

Spectro-temporal receptive fields of subthreshold responses in auditory cortex

Christian K. Machens, Michael S. Wehr, Carlos D. Brody,
and Anthony M. Zador

Cold Spring Harbor Laboratory
One Bungtown Rd
Cold Spring Harbor, NY 11724

Abstract

To investigate how cortical neurons represent the acoustic environment, we recorded the activity of single neurons in auditory cortex in response to natural sounds. In a first step, we estimated the linear component—the spectro-temporal receptive field (STRF)—of the transformation from the sound to the neuron’s membrane potential. While the STRF successfully predicts the responses to some of the natural stimuli, it surprisingly fails to predict the responses to others. Hence, much of a neuron’s response, although deterministically related to the stimulus, cannot be predicted by the linear component. In a second step, we now investigate to what extent the STRF can be amended by including the time-varying dynamics of different adaptation mechanisms.

Summary

While it is widely agreed that the primary visual cortex decomposes images into components such as oriented edges, the corresponding decomposition of acoustic sounds in the primary auditory cortex remains somewhat obscure. In correspondence to the successful linear description of early visual neurons, the linear spectro-temporal receptive field (STRF) has been proposed as a candidate framework for characterizing the function of auditory cortical neurons. In this model, neural responses are approximated as linear functions of the sound spectrogram. The STRF has been estimated using a variety of stimulus ensembles, including dynamic ripples and random chord stimuli (Kowalski et al., 1996a; deCharms et al., 1998; Sahani and Linden, 2003). Yet in their natural environment, animals encounter highly complex, dynamically changing stimuli, and the auditory cortex evolved to process such complex sounds. Hence, it remains an open question to what extent the STRF is capable of explaining the functioning of the auditory cortex in its natural mode of operation.

Here we present estimates of the STRF using *in vivo* whole cell recording in rat auditory cortex. Because whole cell recording measures the total synaptic input to a neuron, rather than just its output—a sparse binary spike train—as in conventional single unit physiological recordings, this technique provides a much richer source of information about the neuron’s response. Not only do these whole cell recordings avoid

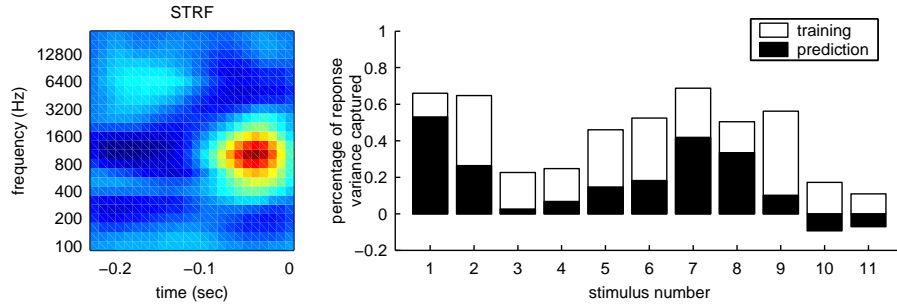


Figure 1: Left: Spectro-temporal receptive field for an example neuron. Right: Percentage of response variance captured by the model for different stimuli. White bars show the predicted response variance as estimated on the data used to train the STRF; black bars show the predicted response variance when estimated on new data.

the problem of sparse spiking responses, they also provide insight into the computations performed by the input to the neuron’s spike generating mechanism. By blocking action potentials pharmacologically, the recordings reveal the subthreshold membrane potential—the total synaptic input to the cell, prior to thresholding by the spike generating mechanism. Neurons typically featured strong subthreshold membrane fluctuations consisting of both stimulus-locked as well as spontaneous (*i.e.* non-stimulus-locked) activity.

In a first step, we sought to estimate the STRF from responses that were averaged over several trials. While STRFs have often been estimated using the reverse-correlation method, the estimation is here complicated due to the fact that the natural stimuli do not sample all directions in stimulus space equally well. To avoid overfitting along the undersampled directions, we used a procedure that incorporated power and smoothness constraints on the STRF parameters. Similar approaches have been used before (Theunissen et al., 2001; Sahani and Linden, 2003).

Typically, the estimated STRFs feature inhibitory and excitatory fields with varying degrees of complexity. An example STRF is shown in the Figure (left panel), displaying an inhibitory field (blue) followed by an excitatory field (red). This STRF predicts strong responses to stimulus onsets within a specific frequency range. Although computed for subthreshold responses, the shapes of the STRFs found approximately agree with those reported previously for spiking neurons in similar preparations. Indeed, to some extent, the estimate of the time-varying firing rate used in these studies might be thought of as a reconstruction of the subthreshold membrane potential. The STRFs are also in accordance with the neuron’s frequency-sensitivity as measured with short sine tones.

To assess the predictive power of the STRF, we tested it on a set of data that was not used in its estimation. To compare all the natural stimuli tested, we quantified the relative prediction success as the percentage of the response variance captured by the STRF. As indicated in the Figure (right panel, black bars), the ability to predict the responses is strongly stimulus-dependent: While the responses to some stimuli (such as

no.3 and 4) are usually partially predicted, the responses to others (such as 10 and 11) cannot at all be predicted. However, the failure to predict the responses of some stimuli cannot be attributed to the absence of stimulus-locked activity; all stimuli feature a high proportion of stimulus-locked activity to noise. Furthermore, the differences in the performance of the STRF cannot be due to the particular trade-offs between the training and prediction sets; even the relative success on the training sets (Figure, right panel, white bars) differs significantly between stimuli. Similar results were found in all other cells recorded.

The widespread failure of the linear model to predict responses for many but not all complex stimuli indicates a high degree of nonlinearity for some stimuli coupled with a lesser degree for others. Since the inclusion of static nonlinearities in the model did not significantly enhance the prediction success, we conclude that the non-linearities encountered must be dynamic. Indeed, there are many situations where the timing of a post-synaptic potential is correctly predicted while its size is strongly underestimated by the linear model. These cases suggest that adaptation (for instance in the form of gain control) dominates the response properties. Moreover, in many cases the model completely fails to predict the occurrence of a post-synaptic potential. These cases might be attributed to dynamic interactions between ON- and OFF-responses, depending on their spectro-temporal relations.

Presently, we seek to quantify these non-linearities by including dynamic components in the STRF model. In particular, we use simple differential equations to model adaptive gain control and intensity adaptation acting on the output of the STRF model. By fitting different candidate models to the natural stimulus data, we ultimately seek to nail down the components that are missing in the linear prediction.

References:

- deCharms RC, Blake DT, Merzenich MM (1998). Optimizing sound features for cortical neurons. *Science* 280:1439–1443.
- Kowalski N, Depireux DA, Shamma SA (1996). Analysis of dynamic spectra in ferret primary auditory cortex. *J Neurophysiol* 76(5):3503–3523.
- Machens CK, Wehr MS, Zador, AM (2003). Spectro-temporal receptive fields of subthreshold responses in auditory cortex. *Neural Information Processing Systems* 15, in press.
- Sahani M, Linden JF (2003). How linear are auditory responses? *Neural Information Processing Systems* 15, in press.
- Theunissen FE, David SV, Singh NC, Hsu A, Vinje WE, Gallant JL (2001). Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. *Network: Comput Neural Syst* 12:289-316