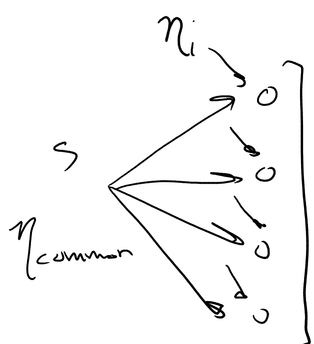
Last time:

1) Encoding & decoding
2) Reverse correlation
3) GLMs
4) Stimulus / noise correlations & pop. coding

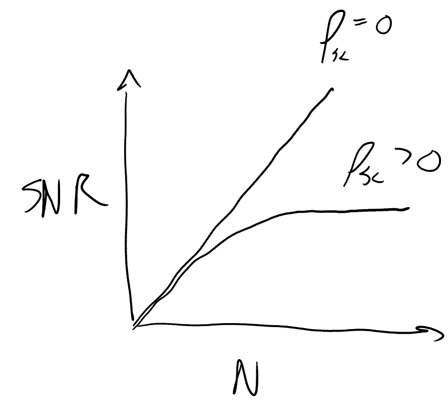
ρ_{sc}

Last time:

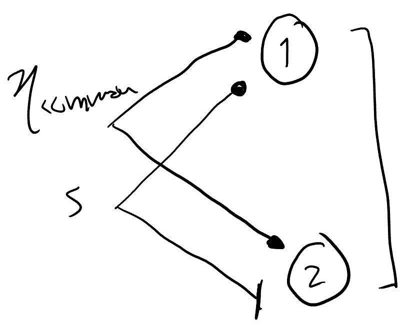


$\Sigma r_i = R$

$$SNR \propto \frac{N^2}{N + N(N-1)\rho_{sc}}$$



Identical tuning

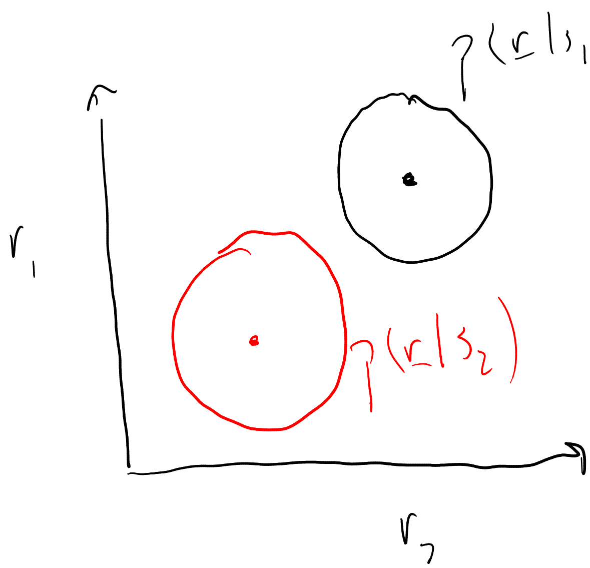


$$r_1 = \theta s + \eta_{common} + b$$

$$r_2 = -\theta s + \eta_{common} + b$$

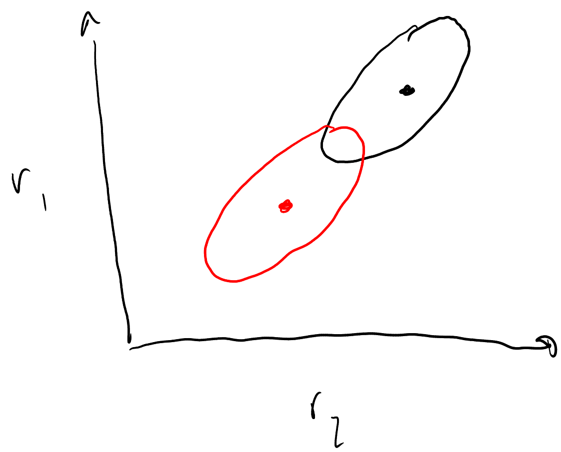
$$R = r_1 - r_2 = 2\theta s$$

Averbeck, Latham, Pouget Nat. Rev. Neurosci 2006



$\rho_{stim} > 0$

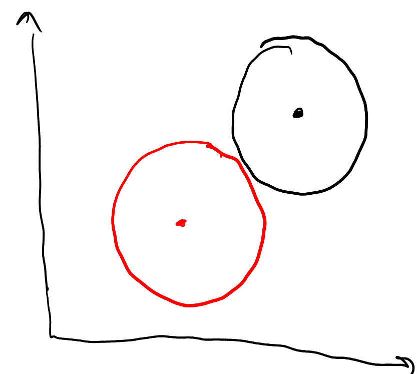
$\rho_{sc} = 0$

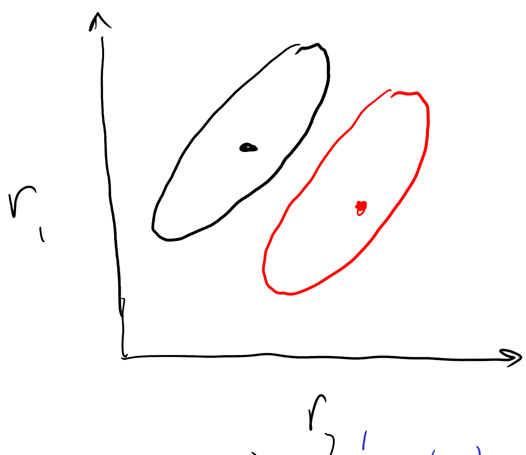


$\rho_{stim} > 0$

$\rho_{sc} > 0$ $\rightarrow$

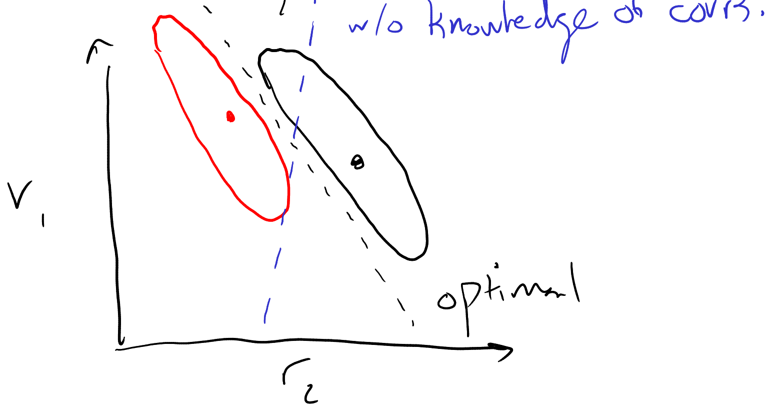
trial-shuffled





$\rho_{stim} < 0$

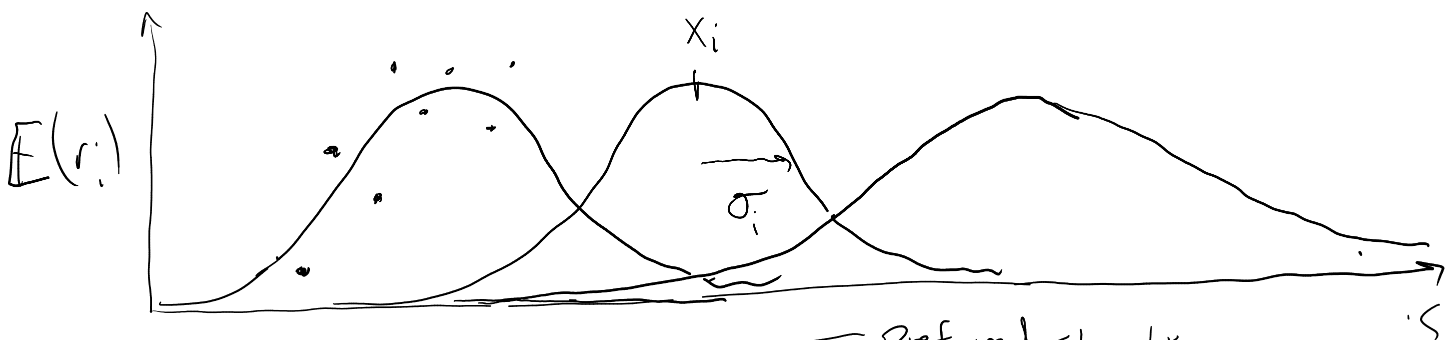
$\rho_{sc} > 0$



Hu, Zylberberg, Shea-Brown PLOS Comp Bio '14:

"Sign rule:" If sign of ρ_{sc} & ρ_{stim} are opposite $\forall$ pairs, Fisher information (& other metrics) are greater than trial-shuffled information.

Population decoding: Example



preferred stimulus

$$\lambda_i = r_{max} \exp\left(\frac{-(s - s_i)^2}{2\sigma_i^2} \right)$$

tuning width

$r_i \sim Pois(\lambda_i)$

If independent,

$$P(\underline{r}|s) = \prod_i P(r_i|s)$$

$$= \prod_i \frac{\lambda_i^{r_i}}{r_i!} e^{-\lambda_i}$$

$$\log P(\underline{r}|s) = \sum_i r_i \log \lambda_i(s) - \log r_i! - \lambda_i(s)$$

Simplifying assumption: $\sum_i \lambda_i(s) \approx$ const.

Maximize $\sum_i r_i \log \lambda_i(s) + C = \sum_i r_i \frac{-(s-s_i)^2}{2\sigma_i^2} + C$

Derivative wrt $s = 0 \Longrightarrow$

$$0 = \sum_i r_i \frac{-\cancel{2}(s-s_i)}{\cancel{2}\sigma_i^2}$$

$$s_{ML} = \sum_i r_i s_i/\sigma_i^2 \Big/ \sum_i r_i/\sigma_i^2$$

Preferred stimulus of neuron i weighted by r_i/σ_i^2

If σ_i constant, reduces to weighted sum

$$s_{ML} = \sum_i r_i s_i \Big/ \sum_i r_i$$

Maximum likelihood estimate.

Bayes rule: $P(s|\underline{r}) = \dfrac{P(\underline{r}|s)\,P(s)}{P(\underline{r})}$

Maximum a posteriori estimate: maximize

$P(s|\underline{r})$ given prior $P(s) \Longleftrightarrow$

Maximize $P(\underline{r}|s)\,P(s)$. If $P(s) = \text{constant}$

(flat prior), MAP estimate = MLE

Example: Gaussian $P(s) = \dfrac{1}{\sqrt{2\pi\sigma_{prior}^2}} \exp\left(\dfrac{-(s - s_{prior})^2}{2\sigma_{prior}^2} \right)$

MAP: $\max \log P(\underline{r}|s)\,P(s)$

$$= \sum_i r_i \frac{-(s-s_i)^2}{2\sigma_i^2} - \frac{(s - s_{prior})^2}{2\sigma_{prior}^2}$$

$$\Longrightarrow s_{MAP} = \frac{\sum_i r_i s_i/\sigma_i^2 + s_{prior}/\sigma_{prior}^2}{\sum_i r_i/\sigma_i^2 + s_{prior}/\sigma_{prior}^2}$$

What about correlated case? Can no longer factorize likelihood.

Bias & variance:

Bias of $\hat{s}$: $b(s) = E[\hat{s} \mid s] - s$

(defined for any parameter θ of a statistical model)

$$\mathrm{Var}(\hat{s} \mid s) = E\{(\hat{s} - E[\hat{s} \mid s])^2 \mid s\}$$

$$E[(\hat{s} - s)^2] = E[(\hat{s} - E[\hat{s} \mid s] + b(s))^2 \mid s]$$

$$= \underbrace{\mathrm{Var}(\hat{s} \mid s)}_{\text{variance}} + \underbrace{b^2(s)}_{\text{bias}}$$

$b(s) = 0$: unbiased estimator

Fisher information (scalar case):

$$I_F(s) = -E\left[\frac{\partial^2 \log p(\underline{r} \mid s)}{\partial s^2}\right]_{p(\underline{r} \mid s)} \qquad (1)$$

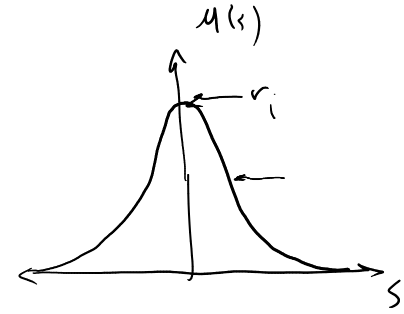
or

$$= -\int d\underline{r}\, p(\underline{r} \mid s) \frac{\partial^2 \log p(\underline{r} \mid s)}{\partial s^2}$$

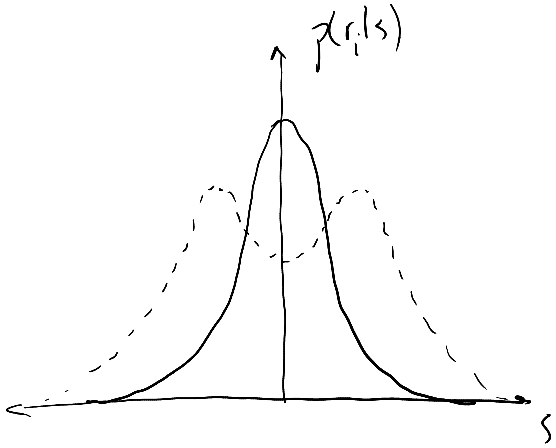
$$E\left[\left(\frac{\partial}{\partial s} \log p(\underline{r} \mid s)\right)^2\right]_{p(\underline{r} \mid s)} \qquad (2)$$

Example: $\lambda_i = \mu(s)$ $\quad r_i \sim \mathrm{Pois}(\lambda_i)$

$\uparrow$ tuning curve



$$p(r_i|s) = \frac{\mu(s)^{r_i} e^{-\mu(s)}}{r_i!}$$



$$\log p(r_i|s) = r_i \log \mu(s) - \mu(s) - \log r_i!$$

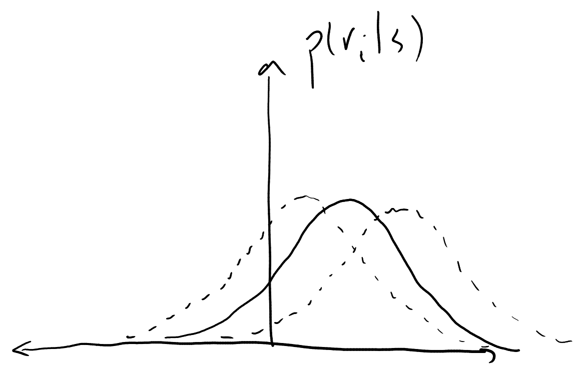
$$\frac{\partial}{\partial s} \log p(r_i|s) = r_i \frac{\mu'(s)}{\mu(s)} - \mu'(s) = \mu'(s)\left[\frac{r_i}{\mu(s)} - 1\right]$$

$$\underset{p(r_i|s)}{E}\left[(\cdot)^2\right] = \left(\mu'(s)\right)^2 \underbrace{E\left[\frac{r_i^2}{\mu^2(s)} - \frac{2r_i}{\mu(s)} + 1\right]}_{\frac{1}{\mu(s)} + 1 - 2 + 1}$$

$E[r_i^2] = \mu(s)(1+\mu(s))$

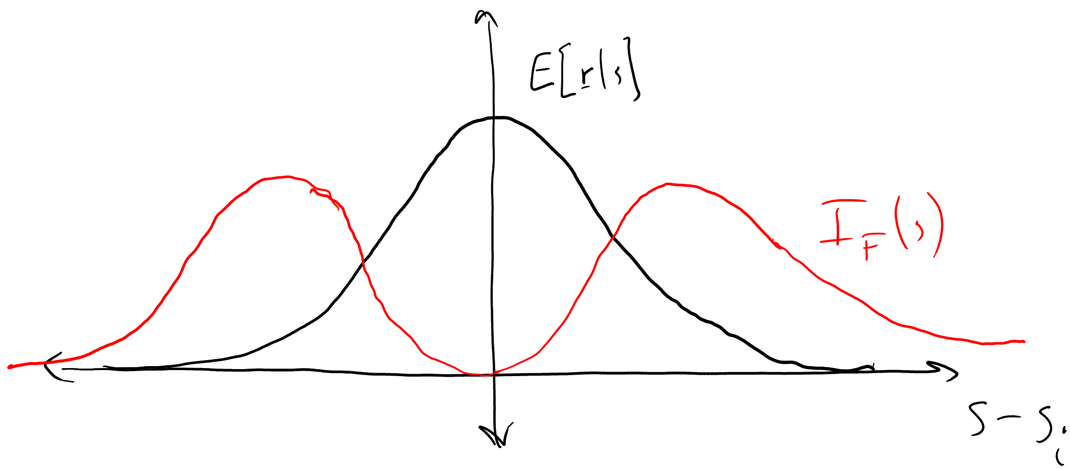
$E[r_i] = \mu(s)$

$$= \frac{\left(\mu'(s)\right)^2}{\mu(s)}$$



$$\lambda_i = r_{max} \exp\left(\frac{-(s_i - s)^2}{2\sigma_i^2} \right) \qquad r_i \sim Pois(\lambda_i)$$

$$I_F(s) = \frac{r_{max}(s_i - s)^2}{\sigma_i^4} \exp\left(\frac{-(s_i - s)^2}{2\sigma_i^2} \right)$$



Intuition (using def 1): Expected curvature of log-likelihood fn.

Intuition (using def. 2): "Score": $\frac{\partial}{\partial s} \log p(\underline{r}|s)$

How much does log-likelihood of observing $\underline{r}$ change when s is varied?

$$E\left[\frac{\partial}{\partial s} \log p(\underline{r}|s)\right]_{p(\underline{r}|s)} = \int d\underline{r}\, \cancel{p(\underline{r}|s)} \cdot \frac{\frac{\partial}{\partial s} p(\underline{r}|s)}{\cancel{p(\underline{r}|s)}}$$

$$= \frac{\partial}{\partial s} \int d\underline{r}\, p(\underline{r}|s) = \frac{\partial}{\partial s}(1) = 0.$$

So $Var(score) = E[score^2] = I_F(s)$

Properties: 1) Local (dependent on value of s)

2) If r_i, r_j independent, $I_F(s) = I_F^i(s) + I_F^j(s)$,

$$\log p(\underline{r}|s) = \log p(r_i|s) + \log p(r_j|s)$$

3) More generally, $I_F^{X,Y} = I_F^X + I_F^{Y|X}$

4) Dependent on stimulus parameterization:

If $u = f(s)$, $I_F(s) = I_F(u)\left(\frac{du}{ds}\right)^2$

5) Related to variance of unbiased estimator (Cramér-Rao bound, later).

Equivalence of 2 definitions:

$$\frac{\partial^2}{\partial s^2}\log p(\underline{r}|s) = \frac{\partial}{\partial s}\left[\frac{\partial}{\partial s}\log p(\underline{r}|s)\right]$$

$$= \frac{\partial}{\partial s}\left[\frac{1}{p(\underline{r}|s)}\frac{\partial}{\partial s}p(\underline{r}|s)\right]$$

$$= \underbrace{\frac{1}{p(\underline{r}|s)}\frac{\partial^2}{\partial s^2}p(\underline{r}|s)} - \underbrace{\left(\frac{\frac{\partial}{\partial s}p(\underline{r}|s)}{p(\underline{r}|s)}\right)^2}$$

$$E[\cdot] = \int d\underline{r}\,\frac{\partial^2}{\partial s^2}p(\underline{r}|s) = \frac{\partial^2}{\partial s^2}\int d\underline{r}\,p(\underline{r}|s) = 0$$

$$= \left[\frac{\partial}{\partial s}\log p(\underline{r}|s)\right]^2$$

So $$I_F(s) = E\left[\left(\frac{\partial}{\partial s}\log p(\underline{r}|s)\right)^2\right]_{p(\underline{r}|s)} \qquad (2)$$