

Representation Of Dynamic Broadband Spectra In Auditory Cortex

S. A. Shamma^{1,2}
D. A. Depireux¹
D. J. Klein^{1,2}
J. Z. Simon¹

¹Institute for Systems Research
²Department 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

also available at <<http://www.isr.umd.edu/CAAR/pubs.html>>

Report Documentation Page

*Form Approved
OMB No. 0704-0188*

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE 1998	2. REPORT TYPE	3. DATES COVERED 00-00-1998 to 00-00-1998			
4. TITLE AND SUBTITLE Representation of Dynamic Broadband Spectra in Auditory Cortex		5a. CONTRACT NUMBER			
		5b. GRANT NUMBER			
		5c. PROGRAM ELEMENT NUMBER			
6. AUTHOR(S)		5d. PROJECT NUMBER			
		5e. TASK NUMBER			
		5f. WORK UNIT NUMBER			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Maryland, Department of Electrical Engineering and Computer Engineering, Institute for Systems Research, College Park, MD, 20742		8. PERFORMING ORGANIZATION REPORT NUMBER			
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)			
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)			
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Same as Report (SAR)	18. NUMBER OF PAGES 17	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

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

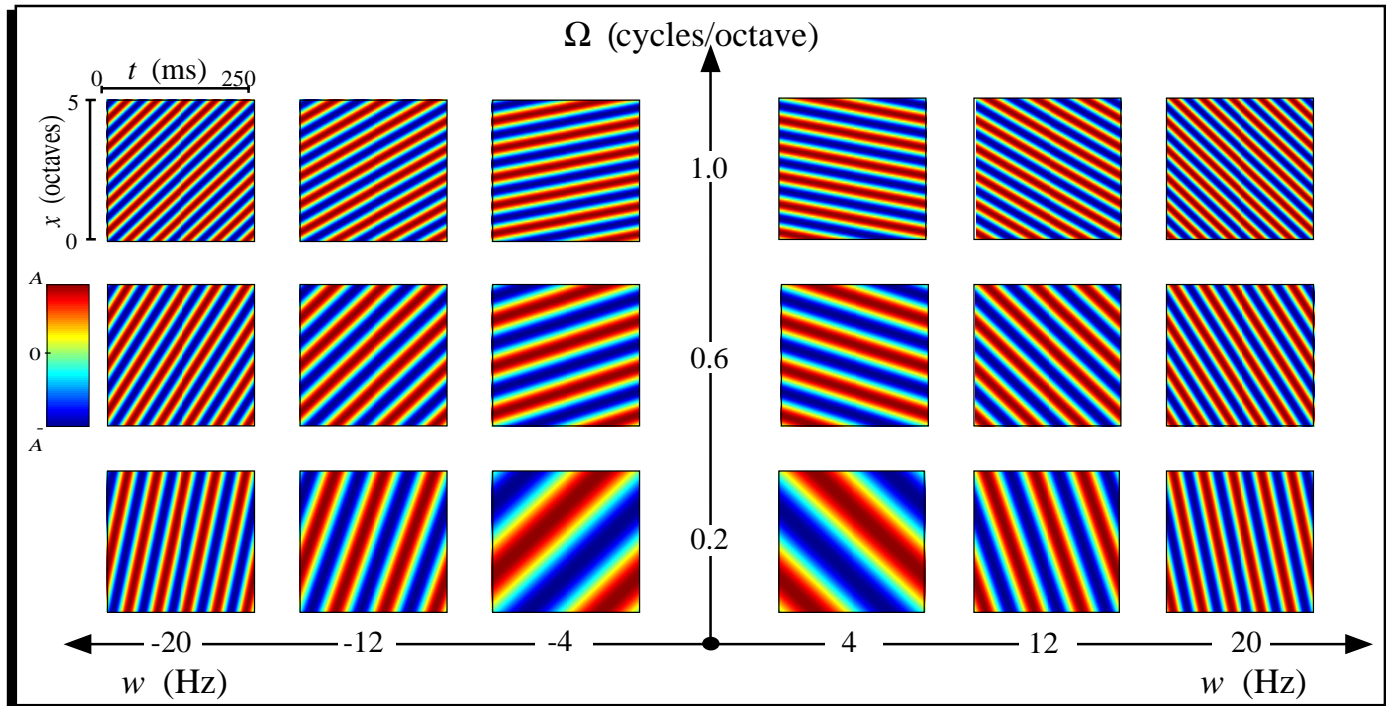
Ripples, or Auditory Gratings

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

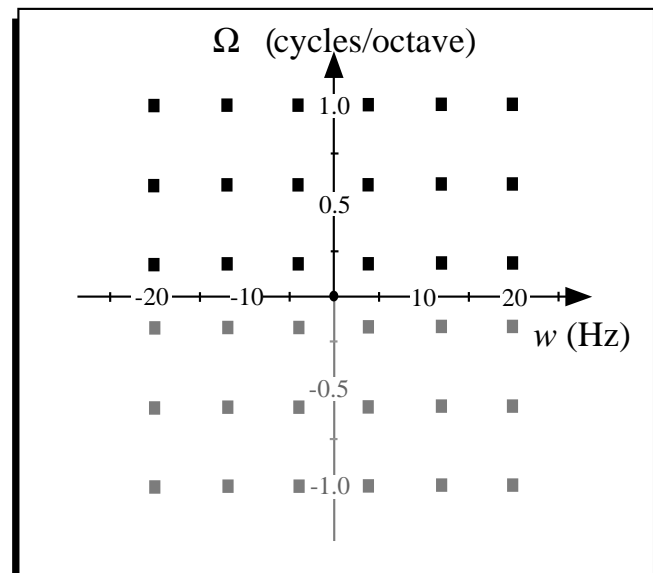
$$S(t,x) = \sin[2\pi w t + 2\pi \Omega x + \phi]$$

$$x = \log_2[f/f_0]$$

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.



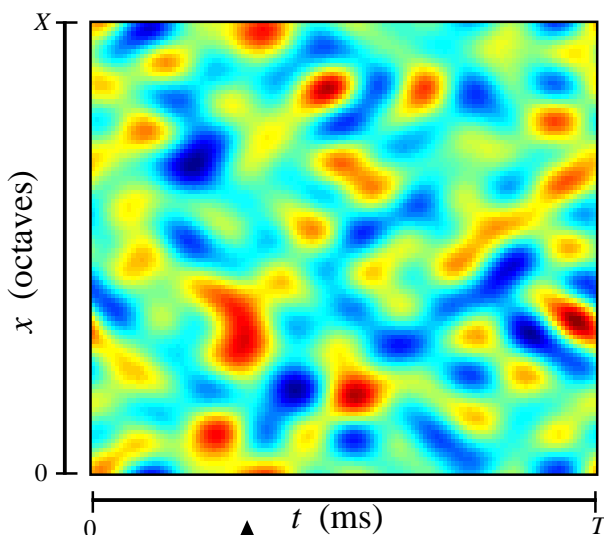
Spectro-Temporal Noise

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.

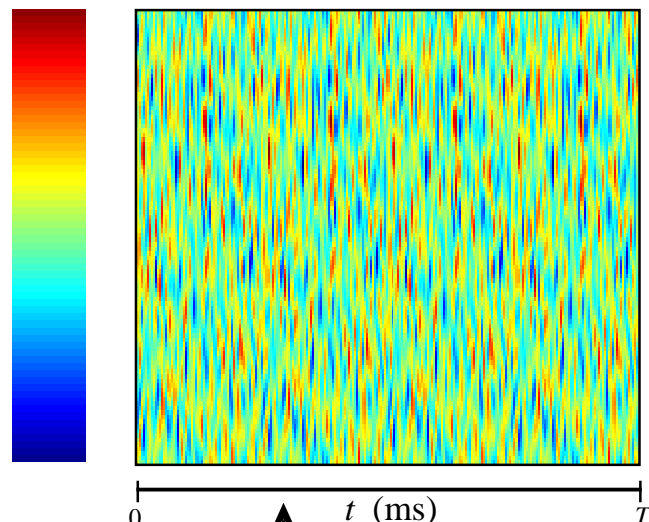
$$S^{\text{noise}}(t,x) = \sum_j \sum_k \sin[2\pi w_j t + 2\pi \Omega_k x + \phi_{j,k}]$$

Examples:

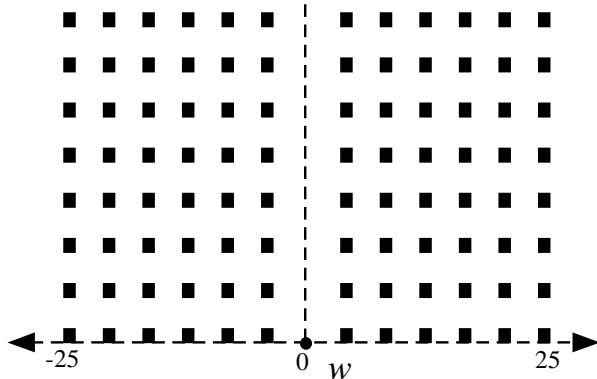
Cortical Stimulus



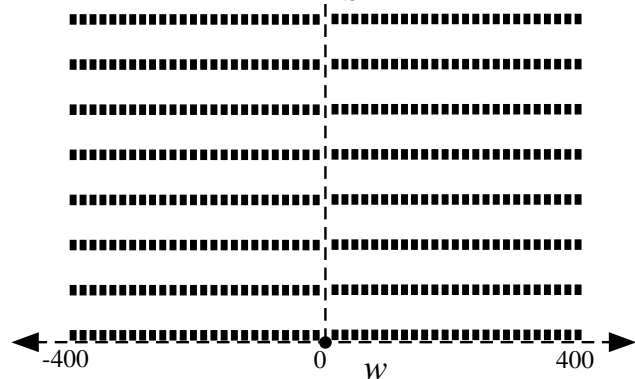
Mid-Brain Stimulus



$\updownarrow |\mathcal{F}\{\cdot\}|$
 $\Omega \uparrow 1.5$



$\updownarrow |\mathcal{F}\{\cdot\}|$
 $\Omega \uparrow 1.5$



STRF from Cross-Correlation with Noise

Stimulus: $S(t, x)$

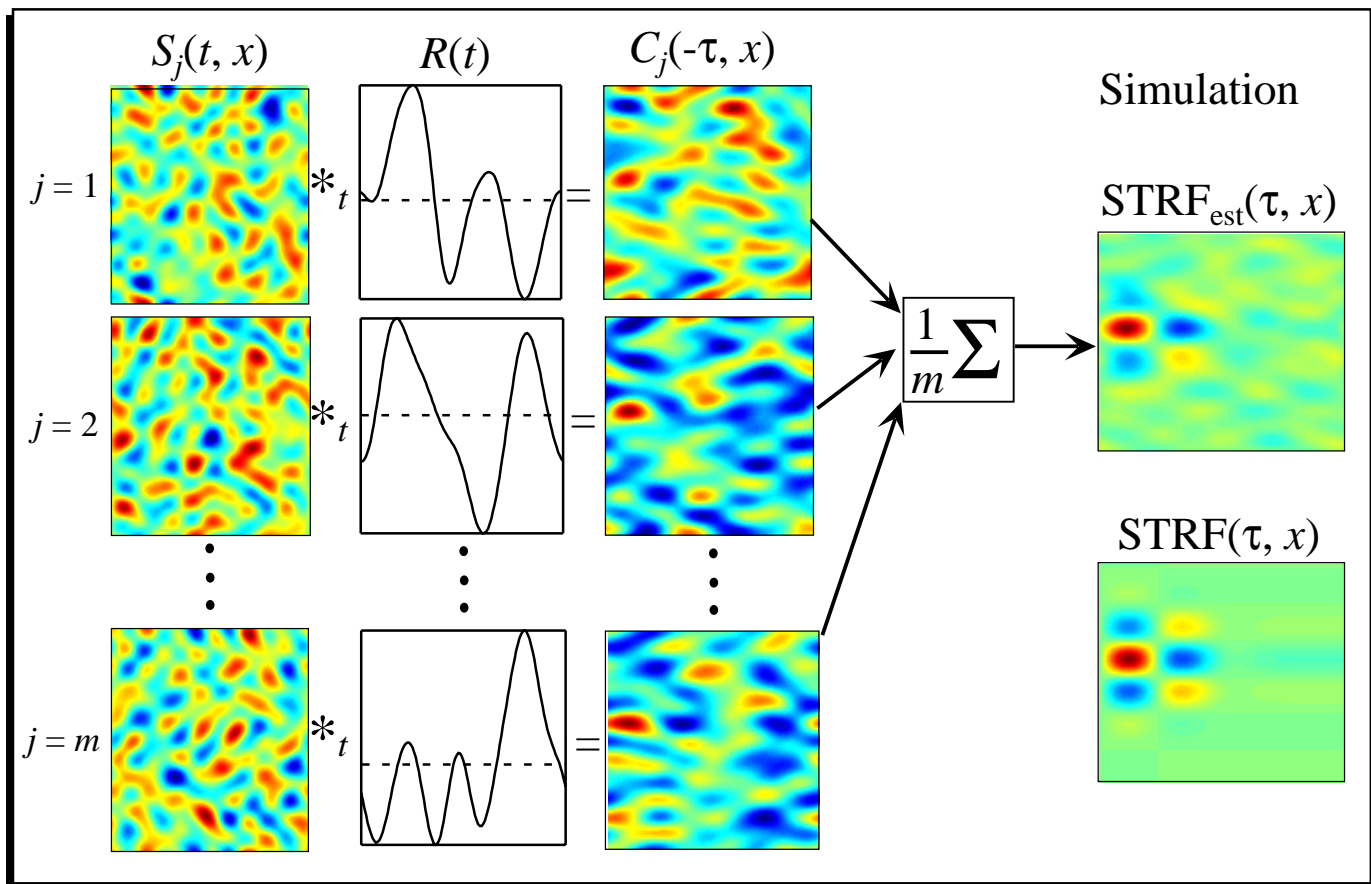
Response: $R(t) = \sum_k \delta(t - t_k)$

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.

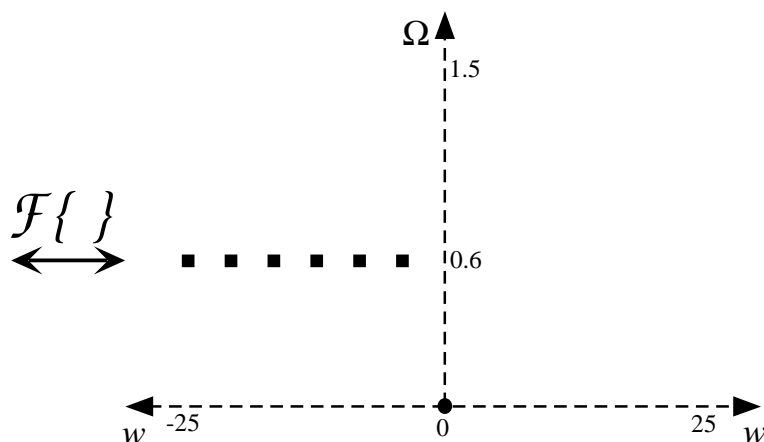
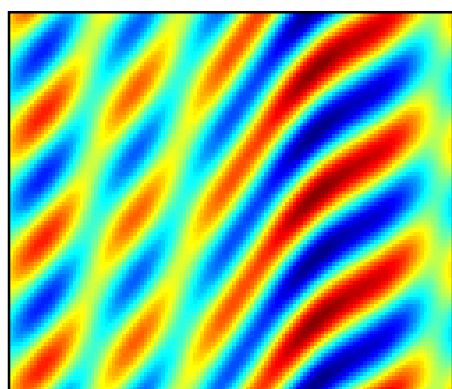
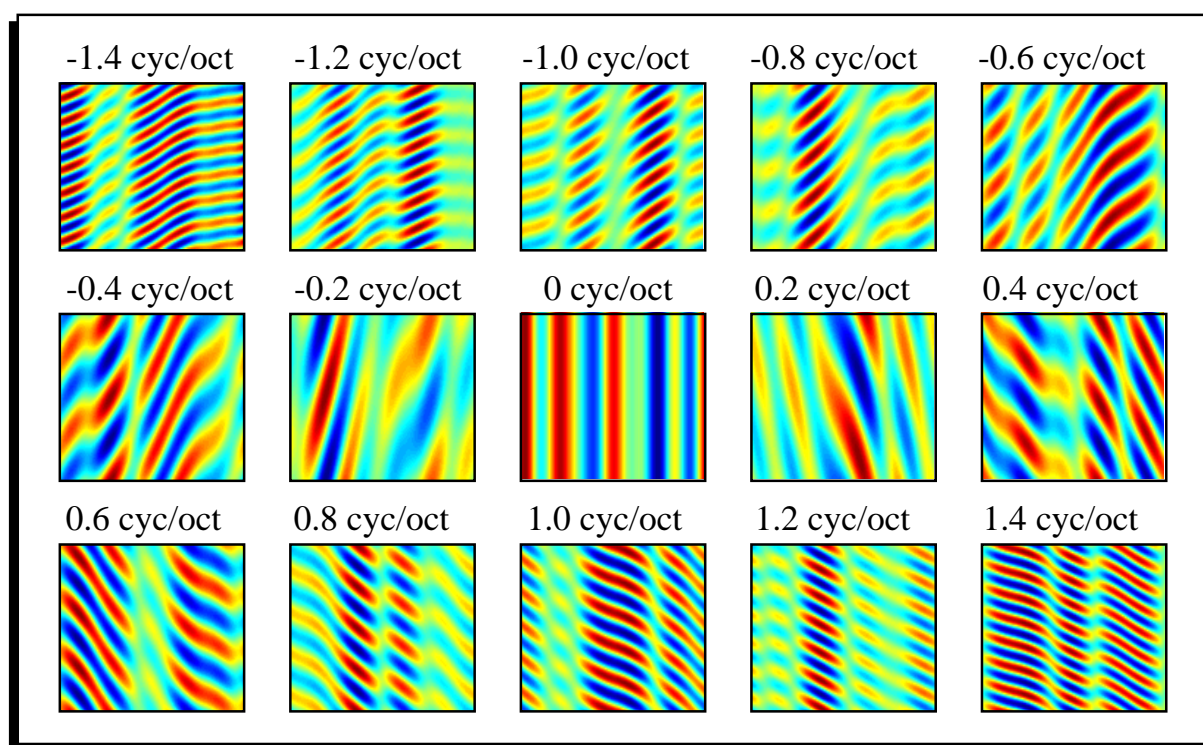
$$\text{STRF}_{\text{est}}^{\text{noise}}(\tau, x) = \frac{1}{m} \sum_{j=1}^m C_j(-\tau, x)$$



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.

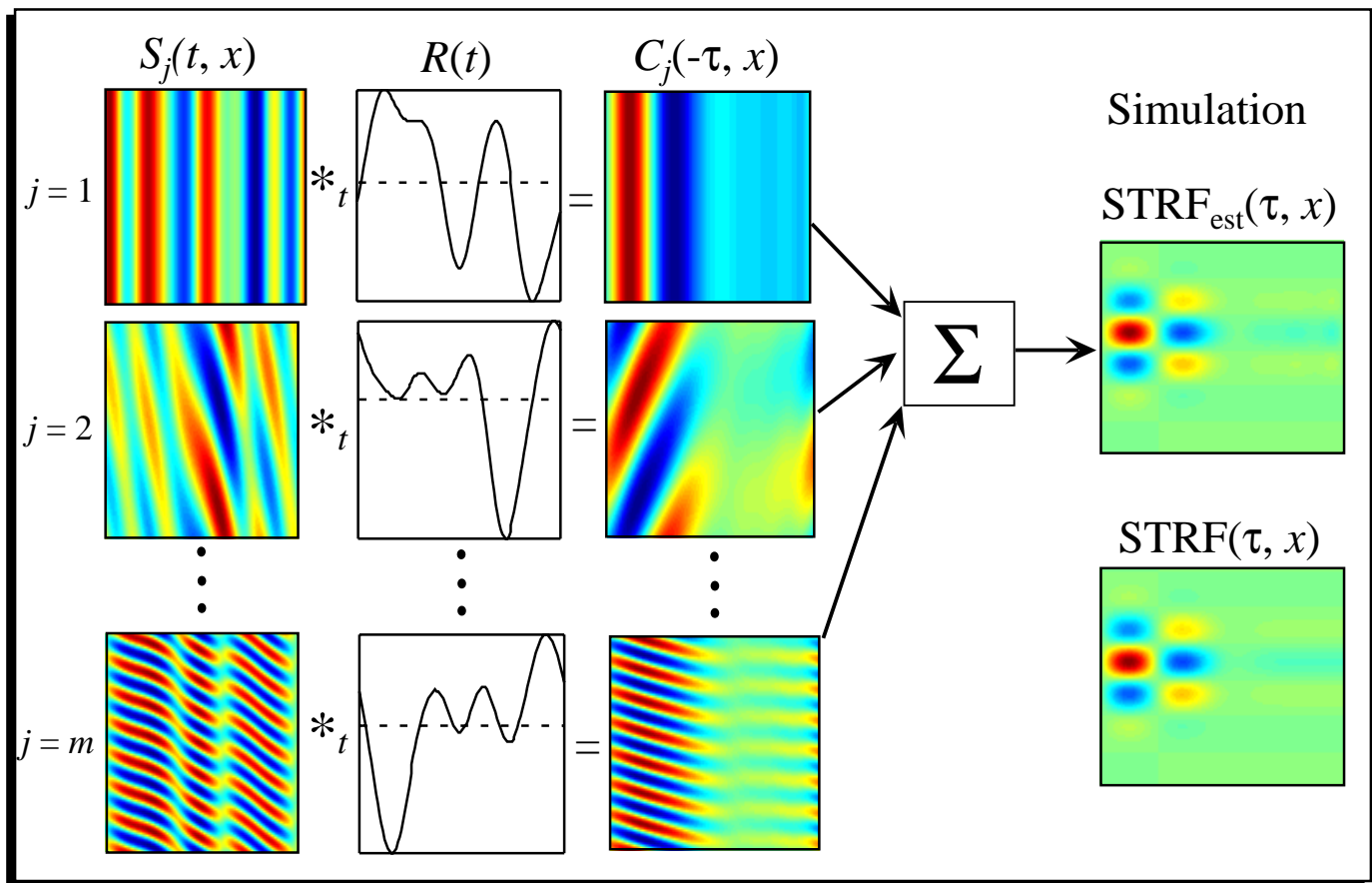
$$S_k^{\text{TORC}}(t, x) = \sum_j \sin[2\pi w_j t + 2\pi \Omega_k x + \phi_{j,k}]$$



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.

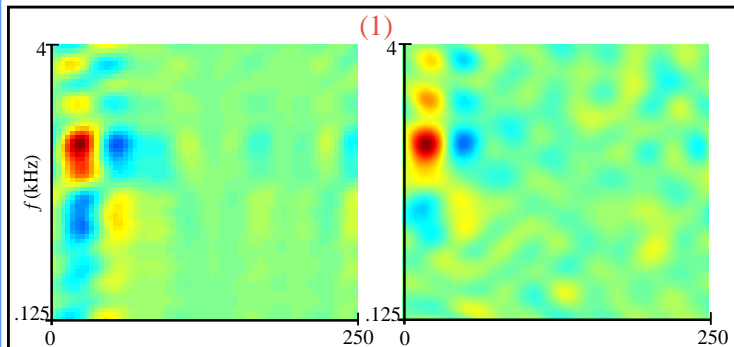
$$\text{STRF}_{\text{est}}^{\text{TORC}}(\tau, x) = \sum_{j=1}^m C_j(-\tau, x)$$



TORC Cross-Correlation vs. Single-Ripple

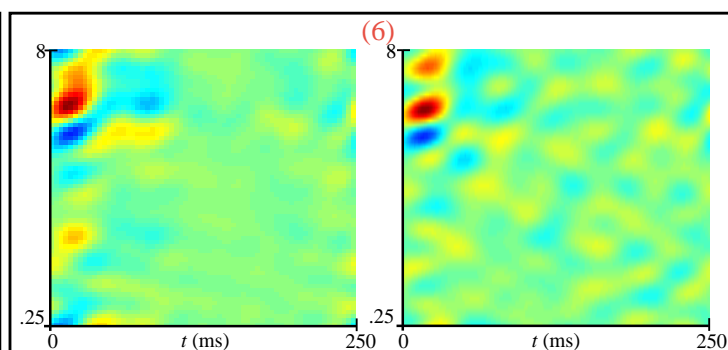
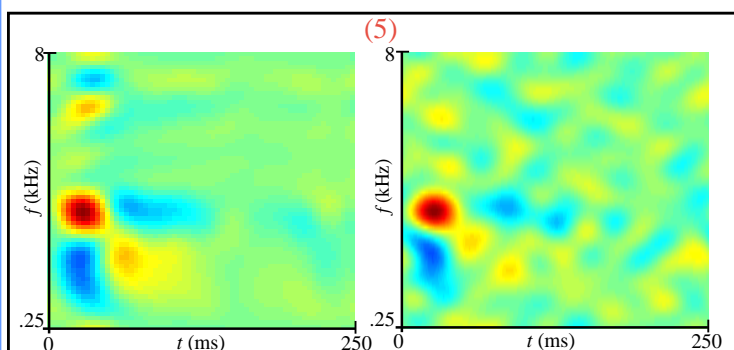
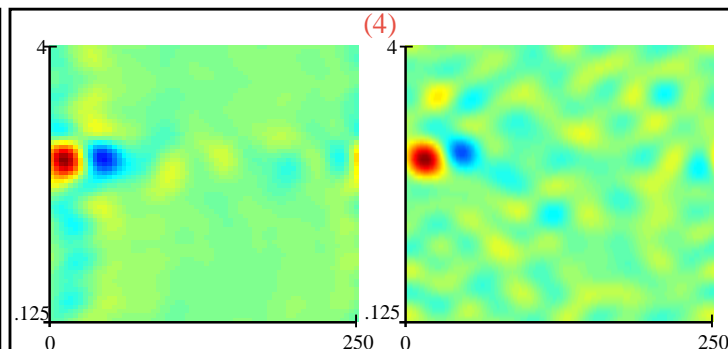
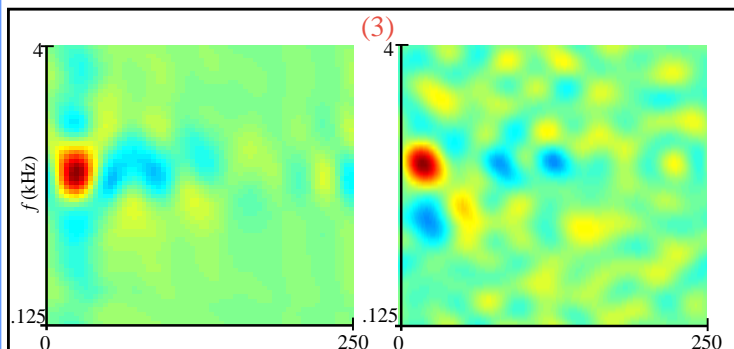
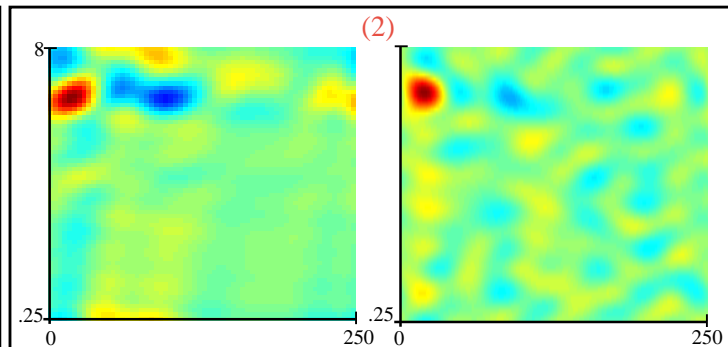
Single Ripple

Cross-Correlation



Single Ripple

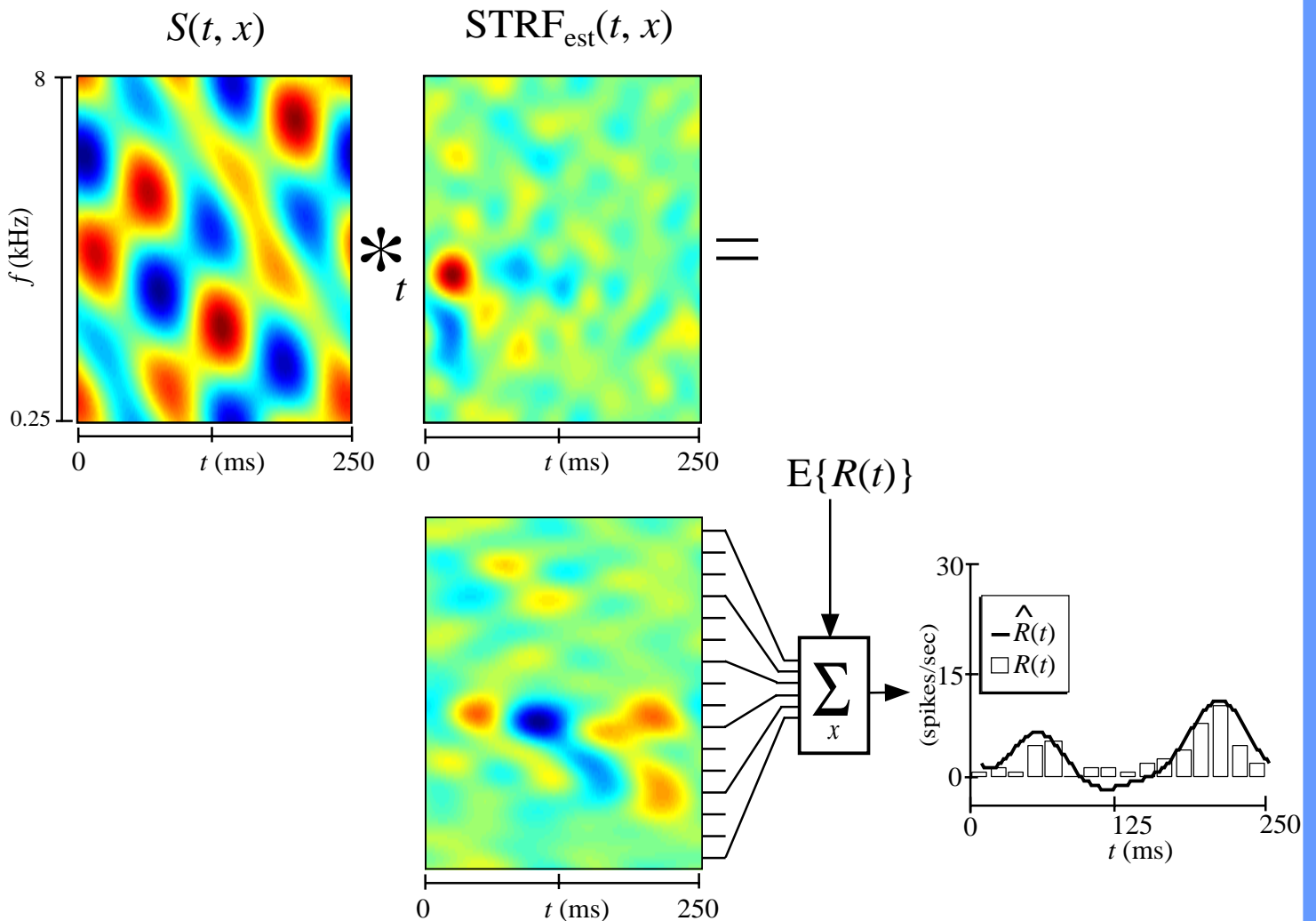
Cross-Correlation



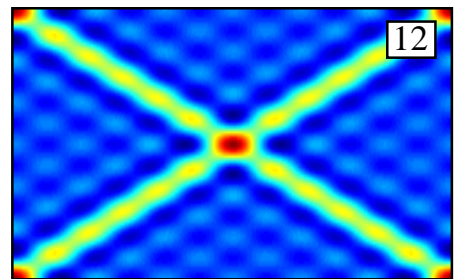
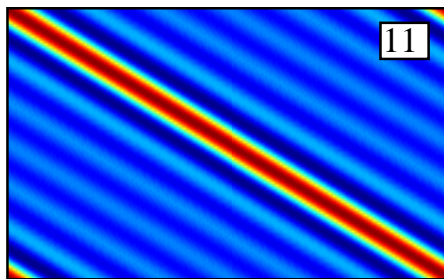
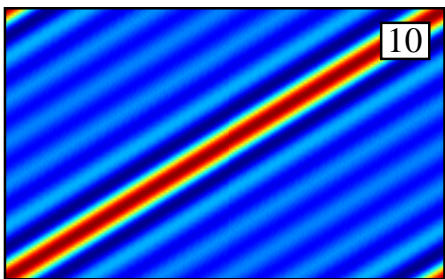
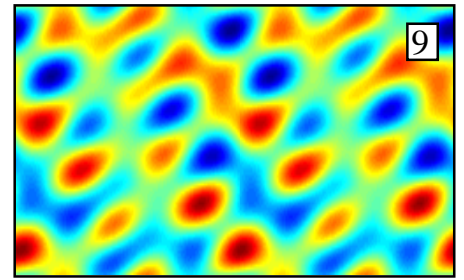
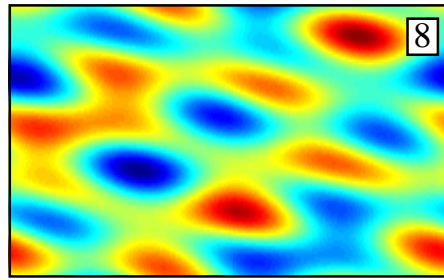
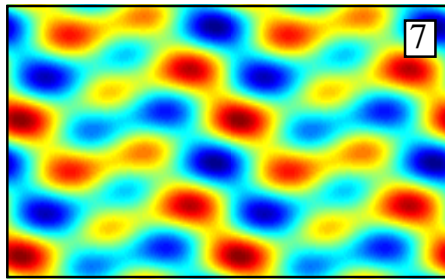
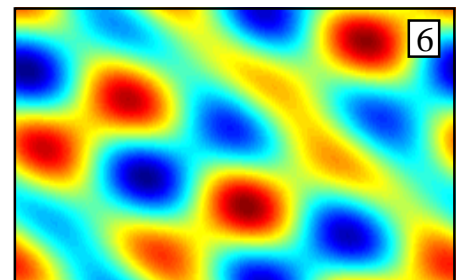
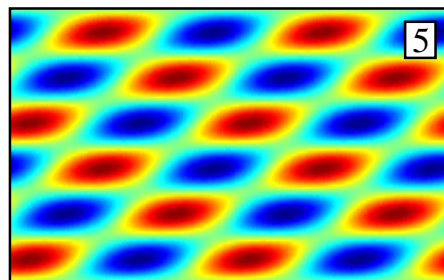
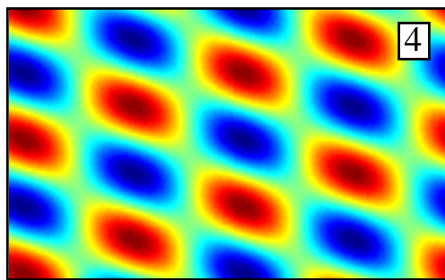
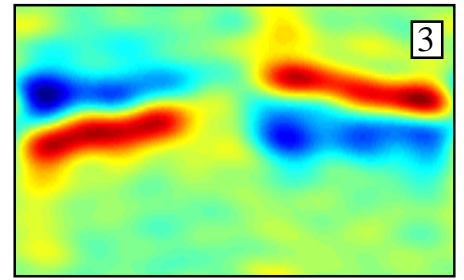
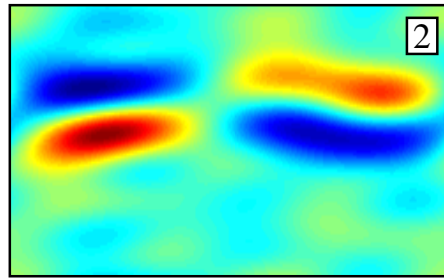
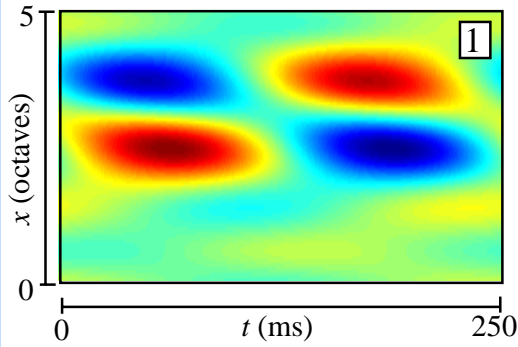
Predicting Responses from STRF

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

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

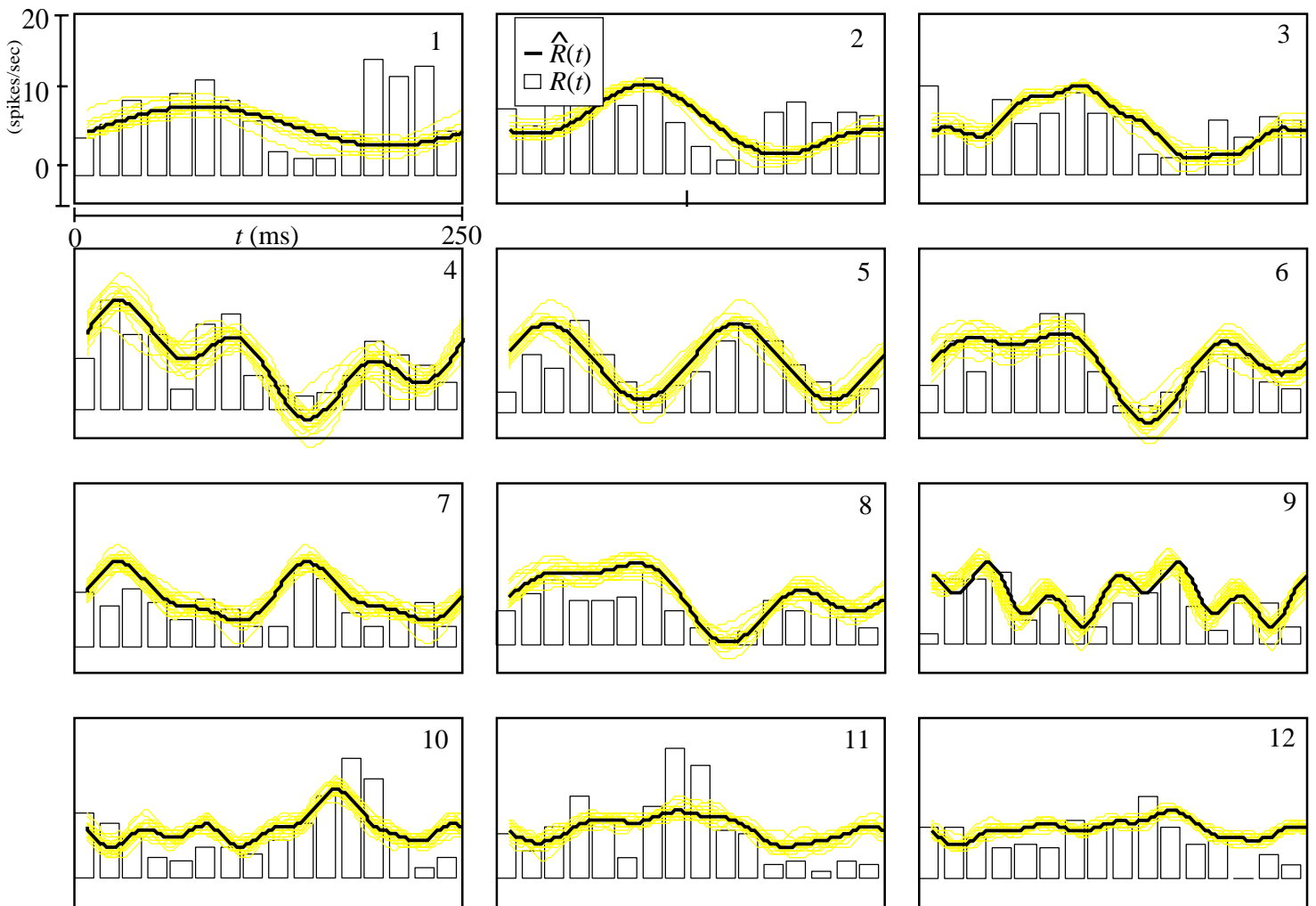
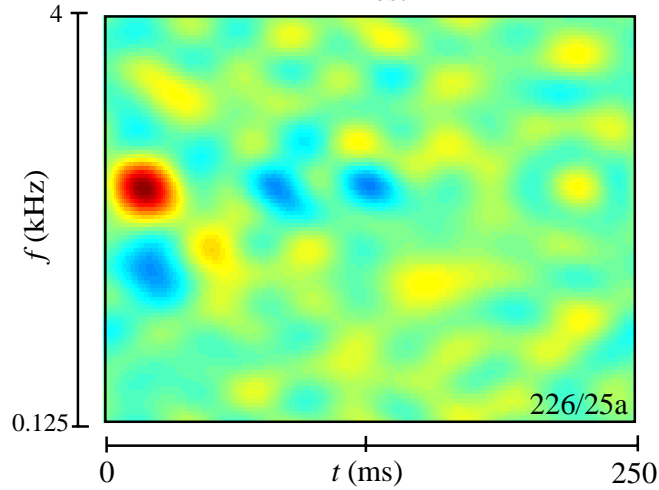


Stimuli Used for Predictions



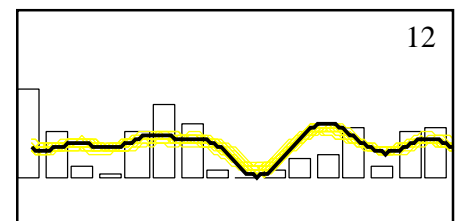
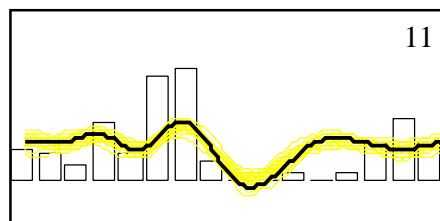
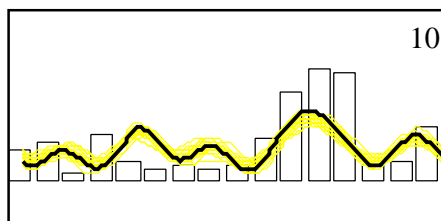
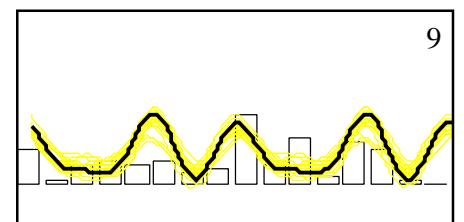
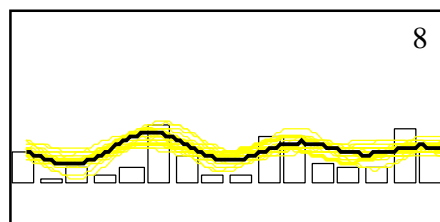
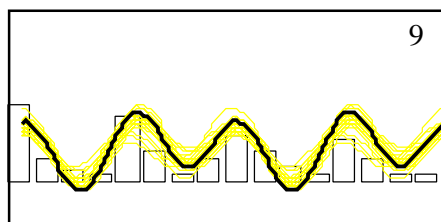
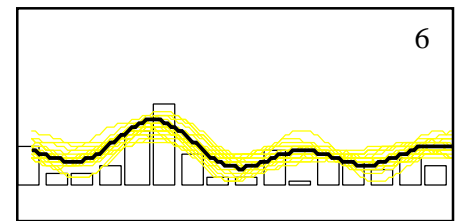
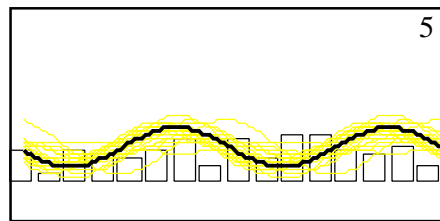
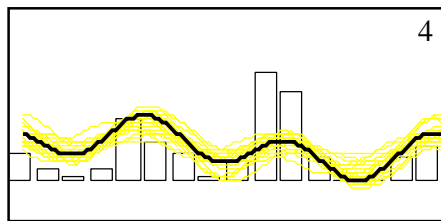
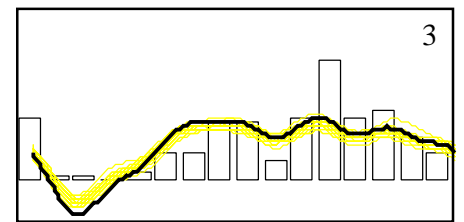
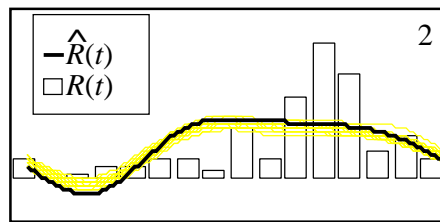
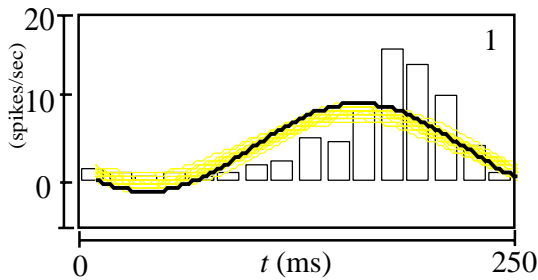
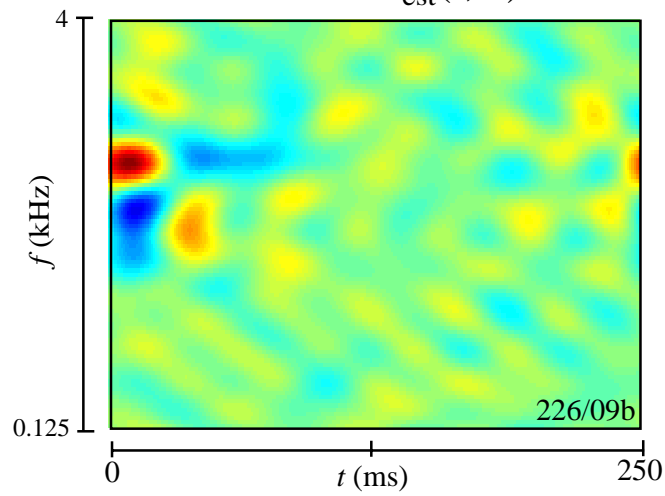
Predictions and Responses I

$$\text{STRF}_{\text{est}}(t, x)$$



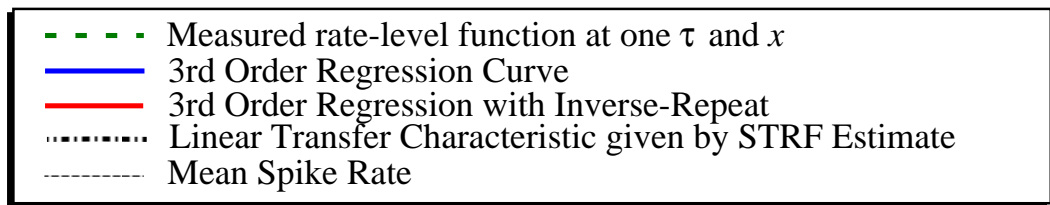
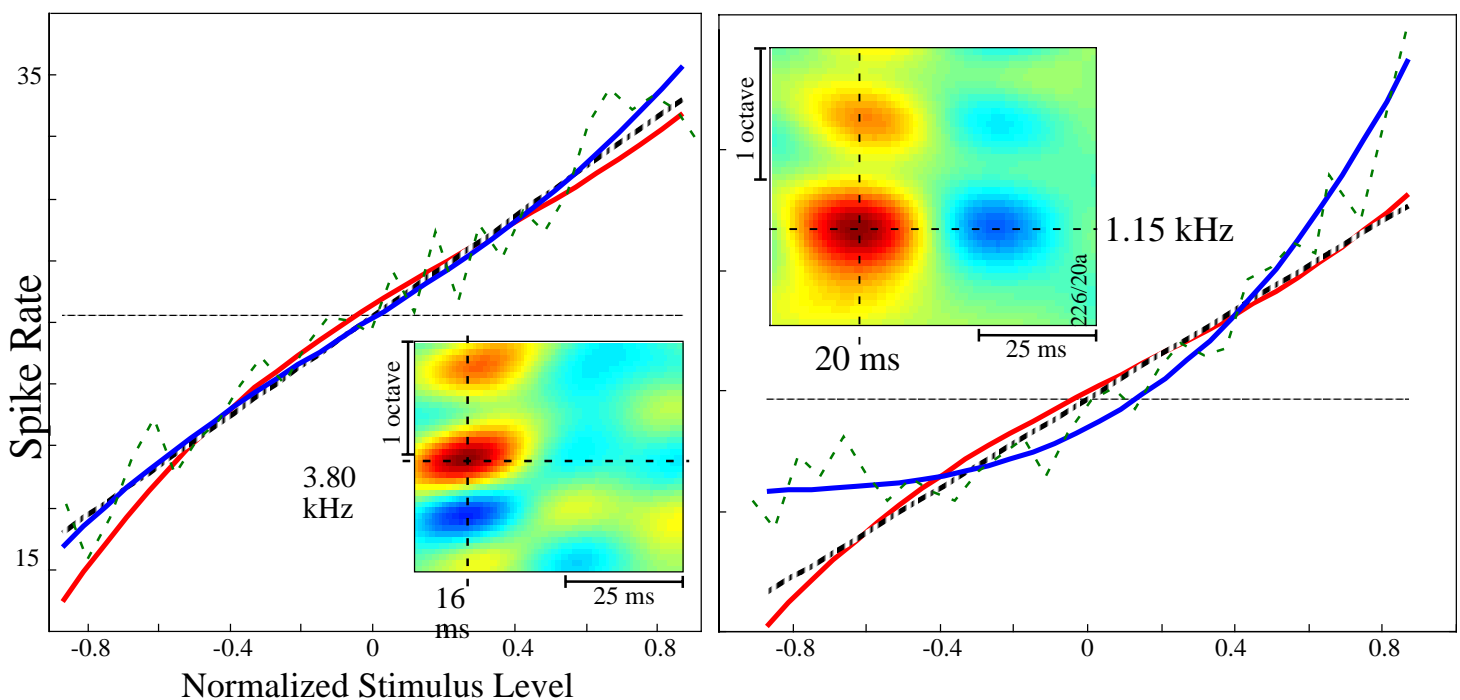
Predictions and Responses II

$$\text{STRF}_{\text{est}}(t, x)$$



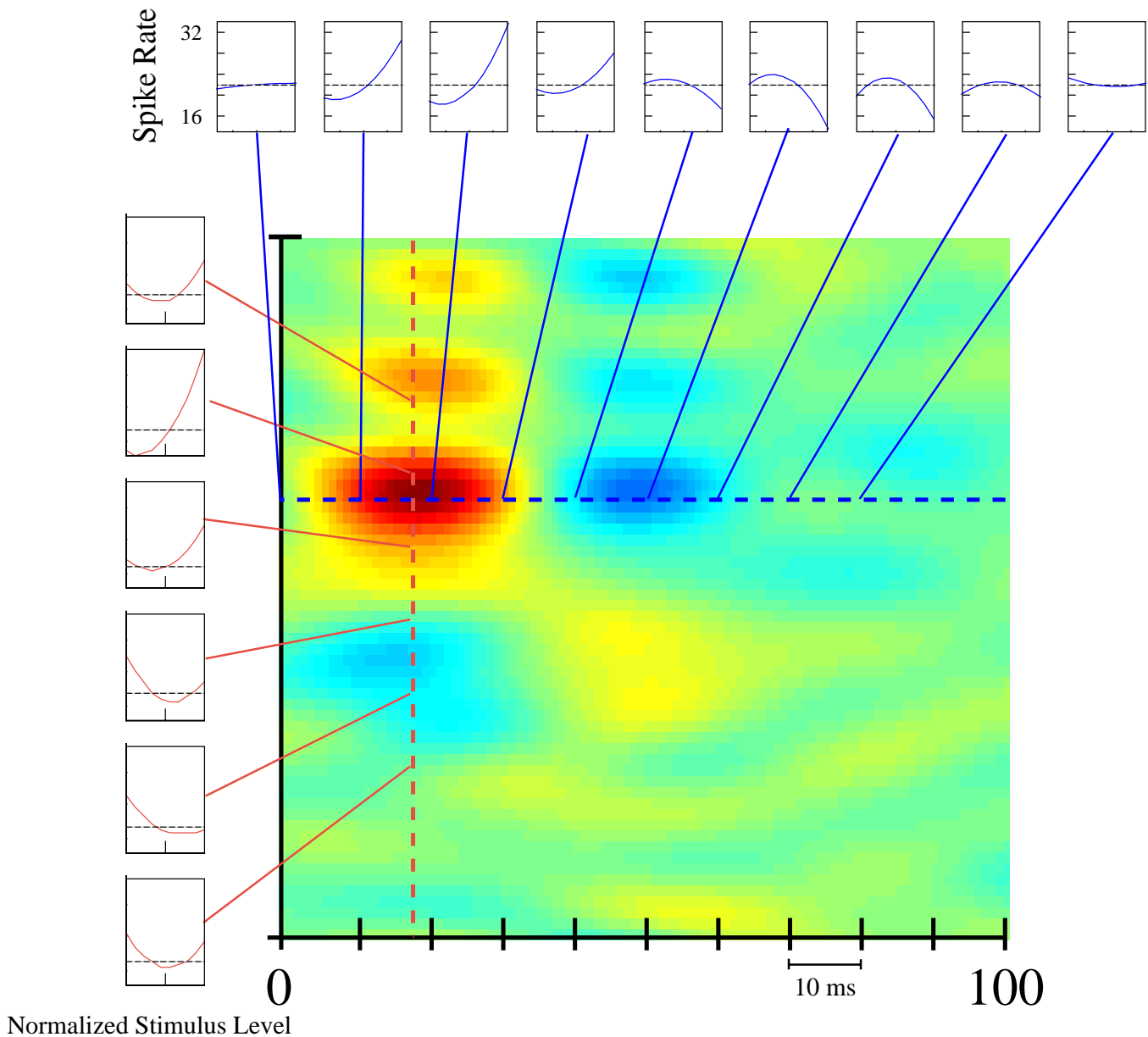
Non-Linearity

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



Spectro-Temporal Rate-Level Functions

Rate-level functions change with τ and x .



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 $S(t, x)$:

$$\begin{aligned}
 R(t) = & v_0 + \int d\tau \int dx \cdot v_1(\tau, x)S(t - \tau, x) \\
 & + \int d\tau_1 \int d\tau_2 \int dx_1 \int dx_2 \cdot v_2(\tau_1, \tau_2, x_1, x_2)S(t - \tau_1, x_2)S(t - \tau_2, x_2) \\
 & + \int d\tau_1 \int d\tau_2 \int d\tau_3 \int dx_1 \int dx_2 \int dx_3 \cdot v_3(\tau_1, \tau_2, \tau_3, x_1, x_2, x_3) \cdot \\
 & \quad S(t - \tau_1, x_1)S(t - \tau_2, x_2)S(t - \tau_3, x_3) + \dots
 \end{aligned}$$

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

$$\begin{aligned}
 g(s; \tau, x) &= E\{R(t) \mid S(t - \tau, x) = s\} \\
 &\approx a_0(\tau, x) + a_1(\tau, x) \cdot s + a_2(\tau, x) \cdot s^2 + a_3(\tau, x) \cdot s^3
 \end{aligned}$$

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

The true STRF is just the linear part...

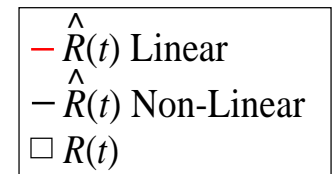
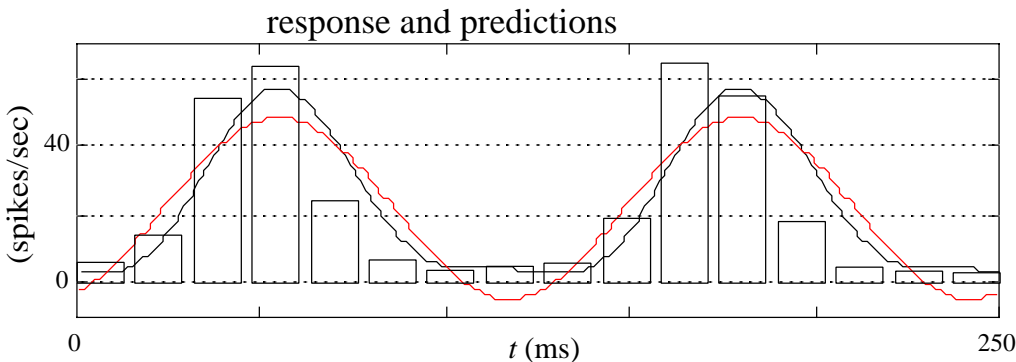
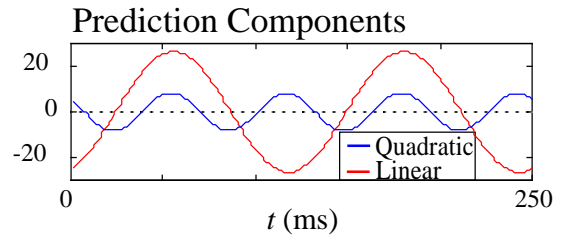
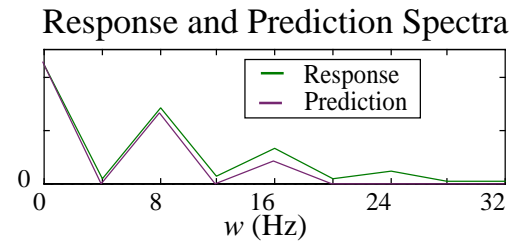
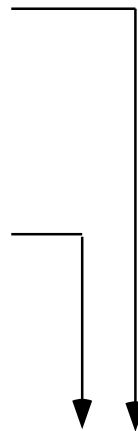
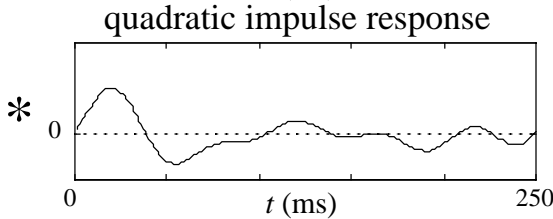
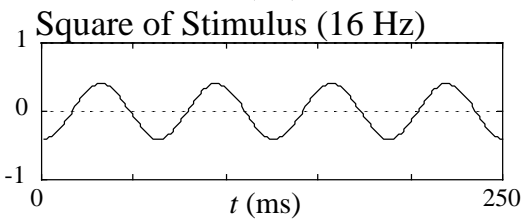
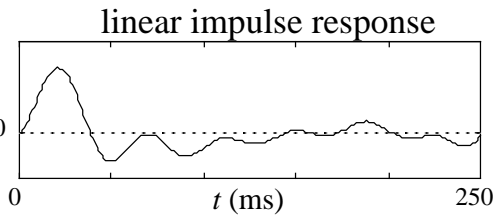
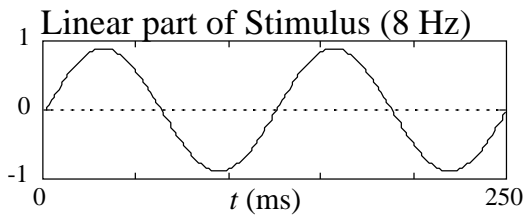
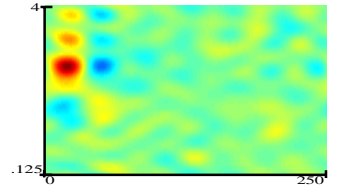
$$STRF(\tau, x) \stackrel{\Delta}{=} a_1(\tau, x)$$

Use of the Regression Function to Predict the Response:

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

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.



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, Wisner PR and Versnel H, J. Neurophys 69(2) (1993) 367-383.