophysical model of a conductance based synapse might be a more accurate description of reality. There are also two technical advantages to this model. First, conductance based synapses allow us to make the fullest use of the connectomic data by using the non-negative synapse counts. Second, separating the optimization of the sign of a synapse from its magnitude might allow for an easier optimization of both. In contrast, the sign and magnitude of a current-based synapse are coupled in the same parameter, and we have observed strong local minima preventing the efficient

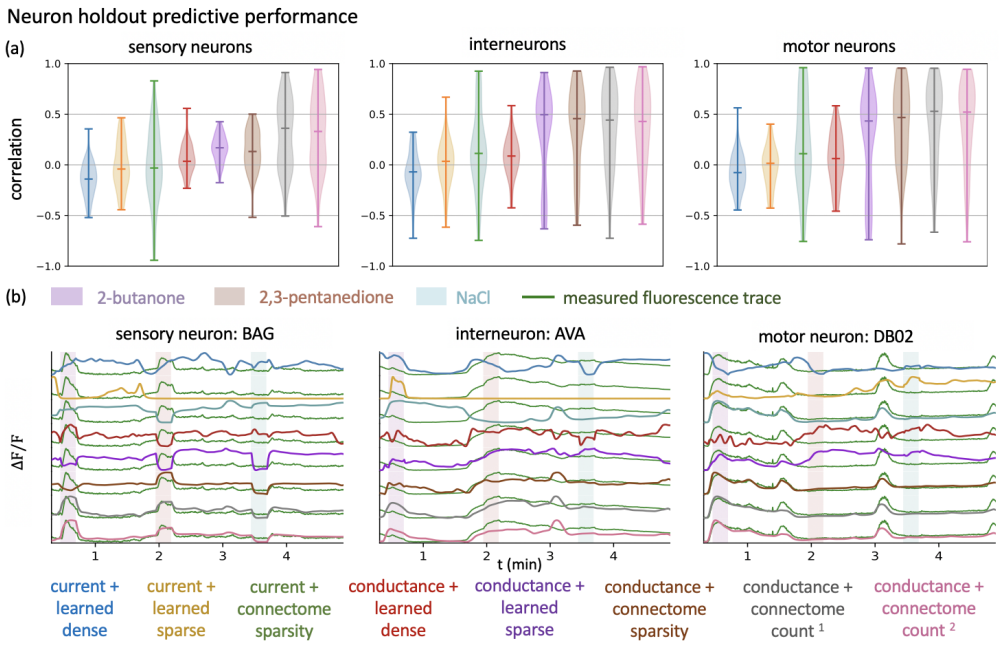

Figure 3: **LVM predictive performance across neuron types. a.** Violin plots comparing the distribution of correlation coefficients between predicted and measured neuron activity for the 8 LVM variants. Neurons are divided into sensory, inter, and motor neurons. **b.** Measured traces and predictions made by the 8 models for three selected neurons: sensory neuron BAG, interneuron AVA, and motor neuron DB02.

optimization of the sign and the magnitude. Accordingly, we found that conductance-based models converge to a connectome closer to the counts given by the original data with less variance due to the removal of the neuron sign parameter (Appendix A.6). We found that a synapse initialized as excitatory is unlikely to change its sign after optimization. For the same reasons, we are unable to effectively use the non-negative synapse counts contained in the connectomic data to inform the magnitude of the current-based synapse while leaving the sign unconstrained.

4 Discussion

In this work, we proposed a connectome-constrained latent variable model (CC-LVM) of the *C.elegans* nervous system. Our latent variable framework enables the estimation of the initial state of the network dynamics, and also captures modeling imperfections, which together enable us to fit models to calcium imaging data with significant trial-to-trial variability. In contrast to popular blackbox generative models used in LVMs in neural data modeling (Pei et al., 2021), we developed a biophysics based, connectome constrained, mechanistically detailed generative model for the neural dynamics. We evaluated two current-based and conductance-based models for synaptic connection. We also explored using connectomic data to constrain the model to different degrees of detail. Since CC-LVMs enable the prediction of neurons which were not used to fit the model, we were able to simulate and validate the process of making biologically relevant and experimentally testable predictions about neurons which have not yet been experimentally characterized.

We found the conductance-based model produces much better-correlated predictions in neuron holdout evaluation and worm holdout evaluation. And using the knowledge of the connectome can aid *in silico* predictions of individual neurons whose activity were not measured. We hope our mechanistic model can be used to predict the effects of perturbations or ablations, leading to a causal understanding of neural computation in neural circuits. The research in this paper was enabled by recent advances in whole brain calcium imaging in identified neurons (Yemini et al., 2019). While we have only analyzed the first datasets to be collected in this manner, we hope that a rich database of such recordings across a diversity of conditions will lead to achieving the dream of standard models of the *C. elegans* nervous system (Izquierdo & Beer, 2016) which generalize across experiments.

Reproducibility: We trained each of our LVMs on 1 Quadro RTX 8000, Training procedures are described in Appendix A.3, A.4. We will release our training software and datasets upon publication.

REFERENCES

Laurence Aitchison, Lloyd Russell, Adam Packer, Jinyao Yan, Philippe Castonguay, Michael Häusser, and Srinivas C Turaga. Model-based bayesian inference of neural activity and connectivity from all-optical interrogation of a neural circuit. In *Advances in Neural Information Processing Systems*, volume 2017, pp. 3487–3496. Neural Information Processing Systems Foundation, 2017.

Cornelia I. Bargmann. Neurobiology of the *Caenorhabditis elegans* genome. *Science*, December 1998. doi: doi:10.1126/science.282.5396.2028.

Jordan H. Boyle, Stefano Berri, and Netta Cohen. Gait modulation in *C. elegans*: an integrated neuromechanical model. *Front. Comput. Neurosci*, March 2012. doi: doi:10.3389/fncom.2012.00010.

Peter Dayan, Laurence F Abbott, et al. Theoretical neuroscience: computational and mathematical modeling of neural systems. *Journal of Cognitive Neuroscience*, 15(1):154–155, 2003.

Padraig Gleeson, David Lung, Radu Grosu, Ramin Hasani, and Stephen D. Larson. c302: a multiscale framework for modelling the nervous system of *Caenorhabditis elegans*. *Philos Trans R Soc Lond B Biol Sci*, September 2018. doi: doi:10.1098/rstb.2017.0379.

Andrew Gordus, Navin Pokala, Sagi Levy, Steven W. Flavell, and Cornelia I. Bargmann. Feedback from network states generates variability in a probabilistic olfactory circuit. *Cell*, April 2015. doi: doi:10.1016/j.cell.2015.02.018.

Eduardo J. Izquierdo and Randall D. Beer. Connecting a connectome to behavior: An ensemble of neuroanatomical models of *C. elegans* klinotaxis. *PLOS Computational Biology*, February 2013. doi: 10.1371/journal.pcbi.1002890.

Eduardo J. Izquierdo and Randall D. Beer. The whole worm: brain-body-environment models of *C. elegans*. *Curr Opin Neurobiol*, October 2016. doi: doi:10.1016/j.conb.2016.06.005.

Saul Kato, Harris S Kaplan, Tina Schrödel, Susanne Skora, Theodore H. Lindsay, Eviatar Yemini, Shawn Lockery, and Zimmer Manuel. Global brain dynamics embed the motor command sequence of *Caenorhabditis elegans*. *Cell*, Oct 2015. doi: doi:10.1016/j.cell.2015.09.034.

Jimin Kim, Julia A. Santos, Mark J. Alkema, and Eli Shlizerman. Whole integration of neural connectomics, dynamics and bio-mechanics for identification of behavioral sensorimotor pathways in *Caenorhabditis elegans*. *biorXiv preprint biorXiv:1711.01846*, August 2019. doi: doi:10.1101/724328.

Judith Kimble and David Hirsh. The postembryonic cell lineages of the hermaphrodite and male gonads in Caenorhabditis elegans. *Dev. Biol.*, 70(2):396–417, 1979. ISSN 00121606. doi: 10.1016/0012-1606(79)90035-6.

Diederik P. Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, May 2014.

Hexuan Liu, Jimin Kim, and Eli Shlizerman. Functional connectomics from neural dynamics: probabilistic graphical models for neuronal network of *Caenorhabditis elegans*. *Philos Trans R Soc Lond B Biol Sci*, September 2018. doi: doi:10.1098/rstb.2017.0377.

Adam Marblestone. Simple *C. elegans*. 2014. URL `https://github.com/adammarblestone/simple-C-elegans`.

Erick Olivares, Eduardo J. Izquierdo, and Randall D. Beer. A neuromechanical model of multiple network rhythmic pattern generators for forward locomotion in *C. elegans*. *Front. Comput. Neurosci*, February 2021. doi: doi:10.3389/fncom.2021.572339.

Chethan Pandarinath, Daniel J. O'Shea, Jasmine Collins, Rafal Jozefowicz, Jonathan C. Kao, Eric M. Trautmann, Matthew T. Kaufman, Stephen I Ryu, Leigh R. Hochberg, Jaimie M. Henderson, Krishna V. Shenoy, L.F. Abbott, and David Sussillo. Inferring single-trial neural population dynamics using sequential auto-encoders. *Nature Methods*, October 2018. doi: doi:10.1038/s41592-018-0109-9.

Felix Pei, Joel Ye, David Zoltowski, Anqi Wu, Raeed H. Chowdhury, Hansem Sohn, Joseph E. O'Doherty, Krishna V. Shenoy, Matthew T. Kaufman, Mark Churchland, Mehrdad Jazayeri, Lee E. Miller, Jonathan Pillow, Il Memming Park, L. Eva Dyer, and Chethan Pandrinath. Neural latents benchmark '21: Evaluating latent variable models of neural population activity. *arXiv preprint arXiv:2109.04463*, September 2021.

Danilo Jimenez Rezende, Shakir Mohamed, and Daan Wierstra. Stochastic backpropagation and approximate inference in deep generative models. *arXiv preprint arXiv:1401.4082*, 2014.

Artur Speiser, Jinyao Yan, Evan Archer, Lars Buesing, Srinivas C. Turaga, and Jakob H. Macke. Fast amortized inference of neural activity from calcium imaging data with variational autoencoders. *arXiv preprint arXiv:1711.01846*, November 2017.

J E Sulston. Neuronal cell lineages in the nematode Caenorhabditis elegans. *Cold Spring Harbor Symposia on Quantitative Biology*, 48 Pt 2:443–452, 1983. ISSN 00917451. doi: 10.1101/sqb.1983.048.01.049.

J. E. Sulston and H. R. Horvitz. Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. *Dev. Biol.*, 56(1):110–156, 1977. ISSN 00121606. doi: 10.1016/0012-1606(77)90158-0.

Mamoru Usuyama, Chisato Ushida, and Ryuzo Shingai. A model of the intracellular response of an olfactory neuron in *Caenorhabditis elegans* to odor stimulation. *PLOS One*, August 2012. doi: 10.1371/journal.pone.0042907.

Lav R. Varshney, Beth L. Chen, Eric Paniagua, David H. Hall, and Dmitri B. Chklovskii. Structural properties of the Caenorhabditis elegans neuronal network. *PLoS Comput. Biol.*, 7(2), 2011. ISSN 1553734X. doi: 10.1371/journal.pcbi.1001066.

Andrew Warrington, Arthur Spencer, and Frank Wood. The virtual patch clamp: Imputing c. elegans membrane potentials from calcium imaging. *arXiv preprint arXiv:1907.11075*, 2019.

J.G. White, E. Southgate, J.N. Thomson, and S. Brenner. The structure of the nervous system of the nematode Caenorhabditis elegans. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 314(1165):1 LP – 340, 11 1986.

Daniel Witvliet, Ben Mulcahy, James K Mitchell, Yaron Meirovitch, Daniel K Berger, Yuelong Wu, Yufang Liu, Wan Xian Koh, Rajeev Parvathala, Douglas Holmyard, et al. Connectomes across development reveal principles of brain maturation in C. elegans. *bioRxiv*, 2020.

Eviatar Yemini, Albert Lin, Amin Nejatbakhsh, Erdem Varol, Ruoxi Sun, Gonzalo E. Mena, Aravinthan D.T. Samuel, Liam Paninski, Vivek Venkatachalam, and Oliver Hobert. NeuroPAL: A Neuronal Polychromatic Atlas of Landmarks for Whole-Brain Imaging in C. elegans. *bioRxiv*, pp. 676312, 2019.

A APPENDIX

A.1 VARIATIONAL INFERENCE

We used bayesian inference (VAE) to compute the posterior probability over the latent variables $\boldsymbol{v}$ (neuron voltages) given the data $\boldsymbol{f}$ (measured calcium fluorescence),

$$P(\boldsymbol{v}|\boldsymbol{f}) = \frac{P(\boldsymbol{f}|\boldsymbol{v})P(\boldsymbol{v})}{P(\boldsymbol{f})} \tag{12}$$

To solve this problem, we need to compute the model evidence,

$$P(\boldsymbol{f}) = \int P(\boldsymbol{f}|\boldsymbol{v})P(\boldsymbol{v})d\boldsymbol{v}. \tag{13}$$

However, computing $P(\boldsymbol{f})$ is an intractable problem, so we use variational inference to approximate the value of $P(\boldsymbol{v}|\boldsymbol{f})$ (Kingma & Welling, 2014). Denoting the approximate posterior $Q(\boldsymbol{v}|\boldsymbol{f})$, we can formulate evidence lower bound (ELBO) as

$$\log P(\boldsymbol{f}) \geq \mathcal{L} = \mathbb{E}_{Q(\boldsymbol{v}|\boldsymbol{f})}[\log P(\boldsymbol{f}, \boldsymbol{v}) - \log Q(\boldsymbol{f}|\boldsymbol{v})], \tag{14}$$

which tightens the lower bound on $\log P(\boldsymbol{f})$ as $\mathcal{L}$ is maximized. This allows us to improve the recognition model and find the posterior such that $Q(\boldsymbol{v}|\boldsymbol{f}) \approx P(\boldsymbol{v}|\boldsymbol{f})$ since the Kullback-Leibler (KL) divergence, which measures the distance between the distributions Q and P, is minimized by maximizing ELBO.

A.2 PRIOR AND POSTERIOR DISTRIBUTION AT A DISCRETE TIME STEP

Using euler discretization with a time step of Δt, the first-order leaky integrator equation in equation 1 in the network model is

$$\tau_i(v_i(t+\Delta t) - v_i(t))/\Delta t + v_i(t) = s_i^c(t) + s_i^e(t) + v_i^{rest} + h_i(t+\Delta t), \tag{15}$$

Then the voltage $v_i(t+\Delta t)$ of neuron i at time step $t+\Delta t$ is formulated as

$$v_i(t+\Delta t) = \Delta t/\tau_i(s_i^c(t) + s_i^e(t) + v_i^{rest} + h_i(t+\Delta t) - v_i(t)) + v_i(t), \tag{16}$$

This equation indicates that the voltage state $v_i(t)$ at the current time step t is determined by a function over multiple components from the last time step $t - \Delta t$, including the past voltage state, the chemical and electric synaptic inputs as well as the sensory input

$$v_i(t+\Delta t) = \mathcal{F}(v_i(t), s_i^c(t), s_i^e(t), h_i(t+\Delta t)), \tag{17}$$

For the posterior distribution $Q_\phi(\boldsymbol{v}(t)|\boldsymbol{f}(t), \boldsymbol{h}(t))$, the inference network outputs $\mu_i(t)$ and $\sigma_i(t)$ for every neuron i at each time step t. The posterior distribution for the voltage generated by the inference network at each time step is $v_i(t) \sim \mathcal{N}(\mu_i(t), \sigma_i(t))$, given a sequence of measured fluorescence data $\boldsymbol{f}(0), ..., \boldsymbol{f}(D\Delta t)$ and sensory input $\boldsymbol{h}(0), ..., \boldsymbol{h}(D\Delta t)$ with total number of steps $D+1$. Then we used the reparameterization trick to sample the voltage $v_i(t)$ from this posterior distribution, to generate a sequence of voltage sample from $v_i(0), ..., v_i(D\Delta t)$.

For the prior distribution, we used the voltage samples $v_i(t)$ from the posterior distribution at the last step as the past voltage state in $\mathcal{F}$ to calculate the prior distribution $P_\theta(v_i(t+\Delta t)|h_i(t+\Delta t))$ at time step $t+\Delta t$. The prior distribution is defined as $\mathcal{N}(\mathcal{F}(\boldsymbol{v}(t), s_i^c(t), s_i^e(t), h_i(t+\Delta t)), \sigma_i^v)$. σ_i^v is the constant standard deviation of the prior distribution that is trainable during optimization. At time step 0, we define another trainable parameter $\mu_i(0)$ to represent the mean of the voltage prior distribution.

Similarly, we also used euler integration to discretize the first-order leaky integrator equation 5. The calcium concentration $[Ca]_i(t+\Delta t)$ of neuron i at time step $t+\Delta t$ is given by

$$[Ca]_i(t+\Delta t) = \Delta t/\tau_{[Ca]}(g(v_i(t)) - [Ca]_i(t)) + [Ca]_i(t), \tag{18}$$

We use a sequence of voltage samples $v_i(t)$, $t = 0, ..., D\Delta t$ as the inputs to this formula to get a sequence of calcium concentration $[Ca]_i(1), ..., [Ca]_i(D\Delta t)$. The initial state of concentration is estimated from the measured fluorescence trace at time step 0 using $[Ca]_i(0) = (F_i(0) - \beta^f)/\alpha^f$. Equation 6 in the generative model P_θ generates the output distributions $P_\theta(\boldsymbol{f}(t)|\boldsymbol{v}(t), \boldsymbol{h})$ of fluorescence trace at time step t as $\mathcal{N}(\alpha^f[Ca]_i(t) + \beta^f, \sigma^f)$. The values $\alpha^f, \beta^f, \sigma^f$ are trainable fluorescence parameters.

A.3 Formulations for objective functions and metrics

Here we provided the mathematical formulations for the objective functions and metrics we used in this work.

The evidence lower bound (ELBO) is a combination of reconstruction log likelihood and the negative KL divergence. The reconstruction log likelihood, given the predicted fluorescence trace $\hat{f}_i(t)$ of neuron i at time step t as output distribution $\mathcal{N}(\hat{f}_i(t), \sigma^f)$. And $\hat{f}_i$ is formulated as

$$\hat{f}_i(t) = \alpha^f[Ca](t) + \beta^f. \tag{19}$$

Given the measured fluorescence trace $f_i(t)$, number of neurons N, and total number of time steps $D+1$, we calculated the reconstruction negative log likelihood as the reconstruction loss

$$\mathcal{L}_{\text{recon}} = -\sum_{t=0}^{D\Delta t}(2\pi(\sigma^f)^2)^{-N/2}\exp\left(-\frac{1}{2(\sigma^f)^2}\sum_{i=1}^{N}(f_i(t) - \hat{f}_i(t))^2\right) \tag{20}$$

For KL divergence between voltage prior and posterior distribution. We represent the posterior distribution generated from the inference network as $\mathcal{N}(\mu_i(t), \sigma_i(t))$, for neuron i at each time step t. We then sample $v_i(t)$ from the posterior distribution. According to equation 17, we represent the mean $\hat{v_i}(t)$ of prior distribution $\mathcal{N}(\hat{v_i}(t), \sigma_i^v)$ as

$$\hat{v_i}(t) = \mathcal{F}(\mu_i(t), s_i^c(t), s_i^e(t), h_i(t+\Delta t)), \tag{21}$$

where σ_i^v is a trainable parameter.

Then we use a closed form of KL divergence defined between two univariate Gaussian distributions

$$D_{\text{KL}} = \sum_{t=0}^{D\Delta t}\sum_{i=1}^{N}\left(\log\left(\frac{\sigma_i^v}{\sigma_i(t)}\right) + \frac{(\sigma_i(t))^2 + (v_i(t) - \hat{v_i}(t))^2}{2(\sigma_i^v)^2} - \frac{1}{2}\right) \tag{22}$$

We minimize the objective function $\mathcal{L} = -ELBO$ during optimization given by

$$\mathcal{L} = \mathcal{L}_{\text{recon}} + D_{\text{KL}} \tag{23}$$

In the worm holdout evaluation, for each neuron pair, we calculate those metrics including *ELBO*, *reconstruction loss* and *KL divergence* on each holdout worm, and in total 6 worms. We also train each model setting under the same constraint with different initialization for 4 times. We calculate the mean and standard error of *ELBO*, *reconstruction loss* and *KL divergence* based on $4 \times 6 = 24$ samples for each model as our metrics.

For both neuron and worm holdout evaluation, we calculate the correlation coefficients between the measured $f_i(t)$ and reconstructed fluorescence traces $\hat{f}_i(t)$, the correlation coefficient for each neuron is simply

$$r_i = \frac{\sum_{t=0}^{D\Delta t} \left(\hat{f}_i(t) - \bar{\hat{f}}_i(t)\right)\left(f_i(t) - \bar{f}_i(t)\right)}{\sqrt{\sum_{t=0}^{D\Delta t} \left(\hat{f}_i(t) - \bar{\hat{f}}_i(t)\right)^2 \sum_{t=0}^{D\Delta t} \left(f_i(t) - \bar{f}_i(t)\right)^2}}. \tag{24}$$

In the neuron holdout evaluation, for each neuron pair, we train a model in which the ground truth of said neuron pair was removed from the training data. In total there are 107 pairs, and 170 measured neurons in total. We also trained 4 times with different initialization for each neuron pair and each of 8 connectome constraints for a total of $8 \times 107 \times 4 = 3424$ different models. We calculate the mean and standard error of r_i based on $4 \times 170 = 680$ samples for each model as our metrics.

In the worm holdout evaluation, we calculate this metrics on each holdout worm, and in total 6 worms. Each worm has 170 measured neurons. We also train each model setting under the same connectomic constraint with different initialization 4 times. We calculate the mean and standard error of r_i based on $6 \times 170 \times 4 = 4080$ samples for each model as our metrics.

A.4 Gradient-based optimization

We used PyTorch to minimize the ELBO using gradient descent with respect to the parameters of both the generative and inference models,

For LVMs which used *connectome count* constraint with α^c and α^e trainable. The objective function is

$$\mathcal{L} = \mathcal{L}(\tau, \tau_{[\text{Ca}]}, \alpha^c, \alpha^e, \alpha^f, \beta^f, \sigma^v, \sigma^f, \boldsymbol{E}, \text{MLP}, \phi) \tag{25}$$

For LVMs which used other constraints including *connectome count* and *learned dense*, *learned sparse*, where $\boldsymbol{M}^c$ and $\boldsymbol{M}^e$ trainable. The objective function is

$$\mathcal{L} = \mathcal{L}(\tau, \tau_{[\text{Ca}]}, \boldsymbol{M}^c, \boldsymbol{M}^e, \alpha^f, \beta^f, \sigma^v, \sigma^f, \boldsymbol{E}, \text{MLP}, \phi) \tag{26}$$

where ϕ is all the parameters of the recognition model. We used Adam with a learning rate scheduler to perform the optimization. We used a clamp in PyTorch to constrain the weight magnitude $\boldsymbol{M}^c$ and $\boldsymbol{M}^e$ to be non-negative in several models. We restrict the $\boldsymbol{M}^e$ to be symmetric by defining it as the sum of an upper triangular matrix and its transpose. Our simulation time step is 6.25ms. Each window size D lasts for 7.5s with 30 imaging steps. We used weights of 1 for both the reconstruction loss and KL divergence. We used a initial learning rate of $3e^{-4}$, with a learning rate schedular with step size of 50, and gamma of 0.5. We also set a gradient clip value of 1.

A.5 Connectome constraints for model weights

We constructed CC-LVMs on both current-based and conductance-based synapses with different levels of connectomic constraints. Given anatomical connectome data with two $N \times N$ matrices: chemical synapse counts ($\boldsymbol{C}^c$), and sizes of electrical synapses ($\boldsymbol{C}^e$). The chemical synaptic weight W^c_{ji} as $W^c_{ji} = \alpha^c T^c_{ji} M^c_{ji}$, where α^c represents a global scaling factor. T^c_{ji} is a binary value which indicates the connectivity, and M^c_{ji} reflects the magnitude of connection strength. The electrical synaptic weights matrix W^e_{ji} is similarly defined as $W^e_{ji} = \alpha^e T^e_{ji} M^e_{ji}$. We provide more implementation details of each constraint below,

conductance-based + *connectome count*[1] We directly assign the synapse counts to the chemical connection magnitudes ($M^c_{ji} = C^c_{ji}$) and the synapse sizes to the eletrical connection magnitudes ($M^e_{ji} = C^e_{ji}$). Under this constraint, both the sparsity T^c_{ji} and T^e_{ji} are fixed to the connectome sparsity, magnitude M^c_{ji} and M^e_{ji} are fixed to the connectome synapse count, and only the global scaling factors α^c and α^e are trainable. M^c_{ji} and M^e_{ji} are restricted to be non-negative.

conductance-based + *connectome count*[2] We initialize all model parameters with the parameters from the trained *conductance-based + connectome count*[1] model. And then the sparsity T^c_{ji} and T^e_{ji} are fixed to the connectome sparsity, global scaling factors α^c and α^e are fixed to 0.01. And

now magnitude M^c_{ji} and M^e_{ji} are trainable during the re-training. M^c_{ji} and M^e_{ji} are restricted to be non-negative.

conductance-based + *connectome sparsity* For chemical and electrical the model sparsities are $T^c_{ji} = \mathbf{1}[C^c_{ji} > 0]$ and $T^e_{ji} = \mathbf{1}[C^e_{ji} > 0]$ respectively. For both chemical and electrical connections, the magnitudes M^c_{ji} and M^e_{ji} are trainable. The sparsity T^c_{ji} and T^e_{ji} are fixed to the connectome sparsity and global scaling factors, α^c and α^e, are fixed to 0.01. M^c_{ji} and M^e_{ji} are restricted to be non-negative.

conductance-based + *learned dense* For chemical connections and electrical connections, the model is a fully connected with sparsity $T^c_{ji} = 1$ and $T^e_{ji} = 1$ respectively. For both chemical and electrical connections, the magnitudes M^c_{ji} and M^e_{ji} are trainable. The sparsity matrices, T^c_{ji} and T^e_{ji} are fixed. Global scaling factors α^c and α^e are also fixed to 0.01. M^c_{ji} and M^e_{ji} are restricted to be non-negative.

conductance-based + *learned sparse* For chemical and electrical connections, the model is fully connected with the sparsity $T^c_{ji} = 1$ and $T^e_{ji} = 1$ respectively. For both chemical and electrical connections, the magnitudes M^c_{ji} and M^e_{ji} are trainable. Meanwhile, it uses an additional L_2 regularization on both magnitude matrices: $|||\boldsymbol{M}^c| - |\boldsymbol{C}^c|||$ and $|||\boldsymbol{M}^e| - |\boldsymbol{C}^e|||$, so that the fraction of nonzero elements in $\boldsymbol{M}^c$ is equal to that of $\boldsymbol{C}^c$, and likewise for $\boldsymbol{M}^e$ and $\boldsymbol{C}^e$. The sparsity T^c_{ji} and T^e_{ji} are fixed to 1 and global scaling factors α^c and α^e are fixed to 0.01. M^c_{ji} and M^e_{ji} are restricted to be non-negative.

current-based + *connectome sparsity* For chemical and electrical connections, the model sparsity is $T^c_{ji} = \mathbf{1}[C^c_{ji} > 0]$ and $T^e_{ji} = \mathbf{1}[C^e_{ji} > 0]$ respectively. For both chemical and electrical connections, the magnitudes M^c_{ji} and M^e_{ji} are trainable. The sparsity T^c_{ji} and T^e_{ji} are fixed to the connectome sparsity and global scaling factors, α^c and α^e, are fixed to 0.01. M^e_{ji} are restricted to be non-negative, while the sign of M^c_{ji} is not restricted.

current-based + *learned dense* For chemical and electrical connections, the model is fully connected with $T^c_{ji} = 1$, and $T^e_{ji} = 1$. For both chemical and electrical connections, the magnitudes M^c_{ji} and M^e_{ji} are trainable. The sparsity T^c_{ji} and T^e_{ji} are fixed to 1 and global scaling factors α^c and α^e are fixed to 0.01. M^e_{ji} are restricted to be non-negative, while the sign of M^c_{ji} is not restricted.

current-based + *learned sparse* For chemical and electrical connections, the model is fully connected with sparsity $T^c_{ji} = 1$ and $T^e_{ji} = 1$. For both chemical and electrical connections, the magnitudes M^c_{ji} and M^e_{ji} are trainable. It uses an additional L_2 regularization on both magnitude matrices, $|||\boldsymbol{M}^c| - |\boldsymbol{C}^c|||$ and $|||\boldsymbol{M}^e| - |\boldsymbol{C}^e|||$, so that the fraction of nonzero elements in $\boldsymbol{M}^c$ is equal to that of $\boldsymbol{C}^c$, and likewise for $\boldsymbol{M}^e$ and $\boldsymbol{C}^e$. The sparsity T^c_{ji} and T^e_{ji} are fixed to 1, global scaling factors, α^c and α^e, are fixed to 0.01. M^e_{ji} are restricted to be non-negative, while the sign of M^c_{ji} is not restricted.

A.6 Model weight consistency analysis

We used the *correlation coefficient* to quantify the inter-consistency of models trained with different random initializations, as well as the connectome-consistency, which is the correlation coefficient between model weights and connectome counts.

In particular, for the connectome consistency, we compare the consistency of our model weights and connectome synapse count. We calculate the correlation coefficient r_i between the absolute value of model weights W_{ij} and connectome synapse count C_{ij} along each row i. Then it is averaged across N rows and 4 models with identical constraints but different random initializations to get $\bar{r}_i$ as our metric for connectome-consistency,

$$r_i = \frac{\sum_{j=1}^{N} \left(W_{ij} - \bar{W_{ij}}\right)\left(C_{ij} - \bar{C_{ij}}\right)}{\sqrt{\sum_{j=1}^{N} \left(W_{ij} - \bar{W_{ij}}\right)^2 \sum_{j=1}^{N} \left(C_{ij} - \bar{C_{ij}}\right)^2}}. \tag{27}$$

model		current-based			conductance-based				
consistency	**weight**	**learned dense**	**learned sparse**	**connectome sparsity**	**learned dense**	**learned sparse**	**connectome sparsity**	**connectome count**[1]	**connectome count**[2]
inter	$\boldsymbol{W}^c$	0.142	0.084	0.497	0.082	0.084	**1.000**	**1.000**	**1.000**
	$\boldsymbol{W}^e$	0.024	0.000	0.748	0.097	0.000	0.997	0.997	**1.000**
connectome	$\boldsymbol{W}^c$	-0.002	-0.001	0.469	-0.001	0.001	**1.000**	**1.000**	**1.000**
	$\boldsymbol{W}^e$	-0.002	0.000	0.654	0.006	0.000	0.998	0.996	**1.000**

Table 3: **Inter-consistency and connectome-consistency**. Correlation between multiple model weights with different initializations, and correlation between model weights and connectome counts. Evaluations are performed for models fitted on a single worm and across current-based and conductance-based models with different levels of connectome constraints. The CC-LVM is more consistent among different random initializations and more highly correlated to the connectome counts.

For inter-consistency, we evaluate how model weights are affected by random initialization after training. We calculate the correlation coefficient r_i between the weights W^1_{ij} of model 1 and weights W^2_{ij} of model 2 along each row i. Model 1 and model 2 have the same constraints but different random initializations. The correlation coefficient between them is

$$r_i = \frac{\sum_{j=1}^{N} \left(W^1_{ij} - \bar{W^1_{ij}}\right)\left(W^2_{ij} - \bar{W^2_{ij}}\right)}{\sqrt{\sum_{j=1}^{N} \left(W^1_{ij} - \bar{W^1_{ij}}\right)^2 \sum_{j=1}^{N} \left(W^2_{ij} - \bar{W^2_{ij}}\right)^2}}. \tag{28}$$

This coefficient, r_i, is then averaged across N rows and each combination of 2 models within the 4 models with random initialization to get $\bar{r_i}$ as our metric for inter-consistency.

As shown in Table 3, the conductance-based model parameters are more consistent under different random initializations. Optimization of the conductance-based model is easier because the chemical weights $\boldsymbol{W}^c$ of the conductance-based synapses are always positive and we do not need a trainable sign variable. Meanwhile, the model parameters are better constrained with the connectome sparsity or counts, which enables it to be consistent across different initialization.

A.7 Predicted traces in neuron holdout evaluation

For each neuron, we performed the neuron holdout evaluation for both current-based and conductance-based models with different levels of connectomic constraint. Here we plot additional predicted traces from 15 representative neurons, including 12 with good performance (top), and 3 with worse performance (bottom) in Figure 4.

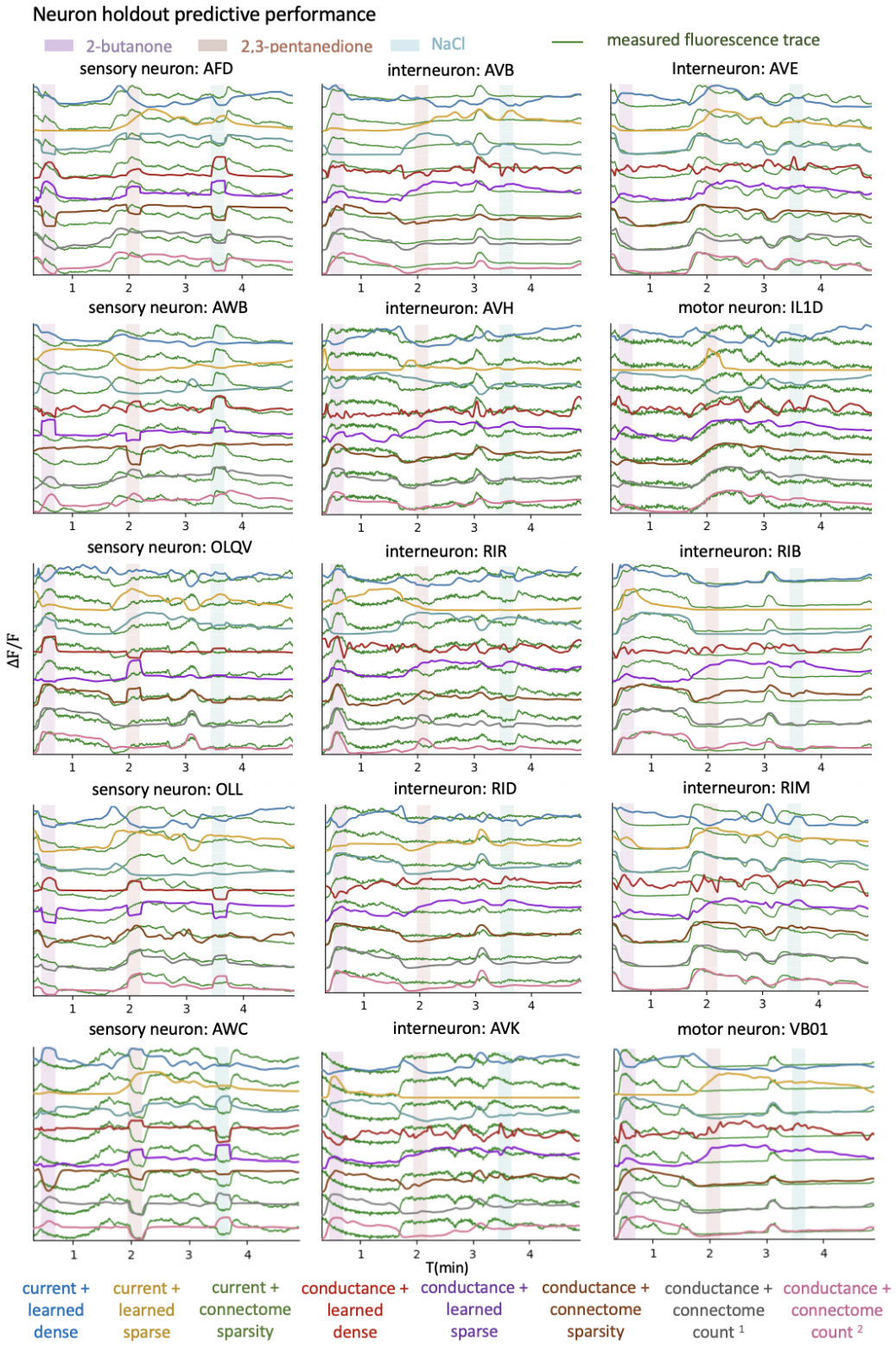

Figure 4: **LVM predictive performance across neuron types.** Measured traces and predictions made by the 8 models for 15 neurons including sensory neurons, interneurons and motor neurons.