Simulated EEG-Driven Audio Information Mapping Using Inner Hair-Cell Model and Spiking Neural Network

Pasquale Mainolfi
Dep. of New Technologies and Musical Languages
"G. Martucci" State Conservatoire of Music
Salerno, Italy
pasquale.mainolfi@consalerno.it

Abstract

This study presents a framework for mapping audio information into simulated neural signals and dynamic control maps. The system is based on a biologically-inspired architecture that traces the sound pathway from the cochlea to the auditory cortex. The system transforms acoustic features into neural representations by integrating Meddis's Inner Hair-Cell (IHC) model with spiking neural networks (SNN).

The mapping process occurs in three phases: the IHC model converts sound waves into neural impulses, simulating hair cell mechano-electrical transduction. These impulses are then encoded into spatio-temporal patterns through an Izhikevich-based neural network, where spike-timing-dependent plasticity (STDP) mechanisms enable the emergence of activation structures reflecting the acoustic information's complexity. Finally, these patterns are mapped into both EEG-like signals and continuous control maps for real-time interactive performance control.

This approach bridges neural dynamics and signal processing, offering a new paradigm for sound information representation. The generated control maps provide a natural interface between acoustic and parametric domains, enabling applications from generative sound design to adaptive performance control, where neuromorphological sound translation explores new forms of audio-driven interaction.

Keywords

Inner Hair-Cell Model (IHC), Spiking Neural Network (SNN), s-EEG, Auditory Signal Processing, Neural Control Signals, STDP, Control Maps, Audio Mapping.

1 INTRODUCTION

1.1 Integrating Neuroscience and DSP in Audio Research

While common artificial neural networks (ANNs) are based on mathematical abstractions that simplify the behavior of biological neurons [17, 28], these are primarily designed to perform specific tasks such as speech recognition, audio classification, or text generation.

In contrast, this research adopts an alternative paradigm within the field of ANNs. Instead of focusing on developing models optimized for specific tasks, it aims to simulate a neurally plausible and biologically inspired system designed to more faithfully replicate real neural structures and processes. This approach focuses on analyzing how these systems respond to various stimuli. Specifically, it examines changes in the system's state in response to input signals, generating biologically compatible signals that can be mapped to observable outputs.

In the field of DSP, similar approaches exist in the literature. However, these approaches mainly focus on the analysis of data collected by monitoring brain activity in the laboratory. They use sophisticated instrumentation such as electroencephalograms (EEG), positron emission tomography (PET), functional magnetic resonance imaging (fMRI), and other brain imaging methods. These approaches tend to emphasize the understanding of neural activation patterns and observed mental or behavioral processes, using empirical data to deduce general principles [2, 15, 34].

1.2 Historical Overview of Experiments

The history of the integration between neuroscience and audio begins in 1934, when Adrian first translated an EEG into sound material [1].

In 1965, Alvin Lucier composed *Music for Solo Performer*, using EEG to vibrate percussion instruments through loudspeakers that amplified low-frequency signals generated by electrodes on the performer's scalp [29].

In the 1970s, D. Rosenboom explored the use of EEG to create works of art and music [29], and in 1990 he presented a system that correlated EEG components with changes in the performer's selective attention [30].

In 2003, Miranda and collaborators published a study on the use of EEG to identify patterns associated with various musical cognitive tasks, using spectral analysis. In 2006, E. R. Miranda described an innovative brain-computer interface (BCI) system for musical composition that used artificial intelligence algorithms to interpret brain patterns and generate music in real time, based on Augmented Transition Networks (ATN) [22].

1.3 Recent Advances

Recent developments in the integration of neuroscience and digital audio signal processing have led to significant discoveries about musical perception in the brain.

Sanyal et al. demonstrated a correlation between sound stimuli and brain activity by sonifying EEG data from 10 participants while they listened to a tanpura drone, showing how stimuli activate different brain regions at varying times [32]. Daly combined EEG and fMRI to improve music decoding, achieving 71.8% accuracy with a biLSTM network, outperforming the 59.2% accuracy of EEG alone, and identifying brain regions involved in music listening [6].

Bellier et al. reconstructed a complete song (*Another Brick in the Wall, Part 1* by Pink Floyd) using intracranial EEG data from 29 patients, showing that nonlinear models improve accuracy by 32% and that high-quality reconstructions are achievable even with a reduced number of electrodes [4].

NIME '25, June 24–27, 2025, Canberra, Australia

Although these advances, Spiking Neural Networks (SNNs) remain underutilized. Studies such as those by Fujii and Oozeki on melody recognition [10] and Liang et al. on music composition have highlighted the potential of SNNs. Liang et al. developed a system based on STDP capable of encoding musical notes and generating melodies in various styles, simulating the behavior of biological neurons [14].

1.4 Research Objectives

In the context of Digital Signal Processing (DSP) and Computer Music, current literature does not document the use of biologically plausible computational models that utilize spiking neural networks (SNNs), either in isolation or in combination with other biologically inspired models, for the processing of complex audio signals. In particular, no studies have been reported that exploit such models for a detailed analysis of the signal processing performed by bio-inspired mechanisms, for instance the use of simulated electroencephalograms (s-EEG), for the generation of control maps.

This study aims to address this gap by integrating two biologically inspired models: SNNs and the inner hair cell (IHC) model proposed by Meddis, Hewitt, and Shackleton [21]. The objective is to develop a system in which the IHC model simulates the transformation of the auditory stimulus, replicating the processes occurring in the inner ear, before this stimulus is further processed by the SNN, which operates analogously to a small brain. The simulation of an EEG will finally allow the analysis of the SNN activity. This final stage of transformation should be understood as the nonlinear translation of a dataset into a different domain, system, or format, preserving and highlighting structural relationships and significant features. Specifically, it involves converting an audio signal into an analytical representation that enables the identification of patterns or the generation of control signals. These outputs can be employed in interactive, artistic, or technological applications.

2 BIO-INSPIRED NEURAL SIMULATION

The model is designed to simulate the pathway an audio signal traverses from the moment it enters the ear until its processing by the brain. The model consists of two principal blocks: the first block replicates the transfer function of the inner hair cells, which perform the crucial role of mechano-electrical transduction of acoustic signals; the second block consists of a spiking neural network, which simulates the complex spatio-temporal processing of the signal within the central nervous system.

Although the existence of numerous studies and computational models for inner hair cells (IHC), such as [16, 24, 31, 35], which introduced improvements in modeling spike dynamics and adaptation, the implementation proposed by Meddis, Hewitt, and Shackleton has been chosen as the basis for this study. This choice is based on several key factors that make the model particularly suitable for the research in question. The model provides a detailed representation of the physiological processes of IHCs, including mechano-electrical transduction and neurotransmitter release, thus offering an accurate and functional description of the conversion of the auditory signal in the early stages of neural processing. Its mathematical formulation is relatively simple, yet effective, allowing for easy implementation and manipulation. This feature, combined with the model's modular structure, facilitates integration with other neural models, making it ideal for more complex simulations of the auditory system. Another strength of the Meddis model is its extensive experimental validation. The model has been extensively tested [3, 18–20] and has proven reliable in replicating the responses of IHCs under various conditions, adding robustness to the results obtained.

Furthermore, despite its accuracy, the model remains computationally efficient, an important aspect when considering implementations in more complex systems or practical applications.

2.1 Meddis Model for IHC

Meddis [19, 20] proposes a probabilistic model of mechano-neural transduction in auditory receptors, structured into three components: neurotransmitter release from hair cells, excitatory postsynaptic potentials (EPSPs) in auditory neurons, and discharge models.

In his work, Meddis presents two models, with the second (Model B) representing a significant improvement over the first through the implementation of a more sophisticated neurotransmitter recycling mechanism.

It is assumed that hair cells contain a certain amount of free neurotransmitter, which leaks into the synaptic cleft through a permeable membrane. The permeability of the membrane fluctuates according to the instantaneous amplitude of the acoustic stimulus, as described in Equation (1),

$$k(t) = \begin{cases} \frac{g[s(t)+A]}{s(t)+A+B}, & \text{for } s(t)+A > 0, \\ 0, & \text{for } s(t)+A \leq 0. \end{cases} \tag{1}$$

where A and B are positive constants with $B > A$, $k(t)$ represents the permeability that oscillates between 0 and g, and $s(t)$ is the acoustic stimulus.

$$\frac{dq(t)}{dt} = y[1 - q(t)] + xw(t) - k(t)q(t), \tag{2}$$

$$\frac{dw(t)}{dt} = rc(t) - xw(t), \tag{3}$$

$$\frac{dc(t)}{dt} = k(t)q(t) - \lambda c(t) - rc(t) \tag{4}$$

Equations (2), (3), and (4) fully describe the model. $q(t)$ is the level of free neurotransmitter in the cell, and y is a constant representing the neurotransmitter production rate (the rate at which neurotransmitter is produced and added to the free neurotransmitter pool). The variable x represents the rate of transfer of neurotransmitter from the reservoir to the pool, while $w(t)$ indicates the amount of neurotransmitter in the reprocessing reservoir. The term $[1 - q(t)]$ represents the remaining capacity fraction of the pool, which is inversely proportional to $q(t)$: when $q(t)$ is high, this term becomes low, and vice versa. In Equation (3), r represents the rate of neurotransmitter reabsorption from the synaptic cleft back into the hair cell, and $c(t)$ denotes the amount of neurotransmitter present in the synaptic cleft at a given time.

Finally, Equation (4) describes the time variation of the neurotransmitter quantity in the synaptic cleft $c(t)$, where $k(t)q(t)$ represents the rate of neurotransmitter release into the cleft, and λ is the irreversible loss rate of neurotransmitter from the cleft.

The probability of an event, as given by an Equation (5), is proportional to $c(t)$, multiplied by a small-time interval dt, and h is a costant in a model.

$$spike = hc(t)dt \tag{5}$$

In summary, the model takes the instantaneous amplitude of the acoustic signal as input and produces an excitation function

that represents the instantaneous fluctuating probability of a spike. It thus generates a signal that reflects the electrical activity in response to the acoustic stimulus.

2.2 Izhikevich Spiking Neuron

Spiking neurons are mathematical models that simulate the electrical activity of biological neurons, focusing on the generation and timing of action potentials. These models aim to capture the discrete and temporal nature of neuronal communication.

The Izhikevich model [13] represents a significant advancement in simulating the activity of cortical neurons, offering an optimal balance between the computational simplicity of integrate-and-fire models and the dynamic richness of Hodgkin and Huxley models [12]. It is a biologically plausible model capable of reproducing a wide range of spiking patterns observed in real neurons using only two differential equations (6) and (7) and four parameters a, b, c, d [13].

$$v' = 0.042v^2 + 5v + 140 - u + I \tag{6}$$

$$u' = a(bv - u) \tag{7}$$

v, the neuron's membrane potential (in millivolts, mV), describes the electrical potential difference between the inside and outside of the neuronal cell and u, the membrane recovery variable (dimensionless), accounts for the activation of ionic K+ currents and the inactivation of Na+ currents. The variable I is an external current injected into the neuron which can vary over time (synaptic input). Parameter a control the speed of u response to changes in v. b determines the influence of v variation on u. The term $[0.042v^2+5v+140-u+I]$, in Equation (6), approximates the spike initiation dynamics of cortical neurons.

When the membrane potential v reaches and exceeds the spike threshold ($V_{spike} = +30mV$) the neuron generates a spike and v and u are reset according to Equation (8).

$$\text{if } v \geq 30mV, then \begin{cases} v \leftarrow c \\ u \leftarrow u + d \end{cases} \tag{8}$$

In conclusion, the Izhikevich model stands out for its ability to capture a wide range of essential features of biological neuronal activity and its computational efficiency, enabling the real-time simulation of large-scale neural networks on standard computers.

3 PROPOSED MODEL

The proposed model (see Figure 1) is structured as a single nonlinear operator that transforms digital audio signal (auditory stimulus) through a temporal sequence of events inspired by biological processes.

This transformation occurs in three distinct phases:

(1) **Cochlear simulation**. The first phase emulates the conversion (based on [21]) of the auditory stimulus into an electrical signal;
(2) **Neural processing**. The second phase processes the electrical signal using spiking neurons interconnected in a 3D space. The network implements a behavioral learning model based on Hebb's rule [11], a fundamental principle of neuroplasticity describing how synaptic strength (connections) can be modified in response to neural activity;
(3) **Analysis of neural activity**. The final phase involves capturing and analyzing the neural activity generated by the network. This is achieved through the s-EEG.

We can express the entire computational framework by defining the nonlinear operator $CNA()$ (an acronym for *Cochlea Simulation, Neural Signal Processing, and Activity Analysis*) as follows:

$$CNA(s) = A(N(C(s))) \tag{9}$$

with s representing the input frame, $A()$ denotes the neural activity analisys operator (s-EEG), $N()$ represents the neural signal processing stage and $C()$ corresponds to the transformation performed by the inner ear.

3.1 Adaptive Neural Processing of IHC Model Outputs

Consider an acquisition system that captures individual audio frames f of length N (in samples).

$$k(t) = \begin{cases} \frac{g(|f[t]|+A)}{|f[t]+A+B|}, & \text{for } |f[t]| + A > 0, \\ 0, & \text{for } |f[t]| + A \leq 0. \end{cases} \tag{10}$$

$$dq[t] = y[M - q[t]] + xw[t] - k[t]q[t], \tag{11}$$

$$dc[t] = k[t]q[t] - \lambda c[t] - rc[t], \tag{12}$$

$$dw[t] = rc[t] - xw[t], \tag{13}$$

$$q[t] = q[t-1]dq\Delta t, \tag{14}$$

$$c[t] = c[t-1]dc\Delta t, \tag{15}$$

$$w[t] = w[t-1]dw\Delta t \tag{16}$$

Each frame f is processed using the model described in [21], suitably adapted for this purpose, as in (10)–(16) (see variables in Table 1).

The transformation is defined as a function that maps each frame f to a signal f_C, represented as a sequence of events such that, based on (5).

$$f_C[t] = \begin{cases} |f[t]| \cdot \gamma, & \text{if } condition, \\ 0.0, & \text{otherwise.} \end{cases} \tag{17}$$

Let $condition$ be defined as $(r < hc[t]\Delta t) \wedge (t == 1 \vee t \cdot \Delta t - \text{last\_spike} > m)$, where r is a random number such that $r \in [0, 1]$, m is the minimum interval (in sec.) between successive spikes and γ is a dynamic factor used to amplify the signal and improve the signal-to-noise-ratio (SNR).

Given the vector f_C generated by the operator $C()$, and a single-layer SNN with N_{neuro} spiking units, the input f_N for the operator $N()$ is defined as in Equation (18), where $interp(f_C, N_{neuro})$ denotes the interpolation operation applied to the vector f_C, yielding a new vector f_N (of length depending on real-time acquisition buffer), α is a scaling factor, and η represents a vector of length N_{neuro} whose elements are sampled from a uniform distribution to simulate background noise in the neural activity.

$$f_N = \alpha \cdot interp(f_C, N_{neuro}) + \eta \tag{18}$$

Equation (18) does not represent the only way to define f_N. Alternatively, f_N can be constructed by randomly assigning each value in f_C to a subset of neurons whose size matches the length of f_C, while assigning the remaining neurons to zero or a neutral value.

All neurons are interconnected and arranged in a 3D space. This spatial organization reflects the biological plausibility of cortical networks and allows for the simulation of spatiotemporal dynamics that emerge from auditory processing. The neural connectivity is typically governed by a probabilistic function

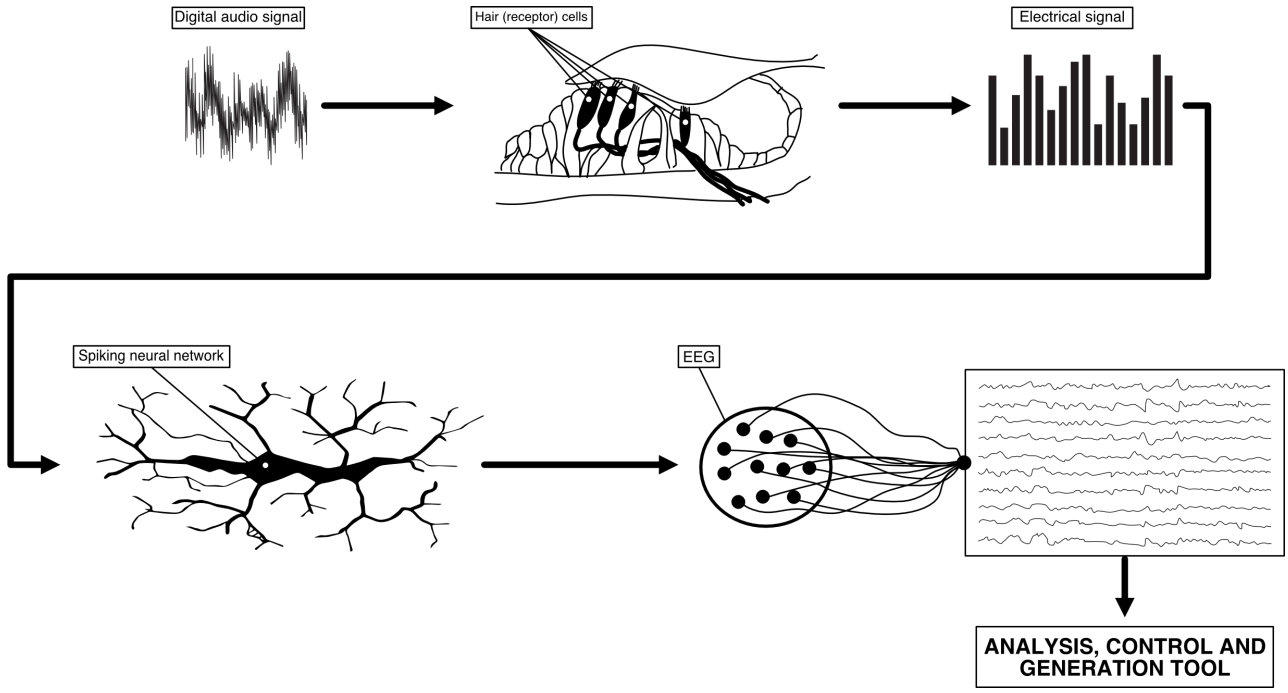

Figure 1: Graphical representation of the proposed model. Three-phase process: Inner-Hair Cell Model, Spiking Neural Network and s-EEG.

that depends on the spatial distance between neurons, promoting the formation of local clusters and long-range connections characteristic of cortical structures.

If we consider a set of neurons (19), where $j_i = (x_i, y_i, z_i)$ represents the coordinates of the i-th neuron.

$$\mathcal{J} = \{j_i \in \mathbb{R}^3 | i = 1, 2, 3, ..., N_{neuro}\} \tag{19}$$

The initial synaptic weight w_{ij} is defined as a function of the euclidean distance d_{ij} between the neurons (Equation (20)), as described in (21).

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} \tag{20}$$

$$w_{ij} = \begin{cases} \frac{1}{d_{ij}^2}, & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} \tag{21}$$

$$W = \begin{bmatrix} w_{00} & w_{01} & \cdots & w_{0N_{neuro}} \\ w_{10} & w_{11} & \cdots & w_{1N_{neuro}} \\ \vdots & \vdots & \vdots & \vdots \\ wN_{neuro}0 & d_{N_{neuro}1} & \cdots & w_{N_{neuro}N_{neuro}} \end{bmatrix} \tag{22}$$

The synaptic weight matrix W, with dimensions $N_{neuro} \times N_{neuro}$, is then formulated as in (22). The membrane potential in (6) is modified to include the effect of synaptic inputs, resulting in (23).

$$v' = v + [0.042v^2 + 5v + 140 - u + (W\varsigma + I)] \tag{23}$$

v and u are vectors of length N_{neuro}. The therm $W\varsigma$ represents the matrix-vector product, where ς is a vector of length N_{neuro} that encodes the spikes of presynaptic neurons. It contains binary values $\varsigma_i \in \{0, 1\}$, indicating whether the i-th neuron has produced a spike (1) or not (0). The variables v and u are initialized following [13].

An unsupervised learning mechanism based on Spike-Timing Dependent Plasticity (STDP) updates the weights in W. STDP adjusts synaptic weights based on the relative timing of spikes between pre- and postsynaptic neurons, strengthening connections when presynaptic spikes precede postsynaptic spikes and weakening them in the opposite case. The STDP update rule [5, 7] is expressed as in (24) and (25):

$$\Delta W = \begin{cases} a_{pre} \cdot e^{-\frac{\Delta t}{t_{pre}}}, & \text{if } \Delta t > 0, \\ -a_{post} \cdot e^{-\frac{\Delta t}{t_{post}}}, & \text{if } \Delta t < 0, \end{cases} \tag{24}$$

$$W \leftarrow W + \Delta W \tag{25}$$

where ΔW is the synaptic weight change, $\Delta t = t_{post} - t_{pre}$ is the timing difference between postsynaptic and presynaptic spikes, a_{pre} and a_{post} are the maximum amplitudes for potentiation and depression, and t_{pre}, t_{post} are time constants for the exponential decay.

The spike threshold V_{spike} is implemented as a dynamic value rather than a fixed constant (e.g. in [12]). Dynamic thresholds adjust the stimulus level required for neuron activation, enhancing adaptability to changing input conditions.

There are various methods for implementing dynamic thresholds. Among these are spike-rate-dependent thresholds, which consider the number of spikes generated by a neuron. In this case, the threshold can increase if the neuron generates a high number of spikes. There are also time-dependent thresholds, which adapt to periods of prolonged inactivity or activity, as well as composite models such as BDETT (Bioinspired Dynamic Energy-Temporal Threshold). BDETT includes two main components: the Dynamic Energy Threshold (DET) and the Dynamic Temporal Threshold (DTT) [9]. All the mentioned methods aim to achieve dynamic and homeostatic adaptation within the network, utilizing different sources of information to modulate the threshold.

In the proposed framework, the threshold is a dynamic value that adapts to the statistical characteristics of f_N (18). Specifically, as illustrated in Equation (26), it corresponds to the standard deviation of f_N multiplied by a scaling factor β, which regulates the sensitivity of the network.

$$V_{spike} = \beta \cdot std(f_N) \tag{26}$$

The formulation adopted proved advantageous during testing. It enhances the network's robustness and sensitivity to input variations, improving adaptability to diverse inputs.

3.2 Modelling EEG Signals From Network Activity

The final phase involves monitoring and recording the network's overall activity, simulating an EEG that detects potential fluctuations through components (electrodes) placed in a 3D (virtual) space.

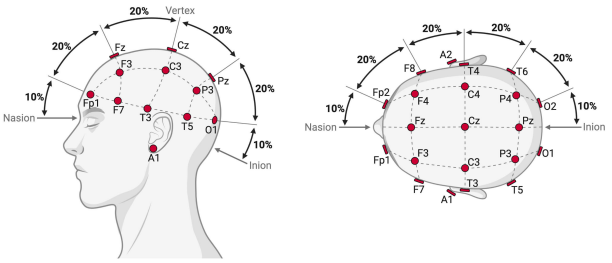

Figure 2: Scalp electrode placement according to the 10-20 system. The number 10-20 refers to the distance between adjacent electrodes is either 10% or 20% of the total front-back or right-left distance of the skull. The lobe is identified by a letter (F = frontal; C = central; P = parietal; T = temporal; O = occipital) and the hemisphere by a number. (image taken from [27].)

An EEG is a non-invasive technique used to record the brain's electrical activity. It is performed by placing electrodes on the scalp. The electric fields generated by the synchronous activity of neurons pass through various tissues (brain, skull, scalp, etc.), which act as conductors, attenuating and diffusing the signals. The electrodes detect these signals and transmit them to a device that amplifies and records them. The collected data is represented as waves that vary in frequency and amplitude.

Similarly, $A()$ in Equation (9), is defined as an operator consisting of a set of points (27), arranged according to the international 10-20 system [26, 33]. This arrangement ensures uniform coverage of relevant brain regions (Figure 2) [37].

$$\mathbb{E} = \{e_k \in \mathbb{R}^3 | k = 1, 2, 3, ..., N_e\} \tag{27}$$

In Equation (27), $e_k = (x_i, y_i, z_i)$ represents the coordinates of electrode in 3D, and N_e represents the total number of electrodes.

EEG signals are primarily generated by the current dipoles formed due to synaptic activity in cortical pyramidal cells. When neurons receive excitatory signal, ionic currents enter the cell (sink) and must be balanced by outward currents (source), forming a current dipole. These dipoles, aligned along the apical dendrites, are the primary source of EEG signals recorded on the scalp. The further one moves from the neuron, the more the extracellular potentials assume a dipolar shape, justifying the single dipole approximation for calculating EEG signals, particularly at large distances from the neuronal source [36].

The contribution of each neuron to the EEG signal measured by the k-th electrode can thus be calculated using the formula for the electric potential V generated by an electric dipole p at a point r (see Equation (28)) [25],

$$p = \Delta V \cdot L \cdot \sigma \cdot o \tag{28}$$

in which ΔV is the potential difference along the neuron, L is the distance between the main signal input and output zones of the neuron, typically ranging between 0.25 and $0.5mm$ (the effective distance over which the neuron's potential difference develops). σ is the intracellular conductivity, and o is the unit vector representing the neuron's orientation in three dimensions.

The potential on the scalp surface can then be approximated using the formula for the electric dipole (29) [25],

$$V(r) = \frac{p \cdot r}{4\pi\sigma r^3} \tag{29}$$

with r representing the vector that defines the position of the dipole relative to the measurement point on the scalp. The term $4\pi\sigma r^3$ accounts for the spherical geometry of the electric field, the tissue conductivity, and the signal attenuation with distance. It is important to emphasize that the decision to omit the dielectric constant is justified by the relatively low frequencies of the electrical signals detected by EEGs.

In this regime, the tissue's conductivity dominates the signal propagation behavior, rendering the effect of permittivity negligible compared to conductivity [25].

Thus, for each electrode e in $\mathbb{E}$, the potential measured on $V(e, t)$ the scalp at time t is calculated by summing the contributions of each neuron j, weighted by the distance between the neuron and the electrode, the neuron's orientation, and the neuron's membrane potential at time t, as in Equation (30),

$$V(e_k, t) = \eta + \sum_{i=1}^{N_{neuro}} \frac{a_i(t) L_i \sigma o_i (e_k - j_i)}{4\pi\sigma ||e_k - j_i||^3} \tag{30}$$

where $a_i(t)$, represents the membrane potential of neuron i at time t and η is a noise term modeled as a normally distributed random variable (simulating the noise generated by the electrode).

4 AUDIO-TO-s-EEG DOMAIN: EXPERIMENTAL SETUP AND MODEL RESPONSE

This section presents the response generated by the proposed model. The implementation of IHC model enabled an accurate representation of the mechano-electrical transduction, providing a bio-realistic input to SNN. The neural network exhibited the ability to learn and adapt to the signal properties through the STDP mechanism. Individual neurons generated complex spatio-temporal activation patterns in response to auditory stimuli.

The analysis of s-EEG signals, obtained by considering the contribution of all neurons in the network, revealed a clear mapping of (fake) brain activity correlated with the characteristics of the input sound. Variations in the amplitude and frequency of s-EEG oscillations emerged in response to changes in intensity, frequency, and the temporal structure of the audio signal.

Finally, adopting a dynamic threshold model capable of adapting to the statistics of the input signal provided the network with

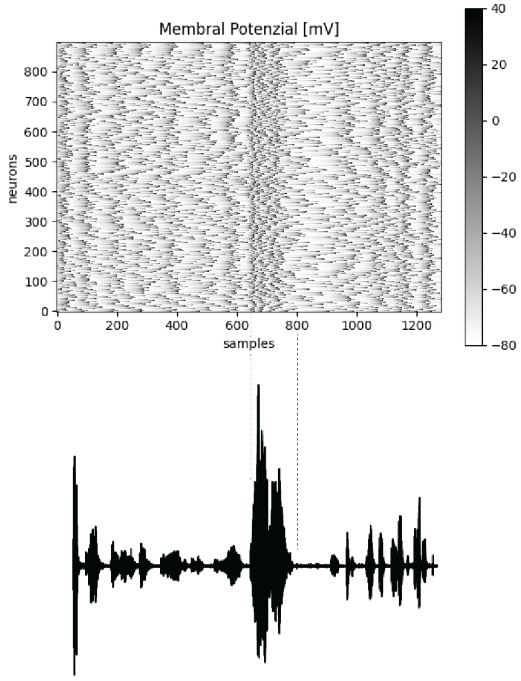

Figure 3: The plot provides an example of the visualization of the neuron's membrane potential over time. The input to the network is an audio signal recorded in a domestic environment, with a buffer size of 512 samples and a sampling rate equal to 12000 Hz. The waveform of the input audio signal is shown below the plot.

greater robustness and adaptability, enhancing the overall quality of the generated data.

To concretely illustrate the capabilities of the proposed model, several experiments were performed using a variety of audio signals as inputs and different model parameter configurations. These examples demonstrate of how the system responds to different types of auditory stimuli and how these reactions are reflected in dynamic and complex control structures.

In Figure 5 and Figure 3 two representative case studies are presented to highlight the model's performance in response to the varying characteristics of the considered stimuli (model parameters in Table 1).

Each neuron position j_i is unique (two neurons cannot occupy the same position) and is randomly selected during the model initialization (see example in Figure 4). This coordinate remains constant throughout all experiments.

The position of each neuron is constrained within a sphere whose radius corresponds to 90% of the average human skull radius, in order to allow for a small tolerance with respect to the surface where the electrodes are placed.

Figure 3 clearly illustrates the significant correlation between the membrane potential graph and the network's input stimulus.

It is evident that variations in the membrane potential reflect the temporal characteristics and amplitude of the stimulus, highlighting a synchronized response of the neural network to the incoming impulses (each vertical colum in Figure 3 represents a single frame consisting of n-samples, as show in Equation (18)).

Table 1: Experimental setup of the model

Description	Variable	Value
Time Interval (audio)	T_c	$8 \cdot 10^{-5}s$
Time Interval (EEG)	dt	$2 \cdot 10^{-3}s$
Total Number of Neurons	N_{neuro}	900
N. of excitatory neurons	exc	$\lfloor 0.91 \cdot N_{neuro} \rfloor$
N. of inhibitory neurons	inh	$N_{neuro} - exc$
Exciters Parameters	a	0.02
	b	0.25
	c	-30.0
	d	8.0
Inhibitory Parameters	a	0.02
	b	0.25
	c	-30.0
	d	2.0
STDP Parameters	a_{pre}	10^{-3}
	a_{post}	10^{-3}
	t_{pre}	dt
	t_{post}	dt
Sparsity	s	0.05 %
Electrodes Model		10-20 3D, 24-ch
Min. Spike Interval (IHC)	m	$10^{-4}s$
Spike Gain Factor (IHC)	γ	10^4
Stimulus Factor	α	10^3
Threshold Factor	β	10^3
Noise Range (stimulus)	η	$[0.0, 1.0]$
Interpolation Mode		No-Interp
Frame Size	N	512 (samples)
IHC Parameters	A	5.0
	B	300.0
	g	2000.0
	y	5.05
	λ	2500.0
	r	6500.0
	x	6631.0
	M	1.0
	h	1.0

This observation suggests that the electrical activity of the membrane is closely tied to changes in the input stimulus, demonstrating a direct link between external stimuli and the generated system response. This indicates, to some extent, that the network behaves consistently with expectations. The transformation applied, first through the IHC and then through SNN, confirms that the s-EEG analysis will be conducted on data that faithfully reproduce the stimulus response.

For the s-EEG, the electrodes were arranged as shown in Table 2. The 3D coordinates of the electrodes (see Table 2) represent their positions on the surface of a sphere expressed in normalized values (from -1 to 1). The electrodes are arranged along the equator of the sphere, an imaginary circle located midway between the north and south poles. Their positions follow the international 10-20 system. The sphere is considered unitary. This standardization simplifies the representation of electrode positions according to the 10-20 system, enabling consistent comparison and analysis of EEG recordings. Each electrode is labeled based on the simulated region where it is positioned (see Figure 2, additional labels include LPA (Left Preauricular) in front of the left ear, RPA (Right Preauricular) in front of the right ear, and

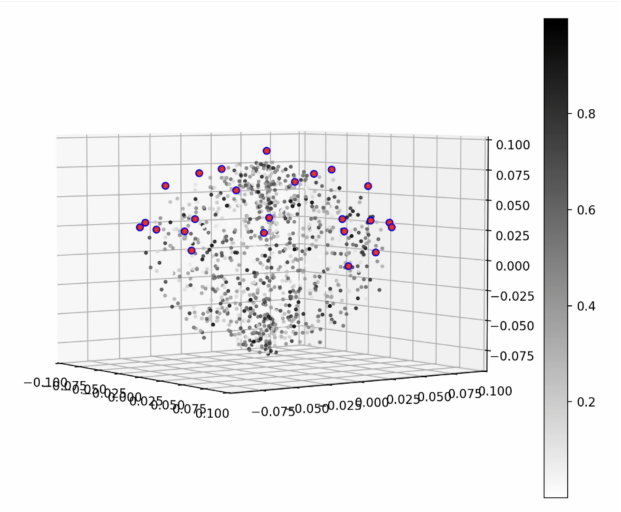

Figure 4: The image shows neurons arranged within a three-dimensional spherical space. The electrodes are positioned on the upper surface of the sphere.

Table 2: 3D representation of EEG electrode locations within normalized range coordinates

Label	x	y	z
C3	0.8090	0.0	-0.5878
C4	0.8090	0.0	0.5878
Cz	1.0	0.0	0.0
F3	0.6730	0.58	-0.4591
F4	0.6730	0.58	0.4591
F7	0.3090	0.5590	-0.7695
F8	0.3090	0.5590	0.7695
Fp1	0.3090	0.9045	-0.2939
Fp2	0.3090	0.9045	0.2939
Fpz	0.3090	0.9511	0.0
Fz	0.8090	0.5878	0.0
LPA	0.0	0.0	-1.0
NAS	0.0	1.0	0.0
O1	0.3090	-0.9045	-0.2939
O2	0.3090	-0.9045	0.2939
Oz	0.3090	-0.9511	0.0
P3	0.6730	-0.58	-0.4591
P4	0.6730	-0.58	0.4591
P7	0.3090	-0.5590	-0.7695
P8	0.3090	-0.5590	0.7695
Pz	0.8090	-0.5878	0.0
RPA	0.0	0.0	1.0
T7	0.3090	0.0	-0.9511
T8	0.3090	0.0	0.9511

NAS (Nasion), located at the point between the forehead and the nose). It is specified that, in the presented experiment, the coordinates listed in the table were appropriately scaled using a factor representing the average radius of a human skull, which was approximated as a sphere for modeling convenience.

In Figure 5, the result of the s-EEG analysis are presented. Graph display raw s-EEG recordings, collected without applying any filter.

Nevertheless, filtering remains a fundamental tool for signal analysis and denoising. They enhance data quality and isolate components relevant to the analysis of brain activity. Similar to human brain activity, the SNN activity while processing an audible stimulus can be segmented into frequency bands. This allows for identifying specific features of the audio signal that might otherwise go unnoticed.

Finally, without delving further into the details, it is important to highlight that the translation of an audio signal into an s-EEG domain would also allow for the use of traditional EEG analysis techniques, commonly applied to study human brain activity, such as coherence analysis, synchronization measures, event-related potential (ERP) analysis, and topographic mapping.

5 FUNCTIONAL PERSPECTIVES AND RESEARCH DIRECTIONS

While the core focus of this research lies in audio-to-s-EEG mapping, the potential applications extend well beyond theory.

The generated s-EEG signals could act as control interfaces for interactive systems, enabling neural patterns triggered by sound to drive actions such as controlling prosthetics or interacting with virtual environments.

This approach transforms simulated brain activity into a command language for digital interaction. The versatility of the proposed system could also be expressed in real-time control of sound synthesis, lighting, and visual effects, with the potential to create an immersive ecosystem where the brain's response dynamically shapes the user's sensory experience.

To demonstrate these possibilities, imagine a sophisticated audio system that captures and analyzes sounds in a specific environment. This system converts audio information into an adaptive framework that interacts with and learns from its environment.

While conceptually similar to Di Scipio's work [8], this approach is fundamentally different in substance. Through the integration of cutting-edge signal processing techniques — including multi-channel spatial filtering and adaptive feature extraction — the system dynamically maps intricate soundscapes onto pseudo-neural control signals. The system creates a two-way relationship with its environment: it uses audio data to adjust its parameters while simultaneously responding to environmental changes. This makes the environment an active part of the creative process rather than just a passive source of input.

Furthermore, the operator (9) has a key feature in its design: it can dynamically adjust the dimensions of the processed information based on specific application needs. The operator $A()$ (see Equation (9)) can adapt the number of s-EEG channels $x : \mathbb{R}^n \rightarrow \mathbb{R}^m$, where n and m represent the number of input and output dimensions. This flexibility allows for both dimensional reduction (for $n > m$), useful when simpler control signals are needed, and dimensional expansion (for $n < m$), beneficial for scenarios requiring higher-order feature spaces. For example, a complex audio signal could be mapped to a smaller set of control parameters for simple interactions, or expanded into a larger set of channels for more detailed analysis and control. This adaptability makes the system particularly versatile for different applications, from basic sound parameter control to complex interactive installations.

These represent just some of the potential applications of the proposed system. Its functional perspectives further expand when considering the possibilities offered by EEG analysis techniques,

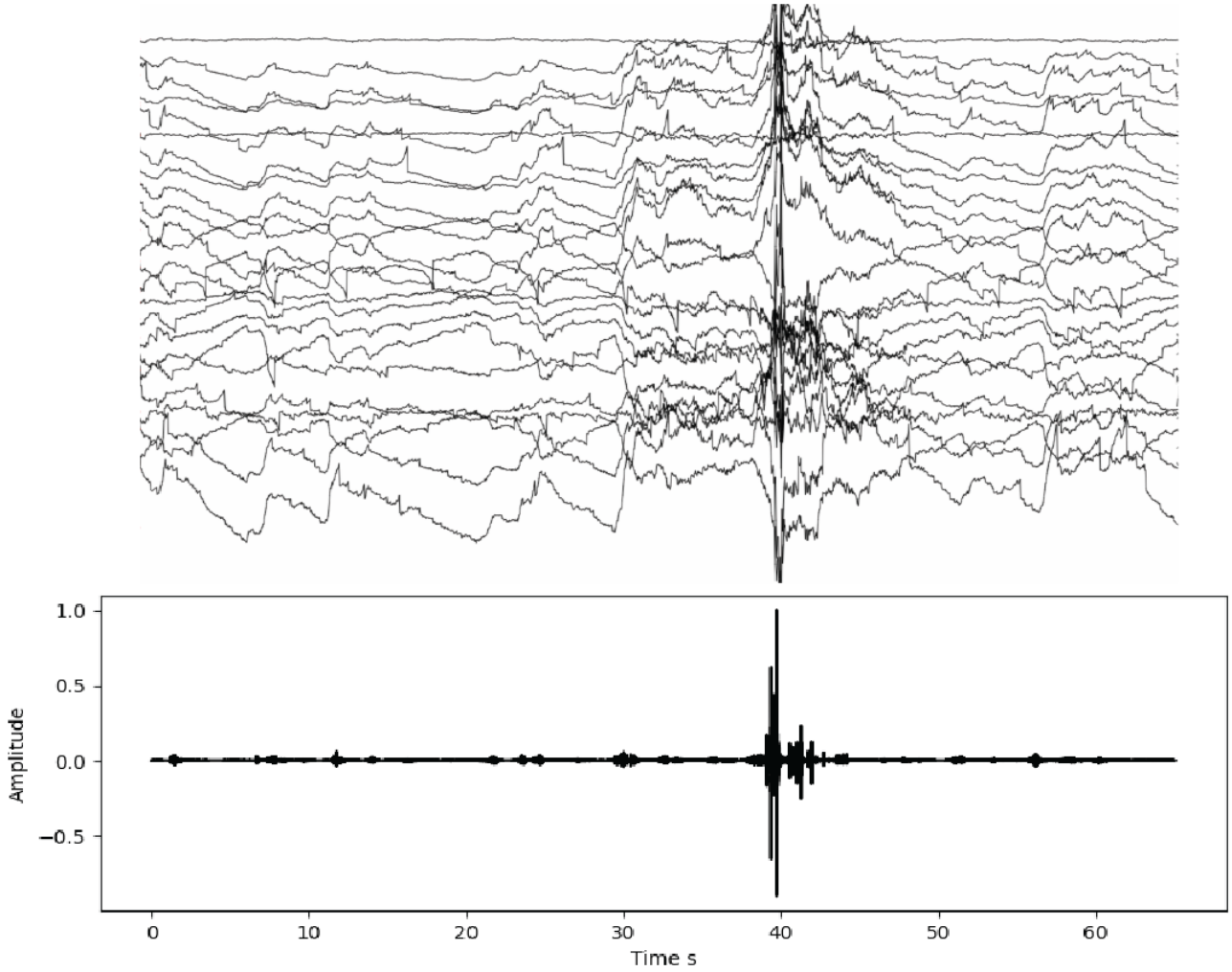

Figure 5: The image shows an example of s-EEG corresponding to an event captured in real time within a domestic environment. The associated audio signal is displayed at the bottom of the figure, while the model parameters are listed in the Table 1.

as mentioned at the end of Section 4. Such techniques would allow for an even deeper exploration of new dimensions in (simulated) neuro-informed interaction and control.

5.1 Future Works

The next step following the development of the audio-to-s-EEG mapping framework will focus on its practical application (as discussed above in Section 5), with the aim of concretely evaluating its effectiveness and versatility in real-world contexts.

This experimental phase will not only validate the proposed framework but will also pave the way for potential optimizations and new development directions, guided by concrete application needs and field experience.

6 CONCLUSION

The proposed method represents an initial step toward a bio-inspired model for audio information analysis and mapping. While the current implementation demonstrates coherence and innovation, several aspects warrant further investigation: the random arrangement of neurons in 3D space, the spike generation threshold function, and the parameter selection criteria. Although these elements may not be critical limitations in artistic applications, their refinement could enhance the model's performance and biological accuracy. A notable observation from the simulations is the imperfect repeatability of results, even with fixed seeds and neural arrangements. This variability likely stems from the network's stochastic components, cumulative numerical errors, and the implemented STDP learning model, which continuously modifies synaptic strengths based on spike timing. The STDP's sensitivity to neural event timing can significantly influence the learning process and network responses, while the network may require extended periods to reach stabilization.

In an artistic context, this non-deterministic behavior can be considered a valuable feature rather than a limitation, contributing to unique and unrepeatable experiences where unpredictability becomes an integral part of artistic creation. Each interaction with the system, influenced by simulated complex brain dynamics, could generate singular sound and visual expressions. With further development and optimization, this approach may develop into a versatile tool for analyzing and generating control maps from complex acoustic phenomena, serving both artistic

and technical applications while offering more precise and reliable data in performance engineering.

7 ETHICAL STANDARDS

This work adheres to the ethical standards of the NIME community [23]. There are no potential conflicts of interest, either financial or non-financial, associated with this research.

References

[1] Edgar Douglas Adrian and Brian HC Matthews. 1934. The Berger rhythm: potential changes from the occipital lobes in man. *Brain* 57, 4 (1934), 355–385.

[2] Miguel Angrick, Christian Herff, Emily Mugler, Matthew C Tate, Marc W Slutzky, Dean J Krusienski, and Tanja Schultz. 2019. Speech synthesis from ECoG using densely connected 3D convolutional neural networks. *Journal of neural engineering* 16, 3 (2019), 036019.

[3] Anthony Au, Yuchen Wang, and Sibai Xie. 2013. *Modeling of Mechanoelectrical Transduction of Hair Cells to Action Potentials in the Auditory Nerve.* Technical Report. University of California, San Diego. https://isn.ucsd.edu/courses/beng260/2013/reports/Au_Wang_Xie.pdf

[4] Ludovic Bellier, Anaïs Llorens, Déborah Marciano, Aysegul Gunduz, Gerwin Schalk, Peter Brunner, and Robert T Knight. 2023. Music can be reconstructed from human auditory cortex activity using nonlinear decoding models. *PLoS biology* 21, 8 (2023), e3002176.

[5] Natalia Caporale and Yang Dan. 2008. Spike timing–dependent plasticity: a Hebbian learning rule. *Annu. Rev. Neurosci.* 31, 1 (2008), 25–46.

[6] Ian Daly. 2023. Neural decoding of music from the EEG. *Scientific Reports* 13, 1 (2023), 624.

[7] Yang Dan and Mu-Ming Poo. 2006. Spike timing-dependent plasticity: from synapse to perception. *Physiological reviews* 86, 3 (2006), 1033–1048.

[8] Agostino Di Scipio. 2003. 'Sound is the interface': from interactive to ecosystemic signal processing. *Organised Sound* 8, 3 (2003), 269–277.

[9] Jianchuan Ding, Bo Dong, Felix Heide, Yufei Ding, Yunduo Zhou, Baocai Yin, and Xin Yang. 2022. Biologically inspired dynamic thresholds for spiking neural networks. *Advances in Neural Information Processing Systems* 35 (2022), 6090–6103.

[10] Robert H. Fujii and Kenjyu Oozeki. 2006. Temporal Data Encoding and SequenceLearning with Spiking Neural Networks. In *Artificial Neural Networks – ICANN 2006*, Stefanos D. Kollias, Andreas Stafylopatis, Włodzisław Duch, and Erkki Oja (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 780–789.

[11] Donald Olding Hebb. 2005. *The organization of behavior: A neuropsychological theory.* Psychology press, New York, NY, USA.

[12] Alan L Hodgkin and Andrew F Huxley. 1952. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology* 117, 4 (1952), 500.

[13] Eugene M Izhikevich. 2003. Simple model of spiking neurons. *IEEE Transactions on neural networks* 14, 6 (2003), 1569–1572.

[14] Qian Liang and Yi Zeng. 2021. Stylistic composition of melodies based on a brain-inspired spiking neural network. *Frontiers in systems neuroscience* 15 (2021), 639484.

[15] Diego Lopez-Bernal, David Balderas, Pedro Ponce, and Arturo Molina. 2022. A state-of-the-art review of EEG-based imagined speech decoding. *Frontiers in human neuroscience* 16 (2022), 867281.

[16] Enrique A Lopez-Poveda and Almudena Eustaquio-Martín. 2006. 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.

[17] Warren S McCulloch and Walter Pitts. 1943. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics* 5 (1943), 115–133.

[18] Alistair McEwan and André van Schaik. 2000. A silicon representation of the Meddis inner hair cell model. In *Proceedings of the ICSC Symposia on Intelligent Systems & Application (ISA'2000).* ICSC Accademy Press, Wollongong, Australia, 1544–078.

[19] Ray Meddis. 1986. Simulation of mechanical to neural transduction in the auditory receptor. *The Journal of the Acoustical Society of America* 79, 3 (1986), 702–711.

[20] Ray Meddis. 1988. Simulation of auditory–neural transduction: Further studies. *The Journal of the Acoustical Society of America* 83, 3 (1988), 1056–1063.

[21] Ray Meddis, Michael J Hewitt, and Trevor M Shackleton. 1990. Implementation details of a computation model of the inner hair-cell auditory-nerve synapse. *The Journal of the Acoustical Society of America* 87, 4 (1990), 1813–1816.

[22] Eduardo Reck Miranda. 2006. Brain-computer music interface for composition and performance. *International Journal on Disability and Human Development* 5, 2 (2006), 119–126.

[23] Fabio Morreale, Nicolas Gold, Cécile Chevalier, and Raul Masu. 2023. NIME Principles & Code of Practice on Ethical Research. https://doi.org/10.5281/zenodo.7545682

[24] David C Mountain and Alan R Cody. 1999. Multiple modes of inner hair cell stimulation. *Hearing research* 132, 1-2 (1999), 1–14.

[25] Solveig Næss, Geir Halnes, Espen Hagen, Donald J Hagler Jr, Anders M Dale, Gaute T Einevoll, and Torbjørn V Ness. 2021. Biophysically detailed forward modeling of the neural origin of EEG and MEG signals. *NeuroImage* 225 (2021), 117467.

[26] Standard Electrode Position Nomenclature. 1991. American electroencephalographic society guidelines for. *Journal of clinical Neurophysiology* 8, 2 (1991), 200–202.

[27] Nathalie Pattyn, Jeroen Van Cutsem, Antonio Martin, Robin Hauffa, and Joern Ungerer. 2024. Scientific Foundation. In *Handbook of Mental Performance.* Routledge, London, UK, 5–51.

[28] F. Rosenblatt. 1957. *The Perceptron, a Perceiving and Recognizing Automaton Project Para.* Cornell Aeronautical Laboratory, Buffalo, NY. https://books.google.it/books?id=P_XGPgAACAAJ

[29] David Rosenboom. 1976. *Biofeedback and the Arts: Results of Early Experiments.* Aesthetic Research Centre of Canada, Vancouver, B. C., Canada.

[30] David Rosenboom. 1990. The performing brain. *Computer Music Journal* 14, 1 (1990), 48–66.

[31] Marek Rudnicki, Oliver Schoppe, Michael Isik, Florian Völk, and Werner Hemmert. 2015. Modeling auditory coding: from sound to spikes. *Cell and tissue research* 361 (2015), 159–175.

[32] Shankha Sanyal, Sayan Nag, Archi Banerjee, Ranjan Sengupta, and Dipak Ghosh. 2019. Music of brain and music on brain: a novel EEG sonification approach. *Cognitive neurodynamics* 13 (2019), 13–31.

[33] Margitta Seeck, Laurent Koessler, Thomas Bast, Frans Leijten, Christoph Michel, Christoph Baumgartner, Bin He, and Sándor Beniczky. 2017. The standardized EEG electrode array of the IFCN. *Clinical neurophysiology* 128, 10 (2017), 2070–2077.

[34] Alborz Rezazadeh Sereshkeh, Rozhin Yousefi, Andrew T Wong, and Tom Chau. 2018. Online classification of imagined speech using functional near-infrared spectroscopy signals. *Journal of neural engineering* 16, 1 (2018), 016005.

[35] Christian J Sumner, Enrique A Lopez-Poveda, Lowel P O'Mard, and Ray Meddis. 2002. A revised model of the inner-hair cell and auditory-nerve complex. *The Journal of the Acoustical Society of America* 111, 5 (2002), 2178–2188.

[36] Brandon J Thio and Warren M Grill. 2023. Relative contributions of different neural sources to the EEG. *Neuroimage* 275 (2023), 120179.

[37] Katrina Wendel-Mitoraj. 2010. *The Influence of Tissue Conductivity and Head Geometry on EEG Measurement Sensitivity Distributions.* Ph. D. Dissertation. Tampere University of Technology.