Stimulus probability driven development of auditory networks

Project-II report submitted to

Indian Institute of Technology Kharagpur

in partial fulfilment for the award of the degree of

Bachelor of Technology

in

Electrical Engineering

by

Sohini Gupta

(20EE38032)

Under the supervision of

Dr. Sharba Bandyopadhyay

Department of Electronics and Electrical Communication Engineering.

Indian Institute of Technology Kharagpur

Spring Semester, 2023-24

April 29, 2024

DECLARATION

I certify that

(a) The work contained in this report has been done by me under the guidance of my supervisor.

(b) The work has not been submitted to any other Institute for any degree or diploma.

(c) I have conformed to the norms and guidelines given in the Ethical Code of Conduct of the Institute.

(d) Whenever I have used materials (data, theoretical analysis, figures, and text) from other sources, I have given due credit to them by citing them in the text of the thesis and giving their details in the references. Further, I have taken permission from the copyright owners of the sources, whenever necessary.

Date: April 29, 2024 (Sohini Gupta)
Place: Kharagpur (20EE38032)

DEPARTMENT OF ELECTRONICS AND ELECTRICAL COMMUNICATION ENGINEERING.

INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR

KHARAGPUR - 721302, INDIA

CERTIFICATE

This is to certify that the project report entitled "**Stimulus probability driven development of auditory networks**" submitted by **Sohini Gupta** (Roll No. 20EE38032) to Indian Institute of Technology Kharagpur towards partial fulfilment of requirements for the award of degree of Bachelor of Technology in Electrical Engineering is a record of bona fide work carried out by him under my supervision and guidance during Spring Semester, 2023-24.

Date: April 29, 2024

Place: Kharagpur

Dr. Sharba Bandyopadhyay
Department of Electronics and Electrical Communication Engineering.
Indian Institute of Technology Kharagpur
Kharagpur - 721302, India

Acknowledgements

I would like to express my special thanks to my Thesis supervisors, **Dr. Sharba Bandyopadhyay**, for his guidance and support throughout the project. He has guided me with ideas and suggestion all the way along the project. They have consistently motivated me in exploring ideas and strategies relevant to the project.

I would also like to thank the IPL Lab members who have always been there for me throughout the project for helping me get accustomed to this field and showing me experiments around the laboratory which has inspired me to know new things.

In conclusion, I would like to acknowledge that this project would not have been possible without the support from the Department of Electrical Engineering and Department of Electronics and Electrical Communication Engineering, IIT Kharagpur. I am sincerely grateful to everyone who helped me in this project.

Abstract

Name of the student: **Sohini Gupta** Roll No: **20EE38032**
Degree for which submitted: **Bachelor of Technology**
Department: **Department of Electronics and Electrical Communication Engineering.**
Thesis title: **Stimulus probability driven development of auditory networks**
Thesis supervisor: **Dr. Sharba Bandyopadhyay**
Month and year of thesis submission: **April 29, 2024**

This work investigates the theoretical underpinnings of how early exposure to rare stimuli can enhance neural responsiveness. We approach this phenomenon using a network model and formulate it as a mathematical optimization problem. The model explores the interplay between a stimuli train containing both frequent and rare events, (oddball train) and the response of a network of interconnected neurons. The objective function comprises two key terms: Mutual Information: This term quantifies the information shared between the network's response (R) and the individual stimulus tokens (T) and Sparsity Constraint: This term contains the mean rates, the sparse activity requires a minimal response by the network of neurons to code the two stimuli. By analyzing how the parameters (weights) connecting the neurons evolve, we aim to understand how the network optimizes its internal representation to maximize the mutual information between the stimuli and the response.

Keywords: Surprise, Mutual Information, Sparsity, Standard, Deviant

Abbreviations

- **CI**-Constrained Information
- **CSI** - Cumulative SSA Index
- **ECO** - Ear Canal Opening
- **EEG** - Electroencephelography
- **L4** - Layer 4 neurons
- **MEG** - Magnetoencephalography
- **MMN** - Mis-match negativity
- **SP** - sub-plate neuron
- **SSA**- Stimulus Specific Adaptation

Contents

Chapter 1

Introduction

The rare sound objects present in the environment carry important information about the physical phenomenon happening in the environment. Being capable of identifying, processing and inferring from those rare sound objects is fundamental for the survival of an organism. The present understanding of neural development states that an important developmental epoch known as the critical period is crucial for the cortical maturation and refinement of the neural circuitry within the auditory pathway. However, a recent study demonstrated that if certain stimuli features are salient enough to stand out in the stimuli space, like a rare event, the auditory circuitry is mature enough to pick it up and extract information from it much before the critical period. A simple framework using computational modeling was proposed by Adarsh Mukesh and we would provide a mathematical optimization to explain how and why an early-developing neural circuit before the critical period can process salient information from the stimuli.

While a network model mechanistically addresses the observations, we theoretically address the observations of early exposure to a low probability stimulus leading to the strengthening of its responses . Our observations and model results are contrary to the general ideas of long-term plasticity; a repeated presentation of a stimulus in the critical period leads to the strengthening of the representation of that particular stimulus in the long-term [1] [3] [6] as theoretically expected from standard Hebbian plasticity. Thus, it is important to reconcile our observations with previous observations from a theoretical perspective as well. Auditory cortical spiking activity is known to be sparse, and sparse coding principles theoretically explain several known receptive field types in the auditory pathway (Hromádka et al., 2008; Smith & Lewicki, 2006). The neural activity generated in response to stimuli encodes a plethora of statistical information contained in the spectral and temporal features of the stimuli. Finding direct mathematical correlates between the generated response and the stimuli becomes difficult. Deterministic methods that try to explain the observed neural activity as a function of stimuli features have not been fruitful. From simple linear and non-linear measures like the reverse correlations (Chichilnisky, 2001), wiener kernels, and Volterra series expansion (Y. W. Lee & Schetzen, 2007) to complex generalised models (McFarland et al., 2013) have only gained limited success. Polynomial approximation methods are only good for neurons whose computations can be approximated with lower-order Taylor representations and for simple low-dimensional stimuli (Barbour & Wang, 2003a; Reiss et al., 2007). Hence, information theoretic measures like mutual information are better tools to study the interactions between stimuli and response [5].

We perform a mathematical optimization in this work to study the interactions between an oddball stimuli train and the response to study the interaction between the two. In this work, we have tried to look into the parameters that will increase the mutual information between the stimuli and response. We also put a sparse coding constraint to see if a sparse regime can force the parameters to take values which would lead to mutual information maximization. Finally, we take the assumptions of higher selectivity towards deviant before the ECO to see the probable range of sparsity over which mutual information maximization can take place.

Chapter 2

Literature Review

2.1 Response to frequent and infrequent stimuli

2.1.1 Mis-match negativity

The consequence of a higher selectivity to rare stimuli features and attenuation of frequent ones is reflected at multiple modalities ranging from sensory perception to motor output. A well-known and well-studied manifestation of selectivity towards rare stimuli and features is the Mis-match negativity (MMN). Mis-match negativity is an event related potential change which is recorded in the brain in response to an odd stimulus embedded within a stream of successively occurring stimulus. This sudden change in the event potential is recorded using electroencephalography (EEG) or magnetoencephalography (MEG). MMN has been widely studied using auditory oddball sound sequences where an oddball sequence of sounds consisting of a more frequent stimulus (known as standard) is presented with a less frequent stimulus (called as deviant) embedded in between [4]. As suggested by the name, the negative component of the wave form obtained by subtracting the event related potentials evoked by standard from that of deviant is characteristic of a normal MMN profile.

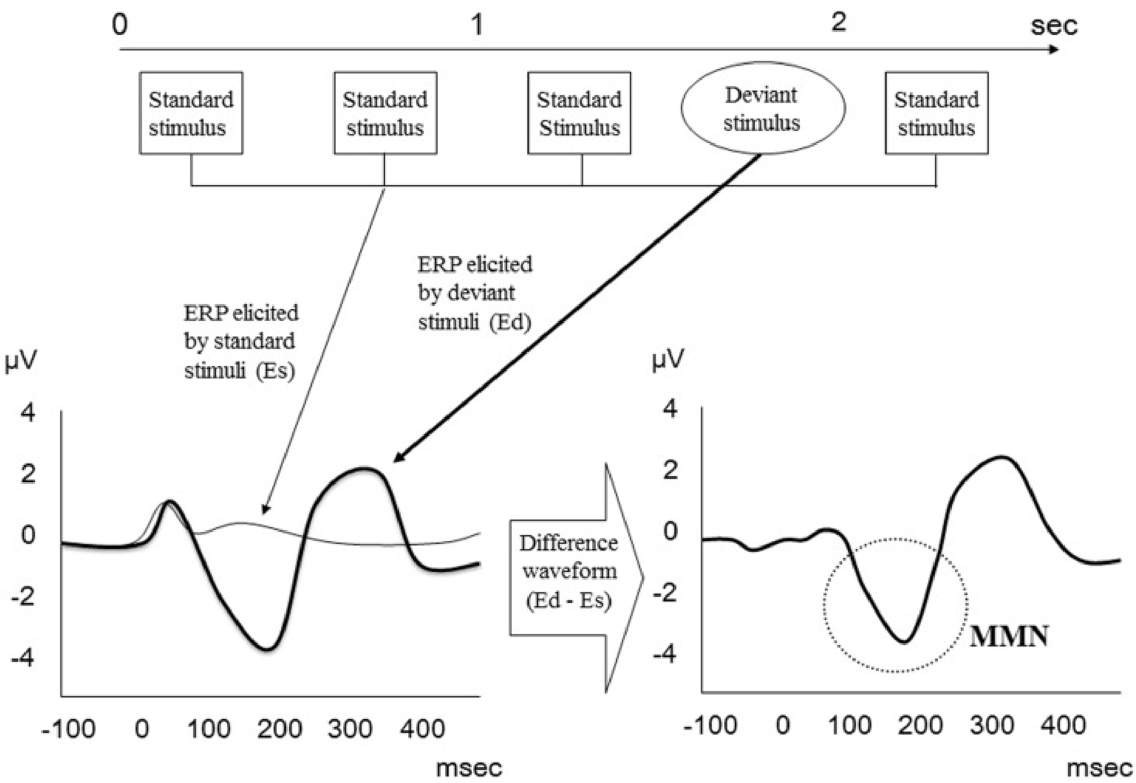

FIGURE 2.1: Mis-match negativity for a healthy subject

2.1.2 Detection of rare event at single neuron level

In the previous section, we discussed the large-scale manifestation of selectivity to deviant stimulus, the mis-match negativity. The selectivity towards the deviant stimulus is reflected also at a micro-scale, at the level of a single neuron [7] [8]. On repeated presentation of a single stimulus, the spike rate keeps on decreasing and later adapts, however when a deviant stimulus is presented, the rate increases sharply, implying the importance of that rare stimulus. This phenomenon is known as Stimulus Specific Adaptation (SSA) and has been extensively studied and characterized throughout the years.

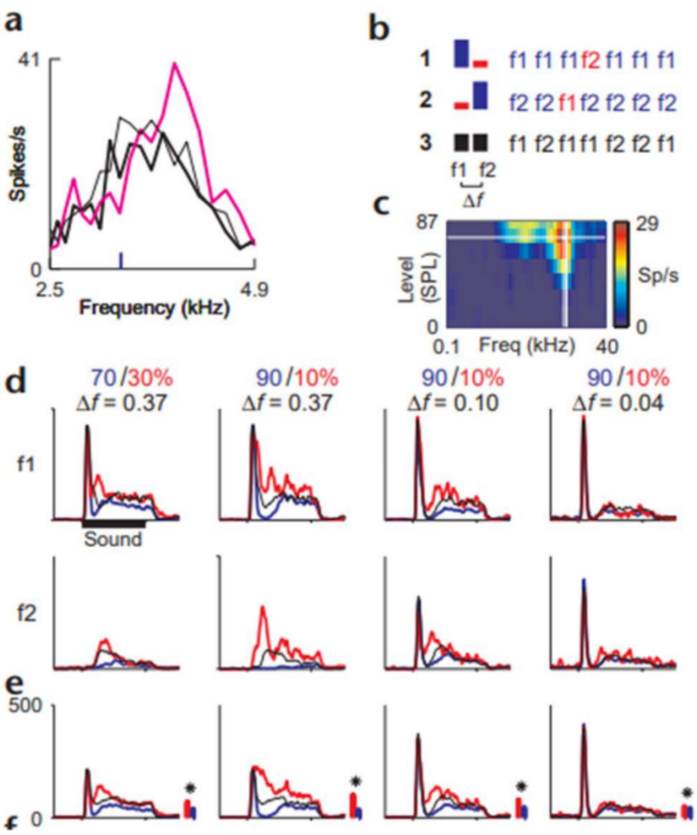

FIGURE 2.2: Response of single neuron demonstrating SSA recorded from primary auditory cortex. a) Tuning curve of a given neuron showing frequency selectivity before adaptation (thick black line), during adaptation (magenta line) and after adaptation (after 30s of recovery, thin black line). b) Oddball stimulus paradigm: 1. f1 is standard and f2 is deviant. 2. f1 is deviant and f2 is standard and 3. Both occur equiprobably at 50% each. c) Receptive field of the neuron. The two vertical white lines show f1 and f2 with a 0.1 octave gap between them. The horizontal white line is the intensity level at which the two frequency tones were presented in the oddball paradigm. d) Response for f1 and f2 (top and bottom) when presented as standard (blue), deviant (red) and equiprobable (black). e) Response averaged for f1 and f2. The four probability conditions with their frequency gaps shown on top. (From Ulanovsky et. al. 2003)

An important aspect of the selectivity towards a stimulus feature shown by the neurons of the auditory cortex is the context in which the stimulus is being presented. The same frequency tone when it is presented as a deviant elicits the highest rate response, followed by an equiprobable presentation and the least response for a standard presentation. When presented as a deviant, the rarer the deviant presentation becomes, the stronger response it can elicit.This is because a deviant which occurs at a 10% probability carries more novelty than when it occurs at a 30% probability. In addition to this, a higher relative difference between the standard and deviant frequencies leads to a higher response difference between deviant and standard [7] [8] [9]. The selectivity shown by a neuron to different frequencies which is reflected on the spike rates and quantified by its tuning curve is dependent on the state of the neuron. If the frequency for which the neuron is showing the highest selectivity (also called as best frequency or preferred frequency) is presented as the standard along with some other frequency as deviant, we see that after repeated presentation of the standard frequency, in an adapted state, the selectivity curve of the neuron changes and the peak shifts towards the deviant frequency [8]. Thus, the selectivity of a neuron for a particular stimulus can be altered by changing the context in which it is being presented. Selectivity towards the context of the stimulus is a fundament property that is novel enough to be identified by the bottom-up filters.

2.1.3 Response to an oddball sound stimulus before the Ear Canal opening

Studies have shown that the presence of responses in early ages before the critical period shows that the auditory circuitry right from the sensory periphery till the cortex is capable of processing simple sound stimuli like pure tones. Robust responses were found both in the SP and L4 neurons, although the response characterization shown by them was poor as compared to the adult case. We had discussed in the previous sections about the SSA where the response to the repeated presentation of a stimulus gets adapted, but when a deviant stimulus is presented, the response gets amplified. This phenomenon was discussed both from a bottom-up SSA framework and a top-down prediction error approach. Phenomena like SSA or prediction error have not been investigated in early developing ages. We had discussed in the earlier sections that a stimulus presented as a deviant is salient enough to be picked up by the bottom-up selectivity filters. The characterization shown by the SP and the L4 neurons for pure frequency tones was poor in the early ages [5]. However, if a stimulus is presented in a salient enough manner (like a deviant), it could be possible to observe more robust responses and higher selectivity in the SP and L4 neuron before ECO also. One recent study reported very robust and temporally synchronized response of SP neurons in response to deviant tones. The SP neurons show a higher selectivity for deviant stimuli as compared to the L4 neurons before the ECO. However, after the ECO and during the critical period, the L4 neurons show a higher selectivity for the deviant stimuli. After the ECO, this trend gets reversed and the L4 neurons start showing a higher deviant selectivity as compared to the SP neurons. Overall, these results show us that much before the ECO and the onset of the accepted critical period, the neural circuitry in the auditory pathway is well active and mature enough to identify and encode novel sound objects. The SP neurons play a pivotal role in this encoding which is reflected in their higher selectivity values.

2.1.4 Exposure to oddball sound sequences before ECO

Perturbing the natural auditory environment of an animal during the critical period can lead to changes in the auditory cortical maturation which can persist till adulthood [3] [2] [10]. The exposure protocol has been previously used to characterize the frequency selectivity during the critical period [3]. Most of these studies have used pure tones or broadband noise for exposure. In the previous section we discussed the high selectivity shown by the SP neurons to stimulus presented as deviant as compared to a control single presentation. Thus, it is intuitive to think that exposure to oddball sound presentation before the critical period may lead to changes that can persist till adulthood. The recent study performed exposure in different developing age groups in mice and found that exposure to an oddball stimulus paradigm just before the onset of critical period in the ages of P6-11 can lead to a higher selectivity being shown by L4 neurons in adulthood for the frequency which was used as the deviant in the oddball exposure protocol[5]. Unlike some previous studies where exposure during critical periods lead to the distortion of tonotopy in adulthood [3] [2], this study did not find any such observation and quantified the increased selectivity by measuring the spike rates corresponding to different frequencies in adulthood. A higher rate response recorded from an L4 neuron for a certain frequency over others could be a representation of stronger thalamic input corresponding to that frequency as compared to others. In the same study, on exposing to the oddball paradigm during the accepted critical period (P11-15), the L4 neurons showed a higher selectivity for the frequency which was presented as standard as compared to the deviant. In addition to this, repeated presentation of just the standard frequency also gave the same result This falls in accordance with the previous studies where exposure to a pure tone during critical period led to its over representation in the adult auditory cortex [3]. No significant difference between selectivity towards standard and deviant frequency was observed in either very early ages (P0-5) and ages after the critical period (P16-21). Thus, these results show that exposure to oddball sound sequences just before the ECO and after the ECO during the accepted critical period strengthens the thalamic to L4 inputs corresponding to the deviant and standard frequency respectively.

Chapter 3

Scope and Objective

Most of the results which we have discussed above regarding the importance of SP neurons and the selectivity before ECO are present in bits and pieces and lack a unified mechanism to explain how before the ECO, the SP neurons show high deviant selectivity, while after the ECO the L4 neurons start showing high deviant selectivity[5]. To address this issue, a computational model had been presented by Adarsh Mukesh in his PhD thesis in Chapter 3 involving SP neurons, L4 neurons and thalamic inputs to demonstrate how selectivity towards deviant stimulus is developed in the L4 neurons and how SP neurons aid in this. Through my work, I am further trying to explain why the deviant stimulus should carry more importance than the standard in for a network of neurons using a mathematical explanation based on mutual information and sparse coding.

Objective:

- To construct a mathematical model for a network of neurons demonstrating the development of selectivity for deviant stimulus by the SP and cortical neurons.
- To mathematically explain the importance of a rare stimulus and why neurons show more selectivity towards it.

Chapter 4

Work Progress and Achievement

4.1 Sparse coding and information maximization principles imply a strengthening of response to low probability stimuli

We assumed that a minimal activity-based coding in network of neurons underlying sparse representation in populations and that such representation is one of developing auditory system's goals while maximizing information about the stimulus in the responses. We use an objective function (eq1) as the desired optimization performed by activity-driven plasticity with these principles. The objective function has a mutual information term between the response (r) and the stimuli token (s) that will increase as successive stimuli tokens come. A sparseness constraint term containing the sum of the mean rate of each neuron limits the increase in mutual information. We call this metric as the Constrained Information (CI).

Constrained Information(CI)= Mutual Information(MI)+ λ (mean rate) ...(eq1)

To keep the solution mathematically tangible and straightforward to understand, we first start to optimize with the most basic network which is a network of two neurons with two tones, one standard(S, with a higher probability) and the other deviant (D, occurring with a lower probability) and then generalize it for m neurons with many standards and many deviants along with recurrence. We also took the

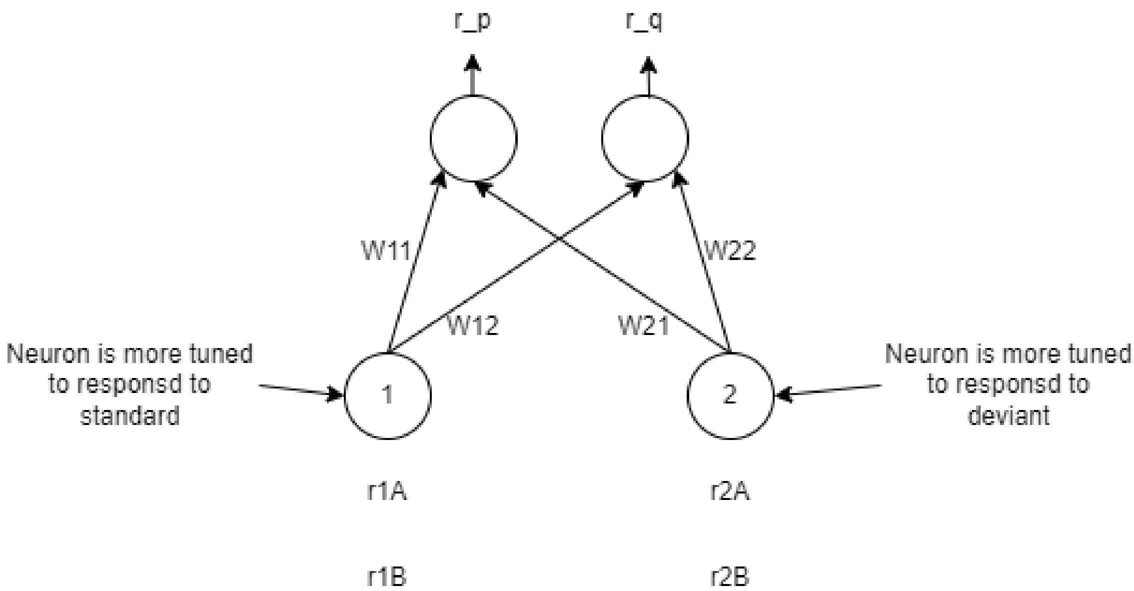

FIGURE 4.1: Diagram of the network of two neurons. Since neuron 1 is more tuned to respond to standard so r1A > r2A. Similarly, neuron 2 is more tuned to respond to deviant so r1B < r2B

liberty of considering the probability distribution of the responses for standard and deviant as Gaussian with a mean and variance to be same. This is because we know that the spiking rate of neurons is a poison distribution.

$$Prob.(response|stimulus) = p(\vec{r}|s_{standard}) \sim \mathcal{N}(\vec{\mu}_A, \, \Sigma_A)\,.$$

$$Prob.(response|stimulus) = p(\vec{r}|s_{deviant}) \sim \mathcal{N}(\vec{\mu}_B, \, \Sigma_B)\,.$$

where

$$\vec{\mu}_A = \begin{bmatrix} W11.r1A + W21.r2A \\ W12.r1A + W22.r2A \end{bmatrix}$$

$$\vec{\mu}_B = \begin{bmatrix} W11.r1B + W21.r2B \\ W12.r1B + W22.r2B \end{bmatrix}$$

$$\Sigma_A = \begin{bmatrix} W11.r1A + W21.r2A & 0 \\ 0 & W12.r1A + W22.r2A \end{bmatrix}$$

$$\Sigma_B = \begin{bmatrix} W11.r1B + W21.r2B & 0 \\ 0 & W12.r1B + W22.r2B \end{bmatrix}$$

We fixed the response distribution for the standard token, while the mean of the response distribution for the deviant token can be higher (or lower) by a factor of a_k, where k denotes the kth token. The stimuli are being presented in a continuous stream of sound tokens where the occurrence probability of deviant is $p(s_{deviant}) = x$, and that of standard is $p(s_{standard}) = 1 - x$ .Thus, the expected response (mean) or the neuron at the kth token will be,
mean rate(neuron p) = $(r_p+a_{kp}).p(s_{deviant})+ r_p.\ p(s_{standard})$ = (r_p+a_{kp}).x+ r_p.(1-x).
mean rate(neuron q) = $(r_q+a_{kq}).p(s_{deviant})+ r_q.\ p(s_{standard})$ = (r_q+a_{kq}).x+ r_q.(1-x).
Thus the total mean rates will be to be sum of the mean responses of each neuron.

Thus, the constrained information at the kth token will be

$$CI_k = I(\vec{r}_k, \vec{s}) + \lambda.(r_p + r_q + x(a_{kp} + a_{kq})) \tag{4.1}$$

We will mathematically see how a_{kp} and a_{kq} are changing with each weight.

We know, $I(\vec{r}_k, \vec{s}) = H(\vec{r}_k) - H(\vec{r}_k|\vec{s})$ We will consider $H(\vec{r}_k)$ to be a constant. And we know $H(\vec{r}_k|\vec{s}) = x[\frac{1}{2}ln|\Sigma_A| + (1 + ln(2\pi))] + (1-x)[\frac{1}{2}ln|\Sigma_B| + (1 + ln(2\pi))]$

Now, $\frac{dln|\Sigma_A|}{dW11} = \frac{r1A}{W11.r1A+W21.r2A}$, $\frac{dln|\Sigma_B|}{dW11} = \frac{r1B}{W11.r1B+W21.r2B}$

Differentiating equation 4.1 wrt W11:

$\frac{dCI_k}{dW11} = -[\frac{(1-x)}{2}.\frac{r1A}{W11.r1A+W21.r2A} + \frac{x}{2}.\frac{r1B}{W11.r1B+W21r2B}] + \lambda r1A + \lambda x\frac{da_k}{dW11}$

$\implies -[\frac{(1-x)}{2}.\frac{r1A}{W11.r1A+W21.r2A} + \frac{x}{2}.\frac{r1B}{W11.r1B+W21r2B}] + \lambda r1A + \lambda x\frac{da_k}{dW11} \geq 0$

$\Longrightarrow \lambda x\frac{da_k}{dW11} \geq [\frac{(1-x)}{2}.\frac{r1A}{W11.r1A+W21.r2A} + \frac{x}{2}.\frac{r1B}{W11.r1B+W21r2B}] - \lambda r1A$

$\Longrightarrow \lambda x\frac{da_k}{dW11} \geq \frac{x}{2}[\frac{r1B}{W11.r1B+W21r2B} - \frac{r1A}{W11.r1A+W21.r2A}+] + \frac{1}{2}.\frac{r1A}{W11.r1A+W21.r2A} - \lambda r1A$

Now since r1A > r2A and r1B < r2B,

We have $\frac{da_k}{dW11} \geq (-)ve$ when $\lambda \geq \frac{1}{2(W11.r1A+W21.r2A)}$

Similarly differentiating equation 4.1 wrt W21:

$\frac{dCI_k}{dW21} = -[\frac{(1-x)}{2}.\frac{r2A}{W11.r1A+W21.r2A} + \frac{x}{2}.\frac{r2B}{W11.r1B+W21r2B}] + \lambda r2A + \lambda x\frac{a_k}{dW21}$

$\Longrightarrow -[\frac{(1-x)}{2}.\frac{r2A}{W11.r1A+W21.r2A} + \frac{x}{2}.\frac{r2B}{W11.r1B+W21r2B}] + \lambda r2A + \lambda x\frac{da_k}{dW21} \geq 0$

$\Longrightarrow \lambda x\frac{da_k}{dW21} \geq [\frac{(1-x)}{2}.\frac{r2A}{W11.r1A+W21.r2A} + \frac{x}{2}.\frac{r2B}{W11.r1B+W21r2B}] - \lambda r2A$

$\Longrightarrow \lambda x\frac{da_k}{dW21} \geq \frac{x}{2}[\frac{r2B}{W11.r1B+W21r2B} - \frac{r2A}{W11.r1A+W21.r2A}+] + \frac{1}{2}.\frac{r2A}{W11.r1A+W21.r2A} - \lambda r2A$

Now since r1A > r2A and r1B < r2B,
We have $\frac{da_k}{dW21} \geq (+)ve$ where $\lambda \leq c$;c can be found from the above equation

Similarly, we will get $\frac{da_k}{dW22} \geq 0$ when $\lambda =$ and $\frac{da_k}{dW12} \leq 0$ and get a specific value of λ which will satisfy all the constraints, this weightage of sparseness which will keep the mutual information increasing.

So from here we mathematically get that the mutual information increases with a sparsity constraint with the change in weights associated with the neuron whose firing rate to the deviant stimulus is more. In other words , there is maximum information extracted between the response and stimulus the stimuluus is a deviant tone. We would now want to extend it to the recurrent network an the model for a two neuron network would look like the following:

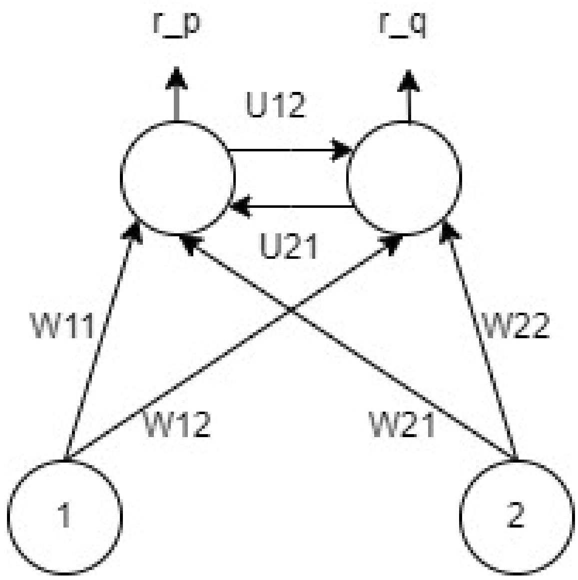

FIGURE 4.2: The figure is for a recurrent network

The following equation is for a recurrent network and the function may be a nonlinear function say sigmoid, ReLU or tanh.

$$\vec{r}_m[k] = f(\vec{W}_m^T.s[k] + \vec{U}_m^T.s[k-1])$$

where $\vec{U}_m$ is the recurrent weight vector, $\vec{W}_m$ is the weight matrix. $\vec{s}[k]$ is the vector that contains the tuning rates of each neuron to the stimulus at the kth instant.

Chapter 5

Future Work

- We wish to extend the proof for a more general network with neurons connected to each other establishing recurrence.
- The stimulus here was a single-tone frequency so we want to establish results wherein the standard and deviant is not just a tone but a sequence.
- Also, we considered one tone as standard and one as deviant, we wish to find results where there are multiple standards and multiple deviants.

Chapter 6

Conclusion

The optimization of mutual information between the stimuli and response subject to a sparsity constraint shows how at later ages the sparsity starts playing a dominant role in maintaining a monotonic increase of the mutual information with the incoming stimuli tokens. Although the mathematical explanation and the intuitive insights give us a decent understanding of why a rare stimulus becomes important for an early developing cortex to attend to, the need for an even sparser coding platform after the ECO has far-reaching consequences. The neurons before ECO are able to significantly respond to deviant stimulus only because the context in which it is presented is novel enough and thus, a sparse framework is not needed for efficient encoding of such stimuli features. However, since the circuitry is not fully mature, stimuli features that are not present in a salient enough context are characterized poorly. Hence, a sparse representation is needed to filter out the +redundant features and amplify the relevant ones.

Bibliography

[1] Tania Rinaldi Barkat, Daniel B Polley, and Takao K Hensch. A critical period for auditory thalamocortical connectivity. *Nature neuroscience*, 14(9):1189–1194, 2011.

[2] Clare L Beasley, Zhi J Zhang, Iain Patten, and Gavin P Reynolds. Selective deficits in prefrontal cortical gabaergic neurons in schizophrenia defined by the presence of calcium-binding proteins. *Biological psychiatry*, 52(7):708–715, 2002.

[3] Etienne de Villers-Sidani, Edward F Chang, Shaowen Bao, and Michael M Merzenich. Critical period window for spectral tuning defined in the primary auditory cortex (a1) in the rat. *Journal of Neuroscience*, 27(1):180–189, 2007.

[4] Teija Kujala, Mari Tervaniemi, and Erich Schröger. The mismatch negativity in cognitive and clinical neuroscience: theoretical and methodological considerations. *Biological psychology*, 74(1):1–19, 2007.

[5] Muneshwar Mehra, Adarsh Mukesh, and Sharba Bandyopadhyay. Separate functional subnetworks of excitatory neurons show preference to periodic and random sound structures. *Journal of Neuroscience*, 42(15):3165–3183, 2022.

[6] Christoph E Schreiner and Daniel B Polley. Auditory map plasticity: diversity in causes and consequences. *Current Opinion in Neurobiology*, 24:143–156, 2014.

[7] Nevo Taaseh, Amit Yaron, and Israel Nelken. Stimulus-specific adaptation and deviance detection in the rat auditory cortex. *PloS one*, 6(8):e23369, 2011.

[8] Nachum Ulanovsky, Liora Las, and Israel Nelken. Processing of low-probability sounds by cortical neurons. *Nature neuroscience*, 6(4):391–398, 2003.

[9] Tohar S Yarden and Israel Nelken. Stimulus-specific adaptation in a recurrent network model of primary auditory cortex. *PLoS computational biology*, 13(3):e1005437, 2017.

[10] Xiaoming Zhou and Michael M Merzenich. Enduring effects of early structured noise exposure on temporal modulation in the primary auditory cortex. *Proceedings of the National Academy of Sciences*, 105(11):4423–4428, 2008.