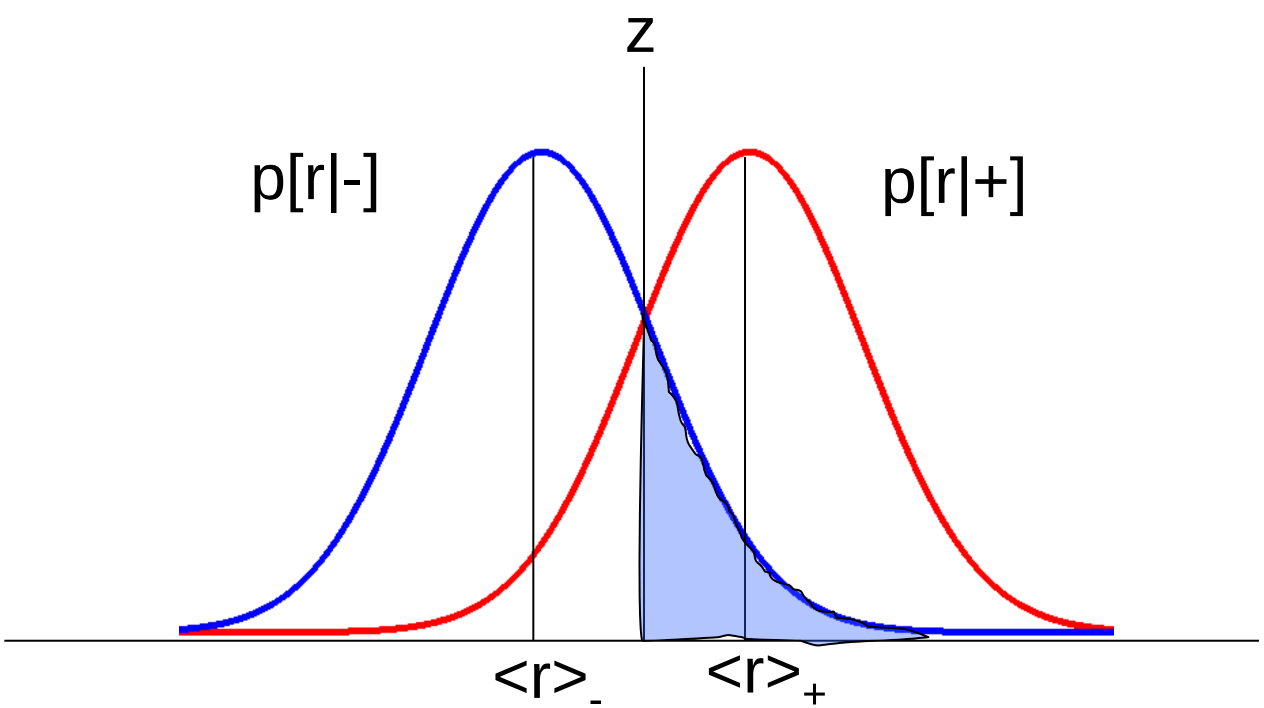
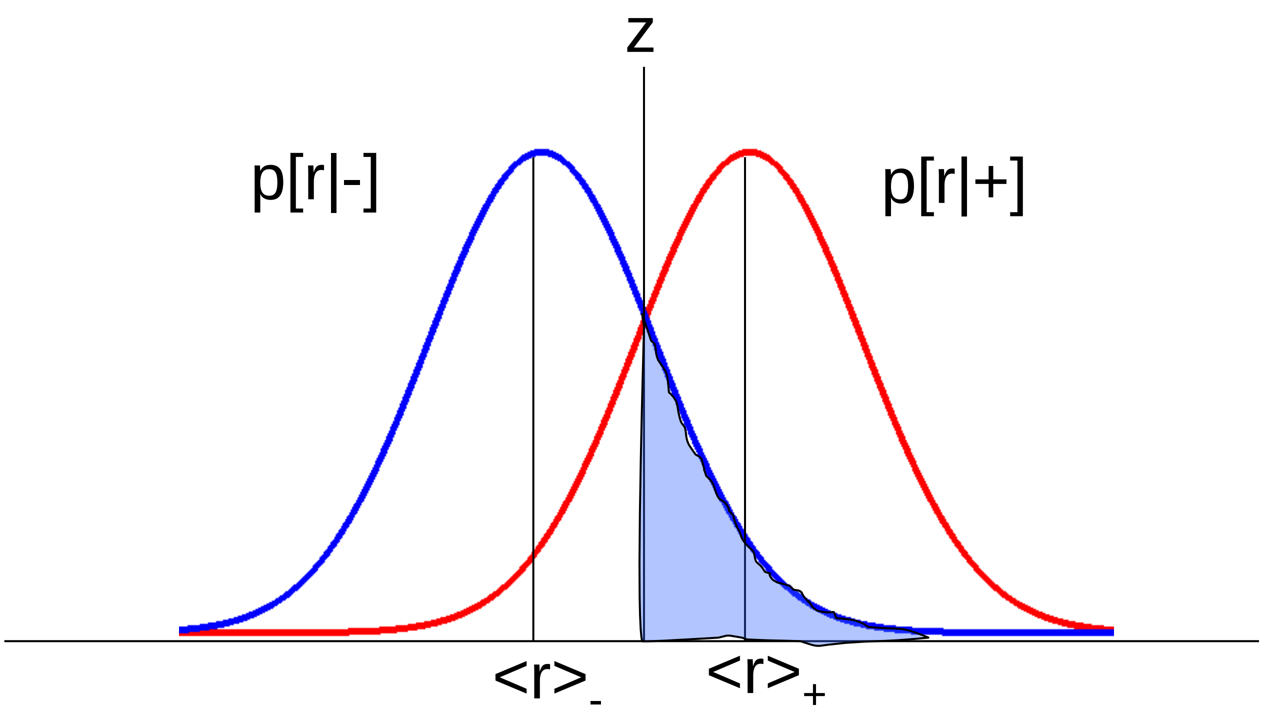
Signal detection theory



Role of *priors*:

Find z by maximizing P[correct] = p[+] $\beta(z)$ + p[-]$(1 - \alpha(z))$

Is there a better test to use than *r*?



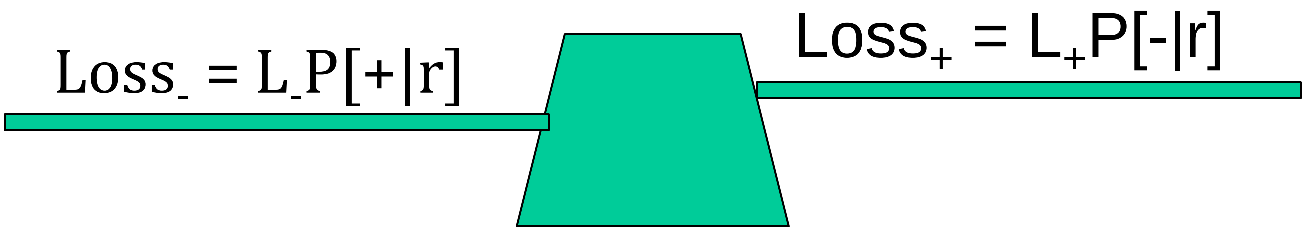
The optimal test function is the *likelihood ratio*,

l(r) = p[r|+] / p[r|-]. (Neyman-Pearson lemma)

Building in cost

Penalty for incorrect answer: L_+, L_-

For an observation r, what is the expected **loss**?



Cut your losses: answer + when $Loss_+ < Loss_-$

i.e. when $L_+ P[-|r] < L_- P[+|r]$.

Using Bayes', $P[+|r] = p[r|+]P[+]/p(r)$;

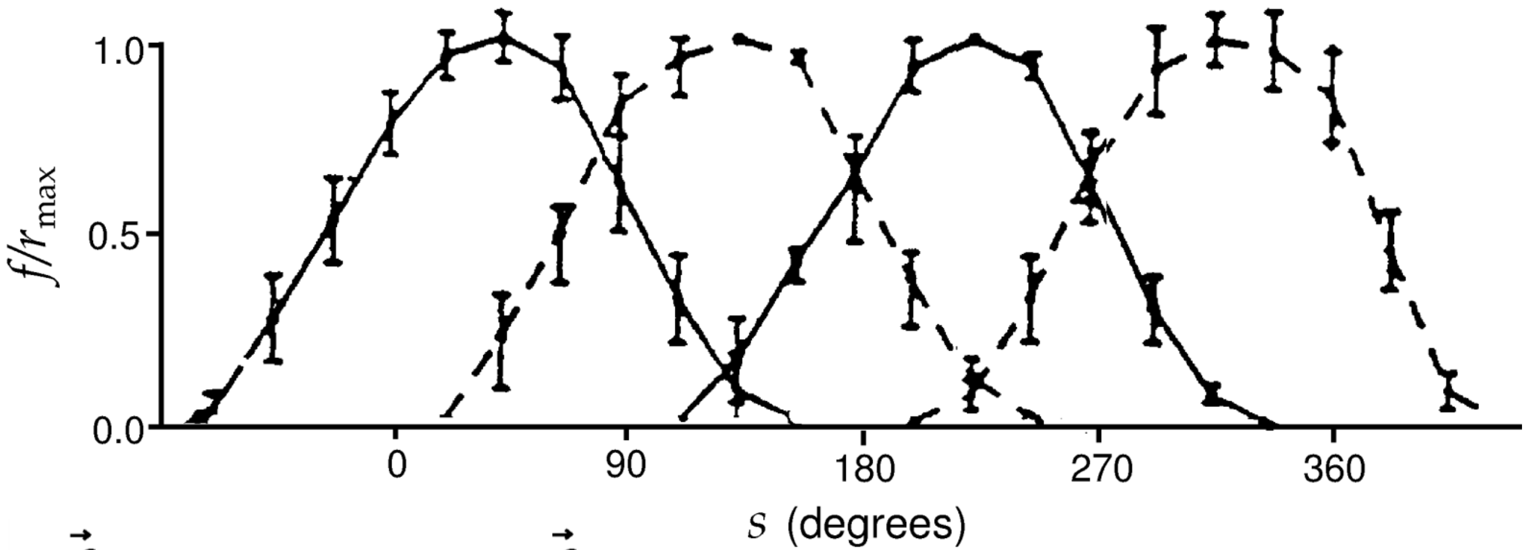
$P[-|r] = p[r|-]P[-]/p(r)$;

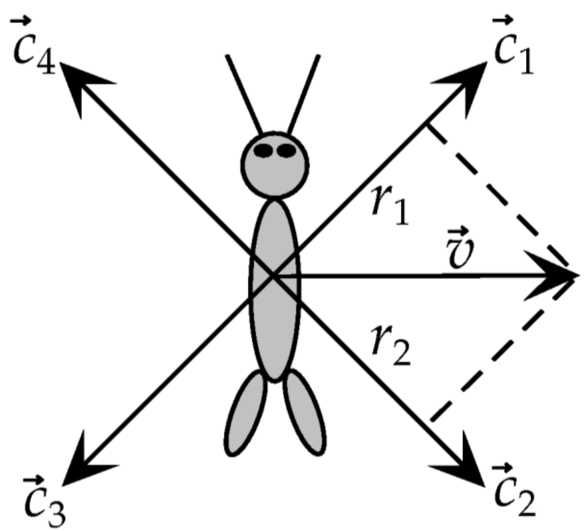
→ $l(r) = p[r|+]/p[r|-] > L_+ P[-] / L_- P[+]$.

Decoding from many neurons: population codes

- Population code formulation
- Methods for decoding:
  - → population vector
  - → Bayesian inference
  - → maximum likelihood
  - → maximum a posteriori
- Fisher information

Cricket cercal cells



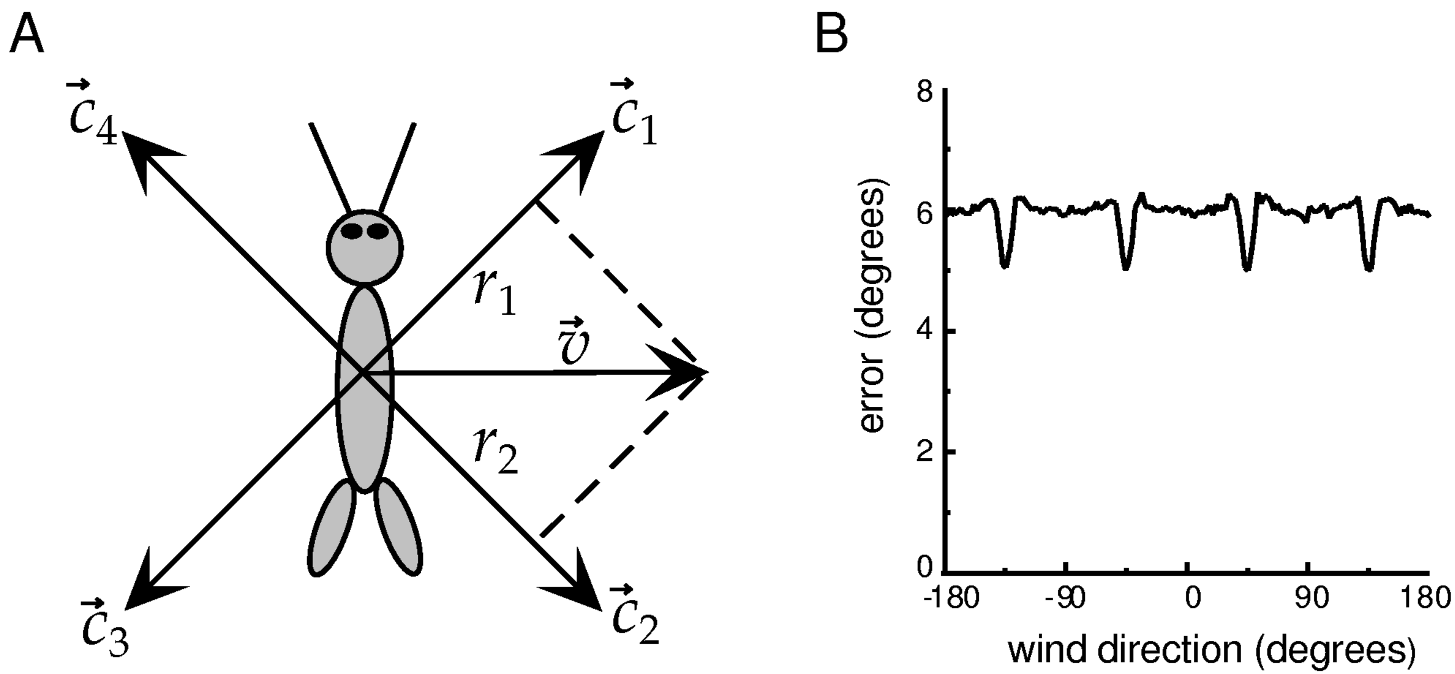


$$\left(\frac{f(s)}{r_{\max}}\right)_a = [\cos(s - s_a)]_+$$

$$\left(\frac{f(s)}{r_{\max}}\right)_a = [\vec{v} \cdot \vec{c}_a]_+$$

Population vector

$$\vec{v}_{\text{pop}} = \sum_{a=1}^{4} \left(\frac{r}{r_{\max}} \right)_a \vec{c}_a$$



RMS error in estimate

Theunissen & Miller, 1991

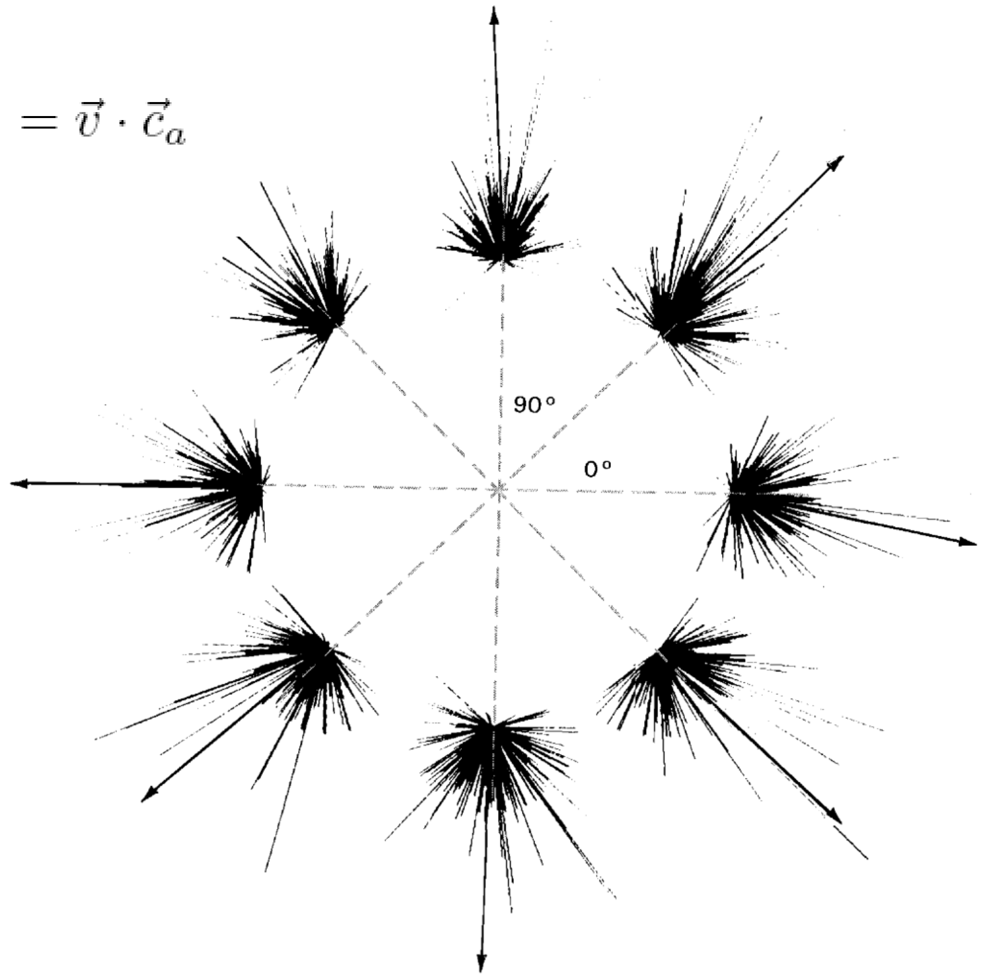
Population coding in M1

Cosine tuning:

$$\left(\frac{\langle r\rangle - r_0}{r_{\max}}\right)_a = \left(\frac{f(s) - r_0}{r_{\max}}\right)_a = \vec{v}\cdot\vec{c}_a$$

Pop. vector:

$$\vec{v}_{\text{pop}} = \sum_{a=1}^{N}\left(\frac{r - r_0}{r_{\max}}\right)\vec{c}_a$$



Is this the best one can do?

The population vector is neither general nor optimal.

“Optimal”:

make use of all information in the stimulus/response distributions

Bayesian inference

Bayes' law:

likelihood function

conditional distribution

$$p[s|\mathbf{r}] = \frac{p[\mathbf{r}|s]\,p[s]}{p[\mathbf{r}]}$$

prior distribution

a posteriori distribution

marginal distribution

Bayesian estimation

Want an estimator s_{Bayes}

Introduce a cost function, $L(s,s_{\text{Bayes}})$; minimize mean cost.

$$\int ds\, L(s, s_{\text{bayes}}) p[s|\mathbf{r}]$$

For least squares cost, $L(s,s_{\text{Bayes}}) = (s - s_{\text{Bayes}})^2$.

Let's calculate the solution..

$$s_{\text{bayes}} = \int ds\, p[s|\mathbf{r}] s$$

Bayesian inference

By Bayes' law,

likelihood function

$$p[s|\mathbf{r}] = \frac{p[\mathbf{r}|s]p[s]}{p[\mathbf{r}]}$$

a posteriori distribution

Maximum likelihood

Find maximum of p[r|s] over s

More generally, probability of the data given the "model"

"Model" = stimulus

assume parametric form for tuning curve

Bayesian inference

By Bayes' law,

likelihood function

$$p[s|\mathbf{r}] = \frac{p[\mathbf{r}|s]p[s]}{p[\mathbf{r}]}$$

a posteriori distribution

MAP and ML

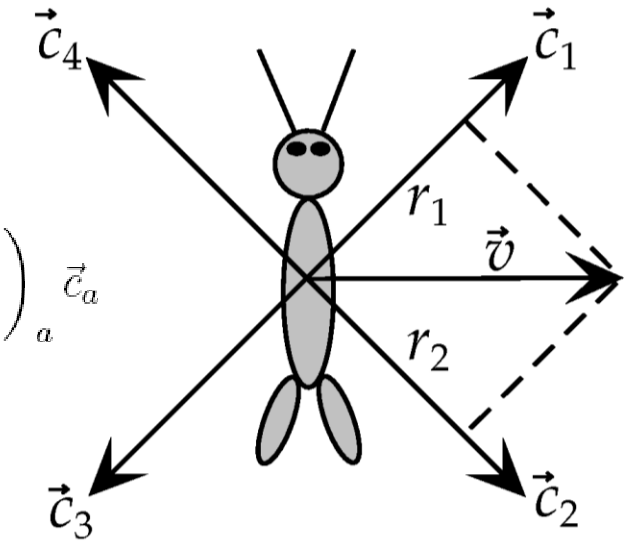
ML: s* which maximizes p[r|s]

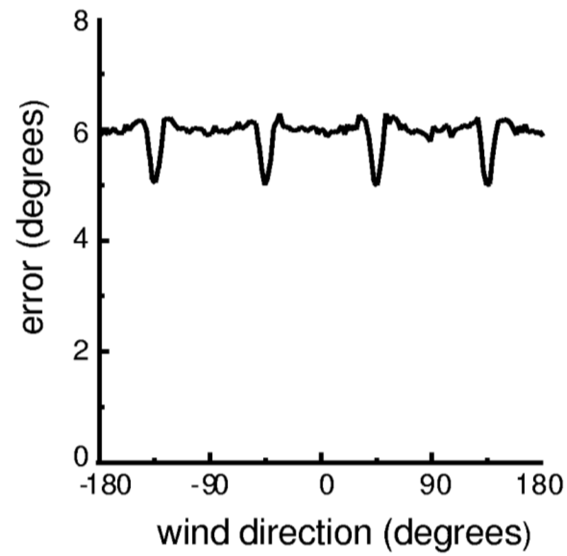
MAP: s* which maximizes p[s|r]

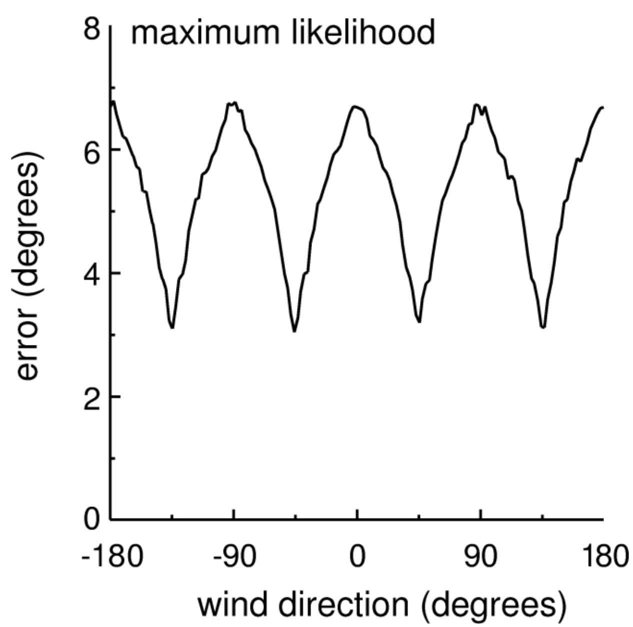
Difference is the role of the prior: differ by factor p[s]/p[r]

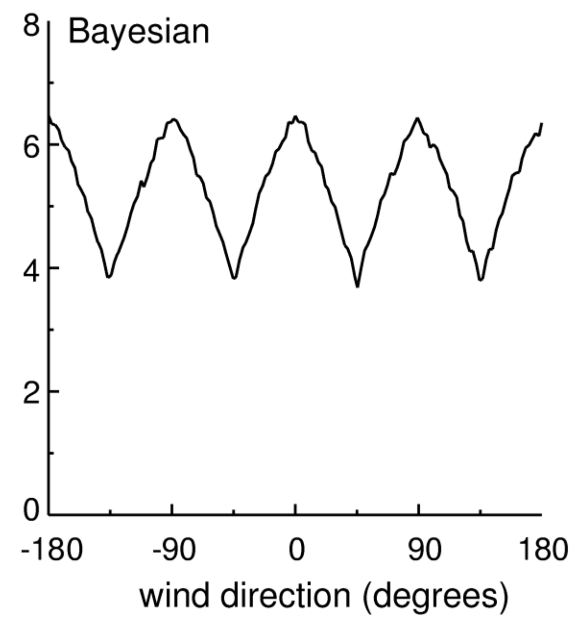
Comparison with population vector

$$\vec{v}_{\text{pop}} = \sum_{a=1}^{4} \left(\frac{r}{r_{\max}} \right)_a \vec{c}_a$$









Decoding an arbitrary continuous stimulus

Many neurons "voting" for an outcome.

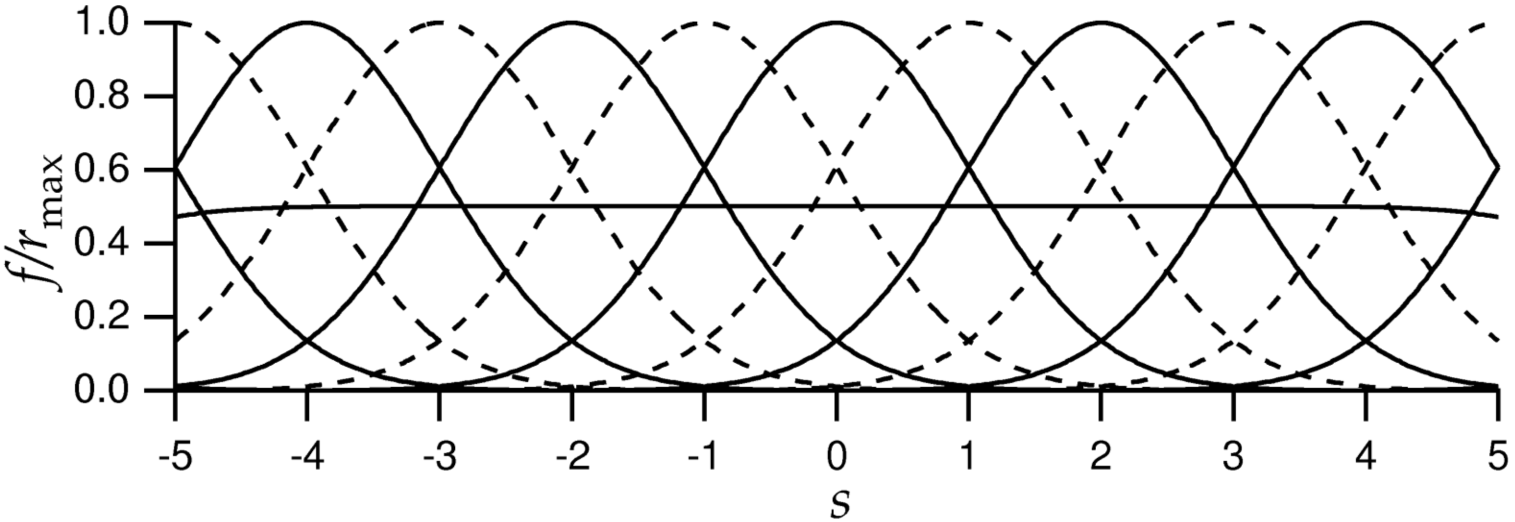
Work through a specific example

- assume independence
- assume Poisson firing

Noise model: Poisson distribution

$$P_T[k] = (\lambda T)^k \exp(-\lambda T)/k!$$

Decoding an arbitrary continuous stimulus



E.g. Gaussian tuning curves

$$f_a(s) = r_{\max} \exp\left(-\frac{1}{2}\left[\frac{(s - s_a)}{\sigma_a}\right]^2\right)$$

$$\sum_{a=1}^{N} f_a(s) \ \text{const.}$$

.. what is $P(r_a|s)$?

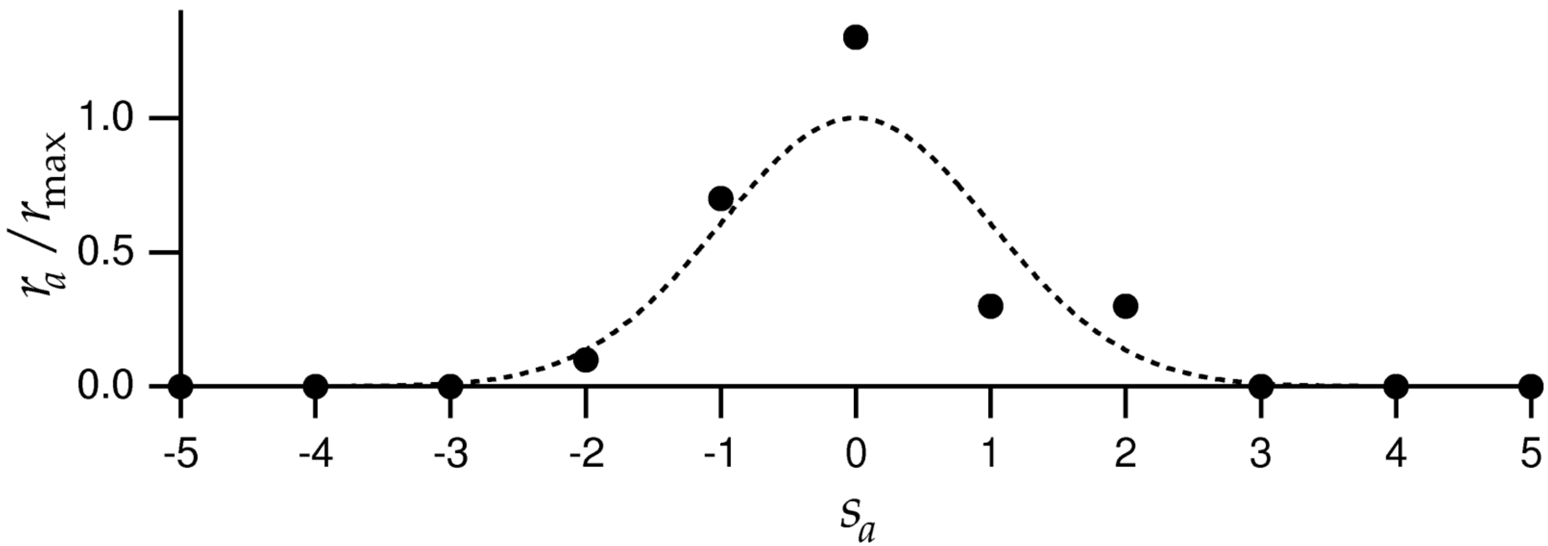
Need to know full P[**r**|s]

Assume Poisson:

$$P[r_a|s] = \frac{(f_a(s)T)^{r_aT}}{(r_aT)!}\exp(-f_a(s)T)$$

Assume independent:

$$P[\mathbf{r}|s] = \prod_{a=1}^{N} \frac{(f_a(s)T)^{r_aT}}{(r_aT)!}\exp(-f_a(s)T)$$



Population response of 11 cells with Gaussian tuning curves

ML

Apply ML: maximize ln P[**r**|s] with respect to s

$$\ln P[\mathbf{r}|s] = T\sum_{a=1}^{N} r_a \ln(f_a(s)) + \ldots$$

Set derivative to zero, use sum = constant

$$\sum_{a=1}^{N} r_a \frac{f'(s^*)}{f(s^*)} = 0$$

From Gaussianity of tuning curves,

$$s^* = \frac{\sum r_a s_a/\sigma_a^2}{\sum r_a/\sigma_a^2}$$

If all σ same

$$s^* = \frac{\sum r_a s_a}{\sum r_a}$$

MAP

Apply MAP: maximise ln p[s|**r**] with respect to s

$$\ln p[s|\mathbf{r}] = \ln P[\mathbf{r}|s] + \ln p[s] - \ln P[\mathbf{r}]$$

$$\ln p[s|\mathbf{r}] = T \sum_{a=1}^{N} r_a \ln(f_a(s)) + \ln p[s] + \ldots$$

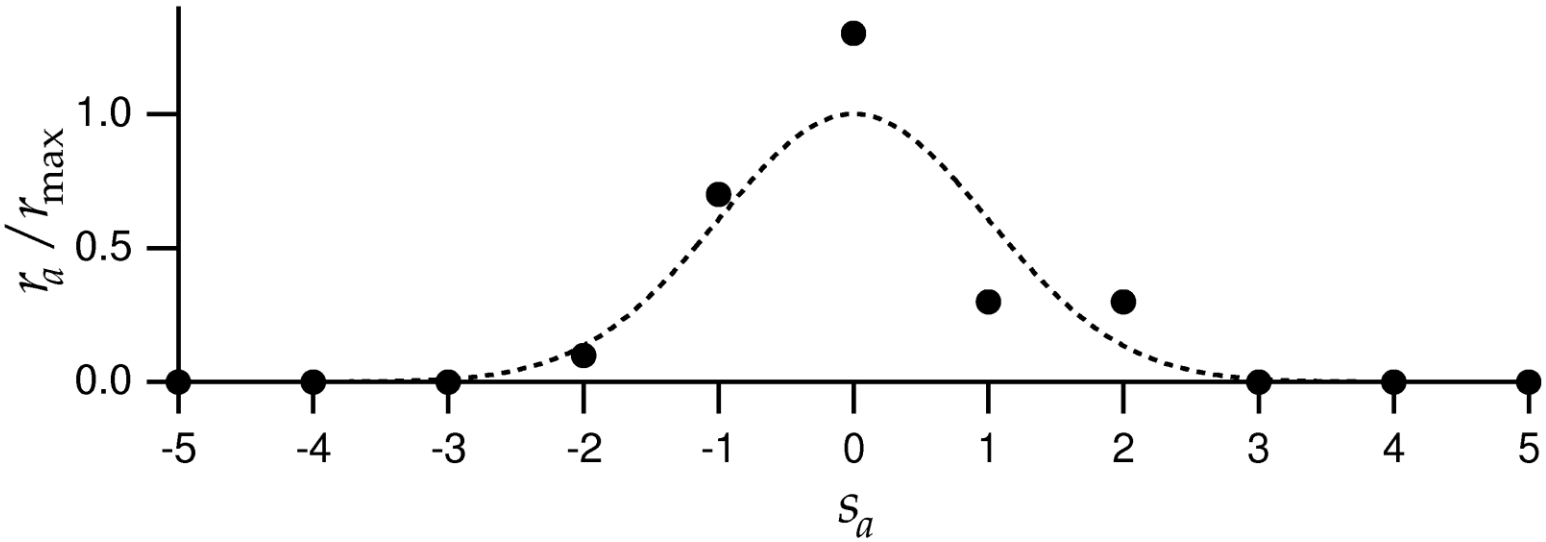
Set derivative to zero, use sum = constant

$$\sum_{a=1}^{N} r_a \frac{f'(s^*)}{f(s^*)} + \frac{p'[s]}{p[s]} = 0$$

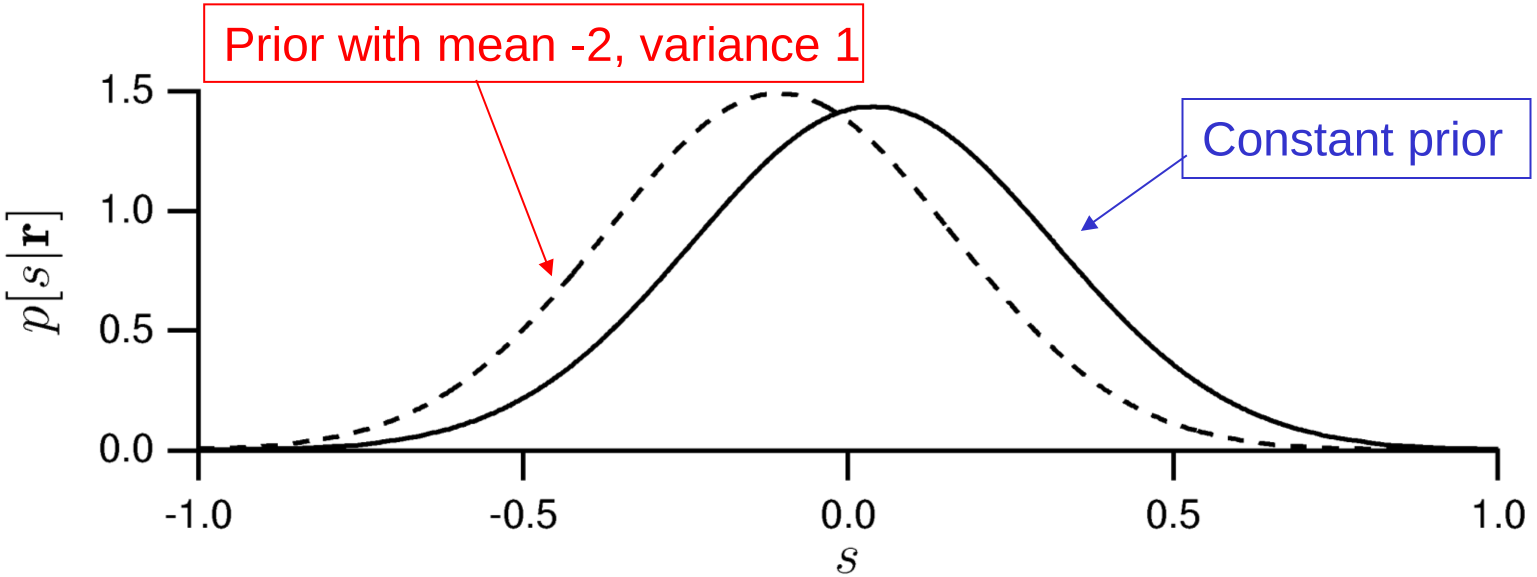
From Gaussianity of tuning curves,

$$s^* = \frac{T \sum r_a s_a / \sigma_a^2 + s_{\text{prior}} / \sigma_{\text{prior}}^2}{T \sum r_a / \sigma_a^2 + 1/\sigma_{\text{prior}}^2}$$

Given this data:



MAP:



How good is our estimate?

For stimulus s, have estimated s_{est}

Bias: $$b_{\text{est}}(s) = \langle s_{\text{est}} - s \rangle$$

Variance: $$\sigma^2_{\text{est}}(s) = \left\langle (s_{\text{est}} - \langle s_{\text{est}} \rangle)^2 \right\rangle$$

Mean square error:

$$\left\langle (s_{\text{est}} - s)^2 \right\rangle = \left\langle (s_{\text{est}} - \langle s_{\text{est}} \rangle + b_{\text{est}}(s))^2 \right\rangle = \sigma^2_{\text{est}}(s) + b^2_{\text{est}}(s).$$

Cramer-Rao bound: $$\sigma^2_{\text{est}} \geq \frac{(1 + b'_{\text{est}})^2}{I_{\text{F}}(s)}$$

Fisher information

(ML is unbiased: b = b' = 0)

Fisher information

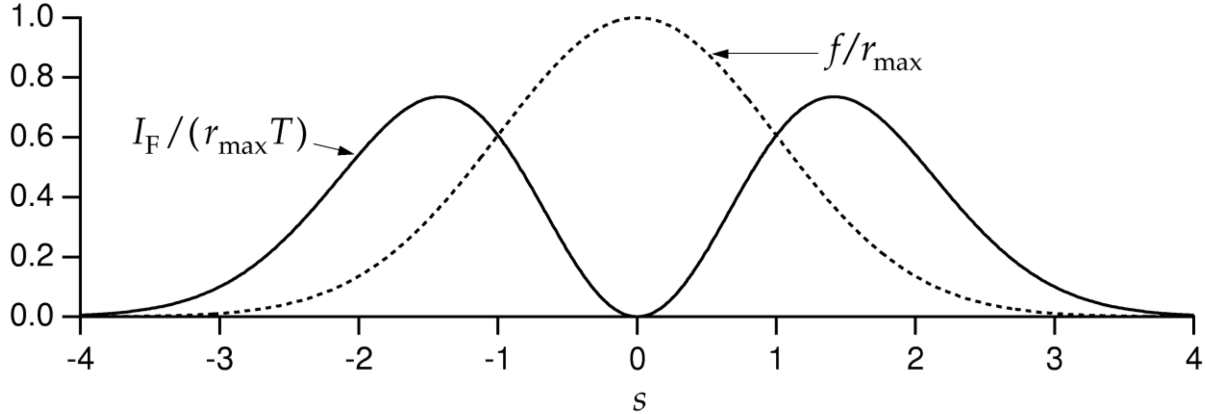
$$I_{\mathrm{F}}(s) = \left\langle -\frac{\partial^2 \ln p[\mathbf{r}|s]}{\partial^2 s} \right\rangle = \int d\mathbf{r}\, p[\mathbf{r}|s] \left(-\frac{\partial^2 \ln p[\mathbf{r}|s]}{\partial s^2} \right)$$

Alternatively:

$$I_{\mathrm{F}}(s) = \left\langle \left(\frac{\partial \ln p[\mathbf{r}|s]}{\partial s} \right)^2 \right\rangle = \int d\mathbf{r}\, p[\mathbf{r}|s] \left(\frac{\partial \ln p[\mathbf{r}|s]}{\partial s} \right)^2$$

Quantifies local stimulus discriminability

Fisher information for Gaussian tuning curves



For the Gaussian tuning curves w/Poisson statistics:

$$I_{\mathrm{F}}(s) = \left\langle \left(\frac{d^2 \ln P[\mathbf{r}|s]}{ds^2} \right) \right\rangle = T \sum_{a=1}^{N} \langle r_a \rangle \left(\left(\frac{f_a'(s)}{f_a(s)} \right)^2 - \frac{f_a''(s)}{f_a(s)} \right)$$

Are narrow or broad tuning curves better?

$$I_{\mathrm{F}} = T \sum_{a=1}^{N} \frac{r_{\max}(s - s_a)^2}{\sigma_r^4} \exp\left(-\frac{1}{2}\left(\frac{s - s_a}{\sigma_r}\right)^2\right)$$

Approximate: $$I_{\mathrm{F}} \sim \frac{\sqrt{2\pi}\rho_s \sigma_r r_{\max} T}{\sigma_r^2}.$$

Thus, $I_{\mathrm{F}} \sim 1/\sigma_r$ → Narrow tuning curves are better

But not in higher dimensions!

$$I_{\mathrm{F}} \sim (2\pi)^{D/2} D \rho_s \sigma_r^{D-2} r_{\max} T$$

..what happens in 2D?

Fisher information and discrimination

Recall d' = mean difference/standard deviation

Can also decode and discriminate using decoded values.

Trying to discriminate s and s+Δs:

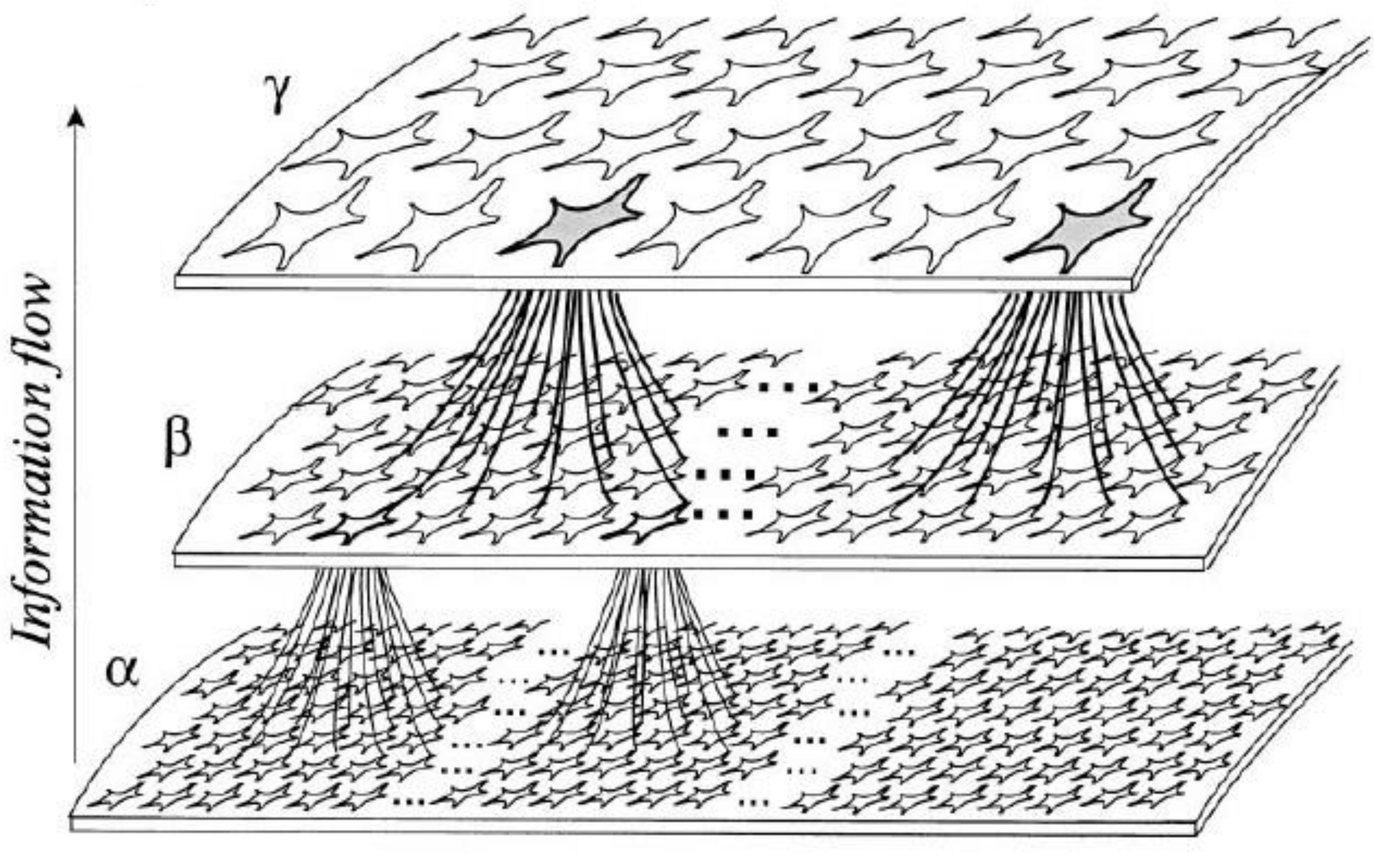
Difference in ML estimate is Δs (unbiased)
variance in estimate is $1/I_F(s)$.

→ $$d' = \Delta s \sqrt{I_{\mathrm{F}}(s)}$$

Limitations of these approaches

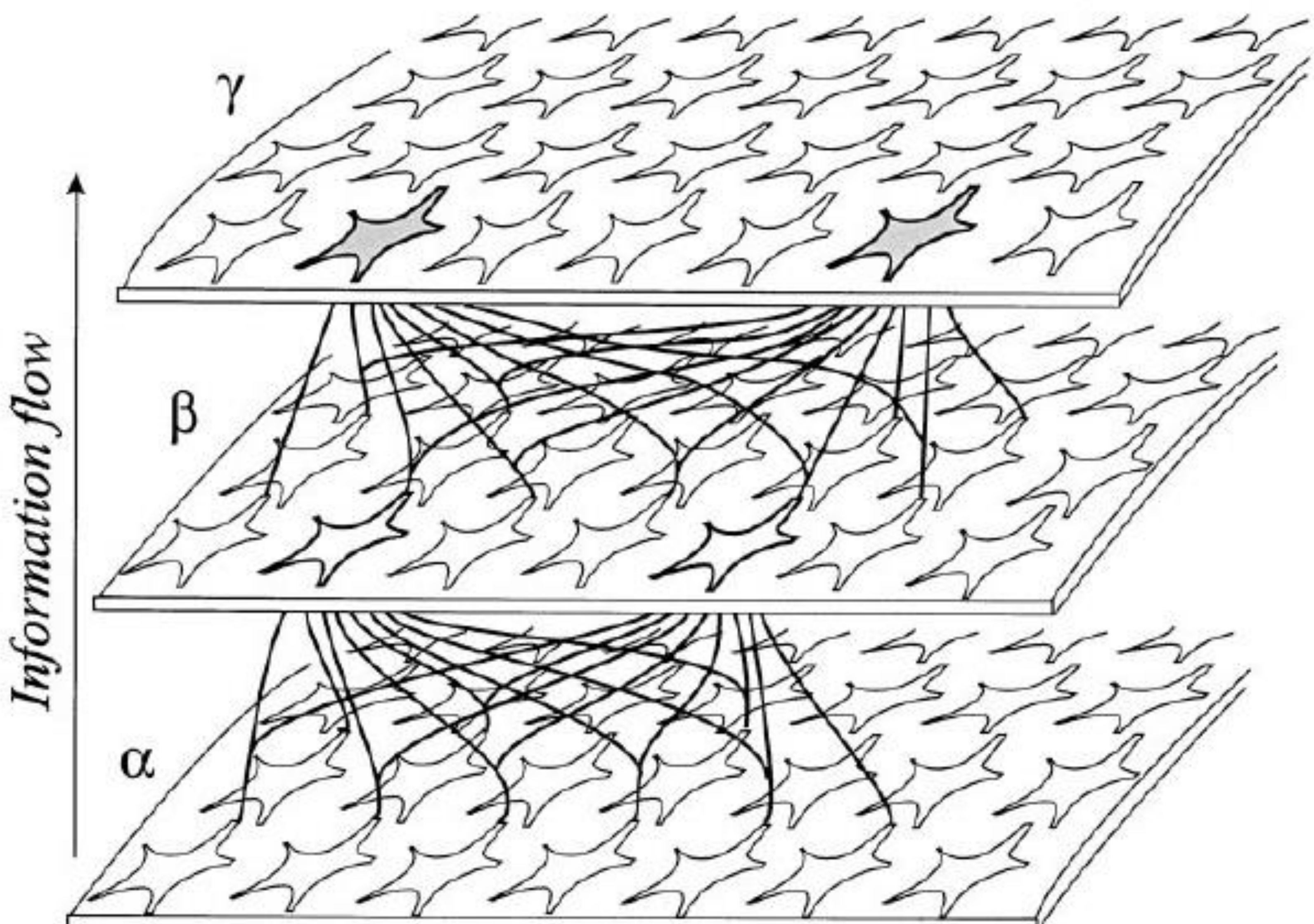
- Tuning curve/mean firing rate
- Correlations in the population

The importance of correlation

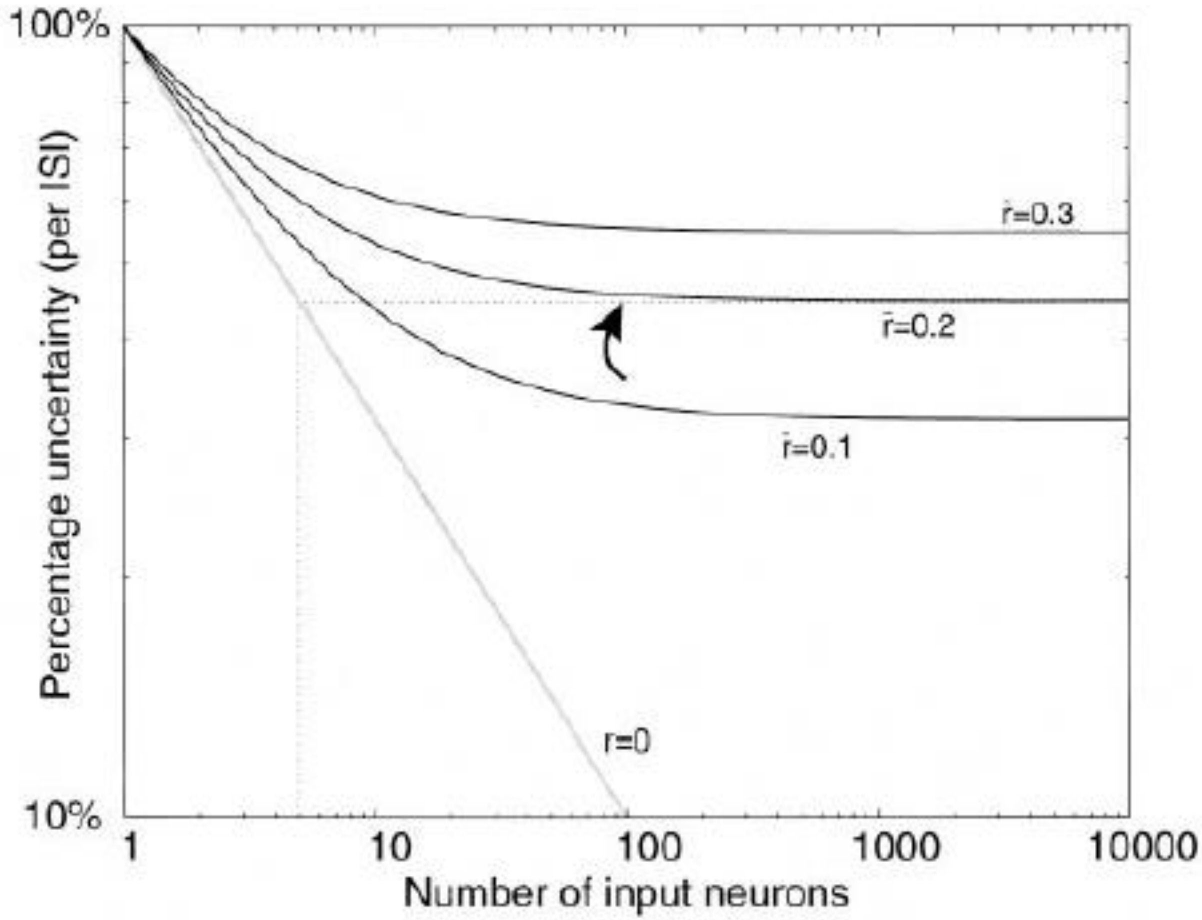


Shadlen and Newsome, '98

The importance of correlation



The importance of correlation



Entropy and Shannon information

Model-based vs model free