Linear and Non-Linear Responses to Dynamic Broad-Band Spectra in Primary Auditory Cortex

D. J. Klein[1,2]
S. A. Shamma[1,2]
J. Z. Simon[1]
D. A. Depireux[1]

[1]Institute for Systems Research
[2]Department of Electrical Engineering
University of Maryland

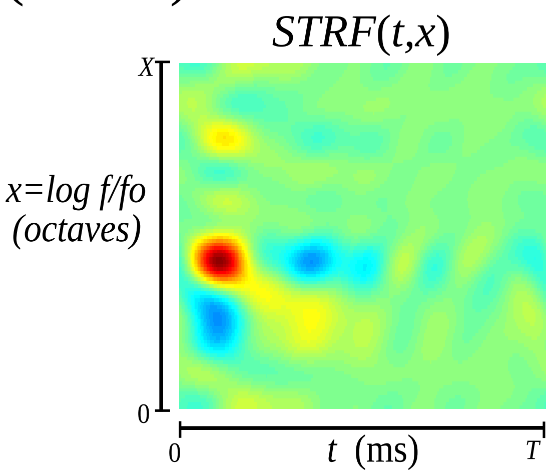
Supported in part by a MURI grant from the Office of Naval Research, a grant from the National Institute on Deafness and Other Communication Disorders and a grant from the National Science Foundation

Center for Auditory and Acoustic Research

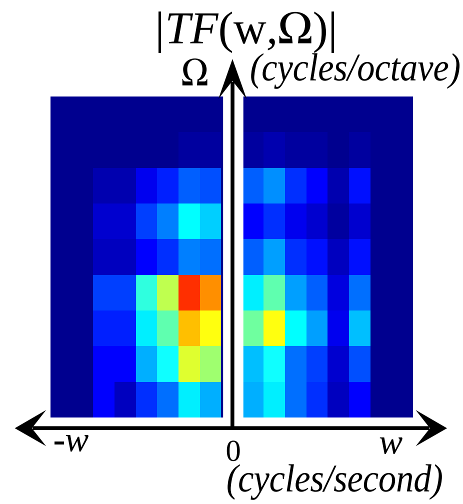
Institute for Systems Research
University of Maryland

By Way of Introduction

Cells in AI are well characterized their spectro-temporal response field (STRF)...



...or equivalently, in the Fourier domain, by the ripple transfer function (TF).

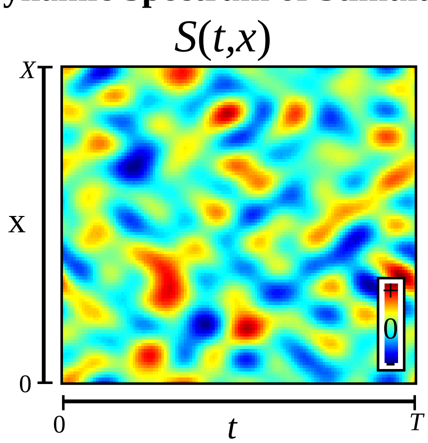


$|\mathcal{F}\{\ \}|$

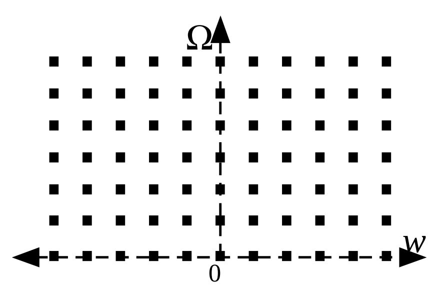
↔

The STRF can be measured by reverse-correlating with the dynamic spectrum of a broad-band stimulus.

Dynamic Spectrum of Stimulus



Ripple Spectrum



The stimuli should contain spectral-peak densities and modulation rates evenly distributed over a relevant range.

Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

Moving Ripples

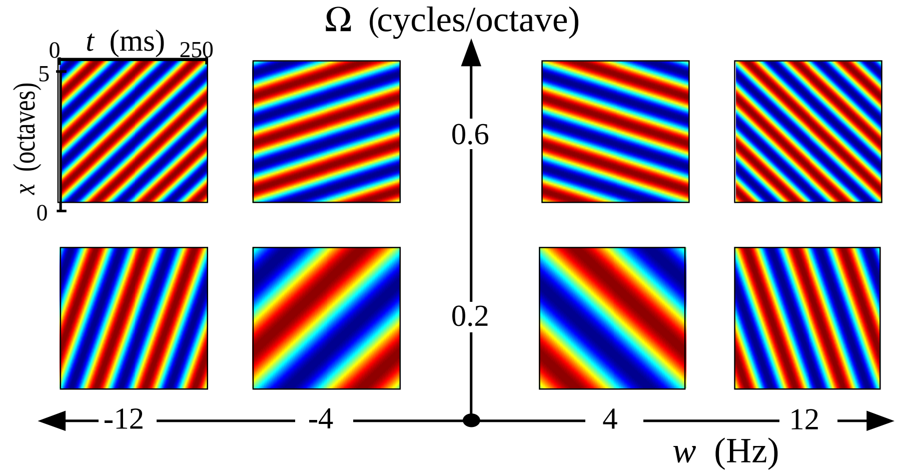
Moving ripples form our basis for the Fourier-domain description of dynamic spectra. At any time ***t*** and any frequency ***x***, their amplitude ***S*(*t*,*x*)** is given by:

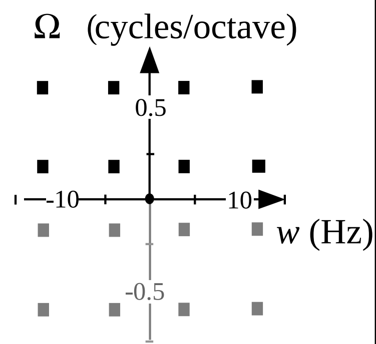
$$S(t,x) = \sin[2\pi wt + 2\pi\Omega x + \phi]$$

$x = \log_2[f / f_0]$

w = ripple velocity, modulation rate

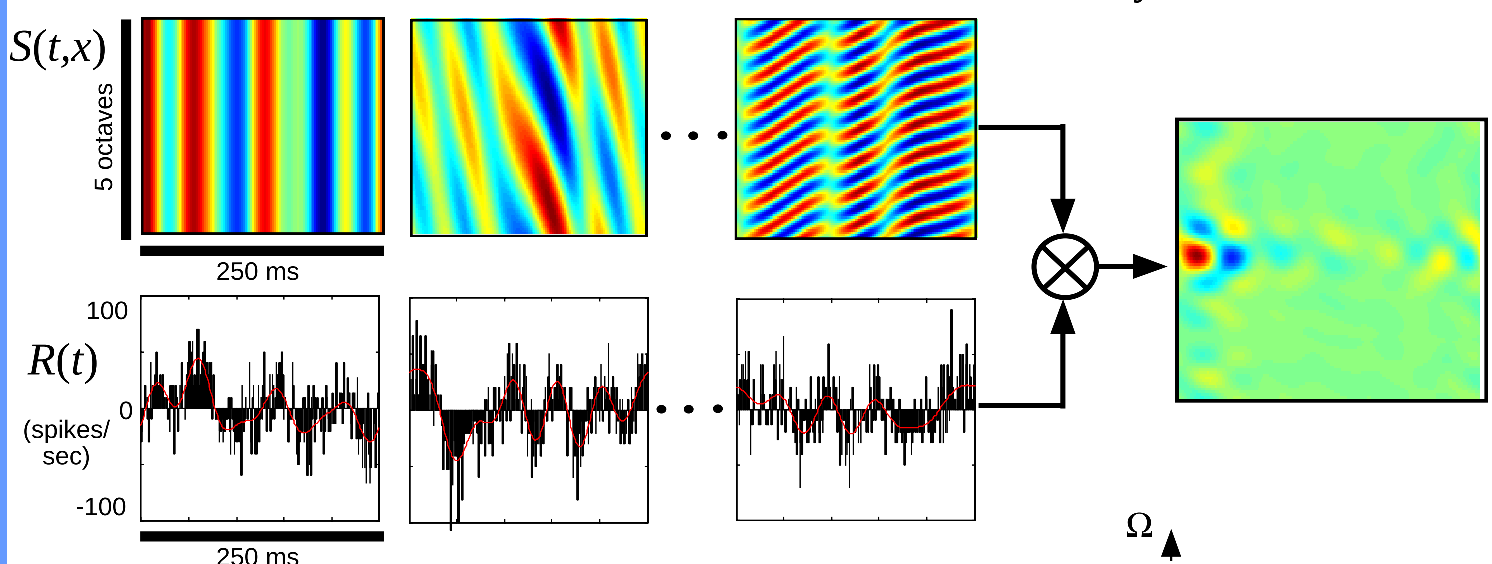
Ω = ripple frequency, spectral-peak density



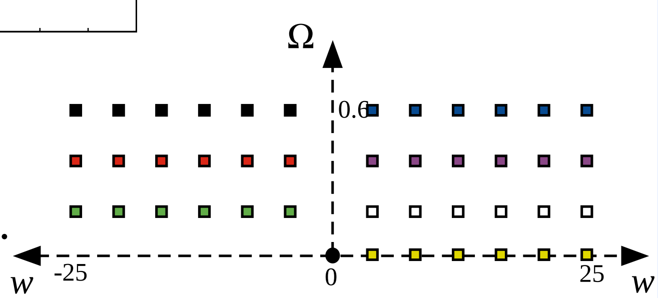


Temporally Orthogonal Ripple Combinations (TORCs)

TORCs are composed only of ripples with different modulation rates.
This allows us to obtain clean STRF estimates with relatively brief stimulation.



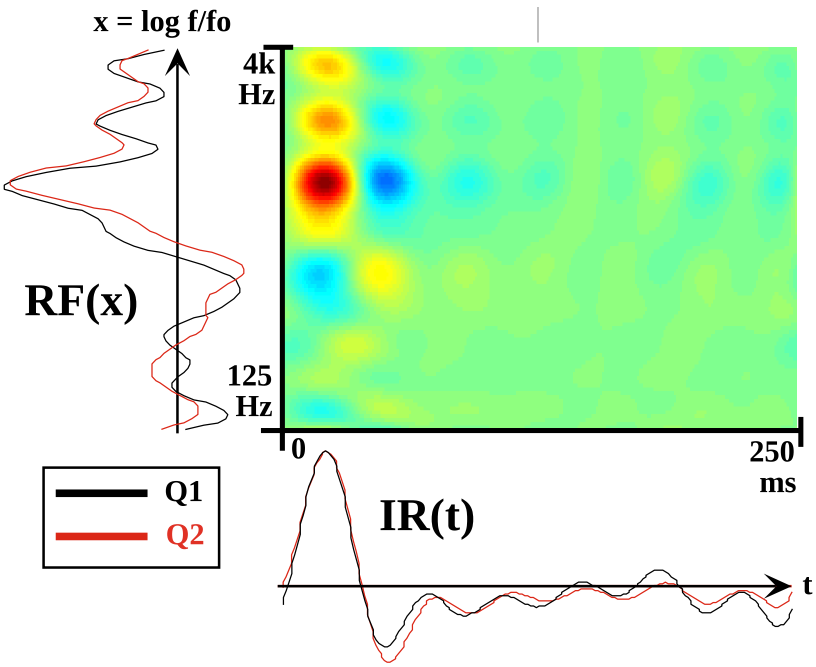
The stimuli shown here contain ripples which cover the same range of ripple velocities, but at different ripple frequencies.

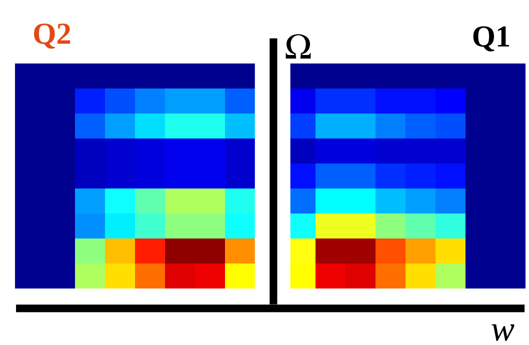


Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

Fully Separable STRF

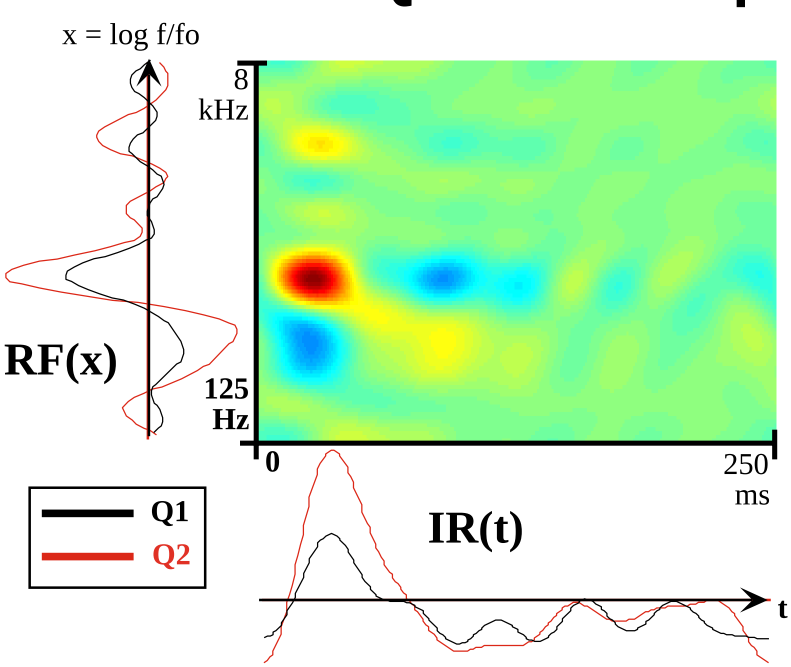


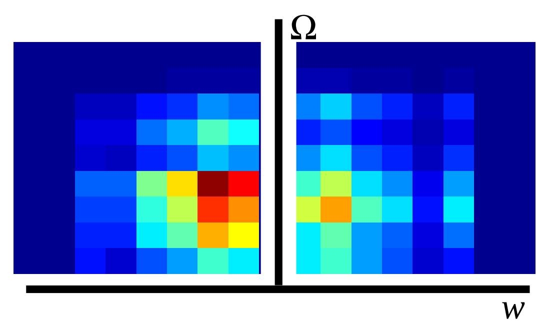


The STRF and TF are a product of a single spectral response function with a single temporal response function.

Shown above are the impulse responses (IR) and receptive fields (RF) derived from quadrant 1 (black) and quadrant 2 (red) of the transfer function by inverse Fourier transformation.

Quadrant Separable STRF





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.

This neuron responded twice as strong to rising frequencies than it did to falling frequencies.

Center for Auditory and Acoustic Research

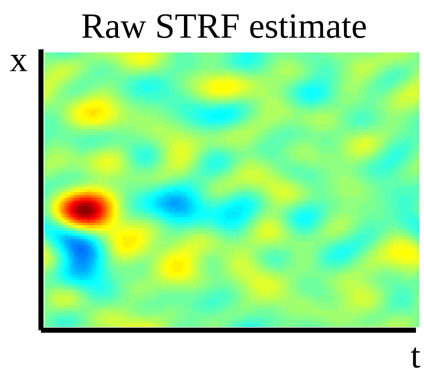
Institute for Systems Research
University of Maryland

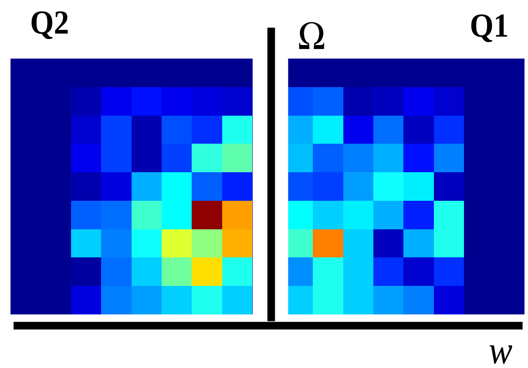
Singular Value Decomposition (SVD)
for Evaluation of Separability and for Noise Reduction

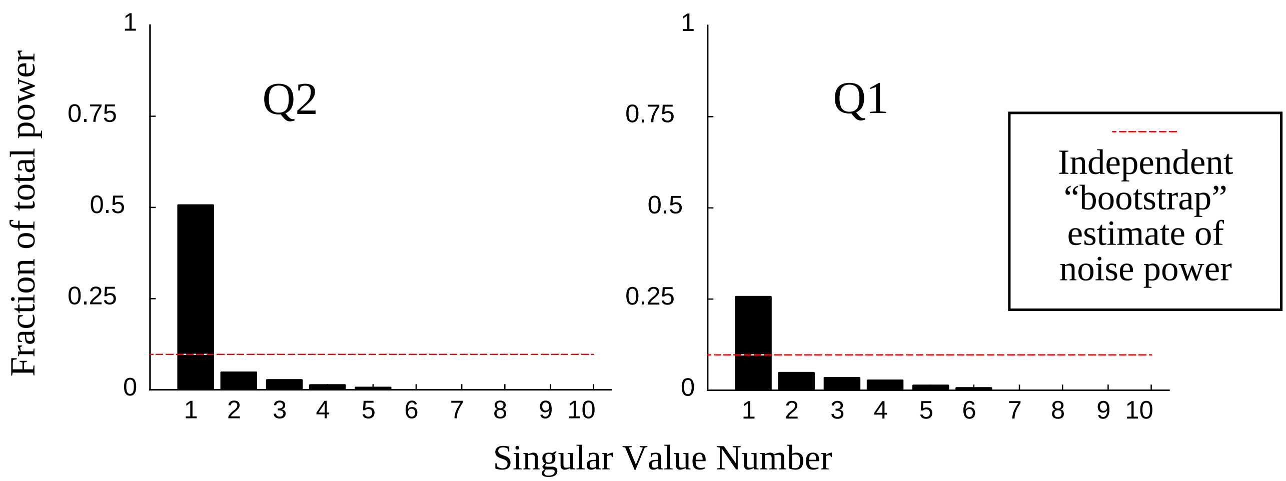
- If a quadrant of the transfer function is separable, every row is a scaled version of every other row. Such a matrix is said to have a rank of one.

- Singular Value Decomposition (SVD) can be used to estimate the rank of a matrix corrupted by noise. It decomposes the matrix into a sum of rank one (separable) matrices, ordered by their overall magnitudes. The first *k* components sum to a matrix of rank *k* which minimizes the power of the remaining components.

- We apply SVD to each quadrant of the transfer function.





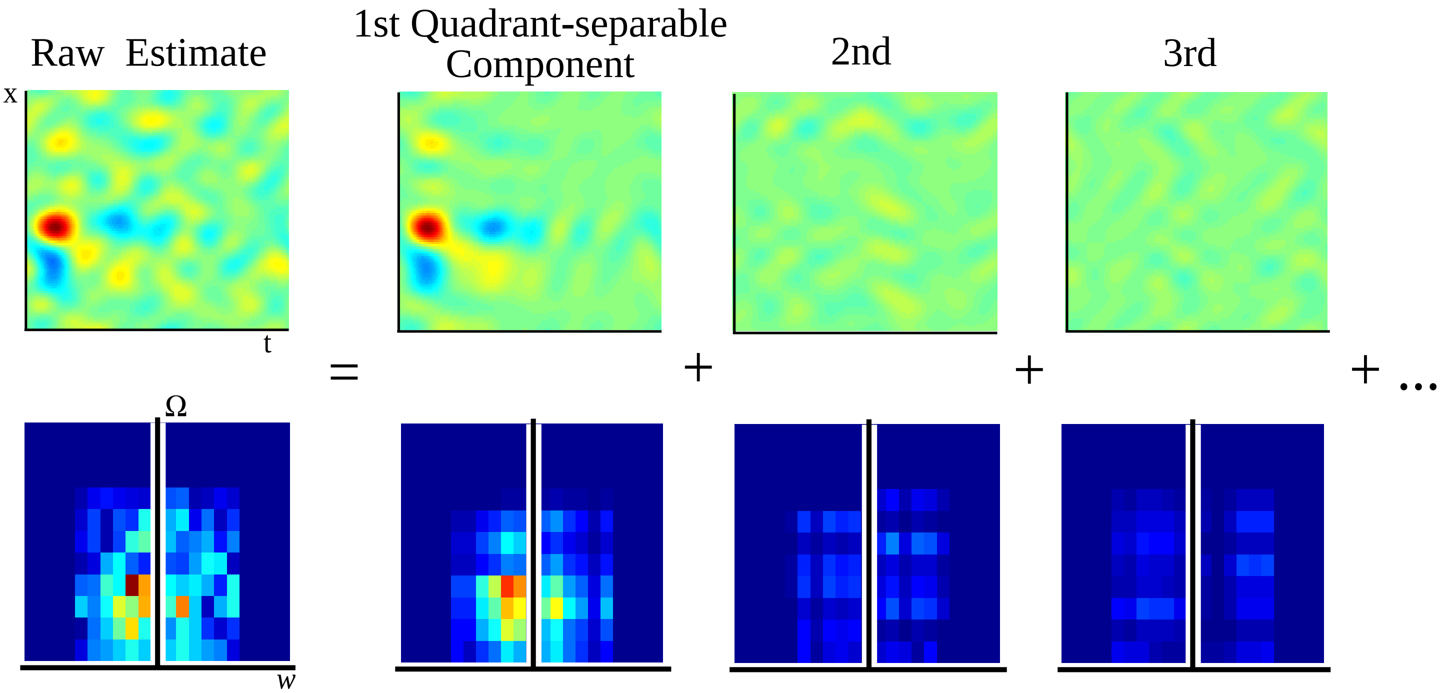


- SVD naturally separates the signal and noise components of a matrix. Typically, large jumps in the singular values indicate where the separation occurs. Noise is removed by discarding the lower-magnitude components.

Center for Auditory and Acoustic Research

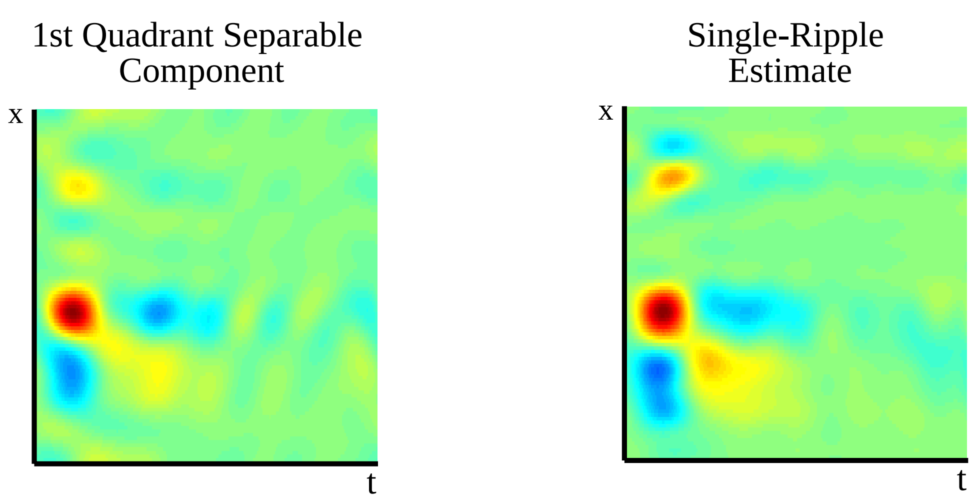
Institute for Systems Research
University of Maryland

Below, STRFs corresponding to the three most significant quadrant-separable components, derived from SVD, are shown.

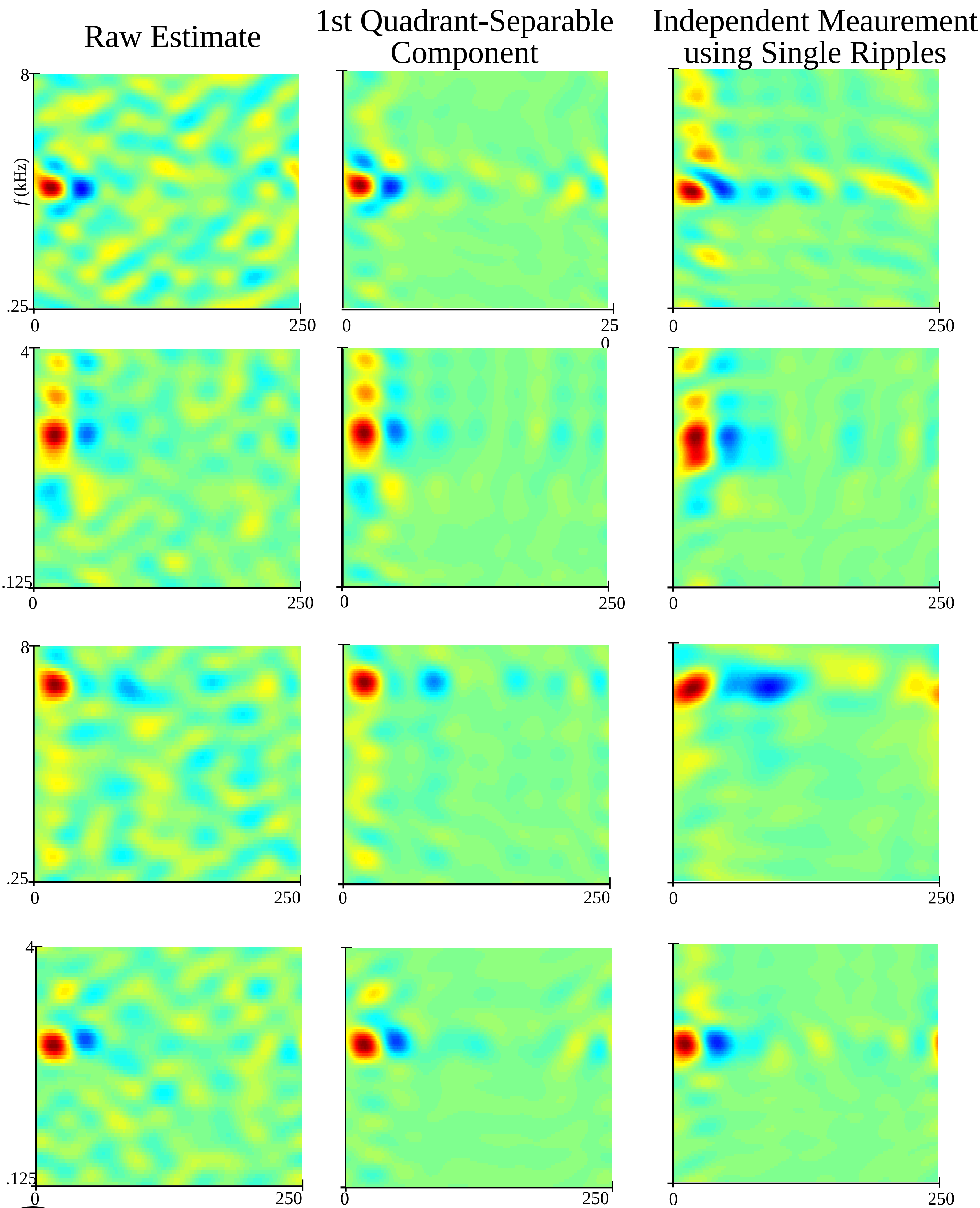


Without prior assumptions, SVD indicates that a large majority of STRFs in AI are quadrant separable.

This finding is further supported by the similarity between the first separable component and the STRF measured with a single-ripple method, for which quadrant separability had been assumed.



Additional Examples



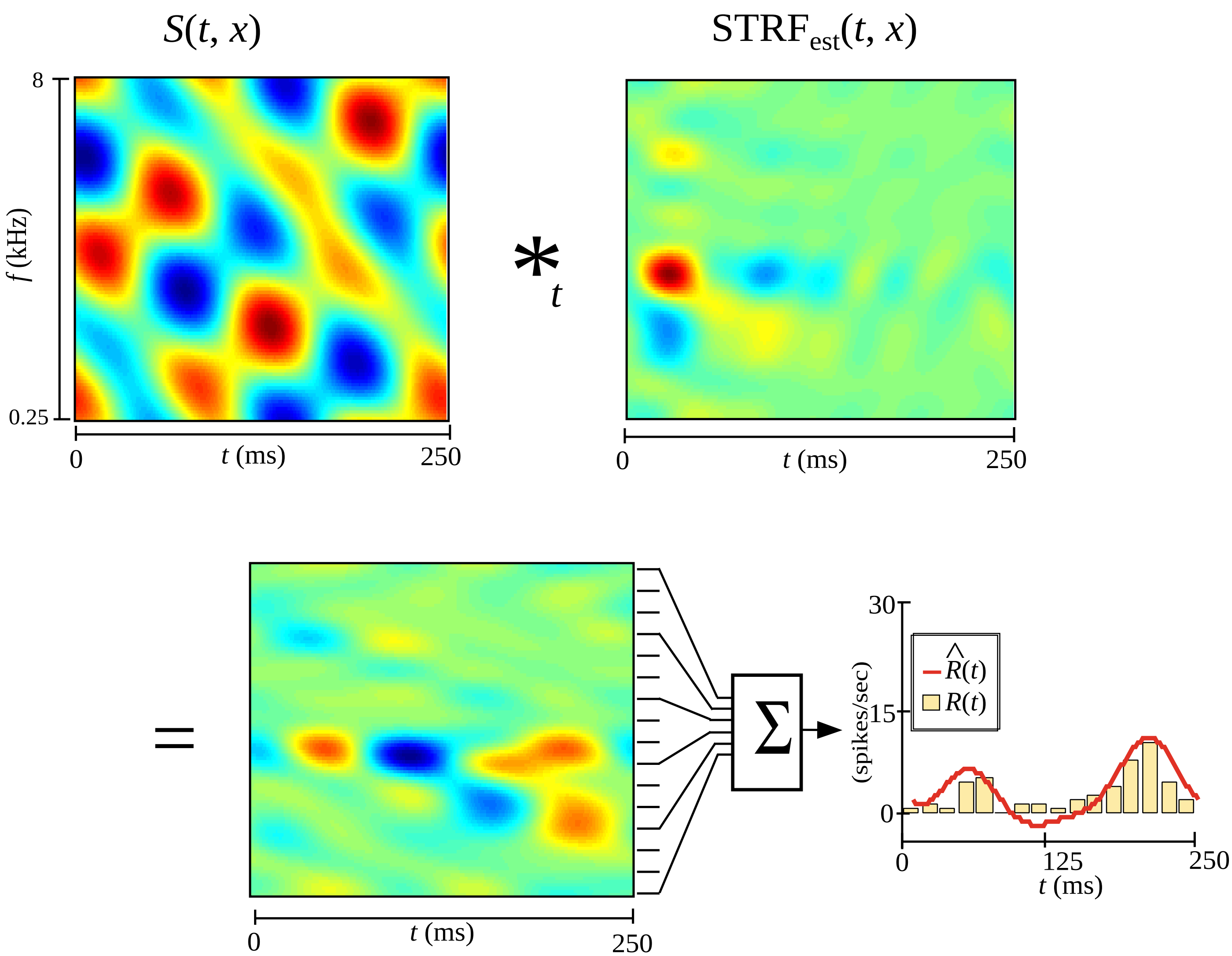
Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

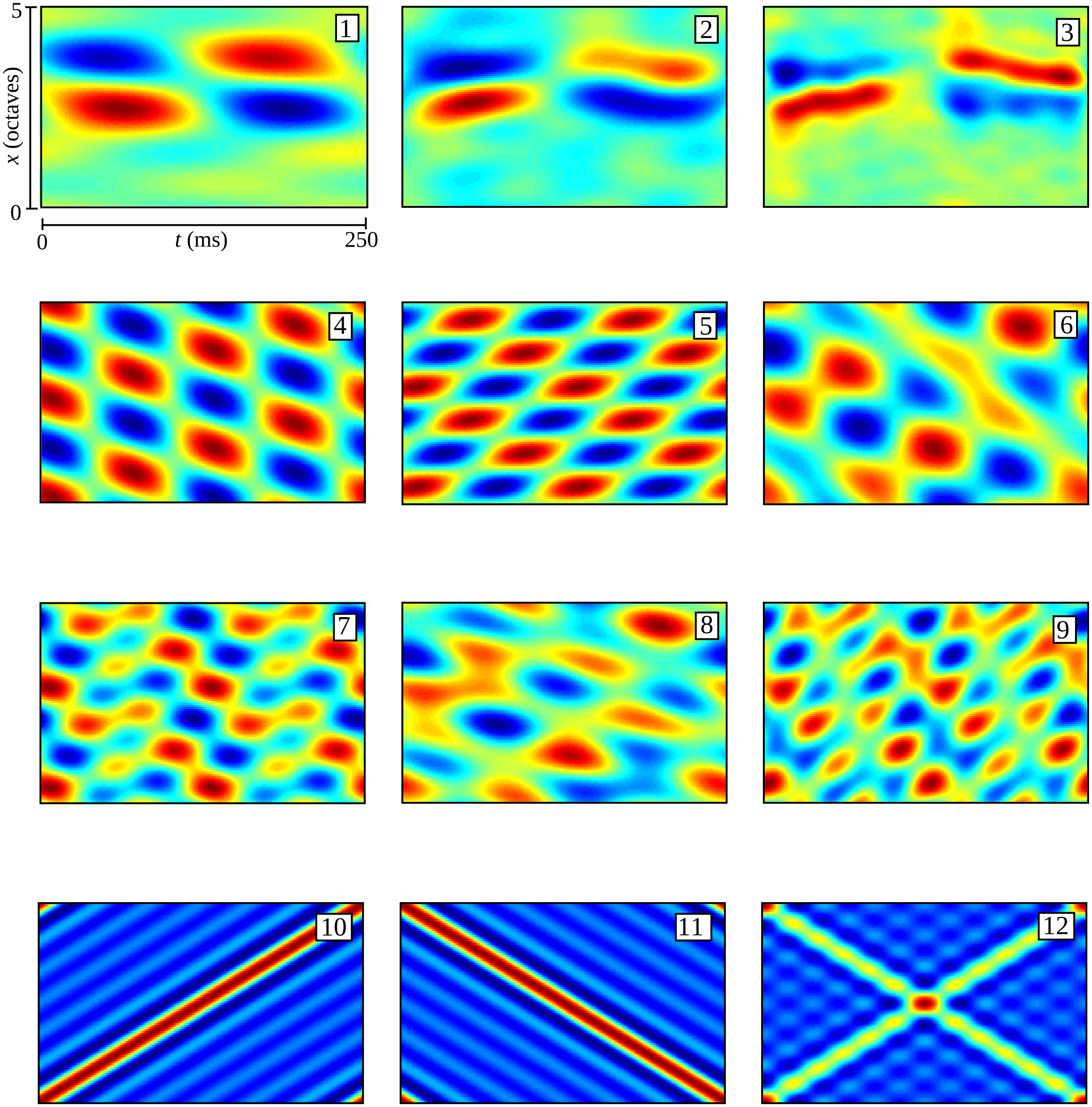
Predicting Responses from STRF

The response to an arbitrary sound is predicted by the convolution of the STRF with the stimulus' spectro-temporal envelope, plus a constant.

$$\hat{R}(t) = \frac{1}{X}\sum_{x}\{\mathrm{STRF}_{est}(t, x) *_t S(t, x)\} + \mathrm{E}\{R(t)\}$$



Stimuli Used for Predictions

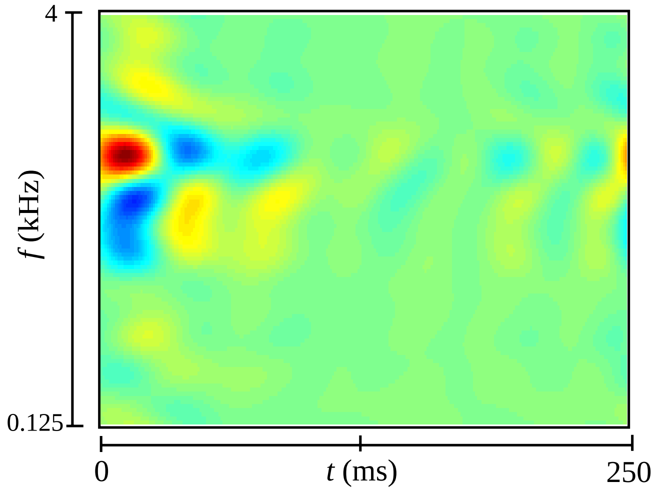


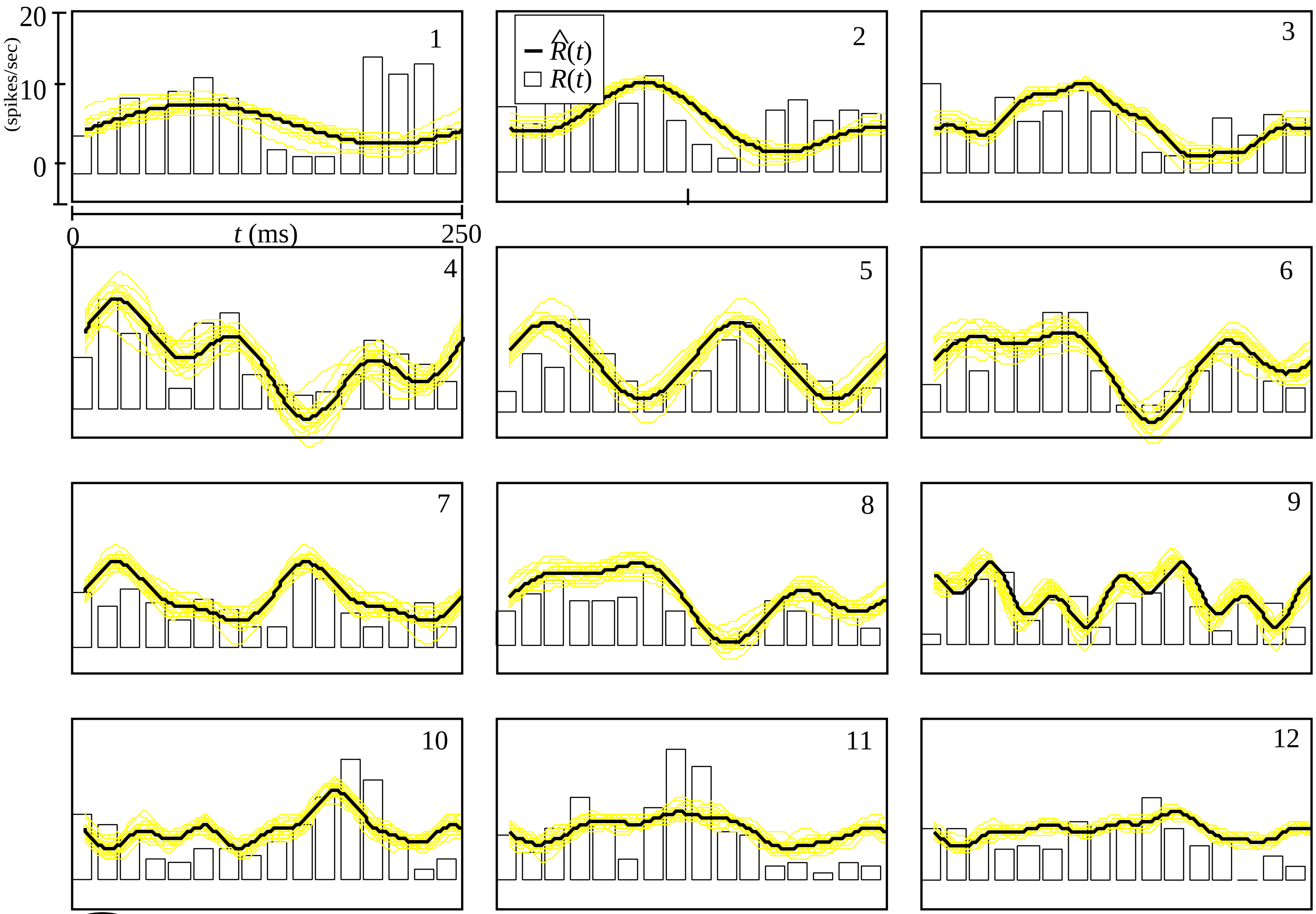
Center for Auditory
and Acoustic Research

Institute for Systems Research
University of Maryland

Validation by Prediction

The STRF estimates often predict the magnitude and dynamics of the response well.





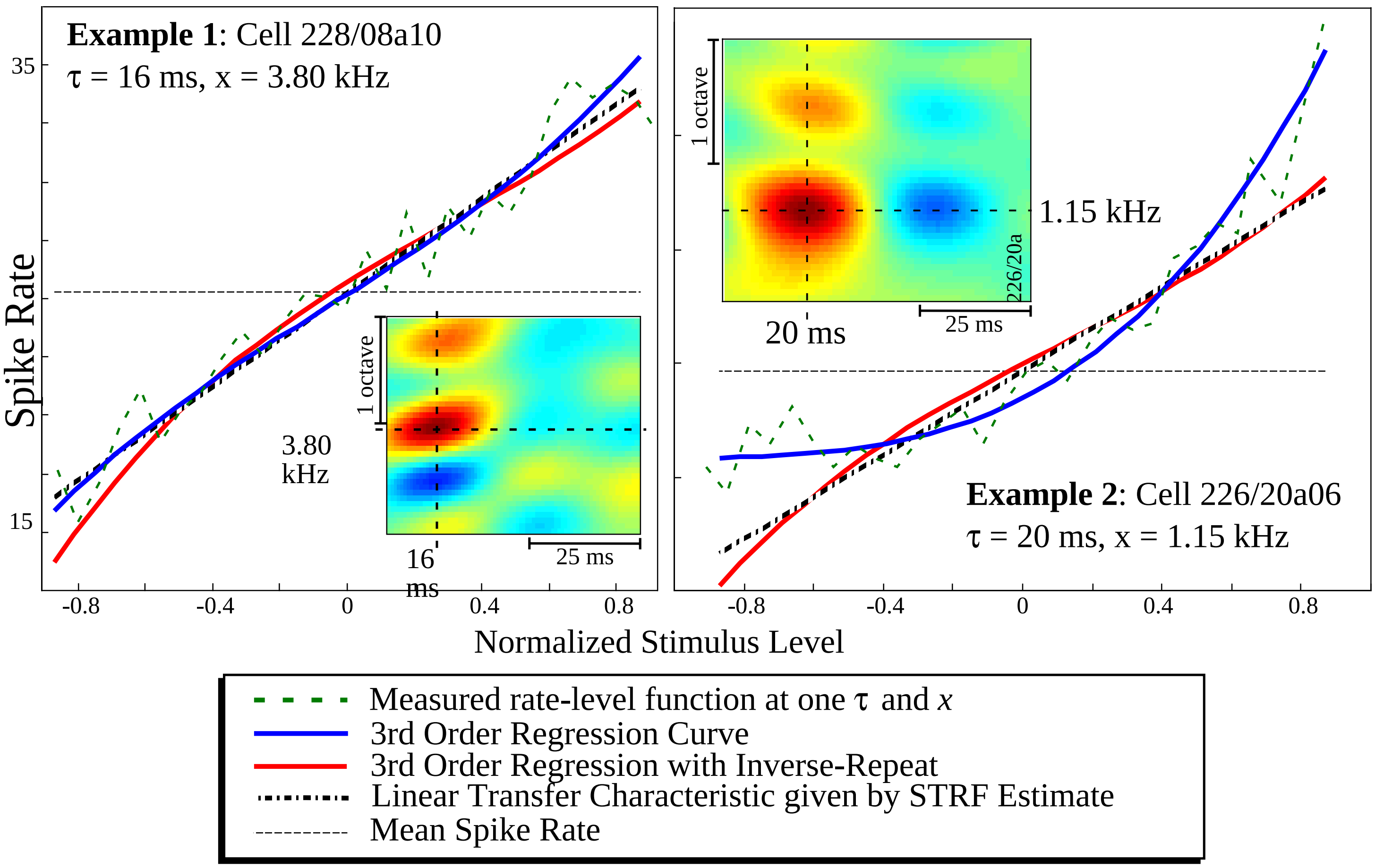
Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

Non-Linearity

- The value of the STRF at each point (τ, x) is the slope of a linear rate-level function: $R_{\tau,x}(t)$ =[STRF(τ, x)] $\cdot$ $S(t-\tau, x)$.

- Polynomial rate-level curves are measured at every (τ, x), and improve the description. These potentially non-linear functions describe the average firing rate given the stimulus level around a frequency x and at a time τ in the past.



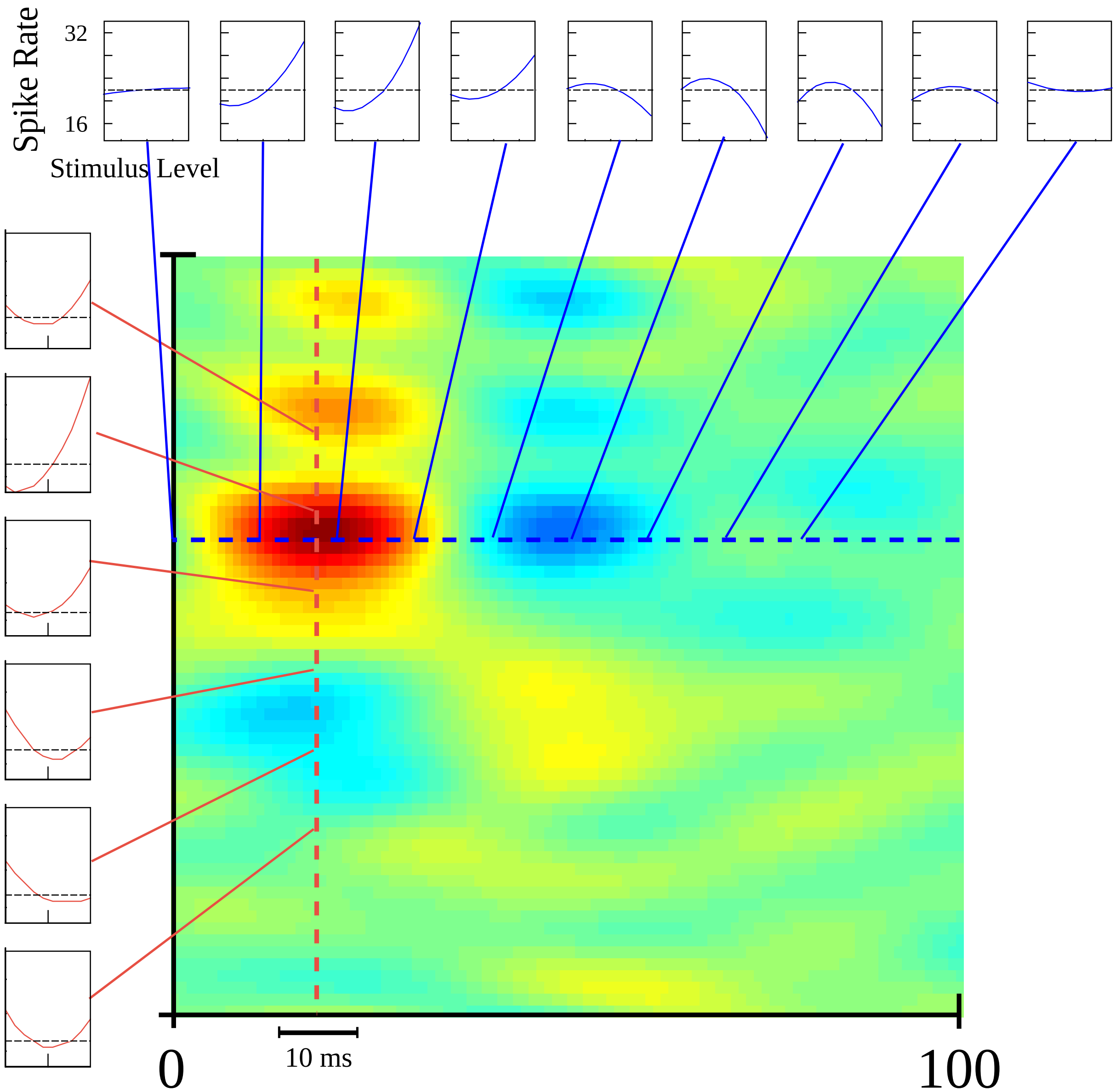
- The power series coefficients of the curves are given by the diagonals of the Volterra kernels, measured using *polynomial correlation*.

- 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 rectifying non-linearity (i.e., purely even order).

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 STRF is the linear time-invariant filter governing the transformation from stimulus to response. Thus it can be identified with the first kernel of a Volterra series expansion of the system.

Third-order Volterra series expansion of an input S(*t*, *x*):

$$R(t) = v_0 + \int d\tau \int dx \cdot v_1(\tau, x) S(t-\tau, x)$$

$$+\int d\tau_1 \int d\tau_2 \int dx_1 \int dx_2 \cdot v_2(\tau_1, \tau_2, x_1, x_2) S(t-\tau_1, x_2) S(t-\tau_2, x_2)$$

$$+\int d\tau_1 \int d\tau_2 \int d\tau_3 \int dx_1 \int dx_2 \int dx_3 \cdot v_3(\tau_1, \tau_2, \tau_3, x_1, x_2, x_3) \cdot S(t-\tau_1, x_1) S(t-\tau_2, x_2) S(t-\tau_3, x_3) \quad + \cdots$$

Form of the Regression Function:
(Third-order Approximation)

$$g(s;\tau,x) = E\left\{R(t) \mid S(t-\tau,x) = s\right\}$$

$$\approx a_0(\tau,x) + a_1(\tau,x)\cdot s + a_2(\tau,x)\cdot s^2 + a_3(\tau,x)\cdot s^3$$

- The regression functions describe the non-linearities within each channel, but not interactions between channels.

The true STRF is just the linear part...

$$STRF(\tau,x) \stackrel{\Delta}{=} a_1(\tau,x)$$

Use of the Regression Function
to Predict the Response:

$$\hat{R}(t) = \int d\tau \int dx \cdot a_0(\tau,x) + \int d\tau \int dx \cdot a_1(\tau,x) S(t-\tau,x)$$

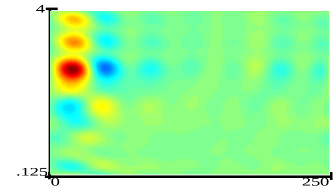
$$+\int d\tau \int dx \cdot a_2(\tau,x) S^2(t-\tau,x) + \int d\tau \int dx \cdot a_3(\tau,x) S^3(t-\tau,x)$$

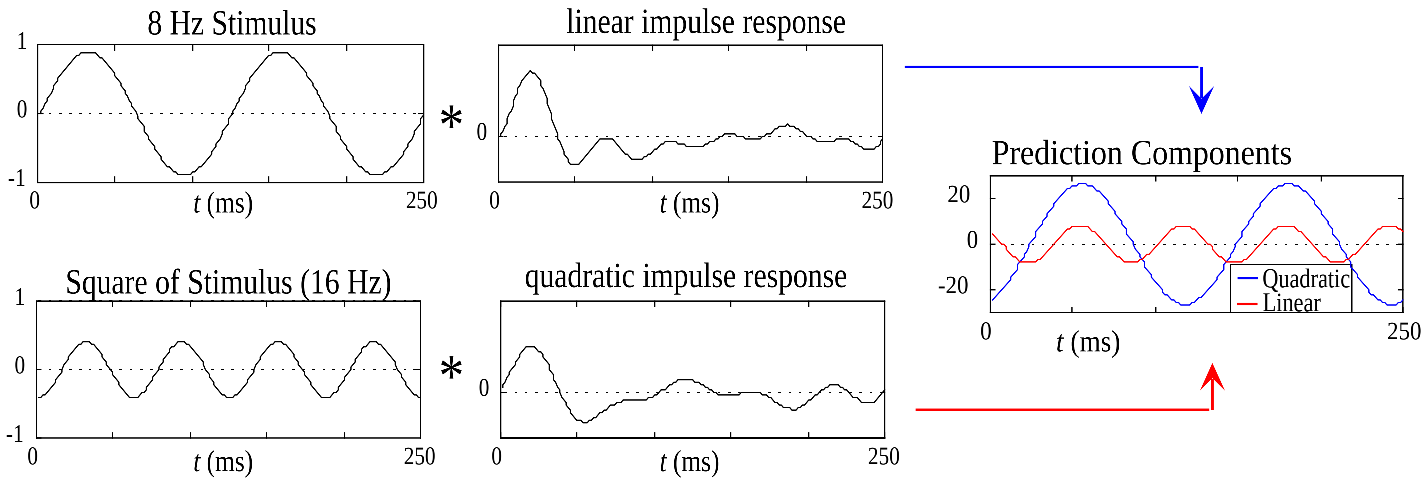
Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

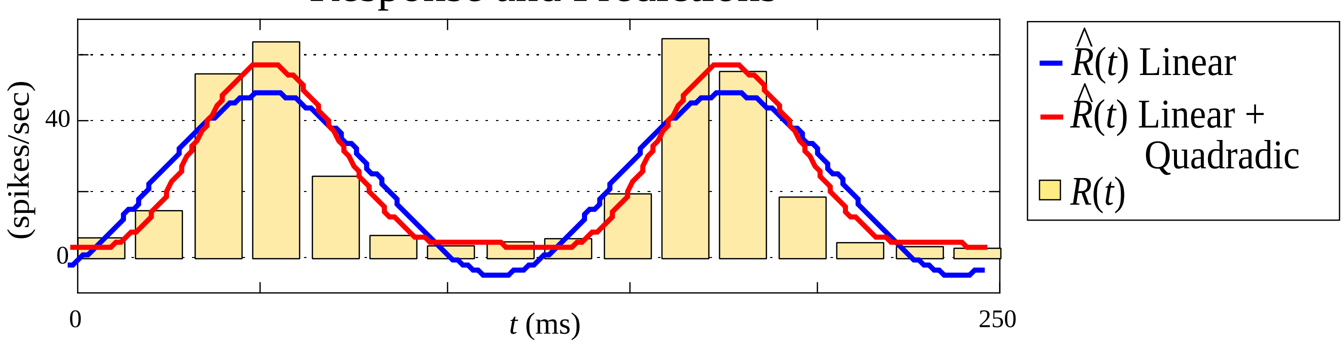
Non-Linear Prediction

- 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.

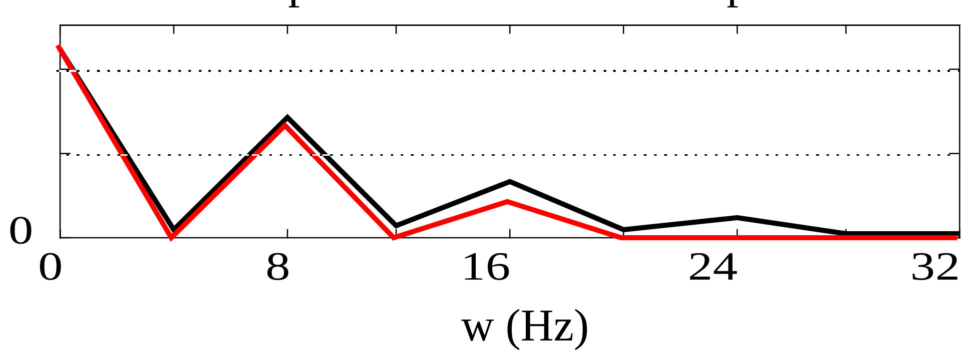




Response and Predictions



Response and Prediction Spectra



Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

Selected References

Moving Ripple / STRF / Ripple Transfer Function

- Kowalski NA, Depireux DA and Shamma SA, J. Neurophys. 76 (5) (1996) 3503-3523, and 3524-3534.
- Depireux DA, Simon JZ and Shamma SA, Comments in Theoretical Biology (1998).

Spectro-Temporal Correlation Methods

- Calhoun BM, Miller RL, Wong JC, and Young ED, 11th International Symposium on Hearing (1997).
- Eggermont JJ, Hearing Research 66 (1993) 177-201.

Singlular Value Decomposition

- Pilgram B, Schappacher W, Int. J. Bifurcation and Chaos 8 (1999) 571-580.

Polynomial Correlation

- van Dijk P, Wit HP, Segenhout JM, Hear. Res. 114 (1997) 229-242.
- de Boer E, Proc. IEEE 67 (1979) 317-318.

Related techniques and models

- Schreiner CE and Calhoun BM. Auditory Neurosci., 1 (1994) 39-61.
- Wang K and Shamma SA, IEEE Trans. on Speech and Audio 2(3) (1994) 421-435, and 3(2) (1995) 382-395.
- Shamma SA, Fleshman JW, Wiser PR and Versnel H, J. Neurophys 69(2) (1993) 367-383.
- Shamma SA, Versnel H and Kowalski NA, J. Auditory Neuroscience (1) (1995) 233-254, and 255-270, and 271-285.

Summary

- The Spectro-Temporal Response Field (**STRF**) embodies the linear component of a neuron's response to a sound's spectro-temporal envelope.
- For neurons in Primary Auditory Cortex (AI) of Ferrets, we have previously characterized the STRF and its Fourier transform: the ripple transfer function. Using single moving ripples, we found best responses to temporal modulations typically from 4 to 16 Hz, and spectral-peak densities from 0.4 to 1.6 cycles/octave.
- By reverse-correlating with temporally orthogonal ripple combinations (TORCs) enriched with these spectral and temporal qualities, we can use brief stimuli and compute clean STRF estimates "on the fly".
- Applying Singular Value Decomposition (SVD) to the measured ripple transfer functions, we found that a vast majority of STRFs in AI are quadrant separable. Moreover, we were able to isolate and discard much of the remaining noise in the STRF estimate.
- The measured STRFs are generally a good predictors of the response to broad-band sounds with arbitrary spectro-temporal structure.
- Using the same experimental paradigm, rate-level non-linearities were measured at numerous frequencies and latencies. These were mild and even-ordered, corresponding to soft-rectification and expansion, and they slightly improved response predictions.

Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland