Washington University in St. Louis
Washington University Open Scholarship

All Theses and Dissertations (ETDs)

1-1-2009

Computational Analysis of Functional Imaging in the Primary Auditory Cortex

Thomas Chen
Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/etd

Recommended Citation
Chen, Thomas, "Computational Analysis of Functional Imaging in the Primary Auditory Cortex" (2009). *All Theses and Dissertations (ETDs)*. 524.
https://openscholarship.wustl.edu/etd/524

This Thesis is brought to you for free and open access by Washington University Open Scholarship. It has been accepted for inclusion in All Theses and Dissertations (ETDs) by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

WASHINGTON UNIVERSITY IN ST. LOUIS
School of Engineering and Applied Science
Department of Biomedical Engineering

Thesis Examination Committee:
Dennis Barbour, Chair
Andreas Burkhalter
David Van Essen

Computational Analysis of Functional Imaging in the Primary Auditory Cortex

by
Thomas Chen

A thesis presented to the School of Engineering of Washington University in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August 2009
Saint Louis, Missouri

ABSTRACT OF THE THESIS

Computational Analysis of Functional Imaging in the Primary Auditory Cortex

by

Thomas Chen

Master of Science in Computer Science

Washington University in St. Louis, 2009

Research Advisor: Dennis Barbour

Functional imaging can reveal detailed organizational structure in cerebral cortical areas, but neuronal response features and local neural interconnectivity can influence the resulting images, possibly limiting the inferences that can be drawn about neural function. Historically, discerning the fundamental principles of organizational structure in the auditory cortex of multiple species has been somewhat challenging with functional imaging as the studies have failed to reproduce results seen in electrophysiology. One difference might result from the way most functional imaging studies record the summed activity of multiple neurons. To test this effect, virtual mapping experiments were run in order to gauge the ability of functional imaging to accurately estimate underlying maps. The experiments suggest that spatial averaging improves the ability to estimate maps with low spatial frequencies or with large amounts of cortical variability, at the cost of decreasing the spatial resolution of the images. Despite the decrease in resolution, the results suggest that current functional imaging studies may be able to depict maps with high spatial frequencies better than electrophysiology can; therefore, the difficulties in recapitulating electrophysiology experiments with imaging may stem from underlying neural circuitry. One possible reason may be the relative distribution of response selectivity throughout the population of auditory cortex neurons. A small percent of neurons have a response type that exhibits a receptive field size that increases with higher stimulus intensities, but they are likely to contribute disproportionately to the activity

detected in functional images, especially if intense sounds are used for stimulation. To evaluate the potential influence of neuronal subpopulations upon functional images of the primary auditory cortex, a model array representing cortical neurons was probed with virtual imaging experiments under various assumptions about the local circuit organization. As expected, different neuronal subpopulations were activated preferentially under different stimulus conditions. In fact, stimulus protocols that can preferentially excite one subpopulation of neurons over the others have the potential to improve the effective resolution of functional auditory cortical images. These experimental results also make predictions about auditory cortex organization that can be tested with refined functional imaging experiments.

Acknowledgments

My thesis would not have been possible without the help of many people. First and foremost, I would like to thank Dennis Barbour, my research advisor. None of this would have been possible if he had not given me the opportunity to join his lab and spent the time to introduce me to the wonders of research. Next, I would like to thank Paul Watkins for his work on the self-organizing feature map and his help with incorporating these maps with my model. I would also like to thank Scott Burns, Kim Kocher, Woosung Kim, Kristy Ratkowski, and the rest of the Barbour lab for all their help in editing this thesis and supporting me during my time in the lab. Lastly, I would like to thank my parents, Yen and Susan Chen, and my siblings, Litton and Tiffany. Without their understanding and support throughout my life, I would not be where I am today.

Thomas Chen

Washington University in St. Louis

August 2009

Table of Contents

Abstract ..... **ii**
Acknowledgments ..... **iv**
Table of Contents ..... **v**
Chapter 1 Introduction **1**
Chapter 2 Experimental Methods **4**
2.1 Self-Organizing Feature Maps 4
2.2 Virtual Imaging Model 5
2.2.1 SOFM Incorporation 5
2.2.2 Map Estimation 7
2.2.3 Map Fit Quantification 9
2.3 Virtual Imaging Model 11
2.3.1 Frequency Response Area Modeling 11
2.3.2 Primary Auditory Cortex Modeling 15
2.3.3 Functional Map Extraction 19
2.3.4 Complex Stimuli 20
Chapter 3 Virtual Mapping **21**
3.1 Results 21
3.1.1 Linear Sampling Density of Estimation 21
3.1.2 Averaged vs. Unaveraged Estimations 24
3.1.3 Extractions with Random Sampling 26
3.1.4 Interpolation 28
3.2 Discussion 28
3.2.1 Sampling Density 28
3.2.2 Averaging 31
3.2.3 Sampling Locations 32
3.2.4 Cortical Variability 32
3.2.4 Display of Estimated Maps 33

3.2.6 Relation to Previous Studies 33
Chapter 4 Virtual Imaging of the Primary Auditory Cortex 35
4.1 Results 35
4.1.1 Model Array 35
4.1.2 Neuronal Subpopulation Responses 40
4.1.3 Functional Map Extraction 48
4.1.4 Responses to Complex Stimuli 50
4.2 Discussion 55
4.2.1 Functional Map Extractions 57
4.2.2 Responses to Complex Stimuli 59
4.2.3 Experimental Techniques Predicted to Improve Imaging Resolution 60
4.2.4 Limitations in Virtual Imaging 60
Chapter 5 Conclusion 63
5.1 Summary of Experimental Results 63
5.2 Functional Imaging as a Neurophysiology Tool 64
5.3 Shortcomings of Functional Imaging 64
5.4 Possible Future Studies 65
References 66
Vita 73

List of Figures

Figure 2.1. Four-feature SOFM. 6
Figure 2.2. Resampled Map. 8
Figure 2.3. Sampling rates affect on R^2 values. 10
Figure 2.4. Self-organization of primary auditory cortex functional properties. 17
Figure 3.1. Examples of mapping estimations with different sampling densities. 22
Figure 3.2. Comparison between averaged and unaveraged estimations. 23
Figure 3.3. Statistical comparison between unaveraged and averaged estimations. 25
Figure 3.4. Comparison between uniform and random sampling. 27
Figure 3.5. Comparisons between estimations using zeroth and linear interpolations. 30
Figure 4.1. Single unit spiking activity reveals 3 response types. 36
Figure 4.2. Sample intrinsic optical imaging experiment from cat A1. 39
Figure 4.3. Virtual imaging of different response types. 41
Figure 4.4. Virtual imaging with a variety of weightings between response types. 43
Figure 4.5. Areas of activation with varying stimulus amplitude 46
Figure 4.6. Virtual imaging experiment with different amounts of cortical variability. .. 47
Figure 4.7. Extracted feature maps. 49
Figure 4.8. Virtual imaging of complex stimuli. 52
Figure 4.9. Differential maps of activity from complex stimuli. 54

Chapter 1 Introduction

Functional neuroimaging has become a powerful tool for evaluating the physiological characteristics of large neuronal populations at high spatial resolution. Functional maps of neocortical sensory areas in particular can often elucidate overall neuronal organization more clearly than even dense electrophysiological mapping studies. Modern functional brain imaging technologies include optical imaging of intrinsic signals (OIS) and functional magnetic resonance imaging of blood-oxygen-level-dependent signal (fMRI-BOLD), both of which exhibit high enough spatial resolution to reveal functional maps within individual cortical areas. Despite having high spatial resolution, however, both techniques still measure summed neuronal activity rather than the responses of individual neurons (Logothetis 2008; Logothetis et al. 2001). The neural activity summed within an individual imaged pixel/voxel is influenced by the correlated activity within the underlying neural circuitry. Presumably, the nature of this neural circuitry can directly influence functional imaging results.

Auditory cortex has typically been challenging to study with functional imaging techniques, particularly for functional organization beyond acoustic frequency. Several factors may have contributed to these difficulties, including the physiological properties of auditory neurons themselves. Pure tone acoustic stimulation in early OIS studies of cat primary auditory cortex (A1) revealed stimulus-driven activity with nearly circular areas of activation that shifted across the cortical surface with changes in tone frequency (Harrison et al. 1998; Spitzer et al. 2001). Electrophysiological mapping studies, however, have demonstrated that isofrequency regions in cat A1 are long and band-like instead of circular (Cheung et al. 2001; Merzenich et al. 1973, 1975; Schreiner and Mendelson 1990). More recent OIS experiments have revealed band-like isofrequency structure and a clear tonotopic organization (Ojima et al. 2005), consistent with the findings from electrophysiology and implying that modern techniques can more closely recapitulate classical electrophysiological findings in A1. Recent FMRI-BOLD experiments in macaque monkey A1 have been able to define tonotopic frequency maps

(Petkov et al. 2006), as well as crude bandwidth maps (Kayser et al. 2007). OIS activity in A1 has been reported to increase monotonically with stimulus intensity (Harrison et al. 1998; Ojima et al. 2005; Sheth et al. 2003; Sheth et al. 2004). Electrophysiology studies have shown that these responses are representative of a subpopulation of A1 neurons whose rate responses increase monotonically with increasing intensity (Pfingst and O'Connor 1981; Sadagopan and Wang 2008). Sophisticated imaging methods therefore appear to be able to discern some features of A1 that are also observed with electrophysiology. These methods have the potential to extend the understanding of A1 functional organization beyond what is possible with electrode mappings that are sparse relative to functional imaging studies. The diversity of A1 receptive field structures and local interconnections may, however, fundamentally limit the inferences that can be drawn from such images.

In this thesis, functional imaging was evaluated as a tool to investigate the primary auditory cortex, and possible limitations that cause the discrepancy between electrophysiology and functional imaging studies were identified. To do this, (1) the ability to accurately estimate A1 maps with electrophysiology and functional imaging was compared, and (2) possible limitations in functional imaging of A1 that result from the neural circuitry were identified. The first objective aims to examine fundamental limitations within the imaging due to technological limitations while the latter objective aims to examine the shortcomings of imaging due to the neural circuitry.

Self-organizing feature maps (SOFMs) have been known to create plausible topographic structures for functional A1 and primary visual cortex maps based upon theoretical arguments (Obermayer et al. 1992; Watkins et al. 2009). For studying the limitations of imaging in mapping, maps were computationally sampled to determine how different mapping features affect the overall accuracy of the reconstructed map. Three different types of mapping characteristics were focused on: sampling density, sampling locations, and single- versus multi-unit mapping.

Examples of these maps were used to construct an array of modeled A1 neurons, which were then probed in virtual imaging experiments to determine how response nonlinearities peculiar to the auditory system may affect the results of functional imaging studies. In particular, some A1 neurons inherit from the cochlea a broadening of frequency sensitivity at increasing sound intensities. This broadening of frequency sensitivity is explored, and the relatively low stimulus selectivity that accompanies this property, may hinder the ability of imaging experiments to detect the underlying physiological activity in A1. Under many circumstances tested, these broadly tuned neurons dominated the imaged responses and led to results similar to those that have been observed physiologically with functional imaging. Different stimulus conditions affected the influence of the each neuronal response in the images, indicating that stimulus design can affect imaging results based upon the neuron population active at a particular time. Strategies to improve functional maps of A1 and extend current imaging results are evaluated.

Chapter 2 Experimental Methods

2.1 Self-Organizing Feature Maps

The self-organizing feature map (SOFM) is a dimensionality-reduction algorithm that projects *n* feature dimensions—in this case physiological features—onto the two anatomical dimensions of the cortical surface. The SOFM model is based on the wiring-minimization principle where neurons that are connected are more likely to be located near each other, and the model assumes that neurons with similar characteristics are more likely to be connected with each other than neurons with less similar characteristics. Using this principle and assumption, the SOFM uses a competitive learning algorithm that incorporates Hebbian learning rates to generate spatial arrangements of neuronal properties well-matched to topographies observed in functional imaging studies of primary visual cortex (V1) (Farley et al. 2007; Obermayer and Blasdel 1993; Obermayer et al. 1992; Yu et al. 2005). The properties of SOFMs relevant to functional topographies in A1 are evaluated more thoroughly in (Watkins et al. 2009). In chapter 3, the SOFM models were used to create possible neuronal maps with varying degrees of spatial frequencies. In chapter 4, the SOFM inputs were adapted to create reasonable models of A1 topographies for neuronal characteristic frequency (CF), bandwidth and threshold. These features are sufficient to determine receptive field shapes and have been determined to be non-randomly distributed in A1 (Cheung et al. 2001; Philibert et al. 2005; Recanzone et al. 1999; Schreiner and Mendelson 1990; Schreiner et al. 1992; Schreiner and Sutter 1992). Higher weightings of a feature lead to greater preservation of mapping uniformity and compactness of that feature. Higher weightings of a particular feature lead to greater preservation of mapping uniformity and compactness of that feature (Watkins et al. 2009). This report uses 3-feature SOFMs with a 10:2:1 weightings and 4-feature SOFMs with a 10:4:2:1 weighting. For the 10:4:2:1 weighting, a uniform distribution of feature 1 values is emphasized 2.5 times more in the SOFM algorithm than the feature 2, five times more than feature 3, and ten times more than feature 4.

2.2 Virtual Imaging Model

2.2.1 SOFM Incorporation

Cortical maps are modeled as an array with each value in the array representing the position and the neuronal properties of each unit in the map. These units are defined from the underlying SOFM. All arrays in this study are square with 150 × 150 units. A 4-feature SOFM with a relative feature weighting of 10:4:2:1, which can be seen in Figure 2.1, is selected for analyzing the estimation of cortical maps. These weightings were chosen to create a variety of order in the maps. In comparison to the other feature maps, the first feature map has relatively low spatial frequencies, and the fourth feature map exhibits high spatial frequency. In this thesis, spatial frequency refers to number of periods or cycles of the extreme values across the space of the map. The first feature map shows a half cycle, the second feature map shows three quarters of a cycle, and the third shows one and a half cycles. The fourth feature shows a numerous number of cycles.

While neurons are organized within a cortical column, the properties of neurons isolated within a cortical column can vary (Atencio and Schreiner 2008; Phillips and Irvine 1981). To model this cortical variability, the SOFM represents the arithmetic mean of a spatial probability distribution instead of the actual values. The particular value of any given unit for any given map was drawn from this distribution. To quantify cortical variability, a multivariate Gaussian distribution with unique standard deviations for each feature (defined in this model as jitter, which is defined as a percentage of the total range of values for that feature) was used. Larger jitter simulates maps in cortical areas with larger variability in the columnar structure. A jitter of 0 represents a deterministic mapping with no variability in columnar structure. Cortical columns were created by layering the arrays. Each layer contained values that were jittered independently from the mean, but the locations of each unit and the underlying SOFM were the same between layers.

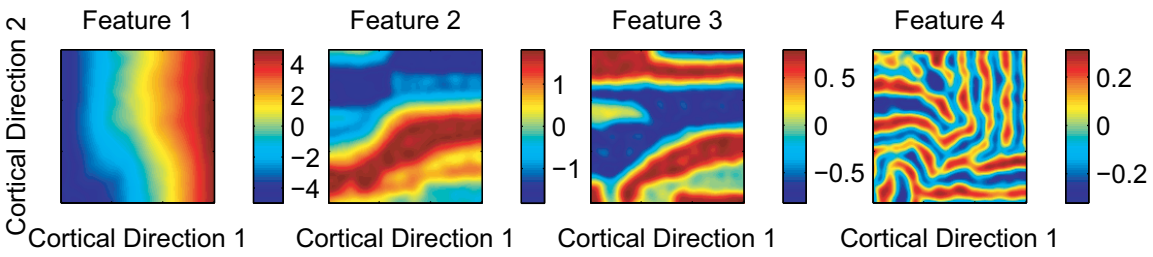

Figure 2.1. Four-feature SOFM.

Self-organization feature map shows different levels of spatial frequency. In this model, four different neuronal features are mapped with a 10:4:2:1 weighting. Maps with higher weightings are more highly organized and uniform than maps with lower weightings. These maps were used as the base maps for the entire study.

2.2.2 Map Estimation

To determine experimental and analytical practices useful for elucidating map structure, the SOFMs are created at the full 150 × 150 resolution and are then probed at lower sampling densities. These map estimations are created with either an averaged or an unaveraged procedure. The unaveraged estimations used sparse samples of actual unit values for a map. This condition was intended to simulate a sparse, single-unit sampling experiment. The averaged estimations are the computed mean feature values in a region around each of the sparsely sampled sites. This condition was intended to simulate a multiunit, local field potential, electrocorticographic or functional imaging experiment where the average activity of a local group of neurons is recorded in each measurement.

Sampling locations in the unaveraged experiments are spaced either uniformly across the array or randomly sampled without replacement from the jittered maps. For the averaged estimations, sampling locations are all uniformly spaced across the array. The number of units or pixels/voxels that were used to estimate the map is represented by the linear sampling density. The linear sampling density is the square root of the total number of total units or pixels/voxels used in the estimation. For example, if 25 units (or a 5 × 5 unit map) are used to estimate the SOFM, then the linear sampling density would be five.

To simulate the contribution of different cortical layers to the recording in the averaged layers, ten independent instantiations of jittered SOFM maps were superimposed upon each other before averaging. For example, if a pixel/voxel corresponded to a 2 × 2 unit square on a map, that pixel/voxel would be assigned the arithmetic mean of 40 different units. Once a sparse map was assembled, it was resampled back to 150 × 150 units using zeroth-order (nearest-neighbor) or first-order (linear) interpolation. The overall process with the linear interpolation is summarized in Figure 2.2. For the estimated maps that were resampled with linear interpolations, units outside the border created from the extracted points are not extrapolated or used for later analysis.

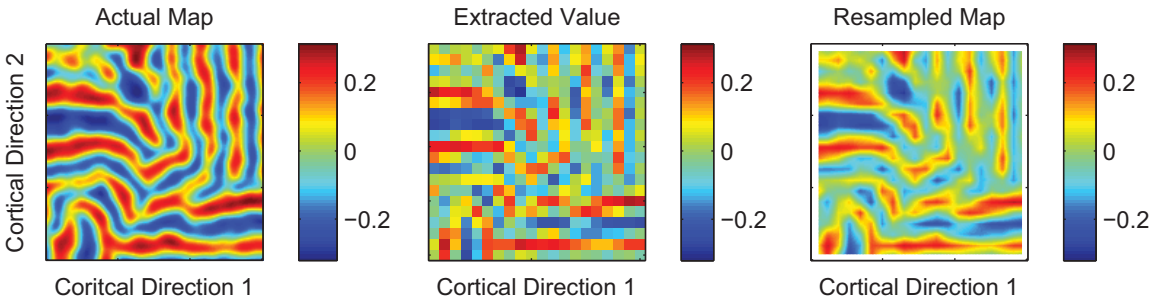

Figure 2.2. Resampled Map.

A unjittered map of feature 4 is estimated with an averaged uniform sampling and a linear sampling density of 20 units. The extracted values (center) shows the map before it was resampled back to 150 × 150 units using linear interpolation to form the estimated map (right). Points shown in white were not interpolated as they were outside the boundary of the sampled locations.

Estimated maps are created with a variety of linear sampling densities and jitters with twenty-four different linear sampling densities, ranging from 2 to 25 points, and seven different jitter percentages, ranging from 0% jitter to 30% jitter. For each combination of linear sampling density and jitter, the estimations are repeated a hundred times, and the coefficient of determination (R^2) is calculated in each set.

2.2.3 Map Fit Quantification

The accuracy of the fits for the estimated maps is quantified using the coefficient of determination (R^2). In this metric, the coefficient of determination measures how the estimated maps fit the underlying SOFM maps in relation to the mean of the SOFM map and is derived from the ratio between the sum of square error and the total sum of squares:

$$R^2 = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \bar{y})^2}, \tag{2.1}$$

where y represents the SOFM values prior to jittering, f represents the interpolated values, and i represents the set of interpolated values and their associated non-jittered SOFM values. A R^2 value of one represents perfect fit with no error while a R^2 value of zero represents a fit that has an equal amount of error as a map with the mean value of the SOFM. While R^2 values typically range from zero to one, R^2 values can be negative when the maps are undersampled. The behavior of a negative R^2 value does not reflect whether the fit of the map is improving or worsening. Figure 2.3 shows the R^2 values of a 10 Hz sine wave sampled at different densities. As the sampling rate decreases, the R^2 values drop, and after the sampling rate drops below the Nyquist frequency, the R^2 values drop into negative values and starts to fluctuate. The lower bound of the R^2 values depends on the sampling method used. For maps created with the averaged sampling, the lower bound is at or near zero because averaged maps with low sampling densities would have similar values to the mean of the underlying SOFM. For the maps created with unaveraged estimates, the R^2 values do not have a lower bound.

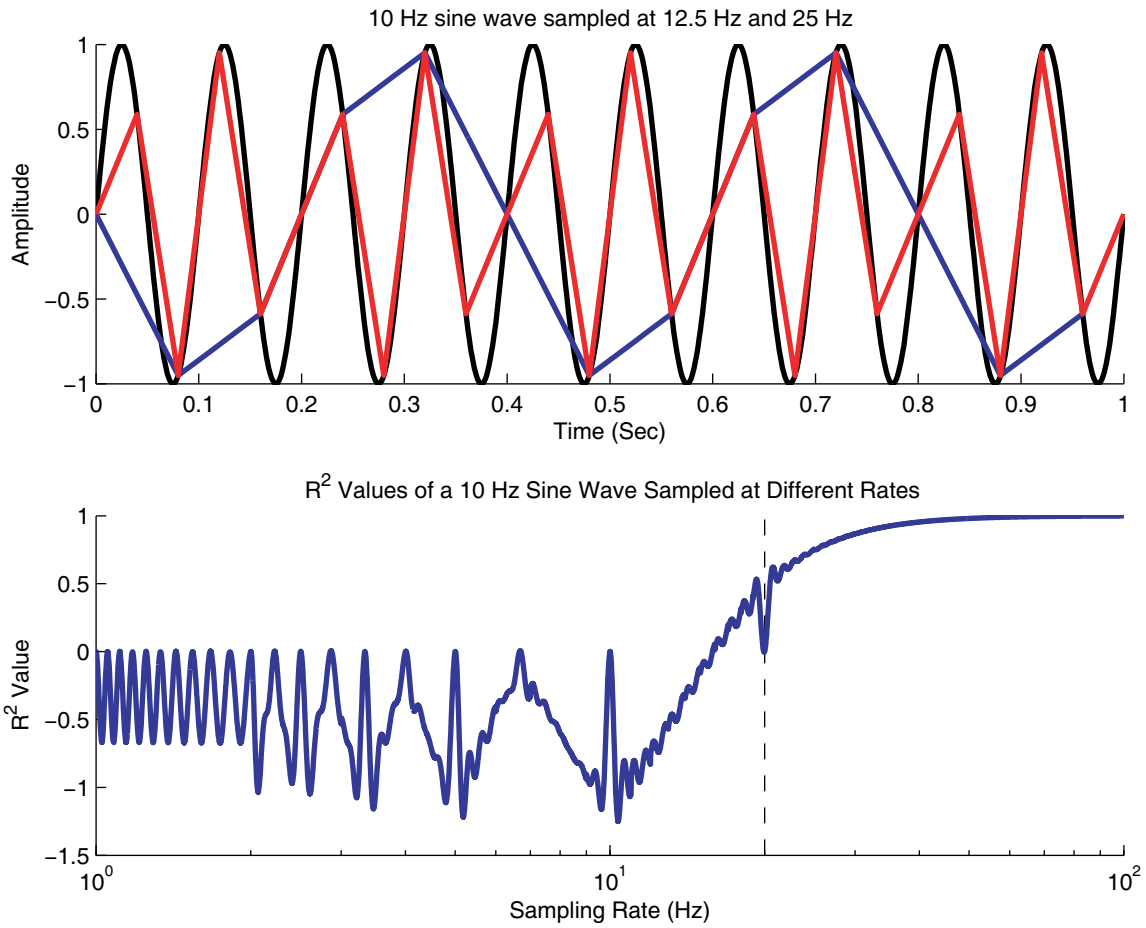

Figure 2.3. Sampling rates affect on R^2 values.

The R^2 values were found for a 10 Hz sine wave that sampled at a various sampling rates. The extracted sine waves were resampled back to 1 kHz using linear interpolation before the R^2 values were calculated. **Top:** A 12.5 Hz sampling rate of the 10 Hz sine wave produces a R^2 value of -0.727 while a 25 Hz sampling rate produces a R^2 value of 0.729. **Bottom:** After the sampling drops below the Nyquist rate (marked with a dotted black line), the sine wave is undersampled. Negative R^2 values are an indicator that the sine wave is undersampled.

To examine the effect of cortical variability, sampling density, and sampling methods on the ability to estimate the underlying map, the mean and the variance of the R^2 are calculated for each distribution set. To determine if any method produced significantly more accurate interpretations of the underlying SOFM, a Wilcoxon rank-sum text is used. To determine if any method produced a more precise estimation, the variance of the R^2 values is tested with a two-sampled F-test.

2.3 Virtual Imaging Model

2.3.1 Frequency Response Area Modeling

Neuronal responses to tone stimuli can be characterized into three major categories that likely represent points along a continuum of frequency selectivity: Type V, Type I, and Type O neurons (Ramachandran et al. 1999; Sadagopan and Wang 2008). Type V and Type I responses are monotonic functions of intensity, meaning their firing rates increase or saturate in response to increasing sound intensity. Type O responses are nonmonotonic functions of intensity and are therefore tuned to a specific sound intensity. Type V responses exhibit wideband receptive fields with increasing bandwidth as stimulus intensity increases, while the Type O and Type I responses exhibit relatively narrowband receptive fields with nearly constant bandwidth constant across their full dynamic range. Type O receptive fields in A1 have been seen with a variety of circular and oval shapes (Sadagopan and Wang 2008).

While spectral responses in A1 as a function of intensity can reveal a great deal of variety (Sutter 2000; Sutter and Schreiner 1991), the model initially focused on the stereotypical responses observed around the major frequency input for any given neuron. Frequency response areas (FRAs) of the three canonical response types described above represent receptive field characteristics as a function of frequency and intensity and were modeled in the current study as combinations of Gaussian functions. Cross-sections of FRAs at fixed intensities demonstrate neuronal tuning properties and are referred to as frequency responses curves (FRCs). These curves were created with a Gaussian density function for

all FRA types, and the increasing bandwidth of the Type V responses were modeled by increasing the standard deviation of the Gaussian as the stimulus intensity increased:

$$FRC(F) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(F-\mu_f)^2/2\sigma_f^2}, \tag{2.2}$$

where μ_f is the mean frequency of the FRC, σ_f is the standard deviation, F represents frequency, and FRC represents the frequency response curve. For Type I and Type O responses, σ_f is a constant while for Type V responses, $\sigma_f(I)$ increases proportionally with intensity.

Intensity response curves (IRCs) were created with a Gaussian density function for Type O responses and a cumulative Gaussian distribution function for Type V and Type I responses:

$$IRC_O(I) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(I-\mu_i)^2/2\sigma_i^2} \text{ and } IRC_{V/I}(I) = \frac{1}{2} erfc\left(\frac{-(I-\mu_i)}{\sigma_i\sqrt{2}}\right), \tag{2.3}$$

where μ_i represents the mean intensity of the IRC, σ_i represents the standard deviation, I represents the stimulus intensity and IRC represents the intensity response curve for either Type O responses or Type V / I responses. The complementary error function is given by

$$erfc(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt. \tag{2.4}$$

For Type I and Type O responses, the outer product of the appropriate IRC and FRC was computed to obtain a template FRA. For Type V responses a unique FRC was computed at each intensity and scaled by the appropriate value of the IRC. The resulting FRA was normalized to take on rate values between 0 and 100 in order to normalize the overall receptive field activity for each unit regardless of its response subtype. The FRA

parameters for any given unit in the model array (characteristic frequency or CF, threshold, frequency bandwidth, and dynamic range) were set to desired values by scaling and shifting the FRCs and IRCs appropriately. The frequency was scaled to define the frequency bandwidth and shifted to set the CF to its desired value:

$$F = f_c \cdot 2^{F_0 \frac{b}{b_0}}, \tag{2.5}$$

where F and F_0 denote new and initial frequencies, respectively, f_c represents the characteristic frequency, and b and b_0 refer to the new and initial bandwidths. The intensity was shifted and scaled differently between the monotonic and the nonmonotonic responses. For the monotonic responses, the intensity was scaled to define the responsive range of the rate-intensity response, then shifted to set the threshold and saturation point to the desired values:

$$I = (I_0 - s_0)\frac{(s-t)}{(s_0 - t_0)} + s, \tag{2.6}$$

where I and I_0 represent the new and the initial intensities, respectively, s and s_0 represent the new and initial intensity maxima, and t and t_0 represents the new and initial intensity thresholds. The responsive range of a unit was considered to extend between its threshold and intensity maximum (i.e., best intensity). For Type I and Type V responses, the response at the point of saturation (threshold + dynamic range at the CF) was then obtained and was set to the maximum response:

$$R = \frac{0.9(R_0 - 0.1)}{(R_S - 0.1)} + 0.1, \tag{2.7}$$

where R and R_0 represent the new and initial response rates, respectively, and R_s is the response at the point of saturation. Any responses greater than 100 where set to 100 after scaling.

For Type O responses, the intensity was scaled to set the dynamic range appropriately and shifted to set the threshold and peak intensity:

$$I = (I_0 - t_0)\frac{a}{2a_0} + t + a, \tag{2.8}$$

where I and I_0 represent the new and the initial intensities, respectively, t represents threshold, and a represents the dynamic range. For the model, the maximum modeled intensity for Type V and Type I responses was set to 100 dB. The threshold was defined as the sound intensity that produces 10% of the maximum firing rate at the CF. The bandwidth was calculated at 10 dB above the threshold by finding the frequency range in octaves between upper and lower frequencies that elicit 10% of the maximum firing rate.

In the model, the dynamic range was defined as the range between threshold and best intensity of the response for Type O responses and the range between threshold and saturation for Type V and Type I responses. For Type V and Type I responses, the dynamic range was normally distributed with a mean of 30 dB and a standard deviation of 20 dB, and for Type O responses, the dynamic range was normally distributed with a mean of 20 dB and a standard deviation of 10 dB. If the actual dynamic range for a particular unit was randomly set to a value below 10 dB, that dynamic range was redefined as 10 dB. If the sum of the threshold and the dynamic range was greater than the maximum modeled intensity, the dynamic range was set to the difference of the maximum modeled intensity and the threshold. These dynamic ranges matched single-unit awake marmoset recordings where the dynamic ranges had a mean of 29.9 and 17.3 for monotonic and nonmonotonic neurons, respectively (Watkins and Barbour 2008). These numbers were slightly higher than the average 80% dynamic ranges (i.e., 10% of maximum firing to 90% of maximum firing) for these monkeys (Watkins and Barbour 2008) and somewhat higher than the average 80% dynamic ranges for barbiturate anesthetized cats, where mean dynamic ranges fell between 12 and 19 dB (Phillips and

Hall 1986; Schreiner et al. 1992). Overall, the total range of dynamic ranges in the model matches the total range of values that has been reported physiologically.

Simulated firing rates were obtained from the model by delivering pure tones or complex sounds to the modeled FRAs. For pure tones, the firing rate was obtained directly from the point on the FRA that corresponds to the frequency and intensity of the stimulus.

2.3.2 Primary Auditory Cortex Modeling

A1 was modeled by three fully overlapping arrays representing the three different canonical FRA classes. An 3-feature SOFM was used to define the position and the FRA properties (i.e., CF, bandwidth and threshold) of each unit in each array. The arrays were square with 150 × 150 units. The particular SOFM selected for the primary auditory cortex model had relative feature weightings of 10:2:1 for CF, bandwidth, and threshold, respectively because these weightings created maps that demonstrate similarity to those seen in sparse electrophysiological recordings in A1 of anesthetized cats (Schreiner and Mendelson 1990; Schreiner et al. 1992; Schreiner and Sutter 1992). Frequency is mapped smoothly in the resulting SOFM with the tonotopic axis running along the diagonals of the underlying grid. CFs ranged from 500 Hz to 32 kHz unless otherwise noted and were originally distributed uniformly on a logarithmic scale. Bandwidths ranged from 0.1 to 0.5 octaves with a median of 0.3 octaves, and thresholds ranged from 0 to 75 dB with a median of 35.8, which are values consistent with those observed in awake marmoset A1 (Sadagopan and Wang 2008). The functional maps used for this study can be see in Figure 2.4.

Units in the arrays were assigned FRAs using two different methods for different experiments: the combined method and the jittered method. Both methods yielded multiple array layers indexed by the same grid coordinates. Each layer contained a single type of FRA with an SOFM defining the locations and characteristics of the units. Stimulus response for each layer was normalized to the maximum firing rate of its corresponding type. Each layer was then weighted to adjust for differing relative numbers

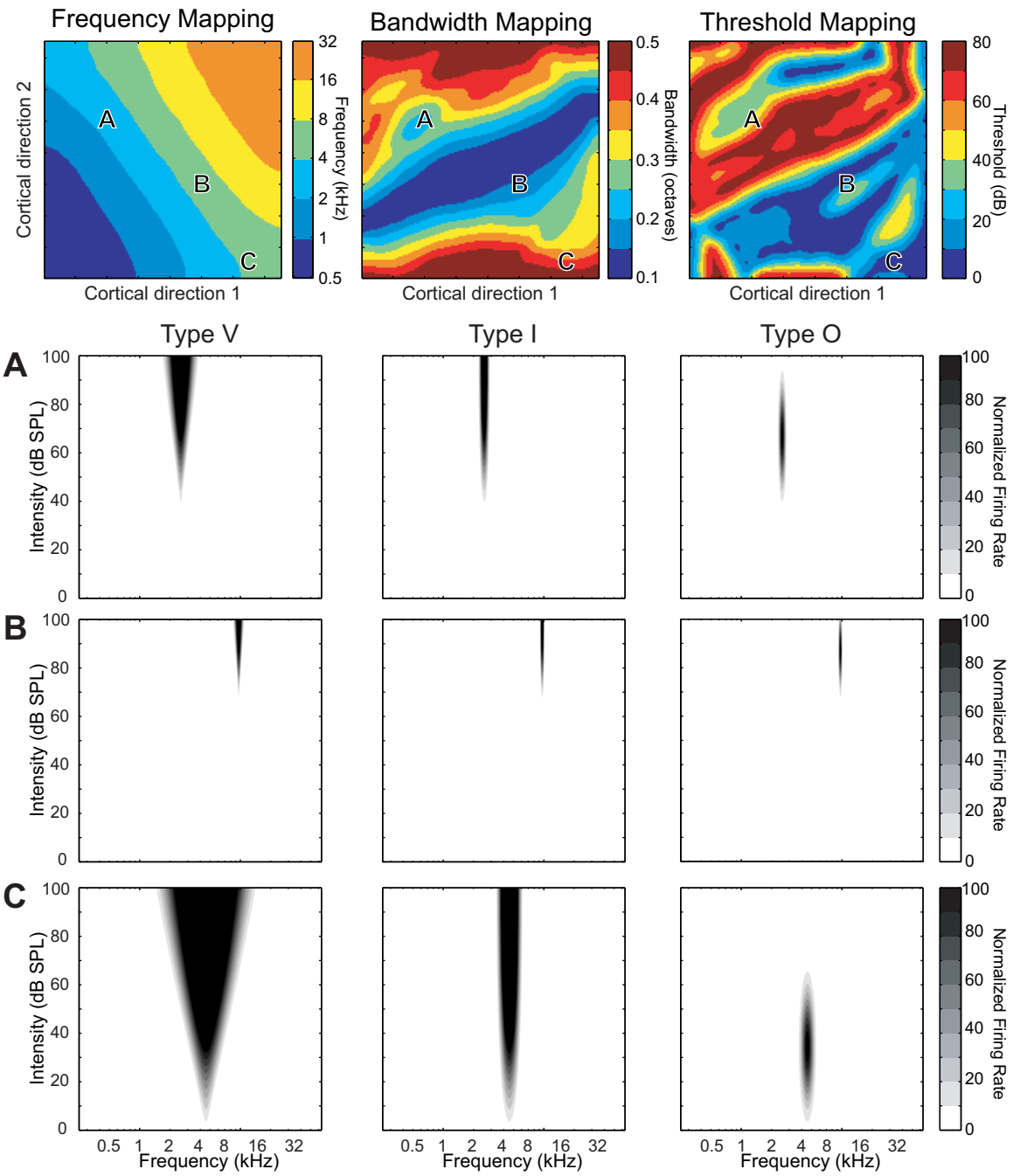
Frequency Mapping
Bandwidth Mapping
Threshold Mapping
Cortical direction 2
Cortical direction 1
Frequency (kHz)
Bandwidth (octaves)
Threshold (dB)
A
B
C
Type V
Type I
Type O
Intensity (dB SPL)
Normalized Firing Rate
Frequency (kHz)

Figure 2.4. Self-organization of primary auditory cortex functional properties.
In the model, three relevant neuronal receptive field features are mapped: center frequency, bandwidth, and threshold. Candidate maps for each of these features are presented in the top row. Each map reflects the same physical array of artificial cortical neurons. Three different sites in the array are labeled A, B and C, reflecting three different values for each of the three features. Below, each of these feature values is applied to the three major types of receptive fields observed in auditory cortex. Even at the same frequency, bandwidth and threshold, Type V neurons have larger receptive fields than Type I neurons, which in turn have larger receptive fields than Type O neurons.

of units belonging to each of the response types. Activity of all of the layers at the same index point (pixel) in response to a given stimulus were summed together to create an overall response. The combined method included three layers each with different FRA response classes and with no randomization of CF, bandwidth, and threshold maps.

The jittered method used nine total unit layers with three different layers for each of the three different FRA type. The CF, bandwidth and threshold maps were randomized to create pixel-by-pixel misalignment between each layer in terms of functional properties. This procedure was intended to simulate the variability in neuronal FRA characteristics observed within individual cortical columns. Jitter was quantified at a single array location for each feature as a percentage of the total feature range. Each of the three mapped features (CF, bandwidth, and threshold) at each point in the SOFM grid was remapped from the values indicated in the map to a random value within $[m - \frac{1}{2}(j \cdot r)$, $m + \frac{1}{2}(j \cdot r)]$, where m is the mapped feature value, j is the percentage of jitter and r is the feature range. For frequency and bandwidth the jitter was calculated in terms of octaves. For threshold the jitter was calculated in terms of dB. If any response feature was jittered outside the absolute lower or upper limits of the relevant defined values, then that feature was assigned the limit value. For both combined and jittered responses, the sum of all the response rates at each pixel was assigned to that pixel in the resulting population map. The summed responses across all pixels of the array were then normalized to a maximum spiking rate of 100 spikes/s.

To examine the response characteristics of the model, five logarithmically spaced pure tones were delivered (1, 2, 4, 8, and 16 kHz) separately and the resulting areas of activation in the model array were examined. Areas of activation were identified by visualizing all pixels with response rates above criterion values of 10% or 50% of the maximum rate. The average firing rate of the array was obtained by calculating the arithmetic mean of all the pixels. To examine the overlapping areas of activation, the mapped bandwidth and threshold values for the regions with overlapping areas of activation at the 10% response criterion were compared to the bandwidth and threshold

values for the activated regions without overlapping activation. Statistical comparisons between the bandwidths and the thresholds in the overlapping and the non-overlapping areas of activation were made using a student's *t*-test.

2.3.3 Functional Map Extraction

Functional maps with CF in the range [31.25 Hz, 8000 Hz] were extracted using pure tones. To extract frequency, twenty-one pure tones ranging from 15.625 Hz to 16 kHz were delivered at 80 dB and half-octave frequency intervals to combined arrays with either 33:33:33 or 10:10:80 relative weightings of Type V: Type I: Type O. For each point on the array, a Gaussian was fitted to the FRC resulting from the tone delivery. The mean of the Gaussian was designated as that point's CF. Bandwidth and threshold maps were extracted using 2121 pure tones at tenth-octave frequency intervals ranging from 15.625 Hz to 16 kHz and at 5 dB intensity intervals ranging from 0 to 100 dB. To extract the threshold maps for each indexed point, the maximum response for each stimulus intensity was obtained to create a rate-intensity response curve. This curve was then linearly interpolated and 10% of the maximum response was set as the threshold. Using the extracted threshold maps, the bandwidth at each point in the array was estimated at 10 dB above the threshold previously estimated for that point. The bandwidth was calculated by finding the frequency distance in octaves between the lower and upper frequencies that activated the unit at 10% of the maximum firing rate. The values were then capped with lower and upper limits of 0.1 and 0.5 octaves.

To quantify the accuracy of the extracted maps, the error between the SOFM maps and the extracted maps was calculated. For frequency, the error at each pixel was calculated by taking the difference in octaves between the mapped and the extracted frequency. For bandwidth and threshold, the pixel error was calculated by taking the difference between the mapped and the extracted values. The mean error was then obtained by averaging the error for each pixel over the entire array.

2.3.4 Complex Stimuli

Two vowels from one speaker, /æ/ and /ɛ/, were extracted from a single utterance of the TIMIT database (Linguistic Data Consortium, University of Pennsylvania). These vowels were delivered at 25 dB and 75 dB to model arrays of all three types of FRA responses, as well as a combined array with equal proportions of the three response types. The Fourier transforms of the vowels were computed and the resulting spectra modeled as a sum of sinusoids of various frequencies. The response of each array unit was computed by summing the individual responses of each of the constituent tones from the unit's FRA. The responses were then normalized such that the maximum response of each FRA type equaled 100 spikes/s. The absolute difference between two responses at the same stimulus intensity and of the same FRA type was calculated pixel-by-pixel. Average responses and average differences were obtained by calculating the arithmetic mean of either the array responses themselves or the pixel-by-pixel difference of two array responses, respectively.

Chapter 3 Virtual Mapping

3.1 Results

3.1.1 Linear Sampling Density of Estimation

The estimations of the four-feature map (depicted in Figure 2.1) are created using a total of 67,200 different map instantiations at 24 different sampling densities, 7 different jitter values, and 4 different features with 100 different instantiations for both averaged and unaveraged samplings. To compare the accuracy of maps with different linear sampling densities, the estimated maps were resampled back to 150 × 150 pixels with either nearest-neighbor or linear interpolation. As the sampling density increases, the true structure of the map becomes more apparent and the R^2 increases. Figure 3.1 shows with a series of estimated maps with unaveraged sampling and increasing sampling densities. These estimated maps demonstrate the improvement with the R^2 values reflects the visual improvement. Since the estimated values are only interpolated and not extrapolated, the units outside the sampled points are not estimated in the resampling process, so smaller sampling densities have a larger unestimated border. These unestimated borders are not included in the coefficient of determination calculations.

Estimated maps for feature 1 which has a low spatial frequency exhibit a low amount of error at even relatively low sampling densities. As the spatial frequency of the maps increases, a higher sampling density is required to determine the general structure of the map. For feature 4, for example, a linear sampling density of at least 10 was needed to produce maps with an mean R^2 of 0.3 while for feature 1, a map with a much lower spatial frequency, a linear sampling density of 2 produced a R^2 greater than 0.9 (Figure 3.2).

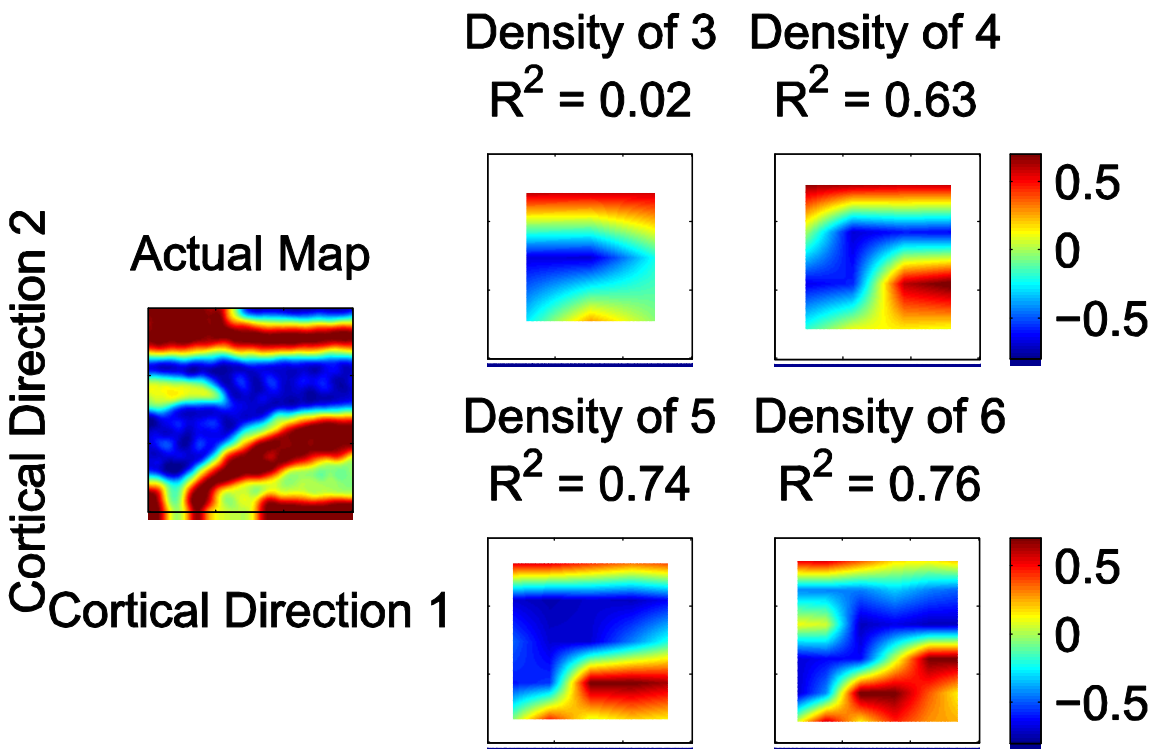

Figure 3.1. Examples of mapping estimations with different sampling densities. Unaveraged estimates were made with unjittered maps to examine the effect of different sampling densities. As the sampling density increased, the R^2 values increased. The estimated maps (four right panels) were derived from the 3rd feature map (left panel).

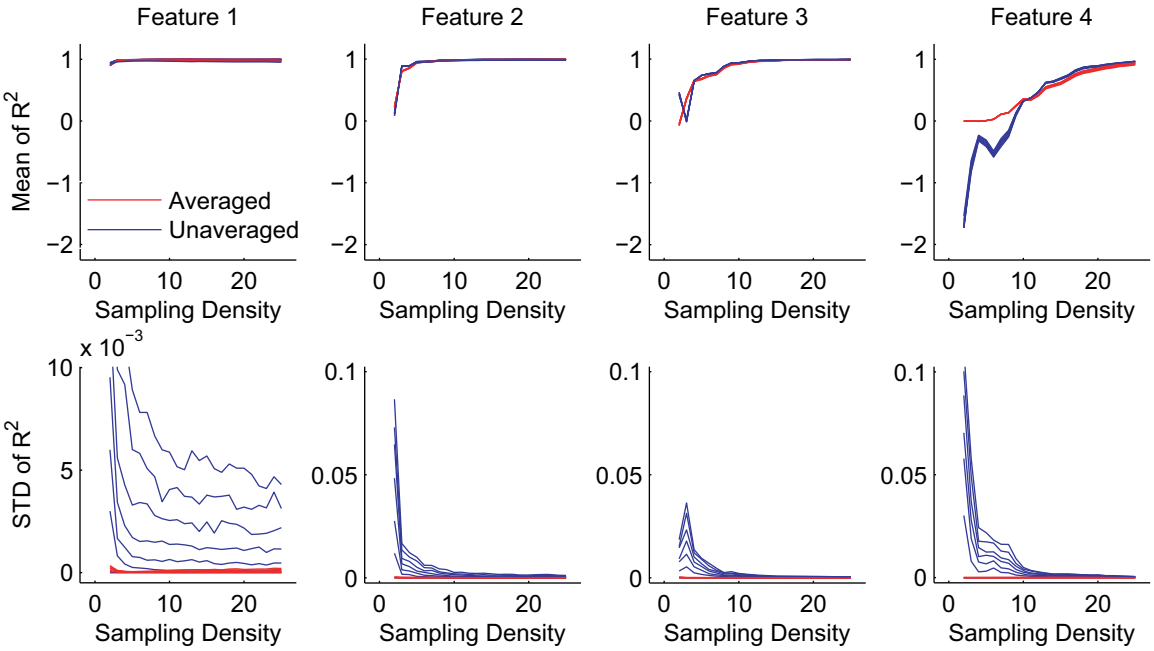

Figure 3.2. Comparison between averaged and unaveraged estimations.

Averaged and unaveraged estimations with uniform samplings were compared. The mean R^2 value of 100 runs for each case was calculated (top row) and the standard deviations of the R^2 values are shown in the bottom row.

At low sampling densities for feature 4, the R^2 values for the estimated maps with unaveraged uniform sampling displayed nonmonotonic behavior and decreased as the sampling density increased. The values of the regions with decreasing R^2 were all negative, suggesting that the estimated maps were undersampled. This nonmonotonic behavior of R^2 values can also be seen in Figure 2.3 when the sine wave was sampled below the Nyquist rate. For both Figure 2.3 and feature 4 after the spatial sampling density passed a threshold resolution, the R^2 started to increase as the spatial sampling density became high enough to begin to identify the structures in the map.

3.1.2 Averaged vs. Unaveraged Estimations

To examine the differences between mapping studies using single unit recordings and studies that result from averaging the activity of many neurons, estimated maps were sampled using two different methods. The first method samples the underlying maps by choosing single points that were uniformly spaced. The second method sampled the underlying maps by averaging all the points within a voxel. To compare the two sampling methods, the R^2 values were calculated for both methods, and the values suggest that the averaged sampling produces more accurate estimations for maps that were sampled at low sampling densities, maps with low spatial frequencies or for maps with high jitter (Figure 3.3). The unaveraged sampling produced more accurate maps with high spatial frequency maps or maps with low amounts of jitter. However, for maps with high spatial frequencies, such as the map for feature 4, the averaged estimations provided better estimations at low sampling densities.

The unaveraged estimations, in general, produces better estimations for maps with higher spatial frequencies because averaging acts as a low pass filter. The low pass filtering smooths out maps created by the jitter. At low sampling densities, the averaged maps closely reflect the mean of the map. This averaging prevents the R^2 from falling significantly below 0 because R^2 values of 0 represent when the sum of squared error is equal to the mean value of the SOFM. The unaveraged maps, however, can fall well below zero as seen in the R^2 values for the feature 4.

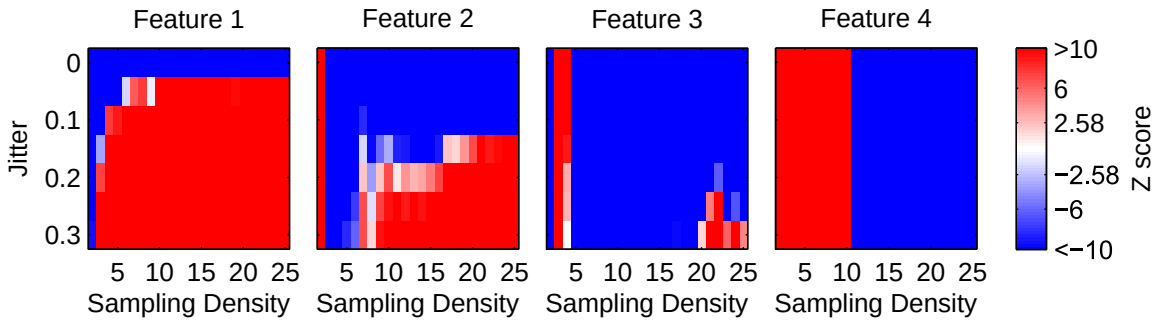

Figure 3.3. Statistical comparison between unaveraged and averaged estimations. The Wilcoxon rank-sum test shows in which cases the averaged estimation or unaveraged estimation provided better estimates of the underlying map. The blue shades represent cases where the unaveraged estimates produced higher R^2 values, and the red shades represent cases where the averaged estimates produced higher R^2 values than the unaveraged estimates. Z scores of 2.58 or greater and Z scores of -2.58 or less represent significant differences ($p < 0.01$) in the median.

The averaged estimations has a significantly smaller varience in R^2 values than the R^2 values of the unaveraged estimations ($p = 8.00 \times 10^{-70} \pm 1.53 \times 10^{-12}$ [sd.], two sample F-test) (Figure 3.2 bottom panels). Out of 672 runs, 625 had significantly lower variances in the R^2 values for the averaged estimations, 30 had significantly lower variances in R^2 values for the unaveraged estimations, and 17 had insignificant differences in R^2 values between the two types of estimations. For all cases with jitter, the averaged estimations produced significantly lower variances in R^2 values than the unaveraged estimations. Due to jitter, the maps estimated using the averaged procedures did not vary as much as the maps extracted using the unaveraged procedures. The averaging during the extraction provides a better estimate of the mean, thereby increasing the precision of the estimation.

3.1.3 Extractions with Random Sampling

The effects of random sampling on the ability to estimate maps were examined in Figure 3.4. In most cases, uniform sampling yielded a significant improvement over random sampling ($p = 7.67 \times 10^{-13} \pm 1.53 \times 10^{-12}$ [s.d.], Wilcoxon rank-sum test). Out of 672 sets, 627 had significantly higher R^2 for uniform sampling, 35 had significantly higher R^2 for random sampling, and 10 had no significant difference in R^2 values ($p < 0.01$, Wilcoxon rank-sum test). The uniformly sampled map had a significantly lower variance in their R^2 values than the random sampling ($p = 2.06 \times 10^{-4} \pm 4.12 \times 10^{-4}$ [s.d.]) with 664/672 incidents showing significantly lower variance with uniform sampling and 8/672 incidents showing insignificant differences ($p < 0.01$, two-sampled F-test). Because the local sampling density can increase and decrease in certain areas of the map, the random sampling may provide better map estimation when the sampling density is low. However, as the sampling density increases, this effect is diminished. The maximum R^2 value does not seem to be affected by sampling randomly, as the R^2 converge at high extraction densities.

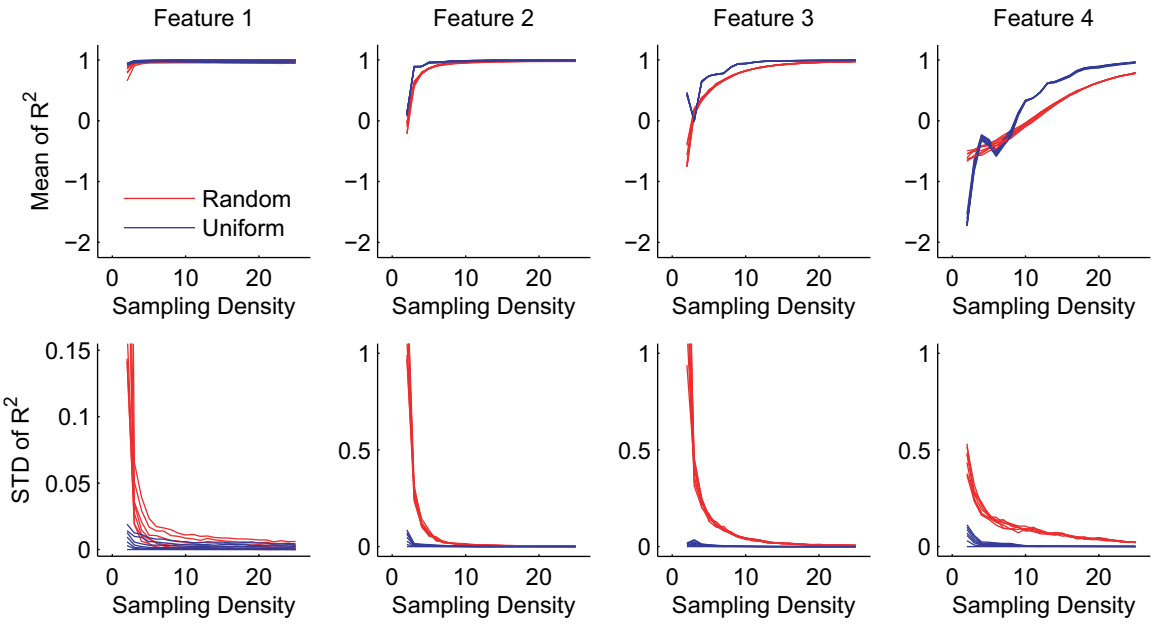

Figure 3.4. Comparison between uniform and random sampling.

The R^2 values of unaveraged estimations of uniform and random sampling were compared. The top row shows the mean of 100 R^2 values for each feature map for different linear sampling densities and different amounts of jitter. The bottom row shows the standard deviation of R^2 values.

3.1.4 Interpolation

Estimated maps created using zeroth-order and first-order interpolation were compared. For the plot in Figure 3.5, the estimates derived from linear interpolation produced higher R^2 values than estimates derived from zeroth-order interpolation. The improvement from the linear interpolation can be seen visually in Figure 2.2, where the mapped features are easier to identify in the resampled map (first-order interpolation) than the extracted values (zeroth-order interpolation). If the sampling density is high enough, a first-order interpolation will improve the quality of the map because a nonrandom map feature would have some continuity across the map. The continuity combined with adequate sampling density allows the first-order interpolation to improve the quality of the map when compared to the zeroth order.

3.2 Discussion

By taking a number of samples within an area of interest, mapping studies show the areas of the brain that are responsible for processing a certain task. This sampling is accomplished through either recording the activity from a single neuron or recording the summation or averaged activity over a number of neurons. The locations and the values of these samples are then combined together to create a map of how the area processed the task. This study seeks to simulate the process of mapping under optimal conditions to examine the effects of sampling density, averaging, sampling locations, and cortical variability on mapping. The model assumes perfect sampling, in that the single units represent a single neuron and that averaged units represent the averaged response over an area.

3.2.1 Sampling Density

To examine the sampling density needed to accurately estimate the underlying map, the sampling density is increased under various conditions. The calculated R^2 value compares

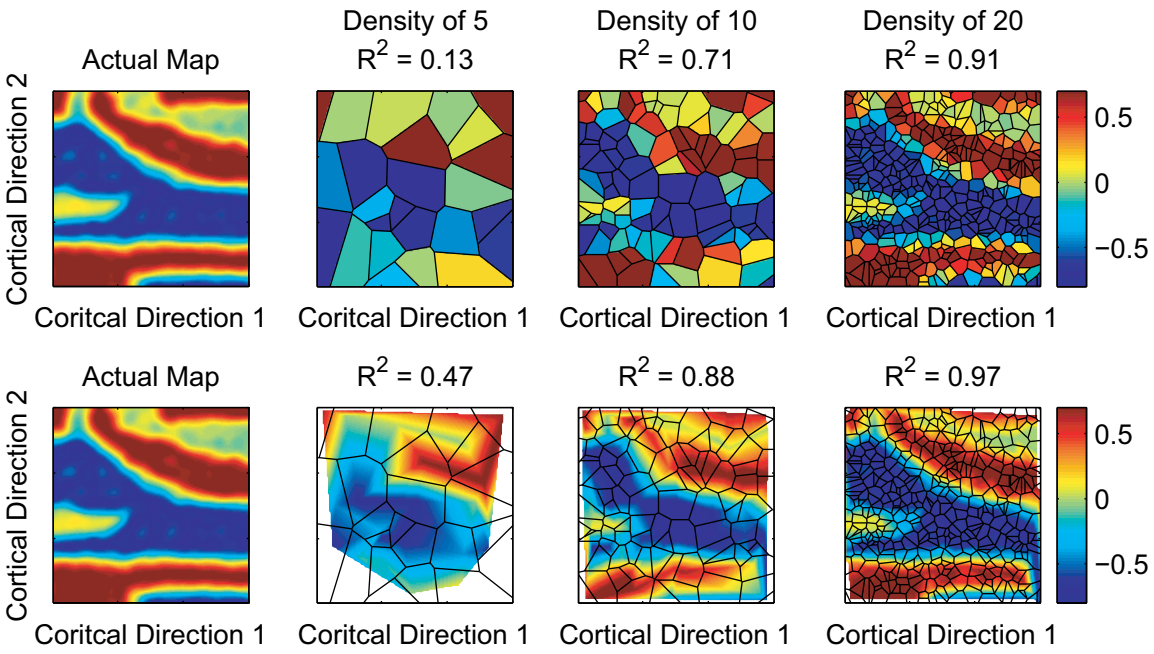

Figure 3.5. Comparisons between estimations using zeroth and linear interpolations. Averaged estimates of an unjittered feature 3 using zeroth-order (top row) and first-order interpolation (bottom row) for resampling are shown. For the first-order interpolated maps, the white pixels denotes locations outside the convex boundary which were not interpolated. The black lines represents boundaries where between the nearest-neighbor for the zeroth-order interpolation (top row), but are only used for comparison in the first-order interpolated plots (bottom row)

the estimated map to the underlying map and quantifies the quality of the estimation. The results show that the spatial frequency of the underlying map is the biggest factor in determining the necessary sampling density. The averaging of units, sampling locations, and cortical variability are minor contributors in comparison to sampling density. Maps with higher spatial frequencies require higher sampling densities to create accurate estimations of the underlying map. The increased sampling density is needed to capture the change within the map and if the required sampling density is not meet, the estimated map is undersampled. Evidence of undersampling can be seen in Figure 2.3. For a 2-dimensional sine wave, an undersampled sampling rate is considered to be any rate below the Nyquist rate where at least two samples are needed for every period. After the sampling rate drops below the Nyquist rate, the R^2 values drop below zero where the values start to fluctuate. This behavior with the R^2 values is also seen in the virtual mapping model. For feature 4 in Figure 3.2, the R^2 values for the unaveraged samples initially behave nonmonotonically because the estimated map is undersampled as indicated by the negative R^2 values. After the sampling density is increased, the R^2 values increase monotonically.

3.2.2 Averaging

To model single-unit studies and studies that summed or averaged activity, single points or averages points are sampled within a pixel/voxel. The results suggest that single unit studies produce more accurate estimated maps for underlying maps with high spatial frequencies, while studies that average activity perform better for estimated maps with low spatial frequencies or with cortical variability. Averaging units together acts as a low-pass filter which lowers the effect of cortical variability but reduces the ability to resolve sharp changes across the map. This limitation may be reduced if the sampling density is increased. At high sampling densities, this effect is diminished because the smoothing occurs over a smaller area.

3.2.3 Sampling Locations

For electrophysiology studies, the location of the sampled units for mapping plays an important part in estimating maps. The results shows that sampling in a uniform manner improves the accuracy and precision in estimating maps. It is theoretically possible for random sampling to create more accurate maps if the sampling density is highest where the spatial frequency is the highest. However, for an unknown map, controlled sampling according to spatial frequency is not possible, so uniform sampling ultimately improves the accuracy. While electrophysiology studies have sampled uniformly with the use of a grid (Recanzone et al. 2000), most electrophysiology studies do not use grids to ensure uniform sampling. Even with the use of grids, uniform sampling may not be possible due to blood vessels, the geometry of the brain, and other potential factors. Functional imaging studies typically sample uniformly because the pixels/voxels are evenly spaced. However for certain imaging modalities such has optical imaging, the geometry of the brain may prevent recording over the entire region of interest.

3.2.4 Cortical Variability

Neurons within a cortical column have been known to have varying characteristics (Atencio and Schreiner 2008; Phillips and Irvine 1981). These varying characteristics were modeled as cortical variation in this study. The results suggested that studies which average activity may perform better than single-unit estimations when cortical variability is present. Surprisingly, when compared to the effect of spatial frequency in the estimating map, the amount of cortical variability has a negligible effect on the R^2 values. This effect is most likely reduced in the model because the values obtained in this model were from a Gaussian distribution. Averaging the values between layers of arrays negates the randomness created by the jitter. When samples are not averaged, the effect of the jitter is overshadowed by the undersampling at low sampling densities or is limited due to high sampling densities. Therefore, while averaging neurons within a cortical column may reduce the effect of cortical variability within a column, the averaging does not reduce the sampling density needed to resolve maps

3.2.4 Display of Estimated Maps

A popular method to visual maps with randomly sampled points is with a zeroth order interpolation, which forms plots known as Voronoi tessellation plot (Figure 3.5) (Cheung et al. 2001; Godey et al. 2005; Read et al. 2001). Voronoi plots allow the reader to identify both the sampling density and the mapped feature values at the same time. The linear interpolation allows the maps to be smoothed out, but the points outside the convex boundary of the sampled points are lost. The linear interpolation produces a more accurate estimation of the underlying map. In addition, linear interpolations improve the ability to visibly discern maps because the maps allow the eyes to identify extreme values more easily.

3.2.6 Relation to Previous Studies

The sampling density required for estimating a map is based on the spatial frequency of the map. As seen in Figure 3.2, maps with low spatial frequencies, such as feature 1, can produce relatively accurate maps at even the lowest sampling densities. However, maps with high spatial frequencies (such as feature 4) require a linear sampling density of well over 15 points to produce estimations with R^2 higher than 0.7. Electrophysiology studies in the primary auditory cortex sometimes sample at or near these densities. Electrophysiology studies for individual animals have been sample maps using 179 (Cheung et al. 2001), 289 (Bonham et al. 2004), and 352 penetrations (Philibert et al. 2005). However, the averaged number of penetrations per experimental animal for the studies is much lower running around 80 to 100 penetrations (linear sampling density of 9 to 10 [see *Section 2.2* for conversion]). Thus, these studies may be able to resolve maps with low spatial frequencies but may have difficultly resolving maps with high spatial frequencies. Imaging studies use linear sampling densities in excess of 80 points (Ojima et al. 2005; Petkov et al. 2006); however, imaging studies still have difficulty in reproducing maps obtained from electrophysiology. Since the resolution of the imaging studies greatly exceeds the sampling density needed to accurately resolve maps, the problems with functional imaging recapitulating electrophysiology studies in the auditory cortex most likely stem from limitations outside of the properties test in the chapter. One

possible limitation is that neurovascular coupling degrades functional imaging resolution. Functional imaging modalities like optical imaging of intrinsic signals and fMRI-BOLD average signals by measuring the hemoglobin concentrations (Logothetis et al. 2001; Ojima et al. 2005). By measuring hemoglobin concentrations, these functional imaging studies measure both the subthreshold and suprathreshold activity is measured while single-unit electrophysiology studies predominately measure firing rates or suprathreshold activity. Another possible fundamental limitation may lie within the neural circuitry of auditory cortex. Chapter 4 examines possible limitations that may be created from auditory cortex neural circuitry.

Chapter 4 Virtual Imaging of the Primary Auditory Cortex

4.1 Results

4.1.1 Model Array

The spectral receptive fields of auditory neurons in response to pure tones come in three canonical forms (Ramachandran et al. 1999; Sadagopan and Wang 2008). Type V neurons respond best at a single frequency, but their response bandwidth (i.e., receptive field size) widens in response to increasing intensity. These responses are reflective of the spectral properties in auditory nerve fibers (Kiang et al. 1965; Palmer and Evans 1980; Sachs and Abbas 1974). Type I neurons respond best at a single frequency and maintain a consistent bandwidth as intensity increases. Type O neurons respond best at a single frequency and have constant bandwidth as intensity increases, but they stop responding or are inhibited at the highest intensities. Because Type I and Type O neurons do not exist in the auditory nerve, they must be created by central auditory circuits, at least some of which are located within A1 (Tan et al. 2007; Wehr and Zador 2003).

Figure 4.1A shows a spike raster plot of one awake marmoset monkey auditory cortical neuron's response to increasing intensity of a tone at its characteristic frequency or CF (4.2 kHz). The shaded area represents the tone duration, and each dot represents a single action potential spike. Stimulus intensity increases from bottom to top. Note that tone-stimulated spiking during the tone interval is evident only over a particular intensity range. Following tone offset, however, an additional spiking pattern is evident at a different latency and over a different intensity range. This offset spiking rate increases

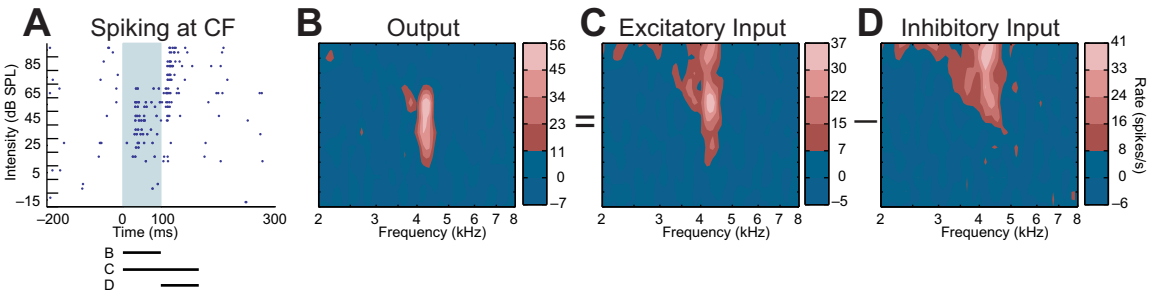

Figure 4.1. Single unit spiking activity reveals 3 response types.

Pure tones varied in frequency and amplitude elicit the three canonical receptive field categories in marmoset primary auditory cortex. **A.** Spike raster plot of one neuron's response to a tone delivered at 4.2 kHz and varied in amplitude. This plot reveals different spiking during the tone delivery period (shaded region) and afterwards—a relatively rare feature that appears to reflect the input contribution of both inhibition and excitation in this neuron. Inhibition that occurs during the stimulus interval appears to rebound following stimulus offset, resulting in spiking of the output neuron after the stimulus is removed. **B.** Spikes occurring during the stimulus interval reflect the output of the neuron (excitation plus inhibition). **C.** Spikes occurring during and following the stimulus interval reflect the excitatory input to the neuron because the inhibition decreases spiking during the stimulus interval and increases spiking afterward, thereby canceling out when these two intervals are analyzed together. **D.** Spikes occurring after the stimulus interval reflect inhibition alone. Intervals of analysis are depicted schematically below the raster plot in A.

proportionally with the degree of inhibition observed during the tone interval (evidenced by the decreasing spike rate during the tone delivery at higher intensities, including complete suppression of all spiking at the highest intensities). In this neuron, inhibition during the stimulus interval is measurable directly by a rebound excitation following stimulus removal, a phenomenon referred to as anode-break excitation (Hodgkin and Huxley 1990). The anode break is most prominent for high stimulus intensities when spiking rate during the stimulus is lower than the spontaneous spiking rate. Spiking below spontaneous rates implies local inhibition within A1 because A1 does not receive inhibitory projections from other areas.

Figure 4.1B shows the receptive field or frequency response area (FRA) measured during the tone interval for tones varied in frequency and intensity. This response class is Type O, also commonly referred to as "nonmonotonic" because the spike rate at CF is a nonmonotonic function of intensity. This response can be thought of as the output of this cell, which is formed from the balance of excitatory and inhibitory input. The excitatory input can be calculated by measuring the spiking rate over both the stimulus and the post-stimulus interval because the inhibitory rate measured during the tone interval is canceled by the rate measured during the rebound excitation. The FRA of this excitatory input appears much like a Type I response, as shown in Figure 4.1C. Finally, if rates are calculated using only the spikes following stimulus offset, the contribution of the inhibitory input alone can be evaluated. As seen in Figure 4.1D, the inhibitory input into this neuron has a Type V FRA.

This example A1 neuron indicates indirectly that specialized responses such as Type O FRAs can be created by inputs tuned to the same CF but different bandwidths and thresholds. Systematic physiological studies yield findings consistent with the circuits inferred from this example (Caspary et al. 1994; Kaur et al. 2004; Tan et al. 2007; Wang et al. 2002; Wehr and Zador 2003). Type V, I and O neurons within A1 therefore appear to combine in very precise circuits to produce specialized sensory coding elements, the specific functions of which are not yet fully understood. The impact of these findings for functional maps is more apparent, however. The existence of specialized circuits repeated

across the cortical surface implies the need for precise spatial arrangements of neurons in order to achieve efficient feature space representations and interconnections (Chklovskii and Koulakov 2004; Koulakov and Chklovskii 2001; Kozloski et al. 2001; Watkins and Barbour 2008). For this reason, virtual functional images of A1 were investigated with plausible maps of frequency, bandwidth and threshold. These neuronal measures reflect the parameters needed to describe FRAs and are non-randomly distributed in A1 (Bonham et al. 2004; Cheung et al. 2001; Merzenich et al. 1975; Philibert et al. 2005; Schreiner and Mendelson 1990; Schreiner and Sutter 1992). Outside of frequency, however, the fundamental principles governing physiological feature organization within A1 have been more challenging to discern with electrophysiology. A neuronal array representing spatially organized A1 neurons was modeled to determine how effective functional imaging may be in extracting auditory maps of multiple features.

The maps of frequency, bandwidth, and threshold that were used in the virtual imaging experiments can be seen in the top row of Figure 2.4. Each point in the array (also referred to as a pixel or a unit) represents a single frequency, bandwidth, and threshold combination used to determine the properties of the three relevant types of receptive fields, depicted as FRAs. Examples indicated by A, B, and C in Figure 2.4 depict nine units divided into classes by column, but with identical features in each row. Differences in FRA shapes across each row derive solely from the nature of the receptive fields themselves while the differences down each column derive solely from different parameter values.

This model system was used to explore a common experimental observation from imaging studies of A1: extensive activation of large areas of cortex by pure tones. A pure tone represents the stimulus that activates the smallest proportion of sensory epithelium in the cochlea. Equivalent stimuli in the visual and somatosensory systems would be a point of light or a pinprick, respectively. As seen in Figure 4.2, OIS experiments reveal that a single pure tone can activate a sizable portion of A1 (Harel et al. 2000; Harrison et al. 1998; Ojima et al. 2005; Spitzer et al. 2001), which contrasts with imaging results from other sensory modalities (Das and Gilbert 1995; Grinvald et al. 1994). Even tones

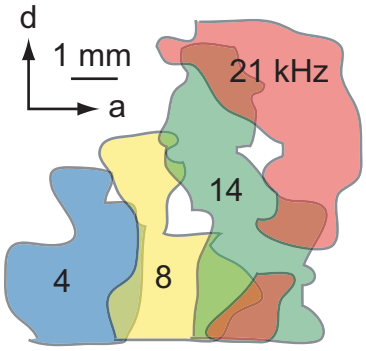

Figure 4.2. Sample intrinsic optical imaging experiment from cat A1.
In this experiment pure tones were delivered sequentially to the animal at a sound intensity of 60 dB and over a total frequency range of about 2.5 octaves. The areas with statistically significant change in absorbance with respect to baseline shown in color. Tones separated even by an octave of frequency can activate overlapping regions. Collectively, these four simple stimuli activated nearly all of primary auditory cortex. Adapted from (Ojima et al. 2005) by permission of Oxford University Press.

spaced an octave apart still elicit overlapping areas of A1 activation. Collectively, the four tones depicted in this figure activate most of A1. The model was evaluated to see if this extensive spread of activation could be attributed to the underlying physiology of A1.

4.1.2 Neuronal Subpopulation Responses

Virtual imaging experiments were initially conducted using the computational arrays for each neuronal subtype separately, as shown in Figure 4.3. Five 80 dB tones spaced one octave apart in frequency were delivered consecutively to each of three arrays consisting entirely of Type V, I or O units. The response criterion for attributing a driven response to a particular unit was set at either 10% or 50% of maximum firing (top row or bottom row, respectively). All array elements responding at rates greater than the appropriate criterion are color coded according to the frequency of the tone delivered.

Differences between the response types are rather striking. Type V responses (leftmost panels) showed the greatest overall activity and extensive overlap in some activated areas for tones of different frequencies. Units in areas of overlapping activity showed significantly greater bandwidths and lower thresholds than those in nonoverlapping areas ($p < 0.001$, student's t test). These areas corresponded to the upper left and lower right of the Type V array. The combination of high bandwidth and low threshold creates a large FRA for monotonic responses, allowing for a greater variety of stimuli to activate these units than would activate units with a low bandwidth or a high threshold or both. Type O responses (rightmost panels) showed a very different response profile with much smaller, nonoverlapping, punctate areas of activation. A single pure tone at 4 kHz and 80 dB, for example, activated 24% of the Type V array elements above the 10% response criterion but activated only 3.0% of the Type O array elements. Type I responses showed trends between those of Type V and Type O units with no activation overlap for the tones spaced one octave apart (middle panels). Raising the response criterion (e.g., from 10% to 50%) is a manipulation an experimenter could perform in order to attempt to localize frequency responses more accurately, which these virtual imaging experiments demonstrate would yield a visible improvement in frequency map determination. Type V

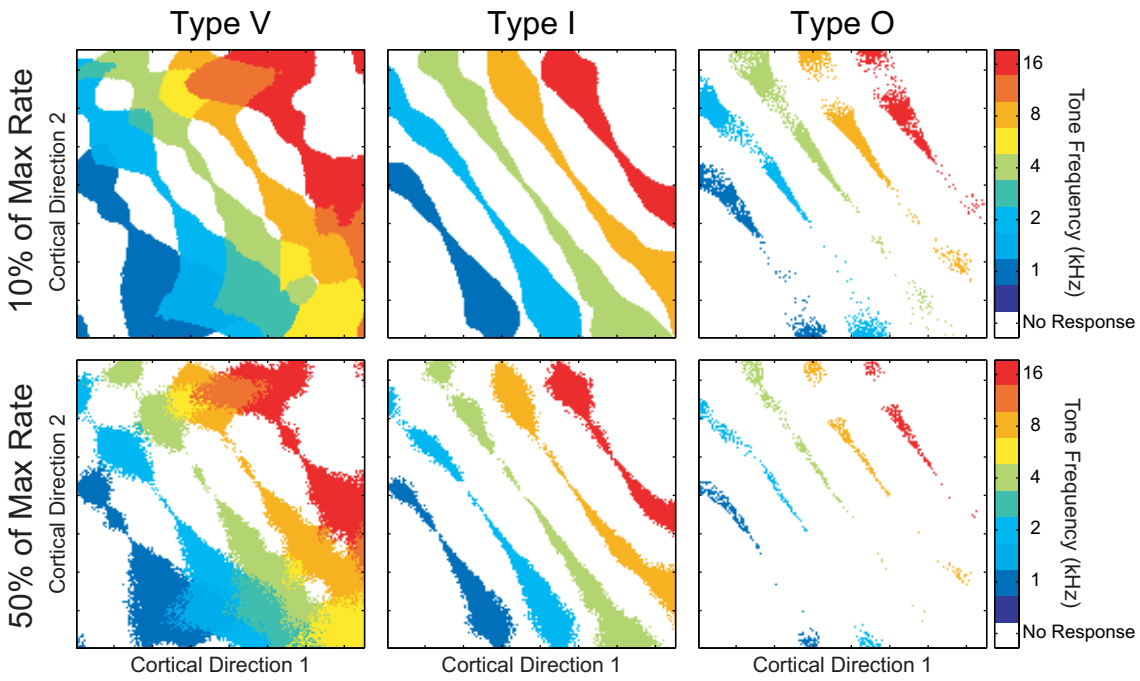

Figure 4.3. Virtual imaging of different response types.

Virtual imaging experiments reveal distinct patterns of primary auditory cortical activity for each of the three main receptive field categories. Five pure tones delivered to an array of Type V neurons at 80 dB and spaced one octave apart in frequency collectively activate a large proportion of the underlying area, including substantial overlapping activity (left column). This pattern is evident both when the neuronal response criterion for inclusion is set at 10% of the maximum response for each pixel (top row) or 50% (bottom row). Similar patterns are seen for Type I neurons, although with less overlap and overall activation at map areas of higher bandwidth (middle column). In contrast, smaller isolated patches of cortex are activated by pure tones for Type O neurons (right column).

responses to octave-spaced tones are likely to overlap at regions of high bandwidth, however, even at relatively high response criteria.

The experiments shown so far were conducted with arrays containing only one type of response FRA. Neurons with different FRAs are generally commingled in A1, however, some combinations of different unit types must be evaluated to understand the expected behavior of the entire cortical area. These commingled arrays did not explicitly take inhibition into account, but imaging methodologies based upon correlates of neuronal activity such as blood oxygenation should reveal activity regardless whether the neurons are excitatory or inhibitory. Figure 4.4 shows five different relative proportions of the three FRA subtypes within each pixel of the corresponding array. In this case, each pixel can be thought of as a simple average of its constitutive units. If 100 units contributed to each pixel of the 50:25:25 array, for example, then 50 of them would be Type V, 25 would be Type I, and 25 would be Type O. The stimuli and analysis are the same as in Figure 4.3 with response criteria applied to the mean response for each pixel. At the 10% response criterion (top row), Type V responses tended to dominate the appearance of the population responses when they represented as little as 20% of the total. Their influence can still be seen even when they represent only 10% of the total (upper right), although it diminished dramatically at higher response criteria, as can be seen with the 50% criterion (bottom row). Both Type V and Type I responses lost all influence on the activated area at the 50% criterion when their individual proportions dropped below 25% (bottom right two panels). In those cases only Type O responses can be discerned, and they are relatively robust. This phenomenon occurs because Type O FRAs contribute the most to pixel averages in those situations, and once the response criterion is set below the contribution of Type O FRAs and above the contributions of Type V/I FRAs, the latter disappear from the resulting image. Depending upon the relative proportion of the receptive field subpopulations within A1, therefore, different response criteria are likely to alter the gross appearance of areas activated by tones. In fact, under the conditions used to create the model array, the response criterion at which tone activation areas fracture into multiple noncontiguous regions exactly matches the relative proportion of

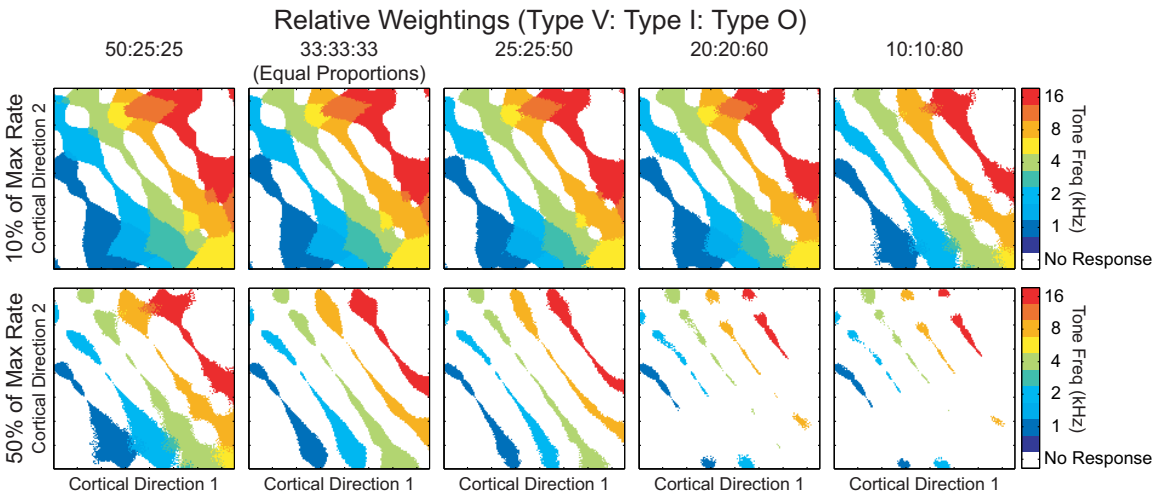

Figure 4.4. Virtual imaging with a variety of weightings between response types. When the overall neuronal population contains a mixture of receptive field types, the resulting functional maps depend upon the relative activity of the subpopulations. Relative proportion of Type V/I responses decrease from left to right, while relative proportion of Type O responses increases. At the lower detection criterion of 10% (top row), Type V responses tend to dominate the maps even when they represent a minority of responses. At the higher detection criterion of 50% (bottom row), the contribution of Type V/I responses to the overall functional map diminishes as their relative proportions diminish.

Type O FRAs. A similar observation in real functional imaging experiments may provide an estimate of the actual Type O prevalence *in vivo*.

Given the unique receptive field dependences of Type V and Type O responses upon stimulus intensity, one might expect areas of activation to be functions of tone intensity. Figure 4.5 shows activated array areas at 10% response criterion for a pure tone at 4 kHz with intensity stepped from 20 dB through 100 dB in 10 dB steps. As shown in Figure 4.5A, activated areas for Type V responses generally increased both along isofrequency contours (upper left to lower right) and iso-bandwidth contours (lower left to upper right) as a function of increasing tone intensity. Activated areas for Type I responses, on the other hand, increased only along isofrequency contours until a narrow, contiguous band became active (Figure 4.5B). Type O responses were different still: activated areas migrated along isofrequency contours from a low-threshold region toward a high-threshold region (Figure 4.5C). While collective Type V and Type I array response areas increased at higher tone intensities, array response areas remained constant for Type O responses. Upon combining the three response classes in equal proportions into a single array, the activation patterns can be seen to mirror Type V responses closely (Figure 4.5D). This same trend is apparent at other relative unit proportions (data not shown). Smaller population activation areas, as well as more accurate frequency estimation, can therefore be expected in A1 at lower tone intensities.

Neurons isolated at different depths within a cortical column of A1 are known to vary somewhat in their response characteristics (Atencio and Schreiner 2008; Phillips and Irvine 1981). In other words, individual neurons within a column may have features that deviate from the relevant feature map. This map variability may contribute to challenges in discerning functional A1 maps, both electrophysiologically and through functional imaging. To assess the effects of columnar variation upon functional imaging experiments, a population array was modified to map features randomly instead of using deterministic values as was done for the previous experiments. In the modified array, the maps refer to mean frequency, bandwidth and threshold at any given point. The three FRA response type subpopulations were combined in equal proportions. Each

subpopulation was further divided into three additional groupings reflecting unique instantiations of the relevant random variables. If, for example, 100 units contributed to the response of each pixel, then (in round numbers) 33 neurons would belong to each of the three responses types, and every 11 neurons would have feature values drawn from different instantiations of the relevant uniform distributions. These arrays were referred to as jittered arrays.

To examine the effects of variations within a cortical column on functional imaging results, tones were delivered at 4 kHz and 80 dB to a jittered array comprising equal proportions of the three neuronal response classes (i.e., 33:33:33). The resulting responses are shown in Figure 4.6. The no-jitter case (Figure 4.6A) reflects no randomization, thereby matching the configuration of all the previous arrays. Contour lines representing deciles of maximum firing rate can be seen, and the overall structure is noticeably influenced by Type V neurons. As jitter is increased to 10%, the response contours are disrupted, but little overall change in the activated area is evident to the naked eye (Figure 4.6B) with one exception: jitter appears to have the greatest impact upon regions of the map with the lowest bandwidths and the highest thresholds. This trend continues at 20% jitter (Figure 4.6C) and beyond (data not shown) and implies that even low-resolution functional imaging results may be relatively robust in the absence of a strict cortical columnar structure of functional similarity, at least at map regions of higher thresholds. To put this finding into perspective, an electrode penetrating a cortical column with 20% jitter at a 4000 Hz point in the jittered map would be equally likely to encounter any frequency in the range of 2639 Hz to 6063 Hz. Even with a clear map of parameter means, such jitter could make discerning the structure with sparse electrode penetrations quite challenging, while the natural neuronal averaging inherent to imaging could enable a fairly clear depiction of the underlying map.

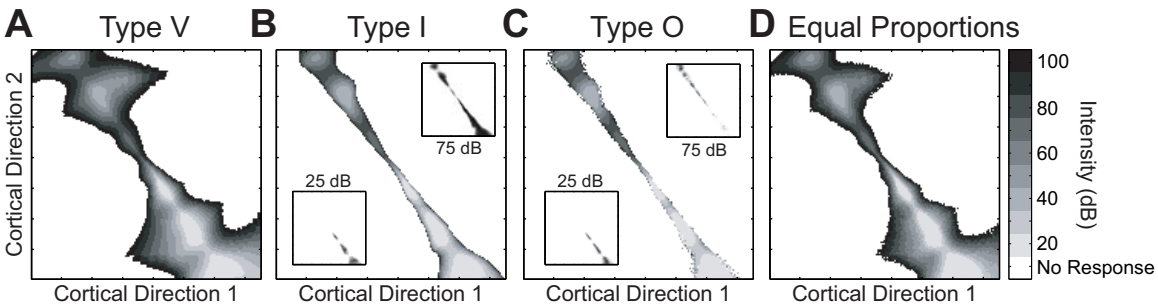

Figure 4.5. Areas of activation with varying stimulus amplitude

Stimulus amplitude predominantly affects the maps of Type V responses. **A**. As the amplitude of a pure tone at 4 kHz increases, Type V subpopulation responses increase in area parallel to the frequency axis. **B**. Type I response areas increase orthogonally to the frequency axis. **C**. Type O response areas do not increase. **D**. Population responses comprising equal proportions of the three subtypes demonstrate increasing activation area parallel to the frequency axis as tone amplitude increases because of the Type V contribution. For all large plots, each pixel is assigned the gray value of the lowest intensity to which its underlying array element responded. Insets show array responses to a tone at one intensity with grayscale values corresponding to response rate.

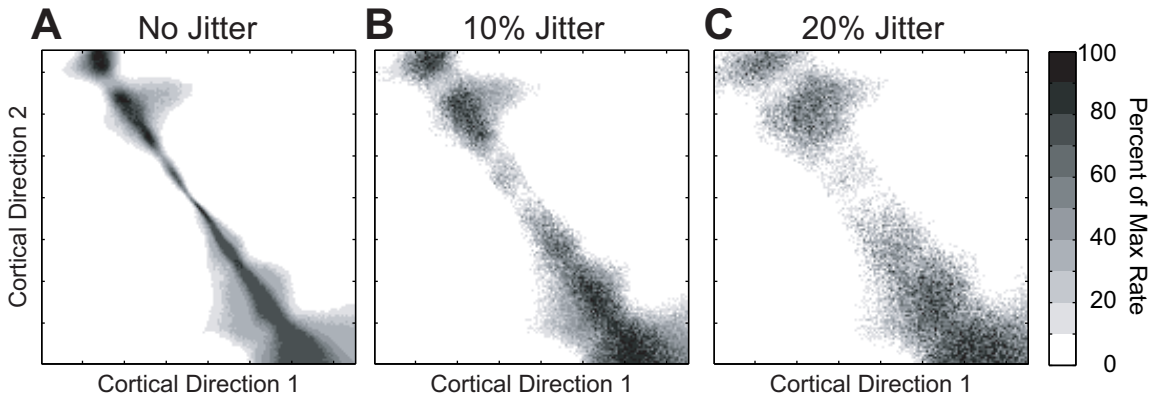

Figure 4.6. Virtual imaging experiment with different amounts of cortical variability. Variability in map fidelity predominantly affects regions of low receptive field bandwidth. A pure tone at 4 kHz and 80 dB was delivered to an array with equal proportions of the three neuronal subtypes. The value of each of the three mapped stimulus features was mapped to each element of the array (i.e., pixel) as uniformly distributed random variables whose means were the feature values for the exact map. Variability in the random maps is presented as a percentage of the mean (i.e., jitter). The exact map is shown on the left (**A**) and standard deviations of the random variables increase toward the right to 10% jitter (**B**) and 20% jitter (**C**). Deciles of rate response are depicted in grayscale. The map areas most clearly affected by randomized mappings corresponds to the narrowest receptive fields and, secondarily, the highest thresholds.

4.1.3 Functional Map Extraction

The feasibility of extracting the original feature maps of frequency, bandwidth and threshold from the virtual imaging results using only pure tone stimuli was examined next. For these models, a frequency range of 31.25 Hz to 8 kHz was used. Figure 4.7 shows maps extracted using relatively simple procedures (see Chapter 2.3.3). CF maps were extracted fairly accurately using only 21 tones (Figure 4.7A). A small portion of the map in the center of the array (<0.4% of the total map) did not produce any response because the CFs of the units located in these regions fell between the tone frequencies presented, and the bandwidths of these units were very narrow. With a larger number of tone stimuli, these unresponsive areas disappeared (data not shown). Ignoring these areas for the case of 21 tones, the estimated CF was on average 0.058 octaves away from the mapped CF values for arrays with equal proportions of FRA types and 0.072 for arrays with 10:10:80 proportions of Type V:Type I:Type O responses (data not shown). Both threshold maps and bandwidth maps were extracted using a large number of pure tones. Mapping both features required tones to be spread out over a variety of sound intensities and frequencies instead of just frequency as in mapping out CF. For threshold, a large number of tones must be spaced out in frequency to predict an accurate rate-intensity curve at the CF. If 2121 pure tones are used, threshold maps were also extracted very accurately (Figure 4.7C). The mean difference between the extracted and the mapped thresholds was 0.84 dB for the arrays with equal proportions of FRA types and 0.69 dB for 10:10:80 arrays. Bandwidth was not easily predicted, however (Figure 4.7B). In order to resolve the bandwidth map, the tones must be spaced in frequency no farther apart than least the lowest neuronal bandwidth. Furthermore, the intervals of tone amplitudes must be small enough to allow for accurate prediction of the bandwidth at the desired intensity (in this study, always measured at 10 dB above threshold). With pure tones presented at tenth-octave intervals of frequency and 5 dB intervals of amplitude, the mean difference between the extracted and actual bandwidths was 0.054 octaves for arrays with evenly distributed proportions of FRA types and 0.052 octaves for arrays with 10:10:80 arrays. These errors represent 18% of the range of mapped bandwidths (0.1 to 0.5 octaves), which is considerably higher than the errors for the other two features. While this degree

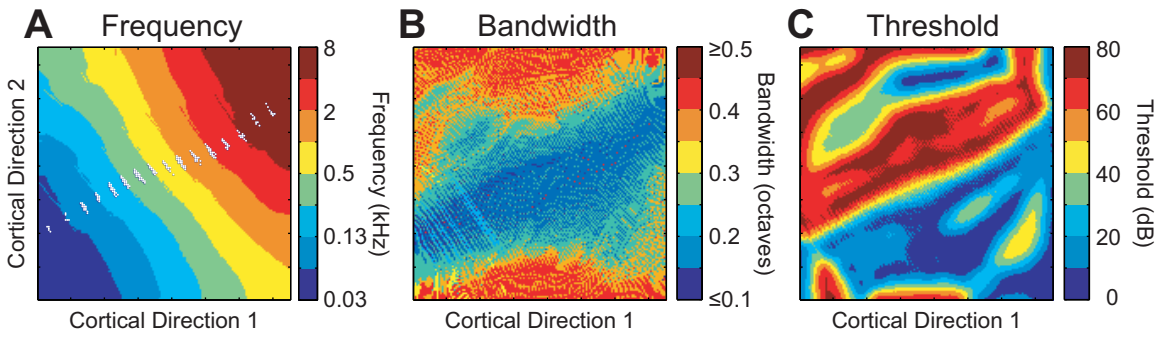

Figure 4.7. Extracted feature maps.

Feature maps can be extracted using sequentially delivered pure tones. **A**. Twenty-one pure tones were delivered at half-octave intervals at 80 dB, and the fitted frequency eliciting the greatest response at a given pixel was assigned to that pixel as its characteristic frequency (CF). CF extraction using pure tones produced an accurate estimation of CF, except for small regions in the center of the map where receptive field bandwidth was narrower than the frequency spacing. **B.** Bandwidth maps were extracted using many pure tones at tenth-octave frequency intervals and 5 dB amplitude intervals. Using the extracted threshold maps, the bandwidth was measured at 10 dB above the estimated threshold. Relatively large errors were generated with this technique. **C**. Threshold maps also extracted using the same stimuli produced relatively accurate map estimates. Each pixel in the extracted threshold map was assigned the lowest stimulus intensity that elicited a spiking rate above the response criterion.

of error does not allow for a fine bandwidth map to be discerned even for large numbers of stimuli, the gross organization of the map can be identified, especially the highest- and lowest-bandwidth regions. Smaller numbers of stimuli may also be just as effective at providing crude estimates of "high-bandwidth" and "low-bandwidth" regions.

4.1.4 Responses to Complex Stimuli

Distributed neuronal activity can be represented sufficiently well with functional imaging to classify sensory stimuli based solely upon the functional images (Kay et al. 2008). To examine the robustness of complex stimulus representation in the model arrays, complex spectra were delivered for the vowels /æ/ as in "bad" and /ɛ/ as in "bed" to arrays of individual neuronal subtypes as well as an array of equal proportions of the three FRA response subtypes (Figure 4.8). Frequency ranges of these arrays spanned 31.25 Hz to 8 kHz, and to test the robustness across intensity in particular, the vowels were delivered at both 25 dB and 75 dB. Similarly to the pure-tone responses, Type V and Type I responses increased their overall activation areas as sound intensity increased while Type O responses simply shifted their activation areas across the array. The mean firing rate of all the Type O responses was independent of the sound intensity of the stimulus (25 spikes/s for both /æ/ and /ɛ/ at both intensities), while for the Type I and Type V responses the mean firing rate depended upon the sound intensity (e.g., for stimulus /æ/, mean firing rates for Type V responses of 4 and 60 spikes/s at 25 and 75 dB, respectively, and for Type I responses of 11 and 72 spikes/s at 25 and 75 dB, respectively). The combined responses for all neuronal types appeared to be dominated by Type O and Type I responses at low sound intensities (bottom left two panels). At high intensities, however, the combined responses appeared to be dominated by the saturated Type V responses (bottom right two panels), although the highest spiking rates at individual pixels were attributable to Type O responses.

Notably in Figure 4.8, differences in sound intensity of a particular vowel appeared to elicit greater distinctions in overall array activity than did two different vowels at the same sound intensity. This characteristic was apparent for all three response classes, as

well as the population response. To visualize these differences in array activity, a differential image was created between the two vowels at 25 dB and 75 dB for populations of each response class, as well as the combined population. Remarkably, differential images for arrays of Type V and Type I responses showed virtually no distinction between these vowels at low intensities, although obvious distinctions were apparent at high intensities (Figure 4.9, top two rows). Type O responses, on the other hand,

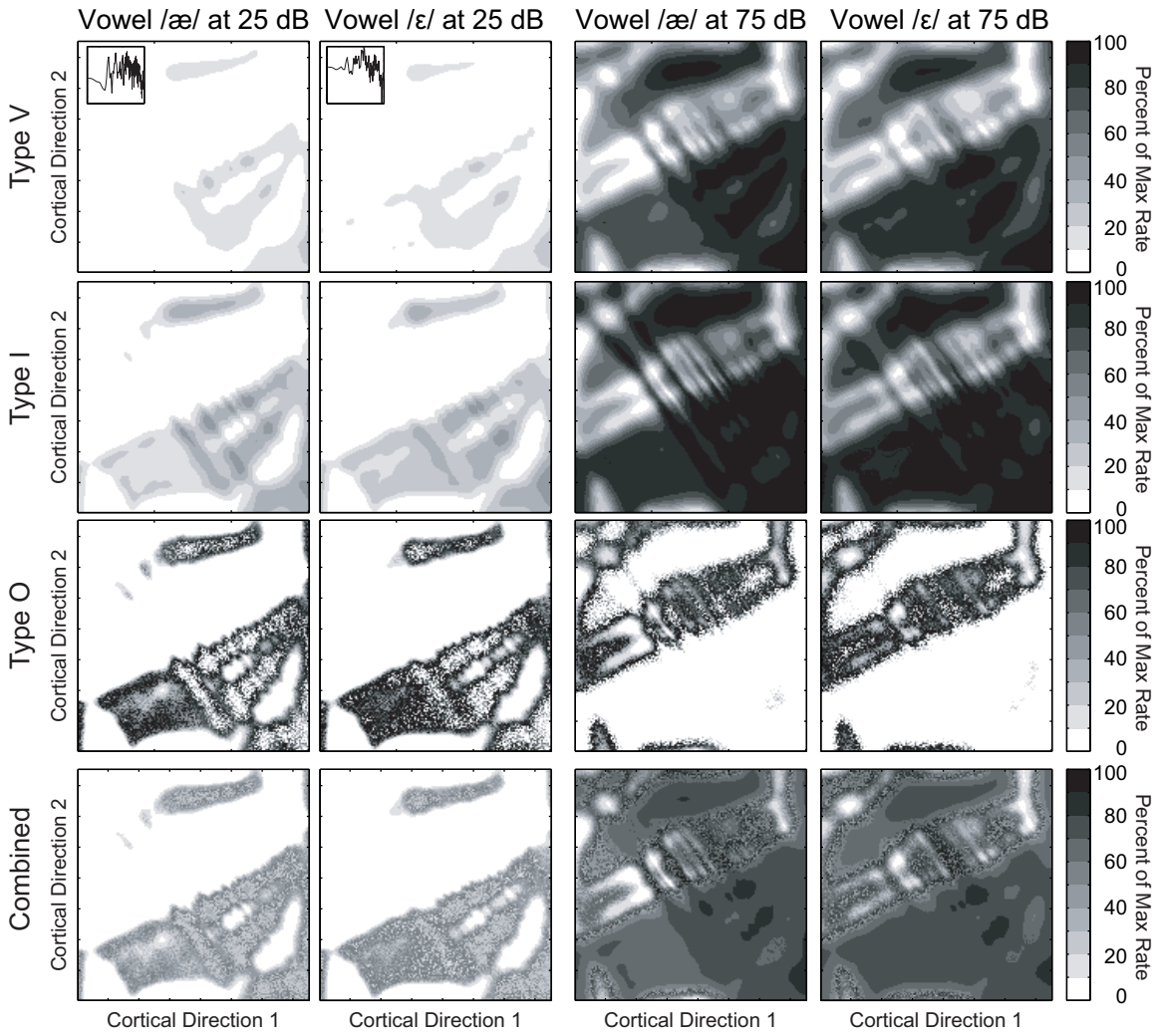

Figure 4.8. Virtual imaging of complex stimuli.

Complex stimuli such as vowels show fundamental response differences between Type V/I and Type O responses. Type V and I arrays showed responses that increased with vowel intensity until nearly the entire map was saturated with high responses. The area of activation in Type O arrays, on the other hand, was independent of the stimulus intensity, although the location of the activation migrated with changes to stimulus intensity. Type O responses can be seen to contribute substantially to the combined arrays at lower intensities, but at higher intensities Type V responses dominate. Vowel spectra are shown in the insets.

showed a similar magnitude of map differences between the two intensities, though largely nonoverlapping populations of neurons appeared to be responsible for this difference (Figure 4.9, third row). Because the magnitudes of the differences are greatest in the Type O case, the differential image created with the equal-proportions array most closely matched the type O responses (Figure 4.9, bottom row). Perhaps just as remarkably, Type V units, despite their extensive overall responses, influenced the combined population differential response very little at any intensity. This result comes about because the Type V units respond to more stimuli than the other response classes do (i.e., they are less selective), so differential images tend to cause more cancellation for Type V than for the other FRA classes. Differential functional images of A1 in response to complex spectra may therefore preferentially reveal Type O activation.

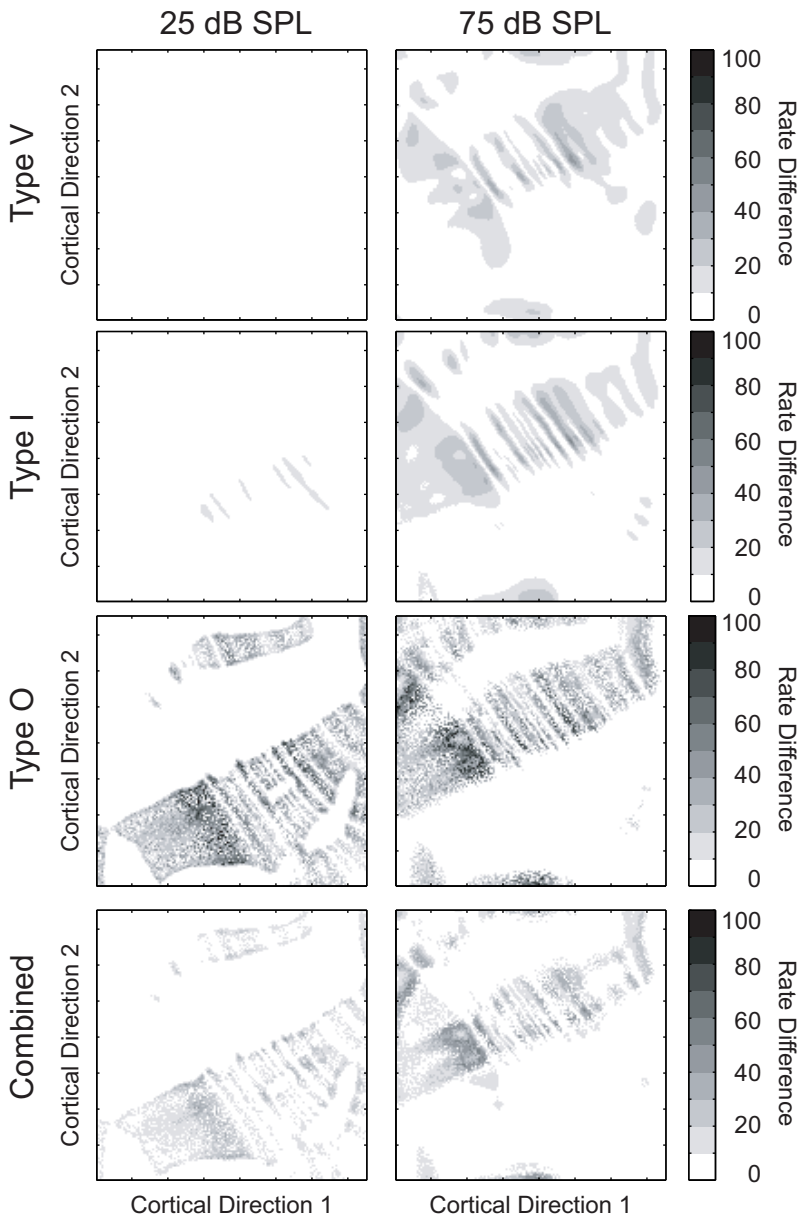

Figure 4.9. Differential maps of activity from complex stimuli.

Differential array activities for vowels at the same stimulus intensity. Type V/I responses indicate little difference between the two vowel activation patterns at low intensities and greater difference at higher intensities. The difference in activation areas for Type O responses appears to be fairly constant at both intensities, although the location of the active neurons shifts across the array. Because the magnitude of the rate differences is greatest in the Type O population, the combined response differences resemble the Type O response differences closely.

4.2 Discussion

Functional imaging studies allow the response properties of large portions of the brain to be analyzed in parallel. These results are achieved by examining combined behavior of many neurons located near one another performing similar functions. Limitations in these studies, however, may stem from technological considerations, such as poor spatiotemporal resolution, or from neuronal characteristics, such as lack of a topographic feature map. Functional imaging has traditionally been less successful at inferring novel neuronal behavior in the auditory system than in other sensory modalities. Some of the known and postulated functional map properties of primary auditory cortex (A1) were examined to determine which imaging limitations can be attributed to the neurophysiology and how these limitations might be overcome using modified image acquisition techniques.

The creation and refinement of Type O responses locally within primary auditory cortex (A1) by neurons with similar characteristic frequencies but different bandwidths and thresholds (Tan et al. 2007; Wehr and Zador 2003) implies that the mapping of these three neuronal features may be important for constructing local microcircuits in A1 that give rise to these responses. The results from a recent A1 optical imaging study that in turn matched classical electrophysiological findings (Ojima et al. 2005) were replicated by constructing theoretical maps of these three features. In that study large portions of A1 were activated with relatively few pure tone stimuli such that overlapping areas of activation were smallest in the center of A1 and greatest dorsally and ventrally. The virtual imaging experiments showed that the overlapping regions of activation observed in the model arrays when stimulated by octave-spaced tones were associated with high-bandwidth and low-threshold map regions. Electrophysiological bandwidth mapping studies in cat A1 indicate that high-bandwidth neurons are located towards the dorsal and ventral areas of A1 (Bonham et al. 2004; Schreiner and Sutter 1992). Threshold mapping studies have not yielded consistent maps as threshold seems to be mapped differently between animals (Schreiner et al. 1992). Because the computational feature maps were

selected based upon their gross similarities to the organization of those found in physiological mapping studies, at least in terms of frequency and bandwidth (Watkins and Barbour 2008), the model array reproduced the dumbbell shapes and overlapping activations seen in images from the Ojima studies.

As pure-tone stimuli of higher intensity are presented to the model arrays, the three response classes exhibit different behaviors that translate into different functional mappings. Type V arrays are the only ones that exhibit expansion along the tonotopic axis as a function of intensity, which would tend to lower effective frequency resolution of any imaging technique at higher intensities. When the three response types are combined into a single array, Type V responses dominate the resulting functional maps at most relative proportions of the three response types. The influence of Type V responses can be dramatically reduced by adopting a higher response criterion, however. This effect is achieved because the Type I and Type O elements responding at the highest rates are more compactly organized than the Type V responses.

The percentage of Type O, intensity-tuned, or "nonmonotonic" neurons isolated in electrophysiological studies of A1 has been reported to be as little as 8% in anesthized cat (Sutter and Schreiner 1995) to as high as 78% in one study of macaque monkeys (Pfingst and O'Connor 1981), and many values in between. The true relative proportion of Type O neurons detectable in A1 is likely to be dependent upon species, anesthetic state, neuronal sampling methodology, stimulus selection and possibly other factors. Physiological data from awake marmoset monkeys indicate nonmonotonic proportions in excess of 50% in A1 (Watkins and Barbour 2008). Given uncertainties regarding the true proportion of Type O neurons across different preparations, however, I elected to evaluate model arrays with various relative proportions of the three response types. It may be possible that simple analytic methods applied to functional images derived from tone stimulation of A1 will enable a more accurate estimate of the true proportion of nonmonotonic neurons. One such technique suggested by the virtual imaging results would be to increase the response criterion until responses to pure tones fracture into multiple disconnected islands along the isofrequency axis. The response criterion as a percentage of maximum

response should approximately equal the relative proportion of Type O neurons under such conditions (c.f., Figure 4.4).

While the properties of neurons within a cortical column are similar to one another, they are not identical and can in some cases exhibit considerable variability, even when a topographic map is known to exist (Abeles and Goldstein 1970; Atencio and Schreiner 2008; Imig and Adrian 1977; Phillips and Irvine 1981). A rather surprising finding is that even with considerable jitter in all three mapped features, the gross structure of A1 activation by a single tone remains relatively constant. Indeed, if the spatial resolution of an imaging modality is lower than the resolution of the array (or, presumably, the columnar spacing in cortex), even extensive feature jitter may not alter functional images substantially. The one likely exception to this observation is in regions of narrow bandwidth (i.e., small receptive field size). Activity in those regions appeared to dilute along the tonotopic axis and diminish in prominence as jitter increased. Thus, even if considerable variability exists in the functional map(s) around some "ideal" map, well-designed imaging experiments may be able to discern the ideal map structure more clearly than electrophysiological experiments, which could be misleading should the number of neurons sampled be too low in the face of map variability. Such a situation may have contributed to early observations from single-unit mapping studies that stimulus frequency is not mapped in A1 (Abeles and Goldstein 1970).

4.2.1 Functional Map Extractions

Pure-tone stimulus protocols were used effectively to extract estimated frequency and threshold maps from the model array that closely matched the actual feature topographies. Tone-based protocols elicited the most error for these two features in map regions of lowest bandwidth. Bandwidth itself proved to be an extremely challenging feature to extract with a reasonable number of tone stimuli. Mapping bandwidth requires high-resolution frequency sampling in order to resolve adequately the activity of low-bandwidth neurons, as well as high-resolution threshold mapping in order to attribute the estimated bandwidth to the proper intensity (e.g., 10 dB above threshold). Since the

bandwidth of Type V responses is a function of stimulus intensity, errors in estimated threshold maps will greatly increase the error in bandwidth estimation for these units in particular. Furthermore, even with a large number of stimuli (over 2000 pure tones for the estimated map in Figure 4.7B), the extracted maps of bandwidth were crude. While they revealed the overall shape and the organization of the actual map, the estimated maps had a relatively high degree of error compared with the frequency and threshold maps. Attempts to map bandwidth in A1 based upon functional imaging experiments have generally attempted simply to identify "high-" and "low-" bandwidth regions (Kayser et al. 2007). In the simulations, identifying high-bandwidth regions was relatively straightforward based upon overlapping areas of activation from tones of different frequencies. Low-bandwidth regions can then be identified as pixels that do not demonstrate overlap at a particular tone frequency spacing and intensity. Any method of mapping a distinction between "high-" and "low-" bandwidth regions likely represents a reasonable approach for determining the general structure of the A1 bandwidth map.

While Type V responses may diminish imaging resolution in general, they paradoxically appear to improve the extraction of frequency maps. Their large bandwidth in response to high sound intensity helps decrease the number of stimuli needed to resolve the frequency map for the entire A1. Because only the maximum firing rate of the fitted frequency response curve was used to assign a frequency to a given pixel, the saturation from the Type V units did not affect the ability to extract the CF. The increased receptive field size allowed the neuron to respond to a larger number of stimuli to provide more points on the FRC to create a better curve fit. This finding may be attributable to the knowledge of the FRC shapes in advance and may not be applicable under physiological conditions when Type V FRCs are not perfect Gaussians. While estimated map accuracy was consistent across all the relative proportions of response types that were tested, the absolute firing rates for the equal-proportion arrays were much higher than for the 10:10:80 Type V: Type I: Type O arrays. The physiological accuracy of estimated maps may vary from what was predicted, therefore, depending upon the noisiness of the neuron activity and the actual response type proportions.

4.2.2 Responses to Complex Stimuli

Vowel spectra are composed of many spectral components dispersed in frequency. Such stimuli could be expected to activate a relatively large portion of A1, and this was also true in the model array, particularly at higher intensities. Different vowel spectra delivered at the same intensity to arrays of Type V and Type I responses elicited very similar activation patterns. Patterns of activity in response to the same vowel spectrum delivered at different intensities, on the other hand, demonstrated substantially greater variability. Most of this variability across intensity can be attributed to the threshold map and indicates that neither Type V nor Type I neurons represent a intensity-invariant code for complex spectra.

Type O responses also exhibit considerably more similarity between activity patterns in response to different spectra than to the same spectrum at different intensities. Unlike with Type V and Type I responses, however, different populations of Type O units are active at different intensities. The resulting activation patterns shift across the threshold axis, which is roughly perpendicular to the tonotopic axis in the model. Creation of a intensity-invariant representation of complex sounds with Type O neurons also seems unlikely based upon these findings. A special contribution of Type O responses does become apparent in the differential response maps, however. At low intensities, most of the Type V/I units responding to one vowel respond to another at the same intensity. The consequence of this lack of selectivity can be seen the activation difference maps, which show little differential activity for these response types between two vowels at low intensity. The difference maps for Type O neurons are prominent across intensities, however, and the combined population maps are therefore dominated by Type O responses at the lowest intensities. This finding provides further evidence that Type O neurons may be the most useful sound encoders at lower intensities (Watkins and Barbour 2008). Actual functional imaging experiments of A1 during complex stimulus delivery may be able to shed additional light on this hypothesis.

4.2.3 Experimental Techniques Predicted to Improve Imaging Resolution

The simulations indicate that any stimulation paradigm preferentially emphasizing Type O neurons is likely to improve the resolution of functional images in A1. First, raising the response criterion for attributing pixels to a given frequency under pure-tone stimulation is one technique that preferentially emphasizes Type O neurons. Second, because the total number of Type O responses elicited as a function of stimulus intensity is relatively constant for the total number of Type V/I responses increases with stimulus intensity. Thus, delivering stimuli at lower intensities raises the relative proportion of Type O responses. Finally, the temporal response properties of Type O neurons appear to differ from Type V/I neurons (Watkins and Barbour 2008). Exploiting differential adaptation may allow the creation of stimuli designed in temporal blocks that preferentially activate Type O neurons at a particular point in time.

4.2.4 Limitations in Virtual Imaging

Several caveats exist regarding the physiological applicability of the virtual imaging model used for this study. First, the feature maps used here were based upon theory rather than specific findings within A1, although under appropriate constraints the theoretical maps match known topographies of A1 (Watkins et al. 2009). While the theoretical maps assume independence of different features mapped within A1, these features may not be mapped completely independently of one another (Cheung et al. 2001; Philibert et al. 2005), which may influence functional images. Furthermore, thresholds are not evenly distributed when the auditory system is adapted to silence and tend to be concentrated toward lower values throughout the auditory system (Watkins and Barbour 2008). Absolute thresholds are also dependent upon frequency because of the filtering effects of the outer and middle ears, leading to higher neuronal thresholds at lower and higher frequencies. These phenomena could affect the activated areas of the simulated functional images. However, the trends that were observed are robust under many conditions that were tested, and they are unlikely to be disrupted by minor variations in the underlying feature distributions. In any case, the main goal of this work has been to devise hypotheses about A1 organization that are testable with real imaging experiments. Such

experiments would provide more convincing tests of these hypotheses than additional model manipulation.

Although not explicitly taken into account in the computational models, real neurons have responses shaped by inhibition as well as excitation (Ojima and Murakami 2002; Tan et al. 2007; Wang et al. 2002; Wehr and Zador 2003), and inhibition has the potential to alter responses to complex stimuli in particular. One likely role of this inhibition is to create selective neuronal responses. Most of the virtual imaging experiments described here exploited the delivery of pure tones. The three types of FRAs described reflect canonical compact A1 receptive field behavior in response to pure tones. Neuronal responses to wideband stimuli exhibit more variety than is apparent with the three FRAs presented here. The responses most typically display increased selectivity, which is most likely created through inhibitory mechanisms. When activity from selective and nonselective neurons is averaged together, the resulting activity most closely resembles the nonselective neurons. This phenomenon was exhibited for combinations of Type V (nonselective) and Type O (selective) neurons in Figure 4.4. Furthermore, imaged activity likely reflects as much sub-threshold activity as supra-threshold (or more). Given that Type O neurons are created and refined by local circuits within A1, Type V, I and O neurons are likely to be intermixed physically in A1, leading to the relatively nonselective response behavior that were explicitly modeled.

Finally, the procedures described here represent idealized imaging experiments. Neuronal noise, imaging noise, motion artifacts, etc., were not introduced into the virtual images of this study. It is possible, though unlikely, that noise sources commonly encountered during functional imaging experiments may systematically bias the resulting images. If that were the case, then real functional images may not recapitulate features of the idealized experiments because of this noise. Instead of modeling noise, it was chosen to focus upon inferences that could be drawn regarding neuronal circuitry from clean images. The degree to which actual functional A1 images correspond with the predicted findings, given the inevitability of noise in the measurements, is best evaluated directly in physiological experiments. It is likely that the results of such experiments will allow

refinements of the virtual imaging protocol to extend the neural network hypotheses presented here.

Chapter 5 Conclusion

5.1 Summary of Experimental Results

This thesis has reported novel results in the evaluation of functional imaging as a mapping tool for the primary auditory cortex. Chapter 3 used virtual mapping to describe how functional imaging studies have the necessary sampling density to accurately resolve maps with high spatial frequencies. At the same sampling densities, single-unit studies can more accurately estimate the underlying maps than multi-unit studies, which average the responses of multiple units together. Even though functional imaging studies such as optical imaging of intrinsic signals and functional-MRI of blood oxygen level-dependent contrast sample their responses by averaging neural responses, these functional imaging studies have sufficiently high resolution to accurately estimate the maps.

Chapter 4 introduces a virtual imaging model that simulates the neural circuitry found in A1. This model is created from the three simplest response types seen in A1: Type O (nonmonotonic), Type I (narrowband monotonic), and Type V (wideband monotonic) responses. The model suggests that the Type V responses may diminish imaging resolution because their large receptive fields obscure the other two types of responses. While Type O responses may account for up to 78% of the neurons in the auditory cortex (Schreiner et al. 1992), a relatively small percentage of Type V responses can obscure the Type O responses. However, the Type V responses may actually improve the ability of functional imaging modalities to map the auditory cortex because their large receptive fields reduce the number of stimuli that are needed to activate and map A1. While the Type V responses may obscure the other response types, it may be possible to reduce the saturation by raising the response criterion, by reducing the intensity of the stimulus, and by using the difference in temporal adaptation to selectively activate Type O responses.

5.2 Functional Imaging as a Neurophysiology Tool

Most functional imaging studies average the activity of a number of neurons. This averaging has both advantages and disadvantages. In mapping studies, averaging may resist varability within the cortical columns. However, this averaging reduces the overall accuracy of the maps compared to single unit mapping performed at the same resolutions. The sampling density of functional imaging is typically much higher than the sampling density of electrophysiology experiments. This increased sampling density helps compensate in resolution lost from the averaging and should allow for functional imaging to accurately map A1.

The second advantage of mapping A1 with functional imaging is that the large receptive fields of Type V responses may increase the accuracy of the maps. The large receptive fields allow the brain to be mapped with fewer responses, at the expense of saturating out the responses from less selective encoders. The obscured responses of the Type O and Type I neurons, however, should not affect the ability to map A1. The organization of physiological properties in cortical columns reduces the need to record all the responses within a column if the average response reflects the distribution of neurons within a column.

5.3 Shortcomings of Functional Imaging

With the exception of multi-photon imaging with calcium indicators, functional imaging lacks the ability to record at single cell resolution and the ability to uniquely detect supra-threshold activity. Because of this, most functional imaging experiments may not be able to measure response from a large number of neurons in A1. Such neurons include Type O neurons, which can account for up to 70% of the isolated neurons in A1. Type O responses are created within the central nervous system since they are not observed within the auditory nerve fibers. Without the use of multiphoton imaging with calcium indicators to isolate the Type O responses from the saturation produced from Type V responses, functional imaging studies may have trouble replicating many electrophysiology studies, unless other novel techniques are developed.

5.4 Possible Future Studies

The next step is to test the techniques suggested in this thesis to overcome the limitations of functional imaging. These techniques may improve resolution by selectively activating Type O responses. The first method is to increase the response criterion of the image. Increasing the response criterion may require both the Type V and the Type O responses to fire together in order to reach the increased response criterion. The next method is to present stimulus at low sound intensities. Imaging studies typically use loud stimuli (>60 dB SPL) in order to improve the signal to noise ratio. In doing so, they preferentially activate Type V responses since Type O responses are inhibited at high sound intensities, and Type V responses have their highest firing rates at loud sound intensities. Another method is to use temporal adaptation to selectively activate Type O responses. In comparison to Type V and Type I responses, Type O responses slowly adapt to sound intensity. Thus, playing a loud sound before a quieter stimulus may cause the Type V and I responses to raise their thresholds, allowing the Type O responses to be selectively activated by the quieter stimulus.

References

Abeles M and Goldstein MH, Jr. Functional architecture in cat primary auditory cortex: columnar organization and organization according to depth. *J Neurophysiol* 33: 172-187, 1970.

Atencio CA and Schreiner CE. Spectrotemporal processing differences between auditory cortical fast-spiking and regular-spiking neurons. *J Neurosci* 28: 3897-3910, 2008.

Bonham BH, Cheung SW, Godey B, and Schreiner CE. Spatial organization of frequency response areas and rate/level functions in the developing AI. *J Neurophysiol* 91: 841-854, 2004.

Caspary DM, Backoff PM, Finlayson PG, and Palombi PS. Inhibitory inputs modulate discharge rate within frequency receptive fields of anteroventral cochlear nucleus neurons. *J Neurophysiol* 72: 2124-2133, 1994.

Cheung SW, Bedenbaugh PH, Nagarajan SS, and Schreiner CE. Functional organization of squirrel monkey primary auditory cortex: responses to pure tones. *J Neurophysiol* 85: 1732-1749, 2001.

Chklovskii D and Koulakov A. MAPS IN THE BRAIN: What Can We Learn from Them? *Annual Review of Neuroscience* 27: 369-392, 2004.

Das A and Gilbert CD. Long-range horizontal connections and their role in cortical reorganization revealed by optical recording of cat primary visual cortex. *Nature* 375: 780-784, 1995.

Farley BJ, Yu H, Jin DZ, and Sur M. Alteration of visual input results in a coordinated reorganization of multiple visual cortex maps. *J Neurosci* 27: 10299-10310, 2007.

Godey B, Atencio CA, Bonham BH, Schreiner CE, and Cheung SW. Functional organization of squirrel monkey primary auditory cortex: responses to frequency-modulation sweeps. *J Neurophysiol* 94: 1299-1311, 2005.

Grinvald A, Lieke EE, Frostig RD, and Hildesheim R. Cortical point-spread function and long-range lateral interactions revealed by real-time optical imaging of macaque monkey primary visual cortex. *J Neurosci* 14: 2545-2568, 1994.

Harel N, Mori N, Sawada S, Mount RJ, and Harrison RV. Three distinct auditory areas of cortex (AI, AII, and AAF) defined by optical imaging of intrinsic signals. *Neuroimage* 11: 302-312, 2000.

Harrison RV, Harel N, Kakigi A, Raveh E, and Mount RJ. Optical imaging of intrinsic signals in chinchilla auditory cortex. *Audiol Neurootol* 3: 214-223, 1998.

Hodgkin AL and Huxley AF. A quantitative description of membrane current and its application to conduction and excitation in nerve. 1952. *Bull Math Biol* 52: 25-71; discussion 25-23, 1990.

Imig TJ and Adrian HO. Binaural columns in the primary field (A1) of cat auditory cortex. *Brain Res* 138: 241-257, 1977.

Kaur S, Lazar R, and Metherate R. Intracortical pathways determine breadth of subthreshold frequency receptive fields in primary auditory cortex. *J Neurophysiol* 91: 2551-2567, 2004.

Kay KN, Naselaris T, Prenger RJ, and Gallant JL. Identifying natural images from human brain activity. *Nature* 452: 352-355, 2008.

Kayser C, Petkov CI, Augath M, and Logothetis NK. Functional imaging reveals visual modulation of specific fields in auditory cortex. *J Neurosci* 27: 1824-1835, 2007.

Kiang NYS, Watanabe T, Thomas EC, and Clark LF. *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve*. Cambridge, MA: The MIT Press, 1965.

Koulakov AA and Chklovskii DB. Orientation preference patterns in mammalian visual cortex: a wire length minimization approach. *Neuron* 29: 519-527, 2001.

Kozloski J, Hamzei-Sichani F, and Yuste R. Stereotyped position of local synaptic targets in neocortex. *Science* 293: 868-872, 2001.

Logothetis NK. What we can do and what we cannot do with fMRI. *Nature* 453: 869-878, 2008.

Logothetis NK, Pauls J, Augath M, Trinath T, and Oeltermann A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412: 150-157, 2001.

Merzenich MM, Knight PL, and Roth GL. Cochleotopic organization of primary auditory cortex in the cat. *Brain Res* 63: 343-346, 1973.

Merzenich MM, Knight PL, and Roth GL. Representation of cochlea within primary auditory cortex in the cat. *J Neurophysiol* 38: 231-249., 1975.

Obermayer K and Blasdel GG. Geometry of orientation and ocular dominance columns in monkey striate cortex. *J Neurosci* 13: 4114-4129, 1993.

Obermayer K, Blasdel GG, and Schulten K. Statistical-mechanical analysis of self-organization and pattern formation during the development of visual maps. *Phys Rev A* 45: 7568-7589, 1992.

Ojima H and Murakami K. Intracellular characterization of suppressive responses in supragranular pyramidal neurons of cat primary auditory cortex in vivo. *Cereb Cortex* 12: 1079-1091, 2002.

Ojima H, Takayanagi M, Potapov D, and Homma R. Isofrequency band-like zones of activation revealed by optical imaging of intrinsic signals in the cat primary auditory cortex. *Cereb Cortex* 15: 1497-1509, 2005.

Palmer AR and Evans EF. Cochlear fibre rate--intensity functions: no evidence for basilar membrane nonlinearities. *Hear Res* 2: 319-326, 1980.

Petkov CI, Kayser C, Augath M, and Logothetis NK. Functional imaging reveals numerous fields in the monkey auditory cortex. *PLoS Biol* 4: e215, 2006.

Pfingst BE and O'Connor TA. Characteristics of neurons in auditory cortex of monkeys performing a simple auditory task. *J Neurophysiol* 45: 16-34, 1981.

Philibert B, Beitel RE, Nagarajan SS, Bonham BH, Schreiner CE, and Cheung SW. Functional organization and hemispheric comparison of primary auditory cortex in the common marmoset (Callithrix jacchus). *J Comp Neurol* 487: 391-406, 2005.

Phillips DP and Hall SE. Spike-rate intensity functions of cat cortical neurons studied with combined tone-noise stimuli. *J Acoust Soc Am* 80: 177-187, 1986.

Phillips DP and Irvine DR. Responses of single neurons in physiologically defined primary auditory cortex (AI) of the cat: frequency tuning and responses to intensity. *J Neurophysiol* 45: 48-58., 1981.

Ramachandran R, Davis KA, and May BJ. Single-unit responses in the inferior colliculus of decerebrate cats. I. Classification based on frequency response maps. *J Neurophysiol* 82: 152-163, 1999.

Read HL, Winer JA, and Schreiner CE. Modular organization of intrinsic connections associated with spectral tuning in cat auditory cortex. *Proc Natl Acad Sci U S A* 98: 8042-8047, 2001.

Recanzone GH, Guard DC, and Phan ML. Frequency and intensity response properties of single neurons in the auditory cortex of the behaving macaque monkey. *J Neurophysiol* 83: 2315-2331, 2000.

Recanzone GH, Schreiner CE, Sutter ML, Beitel RE, and Merzenich MM. Functional organization of spectral receptive fields in the primary auditory cortex of the owl monkey. *J Comp Neurol* 415: 460-481, 1999.

Sachs MB and Abbas PJ. Rate versus level functions for auditory-nerve fibers in cats: tone-burst stimuli. *J Acoust Soc Am* 56: 1835-1847, 1974.

Sadagopan S and Wang X. Level invariant representation of sounds by populations of neurons in primary auditory cortex. *J Neurosci* 28: 3415-3426, 2008.

Schreiner CE and Mendelson JR. Functional topography of cat primary auditory cortex: distribution of integrated excitation. *J Neurophysiol* 64: 1442-1459, 1990.

Schreiner CE, Mendelson JR, and Sutter ML. Functional topography of cat primary auditory cortex: representation of tone intensity. *Exp Brain Res* 92: 105-122, 1992.

Schreiner CE and Sutter ML. Topography of excitatory bandwidth in cat primary auditory cortex: single-neuron versus multiple-neuron recordings. *J Neurophysiol* 68: 1487-1502, 1992.

Sheth S, Nemoto M, Guiou M, Walker M, Pouratian N, and Toga AW. Evaluation of coupling between optical intrinsic signals and neuronal activity in rat somatosensory cortex. *Neuroimage* 19: 884-894, 2003.

Sheth SA, Nemoto M, Guiou M, Walker M, Pouratian N, and Toga AW. Linear and nonlinear relationships between neuronal activity, oxygen metabolism, and hemodynamic responses. *Neuron* 42: 347-355, 2004.

Spitzer MW, Calford MB, Clarey JC, Pettigrew JD, and Roe AW. Spontaneous and stimulus-evoked intrinsic optical signals in primary auditory cortex of the cat. *J Neurophysiol* 85: 1283-1298, 2001.

Sutter ML. Shapes and level tolerances of frequency tuning curves in primary auditory cortex: quantitative measures and population codes. *J Neurophysiol* 84: 1012-1025, 2000.

Sutter ML and Schreiner CE. Physiology and topography of neurons with multipeaked tuning curves in cat primary auditory cortex. *J Neurophysiol* 65: 1207-1226, 1991.

Sutter ML and Schreiner CE. Topography of intensity tuning in cat primary auditory cortex: single-neuron versus multiple-neuron recordings. *J Neurophysiol* 73: 190-204, 1995.

Tan AY, Atencio CA, Polley DB, Merzenich MM, and Schreiner CE. Unbalanced synaptic inhibition can create intensity-tuned auditory cortex neurons. *Neuroscience* 146: 449-462, 2007.

Wang J, McFadden SL, Caspary D, and Salvi R. Gamma-aminobutyric acid circuits shape response properties of auditory cortex neurons. *Brain Res* 944: 219-231, 2002.

Watkins PV and Barbour DL. Specialized neuronal adaptation for preserving input sensitivity. *Nat Neurosci* 11: 1259-1261, 2008.

Watkins PV, Chen TL, and Barbour DL. A computational framework for topographies of cortical areas. *Biol Cybern* 100: 231-248, 2009.

Wehr M and Zador AM. Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature* 426: 442-446, 2003.

Yu H, Farley BJ, Jin DZ, and Sur M. The coordinated mapping of visual space and response features in visual cortex. *Neuron* 47: 267-280, 2005.

Vita

Thomas Leewei Chen

Date of Birth	December 29, 1985
Place of Birth	San Francisco, CA
Degrees	B.S., Biomedical Engineering, May 2008 B.S., Applied Sciences in Chemical Engineering, May 2008
Professional Societies	Biomedical Engineering Society The Association for Research in Otolaryngology
Publications	Watkins PV, Chen TL, and Barbour DL. A computational framework for topographies of cortical areas. *Biol Cybern* 100: 231-248, 2009.

Aug 2009

Analysis of Functional Imaging, Chen, M.S. 2009