From Connectome to Computation: Predicting Neural Function with Machine Learning

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Janne K. Lappalainen

aus Buchholz in der Nordheide

Tübingen

2025

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation: 16.01.2026
Dekan: Prof. Dr. Thilo Stehle
1. Berichterstatter: Prof. Dr. Jakob Macke
2. Berichterstatterin: Prof. Dr. Julijana Gjorgjieva
3. Berichterstatter: Prof. Dr. H. Sebastian Seung

To the people I love.

ABSTRACT

Connectomics maps synapse-level wiring diagrams with electron microscopy, creating detailed 3D reconstructions of information processing neural networks in the brain. However, such connectome measurements cannot directly reveal brain function — electrical dynamics of single neurons and synapses — making their utility controversial for understanding how the brain computes.

While computational neuroscience has traditionally focused on either mechanistically detailed models of small circuits or models that learn patterns from neural recordings without simulating the underlying biology, recent deep learning approaches have taken a different angle — task-optimizing artificial neural networks and establishing that artificial neural network activity matches patterns in brain measurements. However, a main gap remains: models that integrate single-neuron mechanistic detail with end-to-end task computation and accurately match brain activity.

We built a differentiable neural network model from the fly visual system connectome to test whether connectome-constrained networks can accurately predict neural electrical activity. Each neuron in the model corresponds to a neuron of the motion detection pathways in the real fly. We parametrized unknown neuron and synapse properties and trained them on a motion detection task from computer vision using backpropagation through time.

Despite equal connectome constraints and task, equivalently trained models generated variable predictions for responses of corresponding neurons. We therefore trained a model ensemble to characterize predictions statistically. Our ensemble accurately predicts neural tuning properties across 26 studies spanning decades of experimental research. Control experiments showed that models with full connectomic constraints *and* task-optimization predict neural function best.

Using dimensionality reduction and clustering of model responses to naturalistic stimuli, we found model predictions for the same neuron type grouped into competing hypotheses across the ensemble, such as upward versus downward motion selectivity. We found that hypotheses from the best task-performing clusters typically matched experimental data, and that clustering combined with a few measurements of neural activity sufficed to effectively identify correct function predictions.

Our work demonstrates how connectomes can be useful for predicting neural circuit function when combined with task constraints and machine learning. Our approach — parametrizing unknowns, optimizing via tasks, and distilling predictions from a model ensemble — can be broadly applied to end-to-end connectome-constrained modeling. We published pretrained models and software as research tools.

In ongoing work, we develop new models and insights based on, for instance, new connectome data, more realistic synapse models, neural activity measurements, and systematic evaluations of different task optimizations.

ZUSAMMENFASSUNG

Die Konnektomik kartiert Schaltpläne auf Synapsenebene mittels Elektronenmikroskopie und erstellt detaillierte 3D-Rekonstruktionen neuronaler Netze im Gehirn. Solche Konnektom-Messungen geben jedoch keinen direkten Aufschluss über die Gehirnfunktion – die elektrische Dynamik von Neuronen und Synapsen – sodass ihr Nutzen für das Verständnis der Rechenvorgänge im Gehirn umstritten ist.

Während sich die computergestützte Neurowissenschaft traditionell entweder auf mechanistisch detaillierte Modelle kleiner Schaltkreise oder auf Modelle konzentriert hat, die Muster aus neuronalen Aufzeichnungen lernen ohne die zugrundeliegende Biologie zu simulieren, haben aktuelle Deep-Learning-Ansätze einen anderen Weg eingeschlagen —- sie optimieren künstliche neuronale Netze für anspruchsvolle Aufgaben und zeigen, dass künstliche neuronale Netze ähnliche Aktivitätsmuster wie Gehirnmessungen aufweisen. Eine Hauptlücke bleibt jedoch bestehen: Modelle, die mechanistische Details einzelner Neuronen mit anspruchsvoller Aufgabenberechnung ganzheitlich integrieren.

Wir haben ein differenzierbares neuronales Netzwerkmodell aus dem Konnektom des visuellen Systems der Fliege erstellt, um zu testen, ob konnektombasierte Netzwerke neuronale elektrische Aktivität vorhersagen können. Jedes Modellneuron entspricht einem Neuron der Bewegungserkennungsschaltkreise der echten Fliege. Wir haben unbekannte Neuronen- und Synapseneigenschaften parametrisiert und sie mit einer Bewegungserkennungsaufgabe aus dem Computersehen mittels Fehlerrückführung durch die Zeit trainiert.

Trotz identischer Konnektivitätsbedingungen und Aufgabe haben gleich trainierte Modelle unterschiedliche Vorhersagen für entsprechende Neurone erzeugt. Daher haben wir ein Modellensemble zur statistischen Charakterisierung der Vorhersagen trainiert. Unser Ensemble liefert genaue Vorhersagen neuronaler Eigenschaften über 26 Studien aus Jahrzehnten experimenteller Forschung hinweg. Kontrollexperimente haben gezeigt, dass Modelle mit vollständigen konnektomischen Einschränkungen *und* Aufgabenoptimierung die neuronale Funktion am besten vorhersagen.

Mittels Dimensionsreduktion und Clustering der Modellantworten auf naturalistische Reize haben wir über das Ensemble hinweg konkurrierende Hypothesen für die Funktion desselben Neuronentyps gefunden, wie Aufwärts- versus Abwärtsbewegungsselektivität. Hypothesen aus den aufgabenbesten Clustern haben typischerweise mit experimentellen Daten übereingestimmt, und Clustering kombiniert mit wenigen neuronalen Messungen reicht aus, um korrekte Funktionsvorhersagen zu identifizieren.

Unsere Arbeit zeigt, wie Konnektome mit Aufgabenoptimierung zur Vorhersage neuronaler Schaltkreisfunktionen genutzt werden können. Unser Ansatz — Parametrisierung von Unbekannten, Aufgabenoptimierung und Destillierung von Vorhersagen im Modellensemble —- ist breit auf ganzheitliche, konnektombasierte Modellierung anwendbar. Wir haben vortrainierte Modelle und Software als Forschungswerkzeuge veröffentlicht.

In laufenden Arbeiten entwickeln wir neue Modelle und Erkenntnisse basierend auf beispielsweise neuen Konnektomdaten, realistischeren Synapsenmodellen, neuronalen Aktivitätsmessungen und systematischen Evaluierungen verschiedener Aufgabenoptimierungen.

ACKNOWLEDGEMENTS

It has been an honor and a scientific challenge. Surely a rollercoaster at times. What mattered to me was to make the most of it, which I think together we did.

Jakob, thank you for encouraging me when I felt sometimes beaten up by my unclear intermediate results or long stretches of intense work periods with, for me, no visible end in sight. At times you instilled trust in me and helped me navigate chaotic parts of the work with your experience. I appreciate the clarity and simplicity that you sought to provide when giving feedback on complex problems but also the slack you gave me to try things my way.

Srini, thank you for the years of working together. You have held me accountable very well and I am happy about how we developed our ability to work together over time towards being quite a routined team. The times when we routinely crunched analyses together, week-by-week, feel retrospectively most rewarding. I learned to appreciate the unwavering skepticism that you put forward when looking in close detail at my data. Thanks also for always welcoming me to visit your lab, I could always take away inspirations and specific ideas from those visits!

Thank you both for your supervision, mentoring, and support!

I thank all my collaborators spread across Germany, the US, and Canada for working together on parts of the work I present, and for working together on new things not addressed in here. Moreover, I want to thank my thesis examiners, Julijana, Sebastian, and Jakob.

I thank all colleagues at the University of Tübingen and at Janelia Research Campus for being friendly, motivated, and constructive in our interactions, creating a welcoming community to work with. Specifically, I want to thank my colleagues in the Mackelab and Turagalab for being fun, kind, and supportive crowds. Franzi, Shelby, Emilia for being highly professional administrative assistants, representing a huge support for this academic work! Also thank you Franzi, Auguste, Richard, Michael, and Poornima, for going connected through the Covid pandemic with many joyful moments. For this thesis, I thank Auguste for co-accountability meetings for thesis writing, and Richard, Linda, and Zina for proof-reading my thesis chapters.

I want to thank my friends and family, whom I love. Thank you for supporting me, in particular when I sometimes lost myself a bit in this work. You have helped me sustain my resilience and balance throughout this time. I want to thank Yağmur specifically for co-accountability meetings for thesis writing, proof-reading, and loving care during intense stretches of writing!

Looking forward to the years we have ahead of us!

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AI	artificial intelligence
Am	amacrine neuron type
ANN	artificial neural network
API	application programming interface
BERT	Bidirectional Encoder Representations from Transformers
BPTT	backpropagation through time
C2-C3	centrifugal neuron types
CDE	controlled differential equation
CNN	convolutional neural network
Cm	central medulla interneurons
CTRNN	continuous-time recurrent neural network
CT1	complex tangential neuron type
Dm	distal medulla interneuron types
DMN	deep mechanistic network (also: differentiable mechanistic network)
DNN	deep neural network
DSI	direction selectivity index
FIB-SEM	focused ion beam scanning electron microscopy
FRI	flash response index
GPU	graphics processing unit
kPCA	kernel principal component analysis
L1-L5	lamina monopolar neuron types
Lawf	lamina wide-field neuron type
LC	lobula columnar neuron type
Lo1	lobula layer 1
LoVC	lobula ventral centrifugal neuron type
M10	medulla layer 10

MAP	maximum a posteriori
Mi	medulla intrinsic neuron type
MLP	multilayer perceptron
ND	null direction
ODE	ordinary differential equation
OSI	orientation selectivity index
PD	preferred direction
Pm	proximal medulla interneuron types
R1-R8	photoreceptor neuron types
ReLU	rectified linear unit
RNN	recurrent neural network
T4	motion-detector neuron type encoding light-increment edges
T5	motion-detector neuron type encoding light-decrement edges
TEM	transmission electron microscopy
Tm	transmedullary neuron type
TmY	transmedullary neuron types with Y-shape
t-SNE	t-distributed stochastic neighbor embedding
UMAP	uniform manifold approximation and projection
VNC	ventral nerve cord

LIST OF SYMBOLS

Q	membrane charge
C (C_i)	membrane capacitance (of neuron i)
V ($V_i(t)$)	membrane voltage (of neuron i at time t)
V_{rest} ($V^{rest}_{t_i}$)	resting potential (for neuron type t_i)
I ($I_i(t)$)	external current input (to neuron i at time t)
I_C	capacitor current
I_{Ion}	ionic current
g_{Ion} (g_{ij})	conductance (ion channel or synaptic)
E_{Ion} (E_{ij}, E_j)	reversal potential
R (R_i)	leak resistance
τ_i (τ_{t_i})	membrane time constant (for neuron i or type t_i)
γ_{ij}	relative conductance (unitless), $\gamma_{ij} = g_{ij}R_i$
$e_i(t)$	external input potential, $e_i = I_iR_i$
w_{ij}	synaptic weight from neuron j to neuron i
$\mathbf{W} = [w_{ij}]$	weight matrix
$\sigma_{t_it_j}$	synapse sign (-1 inhibitory, $+1$ excitatory) between neuron types t_i, t_j
$\alpha_{t_it_j}$	synapse scaling factor between neuron types t_i, t_j
$N_{t_it_j,\Delta u,\Delta v}$	synapse counts between neuron types at retinotopic offset $(\Delta u, \Delta v)$
(u_i, v_i)	retinotopic coordinates of neuron i
$(\Delta u, \Delta v)$	retinotopic coordinate offset between pre- and postsynaptic neurons
Δt	discretization time step
$f(\cdot)$	nonlinearity (activation or synapse transfer function)
$h_i(t)$ ($\mathbf{h}(t)$)	RNN hidden state (single neuron or vector form)
$\mathbf{x}$, $\mathbf{y}$	model input and target output
ϕ_θ	neural network (function) parametrized by θ

θ	model parameters
η	learning rate
L (L_{task}, L_{neural})	loss (task or neural activity term)
$\mathcal{D}$	dataset
S	number of samples in dataset
M	number of models in an ensemble
K	number of mixture components or functional clusters
$p(\theta \mid \mathcal{D}, W)$	(approx.) posterior over parameters given data and constraints W
$u_i^{(m)}$	2D UMAP embedding of responses of neuron i in model m
π_k	mixture weight of cluster orcomponent k
$\mathcal{N}(u \mid \mu_k, \Sigma_k)$	Gaussian component with mean μ_k and covariance Σ_k
ϵ	filtering threshold (e.g., for model selection under new data)

1

INTRO AND OUTLINE

Philosophers have debated the relationship between body and mind for several millennia *(Lokhorst, 2005)*. René Descartes proposed that a small structure near the center of the brain — the pineal gland — serves as an interface between the physical body and an immaterial soul, a view today known as *mind-body dualism* *(Lokhorst, 2005; Descartes, 1948)*. Fast-forward to the present: neuroscience largely rejects dualism and holds that the mind could be a material property of the brain's neurophysiology *(Churchland, 1986)*, an *emergent property* of large-scale network dynamics *(Bunge, 1977; Sperry, 1980)*, or a fundamental physical property of certain systems *(Tononi, 2012)*. These philosophical controversies of how the mind might be implemented converge to a central question in empirical neuroscience research: How does the brain give rise to behavior (Fig. 1.1a)?

To answer this central question, neuroscientists have studied the detailed structure and function of the brain for over a century. Most fundamentally, anatomists discovered that the brain is made of diverse nerve cells called *neurons* (Fig. 1.1b), famously described by Cajal *(Cajal & Sánchez, 1915)*. Neurophysiologists measured the electrical signals traveling through these cells *(von Helmholtz, 1850; Bernstein, 1902)*. Others identified *synapses* — specialized nanoscopic cell structures which exchange chemicals, called *neurotransmitters* — to transmit electrical signals between neurons *(Sherrington, 2023; Loewi, 1945; Kandel et al., 2000)*. Neurons connected via synapses form *neural networks* (Fig. 1.1c).

What mechanisms allow neural networks to orchestrate behavior? Neural networks integrate information from different sensory modalities and previous experience, and route decisions to motor systems or other subnetworks for further processing and refinement. The laws of biophysics, the properties of individual neurons and synapses — such as their morphology and molecular compositions, and their connectivity combine to a vast diversity of specializations in how electrical and chemical signals travel within neurons and neural networks to compute. However, subnetworks are usually not clearly separable, and sometimes even an individual neuron can be central for core behavior of an animal *(Ache et al., 2019)*. In general, the computation underlying behavior using sensory signals is complex. Processing depends not only on neuron and network properties, but also on influences from neurotransmitters, glia cells, and gap junctions *(Bargmann & Marder, 2013; Scheffer & Meinertzhagen, 2021)*. Also the immediate internal states such as hunger or fear *(Lin et al., 2019)*, and stimulus properties, such as light *(Currier et al., 2023)* or sound properties *(Clemens et al., 2018)*, influence behavior. While human and animal nervous systems differ in number of neurons and synaptic connectivity, they share low-level mechanisms *(Borst & Helmstaedter, 2015; Tanaka & Portugues, 2025)*, molecular mechanisms *(Yamamoto et al., 2014)* and high-level computational tasks *(Basu & Nagel, 2024)* for survival, such as locating food, seeking mates, reproducing, and escaping threats (Fig. 1.1d). Understanding how neural networks — from few to billions of neurons — coordinate their activity to produce coherent behavior despite this complexity remains one of neuroscience's fundamental challenges.

To address this challenge, researchers have developed various computational approaches. Computational neuroscience uses models that are classically often theoretical *(Hopfield, 1982)*, small but detailed *(Marder & Bucher, 2007)*, or empirically fit to patterns of neural activity and behavior *(Paninski & Cunningham, 2018)*. They often enable mechanistic hypothesis generation, but cannot compute higher-level tasks. In contrast, artificial neural networks (ANN, also *deep neural networks* or DNN) from computer science, prioritize empirical task-performance as well as architectural scalability over mechanistic interpretability. Surprisingly, task-optimized *convolutional neural networks* (CNN), a class of DNNs, trained to classify visual objects, turned out to be a more powerful model for predicting high-level neural representations of the mammalian visual systems than models previously conceived with detailed anatomical and physiological correspondences *(Yamins et al., 2014; Yamins & DiCarlo, 2016; Kriegeskorte, 2015)*.

Deep learning[1] presents tools to build neuroscience models more generally. While task-optimized deep neural networks are effective models for predicting high-level neural representations also in a variety of neural systems other than the visual system (Section 2.2.6), deep learning models also provide tools for tracking organism behavior *(Robie et al., 2017; Mathis et al., 2018)* and inferring latent variables in computation from data for explaining neural computation and behavior *(Sussillo et al., 2016; Schneider et al., 2023)*. Despite these modeling advances *(Mathis et al., 2024)* — and the ongoing efforts to understand DNNs mechanistically for interpretable AI *(Bereska & Gavves, 2024; Lindsey et al., 2025)* — biological organisms remain unique in their ability to survive and adapt in dynamic, multimodal environments with mini-

---

[1] A subfield of machine learning that uses neural networks with multiple layers to learn complex patterns from data.

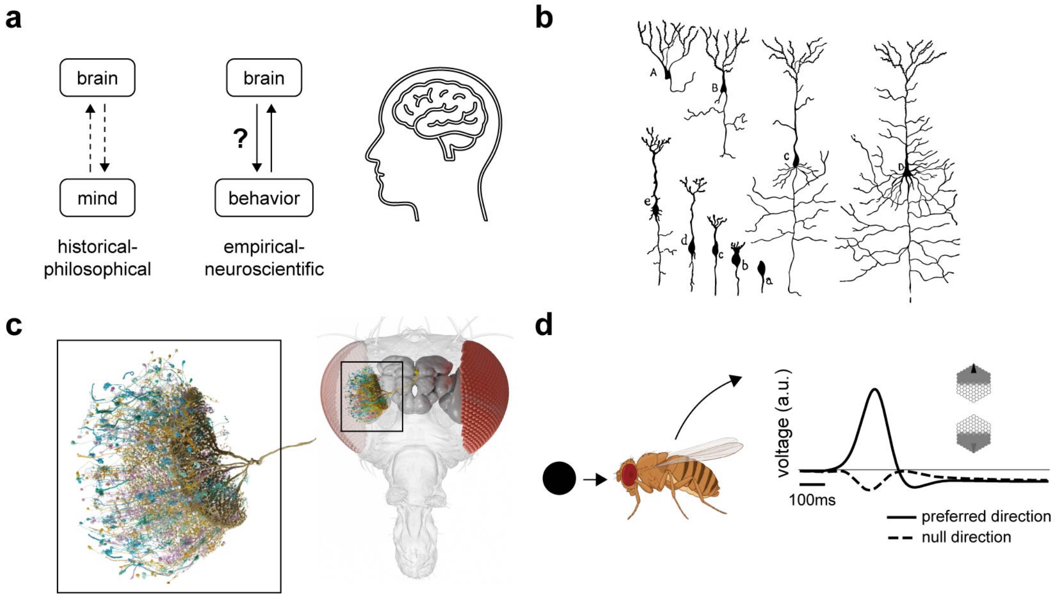


Figure 1.1: **From the brain to behavior. (a)** Perspectives on the relationship between mind, brain, and behavior. The uncertain mind–brain connection (left) reflects centuries of philosophical debate, while the empirically established brain–behavior link (right) forms the foundation of modern neuroscience research. **(b)** The brain is made up of neurons (and other cells) specialized for electrical signal transmission. Neuron morphologies vary, many have tree-like structures — dendrites and axons — to send and receive electrical signals. Original caption: "A-D, nerve-cells or neurons from four Vertebrate animals: A, a fish, B, a reptile, C, a bird, D, a mammal. [...] a-e show five stages in the individual development of the neurons of a bird. [...]". Adapted from *(Thomson, 1925)*. **(c)** Example of a biological neural network. Here, connectomic reconstruction of the fly visual system *(Nern et al., 2025)*. **(d)** Electrical signaling in networks of neurons supports computation underlying behavior. Here, the electrical response of motion sensitive neurons (right) supports detecting an approaching threat represented by the black circle (left) informing the fly to escape *(Card & Dickinson, 2008)*. Motion sensitive neurons respond selectively to the motion direction, with large depolarization to preferred direction motion (here up) and no depolarization to null direction motion (here down).

mal energy and often remarkably few neurons. Discovering *how* with models offers access to unique scientific challenges and opportunities. These results suggest that task-optimized DNNs represent plausible models of neural computation underlying behavior. However, the field lacks models that achieve single-neuron correspondence with real networks while being mechanistically realistic and solving high-level tasks. Can we build accurate models of neural and behavioral computation of organisms that integrate neural network structures, single-neuron mechanisms, and high-level tasks (Fig. 1.2a)?

The key structural property one might wish to reproduce in a precise neural network model of a brain is the *connectome* because it determines much of the brain's dynamic electrical signaling (Section 2.3.2). The connectome represents the measured neurons and synaptic connectivity of the brain (Section 2.3.2). Today, many connectomic measurements of small to large-scale and whole-nervous systems of animals exist *(White et al., 1986; Jarrell et al., 2012; Cook et al., 2019; Witvliet et al., 2021; Zheng et al., 2018; Scheffer et al., 2020; Dorkenwald et al., 2024; Hulse et al., 2021; Wanner et al., 2016; Hildebrand et al., 2017; Svara et al., 2022; Bock et al., 2011; Helmstaedter et al., 2013; Oh et al., 2014; Motta et al., 2019; Loomba et al., 2022; Kasthuri et al., 2015; Shapson-Coe et al., 2024)*.

This leads us to the main goal addressed in this thesis: Can we use a connectome (Fig. 1.1c) to build a computational model that captures dynamic computations of real neurons across a whole neural network with a task-training approach (Fig. 1.2b)?

Theoretically, it is unclear how well a connectome can constrain network computation. The connectome has no direct information of biophysical parameters of neurons and synapses. Moreover, we know that different network parameters can compute similar functions with the same network connectivity *(Prinz et al., 2004)*, but that the same network can approximate different functions too *(Hornik et al., 1989)*. The whole-nervous system of *C.elegans*, with 302 neurons, has been mapped several decades ago *(White et al., 1986)* (Section 2.3), but scientists struggled to replicate animal behavior in a model because crucial molecular, biophysical, and contextual information was missing. This fueled a debate between scientists around whether mapping connectomes was worth the cost *(Jabr, 2012)*.

In this work, we address the question: What can we learn from a connectome by combining dynamical system modeling (Section 2.1 and 2.2.4), with a connectome measurement (Section 2.3.2), and a specified high-level task with tools from deep learning (Section 2.2). The interdisciplinary ap-

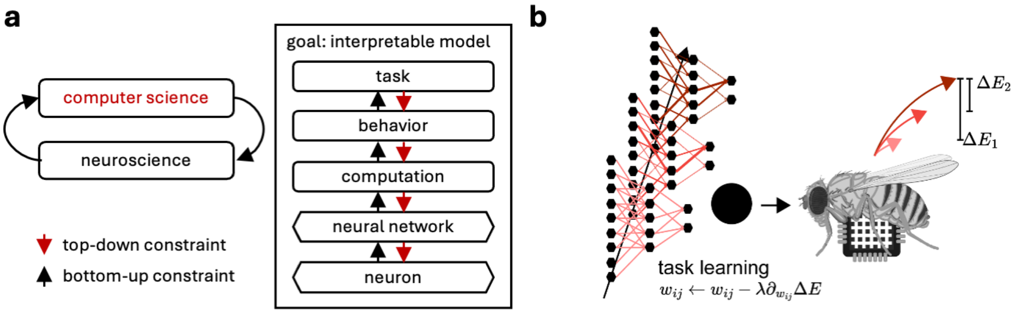


FIGURE 1.2: **Synergy of task-optimized neural networks and neuroscience-constrained modeling.** **(a)** Neuroscience and computer science provide complementary constraints for building interpretable models: Top-down constraints (red arrows) from computer science guide parameter optimization for task performance, while bottom-up constraints (black arrows) from neuroscience ground models in biological plausibility, from neurons up to behavior. **(b)** Example task-training approach to capture behavioral responses to a stimulus. An artificial neural network (left) computes a behavioral escape response of a virtual fly to a looming visual stimulus (right). The initial network weights w_{ij} produce an incorrect response (light red arrow). Task-training adjusts the weights via gradient-based updates ($w_{ij} \leftarrow w_{ij} - \lambda\partial_{w_{ij}}\Delta E$), reducing the task error ΔE over successive trials. After training, the network produces the correct escape behavior, mimicking the biological system.

proach results in a model that we call *deep mechanistic network* (DMN, Section 2.4). Specifically, we built a differentiable neural network model constrained with the connectome of the motion-detection pathway of the fruit fly *(Rivera-Alba et al., 2011; Takemura et al., 2015; 2017; Shinomiya et al., 2019; 2022)* as a continuous-time convolutional recurrent neural network on a graph. We parametrized unknown neuron and synapse properties of the model and then trained the neural network to compute motion in dynamic visual scenes using the deep learning dataset Sintel *(Butler et al., 2012)* using backpropagation through time. Afterwards, we compared dynamic computations predicted by the model to experimental measurements published over the last two decades (Section 2.3.3).

We first found that, despite equal constraints from connectome and task, equivalently trained models produced different predictions for responses of corresponding neurons. Therefore, we trained an ensemble of models to characterize predictions statistically. We found that the model ensemble accurately captures single-neuron activity across the fly visual system, recapitulating experimental measurements from 26 studies. The model predicts neural responses of motion detecting T4 and T5 neurons, critical to animal behavior. Models combining full connectomic constraints and task-optimization agree with known neural function best.

We used dimensionality reduction and clustering of model responses to naturalistic stimuli to explain the variability across the ensemble. We found that individual model predictions from the ensemble for the same neuron type typically cluster into competing hypotheses. For instance, some models predict downwards motion selectivity while others predict upwards motion selectivity for the same neuron. Ranking models and clusters by task-performance and evaluating the best task-performing models, or combining clustering with knowledge of a few measurements of neural activity, typically resulted in correct function predictions in spite of model degeneracies.

Our results demonstrate that a connectome combined with task-training is useful for model-driven prediction of detailed neural computation. Our DMN approach can be broadly applied to use connectome measurements for end-to-end modeling of neural computation and to support addressing the significant remaining challenges for modeling the entire fly comprehensively *(Scheffer & Meinertzhagen, 2021)*. We published our results, model, and software as ressources for the community *(Lappalainen et al., 2024)*.

THESIS OUTLINE

In this thesis, I address the question of whether we can build accurate models of neural computation that integrate neural network structures, single-neuron mechanisms, and high-level tasks using deep learning. I present the modeling approach we developed, the deep mechanistic network, and demonstrate its application to the fly visual system connectome, showing that our model can predict single-neuron dynamics from connectomic reconstructions and a motion-detection task.

This thesis is structured as follows:

- **Chapter 2** provides computational and biological background.

  In the first half (Section 2.1 and 2.2), I establish the computational neuroscience background supporting our work and the mathematical relationship of our simplified neuron model for connectome simulation to a detailed Hodgkin-Huxley-type neuron model for a grounding in established mechanistic biophysical models. Then I introduce the roots of deep learning and basic architectures to establish that our neuron model type simultaneously corresponds to the continuous-time limit of recurrent neural networks — widely used in deep learning for sequence tasks — but that the network structure of the fly visual system does not resemble typical deep neural network architectures. I introduce backpropagation for task-training and review task-training approaches to neuroscience modeling. Last, I bring probabilistic deep learning into perspective for formally addressing the inverse problem of predicting neural parameters and individual neural function from a connectome and high-level task.

In the second half (Section 2.3 and 2.4), I motivate the fly as the right modeling organism for developing new modeling knowledge and tools. Then, I introduce the fly visual system connectome and known functional properties of neurons in this system. Eventually, I summarize what we knew before and after conducting the research described in this thesis, culminating in the methodological approach of the deep mechanistic network for end-to-end modeling and analysis of structure and function using connectomes and deep learning.

- **Chapter 3** presents our main collaborative results in form of the public and printed peer-reviewed journal article *(Lappalainen et al., 2024)*.
- **Chapter 4** explores partially ongoing follow-up work aimed at building better models.

  I project modeling insights onto requirements for building new connectome constrained models of the fly, point out major modeling challenges, and summarize goals from ongoing collaborative work for building more realistic models. Next, I explain heuristics and challenges we encountered for model optimization in the context of the main work, and recapitulate software and infrastructure requirements and insights, for emphasizing requirements to scale the modeling research effort.
- **Chapter 5** considers further questions for interpreting relationships of structure and function with a connectome-constrained model.

  I first address the question of the importance of connectome and high-level task constraint for predicting neural function accurately, summarizing our insights from the main work. I briefly recapitulate our results in context of the (subjectively intriguing) question of task-optimality of connectomic constraints. Next, I point out challenges for establishing model corresponding in face of experimental variability. Then, I analyse the ensemble variability on a higher level, demonstrating our tools to characterize the entire functional space across cells that a model ensemble predicts. Finally, I briefly sketch a probabilistic inference framework with DMNs aimed towards formalizing modeling and making experimental measurements in a closed-loop.

- **Chapter 6** concludes with a summary, discusses the broader implications and limitations of this approach in context of other research, and outlines future directions.

DMNs integrate mechanistic neuron models with task-training using tools from deep learning and connectome measurements. This work assesses whether a connectome can predict neural dynamics that are relevant to behavior and to offer a general strategy to create and interpret mechanistic models of neural systems with connectomes and machine learning. Our approach allows building predictive models for any organism with a mapped connectome with machine learning, potentially accelerating and detailing our understanding of how neural structure gives rise to computation and ultimately behavior through model-driven systems neuroscience research.

2

CONTEXT AND EXISTING THEORY

This chapter establishes the multidisciplinary background ranging from neuroscience to machine learning research that underpins the deep mechanistic network.

2.1 A COMPUTATIONAL MODEL FOR CONNECTOME SIMULATION

Computational neuroscience addresses the gap between neurophysiological and behavioral mechanisms quantitatively for an understanding that manifests in mathematical principles and results in hypothesis-generating theory for more than a century *(Dayan & Abbott, 2005; Gerstner et al., 2014)*. Today's mathematical formalisms trace back through several important multidisciplinary insights. To name a few: *Bernstein (1902)* proposed the *membrane hypothesis*, explaining how ionic gradients across the neuron membrane contribute to electrical neural signaling (Fig. 2.1a). Louis Lapicque (1907) introduced the leaky integrate-and-fire neuron model, a threshold-based, spike generating mathematical model of electrical signaling in neurons *(Abbott, 1999)*. In the 1930s, Kenneth Cole and Howard Curtis demonstrated they could measure the conductance changes during action potentials in the squid's giant axon with electrical feedback control circuits *(Curtis & Cole, 1938)*. These insights culminated in the Hodgkin-Huxley neuron model (Fig. 2.1b), a system of differential equations that predicted the precise biophysical mechanisms generating action potentials in the giant axon of a squid *(Hodgkin & Huxley, 1952)*. Models like Hodgkin-Huxley and leaky integrate-and-fire neurons form the backbone of theoretical neuroscience: The insights and theory underpins not only specialized mechanistic circuit modeling based on first principles to interpret scientific discovery *(Marder & Bucher, 2007; Turner et al., 2008; Gruntman et al., 2018; Groschner et al., 2022; Romani & Tsodyks, 2015)*, but also landmark initiatives that pursue large-scale simulations of mechanistic neural networks *(Markram et al., 2015; Billeh et al., 2020)*. Mechanistic circuit models of neural circuits, such as the stomatogastric ganglion *(Marder & Bucher, 2007)* or the Drosophila motion detection circuit *(Gruntman et al., 2018; Groschner et al., 2022)*, demonstrate how models can match detailed biological computation in highly interpretable systems.

2.1.1 *From neuron biophysics to conductance-based neurons*

I briefly explain the biophysics of a neuron to derive the computational neuron model of the main work (Section 3.5.3) from a detailed Hodgkin-Huxley-type, conductance-based neuron model. I explain the reductionist approach leading to the precise system of differential equations — one of several applicable choices — describing our connectome-constrained network model. In the main work's methods section I did not use the same detail as in this derivation, but I find it worthwhile to attempt closing the derivation gaps in context of existing work, because this knowledge remains fundamental for the naive reader and in follow-up research.

Within the body, electrical signals travel between physical compartments in the form of charged elements that exist in solution, like sodium (Na^+), potassium (K^+), or chloride (Cl^-). Fundamental electromagnetic forces cause attraction between charges of opposite polarity and repulsion between particles of the same polarity, following Coulomb's law *(Coulomb, 1785)*. More generally, charged particles move along gradients of the *electric potential* *(Griffiths, 2023)*. Moving charges create *currents*.

The neuron is a cell that specializes in electrical signal messaging through currents. The neuron's membrane controls electrical currents as they pass as signals through the cell (Fig. 2.1a). The membrane maintains charge *im*balances between the inside and the outside of a neuron, and thus an electric potential called *membrane potential*, by using energy to pump positively charged ions into the extracellular solution *(Kandel et al., 2000)*. Simultaneously, the membrane implements ion channels that allow selective influx of charged particles along the gradient of the membrane potential.

The neural membrane acts similar to a capacitor: It stores charges $Q = CV$ across its lipid-bilayer and maintains a membrane potential or *membrane voltage*, V. Currents I_C flowing through the membrane (dis-)charge the capacitor:

$$I_C = \dot{Q} = C\dot{V}.$$

Note, that I will mostly use the shorter dot-notation for time-derivatives $\dot{x} = \frac{dx}{dt}$. The membrane capacitance C measures how much charge the membrane can store per unit voltage.

The neural membrane has specific *ion channels*. Specific channel types conduct specific ion types, like sodium (Na^+), potassium (K^+), or chloride (Cl^-). The current through an ion channel, following ohmic behavior, is

$$I_{Ion} = g_{Ion}(V - V_{Ion}).$$

The channel conductance $g_{Ion} = g_{Ion}(t, V)$, measures the variable amount of current that flows per unit voltage nonlinearly as a time-varying function of membrane voltage and other synaptic mechanisms. The channel-specific reversal potential V_{Ion}, determines at which voltage the ion current will be zero.

Small neurons can be idealized as points, called *point neurons*, because their membrane voltage is approximately the same across each patch of their membrane. Approximating the neuron as a point neuron, by Kirchhoff's law, the total current flowing in and out of the cell adds up to zero. An external current source I, a membrane leak current $I_{leak} = (V_{rest} - V)/R$, and ionic currents $\sum_{Ions} g_{Ion}(V - E_{Ion})$, charge the membrane potential as:

$$C\dot{V} = \frac{V_{rest} - V}{R} + \sum_{Ions} g_{Ion}(E_{Ion} - V) + I$$

We generalize to N synaptic inputs with their own time-varying, nonlinear conductances $g_1, ..., g_N$, and reversal potentials $E_1, ..., E_N$:

$$C\dot{V} = \frac{V_{rest} - V}{R} + \sum_{j} g_j(E_j - V) + I$$

An equivalent electrical circuit with basic electronic elements represents the dynamic behavior of this differential equations accurately and makes the biophysics of neural computation accessible through the lens of signal processing theory (Fig. 2.1b) *(Hodgkin & Huxley, 1952; Koch, 2004)*. The original Hodgkin-Huxley model represents the case of current flow from sodium and potassium ions, and describes channel conductances as time-varying, nonlinear functions of the membrane voltage with mathematical terms that represent molecular mechanisms of channel opening and closing *(Hodgkin & Huxley, 1952)*.

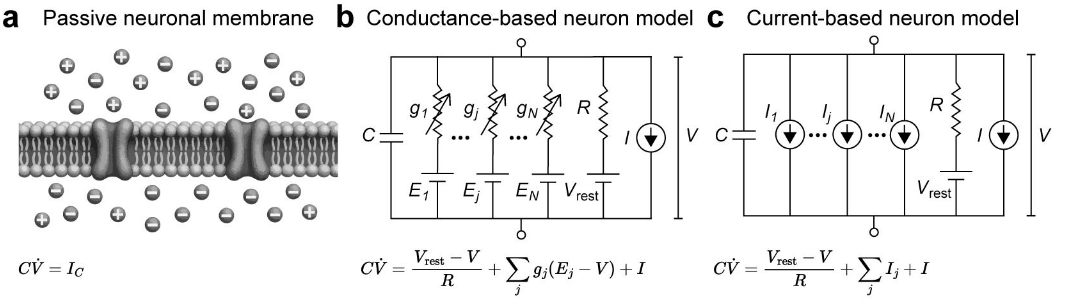


FIGURE 2.1: **Reduction of neural signal processing. (a)** A lipid-bilayer structure embedded with ion channels forms a passive neuronal membrane. Ionized elements, like sodium (Na^+), potassium (K^+), or chloride (Cl^-) exist in intracellular (bottom) and extracellular (top) solution around the membrane. The membrane regulates current flow through channels, which causes changes in membrane voltage. Lipid-bilayer generated by Dall-E (OpenAI), 2025 *(OpenAI, 2025)* and edited in illustration program. **(b)** A conductance-based neuron model (Hodgkin-Huxley type). The membrane voltage V, changes as a function of input currents I, leak currents $(V_{\text{rest}} - V)/R$, and time-varying, nonlinear synaptic currents, characterized by the synaptic conductances g_j, and synaptic reversal potentials E_j. **(c)** A current-based neuron model. The synapse inputs can be reduced to current sources $I_j = g_j E_j$, if the synaptic reversal potential is large compared to the change in membrane voltage *(Koch, 2004; Shriki et al., 2003; Miller & Fumarola, 2012)*.

2.1.2 *From conductance-based neurons to a continuous-time RNN*

In a neural network, the voltage of each neuron i can follow the same mathematical equation:

$$C_i \dot{V}_i = \frac{V_i^{\text{rest}} - V_i}{R_i} + \sum_j g_{ij}(E_{ij} - V_i) + I_i. \tag{2.1}$$

This model grants each neuron i individual capacitance values C_i, leak resistances R_i, interactions g_{ij} and E_{ij}, and external inputs I_i to flexibly account for neuron-to-neuron functional differences. Conductance variables g_{ij}, and reversal potentials E_{ij} are synaptic parameters, representing contributions from presynaptic activity of neuron j to postsynaptic voltage of neuron i. We will see later, that with connectome constraints a parameter distinction on neuron-type level, instead of neuron level, suffices to accurately model the fly visual system. This reduces the required number of individual neuron parameters to the number of neuron types, and the number of required individual synaptic parameters to the number of neuron-type-to-neuron-type connections (Section 3.2.1).

One can approximate the synapse inputs as current sources under different assumptions *(Koch, 2004; Shriki et al., 2003; Miller & Fumarola, 2012)*, for instance by assuming the synaptic reversal potential is large compared to the change in membrane voltage (Fig. 2.1c):

$$C_i \dot{V}_i = \frac{V_i^{\text{rest}} - V_i}{R_i} + \sum_j g_{ij} E_{ij} + I_i.$$

We multiply with leak resistance R_i to simplify the formalism: We substitute the resulting terms with $\tau_i = C_i R_i$ a membrane time constant; relative conductances $\gamma_{ij} = g_{ij} R_i$, a unitless measure of synaptic strength; and the voltage potential $e_i = I_i R_i$ created by external input currents:

$$\tau_i \dot{V}_i = V_i^{\text{rest}} - V_i + \sum_j \gamma_{ij} E_{ij} + e_i.$$

Finally, we reinterpret the synapse interaction. We substitute $\gamma_{ij} E_{ij} = w_{ij} f(V_j)$, with threshold nonlinearity f and synapse strength w_{ij}. This model describes nonlinearly varying voltage inputs at synapse ij as a function of presynaptic voltage $f(V_j)$, representing presynaptic neurotransmitter release. Here, $f(V_j)$ has units of neurotransmitter release rate (vesicles/s), so w_{ij} has units of V·s/vesicle to ensure the product results in voltage input. The synapse strength w_{ij} is positive for depolarizing current inputs and negative for hyperpolarizing current inputs:

$$\tau_i \dot{V}_i = V_i^{\text{rest}} - V_i + \sum_j w_{ij} f(V_j) + e_i.$$

In the next section, I will introduce neural networks formulations from computer science and seize the opportunity to approach the same bottom-up derived result (from single biophysically grounded neurons to neural networks) from the top-down perspective of computer science neural networks (from neural networks to single biophysically grounded neurons).

2.2 FROM NEUROSCIENCE TO DEEP LEARNING AND BACK

Deep learning combines neural network architectures with optimization algorithms to identify function approximators from data on well-defined computational tasks. The foundations of modern deep learning were laid out in the mid-twentieth century at the intersection of different research fields, including neuroscience, cognitive science, psychology, mathematics, numerical analysis, and electronics. *McCulloch and Pitts (1943)* developed a first mathematical framework for the logical operations underlying electrical neural activity. *Rosenblatt (1958)* developed the Perceptron, a neural network capable of linearly separable binary classification. Inspired by a single neuron, it possesses sensory units to receive inputs, association units to mix inputs, and response units to send outputs. In his later book, *Rosenblatt (1962)* detailed a comprehensive perceptron framework and forshadowed later advances in the field, including the *multilayer perceptron* (MLP)[1], an advancement that introduced *hidden layers*. Yet, Rosenblatt's networks were limited in their capability to learn parameters for practical applications. Algorithmic learning in neural networks became a new research focus *(Widrow & Hoff, 1960; Minsky & Papert, 1969)*[2]. The introduction of the modern gradient-based parameter-update algorithm, called *backpropagation*, a reverse-mode automatic differentiation algorithm *(Linnainmaa, 1970; Werbos, 1974)*, and its application to MLPs represents one of the most important inflection points in artificial intelligence (AI) research, and made neural networks with hidden layers broadly applicable as function approximators *(Rumelhart et al., 1986)*. Today's MLPs theoretically approximate any well-behaved and bounded function *(Hornik et al., 1989; Cybenko, 1989)*, and are indispensable building blocks of mainstream machine learning applications such as large language models *(Vaswani et al., 2017)*. Moreover, different neural network architectures can compute the same function *(Goodfellow et al., 2016)*, using different computational mechanisms *(Raghu et al., 2021)*.

---

[1] Today, MLPs are also called fully connected feedforward neural networks and use a large variety of nonlinear activation functions, not only the originally used Heaviside-step function.

[2] I will now use the term training instead of learning for a clean separability to learning in animals.

2.2.1 *The fly visual system is no feedforward MLP*

> *"The theory to be presented here is concerned with a class of "brain models" called perceptrons - By "brain model" we shall mean any theoretical system which attempts to explain the psychological functioning of a brain in terms of known laws of physics and mathematics, and known facts of neuroanatomy and physiology."* – *Rosenblatt (1962)*

To cover the foundations, I introduce a modern form of Rosenblatt's MLP architecture in a non-formal way. Because our mechanistic network operates at the level of single fly neurons, to warm up in preparation for it, we benefit from an explicit, and flexible per-neuron formulation of the MLP in addition to the matrix notation commonly used in machine learning.

It is enough to refer to a supervised learning problem here: Given a dataset of S input-output pairs $\{\mathbf{x}_s, \mathbf{y}_s\}_{s=1}^{S}$, we look for a neural network ϕ, and its parameters θ such that $\mathbf{y}_s \approx \phi_\theta(\mathbf{x}_s)\ \forall s$. For other machine learning problems neural networks can be applied to see for instance *Bishop and Nasrabadi (2006)* or *Goodfellow et al. (2016)*.

An MLP is a parametric function $\phi_\theta : \mathbb{R}^{n_0} \longrightarrow \mathbb{R}^{n_L}$ mapping inputs $x \in \mathbb{R}^{n_0}$ to outputs $\mathbf{y} = \phi_\theta(\mathbf{x}) = \mathbf{h}^{(L)} \in \mathbb{R}^{n_L}$, where L is the number of layers.

In matrix notation:

$$\begin{aligned} \mathbf{h}^{(0)} &= \mathbf{x} && \text{(sensory layer or input)} \\ \mathbf{h}^{(\ell)} &= \mathbf{W}^{(\ell)} f^{(\ell-1)}\big(\mathbf{h}^{(\ell-1)}\big) + \mathbf{b}^{(\ell)} && \text{(association or hidden layer)} \\ \mathbf{r}^{(\ell)} &= f^{(\ell)}\big(\mathbf{h}^{(\ell)}\big), \quad \ell = 1, \ldots, L. && \text{(activation or response)} \end{aligned}$$

In single-neuron notation:

$$\begin{aligned} h_i^{(0)} &= x_i && \text{(sensory layer or input)} \\ h_i^{(\ell)} &= \sum_j w_{ij}^{(\ell)} f^{(\ell-1)}(h_j^{(\ell-1)}) + b_i^{(\ell)} && \text{(association or hidden layer)} \\ r_i^{(\ell)} &= f^{(\ell)}\big(h_i^{(\ell)}\big), \quad \ell = 1, \ldots, L. && \text{(activation or response)} \end{aligned}$$

Neuron indices run from $i = 0$ to the total number of neurons in the network $N = \sum_{l=0}^{L} n_l$. The trainable parameters are $\theta = \{\mathbf{W}^{(\ell)}, \mathbf{b}^{(\ell)}\}_{\ell=1}^{L}$. $\mathbf{W}^{(\ell)} = [w_{ij}^{(\ell)}] \in \mathbb{R}^{n_\ell \times n_{\ell-1}}$ describes connectivity (usually fully connected), signs, and connection strengths, and abstracts biological synapses. $\mathbf{b}^{(\ell)} = [b_i^{(\ell)}] \in$

$\mathbb{R}^{n_\ell}$, usually referred to as bias, describes a constant offset. $f^{(\ell)} : \mathbb{R} \to \mathbb{R}$ applies an element-wise (usually nonlinear) activation function. The MLP architecture is *feed-forward* because signals pass from layer $\ell = 1$ to $\ell = L$ without returning connections. The original perceptron is the special case in which $L = 1$ and $f^{(1)}$ is the Heaviside-step function. Neural networks with more than one hidden-layer are called *deep neural networks*.

In single-neuron notation it is important to think of the network as represented by a graph, a table of nodes for neurons and a table of edges for synaptic connections between nodes, which in practice efficiently represents models with sparse brain connectivity. Later, this type of notation simplifies mathematical expression of the inclusion of granular neuron(-type) specific constraints and of the mechanistic interpretation of single-neuron computation (Section 2.4). One can establish the correspondence to the mathematical form of graph neural networks known from machine learning *(Bronstein et al., 2021)*.

Of course, an MLP is not a good model for the fly visual system because the fly visual system structure diverges from an MLP structure in several important properties even on a high-level, such as sparsity, lateral and recurrent connectivity (Fig. 2.2 a vs. c), morphological detail, and its crystalline architecture. Convolutional neural networks have a crystalline architecture and in the next section I cover convolutional neural networks.

2.2.2 *The fly visual system is no conventional CNN*

Fukushima and Miyake (1982) introduced the Neocognitron, a precursor of the modern *convolutional neural network*, that, inspired by the hierarchical and repetitive structure of the visual cortex found by *Hubel and Wiesel (1962)*, performed strongly in recognizing visual patterns. *LeCun et al. (1989)* popularized the modern form of the convolutional neural network, and trained it using backpropagation to do handwritten digit classification. *Krizhevsky et al. (2012)* extended this architecture to large-scale object classification across a huge image database (ImageNet), marking another turning point for deep learning in computer vision.

CNNs are specialized for processing images, high-dimensional inputs with local structure. Each neuron in a CNN processes a local patch of the previous layer (its *receptive field*) using a shared set of weights (called a *convolutional filter* or *kernel*). Equivalent to the MLP, a typical CNN separates into feedforward layers. Each CNN layer consists of multiple *channels*, where each channel applies a distinct filter over the input. Weight sharing introduces translation equivariance because the same computation is applied locally across spatial positions. CNN layers therefore tile identical feedforward subcircuits across the input domain for efficient, translation equivariant parallel computation.

We can adapt the single-neuron notation of the MLP to reflect the parameter-sharing feature of CNNs.

$$\begin{aligned} h_i^{(0)} &= x_i && \text{(input)} \\ h_i^{(\ell)} &= \sum_{j \in \mathcal{N}_{c_i}} w_{c_i c_j}^{(\ell)} f^{(\ell-1)}(h_j^{(\ell-1)}) + b_{c_i}^{(\ell)} && \text{(convolutional layer)} \\ r_i^{(\ell)} &= f^{(\ell)}(h_i^{(\ell)}), \quad \ell = 1, \dots, L. && \text{(activation)} \end{aligned}$$

We assign each neuron i to a channel c_i. $\mathcal{N}_{c_i}$ is the set of presynaptic neurons j belonging to any receptive field of neurons i of channel c_i. Connections from neurons of channel c_j to c_i share a parameter $w_{c_i c_j}^{(\ell)}$. Postsynaptic neurons i share biases b_{c_i}. The convolutional filter defined by input from presynaptic neurons from channel j $\mathcal{N}_{c_i, c_j}$, with $\mathcal{N}_{c_i, c_j} \subset \mathcal{N}_{c_i} \subset \mathcal{N}$ is a small subset of the total set of neurons, inducing sparsity in network connectivity. Such local filter banks and parameter sharing reduces the number of trainable weights in CNNs notably in comparison to the MLP.

An important finding at the intersection of neuroscience and AI is that CNNs trained on large image databases to classify objects capture representations of mammalian vision accurately *(Yamins et al., 2014)*. Important for our model is that connectivity in the fly visual system can be well-approximated by average connectivity, resulting in convolutional filters, because of the crystalline character of the system *(Takemura et al., 2015; Seung, 2024)* (Section 2.3.2). Vice versa, in CNNs, channels can be viewed as neuron types that share the same receptive field structure and identical filter weights.

While both MLPs and CNNs capture key high-level structural properties of the fly visual system, besides many details such as morphology, anisotropy, and diversity of neuron types, they lack lateral and recurrent connectivity (Fig. 2.2 b vs. c). Recurrent connectivity enables a neuron that is *downstream* in a network to send feedback to a neuron that is *upstream*. We continue with covering recurrent neural networks (RNNs).

2.2.3 *The fly visual system is no discrete RNN*

An RNN's intermediate output depends not only on the intermediate input, but also on past inputs. This property is essential for modeling dynamical systems that depend on context information, like language, motion, or brain activity. Amon other early works, *Wilson and Cowan (1972)* present a neural network model based on firing rates rather than individual spikes of neurons, which can be seen as neuroscientific foundations of the RNN. *Hopfield (1982)* developed RNNs with content-addressable memory and demonstrated they could be trained with Hebbian learning. *Jordan (1986)* and *Elman (1990)* expanded RNN theory in general, and trained RNN variants on speech production with backpropagation. Many different RNN variants exist today *(Hopfield, 1982; Jordan, 1986; Elman, 1990; Hochreiter & Schmidhuber, 1997; Cho et al., 2014; Orvieto et al., 2023)*. Importantly, RNNs are universal function approximators for dynamical systems *(Funahashi & Nakamura, 1993; Schäfer & Zimmermann, 2007)*. I will start here with describing discrete RNNs, like the Elman-type RNN, central to machine learning foundations, and then move on to continuous-time RNNs in the next subsection, central to a lot of work in computational neuroscience.

A basic discretized RNN maintains a hidden state vector $\mathbf{h}_n = [h_i(t)] \in \mathbb{R}^N$, which is typically initialized as $\mathbf{h}(0) = \mathbf{0}$ and updated over discrete steps $t = 1, 2, \ldots, T$ using a recurrence relation. The state of an RNN at discretized time t, at the resolution of single neurons, can be described by the equation

$$h_i(t) = \sum_j^N w_{ij} f\left(h_j(t-1)\right) + \sum_j^D u_{ij} x_j(t) + b_i.$$

$\mathbf{W} = [w_{ij}] \in \mathbb{R}^{N \times N}$ represents network connectivity and $\mathbf{U} = [u_{ij}] \in R^{N \times D}$ transforms D-dimensional data to inputs to hidden neurons. The biases $\mathbf{b} \in R^N$ are constant offsets and f represents a nonlinear activation function. Elman-type RNNs apply the activation function outside of the right-handside term. One can achieve this by substituting with $z_i(t) = f(h_i(t))$ and applying f to both sides. Besides, discretized RNNs unrolled over time are equivalent to MLPs of T layers with shared weights and added inputs $\mathbf{U}\mathbf{x}(t)$ per hidden layer t.

To process spatiotemporal data like videos, RNNs can additionally be combined with parameter-sharing principles of CNNs. One can think of such convolutional RNNs as an expansion of MLPs which apply weight sharing across time and locally across space. A single-neuron in such a network evolves according to:

$$h_i(t) = \sum_{j \in \mathcal{N}_{c_i}} w_{c_i c_j} f\left(h_j(t-1)\right) + \sum_j^D u_{ij} x_j(t) + b_{c_i}.$$

This discretized recurrent neural network model still abstracts away the timescales and decay of dynamic neural computation. In the absence of inputs, it just integrates offsets. Real neurons maintain a constant membrane voltage through ion pumps instead. In 2.1, we already established a class of current-based RNNs from electrophysiology in continuous-time. Next, I will show you that one can derive the equivalent equations from the discretized RNN to anchor the DMN simultaneously in electrophysiology and machine learning.

2.2.4 *Continuous-time RNNs bridge biophysics and ML*

For many years, researchers use continuous-time recurrent neural networks (CTRNN) as mechanistic models for the study of dynamical neural network behavior. CTRNNs serve theoretical insights and the construction of bridges between theory and empirical observations of neural activity and behavior *(Prinz et al., 2004; Amarimber, 1972; Hopfield, 1984; Sompolinsky et al., 1988; Beer, 1995; Kleinfeld & Sompolinsky, 1988; Seung et al., 2000; Vogels et al., 2005; Rajan et al., 2010; Sussillo & Abbott, 2009; Sussillo & Barak, 2013; Mante et al., 2013; Sussillo, 2014; Pandarinath et al., 2018)*. *Funahashi and Nakamura (1993)* coined the modern term continuous-time recurrent neural network and showed they are universal approximators of dynamical systems. In ML, CTRNNs without inputs are also known as neural ordinary differential equations (neural ODE) *(Chen et al., 2018)*, whereas CTRNNs with inputs are known as neural controlled differential equations (neural CDE) *(Kidger et al., 2020)*.

We can convert discrete RNNs to continuos-time to anchor our DMN in machine learning fundamentals. Starting from the update

$$h_i(t) = \sum_j^N w_{ij} f(h_j(t-1)) + \sum_j^D u_{ij} x_j(t) + b_i,$$

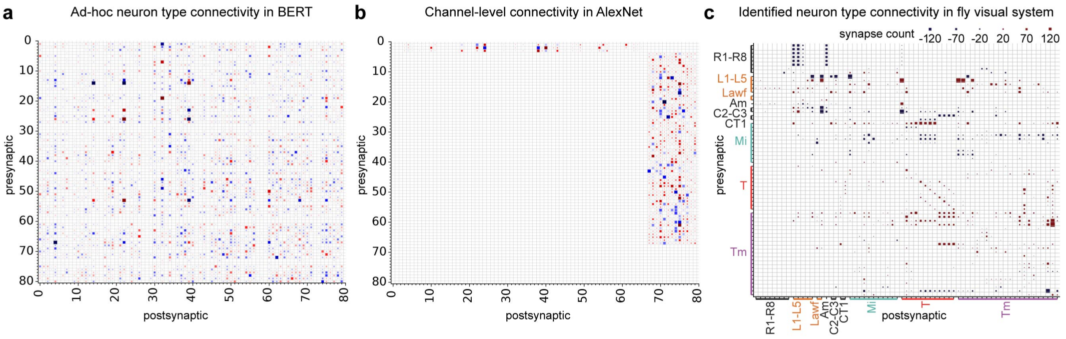


FIGURE 2.2: **Neuron-type level comparison of conventional neural network connectivity vs. connectome constraints. (a)** Ad-hoc feedforward neuron-type connectivity derived by clustering neurons based on their input and output weights in the fully connected 768×768 query matrix of the first attention layer in pretrained BERT (Bidirectional Encoder Representations from Transformers) *(Devlin et al., 2018)*, into 80 connectivity-defined neuron types. Blue (red) indicates weights with positive (negative) sign; square size encodes the average inter-type weight magnitude (arbitrary units). **(b)** Channel connectivity (analogous to neuron-type connectivity), computed by averaging spatial kernel weights between input and output channels across the first 80 channels in the first two convolutional layers of pretrained AlexNet *(Krizhevsky et al., 2012)*. Blue (red) indicates positive (negative) weights; square size represents the channel-averaged weight magnitude. **(c)** Empirically determined connectivity between 64 anatomically defined neuron types in the fly visual system, based on total synapse counts from all presynaptic neurons of one type to postsynaptic targets of another (Section 3.5.1). Blue (red) denotes putative hyperpolarizing (depolarizing) inputs; square size indicates total synapse count.

we subtract $h_i(t-1)$ from both sides, reintroduce the discretization step $\Delta t = 1$, and multiply the left handside with $\frac{\tau_i}{\Delta t} = 1$:

$$\tau_i \frac{h_i(t) - h_i(t-\Delta t)}{\Delta t} = -h_i(t-\Delta t) + \sum_j w_{ij} f(h_j(t-\Delta t)) + \sum_j u_{ij} x_j(t) + b_i. \tag{*}$$

In the finite difference limit $\Delta t \to 0$, we obtain

$$\tau_i \dot{h}_i(t) = -h_i(t) + \sum_j^N w_{ij} f\big(h_j(t)\big) + \sum_j^D u_{ij} x_j(t) + b_i.$$

Doing the reverse (discretizing) is known as the explicit Euler method, which is standard practice for numerically approximating the integral of a differential equation on digital computers. The remaining difference of the discretized continuous-time RNN to the discrete RNNs typically used in ML (Elman-type and others) are discretization steps Δt, and time constants τ_i.

The influence of these parameters on the dynamic updates becomes clear when rearranging the previous discrete equation (*). We obtain

$$h_i(t) = (1 - \frac{\Delta t}{\tau_i})h_i(t - \Delta t) + \frac{\Delta t}{\tau_i}\left(\sum_j^N w_{ij} f(h_j(t - \Delta t)) + \sum_j^D u_{ij} x_j(t) + b_i\right).$$

The bounded fraction $0 < \frac{\Delta t}{\tau_i} \leqslant 1$ determines the ratio of decay versus the ratio of input for the next update of neuron i. In discrete-time RNNs $\frac{\Delta t}{\tau_i} = 1$, and thus the decay is zero, whereas the input is maximal. From that follows that in continuous-time RNNs, neurons with $\Delta t < \tau_i$ can decay slower and maintain longer memory. Generally, CTRNNs with τ_i per neuron can combine variable timescales for computation. The theoretical and neuroscientific relevance of timescales is its own field of study *(Zeraati et al., 2024)*.

Let me finally close the loop to electrophysiology: In Section 2.1, we derived

$$\tau_i \dot{V}_i(t) = V_i^{\text{rest}} - V_i(t) + \sum_j w_{ij} f(V_j(t)) + e_i(t)$$

from a Hodgkin-Huxley-type neuron model. Mapping state symbol $h_i(t) \rightarrow V_i(t)$, bias or leak respectively $b_i \rightarrow V_i^{\text{rest}}$, and external input $\sum_j u_{ij} x_j(t) \rightarrow e_i(t)$ provides equivalence between both equations.

In general, CTRNNs are systems of ODEs parametrized by θ with input:

$$\dot{\mathbf{h}}_i(t) = g_\theta(\mathbf{h}(t), \mathbf{x}(t)).$$

Because the CTRNN ODE is simultaneously the continuous limit of modern RNNs (neural CDE) and a legitimate reduction of biophysics falls within that category, such networks are expressive machine learning tools and reasonably mechanistically interpretable. However, for detailed mechanistic interpretability for neuroscience one would require a model of every single neuron of the portion of the brain one wants to model. Fixing the weight matrix of such a network $W = [w_{ij}]$ to the fly connectome (subsection 2.4) lets us make such a model mechanistically interpretable at single-neuron resolution, because each neuron and each parameter retains physiological meaning. However, this works not without a caveat because we have no access to neural parameters such as the precise synaptic strengths, membrane time constants, and resting potentials.

2.2.5 *The power of backpropagation*

The application of *backpropagation* (backprop) *(Werbos, 1974)* to neural network architectures such as MLPs *(Rumelhart et al., 1986)*, CNNs *(LeCun et al., 1989)*, and RNNs *(Werbos, 1990)* mark inflection points in AI research. What makes backprop so important?

Backprop determines the parameters in a neural network that allow it to accurately compute output y, from input x. Initially, researchers randomly initialized neural network parameters and algorithms adjusted last layer parameters depending on the error between the network's output and the desired target *(Rosenblatt, 1962)*. However, deeper architectures to solve more complex tasks required training algorithms that could adjust each individual parameter of the network.

Backprop is a solution to this *credit assignment problem*: How much does each individual parameter in a deep neural network contribute to the network's output?

Again, given a dataset of S input-output pairs $\{\mathbf{x}_s, \mathbf{y}_s\}_{s=1}^{S}$, we look for parameters θ of a neural network ϕ such that $\hat{\mathbf{y}}_s = \phi_\theta(\mathbf{x}_s) \approx \mathbf{y}_s \; \forall s$. The *loss* $L_s : \mathbb{R}^{\text{out}} \to \mathbb{R}$ is a function that measures how far the prediction $\hat{\mathbf{y}}_s$ is from the true target $\mathbf{y}_s$, for instance $L_s = \|\hat{\mathbf{y}}_s - \mathbf{y}_s\|_2$. Backprop computes the gradient

$$\frac{\partial L_s}{\partial \theta_s} = \frac{\partial L_s}{\partial \hat{\mathbf{y}}_s} \cdot \frac{\partial \hat{\mathbf{y}}_s}{\partial \theta_s},$$

through the entire network with the *chain rule*. The gradient points into the direction of the minimum value of L.

Parameters are then updated after each presentation of a pair of training data, stepping towards the direction of the minimal loss. This is called *gradient descent*:

$$\theta_{s+1} \leftarrow \theta_s - \eta \cdot \frac{\partial L_s}{\partial \theta_s},$$

where η is the *learning rate*, controlling the size of the step. Repeated gradient descent over many input-output pairs improves the network gradually. In practice, the dataset is shuffled such that random pairs of training data are presented to avoid that the training algorithm picks up any unwanted structure in the dataset. This is called stochastic gradient descent.

The chain rule allows step-wise backwards computation of the derivates through the computational graph for automatic differentiation, which is now a standard tool broadly available through libraries like Pytorch and Jax *(Paszke et al., 2017; Bradbury et al., 2018)*.

Backprop allows neural networks to learn hierarchical transformations from input to output. In principle, it can be applied to learn the parameters of any differentiable function, such as deep neural networks with many layers or recurrent neural networks. Training recurrent neural networks requires backpropagation through time (BPTT) *(Werbos, 1990)*. However, analysis shows that gradient computation through time involves exponentiating the weight matrix T times *(Pascanu et al., 2013)*, which can cause gradients to vanish or explode, depending on the properties of the weight matrix. This makes optimization over long sequences challenging.

Critically, backprop requires large amounts of data, that accurately represent the input and output distributions, to incrementally move parameters towards a minimum that is also representative of unseen data. This shows in the fact that other landmarks in deep learning research are represented by the curation of large task-specific datasets in domains like computer vision *(Krizhevsky et al., 2012; Deng et al., 2009; Radford et al., 2021)*, motion perception *(Butler et al., 2012; Dosovitskiy et al., 2015)*, and text understanding *(Vaswani et al., 2017; Devlin et al., 2018; Mikolov et al., 2013; Radford & Narasimhan, 2018)*.

In contrast, the brain brain comes with a wiring that already reflects adaptation to the natural environment and it uses a variety of biological mechanisms for learning. Co-activation of neurons can lead to measurable changes in local connection strength *(Hebb, 1949; Bliss & Lomo, 1973)*, such synaptic plasticity depends on spike-timing *(Bi & Poo, 1998; Abbott & Nelson, 2000; Caporale & Dan, 2008)*, and on top of that neuromodulatory signals strongly alter circuit function *(Marder, 2012)*. While such brain mechanisms may approximate error-driven learning *(Richards et al., 2019; Lillicrap et al., 2020)*, backpropagation remains an engineering solution rather than a neuroscientific theory. More important for the deep mechanistic network here is that backpropagation can be used as a tool to train neuroscience models with many parameters.

2.2.6 *Deep learning is a tool for neuroscience*

Neuroscience circuit models often curate mechanistic detail to reproduce specific computations observed in simple nervous systems *(Prinz et al., 2004)*, similar to how early artificial intelligence research attempted implementing intelligence function-by-function *(Marr, 1982)*. These bottom-up approaches remain central for interpretability and scientific explanation. Yet, alone, they cannot scale to the thousands to billions of neurons and parameters in the brain necessary for high-dimensional, hierarchical computation. Here, top-down optimization approaches from deep learning help to make sense of the complexity and high-dimensionality of neural data *(Richards et al., 2019)*.

Deep neural networks are used today as a high-level modeling framework in neuroscience *(Yamins et al., 2014; Yamins & DiCarlo, 2016; Kriegeskorte, 2015; Jozwik et al., 2023)*, facilitated through advances such as automatic differentiation with neural network libraries *(Bradbury et al., 2018; Paszke et al., 2019)*, measurements of large-scale datasets of neural activity *(Tian et al., 2009; Ahrens et al., 2013; Stirman et al., 2016; Jun et al., 2017)* and behavior *(Mathis et al., 2018; Branson et al., 2009)*. Deep networks trained directly on network inputs and outputs of behavioral variables learn to replicate biological data by optimizing parameters over activity-matching *(Klindt et al., 2017; Cadena et al., 2019; Kindel et al., 2019; Zhang et al., 2019; Walker et al., 2019; Willeke et al., 2023; Ito & Murray, 2023)*, or task-specific *(Yamins et al., 2014; Yamins & DiCarlo, 2016; Kriegeskorte, 2015; Mante et al., 2013; Jozwik et al., 2023)* objectives.

Fundamental to the task-optimization approach in this thesis was the discovery that CNNs trained on object recognition build representations similar to neural activity recordings from primate visual cortex *(Yamins et al., 2014; Khaligh-Razavi & Kriegeskorte, 2014)*. In some cases such task-optimized DNNs explain the brain's visual representations better than a detailed visual system model with realistic anatomical and physiological constraints, demonstrating the power of the top-down approach. Many similar findings for alignment of DNN representations to cortical representations followed: For instance, models optimized for speech and music recognition explain representations in human auditory cortex *(Kell et al., 2018; Tuckute et al., 2023)*, models optimized on cognitive tasks explain representations in primate prefrontal cortex *(Yang et al., 2019; Piwek et al., 2023)*, models trained to predict limb position and velocity explain representations in primate sensory cortex *(Marin Vargas et al., 2024)*, and models trained on speech recognition explain representations in human brain regions associated with language understanding *(Schrimpf et al., 2021)*, suggesting that the representations task-optimization induces into neural networks generally align with brain representations even without accounting for anatomical or physiological details. The conclusion is that if such networks trained for task performance develop brain-like representations, then task-optimized networks can serve as *in-silico* testbeds for neuroscience.

However, the strength of the task-optimized DNN approach not requiring anatomical or physiological details also turns into the approache's biggest weakness in context of neuroscience modeling: While DNN representations often correspond to brain data according to regression scores, such models cannot explain single-neuron computation *(Schaeffer et al., 2022; 2024)*. Therefore DNN may be good models of data deduced from measurements in the brain but they are often bad models of the brain.

Conventional deep networks remain difficult to interpret mechanistically typically simply because they have many neurons. Understanding mechanisms of deep neural network requires *mechanistic interpretability* research *(Bereska & Gavves, 2024)*, which has parallels to neuroscience.

Nevertheless, despite such parallels and DNN's empirical success as models for neuroscience, their neurons have no biological identity and their parameters have limited physiological grounding. Interpretations often remain at the level of representations instead of causally explaining neural circuits *(Mathis et al., 2024)*. Representational similarity to a DNN does not imply that model neurons and real neurons send the same electrical signals. Conversely, in a computational model constructed from a connectome, we can begin to understand unknown mechanistic implementations of (nonlinear) computations that lead to goal-driven behavior at the single-neuron level.

2.2.7 *Probabilistic deep learning to tackle the inverse problem*

In computational neuroscience one wants to find hypotheses for neural mechanisms from observed behavior or recordings of neural activity with models, but generated hypotheses can differ depending on the exact model parameters *(Prinz et al., 2004; Marder & Taylor, 2011; Gonçalves et al., 2020)*. Such ill-defined inverse problems occur when multiple mechanisms can generate the same data, making model parameters *non-identifiable*. While connectomic measurements provide a scaffold of constraints to identify the unknown parameters, the inverse problem in connectome-constrained models usually remains ill-defined *(Biswas & Fitzgerald, 2022)*. We will see that even in connectome-constrained models with additional task constraints competing hypotheses for neural function remain. Below I provide a scaffold to frame our strategy to refine hypothesis generation for identifying specific functional predictions within a probabilistic framework.

Bayesian inference addresses this by seeking a posterior distribution $p(\theta \mid \mathcal{D})$ over parameters θ given data $\mathcal{D} = \{\mathbf{x}_s, \mathbf{y}_s\}_{s=1}^S$ rather than searching for a single best parameter realization. This posterior is used to compute the predictive distribution, i.e., the possible realizations of outputs under the same input, averaged over plausible parameterizations:

$$\begin{aligned} p(\hat{\mathbf{y}} \mid \mathbf{x}, \mathcal{D}) &= \int p(\hat{\mathbf{y}} \mid \mathbf{x}, \theta)\, p(\theta \mid \mathcal{D})\, d\theta \\ &\approx \frac{1}{M} \sum_{m=1}^{M} p(\hat{\mathbf{y}} \mid \mathbf{x}, \theta_m) \quad \text{with } \theta_m \sim p(\theta \mid \mathcal{D}),\ m = 1, \dots, M. \end{aligned}$$

In practice, we approximate posterior expectations via Monte Carlo sampling over the draws θ_m, since posteriors in complex models are typically high-dimensional and analytically intractable. This allows computing weighted averages of random variables that depend on the model parameters, such as internal states of the model itself.

Because exact Bayesian inference in deep neural networks is intractable multiple approximation schemes have emerged. A pragmatic approach is to treat the stochastic training process itself as a source of posterior variation *(Lakshminarayanan et al., 2017; Wilson, 2020)*. Running the same optimization multiple times with different random seeds yields an *ensemble* of solutions that captures variation in the learned functions due to non-identifiability and multimodal posteriors.

A *deep ensemble* is represented by a set of M parameters $\{\theta_m\}_{m=1}^M$ of a deep neural network ϕ, each trained to minimize the same loss, and can be interpreted as samples from an approximate posterior distribution over θ. For each run m, training typically converges to a local mode $\theta_m^\star$ of the posterior $p_\phi(\theta \mid \mathcal{D})$ (a MAP estimate).

Deep ensembles are empirically effective for approximate Bayesian inference, providing more robust predictions and better-calibrated uncertainties than most single-model methods *(Lakshminarayanan et al., 2017; Dietterich, 2000)*.

In DNNs, the parametrized model describing the transition from $\mathbf{x}$ to $\hat{\mathbf{y}}$ is typically deterministic, such that the likelihood $p(\hat{\mathbf{y}} \mid \mathbf{x}, \theta)$ collapses to a Dirac delta function: $p(\hat{\mathbf{y}} \mid \mathbf{x}, \theta) = \delta(\hat{\mathbf{y}} - f_\theta(\mathbf{x}))$. The posterior predictive distribution then becomes a sum over point masses:

$$p(\hat{\mathbf{y}} \mid \mathbf{x}, \mathcal{D}) = \int \delta(\hat{\mathbf{y}} - f_\theta(\mathbf{x}))\, p(\theta \mid \mathcal{D})\, d\theta \approx \frac{1}{M} \sum_{m=1}^{M} \delta(\hat{\mathbf{y}} - f_{\theta_m}(\mathbf{x})).$$

This is a posterior-weighted ensemble of deterministic model outputs. Each θ_m is an approximate draw from the posterior and we use Monte Carlo approximations of the posterior predictive distribution.

Importantly, for mechanistic models, deep ensembles allow us to explore different data-consistent explanations. By analyzing how ensemble members vary, we can quantify uncertainty in inferred mechanisms. For instance, we can average internal variables, such as the predicted voltage $V_i(t)$ of a specific neuron, over the ensemble:

$$\mathbb{E}_{\theta \sim p(\theta|\mathcal{D})}[V_i(t)] \approx \frac{1}{M} \sum_{m=1}^{M} V_i^{(m)}(t).$$

This gives a posterior-averaged estimate of neuron function, useful for interpreting what a neuron "does" in expectation over the model space.

However, when either the posterior or its induced distributions over neural responses are multimodal, a single average may obscure functional differences across modes. In such cases, clustering can be performed in function space — based on neural responses $V_i^{(m)}$ rather than parameters — to identify computationally different solutions. For resulting functional modes $\{\mathcal{M}_1, \mathcal{M}_2, ..., \mathcal{M}_K\}$ mode-specific averages are:

$$\mu_i^{(k)}(t) = \frac{1}{|\mathcal{M}_k|} \sum_{m \in \mathcal{M}_k} V_i^{(m)}(t),$$

where each $\mu_i^{(k)}(t)$ captures the response of neuron i under a distinct explanatory hypothesis. Because clustering in high-dimensional function spaces of neural activity is computationally challenging and can behave unintuitively, one can perform (nonlinear) dimensionality reduction to the functional data *(Cunningham & Yu, 2014)* using established tools such as uniform manifold approximation and projection (UMAP) *(Becht et al., 2019)*.

Mode-conditional expectations reflect the functional diversity of response predictions, supporting nuanced mechanistic interpretation of neural computations in the model.

2.3 THE FLY

> *"9. Dros. melanogaster.*
> *Kopf, Rückenschild und Beine lehmgelb; Hinterleib schwarz.*

> *Capite thorace pedibusque luteis; abdomine nigro.*
> *Die Schwinger sind weiß, die Flügel ungefärbt. – Aus Oesterreich, von Kiel und von Hamburg. – 1 Linie."* – *Meigen (1830)*

Johann Wilhem Meigen's brief characterization may represent the first historical recording of the species. Today, *Drosophila melanogaster* is a central model organism in biology. The common fruit fly — I refer to it simply as fly throughout this thesis — is an invertebrate insect, roughly two to five millimeters in size (not much bigger than a pinhead), that feeds on sugar and fermentation products and reproduces within 10 days from egg to adult. Its adult lifespan is 30 to 40 days under optimal lab conditions, but depends on factors such as temperature, diet, mating status, genetics, and environment.

Its availability, fast reproduction, and genetic advantages position the fly as a popular model system in biology research *(St Johnston, 2002)*. Among others, it is used as a model for studying neurodegenerative disease *(Muqit & Feany, 2002; Prüßing et al., 2013)*, as a model for cancer research *(Gonzalez, 2013)*, and as a model of organ dysfunction *(Dow et al., 2022)*. Consortia of researchers spanning many institutions and countries worldwide collaborate on fly science. Repositories of the organism's genetics, molecular biology *(FlyBase Consortium, 1998; Drysdale et al., 2005)*, nervous system structure *(Dorkenwald et al., 2022; Plaza et al., 2022; Ukani et al., 2016)*, function and other properties *(Court et al., 2023)*, are made publicly available for effective research progress *(Matthews et al., 2005)*.

The fly represents a useful model organism, for neuroscience research specifically. Genetic tools allow for example to image large-field neural activity via calcium and voltage indicators *(Fiala et al., 2002)*, to investigate single-neuron properties such as neurotransmitter and receptor expression at synapses *(Davis et al., 2020)*, and to remove neural circuit parts for studying their causal role in behavioral assays *(Branson et al., 2009; Sato et al., 2025)*. Tools such as FlyLight image-based Split-GAL4 driver line collections *(Meissner et al., 2024)* and optogenetic control of neural activity *(Inagaki et al., 2014)* facilitate precise manipulation and observation of individual circuit elements.

Fundamental research on the invertebrate fly leads to key discoveries that also impact vertebrate neuroscience *(Bellen et al., 2010)*. Often, neural systems in the fly share organizational and computational principles with corresponding systems in vertebrates *(Borst & Helmstaedter, 2015; Tanaka & Portugues, 2025)*. The toolset surrounding the fly facilitates the scientific process itself, allowing the discovery of methods that surpass the boundaries of individual model organisms.

2.3.1 *The right organism for the discovery of methods for connectome simulation*

A main question in this thesis is whether a connectome can be used to derive a functional, dynamic brain model. Selecting an appropriate model organism makes this question tangible by enabling specification and validation of computational implementations.

Connectome

The term **connectome** refers to different types of connectivity in nervous systems depending on scientific context.

- **Structural connectome:** Anatomical map of neural elements and connections, spanning multiple spatial scales. *Nanoscale* connectomes **(used here)** resolve synaptic connectivity via electron microscopy *(White et al., 1986; Denk & Horstmann, 2004; Knott et al., 2008)*; *mesoscale* connectomes map long-range projections between brain regions using high-resolution imaging *(Oh et al., 2014; Winnubst et al., 2019)*; *macroscale* connectomes estimate region-to-region white matter pathways from diffusion tractography *(Basser et al., 2000)*. Although the term originally implied completeness *(Sporns et al., 2005; Hagmann, 2005)*, it is used here also for partial reconstructions.

- **Functional connectome:** Statistical correlations of neural activity between brain regions at a macroscale, commonly derived from population-level imaging methods such as fMRI *(Biswal et al., 1995)*.

- **Neuropeptide connectome:** Network of neuromodulatory signaling pathways mediated by neuropeptides, involving volume transmission beyond synaptic transmission *(Ripoll-Sánchez et al., 2023)*.

Extended definitions may include gap junctions, glial networks, or other brain data.

Connectome simulation is not a new concept. The nematode C. elegans, with only 302 neurons *(White et al., 1986)*, was the first organism for which a connectome was mapped and simulated *(Varshney et al., 2011)*. While these models could partly explain behaviors, they remain limited in providing mechanistic explanations. Worm behavior often arises from closed-loop interactions with its environment, so purely neural simulations fall short. Even embodied models *(Chen et al., 2023; Sarma et al., 2018)* fail to fully capture behavior because they miss physiological detail — especially the dense extrasynaptic network of neuropeptide signaling, which profoundly influences worm behavior *(Ripoll-Sánchez et al., 2023; Watteyne et al., 2024; Taylor et al., 2021; Beets et al., 2023)*.

Beyond worms, there are efforts to simulate portions of the rat *(Markram et al., 2015)*, mouse *(Billeh et al., 2020)*, and human brain *(Amunts et al., 2016)*. However, these models face additional challenges: they are computationally demanding, lack behavioral grounding, and have no completion of connectomic data in sight *(Eliasmith & Trujillo, 2014)*. To cope with data limitations, these models often rely on statistical assumptions about synaptic connectivity *(Hill et al., 2012)*, or mean-field approximations to simplify network dynamics *(Nakagawa et al., 2013)*. Parameter tuning is usually manual and based on plausible physiological ranges *(Nakagawa et al., 2013; Potjans & Diesmann, 2014)*, rather than being driven by input-output computations.

Simulation of intrinsic, cortical circuits requires abstract models of sensory circuits for input to the system and abstract models of behavioral targets for system output *(Billeh et al., 2020; Potjans & Diesmann, 2014)*, making it difficult to formulate testable input-output computations mechanistically in relation to behavior *(Eliasmith & Trujillo, 2014)*. Much reconstruction work focused on cortical circuits *(Motta et al., 2019; Shapson-Coe et al., 2021; Sievers et al., 2024)*, which have no universally defined function *(Horton & Adams, 2005)*, complicating efforts to derive and validate mechanistic models. In summary, mechanistic modeling with portions of rat, mouse, or human brain remains challenging and explorative for the above reasons, while large-scale multimodal measurements of connectivity and function *(MICrONS Consortium, 2025)* at high-resolution for integrated model building and validation with tools from deep learning holds promise *(Willeke et al., 2023)*.

Although fly connectome simulations face similar challenges as those in other organisms, Drosophila offers distinct advantages that make it especially well-suited for developing simulation tools and exploring what is required for mechanistic models of brain function grounded in behavior (Fig. 2.3).

In a recent scientific revolution, several comprehensive and high-quality reconstructions of the fly connectome became publicly available *(Nern et al., 2025; Dorkenwald et al., 2024; Marin et al., 2023; Takemura et al., 2024; Schlegel et al., 2024)*. The entire fly brain comprises approximately $140,000$ neurons and $130,000,000$

chemical synapses *(Dorkenwald et al., 2024)* and the ventral nerve cord comprises 23, 000 neurons and 10, 000, 000 chemical synapses *(Takemura et al., 2024)* — a scale within a computationally practical regime for simulation and optimization.

Moreover, neurotransmitter and receptor expressions can be measured using genetic tools *(Scheffer & Meinertzhagen, 2021; Davis et al., 2020)* and inferred from EM images at large scale with machine learning *(Eckstein et al., 2024)* for determining signs of synaptic connections. Decades of research provide circuit-level mechanistic insights into the visual system (Section 2.3.3), central complex *(Pfeiffer & Homberg, 2014; Franconville et al., 2018)*, and mushroom body *(Modi et al., 2020)*. Crucially, these circuits are linked to complex behaviors such as motion responses, navigation, and learning. Altogether, the availability of detailed data, the manageable scale of the nervous system, and the presence of well-defined behaviors make the fruit fly an exceptionally strong candidate for advancing connectome-constrained brain modeling.

The fly visual system, particularly the motion-detection pathway, has several properties that make it uniquely advantageous for computational modeling and optimization. First, fly visual system neurons often rely on graded potentials rather than on spikes *(Laughlin & Hardie, 1978; Zettler & Järvilehto, 1971; Kretzberg et al., 2001)*, simplifying modeling approaches that rely on differentiability. The system also benefits from clearly characterized sensory input, namely light stimuli, enabling precise definition of input-output relationships. Additionally, the motion detection circuit receives minimal feedback from other brain regions or physical variables beyond light itself, allowing that its isolated analysis is meaningful. Finally, the availability of deep learning tools and datasets from computer vision research offers a solid basis for parameter inference and optimization for a visual system model.

2.3.2 *The almost-convolutional fly visual system*

Anatomical studies of visual system neurons in fly species date back more than a century *(Cajal & Sánchez, 1915)*. Mesoscale studies carefully mapped neural anatomy and characterized neuron types based on morphological features and connectivity, historically relying on staining brain tissue and using light-microscopy imaging *(Cajal & Sánchez, 1915; Strausfeld, 1976; Fischbach & Dittrich, 1989)*.

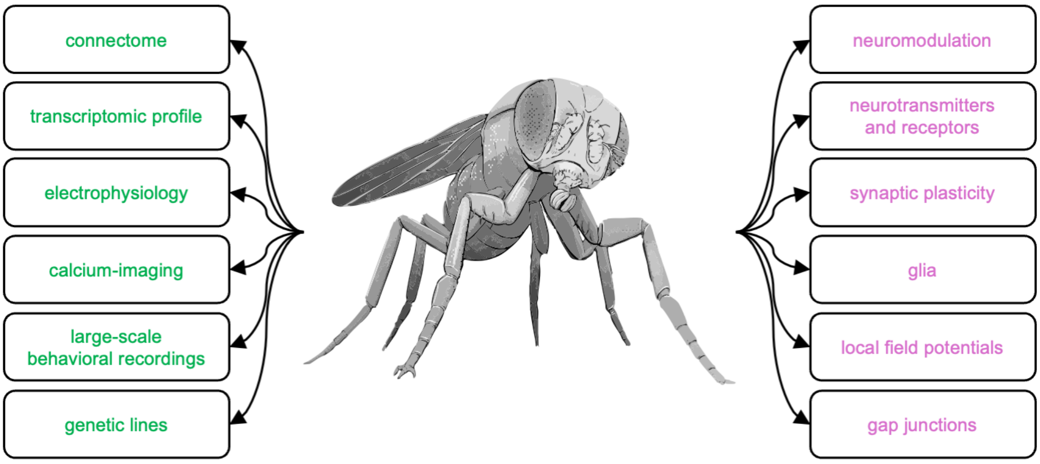


FIGURE 2.3: ***Drosophila melanogaster* in biology research.** Researchers leverage **(left, green)** its fast reproduction cycle and available genetic tools *(Brand & Perrimon, 1993)*, the possibility of making large-scale behavioral recordings *(Branson et al., 2009)*, large-scale calcium imaging of neural activity *(Ahrens et al., 2013; Fiala et al., 2002)*, precise electrophysiological measurements of neural activity *(Wilson et al., 2004)*, transcriptomic characterization of neuron types *(Mortazavi et al., 2008)*, and measurements of the connectome, every neuron and synaptic connection in the fly *(Scheffer et al., 2020; Dorkenwald et al., 2024; Schlegel et al., 2024)*. **(right, magenta)** At the same time, significant gaps remain in our understanding of neuromodulation, neurotransmitter and receptor types and locations, synaptic plasticity, glia cells, the role of local field potentials, and gap junctions across the fly nervous sytem *(Scheffer & Meinertzhagen, 2021)*. Original illustration by Franz Stämmele for the Machine Learning for Science blog *(Lappalainen, 2024)*.

Fast-foward to today: The advancement reaches from nanoscale synapse-level reconstructions initially of the central portion of the fly visual system *(Rivera-Alba et al., 2011; Takemura et al., 2015; 2017; Shinomiya et al., 2019; 2022)*, to two recently finished reconstructions and characterizations of complete visual systems at nanoscale resolution, one from a male fly *(Nern et al., 2025)*, and one from a female fly *(Matsliah et al., 2024)*. Today's reconstructions rely on sophisticated electron microscopy pipelines, such as focused ion beam scanning electron microscopy (FIB-SEM) *(Knott et al., 2008)* or transmission electron microscopy (TEM) *(Zheng et al., 2018)*, semi-automated annotation with deep learning tools *(Januszewski et al., 2018)*, and large-proof reading efforts *(Dorkenwald et al., 2022)* using powerful platform-independent and hardware-accelerated tools such as neuroglancer for real-time 3D rendering and proofreading *(Maitin-Shepard et al., 2021)*.

To give a brief overview of structure with numbers I refer to data from *Nern et al. (2025)* which differ slightly from data in *Matsliah et al. (2024)* because of inter-individual variation and reconstruction quality. *Currier et al. (2023)* comprehensively reviews fly visual system anatomy and function.

The fly visual system consists of the *compound eye*, the external structure supporting vision, and the *optic lobe*, the intrinsic nervous system structure.

The compound eye has ~800 hexagonally arranged ommatidia. Each ommatidium hosts up to eight photoreceptor neuron types (R1-R8). R1-R6 project into the optic lobe's lamina and R7-R8 project into the medulla neuropil *(Fischbach & Dittrich, 1989; Mishra & Knust, 2013)*.

The optic lobe comprises about 53,000 neurons and 49 million connections, grouped into 732 systematically classified neuron types. Both optic lobes combined host 75.7 % of neurons of the entire fly brain *(Nern et al., 2025; Dorkenwald et al., 2024)*, which demonstrates the important role vision plays for the animals survival. It subdivides into five *neuropils*, separate tissues of nerve fibers with synaptic connections.

The first neuropil, *lamina*, hosts about 7,000 neurons (13.2% of the optic lobe neurons and 5% of the brain), grouped into 13 neuron types, with low connectivity *(Nern et al., 2025)*. Fibers in the lamina organize stereotypically into *cartridges*, which inherit the hexagonal arrangement from the retina *(Rivera-Alba et al., 2011)*. The lamina separates ON- and OFF-pathways anatomically and sends projections via lamina monopolar neurons (L1-L5) to the medulla *(Rivera-Alba et al., 2011)*.

The second neurpopil, *medulla*, hosts about 42,000 neurons (79,2 % of the optic lobe neurons and 30% of the brain), which group into 345 neuron types. It contains by far the greatest number of connections between neurons among all five neuropils *(Nern et al., 2025)*. The medulla can be subdivided into ten layers of different innervation patterns and depths from lamina neurons (L1-L5), retina neurons (R7-R8) and others. The medulla hosts intrinsic neurons (Mi), transmedullary neurons (Tm), transmedullary neurons with Y-shape (TmY), distal medulla (Dm) interneurons, and proximal medulla (Pm) interneurons. Neurons in the medulla contribute to ON and OFF luminance pathways, object detection, color vision, and motion detection *(Currier et al., 2023; Matsliah et al., 2024)*. Medulla neuron types have varying spatio-temporal receptive fields *(Arenz et al., 2017)* and serve as asymetric inputs to motion selective neurons *(Takemura et al., 2017; Shinomiya et al., 2019; Takemura et al., 2013)*. Tm and TmY neuron types innervate the lobula and lobula plate *(Nern et al., 2025; Fischbach & Dittrich, 1989; Matsliah et al., 2024)*. The neurites in the medulla inherit retinotopic column coordinates from the ommatidia and cartridge system.

The third neuropil, the *accessory medulla*, is a small and highly specialized neuropil with only 131 neurons but 58 neuron types with moderate connectivity per neuron. It serves as part of the circadian circuit *(Helfrich-Förster et al., 2007; Reinhard et al., 2024)* and has no obvious retinotopic correspondance *(Nern et al., 2025)*.

The fourth neuropil, *lobula*, hosts 26,000 neurons (49 % of the optic lobe neurons and 18.5% of the brain), which group into 422 neuron types and has about half as much connectivity as the medulla. Connectivity in the lobula plate again halfs in comparison to connectivity in the lobula. The lobula and the fifth neuropil, *lobula plate*, form the *lobula complex*. The *lobula plate* hosts about 13,000 neurons (24% of the optic lobe neurons and 9% of the brain), which group into 126 neuron types. Among many innervations from medulla neuron types, the lobula complex contains neurites of the well-studied, columnar ON-edge motion detectors (T4a-T4d) *(Takemura et al., 2017)*, OFF-edge motion detectors (T5a-T5d) *(Shinomiya et al., 2019)*. Moreover, it hosts lobula columnar neurons (LC) which relay motion and object information to the *photocerebrum* in the central brain *(Fischbach & Dittrich, 1989)*. The lobula complex further inherits retinotopic columns through innervations from the medulla *(Nern et al., 2025)*.

The optic lobe resembles an almost crystalline architecture, like a convolutional neural network. The following numbers illustrate this: While the optic lobes together host 75.7 % of neurons of the entire fly brain, they only host 8.7 % of neuron types in the fly brain *(Nern et al., 2025; Schlegel et al., 2024)*. Neuron typing in these studies is based on connectivity and morhpology of neurons. Neurons with similar connectivity and morhpology belong to the same neuron type. Therefore, these fractions illustrate that neurons are abundant but less varied in the optic lobe in comparison to other parts of the nervous system. This mirrors a core property of the visual system: To enable equivariant processing, the optic lobe consists of a consensus circuit of neurons from equivalent neuron types that tiles the optic lobe *(Fischbach & Dittrich, 1989; Nern et al., 2015)* and carries out nearly the same mechanistic computation in each of many columns of the visual system in parallel. Detailed reconstructions show that the majority of synaptic contacts in the motion detection subsystem can be mapped onto a consensus circuit *(Takemura et al., 2008)*. Moreover, the polarity of synaptic connections (excitatory or inhibitory), determined by a combination of transcriptomic profiling of neurotransmitters and receptors with deep learning *(Nern et al., 2025; Davis et al., 2020; Eckstein et al., 2024)*, is the same for connected neurons from the same neuron-type pair. The resulting consensus circuit approximates weight sharing as in convolutional neural networks (Section 2.2.2) and enables equivariant local motion detection across the visual system.

2.3.3 *Circuit mechanisms*

A significant percentage of the fly brain is dedicated to vision. To study visual mechanisms that are fundamental for behavior, the fly is ideally suited as a model organism because of the available toolsets for genetic control, anatomical reconstruction, and measurement of neural activity and behavior *(Currier et al., 2023; Ryu et al., 2022)*.

On a functional high-level, the fly relies on visual information to navigate, procreate, and survive. For instance, visual computation underlies optomotor responses *(Bahl et al., 2013)*, walking *(Creamer et al., 2018)*, and flight behavior *(Fry et al., 2009)*, its courtship relies on visual mechanisms to target mates *(Yamamoto & Koganezawa, 2013)*, and its survival hinges on a recurrent visual machinery for fast motor planning to escape predators *(Card & Dickinson, 2008)*.

On a lower-level, the information that the fly visual system extracts from light contains cues such as luminances and colors in different parts of the visual field, informing for instance phototaxis behavior *(Carpenter, 1905; Heisenberg & Buchner, 1977; Miller et al., 1981)* and circadian rythm *(Helfrich-Förster, 2020)*. Importantly, the fly visual system extracts *optic flow*, local motion across the field of view *(Gibson, 1950)*. Perceiving optic flow is essential for the fly's ability for locomotion control, including its ability to walk and fly *(Fry et al., 2009; Borst & Haag, 2002; Mauss & Borst, 2020)*. Combined with its ability to perceive object shapes and sizes *(Keleş & Frye, 2017; Klapoetke et al., 2022)*, and its own motion *(Chiappe, 2023)*, the fly visual system encodes the behaviorally relevant information to effectively identify, detect, and navigate around moving objects in its environment *(Turner et al., 2022)*.

Theoretical research forwarded different models to explain how neural circuits could implement motion detection *(Hassenstein & Reichardt, 1956; Reichardt, 1961; Barlow & Levick, 1965; Borst, 2000)*. Fundamentally, the Hassenstein-Reichardt model explains local motion detection in a neuron as a multiplicative interaction between neighboring signals, where a temporal delay on one side introduces directional asymmetry, producing a motion-selective response *(Hassenstein & Reichardt, 1956)*. The Barlow-Levick model explains local motion detection as the result of an inhibitory input that suppresses responses to motion in the null direction, producing direction selectivity through asymmetric inhibition *(Barlow & Levick, 1965)*. While these results were particularly formative of much follow-up research, today's experimental measurements offer a more nuanced view.

Today, an unprecedented wealth of knowledge on visual circuit mechanisms exists, after about thirty years of experimentally measuring visual neuron type responses in the fly to make mechanistic explanations *(Gruntman et al., 2018; Groschner et al., 2022; Arenz et al., 2017; Hardie, 1991; Ranganathan et al., 1991; Peretz et al., 1994; Nikolaev et al., 2009; Reiff et al., 2010; Clark et al., 2011; Freifeld et al., 2013; Silies et al., 2013; Maisak et al., 2013; Strother et al., 2014; Meier et al., 2014; Behnia et al., 2014; Fisher, Leong, et al., 2015; Hardie & Juusola, 2015; Leonhardt et al., 2016; Fisher, Silies, & Clandinin, 2015; Yang et al., 2016; Serbe et al., 2016; Strother et al., 2017; Meier & Borst, 2019; Gruntman et al., 2019; Drews et al., 2020; Matulis et al., 2020; Ramos-Traslosheros & Silies, 2021a; Gruntman et al., 2021; Ketkar et al., 2022; Braun et al., 2023; Ammer et al., 2023; Pirogova & Borst, 2023; Gür et al., 2024; Pang et al., 2025; Currier & Clandinin, 2025)*. Because this research — in particular, the experimental work focusing on neuron types central to motion detection mechanisms — is indispensable for validating our method and model of the fly visual system, I will review the most important insights on a high-level to provide a scaffold for understanding the description of the modeling work that follows.

Motion detection in the fly: From the retina to T4 and T5

My description in the following focuses on aggregating core response properties of motion detection subcircuit mediated by R1-R6, such as

ON-selective, OFF-selective, and motion-selective responses across the literature, even though this can only represent a coarse introduction to the nuanced spatio-temporal response properties of individual neuron types described in the cited literature. Moreover, while knowing the inhibitory and excitatory properties of synaptic connections which lead to these responses is crucial for making mechanistic explanations, I will factor out explanations of underlying neurotransmitter and receptor interactions *(Davis et al., 2020)*.

In the retina, photoreceptor neurons depolarize in response to photons that pass through the ommatidial lenses and activate a biophysical machinery of *phototransduction*, the translation of information encoded by photon number and wavelengths to voltage signals *(Hardie, 1991; Ranganathan et al., 1991; Peretz et al., 1994; Hardie & Raghu, 2001)*. These responses are fast: Fly photoreceptors have a 10 times faster molecular machinery to transduce photons than mammalian rods, resulting in sensitivity to up to 100 Hz flicker *(Hardie & Raghu, 2001)*. The photoreceptors transduce light in ultraviolett, blue, and spectral ranges. On a high-level, receptor types R1-R6 primarily feed into the motion detection circuit and are sensitive to a broad spectrum of light *(Heisenberg & Buchner, 1977; Yamaguchi et al., 2008)*. R7-R8 primarily feed into the color subsystem and are sensitive to UV, green, and blue wavelengths *(Morante & Desplan, 2008; Yamaguchi et al., 2010)*. However, interactions across photoreceptor types and functional subcircuits make this functional separation hold only coarsely *(Wardill et al., 2012; Schnaitmann et al., 2013; Longden et al., 2023)*.

In the Drosophila visual system, photoreceptors and many downstream neurons communicate via graded, non-spiking potentials. Voltage signals propagate passively along membranes without triggering action potentials, enabling fast and energy-efficient analog computation *(Kretzberg et al., 2001; Juusola & Hardie, 2001; de Ruyter van Steveninck & Laughlin, 1996)*.

Lamina monopolar neurons forward spatio-temporally filtered signals containing local information of luminance and contrast to medulla neurons *(Clark et al., 2011; Silies et al., 2013)* and adapt in concert with photoreceptors to varying light conditions *(Nikolaev et al., 2009; Zheng et al., 2009)*. Importantly, in the lamina, the fly visual system splits into ON and OFF pathways *(Joesch et al., 2010)* — a common circuit design across species that allows full dynamic range encoding without increased metabolic cost *(Borst & Helmstaedter, 2015; Gjorgjieva et al., 2014)*. L1 and L5 feed into ON-pathway medulla neurons and L2 and L4 feed into OFF-pathway medulla neurons. L3 sends signals to neuron types from both pathways.

The spatio-temporal filtering properties across medulla neuron types vary. Their responses are spatially and temporally filtered versions of the input from lamina neurons *(Groschner et al., 2022; Arenz et al., 2017; Strother et al., 2014; Meier et al., 2014; Behnia et al., 2014; Fisher, Silies, & Clandinin, 2015; Yang et al., 2016; Serbe et al., 2016; Strother et al., 2017; Ramos-Traslosheros & Silies, 2021a)*.

Abundant neurons in the medulla provide inputs to the motion-detector neurons of type T4 and T5. T4 encode motion of light increment edges *(Gruntman et al., 2018; Groschner et al., 2022; Maisak et al., 2013; Leonhardt et al., 2016; Fisher, Silies, & Clandinin, 2015; Strother et al., 2017; Gruntman et al., 2019; Ramos-Traslosheros & Silies, 2021a; Gruntman et al., 2021)*, and T5 neurons encode motion of light decrement edges *(Maisak et al., 2013; Fisher, Silies, & Clandinin, 2015; Gruntman et al., 2019; Ramos-Traslosheros & Silies, 2021a)*. T4 and T5 neuron types split into four subtypes each, encoding motion to each of four caridnal directions. Their motion selectivity results both from anatomical and functional asymmetries of their inputs *(Borst et al., 2020)*.

In addition to the coarse summary of 32 neuron type responses we refer to in our main work for model validation, the majority of the recently discovered 732 neuron type *(Nern et al., 2025)* responses are still uncharacterized. It is experimentally challenging to record with specificity from such a large number of neuron types, which poses fundamental challenges to comprehensively describing circuit mechanisms at single-neuron resolution across the circuit. Even with full measurement of 732 neuron type responses to specific stimuli, no predictions could be made for unseen stimuli. Therefore, our goals include combining existing theoretical and experimental neuroscience knowledge with tools from deep learning to create a comprehensive model of neural computation for hypothesis generation for instance for such uncharacterized neuron type responses to arbitrary stimuli.

The fly visual system is now one of the best studied portions of an animal brain, both structurally and functionally. Yet, many questions remain, including, how do the 600+ uncharacterized neuron types contribute to computation? We create a framework for connectome-constrained modeling to discover and refine hypotheses for mechanistic computation of single neurons inside biological neural networks.

2.4 THE DEEP MECHANISTIC NETWORK

> *"Guided by connectomic and other data to optimize thousands or even billions of parameters, machine-learning models could be trained to produce neural-network behaviour that is consistent with the behaviour of real neural networks - measured using cellular-resolution functional recordings."* – *Jain (2023)*

The deep mechanistic network (DMN) integrates knowledge from different disciplines, such as connectomics and experimental neuroscience, computational neuroscience, deep learning, and probability theory (Fig. 2.4) into a comprehensive computational model.

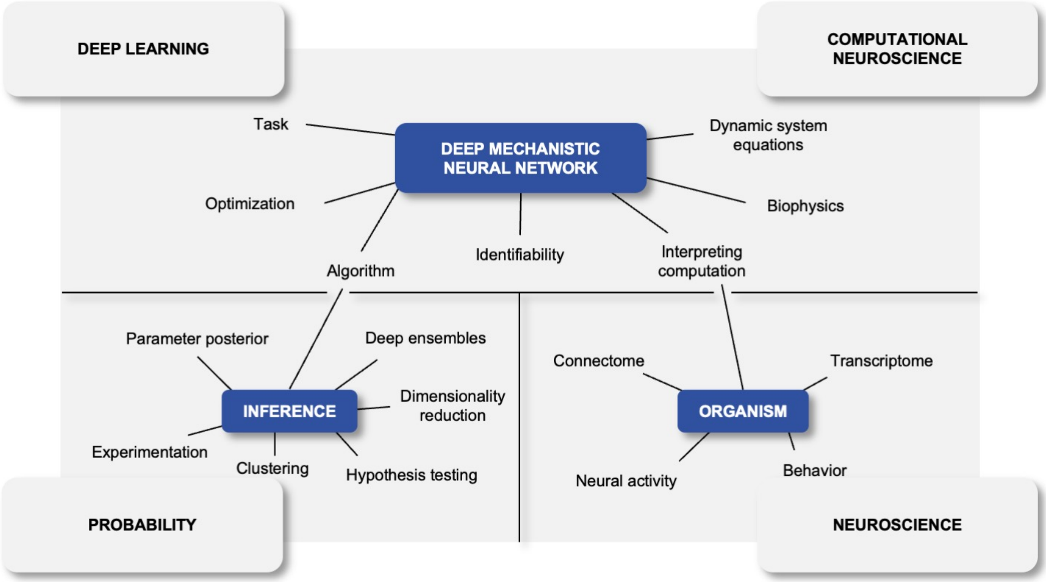


FIGURE 2.4: **Deep mechanistic networks (DMN) at the intersection of deep learning, computational neuroscience, probability, and biological data.** The DMN is a machine learning framework that combines dynamic system formulations rooted in biophysics with task-optimization algorithms from deep learning for training unknown circuit parameters to infer computational mechanisms. The inference leans on methods for approximate Bayesian inference, such as deep ensembles, nonlinear dimensionality reduction, and clustering. DMNs are constrained by biological data of connectome, transcriptome, neural activity and behavior. The interpretation of computation in the DMN can be done side-by-side with experiments in the organism because the connectome establishes a correspondence of model neurons to organism neurons.

Our DMN of the fly visual system follows the dynamical system description derived from the biophysical background (Section 2.1):

$$\tau_{t_i} \dot{V}_i = V_{t_i}^{\text{rest}} - V_i + \sum_j \sigma_{t_i t_j} \alpha_{t_i t_j} N_{t_i t_j, u_i - u_j, v_i - v_j} f(V_j) + e_i. \tag{2.2}$$

This combines convolutional weight sharing and sparse connectivity (Section 2.2.2) in a recurrent neural network (Section 2.2.3) with continous-time modeling (Section 2.2.4) of graded potential neurons, like the ones in the fly motion pathways (Section 2.3.3).

We assign parameters τ_{t_i} and $V_{t_i}^{\text{rest}}$ to each neuron of a neuron type t_i in the model for each neuron type identified in the fly connectome (Section 2.3.2 and 3.2.1). The parameters are unknown but required to simulate the system dynamics.

We used the connectome to constrain the weights

$$w_{ij} = \sigma_{t_i t_j} \alpha_{t_i t_j} N_{t_i t_j, u_i - u_j, v_i - v_j} \tag{2.3}$$

of the model, assuming synaptic weight is proportional to synapse count *(Liu et al., 2022)* (Section 3.5.1). Three components determine each weight w_{ij} from presynaptic neuron j (of type t_j) to postsynaptic neuron i (of type t_i):

- *sign* $\sigma_{t_i t_j} \in \{-1, +1\}$ determined by transcriptomic profiling of neurotransmitters and receptors,
- *connectivity and average synapse count* $N_{t_i t_j, \Delta u, \Delta v}$ between neuron types at retinotopic offset $(\Delta u, \Delta v) = (u_i - u_j, v_i - v_j)$ from EM reconstructions, and
- *scaling* $\alpha_{t_i t_j}$, trainable unitary strength factor (in $V \cdot s$/vesicle, Section 2.1.2).

Connections between neurons of the same types share parameters, implementing the crystalline structure of the fly visual system as a convolutional architecture (Section 2.2.2). Our choice to use one scaling factor $\alpha_{t_i t_j}$ per neuron-type pair balances connectome constraint with optimization flexibility. Stronger constraints (such as a single global α or per-neuron-type α_{t_i}) would provide less flexibility to the model to solve the task, while weaker constraints (such as per-connection α_{ij}) would discard the connectome structure.

We compile a consensus circuit model spanning 721 columns from the average filters, corresponding to the count of 700-900 ommatidia typically found in the fly eye *(Götz, 1964)* and the connectivity of the visual system's central columns. The resulting model comprises 45,669 neurons and 1,513,231 neuron-to-neuron connections, belonging to 64 identified neuron types (Section 2.3.2).

We trained the model using backpropagation through time (Section 2.2.5 and 3.5.5) on the computer vision task Sintel for computing optic flow in naturalistic scenes *(Butler et al., 2012)* (Sections 2.2.6 and 3.5.5). Because of the non-identifiability of solutions, we train deep ensembles of equivalent models and analyze them jointly (Section 2.2.7). The model allows us to compare simulated voltage and tuning properties at single neuron resolution (Sections 2.3.3) to existing experimental measurements, an advancement over creating representation level correspondences only (Section 2.2.6). We characterize all our comparisons statistically across a deep ensemble of 50 task-optimized models (Section 2.2.7 and 3.2.2).

We generated 478 pages of dynamical system characterization for single neurons and different input stimuli across the whole ensemble as a tool for mechanistic discovery with the model (Section 7.1). We publish the software and pretrained models with tutorials as tools for computational experiments, modeling, and generating hypotheses to inform experiments [3].

While the present model does not account for all biological details, as for instance information of gap junctions *(Ammer et al., 2015)* or neuromodulation *(Strother et al., 2018)* is still missing, the here developed DMN framework can be applied to end-to-end modeling of connectome and task-constrained neural networks in general. It is for instance already successfully applied in modeling of other neural system connectomes and tasks *(Özdil et al., 2024)*, to different neuron models with more realistic synaptic dynamics (Section 4.2), or in combination with mechanical simulation of whole-bodies *(Wang-Chen et al., 2024; Vaxenburg et al., 2025)*.

---

[3] https://github.com/TuragaLab/flyvis

3

CONNECTOME-CONSTRAINED NETWORKS PREDICT NEURAL ACTIVITY ACROSS THE FLY VISUAL SYSTEM

Paper: Lappalainen, J. K., Tschopp, F. D., Prakhya, S., McGill, M., Nern, A., Shinomiya, K., Takemura, S.-Y., Gruntman, E., Macke, J. H., & Turaga, S. C. (2024). Connectome-constrained networks predict neural activity across the fly visual system. *Nature*, *634*(8036), 1132

CONTRIBUTION STATEMENT (§6.2 EXAMINATION REGULATIONS)

Note: The following statement signed by J.H.M. and approved by S.C.T. was handed in alongside the thesis per requirements of examination regulations §6(2).

Author Contributions Table

Author	Position	Scientific Ideas %	Data Generation %	Analysis & Interpretation %	Paper Writing %
J.K.L.	1	40	35	60	50
F.D.T.	2	10	35	5	0
S.P.	3	5	5	5	0
M.M.	4	10	5	0	0
A.N.	5	0	5	0	0
K.S.	6	0	5	0	0
S.-y.T.	7	0	5	0	0
E.G.	8	0	0	5	0
J.H.M.	9	15	0	10	25
S.C.T.	10	20	5	15	25

Summary of Individual Contributions

Primary contributions of J.K.L.:

1. *Conceptualization & Methodology (with F.D.T., M.M., J.H.M., S.C.T.):* Co-conceived the project framework, developed the neuron/synapse model, training and analysis pipeline.
2. *Software & Implementation:* Led the development of the computational framework, including the PyTorch model implementation, training infrastructure, and analysis codebase[1].
3. *Model Training & Evaluation:* Trained and evaluated all model ensembles, optimization of hyperparameters, validation of model performance, and selection of models for detailed analysis.
4. *Investigation & Analysis:*
    a. Conducted all computational experiments for Figures 1-4 and Extended Data Figures 1-8
    b. Developed model clustering and mechanistic analyses
    c. Synthesized experimental measurements from literature for model validation
    d. Identified TmY3 as a candidate neuron type for motion computation
5. *Data Curation:* Revised the connectome manually with feedback from A.N. and S.C.T.
6. *Visualization:* Created main figures (1-4) and supplementary visualizations (ED Fig. 1-8; Supp Fig. 2-7)
7. *Writing:* Wrote the first draft and co-wrote the final manuscript with J.H.M. and S.C.T.

Contributions of co-authors:

- *S.C.T.:* Supervised the project as lead senior author, co-conceived the framework, provided funding, provided critical feedback on analyses and interpretation, and co-wrote the manuscript
- *J.H.M.:* Supervised the project, co-conceived methodology, provided funding, provided critical feedback on analyses and interpretation, and co-wrote the manuscript

---

[1] https://github.com/TuragaLab/flyvis

- *F.D.T.:* Built initial connectome from local reconstructions, contributed to conceptualization, created Supp Fig. 1 and 9
- *M.M.:* Co-built the hexagonal lattice PyTorch model
- *S.P.:* Developed synthetic connectome framework methods (Figure 5; Supp Fig. 8)
- *Others:* Provided feedback on connectome reconstruction, comparisons between model and experimental findings, and manuscript revisions

Significance of J.K.L.'s contributions: J.K.L.'s work was central to building a functional computational model from biological connectome data, developing the analysis framework to validate and understand model behavior, and synthesizing the findings into a coherent scientific narrative. He carried out (virtually) all training runs, simulations and their analyses in the paper, with the exception of the synthetic connectome analyses (Fig 5.). The computational infrastructure developed by J.K.L. enabled all model-based discoveries in the paper.

Date: ____________________

Supervisor signature: ______________________________

Connectome-constrained networks predict neural activity across the fly visual system

https://doi.org/10.1038/s41586-024-07939-3
Received: 16 March 2023
Accepted: 9 August 2024

Published online: 11 September 2024
Open access
Check for updates

Janne K. Lappalainen[1,2,3], Fabian D. Tschopp[3], Sridhama Prakhya[3], Mason McGill[3,4], Aljoscha Nern[3], Kazunori Shinomiya[3], Shin-ya Takemura[3], Eyal Gruntman[3,5], Jakob H. Macke[1,2,6] & Srinivas C. Turaga[3]✉

We can now measure the connectivity of every neuron in a neural circuit[1–9], but we cannot measure other biological details, including the dynamical characteristics of each neuron. The degree to which measurements of connectivity alone can inform the understanding of neural computation is an open question[10]. Here we show that with experimental measurements of only the connectivity of a biological neural network, we can predict the neural activity underlying a specified neural computation. We constructed a model neural network with the experimentally determined connectivity for 64 cell types in the motion pathways of the fruit fly optic lobe[1–5] but with unknown parameters for the single-neuron and single-synapse properties. We then optimized the values of these unknown parameters using techniques from deep learning[11], to allow the model network to detect visual motion[12]. Our mechanistic model makes detailed, experimentally testable predictions for each neuron in the connectome. We found that model predictions agreed with experimental measurements of neural activity across 26 studies. Our work demonstrates a strategy for generating detailed hypotheses about the mechanisms of neural circuit function from connectivity measurements. We show that this strategy is more likely to be successful when neurons are sparsely connected—a universally observed feature of biological neural networks across species and brain regions.

Electrical signals propagating through networks of neurons form the basis of computations such as visual motion detection. The propagation of neural activity is shaped by both the functional properties of individual neurons and their synaptic connectivity. Additional factors[10,13], including electrical synapses, neuromodulation and glia, are known to further influence neural activity on multiple timescales. Volume electron microscopy can now be used to comprehensively measure the connectivity of each neuron in a neural circuit, and even entire nervous systems[1–9]. However, we do not yet have the means to also comprehensively measure all other biological details, including the dynamical properties of every neuron and synapse in the same circuit[13]. For these reasons, there has been considerable debate about the utility of connectome measurements for understanding brain function[14]. It is unclear whether it is possible to use only measurements of connectivity to generate accurate predictions about how the neural circuit functions, especially in the absence of direct measurements of neural activity from a living brain. There is considerable evidence from computer science and neuroscience that there is not necessarily a strong link between the connectivity of a neural network and its computational function. Universal function approximation theorems for artificial neural networks[15] imply that the same computational task can be performed by many different networks with very different neural connectivity. Empirically, there exist many classes of general-purpose artificial neural network architectures that can be trained to perform the same computational task[11]. Such differences in connectivity can correspond to qualitatively different computational mechanisms[16]. Similarly, in neuroscience there have been competing proposals for the same computation (for instance, the computation of visual motion)[17,18]. Furthermore, even circuits with the same connectivity can function differently[19]. Thus, neither the connectivity of a circuit alone, nor its computational task alone, can uniquely determine the mechanism of circuit function[20].

Here we show that the connectivity of a neural circuit, together with knowledge of its computational task, enables accurate predictions of the role played by individual neurons in the circuit in the computational task. We constructed a differentiable[21] model neural network with a close correspondence to the brain, whose connectivity was given by connectome measurements and with unknown single-neuron and single-synapse parameters. We optimized the unknown parameters of the model using techniques from deep learning[11], to enable the model to accomplish the computational task[22]. We call such models connectome-constrained and task-optimized deep mechanistic networks (DMNs) (Fig. 1a).

We applied this approach to model the motion pathways in the optic lobe of the *Drosophila* visual system. We constructed a DMN with experimentally measured connectivity[1–5], and unknown parameters for the

---

[1]Machine Learning in Science, Tübingen University, Tübingen, Germany. [2]Tübingen AI Center, Tübingen, Germany. [3]Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA, USA. [4]Computation and Neural Systems, California Institute of Technology, Pasadena, CA, USA. [5]Dept of Biological Sciences, University of Toronto Scarborough, Toronto, Ontario, Canada. [6]Max Planck Institute for Intelligent Systems, Tübingen, Germany. ✉e-mail: turagas@janelia.hhmi.org

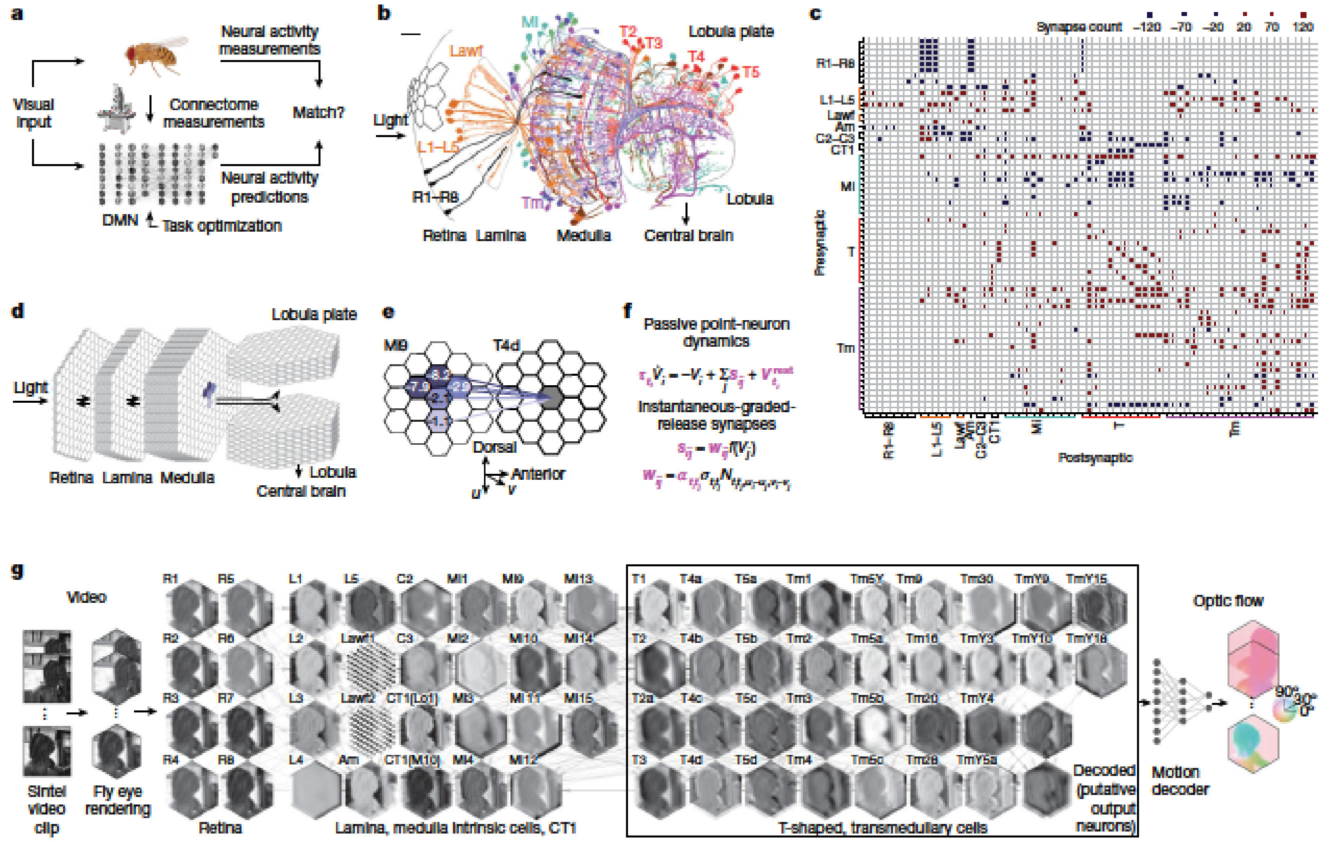


Fig. 1 | Connectome-constrained and task-optimized models of the fly visual system. **a**, DMNs aim to satisfy three constraints: the architecture is based on connectome measurements (**b**–**e**); cellular and synaptic dynamics are given by simple mechanistic models (**f**); and free parameters are task-optimized by training the model to perform optic flow estimation (**g**). Graphics of fruit fly and microscope were created with BioRender.com. **b**, Schematic of optic lobe of *D. melanogaster* with several processing stages (neuropils) and cell types, including photoreceptor (R1–R8), lamina monopolar (L), lamina wide-field (Lawf), medulla intrinsic (Mi), transmedullary (Tm) and T-shaped (T) neurons (adapted from ref. 25, Springer Nature). Scale bar, 10 μm. **c**, Identified connectivity between 64 modelled cell types, represented by total number of synapses from all neurons of a given presynaptic cell type to a postsynaptic cell of a given type. Amacrine (Am), centrifugal (C2–C3) and complex tangential (CT) neurons are also included. Blue (red) colour indicates putative hyperpolarizing (depolarizing) inputs; size of squares indicates number of input synapses. **d**, Retinotopic hexagonal lattice columnar organization of visual system model. Each lattice represents a cell type; each hexagon represents an individual cell. Positions of photoreceptor ommatidia are aligned with downstream columns. The model comprises synapses from all neuropils (Supplementary Fig. 1). **e**, Example of convolutional filter, representing Mi9 inputs onto T4d cells. Values represent the average number of synapses projecting from presynaptic Mi9 cells in columns with indicated offset onto the postsynaptic dendrite of T4d cells. **f**, Single-neuron and synaptic dynamics are given by simple mechanistic models. Free parameters (magenta) are optimized by training the recurrent network model to perform optic flow estimation. **g**, Illustration of DMN performing optic flow estimation. Each hexagonal lattice shows a snapshot of simulated voltage levels of all cells of each type in response to input to the photoreceptors (R1–R8). Edges illustrate connectivity between cell types. A decoder receives the simulated neural activity of all output neurons to compute optic flow. Parameters of DMN and decoder are optimized using deep learning.

single-neuron dynamics and the strength of a unitary synapse. We optimized the model parameters on the computer vision task of detecting motion in dynamic visual stimuli[12]. Visual motion computation in the fly and its mechanistic underpinnings have been extensively studied[23]. Thus, we were able to compare the detailed predictions of our model with experimental measurements of neural activity in response to visual stimuli, on a neuron-by-neuron basis. We found that our connectome-constrained and task-optimized DMN accurately predicts the separation of the visual system into light-increment (ON) and light-decrement (OFF) channels, as well as the generation of direction selectivity in the well-known T4 and T5 motion detector neurons[24]. We release our model as a resource for the community (https://github.com/TuragaLab/flyvis).

DMN of the fly visual system

The optic lobes of the fruit fly are equivalent to the mammalian retina. They comprise several layered neuropils whose columnar arrangement has a one-to-one correspondence with the ommatidia, both possessing a remarkably crystalline organization in a hexagonal lattice. Visual input from the photoreceptors is received by the lamina and medulla, which send projections to the lobula and lobula plate[25] (Fig. 1b). Many components of the optic lobe are highly regular, with columnar cell types appearing once per column, and multicolumnar neurons appearing with only small deviations from a well-defined periodicity in columnar space[25,26]. Several studies have reported on the local connectivity in the optic lobe and its motion pathways[1–5]. We assembled these separate local reconstructions into a coherent local connectome spanning the retina, lamina, medulla, lobula and lobula plate (Fig. 1c, Supplementary Note 1, Supplementary Fig. 1 and Supplementary Data files 1–3).

We approximated the circuitry across the entire visual field as perfectly periodic[2,26], and tiled this local connectivity architecture in a hexagonal lattice across retinotopic space to construct a consensus connectome for 64 cell types across the central visual field of the right eye (Fig. 1d, Methods, Extended Data Fig. 1, Supplementary Fig. 2 and Supplementary Data file 4). By this assumption of translation invariance due to periodic tiling, the synapse count between each pair of neurons was the same across all pairs of neurons with the same presynaptic and postsynaptic cell type and relative location in retinotopic space. For simplicity, we refer here to this partial connectome of the motion pathways as the connectome.

We built a recurrent neural network modelling these first stages of visual processing in the optic lobe based on the connectome for the right eye. Each neuron in this DMN corresponds to a real neuron in the fly visual system, belonging to an identified cell type, and is connected to other neurons only if they are connected by synapses in the connectome (Fig. 1e). We constructed a model with detailed connectivity, but simplified models of single neurons and chemical synapses (Fig. 1f). We used passive leaky linear non-spiking voltage dynamics to model the time-varying activity of single neurons, as many neurons in the early visual system are non-spiking. We modelled neurons with a single electrical compartment, as this has previously been shown to be a good approximation given the small size of many neurons in the optic lobe[27]. The CT1 (complex tangential) neuron, which is among the largest in the brain, spanning the entire optic lobe, was modelled with one compartment per column in the medulla and lobula, as it is highly electrotonically compartmentalized[28] (Supplementary Note 2). We modelled the graded-release chemical synapses between non-spiking neurons with a threshold-linear function to approximate the nonlinear voltage-gated release of neurotransmitters. The resulting network model follows well-known threshold-linear dynamics and is piece-wise differentiable. Such dynamics are typically used to approximate the firing rates of a network of spiking neurons with the nonlinearity arising from spike generation, whereas in our network, the nonlinearity represents the voltage-gated neurotransmitter release. We used the cell-type structure of the connectome to reduce the number of free parameters in the model (Fig. 1f). We assumed that neurons of the same cell type shared the same neuron time constant and resting membrane potential. We modelled synaptic weights as proportional to the discrete number of synapses as reported in the connectome between a connected neuron pair[29], with a scale factor representing the strength of a unitary synapse. The unitary synapse scale factor and the sign of each synapse was the same for all pairs of neurons with the same pre- and postsynaptic cell type. In other words, a connection of five synapses from an Mi1 (medulla intrinsic) neuron to a T4 (T-shaped) neuron is assumed to be exactly half as strong as ten synapses between another pair of neurons of the same presynaptic and postsynaptic cell types, but could be stronger or weaker than five synapses between neurons of a different cell-type pair, for instance, from a Tm3 (transmedullar) neuron to a T4 neuron. The sign of each cell-type connection was determined by neurotransmitter and receptor expression profiling[30] (Methods and Supplementary Data file 2). In total, the connectome-constrained model comprises 45,669 neurons and 1,513,231 connections, across 64 cell types arranged in a hexagonal lattice consisting of 721 columns, modelling the central visual field of the roughly 700–900 ommatidia typically found in the fruit fly retina[31]. Connectome constraints, and our assumption of spatial homogeneity (that is, the hexagonally convolutional structure of the network), result in a marked reduction to just 734 free parameters for this large network model. The only free parameters in our model are the single-neuron time constants and resting membrane potentials (two parameters per cell type), and the unitary synapse strengths (one parameter per type-to-type connection). In the absence of connectome measurements, we would have needed to estimate well in excess of a million parameters corresponding to the weights of all possible connections (Methods).

We used task optimization[22] to further constrain the parameters of the model (that is, by training the model to perform a computational task that is thought to approximate the computations carried out by the circuit). We therefore implemented our recurrent DMN using the PyTorch library[21] (Methods) and used automatic differentiation to optimize the model using gradient-based deep learning training methods[11].

As the computational task constraining the input–output function of the circuit, we chose the computation of visual motion from naturalistic stimuli[12]. Motion computation in the fly visual system and its mechanistic underpinnings have been extensively studied[23]. This computation requires the neural circuit to compare visual stimuli across space and time, and thereby critically relies on temporal integration of visual information by the dynamics of the network. We reasoned that training our model to perform the computer vision task of optic flow computation[12] could help us identify circuit elements involved in motion computation. As our model contains many of the circuit elements that have been experimentally characterized and implicated in the computation of visual motion, we could then validate our model predictions.

To decode optic flow from the DMN, we used a decoding network to map the representation of motion used in the fly nervous system to the representation of optic flow specified by the computer vision task. This two-layer convolutional decoding network is given only the instantaneous neural activity of the medulla and downstream areas as input. Importantly, the decoding network cannot by itself detect motion, which requires the comparison of current and past visual stimuli, but must instead rely on the temporal dynamics of the DMN to compute motion-selective visual features. The resulting combination of our recurrent connectome-constrained DMN model and the feedforward decoding network was then trained end-to-end: we rendered video sequences from the Sintel database[12] as direct input to the photoreceptors of the connectome-constrained model, and used gradient descent (backpropagation through time[11]) to minimize the task error in predicting optic flow (Fig. 1g and Methods).

DMN ensemble predicts known activity

We used only connectome and task constraints to construct our DMN, without any measurements of neural activity. We can therefore validate the model by comparing predictions of neural activity for each of the 64 identified cell types to experimental measurements. As it is possible that these constraints might not uniquely constrain model parameters[32], we generated an ensemble of 50 models, all constrained with the same connectome, and optimized to perform the same task. Each model in the ensemble corresponds to a local optimum of task performance. As the models achieved similar (but not identical) task performance, the ensemble reflects the diversity of possible models consistent with these constraints. The ensemble of models found a variety of parameter configurations (Supplementary Fig. 3), and exhibited superior task performance to both the decoder network alone and models with random parameter configurations (Extended Data Fig. 2a). We focused on the ten models that achieved the best task performance (Fig. 2a). We simulated neural responses to multiple experimentally characterized visual stimuli, and comprehensively compared model responses for each cell type to experimentally reported responses from 26 previously reported studies (Supplementary Note 3 and Supplementary Data files 5 and 6).

First, neural responses in the fly visual system are known to segregate into ON- and OFF-channels[33] defined by whether a neuron depolarizes more strongly to an increase or decrease in stimulus intensity, respectively, a hallmark of visual computation across species[34,35]. We probed the contrast preference of each cell type using flash stimuli[36] and found that the ensemble predicts the segregation into ON- and OFF-pathways with high accuracy: the median flash response index (FRI) across the ensemble predicts the correct ON- and OFF-preferred contrast selectivity for all 32 cell types for which contrast selectivity has been experimentally established (terminals of CT1 in the medulla and lobula are listed as CT1(M10) and CT1(Lo1)). This is also the case for the model with the best task performance (task-optimal model), which correctly predicts the preferred contrast of 30 of 32 cells (Fig. 2b). Furthermore, the ensemble provides predictions for the remaining

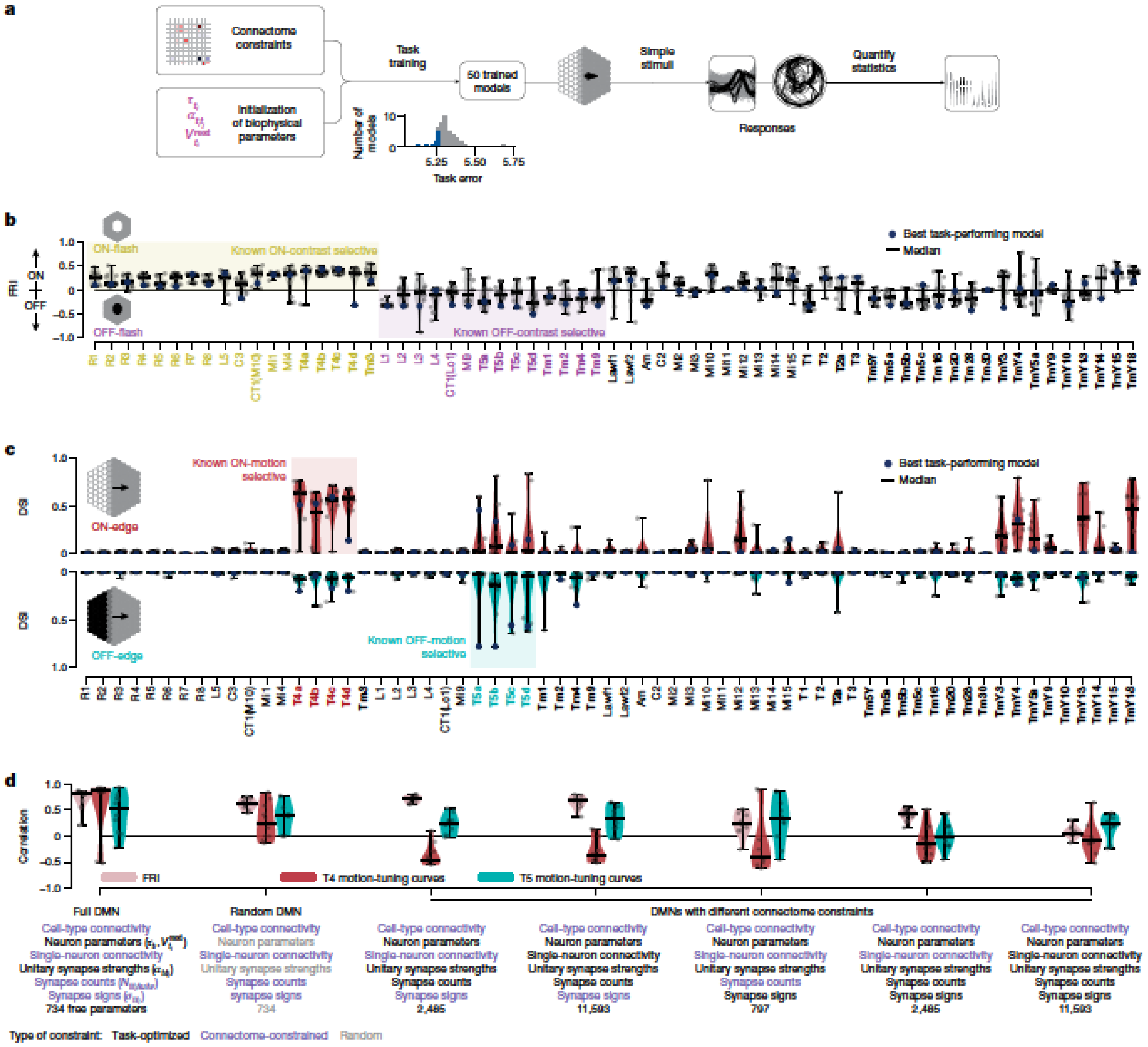


Fig. 2 | Ensembles of DMNs predict tuning properties. a, We task-optimized 50 connectome-constrained DMNs, yielding different solutions for the biophysical parameters (magenta), and compared the tuning properties of their cell types to experimental measurements. Inset: distribution of task errors. Blue: ten best models; also shown in **b**–**d**. **b**, ON- and OFF-contrast selectivity indices (FRI) for each cell type from ten models with best task performance (ten worst models in Extended Data Fig. 6). Yellow: cell types known to be ON-selective. Violet: known OFF-selective types. Black: selectivity not yet established experimentally. Bold: inputs to optic flow decoder. **c**, Direction selectivity index (DSI) from neural responses to moving edges; same ten models as above. **d**, Correlations between measurements and predictions of neural activity from seven types of DMN with different connectome and parameter constraints: FRI (pink); motion-tuning curves for T4 (red) and T5 (turquoise); dashes indicate median correlation across models. The first DMN type on the left corresponds to the main DMNs analysed in **b**,**c** and all subsequent figures; the remaining six DMNs incorporate fewer constraints. Ten best task-performing models from each DMN type.

33 cell types, and consistency across the ensemble provides a measure of confidence in the predictions (Fig. 2b).

Second, a major result in fly visual neuroscience has been the identification of the T4 (ON) and T5 (OFF) neurons as the first direction-selective neurons with four subtypes (T4a, T4b, T4c and T4d, and T5a, T5b, T5c and T5d), each responding to motion in the four cardinal directions[37]. We characterized the motion selectivity of all 64 cell types by their responses to ON- and OFF-edges moving in 12 different directions. We found that the ensemble of models correctly predicts that T4 neurons are ON-motion selective, and T5 neurons are OFF-motion selective (Fig. 2c). The ensemble also correctly predicts the lack of motion tuning in the input neurons to T4 and T5 motion detector neurons (Mi1, Tm3, Mi4, Mi9, Tm1, Tm2, Tm4, Tm9 and CT1; Methods and Supplementary Data file 7).

Our models also suggest the possibility that the transmedullary cell types, TmY3, TmY4, TmY5a, TmY13 and TmY18, might be tuned to ON-motion. Of these cell types, TmY3 neurons do not receive inputs from other known motion-selective neurons suggesting the possibility that these neurons constitute a parallel motion computation pathway from T4 and T5 neurons (Extended Data Figs. 3 and 4 and Supplementary Note 4). We questioned whether our model predicted motion selectivity for all cell types with asymmetric, multicolumnar inputs, as this is a necessary connectivity motif for direction selectivity. On the basis of their local spatial connectivity profiles, we estimated that 19 cell types receive asymmetric, multicolumnar inputs (Methods and Extended Data Fig. 5b), but found that only 12 are predicted to be motion selective by the ensemble (Methods). Spatial offset of excitatory and/or inhibitory inputs did not correlate strongly with direction selectivity (Extended Data Fig. 5c,d). This suggests that our model integrates connectivity across the entire network with the task constraint to determine which neurons are most likely to be motion selective, rather than simply focusing on local connectivity.

Connectome and task are both necessary

We investigated the importance of connectome constraints and task optimization to enable accurate predictions of neural activity. We found that both task optimization and detailed connectome constraints at the single-neuron resolution were critical to the prediction of the preferred contrast of the 32 characterized cell types, and the preferred direction of motion for the T4 and T5 subtypes (Fig. 2d and Extended Data Fig. 2).

We conducted ‘ablation’ studies comparing the DMN ensemble studied in this paper (full DMN) with a range of models with other modelling assumptions. First we verified the importance of task optimization by constructing an ensemble (random DMN) with random single-cell and synapse parameters, and full connectome constraints. This ensemble yielded accurate predictions of preferred contrast, but poor predictions of direction selectivity and preferred direction.

We then studied which aspects of the connectome must be measured accurately to lead to accurate predictions. We found that task-optimized models with access to only cell-type connectivity predicted neural activity poorly. We then considered several scenarios adding more connectome measurements beyond cell-type connectivity. We considered several scenarios: access to synapse signs and single-neuron connectivity but not strength, requiring task optimization of synapse counts; access to signs, but requiring optimization of single-neuron connectivity and synapse counts; full connectome measurements but optimization of synapse signs; access to connectivity but optimization of both synapse counts and signs. Across these modelling assumptions, we found that accurate predictions of contrast preference (FRI) were possible as long as measurements of the connection-signs were available, and that accurate predictions of the direction selectivity—but not preferred direction—could be achieved with measurements of cell connectivity, without the need for synapse count measurements (Extended Data Fig. 2c). This demonstrates the importance of both detailed connectome measurements and task optimization to achieve accurate predictions of neural activity.

Across the ensemble constrained by all connectome measurements and task training, we found that models that exhibited lower task error (Methods) also had more realistic tuning: models with higher task performance predict the direction selectivity index of T4 and T5 cells and their inputs better ($r = 0.60$, $P = 2.6 \times 10^{\ 6}$; Extended Data Figs. 2e and 6b). This suggests the possibility of using task error to rank models in terms of their likelihood to accurately predict neural activity.

Our model relies on an accurate classification of neurons into cell types to share single-neuron and synapse parameters across all neurons of the same cell type. We investigated the degree to which we could coarse-grain the cell-type categorization, leading to fewer cell types and fewer parameters (Extended Data Fig. 2a–d). We found that grouping the four T4 subtypes into a single T4 cell type, and grouping the four T5 subtypes into a single T5 cell type, had no negative impact on the quality of ensemble predictions. However, grouping all 37 excitatory cell types into a single E-type, 22 inhibitory cell types into a single I-type, and 4 mixed cell types into a single mixed type, led to poor performance on par with the random DMN.

Predictions cluster across DMN ensemble

We sought to determine how similar or dissimilar the predictions of different models in an ensemble with the same connectome constraints and task optimization are. To address this for each cell type, we simulated neural activity in response to naturalistic video sequences from the Sintel dataset. We then used uniform manifold approximation and projection[38] to perform nonlinear dimensionality reduction on high-dimensional activity vectors of a representative neuron of each cell type across the model ensemble, and clustered the models in the resulting two-dimensional projections (Fig. 3a and Methods). For many cell types, we found that models predict strongly clustered neural responses (Supplementary Data file 7). For T4c neurons, for example, we found three clusters corresponding to qualitatively distinct responses of this cell type for naturalistic inputs: two clusters contain models with direction-selective T4c cells (Fig. 3a,b) with up- and down-selective cardinal tuning, respectively, whereas neurons in the third cluster are not direction-tuned. The direction-selective cluster with the (correct) upward preference has the lowest average task error (circular marker, average task error 5.297), followed by the cluster with the opposite preference (triangular marker, average task error 5.316). The non-selective cluster has the worst performance (square marker, average task error 5.357), suggesting that models with accurate tuning correlate with lower task error (see also Extended Data Fig. 2e).

We sought to determine what differences in circuit mechanisms underlie the different predictions for direction selectivity in the three clusters (Fig. 3c). Our results showed that direction selectivity in the two tuned clusters is associated with opposite preferred contrast tuning of Mi4 and Mi9 neurons, which provide direct flanking inhibitory input to T4 neurons (Fig. 3d). Models with the correct direction selectivity for T4 neurons also predict the correct contrast selectivity for Mi4 and M9 neurons, and vice versa (Fig. 3e).

Thus, the ensemble can be used to provide hypotheses about different circuit mechanisms that might underlie the response properties of individual cells. Furthermore, it shows that experimentally measuring the tuning of one neuron automatically translates to constraints on other neurons in the circuit. Here, filtering models in the ensemble with the experimentally measured direction selectivity for the T4c neurons (by only selecting models from the correct cluster) is sufficient to correctly constrain the tuning of both Mi4 and Mi9 neurons.

Predicted mechanism of T4 and T5 tuning

Our DMN modelling approach enables a large number of model-based analyses, which can illuminate the mechanistic basis of computation in a circuit and suggest new visual stimuli for experimental characterization. We illustrate these analyses using averages from the model cluster with the best task performance (task-optimal cluster), focusing on the well-studied T4 and T5 neurons (Fig. 4). See Supplementary Data file 7 for a comprehensive set of analyses for all cell types and models. In the task-optimal cluster, the four subtypes of the T4 neurons respond strongly to bright (ON) edges, and the four subtypes of the T5 neurons to dark (OFF) edges, moving in the four cardinal directions, in agreement with experimental findings[27,37,39,40] (Fig. 4a). We probed the mechanism of direction selectivity in T4 and T5 neurons (Fig. 4b and Extended Data Figs. 7 and 8). Examining the input currents to a single T4 neuron (Extended Data Fig. 7a), we found that fast excitatory input and offset delayed inhibitory input currents enable T4 in the model to detect motion, in agreement with experimental findings[27]. The differential response of T4 neurons to motion in the preferred versus null direction is primarily produced by the differential timing of inhibition from Mi4. Additionally, excitatory T4-to-T4 currents between neurons with the same preferred direction lead to an increased response to coherent motion across the visual field. Although research into T4 motion selectivity has largely focused on the role of feedforward inputs,

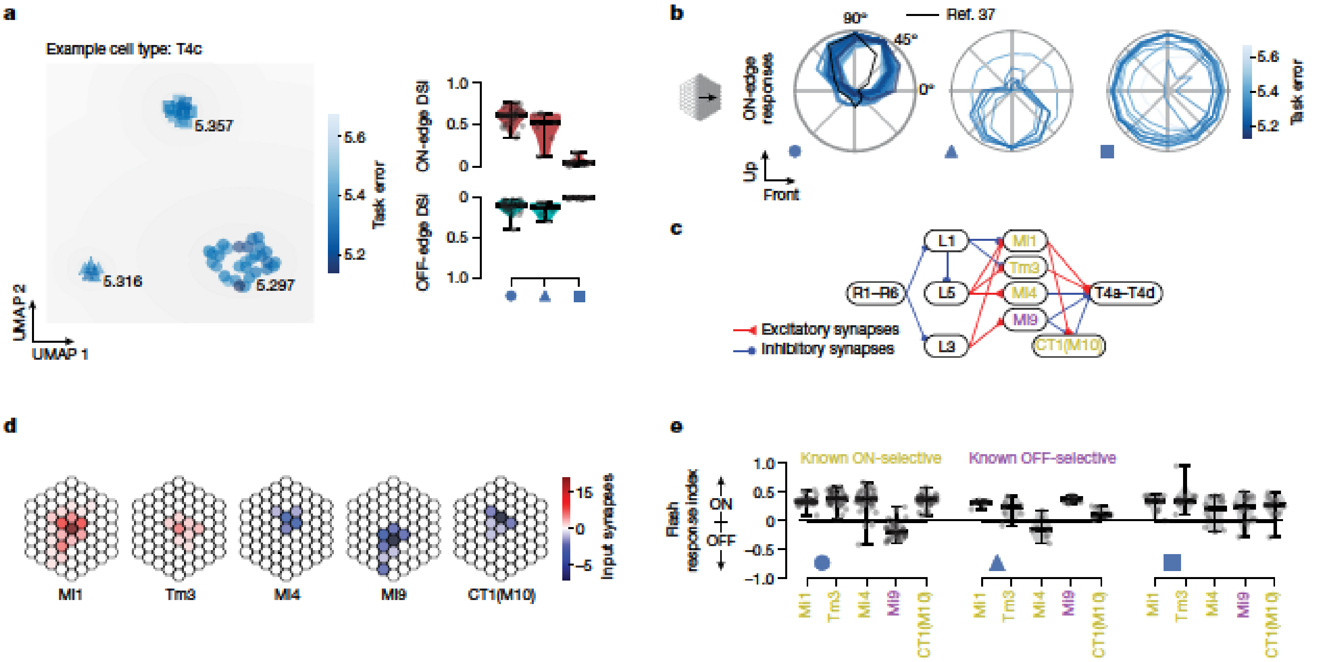


Fig. 3 | Cluster analysis of DMN ensembles enables hypothesis generation and suggests experimental tests. We clustered 50 DMNs after carrying out nonlinear dimensionality reduction of their responses to naturalistic scenes for each cell type, and aimed to identify whether clusters correspond to qualitatively different tuning mechanisms. **a**, Responses of T4c cells exhibit three clusters, two with ON-motion direction selectivity (circular and triangular marker) and one (square marker) without. **b**, T4c tuning in the three clusters. Circular marker: upwards tuning (cluster with lowest average task error 5.297; black: known tuning of T4c). Triangular marker: downwards (5.316 error). Square marker: no motion tuning (5.357 error). **c**, Schematic of corresponding ON-motion detection pathway. **d**, Connectivity of major input elements to T4c. Blue and red colour: putative hyper- and depolarizing inputs. Saturation: average number of input synapses for each offset location. **e**, Tuning properties within each cluster reveal dependencies between T4 tuning and that of Mi4 and Mi9 cells in the ensemble: switching Mi4 (known ON-contrast selective) and Mi9 (known OFF-contrast selective) contrast preferences results in directionally opposite motion-tuning solutions in T4. DMNs in first cluster (T4c in DMN upwards tuned, circle) exhibit ON-selectivity for Mi1, Tm3, Mi4 and CT1(M10), and OFF-selectivity for Mi9. For ON-motion stimuli, in these DMNs T4c receives central depolarizing input from Mi1 and Tm3 and dorsal hyperpolarizing input from Mi4 and CT1(M10).

our modelling predicts an important role for the lateral connectivity between T4 neurons. The mechanisms in our model for T5 motion computation are similar (Extended Data Fig. 8a), with differential timing of inhibition from CT1, as well as excitation from Tm9, contributing to motion-selective responses.

To relate the mechanism of direction selectivity to the well-studied mechanisms of preferred direction enhancement and null direction suppression, we compared the responses of T4 and T5 neurons to moving bars and static bars as in ref. 27. Consistent with voltage measurements[27,40], voltage response predictions by our model show null direction suppression but no preferred direction enhancement (Extended Data Figs. 7b and 8b).

We computed and compared the spatial and temporal receptive fields of the major columnar input neurons to T4 and T5 neurons. These input neurons have been the focus of multiple experimental studies of the motion detection pathways[28,41–45] (Fig. 4d). In agreement with experimental findings[28,43], the prediction of the DMNs is that Tm3 and Tm4 have broad spatial receptive fields (two-column radius, 11.6°), whereas Mi1, Mi4, Mi9, Tm1, Tm2, Tm9 and CT1 compartments in both medulla and lobula have narrow spatial receptive fields (single-column radius, 5.8°).

We characterized the temporal response properties of cells in the motion pathways, including the lamina monopolar cells (L1–L5) and direct inputs to the T4 and T5 neurons. We simulated neural responses to single-ommatidium flashes of varying contrast and duration and compared them to empirically characterized temporal responses (Fig. 4e). The model accurately predicts the preferred contrast of each cell type[33] (that is, whether they depolarize more strongly to ON or OFF single-ommatidium flashes; 5 ms to 300 ms duration; Methods). These cells either depolarize (which we call ON-selective) or hyperpolarize (which we call OFF-selective) in response to light-increment flashes. The temporal response properties are correctly predicted for all except the Tm4 cell in this model: for major T4 inputs, Mi1, Tm3 and Mi4 respond with transient depolarization to ON-flashes. By contrast, CT1(M10) responds with a longer sustained depolarization. Mi9 hyperpolarizes. For major T5 inputs, Tm1, Tm2, Tm9 and CT1(Lo1) respond with transient hyperpolarization. Tm4 is incorrectly predicted to depolarize. For lamina cell types, the DMNs predict biphasic hyperpolarization in L1 and L2 and monophasic hyperpolarization in L3 and L4, as well as depolarization in L5.

For motion-selective neurons such as T4 and T5, the spatio-temporal receptive fields are not separable in space and time. We characterized the full spatio-temporal receptive field for T4c and T5c neurons (Fig. 4f) using single-ommatidium ON- and OFF-flashes (20 ms; Methods). ON-flashes on the leading side of the receptive field of the ON-contrast, upwards-direction-selective T4c cell lead to fast depolarization, whereas ON-flashes on the trailing side lead to delayed hyperpolarization, again matching experimental findings[27]. As T5c is OFF-selective, its OFF-impulse responses are inverted, resembling the T4c spatio-temporal receptive field (Extended Data Fig. 9a). This reflects that T5c implements a similar motion-tuning mechanism to OFF-edges as T4c to ON-edges, in agreement with experimental findings[40].

Finally, we show that the model can be used to design optimized stimuli. We used the task-optimal model to screen for video sequences from the Sintel dataset that elicited the largest responses in the

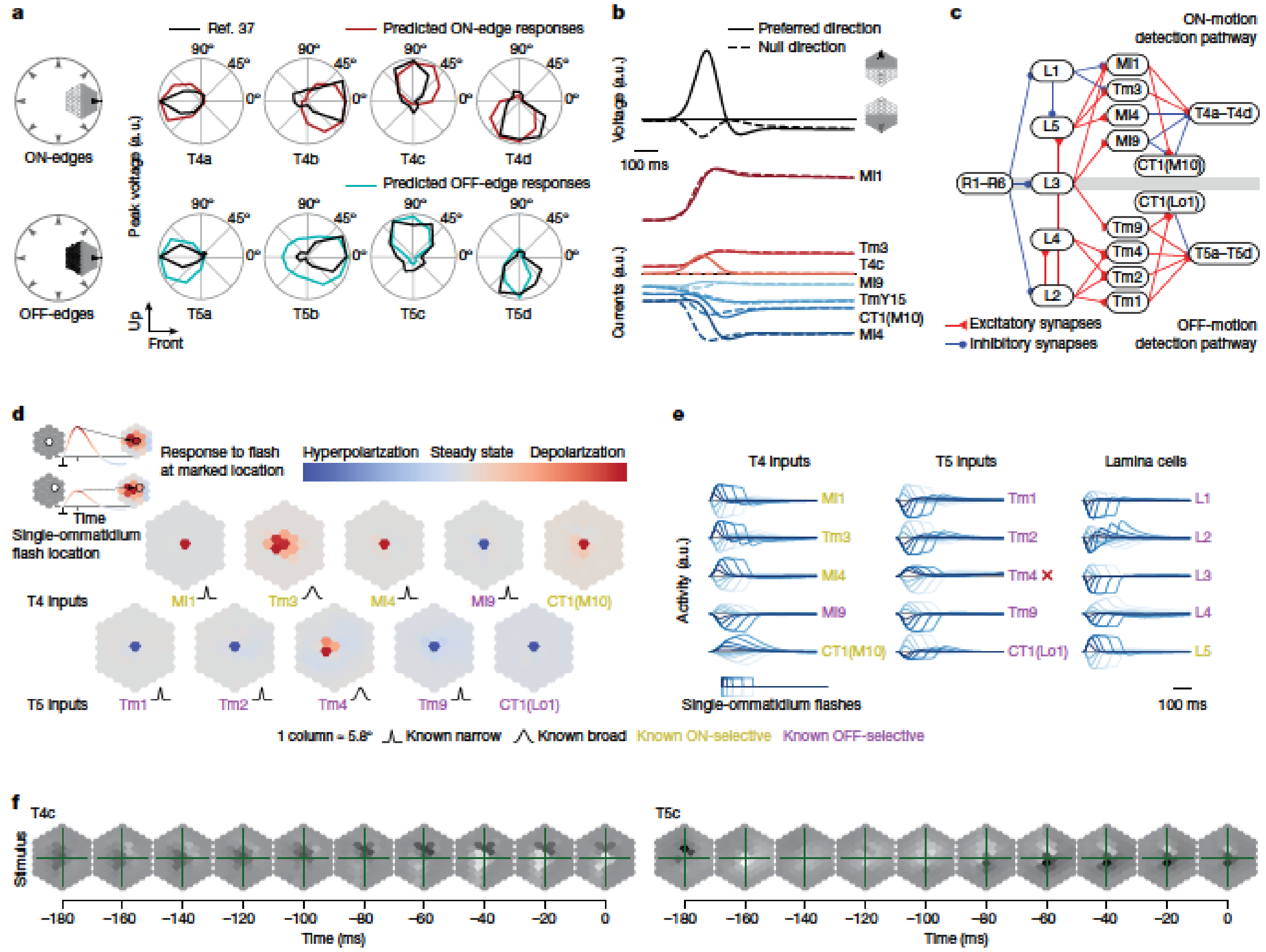


Fig. 4 | Task-optimal DMNs largely recapitulate known mechanisms of motion computation. **a**, Responses to moving edges for T4 and T5 subtypes from task-optimal model clusters, and comparison with experimental measurements[37,39] (null-contrasts in Supplementary Fig. 4). a.u., arbitrary units. **b**, Voltage of T4c neuron (top) and contributions from major input cells (bottom) during movement of an ON-edge across the visual field in preferred (solid) and null (dashed) direction. **c**, Major cell types and connectivity in the ON-motion (T4) and OFF-motion (T5) detection pathways (simplified). **d**, Spatial receptive fields of major motion detector input neurons revealed by single-ommatidium flashes and comparison with experimental measurements[28,43]. **e**, Single-ommatidium flash responses agree with experimental measurements[41,43], with the exception of Tm4 (red cross). **f**, Stimulus sequence predicted to elicit the strongest responses in T4c and T5c cells. A central OFF-disc followed by an ON-edge moving upwards elicits the strongest response in a T4c cell (ON-disc followed by OFF-edge for T5c).

motion-selective neurons (Fig. 4f, Methods and Supplementary Data file 7 for all cell types). One might expect that pure ON- or OFF-stimuli would elicit the largest responses in T4 and T5, respectively. However, we found both ON- and OFF-elements in optimized stimuli, suggesting an interplay between ON- and OFF-pathways. The stimulus that elicited the strongest response in the T4c cell was a central OFF-disc followed by an ON-edge moving upwards, matching the preferred direction of the cell. Similarly, for the T5c cell, the stimulus that elicits the strongest response is a central ON-disc followed by an OFF-edge moving upwards in the preferred direction of the cell (Extended Data Fig. 9c for corresponding full-field naturalistic stimuli, numerically optimized stimuli and preferred moving-edge stimuli). Taken together, these findings show that the model predicts a large number of tuning properties for the T4 and T5 cells and their inputs.

Sparsity leads to accurate predictions

Here we consider when connectome-constrained and task-optimized DMN models might accurately predict neural responses at single-neuron resolution. Sparse connectivity is a hallmark of biological neural circuits. We questioned whether sparse connectivity enables DMNs to make accurate predictions of neural activity. For sparsely connected circuits—assuming the connectome is known—there are fewer synapse parameters left to estimate using task optimization. We reasoned that such networks might support fewer possible mechanisms by which to perform a given task, and so that a task-optimized DMN model is more likely to find the true mechanism and accurately predict single-neuron activity.

We tested this hypothesis in a simulation (Fig. 5), by constructing feedforward artificial neural networks solving the classic MNIST (Modified National Institute of Standards and Technology) handwritten digit classification task. These networks had varying degrees of sparse connectivity, and random assignment of neurons as excitatory and inhibitory respecting Dale's law (25 ground-truth networks for each sparsity level, Methods). We then simulated the process of making connectome measurements from these ground-truth networks, and building connectome-constrained task-optimized DMN simulations of each ground-truth network (Fig. 5a).

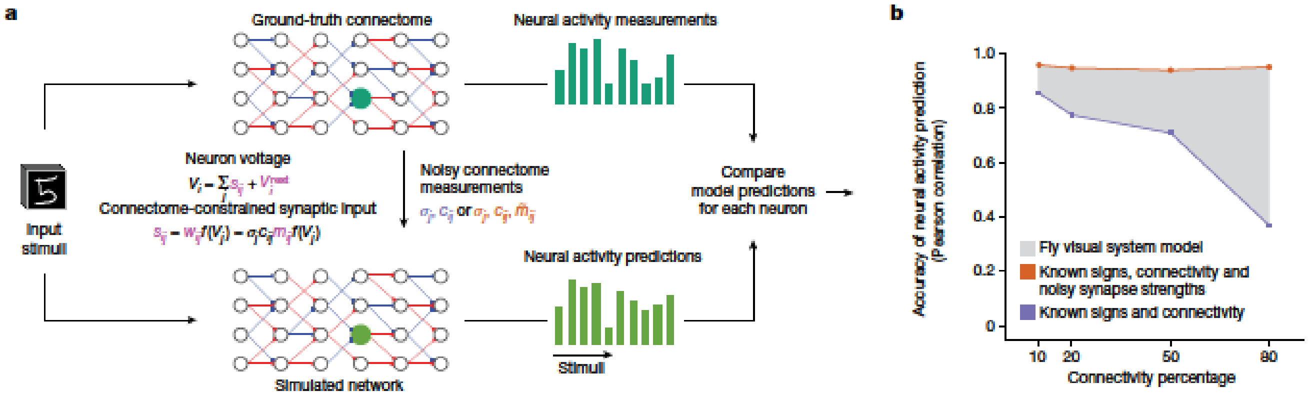


Fig. 5 | Connectome measurements constrain neural networks in circuits with sparse connectivity. a, We constructed synthetic 'ground-truth connectome' networks with varying degrees of sparse connectivity for classifying handwritten digits. For each ground-truth connectome network, we simulated connectome measurements and constructed a connectome-constrained and task-optimized the unknown parameters (magenta) in the 'simulated network' (Methods). We measured the correlation of the neural response vector, across all stimuli, between a ground-truth (dark green) and a simulated (light green) network. **b**, Median neural response correlation coefficients from 100 randomly sampled neuron pairs from each layer and across 25 network pairs. Two conditions were considered in which connectome measurements revealed either only binary connectivity (blue) or also connection strength (orange). The fly visual system model presented here probably falls in the region between the two curves, as measured synapse counts inform relative connection strengths between pairs of neurons for the same pair of cell types, but not absolute connection strength.

As the degree to which connection strength can be inferred from noisy connectome measurements is still unknown, we simulated two settings. In the first, we assumed that connectome measurements reveal connectivity but not connection strength. In this setting, DMNs were task-optimized to infer both the resting membrane potential of each neuron, and the connection strength of each connected pair of neurons. In the second, we assumed that the measurements additionally reveal a noisy estimate of strength, which was used as a soft constraint during task optimization.

Consistent with our hypothesis, our results showed that sparsity in the connectome greatly improves the accuracy of neural activity predictions with measurements of connectivity alone (Fig. 5b and Supplementary Fig. 8; median Pearson correlations of 0.85 for 10% connectivity versus 0.38 for 80% connectivity, 100 randomly selected neurons from 25 randomly generated ground-truth networks). However, with the additional availability of connection strength estimates, we find that DMN simulations accurately predict neural activity even in the absence of sparse connectivity (median Pearson correlation of >0.9 across all connectivities).

Our model of the fly visual system lies in an intermediate regime with regards to our knowledge of connection strength. We assumed that connectome measurements provided relative connection strength but not absolute connection strength, as we assumed that the unitary synaptic strength was unknown but the same for connections of the same cell-type pair. Thus, we attribute the success of our visual system model at predicting neural activity to both the sparse structure of connectivity in this circuit and also the estimates of connection strength from the synapse count.

Discussion

We constructed a neural network with connectivity measured at the microscopic scale. We also required that, at the macroscopic scale, the collective neural activity dynamics across the entire network result in an ethologically relevant computation. This combination of microscopic and macroscopic constraints enabled us to constrain a large-scale computational model spanning many tens of cell types and tens of thousands of neurons. We showed that such large-scale mechanistic models could accurately make detailed predictions of the neural responses of individual neurons to dynamic visual stimuli, revealing the mechanisms by which computations are performed. Knowledge of the connectome played a critical role in this success, in part by leading to a massive reduction in the number of free model parameters.

We have taken a reductionist modelling approach, simplifying the modelling of individual neurons and synapses, to focus on the role played by the connectivity of a neural network. We found that for the motion pathways of the fruit fly visual system, this model correctly predicts many aspects of visual selectivity. We considered only the role of this circuit in detecting motion, which is but one of many computations performed by the visual system[24]. Our reductionist model cannot, for example, account for the role played in this circuit by electrical synapses[46], nonlinear chemical synapses[47] and neuromodulation[48]. However, richer models of neurons, synapses, plasticity and extrasynaptic modulation, along with a broader range of ethologically relevant tasks, can enable accurate modelling of these and other effects in the fly visual system and beyond.

Task-optimized artificial neural network models, for instance of mammalian visual pathways[22], have previously demonstrated only a coarse correspondence of the population neural activity between model layers and brain regions. By contrast, every neuron and synapse in our connectome-constrained model[49–53] has a direct correspondence to neurons and synapses in the brain. This correspondence enables highly detailed experimentally testable predictions at the single-neuron resolution. Thus, our study more directly links artificial neural network models to the biological neural network.

Our modelling approach provides a discovery tool, aimed at using connectome measurements to generate detailed, experimentally testable hypotheses for the computational role of individual neurons. Measurements of neural activity are necessarily sparse and involve difficult trade-offs. Activity can be measured in limited contexts, and either for a limited number of neurons or for a larger number of neurons with poorer temporal resolution. Connectome-constrained DMN models generate meaningful predictions even in the complete absence of neural activity measurements, but can be further constrained by sparse measurements of neural activity as we showed (Fig. 3), or indeed directly fitted to measured neural activity[51] and behaviour[54].

Whole-brain connectome projects have just been completed for the larval and adult fruit fly[7,8,55,56], including two new connectomes of the entire fruit fly optic lobe[6,9], and whole-mouse-brain connectome projects are now being discussed[57]. Large-scale whole-nervous-system models[51,52,58] will be of critical importance for integrating connectomic, transcriptomic, neural activity and animal behaviour measurements across laboratories, scales and the nervous system[13]. Furthermore, with the recent development of detailed biomechanical body models for the fruit fly[59,60], we can now contemplate constructing whole-animal models spanning brain and body.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-024-07939-3.

1. Rivera-Alba, M. et al. Wiring economy and volume exclusion determine neuronal placement in the *Drosophila* brain. *Curr. Biol.* **21**, 2000–2005 (2011).
2. Takemura, S. et al. Synaptic circuits and their variations within different columns in the visual system of *Drosophila*. *Proc. Natl Acad. Sci. USA* **112**, 13711–13716 (2015).
3. Takemura, S. et al. The comprehensive connectome of a neural substrate for 'ON' motion detection in *Drosophila*. *eLife* **6**, e24394 (2017).
4. Shinomiya, K. et al. Comparisons between the ON- and OFF-edge motion pathways in the *Drosophila* brain. *eLife* **8**, e40025 (2019).
5. Shinomiya, K., Nern, A., Meinertzhagen, I. A., Plaza, S. M. & Reiser, M. B. Neuronal circuits integrating visual motion information in *Drosophila melanogaster*. *Curr. Biol.* **32**, 3529–3544 (2022).
6. Matsliah, A. et al. Neuronal "parts list" and wiring diagram for a visual system. Preprint at *bioRxiv* https://doi.org/10.1101/2023.10.12.562119 (2023).
7. Dorkenwald, S. et al. Neuronal wiring diagram of an adult brain. Preprint at *bioRxiv* https://doi.org/10.1101/2023.06.27.546656 (2023).
8. Schlegel, P. et al. Whole-brain annotation and multi-connectome cell typing quantifies circuit stereotypy in *Drosophila*. Preprint at *bioRxiv* https://doi.org/10.1101/2023.06.27.546055 (2023).
9. Nern, A. et al. Connectome-driven neural inventory of a complete visual system. Preprint at *bioRxiv* https://doi.org/10.1101/2024.04.16.589741 (2024).
10. Bargmann, C. I. & Marder, E. From the connectome to brain function. *Nat. Methods* **10**, 483–490 (2013).
11. Goodfellow, I., Bengio, Y. & Courville, A. *Deep Learning* (MIT Press, 2016).
12. Butler, D.J., Wulff, J., Stanley, G.B. & Black, M.J. A naturalistic open source movie for optical flow evaluation. In *Computer Vision – ECCV 2012. ECCV 2012. Lecture Notes in Computer Science* Vol. 7577 (eds Fitzgibbon, A. et al.) 611–625 (Springer, 2012); https://doi.org/10.1007/978-3-642-33783-3_44.
13. Scheffer, L. K. & Meinertzhagen, I. A. A connectome is not enough–what is still needed to understand the brain of *Drosophila*? *J. Exp. Biol.* **224**, jeb242740 (2021).
14. Jabr, F. The connectome debate: Is mapping the mind of a worm worth it? *SciAm* https://www.scientificamerican.com/article/c-elegans-connectome/ (2 October 2012).
15. Hornik, K., Stinchcombe, M. & White, H. Multilayer feedforward networks are universal approximators. *Neural Netw.* **2**, 359–366 (1989).
16. Raghu, M., Unterthiner, T., Kornblith, S., Zhang, C. & Dosovitskiy, A. Do vision transformers see like convolutional neural networks? *Adv. Neural Inf. Process. Syst.* **34**, 12116–12128 (2021).
17. Reichardt, W. in *Sensory Communication* (ed. Rosenblith, W. A.) Ch. 17 (MIT Press, 1961).
18. Barlow, H. B. & Levick, W. R. The mechanism of directionally selective units in rabbit's retina. *J. Physiol.* **178**, 477–504 (1965).
19. Marder, E. Neuromodulation of neuronal circuits: back to the future. *Neuron* **76**, 1–11 (2012).
20. Biswas, T. & Fitzgerald, J. E. Geometric framework to predict structure from function in neural networks. *Phys. Rev. Res.* **4**, 023255 (2022).
21. Paszke, A. et al. PyTorch: an imperative style, high-performance deep learning library. In *Proc. 33rd International Conference on Neural Information Processing Systems* (eds Wallach, H. et al.) 8026–8037 (Curran Associates, 2019).
22. Yamins, D. L. K. et al. Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proc. Natl Acad. Sci. USA* **111**, 8619–8624 (2014).
23. Borst, A., Haag, J. & Mauss, A. S. How fly neurons compute the direction of visual motion. *J. Comp. Physiol. A* **206**, 109–124 (2020).
24. Currier, T. A., Pang, M. M. & Clandinin, T. R. Visual processing in the fly, from photoreceptors to behavior. *Genetics* **224**, iyad064 (2023).
25. Fischbach, K. F. & Dittrich, A. P. M. The optic lobe of *Drosophila melanogaster*. I. A Golgi analysis of wild-type structure. *Cell Tissue Res.* **258**, 441–475 (1989).
26. Nern, A., Pfeiffer, B. D. & Rubin, G. M. Optimized tools for multicolor stochastic labeling reveal diverse stereotyped cell arrangements in the fly visual system. *Proc. Natl. Acad. Sci. USA* **112**, E2967–E2976 (2015).
27. Gruntman, E., Romani, S. & Reiser, M. B. Simple integration of fast excitation and offset, delayed inhibition computes directional selectivity in *Drosophila*. *Nat. Neurosci.* **21**, 250–257 (2018).
28. Meier, M. & Borst, A. Extreme compartmentalization in a *Drosophila* amacrine cell. *Curr. Biol.* **29**, 1545–1550 (2019).
29. Liu, T. X., Davoudian, P. A., Lizbinski, K. M. & Jeanne, J. M. Connectomic features underlying diverse synaptic connection strengths and subcellular computation. *Curr. Biol.* **32**, 559–569 (2022).
30. Davis, F. P. et al. A genetic, genomic, and computational resource for exploring neural circuit function. *eLife* **9**, e50901 (2020).
31. Götz, K. G. Optomotorische Untersuchung des visuellen systems einiger Augenmutanten der Fruchtfliege *Drosophila*. *Kybernetik* **2**, 77–92 (1964).
32. Marder, E. & Taylor, A. L. Multiple models to capture the variability in biological neurons and networks. *Nat. Neurosci.* **14**, 133–138 (2011).
33. Joesch, M., Schnell, B., Raghu, S. V., Reiff, D. F. & Borst, A. ON and OFF pathways in *Drosophila* motion vision. *Nature* **468**, 300–304 (2010).
34. Borst, A. & Helmstaedter, M. Common circuit design in fly and mammalian motion vision. *Nat. Neurosci.* **18**, 1067–1076 (2015).
35. Gjorgjieva, J., Sompolinsky, H. & Meister, M. Benefits of pathway splitting in sensory coding. *J. Neurosci.* **34**, 12127–12144 (2014).
36. Strother, J. A., Nern, A. & Reiser, M. B. Direct observation of ON and OFF pathways in the *Drosophila* visual system. *Curr. Biol.* **24**, 976–983 (2014).
37. Maisak, M. S. et al. A directional tuning map of *Drosophila* elementary motion detectors. *Nature* **500**, 212–216 (2013).
38. Becht, E. et al. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* **37**, 38–44 (2019).
39. Fisher, Y. E., Silies, M. & Clandinin, T. R. Orientation selectivity sharpens motion detection in *Drosophila*. *Neuron* **88**, 390–402 (2015).
40. Gruntman, E., Romani, S. & Reiser, M. B. The computation of directional selectivity in the *Drosophila* OFF motion pathway. *eLife* **8**, e50706 (2019).
41. Behnia, R., Clark, D. A., Carter, A. G., Clandinin, T. R. & Desplan, C. Processing properties of ON and OFF pathways for *Drosophila* motion detection. *Nature* **512**, 427–430 (2014).
42. Yang, H. H. et al. Subcellular imaging of voltage and calcium signals reveals neural processing in vivo. *Cell* **166**, 245–257 (2016).
43. Arenz, A., Drews, M. S., Richter, F. G., Ammer, G. & Borst, A. The temporal tuning of the *Drosophila* motion detectors is determined by the dynamics of their input elements. *Curr. Biol.* **27**, 929–944 (2017).
44. Strother, J. A. et al. The emergence of directional selectivity in the visual motion pathway of *Drosophila*. *Neuron* **94**, 168–182 (2017).
45. Ramos-Traslosheros, G. & Silies, M. The physiological basis for contrast opponency in motion computation in *Drosophila*. *Nat. Commun.* **12**, 4987 (2021).
46. Ammer, G., Leonhardt, A., Bahl, A., Dickson, B. J. & Borst, A. Functional specialization of neural input elements to the *Drosophila* ON motion detector. *Curr. Biol.* **25**, 2247–2253 (2015).
47. Groschner, L. N., Malis, J. G., Zuidinga, B. & Borst, A. A biophysical account of multiplication by a single neuron. *Nature* **603**, 119–123 (2022).
48. Strother, J. A. et al. Behavioral state modulates the ON visual motion pathway of *Drosophila*. *Proc. Natl Acad. Sci. USA* **115**, E102–E111 (2018).
49. Tschopp, F. D., Reiser, M. B. & Turaga, S. C. A connectome based hexagonal lattice convolutional network model of the *Drosophila* visual system. Preprint at https://arxiv.org/abs/1806.04793 (2018).
50. Mano, O., Creamer, M. S., Badwan, B. A. & Clark, D. A. Predicting individual neuron responses with anatomically constrained task optimization. *Curr. Biol.* **31**, 4062–4075 (2021).
51. Mi, L. et al. Connectome-constrained latent variable model of whole-brain neural activity. In *Proc. International Conference on Learning Representations* https://openreview.net/forum?id=CJzi3dRlJE- (2022).
52. Shiu, P. K. et al. A leaky integrate-and-fire computational model based on the connectome of the entire adult *Drosophila* brain reveals insights into sensorimotor processing. Preprint at *bioRxiv* https://doi.org/10.1101/2023.05.02.539144 (2023).
53. Beiran, M. & Litwin-Kumar, A. Prediction of neural activity in connectome-constrained recurrent networks. Preprint at *bioRxiv* https://doi.org/10.1101/2024.02.22.581667 (2024).
54. Cowley, B. R. et al. Mapping model units to visual neurons reveals population code for social behaviour. *Nature* **629**, 1100–1108 (2024).
55. Zheng, Z. et al. A complete electron microscopy volume of the brain of adult *Drosophila melanogaster*. *Cell* **174**, 730–743 (2018).
56. Winding, M. et al. The connectome of an insect brain. *Science* **379**, eadd9330 (2023).
57. Abbott, L. F. et al. The mind of a mouse. *Cell* **182**, 1372–1376 (2020).
58. Warrington, A., Spencer, A. & Wood, F. The virtual patch clamp: imputing *C. elegans* membrane potentials from calcium imaging. Preprint at https://arxiv.org/abs/1907.11075 (2019).
59. Vaxenburg, R. et al. Whole-body simulation of realistic fruit fly locomotion with deep reinforcement learning. Preprint at *bioRxiv* https://doi.org/10.1101/2024.03.11.584515 (2024).
60. Lobato-Rios, V. et al. NeuroMechFly, a neuromechanical model of adult *Drosophila melanogaster*. *Nat. Methods* **19**, 620–627 (2022).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2024

Methods

Construction of spatially invariant connectome from local reconstructions

We built a computational model of the fly visual system that is consistent with available connectome data[1–5], has biophysically plausible neural dynamics, and can be computationally trained to solve an ethologically relevant behavioural task, namely the estimation of optic flow. To achieve this, we developed algorithms to blend annotations from two separate datasets by transforming, sanitizing, combining and pruning the raw datasets into a coherent connectome spanning all neuropils of the optic lobe (Supplementary Note 1 and Supplementary Data files 1–3).

The original data stem from focused ion beam scanning electron microscopy datasets from the FlyEM project at Janelia Research Campus. The FIB-25 dataset volume comprises seven medulla columns and the FIB-19 dataset volume comprises the entire optic lobe and, in particular, detailed connectivity information for inputs to both the T4 and T5 pathways[2–4]. The data available to us consisted of 1,801 neurons, 702 neurons from FIB-25 and 1,099 neurons from FIB-19. For about 830 neurons, the visual column was known from hand annotation. These served as reference positions. Of the 830 reference positions, 722 belong to neuron types selected for simulation. None of the T5 cells, whose directional selectivity we aimed to elucidate, was annotated. We therefore built an automated, probabilistic expectation maximization algorithm that takes synaptic connection statistics, projected synapse centre-of-mass clusters and existing column annotations into account. We verified the quality of our reconstruction as described in Supplementary Note 1. Only the neurons consistently annotated with both 100% and 90% of reference positions used were counted to estimate the number of synapses between cell types and columns, to prune neuron offsets with low confidences.

Synaptic signs for most cell types were predicted on the basis of known expression of neurotransmitter markers (primarily the cell-type-specific transcriptomics data from ref. 30). For a minority of cell types included in the model, no experimental data on transmitter phenotypes were available. For these neurons, we used guesses of plausible transmitter phenotypes. To derive predicted synaptic signs from transmitter phenotypes, we assigned the output of histaminergic, GABAergic and glutamatergic neurons as hyperpolarizing and the output of cholinergic neurons as depolarizing. In a few cases, we further modified these predictions on the basis of distinct known patterns of neurotransmitter receptor expression (see ref. 30 for details). For example, output from R8 photoreceptor neurons, predicted to release both acetylcholine and histamine, was treated as hyperpolarizing or depolarizing, respectively, depending on whether a target cell type is known to express the histamine receptor gene *ort* (which encodes a histamine-gated chloride channel).

Representing the model as a hexagonal convolutional neural network

Our end-to-end differentiable[61] DMN model of the fly visual system can be interpreted as a continuous-time neural ordinary differential equation[62] with a deep convolutional recurrent neural network[63] architecture that is trained to perform a computer vision task using backpropagation through time[64,65]. Our goal was to optimize a simulation of the fly visual system to perform a complex visual information processing task using optimization methods from deep learning. One hallmark of visual systems that has been widely exploited in such tasks is their convolutional nature[66–69] (that is, the fact that the same computations are applied to each pixel of the visual input). To model the hexagonal arrangement of photoreceptors in the fly retina, we developed a hexagonal convolutional neural network (CNN) in the widely used deep learning framework PyTorch[21] (ignoring neuronal superposition[70]), which we used for simulation and optimization of the model. We model columnar cell types, including retinal cells, lamina monopolar and wide-field cells, medulla intrinsic cells, transmedullary cells and T-shaped cells, as well as amacrine cells. The model comprises synapses from all neuropils and downstream- and upstream-projecting connections from the retina, lamina and medulla.

Neuronal dynamics

In detail, we simulated point neurons with voltages V_i of a postsynaptic neuron i, belonging to cell type t_i using threshold-linear dynamics, mathematically equivalent to commonly used formulations of firing-rate models[71]

$$\tau_{t_i}\dot{V}_i = -V_i + \sum_j s_{ij} + V_{t_i}^{\mathrm{rest}} + e_i. \tag{1}$$

Neurons of the same cell type share time constants, τ_{t_i}, and resting potentials, $V_{t_i}^{\mathrm{rest}}$. Dynamic visual stimuli were delivered as external input currents e_i to the photoreceptor (R1–R8), for all other cell types, $e_i = 0$. In our model, instantaneous graded synaptic release from presynaptic neuron j to postsynaptic neuron i is described by

$$s_{ij} = w_{ij} f(V_j) = \alpha_{t_it_j}\sigma_{t_it_j} N_{t_it_j\Delta u\Delta v} f(V_j), \tag{2}$$

comprising the anatomical filters in terms of the synapse count from electron microscopy reconstruction, $N_{t_it_j\Delta u\Delta v}$, at the offset location $\Delta u = u_i - u_j$ and $\Delta v = v_i - v_j$ in the hexagonal lattice between two types of cells, t_i and t_j, and further characterized by a sign, $\sigma_{t_it_j} \in \{-1, +1\}$, and a non-negative scaling factor, $\alpha_{t_it_j}$.

The synapse model entails a trainable non-negative scaling factor per filter that is initialized as

$$\alpha_{t_i,t_j} = \frac{0.01}{\langle N_{t_i,t_j\,\Delta u,\,\Delta v}\rangle},$$

with the denominator describing the average synapse count of the filter. Synapse counts, $N_{t_it_j\Delta u\Delta v}$ from the connectome, and signs, $\sigma_{t_it_j}$ from the neurotransmitter and receptor profiling, were kept fixed. The scaling factor was clamped during training to remain non-negative.

Moreover, at initialization, the resting potentials were sampled from a Gaussian distribution

$$V_{t_i}^{\mathrm{rest}} \sim \mathcal{N}(\mu_{V^{\mathrm{rest}}}, \sigma^2_{V^{\mathrm{rest}}})$$

with mean $\mu_{V^{\mathrm{rest}}} = 0.5$ (a.u.) and variance $\sigma^2_{V^{\mathrm{rest}}} = 0.05$ (a.u.). The time constants were initialized at $\tau_{t_i} = 50$ ms. The 50 task-optimized DMNs were initialized with the same parameter values. During training, in Euler integration of the dynamics, we clamped the time constants as $\tau_i = \max(\tau_i, \Delta t)$, so that they remain above the integration time step Δt at all times.

In total, the model comprises 45,669 neurons and 1,513,231 synapses, across two-dimensional (2D) hexagonal arrays 31 columns across. The number of free parameters is independent of the number of columns: 65 resting potentials, 65 membrane time constants, 604 scaling factors; and connectome-determined parameters: 604 signs and 2,355 synapse counts. Thus, the number of free parameters in the visual system model is 734.

In the absence of connectome measurements, the number of parameters to be estimated is much larger. With $T = 65$ cell types (counting CT1 twice for the compartments in the medulla and lobula) and $C = 721$ cells per type for simplicity, the number of cells in our model would be $TC = 46{,}865$. Assuming a recurrent neural network with completely unconstrained connectivity and simple dynamics $\tau_i\dot{V}_i = -V_i + \sum_j w_{ij} f(V_j) + V_i^{\mathrm{rest}}$, we would have to find $(TC)^2 + 2(TC) = 2{,}196{,}421{,}955$ free parameters. Assuming a convolutional recurrent neural network with shared filters between cells of the same postsynaptic type, shared time constants and shared resting potentials, the amount of parameters reduces markedly to $T^2C + 2T = 3,046,355$. Further assuming the same convolutional recurrent neural network but additionally that convolutional filters are constrained to $F = 5$ visual columns (that is, the number of presynaptic input columns in the hexagonal lattice is $P = 3F\,(F + 1) + 1$), the amount of parameters reduces to $T^2P + 2T = 384,605$. Assuming as in our connectome that only $Q = 604$ connections between cell types exist, this reduces the number of parameters further to $QP + 2T = 55,185$. Instead of parametrizing each individual synapse strength, we assume that synapse strength is proportional to synapse count from the connectome times a scalar for each filter, reducing the number of parameters to $Q + 2T = 734$ while providing enough capacity for the DMNs to yield realistic tuning to solve the task.

Convolutions using scatter and gather operations. For training the network, we compiled the convolutional architecture specified by the connectome and the sign constraints to a graph representation containing: a collection of parameter buffers shared across neurons and/or connections; a collection of corresponding index buffers indicating where the parameters relevant to a given neuron or connection can be found in the parameter buffers; and a list of pairs (presynaptic neuron index, postsynaptic neuron index) denoting connectivity. This allowed us to efficiently simulate the network dynamics through Euler integration using a small number of element-wise, scatter and gather operations at each time step. We found that this is more efficient than using a single convolution operation or performing a separate convolution for each cell type as each cell type has its own receptive field—some much larger than others—and the number of cells per type is relatively small.

Optic flow task

Model training. An optic flow field for a video sequence consists of a 2D vector field for each frame. The 2D vector at each pixel represents the magnitude and direction of the apparent local movement of the brightness pattern in an image.

We frame the training objective as a regression task

$$\hat{\mathbf{Y}}[n] = \mathrm{Decoder}\left(\mathrm{DMN}\left(\mathbf{X}[0], \ldots, \mathbf{X}[n]\right)\right),$$

with $\hat{\mathbf{Y}}$ being the optic flow prediction, and $\mathbf{X}$ being the visual stimulus sequence from the Sintel dataset, both sampled to a regular hexagonal lattice of 721 columns. With the objective to minimize the square error loss between predicted optic flow and target optic flow fields, we jointly optimized the parameters of both the decoder and the visual system network model described above.

In detail, for training the network, we added randomly augmented, greyscaled video sequences from the Sintel dataset sampled to a regular hexagonal lattice of 721 columns to the voltage of the 8 photoreceptor cell types (Fig. 1f and equation (1)). We denote a sample from a minibatch of video sequences as $\mathbf{X} \quad \mathbb{R}^{N,C}$, with N being the number of time steps, and C being the number of photoreceptor columns. The dynamic range of the input lies between 0 and 1. Input sequences during training entailed 19 consecutive frames drawn randomly from the dataset and resampled to match the integration rate. At the original frame rate of 24 Hz, this corresponds to a simulation of 792 ms. We did not find that an integration time step smaller than 20 ms (that is, a frame rate of 50 Hz after resampling) yielded qualitatively superior task performance or more realistic tuning predictions. We interpolated the target optic flow in time to 50 Hz temporal resolution. To increase the amount of training data for better generalization, we augmented input and target sequences as described further below. At the start of each epoch, we computed an initial state of the network's voltages after 500 ms of grey stimulus presentation to initialize the network at a steady state for each minibatch during that epoch. The network integration for a given input $\mathbf{X}$ results in simulated sequences of voltages $\mathbf{V} \quad \mathbb{R}^{N,T_C}$, with T_C being the total number of cells. The subset of voltages, $\mathbf{V}_{\mathrm{out}} \quad \mathbb{R}^{N,D,C}$, of the D cell types in the black rectangle in Fig. 1g was passed to a decoding network. For decoding, the voltage was rectified to avoid the network finding biologically implausible solutions by encoding in negative dynamic ranges. Furthermore, it was mapped to Cartesian coordinates to apply PyTorch's standard spatial convolution layers for decoding and on each time step independently. In the decoding network, one layer implementing spatial convolution, batch normalization, softplus activation and dropout, followed by one layer of spatial convolution, transforms the D feature maps into the 2D representation of the estimated optic flow, $\hat{\mathbf{Y}} \quad \mathbb{R}^{N,2,C}$.

Using stochastic gradient descent with adaptive moment estimation ($_1 = 0.9$, $_2 = 0.999$, learning rate decreased from $5 \times 10^{\ 5}$ to $5 \times 10^{\ 6}$ in ten steps over iterations, batch size of four) and the automatic gradient calculation of the fully differentiable pipeline, we optimized the biophysical parameters through backpropagation through time such that they minimize the L2-norm between the predicted optic flow, $\hat{\mathbf{Y}}$, and the ground-truth optic flow, $\mathbf{Y}$:

$$L(\mathbf{Y}, \hat{\mathbf{Y}}) = ||\mathbf{Y} \quad \hat{\mathbf{Y}}||.$$

We additionally optimized the shared resting potentials for 150,000 iterations, using stochastic gradient descent without momentum, with respect to a regularization function of the time-averaged responses to naturalistic stimuli of the central column cell of each cell type, t_{central}, to encourage configurations of resting potentials that lead to non-zero and non-exploding activity in all neurons in the network. We weighted these terms independently with $\quad = 1$, encouraging activity greater than a, and $\quad = 0.01$, encouraging activity less than a. We chose $_V = 0.1$ and $a = 5$ in arbitrary units. With B being the batch size and T being the number of all cell types, the regularizer is

$$R(V) = \frac{\ _V}{BT} \quad _b \quad _{t_{\mathrm{central}}} \begin{cases} (\bar{V} \quad a)^2, & \text{if } \bar{V} = \frac{1}{N} \quad _n \quad V_{bt_{\mathrm{central}}}[n] \quad a \\ (\bar{V} \quad a)^2, & \text{if } \bar{V} > a. \end{cases}$$

We regularly checkpointed the error measure $L\,(\mathbf{Y}, \hat{\mathbf{Y}})$ averaged across a held-out validation set of Sintel video clips. Models generalized on optic flow computation after about 250,000 iterations, yielding functional candidates for our fruit fly visual system models that we analysed with respect to their tuning properties.

Task-optimization dataset. We optimized the network on 23 sequences from the publicly available computer-animated film Sintel[12]. The sequences have 20–50 frames, at a frame rate of 24 frames per second and a pixel resolution of 1,024 × 436. The dataset provides optical flow in pixel space for each frame after the first of each sequence. As the integration time steps we use are faster than the actual sampling rate of the sequences, we resample input frames accordingly over time and interpolate the optic flow.

Fly-eye rendering. We first transformed the RGB pixel values of the visual stimulus to normalized greyscale between 0 and 1. We translated Cartesian frames into receptor activations by placing simulated photoreceptors in a 2D hexagonal array in pixel space, 31 columns across resulting in 721 columns in total, spaced 13 pixels apart. The transduced luminance at each photoreceptor is the greyscale mean value in the 13 × 13-pixel region surrounding it.

Augmentation. We used: random flips of input and target across one of the three principal axes of the hexagonal lattice; random rotation of input and target around its six-fold rotation axis; adding element-wise Gaussian noise with mean zero and variance $_n = 0.08$ to the input X (then clamped at 0); random adjustments of contrasts,

$\log c \sim \mathcal{N}(0,\ {}_c^2 = 0.04)$ and brightness, $b \sim \mathcal{N}(0,\ {}_b^2 = 0.01)$, of the input with $X\ = c\,(X\quad 0.5) + 0.5 + cb$.

In addition, we 'strided' the fly-eye rendering across the rectangular raw frames in width, subsampling multiple scenes from each. We ensured that such subsamples from the same scene were not distributed across training and validation sets. Input sequences in chunks of 19 consecutive frames were drawn randomly in time from the full sequences.

Black-box decoding network. The decoding network is feedforward, convolutional and has no temporal structure. Aspects of the architecture are explained in the section entitled Model training. The spatial convolutions have a filter size of 5 × 5. The first layer transforms the $D = 34$ feature maps to an eight-channel intermediate representation, which is further translated by an additional convolutional layer to a three-channel intermediate representation of optic flow. The third channel is used as shared normalization of each coordinate of the remaining 2D flow prediction. The decoder uses PyTorch-native implementations for 2D convolutions, batch normalization, softplus activation and dropout. We initialized its filter weights homogeneously at 0.001.

Model characterization

Task error. To rank models on the basis of their task performance, we computed the standard optic flow metric of average end-to-end point error (EPE)[72], which calculates the average over all time steps and pixels (that is, here columns) of the error

$$\mathrm{EPE}(\mathbf{Y}, \hat{\mathbf{Y}}) = \frac{1}{NC} \ {}_{n} \ {}_{c} \sqrt{(y_{1c}[n] \quad \hat{y}_{1c}[n])^2 + (y_{2c}[n] \quad \hat{y}_{2c}[n])^2}$$

between predicted optic flow and ground-truth optic flow and averaged across the held-out validation set of Sintel sequences.

Importance of task optimization and connectome constraints. We generated DMNs with different constraints to assess their relative importance for predicting tuning properties. First, we studied the importance of task optimization of DMN parameters. We created 50 DMNs with random Gaussian-distributed parameters, and task-optimized only their decoding network, to obtain baseline values for both the task error and the accuracy of predicting tuning curves without task optimization of the DMN.

In the full DMN, we constrained single synapses by connectome cell-type connectivity, cell connectivity, synapse counts and synapse signs (equation (2)) and task-optimized the non-negative type-to-type unitary synapse scaling factor $\boldsymbol{\alpha}_{t_i,t_j}$. Next, we trained five additional task-optimized DMNs with different connectome constraints (Fig. 2d and Extended Data Fig. 2a–d).

In these five additional types of DMN, we additionally task-optimized the terms in bold, rather than using connectome measurements, related to synaptic currents from a presynaptic cell j to a postsynaptic cell i: known single-cell connectivity, unknown synapse count: $w_{ij} = \ {}_{t_i,t_j}\mathbf{m}_{t_i,t_j,\ u,\ v}$, in which $\mathbf{m}_{t_i,t_j,\ u,\ v}$ is non-negative; known cell-type connectivity, unknown single-cell connectivity and synapse counts: $w_{i,j} = \ {}_{t_i,t_j}\mathbf{m}_{t_i,t_j,\ 3<\ u,\ v,\ u+\ v<3}$ (that is, for all connected cell types, a connection weight was learned for all cells up to a distance of three columns in hexagonal coordinates, with known signs); known single-cell connectivity and synapse counts, but unknown synapse signs: $w_{i,j} = \boldsymbol{\alpha}_{t_i,t_j}\boldsymbol{\sigma}_{t_j}N_{t_i,t_j,\ u,\ v}$ (that is, connection weights were fixed by measurements, but signs optimized); known single-cell connectivity, but unknown synapse signs and synapse counts: $w_{i,j} = \mathbf{w}_{t_i,t_j,\ u,\ v}$, (that is, all non-zero connection weights were optimized, including their signs); or known cell-type connectivity, unknown single-cell connectivity, synapse counts and synapse signs: $w_{i,j} = \mathbf{w}_{t_i,t_j,\ 3<\ u,\ v,\ u+\ v<3}$ (that is, for all connected cell types, a connection weight and sign was learned for all cells up to distance of three columns). We trained 50 models per DMN type. The task-optimized parameters in each case are highlighted using bold symbols. We randomly initialized the models with $\mathbf{m}_{t_i,t_j}, \mathbf{w}_{t_i,t_j} \sim \mathcal{N}\left(0, \frac{2}{n_{\mathrm{in}}}\right)$, in which n_{in} is the number of cell connections and $\mathbf{m}$ is non-negative, and $\boldsymbol{\sigma}_{t_j}\quad \{\ 1, 1\}$ with equal probability.

Unconstrained CNN. We trained unconstrained, fully convolutional neural networks on the same dataset and task to provide an estimate of a lower bound for the task error of the DMN. Optic flow was predicted by the CNN from two consecutive frames

$$\hat{Y}[n] = \mathrm{CNN}(X[n], X[n \quad 1]).$$

with the original frame rate of the Sintel film. We chose 5 layers for the CNN with 32, 92, 136, 8 and 2 channels, respectively, and kernel size 5 for all but the first layer, for which the kernel size is 1. Each layer performs a convolution, batch normalization and exponential linear unit activation, except the last layer, which performs only a convolution. We optimized an ensemble of 5 unconstrained CNNs with 414,666 free parameters each using the same loss function, $L(Y, \hat{Y})$, as for the DMN. We used the same dataset (that is, hexagonal sequences and augmentations from Sintel) for training and validating the CNN as that used for training and validating the DMN, enabled by two custom modules mapping from the hexagonal lattice to a Cartesian map and back.

Circular flash stimuli. To evaluate the contrast selectivity of cell types in task-optimized models, we simulated responses of each DMN to circular flashes. The networks were initialized at an approximate steady state after 1 s of grey-screen stimulation. Afterwards the flashes were presented for 1 s. The flashes with a radius of 6 columns were ON (intensity $I = 1$) or OFF ($I = 0$) on a grey ($I = 0.5$) background. We integrated the network dynamics with an integration time step of 5 ms. We recorded the responses of the modelled cells in the central columns to compute the FRI.

FRI. To derive the contrast selectivity of a cell type, t_i, we computed the FRI as

$$\mathrm{FRI}_{\mathrm{t_i}} = \frac{r^{\mathrm{peak}}_{\mathrm{t}_{\mathrm{central}}}(I=1) \quad r^{\mathrm{peak}}_{\mathrm{t}_{\mathrm{central}}}(I=0)}{r^{\mathrm{peak}}_{\mathrm{t}_{\mathrm{central}}}(I=1) + r^{\mathrm{peak}}_{\mathrm{t}_{\mathrm{central}}}(I=0)}$$

from the non-negative activity

$$r^{\mathrm{peak}}_{\mathrm{t}_{\mathrm{central}}}(I) = \max_{n} V_{t_{\mathrm{central}}}[n](I) + |\min_{n,I} V_{t_{\mathrm{central}}}[n](I)|,$$

from voltage responses $V_{t_{\mathrm{central}}}[n](I)$ to circular flash stimuli of intensities $I\quad \{0, 1\}$ lasting for 1 s after 1 s of grey stimulus. We note that our index quantifies whether the cell depolarizes to ON- or to OFF-stimuli. However, cells such as R1–R8, L1 and L2 can be unrectified (that is, sensitive to both light increments and light decrements), which is not captured by our index.

For the P values reported in the results, we carried out a binomial test with probability of correct prediction 0.5 (H0) or greater (H1) to test whether both the median FRI from the DMN ensemble and the task-optimal model can predict the contrast preferences. Additionally, we found for each individual cell type across 50 DMNs that predictions for 29 out of 31 cell types are significant ($P < 0.05$, binomial).

Moving-edge stimuli. To predict the motion sensitivity of each cell type in task-constrained DMNs, we simulated the response of each network, initialized at an approximate steady state after 1 s of grey-screen stimulation, to custom generated edges moving to 12 different directions, $\quad [0°, 30°, 60°, 90°, 120°, 150°, 180°, 210°, 240°, 270°, 300°, 330°]$. We integrated the network dynamics with an integration time step of 5 ms. ON-edges ($I = 1$) or OFF-edges ($I = 0$) moved on a grey ($I = 0.5$) background. Their movement ranged from $\quad 13.5°$ to 13.5°

visual angle and we moved them at six different speeds, ranging from 13.92° s ^{1} to 145° s ^{1} (S [13.92° s ^{1}, 27.84° s ^{1}, 56.26° s ^{1}, 75.4° s ^{1}, 110.2° s ^{1}, 145.0° s ^{1}]). In Fig. 2d, we report the correlation between predicted motion-tuning curves to the single experimentally measured tuning curve. We take the maximum correlation across six investigated speeds to make the correlation measure robust to potential variations in preferred speeds.

DSI. We computed a DSI[73] of a particular type t_i as

$$\mathrm{DSI}_{\mathrm{t_i}}(I) = \frac{1}{|\mathbb{S}|}_{S \quad \mathbb{S}} \frac{| \quad r^{\mathrm{peak}}_{\mathrm{t_{central}}}(\mathrm{I}, \mathrm{S}, \quad) \exp(i \quad)|}{\max\limits_{I \quad \mathbb{I}} | \quad r^{\mathrm{peak}}_{\mathrm{t_{central}}}(\mathrm{I}, \mathrm{S}, \quad)|}$$

from rectified peak voltages

$$r^{\mathrm{peak}}_{\mathrm{t_{central}}}(\mathrm{I}, \mathrm{S}, \quad) = \max_n V^{+}_{t_{\mathrm{central}}}[n](\mathrm{I}, \mathrm{S}, \quad),$$

elicited from moving-edge stimuli. We rectify the voltage to quantify the tuning of the effective output of the cell, and to avoid the denominator becoming zero. We parameterized movement angle , intensities $I \quad \mathbb{I}$, and speeds $S \quad \mathbb{S}$ of the moving edges. To take the response magnitudes into account for comparing the DSI for ON- and for OFF-edges, we normalized by the maximum over both intensities in the denominator. To take different speeds into account, we averaged over $\mathbb{S}$.

Normalization of model neural activity for averaging across models in a cluster. Threshold-linear networks have arbitrary units for the voltages and currents. Therefore, we normalized the neural activity before averaging the neural activity predictions from different models. For a single cell or cell type t, we first divided responses (voltages or rectified voltages) by the root mean square across the cell's responses to the naturalistic stimuli:

$$r_t^{\| \quad \|^{\sim 1}}[n] = \frac{r_t[n]}{\left\| \frac{1}{SN} \mathbf{R}_t^{\mathrm{nat.}} \right\|_2},$$

in which $\mathbf{R}_t^{\mathrm{nat.}} \quad \mathbb{R}^{S,N}$ is the cell's response vector to S sequences from the Sintel dataset with N time steps and $r_t[n]$ is the cell's response to any stimuli. This normalization makes averages (Fig. 4a,b,d–e and Extended Data Figs. 4, 7, 8, and 9a,b) independent to variation in the scale of neural activity from model to model. We normalized input currents equivalently (Fig. 4b and Extended Data Figs. 4, 7, and 8) by the same normalization factor. We exclude solutions for which the denominator becomes zero.

Determining whether a cell type with asymmetric inputs counts as direction selective. We counted a cell type as direction selective if the DSIs from its synthetic measurements were larger than 99% of DSIs from non-motion selective cell types (that is, those with symmetric filters). We note, however, that estimates of the spatial asymmetry of connectivity from existing connectome reconstructions can be noisy.

For deriving the 99% threshold, we first defined a distribution $p(d^*|\mathrm{t_{sym}})$ over the DSI for non-direction-selective cells, from peak responses to moving edges of cell types with symmetric inputs, $\mathrm{t_{sym}}$. We computed that distribution numerically by sampling

$$d^* = \frac{| \quad {}_{*} r^{\mathrm{peak}}_{\mathrm{t_{central}}}(\mathrm{I}, \mathrm{S}, \quad ^{*}) \exp i \quad |}{| \quad r^{\mathrm{peak}}_{\mathrm{t_{central}}}(\mathrm{I}, \mathrm{S}, \quad)|}$$

for 100 independent permutations of the angle *. We independently computed d^* for all stimulus conditions, models and cell types with symmetric inputs. From $p(d^*|\mathrm{t_{sym}})$, we derived the threshold $d_{\mathrm{thresh}} = 0.357$ as the 99% quantile of the random variable d^*, meaning that the probability that a realization of $d^* > d_{\mathrm{thresh}}$ is less than 1% for cell types with symmetric inputs. To determine whether an asymmetric cell type counts as direction selective, we tested whether synthetically measuring direction selectivity larger than d_{thresh} in that cell type is binomial with probability 0.1 (H0) or greater (H1). We identified 12 cell types with asymmetric inputs (T4a, T4b, T4c, T4d, T5a, T5b, T5c, T5d, TmY3, TmY4, TmY5a and TmY18) as direction selective ($P < 0.05$) from our models, and 7 cell types with asymmetric inputs to not count as direction selective (T2, T2a, T3, Tm3, Tm4, TmY14 and TmY15; see Extended Data Fig. 5 as reference for cell types with symmetric and asymmetric inputs).

Uniform manifold approximation and projection and clustering. We first simulated central column responses to naturalistic scenes (24 Hz Sintel video clips from the full augmented dataset) with an integration time step of 10 ms. We clustered models in feature space of concatenated central column responses and sample dimension. Next, we computed a nonlinear dimensionality reduction to two dimensions using the UMAP (uniform manifold approximation and projection) algorithm, and fitted Gaussian mixtures of 2 to 5 components, with the number of components that minimize the Bayesian information criterion, using the Python libraries umap-learn and scikit-learn[38,74].

Single-ommatidium flashes. To derive spatio-temporal receptive fields of cells, we simulated the response of each network to single-ommatidium flashes. Flashes were ON ($I = 1$) or OFF ($I = 0$) on a grey ($I = 0.5$) background and presented for [5, 20, 50, 100, 200, 300] ms after 2 s of grey-screen stimulation and followed by 5 s of grey-screen stimulation.

Spatio-temporal, spatial and temporal receptive fields. We derived the spatio-temporal receptive field (STRF) of a cell type t_i as the baseline-subtracted responses of the central column cell to single-ommatidium flashes $J(u, v)$ at ommatidium locations (u, v):

$$\mathrm{STRF}_{t_{\mathrm{central}}}[n](u, v) = V_{t_{\mathrm{central}}}[n](J(u, v)) \quad V_{t_{\mathrm{central}}}[n = 0](J(u, v)).$$

We derived spatial receptive fields (SRFs) from the responses to flashes (20 ms in Fig. 4d) $J(u, v)$ at the point in time at which the response to the central ommatidium impulse is at its extremum:

$$\mathrm{SRF}(u, v) = \mathrm{STRF}(n = \mathrm{argmax}_n \, |\mathrm{STRF}[n](0, 0)|, u, v).$$

We derive temporal receptive fields (TRFs) from the response to a flash $J(0, 0)$ at the central ommatidium: $\mathrm{TRF}[n] = \mathrm{STRF}[n](0, 0)$. For averaging receptive fields across multiple models, we first normalize the voltages as described above.

Maximally excitatory naturalistic and artificial stimuli. First, we found the naturalistic maximally excitatory stimulus, X*, by identifying the Sintel video clip, **X**, from the full dataset with geometric augmentations that elicited the highest possible response in the central column cell of a particular cell type in our models.

$$\mathrm{X}^* = \underset{\mathrm{X} \quad \mathrm{Sintel}}{\mathrm{argmax}} V_{t_{\mathrm{central}}}(\mathrm{X}).$$

Next, we regularized the naturalistic maximally excitatory stimulus, to yield X , capturing only the stimulus information within the receptive field of the cell, with the objective to minimize

$$L(\mathrm{X} \;) = \;_{n} \; \|V_{t_{\mathrm{central}}}(\mathrm{X}^*)[n] \quad V_{t_{\mathrm{central}}}(\mathrm{X} \;)[n]\|^2 + \frac{1}{C} \;_{c} \; \|\mathrm{X} \; [n, c] \quad 0.5\|^2.$$

The first summand preserves the central response to X*, and the second regularizes the irrelevant portions of the stimulus outside the receptive field to grey ($I = 0.5$).

In addition, we computed artificial maximally excitatory stimuli[75].

Model selection. To describe the most data-consistent motion tuning mechanisms predicted by the ensemble at the level of single-cell currents, for Extended Data Figs. 4, 7 and 8, we automatically selected those models from the ensemble with tuning matching to empirical data. Specifically, we selected models with correct contrast tuning in the respective target cells and their inputs (Fig. 4c and Extended Data Fig. 3d), with the DSI larger than the threshold d^* derived above, and with a correctly predicted preferred direction (45° acceptance angle, assuming 225° for TmY3).

Training synthetic connectomes

Training feedforward synthetic ground-truth connectome networks. Sparsified feedforward neural networks with six hidden layers (linear transformations sandwiched between rectifications) with equal number of neurons in each hidden layer functioned as ground-truth connectome networks. The main results describe networks with 128 neurons per hidden layer. We interpret the individual units as neurons with voltage

$$V_i = {}_{j}\, s_{ij} + V_i^{\text{rest}} = {}_{j}\quad {}_j c_{ij} m_{ij} f(V_j) + V_i^{\text{rest}},$$

with presynaptic inputs s_{ij} and resting potentials V_i^{rest}. The connectome-constrained synapse strength, w_{ij}, is characterized by the adjacency matrix c_{ij}, the signs $\ _j$, and the non-negative weight magnitudes m_{ij}. $c_{ij} = 1$ if the connection exists, else $c_{ij} = 0$. To respect Dale's law, the signs were tied to the presynaptic identity, j.

We identified the parameters $\ _j$, m_{ij} and V_i^{rest} by task optimization on handwritten digit classification (Modified National Institute of Standards and Technology (MNIST) database)[76]. We determined adjacency matrices, c_{ij}, for a given connectivity percentage using an iterative local pruning technique, the lottery ticket hypothesis algorithm[77]. The algorithm decreases the connectivity percentage of the ground-truth connectome networks while maintaining high task accuracy.

We optimized the ground-truth connectome networks and all simulated networks described below in PyTorch with stochastic gradient descent with adaptive moment estimation (ADAM with AMSGrad), learning rate 0.001, batch size 500, and an exponentially decaying learning rate decay factor of 0.5 per epoch. To constrain the weight magnitudes to stay non-negative, we clamped the values at zero after each optimization step (projected gradient descent). The parameters after convergence minimize the cross-entropy loss between the predicted and the ground-truth classes of the handwritten digits. More implementation detail is available in Supplementary Note 5.

Simulated networks with known connectivity and unknown strength. Simulated networks inherited connectivity, c_{ij}, and synapse signs, $\ _j$, from their respective ground-truth connectome networks. In simulated networks, signs and connectivity were held fixed. Weight magnitudes, m_{ij}, and resting potentials, V_i^{rest}, were initialized randomly and task-optimized. Just like ground-truth connectome networks, simulated networks were trained on the MNIST handwritten digit classification task until convergence.

Simulated networks with known connectivity and known strength. Alternatively, we imitate measurements of synaptic counts from the ground-truth weight magnitudes:

$$\widetilde{m}_{ij} = m_{ij}\ {}_{ij} \text{ with } {}_{ij} \sim \mathcal{U}(1 \quad , 1 + \quad),$$

with multiplicative noise to imitate spurious measurements. We used $\ = 0.5$ for the main results. Weight magnitudes were initialized at the measurement, $\widetilde{m}_{ij}$, and task-optimized on MNIST with the additional objective to minimize the squared distance between optimized and measured weight magnitudes, $\widetilde{m}_{ij}$ (L2 constraint, Gaussian weight magnitude prior centred around the simulated network's initialization). We weighted the L2 constraint ten times higher than the cross-entropy objective to keep weight magnitudes of the simulated networks close to the noisy connectome measurements. Resting potentials, V_i^{rest}, were again initialized randomly and task-optimized.

Measuring ground-truth-simulated network similarity. Ground-truth-simulated network similarity was measured by calculating the median Pearson's correlation of tuning responses (rectified voltages) of corresponding neurons in the ground-truth-simulated network pair. In each of the 6 hidden layers, $N = 100$ randomly sampled neurons were used for comparison. Response tuning was measured over input stimuli from the MNIST test set ($N = 10{,}000$ images). Results are medians over all hidden layers and over 25 ground-truth-simulation network pairs.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Data, trained models and interactive notebooks are available at https://www.github.com/TuragaLab/flyvis.

Code availability

Code is available at https://www.github.com/TuragaLab/flyvis.

61. AlQuraishi, M. & Sorger, P. K. Differentiable biology: using deep learning for biophysics-based and data-driven modeling of molecular mechanisms. *Nat. Methods* **18**, 1169–1180 (2021).
62. Chen, R. T. Q., Rubanova, Y., Bettencourt, J. & Duvenaud, D. K. Neural ordinary differential equations. In *Proc. Advances in Neural Information Processing Systems* Vol. 31 (eds Bengio, S. et al.) 6571–6583 (Curran Associates, 2018).
63. Shi, X. et al. Convolutional LSTM network: a machine learning approach for precipitation nowcasting. In *Proc. Advances in Neural Information Processing Systems* Vol. 28 (eds Cortes, C. et al.) 802–810 (Curran Associates, 2015).
64. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *Nature* **323**, 533–536 (1986).
65. Werbos, P. J. Backpropagation through time: what it does and how to do it. *Proc. IEEE* **78**, 1550–1560 (1990).
66. Fukushima, K. & Miyake, S. in *Competition and Cooperation in Neural Nets* (eds Amari, S. et al.) 267–285 (Springer, 1982).
67. LeCun, Y. et al. Backpropagation applied to handwritten zip code recognition. *Neural Comput.* **1**, 541–551 (1989).
68. Riesenhuber, M. & Poggio, T. Hierarchical models of object recognition in cortex. *Nat. Neurosci.* **2**, 1019–1025 (1999).
69. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet classification with deep convolutional neural networks. In *Proc. Advances in Neural Information Processing Systems* (eds Pereira, F. et al.) 1097–1105 (Curran Associates, 2012).
70. Braitenberg, V. Patterns of projection in the visual system of the fly. I. Retina-lamina projections. *Exp. Brain Res.* **3**, 271–298 (1967).
71. Miller, K. D. & Fumarola, F. Mathematical equivalence of two common forms of firing rate models of neural networks. *Neural Comput.* **24**, 25–31 (2012).
72. Dosovitskiy, A. et al. FlowNet: Learning optical flow with convolutional networks. In *Proc. IEEE International Conference on Computer Vision* 2758–2766 (IEEE, 2015).
73. Mazurek, M., Kager, M. & Hooser, S. D. V. Robust quantification of orientation selectivity and direction selectivity. *Front. Neural Circuits* **8**, 92 (2014).
74. Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
75. Walker, E. Y. et al. Inception loops discover what excites neurons most using deep predictive models. *Nat. Neurosci.* **22**, 2060–2065 (2019).
76. LeCun, Y., Cortes, C., & Burges, C. J. The MNIST database of handwritten digits. http://yann.lecun.com/exdb/mnist (accessed 4 September 2024).
77. Frankle, J. & Carbin, M. The lottery ticket hypothesis: finding sparse, trainable neural networks. In *Proc. International Conference on Learning Representations* https://openreview.net/forum?id=rJl-b3RcF7 (2018).

Acknowledgements We thank L. Scheffer and L. Umayam for assistance with accessing connectome reconstructions; A. Borst, J. Fitzgerald, N. Klapoetke, G. Rubin, M. Reiser, K. Svoboda and members of the laboratories of S.C.T. and J.H.M. for discussions; J. Fitzgerald, D. Stern, N. Klapoetke, A. Lee, R. Gao, J. Voigts, B. Mensh, A. Schulz, P. Ramesh, M. Deistler, Z. Stefanidi and J. Joyce for feedback on the manuscript; and T. Herman for creating and sharing the colourization of the optic lobe figure[25] (Fig. 1b). The graphics of the fruit fly and the electron microscope in Fig. 1a were created with BioRender.com. This article is subject to HHMI's Open Access to Publications policy. HHMI laboratory heads have previously granted a nonexclusive CC BY 4.0 license to the public and a sublicensable licence to HHMI in their research articles. Pursuant to those licences, the author-accepted manuscript of this article can be made freely available under a CC BY 4.0 licence immediately on publication. This project was supported by the HHMI. J.K.L. and J.H.M. were supported by the German Research Foundation (DFG) through Germany's Excellence Strategy (EXC-Number 2064/1, Project number 390727645), the German Federal Ministry of Education and Research (BMBF; Tübingen AI Center, FKZ: 01IS18039A) and the European Union (ERC, DeepCoMechTome, 101089288). Views and opinions expressed are however those of the author(s) alone and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. J.H.M. is the principal investigator of the DFG-financed SFB 1233. J.K.L. is a member of the International Max Planck Research School for Intelligent Systems.

Author contributions Conceptualization, methodology: J.K.L., F.D.T., M.M., J.H.M. and S.C.T. Data curation: J.K.L., F.D.T., A.N., K.S. and S.-y.T. Software and investigation: J.K.L., M.M., S.P. and F.D.T. Analysis: J.K.L., E.G., A.N., S.P. and S.C.T. Writing: J.K.L., J.H.M. and S.C.T. Writing (review and editing): E.G., A.N., K.S., M.M., S.P. and F.D.T. Supervision and funding: J.H.M. and S.C.T.

Competing interests The authors declare no competing interests.

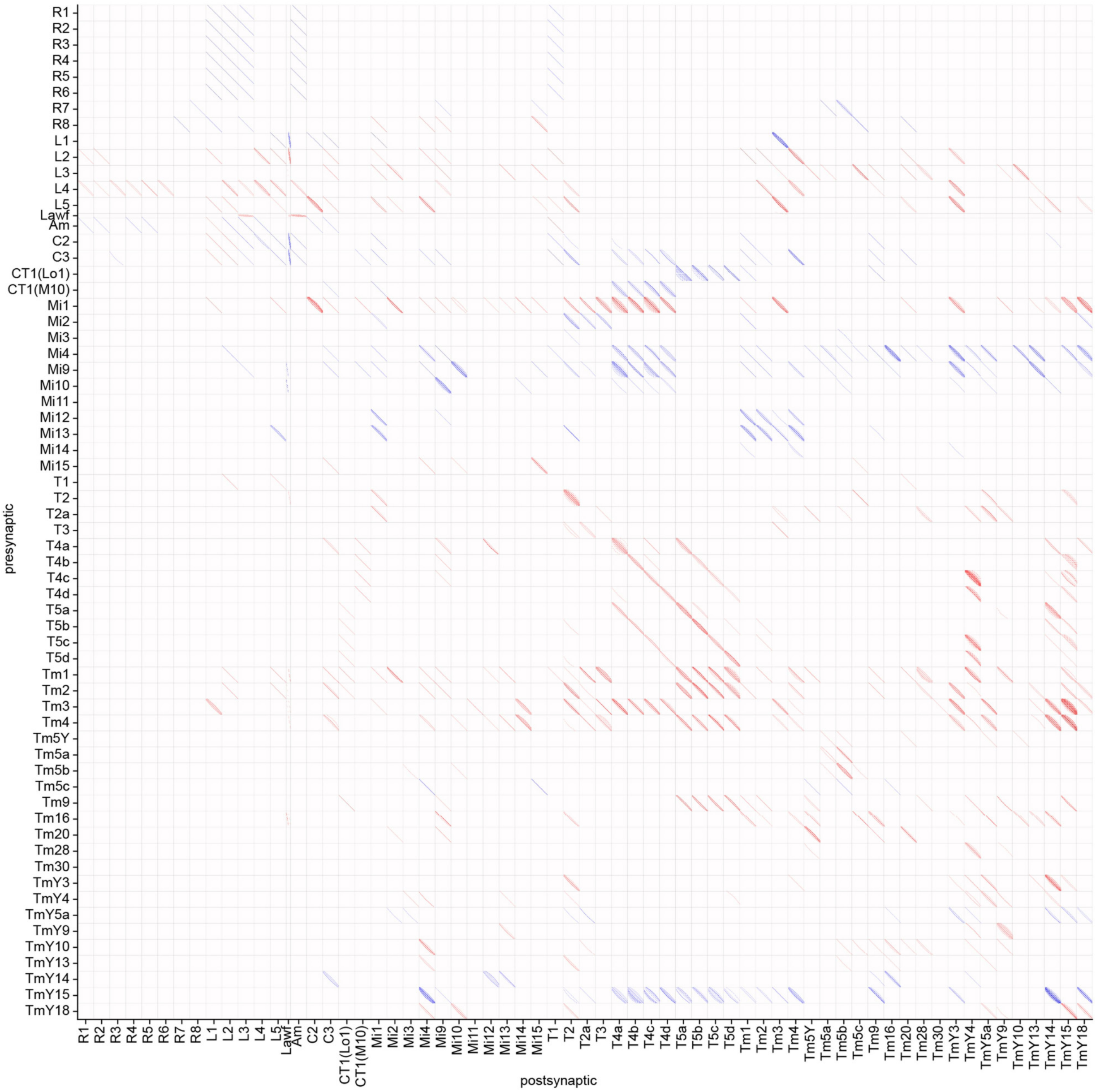
Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-024-07939-3.

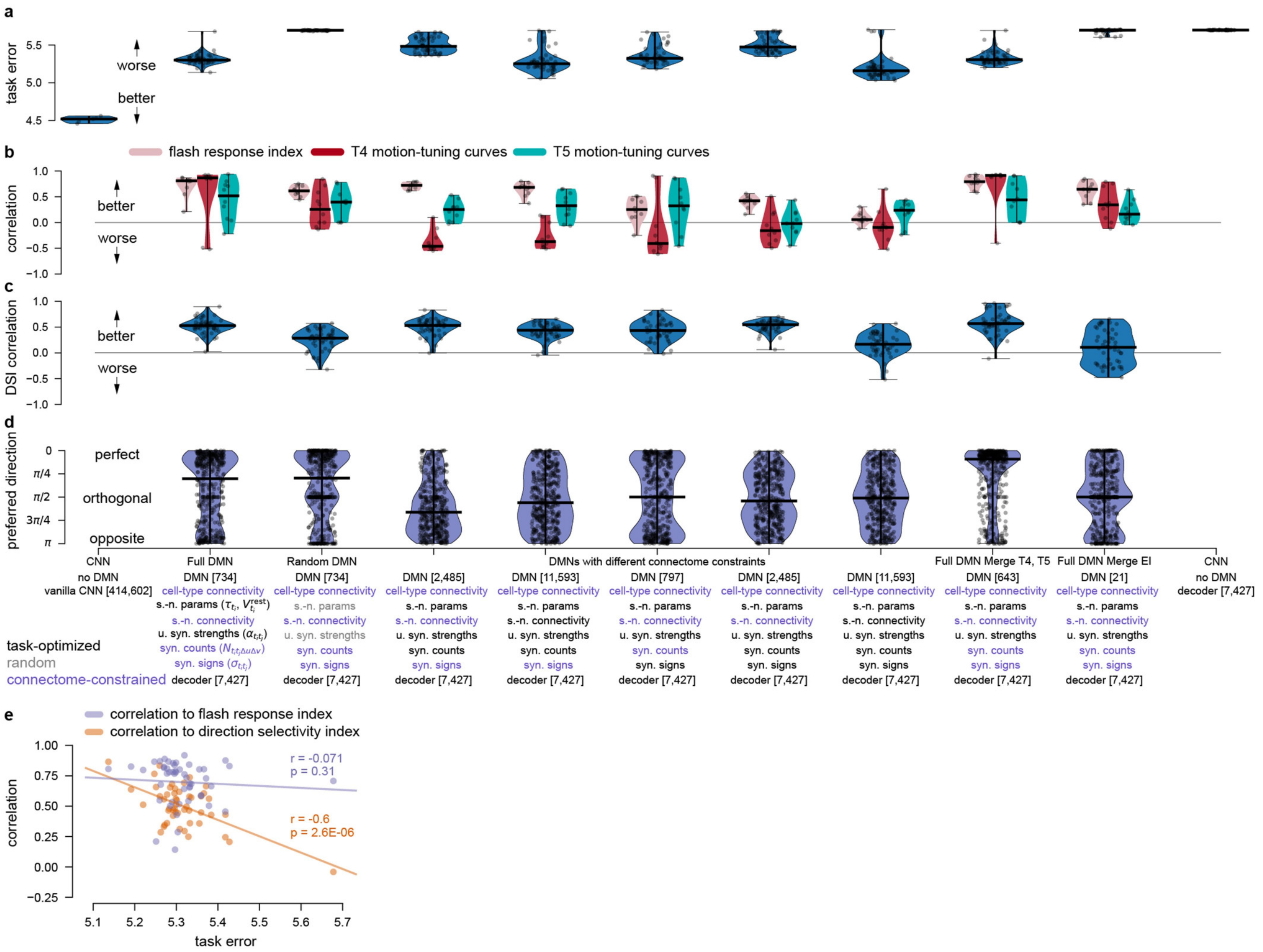
Correspondence and requests for materials should be addressed to Srinivas C. Turaga.

Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

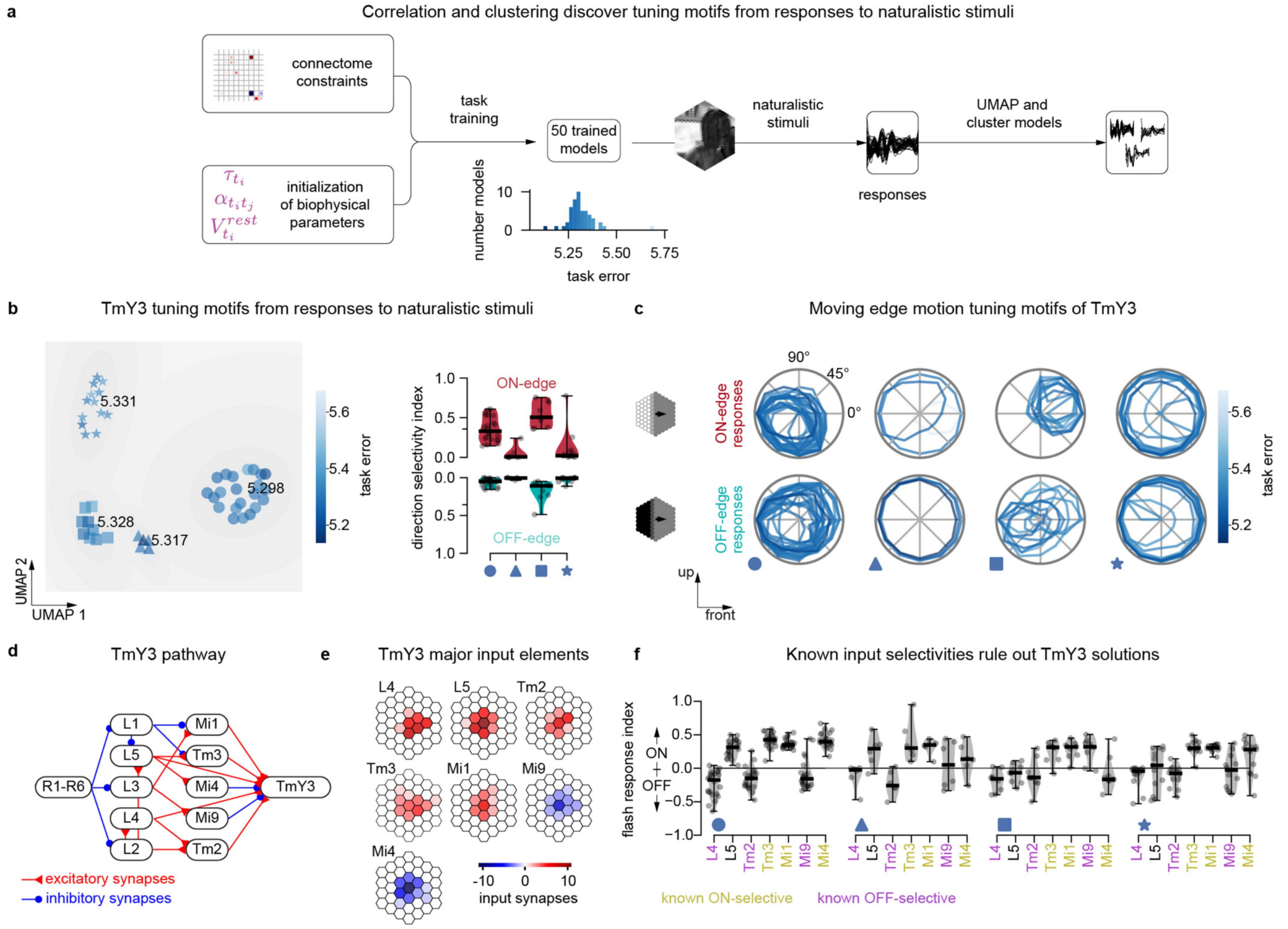
Reprints and permissions information is available at http://www.nature.com/reprints.



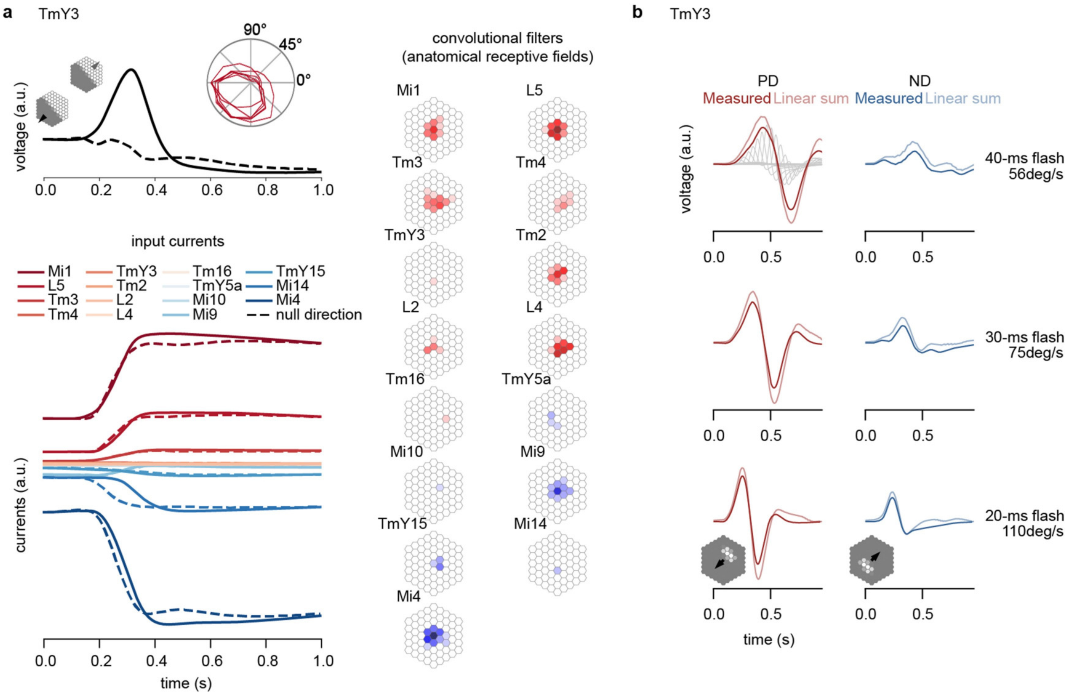
Extended Data Fig. 1 | Cell connectivity. The matrix shows how cells of the 64 cell types within the inner 91 columns (of 721) of the recurrent convolutional DMN connect (Supplementary Data file 1), either by excitatory connections (red) or inhibitory connections (blue).



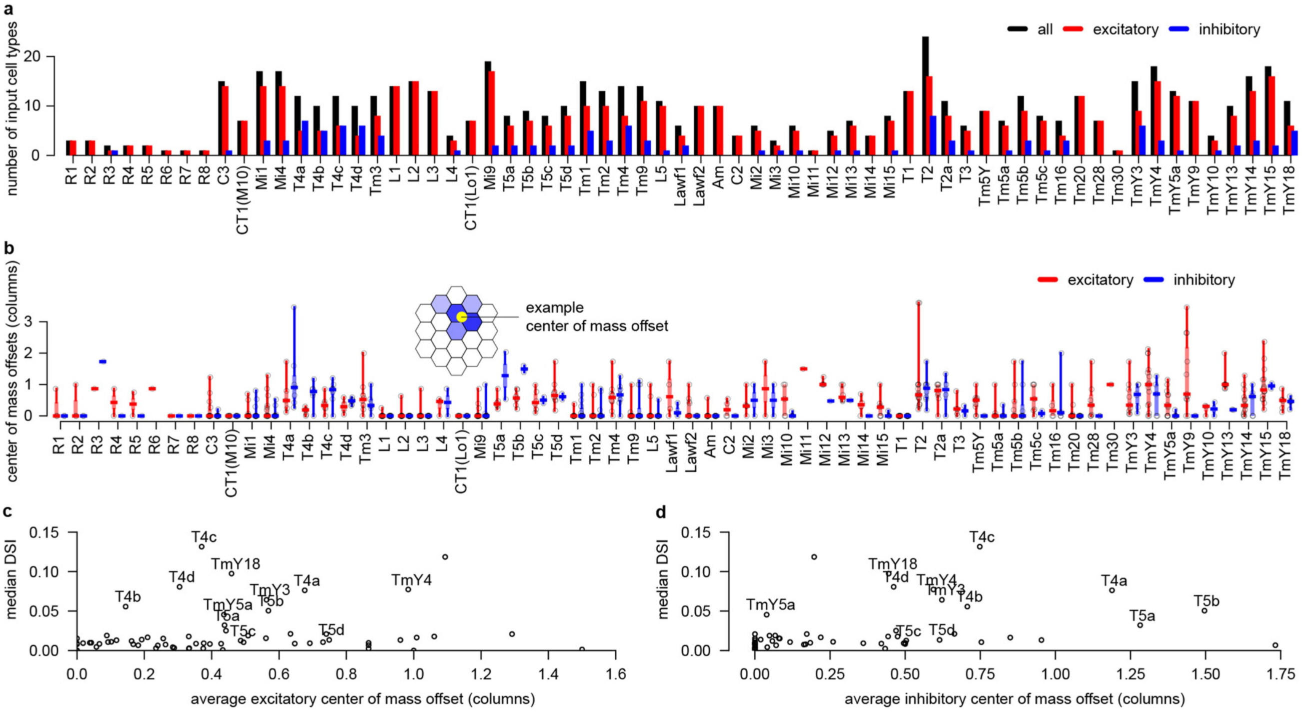
Extended Data Fig. 2 | DMN benchmark of connectomic constraints. (**a**-**d**) **How would incomplete knowledge of connectome affect the tuning predictions?** We artificially varied DMNs with random parameters, connectome-constrained or task-optimized parameters. Five experiments: Four 'Synapse-optimized models', one 'Fully optimized'. Details in Methods. **How would incomplete knowledge of cell types affect the tuning predictions?** We artificially assumed some cell types to be indistinguishable, with shared physiological parameters (resting potentials, time constants, and unitary synapse strengths). Two experiments: (1) 'Full DMN Merge T4, T5' assumes that T4 and T5 subtypes were indistinguishable, reducing the number of cell types to 58. (2) 'Full DMN Merge E/I' assumes that we had three cell types, excitatory (37 cell types), inhibitory (22 cell types) or both (4 cell types), based on our knowledge of synapse signs. Tuning predictions are shown in comparison to the Full DMN and the DMN with random parameters. (**a**) Task error. (**b**) Predicted correlations to flash response indices, T4, and T5 motion-tuning curves (10 best models). (**c**) Predicted correlations to known direction selectivity indices. (**d**) Distances between known preferred directions and predicted preferred directions for T4 and T5 neurons. (**e**) **Better task-performing models predict motion-tuning neurons better.** We correlate predicted tuning metrics from each model to the known tuning properties to understand when better performing models give us better tuning predictions. **(orange)** When correlating the direction selectivity index of each model to the binary known properties for T4 and T5 and their input cell types, we find that this correlation is higher for better performing models (Pearson correlation, $r = \ \ 0.60$, p = $2.6 \times 10^{\ 6}$, $t = r\sqrt{\frac{\mathrm{df}}{1 \ \ r^2}}$, 95% CI = [1, 0.42], df = 48). **(magenta)** While the models predicted the known contrast preferences generally well, the correlation of flash response index to the binary known contrast preferences of 31 cell types did not significantly increase with better performing models.



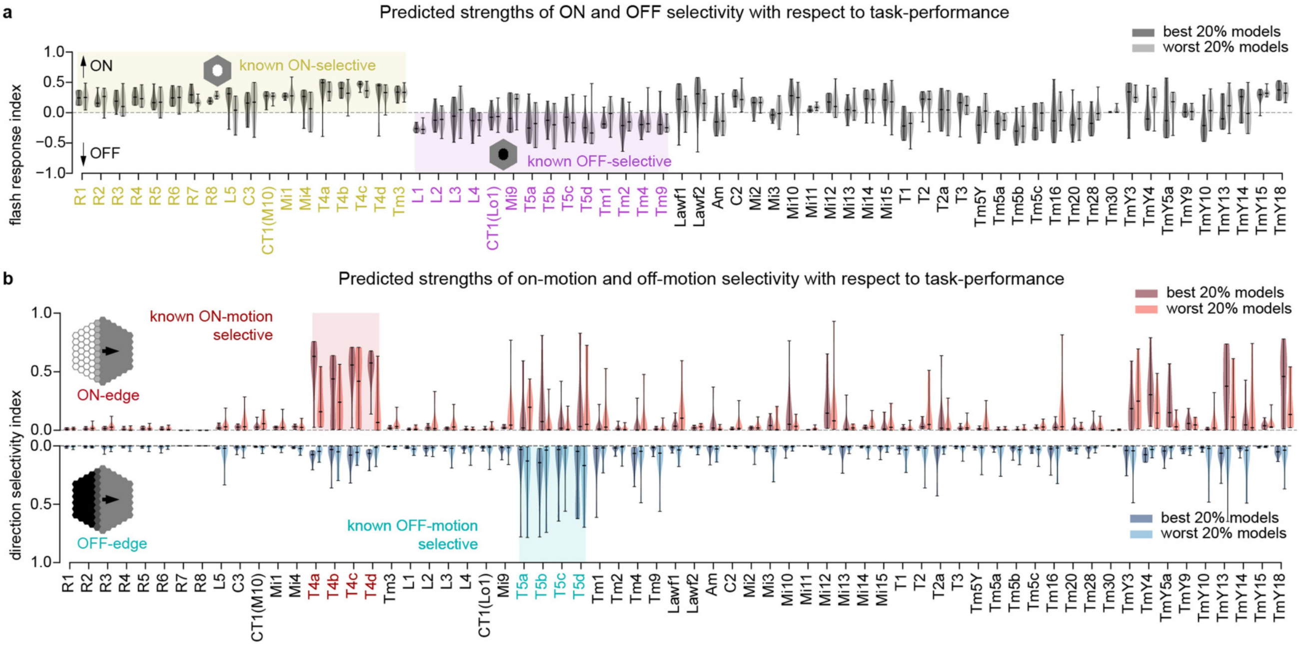
Extended Data Fig. 3 | DMNs suggest that TmY3 neurons compute motion independently of T4 and T5 neurons. (**a**) We clustered 50 DMNs after performing nonlinear dimensionality reduction of their responses to naturalistic scenes for each cell type, and aimed to identify whether clusters correspond to qualitatively different tuning mechanisms. (**b**) Dimensionality reduction on TmY3 responses to naturalistic stimuli reveals 4 clusters of DMNs with average task errors 5.298 (circle), 5.317 (triangle), 5.328 (square) and 5.331 (star). Across clusters, TmY3 shows different strengths of direction selectivity (evaluated with moving edge stimuli). ON-edge direction selectivity is strong in the first and the third cluster. (**c**) Normalized peak responses of TmY3 to moving edge stimuli in the DMNs of each cluster. (**d**) Major cell types and synaptic connections in the pathway that projects onto TmY3 (simplified). (**e**) The input elements of TmY3 with the highest amount of synapses are L4, L5, Tm2, Tm3, Mi1, Mi9, and Mi4. The asymmetries of their projective fields could allow TmY3 to become motion selective. (**f**) Dependencies between TmY3 tuning and the contrast preference of its input cells. For clusters in which TmY3 is motion selective, cluster 1 (TmY3 tuning to downwards/front-to-back motion, circular marker) indicates ON-selectivity for Tm3, Mi1, and Mi4 cells, and OFF-selectivity for L4, Tm2, and Mi9 cells, in agreement with known selectivities. In contrast, cluster 3 (TmY3 tuning to upwards/back-to-front motion, square marker) indicates ON-selectivity for Mi9 in contradiction to the known selectivities and hence ruling out the third TmY3 tuning solution.



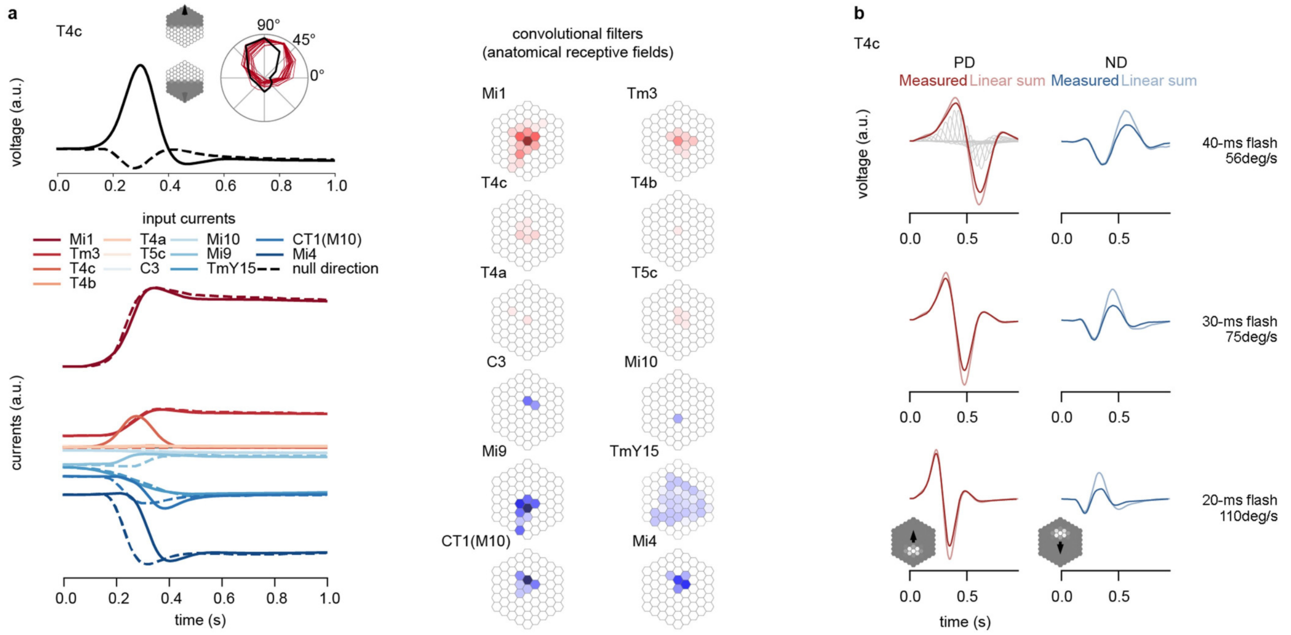
Extended Data Fig. 4 | TmY3 motion detection mechanisms hypothesized by the model. (**a**) Responses to PD and ND ON-edge motion and contributions from input elements. (**b**) PD enhancement and ND suppression in TmY3 in the model.



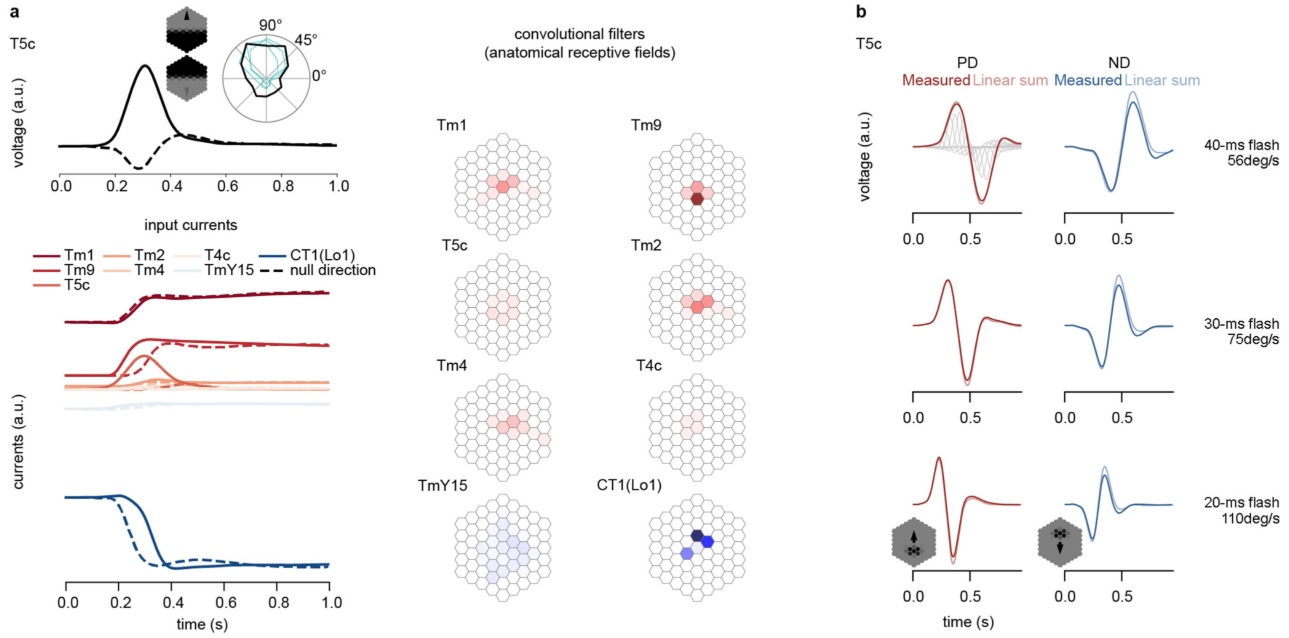
Extended Data Fig. 5 | Statistics of inhibitory and excitatory synapse inputs. (**a**) Number of input cell types per cell type. (**b**) Center of mass offsets of synaptic input. (**c**) Average excitatory and (**d**) inhibitory center of mass offset of synaptic inputs against median predicted direction selectivity index for all cell types. Datapoints for cell types that were predicted as significantly motion selective are labeled.



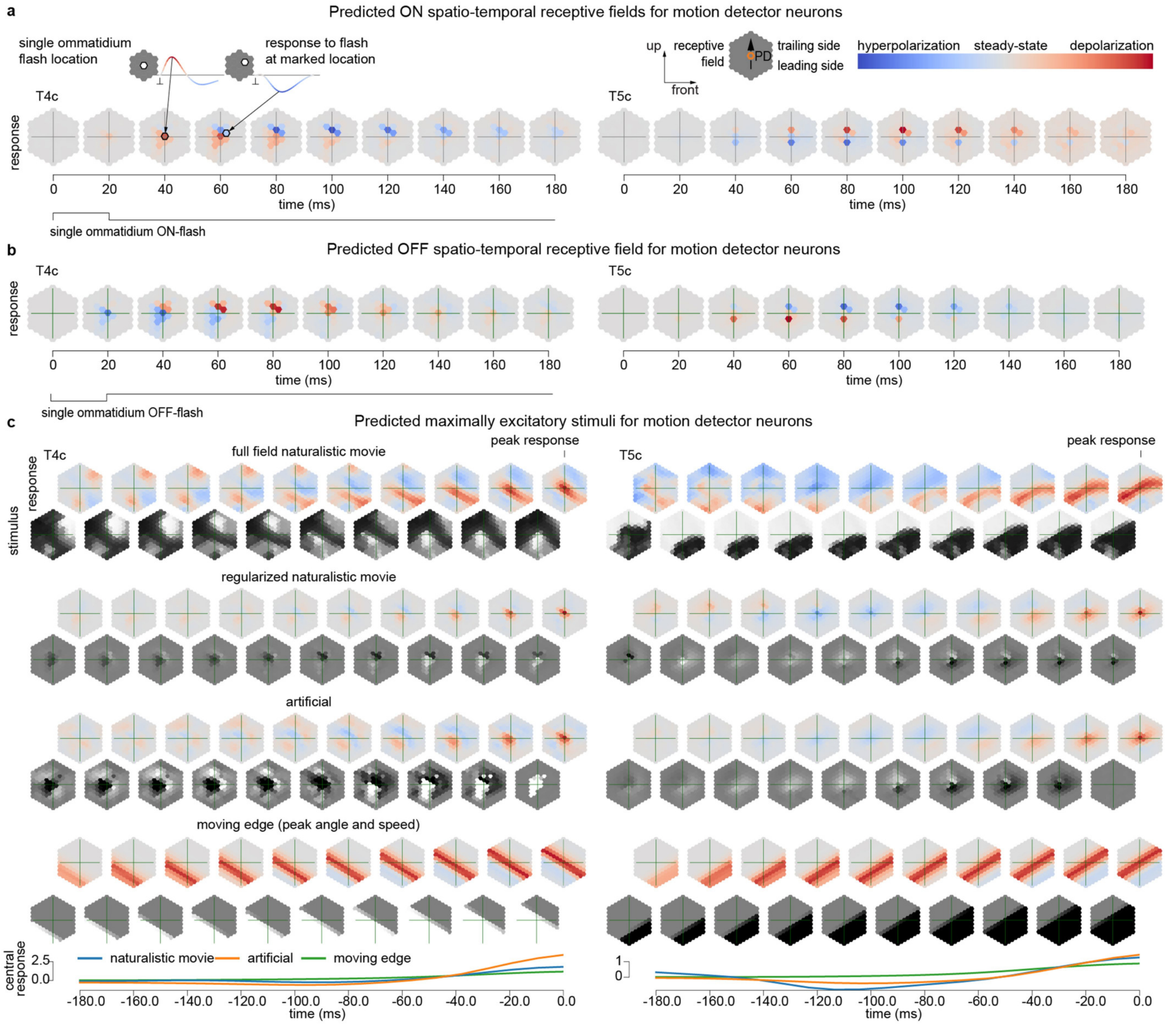
Extended Data Fig. 6 | Predicted tuning with respect to task-performance. (**a**) Flash response index computed as the max-abs-scaled peak response to an off flash subtracted from the max-abs-scaled peak response to an on flash – both of approximately 35° radius and presented for 1 s after 2 s of grey input. Values above 0 indicate on-polarity, values below zero indicate off-polarity. Known on-polar and off-polar cell types are colored in yellow and magenta. (**b**) Single-cell type direction selectivity of best 20% task-performing models versus worst 20% task-performing models of an ensemble of 50 models as a result of peak voltage responses in central columns to on-edges and off-edges moving towards all possible directions on grey background.



Extended Data Fig. 7 | T4 motion detection mechanisms hypothesized by the model. (**a**) Across all T4 cell types (here T4c, Supplementary Fig. 5 for other T4 types), our model predicts that T4 depolarization in response to PD ON-motion (black, solid) is driven by excitatory Mi1 current inputs (darkest red, solid) from roughly a two-column radius of Mi1 cells. Excitatory inputs from neighboring T4 cells of the same type increase the T4 PD-motion response (third darkest red, solid). Tm3 and Mi1 cells excite T4 agnostic to PD vs. ND motion. For ND-motion, Mi4 cells cancel excitatory currents from Mi1 with matching inhibition from the trailing side of the receptive field (darkest blue, dashed). The inhibition from Mi4 cells is delayed for PD-motion (darkest blue, solid), allowing strong depolarization of T4. CT1 shadows the Mi4 mechanism with similar but weaker inhibition from the same location of the receptive field (second darkest, blue). Our model suggests mechanisms involving Mi9 cells and TmY15 cells: both contribute to T4 motion detection by different inhibitory mechanisms for PD-motion with respect to ND-motion. (**b**) 'Measured': Predicted T4c responses to bars moving in PD (left column) and in ND (right column) at varied speeds (saturated red and blue). 'Linear sum': linear sum of responses to individually flashed frames that constitute the moving bar video (faint red and blue). Faint grey traces in background of first panel show individual flash responses before linear summation. Flash duration in each location matched to length of stay at the location in the moving bar video. Bars were approximately 9° wide and 20.25° high and moved across 45° with respect to receptive field in the center. This figure should be compared to Gruntman et al.[27], Fig. 4f.



Extended Data Fig. 8 | T5 motion detection mechanisms hypothesized by the model. (**a**) Across all T5 cell types (here T5c, Supplementary Fig. 6 for other T5 types), our model predicts that T5 depolarization in response to PD OFF-motion (black, solid) is driven by excitatory Tm1 and Tm9 input currents (darkest and second darkest red). Tm1 currents come from a centered, two-column radius of Tm1 cells. Tm9 inputs come from cells offset by one column towards the leading side of the receptive field. We observe delayed excitation from Tm9 cells for ND-motion. The PD-motion response is increased through excitatory inputs from the neighboring T5 cells of the same type (as for T4 cells), not providing excitation for ND-motion. CT1(Lo1) cells cancel excitatory currents with strong inhibitory currents from the trailing side of the receptive field leading to the weak ND response. For PD-motion, inhibition from CT1(Lo1) cells is delayed allowing strong T5 depolarization. (**b**) 'Measured': Predicted T5c responses to bars moving in PD (left column) and in ND (right column) at varied speeds (saturated red and blue). 'Linear sum': linear sum of responses to individually flashed frames that constitute the moving bar video (faint red and blue). Faint grey traces in background of first panel show individual flash responses before linear summation. Flash duration in each location matched to length of stay at the location in the moving bar video. Bars were approximately 9° wide and 20.25° high and moved across 45° with respect to receptive field in the center.



Extended Data Fig. 9 | Spatio-temporal receptive fields mapped with ON- and OFF-impulses and maximally excitatory stimuli. (**a**) Spatiotemporal receptive fields for motion detector neurons agree with experimental measurements (Gruntman et al.[27]). (**b**) Spatio-temporal receptive field mapping with single ommatidium OFF-impulses. (**c**) Maximally excitatory stimuli and baseline-subtracted responses. Including full-field naturalistic, regularized naturalistic, artificial, and moving edge stimuli and responses. Moving edge angle and speed maximize the central cell peak response. Artificial stimuli are optimized from initial noise to maximize the central cell activity using gradient ascent plus full-field regularization towards grey. The last row shows the baseline-subtracted central cell responses. Peak central cell responses at time point zero.

4

EXTENDED RESULTS I: BUILDING BETTER MODELS

In this chapter, I emphasize new opportunities, discuss advanced modeling techniques, and share insights gained from model development.

Part of the work presented here was created in close collaboration with a team in the Mackelab. The active, collaborative teamwork has started with L. Ulmer who joined us on the project as a M.Sc. student in March 2023, around the time when we preprinted the manuscript. Over the past two years, as part of this team usually consisting of four or five members, I contributed much of my work time to supervising and guiding peers, establishing shared software and infrastructure, and fostering a collaborative research culture aimed at generating quality-controlled modeling results.

That said, many results are still work-in-progress such that I can only provide outlooks rather than preliminary results.

4.1 TOWARDS NEW CONNECTOME MODELS

Can our techniques for connectome-constrained simulation be applied to other nervous systems?

The advent of full nervous system reconstructions of the fly (Section 2.3) shortly after we published our techniques for connectome-constrained simulation, optimization, and hypothesis generation represents a wealth of opportunities for testing the connectome-constrained modeling approach and make new model-driven neuroscientific discoveries in the fly. Initial comparable modeling results are fruitful. For instance, a model of a portion of the central brain underlying antennal grooming behavior suggests new circuit motifs *(Özdil et al., 2024)* and a model of the head-direction circuit underlying self-motion perception demonstrates that self-supervised and unsupervised training can be viable alternatives to supervised task-optimization of neural parameters *(Duan et al., 2025)*. Others have created a simulation of most of the connectome and predict behaviorally relevant neural computation by simply using a leaky integrate-and-fire neuron model with manually set parameters without any machine learning *(Shiu et al., 2024)*.

The new vastness of connectome data allows unprecedented analyses, even across individuals and sex. Over the next decade, while new connectome data is on the way, many novel structure-function relationships underlying behavior will be discovered.

4.1.1 *Towards a whole fly nervous system model*

Given the recent data, we can make first estimates of how many unknown parameters a DMN of the whole fly nervous system has.

To simulate the whole nervous system, we would have to simulate about 140,000 neurons for the central brain and 23,000 neurons for the ventral nerve cord (Fig 4.1a, Section 2.3). A DMN with 163,000 neurons and of the form presented here (Section 2.4) will have an order of magnitude more parameters than the visual system model. Compared to the 64 neuron types of only the fly visual motion pathway we modeled, the whole brain has about 8,500 neuron types *(Schlegel et al., 2024)* and the VNC about 110 *(Allen et al., 2020)*. A full nervous system DMN will therefore have at least 17,220 unknown parameters for neuron type time constants and resting potentials.

Compared to the 1,500,000 synapses of our fly visual system model, a model of the whole nervous system has approximately 140,000,000 chemical synapses (Section 2.3). However, the count of neuron type to neuron type connections determines parameter count rather than individual contacts as long as we can continue to estimate unitary scaling factors for synapse strength on neuron type level. Schlegel et al. find 572,980 neuron type to neuron type connections in the brain of which 16%, 91,677, were found to have a 90% chance of persisting across hemispheres but account for 79% of all synapses. In other words, we would require something between 91,677 to 572,980 unitary scaling factors for the brain of the simulation, depending on cut-offs, and likely a couple hundred more for the VNC.

Therefore, the total count lies between $108,897$ and $540,200$ required unknown parameters in a DMN of the whole nervous system — two to three orders of magnitude higher than in the present model with 735 unknown parameters. In this context, it is somewhat surprising that a model of the whole-brain connectome with only 10 distinct parameters could make accurate predictions for a part of the system *(Shiu et al., 2024)*. In our model, we found that we can predict individual subtype selectivity across T4 and T5 subtypes well and potentially even better when we constrained them to share parameters (Fig 4.1b). Taken together, this suggests that we can likely share parameters more coarsely, for instance across cells of the same neuron type family or hemilineage, which would reduce the amount of parameters needed to simulate the system.

What are the memory requirements for such a model? Training a single model (from an ensemble) at 20 ms integration timesteps currently requires about 3GB of GPU memory, or about 0.004 GB per free parameter. With the above parameter count, and the calculated ratio, assuming task-independent requirements, we can estimate a full nervous system model coarsely requires 435 to 2161 GB GPU memory, equivalent to 6-27 Nvidia A100 graphic cards. A costly but feasible hardware requirement.

For a statistically powerful characterization of possible structure-function relationships (Section 2.2.7), one will need to train ensembles of equivalent models to characterize degenerate solutions. Assuming 50 equivalent models would be sufficient (which could be an underestimate), a whole nervous system DMN will already require 271 to 1,350 A100's for training one ensemble in parallel. From our previous experience, we can roughly estimate the speed-up that model-based function characterization per neuron type offers. This is a strong oversimplification, might miss data points, and should be taken with a grain of salt:

Over the last three decades, researchers who made the 26 studies we cited centrally in our paper, contributed to characterizing the function of 32 neuron types. This represents about 1.07 neuron type function characterizations per year for the whole system. With our connectome-constrained model, we characterized 64 neuron types within approximately 5 years, or in other words we made 13 neuron type characterizations per year — an order of magnitude increase over the previous number and others will be faster in future. While a model hypothesis is not the same as an actual experimental measurement, we will be able to generate functional hypotheses for hundreds to thousands of neuron types a year with future connectome-constrained models. Likely no human researcher will be able to manually screen all these model hypotheses, but also not validate them with experimental measurements in the animal. While coarse, these simple calculations suggest that and why the ambition of brain modeling will require reliable machine learning methods and analysis software for massive industrial-type scale and automation.

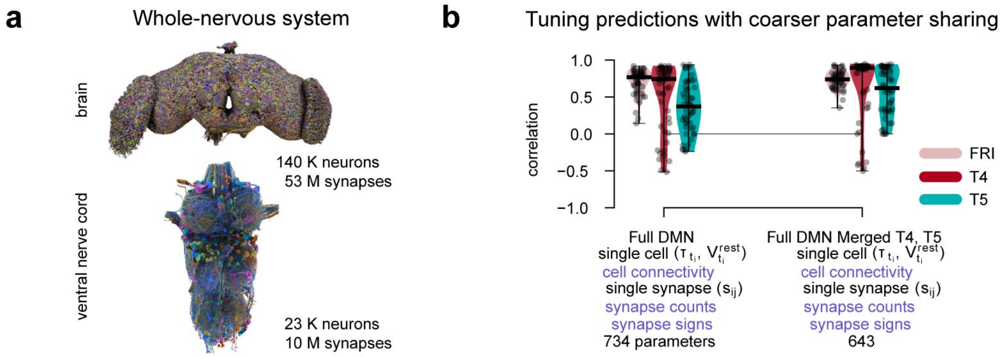


FIGURE 4.1: **Towards whole-nervous system understanding with model-driven function discovery on neuron type level. (a)** Neuronal wiring diagrams of the adult *Drosophila* brain and ventral nerve cord, adapted from *(Dorkenwald et al., 2024)* and *(Takemura et al., 2024)*. *Brain* refers to the two optic lobes on the left and right and the central brain in the center. **(b)** Tuning predictions of two different DMN ensembles with old connectome (50 models each). "Full DMN" infers time constants, resting potential, and synaptic scaling factors for each T4 and T5 subtype separately. "Full DMN Merged T4, T5" shares these parameters across T4 and T5 subtypes. Higher correlation means better predictions.

4.1.2 *Towards new optic lobes models*

As a second step towards the whole-connectome simulation, we are actively working on simulating the new optic lobe connectome *(Nern et al., 2024)*, now one of several comprehensive reconstructions *(Dorkenwald et al., 2024; Schlegel et al., 2024; Matsliah et al., 2024)*. This represents its own opportunities and challenges.

The new male connectome dataset (Fig. 4.2a) comes at isotropic resolution of 8 nm, with complementary mesoscale light microscopy images and is densely reconstructed *(Nern et al., 2025)*. However, truncations of retina and lamina represented the first challenge for faithful connectome-constrained simulation (Fig. 4.2b). In preparation for simulation, one has to manually reconstruct these parts of the fly visual system based on partial EM reconstruction and mesoscale anatomical data from light microscopy imaging.

Big neurons represent the second challenge. Neurites of most neuron types contained in the motion circuit connectome are local, but neurites of interneurons and neuron types downstream of the motion detection circuit often span many visual columns. 50% of the anatomical receptive fields in the model with the old motion pathway connectome have a radius of less than two columns. The neuron types from this model still account

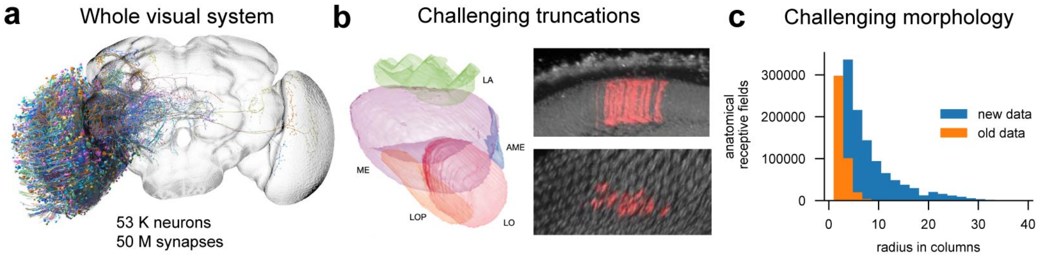


FIGURE 4.2: **New optic lobe. (a)** Neuronal wiring diagram of the left optic lobe. Image adapted from FlyEM / HHMI Janelia Research Campus, licensed under CC BY. **(b)** Truncated optic lobe volume and complementary light microscopy data, adapted from *(Nern et al., 2025)*. **(c)** Radius (maximum extent) of anatomical receptive fields of the old fly visual system data limited to the motion detection pathways vs. the new whole-visual system data.

for 67.4% of cells in the optic lobe[1]. In contrast, with the full visual system connectome, 50% of all anatomical receptive fields extend beyond five columns, with a long tail corresponding to connections spanning tens of columns (Fig. 4.2c). Note, that these numbers might vary depending on the applied synapse thresholds. We could previously model columnar terminals of the large amacrine cell CT1 that span the entire field of view as independent columnar point neurons, because they are electrotonically compartmentalized *(Meier & Borst, 2019)*. However, it is unknown which other neurons are compartmentalized too or for which neurons morphological properties support function and are thus required in modeling. One might be able to guess from EM data based on anatomical features whether terminals are electrotonically compartmentalized (S. Seung private communication), however a comprehensive characterization would be useful.

The third challenge are spiking neurons. While many neurons in the fly visual system communicate via graded potentials, certain neuron types are known to spike, but this has not been catalogued extensively. A comprehensive dataset reporting spiking vs. non-spiking neuron types would help for combined modeling of spiking and graded potential neurons within the same neural network.

New models might represent point neurons with graded potential together with spiking multi-compartment neurons and vice versa. This poses challenges to numerical integration and optimization. Progress in large-scale and automated modeling will therefore hinge upon comprehensive characterization of these properties and better machine learning tools for nu-

---

[1] tinyurl.com/flyvis

merical integration and optimization. Good computational infrastructure can compensate prospective demands for simulating heterogeneous graph-neural networks of stiff differential equations. While facing significant challenges, I think as demonstrated by recent work on connectome analysis and simulation *(Lappalainen et al., 2024; Seung, 2024; Shiu et al., 2024; Seung, 2025; Zhao et al., 2024)*, the scientific opportunities for discovery are exciting and we have more questions than ever before — which justifies the expenditure.

4.2 TOWARDS MODELING BETTER NEURAL AND SYNAPTIC DYNAMICS

The main work describes electrical signals in neurons as passive-point voltages and chemical synaptic input as instantaneous graded potentials (Section 2.1.2). These simplifications have functional limitations. This section presents extensions to the DMN framework: conductance-based synapses, morphologically detailed active neurons, and integration of neural activity measurements.

4.2.1 *DMNs with conductance-based synapses*

The molecular machinery at a real synapse is complex as its local molecular mechanisms, such as ion channel trafficking, interact with more global mechanisms such as voltage-dependent ion channel opening, and consequences for network function and behavior are often opaque *(Destexhe & Marder, 2004; Scott & Frank, 2023)*.

Importantly, voltage dependence causes postsynaptic potentials from multiple inputs to interact nonlinearly, rather than summing linearly as in current-based models. The real biological complexity allows the brain to implement interactions of inputs resulting in nonlinear computations in different ways *(Koch, 2004)*.

For instance, the fly motion detecting neurons implement a nonlinear multiplication mechanism: Coincidence of glutamatergic inhibition and cholinergic excitation leads to sharpened directional tuning curves *(Groschner et al., 2022)*. This motivates to model the DMN with conductance-based synapses in future (Equation 2.1). Unlike current-based DMNs, which assume fixed postsynaptic potentials regardless of voltage state, conductance-based models dynamically modulate current flow based on both synaptic input and membrane potential (Section 2.1.1).

Conductance-based models capture multiplicative amplifications *(Groschner et al., 2022)* in ways current-based models cannot, however, demonstration of such effects with conductance-based dynamics trained on the motion detection task is part of future work.

In general, modeling more biological complexity requires to introduce more free parameters which can affect the degenerate solution space in unforeseen ways. More free parameters can introduce more equivalent solutions with the same loss *(Zhao et al., 2025)*. While we have a general strategy to characterize such solution spaces (Sections 2.2.7 and 3.5.6), a theoretical framework to analyze the symmetries and degeneracies induced by these dynamics would help interpret learned solutions.

4.2.2 *DMNs with spiking neurons, morphology, and* JAXLEY

Passive point neurons with graded potentials are a good reduced model of the fly visual system because its neurons are small and potentials are graded. Detailed single neuron simulation of a T4 cell with and without morphological detail results in equivalent dynamics *(Gruntman et al., 2018)*. However, other neural systems implement mechanisms that rely on the precise morphology of individual cells to support specific functions.

Perhaps the most related example to this work where morphological computation is essential is motion detection in the mammalian retina. Different locations in the dendrites of starburst amacrine cells tune to different cardinal directions *(Euler et al., 2002)*, a computation that is implemented in the fly by T4 and T5 subtypes *(Borst & Helmstaedter, 2015)*.

How to model morphological detail? To capture morphology-dependent computations, we need modeling approaches that go beyond point neurons. Cable theory describes voltage propagation over geometries using partial differential calculus. It results in a recipe for how neural morphologies can be segmented and recombined to a morphologically detailed neuron simulation *(Dayan & Abbott, 2005; Rall, 1989)*.

To use a gradient-based optimization approach as in DMNs, such a simulator must be differential. Existing detailed neuron simulators provide substantial functionality to configure detailed networks of neurons with different ODEs but they are non-differentiable *(Hines & Carnevale, 1997; Goodman & Brette, 2008; Gewaltig & Diesmann, 2007)*. Gradient-based training of spiking Hodgkin-Huxley-type models works for single neurons *(Doya et al., 1993)*, but whether connectome-constrained models of morphologically detailed, conductance-based neurons work under gradient-based task-optimization is not well understood. While we built FLYVIS as a solution in Pytorch for connectome-constrained optimization of differentiable networks, its extended features focus more on custom and flexible ensemble analysis, rather than flexible configuration of network structure and detailed synaptic dynamics as in NEURON.

This, among others, motivates a NEURON-like simulator that is differentiable and built with modern deep learning tools. In recent collaborative work, led by M. Deistler, we built a differentiable simulator called JAXLEY that is flexibly configurable and capable of simulating large networks of spiking and morphologically detailed neurons *(Deistler et al., 2024)*. Like NEURON it handles numerically integrating morphologically detailed neurons, but also supports automatic differentiation via JAX *(Bradbury et al., 2018)* for training of neural parameters.

This capability is also relevant for fly brain modeling, since the fly brain combines graded-potential and spiking neurons. Accurate connectome-constrained modeling of the full fly brain therefore requires integrating different neuron and synapse models within a single computational framework.

4.2.3 *DMNs trained on measurements of neural activity*

When building a DMN, the ultimate goal is a testable, model-driven theory for an *unknown* mechanistic computation underlying behavior. The model should therefore "know" all *known* mechanisms. After task-optimization and clustering of the DMN ensemble, we use known mechanisms to filter models to refine their generated predictions (Section 2.2.7 and 3.2.4). However, our differentiable simulator is technically capable to fit directly to measurements of neural activity. Training DMNs on measurements of neural activity enables data-driven improvement of neural and synaptic dynamics.

Such measurements of neural activity are abundant, however, inclusion of them into DMNs face a number of challenges, for instance:

1. Measurements of neural activity have no standards. For instance, different labs use different recording methods from electrophysiology to calcium and voltage imaging. Experimental setups and data formats differ.

2. Measurements of neural activity are sparse. To correctly constrain distributed computations, we may need recordings of the coordinated activations of many neuron types, but only few neurons can be recorded at once with neuron type specificity.
3. Measurements of neural activity are usually responses to simple laboratory stimuli. Most informative for model building would be recordings of responses to naturalistic stimuli, because they broadly cover the support of the natural distribution of inputs a fly encounters. However, most existing measurements are responses to simple laboratory stimuli like flashes or moving bars because they are not made for model building but for mechanistic interpretation.

Our naturalistic task-optimization approach compensates for a lack of recordings of neuron type specific responses, allowing the model to distribute computation that is still plausible on a high-level. But, as we found, task-optimized models are also degenerate at single-neuron level (Section 3.2.4). How to integrate activity measurements into task-optimized DMNs with machine learning?

A practically convenient solution would be to be able to train the DMN to integrate neural activity measurements as they are becoming available, gradually fine-tuning the model. However, as demonstrated in Linda Ulmer's M.Sc. thesis work *(Ulmer, 2024)* (co-supervised by me and Prof. Macke), sequentially learning different objectives (task and neural activity measurements) leads to decrease in performance on the previous objective. This is a problem commonly known as *catastrophic forgetting* *(French, 1999)*. Sequential fine-tuning does not work, but training the DMN simultaneously on multiple objectives leads to DMNs that "know": Training on optic flow from Sintel simultaneously to tuning curves of T4 subtypes from *(Maisak et al., 2013)*, directly results in all models belonging to the cluster that makes the correct tuning prediction *(Ulmer, 2024)*.

However, the aforementioned challenges in terms of the data foundation remain. Moreover, detailed alignment between stimulated real neurons and simulated model neurons is difficult. More robust data foundations and methods will facilitate building a closed-loop between model and experiment in future.

4.3 TOWARDS BETTER OPTIMIZATION

While the theory of DNN optimization is quickly sketched, actually developing DMNs and optimizing their parameters with tasks and measurements of neural activity requires detailed development work. The hypothetical ideal optimization algorithm works equivalently well across different neuron and synapse models, initializations, network structures, numerical integrators, and objectives, such as tasks or neural activity measurements. It allows refining the model with new measurements without training on all data to facilitate building a closed-loop between modeling and experiment. However, in reality, different combinations of modeling and optimization choices result in different DMN solutions. I discuss relevant considerations below.

4.3.1 *Optimization strategies and training challenges*

For building a DMN from scratch, it is advisable to first overfit the network on a single sequence or batch to debug the pipeline. Next, ideally establish the analysis pipeline for accurate prediction of known mechanisms and further mechanistic interpretation. A stable analysis pipeline ensures that effects of varying hyperparameters in training experiments become tractable.

Once a model trains on the task and generates mechanistic predictions, one usually wants to test different conditions as controls. Would the model predict the same mechanism if we were to change X? Many choices go into the modeling, such as neuron and synapse model, initialization, synapse thresholding, numerical integration, loss function, but also detailed technical choices such as learning rate, learning rate scheduling, parameter penalization, gradient clipping, training sequence length, type of optimizer, type of steady-state accomodation, augmentation, parametrization, or decoder architecture (Section 3.5). To characterize the influences of modeling choice, first pick a baseline model to compare to. Then train a new model under only one changed condition a time. While changing more than one condition at once has a chance of resulting in a better model in some metrics, it will preclude to explain why. Note that tracking the effect conditions have on training results requires to train enough equivalent models (a big ensemble) for statistical power against the baseline, because the optimization is stochastic.

In conclusion, optimizing a connectome-constrained DMN and derive predictions from it is hardly self-propelled. One might discover many pitfalls in the process, from unbound parameters (e.g., $\tau_{t_i} < \Delta t$) leading to state divergence to more complex challenges, such as vanishing and exploding gradients (Section 2.2.5) when optimizing over longer sequences. Eventually, rigorous experimentation making only one change at a time can result in a higher-level intuition and confidence of the predictions that are robust or sensitive to modeling choices.

Once our training and analysis pipelines were stable, we found that predictions varied in the details such as preferred directions, but higher-level predictions, such as general T4 and T5 motion sensitivity, or FRI predictability, were robust across differently (but reasonably) configured ensembles, because of connectome and task-constraints (Section 5.1.1).

4.3.2 *Parameter initialization*

How to initialize the DMN for optimization? Parameter initialization plays a critical role for stable network and optimization dynamics.

Training of deep feedforward neural networks common in deep learning relies on specialized weight initialization to enable successful task optimization *(Glorot & Bengio, 2010; He et al., 2015)*. Parameters must be initialized within a functional regime where activations neither explode nor vanish. Effective initialization strategies achieve this by controlling the variance of layer activations.

However, these insights — while empirically validated in feedforward networks — do not directly extend to recurrent continuous-time dynamical systems such as DMNs. These systems can exhibit a wide range of dynamical behaviors *(Strogatz, 2001)*, including nondeterministic and chaotic dynamics. Such behavior depends on the spectral radius (i.e., the largest eigenvalue) of the recurrent weight matrix and its interaction with initializations of biases and time constants.

For the main work, we developed simple heuristics, including to initialize synapse strengths to small values to avoid runaway excitation, biases in the center of the dynamic range of the input to ensure baseline activity, and time constants 2-3 times higher than the integration timestep to balance responsiveness with temporal stability.

As experimentalists do with the real fly, we present gray-scale stimuli to the model to drive it toward a steady state, where $\dot{\mathbf{V}} \approx 0$. The goal is to avoid arbitrary transient dynamics during stochastic gradient descent by always starting from a consistent reference state — neutral to input centered in the dynamic range. However, depending on the initialization, networks may initially exhibit unstable, vanishing, or oscillatory dynamics.

Heuristically, in our model, favorable initial steady-state dynamics resemble exponential growth or decay toward a sustained response, or transient activity followed by a stable plateau — consistent with the bandpass filtering properties of neuron types in the fly visual system (Section 2.3.3).

Models initialized with random parameters such that they exhibited these favorable dynamics, tended towards predicting tuning properties accurately, but less accurately than task-optimized models and we were unable to decode motion from them (Section 3.2.2).

Parameter initializations subtly influence the generated neural tuning predictions after task-training. Characterizing solutions across broadly initialized ensembles accounts for this uncertainty, but in principle without guarantee that the number of models and the spread of the initialization is sufficient to capture the correct solutions. How predicted tuning distributions across an ensemble vary depending on initialization is subject to further investigation. Another future direction can be the development of algorithms that discover good initializations for mechanistic models automatically, for instance through self-supervised pretraining, or train well directly, largely agnostic to initialization.

4.3.3 *Dead neuron types and auxiliary activity targets*

When visualizing whole-network activity through animated movies, we found that individual trained networks often contain a subset of neuron types with zero activity across all naturalistic stimuli, typically around 5% of all types. We call them *dead neuron types*. The specific neuron types that are dead vary across models, and some models exhibit no dead types at all (Fig. 4.3a).

To reduce the number of dead neuron types, we introduced an activity penalty that adapts resting potentials based on whether activity lies below or above a set threshold (Section 3.5.5). This penalty shifts resting potentials to bring postsynaptic voltage into a regime more favorable for ReLU activation, thereby reducing the prevalence of dead neuron types.

Across two ensembles, the activity penalty reduced the percentage of dead neuron types from 3.7% to 2.0% (Fig. 4.3).

Since the activity penalty constitutes an objective independent of the main task objective yet facilitates training, it serves as an *auxiliary target*. More generally, auxiliary targets act as regularizers that guide the network toward desirable dynamical system behavior or target statistics. In this sense, incorporating sparse measurements of neural activity during training alongside the task objective is a form of auxiliary target training.

A future direction is to design auxiliary targets that accelerate training toward better task performance and neural tuning predictions. Such targets could also support pretraining by steering the system toward favorable initial parameter states.

4.3.4 *Tasks and decoder*

An advantage for modeling the fly visual system is its well-defined high-level tasks, primarily to detect motion for navigation, and our extensive knowledge of its implementation (Section 2.3.3). The local reconstructions that we combined into a coherent connectome originate from research that explicitly focused on reconstructing the motion detection pathways for structural analysis (Section 2.3.2). Besides motion, the fly visual system processes color and object shapes and these high-level functions have recently been associated with hypothetical subcircuits in the complete brain reconstruction *(Matsliah et al., 2024)*.

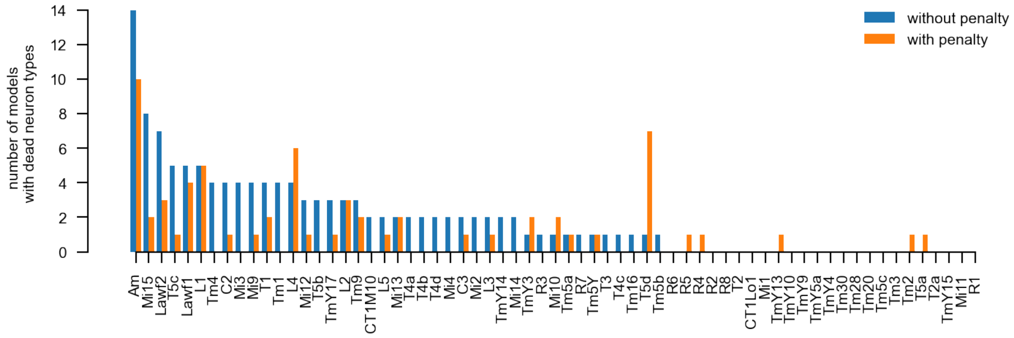


FIGURE 4.3: **Dead neuron problem.** Comparison of dead neuron types across two ensembles of 50 models each trained without activity penalty (blue) or with activity penalty (orange). Sorted from left to right by highest to lowest counts across ensemble without activity penalty.

Our model corresponds to the first stages of the motion detection subsystem. Its circuit structure and high-level function align. How well can connectome-constrained DMNs predict neural activity when the task is misaligned to circuit structure?

To address this question, we trained the DMN on other plausible visual tasks, including the autoencoding of a scene. Tuning predictions from such DMNs were less accurate than tuning predictions from DMNs trained on the motion detection task and as accurate as tuning predictions from models with random parameters (Fig. 4.4).

We extended this analysis in follow-up work with F. Pei and in ongoing work with I. Omolayo. Our results corroborated that only training on motion targets would accurately predict tuning. Additionally, we showed that training DMNs on the same motion targets but with different input types (shapes, gratings, MNIST, textures) would result in better predictions of DSIs than from DMNs with random parameters. However, the choice of the input type strongly influences the directions which T4 and T5 neuron types predominantly tune towards, which cannot be captured by the direction-agnostic DSI. In other words, depending on the input type, predictions fall into different predominant clusters among the degenerate solution space representing different predictions of preferred directions of neurons.

The choice of high-level task is always accompanied by the technical choice of decoder. Typically, task variables are not directly represented in the state space of neural activity. Therefore, decoders translate from neural activity space to task variable representation. For decoder design, we considered four things important:

- The decoder is small, yet allows the combination of connectome-constrained network and decoder to generalize on the task.
- Decoder weights are isotropic at initialization to avoid introducing a structural bias that confounds the connectome-constraint.
- The decoder cannot solve the task alone.
- It decodes from a large superset of neuron types known to be motion selective, to avoid tapping into a circular argument fallacy.

Specifically, our decoder is a CNN with 7,400 parameters. It reads out 34 neuron types with 8 filters, nonlinearity, normalization to improve trainability, dropout for generalization, and a second linear convolutional layer translating to the two dimensional optic flow representation. To prevent

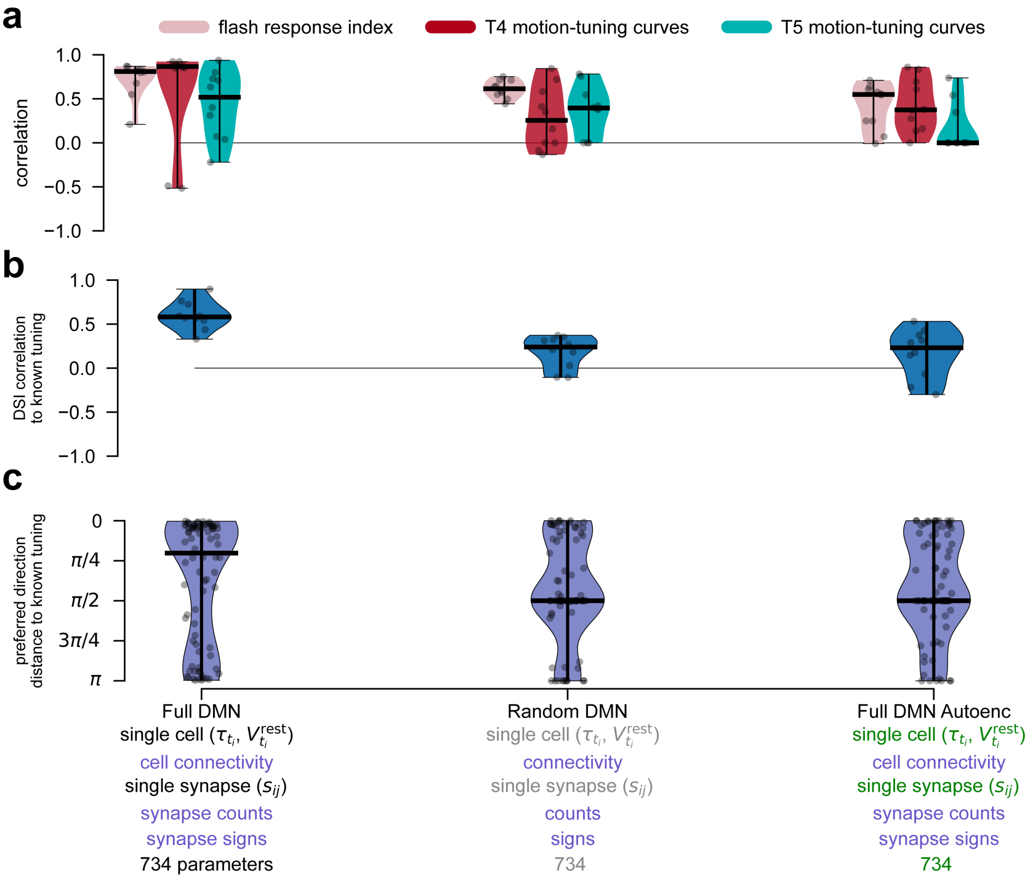


FIGURE 4.4: **Tuning predictions from DMNs trained on motion, with random parameters, and for autoencoding scenes. (a)** DMNs trained on the motion task predict FRIs, T4 and T5 tuning most accurately. Best 20% task-performing models from each ensemble. Each datapoint represents one model of the ensemble. **(b)** The same holds for the predicted direction selectivity index and **(c)** predicted preferred directions.

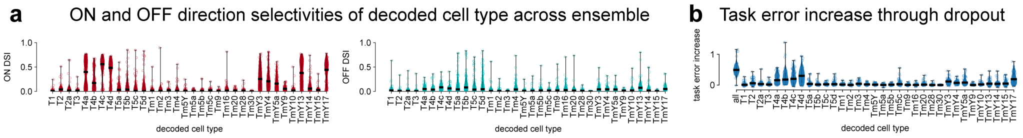


Figure 4.5: **Decoded neuron type influence on task performance. (a)** DSIs for each decoded neuron type across ensemble. (left) ON-edge DSIs. (right) OFF-edge DSIs. **(b)** Increase in task error when neuron type activity is replaced by average response to naturalistic scenes at the decoding stage. As a baseline, in the left-most condition "all", all responses are replaced by the neuron type's average responses such that motion detection is impossible.

that the decoder introduces a structural bias, we homogenously initialize its weights.

To confirm that the trained decoder uses the motion detecting neuron types in the fly visual system model, we analyzed the task error increase when replacing the dynamic neuron type activity with the average responses at the decoding stage (Fig. 4.5). We observed that task error increases for each neuron type, but the largest increase in task loss when replacing T4 activity (Fig. 4.5c). This suggests that the decoder indeed uses the motion detecting neurons and primarily attends to ON-motion selective T4 neurons.

In summary, the task-optimization approach requires careful consideration of the alignment of task and neural circuit. Choices in decoder design can influence the neural tuning predictions. Mechanistic assumptions on the interplay between network and decoder can be tested via post-hoc analysis. While it worked here, we have no general recipe for decoding from connectome-constrained neural networks. Future work should be guided towards a benchmark across different combinations of connectome, task, and decoder networks.

4.3.5 *Parameter degeneracy and hypothesis generation under non-identifiability*

We trained DMNs on naturalistic visual tasks with rich data augmentation, enabling them to generalize the motion detection function across a broad distribution of inputs (Section 2.2.5). The resulting models perform consistently on the high-level motion detection task and reliably reproduce known tuning properties across the ensemble. However, the models are *non-identifiable*: For example, the predicted function of T4c neurons separates into three distinct clusters across the ensemble (Fig. 4.6a,b), each supported by many different trained parameter combinations of T4c time constants, resting potentials, and input scalings (Fig. 4.6c). Clustering and parameter degeneracy across all other neuron types are illustrated in the 478-page supplement accompanying the main publication *(Lappalainen et al., 2024)*.

In summary, despite non-identifiability at the parameter level, functional structure can emerge at the ensemble level. This highlights the need for ensemble-based approaches when using connectome-constrained models to generate mechanistic hypotheses. Rather than aiming for a single "correct" model, our strategy identifies classes of plausible mechanistic solutions - each representing a distinct, testable hypothesis - under fixed structural constraints but variable dynamics.

Such ensemble-based hypothesis generation turns non-identifiability from a modeling limitation into a discovery opportunity. While convergence to different solutions across equivalent runs might be mistaken for suboptimal optimization, it instead reflects the inherent underdetermination of the system. Embracing this variability through ensemble modeling allows us to recover structure in function space - revealing multiple plausible solutions that are consistent with the data and structural constraints, and that can guide experimental investigation in a closed-loop research process with the model.

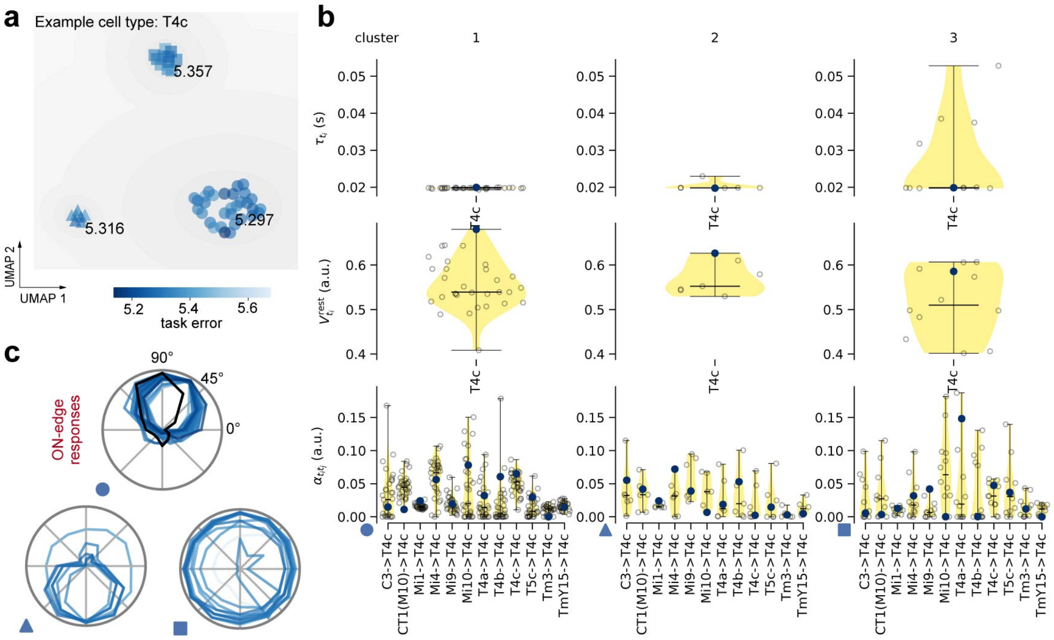


FIGURE 4.6: **Function-space clustering, parameter degeneracy, and hypothesis generation under non-identifiability. (a)** Responses of T4c cells reveal three clusters. **(b)** T4c tuning in the three clusters. Circular marker: upwards tuning (cluster with lowest average task error 5.297; black: known tuning of T4c). Triangular marker: downwards (5.316 error). Square marker: no motion tuning (5.357 error). **(c)** Task-constrained parameters of T4c across clustered models. Clusters along columns and parameter types along rows: time constants, resting potentials, and scaling factors. Blue scatter shows the task-optimal model within the cluster.

4.4 SOFTWARE & INFRASTRUCTURE

In computational science, software and infrastructure define the boundaries of what is feasible, scalable, and shareable. Because the software tools determine what questions we can ask and how quickly we can answer them, they become inseparable from the scientific contribution itself. This section analyzes the software and infrastructure underlying the modeling work in this thesis and how they shaped the research trajectory and enabled key scientific results.

4.4.1 *The research-engineering tension: Infrastructure enables science*

In machine learning for science, we navigate a fast-paced and multi-faceted research environment. Our work involves iterating rapidly between exploring data and models, building infrastructure, hypothesis-driven modeling, analyzing and communicating results. This creates a constant tension between investing in intentional software engineering and prioritizing new experiments for timely results.

The impact of infrastructure on scientific outcomes became clear through specific examples. Early experiments had to be rerun due to missing configuration tracking, delaying insights. More importantly, training single models obscured key mechanistic patterns because of the solution degeneracy (Section 4.3.5). Once we built methods for ensemble training and analysis, these patterns became visible, enabling connectome-constrained models to make consistent and interpretable mechanistic predictions (Section 3.2.4).

Infrastructure development directly enabled several key scientific outcomes (Table 4.1), while delays in scaling or robustness stalled progress.

4.4.2 *From explorative code to refined research tools*

The evolution of our infrastructure needs is best illustrated through two codebases that represent different phases of the research process: First the exploratory development repository DVS-SIM[2], and later a refined public release and the parent for our new models, FLYVIS[3]. This evolution from exploration to communication and collaboration illustrates how infrastruc-

---

[2] https://github.com/TuragaLab/dvs-sim

[3] https://github.com/TuragaLab/flyvis

Table 4.1: Infrastructure decisions and their impact on research capabilities.

Infrastructure Decision	**Motivation**	**Impact on Research**
Tracking experiment configuration comprehensively and flexibly with HYDRA	Prevent experiment loss	Enabled reproducibility and explainability of training runs and analyses due to clear parameter settings
Building ensemble training and analysis pipeline	Scale beyond individual models	Revealed consistent patterns invisible in single runs
Automating analysis and visualization pipeline with PAPERMILL	Analyzing and documenting generated model hypotheses per neuron type	Generated large-scale analysis (Section 7.1) as research tool, aiding pattern discovery
	Running comparisons across differently configured ensembles	Revealed consistent patterns and variability across differently configured ensembles, strengthened confidence in main results
Integrating XARRAY + JOBLIB	Manage high-dimensional data and storage efficiently	Scaled analysis pipeline for faster modeling research and development

ture needs change as research matures. DVS-SIM remains private and is not intended for release, but anyone looking for analysis code missing from FLYVIS can contact us via email.

4.4.2.1 DVS-SIM: *Experimentation and automation*

The DVS-SIM repository documents the project's exploratory phase. Its *code churn (Munson & Elbaum, 2002)* — more than 230,000 lines added and 120,000 deleted (excluding notebooks) — reflect frequent iteration from initial prototype to final result (Table 4.2). The notebook content, close to one million lines across 1,908 notebooks, reflects both manual prototyping and exploration and the computational artifacts created from a custom, parametrically executed ensemble pipeline after running many comparisons.

Among others, the automated analysis and visualization pipeline allowed us to generate a supplement of 478 pages detailing individual neuron type properties (structural and functional) in a systematic way (Section 7.1). We created this not only for providing a comprehensive supplement but heavily used it during the scientific process. The systematic documentation enabled us to identify patterns across models and neuron types that would have been rather difficult to detect through manual analysis, for instance, the relationship between asymmetric and symmetric inputs and the model predictions for motion tuning neuron types across the system. DVS-SIM integrates PAPERMILL (a toolbox for parametrizing and automatically executing Jupyter notebooks) and establishes template notebooks for analysis to avoid duplicating notebooks and running them manually for each model run and for producing standardized analysis outputs. In DVS-SIM, notebooks play a central role not only for prototyping, but also for managing experimental workflows and computationally intensive distributed workloads.

4.4.2.2 FLYVIS: *Communicability and collaboration*

The follow-up repository, FLYVIS, was built for clarity, reusability, and collaboration. The notebook count was reduced from 1,908 to 13 and the Python codebase condensed by two-thirds (Table 4.2), marking a shift from exploratory infrastructure to a focused research tool. We refactored and renamed code for consistency with the publication, emphasizing documentation and usability over feature breadth — a time-consuming but crucial step for scaling and teamwork.

FLYVIS also addressed core limitations of DVS-SIM, particularly in model evaluation and data handling. Evaluating ensembles required managing large, labeled multi-dimensional datasets efficiently. We replaced the custom solution in DVS-SIM with a cleaner backend combining XARRAY *(Hoyer & Hamman, 2017)* for labeled arrays and JOBLIB *(Joblib Development Team, 2020)* for caching and parallel computation. The resulting system returns xarray datasets from native function calls, either computing or loading from disk, and integrates seamlessly with the Python data stack (NUMPY, PANDAS, MATPLOTLIB).[4]

This refactor removed thousands of lines of custom code and established a consistent API for stimulus-response datasets, turning a major bottleneck into a streamlined analysis process. Together with improved documentation, it provides a strong foundation for effective collaboration and future model development.

Table 4.2: Repository statistics showing the evolution from exploratory (DVS-SIM, commit b1eca51e) to communicable (FLYVIS, commit ce555ee) research infrastructure (my contribution percentages in parentheses).

Metric	**dvs-sim**	**flyvis**
Commits	401 (91%)	263 (93%)
Lines added (.py)	230,431 (97%)	101,137 (93%)
Lines deleted (.py)	120,351 (99%)	65,295 (98%)
Current lines (.py)	113,815 (97%)	35,752 (91%)
Notebook lines (code/markdown)	973,713 (100%)	2,594
Notebook files with commits	1,908 (100%)	13

4.4.3 *Collaborative research through processes and infrastructure*

Developing this infrastructure required iterative refinement and occasional missteps (e.g., initially overlooking XARRAY), but ultimately established a robust and collaborative development process. Automatic GitHub workflows, issue tracking, and code review enables us now to make distributed contributions and perform reproducible analyses.

---

[4] Thanks G.-J.B. for the hint.

We created the GitHub organization *flyvis* intended for inter-institutional collaboration on our software tools for connectome-constrained modeling. In addition, we published some self-contained software byproducts of this work, such as DATAMATE[5] for experiment data tracking and MPLIMATE[6] for creation of animations with matplotlib, which have proven useful beyond the main project.

Complementary to the main theme, I could contribute with software and infrastructure concepts to JAXLEY (Section 4.2.2). For instance, for finding better ways of optimizing the MNIST network with over 100,000 parameters, we used DATAMATE for tracking step-by-step cause and effect of configuration changes and model improvements, and eventually finding good model solutions. This demonstrates how following systematic processes and infrastructure approaches can facilitate contributions across computational research.

Lastly, software projects vary fundamentally in nature — and it is easy to accidentally dismiss the different nature of computational research projects and their requirements. To provide two differing examples (but many examples exist): JAXLEY represents a software project with a clear technical scaffold, scope, and goal — in part inherited from NEURON. FLYVIS distills step-by-step model training and analysis from originally many design artifacts from DVS-SIM to understand complex cause and effect relationships in building connectome-constrained models.

4.4.4 *Lessons for computational research*

The following lessons, though not claiming to be exhaustive, summarize how infrastructure choices directly influenced scientific progress in this project.

Ensemble infrastructure reveals patterns invisible in individual runs. Our experience shows that training individual models can be misleading — the degeneracy and consistency patterns that enabled mechanistic predictions only became apparent when we built infrastructure to systematically analyze ensembles. Scalable, automated approaches can fundamentally change what we see and discover.

---

[5] https://github.com/flyvis/datamate

[6] https://github.com/lappalainenj/mplimate

Scale and document systematically, not just finally. The 478-page automated supplement became a research tool that enables pattern recognition across models and neuron types. Infrastructure that generates systematic result documentation during the research process can accelerate discovery itself and provide unexpected analytical capabilities.

Configuration tracking enables reproducibility and explainability. Systematic tracking ensures reproducibility and explainability, preventing setbacks by being unable to recreate insightful or promising preliminary findings.

Understand the project and software requirements. A neural simulator extending established paradigms has different software requirements than a framework for exploratory connectome-constrained modeling. One benefits from defined scopes, the other requires the right balance of flexibility and robustness in designs that can evolve as the scientific questions themselves develop. Recognizing such distinctions allows aligning design decisions and timelines constructively with project goals.

Leverage existing tools early. Before building from scratch, research existing alternatives. Established open source libraries such as PYTORCH *(Paszke et al., 2017)*, NUMPY *(Pedregosa et al., 2011)*, MATPLOTLIB *(Hunter, 2007)*, and SCIKIT LEARN *(Pedregosa et al., 2011)* and many others provide more robust foundations than custom solutions, even when they require learning new approaches. By learning to build efficiently on established tools, one can improve code quality and contribute to the broader community through existing open source libraries. Time invested in understanding existing tools can pay dividends in reduced future maintenance and better integration with the broader scientific ecosystem. Eliminating thousands of lines of code by combining XARRAY *(Hoyer & Hamman, 2017)* and JOBLIB *(Joblib Development Team, 2020)* in a novel way for caching high-dimensional labeled tensors on the disk illustrates this broader principle.

Infrastructure decisions shaped the questions we could ask and how quickly we could answer them. Moving from individual model results to ensemble methods exemplifies how scalable methods can make the difference between scientific dead ends and breakthroughs. Computational research therefore requires deliberate attention to software engineering practices, institutional support for such efforts, and recognition that building sustainable infrastructure is itself a scientific contribution.

5

EXTENDED RESULTS II: INTERPRETING STRUCTURE-FUNCTION RELATIONSHIPS

This second extended results chapter investigates open questions and techniques for extracting discoveries and mechanistic insights from connectome-constrained models.

5.1 WHICH CONSTRAINTS MATTER MOST FOR NEURAL PREDICTIONS?

The success of deep learning approaches to high-level brain modeling in the absence of connectivity data (Section 2.2.6), suggests evaluating the method validity when actually integrating biological structure — and more generally to ask which constraints are (more) important for predicting neural activity — structural or functional constraints?

5.1.1 *Dissecting connectome and task constraints for neural activity predictions*

DMNs combine the top-down training on a deep learning task with the bottom-up construction of the network from the connectome and mechanistic equations for neural activity (Section 2.4). Our results demonstrate that networks whose parameters are connectome-constrained and trained to perform the circuit task accurately predict neural activity across the literature (Section 3.2.2).

For understanding the relative importance of different constraints, we developed additional weight expressions (Table 5.1) of the synapse weights (Equation 2.3) to integrate constraints selectively for comparison groups we found reasonable. The results from Section 3.2.3 are summarized here.

The constraints from the connectome can be separated into four elements: Neuron-type connectivity, synapse signs, single-neuron connectivity, and synaptic counts. All our comparisons integrate knowledge of neuron-type connectivity, because our evaluation metrics rely on predicting function based on the identified neuron types. We evaluated five recombinations (Table 5.1), but ruled out evaluating the case in which one knows the synapse

#	**Synapse signs**	**Single-neuron connectivity**	**Synapse counts**	**Task-trained constraint(s)**	**Weight expression**
1	Known ($\sigma_{t_i t_j}$)	Known ($\Delta u, \Delta v$)	Unknown	Synapse counts and unitary scaling	$w_{ij} = \sigma_{t_i t_j} \mathbf{m}_{t_i t_j \Delta u, \Delta v}$
2	Known	Unknown	Unknown	Average convolutional filter*	$w_{ij} = \sigma_{t_i t_j} \mathbf{m}_{t_i t_j, -3 < \Delta u, \Delta v, \Delta w < 3}$
3	Unknown	Known	Known ($N_{\ldots}$)	Synapse sign (Dale's principle) and unitary scaling	$w_{ij} = \boldsymbol{\alpha}_{t_i t_j} \boldsymbol{\sigma}_{t_j} N_{t_i t_j, \Delta u, \Delta v}$
4	Unknown	Known	Unknown	All weights	$w_{ij} = \mathbf{w}_{t_i t_j \Delta u \Delta v}$
5	Unknown	Unknown	Unknown	All weights	$w_{ij} = \mathbf{w}_{t_i t_j, -3 < \Delta u, \Delta v, \Delta w < 3}$

Table 5.1: DMN with different connectome constraints for testing importance of constraints for neural activity predictions. All models include neuron-type connectivity; shown are different combinations of constraints and which were task-trained. *Average convolutional filters represent the anatomical receptive fields of neuron i of type t_i.

count but does not know the single-neuron connectivity, because in practice we know the mesoscale single-neuron connectivity when we know the synapse counts from nanoscale measurements.

We found that DMNs with full connectome constraints and task-training predict neural activity most accurately (with median correlation: 0.82, Fig. 5.1), suggesting both the connectome and functional constraints from the task are necessary for the best predictions. DMNs with random parameters instead of gradient-based task-optization predict neural activity worse (median correlation: 0.42). Note that the distribution parameters must still be chosen reasonably for random parameter sampling. In contrast, task-trained DMNs with only one less element of connectome constraints, predict neural activity even worse than models without task-training: DMNs without synapse counts result in a median correlation of 0.253 (DMN #1), and without signs in a median correlation of 0.322 (DMN #3). DMNs without connectivity and synapse counts but with signs result in a median correlation of 0.326 (DMN #2), indicating that the signs of connections are the most constraining element besides neuron-type connectivity. Finally, task-trained DMNs without signs and synapse counts or additionally without single-neuron connectivity hardly predict any neural

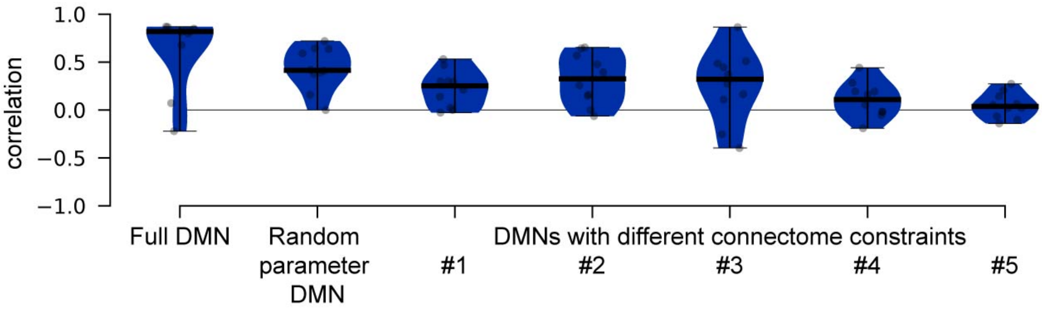


FIGURE 5.1: **Predicted neural activity correlations to known tunings across seven configurations.** *Full DMN* refers to an ensemble of models with full connectome constraints and task-training. *Random parameter DMN* uses full connectome constraints but random parameters instead of task-training. *DMNs #1-5* use only selected connectome-constraints listed in Table 5.1. Each scatter point represents one of the top 20% best task-performing models of each ensemble. Correlations are median values across FRI, T4 and T5 tuning correlations as specified in Section 3.2.3.

activity (median correlation: 0.109, DMN #4; median correlation: 0.075, DMN #5).

Our result highlights the importance of both bottom-up structural constraints and top-down functional constraints in theoretical models of neural circuits for making mechanistic predictions at the single-neuron level. However, based on this analysis and results from *Shiu et al. (2024)*, I suggest that for making predictions of neural mechanisms one should start with evaluating connectome-constrained networks with manually selected or randomly sampled parameters before developing the deep learning machinery for task-training.

5.1.2 *Is the brain's wiring task-optimal?*

Among others, evolutionary pressures, such as survival and reproduction, shape how biological neural networks must function. In contrast, in machine learning, we minimize specific mathematical expressions to shape artificial neural networks for task performance. What constitutes optimal neural network design? While we cannot answer this question with this model, our results provide a partial answer.

We trained several neural networks with selected connectome constraints or without connectome constraints on the same task and compared their task error (Fig. 5.2).

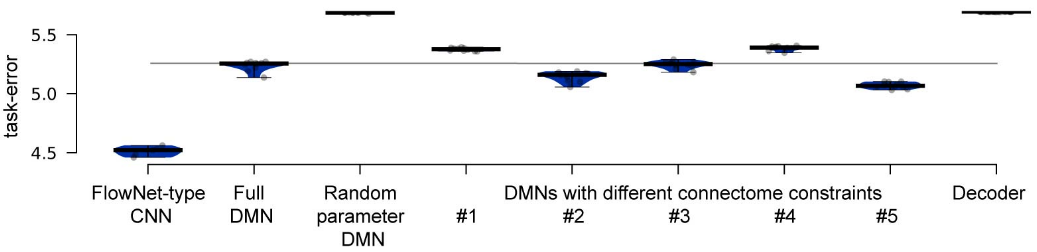


Figure 5.2: **Task error comparison across nine model types.** Lower is better. *FlowNet-type CNN* refers to an ensemble of deep hierarchical CNNs with 414,602 parameters that process two subsequent frames to predict the optic flow. As previously, *Full DMN* refers to an ensemble of models with full connectome constraints and task-training. Grey line indicates its median task error. *Random parameter DMN* uses full connectome constraints but random parameters instead of task-training. *DMNs #1-5* use only selected connectome constraints listed in Table 5.1. For DMNs, each scatter point represents one of the top 20% best task-performing models of each ensemble. *Decoder* refers to an ensemble of the CNN that decodes the DMN activity.

We found that the deep hierarchical *FlowNet-type CNN* (*Dosovitskiy et al., 2015*) with unconstrained convolutional filters (414,602 free parameters) and which compares only two subsequent frames outperforms the *Full DMN* (735 free parameters) on the task (Fig. 5.2). DMN #5 without synapse signs, counts, and single-neuron connectivity achieves the second best task performance, closely followed by DMN #2 whose connectivity is still constrained by the signs of connections. DMNs #3 without sign constraints compute motion as well as the *Full DMN*. DMNs with known connectivity but unknown synapse counts compute motion worse (#1 and #4). Our results also confirm that the DMNs with random parameters and the decoder (7427 free parameters) alone cannot compute motion.

We do not know how the real fly ranks on the task error axis among these comparisons, but the gap between the FlowNet-type CNN and the Full DMN suggests that the structure of the reconstructed partial fly connectome is worse at computing the input and output relationship of the Sintel data — however, alternative causes such as choices in model design and optimization are not ruled out by this comparison. Additionally, the better task-performance of DMN #5 and #2, which train the average convolutional filters instead of using the fly visual system reconstruction, suggests that better filters than the reconstructed exist.

However, the *Full DMN* lies on the Pareto front of the number of unknown (and therefore task-trained parameters) and the task error among these

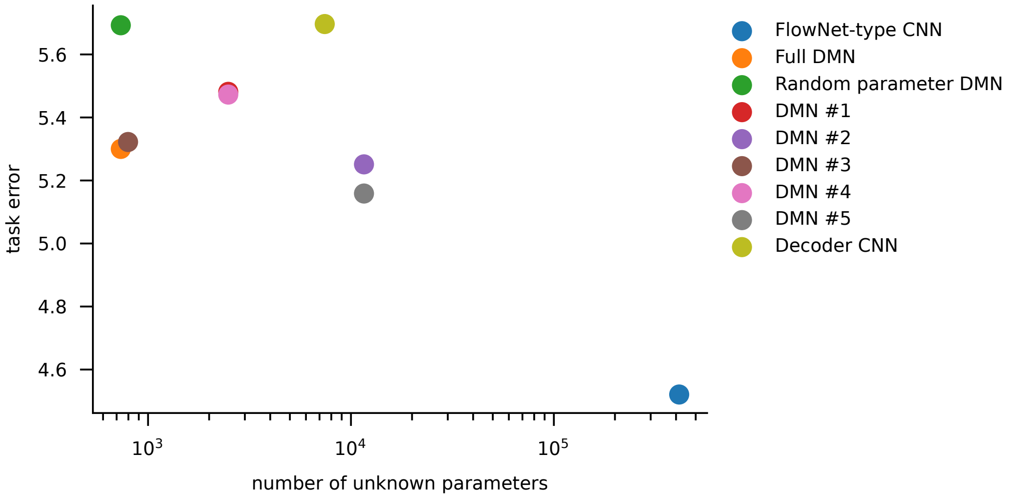


FIGURE 5.3: **Task error against number of free parameters.** Fewer unknown parameters and lower task error are generally better. In this comparison, the *Full DMN* lies on the Pareto front of these two dimensions. Note that parameter types differ across the model types in this high-level comparison.

comparisons (Fig. 5.3). Of course, the real fly brain is additionally exposed to metabolic budget constraints that compete with computational goals. Its nervous system computes on a power budget of approximately 120 nanowatts *(Scheffer, 2021)* — several times less than the power budget of a quartz watch.

While these results provide insights into optimal neural network design, they do not conclusively answer this question. Optimization algorithms play a role in this comparison and RNNs are known to be harder to optimize than CNNs (Section 2.2.5). Moreover, the free parameter types differ across these model types, making it speculative to establish a comparison and Pareto front. Similar computational experiments with different model types and task datasets (including motion detection tasks) can be done to consolidate or refute these results. Ideally, one can close the loop with experimental measurements that allow us to place the real fly and other organisms for comparison on the axis of task error on the Sintel task. While task performance of neural networks is frequently compared to human task performance in deep learning *(Jacobs & Bates, 2019)*, experiments for evaluating motion detection performance in the fly across a deep learning dataset have not yet been developed.

5.2 BRIDGING MODEL AND EXPERIMENTAL VARIABILITY

Different experimental labs use different LED displays for stimulus presentation, stimulus types, neural activity recording techniques, and postprocessing methods for investigating the nervous system function of the fly. Validating our model with respect to the variability in existing experimental measurements has been challenging.

For instance, calcium measurements can lead to nonlinear amplifications of certain neural tuning that cannot be observed in equivalent voltage recordings *(Wienecke et al., 2018; Mishra et al., 2023)*. This is particularly relevant when interpreting underlying mechanisms of observed tuning such as preferred direction (PD) enhancement and null direction (ND) suppression in motion detection. Calcium indicators amplify depolarizations nonlinearly but have little effect on hyperpolarizing responses. As a result, PD enhancement-driven by excitation-can appear overemphasized, while ND suppression-driven by inhibition-can be underestimated in calcium measurements. Our model revealed ND suppression but no PD enhancement (Section 3.2.5), which eventually aligned with voltage recordings *(Gruntman et al., 2018; 2019)*, even if early calcium imaging studies had suggested otherwise.

Evaluating how well our model matched the fly often required critically reflecting on our own modeling and evaluation assumptions and taking the detailed experimental methods into account. We found we can more often than not use our own stimuli for accurate mechanistic interpretation, aligning mechanistic interpretations between our model and different experiments.

For instance, motion sensitivity can be characterized with moving gratings *(Groschner et al., 2022; Maisak et al., 2013; Strother et al., 2017; Gruntman et al., 2019; Ramos-Traslosheros & Silies, 2021b)*, moving bars *(Gruntman et al., 2018; Leonhardt et al., 2016; Fisher, Silies, & Clandinin, 2015; Gruntman et al., 2019; 2021)*, or moving edges *(Maisak et al., 2013; Fisher, Silies, & Clandinin, 2015; Strother et al., 2017; Gruntman et al., 2021)* but we found moving edges are a good common denominator for establishing coarse model correspondences across the literature, even though such alignment may break down in finer detail. As another example, spatiotemporal receptive fields are typically characterized using white-noise stimuli *(Arenz et al., 2017; Currier & Clandinin, 2025)* and reverse correlation, whereas our method uses the ideal single-ommatidium excitability of our model.

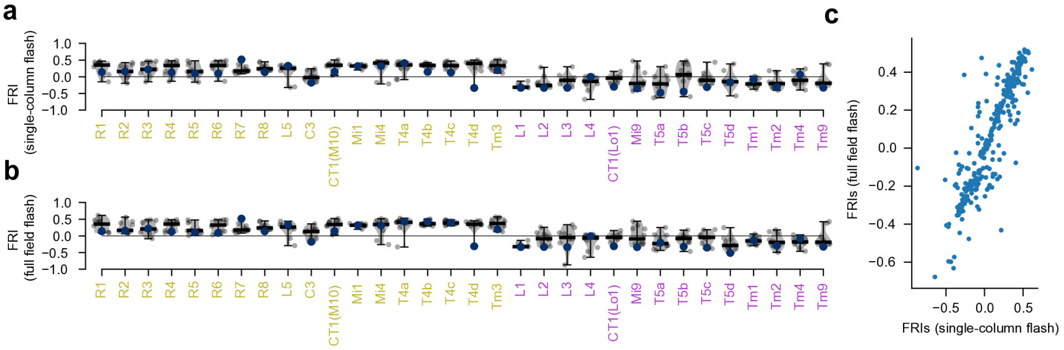


FIGURE 5.4: **Precise mechanistic interpretation is sensitive to stimulus type and metric. (a)** Flash response indices to single-column flashes of neuron types in the model for which such selectivity was known at the time of modeling (new measurements now exist *(Currier & Clandinin, 2025)*). **(b)** Equivalent flash response indices to full-field flashes. **(c)** Comparison scatter plot of FRIs from (a) and (b). Neuron types known to depolarize to central light-increment flashes in yellow, neuron types known to hyperpolarize in magenta.

Similarly, different future computational projects to model the same connectome-constrained system may evaluate models with different stimulus types or postprocessing, too.

For instance, our flash response index is not a comprehensive measure to characterize neuron type function as its outcome depends on flash parameters. A neuron with a center-surround receptive field would be classified as ON-selective with a small flash that only hits the excitatory center but classified as OFF-selective with a larger flash that also hits the inhibitory surround. Therefore, the flash response index is in fact a function of the flash size, shape, and location (Fig. 5.4).

Similarly, orientation selectivity measurements illustrate how stimulus choice affects model validation. T4 and T5 neurons in the fly visual system detect motion and orientation of edges orthogonal to the preferred direction of motion *(Fisher, Silies, & Clandinin, 2015)*. When we evaluated the orientation selectivity predicted by our model (Fig. 5.5), we found both agreement and disagreement with experimental results depending on the stimulus type used. The model accurately predicts high orientation selectivity in T4 and T5 neurons, and orientation selectivity axes in response to oriented edges correspond to experimental measurements for T4 and T5 (Fig. 5.5b). However, with oriented bars, the model accurately predicts orientation selectivity axes for T4 but not for T5 (Fig. 5.5c). Additionally, the model incorrectly predicts no orientation selectivity for TmY9 *(Seung,*

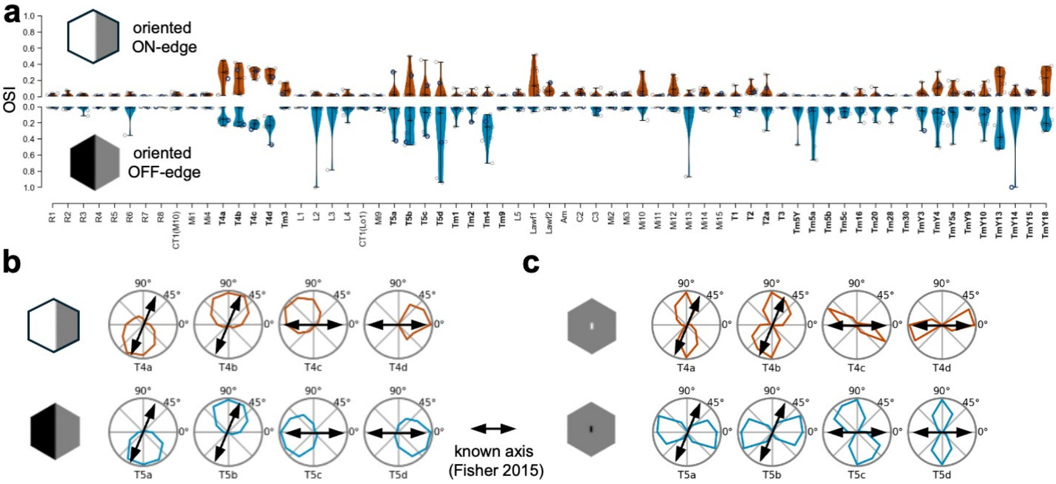


FIGURE 5.5: **Orientation selectivity predicted by the model.** **(a)** Orientation selectivity index (OSI) *(Mazurek et al., 2014)* across neuron types from 10 best task-performing models from the ensemble, computed from responses to oriented edges. **(b)** Orientation tuning curves to edges of T4 (red) and T5 (blue) subtypes. **(c)** Orientation tuning curves to bars of T4 (red) and T5 (blue) subtypes.

2024), correctly predicts orientation selectivity for T2 and TmY5a, but falsely predicts none for Mi1 *(Currier & Clandinin, 2025)*. To better align model analysis with experiments, one can use full-field gratings *(Fisher, Silies, & Clandinin, 2015)* as commonly used in experimental studies.

Our model also enables validation through in-silico silencing experiments that can be qualitatively compared to experimental perturbations. We simulated the silencing experiments reported by *Strother et al. (2017)* by clamping inputs to T4c neurons to zero and comparing the effects to their experimental results using shibire-mediated silencing. Despite methodological differences — our voltage measurements versus their calcium imaging, and our complete input clamping versus their synaptic transmission blocking — the model correctly captures experimental finding qualitatively: Removal of Mi1 excitation eliminates the T4c response, removal of Tm3 excitation decreases but does not eliminate the T4c response, Mi4 removal does not eliminate the T4c response, and Mi9 silencing has no effect on T4c (Section 7.1). The overall agreement demonstrates how connectome-constrained models can predict causal experimental outcomes and be validated beyond correlative measures.

Work targeting the standardization of experiments and computational methods accelerates closing the modeling and experiment loop by reducing epistemic uncertainties. In turn, existing models can help guide the choice of experimental methods.

5.3 UNDERSTANDING MODEL VARIABILITY THROUGH ENSEMBLE ANALYSIS

As seen previously, connectome-constrained models trained on the same task can converge to different solutions with similar performance. Therefore, we train an ensemble of equivalent models which spans a range of task-compatible neural mechanisms. Understanding consensus from the variability in the ensemble is crucial for generating reliable mechanistic hypotheses from the model and identifying which predictions are robust versus which reflect degeneracies in the solution space.

5.3.1 *Distilling predictions from ensembles: Finding consensus in model diversity*

We found specific symmetries in our DMN such as T4 and T5 subtypes that can have the other subtype's preferred direction selectivity after model training. We demonstrated how constraints from measurements of neural activity of one neuron type ruled out symmetric solutions, also across several other neuron types (Section 3.2.4). While such promising strategies for hypothesis refinement exist, dealing with symmetries in model identification remains inevitable because of the nature of the inverse problem (Section 2.2.7).

To find consensus across models, we used nonlinear dimensionality reduction and clustering. Specifically, we used uniform manifold approximation and projection (UMAP) *(McInnes et al., 2018)*, but several other applicable methods exist *(Cunningham & Yu, 2014; Maaten & Hinton, 2008; Hinton & Salakhutdinov, 2006; Schölkopf et al., 1998)*, including t-distributed stochastic neighbor embedding (t-SNE) *(Maaten & Hinton, 2008)*, kernel principal component analysis (kPCA) *(Schölkopf et al., 1998)*, and autoencoder-based embeddings *(Hinton & Salakhutdinov, 2006)*. These dimensionality reduction methods trade off local and global structure preservation, computational speed, and parameter optimization complexity.

However, UMAP itself involves stochastic optimization, leading to variability in embeddings across different random seeds. Subsequent unsupervised clustering of models with Gaussian mixture models based on the UMAP embeddings can therefore vary in their consistency. Note that, because of an uncharacterized implementation issue in the current UMAP package, even results with set random seeds can differ across reruns — which admittedly caused confusion and delays during our model evaluation[1]. Therefore, we carefully evaluate clusters found by nonlinear dimensionality reduction and unsupervised clustering by analyzing their sensitivity to hyperparameters and investigating the data features that separate them.

Eventually, we found that average neuron-type activity predictions from the best average task-performing model clusters accurately capture experimental measurements (Background Section 2.2.7 and Manuscript Section 3.2.5).

As another strategy for distilling predictions from several models, we filtered models based on known neuron selectivity of motion detector neuron types and their inputs (Section 3.5.6).

While this results in several viable models per motion detector neuron type, no single model instance from the main paper ensemble captures all input properties and all motion detector neuron type selectivities accurately simultaneously. Since hypothesis generation is faster with a single model, being able to efficiently distill clustered solutions into a single model instance, for instance using a supervised teacher-student framework where the teacher signal corresponds to averages from the best task-performing clusters, would represent a future method improvement for making mechanistic predictions from the model.

More broadly, ongoing research is concerned with mechanistic analyses of symmetric solutions within a teacher-student framework, uncovering a variety of symmetry types in artificial neural networks at single-neuron resolution *(Şimşek et al., 2021; Martinelli, Simsek, Brea, & Gerstner, 2023)*. For instance, groups of neurons are found to approximate the same output as a single neuron and therefore function as its proxy. Such proxy relationships represent symmetric solutions where different neuron types perform equivalent functions, maintaining the same task performance despite different neural computation.

In a similar spirit, we also asked which neuron types can act as functional proxies for other neuron types across the ensemble.

---

[1] https://github.com/lmcinnes/umap/issues/1080

5.3.2 *Neural function degeneracies and symmetries across the ensemble*

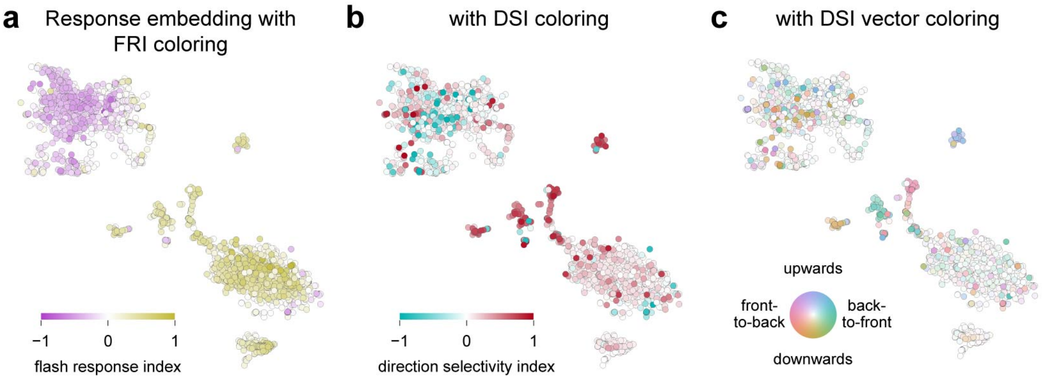


FIGURE 5.6: **Characterizing the functional space predicted by the whole ensemble.** Simulated voltage responses to naturalistic stimuli nonlinearly reduced to two dimensions. Response embedding of voltage from one representative neuron per neuron type and from each of 50 models, resulting in 3,250 scatter points. **(a)** Embedding points colored by flash response index (FRI) (Section 3.5.6). Yellow means ON-selective, magenta means OFF-selective. **(b)** Colored by direction selectivity index (DSI) (Section 3.5.6). Red means ON-direction selective, petrol means OFF-direction selective. **(c)** Colored by "DSI vector": That is the vector pointing towards the predicted preferred motion direction for the preferred motion-contrast of length of the normalized DSI.

In future, one can build a connectome-constrained model of billions of neurons and optimize it on a high-level objective, like a task or behavioral data. Like the real brain, the in-silico model remains complex and must be evaluated step-by-step. Our model has already 45k neurons and 1.5M neuron-to-neuron connections. While we can rely on neuron types, is there a more general strategy on how to start analyzing such a model, potentially providing overview insights even in the absence of neuron types?

What is the functional space spanned by neurons in the trained ensemble of the connectome-constrained model? How degenerate are individual neurons across the ensemble? Which neurons can act as proxies for other neurons because of a lack of constraints?

To address these questions, we expand our nonlinear dimensionality reduction and unsupervised clustering analysis to be agnostic to the neuron type: Treating responses from a representative neuron of each neuron type as independent allows us to embed and cluster the entire functional space spanned by the model ensemble (Fig. 5.6).

One can expect around ten different functional clusters: ON-selective neurons, OFF-selective neurons, ON-selective motion-tuned neurons to either of four cardinal directions, and OFF-selective motion-tuned neurons to either of four cardinal directions. As expected, we find two high-level clusters: ON-selective and OFF-selective neurons (Fig. 5.6a). Moreover, unambiguously, we find four ON-selective, motion-tuned clusters to each cardinal direction (Fig. 5.6 a-c). OFF-motion selective clusters are less clearly pronounced in the embedding, likely because OFF-motion selectivity is generally less pronounced in this ensemble, consistent with the main paper results.

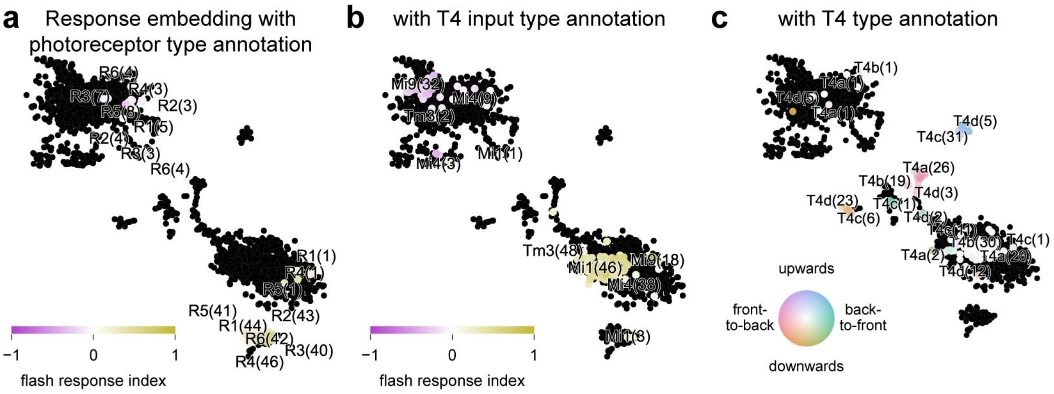


FIGURE 5.7: **Identifying function degeneracy of ON-pathway neuron type in model.** As in Fig. 5.6: Simulated voltage responses to naturalistic stimuli nonlinearly reduced to two dimensions. Response embedding of voltage from one representative neuron per neuron type and from each of 50 models, resulting in 3,250 scatter points. Clusters labeled by occurrence of neuron type from ensemble, counts in parenthesis adding up to 50 models across entire embedding. Only scatterpoints from investigated neuron types are colored, remaining scatter points are black. **(a)** Embedded and clustered responses of photoreceptor neuron types R1-R6 are highlighted. Points colored by flash response index (FRI) (Section 3.5.6). Yellow means ON-selective, magenta means OFF-selective. **(b)** Equivalently, embedded and clustered responses of major T4 input types: Mi1, Mi4, Tm3, and Mi9. **(c)** Embeeded and clustered responses of T4 neuron types, colored by "DSI vector". The vector points towards the predicted preferred motion direction for the preferred motion-contrast of length of the normalized DSI.

While the embedding and clustering is agnostic to the neuron type, in fact we know the underlying neuron types and can label each data point by it. This gives an impression of how degenerate predictions for function of specific neuron types are and clarifies the identity of less expected clusters. For instance, a large functional cluster belongs to the photoreceptor neurons (Fig. 5.7a), likely because they are the only neurons in the circuit that receive external input. While the vast majority of photoreceptor neurons concentrates in that ON-selective cluster, some models predict photoreceptor function as OFF-selective neuron type — even though with near-zero flash response index.

Expanding our analysis to the major T4 input neuron types, we observe a pattern consistent with earlier observations (Section 3.2.4): Some models predict the opposite contrast selectivity for Mi4 and Mi9 neurons (Fig. 5.7b), while Tm3 and Mi1 are accurately predicted with consistency. Eventually, we also observe the clear functional separation into neurons tuning towards all four cardinal direction belonsg to T4 subtypes and we observe the known model degeneracy between pairs of T4 subtypes (Fig. 5.7c) for detecting either upwards or downwards motion or front-to-back or back-to-front motion.

Of course, this analysis is not limited to these neuron types: We find for instance accurate clustering of L1-L5 neuron types, while models are most degenerate in L3 neurons (Fig. 7.1a). We find that most of the major T5 input types accurately fall into the OFF-selective cluster (Fig. 7.1b). Additionally, we find that T5 neuron types mostly distribute in the OFF-selective cluster without clear clustering here, and have symmetries with T4 subtypes (Fig. 7.1c).

Characterizing the function space spanned by the ensemble through nonlinear dimensionality reduction and clustering in conjunction with annotation of neuron labels and computed tuning indexes, represents a general strategy to gain an overview over the function of all neurons in a connectome-constrained model. While the analysis here is exemplary and uses neuron type knowledge, it does not rely on it and can be applied to any neural activity tensor — whether modeled or measured in the living system. However, we do not discover a completely novel functional cluster, indicating that either: (1) the structure to support such a functional cluster exists inside our connectome reconstruction, but we cannot predict the cell function from the task training, or, (2) structures to support additional functional clusters do not exist inside our connectome reconstruction.

5.3.3 *Full characterization of ON-motion-detection degeneracy*

While our model predicts direction-selectivity properties accurately across T4 and T5 subtypes, it suggests direction-selectivity with statistical signifi-

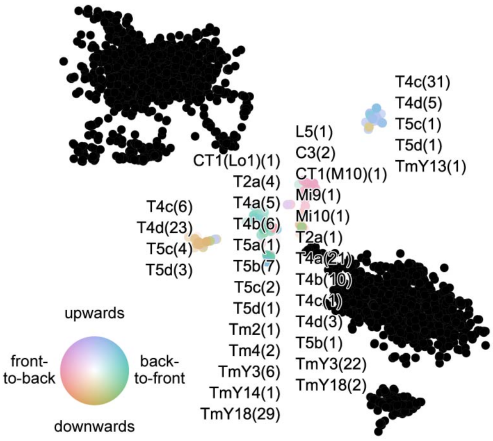


FIGURE 5.8: **Full characterization of ON-motion detection degeneracy in the model across ensemble.** As in Fig. 5.6 and 5.7: Simulated voltage responses to naturalistic stimuli nonlinearly reduced to two dimensions. Response embedding of voltage from one representative neuron per neuron type and from each of 50 models, resulting in 3,250 scatter points. Clusters labeled by occurrence of neuron types, annotated in columnar fashion. Only scatterpoints from investigated neuron types are colored, remaining scatter points are black. From left to right: Downwards ON-motion selective cluster (yellow) contains T4c, T4d, T5c, T5d; Back-to-front ON-motion selective cluster (green) contains CT1(Lo1), T2a, T4a, T4b, T5a, T5b, T5c, T5d, Tm2, Tm4, TmY3, TmY14, and TmY18; Front-to-back ON-motion selective cluster (red) contains L5, C3, CT1(M10), Mi9, Mi10, T2a, T4a, T4b, T4c, T4d, T5b, TmY3, and TmY18; Upwards ON-motion selective cluster (blue) contains T4c, T4d, T5c, T5d, and TmY13.

cance across the ensemble also for neuron types TmY3, TmY4, TmY5a, and TmY18 (Section 3.2.2 and 3.5.6). We reported a full analysis on TmY3 (Section 3.2.2), because the other neuron types receive direct or second-order input from motion selective T4 and T5 neurons in our model.

While the incompleteness of the connectome reconstruction we worked with for this model version (Section 2.3.2) limits the mechanistic inferences that we can make confidently, we developed strategies for generating further mechanistic insights.

As an example, one may wonder what other neuron types function as proxies for T4 neurons. We discovered that T4 neurons clustered in function space into four distinct clusters with high direction selectivity (Fig. 5.7). By definition of the clusters, other neuron types in these clusters have similar function and can therefore act as direct proxies of the corresponding T4 subtype in the models. Analysis reveals moderate degeneracy in upwards and downwards ON-motion selective clusters with occurrence of T4c, T4d, T5c, and T5d neuron types and one occurrence of a TmY13 neuron type in the upwards selective cluster (Fig. 5.8). Front-to-back and back-to-front motion selective clusters are more degenerate, with prevalent occurrence of T4a and T4b types, TmY3, TmY18, and irregularly T5 subtypes. Occasionally, known non-motion selective neuron types, including inputs to T4 and T5 such as CT1, Mi9, Tm2, and Tm4 fall into these clusters. However, such stray occurrences are not significantly observed across the ensemble such that these hypotheses are ruled out already by neuron-type specific clustering and task error sorting of clusters (Section 5.3.1).

This analysis provides a comprehensive high-level overview of neuron function in the model, enabling discovery of new neuron type functions while taking existing functions into account. This approach may be particularly useful when large models hinder manual analysis of every neural response time course, for instance with future connectome-constrained models based on the full reconstructions of the optic lobe *(Nern et al., 2025; Matsliah et al., 2024)*.

5.4 FORMALIZING INFERENCE IN DEEP MECHANISTIC NETWORKS

Because many structures can have the same function and the same structure can have many functions, we face an inverse problem when trying to infer neural function with the connectome and a high-level task. We realized that we need to infer distributions over possible neural functions that are consistent with all measurements to make mechanistic predictions.

We found deep ensembles of connectome-constrained neural networks can generate hypothesis for neural function effectively and allow refinement. Using nonlinear dimensionality reduction and clustering, we found several proposed solutions per neuron type and analyzed them jointly for certain neuron types to make informed guesses of how neural structure and function are related within the constraints our model has (Section 3.2.4). Across neuron types, predicted responses clustered strongly, boiling down the inverse problem to only a few possible solutions. We found that model clusters with the best average task performance resulted in accurate predictions for many neuron types (Section 3.2.5).

5.4.1 *Towards a probabilistic framework for function inference with connectome-constrained models*

Building on the probabilistic deep learning framework (Section 2.2.7, using the same notation here) and our insights using the DMN (Section 3.2.4 and 5.3.1), we specify the inference approach in the following.

Treating the stochastic optimization as the source of posterior variation, for our DMN ensemble with M models, each converged solution θ^*_m represents a sample from the approximate posterior $p(\theta|\mathcal{D}, W)$, where $\mathcal{D} = \{(\mathbf{x}_s, \mathbf{y}_s)\}_{s=1}^{S}$ denotes our training data and W specifically denotes the connectome constraints. The ensemble provides Monte Carlo estimates of neural functions:

$$\mathbb{E}_{\theta \sim p(\theta|\mathcal{D},W)}[V_i(\mathbf{x})] \approx \frac{1}{M} \sum_{m=1}^{M} V_i^{(m)}(\mathbf{x}, \theta^*_m)$$

However, when the posterior is multimodal or when different parameter configurations produce multimodal functional outputs, averaging across all models can obscure differences. We found that functional clustering in model activity space allows us to detect distinct hypotheses. We employed a two-stage analysis:

1. *Functional clustering.* We cluster models based on neural response similarity (from responses to naturalistic stimuli $\mathbf{x}_{\text{nat.}}$) using dimensionality reduction (UMAP) based on pairwise correlation, followed by Gaussian mixture modeling:

$$\begin{aligned} u_i^{(m)} &= \text{UMAP}_{2D}(V_i^{(m)}(\mathbf{x}_{\text{nat.}})) \\ p(u_i) &= \sum_{k=1}^{K} \pi_k \mathcal{N}(u_i|\mu_k, \Sigma_k) \end{aligned}$$

The main drawback of this specific method is that nonlinear dimensionality reduction can be sensitive to hyperparameter settings, which is data dependent (Section 5.3.1).

Choosing the number of mixtures K that minimizes the negative log-likelihood results in K distinct functional clusters $\{\mathcal{M}_1, \mathcal{M}_2, ..., \mathcal{M}_K\}$, each representing a different computational hypothesis.

2. *Task-error ranking.* We rank clusters by their average task performance, usually selecting the best-performing cluster(s):

$$\mathcal{M}^* = \arg\min_k \mathbb{E}_{m \in \mathcal{M}_k}[L_{\text{task}}(\theta_m^*)]$$

where $L_{\text{task}}(\theta_m^*) = -\log p(\mathbf{y}|\mathbf{x}, \theta_m^*)$ can be interpreted as the negative log-likelihood under the model, connecting our approach to maximum likelihood estimation.

3. *Mode-conditional inference.* For the selected cluster $\mathcal{M}^*$, we compute mode-specific neural function estimates. Since models produce neural responses at different scales, we first normalize each model's responses by their mean absolute activity (Section 3.5.6):

$$\mu_i^{(k)}(\mathbf{x}) = \frac{1}{|\mathcal{M}^*|} \sum_{m \in \mathcal{M}^*} \frac{V_i^{(m)}(\mathbf{x}, \theta_m^*)}{\mathbb{E}_{\mathbf{x}' \sim p(\mathbf{x})}[|V_i^{(m)}(\mathbf{x}', \theta_m^*)|]}$$

This approach allows us to identify and characterize distinct mechanistic hypotheses while avoiding the averaging artifacts that can occur when combining fundamentally different solutions.

5.4.2 *Towards applications for experimental design*

Our ensemble clustering framework provides a principled approach for designing experiments that efficiently resolve mechanistic uncertainties. By analyzing disagreements across clusters in our model ensemble, we can identify specific computational mechanisms where models diverge, translating these divergences into testable experimental predictions.

For example, as an illustrative case using already known properties, if our clustering reveals that models disagree about whether Mi4 neurons depolarize or hyperpolarize to light-increment stimuli, experimentalists can prioritize measuring Mi4 responses to light-increment and light-decrement flashes for distinguishing between these hypotheses. Such targeted measurements can efficiently resolve which computational hypothesis is correct. For instance, if Mi4 is indeed depolarizing to light-increment stimuli, as experiments suggest, this supports models where T4c detects upward motion rather than downward motion.

More generally, by identifying which aspects of neural activity vary most across clusters, DMNs can guide experimentalists toward the measurements that will be most informative for resolving mechanistic uncertainties.

For instance, ensemble disagreements about depolarization or hyperpolarization suggests using flashes of different contrasts, while disagreements about temporal dynamics suggest testing with stimuli at specific frequencies.

To close the loop between modeling and experiments, models need to integrate new experimental measurements $\mathcal{D}_{\text{new}}$. While exact Bayesian updating is intractable, we can approximate this process. Given a posterior approximation $p(\theta|\mathcal{D}, W)$ from our ensemble $\mathcal{M}^*$ and new neural activity measurements, we can update the model:

$$p(\theta|\mathcal{D}, \mathcal{D}_{\text{new}}, W) \propto p(\theta|\mathcal{D}, W) \cdot p(\mathcal{D}_{\text{new}}|\theta, W)$$

In practice, on a high-level, we perform such updates through two complementary approaches:

1. *Training the ensemble* with an augmented loss function that includes both task performance and neural activity matching:

$$L_{\text{combined}}(\theta) = L_{\text{task}}(\theta) + \lambda L_{\text{neural}}(\theta, \mathcal{D}_{\text{new}})$$

Training is principled as it can discover new solutions consistent with the data, but to integrate new data, we usually need to retrain, requiring significant computational resources and potentially careful tuning of loss balancing (Section 4.2.3).

2. *Filtering the existing ensemble* to retain only models consistent with the new measurements (Section 3.5.6):

$$\mathcal{M}_{\text{filtered}} = \{m \in \mathcal{M}^* : L_{\text{neural}}(\theta^*_m, \mathcal{D}_{\text{new}}) < \epsilon\}$$

Filtering a task-trained ensemble of $M = 50$ models is computationally efficient but can reduce ensemble size significantly if new measurements are highly constraining, with no guarantee that a compatible model exists. Thresholds ϵ can be determined through percentiles of computationally derived null-distributions (see for instance, Section 3.5.6).

We recommend filtering for initial exploration and validation, reserving retraining for cases where filtering eliminates too many models or when fundamentally new neural responses need to be included.

Future work can explore active learning approaches to automatically propose the most informative experiments based on ensemble uncertainty. Active learning approaches rely on calibrated uncertainties, because the model uncertainty must correlate to its inaccuracy to request the most informative new experiments. In follow-up work with P. von Bachmann *(von Bachmann, 2023)*, we analyzed how calibrated deep ensembles of deep mechanistic synthetic connectome networks are and found they were well-calibrated, aligning with other empirical findings *(Lakshminarayanan et al., 2017)*. Moreover, mixtures of laplace approximations *(Eschenhagen et al., 2021)* can effectively approximate posterior modes from a smaller number of ensemble members (such as $M = 5$).

The iterative refinement process presented here can guide experimental design and enables model ensembles to progressively converge toward the true biological function.

FINAL DISCUSSION

6

6.1 CORE FINDINGS AND CONTRIBUTIONS

In this work, we asked: Can we build dynamic computational models of the brain from its connectome, integrating single-neuron mechanisms and task-computation with machine learning? How can we construct such models, and what tools enable us to evaluate them?

To that end, we specifically constructed a computational model of the fly visual system from a nanoscale connectome reconstruction. We built the model based on 64 identified neuron types and the nearly convolutional nature of the fly visual system (Section 2.3.2 and 3.2.1). We assumed passive-point neurons, because the small neurons in the fly that we model are non-spiking. We equipped each of 45,669 model neurons in the connectome-constrained network graph with continuous-time dynamics (Section 2.1.1 and 2.4). We constrained the 1,513,231 neuron-to-neuron connections with corresponding synapse counts from the connectomic measurement.

We implemented the network in PYTORCH *(Paszke et al., 2019)* to obtain gradients for training the neural parameters of the system of differential equations with backpropagation through time (Section 2.2.5 and 3.2.1). We represent the network as a sparse graph, and the convolutions using scatter and gather operations for efficient task-training of the model with sparse connectome-constrained connectivity (Section 3.2.1 and 3.5.4). We trained it on the Sintel optic flow task *(Butler et al., 2012)*, because motion detection underlies the fly's ability to navigate (Section 2.3.3 and 3.5.5).

While developing the model, training paradigm, methods and carrying out training and analysis, we gradually built a comprehensive compute software and infrastructure for careful experimentation (Section 4.4). Systematic experimentation with this framework enabled consolidation of artifacts of model design, such as the neuron dynamic equations (Section 2.1 and 3.5.3), the precise parametrization of the synaptic strengths (Section 2.4 and 3.5.3), the custom stimuli and tuning indices for evaluating the model (Section 3.5.6 and 3.5.7), ways of training (Section 4.3.1 and 3.5.5), and of paralleling model computation to knowledge from the literature (Section 2.3.3

and 3.2.4). We curated and open-sourced our code[1] and provide a full characterization of the model ensemble for each cell type across 478 pages of supplementary information.

The trained model ensemble predicts single-neuron tuning properties across the fly visual system, accurately capturing functional properties across experimental literature that comprises several decades of experimental work at the level of neuron types (Section 3.2.2). The model predicts the segregation into ON and OFF-pathways for motion detection (Section 3.2.2), T4 and T5 motion tuning to all cardinal directions and hypotheses for detailed underlying synaptic mechanisms (Section 3.2.5). Moreover, we derived hypotheses from it for selectivity in 32 uncharacterized neurons, most notably motion-selectivity in TmY3 (Section 3.2.5 and Fig. 3.ED2) — a neuron type that was previously not considered to be motion selective .

This work presents a novel and general strategy for building a computational brain model from a connectome using machine learning. Careful model design with connectome constraints and suitable neuron dynamics combined with a high-level computational task allow our model to generate mechanistic insights about neural computation in the fly — even without explicitly requiring knowledge of the dynamic computation of each neuron. Our results suggest it is generally possible to derive accurate dynamical models of neural circuits from bottom-up and top-down constraints despite the challenging nature of the high-dimensional inverse problem. We demonstrate, with nonlinear dimensionality reduction and clustering, that predicted neural responses cluster across degenerate predictions from the ensemble. Our method allows to generate and refine hypotheses about neural computation across variable and high-dimensional solution spaces. This represents a strategy to address the challenging inverse problem within a framework that can be tied to Bayesian inference for iterative model and experimental design, for a principled model-based scientific discovery loop.

Our work demonstrates how connectomes can be useful for predicting neural circuit function when combined with task constraints and machine learning. Our approach can be broadly applied to end-to-end connectome-constrained modeling and enables new future model-based research of nervous systems function using connectomes.

---

[1] https://github.com/TuragaLab/flyvis

6.2 CONTEXT

Model neurons, in this work, have a direct correspondence to neurons in the fly brain, enabling detailed mechanistic model interpretation and generation of experimentally testable hypotheses about neural computation. Previous studies of task-optimizing artificial neural networks as models for neural systems, only demonstrated high-level similarity between population level representations of brain regions and model layers (Section 2.2.6).

Other task-optimized models with unit correspondence to neurons in the fly visual system exist. They provided intriguing early evidence that one could predict function from structure with task-training of connectome-constrained neural networks. These were either based on a small circuit, linear-nonlinear receptive field model *(Mano et al., 2021)*, or a discrete RNN constrained by partial reconstruction of the motion pathways *(Tschopp et al., 2018)*. Instead, our neuron model derives from a detailed biophysical model of continuous-time neural dynamics (Section 2.1.2) and simulates voltages across a comprehensive connectome reconstruction of the motion pathway (Section 2.3.2), allowing establishing detailed comparisons of circuit dynamics to experimental measurements (Section 3.2.5).

We found that task-training resulted in diverse neural tuning predictions across an ensemble of equivalently trained models. Previous work made a similar observation in a single-neuron model constrained by a neuron's morphology and its high-level task *(Zhou et al., 2022)*. Here we find degeneracies in most neuron types across the system and developed a general strategy to characterize such degeneracies and distill mechanistic hypotheses from the consensus across model ensembles (Section 2.2.7 and 3.2.4).

We demonstrated that degenerate solutions could be refined effectively by filtering the ensemble with knowledge of neural activity of one neuron, resulting in accurate tuning predictions from the connectome-constrained model (Section 3.2.4). A more principled approach, though more computationally intensive, integrates knowledge of neural activity directly through simultaneous gradient-based optimization on the task objective and a neural activity objective (Section 4.2.3 and 5.4). While previous work demonstrates integration of neural activity measurements into connectome-constrained models of the full connectome of *C.elegans* with machine learning approaches *(Warrington et al., 2019; Mi et al., 2022)*, and investigates requirements for identifying dynamical system computation at single-neuron level with connectome-constrained models and neural activity measurements *(Beiran & Litwin-Kumar, 2024)*, simultaneous task-optimization and training on neural activity of a connectome-constrained neural network — integrating both single-neuron dynamic constraints and system-level computation — represents a novelty.

Complementarily, measurements of behavior *(Branson et al., 2009)* can be used to build connectome-constrained models, which can lead to insights about single-neuron computation even in the absence of detailed connectomic constraints *(Cowley et al., 2024)*. Such models rely on artificial neural network architectures for encoding and decoding to bridge gaps in the connectome reconstruction, particularly in modeling sensory circuits and motor outputs.

The recent arrival of whole-brain reconstructions of the fly (Section 2.3) allows unprecedented scale of connectome-constrained modeling for model-driven discovery of circuit and behavioral computation in the fly. As an early example, *Shiu et al. (2024)* built a point-neural network with a leaky-integrate and fire neuron model with ten parameters. Nine parameters were selected from the literature and the tenth parameter, synaptic strength, was manually tuned to match the measurements of spiking statistics of a target neuron to controlled input. While "precise dynamics may be poorly simulated by the computational model" *(Shiu et al., 2024)*, the model predicted neurons activated by sugar, neurons that are part of the antennal grooming circuit, and neurons that drive proboscis extension. While this work did not analyze degenerate solutions across model ensembles as we do (Section 2.2.7), their successful identification of behaviorally relevant neural activity suggests they may have targeted relatively less degenerate subcircuits or that the degeneracy is an artifact of our modeling choices, whereas the latter explanation seems unlikely given the theoretically established non-identifiability and symmetry properties of neural networks (Section 6.3.4).

6.3 STRENGTHS, LIMITATIONS, AND WHAT'S NEXT?

6.3.1 *Realism of anatomical constraints*

We focused on simulating the fly visual system, specifically the motion pathways, based on the only available reconstruction at the time (Section 2.3.1 and 2.3.2). The fly visual system offered extensive physiological characterization, high precision reconstruction, and defined tasks with overlap to existing deep learning for computer vision datasets.

For model input, we rendered video sequences with a flat hexagonal box filter. This is arguably a crude but fast abstraction of phototransduction by retina lenses in the central portion of the compound eye. For more anatomical realism with nuanced functional consequences, future models may require modeling the full spherical extend of the eye *(Nern et al., 2025; Zhao et al., 2022)* with ray tracing *(Millward et al., 2022)* and optical lens properties that are aligned with the ecological demands of the species *(Gonzalez-Bellido et al., 2011; Fenk et al., 2022)*.

We tiled the optic lobe into retinotopic, hexagonally arranged columns and used average synapse counts from reconstructions within these columns to define a local convolutional filter for each connected pre- and postsynaptic neuron type. To build a comprehensive optic lobe model from such local reconstructions, we assumed perfect convolutional symmetry and compiled a graph for describing full optic lobe connectivity from average convolutional filters across hexagonal columns corresponding to retinal columns of the central portion of the compound eye. The optic lobe model represents the almost-convolutional consensus circuit of the motion detection pathways in the fly (Section 2.3.2).

According to new data, the neuron types in our model account for 29K neurons[2] of 43K optic lobe neurons with inputs in the optic lobe, corresponding to 67.4% of all optic lobe neurons and 41.7% of neurons (assuming right and left visual system equivalence) of the whole fly brain *(Dorkenwald et al., 2024)*. From the new data, we build both new models with the same neuron types and average convolutional filters for comparison to the previous model, and new models that will integrate all reconstructed neurons and synapses of the visual system, such as new instrinsic neurons types, including for instance interneuron types Pm, Dm, and Cm; new connecting neuron types, including, for instance, Y neuron types connecting medulla, lobula and lobula plate; output neuron types, including for instance LC neuron types projecting to the central brain; and centrifugal neuron types, including for instance LoVC neuron types projecting from the central brain to the optic lobe *(Nern et al., 2025)*.

A model integrating the full eye anatomy and optic lobe reconstruction supports the model-driven investigation of questions such as the functional implications of retinal movement *(Fenk et al., 2022)* or anatomical differences in peripheral regions of the fly visual system *(Nern et al., 2025; Zhao et al., 2022)*, the plausibility of fundamental motion detection along six axes instead

---

[2] https://tinyurl.com/flyvis

of four cardinal directions *(Henning et al., 2022)*, the precise structure-function relationship of polarized vision *(Wernet et al., 2012)*, and generally the existence of hypothetical subcircuits specialized for different functions *(Matsliah et al., 2024)*.

6.3.2 *Dynamical system realism*

We model neurons with passive, continuous-time graded voltages and current-based synaptic transmission (Section 2.1). Our model integrates synaptic connectivity, synapse counts and signs from neurotransmitter and receptor profiling. While this model captures the key dynamic computational mechanisms for motion detection (Secion 2.3.3), it still lacks integration of important elements known to contribute to dynamic signaling between neurons, such as gap junctions, glia cells, and neuromodulators *(Bargmann & Marder, 2013)*.

Efforts for systematic mapping of gap junctions, glia cells, neurotransmitter, and receptor types are on the way *(Scheffer & Meinertzhagen, 2021)*, while systematic methods for mapping neuromodulator signaling are still lacking *(Nern et al., 2025)*. Moreover, synaptic strength changes at the timescale of minutes *(Hige et al., 2015)* to hours *(Baltruschat et al., 2021)*, and changes in synaptic connectivity numbers on the timescale of hours *(Bushey et al., 2011; Sugie et al., 2018)* are incompletely understood *(Scheffer & Meinertzhagen, 2021)*. Moreover, the fly nervous system adopts a mixed code of signaling via graded potential and spikes (Section 2.3.3), but also oscillatory mechanisms play a role in computation *(Grabowska et al., 2020)*.

As demonstrated here and elsewhere, modeling with limited measurements offers chances to generate plausible hypotheses for yet unknown or unmeasurable biological quantities and dynamics. New connectome-constrained models (Section 4.2) can combine graded potential neurons with spiking HH-type neurons as a more realistic dynamical system description for many real neurons in the fly brain. More detailed models can serve for instance to infer unknown signs of connections *(Mi et al., 2022)*, channel dynamics *(Gonçalves et al., 2020)*, or synaptic plasticity mechanisms *(Lappalainen et al., 2019; Confavreux et al., 2023)* with machine learning methods.

While we can constrain models with the limited data that we have and flexibly parametrize the unknowns for inference, this comes with increased computational requirements, in practice requiring to find the right balance between detail and reduction *(Herz et al., 2006)*. The choice of exact parametrization and biophysical details to include remains hypothesis-driven, suggesting that iterative refinement should encompass not only parameter optimization with neural activity constraints, but also the systematic evaluation of different dynamical system components *(Schröder & Macke, 2023)* to identify those essential for neural computation.

6.3.3 *Task constraints*

To find plausible solutions for the unknown neural parameters, we trained the connectome-constrained model on the optic flow computer vision task Sintel (Section 3.5.5). The connectomic reconstruction of the fly visual system focuses on the motion detection pathway (Section 2.3.2), such that the task we chose directly corresponds to the function of the system that we model. Our control experiments, evaluating models with random parameters, training the model on an autoencoding task (Section 4.3.4), and preliminary results to train it on a larger variety of task representations to understand how different task elements constrain the same system, suggest that a motion computation task is required for accurate tuning predictions. Conversely, the high performance of FlowNet-type CNNs (Section 5.3) with temporal context of only two frames, corresponding to 83 ms, demonstrates that good performance on the optic flow task does not require temporal computation from long contexts — which could be an advantage for RNN optimization in face of vanishing and exploding gradients.

Other systems within the fly, including the central brain and its various subcircuits for olfaction, gustation, audition and mechanosensation might have longer context requirements, which could add an additional challenge in terms of network optimization to finding a suitable task definition. For modeling other nervous systems, labeled task datasets containing input–output pairs might be harder to come by. Deep learning offers self-supervised tools, such as contrastive learning *(Hadsell et al., 2006)*, that can train connectome-constrained models using only input data without explicit task labels *(Duan et al., 2025)*. Contrastive objectives act directly on the activity patterns or learned representations, typically encouraging similar representations for similar inputs and dissimilar representations for dissimilar ones *(Chen et al., 2020; He et al., 2019)*, aligning representation correlations *(Zbontar et al., 2021)*, or maximizing information-theoretic quantities *(van den Oord et al., 2018)* or manifold capacity *(Yerxa et al., 2023)*. While models trained with self-supervised objectives can capture neural representations *(Nejad et al., 2025)*, such training usually requires careful input augmentations to produce meaningful pairs, a principled choice of objective function, and thoughtful selection of model units over which to apply the objective.[3]

Straightforward and complementary to task-optimization or self-supervised training of connectome-constrained models is the use of constraints from measurements of neural activity *(Beiran & Litwin-Kumar, 2024)* (Section 4.2.3) and behavior *(Özdil et al., 2024)*. Integrated in whole-body simulations *(Wang-Chen et al., 2024; Vaxenburg et al., 2025)*, future connectome-constrained models integrated with whole-body simulations will have access to unlimited high-level behavioral task data, complemented by sparse ground-truth neural measurements.

6.3.4 *Non-identifiability*

We refine functional predictions despite parameter non-identifiability to generate accurate predictions (Section 4.3.5) by focusing on the functional consensus of an ensemble of equivalent models. Our ensemble approach thus transforms the challenge of non-identifiability into an opportunity to explore the full space of functionally plausible solutions (Section 5.3). Our method for discovering functional clusters and resolving symmetries with sparse knowledge of neural activity is generally applicable. However, as a possible consequence of non-identifiability and potentially model misspecification, an ensemble comes with no guarantee of discovering all functional clusters, and the correct tuning prediction might not lie within the hypotheses generated by the models.

Non-identifiability emerges when two or more parameter sets result in the same loss values. The loss function is symmetric under such parameter sets, therefore they are called symmetries. The ability of neural networks to exploit symmetries, for instance through overparametrization — giving a neural network many more parameters than required to solve a task — facilitates task generalization in certain instances *(Advani & Saxe, 2017; Neyshabur et al., 2018; Martinelli, Simsek, Gerstner, & Brea, 2023)* but hinders identifiability.

While the relationship between non-identifiability, symmetries, and generalization requires further research, recent work suggests that overparametrization combined with symmetry reduction can be used to iden-

---

[3] Preliminary follow-up results on contrastive learning in DMNs with I. Omolayo are not addressed in this thesis.

tify a neural network from its inputs and outputs *(Martinelli, Simsek, Gerstner, & Brea, 2023)*. In general, deep learning benefits from using overparametrization as a soft inductive bias for finding "simple" data-consistent parameter solutions that achieve good task generalization *(Wilson, 2025)*.

Similarly, future studies could apply soft inductive biases to connectome-constrained models, allowing flexible model definition that could potentially compensate for model misspecification while expanding the space of generated hypotheses. The resulting functional and parameter symmetries in connectome-constrained modeling can then be refined through clustering and analysis to generate precise hypotheses about neural function.

6.3.5 *Closing the loop to experimental measurements of neural activity and behavior*

Our model accurately predicts neural activity across measurements from 26 papers encompassing several decades of experimental research (Section 3.2.2). Predicted tuning properties include direction selective responses of T4 and T5 neurons and detailed hypotheses for the underlying synaptic, linear, and nonlinear mechanisms (Section 3.2.5). The T4 and T5 responses in the fly underlie optomotor response *(Bahl et al., 2013)*, walking *(Creamer et al., 2018)*, and flight behavior *(Fry et al., 2009)*, directly providing a bridge for linking model to fly behavior. In addition to demonstrating that the model predicts voltage responses and tuning properties that highly correlate to measurements of neural activity underlying behavior, we demonstrated that the model can accurately predict key observations from silencing *(Strother et al., 2017)* (Section 7.3)[4].

In terms of candidate experiments, the model predicts ON-sensitivity to front-to-back motion for TmY3 (Section 3.2.5 and Fig. 3.ED2) and TmY5a neurons (Section 3.2.5 and 7.1), suggesting potential motion sensitive neuron types besides T4 and T5 neurons. Latest large-scale characterization *(Currier & Clandinin, 2025)* includes characterization of direction selectivity in TmY3 and TmY5a neuron types from spatio-temporal receptive field mappings convolved with moving gratings, which is a linear response model. This empirically derived model does not corroborate our finding of potential TmY3 motion sensitivity but predicts motion sensitivity for TmY5a neurons. This does align with predictions from our connectome-

---

[4] Preliminary follow-up results on targeted silencing of neurons and connections with Z.Stefanidi are not addressed in this thesis.

constrained and task-optimized model ensemble. Future work will show whether these predictions also hold in models using the new connectome data, and whether they can be confirmed through experiments in the real system. Intriguingly, TmY5a neurons are the most connected hubs in the visual system in terms of input perplexity times output perplexity *(Matsliah et al., 2024)* and their precise functional role in the circuit might be more nuanced.

In general, the model can generate neuron specific response predictions to any visual stimulus or perturbation. Conversely, it enables the design of visual stimuli and perturbations that elicit targeted responses in specific model neurons. We demonstrate this by reporting maximally excitatory stimuli for each neuron type, predicting for instance that T4 and T5 neurons would depolarize more strongly by their preferred stimulus if it is preceded by a null-contrast luminance component. However, we have only touched upon the possible *in-silico* analyses and hypothesis generations that a dynamical system model with unit correspondence offers for investigating neural network function. New models will integrate new connectomic data to generate hypotheses for neural activity of less understood pathways, such as the output neurons of the visual system *(Nern et al., 2025)* and integrate connectome-constrained models with whole-body simulations for direct simulation of behavior. In new analyses we can then adopt targeted silencing strategies to causally investigate how neural activity may underlie behavior.

Besides, testing the predictive capabilities of connectomes represents a new incentive for experimentalists to record neural activity data at a large-scale, across many neurons and of the whole system simultaneously *(Currier & Clandinin, 2025)*. While such large-scale response measurements will greatly benefit new connectome-constrained models, ideally these measurements would be designed specifically for connectome-constrained modeling requirements. This could be done by measuring neural responses based on a vast set of standardized naturalistic stimuli that evoke the majority of possible responses of the system.

Eventually, an important goal is to create a closed loop between modeling and experimentation. Calibrated models can inform the most informative experiments and experiments can be used to update models. While we already create these loops conceptually through our analyses (Section 3.2.4), practical application will benefit from principled methods including a seamless software solution that directly enables experimentalists to work and update the model.

6.3.6 *Computational infrastructure and scalability*

We have not only published our pretrained models, but also published FLYVIS as a tool for abstracting and automating the computational research process from model training to model analysis. While flyvis is still undergoing a lot of enhancements internally, the model *(Wang-Chen et al., 2024)* and general framework *(Özdil et al., 2024)* were already adopted by several other computational and experimental researchers. Currently, our graphical user interface for using the model for hypothesis generation is limited to tutorial notebooks. In future, building a graphical user interface for integrating connectome and model with all hypothesis generation and model refinement tools holds enormous potential to make the model more easily accessible to experimentalists and theorists alike for transforming neuroscience to a more model-driven discipline.

A graphical frontend with a generic API can use various neural simulators as a backend. For instance, FLYVIS, JAXLEY, or SSNTORCH *(Eshraghian et al., 2023)*. The latter supports training spiking neural networks using surrogate gradients, which currently neither FLYVIS nor JAXLEY support.

Future large-scale neural circuit simulation research benefits from methods and interfaces for combined modeling and optimization of graded potential subnetworks with spiking subnetworks (Section 4.2.2). Similarly, it needs robust methods to integrate new neural activity measurements into models (Section 5.4). Future frameworks should ideally report and use model uncertainty and calibration for model refinement.

Connectome-constrained modeling could benefit from moving from a monolithic approach of whole-nervous system modeling to more modularized models that interface with each other — possibly also agnostic to simulator libraries — for reasons such as robust parallelization of research and scalability. This parallels how software engineering in general has moved from monolithic architectures to microservices *(Newman, 2022)*. Sensible first steps towards compatible models simulating subcircuits that can be combined to bigger systems would be creating models whose neural activity is confined to a desired unit system, and, for models agnostic to simulator frameworks, establishing API formats for simulated state variables and gradients. While recent research demonstrates that the fly brain consists of many small circuits largely operating independently of each other *(Pospisil et al., 2024)*, likely more work investigating the structural *(Seguin* *et al., 2025)* and functional modularization, or for building agnostic interfaces, might be required.

Beyond computational infrastructure, another practical challenge we faced was literature integration, for which a recent technological breakthrough might be a straigthforward solution. The comparison across experimental work required a lot of painstaking literature research. Today's LLM assistants can search publication databases and retrieve relevant findings about the fly based on text similarity. Such a database will help use LLM assistance for significantly speeding up model validation across heterogeneous experimental literature.

6.4 SO WHAT?

Neuroscience asks how the brain gives rise to behavior — a question posed at the beginning of this thesis. Even a biological organism as small as the fly thrives in the real world, efficiently integrating multi-modal sensory information to produce behavior with only 140,000 neurons and negligible energy requirements. Understanding how, at single-neuron resolution, remains fundamentally challenging.

The substantial investments in reconstructing nanoscale connectomes have been controversial precisely because these wiring diagrams alone cannot reveal brain function — the electrical dynamics that underlie behavior. Our work directly addresses this controversy by demonstrating that connectomes, when combined with task constraints and machine learning, can accurately predict neural circuit function.

Our deep mechanistic network approach bridges the gap we identified between mechanistically detailed models of small circuits and task-optimized artificial neural networks. By building models where each unit corresponds to an actual neuron in the fly visual system and training them on ecologically relevant tasks, we show that connectomes provide useful scaffolds for understanding neural computation. This represents a crucial first step toward addressing the central challenge: understanding how large brain networks coordinate single-neuron activity to produce behavior.

6.5 CONCLUSION

Our work demonstrates that we can predict neural function from connectome structure by combining anatomical constraints with machine learning. Our connectome-constrained model of the fly visual system — representing 64 neuron types, 45 thousand neurons and 1.5 million neuron-to-neuron connections, equipped with biologically founded dynamics, and trained on a motion detection task — accurately reproduces functional properties from decades of experimental literature and generates novel, testable hypotheses about uncharacterized neurons.

We developed methods to handle non-identifiability through ensemble analysis and refinement with neural data, establishing a general computational framework that transforms connectomes from static diagrams into accurate dynamic models of neural computation. Our DMN approach bridges the gap between detailed biophysical models and task-trained artificial neural networks, enabling mechanistic understanding at single-neuron resolution across a large task-performing nervous system model.

Looking forward, several directions emerged through our work towards building comprehensive whole-brain or even whole-animal models in future work. Including for instance and on a high level: incorporating new connectome data and more realistic biophysical components, developing principled methods for integrating neural activity measurements and behavioral constraints to refine predictions, and creating accessible computational infrastructure that enables a closed-loop between modeling and experimentation, accelerating discovery in systems neuroscience.

The impact of connectome-constrained modeling becomes clear when we consider the scale of discovery it enables. Over the past three decades, experimental studies have characterized the responses of approximately one neuron type per year in the fly visual system. With our approach, we characterized 64 neuron types in five years — an order of magnitude increase (Section 4.1.1). As we scale to whole-brain models with thousands of neuron types, this acceleration in hypothesis generation will fundamentally change how we approach systems neuroscience, necessitating new automated analysis methods and experimental validation strategies.

While we demonstrated our approach using the fly visual system — which has favorable properties for modeling — the deep mechanistic network framework provides a general strategy for connectome-constrained modeling. As discussed earlier (Section 2.3.1), organisms like *C. elegans* with its complete connectome and emerging whole-brain reconstructions of other species each present unique modeling challenges and opportunities.

The principles of our approach — parametrizing unknowns, optimizing via tasks, and distilling predictions from model ensembles — provide a foundation for using connectomes with task constraints and machine learning to reveal how neural structure gives rise to function. As more data becomes available and tools for iterative model refinement advance, these principles can guide future connectome-constrained modeling efforts. Our work represents a crucial step toward more comprehensive and detailed model-driven systems neuroscience.

7

APPENDIX

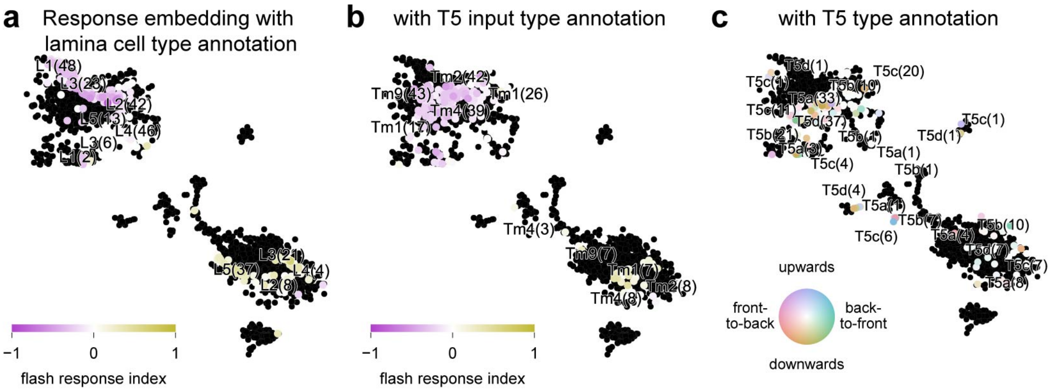


Figure 7.1: **Response embedding, like Fig. 5.7.** With **(a)** labeled lamina neuron types L1-L5, **(b)** labeled T5 input types, and **(c)** labeled T5 neuron types.

7.1 LINK TO 478-PAGE SUPPLEMENT

We generated a supplement *(Lappalainen et al., 2024)* publicly available at https://tinyurl.com/478-pages *(this shortened link redirects to the static ZIP file hosted by the Springer Publishing Group)*, containing the following analyses for each cell type:

1. Anatomical receptive fields
2. Anatomical projective fields
3. Clustering of the responses to naturalistic stimuli
4. Responses to flashes
5. Cluster-average responses to single-ommatidium flashes
6. Peak responses to moving edges
7. Peak responses to moving edges from task-optimal models
8. Responses to moving edges from task-optimal models
9. Peak responses to moving bars
10. Peak responses to moving bars from task-optimal models
11. Responses to moving bars from task-optimal models

12. Spatio-temporal receptive field
13. Maximally excitatory stimuli
14. Task-constrained parameters

7.2 PUBLIC SUPPLEMENTARY INFORMATION

nature portfolio

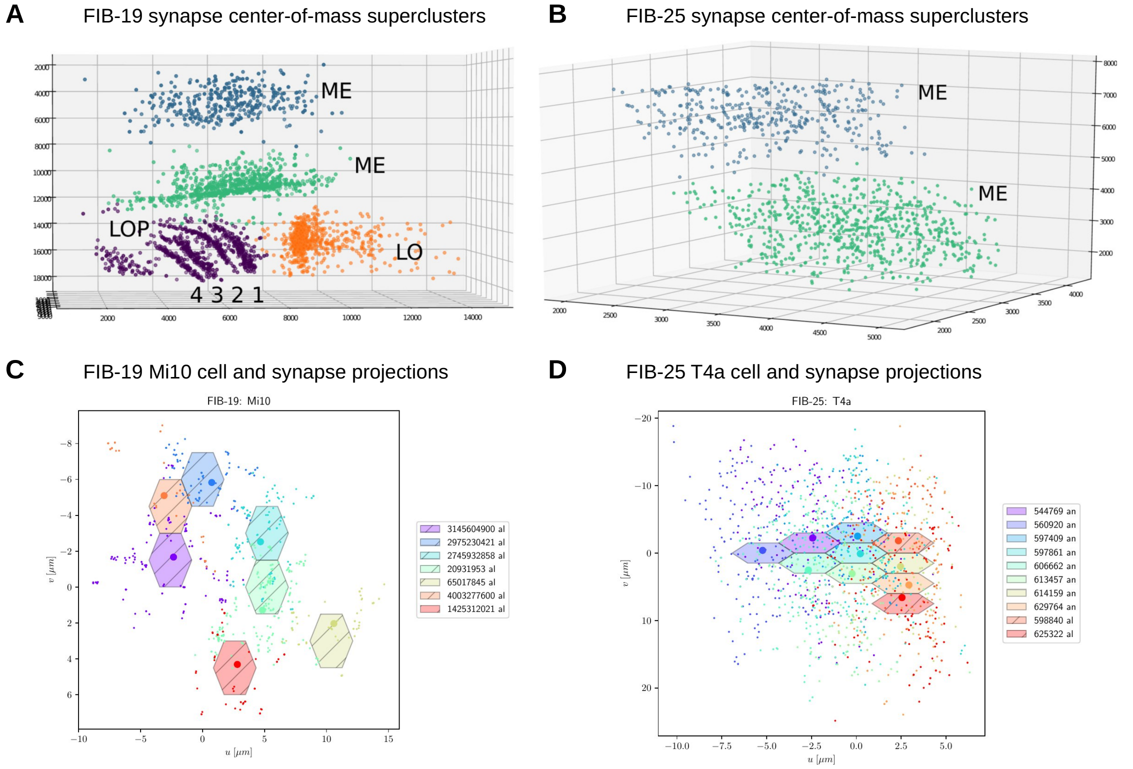
https://doi.org/10.1038/s41586-024-07939-3

Supplementary information

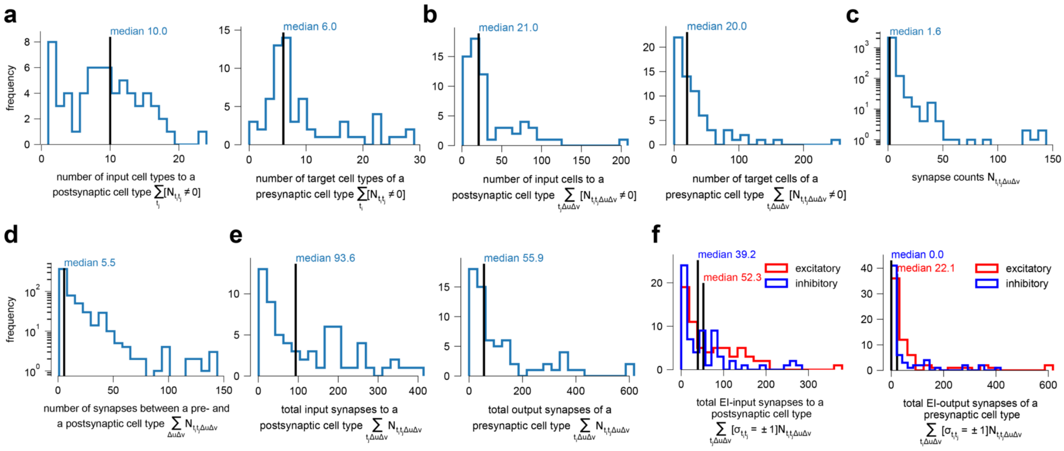
Connectome-constrained networks predict neural activity across the fly visual system

In the format provided by the authors and unedited

Supplementary Information

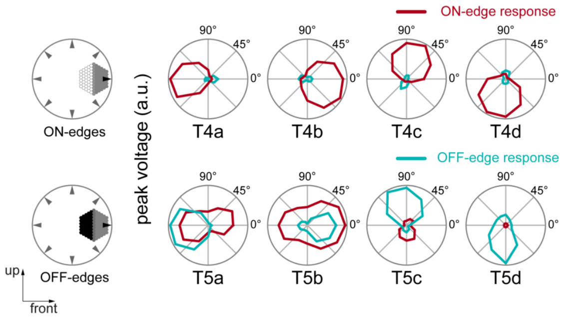


Supplementary Figure 1: **(A)** FIB-19 synapse center-of-mass superclusters. The clusters form two strata in the medulla (ME), and in the lobula (LO) and lobula plate layers (LOP, 1-4) additionally. Each dot corresponds to the center-of-mass of all synapses belonging to the super-cluster. Typically, each diverging arborization of a cell becomes a distinct location, which helps our probabilistic model to project 3D positions of synapses into retinotopic 2D planes, despite the lobula having a different spatial orientation (perpendicular) than the medulla and lobula (Fig. 1a and c). **(B)** FIB-25 synapse center-of-mass superclusters. The clusters form two strata (ME) in the medulla. **(C)** Mi10 cell type in FIB-19 with no pre-annotated lattice positions. The seven cell specimen (hexagons) are recovered by our probabilistic algorithm. Individual synapses and synapse center-of-mass projections are superimposed. **(D)** T4a cell type in FIB-25 with eight pre-annotated (an) and two recovered (al) lattice positions. The projected synapse positions show directional displacement consistent with the direction selectivity of T4a cells (Fig. 3a).

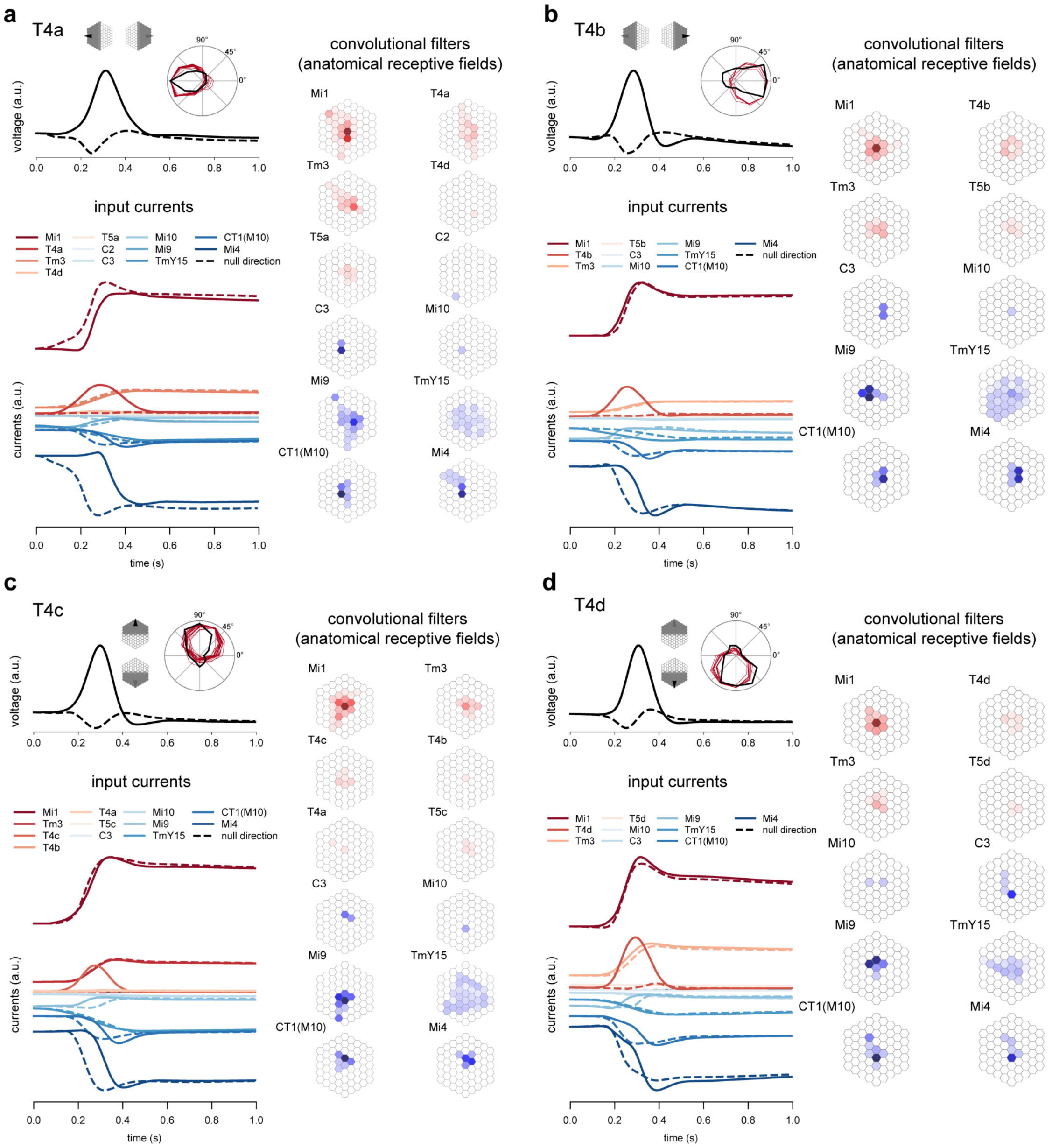


Supplementary Figure 2: **Statistics of derived connectome.** **(a)** (left) Half of the 65 cell types receive input from more than ten other cell types, while the other half receives input from less than ten. (right) Half of the 65 cell types project onto more then six other cell types, while the other half projects onto less than six. **(b)** (left) Half of the 65 cell types receive input from 21 up to 200 cells, while the other half receives input from less than 21 cells. (right) Half of the 65 cell types project output onto 20 up to 200 cells, while the other half projects output onto less than 20 cells. **(c)** Half of the connections are characterized by less than 1.6 synapses while the other half are characterized by 1.6 up to hundreds of synapses. **(d)** A pair of presynaptic and postsynaptic cell type is connected by 5.5 synapses in half of the cases and by more than 5.5 up to hundreds in the other half of the cases. **(e)** (left) Half of the 65 cell types receive input from less than 93.6 synapses and the other half between 93.6 to 400 synapses. (right) Half of the 65 cell types project less than 55.9 synapses and the other half projects between 55.9 to 600 synapses. **(f)** Separating (e) into excitatory and inihibitory synapses, (left) we see that half of the 65 cell types receive excitatory inputs from less than 52.3 synapses and the other half from 52.3 to hundreds. Half of the 65 cell types receive inhibitory inputs from less than 39.2 synapses and the other half from 39.2 to hundreds. (right) Half of the 65 cell types project less than 22.1 excitatory synapses and the other half from 22.1 to hundreds. At least half of the 65 cell types project no inhibitory synapses and the rest project between zero to hundreds.

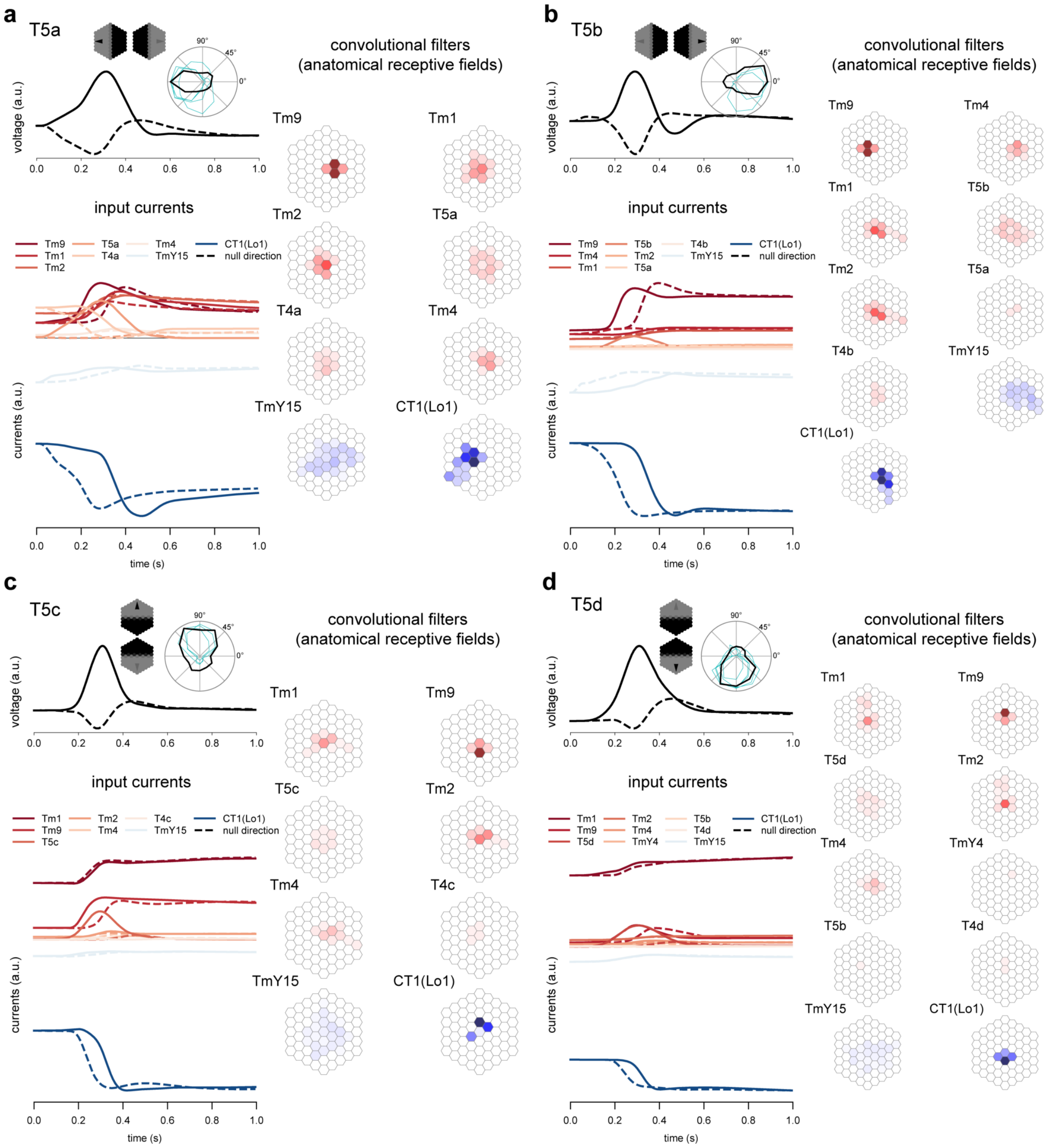
Supplementary Figure 3: **Statistics of learned parameters of best 20% models vs. worst 20% models.** **(a)** Task-optimized resting potentials. **(b)** Task-optimized time constants. **(c)** Task-optimized filter scaling factors.



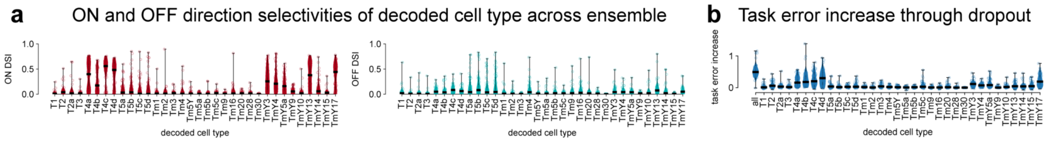
Supplementary Figure 4: **Motion tuning predictions for T4 and T5 subtypes to preferred and null contrast edges in the task-optimal model.**



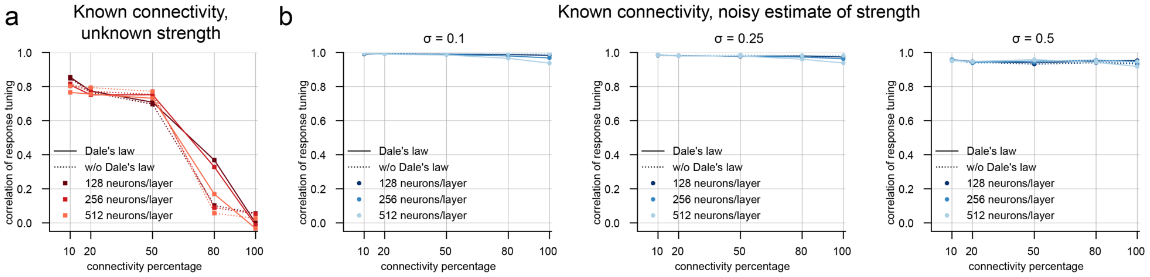
Supplementary Figure 5: **T4 motion detection mechanisms.**



Supplementary Figure 6: **T5 motion detection mechanisms.**



Supplementary Figure 7: **Contribution of all decoded cell types to the motion detection task.** **(a)** Direction selectivity indexes for each decoded cell type and each model from the ensemble of 50 models. (left) Direction selectivity indices from responses to moving ON-edges. (right) Direction selectivity indices from responses to moving OFF-edges. **(b)** Task error increase if all respective decoded cell responses are replaced by their cell type's spatio-temporal averages to the naturalistic input at the decoding stage. This measures the cell type's contribution to the motion detection task. As a baseline, the left-most column shows the per model task error increase if all decoded cell responses are replaced by their cell type's spatio-temporal averages, i.e. no motion can be detected. The decoder attends primarily to the ON-motion selective T4 cell types.



Supplementary Figure 8: **Investigating the role of sparse connectivity with synthetic networks for MNIST handwritten digit recognition.** **(a)** Median hidden-layer response correlation as a function of synthetic network connectivity percentage for task and connectome constrained models that had access to only connectivity information but not connection strength. **(b)** Median hidden-layer response correlation as a function of synthetic network connectivity percentage for task and connectome constrained models with access to noisy estimates of connection strength (multiplicative noise levels of $\sigma = 0.1, \sigma = 0.25$, and $\sigma = 0.5$, respectively). Connectome constrained models were task optimized with a soft (L2) constraint with the noisy connectome measurements.

Supplementary Note 1. Probabilistic model for automatic construction of connectome

The EM-datasets primarily contain the lamina projections, medulla (FIB-25), lobula and lobula plate (FIB-19) cells, and the important cell types of the primary motion detection circuit (T4, T5). In total, they contain 1801 neurons (702 from FIB-25 and 1099 from FIB-19), with hand-annotated positions available for 830 of these neurons (SI Figure 1). To accurately localize the remaining neurons and synapses and to derive cell-type connectivity (Fig. 1b), we build a probabilistic expectation maximization algorithm that takes synaptic connection statistics, projected synapse center-of-mass clusters and existing column annotations into account. We verified the quality of our reconstruction, concluding that even in the absence of 90% of the hand-annotations available to us, we could accurately position the majority of the neurons in our circuit reconstruction (SI Table 1, SI Figure 9). In the absence of ground-truth annotations, we verified the quality of our reconstruction by the *recovery* and *consistency* rates (Table 1). The *recovery* rate is defined as the ratio of reference positions successfully recovered by our algorithm after removing a random proportion of reference positions from the data. For each 10% of the reference positions removed, on average, only 2.5% are not correctly recovered. The *consistency* rate is defined as the fraction of neurons obtaining the same position between evaluations of the algorithm starting with a different fraction of reference positions. For each 10% of the reference positions removed, on average, an additional 3.9% of neurons are not consistently estimated. We found that even just 10% (83 positions) of the available ground truth was sufficient to robustly position the majority of the neurons (64.3%, 534 positions) into the correct columns, and annotate 48% (865 neurons) perfectly consistent.

Probabilistic expectation maximization for unassigned neurons. Each neuron is either annotated in the dataset (in $\mathcal{K}$), assigned to a position (in $\mathcal{A}$) by our algorithm, or still under evaluation (in set $\mathcal{U}$). Iteratively, the EM-expectation step updates the normal distribution ($\mu_{s,t,(y,x)}$, $\sigma_{s,t,(y,x)}$) of expected synapse counts between neurons, while the EM-maximization step updates the positions (y, x) of all neurons not yet assigned to a column (set $\mathcal{U}$).

Synapse center-of-mass as a neuron column position proxy. For offset assignment, we take the center-of-mass of all synapses belonging (pre- or postsynaptic) to a neuron into account. These are generally a more useful hint than the physical location of the cell body, as the cell bodies are mostly positioned on the side of the neuropiles, and not near the column to which the neuron belongs to, but most cells have a majority of synapses in close proximity to their own column. Since many neurons span more than one layer in a neuropile, or even multiple neuropiles, we first group all synapses per neuron into clusters, and then assign the center-of-mass of these clusters to one of up to $N = 5$ super-clusters (approximately matching the medulla, lobula and lobula plate). Clustering is done via k-means with the ideal number of clusters determined by silhouette scores. The super-clusters allow to project 3D synapse coordinates onto a retinotopic 2D hexagonal lattice with a simple projection and affine transformation.

Hybrid cost-model for neuron-position likelihood estimates. Prior knowledge about the normal distribution ($\mu_{s,t,(y,x)}$, $\sigma_{s,t,(y,x)}$) of expected synapse counts between neurons from existing annotations is required to express the probability of any cell specimen c to be located at position (y, x). This metric correlates a neuron to all pre- and postsynaptic neurons it is connected to, of which some already have a fixed, known position. Thereby, the neighbouring neurons with known position (in $\mathcal{K}$ and $\mathcal{A}$) contribute to stabilize the probabilities of unassigned neurons (in set $\mathcal{U}$). For columnar, spatially repeated neurons, we can also assume that only one neuron per position is present. This prior rapidly discounts the number of possible positions an unassigned neuron can have, each time another neuron becomes assigned (moves from $\mathcal{U}$ to $\mathcal{A}$).

Fusion of separately evaluated datasets. Since FIB-19 and FIB-25 are evaluated separately, we have to combine the estimated parameters of both models to a single, coherent model. The datasets overlap partially, in terms of the neurpoiles and cell types covered, and our method therefore fuses the model by always taking the larger estimated parameter.

$$\mu_{s,t,(y,x)} = \max(\mu^{\text{FIB-19}}_{s,t,(y,x)}, \mu^{\text{FIB-25}}_{s,t,(y,x)}) \tag{7}$$

This method of model fusion only underestimates the number of synapses between two neurons if they have connections in two different neuropiles, and if each neuropile is exclusively covered by only one dataset.

Pruning spurious synapses. Some automatic annotations, which are not proof read in the FlyEM DVID data, contain a large number of autapses per neuron on most neuron types, arising from wrongly detected synapses in the cell bodies themselves. Additionally, there are many statistically insignificant single synapses left from the assignment algorithm. We imposed the following additional filter on our estimated model parameters, to remove both autapses and spurious connections with less than one synapse on average.

$$\mu_{s,t,(y,x)} = \begin{cases} \emptyset & \text{if } s = t \wedge (y,x) = (0,0) \\ \emptyset & \text{if } \mu_{s,t,(y,x)} < 1 \\ \mu_{s,t,(y,x)} & \text{otherwise} \end{cases} \tag{8}$$

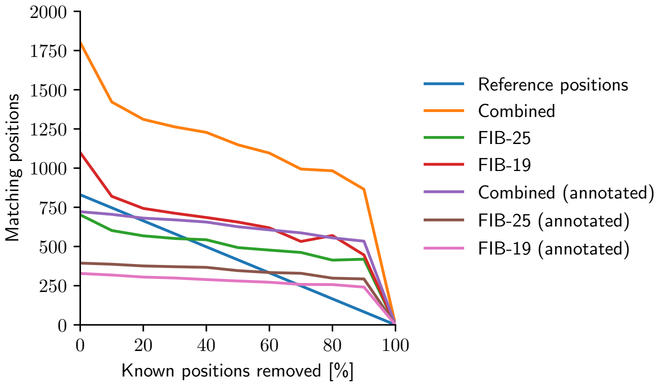
Finally, neuron types without connections and synapses with either missing target or source can be removed. The resulting mean synapse counts $\mu_{s,t,(y,x)}$ form the convolutional filters for our simulation.

Reference positions	1.0	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	average
Recovery rate	**0.870**	0.849	0.820	0.807	0.790	0.754	0.730	0.707	0.669	0.643	
Recovery rate delta		0.020	0.029	0.013	0.017	0.036	0.024	0.023	0.039	0.025	**0.025**
Consistency rate		0.790	0.728	0.701	0.682	0.638	0.609	0.552	0.546	0.480	
Consistency rate delta			0.062	0.027	0.019	0.044	0.029	0.057	0.006	0.066	**0.039**

Table 1: Recovery and consistency of columnar cell position estimation.

Pos. removed [%]	0	10	20	30	40	50	60	70	80	90	100
Reference positions	830	747	664	581	498	415	332	249	166	83	0
Combined	1801	1459	1305	1346	1248	1152	1229	1081	972	944	0
FIB-25	702	592	573	566	538	509	491	464	445	399	0
FIB-19	1099	867	732	780	710	643	738	617	527	545	0
Combined (annotated)	721	708	694	684	658	643	626	576	566	499	0
FIB-25 (annotated)	393	387	379	375	360	353	339	307	304	272	0
FIB-19 (annotated)	328	321	315	309	298	290	287	269	262	227	0

Table 2: Results of the probabilistic model construction.



Supplementary Figure 9: Results of the probabilistic model construction.

Supplementary Note 2. Manually constructed connectome components

Lamina and ommatidia model Since neither FIB-19 nor FIB-25 contain the connections of the ommatidia or first neuropile, the lamina [89, 90], we reused and refined the existing hand-crafted model from our previous work [91], which is based on data from Rivera-Alba *et al.* [92] and Tuthill *et al.* [93, 94].

Non-columnar single CT1 cell model While we did in general not model any neurons with large tangential branches, such as Mt, Mt, Pm, Dm, which span many columns and are therefore insufficiently segmented in FIB-19 and FIB-25, we did model the single CT1 cell present in the lobula CT1(Lo1) and medulla CT1(M10) [95]. Because a multi-compartment model with bidirectional electrical synapses resulted often in oscillatory dynamics in earlier modeling attempts and because CT1 terminals were found to act as functionally independent units [96] we modelled CT1 as two anatomically separate cell types CT1(Lo1) and CT1(M10).

Non-columnar periodic cells Because lamina-wide-field neurons, Lawf1 and Lawf2, do not occur in each individual column but more sparsely, we modeled them with an inferred spatial stride to occur more sparsely resulting in 123 cell of each type in our model (there are approx. 140 Lawf2 neurons per optic lobe, and each of approx. 700 columns is innervated by approx. 5 Lawf2 cells [94]).

Hexagonal lattice rendering of connectome For compilation into the hexagonal grid, the convex hull of the filters is filled with ones to remove spatial discontinuities. Although these are considered mostly false positives from the connectome reconstruction [97], this allowed for weak autapses in our hexagonal model that did not affect the tuning predictions.

Additional proofreading We manually proofread filters on the hexagonal lattice and compared them to the reported filters in the literature to ensure overall correspondence. We found that the reconstruction did not fully capture the asymmetry reported in [90] of the T5 anatomical receptive field of Tm9, which we then substituted by a Gaussian at the reported offset column scaled by the reported number of input synapses. For few T4 and T5 inputs the number of input synapses reported in the literature[90] slightly deviated from our reconstruction. To get a better initialization of our filter scale we scaled them to closely match the number of input synapses reported[90].

Supplementary Note 3. Model predictions agree with experimental measurements of neural activity across 26 studies

We compared simulated model responses for each cell type to experimentally reported responses. We found that the 10 models which achieved the best task performance predict neural responses in consensus with experimental measurements across 26 studies[88, 96, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121].

Supplementary Note 4. Models predict motion tuning for TmY3

Amongst models with the best task performance, TmY3, TmY4, and TmY18 are often ON-motion selective (Fig. 2c). As these neurons have yet to be experimentally characterized, we analyzed these prediction in our models. Since TmY3 neurons do not receive inputs from other known motion selective neurons, we were intrigued by the possibility that they might directly compute a motion signal and possibly constitute a parallel pathway to the well-known T4 and T5 neurons. In contrast, TmY4 and TmY18 cells receive inputs from T4 cells, potentially inheriting their motion tuning.

In the model ensemble, we found four distinct clusters for TmY3 (Fig. 4b). In the task-optimal cluster (circular marker) TmY3 responds to ON-edges from front to back or downwards (Fig. 4c). In contrast, in the second cluster (triangular marker), TmY3 is not direction selective. In the third cluster (square marker) TmY3 is direction selective to ON-edges moving from the back to the front. In the fourth cluster (star marker), TmY3 is, again, not direction selective. Together, the ensemble suggests ON-motion sensitivity for TmY3, but different clusters disagree in their predictions for direction and contrast selectivity.

In our connectome data, the strongest input elements of TmY3 by number of synapses are L4, L5, Tm2, Tm3, Mi1, Mi9, and Mi4 (Fig. 4d and e). While none of these input neurons are motion-selective, the asymmetries in their connectivity to TmY3 might allow it to detect motion. We asked if we could better constrain our predictions by asking which clusters also predicted the correct preferred contrast for these input neurons. We found that the first model cluster (Fig. 4b, circular marker), in which TmY3 is tuned to front-back or downwards motion, most accurately captures the known contrast selectivity of all TmY3 input cells (Fig. 4f). In contrast, all three other clusters fail to consistently capture the OFF-selectivity of Mi9. Thus our model proposes TmY3 as a novel candidate motion detector independent of the well-known T4 and T5 motion pathways with the putative mechanism for the task-optimal cluster shown in Extended Data Fig. 5. Examining the mechanism of direction selectivity, we find null direction suppression but no preferred direction enhancement, consistent with previous experimental findings for T4 [88] and T5 [115].

Supplementary Note 5. Investigating the role of sparse connectivity with synthetic networks for MNIST digit recognition

Training feedforward synthetic networks The weight matrix for each layer in Dale's-law-based synthetic networks (DLTrues) is decomposed into three components: binary adjacency matrix, non-negative weight magnitudes, and a sign vector.

$$\begin{aligned} W_{\texttt{DLTrue}} &= C_{\texttt{DLTrue}} \odot W_{\texttt{DLTrue}} \odot [\mathbf{1} \otimes \mathbf{s}_{\texttt{DLTrue}}], \\ \text{where } C_{\texttt{DLTrue}} &= \text{binary adjacency matrix}, \\ W_{\texttt{DLTrue}} &= \text{non-negative weight magnitudes of true network}, \\ \mathbf{s}_{\texttt{DLTrue}} &= \text{sign vector} \end{aligned}$$

By means of projected gradient descent, $W_{\texttt{DLTrue}}$ is enforced to be non-negative and is initialized from the absolute value of the He initialization distribution [122]. Although sign vector $\mathbf{s}_{\texttt{DLTrue}}$ is randomly initialized with equal probability to be either -1 or +1 to represent inhibitory and excitatory synapses respectively, its elements are allowed to assume values in $\mathbb{R}$ over the course of training.

Inducing sparsity Binary adjacency matrix $C_{\texttt{DLTrue}}$ is initialized to be a unit matrix and is later updated according to the desired true network connectivity level. Following the LTH algorithm, a portion of the lowest-magnitude weights, designated to be pruned, were identified from $W_{\texttt{DLTrue}} \odot [\mathbf{1} \otimes \mathbf{s}_{\texttt{DLTrue}}]$. For pruning synapses, the corresponding entries in the adjacency matrix $C_{\texttt{DLTrue}}$ were then set to zero. After each pruning iteration, weight magnitudes were reset back to their original initialization, followed by a final training run post-pruning.

Training with/without sign constraints In addition to a Dale's-law-based sign constraint, we also experimented with networks trained without any sign constraint. No restrictions were imposed on the nature of outgoing synapses i.e., a neuron can have both excitatory and inhibitory outgoing synapses. For true network variants trained without a sign constraint, $\mathbf{1} \otimes \mathbf{s}_{\texttt{DLTrue}}$ was simply replaced by a sign matrix $S_{\texttt{nonDLTrue}}$ initialized in a similar fashion; that is, all entries were initialized to be in $\{-1, +1\}$ with equal probability. We will refer to true networks trained without a sign constraint as `nonDLTrues`.

$$\begin{aligned} W_{\texttt{nonDLTrue}} &= C_{\texttt{nonDLTrue}} \odot W_{\texttt{nonDLTrue}} \odot S_{\texttt{nonDLTrue}}, \\ \text{where } C_{\texttt{nonDLTrue}} &= \text{binary adjacency matrix}, \\ W_{\texttt{nonDLTrue}} &= \text{non-negative weight magnitudes of true network}, \\ S_{\texttt{nonDLTrue}} &= \text{sign matrix} \end{aligned}$$

Training simulated networks As elements in a true network's $\mathbf{s}_{\texttt{DLTrue}}$ or $S_{\texttt{nonDLTrue}}$ are allowed to assume values in $\mathbb{R}$ while training, only the signs of these elements are inherited by the true networks's respective simulated network.

$$\begin{aligned}
W_{\texttt{DLSimulated}} &= C_{\texttt{DLTrue}} \odot W'_{\texttt{DLSimulated}} \odot \left[\mathbf{1} \otimes \mathbf{s}'_{\texttt{DLTrue}}\right], \\
\text{where } C_{\texttt{DLTrue}} &= \text{binary adjacency matrix from corresponding DLTrue}, \\
W'_{\texttt{DLSimulated}} &= \text{non-negative weight magnitudes of simulated network}, \\
\mathbf{s}'_{\texttt{DLTrue}} &= \text{signs of elements in } \mathbf{s}_{\texttt{DLTrue}}
\end{aligned}$$

$$\begin{aligned}
W_{\texttt{nonDLSimulated}} &= C_{\texttt{nonDLTrue}} \odot W'_{\texttt{nonDLSimulated}} \odot S'_{\texttt{nonDLTrue}}, \\
\text{where } C_{\texttt{nonDLTrue}} &= \text{binary adjacency matrix from corresponding NonDLTrue}, \\
W'_{\texttt{nonDLSimulated}} &= \text{non-negative weight magnitudes of simulated network}, \\
S'_{\texttt{nonDLTrue}} &= \text{signs of elements in } S_{\texttt{nonDLTrue}}
\end{aligned}$$

Extended Data Fig. 8 A shows median hidden-layer tuning correlations for networks trained with and without Dale's law sign constraint for three different architectures.

Training simulated networks that had access to weight magnitudes Three levels of multiplicative noise $\sigma = 0.1, 0.25, 0.5$ were explored, inducing low-noise, medium-noise, and high-noise weight estimates, respectively. Each noise level represents the maximum percentage by which a weight magnitude could be perturbed.

$$\begin{aligned}
W_{\texttt{Simulated\_Init}} &= X \odot W_{\texttt{True}}, \\
\text{where } X_{ij} &= \text{Uniform}(1-\sigma,\ 1+\sigma)
\end{aligned}$$

Simulated networks were trained with a Gaussian prior on the weights centered around the noisy initialization. In effect, this additional loss term penalizes trainable weights for deviating from their noisy initialization.

$$\mathcal{L}_{\texttt{weight\_prior}} = \mathcal{L}_{\texttt{cross\_entropy}} + \lambda \sum \left[W_{\texttt{Simulated}} - W_{\texttt{Simulated\_Init}}\right]^2$$

Extended Data Fig. 8b shows median hidden-layer tuning correlations for networks with low-, medium-, and high-noise weight perturbations.

References for appendix

[88] Gruntman, E., Romani, S. & Reiser, M. B. Simple integration of fast excitation and offset, delayed inhibition computes directional selectivity in Drosophila. *Nature Neuroscience* **21**, 250–257 (2018).

[89] Takemura, S.-y. *et al.* Synaptic circuits and their variations within different columns in the visual system of Drosophila . *Proceedings of the National Academy of Sciences* **112**, 13711–13716 (2015).

[90] Shinomiya, K. *et al.* Comparisons between the ON- and OFF-edge motion pathways in the Drosophila brain. *Elife* **8**, 2431 (2019).

[91] Tschopp, F. D., Reiser, M. B. & Turaga, S. C. A Connectome Based Hexagonal Lattice Convolutional Network Model of the Drosophila Visual System. *arXiv preprint arXiv:1806.04793* (2018). `1806.04793`.

[92] Rivera-Alba, M. *et al.* Wiring economy and volume exclusion determine neuronal placement in the drosophila brain. *Current Biology* **21**, 2000–2005 (2011).

[93] Tuthill, J. C., Nern, A., Holtz, S. L., Rubin, G. M. & Reiser, M. B. Contributions of the 12 Neuron Classes in the Fly Lamina to Motion Vision. *Neuron* **79**, 128–140 (2013).

[94] Tuthill, J. C., Nern, A., Rubin, G. M. & Reiser, M. B. Wide-field feedback neurons dynamically tune early visual processing. *Neuron* **82**, 887–895 (2014).

[95] Takemura, S.-y. *et al.* The comprehensive connectome of a neural substrate for 'ON' motion detection in Drosophila. *eLife* **6**, 1–16 (2017).

[96] Meier, M. & Borst, A. Extreme Compartmentalization in a Drosophila Amacrine Cell. *Current Biology* **29**, 1545–1550.e2 (2019).

[97] Scheffer, L. K. *et al.* A connectome and analysis of the adult drosophila central brain. *Elife* **9**, e57443 (2020).

[98] Peretz, A. *et al.* The light response of drosophila photoreceptors is accompanied by an increase in cellular calcium: effects of specific mutations. *Neuron* **12**, 1257–1267 (1994).

[99] Reiff, D. F., Plett, J., Mank, M., Griesbeck, O. & Borst, A. Visualizing retinotopic half-wave rectified input to the motion detection circuitry of drosophila. *Nature neuroscience* **13**, 973–978 (2010).

[100] Clark, D. A., Bursztyn, L., Horowitz, M. A., Schnitzer, M. J. & Clandinin, T. R. Defining the computational structure of the motion detector in drosophila. *Neuron* **70**, 1165–1177 (2011).

[101] Freifeld, L., Clark, D. A., Schnitzer, M. J., Horowitz, M. A. & Clandinin, T. R. Gabaergic lateral interactions tune the early stages of visual processing in drosophila. *Neuron* **78**, 1075–1089 (2013).

[102] Silies, M. *et al.* Modular use of peripheral input channels tunes motion-detecting circuitry. *Neuron* **79**, 111–127 (2013).

[103] Maisak, M. S. *et al.* A directional tuning map of Drosophila elementary motion detectors. *Nature* **500**, 212–216 (2013).

[104] Strother, J. A., Nern, A. & Reiser, M. B. Direct observation of on and off pathways in the drosophila visual system. *Current Biology* **24**, 976–983 (2014).

[105] Meier, M. *et al.* Neural circuit components of the drosophila off motion vision pathway. *Current Biology* **24**, 385–392 (2014).

[106] Behnia, R., Clark, D. A., Carter, A. G., Clandinin, T. R. & Desplan, C. Processing properties of ON and OFF pathways for drosophila motion detection. *Nature* **512**, 427–430 (2014).

[107] Fisher, Y. E. *et al.* A Class of Visual Neurons with Wide-Field Properties Is Required for Local Motion Detection. *Current Biology* **25**, 3178–3189 (2015).

[108] Hardie, R. C. & Juusola, M. Phototransduction in drosophila. *Current opinion in neurobiology* **34**, 37–45 (2015).

[109] Leonhardt, A. *et al.* Asymmetry of drosophila on and off motion detectors enhances real-world velocity estimation. *Nature neuroscience* **19**, 706–715 (2016).

[110] Fisher, Y. E., Silies, M. & Clandinin, T. R. Orientation Selectivity Sharpens Motion Detection in Drosophila. *Neuron* **88**, 390–402 (2015).

[111] Yang, H. H. *et al.* Subcellular imaging of voltage and calcium signals reveals neural processing in vivo. *Cell* **166**, 245–257 (2016).

[112] Serbe, E., Meier, M., Leonhardt, A. & Borst, A. Comprehensive characterization of the major presynaptic elements to the drosophila off motion detector. *Neuron* **89**, 829–841 (2016).

[113] Strother, J. A. *et al.* The emergence of directional selectivity in the visual motion pathway of drosophila. *Neuron* **94**, 168–182 (2017).

[114] Arenz, A., Drews, M. S., Richter, F. G., Ammer, G. & Borst, A. The Temporal Tuning of the Drosophila Motion Detectors Is Determined by the Dynamics of Their Input Elements. *Current Biology* **27**, 929–944 (2017).

[115] Gruntman, E., Romani, S. & Reiser, M. B. The computation of directional selectivity in the drosophila off motion pathway. *Elife* **8**, e50706 (2019).

[116] Drews, M. S. *et al.* Dynamic signal compression for robust motion vision in flies. *Curr. Biol.* **30**, 209–221.e8 (2020).

[117] Matulis, C. A., Chen, J., Gonzalez-Suarez, A. D., Behnia, R. & Clark, D. A. Heterogeneous temporal contrast adaptation in drosophila Direction-Selective circuits. *Curr. Biol.* **30**, 222–236.e6 (2020).

[118] Ramos-Traslosheros, G. & Silies, M. The physiological basis for contrast opponency in motion computation in drosophila. *Nature communications* **12**, 1–16 (2021).

[119] Gruntman, E., Reimers, P., Romani, S. & Reiser, M. B. Non-preferred contrast responses in the drosophila motion pathways reveal a receptive field structure that explains a common visual illusion. *Current Biology* **31**, 5286–5298 (2021).

[120] Groschner, L. N., Malis, J. G., Zuidinga, B. & Borst, A. A biophysical account of multiplication by a single neuron. *Nature* **603**, 119–123 (2022).

[121] Ketkar, M. D. *et al.* First-order visual interneurons distribute distinct contrast and luminance information across on and off pathways to achieve stable behavior. *Elife* **11**, e74937 (2022).

[122] He, K., Zhang, X., Ren, S. & Sun, J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*, 1026–1034 (2015).

7.3 PUBLIC PEER REVIEW FILE

Peer Review File

Manuscript Title: Connectome constrained networks predict neural activity across the fly visual system

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referees' comments:

Referee #1 (Remarks to the Author):

This manuscript addresses the fundamental question whether connectomes (mutual connectivity between many neurons) can predict functions of individual neurons in a network when the computation of the network is known. The authors address this question in the visual system of Drosophila. They use recent connectomics datasets to create simulations of the visual neuropils with biologically realistic neuron types, spatial arrangement, and connectivity. Parameters describing individual neurons and synapses were adjusted by training networks to detect optic flow. Training improved model performance compared to untrained models with realistic connectivity. Some networks, but not others, reproduced known functional properties of individual neurons such as directional motion sensitivity of T4 and T5 cells. Model neurons showed more biologically realistic properties when connectivity was sparse.
The authors conclude that network models reproducing biological mechanisms of computation can be found by including connectomes as constraints in networks trained end-to-end on a computational task (“DMNs”). They emphasize that it is not necessary to include additional information from neuronal recordings or biophysical measurements. They further propose connectome-constrained network modeling as a tool to discover computational functions and meaningful experiments. These conclusions are based on the observation that some models reproduced functional properties of single neurons such as the motion-sensitive T4 and T5 cells and therefore reproduce the biological mechanism of motion detection. This is indeed a remarkable result. But other models also detect optical flow and fail to reproduce single-neuron properties. Additional knowledge from single-neuron recording experiments (or other sources) is therefore still required to select biologically realistic models. It is also not clear how the approach may or may not generalize to other systems given that the Drosophila visual system is a highly specialized, repetitive, extensively studied network of graded potential neurons. In summary, I believe that the results of this study are remarkable and demonstrate that including knowledge about synaptic connectivity into network simulations can take us a huge step forward towards mechanistic modeling of biological neural networks. This is clearly an important result but it remains open whether the approach can be generalized with similar success to other brain circuits.
The main result of the study is the general notion that mechanistic modeling of biological neural networks can be achieved by including connectomes as constraints. So far, the study did not produce major new insights into the function or structure of visual processing in Drosophila. A prediction from this study is that neuron TmY3 may be a previously unrecognized motion-sensitive neuron, but this prediction remains to be tested experimentally.

Specific points:
1. The main message of this manuscript is not a deeper understanding of visual processing in Drosophila but the general notion that connectomes help (massively) to create biologically realistic network models. Indeed, the authors emphasize that their "…modeling approach provides a discovery tool…". It is thus important to get a good idea how the DMN approach generalizes to other brain circuits. The DMN approach is likely to be facilitated by features of the Drosophila visual system such as graded synaptic transmission and a highly repetitive architecture. More insight into the potential for generalization of this approach would be useful (see also below).

2. The graded potential neurons used in this study are biologically realistic for the Drosophila visual system but not for most other brain circuits. It may be more difficult to get DMNs to reproduce biological mechanisms of computation with spiking neurons. This may be a limitation of the approach that should be discussed (or explored, if possible).

3. The repetitive layout of the Drosophila visual system facilitates network modeling, and so does the extensive knowledge of cell types in this system. How would the DMN approach be affected if knowledge of the connectome were less complete, as is often the case in other brain circuits? How would it be affected if cells were divided into fewer distinct types?

4. An interesting observation is that the same neurons in different networks do not always show the same functional properties and form discrete clusters in functional space. For T4c cells, for example, 3 clusters were found but only one represents biologically realistic neurons with correct motion sensitivity. This observation is interesting because it can, in principle, be a starting point to explore general principles of network design. On the other hand, it means that the DMN approach alone is not sufficient to predict mechanisms of computation, even in this well-established system. Additional knowledge (here: true motion sensitivity of T4 cells) is necessary to distinguish biologically realistic from unrealistic networks. Such knowledge may be hard to come by in other systems. This is a (potentially serious) limitation of the DMN approach that needs to be discussed more. What type of additional information would be most useful to resolve "cluster ambiguities"?

5. The authors emphasize that DMNs can reproduce biological mechanisms of computation, but they do not go very far in analyzing computational mechanisms beyond current knowledge. So far, they mainly asked whether known mechanisms are reproduced in the DMN. For example, motion sensitivity of T4 cells involves direction-dependent temporal shifts between excitatory and inhibitory input currents, which is reproduced by the model. However, the authors could go further and manipulate specific connections to verify that motion sensitivity is indeed generated by the expected implementation of a computational strategy (combination of Hassenstein-Reichardt and Barlow-Levick models) in neural circuitry. Such an analysis should have potential to uncover novel, unknown functions. Similarly, they could use specific manipulations of connectivity to analyze the mechanisms of motion sensitivity in TmY3, following up on the speculations put forward in the text. Generally, the ability to manipulate connections in a biologically realistic simulation has interesting potential because this is often not possible experimentally.

6. The authors suggest that TmY3 is a novel motion-sensitive neuron that has not been recognized previously and computes motion independently from T4 and T5 cells. It is also predicted that other

neurons should be motion sensitive (TmY4, TmY18), probably because they receive input from T4 and T5. So far, these predictions have not been tested experimentally. Doing so could substantially enhance the impact of this study.

7. Abstract: "…we show that with only measurements of the connectivity of a biological neural network, we can predict the neural activity underlying neural computation". This statement is not correct. DMNs also use knowledge of the computation (input-output) for training, and additional knowledge is required for model selection (for example activity/tuning of T4/T5 cells).

8. How good is the optic flow detection achieved by DMNs? The quantification by the error measure is not very intuitive. It shows that training improves performance, but unconstrained CNNs still achieve much better performance than trained DMNs. It would be good to provide more information to get a better intuition how well a DMN is performing in comparison to a real fly.

Minor comments:
9. Ln 481: Fig 1g

Referee #2 (Remarks to the Author):

The authors use connectome data from the fly visual system combined with optic flow training to produce a task-performing mechanistic model with interpretable parameters. Comparing to previous data, the model captures many of the tuning properties of fly visual neurons.

There has been some previous work that uses connection data to define model architectures for task-training. The authors may want to cite some of this work from C. elegans (e.g. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8253844/ and https://arxiv.org/pdf/2201.05242.pdf). The present work is novel in my opinion in the extent to which it compares the neural parameters of the trained model to data.

It would probably be good to provide a contextualization of the DMN task performance and parameter recovery by pulling in some more of the supplementary results (at least more quantified descriptions of them in the text). For example, the main text does not convey how relatively minor the enhancement in performance on the task is for the DMN vs random model, in the context of how well the unconstrained CNN can perform. Also, it seems relevant to note that the random models are still positively correlated with cell tuning and the flash response results can be captured by even the poorly performing DMNs.

Are the T5 off-motion selective neurons in 2c supposed to be tuned to on-motion as well? Or is this a way in which the model does not fit the data? This should be spoken to in the paper.

I have several concerns/questions regarding the synthetic connectome experiments in the MNIST-trained networks. First, I do not understand the motivation behind the version with noisy weight estimates where getting the correct weights is baked into the objective function. What do we learn

from seeing that a network initialized with roughly the correct weights can explicitly learn to recover those weights (regardless of sparsity)?
For the connectome-only version, this still does not seem to necessarily support what the authors seem to be claiming about it. Specifically in the absence of any information that leads to unique cell IDs, comparison of tuning across networks is meaningless. With a sparse network (including the fly connectome) the pattern of connections a cell makes can be a unique identifier for it, and therefore these cells can be labeled as the same and their tuning can be compared across networks. Such unique identities are not possible in densely connected networks. Therefore, the tuning comparisons done here are essentially as if two random neurons were picked across models and expected to have the same tuning. The fact that a random pairing of neurons does not display the same tuning does not mean these networks are not learning the same mechanisms. In fact, ED Fig 9a shows that Dale's law helps with the correlation, which is likely because having a weight constraint offers some kind of (weakly) unique identifier. Also, how were the signs decided for units in these networks?

Clarifications:
Can the authors better explain the differentiation between synapse count and the scalar? It seems the scalar and the count are held constant for all pairs of cells with the same pre- and post-synaptic cell type, so what extra information does having the count as part of the equation provide?

How do the 50 ensemble models differ? Just different draws from the same distribution of resting potentials?

Referee #4 (Remarks to the Author):

Lappalainen et al. build an optic lobe connectome-constrained neural network called a task-optimized deep mechanistic network (DMN), and optimize for a computation performed by that biological circuit (motion detection). They show that constraining both the connectivity and computational task reproduces some of the experimentally determined tuning of specific neurons, and makes predictions about the tuning properties of other neurons in the network that have yet to be experimentally measured. Finally, the paper argues that, for sparse networks (such as some biological neural networks), knowledge of the connectivity, signs of connections, and an estimate of connection strength may be sufficient to predict the mechanism by which the circuit performs a known computational task.

I am in general enthusiastic about the study - it is a useful simulation of a portion of real, complex neural network (64 cell types and 721 columns, plus 1 inhibitory cell type that extends across columns, but missing all of the feedback and neuromodulatory connections) and shows how connectivity shapes many of the known properties of a neural network - the optic lobe is an ideal test case as its cell types have been studied extensively over the past 60 years. Figures 3 and 4 in particular are quite nice - 1) comparisons between the best performing model's T4/T5 cells and the response properties of their inputs to known tuning curves and responses (although I have concerns about some of the details - see below) and 2) comparisons between different models to understand what properties of particular cell types define the best performing models. This is a nice study that will form the basis for simulations of larger biological networks, for fly, and other species. However, I have several concerns about the modeling, model predictions, and interpretations of the results that should be addressed.

The authors build a hybrid optic lobe connectome from several different datasets - the choices they made in how to combine these datasets must be made transparent in the paper. Ideally, they would present a Supplemental Figure devoted to how this was done and how they handled any discrepancies or differences between the datasets. If there were no discrepancies or differences this should also be explained. It is difficult to interpret the findings from the model without a thorough understanding of how the connectome-constrained model was built. Related to this point, I assume the signs of connections were taken from the literature? Can the authors provide citations for these (and can they compare against the same cell type in the open FlyWire/FAFB whole-brain connectome dataset, for which neurotransmitter predictions (from Eckstein et al.) are available in the optic lobe)?
Related to the point above, the authors perform a sort of normalization step with the data so that they can model every column identically (making sure synapse numbers are the same in each column) - this ignores heterogeneity across columns (that might be important for motion detection). Can the authors provide more detail on how this simplification deviates from the actual connectivity (how much heterogeneity is there across columns?) and show how this choice affects modeling results (if they incorporate some of the heterogeneity into the model, how do the results change)?
The manuscript claims that the actual OL connectivity was critical. Extended Data Figure 3 shows that a task-optimized DMN outperforms a DMN with random parameters. But what is missing is a demonstration that the specific connectivity of the fly visual system is what enables optimal performance, rather than a generic neural network with the same level of sparseness and gross connectivity statistics of the biological network. How would a task-optimized artificially generated network (constrained by biological connectivity statistics, rather than the exact connectivity of the fly visual system) perform compared to the task-optimized DMN and the random DMN?
Neurons are modeled with leaky linear non-spiking voltage dynamics and as point-neurons with a single electrical compartment. The authors show abundant tuning/response data for each model cell type, but they should compare the detailed temporal dynamics and delays (critical for motion detection) of model responses to real recordings of these same neurons (if this is present somewhere in the supplement, apologies if I missed it). Many optic lobe neurons have been recorded via Ca++ imaging, voltage imaging, or electrophysiology (e.g., compare with published responses in Behnia et al. Nature 2014 or Yang et al. Cell 2016).
The L1 and L2 neurons are categorized as “known OFF selective” (Fig. 2b), but this deviates from my reading of the literature - shouldn’t they show similar responses to both on and off flashes? Also, the responses of L1 and L2 in Fig. 3e are shown as monophasic and producing an off response - shouldn’t they be biphasic (in contrast with the responses for L3 and L4, which do look biphasic)? These potential mismatches between the literature and model results have me concerned that the model is not producing the expected responses for cell types that have been extensively studied (like L1 and L2).
Given the constraints provided by the connectome, there are only 734 free parameters in the model

(resting membrane potential for each cell type (65) and unitary synapse strength (604)) - but the authors could have also included the synapse NL as a free parameter (varying across synapses) - how would this have affected modeling results? Is there a reason they need to limit to ~700 free parameters? The question has to do with the choice of which parameters to fix across the model and which to vary - how much do these choices affect results?

The authors optimize the model network to perform a particular motion vision task - this makes a lot of sense given that the major function of the optic lobe is to detect motion, but it is not clear how the results depend on this specific task - this needs to be addressed. What differences might they observe or expect if the network was optimized to perform a different task that the fly optic lobe mediates (for example, color or shape detection) or a specific fly behavioral task (like the optomotor response) - how would optimizing for a different task change the responses properties of neurons in the network? Addressing this question is critical for understanding the constraints of the connectome.

Model responses to moving edges in Fig. 2c: Why do many of the models show that T5 is responsive to ON moving edges (which it is not)? Could this be expanded upon in Figure 4 (that compares different models)?

I'm concerned about claims that the mechanism of motion detection matches experimental findings: The mechanism of how direction selectivity in T4 and T5 cells emerges is still debated. Haag et al. suggest there is both PD enhancement and ND suppression; Gruntman et al. argue for ND suppression only; Wienecke et al. argue for neither. It seems likely that the mechanism is different for the ON and OFF pathways. The manuscript shows that the tuning of T4 and T5 cells and their inputs (in some models) qualitatively matches the experimentally measured tuning, but can the authors clarify which mechanism of direction selectivity is matched/favored by their models?

The authors have missed an opportunity to test their model through silencing experiments - inputs to T4 and T5 cells have been silenced experimentally (Strother et al. 2017; Serbe et al., 2016) - does silencing these neurons in the trained models recapitulate experimental findings/impair motion detection by the decoder network?

Of the 19 cell types with asymmetric inputs, only 12 are predicted to be motion selective - can the authors comment on differences between the inputs of these 12 and the other 7?

One of the most exciting findings is the prediction that TmY3 could possess direction selectivity independent of the T4/T5 pathway. This is a bold claim - that the model can be used to identify new motion sensors in the fly visual system that has been studied for more than 60 years. It seems reasonable (and feasible given that the authors' local collaborators) to ask the authors to validate this prediction with experimental data - whereas elsewhere the authors rely on published experimental data for comparisons, this prediction would require new experiments. Further, the authors could discuss if TmY3 is expected to function like an HR detector or BL detector, or is direction selectivity predicted to arise via a different mechanism?

I would like to see some of the statistics of the decoder network. The manuscript argues that it cannot compute motion itself, but is this true? Does the decoder only depend on the output of direction-tuned neurons? Does it perform equally well when the weights for non-direction tuned neurons are forced to zero? What is the minimal set of T and Tm cells required to detect optic flow?

Minor comments

Figure 1 depicts a male fly, though much of the connectome data used, I believe, comes from female

flies.
Winding et al. 2023 larval connectome paper should also be cited at line 7.
The meaning of lines 99-101 is not clear. The text argues that the DMN was optimized for a computational task performed by the real biological network (i.e. fly visual system), but citations 44 & 76 refer to mammalian cortex.
Line 119: typo in “backpropagation”
Line 105 says that motion detection is a “challenging” computation. “Challenging” is too subjective a term and can be removed. Arguably, theoretical models for motion detection proposed in the 1960’s by Hassenstein & Reichert and Barlow & Levick are relatively simple.
Lines 283-285 claim that networks with different sparsity must use different computations. Why must this be true?
Lines 333-336 claim that “DMN models generate meaningful predictions in absence of neural activity measurements… (Fig 4)” feels too strong, since knowing actual tuning of Mi9 was critical in determining “correct” tuning of T4.
The limitations of the approach (e.g. simplistic modeling of neural dynamics, and lack of electrical synapses, neuromodulation and glia) are introduced at the outset. It would be nice if these limitations were further elaborated/explored in the discussion.
Line 1038: Missing reference to extended data figure.

Author Rebuttals to Initial Comments:

We thank Editor and the reviewers for the constructive and detailed comments, and for appreciating the importance of the study. We are excited about the assessment of our results as *remarkable* (R1) and *a huge step towards mechanistic modeling of biological neural networks* (R1) , as *novel* (R2), and that our *study will form the basis for simulations of larger biological networks, for fly, and other species* (R4). In our study, we provide an approach for turning a connectome into detailed hypotheses of how the neural circuit works, and which neurons are involved in which computations. We show that our approach makes concrete predictions at the level of single-cell responses, which are surprisingly accurate, as we showed by testing its predictions against measurements made across 26 studies (and our study makes a large number of additional predictions).

Remarkably, connectome-constrained neural networks seem to learn solutions which are surprisingly similar to the ones implemented in the fly, at the level of yielding predictions for individual cell tuning and even circuit mechanisms. This finding goes vastly beyond previous 'NeuroAI' approaches e.g. in the mammalian cortex, in which the correspondence between artificial and biological networks was weaker, largely at the level of brain regions, and which provided limited mechanistic circuit-level insights.

Importantly, our methodology provides a new approach to generate meaningful hypotheses before making any activity measurements. The timeliness and importance of our approach and findings is underscored by recent developments: High-profile releases of connectomes have provided an unprecedented abundance of neural connectivity measurements, including many neurons from which activity measurements are unavailable (and likely will not be available for some time). This data gap highlights the dire need for frameworks to extract an understanding of how neural systems perform computations, and hence the importance of our approach.

Indeed, in a comment in Nature last month ("How AI could lead to a better understanding of the brain", 07/11/23), Viren Jain writes that " *Guided by connectomic and other data to optimize thousands or even billions of parameters, machine-learning models could be trained to produce neural-network behaviour that is consistent with the behaviour of real neural networks — measured using cellular-resolution functional recordings." and "Researchers could evaluate such models, for instance, by comparing their predictions about the neural activity of a system with recordings from the actual biological system.* " This is, precisely, what our study is successfully doing. We believe that the reviewers appreciated both the importance of the study, and the central messages.

At the same time, their comments also revealed opportunities for strengthening our manuscript. In addition to multiple clearer explanations, the updated manuscript now includes substantial additional analyses (details below):

1. To show that our models can be used to study circuit mechanisms underlying specific computations, we analyzed circuit mechanisms of direction selectivity in T4, T5 and TmY3, both by inspecting the current distributions from each cell type and by simulating inactivation experiments which we compared to published experimental measurements, finding yet more agreement. For instance, while research into the mechanism of T4 motion selectivity has largely focused on the role of feedforward inputs, our new analysis shows a novel prediction suggesting an important role for the significant lateral connectivity between T4 neurons enhancing responses to coherent motion across the visual field.
2. We conducted a large set of additional numerical experiments to analyze how specific constraints [cell connectivity, synapse counts, synapse signs] contribute to the prediction of tuning properties. These results provide a clear account that all of these constraints are necessary for obtaining a close match between model activity and empirically measured tuning.
3. We now also provide a detailed analysis of the temporal properties of major inputs to T4 and T5 cells, providing an example of how our modeling approach can be used as a hypothesis generated for detailed dissection of neural circuits.

Referee #1 (Remarks to the Author):

This manuscript addresses the fundamental question whether connectomes (mutual connectivity between many neurons) can predict functions of individual neurons in a network when the computation of the network is known. The authors address this question in the visual system of Drosophila. They use recent connectomics datasets to create simulations of the visual neuropils with biologically realistic neuron types, spatial arrangement, and connectivity. Parameters describing individual neurons and synapses were adjusted by training networks to detect optic flow. Training improved model performance compared to untrained models with realistic connectivity. Some networks, but not others, reproduced known functional properties of individual neurons such as directional motion sensitivity of T4 and T5 cells. Model neurons showed more biologically realistic properties when connectivity was sparse.

The authors conclude that network models reproducing biological mechanisms of computation can be found by including connectomes as constraints in networks trained end-to-end on a computational task ("DMNs"). They emphasize that it is not necessary to include additional information from neuronal recordings or biophysical measurements. They further propose connectome-constrained network modeling as a tool to discover computational functions and meaningful experiments. These conclusions are based on the observation that some models reproduced functional properties of single neurons such as the motion-sensitive T4 and T5 cells and therefore reproduce the biological mechanism of motion detection. This is indeed a remarkable result. But other models also detect optical flow and fail to reproduce single-neuron properties. Additional knowledge from single-neuron recording experiments (or other sources) is therefore still required to select biologically realistic models.

It is also not clear how the approach may or may not generalize to other systems given that the Drosophila visual system is a highly specialized, repetitive, extensively studied network of graded potential neurons. In summary, I believe that the results of this study are remarkable and demonstrate that including knowledge about synaptic connectivity into network simulations can take us a huge step forward towards mechanistic modeling of biological neural networks. This is clearly an important result but it remains open whether the approach can be generalized with similar success to other brain circuits.

The main result of the study is the general notion that mechanistic modeling of biological neural networks can be achieved by including connectomes as constraints. So far, the study did not produce major new insights into the function or structure of visual processing in Drosophila._A prediction from this study is that neuron TmY3 may be a previously unrecognized motion-sensitive neuron, but this prediction remains to be tested experimentally.

Thank you for your comments, and for appreciating the importance of the central question of our study and its results-- indeed, we also find it remarkable that a large number of single-cell properties can be predicted from connectome- and task-constraints alone. Our large circuit model can be mapped onto individual cells and performs motion computation while capturing an impressive amount of biological realism.

We should clarify that our goal with this study was to evaluate the utility of a connectome by demonstrating how far we can get while only relying on measurements of connectivity---essentially a show-case of how powerful connectomic measurements can be. While we do still find the degree of accuracy and specificity of the predictions remarkable, we also emphasize that it would be unrealistic to expect such a model to be correct in all its details.

We did not intend to claim that experimental measurements of neural activity are not needed. We wholeheartedly agree that measurements and perturbations of neural activity and behavior will be essential for going beyond the current results. And in our paper, we show an example of how such measurements can be incorporated to refine our hypotheses [l. 197-222, Fig. 3].

We will respond to your additional questions and concerns (how important are single-neuron recordings for selecting realistic models, how well does this generalize to other circuits, predictions about TmY3) below.

Specific points:
1. The main message of this manuscript is not a deeper understanding of visual processing in Drosophila but the general notion that connectomes help (massively) to create biologically realistic network models. Indeed, the authors emphasize that their "...modeling approach provides a discovery tool...". It is thus important to get a good idea how the DMN approach generalizes to other brain circuits. The DMN approach is likely to be facilitated by features of the Drosophila visual system such as graded synaptic transmission and a highly repetitive architecture. More insight into the potential for generalization of this approach would be useful (see also below).

Thank you for appreciating our central results. Indeed we have demonstrated our method in the fruit fly optic lobe, where until recently our understanding of circuit connectivity was most comprehensive, with 64 cell types and their connectivity mapped. The study of the optic lobe is supplemented by numerical experiments with MNIST. In principle, we cannot rule out the possibility that our method will only work in the fruit fly visual system (note, however, that even then the recent full-brain connectomes would still yield a deluge of data to apply our method on…). But we can clarify why we believe the fly visual system will not be unique:

1. repetitive architecture: in our modeling, the highly repetitive architecture was exploited to extrapolate a connectome for the whole eye. With a full connectome, the number of unknown parameters in our model would still remain identical, and the model fitting procedure would remain largely the same (very minor difference: convolutions would be replaced by sparse matrix multiplications). The parameters in our model only depend on the number of cell types, and not on their spatial organization. With full connectomes for the fly visual system now becoming available, it will become possible to directly compare the predictions of convolutional models to non-convolutional ones. [However, we note that most published activity measurements also aggregate across columns, so fine-grained validation of non-convolutional models will be challenging.] Finally, we note that the mammalian retina is another model system for which our architecture will be directly applicable, and for which one will also have to use repetitive structure to extrapolate incomplete connectome reconstructions.
2. graded synapses: Yes, our model is based on graded synapses. However, we note that the network equations resulting from our modeling choices result in overall dynamics which are given by threshold linear network dynamics. Threshold linear dynamics have been used extensively to approximate the firing rates in spiking neurons with non-graded (quantal) synapses on a wide range of circuits, and our formalism will likewise be useful for building connectome-constrained models of such circuits.

3. More generally, the approach of constructing a connectome constrained deep mechanistic network is not confined to the specific single neuron and synapse model used in this study. Indeed we believe a more expansive approach where we also search over the space of single neuron and synapse models would be productive and enable the inference of the best model class for each system. The main idea behind our work is to use machine learning to constrain a model simultaneously with the connectome and a computational task.

We agree that it is important to understand the generality of the DMN technique. We eagerly await the mapping of more connectomes and future work modeling other circuits and other model organisms. In the revised paper, we have provided additional clarifications and explanations of these points in the Discussion [l. 360-371].

2. *The graded potential neurons used in this study are biologically realistic for the Drosophila visual system but not for most other brain circuits. It may be more difficult to get DMNs to reproduce biological mechanisms of computation with spiking neurons. This may be a limitation of the approach that should be discussed (or explored, if possible).*

We agree that standard backpropagation through time is not readily applicable as a gradient estimator for spiking neural networks, so that optimization of spiking neurons might pose additional computational challenges. However, there has been much recent progress in the training of spiking neural networks with surrogate gradient methods, (e.g. Zenke et al 2018, Neftci et al 2019, Wang et al 2020) Further, reinforcement learning algorithms which have been used to optimize robotics/physics simulations without the need for gradients are also readily applicable to train spiking networks. We will note that, in Mi et al ICLR 2022, we already demonstrated the training of a more complex (albeit still non-spiking) biophysical synapse model. Therefore, while this is definitely an area where more research will be important for future studies modeling spiking circuits, we do not think that this poses a limitation to the applicability of the DMN approach. We have clarified this in the Discussion.

3. The repetitive layout of the Drosophila visual system facilitates network modeling, and so does the extensive knowledge of cell types in this system. How would the DMN approach be affected if knowledge of the connectome were less complete, as is often the case in other brain circuits?

The hexagonally convolutional structure of the optic lobe, and the extensive knowledge of cell-types in it, was primarily useful for two reasons: First, it allowed us to construct a network model even from incomplete connectomic measurements, as we were able to use the assumption of columnarity to ‘fill in’ missing measurements. Second, it allowed us to validate the model by comparing predictions of the model with the extensive literature on estimates of single-cell selectivity. Third, the availability of cell-types reduces the number of free parameters (if one assumes parameters to be shared across cell-types).

However, the approach could equally well be applied in a setting one has no spatial structure in the connectome whatsoever, and possibly even not (agreed-upon) cell-types: Indeed, in Figure

6, we demonstrate (on the MNIST dataset) that single-tuning in a neural network can be recovered from dense connectomic measurements-- this example neither has convolutional (i.e., columnar) structure, nor does it rely on any notion of cell types. Therefore, a repetitive structure is not critically needed for the DMN approach. In particular, the approach can be directly applied, e.g., to any other brain area in the fruit fly brain/VNC, or dense reconstructions of cortical tissue in the mouse or zebrafish, or the mammalian retina (which even also follows a stereotyped spatial arrangement).

Obviously, *some* way to link model-predictions with experimental measurements is required to validate DMNs--- this could be a notion of cell-types (to compare model predictions to either single-cell tuning, or aggregate statistics across many types of a cell), or a means to directly perturb single-cells. However, this is true for any computational network model in neuroscience.

To more concretely illustrate the importance of connectomic constraints and cell-types for scaling-up brain models with single-cell fidelity, it is useful to consider the case of a whole-Drosophila DMN: This would have about 130,000 neurons (Lin et al 2023). Assuming an unconstrained RNN with passive point neuron voltage dynamics and instantaneous graded release synapses, we would have to find 16,952,040,000 connection strengths and 260,400 neuron parameters in this case (time constant and resting potential per cell). With the approximately 4,200 cell types (Schlegel et al 2023) reported from the connectome and approximately 30,435 cell-type to cell-type connections that are conserved across the hemispheres, our DMN approach would reduce the number of free parameters in a whole-Drosophila DMN to approximately 38,835 (time constant and resting potential for each cell type and 30,435 synapse count scaling factors). I.e. only 0.00023% of the number of parameters that need to be estimated in the naive setting without the connectome-- we posit the DMNs will make it possible to build whole-brain models of behavior from connectomes.

We tested these ideas explicitly and now address this question in Fig 2d and in the Results [l. 171-191].

How would it be affected if cells were divided into fewer distinct types?

The DMN approach is independent of the number of distinct cell types. Cell types facilitate the understanding of neural circuits in general, and our approach leverages cell types to reduce the number of parameters in our model, by assuming that cells of the same type share the same parameters. If there are fewer cell types, we would thus have fewer (free) parameters -- conversely, if there are more cell types (or cells which can not be assigned to any type, and therefore need to have their own parameters) we would have more parameters. Of course, the degree to which this assumption (cells of the same type share parameters) is an empirical question which might have different answers for different model systems. Note that, in Figure 6, we demonstrate a setting in which we do not have any cell types at all and still derive accurate neural tuning predictions due to the sparse nature of synaptic connectivity.

In the future, it is likely that a deeper understanding of the function of specific neurotransmitters, receptors, and synaptic morphology, combined with the ability to infer such synaptic functional parameters from electron microscopy (Eckstein 2020) will enable us to share synaptic parameters based on these, rather than cell types.

4. An interesting observation is that the same neurons in different networks do not always show the same functional properties and form discrete clusters in functional space. For T4c cells, for example, 3 clusters were found but only one represents biologically realistic neurons with correct motion sensitivity. This observation is interesting because it can, in principle, be a starting point to explore general principles of network design. On the other hand, it means that the DMN approach alone is not sufficient to predict mechanisms of computation, even in this well-established system. Additional knowledge (here: true motion sensitivity of T4 cells) is necessary to distinguish biologically realistic from unrealistic networks. Such knowledge may be hard to come by in other systems. This is a (potentially serious) limitation of the DMN approach that needs to be discussed more. What type of additional information would be most useful to resolve "cluster ambiguities"?

We agree that connectome+task-constraints, by themselves, are unlikely to be able to identify a *unique* model in general. We should clarify, we deliberately did not utilize the "well-established" nature of this circuit in constructing the model, since we did not use the extensive neural activity measurements in the model construction. In that sense, our work attempts to show how far we can get with as little experimentation in the living animal as possible. Indeed, it is very easy to construct DMN models to be additionally constrained with neural activity and perturbation experiments, as well as measurements of behavior. We propose that connectome constrained DMN models should serve as an integral part of the hypothesis-experiment loop. As we show in Fig 3, DMN ensemble hypotheses can be used to suggest targeted experiments distinguishing between model classes. And these new experiments would then be used to refine hypotheses via new DMN models constrained with all available measurements of connectivity, neural activity, and behavior.

We clarify that many predictions, for instance the contrast preferences of most neurons, are correctly predicted without the need for further experimental measurements. In particular, all analyses in Figure 2 [prediction of on/off tuning on flash responses, direction selectivity] and Figure 4 [DMNs largely recapitulate known mechanisms of motion computation], are across the entire ensemble of networks, and are not conditioned on the true motion sensitivity of T4 cells.

We do use true contrast preferences to show (in Figure 3~~5~~) how to deal with cases in which the ensemble is not uniquely constrained, and show (arguably remarkably) how these constraints do identify a small number of highly specific and experimentally testable hypotheses. The goal of our analysis was to highlight how our general approach which generates multiple hypotheses can be combined with further experimental measurements. The DMNs demonstrate that these constraints lead to a handful of highly specific and experimentally testable predictions for neural tuning: One single tuning measurement (in this case T4c) is sufficient to identify the correct cluster, and thereby to also constrain the selectivity of the other cells in the circuits. [Even for T4c, the hypothesis cluster with the 'correct' tuning actually has the best average task-performance, which is used in Figure 4]. We use a similar approach in Figure 5, to narrow down predictions for the tuning of TmY3-- while, across the ensemble, there is a variety of predictions for TmY3, selecting only ensemble-members with the correct T4c tuning leads to clear predictions for TmY3.

We do not agree that the inability to perform experimental measurements of neural activity in a given circuit is a “(potentially serious) limitation of the DMN approach”. With the connectome and our DMN approach, at least we can make a small number of hypotheses for such a circuit [as demonstrated in Figures 3 and 5]. In contrast, without the connectome+DMN we would have no hypotheses at all. Obviously, the issue of not being able to make physiological measurements is unrelated to DMNs---in the end measurements will always be useful or even critical, but we do expect that DMNs will be a very powerful way to dramatically reduce the number of neurophysiological measurements needed to characterize complex neural systems.

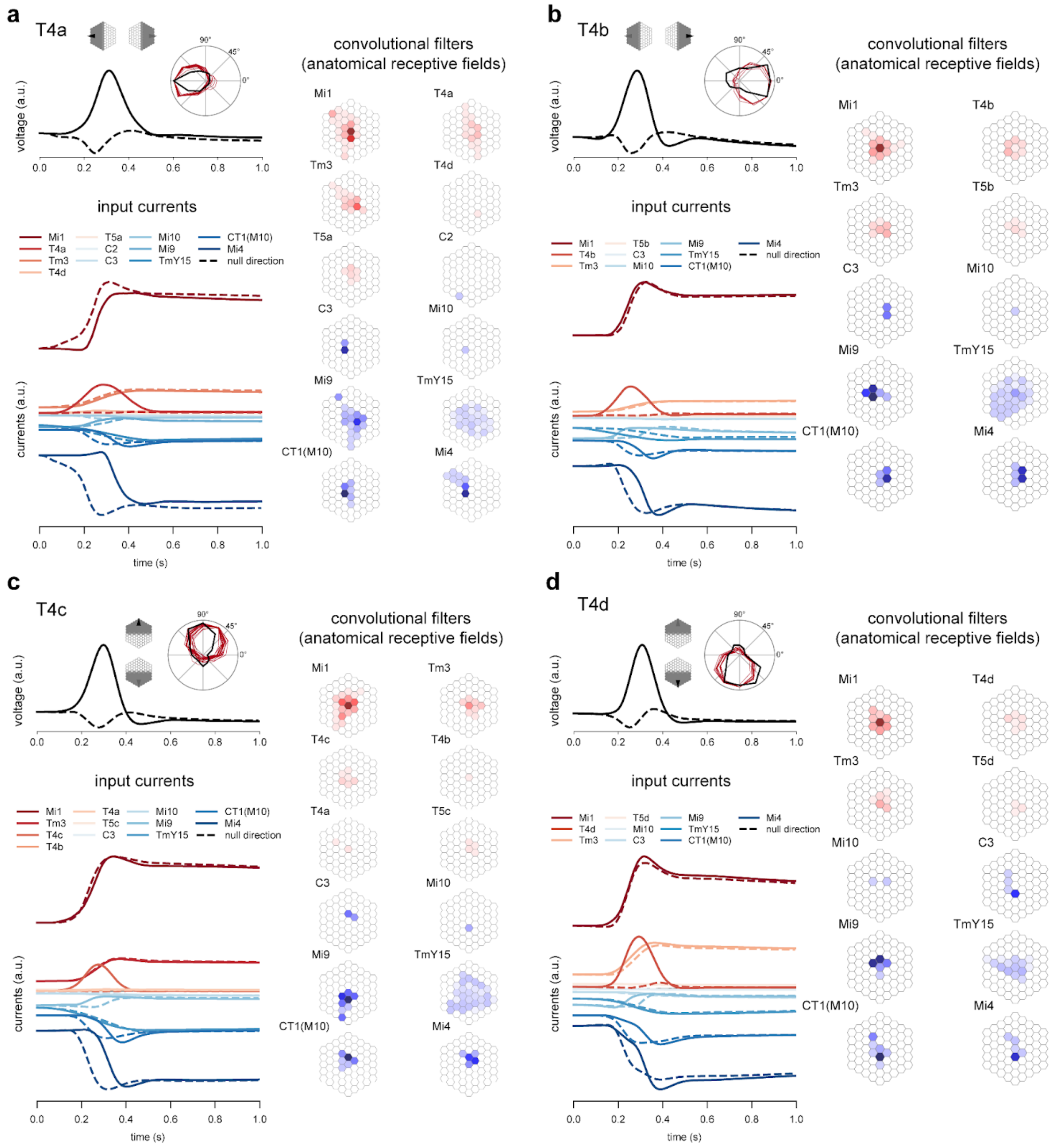
We have clarified this point in the Discussion [l. 381-388].

5. The authors emphasize that DMNs can reproduce biological mechanisms of computation, but they do not go very far in analyzing computational mechanisms beyond current knowledge. So far, they mainly asked whether known mechanisms are reproduced in the DMN. For example, motion sensitivity of T4 cells involves direction-dependent temporal shifts between excitatory and inhibitory input currents, which is reproduced by the model. However, the authors could go further and manipulate specific connections to verify that motion sensitivity is indeed generated by the expected implementation of a computational strategy (combination of Hassenstein-Reichardt and Barlow-Levick models) in neural circuitry. Such an analysis should have potential to uncover novel, unknown functions. Similarly, they could use specific manipulations of connectivity to analyze the mechanisms of motion sensitivity in TmY3, following up on the speculations put forward in the text. Generally, the ability to manipulate connections in a biologically realistic simulation has interesting potential because this is often not possible experimentally.

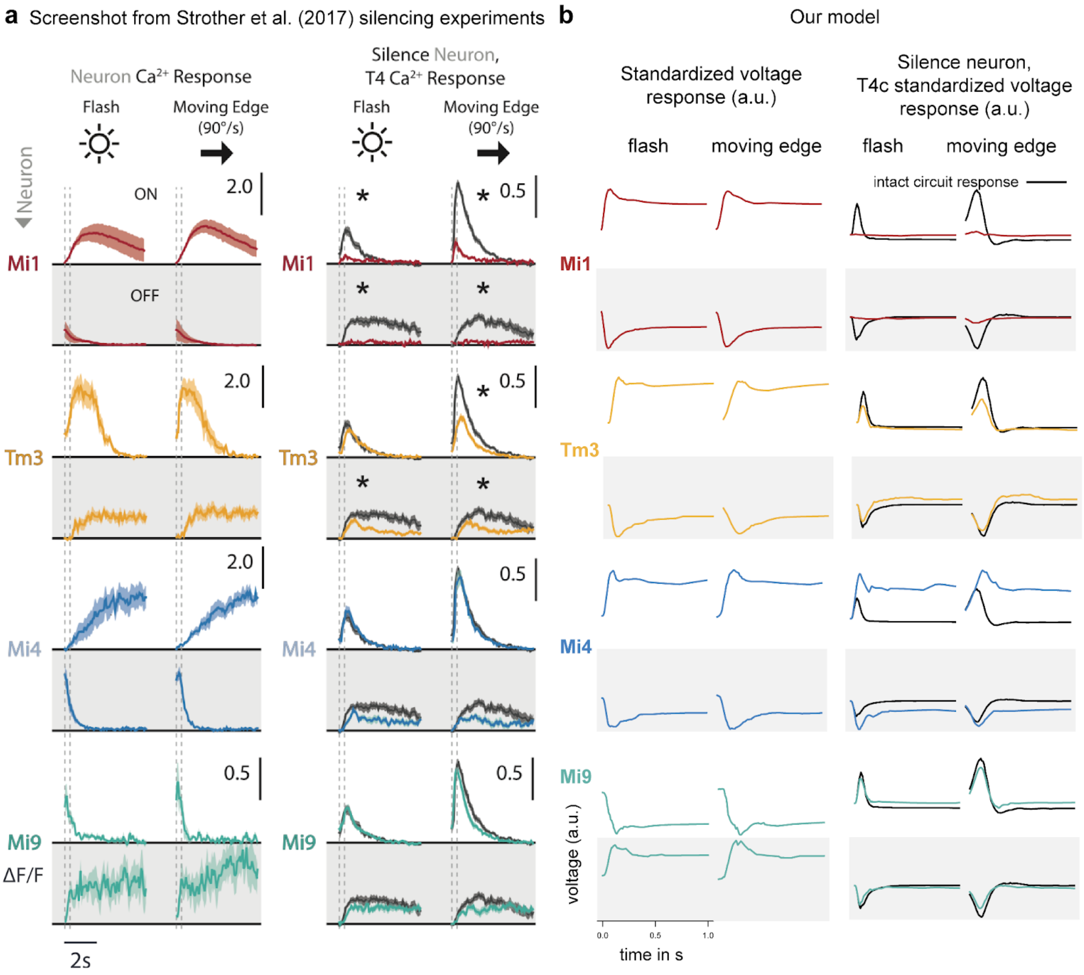
We are grateful for this suggestion. Based on your suggestion, we have now added extensive new analyses and a completely new section that explores these questions: We have further studied the circuit mechanism of direction selectivity in T4, T5, and TmY3 neurons through a combination of techniques and indeed find a combination of preferred direction enhancement (Hassenstein-Reichardt) and null direction suppression (Barlow-Levick) mechanisms for T4, T5, and TmY3.

In our mechanistic model, we can inspect the input current contributions from each cell type and study the differences in these contributions for motion stimuli in the preferred and null directions. This allows us to directly inspect the contribution of neurons to a computation without needing to perturb the circuit through inactivation [however, we also report predictions for inactivation experiments below, see below]. For T4 neurons, we observe null direction suppression mediated by inhibition from Mi4, and also significant enhancement of coherent visual motion mediated by excitatory T4 to T4 connectivity, attributing a role for this lateral connectivity. For T5 neurons, we observe null direction suppression mediated by CT1 inhibition and consistent excitatory input from neighboring T5 to T5 neurons and by excitatory input from Tm9. For TmY3, we see null direction suppression predominantly via Mi14 and Mi4 inhibition and excitation via Mi1 and L5.

We summarize these findings now in the main manuscript by adding text in the Results [l. 237-249], Fig 4b showing input current contributions of T4c. And Extended Data Figs 4, 5, 6 showing detailed input current analysis for preferred and null direction stimuli for all T4 and T5 subtypes, and for TmY3. Extended Data Fig 4 showing input current analysis for T4 subtypes is copied below.



In addition, we now simulated the requested silencing of T4 inputs, and compared model predictions to experimental results reported in Strother et al, Neuron, 2017, finding good agreement for most effects, as shown below. For these results, we averaged model-predictions over all models in the model-cluster with correct T4c tuning (which is also the best task-performing cluster). The two columns in panel a below are direct copies of Strother et al 2017 Figures 3b [left column] and 3c [right column], and panel b shows, for comparison, the output of our analyses. We emphasize that, for these analyses, only a qualitative comparison of the effect is meaningful-- for example, Strother et al report Delta F/F from calcium imaging, whereas we show standardized voltage responses. In addition, our 'silencing' analyses are based on simply clamping the respective T4c inputs to 0, which is likely a crude approximation of the effect of blocking synaptic transmission with shibire(ts1) and temperature increases. Nevertheless, the models correctly capture that: i) Removal of Mi1 excitation removes the T4c response ii) Tm3: Removal of Tm3 excitation decreases, but does not remove, the T4c response. iii) Mi4: Removal of Mi4 does not remove the T4c response [it does, however, lead to a prolonged response in the network model which is not observed experimentally] iv). Mi9: no effect on T4c.



6. The authors suggest that TmY3 is a novel motion-sensitive neuron that has not been recognized previously and computes motion independently from T4 and T5 cells. It is also predicted that other neurons should be motion sensitive (TmY4, TmY18), probably because they receive input from T4 and T5. So far, these predictions have not been tested experimentally. Doing so could substantially enhance the impact of this study.

This study has been conducted by computational labs enabled by the availability of this rich and large dataset. We agree that experimental validation of this prediction would be very exciting--but as computational labs, we do not have direct access to experimental resources, and are therefore eagerly awaiting whether this prediction will bear out. Nonetheless, we have been informed by our colleague Michael Reiser, that unpublished work in progress in his lab does indeed hint at motion selectivity for TmY3. He says the whole picture for this cell type is more interesting and complex and they plan to report the results of their study in the near future.

In the meantime, we do wish to emphasize that no neural activity measurements were used in constructing our model, therefore we have already tested the predictions of our model against experimental measurements of neural activity across 26 studies. Furthermore, we provide a large supplement with hundreds of pages worth of experimentally testable predictions, of which this is but one prediction. We do hope that these predictions, and the fact that we are willing to publish them and ask the community to verify or refute them, already strongly speak to the utility and realism of our model approach.

7. Abstract: "...we show that with only measurements of the connectivity of a biological neural network, we can predict the neural activity underlying neural computation". This statement is not correct. DMNs also use knowledge of the computation (input-output) for training, and additional knowledge is required for model selection (for example activity/tuning of T4/T5 cells).

We apologize for the confusion. We meant to highlight the fact that our model was constructed using experimental measurements of only connectivity, and not also neural activity, etc. Indeed, knowledge of the computation was also necessary. We have now updated the sentence to read "We show that with **experimental measurements of only the connectivity** of a biological neural network, we can predict the neural activity underlying **a specified** neural computation."

8. How good is the optic flow detection achieved by DMNs? The quantification by the error measure is not very intuitive. It shows that training improves performance, but unconstrained CNNs still achieve much better performance than trained DMNs. It would be good to provide more information to get a better intuition how well a DMN is performing in comparison to a real fly.

We agree that comparisons of optic flow estimation between the model and the real fly would be interesting. One challenge, however, is that our model is focused on computation of local motion (in particular, the model does not include LPTCs which spatially integrate local motion signals), whereas experimental characterization (e.g. optomotor responses) has focused on *global* computation of motion by the real fly on relatively simple stimuli (moving gratings). Indeed, the qualitative tuning of behavioral responses depends strongly on the actual behavioral task (see e.g. Creamer et al 2018).

Please note that the goal of task optimization is simply to constrain the parameters of the connectome network enough to make good predictions of neural activity. The accuracy of neural activity predictions is our main goal and the absolute accuracy of the optic flow estimation is not as relevant. Indeed, because we use a blackbox motion decoder network to predict optic flow, the nature of this decoder could lead to better or worse detection of optic flow compared to the real fly, even if the resulting connectome network predicted neural activity with perfect accuracy.

Minor comments:

9. Ln 481: Fig 1g

Thank you, fixed.

Referee #2 (Remarks to the Author):

The authors use connectome data from the fly visual system combined with optic flow training to produce a task-performing mechanistic model with interpretable parameters. Comparing to previous data, the model captures many of the tuning properties of fly visual neurons. There has been some previous work that uses connection data to define model architectures for task-training. The authors may want to cite some of this work from C. elegans (e.g. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8253844/ and https://arxiv.org/pdf/2201.05242.pdf). The present work is novel in my opinion in the extent to which it compares the neural parameters of the trained model to data.

Thank you for your comments and appreciating the novel contributions of our work. We agree that the two studies the reviewer mentioned are relevant work, and have now cited them in the revised manuscript [l. 102]. At the same time, we do want to emphasize that our work goes substantially beyond these two studies: Sakamoto et al trains a model of 69 motor cells and 95 muscle cells to reproduce realistic locomotion patterns in C elegans, and shows that a connectome-constrained network can be successfully trained to solve this task. Bhattasali et al train a neural network architecture inspired by C elegans locomotion circuits, and analyzes the properties of the resulting networks (e.g. in terms of inductive bias).

However, *neither* of these two studies perform any comparison of the neural activity predicted by the model with experimental measurements. Thus, the primary contribution of our paper-- namely, that task-trained connectome-constrained models can predict neural activity remarkably well-- is a clear advance over both of these studies. We achieved these results in a vastly more complex model system (fruit fly visual system vs. C elegans), which also required us to address substantial engineering challenges (e.g., implementing differentiable simulations of convolutional recurrent networks defined on hexagonal grids).

It would probably be good to provide a contextualization of the DMN task performance and parameter recovery by pulling in some more of the supplementary results (at least more quantified descriptions of them in the text). For example, the main text does not convey how relatively minor the enhancement in performance on the task is for the DMN vs random model, in the context of how well the unconstrained CNN can perform. Also, it seems relevant to note that the random models are still positively correlated with cell tuning and the flash response results can be captured by even the poorly performing DMNs.

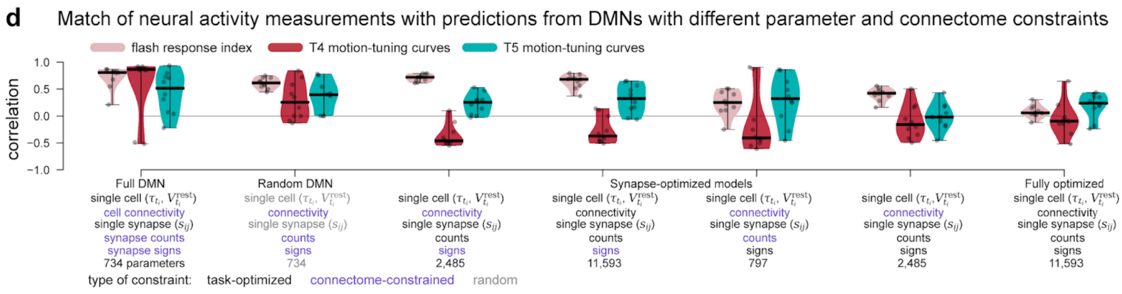
Thank you for this excellent suggestion which has helped us to considerably expand our study:-- You are raising an important point about how (quantitatively) good different models are at predicting different tuning properties, and how specific constraints [cell connectivity, synapse counts, synapse signs] contribute to the prediction of tuning properties.

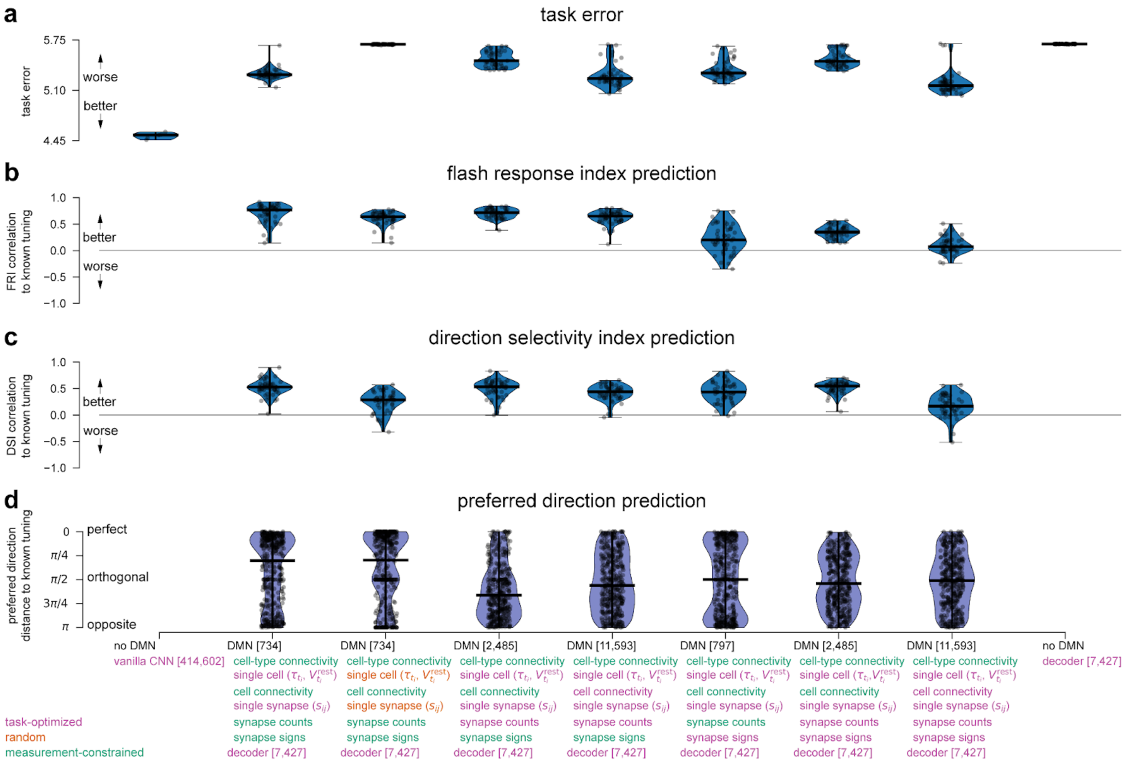
Your comment (and a related comment by reviewer 4, see below) inspired us to perform a large set of additional experiments to tackle this question more systematically. Briefly, we constructed a set of 9 different model ensembles for which we systematically varied which parameters were constrained by connectomic measurements, or which were set by task-optimization. We find that:

1. The preferred contrast as measured by flash response index (FRI) is generally well predicted across model ensembles provided with connectome derived connectivity at the cell-type resolution and synapse signs. Thus, cell-type connectivity and synapse signs are crucial, but all other parameters --- including synapse counts and single cell parameters --- can be randomized or task optimized, and are *not* critical for predicting FRIs.
2. DSI: Accurate predictions of the direction selectivity index (DSI) requires cell-connectivity. However, it does *not require synapse counts or* synapse signs.
3. Preferred directions: To achieve the correct cardinal direction tuning, we need *all* connectome constraints [i.e., all of cell-connectivity, synapse counts, synapse signs].

Thus, these results are still entirely consistent with our overall findings, but provide a detailed picture of the relative importance of different constraints. We are very grateful for your suggestion which we believe to have considerably strengthened the paper. A summary of these additional results has now been added to Figure 2 of the main paper (new Figure 2d, see below), a brief description of these results is now described in the corresponding section in Results [l. 171-191], and a new Extended Data figure 9 (see below) includes a full summary of results.

Your second question as to why unconstrained CNNs outperform the DMNs-- we do believe that this is simply because of the flexibility of the unconstrained network, as they have 414,602 free parameters to transform the naturalistic movie sequences to pixel-wise motion (instead of only 734 free parameters plus 7,427 decoder parameters for the DMNs). This is consistent with general findings in deep learning that, in most settings, task-training models with more parameters typically leads to better performance, given enough training data.

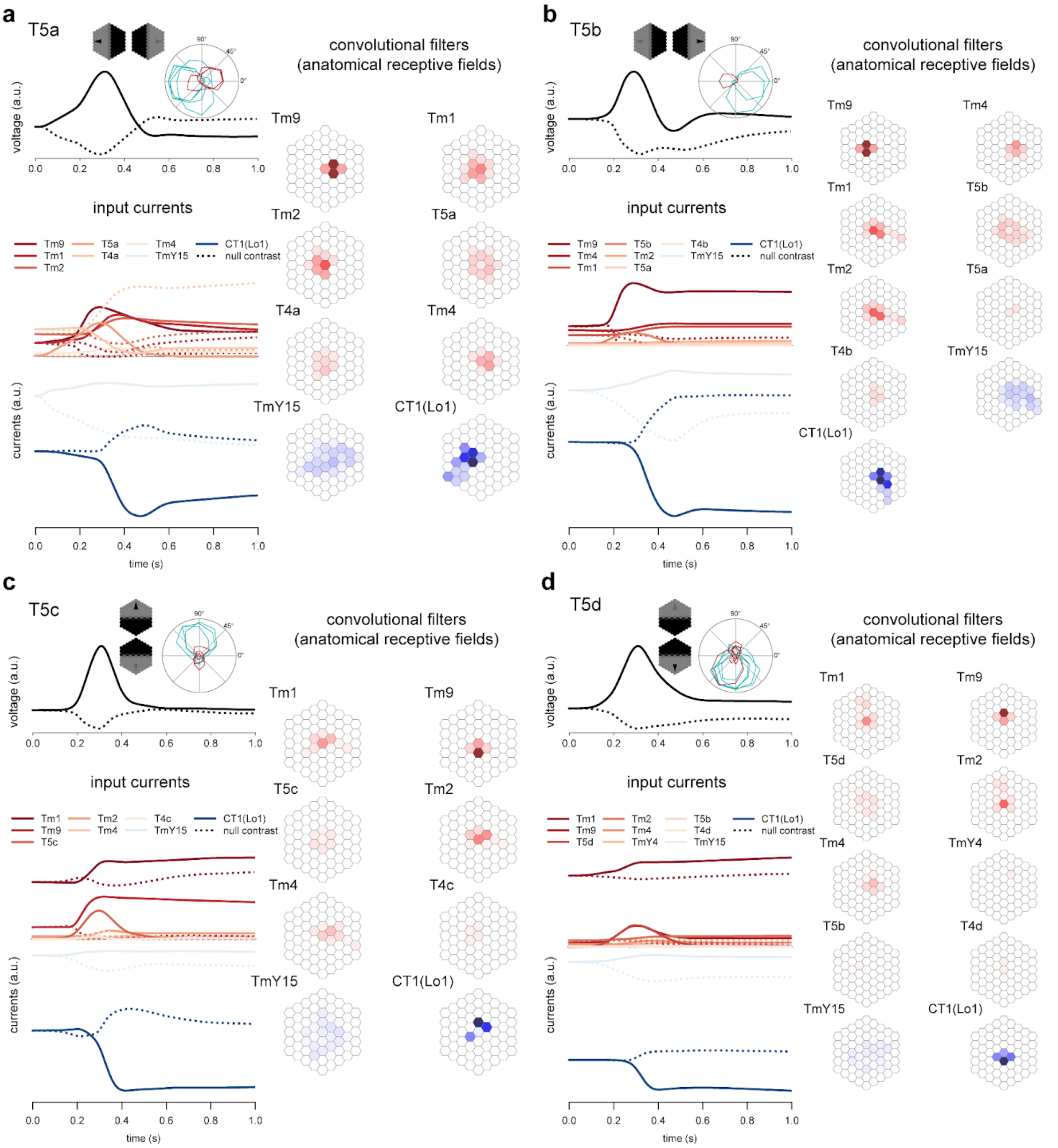




Are the T5 off-motion selective neurons in 2c supposed to be tuned to on-motion as well? Or is this a way in which the model does not fit the data? This should be spoken to in the paper.

We have clarified this in the updated version of the manuscript in the Results [l. 231-236] and added Extended Data Fig 13. Comparing the relative strength of direction selectivity to edges of ON vs OFF contrast, we find that all four T4 subtypes are correctly predicted to respond more strongly to ON edges than OFF, and three out of four T5 subtypes are correctly predicted to respond more strongly to OFF than ON edges (Extended Data Fig. 13).

Across our model ensemble, we find that T5 neurons are predicted to be off-motion selective in more models than on-motion selective (p=0.0009). However, we do find substantial variability in the prediction for T5 on-motion selectivity. When we also filter the ensemble to only select models which exhibit the correct contrast-tuning [to flashes] *and* the correct preferred directions, then these models only predict very weak on-motion tuning for T5 cells. For this selection of models, we show below the model responses to motion in the preferred contrast vs non-preferred/null contrast.



I have several concerns/questions regarding the synthetic connectome experiments in the MNIST-trained networks. First, I do not understand the motivation behind the version with noisy weight estimates where getting the correct weights is baked into the objective function. What do we learn from seeing that a network initialized with roughly the correct weights can explicitly learn to recover those weights (regardless of sparsity)?

For the connectome-only version, this still does not seem to necessarily support what the authors seem to be claiming about it. Specifically in the absence of any information that leads to unique cell IDs, comparison of tuning across networks is meaningless. With a sparse network (including the fly connectome) the pattern of connections a cell makes can be a unique identifier for it, and therefore these cells can be labeled as the same and their tuning can be compared across networks. Such unique identities are not possible in densely connected networks. Therefore, the tuning comparisons done here are essentially as if two random neurons were picked across models and expected to have the same tuning. The fact that a random pairing of neurons does not display the same tuning does not mean these networks are not learning the same mechanisms. In fact, ED Fig 9a shows that Dale's law helps with the correlation, which is likely because having a weight constraint offers some kind of (weakly) unique identifier. Also, how were the signs decided for units in these networks?

Thank you for pointing out that this section was difficult to follow. We have rewritten this whole section [l. 313-350] to clarify the motivation and the results. We also clarify below:

When might connectome constrained and task-optimized DMN models accurately predict neural activity? Sparse connectivity is a hallmark of biological neural circuits, and in this section, we ask whether sparse connectivity enables DMN models to make accurate predictions of neural activity. For sparsely connected circuits---assuming the connectome is known---there are fewer synapse parameters left to estimate using task-optimization. We hypothesized that such networks might support fewer possible mechanisms by which to perform a given task, compared to more densely connected circuits, and so a task-optimized DMN model is more likely to find the true mechanism and accurately predict neural activity.

We addressed this hypothesis in simulation, by constructing feedforward artificial neural networks solving the classic MNIST handwritten digit classification tasks (Fig 6a). These networks had varying degrees of sparse connectivity, and random assignment of neurons as excitatory and inhibitory respecting Dale’s law (25 groundtruth networks for each sparsity level, Methods). We simulated the process of making connectomic measurements from these groundtruth networks, and used those measurements to build connectome-constrained task-optimized DMN simulations of each groundtruth network. Since there is still uncertainty about the degree to which connection strength can be inferred from noisy connectomic measurements of synapse count, we simulated two settings. First, that connectomic measurements reveal connectivity but not connection strength. Or second, that connectomic measurements reveal connectivity and additionally a *noisy* estimate of strength. We then asked how well each task-optimized DMN simulation predicted the neural activity of its corresponding groundtruth network, as a function of the sparseness of the connectivity.

When connectivity is assumed to be known but not connection strength, DMN simulations were connectome-constrained and task-optimized to estimate both the resting membrane potential of each neuron, as well as the connection strength of each pairs of neurons (provided they were connected in the in the corresponding groundtruth network, all other connections were kept at zero). When connectomic measurements can be assumed to also provide noisy estimates of strengths, task-optimization was used to **denoise the noisy estimates**: We used the noisy estimates as a prior on the strength of each connection, to regularize the task-optimized connection strength **towards the noisy measurement**.

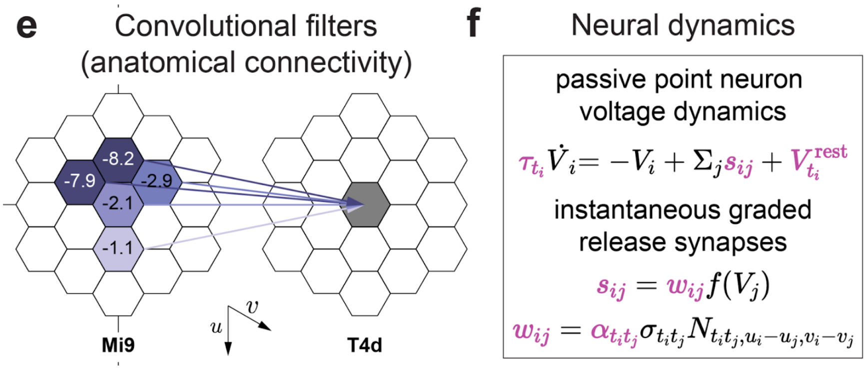
Consistent with our hypothesis, we found that when connection strengths cannot be inferred from connectomic measurements, sparsity in the connectome greatly improves the correlation of neural activity between a groundtruth network and its DMN simulation at the single neuron level (Fig 6b, median Pearson correlations of 0.85 for 10% connectivity vs 0.38 for 80% connectivity, 100 randomly selected neurons from 25 randomly generated groundtruth networks). Conversely, when noisy estimates of connection strengths can additionally be inferred from the connectome, we find that DMN predictions of neural activity correlate well independent of connection sparsity (median Pearson correlation >0.9 for all connectivities).

In our fly visual system model, we assumed an intermediate regime: Pairs of connected neurons of the same pre- and post-synaptic cell type likely express the same neurotransmitter and receptor combination leading to a common unitary synapse strength for all such connections. So our model assumes the synapse count reveals the relative magnitude of connection strength between all such connections. However, our model assumes that absolute connection strength is unknown from the connectome, since the strength of a unitary synapse might vary across cell types expressing different neurotransmitters and receptors. In other words, 5 synapses between Mi1 and T4 neurons could be stronger or weaker than 5 synapses between Mi9 and T4 neurons, but 5 synapses between Mi1 and T4 neurons is assumed to be exactly half as strong as 10 synapses between Mi1 and T4 neurons.

Finally, we agree with your point about the correspondence of neurons between groundtruth networks and their DMN simulations. In our modeling, we assumed that neurons can be uniquely identified and put in correspondence. Experimentally, morphology and gene expression genes are frequently used in addition to connectivity, in order to identify neurons uniquely. However, if we restrict ourselves to identifying and corresponding neurons solely based on connectivity, this is still possible, except in the extreme case of dense all-to-all 100% connectivity, since all neurons will have the same connectivity profiles. For simplicity, we now drop this last data point in our figure, which is not needed to show the overall trend that sparseness improves the accuracy of DMN predictions of neural activity.

Clarifications:

Can the authors better explain the differentiation between synapse count and the scalar? It seems the scalar and the count are held constant for all pairs of cells with the same pre- and post-synaptic cell type, so what extra information does having the count as part of the equation provide?



For each pair of cells i and j, the corresponding filter weight wij is determined by three factors: The synapse sign sigma_ti_tij and the scaling coefficient alpha_ti_tj which are indeed the same for each pair of pre- and post-synaptic cells. The third contribution comes from the synapse counts-- however, and importantly, the synapse counts do not only depend on the cell-types, but also on the relative spatial offsets of the two specific cells. This is denoted by the coordinate subscripts \Deltau=u_i - u_j, \Deltav = v_i - v_j which are the relative columnar offsets of a postsynaptic target cell i of type t_i and a presynaptic source cell j of type t_j in the two-dimensional hexagonal coordinate system.

Thus, the synapse counts determine the *shape* of the filter [as visualized in Figure 1e, copy-pasted above for easy reference], whereas the scalar scales the overall strength of the filter. In total, the model is based on 2355 (non-zero) synapse counts, and 604 (non-zero) connection-filters, and hence there are 604 scalars that are trained. We have updated the description of the model in Results [l. 81-101] to clarify this.

How do the 50 ensemble models differ? Just different draws from the same distribution of resting potentials?

The 50 models are initialized at a random location in the parameter space and also differ in the stochastic optimization. This includes the order by which samples are drawn from the Sintel dataset and their randomized augmentation (random flips, rotations, pixel-wise gaussian noise, random contrast and brightness). We have updated the description of the model in Methods [l. 491-493 and 580-588] to clarify this.

Referee #4 (Remarks to the Author):

Lappalainen et al. build an optic lobe connectome-constrained neural network called a task-optimized deep mechanistic network (DMN), and optimize for a computation performed by that biological circuit (motion detection). They show that constraining both the connectivity and computational task reproduces some of the experimentally determined tuning of specific neurons, and makes predictions about the tuning properties of other neurons in the network that have yet to be experimentally measured. Finally, the paper argues that, for sparse networks (such as some biological neural networks), knowledge of the connectivity, signs of connections, and an estimate of connection strength may be sufficient to predict the mechanism by which the circuit performs a known computational task.

I am in general enthusiastic about the study - it is a useful simulation of a portion of real, complex neural network (64 cell types and 721 columns, plus 1 inhibitory cell type that extends across columns, but missing all of the feedback and neuromodulatory connections) and shows how connectivity shapes many of the known properties of a neural network - the optic lobe is an ideal test case as its cell types have been studied extensively over the past 60 years. Figures 3 and 4 in particular are quite nice - 1) comparisons between the best performing model's T4/T5 cells and the response properties of their inputs to known tuning curves and responses (although I have concerns about some of the details - see below) and 2) comparisons between different models to understand what properties of particular cell types define the best performing models. This is a nice study that will form the basis for simulations of larger biological networks, for fly, and other species. However, I have several concerns about the modeling, model predictions, and interpretations of the results that should be addressed.

Thank you for enthusiastic support of our work, and for your constructive suggestions which have helped us to substantially strengthen our work. Based on your suggestions, we have now exhaustively detailed how we constructed the DMNs from connectomic data with new tables and descriptions [Supplementary Information, Supplementary Data: connectome_constructions.csv, connectome_construction_merge_fib19_fib25.json]. We have also provided extensive new analyses to characterize which aspects of the connectomic data (e.g., cell-type connectivity, synapse counts, synapse signs) are important for achieving a close match between model predictions and neural activity measurements [new main paper Fig. 2d, Extended Data Fig. 9], described in detail in the response to R2. In addition, we performed additional analyses of the L1-5 cell types. Additionally, we characterized the input asymmetries and their relationship to the predicted direction selectivity indices [Extended Data Fig. 2].

The authors build a hybrid optic lobe connectome from several different datasets - the choices they made in how to combine these datasets must be made transparent in the paper. Ideally, they would present a Supplemental Figure devoted to how this was done and how they handled any discrepancies or differences between the datasets. If there were no discrepancies or differences this should also be explained. It is difficult to interpret the findings from the model without a thorough understanding of how the connectome-constrained model was built.

Related to this point, I assume the signs of connections were taken from the literature? Can the authors provide citations for these (and can they compare against the same cell type in the open FlyWire/FAFB *whole-brain connectome dataset, for which neurotransmitter predictions (from Eckstein et al.) are available in the optic lobe)?*

Thank you for pointing out this omission. We have now included extensive tables documenting the source of the connectivity, signs, and neural activity measurements for each cell type in the Supplement. A description of the algorithm used to generate the filters (spatial connectivity between cell types) is also given in the Supplement.

Regarding the comparison to the neurotransmitter predictions from Eckstein et al.: the signs of our connectivity are not computational predictions but actually based on direct experimental measurements of the neurotransmitter and receptors expressed by each cell type, as given by Davis et al. 2020. We have now clarified this in the Results [l. 89-90] and Methods [l. 455-464].

Related to the point above, the authors perform a sort of normalization step with the data so that they can model every column identically (making sure synapse numbers are the same in each column) - this ignores heterogeneity across columns (that might be important for motion detection). Can the authors provide more detail on how this simplification deviates from the actual connectivity (how much heterogeneity is there across columns?) and show how this choice affects modeling results (if they incorporate some of the heterogeneity into the model, how do the results change)?

At present, the best investigation of synapse count variability across columns in the fly optic lobe comes from only a local analysis of 7 columns in Takemura 2015, showing a variability of synapse count of roughly 10% between neighboring columns. It is also known that the color and polarization pathways introduce larger-scale deviations in the presence or absence of neurons in a perfect hexagonal lattice beginning with the random organization of the pale and yellow photoreceptors and the dorsal rim specialization of polarization sensitive photoreceptors (Kind et al 2021). More comprehensive studies of variability across the entire optic lobe are a substantial undertaking and await the full completion and analysis of Flywire and Janelia optic lobe connectomes.

We agree that with these new datasets, it will be possible to explicitly study the importance of spatial/retinotopic deviations from the simplifying assumption of perfect homogeneity of connectivity across the connectome which was essential for the present study. It is worth pointing out that not only our model, but also the *experimental literature* characterizing the neural activity patterns of each cell type, which we use as validation, invariably averages across neurons. Thus both our model and experiment make the same assumption of spatial homogeneity.

The manuscript claims that the actual OL connectivity was critical. Extended Data Figure 3 shows that a task-optimized DMN outperforms a DMN with random parameters. But what is missing is a demonstration that the specific connectivity of the fly visual system is what enables optimal performance, rather than a generic neural network with the same level of sparseness and gross connectivity statistics of the biological network. How would a task-optimized artificially generated network (constrained by *biological connectivity statistics, rather than the exact connectivity of the fly visual system) perform compared to the task-optimized DMN and the random DMN?*

Thank you for this suggestion. In response to your comment and a similar issue raised by Reviewer 2, we performed a large set of experiments to characterize the importance of different aspects of the connectivity [cellular resolution connectivity, synapse counts, synapse signs] and how they relate to both task performance and accuracy of predicting neural activity, (for details also see our reply to referee 2 above) and added Fig 2d and Extended Data Fig 9 and to the text in Results [l. 171-191]. This analysis based on your comment substantially clarifies our understanding of the importance of connectomic constraints as well as task optimization.

In summary, we found that both task optimization and detailed connectomic measurements at the single neuron resolution were critical to prediction of preferred contrast of the 32 characterized cell types, and preferred direction of motion for the T4 and T5 subtypes, at the single neuron resolution (Fig 2d, Extended Data Fig 9). A DMN model ensemble with full connectomic constraints but no task optimization (randomized unknown single cell and synapse parameters) led to accurate predictions of preferred contrast, but poor predictions of direction selectivity and preferred direction. Task-optimized DMN ensembles which assumed the absence of full connectomic measurements of single cell resolution connectivity, synapse counts, or synapse signs, and therefore used task optimization to infer any of these parameters struggled to correctly predict the preferred direction of motion. However, accurate predictions of the direction selectivity --- but not preferred direction --- could be achieved with measurements of cell-connectivity without synapse counts (Extended Data Fig 9c). This demonstrates the importance of both detailed connectomic measurements and task optimization to achieve the best predictions of neural activity.

Regarding the high-level task performance of differently constrained models: consistent with general findings in deep learning, we find that task-training models with more parameters typically leads to better asymptotic performance.

In addition to the task-performance of the unconstrained CNN with 414,602 free parameters, we now show task-performances for the other DMN variations we describe in Figure 2d and in Extended Data Figure 9, showing e.g. that an unconstrained DMN (only constrained by cell-type connectivity) with 11,593 free parameters (+ 7427 decoder parameters) performs only marginally better than the DMN constrained by all known connectomic-constraints with only 734 free parameters (+ 7427 decoder parameters).

While gradient-based methods on unconstrained networks find better solutions in terms of task-performance, our comparison suggests that the full connectomic constraints of the optic lobe provide a good structural initialization for the high-level task and for predicting single-neuron tuning properties.

Neurons are modeled with leaky linear non-spiking voltage dynamics and as point-neurons with a single electrical compartment. The authors show abundant tuning/response data for each model cell type, but they should compare the detailed temporal dynamics and delays (critical for motion detection) of model responses to real recordings of these same neurons (if this is present somewhere in the supplement, apologies if I missed it).

Many optic lobe neurons have been recorded via Ca++ imaging, voltage imaging, or electrophysiology (e.g., compare with published responses in Behnia et al. Nature 2014 or Yang et al. Cell 2016).

The L1 and L2 neurons are categorized as "known OFF selective" (Fig. 2b), but this deviates from my reading of the literature - shouldn't they show similar responses to both on and off flashes?

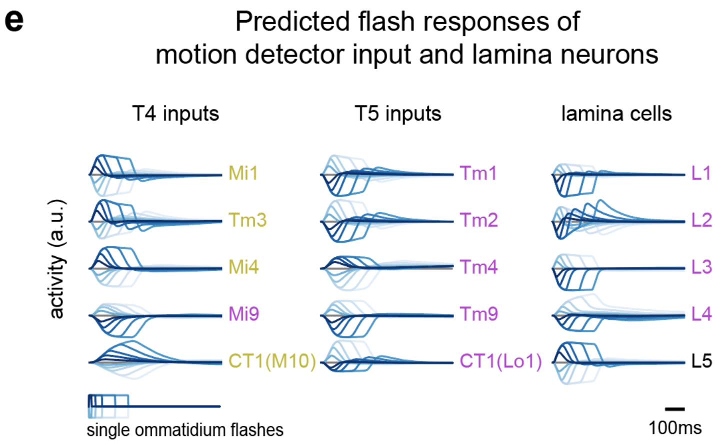
Also, the responses of L1 and L2 in Fig. 3e are shown as monophasic and producing an off response - shouldn't they be biphasic (in contrast with the responses for L3 and L4, which do look biphasic)? These potential mismatches between the literature and model results have me concerned that the model is not producing the expected responses for cell types that have been extensively studied (like L1 and L2).

Thank you for pointing out these details. We indeed agree that it is important to validate our model predictions for the function of the critical L1-L5 visual neurons.

First, we apologize for being unclear in our terminology. With "known OFF selective", we meant to say that cells of these cell types depolarize in response to a light-decrement (and hyperpolarize in response to a light-increment) which is in agreement with the literature [Reiff 2010, Clark 2011, Freifeld 2013, Strother 2014, Fisher 2015, Yang 2016, Drews 2020, Matulis 2020, Kettkar 2022]. We clarified our terminology in the manuscript [l. 142-143].

Regarding your second question: thank you very much for pointing out this fine-grained mismatch between the model presented in Fig. 4e (former Fig. 3e) and the literature. We found that other models in the ensemble correctly capture L1 and L2's biphasic tuning. You helped us realize the limitation of grounding our current Figure 4 only on the single model with the best performance on the optic flow task. Instead, now we updated Figure 4 to show average tunings from the best-task-performing cluster of models from the ensemble of 50 models for each of these cell types. We found that the average response of models from this task optimal cluster consistently provide better predictions.

We have included a new version of Figure 4e with additional L1-L5 tuning predictions from our model: L1 and L2 are indeed correctly predicted as biphasic on average and across a range of single-ommatidium flash durations of both ON and OFF contrast (see below).
We believe that this substantially strengthens the validity of our model predictions and our claims. We emphasize once more that our models are not constrained on any neural activity measurements (but only on the connectome and the optic flow task) -- thus while we do not expect that they capture all detailed tuning properties, they nevertheless yield remarkably accurate predictions for single-cell tuning.



Given the constraints provided by the connectome, there are only 734 free parameters in the model (resting membrane potential for each cell type (65) and unitary synapse strength (604)) - but the authors could have also included the synapse NL as a free parameter (varying across synapses) - how would this have affected modeling results?

Is there a reason they need to limit to ~700 free parameters? The question has to do with the choice of which parameters to fix across the model and which to vary - how much do these choices affect results?

In this work, we focused on building the simplest possible model consistent with measured connectomes, and evaluating to what extent it is sufficient to account for neural tuning. We agree that our framework would be the perfect testing ground to explore a larger class of mechanistic single neuron and synapse models and this would be exciting future work.

Our deliberate choice of the threshold-nonlinearity (ReLu) reduces the number of possible parameters because it is scale invariant. Rescaling the nonlinearity in this model is equivalent to rescaling the unitary synapse strength parameter \alpha, and shifting the nonlinearity corresponds to changing the resting membrane potential. For this reason we did not add additional parameters to the nonlinearity. But for other nonlinear functions, additional parameters could indeed be explored.

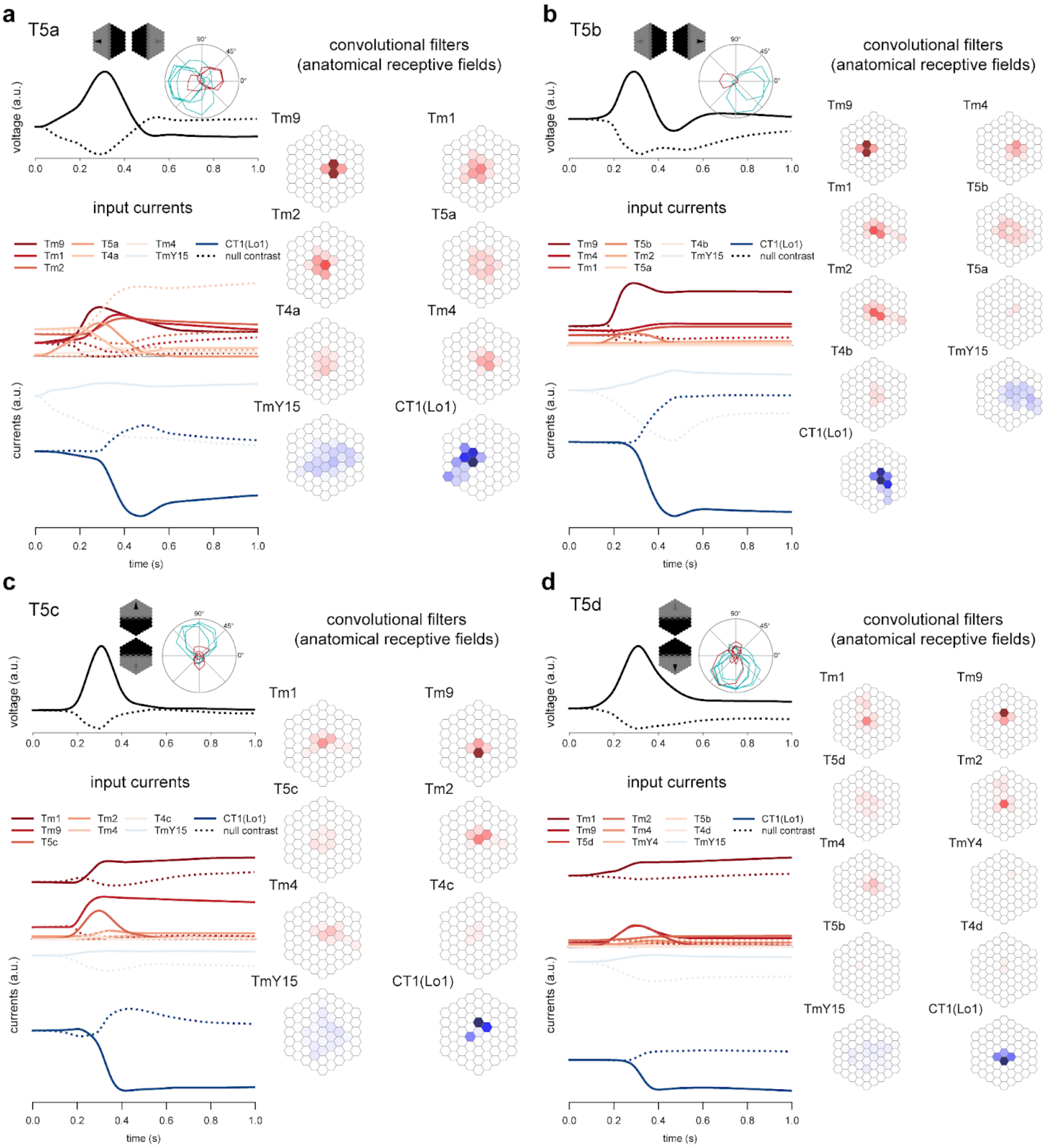
Relevant to this question, we also characterized the number of parameters in the models we described above, and in Fig 2d and Extended Data Fig. 9 to characterize the relationship between connectome-constraints and task-optimization — which includes models with different parametrization of synapses and their numbers of free parameters. We characterize in detail how much the choices of these constraints affect the results above. We do emphasize that our approach is not computationally limited to ~700 free parameters (and this is also shown in these additional analyses in which we optimize over much bigger parameter sets), but a modeling choice.

The authors optimize the model network to perform a particular motion vision task - this makes a lot of sense given that the major function of the optic lobe is to detect motion, but it is not clear how the results depend on this specific task - this needs to be addressed. What differences might they observe or expect if the network was optimized to perform a different task that the fly optic lobe mediates (for example, color or shape detection) or a specific fly behavioral task (like the optomotor response) - how would optimizing for a different task change the responses properties of neurons in the network? Addressing this question is critical for understanding the constraints of the connectome.

We focused on a local motion detection task since the connectome available to us largely focused on reconstructing the circuitry of local motion processing the optic lobe. We agree that an exciting future direction would be to investigate and reverse engineer the panel of tasks for which the circuitry of the optic lobe is optimally evolved over evolutionary time-scales. As you suggest, task optimization across a large variety of behaviorally relevant tasks will likely be critical for constructing accurate future models of the entire optic lobe, which must support all visually guided behavior.

Model responses to moving edges in Fig. 2c: Why do many of the models show that T5 is responsive to ON moving edges (which it is not)? Could this be expanded upon in Figure 4 (that compares different models)?

In our model ensemble, we find that T5 neurons are more often and more strongly predicted to be off-motion (p=0.0009). However, it is the case that across the entire ensemble, we do find substantial variability in the prediction for T5 on-motion selectivity. When we also filter the ensemble to only select models which exhibit the correct contrast-tuning [to flashes] *and* the correct preferred directions, then these models only predict very weak on-motion tuning for T5 cells (see details in the Figure below, which is replicated here again). Thus, while the ensemble as a whole is not inconsistent with the data, there is considerable variability. We have clarified this in the Results [l. 231-236].



I'm concerned about claims that the mechanism of motion detection matches experimental findings: The mechanism of how direction selectivity in T4 and T5 cells emerges is still debated. Haag et al. suggest there is both PD enhancement and ND suppression; Gruntman et al. argue for ND suppression only; Wienecke et al. argue for neither. It seems likely that the mechanism is different for the ON and OFF pathways. The manuscript shows that the tuning of T4 and T5 cells and their inputs (in some models) qualitatively matches the experimentally measured tuning, but can the authors clarify which mechanism of direction selectivity is matched/favored by their models?

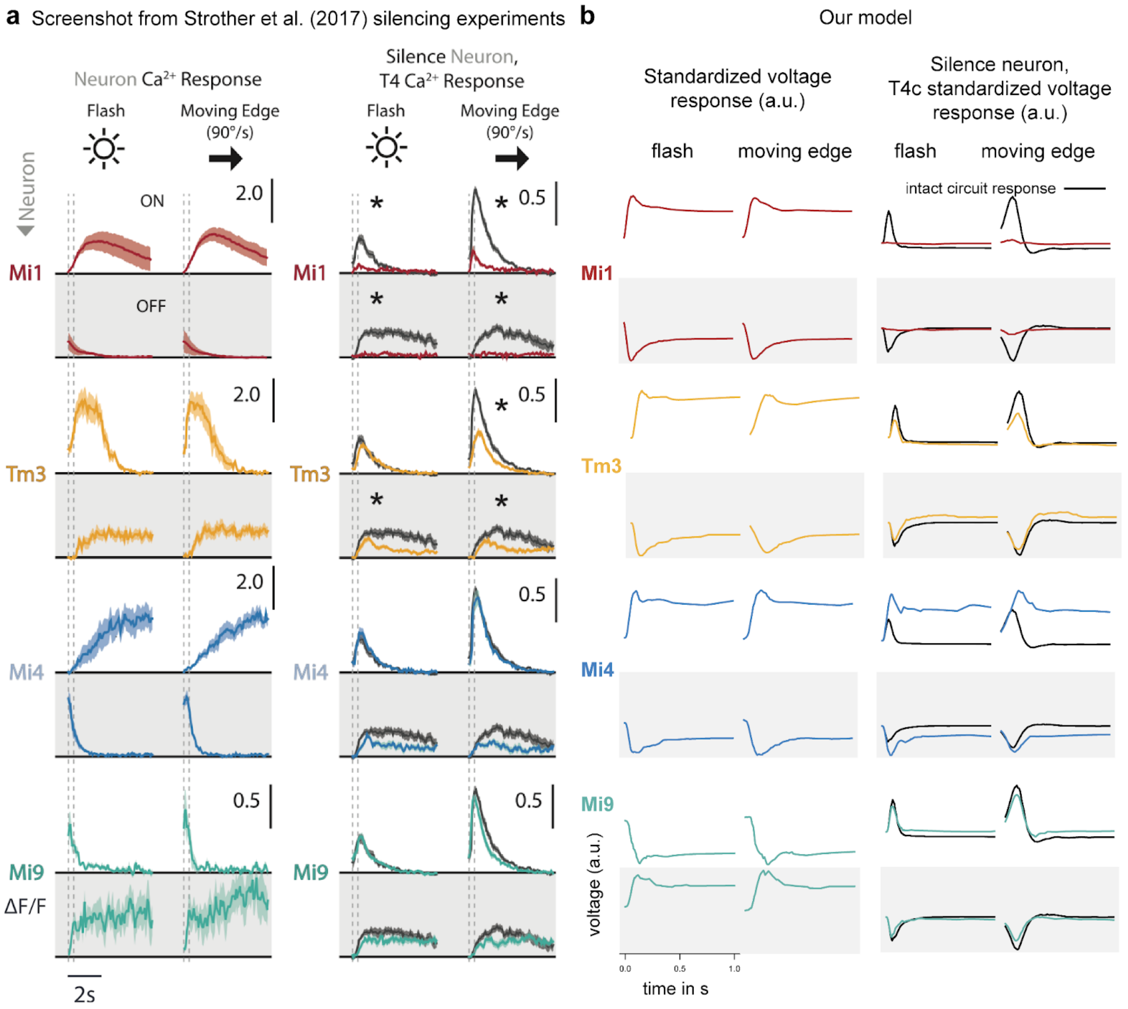
Thank you for this suggestion. We have further studied the circuit mechanism of direction selectivity in T4, T5, and TmY3 neurons through a combination of techniques and indeed predominantly find null direction suppression (Barlow-Levick) mechanisms for T4, T5, and TmY3 but also enhancement of coherent motion across the visual field through lateral interactions between T4 and T5 neurons of the same subtype.

In our mechanistic model, we can inspect the input current contributions from each cell type and study the differences in these contributions for motion stimuli in the preferred and null directions. This allows us to directly inspect the contribution of neurons to a computation without needing to perturb the circuit through inactivation. For T4 neurons, we observe null direction suppression mediated by inhibition from Mi4, and also surprisingly significant excitatory T4 to T4 input attributing a role for this connectivity for the first time. For T5 neurons, we observe null direction suppression mediated by CT1 inhibition and consistent excitatory input from neighboring T5 to T5 neurons and by excitatory input from Tm9. For TmY3, we see null direction suppression predominantly via Mi14 and Mi4 inhibition and excitation via Mi1 and L5.

We summarize these findings now in the main manuscript in the Results [l. 237-249] and in Fig 4. In addition, we show full results for all T4 and T5 subtypes and for TmY3 as Extended Data Figures.

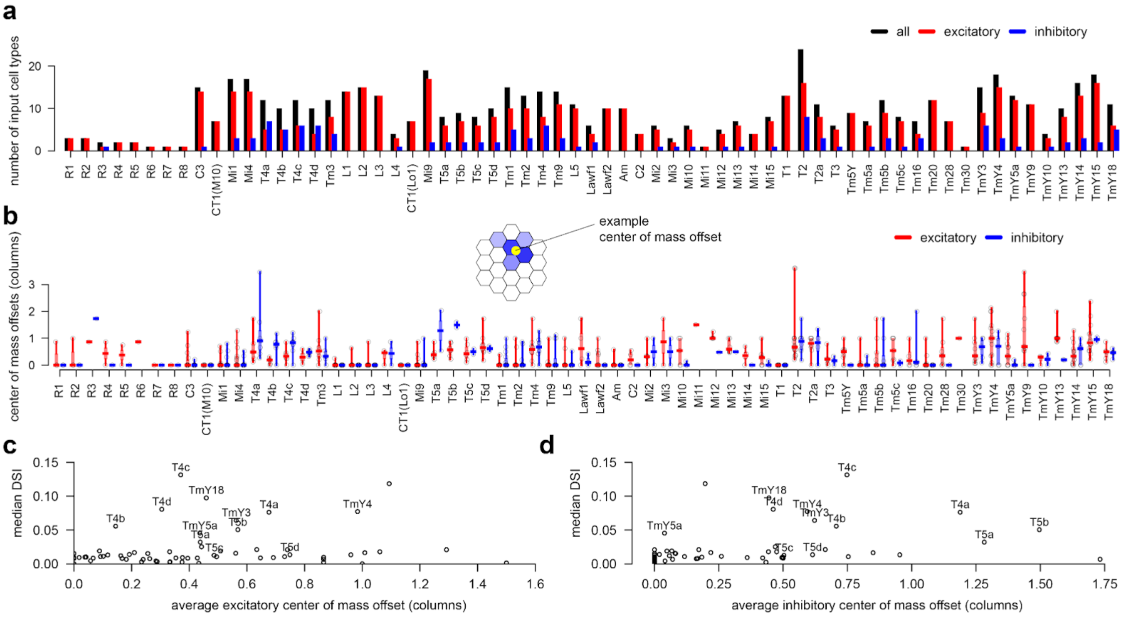
The authors have missed an opportunity to test their model through silencing experiments - inputs to T4 and T5 cells have been silenced experimentally (Strother et al. 2017; Serbe et al., 2016) - does silencing these neurons in the trained models recapitulate experimental findings/impair motion detection by the decoder network?

Thank you for this suggestion. In addition to the analyses of input current contributions, we now also simulated silencing experiments, and compared the model predictions with experimental results reported in Strother et al, Neuron, 2017, finding good agreement for most effects, as shown below. For these results, we averaged model-predictions over all models in the model-cluster with correct T4c tuning (which, as explained above, is also the cluster with the best task-performing model). The two column in panel a below are direct copies of Strother et al 2017 Figures 3b [left column] and 3c [right column], panel b shows, for comparison, the output of our analyses. We emphasize that, for these analysis, only a qualitative comparison of the effect is meaningful-- for example, Strother et al report Delta F/F from calcium imaging, whereas we show standardized voltage responses. In addition, our ‘silencing’ analyses are based on simply clamping the respective T4c inputs to 0, which is likely a crude approximation of the effect of blocking synaptic transmission with shibire(ts1) and temperature increases. Nevertheless, the models correctly capture that: i) Removal of Mi1 excitation removes the T4c response ii) Tm3: Removal of Tm3 excitation decreases, but does not remove, the T4c response. iii) Mi4: Removal of Mi4 does not remove the T4c response [it does, however, lead to a prolonged response in the network model which is not observed experimentally] iv). Mi9: no effect on T4c.



Of the 19 cell types with asymmetric inputs, only 12 are predicted to be motion selective - can the authors comment on differences between the inputs of these 12 and the other 7?

Thanks for this excellent suggestion. In response to your comment, we have now characterized the asymmetries of the inputs and in comparison to the different direction selectivities that our model predicts for these cells to understand the relationship (Fig. below). We find no trivial relationship between the asymmetry of the inputs and the predicted direction selectivity. We have included this new analysis as an Extended Data Figure and summarized in the Results.



One of the most exciting findings is the prediction that TmY3 could possess direction selectivity independent of the T4/T5 pathway. This is a bold claim - that the model can be used to identify new motion sensors in the fly visual system that has been studied for more than 60 years. It seems reasonable (and feasible given that the authors' local collaborators) to ask the authors to validate this prediction with experimental data - whereas elsewhere the authors rely on published experimental data for comparisons, this prediction would require new experiments.

This study has been a substantial undertaking, conducted by computational labs enabled by the availability of this rich and large dataset. This is a new era for computational neuroscience, one in which an enormous number of detailed, comprehensive, and experimentally testable predictions can be made through detailed and comprehensive modeling.

As we have responded earlier, we wish to emphasize that no neural activity measurements were used in constructing our model, and we already tested the predictions of our model against experimental measurements of neural activity across 26 studies. This is already a highly nontrivial validation of our model. Further, our model makes hundreds of pages of experimentally testable predictions (see supplement), of which this is but one prediction. We do not believe that just one more experimental validation of just one of these hundreds of predictions is necessary to substantiate the main claims of our paper.

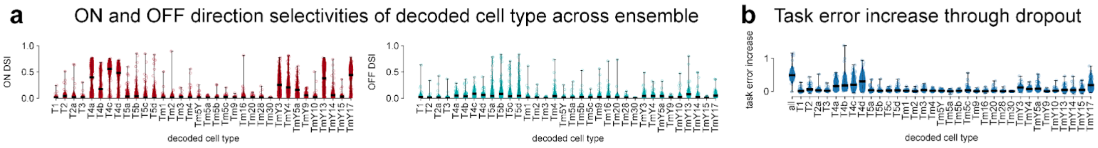
Nonetheless, we are informed by our colleague Michael Reiser (HHMI Janelia), that unpublished work in progress in his lab does indeed hint at motion selectivity for TmY3. He says the whole picture for this cell type is more interesting and complex and they plan to report the results of their own study in the near future.

Further, the authors could discuss if TmY3 is expected to function like an HR detector or BL detector, or is direction selectivity predicted to arise via a different mechanism?

As described in detail above, for TmY3, we see null direction suppression predominantly via Mi14 and Mi4 inhibition and excitation via Mi1 and L5.

I would like to see some of the statistics of the decoder network. The manuscript argues that it cannot compute motion itself, but is this true? Does the decoder only depend on the output of direction-tuned neurons? Does it perform equally well when the weights for non-direction tuned neurons are forced to zero? What is the minimal set of T and Tm cells required to detect optic flow?

Thanks for the suggestion. We now analyzed the relevance of each individual decoded cell type for the optic flow task. We evaluated the increase in task-error when we replaced responses of individual cells from a cell type by their spatio-temporal mean. We find that removing T4 subtypes strongly increases the task error. The decoder therefore most strongly relies on T4 cells which are predicted as on-motion detectors, but typically draws information from each decoded cell type to encode the synthetic flow field. We have included this figure below as an Extended Data Figure.



Minor comments

Figure 1 depicts a male fly, though much of the connectome data used, I believe, comes from female flies.

We updated figure 1.

Winding et al. 2023 larval connectome paper should also be cited at line 7.

Thank you, we have fixed this, and additionally also cited the newly published Flywire pre-prints.

The meaning of lines 99-101 is not clear. The text argues that the DMN was optimized for a computational task performed by the real biological network (i.e. fly visual system), but citations 44 & 76 refer to mammalian cortex.

Thank you, fixed. The citations were meant for the general idea of “task optimization”.

Line 119: typo in “backpropagation”

Thank you, fixed.

Line 105 says that motion detection is a "challenging" computation. "Challenging" is too subjective a term and can be removed. Arguably, theoretical models for motion detection proposed in the 1960's by Hassenstein & Reichert and Barlow & Levick are relatively simple.

Thank you, fixed.

Lines 283-285 claim that networks with different sparsity must use different computations. Why must this be true?

We have no mathematical proof that this statement is true, so we have removed this claim.

Lines 333-336 claim that "DMN models generate meaningful predictions in absence of neural activity measurements… (Fig 4)" feels too strong, since knowing actual tuning of Mi9 was critical in determining "correct" tuning of T4.

The tunings of T4 and Mi9 are correctly predicted by the cluster with the best task performance. While our DMN ensemble generates multiple hypotheses, task performance can be used to rank these hypotheses, and as we show in Extended Data Fig. 11, DMN models with better task performance are also more likely to correctly predict neural activity. Even if this were not the case, our connectome-constrained and task-optimized DMN framework is able to generate just a small number of reasonable hypotheses for the neural activity of each cell type, without the use of any neural activity. In the absence of our method and the connectome, this hypothesis space would be extremely large since any neuron could potentially take on any visual tuning. We feel that we have proven that our model ensemble generates "meaningful predictions in the absence of neural activity", even if multiple hypotheses are generated.

The limitations of the approach (e.g. simplistic modeling of neural dynamics, and lack of electrical synapses, neuromodulation and glia) are introduced at the outset. It would be nice if these limitations were further elaborated/explored in the discussion.

We have now expanded discussion of this issue.

Reviewer Reports on the First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

The authors have addressed most of my comments but one issue remains open. This concerns the question how incomplete knowledge of the connectome or the catalog of cell types would affect model performance. The importance of cell type information is of particular interest. Cell type information in the Drosophila visual system is more detailed and more complete than in almost all other circuits and species. How would model performance be affected if not all cell types were known, or if similar cell types were indistinguishable? This could be simulated by pooling some cell types, or by merging similar cell types. This knowledge would be valuable to assess the requirements to generalize the DMN approach because cell type information in other systems is usually incomplete or inaccurate. I recommend to the authors to address this issue, assuming that it can be addressed by a modest amount of additional work.

Referee #2 (Remarks to the Author):

I thank the authors for their substantial work revising the paper. My concerns have been sufficiently addressed and I think the new analyses determining which elements of the model and training process are crucial for replicating specific response properties is insightful.

Referee #4 (Remarks to the Author):

Many of our comments and concerns have been adequately addressed by the inclusion of new data and analyses - these include:

Additional supplemental data to document how the connectome was constructed from disparate datasets.
Figure 2D and extended data figure 9 are a welcome addition to the manuscript. These show that constraints from the connectome along with task optimization are required to accurately predict direction selectivity, but that connectomic constraints alone are sufficient to predict ON/OFF selectivity of cell types.
Updated analysis in Figure 4e, taking into account the best-performing cluster of models, captures the experimentally measured tuning of lamina cells.
Extended data figure 2 addresses differences in asymmetric input to motion selective and non-motion selective cell types, finding no strong correlation.

Some of our comments and concerns have been addressed, but have raised further questions:

Figure 4 and extended data figures uncovers a new potential mechanism for how direction selectivity arises in T4 and T5 cells, as well as the hypothesized direction selectivity of TmY3. The

DMN suggests that ND suppression is involved in computing motion in all three cell types. In addition, lateral interactions between T4 and T5 neurons of the same subtype is suggested as a novel mechanism for PD enhancement. This result seems to contradict experimental findings showing that PD enhancement contributes significantly to direction selectivity in T4 and T5 cells (e.g. Fisher et al., 2015; Haag et al., 2016 & 2017; Arenz et al., 2017). This discrepancy should be discussed.

In silico silencing of inputs to T4c qualitatively recapitulate experimental findings from Strother et al (2017). The authors say that for these simulations they “clamped the respective T4c inputs to zero”. It is not clear if only the inputs to T4c were clamped to zero, or all outputs of the silenced presynaptic cell type. The latter is a better approximation of the effect of a shibire(ts) experiment, which silences all synaptic transmission, not only the synaptic transmission to the downstream neurons of interest.

Extended data figure 3 is a welcome addition which shows some properties of the decoder network. However, this has raised some further concerns. These data show that the decoder network predominantly uses the output of T4 to compute motion. Incidentally, the DMN most accurately predicts the experimentally measured tuning of T4 (compared to T5). Could it be that the mechanism of direction selectivity in T4 is “simpler” or more directly constrained by the connectome than, e.g., direction selectivity in T5 (or TmY3)? Therefore, the DMN and decoder network learn this first, while direction selectivity in T5 is learned more sloppily (explaining the aberrant ON responsiveness in T5 even in the best performing models)? This seems to be a limitation of the whole approach, as it suggests that there are certain computations that the DMN is particularly good at learning, or certain computations are more constrained by the connectome than others (this also relates to our point below - how much of this is dependent on the specific task chosen?). If T4 is excluded from the decoder, is the network better at learning the experimentally measured tuning of T5? Does it affect the predicted tuning of TmY3?

One major concern is not addressed:

The authors explain that they chose local motion detection to optimize their model due to it being an important and well-studied computation of the optic lobe. However, given the results presented in extended data figure 3, it seems that training on one specific task allows the model to learn the tuning of one cell type to the detriment of others. We still suggest that it would be useful to readers (and within scope) to provide analysis to address how choice of decoder or task optimization affects the learned tuning of neurons in the network. The authors make strong claims that connectivity alone can be used to predict the function of individual neurons, but remains open as to how much this depends on the specific task used.

Minor:

“Whole brain connectome projects have just been completed for the larval and adult fruit fly” - larval connectome is not whole brain exactly (missing the SEZ) and adult whole brain connectome should point to refs 11, 16, and 17.

The authors should reference the new publication https://www.biorxiv.org/content/10.1101/2023.10.12.562119v2.full.pdf - this paper contains connectivity of all neurons of one optic lobe (along with cell typing). The authors work here preceded the completion of the Matsliah et al. wiring diagram, but it would be useful for the authors to note that going forward full connectivity could be used (rather than constructing an RNN from disparate datasets), and that there are important differences in cell type classifications between Matsliah et al. and this paper (for example, Tm5 and TmY neurons).

Relative to the points above, there are 148 references in this paper! Given that almost 100 references will need to be cut for publication (and choices will need to be made), the authors should do this trimming for any further revision.

Connectome is used rather liberally throughout the manuscript (for example, in this sentence in the Discussion “Knowledge of the connectome played a critical role in this success, in part by leading to a massive reduction in the number of free model parameters.”) - however, the authors have used wiring diagrams of portions of the optic lobe and generated an RNN from these constraints. Since connectome, to my knowledge, implies completeness (one would not use the term genome to refer to the sequencing of a subset of genes…), it would be less confusing for the authors to distinguish their data sources from newer more complete wiring diagrams (suggestions: refer to the data sources as partial connectomes or columnar connectomes or wiring diagrams).

Author Rebuttals to First Revision:

We thank you and the reviewers for their helpful comments. In this revision, we have performed new experiments and analysis to address three main questions:

1. The importance of having detailed cell type annotations. With new experiments, we show the results of coarse-graining our knowledge of the cell type of each neuron.
2. The mechanism of direction selectivity in T4, T5, and TmY3. With new analysis, we shed light on whether we see signatures of preferred direction enhancement (no) or null direction suppression (yes) and relate this to the literature.
3. The role of the computational task in task optimization. With new experiments, we show the importance of choosing ethologically relevant tasks for task optimization.

We believe that these additional experiments and analyses strengthen our main claim, that with (appropriate) task optimization, one can use connectomic measurements to construct remarkably predictive mechanistic simulations of the fruit fly visual system at single cell resolution.

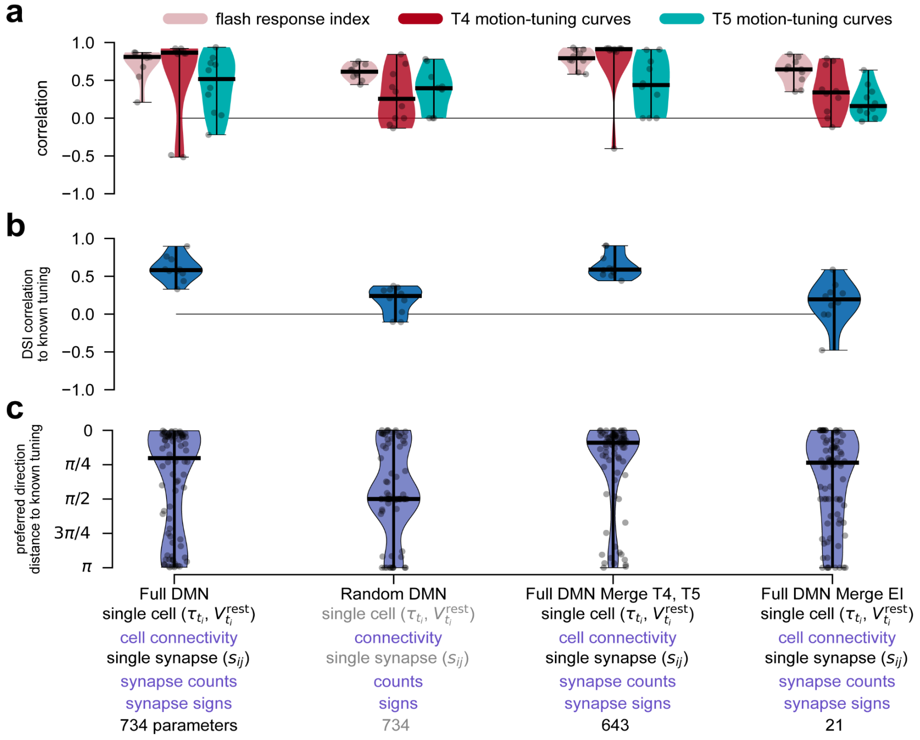
A detailed response to the reviewers is attached.

Referee #1 (Remarks to the Author):

The authors have addressed most of my comments but one issue remains open. This concerns the question how incomplete knowledge of the connectome or the catalog of cell types would affect model performance. The importance of cell type information is of particular interest. Cell type information in the Drosophila visual system is more detailed and more complete than in almost all other circuits and species. How would model performance be affected if not all cell types were known, or if similar cell types were indistinguishable? This could be simulated by pooling some cell types, or by merging similar cell types. This knowledge would be valuable to assess the requirements to generalize the DMN approach because cell type information in other systems is usually incomplete or inaccurate. I recommend to the authors to address this issue, assuming that it can be addressed by a modest amount of additional work.

We agree that an incomplete knowledge of the connectome could impact the usefulness of connectomic constraints. We thank you for the suggestion of pooling cell types. Based on this suggestion, we performed two experiments training new models with pooled cell types. We artificially assumed some cell types to be indistinguishable, and thus modeled them with shared single neuron and synapse parameters (resting potentials, time constants, and unitary synapse strengths). The figure below summarizes these results and is now added as Extended Data Figure 15 of the revised manuscript.

1. **Full DMN Merge T4, T5**: We assumed that the four 'T4' subtypes were indistinguishable from each other, and also that the four 'T5' subtypes were indistinguishable, pooling them into one T4 type and one T5 type. This reduced the number of cell types to 58. We found that this model accurately predicts neural tuning, presumably because these highly similar subtypes share the same neurotransmitters, receptors, and biophysical properties.
2. **Full DMN Merge E/I:** We assumed that we had only three cell types, 'excitatory' (merging 37 cell types), 'inhibitory' (merging 22 cell types) or 'mixed' (merging 4 cell types), based on our knowledge of expressed presynaptic neurotransmitters and postsynaptic receptor types. This model performs similarly to the random DMN model, poorly predicting tuning curves and direction selectivity index.



Referee #2 (Remarks to the Author):

I thank the authors for their substantial work revising the paper. My concerns have been sufficiently addressed and I think the new analyses determining which elements of the model and training process are crucial for replicating specific response properties is insightful.

Referee #4 (Remarks to the Author):

Many of our comments and concerns have been adequately addressed by the inclusion of new data and analyses - these include:

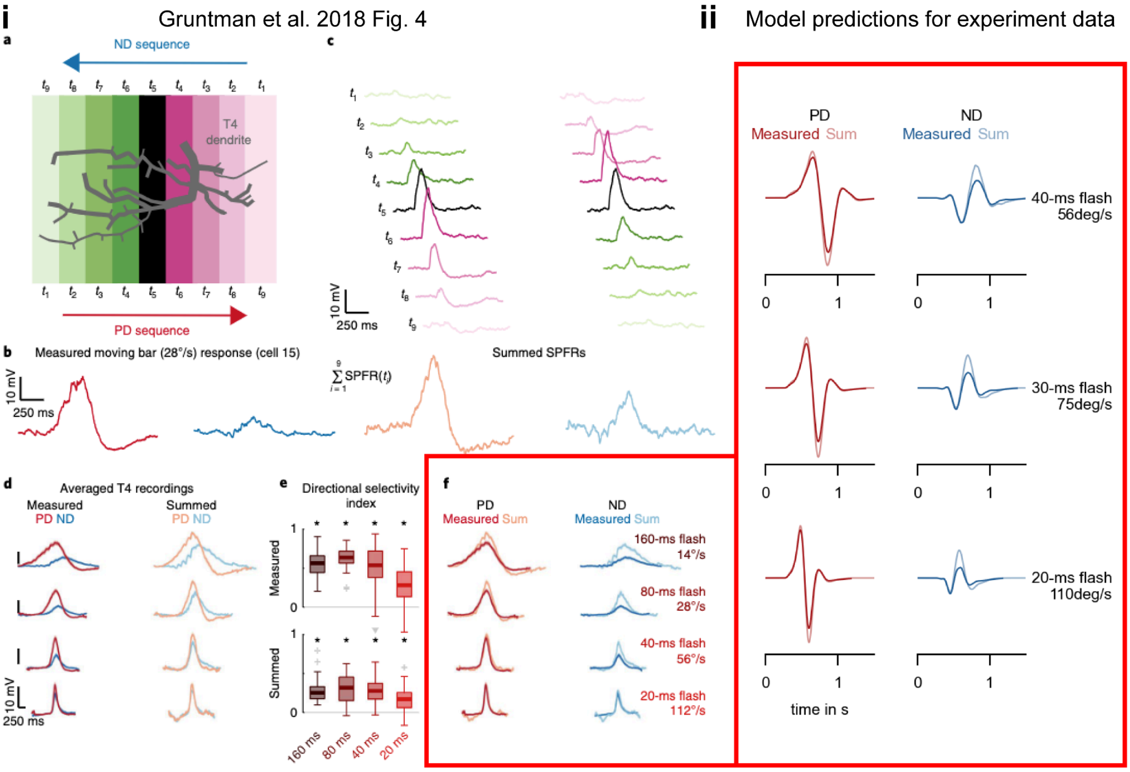
Additional supplemental data to document how the connectome was constructed from disparate datasets.
Figure 2D and extended data figure 9 are a welcome addition to the manuscript. These show that constraints from the connectome along with task optimization are required to accurately predict direction selectivity, but that connectomic constraints alone are sufficient to predict ON/OFF selectivity of cell types.
Updated analysis in Figure 4e, taking into account the best-performing cluster of models, captures the experimentally measured tuning of lamina cells.
Extended data figure 2 addresses differences in asymmetric input to motion selective and non-motion selective cell types, finding no strong correlation.

Some of our comments and concerns have been addressed, but have raised further questions:

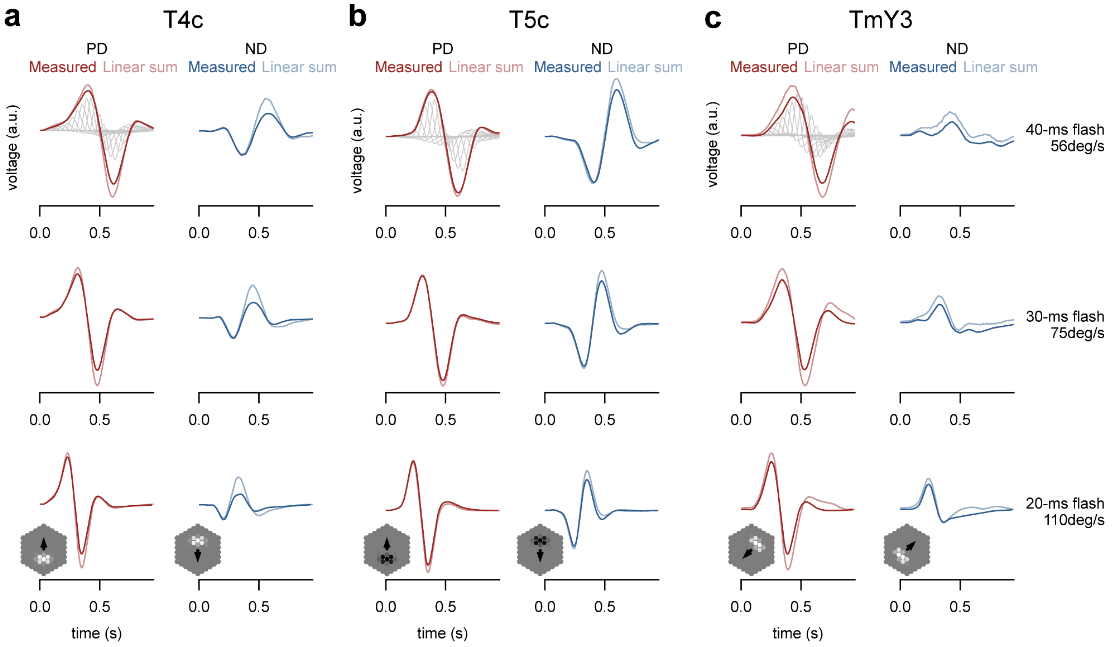
Figure 4 and extended data figures uncovers a new potential mechanism for how direction selectivity arises in T4 and T5 cells, as well as the hypothesized direction selectivity of TmY3. The DMN suggests that ND suppression is involved in computing motion in all three cell types. In addition, lateral interactions between T4 and T5 neurons of the same subtype is suggested as a novel mechanism for PD enhancement. This result seems to contradict experimental findings showing that PD enhancement contributes significantly to direction selectivity in T4 and T5 cells (e.g. Fisher et al., 2015; Haag et al., 2016 & 2017; Arenz et al., 2017). This discrepancy should be discussed.

Thank you for this comment. We now realize that our wording was potentially misleading due to the loaded meaning of the words “enhancement” and “suppression” in the motion detection subfield. These terms are used to refer to the *nonlinear* enhancement or suppression of the neural response to motion stimuli in the preferred or null direction relative to a “*linear* sum” null model. We have now reworded these sentences in our paper to make clear our original intent, which was to comment on a different effect.

Further, we now directly address the question of preferred direction enhancement (PDE) and null direction suppression (NDS) relative to a “linear sum” model by performing a new analysis which exactly replicates the voltage recordings and analysis of Fig 4 of Gruntman 2018 (replicated below). In our models, for all three cell types T4c, T5c, and TmY3, we only observe NDS and no (nonlinear) PDE. These results for T4 and T5 are in excellent agreement with Gruntman 2018 and Gruntman 2019. We should note that there is considerable experimental disagreement on this topic, largely due to differences in experimental methodology (calcium imaging vs voltage recordings). Therefore we emphasize that the predictions of voltage responses in our model are in excellent agreement with measurements of voltage responses by Gruntman 2018 and Gruntman 2019.



We have added these new results to the revised manuscript as Extended Data Fig 14.



In silico silencing of inputs to T4c qualitatively recapitulate experimental findings from Strother et al (2017). The authors say that for these simulations they "clamped the respective T4c inputs to zero". It is not clear if only the inputs to T4c were clamped to zero, or all outputs of the silenced presynaptic cell type. The latter is a better approximation of the effect of a shibire(ts) experiment, which silences all synaptic transmission, not only the synaptic transmission to the downstream neurons of interest.

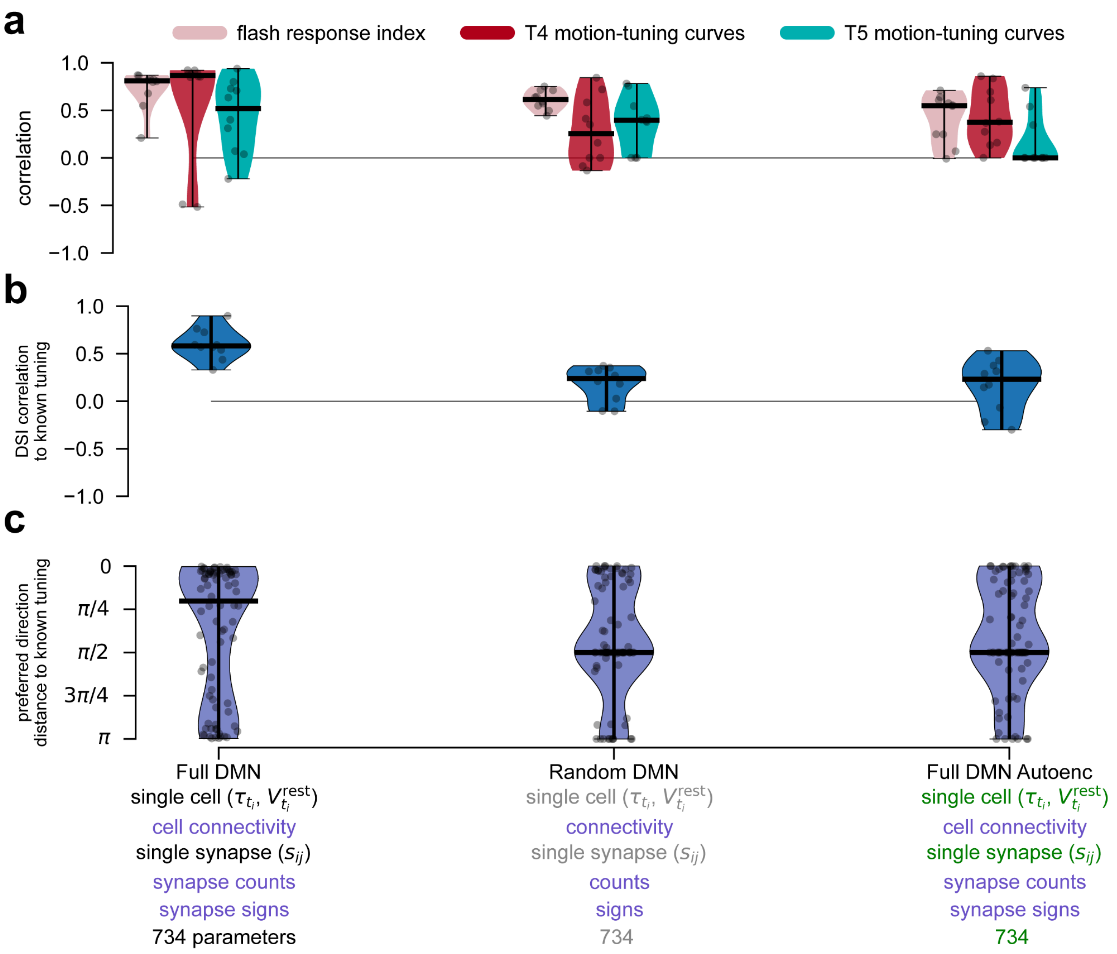
Thank you for requesting this clarification. Indeed we clamped all outputs of the silenced presynaptic cell type to zero, exactly as expected of a shibire(ts) silencing experiment. We have clarified this.

Extended data figure 3 is a welcome addition which shows some properties of the decoder network. However, this has raised some further concerns. These data show that the decoder network predominantly uses the output of T4 to compute motion. Incidentally, the DMN most accurately predicts the experimentally measured tuning of T4 (compared to T5). Could it be that the mechanism of direction selectivity in T4 is "simpler" or more directly constrained by the connectome than, e.g., direction selectivity in T5 (or TmY3)? Therefore, the DMN and decoder network learn this first, while direction selectivity in T5 is learned more sloppily (explaining the aberrant ON responsiveness in T5 even in the best performing models)? This seems to be a limitation of the whole approach, as it suggests that there are certain computations that the DMN is particularly good at learning, or certain computations are more constrained by the connectome than others (this also relates to our point below - how much of this is dependent on the specific task chosen?). If T4 is excluded from the decoder, is the network better at learning the experimentally measured tuning of T5? Does it affect the predicted tuning of TmY3?

One major concern is not addressed:
The authors explain that they chose local motion detection to optimize their model due to it being an important and well-studied computation of the optic lobe. However, given the results presented in extended data figure 3, it seems that training on one specific task allows the model to learn the tuning of one cell type to the detriment of others. We still suggest that it would be useful to readers (and within scope) to provide analysis to address how choice of decoder or task optimization affects the learned tuning of neurons in the network. The authors make strong claims that connectivity alone can be used to predict the function of individual neurons, but remains open as to how much this depends on the specific task used.

We believe that we have convincingly demonstrated our central claim: that task optimization (with an appropriate task) leads to significantly more accurate predictions of neural activity, than without task optimization (random parameters). We are not convinced that an expansive exploration of the choice of task and decoder architectures will further strengthen this claim.

Nevertheless, we have now trained an additional DMN variation with a new autoencoding task (predicting the input from the input) to address this concern. We do not expect this to be an ethologically relevant task, and do not believe that the fly visual system preserves all the information from the visual inputs, but rather computes behaviorally relevant features in a lossy manner. Further, this task does not force the visual system to perform temporal computations on visual inputs in the same way as the optic flow task. Consistent with this, we find that the DMN trained to perform autoencoding does not predict neural tuning more accurately than the DMN with random parameters. This supports our claim that it is important to choose a behaviorally relevant task for task optimization.



Minor:
"Whole brain connectome projects have just been completed for the larval and adult fruit fly" - larval connectome is not whole brain exactly (missing the SEZ) and adult whole brain connectome should point to refs 11, 16, and 17.

The authors should reference the new publication https://www.biorxiv.org/content/10.1101/2023.10.12.562119v2.full.pdf - this paper contains connectivity of all neurons of one optic lobe (along with cell typing). The authors work here preceded the completion of the Matsliah et al. wiring diagram, but it would be useful for the authors to note that going forward full connectivity could be used (rather than constructing an RNN from disparate datasets), and that there are important differences in cell type classifications between Matsliah et al. and this paper (for example, Tm5 and TmY neurons).

Relative to the points above, there are 148 references in this paper! Given that almost 100 references will need to be cut for publication (and choices will need to be made), the authors should do this trimming for any further revision.

Connectome is used rather liberally throughout the manuscript (for example, in this sentence in the Discussion "Knowledge of the connectome played a critical role in this success, in part by leading to a massive reduction in the number of free model parameters.") - however, the authors have used wiring diagrams of portions of the optic lobe and generated an RNN from these constraints. Since connectome, to my knowledge, implies completeness (one would not use the term genome to refer to the sequencing of a subset of genes…), it would be less confusing for the authors to distinguish their data sources from newer more complete wiring diagrams (suggestions: refer to the data sources as partial connectomes or columnar connectomes or wiring diagrams).

Thank you. We have addressed these minor comments. Regarding the larval connectome, we note that Winding et al. 2023 claim to have mapped the "*complete CNS*", writing on page 10: *"Our EM volume contains the complete CNS (brain, SEZ, and nerve cord), allowing us to assess communication between the brain and the rest of the CNS."*

Reviewer Reports on the Second Revision:

Referees' comments:

Referee #1:
Remarks to the Author:
The authors have addressed my concerns. I have no further points. Congratulations!

Referee #4:
Remarks to the Author:
The majority of minor comments have been addressed and terms "enhancement" and "suppression" have been sufficiently clarified in the text. However, the reference list is still too long.

The addition of Extended Data Fig 14 and comparison to Gruntman et al. 2018 sufficiently addresses our questions about the mechanism of direction selectivity in T4, T5 and TmY3 in the model.

Training the model on an autoencoding task addresses our concern about task optimization, and does demonstrate that task optimization is required to predict known tuning of direction selective cell types.

One comment is not yet addressed: the authors should explain (in the Discussion) why the DMN seems to be better at predicting direction selectivity in T4 over T5. Could this be a limitation of the model in the kinds of computations it can learn?

Lokhorst, G.-J. (2005). Descartes and the pineal gland.

Descartes, R. (1948). The passions of the soul, 1650. In *Readings in the history of psychology* (pp. 25–31). Appleton-Century-Crofts.

Churchland, P. S. (1986, Apr.). *Neurophilosophy: Toward a unified science of the mind-brain*. The MIT Press.

Bunge, M. (1977). Emergence and the mind. *Neuroscience*, *2*(4), 501.

Sperry, R. W. (1980). Mind-brain interaction: Mentalism, yes; dualism, no. *Neuroscience*, *5*(2), 195.

Tononi, G. (2012). Integrated information theory of consciousness: An updated account. *Arch. Ital. Biol.*, *150*(2-3), 56.

Cajal, S. R., & Sánchez, D. B. (1915). Contribución al conocimiento de los centros nerviosos de los insectos. *Trab. Lab. Invest. Biol. Univ. Madrid*.

von Helmholtz, H. (1850). Vorläufiger bericht über die fortpflanzungsgeschwindigkeit der nervenreizung. *Archiv für Anatomie, Physiologie und wissenschaftliche Medicin*, 71.

Bernstein, J. (1902). Untersuchungen zur thermodynamik der bioelektrischen ströme. *Pflugers Arch.*, *92*(10), 521.

Sherrington, C. S. (2023, Apr.). The integrative action of the nervous system. In *Scientific and medical knowledge production, 1796-1918* (pp. 217–253). Routledge.

Loewi, O. (1945). Chemical transmission of nerve impulses. *American Scientist*, *33*(3), 159.

Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S. A., Hudspeth, A. J., & Mack, S. (2000). *Principles of neural science* (4th). McGraw-Hill.

Ache, J. M., Polsky, J., Alghailani, S., Parekh, R., Breads, P., Peek, M. Y., Bock, D. D., von Reyn, C. R., & Card, G. M. (2019). Neural basis for looming size and velocity encoding in the drosophila giant fiber escape pathway. *Curr. Biol.*, *29*(6), 1073.

Bargmann, C. I., & Marder, E. (2013). From the connectome to brain function. *Nat. Methods*, *10*(6), 483.

Scheffer, L. K., & Meinertzhagen, I. A. (2021). A connectome is not enough - what is still needed to understand the brain of drosophila? *J. Exp. Biol.*, *224*(21), jeb242740.

Lin, S., Senapati, B., & Tsao, C.-H. (2019). Neural basis of hunger-driven behaviour in drosophila. *Open Biol.*, *9*(3), 180259.

Currier, T. A., Pang, M. M., & Clandinin, T. R. (2023). Visual processing in the fly, from photoreceptors to behavior. *Genetics*, *224*(2).

Clemens, J., Coen, P., Roemschied, F. A., Pereira, T. D., Mazumder, D., Aldarondo, D. E., Pacheco, D. A., & Murthy, M. (2018). Discovery of a new song mode in drosophila reveals hidden structure in the sensory and neural drivers of behavior. *Curr. Biol.*, *28*(15), 2400.

Borst, A., & Helmstaedter, M. (2015). Common circuit design in fly and mammalian motion vision. *Nat. Neurosci.*, *18*(8), 1067.

Tanaka, R., & Portugues, R. (2025). On analogies in vertebrate and insect visual systems. *Nat. Rev. Neurosci.*, 1.

Yamamoto, S., Jaiswal, M., Charng, W.-L., Gambin, T., Karaca, E., Mirzaa, G., Wiszniewski, W., Sandoval, H., Haelterman, N. A., Xiong, B., Zhang, K., Bayat, V., David, G., Li, T., Chen, K., Gala, U., Harel, T., Pehlivan, D., Penney, S., . . . Bellen, H. J. (2014). A drosophila genetic resource of mutants to study mechanisms underlying human genetic diseases. *Cell*, *159*(1), 200.

Basu, J., & Nagel, K. (2024). Neural circuits for goal-directed navigation across species. *Trends Neurosci.*, *47*(11), 904.

Thomson, J. A. (1925). *Concerning evolution*. Yale University Press.

Nern, A., Loesche, F., Takemura, S.-Y., Burnett, L. E., Dreher, M., Gruntman, E., Hoeller, J., Huang, G. B., Januszewski, M., Klapoetke, N. C., Koskela, S., Longden, K. D., Lu, Z., Preibisch, S., Qiu, W., Rogers, E. M., Seenivasan, P., Zhao, A., Bogovic, J., . . . Reiser, M. B. (2025). Connectome-driven neural inventory of a complete visual system. *Nature*, 1.

Card, G., & Dickinson, M. H. (2008). Visually mediated motor planning in the escape response of drosophila. *Curr. Biol.*, *18*(17), 1300.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U. S. A.*, *79*(8), 2554.

Marder, E., & Bucher, D. (2007). Understanding circuit dynamics using the stomatogastric nervous system of lobsters and crabs. *Annu. Rev. Physiol.*, *69*(1), 291.

Paninski, L., & Cunningham, J. P. (2018). Neural data science: Accelerating the experiment-analysis-theory cycle in large-scale neuroscience. *Curr. Opin. Neurobiol.*, *50*, 232.

Yamins, D. L. K., Hong, H., Cadieu, C. F., Solomon, E. A., Seibert, D., & DiCarlo, J. J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proc. Natl. Acad. Sci. U. S. A.*, *111*(23), 8619.

Yamins, D. L. K., & DiCarlo, J. J. (2016). Using goal-driven deep learning models to understand sensory cortex. *Nat. Neurosci.*, *19*(3), 356.

Kriegeskorte, N. (2015). Deep neural networks: A new framework for modeling biological vision and brain information processing. *Annual Review of Vision Science*, *1*, 417.

Robie, A. A., Hirokawa, J., Edwards, A. W., Umayam, L. A., Lee, A., Phillips, M. L., Card, G. M., Korff, W., Rubin, G. M., Simpson, J. H., Reiser, M. B., & Branson, K. (2017). Mapping the neural substrates of behavior. *Cell*, *170*(2), 393.

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.*, *21*(9), 1281.

Sussillo, D., Jozefowicz, R., Abbott, L. F., & Pandarinath, C. (2016). LFADS-latent factor analysis via dynamical systems. *arXiv:1608. 06315*.

Schneider, S., Lee, J. H., & Mathis, M. W. (2023). Learnable latent embeddings for joint behavioural and neural analysis. *Nature*, *617*(7960), 360.

Mathis, M. W., Rotondo, A. P., Chang, E. F., Tolias, A. S., & Mathis, A. (2024). Decoding the brain: From neural representations to mechanistic models. *187*(21), 5814.

Bereska, L., & Gavves, E. (2024). Mechanistic interpretability for AI safety – a review. *arXiv [cs.AI]*.

Lindsey, J., Gurnee, W., Ameisen, E., Chen, B., Pearce, A., Turner, N. L., Citro, C., Abrahams, D., Carter, S., Hosmer, B., Marcus, J., Sklar, M., Templeton, A., Bricken, T., McDougall, C., Cunningham, H., Henighan, T., Jermyn, A., Jones, A., … Batson, J. (2025). On the biology of a large language model. *Transformer Circuits Thread*.

White, J. G., Southgate, E., Thomson, J. N., Brenner, S., et al. (1986). The structure of the nervous system of the nematode caenorhabditis elegans. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, *314*(1165), 1.

Jarrell, T. A., Wang, Y., Bloniarz, A. E., Brittin, C. A., Xu, M., Thomson, J. N., Albertson, D. G., Hall, D. H., & Emmons, S. W. (2012). The connectome of a decision-making neural network. *Science*, *337*(6093), 437.

Cook, S. J., Jarrell, T. A., Brittin, C. A., Wang, Y., Bloniarz, A. E., Yakovlev, M. A., Nguyen, K. C. Q., Tang, L. T.-H., Bayer, E. A., Duerr, J. S., et al. (2019). Whole-animal connectomes of both caenorhabditis elegans sexes. *Nature*, *571*(7763), 63.

Witvliet, D., Mulcahy, B., Mitchell, J. K., Meirovitch, Y., Berger, D. R., Wu, Y., Liu, Y., Koh, W. X., Parvathala, R., Holmyard, D., et al. (2021). Connectomes across development reveal principles of brain maturation. *Nature*, *596*(7871), 257.

Zheng, Z., Lauritzen, J. S., Perlman, E., Robinson, C. G., Nichols, M., Milkie, D., Torrens, O., Price, J., Fisher, C. B., Sharifi, N., et al. (2018). A complete electron microscopy volume of the brain of adult drosophila melanogaster. *Cell*, *174*(3), 730.

Scheffer, L. K., Xu, C. S., Januszewski, M., Lu, Z., Takemura, S.-Y., Hayworth, K. J., Huang, G. B., Shinomiya, K., Maitlin-Shepard, J., Berg, S., et al. (2020). A connectome and analysis of the adult drosophila central brain. *Elife*, *9*, e57443.

Dorkenwald, S., Matsliah, A., Sterling, A. R., Schlegel, P., Yu, S.-C., McKellar, C. E., Lin, A., Costa, M., Eichler, K., Yin, Y., Silversmith, W., Schneider-Mizell, C., Jordan, C. S., Brittain, D., Halageri, A., Kuehner, K., Ogedengbe, O., Morey, R., Gager, J., … Zandawala, M. (2024). Neuronal wiring diagram of an adult brain. *Nature*, *634*(8032), 124.

Hulse, B. K., Haberkern, H., Franconville, R., Turner-Evans, D., Takemura, S.-Y., Wolff, T., Noorman, M., Dreher, M., Dan, C., Parekh, R., Hermundstad, A. M., Rubin, G. M., & Jayaraman, V. (2021). A connectome of the drosophila central complex reveals network motifs suitable for flexible navigation and context-dependent action selection (R. L. Calabrese, S. Heinze, & J. Pipkin, Eds.). *Elife*, *10*, e66039.

Wanner, A. A., Genoud, C., Masudi, T., Siksou, L., & Friedrich, R. W. (2016). Dense EM-based reconstruction of the interglomerular projectome in the zebrafish olfactory bulb. *Nat. Neurosci.*, *19*(6), 816.

Hildebrand, D. G. C., Cicconet, M., Torres, R. M., Choi, W., Quan, T. M., Moon, J., Wetzel, A. W., Scott Champion, A., Graham, B. J., Randlett, O., et al. (2017). Whole-brain serial-section electron microscopy in larval zebrafish. *Nature*, *545*(7654), 345.

Svara, F., Förster, D., Kubo, F., Januszewski, M., Dal Maschio, M., Schubert, P. J., Kornfeld, J., Wanner, A. A., Laurell, E., Denk, W., et al. (2022). Automated synapse-level reconstruction of neural circuits in the larval zebrafish brain. *Nat. Methods*, 1.

Bock, D. D., Lee, W.-C. A., Kerlin, A. M., Andermann, M. L., Hood, G., Wetzel, A. W., Yurgenson, S., Soucy, E. R., Kim, H. S., & Reid, R. C. (2011). Network anatomy and in vivo physiology of visual cortical neurons. *Nature*, *471*(7337), 177.

Helmstaedter, M., Briggman, K. L., Turaga, S. C., Jain, V., Seung, H. S., & Denk, W. (2013). Connectomic reconstruction of the inner plexiform layer in the mouse retina. *Nature*, *500*(7461), 168.

Oh, S. W., Harris, J. A., Ng, L., Winslow, B., Cain, N., Mihalas, S., Wang, Q., Lau, C., Kuan, L., Henry, A. M., et al. (2014). A mesoscale connectome of the mouse brain. *Nature*, *508*(7495), 207.

Motta, A., Berning, M., Boergens, K. M., Staffler, B., Beining, M., Loomba, S., Hennig, P., Wissler, H., & Helmstaedter, M. (2019). Dense connectomic reconstruction in layer 4 of the somatosensory cortex. *Science*, *366*(6469), eaay3134.

Loomba, S., Straehle, J., Gangadharan, V., Heike, N., Khalifa, A., Motta, A., Ju, N., Sievers, M., Gempt, J., Meyer, H. S., et al. (2022). Connectomic comparison of mouse and human cortex. *Science*, *377*(6602), eabo0924.

Kasthuri, N., Hayworth, K. J., Berger, D. R., Schalek, R. L., Conchello, J. A., Knowles-Barley, S., Lee, D., Vázquez-Reina, A., Kaynig, V., Jones, T. R., et al. (2015). Saturated reconstruction of a volume of neocortex. *Cell*, *162*(3), 648.

Shapson-Coe, A., Januszewski, M., Berger, D. R., Pope, A., Wu, Y., Blakely, T., Schalek, R. L., Li, P. H., Wang, S., Maitin-Shepard, J., Karlupia, N., Dorkenwald, S., Sjostedt, E., Leavitt, L., Lee, D., Troidl, J., Collman, F., Bailey, L., Fitzmaurice, A., . . . Lichtman, J. W. (2024). A petavoxel fragment of human cerebral cortex reconstructed at nanoscale resolution. *Science*, *384*(6696), eadk4858.

Prinz, A. A., Bucher, D., & Marder, E. (2004). Similar network activity from disparate circuit parameters. *Nat. Neurosci.*, *7*(12), 1345.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Netw.*, *2*(5), 359.

Jabr, F. (2012). The connectome debate: Is mapping the mind of a worm worth it? *Sci. Am.*

Rivera-Alba, M., Vitaladevuni, S. N., Mishchenko, Y., Lu, Z., Takemura, S.-Y., Scheffer, L., Meinertzhagen, I. A., Chklovskii, D. B., & de Polavieja, G. G. (2011). Wiring economy and volume exclusion determine neuronal placement in the drosophila brain. *Curr. Biol.*, *21*(23), 2000.

Takemura, S.-Y., Xu, C. S., Lu, Z., Rivlin, P. K., Parag, T., Olbris, D. J., Plaza, S., Zhao, T., Katz, W. T., Umayam, L., Weaver, C., Hess, H. F., Horne, J. A., Nunez-Iglesias, J., Aniceto, R., Chang, L.-A., Lauchie, S., Nasca, A., Ogundeyi, O., … Scheffer, L. K. (2015). Synaptic circuits and their variations within different columns in the visual system of Drosophila. *Proceedings of the National Academy of Sciences*, *112*(44), 13711.

Takemura, S.-Y., Nern, A., Chklovskii, D. B., Scheffer, L. K., Rubin, G. M., & Meinertzhagen, I. A. (2017). The comprehensive connectome of a neural substrate for 'ON' motion detection in Drosophila. *Elife*, *6*, 1.

Shinomiya, K., Huang, G., Lu, Z., Parag, T., Xu, C. S., Aniceto, R., Ansari, N., Cheatham, N., Lauchie, S., Neace, E., Ogundeyi, O., Ordish, C., Peel, D., Shinomiya, A., Smith, C., Takemura, S., Talebi, I., Rivlin, P. K., Nern, A., … Meinertzhagen, I. A. (2019). Comparisons between the ON- and OFF-edge motion pathways in the Drosophila brain. *Elife*, *8*, 2431.

Shinomiya, K., Nern, A., Meinertzhagen, I. A., Plaza, S. M., & Reiser, M. B. (2022). Neuronal circuits integrating visual motion information in drosophila melanogaster. *Curr. Biol.*, *32*(16), 3529.

Butler, D. J., Wulff, J., Stanley, G. B., & Black, M. J. (2012). A naturalistic open source movie for optical flow evaluation. *European conference on computer vision*, 611.

Lappalainen, J. K., Tschopp, F. D., Prakhya, S., McGill, M., Nern, A., Shinomiya, K., Takemura, S.-Y., Gruntman, E., Macke, J. H., & Turaga, S. C. (2024). Connectome-constrained networks predict neural activity across the fly visual system. *Nature*, *634*(8036), 1132.

Dayan, P., & Abbott, L. F. (2005). Theoretical neuroscience: Computational and mathematical modeling of neural systems.

Gerstner, W., Kistler, W. M., Naud, R., & Paninski, L. (2014). *Neuronal dynamics: From single neurons to networks and models of cognition*. Cambridge University Press.

Abbott, L. F. (1999). Lapicque's introduction of the integrate-and-fire model neuron (1907). *Brain Res. Bull.*, *50*(5-6), 303.

Curtis, H. J., & Cole, K. S. (1938). Transverse electric impedance of the squid giant axon. *J. Gen. Physiol.*, *21*(6), 757.

Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.*, *117*(4), 500.

Turner, G. C., Bazhenov, M., & Laurent, G. (2008). Olfactory representations by drosophila mushroom body neurons. *J. Neurophysiol.*, *99*(2), 734.

Gruntman, E., Romani, S., & Reiser, M. B. (2018). Simple integration of fast excitation and offset, delayed inhibition computes directional selectivity in Drosophila. *Nat. Neurosci.*, *21*(2), 250.

Groschner, L. N., Malis, J. G., Zuidinga, B., & Borst, A. (2022). A biophysical account of multiplication by a single neuron. *Nature*, *603*(7899), 119.

Romani, S., & Tsodyks, M. (2015). Short-term plasticity based network model of place cells dynamics: Short-term plasticity based network model of place cells dynamics. *Hippocampus*, *25*(1), 94.

Markram, H., Muller, E., Ramaswamy, S., Reimann, M. W., Abdellah, M., Sanchez, C. A., Ailamaki, A., Alonso-Nanclares, L., Antille, N., Arsever, S., Kahou, G. A. A., Berger, T. K., Bilgili, A., Buncic, N., Chalimourda, A., Chindemi, G., Courcol, J.-D., Delalondre, F., Delattre, V., ... Schürmann, F. (2015). Reconstruction and simulation of neocortical microcircuitry. *Cell*, *163*(2), 456.

Billeh, Y. N., Cai, B., Gratiy, S. L., Dai, K., Iyer, R., Gouwens, N. W., Abbasi-Asl, R., Jia, X., Siegle, J. H., Olsen, S. R., Koch, C., Mihalas, S., & Arkhipov, A. (2020). Systematic integration of structural and functional data into multi-scale models of mouse primary visual cortex. *Neuron*, *106*(3), 388.

Coulomb, C. A. (1785). Premier mémoire sur l'électricité et le magnétisme. *Histoire de l'Academie royale des sciences*, *569*.

Griffiths, D. J. (2023, Nov.). *Introduction to electrodynamics* (5th ed.). Cambridge University Press.

Koch, C. (2004). *Biophysics of computation: Information processing in single neurons*. Oxford university press.

OpenAI. (2025). DALL· E (2025 version).

Shriki, O., Hansel, D., & Sompolinsky, H. (2003). Rate models for conductance-based cortical neuronal networks. *Neural Comput.*, *15*(8), 1809.

Miller, K. D., & Fumarola, F. (2012). Mathematical equivalence of two common forms of firing rate models of neural networks. *Neural Comput.*, *24*(1), 25.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, *5*(4), 115.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev.*, *65*(6), 386.

Rosenblatt, F. (1962). Principles of neurodynamics. *Perceptrons and the theory of brain mechanisms.*

Widrow, B., & Hoff, M. E. (1960). Adaptive switching circuits. *IRE WESCON Convention Record*, *4*, 96.

Minsky, M. L., & Papert, S. A. (1969). *Perceptrons: An introduction to computational geometry* (1st). MIT Press.

Linnainmaa, S. (1970). *The representation of the cumulative rounding error of an algorithm as a taylor expansion of the local rounding errors* [Master's thesis, University of Helsinki].

Werbos, P. J. (1974). *Beyond regression: New tools for prediction and analysis in the behavioral sciences* [Doctoral dissertation, Harvard University].

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533.

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, *5*(4), 455.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *arXiv [cs.CL].*

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Raghu, M., Unterthiner, T., Kornblith, S., Zhang, C., & Dosovitskiy, A. (2021). Do vision transformers see like convolutional neural networks? *Adv. Neural Inf. Process. Syst.*, *34*, 12116.

Bishop, C. M., & Nasrabadi, N. M. (2006, Aug.). *Pattern recognition and machine learning* (1st ed.). Springer.

Bronstein, M. M., Bruna, J., Cohen, T., & Veličković, P. (2021). Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *arXiv [cs.LG].*

Fukushima, K., & Miyake, S. (1982). Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. In *Competition and cooperation in neural nets* (pp. 267–285). Springer.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.*, *160*(1), 106.

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Comput.*, *1*(4), 541.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks (F. Pereira, C. J.

Burges, L. Bottou, & K. Q. Weinberger, Eds.). *Commun. ACM*, *60*, 84.

Seung, H. S. (2024). Predicting visual function by interpreting a neuronal wiring diagram. *Nature*, *634*(8032), 113.

Wilson, H. R., & Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.*, *12*(1), 1.

Jordan, M. I. (1986). Attractor dynamics and parallelism in a connectionist sequential machine. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *8*(0).

Elman, J. (1990). Finding structure in time. *Cogn. Sci.*, *14*(2), 179.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Comput.*, *9*, 1735.

Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv [cs.CL]*.

Orvieto, A., Smith, S. L., Gu, A., Fernando, A., Gulcehre, C., Pascanu, R., & De, S. (2023). Resurrecting recurrent neural networks for long sequences. In A. Krause, E. Brunskill, K. Cho, B. Engelhardt, S. Sabato, & J. Scarlett (Eds.), *Proceedings of the 40th international conference on machine learning* (pp. 26670–26698, Vol. 202). PMLR.

Funahashi, K.-I., & Nakamura, Y. (1993). Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Netw.*, *6*(6), 801.

Schäfer, A. M., & Zimmermann, H.-G. (2007). Recurrent neural networks are universal approximators. *Int. J. Neural Syst.*, *17*(4), 253.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv [cs.CL]*.

Amarimber, S.-I. (1972). Characteristics of random nets of analog neuron-like elements. *IEEE Trans. Syst. Man Cybern.*, *SMC-2*(5), 643.

Hopfield, J. J. (1984). Neurons with graded response have collective computational properties like those of two-state neurons. *Proc. Natl. Acad. Sci. U. S. A.*, *81*(10), 3088.

Sompolinsky, H., Crisanti, A., & Sommers, H. J. (1988). Chaos in random neural networks. *Phys. Rev. Lett.*, *61*(3), 259.

Beer, R. D. (1995). A dynamical systems perspective on agent-environment interaction. *Artif. Intell.*, *72*(1-2), 173.

Kleinfeld, D., & Sompolinsky, H. (1988). Associative neural network model for the generation of temporal patterns. theory and application to central pattern generators. *Biophys. J.*, *54*(6), 1039.

Seung, H. S., Lee, D. D., Reis, B. Y., & Tank, D. W. (2000). Stability of the memory of eye position in a recurrent network of conductance-based model neurons. *Neuron*, *26*(1), 259.

Vogels, T. P., Rajan, K., & Abbott, L. F. (2005). Neural network dynamics. *Annu. Rev. Neurosci.*, *28*(1), 357.

Rajan, K., Abbott, L. F., & Sompolinsky, H. (2010). Stimulus-dependent suppression of chaos in recurrent neural networks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.*, *82*(1 Pt 1), 011903.

Sussillo, D., & Abbott, L. F. (2009). Generating coherent patterns of activity from chaotic neural networks. *Neuron*, *63*(4), 544.

Sussillo, D., & Barak, O. (2013). Opening the black box: Low-dimensional dynamics in high-dimensional recurrent neural networks. *Neural Comput.*, *25*(3), 626.

Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *nature*, *503*(7474), 78.

Sussillo, D. (2014). Neural circuits as computational dynamical systems. *Curr. Opin. Neurobiol.*, *25*, 156.

Pandarinath, C., O'Shea, D. J., Collins, J., Jozefowicz, R., Stavisky, S. D., Kao, J. C., Trautmann, E. M., Kaufman, M. T., Ryu, S. I., & Hochberg, L. R. (2018). Inferring single-trial neural population dynamics using sequential auto-encoders. *Nat. Methods*, *15*(10), 805.

Chen, R. T. Q., Rubanova, Y., Bettencourt, J., & Duvenaud, D. K. (2018). Neural ordinary differential equations. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, & R. Garnett (Eds.), *Advances in neural information processing systems* (Vol. 31). Curran Associates, Inc.

Kidger, P., Morrill, J., Foster, J., & Lyons, T. (2020). Neural controlled differential equations for irregular time series. *arXiv [cs.LG]*.

Zeraati, R., Levina, A., Macke, J. H., & Gao, R. (2024). Neural timescales from a computational perspective. *arXiv [q-bio.NC]*.

Werbos, P. J. (1990). Backpropagation through time: What it does and how to do it. *Proc. IEEE*, *78*(10), 1550.

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., & Lerer, A. (2017). Automatic differentiation in pytorch. *AutoDiff Workshop, Neural Information Processing Systems 2017*.

Bradbury, J., Frostig, R., Hawkins, P., Johnson, M. J., Leary, C., Maclaurin, D., Necula, G., Paszke, A., VanderPlas, J., Wanderman-Milne, S., & Zhang, Q. (2018). JAX: Composable transformations of Python+NumPy programs.

Pascanu, R., Mikolov, T., & Bengio, Y. (2013). On the difficulty of training recurrent neural networks, 1310.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248.

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., & Sutskever, I. (2021). Learning transferable visual models from natural language supervision. *arXiv [cs.CV]*.

Dosovitskiy, A., Fischer, P., Ilg, E., Hausser, P., Hazirbas, C., Golkov, V., Van Der Smagt, P., Cremers, D., & Brox, T. (2015). Flownet: Learning optical flow with convolutional networks. *Proceedings of the IEEE international conference on computer vision*, 2758.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv [cs.CL]*.

Radford, A., & Narasimhan, K. (2018). Improving language understanding by generative pre-training.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. Wiley.

Bliss, T. V., & Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J. Physiol.*, *232*(2), 331.

Bi, G. Q., & Poo, M. M. (1998). Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *J. Neurosci.*, *18*(24), 10464.

Abbott, L. F., & Nelson, S. B. (2000). Synaptic plasticity: Taming the beast. *Nat. Neurosci.*, *3 Suppl*(S11), 1178.

Caporale, N., & Dan, Y. (2008). Spike timing-dependent plasticity: A hebbian learning rule. *Annu. Rev. Neurosci.*, *31*(1), 25.

Marder, E. (2012). Neuromodulation of neuronal circuits: Back to the future. *Neuron*, *76*(1), 1.

Richards, B. A., Lillicrap, T. P., Beaudoin, P., Bengio, Y., Bogacz, R., Christensen, A., Clopath, C., Costa, R. P., de Berker, A., Ganguli, S., Gillon,

C. J., Hafner, D., Kepecs, A., Kriegeskorte, N., Latham, P., Lindsay, G. W., Miller, K. D., Naud, R., Pack, C. C., … Kording, K. P. (2019). A deep learning framework for neuroscience. *Nature Neuroscience*, *22*(11), 1761.

Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., & Hinton, G. (2020). Backpropagation and the brain. *Nat. Rev. Neurosci.*, *21*(6), 335.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. W. H. Freeman; Company.

Jozwik, K. M., Kietzmann, T. C., Cichy, R. M., Kriegeskorte, N., & Mur, M. (2023). Deep neural networks and visuo-semantic models explain complementary components of human ventral-stream representational dynamics. *J. Neurosci.*, *43*(10), 1731.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al. (2019). Pytorch: An imperative style, high-performance deep learning library. *Adv. Neural Inf. Process. Syst.*, *32*.

Tian, L., Hires, S. A., Mao, T., Huber, D., Chiappe, M. E., Chalasani, S. H., Petreanu, L., Akerboom, J., McKinney, S. A., Schreiter, E. R., Bargmann, C. I., Jayaraman, V., Svoboda, K., & Looger, L. L. (2009). Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nat. Methods*, *6*(12), 875.

Ahrens, M., Orger, M., Robson, D., Li, J. M., & Keller, P. J. (2013). Whole-brain functional imaging at cellular resolution using light-sheet microscopy. *Nat. Methods*, *10*, 413.

Stirman, J. N., Smith, I. T., Kudenov, M. W., & Smith, S. L. (2016). Wide field-of-view, multi-region, two-photon imaging of neuronal activity in the mammalian brain. *Nat. Biotechnol.*, *34*(8), 857.

Jun, J. J., Steinmetz, N. A., Siegle, J. H., Denman, D. J., Bauza, M., Barbarits, B., Lee, A. K., Anastassiou, C. A., Andrei, A., Aydın, Ç., Barbic, M., Blanche, T. J., Bonin, V., Couto, J., Dutta, B., Gratiy, S. L., Gutnisky, D. A., Häusser, M., Karsh, B., … Harris, T. D. (2017). Fully integrated silicon probes for high-density recording of neural activity. *Nature*, *551*(7679), 232.

Branson, K., Robie, A. A., Bender, J., Perona, P., & Dickinson, M. H. (2009). High-throughput ethomics in large groups of drosophila. *Nat. Methods*, *6*(6), 451.

Klindt, D. A., Ecker, A. S., Euler, T., & Bethge, M. (2017). Neural system identification for large populations separating "what" and "where". *arXiv [stat.ML]*.

Cadena, S. A., Denfield, G. H., Walker, E. Y., Gatys, L. A., Tolias, A. S., Bethge, M., & Ecker, A. S. (2019). Deep convolutional models improve predictions of macaque V1 responses to natural images. *PLoS Comput. Biol.*, *15*(4), e1006897.

Kindel, W. F., Christensen, E. D., & Zylberberg, J. (2019). Using deep learning to probe the neural code for images in primary visual cortex. *J. Vis.*, *19*(4), 29.

Zhang, Y., Lee, T. S., Li, M., Liu, F., & Tang, S. (2019). Convolutional neural network models of V1 responses to complex patterns. *J. Comput. Neurosci.*, *46*(1), 33.

Walker, E. Y., Sinz, F. H., Cobos, E., Muhammad, T., Froudarakis, E., Fahey, P. G., Ecker, A. S., Reimer, J., Pitkow, X., & Tolias, A. S. (2019). Inception loops discover what excites neurons most using deep predictive models. *Nat. Neurosci.*, *22*(12), 2060.

Willeke, K. F., Restivo, K., Franke, K., Nix, A. F., Cadena, S. A., Shinn, T., Nealley, C., Rodriguez, G., Patel, S., Ecker, A. S., Sinz, F. H., & Tolias, A. S. (2023). Deep learning-driven characterization of single cell tuning in primate visual area V4 unveils topological organization. *bioRxiv*, 2023.05. 12.540591.

Ito, T., & Murray, J. D. (2023). Multitask representations in the human cortex transform along a sensory-to-motor hierarchy. *Nat. Neurosci.*, *26*(2), 306.

Khaligh-Razavi, S.-M., & Kriegeskorte, N. (2014). Deep supervised, but not unsupervised, models may explain IT cortical representation. *PLoS Comput. Biol.*, *10*(11), e1003915.

Kell, A. J. E., Yamins, D. L. K., Shook, E. N., Norman-Haignere, S. V., & McDermott, J. H. (2018). A task-optimized neural network replicates human auditory behavior, predicts brain responses, and reveals a cortical processing hierarchy. *Neuron*, *98*(3), 630.

Tuckute, G., Feather, J., Boebinger, D., & McDermott, J. H. (2023). Many but not all deep neural network audio models capture brain responses and exhibit correspondence between model stages and brain regions. *PLoS Biol.*, *21*(12), e3002366.

Yang, G. R., Joglekar, M. R., Song, H. F., Newsome, W. T., & Wang, X.-J. (2019). Task representations in neural networks trained to perform many cognitive tasks. *Nat. Neurosci.*, *22*(2), 297.

Piwek, E. P., Stokes, M. G., & Summerfield, C. (2023). A recurrent neural network model of prefrontal brain activity during a working memory task. *PLoS Comput. Biol.*, *19*(10), e1011555.

Marin Vargas, A., Bisi, A., Chiappa, A. S., Versteeg, C., Miller, L. E., & Mathis, A. (2024). Task-driven neural network models predict neural dynamics of proprioception. *Cell*, *187*(7), 1745.

Schrimpf, M., Blank, I. A., Tuckute, G., Kauf, C., Hosseini, E. A., Kanwisher, N., Tenenbaum, J. B., & Fedorenko, E. (2021). The neural architecture of language: Integrative modeling converges on predictive processing. *Proc. Natl. Acad. Sci. U. S. A.*, *118*(45), e2105646118.

Schaeffer, R., Khona, M., & Fiete, I. R. (2022). No free lunch from deep learning in neuroscience: A case study through models of the entorhinal-hippocampal circuit. *bioRxiv*, 2022.08.07.503109.

Schaeffer, R., Khona, M., Chandra, S., Ostrow, M., Miranda, B., & Koyejo, S. (2024). Position: Maximizing neural regression scores may not identify good models of the brain. *UniReps: 2nd Edition of the Workshop on Unifying Representations in Neural Models*.

Marder, E., & Taylor, A. L. (2011). Multiple models to capture the variability in biological neurons and networks. *Nat. Neurosci.*, *14*(2), 133.

Gonçalves, P. J., Lueckmann, J.-M., Deistler, M., Nonnenmacher, M., Öcal, K., Bassetto, G., Chintaluri, C., Podlaski, W. F., Haddad, S. A., Vogels, T. P., Greenberg, D. S., & Macke, J. H. (2020). Training deep neural density estimators to identify mechanistic models of neural dynamics. *Elife*, *9*, e56261.

Biswas, T., & Fitzgerald, J. E. (2022). Geometric framework to predict structure from function in neural networks. *Phys. Rev. Res.*, *4*(2), 023255.

Lakshminarayanan, B., Pritzel, A., & Blundell, C. (2017). Simple and scalable predictive uncertainty estimation using deep ensembles. *Adv. Neural Inf. Process. Syst.*, *30*.

Wilson, A. G. (2020). The case for bayesian deep learning. *arXiv preprint arXiv:2001. 10995*.

Dietterich, T. G. (2000). Ensemble methods in machine learning. In *Multiple classifier systems* (pp. 1–15). Springer Berlin Heidelberg.

Cunningham, J. P., & Yu, B. M. (2014). Dimensionality reduction for large-scale neural recordings. *Nat. Neurosci.*, *17*(11), 1500.

Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I. W. H., Ng, L. G., Ginhoux, F., & Newell, E. W. (2019). Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.*, *37*(1), 38.

Meigen, J. W. (1830). *Systematische beschreibung der bekannten europäischen zweiflügeligen insekten. sechster theil*. Schulzische Buchhandlung.

St Johnston, D. (2002). The art and design of genetic screens: Drosophila melanogaster. *Nat. Rev. Genet.*, *3*(3), 176.

Muqit, M. M. K., & Feany, M. B. (2002). Modelling neurodegenerative diseases in drosophila: A fruitful approach? *Nat. Rev. Neurosci.*, *3*(3), 237.

Prüßing, K., Voigt, A., & Schulz, J. B. (2013). Drosophila melanogaster as a model organism for alzheimer's disease. *Mol. Neurodegener.*, *8*(1), 35.

Gonzalez, C. (2013). Drosophila melanogaster: A model and a tool to investigate malignancy and identify new therapeutics. *Nat. Rev. Cancer*, *13*(3), 172.

Dow, J. A. T., Simons, M., & Romero, M. F. (2022). Drosophila melanogaster: A simple genetic model of kidney structure, function and disease. *Nat. Rev. Nephrol.*, *18*(7), 417.

FlyBase Consortium. (1998). FlyBase: A drosophila database. *Nucleic Acids Res.*, *26*(1), 85.

Drysdale, R. A., Crosby, M. A., & FlyBase Consortium. (2005). FlyBase: Genes and gene models. *Nucleic Acids Res.*, *33*(Database issue), D390.

Dorkenwald, S., McKellar, C. E., Macrina, T., Kemnitz, N., Lee, K., Lu, R., Wu, J., Popovych, S., Mitchell, E., Nehoran, B., Jia, Z., Bae, J. A., Mu, S., Ih, D., Castro, M., Ogedengbe, O., Halageri, A., Kuehner, K., Sterling, A. R., . . . Seung, H. S. (2022). FlyWire: online community for whole-brain connectomics. *Nat. Methods*, *19*(1), 119.

Plaza, S. M., Clements, J., Dolafi, T., Umayam, L., Neubarth, N. N., Scheffer, L. K., & Berg, S. (2022). neuPrint: An open access tool for EM connectomics. *Front. Neuroinform.*, *16*, 896292.

Ukani, N. H., Yeh, C.-H., Tomkins, A., Zhou, Y., Florescu, D., Ortiz, C. L., Huang, Y.-C., Wang, C.-T., Richmond, P., Lo, C., Coca, D., Chiang, A., & Lazar, A. (2016). The fruit fly brain observatory: From structure to function. *bioRxiv*.

Court, R., Costa, M., Pilgrim, C., Millburn, G., Holmes, A., McLachlan, A., Larkin, A., Matentzoglu, N., Kir, H., Parkinson, H., Brown, N. H., O'Kane, C. J., Armstrong, J. D., Jefferis, G. S. X. E., & Osumi-Sutherland, D. (2023). Virtual fly brain-an interactive atlas of the drosophila nervous system. *Front. Physiol.*, *14*, 1076533.

Matthews, K. A., Kaufman, T. C., & Gelbart, W. M. (2005). Research resources for drosophila: The expanding universe. *Nat. Rev. Genet.*, *6*(3), 179.

Fiala, A., Spall, T., Diegelmann, S., Eisermann, B., Sachse, S., Devaud, J.-M., Buchner, E., & Galizia, C. G. (2002). Genetically expressed cameleon in drosophila melanogaster is used to visualize olfactory information in projection neurons. *Curr. Biol.*, *12*(21), 1877.

Davis, F. P., Nern, A., Picard, S., Reiser, M. B., Rubin, G. M., Eddy, S. R., & Henry, G. L. (2020). A genetic, genomic, and computational resource for exploring neural circuit function. *Elife*, *9*.

Sato, D. X., Okuyama, T., & Takahashi, Y. (2025). Multifaceted and extensive behavioral trajectories of genomically diverse drosophila lines. *Sci. Data*, *12*(1), 400.

Meissner, G. W., Vannan, A., Jeter, J., Close, K., DePasquale, G. M., Dorman, Z., Forster, K., Beringer, J. A., Gibney, T. V., Hausenfluck, J. H., He, Y., Henderson, K., Johnson, L., Johnston, R. M., Ihrke, G., Iyer, N., Lazarus, R., Lee, K., Li, H.-H., … FlyLight Project Team. (2024). A split-GAL4 driver line resource for drosophila CNS cell types. *eLife*.

Inagaki, H. K., Jung, Y., Hoopfer, E. D., Wong, A. M., Mishra, N., Lin, J. Y., Tsien, R. Y., & Anderson, D. J. (2014). Optogenetic control of drosophila using a red-shifted channelrhodopsin reveals experience-dependent influences on courtship. *Nat. Methods*, *11*(3), 325.

Bellen, H. J., Tong, C., & Tsuda, H. (2010). 100 years of drosophila research and its impact on vertebrate neuroscience: A history lesson for the future. *Nat. Rev. Neurosci.*, *11*(7), 514.

Denk, W., & Horstmann, H. (2004). Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. *PLoS Biol.*, *2*(11), e329.

Knott, G., Marchman, H., Wall, D., & Lich, B. (2008). Serial section scanning electron microscopy of adult brain tissue using focused ion beam milling. *J. Neurosci.*, *28*(12), 2959.

Winnubst, J., Bas, E., Ferreira, T. A., Wu, Z., Economo, M. N., Edson, P., Arthur, B. J., Bruns, C., Rokicki, K., Schauder, D., Olbris, D. J., Murphy, S. D., Ackerman, D. G., Arshadi, C., Baldwin, P., Blake, R., Elsayed, A., Hasan, M., Ramirez, D., … Chandrashekar, J. (2019). Reconstruction of 1,000 projection neurons reveals new cell types and organization of long-range connectivity in the mouse brain. *Cell*, *179*(1), 268.

Basser, P. J., Pajevic, S., Pierpaoli, C., Duda, J., & Aldroubi, A. (2000). In vivo fiber tractography using DT-MRI data. *Magn. Reson. Med.*, *44*(4), 625.

Sporns, O., Tononi, G., & Kötter, R. (2005). The human connectome: A structural description of the human brain. *PLoS Comput. Biol.*, *1*(4), e42.

Hagmann, P. (2005). *From diffusion MRI to brain connectomics* [Doctoral dissertation, EPFL]. Lausanne, EPFL.

Biswal, B., Yetkin, F. Z., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.*, *34*(4), 537.

Ripoll-Sánchez, L., Watteyne, J., Sun, H., Fernandez, R., Taylor, S. R., Weinreb, A., Bentley, B. L., Hammarlund, M., Miller, D. M., 3rd, Hobert, O., Beets, I., Vértes, P. E., & Schafer, W. R. (2023). The neuropeptidergic connectome of C. elegans. *Neuron*, *111*(22), 3570.

Varshney, L. R., Chen, B. L., Paniagua, E., Hall, D. H., & Chklovskii, D. B. (2011). Structural properties of the caenorhabditis elegans neuronal network. *PLoS Comput. Biol.*, *7*(2), e1001066.

Chen, Z., Yu, Y., & Xue, X. (2023). A connectome-based digital twin caenorhabditis elegans capable of intelligent sensorimotor behavior. *Mathematics*, *11*(11), 2442.

Sarma, G. P., Lee, C. W., Portegys, T., Ghayoomie, V., Jacobs, T., Alicea, B., Cantarelli, M., Currie, M., Gerkin, R. C., Gingell, S., Gleeson, P., Gordon, R., Hasani, R. M., Idili, G., Khayrulin, S., Lung, D., Palyanov, A., Watts, M., & Larson, S. D. (2018). OpenWorm: Overview and recent advances in integrative biological simulation of caenorhabditis elegans. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, *373*(1758), 20170382.

Watteyne, J., Chudinova, A., Ripoll-Sánchez, L., Schafer, W. R., & Beets, I. (2024). Neuropeptide signaling network of caenorhabditis elegans: From structure to behavior. *Genetics*, *228*.

Taylor, S. R., Santpere, G., Weinreb, A., Barrett, A., Reilly, M. B., Xu, C., Varol, E., Oikonomou, P., Glenwinkel, L., McWhirter, R., Poff, A., Basavaraju, M., Rafi, I., Yemini, E., Cook, S. J., Abrams, A., Vidal, B., Cros, C., Tavazoie, S., … Miller, D. M., 3rd. (2021). Molecular topography of an entire nervous system. *Cell*, *184*(16), 4329.

Beets, I., Zels, S., Vandewyer, E., Demeulemeester, J., Caers, J., Baytemur, E., Courtney, A., Golinelli, L., Hasakioğulları, İ., Schafer, W. R., Vértes, P. E., Mirabeau, O., & Schoofs, L. (2023). System-wide mapping of peptide-GPCR interactions in C. elegans. *Cell Rep.*, *42*(9), 113058.

Amunts, K., Ebell, C., Muller, J., Telefont, M., Knoll, A., & Lippert, T. (2016). The human brain project: Creating a european research infrastructure to decode the human brain. *Neuron*, *92*(3), 574.

Eliasmith, C., & Trujillo, O. (2014). The use and abuse of large-scale brain models. *Curr. Opin. Neurobiol.*, *25*, 1.

Hill, S. L., Wang, Y., Riachi, I., Schürmann, F., & Markram, H. (2012). Statistical connectivity provides a sufficient foundation for specific functional connectivity in neocortical neural microcircuits. *Proc. Natl. Acad. Sci. U. S. A.*, *109*(42), E2885.

Nakagawa, T. T., Jirsa, V. K., Spiegler, A., McIntosh, A. R., & Deco, G. (2013). Bottom up modeling of the connectome: Linking structure and function in the resting brain and their changes in aging. *Neuroimage*, *80*, 318.

Potjans, T. C., & Diesmann, M. (2014). The cell-type specific cortical microcircuit: Relating structure and activity in a full-scale spiking network model. *Cereb. Cortex*, *24*(3), 785.

Shapson-Coe, A., Januszewski, M., Berger, D. R., Pope, A., Wu, Y., Blakely, T., Schalek, R. L., Li, P. H., Wang, S., Maitin-Shepard, J., et al. (2021). A connectomic study of a petascale fragment of human cerebral cortex. *BioRxiv*.

Sievers, M., Motta, A., Schmidt, M., Yener, Y., Loomba, S., Song, K., Bruett, J., & Helmstaedter, M. (2024). Connectomic reconstruction of a cortical column. *bioRxiv*, 2024.03.22.586254.

Horton, J. C., & Adams, D. L. (2005). The cortical column: A structure without a function. *Philos. Trans. R. Soc. Lond. B Biol. Sci.*, *360*(1456), 837.

MICrONS Consortium. (2025). Functional connectomics spanning multiple areas of mouse visual cortex. *Nature*, *640*(8058), 435.

Marin, E. C., Morris, B. J., Stürner, T., Champion, A. S., Krzeminski, D., Badalamente, G., Gkantia, M., Dunne, C. R., Eichler, K., Takemura, S.-Y., Tamimi, I. F. M., Fang, S., Moon, S. S., Cheong, H. S. J., Li, F., Schlegel, P., Ahnert, S. E., Berg, S., Card, G. M., . . . Janelia FlyEM Project Team. (2023). Systematic annotation of a complete adult male *Drosophila* nerve cord connectome reveals principles of functional organisation. *bioRxiv*.

Takemura, S.-Y., Hayworth, K. J., Huang, G. B., Januszewski, M., Lu, Z., Marin, E. C., Preibisch, S., Xu, C. S., Bogovic, J., Champion, A. S., Cheong, H. S. J., Costa, M., Eichler, K., Katz, W., Knecht, C., Li, F., Morris, B. J., Ordish, C., Rivlin, P. K., . . . Berg, S. (2024). A connectome of the male drosophila ventral nerve cord. *eLife*.

Schlegel, P., Yin, Y., Bates, A. S., Dorkenwald, S., Eichler, K., Brooks, P., Han, D. S., Gkantia, M., dos Santos, M., Munnelly, E. J., Badalamente, G., Serratosa Capdevila, L., Sane, V. A., Fragniere, A. M. C., Kiassat, L., Pleijzier, M. W., Stürner, T., Tamimi, I. F. M., Dunne, C. R., . . . Jefferis, G. S. X. E. (2024). Whole-brain annotation and multi-connectome cell typing of drosophila. *Nature*, *634*(8032), 139.

Eckstein, N., Bates, A. S., Champion, A., Du, M., Yin, Y., Schlegel, P., Lu, A. K.-Y., Rymer, T., Finley-May, S., Paterson, T., Parekh, R., Dorkenwald, S., Matsliah, A., Yu, S.-C., McKellar, C., Sterling, A., Eichler, K., Costa, M., Seung, S., . . . Funke, J. (2024). Neurotransmitter classification from electron microscopy images at synaptic sites in drosophila melanogaster. *Cell*, *187*(10), 2574.

Pfeiffer, K., & Homberg, U. (2014). Organization and functional roles of the central complex in the insect brain. *Annu. Rev. Entomol.*, *59*(1), 165.

Franconville, R., Beron, C., & Jayaraman, V. (2018). Building a functional connectome of the drosophila central complex. *eLife*, *7*, e37017.

Modi, M. N., Shuai, Y., & Turner, G. C. (2020). The drosophila mushroom body: From architecture to algorithm in a learning circuit. *Annu. Rev. Neurosci.*, *43*(1), 465.

Brand, A. H., & Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development*, *118*(2), 401.

Wilson, R. I., Turner, G. C., & Laurent, G. (2004). Transformation of olfactory representations in the drosophila antennal lobe. *Science*, *303*(5656), 366.

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-seq. *Nat. Methods*, *5*(7), 621.

Lappalainen, J. K. (2024, Dec.). How a tiny animal helps us improve brain simulations with AI.

Laughlin, S. B., & Hardie, R. C. (1978). Common strategies for light adaptation in the peripheral visual systems of fly and dragonfly. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.*, *128*(4), 319.

Zettler, F., & Järvilehto, M. (1971). Decrement-free conduction of graded potentials along the axon of a monopolar neuron. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.*, *75*(4), 402.

Kretzberg, J., Warzecha, A.-K., & Egelhaaf, M. (2001). Neural coding with graded membrane potential changes and spikes. *J. Comput. Neurosci.*, *11*(2), 153.

Strausfeld, N. J. (1976). *Atlas of an insect brain*. Springer.

Fischbach, K. F., & Dittrich, A. P. M. (1989). The optic lobe of Drosophila melanogaster. I. A Golgi analysis of wild-type structure. *Cell Tissue Res.*, *258*(3), 441.

Matsliah, A., Yu, S.-C., Kruk, K., Bland, D., Burke, A. T., Gager, J., Hebditch, J., Silverman, B., Willie, K. P., Willie, R., Sorek, M., Sterling, A. R., Kind, E., Garner, D., Sancer, G., Wernet, M. F., Kim, S. S., Murthy, M., Seung, H. S., ... Jefferis, G. S. X. E. (2024). Neuronal parts list and wiring diagram for a visual system. *Nature*, *634*(8032), 166.

Januszewski, M., Kornfeld, J., Li, P. H., Pope, A., Blakely, T., Lindsey, L., Maitin-Shepard, J., Tyka, M., Denk, W., & Jain, V. (2018). High-precision automated reconstruction of neurons with flood-filling networks. *Nat. Methods*, *15*(8), 605.

Maitin-Shepard, J., Baden, A., Silversmith, W., Perlman, E., Collman, F., Blakely, T., Funke, J., Jordan, C., Falk, B., Kemnitz, N., Tingzhao, Roat, C., Castro, M., Jagannathan, S., Moenigin, Clements, J., Hoag, A., Katz, B., Parsons, D., ... Li, P. H. (2021, Oct.). Google/neuroglancer:

Mishra, M., & Knust, E. (2013). Analysis of the drosophila compound eye with light and electron microscopy. *Methods Mol. Biol.*, *935*, 161.

Arenz, A., Drews, M. S., Richter, F. G., Ammer, G., & Borst, A. (2017). The Temporal Tuning of the Drosophila Motion Detectors Is Determined by the Dynamics of Their Input Elements. *Curr. Biol.*, *27*(7), 929.

Takemura, S.-Y., Bharioke, A., Lu, Z., Nern, A., Vitaladevuni, S., Rivlin, P. K., Katz, W. T., Olbris, D. J., Plaza, S. M., Winston, P., Zhao, T., Horne, J. A., Fetter, R. D., Takemura, S., Blazek, K., Chang, L. A., Ogundeyi, O., Saunders, M. A., Shapiro, V., ... Chklovskii, D. B. (2013). A visual motion detection circuit suggested by Drosophila connectomics. *Nature*, *500*(7461), 175.

Helfrich-Förster, C., Shafer, O. T., Wülbeck, C., Grieshaber, E., Rieger, D., & Taghert, P. (2007). Development and morphology of the clock-gene-expressing lateral neurons of drosophila melanogaster. *J. Comp. Neurol.*, *500*(1), 47.

Reinhard, N., Fukuda, A., Manoli, G., Derksen, E., Saito, A., Möller, G., Sekiguchi, M., Rieger, D., Helfrich-Förster, C., Yoshii, T., & Zandawala, M. (2024). Synaptic connectome of the drosophila circadian clock. *Nature Communications*, *15*(1), 1.

Nern, A., Pfeiffer, B. D., & Rubin, G. M. (2015). Optimized tools for multicolor stochastic labeling reveal diverse stereotyped cell arrangements in the fly visual system. *Proc. Natl. Acad. Sci. U. S. A.*, *112*(22), E2967.

Takemura, S. Y., Lu, Z., & Meinertzhagen, I. A. (2008). Synaptic circuits of the drosophila optic lobe: The input terminals to the medulla. *J. Comp. Neurol.*, *509*(5), 493.

Ryu, L., Kim, S. Y., & Kim, A. J. (2022). From photons to behaviors: Neural implementations of visual behaviors in drosophila. *Front. Neurosci.*, *16*, 883640.

Bahl, A., Ammer, G., Schilling, T., & Borst, A. (2013). Object tracking in motion-blind flies. *Nat. Neurosci.*, *16*, 730.

Creamer, M. S., Mano, O., & Clark, D. A. (2018). Visual control of walking speed in drosophila. *Neuron*, *100*(6), 1460.

Fry, S. N., Rohrseitz, N., Straw, A. D., & Dickinson, M. H. (2009). Visual control of flight speed in drosophila melanogaster. *J. Exp. Biol.*, *212*(Pt 8), 1120.

Yamamoto, D., & Koganezawa, M. (2013). Genes and circuits of courtship behaviour in drosophila males. *Nat. Rev. Neurosci.*, *14*, 681.

Carpenter, F. W. (1905). The reactions of the pomace fly (drosophila ampelophila loew) to light, gravity, and mechanical stimulation. *Am. Nat.*, *39*(459), 157.

Heisenberg, M., & Buchner, E. (1977). The role of retinula cell types in visual behavior of drosophilamelanogaster.

Miller, G. V., Hansen, K. N., & Stark, W. S. (1981). Phototaxis in drosophila: R1–6 input and interaction among ocellar and compound eye receptors. *J. Insect Physiol.*, *27*(11), 813.

Helfrich-Förster, C. (2020). Light input pathways to the circadian clock of insects with an emphasis on the fruit fly drosophila melanogaster. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.*, *206*(2), 259.

Gibson, J. J. (1950). *The perception of the visual world*. Houghton Mifflin.

Borst, A., & Haag, J. (2002). Neural networks in the cockpit of the fly. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.*, *188*(6), 419.

Mauss, A. S., & Borst, A. (2020). Optic flow-based course control in insects. *Curr. Opin. Neurobiol.*, *60*, 21.

Keleş, M. F., & Frye, M. A. (2017). Object-detecting neurons in drosophila. *Curr. Biol.*, *27*(5), 680.

Klapoetke, N. C., Nern, A., Rogers, E. M., Rubin, G. M., Reiser, M. B., & Card, G. M. (2022). A functionally ordered visual feature map in the drosophila brain. *Neuron*, *110*(10), 1700.

Chiappe, M. E. (2023). Circuits for self-motion estimation and walking control in drosophila. *Curr. Opin. Neurobiol.*, *81*(102748), 102748.

Turner, M. H., Krieger, A., Pang, M. M., & Clandinin, T. R. (2022). Visual and motor signatures of locomotion dynamically shape a population code for feature detection in drosophila. *Elife*, *11*.

Hassenstein, B., & Reichardt, W. (1956). Systemtheoretische analyse der zeit-, reihenfolgen-und vorzeichenauswertung bei der bewegungsperzeption des rüsselkäfers chlorophanus. *Z. Naturforsch. B J. Chem. Sci.*, *11*(9-10), 513.

Reichardt, W. (1961). Autocorrelation, a principle for evaluation of sensory information by the central nervous system. *Principles of sensory communications*.

Barlow, H. B., & Levick, W. R. (1965). The mechanism of directionally selective units in rabbit's retina. *J. Physiol.*, *178*(3), 477.

Borst, A. (2000). Models of motion detection. *Nat. Neurosci.*, *3*(11s), 1168.

Hardie, R. C. (1991). Whole-cell recordings of the light induced current in dissociated drosophila photoreceptors: Evidence for feedback by calcium permeating the light-sensitive channels. *Proc. Biol. Sci.*, *245*(1314), 203.

Ranganathan, R., Harris, G. L., Stevens, C. F., & Zuker, C. S. (1991). A drosophila mutant defective in extracellular calcium-dependent photoreceptor deactivation and rapid desensitization. *Nature*, *354*(6350), 230.

Peretz, A., Suss-Toby, E., Rom-Glas, A., Arnon, A., Payne, R., & Minke, B. (1994). The light response of drosophila photoreceptors is accompanied by an increase in cellular calcium: Effects of specific mutations. *Neuron*, *12*(6), 1257.

Nikolaev, A., Zheng, L., Wardill, T. J., O'Kane, C. J., de Polavieja, G. G., & Juusola, M. (2009). Network adaptation improves temporal representation of naturalistic stimuli in drosophila eye: II mechanisms. *PLoS One*, *4*(1), e4306.

Reiff, D. F., Plett, J., Mank, M., Griesbeck, O., & Borst, A. (2010). Visualizing retinotopic half-wave rectified input to the motion detection circuitry of drosophila. *Nat. Neurosci.*, *13*(8), 973.

Clark, D. A., Bursztyn, L., Horowitz, M. A., Schnitzer, M. J., & Clandinin, T. R. (2011). Defining the computational structure of the motion detector in drosophila. *Neuron*, *70*(6), 1165.

Freifeld, L., Clark, D. A., Schnitzer, M. J., Horowitz, M. A., & Clandinin, T. R. (2013). GABAergic lateral interactions tune the early stages of visual processing in drosophila. *Neuron*, *78*(6), 1075.

Silies, M., Gohl, D. M., Fisher, Y. E., Freifeld, L., Clark, D. A., & Clandinin, T. R. (2013). Modular use of peripheral input channels tunes motion-detecting circuitry. *Neuron*, *79*(1), 111.

Maisak, M. S., Haag, J., Ammer, G., Serbe, E., Meier, M., Leonhardt, A., Schilling, T., Bahl, A., Rubin, G. M., Nern, A., Dickson, B. J., Reiff, D. F., Hopp, E., & Borst, A. (2013). A directional tuning map of Drosophila elementary motion detectors. *Nature*, *500*(7461), 212.

Strother, J. A., Nern, A., & Reiser, M. B. (2014). Direct observation of ON and OFF pathways in the drosophila visual system. *Curr. Biol.*, *24*(9), 976.

Meier, M., Serbe, E., Maisak, M. S., Haag, J., Dickson, B. J., & Borst, A. (2014). Neural circuit components of the drosophila OFF motion vision pathway. *Curr. Biol.*, *24*(4), 385.

Behnia, R., Clark, D. A., Carter, A. G., Clandinin, T. R., & Desplan, C. (2014). Processing properties of ON and OFF pathways for drosophila motion detection. *Nature*, *512*(7515), 427.

Fisher, Y. E., Leong, J. C. S., Sporar, K., Ketkar, M. D., Gohl, D. M., Clandinin, T. R., & Silies, M. (2015). A Class of Visual Neurons with Wide-Field Properties Is Required for Local Motion Detection. *Curr. Biol.*, *25*(24), 3178.

Hardie, R. C., & Juusola, M. (2015). Phototransduction in drosophila. *Curr. Opin. Neurobiol.*, *34*, 37.

Leonhardt, A., Ammer, G., Meier, M., Serbe, E., Bahl, A., & Borst, A. (2016). Asymmetry of drosophila ON and OFF motion detectors enhances real-world velocity estimation. *Nat. Neurosci.*, *19*(5), 706.

Fisher, Y. E., Silies, M., & Clandinin, T. R. (2015). Orientation Selectivity Sharpens Motion Detection in Drosophila. *Neuron*, *88*(2), 390.

Yang, H. H., St-Pierre, F., Sun, X., Ding, X., Lin, M. Z., & Clandinin, T. R. (2016). Subcellular imaging of voltage and calcium signals reveals neural processing in vivo. *Cell*, *166*(1), 245.

Serbe, E., Meier, M., Leonhardt, A., & Borst, A. (2016). Comprehensive characterization of the major presynaptic elements to the drosophila OFF motion detector. *Neuron*, *89*(4), 829.

Strother, J. A., Wu, S.-T., Wong, A. M., Nern, A., Rogers, E. M., Le, J. Q., Rubin, G. M., & Reiser, M. B. (2017). The emergence of directional selectivity in the visual motion pathway of drosophila. *Neuron*, *94*(1), 168.

Meier, M., & Borst, A. (2019). Extreme Compartmentalization in a Drosophila Amacrine Cell. *Curr. Biol.*, *29*(9), 1545.

Gruntman, E., Romani, S., & Reiser, M. B. (2019). The computation of directional selectivity in the drosophila OFF motion pathway. *Elife*, *8*, e50706.

Drews, M. S., Leonhardt, A., Pirogova, N., Richter, F. G., Schuetzenberger, A., Braun, L., Serbe, E., & Borst, A. (2020). Dynamic signal compression for robust motion vision in flies. *Curr. Biol.*, *30*(2), 209.

Matulis, C. A., Chen, J., Gonzalez-Suarez, A. D., Behnia, R., & Clark, D. A. (2020). Heterogeneous temporal contrast adaptation in drosophila Direction-Selective circuits. *Curr. Biol.*, *30*(2), 222.

Ramos-Traslosheros, G., & Silies, M. (2021a, Apr.). *The physiological basis for the computation of direction selectivity in the Drosophila OFF pathway* (tech. rep.). Neuroscience.

Gruntman, E., Reimers, P., Romani, S., & Reiser, M. B. (2021). Non-preferred contrast responses in the drosophila motion pathways reveal a receptive field structure that explains a common visual illusion. *Curr. Biol.*, *31*(23), 5286.

Ketkar, M. D., Gür, B., Molina-Obando, S., Ioannidou, M., Martelli, C., & Silies, M. (2022). First-order visual interneurons distribute distinct contrast and luminance information across ON and OFF pathways to achieve stable behavior. *Elife*, *11*, e74937.

Braun, A., Borst, A., & Meier, M. (2023). Disynaptic inhibition shapes tuning of OFF-motion detectors in drosophila. *Curr. Biol.*, *33*(11), 2260.

Ammer, G., Serbe-Kamp, E., Mauss, A. S., Richter, F. G., Fendl, S., & Borst, A. (2023). Multilevel visual motion opponency in drosophila. *Nat. Neurosci.*, *26*(11), 1894.

Pirogova, N., & Borst, A. (2023). Contrast normalization affects response time-course of visual interneurons. *PLoS One*, *18*(6), e0285686.

Gür, B., Ramirez, L., Cornean, J., Thurn, F., Molina-Obando, S., Ramos-Traslosheros, G., & Silies, M. (2024). Neural pathways and computations that achieve stable contrast processing tuned to natural scenes. *Nat. Commun.*, *15*(1), 8580.

Pang, M. M., Chen, F., Xie, M., Druckmann, S., Clandinin, T. R., & Yang, H. H. (2025). A recurrent neural circuit in drosophila temporally sharpens visual inputs. *Curr. Biol.*, *35*(2), 333.

Currier, T. A., & Clandinin, T. R. (2025). Infrequent strong connections constrain connectomic predictions of neuronal function. *Cell*.

Hardie, R. C., & Raghu, P. (2001). Visual transduction in drosophila. *Nature*, *413*(6852), 186.

Yamaguchi, S., Wolf, R., Desplan, C., & Heisenberg, M. (2008). Motion vision is independent of color in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.*, *105*(12), 4910.

Morante, J., & Desplan, C. (2008). The color-vision circuit in the medulla of drosophila. *Curr. Biol.*, *18*(8), 553.

Yamaguchi, S., Desplan, C., & Heisenberg, M. (2010). Contribution of photoreceptor subtypes to spectral wavelength preference in *Drosophila*. *Proc. Natl. Acad. Sci. U. S. A.*, *107*(12), 5634.

Wardill, T. J., List, O., Li, X., Dongre, S., McCulloch, M., Ting, C.-Y., O'Kane, C. J., Tang, S., Lee, C.-H., & Hardie, R. C. (2012). Multiple spectral inputs improve motion discrimination in the drosophila visual system. *Science*, *336*(6083), 925.

Schnaitmann, C., Garbers, C., Wachtler, T., & Tanimoto, H. (2013). Color discrimination with broadband photoreceptors. *Curr. Biol.*, *23*(23), 2375.

Longden, K. D., Rogers, E. M., Nern, A., Dionne, H., & Reiser, M. B. (2023). Different spectral sensitivities of ON- and OFF-motion pathways enhance the detection of approaching color objects in drosophila. *Nat. Commun.*, *14*(1), 7693.

Juusola, M., & Hardie, R. C. (2001). Light adaptation in drosophila photoreceptors: I. response dynamics and signaling efficiency at 25 degrees C. *J. Gen. Physiol.*, *117*(1), 3.

de Ruyter van Steveninck, R. R., & Laughlin, S. B. (1996). The rate of information transfer at graded-potential synapses. *Nature*, *379*(6566), 642.

Zheng, L., Nikolaev, A., Wardill, T. J., O'Kane, C. J., de Polavieja, G. G., & Juusola, M. (2009). Network adaptation improves temporal representation of naturalistic stimuli in drosophila eye: I dynamics. *PLoS One*, *4*(1), e4307.

Joesch, M., Schnell, B., Raghu, S. V., Reiff, D. F., & Borst, A. (2010). ON and OFF pathways in drosophila motion vision. *Nature*, *468*(7321), 300.

Gjorgjieva, J., Sompolinsky, H., & Meister, M. (2014). Benefits of pathway splitting in sensory coding. *Journal of Neuroscience*, *34*(36), 12127.

Borst, A., Haag, J., & Mauss, A. S. (2020). How fly neurons compute the direction of visual motion. *Journal of Comparative Physiology A*, *206*(2), 109.

Jain, V. (2023). How AI could lead to a better understanding of the brain. *Nature*, *623*(7986), 247.

Liu, T. X., Davoudian, P. A., Lizbinski, K. M., & Jeanne, J. M. (2022). Connectomic features underlying diverse synaptic connection strengths and subcellular computation. *Curr. Biol.*, *32*(3), 559.

Götz, K. G. (1964). Optomotorische untersuchung des visuellen systems einiger augenmutanten der fruchtfliege drosophila. *Kybernetik*, *2*(2), 77.

Ammer, G., Leonhardt, A., Bahl, A., Dickson, B. J., & Borst, A. (2015). Functional Specialization of Neural Input Elements to the Drosophila on Motion Detector. *Curr. Biol.*, *25*(17), 2247.

Strother, J. A., Wu, S.-T., Rogers, E. M., Eliason, J. L. M., Wong, A. M., Nern, A., & Reiser, M. B. (2018). Behavioral state modulates the ON visual motion pathway of drosophila. *Proceedings of the National Academy of Sciences*, *115*(1), E102.

Özdil, P. G., Arreguit, J., Scherrer, C., Ijspeert, A., & Ramdya, P. (2024). Centralized brain networks underlie body part coordination during grooming. *bioRxiv*, 2024.12. 17.628844.

Wang-Chen, S., Stimpfling, V. A., Lam, T. K. C., Özdil, P. G., Genoud, L., Hurtak, F., & Ramdya, P. (2024). NeuroMechFly v2: Simulating embodied sensorimotor control in adult drosophila. *Nat. Methods*, 1.

Vaxenburg, R., Siwanowicz, I., Merel, J., Robie, A. A., Morrow, C., Novati, G., Stefanidi, Z., Both, G.-J., Card, G. M., Reiser, M. B., Botvinick, M. M., Branson, K. M., Tassa, Y., & Turaga, S. C. (2025). Whole-body physics simulation of fruit fly locomotion. *Nature*, 1.

Duan, S., Dong, L. L., & Fiete, I. (2025). From synapses to dynamics: Obtaining function from structure in a connectome constrained model of the head direction circuit. *bioRxiv*, 2025.05.26.655406.

Shiu, P. K., Sterne, G. R., Spiller, N., Franconville, R., Sandoval, A., Zhou, J., Simha, N., Kang, C. H., Yu, S., Kim, J. S., Dorkenwald, S., Matsliah, A., Schlegel, P., Yu, S.-C., McKellar, C. E., Sterling, A., Costa, M., Eichler, K., Bates, A. S., … Scott, K. (2024). A drosophila computational brain model reveals sensorimotor processing. *Nature*, *634*(8032), 210.

Allen, A. M., Neville, M. C., Birtles, S., Croset, V., Treiber, C. D., Waddell, S., & Goodwin, S. F. (2020). A single-cell transcriptomic atlas of the adult drosophila ventral nerve cord. *Elife*, *9*.

Nern, A., Lösche, F., Takemura, S.-Y., Burnett, L. E., Dreher, M., Gruntman, E., Hoeller, J., Huang, G. B., Januszewski, M., Klapoetke, N. C., Koskela, S., Longden, K. D., Lu, Z., Preibisch, S., Qiu, W., Rogers, E. M., Seenivasan, P., Zhao, A., Bogovic, J., … Reiser, M. B. (2024). Connectome-driven neural inventory of a complete visual system. *bioRxivorg*, 2024.04.16.589741.

Seung, H. S. (2025). Multi-output computation by single neuron biophysics in a visual system. *bioRxiv*, 2025.02.02.636137.

Zhao, A., Gruntman, E., Nern, A., Iyer, N., Rogers, E., Koskela, S., Dreher, M., Siwanowicz, I., Flynn, M., Laughland, C., Henrique, L., Thomson, A., Moran, C., Gezahgen, B., Erginkaya, M., Parekh, R., Hoeller, J., Romani, S., Chiappe, E., … Reiser, M. (2024). Connecting the eye to the brain: Retina structure shapes neuron function in drosophila motion vision. *2024 APS March Meeting*.

Destexhe, A., & Marder, E. (2004). Plasticity in single neuron and circuit computations. *Nature*, *431*(7010), 789.

Scott, D. N., & Frank, M. J. (2023). Adaptive control of synaptic plasticity integrates micro- and macroscopic network function. *Neuropsychopharmacology*, *48*(1), 121.

Zhao, B., Walters, R., & Yu, R. (2025). Symmetry in neural network parameter spaces. *arXiv [cs.LG]*.

Euler, T., Detwiler, P., & Denk, W. (2002). Directionally selective calcium signals in dendrites of starburst amacrine cells. *Nature*, *418*, 845.

Rall, W. (1989). Cable theory for dendritic neurons, 9.

Hines, M. L., & Carnevale, N. T. (1997). The NEURON simulation environment. *Neural Comput.*, *9*(6), 1179.

Goodman, D., & Brette, R. (2008). Brian: A simulator for spiking neural networks in python. *Front. Neuroinform.*, *2*, 5.

Gewaltig, M., & Diesmann, M. (2007). NEST (NEural simulation tool). *Scholarpedia J.*, *2*, 1430.

Doya, K., Selverston, A., & Rowat, P. (1993). A hodgkin-huxley type neuron model that learns slow non-spike oscillations. *Neural Inf Process Syst*, *6*, 566.

Deistler, M., Kadhim, K. L., Pals, M., Beck, J., Huang, Z., Gloeckler, M., Lappalainen, J. K., Schröder, C., Berens, P., Gonçalves, P. J., & Macke, J. H. (2024). Differentiable simulation enables large-scale training of detailed biophysical models of neural dynamics. *bioRxiv*, 2024.08.21.608979.

Ulmer, L. (2024, Mar.). *Integrating knowledge of neural tuning into connectome-constrained and task-optimized models* [Master's thesis, Eberhard Karls Universität Tübingen, Mathematisch-Naturwissenschaftliche Fakultät, Wilhelm-Schickard-Institut für Informatik].

French, R. (1999). Catastrophic forgetting in connectionist networks. *Trends Cogn. Sci.*, *3*, 128.

Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. *In Proceedings of the International*

Conference on Artificial Intelligence and Statistics (AISTATS'10). Society for Artificial Intelligence and Statistics.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *Proceedings of the IEEE international conference on computer vision*, 1026.

Strogatz, S. H. (2001). Exploring complex networks. *Nature*, *410*(6825), 268.

Munson, J. C., & Elbaum, S. G. (2002). Code churn: A measure for estimating the impact of code change. *Proceedings. International Conference on Software Maintenance (Cat. No. 98CB36272)*, 24.

Hoyer, S., & Hamman, J. (2017). Xarray: N-D labeled arrays and datasets in python. *J. Open Res. Softw.*, *5*(1), 10.

Joblib Development Team. (2020). Joblib: Running python functions as pipeline jobs [Accessed: 2025-6-30].

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.*, *12*, 2825.

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, *9*(3), 90.

Scheffer, L. K. (2021). The physical design of biological systems - insights from the fly brain. *Proceedings of the 2021 International Symposium on Physical Design*, 101.

Jacobs, R. A., & Bates, C. J. (2019). Comparing the visual representations and performance of humans and deep neural networks. *Curr. Dir. Psychol. Sci.*, *28*(1), 34.

Wienecke, C. F. R., Leong, J. C. S., & Clandinin, T. R. (2018). Linear summation underlies direction selectivity in drosophila. *Neuron*, *99*(4), 680.

Mishra, A., Serbe-Kamp, E., Borst, A., & Haag, J. (2023). Voltage to calcium transformation enhances direction selectivity in drosophila T4 neurons. *J. Neurosci.*, *43*(14), 2497.

Ramos-Traslosheros, G., & Silies, M. (2021b). The physiological basis for contrast opponency in motion computation in drosophila. *Nat. Commun.*, *12*(1), 4987.

Mazurek, M., Kager, M., & Van Hooser, S. D. (2014). Robust quantification of orientation selectivity and direction selectivity. *Front. Neural Circuits*, *8*.

McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv [stat.ML]*.

Maaten, L., & Hinton, G. E. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, *9*, 2579.

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, *313*(5786), 504.

Schölkopf, B., Smola, A., & Müller, K.-R. (1998). Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput.*, *10*(5), 1299.

Şimşek, B., Ged, F., Jacot, A., Spadaro, F., Hongler, C., Gerstner, W., & Brea, J. (2021). Geometry of the loss landscape in overparameterized neural networks: Symmetries and invariances. *arXiv [cs.LG]*.

Martinelli, F., Simsek, B., Brea, J., & Gerstner, W. (2023). Expand-and-cluster: Exact parameter recovery of neural networks. *arXiv [cs.NE]*.

von Bachmann, P. (2023, Apr.). *Uncertainty estimation in connectome-constrained neural networks using deep ensembles* [Master's thesis, Eberhard Karls Universität Tübingen, Mathematisch-Naturwissenschaftliche Fakultät, Wilhelm-Schickard-Institut für Informatik].

Eschenhagen, R., Daxberger, E., Hennig, P., & Kristiadi, A. (2021). Mixtures of laplace approximations for improved post-hoc uncertainty in deep learning. *arXiv [cs.LG]*.

Mano, O., Creamer, M. S., Badwan, B. A., & Clark, D. A. (2021). Predicting individual neuron responses with anatomically constrained task optimization. *Curr. Biol.*, *31*(18), 4062.

Tschopp, F. D., Reiser, M. B., & Turaga, S. C. (2018). A Connectome Based Hexagonal Lattice Convolutional Network Model of the Drosophila Visual System. *arXiv preprint arXiv:1806. 04793*.

Zhou, B., Li, Z., Kim, S., Lafferty, J., & Clark, D. A. (2022). Shallow neural networks trained to detect collisions recover features of visual loom-selective neurons. *Elife*, *11*, e72067.

Warrington, A., Spencer, A., & Wood, F. (2019). The virtual patch clamp: Imputing C. elegans membrane potentials from calcium imaging. *arXiv [q-bio.NC]*.

Mi, L., Xu, R., Prakhya, S., Lin, A., Shavit, N., Samuel, A., & Turaga, S. C. (2022). Connectome-constrained latent variable model of whole-brain neural activity.

Beiran, M., & Litwin-Kumar, A. (2024). Prediction of neural activity in connectome-constrained recurrent networks. *bioRxivorg*, 2024.02.22.581667.

Cowley, B. R., Calhoun, A. J., Rangarajan, N., Ireland, E., Turner, M. H., Pillow, J. W., & Murthy, M. (2024). Mapping model units to visual neurons reveals population code for social behaviour. *Nature*, *629*(8014), 1100.

Zhao, A., Gruntman, E., Nern, A., Iyer, N. A., Rogers, E. M., Koskela, S., Siwanowicz, I., Dreher, M., Flynn, M. A., Laughland, C. W., Ludwig, H. D. F., Thomson, A. G., Moran, C. P., Gezahegn, B., Bock, D. D., & Reiser, M. B. (2022). Eye structure shapes neuron function in*Drosophila*motion vision. *bioRxiv*, 2022.12.14.520178.

Millward, B., Maddock, S., & Mangan, M. (2022). CompoundRay, an open-source tool for high-speed and high-fidelity rendering of compound eyes. *Elife*, *11*.

Gonzalez-Bellido, P. T., Wardill, T. J., & Juusola, M. (2011). Compound eyes and retinal information processing in miniature dipteran species match their specific ecological demands. *Proc. Natl. Acad. Sci. U. S. A.*, *108*(10), 4224.

Fenk, L. M., Avritzer, S. C., Weisman, J. L., Nair, A., Randt, L. D., Mohren, T. L., Siwanowicz, I., & Maimon, G. (2022). Muscles that move the retina augment compound eye vision in drosophila. *Nature*, *612*(7938), 116.

Henning, M., Ramos-Traslosheros, G., Gür, B., & Silies, M. (2022). Populations of local direction-selective cells encode global motion patterns generated by self-motion. *Sci Adv*, *8*(3), eabi7112.

Wernet, M. F., Velez, M. M., Clark, D. A., Baumann-Klausener, F., Brown, J. R., Klovstad, M., Labhart, T., & Clandinin, T. R. (2012). Genetic dissection reveals two separate retinal substrates for polarization vision in drosophila. *Curr. Biol.*, *22*(1), 12.

Hige, T., Aso, Y., Modi, M. N., Rubin, G. M., & Turner, G. C. (2015). Heterosynaptic plasticity underlies aversive olfactory learning in drosophila. *Neuron*, *88*(5), 985.

Baltruschat, L., Prisco, L., Ranft, P., Lauritzen, J. S., Fiala, A., Bock, D. D., & Tavosanis, G. (2021). Circuit reorganization in the drosophila mushroom body calyx accompanies memory consolidation. *Cell Rep.*, *34*(11), 108871.

Bushey, D., Tononi, G., & Cirelli, C. (2011). Sleep and synaptic homeostasis: Structural evidence in drosophila. *Science*, *332*(6037), 1576.

Sugie, A., Marchetti, G., & Tavosanis, G. (2018). Structural aspects of plasticity in the nervous system of drosophila. *Neural Dev.*, *13*(1), 14.

Grabowska, M. J., Jeans, R., Steeves, J., & van Swinderen, B. (2020). Oscillations in the central brain of drosophila are phase locked to attended visual features. *Proc. Natl. Acad. Sci. U. S. A.*, *117*(47), 29925.

Lappalainen, J., Herpich, J., & Tetzlaff, C. (2019). A theoretical framework to derive simple, firing-rate-dependent mathematical models of synaptic plasticity. *Front. Comput. Neurosci.*, *13*, 26.

Confavreux, B., Ramesh, P., Gonçalves, P. J., Macke, J., & Vogels, T. (2023). Meta-learning families of plasticity rules in recurrent spiking networks using simulation-based inference. *Neural Inf Process Syst*, *36*, 13545.

Herz, A. V. M., Gollisch, T., Machens, C. K., & Jaeger, D. (2006). Modeling single-neuron dynamics and computations: A balance of detail and abstraction. *Science*, *314*(5796), 80.

Schröder, C., & Macke, J. H. (2023). Simultaneous identification of models and parameters of scientific simulators. *arXiv [cs.LG]*.

Hadsell, R., Chopra, S., & LeCun, Y. (2006). Dimensionality reduction by learning an invariant mapping. *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2 (CVPR'06)*, *2*, 1735.

Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). A simple framework for contrastive learning of visual representations. *arXiv [cs.LG]*.

He, K., Fan, H., Wu, Y., Xie, S., & Girshick, R. (2019). Momentum contrast for unsupervised visual representation learning. *arXiv [cs.CV]*.

Zbontar, J., Jing, L., Misra, I., LeCun, Y., & Deny, S. (2021). Barlow twins: Self-supervised learning via redundancy reduction. *International Conference on Machine Learning*, 12310.

van den Oord, A., Li, Y., & Vinyals, O. (2018). Representation learning with contrastive predictive coding. *arXiv [cs.LG]*.

Yerxa, T. E., Kuang, Y., Simoncelli, E. P., & Chung, S. (2023). Learning efficient coding of natural images with maximum manifold capacity representations (A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, & S. Levine, Eds.). *Neural Inf Process Syst*, *abs/2303.03307*, 24103.

Nejad, K. K., Anastasiades, P., Hertäg, L., & Costa, R. P. (2025). Self-supervised predictive learning accounts for cortical layer-specificity. *Nat. Commun.*, *16*(1), 6178.

Advani, M. S., & Saxe, A. M. (2017). High-dimensional dynamics of generalization error in neural networks. *arXiv [stat.ML]*.

Neyshabur, B., Li, Z., Bhojanapalli, S., LeCun, Y., & Srebro, N. (2018). The role of over-parametrization in generalization of neural networks. *International Conference on Learning Representations*.

Martinelli, F., Simsek, B., Gerstner, W., & Brea, J. (2023). Expand-and-cluster: Parameter recovery of neural networks. *arXiv [cs.NE]*, 34895.

Wilson, A. G. (2025). Deep learning is not so mysterious or different. *arXiv [cs.LG]*.

Eshraghian, J. K., Ward, M., Neftci, E. O., Wang, X., Lenz, G., Dwivedi, G., Bennamoun, M., Jeong, D. S., & Lu, W. D. (2023). Training spiking neural networks using lessons from deep learning. *Proc. IEEE Inst. Electr. Electron. Eng.*, *111*(9), 1016.

Newman, S. (2022, Jan.). *Building microservices: Designing fine-grained systems*. Ascent Audio.

Pospisil, D. A., Aragon, M. J., Dorkenwald, S., Matsliah, A., Sterling, A. R., Schlegel, P., Yu, S.-C., McKellar, C. E., Costa, M., Eichler, K., Jefferis, G. S. X. E., Murthy, M., & Pillow, J. W. (2024). The fly connectome reveals a path to the effectome. *Nature*, *634*(8032), 201.

Seguin, C., Puxeddu, M. G., Faskowitz, J., Betzel, R. F., & Sporns, O. (2025). Connectome architecture favours within-module diffusion and between-module routing. *bioRxiv*.