The linearity of emergent spectro-temporal receptive fields in a model of auditory cortex

M. Coath *, E. Balaguer-Ballester, S.L. Denham, M. Denham

Centre for Theoretical and Computational Neuroscience, University of Plymouth, Plymouth PL4 8AA, UK

ABSTRACT

The responses of cortical neurons are often characterized by measuring their spectro-temporal receptive fields (STRFs). The STRF of a cell can be thought of as a representation of its stimulus 'preference' but it is also a filter or 'kernel' that represents the best linear prediction of the response of that cell to any stimulus. A range of *in vivo* STRFs with varying properties have been reported in various species, although none in humans. Using a computational model it has been shown that responses of ensembles of artificial STRFs, derived from limited sets of formative stimuli, preserve information about utterance class and prosody as well as the identity and sex of the speaker in a model speech classification system. In this work we help to put this idea on a biologically plausible footing by developing a simple model thalamo-cortical system built of conductance based neurons and synapses some of which exhibit spike-time-dependent plasticity. We show that the neurons in such a model when exposed to formative stimuli develop STRFs with varying temporal properties exhibiting a range of heterotopic integration. These model neurons also, in common with neurons measured *in vivo*, exhibit a wide range of non-linearities; this deviation from linearity can be exposed by characterizing the difference between the measured response of each neuron to a stimulus, and the response predicted by the STRF estimated for that neuron. The proposed model, with its simple architecture, learning rule, and modest number of neurons (< 1000), is suitable for implementation in neuromorphic analogue VLSI hardware and hence could form the basis of a developmental, real time, neuromorphic sound classification system.

© 2008 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Since Hubel and Wiesel (1962) showed that, for neurons in visual cortex, there were '*preferred stimuli*' which evoked a more vigorous response than all other stimuli, it has become commonplace to think of discrete neural units as having stimulus preferences. The quantification of this idea through the use of reverse, or triggered correlation (de Boer and Kuyper, 1968) has led to the concept of the *spatio*-temporal receptive field or spectro-temporal receptive fields (STRF) in visual neuroscience. For auditory stimuli the principal is similar, but the representational dimensions are time and frequency. In the auditory system the term *spectro*-temporal receptive field, also referred to as STRF, has been adopted. The STRF is the linear filter, or kernel, that best explains or predicts the response of a cell to any given stimulus. As the STRF is constructed on an assumption of linearity, the extent to which it predicts the cells response can be interpreted as a measure of how linear the cell is (Machens et al., 2004). Cells with a range of non-linearities have been reported *in*

vivo, e.g. Theunissen et al. (2000), Machens et al. (2004), and it has been suggested that these non-linearities are important in interpreting the highly selective response of some neurons to specific natural stimuli.

Although it is widely believed that auditory perception is based on the responses of cortical neurons that are tuned to spectro-temporal 'features' it is not clear how these features might come in to existence. There is evidence that cortical responses develop to reflect the nature of stimuli in the early post-natal period (Illing, 2004; Zhang et al., 2001, 2002) and that this plasticity persists beyond early development (Wang, 2004). It has been suggested (Coath and Denham, 2005; Coath et al., 2005) that the spectro-temporal patterns found in a limited number of stimuli, which reflect some putative early auditory environment, may bootstrap the formation of these responses. Here we extend this idea to demonstrate that unsupervised, correlation based learning, implemented using conductance based synapses exhibiting *spike timing dependent plasticity* (STDP), leads to responses similar to those reported from measurements *in vivo* by characterizing the STRFs in our model of auditory cortex. Also we show that the resulting STRFs depend on the stimuli chosen to represent the formative environment.

* Corresponding author.
E-mail address: mcoath@plymouth.ac.uk (M. Coath).

If the response of a neuron is linear then convolution of the STRF with the representation of the stimulus used to estimate the STRF would exactly reproduce the observed response to that stimulus. Not only would it reproduce this response but, using the same method, it would be possible to predict responses to other, novel stimuli. It has been shown, e.g. Machens et al. (2004), that some cortical responses, or at least their gross features, can be predicted in this way. However many cortical responses are highly non-linear and the estimated STRFs fail not only to predict cortical responses to novel stimuli, but fail to reproduce the responses used in their estimation. In this work we investigate the linearity of the responses of the neurons in the thalamo-cortical model by calculating the linear correlation between the response of the model cells and that predicted by the STRF.

2. Methods

2.1. Sub-thalamic processing

The auditory system performs a spectral decomposition which can be modelled by a finite number of band pass filters (Patterson et al., 1992), and many cells in the auditory periphery of many different types exhibit well defined characteristic frequencies (Trussel, 2002). In addition there is a great deal of evidence that the auditory pathway is arranged tonotopically and little evidence to support integration across frequency channels in sub-cortical areas (Trussel, 2002).

It is also well documented that the auditory system is sensitive to the temporal structure of the amplitude envelope, particularly rising, or *onset* transients. This has been shown both in physiological and psychophysical measurements (e.g. Phillips et al., 2002; Heil, 1997). This sensitivity increases as measurements are made at successively higher levels in the auditory pathway. Units that detect onsets are found throughout the auditory system: in the cochlear nucleus (Frisina et al., 1985; Rhode and Greenberg, 1994), inferior colliculus (Langner and Schreiner, 1988), thalamus (Rouiller et al., 1981; Rouiller and de Ribaupierre, 1982), and cortex (Eggermont, 2002). It has been suggested that the emphasis given to transients in neural representation may reduce correlations in the stimulus representation and have a role in figure-ground separation (Coath and Denham, 2007). In addition physiological measurements suggest that information in different parts of the tonotopically arranged auditory system is extracted on different, frequency dependent time scales (Krumbholz et al., 2003). In the light of this evidence our model consists of three pre-processing stages.

2.1.1. Cochlear model

The first stage approximates processing in the cochlea. Sounds are processed using a bank of 30 Gammatone filters (Slaney, 1994), with centre frequencies ranging from 100 to about 8000 Hz arranged evenly on an equivalent rectangular bandwidth scale (Glasberg and Moore, 1990). The output in each frequency channel is half-wave rectified as a simple model of inner hair cell function.

2.1.2. Transient enhancement

The next stage of processing involves enhancement of the transients in the signal derived from each cochlear filter. The model of transient responses used here is based on the skewness of the distribution of energy in a time window (Coath and Denham, 2005, 2007; Coath et al., 2005). Short-term skewness is a sensitive indicator of rising and falling energy and has a value near zero when the energy is approximately unchanging. Although this representation identifies both onsets and offsets (positive and negative skewness), these two responses may be characteristic of separate cell populations. In the experiments described here only the onsets, or regions of rising energy, have been used.

2.1.3. Spike generation

The last stage of the sub-thalamic processing involves the generation of spike trains suitable as inputs for the conductance based synapses used in the model. Sequences of spikes are generated where the inter-spike intervals are chosen from a gamma function probability density (Dayan and Abbot, 2001) which reduces the probability of short inter-spike intervals and models refractory effects. The resulting spike train is converted to a time dependent rate using the estimated firing rate, which is taken to be the analogue output of the transient sensitive processing, by a process of spike 'thinning' (Dayan and Abbot, 2001).

The whole of the subcortical processing can be succinctly summarized as in Fig. 1.

2.2. The network

2.2.1. Neurons

The neurons used are *adaptive exponential integrate-and-fire* (aEIF) units (Brette and Gerstner, 2005), this approach substitutes the strict voltage threshold by a more
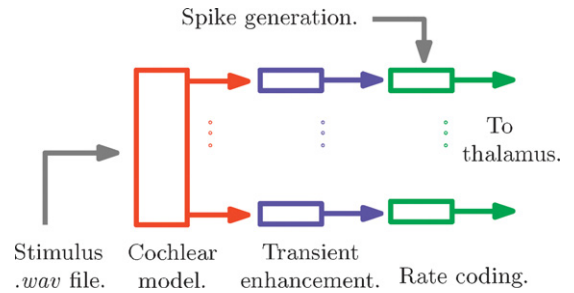
Fig. 1. Summary of the sub-thalamic processing used in the model. The sound file is first decomposed in to frequency channels using a gammatone filter bank, then regions of rising energy within each channel are identified. Spikes trains are then generated which are thinned to rate code the amplitude of each channel. See Section 2.1 for details.

realistic, smooth spike initiation zone (Fourcaud-Trocm and Brunel, 2005). It also includes a variable which allows modelling of subthreshold resonances or adaptation (Richardson et al., 2003). Most importantly it uses a stimulation paradigm not of current injection, but of conductance injection which moves integrate-and-fire models closer to a situation that cortical neurons would experience *in vivo* (Destexhe et al., 2003). This modification also allows the use of conductance based synapses as described below.

2.2.2. Synapses

There are four types of synapse present in the model. The first three representing AAMPA, GABA, and NMDA synapses are modelled by an exponential rise and fall of conductance as a function of the pre-synaptic spike time (Eq. (1)). The fourth synapse type, which is based on an AAMPA synapse that exhibits synaptic depression (referred to as dAAMPA here) is described in Section 2.2.3 below. The time course of the conductance g for AAMPA, GABA, and NMDA synapses as a function of pre-synaptic spike time t_f is given in Eq. (1).

$$g(t) = \sum_f \mathcal{W} \bar{g} \mathcal{N} [e^{-(t-t_f)/\tau_d} e^{-(t-t_f)/\tau_r}] \Theta(t - t_f) \tag{1}$$

In Eq. (1) τ_r and τ_d are the time constants for the rise and decay of the synaptic conductance, $\bar{g}$ is the maximum conductance, and $\mathcal{N}$ is a normalizing factor. $\mathcal{W}$ is the synaptic weight which is adjusted by the learning rule (see Section 2.3). The values used are collated in Table 1 (Gerstner and Kistler, 2002).

2.2.3. Depressing synapses

The dynamical properties of cortical synapses can influence the temporal sensitivity of cortical circuitry (Tsodyks and Markram, 1997). Synaptic responses are context dependent, and may develop depression or facilitation, depending on the cells involved. The model of synaptic depression (Denham, 2001) adopted is based on a system containing three component subsystems (Ostergaard, 1990; Tsodyks and Markram, 1997)(Fig. 2) representing:
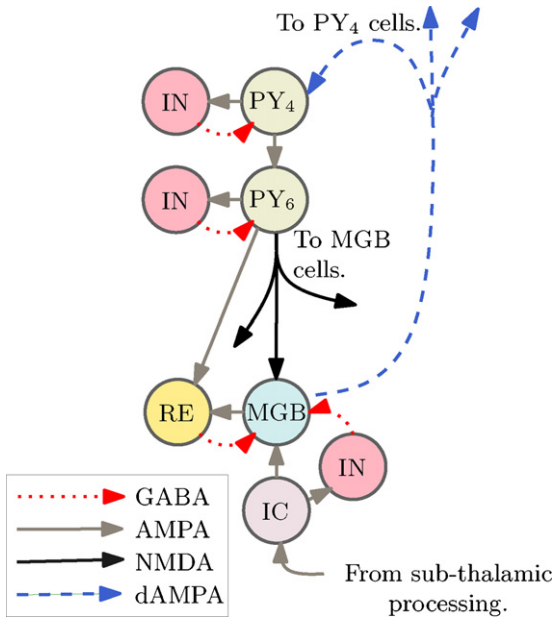
- [-]the pool of available transmitter y,

Table 1
Time constants and conductances used in synapse models described in Eq. (1)

	τ_r (ms)	τ_d (ms)	$\bar{g}$ (S)
AAMPA	0.09E−3	1.50E−3	720.0E−12
GABA	0.01E−3	5.00E−3	40.0E−12
NMDA	3.00E−3	40.00E−3	1.2E−09
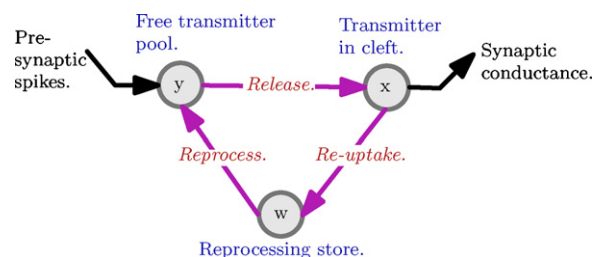
Fig. 2. Representation of the model of synaptic depression showing the three-centred approach described in Eq. (2)

Fig. 3. Each 'column' of the network consists of eight neurons divided in to two sections. The sub-cortical section receives input from one tonotopic channel of the sub-thalamic processing. Each thalamic (MGB) cell is connected to a number of cortical cells representing layer IV, the principal receiving layer. Layer VI cells recurrently connect the cortex to the thalamus VIA NMDA synapses which exhibit STDP which is the origin of the correlation based learning in the network.

- [-]the transmitter released in to the cleft x, and
- [-]the store of transmitter waiting to be reprocessed w.

Transitions between these states are controlled by time constants α and β, by the synaptic efficacy ϵ, and by the function $f[y(t)]$ which is stochastic in that it incorporates a random variation in the amount of transmitter released. $I(t)$ has values unity or zero indicating the presence or absence of a pre-synaptic spike at time t

$$
\begin{aligned}
z(t) &= \epsilon \cdot I(t) \cdot f[y(t)] \\
\frac{dx}{dt} &= z(t) - \alpha \cdot x(t) \\
\frac{dy}{dt} &= \beta \cdot w(t) - z(t) \\
\frac{dw}{dt} &= \alpha \cdot x(t) - \beta \cdot w(t)
\end{aligned}
\tag{2}
$$

The synaptic parameters for Eq. (2) were adjusted so as to replicate rise and fall times of conductances and paired pulse ratios reported in *in vivo* studies of pyramidal neurons (Atzori et al., 2001).

2.2.4. Network architecture

The model auditory cortex consists of one hundred 'columns' eachconsisting of eight aEIF units as illustrated in Fig. 3. The lower, sub-cortical, section represents the junction of the inferior-colliculus (IC) with the medial geniculate body of the thalamus (MGB). The upper section represents a two layer cortical structure consisting of a receiving layer of pyramidal (PY) cells (layers iii and iv (Winer et al., 2005) marked simply as PY_4 in Fig. 3) and a second layer (layer VI marked as PY_6 in the figure) providing a recurrent excitatory connection to the thalamus (Guillery and Sherman, 2002), and recurrent inhibitory connection to the thalamus via the thalamic reticular nucleus (RE) (Guillery and Sherman, 2002). The excitatory afferents from the thalamus to the receiving layer of the cortical part of the model come from a number of MGB cells. These are selected based on a Gaussian distribution of probabilities. The probability distribution is arranged such that there is an 80% chance a PY_4 cell receiving excitation from an IC cell within its own column. This falls as the inter-column distance d increases and is close to zero at $d = 10$, see Fig. 4. Connections at $d > 10$ are rejected. These connections give the opportunity for cortical neurons to integrate information from heterotopic areas of thalamus. They also stand as surrogates for the cortico-cortical connections (Thomson and Bannister, 2003) which have no explicit representation in this model.

Excitatory inputs to the thalamus come from two sources: first, the IC, that is the output of the sub-thalamic processing stage of the model, and second, recurrent connections from layer VI. The cortico-thalamic connections are mediated via NMDA type synapses which are the loci of the STDP and hence the correlation based learning in the network, see Section 2.3. These connections come from a number of PY_6 cells selected from columns based on a Gaussian distribution of probabili-
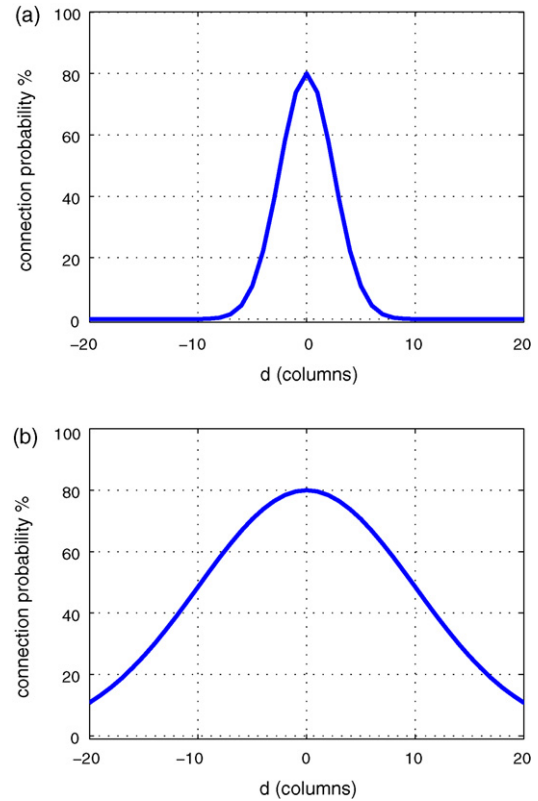

Fig. 4. Probability distributions for connections between thalamic and cortical sub-sections of the network.

ties. The probability distribution is arranged such that there is an 80% chance of a thalamic cell receiving excitation from a PY_6 cell within its own column falling as the inter-column distance (d) increases to $\approx 10\%$ at $d = 20$ see Fig. 4. Connections at $d > 20$ are rejected. Inhibitory inputs to the thalamus also come from two sources: first, the IC, in this case via a GABA type interneuron (although there is evidence for direct connections from GABAergic cells in IC (Winer et al., 1996; Marie et al., 1997), and second, from recurrent connection with the thalamic reticular nucleus (Guillery et al., 1998; Guillery and Sherman, 2002).

2.3. Synaptic plasticity

STDP is the modification of synaptic weights based on the correlation between pre- and post-synaptic firing times. Evidence for this has been gathered *in vitro*, and is beginning to emerge *in vivo* (Jacob et al., 2007), and is believed to be a feature of synapses which have NMDA receptors which regulate genes required for long term maintenance of these changes (Rao and Finkbeiner, 2007). The degree of synaptic modification ($\mathcal{M}$ as a percentage, potentiation or depression) is typically a function of the number of pre- and post-synaptic action potentials, and a function of the time
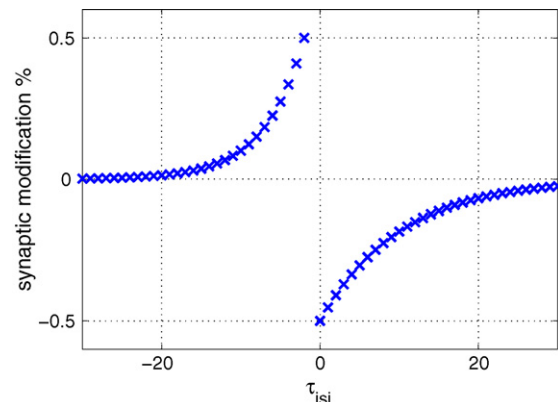
Fig. 5. The window function used in STDP learning.

interval between them, the *inter-spike interval* τ_{isi}. The function of τ_{isi} which yields $\mathcal{M}$ is called the STDP window function (Bi and Rubin, 2005), see Fig. 5.

In general, if a pre-synaptic spike precedes a post-synaptic spike then the synapse is potentiated; if the timing of the spikes is reversed then the synapse is depressed. However, if the firing rate is relatively low then it is not necessary to look at individual spike pairs. The synaptic modification can be calculated by multiplying the cross-correlation function of the pre- and post-synaptic spike trains by the window function chosen and integrating (or summing in the discrete case) the result (Drew and Abbott, 2006). It is this approach that is adopted here for reasons of computational efficiency.

One problem with correlation based learning is that the weight changes are unstable and additional mechanisms have to be invoked to ensure that weights do not increase in an uncontrolled manner. Our approach in this initial work was to start with very low weights and keep the training short. In this way we see how the pattern of weight changes establishes itself in the early stages of training. Another possibility is to implement a form of homeostatic normalization, this is being investigated as part of the next stage in the development of the model.

The window function used in this work is shown in Eq. (3) and illustrated in Fig. 5. The synaptic modification ($\mathcal{M}$) decays exponentially away from the maximum at $\tau_{isi} = -2$ where $\mathcal{M} = 0.5$. This function is not continuous because the cross-correlation of the pre- and post-synaptic spike trains is calculated with the spike times in 1 ms bins.

$$
\begin{aligned}
\tau_{isi} \in] -\infty, -2] &: \mathcal{M} = \exp\left(\frac{\tau_{isi} + 2}{5}\right) \cdot 0.5 \\
\tau_{isi} \in [-1, \infty[&: \mathcal{M} = -\exp\left(\frac{-\tau_{isi}}{10}\right) \cdot 0.5
\end{aligned}
\tag{3}
$$

2.4. Training regime

The network was trained three times, once with each of three stimulus sets. Each set consisted of five stimuli all belonging to the same class. The three classes were
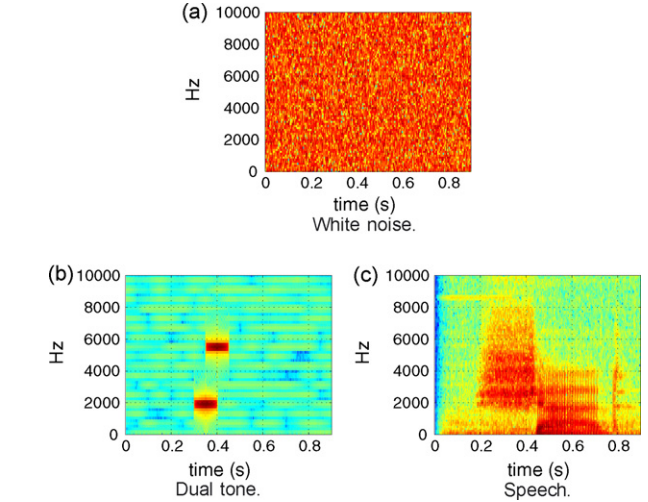
Fig. 6. Examples from each of the three classes of stimuli used in training. Each class consisted of five similar stimuli with each stimulus being presented five times in a random order.

- [-]White noise: five samples of white noise in one second bursts,
- [-]Dual tone: five examples of two overlapping tones each of 100 ms duration at 2.0 and 5.5 kHz preceded by varying amounts of silence with the whole stimulus padded to approximately one second,
- [-]Speech: five speech stimuli ('shad', 'mad', 'dad', 'lad', 'tad') preceded by varying amounts of silence and padded as in the previous stimulus set.
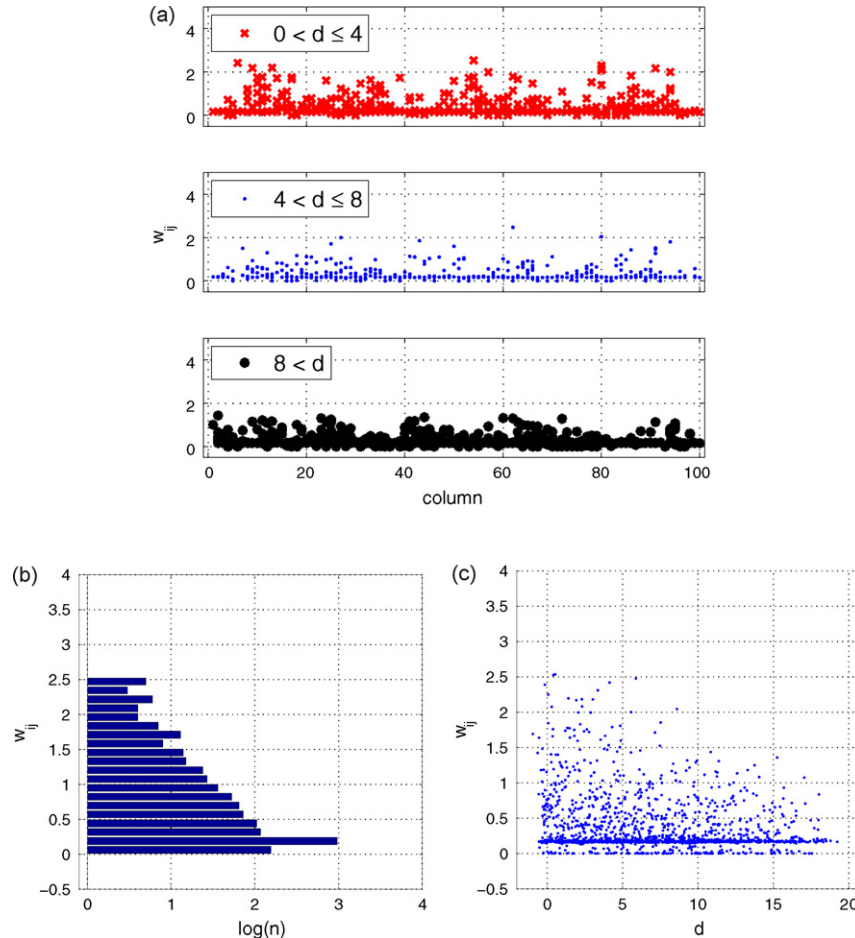
Fig. 7. Results for white noise training. (a) Column number (i.e. position on the tonotopic axis) plotted against final synaptic weight (w_{ij}) after training. The upper, middle, and lower axes indicate short, medium, and long range connections respectively. (b) Synaptic weights divided in to 20 bins indicating $\log_{10}$ of the number of connections n in each weight bin. (c) The number of columns traversed by the NMDA cortico-thalamic connection d against the weight of this connection.

Examples from each of these stimulus sets are illustrated as short term Fourier transform spectrograms in Fig. 6. Each stimulus in the set was presented 25 times giving a total of 125 stimulus presentations per training cycle with the order of stimuli randomized. After each presentation the cross-correlation of the pre- and post-synaptic spike trains for each of the NMDA type synapses was calculated (in 1 ms time bins) and the value at each lag multiplied by the weighting function to determine the synaptic modification (see Section 2.3). The synaptic weights were then modified before the next presentation.

2.5. STRF estimation

A popular method for characterizing the responses of cortical and sub-cortical cells is the spectro-temporal receptive field or STRF, e.g. Kowalski et al. (1996a,b), deCharms et al. (1998), Depireux et al. (2001), Theunissen et al. (2001), Miller et al. (2002), Elhilali et al. (2004). Some time ago it was shown how reverse correlation in response to white noise could be used to characterize STRFs (Aertsen and Johannesma, 1981).The STRFs in this report have been derived using a MATLAB toolbox called STRFPak which is developed by the Theunissen and Gallant laboratories at UC Berkeley (Zhang and Gill, 2006). This software incorporates methods to remove stimulus correlations from STRF estimation (Theunissen et al., 2001; Miller et al., 2002) allowing a broader range of stimuli (including natural stimuli) to be used for this purpose. Although there are no constraints on the stimuli that can be used, a large number of spikes are needed for the noise in the estimate to be reduced. Given the dynamics of the synaptic depression (see Section 2) some types of stimuli will evoke very few spikes and necessitate the processing of unmanageably large stimulus files. For the results in this report we have used stimuli consisting of sequences of tone combinations, or 'random chords', that are based on the idea of an auditory checkerboard (deCharms et al., 1998). These consist of a continuous sequence of up to six 20 ms tone bursts from a randomly selected range of frequencies presented simultaneously.

The spike trains used in STRF estimation were obtained after training and without further adjustment of the synaptic weights. The random chord stimuli were presented to each of the three trained networks for a total of 300 s and the time of each spike in each of the 100 PY_6 neurons recorded, layer VI being regarded as

the 'output' layer for the purposes of this work. The STRFs illustrated in the results section are derived from these spike trains.

2.6. Linearity

The STRF is a linear kernel. To illustrate the degree to which this linear approximation captures the behaviour of the model we calculate the correlation coefficient between the predicted response (obtained by convolving the STRF with the spectro-temporal representation of the stimulus used in its estimation) and the actual response of the model. To do this we estimate the instantaneous firing rate of the PY_6 cells in the model from the output spike times using a weighted sum over a 20 ms Gaussian window. The network weights used were those resulting from the training with speech stimuli. Of the 100 PY_6 cells in the model 50 exhibited less than 1000 spikes during the estimation period and these cells were excluded from the analysis. The correlation coefficient does not take into account that the STRF response and the experimental responses might exhibit a linear relationship but with one or more delays between them. A general method to address this problem is the well-known *linear autoregressive models using exogenous inputs* (ARX) in signal processing theory (Ljung, 1986). We investigated this potential time-delayed linear relationship between each STRF prediction and the corresponding measured model neuron response by finding the linear FIR filter (i.e. an ARX model having no autoregressive terms) of the prediction that most closely resembles the measured response. To assess the degree to which these time delayed terms might be important we also calculate the correlation coefficient between this filtered STRF response and the measured response for each neuron.

3. Results

3.1. Synaptic weights

The final synaptic weights after training using each of the stimulus sets described in Section 2.4 can be seen in Figs. 7–9. All weights
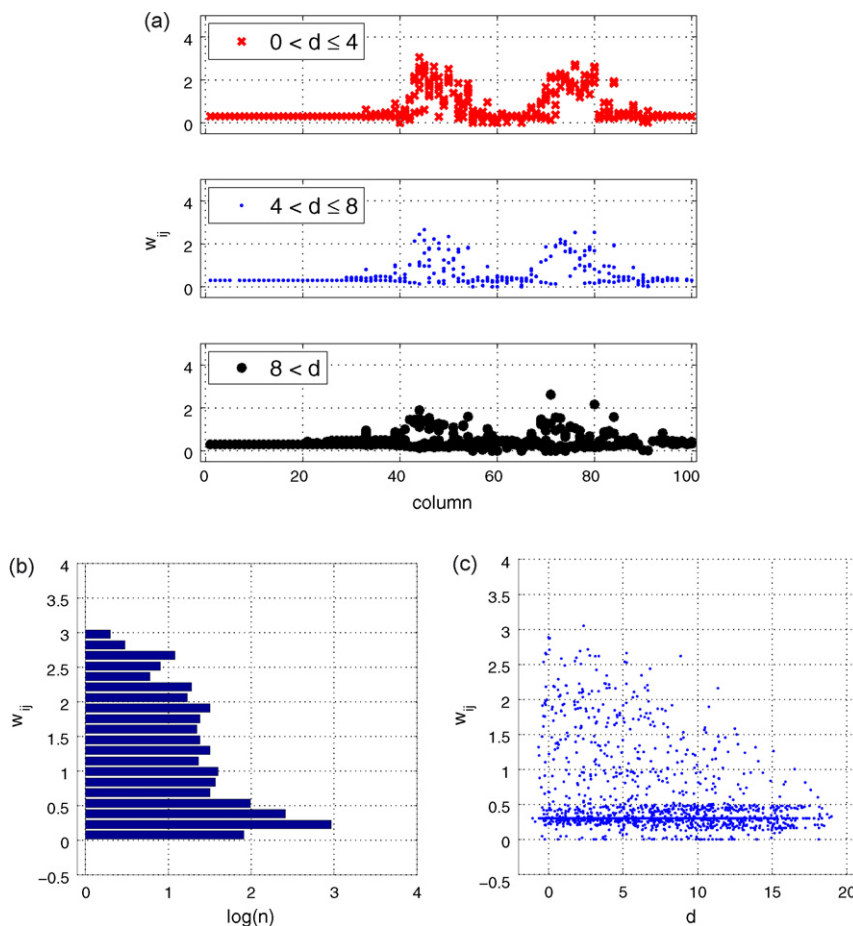

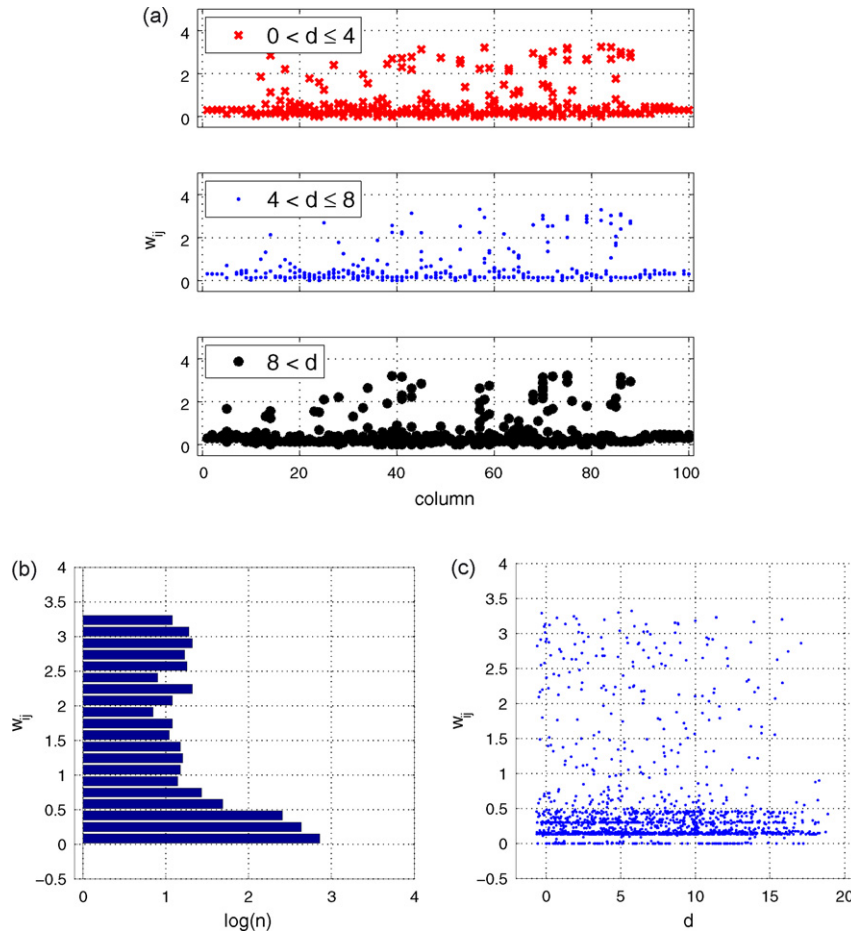**Fig. 8.** Results for dual tone training. See Fig. 7 for explanation.

Fig. 9. Results for speech training. See Fig. 7 for explanation.

in NMDA synapses were initialized to 0.3 and the results in Figs. 7–9 show the weight distribution after training.

The first sub-figure of each set of three shows the column number (the tonotopic axis) on the abscissa against the weight of the synaptic connection (w_{ij}) for all NMDA connections originating at that column. The distance to the destination column is also indicated: short-range ($0 < d \leq 4$) connections marked with red crosses on the upper axis, medium-range connections ($4 < d \leq 8$) marked with blue dots on the centre axis and finally, long-range connections ($8 < d$) marked with black plus signs on the lower axis.

The second sub-figure shows $\log_{10}$ of the total number of connections (n) having a particular weight, with the weight values divided in to 20 bins.

The third sub-figure plots the distance d in columns, traversed by the NMDA connection on the abscissa against the final weight after training.

The white noise stimulus produces a pattern of high weights in short range connections across most of the tonotopic axis. A large number of weights remain approximately unchanged over the course of the training implying un-correlated firing in the network.

The dual tone stimulus produces high weights in the regions of the tonotopic axis corresponding to energy in the stimulus but also produces a greater number of medium-range connections. This reflects the temporal correlations between frequency channels inherent in the stimulus.

It is clear that the final weights for the speech stimuli, with their richer spectro-temporal content, produce a connection pattern with a greater number of long-range, high-strength connections than the simpler dual-tone and white noise stimulus classes. This is accompanied by a greater proportion of synaptic strengths that have been forced to (or very near to) zero.

3.2. STRFs

The net effect of a range of homotopic and heterotopic projections of varying weights from across the tonotopic axis is that the responses of the model cortical neurons will exhibit responses that can be interpreted as STRFs in a way which is comparable to measurements made *in vivo*. These were calculated using STRFPak (Zhang and Gill, 2006) for all pyramidal neurons in layer VI. As each of the training regimes was initiated from precisely the same starting weights and connectivity it is interesting to compare the final response of the same neuron after training with each of the three stimulus classes. This can be seen in two representative cases in Figs. 10 and 11. In the first case (column 55) the final responses are similar for all three training sets. It is clear, however, that there is a minimum in the response at ≈4 kHz which is above the best frequency of ≈2 kHz. This implies the neuron will '*respond preferentially to narrow frequency bands or to constant-frequency edges that correspond to the excitatory region*' (King and Schnupp, 1998). In the second case (column 72) in contrast, the three final states of the network show (a) integration across a broad range of frequencies in temporally imprecise maxima for the noise trained network, (b) two well defined temporally precise maxima in low-to-high order for the dual-tone stimuli, and (c) a single maximum for the speech stimuli.
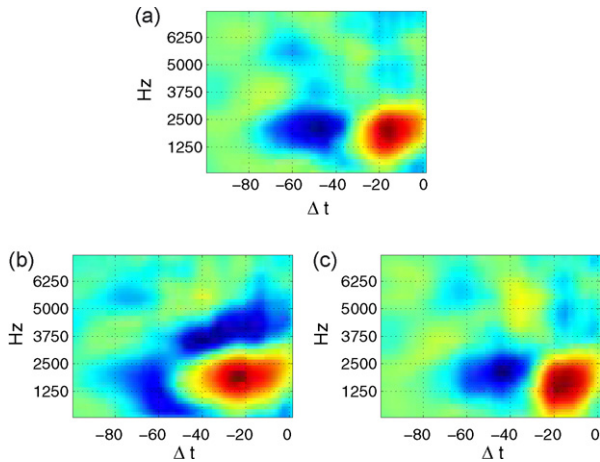
Fig. 10. STRFs for $PY_6 55$ showing simple geometry from each of the three stimulus classes. (a) White noise, (b) dual tone, and (c) speech.
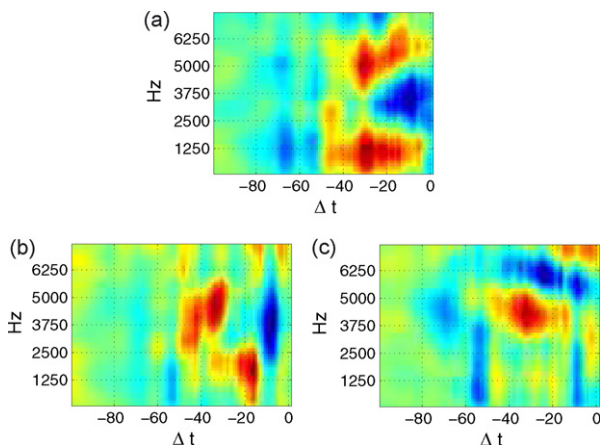
Fig. 11. STRFs for $PY_6 72$ showing more complicated heterotopic integration which is markedly different for each of the three stimulus classes. (a) White noise, (b) dual tone, and (c) speech.

3.3. Linearity

Fig. 12 shows the values of the correlation coefficient between the response of each of the model cells and the prediction by the STRF for that cell—these values have been sorted and arranged in ascending order of correlation coefficient on the abscissa. The upper solid line illustrates the values from the ARX filtered prediction and hence represents, for each neuron, an upper bound on the value of the correlation when linear combinations with time delays are
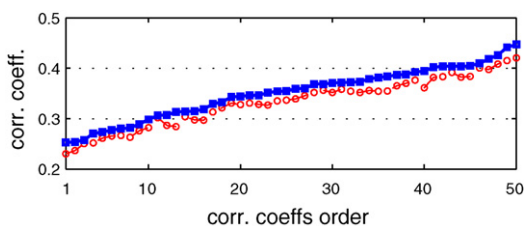
Fig. 12. The correlation coefficients between the model response and (lower broken line) the STRF prediction and (upper solid line) the ARX filtered STRF prediction. These represent a direct measure of the linearity of each of the PY_6 cells. The cells are arranged in increasing order of correlation coefficient along the abscissa. They exhibit a wide range of values from ≈ 0.22 to ≈ 0.42 where a significant change would be 0.03 (95% confidence).

included. The values for the un-filtered STRF prediction (lower broken line) are very close to these values indicating that the majority of the predictive capacity can be accounted for without taking in to account delays. The neurons exhibit a highly significant range of linearities, as do *in vivo* cortical responses, e.g. Machens et al. (2004), however it is difficult to compare these results directly as physiological measurements are frequently made on the basis of membrane potentials, with action potentials blocked pharmacologically, rather than using spike rates.

4. Discussion

The simple nature of the neuron and synapse models used here, and the implementation of the STDP learning rule, make this approach eminently suitable for implementation in neuromorphic analogue VLSI hardware (Indiveri et al., 2006). For the hardware implementation the learning rule would be modified to one that is based on single pairs of pre- and post-synaptic action potentials localized in time as well as space. In this way the model could form the basis of a developmental, real time, neuromorphic sound classification system. It has previously been shown that ensembles of STRF-like kernels derived from formative stimuli respond to stimuli in such a way as to preserve information about what the stimulus is, who is saying it, and in what manner it is being said (Coath and Denham, 2005; Coath et al., 2005). Useful kernels exhibited a range of response types and were of intermediate temporal extent, these ideas are related to interesting work in the field of vision (Ullman et al., 2002).

The work presented here goes some way towards putting these ideas on a biologically plausible footing by showing that correlation based learning in a model thalamo-cortical system can lead to an ensemble of responses whose STRFs also exhibit a wide range of simple and integrative responses that are, in some way, related to properties of the formative stimuli. The model cortex, with its range of spectro-temporal preferences exhibited by the PY_6 neurons, can be seen as an 'ensemble of feature extractors' (Coath and Denham, 2005). However, in this case the formation of the ensemble is stimulus driven, and by mechanisms that could be implemented in the neural substrate.

The weights resulting from training the network with speech result in responses for each of the model cortical cells that exhibit a wide range of linearities. This is revealed by the fact that there are a large range of values for the correlation coefficient between the response of each model neuron and the prediction on the basis of its STRF. This is despite the fact that the response being predicted is precisely the one that was used to estimate the STRF. This is consistent with the situation revealed by physiological measurements although, as mentioned in Section 3.3 direct comparisons are difficult. The origins of these non-linearities could be in the long-range, high-strength connections that are evident in the network trained using more complex stimuli (see Fig. 9(c)); work is now underway to investigate this. These non-linearities are important in interpreting results that show that while STRFs appear to predict responses to certain classes of stimuli well, e.g. Klein et al. (2000), it is clear that neurons can exhibit a high degree of selectivity in response to certain natural stimuli (Theunissen et al., 2000; Machens et al., 2004). Further work is now underway which, it is hoped, will lead to results which will allow a fuller account of the nature and effects of non-linearities in these model cortical responses.

References

Aertsen, A., Johannesma, P., 1981. A comparison of the spectro-temporal sensitivityof auditory neurons to tonal and natural stimuli. Biol. Cybern. 42 (2), 145–156.

Atzori, M., Lei, S., Evans, D.I., Kanold, P.O., Phillips-Tansey, E., McIntyre, O., McBain, C.J., 2001. Differential synaptic processing separates stationary from transient inputs to the auditory cortex. Nat. Neurosci. 4 (12), 1230–1237.

Bi, G.-Q., Rubin, J., 2005. Timing in synaptic plasticity: from detection to integration. Trends Neurosci. 28 (5), 222–228.

Brette, R., Gerstner, W., 2005. Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. J. Neurophysiol. 94 (5), 3637–3642.

Coath, M., Brader, J.M., Fusi, S., Denham, S.L., 2005. Multiple views of the response of an ensemble of spectro-temporal features support concurrent classification of utterance, prosody, sex and speaker identity. Network 16 (2–3), 285–300.

Coath, M., Denham, S.L., 2005. Robust sound classification through the representation of similarity using response fields derived from stimuli during early experience. Biol. Cybern. 93 (1), 22–30.

Coath, M., Denham, S.L., 2007. The role of transients in auditory processing. Biosystems 89 (1–3), 182–189.

Dayan, P., Abbot, L.F., 2001. Theoretical neuroscience: Computational and mathematical modeling of neural systems. MIT Press.

de Boer, R., Kuyper, P., 1968. Triggered correlation. IEEE Trans. Biomed. Eng. 15 (3), 169–179.

deCharms, R., Blake, D., Merzenich, M., 1998. Optimizing sound features for cortical neurons. Science 280 (5368), 1439–1443.

Denham, S.L., 2001. Cortical Synaptic Depression and Auditory Perception. In: Computational Models of Auditory Function. IOS Press, Amsterdam, pp. 281–296.

Depireux, D.A., Simon, J.Z., Klein, D.J., Shamma, S.A., 2001. Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. J. Neurophysiol. 85 (3), 1220–1234.

Destexhe, A., Rudolph, M., Par, D., 2003. The highconductance state of neocortical neurons in vivo. Nat. Rev. Neurosci. 4 (9), 739–751.

Drew, P.J., Abbott, L.F., 2006. Extending the effects of spike-timing-dependent plasticity to behavioral timescales. Proc. Natl. Acad. Sci. U.S.A. 103 (23), 8876–8881.

Eggermont, J.J., 2002. Temporal modulation transfer functions in cat primary auditory cortex: separating stimulus effects from neural mechanisms. J. Neurophysiol. 87 (1), 305–321.

Elhilali, M., Fritz, J.B., Klein, D.J., Simon, J.Z., Shamma, S.A., 2004. Dynamics of precise spike timing in primary auditory cortex. J. Neurosci. 24 (5), 1159–1172.

Fourcaud-Trocm, N., Brunel, N., 2005. Dynamics of the instantaneous firing rate in response to changes in input statistics. J. Comput. Neurosci. 18 (3), 311–321.

Frisina, R.D., Smith, R.L., Chamberlain, S.C., 1985. Differential encoding of rapid changes in sound amplitude by second order auditory neurons, Exp. Brain Res. 60, 417–422.

Gerstner, W., Kistler, M., 2002. Spiking Neuron Models. Cambridge University Press.

Glasberg, B.R., Moore, B.C., 1990. Derivation of auditory filter shapes from notched noise data. Hear. Res. 47 (1), 103–138.

Guillery, R.W., Feig, S.L., Lozsdi, D.A., 1998. Paying attention to the thalamic reticular nucleus. Trends Neurosci. 21 (1), 28–32.

Guillery, R.W., Sherman, S.M., 2002. Thalamic relay functions and their role in corticocortical communication: generalizations from the visual system. Neuron 33 (2), 163–175.

Heil, P., 1997. Auditory cortical onset responses revisited. I. First-spike timing. J. Neurophysiol. 77 (5), 2616–2641.

Hubel, D.H., Wiesel, T.N., 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol. 160, 106–154.

Illing, R.-B., 2004. Maturation and plasticity of the central auditory system. Acta Otolaryngol. Suppl. 552 (552), 6–10.

Indiveri, G., Chicca, E., Douglas, R., 2006. A VLSI array of low-power spiking neurons and bistable synapses with spiketiming dependent plasticity. IEEE Trans. Neural Netw. 17 (1), 211–221.

Jacob, V., Brasier, D.J., Erchova, I., Feldman, D., Shulz, D.E., 2007. Spike timing-dependent synaptic depression in the in vivo barrel cortex of the rat. J. Neurosci. 27 (6), 1271–1284.

King, A., Schnupp, J., 1998. Sensory neuroscience: visualizing the auditory cortex. Curr. Biol. 8 (22), 784–787.

Klein, D.J., Depireux, D.A., Simon, J.Z., Shamma, S.A., 2000. Robust spectrotemporal reverse correlation for the auditory system: optimizing stimulus design. J. Comput. Neurosci. 9 (1), 85–111.

Kowalski, N., Depireux, D., Shamma, S., 1996a. Analysis of dynamic spectra in ferret primary auditory cortex. I. Characteristics of single-unit responses to moving ripple spectra. J. Neurophysiol. 76 (5), 3503–3523.

Kowalski, N., Depireux, D., Shamma, S., 1996b. Analysis of dynamic spectra in ferret primary auditory cortex. II. Prediction of unit responses to arbitrary dynamic spectra. J. Neurophysiol. 76 (5), 3524–3534.

Krumbholz, K., Patterson, R.D., SeitherPreisler, A., Lammertmann, C., Ltkenhner, B., 2003. Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. Cereb. Cortex 13 (7), 765–772.

Langner, G., Schreiner, C., 1988. Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. J. Neurophysiol. 60 (6), 1799–1822.

Ljung, L., 1986. System Identification: Theory for the User. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

Machens, C.K., Wehr, M.S., Zador, A.M., 2004. Linearity of cortical receptive fields measured with natural sounds. J. Neurosci. 24 (5), 1089–1100.

Marie, R.L.S., Stanforth, D.A., Jubelier, E.M., 1997. Substrate for rapid feedforward inhibition of the auditory forebrain. Brain Res. 765 (1), 173–176.

Miller, L.M., Escabi, M.A., Read, H.L., Schreiner, C.E., 2002. Spectrotemporal receptive fields in the lemniscal auditory thalamus and cortex. J. Neurophysiol. 87 (1), 516–527.

Ostergaard, P.B., 1990. Implementation details of a computational model of the inner hair cell/auditory nerve synapse. J. Acoust. Soc. Am. 87–4, 1813–1816.

Patterson, R.D., Robinson, K., Holdsworth, J., McKeown, D., Zhang, C., Allerhand, M.H., 1992. Auditory Physiology and Perception. Oxford, pp. 429–446.

Phillips, D., Hall, S., Boehnke, S., 2002. Central auditory onset responses, and temporal asymmetries in auditory perception. Hear. Res. 167 (1–2), 192–205.

Rao, V.R., Finkbeiner, S., 2007. NMDA and AMPA receptors: old channels, new tricks. Trends Neurosci. 30 (6) 284–291.

Rhode, W., Greenberg, S., 1994. Encoding of amplitude modulation in the cochlear nucleus of the cat. J. Neurophysiol. 71 (5), 1797–1825.

Richardson, M.J.E., Brunel, N., Hakim, V., 2003. From subthreshold to firing-rate resonance. J. Neurophysiol. 89 (5), 2538–2554.

Rouiller, E., de Ribaupierre, Y., Toros-Morel, A., de Ribaupierre, F., 1981. Neural coding of repetitive clicks in the medial geniculate body of cat. Hear. Res. 5 (1), 81–100.

Rouiller, E., de Ribaupierre, F., 1982. Neurons sensitive to narrow ranges of repetitive acoustic transients in the medial geniculate body of the cat, Exp. Brain Res. 48 (3), 323–326.

Slaney, M., 1994. Auditory toolbox documentation. Technical report 45, Tech. rep., Apple Computers Inc.

Theunissen, F.E., David, S.V., Singh, N.C., Hsu, A., Vinje, W.E., Gallant, J.L., 2001. Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. Netw.: Comput. Neural Syst. 12, 289–316.

Theunissen, F.E., Sen, K., Doupe, A.J., 2000. Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds. J. Neurosci. 20 (6), 2315–2331.

Thomson, A.M., Bannister, A.P., 2003. Interlaminar connections in the neocortex. Cereb. Cortex 13 (1), 5–14.

Trussel, L.O., 2002. In: Oertel, D., Fay, R.R., Popper, A.N. (Eds.), Integrative Functions in the Mammalian Auditory Pathway. Springer.

Tsodyks, M.V., Markram, H., 1997. The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. Proc. Natl. Acad. Sci. U.S.A. 94 (2), 719–723.

Ullman, S., Vidal-Naquet, M., Sali, E., 2002. Visual features of intermediate complexity and their use in classification. Nat. Neurosci. 5 (7), 682–687.

Wang, X., 2004. The unexpected consequences of a noisy environment. Trends Neurosci. 27 (7), 364–366.

Winer, J.A., Miller, L.M., Lee, C.C., Schreiner, C.E., 2005. Auditory thalamo-cortical transformation: structure and function. Trends Neurosci. 28 (5), 255–263.

Winer, J.A., Marie, R.L.S., Larue, D.T., Oliver, D.L., 1996. GABAergic feedforward projections from the inferior colliculus to the medial geniculate body. Proc. Natl. Acad. Sci. U.S.A. 93 (15), 8005–8010.

Zhang, L., Bao, S., Merzenich, M., 2001. Persistent and specific influences of early acoustic environments on primary auditory cortex. Nat. Neurosci. 4 (11), 1123–1130.

Zhang, L.I., Bao, S., Merzenich, M.M., 2002. Disruption of primary auditory cortex by synchronous auditory inputs during a critical period. Proc. Natl. Acad. Sci. U.S.A. 99 (4), 2309–2314.

Zhang, J., Gill, P., 2006. Strfpak documentation, Tech. rep., Theunissen and Gallant Laboratories, UC Berkeley.