9 The Ring Model of Recurrent Interactions to Achieve Simulus-Invariant Tuning

An interested puzzle is posed by the response of neurons in V1 cortex to oriented bars and edges that move across the visual field. More commonly in a laboratory, experimenters use oriented and moving gratings. Different cells respond to different angles of the edge, which are most simply described in terms of a peak, a baseline, and a width to the angular modulation. This composite information defines the tuning curve. The same cells 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. Two conundrums, summarized, in the 1997 Current Opinions in Neurobiology article by Shapley and Sompolinsky, arise.

Contrast invariant tuning: The width of the tuning curve is independent of contrast. 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 aspect 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 invaraince in the representation of a feature in the stimulus.

Lets see if a recurrent network with input tuned to orientation can use feedback connections to surmount these challenges. In the sense, the stable states of the network are now representations of features, i.e., preferred orientations of edges in the visual field, as opposed to abstract memories.

9.1 Rate model

In the past we started with discrete modes, in which every neuron is labeled with an index, "i". Here the index refers to the angle of the stimulus that is most likely to cause the cell to spike. This is the "preferred orientation" and we assume that these are unifomly distributed, so that

$$\theta_i = \frac{\pi}{N}i \ \ \forall i \tag{9.9}$$

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

$$\tau \frac{r_i(t)}{dt} + r_i(t) \ = \ f\left[\sum_i^N W(\theta_i, \theta_j) \, r_j(t) \ + \ I_{ext}(\theta_i, \theta_0, t) \ - \ T\right] \tag{9.10}$$

where $W(\theta_i, \theta_j)$ is the interaction between cell i and cell j, θ_0 is the orientation of the external edge, and T is the threshold for spiking. The function f is a nonlinear function that saturates at zero and at a maximum firing rate. Without loss of generality, we take the maximum rate to be 1.

Motivated by experiment, we take the interactions to be a function of the difference in orientation preference angles. Thus $W(\theta_i, \theta_j) = W(\theta_i - \theta_j)$ and $I_{ext}(\theta_i, \theta_0, t) = I_{ext}(\theta_i - \theta_0, t)$. Further, we will write the interaction in terms of a constant term plus one term that varies as aa function of the angular orientation preference between two cells. Thus

$$W(\theta_i - \theta_j) = W_0 + W_2 \cos\left(2(\theta_i - \theta_j)\right) \tag{9.11}$$

where W_0 and W_2 are constants, the factor of two in the *cosine* term comes from the degeneracy of orientation on the circle, i.e., orientalon is modulo π and not 2π. Further, we consider only the *cosine* term and thus the connections should be symmetric with respect to the difference in orientation preference. Similarly, the experimental stimulus can be written in terms of a constant and an orientation dependent term

$$I(\theta_i - \theta_0, t) = \hat{I}_0(t) + \hat{I}_0(t) \cos(2(\theta_i - \theta_0)). \tag{9.12}$$

It will be useful to re-express this in terms of an overall drive and a modulation, $\epsilon(t)$, of the drive, i.e.,

$$I(\theta_i - \theta_0, t) = I_0(t)\left[1 + \epsilon(t) + \epsilon(t) \cos\left(2(\theta_i - \theta_0)\right)\right] \tag{9.13}$$

where, for completeness, $\hat{I}_0(t) = I_0(t)[1 - \epsilon(t)]$ and $\hat{I}_2(t) = I_0(t)\epsilon(t)$. Putting all of this together yields a rate equation as a function of orientation and time

$$\tau\, \frac{r(\theta, t)}{dt} + r_i(\theta, t) = f[\, W_0 \frac{1}{2\pi} \int_{-\pi}^{\pi} d\theta' r(\theta', t) + W_2 \frac{1}{2\pi} \int_{-\pi}^{\pi} d\theta' r(\theta', t) \cos\left(2(\theta - \theta')\right)$$
$$+ I_0(t)\left[1 - \epsilon(t)\right] + I_2(t)\epsilon(t) \cos\left(2(\theta - \theta_0)\right) - T] \tag{9.14}$$

9.1.1   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 the network and the modulation of the activity. This will allow us to write a single equation for the network in terms of the behavior of one neurons relative to the mean. These new parameters must evaluated in a self consistent manner.

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

$$r_0(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} d\theta' r(\theta', t) \tag{9.15}$$

Thus the W_0 term is just $r_0(t)$,

Modulated rate: We define $r_2(t)$ as the average modulation of the firing rate of neurons in the network. $r_2(t)$ is a complex number, so we write it as $|r_2(t)|e^{-i2\psi(t)}$ where

$$|r_2(t)|e^{-i2\psi(t)} \;=\; \frac{1}{2\pi}\int_{-\pi}^{\pi} d\theta' r(\theta',t)\, e^{i2\theta'} \tag{9.16}$$

This allows us to evaluate the W_2 term as

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} d\theta' r(\theta',t)\cos\left(2(\theta-\theta')\right) \;=\; \Re\{\frac{1}{2\pi}\int_{-\pi}^{\pi} d\theta' r(\theta',t)e^{i(2(\theta-\theta'))}\} \tag{9.17}$$

$$= \; \Re\{e^{i2\theta}\frac{1}{2\pi}\int_{-\pi}^{\pi} d\theta' r(\theta',t)e^{-i2\theta'}\}$$

$$= \; \Re\{e^{i2\theta}\,|r_2(t)|e^{-i2\psi(t)}\}$$

$$= \; |r_2(t)|\,\cos\left(2(\theta-\psi(t))\right)$$

Relation to Fourier series: Note that $r_0(t)$ and $r_2(t)$ are the coefficients for the constarnt terms and second harmonic terms in a Fourier series of $r(\theta,t)$.

The mean field rate equation is

$$\tau\frac{r(\theta,t)}{dt}+r(\theta,t) \;=\; f[\,W_0 r_0(t)\;+\;W_2|r_2(t)|\,\cos\left(2(\theta-\psi(t))\right)+\;I_0(t)\left(1\;-\;\epsilon(t)\right)$$

$$+\;I_0(t)\epsilon(t)\,\cos\left(2(\theta-\theta_0)\right)\;-\;T] \tag{9.18}$$

9.2   Steady state

A goal is to understand how the network dynamics can amplify a signal so that a weak input can drive a full cortical response. This goal can be achieved in steady state. The rate equation becomes

$$r(\theta) = f\left[W_0 r_0 + W_2|r_2|\cos\left(2(\theta-\psi)\right) + I_0\left(1-\epsilon\right) + I_0\epsilon\cos\left(2(\theta-\theta_0)\right) - T\right]. \tag{9.19}$$

So long as the gain function "f" is monotonic, the output will be maximized by maximizing the operant. We make the assumption that ψ is chosen to maximize the firing rate. This gives $\psi = \theta_0$ and the steady state rate equation becomes

$$r(\theta) = f\left[(W_0 r_0 + I_0(1-\epsilon) - T)\;+\;(W_2|r_2| + I_0\epsilon)\cos(2(\theta-\theta_0))\right] \tag{9.20}$$

where we have clustered the input into constant pieces and pieces that are modulated by orientation.

9.2.1   Superthreshold (linear) limit

We previously learned how, in the context of the line attractors model, feedback can help in the linear case. Lets see what happens here when the inputs are sufficiently

large so that the neuron operates solely above threshold. We thus take $f[x] = x$. Then

$$r(\theta) = (W_0 r_0 + I_0(1 - \epsilon) - T) + (W_2|r_2| + I_0\epsilon)\cos(2(\theta - \theta_0)). \qquad (9.21)$$

The functional dependence of $r(\theta)$ must follow the drive and thus vary as $\theta - \theta_0$. If we expend $r(\theta)$ as a Fourier series, we get

$$\tilde{r}(\theta') = \frac{1}{2\pi}\int_{-\pi}^{\pi} d\theta'\ r(\theta') + \frac{1}{2\pi}\int_{-\pi}^{\pi} d\theta'\ r(\theta')\ e^{i2\theta'} \qquad (9.22)$$
$$= r_0 + r_2.$$

Noting that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} d\theta'\ \cos(2\theta')\ e^{i2\theta'} = \frac{1}{2} \qquad (9.23)$$

we can equate terms for the average and for the second harmonic, i.e.,

$$r_0 = W_0 r_0 + I_0(1 - \epsilon) - T \qquad (9.24)$$

and

$$r_2 = \frac{W_2 r_2 + I_0\epsilon}{2}. \qquad (9.25)$$

This leads to

$$r_0 = \frac{I_0(1 - \epsilon) - T}{1 - W_0} \qquad (9.26)$$

and

$$r_2 = \frac{I_0\epsilon}{2 - W_2}. \qquad (9.27)$$

We see that, even for the linear case, there is the potential for gain in the modulation term when $W_2 \to 2$.

How does this gain help is altering the output of the network? To make a bit more progress, we can write the selectivity of the input for modulated activity as

$$\textbf{Selectivity of input} \equiv \frac{\hat{I}_2}{\hat{I}_0} \qquad (9.28)$$
$$= \frac{\epsilon}{1 - \epsilon}$$

and note that we can write the selectivity of the output, taking $T = 0$ for simplicity, as

$$\textbf{Selectivity of output} \equiv \frac{|r_2|}{r_0}$$
$$= \frac{I_0\ \epsilon/2}{1 - W_2/2}\frac{1 - W_0}{I_0(1 - \epsilon)} \qquad (9.29)$$
$$= \frac{1 - W_0}{2 - W_2} \times \textbf{Selectivity of input}.$$

This is as far as linearity gets you. Gain, but no invariance.

4

9.2.2 Marginal (spontaneous bump) state

In the linear case, the input determines the output. Thus the choice $\epsilon = 0$ will lead to $r_2 = 0$. Now suppose we increase the interaction term W_2 so that $W_2 ¿ 2$. Clearly we have to allow for a nonlinear gain of the input so that 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., $f[x] = [x]_+$. Then $|r_2| > 0$ even if $\epsilon = 0$, as see by

$$
\begin{aligned}
r_2 &= \frac{1}{2\pi} \int_{-pi}^{\pi} d\theta' [r(\theta')]_+ cos(2\theta') \qquad (9.30) \\
&= \frac{1}{2\pi} \int_{-\theta_C i}^{\theta_C} d\theta' \left([W_0 r_0 + I_0 - T + W_2 r_2 cos(2(\theta' - \psi))] - [W_0 r_0 + I_0 - T + W_2 r_2 cos(2\theta_C)] \right) cos(2\theta') \\
&= W_2 r_2 \frac{1}{2\pi} \int_{-\theta_C}^{\theta_C} d\theta' \left(cos(2(\theta' - \psi)) - cos(2\theta_C) \right) cos(2\theta').
\end{aligned}
$$

This will yield a constraint on W_2 that must reduce to $W_2 > 2$. We take $\psi = 0$ for simplicity, which makes r_2 real and positive. Then

$$
\begin{aligned}
r_2 &= W_2 r_2 \frac{1}{2\pi} \int_{-\theta_C}^{\theta_C} d\theta' [cos(2\theta') - cos(2\theta_C)] cos(2\theta') \qquad (9.31) \\
&= W_2 r_2 \frac{1}{2\pi} \left[\int_{-2\theta_C}^{2\theta_C} d\theta' cos^2(\theta') - cos(2\theta_C) \int_{-2\theta_C}^{2\theta_C} d\theta' cos(\theta') \right] \qquad (9.32) \\
&= W_2 r_2 \frac{1}{2\pi} \left[\frac{4\theta_C}{4} + cos(2\theta_c) sin(2\theta_c) - 2cos(2\theta_c) sin(2\theta_c) \right] \\
&= W_2 r_2 \frac{4\theta_C - sin(4\theta_C)}{4\pi}
\end{aligned}
$$

so that

$$
W_2 = \frac{4\pi}{4\theta_C - sin(4\theta_C)}. \qquad (9.33)
$$

Our result means that the network will form a bump of activity with width $\pm\theta_C$. The minimum value of the connectivity, for a wide bump with $\theta_C = \pi/2$, is $W_2 = 2$. This is just where the linear network fails. Further, $\theta_C \to 0$ as $W_2 \to \infty$, i.e., stronger connections yield a narrower bump. In the absence of an input, the phase of the bump is arbitrary.

9.2.3 Symmetry breaking by a weak input

A weak input will pin the phase of the bump. Weak means that $0 < \epsilon << 1/2$. So long as the stimulus is weak, the tuning does not depend on the stimulus parameters. 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_2. This satisfies the goal of invariance. This the relation of W_2 to the width of the tuning curve constitutes a design rule for invariant tuning.

9.3   Epilog

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

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

Moving bump: A number of recordings from the colliculus for eye position and from the anterior thalamus for heading suggest the notion of a moving bump of activity; we will return to this point in a moment. 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. This was not found.

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. Neurons will code their preferred heading relative to the direction - call it θ_0 - of a landmark.