Neural Coding: Population Coding

Angus Chadwick

School of Informatics, University of Edinburgh, UK

Computational Neuroscience (Lecture 10, 2023/2024)

Outline of Lecture

- Encoding and decoding in populations of neurons

- Fisher Information

- Noise correlations

- Relationship between neural coding and behavioural performance

- Impact of learning and attention on population codes

Representations in Populations of Neurons

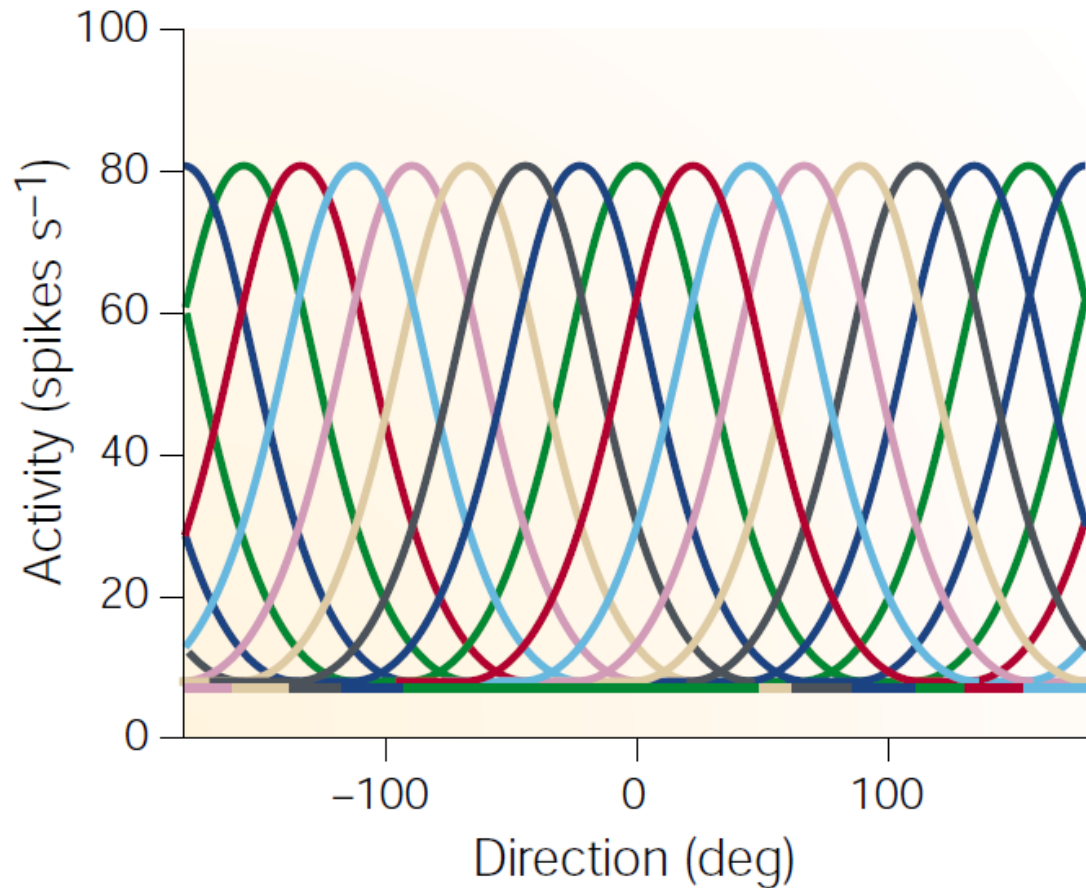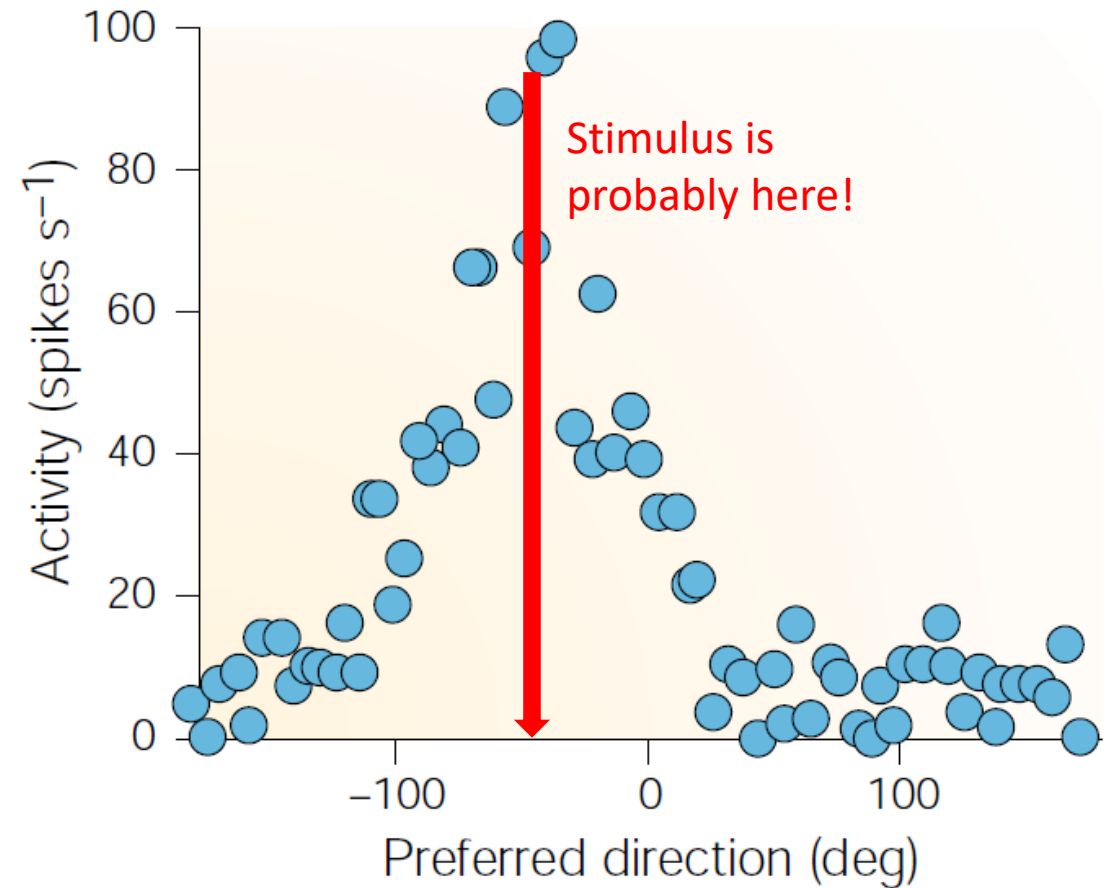Many neurons, each with a different preferred orientation
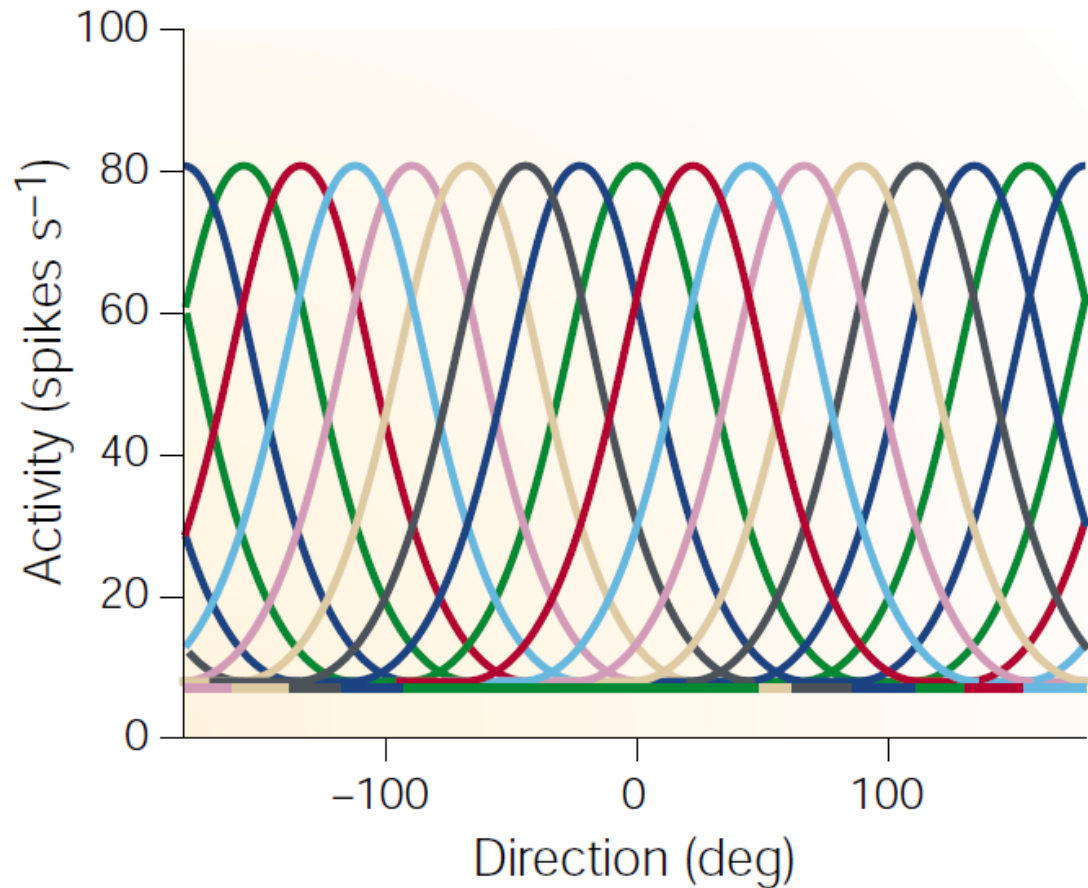
Noisy response of population to a single stimulus presentation

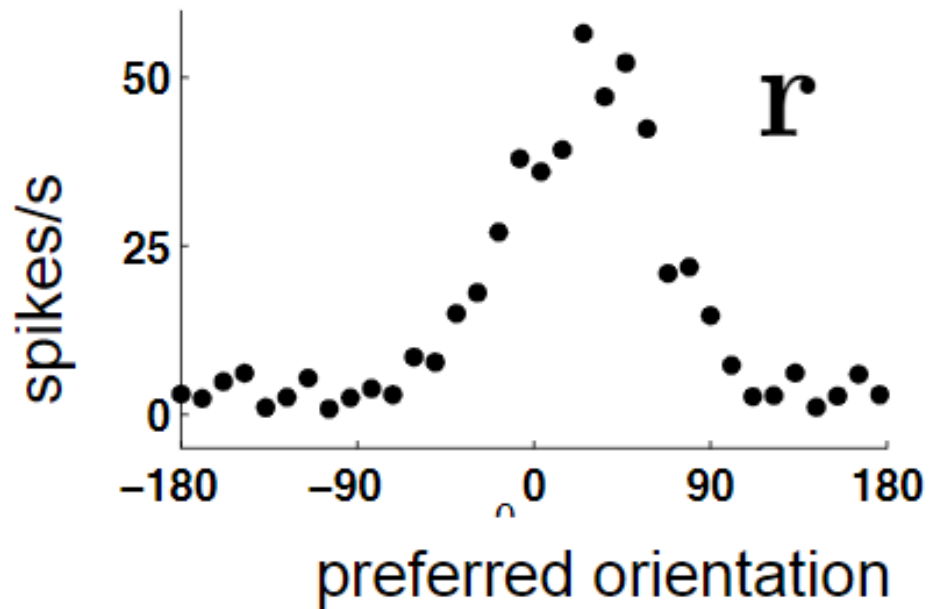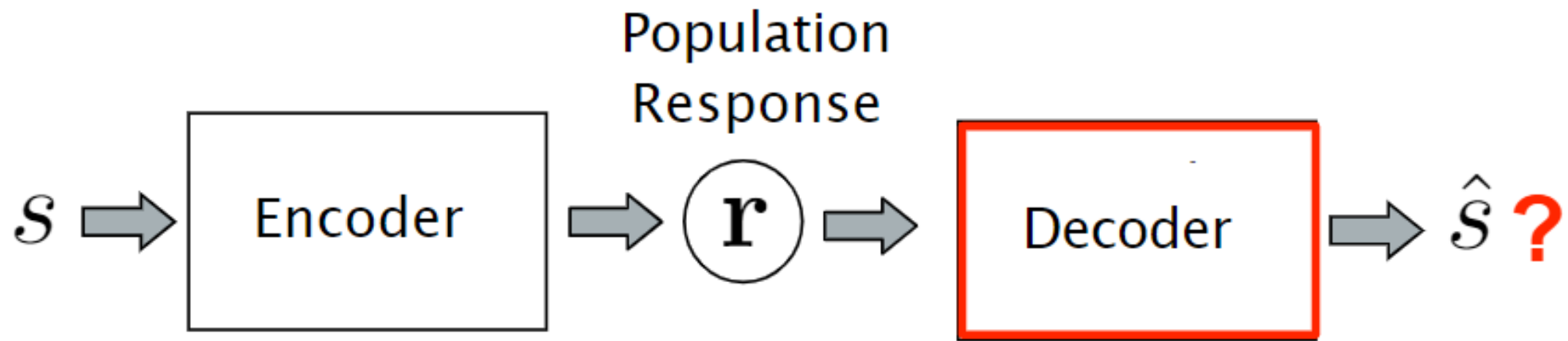# Representations in Populations of Neurons

Many neurons, each with a different preferred orientation

Noisy response of population to a single stimulus presentation

Stimulus is probably here!
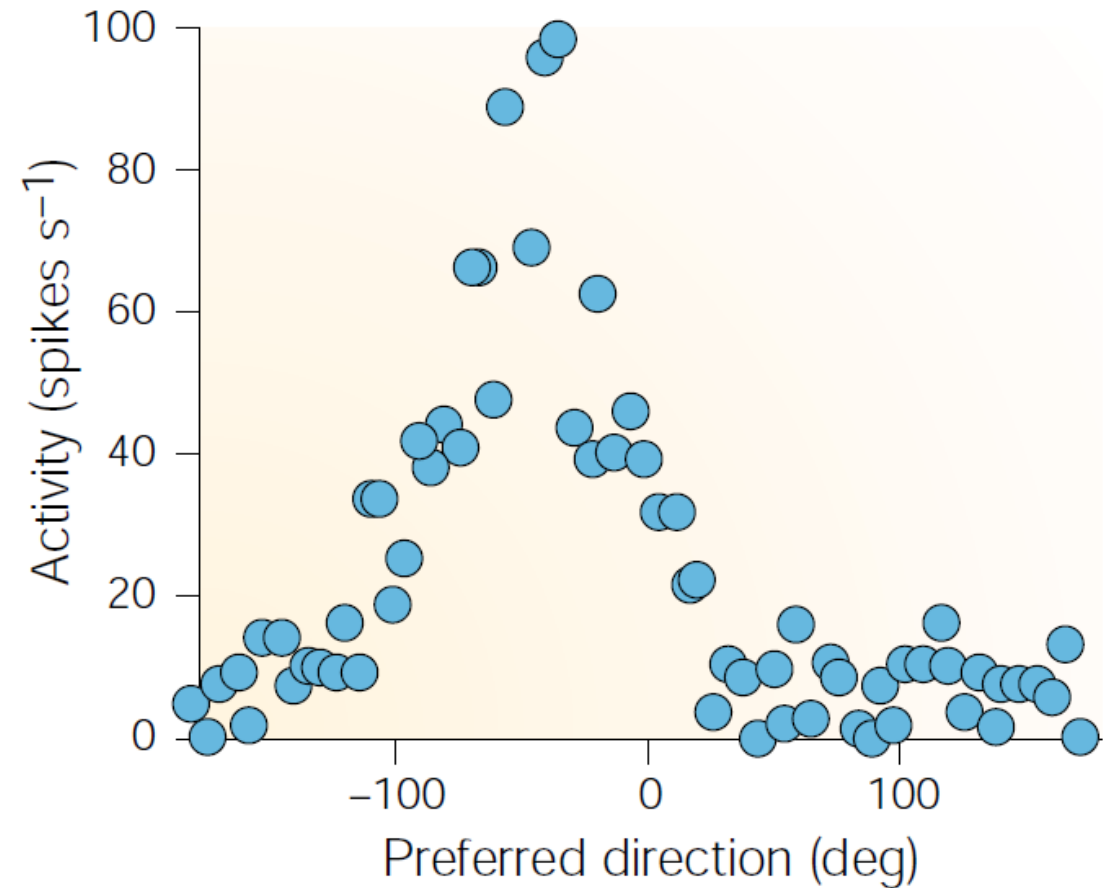
The Decoding Problem

Why Decode?

Decoding has several purposes:

1) **It sheds light on fundamental scientific questions.** How much information is there in a population of neurons? Are perceptual limits dictated by neural coding or something else? Does the brain encode and decode optimally or suboptimally?

2) **It has biomedical applications.** Brain machine interfaces, neural prosthetics, etc.

3) **It's something the brain does all the time.** How does the brain make use of its own activity patterns? For example, in order to use of visual information to make a motor decision, motor brain areas must "decode" the activity of visual brain areas.

Decoding with Bayes Rule

Given a population of *N* neurons with responses $\mathbf{r} = [r_1, r_2, ..., r_N] \sim p(\mathbf{r}|s)$

How can we estimate the stimulus *s* from the response *r*?

Optimal decoder uses Bayes rule: $p(s|\mathbf{r}) = p(\mathbf{r}|s)\dfrac{p(s)}{p(\mathbf{r})}$

Bayes Decoders

Bayes decoders are an "ideal observer" - they give the best possible readout.

We need to know likelihood *p(r|s)* and prior *p(s)* - hard to estimate from data! And perhaps hard for the brain to learn/store.

To make a decision, need to convert to a point estimate, e.g.:
This is typically hard to do!

$$\hat{s}_{ML} = \operatorname{argmax}\left[p(\mathbf{r}|s)\right]$$
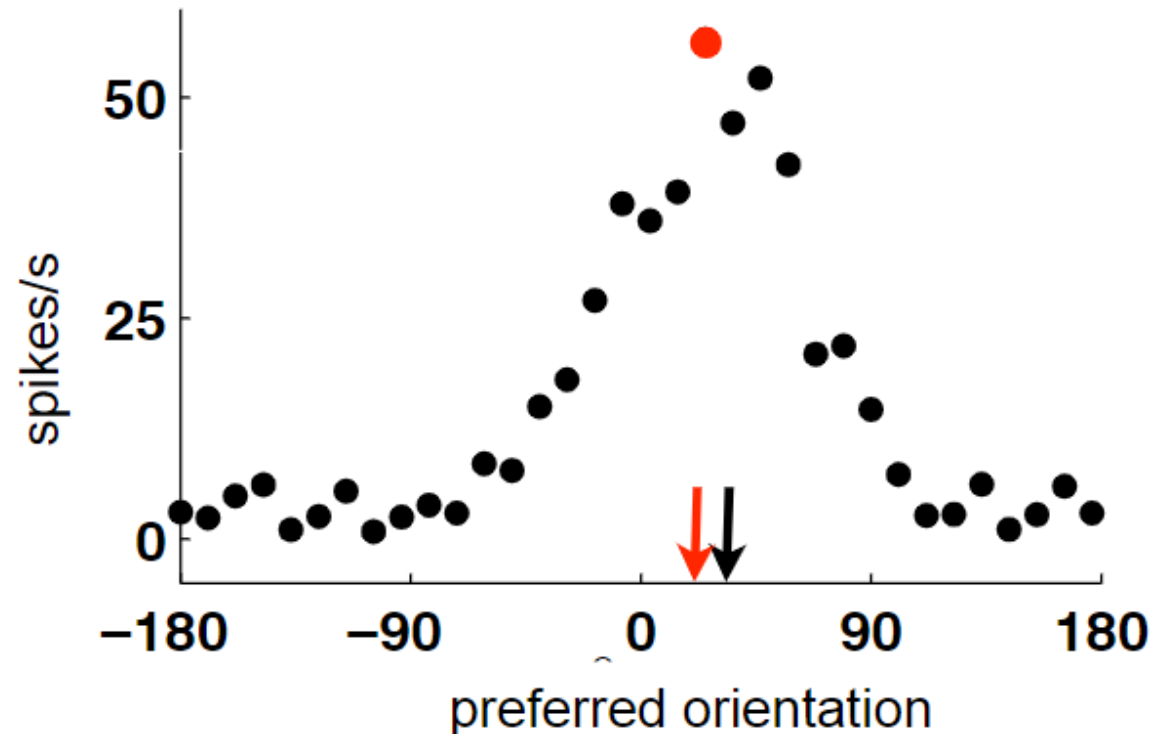
$$\hat{s}_{MAP} = \operatorname{argmax}\left[p(s|\mathbf{r})\right]$$

Alternative: use a simpler, but suboptimal decoder. (Maybe the brain does this too...)

The Winner Takes All Decoder

Decoded stimulus value = stimulus preference of neuron with highest firing rate.
Simple, but highly suboptimal!

The Population Vector Decoder

A kind of weighted average of neurons' stimulus preferences. Still suboptimal, but better!

Population Vector

$$\hat{\mathbf{z}} = \sum_i r_i \left[\cos s_i, \sin s_i\right]$$

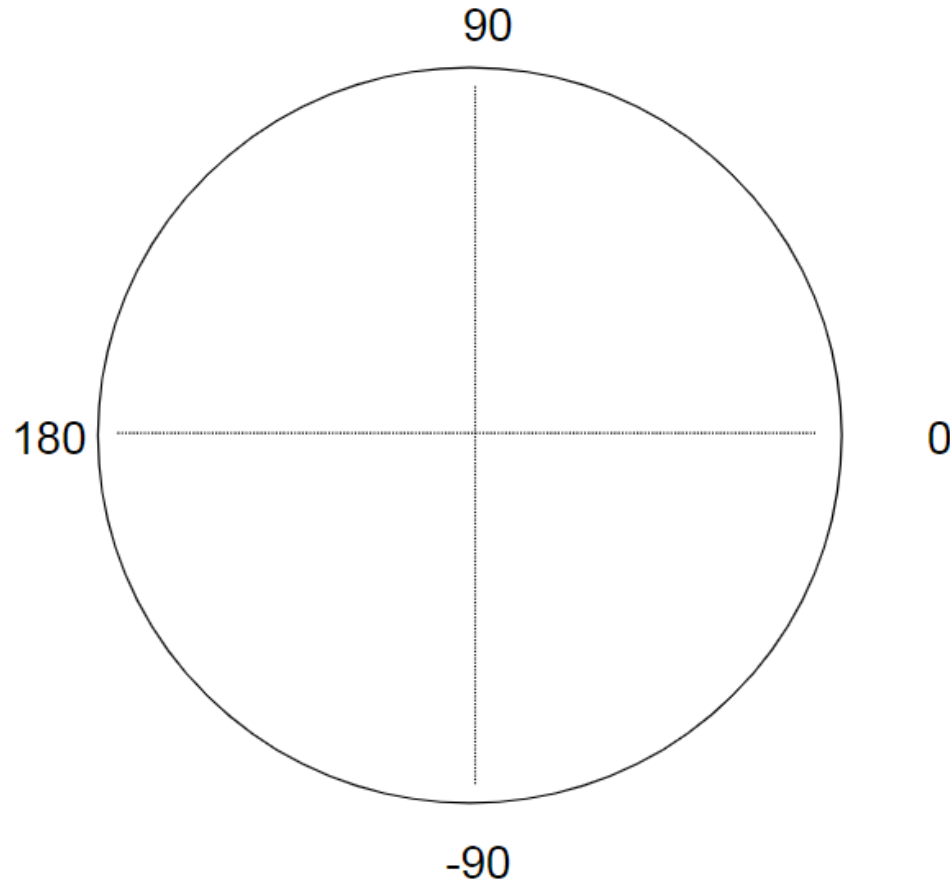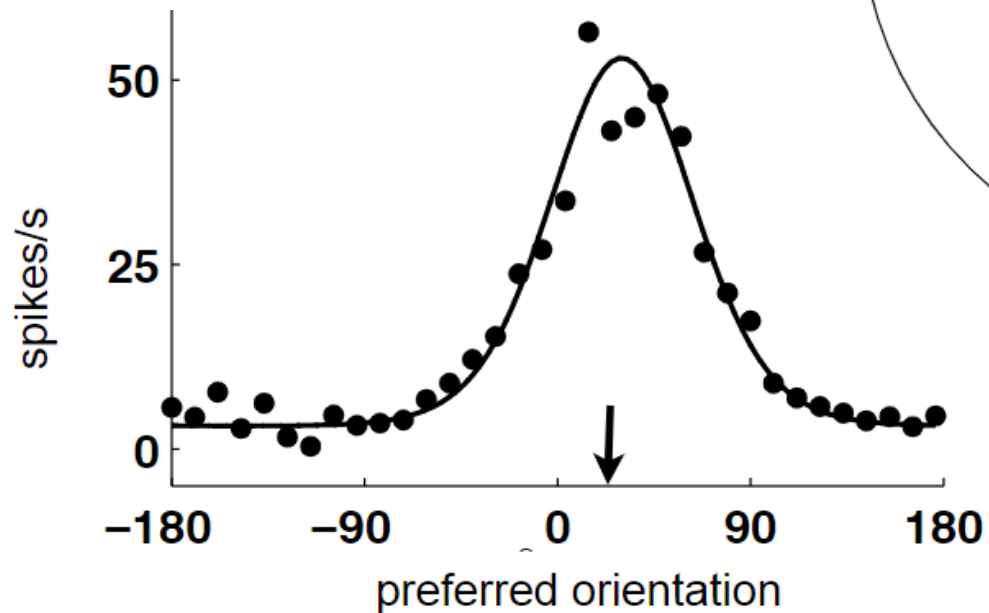$$\hat{s} = \arctan\left(\frac{z_2}{z_1}\right)$$

The Population Vector Decoder

A kind of weighted average of neurons' stimulus preferences. Still suboptimal, but better!

Population Vector

$$\hat{\mathbf{z}} = \sum_i r_i \left[\cos s_i, \sin s_i\right]$$

$$\hat{s} = \arctan\left(\frac{z_2}{z_1}\right)$$
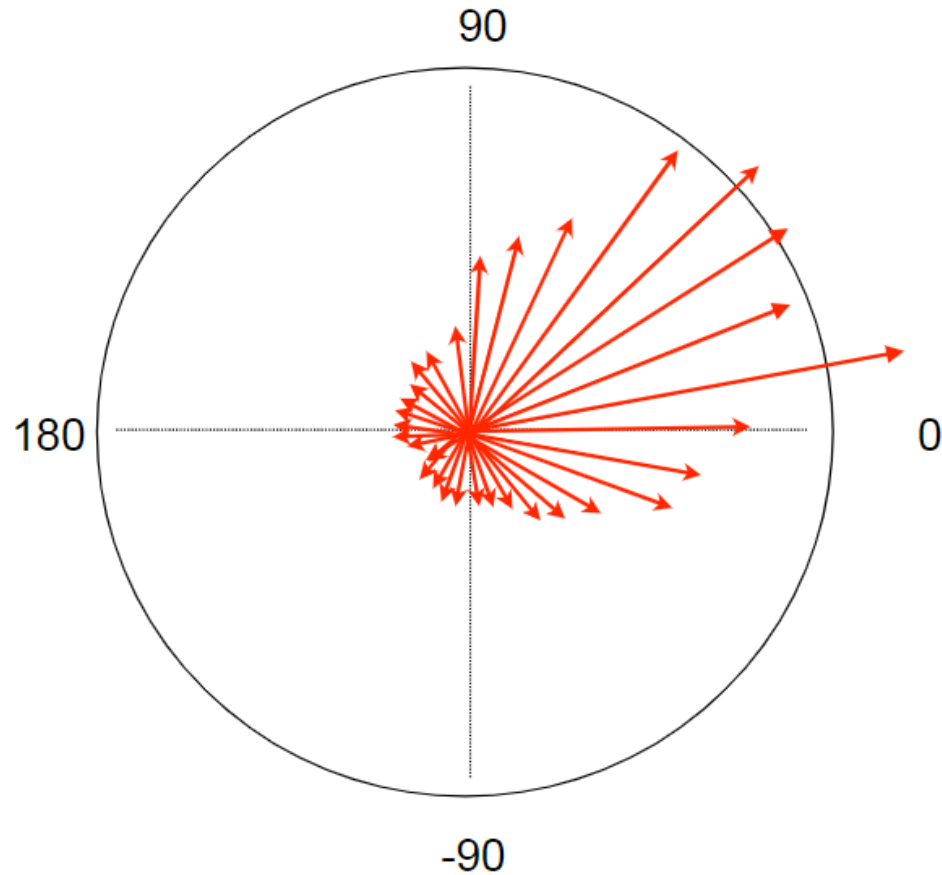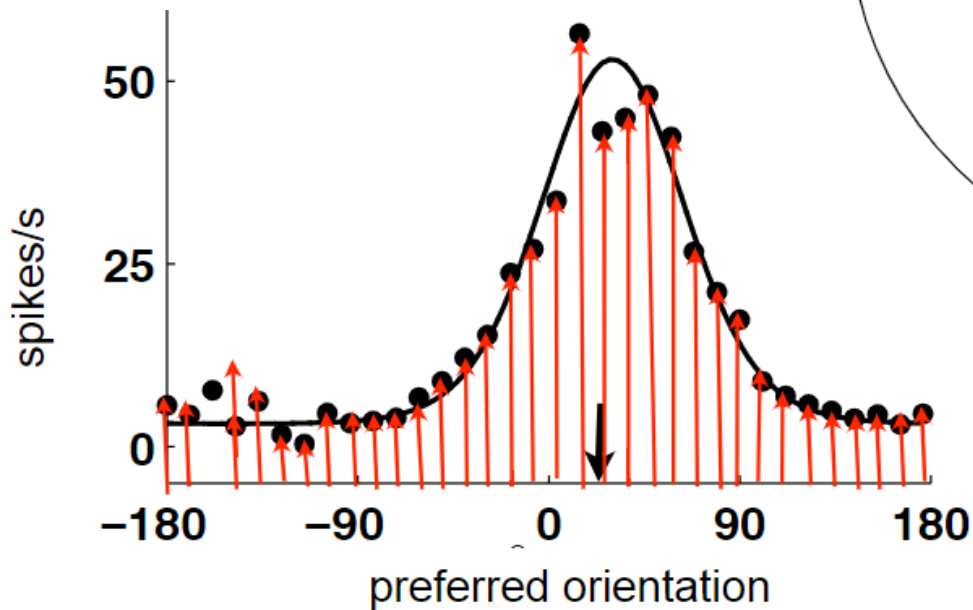
Population Vector vs Maximum Likelihood
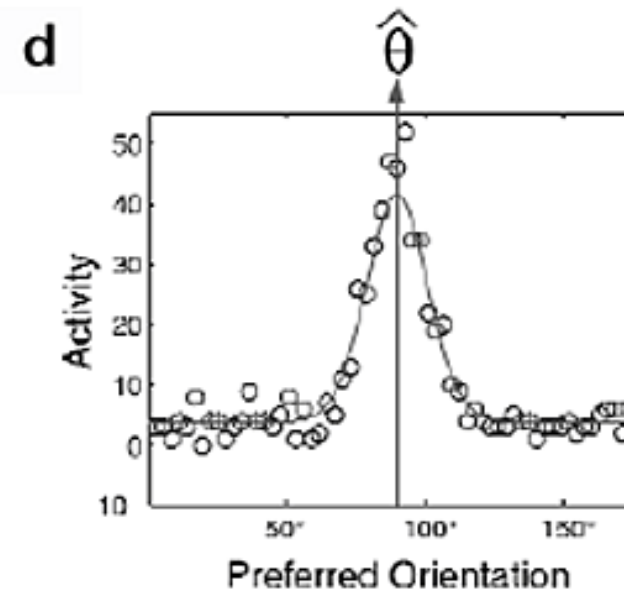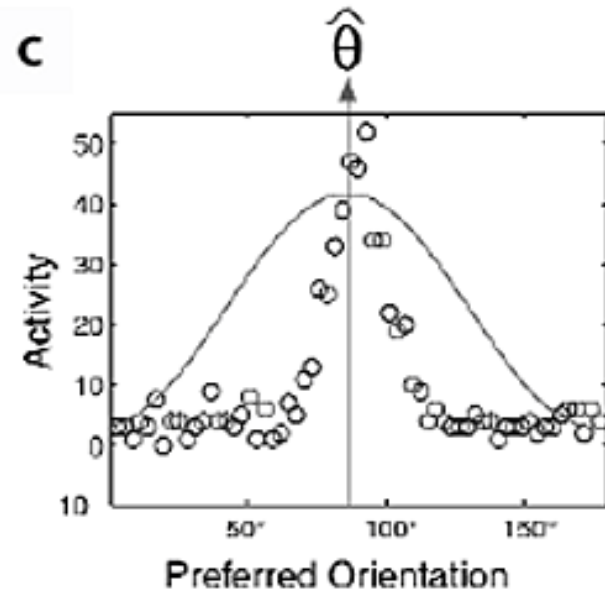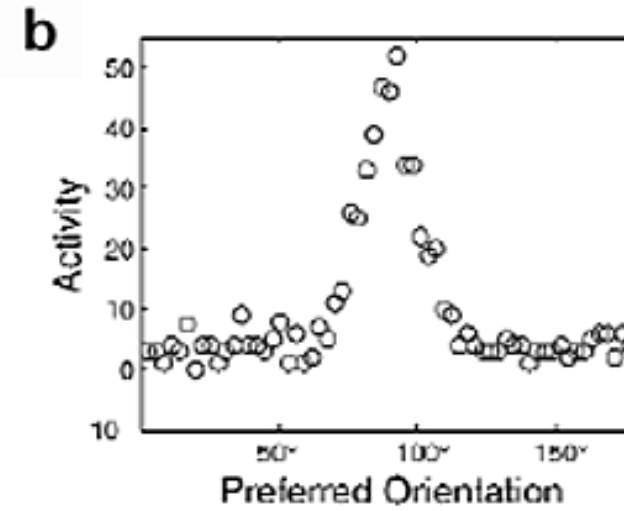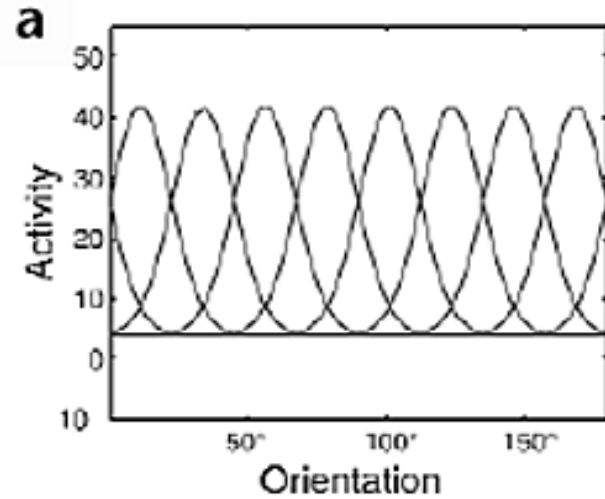
Population vector is equivalent to fitting a cosine to the population response

Maximum likelihood fits a template that matches the true response shape

Gaussian Tuning Curves

$$\langle n_i \rangle = f_i(s) \qquad\qquad f_i(s) = Ae^{-(s-s_i)^2/(2\sigma^2)}$$

Maximum Likelihood Decoding: Independent Poisson Neurons

Gaussian Tuning Curves

$$\langle n_i \rangle = f_i(s) \qquad\qquad f_i(s) = Ae^{-(s-s_i)^2/(2\sigma^2)}$$

Independent Poisson
Spike Counts

$$p(n_i|s) = \frac{(f_i(s))^{n_i} e^{-f_i(s)}}{n_i!} \qquad\qquad p(\mathbf{n}|s) = \prod_{i=1}^{N} p(n_i|s)$$

Maximum Likelihood Decoding: Independent Poisson Neurons

Gaussian Tuning Curves

$$\langle n_i \rangle = f_i(s) \qquad f_i(s) = Ae^{-(s-s_i)^2/(2\sigma^2)}$$

Independent Poisson Spike Counts

$$p(n_i|s) = \frac{(f_i(s))^{n_i} e^{-f_i(s)}}{n_i!} \qquad p(\mathbf{n}|s) = \prod_{i=1}^{N} p(n_i|s)$$

Log likelihood

$$\log p(\mathbf{n}|s) = \sum_{i=1}^{N} \left[n_i \log(f_i(s)) - f_i(s) - \log n_i! \right]$$

Maximum Likelihood Decoding: Independent Poisson Neurons

Gaussian Tuning Curves

$$\langle n_i \rangle = f_i(s) \qquad\qquad f_i(s) = Ae^{-(s-s_i)^2/(2\sigma^2)}$$

Independent Poisson Spike Counts

$$p(n_i|s) = \frac{(f_i(s))^{n_i} e^{-f_i(s)}}{n_i!} \qquad p(\mathbf{n}|s) = \prod_{i=1}^{N} p(n_i|s)$$

Log likelihood

$$\log p(\mathbf{n}|s) = \sum_{i=1}^{N} \left[n_i \log(f_i(s)) - f_i(s) - \log n_i! \right]$$

Assume $\sum_i f_i(s) = \text{const}$

$$\approx -\sum_{i=1}^{N} n_i \frac{(s-s_i)^2}{2\sigma^2} + \text{const}$$

Maximum Likelihood Decoding: Independent Poisson Neurons

Gaussian Tuning Curves

$$\langle n_i \rangle = f_i(s) \qquad\qquad f_i(s) = Ae^{-(s-s_i)^2/(2\sigma^2)}$$

Independent Poisson Spike Counts

$$p(n_i|s) = \frac{(f_i(s))^{n_i} e^{-f_i(s)}}{n_i!} \qquad p(\mathbf{n}|s) = \prod_{i=1}^{N} p(n_i|s)$$

Log likelihood

$$\log p(\mathbf{n}|s) = \sum_{i=1}^{N} \left[n_i \log(f_i(s)) - f_i(s) - \log n_i! \right]$$

Assume $\sum_i f_i(s) = \text{const}$

$$\approx -\sum_{i=1}^{N} n_i \frac{(s-s_i)^2}{2\sigma^2} + \text{const}$$

$$\implies \hat{s}_{ML} = \text{argmax}(\log p(\mathbf{n}|s)) = \frac{\sum_i n_i s_i}{\sum_i n_i}$$

$$\hat{s}_{ML} = \mathrm{argmax}\left[p(\mathbf{r}|s)\right]$$

- Finding the ML solution requires finding the maximum of a complicated function – often requires numerical optimisation methods.

- One can build a recurrent network of neurons that performs such an optimisation via gradient ascent (see figure).

- Whether/how the brain decodes activity encoded in neural populations is not known. The ML (or MAP) decoder gives an upper bound on what the brain could do.
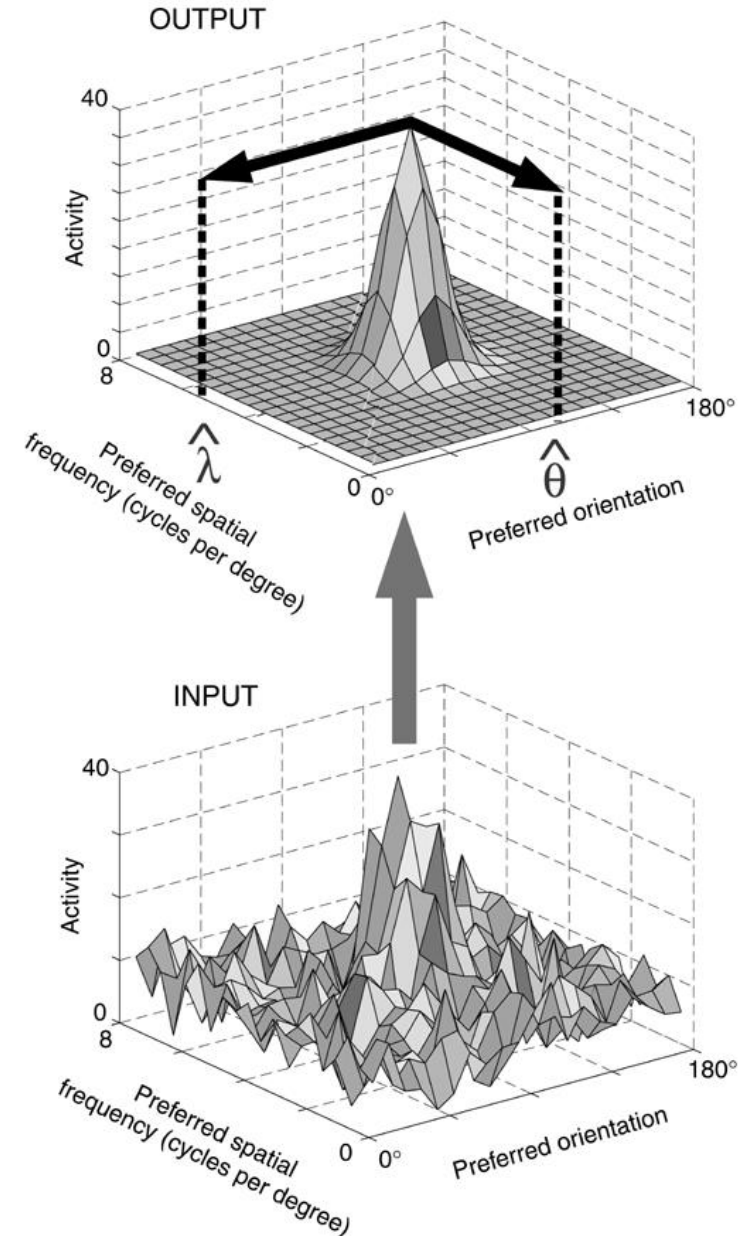
Fisher Information

- How can we quantify the quality of a given population code?

- Imagine you record a population of neurons. You want to decode the stimulus s, and you know their response statistics $p(r|s)$.

- How accurately can you decode s from r? What will be the decoding error? How does it depend on the tuning curves $f(s)$? What about the trial-to-trial variability?

- We use Fisher information to answer questions such as these. Fisher information quantifies the performance of an optimal estimator of s from r.

Fisher Information

- Fisher information is the expected curvature of the likelihood function about the true stimulus *s*:

$$I_F(s) = \left\langle \left(\frac{d \log p(\mathbf{r}|s)}{ds} \right)^2 \right\rangle = - \left\langle \frac{d^2 \log p(\mathbf{r}|s)}{ds^2} \right\rangle$$

- The expectation is taken over the distribution of responses *r*, for a fixed stimulus *s*. Fisher information is therefore a local measure (mutual information is global)

- Intuitively, if the likelihood is very curved about *s*, then a small change in the stimulus *s* will invoke a large change in response *r*, therefore enabling *s* to be well estimated from *r*

- This intuition is formalised by the Cramer-Rao lower bound

The Cramer-Rao Lower Bound

- Definition: an estimator has bias $b(s)$ and variance

$$b(s) = \langle \hat{s} \rangle - s$$

$$var(s) = \langle (\hat{s} - \langle \hat{s} \rangle)^2 \rangle$$

An estimator is unbiased if $b(s) = 0$ for all s.

- Theorem: the variance of any unbiased estimator satisfies $var(s) \geq \dfrac{1}{I_F(s)}$

- Thus, Fisher information quantifies the precision of an optimal estimator.

Decoders and Fisher Information: Summary

- Stimuli can be decoded from the response of a population of neurons in many ways

- Optimal decoders follow Bayes rule. Maximum likelihood is optimal under a uniform prior over stimuli

- Suboptimal decoders are often easier to use, but involve either bias or increased decoding error (or both)

- Fisher information sets a lower bound on the variance of an unbiased estimator

- Thus, we can use Fisher information as a measure of the quality of stimulus encoding in a population of neurons

Fisher Information: Tuning Curve + Noise Model

- Consider a single neuron with tuning curve *r=f(s)* and Poisson spike count *n:*

$$p(n|s) = \mathrm{Poiss}(n|f(s)T) = (f(s)T)^n e^{-f(s)T}/n!$$

Fisher Information: Tuning Curve + Noise Model

- Consider a single neuron with tuning curve *r=f(s)* and Poisson spike count *n:*

$$p(n|s) = \text{Poiss}(n|f(s)T) = (f(s)T)^n e^{-f(s)T}/n!$$

- The Fisher information is:

$$I_F(s) = -\left\langle \frac{d^2 \log p(n|s)}{ds^2} \right\rangle = -\left\langle \frac{d^2}{ds^2}(n\log(f(s)T) - f(s)T - \log n!) \right\rangle$$

Fisher Information: Tuning Curve + Noise Model

- Consider a single neuron with tuning curve *r=f(s)* and Poisson spike count *n:*

$$p(n|s) = \text{Poiss}(n|f(s)T) = (f(s)T)^n e^{-f(s)T}/n!$$

- The Fisher information is:

$$I_F(s) = -\left\langle \frac{d^2 \log p(n|s)}{ds^2} \right\rangle = -\left\langle \frac{d^2}{ds^2}(n \log(f(s)T) - f(s)T - \log n!) \right\rangle$$

$$= -\left\langle \frac{d}{ds}\left(n\frac{f'(s)T}{f(s)T} - f'(s)T \right) \right\rangle$$

Fisher Information: Tuning Curve + Noise Model

- Consider a single neuron with tuning curve *r=f(s)* and Poisson spike count *n:*

$$p(n|s) = \text{Poiss}(n|f(s)T) = (f(s)T)^n e^{-f(s)T}/n!$$

- The Fisher information is:

$$I_F(s) = -\left\langle \frac{d^2 \log p(n|s)}{ds^2} \right\rangle = -\left\langle \frac{d^2}{ds^2}(n\log(f(s)T) - f(s)T - \log n!) \right\rangle$$

$$= -\left\langle \frac{d}{ds}\left(n\frac{f'(s)T}{f(s)T} - f'(s)T\right) \right\rangle$$

$$= -\left\langle \left(n\frac{f''(s)}{f(s)} - n\left(\frac{f'(s)}{f(s)}\right)^2 - f''(s)T\right) \right\rangle$$

Fisher Information: Tuning Curve + Noise Model

- Consider a single neuron with tuning curve *r=f(s)* and Poisson spike count *n:*

$$p(n|s) = \text{Poiss}(n|f(s)T) = (f(s)T)^n e^{-f(s)T}/n!$$

- The Fisher information is:

$$I_F(s) = -\left\langle \frac{d^2 \log p(n|s)}{ds^2} \right\rangle = -\left\langle \frac{d^2}{ds^2}(n\log(f(s)T) - f(s)T - \log n!) \right\rangle$$

$$= -\left\langle \frac{d}{ds}\left(n\frac{f'(s)T}{f(s)T} - f'(s)T\right) \right\rangle$$

$$= -\left\langle \left(n\frac{f''(s)}{f(s)} - n\left(\frac{f'(s)}{f(s)}\right)^2 - f''(s)T\right) \right\rangle$$

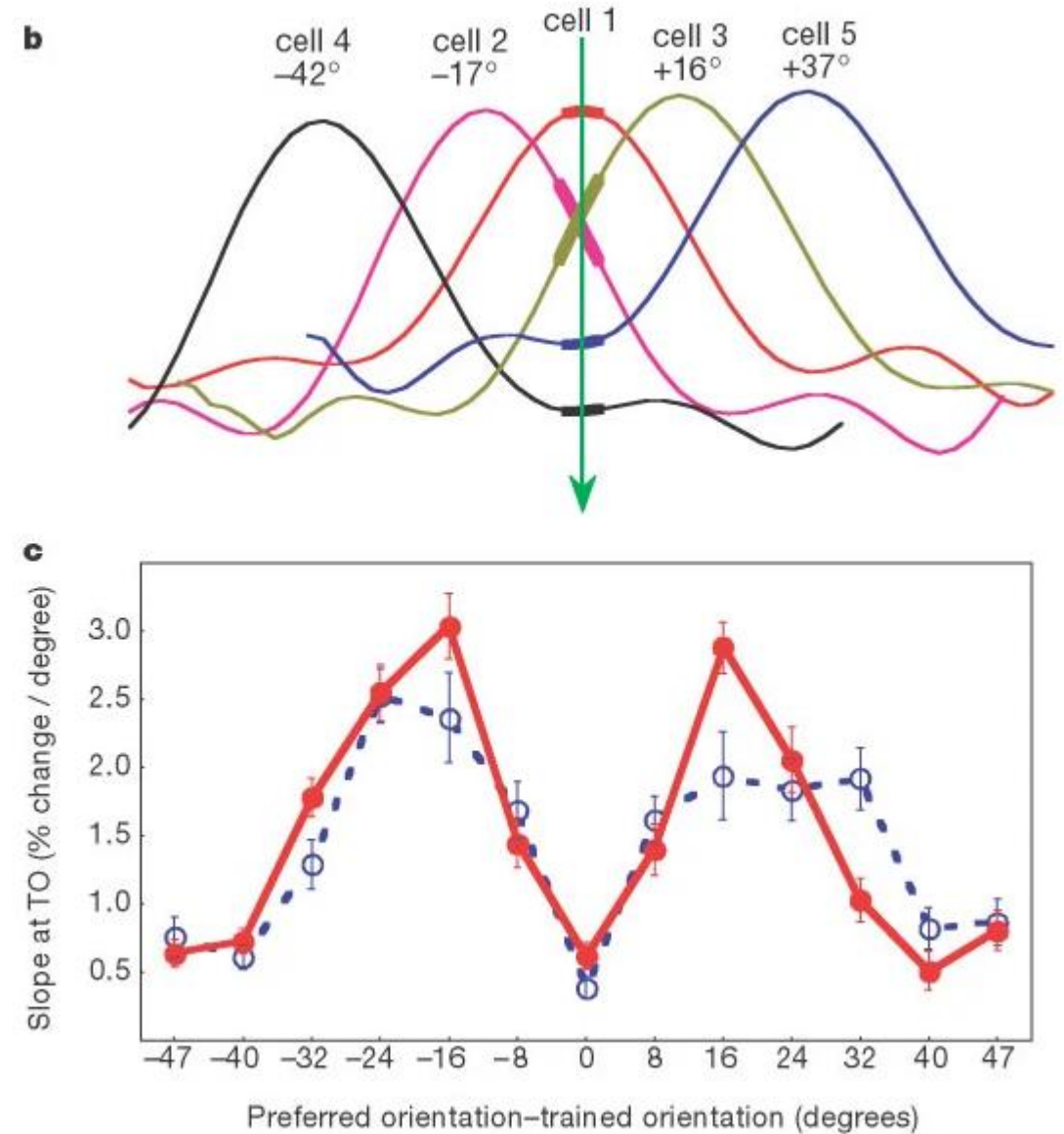$$= \frac{f'(s)^2}{f(s)}T \qquad (\text{since } \langle n \rangle = f(s)T)$$

Fisher Information and Tuning Curves

- The Fisher information is highest at flanks of tuning curve, and zero at the peak!

- Fisher information increases with *T*

- Fisher information increases with *f'(s)^2*, decreases with *f(s)*

- This is because Poisson has variance=mean. In general, Fisher info is a signal-to-noise ratio (mean^2/variance)
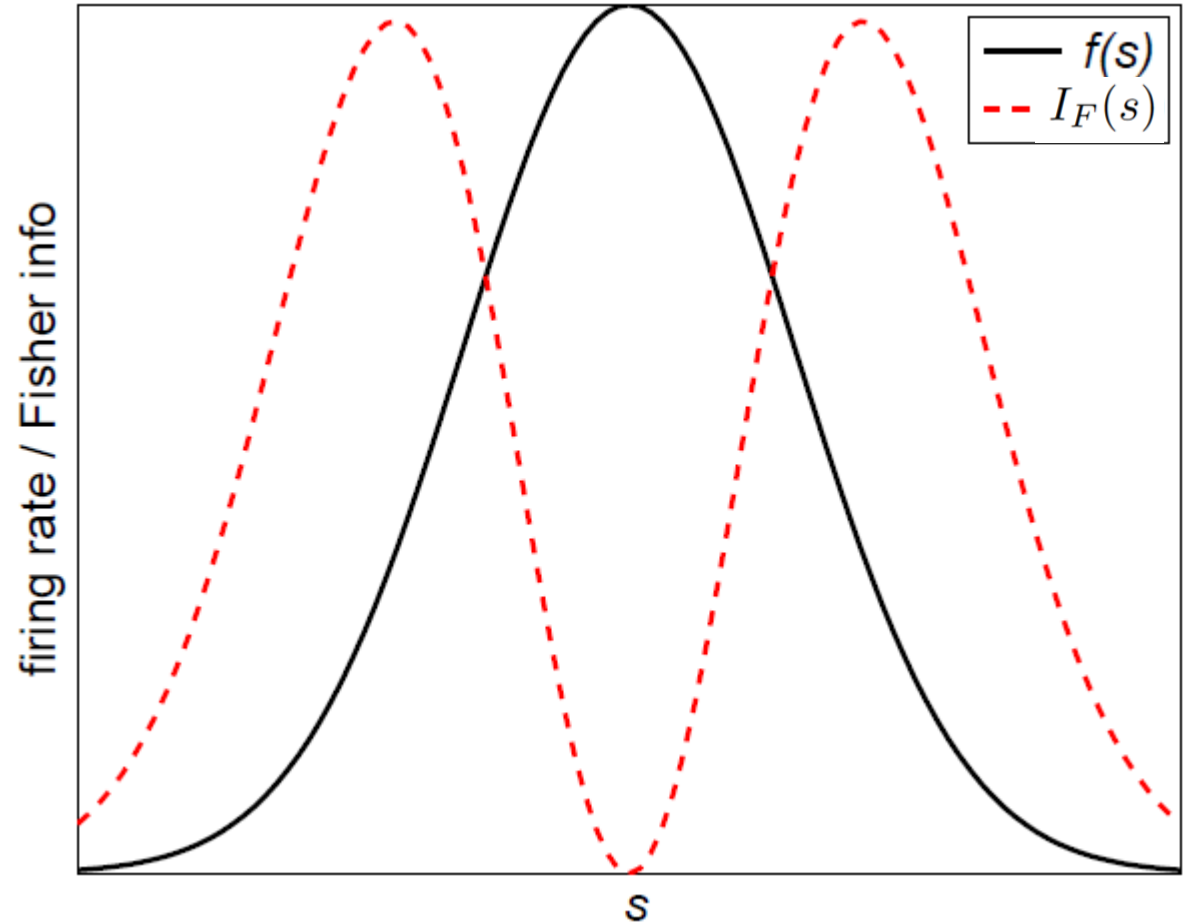
Effects of Perceptual Learning on Tuning Curves

- Monkeys were trained to discriminate between very similar orientations

- With practice, their performance improved

- The slope of tuning curves flanking the trained orientation increased

- This suggests changes with learning increase Fisher information

- Influential study, but some failed replications... still debated

Schoups et al., 2001

Summary: Fisher Information in Single Neurons

- For single neurons, Fisher information depends on the slope of the tuning curve and the response variability

- We derived the Fisher information for Poisson noise (see lecture notes for the Gaussian noise case)

- For Gaussian tuning curves, Fisher information is highest at flanks and lowest at peak (exercise: how would this look for a sigmoidal tuning curve?)

- Learning a fine-scale stimulus discrimination task seems to alter tuning curves to enhance Fisher information

- How learning and attention influence neural responses and Fisher information is an ongoing debate

Fisher Information for Independent Poisson Neurons

- Consider a population of independent Poisson neurons with rates *r=f(s-si)*:

$$p(n_i|s) = \text{Poiss}(n_i|f(s-s_i)T) = (f(s-s_i)T)^{n_i} e^{-f(s-s_i)T}/n_i!$$

$$p(\mathbf{n}|s) = \prod_{i=1}^{N} p(n_i|s)$$

Fisher Information for Independent Poisson Neurons

- Consider a population of independent Poisson neurons with rates *r=f(s-si)*:

$$p(n_i|s) = \text{Poiss}(n_i|f(s - s_i)T) = (f(s - s_i)T)^{n_i} e^{-f(s-s_i)T}/n_i!$$

$$p(\mathbf{n}|s) = \prod_{i=1}^{N} p(n_i|s)$$

- The population Fisher information is:

$$I_F(s) = -\left\langle \frac{d^2 \log p(\mathbf{n}|s)}{ds^2} \right\rangle = \sum_i -\left\langle \frac{d^2 \log p(n_i|s)}{ds^2} \right\rangle = \sum_i \frac{f'(s - s_i)^2}{f(s - s_i)} T$$

- Note: the Fisher information increases linearly with the number of neurons - unlimited information is unrealistic (recall the Data Processing Inequality).

Dependence of Fisher Information on Tuning Curve Width

- In the limit of many neurons, assuming homogenous and dense tuning curves, we have:

$$I_F(s) = \sum_i \frac{f'(s - s_i)^2}{f(s - s_i)} T \rightarrow \rho T \int_{-\infty}^{\infty} \frac{f'(s - s')^2}{f(s - s')} ds'$$

- Where ρ is the density of neurons in stimulus space.

Dependence of Fisher Information on Tuning Curve Width

- In the limit of many neurons, assuming homogenous and dense tuning curves, we have:

$$I_F(s) = \sum_i \frac{f'(s - s_i)^2}{f(s - s_i)} T \to \rho T \int_{-\infty}^{\infty} \frac{f'(s - s')^2}{f(s - s')} ds'$$

- Where ρ is the density of neurons in stimulus space. For Gaussian tuning curves we have:

$$f(s - s_i) = A e^{-(s - s_i)^2 / (2\sigma^2)}$$

Dependence of Fisher Information on Tuning Curve Width

- In the limit of many neurons, assuming homogenous and dense tuning curves, we have:

$$I_F(s) = \sum_i \frac{f'(s - s_i)^2}{f(s - s_i)} T \rightarrow \rho T \int_{-\infty}^{\infty} \frac{f'(s - s')^2}{f(s - s')} ds'$$

- Where ρ is the density of neurons in stimulus space. For Gaussian tuning curves we have:

$$f(s - s_i) = A e^{-(s - s_i)^2 / (2\sigma^2)}$$

$$\Longrightarrow I_F(s) = \frac{\rho T A}{\sigma^4} \int_{-\infty}^{\infty} (s - s')^2 e^{-(s - s')^2 / (2\sigma)^2} ds'$$

- In the limit of many neurons, assuming homogenous and dense tuning curves, we have:

$$I_F(s) = \sum_i \frac{f'(s - s_i)^2}{f(s - s_i)} T \to \rho T \int_{-\infty}^{\infty} \frac{f'(s - s')^2}{f(s - s')} ds'$$

- Where ρ is the density of neurons in stimulus space. For Gaussian tuning curves we have:

$$f(s - s_i) = A e^{-(s - s_i)^2/(2\sigma^2)}$$

$$\implies I_F(s) = \frac{\rho T A}{\sigma^4} \int_{-\infty}^{\infty} (s - s')^2 e^{-(s - s')^2/(2\sigma)^2} ds'$$

$$= \frac{\sqrt{2\pi} \rho T A}{\sigma}$$

- So narrower tuning curves yield higher Fisher information!

Evidence for Tuning Curve Sharpening in Data

- Monkeys were trained to discriminate orientations. For difficult decisions, monkeys paid closer attention

- Attention sharpened and amplified orientation tuning curves in visual cortex (area V4)

- This should cause Fisher information to increase, enabling better discrimination performance

- Reality is more complicated - attention also influences noise, and its hard to measure Fisher information directly. How learning and attention shape population codes is still debated

Tuning curve of neuron

Spitzer et al., 1988
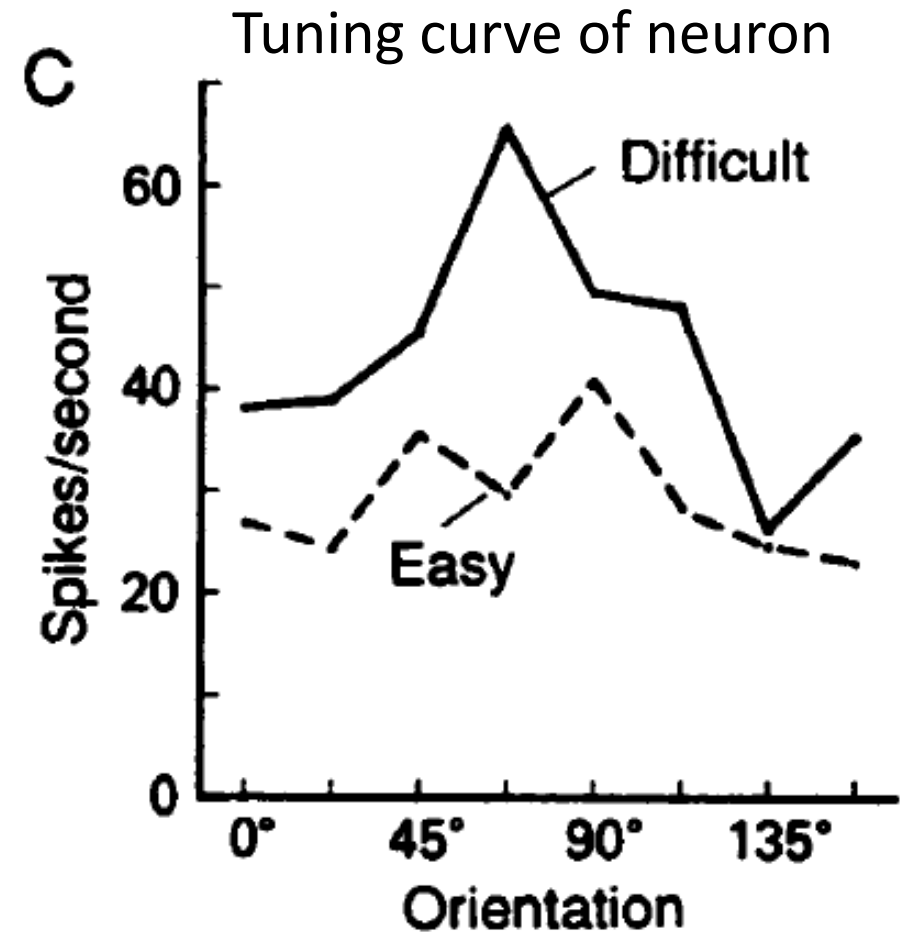
Narrow vs Broad Tuning Curves in Higher Dimensions

- When s is 1-dimensional, narrow tuning curves are best

- For 2-D stimuli $s=[s1,s2]$ (e.g., orientation+speed), Fisher information is independent of tuning width

- For 3-D and higher, broader tuning curves are best

- Other factors also matter, e.g. noise model, correlations between neurons, etc.

Zhang and Sejnowski, 1999
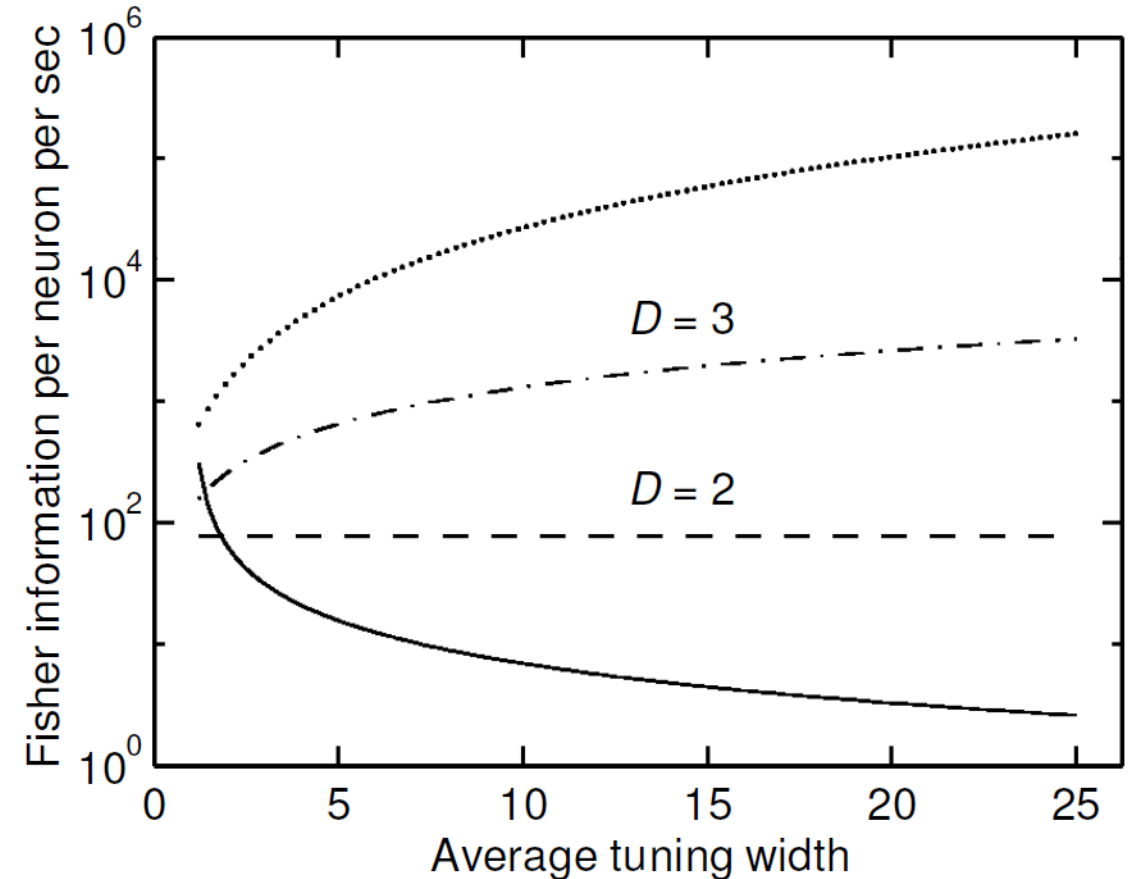
Summary: Fisher Information in Independent Populations

- We assumed a population of neurons with independent Poisson noise

- For independent neurons, Fisher information of each neuron simply adds to give the population information

- This can lead to some implausible results, e.g. unlimited information with increasing N

- For a large population of neurons with homogeneous Gaussian tuning curves, information increases as tuning width decreases (for 1D stimuli…)

- This suggests that sharpening of tuning curves might increase information – whether learning or attention cause tuning curve sharpening has been long debated

Fisher Information in Correlated Populations

- Assuming independent responses of neurons predicted an unrealistic scaling of Fisher information with population size…

- But are responses of neurons really independent?

- How would correlations between responses of different neurons influence population Fisher information?

- Take home message: 1) responses of neurons are actually correlated 2) this limits the amount of information in the population 3) the influence of correlations on Fisher information depends on the specific pattern of correlations

Noise Correlations

- Responses of pairs of neurons may be correlated over repeats of a fixed stimulus
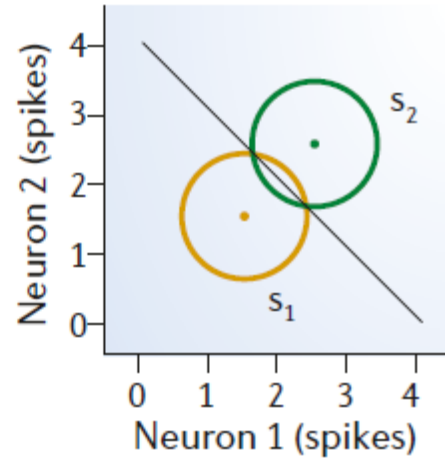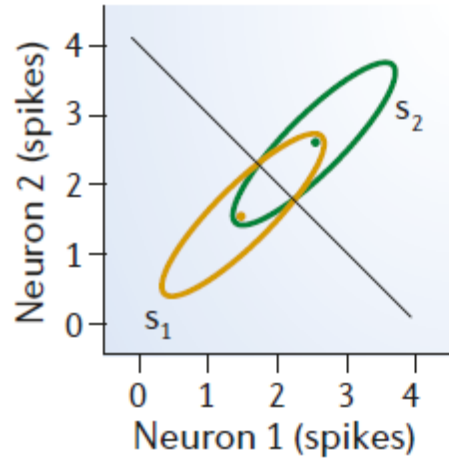
- These corelations are called noise correlations

a Tuning curves

b Examples of noise correlation at s_1

Noise Correlations

- Responses of pairs of neurons may be correlated over repeats of a fixed stimulus

- These correlations are called noise correlations
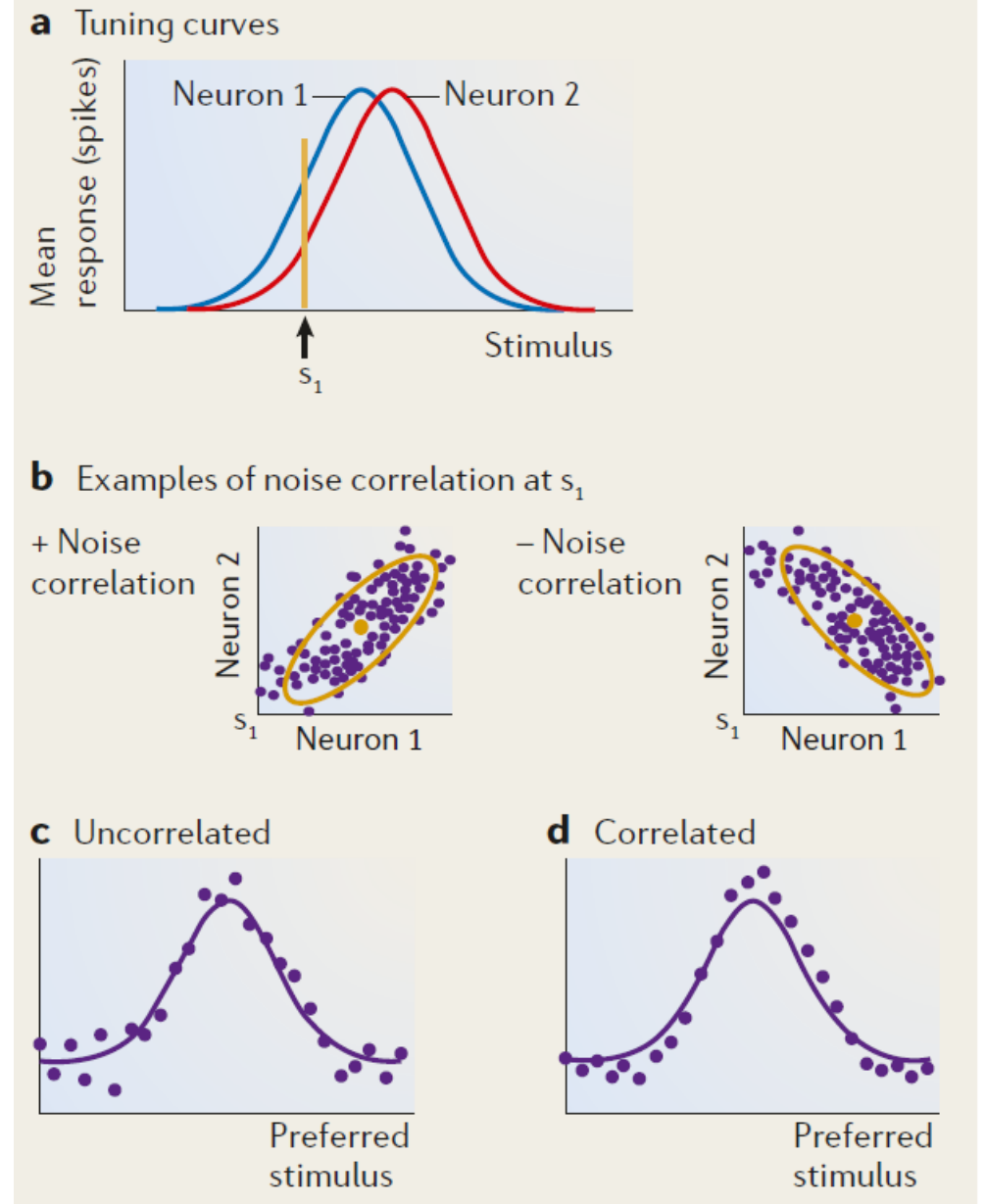
a Tuning curves

b Examples of noise correlation at s_1
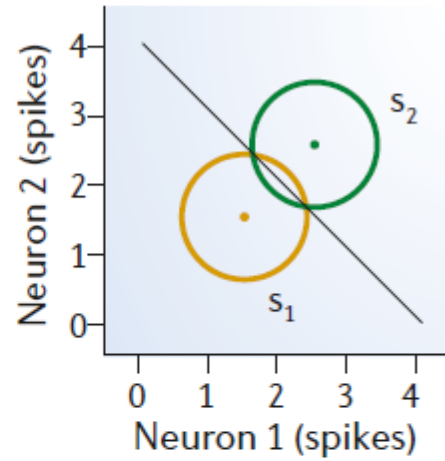
c Uncorrelated

d Correlated

Noise Correlations Influence Coding

Information (I) in unshuffled responses

$\Delta I_{shuffled} < 0$

Neuron 2 (spikes)

S_2

S_1

Neuron 1 (spikes)

Information ($I_{shuffled}$) in shuffled responses
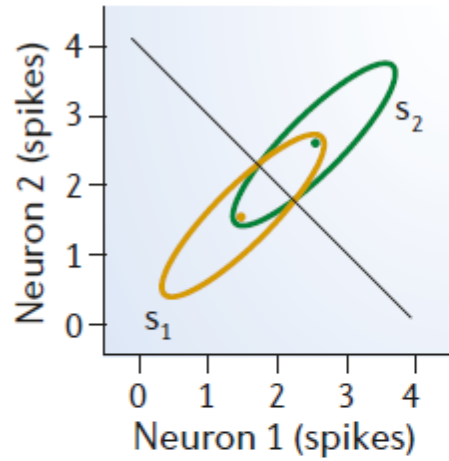
Neuron 2 (spikes)

S_2

S_1

Neuron 1 (spikes)

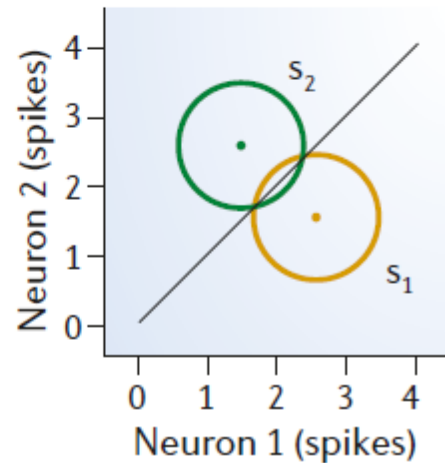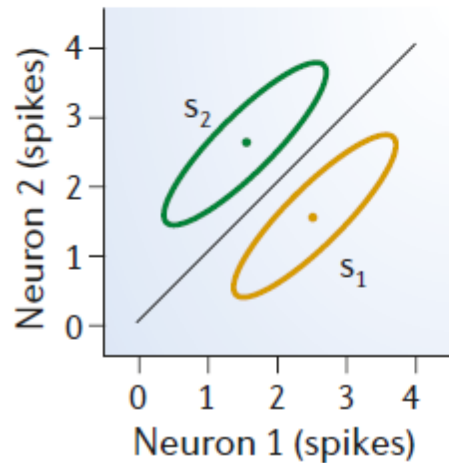# Noise Correlations Influence Coding

Information (I) in unshuffled responses

Information ($I_{shuffled}$) in shuffled responses
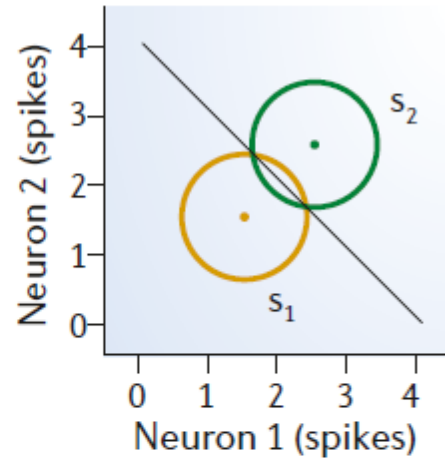
$\Delta I_{shuffled} < 0$
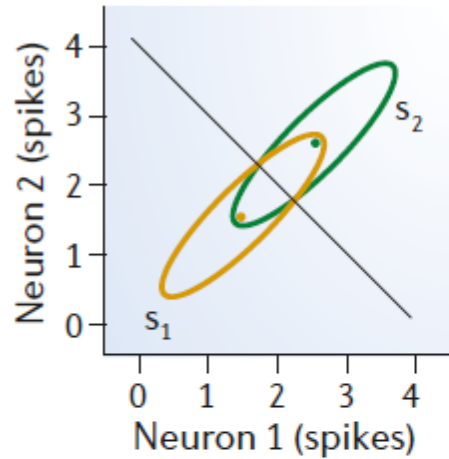
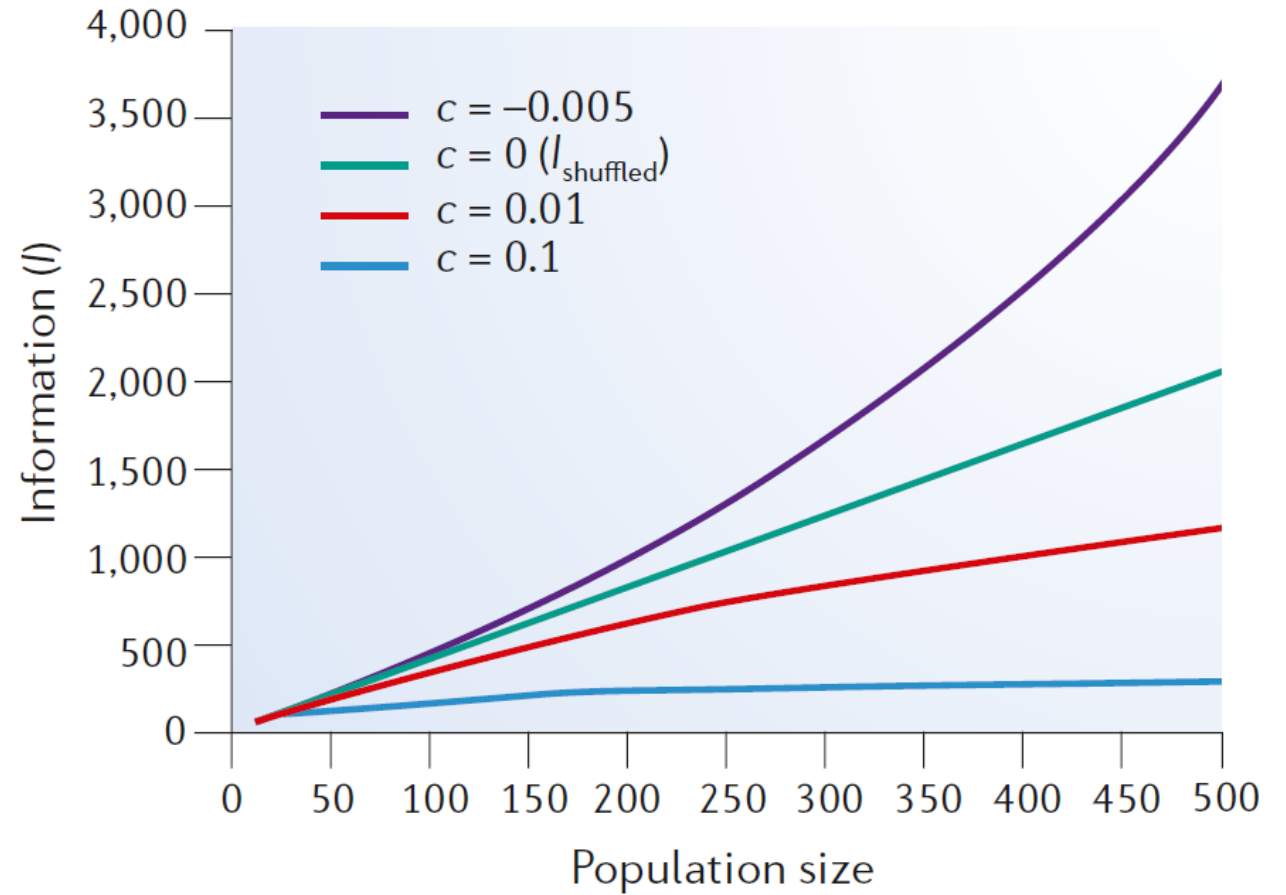$\Delta I_{shuffled} > 0$

Noise Correlations Influence Coding

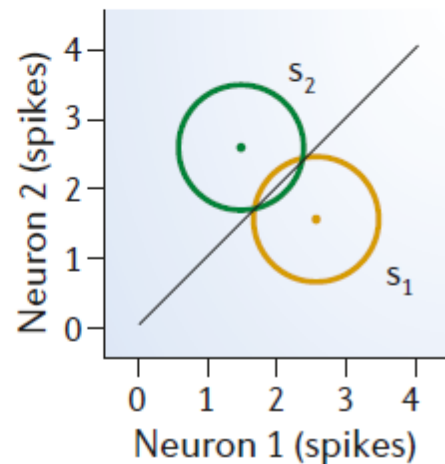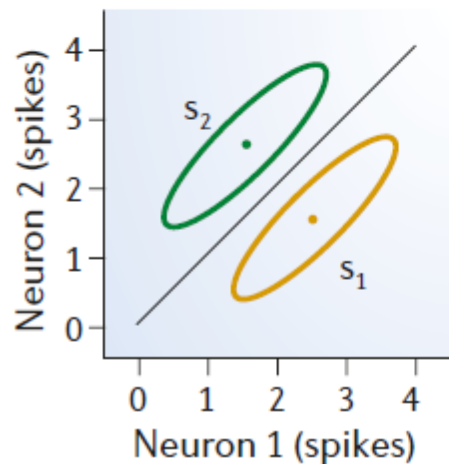Information (I) in unshuffled responses

Information ($I_{shuffled}$) in shuffled responses

$\Delta I_{shuffled} < 0$

$\Delta I_{shuffled} > 0$

- $c = -0.005$
- $c = 0$ ($I_{shuffled}$)
- $c = 0.01$
- $c = 0.1$

Noise Correlations in Neural Data

- Noise correlations in cortex tend to be weak but positive
- Suggested to limit information in populations

Correlated neuronal discharge rate and its implications for psychophysical performance

Ehud Zohary, Michael N. Shadlen & William T. Newsome*

Department of Neurobiology, Stanford University School of Medicine, Stanford, California 94305, USA

SINGLE neurons can signal subtle changes in the sensory environment with surprising fidelity, often matching the perceptual sensitivity of trained psychophysical observers[1–10]. This similarity poses an intriguing puzzle: why is psychophysical sensitivity not greater than that of single neurons? Pooling responses across neurons should average out noise in the activity of single cells, leading to substantially improved psychophysical performance. If, however, noise is correlated among these neurons, the beneficial effects of pooling would be diminished[10–12]. To assess correlation within a

Similar orientation preferences
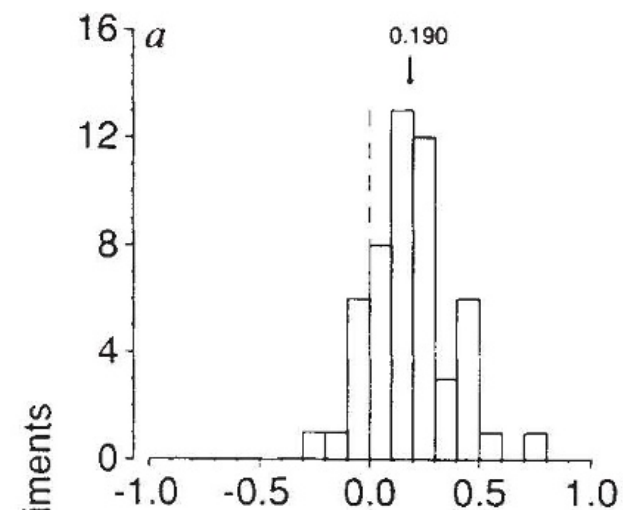
Different orientation preferences
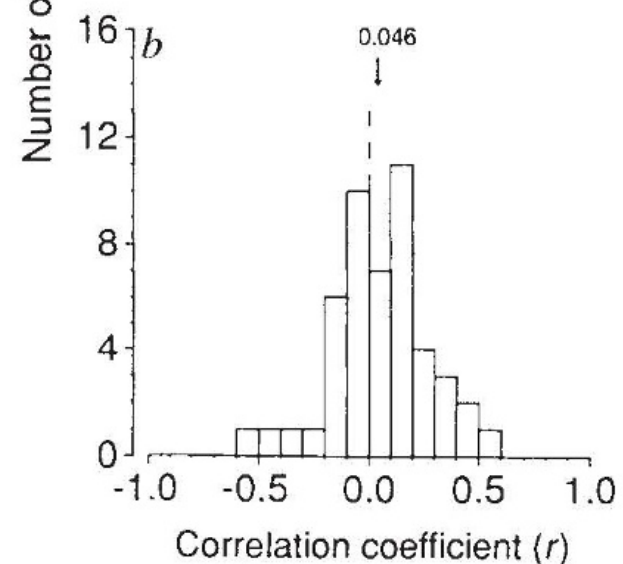
Fisher Information for Correlated Gaussian Neurons

- Model correlated population response as a multivariate normal distribution:

$$p(\mathbf{r}|s) = N(\mathbf{f}(s), \Sigma(s)) = \frac{1}{\sqrt{2\pi\det(\Sigma(s))}} \exp\left[-(\mathbf{r}-\mathbf{f}(s))^T \Sigma^{-1}(s)(\mathbf{r}-\mathbf{f}(s))\right]$$

Fisher Information for Correlated Gaussian Neurons

- Model correlated population response as a multivariate normal distribution:

$$p(\mathbf{r}|s) = N(\mathbf{f}(s), \Sigma(s)) = \frac{1}{\sqrt{2\pi \det(\Sigma(s))}} \exp\left[-(\mathbf{r} - \mathbf{f}(s))^T \Sigma^{-1}(s)(\mathbf{r} - \mathbf{f}(s))\right]$$

- The Fisher information is: $I_F(s) = \mathbf{f}'^T(s)\Sigma^{-1}(s)\mathbf{f}'(s) + \frac{1}{2}\text{Trace}\left[\left(\Sigma^{-1}(s)\frac{d\Sigma(s)}{ds}\right)^2\right]$

Fisher Information for Correlated Gaussian Neurons

- Model correlated population response as a multivariate normal distribution:

$$p(\mathbf{r}|s) = N(\mathbf{f}(s), \Sigma(s)) = \frac{1}{\sqrt{2\pi \det(\Sigma(s))}} \exp\left[-(\mathbf{r} - \mathbf{f}(s))^T \Sigma^{-1}(s)(\mathbf{r} - \mathbf{f}(s))\right]$$

- The Fisher information is: $I_F(s) = \mathbf{f}'^T(s)\Sigma^{-1}(s)\mathbf{f}'(s) + \frac{1}{2}\text{Trace}\left[\left(\Sigma^{-1}(s)\frac{d\Sigma(s)}{ds}\right)^2\right]$

- The first term is called the linear Fisher information, and depends on the relationship between the tuning curves and noise (it is a signal-to-noise ratio)

Fisher Information for Correlated Gaussian Neurons

- Model correlated population response as a multivariate normal distribution:

$$p(\mathbf{r}|s) = N(\mathbf{f}(s), \Sigma(s)) = \frac{1}{\sqrt{2\pi\det(\Sigma(s))}} \exp\left[-(\mathbf{r} - \mathbf{f}(s))^T \Sigma^{-1}(s)(\mathbf{r} - \mathbf{f}(s))\right]$$

- The Fisher information is: $\quad I_F(s) = \mathbf{f}'^T(s)\Sigma^{-1}(s)\mathbf{f}'(s) + \frac{1}{2}\mathrm{Trace}\left[\left(\Sigma^{-1}(s)\frac{d\Sigma(s)}{ds}\right)^2\right]$

- The first term is called the linear Fisher information, and depends on the relationship between the tuning curves and noise (it is a signal-to-noise ratio)

- The second term involves stimulus-dependent noise, and is less intuitive (but interesting nonetheless!)

Information-Limiting Correlations

- What kinds of noise correlations are harmful for stimulus coding?
- Correlations which limit information take on a precise form: $\Sigma_{ILC} \propto \mathbf{f}'(s)\mathbf{f}'^T(s)$

Information-Limiting Correlations

- What kinds of noise correlations are harmful for stimulus coding?

- Correlations which limit information take on a precise form: $\Sigma_{ILC} \propto \mathbf{f}'(s)\mathbf{f}'^{T}(s)$

- These correlations cause the population tuning curve to shift from side to side, which looks like a change in stimulus to a decoder

Moreno-Bote et al., 2014
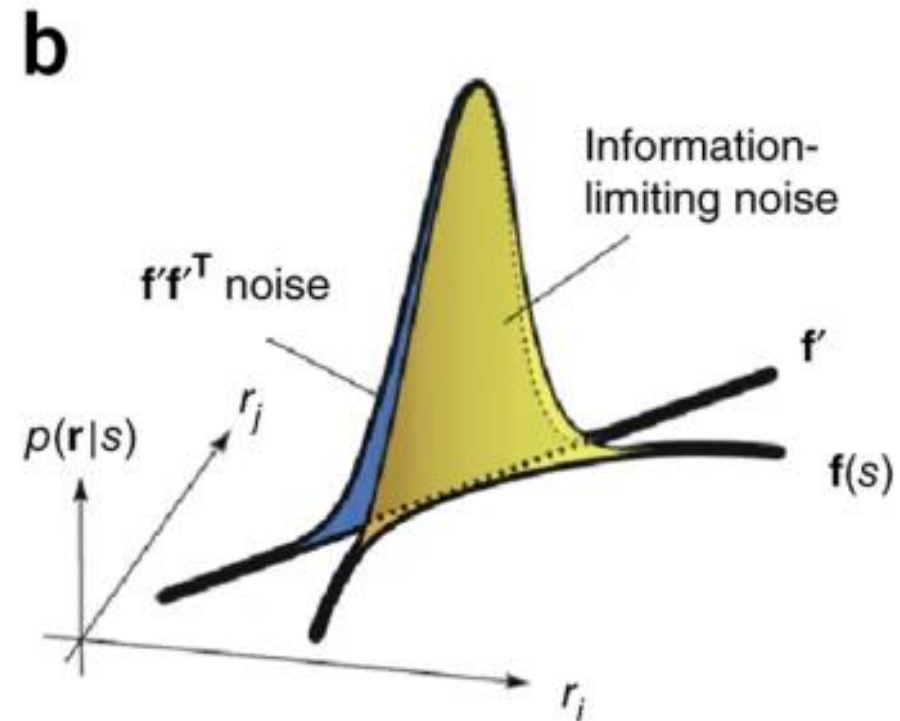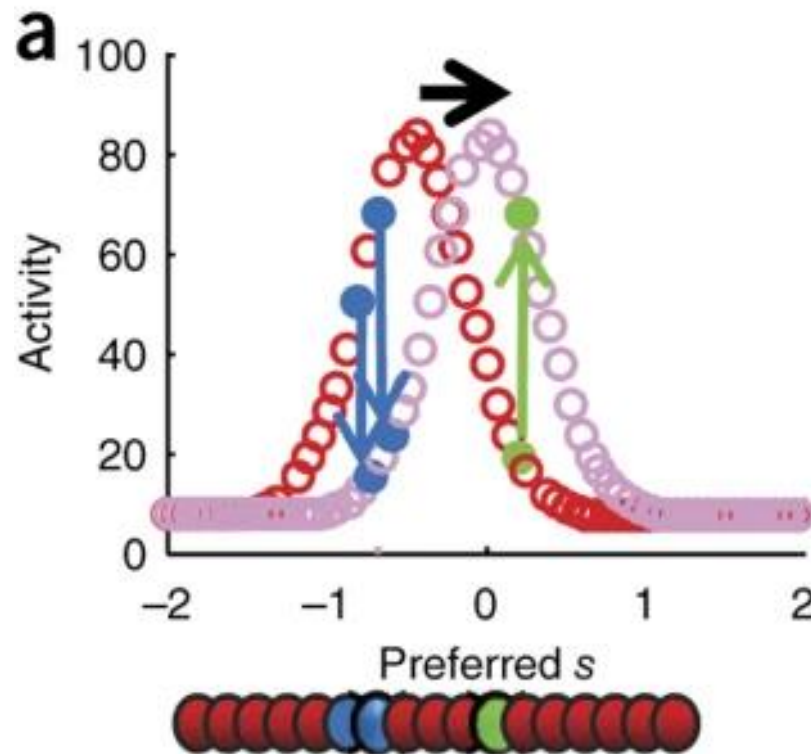
Information-Limiting Correlations

- Adding information-limiting correlations causes population Fisher information to decrease:

$$\Sigma \to \Sigma + \epsilon \mathbf{f}'(s)\mathbf{f}'^T(s) \implies I_F \to \frac{I_F}{1 + \epsilon I_F} \leq \frac{1}{\epsilon}$$

- For example, information increases linearly with *N* for independent neurons, but asymptotes to 1/epsilon when ILCs are added

- ILCs are hard to accurately estimate from data, so there is some debate on how strong they are in real neural populations

Moreno-Bote et al., 2014

Does Fisher Information Determine Perceptual Performance?

Recent study: argued that neural coding is 100x more accurate than perceptual threshold!

High-precision coding in visual cortex

Graphical abstract
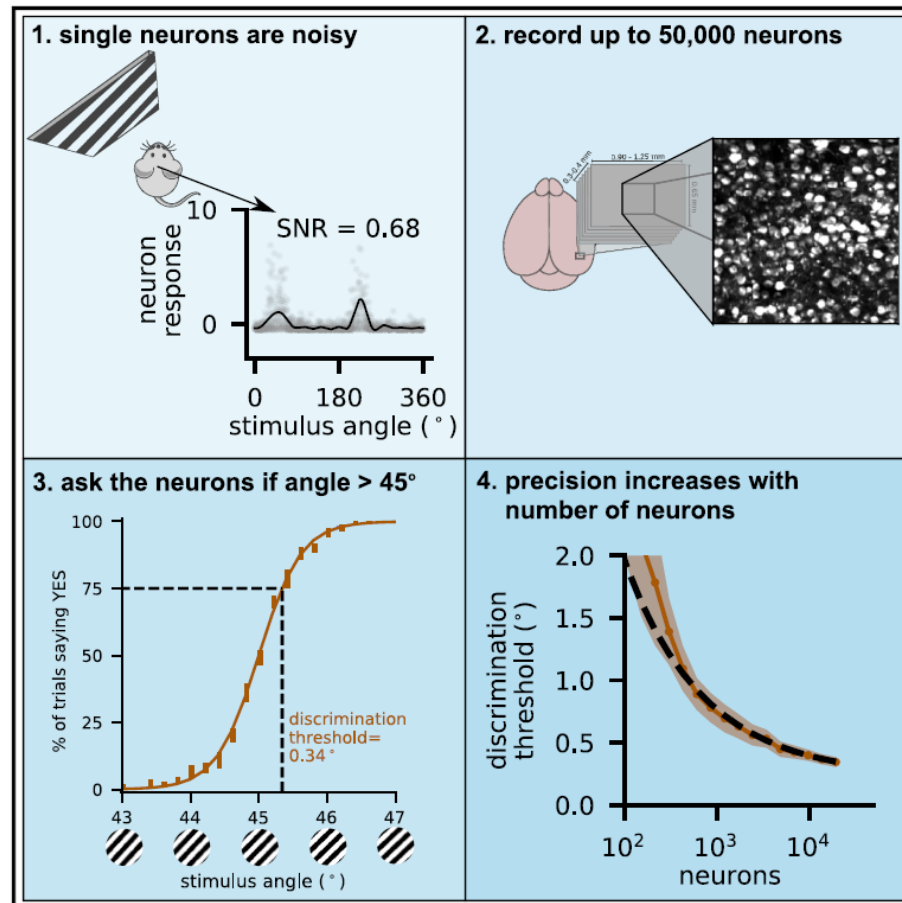
Authors

Carsen Stringer, Michalis Michaelos, Dmitri Tsyboulski, Sarah E. Lindo, Marius Pachitariu

Correspondence

pachitarium@janelia.hhmi.org

In brief

Large-scale recordings in mouse primary visual cortex and higher order visual areas uncover neural representations more precise than behavioral discrimination thresholds, suggesting visual perception is limited by non-sensory brain networks.

Does Fisher Information Determine Perceptual Performance?

Another recent study: argued that neural coding threshold = perceptual threshold...

Article

# Fundamental bounds on the fidelity of sensory cortical coding

Oleg I. Rumyantsev[1,2,3], Jérôme A. Lecoq[1,2,4,5], Oscar Hernandez[1,2], Yanping Zhang[1,2,6], Joan Savall[1,2,6], Radosław Chrapkiewicz[1,2], Jane Li[1,4], Hongkui Zeng[5], Surya Ganguli[1,3] & Mark J. Schnitzer[1,2,3,4,6]

How the brain processes information accurately despite stochastic neural activity is a longstanding question[1]. For instance, perception is fundamentally limited by the information that the brain can extract from the noisy dynamics of sensory neurons. Seminal experiments[2,3] suggest that correlated noise in sensory cortical neural ensembles is what limits their coding accuracy[4–6], although how correlated noise affects neural codes remains debated[7–11]. Recent theoretical work proposes that how a neural ensemble's sensory tuning properties relate statistically to its correlated noise patterns is a greater determinant of coding accuracy than is absolute noise strength[12–14]. However, without simultaneous recordings from thousands of cortical neurons with shared sensory inputs, it is unknown whether correlated noise limits coding fidelity. Here we present a 16-beam, two-photon microscope to monitor activity across the mouse primary visual cortex, along with analyses to quantify the information conveyed by large neural ensembles. We found that, in the visual cortex, correlated noise constrained signalling for ensembles with 800–1,300 neurons.

Summary: Fisher Information in Correlated Populations

- We used a multivariate normal distribution to investigate the impact of correlations on Fisher information

- Noise correlations can strongly influence Fisher information

- Noise correlations in cortex are typically weak but positive

- Such noise correlations should cause information to saturate as population size increases, resolving the unrealistic scaling of the independent population model

- There is debate about the extent to which noise correlations impact population information, and the relationship between Fisher information and perceptual performance

Fisher Information: Limitations

- Fisher information quantifies the information available to an optimal estimator, but it is not clear that the brain uses an optimal estimator (indeed, the brain rarely needs to "decode")

- Fisher information is parametric – we have to define stimulus features of interest. Does the brain care about orientation for example?

- Fisher information is a local measure. Sometimes global errors may occur, which Fisher information is blind to (e.g., perceptual illusions)

- Ignores mechanistic/physical constraints. Can get infinite information. In practice, response statistics depend on connectivity and external inputs to a neural circuit, and infinite information is impossible.

Fisher Information vs Mutual Information

- Mutual information is the reduction in uncertainty about *s* upon observing *r*

- Fisher information is the decoding error on *s* from *r*.

- Mutual information is a global measure over all *s*. Fisher information is local to a given value of *s*.

- In practice: Fisher information is much easier to compute and estimate from data, so most calculations involving population codes use Fisher information.

Bibliography

Lecture notes chapter 9

Dayan and Abbott Chapter 3. [Neural Decoding]