An Optimization framework for Nonspiking Neuronal Networks and Aided Discovery of a Model for the Elementary Motion Detector

Arunava Banerjee

Computer and Information Science and Engineering, University of Florida,

PO Box 116120, Gainesville, FL 32611, USA

E-mail: arunava@ufl.edu

Abstract

We present a general optimization procedure that given a parameterized network of nonspiking conductance based compartmentally modeled neurons, tunes the parameters to elicit a desired network behavior. Armed with this tool, we address the elementary motion detector problem. Central to established theoretical models, the Hassenstein-Reichardt and Barlow-Levick detectors, are delay lines whose outputs from spatially separated locations are prescribed to be nonlinearly integrated with the direct outputs to engender direction selectivity. The neural implementation of the delays—which are substantial as stipulated by interomatidial angles—has remained elusive although there is consensus regarding the neurons that constitute the network. Assisted by the optimization procedure, we identify parameter settings consistent with the connectivity architecture and physiology of the Drosophila optic lobe, that demonstrates that the requisite delay and the con-

comitant direction selectivity can emerge from the nonlinear dynamics of small recurrent networks of neurons with simple tonically active synapses. Additionally, although the temporally extended responses of the neurons permit simple synaptic integration of their signals to be sufficient to induce direction selectivity, both preferred direction enhancement and null direction suppression is necessary to abridge the overall response. Finally, the characteristics of the response to drifting sinusoidal gratings are readily explained by the charging-up of the recurrent networks and their low-pass nature.

Introduction

A major goal of Neuroscience is to understand how the activity of identified neural circuits relate to behavior. Recent advances in (semi)automated reconstruction from EM data and immunolabeling have generated a wealth of information regarding the neural connectivity architecture and polarity of synapses in the brains of model organisms. For example, [33] has released an EM volume of the complete Drosophila brain, parts of which have been reconstructed for neural connectivity. These notable advances have, however, not been followed by the anticipated spate of neural implementation solutions to well characterized high level operations. This discrepancy stems from the fact that the dynamics of a neural circuit is determined not only by connectivity and synaptic polarities, but also by the synaptic gain profiles, information which the current techniques do not reveal at the necessary level of detail. Lacking synaptic profile information, relating circuits to behavior has therefore been difficult.

The observation that network behavior is influenced by synaptic profiles also implies that it is, in principle, possible to infer synaptic profiles from the recorded behavior of a network. Such a data driven approach, where a parameterized model is adapted to the characteristics of a dataset, has been the cornerstone of recent successes in Machine Learning. The general

problem in the current context, where a network comprises both spiking and nonspiking neurons, is likely both formally and computationally intractable. However, as we demonstrate in this article, a more restricted class, that of nonspiking neuronal networks, lends itself to this approach. Specifically, we show that local analysis of the trajectories of the corresponding dynamical system is tractable: in operational terms, given a parameterized network of compartmentally modeled conductance based nonspiking neurons interacting with tonically active synapses, computational optimization can tune the parameters to cause the network's behavior to display desired properties. The optimization procedure developed is general and applies to a large class of parameters, including synaptic profiles as well as morphological properties of the constituent neurons.

Nonspiking neuronal networks abound in the sensory and motor periphery, the set of circuits composed of photoreceptors, horizontal, bipolar, and amacrine cells in the vertebrate retina [10] being among the most widely cited examples. In order to showcase the strength of the optimization driven approach, we have however chosen a network whose (a) architecture is conserved across multiple insect species, (b) has been characterized in substantial detail, and (c) has well established theoretical models whose relationship to the neural implementation remains to be reconciled. In resolving this problem, we show that computational optimization of parameterized synaptic profiles to fit a network's behavior can be a potent tool when combined with connectivity and synaptic polarity information.

The elementary motion detector (EMD) in the fly brain is a paradigmatic neural computation that has been the subject of intense investigation over decades [6]. Established theoretical models, the Hassenstein-Reichardt (HR) [15] and the Barlow-Levick (BL) [4] detectors prescribe delay lines, the neural implementation of which has thus far remained elusive. Much is however known about the neural circuits that implement the EMD, that we summarized here.

The compound eyes of flies consist of anatomically identical units, called omatidia, laid

out in a hexagonal lattice. Visual information processing begins at the photoreceptors in the omatidium, advancing thereafter through neurons in four retinotopically organized neuropile, the lamina, medulla, lobula, and lobula plate. Broad interest in fly motion vision, stemming from its status as a canonical computation, has led to the accumulation of a wealth of data—particularly with regard to the Drosophila—concerning the connectivity architecture and physiology of the neurons in the repeating modules associated with each omatidium, in the lamina [20, 30] as well as the medulla and lobula [29, 26]. Briefly, axons of photoreceptors R1-R6 innervate lamina monopolar cells (LMC) L1-L3 in the corresponding lamina module. Of the cells identified in the lamina, only the joint silencing of the similarly responding [8] L1 and L2 abolishes direction selectivity [30]. Exiting the lamina, motion information is extracted in parallel pathways with L1 feeding the brightness increment (ON) and L2 feeding the brightness decrement (OFF) circuitry [16]. The connectome of the ON module in the medulla has been elucidated in substantial detail [29]. Based on the preponderance of different synaptic contacts, the following core circuit emerges (Fig.1a). The primary targets of L1 are Mi1, Tm3, L5, and C3. L5 drives Mi4, C3 drives Mi9, and Mi4 and Mi9 are reciprocally connected. Lastly, Mi1, Tm3, Mi4, and Mi9 constitute the primary inputs to T4 which is the first cell on the pathway to exhibit direction selectivity. Although Fig.6 in [29] contains additional cells and synaptic contacts, our results demonstrate that this parsimonious circuit is sufficient to manifest direction selectivity, with Mi1/Tm3 providing the direct input and Mi4/Mi9 providing the delayed input into T4. The narrow receptive field T4 comes in four flavors: T4a-d each tuned to one of four cardinal directions. Whole-cell recordings [5] indicate that all of the noted cells are graded potential neurons, and therefore communicate using tonically active synapses [17]. The connectome of the OFF module exhibits strong parallels [26].

We chose to computationally model the ON pathway because data pertaining to the dimensions of neurites and the polarity of synapses is largely available in this case [29]. The principles

we have discovered, however, apply to the OFF pathway as well.

Results

Parameterization of Network and Intractability of Global Analysis

We begin with the standard conductance based compartmental modeling methodology, where each neuron is partitioned into equipotential segments whose terminals are then linked to assemble the network (Fig.1c). Each compartment i is modeled using the differential equation,

$$\tau^i \frac{dV^i}{dt} = (E_l^i - V^i) + 2(V_{in}^i - V^i)\pi^i + 2(V_{out}^i - V^i)\pi^i + (E_{syn}^i - V^i)\rho^i \tag{1}$$

where V^i, V_{in}^i and V_{out}^i denote the time varying membrane potential at the middle and the two terminals of a compartment, and E_l^i, E_{syn}^i denote the constant leak and constant synaptic potentials. Also, $\tau^i = C^i/g_l^i, \pi^i = g_a^i/g_l^i$, and $\rho^i = g_{syn}^i/g_l^i$, where C^i, g_l^i, g_a^i and g_{syn}^i denote the constant capacitance, constant leak, constant axial, and time varying synaptic conductances. The scaling by $(1/g_l^i)$ makes π^i and ρ^i dimensionless quantities, which in turn aid in the interpretability of the equation.

Compartments are linked as mandated by the modeled cellular morphology of the neurons (Fig.1a and c). Each terminal of each compartment satisfied one current balance algebraic constraint: if the terminal corresponding to V_{in}^* of compartment $*$ is linked to the terminals corresponding to V_{out}^i of compartments $i = 1 \ldots m$, then setting $V_{out}^i = V_{in}^*$ for all $i = 1 \ldots m$, we have

$$V_{in}^* = (V^*\pi^* g_l^* + \sum_{i=1}^{m} V^i \pi^i g_l^i)/(\pi^* g_l^* + \sum_{i=1}^{m} \pi^i g_l^i) \tag{2}$$

and likewise, if the terminal corresponding to V_{out}^* of compartment $*$ is linked to the terminals corresponding to V_{in}^i of compartments $i = 1 \ldots m$, then setting $V_{in}^i = V_{out}^*$ for all

$i = 1 \ldots m$, we have

$$V^*_{out} = (V^* \pi^* g^*_l + \sum_{i=1}^{m} V^i \pi^i g^i_l)/(\pi^* g^*_l + \sum_{i=1}^{m} \pi^i g^i_l) \tag{3}$$

The V^i_{in} and V^i_{out} in Eq. 1 can be eliminated by plugging in Eq. 2 and Eq. 3, resulting in a system of ordinary differential equations (ODE) involving only the V^i of each compartment.

It is instructive at this juncture to note that the simplest form this system of equations can take is when the time varying trajectories of the synaptic drives, ρ^i, are predetermined independent of the V^is, such as in a feedforward network where the dynamics of the system is solved layer by layer. Even so, what results is a nonhomogeneous linear time-varying system of ODEs of the form $\mathbf{V}' = \mathbf{A}(t)\mathbf{V} + \mathbf{b}(t)$, where $\mathbf{V}'$ is the vector of time derivatives of the V^i of the compartments, $\mathbf{A}(t)$ is a time varying matrix whose elements are composed of π^i, g^i_l, and the time varying ρ^i, with the structure of $\mathbf{A}(t)$ determined by the connectivity between the compartments, and $\mathbf{b}(t)$ is a time varying vector with elements composed of E^i_l, E^i_{syn} and ρ^i.

The general solution to $\mathbf{V}' = \mathbf{A}(t)\mathbf{V} + \mathbf{b}(t)$ is achieved [23] by solving for the state transition matrix $\mathbf{\Phi}(t, \tau)$ for the homogeneous system $\mathbf{V}' = \mathbf{A}(t)\mathbf{V}$ using the Peano-Baker series

$$\mathbf{\Phi}(t, \tau) = \mathbf{I} + \int_{\tau}^{t} \mathbf{A}(\sigma_1) d\sigma_1 + \int_{\tau}^{t} \mathbf{A}(\sigma_1) \int_{\tau}^{\sigma_1} \mathbf{A}(\sigma_2) d\sigma_2 d\sigma_1 + \ldots \tag{4}$$

and then incorporating $\mathbf{b}(t)$ to arrive at the solution for initial condition $\mathbf{V}(t_0)$ given by

$$\mathbf{V}(t) = \mathbf{\Phi}(t, t_0)\mathbf{V}(t_0) + \int_{t_0}^{t} \mathbf{\Phi}(t, \tau)\mathbf{b}(\tau) d\tau \tag{5}$$

In cases where the fundamental matrix $\mathbf{\Psi}(t)$, satisfying the matrix ODE $\mathbf{\Psi}'(t) = \mathbf{A}(t)\mathbf{\Psi}(t)$ can be computed, $\mathbf{\Phi}(t, \tau)$ can be solved without resorting to the Peano-Baker infinite series, as $\mathbf{\Phi}(t, \tau) = \mathbf{\Psi}(t)\mathbf{\Psi}^{-1}(\tau)$. With the reasonable assumption that $\mathbf{A}(t)$ is periodic with a suitably long period, solving for $\mathbf{\Psi}(t)$ requires the Floquet decomposition of $\mathbf{A}(t)$. Unfortunately, this

approach too eludes analytic solutions even for simple periodic parametric functions of time, ρ^i, since it involves matrix exponentiation as a function of the parameters [23], among other reasons.

Finally, in the case of recurrent networks, the study of which is our primary objective, the ρ^is can not be predetermined independent of the V^is; each ρ^i is a nonlinear function of a V^j, the membrane potential of the appropriate compartment of the presynaptic neuron. The immediate consequence is that the system of ODEs is now nonlinear and none of the above approaches apply. Recognizing then that global analysis of the dynamics of the network is intractable, that is, parameterized closed form solutions may not exist, we are left with the possibility of conducting a local analysis. To elaborate, given a solution trajectory $\mathbf{V}(t)$ of the compartments for some setting of all the parameters, is it possible to characterize solution trajectories in a local neighborhood of those parameter settings? That this is feasible is shown in the next section.

Returning to the network parameterization, what remains is the specification of the model for the time varying relative synaptic conductances $g^i_{syn}/g^i_l = \rho^i$ as a function of the presynaptic potential, i.e., $\rho^i(V_{pre})$, where V_{pre} denotes the V^j of the presynaptic compartment. Noting that the synapses under consideration are tonically active, the time scales at which the synapses operate vis-à-vis the time scale of the network phenomenon under consideration, dictate the choice of the synaptic model. We consider two models. The first is where the network phenomenon operates at a time scale much larger than that of the synapse, so that the synapse can be modeled to be essentially instantaneous. The second is where their time scales are comparable, in which case the synaptic conductance is modeled using an ODE.

In the first model, $\rho^i(V_{pre})$ is parameterized as a monotonically increasing and saturating function of the instantaneous potential V_{pre} of the appropriate presynaptic compartment:

$$\rho^i(V_{pre}) = \frac{\mu^i}{1 + e^{-(\nu^i V_{pre} + \eta^i)}} \tag{6}$$

where the parameters μ^i, ν^i, and η^i, set to be strictly positive, determine the gain, sensitivity, and baseline relative conductance (Fig.1d). The second model specifies the time varying synaptic conductance via standard receptor channel kinetics [9]:

$$\rho^i = \gamma^i P^i \quad and \quad \frac{dP^i}{dt} = \alpha^i(V_{pre})(1 - P^i) - \beta^i P^i \tag{7}$$

where γ^i denotes the maximum relative conductance, P^i is the time varying fraction of open channels, β^i, modeling the closing rate of channels, is a positive constant, and α^i, modeling neurotransmitter concentration, is again a monotonically increasing and saturating function of V_{pre} set as $\alpha^i(V_{pre}) = \mu^i/(1 + e^{-(\nu^i V_{pre} + \eta^i)})$.

The entire network is thus specified by the set of parameters $\{\pi^i, g_l^i, \mu^i, \nu^i, \eta^i\}$ (for the first synaptic model Eq. 6) or $\{\pi^i, g_l^i, \gamma^i, \beta^i, \mu^i, \nu^i, \eta^i\}$ (for the second synaptic model Eq. 7), with i ranging over all compartments. We emphasize that the tractability of the upcoming local analysis hinges on each parameter being a scalar quantity so that the entire network is parameterized as a vector in a finite dimensional Euclidean space. In addition, although the analysis applies to all parameters, the morphological parameters, π^i and g_l^i, may be fixed based on knowledge of the morphology of the neurons in the circuit under consideration. In the case of the EMD, these were computed for each cylindrical compartment of diameter d^i and length l^i using reported values in [22, 29], and universal membrane constants [12] (Methods). τ, being invariant to d and l, had the value $20ms$ at all compartments. We also set $E_l = -55mV$, and $E_{syn} = -85mV$ for hyperpolarizing and $E_{syn} = 0mV$ for depolarizing synapses.

The input drive into a network is specified as time varying trajectories of presynaptic potentials at the input synapses of the network. Given initial conditions $V^i(0)$ (and $P^i(0)$ for the second synaptic model) for all compartments i, the entire differential-algebraic system of equations traces out a trajectory for $V^i(t)$, $V_{in}^i(t)$, $V_{out}^i(t)$ (and $P^i(t)$). We restrict initial conditions to respective equilibrium resting potentials (and equilibrium resting fraction of open chan-

nels for synapses), thereby eliminating them as free parameters, to focus the analysis solely on the network parameterization. Assuming baseline input drive to be a canonical constant over all time, the network is at its equilibrium resting state when $dV^i/dt = 0$ for all compartments and $dP^i/dt = 0$ for all synapses. Setting $dV^i/dt = dP^i/dt = 0$ for all i turns the differential-algebraic system of equations Eq. 1, Eq. 2, Eq. 3, and Eq. 6 or Eq. 7, into an algebraic system of equations. We initialize this system at $V^i = V^i_{in} = V^i_{out} = E^i_l$ for all i (and $P^i = \alpha^i(E^i_l)/(\alpha^i(E^i_l) + \beta^i)$ for the synapses) and solve it using a fixed point Jacobi iteration scheme: at each step the V^is are recomputed from the differential-turned-algebraic equations, following which the V^i_{in}s, V^i_{out}s, and P^is are recomputed from their respective equations using the new V^is. The iterations are continued until convergence, set as all new values evaluating to within an error bound of their previous values.

We make two final observations. First, although we have used a specific function, the sigmoid, in the synaptic models, the upcoming analysis is general and applies to any class of parameterized smooth functions. Second, voltage gated channels have not been included in the compartmental model ODE to eliminate unnecessary clutter. Mathematically speaking, synaptic channels transform into voltage gated channels when V_{pre} is replaced by the V^i of its own compartment. The forthcoming analysis is therefore without loss of generalization.

Local Analysis: Gradient of Voltage Trajectory and Optimization

For any given time varying drive at the input synapses of a network that begins with the canonical baseline value, the differential-algebraic system of equations for the network traces out a corresponding trajectory starting at its equilibrium state. We linearize the system of ODEs with respect to all parameters in a local neighborhood of this trajectory. Assuming infinitesimal perturbations Δ, and dropping higher order terms, we get from Eq. 1, for each compartment i, the ODE

$$
\begin{aligned}
\tau^i \frac{d\Delta V^i}{dt} =& (-\Delta V^i) \\
&+ 2(V^i_{in} - V^i)\Delta\pi^i + 2(\Delta V^i_{in} - \Delta V^i)\pi^i + 2(V^i_{out} - V^i)\Delta\pi^i + 2(\Delta V^i_{out} - \Delta V^i)\pi^i \\
&+ (E^i_{syn} - V^i)\Delta\rho^i + (-\Delta V^i)\rho^i
\end{aligned} \tag{8}
$$

where, for the particular functional form of the first synaptic conductance model Eq. 6,

$$
\Delta\rho^i = \rho^i \left(\frac{\Delta\mu^i}{\mu^i} + (1 - \frac{\rho^i}{\mu^i})(V_{pre}\Delta\nu^i + \Delta\eta^i + \nu^i \Delta V_{pre}) \right) \tag{9}
$$

and, for that of the second synaptic conductance model Eq. 7,

$$
\begin{aligned}
\Delta\rho^i =& \Delta\gamma^i P^i + \gamma^i \Delta P^i \\
\frac{d\Delta P^i}{dt} =& (-\Delta P^i)(\alpha^i + \beta^i) + (1 - P^i)\Delta\alpha^i + (-P^i)\Delta\beta^i \\
\Delta\alpha^i =& \alpha^i \left(\frac{\Delta\mu^i}{\mu^i} + (1 - \frac{\alpha^i}{\mu^i})(V_{pre}\Delta\nu^i + \Delta\eta^i + \nu^i \Delta V_{pre}) \right)
\end{aligned} \tag{10}
$$

In addition, from the algebraic equations, Eq. 2 and Eq. 3, we get, respectively,

$$
\begin{aligned}
&\Delta V^*_{in} \left((\pi^* g^*_l) + \sum_{i=1}^{m} (\pi^i g^i_l) \right) + V^*_{in} \left((\pi^* \Delta g^*_l + \Delta\pi^* g^*_l) + \sum_{i=1}^{m} (\pi^i \Delta g^i_l + \Delta\pi^i g^i_l) \right) = \\
&(\Delta V^* \pi^* g^*_l + V^* \pi^* \Delta g^*_l + V^* \Delta\pi^* g^*_l) + \sum_{i=1}^{m} (\Delta V^i \pi^i g^i_l + V^i \pi^i \Delta g^i_l + V^i \Delta\pi^i g^i_l)
\end{aligned} \tag{11}
$$

and

$$
\begin{aligned}
&\Delta V^*_{out} \left((\pi^* g^*_l) + \sum_{i=1}^{m} (\pi^i g^i_l) \right) + V^*_{out} \left((\pi^* \Delta g^*_l + \Delta\pi^* g^*_l) + \sum_{i=1}^{m} (\pi^i \Delta g^i_l + \Delta\pi^i g^i_l) \right) = \\
&(\Delta V^* \pi^* g^*_l + V^* \pi^* \Delta g^*_l + V^* \Delta\pi^* g^*_l) + \sum_{i=1}^{m} (\Delta V^i \pi^i g^i_l + V^i \pi^i \Delta g^i_l + V^i \Delta\pi^i g^i_l)
\end{aligned} \tag{12}
$$

Careful inspection of the system Eq. 8, Eq. 11, Eq. 12, and Eq. 9 (for the first synaptic model) or Eq. 10 (for the second synaptic model) reveals what has been achieved. First, the

ΔV^is (and ΔP^is in the case of the second synaptic model) are now defined as a system of linear, albeit nonhomogeneous and time-varying ODEs. Second, the ΔV^is are linear in the parameters, $\{\Delta\pi^i, \Delta g_l^i, \Delta\gamma^i, \Delta\beta^i, \Delta\mu^i, \Delta\nu^i, \Delta\eta^i\}$.

We can now determine how the trajectory $\mathbf{V}(t)$ composed of the $V^i(t)$ of all compartments i, changes with infinitesimal changes in the parameters that specify a network. Let X be a proxy for any of the parameters $\{\pi, g_l, \gamma, \beta, \mu, \nu, \eta\}$. Then, $\partial V_{rest}^i/\partial X^j$—the change in the equilibrium resting potential of V^i per unit change in parameter X^j—is obtained by solving the differential-algebraic system of equations with exactly one perturbation $\Delta X^j = 1$ and all other perturbations $= 0$, setting $d\Delta V^i/dt = 0$ (and $d\Delta P^i/dt = 0$) for all compartments i, and setting the input drive to the canonical baseline. That is, $\partial V_{rest}^i/\partial X^j$ is the fixed point of ΔV^i (calculated as in the previous section) when $\Delta X^j = 1$ is the only nonzero parameter perturbation. Likewise, $\delta V^i/\partial X^j$—the time varying change in the trajectory of $V^i(t)$ per unit change in parameter X^j—is obtained by solving the differential-algebraic system of equations for ΔV^i, initialized at $\partial V_{rest}^i/\partial X^j$, setting that perturbation $\Delta X^j = 1$ and all other perturbations $= 0$. That is, $\delta V^i/\partial X^j$ is the solution to $\Delta V^i(t)$ when initialized at $\Delta V^i(0) = \partial V_{rest}^i/\partial X^j$ and solved for $\Delta X^j = 1$ as the only nonzero parameter perturbation.

In this general setting, tuning the parameters of a network so as to elicit a desired property of the trajectory $\mathbf{V}(t)$ involves incrementally updating the parameters X^j informed by properties of the trajectory $\delta V^i/\partial X^j$. Given an appropriate functional G applied to $\delta V^i/\partial X^j$, gradient ascent updates can be computed for each parameter X^j as $\Delta X^j = \zeta G[\delta V^i/\partial X^j]$, where ζ is the learning rate. The choice of the functional G, of course, depends on the specific scenario under consideration. We consider two scenarios; others can be worked out in like fashion.

Let $E[f]$ be a functional defined on function f. Then, the functional derivative, $\delta E/\delta f$ is, by definition, the linear functional such that $E[f + \epsilon g] - E[f] = \epsilon \times (\delta E/\delta f)[g]$ as $\epsilon \to 0$. This linear functional can be specified in multiple ways: as a function in the integral operator,

$(\delta E/\delta f)[g] = \int(\delta E/\delta f) \times g$, as a point evaluation on a linear map on g, etc.

In a scenario where the desired trajectory $U^i(t)$ of compartment i is known for time period $[0, T]$, the error functional, $E[V^i]$, can be defined as:

$$E[V^i] = ||(U^i - V^i)||^2 = \langle (U^i - V^i), (U^i - V^i) \rangle = \int_0^T (U^i(t) - V^i(t))^2 dt \tag{13}$$

It then follows from the definition above that $\delta E/\delta V^i$ is the integral operator specified by $-2(U^i - V^i)$, and therefore, $\partial E/\partial X^j = \int_0^T -2(U^i - V^i)(\delta V^i/\partial X^j)dt$. In this case then, the functional $G[\delta V^i/\partial X^j]$ referred to above corresponds to the inner product in function space between the trajectories $2(U^i - V^i)$ and $\delta V^i/\partial X^j$, the dropping of the negative sign enforcing descent on E.

In the case of the EMD however, the desired trajectory $U^i(t)$ is unknown. What is known is that the time to maximum (minimum) response of $V^i(t)$ of the compartments of Mi4 (Mi9) are to be delayed by a specific amount with reference to that of Mi1. Setting t^* to denote the desired time to reach the extremum, the error functional $E[V^i]$ can be defined as:

$$E[V^i] = (t^* - \arg\underset{t \in [t_1, t_2]}{\text{extremum}}\, V^i(t))^2 = (t^* - t \in [t_1, t_2] \left| \frac{dV^i(t)}{dt} = 0 \right.)^2 \tag{14}$$

where the second equality follows from the assumption that V^i is unimodal over the range $[t_1, t_2]$ and therefore attains its extremal value at the unique time at which $dV^i/dt = 0$.

We can now analyze $E[V^i + \epsilon W^i]$ as follows. Let $\widehat{t}, \widetilde{t} \in [t_1, t_2]$ be the times at which $d/dt[V^i](\widehat{t}) = 0$ and $d/dt[V^i + \epsilon W^i](\widetilde{t}) = 0$. Then applying a Taylor expansion to the second equation around $\widehat{t}$ and dropping higher order terms noting that $\widetilde{t} \to \widehat{t}$ as $\epsilon \to 0$, we get

$$\widetilde{t} = \widehat{t} + \frac{-\epsilon \frac{dW^i}{dt} \mid_{\widehat{t}}}{\frac{d^2V^i}{dt^2} \mid_{\widehat{t}} + \epsilon \frac{d^2W^i}{dt^2} \mid_{\widehat{t}}} = \widehat{t} + \epsilon \frac{-\frac{dW^i}{dt} \mid_{\widehat{t}}}{\frac{d^2V^i}{dt^2} \mid_{\widehat{t}}} \text{ as } \epsilon \to 0 \tag{15}$$

Based on the definition of functional derivative, we then have

$$\frac{\delta E}{\delta V^i}[W^i] = -2(t^* - \widehat{t})\frac{-\frac{dW^i}{dt}\mid_{\widehat{t}}}{\frac{d^2V^i}{dt^2}\mid_{\widehat{t}}}\text{, hence, } \frac{\partial E}{\partial X^j} = \frac{\delta E}{\delta V^i}[\frac{\delta V^i}{\partial X^j}] = 2(t^* - \widehat{t})\frac{\frac{d}{dt}\frac{\delta V^i}{\partial X^j}\mid_{\widehat{t}}}{\frac{d^2V^i}{dt^2}\mid_{\widehat{t}}} \tag{16}$$

We now make two observations. First, both $(t^* - \widehat{t})$ and $d^2V^i/dt^2 \mid_{\widehat{t}}$ do not depend on the choice of X^j and thus amount to a scaling factor in the gradient. Second, the sign of $(t^* - \widehat{t})$ is positive so long as $\widehat{t} < t^*$, and the sign of $d^2V^i/dt^2 \mid_{\widehat{t}}$ is negative if we choose the argmax of the trajectory of an Mi4 compartment and is positive if we instead choose the argmin of the trajectory of an Mi9 compartment. Therefore, if we choose to push the argmax of the V^i of an Mi4 compartment to the future, the functional $G[\delta V^i/\partial X^j] = d/dt[\delta V^i/\partial X^j](\widehat{t})$, that is, the time derivative of the trajectory $\delta V^i/\partial X^j$ at the instant when V^i attains its maximum. The sign is flipped if we choose to push the argmin of the V^i of an Mi9 compartment into the future.

The Substantial Requisite Delay in the EMD and its Deduced Origin

What makes constructing a model neural circuit implementation of the EMD difficult is not the mere presence of the delay lines as prescribed by the HR and BL models, but the magnitude of the requisite delay as can be inferred from available stimulus-response recordings. The authors in [13], for example, presented a moving 2° ON bar at various orientations and velocities in the receptive field of a T4 cell of a Drosophila, while concurrently recording the T4's membrane potential response. They found robust directionally selective responses in the behaviorally relevant velocity range of $14^\circ - 112^\circ/s$. The delay necessary for a robust response at $14^\circ/s$ can be determined from numerical experiments based on anatomical data as described next.

The ~ 750 omatidia in the Drosophila eye [11] are arranged with interomatidial angle of $\approx 5.1^\circ$ [27]. The normalized angular sensitivity of photoreceptors R1-R6 in each omatidium is well approximated by 2-D isotropic Gaussians with full width at half maximum (FWHM) estimated at 4.2° when the combined effects of geometric and wave optics are invoked [27].

Fig.1b-i presents numerical simulations of the temporal profiles of the normalized stimulus intensity sensed by the photoreceptors of seven contiguous omatidia in the lattice of the compound eye of the Drosophila for a 2° ON bar traveling at $12^\circ/s$. The peak to peak delay between the profiles of neighboring photoreceptors for the two lattice orientations are 212 and $368ms$, respectively, with a FWHM of $\approx 360ms$ for each profile. Clearly, a delay of the order of hundred milliseconds or more would be necessary to induce robust direction selectivity at this stimulus velocity. Such a delay can either arise from complex synaptic dynamics, recurrent network dynamics, or a combination thereof. Our optimized model demonstrates that recurrent network dynamics suffices as a parsimonious explanation (Fig.1b-iii).

Central to the model is a Mi4-Mi9 recurrent network whose synapses, when tuned via the optimization procedure, cause the response of both Mi4 and Mi9 to be delayed and extended in time (Fig.1b-ii). In order to guarantee that the optimization procedure did not simply raise the synaptic time constants to precipitate the delay, we used the first synaptic model Eq. 6 whose response to changes in the presynaptic potential is instantaneous. By eschewing delays, we were assured that the synaptic dynamics did not pose a confound for the delays generated via recurrent network dynamics. Furthermore, the parameters of the network model optimized were restricted to the $\{\mu^i, \nu^i, \eta^i\}$ for all synapses on Mi4 and Mi9; all other parameters were fixed to values determined from the anatomy and physiology of the Drosophila optic lobe. To achieve the delay, we chose the functional $G[\cdot]$, introduced in the previous section, to be the time derivative of $\delta V^i/\partial X^j$ for a Mi4 compartment at the instant when its $V^i(t)$ reached its maximum in response to a bar stimulus. This allowed us to make progressive infinitesimal updates to the parameters that pushed Mi4's and Mi9's time to peak to the future (Fig.2a-iv).

The resultant optimized behavior of Mi4 and Mi9 is, crucially, consistent with the neurons' temporal kernels deduced from reverse correlation experiments [3]. Additionally, we found that although the temporally extended responses of the neurons permit simple synaptic integration

of their signals at T4 to be sufficient to induce direction selectivity [13], both preferred direction enhancement and null direction suppression [14, 19] is necessary to abridge the overall response. Finally, the characteristics of the response to drifting sinusoidal gratings are readily explained by the charging-up of the recurrent networks and their low-pass nature (Fig.2b). We elaborate on these results below.

The End-to-End EMD Circuit Implementation

Phototransduction in the Drosophila photoreceptor has been investigated extensively [21, 18]. We used the Wong and Knight model with its reported parameters [18] (order $n = 8, \tau = 2ms$) without static nonlinearities to model the photoreceptor potential (Methods). The histaminergic photoreceptor-LMC synapse, responsible for the sign inverted response in L1 (Fig.1b-ii), has likewise been studied in detail [32], its high pass nature attributed to presynaptic subtraction of the variable extracellular field potential at the synapse [31]. We modeled the field potential as low pass filtered photoreceptor potential ($\tau = 2ms$) and set the difference as the V_{pre} for the hyperpolarizing synapse (Methods).

The lengths, diameters, and synaptic polarities of L1, Mi1, Tm3, C3, L5, Mi4, Mi9, and T4 were drawn from [29, 22] (Table.1). L1 was assigned a $15\mu m$ high conductance initial synaptic zone [24]. Simulations were conducted using $l = 30\mu m$ compartments with a $0.01ms$ time step size and potentials specified in mV. We verified that the results were identical to within precision bounds when using $l = 10\mu m$ compartments (Fig.S1-2, S4). Notably, extensive simulations with randomly sampled diameters and lengths indicated that the neurons in the circuit were approximately equipotential across their entire cell body (Methods). The length of the neurites could not therefore be implicated in the generation of the requisite delays.

L1's hyperpolarizing synapses onto Mi1, L5, and C3, responsible for their sign inverted back response, were modeled to be identical (Table.1) with μ set to fit the $10 - 20mV$ recorded

response to flashes [5], and η set near saturation to induce the requisite half wave rectification (Fig.1b-ii and Methods). Since Mi1 and Tm3 have similar response profiles for moderate stimulus velocities, and both depolarize T4 [28], we chose to use only Mi1 in our parsimonious EMD model, consistent with the observed impact of silencing Tm3 [2, 28].

The mechanism that can potentially engender a delayed and extended response in a recurrent Mi4-Mi9 network can be conceptualized as follows. Mi4 is depolarized by L5's synaptic input whereas Mi9 is hyperpolarized by C3's synaptic input [29]. The reciprocal Mi4-Mi9 synapses are, however, both hyperpolarizing [28, 29]. A hyperpolarized Mi9 in turn anti-hyperpolarizes Mi4, while a depolarized Mi4 hyperpolarizes Mi9, triggering a positive feedback loop. With appropriately set μ's, ν's and η's, this feedback can be tuned to both commence and return cells to equilibrium slowly. We searched for such parameters using the optimization procedure (Fig.2a-iv) that incrementally pushed the peak response time of Mi4 and Mi9 to the future, when the photoreceptors were driven by a bar stimulus. We found a wide range of parameters for which this was achievable (Fig.2a-i, iii and Fig.3a), demonstrating that the phenomenon was robust. Table.1 reports the parameters of the particular network presented here. Importantly, we found that the dynamics could be replicated using a single drive: eliminating either the L5 or C3 synapse and raising the μ of the other (Methods).

We found that setting the peak response time of Mi4/Mi9 to be delayed by $\approx 125ms$ with reference to that of Mi1 (Fig.1b-ii) was sufficient to induce robust direction selectivity in T4 (Fig.1b-iii), owing to the slow decay ($> 4s$) of the Mi4/Mi9 response (Fig.2a-i). However, to abridge the T4 response, it was necessary to have the postsynaptic impact of the decays counteract one another symmetrically: we built a T4 cell that received a depolarizing synapse from the Mi1 of the home module and hyperpolarizing synapses from three Mi9s and three Mi4s [28] from the six neighboring modules (Fig.1a), such that a preferred direction (PD) stimulus triggered, in order, Mi9-Mi1-Mi4 and a null direction (ND) stimulus triggered them in reverse.

The μ's, ν's, and η's of the synapses were set to evoke commensurate effect from each synaptic input, operate in the linear regime, and be symmetric for each Mi4-Mi9 opposing pair to elicit cancellation (Table.1).

The robust directional response of T4 to bar stimuli traveling at $> 100°/s$ [13] (Fig.2a-ii) seems at odds with its equilibrium response to drifting sinusoidal gratings. Grating associated direction selectivity of wide field horizontal cells, the postsynaptic target of T4, has been observed to be a function of temporal frequency [25]—grating velocity divided by wavelength—peaking at a mere $1Hz$. We found that the Mi4-Mi9 network charges-up when driven by gratings, with respective baseline potentials equilibrating at values determined by the grating's frequency (Fig.2b-i). Setting the sensitivity ν of the Mi4/Mi9 synapses on T4 to saturate/close at the peak of the $1Hz$ charged-up equilibrium response of Mi4/Mi9 (Methods) led the synapses to either be persistently open at near full conductance, or be closed, at $> 1Hz$. This resulted in the progressive eradication of grating associated directional responses at higher frequencies (Fig.2b-ii), without compromising bar stimuli associated responses (Fig.2a-ii).

Discussion

Synaptic Profiles can be the Key Determinant of Behavior

Recurrent connectivity is ubiquitous in animal brains. Without precise knowledge of the corresponding synaptic profiles, however, it is difficult to assess what role those recurrent connections play in the overall behavior of a given network. The optimized EMD circuit drives home this message definitively. Careful analysis of the dynamics of the network has shown us that the profiles of just two synapses, Mi4's on Mi9 and Mi9's on Mi4, in relation to the profiles of the other synapses, determine the emergence of the intended behavior of the network. Fig.3a presents the profiles $\rho(V_{pre})$ of the pair of synapses over the operational range of the V_{pre} of Mi4 and Mi9, before and after optimization, and Fig.3b presents the corresponding voltage pro-

files of Mi4 and Mi9 when the photoreceptors are driven by the bar stimulus. As is clear from the voltage profiles of T4 in Fig.3c (as compared to that in Fig.1b-iii) directionally selective response is entirely absent in the pre-optimized network. The critical dynamical characteristics of a network can therefore be determined by the profiles of a select few constituent synapses; broad connectivity information is inadequate in such cases.

Whence the Delay

Although the nonlinear dynamics of recurrent networks belie simplistic explanations in the general case, several fortuitous features of the optimized Mi4-Mi9 network allow for the identification of the source of its delayed response. First, the observation that the neurons are approximately equipotential across their entire cell bodies allows us to model both Mi4 and Mi9 as single compartments. Second, the profiles of $\rho(V_{pre})$ for Mi4 and Mi9 in Fig.3a, for their respective operational ranges, suggest that they can be approximated by affine functions: $\rho(V_{pre}) = a + bV_{pre}$. Third, the symmetric dynamics of Mi4 and Mi9, as displayed in Fig.1b-ii and Fig.3b, permits the approximation $(V^{Mi4} - V^{Mi4}_{rest}) = -(V^{Mi9} - V^{Mi9}_{rest})$. When the ODEs for Mi4 and Mi9 are simplified using these substitutions, one finds that the ODEs, shifted by respective equilibrium resting potentials, are identical to first approximation.

Equivalently, the Mi4-Mi9 recurrent network can be approximated as a single, one compartment, neuron with an *inverted autapse* following the now quadratic ODE:

$$\tau \frac{dV}{dt} = (E_l - V) + (E_{syn} - V)(-a - bV) + V_{drive} \tag{17}$$

where V_{drive} approximates the L5 to Mi4 or C3 to Mi9 drive. Assuming equilibrium V_{rest} is at $V_{drive} = 0$, the right hand side of the reduced model Eq. 17 can the rewritten in terms of $V_{rest} + \overline{V}$ where $\overline{V}$ is the deviation of the potential from V_{rest}. We then get $\tau d\overline{V}/dt = -\overline{V}(1 - a - 2bV_{rest} + bE_{syn}) + \overline{V}^2 b + V_{drive}$. To a first approximation, the time constant near

equilibrium therefore scales by $1/(1 - a - 2bV_{rest} + bE_{syn})$, which depending on the values of a, b, V_{rest}, E_{syn} can be very large, causing the response to have a slow rise and slow decay.

A More Nuanced Description of the EMD Mechanism

The HR and BL models prescribe delay lines that are nonlinearly integrated with the direct lines to engender direction selectivity. We have however found that the mechanism that leads to direction selectivity in the model EMD neural circuit is more subtle and emerges by way of several interacting processes. The delay originating from the nonlinear dynamics of the recurrent Mi4-Mi9 network comes with an associated broadening of the Mi4/Mi9 response (Fig.1b-ii). The magnitude and width of this response (Fig.2a-i and iii) vary nonlinearly with the velocity of the input stimulus. The benefit of this broadening of response is that it functions as an additional delay. This is why a delay of $\approx 125ms$ between the peak response time of Mi1 and that of Mi4/Mi9 was found to be sufficient to induce robust direction selectivity, even when the peak to peak delay between the profiles of neighboring photoreceptors in the lattice ranged between 212 to $368ms$ (Fig.1b-i). The drawback of this broadening is that the tail of the response can persist for longer than is desirable. The EMD circuit eliminates the tail response via opponency: the Mi4 and Mi9 of diametrically opposed lattice points feed the home network. The result is that for a PD stimulus the Mi1 response of the home module sits atop a depolarizing bump generated by the Mi9-then-Mi4 response. Contrastingly, for a ND stimulus the Mi1 response sits atop a hyperpolarizing dip generated by the Mi4-then-Mi9 response. Simple synaptic integration of these signals at T4 therefore suffices to induce direction selectivity.

Our model has thus shown that recurrent dynamics can explain the EMD. Importantly, the optimization procedure that we have leveraged to arrive at the model parameters holds promise for networks beyond the one addressed here.

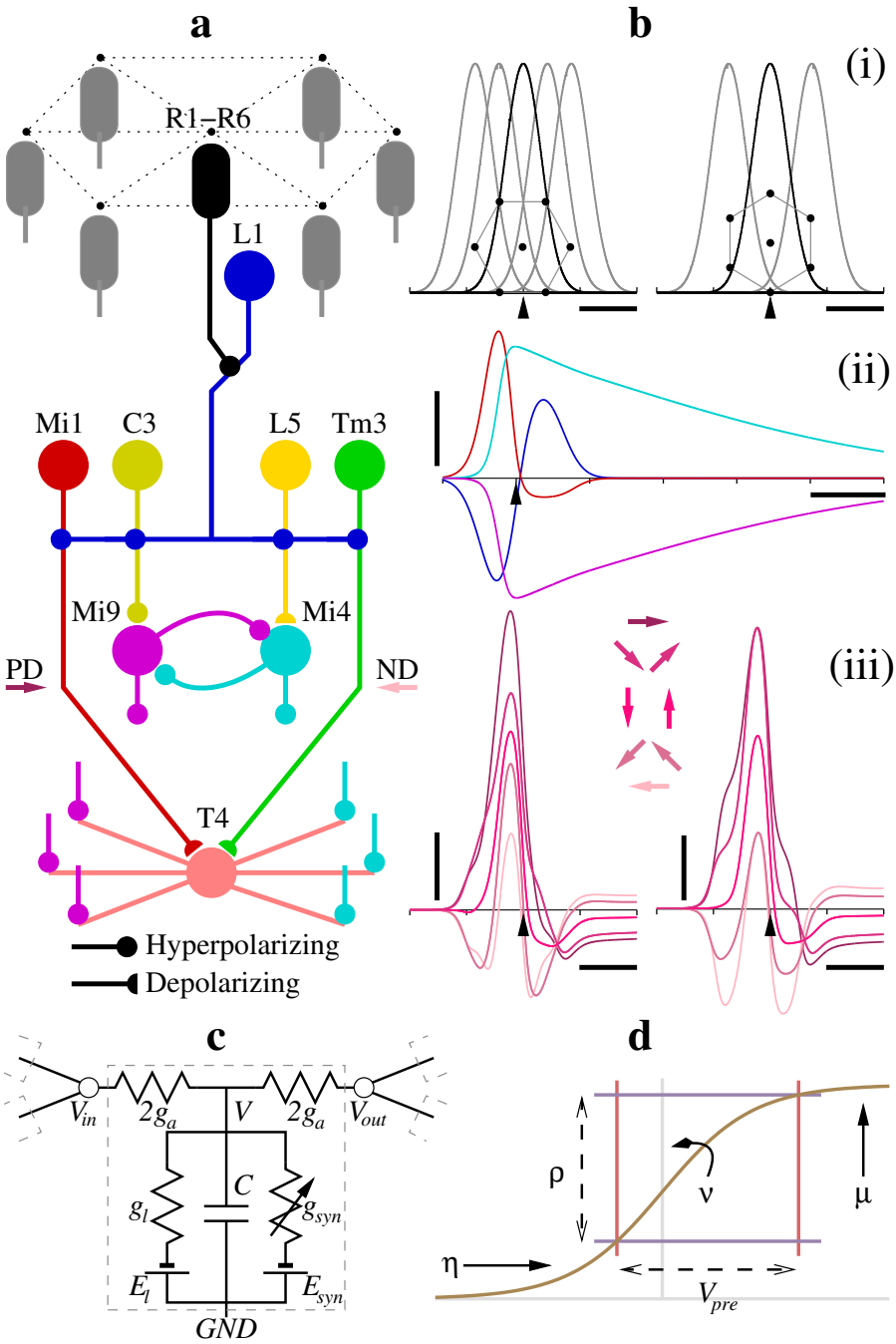

a
b
R1–R6
L1
(i)
(ii)
Mi1
C3
L5
Tm3
Mi9
Mi4
PD
ND
(iii)
T4
Hyperpolarizing
Depolarizing
c
d
V_{in}
$2g_a$
V
$2g_a$
V_{out}
C
g_l
g_{syn}
E_l
E_{syn}
GND
ρ
ν
μ
η
V_{pre}

Figure 1:

Figure 1: Neural circuit and representative behavior. a, Circuit for a single module. The circuit is replicated for all modules in the hexagonal lattice, each receiving input from its corresponding photoreceptors (gray). The home module T4 receives input from three Mi9 (from contiguous neighboring modules), three Mi4 (from symmetrically opposing modules), and the Mi1 and Tm3 of the home module. Synaptic parameters, set symmetrically for each Mi4/Mi9 opposing pair, differ for T4s tuned to different cardinal directions (Methods). **b,** Responses of neurons to a 2° wide bar of light traveling at $12^\circ/s$. Responses from circuits tuned to two cardinal directions are shown; responses for the other two cardinal directions are symmetric. (i) Temporal profiles of the normalized stimulus intensity sensed by the photoreceptors of seven contiguous omatidia in the lattice for the two cardinal circuits. Insets display the orientation of the lattice points with respect to the bar stimulus traveling left to right. (ii) Responses of L1, Mi1, Mi4, and Mi9, color coded and shifted along the ordinate by their respective equilibrium resting potentials. The responses of Mi4 and Mi9 are also scaled $\times 10$. (iii) Responses of the home module T4s tuned to the two cardinal directions. Stimulus directions PD, PD$\pm\pi/4$, PD$\pm\pi/2$, ND$\pm\pi/4$, and ND, are with respect to the insets in (i). Horizontal scale bars in (i), (ii), (iii) $= 500ms$. Vertical scale bar in (ii) $= 5mV$, in (iii) $= 1mV$. Pointers on abscissa mark time alignment for the home module. **c,** Conductance based model for a single compartment. Linked compartments i, j satisfy $V_{in}^i = V_{out}^j$. **d,** First model of synapse.

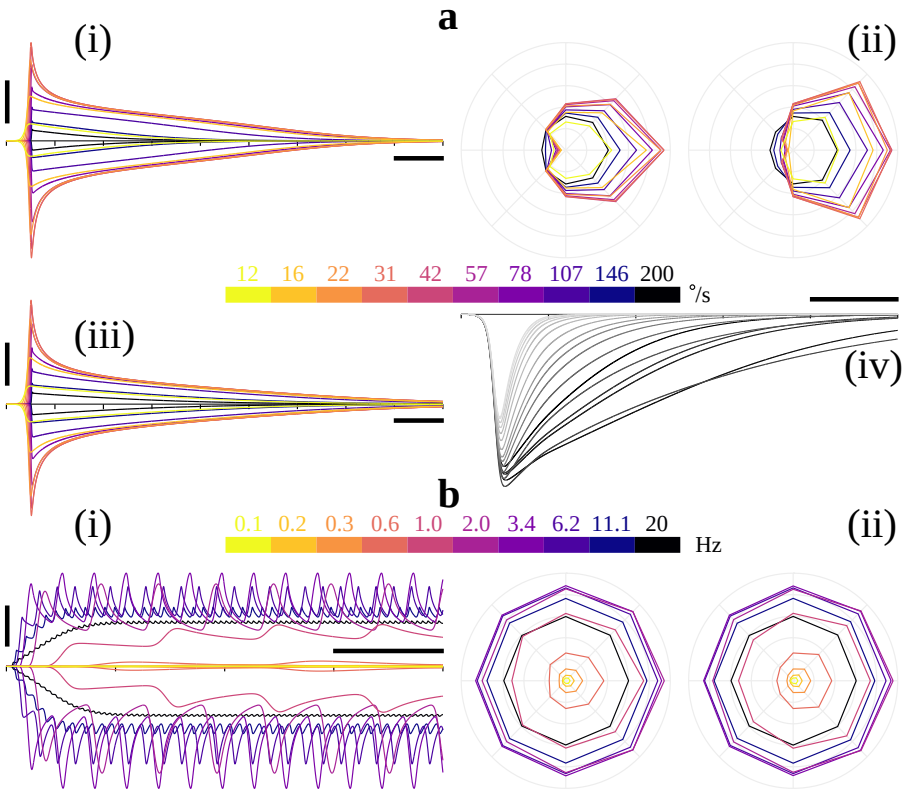

a
(i)
(ii)
12 16 22 31 42 57 78 107 146 200
°/s
(iii)
(iv)
b
(i)
0.1 0.2 0.3 0.6 1.0 2.0 3.4 6.2 11.1 20
Hz
(ii)

Figure 2:

Figure 2: Responses of Mi4, Mi9, and T4 to bar and drifting sinusoidal grating stimuli. **a,** Bar stimuli. (i) Responses of Mi4 and Mi9 to a 2° wide bar traveling at different velocities. The responses are color coded and shifted along the ordinate by their respective equilibrium resting potentials. (ii) Corresponding responses of the home module T4s tuned to the two cardinal directions. Peak response above equilibrium resting potential is displayed for each stimulus direction PD, PD$\pm\pi/4$, PD$\pm\pi/2$, ND$\pm\pi/4$, and ND, for the different velocities. (iii) The Mi4-Mi9 dynamics is robust: response of a network with different synaptic parameters (Methods) from another optimization run. (iv) The optimization in action: the delayed and extended dynamics (only Mi9 shown) evolved gradually with changing synaptic parameters for the same stimulus. **b,** Sinusoidal grating stimuli. (i) Responses of Mi4 and Mi9 to a 20° wavelength sinusoidal grating traveling at different velocities. The responses are color coded and shifted along the ordinate by their respective pre-stimulus equilibrium resting potentials. Both the charging-up and the low-pass nature of the circuit are manifest. (ii) Same as a(ii) for this stimuli. Horizontal scale bars in a(i), a(iii), a(iv), b(i) $= 1s$. Vertical scale bar in a(i), a(iii) $= 2mV$, in b(i) $= 3mV$. Voltage increment of concentric circles in a(ii) $= 2mV$, in b(ii) $= 1mV$.

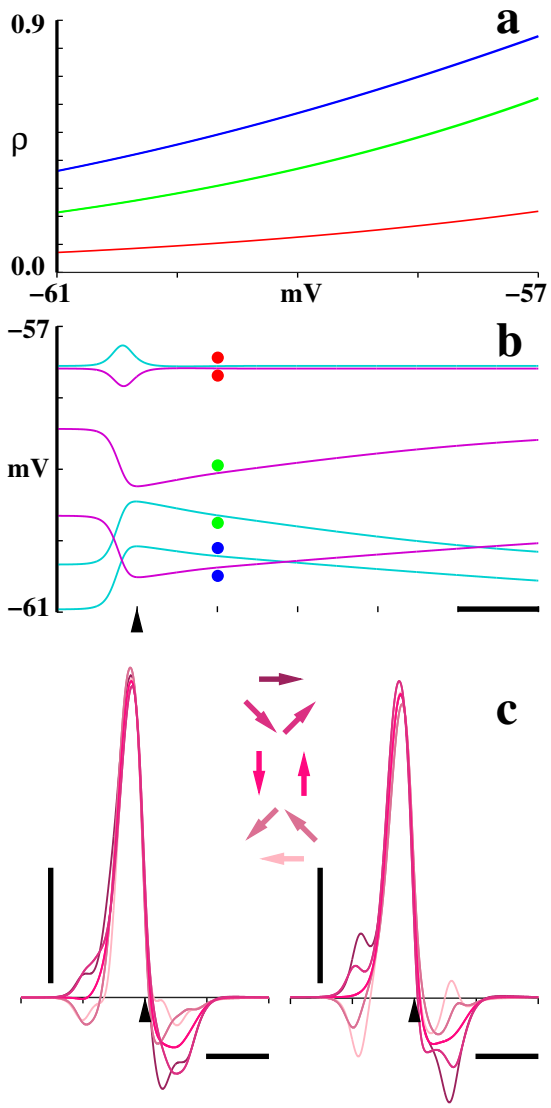


Figure 3:

Figure 3: Pre- and Post Optimized Mi4-Mi9 network and corresponding T4 behavior. **a,** $\rho(V_{pre})$ of the pair of synapses. Initialization, i.e., pre-optimized synapse pair labeled in red. Optimized synapse pair of the network presented in the article labeled in green. A second optimized pair, corresponding to the network in Fig.2a-iii labeled in blue. Note that the range $[-61, -57]mV$ subsumes the operating range of all three synapse pairs as shown in (b). The optimization was not constrained to update the synapses symmetrically. However, it did to within precision bounds. All parallel graphs of $\rho(V_{pre})$ between the green and blue curves manifest the delay necessary to implement the EMD, demonstrating that the phenomenon is robust. **b,** Responses of Mi4 (cyan) and Mi9 (magenta) in the three networks for a 2° wide bar of light traveling at $12^\circ/s$. Traces are labeled with adjacent color coded dots to distinguish between the three networks. Not only does the optimization delay the time to peak for both Mi4 and Mi9, it also substantially slows the tail decay. Note that both V_{rest} and $V(t)$ change as the synaptic parameters are changed. **c,** Responses of the home module T4s tuned to the two cardinal directions (as in Fig.1b-i) when the Mi4/Mi9 synapses are in the pre-optimized parameter setting in red from (a). Compare these to Fig.1b-iii which corresponds to the Mi4/Mi9 synapses in the post-optimized parameter setting in green from (a). Stimulus directions PD, PD$\pm\pi/4$, PD$\pm\pi/2$, ND$\pm\pi/4$, and ND, are with respect to the insets in Fig.1b-i. Note that directionally selective response is almost absent. Horizontal scale bars in (b), (c), $= 500ms$. Vertical scale bar in (c) $= 1mV$. Pointers on abscissa mark time alignment for the home module as in Fig.1b.

Table 1: Parameters for the full network using $l = 30\mu m$ compartments.

Presynaptic Neuron (Diameter d, Length l)	Postsynaptic Neuron	E_{syn}	μ	ν	η
R1-R6 $(-, -)$	L1	$-85mV$	40.0	3.0	-3.82
L1 $(1\mu m, 100\mu m)$	Mi1/C3/L5	$-85mV$	5.0	0.4	27.4
L5 $(0.6\mu m, 60\mu m)$	Mi4	$0mV$	0.6	0.3	15.0
C3 $(0.6\mu m, 60\mu m)$	Mi9	$-85mV$	0.6	0.3	15.0
Mi4 $(0.6\mu m, 60\mu m)$	Mi9	$-85mV$	4.03555	0.29494	15.10545
Mi9 $(0.6\mu m, 60\mu m)$	Mi4	$-85mV$	4.03574	0.29485	15.10667
Mi1 $(0.6\mu m, 60\mu m)$	T4a(b)	$0mV$	1.5 (1.5)	0.1	7.4912
Mi9 $(0.6\mu m, 60\mu m)$	T4a(b)	$-85mV$	1.5 (2.25)	1.0	58.435
Mi4 $(0.6\mu m, 60\mu m)$	T4a(b)	$-85mV$	1.5 (2.25)	1.0	60.317

Materials and Methods

Custom code written in C was used both for simulating the dynamics of the model as well as for the optimization. The programs were run on a 10-Core, 2.3GHz Linux server with 64 GB RAM.

Compartment Model

Each compartment i was modeled to be cylindrical of diameter d^i and length l^i, following $C^i dV^i/dt = (E_l^i - V^i)g_l^i + 2(V_{in}^i - V^i)g_a^i + 2(V_{out}^i - V^i)g_a^i + (E_{syn}^i - V^i)g_{syn}^i$, (Fig.1c) with the differential equation divided on both sided by g_l^i to ease the derivation of parameters as well as to ensure numerical stability. Universal membrane constants [12] were set as: specific membrane capacitance $C_m = 1\mu F/cm^2$, specific membrane resistance $R_m = 20K\Omega cm^2$, and specific axial resistance $R_a = 100\Omega cm$. For a cylindrical compartment i with surface area A^i and cross sectional area a^i, $C^i = C_m \times A^i$, $g_l^i = (1/R_m) \times A^i$, and $g_a^i = (1/R_a) \times a^i \times (1/l^i)$. Parameters were first computed for a canonical cylinder $i = 0$ of diameter $d^0 = 1\mu m$ and length $l^0 = 100\mu m$. For all compartments i, $C^i/g_l^i = (C_m \times A^i) \times (R_m/A^i) = 20ms$. For canonical cylinder $i = 0$, calculations yield $g_l^0 = 0.15708nS$, and the unitless $\pi^0 = g_a^0/g_l^0 = 50$. Hence for cylinder i, $g_l^i = g_l^0 \times \hat{d} \times \hat{l}$, and $\pi^i = g_a^i/g_l^i = \pi^0 \times \hat{d}/\hat{l}^2$, where $\hat{d} = d^i/d^0$ and $\hat{l} = l^i/l^0$. The unitless time varying $\rho^i = g_{syn}^i/g_l^i$ which is a function of V_{pre}—the V^i of the presynaptic compartment—was modeled as $\rho^i(V_{pre}) = \mu^i/(1 + e^{-(\nu^i V_{pre} + \eta^i)})$. ρ^i therefore ranged between 0 and μ^i depending on V_{pre}. E_l was set at $-55mV$, and E_{syn} at $-85mV$ for hyperpolarizing and $0mV$ for depolarizing synapses.

The entire differential-algebraic system of equations was simulated using the forward Euler method. Simulations were initially conducted with compartments of length $l = 10\mu m$ and a time step size of $0.001ms$. However, it was observed that for the range of diameters used, the results were identical to within precision bounds with $l = 30\mu m$ compartments simulated at a

$0.01ms$ time step size (Fig.S1 and S2).

Two idiosyncrasies of tonically active synapses were immediately evident. Firstly, the equilibrium resting potential of a neuron varied markedly [12] depending on how conductive its input synapses were at equilibrium, which then depended on the μ, ν, η and equilibrium resting potential of each of its presynaptic neurons. Secondly, appropriate adjustments to μ, ν and η could replicate the qualitative dynamics for different settings of E_l and E_{syn}.

Morphology of individual Neuron Models

The Drosophila L1 has diameter $d \approx 1\mu m$ [22] and length $l \approx 100\mu m$ with an initial high conductance synaptic zone. We modeled a $15\mu m$ synaptic zone [22] for which g_l and the equilibrium g_{syn} were scaled up ten-fold [24]. A $d = 1\mu m$, $l = 15\mu m$, high conductance synaptic zone compartment for L1 was modeled as follows. Since g_l scales by (dl) and g_a scales by (d^2/l), to have g_l increase ten-fold [24] while maintaining the value of g_a fixed, d for the compartment was scaled $\times 10^{1/3}$ and l was scaled $\times 10^{2/3}$, to the new values $d = 2.1544347\mu m$ and $l = 69.623832\mu m$.

When this L1 synaptic compartment was linked to a $100\mu m$ axon built out of $d = 1\mu m$, $l = 10\mu m$ compartments, we found little difference between the V^i at the synaptic versus the most distal compartment (Fig.S1). We conducted a large set of simulation experiments where the diameters of the ten compartments were chosen randomly from an uniform distribution over the range $[-0.7, 1.3]\mu m$. In all cases, we found little difference between the V^i at the synaptic versus the most distal compartment (almost identical to Fig.S1 and hence not shown). Nor did we find a difference when appropriately fewer compartments of longer l were used (Fig.S2). Even with a $500\mu m$ axon built out of $l = 10\mu m$ compartments, the distal V^i was delayed by a mere $5ms$, although the equilibrium V^i value shifted closer to E_l (Fig.S3). In the final model, L1 was built out of three compartments linked in series: the high conductance synaptic zone

compartment, an intermediate $d = 1\mu m$, $l = 60\mu m$ compartment, and a distal $d = 1\mu m$, $l = 40\mu m$ compartment. The V^i of the distal compartment was set as the V_{pre} of L1's synapses on Mi1, C3, and L5.

The Drosophila C3 has diameter $d \approx 0.63\mu m$ [22]. For C3, Mi1, L5, Mi4, and Mi9, we set each $d = 0.6\mu m$, and since they span the $\sim 50\mu m$ medulla [29], we set each $l = 60\mu m$. The cells Mi1, C3, L5, Mi4, and Mi9 were all built to be identical in their morphology: each neuron was built as two compartments each of $d = 0.6\mu m$, $l = 30\mu m$, linked in series. We verified that the recurrent Mi4-Mi9 network had identical properties when they were each assembled out of six $l = 10\mu m$ compartments (Fig.S4). The same was true for Mi1, C3 and L5. Finally, the entire network was replicated (Fig.1a) for each of the seven contiguous modules in the hexagonal lattice.

T4 was built out of seven compartments each of $d = 0.6\mu m$, $l = 30\mu m$. The home compartment was linked to each of three compartments at one terminal and each of three other compartments at the other. The home compartment received a synaptic input from the Mi1 of the home module; each of the remaining six compartments received a synaptic input from either the Mi4 or the Mi9 from a distinct surrounding module in the hexagonal lattice (Fig.1a).

Photoreceptor and the Photoreceptor-L1 synapse

The Drosophila has neural superposition eyes where each module in the lamina receives the axon of one of each photoreceptor R1-R6, all of which manifest the same visual axis [1]. Since each omatidium contains one of each R1-R6, the angle between the visual axes of the R1-R6 in an omatidium matches the angle between the axes of neighboring omatidia [7]. Hence the visual axis of neighboring modules in the lamina are separated by the same angle as that of the omatidia. The omatidia in the Drosophila eye are arranged with interomatidial angle of $\approx 5.1°$ [27]. The normalized angular sensitivity of R1-R6 in each omatidium is well approximated

by 2-D isotropic Gaussians with full width at half maximum (FWHM) estimated at $4.2°$ when the combined effects of geometric and wave optics are invoked [27]. Hence a photoreceptor centered at $\langle \bar{x}, \bar{y} \rangle$ has normalized angular sensitivity $S_{\langle \bar{x},\bar{y}\rangle}(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x-\bar{x})^2+(y-\bar{y})^2}{2\sigma^2}}$, where $\sigma = 4.2/(2\sqrt{2\ln 2}) = 1.7835$. We modeled seven contiguous omatidia as individual photoreceptors centered on lattice points $5.1°$ apart on a hexagonal lattice (Fig.1b-i), with angular sensitivity $S_{\langle \bar{x},\bar{y}\rangle}(x, y)$. For any given time varying input stimulus $I(x, y, t)$, the input drive for the photoreceptor centered at $\langle \bar{x}, \bar{y} \rangle$ was computed as $I_{\langle \bar{x},\bar{y}\rangle}(t) = \int_{\mathbb{R}^2} S_{\langle \bar{x},\bar{y}\rangle}(x, y) \times I(x, y, t)dxdy$. For both the bar and the drifting sinusoidal grating stimuli, the X-Y plane of integration was first rotated so as to align one of the axes to be orthogonal to the stimulus direction. This allowed the integral to be simplified in closed form to a 1-D integral over $\mathbb{R}$. Numerical integration was then performed with a spatial step size of $0.04°$.

The Drosophila photoreceptor response has previously been modeled as a NLN (nonlinear static-linear dynamic-nonlinear static) cascade, where the two static nonlinearities are third order polynomial functions and the linear component is the Wong and Knight photoreceptor model $g(t) = \frac{1}{n!\tau}(\frac{t}{\tau})^n e^{-t/\tau}$, with the model parameters fit to recorded data [18]. Since the two polynomial fits were found to be fairly close to linear for wild type Drosophila [18], we used only the linear component with the parameters [18] $n = 8$ and $\tau = 2ms$. For a photoreceptor centered at $\langle \bar{x}, \bar{y} \rangle$, the photoreceptor potential was therefore modeled as $V^{pr}_{\langle \bar{x},\bar{y}\rangle}(t) = \int g(u) I_{\langle \bar{x},\bar{y}\rangle}(t - u)du$.

The photoreceptor-LMC synapses being histaminergic, we set $E_{syn} = -85mV$ for the photoreceptor-L1 synapse. It has been demonstrated [31] that at this specialized synapse, the extracellular field potential changes with illumination, and the presynaptic transmembrane potential in the photoreceptor axon is well approximated by subtracting the field potential from the intracellular potential. We modeled the field potential as low-pass filtered photoreceptor potential: $\tau dV^{fp}/dt = -V^{fp} + V^{pr}$ with $\tau = 2ms$, and set the V_{pre} of the photoreceptor-L1

synapse as $V_{pre} = V^{pr} - V^{fp}$.

Synaptic Parameters

Key to most of the parameter settings are the observations that $1/(1+e^{-x})$ is convex for $x < 0$, concave for $x > 0$, approximately linear over $x \in [-1.5, 1.5]$, and saturates beyond $x = \pm 5$. For all feedforward synapses, a range of x was first chosen that induced the desired output of ρ, following which the scale and shift parameters ν and η were computed so as to match it to V_{pre}'s range. The following parameters of the photoreceptor-L1 synapse relate to a stimulus intensity for which a 2° wide bar traveling at $12^\circ/s$ generated a $V_{pre} = V^{pr} - V^{fp}$ at the synapse of range $\pm 0.33mV$ about a baseline of $0mV$. Since the range of the V_{pre} scales linearly with the intensity, the value of ν can be scaled to evoke the same postsynaptic response for a different intensity value. We set $\langle \mu, \nu, \eta \rangle = \langle 40, 3, -3.82 \rangle$ for the photoreceptor-L1 synapse. The resultant postsynaptic response to the bar in the high conductance synaptic compartment was a hyperpolarization of $6.98mV$ followed by a depolarization of $5.31mV$ about an equilibrium resting potential of $-65.22mV$ (Fig.1b-ii). The η offset was chosen to model the slightly larger hyperpolarization reported in experiments [32]. As noted earlier, the entire three-compartment L1 was equipotential to within $0.3mV$.

L1's synapse on Mi1, C3, and L5 were modeled to be identical. Since the synapses are hyperpolarizing, we set $E_{syn} = -85mV$. We set $\langle \mu, \nu, \eta \rangle = \langle 5, 0.4, 27.4 \rangle$, which for the noted bar stimulus resulted in a depolarization of $10mV$ followed by a much smaller hyperpolarization of $1.3mV$ about an equilibrium resting potential of $-74.97mV$ (Fig.1b-ii). The values of μ and ν were set to mimic the response of Mi1 to flash stimuli [5]. The value of η, set close to saturation at the equilibrium $V_{pre} = -65.11mV$ of L1's distal compartment, yielded the intended half wave rectification.

L5's synapse on Mi4 being depolarizing and C3's synapse on Mi9 being hyperpolarizing,

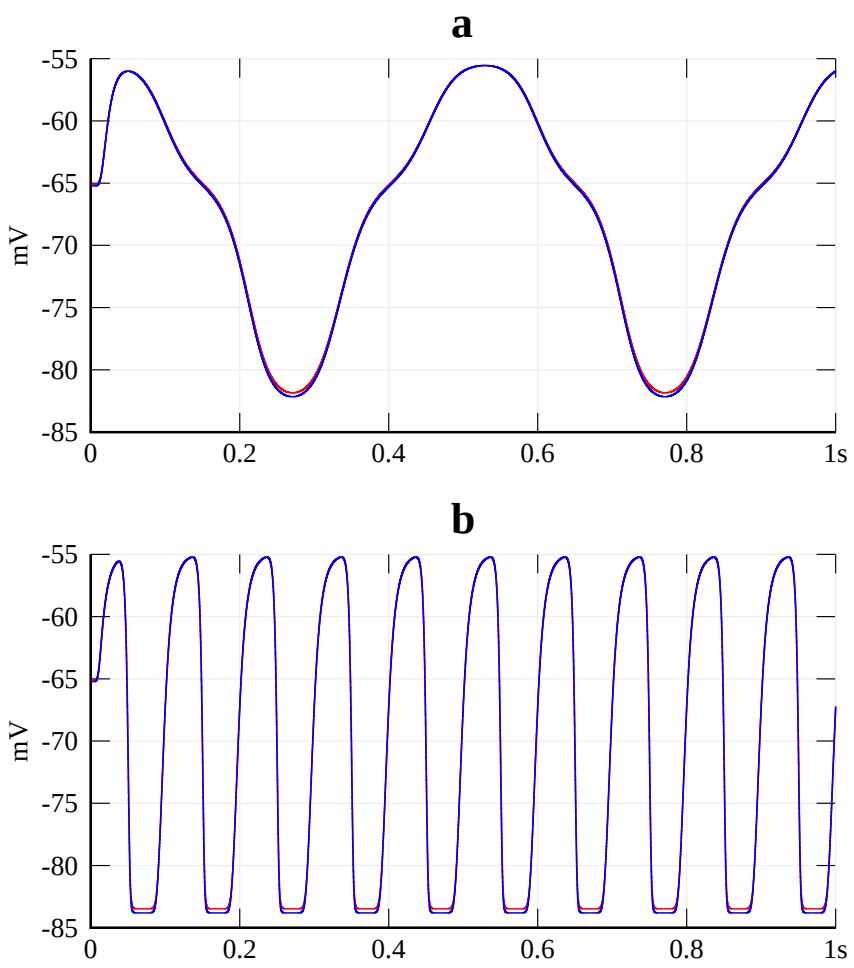
we set their $E_{syn} = 0mV$ and $-85mV$, respectively. Both synapses were assigned $\langle \mu, \nu, \eta \rangle = \langle 0.6, 0.3, 15 \rangle$. We found that either synapse could be eliminated and the other set at $\langle \mu, \nu, \eta \rangle = \langle 2, 0.3, 15 \rangle$ to replicate the dynamics in Fig.2a-i. The reciprocal synapses between Mi4 and Mi9 are both hyperpolarizing, hence we set both $E_{syn} = -85mV$. Their synaptic parameters were then arrived at via the optimization procedure. We found a large range of parameter settings for which the Mi4-Mi9 recurrent dynamics was delayed and extended (Fig.2a-i and iii). The model reported here (Fig.2a-i, Table.1) has $\langle \mu, \nu, \eta \rangle = \langle 4.03555, 0.29494, 15.10545 \rangle$ for Mi4's synapse on Mi9, and $\langle \mu, \nu, \eta \rangle = \langle 4.03574, 0.29485, 15.10667 \rangle$ for Mi9's synapse on Mi4. For the noted bar stimulus, this resulted in Mi4 depolarizing by $0.88mV$ about an equilibrium resting potential of $-60.32mV$, and Mi9 hyperpolarizing by $0.81mV$ about $-58.44mV$ (Fig.1b-ii). The delayed and extended dynamics evolved gradually with changing parameters during the optimization procedure (Fig.2a-iv). The stated parameter values were chosen to have the peak response of Mi4/Mi9 be delayed by $\approx 125ms$ with reference to that of Mi1. A different optimization run that achieved an even larger delay (Fig.2a-iii) gave $\langle \mu, \nu, \eta \rangle = \langle 2.34727, 0.28076, 15.42189 \rangle$ for Mi4's synapse on Mi9, and $\langle \mu, \nu, \eta \rangle = \langle 2.35028, 0.28072, 15.42252 \rangle$ for Mi9's synapse on Mi4.

Mi1's synapse on T4 being depolarizing, we set its $E_{syn} = 0mV$. Mi4's and Mi9's synapses on T4 being hyperpolarizing, we set their $E_{syn} = -85mV$. The home module Mi1's synapse on T4 was set at $\langle \mu, \nu, \eta \rangle = \langle 1.5, 0.1, 7.49 \rangle$. For cardinal direction 'a' (Fig.1b-i), the three Mi9 synapses on T4—from Mi9s of three contiguous modules—were set at $\langle \mu, \nu, \eta \rangle = \langle 1.5, 1.0, 58.44 \rangle$. Additionally, the three Mi4 synapses on T4—from Mi4s of the three symmetrically opposing contiguous modules—were set at $\langle \mu, \nu, \eta \rangle = \langle 1.5, 1.0, 60.32 \rangle$. For cardinal direction 'b' (Fig.1b-i), the μ's were raised to 2.25 along the PD-ND axis and dropped to 0 orthogonal to it. The μ's in both cases were set to mimic the reported response of T4 to the noted bar stimulus [13]. The ν's for the Mi4 and Mi9 synapses on T4 were set such that at

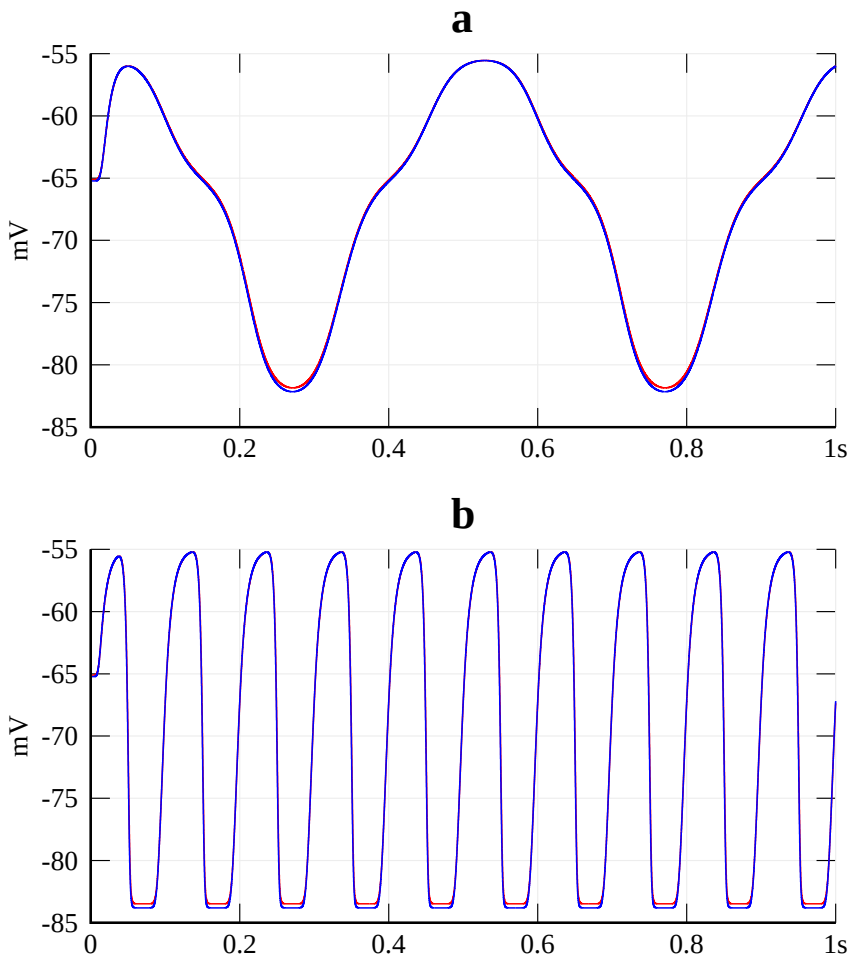
their charged-up peak response to a $1Hz$ drifting sinusoidal grating, amounting to a shift of $\approx 2.8mV$ (Fig.2b-i), the synapses reached saturation and closure, respectively.

Equilibrium Resting Potentials

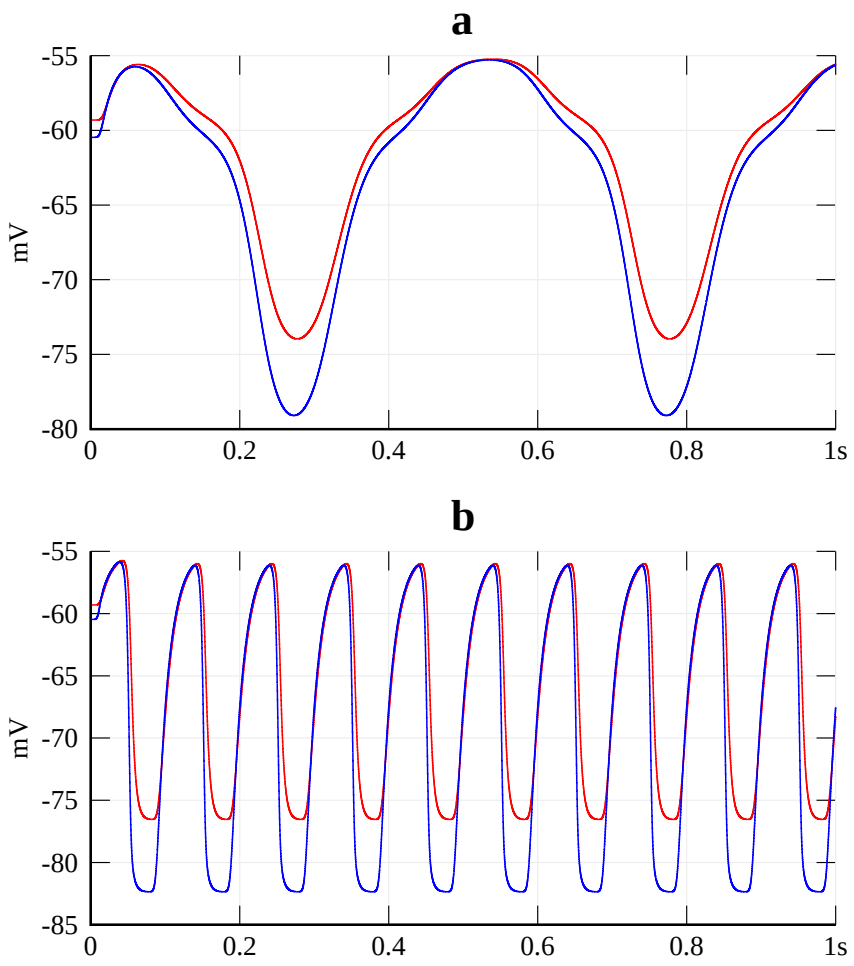
We began by setting the photoreceptor potential to a constant 0 over all time. The network is at its equilibrium resting state when $dV^i/dt = 0$ for all compartments i. Setting $dV^i/dt = 0$ for all i turns the differential-algebraic system of equations into an algebraic system of equations. We initialized this system at $V^i = V^i_{in} = V^i_{out} = E^i_l$ for all i, and solved it using the fixed point Jacobi iteration scheme. The iterations were continued until convergence, set as all new values evaluating to within $10^{-9}mV$ of their previous values.



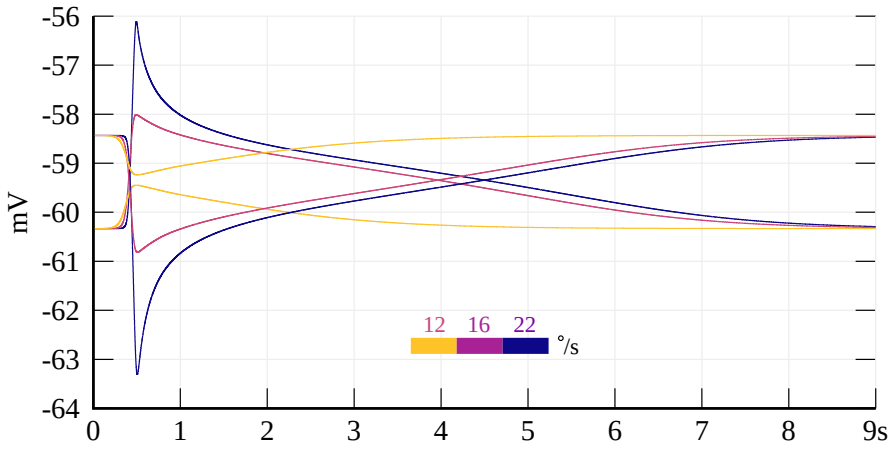
Supplementary Figure 1: **Neurons were almost equipotential across their entire cell body.** A $d = 1\mu m$, $l = 115\mu m$ L1, built from ten $l = 10\mu m$ and one $l = 15\mu m$ high conductance synaptic zone compartment. Responses of the most distal of the ten (red) and the initial synaptic compartment (blue), to a wavelength $= 20°$ square wave traveling at **a**, $2°/s$ and **b,** $10°/s$. Responses of all other compartments were in between. All other neurons, modeled with $d = 0.6\mu m$, $l = 60\mu m$, were even closer to being equipotential due to their larger axial to leak conductance ratios.



Supplementary Figure 2: **Assembly using fewer longer compartments is viable.** The L1 in Fig.S1 built from a $l = 40\mu m$ linked to a $l = 60\mu m$ linked to the $l = 15\mu m$ high conductance synaptic zone compartment. Responses of the $l = 40\mu m$ (red) and the initial synaptic compartment (blue), to the same square wave stimuli in Fig.S1, i.e., traveling at **a**, $2°/s$ and **b,** $10°/s$. To aid in comparison, we note that the responses of the high conductance synaptic zone compartment (blue) are identical to those in Fig.S1.



Supplementary Figure 3: **Cell body does not generate the requisite delay.** A $d = 1\mu m$, $l = 515\mu m$ L1, built from fifty $l = 10\mu m$ and one $l = 15\mu m$ high conductance synaptic zone compartment. Responses of the most distal of the fifty (red) and the initial synaptic compartment (blue), to a wavelength $= 20°$ square wave traveling at **a**, $2°/s$ and **b,** $10°/s$. Responses of all other compartments were in between. Although the responses are shifted toward $E_l = -55mV$, the peak to peak delay between the proximal and distal compartments is $\approx 5ms$.



Supplementary Figure 4: **Mi4-Mi9 network response was identical to when neurons were assembled using shorter compartments.** Neurons C3, L5, Mi4, and Mi9, modeled as $d = 0.6\mu m$, $l = 60\mu m$, were each built from six $l = 10\mu m$ compartments. The responses were identical to within precision bounds as compared to the models where the neurons were each built out of two $l = 30\mu m$ compartments. Raw responses of Mi4 and Mi9 are shown for a subset of velocities of the bar stimuli, for comparison against Fig.2a-i.

Acknowledgments

Funding: This work was supported by a grant from the U.S Air Force Office of Scientific Research (FA9550-16-1-0135).

Competing Interests: The author declares that he has no competing financial interests.

Data availability: The data that support the findings of this study are publicly and openly available at https://doi.org/10.6084/m9.figshare.c.4503218.v2

References

[1] Egemen Agi, Marion Langen, Steven J. Altschuler, Lani F. Wu, Timo Zimmermann, and Peter Robin Hiesinger. The evolution and development of neural superposition. *J. Neuro-*

genet., 28:216–232, 2014.

[2] Georg Ammer, Aljoscha Leonhardt, Armin Bahl, Barry J. Dickson, and Alexander Borst. Functional specialization of neural input elements to the drosophila on motion detector. *Curr. Biol.*, 25:2247–2253, 2015.

[3] Alexander Arenz, Michael S. Drews, Florian G. Richter, Georg Ammer, and Alexander Borst. The temporal tuning of the drosophila motion detectors is determined by the dynamics of their input elements. *Curr. Biol.*, 27:929–944, 2017.

[4] H. B. Barlow and W. R. Levick. The mechanism of directionally selective units in rabbits retina. *Journal of Physiology*, 178:477–504, 1965.

[5] Rudy Behnia, Damon A. Clark, Adam G. Carter, Thomas R. Clandinin, and Claude Desplan. Processing properties of on and off pathways for drosophila motion detection. *Nature*, 512:427–430, 2014.

[6] Alexander Borst, Juergen Haag, and Dierk F. Reiff. Fly motion vision. *Annual Review of Neuroscience*, 33:49–70, 2010.

[7] V. Braitenberg. Patterns of projection in the visual system of the fly. i. retina-lamina projections. *Exp Brain Res*, 3:271–298, 1967.

[8] Damon A. Clark, Limor Bursztyn, Mark A. Horowitz, Mark J. Schnitzer, and Thomas R. Clandinin. Defining the computational structure of the motion detector in drosophila. *Neuron*, 70:1165–1177, 2011.

[9] Peter Dayan and L. F. Abbott. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. The MIT Press, 2005.

[10] John E Dowling. *The retina: an approachable part of the brain*. Harvard University Press, 1987.

[11] N. Franceschini and K. Kirschfeld. Étude optique in vivo des éléments photorécepteurs dans l'oeil composé de drosophila. *Kybernetik*, 8:1–13, 1971.

[12] Nathan W. Gouwens and Rachel I. Wilson. Signal propagation in drosophila central neurons. *J. Neurosci.*, 29(19):6239–6249, 2009.

[13] Eyal Gruntman, Sandro Romani, and Michael B. Reiser. Simple integration of fast excitation and offset, delayed inhibition computes directional selectivity in drosophila. *Nat. Neurosci.*, 21:250–257, 2018.

[14] Juergen Haag, Alexander Arenz, Etienne Serbe, Fabrizio Gabbiani, and Alexander Borst. Complementary mechanisms create direction selectivity in the fly. *eLife*, 5:e17421, 2016.

[15] B. Hassenstein and W. Reichardt. Systemtheoretische analyse der zeit-, reihenfolgen- und vorzeichenauswertung bei der bewegungsperzeption des rüsselkäfers chlorophanus. *Zeitschrift Für Naturforschung*, 11:513–524, 1956.

[16] M. Joesch, B. Schnell, S. V. Raghu, D. F. Reiff, and A. Borst. On and off pathways in drosophila motion vision. *Nature*, 468:300–304, 2010.

[17] Mikko Juusola, Andrew S. French, Raimo O. Uusitalo, and Matti Weckström. Information processing by graded-potential transmission through tonically active synapses. *Trends Neurosci.*, 19(7):292–297, 1996.

[18] Mikko Juusola, Jeremy E. Niven, and Andrew S. French. Shaker K^+ channels contribute early nonlinear amplification to the light response in drosophila photoreceptors. *J. Neurophysiol.*, 90:2014–2021, 2003.

[19] Jonathan Chit Sing Leong, Jennifer Judson Esch, Ben Poole, Surya Ganguli, and Thomas Robert Clandinin. Direction selectivity in drosophila emerges from preferred direction enhancement and null-direction suppression. *J. Neurosci.*, 36(31):8078–8092, 2016.

[20] Ian A. Meinertzhagen and S. D. ONeil. Synaptic organization of columnar elements in the lamina of the wild type in drosophila melanogaster. *J Comp Neurol*, 305(2):232–263., 1991.

[21] Jeremy E. Niven, Mikko Vähäsöyrinki, Mika Kauranen, Roger C. Hardie, Mikko Juusola, and Matti Weckström. The contribution of shaker K^+ channels to the information capacity of drosophila photoreceptors. *Nature*, 421:630–634, 2003.

[22] Marta Rivera-Alba, Shiv N. Vitaladevuni, Yuriy Mishchenko, Zhiyuan Lu, Shin ya Takemura, Lou Scheffer, Ian A. Meinertzhagen, Dmitri B. Chklovskii, and Gonzalo G. de Polavieja. Wiring economy and volume exclusion determine neuronal placement in the drosophila brain. *Curr. Biol.*, 21:2000–2005, 2011.

[23] Wilson J. Rugh. *Linear system theory*. Prentice Hall,, Upper Saddle River, N.J., 2nd ed. edition, 1996.

[24] Juha Rusanen and Matti Weckström. Frequency-selective transmission of graded signals in large monopolar neurons of blowfly calliphora vicina compound eye. *J. Neurophysiol.*, 115:2052–2064, 2016.

[25] B. Schnell, M. Joesch, F. Forstner, S. V. Raghu, H. Otsuna, K. Ito, A. Borst, and D. F. Reiff. Processing of horizontal optic flow in three visual interneurons of the drosophila brain. *J. Neurophysiol.*, 103:1646–1657, 2010.

[26] Kazunori Shinomiya, Gary Huang, Zhiyuan Lu, Toufiq Parag, C. Shan Xu, Roxanne Aniceto, Namra Ansari, Natasha Cheatham, Shirley Lauchie, Erika Neace, Omotara Ogundeyi, Christopher Ordish, David Peel, Aya Shinomiya, Claire Smith, Satoko Takemura, Iris Talebi, Patricia K. Rivlin, Aljoscha Nern, Louis K. Scheffer, Stephen M. Plaza, and Ian A. Meinertzhagen. Comparisons between the on- and off- edge motion pathways in the drosophila brain. *eLife*, 8:e40025, 2019.

[27] D. G. Stavenga. Angular and spectral sensitivity of fly photoreceptors. ii. dependence on facet lens f-number and rhabdomere type in drosophila. *J Comp Physiol A*, 189:189–202, 2003.

[28] James A. Strother, Shiuan-Tze Wu, Allan M. Wong, Aljoscha Nern, Edward M. Rogers, Jasmine Q. Le, Gerald M. Rubin, and Michael B. Reiser. The emergence of directional selectivity in the visual motion pathway of drosophila. *Neuron*, 94(1):168–182.e.10, 2017.

[29] Shin Ya Takemura, Aljoscha Nern, Dmitri B. Chklovskii, Louis K. Scheffer, Gerald M. Rubin, and Ian A. Meinertzhagen. The comprehensive connectome of a neural substrate for 'on' motion detection in drosophila. *eLife*, 6:e24394, 2017.

[30] John C. Tuthill, Aljoscha Nern, Stephen L. Holtz, Gerald M. Rubin, and Michael B. Reiser. Contributions of the 12 neuron classes in the fly lamina to motion vision. *Neuron*, 79(10):128–140, 2013.

[31] Matti Weckström and Simon Laughlin. Extracellular potentials modify the transfer of information at photoreceptor output synapses in the blowfly compound eye. *J. Neurosci.*, 30(28):9557–9566, 2010.

[32] Lei Zheng, Anton Nikolaev, Trevor J. Wardill, Cahir J. O'Kane, Gonzalo G. de Polavieja, and Mikko Juusola. Network adaptation improves temporal representation of naturalistic stimuli in drosophila eye: I dynamics. *PLoS One*, 4(1):e4307, 2009.

[33] Zhihao Zheng, J Scott Lauritzen, Eric Perlman, Camenzind G Robinson, Matthew Nichols, Daniel Milkie, Omar Torrens, John Price, Corey B Fisher, Nadiya Sharifi, et al. A complete electron microscopy volume of the brain of adult drosophila melanogaster. *Cell*, 174(3):730–743, 2018.