Last time: 1) Random rate network
2) SVD
3) Tensor decomposition

Notation

$P(s)$ (discrete) or $p(s)$ (continuous) (prior)

$P(\underline{r}, s)$ joint distribution

$P(\underline{r}|s)$ conditional firing rate distribution (encoding)

$P(s|\underline{r})$ posterior (decoding)

$$P(\underline{r}, s) = \sum_s P(\underline{r}|s) P(s) = \sum_{\underline{r}} P(s|\underline{r}) P(\underline{r})$$

Bayes Thm: $$P(s|\underline{r}) = \frac{P(\underline{r}|s) P(s)}{P(\underline{r})}$$

$s \sim N(\mu, \sigma^2)$ "s distributed according to normal dist. w/ parameters μ, σ^2."

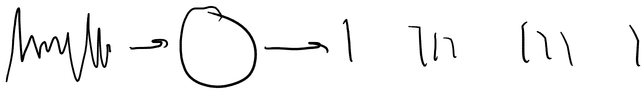
$$p(s) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}$$

$$E[s] = \sum_i s_i P(s_i) \quad \text{or} \quad \int ds\, s\, p(s)$$

$$\mathrm{Cov}(x, y) = E[(x - E[x])(y - E[y])]$$

$$\mathrm{Var}(x) = \mathrm{Cov}(x, x)$$

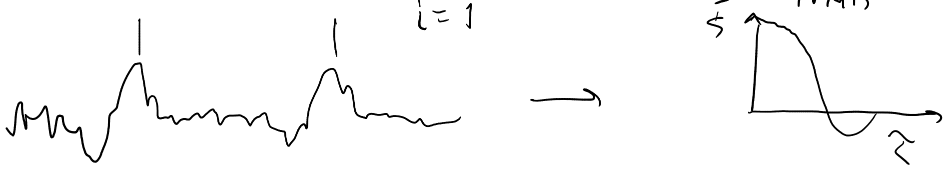
Estimating receptive fields



1) Spike-triggered average
2) Reverse correlation } Dayan & Abbott Ch. 1, 2
3) Max. likelihood.

1) Let $s(t)$ be stimulus and $\{t_i\}$ be spike times

$$\text{STA}\quad C(\tau) = E\left[\frac{1}{n}\sum_{i=1}^{n} s(t_i - \tau)\right]$$



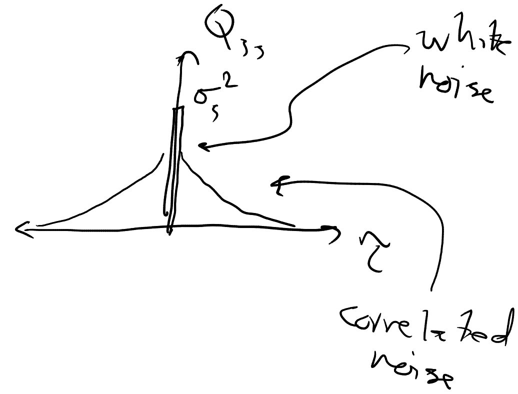
2) Let $r(t)$ be avg firing rate over trials, rate-stim. correlation fn. is:

$$Q_{rs}(\tau) = \frac{1}{T}\int_0^T dt\, r(t)\, s(t+\tau) = E[r]\, C(-\tau)$$

(replace sum in STA w/ integral of δ-functions)

Stim-stim. correlation fn. is:

$$Q_{ss}(\tau) = \frac{1}{T}\int_0^T dt\, s(t)\, s(t+\tau)$$



Predicted neural activity (linear):

$$r_{est} = r_0 + \int_0^\infty d\tau\, D(\tau)\, s(t-\tau)\,.\quad \text{What is } D(\tau)?$$

Minimize $E = \frac{1}{T}\int_0^t dt \, \left(r_{est}(t) - r(t)\right)^2$

Functional derivative wrt $D = 0 \implies$

$$\underbrace{\int_0^\infty d\tau' \, Q_{ss}(\tau - \tau') \, D(\tau')} = Q_{rs}(-\tau)$$

"reverse correlation"

If white noise, only $\tau = \tau'$ contributes

$= \sigma_s^2 \, D(\tau)$

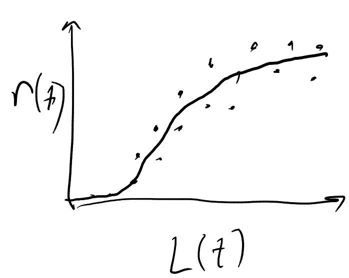
$$\implies D(\tau) = \frac{1}{\sigma_s^2} Q_{rs}(-\tau) = \frac{E[r]}{\sigma_s^2} \overset{\text{STA}}{C(\tau)}$$

Best linear estimate of r.

What if $r_{est} = r_0 + F(L(t))$, $L(t) = \int_0^t d\tau \, D(\tau) \, s(t-\tau)$?

Bussgang's theorem: If s is white Gaussian noise, the above $D(\tau)$ is still optimal.

What is F ?



Reverse correlation:

1) Apply Gaussian white noise stim.
2) Calculate optimal linear filter
3) Determine nonlinearity

Generalized linear models & MLE (J. Pillow, L. Paninski)

Linear Poisson neuron: $\underset{\text{rate}}{\lambda} = \theta s$, $\underset{\text{spk count}}{y} \sim Pois(\lambda)$

Recall Poisson distribution, $p(x|\lambda) = \frac{1}{x!}\lambda^x e^{-\lambda}$

$$P_\theta(x|s) = \frac{1}{x!}(\theta s)^x e^{-(\theta s)}$$

Suppose we have data X, S, $X = \{x_i\}$, $S = \{s_i\}$

$$p_\theta(X|S) = \prod_i p_\theta(x_i|s_i) \quad \text{(conditional independence)}$$

Find θ to max $p_\theta(X|S) \iff \max \log p_\theta(X|S)$.

$$\log p_\theta(X|S) = \log \prod_i p_\theta(x_i|s_i) = \sum_i \log p_\theta(x_i|s_i).$$

$$\log p_\theta(x_i|s_i) = \underbrace{-\log(x_i!)}_{} + x_i(\log\theta + \underline{\log s_i}) - \theta s_i$$

don't depend on θ.

$$\log p_\theta(X|S) = \sum_i x_i \log\theta - \theta s_i + C$$

$$= \log\theta\left[\sum_i x_i\right] - \theta\left[\sum_i s_i\right] + C$$

To max. wrt θ, $\frac{d}{d\theta} P_\theta(X|S) = 0$

$$\Rightarrow 0 = \frac{1}{\theta} \sum_i x_i - \sum_i s_i$$

$$\Rightarrow \theta_{ML} = \sum_i x_i \Big/ \sum_i s_i$$

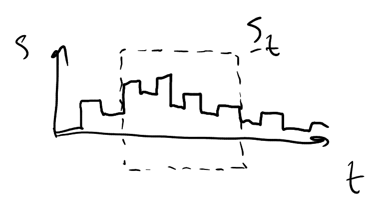
GLMs:

Linear-nonlinear-Poisson model:

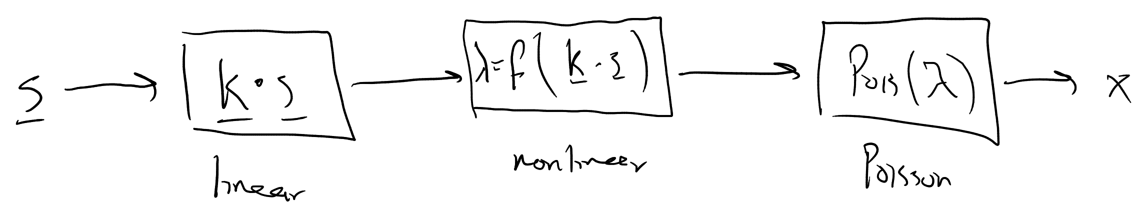
$$\lambda = f(\underline{k} \cdot \underline{s}), \qquad x = \text{Pois}(\lambda).$$

f: nonlinearity; $\underline{k}$: linear kernel; $\underline{s}$: stimulus

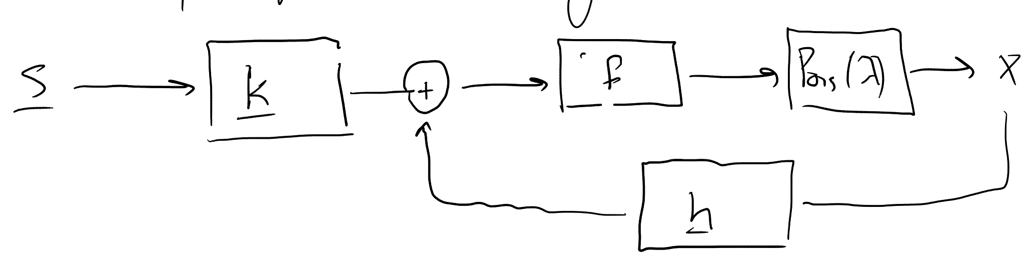
$\underline{s}$ can be vector of times & stim. dimensions.



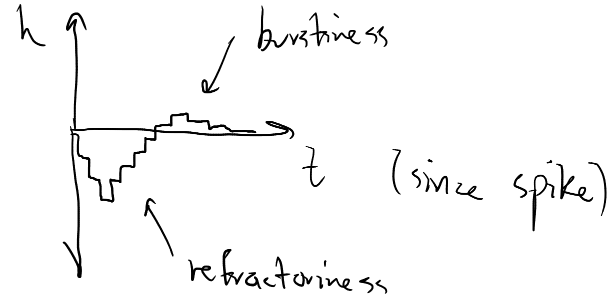
HW: Compute θ_{ML} for LNP model.



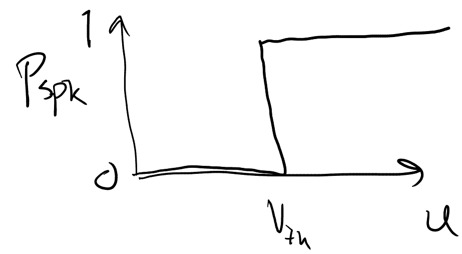
Spike-history dependence (bursting, refractoriness):

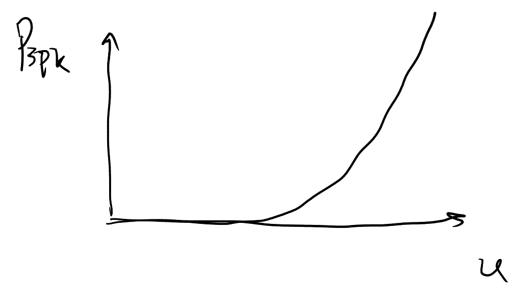


$\underline{h}$: post-spike filter:



Typically, $f(u)=e^u$. Can be thought as "soft-threshold" integrate & fire neuron:





Multiple GLMs can be coupled to infer $\underline{k}_i$, $\underline{h}_i$, and $\underline{W}_{ij}$ simultaneously. Inferred coupling $\neq$ synaptic coupling.

Combining (Macke, Büsing, Cunningham, Yu, Shenoy, Sahani 2011)

Ex: <u>Poisson LDS</u>:

$$\underline{z}_t \sim A\underline{z}_{t-1} + \underline{\eta}_t$$

$$x_{t,i} \sim \text{Pois}\left(\exp\left(C\underline{z}_t + H\underline{x}^{hist} + b\right)_i\right)$$

Population coding, correlations

Spike count correlations:

$$\rho = \frac{\mathrm{Cov}(r_1, r_2)}{\sqrt{\mathrm{Var}(r_1)\,\mathrm{Var}(r_2)}}$$

Stimulus correlation:

Let $\bar{r}_1(s) = E[\,r_1 | s\,]$

$$\rho_{stim} = \frac{\mathrm{Cov}_s(\bar{r}_1(s), \bar{r}_2(s))}{\sqrt{\mathrm{Var}_s(\bar{r}_1(s))\,\mathrm{Var}_s(\bar{r}_2(s))}}$$

"similarity in tuning"

Noise correlation:

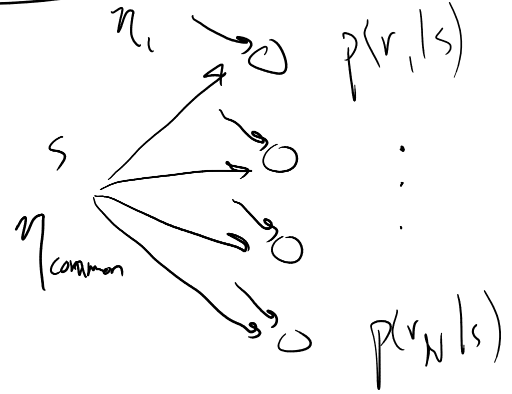
$$\rho_{sc}(s) = \frac{\mathrm{Cov}(r_1, r_2 | s)}{\sqrt{\mathrm{Var}(r_1 | s)\,\mathrm{Var}(r_2 | s)}}$$

"trial-to-trial covariability"

Often assumed $\rho_{sc} = E_s[\rho_{sc}(s)]$.

Dependence on s: "stimulus-dependent noise correlations"

Implications for stimulus decoding (Zohary, Shadlen, Newsome 1998)



$$p(r_i|s) = p(r|s)$$

$$R = \sum_i r_i$$

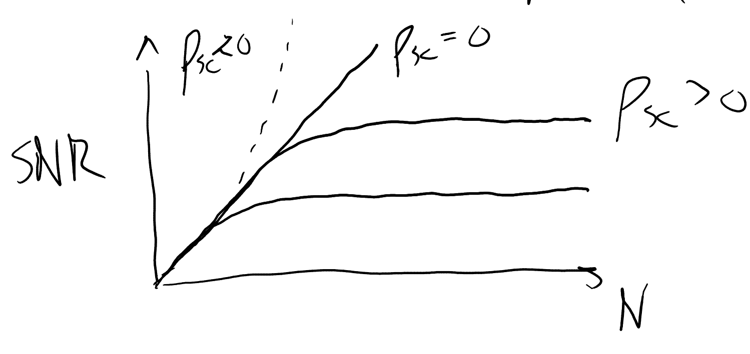
$$E[R|s] = N\,E[r|s] \qquad \text{"Signal"}$$

$$Var(R|s) = Var\left(\sum_i r_i \Big| s\right) = E\Big[\sum_i (r_i - E[r_i])^2 + \underbrace{\sum_{i \neq j} (r_i - E[r_i])(r_j - E[r_j])\Big]}_{\text{conditioned on } s}$$

$$= N\,Var(r|s) + N(N-1)\,Cov(r_i, r_j|s)$$

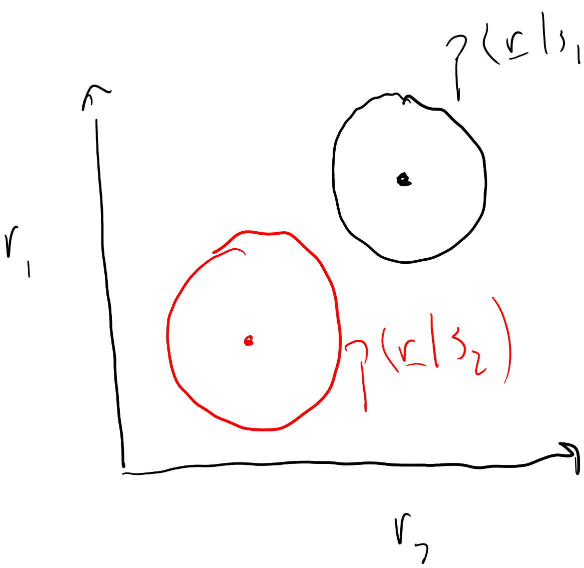
$$= Var(r|s)\left[N + N(N-1)\rho_{sc}\right]$$

$$SNR = \frac{E^2[r|s]}{Var(r|s)} \cdot \frac{N^2}{N + N(N-1)\rho_{sc}}$$



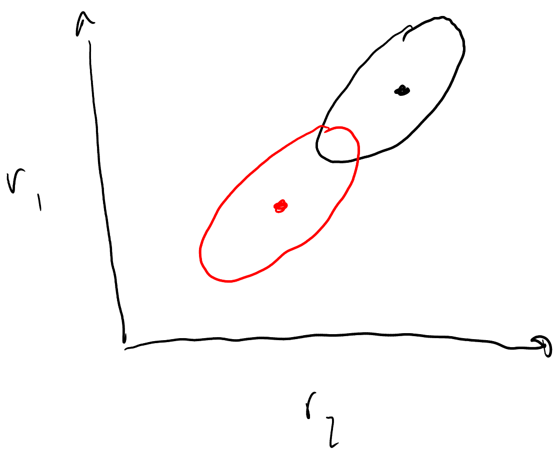
$\rho_{sc} \sim 1/N$ detrimental

Averbeck, Latham, Pouget Nat. Rev. Neurosci 2006



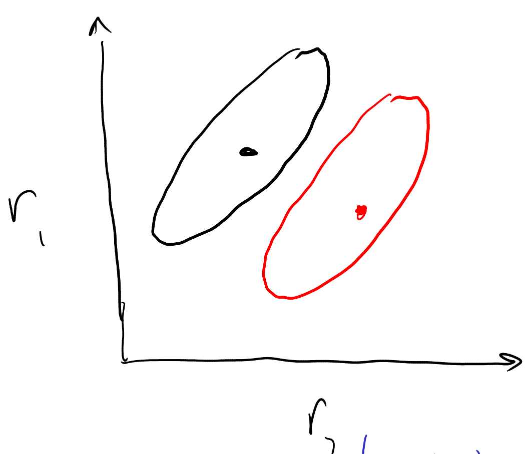
$\rho_{stim} > 0$

$\rho_{sc} = 0$



$\rho_{stim} > 0$

$\rho_{sc} > 0$



$\rho_{stim} < 0$

$\rho_{sc} > 0$