Representation Of Dynamic Broadband Spectra In Auditory Cortex

S. A. Shamma[1,2]
D. A. Depireux[1]
D. J. Klein[1,2]
J. Z. Simon[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 and by a grant from the National Institute on Deafness and Other Communication Disorders

Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

Summary

- In Primary Auditory Cortex (AI) of Ferrets, we have previously characterized cells' responses to dynamic broad-band sounds. We found best responses to temporal modulations from 4 to 16 Hz, and spectral modulations from 0.4 to 1.6 cycles/octave in the stimulus's spectro-temporal envelope.

- The Spectro-Temporal Response Field (**STRF**) explains the linear component of the response to the spectro-temporal envelope of a broad-band sound.

- The STRF is often a good predictor of the response to an arbitrary sound. However, previous measurements of the STRF using sinusoidal spectro-temporal envelopes were hampered by the time required to accumulate data from a cell.

- We use sums of spectro-temporal sinusoids as stimuli: these
  - reduce recording time
  - confirm quadrant spectro-temporal separability
  - can be used to explore non-linearities

Ripples, or Auditory Gratings

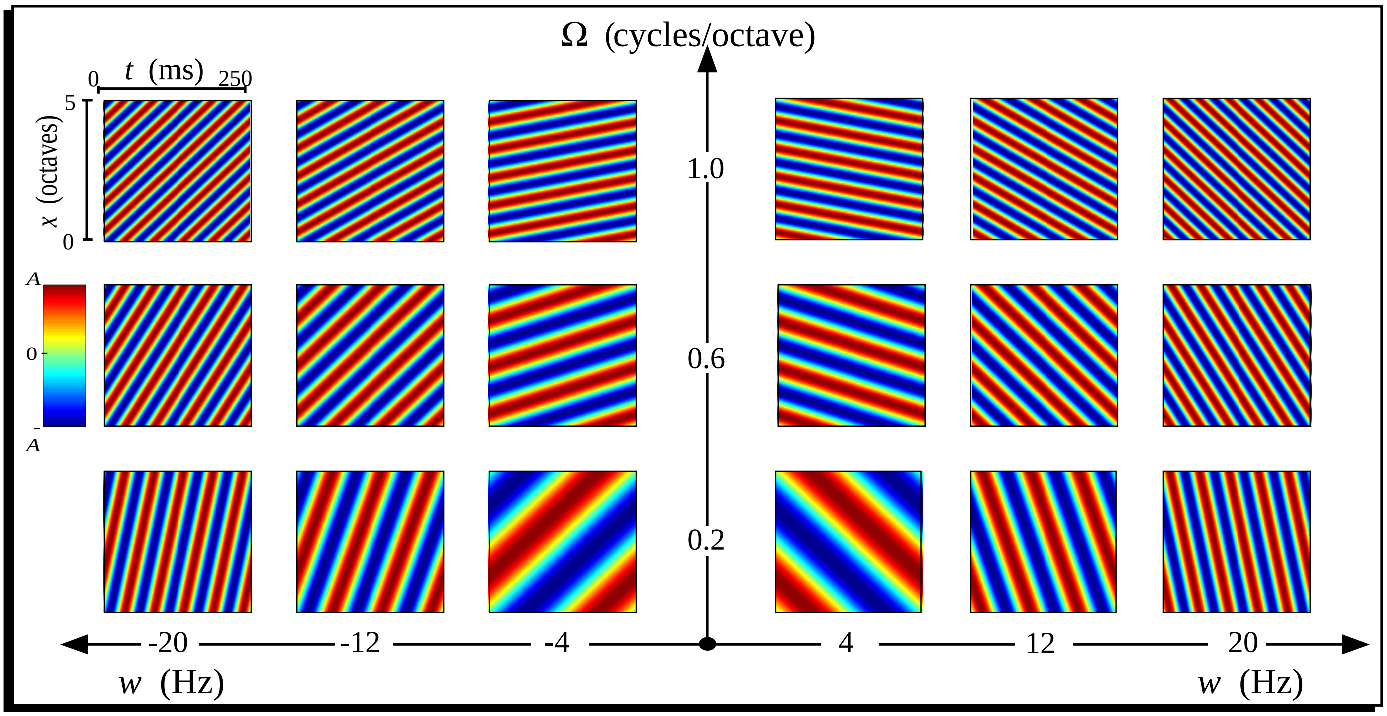
Ripples are auditory gratings whose spectral envelope is a sinusoid along the log(frequency) axis. At any time ***t*** and any frequency ***x***, the amplitude ***S*(*t*,*x*)** is given by:

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

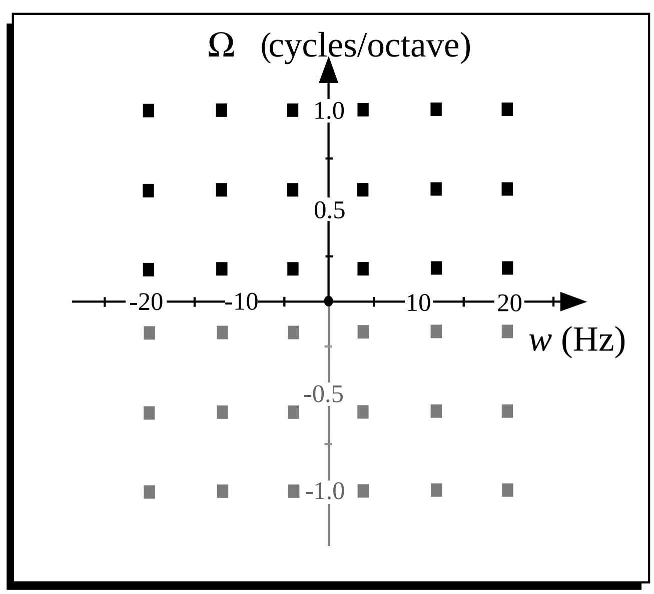
$x = \log_2[f / f_0]$

w = ripple velocity

Ω = ripple density



The **Ripple Domain** is the Fourier space of the spectrograms. We probe a cell at different velocities w and different densities Ω, and quantify the response for up and down-moving sounds.
Any ripple in the lower half-plane is equivalent to a ripple in the upper-half plane.



Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

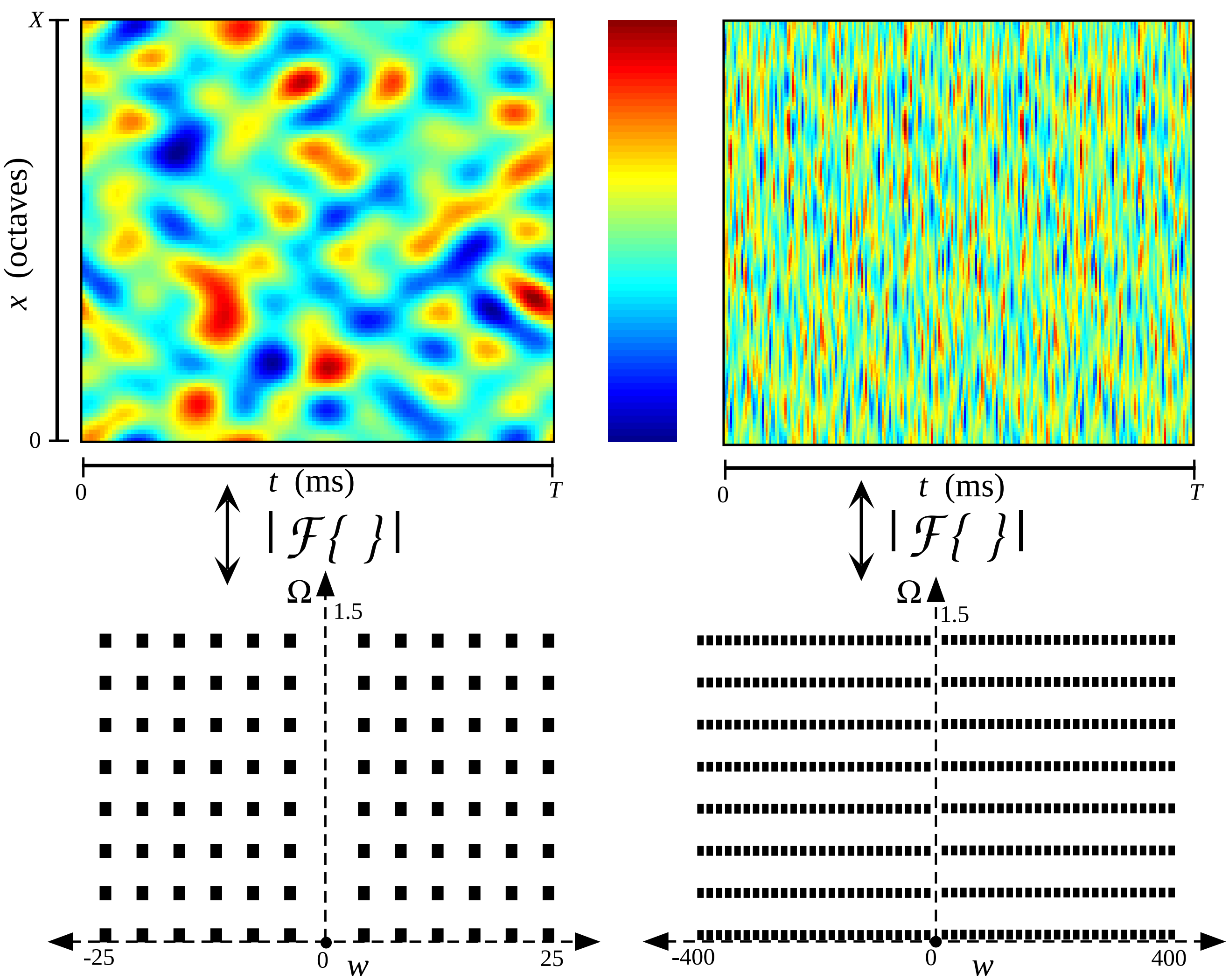
Spectro-Temporal Noise

In order to speed up the characterization of a cell's response, we have used multiple combinations of ripples, of all velocities w and densities Ω, with random phases. Different combinations have different choices of individual ripple phases. The range of frequencies and/or velocities is adjusted to match the range of interest to the cells being studied. We generally use –24 Hz to 24 Hz for cortex, and –400 Hz to 400 Hz for the Inferior Colliculus.
This is a Spectro-Temporal generalization of white noise.

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

Examples:

Cortical Stimulus **Mid-Brain Stimulus**



Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

STRF from Cross-Correlation with Noise

Stimulus: $S(t, x)$

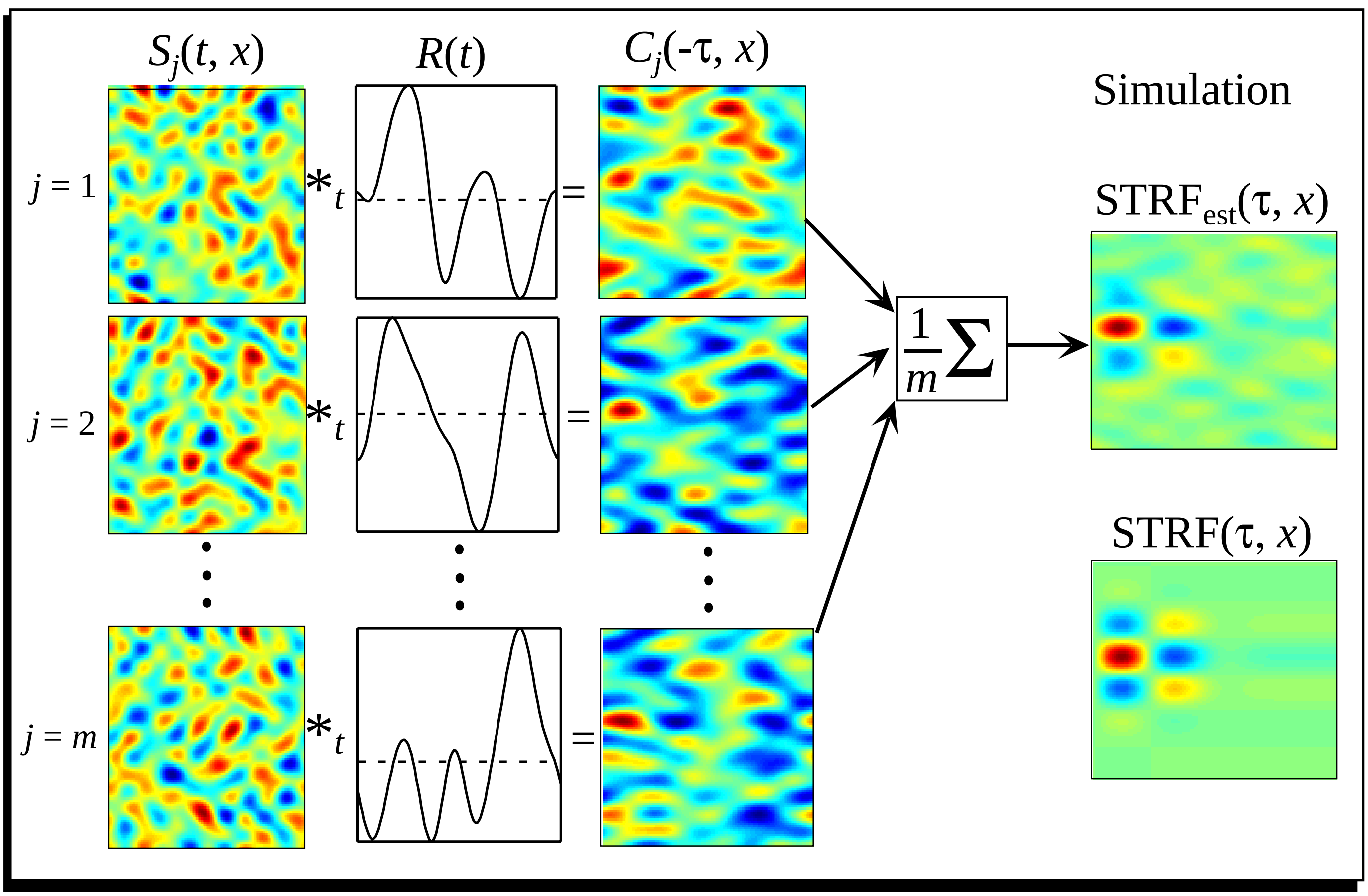
Response: $R(t) = \sum_k \delta(t - t_k)$

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

(= Spike-Triggered Average)

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

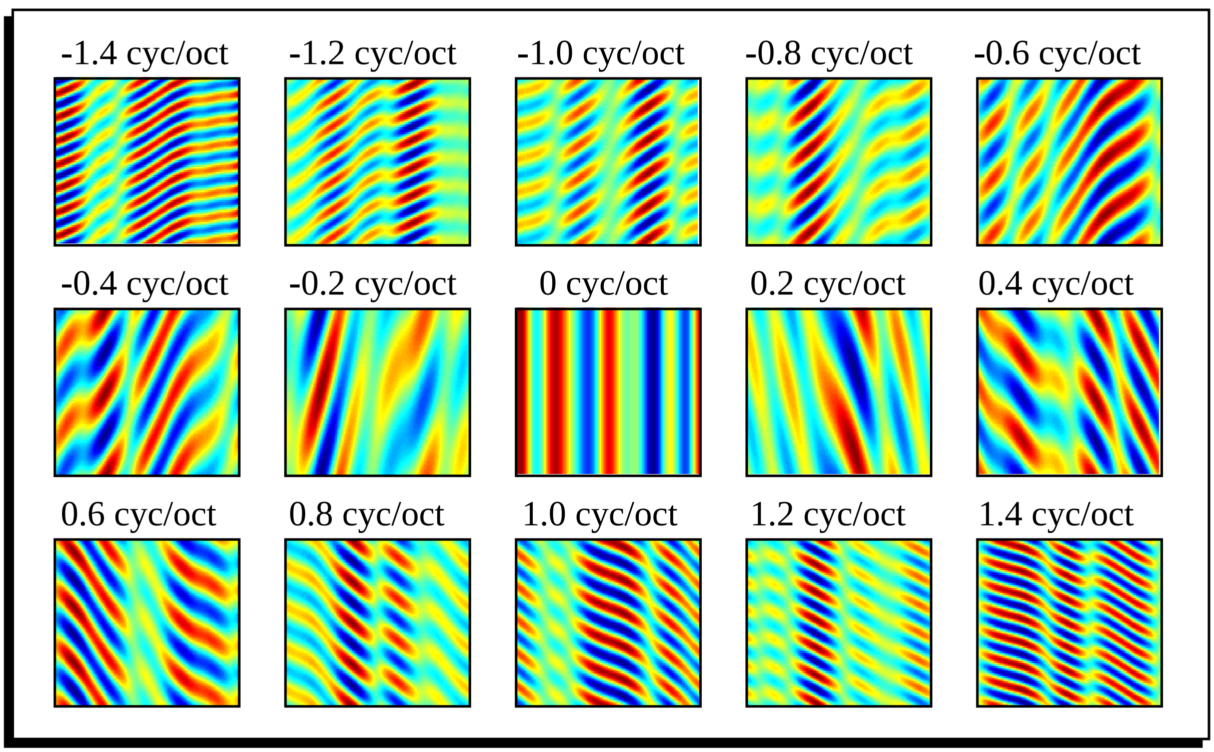
$$\mathrm{STRF}_{\mathrm{est}}^{\mathrm{noise}}(\tau, x) = \frac{1}{m}\sum_{j=1}^{m} C_j(-\tau, x)$$

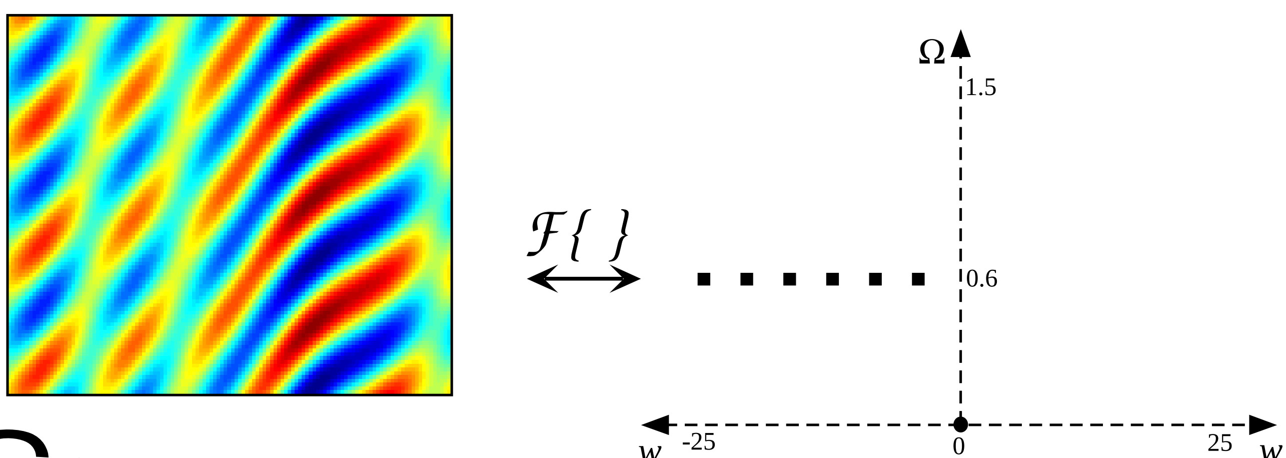


Temporally Orthogonal Ripple Combinations (TORCs)

- Stimuli are composed only of ripples with different ripple velocities.
- Each stimulus contains ripples which cover the same range of ripple velocities, but at different ripple frequencies.
- Multiple stimuli are still needed to present a complete set of ripples.

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





Center for Auditory and Acoustic Research

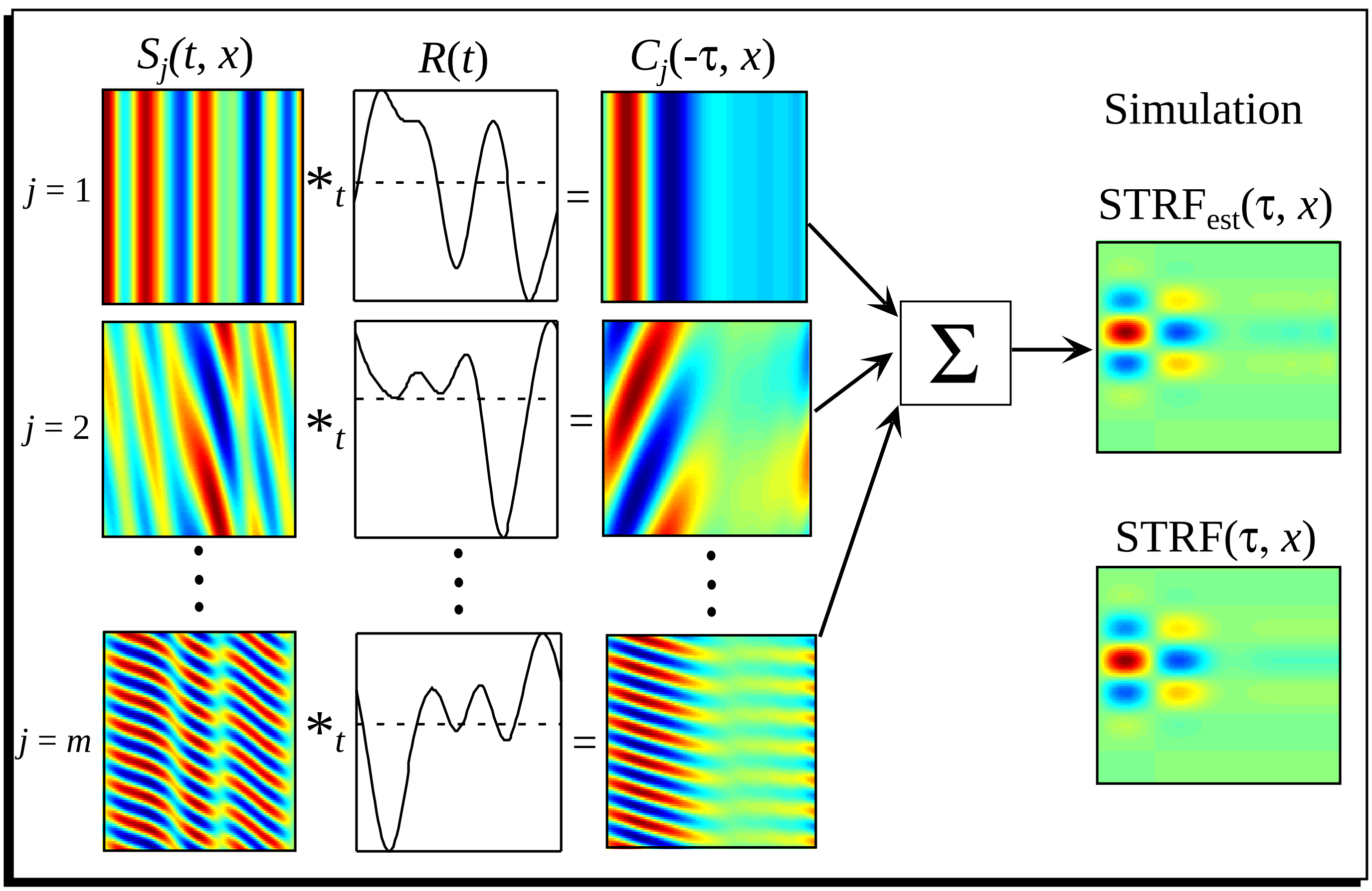
Institute for Systems Research
University of Maryland

STRF from Cross-Correlation with TORCs

- TORCs are better suited for temporal cross-correlation because there are no cross terms.

- The resulting estimates are robust, use short-duration stimuli, and are quickly computed.

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

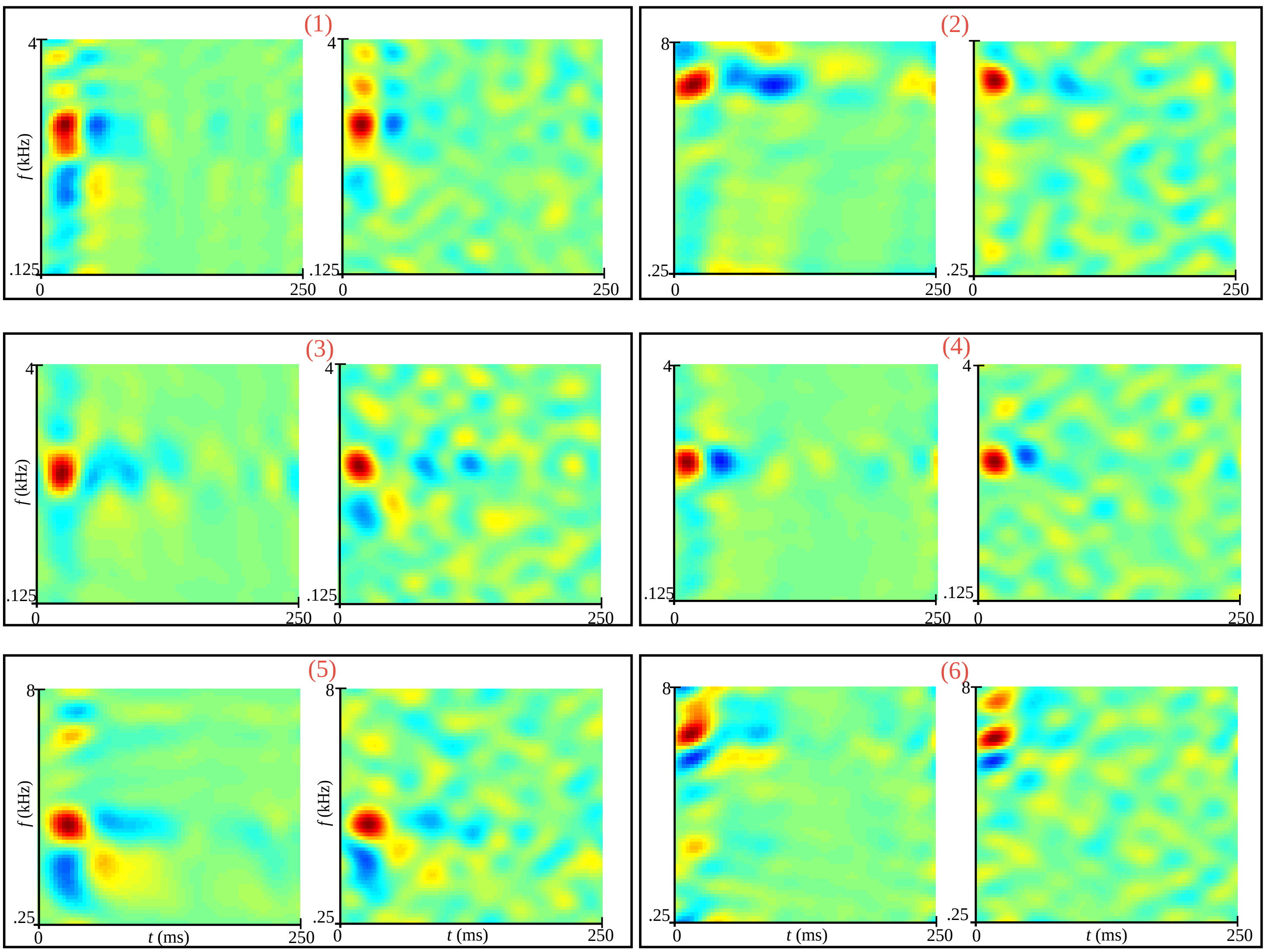


Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

TORC Cross-Correlation vs. Single-Ripple

Single Ripple | Cross-Correlation | Single Ripple | Cross-Correlation



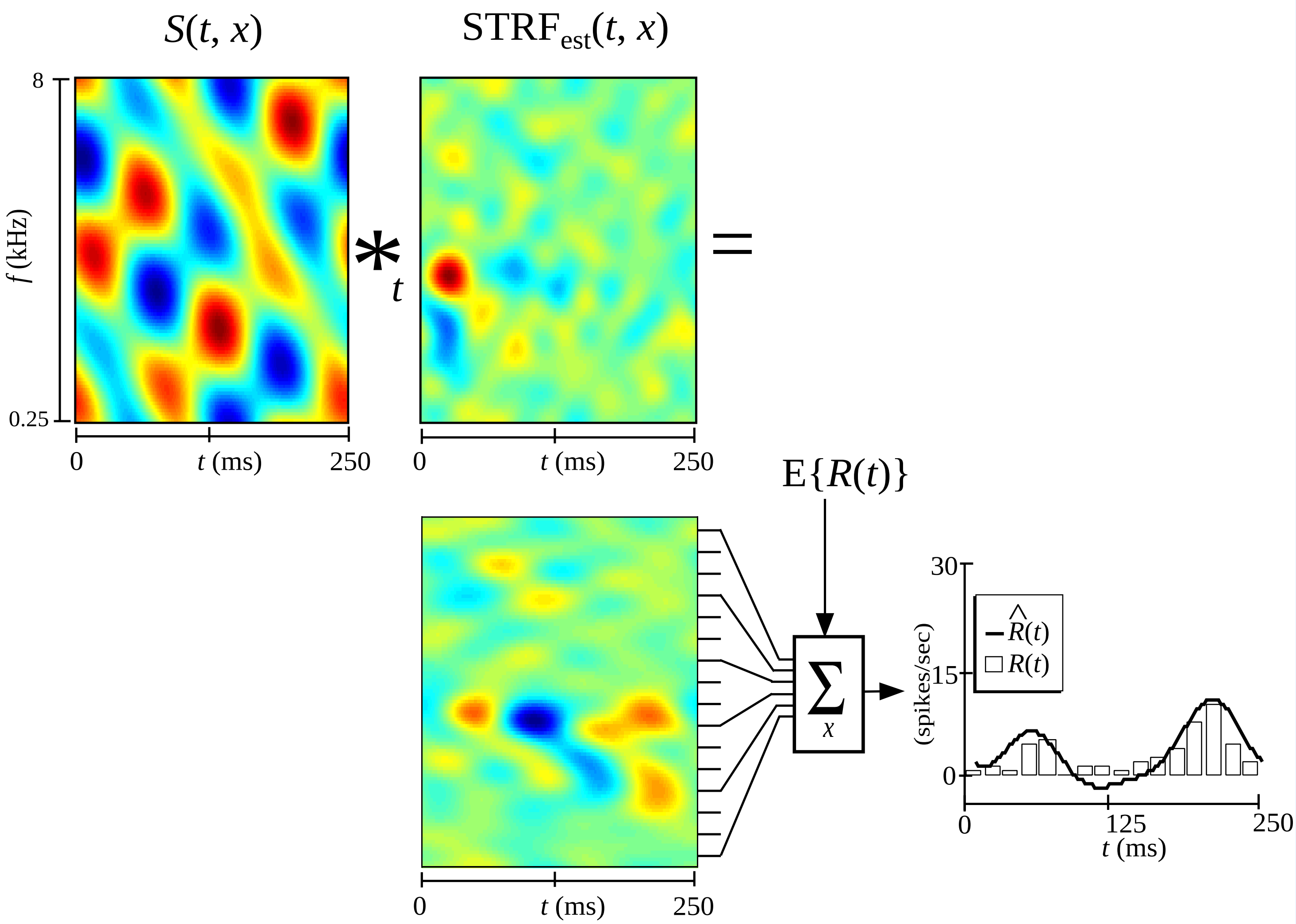
Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

Predicting Responses from STRF

The response to an arbitrary sound is given by the convolution of the STRF with the stimulus envelope, plus a constant.

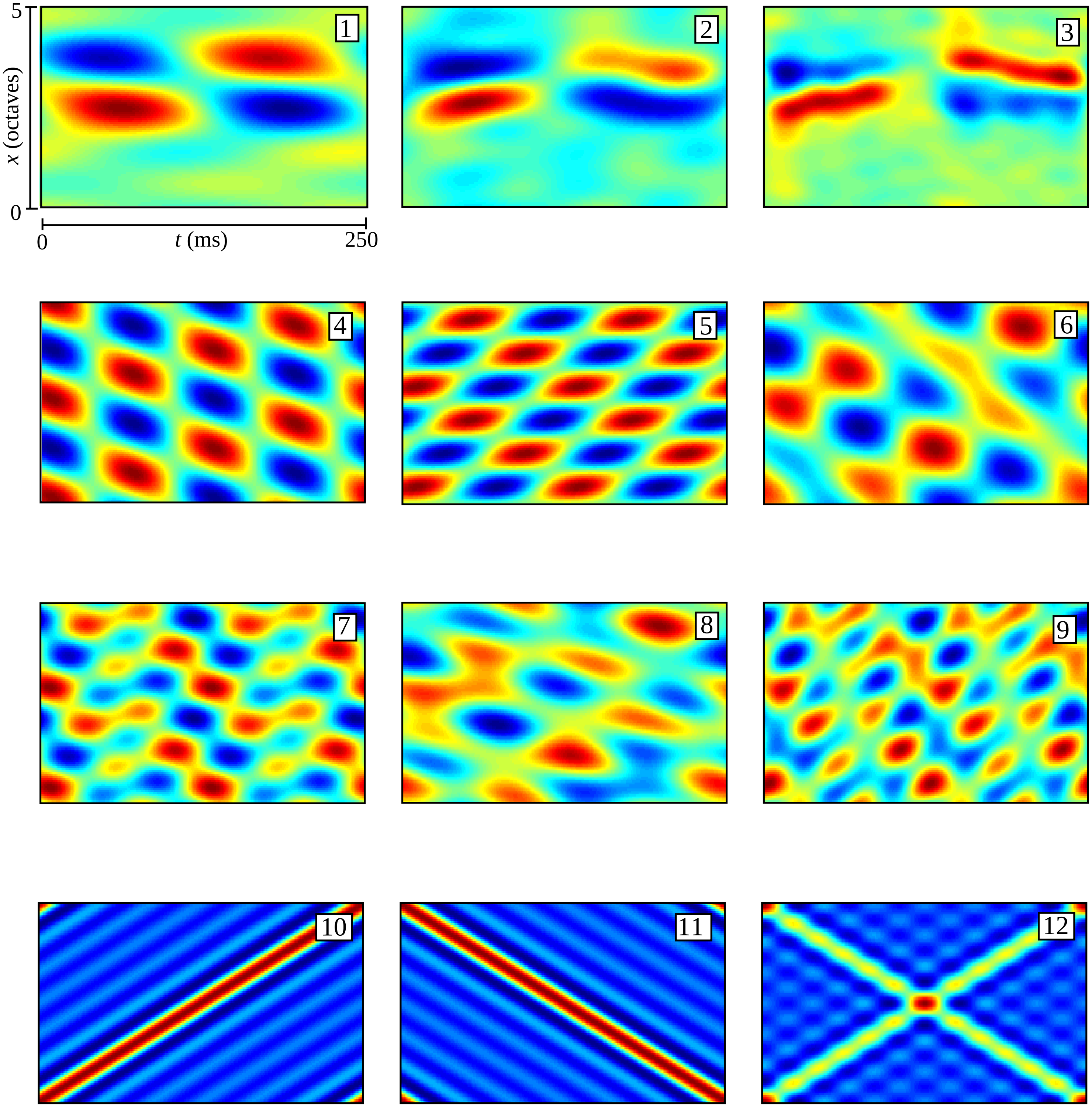
$$\boxed{\hat{\boldsymbol{R}}(\boldsymbol{t}) = \frac{\mathbf{1}}{\boldsymbol{X}} \sum_{\boldsymbol{x}} \{\mathbf{STRF_{est}}(\boldsymbol{t}, \boldsymbol{x}) *_t \boldsymbol{S}(\boldsymbol{t}, \boldsymbol{x})\} + \mathbf{E}\{\boldsymbol{R}(\boldsymbol{t})\}}$$



Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

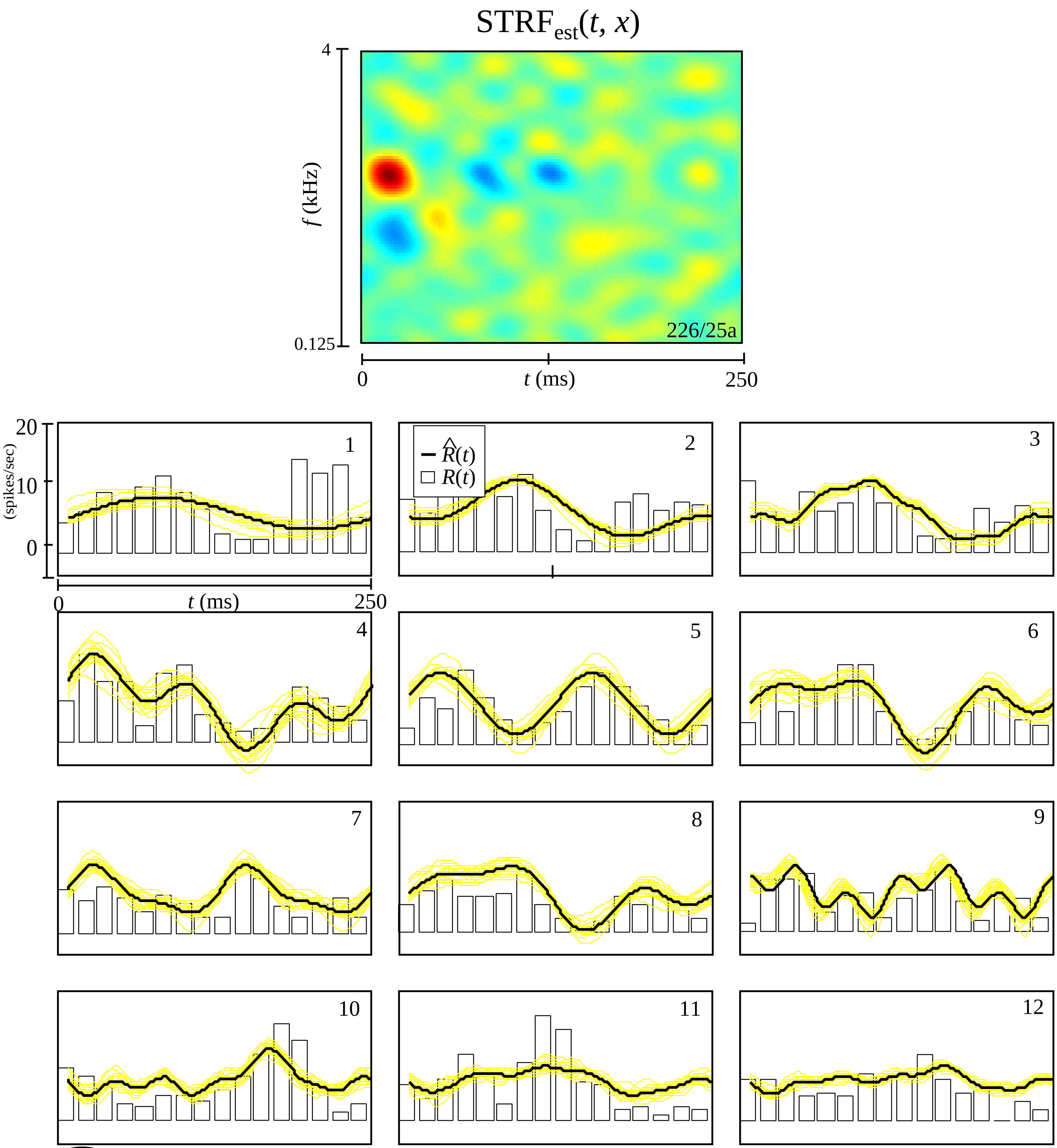
Stimuli Used for Predictions



Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

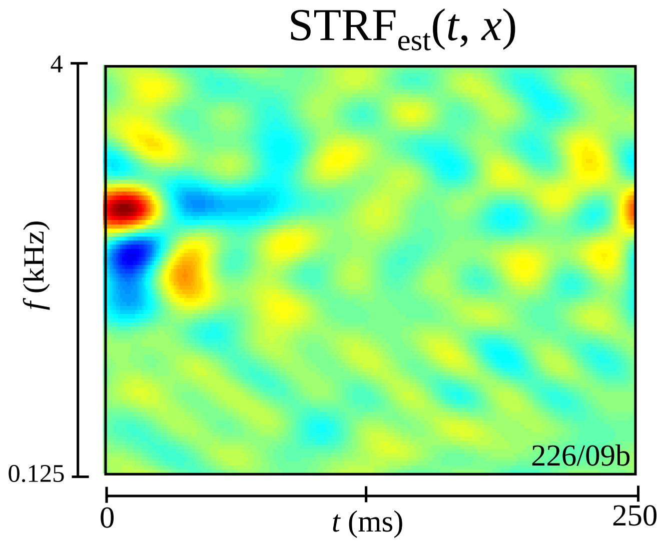
Predictions and Responses I

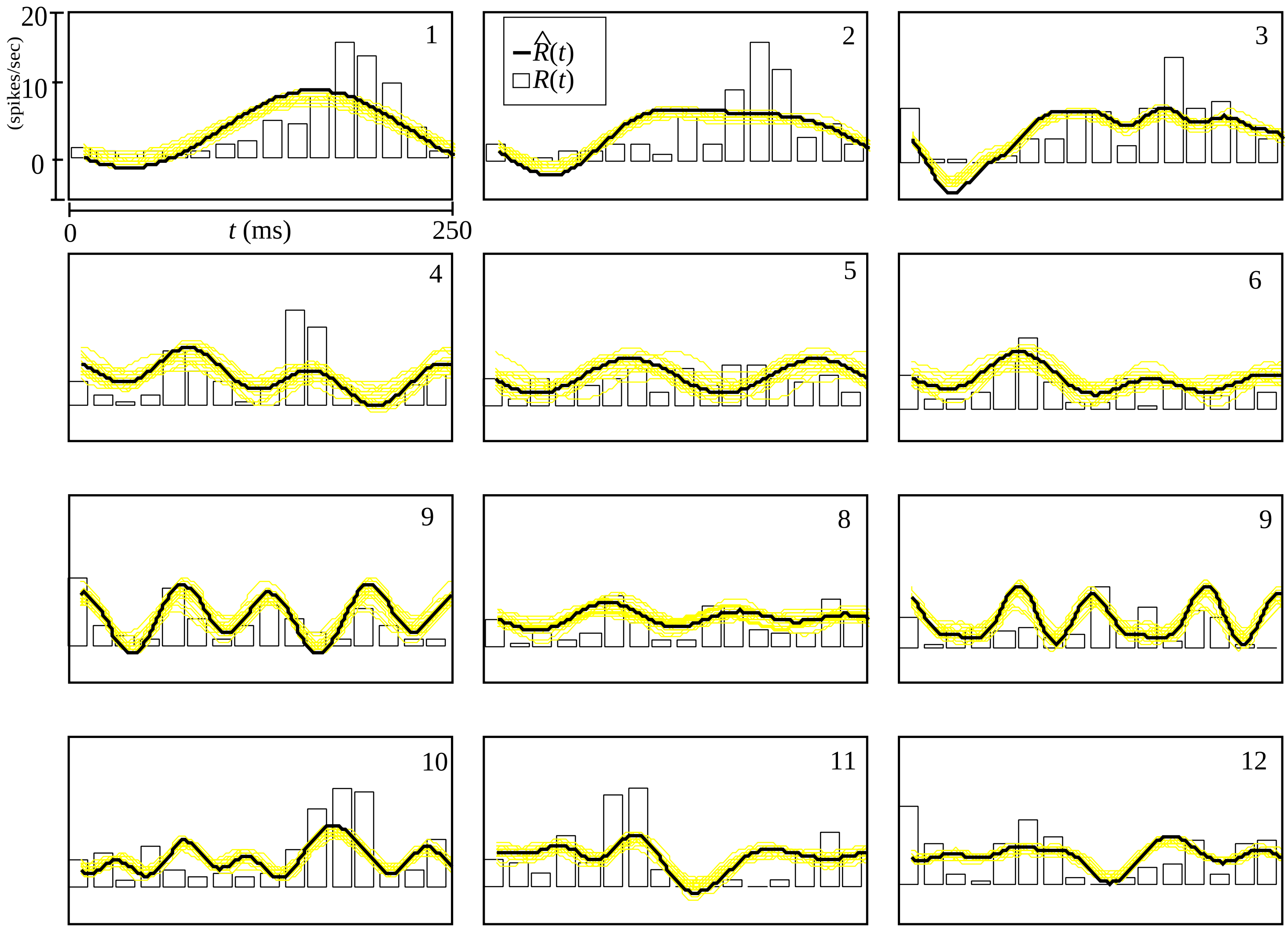


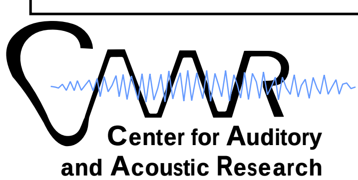
Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

Predictions and Responses II



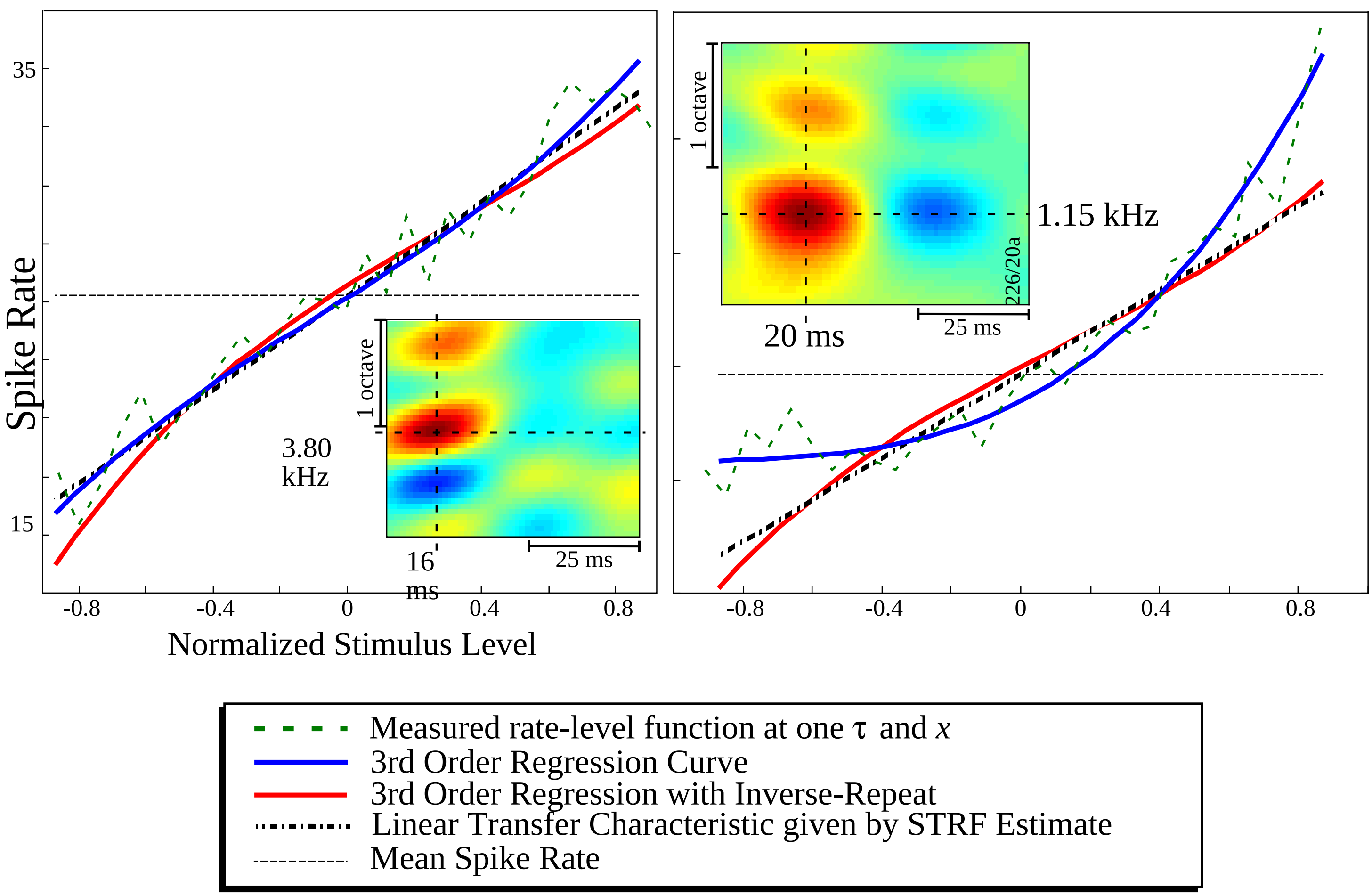




Institute for Systems Research
University of Maryland

Non-Linearity

- The value of the STRF at each point is the slope of a linear rate-level function: $R_{\tau,x}(t) = \mathrm{STRF}(\tau, x) \cdot S(t\text{-}\tau, x)$. The instantaneous rate is given by a **linear** operation between the stimulus envelope and the STRF.
- Polynomial rate-level curves can be measured at every (τ, x), and improve the description.
- The coefficients of a power series expansion of the curves are given by the diagonals of the Volterra kernels.
- Using cubic polynomials, and inverse-repeat stimuli, we have shown that either the non-linearities are absent, or they are restricted to second order.
- Subtraction of the response to the inverted envelope from the response to the non-inverted envelope gives a polynomial fit dominated by the linear term. This would be expected, for example, from a rectifying non-linearity.

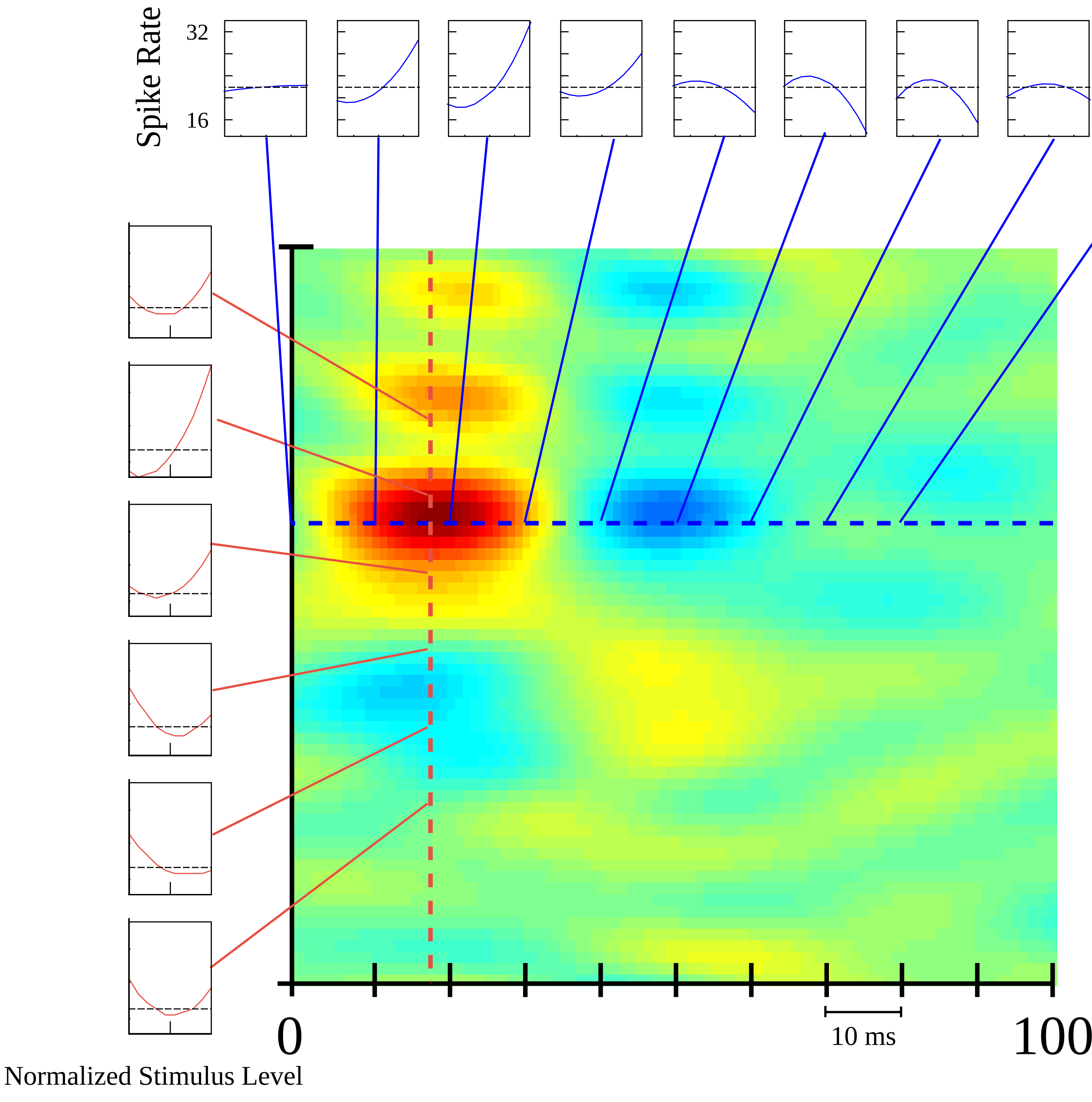


Center for Auditory and Acoustic Research

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 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...

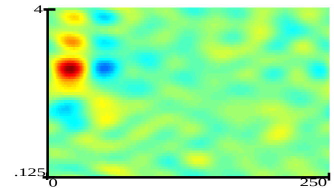
$$STRF(\tau, x) \stackrel{\Delta}{=} a_1(\tau, x)$$

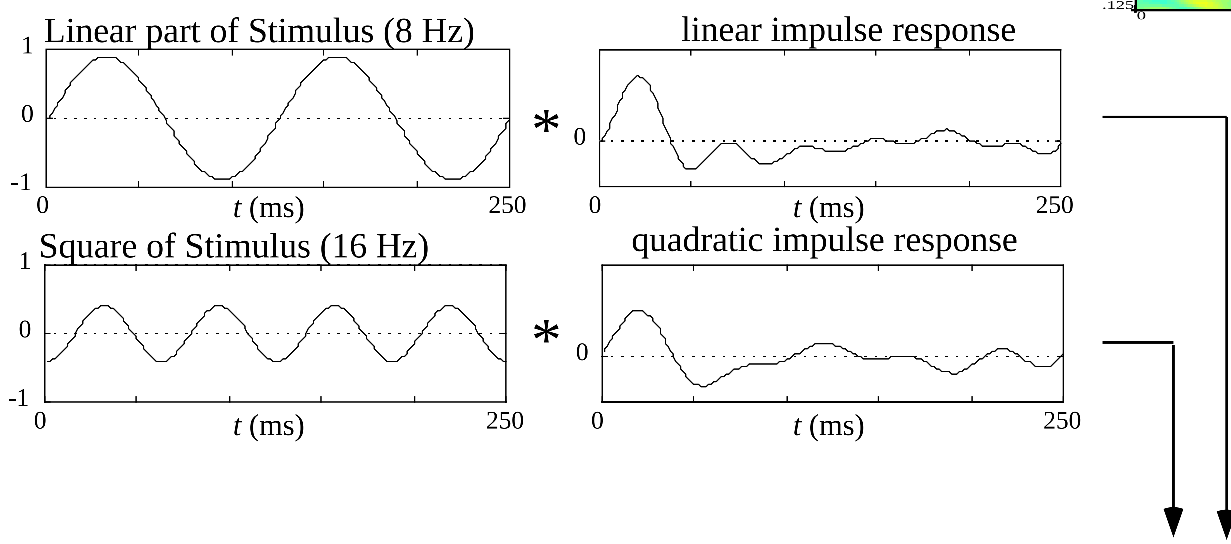
Use of the Regression Function
to Predict the Response:

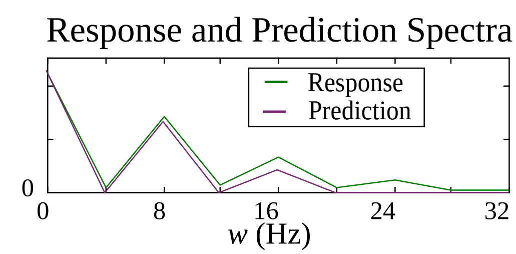
$$\hat{R}(t) = \int d\tau \int dx \cdot a_0(\tau, x) + \int d\tau \int dx \cdot a_1(\tau, x) S(t-\tau, x)$$
$$+ \int d\tau \int dx \cdot a_2(\tau, x) S^2(t-\tau, x) + \int d\tau \int dx \cdot a_3(\tau, x) S^3(t-\tau, x)$$

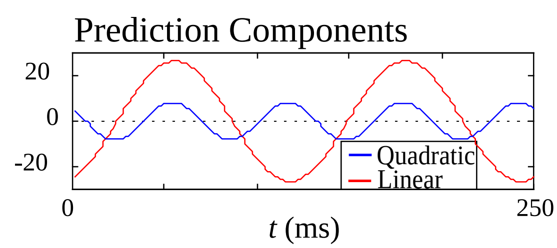
Non-Linear Prediction

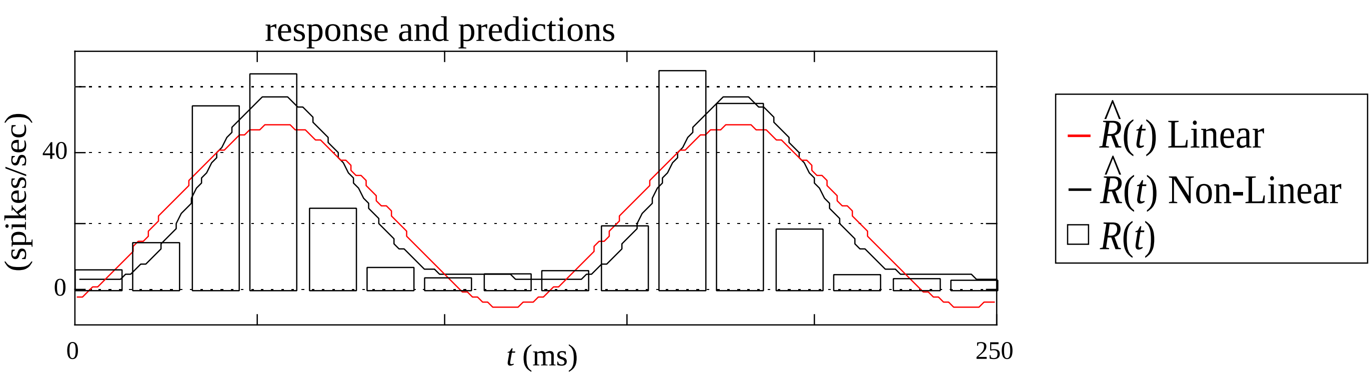
- Preliminary results indicate that the non-linear predictions often fit the responses more accurately than the linear predictions, although the differences between the two may be subtle.











Center for Auditory and Acoustic Research

Institute for Systems Research
University of Maryland

Selected References

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.

Sum-of-Sinusoids Stimuli

- Victor JD, Shapley R, Biophys. J. 29 (1980) 459-484.

Polynomial Correlation

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

Dynamical Transfer Function papers

- 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).

Stationary Transfer Function papers

- Shamma SA, Versnel H and Kowalski NA, J. Auditory Neuroscience (1) (1995) 233-254, and 255-270, and 271-285.
- Schreiner CE and Calhoun BM. Auditory Neurosci., 1 (1994) 39-61.

Related analysis techniques and models

- 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.