

A minimal synaptic model for direction selective neurons in *Drosophila*

Jacob A. Zavatone-Veth

Department of Molecular, Cellular, and Developmental
Biology, Yale University, New Haven, CT, USA
Department of Physics,
Yale University, New Haven, CT, USA

Bara A. Badwan

School of Engineering and Applied Science,
Yale University, New Haven, CT, USA

Damon A. Clark

Department of Molecular, Cellular, and Developmental
Biology, Yale University, New Haven, CT, USA
Department of Physics,
Yale University, New Haven, CT, USA
Interdepartmental Neuroscience Program,
Yale University, New Haven, CT, USA
Department of Neuroscience,
Yale University, New Haven, CT, USA

Visual motion estimation is a canonical neural computation. In *Drosophila*, recent advances have identified anatomic and functional circuitry underlying direction-selective computations. Models with varying levels of abstraction have been proposed to explain specific experimental results but have rarely been compared across experiments. Here we use the wealth of available anatomical and physiological data to construct a minimal, biophysically inspired synaptic model for *Drosophila*'s first-order direction-selective T4 cells. We show how this model relates mathematically to classical models of motion detection, including the Hassenstein-Reichardt correlator model. We used numerical simulation to test how well this synaptic model could reproduce measurements of T4 cells across many datasets and stimulus modalities. These comparisons include responses to sinusoid gratings, to apparent motion stimuli, to stochastic stimuli, and to natural scenes. Without fine-tuning this model, it sufficed to reproduce many, but not all, response properties of T4 cells. Since this model is flexible and based on straightforward biophysical properties, it provides an extensible framework for developing a mechanistic understanding of T4 neural response properties. Moreover, it can be used to assess the sufficiency of simple biophysical mechanisms to describe features of the direction-selective computation and identify where our understanding must be improved.

Introduction

Motion estimation is a canonical visual computation that requires integrating information nonlinearly over both time and space. Direction-selective signals are tuned to motion in a preferred-direction (PD), which elicits the strongest responses, while motion in the opposite, null-direction (ND), elicits a weaker response. This directional computation has been described by a wide variety of computational models. Classical models, such as the Hassenstein-Reichardt correlator (HRC) (Hassenstein & Reichardt, 1956) and motion energy model (Adelson & Bergen, 1985), rely on sensing correlations between pairs of points separated in space and time. These phenomenological models have provided striking insights into neural and behavioral responses in a variety of species, including in flies (Yang & Clandinin, 2018).

In the last decade, advances in defining the anatomical and functional connectivity of *Drosophila*'s visual circuits suggest that we may move toward more mechanistic, biophysical descriptions of this computation. Here, we follow previous work (Gruntman, Romani, & Reiser, 2018; Torre & Poggio, 1978) to propose a simple, biophysically plausible synaptic model for direction selectivity in *Drosophila*'s ON-edge-sensitive motion pathway. We compare its

Citation: Zavatone-Veth, J. A., Badwan, B. A., & Clark, D. A. (2020). A minimal synaptic model for direction selective neurons in *Drosophila*. *Journal of Vision*, 20(2):2, 1–22, <https://doi.org/10.1167/jov.20.2.2>.



predictions to measurements made by several research groups in response to many stimuli, giving us a tool for understanding which features are sufficient to describe different response properties.

The most peripheral direction-selective neurons in the *Drosophila* optic lobe are the T4 and T5 cells, which are sensitive to moving ON-edges (consisting of contrast increments) and OFF-edges (consisting of contrast decrements), respectively (Clark, Bursztyn, Horowitz, Schnitzer, & Clandinin, 2011; Joesch, Schnell, Raghu, Reiff, & Borst, 2010; Maisak, Haag, Ammer, Serbe, Meier, Leonhardt, Schilling, Bahl, Rubin, Nern, 2013). Electron microscopy and genetic silencing have identified the primary inputs to T4 and T5 cells (Serbe, Meier, Leonhardt, & Borst, 2016; Shinomiya, Huang, Lu, Parag, Xu, Aniceto, Ansari, Cheatham, Lauchie, & Neace, 2019; Strother, Wu, Wong, Nern, Rogers, Le, Rubin, Reiser, 2017; Takemura, Nern, Chklovskii, Scheffer, Rubin, & Meinertzhagen, 2017). These studies suggest that T4 cells receive input from three distinct colinear spatial locations, with the neurons Mi1 and Tm3 both relaying information about the central point, and the neurons Mi9 and Mi4 acting as relays for the two flanking points (Takemura et al., 2017) (Figure 1A). The neuron T5 appears to have a similar spatial structure, with different input neurons (Shinomiya et al., 2019). Both cell types also receive spatially-localized inputs from other neurons, whose functions remain less well understood (Shinomiya et al., 2019; Takemura et al., 2017).

The physiological properties of the inputs to T4 and T5 cells have also been characterized. At their receptive field centers, Mi1 and Tm3 cells respond quickly to visual stimuli and provide excitatory input to T4 (Arenz, Drews, Richter, Ammer, & Borst, 2017; Behnia, Clark, Carter, Clandinin, & Desplan, 2014; Gruntman et al., 2018; Strother et al., 2017; Takemura et al., 2017). On the preferred direction side of the receptive field, the cells Mi4 and CT1 show ON responses with slower kinetics, likely inhibiting T4 cells (Arenz et al., 2017; Meier & Borst, 2019; Shinomiya et al., 2019; Takemura et al., 2017). On the null direction side of the receptive field, Mi9 cells are delayed OFF cells, which likely provide inhibitory, glutamatergic input to T4 cells (Arenz et al., 2017; Salazar-Gatzimas, Agrochao, Fitzgerald, & Clark, 2018). The inputs to T5 cells similarly appear to be arranged with a fast central input and delayed flanking inputs, but whether these inputs excite or inhibit T5 is less clear (Arenz et al., 2017; Behnia et al., 2014; Shinomiya et al., 2019; Wienecke, Leong, & Clandinin, 2018).

This wealth of data about T4 and T5 cells and their inputs has led to many different models that seek to describe the response properties of T4 and T5 cells (Arenz et al., 2017; Badwan, Creamer, Zavatone-Veth, & Clark, 2019; Behnia et al., 2014; Clark et al., 2011; Creamer, Mano, & Clark, 2018;

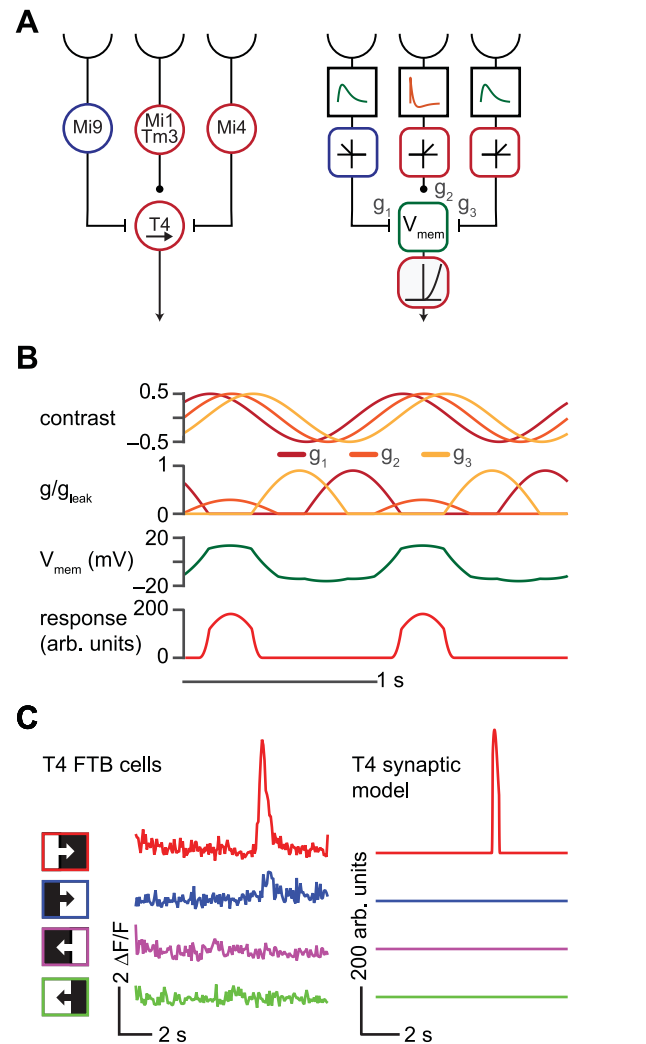


Figure 1. An anatomically constrained synaptic model for T4 cells. (A) *Left*: Diagram of proposed inputs to *Drosophila* T4 first-order direction-selective cells based on anatomic and physiological measurements. Mi1 and Tm3 cells provide ON excitatory input at the center of the receptive field of each T4 cell, whereas Mi9 provides delayed OFF inhibitory input offset in the null direction, and Mi4 provides delayed ON inhibitory input offset in the preferred direction. *Right*: Synaptic model based on the anatomic structure shown at left. (B) Responses of each component of the synaptic model to a 1 Hz, 45-degree sinusoidal grating drifting in the preferred (rightward) direction. *Top*: Input contrasts to each of the 3 presynaptic units of the model. *Upper middle*: Conductances of excitatory and inhibitory currents corresponding to each input in response to the sinusoidal stimulus. *Lower middle*: Membrane voltage. *Bottom*: Calcium signal. (C) *Left*: Responses of T4 cells sensitive to front-to-back (FTB) motion to ON and OFF edges moving FTB and back-to-front (BTF) at 30 degrees/s, measured using 2-photon calcium imaging (data from (Salazar-Gatzimas et al., 2016)). *Right*: As at left, but for the T4 synaptic model.

Eichner, Joesch, Schnell, Reiff, & Borst, 2011; Gruntman et al., 2018; Haag, Arenz, Serbe, Gabbiani, & Borst, 2016; Leong, Esch, Poole, Ganguli, & Clandinin, 2016; Leonhardt, Ammer, Meier, Serbe, Bahl, & Borst, 2016; Salazar-Gatzimas et al., 2018; Salazar-Gatzimas, Chen, Creamer, Mano, Mandel, Matulis, Pottackal, & Clark, 2016; Serbe et al., 2016; Strother et al., 2017; Wienecke et al., 2018). Many of the measurements of T4 and T5 have demonstrated phenomenology that could not be produced by the classical HRC model. However, proposed models were most often evaluated by how they reproduced the associated dataset, rather than the full range of phenomena documented in the literature. Here, we ask how a minimal, constrained model reproduces T4 phenomenology (and some T5 phenomenology) from many different experiments. We compare the model to data in response to moving edges, to sinusoids, to apparent motion stimuli, to stochastic stimuli, and to natural scenes.

In this minimal model, the spatially-separated inputs to T4 are represented as three linear-nonlinear (LN) transformations of the input contrast (Dayan & Abbott, 2001). These model neurons then interact with T4 by altering the conductance of excitatory and inhibitory currents (Gruntman et al., 2018; Torre & Poggio, 1978). This construction is simple enough to allow some algorithmic intuition but incorporates greater biophysical realism than most phenomenological models. We do not fit the model to every dataset. Rather, our goal is to test the sufficiency of a minimal circuit model to account for different measured phenomena in T4 cells. This model does not contain any exotic channels or receptors, and it biophysically models the membrane voltage and intracellular calcium concentration in T4 neurons. It does not reproduce all functional properties of T4 cells, but it provides a flexible framework for understanding the sufficiency of simple circuit properties and mechanisms to describe the processing properties of T4 neurons. In cases where this model is insufficient to describe data, we suggest how model parameters might be changed to better account for it.

Methods

Constructing an anatomically constrained synaptic model for T4 cells

Following proposed synaptic architectures for direction-selective computations (Gruntman et al., 2018; Torre & Poggio, 1978), we constructed an elementary motion detector based on the connectome of the *Drosophila* optic lobe. We simplified this structure to consider three inputs to a T4 cell: a delayed ND-offset

OFF inhibitory input representing Mi9, a centered ON excitatory input representing Mi1 and Tm3, and a delayed PD-offset ON inhibitory input representing Mi4 (or CT1 or both) (Figure 1A) (Strother et al., 2017; Takemura et al., 2017).

We will model these inputs to T4 cells as simple linear-nonlinear (LN) transformations of the input contrast (Behnia et al., 2014). We will further model effects of these synaptic inputs on the membrane potential of the T4 cell by changes in the conductance of excitatory and inhibitory currents (Torre & Poggio, 1978). For notational convenience, we define our model below in continuous space and time, noting where adjustments are made for the discretization used in numerical simulation. We take the inputs to the model to be contrasts. We take each input to the motion detector to have an L_1 -normalized Gaussian spatial acceptance function

$$h(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}.$$

To represent the delayed inputs to the motion detector, we use the L_2 -normalized lowpass temporal filter

$$f(t) = 2\tau^{-\frac{3}{2}} t \Theta(t) e^{-\frac{t}{\tau}},$$

where $\Theta(x)$ is the Heaviside step function. To represent the nondelayed central input to the motion detector, we replace the temporal filter f by its normalized distributional derivative $\dot{f} = 2\tau^{-\frac{3}{2}}(\tau - t)\Theta(t)e^{-\frac{t}{\tau}}$. Using these filters, we define the filtered contrast signal s at each point in spacetime:

$$s(t, x) := (fh * c)(t, x),$$

where $c(t, x)$ is the input contrast and $*$ denotes spatiotemporal convolution with the functions f and h over the appropriate domain. Because taking the temporal derivative of the filtered contrast signal is equivalent to filtering with the derivative of the temporal filter, we will use the notation \dot{s} for the high-pass-filtered signal throughout. For convenient handling of spatial boundary conditions, we numerically simulate the full 360 degrees of visual space, which is a periodic interval.

Then, we define the three inputs to the motion detector as rectified-linear functions of the filtered contrast signal at three points in space, mimicking the polarity-selectivity of the inputs to T4 cells:

$$g_1(t, x) := g_{\text{inh}} R(-s(t, x - \Delta)),$$

$$g_2(t, x) := g_{\text{exc}} R(\dot{s}(t, x)),$$

$$g_3(t, x) := g_{\text{inh}} R(s(t, x + \Delta)),$$

where Δ is the spacing between neighboring inputs, $R(x) := \max\{0, x\}$ is the ramp function, and g_{inh} and g_{exc} are parameters scaling the effects of each input on the postsynaptic conductances (Figure 1A–B). Thus we represent the conductances as linear-nonlinear (LN) transformations of the input contrast (Dayan & Abbott, 2001).

We define the membrane potential V_m of the postsynaptic cell such that the reversal potential for leak currents is 0 mV. The cell's membrane voltage dynamics are then given as (Torre & Poggio, 1978)

$$c_m \dot{V}_m + V_m (g_{\text{leak}} + g_1 + g_2 + g_3) \\ = g_1 E_{\text{inh}} + g_2 E_{\text{exc}} + g_3 E_{\text{inh}},$$

where c_m is the membrane capacitance, g_{leak} is the leak conductance, and E_{inh} and E_{exc} are the reversal potentials for inhibitory and excitatory currents, respectively. We follow previous work to neglect capacitive currents in T4 cells (Gruntman et al., 2018), and we solve for the pseudo-steady-state membrane voltage (Gruntman et al., 2018; Torre & Poggio, 1978)

$$V_m = \frac{g_1 E_{\text{inh}} + g_2 E_{\text{exc}} + g_3 E_{\text{inh}}}{g_{\text{leak}} + g_1 + g_2 + g_3}.$$

Then, we model the transformation from membrane voltage to calcium concentration C as a positively rectifying half-quadratic function $R^2(x) := (R(x))^2$:

$$C(t, x) := R^2(V_m(t, x)),$$

which qualitatively captures the expansive nonlinear effect of the transformation between voltage and calcium (Kato, Xu, Cho, Abbott, & Bargmann, 2014; Leong et al., 2016) (Figure 1B).

Visual stimuli

We presented this model with spatiotemporal contrast patterns to mimic a variety of visual stimuli used in the field. Detailed mathematical descriptions of each stimulus are given in Appendix A. Briefly, we presented the model with moving and stationary sinusoidal gratings, with apparent motion stimuli, with stochastic stimuli including those with imposed correlations, and with natural scenes. In each case, we compared how the model responds to the published responses of T4 and T5 neurons.

Selecting model parameters

We divide the parameters of the T4 synaptic model into two conceptual classes: those that describe the spatiotemporal properties of the inputs to the model

cell, and those that describe how the inputs interact postsynaptically. We sought to maximally constrain our choices of model parameters by published physiological measurements and to show that the model is relatively robust to the specific values of remaining unconstrained parameters. The set of input parameters used here is equal to that used in a model developed to explain direction-opponency in T4 cells (Badwan et al., 2019). To approximately match the spatial acceptance functions of photoreceptors in the fly eye (Stavenga, 2003), we set the full width at half maximum of the input spatial filters to 5.7 degrees. We use 5 degree spacing between inputs such that they evenly tile 360 degrees of visual space, which approximately matches the 5.1-degree spacing between adjacent ommatidia (Stavenga, 2003). Broadening the input spatial receptive fields or decreasing the input spacing would increase the correlation between the input conductances in this model. We would expect that to affect the spatiotemporal correlations to which the model is sensitive (Badwan et al., 2019; Fitzgerald & Clark, 2015). Previous functional studies of the inputs to T4 cells have suggested that their spatial receptive fields may be somewhat wider than that of one photoreceptor (Arenz et al., 2017; Behnia et al., 2014; Gruntman et al., 2018), but we use this simplified spatial processing because our goal is to construct a minimal model. To produce peak responses to PD sinusoidal gratings at approximately 1 Hz (Badwan et al., 2019; Maisak et al., 2013; Strother et al., 2017), we fixed the time constant of the temporal filters to 150 ms. The true temporal filters are unlikely to have same time constants or be perfect derivatives of one another (Arenz et al., 2017; Behnia et al., 2014). However, this tuning suffices to explain some basic properties and allows the construction of a model with minimal fitting. These spatial and temporal filters set the tuning of the model inputs, so changing the parameters or functional forms will alter the spatiotemporal tuning of the model.

The remaining model parameters set the weighting of the inputs in their nonlinear postsynaptic interactions. We fix the excitatory and inhibitory reversal potentials to values of $E_{\text{exc}} = 60$ mV and $E_{\text{inh}} = -30$ mV, which are plausible based on electrophysiological experiments (Gruntman et al., 2018). The model membrane potential can be written as

$$V_m = \left(\frac{g_1}{g_{\text{leak}}} E_{\text{inh}} + \frac{g_2}{g_{\text{leak}}} E_{\text{exc}} + \frac{g_3}{g_{\text{leak}}} E_{\text{inh}} \right) \\ \times \left(1 + \frac{g_1}{g_{\text{leak}}} + \frac{g_2}{g_{\text{leak}}} + \frac{g_3}{g_{\text{leak}}} \right)^{-1};$$

hence, only the ratios of g_1 , g_2 , and g_3 to g_{leak} , rather than their absolute magnitudes, are relevant. We therefore express the postsynaptic conductances as nondimensional quantities in units of g_{leak} , leaving $g_{\text{exc}}/g_{\text{leak}}$ and $g_{\text{inh}}/g_{\text{leak}}$ as the model's two free

Parameter	Value
Photoreceptor spacing Δ	5°
Photoreceptor spatial FWHM	5.7°
Temporal filter time constant τ	0.150 s
Leak current reversal potential E_{leak}	0 mV
Excitatory current reversal potential E_{exc}	+60 mV
Inhibitory current reversal potential E_{inh}	−30 mV
Excitatory to leak conductance ratio $g_{\text{exc}}/g_{\text{leak}}$	0.1
Inhibitory to leak conductance ratio $g_{\text{inh}}/g_{\text{leak}}$	0.3

Table 1. Model parameter values used in all simulations.

parameters. The procedure used to select the values of these parameters is described in detail in Appendix B. As shown previously (Badwan et al., 2019), there exists a broad region of parameter space for which this model displays strong directional responses to sinusoid gratings with a temporal frequency of 1 Hz and a spatial wavelength of 45 degrees. We note that our choice of filter normalization, which differs from that in the previous use of this model (Badwan et al., 2019), affects the parameter values chosen, as it scales g_1 , g_2 , and g_3 relative to g_{leak} . Table 1 summarizes the model parameter values used in all simulations.

Numerical methods

Numerical simulations were conducted using Matlab 9.6 (R2019a) (The MathWorks, Natick, MA, USA). In our simulations, we discretized space and time using a spatial sampling interval of 0.5 degrees and a temporal sampling interval of 1/240 s. Our results are robust to the simulation resolution so long as the shortest spatial and temporal scales represented in the stimulus set can be resolved. For stimuli containing randomly-generated components, responses were averaged over 1000 realizations, and bootstrapped 95% confidence intervals for the mean were computed using the bias-corrected and accelerated percentile method (Efron, 1987). All code is publicly available from GitHub at <https://www.github.com/ClarkLabCode/SynapticModel>.

In vivo two-photon calcium imaging in T4 cells

Most of our comparisons relate the synaptic model's responses to published data, but we also compare the model to a new dataset of T4 cell responses to glider stimuli. The protocol for 2-photon calcium imaging in T4 cells matches published methods (Badwan et al., 2019) and used Psychtoolbox (Brainard, 1997; Kleiner, Brainard, Pelli, Ingling, Murray, & Broussard, 2007; Pelli, 1997) to present stimuli on a panoramic visual display (Creamer, Mano, Tanaka, & Clark, 2019). The glider stimuli presented during these measurements

are described in Appendix A. Net responses were computed as the difference in responses to stimuli moving in the preferred and null directions of each T4 region of interest and then averaged within each fly. Nonparametric 2-sided Wilcoxon signed-rank tests were used to test whether median net responses differed significantly from zero (Hollander, Wolfe, & Chicken, 2013). For statistical purposes, each individual fly was considered to be an independent sample.

Results

The synaptic model reduces to HRC-like terms

To gain intuition about the operation of the T4 synaptic model, we consider its expansion in the small-input limit. To do so, we approximate the ramp function nonlinearity with a smooth function that represents a soft rectifier, which can be approximated by a linear function for small inputs (Fitzgerald & Clark, 2015). In particular, letting $s_1(t) = s(t, x - \Delta)$, $s_2(t) = s(t, x)$, and $s_3(t) = s(t, x + \Delta)$, and defining the non-negative constants $\alpha = |g_{\text{inh}}E_{\text{inh}}/g_{\text{leak}}|$ and $\gamma = |g_{\text{exc}}E_{\text{exc}}/g_{\text{leak}}|$, we have, to lowest order in the inputs,

$$C \approx \frac{1}{16}(\alpha s_1 + \gamma \dot{s}_2 - \alpha s_3)^2,$$

which may be rewritten as

$$C \approx \frac{\alpha^2}{16}(s_1 - s_3)^2 + \frac{\gamma^2}{16}(\dot{s}_2)^2 + \frac{\alpha\gamma}{8}(s_1\dot{s}_2 - \dot{s}_2s_3).$$

This expansion represents a motion-energy approximation of the model. The first term in this expansion is a finite-difference approximation to a spatial derivative, whereas the second term is a temporal derivative at the center of the model's receptive field. The third term, which is the only direction-selective term, is the subtraction of two offset correlators with opposite directional tuning. This subtraction step provides some intuition for why this model mimics some properties of a fully-opponent HRC model (Badwan et al., 2019). This same direction-selective term also appears in the second-order expansion of the membrane voltage. Because this expansion of the model is only to second-order, it is invariant under contrast inversions and cannot account for properties such as ON-edge selectivity or responses to triplet correlations (Clark et al., 2011; Clark et al., 2014; Fitzgerald & Clark, 2015; Fitzgerald, Katsov, Clandinin, & Schnitzer, 2011; Joesch et al., 2010; Maisak et al., 2013). Although this simple description does not capture all properties

of the synaptic model, it provides intuition for the sensitivity of the model to certain stimulus features.

The synaptic model is strongly ON-edge- and direction-selective

T4 and T5 neurons are distinguished by the fact that T4 cells respond to ON-edges whereas T5 cells respond to OFF-edges (Maisak et al., 2013). We first compared the ON/OFF-edge- and direction-selectivity of our T4 synaptic model to responses measured using 2-photon calcium imaging in T4 cells sensitive to front-to-back (FTB) motion. Like T4 FTB cells, our synaptic model responded strongly to an ON-edge moving in the FTB direction but displayed little or no response to OFF-edges moving in the FTB direction or to edges of either polarity moving in the back-to-front direction (Figure 1C) (Maisak et al., 2013; Salazar-Gatzimas et al., 2016).

The spatiotemporal tuning of the synaptic model is consistent with that of T4 cells

Sinusoid grating stimuli are a common tool for characterizing direction-selective computations. Responses to these stimuli have been used to suggest that the membrane voltage of T5 cells is a nearly linear transformation of the visual input (Figure 2A) (Wienecke et al., 2018). In the synaptic model, the membrane voltage is a nonlinear function of the input contrast because the inputs are first rectified and then interact nonlinearly. We applied the same linearity testing protocol to our model membrane voltage, constructing predictions for responses to PD and ND drifting gratings from the responses to counterphase gratings (see Appendix A) (Jagadeesh, Wheat, & Ferster, 1993; Wienecke et al., 2018). The responses of the T4 synaptic model to drifting gratings were similar to those predicted by a linear model for membrane voltage (Figure 2A). Thus even a nonlinear system like the T4 synaptic model may appear reasonably linear by this protocol.

T4 and T5 cells display direction-opponent average calcium responses to sinusoid gratings (Figure 2B) (Badwan et al., 2019). This property means that the average response to PD motion is reduced by the addition of ND motion, imposing a strong constraint on models for the direction-selective computation. In particular, it implies that linear-nonlinear models with expansive nonlinearities cannot account for the response properties of these cells (Badwan et al., 2019). A variant of this synaptic model was proposed to account for these direction-opponent responses (Badwan et al., 2019). This model reproduces the strong suppression when ND motion is added to

PD motion without substantial enhancement when orthogonal-direction (OD) motion is added to PD motion (Figure 2B).

T4 and T5 cells are tuned to the temporal frequency of sinusoidal stimuli (Figure 2C) (Creamer et al., 2018). This means that the mean neural response is maximal at a single temporal frequency, independent of the wavelength. This property also applies to measurements of fly behavior (Creamer et al., 2018; Kunze, 1961) and is consistent with the classical, fully-opponent HRC. We presented the T4 synaptic model with drifting gratings of different spatial and temporal frequencies to find the mean response to each. The model response was strongly temporal-frequency-tuned (Figure 2C). To quantify the temporal-frequency-tuning, we asked how much of the variance in this surface was accounted for by the product of one function of temporal frequency and 1 function of spatial frequency response (Creamer et al., 2018; Priebe, Lisberger, & Movshon, 2006). Such a separable model accounted for 99% of the variance in the response (see Appendix A, Figure 2C). Because of our choice of parameters, the input temporal filters in this model produce peak responses at around 1 Hz, lower than the roughly 2–4 Hz peak measured in these T4 cells.

T4 and T5 cells respond to static gratings with amplitudes that depend on the grating orientation (Fisher, Silies, & Clandinin, 2015) (Figure 2D). The preferred orientation (defined by the vector normal to the edges in a static grating) approximately matches the preferred direction of motion of these cells (Maisak et al., 2013). The convention we use here for defining the orientation of a static grating is rotated 90 degrees relative to that used in the original study, which defined orientation in terms of vectors parallel, rather than normal, to the edges (Fisher et al., 2015) (see Appendix A). When the T4 synaptic model was presented with both static and drifting gratings of many different orientations, it reproduced the orientation tuning observed experimentally for both static and moving gratings (Figure 2D). The model was more selective for both orientation and direction than the T4 cell measurements.

The synaptic model reproduces the selectivity of apparent motion responses in T4 cells

In addition to sinusoid gratings, apparent motion stimuli are a useful tool for investigating direction-selective systems. These stimuli decompose visual motion into summations of simpler spatiotemporal patterns, which can provide strong intuition about the motion computation (Barlow & Levick, 1965).

Electrophysiological measurements of T4 cells have shown fast depolarization and delayed, offset hyperpolarization in response to a small flashed white bar placed on a gray background (Gruntman et al.,

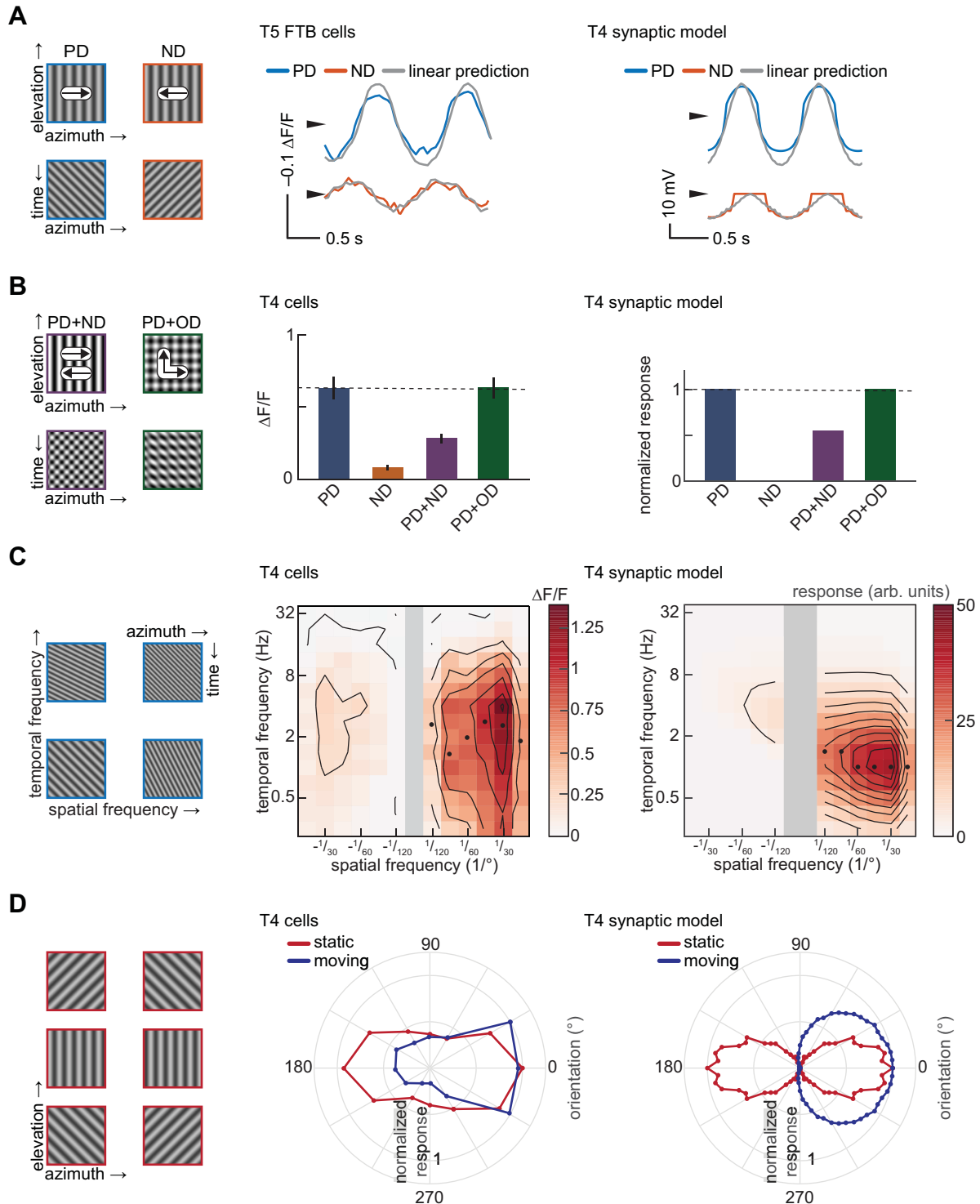


Figure 2. The T4 synaptic model reproduces sinusoidal grating responses measured in T4 cells. (A) *Left*: Images and kymographs of sinusoid gratings drifting in the preferred (PD) and null (ND) directions. *Center*: Membrane voltage of T5 FTB cells to 1 Hz, 25° drifting gratings compared with linear predictions from contrast-modulated counterphase gratings, measured using voltage indicators (data from Wienecke et al., 2018). *Right*: As at center, but for voltage responses of the T4 synaptic model. The model has coefficients of determination for PD and ND of 0.92 and 0.82. (B) Mean responses to 1 Hz, 45° sinusoid gratings. *Left*: Images and kymographs of composite sinusoid gratings containing PD and ND motion or PD and orthogonal direction (OD) motion. *Center*: Mean responses of T4 cells to drifting gratings, measured using a calcium indicator (data from Badwan et al., 2019). Error bars indicate ± 1 SEM. *Right*: As at center, but for calcium responses of the T4 synaptic model. (C) Spatiotemporal frequency tuning. *Left*: Kymographs of sinusoid gratings with varying spatiotemporal frequency content. *Center*: Spatiotemporal frequency tuning of T4 cells (data from

←
(Creamer et al., 2018)). Black circles indicate the temporal frequency at which the maximum response at a given spatial frequency is attained. *Right*: As at center, but for the T4 synaptic model. (D) Orientation and direction tuning. *Left*: Images of oriented sinusoid gratings. *Center*: Orientation tuning of T4 and T5 cells with static gratings (data from (Fisher et al., 2015)) and direction tuning of T4 cells with drifting gratings (data from Maisak et al., 2013). The orientation of a static grating is defined by the vector normal to the apparent edges, the same definition as for moving gratings (see Appendix A). *Right*: As at center, but for the T4 synaptic model.

2018) (Figure 3A). The synaptic T4 model displayed qualitatively consistent responses to the same stimulus (Figure 3A). The positive lobe in the model is narrower than in the electrophysiological recording; this is likely because the true central input to T4 has a wider receptive field than in our model (Behnia et al., 2014; Takemura, Bharioke, Lu, Nern, Vitaladevuni, Rivlin, Katz, Olbris, Plaza, & Winston, 2013). Consistent with electrophysiology, the OFF input to the T4 model is not visible under this analysis because it was rectified with a threshold at mean gray (zero contrast).

Because the synaptic model reproduced T4 cell voltage responses to flashed bars, we sought to characterize its responses to apparent motion stimuli composed of pairs of bars offset in spacetime (Salazar-Gatzimas et al., 2018). These stimuli can induce in humans and in flies the “reverse-phi” motion illusion, in which a reversal of contrast polarity induces a motion percept in the direction opposite the stimulus displacement (Anstis, 1970; Clark et al., 2011; Hassenstein & Reichardt, 1956). We aligned these stimuli so that the temporally delayed bar is placed at the center of the receptive field (Figure 3B) (Salazar-Gatzimas et al., 2018). T4 cells respond maximally to one phi and one reverse-phi apparent motion stimulus out of eight possible pairings (Salazar-Gatzimas et al., 2018). The synaptic model reproduced this selectivity (Figure 3C–D).

Various groups have assessed nonlinear enhancement or suppression of PD and ND apparent motion stimuli relative to linear decompositions. This analysis can be misleading because it does not allow one to uniquely characterize the nonlinearity as “enhancing” or “suppressing,” because there exist an infinite number of linear decompositions of a given stimulus (Salazar-Gatzimas et al., 2018). Despite this difficulty, such analyses have been applied as an intuitive way to try to understand direction-selective computations (Barlow & Levick, 1965; Fisher et al., 2015; Gruntman et al., 2018; Haag et al., 2016).

In T4 cells, an analysis of responses to sequential bars has indicated that calcium signals include both PD enhancement and ND suppression relative to a linear prediction from the responses to individual bars (Haag et al., 2016) (Figure 3E). Our model failed to reproduce this result, showing only suppression of ND motion under this analysis (Figure 3E). This discrepancy is likely influenced by the timescale of this stimulus, which is far longer than the 150 ms offset

used in the apparent motion stimuli in (Figure 3C–D). Additionally, previous theoretical work has shown that disinhibition can generate PD enhancement in similar models (Borst, 2018; Koch & Poggio, 1992); the choice of thresholds in this model did not permit flanking disinhibition with ON stimuli.

The synaptic model does not reproduce the fast timescale tuning of T4 cells

A third approach to characterizing direction-selective signals has been to apply stochastic stimuli with specified correlation structure. Responses to uncorrelated stimuli can be used to generate an unbiased estimate of a system’s linear receptive field (Chichilnisky, 2001). By using reverse-correlation and uncorrelated stimuli to extract spatiotemporal receptive fields, T4 cells have been characterized by oriented linear receptive fields with a central excitatory lobe and a delayed, offset inhibitory lobe (Figure 4A) (Leong et al., 2016; Salazar-Gatzimas et al., 2016). The T4 synaptic model generates the same shape of receptive field (Figure 4A). However, in the model, the inhibitory lobe lasts longer than that measured in T4 cells, and the tuning of the model was slower overall.

Responses to stochastic stimuli containing precise pairwise spatiotemporal correlations have revealed fast-timescale tuning in T4 cells (Salazar-Gatzimas et al., 2016). In measurements of T4 and T5, the cells could discriminate between spatiotemporal correlations with delays of 0 and 15 ms (Figure 4B). We presented the synaptic model with stimuli containing pairwise spatiotemporal correlations at different temporal delays. The model was direction-selective and responded to both positive and negative correlations, as in the cellular measurements. However, the model did not reproduce the fast timescale discrimination between delays (Figure 4B). Furthermore, the synaptic model showed strong suppression of ND-oriented positive correlations and enhancement of ND-oriented negative correlations, which was not observed in this dataset.

Behaving *Drosophila* respond direction-selectively to correlations higher than second-order (Clark et al., 2014; Leonhardt et al., 2016). This cannot be explained by models that compute pairwise correlations in the stimulus, such as the HRC and motion energy model. The sensitivity to higher-order correlations has been assessed using 3-point glider stimuli, which contain

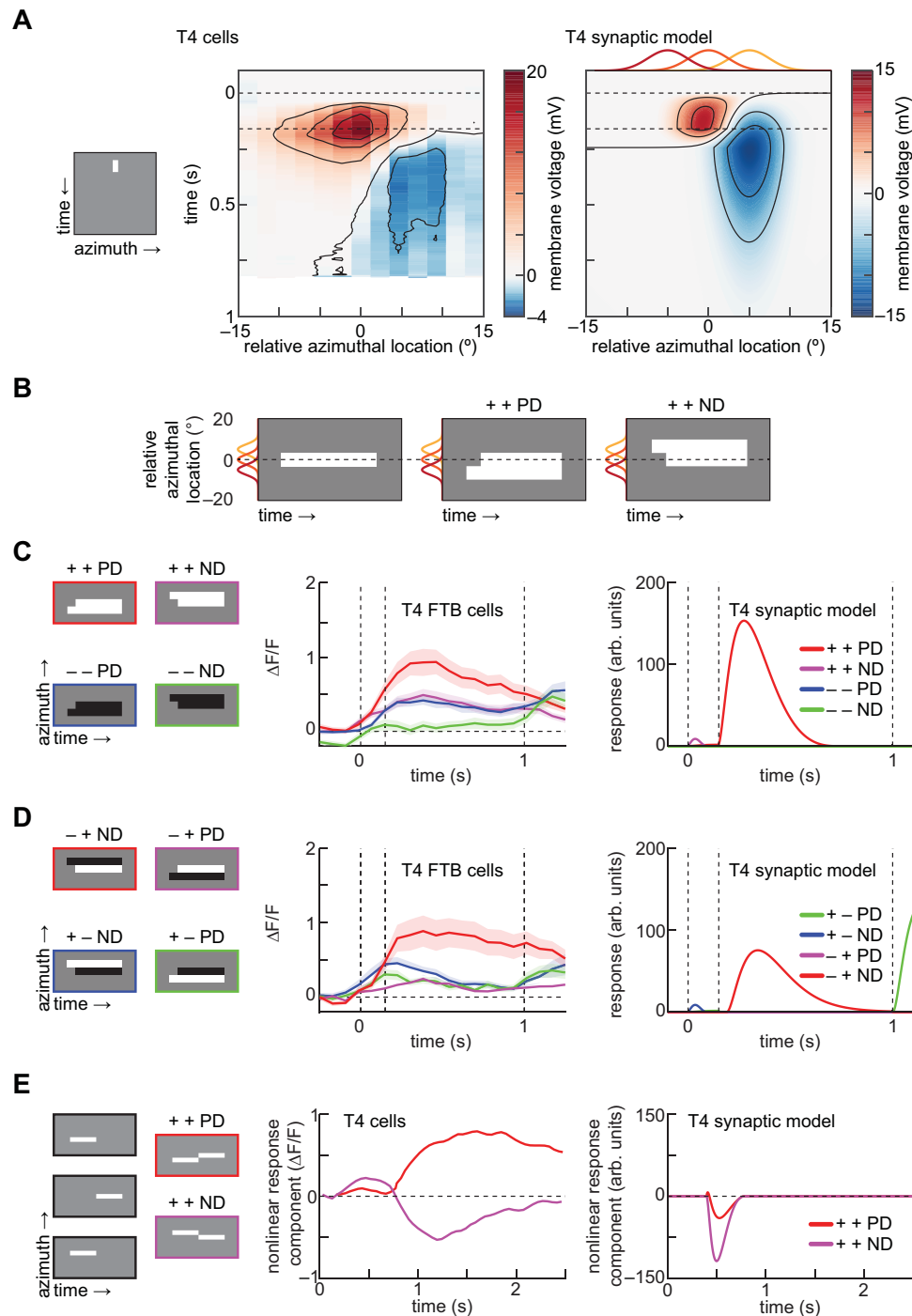


Figure 3. The T4 synaptic model reproduces the spatial organization and selectivity of apparent motion responses in T4 cells. (A) Responses to a single white bar flashed at different spatial locations. *Left*: Kymograph of 2-degree white bar presented for 160 ms. *Center*: Membrane voltage of T4 cells to flashed white bars, measured using electrophysiology (data from Gruntman et al., 2018). *Right*: As at center, but for the T4 synaptic model. Red, orange, and yellow lines indicate the spatial acceptance functions of the 3 model inputs. (B) Apparent motion stimuli are aligned such that the lagging bar is located at the center of the receptive field (Salazar-Gatzimas et al., 2018). The leading bar is presented at time 0 and lasts for 1 second, and the lagging bar is presented 150 ms later. Each bar subtends 5 degrees of visual angle. Red, orange, and yellow lines indicate the spatial acceptance functions of the 3 inputs. (C) Responses to phi apparent motion stimuli, aligned as in (B). *Left*: Kymographs of all four possible phi apparent motion stimuli. *Center*: Responses of T4 FTB cells to all 4 phi apparent motion stimuli, measured using 2-photon calcium imaging (data from Salazar-Gatzimas et al., 2018). Error patches indicate ± 1 SEM. *Right*: As at center, but for the T4 synaptic model. (D) As in panel C, but for reverse-phi apparent motion stimuli, in which the sequentially presented bars have opposite contrasts. (E) Assessing PD

←
enhancement and ND suppression. *Left*: Kymographs of linear decomposition of flashed apparent motion stimuli, with 4.5-degree-wide white bars presented sequentially for 400 ms each. *Center*: Nonlinear response component, defined as the residual of the linear prediction, measured using a calcium indicator (data from (Haag et al., 2016)). *Right*: As at center, but for the T4 synaptic model.

precise third-order correlations (Hu & Victor, 2010) (Figure 4C). The net responses of T4 cells to these stimuli have previously been inferred from behavioral measurements in *Drosophila* with the synaptic outputs of T5 cells silenced, using gliders updated at 24 Hz (Leonhardt et al., 2016). We used in vivo 2-photon calcium imaging to directly measure the responses of T4 cells to 3-point gliders updated at 5 Hz and found that the signs of the net responses were consistent with those measured in behavior with T5 cells silenced (Figure 4C, see Methods and Appendix A for details).

With an update rate of 24 Hz, the synaptic model correctly predicted the signs of net responses to diverging gliders measured in imaging and behavior, but predicted the wrong converging glider responses (Figure 4C). At 5 Hz, the synaptic model correctly predicted the signs of both converging and diverging glider responses but not the relative magnitudes. Thus the glider responses in T4 cells appear relatively insensitive to the glider timescale (24 vs. 5 Hz), but the model's response depends strongly on the input timescales.

T4 and T5 cells have been shown to display strongly direction-selective responses to rigidly-translating stimuli consisting of black and white squares placed at random on a gray background (see Appendix A) (Badwan et al., 2019). When two such stimuli that move in opposite directions are superimposed, they generate transparent motion percepts in primates (Qian, Andersen, & Adelson, 1994), and they generate responses in T4 that are reduced compared to presenting PD stimuli alone (Badwan et al., 2019). The synaptic model qualitatively reproduced these responses (Figure 4D). In particular, the responses of T4 cells are suppressed more strongly under the addition of ND motion than under the addition of OD motion, a feature that is reproduced by the synaptic T4 model (Figure 4D). Therefore, as in T4 cells, the selective direction-opponency observed in the model persists even with stimuli containing multiple spatiotemporal frequencies.

The T4 synaptic model provides decorrelated channels for naturalistic motion

Beyond artificial stimuli, natural scenes have been used to investigate the performance of direction-selective signals in flies (Badwan et al., 2019; Brinkworth & O'Carroll, 2009; Chen, Mandel, Fitzgerald, & Clark,

2019; Dror, O'Carroll, & Laughlin, 2001; Fitzgerald & Clark, 2015; Leonhardt et al., 2016; Salazar-Gatzimas et al., 2018; Straw, Rainsford, & O'Carroll, 2008). We therefore sought to investigate the performance of the T4 synaptic model in natural motion processing. To do so, we presented it with rigidly-translating scenes from a database of natural images (see Appendix A, Figure 5A) (Meyer, Schwegmann, Lindemann, & Egelhaaf, 2014). Although the structure and properties of inputs to T5 cells are known to differ from the inputs to T4 (Serbe et al., 2016; Shinomiya et al., 2019), to make an OFF-edge selective channel we created a "T5" model by simply inverting the ON/OFF selectivity of the inputs to our T4 synaptic model. This is intended merely to be an OFF-selective channel for the purposes of comparing T4 and potential T5 cell responses. The resulting four channels displayed strongly direction-selective average responses to translating natural scenes (Figure 5B).

Measured responses of T4 and T5 cells to translating natural scenes are decorrelated, so that only one channel is active at once (Salazar-Gatzimas et al., 2018). The synaptic models of T4 and "T5" also generated highly decorrelated responses, with the coactivation matrix of the four channels being nearly diagonal (Figure 5C). Such decorrelated parallel channels may provide a convenient representation of motion signals (Salazar-Gatzimas et al., 2018).

Discussion

An anatomically constrained synaptic model suffices to reproduce many, but not all, of the properties of *Drosophila* T4 cells. This model reproduces the direction-opponency, temporal-frequency-tuning, orientation-tuning, and phi/reverse-phi selectivity measured in T4 cells (Figures 2–4). When applied to a naturalistic velocity estimation task, it produces decorrelated signals similar to those measured in T4 and T5 neurons (Figure 5). However, it fails to reproduce the PD enhancement and fast-timescale tuning observed in T4 cells (Figures 3–4). Moreover, although it is sensitive to triplet correlations in its input, it fails to reproduce responses on the same timescales as observed in the data (Figure 4). In short, this simple synaptic model is sufficient to reproduce several distinct properties of T4 cells but cannot account for several observations.

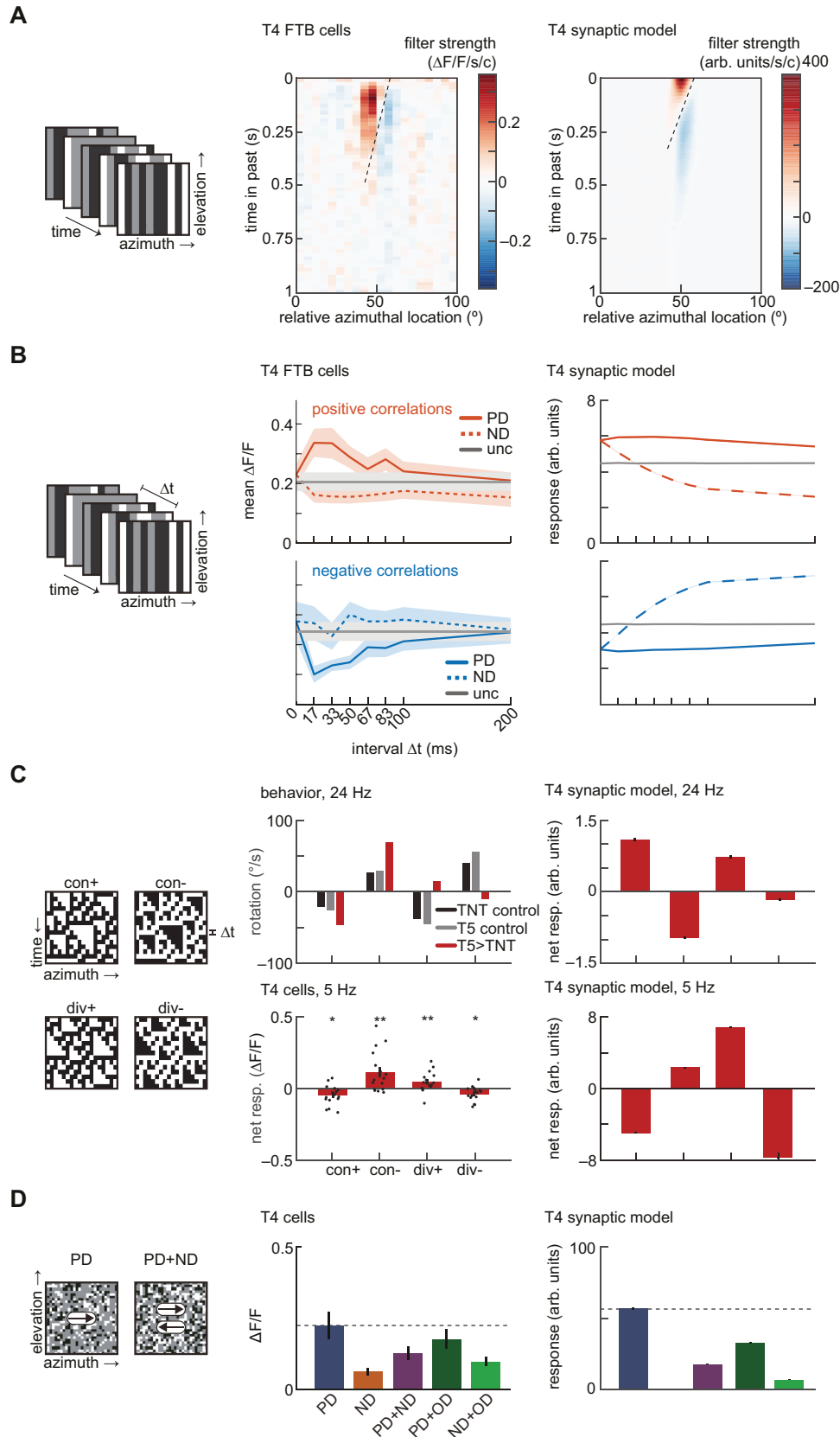


Figure 4. The minimal T4 synaptic model is not sufficient to reproduce the fast-timescale tuning of T4 cells. (A) Linear receptive field measurements. *Left*: Schematic depiction of binary, uncorrelated spatiotemporal noise. *Center*: Linear receptive field of T4 FTB cells (data from Salazar-Gatzimas et al., 2016). *Right*: As at center, but for the T4 synaptic model. (B) Correlation interval receptive field

←
 measurements. *Left*: Schematic depiction of ternary noise containing pairwise correlations at a specified interval Δt . *Center*: Responses of T4 FTB cells to positive and negative correlations, measured using a calcium indicator (data from Salazar-Gatzimas et al., 2016). Error bars indicate ± 1 SEM. *Right*: As at center, but for the T4 synaptic model. Error patches, which are barely visible, indicate 95% confidence intervals of the mean, which is variable due to the stochastic stimulus. (C) Triplet correlation sensitivity. *Left*: Kymographs of 3-point glider stimuli containing positive and negative triplet correlations. *Center top*: Turning behavioral responses to 3-point gliders updated at 24 Hz of flies with the synaptic outputs of T5 cells silenced (data from Leonhardt et al., 2016). Positive rotations correspond to the direction of the displacement of the spatial mean location of each triplet. *Center bottom*: Net responses of T4 cells to 3-point gliders updated at 5 Hz, measured using a calcium indicator (see Methods). Asterisks indicate that median net response differs from zero at the $p < 0.05$ (*) or $p < 0.01$ (**) level by a Wilcoxon signed-rank test with $N = 16$ flies. Exact p values are $p = 0.0174, 0.0061, 0.0097,$ and 0.0131 for con+, con−, div+, and div−, respectively. Error bars indicate ± 1 SEM over flies, and black circles indicate individual per-fly means. *Right*: As at center, but for the T4 synaptic model. Error bars indicate 95% confidence intervals of the mean. (D) Responses to rigidly translating stimuli with stochastic checkerboard patterns. *Left*: Images of random checkerboard stimuli. *Center*: Mean responses of T4 cells to checkerboard stimuli translating at 100 degrees/s, measured using a calcium indicator (data from Badwan et al., 2019). Error bars indicate ± 1 SEM. *Right*: As at center, but for the T4 synaptic model. Error bars indicate 95% confidence intervals of the mean.

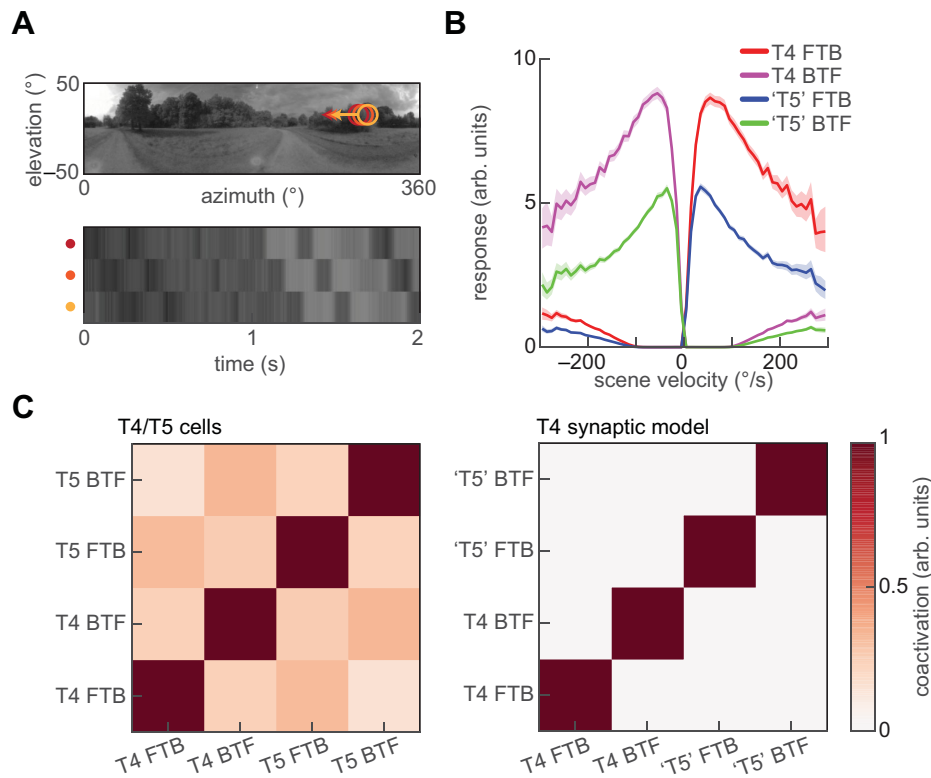


Figure 5. The T4 synaptic model generates decorrelated signals in response to naturalistic motion. (A) A rigidly rotating panoramic natural scene, with three spatially offset input signals as a function of time. (B) Average responses of the T4 synaptic model and a “T5” variant to naturalistic motion constructed by rigidly translating natural scenes at a variety of velocities. Error patches indicate 95% confidence intervals of the mean. “T5” cells were constructed by sign-inverting the inputs to the minimal T4 synaptic model. (C) Decorrelation of channels with naturalistic motion. *Left*: Coactivation matrix of T4 and T5 cells in response to rigidly translating natural scenes (data from Salazar-Gatzimas et al., 2018). *Right*: As at left, but for the T4 synaptic model and “T5” variant. Coactivation was computed for an ensemble of (image, velocity) pairs, in which the velocity was chosen from a Gaussian distribution with zero mean and 100 degrees/s standard deviation.

Minimal models and levels of understanding

Here, we asked whether a minimal synaptic model could qualitatively reproduce features of T4 cell

responses. The minimal model required no exotic neurotransmitter receptors or interactions and was based on simple synaptic conductances. The simplifications sufficient to explain different phenomena

will depend strongly on the features one seeks to reproduce and on the desired level of fidelity. However, minimal models are useful precisely because they can be relatively straightforward to analyze.

Marr famously proposed different levels of understanding neural circuitry, including an algorithmic level and a mechanistic level (Marr & Poggio, 1976). As we drive toward a deeper understanding of the visual motion circuit in the fly, the levels of algorithm and mechanism can appear increasingly blurred. It is hard to define what distinguishes the mechanistic circuit description presented here from a detailed algorithm-level description of the computation. However, it remains important to connect proposed mechanistic models to high-level descriptions of the system, like the HRC. This is because the high-level descriptions of computations provide a level of intuition for the behavior of the system that a more intricate model cannot. Moreover, the HRC explains a wide variety of neural and behavioral data in flies (Borst & Egelhaaf, 1989; Yang & Clandinin, 2018), so an HRC-like algorithm must be a limiting case of any proposed model (Potters & Bialek, 1994).

Sufficiency of models

Many details of the function of the early visual system were neglected in this model. For instance, the filter shapes in neurons leading into the model T4 cell have been well characterized (Arenz et al., 2017; Behnia et al., 2014), but this model used simple exponential filters. Lateral inhibition is widely documented in the early fly visual system (Arenz et al., 2017; Freifeld, Clark, Schnitzer, Horowitz, & Clandinin, 2013; Meier, Serbe, Maisak, Haag, Dickson, & Borst, 2014), but this model used simple Gaussian spatial acceptance functions without lateral inhibition. The synapses that feed into the medulla neurons that synapse onto T4 are likely to have complex, nonlinear processing properties (Matulis et al., 2020; Drews et al., 2020; Molina-Obando, Vargas-Fique, Henning, Gür, Schladt, Akhtar, Berger, & Silies, 2019; Yang, St-Pierre, Sun, Ding, Lin, & Clandinin, 2016), yet we modeled the entire input pathway as a purely linear filter. The rectifications of neural responses upstream of T4 are imperfect (Behnia et al., 2014; Salazar-Gatzimas et al., 2018), but this model used simple threshold-linear rectifiers. The fly eye possesses neurons that feedback onto earlier stages and create reciprocal interactions between neurons (Takemura et al., 2013; Takemura et al., 2017; Zheng, de Polavieja, Wolfram, Asyali, Hardie, & Juusola, 2006), but this model is entirely feedforward.

Despite these approximations, the synaptic T4 model presented here is sufficient to qualitatively match a variety of T4 neuron responses. Adding some of these

neglected details into a model may make it sufficient to reproduce other features of T4 responses. This provides a method for understanding which details of processing are related to which response features in T4 cells: one may ask how different details of the system affect the sufficiency of a model to reproduce specific downstream response properties. As the field acquires more and more detailed information about the motion detection circuitry, this sort of analysis will be critical to understand the functional role of different properties.

One might naturally ask whether the synaptic model presented here might be further simplified without sacrificing its ability to account for the response properties of T4 cells. As described in Appendix C, a simplified linear-nonlinear cascade representing the numerator of the biophysical nonlinearity can generate some, but not all, of the properties of the full model.

Flexibility in extending this minimal synaptic model

In selecting parameter values for this synaptic model, we sought to reproduce only a few properties of T4 cells: a temporal frequency maximum of 1 Hz and direction-opponent average responses to sinusoid gratings with a temporal frequency of 1 Hz and a spatial wavelength of 45 degrees (Appendix B) (Badwan et al., 2019). To capture a larger subset of the measured properties of T4 cells, one could fit the model's parameters to optimize its predictions of responses to many different stimuli (Deb, 2014). Such a solution would provide information about the maximal ability of this synaptic model to reproduce the properties of T4 cells, but it seems unlikely to provide insight into the predictive power of the core features of the model.

The organization of this model allows for several clear tuning mechanisms. First, the temporal filters could be modified to better match measured filters, which would change the tuning of the system in response to sinusoids (Figure 2). Second, the degree to which inhibition is shunting or hyperpolarizing can be adjusted by changing the reversal potential of inhibitory currents. This could effectively hide inhibition under some stimuli and measurements. Third, the threshold for the OFF-inhibitory input could be changed to better represent preferred direction enhancement (Figure 3) (Borst, 2018). This would allow disinhibition of Mi4 to change the gain of the central input.

In the model analyzed here, we chose all thresholds of the input LN models to be zero. This effectively ignores contrast asymmetries in the natural world (Geisler, 2008), which have been used to understand many functional properties of motion detectors in flies (Chen et al., 2019; Clark, Fitzgerald, Ales, Gohl, Silies,

Norcia, & Clandinin, 2014; Fitzgerald & Clark, 2015; Fitzgerald et al., 2011; Leonhardt et al., 2016; Salazar-Gatzimas et al., 2018). Changing these thresholds to optimize for natural scene motion estimation might also generate a parameter set that better captures responses to triplet correlations (Figure 4) (Chen et al., 2019; Fitzgerald & Clark, 2015). In short, the synaptic model presented here is highly flexible and extensible, and uses only simple, known biophysical mechanisms.

For the sake of simplicity, we have used single delay and nondelay lines in this work (Figure 1A). However, T4 cells receive fast excitatory input at the center of their receptive fields from both Mi1 and Tm3 cells, and delayed OFF inhibitory input offset in the PD from both Mi4 and CT1 cells (Shinomiya et al., 2019; Takemura et al., 2017). Dissecting how information from these parallel channels is used, particularly if it is nonlinearly combined (Strother et al., 2017), will be important in developing a full understanding of the direction-selective computation performed by T4 cells.

Modelling temporal processing

The model presented here failed to capture some of the fast-timescale tuning measured in T4, including in its responses to pairwise and triplet spatiotemporal correlations (Figure 4). In this minimal model, we represented all temporal processing by linear filters. However, the temporal processing upstream of T4 cells involves nonlinear and adaptive mechanisms, which can affect temporal response properties (Matulis et al., 2020; Drews et al., 2020; Howard, Dubs, & Payne, 1984; Zheng et al., 2006). Thus far, the study of nonlinear mechanisms in the fly visual system has focused on static nonlinear effects such as rectification (Behnia et al., 2014; Yang et al., 2016) and on nonlinear interactions between linearly filtered signals (Borst, Flanagin, & Sompolinsky, 2005; Fitzgerald & Clark, 2015). The inclusion of nonlinear effects on the dynamics themselves may be necessary to accurately capture the temporal processing upstream of T4 cells. As a first step toward understanding adaptation in this circuit, one might characterize the temporal kernels of the inputs to T4 cells with high resolution (Mano, Creamer, Matulis, Salazar-Gatzimas, Zavatone-Veth, & Clark, 2019; Yang et al., 2016) and study how their properties depend on stimulus statistics and history (Baccus & Meister, 2002; Kim & Rieke, 2001; Rieke, 2001). Only a few models have focused on these sorts of adaptive changes in processing dynamics (Clark, Benichou, Meister, & da Silveira, 2013). Although the analysis of dynamic temporal nonlinearities is complex, incorporating them into models may provide insight into how fast timescale tuning of T4 cells arises.

Direction-selective signals emerge at multiple sequential processing stages

Direction-selective time-averaged signals require a nonlinear interaction between inputs over space and time (Adelson & Bergen, 1985; Borst & Egelhaaf, 1989). In the minimal model we presented here, there are two distinct steps at which such nonlinearities could arise. First, the postsynaptic interaction between conductances permits nonlinear interactions that appear in the membrane potential. Second, the transformation from membrane potential into calcium, which we represent as a point nonlinearity, can generate nonlinear interaction terms and is in principle very similar to the motion energy model if the membrane potential is a linear transformation of the stimulus (Adelson & Bergen, 1985). Thus this model permits two sequential processing steps that each contribute to direction selectivity. This model can create small direction-selective biases in the membrane voltage, similar to those observed in T4 cells (Gruntman et al., 2018), but can also use the calcium transformation to generate strong direction-selective signals even if the membrane potential is nearly linearly related to the stimulus (Wienecke et al., 2018). Similarly, the biophysical properties of T4 cells also likely permit direction-selective signals to arise at multiple stages during processing; a third stage might be the transformation from calcium concentration to synaptic release. We hypothesize that having direction-selective signals emerge at multiple stages of processing might make the computation robust to perturbations.

A T5 synaptic model

In this work, we used a sign-inverted version of our T4 synaptic model to represent the OFF-edge-selective T5 cells. This representation would correspond to a first-order direction-selective cell that receives OFF excitatory input at the center of its receptive field, delayed OFF inhibitory input offset in its preferred direction, and delayed ON input offset in the null direction. Such a model would correctly predict the selective responses of T5 cells to phi and reverse-phi apparent motion stimuli (Salazar-Gatzimas et al., 2018). However, the functional and anatomical structure of the inputs to T5 cells suggests that it receives only OFF inputs (Serbe et al., 2016; Shinomiya et al., 2019). Somehow, however, signals in T5 cells are sensitive to both contrast increments and decrements (Salazar-Gatzimas et al., 2018; Wienecke et al., 2018). Further study of the functional properties of the OFF-edge motion pathway will be required to elucidate how the direction-selective computations in T4 and T5 cells differ.

Relationships to mammalian visual systems

The fly's visual motion detection circuits bear striking similarities to those in mammalian retina in their anatomy, circuitry, and algorithmic processing (Borst & Helmstaedter, 2015; Clark & Demb, 2016; Sanes & Zipursky, 2010). In mammalian retina, the earliest direction-selective signals are generated by starburst amacrine cells (SACs), which are also tuned to ON- and OFF-edges (Euler, Detwiler, & Denk, 2002; Famiglietti Jr, 1983). It appears that SACs may receive inputs that are differentially delayed (Fransen & Borghuis, 2017; Ichinose & Hellmer, 2016; Kim, Greene, Zlateski, Lee, Richardson, Turaga, Purcaro, Balkam, Robinson, & Behabadi, 2014), similar to the inputs to T4 cells. It would be interesting to investigate how much SAC phenomenology that mechanism alone could account for, when linked to simple biophysical mechanisms. As in this study, it could provide insight into where the circuit understanding is lacking, especially when complex stimuli are used to probe SAC function (Chen, Pei, Koren, & Wei, 2016).

It is notable that the ON-ON-OFF spatial organization of T4 inputs (Figure 1A) is almost identical to a model proposed to explain cortical responses to pairwise correlations (Mo & Koch, 2003). This suggests there may be deep parallels between T4 and T5 and cortical motion processing. Models for fly and cortical direction selectivity have traditionally differed in whether they assume discrete inputs (fly, HRC-like models) or more continuous inputs (cortex, motion-energy-like spatiotemporal filtering). If synaptic interactions are considered, then continuous linear filters cannot be applied, and models must incorporate the discrete receptive fields of the inputs to a cell. It would be interesting to ask how such conductance models fare in predicting cortical responses; the statistical nature of cortical connections make it more difficult to construct a general model of this type.

In this synaptic model of T4 cell function, we have paired known connectivity with measured physiology and simple biophysics to predict many circuit processing properties. This allows us to define where such a model succeeds and where it fails. This represents progress toward understanding this circuit at all levels, from utility to algorithm to mechanism.

Keywords: Keywords: direction selectivity, drosophila, synaptic model, circuit model, receptive fields

Acknowledgments

We thank J. E. Fitzgerald and members of the Clark lab for helpful conversations. DAC and this research were supported by NIH R01EY026555, NIH P30EY026878, NSF IOS1558103, a Searle Scholar

Award, a Sloan Fellowship in Neuroscience, the Smith Family Foundation, and the E. Matilda Ziegler Foundation.

Author contributions: JAZV and DAC conceived of numerical experiments. JAZV performed numerical simulations and analyzed the model. BAB acquired and analyzed calcium imaging data. JAZV and DAC wrote the paper.

Commercial relationships: none.

Corresponding author: Damon A. Clark.

Email: damon.clark@yale.edu.

Address: Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, CT, USA.

Present address of JAZV: Department of Physics, Harvard University, Cambridge, MA, USA.

References

- Adelson, E., & Bergen, J. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2, 284–299.
- Anstis, S. (1970). Phi movement as a subtraction process. *Vision Research*, 10, 1411–1430.
- 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. *Current Biology*, 27, 929–944.
- Baccus, S., & Meister, M. (2002). Fast and slow contrast adaptation in retinal circuitry. *Neuron*, 36, 909–919.
- Badwan, B. A., Creamer, M. S., Zavatone-Veth, J. A., & Clark, D. A. (2019). Dynamic nonlinearities enable direction opponency in *Drosophila* elementary motion detectors. *Nature Neuroscience*, 22, 1318–1326.
- Barlow, H., & Levick, W. R. (1965). The mechanism of directionally selective units in rabbit's retina. *The Journal of Physiology*, 178, 477.
- Barry, S. (2015). *A comprehensive course in analysis*. Providence, Rhode Island: American Mathematical Society.
- 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, 427–430.
- Borst, A. (2018). A biophysical mechanism for preferred direction enhancement in fly motion vision. *PLoS Computational Biology*, 14, e1006240.
- Borst, A., & Egelhaaf, M. (1989). Principles of visual motion detection. *Trends in Neurosciences*, 12, 297–306.

- Borst, A., Flanagan, V. L., & Sompolinsky, H. (2005). Adaptation without parameter change: Dynamic gain control in motion detection. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 6172–6176.
- Borst, A., & Helmstaedter, M. (2015). Common circuit design in fly and mammalian motion vision. *Nature Neuroscience*, *18*, 1067–1076.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.
- Brinkworth, R. S., & O’Carroll, D. C. (2009). Robust models for optic flow coding in natural scenes inspired by insect biology. *PLoS Computational Biology*, *5*, e1000555.
- Chen, J., Mandel, H. B., Fitzgerald, J. E., & Clark, D. A. (2019). Asymmetric ON-OFF processing of visual motion cancels variability induced by the structure of natural scenes. *eLife*, *8*, e47579.
- Chen, Q., Pei, Z., Koren, D., & Wei, W. (2016). Stimulus-dependent recruitment of lateral inhibition underlies retinal direction selectivity. *eLife*, *5*, e21053.
- Chichilnisky, E. (2001). A simple white noise analysis of neuronal light responses. *Network: Computation in Neural Systems*, *12*, 199–213.
- Clark, D. A., Benichou, R., Meister, M., & da Silveira, R. A. (2013). Dynamical Adaptation in Photoreceptors. *PLoS Computational Biology*, *9*, e1003289.
- 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*, 1165–1177.
- Clark, D. A., & Demb, J. B. (2016). Parallel computations in insect and mammalian visual motion processing. *Current Biology*, *26*, R1062–R1072.
- Clark, D. A., Fitzgerald, J. E., Ales, J. M., Gohl, D. M., Silies, M., Norcia, A. M., . . . Clandinin, T. R. (2014). Flies and humans share a motion estimation strategy that exploits natural scene statistics. *Nature Neuroscience*, *17*, 296–303.
- Creamer, M. S., Mano, O., & Clark, D. A. (2018). Visual Control of Walking Speed in *Drosophila*. *Neuron*, *100*, 1460–1473.
- Creamer, M. S., Mano, O., Tanaka, R., & Clark, D. A. (2019). A flexible geometry for panoramic visual and optogenetic stimulation during behavior and physiology. *Journal of Neuroscience Methods*, *323*, 48–55.
- Dayan, P., & Abbott, L. F. (2001). *Theoretical neuroscience*. Cambridge, MA: MIT Press.
- DeAngelis, B. D., Zavatone-Veth, J. A., & Clark, D. A. (2019). The manifold structure of limb coordination in walking *Drosophila*. *eLife*, *8*, e46409.
- Deb, K. (2014). Multi-objective optimization. In E. K. Burke, & G. Kendall (Eds.), *Search methodologies*. (pp. 403–449). Berlin, Germany: Springer.
- 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. *Current Biology*, P209–2221.E8.
- Dror, R. O., O’Carroll, D. C., & Laughlin, S. B. (2001). Accuracy of velocity estimation by Reichardt correlators. *Journal of the Optical Society of America A*, *18*, 241–252.
- Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American statistical Association*, *82*, 171–185.
- Eichner, H., Joesch, M., Schnell, B., Reiff, D. F., & Borst, A. (2011). Internal structure of the fly elementary motion detector. *Neuron*, *70*, 1155–1164.
- Euler, T., Detwiler, P. B., & Denk, W. (2002). Directionally selective calcium signals in dendrites of starburst amacrine cells. *Nature*, *418*, 845–852.
- Famiglietti, E. V., Jr (1983). On and off pathways through amacrine cells in mammalian retina: the synaptic connections of “starburst” amacrine cells. *Vision Research*, *23*, 1265–1279.
- Fisher, Y. E., Silies, M., & Clandinin, T. R. (2015). Orientation selectivity sharpens motion detection in *Drosophila*. *Neuron*, *88*, 390–402.
- Fitzgerald, J. E., & Clark, D. A. (2015). Nonlinear circuits for naturalistic visual motion estimation. *eLife*, e09123.
- Fitzgerald, J. E., Katsov, A. Y., Clandinin, T. R., & Schnitzer, M. J. (2011). Symmetries in stimulus statistics shape the form of visual motion estimators. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 12909–12914.
- Fransen, J. W., & Borghuis, B. G. (2017). Temporally Diverse Excitation Generates Direction-Selective Responses in ON-and OFF-Type Retinal Starburst Amacrine Cells. *Cell Reports*, *18*, 1356–1365.
- 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*, 1075–1089.
- Geisler, W. S. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology*, *59*, 167–192.

- Gruntman, E., Romani, S., & Reiser, M. B. (2018). Simple integration of fast excitation and offset, delayed inhibition computes directional selectivity in *Drosophila*. *Nature Neuroscience*, *21*, 250–257.
- Haag, J., Arenz, A., Serbe, E., Gabbiani, F., & Borst, A. (2016). Complementary mechanisms create direction selectivity in the fly. *eLife*, *5*.
- Hassenstein, B., & Reichardt, W. (1956). Systemtheoretische Analyse der Zeit-, Reihenfolgen- und Vorzeichenauswertung bei der Bewegungsperzeption des Rüsselkäfers *Chlorophanus*. *Zeitschrift für Naturforschung*, *11*, 513–524.
- Hollander, M., Wolfe, D. A., & Chicken, E. (2013). *Nonparametric statistical methods, Vol. 751*. Hoboken, NJ: Wiley.
- Howard, J., Dubs, A., & Payne, R. (1984). The dynamics of phototransduction in insects. *Journal of Comparative Physiology A*, *154*, 707–718.
- Hu, Q., & Victor, J. D. (2010). A set of high-order spatiotemporal stimuli that elicit motion and reverse-phi percepts. *Journal of Vision*, *10*, 9.1–16.
- Ichinose, T., & Hellmer, C. B. (2016). Differential signalling and glutamate receptor compositions in the OFF bipolar cell types in the mouse retina. *The Journal of Physiology*, *594*, 883–894.
- Jagadeesh, B., Wheat, H. S., & Ferster, D. (1993). Linearity of summation of synaptic potentials underlying direction selectivity in simple cells of the cat visual cortex. *Science*, *262*, 1901–1904.
- Joesch, M., Schnell, B., Raghu, S., Reiff, D., & Borst, A. (2010). ON and OFF pathways in *Drosophila* motion vision. *Nature*, *468*, 300–304.
- Kato, S., Xu, Y., Cho, C. E., Abbott, L., & Bargmann, C. I. (2014). Temporal responses of *C. elegans* chemosensory neurons are preserved in behavioral dynamics. *Neuron*, *81*, 616–628.
- Katsov, A., & Clandinin, T. (2008). Motion processing streams in *Drosophila* are behaviorally specialized. *Neuron*, *59*, 322–335.
- Kim, J. S., Greene, M. J., Zlateski, A., Lee, K., Richardson, M., Turaga, S. C., ... Behabadi, B. F. (2014). Space-time wiring specificity supports direction selectivity in the retina. *Nature*, *509*, 331–336.
- Kim, K., & Rieke, F. (2001). Temporal contrast adaptation in the input and output signals of salamander retinal ganglion cells. *The Journal of Neuroscience*, *21*, 287–299.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, *36*, 1–16.
- Koch, C., & Poggio, T. (1992). Multiplying with synapses and neurons. In T. McKenna, J. Davis, & S. Zornetzer (Eds.) *Single neuron computation* (pp. 315–345). Philadelphia: Elsevier.
- Kunze, P. (1961). Untersuchung des Bewegungssehens fixiert fliegender Bienen. *Zeitschrift für vergleichende Physiologie*, *44*, 656–684.
- Leong, J. C. S., Esch, J. J., Poole, B., Ganguli, S., & Clandinin, T. R. (2016). Direction selectivity in *Drosophila* emerges from preferred-direction enhancement and null-direction suppression. *The Journal of Neuroscience*, *36*, 8078–8092.
- 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. *Nature Neuroscience*, *19*, 706–715.
- Maisak, M. S., Haag, J., Ammer, G., Serbe, E., Meier, M., Leonhardt, A., ... Borst, A. (2013). A directional tuning map of *Drosophila* elementary motion detectors. *Nature*, *500*, 212–216.
- Mano, O., Creamer, M. S., Matulis, C. A., Salazar-Gatzimas, E., Zavatone-Veth, J. A., & Clark, D. A. (2019). Using slow frame rate imaging to extract fast receptive fields. *Nature Communications*, *10*, 4979.
- Marr, D., & Poggio, T. (1976). From understanding computation to understanding neural circuitry. In *AI Memo (Massachusetts Institute of Technology)*.
- Matulis, C. A., Chen, J., Gonzalez-Suarez, A. D., Behnia, R., & Clark, D. A. (2020). Heterogeneous Temporal Contrast Adaptation in *Drosophila* Direction-Selective Circuits. *Current Biology*, *222–236.e6*.
- Meier, M., & Borst, A. (2019). Extreme Compartmentalization in a *Drosophila* Amacrine Cell. *Current Biology*, *29*, 1545–1550. e1542.
- 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. *Current Biology*, *24*, 385–392.
- Meyer, H. G., Schwegmann, A., Lindemann, J. P., & Egelhaaf, M. (2014). Panoramic high dynamic range images in diverse environments. Available at <https://pub.uni-bielefeld.de/record/2689637>. Accessed March 25, 2019.
- Mo, C.-H., & Koch, C. (2003). Modeling reverse-phi motion-selective neurons in cortex: double synaptic-veto mechanism. *Neural Computation*, *15*, 735–759.
- Molina-Obando, S., Vargas-Fique, J. F., Henning, M., Gür, B., Schladt, T. M., & Akhtar, J., ... Silies, M. (2019). ON selectivity in *Drosophila* vision is a

- multisynaptic process involving both glutamatergic and GABAergic inhibition. *eLife*, 8, e49373.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442.
- Potters, M., & Bialek, W. (1994). Statistical mechanics and visual signal processing. *Journale de Physique*, 4, 1755–1775.
- Priebe, N. J., Lisberger, S. G., & Movshon, J. A. (2006). Tuning for spatiotemporal frequency and speed in directionally selective neurons of macaque striate cortex. *The Journal of Neuroscience*, 26, 2941–2950.
- Qian, N., Andersen, R. A., & Adelson, E. H. (1994). Transparent motion perception as detection of unbalanced motion signals. *The Journal of Neuroscience*, 14, 7357–7366.
- Rieke, F. (2001). Temporal contrast adaptation in salamander bipolar cells. *The Journal of Neuroscience*, 21, 9445–9454.
- Salazar-Gatzimas, E., Agrochao, M., Fitzgerald, J. E., & Clark, D. A. (2018). The Neuronal Basis of an Illusory Motion Percept Is Explained by Decorrelation of Parallel Motion Pathways. *Current Biology*, 28, 3748–3762. e3748.
- Salazar-Gatzimas, E., Chen, J., Creamer, M. S., Mano, O., Mandel, H. B., & Matulis, C. A., ...Clark, D. A. (2016). Direct measurement of correlation responses in *Drosophila* elementary motion detectors reveals fast timescale tuning. *Neuron*, 92, 227–239.
- Sanes, J. R., & Zipursky, S. L. (2010). Design principles of insect and vertebrate visual systems. *Neuron*, 66, 15.
- Serbe, E., Meier, M., Leonhardt, A., & Borst, A. (2016). Comprehensive characterization of the major presynaptic elements to the *Drosophila* OFF motion detector. *Neuron*, 89, 829–841.
- Shinomiya, K., Huang, G., Lu, Z., Parag, T., Xu, C. S., & Aniceto, R., ...Neace, E. (2019). Comparisons between the ON- and OFF-edge motion pathways in the *Drosophila* brain. *eLife*, 8, e40025.
- Stavenga, D. (2003). Angular and spectral sensitivity of fly photoreceptors. II. Dependence on facet lens F-number and rhabdomere type in *Drosophila*. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 189, 189–202.
- Straw, A. D., Rainsford, T., & O'Carroll, D. C. (2008). Contrast sensitivity of insect motion detectors to natural images. *Journal of Vision*, 8, 32–32.
- Strother, J. A., Wu, S.-T., Wong, A. M., Nern, A., Rogers, E. M., & Le, J. Q., ...Reiser, M. B. (2017). The emergence of directional selectivity in the visual motion pathway of *Drosophila*. *Neuron*, 94, 168–182. e110.
- Takemura, S.-Y., Bharioke, A., Lu, Z., Nern, A., Vitaladevuni, S., & Rivlin, P. K., ...Winston, P. (2013). A visual motion detection circuit suggested by *Drosophila* connectomics. *Nature*, 500, 175–181.
- 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, e24394.
- Torre, V., & Poggio, T. (1978). A synaptic mechanism possibly underlying directional selectivity to motion. *Proceedings of the Royal Society B*, 202, 409–416.
- Wienecke, C. F., Leong, J. C., & Clandinin, T. R. (2018). Linear summation underlies direction selectivity in *drosophila*. *Neuron*, 99, 680–688.e4.
- Yang, H. H., & Clandinin, T. R. (2018). Elementary motion detection in *Drosophila*: algorithms and mechanisms. *Annual Review of Vision Science*, 4, 143–163.
- 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, 245–257.
- Zheng, L., de Polavieja, G. G., Wolfram, V., Asyali, M. H., Hardie, R. C., & Juusola, M. (2006). Feedback network controls photoreceptor output at the layer of first visual synapses in *Drosophila*. *The Journal of General Physiology*, 127, 495–510.

Appendix A: Visual stimuli used in simulations and imaging experiments

In this appendix, we describe in detail all stimuli used in this work.

ON and OFF edges (Figure 1)

ON and OFF edges were constructed by placing white (respectively, black) edges on a gray background. All edges translated at 30 degrees/s.

Sinusoid grating stimuli (Figures 1 and 2)

Sinusoid grating stimuli were constructed as in previous studies (Badwan et al., 2019; Creamer et al.,

2018; Maisak et al., 2013). Briefly, rightward- and leftward-drifting gratings were constructed as

$$c(t, x) = c_0 \sin(\omega t \mp \kappa x)$$

where c_0 is the input contrast, ω is the temporal frequency in units of radians per second, κ is the spatial frequency in units of radians per degree, and the negative sign is taken for rightward-drifting gratings. To assess whether our model is temporal-frequency-tuned, we computed the fraction of the total variance in a spatiotemporal frequency sweep of its responses accounted for by a separable approximation resulting from its singular value decomposition (Creamer et al., 2018). Counterphase gratings were constructed as

$$c(t, x) = c_0 \sin(\omega t + \kappa x + \phi_1) + c_0 \sin(\omega t - \kappa x + \phi_2)$$

where ϕ_1 and ϕ_2 are uniformly sampled phase offsets, over which we average in all analyses. Gratings containing preferred- and orthogonal-direction motion were constructed as

$$c(t, x) = c_0 \sin(\omega t - \kappa x + \phi_1) + c_0 \sin(\omega t + \phi_2)$$

The linearity analysis in Figure 2A was applied to T5 cells (Wienecke et al., 2018), following a previously developed protocol (Jagadeesh et al., 1993). This analysis relies on the fact that a drifting sinusoid grating may be decomposed into a sum of counterphase gratings as

$$\begin{aligned} c(t, x) &= c_0 \sin(\omega t \mp \kappa x) \\ &= \frac{c_0}{4} \sum_{n=0}^7 \sin\left(\omega t + \frac{n\pi}{8} \mp \frac{\pi}{2}\right) \sin\left(\kappa x \pm \frac{n\pi}{8}\right). \end{aligned}$$

Therefore, if a system is linear, its scaled, summed response of a linear system to counterphase gratings with these phase shifts will be equal to its response to the corresponding drifting grating. By comparing the linear prediction of the drifting grating response to the actual response, one may assess a system's linearity.

To assess the orientation- and directional-tuning of the model with sinusoid gratings in Figure 2D, we defined a 2-dimensional grating

$$c(t, x, y) = c_0 \sin(\omega t - \kappa(x \cos \theta + y \sin \theta)),$$

where the angle θ defines its orientation. In this analysis, we assume that the ring of detectors is located at $y = 0$, and that the Gaussian spatial filter is symmetric in x and y . Static gratings were formed by setting $\omega = 0$. We note that our convention for the orientation of a static grating differs from the original manuscript (Fisher et al., 2015); we define the orientation as the angle between the normal to the apparent edge and the

preferred direction rather than the angular position of the edge itself. Therefore, in our convention the preferred orientations and directions align.

Apparent motion stimuli (Figure 3)

Single-bar stimuli were constructed as previously published (Gruntman et al., 2018; Salazar-Gatzimas et al., 2018). Briefly, 5 degrees (respectively, 2 degrees) black or white bars were placed on a gray background and presented for 1 second (respectively 160 ms) to match (Salazar-Gatzimas et al., 2018) (respectively, Gruntman et al., 2018). Bar pair apparent motion stimuli were constructed as in (Salazar-Gatzimas et al., 2018). Briefly, 5-degree black or white bars were placed on a gray background and presented for 1 second. To create apparent motion, a second black or white bar was added 150 ms after the onset of the first bar at a neighboring spatial location. Responses to these bar pair apparent motion stimuli were aligned such that the location of the lagging bar matched the location of peak single-bar responses, as in (Salazar-Gatzimas et al., 2018). Flashed apparent motion stimuli were constructed similarly to those presented to T4 and T5 cells in (Gruntman et al., 2018; Haag et al., 2016). Briefly, 4.5 degrees white bars were placed on a gray background and were presented for 400 ms in sequential spatial positions.

Noise stimuli and linear receptive field extraction (Figure 4)

As previously published (Salazar-Gatzimas et al., 2016), we extracted linear receptive fields from responses to uncorrelated binary stimuli composed of 5-degree black or white bars, updated at 60 Hz. We estimated the linear receptive field from these responses using reverse correlation (Chichilnisky, 2001). Ternary noise stimuli with pairwise correlations were constructed as in (Salazar-Gatzimas et al., 2016). Briefly, the contrast of the correlated noise stimulus was given as

$$c(t, x) = \frac{1}{2} (B(t, x) \pm B(t + \delta t, x + \delta x)),$$

where $B(t, x)$ is an uncorrelated binary stimulus composed of 5-degree black or white bars, and addition (respectively, subtraction) generates positive (respectively negative) correlations. The stimulus was updated at a fixed rate, and the temporal offset δt was taken to be one cycle, with its sign determining whether the stimulus was oriented in the preferred or null direction. The spatial offset δx was fixed to be 1 bar width. As shown in (Salazar-Gatzimas et al., 2016), the autocorrelation function of this stimulus, with

spacetime discretized by the bar width and sampling rate, is

$$\begin{aligned} &\langle c(t, x) c(t + \tau, x + \rho) \rangle \\ &= \frac{1}{2} \delta_{\tau, 0} \delta_{\rho, 0} + \frac{1}{4} (\delta_{\tau, \delta t} \delta_{\rho, \delta x} + \delta_{\tau, -\delta t} \delta_{\rho, -\delta x}), \end{aligned}$$

where $\delta_{i,j}$ is the Kronecker delta.

Three-point glider stimuli (Figure 4)

As in previous studies (Clark et al., 2014; Fitzgerald and Clark, 2015), we constructed 3-point glider stimuli following (Hu and Victor, 2010). Briefly, these binary stimuli enforce correlations over space and time among triplets of pixels. Three-point gliders may be categorized into 4 types: converging gliders with positive parity (con+), converging gliders with negative parity (con-), diverging gliders with positive parity (div+), and diverging gliders with negative parity (div-). Letting ρ be the pixel spacing and δ be the frame duration (the inverse of the update rate), the update rules for each of the four 3-point glider types are (see kymographs in Figure 4):

$$c_{\text{con}+}(t, x) c_{\text{con}+}(t, x + \rho) c_{\text{con}+}(t + \delta, x + \rho) = +1,$$

$$c_{\text{con}-}(t, x) c_{\text{con}-}(t, x + \rho) c_{\text{con}-}(t + \delta, x + \rho) = -1,$$

$$c_{\text{div}+}(t, x) c_{\text{div}+}(t + \delta, x) c_{\text{div}+}(t + \delta, x + \rho) = +1,$$

$$c_{\text{div}-}(t, x) c_{\text{div}-}(t + \delta, x) c_{\text{div}-}(t + \delta, x + \rho) = -1.$$

The direction of the displacement of the spatial mean location of each triplet is inverted by inverting the sign of the pixel spacing. Starting from an initial seed state, the values of each pixel at each timepoint are determined by these update rules using the surrounding pixels' values. As we simulate the full 360 degrees of visual space, we use periodic boundary conditions to avoid undetermined edge pixel values. As in previous studies (Clark et al., 2014; Fitzgerald and Clark, 2015; Leonhardt et al., 2016), the pixel spacing was taken to be 5 degrees in both imaging experiments and numerical simulations. In imaging experiments, visual stimuli were generated and presented as described in previous studies (Badwan et al., 2019).

Random checkerboard stimuli (Figure 4)

Random checkerboard stimuli were constructed as in previous studies (Badwan et al., 2019). Briefly, 5-degree black or white bars were placed at random with a density of 40% on a gray background. The resulting checkerboards were then rigidly translated at a velocity of 100 degree/s. When combining rightward-

and leftward-moving stimuli, summation was defined such that 2 white bars summed to white, 2 black bars summed to black, and 1 white and 1 black bar summed to gray. Therefore the contrast of the composite stimulus matched that of the individual components, although its density rose to 64%.

Natural scene stimuli (Figure 5)

Following prior work (Chen et al., 2019; Clark et al., 2014; Fitzgerald and Clark, 2015), we generated a left-right symmetric ensemble of natural scenes by drawing independent row and column samples from the database gathered by (Meyer et al., 2014). In this ensemble, scenes were rigidly-translated at velocities sampled from a Gaussian distribution with a standard deviation of 100°/s, which roughly matches typical rotational velocities of walking flies (DeAngelis, Zavatone-Veth, & Clark, 2019; Katsov & Clandinin, 2008). To convert the scenes to contrast signals, we spatially filtered each image with the photoreceptor kernel to generate blurred images I_{blur} , and then used a Gaussian kernel with a standard deviation of 20° to estimate locally-averaged images I_{mean} . The contrast signal was then defined as (Chen et al., 2019)

$$c(x, y) := \frac{I_{\text{blur}}(x, y) - I_{\text{mean}}(x, y)}{I_{\text{mean}}(x, y)}.$$

As in previous studies of coactivation (Salazar-Gatzimas et al., 2018), the coactivations in Figure 5C were computed as normalized inner products of response timeseries. For all analyses in Figure 5, we used an ensemble with 10^6 elements.

Appendix B: Parameter value selection

In this appendix, we briefly describe how we selected values of the weighting parameters $g_{\text{exc}}/g_{\text{leak}}$ and $g_{\text{inh}}/g_{\text{leak}}$. We evaluated the model solely based on its ability to produce direction-opponent average responses to 1 Hz, 45-degree sinusoid gratings similar to those measured in T4 cells (Badwan et al., 2019). To do so, we considered the direction selectivity index and analogous indexes of direction-opponency and orthogonal direction enhancement, defined as

$$\text{DSI} := \frac{r(\text{PD}) - r(\text{PD})}{r(\text{PD}) + r(\text{PD})},$$

$$I_{\text{PD+ND}} := \frac{r(\text{PD} + \text{ND}) - r(\text{PD})}{r(\text{PD} + \text{ND}) + r(\text{PD})},$$

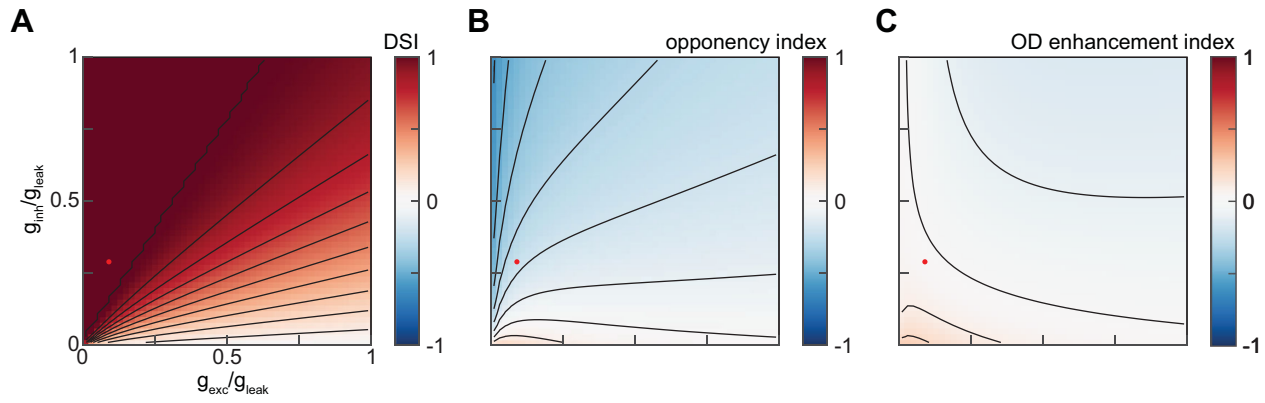


Figure B1. Sweeping the parameters of the T4 synaptic model. **A.** The direction-selectivity index of the T4 synaptic model’s responses to drifting sinusoid gratings as a function of g_{exc}/g_{leak} and g_{inh}/g_{leak} . Red dot indicates the selected values of $g_{exc}/g_{leak} = 0.1$ and $g_{inh}/g_{leak} = 0.3$. **B.** As in (A), but for the opponency index. **C.** As in (A), but for the OD enhancement index.

and

$$I_{PD+OD} := \frac{r(PD + OD) - r(PD)}{r(PD + OD) + r(PD)}.$$

As shown in Figure B1, there exists a broad region of parameter space for which the model produces responses with a similar degree of direction-opponency to that measured in T4 cells without significant PD+OD enhancement. We therefore made a simple choice of round-number values within that region.

Appendix C: LNLN cascade factorization of the T4 synaptic model

In this appendix, we show how our T4 synaptic model may be factorized as a product of linear-nonlinear-

linear-nonlinear (LNLN) cascades representing the numerator and denominator of the biophysical nonlinearity. The response C of the full model at each point in spacetime is given in terms of the filtered contrast signal s as

$$C = R^2 \left(\frac{\tilde{g}_1 E_{inh} + \tilde{g}_2 E_{exc} + \tilde{g}_3 E_{inh}}{1 + \tilde{g}_1 + \tilde{g}_2 + \tilde{g}_3} \right),$$

where we have defined $\tilde{g}_i := g_i/g_{leak}$ for brevity. Noting that the denominator of this expression is always positive, we may re-express the response as

$$C = \frac{R^2 (\tilde{g}_1 E_{inh} + \tilde{g}_2 E_{exc} + \tilde{g}_3 E_{inh})}{(1 + \tilde{g}_1 + \tilde{g}_2 + \tilde{g}_3)^2}.$$

Hence, the full EMD model admits a factorization into a product of LNLN models as

$$C(t, x) = N(t, x) D(t, x)$$

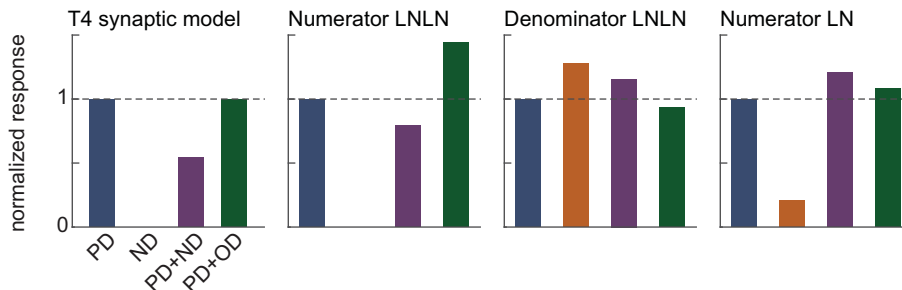


Figure C1. Sinusoid grating responses of different components of the LNLN factorization. *From left to right:* Average responses of the full T4 synaptic model, the numerator LNLN cascade, the denominator LNLN cascade, and the numerator LN cascade to 1 Hz, 45-degree sinusoid gratings. All responses are normalized by the response of the given component to a grating drifting in the PD of the full model.

where

$$N(t, x) := R^2 (\tilde{g}_1 E_{\text{inh}} + \tilde{g}_2 E_{\text{exc}} + \tilde{g}_3 E_{\text{inh}})$$

and

$$D(t, x) := (1 + \tilde{g}_1 + \tilde{g}_2 + \tilde{g}_3)^{-2},$$

which is bounded as $D(t, x) \leq 1$. Because $D(t, x) \leq 1$, $C(t, x) \leq N(t, x)$.

The denominator LNLN cascade D is the result of applying a convex function (x^{-2} for $x > 0$) to a non-negative linear combination of LN models with convex nonlinearities. Therefore it cannot generate direction-opponent (DO) average responses to sinusoid gratings. The proof of this proposition is a minor extension of our previous results on LNLN models with continuously-differentiable convex nonlinearities and non-negative secondary linear filters (Badwan et al., 2019). We define the soft ramp function

$$R_\beta(x) := \beta^{-1} \log(1 + \exp(\beta x)),$$

which is a continuously differentiable, monotone increasing, non-negative, and convex function of x for all positive β . As $\beta \rightarrow \infty$, $R_\beta(x) \rightarrow R(x)$ pointwise. By continuity, defining $D_\beta(t, x)$ using R_β , we have $0 \leq D_\beta(t, x) \rightarrow D(t, x) \leq 1$ as $\beta \rightarrow \infty$. We denote the nonlinear functional corresponding to the spacetime average of $D_\beta(t, x)$ for some input stimulus f as $D_\beta[f]$. As we have the integrable constant dominating function 1, by the Lebesgue dominated convergence theorem, we have $0 \leq D_\beta[f] \rightarrow D[f] \leq 1$ as $\beta \rightarrow \infty$ (Barry, 2015). By

the result of (Badwan et al., 2019), we know that $D_\beta[PD + ND] \geq D_\beta[PD]$ and $D_\beta[PD + ND] \geq D_\beta[ND]$, where $D_\beta[PD]$, $D_\beta[ND]$, and $D_\beta[PD + ND]$ are the average responses to PD, ND, and PD+ND sinusoid gratings, respectively. As these inequalities hold pointwise for all positive β , by taking $\beta \rightarrow \infty$ we may obtain $D[PD + ND] \geq D[PD]$ and $D[PD + ND] \geq D[ND]$. Therefore the denominator LNLN cascade cannot generate DO average responses to sinusoid gratings.

However, as the numerator LNLN model is the result of applying a convex function to a non-convex linear combination of LN models with convex nonlinearities, we cannot analytically exclude the possibility that it could generate DO average responses to sinusoid gratings using the results of (Badwan et al., 2019). In fact, numerical simulation shows that it can generate DO average responses to sinusoid gratings, though it generates strong PD+OD enhancement (Figure C1). It also generates DO responses over a smaller region in spatiotemporal frequency space than the full model. If one replaced the infinitely sharp ramp functions with more biophysically plausible soft rectifiers, the numerator LNLN cascade would be well-approximated for small input contrasts by a LN model with a quadratic nonlinearity. Therefore, it could not generate DO average responses for sufficiently small input contrasts. However, even in the limit in which both the numerator and denominator are represented as LN models with quadratic nonlinearities, the full model could likely generate DO average responses. In particular, this limiting construction would resemble a type of adaptive gain model, which was previously shown to generate DO average responses (Badwan et al., 2019).