# Behavioral/Systems/Cognitive Cortical Processing of Dynamic Sound Envelope Transitions

**Yi Zhou** and **Xiaoqin Wang**

Laboratory of Auditory Neurophysiology, Department of Biomedical Engineering, School of Medicine, Johns Hopkins University, Baltimore, MD 21205

## Abstract

Slow envelope fluctuations in the range of 2-20Hz provide important segmental cues for processing communication sounds. For a successful segmentation, a neural processor must capture envelope features associated with the rise and fall of signal energy, a process that is often challenged by the interference of background noise. This study investigated the neural representations of slowly varying envelopes in quiet and in background noise in the primary auditory cortex (A1) of awake marmoset monkeys. We characterized envelope features based on the local average and rate of change of sound level in envelope waveforms and identified envelope features to which neurons were selective by reverse correlation. Our results showed that envelope feature selectivity of A1 neurons was correlated with the degree of non-monotonicity in their static rate-level functions. Non-monotonic neurons exhibited greater feature selectivity than monotonic neurons in quiet and in background noise. The diverse envelope feature selectivity decreased spike-timing correlation among A1 neurons in response to the same envelope waveforms. As a result, the variability, but not the average, of the ensemble responses of A1 neurons represented more faithfully the dynamic transitions in low-frequency sound envelopes both in quiet and in background noise.

## Keywords

sound envelope; auditory cortex; synchrony; non-monotonicity; segmentation

## Introduction

One task routinely faced by the auditory system is the parsing of a mixture of sounds arriving at the ear(s) into individual streams of perceptual events (Bregman, 1990). As a necessary step, the temporal boundaries of superimposed events must be identified in order to construct a proper ordering of stimulus sequences. Previous studies have proposed that such envelope-transient sensitivity, or "temporal edge detection", presents at the level of single neurons in auditory cortex (Fishbach et al., 2001; Phillips et al., 2002). Despite frequent reports of onset and offset responses to sound envelopes throughout the central auditory system (Bieser and Muller-Preuss, 1996; Kuwada and Batra, 1999; Shaddock Palombi et al., 2001; Liang et al., 2002), a systematic approach for distinguishing and classifying neural responses to various transient components in sound envelope is still lacking.

**Corresponding author:** Dr. Yi Zhou Laboratory of Auditory Neurophysiology Department of Biomedical Engineering School of Medicine, Johns Hopkins University 720 Rutland Avenue, Traylor 419 Baltimore, MD 21205-2195, U.S.A. Tel (410) 614-4547 / Fax (410) 614-9599 zhouyi@jhu.edu.

The purpose of this study is to examine neural selectivity to dynamic envelope features in the primary auditory cortex (A1) of awake marmoset monkeys. Of particular interest is the low-frequency (~4Hz) amplitude modulation (AM) known to provide important segmental cues for processing communication sounds including human speech (Houtgast and Steeneken, 1985; Rosen, 1992) and animal vocalizations (Rose, 1986; Wang et al., 1995; Ghazanfar and Hauser, 2001). Traditionally, low-frequency AM has been studied as part of the continuum of modulation frequency (MF) selectivity found in auditory neurons (See reviews by Langner, 1992; Frisina, 2001; Joris et al., 2004; Wang et al., 2008). MF selectivity can be characterized using a Fourier-based linear systems approach in terms of modulation transfer functions based on firing rate or synchrony (e.g., Schreiner and Urbas, 1986; Eggermont, 1994). However, due to considerable nonlinearity in AM responses (Joris et al., 2004), MF analysis has limited power in predicting neural responses to aperiodic dynamic envelope transitions. One example is the differential neural responses to forward and time-reversed envelope waveforms observed at multiple levels of the ascending auditory pathway (Pressnitzer et al., 2000; Lu et al., 2001; Neuert et al., 2001), suggesting that selectivity to the direction of change in sound level is a preserved feature of auditory neurons in addition to MF selectivity. As such, the directional selectivity (which can be described by the slope of envelope) permits auditory neurons to differentiate non-stationary and/or aperiodic stimuli commonly found in communication signals such as the envelopes of steady-state vowels (Olive et al., 1993).

Considering these factors, the present study used a reverse correlation method to directly extract the spike-triggering dynamic envelope features in low-frequency aperiodic AM sounds. Envelope feature selectivity was compared between groups of neurons showing monotonic or non-monotonic static rate-level functions (RLFs), in an effort to further distinguish functional neuronal groups in A1 by their level-response characteristics (e.g., Sadagopan and Wang, 2008; Watkins and Barbour, 2008). To evaluate the robustness of envelope feature selectivity, responses of the two neuronal populations were compared in background noise. We found that the variability, and not the average, of the ensemble responses of A1 neurons provide a robust representation of envelope transitions both in quiet and in background noise.

## Methods and Materials

### Animal preparation, apparatus, and electrophysiological recordings

A chronic recording preparation (e.g., Lu et al., 2001a) was used to record single-neuron activity in the primary auditory cortex (A1) of awake adult common marmoset monkeys (*Callithrix jacchus*). All experimental procedures were approved by the Institutional Animal Care and Use Committee of the Johns Hopkins University following NIH guidelines.

Experiments were conducted in a double-walled acoustic chamber (IAC-1024, Industrial Acoustics). The internal walls and ceiling were lined with three-inch acoustic absorption foam (Sonex, Illbruck) to reduce acoustic reflections. In the early stages of experiments, sounds were delivered from a loudspeaker (B&W 601) located 90 cm directly in front of the animal. A multi-speaker setup was used in the later stages of experiments. Fifteen loudspeakers (Dome Tweeter, Fostex Co.) were positioned in the semicircular frontal field (−90° to 90° along the horizontal axis and at 0°, 45°, 90° elevations) at a distance of approximately 80 cm from the head of the animal.

Before surgery and chronic recordings began, animals were adapted to sit still in a custom-designed primate chair. After two weeks, two stainless head posts were attached to the animal's skull under sterile conditions with the animal deeply anesthetized by isoflurane (0.5-2.0%, mixed with 50% oxygen and 50% nitrous oxide). The head posts were used to

immobilize the animal's head during chronic recordings. To access the auditory cortex, small craniotomies (~1 mm in diameter) were made on the skull over the superior temporal gyrus to allow for penetration of electrodes (tungsten electrodes, 2-5 MΩ impedance, A-M systems, Carlsborg, WA). A hydraulic microdrive (Trent-Wells, Los Angeles, CA) was used to advance the electrodes slowly through the dura to cortex. Simultaneously, a set of search stimuli was played including tones, band-pass noises, and animal vocalizations. Single-neuron activity was sorted online using a template-based spike-sorting program (MSD, Alpha Omega Engineering) and stored for off-line data analysis in Matlab (Mathworks, Natick, Massachusetts).

## Single-neuron characterization

Single-neuron responses were collected from the primary auditory cortex (A1) in three hemispheres of one male and one female adult marmoset monkey. Pure tones (0.5~32 kHz in 10 steps/octave) were used to characterize the frequency tuning property of a neuron. The best frequency (BF) of a neuron was defined as the frequency evoking the maximal firing rate over the range of sound levels tested. Responses beginning from the tone onset to 50 ms after the tone offset were included for the BF analysis. The median BF of all neurons collected was 10.6 kHz with the lower and higher quartiles at 5.6 kHz and 17.1 kHz. Once the BF of a neuron was determined, the rate-level function (RLF) using 100 ms BF tones was collected over a range of levels from -10dB to 80dB SPL in 10dB steps. If the physiological conditions allowed, an additional RLF in background noise (white noise with a duration of 200 ms temporally surrounding the BF tone) was also collected. The median noise level used was 40dB SPL for neurons reported in this study with the lower and higher quartiles at 30dB and 60dB SPL, respectively. Tone and noise intensities are both expressed in terms of the peak-to-peak equivalent dB SPL. The reference amplitude was set by a 1 kHz tone calibrated at ~90dB SPL (re. 20 $u$Pa) with zero dB attenuation. The rms SPL of noise was 40 dB/Hz at zero dB peak attenuation.

Neural responses were compared between three subpopulations of neurons based on the monotonicity of their static RLFs. Monotonicity Index (MI) was defined as the ratio of the firing rate at the loudest sound level used (80dB SPL) to the maximal firing rate at the best sound level. 156 out of 214 neurons (73%) were classified as non-monotonic neurons (MI<1). This fraction is similar to those previously reported in awake monkeys (macaque: Pfingst and O'Connor, 1981; marmoset: Sadagopan and Wang, 2008) and higher than those reported in anesthetized cats (e.g., Sutter and Schreiner, 1995; Moshitch et al., 2006). The non-monotonic neurons were further divided into two subpopulations: highly non-monotonic ("non", MI  0.2) and moderately non-monotonic ("non$_{mod}$", 0.2<MI<1). This division allowed us to examine the envelope responses of A1 neurons showing rate-level monotonicity at two extreme ends (i.e., monotonic neurons, "mon," vs. highly non-monotonic neurons, "non"). The arbitrary boundary of 0.2 was chosen to match the number of neurons (58 "mon" and 59 "non") and the number of test conditions (95 "mon" and 95 "non") between the monotonic and highly non-monotonic groups for the purpose of statistical analyses. The test conditions included instances of responses of a neuron tested at multiple sound levels. The total numbers of test conditions and numbers of neurons are listed in Table I. The SPL distribution of the data sample is given in later sections.

## aAM stimulus generation

We examined envelope feature selectivity of A1 neurons in quiet and in background noise. Our test stimuli were aperiodically amplitude-modulated tones (aAM) with a carrier frequency at the BF of the neuron under study. The aAM waveforms were generated using a modified formula used for sinusoidal amplitude-modulation (SAM), $A = [1+\cos(2\pi f_m t + \phi))]/2$. For SAM stimuli, the modulation frequency $f_m$ and the starting phase $\phi$ were constant

numbers; for aAM stimuli, the modulation frequency $f_m$ was a time-varying random variable chosen at each sample point from a normal distribution with a mean of 4 Hz and a standard deviation of 1 Hz and the initial phase $\phi$ was a time-varying random variable chosen at each sample point from a uniform distribution between $-\pi$ and $\pi$. In their digital forms, the aAM waveforms were generated at a low sampling rate of 30Hz. The set of discrete numbers were then spanned into continuous time functions at a sampling rate of 100 kHz using spline interpolation in Matlab. A total of ten waveforms (five plus their time inverse versions) that yielded temporally dissimilar patterns were used in this study. The durations of the aAM stimuli were 500 ms when presented alone (T) and presented in background noise (T+N). In the T+N condition, the broadband noise was gated on 50 ms before the aAM onset and gated off 50 ms after the aAM offset. The 50-ms out-of-alignment helped the experimenters disambiguate the onset and offset temporal patterns to aAM stimuli and to broadband noise. Different tokens of frozen noise were used for different neurons. All stimuli were gated on and off with a 5 ms linear ramp. At least 10 repetitions (20 in some cases) of responses were collected for each stimulus with an inter-stimulus interval of 1000 ms. Before delivery to a loudspeaker, the aAM and noise stimuli were converted to analog signals (DA4), passed through two separate attenuation modules (PA4), and summed physically (SM3) using the TDT systems.

For a given neuron, the aAM SPL was chosen at least 10dB above the threshold based on its static RLF. In the data sample, more than half of the neurons were tested at more than one SPL (37 out of 58 monotonic neurons, 72 out of 97 moderately non-monotonic neurons, and 36 out of 59 highly non-monotonic neurons). The lower quartile, median, higher quartile of SPLs were at 40, 60, 70 dB SPL, respectively, for monotonic and moderately non-monotonic neurons and at 30, 40, 60dB SPL, respectively, for highly non-monotonic neurons.

### Using reverse correlation to characterize neural envelope feature selectivity

As a first-order linear approximation, the dynamic transitions in envelope modulation waveforms can be characterized by the local average (in terms of mean) and the rate of change (in terms of slope) of sound level. We described neural selectivity to the joint mean-slope variations using an envelope feature map, which depicted the conditional probability of the firing rate of a neuron for stimulus s with a specific mean-slope combination – P(spike|$s$), which via Bayes' theorem is proportional to the ratio P($s$|spike)/P($s$). We estimated the spike-triggered stimulus feature distribution, P($s$|spike), and the total stimulus feature distribution, P($s$), using the reverse correlation method.

**Envelope feature density and feature types in aAM stimuli**—The means and slopes of aAM envelopes at a given peak SPL were extracted from linear regression analysis on consecutive envelope segments expressed in dB (25 ms in duration) with a 1-ms lag between adjacent segments. To capture the onset and offset dynamics, a 25-ms zero envelope segment was padded before the aAM onset and a 50-ms zero envelope segment was padded after the aAM offset. The zero envelope amplitude corresponded to -10dB SPL in this study. The joint mean-slope distribution was binned with a resolution of 2dB for the mean and 0.24dB/ms for the slope. The joint mean-slope distribution represented in approximate terms the envelope feature distribution P($s$) in the aAM stimuli at a given peak SPL. Since the mean-slope distribution was governed by modulation frequency, and since the ten aAM envelopes differed in modulation phase, but not much in modulation frequency (~4Hz temporal modulation), their P($s$) showed similar patterns and was not individually reported in data analysis. As discussed below, a complete characterization of P(s) requires a full spectrum analysis with an infinitely small time window. This temporal resolution, however, is unlikely achieved by cortical neurons. Using the 1-ms time resolution, the

estimated P($s$) was in essence the feature map of an "all-pass" neuron with a constant firing rate of 1 kHz and with no preference to any of envelope features.

Based on the signs of envelope slope values, the complete set of envelope features were divided into five categories ("onset", "up", "peak", "down", and "offset" types). These feature types described fast ("onset" and "offset" types), slow ("up" and "down" types), and zero ("peak" types) dynamic envelope transitions. In the logarithmic scaling, envelope slope values within each of the five feature types are in theory level invariant, i.e., d[log(AX(t))]/dt=d[logX(t)]/dt, where X(t) is the time-varying sound pressure and A is a scale factor. In other words, envelope slopes should remain in a fixed relationship with time despite variations in sound level. Achieving this theoretical limit requires one to assess the instantaneous means and slopes using an infinitely small time window dt. Any longer dt would lead to underestimated slope values related to rapid changes in sound level, which is inevitable in the linear approximation method we used. Therefore, we used absolute times, as opposed to absolute slope thresholds, to define "onset" and "offset" envelope features. More specifically, the envelope features within the first 25 ms after the stimulus onset were attributed to the "onset" type and the envelope features within a 50 ms window after the stimulus offset were attributed to the "offset" type. For the remaining features, those with the envelope slope values larger than 0.24 dB/ms were attributed to the "up" type; those with the envelope slope values less than −0.24 dB/ms were attributed to the "down" type; those with the envelope slope values between −0.24 dB/ms and 0.24 dB/ms were attributed to the "peak" type.

**Envelope feature selectivity of A1 neurons—**The relationship between the spike times and envelope features was established by reverse correlation as well. To account for the transmission delay between the loudspeaker and the recoding site, all spike times after the aAM stimulus onset were advanced by 15 ms (~2.16 ms between the loudspeaker and the center of an animal's head; ~10 ms was the minimal first-spike latency). Data analyses were conducted based on these modified spike times. The reverse correlation procedure was performed on spike times occurring between 0 ms after the aAM stimulus onset and 50 ms after the aAM stimulus offset in both T and T+N conditions. The mean and slope values of an envelope segment (with a duration of 25 ms) preceding each spike were extracted and binned using the same method for characterizing P($s$). The resultant P($s$|spike) was then normalized by the total stimulus features P($s$) at the sound level tested to get P(spike|$s$). To avoid the small denominator effect for normalization, all values in P($s$) less than 0.5% of the peak were considered insignificant features [which resulted in equal or smaller P($s$|spike)] and therefore replaced with a large number 10000 before normalization. This procedure ensured a near-zero P(spike|$s$) for small P($s$). The final result was denoted as the envelope feature map of a neuron. As the result of stimulus normalization, the magnitude of an envelope feature map is proportional to the probability of the spike rate per stimulus s in terms of a mean-slope combination, P(spike|$s$).

The envelope Feature Preference Index (FPI) of a neuron was defined as FPI = ($CC_{max}$ − $CC_{min}$)/($CC_{max}$ + $CC_{min}$), where $CC_{max}$ and $CC_{min}$, where and were the maximal and minimal correlation coefficients between the envelope feature map of a neuron and mean-slope distributions of five feature types. Considering that envelope responses of neurons were tested across a range of sound levels, we calibrated each FPI value based on the feature distribution in the aAM stimuli at the SPL tested (which ranged from 0 to 80dB SPL) using an all-pass model neuron, as mentioned earlier, whose feature map was simply the stimulus feature distribution P(s). It was observed that when the sound level increased, the relative densities of onset and offset envelope features became sparse due to fast changes in the envelope slope values, resulting in lower $CC_{min}$ and therefore higher FPI values of the all-pass neuron. The FPI of the all-pass model was 0.2657, 0.1581, 0.2270 0.2251, 0.2636,

0.2912, 0.3693, 0.4481, and 0.4297 from 0dB to 80dB SPLs. We corrected this measurement artifact by subtracting the all-pass model FPI value from the neural FPI values at each SPL. After correction, the median FPI values did not show sound-level dependence for either group of neurons ("mon": $R^2=0.0013$; "$non_{mod}$": $R^2=0.0001$; "non": $R^2=0.0004$). This observation enabled us to pool data tested at different SPLs in characterizing the envelope feature selectivity of a neuron.

### Analysis of the ensemble responses of a population of neurons

The analysis of spike-timing patterns was conducted on the peri-stimulus time histograms (PSTHs) of single neurons. The PSTHs were generated by counting spikes in fixed time bins as a function of time and then smoothing these time series using a Gaussian filter (which had a mean of zero and standard deviation of the binwidth) truncated at $+/- 3$ standard deviations with unit energy. The binwidths used were 1, 2, 5, 10, 12.5, 25, and 50 ms. The number of spikes per bin was averaged over 10 or 20 repetitions except for the analysis of between-trial correlation of spike times as described below.

For the between-neuron and within-neuron correlation analyses shown in Fig. 7, the PSTHs to the ten aAM stimuli were concatenated into a single PSTH vector. Only responses between 15 ms after the aAM onset and 65 ms after its offset were analyzed in both T and T +N conditions. The between-neuron correlation analysis measured the Pearson correlation coefficient between two PSTH vectors of two different neurons, while the within-neuron correlation analysis measured the Pearson correlation coefficient between two PSTH vectors of the same neuron tested in the T and T+N conditions.

For the between-trial correlation analysis shown in Fig. 7, the Pearson correlation coefficient was measured between PSTHs of a neuron from two different trials in response to the same stimulus. The average correlation coefficient from all possible pairs of comparisons was obtained for each neuron. The dependence of between-trial correlation on logarithmic average firing rate (in response to ten aAM stimuli) was assessed by the linear regression analysis, which yielded the slope and $R^2$ estimates. To remove the rate dependence of between-trial correlation, the product of slope and the corresponding logarithmic average rate was subtracted from each correlation data point.

For the population response analyses shown in Figs. 8 and 9, the PSTH vectors of all neurons were assembled into a PSTH matrix for the T and T+N conditions, respectively; neurons that were tested at more than one sound level had multiple entries in the PSTH matrices. The average of ensemble responses was simply the mean of all PSTH vectors, and the variability of ensemble responses was measured by the mean-normalized variance of all PSTH vectors, namely, the Fano Factor (FF) at each time bin $t_n$ [$Var(t_n)/Mean(t_n)$]. The level-dependence of neural responses was not taken into account in the ensemble analyses.

To analyze the relationships between the envelopes of aAM stimuli and ensemble responses of neurons, two types of correlation analyses were conducted for results shown in Fig. 9: (1) between the amplitude of the aAM envelopes and the average of ensemble PSTHs, and (2) between the absolute value of the slope of aAM envelopes and the FF of ensemble PSTHs. Stimuli and neural responses were both smoothed with a 10-ms window. The correlation analyses (1) and (2) were made with the aAM stimuli at a peak level of 50dB SPL. Considering that neurons were tested at different SPLs, we measured non-parametric Spearman rank correlation coefficients for (1) and (2). To evaluate whether increasing the population size enhanced envelope encoding, we took repetitive samples from a PSTH matrix through a bootstrap procedure (n=50 repeats), and varied the ensemble size (i.e, the number of PSTHs in a sample) from roughly a quarter to the full data set. At each ensemble size, we calculated the amplitude and FF of the ensemble PSTHs and then conducted the

correlation analyses (1) and (2). The mean and SD of population results were reported in Fig. 9.

### Statistical significance tests

Neural responses were compared among three neural groups ("mon", "non$_{mod}$", "non"). The significance of the bimodality in MI was tested with a Hartigan's Dip test (Hartigan and Hartigan, 1985) using a Matlab algorithm adapted from F. Mechler's original code (downloaded from http://www.nicprice.net/diptest/). For a given data set, we used Lilliefors tests (Lillietest.m in Matlab) to examine whether the data samples were normally distributed. Subsequently, Student-t tests (ttest.m and ttest2.m in Matlab) were used to compare population means that were normally distributed and Wilcoxon rank-sum tests (ranksum.m in Matlab) were used to compare population medians that were not normally distributed. The trend analysis on a given data set was based on a linear regression $t$-test. The $R^2$ and $t$ statistic of the slope were reported. We used an alpha level of 0.05 for all statistical tests.

## Results

This study examined single-neuron responses in the primary auditory cortex (A1) to amplitude-modulated tones presented at each neuron's best frequency (BF). The modulating envelopes were aperiodic waveforms oscillating at a long-term average rate of 4Hz (denoted as aAM stimuli). In total, 214 single neurons with significant tone-driven responses (paired $t$-test, $p<0.05$) were tested with aAM stimuli at multiple sound levels; 102 neurons were further tested with aAM stimuli in the presence of broadband noise.

### Characterizations of envelope features of aAM stimuli

When a sound envelope oscillates at a low rate, neural responses are affected by both the change in sound level as well as the mean sound level. To characterize the densities of these envelope features, we analyzed the local average (in terms of mean) and the rate of change (in terms of slope) of sound intensity in the envelopes of a set of ten aAM stimuli using a linear regression analysis (See Methods). Figure 1A shows an example of the envelope of an aAM stimulus on a decibel scale (left panel) and the corresponding mean-slope trajectory (right panel). Horizontal color bars under the aAM stimulus mark the time durations of envelope features with different slope polarities. Accordingly, the "onset" and "up" features have positive slopes, the "offset" and "down" features have negative slopes, and the "peak" feature has zero slope. As shown on the mean-slope trajectory, these five features are associated with different combinations of mean and slope values of the envelope.

Next, we examined how sound level influenced the envelope feature distribution in the mean-slope plane. Figure 1B compares the averaged mean-slope distribution obtained from all ten aAM stimuli at 80dB and 30dB SPL, respectively. The spatial pattern of the mean-slope distribution is more restricted at 30dB than at 80dB SPL. The difference is also evident when individual envelope features were examined separately (Fig. 1C). Lowering the peak sound level caused a downward shift in the mean values of all envelope features. The relatively robust slope distribution for each feature (except for those of "onset" and "offset" types) can be explained by a unique property of the logarithmic scaling. In this case, the time derivative of sound level (i.e., slope) is inherently level invariant. The reduction in the slope values of "onset" and "offset" at 30dB SPL was caused by zero padding before the onset and after the offset of the aAM stimuli, which reduced the dynamic ranges of the rise and fall of aAM stimuli during the linear regression analysis. Overall, these results provide a quantitative characterization of envelope features in aAM stimuli. Because the five envelope

feature types exhibit largely distinct mean-slope patterns at a given SPL, they were used as classifiers to characterize neural selectivity to different envelope features in aAM envelopes.

## Characterization of neural selectivity to envelope features

We used a reverse correlation method (de Boer and de Jongh, 1978) to characterize the envelope feature selectivity of a neuron. The basic procedure is detailed in the Methods and illustrated in Fig. 2A. The top panel of Fig. 2A shows the envelope of an aAM stimulus (gray line, 500 ms duration, 60dB SPL) and the raster plot of responses of an example neuron. The mean and slope values of the 25-ms envelope segments immediately preceding each spike were extracted via spike triggered average and then normalized by the stimulus mean-slope distribution (e.g., Fig. 1B). The middle panel of Fig. 2A shows the stimulus-normalized mean-slope distribution (denoted as envelope feature map) of the neuron, which demonstrates a preference to envelope segments with negative slope values. The preferred envelope feature type of the neuron was then determined based on the strength of pixel-by-pixel correlations between the envelope feature map of a neuron and the mean-slope distributions of five feature types within aAM stimuli at the sound level tested. As shown in the bottom panel of Fig. 2A, the "down" type yielded the highest correlation coefficient (CC) with neural responses and was then designated as the preferred envelope feature of this neuron.

This procedure was applied to all neurons tested with aAM stimuli. Fig. 2B shows four more example neurons. They exhibited diverse temporal response patterns to the same envelope waveform of an aAM stimulus. The envelope feature maps captured their distinctive preferences to different envelope features in aAM stimuli. Based on the maximal 2-D CC values (highlighted in red), the preferred envelope features of the four neurons from left to right were characterized as "onset", "up", "peak", and "offset" types, respectively. These example neurons also exhibited different amounts of suppression in their RLFs, as indicated by their Monotonicity Index (MI) values. The MI values for the neuron shown in Fig. 2A and the four shown in Fig. 2B were in turn 0.86, 0.33, 0, 0.49, and 0.15. The relationship between envelope feature selectivity and MI is further investigated in the next section.

## A1 neurons respond selectively to different envelope features

To distinguish the functional neural groups in A1, we compared the envelope feature selectivity of neurons showing monotonic and non-monotonic RLFs. Our hypothesis was that inhibition that influences rate-level non-monotonicity might attribute to differences observed in envelope feature preferences among cortical neurons. Figure 3A shows the MI distribution of neurons collected for this study, where MI is defined as the ratio of the firing rate at the loudest sound level used (80dB SPL) to the maximal firing rate at the best sound level (Pfingst and O'Connor, 1981). Similar to the previous report from our laboratory (see Fig. 2D in Sadagopan and Wang, 2008), we observed that in the auditory cortex of awake marmoset, (1) a majority of neurons discharged less at higher BF-tone levels (MI<1, 156 out of 214 neurons, 73%), suggesting increasing inhibitory effects of input with sound level, and (2) the MI distribution was bimodal with a dip around 0.7 (Hartigan's Dip test, $p$<0.001). To distinguish the potential influences of inhibition on envelope responses, in this study, the neural population was divided into three subgroups and neural responses were compared between those classified as monotonic (MI=1, denoted as the "mon" group) and highly non-monotonic (MI<=0.2, denoted as the "non" group). The remaining neurons showing moderate non-monotonicity (0.2<MI<1, denoted as the "non$_{mod}$" group) were used as a control group in all tests (See Methods for inclusion criteria). The BF distributions were indistinguishable among the three neural groups. The median BFs was 10.6 kHz, 11.3 kHz, and 10.1 kHz for "mon", "non$_{mod}$", and "non" neurons, respectively (one-way ANOVA, $p$=0.21). However, non-monotonic neurons were found spatially closer to the depth of the

first spike encountered in penetrations. The median distance from the first spike ($D_{spk}$) was 150, 200, and 275 micron for "non", "non$_{mod}$", and "mon" neurons, respectively. Pair-wise comparisons revealed that the depth difference between "non" and "mon" neurons was significant ($p<0.05$), but not that between "non$_{mod}$" and "mon" neurons ($p>0.05$); *ranksum* test. Nonetheless, the distribution of $D_{spk}$ overlapped considerably among the three subgroups (two-sample Kolmogorov-Smirnov test; $p>0.13$), suggesting that the MI criterion alone was insufficient to separate the spatial patterns of A1 neurons across cortical layers.

Figure 3B summarizes the proportions of preferred envelope feature types of the three groups of neurons. Consistent with previous findings in auditory cortex (Phillips et al., 2002), the majority of A1 neurons responded preferentially to the positive envelope slopes, as shown by the higher percentages of the "onset" and "up" types than the "peak", "down" and "offset" types in all three neuronal groups. By comparison, the preferred envelope features of non-monotonic neurons were more diverse than those of monotonic neurons as manifested by higher percentages of "peak", "down", and "offset" types. Figure 3C plots the medians of the ordered stimulus-response correlations (as illustrated in Fig. 2), which indicate the strength of selectivity to individual envelope features. Due to differences in feature preferences among neurons, the correlation values within the same rank could represent different envelope feature types. Correlations were significantly different among the three neuronal groups for less preferred envelope features (5th: F=7.01, $p<0.001$; 4th: F=9.48, $p<0.0001$; 3rd: F=5.38, $p<0.01$), but not for more preferred envelope features (2nd: F=1.5, $p=0.22$; 1st: F=0.25, $p=0.78$). The pair-wise comparison revealed that the highly non-monotonic neurons had lower selectivity to the three less preferred envelope features than did monotonic neurons (*ranksum* test, $p<0.01$). Since low selectivity arises from low firing rates to a particular envelope feature, these data show that the responses of highly non-monotonic neurons were more tightly associated with their preferred envelope features than those of monotonic neurons.

This observation prompted us to measure the relative strength of neural selectivity to different envelope features using a contrast metric: Feature Preference Index (FPI) defined as FPI = $(CC_{max} - CC_{min})/(CC_{max} + CC_{min})$, where $CC_{max}$ and $CC_{min}$, were the maximal and minimal correlation coefficients (e.g., those ranked 1st and 5th in Fig. 3C), representing the most and least preferred envelope features. This analysis was ensured by non-zero differences between $CC_{max}$ and $CC_{min}$, which were above 0.16 for neurons reported in this study. Considering that responses of neurons were collected across a range of SPLs, we calibrated each FPI value based on the envelope feature distribution in the aAM stimuli at the SPL tested. FPI values showed no SPL dependence after calibration (See the Methods).

The relationship between envelope feature selectivity and RLF monotonicity was evaluated for each envelope feature type. As shown in Fig. 3E, significant negative correlations were observed between FPI and MI for the three most prevalent feature types in Fig. 3B, which accounted for 84% of total observations (onset: $R^2=0.18$, $t_{(178)}=-6.26$, $p<0.001$; up: $R^2=0.11$, $t_{(82)}=-3.16$, $p<0.01$; peak: $R^2=0.17$, $t_{(42)}=-2.95$, $p<0.01$; linear regression $t$-test), and not for "down" ($R^2=0.04$, $t_{(40)}=-1.26$, $p=0.11$) and "offset" ($R^2=0.29$, $t_{(13)}=-2.32$, $p=0.39$) feature types. The difference in the strength of selectivity was evident between neurons with very low and very high MI values. Figure 3D summarizes the FPI values (mean ± SEM) of monotonic and two non-monotonic neuronal groups. The highly non-monotonic neurons had greater selectivity than the monotonic neurons for all feature types ($t$-test, $p<0.05$), while the difference between the moderately non-monotonic and monotonic neurons was only significant for the "onset", "peak", and "down" types ($t$-test, $p<0.05$). Overall, these results revealed a moderate correlation between RLF monotonicity (in response to BF-tone stimuli with a flat envelope) and envelope feature selectivity (in response to BF-tone stimuli with a dynamic envelope), suggesting that similar neural

mechanisms, such as synaptic inhibition, may underlie the distinctive response types among A1 neurons.

Demonstrating that A1 neurons indeed encode distinctive envelope features (Fig. 3B) requires careful examinations of the effect of sound level on envelope selectivity. In this study, the three neural groups were tested with aAM stimuli at slightly different intensity ranges (median SPL was 60dB for monotonic and moderately non-monotonic neurons and 40dB for highly non-monotonic neurons). One may argue that the greater instances of "onset" and "up" types and fewer instances of "peak", "down", and "offset" types for monotonic neurons relative to non-monotonic neurons (Fig. 3B) might be caused by a shift in neural preferences to positive envelope slope values with SPL. This conjecture, however, was not supported by the data. For neurons tested at more than one SPL, nearly half retained their feature preferences ("mon": 20/37, 54%; "non$_{mod}$": 43/72, 60%; "non": 18/36, 50%) with an increase of sound level (median, 20dB SPL). Among those altered their feature preferences, many changed between "onset" and "up" types or between "down" and "offset" types without altering the slope sign ("mon": 10/37, 27%; "non$_{mod}$": 18/72, 25%; "non": 9/36, 25%). In contrast, an inverse of slope preference from negative to positive was not frequently observed ("mon": 4/37, 11%; "non$_{mod}$": 6/72, 8%; "non": 6/36, 17%); equally rare was an inverse of slope preference from positive to negative ("mon": 3/37, 8%; "non$_{mod}$": 5/72, 7%; "non": 3/36, 8%). These ratios were highly preserved across three neuronal groups, suggesting that neural selectivity to diverse envelope features likely arises from level-invariant response properties, such as envelope slope preference.

The envelope slope preference of a neuron can be directly evaluated by the symmetry of an envelope feature map around the abscissa (i.e., zero-slope line). We quantified the degree of the symmetry using Symmetry Index (SI), defined as the 2-D correlation coefficient between the activity on an envelope feature map above the abscissa and that below. Figure 4A shows the feature map of the "onset" neuron shown previously in Fig. 2B. It is highly asymmetric with a SI of 0.17. To ensure that the slope preference shown in Fig. 4A was not caused by random spiking activity during stimulus presentation, we calculated a control feature map using shuffled spike times (Fig. 4B). The activity on the control map shows a symmetric pattern with a SI of 0.9 and no clear preferences for any of the five feature types.

The majority of neurons in the sample exhibited envelope slope preferences to some extent. Figure 4C is the scatter plot of SI values of the original and control feature maps. Of 365 test conditions (which included 214 single neurons tested using more than one sound level), all but one showed increased symmetry in the control feature maps, indicating that randomizing spike times eliminated the slope preference of a neuron. As summarized in Fig. 4D, the original feature maps were much less symmetric than the control feature maps for each feature type (onset: $t_{(358)}$=26.8, $p$<0.001; up: $t_{(166)}$=16, $p$<0.001; peak: $t_{(86)}$=11.7, $p$<0.001; down: $t_{(82)}$=12.4, $p$<0.001; offset: $t_{(28)}$=7, $p$<0.001). Moreover, the symmetry of a feature map was correlated with the feature type of a neuron. Feature maps of slope-sensitive neurons ("onset", "offset", "up", and "down" neurons) were more asymmetric than those of peak-sensitive neurons ($t$-test, p<0.05). This is not seen among the control feature maps derived from shuffled spike times ($t$-test, p>0.5). This finding supports a long-hypothesized notion that sensitivity to envelope transients is related to sensitivity of neurons to the rate of change in sound level in auditory cortex (Phillips and Hall, 1987; Schreiner and Urbas, 1988; Heil and Irvine, 1998).

### Robust neural selectivity to envelope features in background noise

Next, we examined the robustness of envelope feature selectivity in background noise. Figure 5 shows the raster plots and peristimulus time histograms (PSTHs) of three representative neurons (one monotonic and two highly non-monotonic) in response to a pair

of forward and reversed aAM envelopes presented alone (T, left column) and presented in background noise (T+N, right column). In comparison, persistent background noise exerted variable effects on the strength of neural responses to aAM stimuli. Specifically, noise enhanced the neural responses shown in Fig. 5A, suppressed those shown in Fig. 5B, and caused little changes to those shown in Fig. 5C. Nonetheless, the background noise did not switch neural selectivity to non-preferred envelope features, nor did it smear the temporal patterns of aAM responses, suggesting active control of envelope feature preferences by A1 neurons.

Figure 6 summarizes the quantitative comparisons between envelope feature selectivity of individual neurons measured in T and in T+N conditions. The results of monotonic, moderately non-monotonic, and highly non-monotonic neurons are shown in the left, middle, and right columns, respectively. Panels A-C of Fig. 6 compare the preferred envelope features of a neuron in T and T+N conditions. A higher percentage of highly non-monotonic neurons maintained their feature preferences in noise (77.1%, 37/48 conditions on the diagonal line) than moderately non-monotonic (64.6%, 62/96 conditions on the diagonal line) and than monotonic neurons (64.4%, 38/59 conditions on the diagonal line). Panels D-F of Fig. 6 compare the overall selectivity to all feature types in T and T+N conditions. For all three groups, noise reduced neural selectivity to the most preferred envelope feature and appeared to "enhance" neural selectivity to other less preferred envelope features (paired *t*-test, p<0.05) with the exception of 2nd and 5th features associated with the highly non-monotonic neurons in Fig. 6F (paired *t*-test, p>0.05). Such "enhancement", however, was likely associated with noise-driven excitatory responses added to the aAM responses. Nevertheless, the magnitudes of changes were rather small, especially for those of highly non-monotonic neurons. Finally, panels G-I of Fig. 6 compare the FPIs of individual neurons in T and T+N conditions. Noise caused a significant reduction in FPI values of monotonic neurons (mean reduction, $\Delta FPI=FPI_{T+N}-FPI_T=$ −0.091, paired *t*-test, p<0.01) and moderately non-monotonic neurons ($\Delta FPI=FPI_{T+N}-FPI_T=$ −0.049, paired *t*-test, p<0.02), but not in those of highly non-monotonic neurons ($\Delta FPI=FPI_{T+N}-FPI_T=$−0.013, paired *t*-test, *p*=0.469). The gradual decrease of $\Delta FPI$ with RLF non-monotonicity indicated that envelope feature selectivity of non-monotonic neurons was more robust in background noise than that of monotonic neurons. Consequently, non-monotonic neurons exhibited higher feature selectivity than monotonic neurons not only in quiet (as previously shown in Fig. 3E), but also in noise. In Figs. 6G-I, the median FPI was 0.472, 0.542, and 0.718 in the T condition and 0.363, 0.475, and 0.7 in the T+N condition for the "mon", "non$_{mod}$", and "non" neurons, respectively. The pair-wise comparison revealed significant between-group differences (*ranksum* test, p<0.05) except that between the moderately non-monotonic and monotonic neurons in the T condition (*ranksum* test, *p*=0.14).

## Synchronous vs. asynchronous population responses among monotonic and non-monotonic neurons

The observation of differential envelope response properties in A1 prompted us to examine directly the neural output of spikes times and to investigate potential population coding mechanisms for sound envelope. Multiple measurements on spike times were carried out, which included the relationship between spike-timing patterns of two different neurons (denoted as the between-neuron comparison), the robustness of spike-timing patterns between T and T+N conditions for the same neuron (denoted as the within-neuron comparison), and the reliability of spike-timing patterns of individual neurons (denoted as the between-trial comparison). The Pearson correlation coefficients were estimated on all possible pairs of PSTHs for a given comparison and their differences were evaluated by the *ranksum* test.

For the between-neuron comparison, the correlation between responses of monotonic neurons were significantly higher than that between responses of highly non-monotonic neurons, and than that between responses of moderately non-monotonic neurons in both T (Fig. 7A) and T+N conditions (Fig. 7B) at multiple time scales ($p$<1e-7). This shows that monotonic neurons responded more synchronously than non-monotonic neurons when stimulated with the same aAM stimuli. This is not a surprising result in that diversity in envelope feature types among non-monotonic neurons caused their PSTHs to peak at different times and therefore increased population asynchrony.

For the within-neuron comparison (Fig. 7C), the opposite trend was observed. The PSTHs of individual, highly non-monotonic neurons showed greater correlations between T and T+N conditions than did monotonic neurons at multiple time scales (*$p$<0.05), suggesting that the spike-timing patterns of highly non-monotonic neurons were more robust against noise than those of monotonic ones. The statistical strength of the effect was much reduced between monotonic and moderately non-monotonic neurons ($p$>0.27). Background noise also affected to variable extents the trial-to-trial variability of spike timings of neurons. We observed that the between-trial correlation (Figs. 7DE) increased with logarithmic average rates ($R^2$=0.13 in the T condition and $R^2$=0.16 in the T+N condition; $p$<$10^{-5}$). After removing the positive rate trends (See the Methods), it was revealed that the response reliability of highly non-monotonic neurons was significantly greater than that of monotonic neurons (median ΔCC=0.08; $p$<0.01) in the T+ N condition, but not in the T condition ($p$=0.3). The responsible reliability of moderately non-monotonic neurons was greater than that of monotonic neurons in both T and T+N conditions, but the differences were not significant ($p$>0.05).

Collectively, these results show that monotonic neurons yielded more synchronous population responses than non-monotonic neurons and that individual non-monotonic neurons produced more reliable, noise-tolerant spike-timing patterns than individual monotonic neurons.

## Ensemble average and variability as complementary codes for encoding sound envelope by a neural population

Retrieving information from asynchronous population activity makes an intriguing demand on neural encoding, if averaging across the neural population is considered as a coding strategy to enhance the signal-to-noise ratio. The concern is that the population average may fail to capture a signal carried by asynchronous neural activity. This point is illustrated in Fig. 8A by comparing the traces of the averaged PSTH of all monotonic neurons (red) and those of all highly non-monotonic neurons (blue) in response to one aAM envelope in the T and T+N conditions. In comparison to monotonic neurons, the averaged activity of highly non-monotonic neurons entrained less faithfully to envelope amplitude due to their asynchronous response patterns. We thus evaluated an alternative coding strategy based on the converse of the ensemble average - the variability of the ensemble responses. Figure 8B shows the Fano Factor (FF) as a function of time in T and T+N conditions. The FF was calculated as the ratio between the variance and mean of the ensemble PSTHs at each time bin. In sharp contrast to the average results shown in Fig. 8A, the FF values of the highly non-monotonic neurons entrained precisely to the rapid rises and falls of the aAM envelope, whereas the FF values of the monotonic neurons remained fairly flat, insensitive to envelope transitions (except to the envelope onset).

Differential noise effects on the population activity were also observed between the two neuronal groups. (1) The background noise increased the averaged responses of the monotonic neurons, especially during the noise onset before the aAM stimulus was on (as marked by arrow "a" in Fig. 8A), but not the average responses of the highly non-monotonic

neurons. (2) The background noise attenuated the FF values of the highly non-monotonic neurons during the aAM onset, but not those of the monotonic neurons (as marked by arrow "b" in Fig. 8B). It appears that noise onset responses could quench the variability in aAM onset response patterns among non-monotonic neurons without changing the magnitude of their average responses. The contrasting performances of the two neuronal populations were applied to all ten aAM envelopes (Figs. 8C and D).

These data suggest that cortical neurons may encode sound envelope through two complementary coding strategies: using the averaged population responses to encode the amplitude of sound envelope and using the variability in population responses to encode the dynamic transition of sound envelope. Empirically evaluating this hypothesis requires simultaneous recordings of activity of a population of neurons. This technique was not implemented by the current study. As a proof of principle, we pooled data from experiments conducted sequentially and used an accretion process to test the effects of the ensemble size on envelope encoding through a bootstrap procedure (See the Methods). Analyses focused on (1) the correlation between the amplitude of an aAM envelope and the average of ensemble PSTHs (denoted as $CC_{amp-avg}$) and (2) the correlation between the absolute slope values of an aAM envelope and the FF of ensemble PSTHs (denoted as $CC_{slope-FF}$). We proposed that an increase in the stimulus-response correlation with ensemble size would argue for a population-based coding scheme.

Figure 9A shows the results of the correlation metric $CC_{amp-avg}$. The average responses of monotonic neurons (red) encode better envelope amplitude than those of highly non-monotonic neurons (blue). The median increment, $CC_{amp-avg}$(mon)-$CC_{amp-avg}$(non), across the ensemble was 0.13 in the T condition and 0.12 in the T+N condition (*ranksum* test, $p<10^{-5}$). Notably, $CC_{amp-avg}$ values of monotonic neurons were less than those moderately non-monotonic neurons (gray) in both T and T+N condition (*ranksum* test, $p<10^{-5}$). This occurred because encoding envelope amplitude via population average requires not also synchrony between neurons but also reliable envelope responses of individual neurons. Although monotonic neurons were most synchronized among the three groups of neurons (Figs. 7AB), they had lowest envelope feature selectivity (Figs. 6). This weakened their coding capacity relative to those of moderately non-monotonic neurons. Comparing $CC_{amp-avg}$ values between T and T+N conditions, noise greatly reduced the strength of stimulus-response correlation for all three groups of neurons. The median reduction, $CC_{amp-avg}$(T)-$CC_{amp-avg}$(T+N), across the ensemble was 0.24 for monotonic neurons, 0.19 for moderately non-monotonic, and 0.22 for highly non-monotonic neurons.

The opposite performances of neurons were observed in the variability analysis based on $CC_{slope-FF}$ (Fig. 9B). The FF values of highly non-monotonic neurons encoded better envelope transitions than those of monotonic neurons. The median increment, $CC_{slope-FF}$(non)-$CC_{slope-FF}$(mon), across the ensemble was 0.31 in the T condition and 0.25 in the T+N condition (*ranksum* test, $p<10^{-7}$). The differences in $CC_{slope-FF}$ were much reduced between moderately non-monotonic and monotonic neurons. The median increment, $CC_{slope-FF}$(non$_{mod}$)-$CC_{slope-FF}$(mon), across the ensemble was zero in the T condition and 0.04 in the T+N condition. In contrast to results shown in Fig. 9A, noise did not change drastically the stimulus-response correlation $CC_{slope-FF}$ for either neuronal group. The median reduction, $CC_{slope-FF}$(T)-$CC_{slope-FF}$(T+N), was zero for monotonic neurons, -0.04 for moderately non-monotonic neurons, and 0.06 for highly non-monotonic neurons.

The linear regression analysis was further used to examine whether recruiting more neurons improved population coding. Only a few conditions revealed significant size effects. The most sensitive ones were associated with $CC_{slope-FF}$ of the highly non-monotonic neurons in

Fig. 9B, which increased rapidly with ensemble size in the T ($R^2$=0.41) and T+N ($R^2$=0.17) conditions ($p$<0.001). Three conditions associated with $CC_{amp-avg}$ in Fig. 9A also revealed weak but significant size effects - those of the moderately non-monotonic neurons in the T ($R^2$=0.08) and T+N conditions ($R^2$=0.02) and those of monotonic neurons in the T condition ($R^2$=0.1); $p$<0.05. For all other cases, results either remained unchanged or even decreased with ensemble size, especially those associated with $CC_{slope-FF}$ values of moderately non-monotonic and monotonic neurons in Fig. 9B. This indicates that the variability analysis was effective only for a heterogeneous neural population.

Comparing the performance of the two different correlation metrics (i.e., same color lines in Figs. 9A and B for either T or T+N condition), for monotonic and moderately non-monotonic neurons, ensemble average showed a greater predictive power than ensemble variability in both T and T+N conditions. The median difference, $CC_{amp-avg}$(mon)-$CC_{slope-FF}$(mon), across the ensemble was 0.46 in the T condition and 0.22 in the T+N condition (*ranksum* test, $p$<0.001) and that of moderately non-monotonic neurons was 0.52 in the T condition and 0.29 in the T+N condition (*ranksum* test, $p$<0.001). To the contrary, for highly monotonic neurons, the advantage of ensemble variability was more evident, where $CC_{slope-FF}$(non)-$CC_{amp-avg}$(non) across the ensemble was -0.02 in the T condition and 0.15 in the T+N condition (*ranksum* test, $p$<0.01). Together, these results confirmed our initial observations in Fig. 8. Monotonic and highly non-monotonic neurons could use different strategies to encode sound envelope. Notably, their performances differed in background noise - the variability-based metric was more robust and benefited more from a large ensemble size than the average-based metric. For neurons with intermediate RLF non-monotonicity ($non_{mod}$), their performance appeared to favor the average-based metric, suggesting a continuum rather than discrete segregation of envelope coding strategies among A1 neurons.

## Discussion

The main finding of this study is the heterogeneity of selectivity profiles to envelope transitions among neurons in primary auditory cortex and its relationship with static tone rate-level functions. Non-monotonic neurons showed greater envelope feature selectivity than monotonic neurons. Mechanistically, envelope coding based on the variability of ensemble responses is more robust in background noise than that based on the average of ensemble responses.

### Strengths and limitations of using reverse correlation to characterize AM responses

In this study, the envelope feature selectivity of a neuron was identified by reverse correlation. This approach differs from the standard rate- or synchrony-based AM analyses (e.g., Schreiner and Urbas, 1986), in that it captures neural sensitivity to dynamic envelope features associated with the directional change of sound envelope. These features cannot be parameterized by modulation frequency, phase, depth, and sound level - the four parameters that define a periodic envelope such as SAM. Moreover, reverse correlation takes into account the stimulus statistics and results in a stimulus-normalized response probability. In contrast, although modulation depth and ramp/damp times may be more efficient in describing a particular envelope shape (Swarbrick and Whitfield, 1972; Schreiner and Urbas, 1988; Lu et al., 2001; Malone et al., 2007, 2010), these metrics do not provide explicit information on the statistics of various dynamic envelope features in stimuli. For a neuron sensitive to a particular envelope feature type, its average firing rate and temporal response pattern may depend not only on its envelope feature selectivity but also on the density of that feature type in stimuli (Fig. 1). Such a codependence cannot be differentiated by a rate- or synchrony-based response metric. This confound, however, is largely removed by the reverse correlation method. Similar approaches have been used to study the

spectrotemporal receptive fields (STRFs) of cortical neurons (e.g., deCharms et al., 1998; Klein et al., 2000; Miller et al., 2002).

One clear limitation of using the reverse correlation method to characterize AM responses is that its accuracy depends on the temporal precision of neural responses relative to the stimuli. Using the same neurophysiological preparation, previous studies from our laboratory have shown that many neurons in the auditory thalamus and cortex of awake marmoset exhibit significantly driven, non-synchronized firing patterns during the ongoing portions of AM stimuli, suggesting a transformation from temporal to rate representations of modulation frequency (MF) of a sound at higher auditory stages (Wang et al., 2008). If not reliable across trials, the non-synchronized AM responses would yield weak envelope feature selectivity in our analysis, potentially undermining their contributions to AM encoding. However, at low-rate envelope modulation (~4Hz), envelope feature selectivity revealed in this study is primarily associated with a rate representation of local dynamic envelope transitions, such as envelope onset, not MF. These transient response properties are usually present when a neuron is stimulated with effective stimuli. For this reason, only neurons with significant tone-driven responses (at BF) were tested with aAM stimuli and included in the data analysis.

**Implications of diverse envelope feature selectivity for envelope coding**

In the auditory cortex, differential neural responses to onset and offset transients and to ramping and damping profiles of sound envelopes have been reported in both anesthetized (Heil, 1997b, a; Phillips et al., 2002) and awake (Bieser and Muller-Preuss, 1996; Recanzone, 2000; Lu et al., 2001; Malone et al., 2007; Qin et al., 2007) animals. Our results support these findings and further show that the slope preference of a neuron is highly asymmetrical in A1 (Fig. 4). The slope preference has important implications for general sound processing by the auditory system. When a spike is triggered by changes in envelope slope, spike times are insensitive to the absolute sound level due to logarithmic scaling. In contrast, when a spike is trigged by changes in envelope amplitude, changing sound level leads to an advance or delay in spike times (Hopfield, 1995). Since the perceptual quality of sensory stimuli, such as pure-tone frequency, remains mostly unchanged with changes in stimulus amplitude, one would expect a robust, scale-invariant, neural representation to behave accordingly. The slope sensitivity to sound envelope could help preserve the temporal sequences of neural firings in relation to those of other neurons within an ensemble to support stable perception.

In the Results, we show that non-monotonic neurons exhibited greater feature selectivity than monotonic neurons. One parsimonious interpretation is that inhibitory activity is stronger in responses of non-monotonic than monotonic neurons and that inhibition enhances envelope selectivity of a neuron by suppressing its responses to less preferred envelope features in quiet (Figs. 3C) and in background noise (Figs. 6DEF). This conjecture is in part supported by recent extracellular results from the auditory cortex of awake marmoset (Sadagopan and Wang, 2010), showing that on-BF inhibition affects the stimulus selectivity of a neuron in frequency and time. Moreover, intracellular studies in the auditory cortex of anesthetized rats have shown that synaptic excitation and inhibition appear to be matched in their frequency tuning, but are less congruent in their timing and level-dependent strength (Wehr and Zador, 2003; Tan et al., 2004; Wu et al., 2006; Tan et al., 2007). The latter distinction might contribute to diverse envelope feature selectivity among A1 neurons reported here. For future studies that aim to further distinguish the functional neuronal groups in auditory cortex, the anesthesia effects need to be controlled. It has been shown that anesthesia can alter the onset and offset response patterns depending on anesthetic conditions (Zurita et al., 1994; Ter-Mikaelian et al., 2007) and reduce the sustained patterns of neural responses related to general sound processing in auditory cortex (deCharms and

Merzenich, 1996; Wang et al., 2005). Importantly, the onset preferences of monotonic neurons reported here differ from the transient onset responses of monotonic neurons found in the anesthetized cortex (Phillips et al., 2002). Monotonic neurons in the awake condition were not silent to less preferred envelope features at the single-neuron (Fig. 3C) and population levels (Fig. 8C).

### Population encoding of low-frequency sound envelope

Using multielectrode recording, recent studies have shown that responses of neurons in the primary visual cortex are actively decorrelated in the awake condition (Ecker et al., 2010), and that neural decorrelation is affected by local activation of muscarinic acetylcholine receptors after nucleus basalis stimulation (Goard and Dan, 2009). In this study, we observed that decorrelated neural activity was associated with distinctive envelope feature selectivity of neurons with highly non-monotonic RLFs in the A1 of awake marmoset (Fig. 7). This finding suggests that cortical inhibition may play important roles in modulating the level of decorrelation between neurons to enhance sensory coding. In general terms, neurons showing asynchronous spike patterns convey more information about sensory stimuli (Reich et al., 2001; Kayser et al., 2009). Functionally, when sensory information can be transmitted at different time points by different neurons, a coding system has a broader bandwidth, resembling time-division multiplexing in communication theory (Cariani, 1995). The asynchronous responses of non-monotonic neurons shown here might provide a neural substrate for such a time-sharing process undertaken at the level of auditory cortex.

Our analysis further showed that this seemingly less "coordinated" neural assembly has a powerful capacity to encode envelope transitions using the variability in their ensemble responses, especially in background noise (Fig. 9B). It remains to be shown what types of synaptic, cellular, and circuit mechanisms might utilize the variability/asynchrony information in ensemble neuronal responses to encode sound envelope. One testable hypothesis is that a downstream neuron that encodes the boundary information of sound receives phasic synaptic inputs that are time-locked to sound envelope transitions (as a result of convergence of asynchronous inputs from neurons with distinctive envelope feature selectivity). The assumption that implicitly underlies this hypothesis is that the sign of envelope slope is encoded by the weights of synaptic inputs of projection neurons. Conversely, a downstream neuron that encodes envelope amplitude may receive tonic synaptic inputs throughout sound stimulation (as a result of convergence of synchronous inputs from neurons with similar envelope feature selectivity). The differences between these two types of neurons could be evaluated by the dynamics of membrane potentials in relation to sound envelope using in vivo intracellular recordings.

We argue that at higher stages of the auditory system, a major goal of low-frequency sound envelope processing is to extract temporal boundary/segmental information of envelope waveform, as opposed to an isomorphic representation of envelope shape as seen in the peripheral auditory system (Joris and Yin, 1992). Such a transformation might start at loci earlier than auditory cortex, such as inferior colliculus (IC). The IC response to speech utterance is more phasic than that of auditory nerve (AN) and cochlear nucleus (CN) and speech reconstruction is less satisfactory using IC responses than those of AN and CN (Delgutte et al., 1998). It remains to be tested to what extent the distinction between monotonic and non-monotonic neurons reported here (1) reflects the inherited sound processing prior to auditory cortex and (2) applies to broadband sounds such as speech. At present, it is conceivable that monotonic and non-monotonic neurons in A1 might yield different types of information about sound envelope for downstream processing.

## Acknowledgments

## Reference

Bieser A, Muller-Preuss P. Auditory responsive cortex in the squirrel monkey: neural responses to amplitude-modulated sounds. Exp Brain Res. 1996; 108:273–284. [PubMed: 8815035]

Bregman, AS. Auditory scene analysis: the perceptual organization of sound. MIT Press; Cambridge, Massachusetts: 1990.

Cariani P. As if time really mattered: temporal strategies for neural coding of sensory information. Commun Cogn Artif Intell. 1995; 12:161–229.

de Boer E, de Jongh HR. On cochlear encoding: potentialities and limitations of the reverse-correlation technique. J Acoust Soc Am. 1978; 63:115–135. [PubMed: 632404]

deCharms RC, Merzenich MM. Primary cortical representation of sounds by the coordination of action-potential timing. Nature. 1996; 381:610–613. [PubMed: 8637597]

deCharms RC, Blake DT, Merzenich MM. Optimizing sound features for cortical neurons. Science. 1998; 280:1439–1443. [PubMed: 9603734]

Delgutte, B.; Hammond, BM.; Cariani, PA. Neural coding of the temporal envelope of speech: Relation to modulation transfer functions. In: Palmer, AR.; Reese, A.; Summerfield, AQ.; Meddis, R., editors. Psychophysical and Physiological Advances in Hearing. Whurr Publisher Ltd; 1998. p. 595-603.

Ecker AS, Berens P, Keliris GA, Bethge M, Logothetis NK, Tolias AS. Decorrelated neuronal firing in cortical microcircuits. Science. 2010; 327:584–587. [PubMed: 20110506]

Eggermont JJ. Temporal modulation transfer functions for AM and FM stimuli in cat auditory cortex. Effects of carrier type, modulating waveform and intensity. Hear Res. 1994; 74:51–66. [PubMed: 8040099]

Fishbach A, Nelken I, Yeshurun Y. Auditory edge detection: a neural model for physiological and psychoacoustical responses to amplitude transients. J Neurophysiol. 2001; 85:2303–2323. [PubMed: 11387378]

Frisina RD. Subcortical neural coding mechanisms for auditory temporal processing. Hear Res. 2001; 158:1–27. [PubMed: 11506933]

Ghazanfar AA, Hauser MD. The auditory behaviour of primates: a neuroethological perspective. Curr Opin Neurobiol. 2001; 11:712–720. [PubMed: 11741023]

Goard M, Dan Y. Basal forebrain activation enhances cortical coding of natural scenes. Nat Neurosci. 2009; 12:1444–1449. [PubMed: 19801988]

Hartigan JA, Hartigan PM. The Dip Test of Unimodality. The Annals of Statistics. 1985; 13:70–84.

Heil P. Auditory cortical onset responses revisited. I. First-spike timing. J Neurophysiol. 1997a; 77:2616–2641. [PubMed: 9163380]

Heil P. Auditory cortical onset responses revisited. II. Response strength. J Neurophysiol. 1997b; 77:2642–2660. [PubMed: 9163381]

Heil P, Irvine DR. The posterior field P of cat auditory cortex: coding of envelope transients. Cereb Cortex. 1998; 8:125–141. [PubMed: 9542892]

Hopfield JJ. Pattern recognition computation using action potential timing for stimulus representation. Nature. 1995; 376:33–36. [PubMed: 7596429]

Houtgast T, Steeneken HJM. A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. J Acoust Soc Am. 1985; 77:1069–1077.

Joris PX, Yin TC. Responses to amplitude-modulated tones in the auditory nerve of the cat. J Acoust Soc Am. 1992; 91:215–232. [PubMed: 1737873]

Joris PX, Schreiner CE, Rees A. Neural processing of amplitude-modulated sounds. Physiol Rev. 2004; 84:541–577. [PubMed: 15044682]

Kayser C, Montemurro MA, Logothetis NK, Panzeri S. Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. Neuron. 2009; 61:597–608. [PubMed: 19249279]

Klein DJ, Depireux DA, Simon JZ, Shamma SA. Robust spectrotemporal reverse correlation for the auditory system: optimizing stimulus design. J Comput Neurosci. 2000; 9:85–111. [PubMed: 10946994]

Kuwada S, Batra R. Coding of sound envelopes by inhibitory rebound in neurons of the superior olivary complex in the unanesthetized rabbit. J Neurosci. 1999; 19:2273–2287. [PubMed: 10066278]

Langner G. Periodicity coding in the auditory system. Hear Res. 1992; 60:115–142. [PubMed: 1639723]

Liang L, Lu T, Wang X. Neural representations of sinusoidal amplitude and frequency modulations in the primary auditory cortex of awake primates. J Neurophysiol. 2002; 87:2237–2261. [PubMed: 11976364]

Lu T, Liang L, Wang X. Neural representations of temporally asymmetric stimuli in the auditory cortex of awake primates. J Neurophysiol. 2001; 85:2364–2380. [PubMed: 11387383]

Malone BJ, Scott BH, Semple MN. Dynamic amplitude coding in the auditory cortex of awake rhesus macaques. J Neurophysiol. 2007; 98:1451–1474. [PubMed: 17615123]

Malone BJ, Scott BH, Semple MN. Temporal codes for amplitude contrast in auditory cortex. J Neurosci. 2010; 30:767–784. [PubMed: 20071542]

Miller LM, Escabi MA, Read HL, Schreiner CE. Spectrotemporal receptive fields in the lemniscal auditory thalamus and cortex. J Neurophysiol. 2002; 87:516–527. [PubMed: 11784767]

Moshitch D, Las L, Ulanovsky N, Bar-Yosef O, Nelken I. Responses of neurons in primary auditory cortex (A1) to pure tones in the halothane-anesthetized cat. J Neurophysiol. 2006; 95:3756–3769. [PubMed: 16554513]

Neuert V, Pressnitzer D, Patterson RD, Winter IM. The responses of single units in the inferior colliculus of the guinea pig to damped and ramped sinusoids. Hear Res. 2001; 159:36–52. [PubMed: 11520633]

Olive, JP.; Greenwood, A.; Coleman, J. Acoustics of American English Speech. A Dynamic Approach. Springer-Verlag; New York: 1993.

Pfingst BE, O'Connor TA. Characteristics of neurons in auditory cortex of monkeys performing a simple auditory task. J Neurophysiol. 1981; 45:16–34. [PubMed: 7205342]

Phillips DP, Hall SE. Responses of single neurons in cat auditory cortex to timevarying stimuli: linear amplitude modulations. Exp Brain Res. 1987; 67:479–492. [PubMed: 3653310]

Phillips DP, Hall SE, Boehnke SE. Central auditory onset responses, and temporal asymmetries in auditory perception. Hear Res. 2002; 167:192–205. [PubMed: 12117542]

Pressnitzer D, Winter IM, Patterson RD. The responses of single units in the ventral cochlear nucleus of the guinea pig to damped and ramped sinusoids. Hear Res. 2000; 149:155–166. [PubMed: 11033255]

Qin L, Chimoto S, Sakai M, Wang J, Sato Y. Comparison between offset and onset responses of primary auditory cortex ON-OFF neurons in awake cats. J Neurophysiol. 2007; 97:3421–3431. [PubMed: 17360820]

Recanzone GH. Response profiles of auditory cortical neurons to tones and noise in behaving macaque monkeys. Hear Res. 2000; 150:104–118. [PubMed: 11077196]

Reich DS, Mechler F, Victor JD. Independent and redundant information in nearby cortical neurons. Science. 2001; 294:2566–2568. [PubMed: 11752580]

Rose G. A temporal-processing mechanism for all species? Brain Behav Evol. 1986; 28:134–144. [PubMed: 3567537]

Rosen S. Temporal information in speech: acoustic, auditory and linguistic aspects. Philos Trans R Soc Lond B Biol Sci. 1992; 336:367–373. [PubMed: 1354376]

Sadagopan S, Wang X. Level invariant representation of sounds by populations of neurons in primary auditory cortex. J Neurosci. 2008; 28:3415–3426. [PubMed: 18367608]

Sadagopan S, Wang X. Contribution of inhibition to stimulus selectivity in primary auditory cortex of awake primates. J Neurosci. 2010; 30:7314–7325. [PubMed: 20505098]

Schreiner CE, Urbas JV. Representation of amplitude modulation in the auditory cortex of the cat. I. The anterior auditory field (AAF). Hear Res. 1986; 21:227–241. [PubMed: 3013823]

Schreiner CE, Urbas JV. Representation of amplitude modulation in the auditory cortex of the cat. II. Comparison between cortical fields. Hear Res. 1988; 32:49–63. [PubMed: 3350774]

Shaddock Palombi P, Backoff PM, Caspary DM. Responses of young and aged rat inferior colliculus neurons to sinusoidally amplitude modulated stimuli. Hear Res. 2001; 153:174–180. [PubMed: 11223307]

Sutter ML, Schreiner CE. Topography of intensity tuning in cat primary auditory cortex: single-neuron versus multiple-neuron recordings. J Neurophysiol. 1995; 73:190–204. [PubMed: 7714564]

Swarbrick L, Whitfield IC. Auditory cortical units selectively responsive to stimulus 'shape'. J Physiol. 1972; 224:68P–69P.

Tan AY, Zhang LI, Merzenich MM, Schreiner CE. Tone-evoked excitatory and inhibitory synaptic conductances of primary auditory cortex neurons. J Neurophysiol. 2004; 92:630–643. [PubMed: 14999047]

Tan AY, Atencio CA, Polley DB, Merzenich MM, Schreiner CE. Unbalanced synaptic inhibition can create intensity-tuned auditory cortex neurons. Neuroscience. 2007; 146:449–462. [PubMed: 17320296]

Ter-Mikaelian M, Sanes DH, Semple MN. Transformation of temporal properties between auditory midbrain and cortex in the awake Mongolian gerbil. J Neurosci. 2007; 27:6091–6102. [PubMed: 17553982]

Wang X, Merzenich MM, Beitel R, Schreiner CE. Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. J Neurophysiol. 1995; 74:2685–2706. [PubMed: 8747224]

Wang X, Lu T, Snider RK, Liang L. Sustained firing in auditory cortex evoked by preferred stimuli. Nature. 2005; 435:341–346. [PubMed: 15902257]

Wang X, Lu T, Bendor D, Bartlett E. Neural coding of temporal information in auditory thalamus and cortex. Neuroscience. 2008; 157:484–494. [PubMed: 19143093]

Watkins PV, Barbour DL. Specialized neuronal adaptation for preserving input sensitivity. Nat Neurosci. 2008; 11:1259–1261. [PubMed: 18820690]

Wehr M, Zador AM. Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. Nature. 2003; 426:442–446. [PubMed: 14647382]

Wu GK, Li P, Tao HW, Zhang LI. Nonmonotonic synaptic excitation and imbalanced inhibition underlying cortical intensity tuning. Neuron. 2006; 52:705–715. [PubMed: 17114053]

Zurita P, Villa AE, de Ribaupierre Y, de Ribaupierre F, Rouiller EM. Changes of single unit activity in the cat's auditory thalamus and cortex associated to different anesthetic conditions. Neurosci Res. 1994; 19:303–316. [PubMed: 8058206]
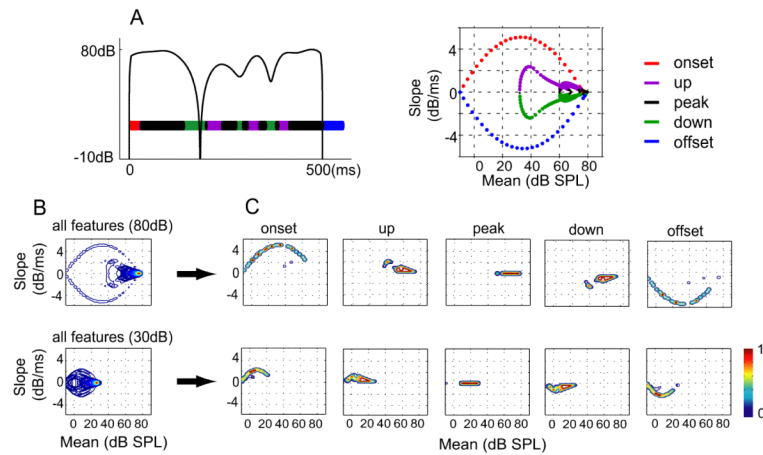
**Figure 1. Envelope features of the aAM stimuli**
(A) Envelope of an aAM stimulus with a peak level of 80dB SPL (left) and its mean-slope trajectory (right). Envelope segments with positive, zero, and negative slope values and their associated mean-slope trajectories are plotted in different colors. Each color dot on the mean-slope trajectory marks the mean-slope values of one 25-ms envelope segment. (B) Contour plots of the mean-slope distributions of total envelope features in ten aAM stimuli played at 80dB and 30dB SPL. (C) Contour plots of the mean-slope distributions of five envelope feature types at 80dB and 30dB SPL. The color of a contour line indicates the peak-normalized magnitude of a distribution.
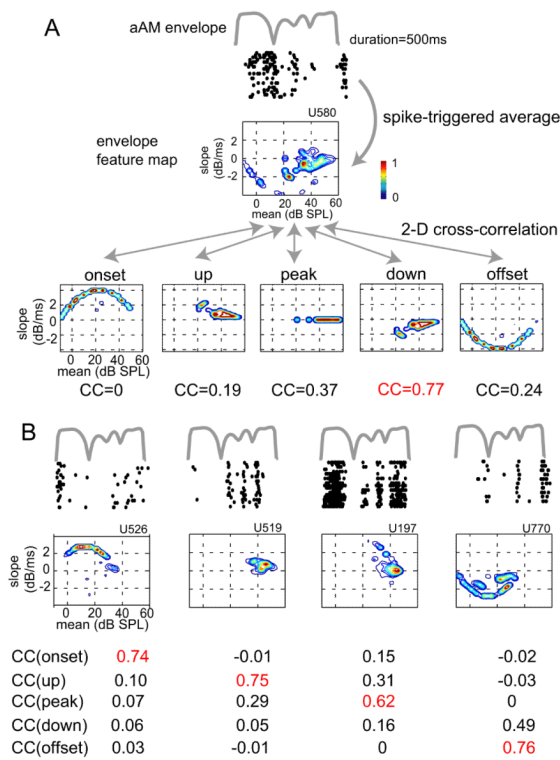
**Figure 2. Characterization of envelope feature selectivity of A1 neurons**
(A) Shown from top to bottom are the envelope of an aAM stimulus at 60dB SPL, the raster plot of corresponding responses of an example neuron, its envelope feature map, and the mean-slope distributions of five feature types of aAM stimuli at 60dB SPL. The preferred feature type of a neuron was determined by the strength of the 2-D correlation between the envelope feature map of a neuron and the mean-slope distributions of five envelope feature types. As highlighted in red, the "down" type yielded the highest correlation coefficient (CC), and was then designated as the preferred envelope feature of this neuron. (B) Raster plots and envelope feature maps of four other example neurons. Based on the maximal strengths of correlation values with five envelope feature types (red), the preferred envelope feature types for the four example neurons from left to right are "onset", "up", "peak", and "offset", respectively. The peak SPLs used were 40dB SPL for the "onset" and "offset" neurons, 60dB SPL for the "up" neuron, and 50dB SPL for the "peak" neuron. The Feature Preference Index (FPI, defined in Methods) of the five neurons was in turn 0.63, 0.56, 0.66, 0.63, and 0.63. The aAM waveform lasted 500 ms and neural responses between 15 ms after the stimulus onset and 65ms after the stimuli offset were analyzed. The raster plots were advanced 15 ms to visually align with the envelope of the aAM stimulus.
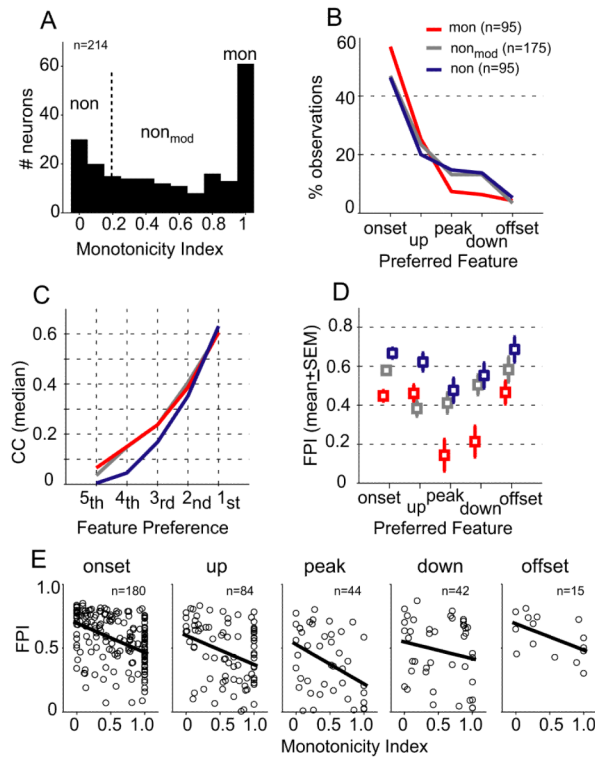
**Figure 3. Comparison of envelope feature selectivity between monotonic and non-monotonic neurons**

(A) Distribution of Monotonicity Index (MI). For data analyses, the neural population was divided into three subgroups based on MI values (Table I). (B) Percentage of observations of feature types for the three neural groups. Neurons that were tested at more than one SPL contributed multiple observations. (C) Ordered correlation coefficients (median) between stimulus-response feature maps as demonstrated in Fig. 2. The highest rank (1st) is associated with the preferred feature type of a neuron. (D) Feature Preference Index (FPI, mean ±SEM) of monotonic (red), moderately non-monotonic (gray) and highly non-monotonic neurons (blue) for each feature type. (E) Scatter plots of FPI as a function of MI for five feature types. Best-fit linear regression is shown in black line.
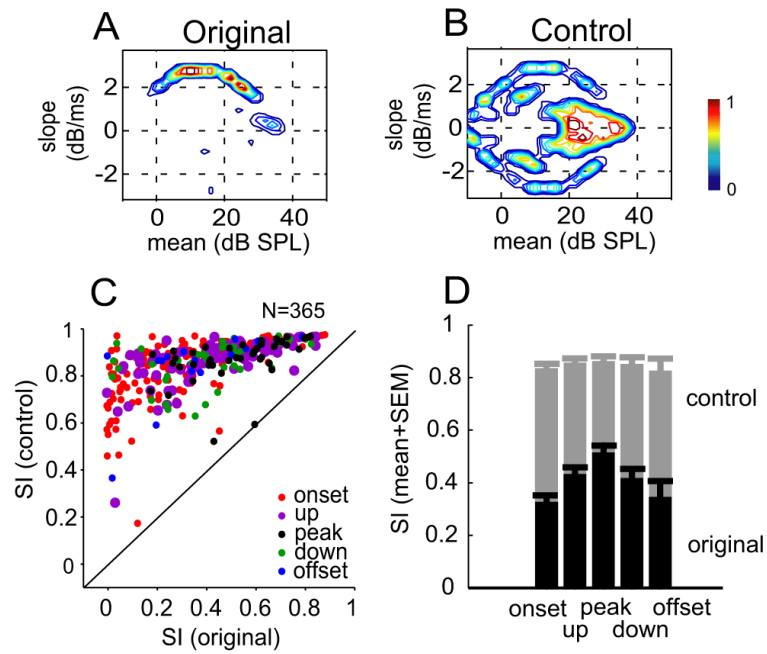
**Figure 4. Envelope slope sensitivity of neurons in auditory cortex**
Original (A) and control (B) feature maps of an example neuron. Highly asymmetric activity was observed in the original map, but not in the control obtained from shuffled spike times. (C) Scatter plot of Symmetry Index (SI) in the original (abscissa) and in the control map (ordinate) for all neurons tested. Different preferred envelope feature types are labeled in different colors. (D) SI values (mean+SEM) of the original (black) and control (gray) feature maps for each feature type.
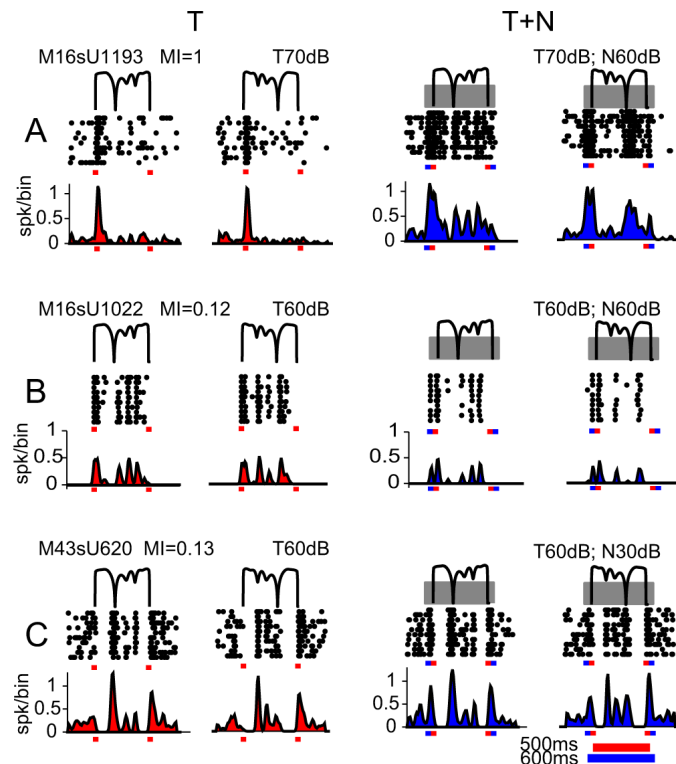
**Figure 5. Spike-timing patterns to aAM envelope in background noise**

(A)-(C) Raster plots and PSTHs of responses of three neurons in T and T+N conditions. The envelopes of two aAM stimuli (forward and time-reversed) are shown in thin black lines above the raster plots. The MI value of a neuron and the SPLs of aAM and noise are indicated above the stimuli. To provide a visible reference for the noise responses, we intentionally chose these neurons showing excitatory responses to the noise onset. In the T +N condition, the broadband noise (depicted as gray blocks overlaying the aAM stimuli) was gated on 50 ms before the aAM onset and gated off 50 ms after the aAM offset. The onset and offset times are indicated by tick marks on each panel (aAM in red and noise in blue). Each aAM waveform lasted 500 ms and noise lasted 600 ms. The binwidth used for PSTHs was 10 ms.
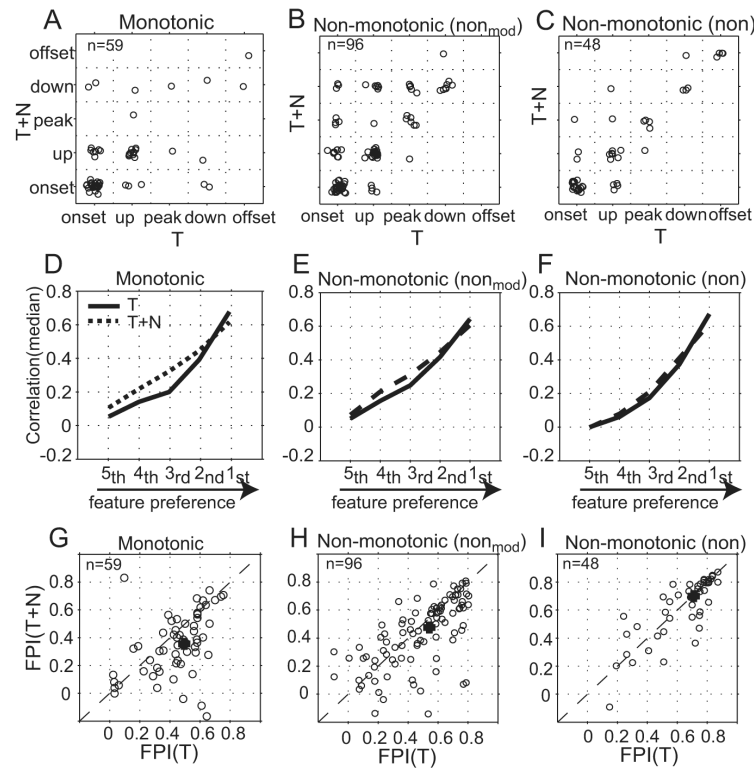
**Figure 6. Comparison of envelope feature selectivity in background noise**
(A)-(C) Scatter-plot of the preferred feature types of monotonic, moderately non-monotonic and highly non-monotonic neurons tested in both T and T+N conditions. Each circle designates the preferred envelope types of one neuron with and without background noise in response to the same aAM stimuli. The median noise level was 40dB SPL for all three neuronal populations. For the purpose of plotting, the five feature types were numbered from 1 to 5. For example, the onset-onset transition was positioned (1,1) on the x-y plane. To visually separate the circles within each subgrid, we jittered the position of each circle by adding a small Gaussian random number (mean=0 and SD=0.1) to its numerical feature type. (D)-(F) Correlation coefficient (median) between stimulus-response feature maps in T and T+N conditions. Data are presented in the same format as those in Fig. 3C. (G)-(I) Scatter-plot of FPI values between T and T+N conditions for individual neurons. The median FPI value is marked by a cross sign on each panel.
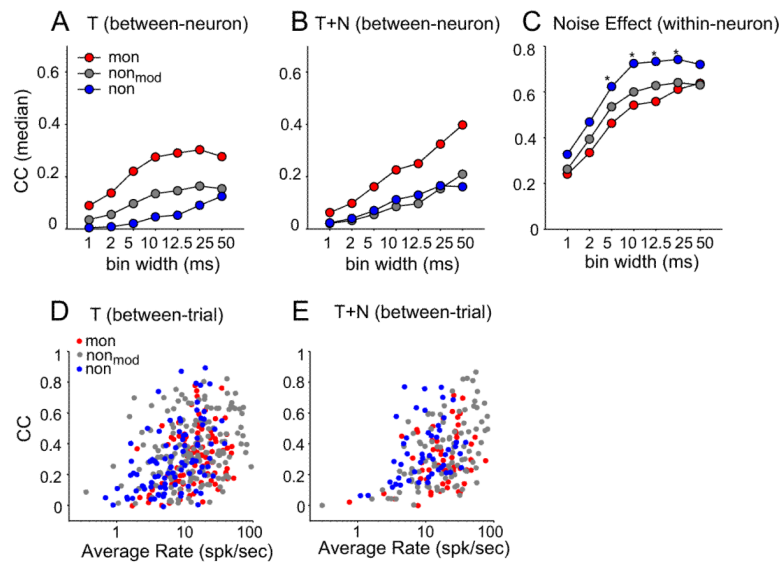
**Figure 7. Correlation analysis of spike-timing patterns in response to aAM stimuli**
(A)(B) The Pearson correlation coefficient (CC, median) of the between-neuron comparison in T and T+N conditions. (C) The Pearson correlation coefficient (CC, median) of the within-neuron comparison between T and T+N conditions. In A-C, the spike times were binned at multiple resolutions ranging from 1 to 50 ms. Due to the temporal smoothing effect of a spike counting window, the correlation computed from PSTHs increased with the binwidth. (D)(E) The Pearson correlation coefficient (CC) of the between-trial comparison as a function of average rate in T and T+N conditions. Each circle represents data from one neuron, averaged over CCs from all pair-wise correlations between different trials of responses. The PSTH binwidth was 10ms. The total numbers of PSTHs (i.e., test conditions) used in T and T+N conditions are listed in Table I.
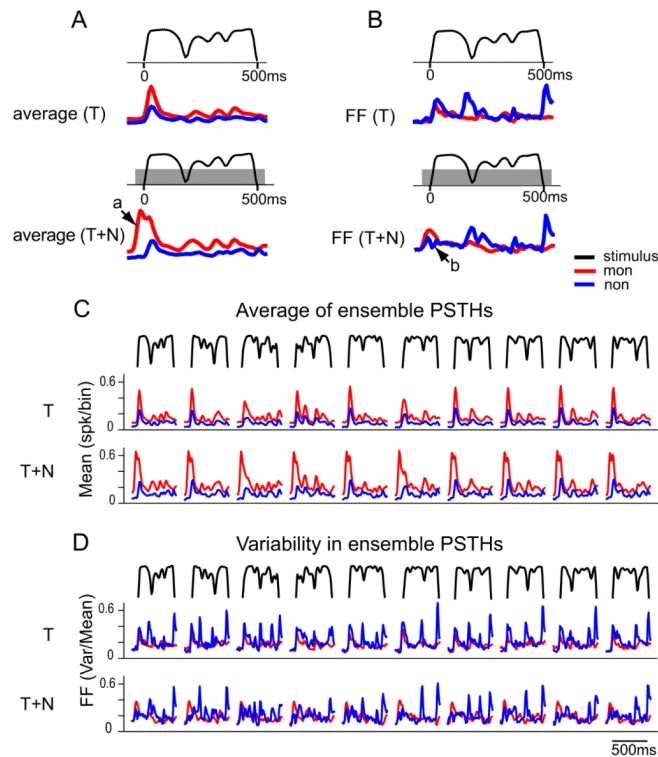
**Figure 8. aAM envelope representation by a population of neurons**
(A) Average and (B) FF of ensemble PSTHs for monotonic (red) and highly non-monotonic neurons (blue) in response to an aAM envelope in T and T+N conditions. (C) Average and (D) FF of ensemble PSTHs to the ten aAM envelopes in T and T+N conditions. The aAM duration was 500 ms and noise duration was 600 ms. The PSTHs show responses occurring from 100 ms before the aAM onset to 50 ms after the aAM offset (spike times were advanced 15 ms in data analyses). The PSTH binwidth was 10 ms. The ensemble sizes in T and T+N conditions were the same as those in Fig. 7.
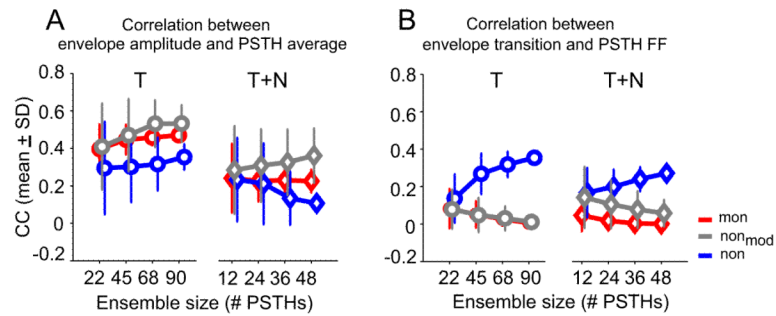
**Figure 9. Representation of envelope amplitude and transition by a population of neurons**
(A) Correlation coefficient (mean±SD) between envelope amplitude and the average of ensemble PSTHs in T and T+N conditions. (B) Correlation coefficient (mean±SD) between envelope transition and the FF of ensemble PSTHs in T and T+N conditions. The envelope transitions were described by absolute values of the time derivative of aAM amplitude with a 1-ms time resolution. Individual PSTHs, envelope amplitude, and envelope transition were smoothed using a 10-ms time window prior to the correlation analysis. A bootstrap method was used to take repeated samples from the total PSTHs to form an ensemble. The ensemble size had roughly an incremental step of a quarter of the total PSTHs collected at each condition for monotonic and highly non-monotonic neurons. The total numbers of PSTHs (i.e., test conditions) for each neuronal group were the same as those in Figs. 7 and 8 and are listed in Table I.

## Table I

The numbers of test conditions and neurons for monotonic ("mon"), moderately non-monotonic ("$non_{mod}$") and highly non-monotonic ("non") groups in response to aAM stimuli alone ("T" condition) and in background noise ("T+N" condition).

| neural groups | T | | T+N | |
|---|---|---|---|---|
| | # condition | # neuron | # condition | # neuron |
| mon (MI=1) | 95 | 58 | 59 | 29 |
| $non_{mod}$ (0.2<MI<1) | 175 | 97 | 96 | 50 |
| non (MI  0.2) | 95 | 59 | 48 | 23 |
| total | 365 | 214 | 203 | 102 |