Using Information Theory to Understand Neural Representation in the Auditory Cortex

Hannah (Qiaochu) Cui

A thesis presented for the degree of
Bachelor of Science

To the Department of Mathematics
and the Robert D. Clark Honors College
University of Oregon
May 2022

An Abstract of the Thesis of

Hannah(Qiaochu) Cui for the degree of Bachelor of Science
in the Department of Mathematics to be taken May 2022

Title: Using Information Theory to Understand Neural Representation

Approved: ____
James Murray

Neurons in the brain face the challenge of representing sensory stimuli in a way that accurately encodes the features of these stimuli while minimizing the effects of noise. This thesis will use the concept of mutual information from information theory, which quantifies the amount of information one variable can tell us about another and vice versa, to better understand neural coding in the auditory cortex. Previous research has been done in maximizing mutual information to better understand neural behavior patterns in the visual cortex, with limited auditory findings. We perform numerical optimization in Python to maximize information that a population of neurons contains about an auditory stimulus within the framework of information theory. This is done by first finding the optimal width and location of tuning curves that characterize neural response to one dimensional stimuli (sound frequency), then updating the optimization algorithm to fit two-dimensional stimuli (sound frequency and intensity). By testing the algorithm with a set of natural sound data, our computations show that in the latter case, optimal stimulus information is represented by a specific homogeneous population with similar response properties. Our findings provide a method to better understand neural representation in the auditory cortex, specifically, the relationship between neural response and natural sound stimuli.

Keywords: neural coding, mutual information, auditory cortex, optimization

Acknowledgments

I would like to thank Dr. James Murray for his wonderful guidance and support throughout my two years of research journey at the University of Oregon. As one of few applied mathematicians in the department, I was grateful to find my direction and future path from being a part of the Murray lab. In addition, I express my gratitude to Dr. Christian Schmid, a Postdoctoral scholar in the Murray lab.

Starting from scratch heading into the field of computational neuroscience, I would be lost if it weren't for Dr. Murray's detailed instructions on Mathematics and Neuroscience concepts, and Dr. Schmid's clear explanations, patiently working with me on my endless coding questions...

Almost paving the path for applied mathematics research at the UO, these mentors helped me step by step in getting to where I am today — finishing an honors thesis and attending a PhD program in applied mathematics in the fall. I want to also extend my gratitude to Dr. Chris Sinclair and Dr. Elizabeth Raisanen, for agreeing to be on my thesis defense committee and providing me with helpful feedback and advice on different perspectives of this research topic.

I want to also give a special shoutout to my friends and family for being there for me throughout this process, pushing me to do my best ever day. From all the late nights to early mornings, from words of encouragement when things get tough, and acts of celebration when small accomplishments are made, I couldn't have done this without everyone's support.

Finally, We did it!

Table of Contents

List of Figures

1 Introduction and Background

One phenomenon in the brain many researchers are interested in understanding is neurons' ability to accurately encode information about the world. This proves to be difficult as neurons are noisy, meaning that when the same stimulus is presented, the neurons do not produce the same response each time. A goal in neuroscience, specifically in the study of how the brain represents sensory information, is to be able to construct models that best explain neural response to sensory stimuli while representing the relationship between the two.

This research is in theoretical and computational neuroscience, which means there will not be elements commonly seen in scientific research, such as conducting experiments, producing the data, and analyzing the data to reach a conclusion. In this thesis, we will use theories, mathematical concepts, formulas, and computer programming to better understand phenomena that have been observed previously in experiments by other researchers. In particular, we will focus on the mathematical description of tuning curves in auditory cortex, where a tuning curve characterizes a neuron's response to different stimuli. The below sections provide important background information and motivation for our study of the auditory cortex, for our research method of mutual information maximization, for our analysis of neural response represented by tuning curves, which help us to answer this work's main research question: how auditory stimuli information is optimally represented in the brain.

1.1 Auditory Cortex

Definition (Auditory Cortex)**.** *The auditory cortex* is a network of areas in the part of the brain that receives inputs from the subcortical auditory pathways in the brainstem and thalamus. Through an elaborate network of intrinsic and extrinsic connections, the auditory cortex is thought to bring about the conscious perception of sound and provide a basis for the comprehension and production of meaningful utterances[8].

There has been much research done in understanding how auditory signals undergo the chain of processing that begins at the ear, through a series of complex subcortical brain regions, then ultimately reaches the auditory cortex[10]; however, the specifics of how neurons in the auditory cortex respond to sound stimuli, specifically natural sounds, is far from understood. This is because of the unique organization of the Auditory cortex— with a longer subcortical pathway between the peripheral receptors compared with other sensory cortices. This suggests that the auditory system may implement different strategies in contrast to other cortices in order to extract behaviorally relevant information from an environment of natural stimuli [21]. In addition, it has been shown that, as opposed to other cortices, auditory neurons have the property of high complexity in their selectivity to sound features. This is crucial when trying to understand neural response to natural sounds. Therefore, analyzing auditory cortical processing is an important step in understanding many overall brain functions such as decision making and learning [10].

Why are we studying neural responses to natural sound stimuli? This is because the structure of natural sounds consists of various and complex temporal patterns of acoustic energy and sound levels extending over a wide range of frequency bands [14]. Therefore, natural sound stimuli are interesting because they are rich and complex in important ways that simple stimuli traditionally used in laboratory studies are not. Since laboratory stimuli are not representative of such complexities encoded in natural stimuli, we are interested in using natural sounds in order to produce more biologically plausible results, meaning how sounds in the real world affect neural response in the auditory cortex. Moreover, studies in the visual cortex, a different brain region, have shown that models derived from natural stimuli are the most robust at predicting responses to other broadband stimulus ensembles that were not used in their estimation and also provide good predictions of tuning curves [18]. This is likely to be applicable to the auditory cortex as well.

In pursuit of finding strategies to represent how the auditory cortex encodes information, researchers have come up with many encoding strategies. One approach is by using the spectrotemporal receptive field (STRF)[10]. The STRF is a computational tool for characterizing the responses of auditory neurons. Most widely used as part of a linear-nonlinear (LN) model, the STRF model is reliable as it chooses random structured stimuli with relatively small amounts of data. Although this method provides a set of tools to represent a single neuron or a population of neurons and describe neural responses, STRFs do not accurately capture the full complexity of the behavior of auditory neurons [19]; this means the model results display too much discrepancy between synthetic/laboratory stimuli and natural sounds. Additionally, researchers have seen that neurons in the auditory cortex are classically thought to primarily encode auditory frequency [20]. Although most neurons are responsive to a singular frequency, a portion do respond to multiple frequencies, and research has shown that the relationship between these frequencies is harmonic [20]. Harmonic template neurons [5] have an important role in processing sounds with harmonic structures, especially those seen in nature. Responses of these neurons show nonlinear facilitation to harmonic complex sounds over inharmonic sounds and selectivity for particular harmonic structures [5]. The existence of these harmonic template neurons further suggests there is a complex harmonic processing organization in the auditory cortex[20]; however, it is unclear of exactly where this stage of processing occurs [5]. Neither model fully captures all of the ways that neurons respond to natural-sound stimuli. Therefore, it is important for us to construct a model that can optimally represent natural auditory stimuli.

1.2 Tuning Curves

Definition (Tuning curve)**.** A *tuning curve* characterizes the response of a neuron. A simple way of representing this is to find the average firing rate r, which is the number of action potentials fired over (in theory, an infinite number of) trials and divide by the trial duration. The average firing rate can be written as a function, $r = f(s)$, where s is a stimulus parameter (e.g. sound frequency), is called the neural response *tuning curve* [4].

The above definition is saying that neurons communicate with other neurons by sending brief electro-chemical pulses known as action potentials or spikes to their neighbors. The number of spikes generated in a given period of time is thought to be the means by which neurons encode information about the world. Tuning curves describe a neuron's response to a stimulus in spikes per second. A set of tuning curves can define a neural code for a population of neurons, which can be understood as the relationship between responses and sensory neurons to sensory stimuli such as the frequency of an auditory tone. As discussed by Kriegeskorte [12], a "tuned" neuron may selectively respond to stimuli within a particular band of some stimulus variable, such as frequency or intensity, and the tuning curves are a quantitative measure of neural responses as a function of a stimulus variable. An example of such 1D tuning curve can be seen in Figure 1

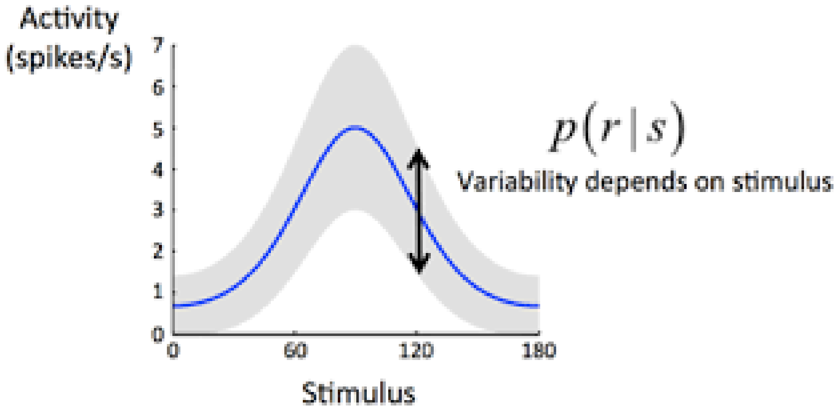

Figure 1: Sample tuning curve (blue) shows mean spikes per second (firing rate) as a function of the stimulus parameter. The shaded gray region illustrates the variability in firing rate upon the administration of repeated, identical trials, also known as *noise*.[11]

A phenomenon that inspired this project is the paper by Sadagopan and Wang [16], in which the authors investigated neural response characteristics in the primary auditory cortex region of awake marmosets. In the Sadagopan and Wang study, two different types of tuning representations were observed: "V-Shaped" and "O-Shaped" Tuning Curves, as seen in Figure 2 [16].

Typically, a neuron's firing rate response increases as sound intensity level increases until saturation. This is called a monotonic rate level function. Conversely, when neurons exhibit a maximum response rate followed by a decrease as sound level increases, this is a nonmonotonic rate level function. Wang and Sadagopan found that neurons with “V” and “I”-shaped tuning curves span across a higher frequency range as sound level increases. This covariance makes it difficult to predict frequency and sound level simply based on firing rate. On the other hand, “O”-shaped tuning curves spanning across a certain frequency range maintain their width at different sound levels, which makes it easy to parse frequency and sound level based on neural responses. A neuron with “O”-shaped turning curves is a non-monotonic rate level function. The idea of modeling frequency tuning in respect to sound level (intensities) is one of the main goals of this project. The hypothesis that this thesis investigates is that optimal stimulus information is represented by a mixed population of these “O”-shaped and “V”-shaped neurons that respond in qualitatively different ways to auditory stimulus features, and not a homogeneous distribution of either all “O”-shaped or all “V”-shaped neurons. We came to this hypothesis because of neural heterogeneity [9]: only a small fraction of neurons in a given population carry significant sensory information in a specific context. Therefore we suspect that optimal tuning occurs with a combination of “V” and “O”-shaped tuning curves.

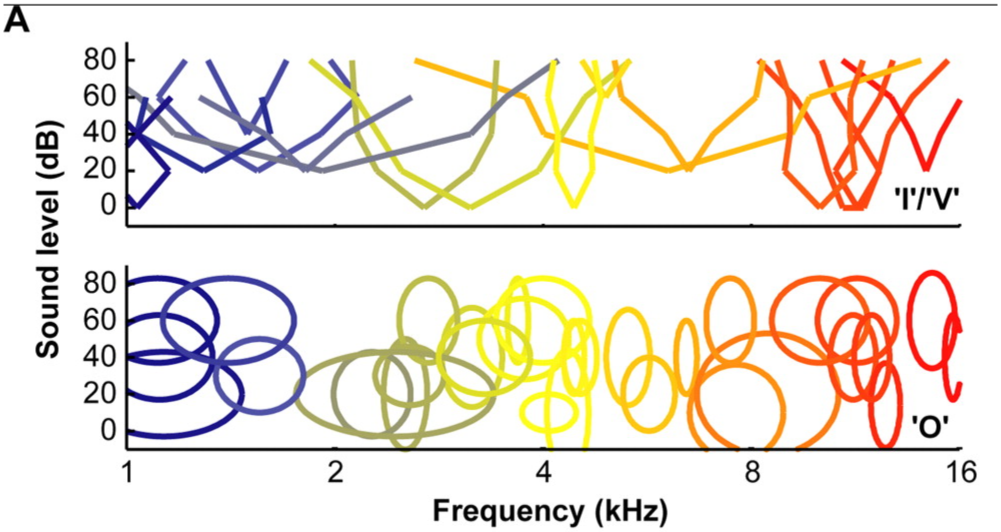

Figure 2: ”O”(bottom) and ”V”(top) shaped tuning curves, Each shape denotes the combination of frequency and sound level that causes a particular neuron to respond with an increased firing rate [16]

1.3 Information Theory

Definition (Information Theory)**.** *Shannon's information theory* is the mathematical treatment of the concepts, parameters and rules governing the transmission of messages through communication systems [13]. It is a means of quantifying information, choice and uncertainty. Additionally, it is a methodology to develop efficient coding and communication of data across noisy channels [3].

Information theory was first introduced to the world by Claude Shannon with his seminal paper "A mathematical theory of communication [17]". As Dayan and Abbott [4] state, a main motivation for using information theory in neuroscience, is to answer the question "How much does the neural response tell us about a stimulus?" The techniques of information theory allow us to answer this question in a quantitative manner. More importantly, information theory provides a framework that quantifies how a coding scheme or a communication channel conveys information and enables us to understand the relationships and interactions between arbitrary multivariate random variables [3]. Besides information theory, there have been many approaches and methodologies to analyzing and understanding neural response behavior to capture specific features of neural activity such as connectivity, effective dimensionality, encoding, and decoding [3].

Information theory has fruitfully been applied in many domains, including neuroscience, statistics, and economics [3]. As mentioned by Borst [2], the rise in popularity of information theory, can be attributed to the theory's rigor, which enables us to precisely measure information transfer by determining the exact probability distribution of outputs (in our case, neural response) given any particular signal or input (in our case, auditory stimuli). Moreover, because of its mathematical completeness, information theory has fundamental theorems on the maximum information transferrable in a particular communication channel.

The key idea from information theory that will be used in this thesis is that information encoded by one variable (the neural response) can be quantified by the reduction of its entropy when we learn about the other variable (the auditory stimulus) [12]. The mutual information(MI) between neural response and stimulus provides a means of quantifying this relationship. We explain mathematically in further details below what mutual information is. In order to understand MI, we first introduce entropy.

Definition (Entropy)**.** The *entropy* of a neural response with a firing rate represented by random variable r and a probability density $p(r)$ of observing that firing rate is a measure of our uncertainty about the variable.[4]

There are three probability distributions associated with entropy and mutual information. First, the probability of a firing rate happening as seen in the above definition. Then, we are interested in the probability of a certain stimulus happening given a set of stimulus values, also known as $p(s)$. Lastly, we want to know the likelihood or probability of observing a neural firing rate r given a stimulus s, denoted $p(r|s)$

In other words, entropy represented by $H(r)$[4] in the below equation, is how "surprised" we are to be seeing that response rate r, then averaging over all possible responses.

$$H(r) = -\int p(r) \log p(r) dr. \tag{1}$$

Conditional entropy describes how "surprised" we are to be seeing that same firing rate r given information about the stimulus parameter s. It is represented by [12]:

$$H(r|s) = -\int p(r|s) \log p(r|s) drds. \tag{2}$$

Finally, mutual information $I(r, s)$ measures the amount of information one variable r can tell us about the other s (and vice versa). We can first think of it as the reduction of entropy in one variable after learning about the other variable, denoted by [4]:

$$I(r, s) = H(r) - H(r|s) = H(s) - H(s|r). \tag{3}$$

Using Equations 1 and 2, the final formula becomes [4]:

$$\int\int p(r,s)log\frac{p(r,s)}{p(r)p(s)}drds. \tag{4}$$

In the case where variables are discrete,

$$I(r,s) = \sum_{s,r} p(s)p(r|s)log\left(\frac{p(r|s)}{p(r)}\right). \tag{5}$$

1.4 Putting Everything Together

This project revolves around this relationship between natural sound stimuli and neural response, and maximizing their mutual information so that we can learn more about neural response if there is a strong relationship between that and natural stimuli. How do we maximize MI? Note that $p(s)$, the probability distribution of the stimuli (in our case frequency and intensity), varies depending on the sound environment (hence the importance of using natural sound stimuli). The logic of maximizing MI comes from minimizing uncertainty, which means the model uses optimization techniques (further explained in later sections) to increase the accuracy of predicting how a neuron responds given a set of stimuli. We want to know through these mutual information computations whether neural response in the auditory cortex matches the results of our optimizations. This idea sits on the basis of the efficient coding hypothesis[1], which states that sensory systems such as the auditory system maximizes the amount of information transferred from the environment to the brain. Modeling a population of noisy neurons characterized by their tuning-curve responses to a sensory variable, Ganguli and Simoncelli mathematically solved the information maximization problem to derive optimal tuning curves as a function of prior probability distribution $p(s)$ of the sensory variable [7]. Assuming tuning curve widths vary inversely with cell density, the specific findings from this study include the following conclusion: for the optimal population, the cell density is proportional to $p(s)$, meaning that more cells with narrower tuning curves are allocated to represent stimuli with a higher prior probability density[7][6]. This is one of the results we will test with our optimization algorithm in Chapter 2.

We first observe how one-dimensional tuning curve characteristics such as width and location change as a function of stimulus (frequency) probability by maximizing mutual information. Then after quantifying this information, we investigate whether the distribution of neural response tuning curves match our hypothesis: more neurons should be dedicated to represent narrow tuning curves at stimuli with high probabilities, while wide tuning curves should present to be more sparse at low probability stimuli. If the resulting tuning curve distribution after MI maximization accurately encode $p(s)$, then we can conclude our algorithm indeed finds the optimal neuron representation in the auditory cortex.

Next, by updating functions to perform optimization on two-dimensional stimuli, we attempt to explain the existence of both “V-shaped” and “O-shaped” tuning curves with respect to sound frequency and intensity. Based on this idea, our hypothesis is that a combination of these particular shapes of tuning curves will maximize the mutual information with natural stimuli, rather than one homogeneous type of neural response tuning curves. Moreover, another question we will be exploring is how the orientations of these two-dimensional tuning curves contribute to the maximization of mutual information. The research outcome is to predict what neural tuning curves should look like for different stimulus distributions under the assumption that their role is to maximize MI between the stimulus and neural response, then to look for correspondences with experimentally observed tuning-curve properties. A potential impact of this study is to embed and connect these research results to how neural representations of stimuli change with learning, for example in mice trained to respond to a particular auditory frequency. Further studies can be done to explore whether more neurons should be devoted to representing stimuli that are important for performing a learned behavior. In the next chapter, we investigate optimal one-dimensional tuning curves and what they tell us about neural response to natural auditory stimuli in the auditory cortex.

2 Maximizing Mutual Information: one-dimensional stimuli

One of our research outcomes is to model neural tuning curves (neural response patterns) and how their properties should differ for different prior stimulus distributions. Note all computations and modeling were done in NumPy (Python library for scientific computing). The data used in our computations and analysis is extracted from Montana State Library digital collections — the Acoustic Atlas, ambient recordings[15]. This collection consists of 2,500 recordings of species and environments from throughout the Western United States. For the purposes of our work, these ambient sounds constitute a representative ensemble of natural stimuli.

The first part of our one-dimensional stimulus parameter optimization, sound frequency, uses the first 150 seconds of one specific sound-track: Crickets at Beaver Dam Wash (Utah) for the purpose of efficient modeling. The cricket sound track is then plotted with a spectrogram, which displays sound intensity at different frequencies at each moment in time, as demonstrated in Figure 3. Next, we convert the spectrogram (with frequencies on the y-axis) to a power spectral density graph seen in Figure 4 (with mel-frequencies on the x-axis), which is the measure of a signal's power content. The mel scale is a logarithmic frequency conversion that allows auditory signals to be represented in a manner that conforms with human auditory perception. The formula we use for this mel scale conversion is the following:

$$m = 2595 \log_{10}\left(1 + \frac{f}{700}\right), \tag{6}$$

where f represents the list of original frequencies and m represents the converted frequencies. By making our stimuli into a more biologically plausible representation, we are able to see more clearly at which frequencies the stimuli display the most intensity (about 200Hz and 2200Hz in the cricket data).

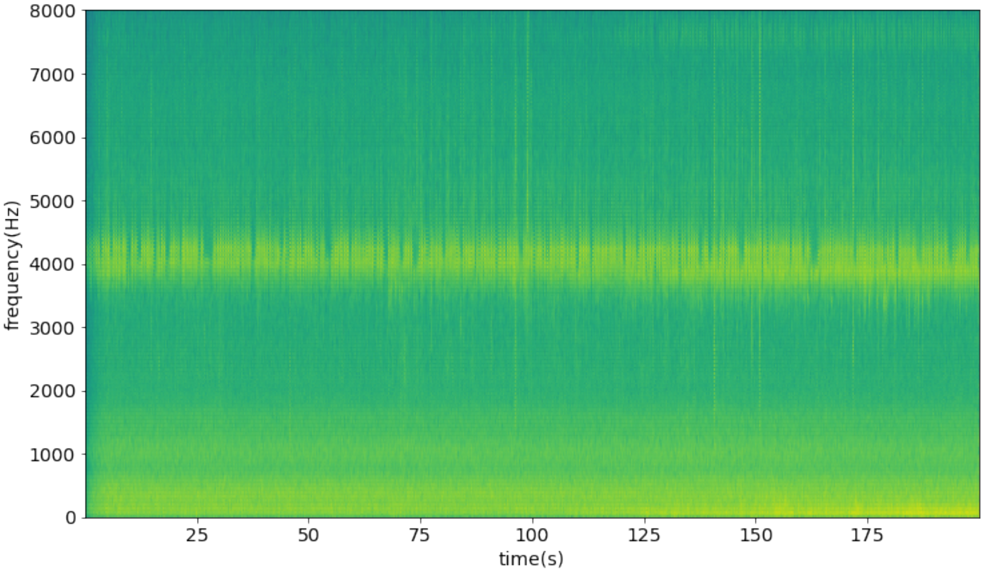

Figure 3: Spectrogram of natural crickets' soundtrack (Montana State University library digital collections)

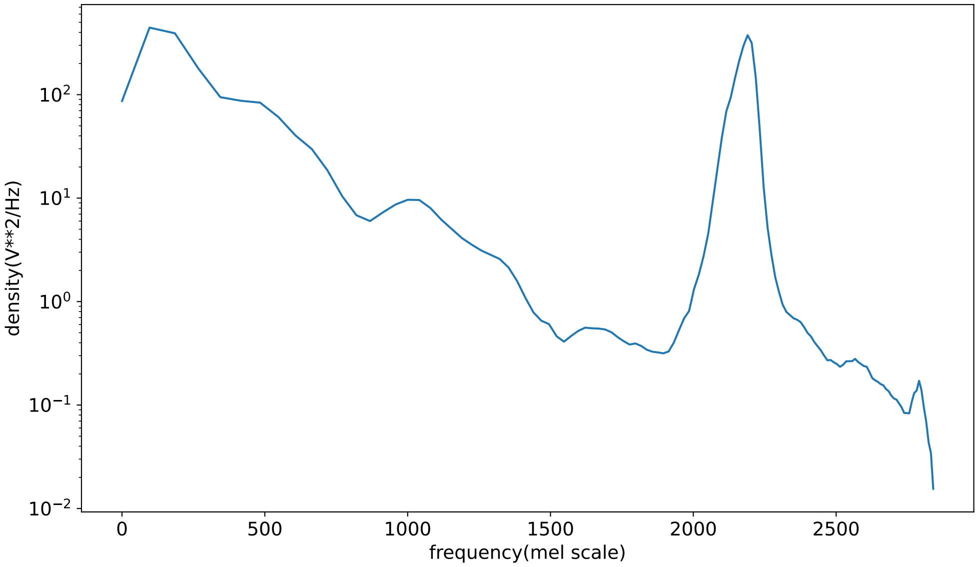

Figure 4: Power Spectral Density diagram of Spectrogram (x-axis is mel scale of frequencies)

To construct our model, given that the formula for mutual information (MI) is

$$I(r,s) = \sum_{s,r} p(s)p(r|s)\log\left(\frac{p(r|s)}{p(r)}\right), \tag{7}$$

we need to first find the stimulus probabilities $p(s)$. Since most auditory stimuli are not uniformly distributed across the frequencies, we utilized the method of binning discrete frequencies to construct a non-uniform prior stimulus distribution. This allows us to generate random stimuli according to the probabilities. The one-dimensional stimuli we use in our equations below is a vector of $10,000$ random frequency stimuli $\mathbf{s}$, based on the distribution of $p(s)$. Next, We need to find the probability distribution of neural firing rates (neural response) $p(\mathbf{r}|\mathbf{s})$ given the stimulus distribution $p(\mathbf{s})$, as well as the probability distribution of firing rates $p(\mathbf{r})$. It's important to note that since we are interested in the behavior of a population of neurons, we initialize our computations with 20 hypothetical neurons. To do this, we first modeled tuning curves prior to optimization as a Gaussian distribution with probability density function:

$$f(s) = c * \exp\left(-\frac{1}{2}\left(\frac{\boldsymbol{s}-\boldsymbol{\mu}}{\boldsymbol{\sigma}}\right)^2\right). \tag{8}$$

Here, $f(\mathbf{s})$ represents the average number of spikes per second that the 20 neurons fire in response to stimuli $\mathbf{s}$. c is a coefficient that gives the maximum firing rate of the neurons (we set this constant to 10, meaning the maximum firing rate of the neuron cannot exceed 10). $\sigma = \{\sigma_1, ..., \sigma_{20}\}$ represents a vector of 20 tuning curve widths, $s = \{s_1, ..., s_{10000}\}$ represents random stimulus frequencies, and $\mu = \{\mu_1, ..., \mu_{20}\}$ represents a vector of 20 stimulus values at which the responses is maximized, or as we like to call it — location of tuning curves. We set our 20 neurons to have equal widths and locations evenly distributed across the stimuli range, thereby each neuron has its own parameters μ_i and σ_i.

Then, the probability of a firing rate given a certain stimulus is modeled with a Poisson distribution with mean given by Equation 8. The Poisson distribution is a probability distribution over discrete events and is often used for modeling spike count distributions in neuroscience [4]. The probability that the neuron emit r spikes in response to stimulus s is then given by:

$$p(\boldsymbol{r}|\boldsymbol{s}) = \frac{f(\boldsymbol{s})^{\boldsymbol{r}} \exp(-f(\boldsymbol{s}))}{\boldsymbol{r}!}. \tag{9}$$

Lastly, we get the probability function of firing rates:

$$p(\boldsymbol{r}) = \frac{1}{n} \sum_{s_1,\ldots,s_n} (p(\boldsymbol{r}|\boldsymbol{s})), \tag{10}$$

by taking the mean of the $p(r|s)$ functions over the stimulus parameter, obtaining the probability of a firing rate averaged over stimuli for each neuron.

After putting together the MI function with the above probability components, the total MI that we are optimizing in Equation 7 includes a sum over our 20 neurons. We use automatic differentiation in JAX NumPy, a Python library designed for high-performance numerical computing especially in machine learning research, to take the partial derivatives of the MI with respect to the parameters that we wish to optimize — in our case, the width σ_i and location μ_i of each tuning curve i. In order to iteratively update the parameters to maximize the MI, these partial derivatives are then used to perform gradient ascent.

Definition (Gradient Ascent)**.** Given an objective function Q, and a set of parameters, θ, gradient-based approaches involve estimating the direction in parameter space to move so as to improve the model performance. This is achieved by estimating the derivative of the objective function with respect to the parameters and iteratively updating the parameters along the gradient, leading to an increase in the value of the objective function. Consequently, this requires that the objective function for *gradient based optimization* is differentiable. Typically, estimating the true gradient of the objective function is intractable. The gradient is often estimated using a random sample of training data leading to stochastic estimates of the true gradient, hence called *stochastic gradient-ascent*. [4]

Using this idea, the formula we use for gradient ascent is:

$$\theta \leftarrow \theta + \alpha \frac{\partial}{\partial \theta} I(\theta), \tag{11}$$

where θ is the parameter vector we are trying to optimize, $\frac{\partial}{\partial \theta} I(\theta)$ is the gradient of MI, and α is the learning rate of the formula, determining the size of the step along the gradient. This learning rate is crucial in our computations because if we choose it to be too small, our MI might never converge; on the other hand, if we choose it to be too large, our MI might oscillate back and forth without ever finding the maximum.

2.1 1D tuning curves – optimal widths

We choose to first run this algorithm for maximizing MI with respect to one parameter: the widths of the tuning curves, represented by σ_i. Using gradient ascent, we find the optimal widths of 20 tuning curves for the crickets dataset, showing that tuning curves tend to be narrower around the frequencies that carry greater power spectral density (around mel-scale values of 200 and 2200 in Figure 5). This matches the peaks of the power spectral density diagram 4. Plotting MI at each step of the gradient ascent algorithm, we find a steady increase and plateau of MI, as seen in Figure 6. This means we are indeed reaching a maximum of MI with our computations.

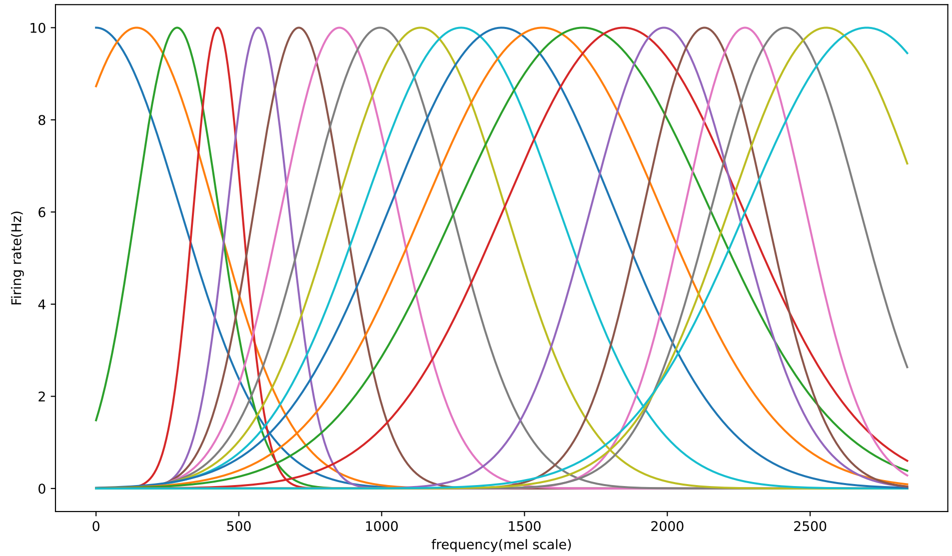

Figure 5: 1D Tuning curves for 20 neurons with optimized width

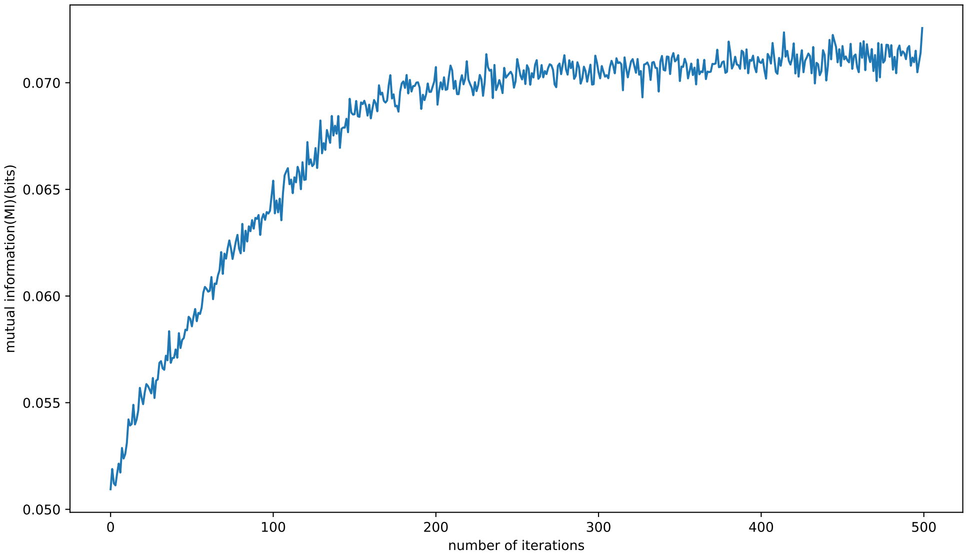

Figure 6: MI at each step of gradient ascent optimization algorithm 1D width

2.2 1Dtuning curves – optimal widths and population density

Next, we use the gradient ascent algorithm to simultaneously optimize with respect to two parameters: tuning curve widths and locations. The result of this optimization can be seen in Figure 7. As seen in the plot, neurons tuned to frequencies that have higher intensities display a high density of narrow tuning curves; conversely, those tuned to lower intensities are wider and sparser. Checking this result by plotting mutual information at each step of the gradient ascent algorithm once more, we see in Figure 8 another steady increase to plateau, meaning that we have successfully maximized MI. This approach of understanding auditory stimuli matches the results of previous literature, specifically in the 2010 paper by Ganguli and Simoncelli [6], where the authors found high density narrow tuning curves corresponding to the frequencies with high probabilities of a heterogeneous prior stimulus distribution.

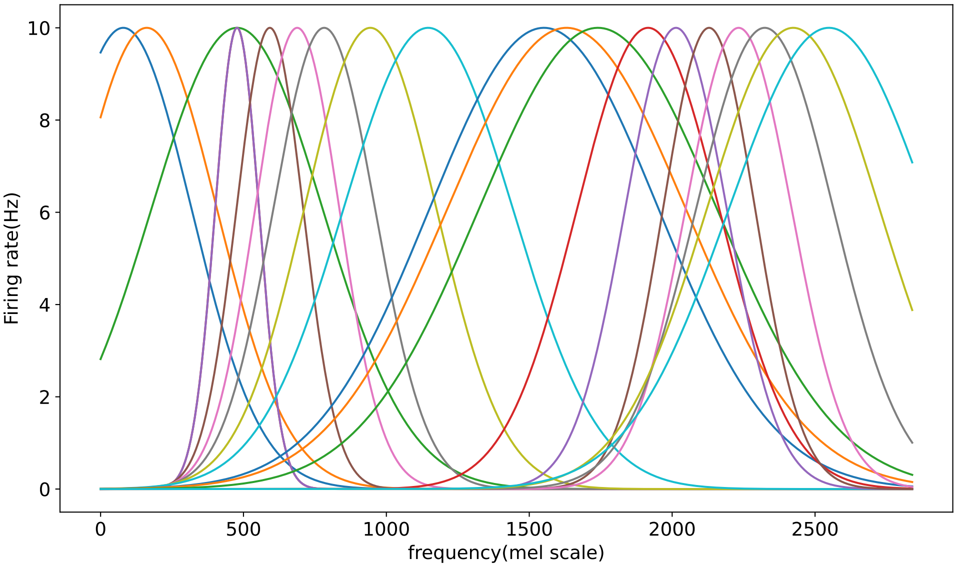

Figure 7: 1D Tuning curves for 20 neurons with optimized width and location

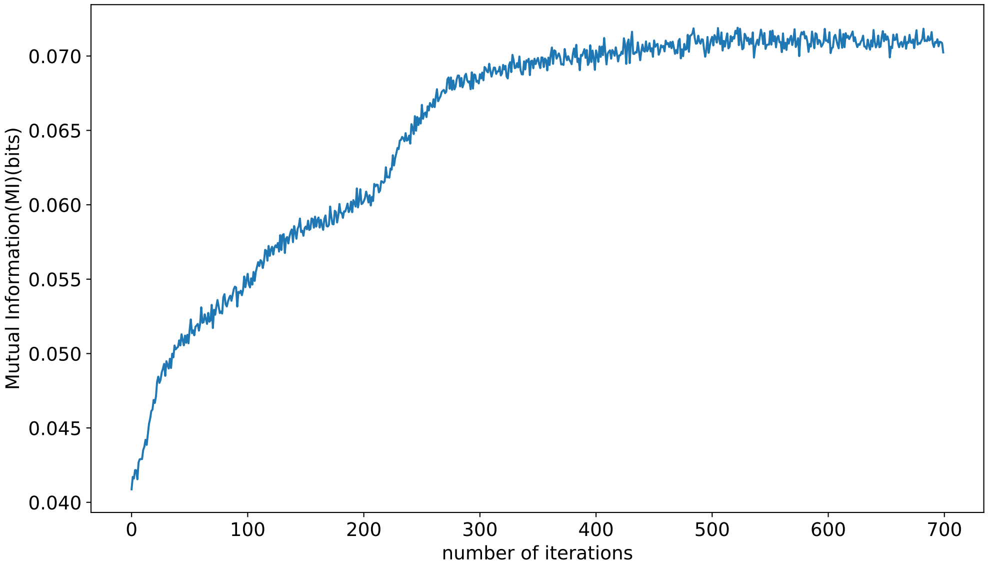

Figure 8: MI at each step of gradient ascent optimization algorithm 1D width and location

2.3 1D tuning curves – combined natural sounds

To further understand one-dimensional tuning curves and to generalize the results to explain neural response to a diverse ensemble of natural sound stimuli, we then automate the process of importing 98 natural soundtracks using the web scraping method from the same ambient sounds database [15], combining 100 seconds of each sound into one large sound file. We reason the combined soundtrack better represents the behavior of natural sound stimuli. Applying the same steps of the algorithm on this new set of stimuli, we first find the prior probability distribution $p(s)$ for the combined stimuli, using the same binning method as above for the singular cricket sound-track. We discover the highest probabilities occur around $400 - 800$ mel scale as seen in Figure 9. With this information, we should expect to see a dense population of narrow tuning curves around that frequency band. The results as seen in Figure 10 show that this is mostly the case, with the narrowest tuning curves somewhat densely populated around $400 - 800$ mel scale on the x-axis representing frequencies. Once again, to visually and quantitatively represent the mutual information maximization process as at each step of the gradient ascent, the plot in Figure 11 displays the plateauing trajectory of MI as we reach almost 900 iterations.

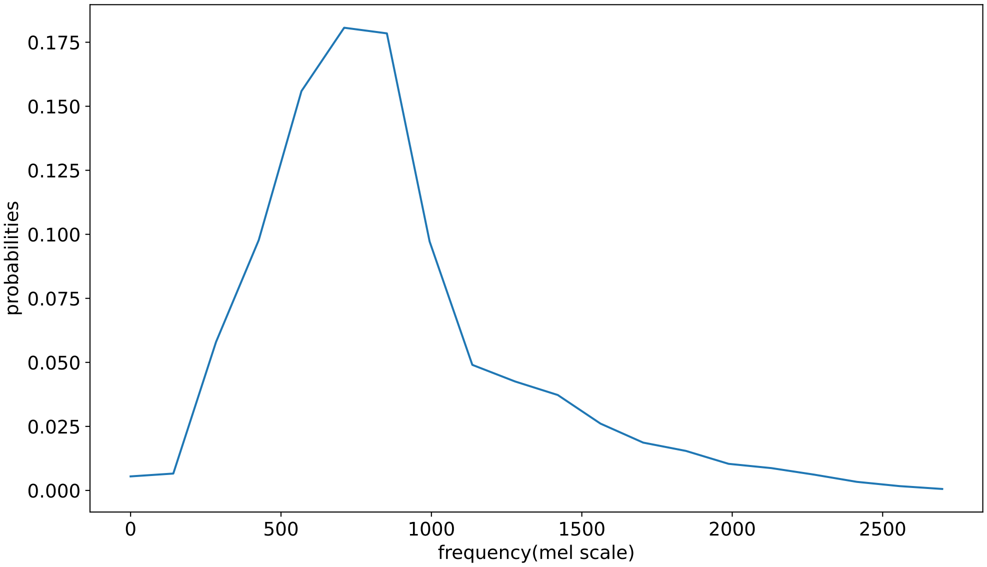

Figure 9: prior probability distribution of 98 combined natural soundtracks

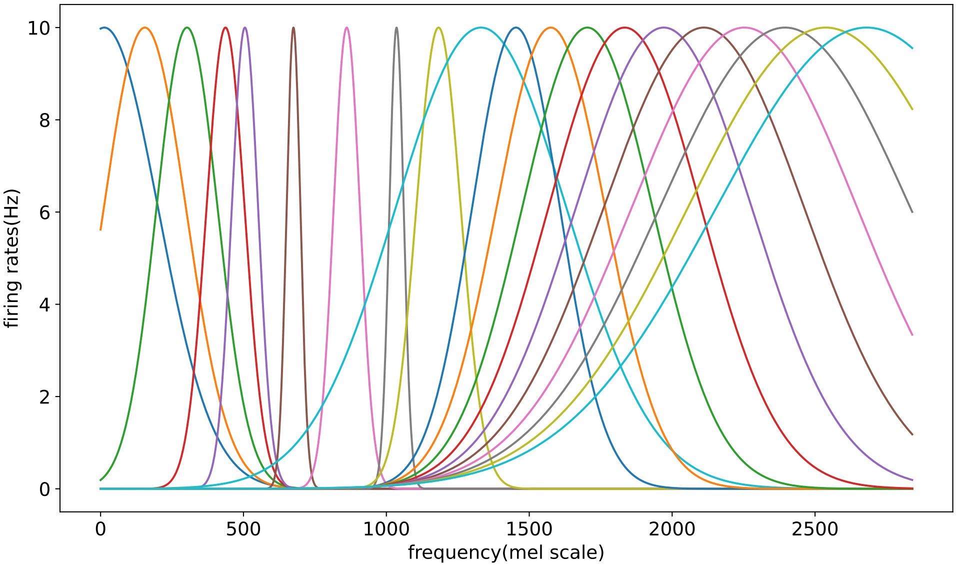

Figure 10: tuning curves of 98 combined natural sound stimuli

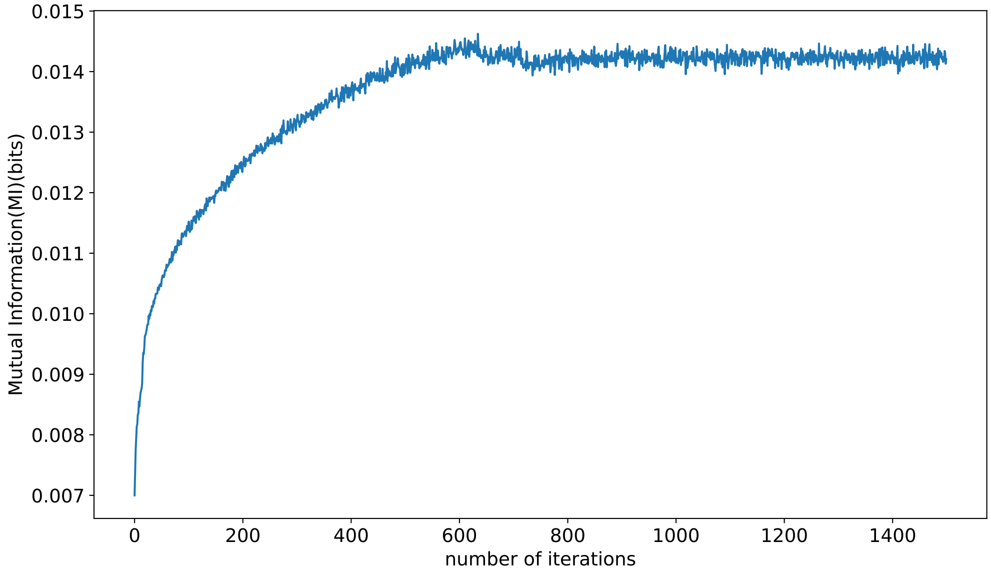

Figure 11: MI at each step of gradient ascent optimization algorithm for 98 natural sounds

So far, the analysis of the 1D neural tuning curves with our algorithm of maximizing MI utilizing stochastic gradient ascent have matched previous findings, specifically tuning curves tend to be narrower and denser in regions where the stimulus probability is larger. However, we have assumed up to this point that all tuning curves have the same functional form, bell-shaped neural responses (in the 1D case). We reason that this is not the case: optimal stimulus information is represented by multiple populations of neurons that respond in qualitatively different ways to auditory stimulus features, rather than one homogeneous population with similar response properties. Therefore, with highly suggestive evidence in the 1D case that our algorithm is a robust model in aiding us to further understand neural coding in the auditory cortex, considering heterogeneous populations of sensory neurons with two-dimensional stimuli will be the subject of the next chapter.

3 Maximizing Mutual Information: two-dimensional stimuli

While the algorithm from Chapter 2 successfully produces results that matched our hypothesis and paints a rather clear picture of the behavior of one-dimensional tuning curves in response to one stimulus parameter, we are still interested in understanding the existence of "O"-shaped and "V"-shaped tuning curves [16], which are in response to two auditory stimuli parameters: frequency and intensity (also known as sound level). Using the cricket soundtrack once again, we transform the frequencies into mel scale, and we also take the logarithm of the intensity values to better accommodate the large dynamic range of sound intensities. The prior stimulus probability distribution $p(s)$ after binning is now two dimensional, with a single probability assigned to each pair of values, $(\text{frequency}_i, \text{intensity}_j)$. To obtain this probability distribution, we utilize the histogram function to quantify intensity values, so that the sum of each frequency bin adds up to the marginal probability density function of frequency values, while the whole probability matrix is normalized to add up to 1. After constructing the prior probability distribution, we follow the steps similar to the 1D case, and we can model the tuning function with the two-dimensional Gaussian function:

$$f(s_1, ...s_k) = \frac{\exp\left(-\frac{1}{2}(\boldsymbol{s}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{s}-\boldsymbol{\mu})\right)}{\sqrt{(2\pi)^k|\boldsymbol{\Sigma}|}}. \tag{12}$$

3.1 2D O-shaped Tuning Curves

To model 2-dimensional O-shaped tuning curves, we represent the tuning widths σ_1, σ_2 in both directions in the following matrix Σ, where the widths of the tuning curves are determined by the eigenvalues of the matrix, which only correspond to σ_1 and σ_2 if ρ is 0:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho \\ \rho & \sigma_2^2 \end{pmatrix}, \tag{13}$$

where ρ represents the orientation of the 2D tuning curves.

Utilizing the fact that $\frac{\rho}{\sigma_1\sigma_2}$ has to be between -1 and 1 and the trigonometry function $\tanh(\gamma)$ is bounded between -1 and 1, we can represent ρ in terms of parameters $\sigma_1, \sigma_2, \gamma$:

$$\rho = \sigma_1\sigma_2 \tanh(\gamma). \tag{14}$$

With the nice property that O-shaped tuning curves are symmetrical across either the x or y-axis when $\rho = 0$, we insert this value of ρ in terms of $\sigma_1, \sigma_2, \gamma$. Now, we can derive the Gaussian tuning function that represents our O-shaped tuning curves, where c is the maximum amplitude of our tuning curves:

$$f(\boldsymbol{s}) = c{*}\exp\left(-\frac{1}{2}\left(\frac{(\boldsymbol{s}_1-\boldsymbol{\mu}_1)^2}{\boldsymbol{\sigma}_1^2(1-\tanh(\boldsymbol{\gamma})^2)} + \frac{(\boldsymbol{s}_2-\boldsymbol{\mu}_2)^2}{\boldsymbol{\sigma}_2^2(1-\tanh(\boldsymbol{\gamma})^2)} + \frac{(\boldsymbol{s}_1-\boldsymbol{\mu}_1)(\boldsymbol{s}_2-\boldsymbol{\mu}_2)\tanh(\boldsymbol{\gamma})}{\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2(1-\tanh(\boldsymbol{\gamma})^2)}\right)\right). \tag{15}$$

Interpreting this function in the context of our problem, the $\boldsymbol{s_1}, \boldsymbol{s_2}$ values are stimuli pairs frequency s_1 and intensity s_2, and there are 5 parameters we are interested in optimizing to investigate how each of them contribute to maximum MI: $\boldsymbol{\mu_1}, \boldsymbol{\mu_2}$ values are the centers of each tuning curve, this should be 20 pairs of values on the frequency and intensity axes for our 20 neurons; $\boldsymbol{\sigma_1}, \boldsymbol{\sigma_2}$ are vectors of 20 tuning widths; $\boldsymbol{\gamma}$ is a vector of 20 values determining the orientation of our tuning curves. We want to try to understand what combination of these parameters contribute to optimal 2-dimensional neural coding.

Initializing each parameter similar to the 1D case, we begin the optimization process with equal widths in both directions, evenly distributed peak locations across both frequency and intensity values, and a vector of 0s for the orientations of 20 neurons. To further optimize our results, we perform the gradient ascent to maximize MI with respect to $\boldsymbol{\mu_1}, \boldsymbol{\mu_2}, \boldsymbol{\sigma_1}, \boldsymbol{\sigma_2}, \boldsymbol{\gamma}$ on a Graphical Processing Unit (GPU) in the lab. As opposed to CPU on a laptop computer, GPUs have more cores than CPU and hence when it comes to parallel computing of data, GPUs perform exceptionally better than CPU. We are able to run more iterations at a faster time this way, and obtain the 2D O-shaped tuning curves of 20 neurons, as seen in Figure 12. As demonstrated by the plot of MI at each step of the gradient ascent algorithm in Figure

14, we see as the number of iterations approach 20000, our MI increases fast then slow to approach a constant value, suggesting we have found the maximum MI for this set of 20 neurons. More importantly, if we compare this result to the cricket soundtrack probability distribution plot as seen in Figure 13, we are able to see that if a stimulus exhibit a high probability at a high frequency, then it has a high probability of having high intensity (this is based on the marginal probability density construction of our 2-dimensional $p(s)$). Here, the O-units that span across the smallest frequency range and intensity range (top right corner), and the highest density of O-units, are positioned at the relative high frequency and high intensity values, which corresponds to the high probability region as determined by the prior probability distribution $p(s)$.

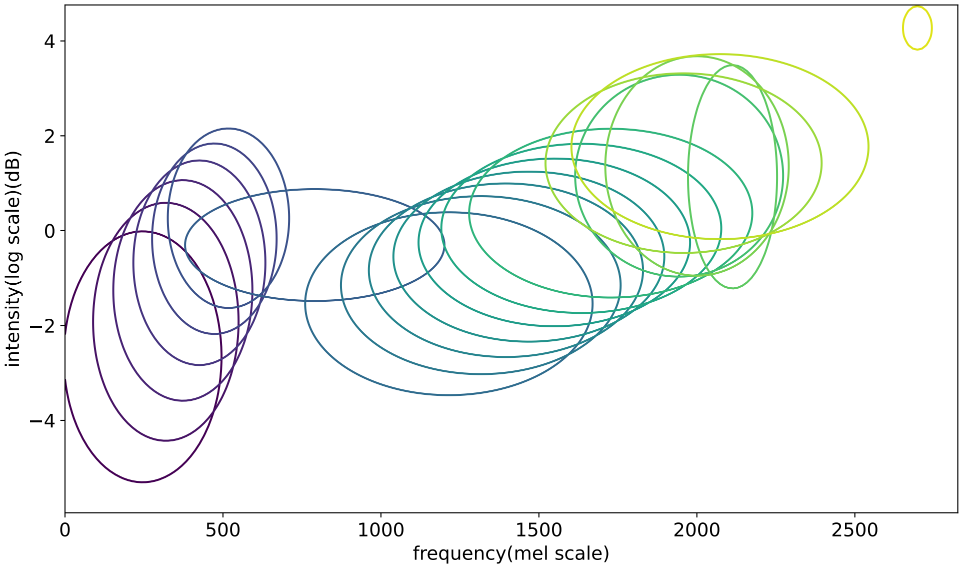

Figure 12: 2D O shaped tuning curves

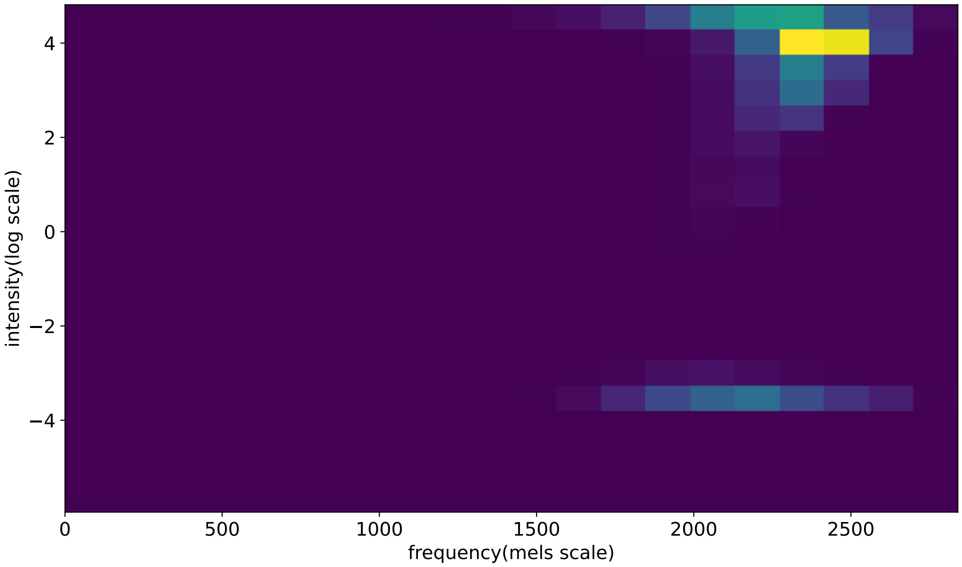

Figure 13: 2D crickets natural stimuli probability distribution

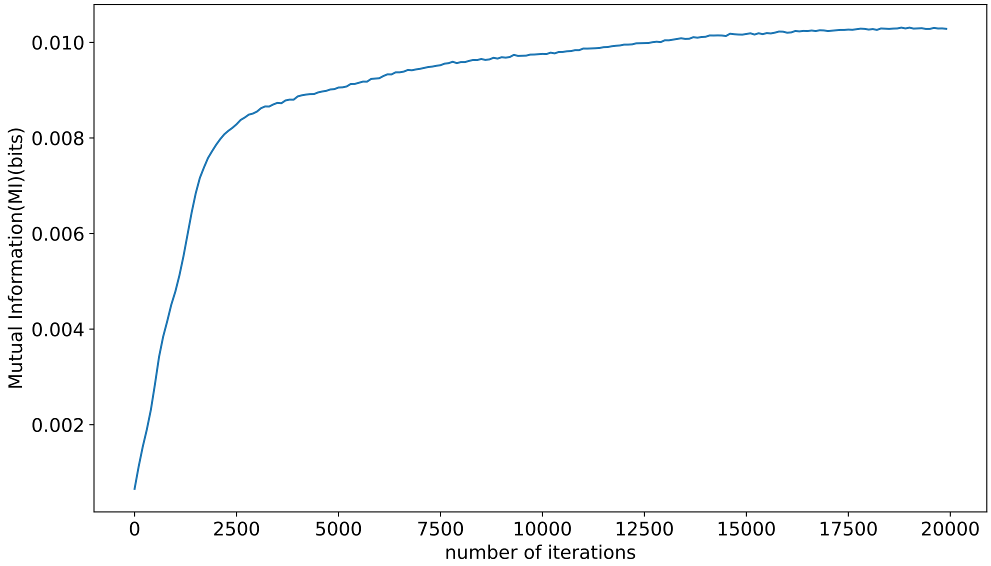

Figure 14: mutual information at each step of gradient ascent optimization algorithm for 2D O-shaped tuning curves

One potential interpretation of the distribution of our O-shaped tuning curves is that the optimal O shaped tuning curve should span across a small range of frequency and intensity values, so that we can derive accurate and specific information at that stimuli pair. Additionally, the orientations of the O-units appear to be most optimally represented aligned with respect to the vertical and horizontal axes of the two-dimensional stimulus space. While values of MI such as the ones in Figure 14 are difficult to interpret in absolute terms, we can gain insight by comparing maximum MI values computed in different conditions. Thus, in the following sections we want to compare the approximated maximal mutual information value of 20 neurons with purely O-shaped tuning, purely V-shaped tuning, or a mix of both shapes in order to determine how optimal stimulus information is represented in a population of neurons.

3.2 2D V-shaped Tuning Curves

Next, we model "V"-shaped tuning curves with a similar 2D Gaussian function as before, but now with a minor difference. This is because as the stimulus parameter in the y-direction (sound intensity) increases, the V-shaped tuning curves span across a larger range of frequency values; therefore, the tuning width in the x direction is dependent of the intensity stimulus value. We represent this by writing σ_1 as a function of s_2: $\sigma_1(s_2) = \boldsymbol{\beta_1}(\boldsymbol{s_2} - \boldsymbol{\alpha})$, where $\boldsymbol{\alpha}$ is a vector of threshold values for each neuron to constrain $\boldsymbol{s_2}$ from falling below a certain value, and $\boldsymbol{\beta_1}$ is a multiplication parameter determining the widths of the tuning curves. Then, putting all the necessary components into the tuning function, we get:

$$t = c * \exp\left(-\frac{1}{2}\frac{(\boldsymbol{s_1} - \boldsymbol{\mu_x})^2}{(\boldsymbol{\beta_1}(\boldsymbol{s_2} - \boldsymbol{\alpha}))^2}\right)\tanh(\boldsymbol{\beta_2}(\boldsymbol{s_2} - \boldsymbol{\alpha})). \tag{16}$$

Here, c once again controls the maximum firing rate of the tuning curve; s_1, s_2 are the frequency and intensity of a stimulus randomly drawn from the prior probability distribution and $\boldsymbol{\mu_x}$ is the center of each of the 20 V-shaped tuning curve. Our goal will be to optimize the MI with respect to the parameters $\boldsymbol{\mu_x}, \boldsymbol{\beta_1}, \boldsymbol{\alpha},$ and $\boldsymbol{\beta_2}$.

The tanh in Equation 16 restricts the function to have a maximum firing rate, preventing the MI from growing indefinitely with increasing neural firing rate. Adapting each probability density function to reflect the V-shaped tuning curve behavior, We perform gradient ascent as before to optimize the parameters in this new model. The result of the optimization can be seen in Figure 15. As we can see, we derive our desired V-shaped tuning curves. It is interesting to see the tuning widths to be so narrow across the frequency stimuli, and rather evenly spaced locations spanning across the intensity stimuli space.

As we plot MI again at each step of the gradient ascent iterations, we find the plot reaching a maximum of approximately 0.0013 (see Figure 16). This relatively small value means that, given our relatively small number of neurons, we can tell little about a stimulus given the neural response. Additionally, comparing this value to the maximum MI from the O-shaped tuning curves, we can conclude that so far, O-shaped tuning is the model that encodes more information about the natural stimuli. In the next section, we investigate the case of a neural population with a heterogeneous combination of O and V-shaped tuning curves, how this affects MI, and what the results tell us compared to homogeneous population of O-shaped tuning curves or V-shaped tuning curves.

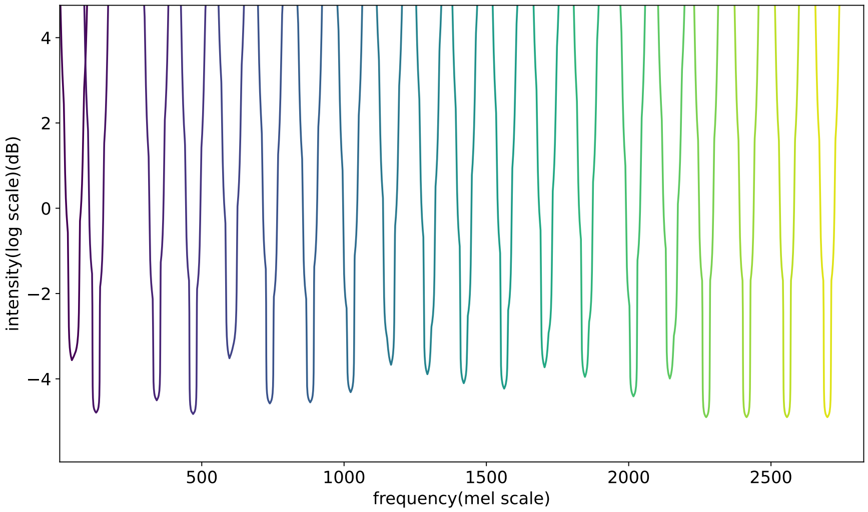

Figure 15: 2D V shaped tuning curves

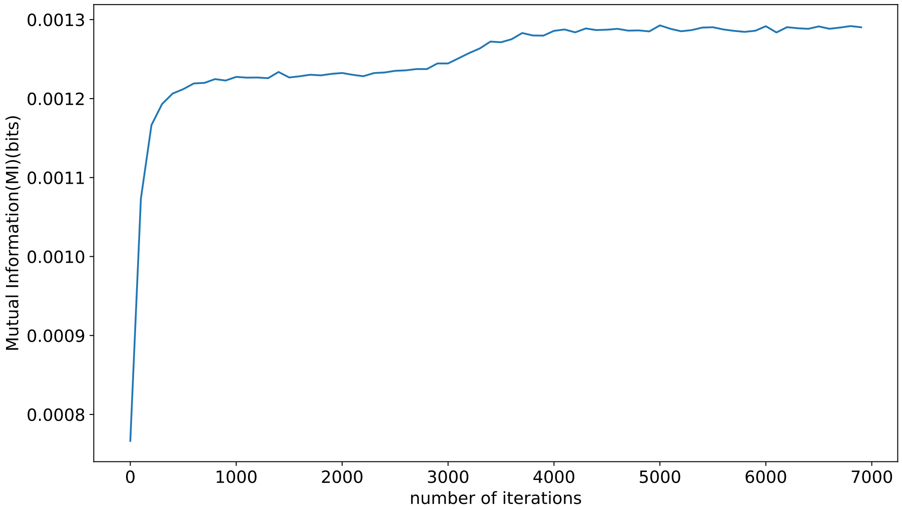

Figure 16: mutual information at each step of gradient ascent optimization algorithm for 2D V-shaped tuning curves

3.3 2D Mixed O/V-shaped Tuning Curves

To construct a tuning function that models a combination of "O" and "V"-shaped tuning curves, we achieve this by combining 10 neurons with O-shaped tuning function 15 and 10 neurons with V-shaped tuning function 16. Now, we need to optimize our mutual information with respect to 9 parameters: $\boldsymbol{\sigma_1}$, a vector of 10 tuning widths across the frequency axis; $\boldsymbol{\sigma_2}$, a vector of 10 tuning widths across the intensity axis; $\boldsymbol{\mu_1}$, peaks of 10 O-shaped tuning curves on the frequency axis; $\boldsymbol{\mu_2}$, peaks of 10 O-shaped tuning curves on the intensity axis; $\boldsymbol{\mu_3}$, peaks of 10 V-shaped tuning curves on the frequency axis; $\boldsymbol{\gamma}$, a vector of 10 values determining the orientation of O-tuning curves; $\boldsymbol{\beta_1}$, a vector of 10 values determining the width of V-tuning curves; $\boldsymbol{\beta_2}$, a vector of 10 multiplication factors for V-shape tuning curves; $\boldsymbol{\alpha}$, a vector of 10 intensity threshold values for V-shaped tuning curves. We proceed to update all functions in the algorithm once more to analyze the mixed tuning curve distribution and MI maximization with respect to these parameters. As plotted in Figure 17, we see a distribution of O-shaped and V-shaped tuning curves obtained after performing this optimization. The convergence of MI during gradient ascent is shown in Figure 18. The final value of 0.005 is the second greatest MI out of the three cases that we have considered, as visually represented in Figure 19, suggesting that our optimal neural population might not take form in a heterogeneous population of neurons, but rather a homogeneous distribution of all O-shaped tuning curves.

Another aspect of understanding these two-dimensional tuning curves that has not been investigated by previous literature is how the orientations of O-shaped tuning curves contribute to maximization of mutual information, and what amount of auditory information is encoded in the orientation parameter ρ from Equation 14. While previous experimental work [16] appears to suggest that tuning curves tend to be aligned with respect to the vertical and horizontal axes of the two-dimensional stimulus space, as shown in Figure 2, this does not appear to have been carefully quantified. By adding this parameter ρ modeled in Equation 14 to the function characterizing O-shaped tuning curves in Equation 15, we find that the optimal tuning curves orientation that maximizes MI is not necessarily at a 90 degree angle,

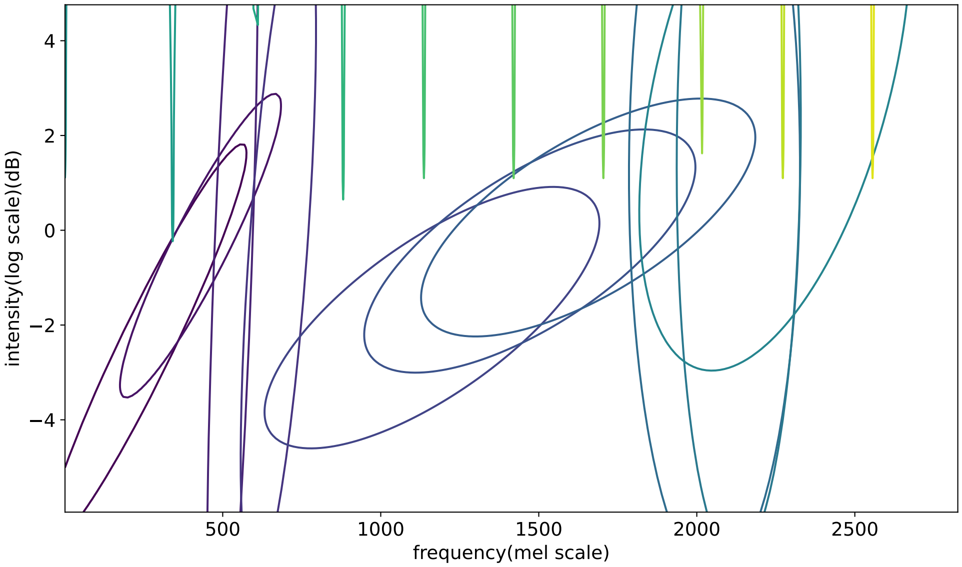

Figure 17: 2D mixed shaped tuning curves

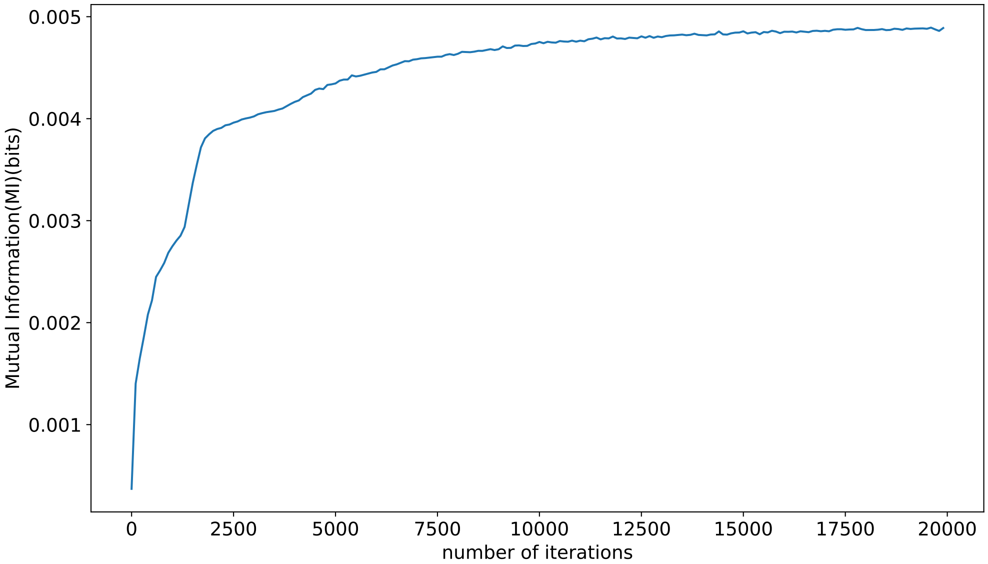

Figure 18: mutual information at each step of gradient ascent algorithm for 2D mixed shaped tuning curves

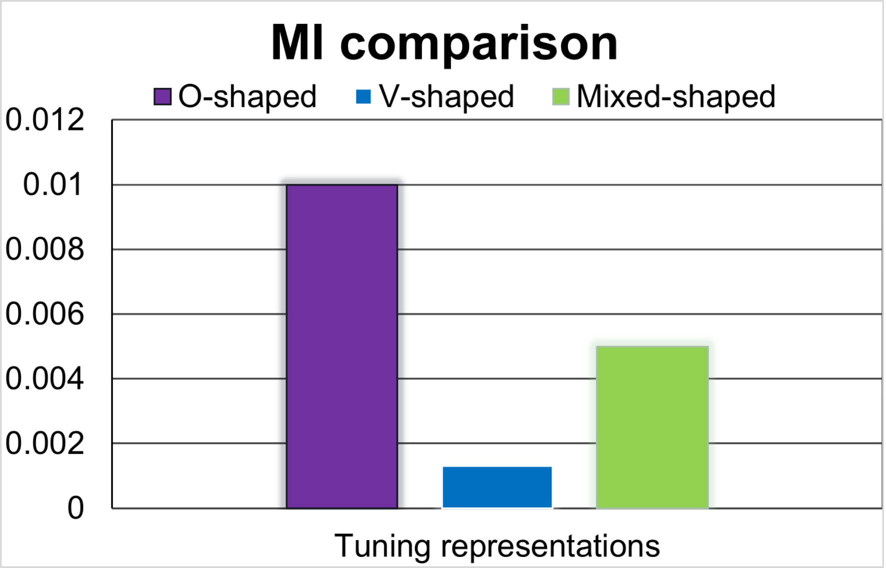

Figure 19: mutual information comparison for each of the three neural representations we considered

under the condition of a heterogeneous population of qualitatively different tuning curves. As we see in our 2D plot in Figure 17, the optimized tuning curves in many cases display a tilted angle. Precise quantification of the specific degree of the tilt and how it relates to optimal neural response representation is an area that would require further research.

4 Discussion

This thesis approaches a rather complex problem: understanding the neural response representation of two-dimensional tuning curves, with a step-by-step analysis. Starting with modeling 1D tuning curves with natural sound stimuli, we successfully find the optimal tuning width and tuning density associated with 20 arbitrary neurons with a max firing rate of 10. In the case of 1D neural response to natural sound stimuli, we find frequently occurring stimuli are optimally represented by high densities of neurons with narrow tuning. This result not only confirms the findings of previous studies [6], but it also provides the modeling framework for optimizing neural representations for two-dimensional stimuli.

In the 2D case, motivated by experimental observations, we consider both O- and V-shaped tuning curves in our model. Our mutual information maximization and gradient ascent algorithm display two important findings: first, compared to V-shaped neurons and a combination of V-shaped and O-shaped neurons, a distribution of entirely O-shaped tuning appear to encode more information about natural stimuli. This can be seen from the MI comparison results in Figure 19 after performing the optimization process for the three cases we considered. Second, the nuance observation that the orientations of O-shaped tuning curves in a population of neurons with mixed response properties are not optimally represented by the aligned vertical or horizontal positions, but rather at a tilted angle, sheds new light into further investigations on this topic.

The result of our optimizations of the two-dimensional stimuli did not match our hypothesis; therefore, we propose two possible reasons for this outcome: first, we should more carefully update our models and repeat experiments to investigate the result of our optimal tuning curves. Second, the auditory cortex does not actually employ mutual information maximization technique to optimally represent natural sound stimuli. Nevertheless, the methods and results of this research project help us better understand neural representation in the auditory cortex. This is because our method produced satisfactory results in the one-dimensional case, as well as the possible conclusion that O-shaped tuning curves truly serve as the optimal two-dimensional stimulus representation. This is important to further research in auditory neuroscience research, as a faithful representation of auditory stimuli in the auditory cortex is an essential first step in auditory processing, especially when trying to understand the behavior of a large population of neurons.

We recognize the limitations of these results in addition to the discrepancy between our second hypothesis and result, and we propose possible solutions. First, a larger data collection and more fine-tuned sound stimuli selection might increase the accuracy of the results. Selecting 200s of a random natural sound from the Montana State University database as stimuli is a good method for the time being; however, it does not account for times where there are little to non-existent sound activity; furthermore, a larger quantity of data would be even more representative of the natural sound environment, for both the 1D and 2D cases. Next, modeling more than the arbitrary 20 neurons we utilized in our computations, and investigating the robustness of our optimizations by repeating each maximization many times to make sure that the algorithm finds similar solutions each time, are both strategies we can implement to improve the results.

An immediate future direction of this research project is to better quantify the orientation parameter as well as any additional factors that impact neural response, in addition to optimizing the Python code that contains the algorithm to fit larger datasets. Furthermore, it will be interesting to apply our algorithm to explore how neural representations of stimuli change with learning, for example in mice trained to respond to a particular auditory frequency. One potential hypothesis is that more neurons should be devoted to representing stimuli that are important for performing a learned behavior. The outcome of this line of future work is to be able to predict what neural tuning curves should look like and how they should differ for different stimuli before and after learning, which would be an exciting connection to not only the field of neuroscience, but also other areas of studies such as linguistics and education. For example, in understanding the role of active and passive language learning.

References

[1] Horace B Barlow et al. “Possible principles underlying the transformation of sensory messages”. In: *Sensory communication* 1.01 (1961).

[2] Alexander Borst and Frédéric E. Theunissen. “Information theory and neural coding”. In: *Nature Neuroscience* 2 (1999), pp. 947–957.

[3] Madhavun Candadai. “Information theoretic analysis of computational models as a tool to understand the neural basis of behaviors”. In: *ArXiv* abs/2106.05186 (2021).

[4] Peter Dayan and Laurence F Abbott. “Theoretical neuroscience: computational and mathematical modeling of neural systems”. In: *MIT press* (2005).

[5] Lei Feng and Xiaoqin Wang. “Harmonic template neurons in primate auditory cortex underlying complex sound processing”. In: *Proceedings of the National Academy of Sciences* 114 (Jan. 2017), p. 201607519. DOI: `10.1073/pnas.1607519114`.

[6] Deep Ganguli and Eero Simoncelli. “Implicit encoding of prior probabilities in optimal neural populations”. In: *Advances in neural information processing systems* (Jan. 2010), pp. 658–666.

[7] Deep Ganguli and Eero Simoncelli. “Neural and perceptual signatures of efficient sensory coding”. In: *arXiv:1603.00058: Neurons and Cognition* (Feb. 2016).

[8] Troy A. Hackett. “Chapter 2 - Anatomic organization of the auditory cortex”. In: Handbook of Clinical Neurology 129 (2015). Ed. by Michael J. Aminoff, François Boller, and Dick F. Swaab, pp. 27–53. ISSN: 0072-9752. DOI: `https://doi.org/10.1016/B978-0-444-62630-1.00002-0`.

[9] Robin AA Ince, Stefano Panzeri, and Christoph Kayser. “Neural codes formed by small and temporally precise populations in auditory cortex”. In: *Journal of Neuroscience* 33.46 (2013), pp. 18277–18287.

[10] Andrew King, Sundeep Teki, and Ben Willmore. “Recent advances in understanding the auditory cortex”. In: *F1000Research* 7 (Sept. 2018), p. 1555. DOI: `10.12688/f1000research.15580.1`.

[11] Konrad P Kording. *Short Course 2: The Basics of Neural Coding*. Accessed: 5-13-2022. 2016.

[12] Nikolaus Kriegeskorte and Xue-Xin Wei. “Neural tuning and representational geometry.” In: *Nature reviews. Neuroscience* (2021).

[13] L. Martignon. “Information Theory”. In: (2001). Ed. by Neil J. Smelser and Paul B. Baltes, pp. 7476–7480. DOI: `https://doi.org/10.1016/B0-08-043076-7/00608-2`.

[14] Daniel B. Polley, Elizabeth E. Steinberg, and Michael M. Merzenich. “Perceptual Learning Directs Auditory Cortical Map Reorganization through Top-Down Influences”. In: *Journal of Neuroscience* 26.18 (2006), pp. 4970–4982. ISSN: 0270-6474. DOI: `10.1523/JNEUROSCI.3771-05.2006`.

[15] Researchers and National Park Service. “Acoustic Atlas Ambient Recordings”. In: *Montana State University Library Digital Collections* (2022). URL: `https://acousticatlas.org/ambient.php`.

[16] Srivatsun Sadagopan and Xiaoqin Wang. “Level Invariant Representation of Sounds by Populations of Neurons in Primary Auditory Cortex”. In: *Journal of Neuroscience* 28.13 (2008), pp. 3415–3426. ISSN: 0270-6474. DOI: `10.1523/JNEUROSCI.2743-07.2008`.

[17] Claude E Shannon. “A Mathematical Theory of Communication, Bell System Technical Journal, July”. In: (1948).

[18] Vargha Talebi and Curtis L. Baker. “Natural versus Synthetic Stimuli for Estimating Receptive Field Models: A Comparison of Predictive Robustness”. In: *Journal of Neuroscience* 32.5 (2012), pp. 1560–1576. ISSN: 0270-6474. DOI: `10.1523/JNEUROSCI.4661-12.2012`.

[19] Frédéric E Theunissen, Kamal Sen, and Allison J Doupe. “Spectral-temporal receptive fields of nonlinear auditory neurons obtained using natural sounds”. In: *Journal of Neuroscience* 20.6 (2000), pp. 2315–2331.

[20] Xiaoqin Wang. “Cortical Coding of Auditory Features”. In: *Annual Review of Neuroscience* 41.1 (2018). PMID: 29986161, pp. 527–552. DOI: `10.1146/annurev-neuro-072116-031302`.

[21] Xiaoqin Wang. “Neural coding strategies in auditory cortex”. In: *Hearing Research* 229.1 (2007). Auditory Cortex 2006 - The Listening Brain, pp. 81–93. ISSN: 0378-5955. DOI: `https://doi.org/10.1016/j.heares.2007.01.019`.