10 Receptive Fields and Reliability

10.1 General description of receptive fields

We consider a phenomenological description of the stimulus that causes a neuron to fire. Our description will be general, although as a matter of practice it is simplest to think in terms of visual objects, *i.e.*, a pattern of illumination that evolves over time and space. The receptive field forms a kernel, such that the spike rate of the cell is the temporal convolution of the stimulus with the receptive field and the spatial overlap of the stimulus with the receptive field.

We define the inhomogeneous spike rate as r(t). This is the rate that goes into, per se, a Poisson rate expression where the probability of no spikes in the interval [0,t] and one spike in the interval;(t,t+dt] is $P(t) = r(t) \cdot exp\left(-\int_{0}^{t} dt'r(t')\right)$. Then

$$r(t) = f \left[I_o + \int_{-\infty}^{\infty} d^2 \vec{x} \int_{-\infty}^{t} dt' S(\vec{x}, t') R(\vec{x}, t - t') \right]$$
(10.10)

where $f[\cdot]$ is the nonlinear input-output relation, $S(\vec{r}, t)$ is the stimulus, $R(\vec{x}, t)$ is the receptive field with \vec{x} the two-dimensional spatial vector, and I_o is the baseline input.

When the stimuli driven part of the input is small compared to I_o , we can expand $g[\cdot]$ in a Taylor series and write

$$r(t) \simeq r_o + f' \int_{-\infty}^{\infty} d^2 \vec{x} \int_{-\infty}^{t} dt' S(\vec{x}, t') R(\vec{x}, t - t')$$
(10.11)

where $r_o = f[I_o]$ and

$$f' = \frac{df}{dI} \mid_{I=I_o} \tag{10.12}$$

so that the firing rate is a linear function of the stimulus. This allows us to focus on the receptive field without worrying about the nonlinearity $f[\cdot]$. The review by Chichilnisky (2001) shows a nice example in Figure 4 for the case where the response is separable, *i.e.*, where it can be written as the product of a spatial pattern times a temporal waveform, *e.g.*, $R(\vec{r},t) = u(\vec{r})v(t)$; this review also addresses the assignment of both $R(\vec{x},t)$ and $f[\cdot]$ when the stimulus driven part of the input is not small compared to I_o .

To gain some insight into the general response properties of neurons, we recall that a matrix can always be expanded in terms of it's eigenvectors by a singular valued decomposition. In terms of the notion for the receptive field, we have

$$R(\vec{x},t) \equiv \sum_{n=1}^{rank(R)} \lambda_n u_n(\vec{x}) v_n(t)$$
(10.13)

where the functions $u_n(\vec{x})$ form an orthonormal basis set in space and $v_n(t)$ for an orthonormal basis set in time. The eigenvalues for these basis sets are given by λ_n^2 and. of course, are ordered so that $\lambda_1 > \lambda_2 > \lambda_3 \cdots$ In this case the receptive field is not separable, as first discussed by the work of McClean and Palmer (1989) and analyzed in some detail in the work of Golomb, Kleinfeld, Reid, Shapley and Shraiman (1994).

FIGURE - SVD reduction of RGC data

Then

$$r(t) \simeq r_o + f' \sum_{n=1}^{rank(R)} \lambda_n \int_{-\infty}^{\infty} d^2 \vec{x} \, u_n(\vec{x}) \int_{-\infty}^{t} dt' \, S(\vec{x}, t') \, v_n(t - t').$$
(10.14)

Now suppose that the stimulus is separable, as is often the case in primary sensory areas. For example, in vision our eyes shift from position to position about five times a second. In this case we may write

$$S(\vec{x},t) \equiv X(\vec{x})T(t). \tag{10.15}$$

The spatial part of the stimulus that each mode "sees" is given by the overlap integral of the spatial pattern of the stimulus with the spatial pattern of each mode, *i.e.*,

$$U_n = \int_{-\infty}^{\infty} d^2 \vec{x} \ X(\vec{x}) \ u_n(\vec{x}).$$
(10.16)

where the U_n are scalars. The time dependence of the stimulus is convoluted with each of the associated temporal modes to form the temporal evolution for that mode, *i.e.*,

$$V_n(t) = \int_{-\infty}^t dt' \ T(t') \ v_n(t-t').$$
(10.17)

where the $V_n(t)$ are functions. We thus find

$$r(t) = r_o + f' \sum_{n=1}^{rank(R)} \lambda_n U_n V_n(t).$$
 (10.18)

so that each temporal waveform is weighted by the expansion coefficient for the receptive field and the spatial overlap of the mode with the stimulus. The point is that the temporal response of the neuron, given by r(t), depends on the spatial pattern of the input as well as the temporal evolution of the stimulus. This is what one calls, or some call, a "temporal code", *i.e.*, the coding of different stimuli, even

quasi-static stimuli, by different temporal patterns of spike rates. The inhomogeneous rate r(t) may evolve in time as fast as the response of the sensory cells, such at retinal ganglion cells for the case of vision.

A final point is that the summation over modes typically contains only a few terms, not the full rank of the matrix R. The spatial coefficient U_n has a signal-tonoise ratio that varies in proportion to λ_n for the n - th mode. Thus the above series is cut off after two or three terms as the signal dives below the noise. The SVD expansion can be used as a data compression scheme in the description of the receptive field. For magnocellular cells,

$$r(t) \approx r_o + [f'\lambda_1 U_1] V_1(t) + [f'\lambda_2 U_2] V_2(t).$$
(10.19)

FIGURE - Nonseparability of RGC data

10.2 Digression on singular values decomposition

In the expansion

$$R(\vec{x},t) \equiv \sum_{n=1}^{rank(R)} \lambda_n u_n(\vec{x}) v_n(t)$$
(10.20)

the functions satisfy the orthonormality constraints

$$\int_{-\infty}^{\infty} d^2 \vec{x} \, u_n(\vec{x}) u_m(\vec{x}) = \delta_{nm} \tag{10.21}$$

and

$$\int_{-\infty}^{\infty} dt' v_n(t')v_m(t') = \delta_{nm}.$$
(10.22)

We now consider the contraction of the receptive field matrices to form a symmetric correlation matrix, i.e.,

$$C(t,t') \equiv \int_{-\infty}^{\infty} d^{2}\vec{x}R(\vec{x},t)R(\vec{x},t')$$
(10.23)
$$= \sum_{n=1}^{rank(R)} \sum_{m=1}^{rank(R)} \lambda_{n}\lambda_{m} \int_{-\infty}^{\infty} d^{2}\vec{x} \, u_{n}(\vec{x})u_{m}(\vec{x}) \, v_{n}(t)v_{m}(t')$$
$$= \sum_{n=1}^{rank(R)} \sum_{m=1}^{rank(R)} \lambda_{n}\lambda_{m} \, \delta_{nm} \, v_{n}(t)v_{m}(t')$$
$$= \sum_{n=1}^{rank(R)} \lambda_{n}^{2} \, v_{n}(t)v_{n}(t').$$

Then $v_n(t)$ solves the eigenvalue equation

$$\int_{-\infty}^{\infty} dt' C(t,t') v_n(t') = \sum_{m=1}^{rank(R)} \lambda_m^2 v_m(t) \int_{-\infty}^{\infty} dt' v_n(t') v_m(t') \qquad (10.24)$$
$$= \lambda_n^2 v_n(t)$$

and the $u_n(\vec{x})$ are found from

$$\int_{-\infty}^{\infty} dt' R(\vec{x}, t') v_n(t') = \sum_{m=1}^{rank(R)} u_m(\vec{x}) \int_{-\infty}^{\infty} dt' v_m(t') v_n(t')$$
(10.25)
= $u_n(\vec{x}).$

10.3 Spikes and Reliability

We now switch gears and consider how well animals do when their neurons communicate and compute with action potentials. In particular, we examine how well neurons in various mammals faithfully reproduce aspects of their sensory environment. We are clearly jumping over a large philosophical gap by ignoring the issue of whether the nervous system actually wants all the information in an environment, since one thing animals due best is to make binary behavioral decisions based on a complex sensory stream that reports features in the environment.

10.3.1 Coding of Angular Velocity by H1 in the Fly

The first system is the encoding of angular position in the fly. We consider the response of cell H1 in the lobular plate. These are neurons with very large receptive fields that are excited by back-to-front motions across the visual field and inhibited by front-to-back motions, *i.e.*, yaw sensitive. They have essentially no response to vertical motion, *i.e.*, pitch insensitive. There are two H1 cells, one for each side of the head, and these are arranged in antagonism so that rigid rotation of the fly in a static background excites one cell and inhibits the other.

FIGURE - Fly-1.eps

FIGURE - Fy-2.eps

These cells function to convey information about the rotation of the entire fly during flight. They are three synapses removed from the photoreceptors, *i.e.*, optic ganglion \rightarrow lamina \rightarrow medulla \rightarrow lobula. There is a strong resemblance of this cell and its input structure to type 1 cells in the colliculus of vertebrates; thus even the mammal-centric among us can love them. While we focus only on the sensory part, the output from H1 is ultimately combined with sensory input from wind hairs in the segmental ganglia in the cord of the fly. The error signal controls the amplitude of the wing beats and thus the steering behavior of the animal.

FIGURE - Fly-h1.eps

We consider the experiments of Rob deRuyter, who rotated the visual world of a fly that was tethered, *i.e.*, pasted to a stick, so that fly became a function! Rob measured the spike output from H1 for many epochs of the stimulus. Thus he had two pieces of information,

- $T_{app}^k(t)$ is the applied visual stimulus for the k-th trial. This is just the same T(t) that defined the temporal part of the space-time stimulus above.
- $S_{meas}^k(t) = \sum_{s}^{\text{spike times}} \delta(t t_S^k)$ is the measured spike time for the k-th trial.

The key is to ask if we can use this measured information to construct a filter that allows us to predict the stimulus for an unknown trial from the measured spike train. In a sense, we ask, "How well can we reconstruct the stimulus from the spikes". We consider this via a linear transfer function that minimizes the leastsquare-error - a.k.a. optimal linear filter or linear kernel, an idea that is at least a century old, although it came into use only at the time of WW II when there was a big push at the MIT Radar Laboratory to formulate the mathematics of optimal filtering and prediction. This is described in "Threshold Signals" by Lawson and Uhlenbeck. The procedure is as follows:

- We define W(t) as the sought after transfer function.
- We define $T_{pred}^k(t)$ as the predicted stimulus for the k-th trial, based on the measured spike train, where

$$T_{pred}^{k}(t) = \int_{-\infty}^{t} dt' W(t - t') S_{meas}^{k}(t')$$
(10.26)
$$= \int_{-\infty}^{t} dt' W(t - t') \sum_{s}^{\text{spike times}} \delta(t' - t_{S}^{k})$$

$$= \sum_{s}^{spike times} \int_{-\infty}^{t} dt' W(t - t') \delta(t' - t_{S}^{k})$$

$$= \sum_{s}^{spike times} W(t - t_{S}^{k})$$

is the predicted output.

FIGURE -convolution.eps

To get W(t), we minimize the difference between the actual and the predicted stimulus, averaged over all trials and time, *i.e.*,

$$Error = \sum_{k}^{\text{all trials}} \int_{-\infty}^{\infty} dt \left[T_{pred}^{k}(t) - T_{app}^{k}(t) \right]^{2}$$

$$= \sum_{k}^{\text{all trials}} \int_{-\infty}^{\infty} dt \left[\int_{-\infty}^{t} dt' W(t-t') S_{meas}^{k}(t') - T_{app}^{k}(t) \right]^{2}$$
(10.27)

The error is computed in terms of measured quantities, except for W(t), which we find by the criteria that we want to choose W(t) to minimize the error. This is much easier to solve in the frequency domain, where convolutions turn into products. As a mathematical aside, we consider the Fourier transformed variables:

$$T^{k}_{app}(t) \iff \tilde{T}^{k}_{app}(f)$$
(10.28)

$$S_{meas}^{k}(t) \iff \tilde{S}_{meas}^{k}(f) \tag{10.29}$$

$$T_{pred}^{h}(t) \iff T_{pred}^{h}(f) \tag{10.30}$$

$$W(t) \iff W(f) \tag{10.31}$$

(10.32)

where

$$\tilde{W}(f) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt e^{i2\pi f t} W(t)$$
(10.33)

$$W(t) = \sqrt{2\pi} \int_{-\infty}^{\infty} df e^{-i2\pi f t} \tilde{W}(f) \qquad (10.34)$$

so that (ignoring causality for the moment) the convolution becomes

$$\int_{-\infty}^{\infty} dt' W(t - t') S(t') = \tilde{W}(f) \tilde{S}(f)$$
 (10.35)

and the other relation we need is Parseval's theorem, effectively a conservation of energy, *i.e.*,

$$\int_{-\infty}^{\infty} dt |W(t)|^2 = \int_{-\infty}^{\infty} df |\tilde{W}(f)|^2$$
(10.36)

where $|\tilde{W}(f)|^2 = \tilde{W}(f)\tilde{W}^*(f)$.

We put the above together to write:

$$Error = \sum_{k}^{\text{all trials}} \int df \left| \tilde{T}_{pred}^{k}(f) - \tilde{T}_{app}^{k}(f) \right|^{2}$$

$$= \int df \sum_{k}^{\text{all trials}} \left| \tilde{T}_{pred}^{k}(f) - \tilde{T}_{app}^{k}(f) \right|^{2}$$

$$= \int df \sum_{k}^{\text{all trials}} \left| \tilde{W}(f) \tilde{S}_{meas}^{k}(f) - \tilde{T}_{app}^{k}(f) \right|^{2}$$
(10.37)

$$= \int df \sum_{k}^{\text{all trials}} \left[\tilde{W}(f) \tilde{S}_{meas}^{k}(f) - \tilde{T}_{app}^{k}(f) \right] \left[\tilde{W}^{*}(f) \tilde{S}_{meas}^{k*}(f) - \tilde{T}_{app}^{k*}(f) \right]$$

$$= \int df \sum_{k}^{\text{all trials}} \left[\tilde{W}(f) \tilde{W}^{*}(f) |\tilde{S}_{meas}^{k}(f)|^{2} - \tilde{W}(f) \tilde{S}_{meas}^{k}(f) \tilde{T}_{app}^{k*}(f) - \tilde{T}_{app}^{k}(f) \tilde{W}^{*}(f) \tilde{S}_{meas}^{k*}(f) + |\tilde{T}_{app}^{k}(f)|^{2} \right]$$

$$= \int df \tilde{W}(f) \tilde{W}^{*}(f) \sum_{k}^{\text{all trials}} |\tilde{S}_{meas}^{k}(f)|^{2} - \int df \tilde{W}(f) \sum_{k}^{\text{all trials}} \tilde{S}_{meas}^{k} \tilde{T}_{app}^{k*}(f) - \int df \tilde{W}^{*}(f) \sum_{k}^{\text{all trials}} \tilde{T}_{app}^{k} \tilde{S}_{meas}^{k*}(f) + \int df \sum_{k}^{\text{all trials}} |\tilde{T}_{app}^{k}(f)|^{2}.$$

The next step is to minimize the error with respect to the transfer function. We compute the functional derivative

$$\frac{\delta(Error)}{\delta\tilde{W}^*(f)} = 0 \tag{10.38}$$

so that

$$0 = \int df \tilde{W}(f) \sum_{k}^{\text{all trials}} |\tilde{S}_{meas}^{k}(f)|^{2} - \int df \sum_{k}^{\text{all trials}} \tilde{T}_{app}^{k} \tilde{S}_{meas}^{k*}(f) \qquad (10.39)$$
$$= \int df \left[\tilde{W}(f) \sum_{k}^{\text{all trials}} |\tilde{S}_{meas}^{k}(f)|^{2} - \sum_{k}^{\text{all trials}} \tilde{T}_{app}^{k} \tilde{S}_{meas}^{k*}(f) \right].$$

The expression for W(f) must be valid at each frequency. Thus the frequency representation of the transfer function is

$$\tilde{W}(f) = \frac{\sum_{k}^{\text{all trials}} \tilde{T}_{app}^{k}(f) \tilde{S}_{meas}^{k*}(f)}{\sum_{k}^{\text{all trials}} |\tilde{S}_{meas}^{k}(f)|^{2}}.$$
(10.40)

This is the central result.

For the case of measured signal that is a spike train,

$$\tilde{W}(f) = \frac{\sum_{k}^{\text{all trials}} \tilde{T}_{app}^{k}(f) \sum_{s}^{\text{spike times}} e^{i2\pi f t_{S}^{k}}}{\sum_{k}^{\text{all trials}} \sum_{s,s'}^{\text{spike times}} e^{i2\pi f (t_{S}^{k} - t_{S'}^{k})}}.$$
(10.41)

In the time domain, this is just

$$W(t) = \int df e^{-i2\pi ft} \frac{\sum_{k}^{\text{all trials}} \tilde{T}_{app}^{k}(f) \sum_{s}^{\text{spike times}} e^{i2\pi f t_{S}^{k}}}{\sum_{k}^{\text{all trials}} \sum_{s,s'}^{\text{spike times}} e^{i2\pi f (t_{S}^{k} - t_{S'}^{k})}}$$
(10.42)

Ugly! But simple to compute numerically. However, we notice that this has a simple form when the spike arrival times may be taken to be a random, *e.g.*, Poisson variable. This occurs if the spike rate is not too high, so that the refractory period

plays no role. In this case the phase term $e^{i2\pi f(t_S^k - t_{S'}^k)}$ in the denominator appears as a random vector with a length of one that averages to zero except for the contributions with s = s', so that the sum over all spike times in all trials gives

$$\sum_{k}^{\text{all trials spike times}} \sum_{s,s'} e^{i2\pi f(t_S^k - t_{S'}^k)} \approx N \tag{10.43}$$

where N is the total number of spikes across all trials. With the denominator equal to a constant, the Fourier transform can be completed on just the numerator, *i.e.*,

$$W(t) \approx \int df e^{-i2\pi ft} \frac{\sum_{k}^{\text{all trials}} \tilde{T}_{app}^{k}(f) \sum_{s}^{\text{spike times}} e^{i2\pi ft_{S}^{k}}}{N}$$
$$\approx \frac{1}{N} \sum_{k}^{\text{all trials spike times}} \sum_{s}^{\int df e^{-i2\pi f(t-t_{S}^{k})} \tilde{T}_{app}^{k}(f)$$
$$\approx \frac{1}{N} \sum_{k}^{\text{all trials spike times}} \sum_{s}^{K} T_{app}^{k}(t-t_{S}^{k}).$$
(10.44)

Thus W(t) is just the spike triggered average of the stimulus waveform. At last, we see that all that happens is that the transfer function reports the waveform of the stimulus that is most likely to cause the neuron to fire.

FIGURE - fly-h1.eps (again!)

Rob van Steveninck and Bill Bialek calculated for transfer function for Rob's data, It looks largely like a derivative in time; this is not unreasonable for a system that is part of looking for deviations in the visual field. The fit of the prediction to the measured data is good. It is better for low frequencies than high frequencies. In fact, we expect that the filter should attenuate at time scales shorter than it's width of ≈ 40 ms, or in other words, at frequencies above the inverse of the width, or 2 Hz. Bialek did estimate. He quantified, as an average over trials in the frequency domain, the error between the spectral response and the observed behavior. Not surprisingly, the error exceeds the signal at high frequencies, above 25 Hz in this case.

FIGURE - fly-h1-error.eps

In general, the fly can initiate visually guided course corrections in 30 ms, so the ≈ 20 ms width of W(t) is consistent with this. It is also interesting to point out that the maximum rate that this cell fires is about 100 - 200 Hz. This gives 2 to 4 spikes per width, which is about the same as the correlation time. More typically, we are down to one spike per correlation time.

10.3.2 Causality in Transfer Functions

The filter here has been shifted away from the origin. so that W(t < 0) = 0. This is a potentially dirty trick. The critical thing is that there is a delay until the information gets out. It is likely that this delay is the time required to initiate a visually guided flight correction. A side remark is that a related technique, called reverse correlation, gives a causal filter directly.

10.3.3 Uncertainty in Transfer Functions

The error between the predicted performance and the measured performance may be quantified in terms of the variance of the transfer function $\tilde{W}(f)$. In particular, only a limited spectral range of stimuli may be used. One then calculates the error band for the transfer function, something one can do by a brute force method, like jackknife statistics, or in terms of asymptotic estimates of the variance.

10.3.4 Coding of Vibrissa Motion in S1 of Rat

The above example shows that a single neuron may have a high degree of reliability of relaying the status of the sensory world. One obvious possibility is that H1 in the fly is a very special cell, especially given that it is a singular neuron. We now consider an alternate example from the mammalian cortex, for which thousands of cells may respond to the same stimulus. In particular, we consider the case of neurons in the primary somatosensory (S1) cortex of rat that respond to motion of the vibrissae during a task in which the animal whisks in air. It was shown that the probability of spiking by these cells in proportional to the position of the vibrissae. Here, in analogy with the case for H1, we ask how well the spike trains from such cells can be used to predict the position of the vibrissae.

FIGUREs - rat-1-3.eps

The first result is that the transfer function appears tuned, *i.e.*, rather than a flat function or a bump, the transfer function appears as a dampened oscillation. The damping, or correlation time, is about 150 ms, slightly longer than the period of whisking. The second result is a single spike train can be used to predict the position of the vibrissae with excellent fidelity. In fact, if one looks only at the frequency for whisking, the magnitude of the coherence between the predicted and the measured signal is greater than 0.9. From a functional point of view, this results shows that some cells in S1 (about 1 out of 20) have high fidelity information about the position of the vibrissae. Further, it speaks against the necessity of a "distributed code" as a means to represent the sensory stream.

FIGURE - rat-4.eps

As in the case of H1, the high degree of reliability can be understood in terms of a sufficiently high spike rate, *i.e.*, about one spike per correlation time, and a sufficiently high correlation between spiking and the EMG for whisking. In fact, an additional measurement, in which rats were trained to use their vibrissae to detect the presence of a edge, demonstrates that a neurons in S1 produce about 1 spike within the time it takes the rat to make a decision. In this case, we plot the integrated number of spikes as the rat probes the edge, and note that she turns to make her decision after close to a single spike has accumulated.

FIGURE - rat-contact.eps

10.3.5 Spike Coding and Behavioral Choice for Random Dot Motion

We now turn to a final example and ask how a sensory spike signal may be related to the behavior, or psychometric judgment, of an animal. The importance of this is that the animal is likely to pool the response of many sensory input neurons if the behavioral response is much better than the response of a given cell. This is a vote in favor of a "distributed code" across primary sensory cells. On the other hand, it is certainly possible that the output of a single neuron further down the sensory stream is an excellent predictor of behavior. This was hinted at in our discussion of the rat experiments, but there was no measure of correct or incorrect behavior in that case.

We consider Newsome's experiments on directional tuning in the visual cortex of monkey, and focus of the response of neurons in area MT. These cells receive input via two pathways, one is the retina \rightarrow thalamus \rightarrow primary visual (V1) cortex $\rightarrow \cdot$ $\cdot \cdot \rightarrow$ medial temporal (MT) cortex pathway. The other is the retina \rightarrow colliculus \rightarrow pulvinar \rightarrow MT pathway.

FIGURES - monkey-1,2.eps

The monkey is trained to fixate on a spot and then is presented with a random dot pattern. The motion of the dots has a large random component and a small biased component in one of two directions. The monkey must decide which direction the bias is in, and at the same time the spiking output from neurons in MT is recorded and used to calculate the preference.

FIGURE - monkey-3.eps

In this experiment, only cells that had a strong directional preference were selected. From about 50 to 100 trials of data, both psychometric and physiometric decisions curves were constructed. For the psychometric case, the curve was simply the probability of a correct choice. For the physiometric case, it was (very) roughly the probability of the number of spikes in a 2 second internal coming from the null or unresponsive direction versus the preferred direction. Similar in spirit to the above cases, the data from all trials was used to construct the distribution of response. Then, a maximum likelihood measure was used to determine if the response on a given trial best fit the distribution for the null direction versus the preferred direction. This was matched against the psychometric response.

FIGURE - monkey-4.eps

The essential result is that the average output from a single neuron could be used to predict the direction of the dots with a reliability that essentially equaled the psychometric choice of the animal.





chapt_2_convolution.eps

American Association for the Advancement of Science



28 JUNE 1991 Vol. 252 **PAGES 1757-1886**

chapt_2lyflig51pieps

\$6.00





chapt_2_fly_2.eps

Linear Modeling of the Fly H1 Response



chapt_2_fly_h1.eps

Sensitivity Analysis of the Fly H1 Response







chapt_2_rat_1.eps



Simultaneous Recording of EMG and Single Units in S1 Cortex



chapt_2_rat_3.eps

Predicting the Position of the Vibrissae (Mystatial EMG) on a Single-Trial Basis









chapt_2_monkey_4.eps