Empirical estimates of STRFs are influenced by higher-order stimulus statistics

Jörn Anemüller
+Arne Meyer, +Jörg-Hendrik Bach
+Jan-Philipp Diepenbrock, +Frank Ohl

Computational Audition Group
Medical Physics Section and Cluster of Excellence Hearing4all
Carl von Ossietzky University
Oldenburg, Germany
Listening in Acoustic Scenes
Listening in Acoustic Scenes
Listening in Acoustic Scenes

Figure 9. Intersection where traffic noise recording was carried out.

Figure 10. Example spectrograms of typical events in recorded scenes: chair scratching on the floor of the cafeteria (left panel) and accelerating motorbike at a green traffic light (right panel). Abscissa represents time, ordinate frequency, light colours high spectro-temporal signal power, dark colours low power.
3.5 Downtown pedestrian area

This recording was conducted using the mobile setup on a Saturday around noon time in the city centre of Oldenburg with many pedestrians in the area as seen in Fig. 11. The whole recording covers about one hour and contains a round tour through the pedestrian area with two stops of about five minutes each for recordings from fixed positions. Ambient noise mainly consists of pieces of conversation and babbling in the background, music from stores and from outside presentations.

3.6 Audio-visual pilot recordings

At the DIRAC project meeting in Leuven (September 5/6, 2006) audio-visual test recordings were done in cooperation with the research group from the Center for Machine Perception at the Czech Technical University in Prague. Together with the mobile visual recording setup from CTU the recordings were performed in an office room at the building of the Katholike Universiteit Leuven and outside in a parking lot. Different setups of the visual equipment were tested, e.g., the frame rate for image acquisition was varied and pictures were taken in colour and black and white. Overall six audio-visual recordings with a total duration of about seven minutes were captured. At the same venue, about 80 minutes of audio-only recordings were performed, capturing a meeting in a large conference room using the HATS with 6-channel hearing aid microphones (no inner ear microphone channels).
Listening in Acoustic Scenes

Figure 9. Intersection where traffic noise recording was carried out.

Figure 10. Example spectrograms of typical events in recorded scenes: chair scratching on the floor of the cafeteria (left panel) and accelerating motorbike at a green traffic light (right panel). Abscissa represents time, ordinate frequency, light colours high spectro-temporal signal power, dark colours low power.

3.5 Downtown pedestrian area
This recording was conducted using the mobile setup on a Saturday around noon time in the city centre of Oldenburg with many pedestrians in the area as seen in Fig. 11. The whole recording covers about one hour and contains a round tour through the pedestrian area with two stops of about five minutes each for recordings from fixed positions. Ambient noise mainly consists of pieces of conversation and babbling in the background, music from stores and from outside presentations.

Figure 11. City centre of Oldenburg where recordings were performed.

3.6 Audio-visual pilot recordings
At the DIRAC project meeting in Leuven (September 5/6, 2006) audio-visual test recordings were done in cooperation with the research group from the Center for Machine Perception at the Czech Technical University (CTU) in Prague. Together with the mobile visual recording setup from CTU the recordings were performed in an office room at the building of the Katholike Universiteit Leuven and outside in a parking lot. Different setups of the visual equipment were tested, e.g., the frame rate for image acquisition was varied and pictures were taken in colour and black and white. Overall six audio-visual recordings with a total duration of about seven minutes were captured. At the same venue, about 80 minutes of audio-only recordings were performed, capturing a meeting in a large conference room using the HATS with 6-channel hearing aid microphones (no inner ear microphone channels).
auditory object properties

spectro-temporal  spatial

machine learning / statistics

biological data (spectro-temporal)

audio data

psychoacoustics & signal processing (spatial)
Historical approaches to neuron characterization
Historical approaches to neuron characterization

(a) Classical analytical approach

Pure tones

Frequency tuning curve

Threshold (dB)

Frequency (kHz)

(b) Neuroethological approach

Vocalization

Frequency

Time

Mean rate (z value)

Pure tones, Compound tones, White noise, Conspecific song

(Theunissen and Elie, 2014)
Hubel and Wiesel’s visual receptive fields
Hubel and Wiesel’s visual receptive fields
Statistical approach of neural characterization

- Goal: infer response properties from stimulus and evoked response
Simple cell linear receptive field model

\[
y = g \left( \sum_{j=1}^{d} w_j x_j + w_0 \right)
\]

(Kandel et al., 2000)
Cies. Rates of 10 to 100 chords/s have been used, as well as stimuli with
comprised of a variable number of tones of randomly selected frequen-
ty behavior, such as bats (R. Christopher deCharms, David T. Blake, 5
rons in the awake primary auditory cortex 5
vies from 84 possible val-

The density of tones in 1998), where stimuli have been se

studies have consisted of rapidly presented

perceived sounds? A reverse correlation technique demonstrates that neurons in the

cortex. In 1998, this unbiased method constructs the av

discern the full spectral and temporal fea

tory domain consisted of 1998, which has demonstrated that au

A Visual Cortex: Reverse Correlation

Using 2D Visual Patterns in Time

t = 0 ms

t = 20 ms

t = 40 ms

Spatiotemporal Receptive Field

B Auditory Cortex: Reverse Correlation

Using 1D Auditory Patterns (Chords) in Time

t = 0 ms

t = 200 ms

t = 400 ms

Spectrotemporal Receptive Field

deCharms, Blake, Merzenich (1998)

The reverse correlation method is similar to the methods used

tory neurons are tuned for a number of

independent feature parameters of simple

Auditory neurons are tuned for a number of

independent feature parameters of simple

dependent manner (Liu et al., 2007), where stimuli have been se

specific vocalizations (Kawahara et al., 2007), which has demonstrated that au

or both. Fewer frequencies are shown in the diagram than were actually

A reverse correlation technique demonstrates that neurons in the

edges, direction and velocity information, and color. How does the cortex decompose

visual system, re-

selectivity that indicate sensitivity to stimulus edges in frequency or in time, stimulus

edges, direction and velocity information, and color. How does the cortex decompose

ural decomposition of visual images.

cortical neurons reflect an underlying

processing characteristics of auditory cortical

stimuli have been applied to more peripheral struc

have only recently begun to succeed in ap

vice neurons in

discern the full spectral and temporal fea

tonal stimuli are typically much lower than

measurement of the one

mention of the one

t = 0 ms t = 0 ms

t = 200 ms t = 200 ms

t = 400 ms t = 400 ms

visual space

visual space

frequency space

frequency space

Spike

X

X

Trains

1998, this unbiased method constructs the av

deCharms, Blake, Merzenich (1998)

A Visual Cortex: Reverse Correlation

Using 2D Visual Patterns in Time

t = 0 ms

t = 20 ms

t = 40 ms

Spatiotemporal Receptive Field

B Auditory Cortex: Reverse Correlation

Using 1D Auditory Patterns (Chords) in Time

t = 0 ms

t = 200 ms

t = 400 ms

Spectrotemporal Receptive Field

deCharms, Blake, Merzenich (1998)
Stimulus = Point in feature space

A  
Spike-triggered ensemble

Stimulus ensemble

Response

B  
Stixel 1

Stixel 2

Eggermont, Johannesma, Aertsen (1983)
Schwartz, Pillow, Rust, Simoncelli (2006)
Spike triggered average (STA)

2. Compute the filter responses for the stimulus, and estimate the nonlinear firing rate function based on these responses. As noted earlier, typical physiological data sets allow nonparametric estimates of the nonlinearity for one or two filters but require more model restrictions as the number of filters increases.

In the following subsections, we describe these steps in detail. In the Experimental issues section, we also stress the importance of an additional step: validating the resulting model by comparing it to neural responses from other stimuli.

Subspace (filter) estimation

In general, one can search for any deviation between the raw and spike-triggered stimulus ensembles. This can be done, for instance, using measures of information theory (Paninski, 2003; Sharpee et al., 2003, 2004). Another natural approach is to consider only changes in low-order moments between the raw and spike-triggered stimulus. Here, we focus on changes in the first and second moments, which may be computed efficiently and manipulated using a set of standard linear algebraic techniques. We also briefly discuss how the analysis relates to the Wiener/Volterra approach.

Spike-triggered average

The simplest deviation between the spike-triggered and raw stimulus distributions is a change in the mean. Assuming that the raw stimuli have zero mean, this can be estimated by computing the average of the spike-triggered ensemble (STA):

$$
\hat{A} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{s}_Y(t_n);
$$

where $t_n$ is the time of the $n$th spike, $\mathbf{s}_Y(t_n)$ is a vector representing the stimuli presented during the temporal window preceding that time, and $N$ is the total number of spikes. In practice, the times $t_n$ are binned. If there is more than one spike in a bin, then the stimulus vector for that time bin is multiplied by the number of spikes that occurred. The STA is illustrated in Figure 3A.

For an LNP model with a single linear filter, the STA provides an unbiased estimate of this filter, provided that the input stimuli are spherically symmetric (Bussgang, 1952; Chichilnisky, 2001; Paninski, 2003), and the nonlinearity of the model is such that it leads to a shift in the mean of the spike-triggered ensemble relative to the raw ensemble (see Limitations and potential failures section and Experimental issues section). This last requirement rules out, for example, a model with a symmetric nonlinearity such as full-wave rectification or squaring.

For an LNP model with multiple filters, the STA provides an estimate of a particular linear combination of the model filters, subject to the same restrictions on input stimuli and the form of the nonlinearity given above (Paninski, 2003; Schwartz et al., 2002). That is, the STA lies in the subspace spanned by the filters, but one cannot assume that it will exactly represent any particular filter in the model.

Figure 3.

Two alternative illustrations of STA. (A) The STA is constructed by averaging the spike-triggered stimulus segments (red boxes), and subtracting off the average over the full set of stimulus segments. (B) Geometric (vector space) depiction of spike-triggered averaging in two dimensions. Black points indicate raw stimuli. White points indicate stimuli eliciting a spike. The STA, indicated by the line in the diagram, corresponds to the difference between the mean (center of mass) of the spike-triggered ensemble and the mean of the raw stimulus ensemble.
Response characteristics in A1

deCharms, Blake, Merzenich (1998)
Stimuli: Dynamic moving ripples (DMR) and ripple noise
STRF estimation with STA

Escabi, Schreiner (2002)
Response characteristics in IC

- Image of a graphical representation showing spectrotemporal receptive fields (STRFs) for different types of responses in IC neurons.
- The image includes color-coded maps indicating the firing rate and delay, with examples labeled as type I, II, and III.
- The graphs show how neurons respond to dynamic ripple, ripple noise, and different frequency and octave combinations.
- The maps are color-coded, with red indicating higher firing rates and blue indicating lower rates.
- The data is consistent with previous studies by Escabi, Schreiner (2002).
The correlation coefficients were 0.45 (range, 0.12–0.66) for the song ensemble (Fig. 10E, "CC-matched"). These two distributions of correlation coefficients were very similar (p < 0.07, Wilcoxon signed rank test). In both cases, there was a wide range of fits showing that the linear model is a good approximation for some neurons and a much poorer approximation for others. However, both the response to tone pips and songs could be modeled with similar effectiveness as long as the STRF used in the fitting was obtained with the same type of stimulus ensemble.

The picture changed radically when the STRFs were switched, so that the predicted response to a stimulus from one type of ensemble was generated with the STRF obtained using the other ensemble. Figure 10A–C contrast the actual data with the predicted responses obtained with the matched STRF and with...
STA not optimal for correlated stimuli

Theunissen, Sen, Doupe (2000)
Correcting for stimulus correlations

- Normalized reverse correlation (NRC)

\[ r = Sh + e \]
\[ h_{\text{ideal}} = \frac{1}{T} C_{ss}^{-1} S^T r \quad C_{ss} = S^T S / T \]

- Ridge Regression
  Regularize \( C_{ss} \) with additive diagonal term
  (Often done on spike rate)

Machens, Wehr, Zador (2004); David, Mesgarani, Shamma (2007); Park, Pillow (2011)
Linear-nonlinear Poission (LNP) model

Equation 8

\[ \text{Firing rate} = P(S|\text{Spike}) \]

Chichilniski, 2001; Schwartz, Pillow, Rust, Simoncelli (2006)
Linear-nonlinear Poission (LNP) model

Chichilniski, 2001; Schwartz, Pillow, Rust, Simoncelli (2006)
LNP estimation: generalized linear models (GLM) and maximally informative dimensions (MID)

**Figure 1** Geometric interpretation of the receptive field (RF) in the context of the linear and linear-nonlinear (LN) models. An example stimulus is a natural image taken from a van Hateren data set (van Hateren & van der Schaaf 1998). The stimulus has \( d \) pixels in the horizontal and vertical dimensions, which yields a stimulus of \( D = d^2 \) dimensions. The RF taken to mimic properties of V1 neurons is also defined in this space. The linear model predicts the spike probability as taking a projection between the stimulus and the RF. The LN model adds a nonlinear gain function to account for such properties as rectification and saturation in the neural response.

**LINEAR MODEL**

To predict the firing rate of a neuron to a novel stimulus using a linear model, one can compare how similar that stimulus is to the optimal pattern, i.e., the RF. Mathematically, this corresponds to multiplying stimulus values pixel by pixel by the RF values and summing across all pixels. In this interpretation, the RF becomes the weighting function according to which stimulus values are combined to obtain the firing rate (Figure 1). The linear model also includes the coefficient of proportionality between the stimulus similarity to the RF and the neural firing rate. This coefficient of proportionality, referred to as the "gain," is the same for all stimuli.

Before discussing various ways for building nonlinear models of neural responses, it is useful to explore other ways of thinking about the RF concept. If the RF has \( D \) pixels (which could include temporal profiles), then it can also be represented as a vector in a \( D \)-dimensional space. To compare each new stimulus to the RF, stimuli should be defined on the same grid of pixel values as the RF. Then, each stimulus can be considered as a vector in the same \( D \)-dimensional space. The mathematical procedure described above of weighting each stimulus value by the RF profile corresponds to the computation of a dot product between the RF and the stimulus for which we would like to obtain the firing rate prediction. In geometrical terms, this corresponds to taking a projection of a vector that describes the stimulus onto the vector that describes the RF (Figure 1).

**GLM:** maximize likelihood based on prior distribution

**MID:** search for directions that preserve stimulus information

Sharpee, Rust, Bialek (2003), Sharpee (2013)
Subspace approach: Spike-triggered covariance (STC)

Model neuron:

STA analysis:

Unstructured

Schwartz, Pillow, Rust, Simoncelli (2006)
Subspace approach: Spike-triggered covariance (STC)

Model neuron:

STA analysis: (Unstructured)

STC analysis:

Schwartz, Pillow, Rust, Simoncelli (2006)
From sounds to spikes

Classic STRF model:

- Separation into **linear part** (= receptive field or kernel $k$) and **static memoryless nonlinearity** (Chichilnisky 2001)
- Once we know $k$ estimation of the nonlinearity is quite simple!
- White noise approach: Estimation of linear part using (normalized) reverse correlation method (Bussgang Theorem 1952)
- BUT: need Gaussian (symmetric) stimuli!
Stimuli: Bank of frequency-modulated tones (FM-Bank)

Meyer, Diepenbrock, Happel, Ohl, Anemüller (2014)
Classification-based receptive field (CbRF) estimation

Training a classifier to predict spike trains

Cooperation with F. Ohl
- Experiments
- Statistical modeling
- Experimental paradigm for influence of stimulus changes

Black Box SVM Classifier

Spike ?
yes / no

Figure 33: STRF for FM-bank stimuli estimated using (a) linear regression and (b) a linear SVM classifier. About 5% of the feature dimensions have been used when training the classifier. The STRFs estimated using DMR stimuli shown in Figure 34 are different latencies.

In Figure 45 in the appendix STRFs are shown that were estimated using SVM classification for the main units of the STRF obtained for FM-bank stimuli. The STA-based STRF is sensitive to different stimulus changes. However, the neurons at different positions in the auditory cortex have different STRFs. The neurons at the same position for different units of the same stimulus are similar to the unit presented here. Especially the STRFs estimated using SVM classification with data from the same electrode but assigned to the tonotopic organization of the auditory system we assume to maximal excitatory regions and red to maximal inhibitory regions and red.

Figure 5: An ensemble of spike trains obtained for five repetitions to repeat our measurements to obtain an ensemble of realizations, e.g. from 5 repetitions of the same stimulus for cross-validation.

The spike rate of a neuron is simply given by the ensemble average of the single spike trains. For an ensemble of M realizations, the spike rate of the neuron, measured as the number of spikes occurring in a given time interval divided by the length of this interval, is given by the ensemble average of the single spike trains. For discrete time signals the discrete and binary variable (spike/no spike) is considered to be a stochastic process we have already considered to be a stochastic process. The ensemble average is given by equation (2.7).
Classification-based receptive field estimation (CbRF)

Spectro-temporal receptive field (STRF) estimation

STRF estimation from gerbil neurons

stimulation with frequency-modulated tone complexes

Meyer, Diepenbrock, Happel, Ohl, Anemüller (2014)
STRF estimation from gerbil neurons

stimulation with frequency-modulated tone complexes

Unit A
Unit B
Unit C
Unit D

CbRF
Ridge

Meyer, Diepenbrock, Happel, Ohl, Anemüller (2014)
Higher-order stimulus statistics influences STRF estimation results

Stimulus
Frequency
Time
DRC

true STRF
Frequency
Time

neuronal non-linearity

Simulated spike response

estimated STRF

Ripple

Christianson, Sahani, Linden, J. Neurosci., 2008
Higher-order stimulus statistics influences STRF estimation results

Stimulus → true STRF → neuronal non-linearity → Simulated spike response → estimated STRF

Ripple

DRC

Figure 5.

Christianson, Sahani, Linden, J. Neurosci., 2008
Higher-order stimulus statistics influences STRF estimation results

- **Stimulus**: DRC
  - Frequency
  - Time

- **True STRF**
  - Neuronal non-linearity
  - Output intensity vs. Input intensity

- **Simulated spike response**
  - Time

- **Estimated STRF**

- **Stimulus**: Ripple
  - Frequency
  - Time

- **Input**

- **Output**

- **Query**

---

Christianson, Sahani, Linden, J. Neurosci., 2008
Higher-order stimulus statistics influences STRF estimation results.

**Stimulus**
- DRC
- Ripple

**true STRF**
- Neuronal non-linearity

**Simulated spike response**

**estimated STRF**
Classification-based method best matches ground-truth
STRF estimation from gerbil IC neurons

- **Figure 1.** Spectrotemporal Sound Analysis in the Auditory Midbrain

- **Table:**
  - **Unit A**
  - **Unit B**
  - **Unit C**
  - **Unit D**

- **Legend:**
  - **Ridge**
  - **MID**
  - **GLM**
  - **CbRF**
Influence of stimulus ensemble on STRF in gerbil IC recordings

FM tone complexes

Unit C

Unit D

CbRF

Ridge
Influence of stimulus ensemble on STRF in gerbil IC recordings

- **FM tone complexes**
  - Unit C
  - Unit D

- **Dynamic moving ripples**
  - Unit C
  - Unit D

**CbRF**

**Ridge**
Influence of stimulus ensemble on STRF in gerbil IC recordings

FM tone complexes

Dynamic moving ripples

CbRF

Unit C

Unit D

Ridge

Unit C

Unit D
Influence of stimulus ensemble on STRF in gerbil IC recordings

**FM tone complexes**
- Unit C
- Unit D

**Dynamic moving ripples**
- Unit C
- Unit D

**CbRF**
- Ridge

**Influence of stimulus ensemble**

To test the possibility that individual auditory neurons in the ICC central nucleus of the inferior colliculus. Specifying the influence of stimulus ensemble on STRF in gerbil IC recordings.

Although this approach was successfully applied for many neurons, it was also useful to characterize population ripple transfer functions (pCRH). To facilitate comparisons, the pCRH was interpolated using the moving ripple and ripple noise spectrotemporal characteristics of the dynamic and coherently modulated (sound is constructed so that its spectrotemporal envelope is structured.

Time (s)
- Center frequency (kHz)
  - 1.2
  - 1.4
  - 1.6
  - 1.8
  - 2

- 0.5
- 1
- 2
- 4
- 8

- Time before spike (ms)
  - 1
  - 2
  - 4
  - 8
  - 16

Unit A
- Unit B
- FM tone complexes

- Ridge
- CbRF

Although this approach was successfully applied for many neurons, it was also useful to characterize population ripple transfer functions (pCRH). To facilitate comparisons, the pCRH was interpolated using the moving ripple and ripple noise spectrotemporal characteristics of the dynamic and coherently modulated (sound is constructed so that its spectrotemporal envelope is structured.

Time (s)
- Center frequency (kHz)
  - 1.2
  - 1.4
  - 1.6
  - 1.8
  - 2

- 0.5
- 1
- 2
- 4
- 8

- Time before spike (ms)
  - 1
  - 2
  - 4
  - 8
  - 16

Unit A
- Unit B
- FM tone complexes

- Ridge
- CbRF

Although this approach was successfully applied for many neurons, it was also useful to characterize population ripple transfer functions (pCRH). To facilitate comparisons, the pCRH was interpolated using the moving ripple and ripple noise spectrotemporal characteristics of the dynamic and coherently modulated (sound is constructed so that its spectrotemporal envelope is structured.

Time (s)
- Center frequency (kHz)
  - 1.2
  - 1.4
  - 1.6
  - 1.8
  - 2

- 0.5
- 1
- 2
- 4
- 8

- Time before spike (ms)
  - 1
  - 2
  - 4
  - 8
  - 16

Unit A
- Unit B
- FM tone complexes

- Ridge
- CbRF

Although this approach was successfully applied for many neurons, it was also useful to characterize population ripple transfer functions (pCRH). To facilitate comparisons, the pCRH was interpolated using the moving ripple and ripple noise spectrotemporal characteristics of the dynamic and coherently modulated (sound is constructed so that its spectrotemporal envelope is structured.

Time (s)
- Center frequency (kHz)
  - 1.2
  - 1.4
  - 1.6
  - 1.8
  - 2

- 0.5
- 1
- 2
- 4
- 8

- Time before spike (ms)
  - 1
  - 2
  - 4
  - 8
  - 16

Unit A
- Unit B
- FM tone complexes

- Ridge
- CbRF

Although this approach was successfully applied for many neurons, it was also useful to characterize population ripple transfer functions (pCRH). To facilitate comparisons, the pCRH was interpolated using the moving ripple and ripple noise spectrotemporal characteristics of the dynamic and coherently modulated (sound is constructed so that its spectrotemporal envelope is structured.

Time (s)
- Center frequency (kHz)
  - 1.2
  - 1.4
  - 1.6
  - 1.8
  - 2

- 0.5
- 1
- 2
- 4
- 8

- Time before spike (ms)
  - 1
  - 2
  - 4
  - 8
  - 16

Unit A
- Unit B
- FM tone complexes

- Ridge
- CbRF

Although this approach was successfully applied for many neurons, it was also useful to characterize population ripple transfer functions (pCRH). To facilitate comparisons, the pCRH was interpolated using the moving ripple and ripple noise spectrotemporal characteristics of the dynamic and coherently modulated (sound is constructed so that its spectrotemporal envelope is structured.

Time (s)
- Center frequency (kHz)
  - 1.2
  - 1.4
  - 1.6
  - 1.8
  - 2

- 0.5
- 1
- 2
- 4
- 8

- Time before spike (ms)
  - 1
  - 2
  - 4
  - 8
  - 16

Unit A
- Unit B
- FM tone complexes

- Ridge
- CbRF

Although this approach was successfully applied for many neurons, it was also useful to characterize population ripple transfer functions (pCRH). To facilitate comparisons, the pCRH was interpolated using the moving ripple and ripple noise spectrotemporal characteristics of the dynamic and coherently modulated (sound is constructed so that its spectrotemporal envelope is structured.

Time (s)
- Center frequency (kHz)
  - 1.2
  - 1.4
  - 1.6
  - 1.8
  - 2

- 0.5
- 1
- 2
- 4
- 8

- Time before spike (ms)
  - 1
  - 2
  - 4
  - 8
  - 16

Unit A
- Unit B
- FM tone complexes

- Ridge
- CbRF

Although this approach was successfully applied for many neurons, it was also useful to characterize population ripple transfer functions (pCRH). To facilitate comparisons, the pCRH was interpolated using the moving ripple and ripple noise spectrotemporal characteristics of the dynamic and coherently modulated (sound is constructed so that its spectrotemporal envelope is structured.

Time (s)
- Center frequency (kHz)
  - 1.2
  - 1.4
  - 1.6
  - 1.8
  - 2

- 0.5
- 1
- 2
- 4
- 8

- Time before spike (ms)
  - 1
  - 2
  - 4
  - 8
  - 16

Unit A
- Unit B
- FM tone complexes

- Ridge
- CbRF

Although this approach was successfully applied for many neurons, it was also useful to characterize population ripple transfer functions (pCRH). To facilitate comparisons, the pCRH was interpolated using the moving ripple and ripple noise spectrotemporal characteristics of the dynamic and coherently modulated (sound is constructed so that its spectrotemporal envelope is structured.
STRF estimation from gerbil IC neurons: Mutual information analysis

Non-linear methods (CbRF, GLM, MID) show highest information-transfer rates for non-Gaussian stimuli

Meyer, Diepenbrock, Happel, Ohl, Anemüller (2014)
Classification-based method converges quickly and with relatively small estimation variance

Meyer, Diepenbrock, Happel, Ohl, Anemüller (2014)
Matching pursuit (MP) analysis of spectral, temporal and spectra-temporal characteristics

Gabor basis functions as MP-atoms

Spectro-temporal distribution of MP-selected atoms

Data from Gill, Zhang, Woolley, Fremouw, Theunissen (2006) + ridge regression

Bach, Kollmeier, Anemüller (2017)
Summary STRF estimation

STRF estimation algorithms need to go beyond second order statistics (e.g., GLM, MID, CbRF)

Data show that estimation algorithm influences obtained STRF pattern qualitatively

FM-bank stimuli to mimic speech-like t-f-transients

Joint spectra-temporal STRFs appear to be rare, even on FM-bank stimuli
Beyond the linear time-invariant STRF model

Cue selection under dynamic stimulus changes

- Acoustic signal
- Dynamic changes
- Cochlear filterbank
- STRF
- Adaptation
Beyond the linear time-invariant STRF model
Beyond the linear time-invariant STRF model
Beyond the linear time-invariant STRF model
Adaptive STRF estimation model

Step 1
“global”

Estimate STRF for all stimuli with sparse ("zero-mean") prior

Meyer, Diepenbrock, Ohl, Anemüller (2014)
Adaptive STRF estimation model

Step 1
“global”

Estimate STRF for all stimuli with sparse (“zero-mean”) prior

Step 2
“local”

Fit model parameters
Use as prior

Meyer, Diepenbrock, Ohl, Anemüller (2014)
Adaptive STRF estimation model

![Image showing true and estimated averages for different conditions: true, sparse, adaptive, with corresponding time series and numerical values]

- True average
- Sparse average
- Adaptive average

Numerical values for each condition:
- True: 0.70, 0.72, 0.75, 0.73, 0.72, 0.67, 0.70, 0.67, 0.60, 0.71
- Sparse: 0.90, 0.91, 0.94, 0.96, 0.97, 0.97, 0.96, 0.96, 0.93, 0.93
- Adaptive: [values not visible in text]
STRF variability across time: Gerbil inferior colliculus

Meyer, Diepenbrock, Ohl, Anemüller (2014)
STRF variability across time: Gerbil inferior colliculus

Meyer, Diepenbrock, Ohl, Anemüller (2014)
STRF variability across time: Gerbil inferior colliculus

Meyer, Diepenbrock, Ohl, Anemüller (2014)
STRF variability across time: Gerbil auditory cortex

Meyer, Diepenbrock, Ohl, Anemüller (2014)
STRF variability across time: Gerbil auditory cortex

Meyer, Diepenbrock, Ohl, Anemüller (2014)
STRF variability across time: Gerbil auditory cortex

Meyer, Diepenbrock, Ohl, Anemüller (2014)
STRF variability across time: 3 units, moderate to strong fluctuation
STRF variability across time: Summary statistics

Meyer, Diepenbrock, Ohl, Anemüller (2014)
STRF variability across time: Summary statistics

Inferior colliculus

Auditory cortex

Meyer, Diepenbrock, Ohl, Anemüller (2014)
Likelihood evidence of static vs. adaptive STRF model

Meyer, Diepenbrock, Ohl, Anemüller (2014)
Likelihood evidence of static vs. adaptive STRF model

Meyer, Diepenbrock, Ohl, Anemüller (2014)
Cluster analysis:
Do the neurons “revisit” discrete states?

Cluster analysis of two A1 units
Summary dynamic STRF model

- dynamic variability in STRF seems unrelated to spike count
- STRF variability higher in A1 than IC (shown quantitatively)
- quality of STRF variability:
  - parts of STRF change dynamically
  - even spectral BF changes in some cases
- dynamic STRF model supported by higher likelihood than static model

- Origin of fluctuations unclear. Randomly on timescale ~10s? Linked to stimuli?