7 The 'Ring' Model of Recurrent Interactions to Achieve Stimulus-Invariant Tuning

An interested puzzle is posed by the seemingly stable response of neurons to incomplete stimuli or stimuli who persistence fluctuates. A classic case is that of the response of neurons n V1 cortex to oriented bars, gratings, and/or edges that move across the visual field. 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 a 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 invariance in the representation of a feature in the stimulus.

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. A second case concerns tuning of neurons toward a heading, i.e., a direction that is defined by one or more landmarks.

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. 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.

7.1 Rate model

For will write our equations for motion over the full range of 2π radians, which is suitable to describe heading. A similar set of equations can be written for the case of orientation, except that this covers π radians. 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, so that

$$\phi_i = \frac{2\pi}{N} i \ \ \forall i \tag{7.7}$$

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

$$\tau \frac{r_i(t)}{dt} + r_i(t) \ = \ f\left[\frac{1}{N}\sum_{j=1}^{N} W(\phi_i,\phi_j)\, r_j(t) \ + \ I^{ext}(\phi_i,\phi_0,t) \ - \ \theta\right] \tag{7.8}$$

where $W(\phi_i,\phi_j)$ is the interaction between cell i and cell j, ϕ_0 is the orientation of the external edge, and θ 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 experimental observations, we take the interactions to be a function of the difference in orientation preference angles. Thus $W(\phi_i,\phi_j) = W(\phi_i - \phi_j)$ and $I^{ext}(\phi_i,\phi_0,t) = I^{ext}(\phi_i - \phi_0,t)$. Further, we will write the interaction in terms of a constant term plus one term that varies as aa function of the in=plane heading preference between two cells. Thus

$$W(\phi_i - \phi_j) = W_0 \ + \ W_1\, cos\,(\phi_i - \phi_j) \tag{7.9}$$

where W_0 and W_1 are constants. 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

$$I(\phi_i - \phi_0,t) = \hat{I}_0(t) \ + \ \hat{I}_1(t)\, cos\,(\phi_i - \phi_0)\,. \tag{7.10}$$

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

$$I(\phi_i - \phi_0,t) = I_0(t)\left[1 \ + \ \epsilon(t)\,(1 \ + \ cos\,(\phi_i \ - \ \phi_0))\right] \tag{7.11}$$

where, for completeness, $\hat{I}_0(t) = I_0(t)[1 + \epsilon(t)]$ and $\hat{I}_1(t) = I_0(t)\epsilon(t)$. Putting all of this together yields a rate equation as a function of orientation and time

$$\begin{aligned}\tau \frac{r(\phi,t)}{dt} + r_i(\phi,t) \ = \ & f\{\frac{W_0}{2\pi}\int_{-\pi}^{\pi} d\phi' r(\phi',t) \ + \ \frac{W_1}{2\pi}\int_{-\pi}^{\pi} d\phi' r(\phi',t) cos\,(\phi - \phi') \\ & + \ I_0(t)\left[1 + \epsilon(t)\right] \ + \ I_0(t)\epsilon(t)\, cos\,(\phi - \phi_0) \ - \ \theta\}\end{aligned} \tag{7.12}$$

where $\frac{1}{2\pi}\int_{-\pi}^{\pi} d\phi'$ replaces $\frac{1}{N}\sum_{j=1}^{N}$.

7.1.1 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 ϕ, i.e.,

$$r_0(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} d\phi' r(\phi', t) \tag{7.13}$$

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

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 number , so we write it as"

$$r_1(t) \equiv |r_1(t)|e^{-i\psi(t)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} d\phi' r(\phi', t) e^{-i\phi'} \tag{7.14}$$

This allows us to evaluate the W_1 term as

$$
\begin{aligned}
\frac{1}{2\pi} \int_{-\pi}^{\pi} d\phi' r(\phi', t) \cos(\phi - \phi') &= \Re\{\frac{1}{2\pi} \int_{-\pi}^{\pi} d\phi' r(\phi', t) e^{i(\phi - \phi')}\} \quad (7.15)\\
&= \Re\{e^{i\phi} \frac{1}{2\pi} \int_{-\pi}^{\pi} d\phi' r(\phi', t) e^{-i\phi'}\}\\
&= \Re\{e^{i\phi} |r_1(t)|e^{-i\psi(t)}\}\\
&= |r_1(t)| \cos(\phi - \psi(t))
\end{aligned}
$$

where $\Re$ means real part.

Relation to Fourier series: Note that $r_0(t)$ and $r_1(t)$ are the coefficients for the constant term and the first harmonic term in a Fourier series of $r(\phi, t)$.

The mean field rate equation is thus

$$
\begin{aligned}
\tau \frac{dr(\phi, t)}{dt} + r(\theta, t) &= f\{ W_0 r_0(t) + W_1 |r_1(t)| \cos(\phi - \psi(t)) + I_0(t)(1 + \epsilon(t))\\
&\quad + I_0(t)\epsilon(t) \cos(\phi - \phi_0)) - \theta\} \tag{7.16}
\end{aligned}
$$

7.2 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

$$r(\phi) = f\{W_0 r_0 + W_1 |r_1| \cos(\phi - \psi)) + I_0(1 - \epsilon) + I_0\epsilon \cos(2(\phi - \phi_0)) - \theta\}. \tag{7.17}$$

3

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 $\psi = \phi_0$ and the steady state rate equation becomes

$$r(\phi) = f\{[W_0 r_0 + I_0(1 + \epsilon) - \theta] + [W_1|r_1| + I_0 \epsilon] \cos(\phi - \phi_0)\} \tag{7.18}$$

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

7.2.1 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

$$r(\theta) = [W_0 r_0 + I_0(1 - \epsilon) - T] + [W_1|r_1| + I_0\epsilon] \cos(2(\phi - \phi_0)). \tag{7.19}$$

The functional dependence of $r(\psi)$ 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, i.e.,

$$\tilde{r}(\phi) = r_0 + r_1 e^{i\phi} + r_{-1} e^{-i\phi} \tag{7.20}$$

where

$$r_0 \equiv \frac{1}{2\pi} \int_{-\pi}^{\pi} d\phi' r(\phi'), \tag{7.21}$$

$$\begin{aligned} r_1 &\equiv |r_1| e^{-i\psi(t)} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} d\phi' r(\phi', t)\, e^{-i\phi'}. \end{aligned} \tag{7.22}$$

and

$$\begin{aligned} r_{-1} &\equiv r_1^* = |r_1| e^{i\psi(t)} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} d\phi' r(\phi', t)\, e^{i\phi'}. \end{aligned} \tag{7.23}$$

Then

$$\tilde{r}(\phi) = r_0 + 2|r_1| \cos(\phi - \psi) \tag{7.24}$$

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

$$r_0 = W_0 r_0 + I_0(1 + \epsilon) - \theta \tag{7.25}$$

or

$$r_0 = \frac{I_0(1 + \epsilon) - \theta}{1 - W_0} \tag{7.26}$$

and

$$r_1 = \frac{W_1 r_1 + I_0\epsilon}{2}. \tag{7.27}$$

4

or

$$r_1 = \frac{I_0 \epsilon}{2 - W_1}. \tag{7.28}$$

We see that, even for the linear case, there is the potential for gain in the modulation term when $W_1 \to 2$.

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

$$\textbf{Selectivity of input} \quad \equiv \quad \frac{\hat{I}_1}{\hat{I}_0} \tag{7.29}$$
$$= \quad \frac{\epsilon}{1 + \epsilon}$$

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

$$\textbf{Selectivity of output} \quad \equiv \quad \frac{|r_1|}{r_0}$$
$$= \quad \frac{I_0 \, \epsilon}{2 - W_1} \frac{1 - W_0}{I_0 (1 + \epsilon)} \tag{7.30}$$
$$= \quad \frac{1 - W_0}{2 - W_1} \times \textbf{Selectivity of input}.$$

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

7.2.2   Marginal (spontaneous bump) state

In the linear case, the input determines the output. Thus the choice $\epsilon = 0$ will lead to $r_1 = 0$. Now suppose we increase the interaction term W_1 so that $W_1 > 2$. 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 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, ψ.

$$r_1 = \frac{1}{2\pi} \int_{\psi - \pi}^{\psi + \pi} d\phi' \, [W_0 r_0 + I_0 - \theta + W_1 |r_1| \cos(\phi' - \psi)]_+ \, e^{-i\phi'}. \tag{7.31}$$

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

$$\begin{aligned} r_1 &= \frac{1}{2\pi} \int_{\psi - \phi_C}^{\psi + \phi_C} d\phi' ([W_0 r_0 + I_0 - \theta + W_1 |r_1| \cos(\phi' - \psi)] \\ &\quad - [W_0 r_0 + I_0 - \theta + W_1 |r_1| \, \cos(\psi \pm \phi_C - \psi)]) e^{-i\phi'} \\ &= W_1 \, |r_1| \frac{1}{2\pi} \int_{\psi - \phi_C}^{\psi + \phi_C} d\phi' \, (\cos(\phi' - \psi) - \cos \phi_C) \, e^{-i\phi'}. \\ &= W_1 \, |r_1| e^{-i\psi} \frac{1}{2\pi} \int_{-\phi_C}^{\phi_C} dx \, (\cos x \, - \, \cos \phi_C) \, e^{-ix} \end{aligned} \tag{7.32}$$

5

$$. \quad = \quad W_1 \, r_1 \, \frac{1}{2\pi} \int_{-\phi_C}^{\phi_C} dx \left(\frac{1}{2} + \frac{e^{-i2x}}{2} - cos \, \phi_C \, e^{-ix} \right)$$

$$= \quad W_1 \, r_1 \, \frac{1}{2\pi} \int_{-\phi_C}^{\phi_C} dx \left(\frac{1}{2} + \frac{e^{-i2x}}{2} - cos \, \phi_C \, e^{-ix} \right)$$

$$= \quad W_1 \, r_1 \, \frac{1}{2\pi} \left(\phi_C - \frac{1}{2} sin \, 2\phi_C \right)$$

so that

$$W_1 = \frac{4\pi}{2\phi_C - sin \, 2\phi_C}. \tag{7.33}$$

Our result relates the synaptic strength to the pattern 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 $W_1 = 2$. 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.

7.2.3 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, i.e., on the selectivity of the input. Noting that a similar analysis to the one above yields

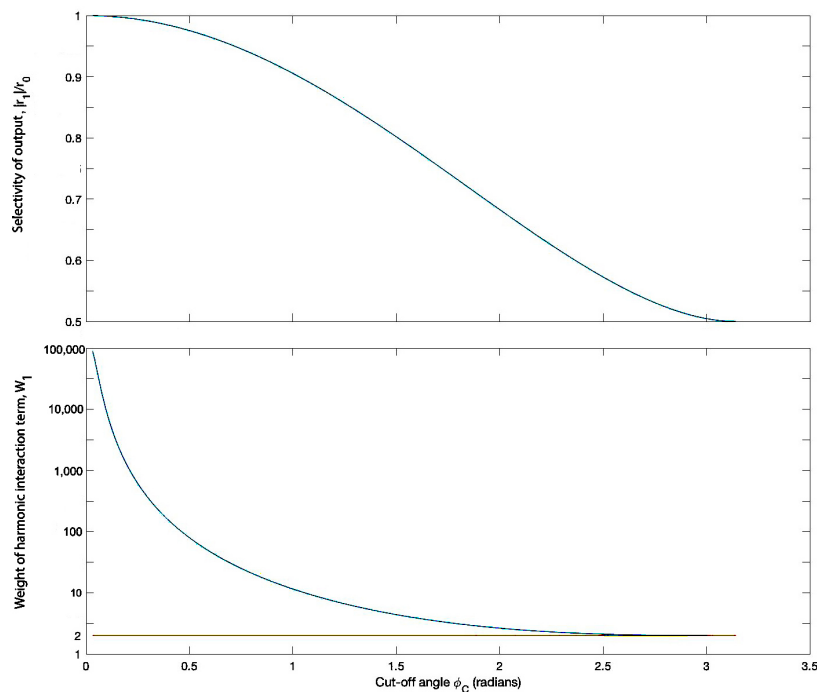$$r_0 = |r_1| \frac{W_1}{\pi} \left(sin\phi_C - \phi_C \, cos\phi_C \right), \tag{7.34}$$

we have

$$\textbf{Selectivity of output} \quad \equiv \quad \frac{|r_1|}{r_0} \tag{7.35}$$

$$= \quad \frac{1}{4} \frac{2\phi_C - sin2\phi_C}{sin\phi_C - \phi_C \, cos\phi_C},$$

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.

7.3 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 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.

Moving 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. 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 ϕ_0 - of a landmark.