# Neural Constraints in Auditory Cortex

Jonathan Z. Simon
Didier A. Depireux
David J. Klein
Shihab A. Shamma

Institute for Systems Research
University of Maryland
College Park

Center for Auditory
and Acoustic Research

Institute for Systems Research
University of Maryland

# Outline

- Introduction to Auditory Spectro-Temporal Processing
- Stimuli—Dynamic & Broadband: Ripples, Ripple Combinations
- Examples & Properties of Spectro-Temporal Response Fields
- Constraints on Neural Connectivity

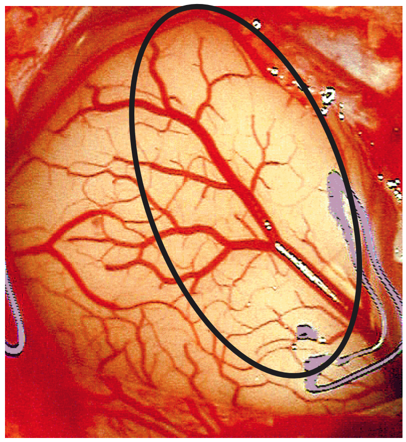
- **Introduction to Auditory Spectro-Temporal Processing**
- **Stimuli—Dynamic & Broadband: Ripples, Ripple Combinations**
- **Examples & Properties of Spectro-Temporal Response Fields**
- **Constraints on Neural Connectivity**

# Basics

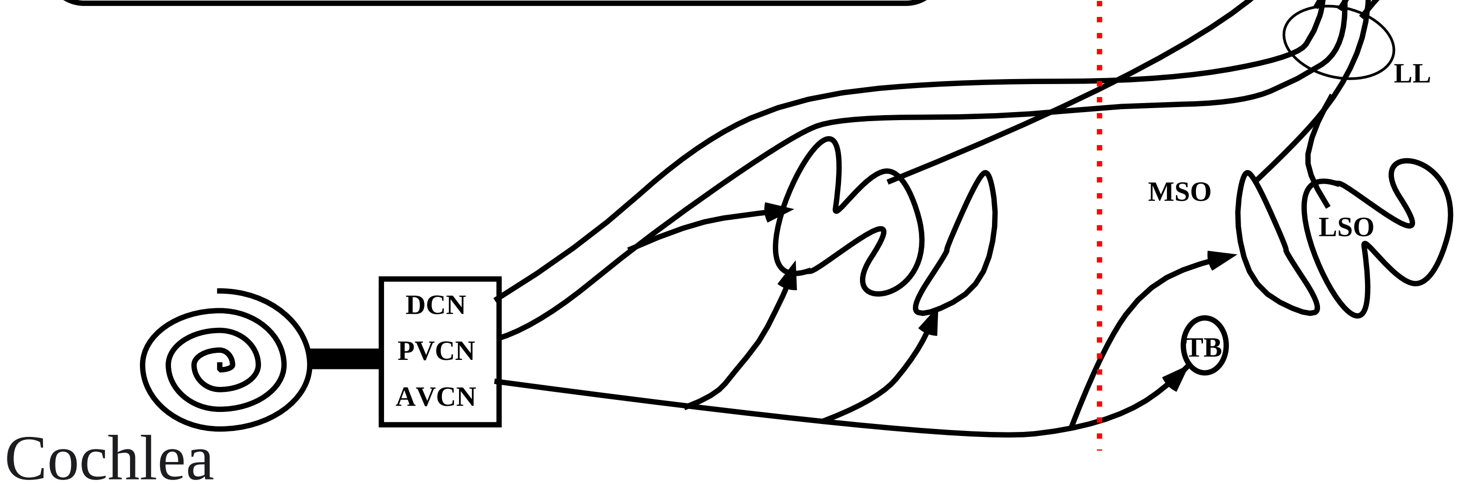
- **Most important auditory cues are acoustically non-trivial**
  - e.g. speech, speaker ID, emotional content, pitch, timbre, sound location, and many, many others
- **Enormous parallel and serial neural processing in multiple stages** from auditory nerve to cortex
- **Neural code is essentially unknown** for almost all auditory features
  - Especially in cortex
  - Much progress in coding near periphery, especially coding of sound location

# Auditory Pathway
(oversimplified)

Primary Auditory Cortex (AI)



*Phase-locks to envelope of acoustic waveform up to ~20 Hz*




*Linear distance ~* log *f*

*Phase-locks to acoustic waveform itself up to ~2 kHz*

# Motivation

- **The Quest**

  Teasing out "function" of Primary Auditory Cortex (AI)

  ≈ { which sounds/features evoke responses?
  how are they encoded into spike trains? }

- **Broadband and dynamic sounds**
  - Evoke strong, sustained, dynamic responses in AI
  - Many natural sounds, e.g. speech, backgrounds

- **Reasonable quest:** Quantitative measure of how spikes encode sound features
  - Quantitative descriptor (and predictor)
  - Qualitative descriptor/Visual tool

# The Path

- **Compromise** from quantitative necessity
  - Restrict broadband and dynamic sounds to mathematically simple subset:
  - **Noise**—strongly modulated in **spectrum** and **time**
  - not a severe compromise

- **Spectro-Temporal Receptive Field (STRF)** succeeds:
  - Quantitative descriptor (and predictor)
  - Qualitative descriptor/Visual Tool

- **Bonus**
  - Constraints on Neural Connectivity

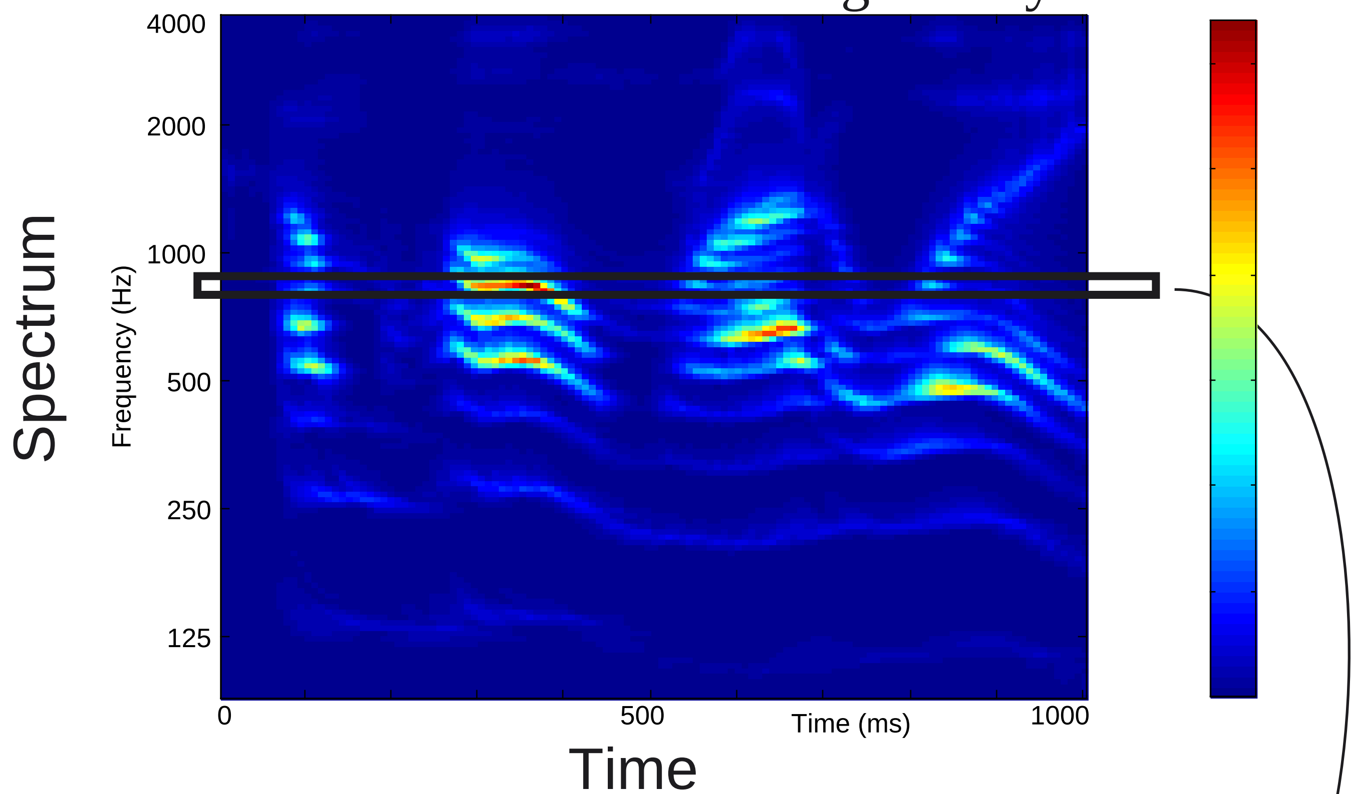
# Sound Features

- **Spectro-Temporal Features of Any Sound**

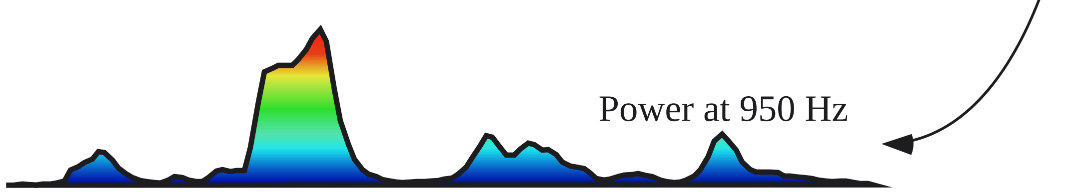
- Spectral content of sound as a function of time.

  Which spectral frequency bands have enhanced power?
  Which spectral frequency bands have diminished power?
  How do these change as a function of time?

"Come home right away."




- log $f \approx$ linear cochlear distance

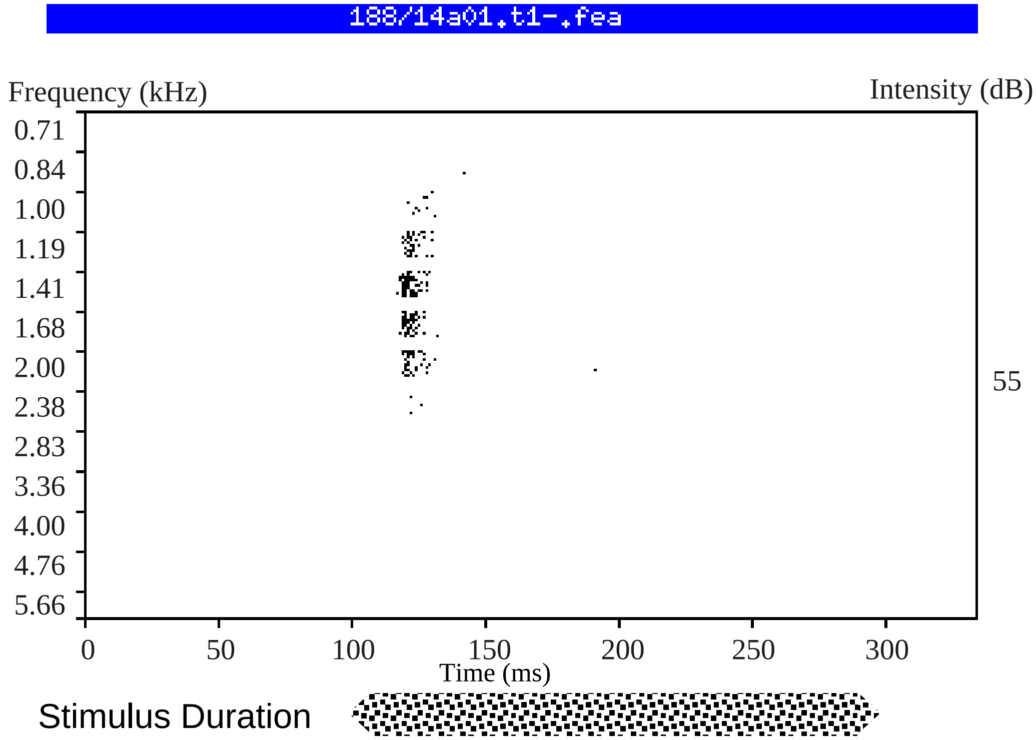


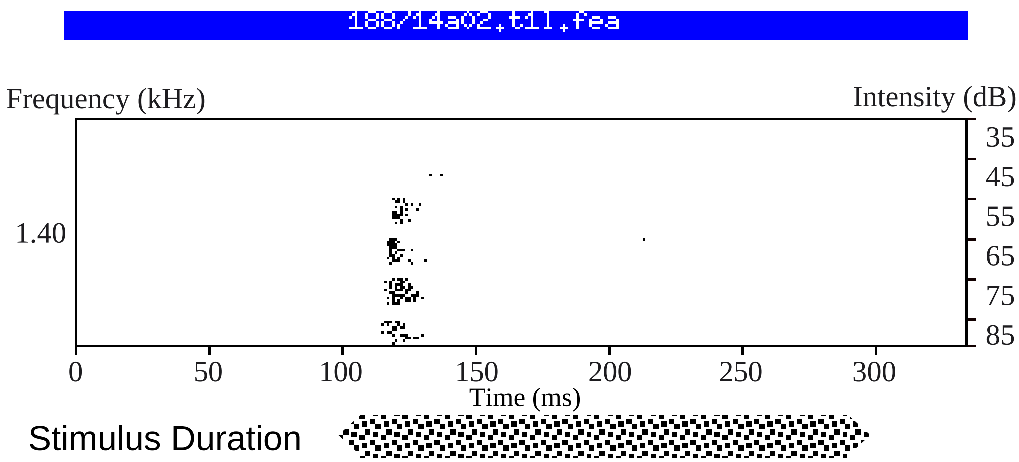

- Characterization from cross-section is limited

Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

# Response to Pure Tones



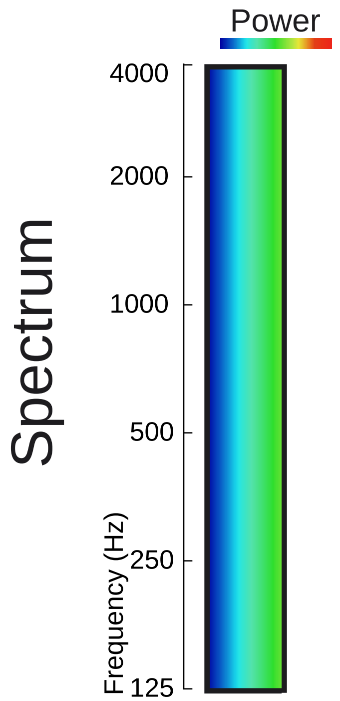



- Introduction to Auditory Spectro-Temporal Processing
- **Stimuli—Dynamic & Broadband: Ripples, Ripple Combinations**
- Examples & Properties of Spectro-Temporal Response Fields
- Constraints on Neural Connectivity

# Stimulus Construction

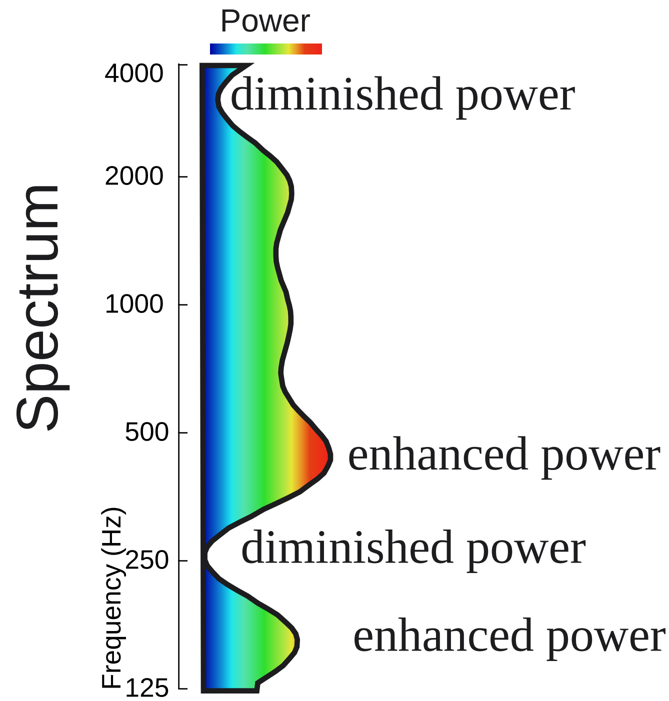
- Pink Noise = flat power density in octaves (log *f*)
  not white

- ***Unmodulated noise** (flat)*

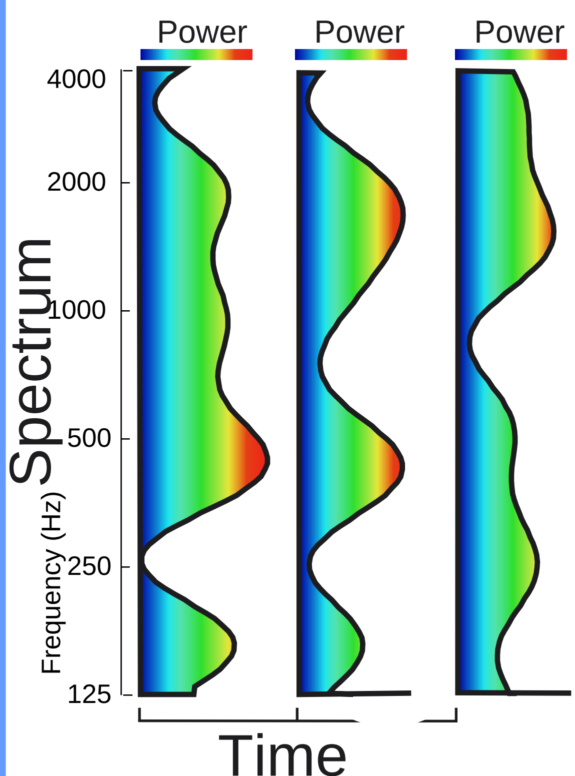



flat
power
over
entire
spectrum

- ***Spectrally modulated noise***




- ***Spectro-Temporally modulated noise***

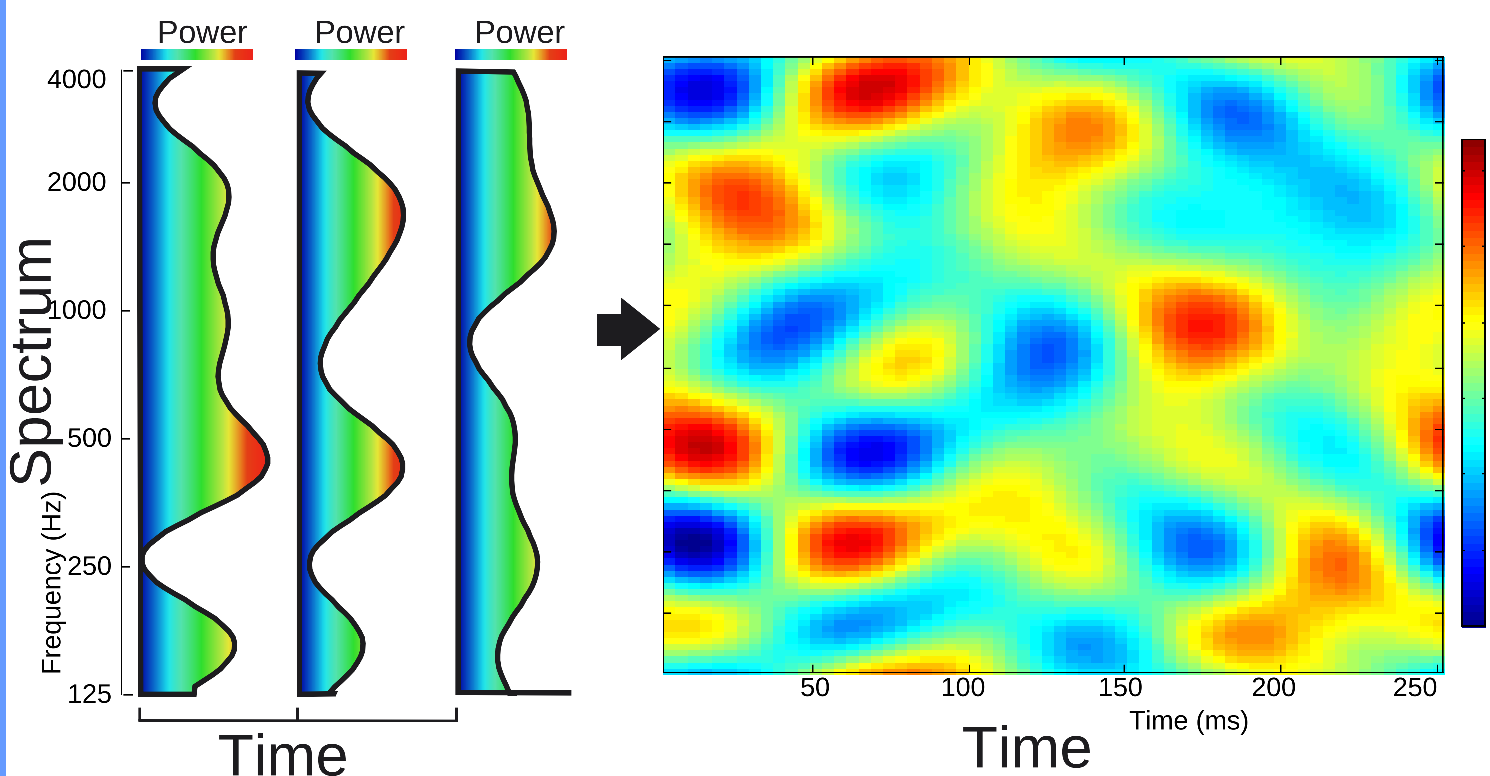



Center for Auditory
and Acoustic Research

Institute for Systems Research
University of Maryland

# Stimulus Construction

- ***Spectro-Temporally modulated noise***



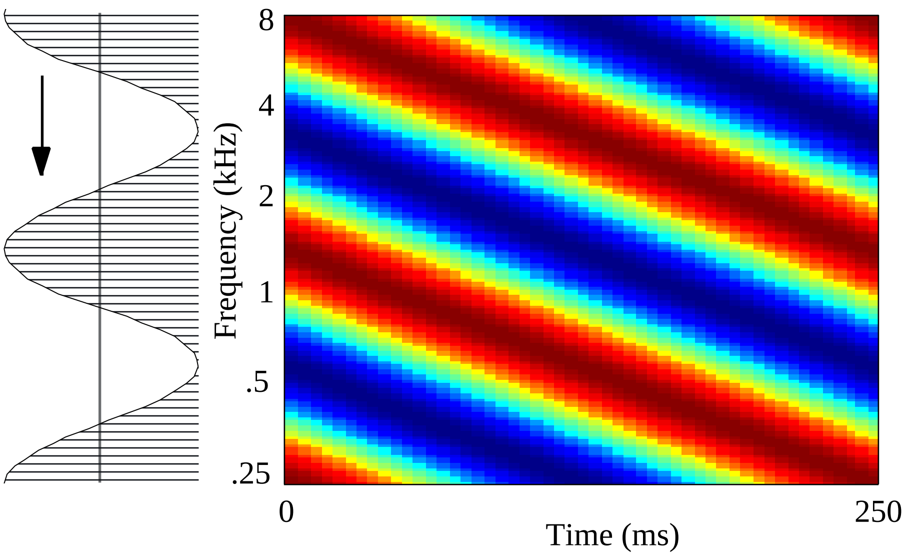
# Single Moving Ripple

Simplest Dynamic Stimulus Used

$$S(t,x) = \sin(2\pi wt + 2\pi\Omega x + \phi)$$

$x = \log_2(f / f_0)$
$w$ = ripple velocity,
e.g. 4 Hz = 4 cycles/s
$\Omega$ = ripple density,
e.g. 0.4 cycles/octave
= 2 cycles/5 octaves

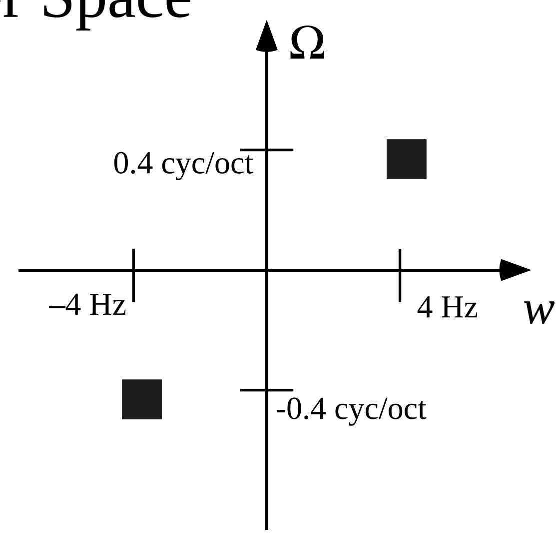
in Spectro-Temporal Space (*Spectrogram*)




c.f. visual contrast gratings

$$\int [.] \exp(\pm 2\pi j\Omega x \pm 2\pi jwt)$$

in Fourier Space




The Fourier transform of a single moving sinusoid has support only on a single point (and its complex conjugate).


Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

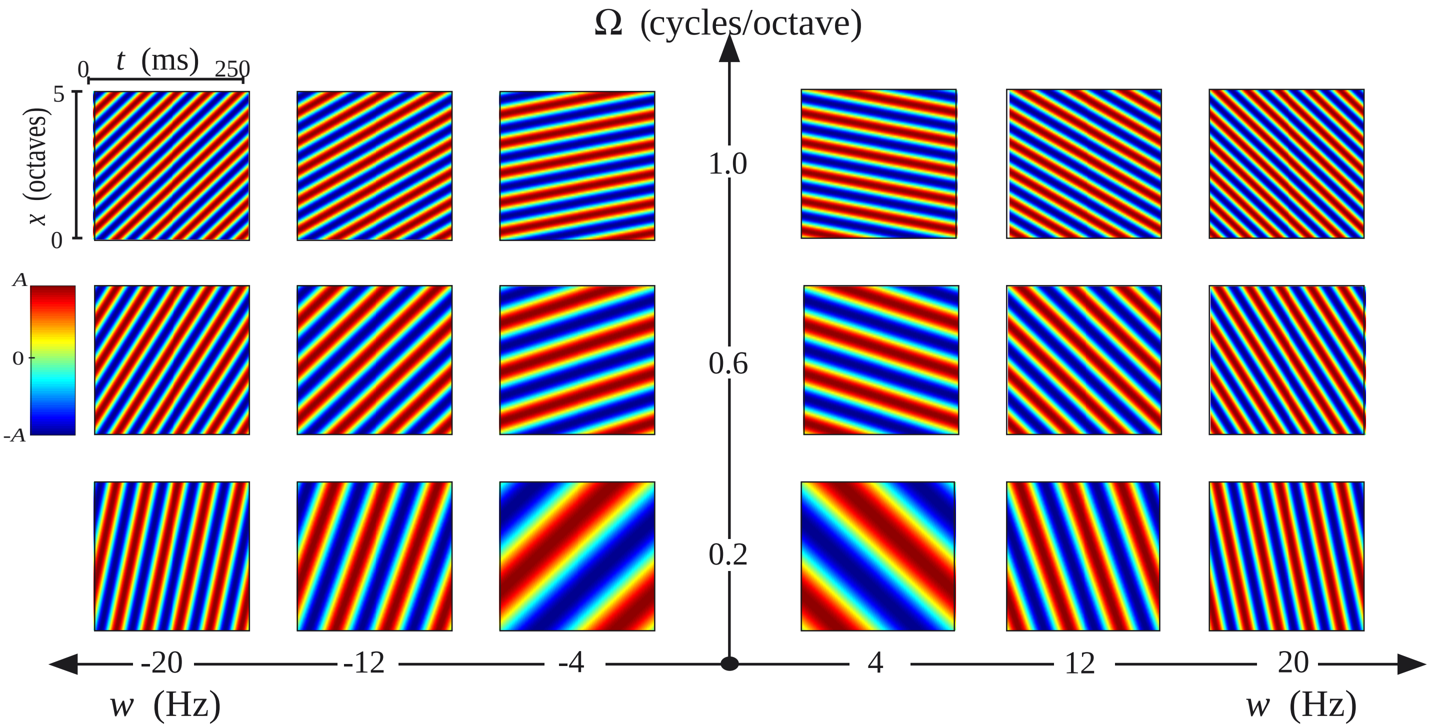
# Multiple Individual Ripples

$S(t,x) = \sin(2\pi w t + 2\pi \Omega x + \phi)$

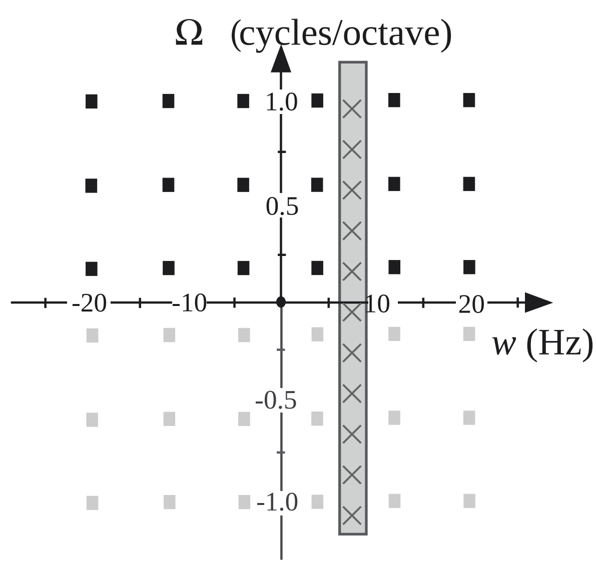
$x = \log_2(f / f_0)$

$w$ = ripple velocity (Hz)

$\Omega$ = ripple density (cyc/oct)




The Transfer function can be obtained by measuring the amplitude and phase of the response to each single ripple.

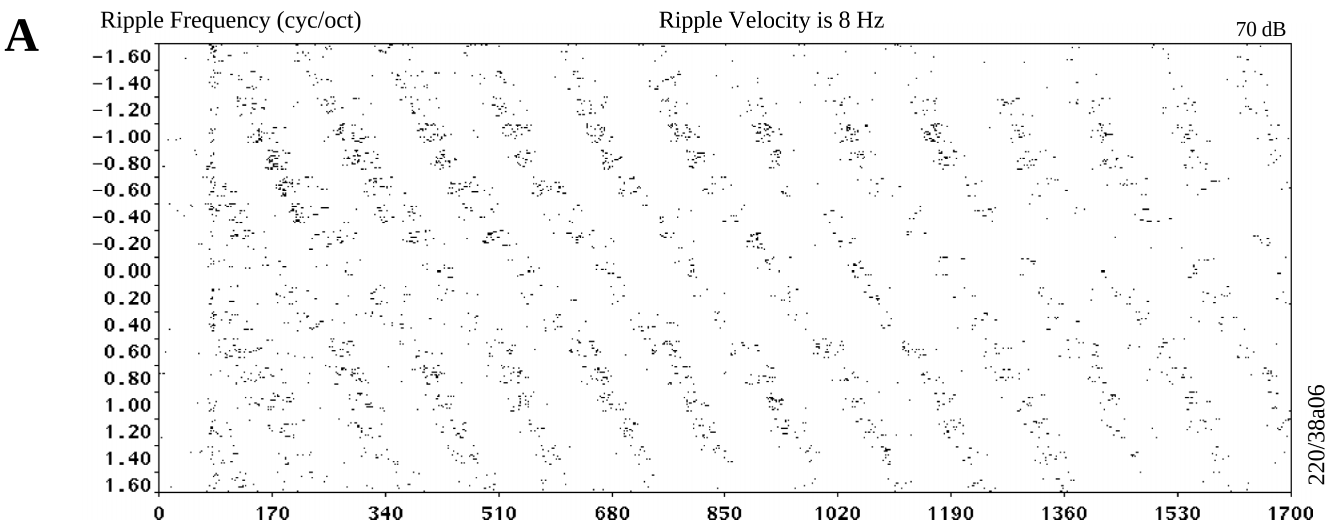


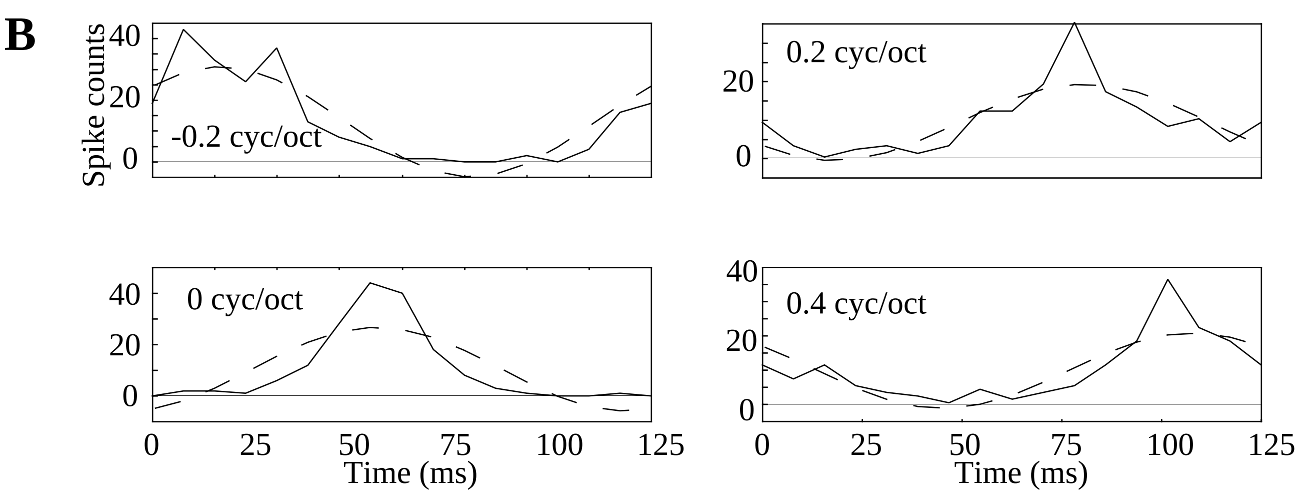

*Audio Demo*

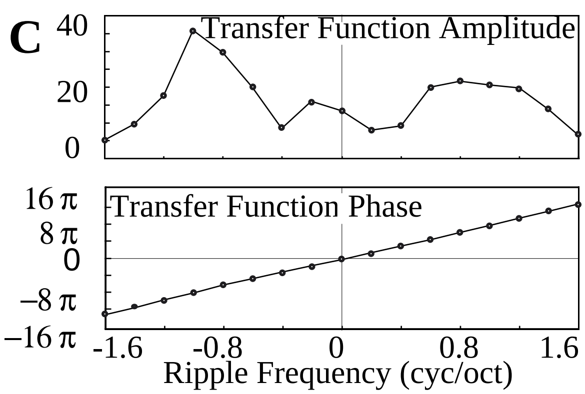

Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

# Spike Train Measurements

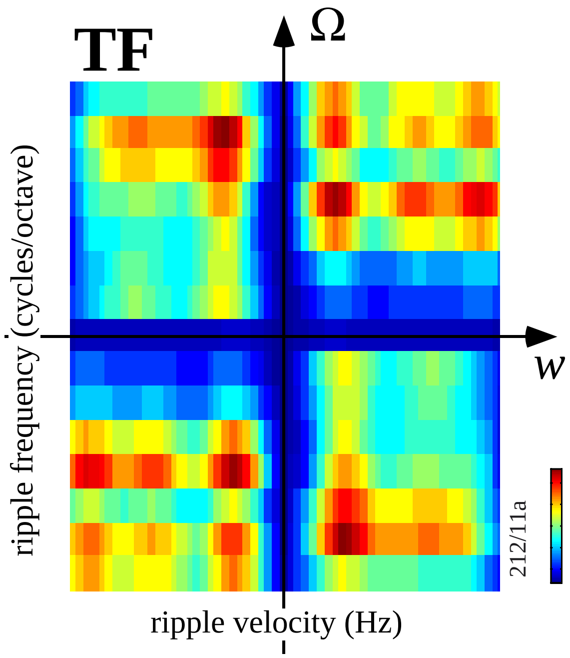



Spike events in (**A**) are turned into period histograms in (**B**). The amplitudes and phases give the transfer function in (**C**).

# Spectro-Temporal Response Field (STRF)



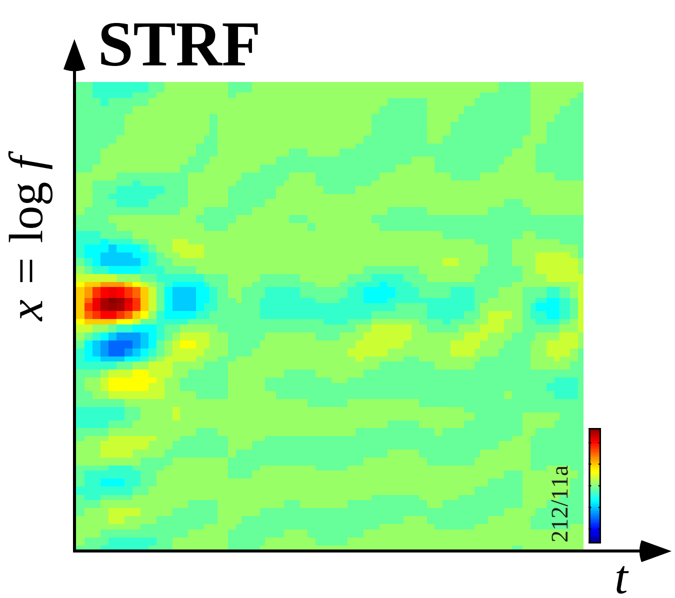

## 2 Dimensional Transfer Function

- Complex conjugate symmetric
- Spectral range: ~ 0 — ~ 2 cycle/octave
- Temporal range: ~ 2 — ~ 20 Hz

*2 Dimensional Inverse Fourier Transform*

$$\int [.] \exp(\pm 2\pi j \Omega x \pm 2\pi j w t)$$

*2 Dimensional Fourier Transform*




## Spectro-Temporal Response Function of the same neuron

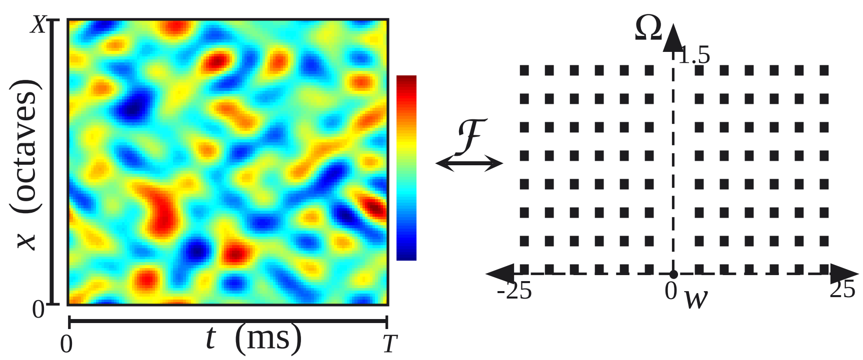
Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

# Spectro-Temporal Noise

To speed up the characterization of a cell's response, we use combinations of ripples of *all* velocities $w$ and densities $\Omega$, with random phases.

$$S^{\text{noise}}(t,x) = \sum_j \sum_k \sin(2\pi w_j t + 2\pi \Omega_k x + \phi_{j,k})$$




Spectro-Temporal generalization of white noise

Cross-Correlation: $C(\tau, x) = \frac{1}{T}\int_0^T S(t, x)R(t-\tau)\mathrm{d}t = \frac{1}{T}\sum_k S(t_k-\tau, x)$

= Spike-Triggered Average

- $C(\tau, x)$ contains cross terms
- Cross terms have random phase and can be attenuated by averaging over multiple, random-phase stimuli $S_j$

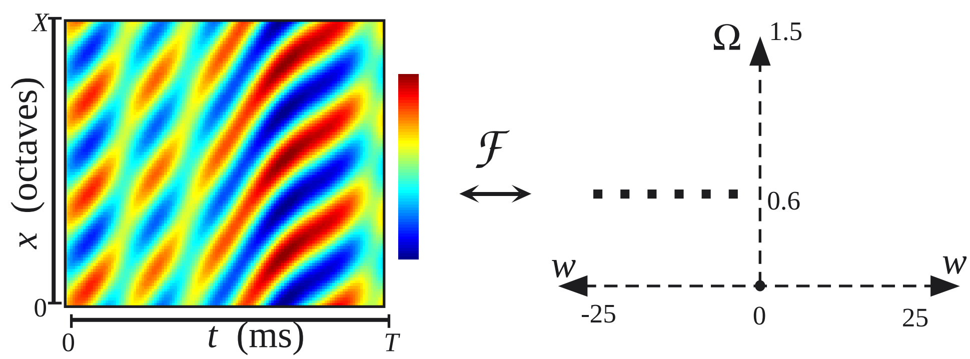
$$\text{STRF}_{\text{est}}(\tau, x) = \frac{1}{m}\sum_{j=1}^{m} C_j(-\tau, x)$$

- Cross terms give noisy estimates without many random-phase stimuli

# Temporally Orthogonal Ripple Combinations (TORCs)

To eliminate interference from cross-terms, we use specific combinations of ripples with differing velocities *w* and random phases.

$$S^{\mathrm{TORC}}(t,x) = \sum_j \sin(2\pi w_j t + 2\pi \Omega_k x + \phi_{j,k})$$




- Stimuli have unique instances of each ripple velocity.
- Multiple stimuli are still needed to present a complete set of ripples.

Cross-Correlation: $C(\tau, x) = \dfrac{1}{T}\int_0^T S(t, x)R(t-\tau)\mathrm{d}t = \dfrac{1}{T}\sum_k S(t_k-\tau, x)$

= Spike-Triggered Average

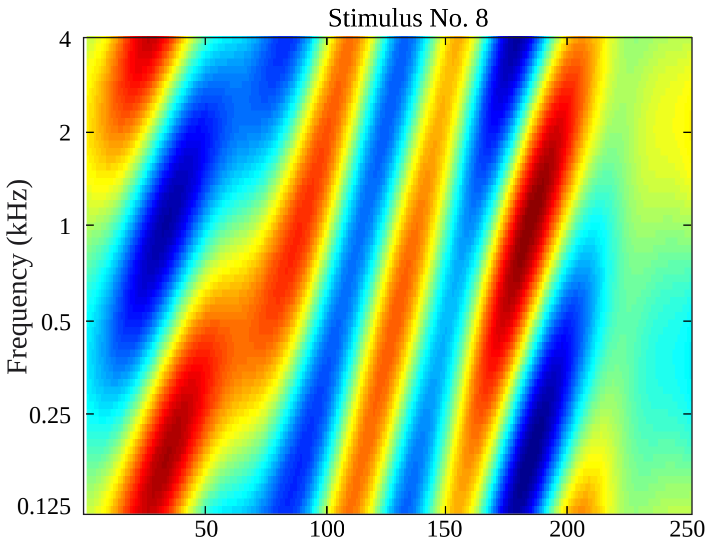
- $C(\tau, x)$ contains no cross terms

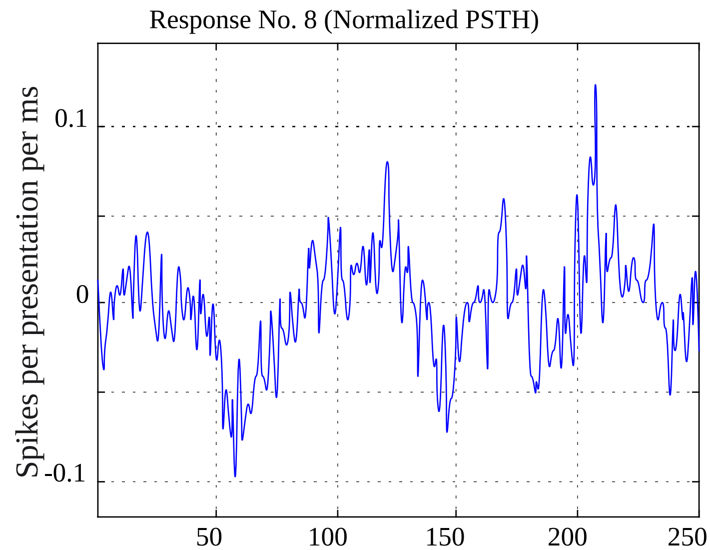
$$\mathrm{STRF}_{\mathrm{est}}(\tau, x) = \sum_{j=1}^{m} C_j(-\tau, x)$$

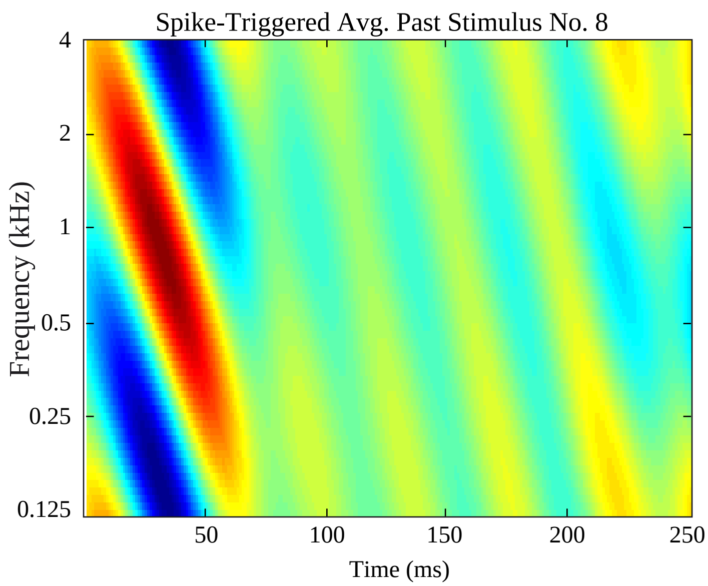

Center for Auditory and Acoustic Research

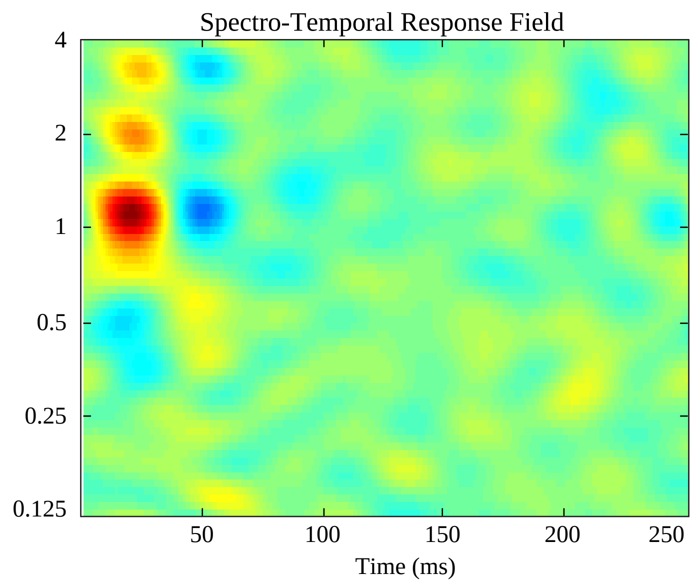
Institute for Systems Research
University of Maryland

# Spike Averaged Data



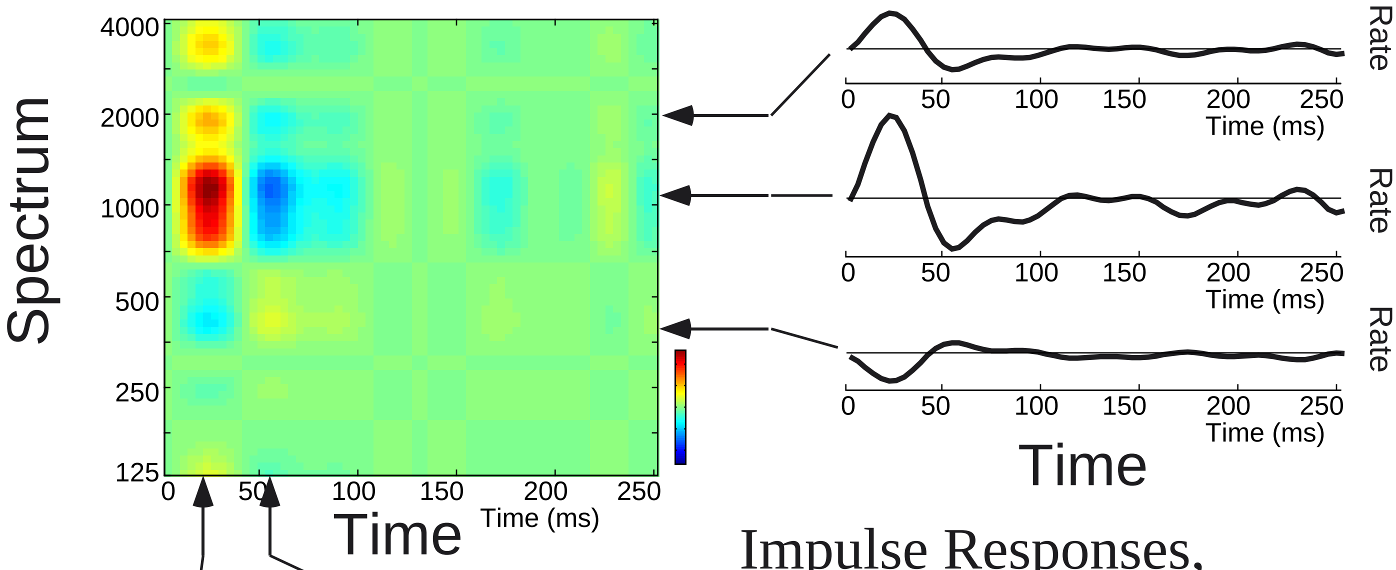

Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

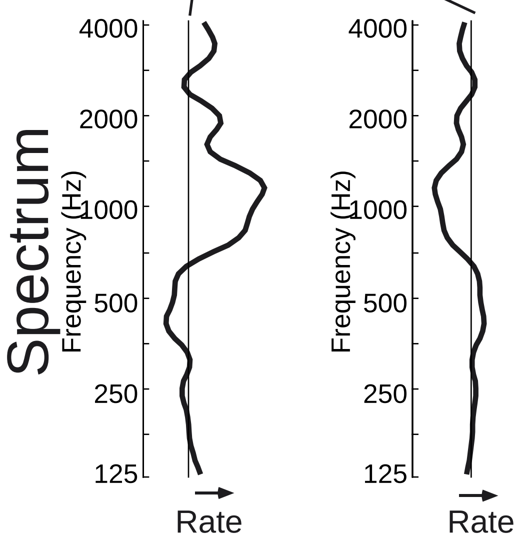
- Introduction to Auditory Spectro-Temporal Processing
- Stimuli—Dynamic & Broadband: Ripples, Ripple Combinations
- **Examples & Properties of Spectro-Temporal Response Fields**
- Constraints on Neural Connectivity

# Interpreting STRFs

## Cross-section interpretations

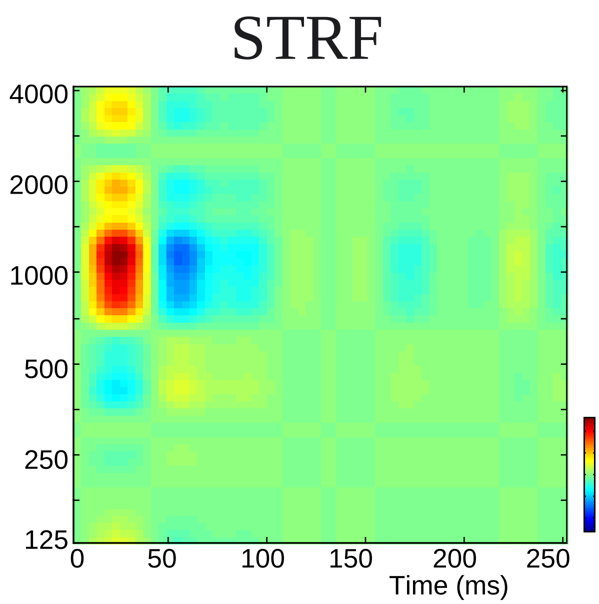


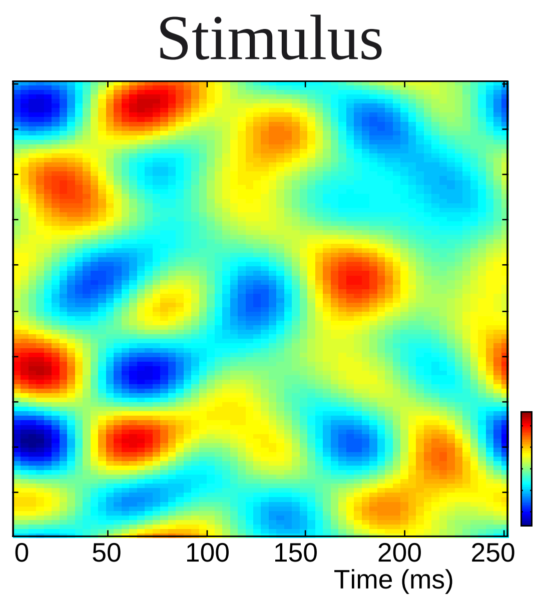

Impulse Responses,
*parametrized by spectral band*




Spectral Response Fields,
*evolving in time*

# Interpreting STRFs



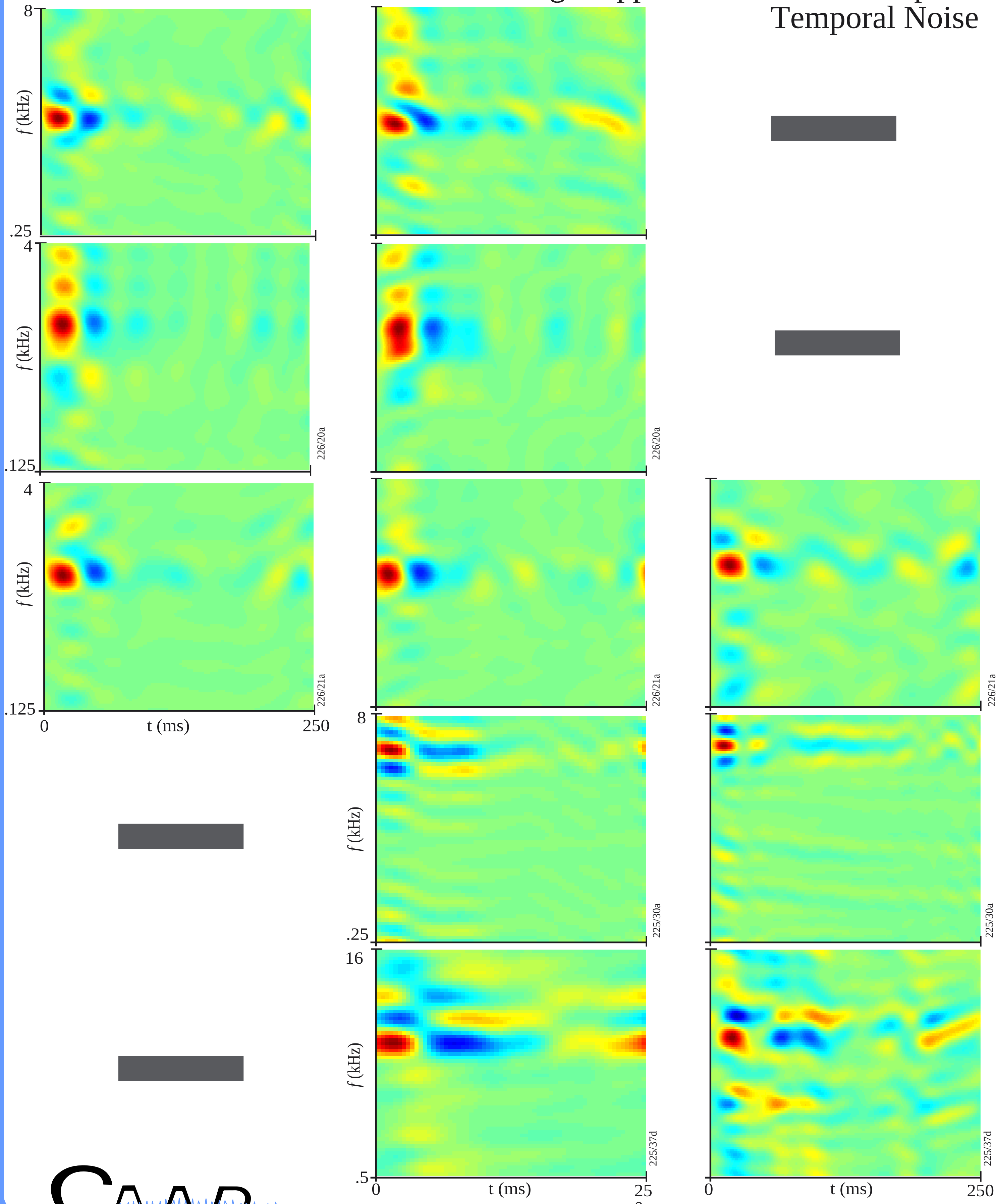

## Stimulus Effect on Rate

| STRF region | Stimulus Power | Spike rate contribution |
|---|---|---|
| Excitatory | Enhanced | Faster |
| Inhibitory | Enhanced | Slower |
| Excitatory | Diminished | Slower |
| Inhibitory | Diminished | Faster (!) |

*Animation Demo*

# STRFs Compared

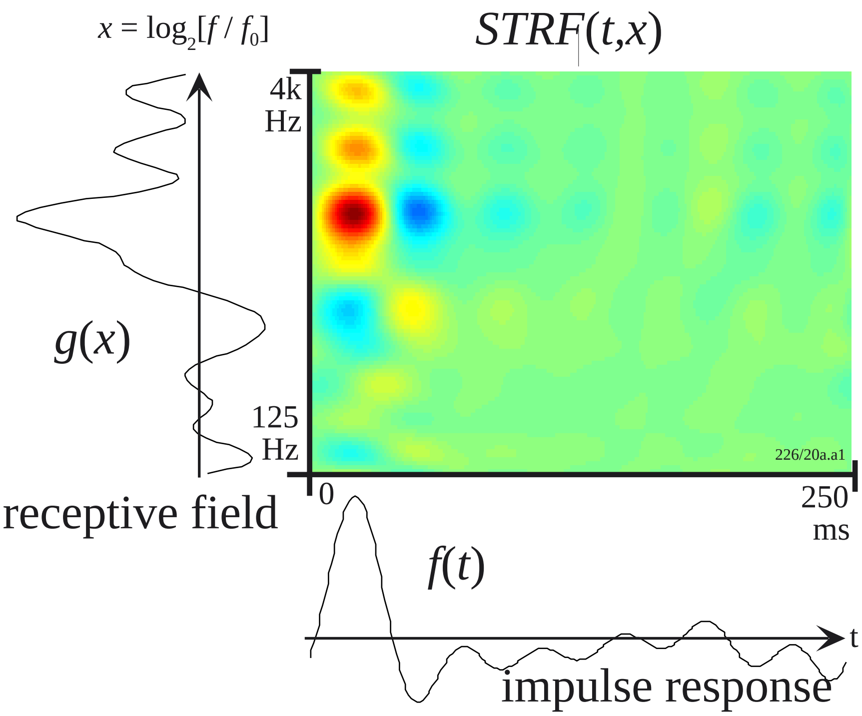
STRF from TORCs | STRF from single ripples | STRF from Spectro-Temporal Noise




Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

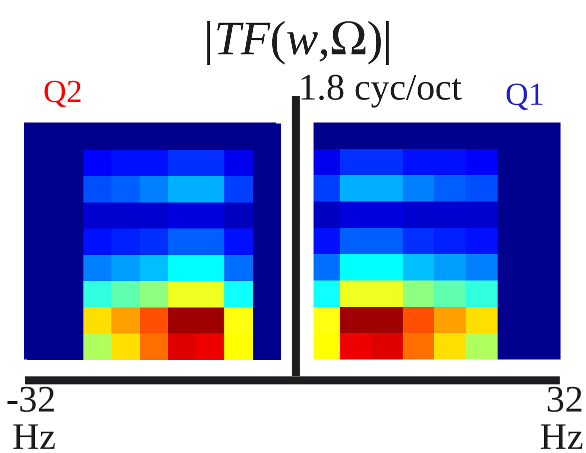
# Full Separability




The STRF is a simple product of a single spectral response function with a single temporal response function.

$$STRF(t,x) = f(t)\, g(x)$$

$|\mathcal{F}\{\ \}| \updownarrow$




| | |
|---|---|
| $f(t)$ | $g(x)$ |
| $\updownarrow$ | $\updownarrow$ |
| $F(w)$ | $G(\Omega)$ |

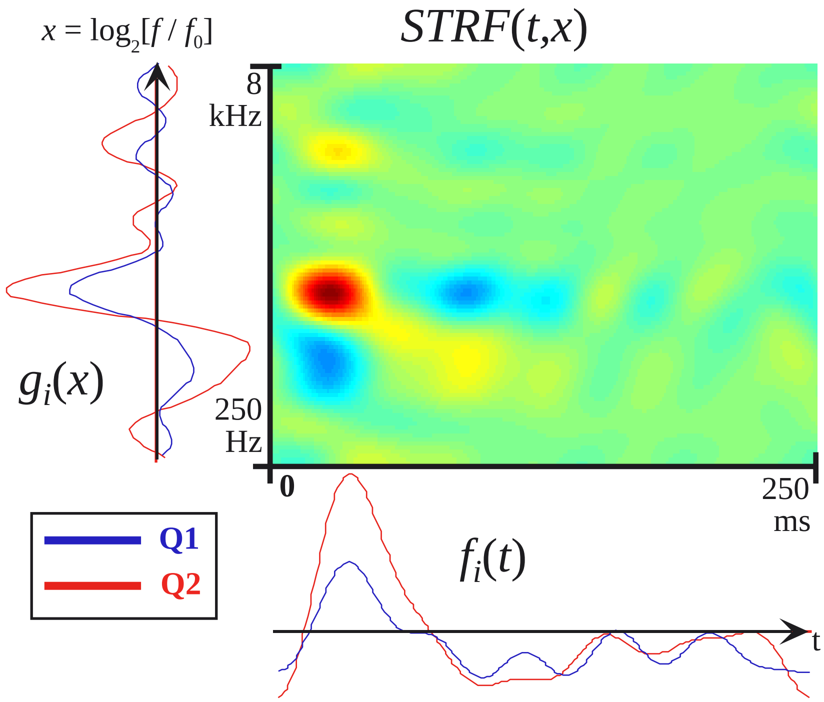
$$TF(w,\Omega) = F(w)\, G(\Omega)$$

Therefore the TF is also a simple product


Center for Auditory and Acoustic Research

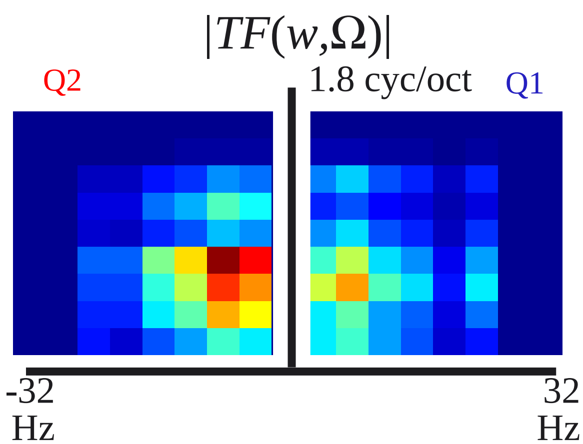
Institute for Systems Research
University of Maryland

# Quadrant Separability




The STRF is not separable, but each quadrant of the transfer function is, i.e., there are different spectral and temporal responses for upwards and downwards frequency modulation.

$|\mathcal{F}\{\ \}|$ ↕




| | |
|---|---|
| $f_i(t)$ | $g_i(x)$ |
| ↕ | ↕ |
| $F_i(w)$ | $G_i(\Omega)$ |

$$T(w,\Omega) = \begin{cases} F_1(w)\, G_1(\Omega) & w > 0, \Omega > 0 \\ F_2(w)\, G_2(\Omega) & w < 0, \Omega > 0 \end{cases}$$

for $\Omega > 0$: $T(w,\Omega) = T^*(-w,-\Omega)$

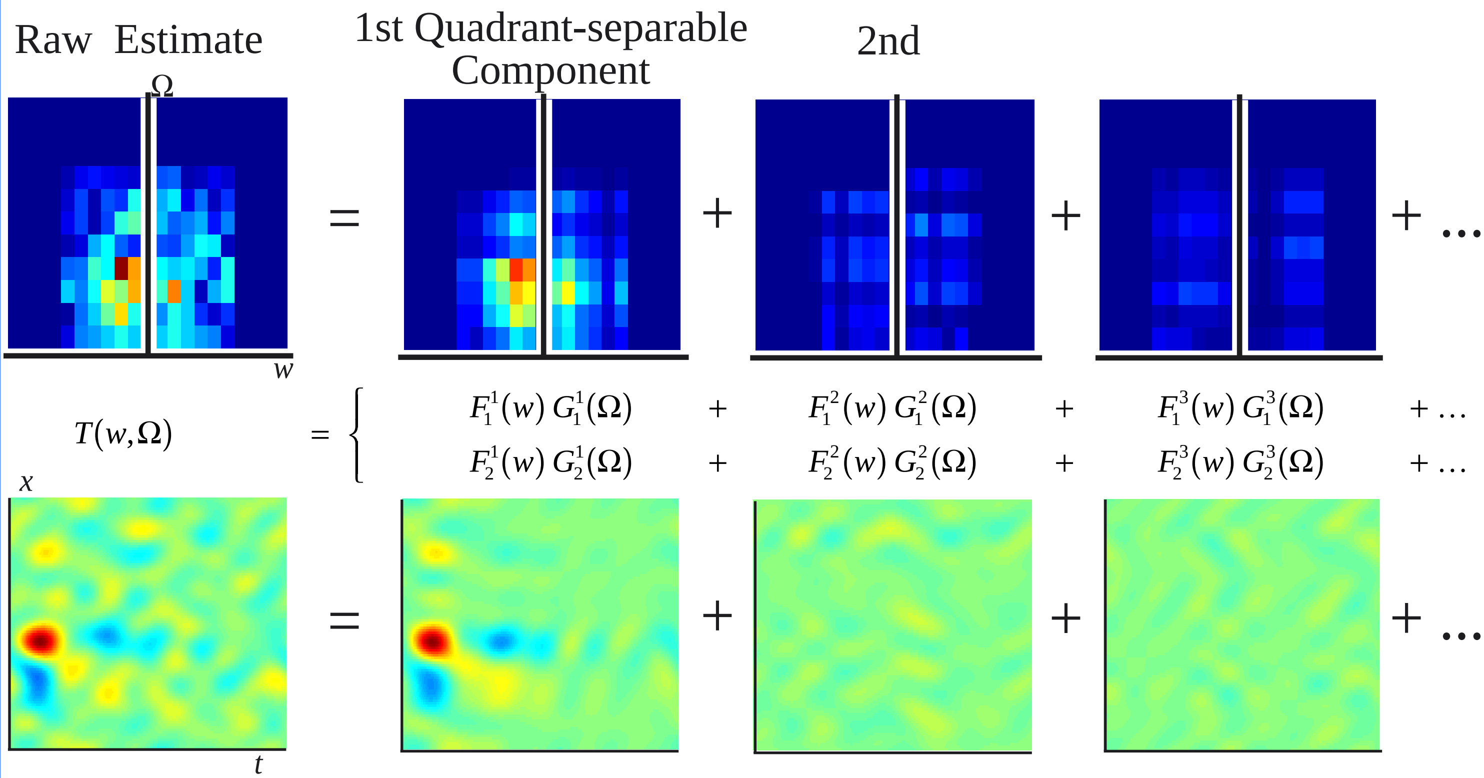

Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

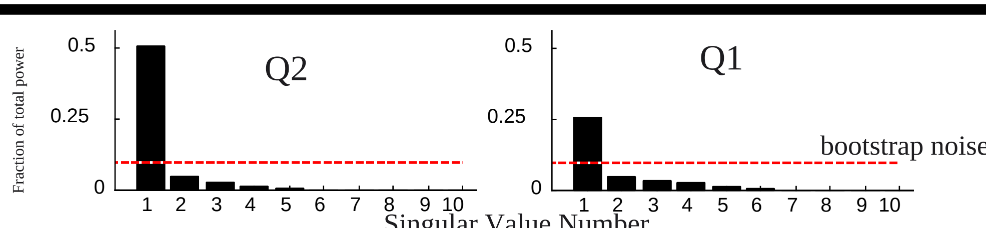
# Measuring Separability with SVD

- Singular Value Decomposition (SVD) can be used to estimate the separability of a Transfer Function (possibly corrupted by noise). It decomposes the Transfer Function into a sum of Quadrant Separable Transfer Functions, ordered by their power.

- Large jumps in the singular values separate signal from noise (& straddle bootstrap estimate of noise).




**All units measured in AI are quadrant separable (or fully separable).**





Center for Auditory and Acoustic Research
Institute for Systems Research
University of Maryland

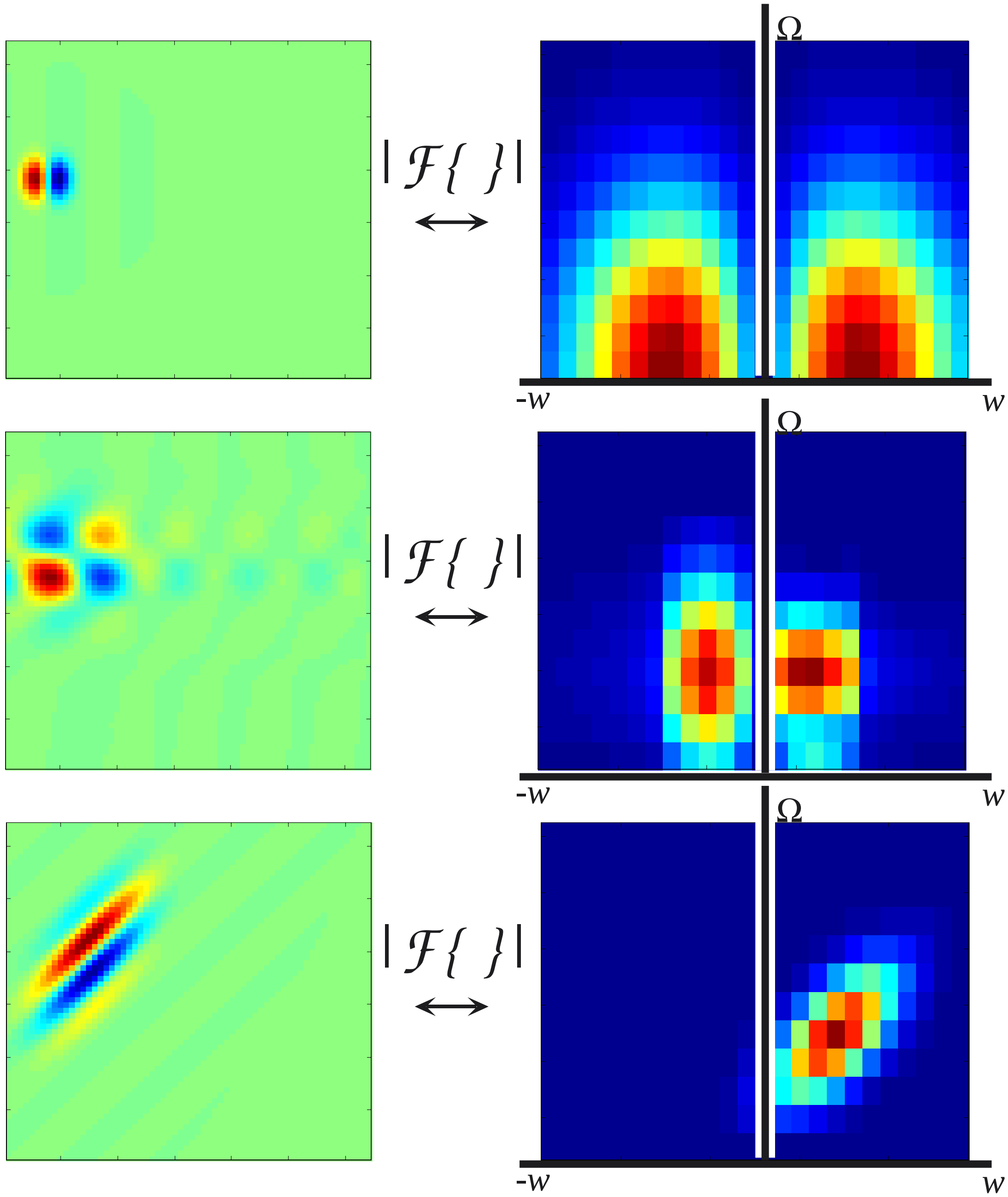
- Introduction to Auditory Spectro-Temporal Processing
- Stimuli—Dynamic & Broadband: Ripples, Ripple Combinations
- Examples & Properties of Spectro-Temporal Response Fields
- **Constraints on Neural Connectivity**

# Separability Examples

Fully Separable

Quadrant Separable

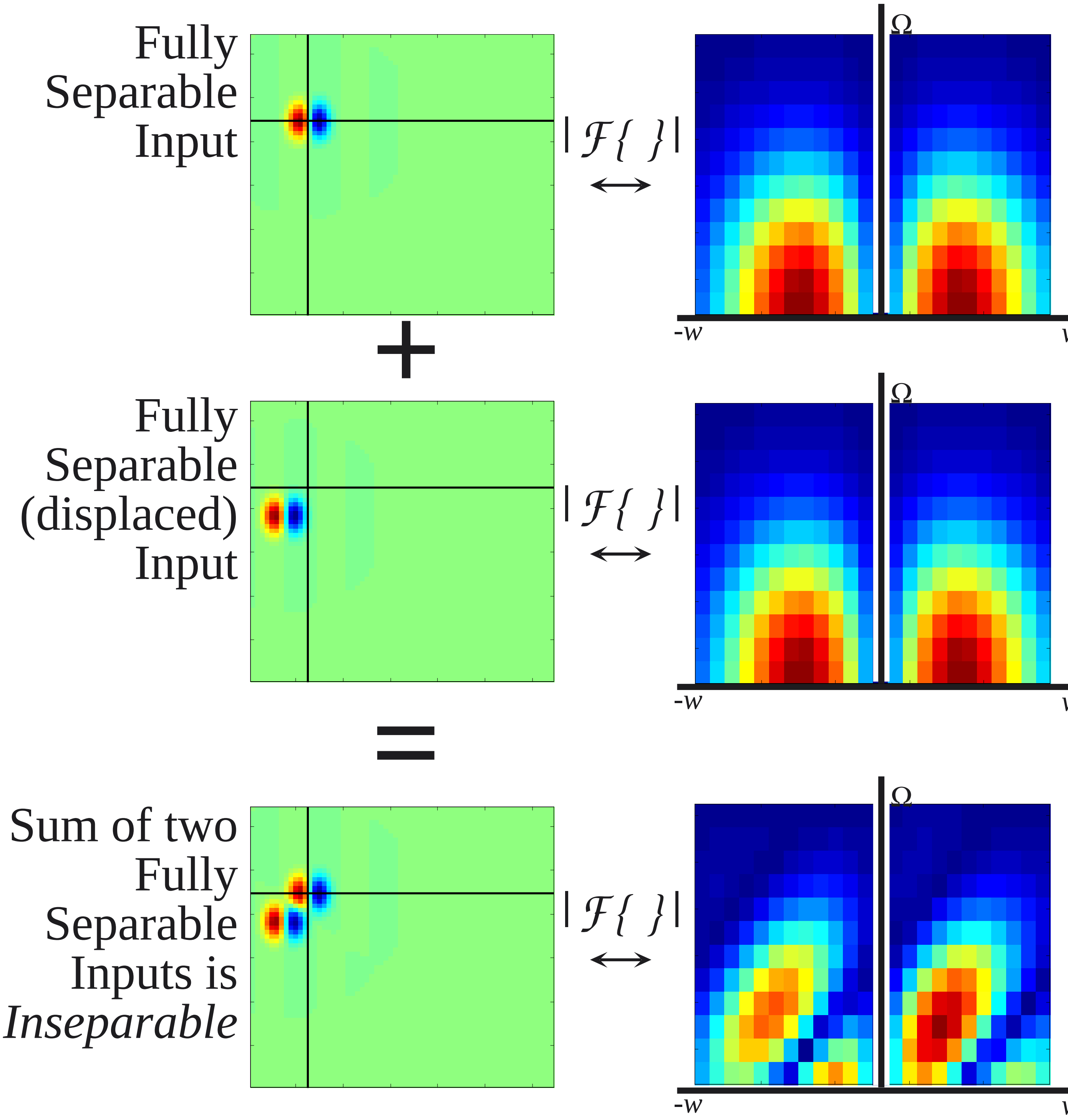
Velocity Selective is *Inseparable*




Quadrant separability is incompatible with velocity selectivity.

*Simulation*

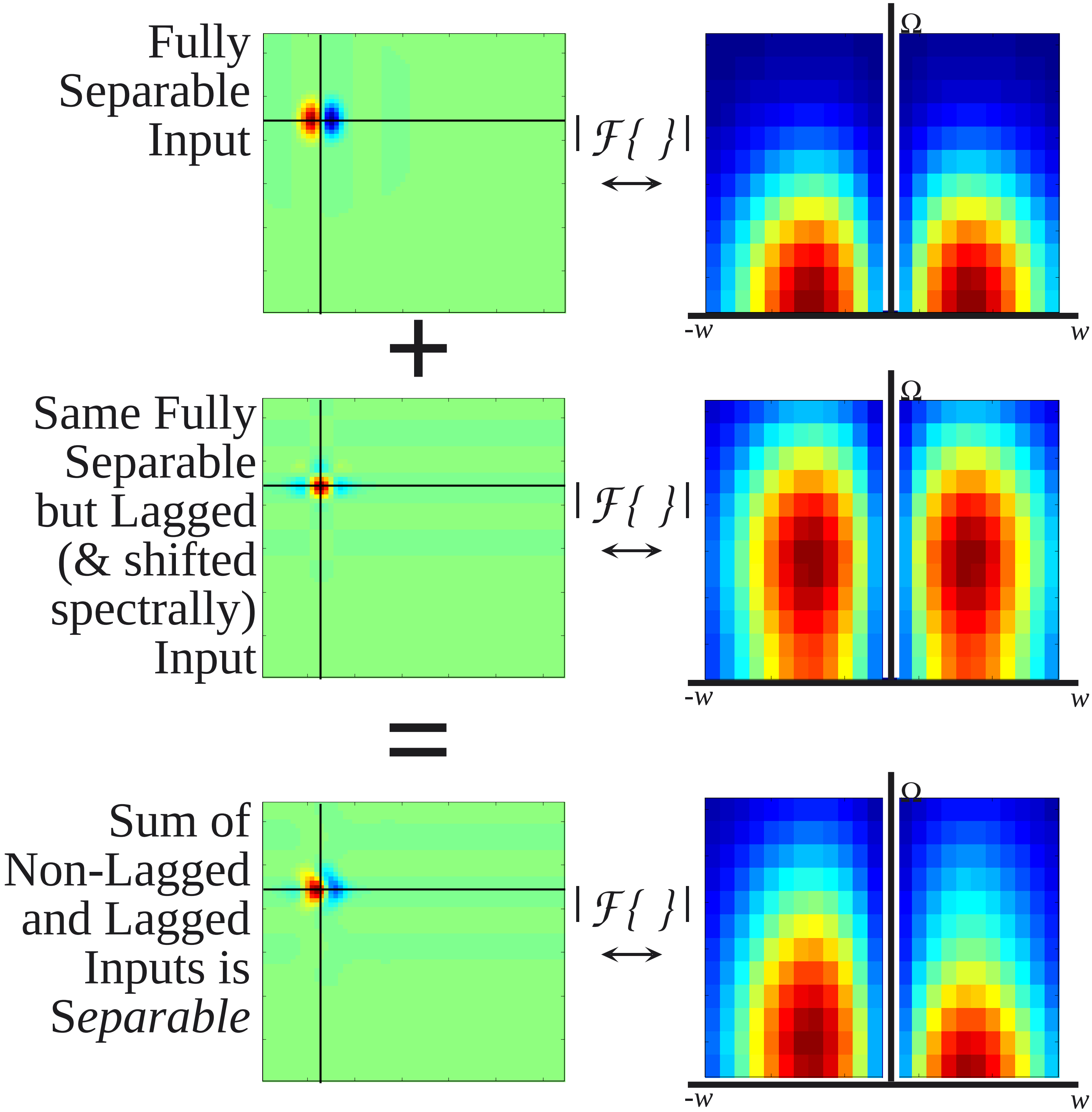
# Separability Counterexample




Naive sum of two fully separable input STRFs is inseparable.

*Simulation*

# Separability Example




Sum of two fully separable input STRFs is separable if the temporal processing is in quadrature.

*Simulation*

Center for Auditory and Acoustic Research

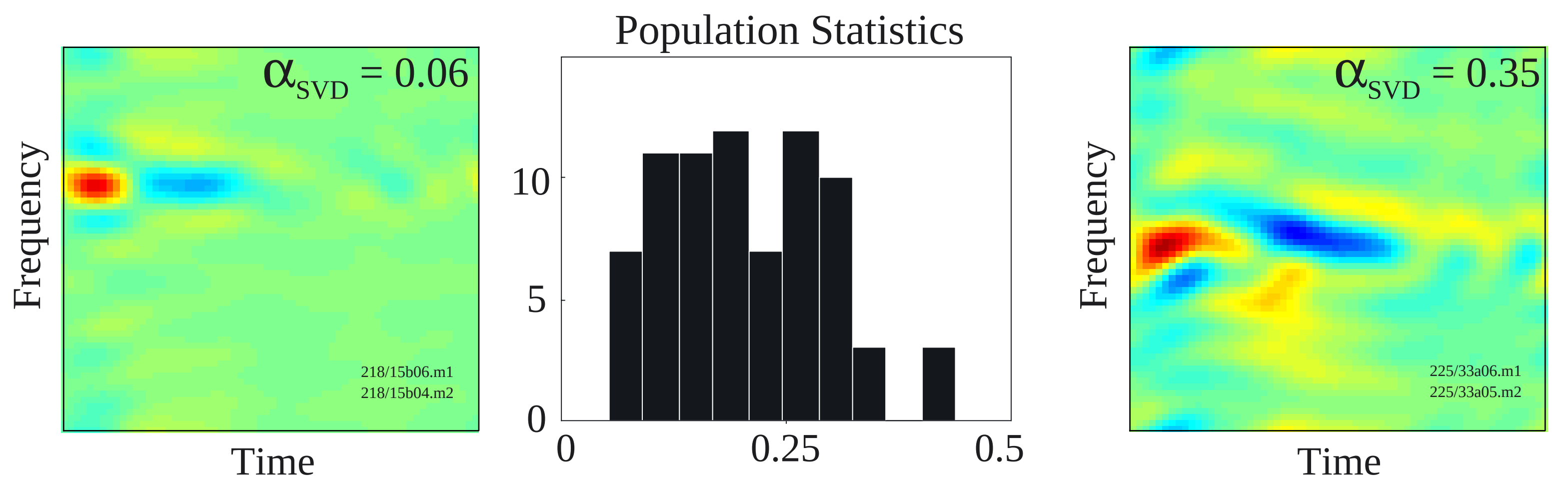
Institute for Systems Research
University of Maryland

# Measure of Separability

- SVD supplies a natural measure of separability, $\alpha_{SVD}$

$$\alpha_{SVD} = \left(1 - \frac{\lambda_1^2}{\sum_i \lambda_i^2}\right)$$

- $\alpha_{SVD} \approx 0$ is fully separable
- $\alpha_{SVD} > 0.3$ is strongly inseparable




Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

# Symmetry by Power

- $\alpha_d$: Power asymmetry breaks full separability, producing quadrant separability

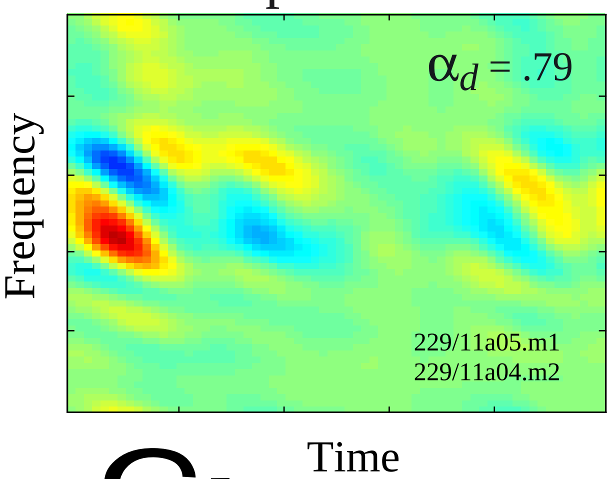
$$\alpha_d = (P_1 - P_2)/(P_1 + P_2)$$

$$P_1 = (\text{Power in quadrant 1}) = (\lambda_1)^2$$

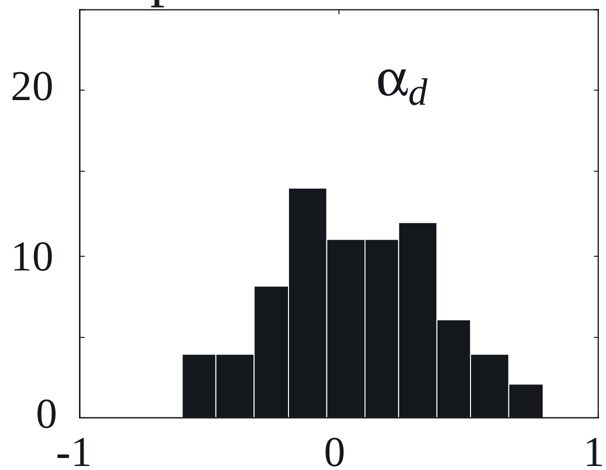
$$P_2 = (\text{Power in quadrant 2}) = (\lambda_2)^2$$

- $\alpha_d \approx 0$ is symmetric in power
- $|\alpha_d| > 0.3$ is quite asymmetric in power—strongly inseparable

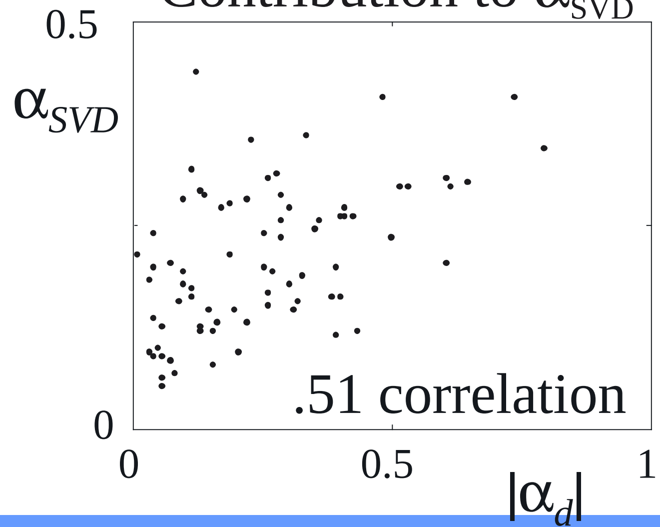
Example STRF




Population Statistics




Contribution to $\alpha_{SVD}$



# Spectral Symmetry

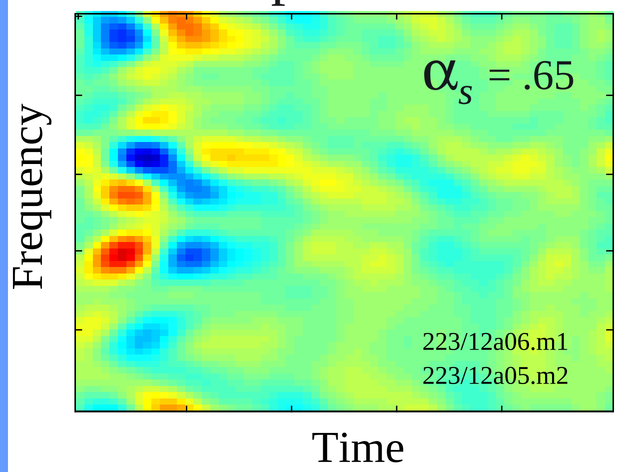
- $\alpha_s$: Asymmetry between spectral cross-sections $G_i(\Omega)$:

$$\alpha_s = 1 - \left| \frac{\sum_{\Omega>0} G_1(\Omega)\, G_2^*(\Omega)}{\sqrt{\sum_{\Omega>0} |G_1(\Omega)|^2 \; |G_2(\Omega)|^2}} \right|$$

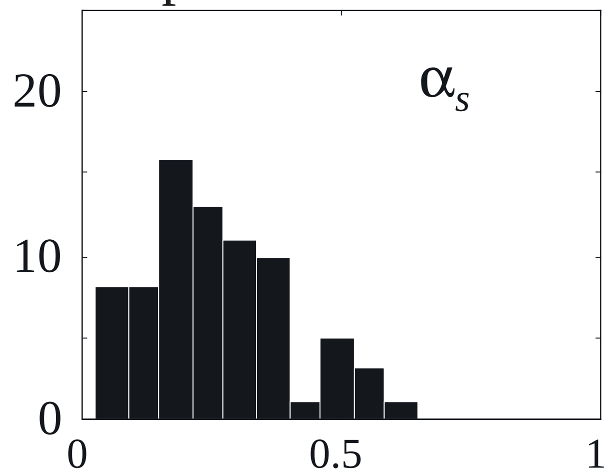
where the quantity inside the big absolute value bars is the (complex) correlation between $G_1(\Omega)$ and $G_2(\Omega)$

- $\alpha_s \approx 0$ is spectrally symmetric
- $\alpha_s > 0.3$ is spectrally asymmetric—strongly inseparable

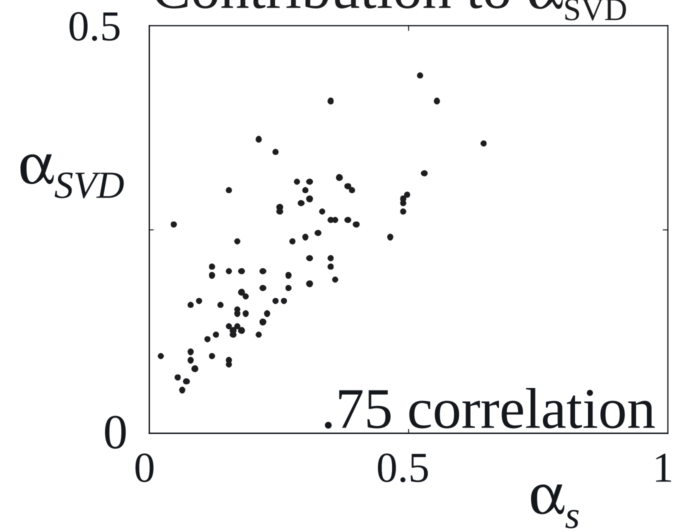
Example STRF




Population Statistics




Contribution to $\alpha_{SVD}$





Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

# Temporal Symmetry

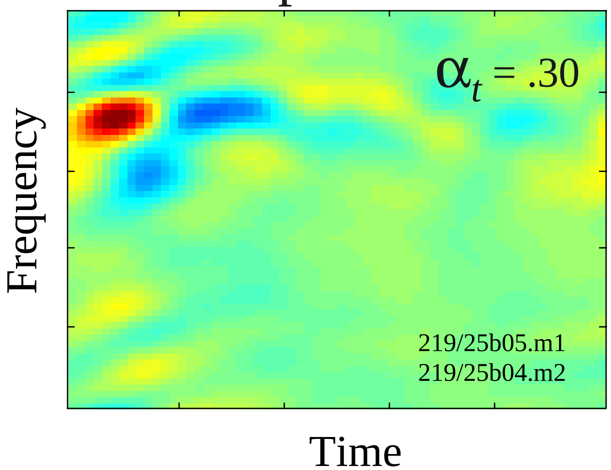
- $\alpha_t$: Asymmetry between temporal cross-sections $F_i(w)$:

$$\alpha_t = 1 - \left| \frac{\sum_{w>0} F_1(w)\, F_2(-w)}{\sqrt{\sum_{w>0} |F_1(w)|^2\, |F_2(-w)|^2}} \right|$$

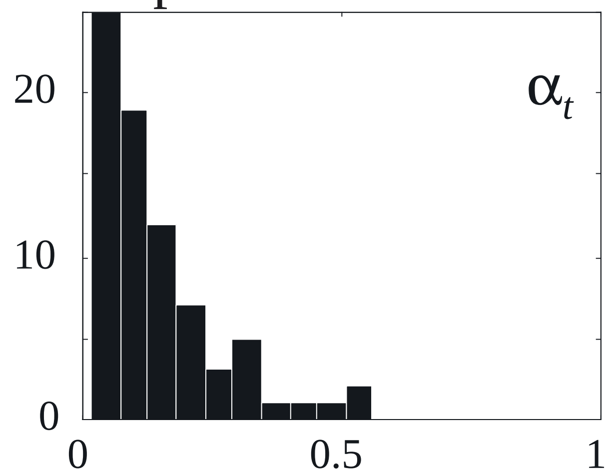
where the quantity inside the big absolute value bars is the (complex) correlation between $F_1(w)$ and $F_2^*(-w)$

- $\alpha_t \approx 0$ is temporally symmetric
- $\alpha_t > 0.3$ is temporally asymmetric—strongly inseparable

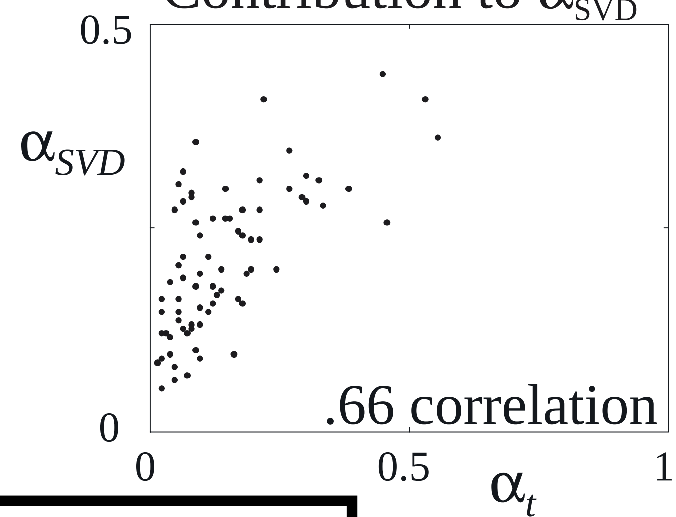
Example STRF




Population Statistics




Contribution to $\alpha_{SVD}$




**Distribution is strongly skewed toward temporal symmetry.**

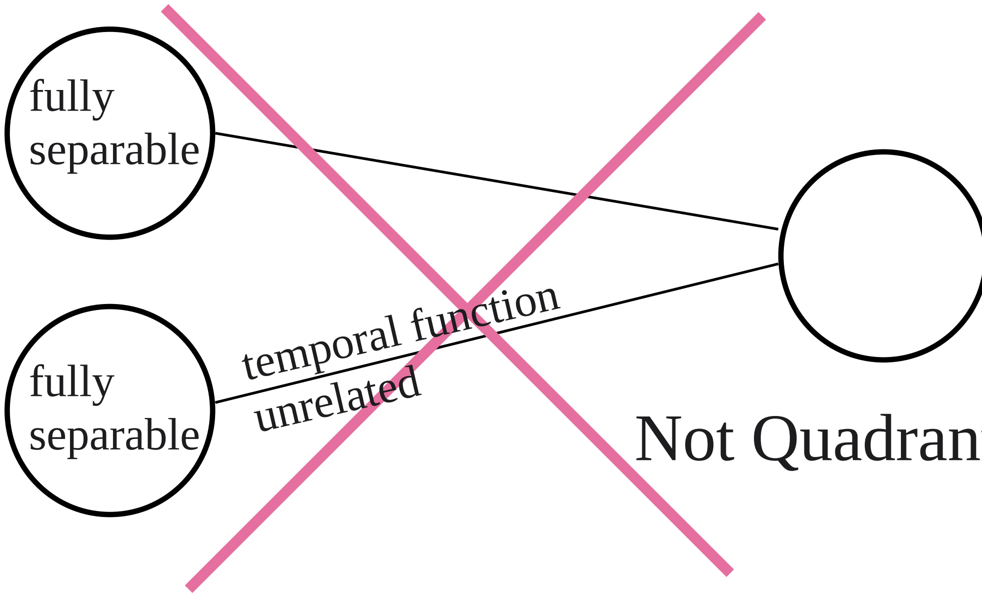

Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

# Neural Connectivity Constraints

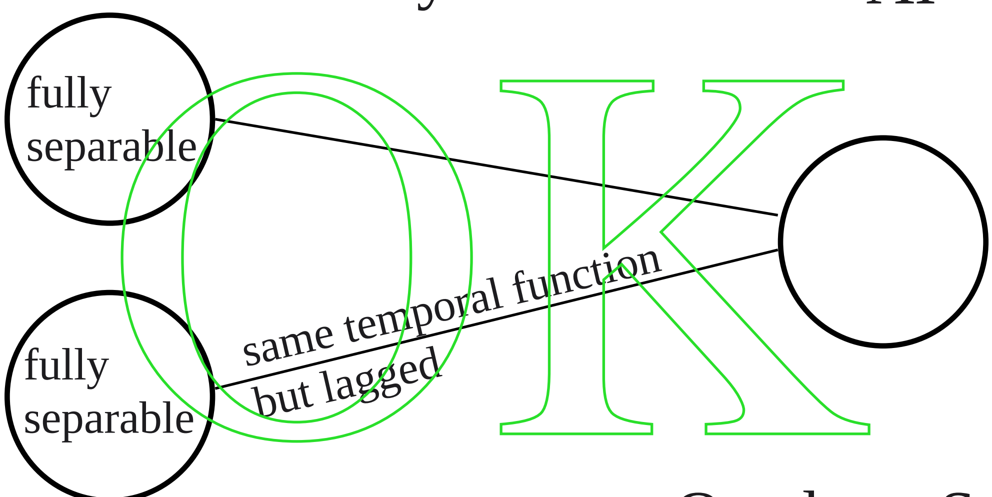
Input to AI =
Medial Geniculate Body

AI




Input to AI =
Medial Geniculate Body

AI




Quadrant Separable

Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

# Summary

- **The function of AI**
  To encode spectro-temporal features of sounds
  spectrally: to ~1 cycles/octave
  temporally: ~2 to ~20 Hz (in ferret)

  *plus encoding other sound features not addressed here*

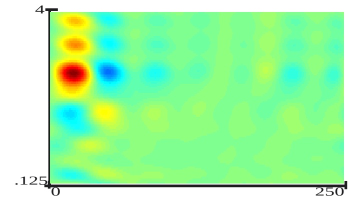
- **Spectro-Temporal Response Field (STRF)**
  - **Descriptor** of response to broadband dynamic stimuli
  - **Predictor** of spike train for stimuli of dynamic, spectral modulations of noise
    - STRFs agree despite measurement method
    - Linear processing conveys most of the information
  - **Visual Tool** conveys spectro-temporal regions of excitation and inhibition

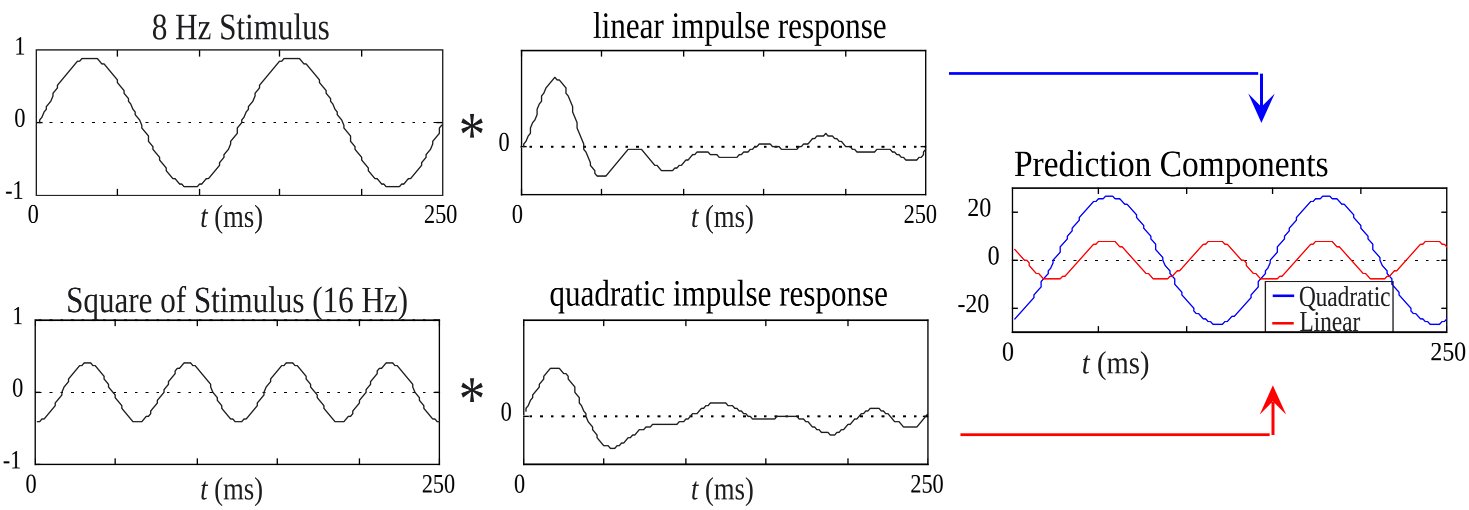
- **Constraints of Quadrant Separability**
  - Limits possible network dynamics

- Introduction to Auditory Spectro-Temporal Processing
- Stimuli—Dynamic & Broadband: Ripples, Ripple Combinations
- Examples & Properties of Spectro-Temporal Response Fields
- Constraints on Neural Connectivity
- MagnetoEncephaloGraphy (MEG)
- Predicting Responses to Novel Stimuli
- **Non-Linearities**

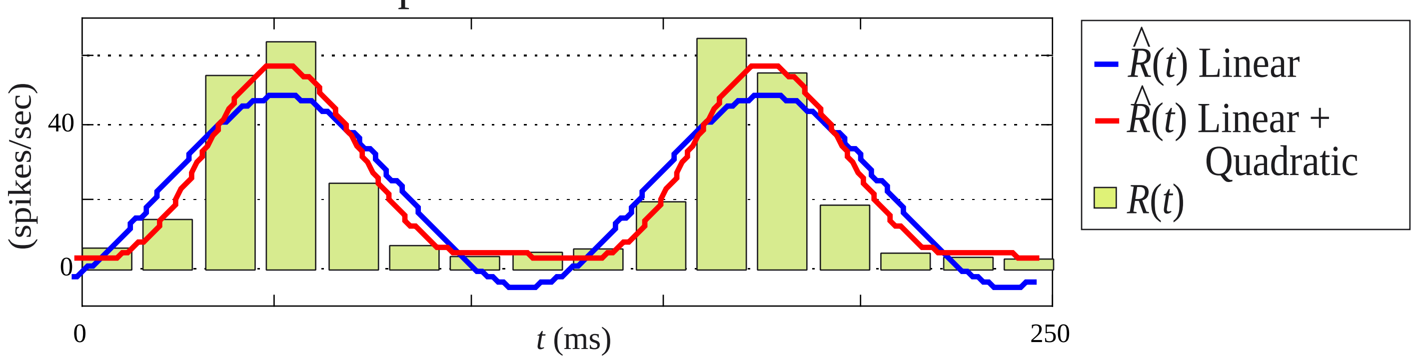
# Non-Linearity—Predictions

- Preliminary results indicate that the non-linear predictions fit the responses more accurately than the linear predictions, although the differences between the two are typically subtle.

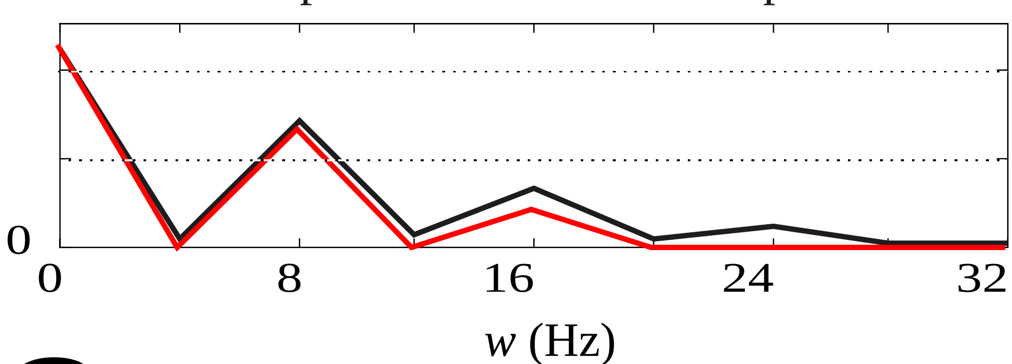



Response and Predictions




Response and Prediction Spectra

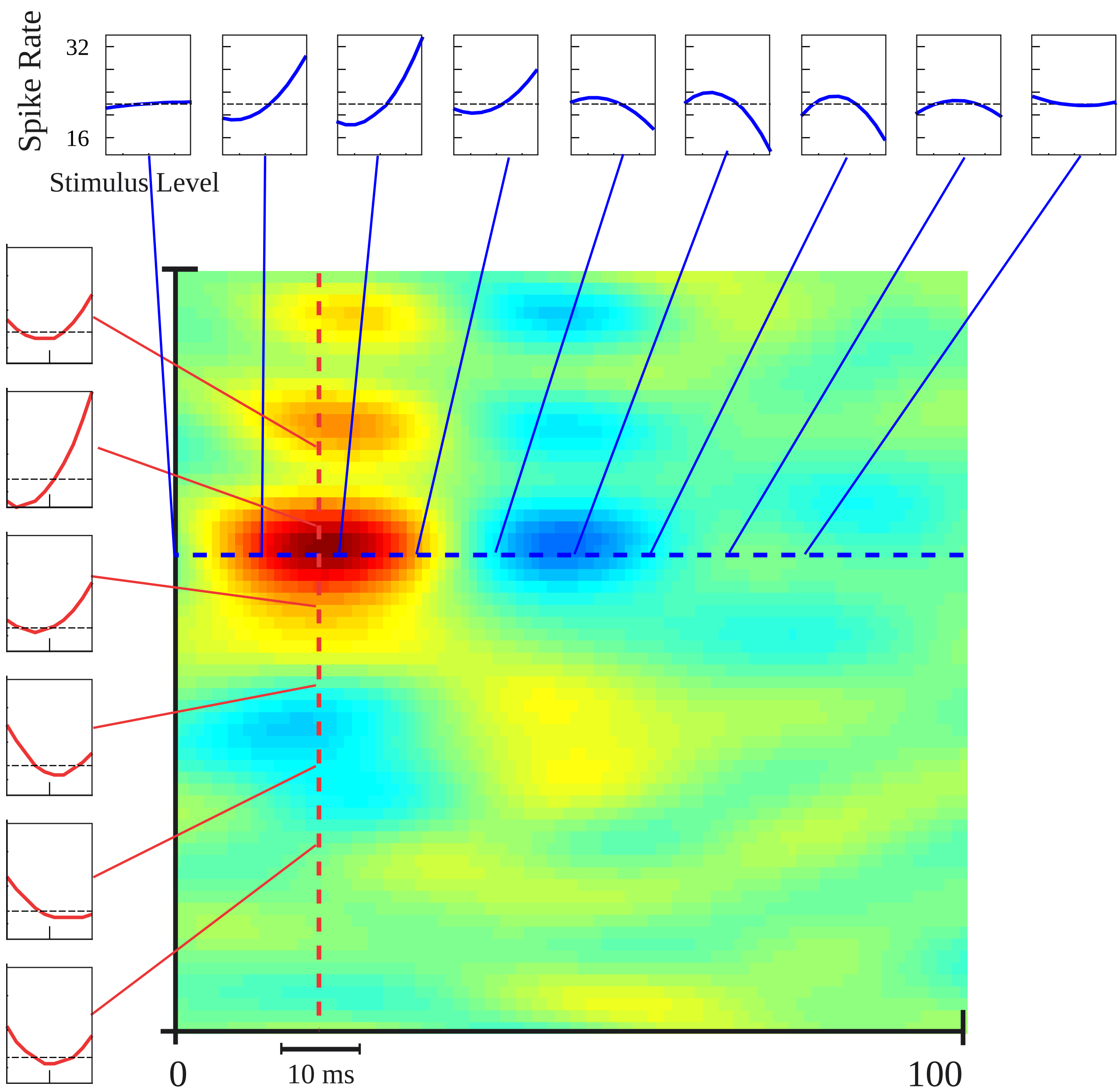



Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

# Spectro-Temporal Rate-Level Functions

Rate-level functions change with τ and *x*.

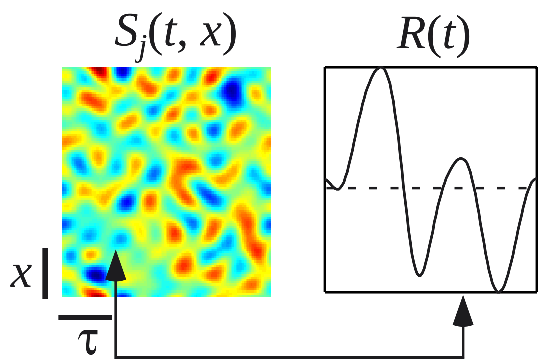


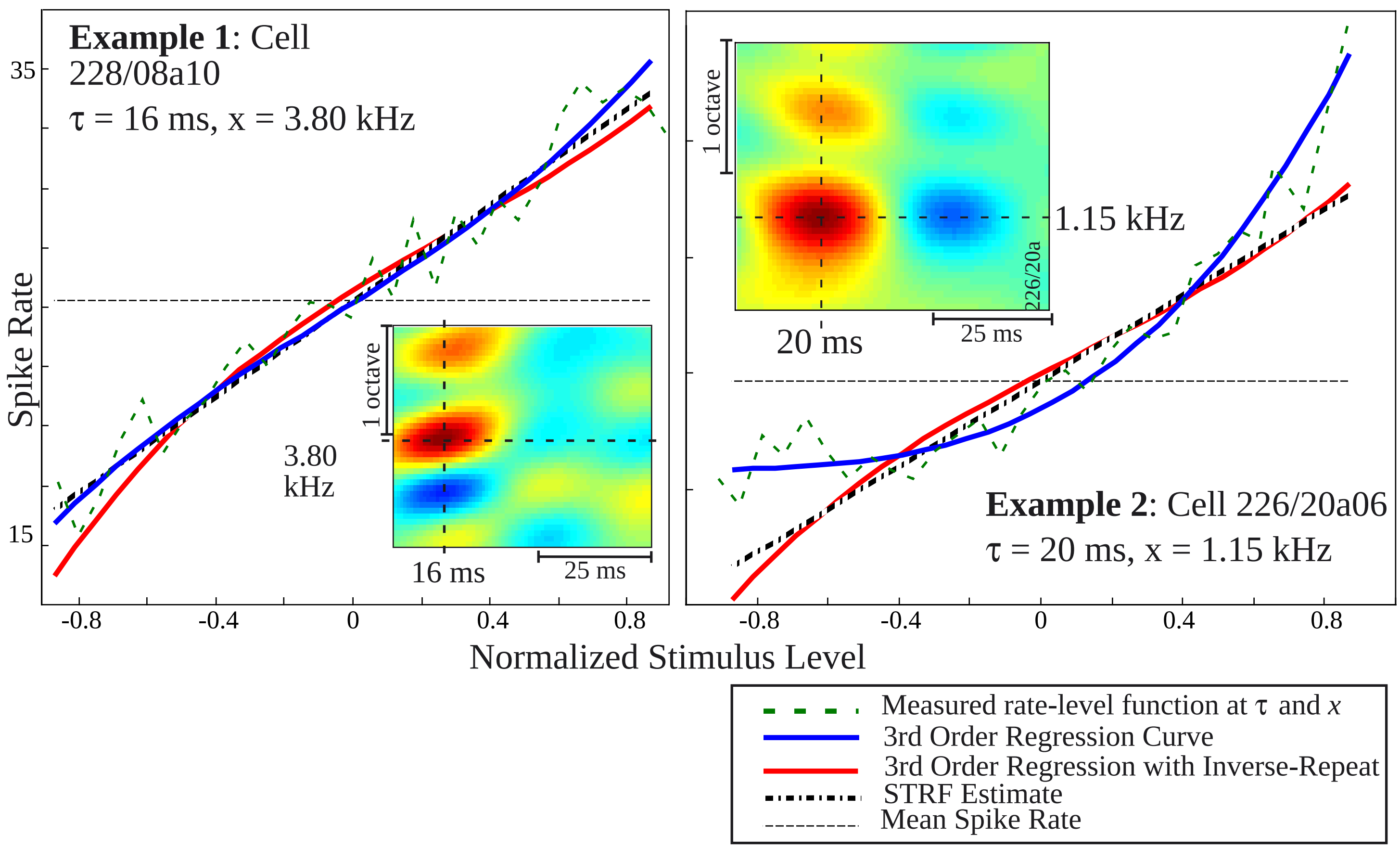

Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

# Non-Linearity—Theory

- The value of the STRF at each point ($\tau$, $x$) is the slope of a linear rate-level function: $R_{\tau,x}(t) = [\mathrm{STRF}(\tau, x)] \cdot S(t-\tau, x)$ .
- Polynomial rate-level curves measured at every ($\tau$, $x$) improve the description. These are potentially non-linear functions.




- **Using cubic polynomials, we have shown that either the non-linearities are absent, or they are dominantly second order.**
- **Subtraction of the response to the inverted envelope gives a nearly linear polynomial fit. This would be expected, for example, from a purely even order (e.g., rectifying non-linearity).**


Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland