

**SPECTRAL INTEGRATION AND NEURAL REPRESENTATION OF HARMONIC
COMPLEX TONES IN PRIMATE AUDITORY CORTEX**

By
Lei Feng

A dissertation submitted to the Johns Hopkins University in conformity with the
requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

November 2013

© 2013 Lei Feng
All Rights Reserved

Abstract

Many natural and man-made sounds, such as animal vocalizations, human speech, and sounds from many musical instruments contain rich harmonic structures. Although the peripheral auditory system decomposes these sounds into separate frequency channels, harmonically related frequency components must be grouped together in order to form a single auditory percept. A central neural process is therefore required to accomplish this perceptual grouping and to integrate information across frequency channels in order to compute spectral properties, such as pitch and timber, which are not explicitly encoded in the auditory periphery. In this dissertation, I investigated whether there are representations of harmonic structures at the single neuron level in auditory cortex beyond pitch and how harmonic sounds are represented by populations of cortical neurons. I systematically tested single neurons in the primary auditory cortex (A1) of awake marmoset monkeys with harmonic and inharmonic complex tones, varying fundamental frequency (f_0) and harmonic composition. I found harmonic template neurons, which were strongly driven by harmonic complex tones but showed weak or no response to single harmonics. Harmonic template neurons were selective to f_0 s and sensitive to harmonic numbers. They also exhibited a reduced firing rate in response to inharmonic complex tones. Other sound features of a harmonic complex tone, such as overall sound level, resolved individual harmonic partials, and temporal envelope were represented by different subpopulations of neurons in A1. Overall, the findings of this dissertation support the existence of a distributed neural code for harmonic complex tones in A1 which represents an important stage in the auditory pathway for robust feature extraction and sound source recognition.

In the study of spectral integration and neural coding of complex tones, searching for preferred stimuli of cortical neurons has also proven challenging because of the high dimensionality of the acoustic space of possible stimuli and limited recording time. In the last part of this dissertation, I presented an online adaptive stimulus design approach based on a neural

network model for studying spectral integration in auditory cortex. The models estimated online helped to build a connection between receptive field structures and diverse spectral selectivity of cortical neurons.

Readers: Xiaoqin Wang (Advisor), Eric Young

Acknowledgements

This dissertation is a culmination of six years of hard work, starting from experiment design, days of recordings, to the data analysis and writing. It has been a difficult yet rewarding long journey for me. I owe tremendous thanks to those who led me down the path of scientific inquiry or supported me all the way.

I was introduced to the research of neuroscience by my master advisor, Bo Hong. It opened a door for me to a completely new world as a newly trained biomedical engineering student. My thesis advisor, Xiaoqin Wang, is a true scientist and scholar. The short-term course taught by Xiaoqin on system neuroscience in Tsinghua many years ago inspired me and encouraged me to come to the States to pursue my Ph.D. in neuroscience related research where I could put my engineering and mathematical trainings into good use. He gives us freedom to try what we want to do, but always provide guidance and support so that we won't get discouraged by our failures. Xiaoqin has taught me how to be a scientist through his wisdom, passion and patience.

My thesis committee- Eric Young, Kechen Zhang, Steven Hsiao and Ed Connor have been very supportive for my academic development as a graduate student, from my Graduate Board Oral Exam (GBO), thesis proposal and thesis defense. Eric Young and Kechen Zhang have been especially helpful in advising me in my research. They always generously and patiently offer their advice for my data analysis and theoretical methods. Ed Connor and Steven Hsiao offered unique insights from outside the field which make me think further.

During the six years in lab, I have had many fruitful discussions and advice from all members of Wang lab. Yi Zhou, has been a terrific colleague and friend to me. Yi patiently taught me how to handle and take care of monkeys and all the details of recording. The trainings that I got from her were the foundation of the success in my own project. She was very generous to

share with me her ideas and experience, also provided lots of insightful comments and discussions for my research. Cory and I had many discussions when I first joined the lab, which helped me to start thinking about the ‘big picture’ of my research. Darik and I made a good team. We set up the new system in Chamber 2 together and were partners for many surgeries. He is a terrific engineer and I have learnt a lot from working with him. I especially would like to thank Mike, Luke and Evan for helping reading my writings and correcting my English. Also I owe many thanks to all other lab members, Kai, Chia-Jung, Lixia, Xindong, Lingyun, Seth, Shanequa, Nate, Jenny, Roy. They were always generous to offer their help when I needed any.

In the last but not the least, I would like to thank my parents and my husband Jiarong Hong for their understanding and unconditional love.

Table of Contents

1. Introduction	1
1.1 Harmonic sounds, perception and auditory system.....	1
1.2 Information transformation in the sensory system.....	3
1.3 An overview of methods for studying frequency selectivity in auditory system.....	5
1.4 Objectives of the dissertation.....	6
2. Methods and experimental design	8
2.1 General experimental procedure	8
2.2 Acoustic Stimuli.....	9
2.2.1 General acoustic stimuli.....	9
2.2.2 Harmonic complex tone and inharmonic complex tones.....	9
2.2.3 Random harmonic stimulus (RHS).....	10
2.3 Data Analysis.....	11
2.3.1 Characterize neural responses to HCTs	11
2.3.3 Neural correlations.....	13
2.3.4, Identification of cortical areas and layers	13
2.3.5 Linear weighting function estimation	14
2.4 Online adaptive experiment design (OED).....	15
2.4.1 A feed-forward neural network model.....	15
2.4.2 Offline Data fitting.....	16
2.4.3 Experimental system design for OED.....	17
3. Spectral templates for encoding harmonic structure in marmoset auditory cortex ..	25
3.1 Introduction.....	25
3.2 Results.....	27
3.2.1 Neurons showed combination sensitivity to harmonic complex tones	27
3.2.2 Harmonic template neurons process spectral information	28
3.2.3 The relationship between BF and preferred f0 (Bf0).....	29
3.2.4 Harmonic template neurons are sensitive to spectral regularity	31
3.2.5 Inhibition plays a role in shaping harmonic templates in A1.....	33
3.2.6 Random harmonic stimuli for studying harmonic template neurons	33
3.3 Discussion.....	35

3.3.1 Compare with previous studies	36
3.3.2 Information transformation and feature detection.....	36
3.3.3 Harmonic template neurons and pitch neurons	37
3.3.4 Harmonic resolvability and pitch computation	38
3.3.5 Implications in auditory perception	39
4. A distributed harmonic process in primary auditory cortex	60
4.1 Introduction.....	60
4.1.1 Spectral analysis and perception	60
4.1.2 Neural representation of harmonic complex tones at auditory nerves	61
4.1.3 Frequency receptive field of neurons in auditory cortex	62
4.1.4 The functional organization of auditory cortex for spectral processing.....	62
4.2 Results.....	63
4.2.1 Frequency selective neurons in A1	63
4.2.2 Band-pass neurons	65
4.2.3 Modulation sensitive neurons	66
4.2.4 Harmonic template neurons	67
4.2.5 A level invariant representation of spectral information.....	68
4.2.6 Spectral information and Resolvability.....	69
4.2.7 The relationship between other properties and response pattern to harmonic complex tones	69
4.2.8 Heterogeneity of spectral selectivity in the large scale tonotopic map	70
4.3 Discussion	71
4.3.1 A distributed harmonic process in A1.....	71
4.3.2 Harmonic resolvability in auditory cortex	72
4.3.3 Comparison between the peripheral auditory system and the central auditory system.....	73
4.3.4 Parallel spectral analysis in A1	74
4.3.5 Harmonic process and comparative aspects of pitch perception	75
5. Online adaptive stimulus design for studying spectral integration in auditory cortex.....	93
5.1 Introduction.....	93
5.2 Results.....	96
5.2.1, A generalized feedforward model for studying spectral integration in auditory cortex	96
5.2.2, Application of OED experiments.....	98

5.2.3 Finding the “optimal stimuli”	98
5.2.4 Receptive field structure underlying the spectral selectivity	100
5.2.5 Evaluation of OED.....	101
5.3 Discussion	102
5.3.1 Summary of experiments	102
5.3.2 Co-tuned excitatory and inhibitory inputs	103
5.3.3 Implication for spectral integration in auditory system	104
5.3.4 Network models used in studying neural response properties in A1	105
6. Conclusion.....	121
6.1 The existence of harmonic template neurons in A1	121
6.1.1 Harmonic template neurons	121
6.1.2 Comparison between harmonic template neurons and pitch selective neurons	122
6.1.3 Harmonic template neurons and harmonic sensitive neurons in bats	123
6.2 A distributed harmonic process in A1.....	123
6.2.1 A parallel process of harmonic complex tones	123
6.2.2 Inhibition and a level-invariant representation of spectral information	124
6.2.3. Functional structures for spectral processing in A1	125
6.3 An online adaptive method for studying spectral integration in A1	125
6.3.1 Finding the ‘optimal’ stimuli	125
6.3.2 Frequency receptive field structure underlying the spectral selectivity	127
6.4 Methodological considerations for studying auditory cortex.....	127
7. Miscellaneous findings	129
7.1 Temporal response pattern	129
7.2 Auditory grouping.....	131
REFERENCES.....	136
Curriculum Vitae	150

List of Tables

Table 5.1 The predication of the model to training and validation set	108
Table 5.2 Comparison between online model and offline model	109

List of Figures

Figure 2.1 Demonstration of Jittered complex tones	20
Figure 2.2 An example of a RHS set	21
Figure 2.3 An example of LFPs and CSD of one recording track	22
Figure 2.4 A feed-forward network model with three excitatory subunits (red) and two inhibitory subunits (blue).....	23
Figure 2.5 A diagram of the online adaptive experiment system	24
Figure 3.1 An examples of a neuron which preferred to harmonic complex tones.	41
Figure 3.2 Another example of harmonic template neuron.	42
Figure 3 Harmonic template neurons showed periodic response patterns to shifted complex tones.	43
Figure 3.4 Normalized responses of 35 different harmonic template neurons to the shifted complex tones.	44
Figure 3.5 Two criteria for harmonic template neurons.	45
Figure 3.6 The relationship between BF and the preferred f_0 for harmonic template neurons.	46
Figure 3.7 The different responses of two neurons to the same harmonic complex tones of the same f_0 but different compositions.....	47
Figure 3.8 Harmonic template neurons' responses to sAM tones	48
Figure 3.9 Distribution of harmonic template neurons.	49
Figure 3.11 Distribution of harmonic template neurons in auditory cortex.....	50
Figure 3.11 A harmonic template neuron's responses to jittered complex tones.	51
Figure 3.12 The normalized firing rate distributions for all jitter level of 29 harmonic template neurons.....	52
Figure 3.13 Examples of a harmonic template neuron and a non-harmonic template neuron.....	53
Figure 3.13 Compressed and stretched complex tones test.....	54

Figure 3.14 The role of inhibition in shaping harmonic templates.....	55
Figure 3.16 RHS and linear weights estimations.....	56
Figure 3.17 RHS and linear weights estimations.....	57
Figure 3.18 An estimation of number of harmonics coded by harmonic template neurons	58
Figure 3.19 The preferred f_0 measured at different sound levels for harmonic template neurons	59
Figure 4.1, An example of the responses of a frequency selective neuron to pure tone and harmonic complex tones.	77
Figure 4.2. Metrics for quantify response pattern to HCTs.	78
Figure 4.3 Distribution of three different measurements for tuning to harmonic complex tones..	79
Figure 4.4 A nonparametric permutation test	80
Figure 4.5 An example of the responses of an band-pass neuron to pure tone and harmonic complex tones.	81
Figure 4.6 An example of how to quantify a band-pass neuron.	82
Figure 4.7 An example of a band-pass neuron which preferred high density and loud stimuli.....	83
Figure 4.8 An example of the responses of a modulation sensitive neuron.....	84
Figure 4.9 An example of a harmonic template neuron.....	85
Figure 4.10 Distributions of different subgroups of neurons in different metrics	86
Figure 4.11, A level independent rate code for spectral information of HCTs.....	87
Figure 4.12 Estimation of resolved harmonic numbers for different.....	88
Figure 4.13 A comparison between different subpopulations.	89
Figure 4.14 Distribution of different subpopulations in auditory cortex.	90
Figure 4.15 Correlations of neuron pairs in responses to complex tones.	91
Figure 4.16 a summary of all 63 pairs in BF and neuron correlations.....	92
Figure 5.1 An example of the fitted feed-forward model	110

Figure 5.2 The distribution of correlation coefficient for model predictions for training and validation sets in four subgroups of neurons	111
Figure 5.3 co-tuned excitatory and inhibitory subunits	112
Figure 5.4 Tone response properties of neurons with co-tuned subunits.....	113
Figure 5.5 An example of one run for the OED experiment.....	114
Figure 5.6 The estimated model online and prediction of responses to pure tones	115
Figure 5.7 An example of the optimal stimuli designed online.	116
Figure 5.8 An example of a neuron which was suppressed by sound stimuli.....	117
Figure 5-9 An example of a neuron which selectively responded to broadband stimuli.	118
Figure 5-10 An example of a neuron pair with different spectral selectivity.	119
Figure 5-11 The distribution of correlation coefficients for online model and offline model in four different subgroup of neurons.....	120
Figure 7.1 Four examples of the temporal response patterns in response to HCTs.....	133
Figure 7.2 onset synchrony test for harmonic grouping	134
Figure 7.3 harmonic grouping across space.....	135

CHAPTER 1:

1. Introduction

Our ability to quickly and accurately identify different sound sources in the environment relies on the analysis of the acoustic information arriving at both ears by our auditory system (Bregman 1990). We can easily distinguish different speakers who talk at the same time at a cocktail party, or identify different musical instruments in an orchestra. For animals, such ability is crucial for survival as it is important for detecting predators, hunting and localizing food. Many species rely on species-specific vocalizations to maintain the social group coherence, or to find mates for reproduction (Feng and Narins 2008, Shen, Feng et al. 2008). It is therefore crucial for the auditory system to analyze sounds related to communication.

1.1 Harmonic sounds, perception and auditory system

Many natural and man-made sounds, such as animal vocalizations, speech and sound played from most musical instruments are broadband with energy distributed across multiple frequencies which are integer multiples of the same fundamental frequency (f_0). Although the peripheral auditory system decomposes these sounds into separate frequency channels, harmonically related frequencies are tended to be grouped together to form a single percept (Bregman 1990, Darwin and Carlyon 1995). A central neural process is required to assemble information from those different channels to create a coherent representation of the sound. In

addition, the spectral properties need to be transformed into perceptual attributes, such as pitch and timbre, which define and distinguish different sound sources.

Pitch allows sounds to be ordered on a scale extending from low to high. For a harmonic complex tone (HCT), the elements of which are integer multiples of a fundamental frequency (f_0), the pitch is perceived at f_0 even the f_0 itself is missing (missing fundamental harmonic complex tone, MFs) (Schouten 1938). Harmonic complex tones of the same f_0 but different spectrums differ in other sound qualities, such as timbre, which can be used to distinguish different musical instruments playing the same note. The fusion of harmonics and perception of f_0 has also been shown on different animal models (Heffner and Whitfield 1976, Cynx and Shapiro 1986, Tomlinson and Schwarz 1988, Walker, Schnupp et al. 2009, Osmanski, Song et al. 2013).

It has been shown in previous studies that the auditory system has evolved to adapt to the statistics of the acoustic environment, adjusting to fulfill functional requirements. The mustached bat is a good example of such adaptation. The auditory cortex of mustached bats is divided into different functional regions for measuring the properties of echoes from emitted biosonar signals to navigate and locate prey in dark. In one of the regions, neurons only respond to tone pairs with specific frequency combinations corresponding to the range of encountered Doppler shifts (Suga, O'Neill et al. 1979). Another example is the barn owl, which has been a good model for studying sound localization. The neurons in the mid-brain auditory nucleus were found arranged systematically to create a physiological map of auditory space (Knudsen and Konishi 1978).

Therefore, it's of great importance to determine whether the auditory system has specialized structures to encode spectral and temporal features of harmonic sounds and how the physical attributes (frequency and level) are transformed into the perception that we experience.

1.2 Information transformation in the sensory system

The hierarchical processing network is common in all sensory systems. In visual system, the representation of objects starts from focal points of light in the visual peripheral (Adrian and Matthews 1927, Adrian and Matthews 1927), to line orientations in primary visual cortex (V1), then curvatures in V4 and finally the multipart shape configuration in inferotemporal cortex (IT) (Hubel 1968, Baker, Behrmann et al. 2002, Pasupathy and Connor 2002, Brincat and Connor 2004).

Similar to visual system, auditory system also consists of a hierarchy of processing stages which transform physical features of acoustical signals into perceptual attributes related to behavior context. If we use a harmonic complex tone as an example, different frequencies within this tone stimulate different regions of the cochlear: low harmonics near the apex of the cochlear while the high harmonics near the base. This stage of analysis can be approximated as a frequency analyzer with a serial of topographically band-pass filters centered at different characteristic frequencies (CFs). When the harmonic number is small, each component falls into a separate filter (resolved) and the filtered waveform is similar to single pure tone at the same frequency. The spectrum could be encoded by the spatial activation pattern of all frequency channels and the preserved temporal information (Cedolin and Delgutte 2005). Because the auditory filters become broader at high frequency, higher harmonics can fall into the same filter (unresolved). The interaction of frequencies generates complex waveform. For example, the output signal will have an envelope modulation at f_0 if all harmonics start from the same phase.

The spectral information and temporal information of the harmonic complex tone are transmitted to the auditory midbrain via auditory nerve fibers (ANs). The frequency-to-place representation in cochlear can be preserved in a rate-place manner by the tonotopic organization maintained throughout the auditory pathway. The temporal information is coded in the temporal firing pattern (phase-locking). Although such representations provide enough information of

acoustic properties of the harmonic complex tone, the central auditory system must integrate information across different frequency channel and represent the global properties, such as which spectral components are from the same complex tone but not from other presented sounds, and also the pitch and the timber of the tone giving the detected spectrum.

Auditory cortex is essential for such computations. Lesion in auditory cortex results in deficits in speech perception, judgments about shift in pitch of missing fundamental sounds, pitch discrimination tasks and vowel identification (Whitfield 1980, Coslett, Brashear et al. 1984, Heffner 1986b, Kudoh, Nakayama et al. 2006). From A1 and beyond, neurons are more selective to combination of features. Harmonic sensitive neurons were found in auditory cortex of mustache bats which showed facilitation when the first harmonic was simultaneously delivered with one or more higher harmonics (Suga, O'Neill et al. 1979). Neurons in belt auditory cortex are found driven by narrow-band noise or conspecific vocalizations (Rauschecker, Tian et al. 1995, Recanzone 2008). A putative pitch region was found near the anterolateral boarder of primary auditory cortex on marmoset monkeys, where pitch-selective neurons were identified (Bendor and Wang 2005, Bendor, Osmanski et al. 2012). Such findings provide evidence that higher order features have been extracted from the rate-place and phase-locking representations at cortical level, although it's still largely unknown where and how exactly such computation happens.

A1 becomes a good candidate, because the extraction of features, like pitch, requires integrating information across different frequency channels. The increased complexity in local neural network connections in A1 enables single neurons to integrate excitatory and inhibitory inputs from a broader frequency range. Multi-peaked neurons in A1 have been reported from studies on different species (Sutter and Schreiner 1991, Kadia and Wang 2003, Qin, Sakai et al. 2005) with separated excitatory peaks at harmonically related frequencies. It also has been shown that stimuli outside the classical frequency receptive field could modulate a neuron's responses to

simple or complex stimuli within the receptive field (Shamma and Symmes 1985, Nelken, Prut et al. 1994, Kadia and Wang 2003, Sutter and Loftus 2003, Qin, Sakai et al. 2005, Sadagopan and Wang 2010). Therefore, it's very important to study how single neurons in A1 represent harmonic complex tones, which could help reveal the neural mechanisms underlying pitch perception, auditory grouping, and identification of ethologically relevant sounds.

1.3 An overview of methods for studying frequency selectivity in auditory system

Frequency selectivity is a fundamental property of auditory neurons. For ANs, the frequency selectivity can be described by the tuning curve: the sound levels needed to produce a significant increase in firing rate at different frequencies. In the central nervous system, single tones can evoke both excitatory and inhibitory responses. The responses of a single neuron to different frequencies are determined by all excitatory and inhibitory inputs. The frequency analysis of complex sounds cannot be fully characterized by responses to pure tones, nor can be measured by creating a look-up table from testing all possible combinations of frequencies. Instead, parametric, broadband stimuli have been used to characterize neurons and a transfer function is used to describe the transformation of sound spectrums to neural responses.

A widely used method in measuring frequency tuning of auditory neurons is the spectro-temporal receptive field (STRF) characterization and weighting function models (Theunissen, Sen et al. 2000, Yu and Young 2000, Barbour and Wang 2003, Linden, Liu et al. 2003, Reiss, Bandyopadhyay et al. 2007). Those approaches have improved our understanding of the frequency selectivity of auditory neurons. They have provided insights to understanding the mechanism underlying the functional properties. However, such approaches have limitations. First of all, STRFs and linear weights can only adequately characterize neurons that are approximately linear (Young, Yu et al. 2005, Christianson, Sahani et al. 2008). Secondly, more

complex models such as the quadratic models and neural network models often require large amount of data and computation (Yu and Young 2000, Prenger, Wu et al. 2004, Bandyopadhyay, Reiss et al. 2007). Third, the selectivity of cortical neurons to complex sounds requires sampling in a high dimensional space in order to get a good estimation of the transfer function, which is difficult in a limited recording time. A combination of models and efficient data-collecting method could potentially be a useful tool to study how cortical neurons process complex sounds.

1.4 Objectives of the dissertation

Two main objectives guide the experiments in this dissertation: 1) to investigate the neural presentation of harmonic complex tones and information transformation in primary auditory cortex (A1); 2) to explore an online adaptive stimulus design approach based on neural network model for studying spectral integration in A1. There are three questions that I have attempted to address:

1, Are there representations of harmonic structures at single neuron level in auditory cortex beyond pitch (Chapter 3)?

The discovery of pitch neurons suggests a possible pre-processing stage in A1 for computing f_0 from spectral information. Although neurons in A1 have been extensively studied by using pure tone, two-tones and multiple tones, there is little evidence of the existence of single neurons suitable for extracting pitch, for example, a harmonic template (Goldstein 1973).

2, Are there any functional structures for harmonic processing in A1 (Chapter 4)?

The heterogeneity within the large scale tonotopic organization in A1 suggests a parallel process for harmonic complex tones. However it also remains unknown how individual neurons encode the spectral and temporal information of harmonic complex tones differently, given the

different selectivity to frequency, level and modulation. A comparison of the representation in A1 and the representation in subcortical region would help us understand the information transformation along the auditory pathway better.

3, Can the optimal experimental design (OED) be applied to study on spectral integration of neurons in auditory cortex? (Chapter 5)

Previous study proposed a general statistical data-collection method (optimal experimental design, OED): a small number of optimally designed stimuli can yield a model estimation as good as a large number of random independent and identically distributed (IID) samples (DiMattina and Zhang 2011). This method has been broadly used in machine learning, psychophysics studies and theoretical simulations (Watson and Pelli 1983, MacKay 1992, Lewi, Butera et al. 2009). Only a few studies have applied this method to actual in-vivo neurophysiology experiment (Tam 2012). When complex tones are used, searching for preferred stimuli of cortical neurons has also proven challenging because of the high dimensionality of the acoustic space of possible stimuli and limited recording time. In Chapter 5, I explore the feasibility of a new method combining neural network model and OED for studying spectral integration in auditory cortex.

CHAPTER 2:

2. Methods and experimental design

2.1 General experimental procedure

All recordings were done in a double-walled soundproof chamber (Industrial Acoustics, Bronx, NY). Single neurons' responses were recorded from four hemispheres of three marmoset monkeys (M79u, left hemisphere; M73v, both hemisphere; M6x, left hemisphere). All experiment procedures were approved by the Johns Hopkins University Animal Use and Care Committee. Details of the chronic recording preparation can be found in previous papers from our laboratory (Lu, Liang et al. 2001). Animals were adapted to sit quietly in a primate chair with the head immobilized. A Tungsten electrode (A-M System, 2-5 M Ω) was inserted into the auditory cortex perpendicularly to the surface through a 1 mm craniotomy on the skull. The electrode was manually advanced by a hydraulic microdrive (Trend Wells). Harmonic complex tones and pure tones were used for searching neurons.

Each recording session lasted 3 – 5 hours. Animals were awake but not performing any task during recordings. All spikes waveforms were high-pass filtered (300Hz – 3.75kHz), digitized and sorted in a template based online sorting software (MSD, alpha omega engineering).

2.2 Acoustic Stimuli

All sound stimuli were generated digitally and delivered by a speaker 1m away from the front of the animal in a free field. All stimuli were sampled at 100 kHz and attenuated to a desired sound pressure level (RX6, PA5, Tucker-Davis Technologies).

2.2.1 General acoustic stimuli

Tones at different frequency (1-40 kHz, 10 steps per octave) were played at a moderate sound level (30-60dB SPL) to measure the frequency selectivity. Best frequency (BF) was defined as the pure-tone frequency which evoked the maximal firing rate. The threshold of a neuron was estimated from the rate-level function of the BF tone (-10 to 80dB SPL, in a step of 10dB). Threshold was defined as the lowest sound level evoke response significantly different from spontaneous firing rate (t-test, $p < 0.05$). Typically, stimuli were 100 ms in duration with a 500ms inter-stimulus interval (ISI) and 5ms onset and offset ramp. Longer durations (200ms and 500ms) with longer ISIs (> 1000 ms) were used for neurons with a long response latency and long lasting offset firing. Every stimulus was presented at least 10 repetitions in a random order.

2.2.2 Harmonic complex tone and inharmonic complex tones

Three types of tone complexes were used to study the spectral selectivity: harmonic complex tones (HCTs), shifted complex tones, spectrally jittered complex tones. For all complex tones, individual components were kept at the same sound intensity level and were usually added with cosine phases (COS). Most of our complex tones were played at modest levels. The sound intensity per component was 10dB above the threshold of BF tone typically. If a neuron did not respond to pure tones at all, a 40dB SPL sound level would be used. For some neurons, HCTs were tested at two or three higher sound intensities. However, no sound intensity per component

was above 60 dB SPL considering the overall sound level. Additional test with alternating phase (ALT) or random phase (RND) were used on some neurons to exclude the effect of temporal cues.

Harmonic complex tones at different fundamental frequencies (f_0 s) were generated but only harmonics within a 3 octave frequency range centered at BF were used. The preferred HCT, the one that evoke the maximal firing rate, were used for generating shifted complex tone and spectrally jittered complex tone. Shifted complex tones were generated by adding a shift frequency to the first six harmonics. The shift frequency was proportional to f_0 in a step of 25% of f_0 . Spectrally jittered inharmonic complex tones were generated by jittering individual component of the preferred harmonic complex tone except the one at BF. The jittered components were randomly and independently drawn from a uniform distribution with mean at the corresponding harmonics and standard deviation proportional to the jitter level (Figure 2.1). Five different jitter levels, 10%, 20%, 30%, 40% and 50% were used. Twenty five stimuli with different spectral contents were generated for each jitter level.

2.2.3 *Random harmonic stimulus (RHS)*

Random harmonic stimulus (RHS) was adapted from random spectrum stimulus (RSS) (Yu and Young 2000, Barbour and Wang 2003). The acoustic stimuli contained simultaneous harmonics of a f_0 within a three octave frequency range centered at BF. The RHS were arranged into sets that determined the levels for each component of individual stimuli such that the set as a whole was “white”, (the stimuli were uncorrelated to each other) and constant-variance (the level distributions of each component were identical). All RHS sets contained ten more stimuli than the number of harmonics. An auxiliary RHS set was created by using the spectral inverse of the original set to increase the frequency space sampling density. Figure 2.2 showed the adjusted level matrix of the RHS set used for one neuron. Each row of the plot is a stimulus. Each column

refers to each harmonics. In the adjusted level matrix, each harmonic has had the mean level subtracted. Three stimuli are plotted on the right to show the different spectral profiles of individual stimulus. There are four independent parameters for RHS: frequency range, f_0 , mean sound level and standard deviation of sound level. For every neuron tested with RHS, the f_0 is chosen as a ratio of the preferred f_0 (Bf_0), typically $Bf_0/2$ and $Bf_0/4$.

2.3 Data Analysis

Firing rates were calculated over the time window between 15ms after onset and 50ms after offset. The preferred harmonic complex tone was defined as the one evoking the maximal firing rate. The f_0 of the preferred harmonic complex tone was referred as the preferred f_0 (Bf_0).

2.3.1 Characterize neural responses to HCTs

A *facilitation index* is defined as $(R_{\text{HCT}} - R_{\text{Tone}})/(R_{\text{HCT}} + R_{\text{Tone}})$, where R_{HCT} is the firing rate to the preferred harmonic complex tone, and R_{Tone} is the maximal firing rate to individual components in the preferred harmonic complex tone at equal sound level. Facilitation index is a measure of the increase in firing rates by presenting a tone in different spectral context. The facilitation index is 1 if the neuron only responds to the harmonic complex tone but not individual components alone. The index is 0 if other simultaneously presented tones do not change the neural response to a single tone.

In our test, because the f_0 of HCTs was changed in small increments, the peak in the frequency tuning (BF) either concurrent with a harmonic or fell in between two harmonics. If harmonic number, the ratio of BF to f_0 , is used to replace f_0 , the rate-harmonic profile should display an oscillating patten with peaks at integer values and valleys between two adjacent integer values. Periodicity in the rate- f_0 tuning can be measured by using discrete Fourier transform

(DFT) of the rate-harmonic number tuning. The peaks in the power spectral density (DFT) are used to determine whether there is oscillation in the tuning. A nonparametric permutation test (Ptitsyn, Zvonic et al. 2006, Fishman, Micheyl et al. 2013) was used to assess the statistical significance of the peaks by randomly shuffling the sample points in the rate-harmonic tuning. For every shuffled trial, the power spectral density was computed and the amplitude at peak frequency in the PSD of the original sequence was measured. The null-hypothesis is that a random point serial will yield a same amplitude or higher peak in the PSD. A null reference distribution of the spectral amplitude at the given peak frequency was generated by repeating this process for 1000 times. The p-value is defined as the probability of observing an amplitude at the given peak frequency equal or larger than the observed value.

Harmonic Tuning width was used to evaluate whether individual harmonics can be resolved in the rate-harmonic number profile. A moving window of 1 was centered at each integer harmonic number and the total response area, where firing rate was larger than 20% of the maximal firing rate in the window, was calculated. A square was used to match the normalized response area with the height at the maximal firing rate in the window. The width of the square was the harmonic tuning width for each integer harmonic number. The less response for non-integer harmonic number, the smaller harmonic tuning width is. The *density preference index* is defined as the ratio of the harmonic tuning width at the maximal harmonic number to the maximal harmonic tuning width.

A *periodic index* is defined below to quantify the tuning to shifted complex tones.

$$\text{Periodic Index} = \sum_{i=1}^N \frac{(2 \times R_{n_i} - R_{n_i+50\%} - R_{n_i-50\%})}{(2 \times R_{n_i} + R_{n_i+50\%} + R_{n_i-50\%})} / N$$

Where R is the firing rate, and n_i for R_{n_i} is the percentage frequency shift at multiples of 100%. Only those shifts which evoke a firing rate significantly different from spontaneous rate were included (t-test, $p < 0.05$). If a neuron only responds to harmonic shifts, the periodicity index will be 1. If a neuron cannot distinguish harmonic shifts from others shifts, the periodic index will be close to 0.

2.3.3 Neural correlations

Signal correlations were computed for each neuron pair as the correlation coefficients between the mean firing rates to all stimuli tested including all complex tones. Noise correlation were computed using a normalized correlation measure (Bair, Zohary et al. 2001, Jeanne, Sharpee et al. 2013) across trials. In this measure, the firing rate on each trial was subtracted by the mean firing and normalized by the standard deviation across all trials to obtain z-score.

2.3.4, Identification of cortical areas and layers

Primary auditory cortex was identified by neural response properties to tones and tonotopic gradient (low frequency at rostral-lateral and high frequency at caudal-medial). The border between A1 and the rostral regions (R and RT) was estimated from the frequency gradient reversal. Low-pass filtered Local field potentials (1Hz – 300Hz) (LFPs) were also recorded. If time permits, at the recording section, LFPs were recorded every $100\mu m$ to a 20dB tone at the BF of the track in a total depth of 2.5mm from dura typically. One dimensional current source density (CSD) was calculated as the second spatial derivatives of the LFPs (Figure 2.3). CSD profile is a representation of the direction, location and intensity of the transmembrane currents underlying the evoked response potentials (Schroeder, Mehta et al. 1998, Hirsch and Martinez 2006). Current sinks (negative peaks) are indicator of depolarizing events, such as the recipient layers for

thalamo-cortical projections (layer III/IV). Current sources (positive peaks) are indicator for passively return currents. Although we did not do CSD for all track to fully categorize the neurons we recorded into different layers, it provides a reference for a general estimation of the layers we recorded from in addition to the response properties. Most of samples in our recordings were from layer II and III.

2.3.5 Linear weighting function estimation

The RHS results can be used to build a linear model of how a cortical neuron responds to harmonic complex tones with arbitrary spectral profile. The model can be written as

$$\vec{R} = R_0 + \Lambda \vec{w},$$

Where R is a vector of firing rate to the set of RHS, R_0 is a constant, the response to flat spectrum stimulus, Λ is the adjusted level matrix and \vec{w} is the linear weighting function, which represents an estimate of rate slopes at each frequency. All rates used in the linear weights estimation are driven rates: averaged firing rate minus spontaneous rate. Weighting functions are computed by solving the synthesis equation. Because of the way the adjusted level matrix is constructed, the equation above can be solved to obtain the unconstrained least-squares estimate:

$$\vec{w} = (\Lambda^T \Lambda)^{-1} \Lambda^T \hat{R}$$

where \hat{R} are rates measured in experiment.

The *Fraction of variance* f_v is used to evaluate how good the model is to predict the responses to different stimuli (Reiss et al. 2007):

$$f_v = 1 - \frac{\sum_{n=1}^N (r_n - \hat{r}_n)^2}{\sum_{n=1}^N (r_n - \bar{r})^2}$$

where r_n is the measured firing rate in response to the n th stimulus, \hat{r}_n is the predicted rate from the model, and \bar{r} is the average rate of r_n .

2.4 Online adaptive experiment design (OED)

2.4.1 A feed-forward neural network model

The model used in the online adaptive experiments was adapted from previous works. We assume that adjacent excitatory and inhibitory neurons in A1 get their excitatory inputs from thalamocortical axons with similar best frequencies, and that an excitatory cortical neuron receives feedforward inhibition from other cortical neurons (Dominguez et al., 2006; Soto et al., 2006; de la Rocha et al., 2008; Wu et al., 2008). We choose feedforward neural network models because they are the simplest models that can capture the general notion that more complex responses should arise by combining simpler responses from the lower levels.

The model is composed of three excitatory subunits (E1, E2 and E3) and two inhibitory subunits (I1 and I2), all converging into a single principal cell at the top which is the supposedly recorded neuron in A1 (Figure 2.4). The rationale for choosing multiple excitatory subunits was to allow the formation of complex frequency receptive fields with more than one excitatory peak.

The input synaptic weights of each subunit are modeled by a Gaussian with only three free parameters:

$$w_i = A \exp\left(-\frac{(f_i - \mu)^2}{2\sigma^2}\right)$$

where f_i is the frequency of the i -th input node (circles at the bottom of Figure 2.4), μ is the center frequency corresponding to the BF of the subunit, σ is the width of the spectral input range, and A is the weight amplitude. The input $\mathbf{x} = (x_1, \dots, x_n)$ to the network is the spectrum of a sound stimulus: x_i is the sound level (dB) in the i -th frequency bin. For simplicity, we did not take into account any subcortical nonlinearity in the inputs. The output of each subunit is given by

$$y = g\left(\sum_{i=1}^n w_i x_i - B\right)$$

where B is the bias and $g(u) = 1/(1 + e^{-u})$ is the sigmoid gain function. The synaptic connection weight between each subunit and the principle cell is always constrained to be positive for an excitatory subunit and negative for an inhibitory subunit. The principal cell sums the weighted outputs from all subunits, passes the sum through the sigmoid function, and then multiplies the outcome by a final gain factor to yield the final response (averaged firing rate). Each subunit includes 5 parameters: 3 for its Gaussian shaped weights (amplitude, mean and variance), 1 for the bias, and 1 for the weight of its connection with the principal cell, yielding a subtotal of 25 parameters for all 5 subunits. After adding the final gain factor, model 1 has a total of 26 free parameters.

2.4.2 Offline Data fitting

The entire stimuli-response set recorded for each neuron was divided randomly into a training set that included 75% of the data and a validation set that included the remaining 25%. Model was fitted to the training data through minimization of the mean

square error (MSE) between the observed average firing rate and the model's prediction for each stimulus in the set:

$$\text{MSE} = \frac{1}{N} \sum_{n=1}^N (r_n - f(\mathbf{x}_n))^2$$

where N is the number of data points in the set, r_n is the recorded response to the n -th stimulus, \mathbf{x}_n and $f(\mathbf{x}_n)$ is the model's predicted mean response to the same stimulus. The observed average firing rate was computed from the spike count within a time window from 15ms after onset and 50ms after offset. The optimization was performed using the function `fmincon` in Matlab (MathWorks, Matick, MA) which finds the minimum of a constrained nonlinear multivariate function (MSE) over the space of network parameters, starting from a single initial guess. In order to avoid local minima of the MSE in the parameter space and increase the probability of finding an optimal solution, each optimization process was repeated 1000 times, starting from random initial guesses. The chosen parameter set was the one that yielded the smallest MSE for the training set.

2.4.3 Experimental system design for OED

This hardware setup is the same as the traditional neurophysiology recording in section 2.1. The different part is, instead of pre-generate a set of stimuli, every stimulus is generated online (Figure 2.5). After one stimulus is played, the neural response is recorded and then the OED algorithm will find the next optimal stimulus and update the model parameters according to the recorded response. The total iteration is 300. After Every 5 iterations, a randomly picked stimulus will be played to avoid local minimum or possible neural adaptation. All stimuli were

200ms long with a 5ms onset and offset ramp. The inter stimulus interval was at least 600ms depending on the time of computation.

2.4.3.1 Stimulus design

In this experiment, the collected responses are used to estimate the parameters θ of a hypothesized model $f(\mathbf{x}, \theta)$ given the distribution $p(r|\mathbf{x}, \theta)$ of a neuron's stochastic response, where r represents the response and \mathbf{x} denotes the stimulus. Given a set of stimulus – response observations $D_n = \{(\mathbf{x}_1, r_1), \dots, (\mathbf{x}_n, r_n)\}$, the posterior distribution of the parameters vector θ can be given by Bayes rule, shown below:

$$p_n(\theta|D_n) \propto p(D_n|\theta)p_0(\theta)$$

where $p(D_n|\theta) = \prod_{i=1}^n p(r_i|\mathbf{x}_i, \theta)$ assuming the responses are independent and identically distributed and $p_0(\theta)$ is the prior distribution of θ .

A possible way to reduce the uncertainty of $p_n(\theta) \stackrel{\text{def}}{=} p_n(\theta|D_n)$ is to quantify it using the differential entropy $H[p_n(\theta)]$ of the posterior density. This way, stimuli that minimize this entropy using the current estimate of the parameters $p_n(\theta)$ can be iteratively chosen (DiMattina & Zhang, 2011). This strategy is implemented in the algorithm presented by designing stimuli that minimize the following utility function:

$$U_{n+1}^{(E)}(\mathbf{x}) = - \int H[p_{n+1}(\theta)]p(r|\mathbf{x})dr$$

where $p(r|\mathbf{x}) = \int p(r|\mathbf{x}, \theta)p_n(\theta)d\theta$ is the response distribution given input \mathbf{x} . More details of the algorithm can be found in the master thesis of Eyal Dekel (Dekel 2012).

In every iteration of the algorithm, 5000 stimuli are generated and passed through the utility function, where the one that maximizes the utility function is chosen to be presented. These

stimuli are composed of 31 frequency bins, spanning a frequency range that is specified by the user at the beginning of the process. Every frequency bin can get any sound level intensity from zero to a maximum value, specified again by the user to make sure the model operates in a reasonable range of sound levels.

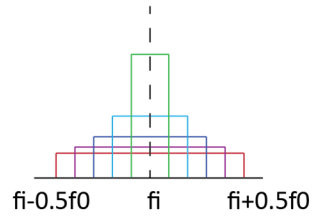
2.4.3.2 Parallel modeling

In our experiment, a modified version of the OED algorithm for parallel estimation of multiple models was developed and implemented by using Matlab “Parallel computing toolbox”. This version of the algorithm is initialized with four different initial parameter estimates (means of the parameters prior distribution) serving as starting points for four different models. During the run of the algorithm, the parameters of each model are updated in parallel (though independently) following each stimulus. Even though the update of the models parameters can be done in parallel at approximately the same time, the optimal stimuli chosen by each one of the models cannot be presented in parallel to the animal. In our implementation, the model which fits the data collected up to the specific iteration the best (the least MSE) will determines the next optimal stimulus.

Despite the fact that the rest of the models are updated based on a stimulus that is not designed according to their specific utility function, this parallelization still has some advantages comparing with single processing. First, starting from four different initial guesses of the parameters increases the chances that one of these guesses is closer to the optimal estimate of the network. Second, as the models evolve, the “favorite” model (i.e. the one that fits the data best) may switch between the four, and so the optimal stimuli are not designed strictly based on one model.

A

Distribution of the i th harmonic



- HCTs
- 10% jitter
- 20% jitter
- 30% jitter
- 40% jitter
- 50% jitter

B

Spectrum of jittered complex tones

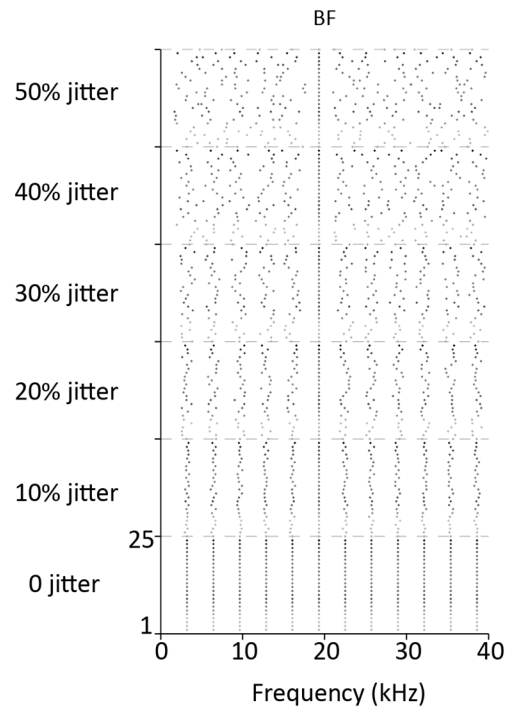


Figure 2.1 Demonstration of Jittered complex tones

A, the uniform distribution for individual component in the jittered complex tones at different jitter levels

B, the spectrum of jittered complex tones

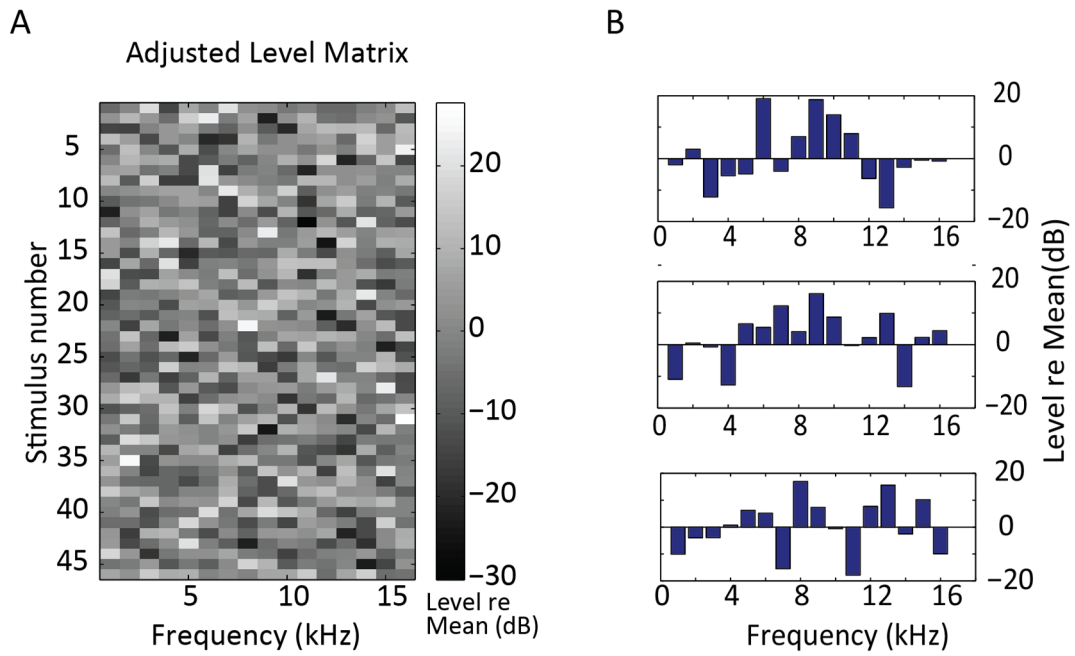


Figure 2.2 An example of a RHS set

A, the adjusted level matrix of a RHS set.

B. three stimuli from the RHS set in A with different spectral profiles.

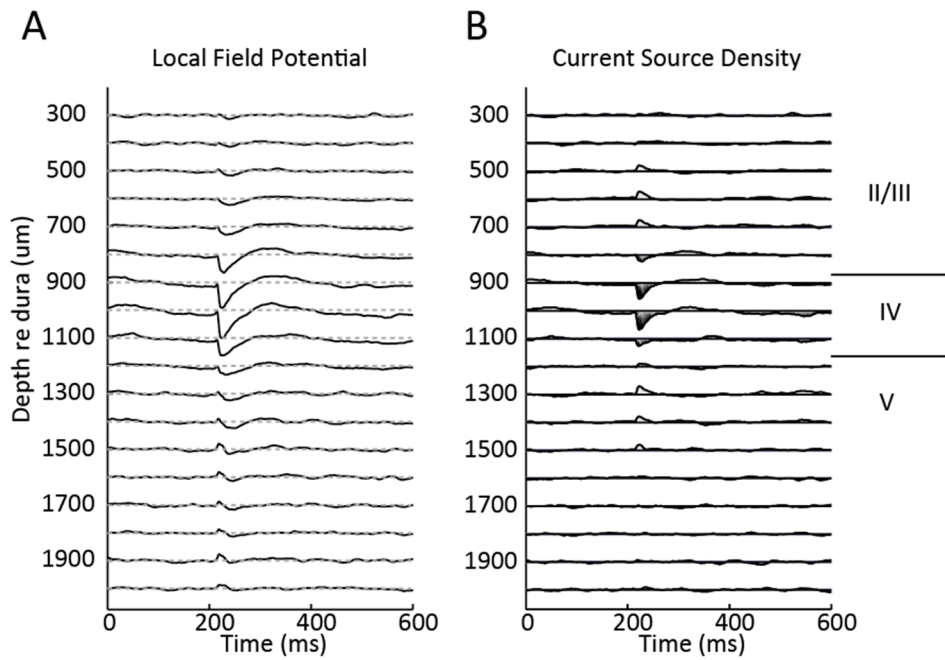


Figure 2.3 An example of LFPs and CSD of one recording track

A, LFPs recorded at different depth from the dura with a step of $100\mu\text{m}$.

B, CSD from the LFPs in A. Negative peaks indicate sinks.

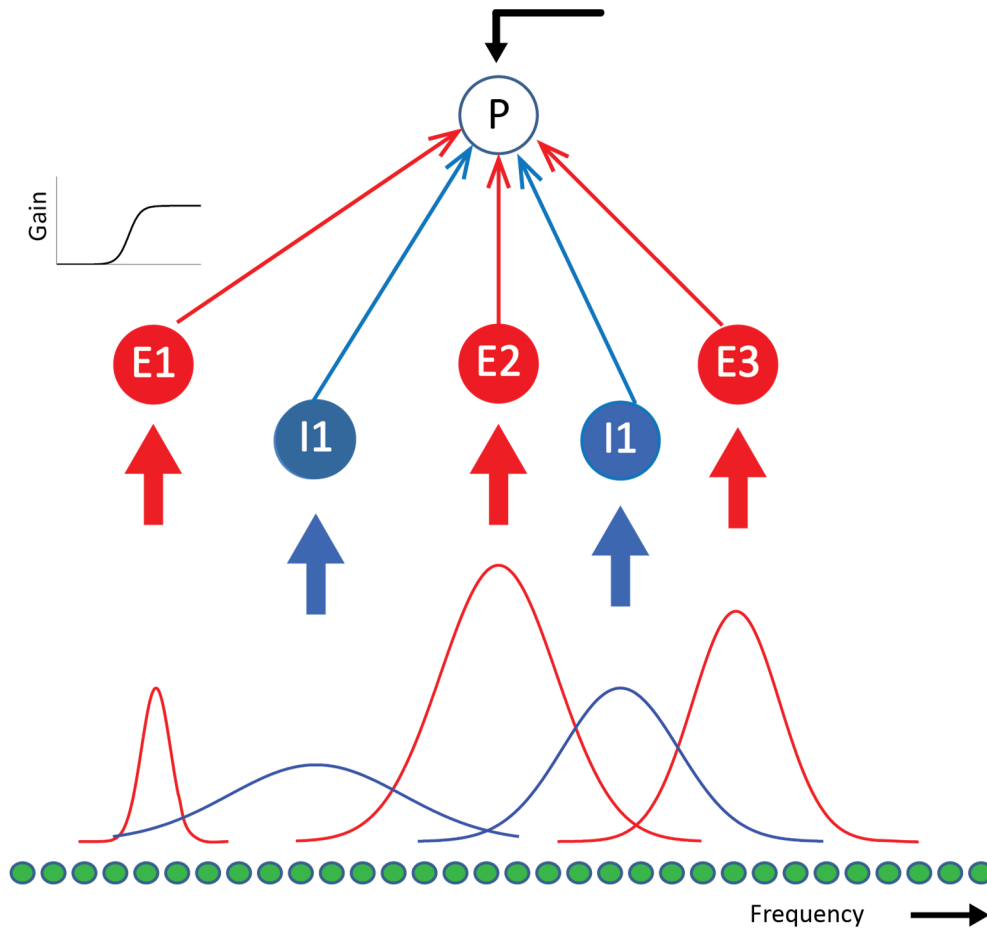


Figure 2.4 A feed-forward network model with three excitatory subunits (red) and two inhibitory subunits (blue).

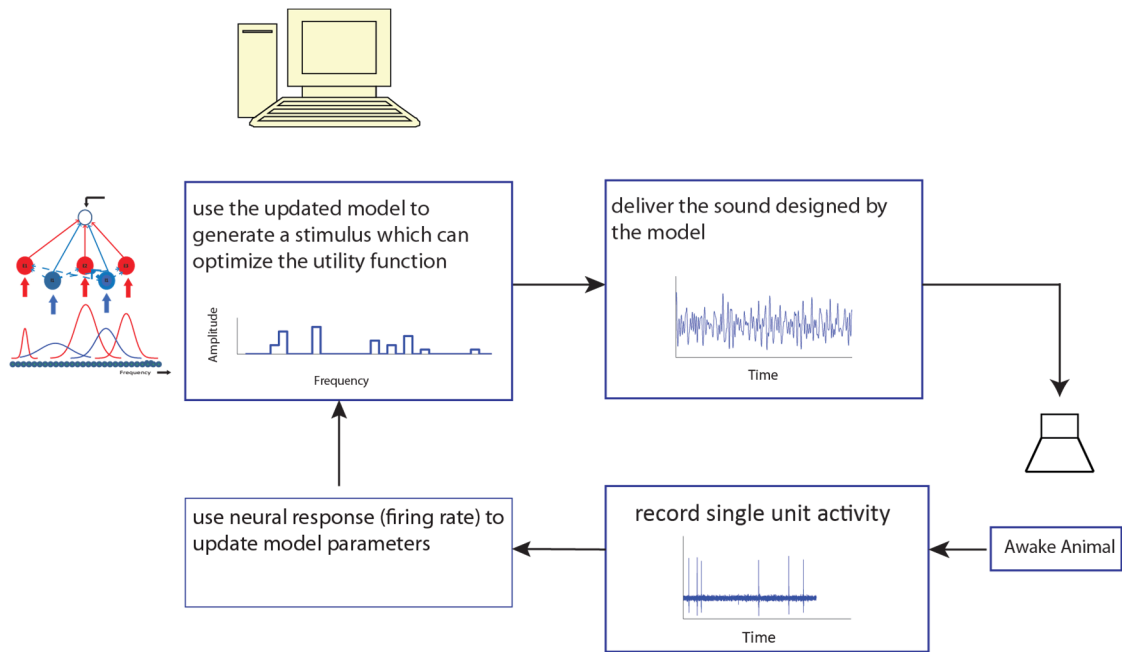


Figure 2.5 A diagram of the online adaptive experiment system

CHAPTER 3:

3. Spectral templates for encoding harmonic structure in marmoset auditory cortex

3.1 Introduction

Many natural sounds, like animal vocalization, human speech, and the sounds from most musical instruments comprise of a set of sinusoids at frequencies which are integer multiples of the same fundamental frequency (f_0). The fusion of harmonics is a natural process which groups all frequencies relevant to the same sound source together to form a single percept (Bregman 1990, Darwin and Carlyon 1995). Perceptual properties such as pitch and timbre can be then computed from the harmonic spectra (Schouten 1938). Such harmonic process can also be associated with the preference to consonant chords over dissonant chord in music perception, because the consonant chords, such as the perfect fifth, are more similar to single notes with harmonic spectra (Terhardt 1974, Tramo, Cariani et al. 2001, Ebeling 2008, McDermott, Lehr et al. 2010). The fusion of harmonics and perception of f_0 has also been shown on different animal models (Heffner and Whitfield 1976, Cynx and Shapiro 1986, Tomlinson and Schwarz 1988, Walker, Schnupp et al. 2009, Osmanski and Wang 2011), which implicate a common neural mechanism for harmonic sound processing across different species.

At the peripheral auditory system, a harmonic sound is decomposed into different frequency channels and can be represented in terms of its constituent parts. Therefore, a central process is required to accomplish the harmonic fusion and following computations. Auditory cortex is essential for processing harmonic sounds. Deficits in pitch discrimination and vowel

identification were observed following cortical lesion (Whitfield 1980, Kudoh and Shibuki 2006). In addition, electrophysiology recordings in primary auditory cortex (A1) found multi-peaked neurons with separated excitatory peaks at harmonically related frequencies (Sutter and Schreiner 1991, Kadia and Wang 2003, Qin, Sakai et al. 2005). Previous studies have also shown a nonlinear two-tone interactions of harmonically related frequencies at single neurons in A1 (Suga, O'Neill et al. 1979, Kadia and Wang 2003). The emergence of such harmonic sensitivity suggests that neural circuits in auditory cortex have evolved to efficiently process harmonic sounds. However, studies using harmonic complex tones showed that populations of neurons in A1 carry sufficient spectral-temporal information for pitch extraction but did not find any combination sensitive neurons (Schwarz and Tomlinson 1990, Fishman, Reser et al. 1998, Fishman, Micheyl et al. 2013). Neurons which selectively responded to harmonic complex tones were found in a tentative pitch region in marmoset auditory cortex and have proven to code pitch (Bendor and Wang 2005, Bendor, Osmanski et al. 2012), but they were only tuned to low frequency pitches (<1kHz). Marmosets can hear up to 36kHz (Osmanski and Wang 2011) and the fundamentals of four major marmoset calls are between 5kHz to 10kHz (DiMattina and Wang 2006). It is possible that there is a more generalized harmonic process in auditory cortex over a broader frequency range in A1. In the present study, we systematically tested neurons in A1 with harmonic and inharmonic complex tones, varying f_0 and harmonic compositions. We found harmonic template neurons in A1, which were strongly driven by harmonic complex tones and showed reduced response to inharmonic complex tones. Responses of harmonic template neurons were selective to fundamental frequency (f_0) and sensitive to harmonic number and spectral regularity. Those templates covered a broad range of f_0 (400Hz – 12kHz). They were distributed in the entire frequency range of A1 and organized tonotopically by their best frequencies (BFs). Our findings indicate that there is a more generalized harmonic process in A1 for harmonic fusion and pitch extraction.

3.2 Results

We recorded from four hemispheres of three awake marmoset monkeys. 372 well isolated and fully tested neurons which responded (the firing rate to at least one tested harmonic complex tone is significantly larger than spontaneous firing rate, t -test, $p < 0.05$) to harmonic complex tones were used in data analysis.

3.2.1 Neurons showed combination sensitivity to harmonic complex tones

Most neurons in auditory cortex had frequency preference with a peak response at their BFs (Figure 3.1A,B). When a second tone was played simultaneously with the BF tone, the response to BF tone could be either facilitated or suppressed depending on the frequency of the second tone (Figure 3.1C). The two facilitatory areas above and below BF were interleaved with suppressive areas around BF (Figure 3.1D). When tested with harmonic complex tones, this example neuron showed a significant increase in responses to some f0s (Figure 3.1E). In the tuning to f0s, the two largest peaks were separated by f0s which evoked less or no response. By examining the spectrum of the complex tones corresponding to the two peak responses (Figure 3.1E, left), we found that both tones included three components near BF and in the two facilitatory areas in two-tone test. In other words, this example neuron selectively responded to the concurrence of components from the three excitatory frequency areas. Moreover, this neuron showed onset responses to pure tones but exhibited sustained firing to the preferred harmonic complex tones (Figure 3.1A and 3.1E). Previous study has showed that sustained responses are usually evoked in auditory cortex by preferred stimuli (Wang, Lu et al. 2005). This example suggests pure tone responses, even two-tone responses might underestimate the spectral selectivity of some neurons in A1.

Previous studies reported 25 ~ 50% neurons in A1, which do not respond to simple stimuli such as pure tones and noise but to complex sound features (Evans and Whitfield 1964, Sadagopan and Wang 2009). In our study, we also found neurons that did not respond to pure tone but only responded to harmonic complex tones at specific f_0 s (Figure 3.2A and B). Since some cortical neurons were narrowly tuned to levels and show a non-monotonic rate-level tuning (Sadagopan and Wang 2008), the preferred harmonic complex tone and its individual harmonics were tested separately at different sound levels. These neurons did not respond to f_0 alone nor to other individual harmonics. But the response to the preferred f_0 at 621Hz was consistent at different sound levels (Figure 3.2C). Since it is not uncommon for A1 neurons to be sensitive to amplitude modulation (Lu, Liang et al. 2001), harmonic complex tones with cosine and alternating phases were tested separately for these neurons. A harmonic complex tone added at cosine phase will have an envelope modulation at f_0 while alternating phase gives rise to a modulation at $2*f_0$ with the same spectral content. The example neuron showed similar tunings to both cosine phase and alternating phase (Figure 3.2D), suggesting the enhanced responses to harmonic complex tones were due to spectral integration rather than temporal modulation. We used a facilitation index (FI) to quantify neurons' preference to harmonic complex tones. A FI is between -1 and 1. A FI larger than 0 indicates a preference to complex tones while a negative FI indicates a preference to pure tones. Only neurons with a FI larger than 0.3 were identified as candidates for harmonic template neurons (Figure 3.5A).

3.2.2 Harmonic template neurons process spectral information

When complex tones are presented, the change in spectral density, total sound level, and temporal modulation might cause the increase in firing rate other than harmonicity itself. In order to separate neurons which were sensitive to harmonic structures rather than other co-varying sound parameters, frequency-shifted complex tones were used to test candidates which preferred

complex tones over pure tones. In this test, the first five or six harmonics of the preferred f_0 (0% shift) were gradually shifted in frequency by an amount proportional to f_0 (Figure 3.3A). When the shifts were multiples of 100%, the complex tones were still harmonic with the same f_0 but different compositions. Other shifts either resulted inharmonic complex tones (25%, 75%, 125%, ...) or odd-harmonic complex tones with a fundamental half octave below the original f_0 (50%, 150%, 250%, ...). The temporal modulation remained the same because of the equal spectral space in all shifts (Figure 3.3A). For humans, the perceived pitch shifts with the inharmonic shifts and ambiguous pitches were perceived for odd-harmonic complex tones if resolved harmonics are included (Patterson and Wightman 1976, Moore and Moore 2003). On the contrary, the shifts had less or negligible effects on perceived pitch of unresolved harmonics because the temporal cue was used (Moore and Moore 2003). Neurons which encode spectral information of harmonic complex tones should be able to distinguish harmonic shifts from inharmonic shifts or the odd-harmonic tones (Figure 3.3B). The example neuron showed a periodic response pattern to the shifted complex tones: large responses at harmonic shifts (100%, 200%, 300%, ...) when the f_0 remained the same, while weaker or no response at inharmonic shifts and odd-harmonic complex tones. More example neurons were shown in Figure 3.4. Based on the responses to harmonic complex tones and shifted complex tones, we had our criteria for harmonic template neurons which: 1, showed larger response to harmonic complex tones than responses to pure tones (Facilitation index > 0.3 , Figure 3.5A); 2, preferred to harmonic complex tones over inharmonic tones or odd-harmonic tones (Periodicity index > 0.5 , Figure 3.5B).

3.2.3 The relationship between BF and preferred f_0 (Bf_0)

Another observation from the response pattern to shifted harmonic complex tones was the selectivity to harmonic compositions. Despite the selectivity to f_0 , the response to the first five harmonic (0% shift in Figure 3.3B) was weak even the f_0 itself was presented and the pitch

salience was high. The largest firing rate was always at harmonic shifts but at different harmonic compositions for different neurons (Figure 3.4). Such response pattern suggests harmonic template neurons encode not only f_0 but also absolute frequency of individual components. The next question is how those absolute frequencies are related to the BFs of harmonic template neurons.

For harmonic template neurons, the preferred f_0 s were linearly proportional to the BFs (Figure 3.6A). The ratios between BF and the preferred f_0 s were close to integer numbers (Figure 3.6B). In other words, the preferred f_0 of a harmonic template neuron equals to BF or subharmonics of BF ($BF/2$, $BF/3$, ...). In order to examine the role of BF in coding harmonic complex tones, we compared responses of two neurons with the same preferred f_0 but different BFs. The same set of harmonic complex tones comprising of different harmonic of the same f_0 were used. The BF of the first neuron was the third harmonic of the preferred f_0 while the second neuron had a BF at the eighth harmonic (Figure 3.7A). Low BF neuron preferred to the lower harmonics whereas the high BF neurons selectively respond to the higher harmonics (Figure 3.7B and C). Harmonic template neurons were selective to absolute frequencies which were determined by their BFs.

Another way to test the frequency region for harmonic template neurons was using sinusoidal amplitude modulated tones (sAMs), which consisted of just three components (Figure 3.8A). A sAM tone becomes harmonic when the carrier frequency (f_c) is an integer multiple of the modulation frequency (f_m). Different carrier frequencies with the modulation frequency at the preferred f_0 of a harmonic template neuron were used, which were similar to the frequency tuning to pure tones. The example neuron in Figure 6B did not respond to pure tone alone. When an amplitude modulation at preferred f_0 was added, this neuron showed responses to carrier frequencies that were harmonics of the f_0 (Figure 3.8B). The second neuron also showed multiple peaks at harmonics of f_0 (Figure 3.8C). Either neuron responded to carrier frequencies between

those harmonics. The results from sAMs experiments further supports that harmonic template neurons encode both f_0 and harmonic numbers.

In all 64 harmonic template neurons we recorded, the preferred f_0 s vary from 400Hz to 12kHz and the BFs cover a broad frequency range from 1kHz to 30kHz (Figure 3.9A). The majority of harmonic template neurons in our study were found in primary auditory cortex (A1) and distributed in the entire frequency range of tonotopic map by BFs (Figure 3.9B). We also found a few neurons in the high frequency rostral region of the core area (Figure 3.10). There was no difference in terms of spectral selectivity so they were also included in our analysis.

3.2.4 Harmonic template neurons are sensitive to spectral regularity

The periodicity in spectrum is essential for harmonic complex tones. Frequencies, which are not integer multiples of a common fundamental frequency do not elicit a clear pitch sensation as harmonic complex tones. More than one pitch may be perceived or the complex tone may sounds like noise. In music, a consonant sounds more pleasant than a dissonant. Here we investigated whether harmonic template neurons were sensitive to spectral regularity by using spectrally jittered complex tones. The harmonic at BF was fixed while all other components were independently and randomly drawn from a uniform distribution with a mean at corresponding harmonics and a standard deviation proportional to a certain jitter level (Figure 2.3). As the jitter level increased, the spectrum became more irregular and the pitch of the complex tone was less clear. For each jitter level, 25 stimuli were generated with different spectral structures. The preferred harmonic complex tone was also repeated 25 times to estimate the internal noise of neural response (Figure 2.3). For harmonic template neurons, firing rate decreased when the spectral regularity was disrupted by jitter (Figure 3.11A and B). The decrease in firing rates was significant for jitter levels larger than 10% by comparing the firing rate distribution at individual

jitter level with the rate distribution to repeated harmonic complex tone (Figure 3.12, Wilcoxon signed-rank test, $p < 0.0001$).

Another way to visualize the firing rate change with the spectrum regularity was to arrange the jittered complex tones according to the responses from high to low (Figure 3.13A). As the spectrum became more irregular, the firing rate decreased. An inharmonic index was defined as the average distance between individual component and the corresponding harmonic in the original harmonic complex tone.

$$\text{Inharmonic Index} = \frac{\sqrt{\frac{(f_1 - f'_1)^2 + (f_2 - f'_2)^2 + \dots + (f_n - f'_n)^2}{n - 1}}}{f_0}$$

Where f_i represents the frequency from the harmonic complex tone and f'_i is the jittered component for that harmonic. The inharmonic index was negatively correlated with the firing rate (Figure 3.13A, bottom figure). As a comparison, a non-harmonic template neuron's responses to jittered complex tones were shown in Figure 3.13B. There was no correlation between the inharmonic index and the firing rate. Non-harmonic template neurons in general were not sensitive to spectral regularity.

Additional test was done for a few harmonic template neurons with the compressed/stretched complex tones. In this test, the harmonic at BF was still fixed while the distance between adjacent harmonic was increased (stretched tones) or decreased (compressed tones) (Figure 3.14A). All 9 harmonic template neurons showed a decreased in firing rate for both stretched and compressed tones. The decrease in firing rate was significant when the change was at least 8% of the f_0 (Figure 3.14B). This result indicates that harmonic template neurons are not only selective to the local spectral regularity but a very specific spectral space.

3.2.5 Inhibition plays a role in shaping harmonic templates in A1

The emergence of harmonic templates could result from a combination of sharply tuned harmonically related inputs. Previous studies have already showed narrowly tuned excitatory responses in A1 (Abeles and Goldstein 1972, Pelleg-Toiba and Wollberg 1989, Sadagopan and Wang 2008). However, excitatory inputs alone cannot distinguish stimuli with overlapping spectral contents. For example, the harmonics of f_0 are also the even harmonics of $f_0/2$. A harmonic template only based on excitatory inputs would match both f_0 and $f_0/2$. However, harmonic template neurons in our study showed distinct responses to the preferred f_0 (Bf_0) and $Bf_0/2$. The firing rate to $Bf_0/2$ was much smaller than the firing rate to Bf_0 (Figure 3.15A). Some neurons did not even respond to the $Bf_0/2$ (circles on the x axis in Figure 3.15A) at all. The decreased firing rates indicate that there were inhibitory areas which may contribute to reject stimuli which have overlapping components with the harmonic template.

More direct evidence for inhibitory areas came from neurons with high spontaneous firing rates. Driven rate was used by subtracting spontaneous rate from the firing rate estimated from the spike count during the stimulus. In the shifted complex tone test, the negative driven rates for the inharmonic tones and odd-harmonic tones indicated inhibitory frequency regions between the components of the harmonic template (Figure 3.15B). Our data suggest interaction between excitatory and inhibitory areas in the frequency receptive field give rise to harmonic templates in A1.

3.2.6 Random harmonic stimuli for studying harmonic template neurons

The design of RHS was introduced in Methods. The rationale for RHS was to explore the possible mechanism for generating such harmonic templates. The most intuitive explanation would be an integration of multiple excitatory inputs and inhibitory inputs given the previous

observations. Because of the limitation of the extracellular recording, the excitation and inhibition cannot be separated in our data. However, by using a linear weighting model (Yu and Young 2000), we could estimate the distribution of possible excitatory and inhibitory inputs. For harmonic template neurons, the f_0 s for random harmonic stimulus usually is chosen to be $Bf_0/2$ or $Bf_0/4$ so that there are components sampling the area between two adjacent harmonics in the preferred template. For each stimulus in the RHS set, the variations in the sound level of each component change the balance of excitatory and inhibitory components and result in a wide variety of response strength (Figure 3.16A and B). For the flat spectrum stimulus, the f_0 is not at the Bf_0 , which make it a sub-optimal stimulus for the neuron. Therefore, the response is low. The estimated linear weights showed an alternating of positive weights at even harmonics and negative weights at odd harmonics (Figure 3.16C). Such linear weight pattern suggests the organized excitatory and inhibitory components in the frequency receptive field creates harmonic templates to accurately detect concurrence of harmonic related frequencies.

Our sampling density for RHS is small, because there will be fewer spikes due to the strong inhibition if there are more components in the inhibitory area. It is difficult to get a good estimation of the weight if the firing rates are too low. For a few neurons, we managed to run the RHS at $Bf_0/4$ so that there were three additional components between adjacent harmonics in the template (Figure 3.17A). We also saw similar pattern of excitatory peaks with inhibitory areas between on the second neuron (Figure 3.17 B and C). We tested RHS at different mean sound levels (20dB and 30dB) but the same standard deviation, the patterns were consistent although the weight strengths were different. This confirmed that the alternating excitation and inhibition pattern were not random, or due to the sparse sampling. It might reflect some general receptive field property of those harmonic template neurons.

Another interesting question to ask is how many harmonics are encoded by harmonic template neurons. In our test, the number of harmonics were determined by the frequency range

and f_0 . In the shifted five harmonics test, we have already showed that harmonic template neurons only responded to harmonics around BF. RHS provided a chance to investigate this question. Although RHS includes a large number of harmonics, the harmonics far away from the BF might not have any contribution to predict neural responses. We started with a single component at BF, adding one more on either side of BF at one time until all harmonics were used. For each time after a component was added, the linear weights were estimated with all the chosen components and the fraction of variance was calculated. As shown in Figure 3.18A, adding more components gradually improved the model performance. The improvement became very small after a few number of components. All the harmonic numbers that improve the model performance significantly were counted as useful harmonics for this neuron (1000 bootstrapping with replacement, t -test, $p < 0.05$). The same approach was applied to 6 out of 19 neurons in our test, which yielded fairly good predication. The number of useful harmonics for all six neurons were plotted in Figure 3.18B. Most neurons encoded at least the two adjacent harmonics of BF. A few neurons encoded more than three harmonics. The alternating excitatory and inhibitory weights pattern was observed for all neurons (Figure 3.18C).

3.3 Discussion

In the experiments described in this chapter, I used harmonic and inharmonic complex tones to systematically examine the selectivity to complex spectral features in auditory cortex. I found harmonic template neurons which maximally respond to a combination of certain harmonics of a preferred f_0 . Those neurons were sensitive to the frequency shift and spectral jitter. They provide templates which can be used for extracting f_0 and detecting harmonic structures in many ecologically relevant sounds.

3.3.1 Compare with previous studies

Previous studies did not find any combination sensitivity neurons in A1 of awake macaque monkey or ferrets (Schwarz and Tomlinson 1990, Kalluri, Depireux et al. 2008). The findings in our study might be due to the sampling bias rather than different animal models. In the macaque study, noise was used for searching neurons. As shown in Figure 3.14, harmonic template neurons have inhibitory inputs which make them non-responsive to broad-band noise. Some neurons have low spontaneous firing rate (Figure 3.2A) and high selectivity to sound parameters. It's easy to miss such neurons by searching with noise. In our study, we used a different searching strategy. The electrode was advanced in steps of 25 μ m. Pure tone and harmonic complex tones would be played before the next movement. Such strategy helped us find harmonic template neurons with extremely low spontaneous rate and high selectivity to acoustic features (Figure 3.2B).

3.3.2 Information transformation and feature detection

The ability to rapidly recognize different sounds in the environment enables us to escape from danger, efficiently communicate within the social group. How the auditory system analyzes the acoustic signals and detects the important spectral features for sound recognition still remains largely unknown. The increased complicity in frequency receptive fields of single neurons along the auditory pathway and the emergence of neurons highly selective to species-specific vocalizations in the auditory cortex (Tian, Reser et al. 2001) suggest that the low-level neural signals for simple frequency partials have been synthesized into integrative representations of sound features, such as pitch and timbre, and eventually form coherent representations of 'objects'. In our study, the finding of harmonic template neurons which exhibit combined selectivity to harmonically related frequency partials suggests A1 is an important intermediate processing stage for extracting sound features. Such feature processing is not unique for the auditory system. In the visual system, the retinal

signals are transformed into orientation information, curvature, and subsequently contour fragments with certain combination of curvature and orientation (Connor, Brincat et al. 2007). Another important question to ask next is where harmonic templates emerge in auditory cortex or in subcortical areas such as inferior colliculus (IC). We showed that inhibitory inputs in addition to the sharply tuned excitatory inputs for harmonic template neurons, given the sharper frequency tuning and local microcircuits with inhibitory neurons and long range projection from other areas in the tonotopic map, auditory cortex may be a good place for the formation of harmonic templates. However, combination sensitive neurons were found in auditory cortex (Suga, O'Neill et al. 1979) as well as some subcortical areas such as IC (Leroy and Wenstrup 2000). It will be interesting to test neurons from inferior colliculus to see whether at least part of the information of harmonic template already encoded by IC neurons. In addition, it will be interesting to test the excitatory and inhibitory inputs for harmonic template neurons using other techniques such as whole cell recording and distinguish the pre-synaptic inputs from intra-cortical connections and from direct thalamo-cortical inputs. It will help us understand more of the network architecture for generating such complex receptive fields.

3.3.3 Harmonic template neurons and pitch neurons

The harmonic template neurons reported in our study are different from pitch-selective neurons in previous studies (Bendor and Wang 2005, Bendor, Osmanski et al. 2012), although both types of neurons respond to harmonic complex tones even if they don't respond to individual component. First of all, pitch selective neurons respond to pure tone and harmonic complex tones with a f_0 near its BF. They respond to stimuli with equal pitch regardless of different spectral contents. In other words, the tuning of f_0 for pitch neurons is along some perceptual dimension instead of physical dimension. For harmonic template neurons, the selectivity to f_0 comes from the combination sensitivity to concurrent frequencies with harmonic spacing. The BF of

harmonic template neurons can either be the preferred f_0 or the harmonic of the preferred f_0 . Secondly, in the shifted harmonic test, pitch neurons showed a higher firing rate to harmonic complex tones with lower harmonics but lower firing rate to those with higher harmonics suggesting a sensitivity to pitch salience, because harmonics complex tones with lower harmonics have a higher pitch salience. For harmonic template neurons, they respond maximally to harmonics including the BF component. The preference for harmonic numbers depends on the ratio between their BFs and the preferred f_0 . Third, pitch neurons are organized by the preferred f_0 s because f_0 s equal their BFs. They are located in low frequency boarder between A1 and rostral region of the core area (R) and the preferred f_0 s are mostly below 1kHz. Harmonic template neurons are scattered in the entire frequency range of A1 (1kHz to 30kHz) and organized by their BFs even the preferred f_0 s can be as low as 400Hz. We did find a few harmonic template neurons in the high frequency rostral region as well (Figure 3.10A). In conclusion, Harmonic template neurons are not pitch neurons. However, harmonic encode absolute frequencies which can be used for extracting pitch. They also preserve the spectral information which can be used for coding timbre. In the hierarchical information pathway, harmonic template neurons might be a pre-processing stage for pitch neurons.

3.3.4 Harmonic resolvability and pitch computation

The peripheral auditory system can be modeled as a series of auditory filters centered from low to high frequencies along the basilar membrane. The filters become broader with frequency whereas the frequency spacing of a harmonic complex tone remains constant (Glasberg and Moore 1990). As a result, lower harmonics are separated into different filters to be ‘resolved’ while higher harmonics can fall into the same filter to be ‘unresolved’. Lower harmonics and higher harmonics can have the same pitch but differ in many other aspects of perception. For example, it’s easier to discriminate the f_0 s of complex tones containing lower harmonics than the

f0s of tones comprising of only higher harmonics (Houtsma and Smurzynski 1990). In addition, there is perceived pitch shift for resolved harmonics but not unresolved harmonics if all harmonics are shifted in frequency (Moore and Moore 2003). It has been proposed that harmonic templates can be used for pitch extraction of resolved harmonics (Goldstein 1973, Terhardt 1974, Shamma and Klein 2000). Our study provides biological evidence for such templates. Pitch and harmonic numbers can be estimated by matching an unknown complex tone to different templates. However, the final decision of pitch requires additional computation because such estimation has certain ambiguity because of the noisy frequency inputs (Goldstein 1973). For example, the example neuron in Figure 3.1 responded largely to two f0s because both the 4, 5, 6th harmonics of the first f0 and the 5, 6, 7th harmonics of the second f0 match the template.

The harmonic template neurons found in our study are biological evidence that how pitch and harmonic numbers are estimated from resolved harmonics. Due to the different hearing range and cochlear structures across species, harmonic template neurons from different animal models might differ in the frequency range and resolved harmonic numbers. We need take into account the physiological difference across species when we use knowledge from animal studies to understand human perception.

3.3.5 Implications in auditory perception

The harmonic template neurons found in our studies cover a broad frequency range, not limited to marmoset vocalization range. The four major marmoset calls have the first harmonic between 3kHz to 8kHz (DiMattina and Wang 2006). This result suggests a general principle of the auditory processing which could take advantage of the probabilistic structures of natural sounds. The following question is how harmonic templates emerge in the auditory processing? One possibility is that harmonic templates form during early development due to the exposure to

the natural acoustic environment full of harmonic sounds. Some studies have shown how harmonic templates can be generated from the exposure to the rich acoustic environment during early development or even to any broadband sounds (Shamma and Klein 2000). One other possibility is our auditory system have evolved to efficiently detect harmonic sounds by using harmonic templates. The ability to quickly and accurately identify different sound sources in the environment is crucial for survival as it is important for detecting predators, hunting and localizing food, reproduction and communication. Future studies across different species are necessary to fully answer this question.

Most sounds from musical instruments are harmonic as well. Harmony is an important discipline in western music, which delimits the structure of chords based on the frequency ratios. Most listeners report the similar consonance ordering of chromatic scale tone combinations (Malmberg 1918, Krumhansl 1990). It has been shown the preference for harmonic spectra was consistently correlated with preferences for consonant over dissonant chord across more than 250 subjects (McDermott, Lehr et al. 2010). Another study has also showed that the statistical acoustics of human speech sounds can successfully predict some widely shared aspect of music perception, which suggests music can be a side effect of auditory mechanism that evolved for other functions(Schwartz, Howe et al. 2003). Therefore, the findings in our study might imply some neural mechanism underlying music perception.

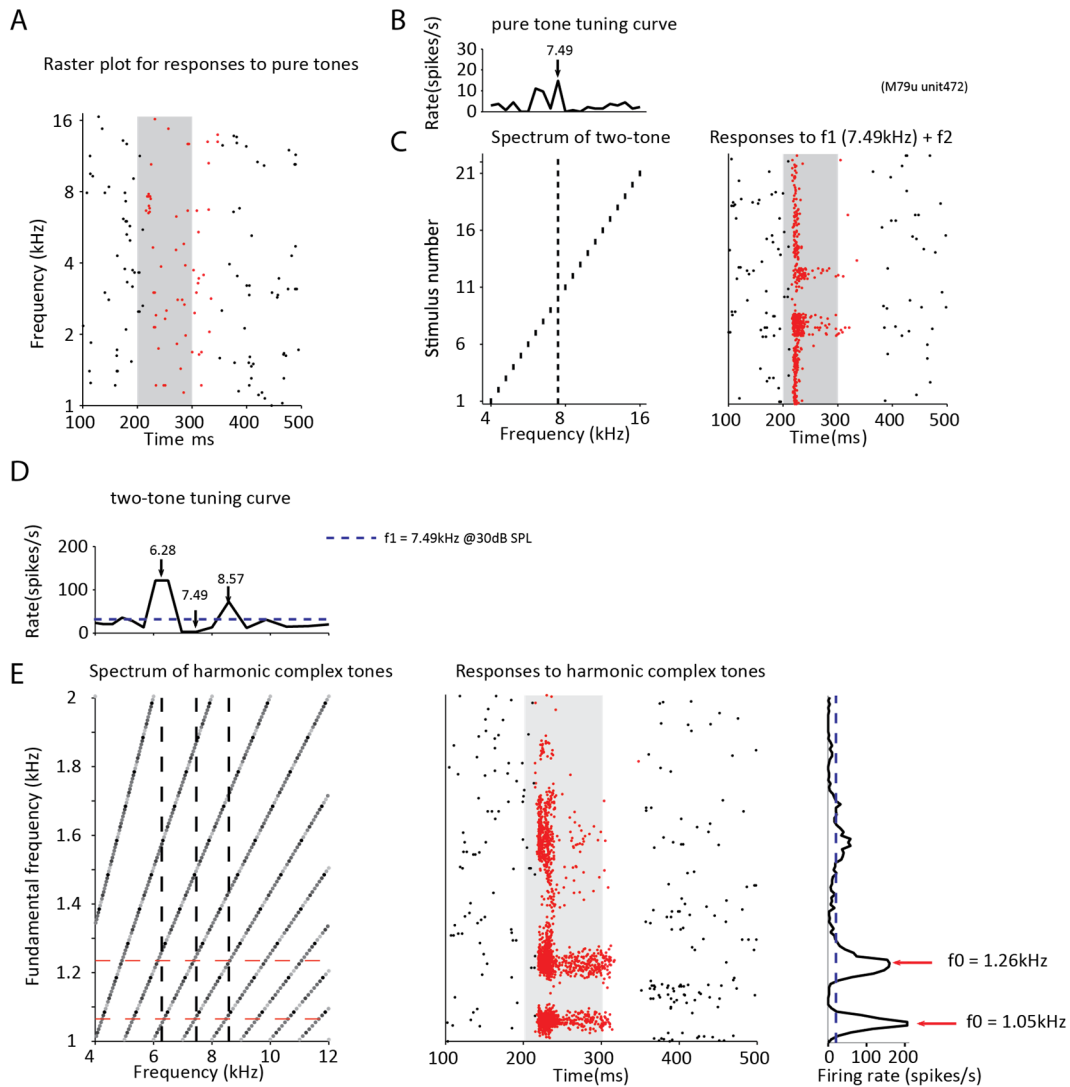


Figure 3.1 An examples of a neuron which preferred to harmonic complex tones.

A, the raster plot of an example neuron responding to pure tone at different frequencies at 40dB SPL sound level).

B, frequency tuning curve to pure tones from the raster plot in A.

C, left: spectrum of the two-tone stimuli. Right: raster plot of neural response to two-tone stimuli.

D, tuning curve to the two-tone stimuli. x axis was the frequency of the second tone. The dashed line was the neural response to BF tone alone. The solid line was the firing rate to BF played simultaneously with a second tone.

E, the raster plot (left) and tuning curve (right) of the neuron's response to harmonic complex tones at different fundamental frequencies. The dashed line indicates the maximal firing rate in pure tone tuning (shown in A).

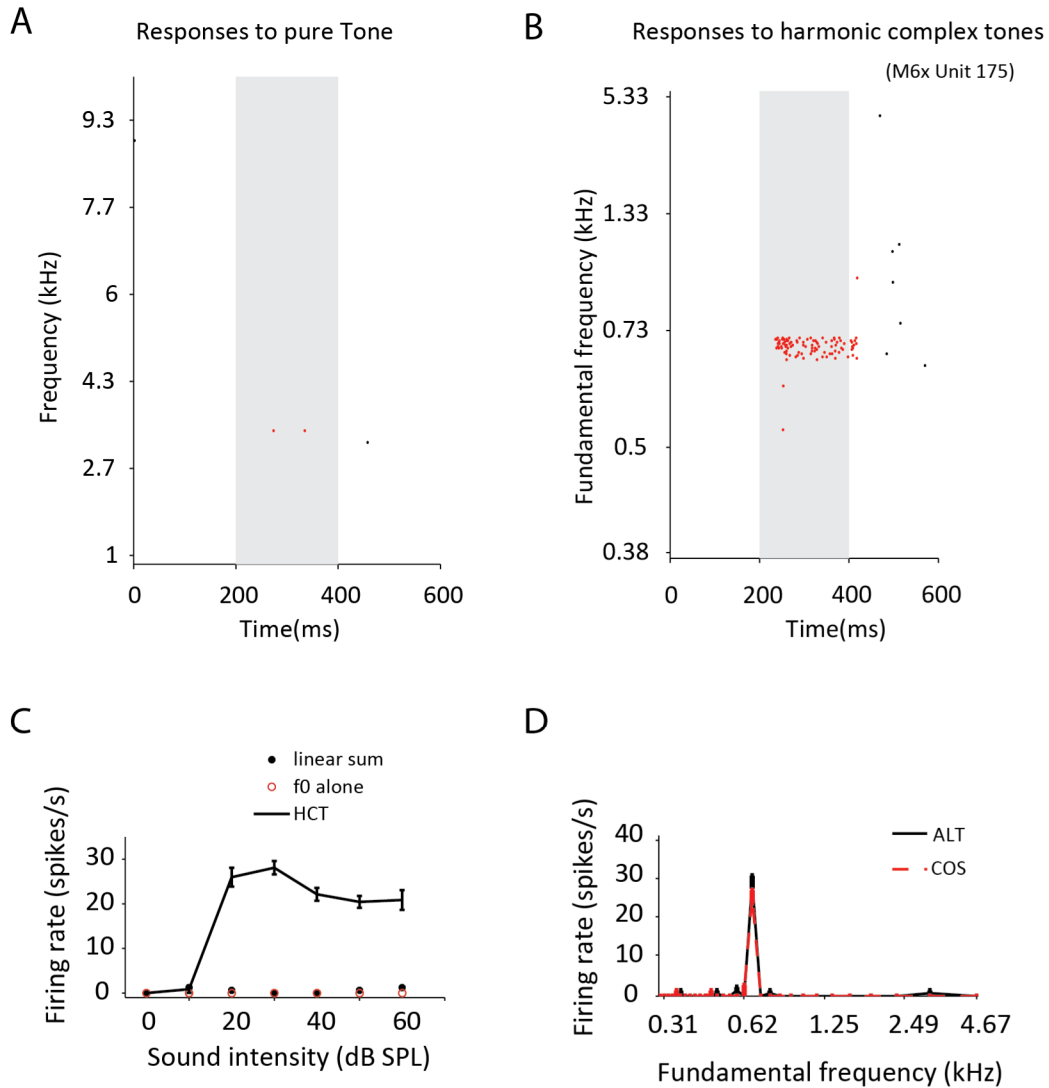


Figure 3.2 Another example of harmonic template neuron.

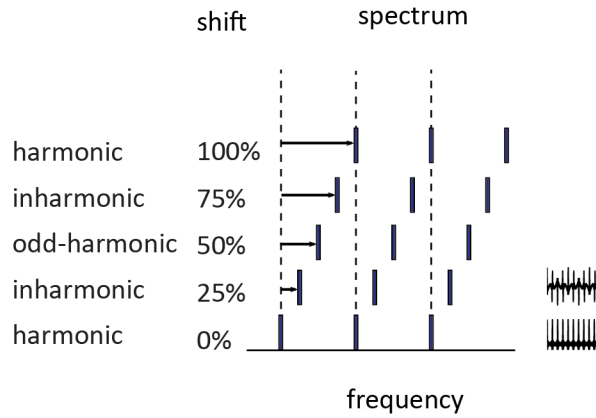
A, the raster plot of the example neuron's response to pure tones at different frequency at 40dB SPL sound level.

B, Responses to harmonic complex tones. All harmonic tone have equal amplitude of 40dB SPL per component.

C, the neuron's responses to the preferred f0 at different sound levels (black line), to f0 alone (black dot) and the linear sum of responses to individual harmonics (red circles).

D, The tunings to f0s with different adding phases (cosine phase: COS and alternating phase: ALT).

A



B

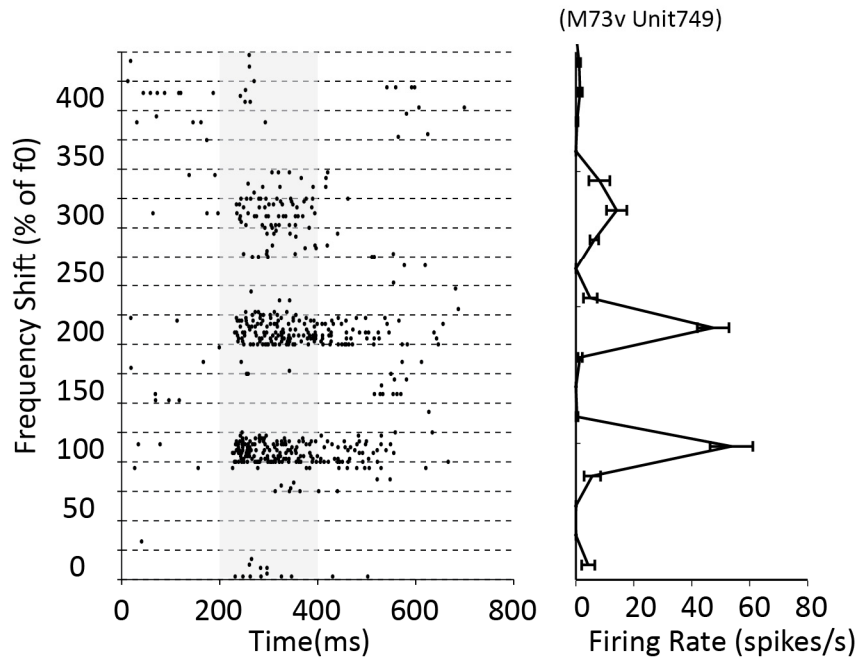


Figure 3 Harmonic template neurons showed periodic response patterns to shifted complex tones.

A, a diagram of shifted complex tones in spectrum.

B, the responses of a harmonic template neuron to the shifted complex tones shown in A.

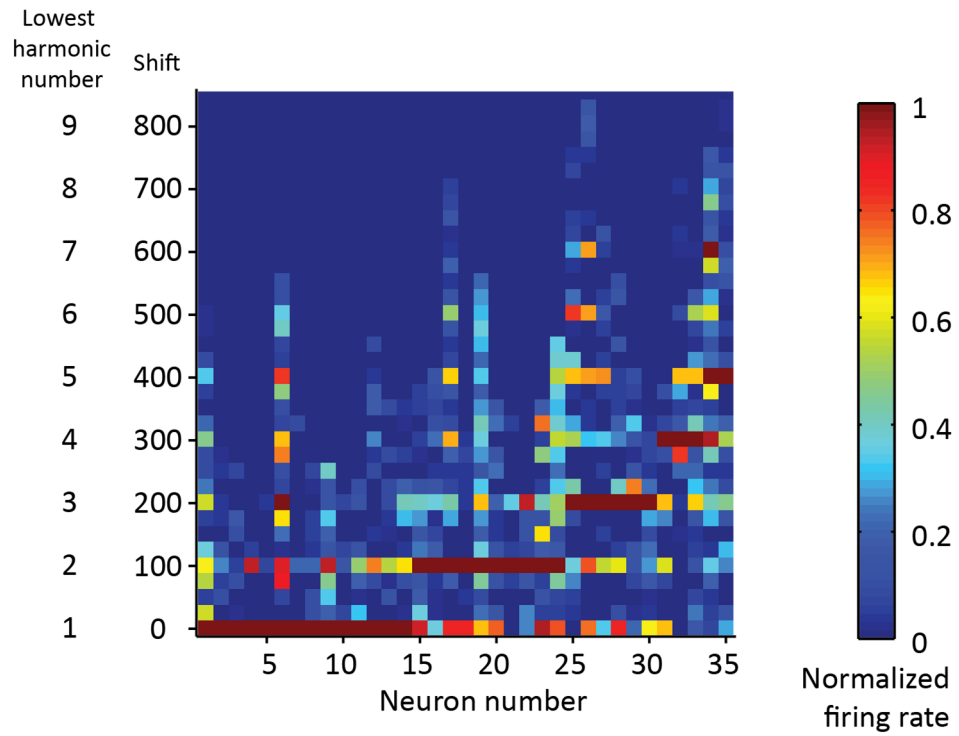


Figure 3.4 Normalized responses of 35 different harmonic template neurons to the shifted complex tones. Each column is the responses from one neuron. The color indicates normalized firing rates from 0 to 1. Different neurons are aligned by the location of the maximum firing rate.

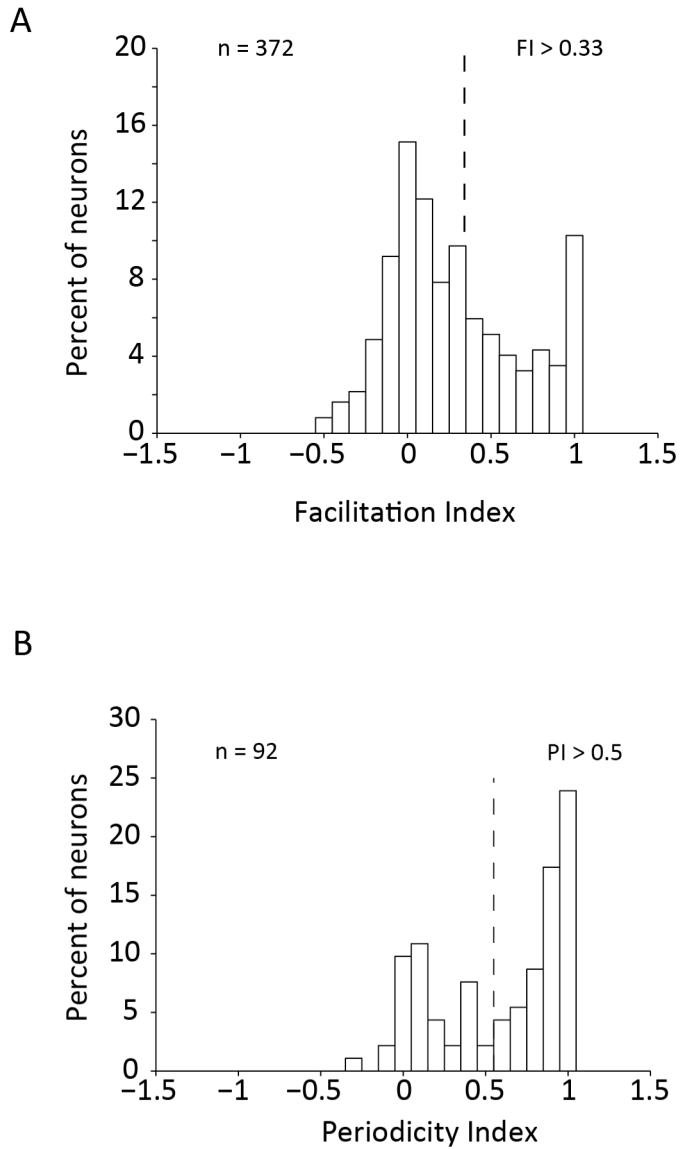


Figure 3.5 Two criteria for harmonic template neurons.

A, the distribution of facilitation index for all 372 neurons. A FI larger than 0.3 will be the first criterion for harmonic template neurons.

B, the distribution of periodicity index of all 92 neurons with a FI larger than 0.3. The second criterion for harmonic template neuron is: have a periodicity index larger than 0.5.

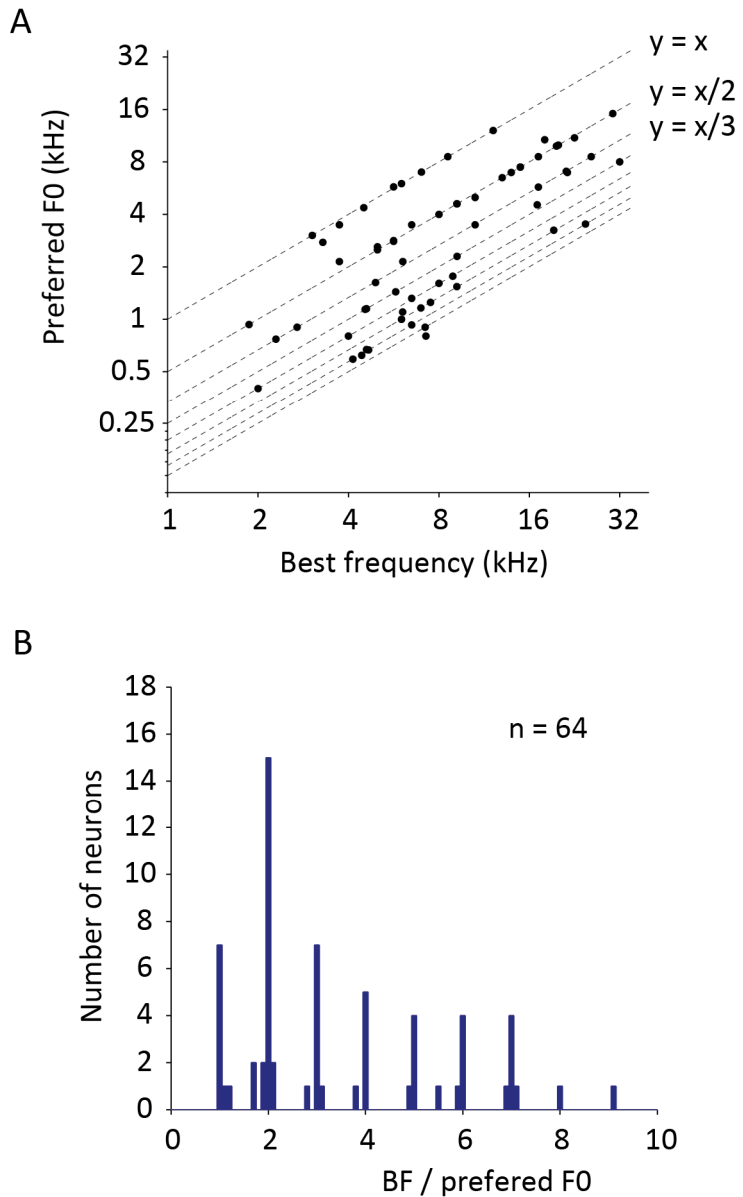
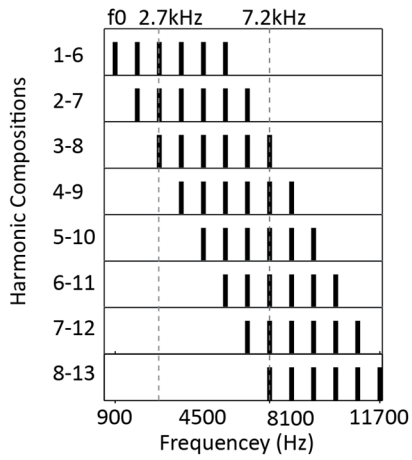


Figure 3.6 The relationship between BF and the preferred f0 for harmonic template neurons.

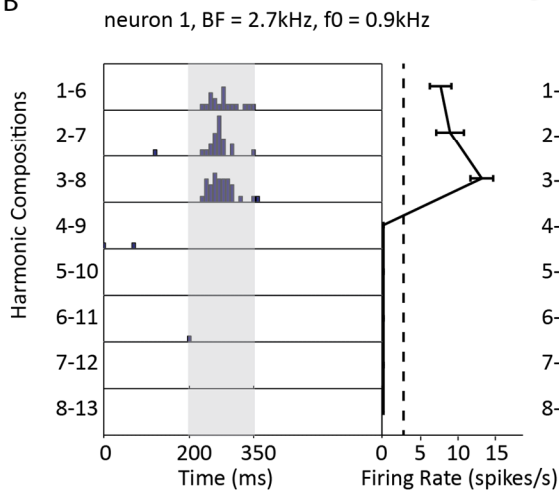
A, a plot of BFs against preferred f0s of all 64 harmonic template neurons in logarithm scale. The preferred f0s are linearly proportional to BFs.

B, the distribution of ratios between BF and Bf0.

A



B



C

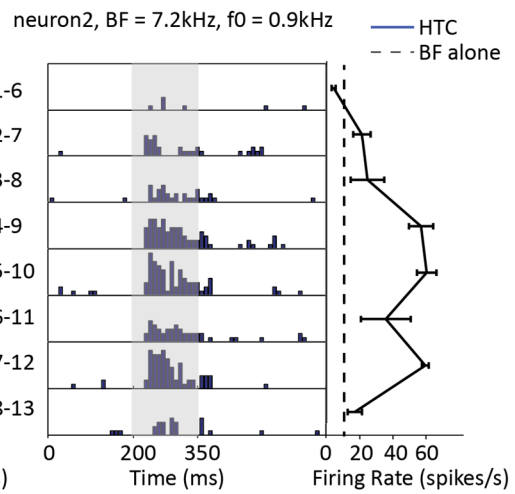


Figure 3.7 The different responses of two neurons to the same harmonic complex tones of the same f_0 but different compositions.

A, a diagram of the spectrum of all harmonic complex tones and the BFs of two neurons.

B, responses of Neuron1 which has a BF at the third harmonic of the preferred f_0 .

C, responses of neuron2, the BF of which is the eighth harmonic.

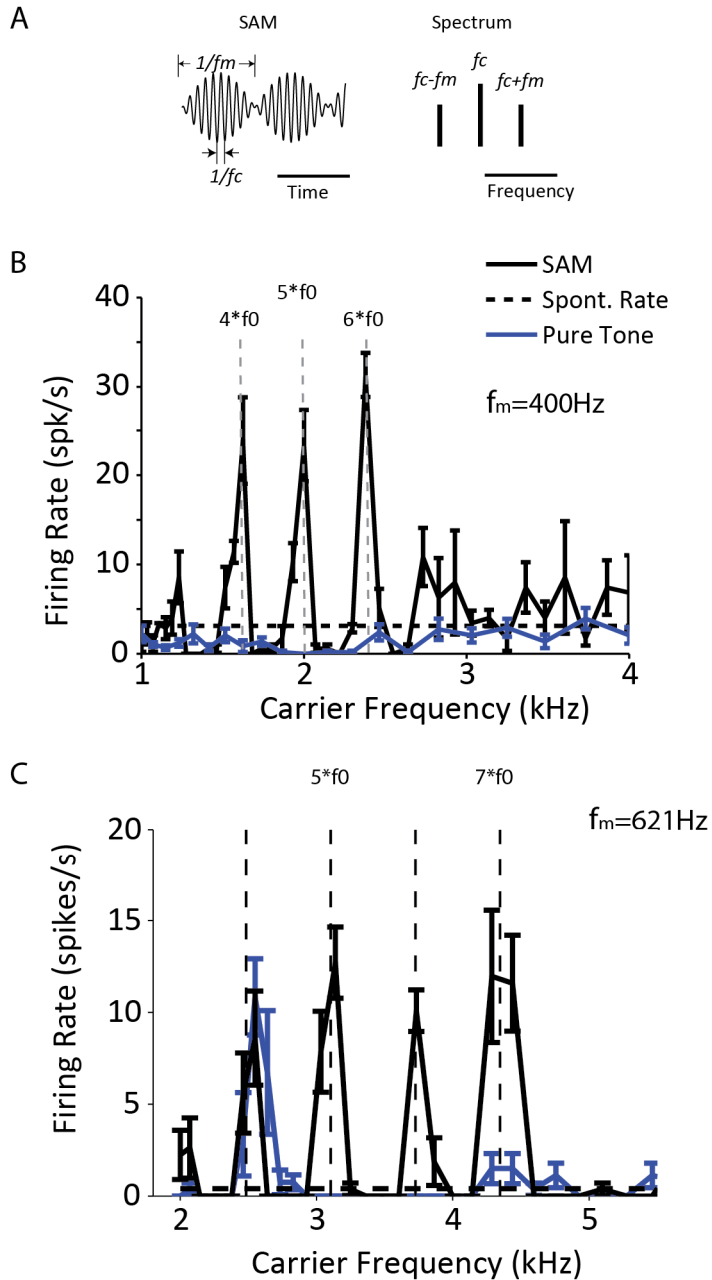


Figure 3.8 Harmonic template neurons' responses to sAM tones

A, a diagram of an amplitude modulate tone in both time domain and frequency domain.
 B, the responses of an example neuron to pure tones and sAM tones at different carriers.
 C, another example of a harmonic template neuron showed responses to carrier at multiple harmonics when it is modulated at 621Hz.

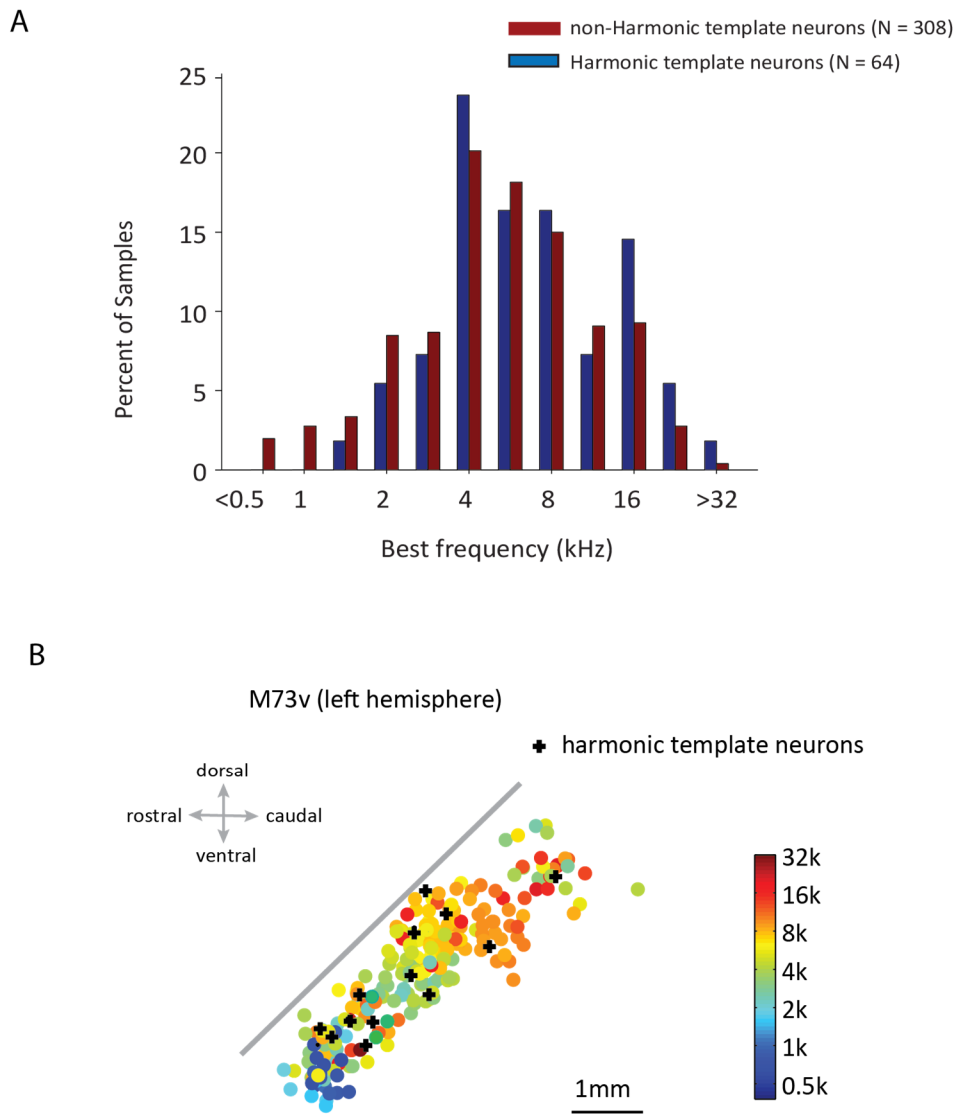


Figure 3.9 Distribution of harmonic template neurons.

A, the distribution of estimated BFs of harmonic template neurons.

B, tonotopic map of marmoset auditory cortex and location of harmonic template neurons.

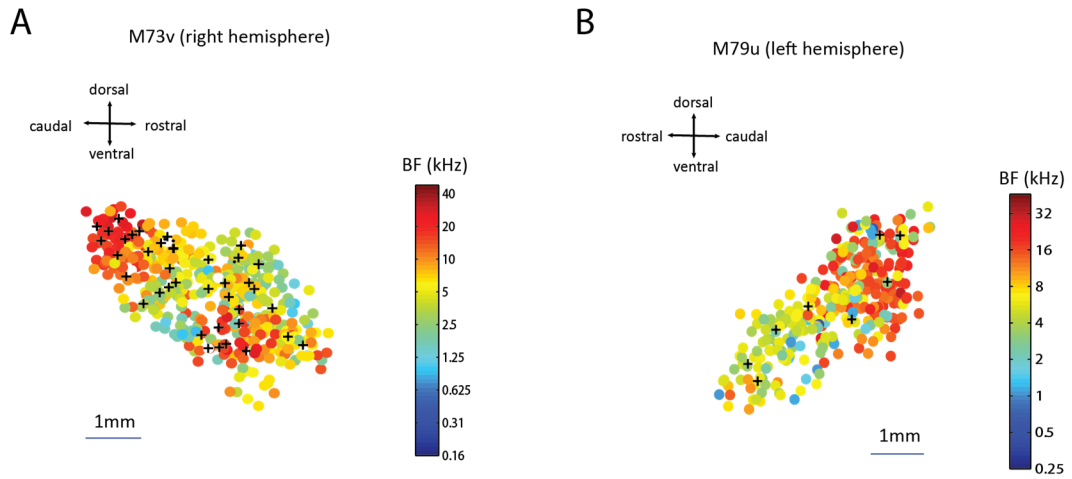


Figure 3.11 Distribution of harmonic template neurons in auditory cortex

A, tonotopic map and locations of harmonic template neurons from the right hemisphere of monkey M73v.

B, tonotopic map and locations of harmonic template neurons from the left hemisphere of monkey M79u.

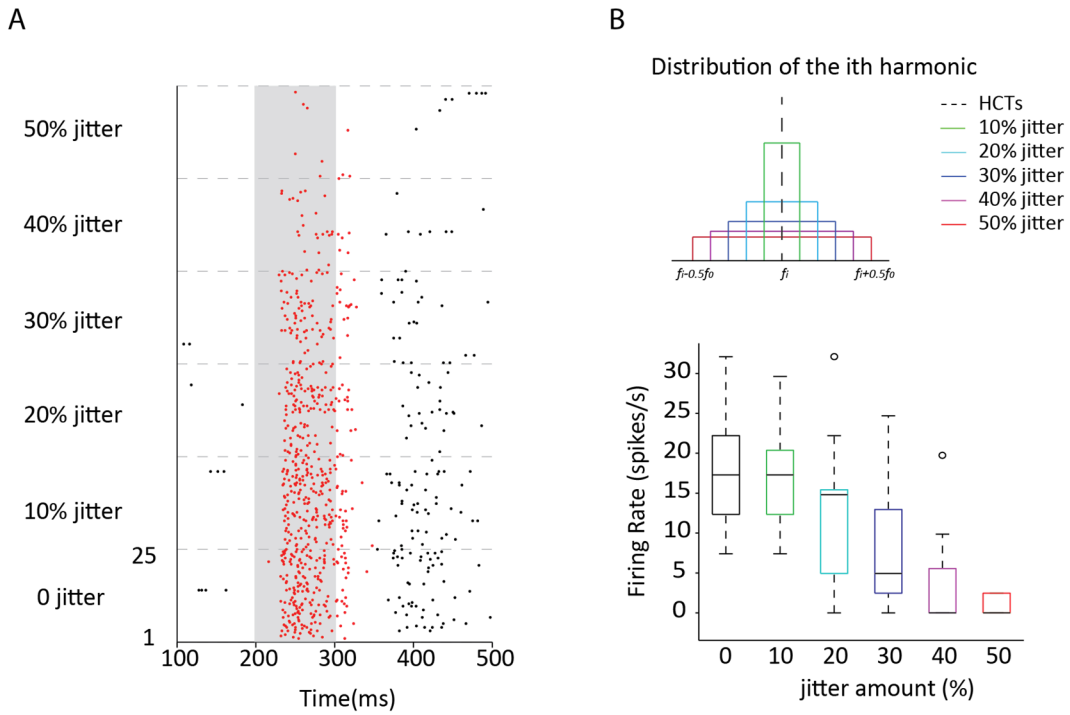


Figure 3.11 A harmonic template neuron's responses to jittered complex tones.

A, the raster plot of a harmonic template neuron's response to spectrally jittered harmonic complex tones.

B, boxplot of the distribution of firing rate for 25 stimuli at each jitter level.

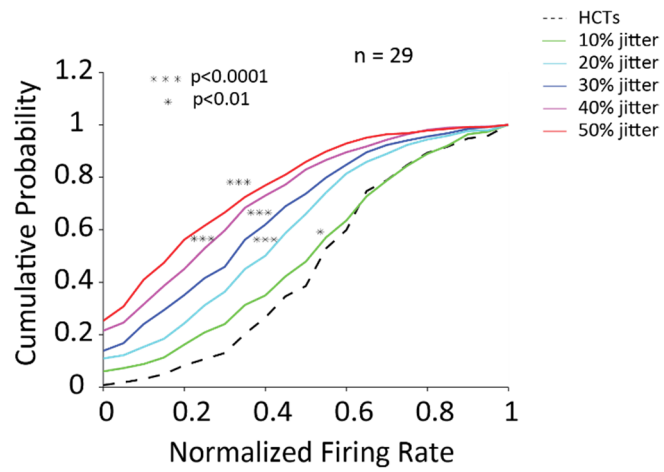


Figure 3.12 The normalized firing rate distributions for all jitter level of 29 harmonic template neurons. For each jitter level, accumulative probability distribution is used. The firing rate is significantly decreased when the jitter level is larger than 10% (t-test, $p < 0.001$).

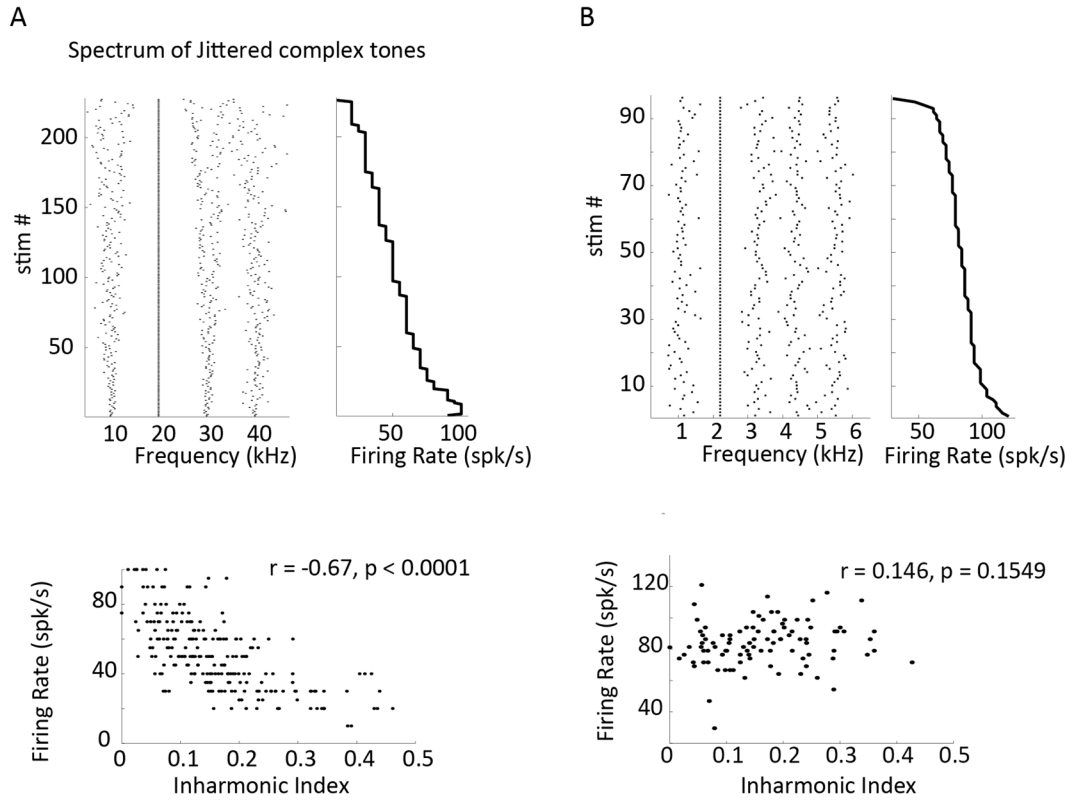


Figure 3.13 Examples of a harmonic template neuron and a non-harmonic template neuron

A, the spectra of jittered complex tones sorted by firing rates of a harmonic template neuron (top). The correlation between firing rate and inharmonic index (bottom)

B, a non-harmonic template neuron did not show sensitivity to spectral regularity.

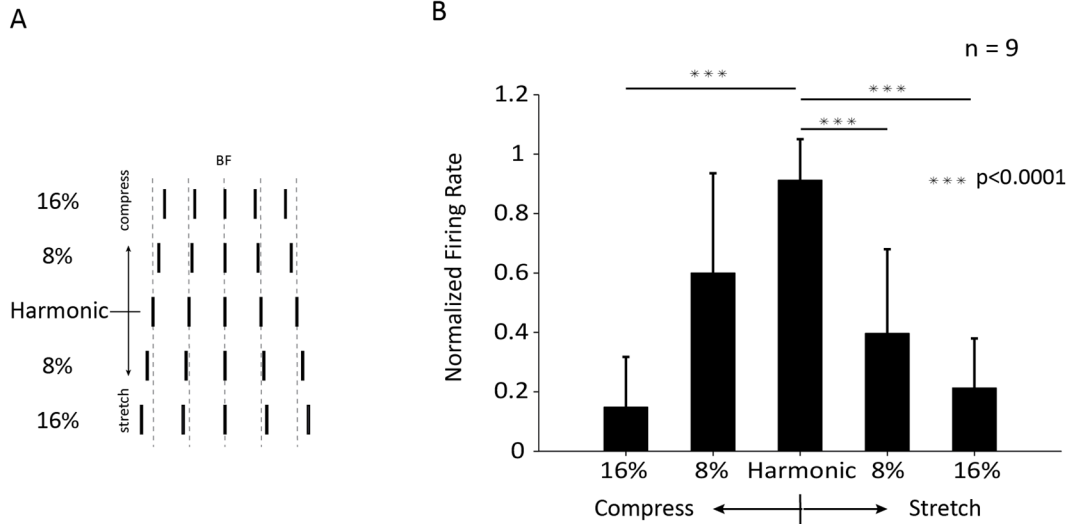


Figure 3.14 Compressed and stretched complex tones test.

A, the spectra of the compressed and stretched complex tones.

B, the average firing rates of 9 neurons to those complex tones shown in A.

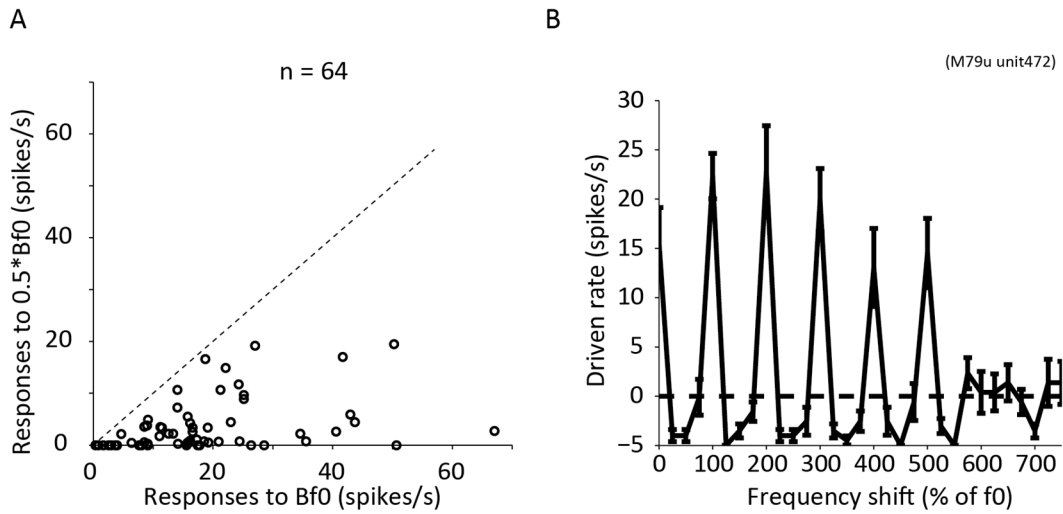


Figure 3.15 The role of inhibition in shaping harmonic templates.

A, a comparison of the responses to the preferred f_0 (Bf_0) and responses to $Bf_0/2$ for all harmonic template neurons.

B, the responses of a harmonic template neuron to shifted complex tones plotted in driven rates. Driven rates were used by subtracting spontaneous rate from the firing rates estimated from the spike count during the stimulus.

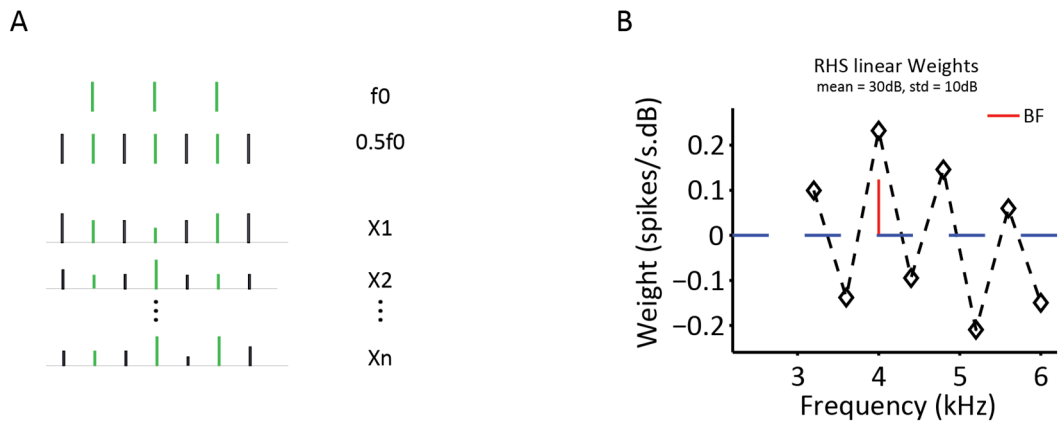


Figure 3.16 RHS and linear weights estimations

A, a diagram of the RHS design for a harmonic template neuron at the sample frequency of $f_0/2$.
 B, the estimated linear weights at different harmonics of $f_0/2$.

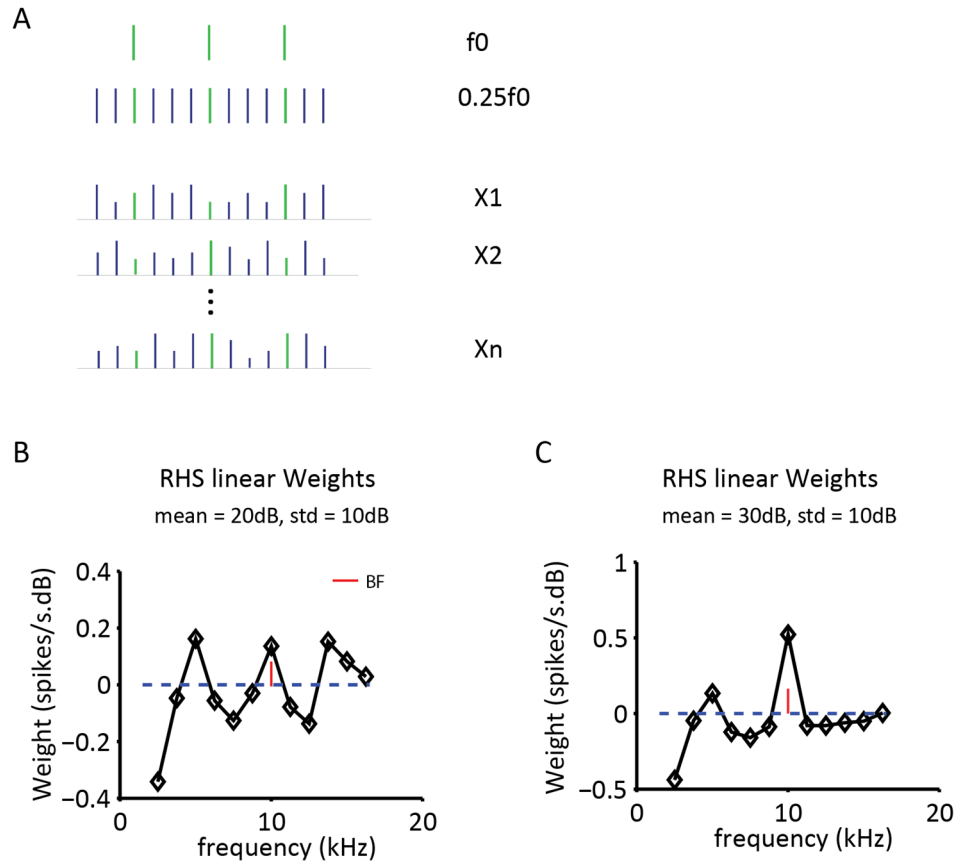


Figure 3.17 RHS and linear weights estimations

A, a diagram of the RHS design for a harmonic template neuron at the sample frequency of $f_0/4$.
 B, the estimated linear weights at 20dB mean sound level with 10dB standard deviation.
 C, the estimated linear weights at 30dB mean sound level with 10dB standard deviation.

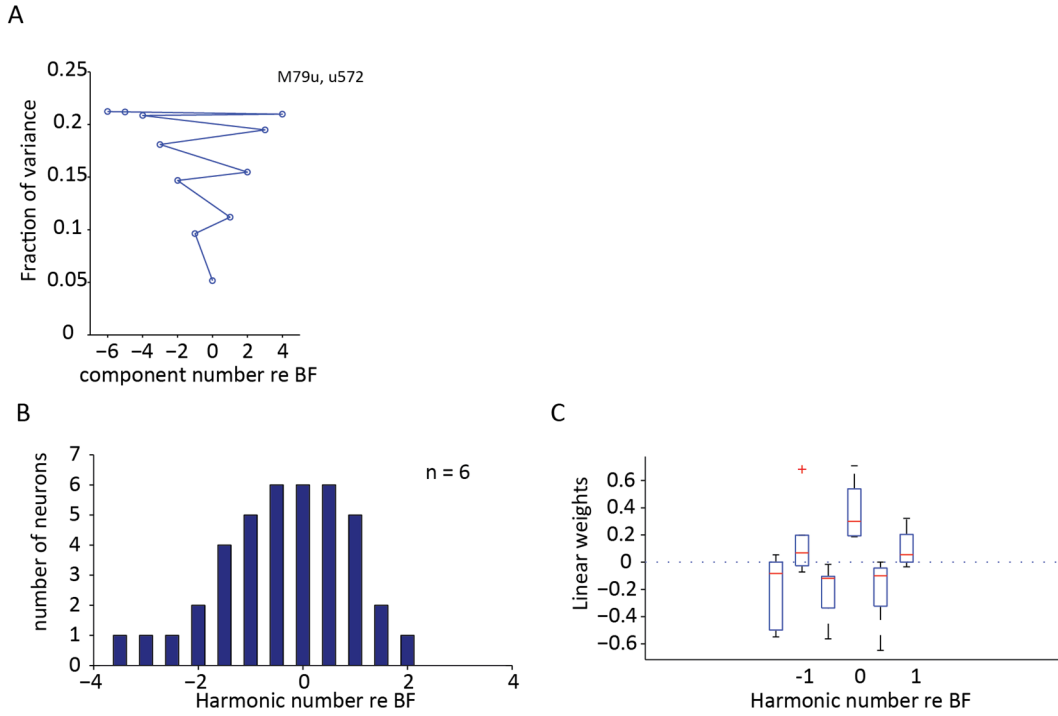


Figure 3.18 An estimation of number of harmonics coded by harmonic template neurons

A, a demonstration of how the model performance changes with more frequency components.

B, a summary of all frequency components for all 6 neurons that are important for predicting neural responses.

C, the distribution of the linear weights of five harmonics around BF at the sample of $f_0/2$.

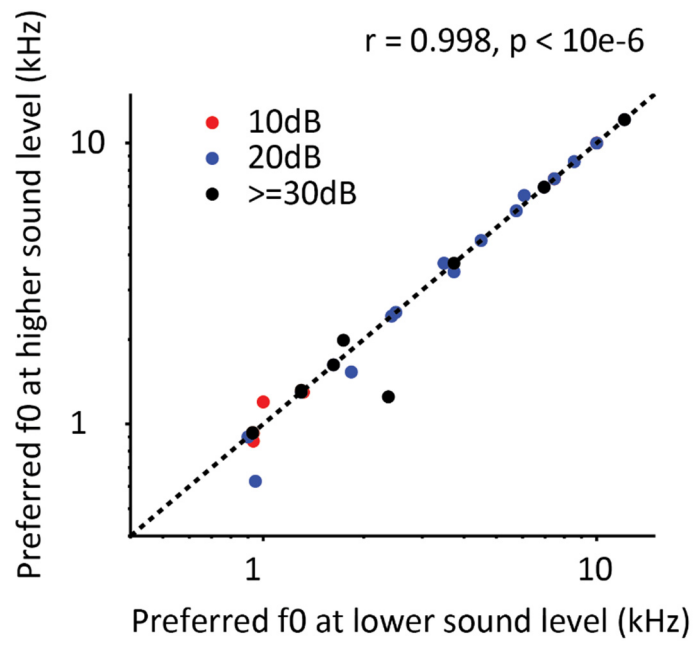


Figure 3.19 The preferred f0 measured at different sound levels for harmonic template neurons

CHAPTER 4:

4. A distributed harmonic process in primary auditory cortex

4.1 Introduction

In chapter 3, I describe a specialized subgroup of neurons in A1: harmonic template neurons. However, this subgroup of neurons are only a small portion of all neurons sampled in A1 (64/372). In this chapter, I analyze at the population level all other neurons to understand how harmonic complex tones are represented.

4.1.1 Spectral analysis and perception

The spectral analysis of a time-varying waveform of the pressure waves can be referred to Fourier analysis in physics. Any sound can be written as a summation of a series of sinusoid waveforms. For our auditory system, the definition of spectral analysis is more complicated. In order to accurately and identify a sound source or to have an efficient communication, the analysis of the acoustical signal has to be done on multiple scales because most natural sound and conspecific vocalizations are broadband with energy distributed at different frequency bands. If we use speech as an example, in order to discriminate a vowel, our auditory system has to analyze the space between each frequency and the location of peaks in the spectral profile. The space is associated with pitch, which carries information of a talker, such as the gender. The location of peaks are formants, the information bearing elements for distinguish different vowels.

Although the auditory filter model has been widely used to understand the spectral processing in perception (Moore 1995), there are many conditions for complex stimuli that the simple peripheral filter model cannot account for performances, such as profile analysis (Bernstein and Green 1987, Bernstein and Green 1987), harmonic grouping (Bregman 1990), co-modulation masking release (Hall, Haggard et al. 1984). In those cases, the perception of a frequency can be influenced by another frequency component located far beyond the filter bandwidth. Those results suggest that the auditory system integrate information across different frequency channels and is sensitive to the global spectral patterns. However, our knowledge of the neural mechanism underlying such cross-channel processing is very limited. Harmonic complex tones can be useful to explore the possible mechanism because they are sophisticated enough to involve the cross-channel processing yet simple enough to systematically control the parameters.

4.1.2 Neural representation of harmonic complex tones at auditory nerves

In the peripheral auditory system, which can be approximated as filter banks organized topographically by the center frequencies, a complex sound is decomposed into its frequency partials (Zhang, Heinz et al. 2001). When a harmonic complex tone is presented, individual AN fibers had a larger firing rate when its characteristic frequencies (CF) was near a low-order harmonic and low firing rate when CF fell in between two harmonics (Cedolin and Delgutte 2005). A model incorporating a band-pass filter could predict the response trend to f_0 changes.

Although previous study has shown a robust representation of vowels combining rate, place and temporal information of populations of AN fibers (Young and Sachs 1979), the spectral information of resolved harmonics at single AN fibers degraded at higher sound level. It's still

largely unknown how neurons in the later stage of auditory system read out the information from AN fibers and represent harmonic complex tones differently.

4.1.3 Frequency receptive field of neurons in auditory cortex

In the primary auditory cortex, single neurons are highly selective to frequencies (Suga 1965, Abeles and Goldstein 1972, Pelleg-Toiba and Wollberg 1989, Sadagopan and Wang 2008) and tonotopically organized. However, the increased complexity in local neural network connections in auditory cortex enables single neurons to integrate excitatory and inhibitory inputs from a broader frequency range. There were evidences that stimuli outside the classical frequency receptive field could modulate a neuron's responses to simple or complex stimuli within the receptive field (Shamma and Symmes 1985, Nelken, Prut et al. 1994, Kadia and Wang 2003, Sutter and Loftus 2003, Qin, Sakai et al. 2005, Sadagopan and Wang 2010). Therefore, single-tone responses are not adequate to characterize how cortical neurons process harmonic complex tones nor their functional significance in sound feature extraction. On the other hand, such extra-classical receptive fields make A1 neurons good candidates for integrating information across frequency and form an integrative representation of the global spectral feature.

4.1.4 The functional organization of auditory cortex for spectral processing

The tonotopic frequency organization in A1 has been confirmed from many studies on many species. It's still under debate whether there is other functional structure beyond the tonotopic map for the purpose of spectral processing. Different functional areas were found in the auditory cortex of mustached bats, which echo-locate by emitting biosonar signal and measure the properties of echoes. There is a CF-CF region, where neurons only respond to tone pairs with specific frequency combinations corresponding to the range of encountered Doppler shifts (Suga,

1988). In cat auditory cortex, a bandwidth map was found in dorsal A1: sharply tuned neurons at the dorsal ventral center, broadly tuned neurons were located more dorsally (Schreiner, 1992). In addition, multi-peaked neurons were found in the dorsal zone (DZ) (Sutter and Schreiner, 1991).

Such functional maps were rarely reported from other species. Some studies showed that single neurons in A1 can have a large diversity in the frequency and level selectivity, even in the same isofrequency column (Schwarz and Tomlinson 1990, Hromadka, Deweese et al. 2008, Bandyopadhyay, Shamma et al. 2010, Rothschild, Nelken et al. 2010). However those studies did not test complex tone or broadband stimuli as the studies on cats. One important question will be whether neurons with different tone response properties, for example monotonic or non-monotonic rate level functions (Brugge and Merzenich 1973, Pfingst and O'Connor 1981, Sadagopan and Wang 2008), process harmonic complex tones differently.

In this chapter, we systematically characterize single neurons' responses to tones and harmonic complex tones covering a broad f_0 frequency range at a moderate sound level to investigate how harmonic complex tones are represented in A1.

4.2 Results

All analyses in this chapter are based on 372 well isolated single units from primary auditory cortex from three hemispheres of two monkeys. All neurons had a significant increase in firing rate to at least one harmonic complex tone.

4.2.1 Frequency selective neurons in A1

Neurons which showed a single peak in tone frequency tuning (Figure 4.1A) could exhibit multiple peaks if the average firing rate was plotted against the f_0 s (Figure 4.1B). Because

when the fundamental frequency (f_0) changes, the corresponding harmonics move in and out of the frequency receptive field (Figure 4.1C). We used a dimensionless ratio of single neuron's best frequency (BF) to f_0 (harmonic number) (Cedolin and Delgutte 2005, Fishman, Micheyl et al. 2013), because the selectivity to harmonic complex tones was constrained by the BF. Harmonic numbers vary inversely with f_0 s. The BF became one of the harmonics of the stimuli when the harmonic number was an integer value. When the harmonic number is odd integer multiples of 0.5, the BF fell in between two adjacent harmonics. Therefore, in the rate- f_0 tuning, there were peaks at integer harmonic numbers and valleys at odd integer multiples of 0.5. The example neuron showed such oscillation pattern for lower harmonic number but stopped responding when the harmonic number was larger than 6 even there were always harmonics in the tone receptive field (Figure 4.1B, spectrum on the left). This response pattern, which was different from the auditory nerve responses (Cedolin and Delgutte 2005), could not be simply predicted from the pure tone tuning.

We used discrete Fourier transform (DTF) to calculate the power spectral density (PSD) in order to detect oscillations in the rate- f_0 tuning profile. If neurons encode precise spectral information so that the response peaks only appear at integer harmonic numbers, there will be a peak at 1 in the PSD (Figure 4.2B). Neurons which showed similar response pattern could be identified because the frequency at the maximal amplitude of PSD would be around 1 (Figure 4.3A). A non-parametric statistical test was used to assess whether there is a significant peak around 1 cycle per Hertz (Methods, Figure 4.4). The firing rate decrease with increased harmonic number was quantified by a suppression index. A suppression index was defined as the normalized firing rate difference between the smallest and largest harmonic number. If only the integer harmonic numbers were used (Figure 4.2C). The distribution of the suppression index was shown in Figure 4.3C. Surprisingly, more than half of the neurons were suppressed by harmonic complex tones. Based on the two criteria: periodic response pattern and decrease in

firing rate (Suppression index < 0), 60/379 neurons were categorized as frequency selective neurons (Figure 4.9). The response pattern that the simultaneously presented components suppress the response to BF tone and the suppression was larger for spectrally denser stimuli suggest surrounding inhibitions for the excitatory receptive field.

4.2.2 Band-pass neurons

Another type of neurons also had a single peak in pure tone tuning (Figure 4.5A) but showed completely different response pattern from the frequency selective neurons. The firing rate increased as the spectral density increased (Figure 4.5B). Moreover, the firing rate tuning curve had one peak at low harmonic number but became flat gradually to high harmonic numbers, which indicates the neuron could not resolve the frequencies at high harmonic numbers. We used a tuning width to characterize change of spectral resolvability in the tuning curve (Methods). When the tuning is completely flat, the tuning width will be 1, which indicates that the neuron cannot distinguish the spectral information (Figure 4.6A and B). A density preference index, the ratio of the tuning width at the largest harmonic number to the maximal tuning width, was used to quantify the change of tuning width. We use density preference index > 0.8 and suppression index ≤ 0 to identify band-pass neurons because such response pattern was very similar to the AN fibers (Figure 4.9). However, there were a few neurons which did not respond to pure tone but to harmonic complex tones and broadband noise also fit the criteria (Figure 4.7A). The firing rate usually increased with noise bandwidth and sound level (Figure 4.7B, C). Because those neurons still showed the similar preference to spectrally dense stimuli and similar trend of firing rate increase as the band-pass neurons, we did not put them into a separate category. Band-pass neurons are less likely to encode spectral information because of the relatively flat tuning. However they could encode other information, such as overall sound level and frequency density, which are still important features of harmonic complex tones.

4.2.3 Modulation sensitive neurons

Another important feature of a harmonic complex tone is the temporal regularity. A harmonic complex tone will have a periodic temporal modulation equal to f_0 if all harmonics start at the same phase. Such envelope modulation could be coded in the temporal phase-locking in the auditory system (Cedolin and Delgutte, 2005; Fishman et al. 2013). The cutoff frequency for phase-locking in auditory cortex is much lower than subcortical areas (Lu, Liang et al. 2001). Some new coding strategy is necessary for coding the temporal feature for larger f_0 s. We found a few neurons which did not respond to pure tones but only to harmonic complex tones at certain f_0 s (Figure 4.8A). We did additional tests by manipulating the starting phases. If the even harmonics started at cosine phase while odd harmonics started at sine phase (ALT), the temporal modulation rate will be twice of the f_0 in this manipulation. If every component starts with a random phase (RND), the peaks in the temporal envelope become less obvious (Figure 4.8C). In all three conditions (COS, ALT, RND), the spectral content remained the same. The example neuron showed a preferred f_0 at around 200Hz (Figure 4.8 B). However, the response peak shifted to around 100Hz when alternating phase was used. The neuron did not respond to harmonic complex tones with random phase.

Although it did not respond to pure tones, the example neuron responded to sinusoidal amplitude modulated tones (sAMs) when the modulation frequency was at the preferred f_0 (Figure 4.8D). The preferred carrier frequency around 7kHz, which was consistent with BFs of neurons recorded from the same recording track. Such phase sensitivity suggested this example neuron encoded the temporal modulation information of harmonic complex tones. Because of limited samples of temporal modulation sensitive neurons (6/372), we did not separate this subpopulation in the population analysis later in this chapter.

4.2.4 Harmonic template neurons

We also found another subpopulation of neurons which responded to harmonic complex tones but showed weak or no response to pure tone (Figure 4.9 A and B). However, different from the temporal modulation sensitive neurons, those neurons were usually not phase sensitive (Figure 4.9C). A facilitation index, the normalized firing rate difference between response to best f_0 and the response to the BF, was used to quantify such combination sensitivity. In addition, shifted harmonic complex tones were used to identify harmonic template neurons in addition to the facilitation index. Shifted harmonic complex tones were generated by shifting the first 5 or 6 harmonics of the preferred f_0 of individual neurons by 25% of f_0 gradually (Figure 4.9D). Only when the shift is 100%, 200%, 300%, ..., the shifted complex tone is still harmonic with the same f_0 . When the shift is odd integer multiples of 50%, all frequencies are odd harmonics of $f_0/2$. The other shifts generated inharmonic tones. When the harmonics are resolved, a change in pitch will be perceived for the inharmonic and odd harmonic shifts (Moore and Moore 2003). Because harmonic templates are for resolved harmonics, a harmonic template neuron should be able to distinguish harmonic shifts from the other shifts (Figure 4.9E). Neurons that met both criteria: $FI > 0.33$ and $PI > 0.5$ were identified as harmonic template neurons. The emergence of such neurons suggests an integrative representation of complex sounds in A1.

More detailed properties of harmonic template neurons were already described in Chapter 3. Because of the complicity in responses to harmonic complex tones, we used different criteria for separating different subgroups of neurons. In Figure 4.10, all neurons were plotted in a three dimensional space: the frequency relative to 1 in periodogram, suppression index and density preference index. With all three measurements, the frequency selective neurons and band-pass neurons could be fairly well separated. However, harmonic template neurons were more scattered in the 3-D distribution. This suggests for highly selective neurons, it's not sufficient to evaluate what features in harmonic complex tones are encoded. Additional tests are needed to fully

understand the functional properties. This was also the reason we did not categorize the rest neurons further and put them in the uncategorized group together.

4.2.5 A level invariant representation of spectral information

In auditory nerves, the rate responses to complex stimuli are strongly dependent on stimulus levels (Sachs and Young 1979). The representation of resolved harmonics in rate response was degraded as the sound level increased which could possibly due to broadened cochlear tuning or rate saturation (Cedolin and Delgutte 2005). We tested some frequency selective neurons and harmonic template neurons at different sound levels. For frequency selective neurons in A1, the peak firing rates changed under different sound levels. However, the tuning curve did not become flat and there were still clear peaks and valleys around the integer values (Figure 4.11A, B). We compared the tuning width change at different harmonic number and did not find a significant change in tuning width for different sound levels (Figure 4.11C).

Such sharp frequency selectivity cannot be simply explained by a sharp frequency tuning shown in pure tone responses. A few neurons with high spontaneous rates provided evidence for the role of inhibition in maintaining sharp frequency tuning at high sound level. The example neuron showed increased firing rate to integer harmonics at soft level (10dB SPL). When the sound level was increase to 40dB, the firing was suppressed for f_0 s when BF fell between two harmonics and increased when BF coincides with one of the harmonics (Figure 4.11B). Inhibition could also explain the decreased firing rate to increased harmonic number for frequency selective neurons.

4.2.6 Spectral information and Resolvability

Harmonic template neurons and frequency selective neurons have been shown to encode spectral information of harmonic complex tones. By pooling responses from those two subpopulations over a broad frequency range, we could have an estimation of the resolved harmonics for different fundamental frequency. The resolved harmonic numbers depend on the fundamental frequency. For f_0 between 400Hz to 1kHz, the 10th harmonic can be resolved (Figure 4.12). This estimation was consistent with the marmoset behavior but different from human perception in which the resolved harmonic number is independent of the f_0 s.

4.2.7 The relationship between other properties and response pattern to harmonic complex tones

Previous studies have shown different response properties of A1 neurons to pure tones. There were monotonic neurons which increase firing rate when the sound level increases. Other neurons show non-monotonic tuning to sound level changes: maximal firing rate to a certain sound level and the firing rate decreases either the sound level gets higher or lower than the preferred level. These different rate-level tunings suggest different excitation and inhibition inputs pattern (Ojima and Murakami 2002, Tan, Atencio et al. 2007, Levy and Reyes 2011). It's still unknown, however, how these neurons differ in the spectral selectivity. We compared the rate-level functions to BF tones of neurons in different groups described in the previous part. A monotonicity index (MI) is defined as the ratio between the firing rate to the maximal sound level and maximal firing rate. A MI is between 0 and 1. A MI close to 0 indicates a non-monotonic rate level function whereas a MI close to 1 indicates a monotonic tuning. Frequency selective neurons and harmonic template neurons are mainly non-monotonic neurons (Figure 4.12B). A majority of band-pass neurons are monotonic neurons. The uncategorized group included both monotonic and non-monotonic neurons.

There was no significant difference in spontaneous firing rates, or distribution of BFs across different subpopulations (Figure 4.12 C and D). The facilitation index was also compared across different groups. Harmonic template neurons have a strong preference to harmonic complex tones, which was also the first criterion for such neurons. Frequency selective neurons responded to pure tones and harmonic complex tones equally well. The facilitation index uses the largest firing rate in HCTs tuning. Earlier in this Chapter, we also showed a trend of firing rate decrease with increasing harmonic number or spectral density for frequency selective neurons. They have a preference to spectrally sparse stimuli. Since pure tones are the sparsest stimuli, frequency selective neurons usually show large responses to BF tones too.

4.2.8 Heterogeneity of spectral selectivity in the large scale tonotopic map

Neurons in A1 are tonotopically organized and neighboring neurons usually have similar BFs because of common thalamo-cortical inputs (Figure 4.14A). If we plot the distribution of neurons within the tonotopic map and also distribution in depth in different categories of spectral selectivity, different subpopulations were mixed together and we did not observe any cluster either in the tonotopic map or in depth (Figure 4.14 B and C). However, the recording depth distribution (Figure 4.14 D) showed a biased sample in supragranular layers (II/IV)

Another more direct way to test whether there is possible functional cluster in A1 is to compare the responses to complex tones of simultaneously recorded neuron pairs. Spikes from different neurons can be sorted by different waveform templates (Figure 4.15B, insert plot). Those simultaneously recorded neurons were usually close to each other physically (< 100um distance). The same set of stimuli, including pure tones, complex tones were tested. We use signal correlation to measure the similarity of the spectral selectivity for neuron pairs. The responses of a neuron pair to the shifted harmonic complex tones were shown in Figure 4.15A.

For shifted tones, the frequency center also moved from low frequency to high frequency when the shift increased. The response areas for both neurons overlapped, which implies similar frequency preference. However, one neuron only responded to harmonic shifts while the other responded to all shifts.

Signal correlation is a measure of the similarity of tunings. The negative signal correlation suggested two neurons barely responded to the same stimulus (Figure 4.15C). For all 63 neuron pairs, the estimated BFs were close to each other (Figure 4.16A). However, the similarity in BF did not result in similar selectivity to complex tones with more complicated spectral contents. For all 63 neuron pairs, both signal correlation and noise correlation were low (Figure 8B and C). Overall, our results suggest within the large scale tonotopic map, there is a parallel process for harmonic complex tones which might be used for extracting different sound features.

4.3 Discussion

4.3.1 A distributed harmonic process in AI

In order to identify a sound source with harmonic structures, a parallel spectral analysis has to be done on different scales to extract features associated with perceptual attributes. For example, the discrete frequency components need to be grouped together to estimate f_0 if the fundamental frequency itself is not presented and to compute the overall sound level. In addition, the accurate representation of individual harmonics is important for timbre perception. The diversity in neural responses to harmonic complex tones suggests different information was extracted and carried by different subpopulations of neurons. The harmonic template neurons and frequency selective neurons encode resolved harmonics and provide spectral information for f_0 extraction and timber. Integrators can be used for coding frequency range, spectral density or

overall sound levels. The temporal envelope neurons mainly code the temporal modulation in the signal. Last but not least, although the uncategorized group cannot be characterized by our measurements, they might encode some other important features. Because of the high dimensionality of complex tones, additional tests are necessary to fully understand the functions of those neurons.

All the different subgroups of neurons form a distributed neural code for harmonic complex tones with pre-processed information which can be used further for extracting perceptual features, such as loudness, pitch and timber to form a code of ‘auditory objects’ at later stages.

4.3.2 Harmonic resolvability in auditory cortex

By examining the response of cortical neurons, we could estimate the lowest f_0 , the harmonics of which can be resolved in a given frequency range (different BFs). By pooling all neurons which encode spectral information of harmonic complex tones, we could estimate the highest harmonic number that could be resolved in auditory cortex for different f_0 s. We found the resolved harmonic numbers changes with the f_0 s. For the f_0 s around 1kHz, up to the 10th harmonic can be resolved. However, our estimation for the upper limit of resolved harmonic numbers is based on two subpopulations of neurons characterized by our measures. It might be an underestimate because of the limited sample size and the fact that we did not take into account the resolvability of neurons in the uncategorized group. Our results are consistent with previous studies. Two different groups showed the maximal resolved harmonic numbers increases with characteristic frequency (CF) of AN fibers, from 2-3 at 200Hz to about 10 at 10kHz and above on different animal models (Evans, Rosenberg et al. 1971, Cedolin and Delgutte 2005). Our result was also comparable with the study in the cat anteroventral cochlear nucleus (AVCN), which

reported the number of harmonics resolved in the rate responses of single units increases from 2 at 250Hz to 13 at 10kHz (Smoorenburg and Linschoten 1977).

4.3.3 Comparison between the peripheral auditory system and the central auditory system

The peripheral auditory system acts like a frequency analyzer which decomposes a harmonic complex tone into a series of topographically organized frequency channels along the basilar membrane of cochlear. When the harmonic number is small, each component falls into different filters (resolved). The frequency contents could be encoded by the spatial activation pattern of all frequency channels. For unresolved harmonics, the interaction of frequencies within a single filter generates an envelope modulation which can be coded by neurons which phase lock to such temporal modulation. The subpopulations of cortical neurons, which encode spectral information only responded to resolved harmonics and cease firing when the harmonics are not resolved. The emergence of those specialized neurons suggest the spectral and temporal information start to separate in A1.

Another important finding in our study is that the resolvability does not change with sound level change. This is very important because our auditory system operates in a broad dynamic range because of the distance to the sound source, background noise. For AN fibers and AVCN neurons, the ability to resolve harmonics in their rate response degrades rapidly with increasing stimulus levels, even the low and moderate stimulus levels (15-20dB above threshold) were used to minimize rate saturation (Smoorenburg and Linschoten 1977, Cedolin and Delgutte 2005). Although the rate-place representation in the AN fibers provide sufficient information to account for the behavior performance at high sound levels as shown in previous studies, it fails to explain the general robust psychophysical performance at high sound levels (Hirsh, Reynolds et al. 1954). In addition, accurate frequency information need to be maintained in order to use a

‘‘harmonic template’’ to extract pitch information(Goldstein 1973). Frequency selective neurons showed a level-invariant representation of the spectral information over a relative wide range (up to 40dB difference). Although we were not able to separate excitation and inhibition in extracellular studies, the suppression in firing rates to spectrally dense stimuli and harmonic complex tones that don’t include BF suggests that inhibition can help maintain the frequency selectivity at higher sound levels.

4.3.4 Parallel spectral analysis in A1

In our study, we did not find any clusters for different subpopulations of neurons characterized by their responses to harmonic complex tones. Instead, they were all scattered in the tonotopic map and intermingled with each other. Our finding was consistent with previous studies, which found tone-unresponsive neurons scattered in the local circuitry and neighboring neurons could have similar or complete different frequency response areas(Hromadka, Deweese et al. 2008, Bandyopadhyay, Shamma et al. 2010, Rothschild, Nelken et al. 2010). Although our criteria for grouping neurons were not based on tone response but responses to harmonic complex tones. We did find some correlation between the tone responses and selectivity to harmonic complex tones. For example, the majority of the frequency selective neurons were tone-responsive and had a non-monotonic rate-level functions. On the contrary, the band-pass neurons tend to have a monotonic rate-level function. Harmonic template neurons and modulation sensitive neurons had weak or no responses to tones.

We also found that on average the signal correlations between neuron pairs were low, consistent with previous studies. However, the noise correlation in our study was on average low as well, which was different. One possible reason could be the different stimulus set. The complex tones used in our study cover a broad frequency range (3 octaves), which might recruit

more neurons in the local circuitry or even long range intra-cortical connections comparing with pure tones. Another possible reason could be the different experiment preparation. In the study by Rosechild and Nelken, recording was done under anesthesia while we recorded from awake animals (Rothschild, Nelken et al. 2010). The smaller noise correlation in our study might be due to the awake state. Previous studies have shown the neural correlation changed in different states or by associative learning (Issa and Wang 2013, Jeanne, Sharpee et al. 2013). The third reason might be the sample sizes. Because we had only 63 neurons pairs, there could be a larger bias in our distribution.

We did not find any organization along the recording track either. However, in our study, we had a biased sample in Layer II and III. Because of the information flow hierarchy across layers, it will be interesting to compare the neural responses to harmonic complex tones across different layers.

In summary, data presented in this chapter show that within the large scale tonotopic map, the local processing for complex tones is heterogeneous, which suggests parallel processing for feature extraction.

4.3.5 Harmonic process and comparative aspects of pitch perception

In this chapter, I systematically examined how cortical neurons represent harmonic complex tones. The perception of complex sounds by animals may be related to human pitch perception. We did not directly associate our study with pitch perception as previous studies, because we were trying to understand a general neural mechanism for complex sound processing, which might also involve in pitch extraction. There has been shown a link between speech and pitch perception. The probability distribution of speech sounds could predict many pitch perception phenomena.

In our study we used f_0 s cover a broad frequency range (400Hz to 15kHz) because marmosets have a hearing range from 300Hz to 38kHz (Osmanski and Wang 2011). The first harmonics of four major marmoset vocalizations are between 3kHz to 8kHz (DiMattina and Wang 2006). We found that the upper limit for resolved harmonics depends on the f_0 s, which is consistent with the estimation from the critical bands of marmosets. We did not test the temporal processing systematically because most of stimuli had f_0 s above 400Hz. Behavior study suggest marmosets rely on spectral cues for f_0 s above 400Hz in a pitch discrimination task. This could also be the reason that we did not get many modulation sensitive neurons with harmonic complex tone search. It will be interesting to compare the spectral and temporal information processing for harmonic complex tones on different animal models given different vocalization ranges and hearing ranges across species. We will have more evidence whether there is a general pitch mechanism common for all vertebrate species.

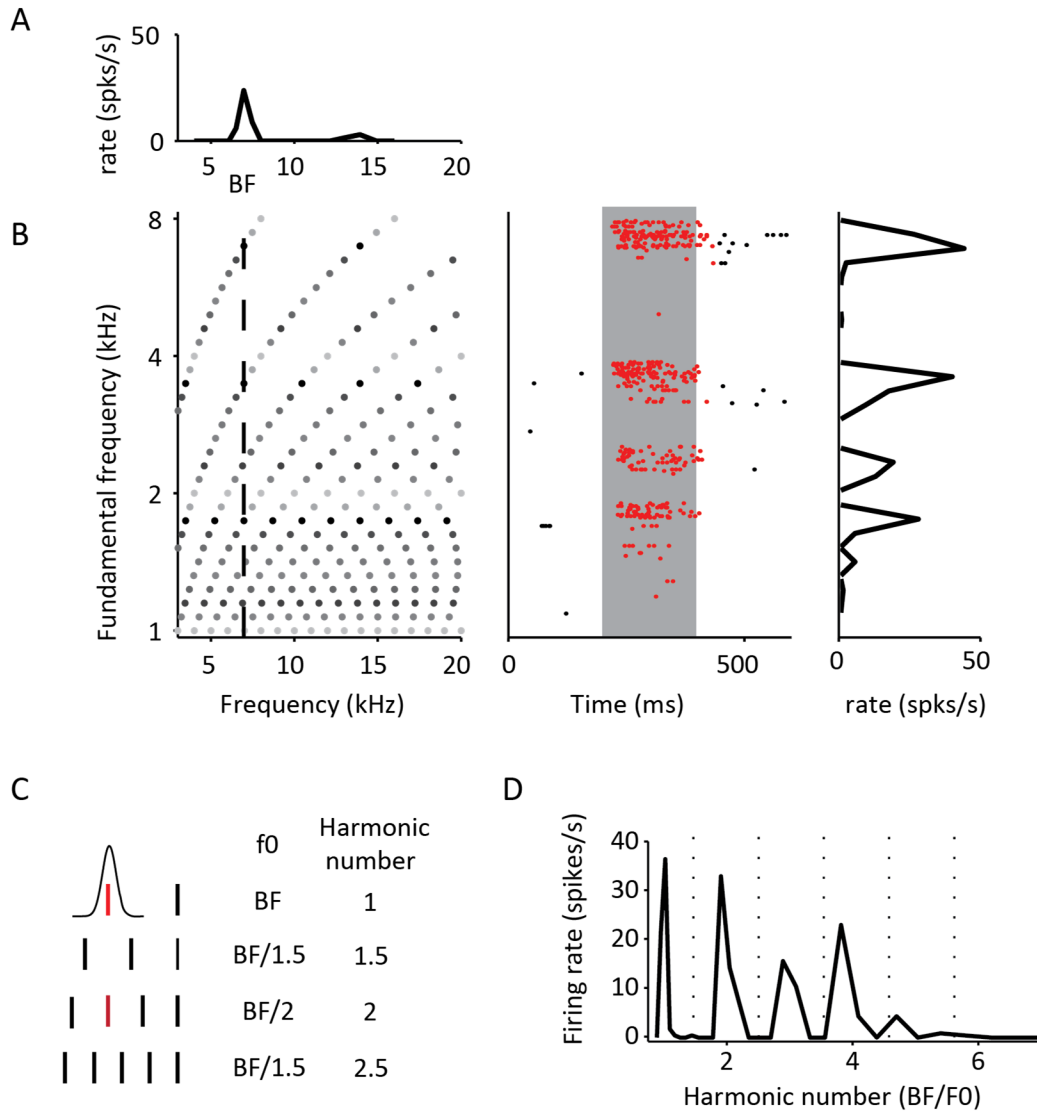


Figure 4.1, An example of the responses of a frequency selective neuron to pure tone and harmonic complex tones.

- A, frequency tuning measured at 30dB SPL showed a peak at the best frequency (BF).
- B, left: the spectrum of harmonic complex tones at different f_0 s with equal amplitude. The level for individual harmonics was 30dB SPL. Middle: the raster plot of the neural responses to the stimuli shown on the left. The shaded area indicates the duration of sound stimuli. Right: the average firing rates to different f_0 s.
- C. The schematic illustration of the HCT stimuli.
- D, firing rate tuning to harmonic number, a dimensionless parameter

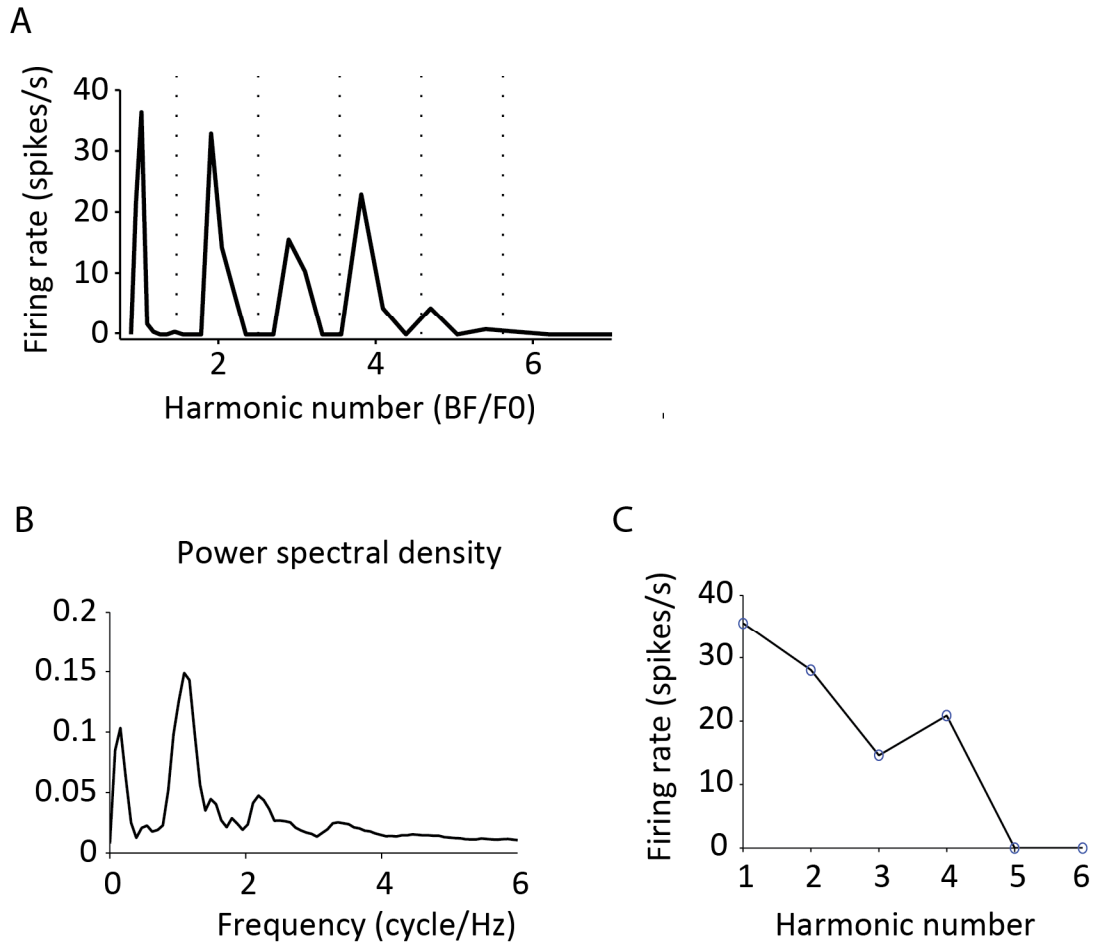


Figure 4.2. Metrics for quantify response pattern to HCTs.

A, the rate-harmonic number tuning.

B, the power spectral density of the rate-harmonic number tuning in A.

C, firing rates of harmonic numbers at integer value only. There was a trend of decrease in firing rate when the harmonic number increased.

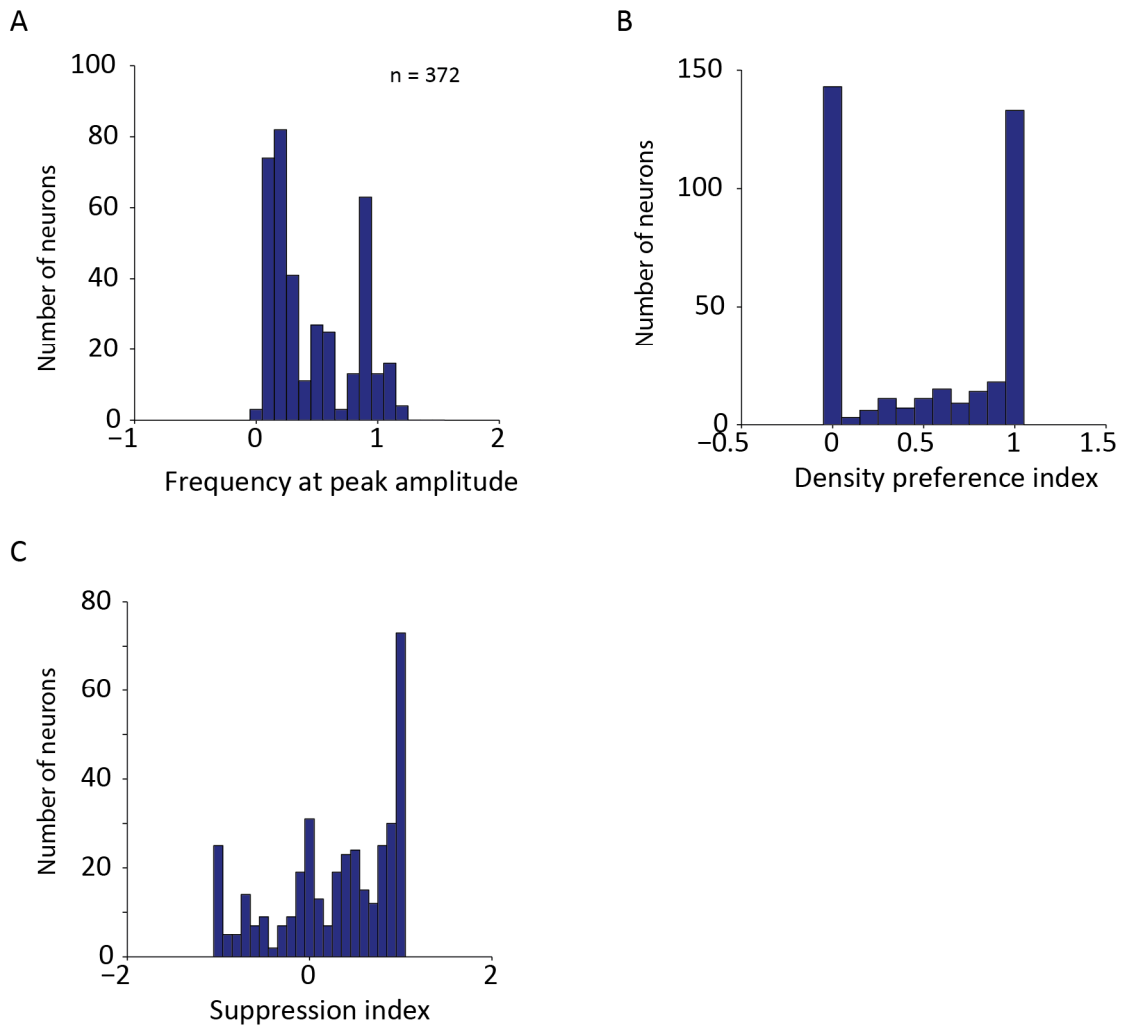


Figure 4.3 Distribution of three different measurements for tuning to harmonic complex tones

A, distribution of frequencies at peak amplitude in the power spectral density function. A value near 1 indicate an oscillating response pattern in Figure 4.1.

B, distribution of the density preference index

C, distribution of the suppression index

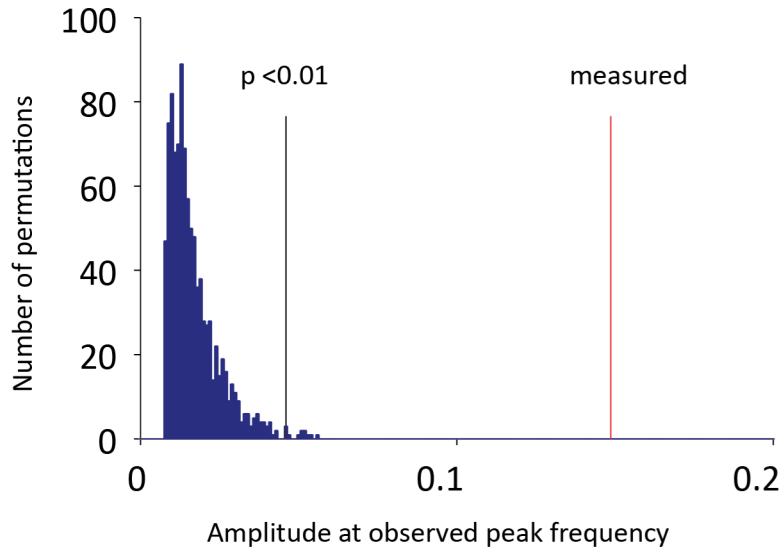


Figure 4.4 A nonparametric permutation test (Ptitsyn, Zvonick et al. 2006, Fishman, Micheyl et al. 2013) to assess the statistical significance of the peaks by randomly shuffling the sample points in the rate-f0 tuning. For every shuffled trial, the power spectral density was computed and the amplitude at the peak frequency of the original trial was measured. The null-hypothesis is that a random point serial will yield a same amplitude or higher peak in the power spectral density distribution. A null reference distribution of the spectral amplitude at the given period was generated by repeating this process for 1000 times. The p-value is defined as the probability of observing an amplitude at a given frequency equal or larger than the observed value.

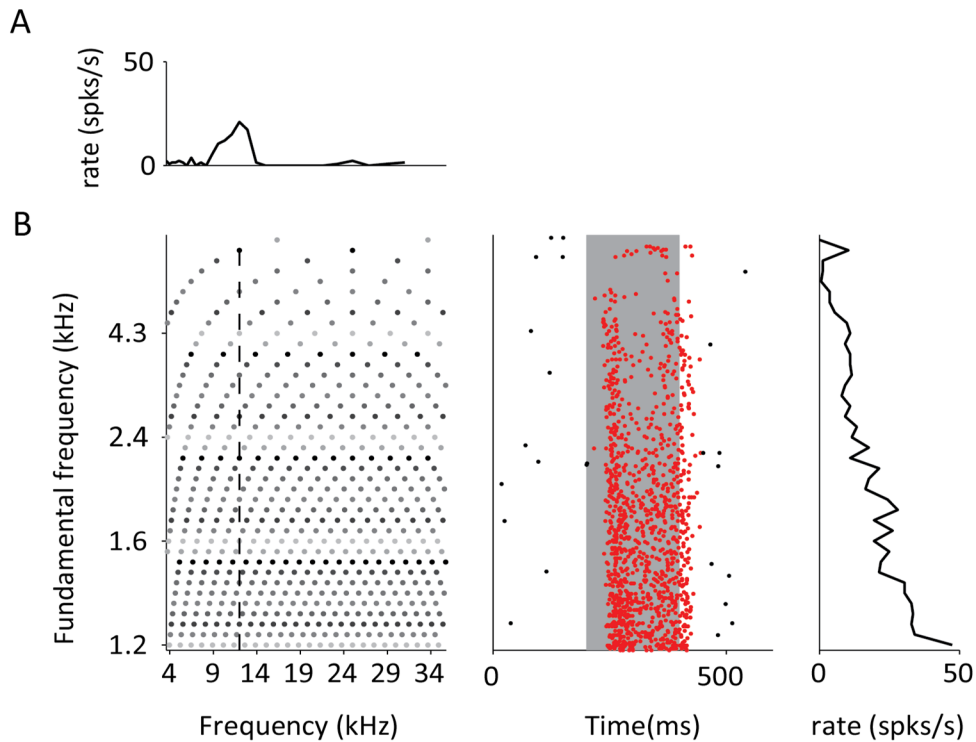
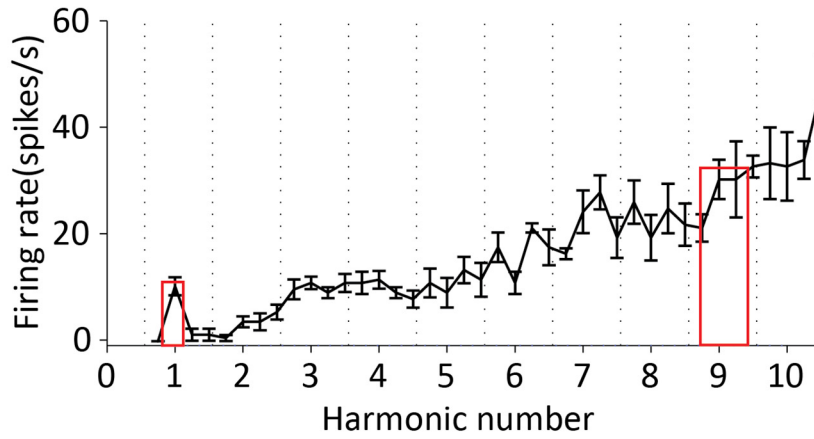


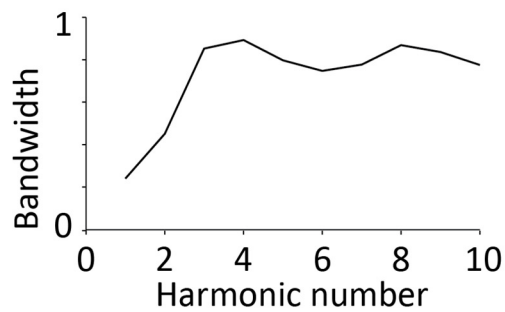
Figure 4.5 An example of the responses of an band-pass neuron to pure tone and harmonic complex tones.

A, frequency tuning measured at 30dB SPL showed a peak at the best frequency (BF).
 B, left: the spectrum of harmonic complex tones at different f_0 s with equal amplitude. The level for individual harmonics was 30dB SPL. Middle: the raster plot of the neural responses to the stimuli shown on the left. The shaded area indicates the duration of sound stimuli. Right: the average firing rates to different f_0 s.

A



B



C

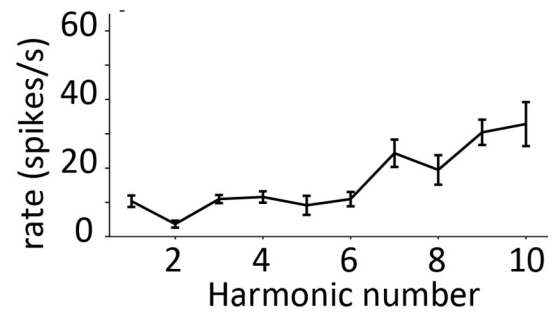


Figure 4.6 An example of how to quantify a band-pass neuron.

A. a demonstration of how the response pattern to HCTs is quantified. The red rectangular squares were used to match the responsive area within each window.

B. the harmonic tuning widths measured at different harmonic number (the width of the red rectangular square). The tuning became more flat to higher harmonic numbers.

C. firing rates of harmonic numbers at integer value only. There was a trend of increase in firing rate when the harmonic number increased.

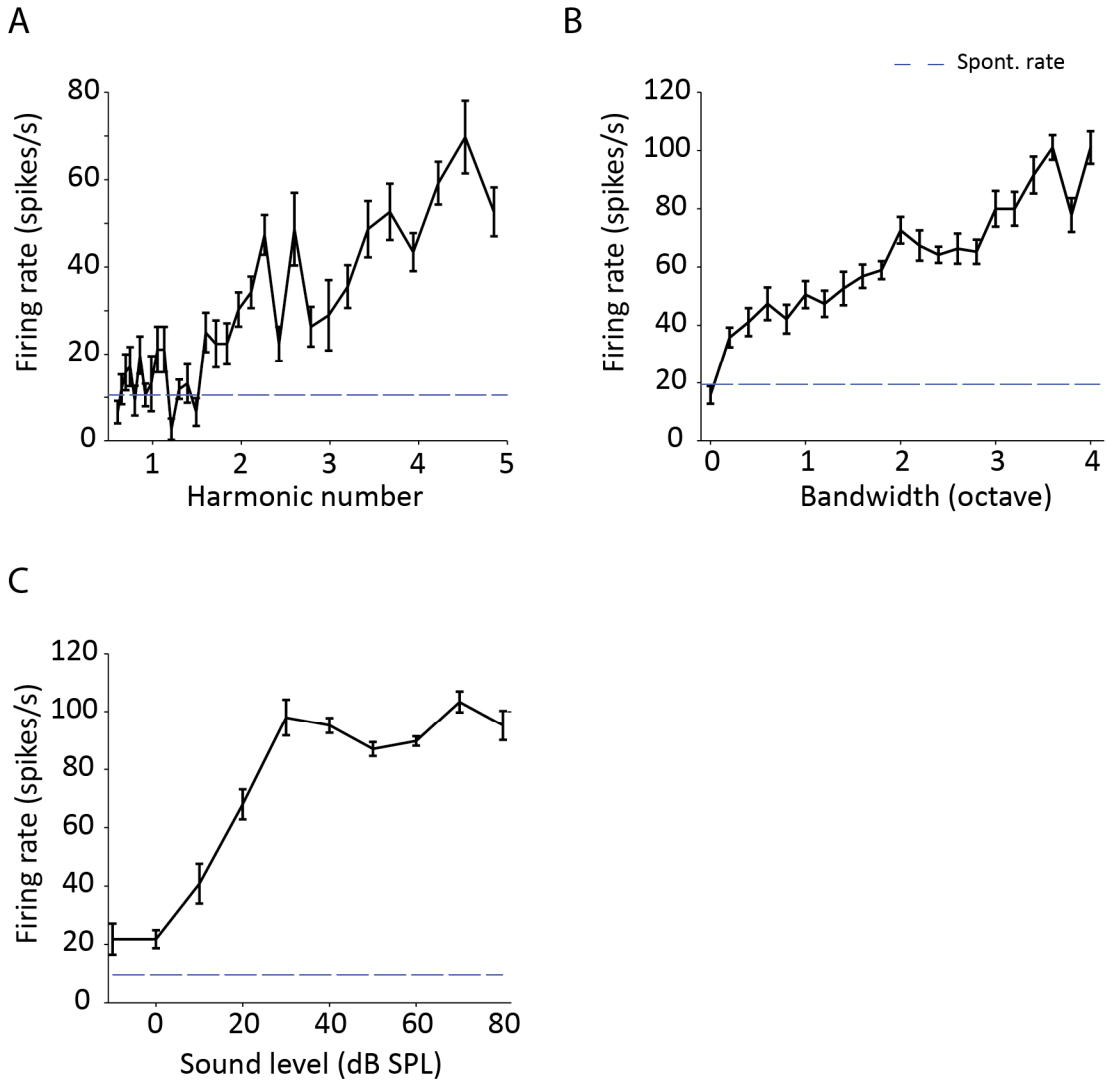


Figure 4.7 An example of a band-pass neuron which preferred high density and loud stimuli.

A, responses to HCTs.

B, firing rates to noise with different bandwidth.

C, rate-level function of a broadband noise.

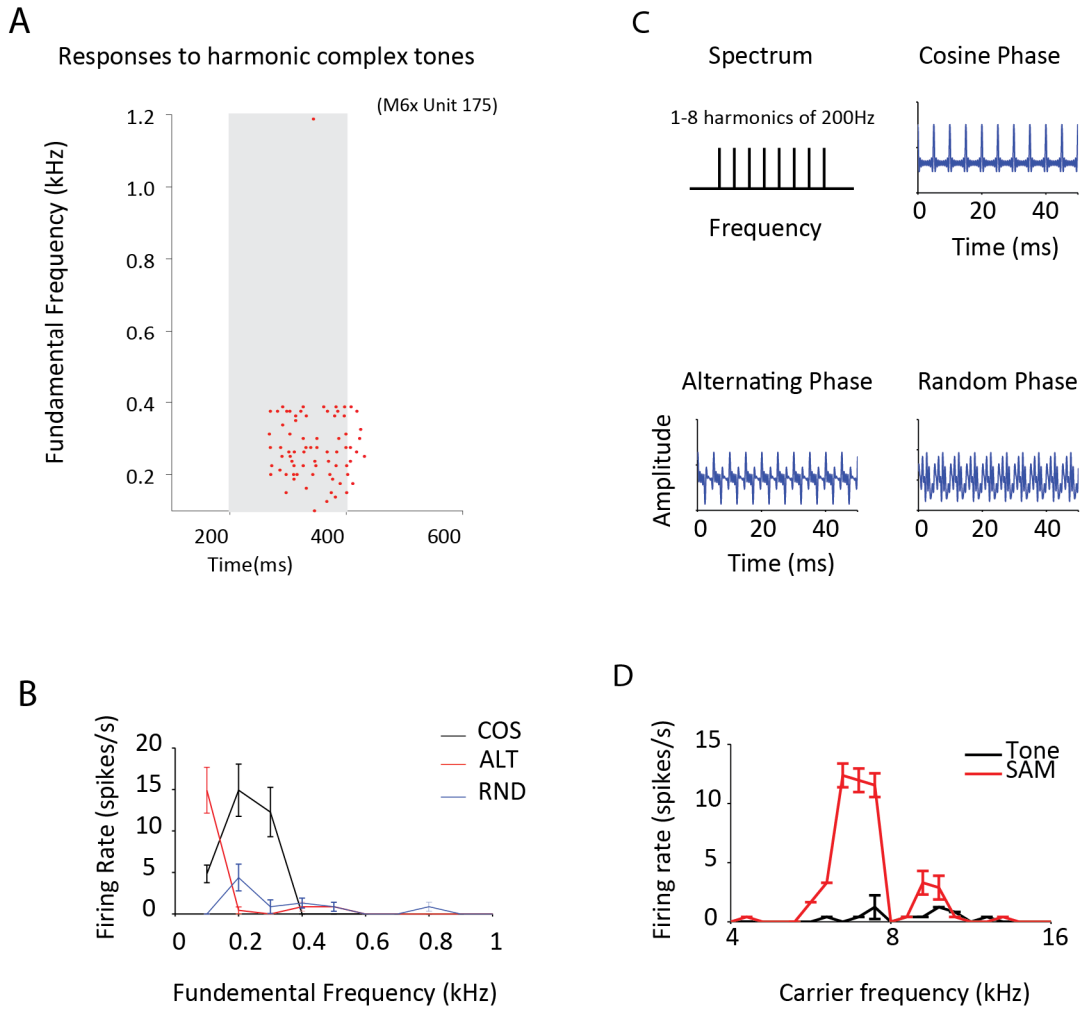


Figure 4.8 An example of the responses of a modulation sensitive neuron.

A, raster plot of responses to f_0 of HCTs.

B. An illustration of different stimulus waveforms for different phase relationships among harmonic of a 200Hz f_0 .

C, The average firing rate tuning to f_0 of HCTs in three different phase relationships.

D, The responses to pure tone at different frequency and a 256Hz sinusoidal amplitude modulated tones.

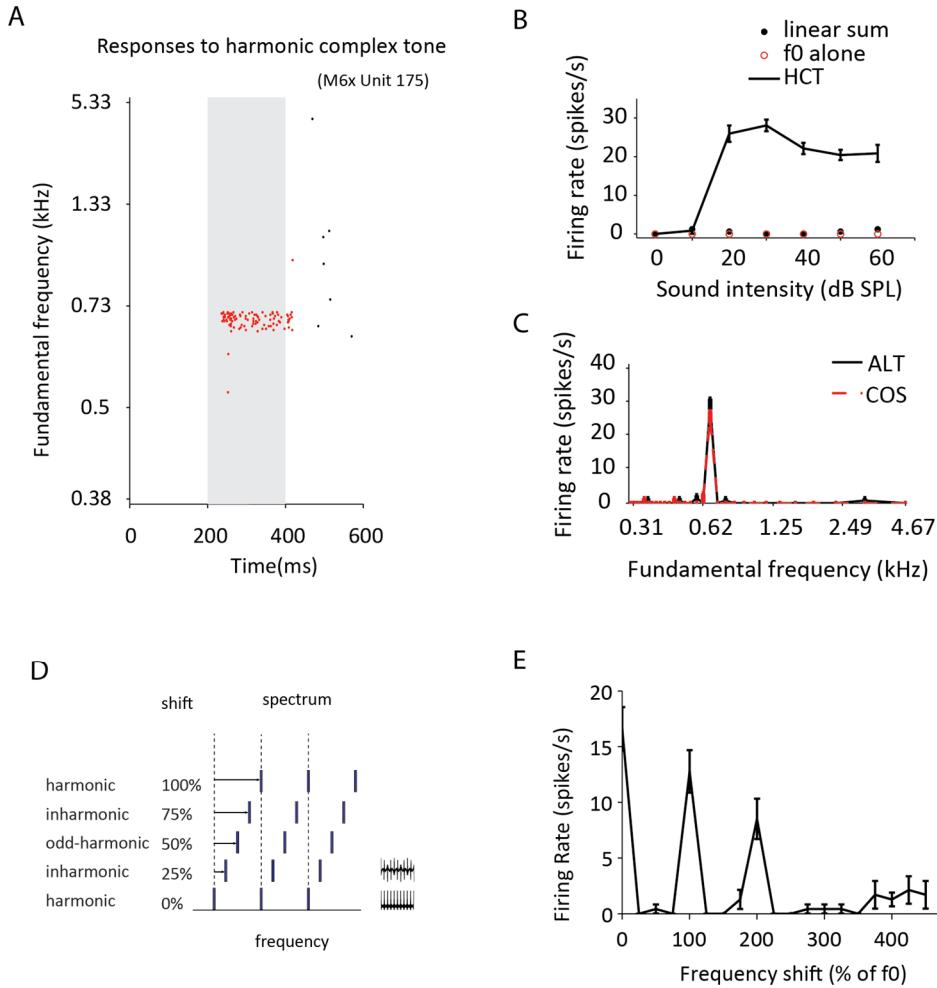


Figure 4.9 An example of a harmonic template neuron.

A, raster plot of responses to f_0 of HCTs.

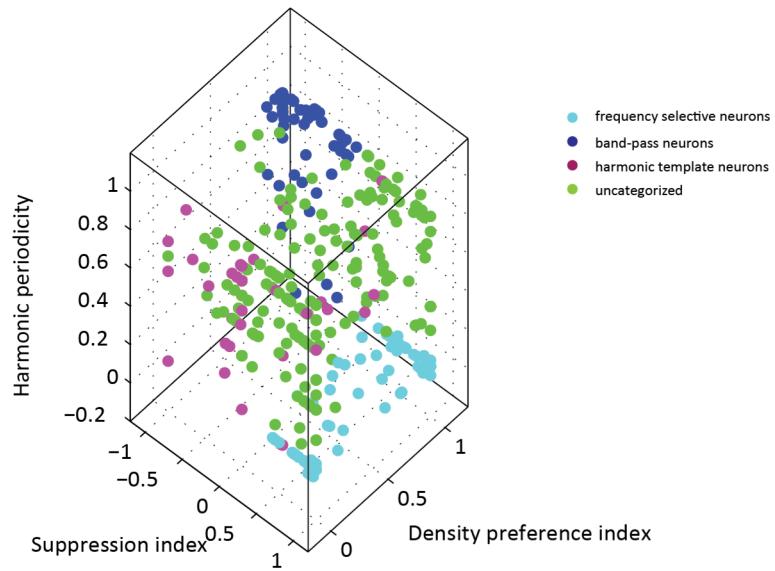
B, Responses to the preferred HCT and individual harmonics alone at different sound levels.

C, the rate- f_0 tuning in two different phase relationships.

D, a diagram of the spectrum for shifted harmonic tones.

E, the responses to shifted harmonic tones. Neuron only responded to harmonic shifts.

A



B

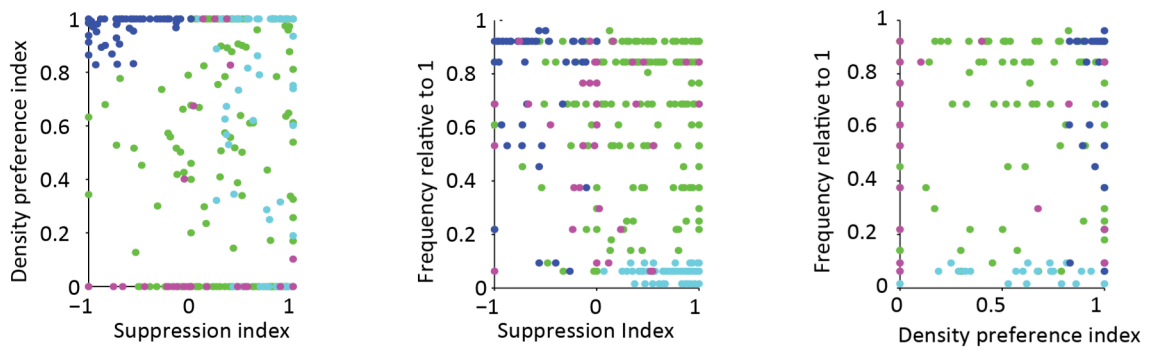


Figure 4.10 Distributions of different subgroups of neurons in different metrics

A, the distribution of different subgroups of neurons in a 3-D space of the three measurements: periodicity, suppression index, and density preference index

B, a projection of the 3-D distribution to 2-D spaces.

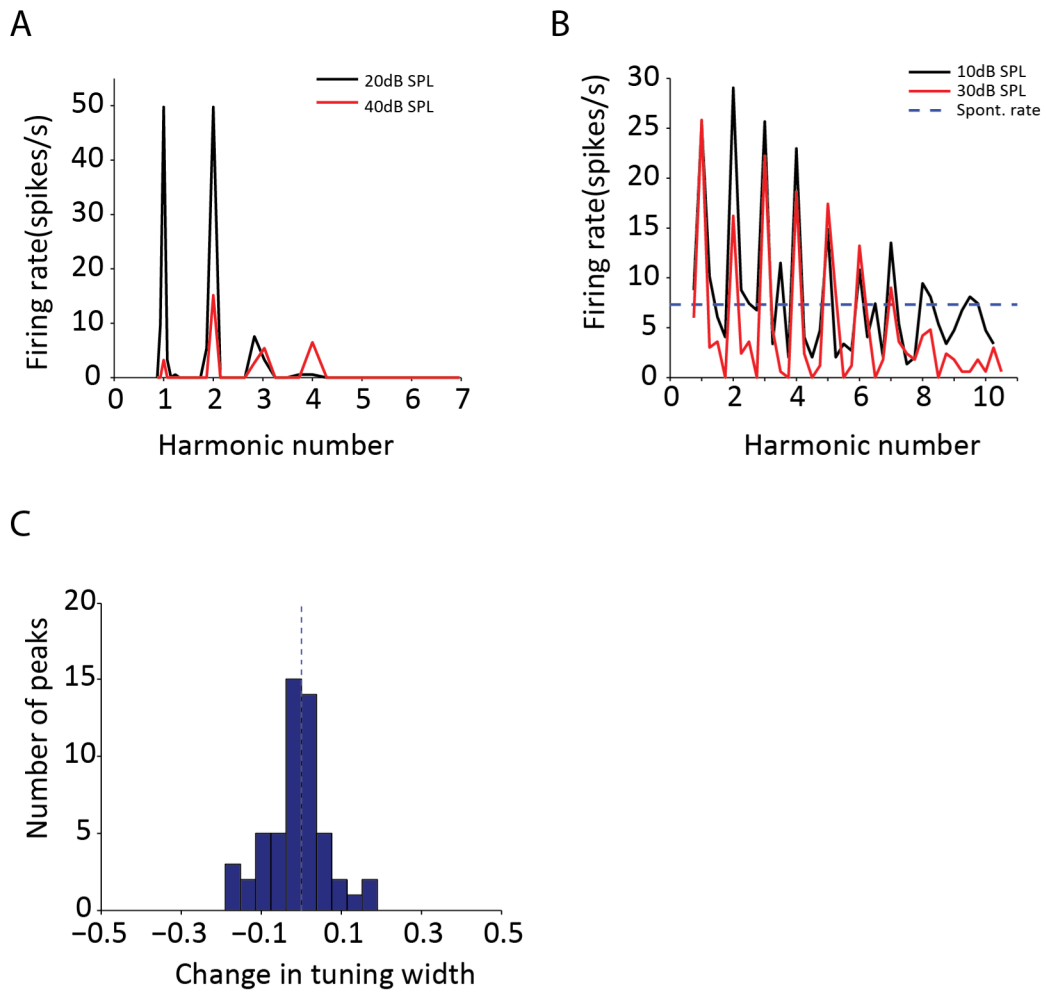


Figure 4.11, A level independent rate code for spectral information of HCTs.

A, an example of a frequency selective neuron's tuning to f_0 s at two sound levels tested.
 B, another example of a frequency selective neuron with a high spontaneous firing rate.
 C, the distribution of change in harmonic tuning width at every peaks in the f_0 tuning for 10 neurons when the sound level was increased by 10 – 40dB SPL.

Harmonic Resolvability on Marmoset

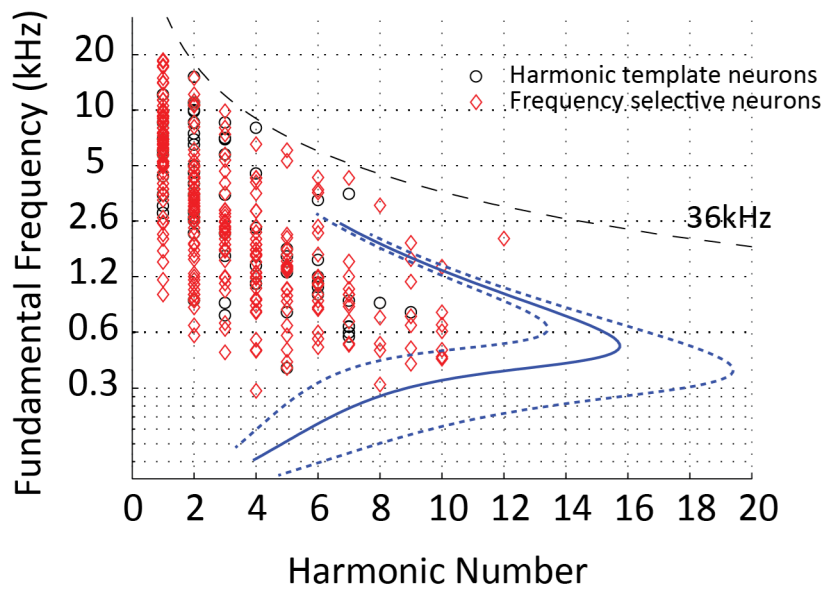


Figure 4.12 Estimation of resolved harmonic numbers for different f_0 s by pooling all frequency selective neurons and harmonic template neurons. The blue lines were the estimated resolvability from the measured critical bandwidth of marmosets (Osmanski, Song et al. 2013).

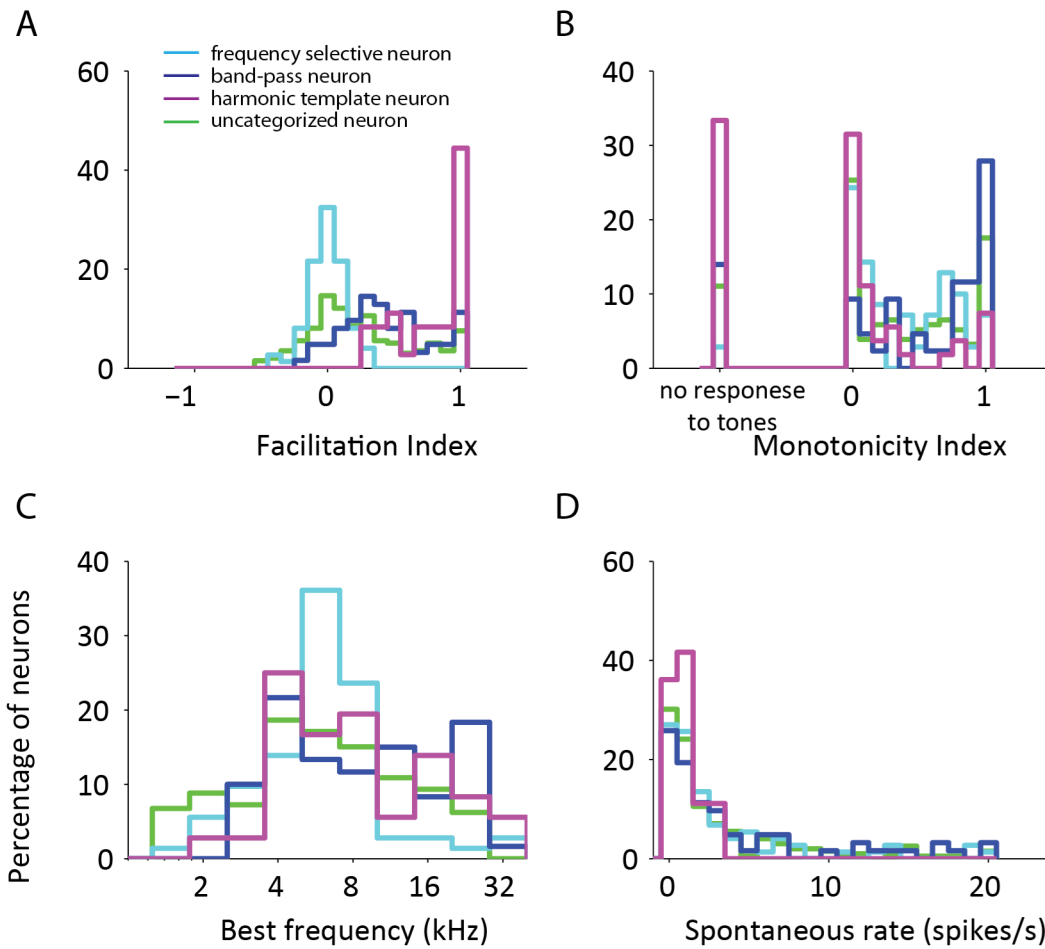


Figure 4.13 A comparison between different subpopulations.

A, the distribution of facilitation index indicates different preference to complex tones and pure tones.

B, the distribution of monotonicity index of rate-level function at BF tones.

C, the distribution of best frequencies

D, the distribution of spontaneous firing rates.

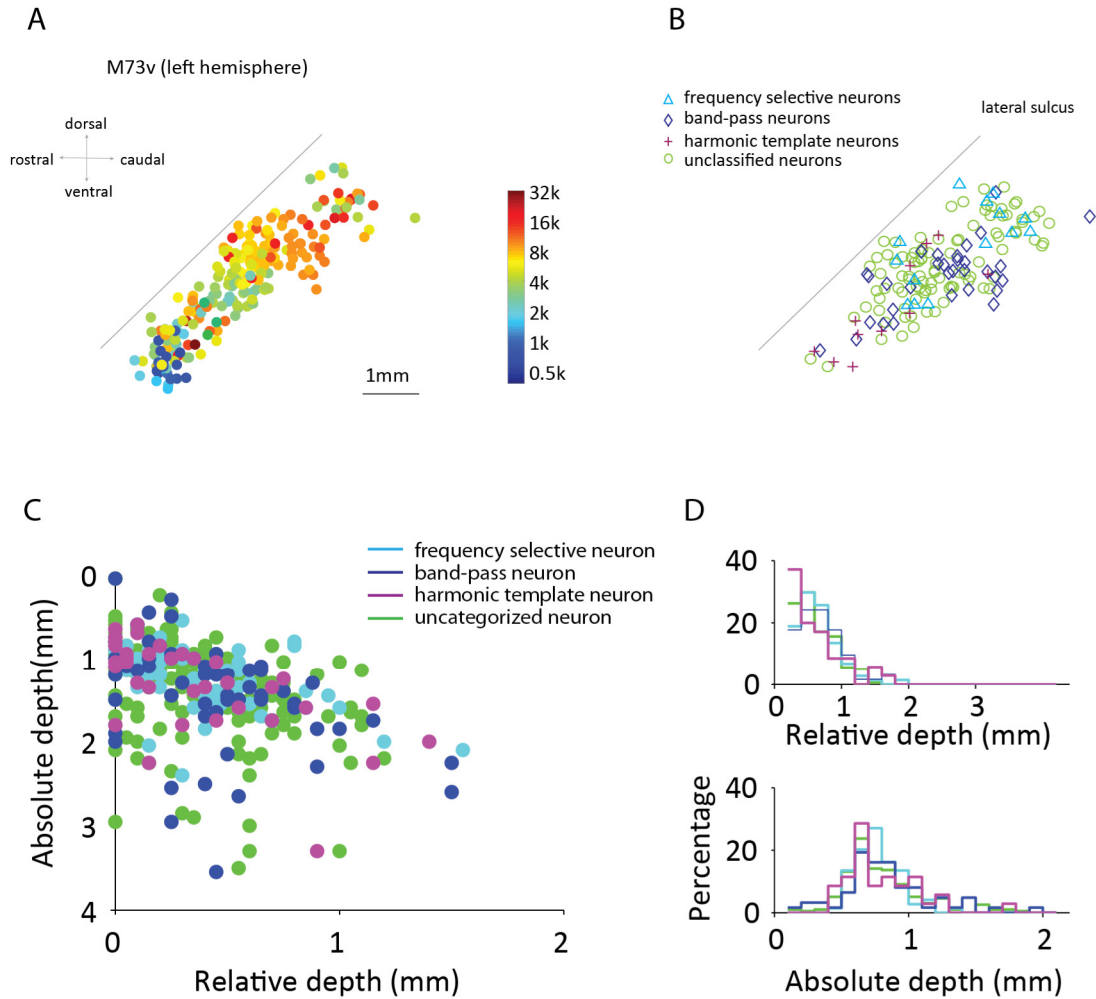


Figure 4.14 Distribution of different subpopulations in auditory cortex.

A, the tonotopic map of A1 from one hemisphere of one monkey (73v).

B, the distribution of different groups in the tonotopic map.

C, the distribution of different groups in depth.

D, the recording depth distribution for different groups. All groups were within a 1mm distance from the first spikes, suggesting that most of our samples are form layer II/III.

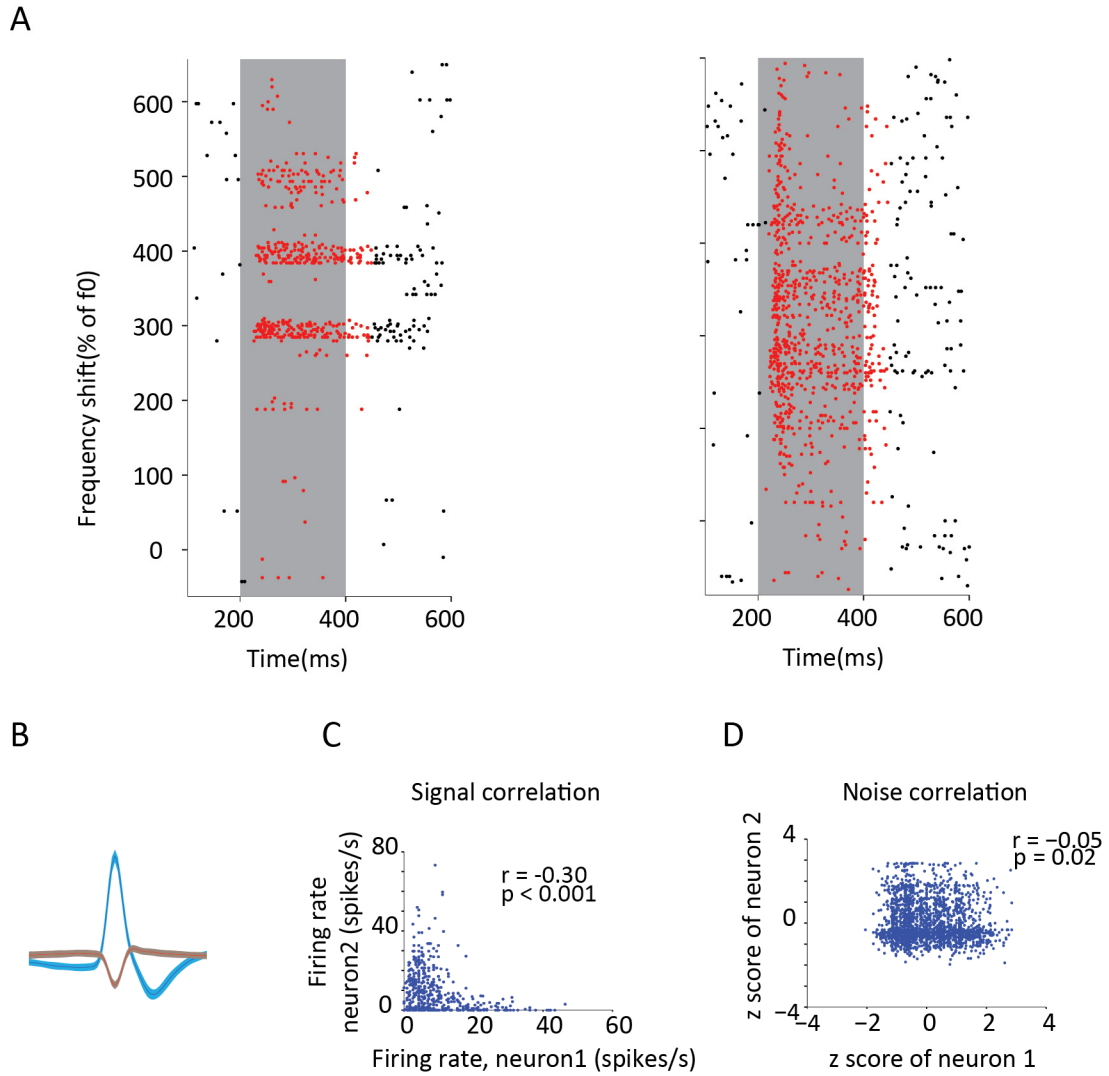


Figure 4.15 Correlations of neuron pairs in responses to complex tones.

A, raster plots of the responses to shifted HCTs of two simultaneously recorded neurons.
 B, average waveform and standard deviation (shaded area) for each neuron
 C, the signal correlation of the neuron pair shown in A in responses to complex tones
 D, the trial to trial noise correlation of the neuron pair in responses to complex tones.

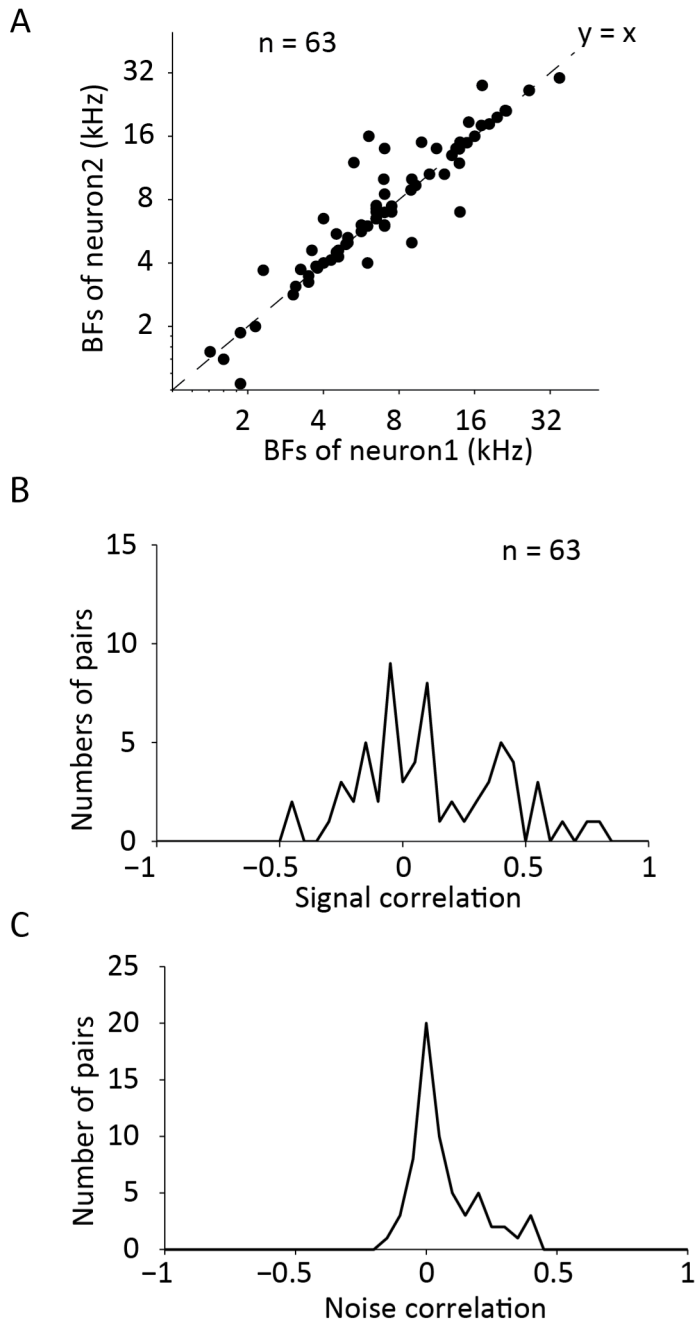


Figure 4.16 a summary of all 63 pairs in BF and neuron correlations

- A, a comparison of estimated BFs for pairs of neurons.
- B, the distribution of signal correlations for all neuron pairs
- C, the distribution of noise correlation for all pairs.

CHAPTER 5:

5. Online adaptive stimulus design for studying spectral integration in auditory cortex

In the previous chapters, I investigated the neural representation of harmonic complex tones in primary auditory cortex using the traditional neurophysiology procedure: pre-determine a set of stimulus with varying parameters, record neural responses, and analyze data offline. In this chapter, I explore a new method: online adaptive stimulus design to study spectral integration in A1. In this approach, every stimulus is generated online according to neural responses to previous stimuli, which can sample a high dimensional feature space more efficiently.

5.1 Introduction

In order to understand how auditory system analyzes sounds and give rise to perception, we are trying to quantify the transfer function between stimuli and neural responses at single neuron level. In most neurophysiology studies, a widely used approach is to play a set of pre-determined stimuli with varying parameters along one feature dimension, for example, the frequency of a pure tone, record the neural responses and analyze the data offline to determine the relationship between stimuli and responses, for example tuning curves. This approach has been successful to characterize auditory neurons to some extent. A good example is the auditory nerve filter model, which provides approximately good prediction for the responses to speech-like

stimuli (Meddis, O'Mard et al. 2001, Zhang, Heinz et al. 2001, Bruce, Sachs et al. 2003) However, the spectral analysis is much more complex in central auditory system (CNS) because of the nonlinear integration of both inhibitory and excitatory responses to a broad frequency and level range. In auditory cortex, for example, a neuron's response to a spectrally complex sound cannot be predicted from the linear sum of its responses to individual partials (Abeles and Goldstein 1972, Shamma and Symmes 1985, Sutter, Schreiner et al. 1999, Kadia and Wang 2003). Combination sensitive neurons which did not respond to simple stimuli but to combination of features make it necessary to study cortex neurons with complex stimuli.

This raises two problems: first, the high dimensionality of complex stimuli makes a lookup table that describe the response of a neuron to all possible combination of stimulus features become impossible. Different models were used other than tuning curves to describe the stimuli and responses relationships in CNS. Commonly used models include the Spectro-Temporal Receptive Fields (STRFs), which are based on spike-triggered average obtained using random noise or natural sounds(Marmarelis 1978, Jones and Palmer 1987, Theunissen, Sen et al. 2000) and frequency weighting function models using random spectrum stimuli (Yu and Young 2000). One of the limitation for STRFs and linear weighting function model is that they can only adequately characterize neurons that are approximately linear (Young, Yu et al. 2005, Wu, David et al. 2006). For cortical neurons, those models can have a poor performance in predicting responses to new stimuli and cannot accurately capture complex receptive fields, such as the multi-peaks (Theunissen, Sen et al. 2000, Christianson, Sahani et al. 2008). More complex models such as the quadratic models and neural network models often require large amount of data and computation (Pfingst and O'Connor 1981, Yu and Young 2000, Bandyopadhyay, Reiss et al. 2007).

The second problem is: the large number of combinations makes it difficult to find the optimal stimulus in a limited recording time. Some studies have tried an adaptive stimulus

optimization technique to maximize the firing rate of single neurons (Nelken, Prut et al. 1994, O'Connor, Petkov et al. 2005). In those studies, an optimal stimulus is defined as the one that produces the highest firing rate. However, from neural coding aspect, the stimulus that maximally changes the firing rate (maximal slope) is also informative.

A general statistical data-collection method (optimal experimental design, OED) has been proposed in a recent work for hierarchical neural networks: a small number of optimally designed stimuli can yield a model estimation as good as a large number of random independent and identically distributed (IID) samples (DiMattina and Zhang 2011). This method could potentially reduce the recording time needed for an accurate estimation. In addition, the responses elicited by this method should potentially span the entire dynamic range of firing rate. This method has been applied to the estimation of generalized linear models, where fewer stimuli chosen adaptively can give an accurate estimation (Lewi, Schneider et al. 2011), but not been broadly used in in-vivo neurophysiology experiments except the study by Tam and Young in inferior colliculus (Tam 2012).

In this chapter, we combined neural modeling and the statistical data-collection method in the neurophysiology experiment to study spectral integration in auditory cortex. In this study, a feed-forward network model with five subunits was used to approximate the computation of cortical neurons. Those subunits represented potential excitatory and inhibitory synaptic inputs. Stimuli were generated adaptively to optimize the parameter estimation of the hypothesized model based on the recorded firing rates from previous stimuli. Our results show that the optimal design algorithm can use less than 300 stimuli for estimating the parameters of the neural network. The model can potentially predict firing rates of various neurons and also capture the nonlinear spectral integration of some cortical neurons. Our preliminary data also revealed possible frequency receptive field structures underlying diverse spectral selectivity.

5.2 Results

A total of 92 neurons were recorded from three hemispheres of two marmoset monkeys (left hemisphere of M79u, both hemispheres of M73v). All neurons were tested with pure tones, two tones, harmonic complex tones, inharmonic complex tones (Methods). The entire data set was divided into training set (75% of all stimuli) and validation set (25% of all stimuli). For each neuron, the feed-forward model as in Figure 2.2 was estimated using the training data set and then the predictive capability was tested on the validation set. We used the correlation coefficient (r) to quantify the performance of the model in predicting neural responses. A correlation coefficient close to 1 indicate good prediction. A total of 66 neurons were tested in the online adaptive stimulus design.

5.2.1, A generalized feedforward model for studying spectral integration in auditory cortex

The rational for choosing multiple excitatory subunits was to allow the formation of complex frequency receptive fields with more than one excitatory peak in order to account for the modulation effect from extra-classical receptive field and combination sensitivity shown on harmonic template neurons in Chapter3 (Shamma and Symmes 1985, Sutter, Schreiner et al. 1999, Kadia and Wang 2003). However, not all subunits will turn to be useful after fitting: the synaptic connection can be close to 0. Therefore, this model can still account for simple receptive fields, such as excitation peak surrounded by lateral inhibition. The example neuron shown in Figure 5.1 is a good demonstration. After fitting, only one excitatory subunit and one inhibitory subunit had significant weights. The estimated model gave a good predication for both training data set and validation dataset. One interesting observation from the estimated model was the overlapping of the excitatory and inhibitory subunits. We will have a detailed discussion later.

As we shown in Chapter 4, there are different subpopulations which showed distinct selectivity to complex tones. The diversity in responses suggests different structure of the frequency receptive fields. Different subpopulations might also have different level of nonlinearity in spectral integration. Therefore, we evaluated the model's performance across all 92 neurons in four different groups defined in chapter 4. Figure 5.2 shows the distributions of the correlation coefficient (r) for four different groups. On average, the validation data were often harder for a model to predict than the training data. This was expected because this simplified model cannot account for the possible nonlinearity. The performance varied greatly from neuron to neuron. However, our model had a better performance on the frequency selective neurons and band-pass neurons for both training set and validation set in general. For harmonic template neurons, the prediction was good on training set but poor on validation set. The correlation coefficients were also listed in Table 5.1. The across group variance suggests that frequency selective neuron and band-pass neurons are more linear than harmonic template neurons and neurons that were not categorized. Overall, the five subunits feedforward network model can account for the spectral selectivity of cortical neuron to some extent.

One surprising result was that co-tuning occurred quite frequently in the estimated network. Figure 5.3A shows examples of different model estimates. In the first example, one excitatory subunit and one inhibitory subunit overlapped with each other. In other words, the two Gaussian weights have almost identical centers and variances. The appearance of co-tuned subunits cannot be dismissed as a mere coincidence at the population level. Figure 5.3B shows the distribution of the co-tuning index over the whole population. Out of the 92 recorded neurons, 18 neurons had a COI lower than 0.2 (20%). Those neurons with co-tuned subunits had a broad BF range (Figure 5.4A). Intuitively, a co-tuned pair would yield a non-monotonic rate-level function. However, half of the 18 neurons had monotonic rate-level function at their BFs (Figure 5.4B).

5.2.2, *Application of OED experiments*

From the offline data fitting, we have shown that the simplified feedforward network model can capture the stimulus-response relationship to some extent. The next question is whether the online adaptive search method could give an estimation as good as the model from offline fitting. During the OED experiment, the model was trained over 300 iterations. Figure 5.5 showed how the optimal stimuli developed during the process. In our design, there are 31 bins covering a broad frequency range of three octaves. The 300 stimuli can be either narrow-band and sparse or broadband. However, pure tones were relatively rare. In the example, the model started with random guess, but after 50 iterations, the designed stimuli had spectral energy focusing around frequencies that the neuron was most sensitive to, which can either strongly excite or inhibit the example neuron. The firing rate to each stimulus was plotted below. Some of the stimuli increased the neural firing rate (above the dashed line) while others suppress the firing rate (below the dashed line).

The estimated model was plotted in Figure 5.6A in terms of the synaptic input weights for all subunits. The excitatory units were spread in different frequency regions. The excitatory subunit with the largest peak was near the estimated BF from other test. The only inhibitory unit left was also centered at BF but slightly wider than the excitatory subunit. The model gave a good prediction of the frequency tuning and non-monotonic rate-level function of the example neuron even though the model was not trained on pure tones (Figure 5.6B, C).

5.2.3 *Finding the “optimal stimuli”*

As mentioned earlier in this chapter, our approach aims to find the “optimal stimuli” which drive a neuron in its full dynamic range. In order to test whether the online model was

optimized to do such task, thirty “optimal stimuli” were used as a separate validation set. The thirty stimuli were picked according to the predicted responses to randomly generated 1,000,000 stimuli with different spectral profiles from the model: ten stimuli which invoked high firing rates, ten median, and ten low. Neural responses to the thirty stimuli were recorded and compared with the model prediction. As shown in Figure 5.7, although the model was not as successful to give a good prediction of the absolute firing rates, it captured how the firing rate changes with different spectral contents.

Previous adaptive methods aimed to find the “optimal stimuli” that has the highest probability to make a neuron fire a spike. Some neurons in auditory cortex are suppressed by sound stimuli. The decrease in firing rate encodes information of sounds. For such neurons, the adaptive methods have to adjust the optimization algorithm in order to find the “optimal stimulus” which maximally suppresses the firing rate. The OED method in our study aims at driving a neuron in its full dynamic range, which by default could not only find stimuli that excite the neuron but also stimuli that suppress the neuron. Therefore, this method can also study neurons which only decreases firing rate to stimuli. An example neuron was shown in Figure 5.8. The final model estimated online only had inhibitory subunits, which explained the suppressed firing rate by most stimuli.

For most neurons that we tested with OED, regular stimuli, such as pure tones and complex tones were also used to characterize individual neuron. For some neurons which did not respond to any simple stimuli tested, the OED method was also able to find stimuli which could drive those neurons. One example is shown in Figure 5.9. The “optimal stimuli” designed by the online model were less structured, but showed a preference to broadband stimuli.

5.2.4 Receptive field structure underlying the spectral selectivity

We applied OED experiment on 66 neurons. The model revealed different receptive field structure with distinct selectivity. As described in Chapter 4, the adjacent neuron pairs have little similarity in their selectivity to complex tones even their BFs are close. It becomes an interesting question whether the OED model is accurate enough in order to capture the details of the local microcircuits within a small frequency range.

In one experiment, we managed to hold two neurons long enough to have the models for each neuron estimated online separately. The first neuron was the example shown in Figure 5.2 with overlapped excitatory and inhibitory subunits centered at BF (Figure 5.9A, left). The model showed a preference of this neuron to sparse tones because of the surrounding inhibitory (Figure 5.9B and C, left). For the second neuron, two excitatory subunits had their centers shifted from the BF and one broadly tuned inhibitory subunit covered the area between the two excitatory peaks including BF, where was the excitatory area for the first neuron (Figure 5.9A). The second neuron responded maximally to two tones simultaneously and did not respond to stimuli with energy at BF.

These examples showed that the OED method has a high accuracy to estimate the receptive field structure even in a small frequency range. It also proved that within 300 iterations, OED can efficiently sample the most interesting areas for individual neuron in a high dimensional space to give a good estimation of the stimulus-response relationship. This method would potentially provide more insights to our observations of the little similarity in the spectral selectivity of adjacent neuron pairs.

5.2.5 Evaluation of OED

In the OED experiment, the more the iterations are, the higher probability is to get a low mean predictive error (MPE) (Sheiner and Beal 1981). However, considering the limited recording time, only 300 iterations were used. Because a simulation has shown that there is no significant improvement in predication when the iteration increases from 300 to 400 given the complicity of our model (Dekel 2012). In a fixed number of stimuli, the estimated model will depend on the stimulus samples used. In order to evaluate the model estimated online, we also did offline fitting combining all the stimuli used for characterizing a neuron and the 30 optimal stimuli designed online excluding the 300 stimuli during the adaptive searching to get another estimation as comparison. The entire data set was divided randomly to a training set with 75% of the total stimuli and the validation set with the rest 25%. For most neurons, the training set included more than 300 stimuli. The parameters of the model were estimated using a multi-start, data-fitting minimization of the MSE between the neural responses and the model, starting from a 1000 initial guesses. This offline fitting presumably will give a more optimized model. However we have to keep in mind that the large amount of computation for the offline fitting and optimization is not feasible for an online experiment.

A comparison between the online model and the offline optimized model was based on 66 neurons. On average, the offline model could predict the neural responses better than the model estimated online (Figure 5.10). We also compared the models on different subsets of data: the optimal stimuli generated online, pure tone, complex tones (harmonic complex tones and inharmonic complex tones). There was no significant difference in the performance in predicting the responses to the 30 optimal stimuli generated online. But the model generated online was not as good as the optimized offline model in predicting pure tone responses. The prediction to complex tones was the worse among all three sets. Table 5.2 summarizes the results of the OED online model and the optimized offline model, including a t-test comparison between the two

models. A possible reason for the poor performance on pure tones might be the training stimuli generated online were mostly broadband. As for complex tones, because of the nonlinear summation of harmonic complex tones, our feed-forward model might not be adequate.

5.3 Discussion

5.3.1 Summary of experiments

We combined neurophysiological recording and computational modeling to investigate the role of the interaction of excitatory and inhibitory receptive field components in shaping the spectral selectivity of neurons in A1 of awake marmosets. We found that simple feedforward networks, composed of a relatively small number of excitatory and inhibitory subunits (Figure 2.2), were capable of capturing the complex responses of many A1 neurons to stimuli with varying spectral contents to some extent, ranging from pure-tones to multi-band stimuli. In particular, the network models were able to predict the responses to novel stimuli that were not used for parameter estimation.

Based on the feed-forward model, we implemented an online adaptive method which is based on an information-theoretic criterion that maximize the mutual information between the presented stimuli and the expected response given the neuron's response history and the current parameters estimated in our neurophysiology experiment. The adaptively designed 300 stimuli can efficiently search in the high dimensional space to drive neurons in their full firing rate range. The estimated model online could accurately capture the difference in receptive field structure even within a small frequency range.

5.3.2 *Co-tuned excitatory and inhibitory inputs*

The best-fitting models for many neurons showed co-tuning; that is, the networks contained excitatory and inhibitory subunits that had almost identical center frequencies and almost identical tuning widths (Figure 5.3). We analyzed those neurons with co-tuned subunits, focusing especially on their underlying network mechanisms. The center of co-tuned excitation and inhibition was spread throughout the hearing range of marmoset. A neuron with co-tuned subunits had either monotonic or non-monotonic rate-level function at its best frequency, depending on the distance between the best frequency and the frequency center of the co-tuned pair (Dekel 2012). When the co-tuned pair was centered far away from the best frequency, the rate-level function tended to be monotonic. By contrast, when the co-tuned pair was centered close to the best frequency, the rate-level function tended to be non-monotonic, peaking at some intermediate sound level. In this situation, the non-monotonicity was produced predominantly by the projections from the co-tuned pair to the principal unit.

We also considered broader theoretical implications of our findings. Co-tuning implies degeneracy of a synaptic weight matrix, which is a necessary condition that allows a network to generate peaked responses around the optimal stimulus (DiMattina and Zhang 2007). Although the relationship between optimal stimulus and weight matrix degeneracy may seem abstract or even counterintuitive, it provides a conceptual bridge from co-tuning to non-monotonic turning curves and robust selectivity to specific high-dimensional stimulus features.

Our results showed that co-tuning could emerge solely from spectral models without involving any temporal role at all. This is because our models were focused strictly on spectral information processing while temporal dynamics was completely ignored. Of course, this finding should by no means be interpreted as a contradiction to any temporal function of co-tuning (Wehr and Zador 2003, Zhang, Tan et al. 2003, Tan, Atencio et al. 2007). In reality, co-tuning probably has both spectral and temporal roles to play in the auditory cortex.

Similar to previous experimental and theoretical results (de la Rocha, Marchetti et al. 2008, Wu, Arbuckle et al. 2008) , our findings do not imply that lateral inhibition and co-tuning are mutually exclusive in the neural architecture underlying A1 neurons. On the contrary, most estimated models containing co-tuning also contained lateral inhibition to some extent. The biological system probably uses both configurations together as building blocks to construct more elaborate circuitries.

5.3.3 Implication for spectral integration in auditory system

Many natural and man-made sounds, such as animal vocalizations, human speech and music have broad spectrum. Although the peripheral auditory system segregates sounds into narrow frequency channels, frequency components of one sound source needed to be grouped together and separated from components of other sound sources, usually known as auditory scene analysis (Bregman 1990). To understand how neurons in central auditory system analyze the spectrum provided by the peripheral auditory system and form representations of individual sound sources is very important for understanding the grouping and segregation in perception. Neurons in A1 showed narrow tuning to pure tones, but intracellular recording and two-tone experiments revealed much broader subthreshold excitatory and inhibitory inputs (Merzenich, Knight et al. 1975, Recanzone, Guard et al. 2000, Linden, Liu et al. 2003, Moshitch, Las et al. 2006, Sadagopan and Wang 2008). These data suggest that neurons in A1 are candidates for spectral integration and coding of complex spectral profiles. By using complex tones in addition to single tones, we were able to study neural response properties and selectivity in a higher dimensional stimulus feature space.

A network model with excitatory and inhibitory subunits covering different frequency areas helped reduce the dimensionality in the description of neural responses. Instead of

specifying a lookup table that describes the response of a neuron to all possible combination of stimulus features, which quickly becomes impossible as the stimulus dimension increases, one only needs to specify the parameters of a network model, which can potentially predict the response to any stimulus and therefore carry full information about the global stimulus-response relationship. Our models also captured the nonlinearity at multiple levels in the network in addition to the linear integration of synaptic inputs which are attainable by methods such as Spectrotemporal receptive fields (Theunissen, Sen et al. 2000) and random spectral stimuli estimations (Yu and Young 2000, Barbour and Wang 2003). There is no inherent theoretical limitation on the complexity of neural responses that can be accommodated by a network model, because these networks are universal approximators that can approximate any input-output function to arbitrary precision when enough subunits are included (Cybenko 1989, Hornik, Stinchcombe et al. 1989). It would be useful to extend our method to exploit the feature selectivity of nonlinear neurons in higher auditory areas.

5.3.4 Network models used in studying neural response properties in AI

The neural network models used in our study were highly simplified but proved to be adequate for the purpose of this work. In the hierarchical networks, each unit at the bottom preferred only a single frequency while the unit at the top integrated excitatory and inhibitory inputs from below to generate more complex spectral selectivity that spanned a wide frequency range, not unlike the real system to a first approximation.

Our current models included only feedforward excitation and inhibition. A recurrent network with additional feedback excitation and inhibition may potentially offer a more expressive model for describing neuronal responses. By incorporating temporal dynamics and recurrent connections into network models (de la Rocha, Marchetti et al. 2008, Schinkel-Bielefeld,

David et al. 2012), we might be able to account for the behaviors of a large proportion of auditory cortical neurons more precisely. Although biological realism is always desirable in modeling, we need to keep in mind that an increase in model complexity should be accompanied by adequate decrease in model fitting error and, more importantly, model prediction error.

There are many biological details that could be accommodated into future versions of our models. For example, thalamocortical inputs and intracortical inputs received by A1 neurons are thought to be separated in both spectral and temporal domains: the thalamocortical pathway preferentially mediates fast response to BF tones, while the intracortical pathway mediates late response to non-BF stimuli. For near-BF tones, inputs come from both pathways, with the thalamocortical inputs arriving before the intracortical inputs (Kaur, Lazar et al. 2004, Liu, Wu et al. 2007, Happel, Jeschke et al. 2010). The thalamic and intracortical projections might contribute differently to the neural selectivity to spectral contents. Besides other limitations mentioned earlier, our current models did not make such fine differentiations, and future improvement should come from incorporating more biologically realistic features into the models.

In addition, there are different types of GABAergic interneurons which have distinct morphology and postsynaptic targets (Markram, Toledo-Rodriguez et al. 2004). An important open question is to model such subtypes of inhibitory neurons differently and clarify their respective functional roles in cortical processing. Now with the development of optogenetic tools, it is possible to target certain types of neurons and manipulate distinct neural circuitry (Lee, Kwan et al. 2012, Wilson, Runyan et al. 2012). It will be interesting in the future to combine our modeling approach with such experimental manipulations. Because of the online adaptive method, it's possible to have an estimated model in 300 iterations, usually a few minutes. The estimated models under different conditions may uncover systematic changes in the functional network structures. It will also be interesting to test how attention or change of behavioral states may alter

network parameters such as the gains, thereby changing the balance of excitation and inhibition and the efficiency of neural coding (Fritz, Shamma et al. 2003).

In summary, in the future work, modeling and experiments should be further integrated to gain a better understanding of cortical functions and underlying anatomic structures.

Table 5.1 The predication of the model to training and validation set

(Median of the correlation coefficient of the neural responses and model predictions)

	Frequency selective	Band-pass	Harmonic template	Uncategorized
Training set	0.808	0.703	0.809	0.693
Validation set	0.647	0.548	0.597	0.256
P-value	< 0.0001	0.0001	0.0048	0.0064

Table 5.2 Comparison between online model and offline model

(Medians of correlation coefficient of neural responses and predicted responses)

	All data set	Optimal stimuli	Pure tones	Complex tones
Online Model	0.362	0.442	0.398	0.2
Offline Model	0.593	0.547	0.655	0.512
P-value	<0.0001	0.97	0.002	<0.0001

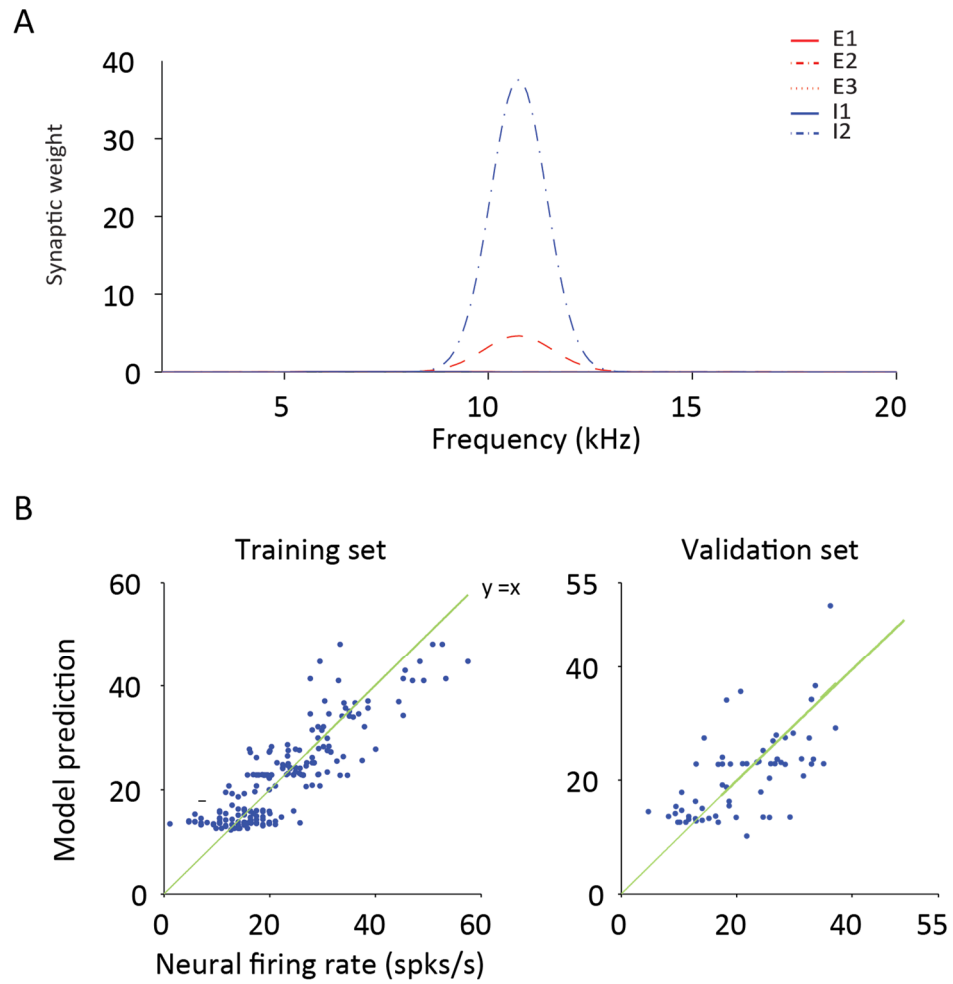


Figure 5.1 An example of the fitted feed-forward model

A, the synaptic weights for five subunits of an estimated model.

B, the model prediction of the firing rates to stimuli in the training set (left) and validation set (right)

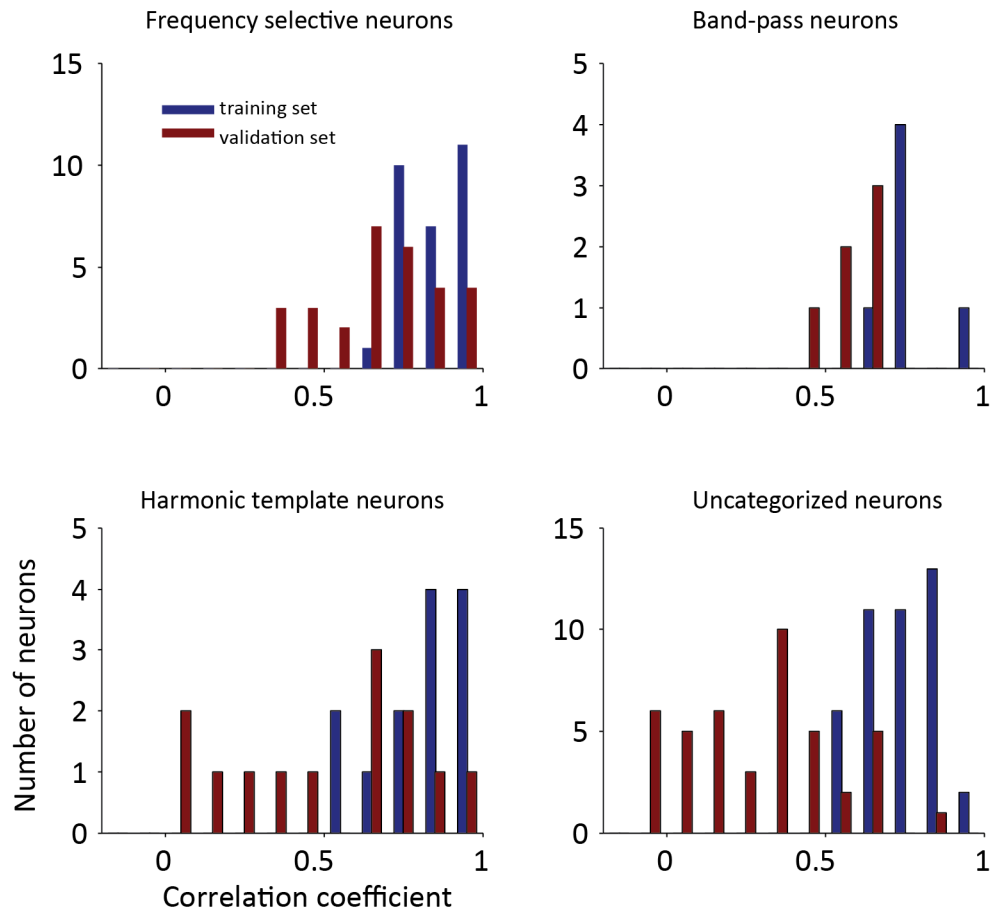


Figure 5.2 The distribution of correlation coefficient for model predictions for training and validation sets in four subgroups of neurons

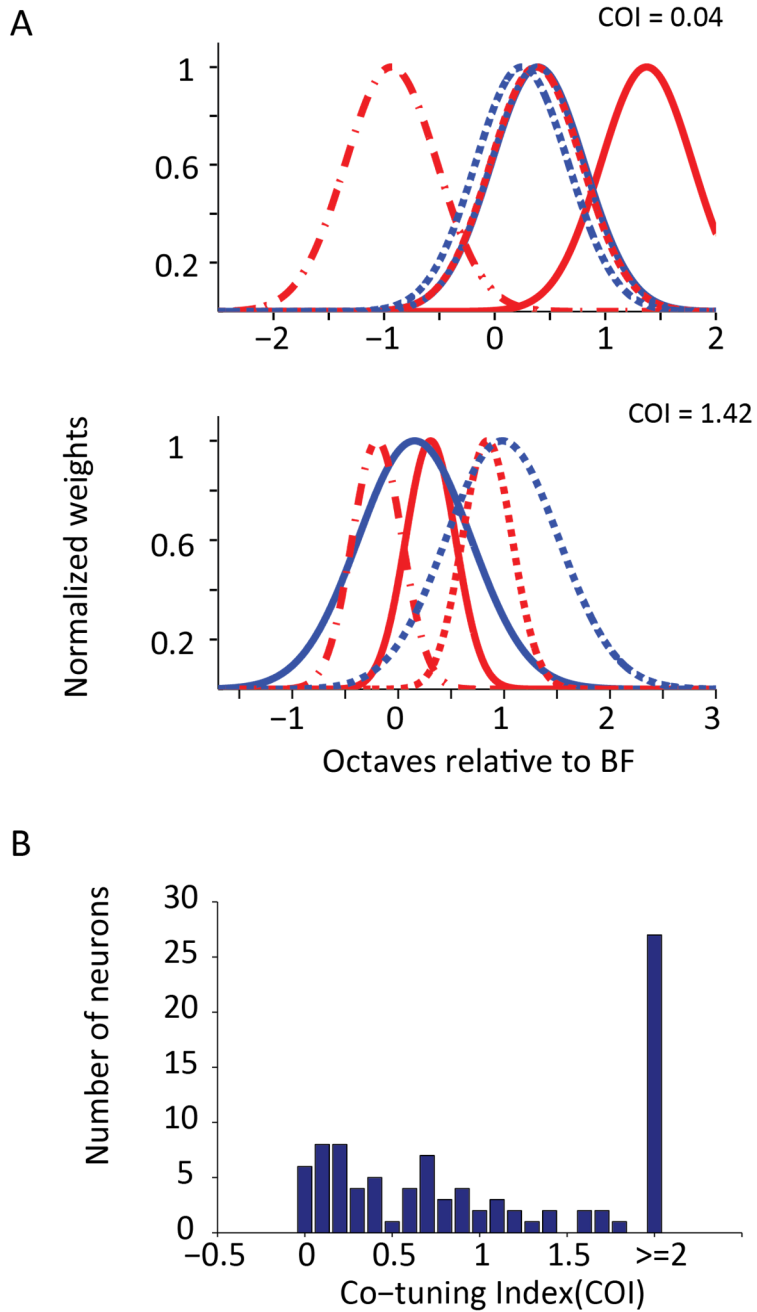


Figure 5.3 co-tuned excitatory and inhibitory subunits

A, examples of subunits synaptic weight functions. The neuron in the top figure showed overlapped excitatory and inhibitory subunits. The other neuron in the bottom figure had a more distributed subunits.

B, The distribution of co-tuning index for all 92 neuron

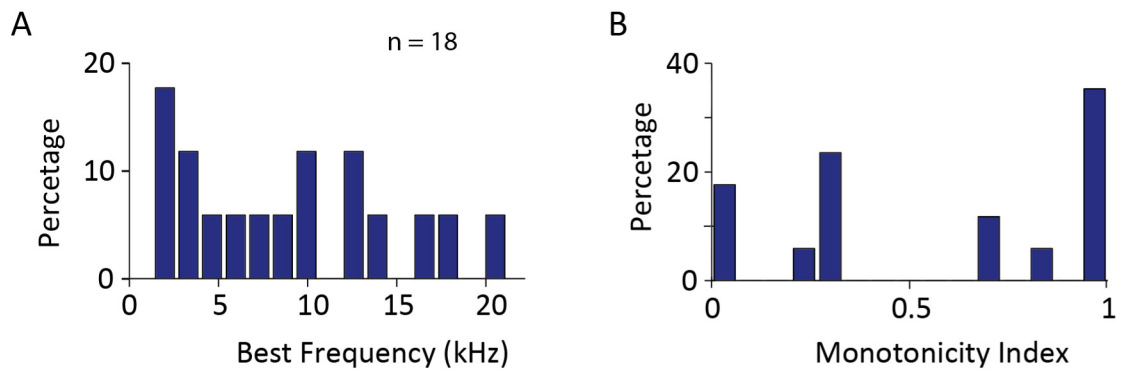


Figure 5.4 Tone response properties of neurons with co-tuned subunits

A, the distribution of BFs of 18 neurons which had co-tuned excitatory and inhibitory subunits in the estimated model.

B, the distribution of the monotonicity index at BF for the same 18 neurons.

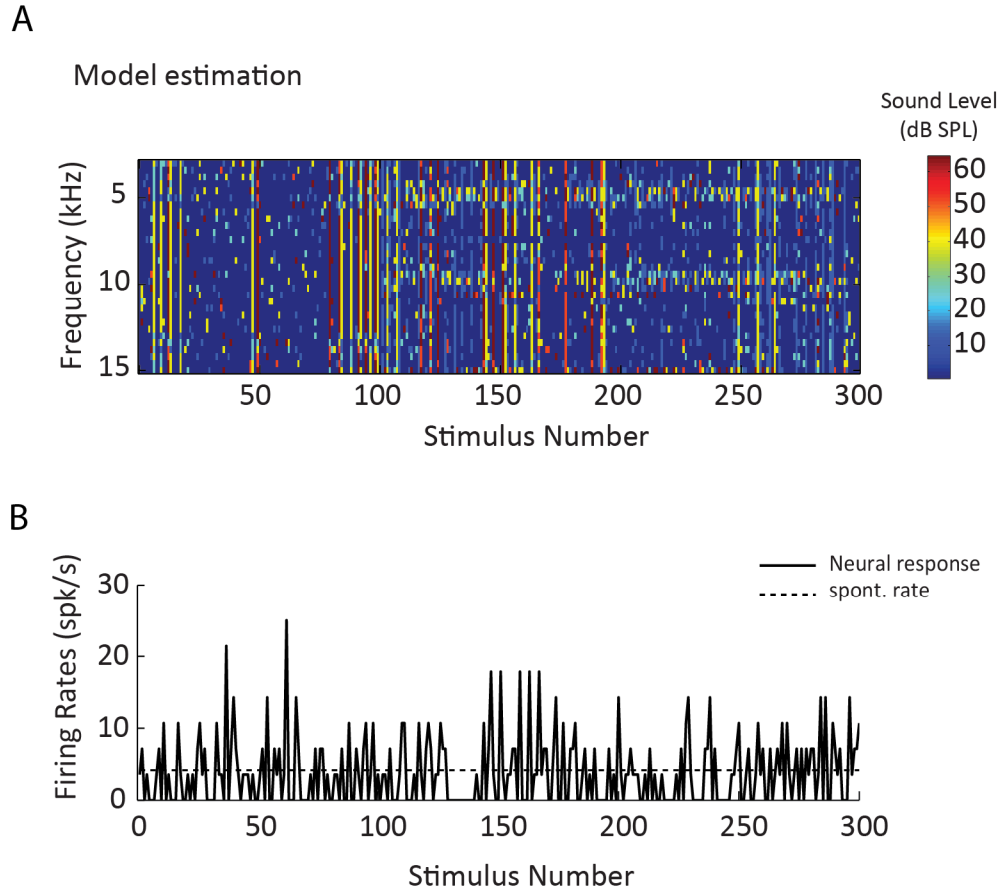


Figure 5.5 An example of one run for the OED experiment

A, the 300 stimuli developed during the OED experiment.
 B, the neural responses to the 300 stimuli during search.

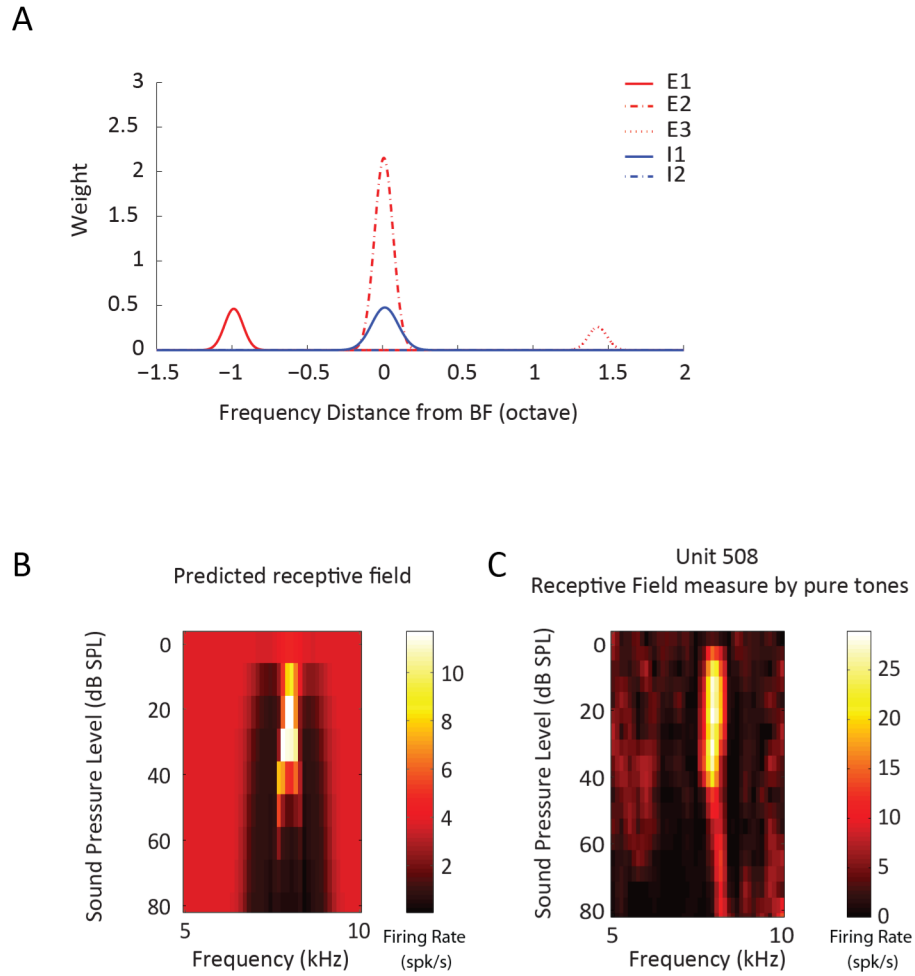


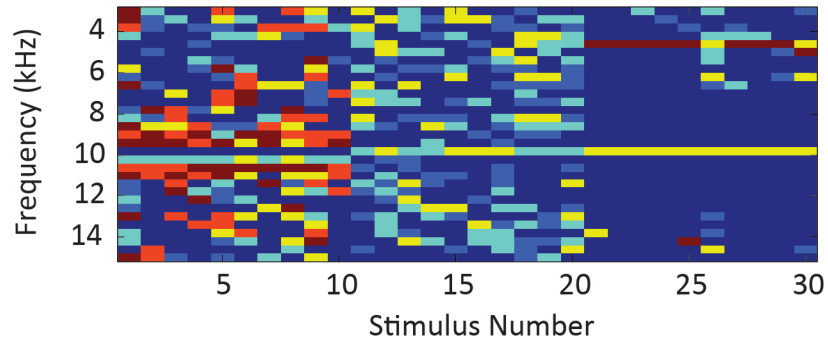
Figure 5.6 The estimated model online and prediction of responses to pure tones

A, The synaptic weighting functions for all subunits in the online estimated model.

B, The tone response map predicted by the online model. C, Response map measured by pure tones

A

Optimal stimuli



B

correlation coefficient = 0.68

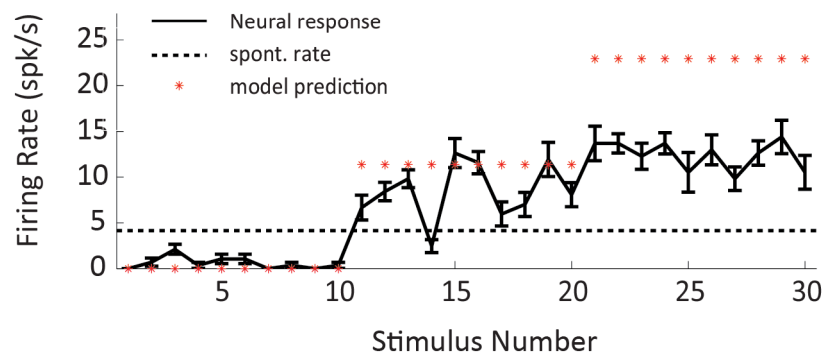


Figure 5.7 An example of the optimal stimuli designed online.

A, the spectrum of thirty stimuli designed by the online model: stimulus 1-10 will give low firing rate, 11-20 median firing rate and 21-30 high firing rate.

B, the predicted responses and real neural responses to the optimal stimuli

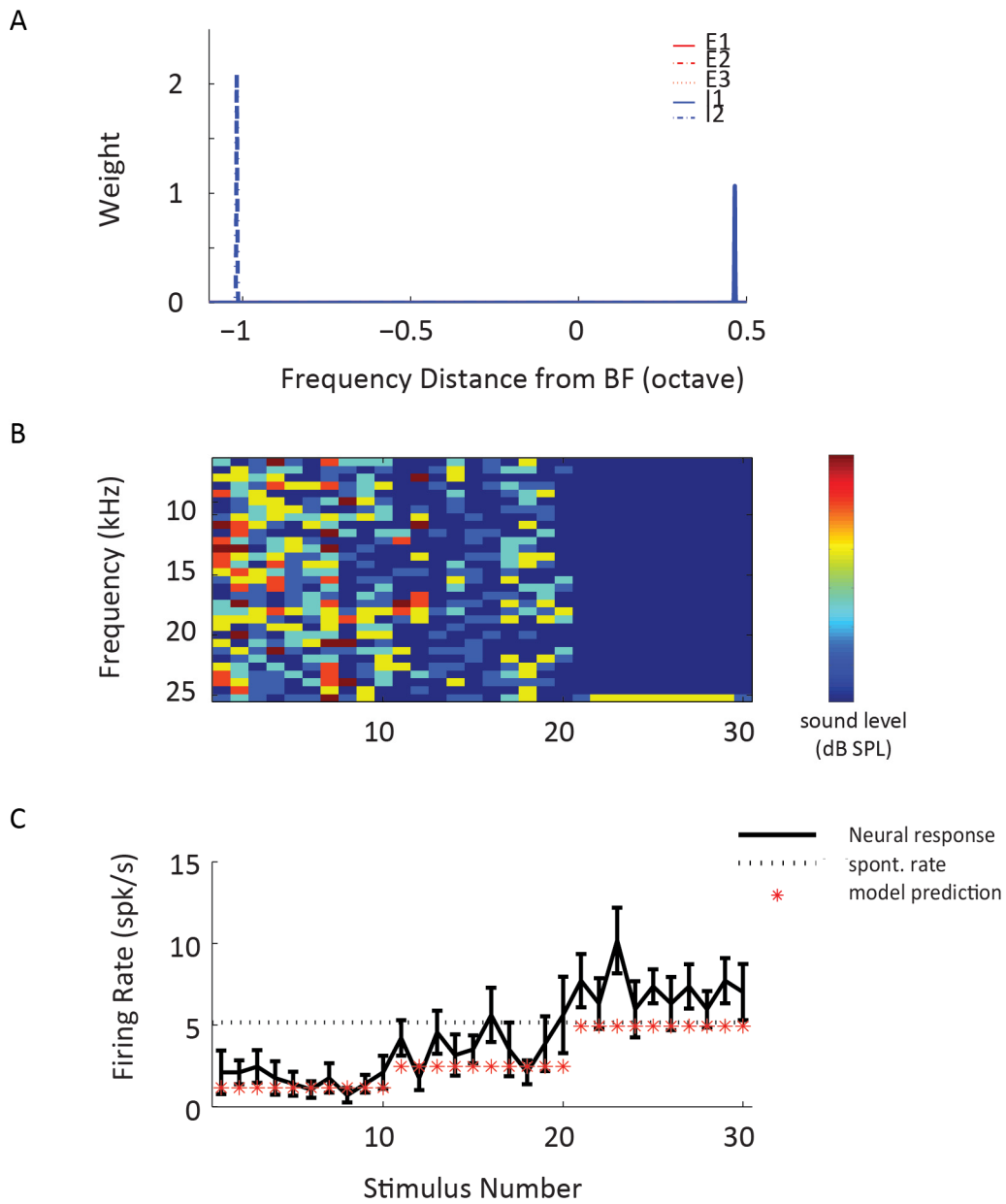


Figure 5.8 An example of a neuron which was suppressed by sound stimuli.

A, the synaptic weighting function of five subunits in the online estimated model.

B, the spectrum of the optimal stimuli designed by the online model.

C, the predicted responses and real neural responses to the optimal stimuli.

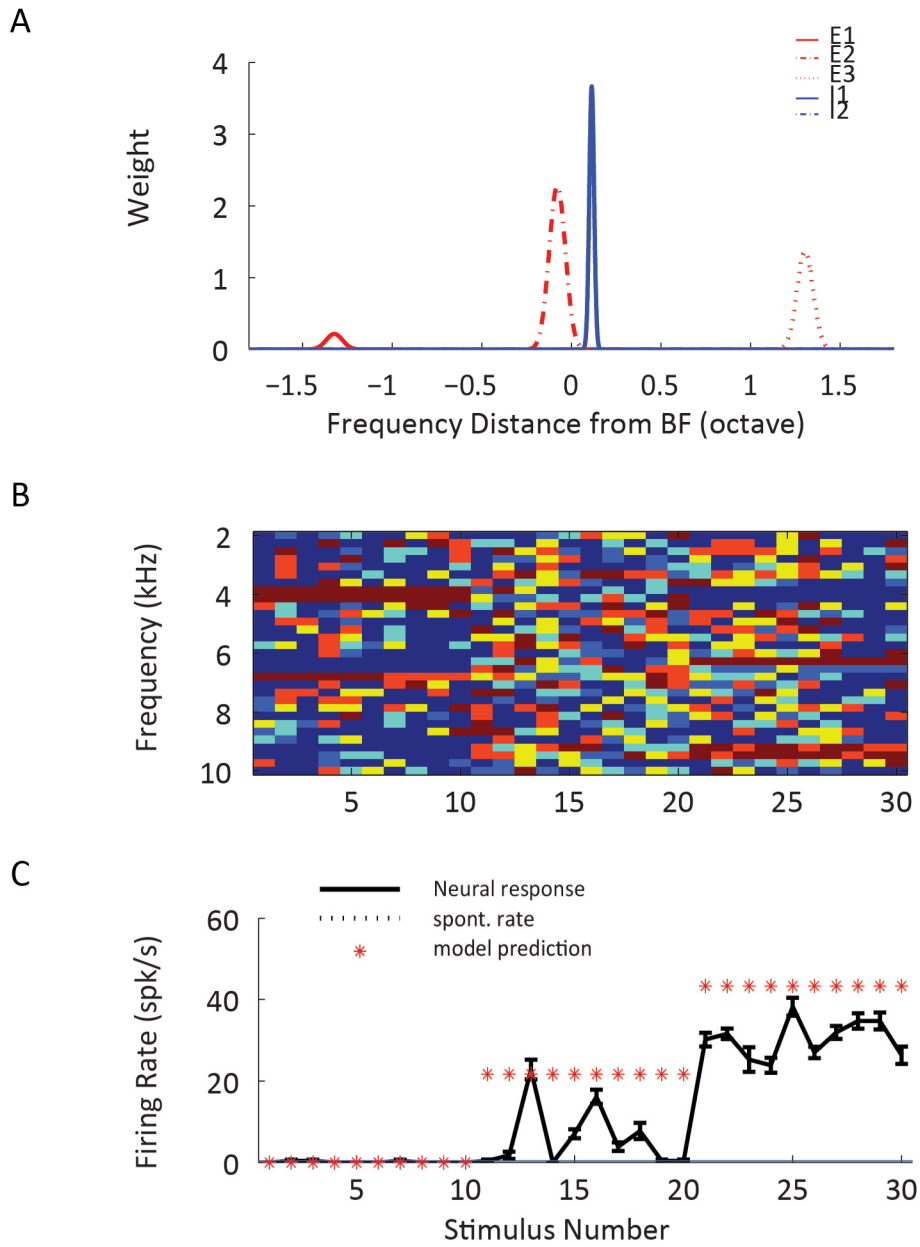


Figure 5-9 An example of a neuron which selectively responded to broadband stimuli.

- A, the synaptic weighting function of five subunits in the online estimated model.
- B, the spectrum of the optimal stimuli designed by the online model.
- C, the predicted responses and real neural responses to the optimal stimuli.

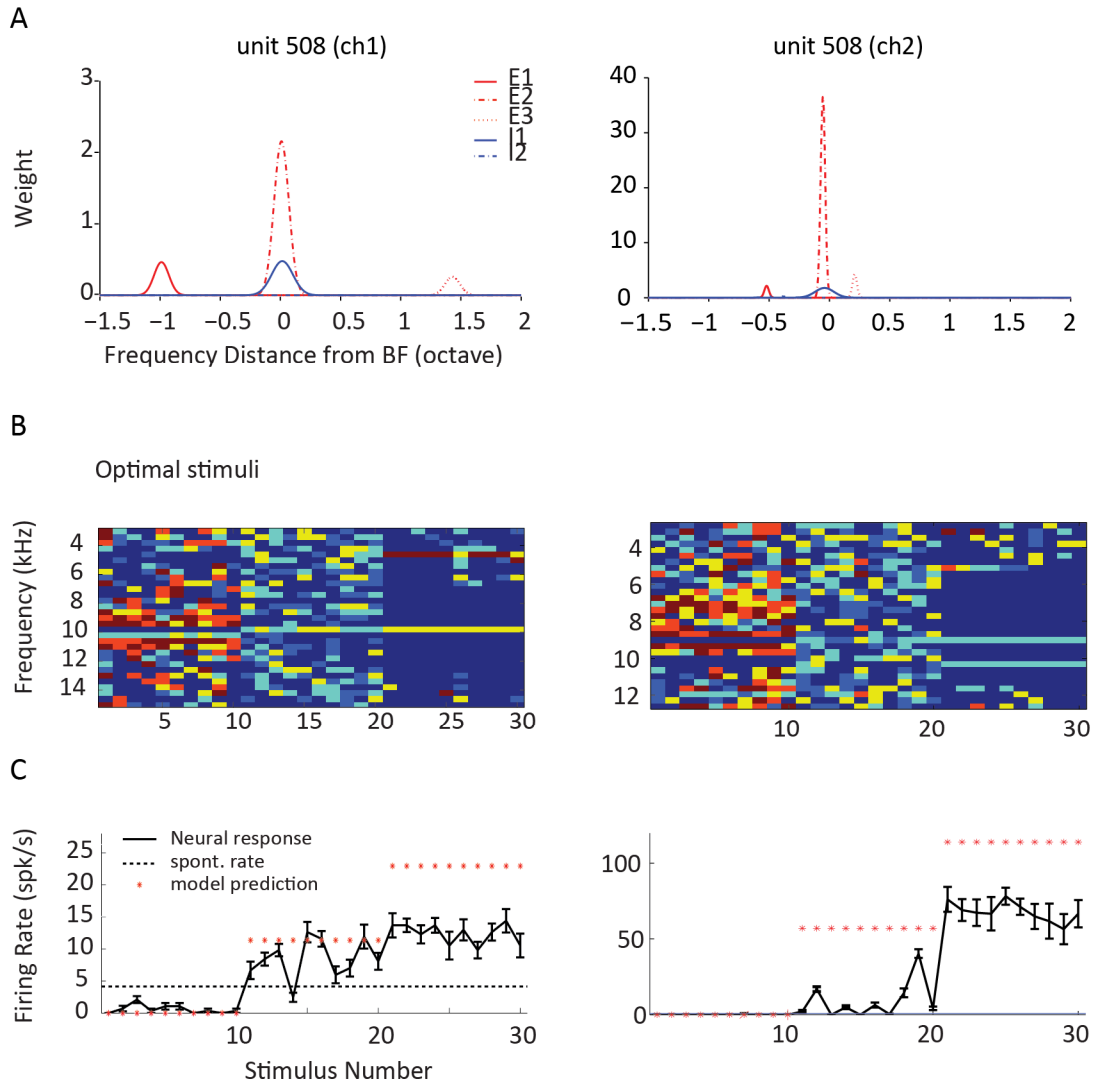


Figure 5-10 An example of a neuron pair with different spectral selectivity.

A, the synaptic weighting function of five subunits in the online estimated model for neuron 1 and 2.

B, the spectrum of the optimal stimuli designed by the online models.

C, the predicted responses and real neural responses to the optimal stimuli.

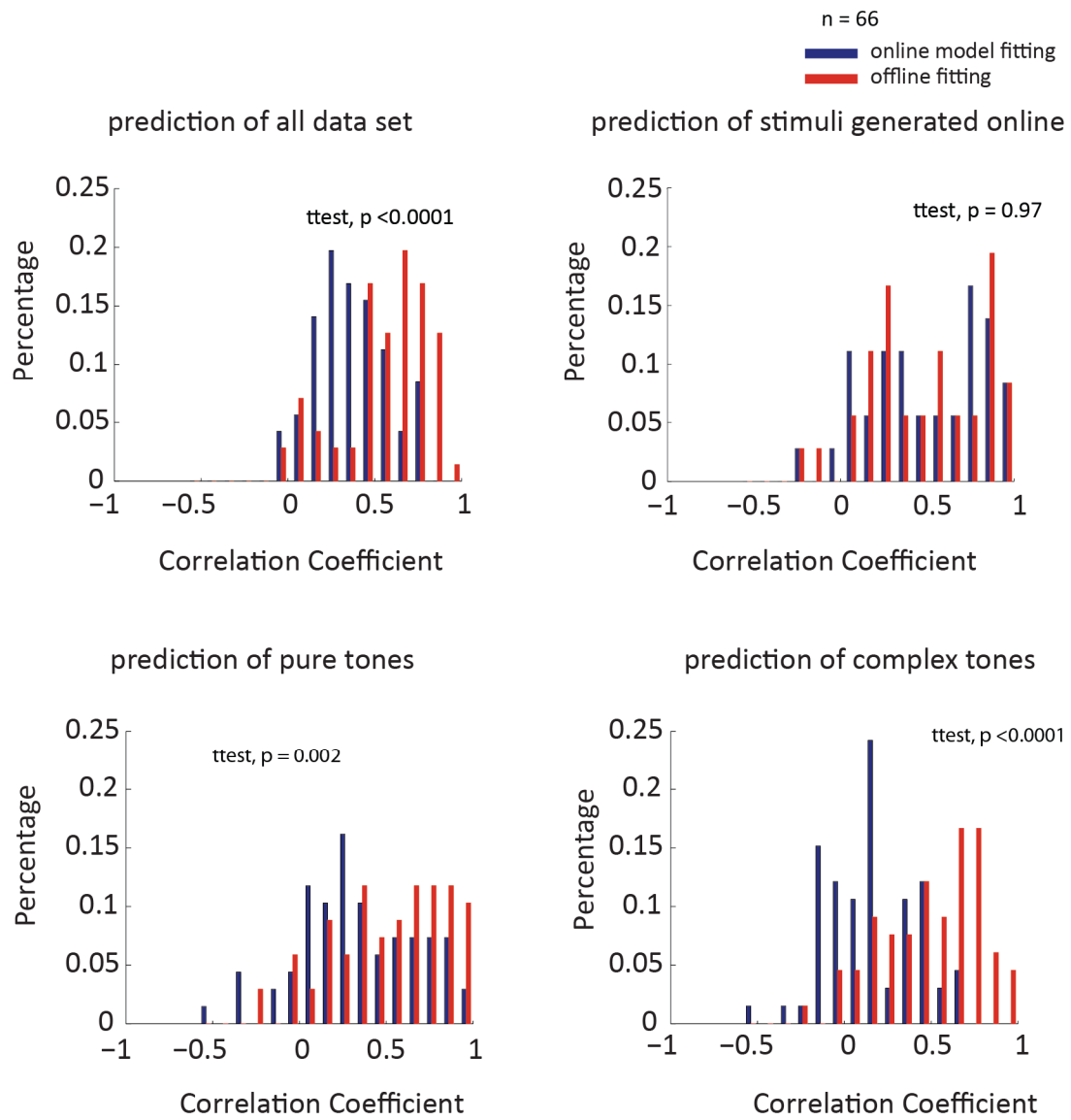


Figure 5-11 The distribution of correlation coefficients for online model and offline model in four different subgroup of neurons

CHAPTER 6:

6. Conclusion

This dissertation has combined neurophysiology and modeling to study the spectral integration and neural representation of harmonic complex tones in the auditory cortex of awake marmosets.

6.1 The existence of harmonic template neurons in A1

6.1.1 Harmonic template neurons

Chapter 3 reported the finding of harmonic template neurons in A1. Harmonic template neurons show strong responses to concurrent harmonics of a preferred f_0 even if individual frequency evokes weak or no response. In the test of shifted harmonic complex tones, harmonic template neurons can distinguish harmonic shifts from inharmonic shift and odd harmonic shifts. The first criterion is for select combination sensitive neurons which encode features specific for harmonic complex tones, rather than information of individual components. The second criterion makes sure those neurons are selective to spectral but not temporal information. A spectral template is for harmonics that can be resolved by the peripheral auditory system. For humans, the perceived pitch shifts with the inharmonic shifts and ambiguous pitches were perceived for odd-harmonic complex tones if resolved harmonics are included (Moore and Moore, 2003; Patterson and Wightman, 1976). On the contrary, the shifts had less or negligible effects on perceived pitch of unresolved harmonics because the temporal cue was used (Moore and Moore, 2003). Therefore,

any candidate for harmonic template should be able to distinguish the inharmonic shifts and the odd-harmonic complex tones.

6.1.2 Comparison between harmonic template neurons and pitch selective neurons

Harmonic template neurons are not pitch-selective neurons reported in previous studies (Bendor and Wang, 2005, Bendor et al. 2012). First of all, pitch neurons respond to all stimuli at the same pitch even the spectrum is different. Whereas harmonic template neurons only respond to stimuli contain frequencies matching the template. A good example would be on the neuron shown in Figure 3-1. There were two separated peaks in the f_0 tuning ($f_{01}=1.2$, $f_{02}=1.05$) and there was no response to any f_0 in between. This discontinuity in f_0 tuning was very different from pitch neurons. It suggested there is some other feature that determines the response rather than pitch itself. By analyzing the spectrums for both f_{01} and f_{02} , we found both the 5 – 7th harmonics of f_{01} and the 6-8th harmonics of f_{02} were close to the spectral template provided. Such ambiguity is also shown in the optimum pitch model (Goldstein, 1979). Secondly, pitch neurons' responses to the harmonic complex tones decrease when the lowest harmonic number increases because the pitch salience decreases. Harmonic template neurons don't respond to a harmonic complex tone that does not include the template even it has high pitch salience. Third, the preferred f_0 s of pitch neurons are at their BFs and they respond equally well to pure tones at BFs. For harmonic template neurons, the preferred f_0 is either at BF or subharmonics of BF ($BF/2$, $BF/3$, $BF/4$,...). In other words, the BF has to be a harmonic of F_0 , but not necessarily the f_0 . For one f_0 , there are several templates distributed in the tonotopic map for detecting different harmonics. Therefore, harmonic template neurons don't encode f_0 per se, but represent a stage where the f_0 can be extracted. Because harmonic template neurons encode the absolute frequencies, they can potentially be used for coding of timbre.

6.1.3 Harmonic template neurons and harmonic sensitive neurons in bats

The response properties of harmonic template neurons are very similar to the harmonic sensitive neurons found in bat auditory cortex, which showed facilitation when the first harmonic was simultaneously delivered with one or more higher harmonics (Suga, 1979). However, harmonic sensitive neurons were found in a specialized region for detecting biosonar signals for echolocation. They are selective to high frequencies around 30kHz. The harmonic template neurons found in this study codes a broad range of f_0 from 400Hz to 12kHz and they were distributed in the entire frequency range of A1 and organized tonotopically by their frequency center. Marmosets can hear up to 36kHz (Osmanski and Wang, 2010) and the fundamentals of four major marmoset calls were between 5kHz to 10kHz (Dimartinna and Wang, 2005). The harmonic template neurons are not limited to the frequency range of marmoset vocalizations. They may represent a common structure in A1 for processing harmonic sounds which are very common in the acoustical environment. Harmonic template neurons provide an integrative representation of harmonic sounds that could be used further for pitch extraction and recognition of conspecifics vocalizations.

6.2 A distributed harmonic process in A1

6.2.1 A parallel process of harmonic complex tones

Chapter 4 describes different subpopulations of neurons in A1 with distinct response patterns to harmonic complex tones. Frequency selective neurons resolve individual harmonics for low harmonic number. Band-pass neurons are less frequency selective but tune to the overall sound intensity. Harmonic template neurons are sensitive to combination of harmonics. Modulation sensitive neurons encode the envelope modulation. There was also a group of neurons which cannot fully characterize by our measure.

In order to respond to a sound source fast, a parallel spectral analysis has to be done on different scales to efficiently extract features associated with perceptual attributes from the physical features (frequency and level). For example, the discrete frequency components need to be grouped together to estimate f_0 if the fundamental frequency itself is not presented and to compute the overall sound level. In addition, the accurate representation of individual harmonics is important for timbre perception.

The diversity in neural responses to harmonic complex tones suggests A1 is an important stage in auditory processing, where pre-processed information of harmonic complex tones can be carried by different population of neurons and can be used further for extracting perceptual features, such as loudness, pitch and timber to form a code of ‘auditory objects’ at later stages.

6.2.2 Inhibition and a level-invariant representation of spectral information

For AN fibers and AVCN neurons, the ability to resolve harmonics in their rate response degrades rapidly with increasing stimulus levels, even the low and moderate stimulus levels (15-20dB above threshold) were used to minimize rate saturation (Smoorenburg and Linschoten 1977, Cedolin and Delgutte 2005). Although the rate-place representation in the AN fibers provide sufficient information to account for the behavior performance at high sound levels as shown in previous studies, it fails to explain the general robust psychophysical performance at high sound levels (Hirsh, Reynolds et al. 1954). Accurate frequency information need to be maintained in order to use a ‘‘harmonic template’’ to extract pitch information (Goldstein 1973) and timbre information. It has been shown that a population code of AN fibers combining rate, space, temporal information provide a robust code for vowels over a 80dB SPL range.

The different finding in our study was the level-invariant representation of the spectral information over a relative wide range (up to 40dB difference) at single neuron level. The multi-

peak response pattern of frequency selective neurons remains even at higher levels when the rate – place representation at AN fibers start to degrade. The preferred f0s for harmonic template neurons did not change at different levels tested. Those observations suggest that cortical processing enhanced the frequency discriminability for complex tones. Although we were not able to measure inhibition with the extracellular recording method, our data have implied that inhibition play a very important role in this processing. This level invariant representation of spectral information is necessary and crucial for the auditory system, because it operates in a broad dynamic range because of the distance to the sound source and background noise.

6.2.3. Functional structures for spectral processing in A1

In Chapter 4, we observed that different subpopulations are intermingled in A1 and distributed in a large frequency range. Simultaneously recorded neuron pairs showed small signal correlations in responses to complex tones even their BFs were close to each other. There was no functional cluster observed yet in this study. This finding was different from the primary visual cortex, which is organized in functional columns. However, from the computation aspect of view, a disruption of the organized rate-place codes from the peripheral auditory system is the start of information transformation from the isomorphic representation of frequency and level to a non-isomorphic representation of perceptual attributes and integrative representation of “object” at higher levels beyond A1.

6.3 An online adaptive method for studying spectral integration in A1

6.3.1 Finding the ‘optimal’ stimuli

The majority of studies in auditory cortex were trying to find the “optimal stimuli” for each neuron, because they help us understand how individual neuron encodes information of

physical spaces. In order to find the true “optimal stimuli”, not the sub-optimal stimuli, it’s not enough to use a subset of stimuli. The combination selectivity requires a search in a high dimensional feature space, which becomes infeasible in a limited experiment time. Chapter 5 describes a new approach to efficiently search in a high dimensional feature space to find the “optimal stimuli”.

The most common definition for optimal stimulus is the stimulus that has the highest probability to firing a spike. However, an optimal stimulus can also be the one to which a neuron is most sensitive. In other words, the change of firing rate will be the largest by changing the stimulus away from the optimal, corresponding to the maximal slope in a tuning curve. In our approach, we used an information-theoretic criterion that maximizes the mutual information between the presented stimuli and the expected response given the neuron’s response history and the current parameters estimated. In this way, our model could design not only the optimal stimuli to maximize firing rate, but also stimuli to minimize the firing rate. In other word, our optimal stimuli can drive neurons at its full dynamic range.

Even with the simplified feed-forward network model, our approach could accurately capture the stimulus-response relationship for some neurons within 300 iteration within a few minutes (5-10min). 300 iterations were only half the size of the entire stimulus set including pure tones, complex tones in experiments in Chapter 3 and 4, which were designed manually by the experimenter. The stimulus design was largely dependent on the experience and knowledge of the experimenters. The online adaptive stimulus design approach is an automatically searching procedure which requires little information of each neuron. For example, the neuron shown in Figure 5.9 which did not respond to simple tone, complex tones tested. The OED was still able to find stimuli which could drive this neuron. Although our program requires to provide BF information to decide the frequency range, the three-octave range was broad enough for most neurons. Therefore it won’t affect the optimization even if BF estimated was not accurate.

6.3.2 Frequency receptive field structure underlying the spectral selectivity

Our model recovered different frequency receptive field structures for neurons with different spectral selectivity. Although the synaptic weighting functions of subunits in the network model might not be the actual synaptic inputs distribution because it's a linear approximation, it provide possible structures for certain functional properties which give more insights to understand the variability of individual neurons in the spectral processing. Those hypothesized structures could potentially to be verified or tested with other technics in the future.

6.4 Methodological considerations for studying auditory cortex

In this thesis, I systematically tested neurons in auditory cortex with harmonic and inharmonic complex tones and explored a new method for neurophysiology experiments combining neural network model and the statistical data-collection method. There were interesting findings in this study but also revealed new challenges. The variability and complexity in neural responses to complex tones made it very difficult to interpolate the results. The boundary for each subgroups was less obvious even in a multi-dimensional metrics. No single criterion is enough to fully characterize the response properties. There are some methodological principles I learnt from my experiments for future studies in auditory cortex:

- 1, It's necessary to use complex stimuli to study neurons in auditory cortex because of the high selectivity. We really need to try to “drive” each neuron to understand the real stimulus-response relationship.

- 2, It's necessary to take into account single neuron variability. Any form of averaging (such as multi-units) will possibly lose some important information when we evaluate the functional properties of neurons in A1.

3, Computational models are very useful tools for understanding the real stimulus-response relationship in a high dimensional feature space.

CHAPTER 7:

7. Miscellaneous findings

7.1 Temporal response pattern

In this dissertation, the neural response was quantified based on the spike count within a time window from 15ms after stimulus onset until 50ms after offset, which ignored temporal firing patterns within this time window. In some neurons, I observed temporal response pattern changes during different stimuli. The first example neuron (Figure 7.1A) was a frequency selective neuron with the typical response pattern: large firing rate at integer harmonic numbers and firing rate decreases for high harmonic number. The firing pattern changed with harmonic number as well. For the low harmonic numbers, the firing started shortly after stimulus onset and lasted through the stimulus. However, starting from the harmonic number at 3, the neuron was suppressed after stimulus onset then started to fire again. The duration for such suppression became longer for higher harmonic numbers. The second neuron (Figure 7.1B) showed similar trend in its response temporal pattern. The sustained response during stimulus degraded rapidly as the harmonic number increased and completely disappeared after harmonic number 3. The offset responses remained until harmonic number 10.

Two other examples were from the subgroup of band-pass neurons (Figure 7.1C and D). They showed rate response patterns like auditory nerve fibers. However, within the subgroup, even neurons change average firing rates in a similar way, the temporal response pattern could be largely different. The first neuron (Figure 7.1C) exhibited onset responses to all HCTs. For large harmonic numbers, suppression started to appear after the onset response and also the offset

rebound. The second neuron showed an opposite trend. For low harmonic numbers, there was onset response followed by suppression lasting until 100ms after offset. There were also response rebounds after the suppression. However, when the spectrum became denser, the suppression was weaker so that the onset response turned into a sustained response.

In the future work, the temporal response pattern should be considered for studying spectral integration because the integration of “object parts” at higher level sensory cortex can be a dynamic process (Brincat and Connor 2006). The temporal pattern, such as the onset, sustained, even suppression might carry different information of the “object”. A model taking account into temporal parameters, such as a recurrent network model, should be considered.

In Chapter 3 and 4, we discussed about the possible role of inhibition in sharpening frequency selectivity. It appears here that inhibition affects a larger population of neurons than just harmonic template neurons and frequency selective neurons in A1. The role of inhibition in neural coding might be underestimated in our study because only the averaged firing rate was used. The examples showed here suggest inhibitory inputs and excitatory inputs have more specific spectral-temporal interaction patterns, which might not be surprising. Because previous studies already showed that inhibition and excitation occurred in a precise temporal sequence (Wehr and Zador 2003, Zhang, Tan et al. 2003, Tan, Zhang et al. 2004), which could shape the time course of the spike responses, frequency tuning and direction selectivity of AI neurons. However, it's still unknown how the stereotypical temporal patterns of tone-evoked inhibition and excitation in those studies can associate with the diverse temporal response pattern to complex tones. Although the integration of simultaneous auditory components is very important in a scene analysis (Bregman 1990), we have to keep in mind that our auditory system has to track time-varying sounds most of the time in a natural environment. Future works need to combine spectral and temporal integration together.

7.2 Auditory grouping

Our auditory system relies on different cues to parse the mixed acoustical signals in order to correctly group all components from the same sound source together and separate them from the ones from a different sound source. Important cues include the onset synchrony, the frequency components of one sound tend to start together; harmonicity, all the components from one sound source are usually harmonically related; space, frequencies coming from different locations are likely from different sources. However, the three cues are not equally weighted our auditory system. For example, a resolved harmonic does not contribute to the pitch of a complex tone if the onset time is different and the difference is larger than 160ms (Darwin and Ciocca 1992). Another study showed that subjects would perceive the same pitch even if the odd harmonics and even harmonics were delivered to different ears via headphone (Bernstein and Oxenham 2008).

Harmonic template neurons were shown to have the ability to group simultaneously presented harmonics. One interesting question would be whether the grouping is constrained by the onset timing as shown in psychoacoustic studies. I tested the onset synchrony on one of the harmonic template neurons. In this test, all even harmonics were shifted in time from leading the odd harmonics to following the odd harmonics (Figure 7.2A). The black bars indicate the overlapping parts of all harmonics. As shown in Figure 7.2B, onset time was very important for the grouping based on harmonicity. This neuron only responded during the overlap parts. When the overlap time was less than 50ms, or the onset time difference was larger than 150ms, there was no response. Another interesting observation is that odd harmonics dominated the responses. The neuron still responded to odd harmonic even the response was much weaker compared to the responses to all harmonic. When the odd harmonics were leading, the responses always stopped after the odd harmonic offset. However, when the even harmonics were leading, the response could still last even after the offset of the even harmonics, until the odd harmonics were off.

Another test was separating even harmonics and odd harmonics in space. There were 16 speakers in the chamber (8 speakers at 0 degree elevation, 30° apart on azimuth; 8 speakers at 45 degree above horizontal plane, 30° apart on azimuth). All speakers were at equal distance (1 m) from the center of the head of the animal. This neuron responded to the harmonic complex tone including all harmonics from all 16 speakers. In the spatial test, the even harmonic were always played from the speaker horizontal plane, 60 degree from the midline contralaterally. The odd harmonics were played from other different speakers (Figure 7.3A). All harmonics were played simultaneously. Although odd harmonics or even harmonics alone evoked no or weak response, the response was largely increased when they were played simultaneously even from different locations. This result suggests that this neuron can integrate harmonics from different locations.

Although only two harmonic template neurons were tested in the spatial separation task, the results were similar. Those preliminary data provide evidence for common grouping rules both in perception and at single neurons in A1. Further tests are needed for investigate the neural mechanism underlying auditory grouping across time, frequency and space.

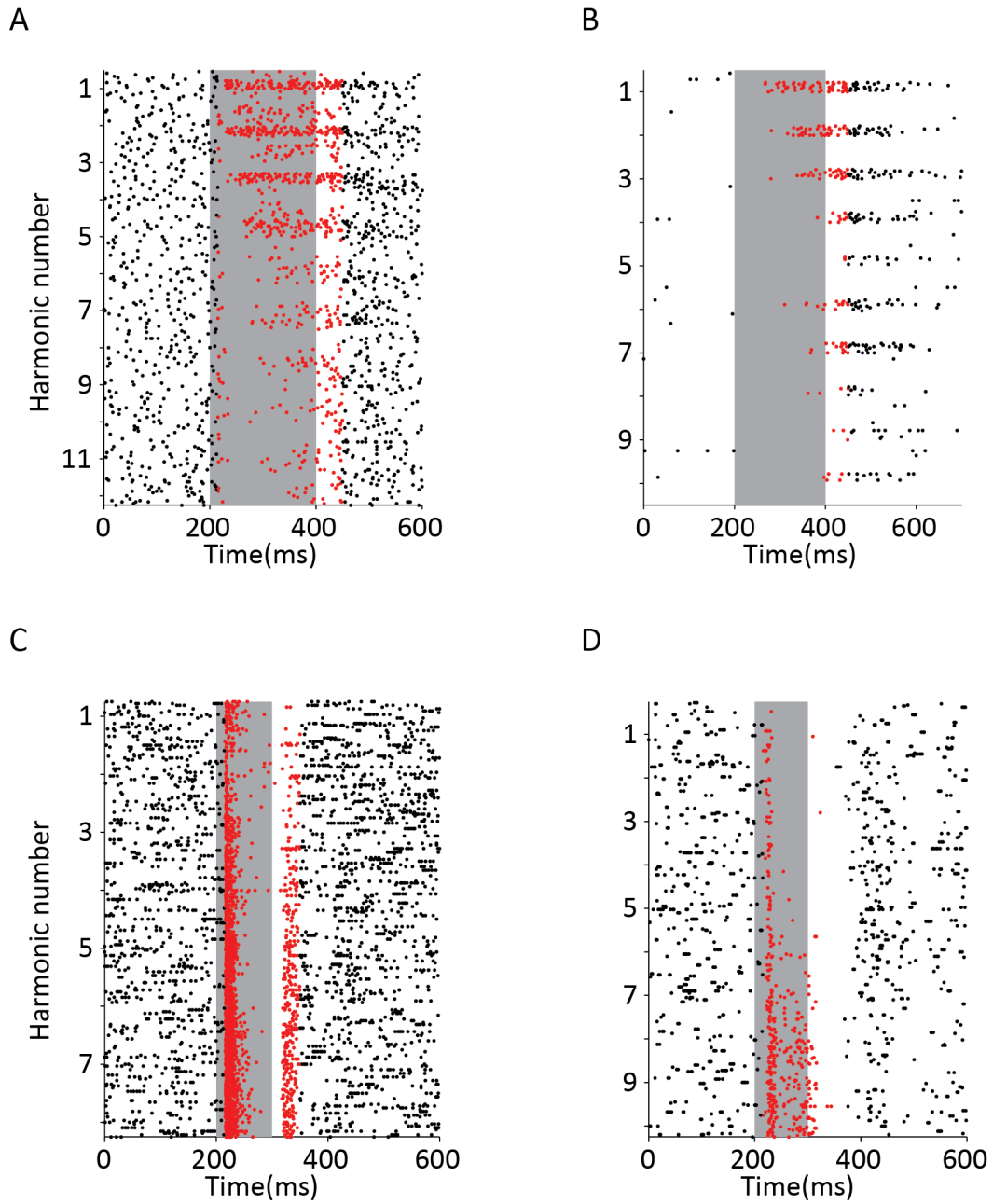


Figure 7.1 Four examples of the temporal response patterns in response to HCTs

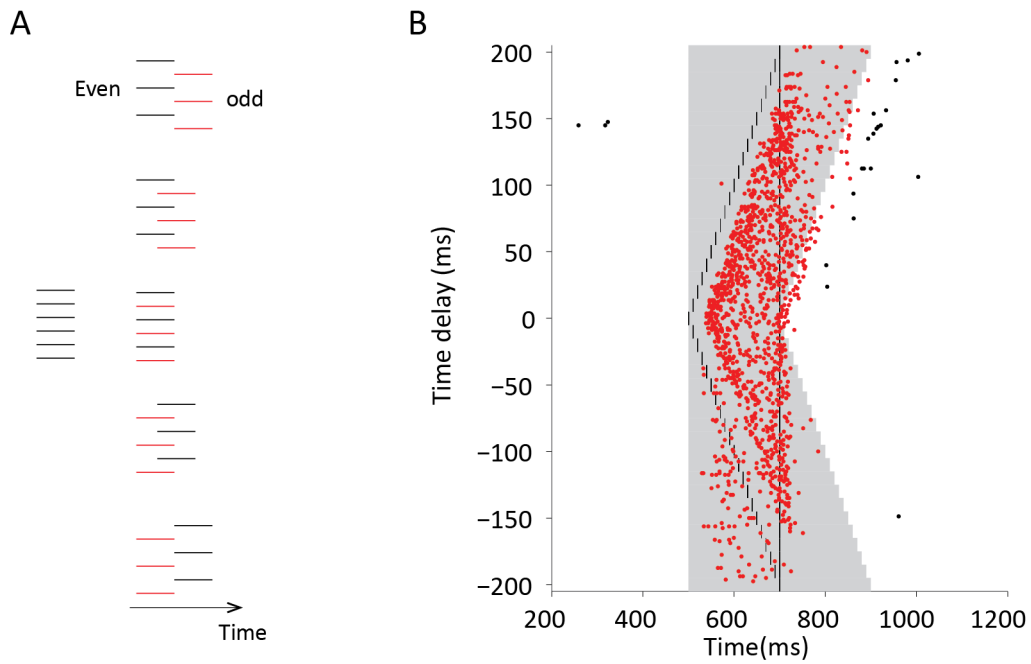


Figure 7.2 onset synchrony test for harmonic grouping

A, a diagram of the unsynchronized harmonic tones.

B, raster plot of a harmonic template neuron's responses to the unsynchronized harmonic tones

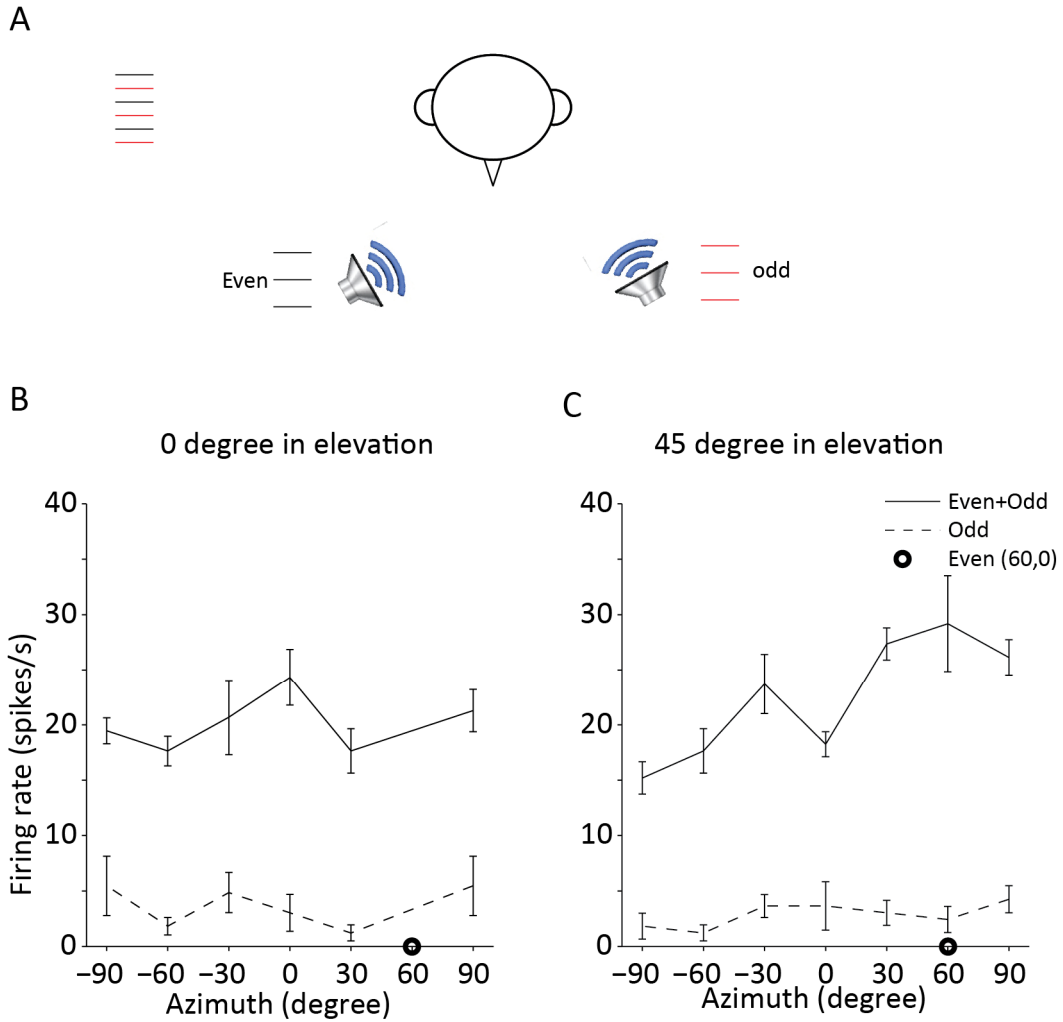


Figure 7.3 harmonic grouping across space

A, a diagram of the spatially separated even and odd harmonics.

B, the responses of a harmonic template neuron when the even harmonics and odd harmonics were separated on azimuth

C, B, the responses of a harmonic template neuron when the even harmonics and odd harmonics were separated both on azimuth and elevation

REFERENCES

1. Abeles, M. and M. H. Goldstein, Jr. (1972). "Responses of single units in the primary auditory cortex of the cat to tones and to tone pairs." Brain research **42**(2): 337-352.
2. Adrian, E. D. and R. Matthews (1927). "The action of light on the eye: Part I. The discharge of impulses in the optic nerve and its relation to the electric changes in the retina." J Physiol **63**(4): 378-414.
3. Adrian, E. D. and R. Matthews (1927). "The action of light on the eye: Part II. The processes involved in retinal excitation." J Physiol **64**(3): 279-301.
4. Bair, W., E. Zohary and W. T. Newsome (2001). "Correlated Firing in Macaque Visual Area MT: Time Scales and Relationship to Behavior." The Journal of Neuroscience **21**(5): 1676-1697.
5. Baker, C. I., M. Behrmann and C. R. Olson (2002). "Impact of learning on representation of parts and wholes in monkey inferotemporal cortex." Nat Neurosci **5**(11): 1210-1216.
6. Bandyopadhyay, S., L. A. Reiss and E. D. Young (2007). "Receptive field for dorsal cochlear nucleus neurons at multiple sound levels." J Neurophysiol **98**(6): 3505-3515.
7. Bandyopadhyay, S., S. A. Shamma and P. O. Kanold (2010). "Dichotomy of functional organization in the mouse auditory cortex." Nature neuroscience **13**(3): 361-368.
8. Barbour, D. L. and X. Wang (2003). "Auditory cortical responses elicited in awake primates by random spectrum stimuli." The Journal of neuroscience **23**(18): 7194-7206.
9. Bendor, D., M. S. Osmanski and X. Wang (2012). "Dual-pitch processing mechanisms in primate auditory cortex." The Journal of neuroscience **32**(46): 16149-16161.
10. Bendor, D. and X. Wang (2005). "The neuronal representation of pitch in primate auditory cortex." Nature **436**(7054): 1161-1165.
11. Bernstein, J. G. and A. J. Oxenham (2008). "Harmonic segregation through mistuning can improve fundamental frequency discrimination." J Acoust Soc Am **124**(3): 1653-1667.

12. Bernstein, L. R. and D. M. Green (1987). "Detection of Simple and Complex Changes of Spectral Shape." Journal of the Acoustical Society of America **82**(5): 1587-1592.
13. Bernstein, L. R. and D. M. Green (1987). "The Profile-Analysis Bandwidth." Journal of the Acoustical Society of America **81**(6): 1888-1895.
14. Bregman, A. S. (1990). Andutiroy Scene Analysis. Cambridge, MA, MIT Press.
15. Brincat, S. L. and C. E. Connor (2004). "Underlying principles of visual shape selectivity in posterior inferotemporal cortex." Nat Neurosci **7**(8): 880-886.
16. Brincat, S. L. and C. E. Connor (2006). "Dynamic shape synthesis in posterior inferotemporal cortex." Neuron **49**(1): 17-24.
17. Bruce, I. C., M. B. Sachs and E. D. Young (2003). "An auditory-periphery model of the effects of acoustic trauma on auditory nerve responses." J Acoust Soc Am **113**(1): 369-388.
18. Brugge, J. F. and M. M. Merzenich (1973). "Responses of neurons in auditory cortex of the macaque monkey to monaural and binaural stimulation." Journal of neurophysiology **36**(6): 1138-1158.
19. Cedolin, L. and B. Delgutte (2005). "Pitch of complex tones: rate-place and interspike interval representations in the auditory nerve." Journal of neurophysiology **94**(1): 347-362.
20. Christianson, G. B., M. Sahani and J. F. Linden (2008). "The consequences of response nonlinearities for interpretation of spectrotemporal receptive fields." J Neurosci **28**(2): 446-455.
21. Connor, C. E., S. L. Brincat and A. Pasupathy (2007). "Transformation of shape information in the ventral pathway." Curr Opin Neurobiol **17**(2): 140-147.
22. Coslett, H. B., H. R. Brashear and K. M. Heilman (1984). "Pure word deafness after bilateral primary auditory cortex infarcts." Neurology **34**(3): 347-352.
23. Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. Mathematics of Control. Signals and Systems: 303-314.

24. Cynx, J. and M. Shapiro (1986). "Perception of missing fundamental by a species of songbird (*Sturnus vulgaris*)." Journal of comparative psychology (Washington, D.C.: 1983) **100**(4): 356-360.
25. Darwin, C. J. and R. P. Carlyon (1995). Auditory grouping. Hearing. B. C. Moore. London, Academic Press. **2nd**: 387-424.
26. Darwin, C. J. and V. Ciocca (1992). "Grouping in pitch perception: effects of onset asynchrony and ear of presentation of a mistuned component." J Acoust Soc Am **91**(6): 3381-3390.
27. de la Rocha, J., C. Marchetti, M. Schiff and A. D. Reyes (2008). "Linking the response properties of cells in auditory cortex with network architecture: Cotuning versus lateral inhibition." Journal of Neuroscience **28**(37): 9151-9163.
28. Dekel, E. (2012). Adaptive on-line modeling in the auditory system - in vivo implementation of the optimal experimental design paradigm. M.S., Johns Hopkins University.
29. DiMattina, C. and X. Wang (2006). "Virtual vocalization stimuli for investigating neural representations of species-specific vocalizations." Journal of neurophysiology **95**(2): 1244-1262.
30. DiMattina, C. and K. Zhang (2007). "How Optimal Stimuli for Sensory Neurons Are Constrained by Network Architecture." Neural Computation **20**(3): 668-708.
31. DiMattina, C. and K. Zhang (2011). "Active data collection for efficient estimation and comparison of nonlinear neural models." Neural Comput **23**(9): 2242-2288.
32. Ebeling, M. (2008). "Neuronal periodicity detection as a basis for the perception of consonance: a mathematical model of tonal fusion." J Acoust Soc Am **124**(4): 2320-2329.
33. Evans, E. F., J. Rosenberg and J. P. Wilson (1971). "The frequency resolving power of the cochlea." J Physiol **216**(2): 58P-59P.

34. Evans, E. F. and I. C. Whitfield (1964). "Classification of Unit Responses in the Auditory Cortex of the Unanaesthetized and Unrestrained Cat." The Journal of physiology **171**: 476-493.
35. Feng, A. S. and P. M. Narins (2008). "Ultrasonic communication in concave-eared torrent frogs (*Amolops tormotus*)." Journal of comparative physiology.A, Neuroethology, sensory, neural, and behavioral physiology **194**(2): 159-167.
36. Fishman, Y. I., C. Micheyl and M. Steinschneider (2013). "Neural representation of harmonic complex tones in primary auditory cortex of the awake monkey." The Journal of neuroscience **33**(25): 10312-10323.
37. Fishman, Y. I., D. H. Reser, J. C. Arezzo and M. Steinschneider (1998). "Pitch vs. spectral encoding of harmonic complex tones in primary auditory cortex of the awake monkey." Brain Res **786**(1-2): 18-30.
38. Fritz, J., S. Shamma, M. Elhilali and D. Klein (2003). "Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex." Nat Neurosci **6**(11): 1216-1223.
39. Glasberg, B. R. and B. C. Moore (1990). "Derivation of auditory filter shapes from notched-noise data." Hearing research **47**(1-2): 103-138.
40. Goldstein, J. L. (1973). "An optimum processor theory for the central formation of the pitch of complex tones." The Journal of the Acoustical Society of America **54**(6): 1496-1516.
41. Hall, J. W., M. P. Haggard and M. A. Fernandes (1984). "Detection in noise by spectro-temporal pattern analysis." J Acoust Soc Am **76**(1): 50-56.
42. Happel, M. F., M. Jeschke and F. W. Ohl (2010). "Spectral integration in primary auditory cortex attributable to temporally precise convergence of thalamocortical and intracortical input." The Journal of neuroscience **30**(33): 11114-11127.
43. Heffner, H. and I. C. Whitfield (1976). "Perception of the missing fundamental by cats." The Journal of the Acoustical Society of America **59**(4): 915-919.

44. Heffner, H. H., R. (1986b). "Effect of unilateral and bilateral auditory cortex lesions on the discrimination of vocalizations by Japanese macaques." J Neurophysiol(56): 683-701.
45. Hirsch, J. A. and L. M. Martinez (2006). "Laminar processing in the visual cortical column." Curr Opin Neurobiol **16**(4): 377-384.
46. Hirsh, I. J., E. G. Reynolds and M. Joseph (1954). "Intelligibility of Different Speech Materials." Journal of the Acoustical Society of America **26**(4): 530-538.
47. Hornik, K., M. Stinchcombe and H. White (1989). "Multilayer Feedforward Networks Are Universal Approximators." Neural Networks **2**(5): 359-366.
48. Houtsma, A. J. M. and J. Smurzynski (1990). "Pitch Identification and Discrimination for Complex Tones with Many Harmonics." Journal of the Acoustical Society of America **87**(1): 304-310.
49. Hromádka, T., M. R. Deweese and A. M. Zador (2008). "Sparse representation of sounds in the unanesthetized auditory cortex." PLoS biology **6**(1): e16.
50. Hubel, D. H. W., T.N. (1968). "Receptive fields and functional architecture of monkey striate cortex." J Physiol(195): 215-243.
51. Issa, E. B. and X. Wang (2013). "Increased neural correlations in primate auditory cortex during slow-wave sleep." Journal of Neurophysiology **109**(11): 2732-2738.
52. Jeanne, James M., Tatyana O. Sharpee and Timothy Q. Gentner (2013). "Associative Learning Enhances Population Coding by Inverting Interneuronal Correlation Patterns." Neuron **78**(2): 352-363.
53. Jones, J. P. and L. A. Palmer (1987). "The two-dimensional spatial structure of simple receptive fields in cat striate cortex." J Neurophysiol **58**(6): 1187-1211.
54. Kadia, S. C. and X. Wang (2003). "Spectral integration in A1 of awake primates: neurons with single- and multi-peaked tuning characteristics." Journal of neurophysiology **89**(3): 1603-1622.

55. Kalluri, S., D. A. Depireux and S. A. Shamma (2008). "Perception and cortical neural coding of harmonic fusion in ferrets." J Acoust Soc Am **123**(5): 2701-2716.
56. Kaur, S., R. Lazar and R. Metherate (2004). "Intracortical pathways determine breadth of subthreshold frequency receptive fields in primary auditory cortex." Journal of neurophysiology **91**(6): 2551-2567.
57. Knudsen, E. I. and M. Konishi (1978). "A neural map of auditory space in the owl." Science **200**(4343): 795-797.
58. Krumhansl, C. L. (1990). The psychological representation of musical pitch in a tonal context, Oxford University Press.
59. Kudoh, M., Y. Nakayama, R. Hishida and K. Shibuki (2006). "Requirement of the auditory association cortex for discrimination of vowel-like sounds in rats." Neuroreport **17**(17): 1761-1766.
60. Kudoh, M. and K. Shibuki (2006). "Sound sequence discrimination learning motivated by reward requires dopaminergic D2 receptor activation in the rat auditory cortex." Learn Mem **13**(6): 690-698.
61. Lee, S. H., A. C. Kwan, S. Zhang, V. Phoumthippavong, J. G. Flannery, S. C. Masmanidis, H. Taniguchi, Z. J. Huang, F. Zhang, E. S. Boyden, K. Deisseroth and Y. Dan (2012). "Activation of specific interneurons improves V1 feature selectivity and visual perception." Nature **488**(7411): 379-383.
62. Leroy, S. A. and J. J. Wenstrup (2000). "Spectral integration in the inferior colliculus of the mustached bat." J Neurosci **20**(22): 8533-8541.
63. Levy, R. B. and A. D. Reyes (2011). "Coexistence of lateral and co-tuned inhibitory configurations in cortical networks." PLoS Comput Biol **7**(10): e1002161.
64. Lewi, J., R. Butera and L. Paninski (2009). "Sequential optimal design of neurophysiology experiments." Neural Comput **21**(3): 619-687.

65. Lewi, J., D. M. Schneider, S. M. Woolley and L. Paninski (2011). "Automating the design of informative sequences of sensory stimuli." J Comput Neurosci **30**(1): 181-200.
66. Linden, J. F., R. C. Liu, M. Sahani, C. E. Schreiner and M. M. Merzenich (2003). "Spectrotemporal structure of receptive fields in areas AI and AAF of mouse auditory cortex." J Neurophysiol **90**(4): 2660-2675.
67. Liu, B. H., G. K. Wu, R. Arbuckle, H. W. Tao and L. I. Zhang (2007). "Defining cortical frequency tuning with recurrent excitatory circuitry." Nature neuroscience **10**(12): 1594-1600.
68. Lu, T., L. Liang and X. Wang (2001). "Neural representations of temporally asymmetric stimuli in the auditory cortex of awake primates." Journal of neurophysiology **85**(6): 2364-2380.
69. Lu, T., L. Liang and X. Wang (2001). "Temporal and rate representations of time-varying signals in the auditory cortex of awake primates." Nat Neurosci **4**(11): 1131-1138.
70. MacKay, D. (1992). "Information based objective functions for active data collection." Neural Comput.(4): 448-472.
71. Malmberg, C. F. (1918). "The perception of consonance and dissonance." Psychological Monographs **25**(2): 93-133.
72. Markram, H., M. Toledo-Rodriguez, Y. Wang, A. Gupta, G. Silberberg and C. Wu (2004). "Interneurons of the neocortical inhibitory system." Nature reviews.Neuroscience **5**(10): 793-807.
73. Marmarelis, P. Z. M., V. Z. (1978). Analysis of physiological systems: The white-noise approach. New York, Plenum Press.
74. McDermott, J. H., A. J. Lehr and A. J. Oxenham (2010). "Individual differences reveal the basis of consonance." Curr Biol **20**(11): 1035-1041.
75. Meddis, R., L. P. O'Mard and E. A. Lopez-Poveda (2001). "A computational algorithm for computing nonlinear auditory frequency selectivity." J Acoust Soc Am **109**(6): 2852-2861.

76. Merzenich, M. M., P. L. Knight and G. L. Roth (1975). "Representation of cochlea within primary auditory cortex in the cat." J Neurophysiol **38**(2): 231-249.
77. Moore, B. C. (1995). Frequency analysis and masking. Hearing. San Diego, CA, Academic Press: 161-206.
78. Moore, G. A. and B. C. Moore (2003). "Perception of the low pitch of frequency-shifted complexes." The Journal of the Acoustical Society of America **113**(2): 977-985.
79. Moshitch, D., L. Las, N. Ulanovsky, O. Bar-Yosef and I. Nelken (2006). "Responses of neurons in primary auditory cortex (A1) to pure tones in the halothane-anesthetized cat." J Neurophysiol **95**(6): 3756-3769.
80. Nelken, I., Y. Prut, E. Vaadia and M. Abeles (1994). "In search of the best stimulus: an optimization procedure for finding efficient stimuli in the cat auditory cortex." Hear Res **72**(1-2): 237-253.
81. Nelken, I., Y. Prut, E. Vaddia and M. Abeles (1994). "Population responses to multifrequency sounds in the cat auditory cortex: four-tone complexes." Hearing research **72**(1-2): 223-236.
82. O'Connor, K. N., C. I. Petkov and M. L. Sutter (2005). "Adaptive Stimulus Optimization for Auditory Cortical Neurons." Journal of Neurophysiology **94**(6): 4051-4067.
83. Ojima, H. and K. Murakami (2002). "Intracellular Characterization of Suppressive Responses in Supragranular Pyramidal Neurons of Cat Primary Auditory Cortex In Vivo." Cerebral Cortex **12**(10): 1079-1091.
84. Osmanski, M. S., X. Song and X. Wang (2013). "The Role of Harmonic Resolvability in Pitch Perception in a Vocal Nonhuman Primate, the Common Marmoset (*Callithrix jacchus*)." The Journal of neuroscience **33**(21): 9161-9168.
85. Osmanski, M. S. and X. Wang (2011). "Measurement of absolute auditory thresholds in the common marmoset (*Callithrix jacchus*)." Hearing research **277**(1-2): 127-133.

86. Osmanski, M. S. and X. Wang (2011). "Measurement of absolute auditory thresholds in the common marmoset (*Callithrix jacchus*)."
Hear Res **277**(1-2): 127-133.
87. Pasupathy, A. and C. E. Connor (2002). "Population coding of shape in area V4." Nat Neurosci **5**(12): 1332-1338.
88. Patterson, R. D. and F. L. Wightman (1976). "Residue pitch as a function of component spacing." The Journal of the Acoustical Society of America **59**(6): 1450-1459.
89. Pelleg-Toiba, R. and Z. Wollberg (1989). "Tuning properties of auditory cortex cells in the awake squirrel monkey." Experimental brain research. Experimentelle Hirnforschung. Experimentation cerebrale **74**(2): 353-364.
90. Pfingst, B. E. and T. A. O'Connor (1981). "Characteristics of neurons in auditory cortex of monkeys performing a simple auditory task." Journal of neurophysiology **45**(1): 16-34.
91. Prenger, R., M. C. Wu, S. V. David and J. L. Gallant (2004). "Nonlinear V1 responses to natural scenes revealed by neural network analysis." Neural Netw **17**(5-6): 663-679.
92. Ptitsyn, A. A., S. Zvonic and J. M. Gimble (2006). "Permutation test for periodicity in short time series data." BMC bioinformatics **7 Suppl 2**: S10.
93. Qin, L., M. Sakai, S. Chimoto and Y. Sato (2005). "Interaction of excitatory and inhibitory frequency-receptive fields in determining fundamental frequency sensitivity of primary auditory cortex neurons in awake cats." Cerebral cortex (New York, N.Y.: 1991) **15**(9): 1371-1383.
94. Rauschecker, J. P., B. Tian and M. Hauser (1995). "Processing of complex sounds in the macaque nonprimary auditory cortex." Science **268**(5207): 111-114.
95. Recanzone, G. H. (2008). "Representation of con-specific vocalizations in the core and belt areas of the auditory cortex in the alert macaque monkey." J Neurosci **28**(49): 13184-13193.
96. Recanzone, G. H., D. C. Guard and M. L. Phan (2000). "Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey." J Neurophysiol **83**(4): 2315-2331.

97. Reiss, L. A., S. Bandyopadhyay and E. D. Young (2007). "Effects of stimulus spectral contrast on receptive fields of dorsal cochlear nucleus neurons." J Neurophysiol **98**(4): 2133-2143.
98. Rothschild, G., I. Nelken and A. Mizrahi (2010). "Functional organization and population dynamics in the mouse primary auditory cortex." Nature neuroscience **13**(3): 353-360.
99. Sachs, M. B. and E. D. Young (1979). "Encoding of steady-state vowels in the auditory nerve: representation in terms of discharge rate." J Acoust Soc Am **66**(2): 470-479.
100. Sadagopan, S. and X. Wang (2008). "Level invariant representation of sounds by populations of neurons in primary auditory cortex." The Journal of neuroscience **28**(13): 3415-3426.
101. Sadagopan, S. and X. Wang (2009). "Nonlinear spectrotemporal interactions underlying selectivity for complex sounds in auditory cortex." The Journal of neuroscience **29**(36): 11192-11202.
102. Sadagopan, S. and X. Wang (2010). "Contribution of inhibition to stimulus selectivity in primary auditory cortex of awake primates." The Journal of neuroscience **30**(21): 7314-7325.
103. Schinkel-Bielefeld, N., S. V. David, S. A. Shamma and D. A. Butts (2012). "Inferring the role of inhibition in auditory processing of complex natural stimuli." J Neurophysiol **107**(12): 3296-3307.
104. Schouten, J. F. (1938). "The perception of subjective tones." Proceedings of the Koninklijke Nederlandse Akademie Van Wetenschappen **41**(6/10): 1086-1093.
105. Schroeder, C. E., A. D. Mehta and S. J. Givre (1998). "A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque." Cereb Cortex **8**(7): 575-592.
106. Schwartz, D. A., C. Q. Howe and D. Purves (2003). "The statistical structure of human speech sounds predicts musical universals." J Neurosci **23**(18): 7160-7168.

107. Schwarz, D. W. and R. W. Tomlinson (1990). "Spectral response patterns of auditory cortex neurons to harmonic complex tones in alert monkey (*Macaca mulatta*)."
Journal of neurophysiology **64**(1): 282-298.
108. Shamma, S. and D. Klein (2000). "The case of the missing pitch templates: how harmonic templates emerge in the early auditory system."
J Acoust Soc Am **107**(5 Pt 1): 2631-2644.
109. Shamma, S. A. and D. Symmes (1985). "Patterns of inhibition in auditory cortical cells in awake squirrel monkeys."
Hearing research **19**(1): 1-13.
110. Sheiner, L. B. and S. L. Beal (1981). "Some suggestions for measuring predictive performance."
J Pharmacokinet Biopharm **9**(4): 503-512.
111. Shen, J. X., A. S. Feng, Z. M. Xu, Z. L. Yu, V. S. Arch, X. J. Yu and P. M. Narins (2008). "Ultrasonic frogs show hyperacute phonotaxis to female courtship calls."
Nature **453**(7197): 914-916.
112. Smoorenburg, G. F. and D. H. Linschoten (1977). A neurophysiological study on auditory frequency analysis of complex tones. Psychophysics and Physiology in Hearing. E. F. Evans and J. P. Wilson. London, Academic Press: 175-184.
113. Suga, N. (1965). "Functional properties of auditory neurones in the cortex of echolocating bats."
The Journal of physiology **181**(4): 671-700.
114. Suga, N., W. E. O'Neill and T. Manabe (1979). "Harmonic-sensitive neurons in the auditory cortex of the mustache bat."
Science **203**(4377): 270-274.
115. Sutter, M. L. and W. C. Loftus (2003). "Excitatory and Inhibitory Intensity Tuning in Auditory Cortex: Evidence for Multiple Inhibitory Mechanisms."
Journal of Neurophysiology **90**(4): 2629-2647.
116. Sutter, M. L. and C. E. Schreiner (1991). "Physiology and topography of neurons with multip peaked tuning curves in cat primary auditory cortex."
Journal of neurophysiology **65**(5): 1207-1226.

117. Sutter, M. L., C. E. Schreiner, M. McLean, N. O'Connor K and W. C. Loftus (1999). "Organization of inhibitory frequency receptive fields in cat primary auditory cortex." J Neurophysiol **82**(5): 2358-2371.
118. Sutter, M. L., C. E. Schreiner, M. McLean, K. N. O'Connor and W. C. Loftus (1999). "Organization of inhibitory frequency receptive fields in cat primary auditory cortex." Journal of neurophysiology **82**(5): 2358-2371.
119. Tam, W. (2012). Adaptive modeling of marmoset inferior colliculus neurons in vivo. Ph.D., Johns Hopkins University.
120. Tan, A. Y., L. I. Zhang, M. M. Merzenich and C. E. Schreiner (2004). "Tone-evoked excitatory and inhibitory synaptic conductances of primary auditory cortex neurons." J Neurophysiol **92**(1): 630-643.
121. Tan, A. Y. Y., C. A. Atencio, D. B. Polley, M. M. Merzenich and C. E. Schreiner (2007). "Unbalanced synaptic inhibition can create intensity-tuned auditory cortex neurons." Neuroscience **146**(1): 449-462.
122. Terhardt, E. (1974). "Pitch, consonance, and harmony." J Acoust Soc Am **55**(5): 1061-1069.
123. Theunissen, F. E., K. Sen and A. J. Doupe (2000). "Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds." J Neurosci **20**(6): 2315-2331.
124. Tian, B., D. Reser, A. Durham, A. Kustov and J. P. Rauschecker (2001). "Functional specialization in rhesus monkey auditory cortex." Science **292**(5515): 290-293.
125. Tomlinson, R. W. and D. W. Schwarz (1988). "Perception of the missing fundamental in nonhuman primates." The Journal of the Acoustical Society of America **84**(2): 560-565.
126. Tramo, M. J., P. A. Cariani, B. Delgutte and L. D. Braida (2001). "Neurobiological foundations for the theory of harmony in western tonal music." Ann N Y Acad Sci **930**: 92-116.

127. Walker, K. M., J. W. Schnupp, S. M. Hart-Schnupp, A. J. King and J. K. Bizley (2009). "Pitch discrimination by ferrets for simple and complex sounds." The Journal of the Acoustical Society of America **126**(3): 1321-1335.
128. Wang, X., T. Lu, R. K. Snider and L. Liang (2005). "Sustained firing in auditory cortex evoked by preferred stimuli." Nature **435**(7040): 341-346.
129. Watson, A. B. and D. G. Pelli (1983). "QUEST: a Bayesian adaptive psychometric method." Percept Psychophys **33**(2): 113-120.
130. Wehr, M. and A. M. Zador (2003). "Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex." Nature **426**(6965): 442-446.
131. Whitfield, I. C. (1980). "Auditory cortex and the pitch of complex tones." The Journal of the Acoustical Society of America **67**(2): 644-647.
132. Wilson, N. R., C. A. Runyan, F. L. Wang and M. Sur (2012). "Division and subtraction by distinct cortical inhibitory networks in vivo." Nature **488**(7411): 343-348.
133. Wu, G. K., R. Arbuckle, B. H. Liu, H. W. Tao and L. I. Zhang (2008). "Lateral sharpening of cortical frequency tuning by approximately balanced inhibition." Neuron **58**(1): 132-143.
134. Wu, M. C., S. V. David and J. L. Gallant (2006). "Complete functional characterization of sensory neurons by system identification." Annu Rev Neurosci **29**: 477-505.
135. Young, E. D. and M. B. Sachs (1979). "Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers." J Acoust Soc Am **66**(5): 1381-1403.
136. Young, E. D., J. J. Yu and L. A. Reiss (2005). "Non-linearities and the representation of auditory spectra." Int Rev Neurobiol **70**: 135-168.
137. Yu, J. J. and E. D. Young (2000). "Linear and nonlinear pathways of spectral information transmission in the cochlear nucleus." Proc Natl Acad Sci U S A **97**(22): 11780-11786.

138. Zhang, L. I., A. Y. Tan, C. E. Schreiner and M. M. Merzenich (2003). "Topography and synaptic shaping of direction selectivity in primary auditory cortex." Nature **424**(6945): 201-205.
139. Zhang, X., M. G. Heinz, I. C. Bruce and L. H. Carney (2001). "A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression." J Acoust Soc Am **109**(2): 648-670.

Curriculum Vitae

Lei Feng

Johns Hopkins University
Dept. of Biomedical Engineering
412 Traylor, 702 Rutland Avenue,
fenglei.thu@gmail.com
Baltimore, Maryland, 21205

Telephone:
(410)502-6019

Email:

Birth

July 24, 1983, Nanchong, Sichuan, China

EDUCATION

Johns Hopkins University, Baltimore, Maryland	Ph.D.	Biomedical Engineering (2013)
Tsinghua University, Beijing, China	M.S.	Biomedical Engineering (2007)
Tsinghua University, Beijing, China	B.S.	Biomedical Engineering (2005)

RESEARCH EXPERIENCE

Laboratory of Auditory Neurophysiology

Johns Hopkins University (2008-2013), Ph.D. Student, Advisor: Dr. Xiaoqin Wang

Single-unit electrophysiology studying spectral integration in the auditory cortex of marmosets.

Institute of Neural Engineering

Tsinghua University (2005-2007), Master Student, Advisor: Dr. Bo Hong and Dr. Shangkai Gao

Single-unit electrophysiology studying stimulus specific adaptation in the inferior colliculus of anesthetized rats.

HONORS/AWARDS

Guanghua Scholarship (2006)

Dekai Wu Couple Scholarship (2002)

Tsinghua University Scholarship for Excellent Freshmen (2001)

PUBLICATIONS

Journal Articles

1. **Feng, L.** and Wang, X., (In preparation) Integrative Representation of Harmonic Structure by Combination Sensitive Neurons in Primary Auditory Cortex of Awake Marmoset.
2. **Feng, L.** and Wang, X., (In preparation) A distributed harmonic processing in primary auditory cortex.
3. **Feng, L.**, Dekel, E., Zhang, K., and Wang, X., (In preparation) An online adaptive stimulus design for studying spectral integration of single neurons in auditory cortex.
4. Dekel, E., **Feng, L.**, Wang, X. and Zhang, K., (to be submitted) Excitation-Inhibition “Co-Tuning” in Primary Auditory Cortex: Evidence from Experiment-driven Neural Modeling in Awake Primates.
5. Zhao, L., Liu, Y., Shen, L., **Feng, L.** and Hong, B., (2011). Stimulus-specific Adaptation and its Dynamics in the Inferior Colliculus of Rat. *Neuroscience*. 181(5), 163-174.

Conference Proceedings

1. Liu, Z., Huang, P., Wang, Z., **Feng, L.**, Zhang, C., Li, D., Gao, S., Song, X. and Ren, T., (2005). Full Custom Design of the Interface for an in vitro Neural Biosensor. *Proceedings of the 48th IEEE International Midwest Symposium on Circuits and Systems*, 2, 1306-1309.

Conference Abstracts

1. **Feng, L.**, Dekel, E., Zhang, K. and Wang, X., Online Adaptive Stimulus Design for Studying Non-linear Spectral Integration in Auditory Cortex, *Society for Neuroscience Abs*, New Orleans, LA, October 2012.
2. **Feng, L.** and Wang, X., Neural Coding of Harmonic Complex Tones in Auditory Cortex, *Gordon Research Seminar and Gordon Research Conference*, Lewiston, MA, July 2012 (oral presentation at GRS).
3. **Feng, L.** and Wang, X., Neural Coding of Harmonic Complex Tones in Auditory Cortex of Awake Marmoset, *Society for Neuroscience Abs*, Washington D.C., November 2011.

4. Dekel, E., **Feng, L.**, Wang, X. and Zhang, K., Modeling of Harmonic Sensitive Neurons in the Primary Auditory Cortex of Marmoset Monkeys, *Society for Neuroscience Abs*, Washington D.C., November 2011.

5. **Feng, L.** and Wang, X., The Role of Interactions between Excitatory and Inhibitory Receptive Field Components in Encoding Harmonic Structures in Auditory Cortex of Awake Marmosets, *34th Midwinter Meeting of the Association for Research in Otolaryngology (ARO)*. Baltimore, MD, February 2011.

6. **Feng, L.** and Wang, X., The Role of Interactions between Excitatory and Inhibitory Receptive Field Components in Encoding Harmonic Structures in Auditory Cortex of Awake Marmosets, *Society for Neuroscience Abs*, San Diego, CA, November 2010.

TEACHING EXPERIENCE

- Teaching assistant, System Bioengineering II at Johns Hopkins University. Fall 2010
- Teaching assistant, System Bioengineering II Lab at Johns Hopkins University. Fall 2011

PROFESSIONAL ASSOCIATIONS

- Society of Neuroscience (SfN)
- Association for Research in Otolaryngology (ARO)