Spectro-Temporal Processing of Dynamic Broadband Sounds In Auditory Cortex

Shihab Shamma
Jonathan Simon*
Didier Depireux
David Klein

Institute for Systems Research & Department of Electrical Engineering
University of Maryland

Supported in part by
Office of Naval Research MURI # N00014- 97- 1- 0501
National Institute on Deafness and Communication Disorders # T32 DC00046-01
and National Science Foundation # NSFD CD8803012

These slides are available at <http://www.isr.umd.edu/CAAR/pubs.html>
Report Documentation Page

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  1999
2. REPORT TYPE
3. DATES COVERED  00-00-1999 to 00-00-1999

4. TITLE AND SUBTITLE
Spectro-Temporal Processing of Dynamic Broadband Sounds in Auditory Cortex

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHOR(S)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Maryland, Department of Electrical 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:
   a. REPORT
      unclassified
   b. ABSTRACT
      unclassified
   c. THIS PAGE
      unclassified

17. LIMITATION OF ABSTRACT
   Same as Report (SAR)

18. NUMBER OF PAGES 29

19a. NAME OF RESPONSIBLE PERSON

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18
• Introductory Material
• Stimulus Types
• Spectro-Temporal Response Fields
• Predicting Responses to Novel Stimuli
• Non-Linearities
• Introductory Material
  • Stimulus Types
  • Spectro-Temporal Response Fields
  • Predicting Responses to Novel Stimuli
  • Non-Linearities
• The Big 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, vocalizations, backgrounds

• Reasonable quest
  Quantitative measure of how spikes encode sound features
  • Quantitative descriptor
  • Quantitative predictor
  • Visual tool—a la Spectral Response Field/Tuning Curve

• 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) meets criteria:
  • Quantitative descriptor
  • Quantitative predictor
  • Visual Tool
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?

“water all year”

• log (f) ≈ linear cochlear distance

• Characterization from cross-section is limited
Stimulus Construction

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

- **Unmodulated noise** (flat)
  - Power
  - flat power over entire spectrum

- **Spectrally modulated noise**
  - Power
  - dimished power
  - enhanced power

- **Spectro-Temporally modulated noise**
  - Power
  - enhanced power
  - dimished power

[Images and diagrams showing various power distributions across different frequencies and time scales.]
Spectro-Temporal Response Field (STRF)

- Spectro-temporal response field of neuron is the standard response field, made time-dependent.
- Frequencies mapped along cochlea on log frequency axis.
### Interpreting STRFs

**Stimulus Effect on Rate**

<table>
<thead>
<tr>
<th>STRF region</th>
<th>Stimulus Power</th>
<th>Spike rate contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excitatory</td>
<td>Enhanced</td>
<td>Faster</td>
</tr>
<tr>
<td>Inhibitory</td>
<td>Enhanced</td>
<td>Slower</td>
</tr>
<tr>
<td>Excitatory</td>
<td>Diminished</td>
<td>Slower</td>
</tr>
<tr>
<td>Inhibitory</td>
<td>Diminished</td>
<td>Faster (!)</td>
</tr>
</tbody>
</table>

**Cross-section interpretations**

**Spectral Response Fields,** evolving in time

**Impulse Responses,** parametrized by spectral band

**Spectral Response Fields,** evolving in time
Spectro-Temporal Response Field (STRF) & its Transfer Function

- Its Fourier transform is the transfer function.
- Analysis is often conceptually simpler in the Fourier domain.
- Either can be used to predict the linear response to any broadband dynamic sound.

2 Dimensional Fourier Transform

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

Inverse Transform

2 Dimensional Transfer Function of the same neuron

\( \Omega \)

ripple frequency (cycles/octave)

ripple velocity (Hz)
• Introductory Material

• **Stimulus Types**

• **Spectro-Temporal Response Fields**

• Predicting Responses to Novel Stimuli

• **Non-Linearities**
Ripples are auditory “gratings” whose spectral envelope is a sinusoid along the log(f) 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, e.g. 4 Hz = 4 cycles/s  
$\Omega$ = ripple density, e.g. 0.4 cycles/octave = 2 cycles/5 octaves

The Fourier transform of a single moving sinusoid has support only on a single point (and its complex conjugate).
The Fourier space of the spectrograms. We probe a cell at different velocities \( w \) and different densities \( \Omega \), 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.
Spike events in (A) are turned into period histograms in (B). The amplitudes and phases give the transfer function in (C), which can be inverse Fourier transformed to give Response Fields in (D).
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.

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

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

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

Simulation
Temporally Orthogonal Ripple Combinations (TORCs)

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

- 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.

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

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.

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

Simulation
• Introductory Material
• Stimulus Types
• **Spectro-Temporal Response Fields**
• Predicting Responses to Novel Stimuli
• Non-Linearities
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.

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

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 strongly to rising frequencies as it did to falling frequencies.
Singular Value Decomposition (SVD) decomposes the matrix into a sum of separable matrices, ordered by their overall magnitudes. The first $k$ components sum to a matrix 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.

Without prior assumptions, SVD indicates that a large majority of STRFs in AI are quadrant separable.
- Introductory Material
- Stimulus Types
- Spectro-Temporal Response Fields
- Predicting Responses to Novel Stimuli
- Non-Linearities
Predicting Responses

- Predictions of Responses to Novel Stimuli
  Good test of linearity

Stimuli Used for Predictions

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>12</td>
</tr>
</tbody>
</table>
Predictions and Responses

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

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

0 250
0 250

1 2 3
4 5 6
7 8 9
10 11 12

0 250
0 250

0.125 f (kHz)
0.125 f (kHz)

(spikes/sec)
(spikes/sec)

R(t)
R(t)

R(t)
R(t)

226/09b
226/09b

Center for Auditory and Acoustic Research
Institute for Systems Research
University of Maryland
• Introductory Material
• Stimulus Types
• Spectro-Temporal Response Fields
• 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.
Non-Linearity—Theory

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

\[ S_j(t, x) \]
\[ R(t) \]

\[ \begin{align*}
\text{Example 1: Cell} & \quad 228/08a10 \\
\tau &= 16 \text{ ms}, \; x = 3.80 \text{ kHz} \\
\text{Example 2: Cell} & \quad 226/20a06 \\
\tau &= 20 \text{ ms}, \; x = 1.15 \text{ kHz}
\end{align*} \]

- Using cubic polynomials, we have shown that either the nonlinearities 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).
Rate-level functions change with $\tau$ and $x$. 
Summary

• The function of AI
  To encode spectro-temporal features of sounds
  spectrally: up to ~ 1 cycles/octave (rarely up to 4 c/o)
  temporally: ~ 2 to ~ 20 Hz (rarely up to 100 Hz)
  (in ferret)
  plus, of course, to encode other sound features not
  addressed here

• Spectro-Temporal Response Field (STRF)
  • Descriptor of spike rate for broadband dynamic stimuli
  • Predictor of spike train for stimuli of dynamic,
    spectral modulations of noise
    • STRFs agree despite measurement method
    • Predictions of responses to novel stimuli

• Visual Tool demonstrates spectro-temporal regions of
  excitation and inhibition

• Physiological Constraints (e.g. Separability) constrain
  possible network dynamics

• Non-linear corrections/generalizations
  realizable and work in progress
Selected References

Moving Ripple / STRF / Ripple Transfer Function

Spectro-Temporal Correlation Methods

Singular Value Decomposition

Polynomial Correlation

Related techniques and models