A Neuromorphic Model of the Peripheral Auditory System Implemented in MATLAB

WPI

A Major Qualifying Project Report Submitted to the Faculty
Of the
WORCESTER POLYTECHNIC INSTITUTE
in Partial Fulfillment of the Requirements for the
Degree in Bachelor of Science

Submitted By:

Austin Aguirre – Computer Science
Hunter Lassard – Biomedical Engineering
Aidan Pereira – Biomedical Engineering
Jack Brazer – Biomedical Engineering

Advised by:

Professor Adam Lammert
Department of Biomedical Engineering
Department of Computer Science

Abstract

Many people are affected by disorders of the auditory system that can make it harder to connect to other people and the rest of the world. Thus, researchers of the auditory system have attempted to replicate it in different ways in order to better understand it. However, these models tend to be limited by their levels of abstraction, which makes it difficult to chain them together for joint study. The model described in this paper sought to find a middle ground in the levels of abstraction, and to keep it as physiologically relevant as possible. To do this, the model contained parts to model the basilar membrane, inner hair cells, and the auditory nerve. Code was written to simulate the various functions of the encoding of a signal performed by the inner ear. The encoded signal was then decoded in order to assess the validity of the model. A constant 2 Hertz (Hz) signal and a short voice recording were both run through the model that produced accurate and audible results.

Table of Contents

1. Introduction

Hearing is essential to our relationships, development, and perception of the world around us. Our ability to hear connects us, with one another in a way that none of our other senses provide, allowing us to communicate through the art of language. A properly functioning auditory system allows for our perception and understanding of complex speech patterns. This allows for the transmission of ideas that encourage problem solving and emotional expression. These aspects of communication are essential to our development as children and our process of learning and understanding the world around us. Hearing also allows us to respond to dangerous stimuli in our environment, or listen to a heartwarming piece of music. Our ability to hear is essential to our physical, mental, and emotional wellbeing. When the ability to hear is inhibited, a person's quality of life diminishes. Hearing impairments are estimated to affect 2.5 billion people by 2050 and 700 million of them will have disabling hearing loss or deafness [1]. Therefore, it is essential for auditory researchers to better understand the auditory system and the causes of hearing loss to find better treatments and hearing aids.

An extremely effective way for researchers to better understand the auditory system is with computational models of the auditory system. These models allow researchers to study the way our auditory system processes and interprets sound in the brain. They also can simulate damaged auditory systems and the impact of hearing aids and cochlear implants, leading to accurate predictions for improving hearing loss and impairments. Hearing impairments present themselves in many different ways. People can experience varying difficulties with hearing and auditory perception through: hearing loss, deafness, tinnitus, auditory processing disorder, and

Meniere's disease. Therefore, models of the auditory system can be extremely useful in modeling dysfunctions of its different components that cause different impairments and symptoms.

Currently there are a wide range of models that capture different sections of the auditory system, notably: the Meddis hair cell model [2] and the Heinz model of the auditory nerve [3]. These models both focus on distinct parts of the broader auditory system, and one important stumbling block for attempts to unify them into a single model has been the fact that these models exist at different levels of abstraction [4]. For instance, the Meddis model exists on the implementational level - i.e., describing the biochemistry leading to nervous system activity - while the Heinz model of the auditory nerve, and models of the central nervous auditory system exist at the algorithmic level - i.e., describing neural activity in aggregate. Attempts to create a single model of the auditory system, for example the MATLAB Auditory Periphery (MAP) model [5], have approached this mismatch in abstraction largely by eliminating implementational details, resulting in an overall more abstract model.

As a result, current algorithmic and computational models of the auditory system do not accurately represent the physiological stimulation and propagation of signals defined at the implementational level, which prevents these models from transmitting a sound relevant to the human case. Differing levels of abstraction prevent different models from being relevant to one another or useful to combine. A strictly algorithmic, computational model is less useful for researchers developing medical devices as these models do not reflect the exact auditory signal transmission process. Therefore, each biological component in the auditory system is not accurately represented, and the model is unable to address the pathologies of auditory disease.

One solution to the problem of creating a unified model of the auditory system that provides sufficient low-level detail is to incorporate recent advancements in neural modeling

from the field of neuromorphic computing. Modern neuromorphic models are capable of representing fairly abstract signal processing functions like those conducted by the human auditory system, while still emulating more realistically the exact stimulation processes that naturally occur in humans. A comprehensive example of such neuromorphic models can be found in the mathematical principles of the Neural Engineering Framework (NEF). The NEF is able to link the algorithmic level of sound signal processing, based on the encoding and decoding of a population of neurons, to the implementational level of the stimulation of single neurons from the transduction of mechanical pressure waves to electrical signals [6]. The neuromorphic model also allows the cohesion of different sections of the auditory system that transmit signals between each other.

There is a need for a more physiologically relevant model of the auditory system that is able to bridge the algorithmic and implementational levels of abstraction. Such a model would allow for multiple sections of the auditory periphery to be simulated with real-time signal processing in one model. Following the principles of the Neural Engineering Framework, our model will be physiologically relevant to the signal processing conducted by the basilar membrane, inner hair cells, and auditory nerve, while simultaneously relating those functions to low-level neural activity (e.g., action potentials). Furthermore, by implementing the model in a software interface using MATLAB, the model itself will be accessible to a researchers with a wide range of technical backgrounds, in addition to being useful for understanding the physiological signal processing of sound, pathologies of specific hearing impairments, and effects of cochlear implants on specific areas in the auditory system.

2. Background

2.1 Sound as a Signal:

Sound is the natural phenomena caused by the vibration of matter. In the context of hearing, sound is propagated through the vibration of air and its changing pressure. The vibration travels away from its origin along an axis where molecules condense and refract about a central resting point to propagate the signal. This signal can be quantified by a sinusoidal waveform that represents the change in pressure for a certain period of time. A sinusoid is represented by three main characteristics:

1. Frequency: the number of times the waveform repeats itself in one second (represented in Hertz).
2. Amplitude: the magnitude of the pressure change (represented in Pascals or other pressure units).

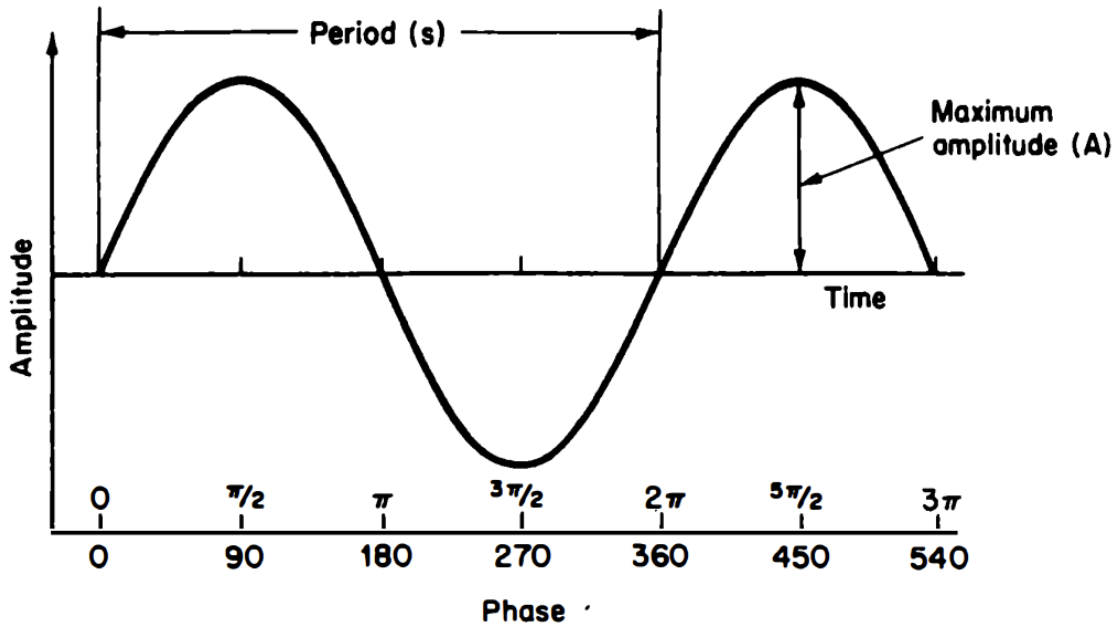


Figure 1: Sound as a Signal. Moore [7]

3. Phase: The distance that the peak of one sine wave is separated from another sine wave.

The perception of the unique pitch of a sound is related to its vibrational frequency. The higher the frequency, the higher the pitch. Sound can be characterized by its magnitude by our perception of loudness. The intensity of a sound is measured in decibels. The intensity of a sound is related to its power, or the energy it transmits over a certain area in a specific amount of time.

Sounds can be composed of many pressure sinusoids that make up a complex signal. These complex signals can be broken down into its component parts through a Fourier transform. The Fourier transform is a mathematical process that deconstructs a complex signal into a series of simple sinusoids, specified by their frequency, amplitude, and phase. The complex waveform can therefore be thought of as the addition of many simple waveforms.

2.2 Anatomy of the Auditory System:

The body's auditory system interprets sound by processing vibrational signals, transmitted by differing pressure waveforms in the air. The body's ability to process sound as raw waveforms and transmit them into useful electrical signals in the nervous system allows for the understanding of our environment and interpersonal communication. The ear is the first and most familiar component of the auditory system. It transduces raw sound pressure waveforms into the electrical patterns that the nervous system can propagate and interpret. The ear has three main structures: the external ear, the middle ear, and the inner ear. These portions of the ear can be visualized in *Figure 2* below.

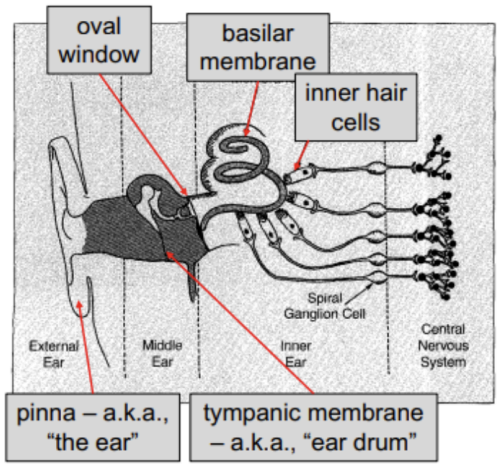


Figure 2: Structure of the ear and nervous auditory system [7]

The external ear is composed of the pinna (visible portion of the ear) and the auditory canal [7]. The structure of the pinna amplifies and modifies raw sound waves and directs them down the auditory canal. The auditory canal directs the sound waves to the tympanic membrane (eardrum), located in the middle ear, causing it to vibrate.

The vibration of the tympanic membrane propagates to the ossicles in the middle ear. The ossicles (the malleus, incus, and stapes) are the smallest bones in the body and further amplify the vibrational signal [8]. The stapes contacts the cochlea's outermost membrane called the oval window. The stapes vibrates onto the oval window, causing a pressure differential between the oval window and the cochlea's second membrane, the round window. The pressure differential effectively transmits sound into the cochlea. The middle ear's propagation of vibrations to the cochlea resembles an impedance matching device or transformer, at frequencies between 500 and 5000 Hz [7].

The inner ear originates with the cochlea, a spiral shaped structure composed of the vestibular (Reissner's) membrane and the basilar membrane. The origin of the cochlea at the oval membrane is referred to as the base and the end of the spiral is known as the apex. This structure

can be depicted as unwound from its spiral shape to better visualize the vibrations of the basilar membrane. The vibration of the oval membrane extends along the basilar membrane, causing it to move. The pressure differential across the basilar membrane causes motion in response to sinusoidal stimulation that travels from the base of the basilar membrane to its apex. The basilar membrane's mechanical properties vary along the length of its narrow and rigid base to the wide and flexible apex. The mechanical differences result in certain vibrational frequencies causing different levels of displacement of the basilar membrane along its length, acting like a Fourier analyzer [7]. The frequency that causes the maximum displacement of a point on the basilar membrane is known as its characteristic frequency [7]. Therefore, each position on the basilar membrane acts like its own bandpass filter with its center frequency being its characteristic frequency.

The basilar membrane is connected to the organ of Corti by inner hair cells (IHC). The IHC are oriented along the length of the basilar membrane and are activated when its corresponding location on the basilar membrane is stimulated by its characteristic frequency. The stereocilia of the IHC are protrusions of the cellular membrane that are displaced by the motion of the basilar membrane caused by its characteristic frequency. The deflection of the stereocilia open mechanically gated ion channels that allow K^+ to enter the cell. The influx of K^+ causes the cell to depolarize, opening voltage-gated calcium channels, initiating the release of the neurotransmitter glutamate at the opposite end of the hair cell. Across this end of the hair cell are auditory nerve fibers that are depolarized and fire an action potential from the release of glutamate. The hair cell effectively transmits the mechanical signal of the basilar membrane into an electrical signal that relays a chemical signal in the form of a neurotransmitter. This signal

transduction converts mechanical signals into chemical and electrical signals that the nervous system propagates further along the auditory system.

The auditory nerve is composed of roughly 30,000 nerve fibers, approximately 20 for every hair cell. This results in population coding, where multiple neurons represent the stimuli of one hair cell. Their combined action potentials are carried to the cochlear nucleus in the brain stem and then to the auditory cortex in the brain where perception occurs.

2.3 Neural Communication in the Auditory System:

Neurons communicate through the transmission of electrical signals that occur as action potentials. Action potentials occur when a stimulus causes the neuron's cell membrane potential to rise above its action potential threshold [9]. Action potentials that are propagated by action potentials from depolarization that cause a spike in the activity of the neuron along its cell membrane.

The neurons in the auditory nerve transmit signals from the cochlea to the central nervous system. These signals travel through the auditory nerve to the brainstem. The auditory nerve consists of many neural fibers that relay the information from one haircell, interpreting the signal from a frequency specific location on the basilar membrane.

Tuning curves are utilized to compare an auditory nerve fiber's threshold at different frequencies. The minima on a frequency-threshold curve (FTC) corresponds to the characteristic frequency that displaces the basilar membrane that causes activation of the corresponding hair cell. Therefore, these fibers are frequency dependent based on its activation from the frequency selectivity of the basilar membrane. This phenomena causes the orientation of the cochlea and auditory nerve to mirror each other by tonotopic organization [7]. Fibers activated by high

frequencies are found on the outside of the auditory nerve, and fibers activated by lower frequencies are located in the middle of the auditory nerve.

Auditory nerve fibers experience background spikes without sound stimulation. The frequency of auditory nerve background spikes are dictated by the orientation of the synapse between the auditory neurons and inner hair cells and the neuron's stimulation threshold. High frequencies of background spikes are caused by large synapses, and low spike frequency is caused by small synapses. Low thresholds cause high spike frequencies, and high thresholds cause low spike frequencies. These firing rates are classified into three categories:

1. High rates: 18-250 spikes / second (61% of fibers)
2. Medium rates: 0.5-18 spikes / second (23% of fibers)
3. Low rates: < 0.5 spikes / second (16% of fibers)

The causes and tendencies of neural stimulation in high levels of the auditory system are very complex and have not been studied extensively.

2.4 Physiological Relevant Modeling:

2.4.1 Overview:

The physiological transformation process from the pressure wave that enters the pinna to the electrical signal in the IHCs and AN is a difficult task for modelers to accurately represent [10]. The transmission and transformation of the auditory signal from one physiological auditory structure to another must be considered when constructing a physiological relevant model. Computational models are utilized to understand and visualize the signal transformations between successive structures in the auditory system [10]. Physiological relevant computational models can be constructed by utilizing measurements of vibrational tendencies of the eardrum,

stapes, and basilar membrane; and their correlation to electrical activity in hair cells. The measurement of electrical activity in IHCs can then be correlated to the action potentials within the auditory nerve. The trends of these physiological measurements can be compared to the inputs and outputs at certain stages of the auditory model to gauge the success of the model [10]. For one frequency characteristic position on the basilar membrane, the following quantifiable data can be related to each other: displacement (meters) of the stapes, displacement (meters) of the basilar membrane, voltage of inner hair cell, probability of vesicular glutamate release, and voltage of the resulting action potentials in the auditory nerve fiber [10].

Auditory models must account for the frequency selectivity of the basilar membrane so that only some hair cells and auditory fibers are affected by an incoming frequency, in a physiologically relevant way.

2.4.2 Nonlinearity of the Auditory System:

In linear systems, there is a proportional relationship between the input signal and output signal that can be correlated by constant value. This would result in a larger input signal producing a correspondingly larger output signal. However, the signal propagation of the auditory system is mostly nonlinear. The peripheral auditory system experiences the nonlinear relationship between the basilar membrane's displacement and the neurotransmitter release in the haircell [21]. Another nonlinear process of two-tone suppression occurs when a neuron's response to a particular tone is dampened by the presence of another tone, even when the second tone does not actively stimulate the neuron itself [7].

2.4.3 The Outer Ear:

The raw sound wave received by the body is first filtered by the head and outer ear. This filtering process can be modeled by the head-related transfer function (HRTF) [10]. An

individual's anthropometric uniqueness and size of their head, pinna, and ear canal can alter the way their head and outer ear process sound. In the modeling scope, HRTFs can be related to general head-related impulse response (HRIR) data. HRIR is measured by a microphone located near the eardrum or along the ear canal. This sinusoidal HRIR response at the eardrum can be compared to the raw sinusoidal vibration wave to determine the HRTF. This relationship is linear, and therefore Fourier transforms are able to relate HTRFs and HRIRs [10].

2.4.4 The Middle Ear:

The middle ear absorbs acoustic energy from differing air pressures and transmits it to pressure differentials over biological fluids and membranes in a linear signal processing method by vibration propagation [10]. The transfer function between the input and output of the middle ear is represented by a ratio measured in decibels as a function of frequency [10]. The vibrational displacement of the stapes directly relates to the force it exerts on the oval membrane, corresponding to the pressure differential between the oval and round membranes [10]. Therefore, the frequency transfer function of the middle ear can be characterized by a stapes displacement or velocity versus frequency function. While experiencing pure tone signals, the stapes velocity and displacement are related in the equation shown below:

$$v = 2\Pi f d$$

Traditionally, the middle ear is modeled through analog electrical circuits [7].

2.4.5 Basilar Membrane:

The stimulus of the basilar membrane can be quantified by its displacement and velocity in respect to its resting position [10]. Positions along the basilar membrane react with a maximum displacement and velocity at their characteristic frequencies. In this way, the mechanical properties of the basilar membrane act as a bandpass filter to be responsive to its

characteristic frequency. The varying mechanical properties along the basilar membrane create a series of overlapping frequency filters [10]. The filters that the basilar membrane replicates are nonlinear and asymmetric.

The asymmetric nature of these filters are evident in that the magnitude of BM displacement is less for frequencies above that location's characteristic frequency than frequencies below its characteristic frequency. The basilar membrane exhibits nonlinear gain responses, higher rates at low levels compared to high sound levels. Two-tone suppression is also experienced by the BM and contributes to its nonlinearity. Two tones primary to each other can also combine and cause distortion in audio perception by changing the location where the characteristic frequency occurs on the BM [10]. The asymmetric and nonlinear nature of the basilar membrane requires accurate physiological representation to be useful in consecutive models of the auditory nerve used in parallel.

2.5 Basilar Membrane Models:

The basilar membrane can be modeled in terms of its displacement and velocity across its length depending on the varying frequency of signals. In the scope of a model for the auditory system, the basilar membrane can be accurately modeled using digital filter signal processing [10].

2.5.1 Goldstein's Multiple BandPass NonLinear Model:

Goldstein's Multiple BandPass NonLinear (MBPNL) model is used to represent the "complex linear phenomena such as compression, suppression, distortion, and simple tone interference" [11]. The MBPNL model accomplishes this using a series of narrowly tuned

bandpass filters followed by a more broadly tuned bandpass filter shown in *Figure 3* below [10].

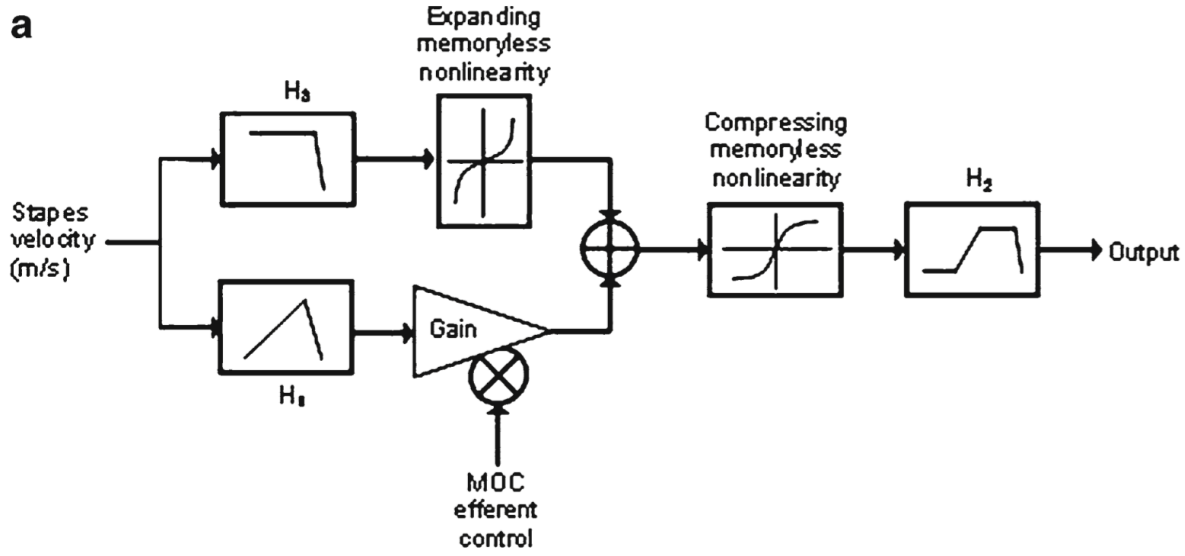


Figure 3: Goldstein's Multiple BandPass NonLinear Model [11]

The structure of this model is successful in reproducing the cycling, nonlinear interactions between a location on the basilar membrane's characteristic frequency and a more intense tone using simple-tone interaction [10]. A series of interacting MBPNL filters is able to reproduce the signal processing capabilities of the sections of the cochlear and account for the propagation of combining tones [10]..

2.5.2 The Gammatone Filter:

Gammatone filters were first utilized for modeling the impulse response of auditory nerve fibers [10]. The gammatone filter is composed of two main parts: a carrier tone that is equal to the characteristic frequency of that location on the BM, and a gamma-distribution function that calculates the impulse response [10]. A series of parallel gammatone filters have been used to approximate the frequency responses of the BM and other frequency dependent physiological systems [10]. However, gammatone filters are linear and therefore do not have the capacity to accurately simulate the nonlinearity of the BM on its own. The implementation of nonlinearly

spaced gammatone filters using the Mel's scale can account for its nonlinear characteristic frequencies (described in further detail in the Methods section). .

2.6 Hair Cell Modeling:

As previously discussed, the inner hair cells (IHC) transduce the mechanical movement of the basilar membrane into electrochemical signals that affect the auditory nerve. This occurs via current due to the movement of K+ ions across the membranes of the IHC. The transduction of the mechanical to electrical signal is nonlinear in two aspects: one related to the saturation of the current, and the other to the membrane's conductance.

Previously, biophysical models of IHC did not include time- and voltage-dependent potassium currents. Also, it was rare to see the models compared to *in vivo* or *in vitro* characteristics. A model created by Lopez-Poveda and Eustaquio-Martín aimed to reproduce in vitro characteristics while considering the time- and voltage-dependent parameters of the IHC membranes' conductance [12]. Their model simplifies the process by assuming the IHC potential is determined by one inward flow of K+ current due to cilial deflections and two outward K+ currents that regulate the cells' potential, disregarding the role of other ions. A figure of their model can be seen in *Figure 4*.

In this diagram, g_A is the K+ conductance of the membrane, $g_{K,f}$ and $g_{K,s}$ are the outward K+ currents, $u(t)$ is the displacement of the stereocilia as a function of time, and V_M is the membrane potential, or the difference between V and V_{OC}, the intra- and extracellular potentials.

For the purposes of our project, the IHC model needs to respond to a simulated stereocilia disturbance and then stimulate the corresponding neuron in the auditory nerve. This model will

likely be closely tied to the basilar membrane model that is developed, and have to respond to a frequency according to its place on the basilar membrane.

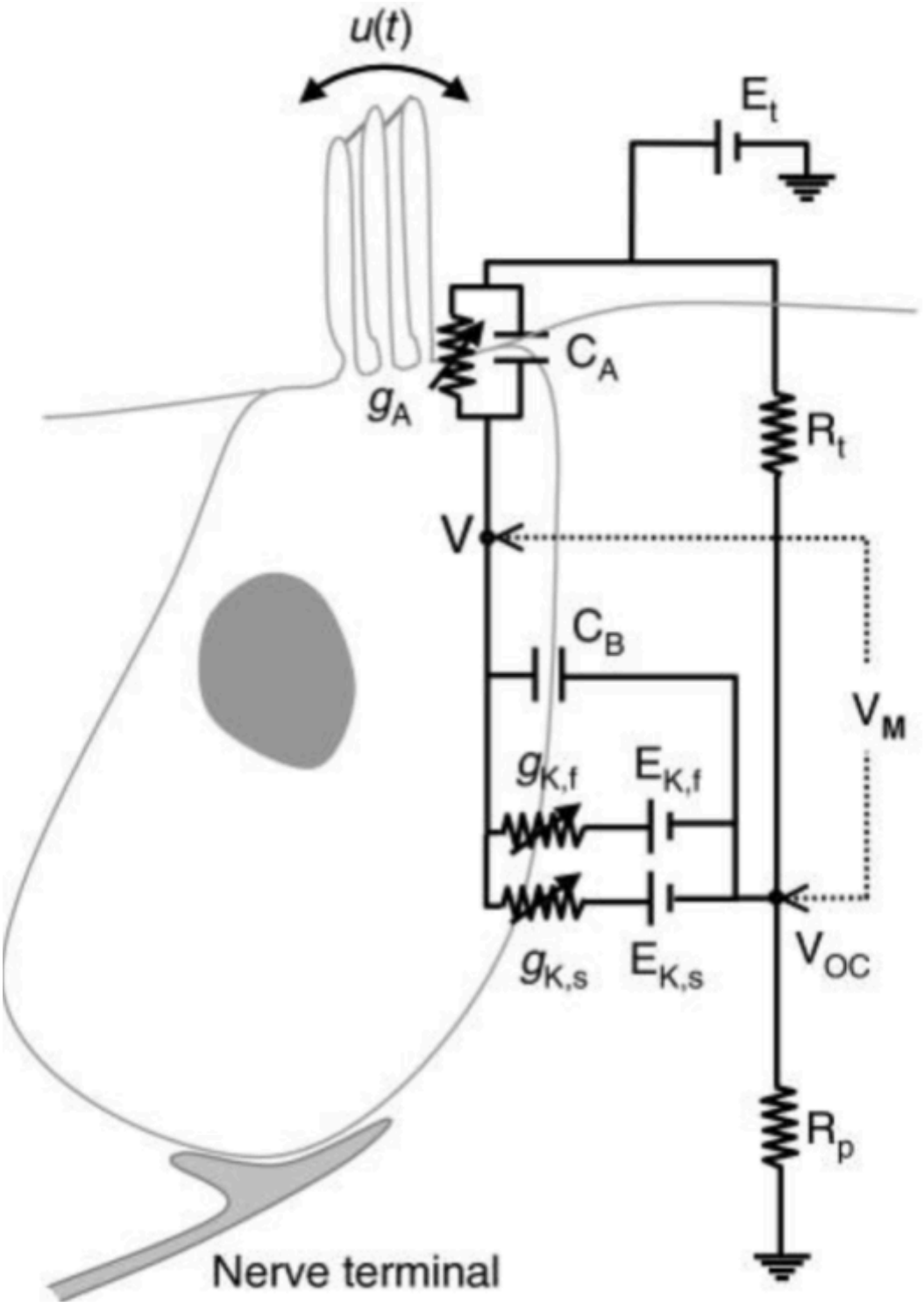


Figure 4: A circuit diagram of the Lopez-Poveda-Eustaquio-Martín model [12].

2.7 Auditory Nerve Modeling:

The auditory nerve (AN) and its behavior in response to stimuli is important for understanding how the auditory system processes sound. The encoding process is important to this project and is why it is focused on through the NEF. Creating an auditory nerve model is complex and is why models usually fall into one of Marr's levels of analysis as discussed previously. The most granular models aim to focus on biological and biophysical processes. Then a step up algorithmically, they take a simplified approach to represent the AN responses without getting into the biophysical details. Lastly and most abstract, computational models aim to be accurate but often leave out the most granular details and are used to be quick and efficient. These are the types of AN models and there are vast differences between each one of them.

The most granular models that are made to be faithful to biological and biophysical details often get into how each individual neuron in the system responds to a sound. These models are seeking to simulate the physiological process and provide an understanding of the neural mechanisms responsible for encoding and transferring the signals. One example of this type of model is the Hodgkin-Huxley model [13].

His model is based on the ideas that electrical behavior of neurons can be explained by the flow of ions across the neuronal membrane. The focus is on three key ion channels: sodium (NA+), potassium (K+), and leaky channels.

The model incorporates voltage dependent conductances for sodium and potassium ions. These conductances describe how the flow of ions through channels in the neuron's membrane depends on the membrane's voltage. When a stimulus depolarizes the neuron's membrane potential, the sodium channels open first, allowing an influx of sodium ions into the cell. This influx results in depolarization and the initiation of an action potential. As the membrane

potential increases, the sodium channels begin to become inactive, preventing further sodium influx. As this is happening potassium channels open, and potassium ions flow out of the cell, leading to repolarization. The model does account for the propagation of the action potential along the neuron's axon, taking into consideration the movement of ions at the refractory period during which the neuron cannot fire another action potential. Finally after the action potential, the membrane potential returns to its resting state as potassium ions continue to flow out of the cell through open channels.

The model is extremely granular considering other models available and by including all the biological and biophysical details the model can be used by researchers who are needing a model like this. This is only one type of model and the other levels of abstraction serve as a guide for the other types of models.

As seen in the above model the granularity can become complex quickly and take away from the efficiency or are not relevant to researchers, but they still want to stay grounded and use a more simplified model. This is where the algorithmic models can be useful since they are in the middle of the levels of abstraction. The Heinz model is the middle between the most abstract and least abstract, there is an attempt to balance being biologically accurate and being computationally efficient [3].

The model starts farther out by simulating the cochlear processing and the inner hair cells representations. Then the model uses different nerve fiber spontaneous rates to capture the dynamics of spike generation in response to the stimuli. There is an emphasis on the importance of precise spike timing and rate coding in the auditory nerve responses. It provides a representation of how different fiber types encode temporal features of sound. As stated in the

paper the model can be extended to incorporate forward masking making it valuable for complex sounds.

Lastly, in terms of abstraction, it is worth noting that the Meddis model of the hair cell includes the AN, where its goal is to capture the important aspect of the AN but simplify the biophysical details.

This model focuses on the important perceptual features that are going on in the AN. It also begins farther out in the ear and doesn't only focus only on the AN since it isn't as granular. The model begins by mimicking the cochlear frequency analysis, and then calculates the basilar membrane's response to the sound input. This simulates the response of the inner hair cells and this is where the AN becomes important. Instead of modeling the detailed process of the biophysical processes of spike generation, this model employs a more simple approach. It uses a rectifier and low pass filter to generate "spike trains" that represent the firing of auditory nerve fibers. The effect of the adaptation and refractory periods are taken into account and by the end the model produces information about the intensity and timing of the spikes, which can be related to perceptual features like volume and pitch.

In this model by leaving out the biophysical details the model is much more simple and useful for researchers who are looking for a model where it doesn't gloss over the details of the AN but still incorporates the key points that cannot be missed.

Each of these models include different levels of abstraction and are important depending on the type of research that is being done. The models have pros and cons depending on what is required, where in some cases the granularity is needed, but in other cases that is not important and not relevant. In this project the goal will include a model of the AN and therefore having scope on what other models have done will provide context for this project.

2.8 Conditions of the Auditory System:

There are many conditions which can affect the auditory system. These conditions can affect any part, including the outer ear, inner ear, auditory nerve, or even parts of the brain. For this project, an emphasis is put on speech recognition problems such as noise and age related hearing loss, and tinnitus. A neuromorphic model would help in this context as researchers could gain a better understanding of the role in the peripheral auditory system in these conditions.

2.8.1 Age Related Hearing Loss

Age related hearing loss, formally known as presbycusis, is a condition that impairs the ability to understand high-frequency components of speech[14]. It is the most common form of hearing loss, affecting approximately two-thirds of Americans above the age of 70. This condition begins to show symptoms at around age 60, and continues to progress as the patient grows older. This condition affects speech recognition, impacting the ability for patients to communicate. This condition occurs due to multiple factors, including ageing of different parts of the cochlea, expression of genetics, ototoxicity, noise exposure, and hormonal imbalances[14]. Using a neuromorphic model would greatly benefit research on this condition, specifically in cases where aging of the cochlea is being studied, as the cochlea is a part of the model.

2.8.2 Noise-Induced Hearing Loss

Noise-induced hearing loss is the second most common cause of hearing loss after presbycusis. It is estimated to affect approximately 5% of the world population[15]. This condition occurs due to either prolonged exposure to noise levels of between 75-115dB or immediately after a single close range exposure at 120 to 160dB[15]. These sound exposures

cause larger than normal distortions in the basilar membrane, resulting in cochlear fluid movement which shears the inner hair cells. Patients' personal and work lives are affected by this, as their speech recognition is impaired and their ability to communicate is heavily impacted. A neuromorphic model would help researchers by understanding how the cochlea is impacted by this condition.

2.8.3 Tinnitus

Tinnitus is a condition which causes perception of a sound in patients with no external source. It is often described as ringing, but it can also present as roaring, rushing, or buzzing. Cases of tinnitus vary from patient to patient, as it may or may not be present at all times, with the same intensity at all times, and the sound may be different. Most cases of tinnitus are subjective, as only the patient can hear it. Approximately 10 to 25% of adults are affected by tinnitus.[16] The causes of tinnitus are not fully understood, but researchers have linked it to noise exposure, hearing loss, medications, ear infections, and head or neck injuries. There are treatments for tinnitus such as sound therapy, though this method requires exposure to the sound the patient is hearing, and it can be difficult to determine the exact frequencies of that sound. A neuromorphic model would help researchers understand potential factors of tinnitus that could occur in the peripheral auditory system.

2.9 Different Models

In the field of auditory modeling, several influential models and theories have expanded our understanding of how humans perceive sound. To provide context for the project, we will discuss the Ray Meddis MAP model and Heinz Auditory Nerve (AN) model. The MAP model is

the closest to what we are developing and is the biggest competitor hence why we want to compare its qualities to this project.

Ray Meddis developed a model of auditory processing known as the "MATLAB Auditory Periphery," model (MAP)[2]. The motivation behind this model is to provide a tool to researchers who want to explore the relationship between physiology of the auditory system and the perception of sound both in normal and impaired hearing, meant to be an all encompassing model that ties all of Meddis models together. The creation of such a model is extremely time consuming and therefore an “off the shelf” model would provide a much more accessible option for researchers to use. The model itself provides the ability to do things like pitch matching, simulating hearing disorders, and can be used in various students such as studying how speech recognition works in quiet and noisy environments. This is the most comprehensive model that tries to fill a similar need as this project but due to a few differences in our mission statement our model strives to achieve a different goal.

2.10 Marr’s Levels of Abstraction

Marr’s Tri-Level Hypothesis defines these levels of abstraction as the computational level, the algorithmic level, and the implementational level. The computational level represents how the system reaches its goal and the purpose of the process [6]. The algorithmic level uses and manipulates representations to process data to achieve the model’s goal. The implementational level is the physical and biological reality of how the system works and propagates signals. For example, the release of neurotransmitters causes an action potential in a neuron. The Meddis model of the inner hair cells (IHCs) and the Hodgkin-Huxley model of

neural communication exist on the implementational level. The Heinz model of the auditory nerve, and models of the central nervous auditory system exist at the algorithmic level.

2.11 Benefits and Knowledge gaps

Existing auditory models have significantly contributed to our understanding of auditory processing, but they often operate at different levels of abstraction as seen in the numerous examples discussed. While these models provide valuable insights, there is a notable gap when it comes to a comprehensive neurophysiological sound model. The challenge arises from the different levels of abstraction in these models, making it difficult to create a unified model that accurately simulates the entire auditory process and gracefully bridging the gap from one level of abstraction to another. This is what the Neural Engineering Framework accomplishes extremely well and is why it is a major focus for this project. This is also why the MAP model was discussed in the previous section since without the NEF it creates a model of the ear that attempts to bridge these gaps in the levels of abstraction. The problem however is that since older models aren't directly compatible they will not always be faithful to how the different sections actually work together. This is useful for researchers who have a dedicated use case but this model does not provide a neuromorphic accurate representation of the correct parts.

To achieve a more neurophysiologically accurate sound model, it is essential to correctly model the processes occurring at the level of auditory nerve fibers, this is where the NEF will come into play. Current models often simplify or overlook the complexities in the auditory nerve and its encoding and decoding. Bridging this gap would not only enhance our understanding of auditory perception but also open doors to more neurobiologically inspired sound models. Such a model would bridge the varying levels of abstraction to provide a more faithful representation of

how the human auditory system processes sound. This would enable more precise and effective simulations of auditory experiences for researchers to make direct use of and that is the goal of this project and is why that is a priority.

2.12 Neuromorphic Computing

To create a model that takes these limitations into account, it is necessary to utilize neuromorphic computing. Neuromorphic computing is a computational approach inspired by the brain. It functions in such a way that mimics the function of neurons and synapses. The first instance of neuromorphic computing was in an analog model of the retina in 1988. This model was composed of a silicon chip with a resistive array network, and its functionality modeled the processing that occurs in the receptors and outer plexiform layer. [2] Although the model was not precise enough to be used in experiments to determine the true importance of the retina, it created discourse about using a different style of computation, resulting in the birth of neuromorphic computing.

Neuromorphic computers differ from traditional Von Neumann computers, which utilize CPUs for processing and separate memory units for storing data and instructions. Instead, this method of computing takes inspiration from the construction of neurons and synapses to store and process this information. This is accomplished by receiving spikes as input, which can then be used to decode numerical information, as opposed to the binary values encoding information for Von Neumann computers. The result of this is fundamental differences in operation. These differences pose a number of benefits for creating a model of the auditory system including parallel operation, collocated processing and memory, inherent scalability, event-driven computation, and stochasticity[17].

In a software-based neuromorphic model, processing of information would be event-based and use spike based input. The code that makes up the model would be organized and made to function in a way that replicates the functionality of the biological parts of the peripheral auditory system. Due to the inherent cost and complexity of neuromorphic computers, our model will be utilizing a Von Neumann computer. In order to replicate the function of the auditory system, we will be utilizing a neural framework in MATLAB.

2.13 Neural Engineering Framework

2.13.1 Overview

The book *Neural Engineering* by Chris Eliasmith and Charles Anderson discusses how to computationally represent a neural system in order to more easily research and understand them[6]. They intended their book to be applicable to different kinds of systems, and thus explain the basis for their model and methodology on how to implement it with the systems. In order to guide the neural engineering framework (NEF) model, there are 4 principles that must be considered. First, is that the system must be represented by both non-linear encoding and linear decoding. This principle assumes that the encoding device of a neural signal is independent of the one that decodes it. Although neurons may not always use a linear decoder, any change in the decoding of a neuron directly affects the decoding and encoding of neurons in the nearby population. The second principle states that neural transformations are based on variables in the neuron population, and an alternately weighted linear decoding. Nonlinear transformations are also very common in neural processes and can even be predicted using the neural network itself. The third principle states that neurons should be characterized with control state variables, so they can be analyzed with control theory. This allows for neural analysis in dynamic

neurobiological systems. Finally, the last principle states that noise must be taken into account when analyzing neural systems. This helps narrow down the uses of a particular system and predict its efficacy.

2.13.2 Neuron Representation

When creating a model of any kind, from physics to biology, it is important to consider how its characteristics will be represented in a meaningful way. A simple but physiologically relevant way to represent neurons can be with a digital computer. The neuron's function of encoding a neural signal can be compared to an analog-to-digital converter (ADC). In this example, there are two steps: Converting a scalar to a series of values (0 or 1), and decoding that series in order to produce a meaningful output.

In order to translate this to neurobiological systems, it is assumed the decoding process is linear. Also, rather than converting a scalar into a temperature code, neural systems convert scalars into firing rates. The relationship between these values is called the neuron's "tuning curve", which is determined by processing input signals producing a current, and then generating a voltage from the soma of the targeted neuron; these are represented by $J(x)$, the current in the neuron, and $G[J(x)]$ respectively. Although currents do not directly cause a neuron to fire, this is an assumption made though the text to simplify the model. Therefore, the encoding and decoding processes for the firing rate and signals of biological systems are

$$a_i(x) = G_i[\alpha_i x + J^{bias}] \quad \text{(Eq. 1)}$$

and

$$\hat{x} = \sum_i a_i(x)\phi_i \quad \text{(Eq. 2)}$$

where

$$\phi = \Gamma^{-1} \Upsilon. \quad \text{(Eq. 3)}$$

In the weighting function, Eq. 5,

$$\Gamma_{ij} = < a_i(x) a_j(x) >_x \quad \text{(Eq. 4)}$$

and

$$\Upsilon_j = < a_j(x) x >_x. \quad \text{(Eq. 5)}$$

In (**Eq. 1**), α represents the gain, as well as a parameter that converts the units of the input scalar, x, and scales the scalar to the tuning curve of a neuron within $G[\bullet]$. The weighting function, The decoding function, $G[\bullet]$, is determined by the properties of the neuron itself.

2.13.3 Temporal Representation

The leaky integrate-and-fire (LIF) neuron is a model used to emulate voltage spikes produced by real neurons. It is a simplified version of the Hodgkin-Huxley model, so it is applicable to neurons all over the human body. In this model, the voltage within the neuron is constantly increasing, until it reaches the threshold, V_{th}, where it spikes for only an instant. It then ceases activity for a time, $\boldsymbol{\tau}_{\text{ref}}$, and then continues to build voltage until it reaches the threshold again. It can also be compared to a simple RC circuit coupled with a short-circuit switch. When the voltage across the circuit reaches V_{th}, the short-circuit switch flips, generating a spike in voltage for a brief moment, before dropping to 0 **V**, and opening after $\boldsymbol{\tau}_{\text{ref}}$. The process can be seen below, in *Figure 5*. This model is physiologically applicable for a few reasons. First, the spike represented in the LIF neuron is similar to one that is produced by a real neuron.

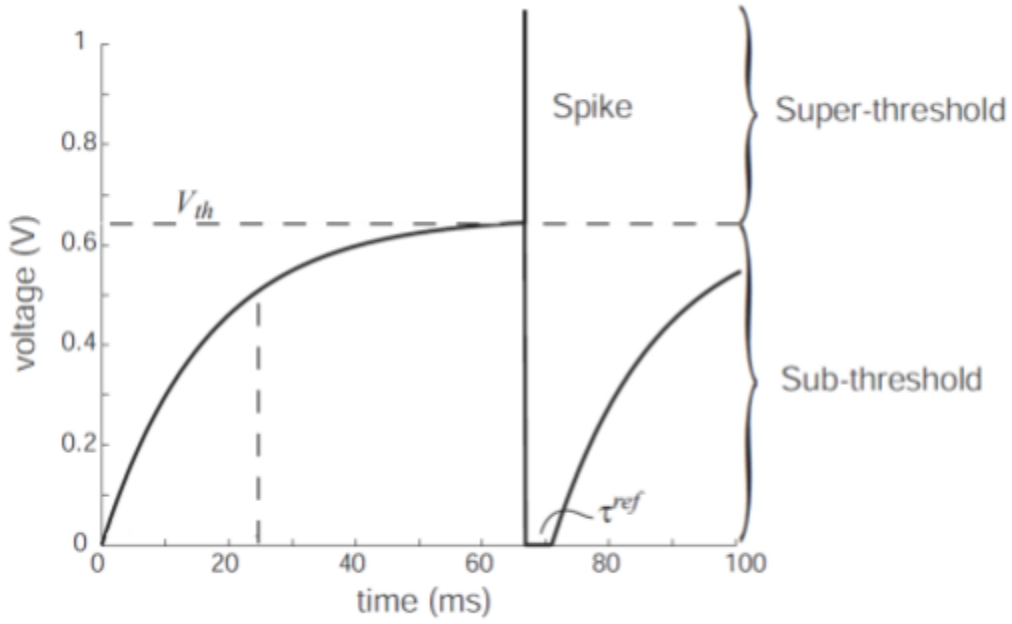


Figure 5: Spiking behavior of LIF neurons from the Neural Engineering Framework [6]

Second, the resting period after a spike is also found in neurons. Finally, this model can be easily represented in a physical circuit. Although there are aspects of this model that are not physiologically relevant, the benefits of the model greatly outweigh them.

Using the LIF model, the firing rate of a neuron can be found in terms of the input current, J_M:

$$a(J_M) = \frac{1}{\tau^{ref} - \tau^{RC} ln(1 - \frac{J_{th}}{J_M(x)})} \quad \text{(Eq. 6)}$$

2.13.4 Decoding Spikes

In the world of representing the timing of neural signals, there is discussion about whether the signal should be processed by a rate code or timing code. However, Eliasmith and Anderson argue that it does not matter much, as they are both time-dependent codes, and developed a method that accounts for both types of codes.

For a signal function based on time, *x(t)*, an encoder will transform it into a series of spikes, $\delta(t\text{-}t_n)$. For peripheral nerves, using a decoder like (**Eq. 2**) will return a signal, $\hat{x}(t)$, that is similar to the original signal function. The auditory nerve, however, is part of the central nervous system. This means the signal will go through an additional encoding into spikes, and then an additional decoding which produces a complex function of *x(t)*. This additional decoding helps characterize the previous encoding steps. Also, because neurons are not single units that operate independently, the decoding process for one neuron is directly related to the encoding process of a neighboring one.

These paired cells can be represented by a push-pull amplifier. This essentially means for a given pair of neurons, the signal can be split in half between the two, and combined later to produce a signal. For each spike train, spikes that are close together represent an increase in amplitude which can be positive or negative based on the neuron. This is possible because of opponency, where two neighboring neurons have opposite tuning curves. This signal ends up being more linear, more symmetrical, and more efficient. When these opposing signals are converted into spikes, or impulses, the approximation of the original signal can be represented as:

$$\widehat{x}(t) = \sum_{i=1}^{N} h(t - t_n) = \sum_{i=1}^{N} a_i(x(t))\phi_i(t) \qquad \text{(Eq. 7)}$$

where the function *h(t)* is a linear decoder, made up of the firing rate and the encoding weight. The final expression comes from the NEF [6].

3. Methods

3.1 Design Criteria

The creation of a complete and physiologically relevant model of the auditory system needs a precise approach that follows specific design criteria. These criteria are important in ensuring that the model not only mirrors the anatomical structure of the auditory system but also simulates its behavioral functions. Our set of criteria, accuracy in anatomical structure and behavioral resemblance, usability as a research tool, real-time signal processing capability, extensibility for future modifications, and efficiency in computational performance—forms the foundation of our methodology.

Biological plausibility and accuracy are our primary objective since we want to create a model that represents the physiological processes of the auditory system. This criteria is essential for simulating the auditory signal processing done by the basilar membrane, inner hair cells, and auditory nerve. By prioritizing biological accuracy, we ensure that the model can accurately simulate normal auditory functions and various pathologies, therefore serving as a valuable tool for researching hearing impairments and the effects of interventions such as cochlear implants.

Usability as a research tool is an important criteria due to being a major motivator for doing this project. Researchers being able to see what is going on and understand and use what is presented is important to the design of the model. Making a model that is overly complex does not benefit researchers, this is why a simple and correctly thought out model is crucial.

Computational efficiency remains a significant concern. A model that is computationally demanding may limit its accessibility and usability, particularly for researchers and clinicians who may not have access to high-performance computing resources. Therefore, efficiency is

crucial for facilitating broader adoption and application of the model, albeit secondary to the criterion of biological accuracy.

The ability to easily modify and understand the model is crucial for its longevity and relevance, especially given the rapid pace of advancements in auditory science and computational modeling techniques. Extensibility ensures that future research can build upon the model without starting from scratch, allowing for the incorporation of new findings and methodologies.

3.2 Model Inputs and Outputs

The neuromorphic model of the human auditory system utilizes a one dimensional vector as an input signal. The input signal represents a sound wave signal that varies in magnitude over the time domain. The input signal can be a raw audio waveform file read using MATLAB's function audioread, or an arbitrary time varying input signal or sinusoid. The input signal is passed through the basilar membrane filters, inner hair cell filters, and auditory nerve populations. The signal is reconstructed after being filtered by the basilar membrane and inner hair cell portion of the model. These signals can be played as audio outputs of the model using MATLAB's *soundsc* function. The time varying function of these processed audio waveforms are also outputted with a visual graph.

The auditory nerve fibers in one population encode the input voltage signal received from their corresponding inner hair cell. These encoded signals are then decoded for each population of neurons and reconstructed into a final decoded signal. The final decoded signal can be compared to the initial input signal and the processed output signals of the basilar membrane and

inner hair cells. These different signals can then be compared graphically and by listening to their audio outputs.

3.3 Model Structure

As mentioned previously, our model aims to represent three major portions of the auditory system: the basilar membrane, the inner hair cells, and the auditory nerve. This will make our model physiologically relevant for researchers. A bank of gammatone filters has been chosen to model the basilar membrane, due to the frequency specificity of the gammatone filters. A leaky integrate-and-fire neuron structure has been chosen and adapted from the Neural Engineering Framework to represent both the inner hair cells and the auditory nerve neuron populations. For each spot on the basilar membrane modeled, there is a corresponding array of cells that will take the output of each individual filter. Subsequently, for each inner hair cell modeled, there is a population of neurons that receive the output from each inner hair cell.

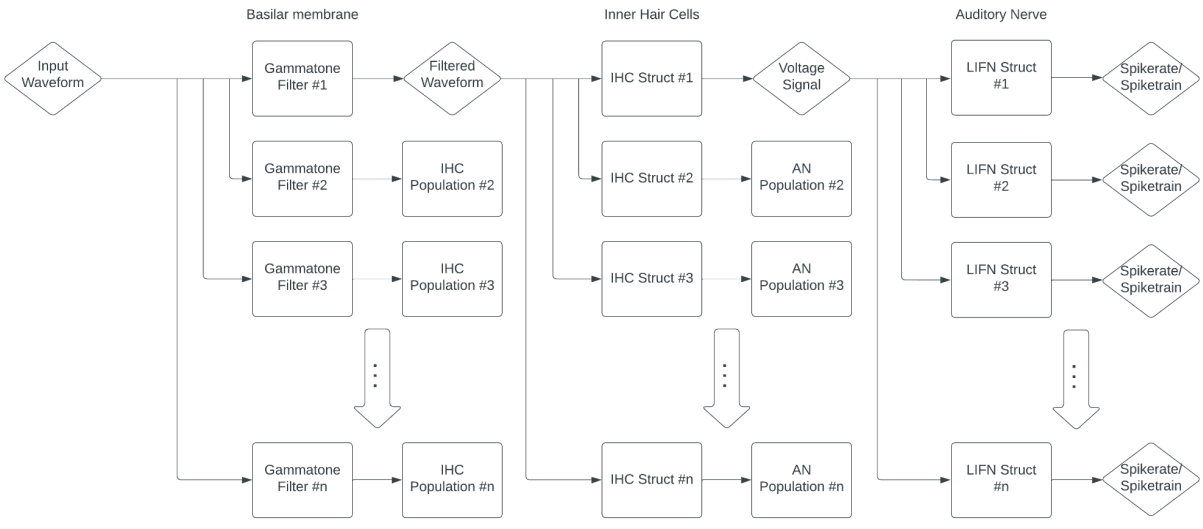


Figure 6. Complete block diagram of the auditory processing model.

To represent the response of the basilar membrane to frequency-specific signals, we designed a bank of Gammatone Filters. This choice was made to keep it physiologically relevant by aligning it with the logarithmic, nonlinear spacing of frequency specificity on the basilar membrane. Our adaptation incorporated Goldstein's Multiple Bandpass nonlinear model to refine the scale and establish frequency bins corresponding to distinct locations on the basilar membrane[18].

For modeling the inner hair cells, reference was taken from the Meddis Model, previously recognized as a dominant framework in auditory research. This selection prioritized physiological relevance, as well as giving the option for scalability in research. Integrating concepts from the Neural Engineering Framework, we tailored the model to our needs, focusing on LIF neurons. To model these inner hair cells, we created a function that created MATLAB structures that contained variables like the membrane voltage and background current of the IHC, a scaling parameter that ensures physiological relevance, the encoding weight, the sampling rate, RC time constant, and voltage leak resistance. This function took in the sampling rate, an encoding weight (phi = 1) and randomly generated background current and scaling parameter values as parameters. The structure for auditory nerves is similar but includes more values: spike state, refractory period, current threshold, and voltage thresholds. These nerve structures were contained in a 3D matrix of cells based on the number of filters, number of inner hair cells per filter, and number of neurons per cell.

3.3 Basilar Membrane Model

The basilar membrane is modeled using a series of gammatone filter banks. Each gammatone filter represents a frequency sensitive location on the basilar membrane with a unique characteristic frequency. The parameters for the gammatone filters are set between the frequency ranges of 200 to 8,000 Hertz. This range was selected to be used for the audible frequencies of human speech and sound and it is the physiological sensitivity range of the basilar membrane from its base to its apex. This sensitivity range can be changed in the model, to allow researchers the ability to vary the sensitivity of the model from 0 to 20,000 Hertz which is the complete frequency range of the basilar membrane.

The boundaries of this frequency domain are converted from Hertz to Mels. The Mels scale provides an accurate representation of the nonlinear frequency sensitivity of the basilar membrane along its length.[19] The frequency domain in Mels is linearly spaced between the boundaries for each of the gammatone filters to be created. The equation utilized to convert hertz to Mels and then back into frequency is shown below:

$$Mel = \frac{1}{log(2)} \cdot log(1 + \frac{Hz}{1000}) \cdot 1000$$; [20].

The resulting evenly spaced vector is converted back to Hertz to represent the nonlinear frequency sensitivity of the basilar membrane. The nonlinear spaced vector in Hertz is then indexed to retrieve the center frequency values for each gammatone filter created. The bandwidth for each gammatone filter also follows the same nonlinearity and is the corresponding index multiplied by a constant of 0.25.

The amount of gammatone filters created and utilized by the model is initiated in the *NumFilters* variable. A filter is created for the amount specified by *NumFilters* using a for loop that runs for the amount of filters being created. This for loop uses the indexed center frequency

and bandwidth values and the sampling rate in the function *makefiltgt.m* to create a unique gammatone filter. The function *makefiltgt.m* uses real time signal processing to analyze one time point at a time that is dependent on its previous output. Each gammatone filter is stored in the array of cells *F*.

Once the gammatone filters are created and stored in *F*, each filter is applied to the input signal. This ensures the physiologic relevancy of real time signal processing. One individual time point at a time is processed by the function. This function applies real time signal processing through the indexed gammatone filter and stores the updated value for each time point in *SignalIn* in the vector *y_realtime*. The *y_realtime* output for every filter is stored in a corresponding indexed cell in the array of cells *FilterValues*. This process is repeated with a cascading set of identical basilar membrane gammatone filters where the *FilterValues* output is filtered through its corresponding center frequency gammatone filter bank. The new output is stored in the cell *FilterValues2.* The output of each of the filters are then plotted against the original signal. These processed signals are also summed together and compared to the original signal.

3.4 Inner Hair Cells

In order to encode the incoming signal, we created a matrix to contain the time for each neuron and cell (TT), and one to store the stimulus of each neuron (XX). Then, the voltage from each inner hair cell is calculated based on the input signal to the IHC. This signal varies from each inner hair cell population depending on what section of the basilar membrane it is in. This calculation is based on various factors that are determined at the creation of each IHC.

Upon initialization there are a few inputs that are randomly varied to create an illusion of randomness for the cell creating. This is to help model the human ear; each cell is not going to be exactly the same in the way it conducts current.

Below is the structure responsible for the IHC.

```
function IHC = makeIHC(alpha, phi, J_bias, fs, tau_RC)
% Create an Inner Hair Cell (IHC) model
% Inputs:
%   alpha - Scaling parameter for IHC
%   phi - Encoding weight for IHC
%   J_bias - Bias or background current for IHC
%   fs - Sampling rate in Hz
% Outputs:
%   IHC - A struct representing the IHC
%     'V' - IHC membrane voltage
%     'alpha' - Scaling parameter
%     'phi' - Encoding weight
%     'J_bias' - Bias or background current
%     'fs' - Sampling rate
%     'tau_RC' - RC time constant
%     'R' - Leak resistance

IHC = struct(...
    'V', 0, ...
    'alpha', alpha, ...
    'phi', phi, ...
    'J_bias', J_bias, ...
    'fs', fs, ...
    'R',10,...
    'tau_RC', tau_RC...
    );

return
%eof
```

Figure 7: Structure of the modeled IHC.

As seen there are 5 hyperparameters that can be controlled upon initialization, these are the hyperparameters that are used to simulate randomness in the IHCs and each plays a role in how the electric current is calculated by each of them. The processing is displayed below and there are a few steps when it comes to going from input signal to outputting the voltage that the neurons are going to use.

```
function IHC = updateIHC(input_signal, IHC)
% Update the IHC model based on an input signal
% Inputs:
%   input_signal - A scalar representing the input signal to the IHC
%   IHC - A struct representing the IHC model
% Outputs:
%   IHC - The updated IHC model
%     'V' - IHC membrane voltage

tau = 1 / IHC.fs; % Simulation time step in seconds

J = IHC.alpha * IHC.phi * input_signal + IHC.J_bias; % Calculate the current

% Update the membrane voltage of the IHC using Euler integration
Vd = -(1 / IHC.tau_RC) * (IHC.V - J * IHC.R);
IHC.V = IHC.V + Vd * tau;

return
%eof
```

Figure 8: Function used to update IHC.

The updateIHC function has been modeled after the NEF where a lot of the chemical complexities of the human ear are lost in low level implementations but the computation is not just simple signal processing. This was the goal of the NEF to bridge the gap between implementation and algorithmic, and that is why this code was based on that same logic. Similar to the LIF neurons the updateIHC runs tick by tick instead of computation in chunks.

In the first calculation of J the hyperparameters from initialization become important because they are responsible for how strong the current is going to be. As seen by the use of Matlab's dot notation and object oriented programming, each IHC has different alpha, phi, and J_bias values that are included in this basic computation.

Upon finishing the current calculation the membrane voltage of the IHC needs to be calculated using Euler integration. The basic idea behind Euler integration is to approximate the next value of the variable by adding the current value to the product of the rate of change of that variable and the time step. This approach is called "forward Euler," and while it is simple to implement, it may introduce numerical errors if the floating point numbers are extremely small,

or if the sampling rate is relatively low. Then after the voltage has been calculated the IHC.v is updated and the function is finished.

This file is very similar to the code from the LIF neuron but without the added complexity of the spiking and refractory periods. This is a much more basic version of a neuron and that is exactly how it has been designed to be. In the human ear, inner hair cells are along the basilar membrane to pick up on vibrations and create voltage that is passed to more complex cells like the ones in the auditory nerve. This is exactly what our model reflects with the IHCs being a more basic version of the LIF neurons, since the neurons have more nuance and complexity to them.

3.5 Auditory Nerve Model

The auditory nerves in the model were based off of the LIF neurons outlined in the Neural Engineering Framework because they are physiologically relevant and easy to model. There is a function, makeLIFN, that was used to generate the neurons. Randomly generated encoding weights, scaling coefficients, and background currents are used as inputs. This function is similar to makeIHC in that it creates a structure to contain the neuron's properties and values; the difference lies in the properties included. The auditory nerves contain all of the same properties as the IHCs, but also have values for the state (sub- or super-threshold), current threshold, refractory period, refractory clock, and voltage threshold.

3.5.1 Encoding

Nested for loops are used to encode the stimulus value of each neuron based on the voltages of each IHC using the updateLIFn function. This function takes in the physical stimulus

from the inner hair cell, the neuron that is being updated, and the sampling rate. If the neuron is in a super-threshold state, it checks if the refractory period has been reset. If yes, it causes the neuron to "spike". Otherwise, the voltage is set to 0. It then updates the simulation clock or resets it. When it is reset, the neuron returns to a sub-threshold state. If the neuron is in a sub-threshold state, it updates the current using the equation

$$J(x(t)) = \alpha\phi x(t) + J^{bias}$$

and calculates the change in voltage based on equation 4.4

$$\frac{dV}{dt} = -\frac{1}{\tau^{RC}}(V - J_M R)$$

from the Neural Engineering Framework.[6] If the updated voltage is over the voltage threshold, the state is changed. *Figure 6* shows how the voltage of the neuron is modeled by this function. The spikes and their timestamps are then encoded in a 1:Fs array in a 3D matrix of size numFilters x numCells x numNeurons, D. These arrays are also called spike trains.

In order to decode the signal from the stimulus recording, each of the populations of neurons needs an activation function. This was done by our function characterizeLIFn. This function was given to us by Prof. Lammert but was edited to fit our needs. This function takes in an auditory nerve structure, the minimum and maximum values to be accepted, resolution of the values selected, and the sampling rate. It then returns a (*numFilt x numCells*) cell, A, containing a matrix with the spiking rates/activation functions for each neuron in the population, and a (*res x 1*) vector of values to be encoded, X, for the specified neuron. Then, for the number of samples, which is determined by the sampling rate, the neuron is updated using updatelifn, and the spike rate for that neuron is calculated. These outputs are then used to help determine the decoders for each neuron in the function determinedecoders.

DetermineDecoders takes in the activation functions and values to be encoded determined by characterizelifn, and returns a *NumNeurons* x 1 array of each neuron's decoder. This function models (**Eq. 3**), which determines the decoding coefficients for each neuron using the covariance between the neuron's activation functions. ϒ (**Eq. 4**) was calculated by summing the product of the physical values and the spiking rates for each population of neurons. Then, Γ (**Eq. 5**) was calculated by summing the spikerates in each population of neurons, creating a *NumNeurons* x *NumNeurons* matrix. There were a few ways this could be calculated with different MATLAB syntaxes, but we chose to use element-wise operations because it was the easiest to use based on the way the neurons were structured.

3.5.2 Decoding

Before the encoded signal could be with the decoders that had just been calculated, a matrix was created to contain all of the activation functions for each neuron. Then, a function called Decodespiketrain finds these activation functions. It takes in a neuron structure, the neuron's spike train, and the frequency of the input signal. There are two possible routes for decoding. One is a time-based, linear decoding that is more physiologically relevant and models each neuron's postsynaptic current (PSC), and the other is a slightly more optimal, non-linear decoder. Eliasmith and Anderson point out that there is not much of a difference in the optimization of the decoders: the optimal decoding only provides a 5% improvement in information transmission, and the linear decoder has a 6% decrease in information transmission [6]. Therefore, the more physiologically relevant method was chosen. Decodespiketrain creates a vector of the neuron's PSC based on the signal that passed through it using the equation

$$h(t) = e^{(-t/\tau_{RC})}, \quad \text{(Eq. 8)}$$

where τ_{RC} is the time constant hardcoded into the neuron. Then, this vector is convolved with the neuron's spike train. This yields a vector that represents the connection between a neuron's spiking rate and postsynaptic current, or how the neuron reacts to stimuli. This can also be referred to as a neuron's activation function. This vector is twice as long as *Fs*, so the first half of the vector is used as the activation function for the neuron. These activation functions are stored in a *Fs* x *NumNeurons* array per hair cell in a *NumFilters* x *NumCells* cell.

Finally, an estimate of the input signal is calculated with the function DecodeSpikeRate. This function takes in the activation functions created by DecodeSpikeTrain and the decoders found with DetermineDecoders. There were a number of ways this could be done, like iteration over each neuron, over the samples, or inner product. However, the best option was iterating over each neuron using (**Eq. 7**) above for each population of auditory neurons. This results in a *Fs* x *1* estimate for each population of neurons. Later, all of these individual vectors were added together into a single signal to give the complete estimate.

3.6 Parameter Optimization

After the team was able to achieve a basic model the parameter optimization problem began. Attention was directed to optimize the models parameters to achieve an ideal balance between accuracy, computational efficiency, and pseudorandomness. This process was important in ensuring that the model could correctly stimulate and process inputs but also be modeled after real neurons and inner hair cells.

Our approach to optimization changed over time, finally landing on a strategy that identifies the best set of parameters for each individual population of LIF neurons. The three primary parameters that were focused on were the scaling factor (ALPHA), the base phase

(PHI0), and the bias current (J_BIAS). These parameters were selected due to their significant impact on the model's ability to accurately simulate the signal.

The parameter search was encapsulated into a single file called Script_determineparams.m. This function takes the input voltage signal and the sampling frequency (Fs) as inputs, utilizing them to adjust the model's internal settings. The goal of this function is to find a set of 3 values for a neuron population that will properly encode and decode the input voltage signal. The optimization was conducted through an iterative process, systematically varying the values of ALPHA, PHI0, and J_BIAS across a predefined range. For each set of values a scaled down set of 5 LIF neurons were created and used to create an output from the input signal, the model's output was evaluated against a performance metric and the highest performing set of parameters would be saved and used by that single population of neurons in the main script.

The performance of each parameter set was assessed based on one main criteria: the correlation value from the input signal against the output signal. Matlab’s corr function was responsible for this heavy lifting and would return an integer metric in the range from -1 to 1 that when approaching 1 means that the two signals are more correlated. The reason for using this metric is because correlation is good at measuring the strength and direction of a linear relationship between two variables. Other metrics like Mean Squared Error for example (MSE) are better for predictive modeling and are oftentimes used in Machine Learning. Its pros are not as valued in this project and therefore using corr is the best option for this type of problem.

With the new optimized parameters for each population of neurons the model's performance made significant improvements and this jump led to much smoother graphs and less spasmic decided signal behavior. By fine tuning the parameters we are now able to balance

physiological relevance and computational pragmatism. This iterative approach of finding suitable random ranges for these parameters shows the importance of these values and why proper planning needs to be taken when creating these models.

4. Results

4.1 Simulation 1: 2 Hertz Tone

All iterations of Simulation 1 were run for an input sound wave of 2 Hertz over a duration of 0.9 seconds with a sampling rate of 44100 samples per second. This input sound wave is plotted in *Figure 9* below.

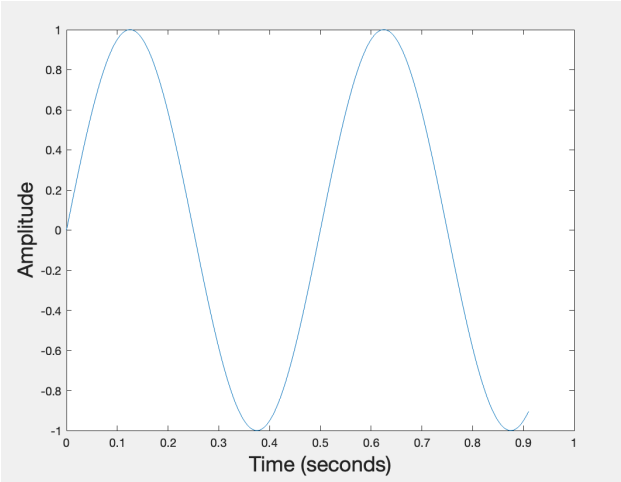


Figure 9: Simulation 1 Input Signal.

Simulation 1A displays the model with the largest amount of elements to validate output correlation and model run time. These performance parameters are linked to the amount of basilar membrane locations, inner hair cells, and neurons in each population that are evaluated between the different simulations within Simulation 1. Every graphical output is plotted for this simulation to highlight all of the data that the model is able to process and present. The results of Simulations 1B-H only present the graphs for the summed basilar membrane filter outputs, summed IHC filter outputs, and the lowpass decoded output. A limited range of graphs were

presented to prevent repetitiveness of the qualitative data that is accurately represented by the outputs of Simulation 1A.

Simulation 1B analyzes the effects of reducing the amount of basilar membrane locations on the output parameters. Simulation 1C analyzes reducing the amount of IHCs on the output parameters. Simulation 1D analyzes reducing the amount of neurons in each population on the output parameters. Simulation 1E analyzes reducing the amount of basilar membrane locations and inner hair cells. Simulation 1F analyzes the model with reduced basilar membrane locations and neurons in each population. Simulation 1G analyzes the model with reduced inner hair cells and less neurons in each population. Finally, Simulation 1H analyzes the lowest scaled model, with reduced amounts of basilar membrane locations, inner hair cells, and neurons.

The exact amount of each model component for each simulation can be seen in *Table 1* below. Each iteration of Simulation 1 accurately compares the effects of changing the quantity of the three element types in the model to understand the importance of their effect on the output.

4.1.1 Simulation 1A: Large Scale Simulation

Simulation 1A was run for the input signal defined in *Figure 9* above in Section 4.1. The model was specified with 10 basilar membrane locations, 3 inner hair cells for each location, and 5 neurons in each neural population. The model produced the graphs and output values shown in the table and figures below.

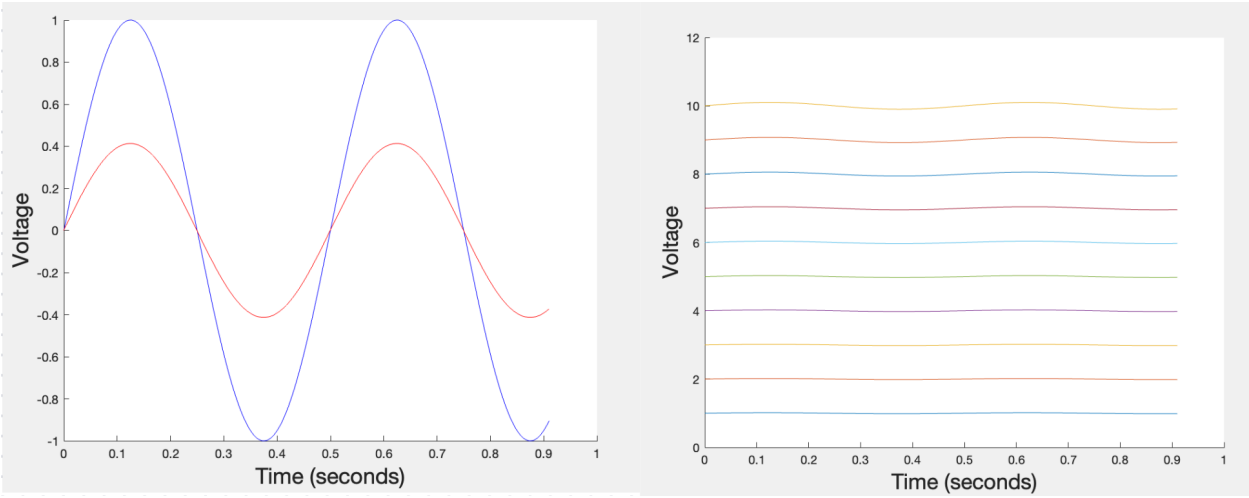


Figure 10: Basilar Membrane Outputs. The graph on the left panel displays the input wave signal in blue, and the summed outputs of each basilar membrane channel in red. The graph on the right panel displays the output of each filter channel after filtering the input signal. The channels outputs at the bottom of the graph represent locations closer to the base of the basilar membrane (lower center frequency) and channels closer to the apex are higher on the graph (higher center frequency).

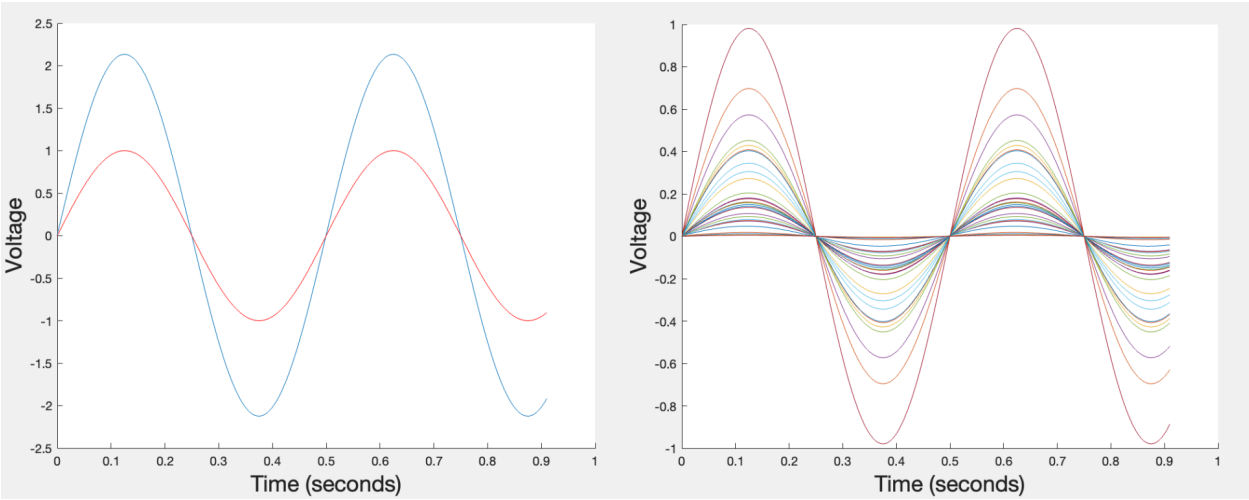


Figure 11: Inner Hair Cell Outputs. The graph on the left panel displays the input waveform in red, and the summed outputs of each inner hair cell in blue. The complete output spectrum of each inner hair cell is shown in the right panel.

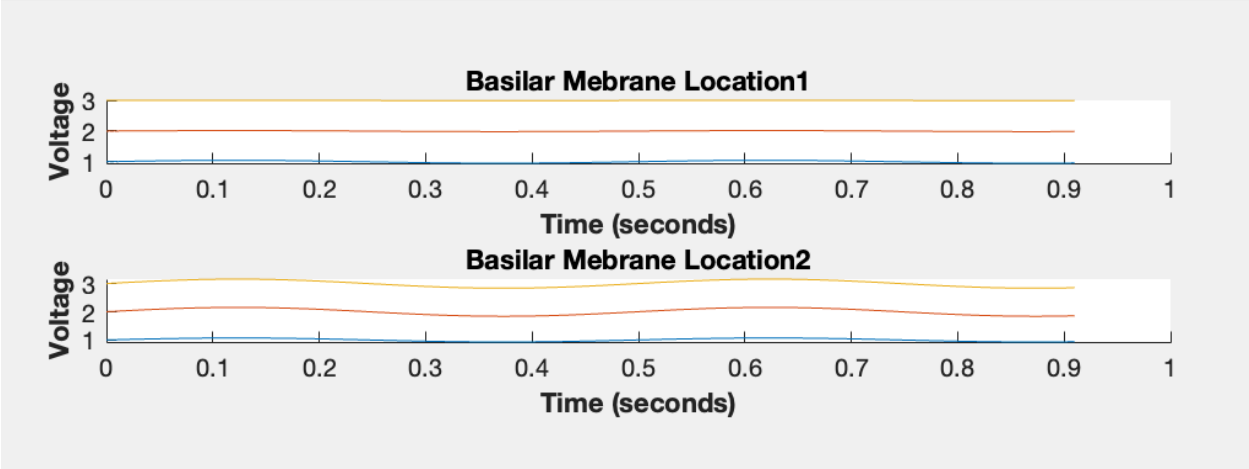


Figure 12: IHC Outputs for Corresponding BM Location. The graph displays the inner hair cell outputs that correspond to the first two basilar membrane locations for Simulation 1A. Both Location 1 and Location 2 have three inner hair cells receiving the same input signal from that location's output.

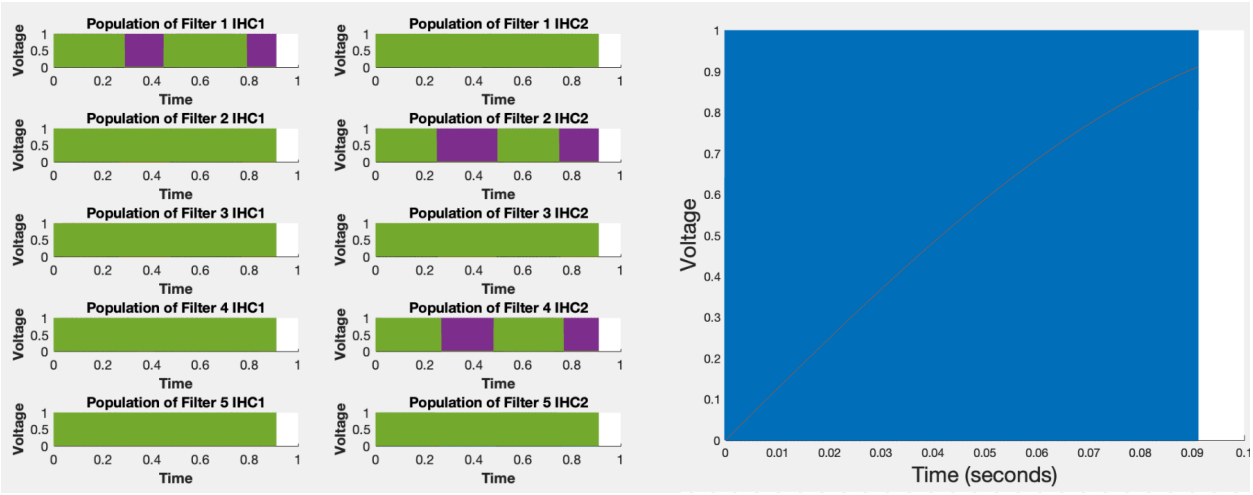


Figure 13: Neural Spike Train. The panel on the left shows the neural spike train for the first 10 neural populations to highlight the different sensitivities that each population has for their responsibility of encoding particular components of their input signal. The panel on the right displays a condensed spike train for just one neuron overlaid on the input signal.

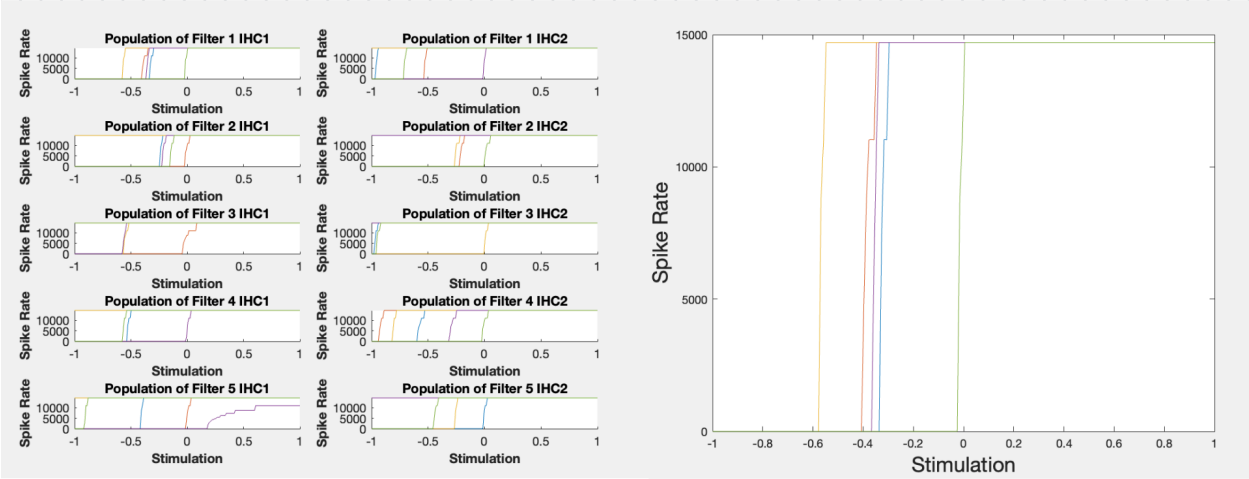


Figure 14: Activation Functions. The left panel displays the activation functions for the first 10 neural populations, with each function corresponding to one of the five neurons in each population. The right panel highlights the graph of the activations for one neural population.

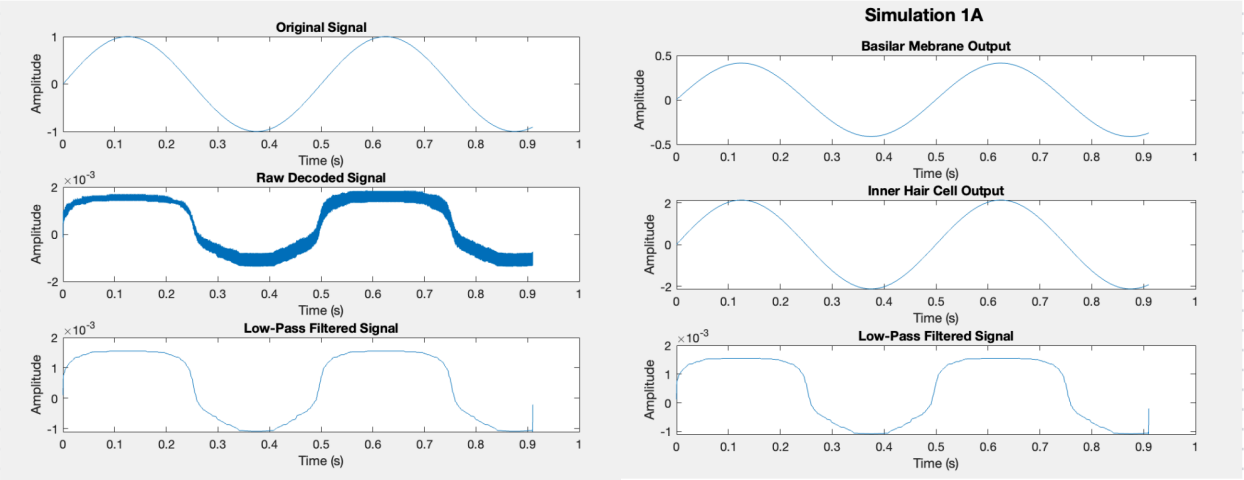


Figure 15: Model Outputs. The graph displayed on the left panel compares the original input signal to the raw signal decoded by the auditory nerve to the final low pass filtered signal output to reduce high frequency noise. The graph displayed on the right panel compares the signal at the 3 major checkpoints throughout the model: Basilar Membrane Output, Inner Hair Cell Output, and Final Low Pass Filtered Decoded Signal.

4.1.2 Simulation 1A-H Outputs

The graphs shown below compile all of the outputs of Simulation 1 to be used as a qualitative comparison between the effects of each model element's effect on the accuracy. This is done by comparing the output of each of the three model components: basilar membrane, inner hair cell, and auditory nerve.

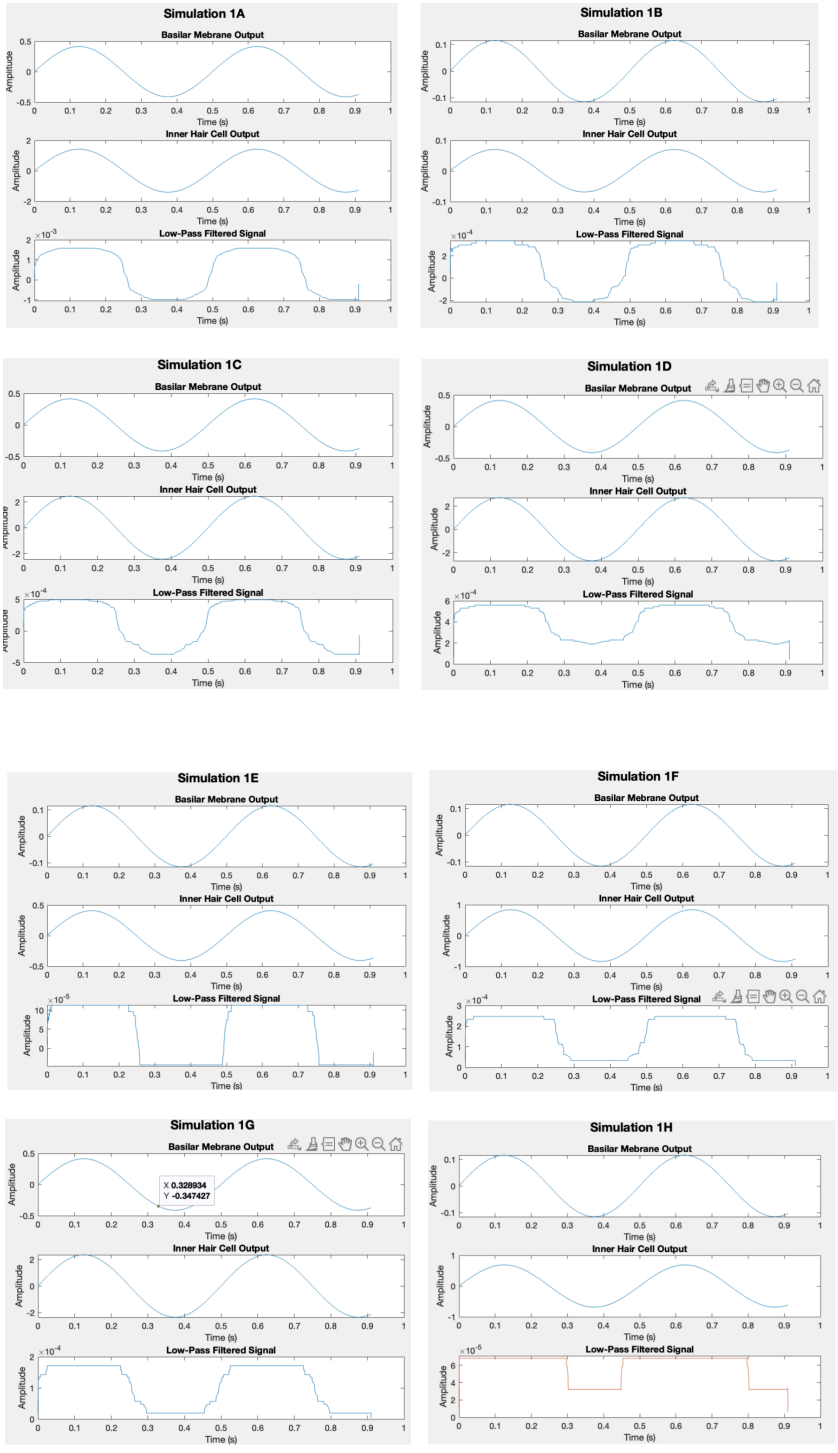


Figure 16: Outputs for Simulations 1-H. The figure displays eight panels, one for each simulation for the 2 Hz signal, each with three subplots displaying the Basilar Membrane Output, Inner Hair Cell Output, and Low Pass Filtered Decoded Output.

4.2 Simulation 2: Speech Signal Processing

The robustness and precision of the model was further tested with speech recordings saved in .wav files. Simulation 2 validated the purpose of the model to be utilized by researchers to better understand physiologically relevant auditory signal processing and the auditory system as a whole. The simulation was run over a 0.91 second duration with a sampling frequency of 48000. The wave file was read by MATLAB's *audioread* function and plotted shown in *Figure 17* below.

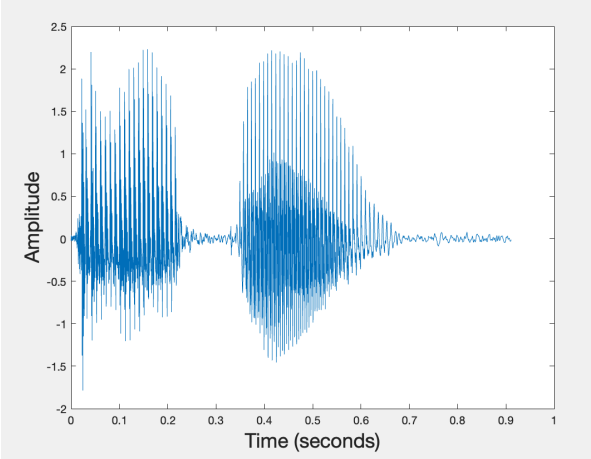


Figure 17: Simulation 2 Input Signal.

Similarly to Simulation 1A, Simulation 2A displays the model with the largest number of elements to validate output correlation and model run time for a speech audio input. These performance parameters are linked to the number of basilar membrane locations, inner hair cells, and neurons in each population that are evaluated between the different simulations within Simulation 2. Every graphical output is plotted for this simulation to highlight all of the data that the model is able to process and present. The results of Simulations 2B-H only present the graphs for the summed basilar membrane filter outputs, summed IHC filter outputs, and the lowpass

decoded output. A limited range of graphs were presented to prevent repetitiveness of the qualitative data that is accurately represented by the outputs of Simulation 1A and 2A.

Simulation 2B analyzes the effects of reducing the amount of basilar membrane locations on the output parameters. Simulation 2C analyzes reducing the amount of IHCs on the output parameters. Simulation 2D analyzes reducing the amount of neurons in each population on the output parameters. Simulation 2E analyzes reducing the amount of basilar membrane locations and inner hair cells. Simulation 2F analyzes the model with reduced basilar membrane locations and neurons in each population. Simulation 2G analyzes the model with reduced inner hair cells and less neurons in each population. Finally, Simulation 2H analyzes the lowest scaled model, with reduced amounts of basilar membrane locations, inner hair cells, and neurons.

The exact amount of each model component for each simulation can be seen in *Table 2* below. Each iteration of Simulation 2 accurately compares the effects of changing the quantity of the three element types in the model to understand the importance of their effect on the signal processing of speech.

4.2.1 Simulation 2A: Large Scale Simulation

Simulation 2A was run for the input signal defined in *Figure 17* above in Section 4.2. The model was specified with 10 basilar membrane locations, 3 inner hair cells for each location, and 5 neurons in each neural population. The model produced the graphs and output values shown in the table and figures below.

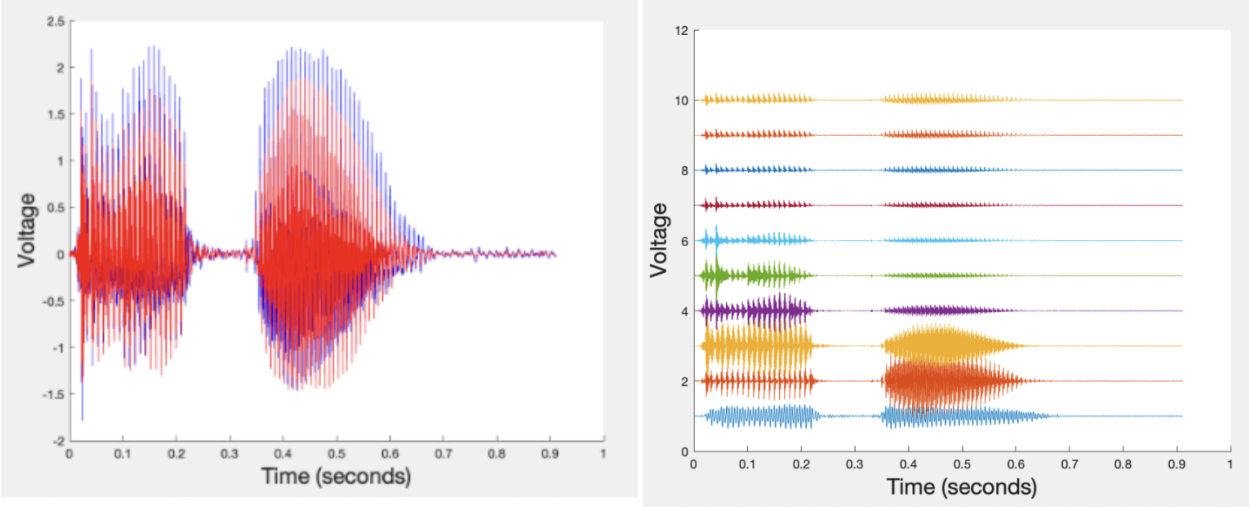


Figure 18: Basilar Membrane Outputs. The graph on the left panel displays the input wave signal in blue, and the summed outputs of each basilar membrane channel in red. The graph on the right panel displays the output of each filter channel after filtering the input signal. The channels outputs at the bottom of the graph represent locations closer to the base of the basilar membrane (lower center frequency) and channels closer to the apex are higher on the graph (higher center frequency).

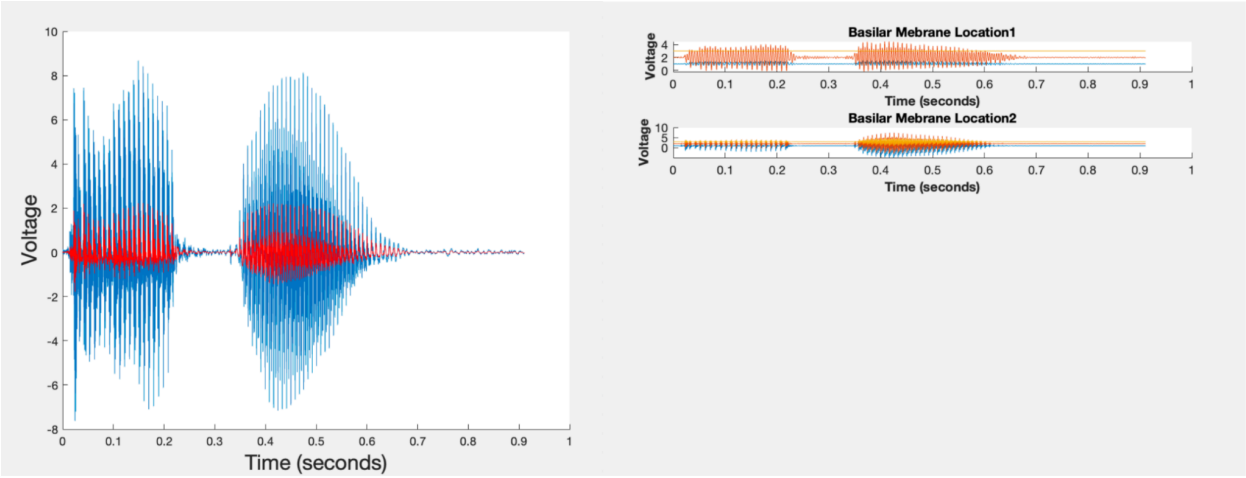


Figure 19: Inner Hair Cell Outputs. The graph on the left panel displays the input waveform in red, and the summed outputs of each inner hair cell in blue. The graph on the right panel displays the inner hair cell outputs that correspond to the first two basilar membrane locations for Simulation 2A. Both Location 1 and Location 2 have three inner hair cells receiving the same input signal from that location's output.

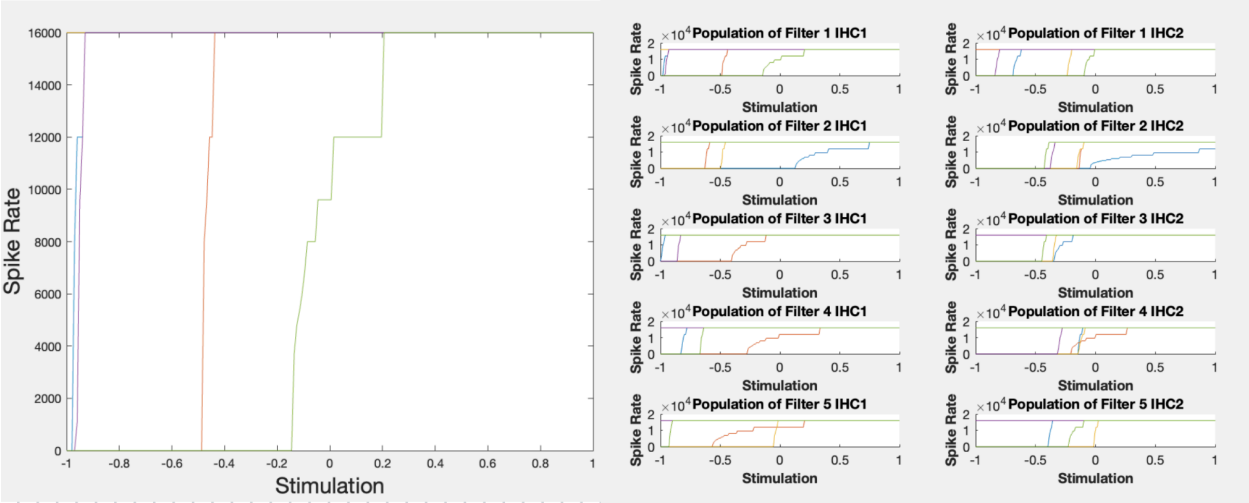


Figure 20: Activation Functions. The left panel displays the activation functions for the first 10 neural populations, with each function corresponding to one of the five neurons in each population. The right panel highlights the graph of the activations for one neural population.

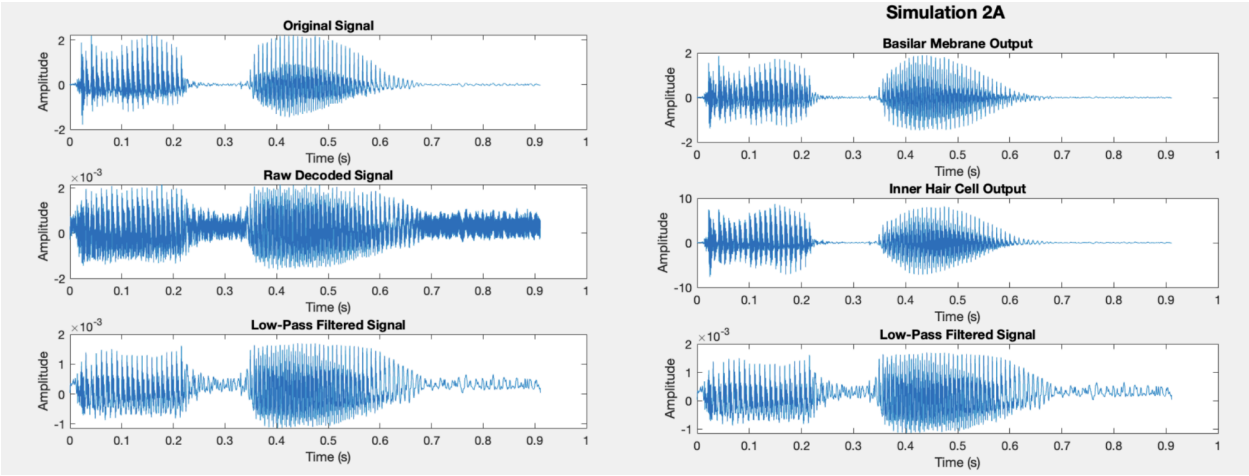
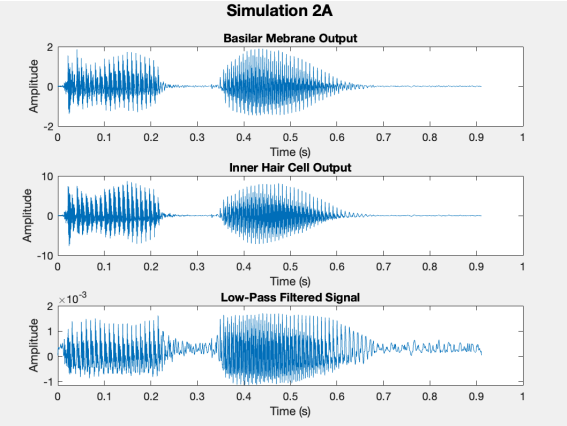
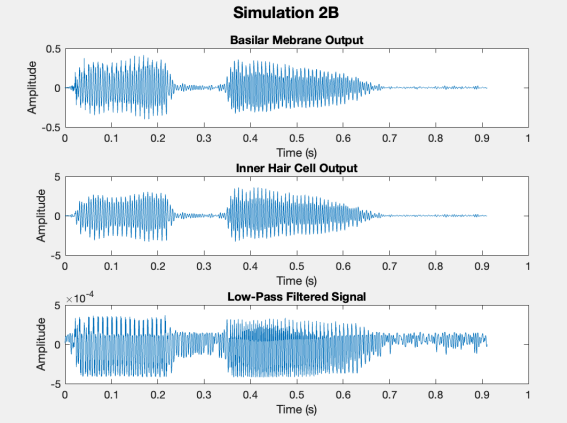


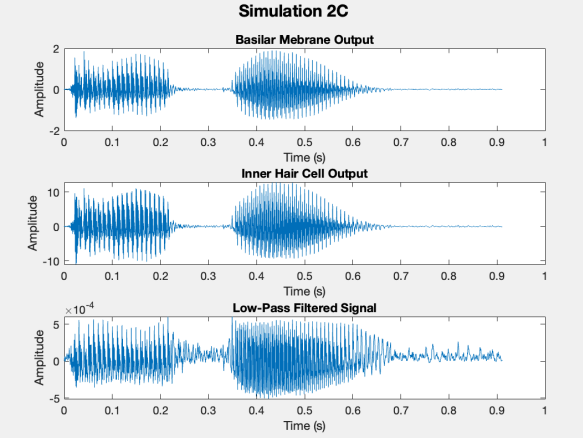
Figure 21: Model Outputs. The graph displayed on the left panel compares the original input signal to the raw signal decoded by the auditory nerve to the final low pass filtered signal output to reduce high frequency noise. The graph displayed on the right panel compares the signal at the 3 major checkpoints throughout the model: Basilar Membrane Output, Inner Hair Cell Output, and Final Low Pass Filtered Decoded Signal.

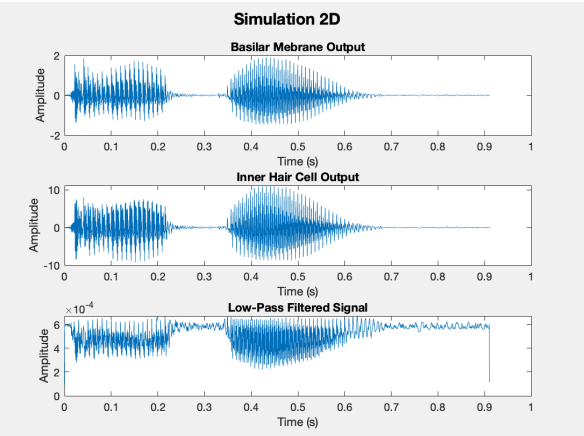
4.2.2 Simulation 2A-H Outputs

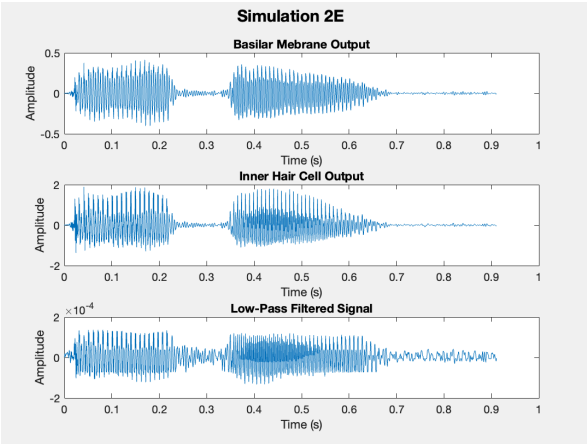
The graphs shown below compile all of the outputs of Simulation 1 to be used as a qualitative comparison between the effects of each model element's effect on the accuracy. This is done by comparing the output of each of the three model components: basilar membrane, inner hair cell, and auditory nerve.

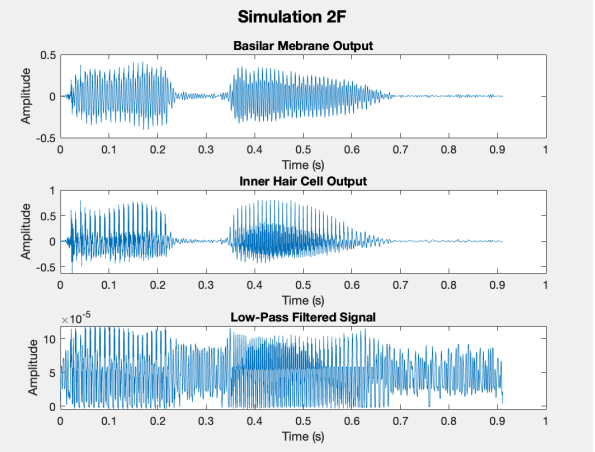

Simulation 2A
Basilar Mebrane Output
Amplitude
Time (s)
Inner Hair Cell Output
Amplitude
Time (s)
Low-Pass Filtered Signal
×10^-3
Amplitude
Time (s)


Simulation 2B
Basilar Mebrane Output
Amplitude
Time (s)
Inner Hair Cell Output
Amplitude
Time (s)
Low-Pass Filtered Signal
×10^-4
Amplitude
Time (s)


Simulation 2C
Basilar Mebrane Output
Amplitude
Time (s)
Inner Hair Cell Output
Amplitude
Time (s)
Low-Pass Filtered Signal
×10^-4
Amplitude
Time (s)


Simulation 2D
Basilar Mebrane Output
Amplitude
Time (s)
Inner Hair Cell Output
Amplitude
Time (s)
Low-Pass Filtered Signal
×10^-4
Amplitude
Time (s)


Simulation 2E
Basilar Mebrane Output
Amplitude
Time (s)
Inner Hair Cell Output
Amplitude
Time (s)
Low-Pass Filtered Signal
×10^-4
Amplitude
Time (s)


Simulation 2F
Basilar Mebrane Output
Amplitude
Time (s)
Inner Hair Cell Output
Amplitude
Time (s)
Low-Pass Filtered Signal
×10^-5
Amplitude
Time (s)

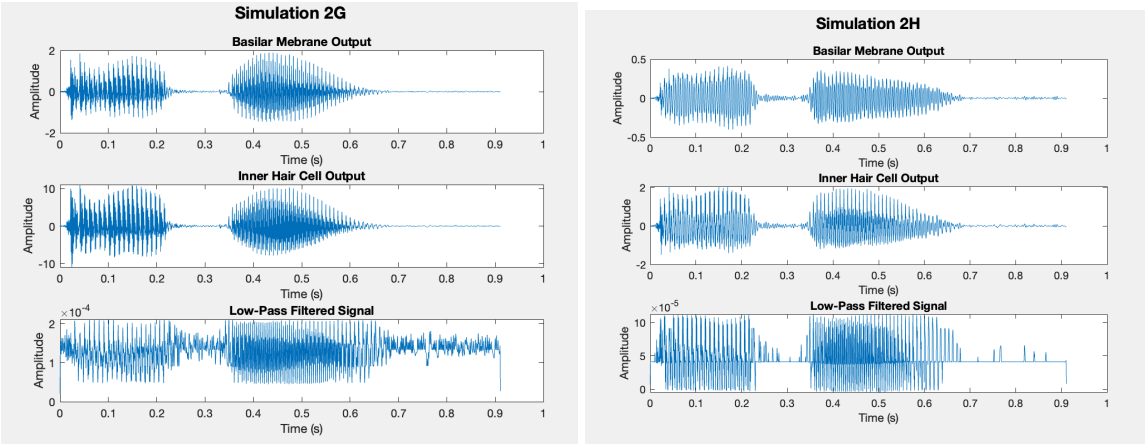


Figure 22: Outputs for Simulations 1-H. The figure displays eight panels, one for each simulation for the 2 Hz signal, each with three subplots displaying the Basilar Membrane Output, Inner Hair Cell Output, and Low Pass Filtered Decoded Output.

4.3 Analysis of Simulations

4.3.1: Comparison of Model Components

The data from each of the simulations run for Simulation 1 and Simulation 2 were compiled into *Table 1 and Table 2* below to compare each simulation's efficiency and accuracy. *Table 3* has the average values and standard deviations for each simulation for easy comparison.

Table 1: Model Correlation Values. displays the correlation between the input signal and the signal after being filtered by the basilar membrane, inner hair cells, neural decoding, and low pass filtering (to reduce noise) for all iterations of Simulation 1. The time it took for the model to run to completion is also recorded.

Simulation #	BM/IHC/Neurons	BM Corr	IHC Corr	Decoded Corr	LP Decoded Corr	Model Duration

1A	10/3/5	0.99998	0.99999	0.96258	0.97002	640.432013
1B	2/3/5	0.99996	0.99985	0.94551	0.96856	132.720302
1C	10/1/5	0.99998	1.00000	0.95712	0.96808	217.667115
1D	10/3/1	0.99998	0.99999	0.91727	0.96948	557.257850
1E	2/1/5	0.99996	0.99987	0.91720	0.92944	47.436233
1F	2/3/1	0.99996	1.00000	0.93188	0.94910	114.910441
1G	10/1/1	0.99998	0.99998	0.94189	0.96324	190.624951
1H	2/1/1	0.99996	0.99998	0.72083	0.82043	41.496254

Table 2: Model Correlation Values. displays the correlation between the input signal and the signal after being filtered by the basilar membrane, inner hair cells, neural decoding, and low pass filtering (to reduce noise) for all iterations of Simulation 2. The time it took for the model to run to completion is also recorded.

Simulation #	BM/IHC/Neurons	BM Corr	IHC Corr	Decoded Corr	LP Decoded Corr	Model Duration
2A	10/3/5	0.82638	0.87689	0.76228	0.83489	695.294174
2B	2/3/5	0.30572	0.01634	0.29494	0.29763	143.869999
2C	10/1/5	0.82638	0.85854	0.69759	0.82426	239.954774

2D	10/3/1	0.82638	0.80727	0.57806	0.57523	617.615783
2E	2/1/5	0.30572	0.67835	0.38387	0.47554	53.846081
2F	2/3/1	0.30572	0.88407	0.39591	0.39766	140.545675
2G	10/1/1	0.82638	0.86321	0.49607	0.51021	210.689034
2H	2/1/1	0.30572	0.67400	0.58391	0.61210	45.082965

+

Table 3. Average values and standard deviations for all trials in both simulations

Simulation	**Basilar Membrane output correlation**		**Inner Hair Cell Output Correlation**		**LP Decoded Output Correlation**		**Model Duration (seconds)**	
	Average	**SD**	**Average**	**SD**	**Average**	**SD**	**Average**	**SD**
2Hz Signal	0.99997	1E-5	0.99996	5.7E-5	0.94229	0.0479	242.818	214.334
Speech	0.56605	0.26033	0.70733	0.27307	0.56594	0.17779	268.3623	233.476

4.3.2: Model Efficiency, Speed, and Accuracy

For each simulation run above, the Low Passed Decoded Signal's correlation to the Input Signal is used as the measure for the model's accuracy. This value for accuracy is plotted against the model duration which is a measure of the models' speed. The accuracy value is plotted against its corresponding speed value for each simulation to understand how these values are correlated for each simulation and give a method to judge the efficiency of the model.

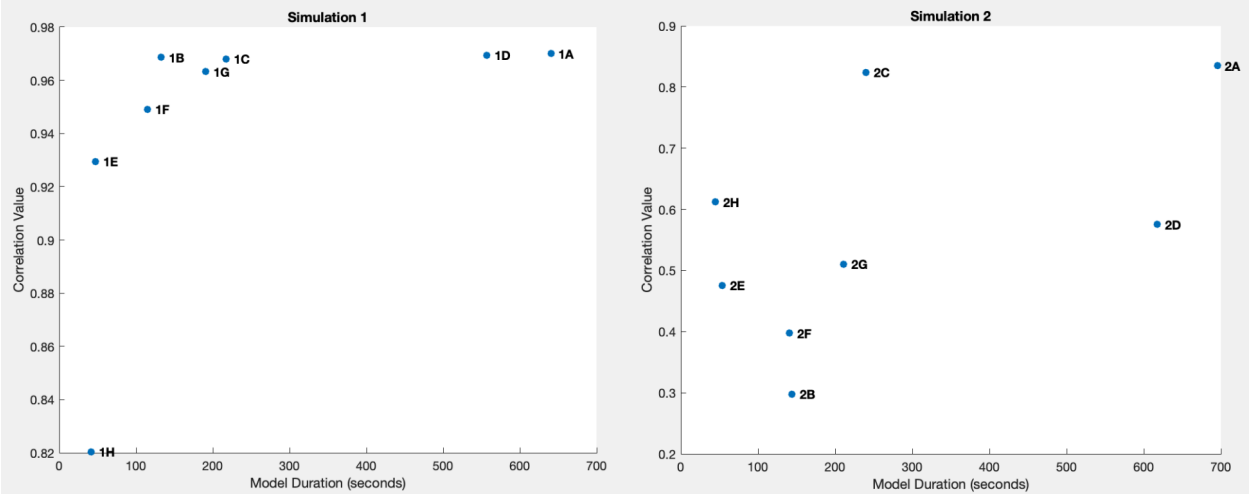


Figure 23: Model Efficiency Analysis. The left panel of the above figure displays the correlation value for each simulation plotted against the time it took to complete for Simulation 1. The same relationship is plotted on the right panel for Simulation 2.

4.4 Audibility Test

For the audibility test, two input audio files were imported into MATLAB with different speech patterns for processing. The first audio file recorded the word “apple” and the second file recorded the word “orange”. These words were chosen to give a wide range in the frequency spectrums composed within the pronunciation of the word. Participants were then asked to repeat the word that they listened to from the model’s output played using MATLAB’s *soundsc* function. Initially the participants were not told what words were inputting into the model to prevent them from inferring the answer and data bias. Each participant listened to each model output 10 times in a random order and their audio perception was recorded and they were given an accuracy score out of 20.

Table 4: Audibility Test Scores. The table shown above displays the scores of the predescribed audibility test to analyze the comprehension rate of the model output.

Participant	Accuracy Score
1	100%
2	100%
3	100%
Average	100%

5. Discussion

5.1 Simulation 1

The basilar membrane and inner hair cells closely resemble the input signal with average correlations of 0.99997 and 0.99996 each. Each of the final outputs with the constant signal contained two distinct peaks and two distinct valleys, which can be considered a success for the model. Across all of the model's simulations, the output signal's amplitude was greatly reduced. Over the simulations with the 2Hz signal, the output's amplitude was reduced by a factor of 1000 to 100000. Passing the output through a low pass filter increased the correlation in all of the simulations. Simulation 1A produced the highest correlated output signal, 0.97, which is to be expected for a model that has higher counts of basilar membrane locations, inner hair cells, and auditory nerves. This is also the most physiologically relevant simulation, as there is a large

number of basilar membrane locations, inner hair cells, and auditory nerve fibers in the auditory system. Simulation 1H, with the lowest amount of membrane locations, inner hair cells and neurons still had a high correlation, 0.82, although it closely resembles a square wave. Apart from this one sub-simulation, the variation in the model's components did not have a large effect on the correlation, averaging at 0.942. The largest difference between the simulations was the runtime with a range of 598.936 seconds.

5.2 Simulation 2

When the model was tested using a speech clip, there were varying degrees of success. The average correlation of all simulations was 0.566 with a standard deviation of 0.19. Each of the outputs somewhat share a similar pattern of high and low amplitudes, mirroring the input signal. The average correlations of the basilar membrane and inner hair cell outputs were much lower than the previous simulation, being 0.566 and 0.707 respectively. Simulations 2A and 2C were the most clear when played back to the user and have correlations of 0.83 and 0.82. Simulation 2B was barely audible and had the lowest correlation, 0.298. This is as expected and adds to the physiological relevance of the model because reducing the number of basilar membrane locations on the model greatly reduces the distinct frequencies that a signal can be broken down into. The other playback audios for the simulation had some drawbacks as well. Simulation 2D's playback was choppy, simulations 2E, 2F, and 2G have lots of background noise, and simulation 2H was muffled when played back. This simulation had a greater range of time values than the constant signal - 650.211 seconds. An average score of 100% on the listening test is extremely helpful when validating the model, and shows that the model's output is accurate.

5.3 Efficiency

In the development and optimization of our auditory model, efficiency has emerged as a big concern, particularly because of parameter optimization efforts. While the optimization process significantly enhanced the model's accuracy, it inadvertently introduced a substantial increase in computational demand. This increase primarily affects the model's runtime, making it a crucial aspect to address in our evaluation of the model's efficiency.

As detailed in Section 4.4 ("Comparison of Simulations") and substantiated by the data within, the structure of our model inherently influences its computational efficiency. The model's architecture, which branches out from basilar membrane (BM) locations to inner hair cells (IHCs) and further to neurons (LIFNs), creates a scenario where the number of BM locations has the most pronounced effect on runtime. This effect is due to the increase in the number of IHCs and LIFNs, which escalates the computational load. For instance, simulation 2A, with a configuration of 10 BM spots, 3 IHCs, and 5 neurons, exhibits a model duration significantly higher than other configurations with fewer BM spots (e.g., Simulation 2B).

Following the BM spots, the second biggest factor on runtime is the number of IHCs. This is because IHCs contribute to the branching complexity of the model, impacting the computational resources required for simulation. The correlation between the model's components and its runtime is evident when examining different configurations and their corresponding durations, as shown in the comprehensive table provided in Section 4.4.1 ("Comparison of Model Components").

Post-optimization, the script's runtime has notably increased, highlighting a trade-off between model accuracy and computational efficiency. This increase is attributed to the optimized parameters enhancing the model's accuracy but also adding to its complexity and, by

extension, its computational demands. The data shows that configurations with a higher number of BM locations endure the most significant impact on runtime due to the amplified branching and consequent increase in IHCs and LIFNs.

Despite these challenges, understanding the dynamics between model configuration and computational efficiency is crucial. It provides insights into how modifications in the model's structure can disproportionately affect its performance and operational viability. For instance, simulation results such as those from 2A versus 2E reveal how varying BM, IHC, and neuron counts impact model duration, offering an idea which to gauge the efficiency of different configurations.

In conclusion, while parameter optimization has undeniably aided in the model's accuracy and capability in simulating auditory processes, it has also spotlighted the critical balance between computational efficiency and simulation accuracy. The insights gained from analyzing the impact of model configurations on runtime are valuable as to how our model performs under different configurations.

6. Conclusion and Recommendations

This project aimed to create the first physiologically relevant neuromorphic model of the auditory system, with the aim of addressing the limitations of current algorithmic and implementational models by integrating both into one unified framework. As mentioned previously, objectives were established to ensure the accurate structuring of the model in alignment with the physiology of the inner ear, successful simulation of each individual part's processing of inputs, generation of graphs for monitoring the results of each model section, production of a decoded output signal matching the input, and facilitation of real-time signal processing.

The model was shown to produce outputs for each model section that matched the expected physiological response. The gammatone filters were successful in modeling the function of the basilar membrane. The frequency specificity was achieved through using the filter banks, the outputs of which are shown in figure 10. The plots representing the outputs of the inner hair cells also matched the expectations of typical hair cell behavior, in figure 11. The auditory nerve section exhibited behavior that matched expectations. The activation functions matched physiological behavior, as shown in figure 14. As mentioned previously in the results and discussion section, the model successfully processed and produced a decoded result for both the simple 2Hz sinusoid and the complex speech waveform. The correlation values before being filtered through a low-pass filter had a high of 0.963 and a low of 0.721 for the sinusoid, and a high of 0.762 and low of 0.384 for speech, shown in Table 1 and Table 2 respectively. This is indicative of the model functioning successfully.

Although the model could successfully process signals in the same way the physiology can, there were a few limitations. During development, the inner hair cell and auditory nerve was

not effective at producing the expected outputs, but through optimizing hyperparameters, a clear decoded output was produced. As well, the model is not efficient enough to facilitate real-time signal processing, which was another objective. Further, the model is published as one complete script, so individual sections can not be run standalone.

Future work on this model should focus on optimizing the efficiency of the model to facilitate real-time signal processing. A better method of producing hyperparameters should be found in order to reduce run-time and increase the correlation value for the decoded output signal. Furthermore, enabling for the tweaking of individual parameters, such as disabling certain regions of the basilar membrane should become a feature, in order to facilitate more functions of the model as a research tool for specific auditory processing disorders. Finally, it would be helpful to separate the model sections, and enable standalone running of individual model sections, so that individual physiological parts can be focused on, rather than the system as a whole.

This project marks the creation of the first neuromorphic model of the auditory system. By utilizing this model as a tool for research, it is possible to better understand how the auditory system processes sound, how hearing disorders affect the processing of the ear, and potentially allowing for better diagnosis and treatment of disease.

Bibliography:

[1] World Health Organization. (2024, February 2). Deafness and hearing loss. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss

[2] Meddis, R., Lecluyse, W., Clark, N. R., Jürgens, T., Tan, C. M., Panda, M. R., & Brown, G. J. (2013). A computer model of the auditory periphery and its application to the study of hearing. Basic aspects of hearing: physiology and perception, 11-20.

[3] Heinz, M. G., Zhang, X., Bruce, I. C., & Carney, L. H. (2001). Auditory nerve model for predicting performance limits of normal and impaired listeners. Acoustics Research Letters Online, 2(3), 91-96.

[4] Marr, David, and Shimon Ullman. "Directional selectivity and its use in early visual processing." Proceedings of the Royal Society of London. Series B. Biological Sciences 211, no. 1183 (1981): 151-180.

[5] Sumner, C. J., Lopez-Poveda, E. A., O'Mard, L. P., & Meddis, R. (2002). A revised model of the inner-hair cell and auditory-nerve complex. The Journal of the Acoustical Society of America, 111(5 Pt 1), 2178–2188. https://doi.org/10.1121/1.1453451

[6]. Eliasmith, C., & Anderson, C. H. (2003). Neural Engineering Computation, Representation, and Dynamics in Neurobiological Systems. The MIT Press.

[7] Moore, B. C. (2008). Basic auditory processes involved in the analysis of speech sounds. Philosophical Transactions of the Royal Society B: Biological Sciences, 363(1493),

947-963.

[8] Anthwal, Neal, and Hannah Thompson. "The development of the mammalian outer and middle ear." Journal of anatomy 228, no. 2 (2016): 217-232.

[9] Zhang, Jiawei. "Basic neural units of the brain: neurons, synapses and action potential." arXiv preprintarXiv:1906.01703 (2019).

[10] Meddis, Ray, and Enrique A. Lopez-Poveda. "Auditory periphery: from pinna to auditory nerve." Computational models of the auditory system (2010): 7-38.

[11] Goldstein, Julius L. "Modeling rapid waveform compression on the basilar membrane as multiple-bandpass-nonlinearity filtering." Hearing research 49, no. 1-3 (1990): 39-60.

[12] Lopez-Poveda, Enrique A., and Almudena Eustaquio-Martín. "A biophysical model of the inner hair cell: the contribution of potassium currents to peripheral auditory compression." Journal of the Association for Research in Otolaryngology 7 (2006): 218-235.

[13] Nelson, Mark, and John Rinzel. "The hodgkin-huxley model." The book of genesis (1995): 29-49.

[14] Huang, Qi, and Jianguo Tang. "Age-related hearing loss or presbycusis." European Archives of Oto-rhino-laryngology 267 (2010): 1179-1191.

[15] Natarajan, Nirvikalpa, Shelley Batts, and Konstantina M. Stankovic. "Noise-Induced Hearing Loss." Journal of Clinical Medicine 12, no. 6 (2023): 2347.

[16] Chari, Divya A., and Charles J. Limb. "Tinnitus." Medical Clinics 102, no. 6 (2018): 1081-1093.

[17] Aimone, James B., Prasanna Date, Gabriel A. Fonseca-Guerra, Kathleen E. Hamilton, Kyle Henke, Bill Kay, Garrett T. Kenyon et al. "A review of non-cognitive applications for

neuromorphic computing." Neuromorphic Computing and Engineering 2, no. 3 (2022): 032003.

[18] Goldstein, Julius L. "Modeling rapid waveform compression on the basilar membrane as multiple-bandpass-nonlinearity filtering." Hearing research 49, no. 1-3 (1990): 39-60.

[19] Kim, Doh-Suk, Soo-Young Lee, and Rhee Man Kil. "Auditory processing of speech signals for robust speech recognition in real-world noisy environments." IEEE Transactions on speech and audio processing 7, no. 1 (1999): 55-69.

[20] Umesh, Srinivasan, Leon Cohen, and D. Nelson. "Fitting the mel scale." In 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No. 99CH36258), vol. 1, pp. 217-220. IEEE, 1999.

[21] Meddis R. Simulation of mechanical to neural transduction in the auditory receptor. J Acoust Soc Am. 1986 Mar;79(3):702-11. doi: 10.1121/1.393460. PMID: 2870094.