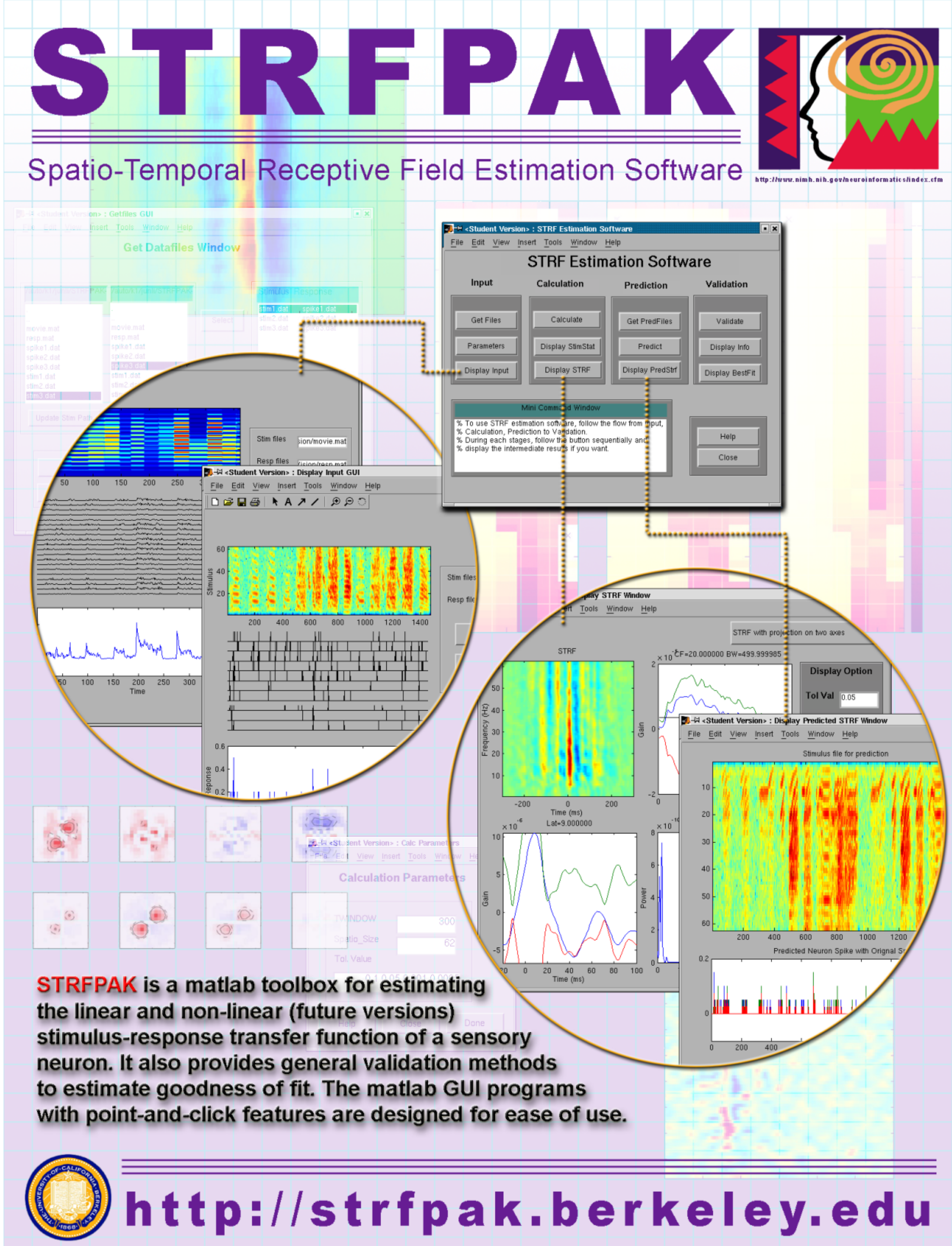
Neural Data Analysis

Tutorial 2: Linear Model. Direct Fit

Rev-Correlations for Nat Stats

- STRFPAK.
  - Non stationary spatial stimulus correlations
  - Regularization in eigen-space
  - Regularization in joint eigen/pixel space
  - Prediction Validation
- Preprocessing for auditory neurons.
  - Compression
  - Spectrograms vs Wavelets and Time-frequency scale
  - Adaptive gain control



STRF
UCBSTRF

Reducing the dimensions: The Classical STRF Model

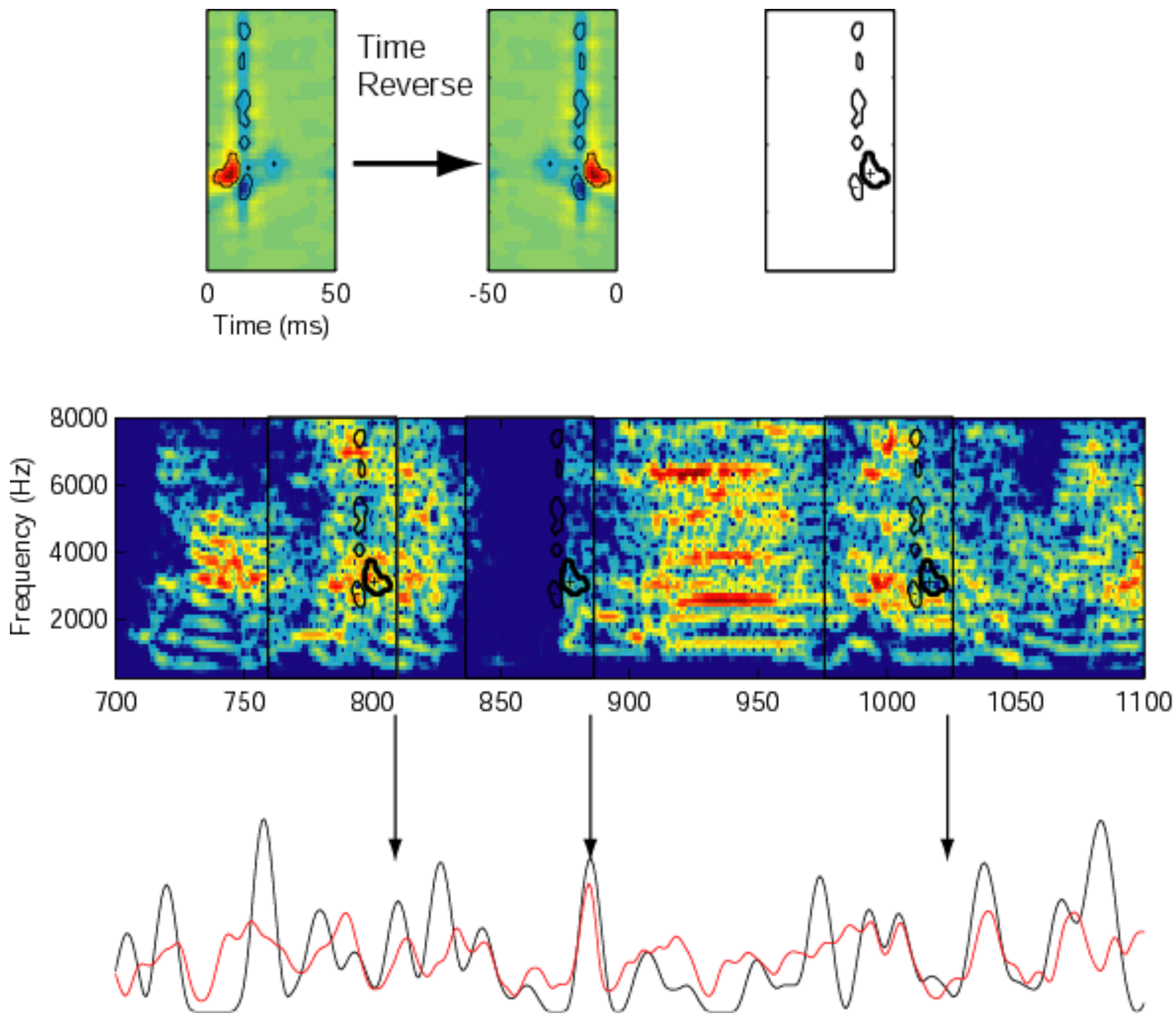
- Encoding Analysis: Linear Model

$$\overbrace{\overline{\vec{r}(t)}}^{Est} = f(\vec{s}(t)) \Rightarrow \overbrace{\overline{\vec{r}(t)}}^{Est} = \vec{h} \cdot \vec{\phi}(\vec{s}(t))$$

- Mutual information: Gaussian Noise

$$I(s,r) = H(r) - H(r \mid s) \quad \Rightarrow \quad p(r(t)) = \frac{1}{\sqrt{2\pi n}} e^{-\frac{1}{2}\frac{(r(t)-\overline{r(t)})^2}{n^2}}$$

Spectro-temporal Receptive Fields of Auditory Neurons



Receptive Fields and Linear Filters in Sensory Neuroscience

- Vision: Spatio-Temporal
  - Hubel & Wiesel
  - deValois
  - Shapley, Ringach, Furster
  - Freeman
  - Gallant, Dan.
  - Pillow
- Audition: Spectro-Temporal
  - De Boer, Eggermont, Aertsen, Young
  - Shamma, Schreiner
  - Linden, Sahani
  - Theunissen

Mathematics of linear STRF Estimation – analytical solution

- White-noise analysis.
  - Revcor. Spike-triggered average (STA).
- Non-white analysis.
  - Linear algebra problem.
  - Stationary. Fourier Transform.
  - LTI

Mathematics of STRF Estimation

- Maximum Likelihood and LMSE (Least Mean Square Error)
- 1-d
- N-d
- Time
- Time/space
- Maximum a posterior (MAP) estimates and Regularization. Ridge Regression.
- Validation

Likelihood Function: Probability of observing data given model

$$L = \prod_i p(\vec{r}_i \mid f(\vec{s}_i))$$

i trials/time=data points
f() encoding model

Maximizing the likelihood = Minimizing the negative log likelihood

$$-\log(L) = -\sum_i \log\left(p(\vec{r}_i \mid f(\vec{s}_i))\right)$$

Likelihood between 0 and 1 but has a ceiling value <= 1 for a saturated model

Negative log Likelihood >= 0 but has a minimum value

$$p(\vec{r}_i \mid f(\vec{s}_i)) = p(\vec{r}_i \mid \vec{h}, \vec{s}_i)$$

$$\hat{\vec{r}}_i = \vec{h} \cdot \vec{s}_i$$ and $$\vec{r}_i = \vec{h} \cdot \vec{s}_i + \vec{n}_i$$

$$p(\vec{r}_i \mid \vec{h}, \vec{s}_i) = N(\vec{h} \cdot \vec{s}_i, n) = \frac{1}{\sqrt{2\pi n}} e^{-\frac{1}{2n^2}(r_i - \vec{h} \cdot \vec{s}_i)^2}$$

$$-\log(L) = \sum_i \left(\log\left(\sqrt{2\pi n}\right) + \frac{1}{2n^2}(r_i - \vec{h} \cdot \vec{s}_i)^2 \right)$$

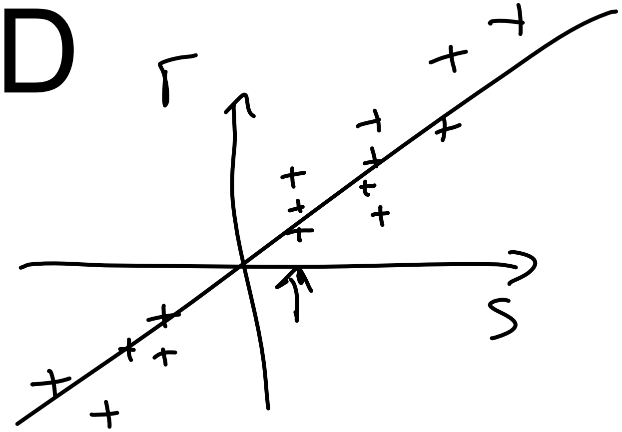
Minimizing $-\log(L)$ ⟷ Minimizing $\sum_i \left((r_i - \vec{h} \cdot \vec{s}_i)^2 \right)$

Maximizing Likelihood ⟷ Minimizing Mean Square Error

1-D

1-D

1-D



s = Stimulus

r = Response

$\hat{r}$ = Predicted Response

$\hat{r} = h\,s$ $\quad$ h = parameters of model = linear filter

$C = \langle (\hat{r} - r)^2 \rangle$ $\quad$ $\langle\ \rangle$ = average $r_i \,, s_i$

$i = 1, \dots N$

minimum

$\langle r \rangle = \frac{1}{N} \sum_{i=1}^{N} r_i$

Noise = $p(r|s) = \mathcal{N}(\bar{r}_s, \sigma_s) = \frac{1}{\sqrt{2\pi\sigma_s^2}} e^{-\frac{1}{2}\frac{(r-\bar{r}_s)^2}{\sigma_s^2}}$

1-D

$$\langle (\hat{r} - r)^2 \rangle \to \min$$

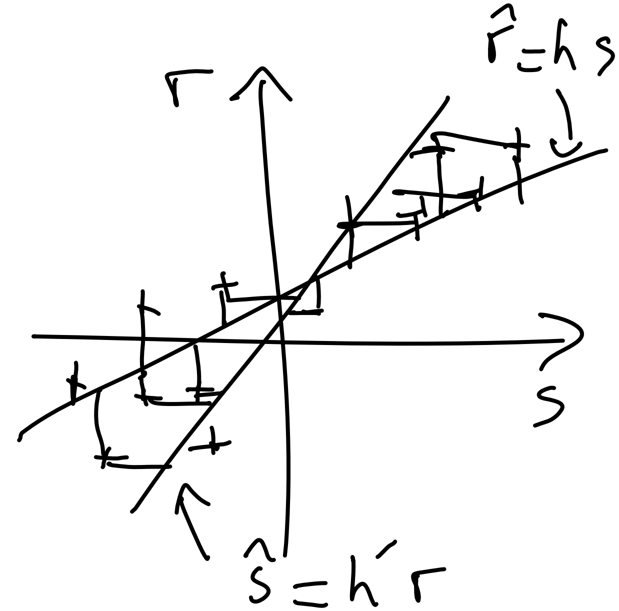
$$\hat{r} = hs$$

$$\frac{d}{dh}\langle (hs - r)^2 \rangle = 0$$

$$\langle 2(hs - r)s \rangle = 0$$

$$h\langle s^2 \rangle = \langle rs \rangle$$

$$\boxed{h = \frac{\langle rs \rangle}{\langle s^2 \rangle}}$$



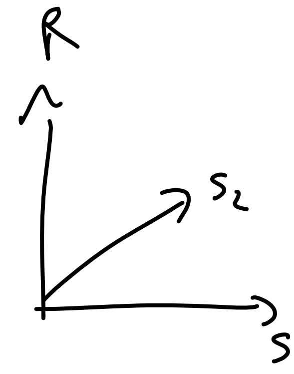
$$h' = \frac{\langle rs \rangle}{\langle rr \rangle}$$

$$h \neq \frac{1}{h'}$$

2-D

2-D

2-D



r

$$(s_1, s_2) = \vec{s}$$

$$(h_1, h_2) = \vec{h}$$

$$\hat{r} = h_1 s_1 + h_2 s_2$$

$$\langle (\hat{r} - r)^2 \rangle \quad min$$

$$\begin{cases} \dfrac{d}{dh_1} \langle (\hat{r} - r)^2 \rangle = 0 \\[2ex] \dfrac{d}{dh_2} \langle (\hat{r} - r)^2 \rangle = 0 \end{cases} \Longleftrightarrow \begin{cases} h_1 \langle s_1 s_1 \rangle + h_2 \langle s_1 s_2 \rangle = \langle r s_1 \rangle \\[2ex] h_1 \langle s_2 s_1 \rangle + h_2 \langle s_2 s_2 \rangle = \langle r s_2 \rangle \end{cases}$$

$$C_{ss} = \begin{pmatrix} \langle s_1 s_1 \rangle & \langle s_1 s_2 \rangle \\ \langle s_2 s_1 \rangle & \langle s_2 s_2 \rangle \end{pmatrix} \qquad \vec{C}_{sr} = \begin{pmatrix} \langle r s_1 \rangle \\ \langle r s_2 \rangle \end{pmatrix}$$

$$C_{ss} \vec{h} = C_{sr} \qquad \boxed{\vec{h} = \frac{\vec{C}_{sR}}{C_{ss}}}$$

2-D

$$C_{ss}\vec{h} = \vec{C}_{sr}$$

$$\vec{h} = C_{ss}^{-1}\vec{C}_{sr}$$

Eigenvectors

$$C_{ss} = Q_{ss}^{-1}\Lambda_{ss}Q_{ss}$$

Eigenvalues

$$\Lambda = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & 0 \\ 0 & & \lambda_3 & \\ & & & \ddots & \\ & & & & \lambda_n \end{pmatrix}$$

$$\boxed{Q_{ss}\vec{h} = \frac{Q_{ss}\vec{C}_{sr}}{\Lambda_{ss}}}$$

The Normal Equation

$$\vec{h} = \frac{\vec{C}_{sr}}{\mathbf{C}_{ss}} = \frac{\langle \vec{s} r \rangle}{\langle \vec{s}\vec{s} \rangle}$$

n = number of data points
k = number of parameters

$$\langle \vec{s} r \rangle = \frac{1}{n} \begin{pmatrix} \sum_{i=1}^{n} s_{i,1} r \\ \sum_{i=1}^{n} s_{i,2} r \\ \vdots \\ \sum_{i=1}^{n} s_{i,k} r \end{pmatrix}$$

$$\langle \vec{s}\vec{s} \rangle = \frac{1}{n} \begin{pmatrix} \sum_{i=1}^{n} s_{i,1} s_{i,1} & \cdots & \sum_{i=1}^{n} s_{i,1} s_{i,k} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} s_{i,k} s_{i,1} & & \sum_{i=1}^{n} s_{i,k} s_{i,k} \end{pmatrix}$$

The Normal Equation

If we write the stimulus as a Matrix, where rows are data points (from 1 to n) and columns are the stimulus dimensions (from 1 to k):

$$\mathbf{S} = \begin{pmatrix} s_{1,1} & s_{1,2} & & s_{1,k} \\ s_{2,1} & s_{2,2} & & s_{2,k} \\ \vdots & & \ddots & \vdots \\ s_{n,1} & s_{n,2} & \cdots & s_{n,k} \end{pmatrix} \quad \text{and} \quad \mathbf{r} = \begin{pmatrix} r_1 \\ r_2 \\ \\ r_n \end{pmatrix} \quad \text{then:}$$

$$\left\langle \vec{s}\vec{s} \right\rangle = \frac{1}{n} \begin{pmatrix} \sum_{i=1}^{n} s_{i,1} s_{i,1} & \cdots & \sum_{i=1}^{n} s_{i,1} s_{i,k} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} s_{i,k} s_{i,1} & & \sum_{i=1}^{n} s_{i,k} s_{i,k} \end{pmatrix} = \frac{1}{n} \mathbf{S}^{\mathbf{T}} \mathbf{S} \qquad \left\langle \vec{s} r \right\rangle = \frac{1}{n} \begin{pmatrix} \sum_{i=1}^{n} s_{i,1} r_i \\ \vdots \\ \sum_{i=1}^{n} s_{i,k} r_i \end{pmatrix} = \frac{1}{n} \mathbf{S}^{\mathbf{T}} \mathbf{r}$$

The Normal Equation

The normal equation in Matrix notation (think Matlab/Python) is then:

$$\vec{h} = \frac{\vec{C}_{sr}}{\mathbf{C}_{ss}} = \frac{\langle \vec{s} r \rangle}{\langle \vec{s}\vec{s} \rangle} = \frac{{}^{1}\!/_{n}\,\mathbf{S}^{\mathrm{T}}\mathbf{r}}{{}^{1}\!/_{n}\,\mathbf{S}^{\mathrm{T}}\mathbf{S}} = \frac{\mathbf{S}^{\mathrm{T}}\mathbf{r}}{\mathbf{S}^{\mathrm{T}}\mathbf{S}}$$

$$\boxed{\vec{h} = \left(\mathbf{S}^{\mathrm{T}}\mathbf{S}\right)^{-1}\mathbf{S}^{\mathrm{T}}\mathbf{r}}$$

Normal Equation in EigenSpace

$$(\mathbf{S}^{\mathbf{T}}\mathbf{S}) = \mathbf{Q} \cdot \Lambda \cdot \mathbf{Q}^{\mathbf{T}}$$

Q is a k by k matrix of eigenvectors in columns

Or from an Singular Value Decomposition of S:

$$\mathbf{S} = \mathbf{U} \cdot \mathbf{W} \cdot \mathbf{V}^{\mathbf{T}}$$

V = k by k input vectors
U = n by k output vectors
W = k by k diagonal singular values (called S in Matlab)

$$(\mathbf{S}^{\mathbf{T}}\mathbf{S}) = \mathbf{V} \cdot \mathbf{W}^{2} \cdot \mathbf{V}^{\mathbf{T}}$$

V = Q
$W^2 = \Lambda$

The Normal Equation in Eigen Space

$$\boxed{\vec{h} = \left(\mathbf{S}^{\mathrm{T}}\mathbf{S}\right)^{-1}\mathbf{S}^{\mathrm{T}}\mathbf{r}}$$

$$\vec{h} = \mathbf{V}^{\mathrm{T}}\left(\Lambda\right)^{\mathbf{-1}}\mathbf{V}\mathbf{S}^{\mathrm{T}}\mathbf{r}$$

Uncorrelate

$$\mathbf{V}\vec{h} = \left(\Lambda\right)^{\mathbf{-1}}\mathbf{V}\mathbf{S}^{\mathrm{T}}\mathbf{r}$$

Whiten

The Normal Equation directly from SVD

$$\boxed{\vec{h} = \left(\mathbf{S}^{\mathrm{T}}\mathbf{S}\right)^{-1}\mathbf{S}^{\mathrm{T}}\mathbf{r}}$$

$$\vec{h} = \mathbf{V}^{\mathrm{T}}\left(\mathbf{W}^{2}\right)^{\mathbf{-1}}\mathbf{V}\mathbf{S}^{\mathrm{T}}\mathbf{r}$$

Time is special

- Causal
- Wide sense stationarity
- Linear algebra can be written in terms of correlations and convolutions.
- The eigenvectors of the correlation function is the discrete Fourier Transform basis.

Time is special

$r(t)$

$s(t-0), s(t-1), s(t-2) \ldots$

$$\hat{r}(t) = h(0)\, s(t-0) + h(1)\, s(t-1) + h(2)\, s(t-2) + \ldots$$

$$\hat{r}(t) = \int_0^\infty h(\tau)\, s(t-\tau)\, d\tau$$

$$C_{ss} \vec{h} = C_{sr} \qquad C_{ss} = \begin{pmatrix} \langle s(t-0)s(t-0)\rangle & \langle s(t-0)s(t-1)\rangle & \ldots \\ \langle s(t-1)s(t-0)\rangle & \langle s(t-1)s(t-1)\rangle & \ldots \\ \vdots & & \end{pmatrix}$$

$$C_{ss} = \begin{pmatrix} a & b & c & d \\ b & a & b & c \\ c & b & a & b \\ d & c & b & a \end{pmatrix}$$

Symetric Toeplitz matrix

Time is special

$$\boxed{Q_{ss} \vec{h}_t = \Lambda_{ss}^{-1} Q_{ss} \vec{C}_{sr_t}}$$

$$\boxed{\vec{h}_t(\omega_t) = \Lambda_{ss}^{-1} \vec{C}_{sr}(\omega_t)}$$

Time-Space

- If space is WSS as well – Fourier domain
- If second order statistics in space are not stationary then mixed: Fourier domain for time and linear algebra in space

Estimating STRFs – Modified Reverse Correlation

Linear Time Invariant Transfer Functions for Non Stationary Spatial Statistics

$$\hat{r}[t] = \sum_{i=0}^{N-1}\sum_{j=0}^{M-1} h[i,j]s[t-i;j] + r_0$$

$$\Downarrow$$

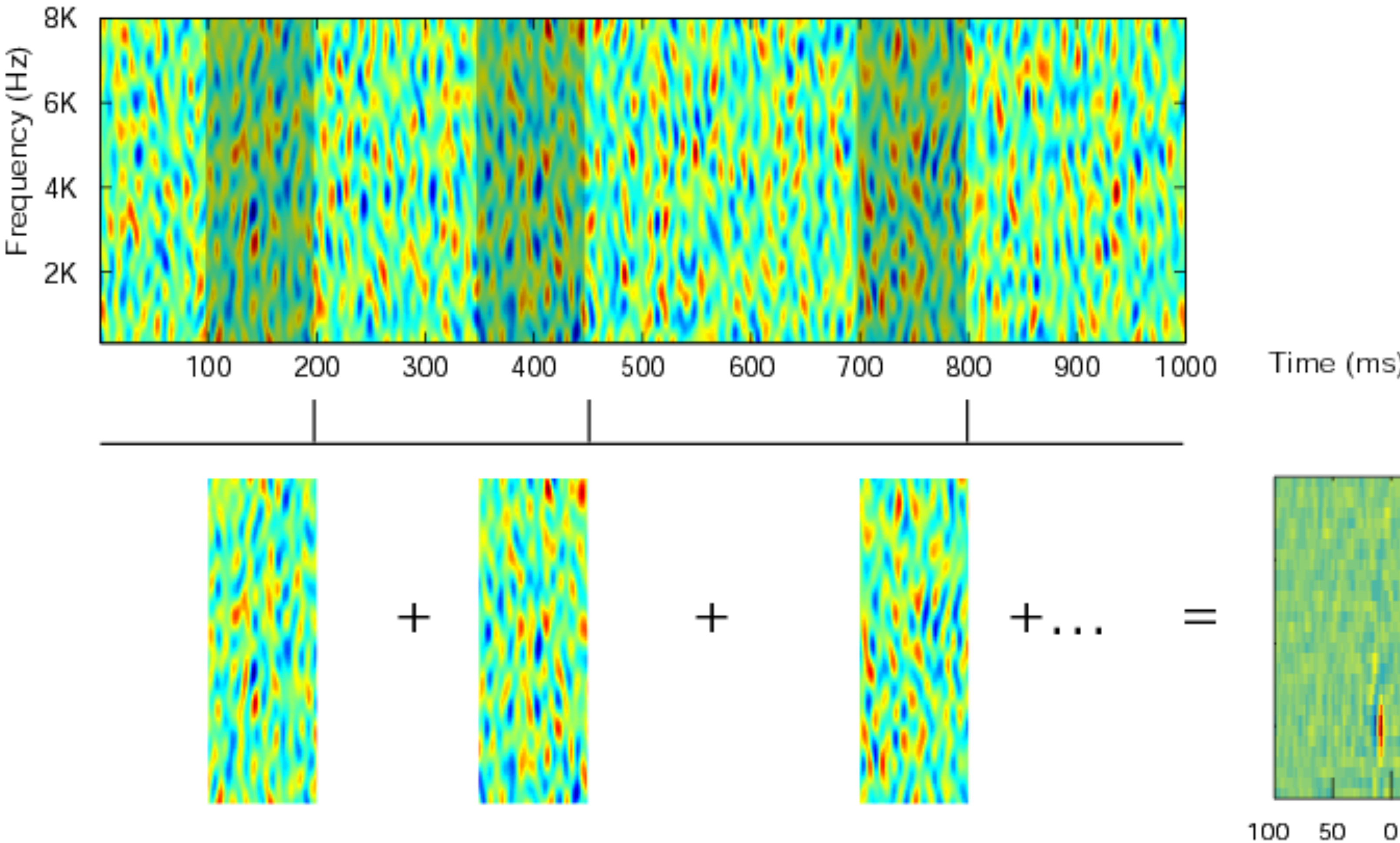
$$\mathbf{k} = \mathbf{C}_{ss}^{-1}\mathbf{C}_{sr} \qquad C_{ss} = \begin{pmatrix} \mathbf{r}_{0,0} & \cdots & \mathbf{r}_{0,M-1} \\ \vdots & \ddots & \vdots \\ \mathbf{r}_{M-1,0} & \cdots & \mathbf{r}_{M-1,M-1} \end{pmatrix}$$

$$\mathbf{r}_{i,j} = \begin{pmatrix} \langle s[t,i]s[t,j]\rangle & \langle s[t,i]s[t-1,j]\rangle & \cdots & \langle s[t,i]s[t-(N-1),j]\rangle \\ \langle s[t-1,i]s[t,j]\rangle & \langle s[t-1,i]s[t-1,j]\rangle & & \vdots \\ \vdots & & \ddots & \\ \langle s[t-(N-1),i]s[t,j]\rangle & \cdots & & \langle s[t-(N-1),i]s[t-(N-1),j]\rangle \end{pmatrix}$$

$$\mathbf{H}\left(\omega_t[i]\right) = \Lambda\left(\omega_t[i]\right)^{-1}\mathbf{C}_{SR}\left(\omega_t[i]\right)$$

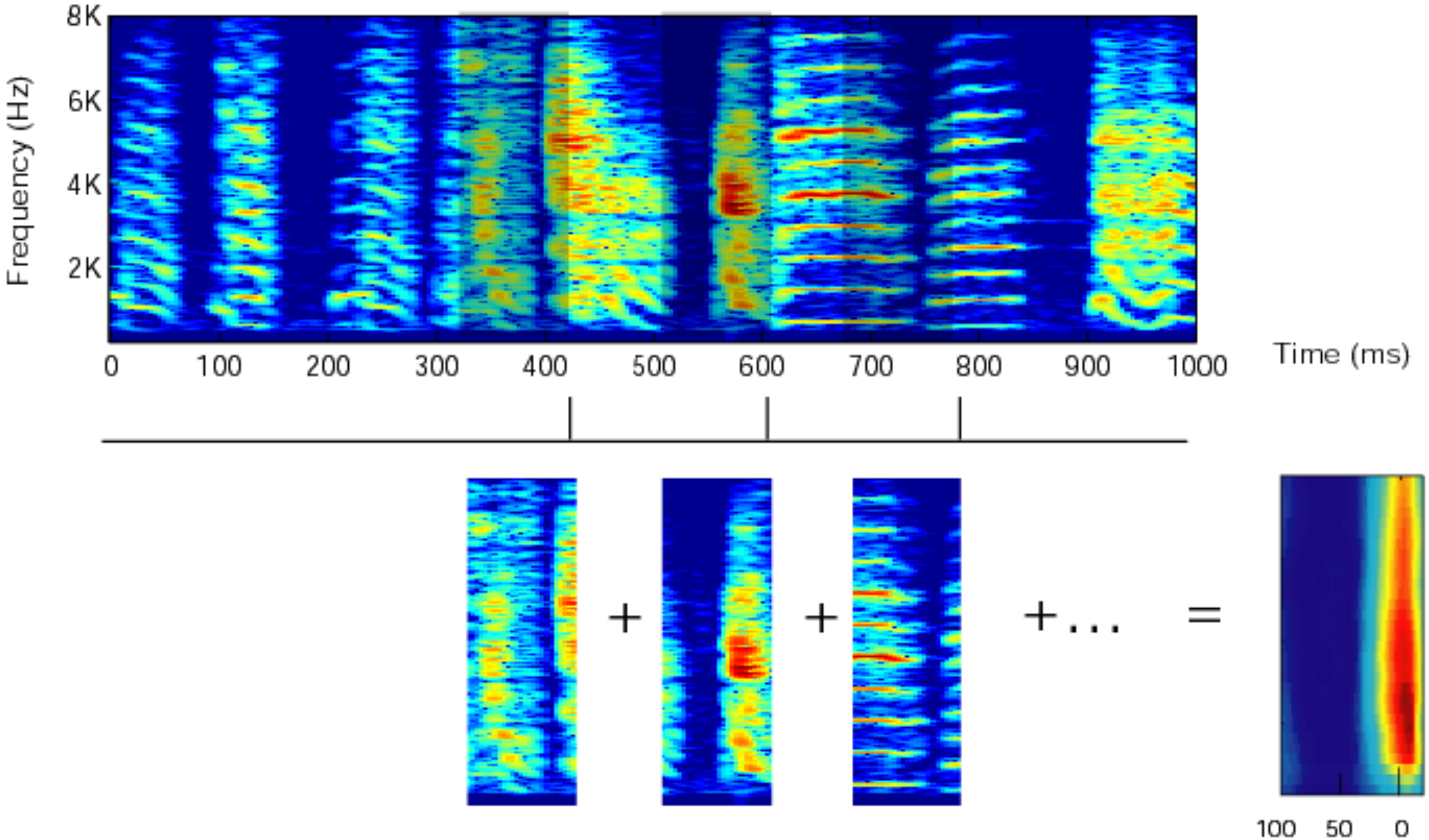
Estimating the STRF

A. Reverse correlation: Spike-triggered average with ripple-noise yields the STRF

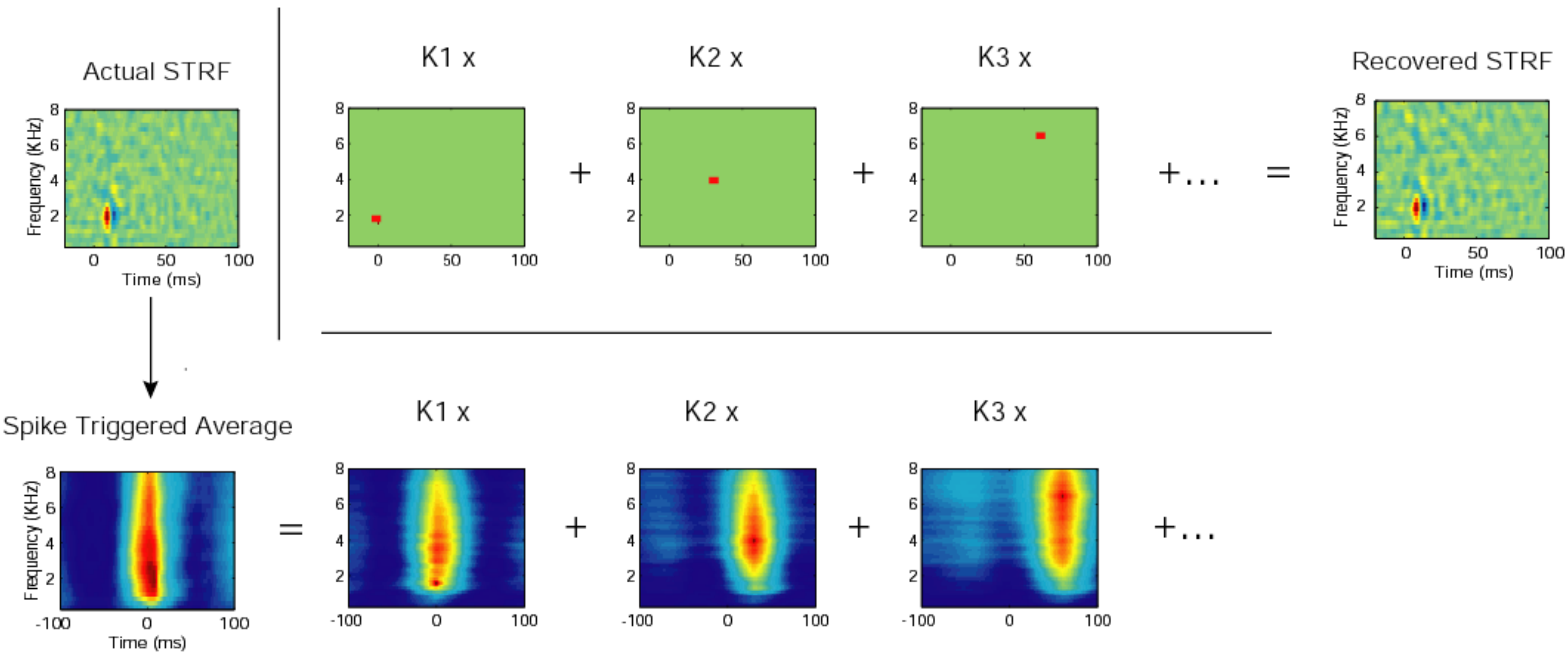


Estimating the STRF

B. Spike-triggered average with song does not recover the STRF



Estimating the STRF



$$\vec{h} \cdot AutoCorrelation = SpikeTriggeredAverage$$

Two point correlation function is translation invariant in time but not in space, Here h=(k1, k2, k3…)

Estimating STRFs – Modified Reverse Correlation

Linear Time Invariant Transfer Functions for Non Stationary Spatial Statistics

$$\hat{r}[t] = \sum_{i=0}^{N-1}\sum_{j=0}^{M-1} h[i,j]s[t-i;j] + r_0$$

$$\Downarrow$$

$$\mathbf{k} = \mathbf{C}_{ss}^{-1}\mathbf{C}_{sr} \qquad C_{ss} = \begin{pmatrix} \mathbf{r}_{0,0} & \cdots & \mathbf{r}_{0,M-1} \\ \vdots & \ddots & \vdots \\ \mathbf{r}_{M-1,0} & \cdots & \mathbf{r}_{M-1,M-1} \end{pmatrix}$$

$$\mathbf{r}_{i,j} = \begin{pmatrix} \langle s[t,i]s[t,j]\rangle & \langle s[t,i]s[t-1,j]\rangle & \cdots & \langle s[t,i]s[t-(N-1),j]\rangle \\ \langle s[t-1,i]s[t,j]\rangle & \langle s[t-1,i]s[t-1,j]\rangle & & \vdots \\ \vdots & & \ddots & \\ \langle s[t-(N-1),i]s[t,j]\rangle & \cdots & & \langle s[t-(N-1),i]s[t-(N-1),j]\rangle \end{pmatrix}$$

$$\mathbf{H}(\omega_t[i]) = \Lambda(\omega_t[i])^{-1}\mathbf{C}_{SR}(\omega_t[i])$$

STRF Estimation

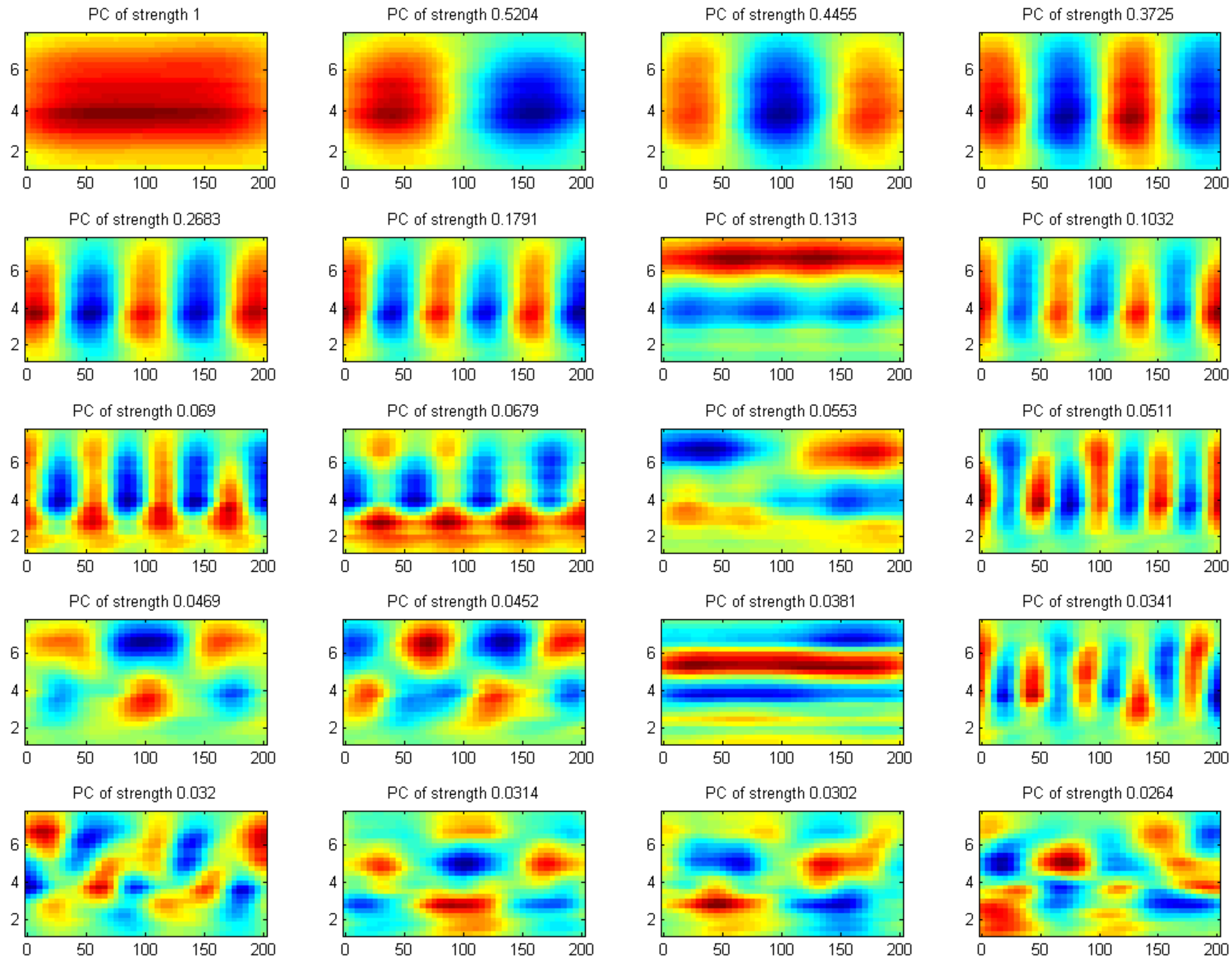
- Calculate STA and Stimulus Covariance Matrix as a function of spatial dimension and time delays.
- Take Fourier Transform of temporal dimension of STA and Stim Cov
- For each temporal frequency solve a set of M linear equations where M is the number of spatial dimensions.
- M equations are not necessarily independent. Perform the inverse in a subspace. Use this step for regularization.
- Take the inverse Fourier Transform of $K(\omega_t)$.

Theunissen, David et al. Network: Comput. Neural Syst. 12 (2001) 289–316

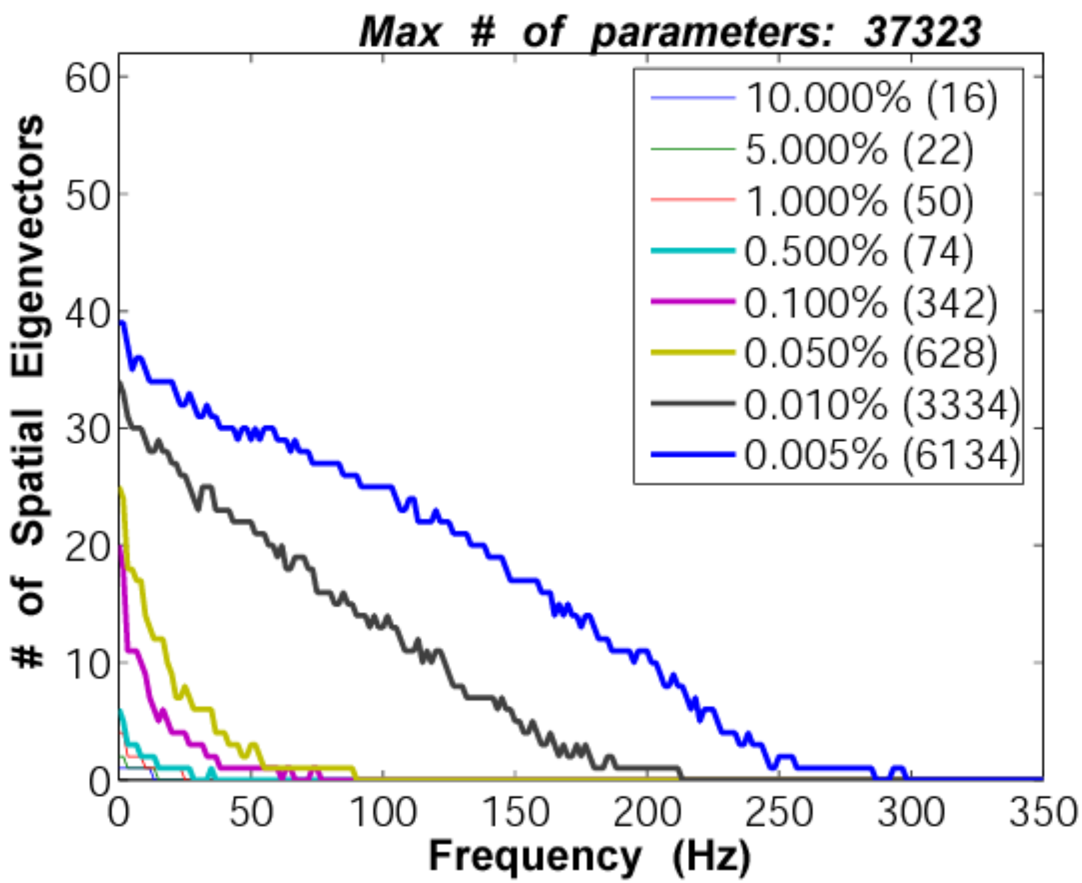
Regularization

- Too many parameters – too little data
- Priors on what the parameters should look like
- Normal Priors give L2 regularization or ridge regression.
- L2 solves the problem into the subspace of the stimulus Principal Components with the largest eigenvalues

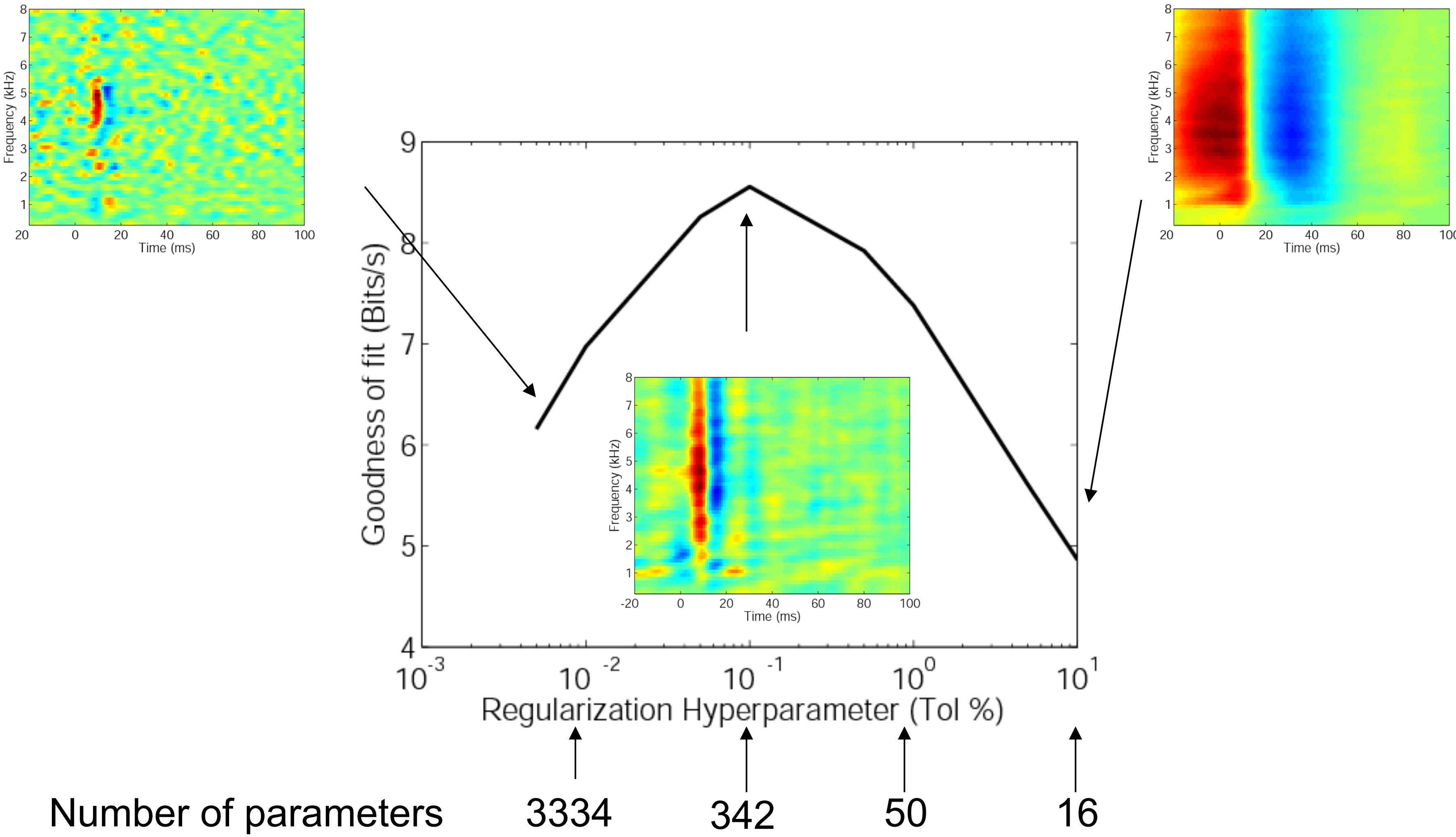
PCAs of Zebra Finch Song



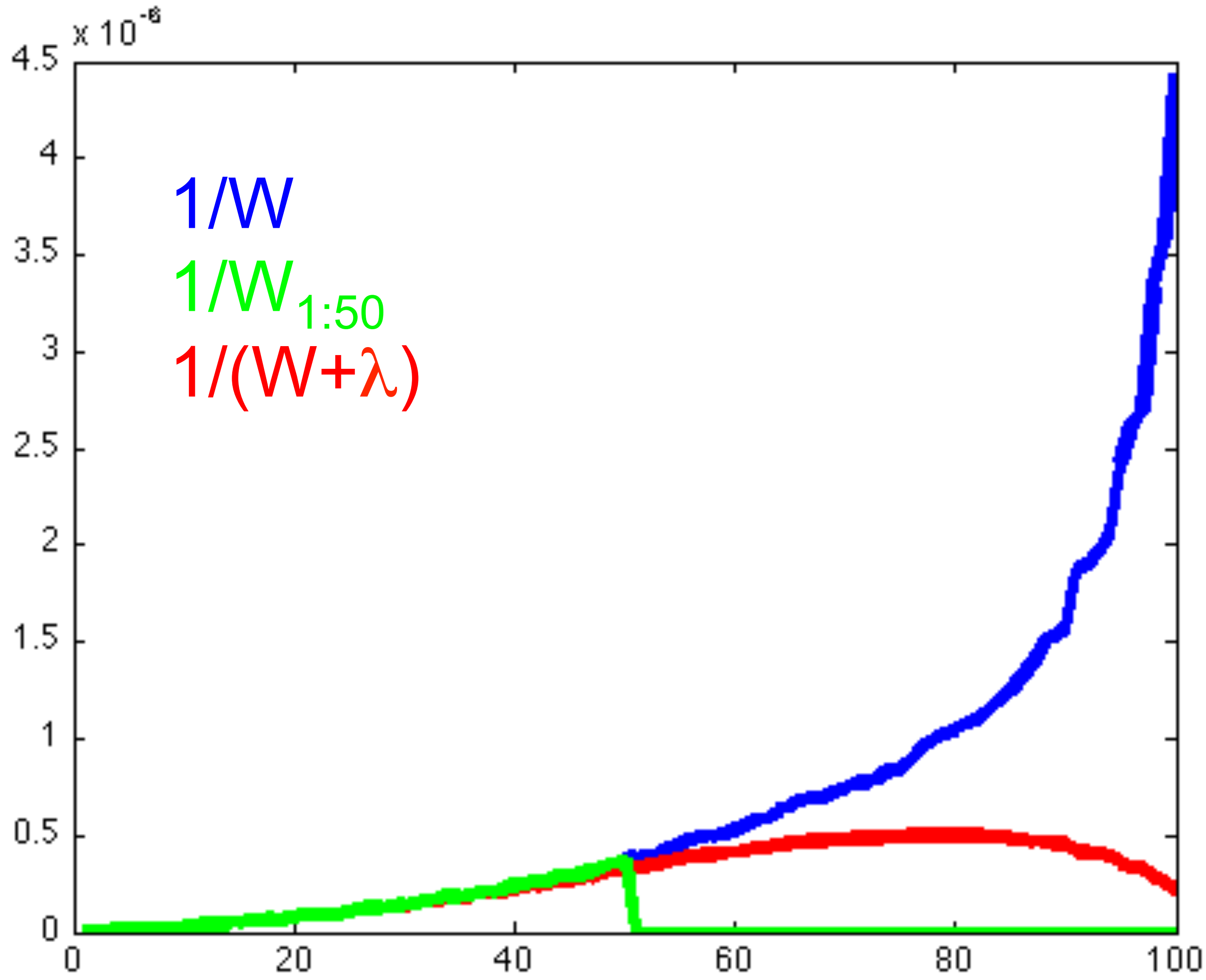
Regularization in Eigen-Space of Stimulus Autocorrelation



Regularization in Eigen-Space of Stimulus Autocorrelation



Ridge vs Subspace Regression



Bayesian Formulation

MAP estimates

Ridge Regression

Neural Response Model

- Model response as a function of stimulus plus noise

$$r = y(S,h) + \varepsilon$$

- For simplicity let's assume a linear model with Gaussian noise

$$r = Sh + \varepsilon \qquad \varepsilon = \mathcal{N}(0,\sigma^2)$$

$$p(r \mid S,h) = \mathcal{N}(r \mid Sh,\sigma^2)$$

How do we find the parameters that determine response?

- We have a response distribution: the probability of response, r, given stimulus, S, and parameters, h

$$p(r \mid S, h) = \mathcal{N}(r \mid Sh, \sigma^2)$$

- How do we determine h? What is the probability of h, given the stimulus, S, and responses, r

$$p(h \mid S, r)$$

Bayes' Rule

$$p(a,b) = p(b,a)$$

$$p(a,b) = p(a \mid b)p(b)$$

$$p(b,a) = p(b \mid a)p(a)$$

$$p(a \mid b) = \frac{p(b \mid a)p(a)}{p(b)}$$

Posterior and Likelihood

Likelihood: Probability of data given the model
Posterior: Probability of model given data

$$p(\vec{h} \mid \vec{s}_i, r_i) = \frac{p(r_i \mid \vec{s}_i, \vec{h})\, p(\vec{h})}{p(r_i)}$$

We can drop p(r) and convert to a proportionality

$$\underbrace{p(\vec{h} \mid \vec{s}_i, r_i)}_{\text{Posterior}} \propto \underbrace{p(r \mid \vec{s}_i, \vec{h})}_{\text{Likelihood}}\, \underbrace{p(\vec{h})}_{\text{Prior}}$$

Case of flat prior

$$p(\vec{h} \mid \vec{s}_i, r_i) \propto p(r \mid \vec{s}_i, \vec{h})$$

Maximizing the Posterior $\Leftrightarrow$ Maximizing the likelihood

The Maximum a Posteriori (MAP) Estimate

$$p(\vec{h} \mid \vec{s}_i, r_i) \propto p(r \mid \vec{s}_i, \vec{h})\, p(\vec{h})$$

Posterior Likelihood Prior

If we choose a Gaussian Prior to combine with the Gaussian Likelihood, the Posterior distribution will also be Gaussian, i.e. it is a conjugate prior

$$L = \prod_{i=1}^{n} \mathcal{N}(r_i \mid \vec{s}_i, \vec{h}, \sigma_n^2)$$

$$p(\vec{h} \mid \lambda) = \mathcal{N}(0, \lambda^{-1}\,\mathrm{I})$$

Negative log of MAP Estimate

$$Posterior \propto \left(\prod_{i=1}^{n} \mathcal{N}(r_i \mid \vec{h}, \vec{s}_i, \sigma^2) \right) \mathcal{N}(0, \lambda^{-1}\,\mathrm{I})$$

$$-\ln(Posterior) \propto \frac{1}{2} \sum_{i=1}^{n} (r_i - \vec{h} \cdot \vec{s}_i)^2 + \frac{\lambda}{2} \left\| \vec{h} \right\|^2$$

Error/Cost Penalty

A Gaussian prior is equivalent to an L2 penalty on the Least Squares Solution

Solving for the MAP Estimate

$$-\ln(Posterior) \propto \frac{1}{2}\sum_{i=1}^{n}(r_i - \vec{h}\cdot\vec{s}_i)^2 + \frac{\lambda}{2}\left\|\vec{h}\right\|^2$$

$$\frac{\partial}{\partial h}\left(\frac{1}{2}\sum_{i=1}^{n}(r_i - \vec{h}\cdot\vec{s}_i)^2 + \frac{\lambda}{2}\vec{h}\cdot\vec{h}\right) = 0$$

$$\mathbf{S}^{\mathrm{T}}\mathbf{r} - \mathbf{S}^{\mathrm{T}}\mathbf{S}\vec{h} + \lambda\mathbf{I}\vec{h} = 0 \qquad \mathbf{S}^{\mathrm{T}}\mathbf{r} = (\mathbf{S}^{\mathrm{T}}\mathbf{S} + \lambda\mathbf{I})\vec{h}$$

$$\vec{h} = (\mathbf{S}^{\mathrm{T}}\mathbf{S} + \lambda\mathbf{I})^{-1}\mathbf{S}^{\mathrm{T}}\mathbf{r}$$

L2=Ridge Regression=Isotropic Gaussian Prior on model parameters Analytical Solution!

$$\boxed{\vec{h} = (\mathbf{S}^{\mathrm{T}}\mathbf{S} + \lambda\mathbf{I})^{-1}\mathbf{S}^{\mathrm{T}}\mathbf{r}}$$

$$(\mathbf{S}^{\mathrm{T}}\mathbf{S}) = \mathbf{Q}\cdot\Lambda\cdot\mathbf{Q}^{\mathrm{T}}$$

$$(\mathbf{S}^{\mathrm{T}}\mathbf{S} + \lambda\mathbf{I})^{-1} = \mathbf{Q}^{\mathrm{T}}\cdot(1/(\Lambda + \lambda))\cdot\mathbf{Q}$$

$$\boxed{\vec{h} = \mathbf{Q}^{\mathrm{T}}(\Lambda + \lambda\mathbf{I})^{-1}\mathbf{Q}\mathbf{S}^{\mathrm{T}}\mathbf{r}}$$

Ridge vs Subspace Regression

$$\vec{h} = \mathbf{Q}^{\mathrm{T}} \left(\frac{1}{\Lambda} \right) \mathbf{Q} \mathbf{S}^{\mathrm{T}} \mathbf{r}$$

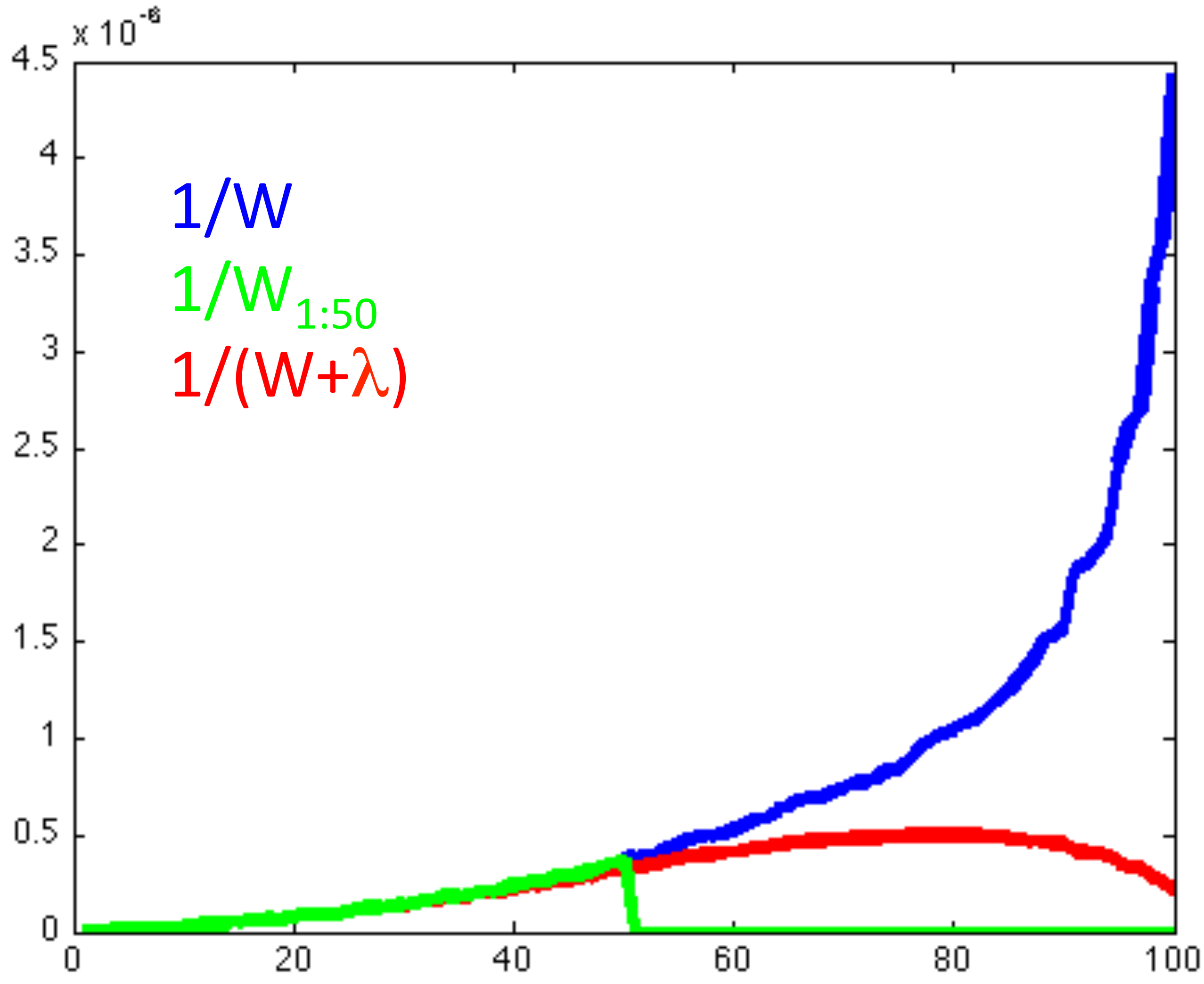
For subspace regression, a subset of the Eigenvalues are used:

$$\vec{h} = \mathbf{Q}^{\mathrm{T}} \left(\frac{1}{\Lambda_{1:50}} \right) \mathbf{Q} \mathbf{S}^{\mathrm{T}} \mathbf{r}$$

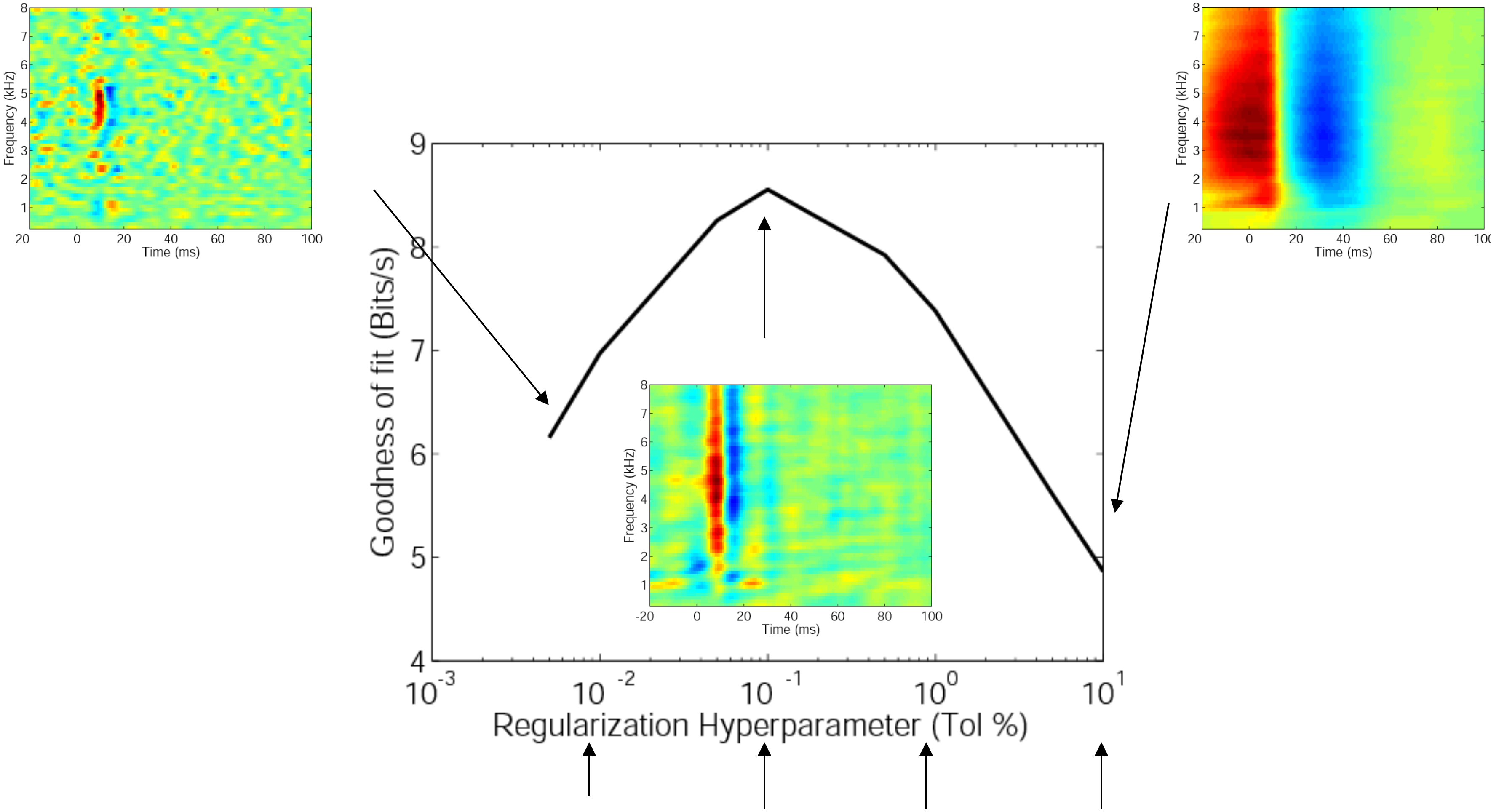
For Ridge regression, all eigenvalues are used, but reweighted:

$$\vec{h} = \mathbf{Q}^{\mathrm{T}} \left(\frac{1}{(\Lambda + \lambda \mathbf{I})} \right) \mathbf{Q} \mathbf{S}^{\mathrm{T}} \mathbf{r}$$

Ridge vs Subspace Regression

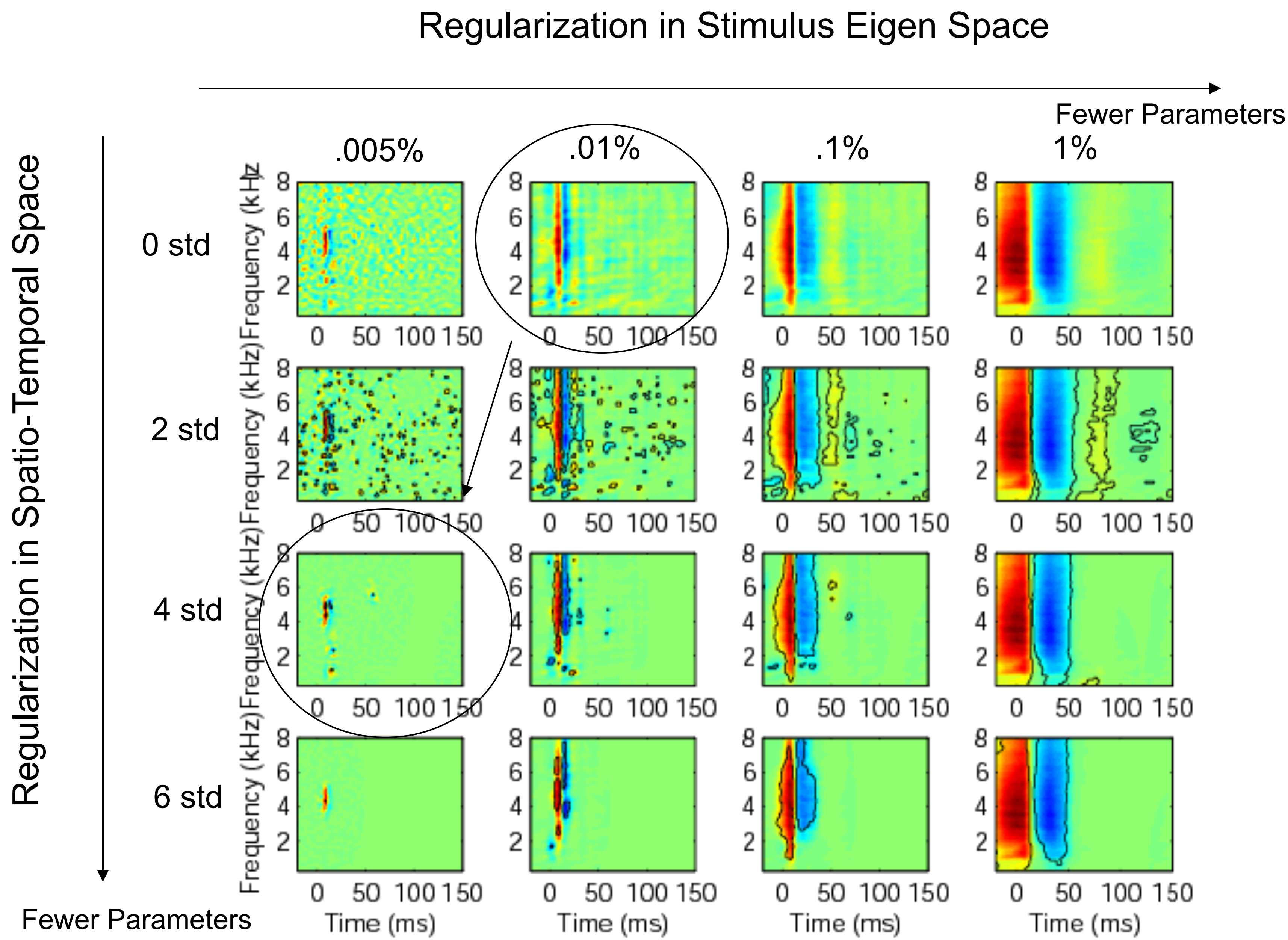


Regularization in Eigen-Space of Stimulus Autocorrelation



Number of parameters 3334 342 50 16

Regularization with Two Hyper Parameters



Validation

Neural Responses are noisy

$$R^2(\omega) = A^2(\omega) + N^2(\omega)$$

PSTH is less noisy but still noisy

$$\overline{R}_M^2(\omega) = A^2(\omega) + N^2(\omega) / M$$

Actual Response A is unknown

Validation

Solution: Compare the coherence function between a single spike train, R, and the actual rate, A, with the coherence function between a single spike train, R, and the predicted rate, B.

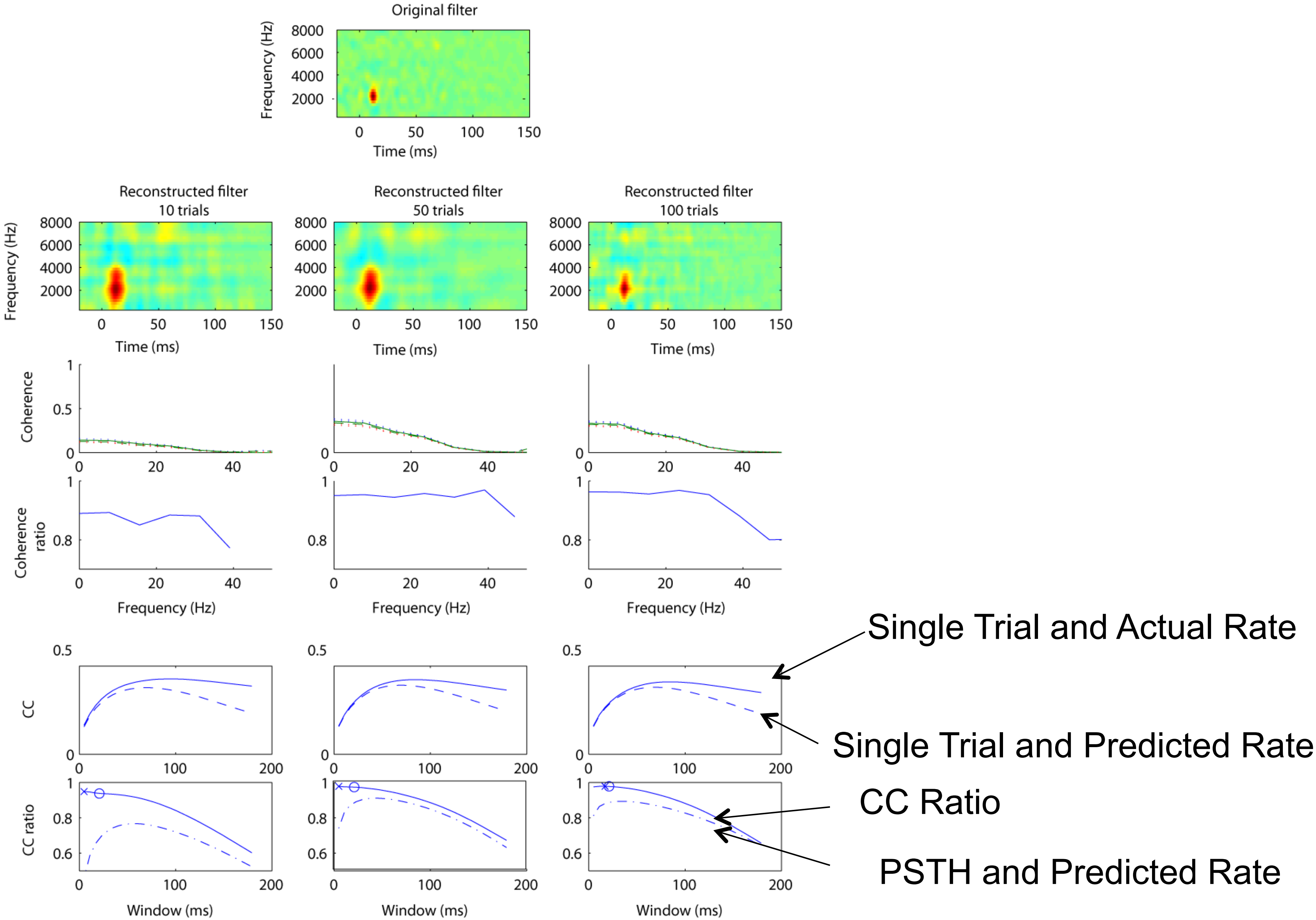
Coherence between R and A is estimated by resampling:

$$\frac{1}{\gamma^2_{AR}} - 1 = \frac{-M + M\sqrt{\left(\frac{1}{\gamma^2_{\bar{R}_{1,M/2}\bar{R}_{2,M/2}}}\right)}}{2}$$

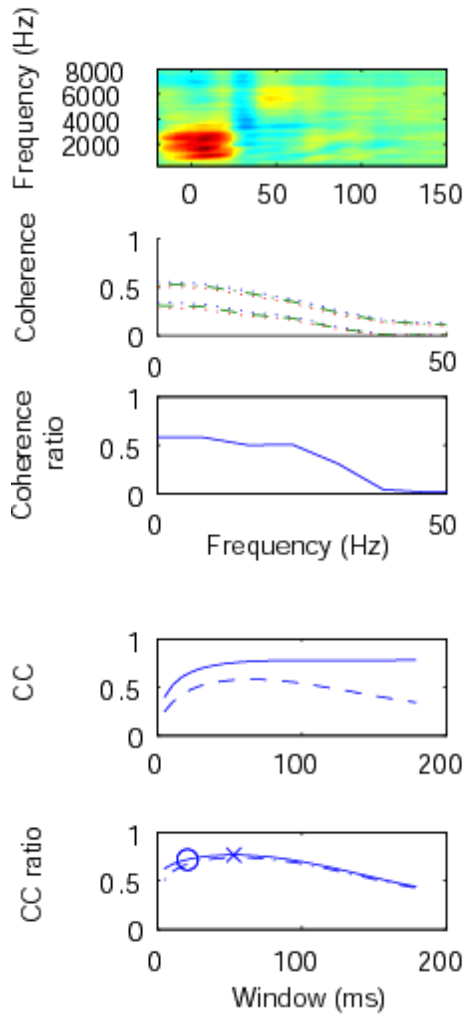
Coherence between R and B can be estimated from the coherence between the PSTH and B:

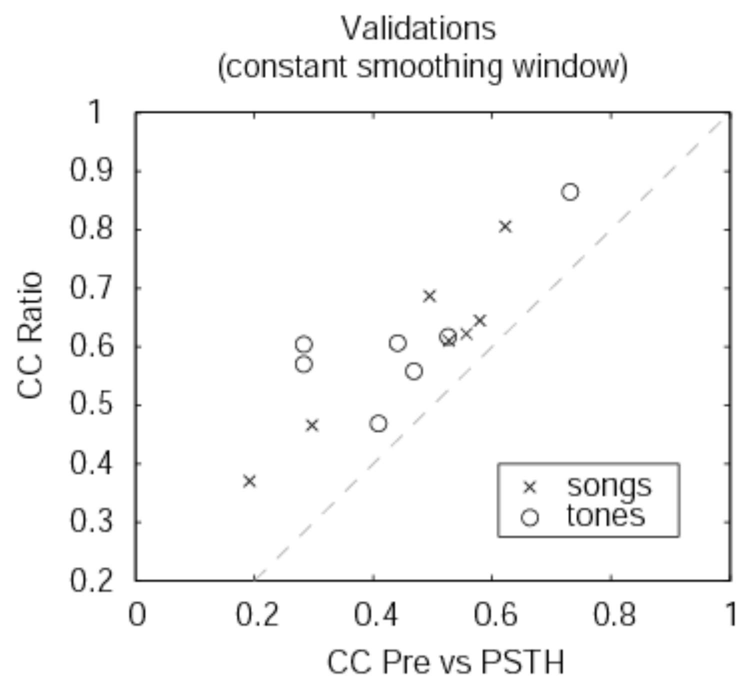
$$\frac{\gamma^2_{BR}}{\gamma^2_{B\bar{R}}} = \frac{1 + \sqrt{\left(\frac{1}{\gamma^2_{\bar{R}_{1,M/2}\bar{R}_{2,M/2}}}\right)}}{-M + M\sqrt{\left(\frac{1}{\gamma^2_{\bar{R}_{1,M/2}\bar{R}_{2,M/2}}}\right)} + 2}$$

Validation: Model Linear Neuron



Validation: Actual Data (Auditory Neurons)





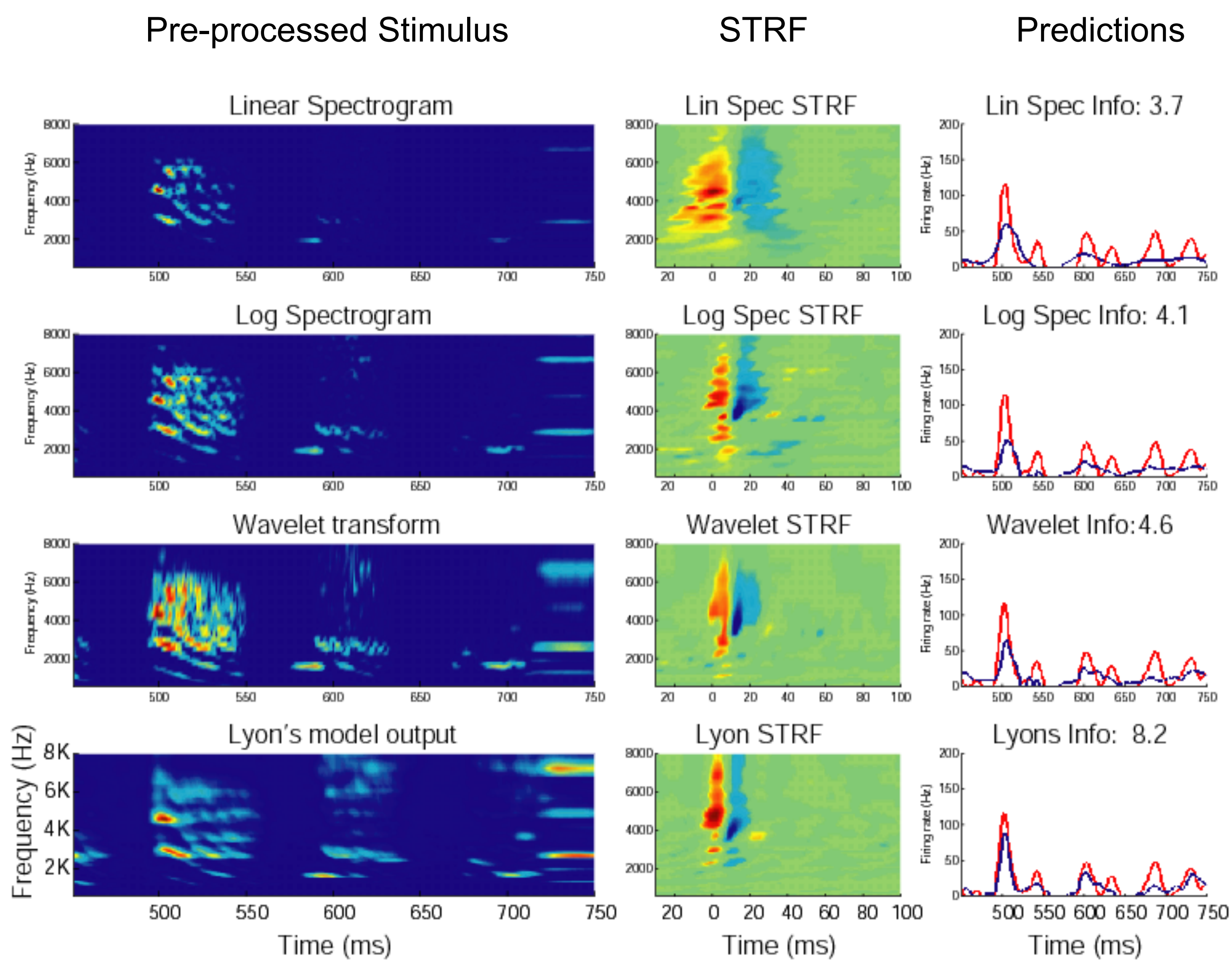
Validation: Summary Lessons

- Figure out your ceiling value
- Use coherence function
- If you use the correlation coefficient (r) specify the time scale (smoothing window).

Hsu, Borst and Theunissen. Network: Comput. Neural Syst. ***15*** *(2004) 91–109*

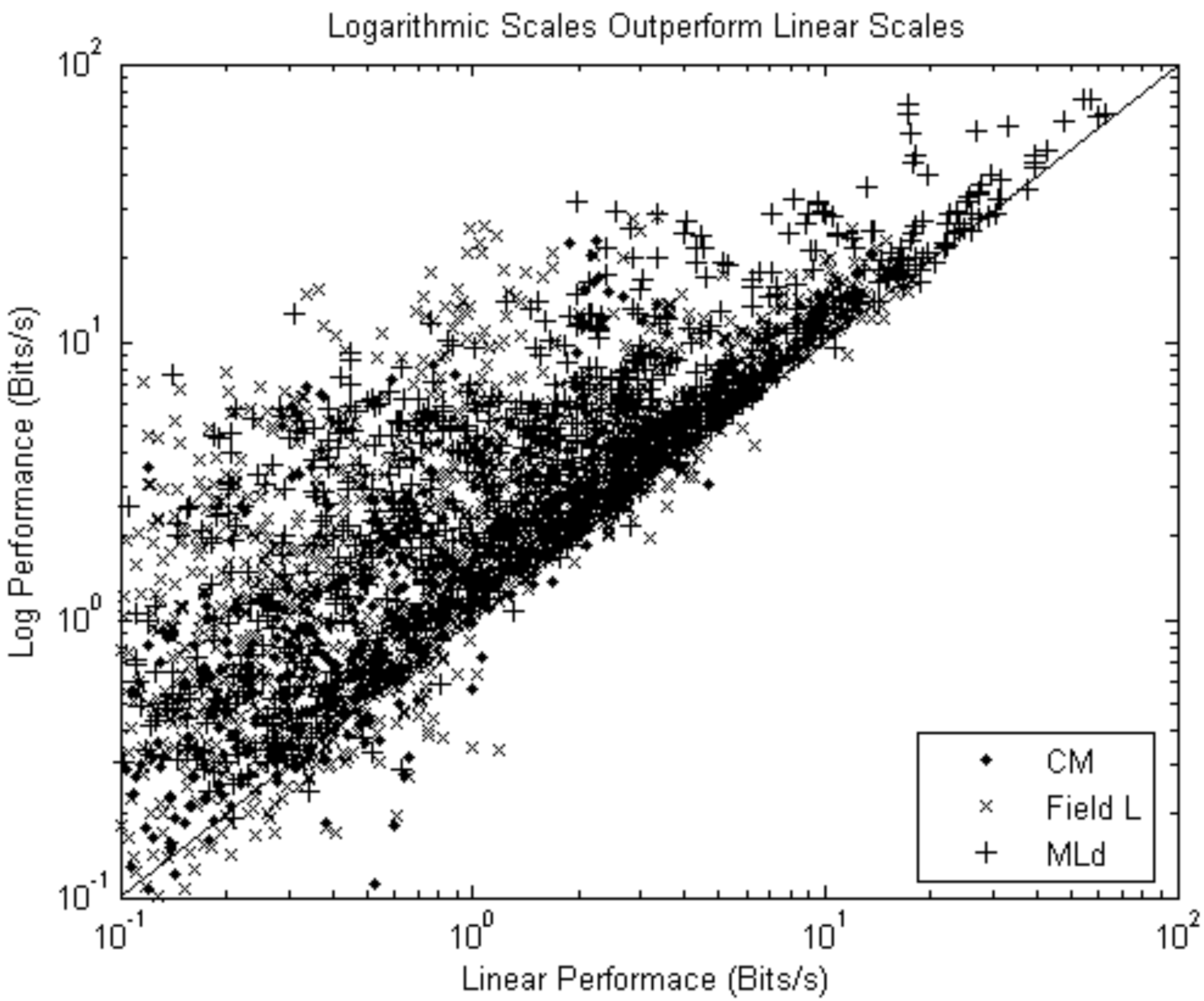
Rev-Correlations for Nat Stats

- STRFPAK.
  - Non stationary spatial stimulus correlations
  - Regularization in eigen-space
  - Regularization in joint eigen/pixel space
  - Prediction Validation
- Non-linear Preprocessing for auditory neurons.
  - Compression
  - Spectrograms vs Wavelets and Time-frequency scale
  - Adaptive gain control

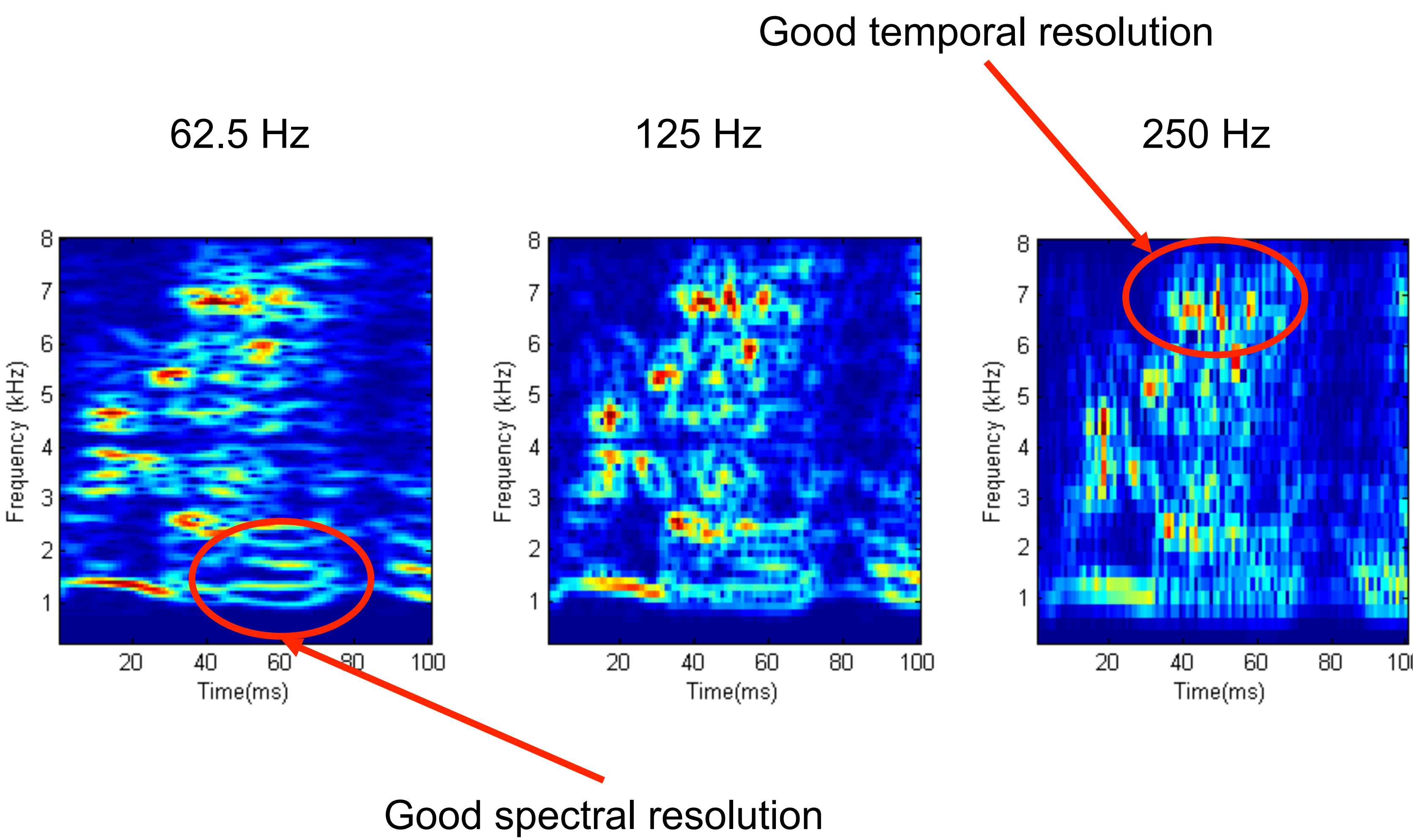


Pre-processed Stimulus
STRF
Predictions
Linear Spectrogram
Lin Spec STRF
Lin Spec Info: 3.7
Log Spectrogram
Log Spec STRF
Log Spec Info: 4.1
Wavelet transform
Wavelet STRF
Wavelet Info:4.6
Lyon's model output
Lyon STRF
Lyons Info: 8.2
Frequency (Hz)
Firing rate (Hz)
Time (ms)

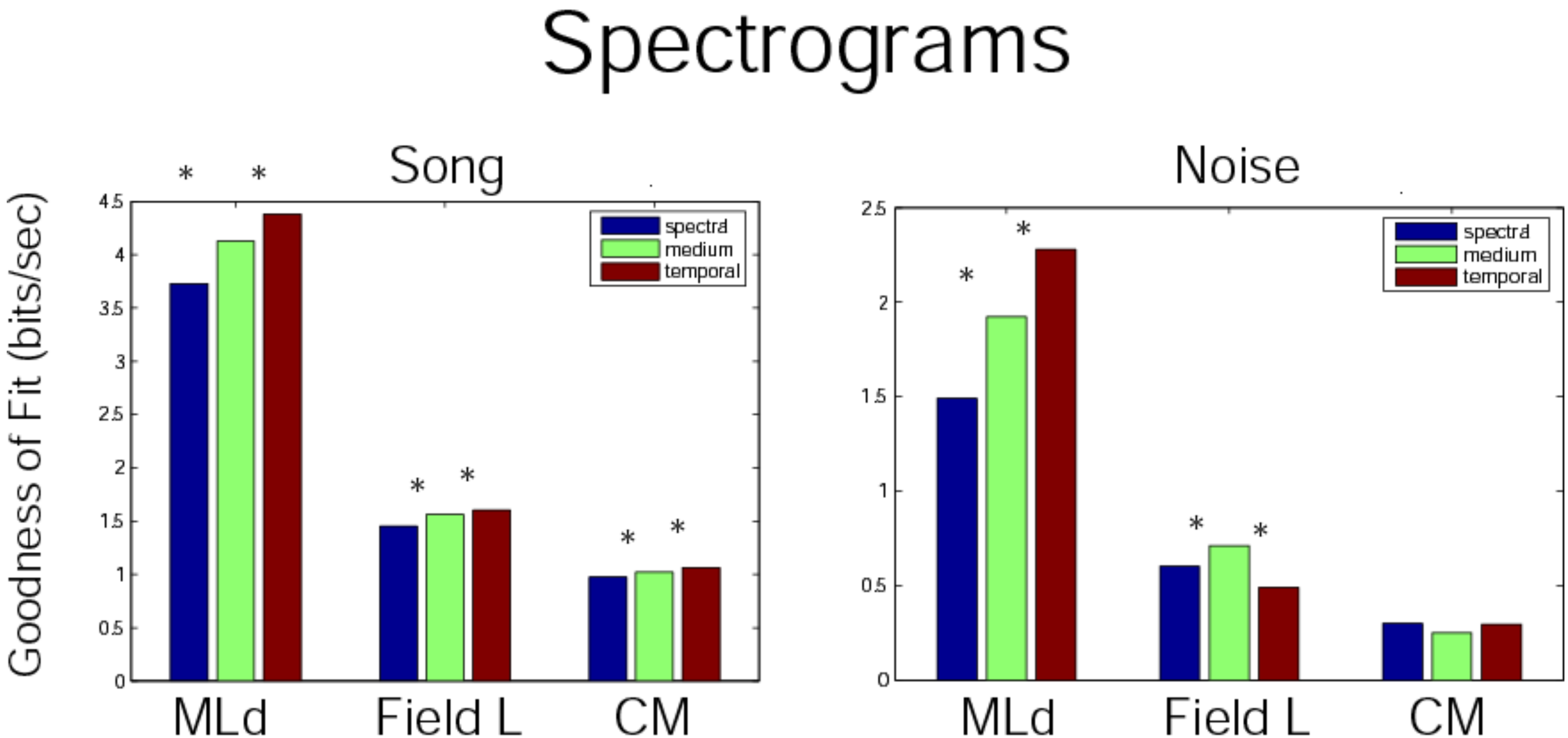
Compression is a must!



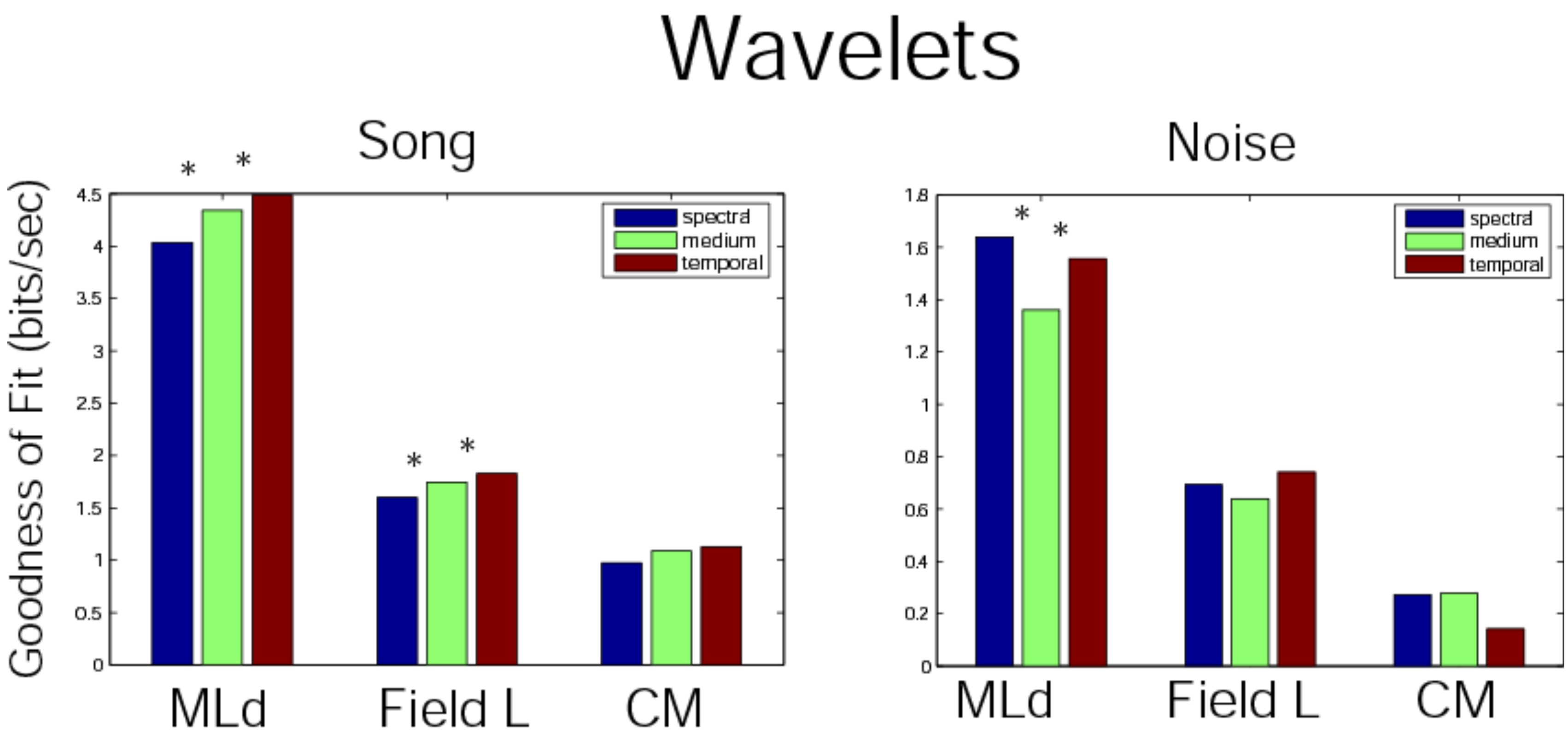
Three Time-Frequency Scales



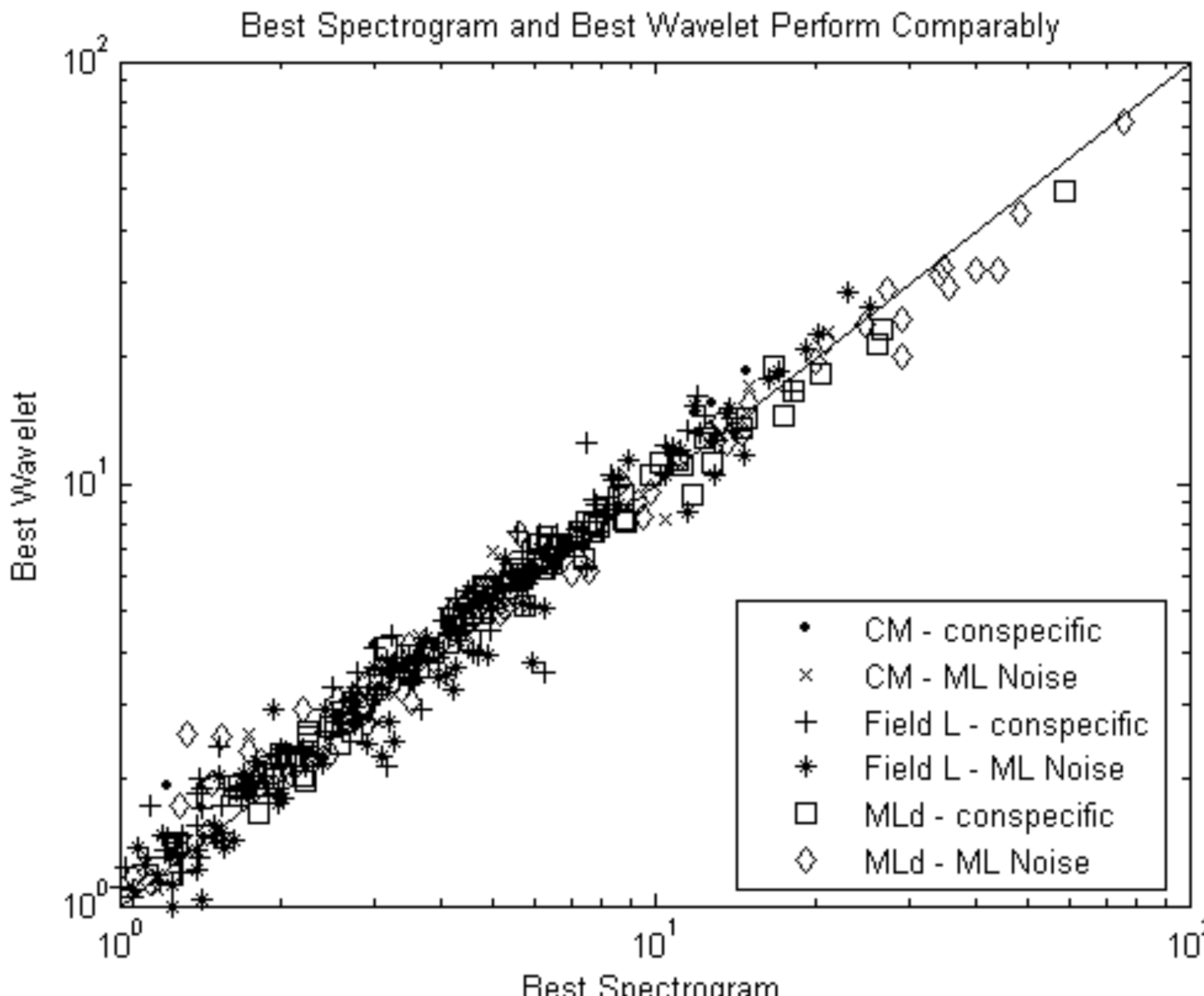
Temporal Representations are favored for song



Temporal Representations are favored for song



Spectrogram and Wavelet yield similar results

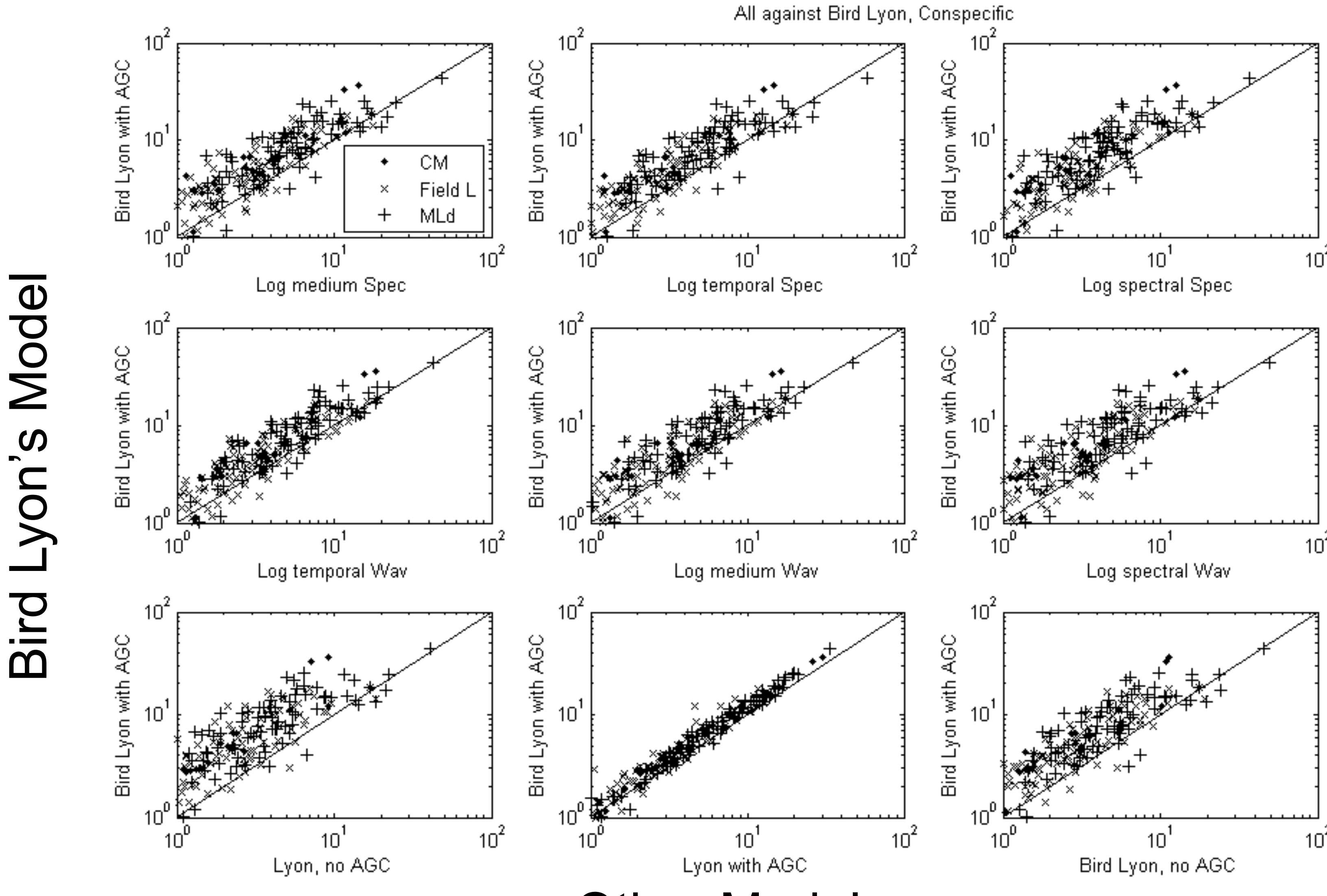


Dynamic Gain Control in Lyon/ Slaney's cochlear model

The model generates a cochleogram. It incorporates:

- Wavelet like filter bank
- Rectification, Low Pass filter, Compression
- Adaptive Gain Control with 4 stages with time constants of 640, 160, 40 and 10 ms.
- We modified Lyon's model to use the time-frequency that gave us the best predictions for the wavelets.

Adaptive gain control is a winner!



Pre-processing for auditory system

- Does NOT affect goodness of fit:
  - Spectrograms vs Wavelets.
- Has small effect on goodness of fit:
  - Time-frequency scale. Stimulus dependant.
- Increases goodness of fit:
  - Input compressive non-linearity
  - Dynamic gain control

References

- Theunissen, F. E., S. V. David, N. C. Singh, A. Hsu, W. Vinje and J. L. Gallant (2001). "Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli." Network: Comp. Neural Syst. **12**: 1-28.
- Hsu, A., A. Borst and F. E. Theunissen (2004). "Quantifying variability in neural responses and its application for the validation of model predictions." Network **15**(2): 91-109.
- Gill, P., J. Zhang, S. M. Woolley, T. Fremouw and F. E. Theunissen (2006). "Sound representation methods for spectro-temporal receptive field estimation." J Comput Neurosci **22**: 22.
- Wu, M. C. K., S. V. David and J. L. Gallant (2006). Complete functional characterization of sensory neurons by system identification. Annual Review of Neuroscience. **29:** 477-505.

Assignment

- directfit_exercise.m
- Finish the day by running: direcfit_tutorial.m