Revised 31 January 2025

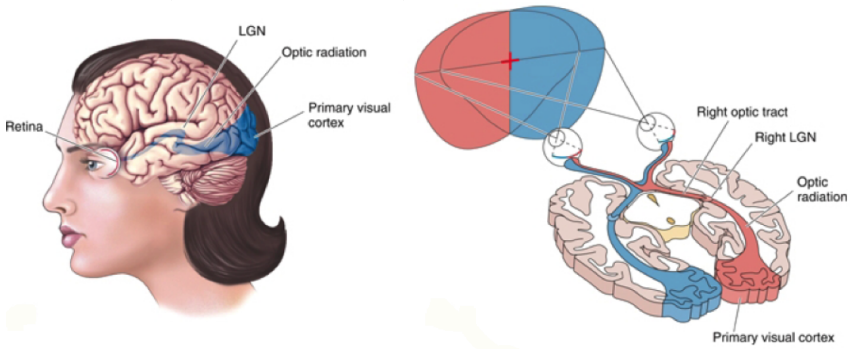
4 Recurrent neuronal networks: Invariant tuning and continuous attractors

We continue in our quest to connect recurrent networks with biological neurobiological phenomena. One central issue is that of invariant tuning by sensory, albeit high order sensory, neurons, as well as by high-order heading cells to chart the direction of locomotion. Invariant tuning refers to the seemingly stable response of neurons to incomplete stimuli, stimuli who persistence fluctuates or whose intensity varies, or stimuli that must be decoded in the presence of multiple distractors. Just as a Hopfield network uses attractor dynamics through recurrence to complete missing information, recurrence in general can be used to complete the missing information in coding a sensory feature or stabilize the description of a feature in the face of distractors. We will start with a few examples through a review of the literature, then constructs a simple model that is motivated by the data.

4.0.1 Invariant tuning to orientation in mammalian primary vision

A classic case is that of the response of neurons in primary visual (V1) cortex to oriented bars, gratings, and/or edges that sweep across the visual field. We start with a quick overview of vision from image formation in the retina to responses in V1 cortex (Figure 1).

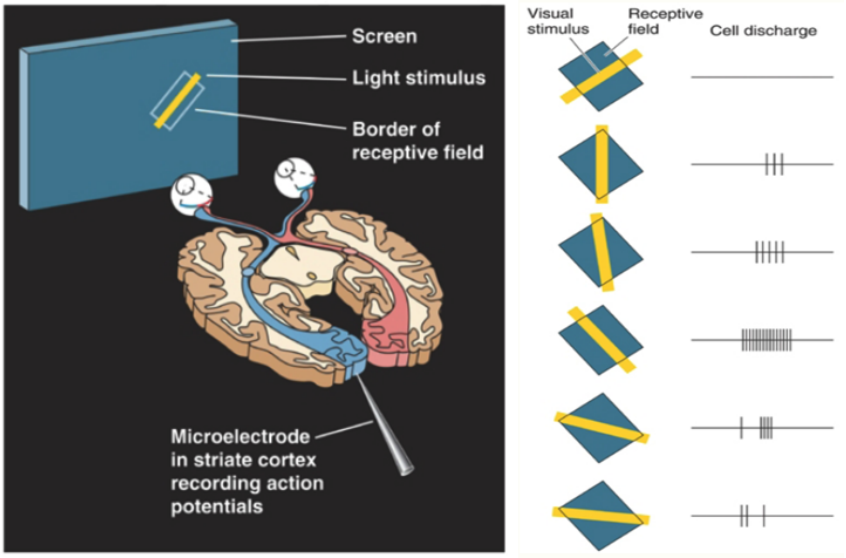
Figure 1: The gross layout of the visual stream in mammals, from the photoreceptors in the retina to primary visual cortex. From the textbook of Bear, Connors and Paradiso, 2007



Different cells respond to different orientations of the edge, which are most simply described in terms of a peak spike rate, a baseline rate, and a width of the angular modulation of the rate (Figure 2). This composite information defines the tuning curve.

Orientation specificity is believed to originate from the geometry of the input. Center-surround cells in the retina and thalamus respond like a Laplacian in all directions (Figure

Figure 2: The phenomenology of orientation specificity in V1 cortex. From the textbook of Bear, Connors and Paradiso, 2007



4.0.1). This symmetry is inconsistent with orientation coding. The inputs of many center-surround cells appear to have their inputs "line up"as they synapse on cortical neurons (Figure 4). This break in symmetry leads to orientation specificity for a moving bar or edge or grating.

The details of the orientation preference can occur solely from feedforward connections or from a mix of feedforward and recurrent connections. The data indicates that the kernel of orientation results from the input but that the width of the tuning curve results from cortical interactions.

By the way, in large brains, e.g., monkey, many neighboring cells have a similar preference for different orientations and thus form a map across the brain. Since space is mapped onto the cortical mantle, the attempt to map the three dimensions of space and orientation onto the two dimensional cortical mantle leads to fissures in the map (Figure 4.0.1). This does not occur in small brains. While a well known feature, we will ignore if in our presentation as it does not impact the circuitry of forming the orientation specificity of individual neurons.

Individual neurons that respond to the orientation of a stimulus also respond to the contrast of the scene; at modest to high light levels the contrast and not the absolute intensity determines the average spike rate so long as the modulation is not too slow nor too fast. They may also respond to other features like the spatial frequency of a patterned input. Three (or more) conundrums arise:

Contrast invariant tuning: The width of the tuning curve is independent of contrast (Figure 6). This appears to be inconsistent with feed-forward models, in which a fixed threshold would cause the width to increase with increasing contrast. This is referred to as the "iceberg" effect.

Size invariant tuning: The width of the tuning curve is largely independent of the as-

Figure 3: Many output cells from the retina, i.e., retinal ganglion cells, have a center-surround receptive field; Center OFF here. From the textbook of Bear, Connors and Paradiso, 2007

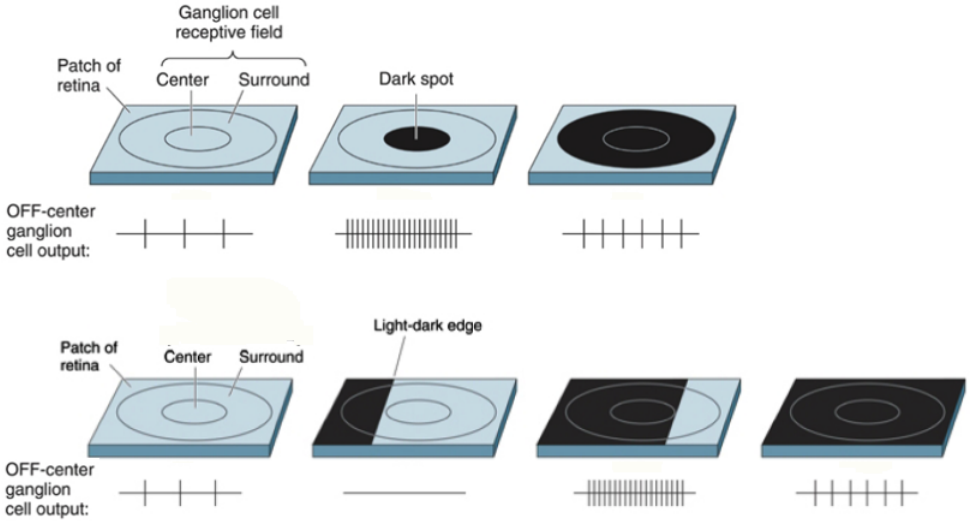
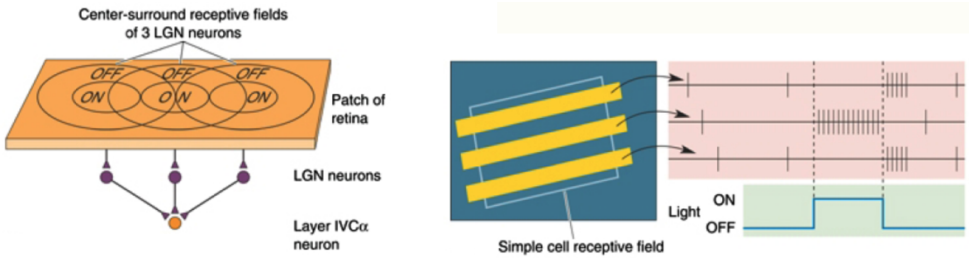


Figure 4: Cortex synthesizes orientation preference from center-surround receptive fields. From the textbook of Bear, Connors and Paradiso, 2007



pect ratio of the oriented bar. For small bars, this is inconsistent with a geometrically-based feed forward model, i.e., the Hubel-Wiesel model. More generally, it points to an invariance in the representation of a feature in the stimulus.

Spatial frequency invariant tuning: The width of the tuning curve is largely independent of the rate of spatial repetition of a grating, like the pickets in a fence.

A recurrent network with input tuned to orientation can use feedback connections to surmount these challenges. The stable states of the network are representations of features, i.e., preferred orientations of edges in the visual field. In fact, removing lateral interactions in cortex leads to a loss of tuning, supporting the notion of feedback for some if not all tuning properties of cortical neurons.

4.0.2 Invariant tuning to the spatial extent of touch

Invariance can refer to a signal that depends on the central location of a stimulus but not the spatial extent. This is seen for the case of vibrissa touch (Figure 8). The extent of the

Figure 5: Overlap of orientation preference for neurons at three different locations obtained with IOS imaging of all of V1 cortex. From Bonhoeffer and Grinvald, 1993

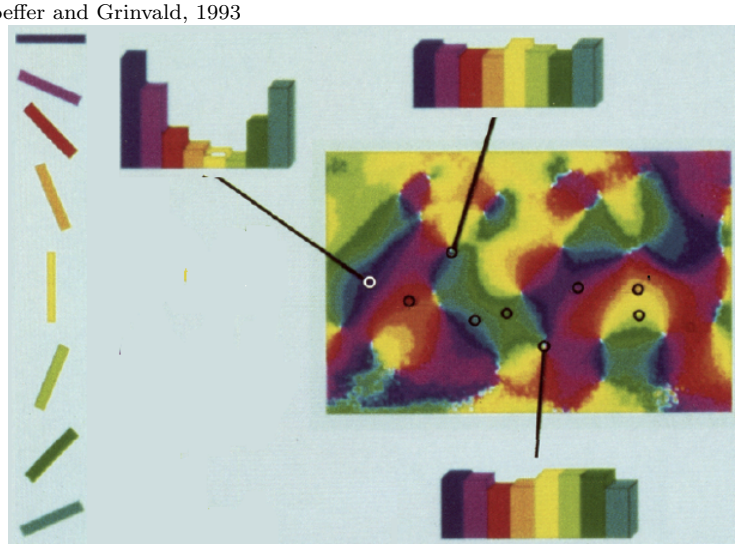
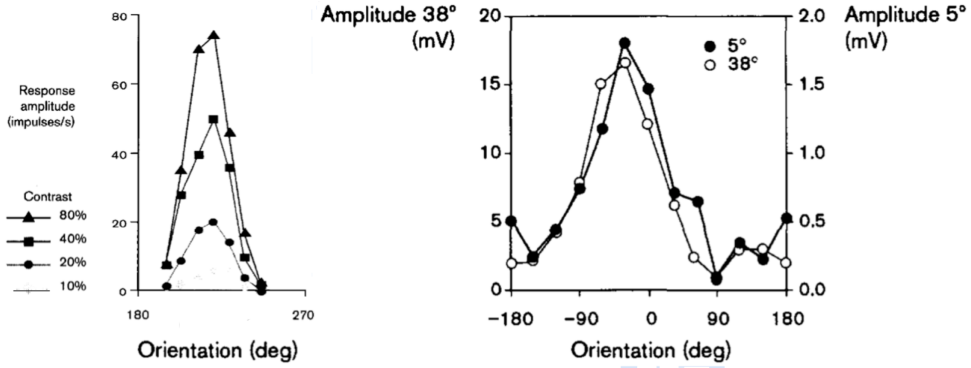


Figure 6: Invariance of the width of the orientation preference to contrast, from Sclar and Freeman 1962, and invariance to spatial frequency, from Ferster, Sooyoung and Wheat 1996



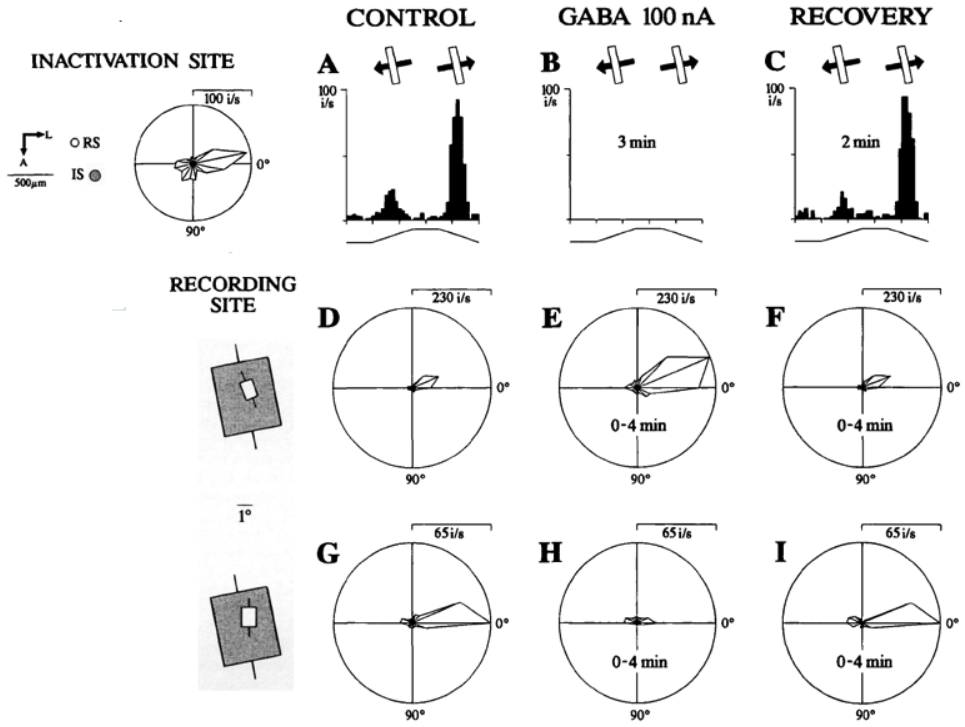
response in a neuron in primary vibrissa somatosensory (vS1) cortex that is sensitive to touch of the vibrissae is largely insensitive to the number of vibrissae that are activated.

4.0.3 Invariant tuning of neurons toward a heading

Neurons that are tuning to a particular heading have been long know, and more recently neurons that change their activity relative to he orientation of an animal toward or away from a landmark have been characterized (Figures 9 and 10). A special feature of these neurons is their immunity to distractors. As in the above case, a recurrent network with input tuned to heading can use feedback connections to surmount the challenge of distractors and incomplete input information. The stable states of the network are a manifold of preferred headings relative to a landmark in the sensory field.

One example is found in the head direction cells in anterior-dorsal (ADn) thalamus. A second example, and one that is particularly dramatic, is found in the ellipsoid body

Figure 7: Cortical interactions, as opposed to solely feedforward features, define the tuning width. From Crook, Kisvarday and Eysel



of the central complex of the fly (Figures 11, 12, and 13).

4.1 A rate model for neuronal firing

The notion of a tuning curve, with a smoothly varying rate of spiking, appears inconsistent with modeling using binary neurons with ON and OFF states. Yet we can maintain the idea of binary neurons but think of modeling where each element, or index, refers to a group of cells with similar coding properties. We thus add one level of complexity and consider that each group, with index "i", can spike with a rate, $r_i(t)$, that is a monotonic function of the input. The ideas of rate is that we are averaging over many spikes and counting spikes per unit time as opposed to just labeling the output ON or OFF. Rate can be be construed as the number of binary neurons within a group that are firing.

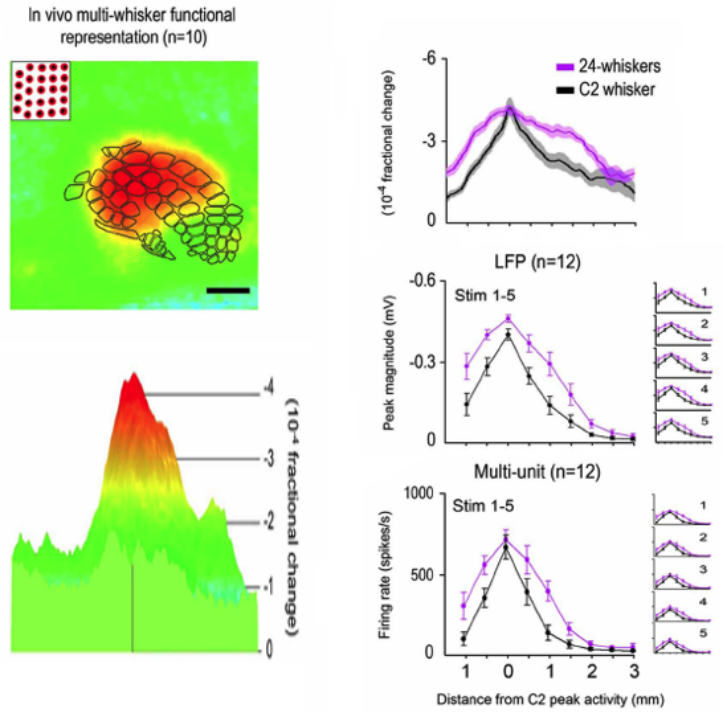
A simple function is motivated by $S_i = tanh\left[2G(\mu_i - \theta_i)\right]$, where G is the gain, is

$$\begin{aligned} r_i &= \frac{S_i + 1}{2} \\ &= \frac{1}{e^{-G(\mu_i - \theta_i)} + 1} \end{aligned} \tag{4.1}$$

where the slope has a value of G in the linear region, i.e.

$$r_i = \begin{cases} 0 & \text{if } \mu_i \ll \theta_i \\ G\left(\mu_i - \theta_i\right) & \text{if } \mu_i \approx \theta_i \\ +1 & \text{if } \mu_i \gg \theta_i. \end{cases} \tag{4.2}$$

Figure 8: Invariance of the amplitude of the vibrissa touch response. Composite data. From Chen-Bee, Zhou, Jacobs, Lim and Frostig. 2012



Another useful function is the piecewise-linear function

$$[r_i]_+ = \begin{cases} 0 & \text{if } \mu_i < \theta_i \\ G(\mu_i - \theta_i) & \text{if } \mu_i > \theta_i. \end{cases} \tag{4.3}$$

where θ_i now refers to the transition from OFF to the onset of spiking.

4.2 The 'Ring' Model of Recurrent Interactions

We now consider a particular model, the so called "ring" model, as a demonstration of how recurrent connections and the threshold in the (piece-wise linear) gain curve can lead to a powerful computation that preserves stimulus invariance. We will write the rate equations for motion over the full range of 2π radians, which is suitable to describe heading, as described previously. A similar set of equations can be written for the case of orientation, except that this covers π radians. There are many substantiations (Figure 15) - we follow the one includes global inhibition and "cosine" tuning.

Every neuron is labeled with an index, "i", that refers to the angle of the heading that is most likely to cause the cell to spike. This is the "preferred heading" and we assume that these are uniformly distributed across a sea of neurons, so that

$$\phi_i = \frac{2\pi}{N} i \quad \forall i \tag{4.4}$$

where N is the total number of neurons. The rate equation for a neuron with preferred

]

Figure 9: Bumps of activity in the heading direction system in rodent thalamus/ From Peyache, Lacroix, Petersen and Buzsaki 2015

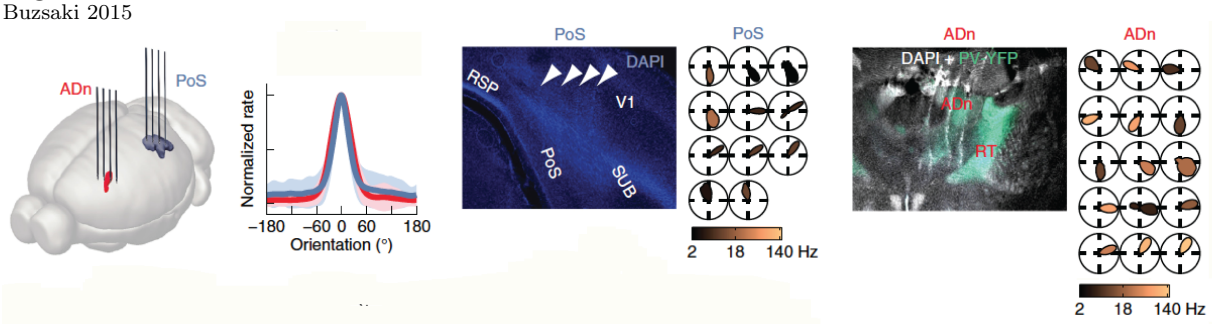
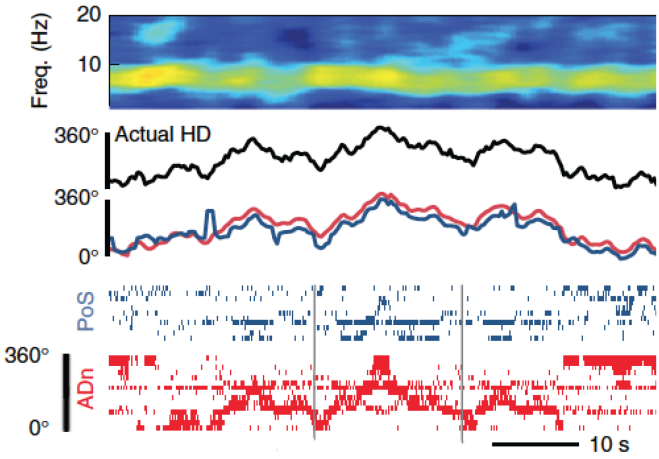


Figure 10: Bumps of activity in the heading direction system in rodent thalamus. From Peyache, Lacroix, Petersen and Buzsaki 2015



heading ϕ_i is

$$\tau \frac{dr_i(t)}{dt} + r_i(t) = f\left[\frac{1}{N}\sum_{j=1}^{N} W(\phi_i,\phi_j)\, r_j(t) + I^{ext}(\phi_i,\phi_0,t) - \theta_i\right] \tag{4.5}$$

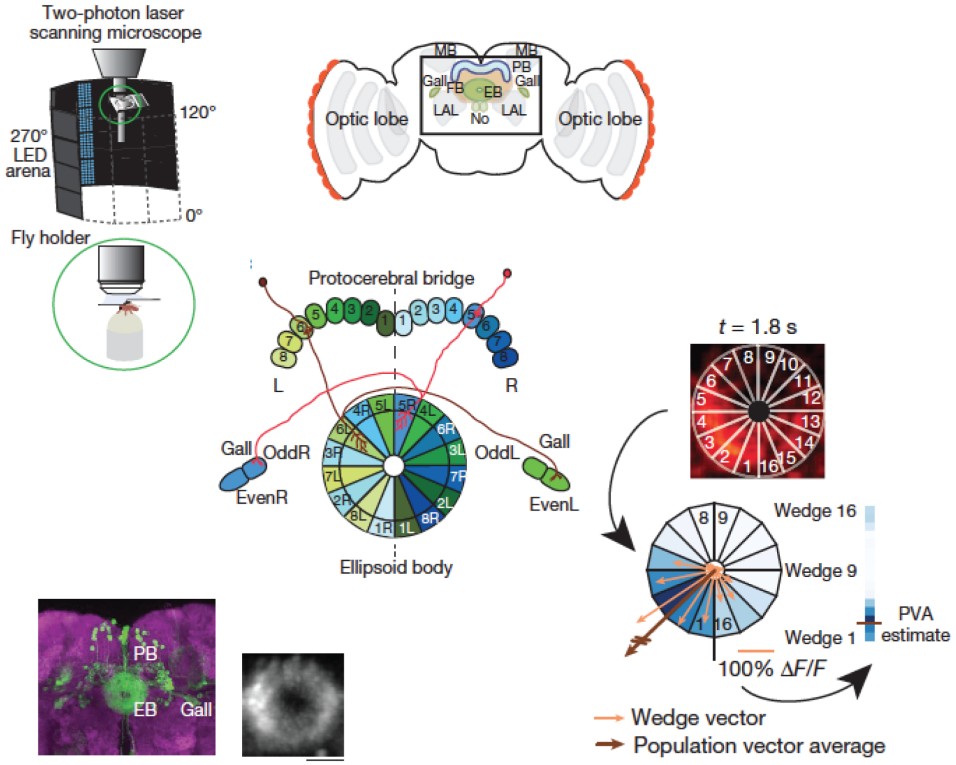
where $W(\phi_i,\phi_j)$ is the interaction between cell i and cell j, ϕ_0 is the orientation of a vector to the landmark (for heading) or of an external edge (for coding in vision), and θ is the threshold for spiking. The function $f[\cdot]$ is a nonlinear function that saturates at zero and possibly at a maximum firing rate. One such model is a logistic function.

We will shift to the continuum limit, where $r_i(t) \rightarrow r(\phi,t)$ and $\frac{1}{2\pi}\int_{-\pi}^{\pi} d\phi'$ replaces $\frac{1}{N}\sum_{j=1}^{N}$. Thus

$$\tau \frac{dr(\phi,t)}{dt} + r(\phi,t) = f\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} d\phi'\, W(\phi,\phi')\, r(\phi',t) + I^{ext}(\phi',\phi_0,t) - \theta(\phi)\right] \tag{4.6}$$

Motivated by experimental observations in visual systems (Figure 16) and heading systems, we take the interactions to be a function of the difference in orientation preference angles, so that neurons with similar orientation preference have relatively stronger

Figure 11: Bumps of activity in the landmark heading system in the fly ellipsoid body of the central complex thalamus. From Seelig and Jayaraman 2015



connections. Thus

$$W(\phi, \phi') = W(\phi - \phi') \tag{4.7}$$

and

$$I^{ext}(\phi, \phi_0, t) = I^{ext}(\phi - \phi_0, t). \tag{4.8}$$

We will write the interaction in terms of a constant term plus one term that varies as a function of the in-plane heading preference between two cells. This is equivalent to the first two terms in a Fourier expansion in $(\phi - \phi_0)$ of the interactions. We have

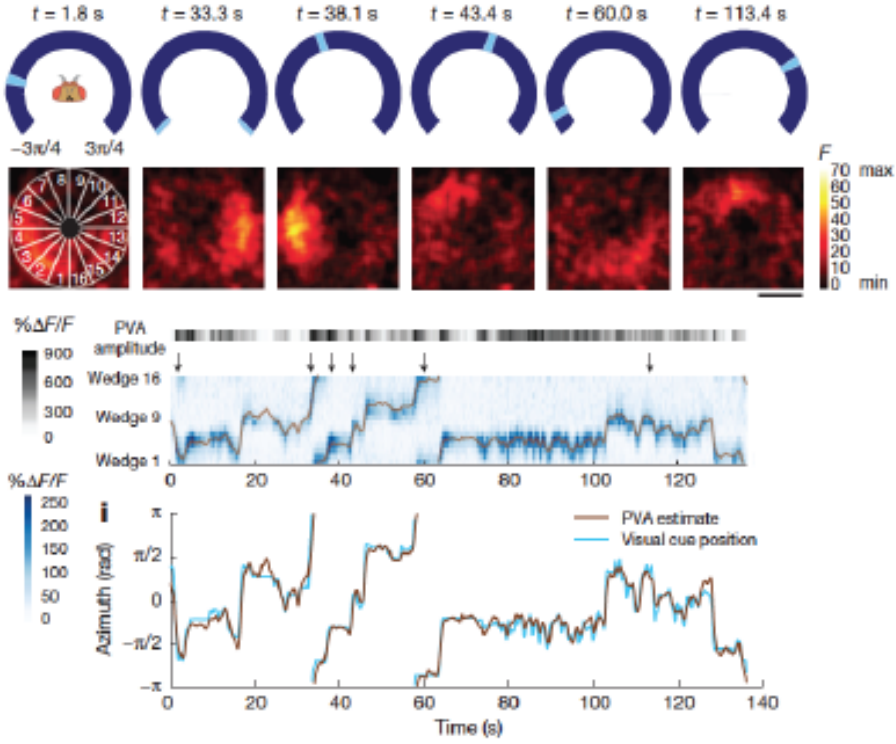
$$W(\phi - \phi') = W_0 + W_1 \cos(\phi - \phi') \tag{4.9}$$

where W_0 and W_1 are constants. We consider only a *cosine* term, and discard the *sine* term, as the connections should be symmetric with respect to the difference in orientation preference (Figure 17). One can add higher order terms to describe more complicated (and realistic) patterns of connectivity, but the basic lessons will be unchanged. Similarly, the experimental stimulus can be written in terms of a constant and an orientation dependent term

$$I(\phi - \phi_0, t) = \hat{I}_0(t) + \hat{I}_1(t) \cos(\phi - \phi_0). \tag{4.10}$$

The *cosine* is the leading term for the projection of a moving bar on a linear array of center-surround detectors. Yet we also need to be careful that the input is always positive,

Figure 12: Bumps of activity in the landmark heading system in the fly ellipsoid body of the central complex thalamus. From Seelig and Jayaraman 2015



so we re-express this as an overall drive and a modulation, $\epsilon(t)$, of the drive, i.e.,

$$\begin{aligned} I(\phi - \phi_0, t) &= I_0(t)\,(\,1\,+\,\epsilon(t)\,[1\,+\,\cos\,(\phi\,-\,\phi_0)]\,) \\ &= I_0(t)\,[1+\epsilon(t)] + I_0(t)\epsilon(t)\,[\cos\,(\phi\,-\,\phi_0)] \end{aligned} \tag{4.11}$$

where, for completeness, $\hat{I}_0(t) = I_0(t)[1+\epsilon(t)]$ and $\hat{I}_1(t) = I_0(t)\epsilon(t)$ and the selectivity of the input for modulated activity is

$$\begin{aligned} \textbf{Selectivity of input} &\equiv \frac{2\hat{I}_1}{\hat{I}_0} \\ &= \frac{2\epsilon}{1+\epsilon} \end{aligned} \tag{4.12}$$

which varies between 0 and 1 as the value of ϵ ranges from 0 to 1.

Putting all of this together yields a rate equation as a function of orientation and time

$$\begin{aligned} \tau\,\frac{dr(\phi,t)}{dt} + r(\phi,t) &= f\Big\{\frac{W_0}{2\pi}\int_{-\pi}^{\pi} d\phi' r(\phi',t)\;+\;\frac{W_1}{2\pi}\int_{-\pi}^{\pi} d\phi' r(\phi',t)\cos\,(\phi-\phi') \\ &+\quad I_0(t)\,[1+\epsilon(t)] +\; I_0(t)\epsilon(t)\cos\,(\phi-\phi_0)\;-\;\theta(\phi)\Big\}. \end{aligned} \tag{4.13}$$

Figure 13: Bumps of activity in the landmark heading system in the fly ellipsoid body of the central complex thalamus in the presence of interfering stimuli. From Seelig and Jayaraman 2015

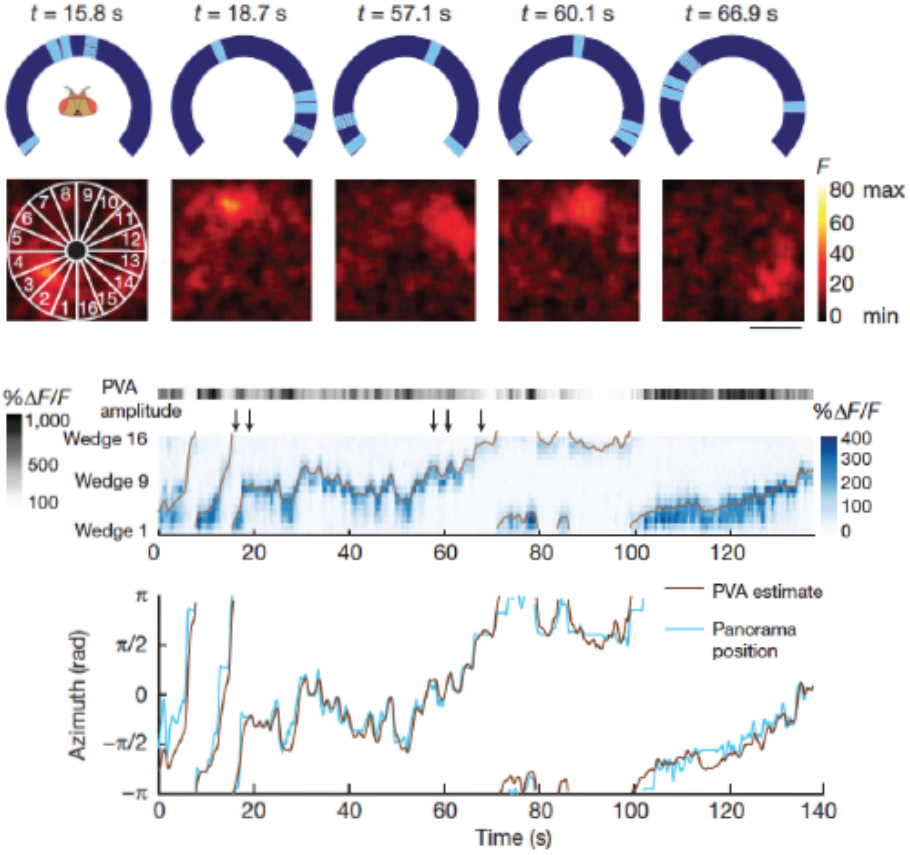
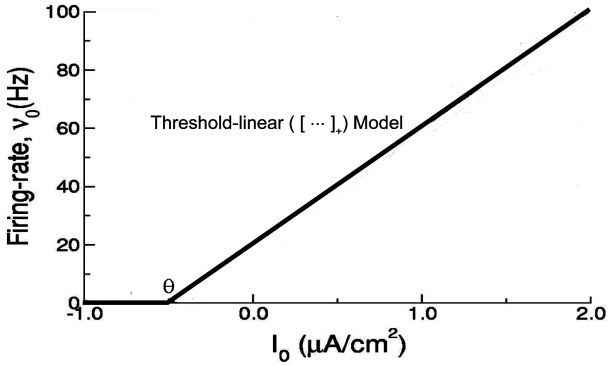


Figure 14: Neuronal piece-wise linear f-I function.



4.2.1 Mean field approach

We solve the coupled rate equations by introducing two parameters, referred to as "order parameters", that will represent the mean activity of neurons in the network and the modulation of the activity of neurons in the network. This will allow us to rewrite equation for the network (Eq. 4.14) in terms of the mean rate of spiking and the modulation of

Figure 15: The global model (presented here) and models with local interactions (better suited for the fly) both lead to a moving bump

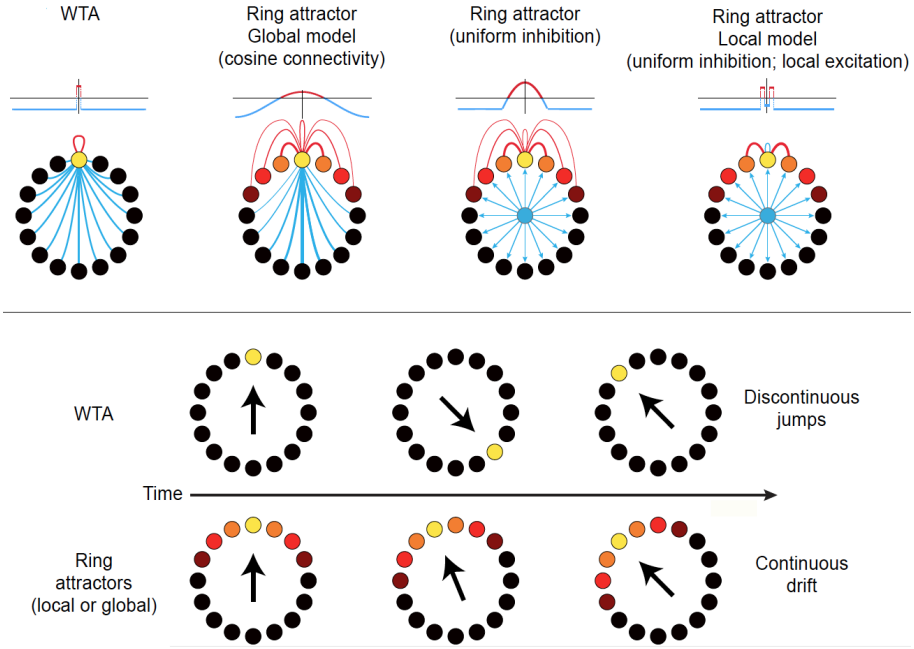
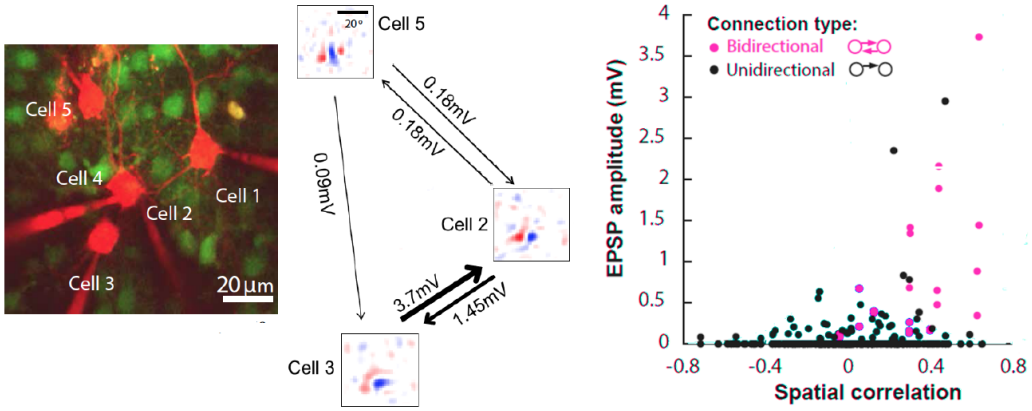


Figure 16: Connectivity among neurons in mouse V1 cortex is stronger for cells with overlapping receptive fields. From Cossell, Iacaruso, Muir, Houlton, Sader, Ko, Hofer and Mrsic-Flogel 2015.

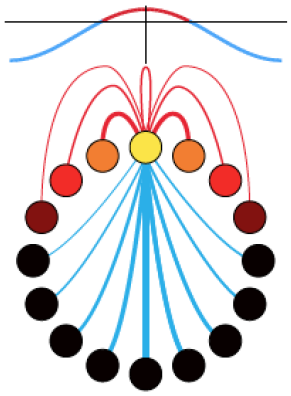


that rate. These two new parameters must evaluated in a self consistent manner. The order parameters are typically measured in the laboratory, although they are usually not identified as such, so providing a connection between theory and experiment.

Mean rate: We define $r_0(t)$ as the average firing rate of neurons in the network. This order parameter is an average over ϕ, i.e.,

$$r_0(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} d\phi' r(\phi', t) \tag{4.14}$$

Figure 17: Connectivity among neurons from W_0 multiplying a constant term and W_1 multiplying a cosine term.



Thus the W_0 term is just a synaptic weight times the mean activity $r_0(t)$.

Modulated rate: We define $r_1(t)$ as the average modulation of the firing rate of neurons in the network. This order parameter is a complex valued function that can be expressed in terms of magnitude $|r_1(t)|$ and phase $\psi(t)$, i.e.,

$$r_1(t) \equiv |r_1(t)|e^{-i\psi(t)}. \tag{4.15}$$

This order parameter is calculated as a weighted average of the activity with respect to the complex phase, or angle, and is given by

$$r_1(t) = \frac{1}{2\pi}\int_{-\pi}^{\pi} d\phi' r(\phi', t)\, e^{-i\phi'} \tag{4.16}$$

This allows us to evaluate the W_1 term as (see Box 1)

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} d\phi' r(\phi', t)\, \cos(\phi - \phi') = |r_1(t)|\, \cos(\phi - \psi(t)). \tag{4.17}$$

The mean field rate equation is thus transformed to

$$\begin{aligned}\tau \frac{dr(\phi,t)}{dt} + r(\phi,t) &= f\Big\{ W_0 r_0(t) + W_1|r_1(t)|\, \cos(\phi - \psi(t)) \\ &\quad + I_0(t)\,[1 + \epsilon(t)] + I_0(t)\epsilon(t)\, \cos(\phi - \phi_0)) - \theta(\phi)\Big\}.\end{aligned} \tag{4.18}$$

Now we have three simpler equations to solve; the mean field equation (Eq. 4.19) and two order parameter equations (Eqs. 4.14 and 4.16).

4.3 Equilibrium

The goal is to understand how the network dynamics can amplify a signal so that a weak input can drive a full cortical response. Can this goal can be achieved under equilibrium conditions? The rate equation becomes

$$r(\phi) = f\Big\{W_0 r_0 + W_1|r_1|\cos(\phi-\psi) + I_0(1+\epsilon) \;+\; I_0\,\epsilon\,\cos(\phi-\phi_0) - \theta(\phi)\Big\}. \quad (4.19)$$

So long as the gain function "$f[\cdot]$" is monotonic, the output will be maximized by maximizing the operand. We make the assumption that ψ is chosen to maximize the firing rate, i.e.,

$$\begin{aligned} \frac{dr(\phi)}{d\psi}|_{\phi=\phi_0} &= -W_1|r_1|\sin(\phi_0-\psi) \quad (4.20)\\ &= 0 \end{aligned}$$

This gives

$$\psi = \phi_0 \quad (4.21)$$

and the equilibriumrate equation becomes

$$r(\phi) = f\Big\{\,[W_0 r_0 + I_0(1+\epsilon) - \theta] \;+\; [W_1|r_1| + I_0\,\epsilon]\cos(\phi-\phi_0)\Big\} \quad (4.22)$$

where we have clustered the input into a constant piece and a piece that is modulated by the orientation of the input. We further take the threshold to be the same for all neurons.

4.3.1 Super-threshold (linear) limit

Let's see what happens when the inputs are sufficiently large so that the neuron operates solely above threshold. We thus take $f[x] = x$. Then

$$r(\phi) = [W_0 r_0 + I_0(1+\epsilon) - \theta] \;+\; [W_1|r_1| + I_0\epsilon]\cos(\phi-\phi_0). \quad (4.23)$$

The functional dependence of $r(\phi)$ must follow the drive and thus vary as $\phi-\phi_0$. We can expend $r(\phi)$ as a Fourier series with coefficients that are identical to the order parameters and then equate terms with Equation 4.22. Thus

$$\tilde{r}(\phi) = r_0 \;+\; r_{+1}e^{i\phi} \;+\; r_{-1}e^{-i\phi} \quad (4.24)$$

where

$$r_0 \;\equiv\; \frac{1}{2\pi}\int_{-\pi}^{\pi} d\phi' r(\phi'), \quad (4.25)$$

$$\begin{aligned} r_{+1} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} d\phi' r(\phi')\, e^{-i\phi'} \quad (4.26)\\ &\equiv |r_{+1}|e^{-i\phi_0}. \end{aligned}$$

and

$$\begin{aligned} r_{-1} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} d\phi' r(\phi')\, e^{i\phi'} \\ &\equiv |r_{+1}|e^{i\phi_0}. \end{aligned} \tag{4.27}$$

Then

$$\begin{aligned} \tilde{r}(\phi) &= r_0 + |r_{+1}|\left(e^{-i(\phi_0-\phi)} + e^{i(\phi_0-\phi)}\right) \\ &= r_0 + 2|r_{+1}|cos(\phi - \phi_0). \end{aligned} \tag{4.28}$$

We now equate terms for the average and for the first harmonic, i.e.,

$$r_0 = W_0 r_0 + I_0\,(1+\epsilon) - \theta \tag{4.29}$$

or

$$r_0 = \frac{I_0\,(1+\epsilon) - \theta}{1 - W_0} \tag{4.30}$$

and

$$r_1 = \frac{W_1 r_1 + I_0\epsilon}{2}. \tag{4.31}$$

or

$$r_1 = \frac{I_0\epsilon}{2 - W_1}. \tag{4.32}$$

We see that, even for the linear case, there is the potential for gain in the modulation term when $W_1 \to 2$. We put all of the above together to write

$$\tilde{r}(\phi) = \left(\frac{I_0\,(1+\epsilon) - \theta}{1 - W_0}\right) + \epsilon\left(\frac{2I_0}{2 - W_1}\cos(\phi - \phi_0)\right). \tag{4.33}$$

How does this response help in altering the output of the network? We first let $\theta = 0$ without loss of generality to get the simplified form

$$\tilde{r}(\phi) = I_0\left[\frac{1+\epsilon}{1 - W_0} + \frac{2\epsilon}{2 - W_1}\cos(\phi - \phi_0)\right] \tag{4.34}$$

and then write the selectivity of the output as

$$\begin{aligned} \textbf{Selectivity of output} &\equiv \frac{|r_1|}{r_0} \\ &= \frac{2\epsilon}{1+\epsilon}\,\frac{1 - W_0}{2 - W_1} \\ &= \textbf{Selectivity of input} \times \left(\frac{1 - W_0}{2 - W_1}\right). \end{aligned} \tag{4.35}$$

recalling Equation 4.13. This is as far as linearity gets you. There is gain, and potentially very large gain, but *no* invariance! In the linear case, the input determines the output. Thus the choice $\epsilon = 0$ will lead to $r_1 = 0$ and no modulation of the neuronal activity, despite the angular dependence of the interactions.

4.3.2 Marginal (spontaneous bump) state

We now introduce an angular dependence to the activity, a bump along ϕ, by allowing for nonlinearity in the gain functions and an increase in the interaction terms so that $W_1 > 2$. Clearly, some neurons will be ON and some OFF so that the average modulation is bounded. For simplicity, we take $f[x]$ as threshold linear, i.e.

$$f[x] = [x]_+. \tag{4.36}$$

Then we expect that $|r_1| > 0$ even if $\epsilon = 0$. In this case, we expect a bump of neuronal activity that is centered around the average direction of phase, denoted ψ. With the choice $\epsilon = 0$, the bump is spontaneous and the choice of direction is random. The idea is that the smallest input, i.e., smallest value of ϵ, will seed the angular position of the bump.

$$\begin{aligned} r_1 &= \frac{1}{2\pi}\int_{\psi-\pi}^{\psi+\pi} d\phi' \left[r(\phi')\right]_+ e^{-i\phi'} \\ &= \frac{1}{2\pi}\int_{\psi-\pi}^{\psi+\pi} d\phi' \left[W_0 r_0 + I_0 - \theta + W_1|r_1|\cos(\phi'-\psi)\right]_+ e^{-i\phi'}. \end{aligned} \tag{4.37}$$

The bump is taken to have an extent with a half width of ϕ_C, which we will have to relate to the synaptic weights. Then we can write a self consistency equation for r_1 (see Box 2 for details) that incorporates the maximization $\psi = \phi_0$ (Eq. 4.21)

$$\begin{aligned} r_1 &= \frac{1}{2\pi}\int_{\psi-\phi_C}^{\psi+\phi_C} d\phi' \{[W_0 r_0 + I_0 - \theta + W_1|r_1|\cos(\phi'-\psi)] \\ &\qquad - [W_0 r_0 + I_0 - \theta + W_1|r_1|\ \cos(\psi \pm \phi_C - \psi)]\} e^{-i\phi'} \\ &= W_1\ r_1\ \frac{1}{2\pi}\left(\phi_C - \frac{1}{2}\sin\ 2\phi_C\right) \end{aligned} \tag{4.38}$$

so that

$$W_1 = \frac{4\pi}{2\phi_C - \sin\ 2\phi_C}. \tag{4.39}$$

Our result relates the synaptic strength to the width of activation. It means that the network will form a bump of activity with width $\pm\phi_C$. The minimum value of the connectivity, for the widest possible bump with $\phi_C = \pi$, is $W_1 = 2$ (Figure 18). This is just where the network is linear. Further, $\phi_C \to 0$ as $W_1 \to \infty$, i.e., stronger connections yield a narrower bump. In the absence of an input, the phase of the bump is arbitrary.

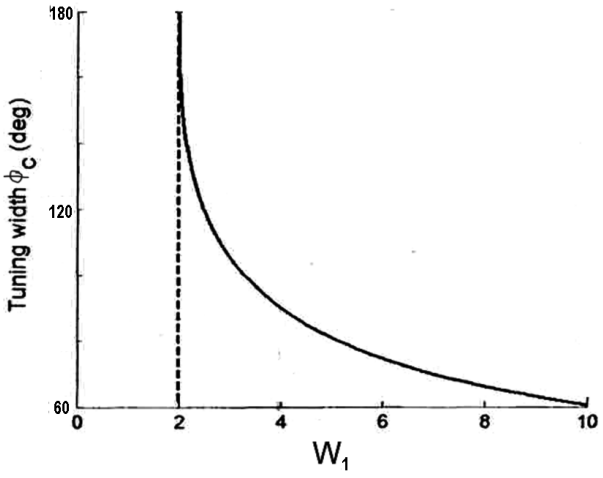
4.3.3 Symmetry breaking by a weak input

A weak input will pin the angular position, or phase, of the bump. So long as the stimulus is weak, the tuning does not depend on the stimulus parameters, i.e., on the selectivity of the input. To calculate the selectivity of the output we first need to calculate r_0 similarly to the analysis for the above analysis for r_1. We have (see Box 3 for details):

$$\begin{aligned} r_0 &= \frac{1}{2\pi}\int_{\psi-\pi}^{\psi+\pi} d\phi' \left[r(\phi')\right]_+ \\ &= r_1\ 4\ \frac{\sin\phi_C - \phi_C\ \cos\phi_C}{2\phi_C - \sin\ 2\phi_C}. \end{aligned} \tag{4.40}$$

]

Figure 18: Tuning width versus W_1



so that unlike the case for linear response, the average activity tracks the modulated activity. Thus we have

$$\begin{aligned} \textbf{Selectivity of output} &\equiv \frac{|r_1|}{r_0} \\ &= \frac{1}{4}\,\frac{2\phi_C - \sin 2\phi_C}{\sin\phi_C - \phi_C\ \cos\phi_C}, \end{aligned} \tag{4.41}$$

which varies between 1/2 and 1, i.e., by very little, as a function of ϕ_C. In this sense, the network will amplify a weak input and drive a response. Unlike the case of a feedforward network, where the width of the tuning curve depends on the input parameters I_0 and ϵ, here the width depends only on W_1. This satisfies the goal of invariance. The relation of W_1 to the width of the tuning curve constitutes a design rule for invariant tuning (Figure 19). There are also a restricted set of values of W_0 and W_1 that satisfy stability of the output (Figure 20).

We can now substitute in our expressions for W_1 (Eq. 4.39) and r_1 (Eq. 4.41) into the mean field equation (Eq. 4.22) to assess the competition between recurrent interactions and external input. We find

$$r(\phi) = f\left\{4W_0 r_0 + I_0\,(1+\epsilon) - \theta + \left(\frac{\pi}{\sin\phi_C - \phi_C\ \cos\phi_C} r_0 + I_0\epsilon\right)\cos(\phi - \phi_0)\right\}. \tag{4.42}$$

To complete the representation in terms of the width of the bump, we cook-up a form for $r(\phi)$ in terms of ϕ_C. A uniform distribution with $\theta = 0$ gives $r_0 = \phi_C/\pi$. A cosine bump gives

$$r_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} d\phi' \left[\cos\left(\frac{\pi}{2}\frac{\phi'}{\phi_C}\right)\right]_+ \tag{4.43}$$

Figure 19: Selectivity of the network as a function of W_1

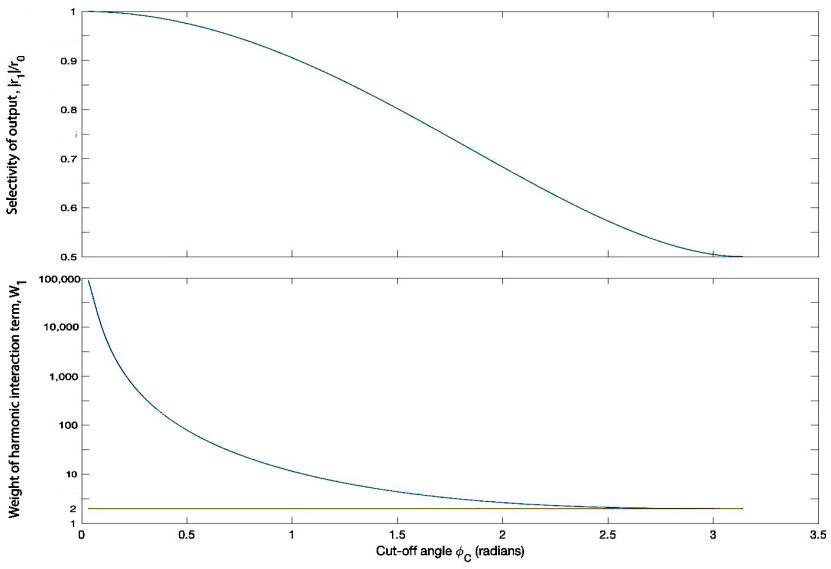
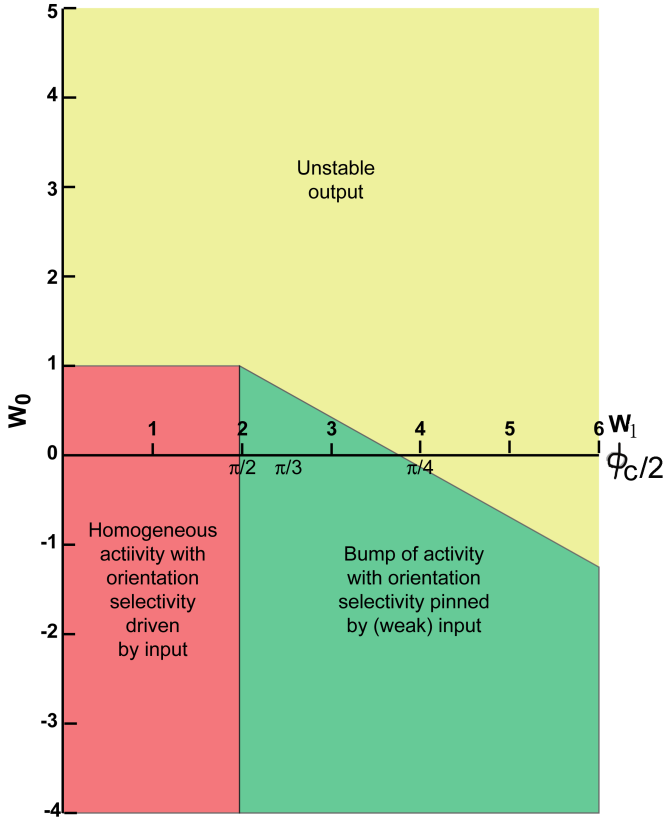


Figure 20: Phase diagram for different ranges of synaptic weights



$$= \frac{1}{2\pi}\int_{-\phi_0}^{\phi_0} d\phi' \cos\left(\frac{\pi}{2}\frac{\phi'}{\phi_C}\right)$$

$$= \quad \frac{2}{\pi^2}\,\phi_C$$

Then

$$r(\phi) = f\left\{\frac{8}{\pi^2}W_0\,\phi_C + I_0\,(1+\epsilon) + \left(\frac{2}{\pi}\,\frac{\phi_c}{\sin\phi_C - \phi_C\,\cos\phi_C} + I_0\epsilon\right) cos(\phi-\phi_0)\right\} \quad (4.44)$$

where the competition is between the recurrent term, i.e.,

$$\frac{2}{\pi}\,\frac{\phi_C}{\sin\phi_C - \phi_C\,\cos\phi_C} \rightarrow \begin{cases} \frac{2}{\pi} & \text{as } \phi_C \rightarrow \pi \\ \frac{1}{3\phi_C^2} & \text{as } \phi_C \rightarrow 0, \end{cases}$$

and the external input, i.e.,

$$I_0\epsilon, \quad (4.45)$$

both of which contribute to the $\cos(\phi-\phi_0)$ term. When recurrency wins, there will be a bump of activity. When the external input wins, the neurons track the input.

4.4 Predictions versus experiment

The ring model was motivated by experiments on the coding of orientation in visual stimuli. A number of predictions were made.

Contrast invariance: This is observed and was a motivation observation (Figure 24).

Transient onset of invariance: The invariance should arise slowly as this depends on recurrent connections. Thus the response of neurons at short times is expected to follow feedforward dynamics, while the response at later times, say after tens of milliseconds, would follow recurrent dynamics. This was not found.

Traveling bump: Recordings from the colliculus for eye position and from the anterior thalamus for heading suggest the notion of a moving bump of activity. This was predicted to occur in the visual system when the angle of the stimulus is rapidly changes. Here, activity would transiently pass through neurons that coding intermediate orientations (Figure 23). This was not found to date in V1 cortex, yet is seem nicely in the fly ellipsoid body (Figure 21).

Angular dependent connectivity: This is really a postdiction. Neurons with similar orientation preference tend to make stronger connections. Ditto for neurons in the same direction.

The great success of the model turns out to be with respect to heading, as seen in the activity of neurons in the ellipsoid body of the central complex of the fly and their manipulation by optogenetics (Figure 22). Neurons will code their preferred heading relative to the direction - call it ϕ_0 - of a landmark. Modifications of the ring model that more closely match the observed anatomy have been analyzed (Figure 15). Other work is trying to see if these models hold, in some fashion, for three-dimensional movement (Figure 25).

Figure 21: Evolution of neuronal activity in response to a shift in landmark. From Kim, Rouault, Druckmann and Jayaraman 2017.

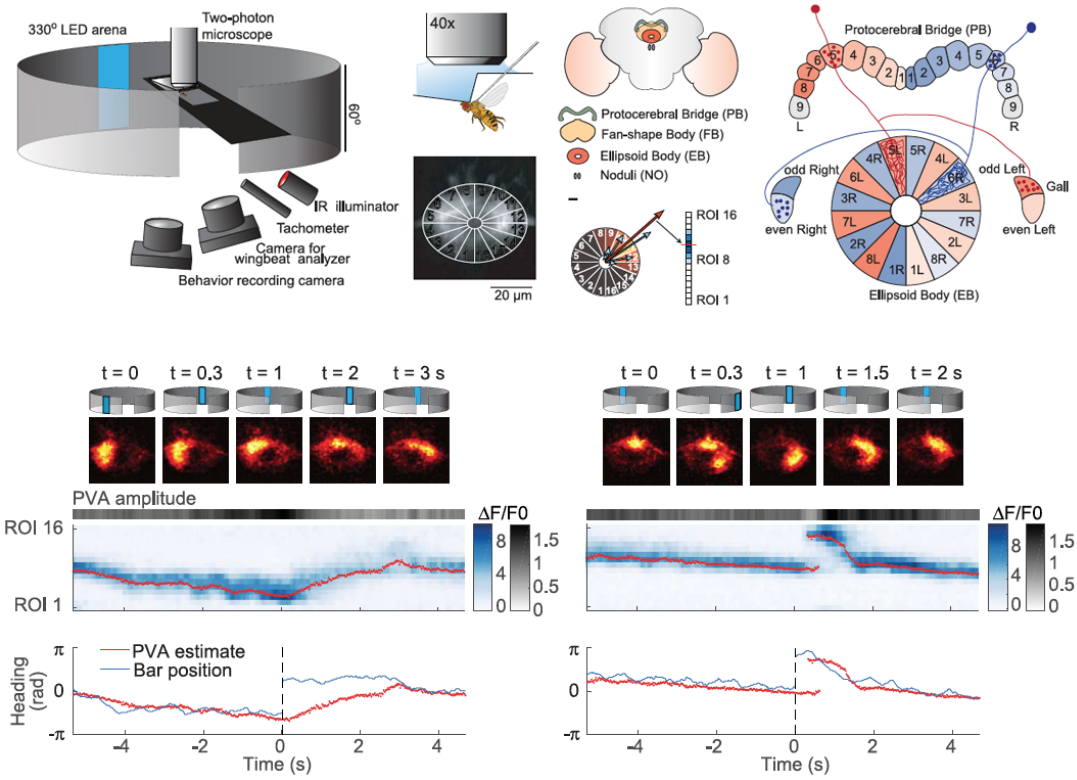
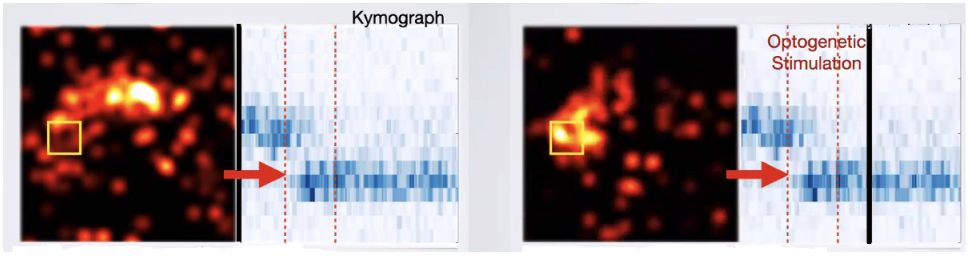


Figure 22: Optogenetic activation of heading cells leads to a stable state. From Kim, Rouault, Druckmann and Jayaraman 2017.



Box 1. General form of the order parameter for modulation

$$
\begin{aligned}
\frac{1}{2\pi}\int_{-\pi}^{\pi} d\phi' r(\phi',t)\ \cos\left(\phi-\phi'\right) &= \Re\{\frac{1}{2\pi}\int_{-\pi}^{\pi} d\phi' r(\phi',t) e^{i(\phi-\phi')}\} \\
&= \Re\{e^{i\phi}\ \frac{1}{2\pi}\int_{-\pi}^{\pi} d\phi' r(\phi',t) e^{-i\phi'}\} \\
&= \Re\{e^{i\phi}\ |r_1(t)| e^{-i\psi(t)}\} \\
&= |r_1(t)|\ \Re\{e^{i(\phi-\psi(t))}\} \\
&= |r_1(t)|\ cos\left(\phi-\psi(t)\right) \qquad (4.46)
\end{aligned}
$$

Figure 23: Evolution of neuronal activity in response to a shift in stimulus orientation from 0 to 60 degrees. From Hansel and Sompolinsky 1998.

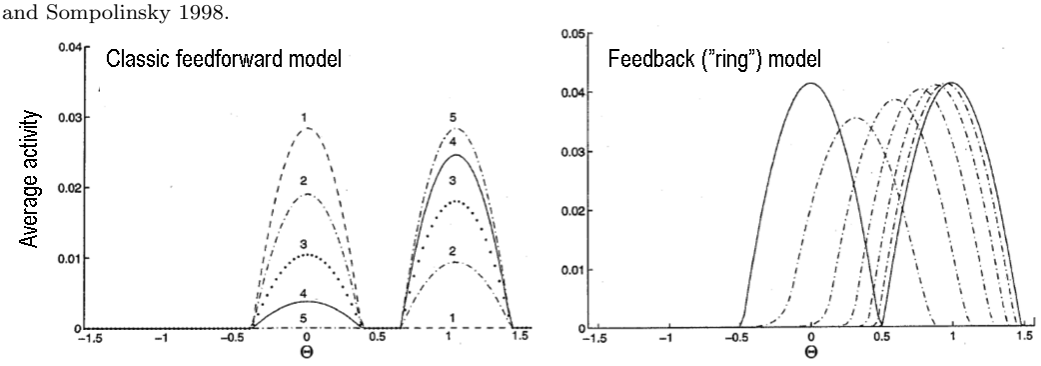


Figure 24: Contrast invariance with feedback. W_0 = -0.4, W_1 = 4.0, and ϵ = 0.09. From Hansel and Sompolinsky 1998.

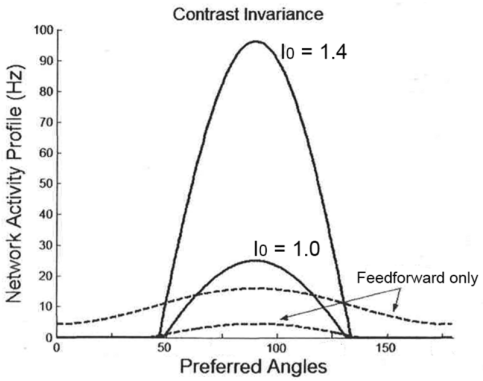
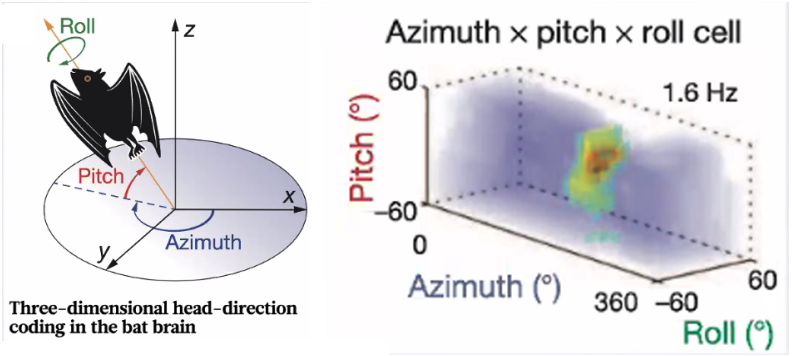


Figure 25: The ring model can be extended to describe coding in two and even three dimensions, as occurs in the heading of flying animals, such as bats



where $\Re$ means real part.

Box 2. The order parameter r_1 for the marginal phase

$$
\begin{aligned}
r_1 &= \frac{1}{2\pi}\int_{\psi-\phi_C}^{\psi+\phi_C} d\phi'([W_0 r_0 + I_0 - \theta + W_1|r_1|\cos(\phi'-\psi)] \\
&\quad - [W_0 r_0 + I_0 - \theta + W_1|r_1|\ \cos(\psi \pm \phi_C - \psi)])e^{-i\phi'} \\
&= W_1\ |r_1|\ \frac{1}{2\pi}\int_{\psi-\phi_C}^{\psi+\phi_C} d\phi'\left(\cos(\phi'-\psi) - \cos\ \phi_C\right) e^{-i\phi'} \\
&= W_1\ |r_1| e^{-i\psi}\frac{1}{2\pi}\int_{-\phi_C}^{\phi_C} dx\left(\cos\ x\ -\ \cos\ \phi_C\right) e^{-ix} \\
&= W_1\ r_1\ \frac{1}{2\pi}\int_{-\phi_C}^{\phi_C} dx\left(\frac{e^{ix}}{2}\ +\ \frac{e^{-ix}}{2}\ -\ \cos\ \phi_C\right) e^{-ix} \\
&= W_1\ r_1\ \frac{1}{2\pi}\int_{-\phi_C}^{\phi_C} dx\left(\frac{1}{2}\ +\ \frac{e^{-i2x}}{2}\ -\ \cos\ \phi_C\ e^{-ix}\right) \\
&= W_1\ r_1\ \frac{1}{2\pi}\left(\phi_C - \frac{1}{2}\sin\ 2\phi_C\right)
\end{aligned}
\tag{4.47}
$$

Box 3. The order parameter r_0 for the marginal phase

$$
\begin{aligned}
r_0 &= \frac{1}{2\pi}\int_{\psi-\pi}^{\psi+\pi} d\phi'\left[r(\phi')\right]_+ \\
&= \frac{1}{2\pi}\int_{\psi-\pi}^{\psi+\pi} d\phi'\left[W_0 r_0 + I_0 - \theta + W_1|r_1|\cos(\phi'-\psi)\right]_+ \\
&= \frac{1}{2\pi}\int_{\psi-\phi_C}^{\psi+\phi_C} d\phi'([W_0 r_0 + I_0 - \theta + W_1|r_1|\cos(\phi'-\psi)] \\
&\quad - [W_0 r_0 + I_0 - \theta + W_1|r_1|\ \cos(\psi \pm \phi_C - \psi)]) \\
&= W_1\ |r_1|\ \frac{1}{2\pi}\int_{\psi-\phi_C}^{\psi+\phi_C} d\phi'\left(\cos(\phi'-\psi) - \cos\ \phi_C\right) \\
&= W_1\ |r_1|\frac{1}{2\pi}\int_{-\phi_C}^{\phi_C} dx\left(\cos\ x\ -\ \cos\ \phi_C\right) \\
&= W_1\ |r_1|\ \frac{1}{\pi}\left(\sin\phi_C - \phi_C\ \cos\phi_C\right) \\
&= |r_1|\ 4\ \frac{\sin\phi_C - \phi_C\ \cos\phi_C}{2\phi_C - \sin\ 2\phi_C}.
\end{aligned}
\tag{4.48}
$$

Box 4 - Travel versus heading

From an ethological perspective the fly needs to compute its direction travel as opposed to just heading (Figure 26). The travel direction, denoted Ω, includes shifts in heading caused by wind. Recent work provides a basis for this calculation. There are four populations of neurons that code for wind direction, the PreFan body neurons (PF-Ns), that target neurons

in the Fan-shaped body. These are tuned to wind direction, but not heading (Figure 27). Heading and wind direction come together in the Fan-shaped body, and the question is how travel is computed (Figure 26). These neurons have responses whose amplitude depend on wind velocity, i.e., speed $|bfV|$ and direction Θ, and that code the direction of travel direction, denoted Ω as well. The four flavors of input can be described as

$$\begin{aligned}
\text{Rate of PF} - \text{N}_{\text{dL}} \text{ input} &= |V| \sin\Omega \cos\left(\Theta - \phi + \pi/4\right) && (4.49)\\
\text{Rate of PF} - \text{N}_{\text{dR}} \text{ input} &= |V| \cos\Omega \cos\left(\Theta - \phi - \pi/4\right) && (4.50)\\
\text{Rate of PF} - \text{N}_{\text{vL}} \text{ input} &= -|V| \sin\Omega \cos\left(\Theta - \phi - 3\pi/4\right) && (4.51)\\
\text{Rate of PF} - \text{N}_{\text{vR}} \text{ input} &= -|V| \cos\Omega \cos\left(\Theta - \phi + 3\pi/4\right) && (4.52)
\end{aligned}$$

Nature supplies a number of phase shifts through anatomical positioning of different copies of the inputs to $h\Delta B$ neurons. in the Fan Shaped Body (Figure 28); this turns out to be critical for summation of trigonometric functions! The output from these four cells are added together in the Fan Shaped Body and used to compute the direction of heading. All we need to recall to simply this summation is that $\cos(A+B) = \cos A \cos B$ - $\sin A \sin B$ and that $\cos(\pi/4) = \sin(\pi/4) = 1/\sqrt{2}$. Then

$$\text{Sum of rates of PF} - \text{N inputs} = \frac{|V|}{\sqrt{2}} \left[\sin\left(\Omega + \Theta - \phi\right) + \cos\left(\Omega + \Theta - \phi\right)\right]. \qquad (4.53)$$

This is maximum when the derivation is zero, or $\cos\left(\Omega + \Theta - \phi\right) = \sin\left(\Omega + \Theta - \phi\right)$, or $\Omega + \Theta - \phi = \pi/4$. Thus if different neurons code different values of Ω, the rate is greatest for the cell with

$$\Omega = \phi - \Theta + \frac{\pi}{4}. \qquad (4.54)$$

So the fly knows its travel direction, using information based on heading and wind direction, by the index of the neuron with the largest rate (Figure 26).

Figure 26: The travel direction complex of fly. From Lyu, Abbott and Maimon, Nature 2021

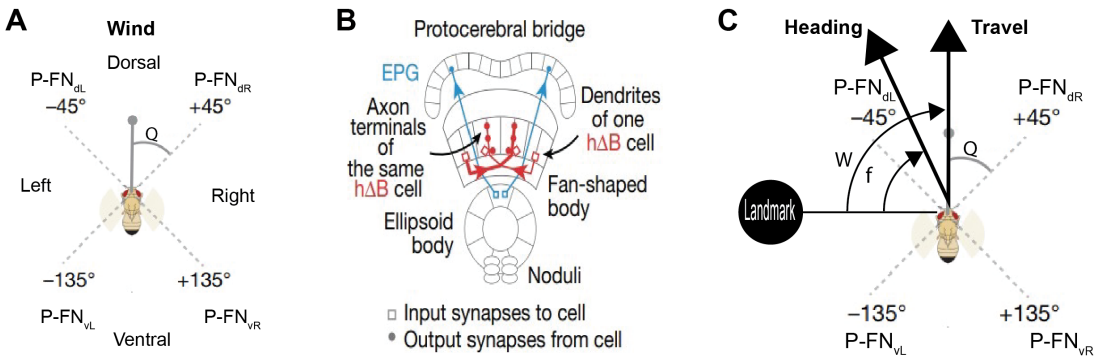


Figure 27: The response of PR-N neurons to changes in wind direction (top and middle row) and the lack of response with respect to changes in heading (bottom row). Note $\pm\pi/4$ and $\pm 3\pi/4$ shifts in physiological output (Figure 28). From Lyu, Abbott and Maimon, Nature 2021

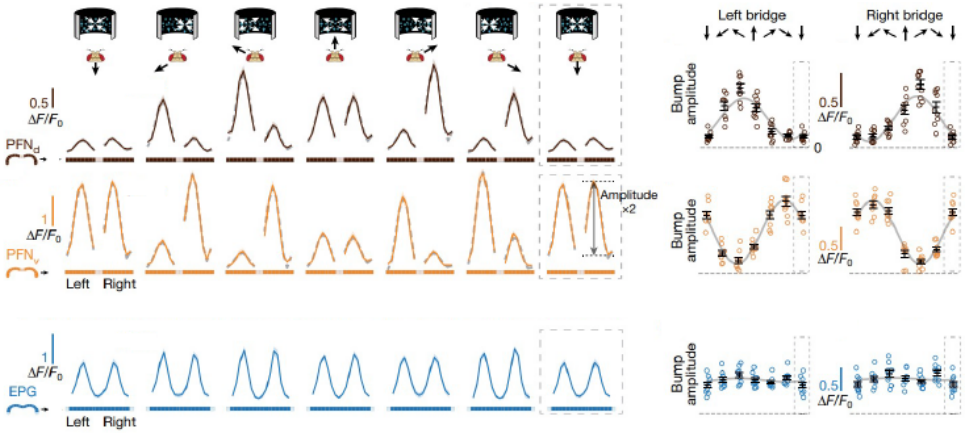


Figure 28: The direction complex of fly showing $\pm\pi/4$ and $\pm 3\pi/4$ anatomical offsets. From Lyu, Abbott and Maimon, Nature 2021