

Aus dem Bereich Theoretische Medizin und Biowissenschaften
der Medizinischen Fakultät der Universität des Saarlandes, Homburg/Saar

**Decoding Electrophysiological Correlates of Selective Attention by
Means of Circular Data**

*Dissertation zur Erlangung des Doktorgrades
der Naturwissenschaften (Dr. rer. nat.)
an der Medizinischen Fakultät
der Universität des Saarlandes
2019*

vorgelegt von:

Zeinab Schäfer (geb. Mortezapouraghdam), M.Sc.
geboren am 21. September 1988 in Tehran, Iran

Dekan: Prof. Dr. med. Michael D. Menger

Referent: Prof. Dr. rer. nat. Dr. rer. med. Daniel J. Strauss

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Bei der Auswahl und Auswertung folgenden Materials haben mir die nachstehend aufgeführten Personen in der jeweils beschriebenen Weise unentgeltlich/entgeltlich geholfen:

- Alle in der Danksagung des Dissertationsmanuskripts erwähnten Personen haben mir in der dort beschriebenen Weise unentgeltlich geholfen.

Weitere Personen waren an der inhaltlichen und materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich nicht die entgeltliche Hilfe von Vermittlungs- bzw. Beratungsdiensten (Promotionsberater/innen oder anderer Personen) in Anspruch genommen. Außer den Angegebenen hat niemand von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form in einem anderen Verfahren zur Erlangung eines Doktorgrades einer anderen Prüfungsbehörde vorgelegt.

In versichere an Eides statt, dass ich nach bestem Wissen die Wahrheit gesagt und nichts verschwiegen habe. Vor Aufnahme der vorstehenden Versicherung an Eides statt wurde ich über die Bedeutung einer eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung belehrt.

Saarbrücken, den 4.1.2019

Zeinab MortezaPouraghdam, M.Sc.

Abstract

Sustaining our attention to a relevant sensory input in a complex listening environment, is of great importance for a successful auditory communication. To avoid the overload of the auditory system, the importance of the stimuli is estimated in the higher levels of the auditory system. Based on these information, the attention is drifted away from the irrelevant and unimportant stimuli. *Long-term habituation*, a gradual process independent from sensory adaptation, plays a major role in drifting away our attention from irrelevant stimuli.

A better understanding of *attention-modulated neural activity* is important for shedding light on the encoding process of auditory streams. For instance, these information can have a direct impact on developing smarter hearing aid devices in which more accurate *objective* measures can be used to reflect the hearing capabilities of patients with hearing pathologies. As an example, an objective measures of long-term habituation with respect to different level of sound stimuli can be used more accurately for adjustment of hearing aid devices in comparison to verbal reports.

The main goal of this thesis is to analyze the *neural decoding signatures* of long-term habituation and neural modulations of selective attention by exploiting circular regularities in electrophysiological (EEG) data, in which we can objectively measure the level of attentional-binding to different stimuli. We study, in particular, the modulations of the instantaneous phase (IP) in event related potentials (ERPs) over trials for different experimental settings. This is in contrast to the common approach where the ERP component of interest is computed through averaging a sufficiently large number of ERP trials. It is hypothesized that a high attentional binding to a stimulus is related to a high level of IP cluster. As the attention binding reduces, IP is spread more uniformly on a unit circle. This work is divided into three main parts.

In the initial part, we investigate the dynamics of long-term habituation with different acoustical stimuli (soft vs. loud) over ERP trials. The underlying temporal dynamics in IP and the level of phase cluster of the ERPs are assessed by fitting circular probability functions (pdf) over data segments. To increase the temporal resolution of detecting times at which a significant change in IP occurs, an abrupt change point model at different pure-tone stimulations is used. In a second study, we improve upon the results and methodology by relaxing some of the constrains in order to integrate the *gradual process* of long-term habituation into the model. For this means, a Bayesian state-space model is proposed. In all of the aforementioned studies, we successfully classified between different stimulation levels, using solely the IP of ERPs over trials.

In the second part of the thesis, the experimental setting is expanded to contain longer and more complex auditory stimuli as in real-world scenarios. Thereby, we study the neural-correlates of attention in spontaneous modulations of EEG (ongoing activity) which uses the complete temporal resolution of the signal. We show a mapping between the ERP results and the ongoing EEG activity based on IP. A Markov-based model is developed for removing spurious variations that can

occur in ongoing signals. We believe the proposed method can be incorporated as an important pre-processing step for a more reliable estimation of objective measures of the level of selective attention. The proposed model is used to pre-process and classify between attending and un-attending states in a seminal dichotic tone detection experiment.

In the last part of this thesis, we investigate the possibility of measuring a mapping between the neural activities of the cortical laminae with the auditory evoked potentials (AEP) in vitro. We show a strong correlation between the IP of AEPs and the neural activities at the granular layer, using mutual information.

Zusammenfassung

Die Aufmerksamkeit auf ein relevantes auditorisches Signal in einer komplexen Hörumgebung zu lenken ist von großer Bedeutung für eine erfolgreiche akustische Kommunikation. Um eine Überlastung des Hörsystems zu vermeiden, wird die Bedeutung der Reize in den höheren Ebenen des auditorischen Systems bewertet. Basierend auf diesen Informationen wird die Aufmerksamkeit von den irrelevanten und unwichtigen Reizen abgelenkt. Dabei spielt die sog. *Langzeit-Habituation*, die einen graduellen Prozess darstellt der unabhängig von der sensorischen Adaptierung ist, eine wichtige Rolle.

Ein besseres Verständnis der *aufmerksamkeits-modulierten neuronalen Aktivität* ist wichtig, um den Kodierungsprozess von sog. auditory streams zu beleuchten. Zum Beispiel können diese Informationen einen direkten Einfluss auf die Entwicklung intelligenter Hörsysteme haben bei denen genauere, *objektive* Messungen verwendet werden können, um die Hörfähigkeiten von Patienten mit Hörpathologien widerzuspiegeln. So kann beispielsweise ein objektives Maß für die Langzeit-Habituation an unterschiedliche Schallreize genutzt werden um - im Vergleich zu subjektiven Selbsteinschätzungen - eine genauere Anpassung der Hörsysteme zu erreichen.

Das Hauptziel dieser Dissertation ist die Analyse *neuronaler Dekodierungssignaturen* der Langzeit-Habituation und neuronaler Modulationen der selektiver Aufmerksamkeit durch Nutzung zirkulärer Regularitäten in elektroenzephalografischen Daten, in denen wir objektiv den Grad der Aufmerksamkeitsbindung an verschiedene Reize messen können.

Wir untersuchen insbesondere die Modulation der Momentanphase (engl. Instantaneous phase, IP) in ereigniskorrelierten Potenzialen (EKPs) in verschiedenen experimentellen Settings. Dies steht im Gegensatz zu dem traditionellen Ansatz, bei dem die interessierenden EKP-Komponenten durch Mittelung einer ausreichend großen Anzahl von Einzelantworten im Zeitbereich ermittelt werden. Es wird vermutet, dass eine hohe Aufmerksamkeitsbindung an einen Stimulus mit einem hohen Grad an IP-Clustern verbunden ist. Nimmt die Aufmerksamkeitsbindung hingegen ab, so ist die Momentanphase uniform auf dem Einheitskreis verteilt. Diese Arbeit gliedert sich in drei Teile. Im ersten Teil untersuchen wir die Dynamik der Langzeit-Habituation mit verschiedenen akustischen Reizen (leise vs. laut) in EKP-Studien. Die zugrundeliegende zeitliche Dynamik der Momentanphase und die Ebene des Phasenclusters der EKPs werden durch die Anpassung von zirkulären Wahrscheinlichkeitsfunktionen (engl. probability density function, pdf) über Datensegmente bewertet. Mithilfe eines sog. *abrupt change-point* Modells wurde die zeitliche Auflösung der Daten erhöht, sodass signifikante Änderungen in der Momentanphase bei verschiedenen Reintonstimulationen detektierbar sind.

In einer zweiten Studie verbessern wir die Ergebnisse und die Methodik, indem wir einige der Einschränkungen lockern, um den *gradualen Prozess* der Langzeit-Habituation in das *abrupt change-point* Modell zu integrieren. Dazu wird ein bayes'sches Zustands-Raum-Modell vorgeschlagen. In

den zuvor genannten Studien konnte erfolgreich mithilfe der Momentanphase zwischen verschiedenen Stimulationspegeln unterschieden werden. Im zweiten Teil der Arbeit wird der experimentelle Rahmen erweitert, um komplexere auditorische Reize wie in realen Hörsituationen untersuchen zu können. Dabei analysieren wir die neuronalen Korrelate der Aufmerksamkeit anhand spontaner Modulationen der kontinuierlichen EEG-Aktivität, die eine zeitliche Auflösung ermöglicht. Wir zeigen eine Abbildung zwischen den EKP-Ergebnissen und der kontinuierlichen EEG-Aktivität auf Basis der Momentanphase. Ein Markov-basiertes Modell wird entwickelt, um störende Variationen zu entfernen, die in kontinuierlichen EEG-Signalen auftreten können. Wir glauben, dass die vorgeschlagene Methode als wichtiger Vorverarbeitungsschritt zur soliden objektiven Abschätzung des Aufmerksamkeitsgrades mithilfe von EEG-Daten verwendet werden kann. In einem dichotischen Tonerkennungsexperiment wird das vorgeschlagene Modell zur Vorverarbeitung der EEG-Daten und zur Klassifizierung zwischen gerichteten und ungerichteten Aufmerksamkeitszuständen erfolgreich verwendet.

Im letzten Teil dieser Arbeit untersuchen wir den Zusammenhang zwischen den neuronalen Aktivitäten der kortikalen Laminae und auditorisch evozierten Potentialen (AEP) *in vitro* im Tiermodell. Wir zeigen eine starke Korrelation zwischen der Momentanphase der AEPs und den neuronalen Aktivitäten in der Granularschicht unter Verwendung der Transinformation.

List of Abbreviations

A1 Primary auditory cortex.

ACP Abrupt change point.

AEP Auditory evoked response.

ALR Auditory late(evoked) response.

ANOVA Analysis of variance.

CA1 Cornu ammonis 1.

CA3 Cornus ammonis 3.

CI Confidence interval.

CWT Continuous wavelet transform.

dB Decibel.

EEG Electroencephalography.

ER Evoked response.

ERP Event related potential.

FD Fascia dentata.

FN False negative.

FP False positive.

F_s Sampling Frequency.

GC Granger causality.

HL Hearing level.

HT Hilbert transform.

IE Instantaneous envelope.

IF Instantaneous frequency.

IP Instantaneous phase.

ISI Inter stimulus Interval.

KS Kalman smoother.

LFP Local Field Potential.

MGB Medial geniculate body.

MI Mutual information.

mm Millimeter.

MSE Mean square error.

NLM Non-local means.

pdf Probability density function.

RN Raphe Nuclei.

s Seconds.

SNR Signal to noise ratio.

SPL Sound pressure level.

TN Thalamic Nuclei.

TRN Thalamic Reticular Nucleus.

WT Wavelet transform.

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Preliminary: The Hearing Brain	4
1.2.1	The Physical Properties of Sound	4
1.2.2	From Ear to The Auditory Cortex	6
1.3	Auditory Scene Analysis	8
1.3.1	Auditory Attention	10
1.3.2	Theories of Selective Attention	11
1.3.3	Habituation	14
1.4	Decoding of Neural Signatures of The Long-term Habituation Process	17
1.4.1	Mapping Between Ongoing EEG and ERPs	18
1.4.2	The classical view	19
1.4.3	Pure Phase-resetting	19
1.4.4	Pure-phase Resetting and Amplitude Enhancement	20
1.4.5	Organization of the Thesis	20
2	Materials & Methods	22
2.1	Synchronized Neural Activity and Attention	22
2.2	Extraction of Instantaneous Phase (IP)	23
2.2.1	Filtering procedure	23
2.2.2	Hilbert Transformation	24
2.2.3	Wavelet Transformation	25
2.3	Statistical Analysis and Decoding of Neural Responses of Attention-Binding	28

2.3.1	Descriptive Circular Statistics	28
2.3.2	Probabilistic Modeling of Instantaneous Phase	30
2.3.3	Hidden Markov Models & Change Detectors During the Course of Long-term Habituation	38
2.3.3.1	Tracking changes in the IP of long-term habituation processes	40
2.3.3.2	Discrete Forward-Backward Bayesian Change Point Model	40
2.3.3.3	Experimental Setting: Data Acquisition	45
2.3.4	Continuous Forward-Backward Change Point Model	45
2.3.4.1	State Transition Model, $p(\mu_t, \kappa_t \mu_{t-1}, \kappa_{t-1})$	48
2.3.4.2	Experimental Setting	48
2.4	Oscillatory EEG Signals and IP Extraction	49
2.4.1	Phase Singularities: Definition of Spurious Phase Slips and Types	50
2.4.2	Modeling the Variations Between Instantaneous Phase and Amplitude	53
2.4.2.1	Model Derivation	54
2.4.3	Neural correlates of Selective Attention & Ongoing EEG	58
2.4.4	Experimental Setting	59
2.5	Laminar Auditory Activities and its Relation to Auditory Evoked Responses	60
2.5.1	The Cortical Laminar Organization in A1	62
2.5.2	Experimental Setting	63
2.5.3	Mutual Information of IP between Laminar Cortex	63
3	Results	66
3.1	Long-term Habituation Decoding Using Von-Mises Model	66
3.2	Abrupt and Continuous Detection of Temporal Changes in ERPs in Long-term Ha- bituation Process	70
3.2.1	Synthetic Data Evaluation	71
3.2.2	Evaluation of CPs on the Measurements	72
3.3	Detecting the Gradual changes of long-term habituation using a State-Space Model	75
3.3.1	Initial Conditions	76
3.3.2	Optimization of Prior Parameters	76
3.3.3	Model Validation on Artificial Circular Data	79
3.3.4	Determining the Model Parameters for Experimental Data	82
3.3.5	Tracking Data Distribution over Trials	82
3.3.6	Average Concentration For Different Auditory Stimulations	85
3.4	Removal of Spurious Phase Variations in Ongoing EEG Signals	86

3.4.1	Validation of Model on Synthetic Data	86
3.5	Analysis of IP of Ongoing EEG Measurements	90
3.5.1	Filtering procedure & IP Extraction	90
3.5.2	Setting of KS factors	91
3.5.3	Quantification of Attentional Effort In Experimental Setting	92
3.5.4	Optimization of Filter and KS Parameters	92
3.6	Relating AER to Laminar Phase Dynamics	95
3.6.1	Data Processing	95
3.6.2	Interaction Between Cortical Layers and AERs Using MI	96
4	Discussions & Limitations	99
4.1	Tracking Changes in Long-term Habituation Over Trials	99
4.1.1	Limitations:	101
4.2	From the Abrupt CPM to Continuous Detection of Changes	103
4.2.1	Limitation With Abrupt Definition of Changes	105
4.3	Continuous Definition of Long-term Habituation and Tracking The Changes	106
4.3.1	The Rate of Change of the concentration Parameter	107
4.3.2	Limitations	108
4.4	Tracking the level of selective attention in ongoing EEG	109
4.4.1	Setting of the α & σ Parameters	109
4.4.2	Setting of the β Parameter	111
4.4.3	Different Choices of β and Λ	111
4.4.4	Limitations	114
4.5	The Mappings Between Laminae Activities and AERs	115
5	Conclusions	117
5.1	Conclusion	117
	Bibliography	120
	Appendices	134
A	Cauchy Principal value	134
B	Fourier Transform of Complex Wavelet Function	134
C	Non-local Means Method	135
D	Derivation of Forward Pass of Abrupt CPM	136
E	Derivation of the Posterior Distribution Over the States in CPM	137

F	Derivation of Weights	137
G	Derivation of Probability Density Functions	138
List of Publications		142
Danksagung		144

Chapter 1

Introduction

1.1 Motivation

Hearing is one of our major senses used for communication and perceptual processing. We can experience fear, alertness and pleasure through *sounds* which are generated by the changes in the air pressure via vibrations of objects¹. However, the only role of the auditory system is not to detect such external changes, but rather to construct an *internal model* that helps us to have an *experience* of what we are able to hear. Our perception and experience of sound depends on processing and interpreting an internal model. For example, at every point in time, many different sound signals arrive simultaneously to the auditory system. The sound signals are broken down into a different set of features and passed to the higher cortical levels for further processing. The features are then combined and processed to achieve a perceptual understanding of the reconstructed sounds. Detection and localization of sound sources, identification of familiar sounds between different stimuli and noises, filling out the missing gaps in familiar songs or an auditory text (see Kraemer et al. (2005)) are all a small part of the examples of the higher level processing of sounds and the experiences we accomplish through listening.

Therefore, hearing is not merely a passive physiological process (bottom-up process) to perceive loudnesses and pitches, but a series of high-level cognitive processes (top-down process) that require cognitive resources and allocation of mental effort (Goldstein and Brockmole, 2016). One of the crucial aspects of auditory processing is dealing with an overwhelming presence of auditory stimuli

¹The human auditory system is capable of detecting a broad range of air pressure changes, ranging from 0.00002 to over 100 Pascals. To understand this scale, we define **atmospheric pressure** which is defined as the environmental pressure in the absence of any sound. According to the SI (Système International) unit, it is called Pascal. one pascal (1 Pa) is equivalent to a force of one Newton acting on a surface of one square meter. The standard atmosphere (standard reference) is defined as 101325Pa. The *sound pressure* is defined as the difference between the instantaneous sound pressure due to sound and the standard atmosphere. The difference is measured in Pascal as well.

at every second, generated from a multitude of sources, different delay and spatio-temporal characteristics. To obtain a meaningful experience of all these sounds at different circumstances and avoid overloading of the system – a phenomenon known as the *cocktail party problem* (Bregman, 1994, Chapter 6), an efficient filtering process is required. The cocktail party problem was first reported by Colin Cherry in 1953 in a series of experiments for understanding how can we select the voice of a particular person in presence of many other voices (Cherry, 1953, 1954; Bregman, 1994). This is necessary as our processing capacity is limited. One of the key methods through which auditory scenes are parsed is *attention*, by which irrelevant stimuli are muted out and only salient stimuli are considered for further allocation of cognitive and sensory resources (Bregman, 1994; Eysenck and Keane, 2000; James, 1890). The saliency of a stimulus may depend on different factors such as its loudness, its behavioral relevance (e.g., one’s own name) as well as a particular associative experience (Conway et al., 2001; Wood and Cowan, 1995).

In the same context of attention, another way of filtering out irrelevant information is through *long-term habituation*, a simple learning mechanisms by which the attention is shifted away from the irrelevant stimuli and allocated to the relevant or salient stimulus (Rankin et al., 2009; Domjan, 2014; Thompson, 2009). Dysfunctions in the filtering mechanisms can have significant implications. The effect of long-term habituation has been investigated in patients with auditory pathologies such as tinnitus. One of the main treatment methodologies for tinnitus patients is to train them to habituate to the phantom noise (loosely speaking, to *ignore* the stimulus). Decompensated patients who are not able to habituate to the phantom noise could suffer severely from the sounds caused by the interferences in the brain’s hearing system. In some previous studies, the level of habituation in different groups of patients has been analyzed by studying the level of habituation in the auditory event related potentials (ERPs) (Walpurger et al., 2003; Carlsson and Erlandsson, 1991).

One of the domains in which the effect of long-term habituation is important, is the adjustment of most comfortable loudness level (MCL) in hearing aid devices and cochlear implantation, particularly in uncooperative patients such as infants. The setting of the loudness levels should illustrate the patient’s comfort in hearing sounds for a long-term in such a way that the filtering process of the irrelevant stimuli are maximized and attention is properly amplified towards the desired target stimulus. We use cerebral recordings, in particular auditory evoked potentials, an established diagnostic tool in audiology for the study of selective attention and the habituation process in the long-term. The selective attention effect has been observed as an endogenous modulation of the N1 component. In Rao et al. (2010); Hillyard et al. (1998, 1973) the amplitude of N1/P2 components were enhanced in the case of attended in comparison to unattended stimuli. In many studies the effect of neural-correlates of attention is studied as an average of amplitudes of sufficiently large numbers of ERPs. However, in the studies conducted in Strauss et al. (2008b); Trenado et al. (2009); Low and Strauss (2011), it has been shown that the correlate of attention states in the averaged ERP originates partly from a jitter in the IP (also known as phase-resets) of subsequent single-trial ERPs

in auditory paradigms. In section 1.4.1, we describe in more details about the genesis of ERPs and its relation to the amplitude and IP of single trial ERPs.

Much less examined in literature, is the dynamics of attentional binding and long-term habituation, e.g., the time course of exogenous and endogenous selective attention during ERP paradigms, even though these dynamics are of major importance in clinical neurodiagnostic procedures as well as in experimental research (Walpurger et al., 2003; Strauss et al., 2008b; Haab et al., 2011; Mariam et al., 2009; Rauschecker et al., 2010). In Walpurger et al. (2003), the long-term habituation is studied in the time domain, by which large-scale partial averages of trials are used. In Strauss et al. (2013) a 2-D denoising scheme to extract habituation correlates in individual ERP trials is used. Time-scale coherence measures which evaluate segmented ERP sequences were used in Mariam et al. (2009) as well as hybrid time-scale machine learning methods as in Mariam et al. (2012). Fuentemilla et al. (2006) discuss the influence of phase-reset and the increase of spectral power on the generation of ERP from ongoing scalp EEG activity. They could already report changes of phase coherence over trials when monitoring the N1 component of consecutive ERPs. All the cited methods use the amplitude information exclusively or combined amplitude-phase measures and thus do not allow for an isolated analysis of the aforementioned IP related to selective attention. In addition, the temporal modulations of IP over trials has not been fully investigated.

Given the current state of the art, the major incentive of this thesis is to address the neural-correlates of long-term habituation and attentional-binding using the circular regularities in IP over single-trial ERPs as well ongoing EEG activities. It is important to highlight the difference between the *subsequent-single trials* of ERPs and the averaging over ERP trials for studying the relevant signal components. In the averaging process, despite increasing the signal-to-noise ration, details regarding the dynamic changes in the N1-P2 component as well temporal precision are lost. By modeling the dynamics of the isolated IP over single trials in different experimental settings, we aim to understand the temporal characteristics of the long-term habituation and the corresponding level of attentional binding to different stimulation. The challenge requires a comprehensive understanding of neural coding and decoding of sensory information, as well as appropriate computational merit for the study of single-trial responses.

The contributions of this thesis can be divided as follows:

1. Analyzing the effect of long-term habituation in measured ERPs over single trials in an invasive and non-invasive experimental setting. The phase modulations of N1 component of ERPs are assessed initially at two distinct stimulation levels of 50dB SPL (soft) and 100dB SPL (aversive, loud). We assess the neural correlates of attention-binding by fitting a circular distribution over different data windows and analyze the changes in the cluster level of the data distributions (Mortezapouraghdam et al., 2015a). The results in Mortezapouraghdam et al. (2015a) are cross examined with the simulations obtained from a well founded quantitative neurofunctional model that covers several spatio-temporal scales of neural processing to generate ERP sequences.

2. To obtain a better resolution in detecting the changes in the habituation process than the method proposed in (1), a change point model is used to detect the underlying significant changes in the neural-correlates of attention. The method compensates for some of the shortcomings in (1), however does not uniquely integrate the *gradual process* of habituation in the model. We therefore develop for the first time, a Bayesian state-space model, that is able to incorporate the gradual changes of the habituation process in the model. The ability of the model to reflect the continuous changes in the long-term habituation is evaluated on a series of pure tone stimuli, ranging from 60dB SPL to 90dB SPL (Mortezapouraghdam et al., 2014, 2016).
3. We analyze the effect of selective attention in a more complex experimental setting than proposed in (1) and (2). The measurements of this study are based on a repetition of a seminal dichotic tone detection presented by Hillyard et al. (1973). Due to the longer stimuli, we are required to analyze the ongoing EEG signals. This makes the study more challenging as noise factors can distort the data, and as no averaging is done, the results are more prone to noise than the other methods. Therefore, we developed a variant of a Kalman smoother for removing spurious phase variations which are not binded to any neurophysiological effect (Mortezapouraghdam and Strauss, 2017). Using the proposed method, we show a mapping between the segmented ERPs and ongoing EEG data suggesting a unified framework to analyze neural correlates of selective attention in ERPs and their ongoing EEG activity.
4. In addition to the studies in (1-3), we investigate the interaction of thalamo-cortical activities and its relations to auditory late responses (ALRs) in-vitro. As invasive experiments are costly and more hazardous in humans, we examine the projection of cortical activities of the auditory cortex onto auditory late response (ALR) measurements (Mortezapouraghdam et al., 2015b) in rats at different frequencies. Of interest is how the in-depth information are correlated to the ALRs and how effectively they can be measured. We employ only the phase information of the measured neural activities.

1.2 Preliminary: The Hearing Brain

In this section we will describe some of the main properties of the auditory system and the translation procedure of sounds to neural-activities in the auditory cortex.

1.2.1 The Physical Properties of Sound

To understand the psychoacoustics of normal hearing, we briefly explain the main attributes of the sounds and its relation to hearing. Sounds are described as the displacement of air (or other medium) molecules which are known as condensation (an increase in the density of air molecules)

and rarefaction (a decrease in the density of air molecules). The resulting oscillations are propagated away from the sound source much like water ripples (with the difference that sound waves have a longitudinal propagation than transversal propagation) (Gelfand, 2001). By repeating this process, a sinusoidal pattern of air molecules is formed, also referred to as **pure tone**.

Loudness is defined as the *perceived* intensity of a sound and it depends both on the amplitude² and the frequency. The amplitude of the sound is closely related to the level or the loudness of a sound and the the frequency of a pure tone is closely related to what is perceived as the **pitch** of the sound. The frequency range that the human auditory system can respond to is extended from 20Hz to as high as 20kHz, however the range diminishes to 15-17 kHz for matured adults, as they slightly lose sensitivity to some frequencies. The relation between the frequency and the perceived loudness is not linearly correlated. The sounds of different SPL can be perceived to have the same loudness at different frequencies. The minimum level of sound (dB SPL) at the particular frequencies that make the sounds audible (the most quietest sound that a normal hearing subject could detect) are referred to as threshold values (known as audibility curve or the threshold of hearing). Each of these thresholds (dB SPL) at the particular frequency refer to a 0 dB hearing level (0 dB HL). The HL scale is used commonly by clinicians to measure the level of hearing loss. For more details see Gelfand 2001, Chapter 4; Goldstein and Brockmole, 2016, Chapter 11.

In perception, it is important to make a distinction between the **subjective loudness** of a sound and its physical properties. Loudness and pitch correspond to how we perceive the physical properties, namely the intensity and frequency of a sound (Gelfand, 2001). The relationship between the physical (i.e., intensity and frequency) and psychological (i.e., loudness and pitch) features of sound in perception are not always directly balanced. For example, a low intensity sound can be perceived loud in the case of patients with hyperacusis (more details and explanations are given in section 1.3). In all our experimental settings, we use subjects with a normal hearing so that they have a comparable HL.

Pitch is also a psychological property of sound and though it physically cannot be measured, it is very closely related to the fundamental frequency of a sound³. We describe the pitch as low (in case of low frequencies) or high (for high frequencies) based on the perception of sound. According to Bendor and Wang (2005) and Goldstein and Brockmole (2016), it is best described as a *sensation that enables us to order the sound from low to high*. The psychological property of pitch is shown by Stevens (1935) where it has been shown that the pitch of a low frequency tone is perceived to get lower when it is made louder, and the pitch of a high frequency tone is perceived to get higher when it is made louder.

²To handle the vast set of sound levels which we are able to hear, the unit *sound pressure level (SPL)* is used. SPL is a logarithmic measure of the effective pressure of a sound stimulus relative to a reference pressure and is described in decibels (dB). The dB is computed as $L_p = 20 \log_{10}(p/p_0)$, where p is the pressure of sound and p_0 is the reference pressure level, commonly set as 20 micropascals (which is near the sound pressure of the hearing threshold at 1kHz). The amplitude is therefore expressed as dB SPL.

³The tone's pitch remains the same if the fundamental frequency or the higher harmonics of a tone are removed.

1.2.2 From Ear to The Auditory Cortex

Most sounds around us are transmitted through airborne vibrations. The transmission of sound through two different mediums, particularly with a different impedance⁴ is rather poor (Alberti, 2001). In case of humans, the airborne vibrations should be transmitted to the brain which is surrounded by fluids with different properties. Therefore, the ear has evolved a set of mechanisms to deal with the problem of impedance mis-match, which is the transfer of sound energy from the outer ear to the inner ear with a minimum amount of energy loss. This process is called the *sound-conduction* mechanism (Alberti, 2001). In the following we briefly describe the transformation process of sounds to nerve impulses and up to the auditory cortex for processing sounds.

The transformation journey of sounds into nerve impulses begin from the outer ear. The outer ear contains the pinna (or earlobes), the auditory canal (external auditory meatus) for transmitting sound to the tympanic membrane. The shape of the pinna and structure of the auditory meatus are such that the sounds around 3kHz are more amplified, which explains the reason that humans are sensitive to a frequency range of 2-5 kHz, encompassing the energy range required to distinguish different phonemes in speech (Purves et al., 2015). The unique shape of pinnae have been shown to play an important role in sound localization as well (Batteau, 1967). In Wenzel et al. (1993), it has been shown how the shape of one's ears relative to his own head can be used to infer the location of sounds⁵.

In the second step, the low impedance airborne vibrations are converted into liquid-borne (higher impedance) vibrations. The conversion process is done through the middle ear and minimizing the amount of energy loss. This is achieved primarily by mechanically boosting the airborne pressure (nearly by 200 fold). The three bone **ossicles** (i.e., malleus, incus, and stapes) transfer the mechanical pressure on the eardrum to smaller diameter bone window (**oval window**) whose membrane is part of the fluid filled **cochlea**. The oval window, is the site at which the ossicles contact the inner ear.

In the third step, the liquid-borne vibrations are converted into neural impulse responses through the cochlea, a coiled structure of length of approximately 35mm in its uncoiled form (Purves et al., 2015). The cochlea not only amplifies the sound waves and translates the sensory input into electrical signals, but it also acts as a frequency analyzer, in which the complex sensory input is decomposed and represented in terms of simpler elements. It contains three fluid filled chambers divided by two thin membranes (the basilar membrane and the tectorial membrane, see Purves et al. (2015) for more details). The chambers are the Scala vestibuli, Scala media, and Scala tympani. The middle tube which is the scala media contains the **organ of Corti**, which sits on the **basilar membrane**.

⁴Impedance is defined as the product of medium density and the speed of sound through it.

⁵This is called the *head-related transfer function* (HRTF). Other factors such as the angle of the head in relation to the sound source is important for sound localization. In Griffiths and Warren (2002) it has been shown that *planum temporale* area located in the posterior site of the primary auditory cortex is responsible for gathering the information from the ears (where the acoustic stimuli are broken into different spectrotemporal pieces and sound patterns distorted by the shape of ears) and the input with respect to the HRTF, for integrating and processing the spatial properties of the sound.

The functionality of the basilar membrane was revealed in 1960 by a series of studies conducted by Georg von Békésy, who won the Noble Prize in medicine and physiology in 1961. He measured the strength of the neural responses at different positions of the basilar membrane that has been stimulated by a sound source. The observations indicated strongly that the basilar membrane doesn't respond uniformly, and despite an overall vibration of the membrane, different positions of the membrane respond more strongly to a given frequency. He stated that the motion of a basilar membrane becomes similar to a *traveling wave* along the path given the tuning of different positions of the membrane. The high frequency components stimulate the base of the cochlea more whereas the lower frequencies stimulate the apex more (Békésy, 1961; Purves et al., 2015). This is also attributed to the structure of the basilar membrane, which is narrower and stiffer at the basal end and gets more flexible and wider at the apical end. The tonotopic structure of the cochlea has been extensively used in applications such as cochlear implants⁶ (see Chapter 12 - (Purves et al., 2015)). On top of the basilar membrane sits the auditory hair cells. There are two types of hair cells, the inner and the outer hair cells, for which most of the afferent dendrites are on the inner hair cells and most efferent on the outer cells. The afferent pathways will lead the neural impulses from the cochlea to the brain, and the efferent pathways bring the impulses from the brain to the cochlea. The hair cells respond to atomic movements with a temporal precision of microseconds. The motion generated from the traveling wave in the basilar membrane will stimulate the hair cells accordingly and ultimately lead to the generation of electrical signals that will be transmitted to the cochlear nuclei in the brainstem via the auditory nerve fibers. The projections of the signals will extend up to the **primary auditory cortex** through several hierarchical levels including the medial geniculate nucleus (describing the networks and layers which take part in the transmission of neural impulses are beyond the scope of this research. Comprehensive details can be found in Purves et al. (2015)). The primary auditory cortex (referred as the A1 region) is the main cortical area that receives the auditory-based thalamic input. It is located in Heschl's gyrus in the temporal lobe and is surrounded by the secondary auditory cortical areas (belt and parabelt regions (Purves et al., 2015)). These peripheral regions also receive input from the medial geniculate nucleus. The auditory nerve fibers carrying the signal are extended up to different parts of the auditory cortex. Therefore, a lesion to the core region will not lead to absolute deafness (Musiek et al., 2007). Similar to the tonotopic organization of the basilar membrane, different neurons within the auditory nerve respond more strongly to certain sound frequencies than others. The same structure has also been observed up to the primary auditory cortex, where the central region responds more strongly to the lower frequencies and the outer regions to the higher frequencies (Kiang, 1965; Ward, 2015).

⁶With the expansion of new technologies, researchers were able to measure the neural responses on an *alive cochlea* which lead to a better understanding of the basilar membrane mechanism (Goldstein and Brockmole, 2016; Narayan et al., 1998; Khanna and Leonard, 1982). The new measurements suggested that the vibrations at different positions of the membrane occur *sharper* and *more localized* in comparison to the tonotopic model of the von Békésy (Purves et al., 2015; Goldstein and Brockmole, 2016). The sharper responses of the new observations were explained by the role of outer hair cells. The expansion and contraction of outer hair cells results in amplifying and sharpening the vibrations on the basilar membrane.

We can measure the neural activities within the auditory system (auditory regions of the brain) by amplifying the changes in the voltage measurements of the electrode located in proximity of the auditory cortex. As EEG recordings represent a mixture of vast number of different neural activities from different regions of the brain, it is difficult to have an accurate spatial resolution of the actual neuro-cognitive processes (Berger, 1929). However, we can measure the neural responses in regard to a specific event or stimuli which is embedded in the recorded EEG activities (Luck, 2014). The measured neural activity within the auditory system (auditory regions of the brain) in relation to a specific auditory stimulus is referred to as late-auditory-*event-related potential (ERP)* (also known as late-auditory evoked response (AER))⁷. They are known to reflect the summation of postsynaptic potentials which is generated when a large number of similarly cortical pyramidal neurons fire in synchrony while processing information (Sur and Sinha, 2009). Despite some of the main drawbacks of ERPs such as lack of spatial resolution, they are shown to strongly correlate with interpreting different cognitive process (Hall, 1992; Woodman, 2010). The peaks and troughs of stimulus-locked ERPs are shown to be useful in understanding different cognitive processes as they develop (Woodman, 2010; Hall, 1992; Luck, 2014). The peaks and troughs, known commonly as the *ERP-components* are defined in terms of their voltage polarity of their responses (positive or negative), latency, order of occurrence, and sensitivity to task manipulations (Sur and Sinha, 2009). In Luck (2014); Hall (1992) the list of different ERP components for visual and auditory sensory stimulation have been fully described. The P1 (P100, positive voltage, 50-80ms after stimulus onset), the N1 (N100, negative voltage, 100-150ms after stimulus onset) and the P2 (P200, positive voltage, 150-200ms after stimulus onset) components of an ERP have been shown in Fig. 1.1. For more information regarding the polarity and latency as well as later components, see Hall (1992). In section 1.4 we will explain the importance of N1-P2 complex in regard to study of long-term habituation.

1.3 Auditory Scene Analysis

In the previous section we explained the most important stages of sound translation to neural impulses in the case of a pure tone. In reality, we usually do not encounter pure tones in our environments, but numerous complex sound stimuli which arrive simultaneously or with different delay times to the ears. Despite the large number of auditory stimuli that we may obtain at every point of time, we are able to filter and process them appropriately. For example, the sounds that enter the cortex have different acoustical characteristics such as frequency, amplitude and phase and they may be generated from different sound sources. Despite all the complexities of the environment, we are able to hear the auditory inputs as *separate entities*. The sound of a fan, typing on a keyboard and a colleague speaking on the phone may all be heard at the same time, but we are still able to

⁷In general, ERPs are the measured voltages (measured in microvolts) generated in different brain structures in regard to a specific event or stimuli.

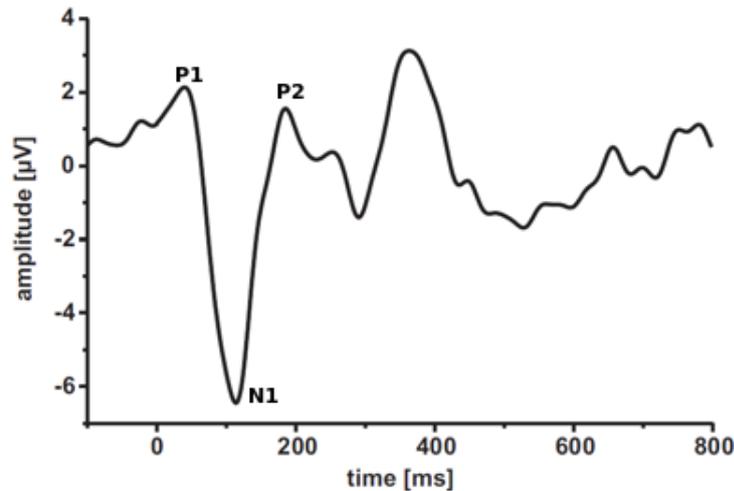


Figure 1.1: An illustration of an averaged auditory late response. The wave components are named after their order of appearance and polarity. The P1 component is expected 50-80ms post-stimulus and has a positive polarity. The N1 component is expected to appear 80-120 ms after the stimulus onset (and a negative (N) polarity) followed by the P2 component (positive polarity and expected ~200ms post-stimulus) (Hall, 1992).

distinguish and localize them. *The process of making sense of the superposition of sound sources arriving to the ears and to separate them into individual sources is known as auditory scene analysis* (Bregman, 1994). To correctly group the features and bits of sound elements that belong to a source and segregate the different sources from one another, are important and challenging tasks for having the right perception of sound.

A series of *Gestalt principles* which are extensively used in the area of vision apply also in the area of hearing. If distinct and perceptible elements in an environment satisfy a set of conditions, they are grouped together in order to build a coherent perceptual organization (Bregman, 1994). The Gestalt principles consist of a set of rules mainly based on the concepts of *proximity*, *similarity*, *continuity*, and *closure* of features and elements for the means of grouping (Bregman, 1994). As an example, if several frequencies of a sound source start and possibly end at the same time, they are expected to belong to the same auditory source. They can possibly grow and decline together and hence those bits of information and elements of the input stimuli are grouped together to represent one specific source. Two sounds which arrive at two different times are possibly from two different sources and based on localization mechanisms, they can be separated from each other and grouped into two different sources. Sounds which remain constant or change smoothly over time resemble the principle of continuation of the Gestalt psychology and hence are considered most probable to arise from the same source (Bregman, 1994). The gestalt principles can be considered as a *heuristic approach* which can reasonably well describe the framework for grouping elements for perception. A list of known techniques that the hearing system employs for separation of sources are based on *onset time*, *pitch*, *auditory continuity* and *experience*⁸(Bregman, 1994).

⁸Auditory localization is one of the main aspects that helps the hearing system for source recognition. Given two sounds with the same frequency and loudness but generated at different locations, we can easily distinguish

1.3.1 Auditory Attention

In section 1.3 it was discussed that sounds may arrive as a mixture of numerous auditory sources to the ears. The auditory system has the task of separating different sources from each other based on grouping the elements and features that belong together. This is followed by the perceptual interpretations of different streams. In the same context, it is evident that when numerous sensory inputs (in the case of the auditory system, we refer to inputs as auditory streams) enter the auditory system, it will be inefficient to *select* all the sensory inputs for higher levels of processing. Many of the auditory streams may not contain relevant information at a time, others may contain more crucial information and hence a better mechanism is required to avoid the overload of the perceptual system given the limited capacity (Kahneman, 1973). The mechanism through which *irrelevant sensory streams are filtered* and more relevant sensory inputs are selected for higher levels of processing is referred to as **selective attention**. Attention was described by the personal experiences of William James, a professor of psychology at Harvard (James, 1890) as follows:

“Millions of items . . . are present to my senses which never properly enter my experience. Why? Because they have no interest for me. My experience is what I agree to attend to. . . . Everyone knows what attention is. It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration, of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others.”

William James, 1890, The Principles of Psychology

Despite a significant amount of research focused on the role of selective attention for vision, it is now believed that the same set of principals can be applied for auditory modality as well (Mayer et al., 2006; Shinn-Cunningham, 2008). Attention can be stimulus oriented (i.e., the physical properties of the stimulus is considered a major drive for attention) which is referred to as *exogenous attention* or more voluntarily by the goal or intention of the person to allocate attention to a specific target stimuli, which is referred to as *endogenous attention*. The passage of an ambulance with a loud siren in the vicinity of a person will capture his/her attention swiftly. The allocation of his/her attention towards the siren is a direct result of the stimulus's unique loudness and salient properties. The person had no voluntary act in selection of the siren sound input for higher cognitive processing. The endogenous attention however can be attained by a *volunteer* attempt/effort to a specific source. This volunteer act regularly happens as we decide to attend to one person speaking to us in a crowded environment. As a consequence of attending to one stimulus, the background noises are

the difference between their positions in the space and hence consider them as two separate entities. The similar sounds indicate that they will trigger the same auditory nerve fibers and the same hair cells in the cochlea and hence the generated patterns are the same. The auditory system should therefore should use other mechanisms for sound localization. The information used from both ears for sound localization are referred to as *binaural cues* and the information used from one single ear are named as *monaural cues*. Using multitude number of cues, the auditory system is able to locate the sounds in space and thereby distinguish them despite their similar tones. For more information see Goldstein and Brockmole (2016)

greatly filtered and the processing of the target stimulus is enhanced. Unlike exogenous effect of attention, the maintenance of the endogenous processes require effort, attention, and are limited in capacity (Kahneman, 1973; Strauss and Francis, 2017). There are a different set of assumptions about the extend to which unattended and filtered messages/stimuli are processed. In the following sections we will highlight some of the main models presented for describing the process of selective attention along with their limitations.

1.3.2 Theories of Selective Attention

In a study conducted by Simons and Chabris (1999), participants had to focus on a specific aspect of a game run by several players (i.e., counting the number of times a ball would pass between the players) which resulted in surprising findings: many subjects with a high probability (50%) would miss to observe the presence of a Gorilla costumed person crossing the playing field. Such an effect is referred to as *inattentional blindness* defined as the failure to observe a visual stimulus because of the diversion of attention to another high demanding task (Rensink, 2009). The experiment was similar to the earlier dichotic⁹ listening task in which people fail to recall the contents presented to one ear while actively focusing their attention to a target stimuli presented to the other ear (Moray, 1959). The basic principles of such filter mechanisms for the auditory system have first been suggested by Broadbent (Broadbent, 1958) as he tried to explain the switch of attention between different messages that would constantly arrive to both ears at an air traffic control station.

The first dichotic listening experiment was designed by Colin Cherry (Cherry, 1953, 1954). In his experiment, he presented auditory words to both ears simultaneously via headphones and participants were instructed to repeat messages presented to one of the channels. The subjects were successful in reporting the shadowed messages specifically if they had to shadow specific physical features such as voice (i.e., shadowing a female instead of a man's voice) or a specific sensory channel (i.e., shadowing messages coming to the right ear). However, the participants had a little amount of knowledge regarding the contents or semantic changes in the unattended messages (See Spence and Santangelo (2010) and references therein). These observations lead to the theory of early selective attention described by Broadbent. He stated that selective attention operates merely as a *filter*. According to this model, the sensory information that arrives simultaneously to both channels is registered in a sensory buffer. The filter that selects one channel for further processing will pick one of the channels based on its *physical properties* such as to what ear the information was coming first, the intensity of incoming streams, time of arrival, etc. The filter will select one of the channels based on its priority or relevance and pass the message to the higher levels for *semantic processing*. According to this model, none of the other messages are semantically processed. The remaining information and messages which are unattended, will be shortly kept in the buffer and rapidly decay. Broadbent

⁹In a dichotic experimental paradigm the subjects are presented with separate messages to both ears simultaneously. They are instructed to repeat one of the messages as soon as possible and ignore the other. Therefore, the messages reported to from one ear is namely the *shadowed ear* and the other side is *unattended ear*. This form of the experiment is also called *shadowing*.

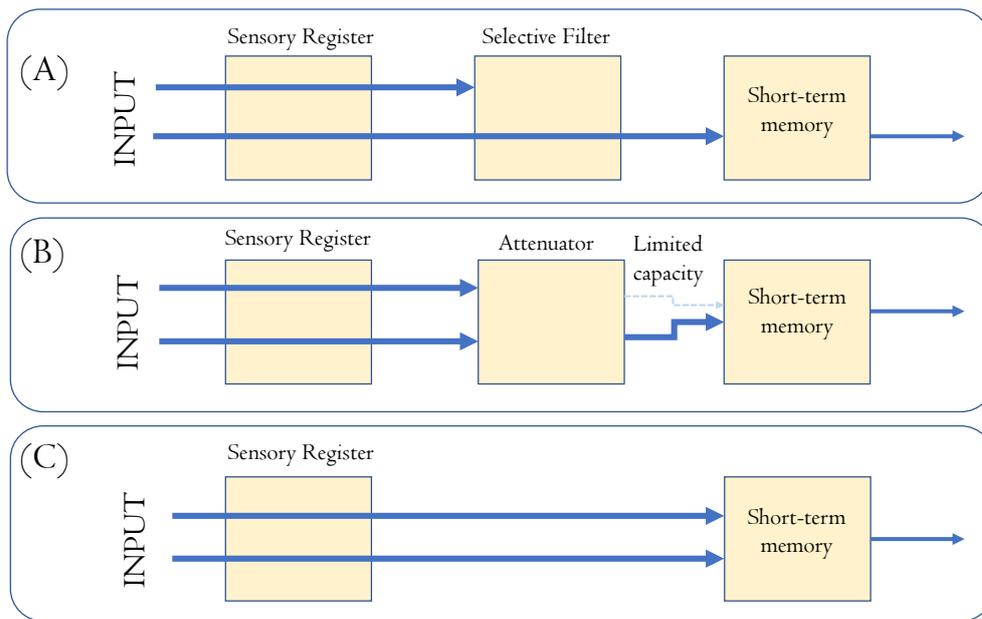


Figure 1.2: (A) Schematic representation of Broadbent's filter model. (B) Treisman's attenuation model of selective attention. (C) Deutsch and Deutsch's theory model. The plot is adapted from Eysenck, M. W., & Keane, M. T. (2000). *Cognitive psychology: A student's handbook*. Taylor & Francis. Pp. 155.

believed that the following scheme explains the mechanism that avoids the overload of the processing units due to the high amount of incoming sensory streams (stimulations). This was one of the earliest models on the theory of selective attention. Fig. 1.2 (A) shows a schematic presentation of the model.

The model of Brodbent was later criticized for its lack of flexibility in explaining the variability in the selection of unattended messages which was later observed. Some critics (Moray, 1959; Underwood, 1974) of the Broadbent's model correctly argued that one of the factors, that the participants performed poorly on, in the shadowing experiment was due to their lack of experience. Moray (1959) showed that by using experienced participants in the shadowing experiment, the detection of messages is significantly increased. In addition he reports that about one-third of the participants were able to report the presence of their own name in the unattended messages. In more developments, in an experiment conducted by von Wright et al. (1973), the appearance of a word previously associated with an electric shock to the unattended ear showed an increase in the galvanic skin response (GSR) for some participants. The latter findings are in contrast with the explanation of Broadbent's filter model. Based on Broadbent's model, the messages of the unattended channel should be blocked at early stages from any further processing after the filtering stage. The increase in GSR suggests that the processing could occur at lower levels, however is not blocked entirely.

In a modification to Broadbent's selective attention model, Moray discerned that selective attention filters most of the irrelevant information, however messages with important or powerful content can *break through* the filter and be passed for higher level processing.

We briefly outline some of the developments in describing the selective attention which were later presented.

- In another model presented by Anne Treisman, the sensory information with less relevance also enters the perceptual processing, however, the extent, to which semantic processing takes place on different sensory inputs varies (see Fig. 1.2). This is in contrast to the early selection model of Broadbent where unattended sensory information is entirely blocked from transition to the higher levels at the early stages. Her model is based on her observations in Treisman (1960). The Treisman model will *attenuate* the unattended sensory inputs instead of eliminate them. The unattended inputs can be selected based on transition probability, given their relevance. For example, if one of the unattended channels include a familiar name (i.e., the person's name), then it is likely that the input will be picked up from the sensory buffer and passed to the higher levels of conscious processing. The model, therefore gives a better account of the 'Cocktail Party Syndrome'. The difference between the Treisman's model and Broadbent's model is the *location* of the bottleneck, that is to what extent the unattended messages are processed (See Fig. 1.2 (B)).
- Diana and J. Anthony Deutsch put forward a different interpretation of the selective model of attention. They argue that the selection of sensory items takes place after some perceptual and semantic processing in the model. Based on this view, all the sensory inputs are processed up to a semantic level and based on the weighting of the content, which is an indicator for its desirability and importance, they will be selected for the higher levels. Otherwise, the sensory inputs are thrown out of the mechanism. The difference between the Deutsch & Deutsch interpretation and the Treisman model is in the position of the filtering that is located much closer to the response end of the processing system. See Fig. 1.2 (C) for a schematic representation of the model. The critics of the model refer to the processing of all sensory items before filtering as redundant. Other neurophysiological evidences based on the study of ERP strength in the case of the presence of a target to the attended and unattended channels, were shown not to be consistent with Deutsch and Deutsch's theory (Coch et al., 2005; Eysenck and Keane, 2000).

There have been other theories proposed to determine if selective attention is a late or early stage process. Lavie and Tsal (1994) argue that the work load of the primary task has an important role on how participants' attention drifts to the unattended messages. If factors such as repetition or speed of messages in the shadowing ear increases, then all attentional resources are allocated to one specific channel. If load of messages or primary task is low, then some of the resources will be allocated to the unattended channel. Therefore, she argues under high-load condition, the filtering process should occur rather at the early stages.

The proposed models of selective attention have their own limitations and supports, mainly due to the fact that the researchers have used different measures for a participant's awareness of unattended

messages (i.e., from self report to physiological data such as galvanic skin response). It has to be noted that in Broadbent's model of selective attention, he argues that there are sensory buffers by which sensory information are stored temporarily. Attention can be shifted occasionally to the stored unattended information. It has been found that there is indeed sensory memory, composed of namely the iconic (visual modality) and echoic (auditory modality) memory. However, the timing shift of the attention has not been specified adequately in Broadbent's model. Therefore, the difference between an endogenous driven shift of attention and an exogenous effect (i.e., a salient stimulus) is not well defined. In Lachter et al. (2004), with a series of experiments applied to study selective attention, a modern variation of Broadbent's model is endorsed, by which the set of experimental results support the concept of slippage of Broadbent's model (allocation of attention to irrelevant items) against the leakage-based models (semantic processing of irrelevant items while attention is allocated somewhere else).

In the studies presented by Strauss et al. (2008a, 2010) and Corona-Strauss et al. (2011), a probabilistic model of the selective attention has been used. The model takes into account the exogenous (i.e., the physical characteristics of the input such as intensity, pitch) and endogenous factors (i.e., template driven) that influence the attention selection mechanism. The endogenous factors correspond to feedback that is received from the top-down processing and exogenous factors that correspond to the bottom-up processing. Fig. 1.3 shows a schematic illustration of this conceptual model of selective attention. The superposition of the endogenous and exogenous effects, defined as a d -dimensional weight vector $\omega = (\omega_1, \omega_2, \dots, \omega_d)$ is assigned to the input sensory streams $s^{(i)} = (s_1, s_2, \dots, s_d)$ where i is the step number and d is the total number of streams. The assigned weights have to be translated for the decision making circuitry for further analysis. The decision making circuitry for picking the relevant stream maps the input weights to a probability distribution, where it presents the likelihood of every stream selection based on its assigned weight value. Here, we assume the weights are sorted according to their selection probability, from highest to lowest probability that can be seen on the right side of the figure (the streams with a higher weight have a higher probability for selection). This step is comparable to the late/early filtering stage of the selective attention model described by Treisman and Deutsch. We denote the probabilities for every stream $P = (p_1, p_2, \dots, p_d)$. By assigning a threshold p_{th} , the sensory streams with a higher threshold than p_{th} will be selected for further analysis. The higher level (top-down) process that corresponds to the endogenous effect incorporates more details. More details regarding the neurophysical mapping of the model has been described in Strauss et al. (2008a, 2010).

1.3.3 Habituation

Habituation is known as a gradual reduction and filtering of irrelevant sensory inputs, and is probably the simplest form of leaning in the animal kingdom (Harris, 1943; Rankin et al., 2009). Through habituation, the animal *learns* to tune out the repetitive stimuli that have no significance, and

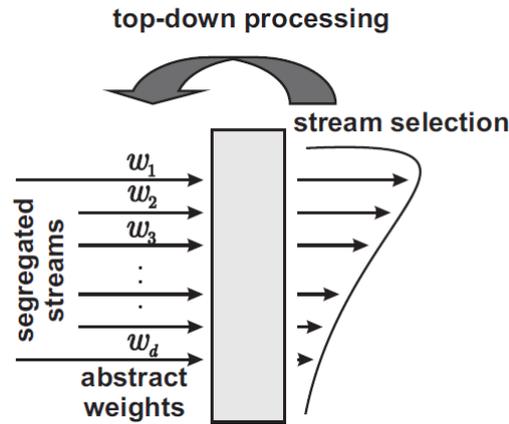


Figure 1.3: A probabilistic model of auditory stream selection. The weights w_i represent the superposition of exogenous and endogenous effects of the segregated sensory streams. The streams are stored based on their selective probability, a mapping from the weights to probability values. Image has been adapted from (Strauss et al., 2010).

instead to react to the novel stimuli (Rosen, 1992; Chandrasekaran et al., 2010).

Habituation has been observed in *Aplysia* where the siphon-elicited gill and siphon withdrawal reflex can undergo habituation that lasts for several weeks. This effect has been described in terms of long-term synaptic depression of the sensorimotor pathway that mediates the withdrawal reflex (Carew and Kandel, 1973). On the other hand, synaptic studies in short-term habituation experiments suggest a reduction in neurotransmitter release in presynaptic terminals (Hawkins et al., 1993; Thompson, 1986).

However, the effect of habituation can also occur without the Pavlovian reinforcement learning. It has to be also differentiated from sensory adaptation (receptor adaptation) or fatigue (effector fatigue) (Harris, 1943; Rankin et al., 2009; Domjan, 2014; Thompson, 2009). Habituation can occur in *short-term* and *long-term* forms. In the short-term habituation, the response strength decreases quickly and the spontaneous recovery is rather fast. The drift of attention in long-term habituation from irrelevant stimulus occurs over a longer time period with a longer lasting effect and has a slower spontaneous recovery.

In Thompson and Spencer (1966), a series of criteria are defined for distinguishing between habituation and other forms of learning. We briefly outline some of the most important criteria that have to be met for habituation process (for more details see Thompson and Spencer (1966) and the references therein).

- If a novel stimulus elicits a response, repetition of the stimulus will lead to decrement in the response.
- If the stimulus is withheld or presence of another novel stimulus, the response tends to recover.
- Rapid stimulations with shorter time intervals lead to more pronounced habituation.
- The weaker the intensity of the stimulus, the more habituation occurs.

- Habituation is a reversible process such that presentation of a novel stimulus will lead to dishabituation (Harris, 1943). It has to be noted that repetition of the dishabituation stimulation can lead to habituation (Crampton and Schwam, 1961).

The effect of long-term habituation influences a wide range of attention-oriented research, either due to its implication for data acquisition in enduring attention paradigms or due to its ability to shed more light on pathological attentional binding in a number of related dysfunctions (Rauschecker (2005); Walpurger et al. (2003); Williams et al. (2013)). The difference between sensory adaptation and habituation is subtle, however an important one. Sensory adaptation is the process of lessening the attention to a stimulus *that is not exerted consciously* (for example, we cannot control how quickly we can adapt to a particular stimulus such as light). Habituation on the other hand can be to some extent consciously *controlled* (Sternberg and Sternberg, 2016). Once habituated, we can consciously divert or shift our attention to the stimulus in which we have habituated, once asked to or decide. As another difference between habituation and sensory adaptation, we can refer to the role of the intensity level. In habituation, the duration of prior exposure to a stimulus is more important than its intensity, whereas sensory adaptation is more tied to intensity level than to prior exposure. We habituate much easier to a sound if we are exposed to it continuously over a period of time, whereas the skin sense receptors will respond the same to a temperature, no matter how long or often has it been exposed to that temperature before. Therefore, it is important to highlight that **sensory adaptation is a process that occurs mostly in our sense organs, orienting a bottom-up process whereas habituation is a cognitive adaptation governed mostly by top-down processes**(Sternberg and Sternberg, 2016).

It is important to point out that in the context of this thesis, long-term habituation is considered as the drift of attention away from an irrelevant stimulus, when the stimulus is frequently presented in a long-term. The physical properties of the stimuli can be salient, however, the endogenous top-down processes will lead to the drift of attention away from the repeated stimulus as no novel information is presented. Such mechanism helps us to allocate our attention on relevant stimuli and prevent cognitive overload.

Based on the probabilistic model of Strauss et al. (2008a, 2010) which was briefly described in the previous section, we can describe the effect of habituation as follow: a target stimulus s has been repeatedly presented to the subject over a long duration of time. The function f maps the stimulus to a probability value based on its corresponding weight, expressing the likelihood of selecting the stimulus for higher levels of cognitive processing. As the repetition of the stimulus increases, the corresponding novelty of the stimulus decreases and hence modifies the weight assigned to the stimulus. As the probability density of selection becomes less, the likelihood of selecting the stimulus decreases and hence the corresponding binding of the attention to the stimulus is reduced.

The two main factors that influence the degree of habituation are the *internal variation within a stimulus* and the subjective *arousal level* (Sternberg and Sternberg, 2016). As an example for the

first factor, we can compare the degree of habituation of the drone of an air conditioner to music presented continuously. As the latter one is subject to more variations and changes in time, it is comparably more difficult to remain habituated for a longer period of time in comparison to the constant noise of an air-conditioner. The subjective arousal level is an indicator for the *degree of physiological excitation and readiness for action* in comparison to the baseline. It can be measured using autonomic responses such as heart rate or central nervous system correlates in the EEG.

1.4 Decoding of Neural Signatures of The Long-term Habituation Process

The endogenous effect of selective attention has been firmly established as the modulation (enlargement) of the N1-P2 component of ERPs (see Hillyard et al. (1973); Hansen and Hillyard (1980); Hillyard et al. (1973, 1978) and references therein). The N1-P2 component is elicited as a result of the attended stimulus (see Fig. 1.1 as an illustration of the N1-P2 component). The main components of the interest are obtained by averaging through a sufficient number of single trials of ERPs, obtained from repetitive stimulation.

However, we are interested in the dynamics of ERP waveforms over time. The effect of N1-P2 has been investigated through the amplitude and phase characterization of ERP waveforms followed by different transformation techniques, such as the Wavelet and Hilbert transformation (Handy, 2005). The behavior of the IP has been particularly used in many studies as a feature for decoding the cognitive processes and biological signals in relation to different types of stimuli (Yeung et al., 2007; Busch et al., 2009; Makeig et al., 2002). Thereby, to assess to what degree the voltage changes of measured neural activities are time-locked to the target-stimulus of interest is one of the main evaluation methods for understanding the level of attention drift.

One neural signature of long-term habituation can be found in the changes of the N1-P2 component of auditory ERPs. Therefore, long-term habituation can be decoded as the gradual reduction in the N1-P2 component of measured ERPs (Thompson, 2009). The reduction of the N1-P2 wave has been studied both in the amplitude and IP of neural activities (Hillyard et al., 1973; Butler et al., 1969; Öhman and Lader, 1972; Rosburg et al., 2006; Babiloni et al., 2002; Busch et al., 2009; Low and Strauss, 2011; Morteza pouraghdam et al., 2015a). However, in the case of long-term habituation, there are not a lot of studies to assess the changes over single trials based solely on IP. Most of the studies mentioned earlier depend on the amplitude or measures based on amplitude-phase (Strauss et al., 2013; Mariam et al., 2012; Fuentemilla et al., 2006; Low and Strauss, 2011). In Strauss et al. (2008b) the IP was used exclusively but only on a trial-to-trial basis, i.e., for fixed points in time the phase of a trial is compared only to its direct neighbors, neglecting long-term changes in the phase dynamics. Consequently there is a pressing need for the development of methods which allow the long-term habituation assessment of the dynamics captured in the instantaneous phase

of ERP single-trial sequences. Moreover phase information are considered to contain less noise in comparison to the amplitude which is very susceptible to various artifacts such as eye/muscle movements or impedance of the skull (Andreas, 2000; Seraj, 2016).

Therefore, the first part of the work deals with applying/developing methods to understand the dynamic of IP of the N1-P2 component of measured ERPs. The experiments are designed such that they reflect the effect of long-term habituation for different sets of stimuli. The objective is to identify whether habituation occurs during the course of time, solely based on the IP of single-trial ERPs. A broad spectrum of stimuli from comfortable (of low-loudness or easy to habituate) to aversive (of high-loudness or difficult to habituate) are used to objectively determine the occurrence of habituation. The proposed methods objectively assess the underlying statistical properties of data and identify the regimes by which a significant change occurs. Using such methods, the incentive is to reveal a mechanism for revealing the changes in the N1 component of IP of ERPs and test for the presence of habituation.

1.4.1 Mapping Between Ongoing EEG and ERPs

One of the main limitations of studying ERPs is the short durations of stimuli, which are usually tone bursts, short sets of words or syllables (Hall, 1992). Therefore, the measured neural responses cannot be studied for longer intervals. Moreover, we are commonly exposed to complex stimuli rather than pure tone bursts. The sounds we encounter on a daily basis are longer in duration and also more complex. Therefore, to understand the decoding mechanism of selective attention for a longer duration of time and a continuous stimulus such as speech, we require a longer time interval. This requires the analysis of long rhythmic activities of ongoing EEG activities.

The relation between ERP and ongoing EEG activities is a debatable issue (Sauseng et al., 2007; Yeung et al., 2007, 2004; Klimesch et al., 2007). By averaging over sufficiently large number of ERP waveforms, the main components N1-P2 are seen as large peaks, reflecting the endogenous effect of attention in neural activities. However, it is assumed that the peaks in the ERP waves are uncorrelated to the background or ongoing EEG activities (Pfurtscheller and Da Silva, 1999; Yeung et al., 2004). This assumption has been challenged in the last three decades as extensive research has shown that the ERP components may as well be formed from synchronized EEG activities. In other words, the phase re-organization of ongoing EEG signals lead to generation of ERP peaks (Klimesch et al., 2007; Sauseng et al., 2007; Sayers et al., 1974; Penny et al., 2002).

In this part of the work, we assess the phase-reorganization of ongoing EEG activities for understanding selective attention for longer time intervals. The neural activities of attended versus unattended channels are analyzed and compared against the results obtained from studying ERPs. We thereby try to assess methodologies to map the phase modulations of ongoing EEG activities to results from the ERP studies.

In the following sections, we briefly examine some of the main theories for explaining the ERP

generation and its relation to ongoing EEG activities. Some of the main references that we resort to are Sauseng et al. (2007); Yeung et al. (2007); Klimesch et al. (2007, 2004b,a); Penny et al. (2002).

1.4.2 The classical view

In the classical view, the ERP is obtained by averaging over a sufficiently large number of single trials of recorded EEG which are produced with respect to an event. Based on this view, the EEG oscillations (often referred to as background EEG) are considered as *noise*, whose effect vanishes (i.e., tends to zero) after averaging (Pfurtscheller and Da Silva, 1999). The term noise merely implies that the background oscillatory activity is not *correlated* or time-locked with the event. It is hence believed that the ERP-waves are *evoked and embedded* in the background EEG, which obscures the informative ERP-wave components (Pfurtscheller and Da Silva, 1999). However, the informative part or the embedded ERP *signal* is present at every trial or sweep, and the effect becomes pronounced by averaging over sufficiently large number of trials. It is then concluded that the evoked activities are consistent across trials, which are described by activity bursts that are *time-locked* or time-coordinated with respect to the onset of a stimulus¹⁰. The mathematical formulation of the classical view can be seen as the linear superposition of the EEG response over a number of single trials (k):

$$ERP_{evoked} = \frac{1}{k} \left[\sum_{i=1}^k s(t, i) + W(t, i) \right]$$

at every time t . The averaged evoked activity is the superposition of all evoked responses s (k trials) at every time t . The ERP is superimposed and has no interaction with the background activities, which is considered as noise $W(t, i)$ for every time t and trial i . The classical view has been the base of the study of ERP analysis for the last decades, covering a broad range of studies in the field of cognitive processing and information processing.

1.4.3 Pure Phase-resetting

The proposed hypothesis against the classical view states that the ERP generation is not independent from the background EEG oscillations (Pfurtscheller and Da Silva, 1999; Sayers et al., 1974). It is assumed that ERPs are generated by the re-organization of stimulus induced phase-resets of ongoing EEG rhythms (Penny et al., 2002; Makeig et al., 2002; Sayers et al., 1974). Thereby, the ERP generation is not solely based on the superposition of evoked, fixed-latency, and fixed-polarity responses that are independent from the ongoing EEG activity (Sauseng et al., 2007). Based on this definition, the background EEG activity comprises an important part of the ERP generation

¹⁰If the recorded oscillation after the onset of a stimulus occurs with varying time-shifts and variations across different sweeps, the oscillations are considered *induced* by the stimulus rather than *evoked* (see Busch et al. (2009)).

process. This view is also referred to as *phase modulation* (PM), in contrast to the classical view, namely the *amplitude modulation* (AM). In Fig. 1.4, an illustration of the classical view of ERP generation versus the phase-resetting model has been illustrated.

To detect if the evoked responses over the trials have an AM behavior or PM dynamics, it is required to look at the spectral changes over single-trials. If there are no significant stimulus-induced power changes or increases over the single-trials, then it is highly likely that it is a PM process (Yeung et al., 2004). In this condition, the peaks and troughs over trials for a period of time are aligned and consistent, demonstrated as phasic bursts around the period related to the stimulus/event. Averaging over the trials will lead to the main peaks and troughs of the main ERP components. For any time outside this region, no consistency or phase alignment exists. If we consider the distribution of phase in the pre-stimulus region, the phase distribution is uniformly distributed, whereas post-stimulus will peak and center around a dominant value (Penny et al., 2002).

1.4.4 Pure-phase Resetting and Amplitude Enhancement

The pure-phase resetting, as was described in section 1.4.3, has been demonstrated to hold true for early auditory and visual EPs, as demonstrated by Jansen et al. (2003); Luo and Poeppel (2007); Makeig et al. (2002). However, some studies have shown that there exists a spectral power increase at specific frequencies as well as phase resetting that lead to the ERP generation (Brandt, 1997; Shah et al., 2004). This view is known as the phase-reset with enhancement and it is difficult to distinguish it from the classical view as both contain an increase in the power spectral of some frequencies. Additional analyses are required to assert if the generated ERP is obtained from the classical view or the phase resetting with enhancement as explained in Makeig et al. (2002) (for more details see Yeung et al. (2004) and references therein). In Yeung et al. (2004), it has been further investigated that the proposed analyses methods still contain ambiguities to clearly dissociate between the ERPs generated by the classical view and the phase-reset with enhancement approach.

1.4.5 Organization of the Thesis

This thesis is organized as follows: In Chapter 2, we describe the methodologies that have been used for individual studies. A brief introduction to the problem is provided at the beginning of every section which determines the study case. The methodology is then fully described and at the end, the experimental setting and data acquisition methods are explained. In Chapter 3, only the results are reported, in the same study order as in Chapter 2. In Chapter 4 we discuss results in Chapter 3 and state the limitations and shortcomings of every method. At the end, we present conclusions and future work.

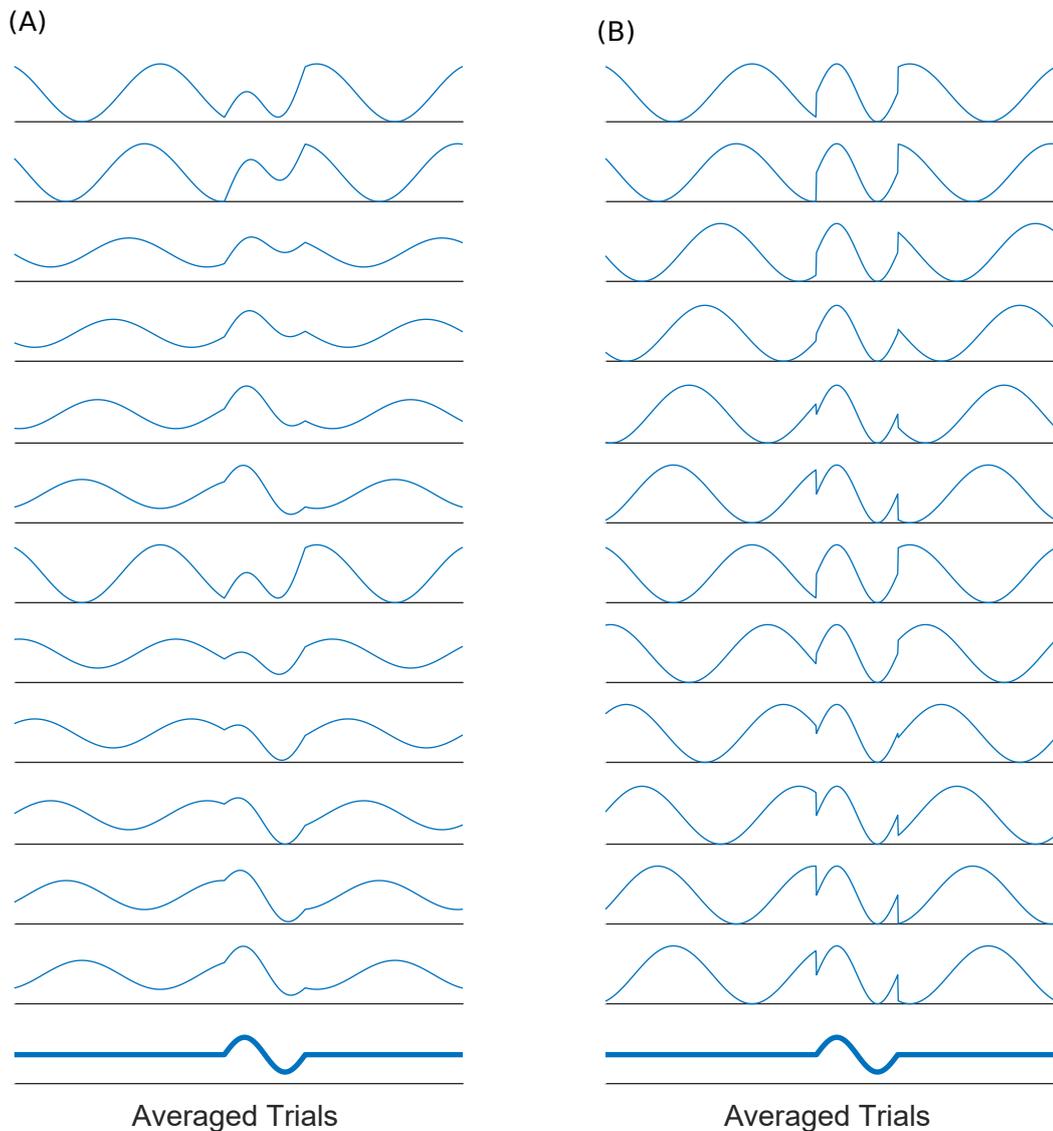


Figure 1.4: (A) The evoked classical model assumes that the evoked components of averaged ERP are generated by a constant evoked response that is added onto EEG activity. The evoked responses have the same latency and polarity for all trials. The average of all responses over all trials will yield the averaged ERP. (B) In other explanations of the ERP genesis model, the average evoked response is based on a phase-resetting of background EEG activity. By averaging over all trials in which phase reset has occurred, the same results as in (A) are produced. A better representation of the concept can also be found in Sauseng et al. (2007).

Chapter 2

Materials & Methods

2.1 Synchronized Neural Activity and Attention

Essential for sensory processing is the orchestration of brain regions responsible for stimulus appraisal (e.g., aversiveness by structures of the limbic system) and the thalamocortical feedback system. The nuclei of the thalamus act as a regulatory circuit for connecting the sensor to the neocortical areas of stimulus processing. As the projections from the thalamus to the cortex are sparse (approximately 15 %, see Benshalom and White (1986); Lübke et al. (2000)) the driving mechanism must be based on the temporal orchestration. Studies in Bruno and Sakmann (2006) demonstrated near-synchronous patterns of action potentials among converging thalamic inputs for strong sensory stimulation. Thus coherent neural activity, locked to the stimulus, may be a key concept for stimulus intensity appraisal (Fries, 2005; Börgers and Kopell, 2008).

The changes in the coherent neural activity in the thalamocortical circuitry should be reflected in non-invasive measurable EEG activities. The coherent activity in neural population can be observed in terms of phase synchrony among different neural assemblies. The study of phase synchronization has received a great deal of attention for decoding neural propagation for different mental/cognitive tasks between different regions (Schnitzler and Gross, 2005; Lachaux et al., 1999). Throughout various phase-related studies, the IP is studied using the analytic form of signals. Examples include the application of different BCI systems using the phase synchrony concept. Phase re-organization (phase resetting) and its relation to EEG has been investigated in many studies (see Penny et al. (2002); Makeig et al. (2002); Sauseng et al. (2007)). In Forte et al. (2017), the empirical mode decomposition technique is used to measure brainstem response to running speech. In addition, measures such as PLV (*phase locking value*) are numerous used to measure the level of phase

locking between different experimental trials or electrodes(See Lachaux et al. (1999) and references therein) and thereby decoding different cognitive tasks.

In all aforementioned studies, to analyze the dynamic of EEG oscillations and synchronization with regard to different stimulations and experimental paradigms, we need appropriate tools to study neural activities. This covers a broad spectrum of analysis tools from appropriate pre-processing approaches to characterizing signals in terms of phase and amplitude and statistical modeling of data. In the following sections, we explain some of the main pre-processing steps such as filtering and IP extraction to obtain meaningful pre-processed data. We then describe the related methods (applied/developed) for the study of our experimental data in relation to decoding auditory attention signatures and the long-term habituation. The utilized methods for every experimental setting as well as the data acquisition procedures are demonstrated. We refer to Chapter 3 for further discussions and analysis.

2.2 Extraction of Instantaneous Phase (IP)

2.2.1 Filtering procedure

Before the extraction of the IP, EEG data is usually narrow-bandpassed around a center frequency f_c , $\omega = 2\pi f_c/f_s$, with a sampling frequency of f_s . To have a narrow-bandpass filter, a small bandwidth is used around the frequency of interest. The main reason for applying a narrow-bandpass filter is to have a meaningful interpretation of the IP, as the IP in a broad frequency spectrum is not easy to interpret (Chavez et al., 2006).

Given the signal $x \in \mathbb{R}^n$, the bandpass filtering can be described as the convolution of the signal $x(t)$ with the filter's impulse response function h defined as follows

$$y[\cdot] = (x * h)[\cdot] = \sum_{k=0}^M h[k]x[\cdot - k]. \quad (2.1)$$

The bandpass filter is in the class of *linear time-invariant (LTI) systems*. The specific design of the filter properties and its parameters (i.e, the bandwidth, center frequency, stop-band attenuation, etc.) depends on the data and the application of interest. However, as we are interested in the analysis of IP, it is important that the *filtering process preserves the original phase of the input signal*.

A common approach for filtering and presenting the original phase is applying linear-phase finite impulse response (FIR) filters, in which the phase of the filter's impulse response is summed with the phase of the input signal. To narrow-bandpass a signal, the order of the FIR can substantially increase, in result increasing the length of the transient response. As these samples have to be dismissed, a long width of the signal is cut, thereby reducing the temporal resolution. To avoid a high order FIR, we can employ an infinite impulse response (IIR) filter, in which lower orders are

shown to narrow-bandpass the signal (Seraj and Sameni, 2017). However, the phase of IIR filters is often nonlinear and hence distorts the resulting input's signal phase. Distortion in case of using FIR can also occur due to phase wrapping of the filter's phase.

To avoid the problem of phase distortion, zero phase forward-backward filtering (FIR/IIR) is used. This standard method ensures a zero phase difference between the input and output's phase, regardless of the linearity or nonlinearity of the filter's phase response. Given an input signal $f(t)$ and a filter response k , the method applies the filtering as $\text{reverse}(k * \text{reverse}(k * f))$ such that *reverse* operator yields the temporal reverse of data. By setting $x = \text{reverse}(k * \text{reverse}(k * f))$ the output of the latter expression in Fourier domain can be described as $(\mathcal{F}x)(\omega) = \|K(\omega)\|^2 \mathcal{F}(\omega)$ for $\omega \in \mathbb{R}$, where $\mathcal{F}(\omega)$ and $K(\omega)$ correspond to the Fourier transform of the signal $f(t)$ and the filter response $k(t)$ respectively. The phase of the forward and the phase of the backward pass cancel out and therefore the resulting response in the frequency domain is the multiplication of the positive coefficients $\|K(\omega)\|^2$ (zero phase) with the Fourier transform of the signal¹. For more details on filters see Oppenheim et al. (1989).

2.2.2 Hilbert Transformation

After narrow-bandpassing the signal, the IP of the analytic signal can be obtained through different transformation techniques. One of the common methods for extraction of IP is using the analytic form of the signal using Hilbert transformation (HT). The analytic form of a narrow-bandpassed signal $\mathbf{x} \in \mathbb{R}^n$ using the Hilbert transform filter is computed as:

$$X(t) = x(t) + i\mathcal{H}(x(t)), \quad (2.2)$$

where the complex valued coefficients are the filtered signal in the real part and the Hilbert transform in the imaginary part. The HT of $x(t)$ is defined as

$$\mathcal{H}(x(t)) = \frac{1}{\pi} PV \int_{-\infty}^{\infty} \frac{x(t')}{t-t'} dt', \quad (2.3)$$

where PV is the *Cauchy Principal Value*. The definition of PV has been included in Appendix. A. The resulting output coefficients at every time step is a complex number, such that analytic amplitude and phase can be obtained as

$$A(t) = \sqrt{x^2(t) + \mathcal{H}(x(t))^2} \quad (2.4)$$

$$\theta(t) = \text{atan2} \left(\frac{\mathcal{H}(x(t))}{x(t)} \right) \quad (2.5)$$

respectively. Here *atan2* denotes the "quadrant-specific" inverse of the tangent function. The definition enforces the 2π -periodic angles $\theta(t)$ to be in $[-\pi, \pi)$.

¹In matlab the function *filtfilt* applies the zero-phase forward backward method

Using Euler's formula, the complex coefficients can be represented as points on a circle. Assume the given complex coefficients as $x(t) + iy(t)$ the polar representation of the data sample is $A(t)e^{i\theta}$ where $A(t)$ is the analytic amplitude and the θ is the analytic phase angle.

2.2.3 Wavelet Transformation

One of the other techniques used to compute the IP of the analytic signal with a high time-frequency localization is wavelet transformation (WT). The transformation is the convolution operation of the signal $x(t) \in \mathbb{R}$ with a function $\psi \in L^2(\mathbb{R})$ known as the *mother wavelet*². A one dimensional signal is described in terms of time and frequency into a two dimensional space by expressing the signal in terms of a series of orthonormal basis functions (in this case the wavelet functions). The minimum requirement imposed on the mother wavelet function $\psi(t) \in L^2(\mathbb{R})$ is the *admissibility condition*, that is:

$$0 < C_\psi := \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|}{|\omega|} d\omega < \infty \quad (2.6)$$

where $\Psi(\omega)$ is the Fourier transform of the function $\psi(t)$, with ω being the frequency. The admissibility constant factor is shown as C_ψ . This condition imposes a mild decaying criteria for the function. From Eq.2.6 it follows that

$$\int_{-\infty}^{\infty} \psi(t) dt = 0. \quad (2.7)$$

The latter criteria can be interpreted as the mother-wavelet behaving similar to a wave with ups and downs in the time domain together with the decaying property. This assumption leads to a better localization and tracing of the changes across the time and frequency. See Strang and Nguyen (1996) for more details on wavelets and explanations.

A family of wavelet functions from the function ψ is formed by scaling (with the dilation parameter $s \in \mathbb{R}, s \neq 0$) and shifting (with the translation parameter $\tau \in \mathbb{R}$) its center in time. The translated and dilated function $\psi_{s,\tau}$ is equivalent to

$$\psi_{s,\tau}(\cdot) := \frac{1}{\sqrt{|s|}} \psi\left(\frac{\cdot - \tau}{s}\right), s, \tau \in \mathbb{R}, s \neq 0. \quad (2.8)$$

Given the form of the wavelet function, the CWT of a signal $x \in L^2(\mathbb{R})$, with respect to a wavelet function $\psi_{s,\tau}$ is defined as:

$$(W_\psi \mathbf{x})(s, \tau) = \langle \mathbf{x}, \psi_{s,\tau} \rangle_{L^2} = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} x(t) \psi^*\left(\frac{t - \tau}{s}\right) dt. \quad (2.9)$$

In Eq. 2.9, the function $x(t)$ is mapped into a new space by a family of basis functions ψ and ψ^*

² $L^2(\mathbb{R})$ is the set of square integrable functions

is the conjugate operation. The new space is described in terms of the parameters s and τ , which gives us the localized information of $x(t)$ in time and frequency domain³. For more details on CWT and inverse CWT see Oppenheim et al. (1989).

Complex Continuous Wavelet Transform One of the methods for extracting IP is to utilize complex wavelet functions. Given a complex wavelet function $\psi_{s,\tau}$ and a real-valued signal $\mathbf{x} \in \mathbb{R}^M$, the resultant CWT will be complex valued ($v_{s,\tau} := \langle \mathbf{x}, \psi_{s,\tau} \rangle, \tau = 1, \dots, M$). The complex valued coefficients are expressed as:

$$v_{s,\tau} = \Re(v_{s,\tau}) + i\Im(v_{s,\tau}) = |v_{s,\tau}| \exp^{i\theta_{s,\tau}}. \quad (2.10)$$

The $\Re(v_{s,\tau})$ is the real and $\Im(v_{s,\tau})$ is the imaginary parts of the complex coefficient $v_{s,\tau}$. The terms $|v_{s,\tau}|$ is the *amplitude* and $\theta_{s,\tau}$ is the *instantaneous phase, IP*. The instantaneous amplitude and IP using the complex coefficients $v_{s,\tau}$ are computed according to Eq. 2.5. We denote the IP as $\theta_{s,\tau}$ (or θ).

In order to separate the amplitude and phase information we are required to use complex wavelet functions with a positive frequency response in the Fourier domain, that is $F(\psi)(\omega) = 0$ for $\omega < 0$. A wavelet with such property is referred to as an *analytic* wavelet and is defined in the *Hardy space*. One of the informative wavelet functions that have been studied thoroughly in the former studies (Strauss et al., 2005, 2008b; Low and Strauss, 2011) is the sixth derivative of a complex Gaussian function. Strauss et al. (2008b) demonstrates that for the chosen wavelet scale $s = 40$ (upper theta band (Matsuoka, 1990)) a significant (p<0.05) correlate of fluctuations in the N1/P2 component of auditory ERPs exists and thus a correlate for attention, with satisfactory temporal resolution. For a relation between the analytical signal of a bandpass-filtered signal and the complex wavelet transform we refer to Stéphane (1999); Bruns (2004).

The information associated with the selected wavelet function are as follows: Fig. 2.1 illustrates the selected wavelet function $\psi_{s,\tau}$, the sixth derivative of a complex Gaussian function defined as follows in Matlab

$$\psi(t) := c \left(e^{it} e^{-t^2} \right)^{(6)}$$

where c is chosen such that $\|\psi\|_2 = \sqrt{\int_{-\infty}^{\infty} |\psi(t)|^2 dt} = 1$. The selected scale $s = 40$ corresponds to a center frequency of $7.6Hz$. We investigated the frequency domain of the wavelet function. In Fig. 2.2, we have plotted the absolute value of the Fourier transform of the wavelet function. The fourier transform of the sixth derivative of a complex Gaussian function is given by (For details regarding the derivation see Appendix. B):

³This is similar to Fourier transform, in which we map the signal into a new space by a set of orthonormal basis functions. The differences are in the selection of basis functions (a series of sines and cosines for Fourier transform and a series of wavelets for WT) and that in the Fourier transform we do not obtain information in the time domain.

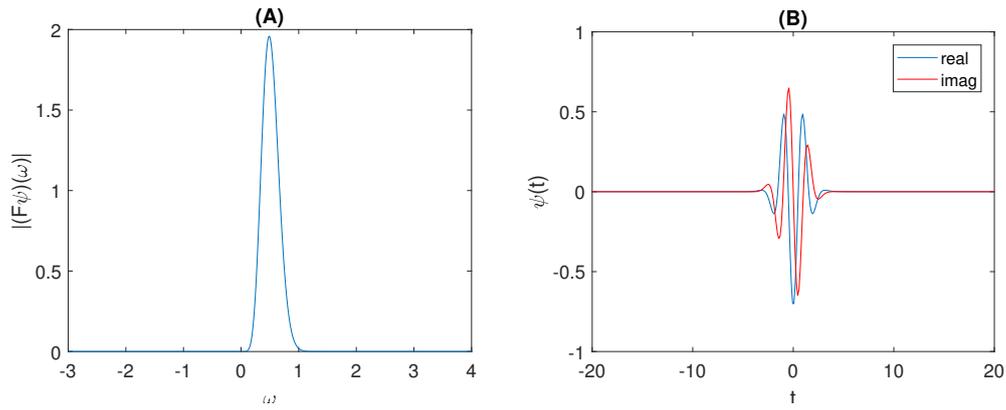


Figure 2.1: (A): The absolute value of the fourier transform of the truncated 6th-derivative of a complex gaussian function $F(\psi)$. (B): The corresponding function ψ in the time domain.

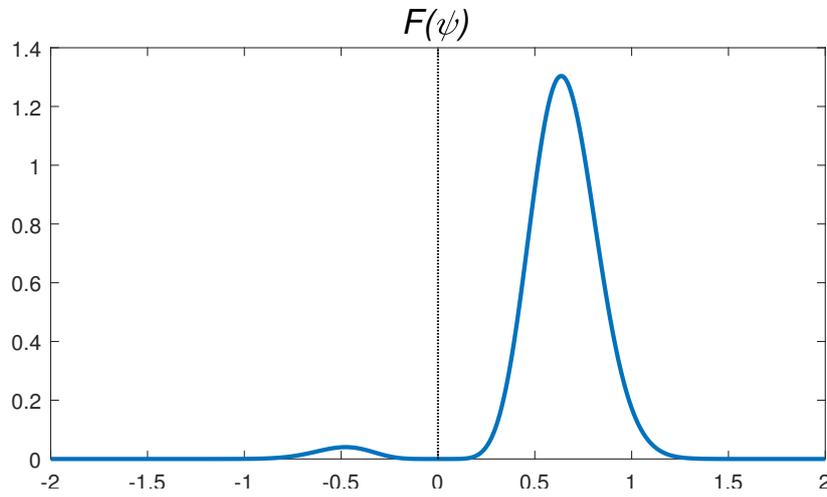


Figure 2.2: The absolute value of the fourier transform of the 6th-derivative of a complex gaussian function $F(\psi)$.

$$(\mathcal{F}\psi)(\omega) \propto -\omega^6 e^{-(\pi\omega - \frac{1}{2})^2} \quad (2.11)$$

where ω is the frequency.

As shown, the information in the negative frequency is significantly diminished (as the order of the complex Gaussian increases, the Fourier transform acts as a filter in the positive frequency and removes the negative frequencies). Although the effect of the negative frequencies is significantly small, to maintain a Hardy space and avoid distortions in the distribution of the extracted phase information, the negative frequencies are set to zero. Hence, we define the truncated sixth-derivative

of a Gaussian wavelet ψ by its Fourier transform as follows:

$$(F\psi)(\omega) = \begin{cases} -c\omega^6 e^{-(\pi\omega - \frac{1}{2})^2} & \text{if } \omega \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where c is chosen such that $\|F\psi\| = \|\psi\|_2 = 1$ (See Fig. 2.1 for the Fourier transform of the truncated version).

The complex wavelet transform of a signal $f(t)$ for a fixed *scale* s is a function $\hat{f}(\tau)$ of the *translation* parameter τ :

$$\hat{f}(\tau) := \frac{1}{|s|^{1/2}} \int_{-\infty}^{\infty} f(t) \psi^* \left(\frac{t - \tau}{s} \right) dt.$$

We can express \hat{f} as the convolution of f with the function $g(t) := \frac{1}{|s|^{1/2}} \psi^*(t/s)$. To extract the analytic signal of our data in the theta band, we use the scale $s = 40$. This transformation is especially simple to implement, since it corresponds exactly to a Hilbert transform followed by a wavelet transform with the ordinary sixth derivative of Gaussian wavelet. We obtain the phase of a signal f as the phase of its complex analytic signal.

2.3 Statistical Analysis and Decoding of Neural Responses of Attention-Binding

2.3.1 Descriptive Circular Statistics

Circular data has two main properties in which arithmetic means and variances cannot be used for describing the circular statistical properties (Fisher, 1995; Mardia, 1975). For example, the arithmetic mean of two angles 1° and 359° will be 180° whereas intuitively the mean angle should be around 0° . The different algebraic structure of the circle in comparison to the line, that is on a circle only one operation (addition modulo 2π) and on a line two operations (addition and multiplication) are defined, leads to different forms of central theorem and stability conditions (see (Mardia and Jupp, 2009)). On a circle, the samples depend on two properties, namely, the zero direction which is the starting point and the direction of rotation, i.e., whether clockwise or anti-clockwise. A definition of the circular mean and variance should be *independent* of the particular choice of these two quantities.

Mean Direction Assume $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ is a series of data points on a unit circle, measured in the direction of anti-clockwise. Every point x_i can be represented by an angle θ which is measured from the positive x -axis to the line connecting the origin to the point. We denote the set of angles as $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$, $\theta_i \in [-\pi, \pi)$ and the direction vector for every point as $v_i = (\cos \theta_i, \sin \theta_i)^T$, $i = 1, \dots, n$. The *resultant vector* \mathbf{R} of the corresponding directions

v_i is defined by

$$\mathbf{R} := \left(\sum_{i=1}^n \cos \theta_i, \sum_{i=1}^n \sin \theta_i \right)^T = (C, S)^T$$

where

$$C := \sum_{i=1}^n \cos \theta_i \quad \text{and} \quad S := \sum_{i=1}^n \sin \theta_i. \quad (2.12)$$

The *circular mean* $\bar{\theta}$ is the angle in $[-\pi, \pi)$ of the direction of \mathbf{R} , i.e.,

$$\bar{\theta} := \text{atan2}(S, C), \quad (2.13)$$

where atan2 is the inverse of the tangent function and its output is confined to the interval $(-\pi, \pi]$.

Similarly as the usual mean \bar{x} of real values $\{x_i : i = 1, \dots, n\}$ fulfills

$$\bar{x} = \underset{x \in \mathbb{R}}{\text{argmin}} \sum_{k=1}^n |x - x_k|^2 = \frac{1}{n} \sum_{i=1}^n x_i,$$

the circular mean satisfies

$$\bar{\theta} = \underset{\theta \in [-\pi, \pi)}{\text{argmin}} \sum_{i=1}^n d(\theta, \theta_i)$$

with the circular distance $d(\theta, \eta) := 1 - \cos(\theta, \eta)$. Note that d does not coincide with the geodesic distance $\min(|\theta - \eta|, 2\pi - |\theta - \eta|)$ on the unit circle.

Mean Resultant Length The length of the resultant mean is defined by $R' = n|\mathbf{R}|$. We define $S = 1 - R'/n$, which is between $(0, 1)$. If the directions are highly clustered around the mean direction (or in the extreme case of coincidence), then the length R' will be almost as large as n and S will be nearly zero, indicating a high clustering of data. In contrast, as R' gets lower, S will be closer to 1 implying a more uniform distribution of data directions around the circle. However, $R' = 0$ does not necessarily imply that the data points are uniformly spread on the unit circle. In Fig. 2.3 we show an example of data points, which despite a high orientation of data samples in two different directions, the resultant mean length is very small (~ 0.024). Visualization of data is necessary in order to obtain a better understanding of data distribution.

Circular Variance The quantity $V = 1 - R'$ is known as sample *circular variance* for $0 \leq V \leq 1$ (Mardia, 1975; Fisher, 1995). A variance of zero and a resultant length of one means a low variability between the data points (i.e., all directing to the same direction), and a high variance of one (with a small resultant length) refers to a uniform distribution of data. Similar to R' , $V = 1$ does not necessarily imply that data is uniformly distributed around the circle. As was described for the mean resultant length (Fig. 2.3), a bimodal distribution of data, pointing to the opposite sides will have a small resultant mean length and low variance, not capable of

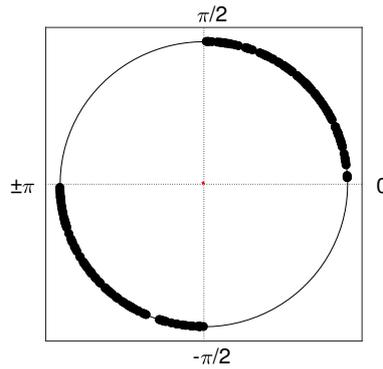


Figure 2.3: An example of a bimodal data distribution, spread at two opposite sides of a circle. The mean resultant length R' is ~ 0.024 . Despite the low value of R' , the measure doesn't reflect the actual data concentration around the circle.

capturing the underlying two modes. For more information see Mardia (1975); Fisher (1995).

2.3.2 Probabilistic Modeling of Instantaneous Phase

Probabilistic modeling of data is considered as an important feature that allows a better understanding of data structures and the underlying generative models. Fitting a model to data (either parametric or non-parametric modeling) enables us to describe the data efficiently by a set of parameters. Using more complex modeling frameworks such as Bayesian approaches, we are able to obtain detailed information from the data structure and corresponding underlying model. As the statistics for circular and spherical data are particularly different than linear data (as explained in section 2.3), modeling data requires appropriate tools that take into account the circularity characteristics of data structures. We are particularly interested in studying of behavior of IP in order to decode neural activity with regard to auditory selective attention. Therefore, we employ and develop new approaches for understanding IP of auditory selective attention for different experiments so as to give us a deeper insight into data. In each of the following sections, we will describe the experimental settings and the methods that have been used for data analysis. Results and discussions are fully elaborated on in chapters 3 and 4 respectively.

Von-Mises Modeling of Windowed Instantaneous Phase

The objective of this study is to assess statistically the dynamic behavior and characteristics of long-term habituation as reflected in the IP of auditory ERP sequences. To accomplish this, we use a theoretically well founded quantitative multiscale modeling and experimental framework for human ERP single-trial data. We apply a quantitative neurofunctional model that covers several spatiotemporal scales of neural processing to simulate human macroscopic ERP sequences. This macroscopic simulation output depends on an oscillatory microscopic model of the cortico-limbo-

thalamocortical circuitry in which dynamics were fitted to invasive electrophysiological data. Thus the model allows us to simulate the macroscopic ERP correlates of long-term habituation as defined by the microscopic dynamics. For reasonable parameters of the time–frequency transform, we assess the IP of the simulated ERP sequences with a von Mises model, providing us with predictions about the circular organization of the phase during attentional binding. Measured habituation data is used to cross-validate the model’s prediction by a circular statistics approach that has not considered for the assessment of long-term habituation thus far.

This method can be used in the experimental/clinical neurodiagnostic assessment of attentional binding and also provides data for fitting time-dependent and phase-related parameters in quantitative models of long-term habituation to different species, exogenous conditions, and modalities. The model also links the macroscopic dynamics of the IP to a microscopic oscillatory model, at least in our multiscale framework.

We first describe the neurofunctional model of attention and the simulation of the ERP responses. Next, we explain about the data measurements, and at the end, the proposed probabilistic model for studying the circular regularities over the measured ERP-trials is described.

Neurofunctional Model of Attentional Regulation

The quantitative framework of the model that has been used in this study is a continuum model of thalamocortical interaction introduced in a series of papers by Robinson et al. (2005) (see also the references therein). The model simulates ERPs based on a set of reasonable physiological parameters. In a study conducted by Trenado et al. (2009), the model was further expanded to simulate deviant levels of attentional binding towards a given auditory stimulus by modulating the internal gains. Furthermore, extrasensory neurofunctional structures were integrated to add behavioral dynamics in a simulation of the effects of habituation onto ERP morphology. A comparator structure based on physiological and anatomical hippocampus data (see Vinogradova (2001)) was added to simulate the response effect of novelty and fading attentional binding during recollection and habituation (Haab et al., 2011).

In particular, we used the theta-regulated attention hypothesis (Vinogradova, 2001) to compare between feature-reduced novel sensory data and preprocessed data in memory. This preprocessed input in memory might contain a prediction of sensory environmental dynamics, but considering a static surround for the sake of convenience, we can use memory synonymously. In accordance with Sato and Yamaguchi (2007); Yamaguchi et al. (2007), we can hypothesize a feature reduction of sensory data to a level of directionality and exogenous weighting in an egocentric environmental mapping.

Fig. 2.4 illustrates the functional interconnections of the model with all interacting neural elements. For a detailed analysis of the importance of all involved neural structures see Vinogradova (2001). The projection to the thalamus might influence attentional gating Crick (1984); McAlonan et al.

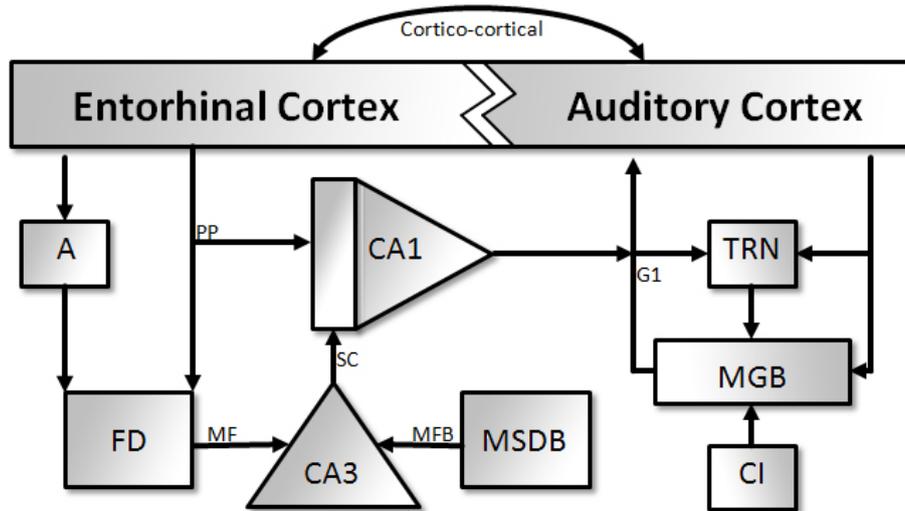


Figure 2.4: Descriptive model of hippocampal functional elements and interconnections used for the numerical implementation of the comparator model; FD (fascia dentata), CA1&CA3 (cornu ammonis 1 & 3), MSDB (medial septum, diagonal band of Broca), RN (raphe nuclei), MFB (medial forebrain bundle), SC (Schaffer's collaterals), PP (perforant pathway), MF (mossy fibres), A (amygdala). The thalamocortical feedback system consists of the TRN (thalamic reticular nucleus), MGB (medial geniculate body) using the cochlear input (CI).

(2000); Zikopoulos and Barbas (2006); Kimura et al. (2012) and recently the convergence of limbic and attentional pathways in the thalamic reticular nucleus (TRN) has been highlighted in Zikopoulos and Barbas (2012). The model outlined above is integrated into the thalamocortical attention model by Trenado et al. (2009) to achieve comparability to recorded ERP data. Simulation data on habituation effects show similar trend over the number of subsequent stimulus presentations as compared to recorded data. The modeling parameters used to perform this study were taken from the same set of physiological neuron properties as described in Strauss et al. (2008b); Haab et al. (2011, 2009).

Data Generation using the Neurofunctional Model

We used the aversive binding model in Mariam et al. (2009) for habituation and non-habituation. In this model, a repetitive, monotonic, and periodic soft/comfortable loud auditory pure tone stimulation of 50dB(SPL) results in habituation due to a lack of novelty and limbic significance. A loud stimulation of 100dB(SPL) of the same type serves as aversive sound that leads to limbically-triggered attentional binding, resulting in non-habituation ERP correlates (see section 2.3.2 for more details and the corresponding experimental setting). The habituation time-course, namely the reduction of the CA1 integration time window, was based on the physiological data of Vinogradova (2001) and follows a sigmoidal function (Haab et al., 2011). We used this sigmoidal function as a limbic input into the model of thalamocortical interaction to simulate decreasing excitatory effects to the thalamocortical (G1) gain (Haab et al., 2011). Note that the model is not fitted to macroscopic

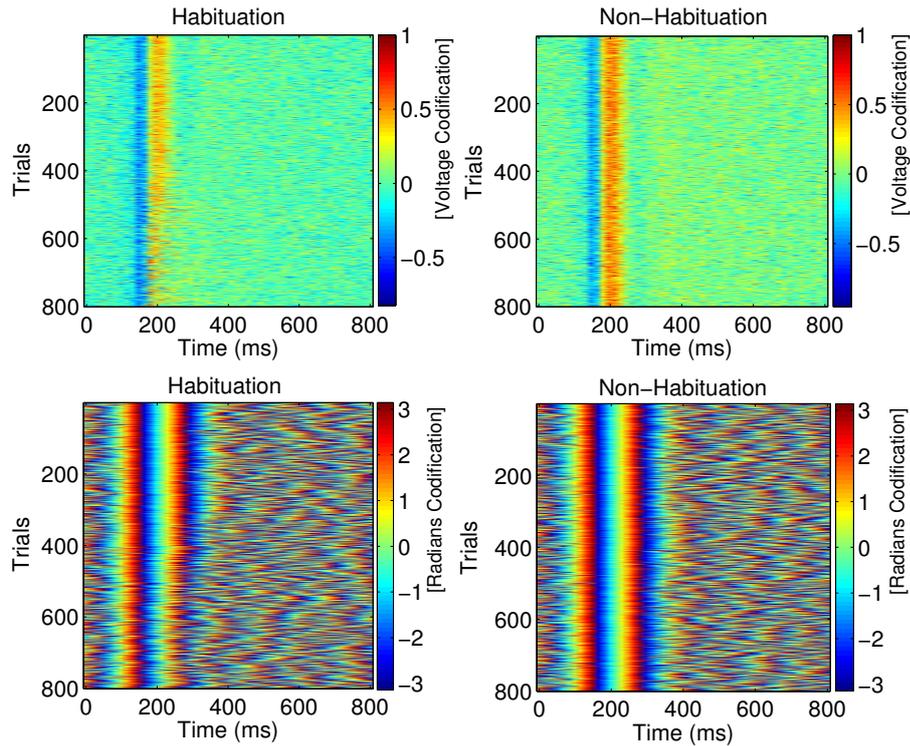


Figure 2.5: Top: Results for simulated habituation and non-habituation data are illustrated for 800 trials. For the simulated habituation data, the N100 wave and P200 wave (between 80ms and 220ms) is highly reproducible for the first 300 trials and then its strength tends to diffuse over subsequent trials. This is in contrast to the simulated non-habituation data that shows a strong N100/P200 reproducibility throughout all the trials. Bottom: the corresponding (scale specific) IP of the matrices in the top row.

ERP data. Only the microscopic oscillatory model is fitted to invasive electrophysiological data, being on a much smaller spatio-temporal scale of neural processing than our macroscopic simulation ERP output.

Fig. 2.5 shows an example of the simulated data for a topographic ERP image (see (Strauss et al., 2013)), i.e., individual sweeps form the rows of a matrix, in which the amplitude is color-coded. It is noticeable that the absolute value of the N100 wave is decreasing (top). Note that this effect results purely from the model used for generating the data. Such a clear trend is not noticeable in measured ERP single-trials due to noise and the superposition effect of other on-going neural processes.

Fig. 2.5 (bottom) shows the IP of the trials. The model predicts that the IP is less coherent at the end of the experiment in the N100/P200 time frame for the habituation setting (left) but coherent for the non-habituation case (right). In other words, the simulated drift of selective attention results in an unstable phase in the long-term habituation setting. The 'phase effect' was never analyzed independently of the amplitude for habituation analysis up to now.

In order to investigate the effect of long-term habituation in IP, we suggested the use of a von Mises modeling over time. The statistical analysis of phase data using a von-Mises model approach has

been applied on synthetic as well as measured ERPs.

Experimental Setting: Data Acquisition and Segmentation

Data from the long-term habituation study was acquired from 10 subjects with no hearing deficit history at the Saarland University. The mean age of the participants was 30 years and 11 months \pm (3 years and 9 months). Each subject received an audiogram test before the experiment and an audiogram check up after the experiment. Subjects were asked to lie on a bed in a sound-proof room with eyes closed and avoid any motion during the experiment. The evoked potential responses to the individual tones were recorded using surface electrodes (Ag/AgCl) which were placed at the right and left mastoid (active), the vertex (reference) and the upper forehead (ground). The signal acquired from each mastoid was referred to vertex (Cz) and processed separately. In this work we analyze the signals that are referred to the ipsilateral side, which is defined as the signal and electrode placement on the same side of the head, thus changing with stimulus direction. The sound stimuli were presented only to the right ear via a headphone (HDA 200, Sennheiser) at two different sound levels of 50dB(SPL) and 100dB(SPL). A stimulation to the right ear would require the analysis of the right-side mastoid electrode versus the reference. As 50dB(SPL) is a low sound intensity and the subject can easily habituate to the sound, contrary to 100dB(SPL) which is a loud and aversive stimulus and hence habituation would be diminished. A break of 3min was given between the 50dB(SPL) and 100dB(SPL) sound exposures. The auditory stimuli consisted of pure tones of 1 kHz with a duration of 40ms, and a constant inter-stimulus interval (ISI) of 0.75s. Stimulus presentation lasted for 8min each with a 3min break in between. High- and low-intensity sessions were presented in a random order to individual subjects to exclude sensory adaptation. During the experiment, subjects were consistently observed to make sure that they were conscious and motionless. The recorded EEG data was sampled at 512Hz.

Data Pre-processing A bandpass FIR filter of order 1000 and cut-off frequencies of 1Hz and 30Hz was applied to the raw EEG signal in order to remove unwanted frequencies. The high filter order leads to having a sharper stop-band. Afterwards, the acquired data from each experiment was divided into segments of 800ms according to a stimulus triggered signal with the sampling rate of 512Hz we obtain $T = 410$ values per segment/trial.

Trials containing artifacts and amplitudes larger than $50\mu V$ were removed. In order to have the same number of trials for each subject, a total of $N = 800$ artifact-free trials $\mathbf{s}_k^T \in \mathbb{R}^T$, $k = 1, \dots, N$ were used in the analysis for each subject. The data is represented as an $N \times T$ ERP image

$$\mathbf{S} := (\mathbf{s}_1, \dots, \mathbf{s}_N)^T,$$

where the trials $\mathbf{s}_k^T \in \mathbb{R}^T$ appear as rows. An example of the acquired data for one subject and 50dB(SPL) responses is shown in Fig. 2.6. The 800×410 matrix \mathbf{S} can be seen on the right hand

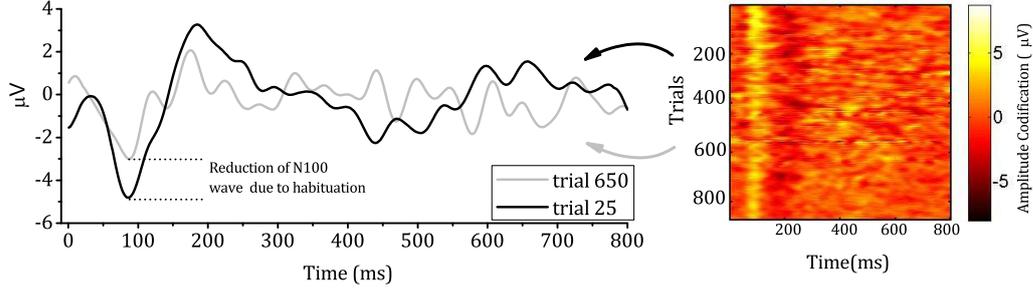


Figure 2.6: Example of ERP image \mathcal{S} of 50dB(SPL) responses is shown on the right side where each row represents a single trial of 800ms. On the left side two different trials are selected. The effect of habituation can be observed as a reduction of the N100 wave amplitude that appears approximately 100ms after the onset of the stimuli.

side. Two different single trials \mathbf{s}_{25}^T and \mathbf{s}_{650}^T are depicted in the left image. The effect of habituation can be observed in terms of a reduction of the N100 wave that appears approximately 100ms after the onset of stimulus.

Finally, we applied a two-dimensional non-local means (NLM) filter to the images \mathcal{S} for denoising (see Appendix.C for details regarding the method). This scheme exploits the self-similarity in the event-related activities \mathcal{S} . The NLM filter was applied as described in (Strauss et al., 2013) with the same parameters (We used a fixed 1×11 similarity patch, an asymmetric Gaussian with $\sigma = (1.0, 5.0)^T$ and $\lambda = 1000$ being the denoising parameter). In Chapter 3, we will deal both with original and the NLM filtered images as to make sure that the filtering process has not invalidated the long-term habituation effect.

IP Extraction We compute the IP of the band-filtered analytical signals of \mathbf{s}_k , $k = 1, \dots, N$ according to section 2.2.3 for a fixed $a > 0$ (as described in section 2.2.3, the section on the complex continuous wavelet transform, $a = 40$ has been shown to present physiologically meaningful correlates of neural data). The resultant complex wavelet coefficients are computed as

$$w_{k,b} := \langle \mathbf{s}_k, \psi_b^a \rangle, \quad b = 1, \dots, M.$$

Here ψ was chosen as the sixth derivative of the complex Gaussian, and ψ_b^a is the sampled vector of $\psi_b^a(\cdot) := |a|^{-1/2} \psi((\cdot - b)/a)$. From the complex coefficients

$$w_{k,b} = \text{Re}(w_{k,b}) + i \text{Im}(w_{k,b}) = |w_{k,b}| \exp(ip_{k,b})$$

we obtain the phase $\theta_{k,b}$ for every trial. The phase matrix is denoted as \mathbf{P} to the corresponding matrix \mathcal{S}

$$\mathbf{P} = (\Theta_1, \Theta_2, \dots, \Theta_N)^T.$$

(See section 2.2.3 for more details). We model the changes in the clustering level of phase data at a specific time over all trials. See Fig. 2.7 as an example.

Parametric Modeling using a Von Mises Distribution One of the most popular parametric models for the analysis of circular data is the von Mises distribution that resembles the normal distribution of real-valued data. It was introduced by (von Mises, 1918) and discussed earlier in a physical context in (Langevin, 1994). The probability density function (pdf) of a von Mises distributed random variable is given for $-\pi \leq \theta < \pi$ by

$$f(\theta; \mu, \kappa) = (2\pi I_0(\kappa))^{-1} \exp(\kappa \cos(\theta - \mu)) \quad (2.14)$$

where $\mu \in [-\pi, \pi]$ is the *mean* and $\kappa \geq 0$ the *concentration*. Further $I_r(\kappa)$ denotes the *modified Bessel function of the first kind and order r* . More details on Bessel functions are given e.g. in (Kent, 1978). As $\kappa \rightarrow \infty$ the data becomes more clustered towards the mean, the probability distribution has a high kurtosis. As $\kappa \rightarrow 0$, the data becomes more spread out around the mean, approaching a uniform distribution. Next we deal with parameter estimations of von Mises pdfs using a maximum likelihood approach. Assume that the data in the set $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ is identically and independently distributed and drawn from a von Mises distribution with unknown mean μ and concentration κ . We want to estimate these parameters as maximizers of the likelihood

$$L(\mu, \kappa | \Theta) := \prod_{i=1}^n \frac{1}{2\pi I_0(\kappa)} \exp(\kappa \cos(\theta_i - \mu)).$$

The log-likelihood $l = \log L$ which has the same maximizer is

$$\ell(\mu, \kappa) = -n \log(2\pi I_0(\kappa)) + \kappa \sum_{i=1}^n \cos(\theta_i - \mu). \quad (2.15)$$

Then the parameters can be found by setting the first derivatives of ℓ with respect to μ and κ to zero, respectively. For the derivative with respect to μ we obtain

$$\frac{\partial \ell}{\partial \mu}(\mu, \kappa) = \sum_{i=1}^n \sin(\theta_i - \mu) = 0$$

Using $\sin(x - y) = \sin x \cos y - \cos x \sin y$ and setting C and S as in (2.12) this can be rewritten as

$$\sum_{i=1}^n (\sin \theta_i \cos \mu - \cos \theta_i \sin \mu) = 0,$$

$$S \cos \mu - C \sin \mu = 0.$$

Hence the estimated mean is

$$\hat{\mu} = \text{atan2}(S, C) = \bar{\theta}, \quad (2.16)$$

and coincides with the circular mean $\bar{\theta}$ of the samples from Θ in (2.13). The differentiation of the log-likelihood with respect to κ gives

$$\frac{\partial \ell}{\partial \kappa}(\mu, \kappa) = -n \frac{I_0'(\kappa)}{I_0(\kappa)} + \sum_{i=1}^n \cos(\theta_i - \mu) = 0$$

and since $I_0'(\kappa) = I_1(\kappa)$ further

$$-n \frac{I_1(\kappa)}{I_0(\kappa)} + \sum_{i=1}^n \cos(\theta_i - \mu) = 0.$$

Substituting the estimated mean (2.16) into this equation and using

$$\begin{aligned} \sum_{i=1}^n \cos(\theta_i - \bar{\theta}) &= \sum_{i=1}^n \cos \theta_i \cos \bar{\theta} + \sin \theta_i \sin \bar{\theta} \\ &= C \cos \bar{\theta} + S \sin \bar{\theta} = |\mathbf{R}|, \end{aligned}$$

where the last equation follows from Eq. 2.13 and we obtain

$$\frac{I_1(\kappa)}{I_0(\kappa)} = \frac{|\mathbf{R}|}{n}. \quad (2.17)$$

Unfortunately Eq. 2.17 does not provide a closed form solution for κ . Instead we use the following common approximation based on a series expansion method of Fisher, see (Fisher, 1995, p. 51):

$$\hat{\kappa} = \begin{cases} 2r + r^3 + \frac{5}{6}r^5 & \text{if } r < 0.53, \\ -0.4 + 1.39r + 0.43(1-r) & \text{if } r \in [0.53, 0.85), \\ (r^3 - 4r^2 + 3r)^{-1} & \text{if } r \geq 0.85, \end{cases}$$

where $r := \frac{|\mathbf{R}|}{n}$.

In Chapter 3, section 3.1, the concentration parameters of phase data \mathbf{P} corresponding to two different sound stimuli (soft vs. aversive) are investigated. The variations of the concentration parameter of a von Mises distribution over different trials at different times are shown to be an indicator for the state of the habituation process in the subject.

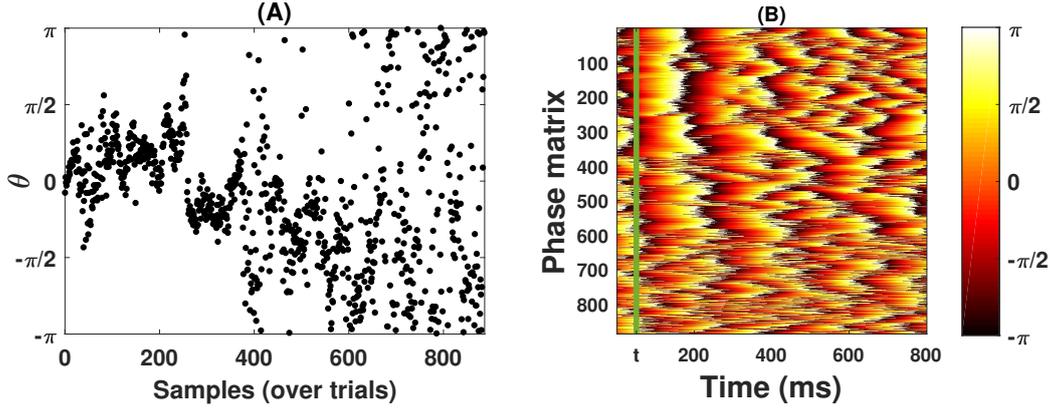


Figure 2.7: An example of phase matrix P in (B) and the variations of phase samples over all trials (denoted as θ) at a specific time t that has been shown in (A).

2.3.3 Hidden Markov Models & Change Detectors During the Course of Long-term Habituation

The von Mises modeling of IP for tracking the changes of long term habituation in section 2.3.2 (Parametric Modeling using a von Mises distribution) is improved by employing a Bayesian Hidden Markov model (HMM). In chapter 3, section 3.2 the limitations and drawbacks of using a windowed estimation of IP in EEG have been fully described. In short, the lack of ability to *track the gradual changes in the habituation process with a higher temporal resolution* can be considered one of the main disadvantages of the model. A model that reflects instant changes in the long-term habituation can be used in real-world applications to detect the level of habituation online and the state of the loudness comfort level over a long period of time.

To improve upon the previous approach in section 2.3.2, we employed a HMM model for modeling the circular data. To the best of our knowledge, no studies had considered the effectiveness of modeling the phase information in the context of long-term habituation using a HMM approach. A HMM is a *sequence model* that *assigns a label or class* to every unit in a *sequence of observations*. More generally, a HMM maps the sequence of observations to a sequence of labels or classes and have been extensively used for modeling time series throughout different domains. Examples include natural language processing, data compression, computational molecular biology and neuroscience for decoding neural processes at different scales. The main objective using a HMM is to *represent a probability distribution over sequences of observations*.

A HMM is specified by a set of *states* $\mathbf{S} = \{s_1, s_2, \dots, s_T\}$ and observations $\mathbf{X} = \{x_1, x_2, \dots, x_T\}$. To illustrate the dependency between the states and observations, a *graphical representation* is used in which the observation and states are considered as random variables. These are known as *graphical models*, in which based on *a priori* information of the problem, the conditional dependencies between random variables can be simplified and defined. In Fig. 2.8 we show an example of a first order HMM, a specific form of a Bayesian network.

The main assumptions in a HMM are as follows:

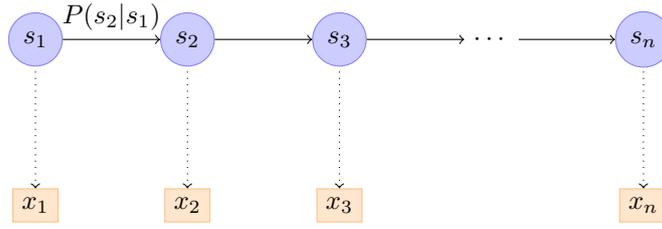


Figure 2.8: Illustration of a HMM with states s and the observations x . The transition between the states are shown as edges between the states $P(s_{k+1}|s_k)$ and the observation likelihood given the state is illustrated as the dotted edges. This is expressed as $P(x_i|s_i)$ (i.e., the likelihood of observing x_i given the state s_i .)

1. The observation at time t was generated by some process whose state s_t is hidden from the observer.
2. The states satisfy the *first-order Markov property*. That is given the value at s_{t-1} , the current state s_t is independent of all states prior to $t - 1$, that is $P(s_t|s_1, \dots, s_{t-1}) = P(s_t|s_{t-1})$.⁴
3. The observations are also conditionally independent given the current state s_t . That is $P(x_t|x_{1:t-1}, s_t) = P(x_t|s_t)$ (the Markov property of data with respect to the states).

Given the set of Markov properties the joint distribution between the states and observations is factorized as

$$P(S_{1:T}, X_{1:T}) = P(s_1)P(x_1|s_1) \prod_{t=2}^T P(s_t|s_{t-1})P(x_t|s_t) \quad (2.18)$$

From the decomposition, to define a probability distribution over the observations, the following set of parameters have to be defined:

- *The prior probabilities π* : $\pi_i = P(s_i), i \in \{1, \dots, T\}$ is the probability of state s_i being the first state in the state sequences. In many cases, the prior probability over states are defined uniformly, that is $P(s_i) = 1/T$. The vector $\pi = \{P(s_1), P(s_2), \dots, P(s_T)\}$ corresponds to the prior information over states s_i .
- *The transition probabilities*: The probability of transiting to state s_{t+1} from the current state s_t . This is expressed in the decomposition formula as $P(s_t|s_{t-1})$.
- *The emission probabilities*: define the likelihood of observing the observation x_t at time t , given the current state s_t . This is characterized as $P(x_t|s_t)$.

In the case of continuous observation, a common feature in many applications, $P(x_t|s_t)$ is usually modeled by a set of probability density functions that are defined over the continuous observations

⁴Similarly, the *n*th order property means that the current state is independent from all other states prior to $t - n$.

given the state at time t to be s_t (Ghahramani, 2001). One of the most common forms for modeling the emission probability is by parameterizing in the form of a Gaussian or a Mixture of Gaussians. Using the three main parameters, we are able to learn a probability distribution over the states and observations. In (Rabiner and Juang, 1986) the three main problems of the HMM are fully described. A comprehensive review of HMM and Bayesian network is presented by Ghahramani (2001). The models presented in the following sections are variants of Bayesian graphical models, equivalent to a HMM. The models are used for learning the *underlying model or hidden states* that generated the data as well as detecting the changes in the data regimes.

2.3.3.1 Tracking changes in the IP of long-term habituation processes

Using the terminologies in section 2.3.3, we model *abrupt changes* in the long-term habituation paradigm. The objective is to estimate the time range in which the habituation process occurs. Fig. 2.9 is an illustration of a time-series with abrupt changes in the mean. The data consists of three different data regimes or partitions, namely g_1, g_2, g_3 with equal number of samples in each regime. A *change point algorithm* identifies times of abrupt variations in the parameters of the underlying generative model.

One of the characteristics of long-term habituation signals as described in Chapter 2, Section 2.2 is that the N1-P2 component of the IP of auditory ERPs enters a precise phase locking mode. As in Fig. 2.7, the phase information enter a phase-locked mode with respect to the repetition of a stimulus in the state of focused attention. With adaptation to the stimulus, attention tends to trail away (long-term habituation), characterized by changes in the phase signature, and become more diffuse across trials. We therefore propose a Bayesian change point algorithm based on the model presented by Paquet (2007) and Adams and MacKay (2007) to be able to detect such transitions in the signal and label the signal as either a habituation or non-habituation signal.

In the next section, we describe the main terminologies of the change point model and data acquisition experiment for testing the algorithm.

2.3.3.2 Discrete Forward-Backward Bayesian Change Point Model

We first describe the main terminologies of the forward-backward Bayesian change point algorithm from Adams and MacKay (2007) for detecting regime changes in the generative process underlying the given directional time series data. A *change point* divides the set of data points $\Theta := \{\theta_1, \dots, \theta_N\}, \theta_i \in [-\pi, \pi)$ into non-overlapping partitions g . This is also known as the *product partitions* (Barry and Hartigan, 1992). The data points in each partition are assumed to be independently and identically distributed (i.i.d) from a corresponding probability distribution $p(\theta_t|\lambda_g)$ with parameters λ_g . Every partition is also called a *segment*. Fig. 2.9 is an example of a time series with abrupt changes.

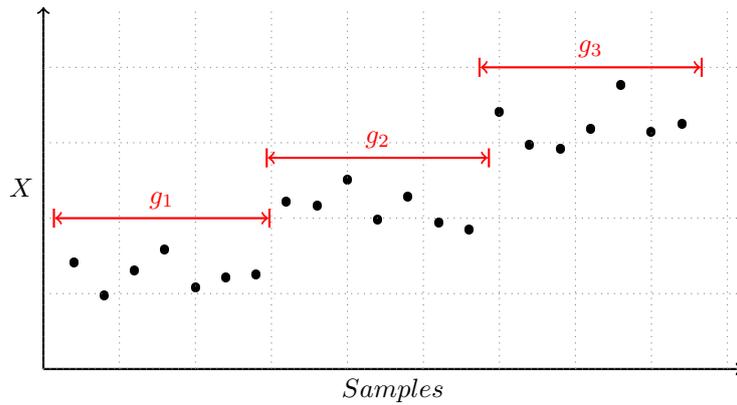


Figure 2.9: An illustration of a time-series with abrupt change points. The data points are separated into three partitions (specified by g representing the gap size between each section). The change points are defined as abrupt changes in the mean of the data. In real data measurements, the change points may not be as simply visible due to noise factors.

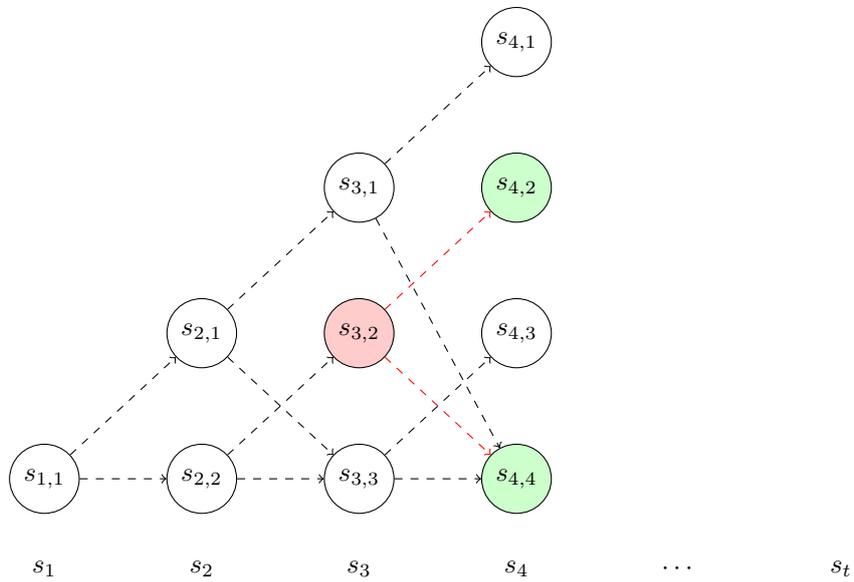


Figure 2.10: The trellis structure of run-states over time. Every run-state is denoted by $s_{t,i}$ or $s_t = i, \forall i \in \{1 \cdots t\}$. Each state takes values of $s_t \in \{1, \cdots t\}$. Given a state $s_{t,i}$, we could either increase by one time step, indicating no change point or transfer to a change point state. Given the state $s_{3,2}$, we can either transit to $s_{4,2}$ or a change point state $s_{4,4}$.

Run state The states of the model are described by the *latent* variables, called the *run state* in Paquet (2007). We denote the run state at time t as s_t , which is either increasing by one time step at every time t , or resets to zero, in which a change point at time t has occurred. The run state at a specific time and number is denoted as $s_{t,i}$ for $i = 1, \dots, t$. The case s_{tt} is the *change state* and other states $s_{t,i}$, for $i = 1, \dots, t-1$ correspond to the non-change states. See Fig. 2.10 for more clarification on the structure of a change point.

Run length For simplicity we denote the length of the current run state at time t as $s_t = i$, for $i \in \{1, \dots, t\}$. That is, at time t , the state s_t can take on any values between one and t . If it is at state t , then a change has occurred; otherwise, it is not a change point. The number of time steps since the last change point is called the *run-length*. The run-length will specify if the current data point θ_i has been generated by the model parameters of the previous regime, or a new change point has happened. The term $\theta_t^{(s_t=i)} = \theta_{i:t}$ indicates the set of observations associated with the run-length $t - i + 1$.

At every time step t , only one change point can occur (i.e., only one of the following states $s_{t,i}$ is on or equal to one, if $i = t$). Given only $s_{t,t} = 1$ for representing a change point at time t , we have $\sum_{i=1}^t s_{ti} = 1$. Figure 2.10 is an illustration of the run length network over time. At time t , the probability of being at any of the run-state is specified by s_t .

We assume a *a priori distribution* over the interval length between change points known as the *hazard rate* (See (Adams and MacKay, 2007; Wilson et al., 2010)). In other words, the hazard rate indicates the *rate of changes* in a time series. The choice of prior is important for the detection of change points and can be defined as a function of the run-length. We will describe in Chapter 3, Section 3.2, the choice of the hazard rate given the specific problem statement.

Given the model terminologies above, our aim is to *estimate the posterior distribution over the states at each time t* . See Fig. 2.10 for a clearer description of the notations. The marginal density over the *run-length* or the posterior probability using the Bayes's equation is

$$\begin{aligned}
 P(s_t = i | \Theta) &= \frac{P(\Theta | s_t = i) P(s_t = i)}{P(\Theta)} \\
 &= \frac{P(\Theta_{t+1:T} | \Theta_{1:t}, s_t = i) P(\Theta_{1:t} | s_t = i) P(s_t = i)}{P(\Theta)} \\
 &= \frac{P(\Theta_{t+1:T} | \Theta_{1:t}, s_t = i) P(s_t = i, \Theta_{1:t})}{P(\Theta)}
 \end{aligned} \tag{2.19}$$

for i between 1 and t .

According to (Paquet, 2007), Eq. 2.19 can be written in terms of the *forward-pass* and *backward-pass* and be solved using a recursive back-propagation technique ((Paquet, 2007)). The reformulation is

$$p(s_t = i | \Theta) = \frac{\alpha_i^t \beta_i^t}{p(\Theta)} \tag{2.20}$$

with the forward pass (α -pass) defined as $\alpha_i^t := p(s_t = i, \Theta_{1:t})$ and the backward pass (β -pass) as $\beta_i^t := p(\Theta_{t+1:T} | s_t = i, \Theta_{1:t})$. In the following, we explain the forward and backward passes as proposed in Adams and MacKay (2007), and the choice for the prior distribution. For more details see also Paquet (2007).

The Forward Pass The α -pass is based on the forward model presented by Adams and MacKay (2007) and is solved recursively using a *message passing* approach. By factorization and marginalization over s_{t-1} , the joint probability density between the states and observations is

$$P(\Theta_{1:t}, s_t = i) = P(\theta_t | \Theta_{1:t-1}, s_t = i) \sum_{j=1}^{t-1} P(s_t = i | s_{t-1} = j) P(s_{t-1} = j, \Theta_{1:t-1}). \quad (2.21)$$

The details of the factorization and simplifications are given in Appendix. D. The term $P(\theta_t | \Theta_{1:t-1}, s_t = i)$ is a prediction of the current data point θ_t using the data samples corresponding to the state $s_t = i$. To solve Eq. 2.21, there are two different transitions from the previous time step to the current one. If the current state at time t is not a change point, that is, $s_t = i$ for $i = \{1, \dots, t-1\}$, then only one transition with a probability of $(1 - H)$ from state $s_{t-1} = i$ to $s_t = i$ is possible. In this case Eq.2.21 can be simplified as

$$\begin{aligned} \alpha_i^{(t)} &= P(s_t = i, \Theta_{1:t}) = P(\theta_t | s_t = i, \Theta_{1:t-1}) \sum_{j=1}^{t-1} P(s_t = i | s_{t-1} = j) P(s_{t-1} = j, \Theta_{1:t-1}) \\ &= P(\theta_t | \Theta_{1:t-1}, s_t = i) (1 - H) \alpha_i^{(t-1)} \end{aligned}$$

The second state corresponds to $s_{t,i} = s_{t,t}$, that is, the case of a change point at time t . As in Fig. 2.11, from any of the states at the previous time $t-1$, we can have a transition to the new state of a change point at time t . Therefore we can simplify Eq. 2.21 as the sum of all the transitions from the previous state to the current state:

$$\alpha_t^{(t)} = P(\theta_t | \Theta_{1:t-1}, s_t = i) \sum_{j=1}^{t-1} H \alpha_j^{(t-1)}. \quad (2.22)$$

The Backward Pass Similar to the forward pass, the β -pass (backward-pass $\beta_{s_t}^t$) is solved using the same method. Given the state at time $t+1$, there are two transitions that can occur from the previous state $s_t = i$. We could either have a change point from the state i to the next state, or increase by one time step to state j at time $t+1$. The other transitions that have not been defined are equal to zero. Therefore, we can write $\beta_i^{(t)}$ in the following form:

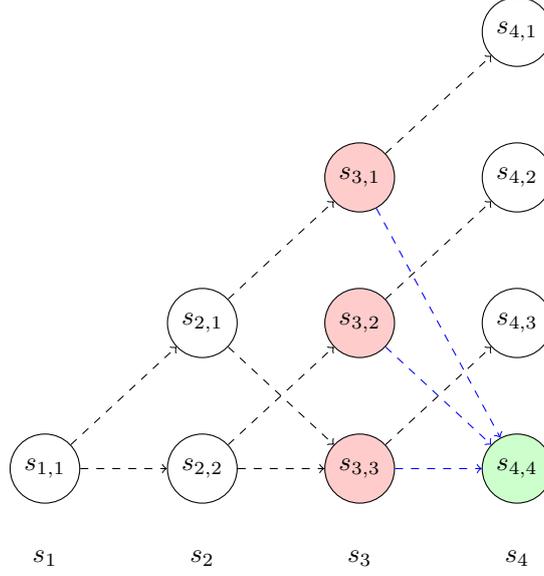


Figure 2.11: From any of the previous states at time $t - 1$, we can transit to a new change point state at time t . This has been shown as the blue arrows at time $t = 3$ from different states to a change point state at time $t = 4$.

$$\begin{aligned} \beta_i^{(t)} &= \sum_{j=1}^{t+1} P(\Theta_{t+2:T} | \Theta_{1:t+1}, s_{t+1} = j) P(\theta_{t+1} | \Theta_{1:t}, s_{t+1} = j) P(s_{t+1} = j | s_t = i) \\ &= \beta_i^{(t+1)} P(\theta_{t+1} | \Theta_{1:t}, s_{t+1} = j) (1 - H) + \beta_{t+1}^{(t+1)} P(\theta_{t+1} | \Theta_{1:t}, s_{t+1} = j) H \end{aligned}$$

The Choice of Parameters We have chosen a *von Mises*(μ, κ) probability density function (pdf) for the predictive distribution $p(\theta_t | s_t = i, \Theta_{1:t-1})$, where $\mu \in [-\pi, \pi]$ and $\kappa > 0$ are respectively the mean and concentration parameters for the relevant data $\theta_t^{(s_t=i)} = (\theta_{i:t})$. As described in the first study, the von Mises distribution is one of the most popular parametric models for the analysis of *circular data* that resembles a normal distribution of real-valued data.

The second most important parameter is the choice of the *hazard rate*, or prior information on the expected number of change points. A conditional prior on the change points, the *hazard rate* $H(x)$, is incorporated as

$$p(s_t | s_{t-1}) = \begin{cases} H(s_{t-1} + 1) & \text{if } s_t = t \\ 1 - H(s_{t-1} + 1) & \text{if } s_t = s_{t-1} \\ 0 & \text{otherwise,} \end{cases}$$

with $H(x)$ formulated according to (Adams and MacKay, 2007) as

$$H(x) = \frac{P_{gap}(g = x)}{\sum_{t=x}^{\infty} P_{gap}(g = t)}. \quad (2.23)$$

where g represents the data partitions. We use a geometric distribution with a probability $P(g = x) = (1 - p)^{x-1}p$ where the probability of having a change point at x is p . Hence, the hazard rate according to Eq.2.23 is $H(x) = p$. The expected value of a geometrically distributed r.v. g is given by $E[g] = \frac{1}{p}$. Therefore, the hazard rate is defined by a constant that represents the average number of samples before we observe a change point.

Finding the Maximum Sequence of States We have used the forward-backward method to estimate the posterior distribution over states $p(s_t = i | \Theta) \forall i \in \{1, \dots, t\}$ at different times t . In this section, we derive a method to determine the single most likely and consistent sequence of run-lengths over the whole time span $t = 1, \dots, N$ of the observation sequence. To extract the sequence of states, we apply the Viterbi algorithm analogously applying the forward-pass α . We define $V_{t,k}$ to be the maximum probability of being in state k at time t given the observation sequence $\Theta_{1:t}$. We initialize the states for the first data observation as $V_{1,1} = 1$ and $V_{1,k} = 0, (\forall k, k \neq 1)$. The transition probability from the current time t and state k to the next state can be written as

$$\begin{cases} V_{t+1,k} = (1 - H) V_{t,k} P(\theta_{t+1} | \Theta_{k:t+1}) & \text{for transition into non-change state} \\ V_{t+1,t+1} = H(\max_{k \leq t} (V_{t,k})) P(\theta_{t+1} | \Theta_{k:t+1}) & \text{for transition into change state} \end{cases} \quad (2.24)$$

By applying the procedure for time T , we can recursively extract the most likely state sequence transitions. This is done by taking the maximum state for time $V_{T,k}, \forall k$ and tracing back the state with the highest probability at the previous time.

2.3.3.3 Experimental Setting: Data Acquisition

The proposed discrete change point detection method is then used for identifying the changes in the long-term habituation process. The same data as in the previous study has been used. For more details see 2.3.2. Results and discussions on applying the method on data are explained in chapter 3, section 3.2.

2.3.4 Continuous Forward-Backward Change Point Model

The abrupt change point detection method presented in section 2.3.3 is not capable of giving a good estimate of temporal changes in the long-term habituation process. Despite observing a significant difference between the run-length of a lower sound stimulus (soft stimulus) in comparison to a louder stimulus (an aversive sound stimulus), the temporal subtle changes in the habituation process are not identified. It is more difficult to apply this method to mid-range loudness intensities, where changing to a habituation state may not be an abrupt transition. See the results and discussions of the previous approach in chapter 3 (Section 3.2) and chapter 4 (section 4.2) respectively.

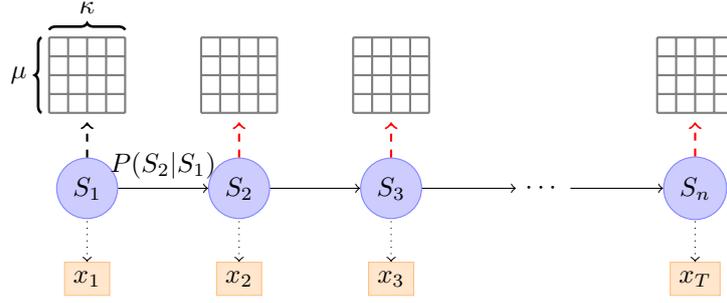


Figure 2.12: Illustration of a HMM with states illustrated as S and the observations as x . The transitions between the states are shown as edges between the states $P(S_{k+1}|S_k)$ and the observation likelihood given the state, expressed as $P(x_i|S_i)$, is illustrated as dotted black lines. In our model the observation sequence is the phase angles $\{\theta_1, \theta_2, \dots, \theta_T\}$ and every state S_t is a discretization of the parameter space of μ and κ .

In order to overcome the limitations of the discrete change point detection method in section 2.3.3.2, we use a different approach to detect habituation process. In the new method we first rely on the general definition of long-term habituation process as a *gradual* process. This implies the transition into a habituation state can be subtle, and hence requires more appropriate tools than an abrupt change point detector. As we are interested in detecting the temporal gradual changes over time, we use a discretized state space model in which the likelihood of all possible values of the random variables at every state is computed. This allows the system to integrate information efficiently over time and space and to propagate information from one stage to another without having to draw concrete conclusions at early stages (Knill and Pouget, 2004).

In our Bayesian model, the sequence of phase observations is defined over phase trials at a specific time t , more precisely $\Theta = \{\theta_1, \theta_2, \dots, \theta_N\}$, $\theta_i \in [-\pi, \pi)$. The data is modeled by assuming that at each time step, θ_i was generated from a certain but unknown state variable S_t . Note that the observation sequence corresponds to the phase modulations over trials at a specific time as illustrated in Fig. 2.7. Here we define $S_t = (\mu_t, \kappa_t)$ by the parameters of a von Mises distribution, where μ_t and κ_t are the mean and concentration parameter respectively. For simplicity, we represent this state on an equidistant grid of discrete values $R_\mu = \{u_1, \dots, u_m\}$, $u_i \in [-\pi, \pi)$ and $R_\kappa = \{k_0, \dots, k_m\}$, $k_i \in [0, \ell]$ where ℓ is the upper-bound of the concentration discretization. We have a HMM with a structure as presented in Fig. 2.12.

We assume that we have prior information on how the states should evolve given the past states $p(S_{t+1}|S_1, \dots, S_t)$, but cannot observe them directly. The available information is the measurements that are dependent on the state with noise: $p(\theta_t|S_t)$.

Our goal is to infer the distribution over hidden states (μ_t, κ_t) based on the phase data from all N trials, $\theta_{1:N}$. Given the conditional independence in our model, this can be represented as:

$$p(\mu_t, \kappa_t | \theta_{1:N}) \propto p(\theta_{t+1:N} | \mu_t, \kappa_t) p(\mu_t, \kappa_t | \theta_{1:t}), \quad (2.25)$$

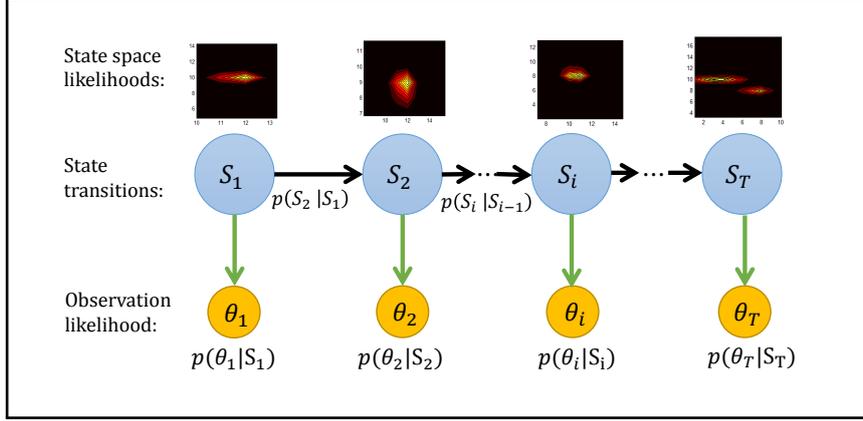


Figure 2.13: The graphical representation of the forward Bayesian model. The state space likelihood in our study are the mean μ and concentration κ parameters of a von Mises distribution. More precisely $S_i = (\mu_i, \kappa_i)$.

where $p(\mu_t, \kappa_t | \theta_{1:t})$ denotes the distribution over states at time t given information in the past and $p(\mu_t, \kappa_t | \theta_{t+1:N})$ is the distribution over the states given information in the future. The details of the derivation is included in Appendix. E. These distributions can be computed using recursive algorithms that sweep forwards (for $p(\mu_t, \kappa_t | \theta_{1:t})$) and backwards (for $p(\mu_t, \kappa_t | \theta_{t+1:N})$) through the data. We describe these sweeps in the following sections.

The forward sweep computes $p(\mu_t, \kappa_t | \theta_{1:t})$, the distribution over the state at time t , given data in the past. Based on the dependencies outlined in Fig. 2.13, this can be computed recursively using Bayes rule as

$$\begin{aligned} p(\mu_t, \kappa_t | \theta_{1:t}) &\propto p(\theta_t | \mu_t, \kappa_t) p(\mu_t, \kappa_t | \theta_{1:t-1}) \\ &= p(\theta_t | \mu_t, \kappa_t) \sum_{\mu_{t-1}, \kappa_{t-1}} p(\mu_t, \kappa_t | \mu_{t-1}, \kappa_{t-1}) p(\mu_{t-1}, \kappa_{t-1} | \theta_{1:t-1}) \end{aligned} \quad (2.26)$$

where $p(\theta_t | \mu_t, \kappa_t)$ is the likelihood of the data given the parameters and $p(\mu_t, \kappa_t | \mu_{t-1}, \kappa_{t-1})$ defines our prior on the dynamics of μ_t and κ_t .

The backward pass computes the probability of the future data given the state at time t , $p(\theta_{t+1:N} | \mu_t, \kappa_t)$. As for the forward pass, this can be computed recursively using Bayes rule

$$\begin{aligned} p(\theta_{t+1:N} | \mu_t, \kappa_t) &= \sum_{\mu_{t+1}, \kappa_{t+1}} p(\mu_{t+1}, \kappa_{t+1} | \mu_t, \kappa_t) p(\theta_{t+1:N} | \mu_{t+1}, \kappa_{t+1}) \\ &= \sum_{\mu_{t+1}, \kappa_{t+1}} p(\mu_{t+1}, \kappa_{t+1} | \mu_t, \kappa_t) p(\theta_{t+1} | \mu_{t+1}, \kappa_{t+1}) p(\theta_{t+2:N} | \mu_{t+1}, \kappa_{t+1}) \end{aligned} \quad (2.27)$$

where $p(\theta_{t+1} | \mu_{t+1}, \kappa_{t+1})$ is the same likelihood and $p(\mu_{t+1}, \kappa_{t+1} | \mu_t, \kappa_t)$ the same prior as used in the forwards pass.

The benefit of using a backward pass and its effect on removing initialization problem has been described in the result's section of chapter 3, section 3.3.1.

2.3.4.1 State Transition Model, $p(\mu_t, \kappa_t | \mu_{t-1}, \kappa_{t-1})$

For the prior distribution over the state transitions $p(\mu_t, \kappa_t | \mu_{t-1}, \kappa_{t-1})$, we first assume that the transition distributions for the mean, μ , and concentration, κ , are independent; i.e.,

$$p(\mu_t, \kappa_t | \mu_{t-1}, \kappa_{t-1}) \propto p(\mu_t | \mu_{t-1}) p(\kappa_t | \kappa_{t-1}).$$

For transition distributions of μ and κ we choose von Mises and Gaussian distributions, respectively. More specifically, for the circular parameter, $\mu_t \in [-\pi, \pi)$ we assume a von Mises distribution with the mean of μ_{t-1} and the concentration of $K \in \mathbb{R}^+$; i.e.,

$$p(\mu_k | \mu_{k-1}, K) = \frac{\exp(K \cos(\mu_t - \mu_{t-1}))}{2\pi I_0(K)}.$$

For the real valued concentration random variable, we assume a Gaussian transition distribution,

$$p(\kappa_t | \kappa_{t-1}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(\kappa_t - \kappa_{t-1})^2}{2\sigma^2}\right)$$

with variance σ^2 .

As the states are discretized for computing $p(\mu_t | \mu_{t-1})$ and $p(\kappa_t | \kappa_{t-1})$, we evaluate the von Mises distribution and Gaussian distribution for all discrete values of μ_t and κ_t and then re-normalize the values. The selection criteria for the prior parameters K and σ^2 and optimization are all described in Chapter 3, Section 3.3.2).

2.3.4.2 Experimental Setting

To have a better understanding of the temporal changes in the long-term habituation process, we use a different range of sound stimuli to test the model. That is, instead of using a soft and aversive stimuli, we employed a range of different stimuli in between in order to track the differences between state transitions as the intensity of the sounds became higher. Twenty participants (16 female and 4 male; mean age: 23 years and 3 months with a standard deviation of 4 years and 1 month) attended the experiment. One measurement (female) was discarded due to data corruption. We use the remaining 19 subjects for our study. All participants had normal hearing as assessed by an audiogram test before and after the experiment. All subjects provided informed consent and the study was conducted in accordance with the declaration of Helsinki.

For the electrophysiological experiment, subjects lay on an examination bed and were instructed to relax with closed eyes. The EEG signals were recorded using surface electrodes (Ag/AgCl) which were placed at the right and left mastoid (active), the vertex (reference) and the upper forehead (ground). The signal acquired from each mastoid was referred to vertex (Cz) and processed separately. The electrode impedance was kept below $5k\Omega$ and the recording EEG was sampled at

512Hz.

The experiment consisted of listening to a series of *puretone beeps* presented in the subjects' right ear via headphones (Sennheiser HDA 200) at four different volumes of 60, 70, 80 and 90dB SPL. Each beep was at a frequency of 1kHz and lasted 40ms. There was a 1s ISI between stimuli and at least 500 tones were presented per stimulation level. Participants were instructed not to pay attention to the stimulus and attempt to ignore the sound during the experiment. To be sure they did not fall asleep, they were constantly monitored.

We analyzed the recorded ERPs over different times for the effect of long-term habituation and distinguishing between different stimuli in terms of the underlying generative process.

2.4 Oscillatory EEG Signals and IP Extraction

Up until this point, we introduced different methods to measure the level of long-term habituation and non-habituation for auditory ERPs. We were able to *objectively* classify between the *attended state* and the drift to the *non-attended (habituation)* state using the neural correlates ((Mortezapouraghdam et al., 2016, 2015a, 2014)). More successfully, we were able to distinguish between the two different states of habituated and non-habituated states with different types of stimuli ((Mortezapouraghdam et al., 2016)). Although the study of ERPs using pure tone stimuli is important in obtaining useful information to decode the behavior of the underlying population of neurons, in nature and daily environments, we rarely encounter pure tone stimuli. Most of the sounds around us are complex and composed of many different components (i.e., superposition of many sine waves with different frequencies and various delays). Additionally, we are exposed to much longer and continuous auditory stimuli. Hence, the *detection of attention* and its drift to a *habituation* state using the neural correlates will be more complicated in such examples as a more complex and rich stimuli content is being used.

To measure the level of attention over a long duration of time, a continuous EEG signal is essential. One of the goals in analyzing the ongoing EEG and objectively determining the level of attention is improving the performance of hearing aid devices. Different approaches and experiments are applied to adjust the hearing aid settings, however not much scrutiny is given to improvement of hearing aid devices using the *endogenous* effects of listening with respect to different stimuli. In some previous studies (See (Bernarding et al., 2013a, 2016) and the references therein), extensive research has been conducted to measure the level of *listening effort* and *objective hearing aid fitting procedures* with regard to different stimuli by decoding the endogenous signatures of EEG signals. These endogenous parameters are then compared against the performance of participants to understand any existing correlations between the measures and the performance and listening effort of the subject. The improvements in this area can lead to better development of hearing aid devices, in which the adjustments of the device can be made based on an individuals' neural responses.

In some of the previous studies, the phase reorganization of signals is used as the main feature to objectively measure the level of attentional effort in patients ((Bernarding et al., 2013a, 2016; MortezaPouraghdam et al., 2017)). Phase synchronization techniques are among the main analysis tools that measure the degree of phase reorganization in signals. The phase reorganization approach, the phase alignment in stimulus-induced potentials aids decoding cognitive processes. Despite the developments in this area, the precision of the measures may vary based on the level of noise. Phase distortion, due to artefacts and pre-processing steps can introduce *artificial phase resets* or alignment. These artificial phase resets that is referred to as spurious phase resets, are not correlated to any stimulus or event.

In many studies, the effect of phase synchronization is analyzed without considering the variations in the amplitude of the signal. However, in a recent study presented by Sameni and Seraj (2017), it has been shown that a low instantaneous envelope lead to abrupt changes in IP and could distort the results of phase synchronization. Therefore, the distortions that appear as abrupt jumps in IP are not related to any underlying physiological effects, but instead are noisy variations due to low envelope of the signal. They present a more reliable method of computing IP using the Monte-Carlo simulation technique. As a continuation to the proposed approach, we present a model that is able to *remove the spurious phase variation* from the phase information of signals as well as determine a standard deviation for the estimated phase. The following de-noising technique aims to provide a better estimate of the level of phase-synchronization with regard to different stimuli.

In the following sections, we describe the problem of spurious phase variation, removal the spurious variations and application on synthetic and ongoing EEG measurements. The level of the attention is computed by assessing the phase reorganization of signals.

2.4.1 Phase Singularities: Definition of Spurious Phase Slips and Types

Before the extraction of IP, EEG data is usually pre-processed. To observe the oscillations at a specific frequency, a narrow bandpass filter is applied. A FIR (finite impulse response) filter is usually applied to avoid distortion of phase information (see chapter 5 of (Handy, 2005) for more references). The IP is obtained by computing the analytic signal, commonly by either applying a wavelet transformation (WT) or Hilbert transformation (HT). The HT of a signal $f(t)$ is represented as

$$\hat{f}(t) = f(t) + i\mathcal{H}(f(t)) = f(t) + iu(t).$$

Measures such as the IE and IP are obtained from $A_t = \sqrt{f(t)^2 + u(t)^2}$ and $\theta(t) = \tan^{-1}\left(\frac{u(t)}{f(t)}\right)$ respectively, where $\theta \in [-\pi, \pi)$, $t \in \mathbb{R}$. The IP resets and drops to $-\pi$ when it reaches π . One of the techniques for tracking the modulations of IP in long durations of time is analyzing the unwrapped version of the IP. The unwrapped IP is obtained by adding 2π every time a reset occurs. We use $p : \mathbb{R} \rightarrow \mathbb{R}^+$ for unwrapped phase.

The slope of $p(t)$ is related to the *mean frequency*. The slope of a signal with center frequency f_c which best determines the activities at that particular frequency is given by $\omega_0 = 2\pi f_c$. If the IP contains no additional resets, it is uniformly distributed. This can be easily illustrated by a sine wave with complete cycle, at which the IP is uniformly distributed. Hence the change to the uniformity can be represented by the difference between $p(t)$ and the line $\omega_0 t$, i.e., $r(t) = p(t) - \omega_0 t$. See Fig. 2.14 which is an example of two sine waves with high and low number of resets along with $r(t)$.

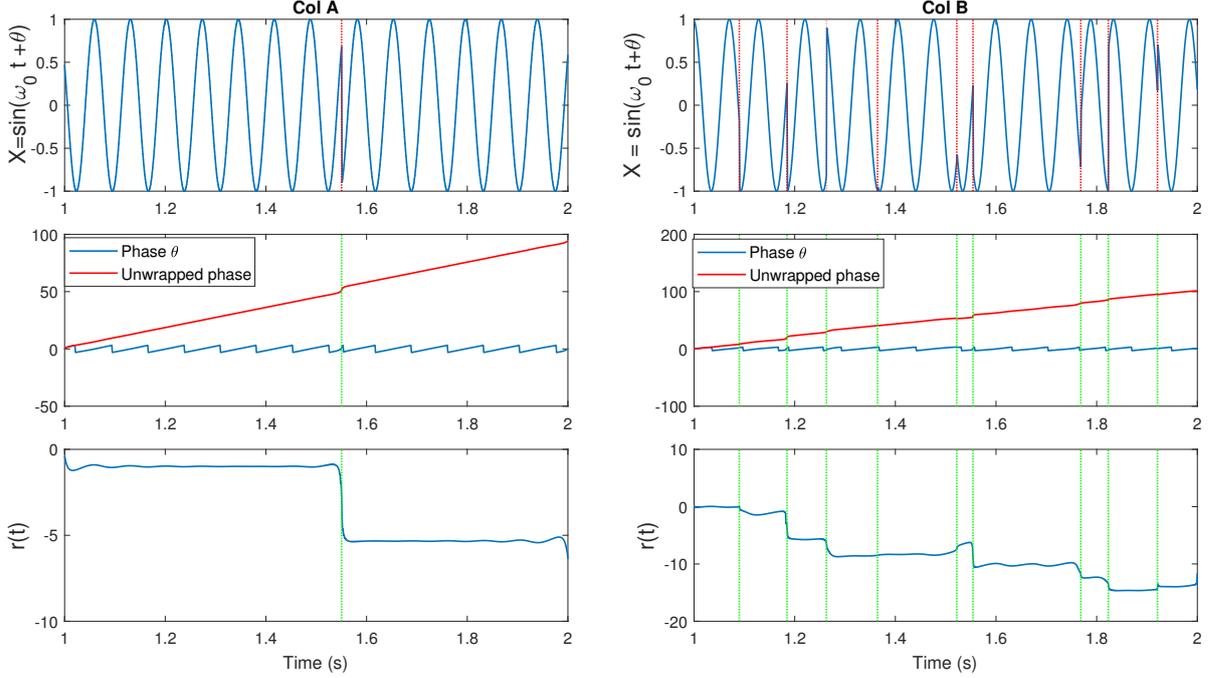


Figure 2.14: Col A - [top]: An example of a sine wave $X = \sin(\omega_0 t + \theta)$ with $f_c = 7$. We introduced one artificial shift in the signal. [Middle]: The unwrapped IP of X (denoted as $p(t)$) and the wrapped version $\theta(t)$. [Bottom]: The residual $r(t)$ is computed as $r(t) = p(t) - \omega_0 t$. The shift in the signal has been shown as an abrupt jump in the residual. Col B -[top]: An example of a sine wave with more random shifts in the signal. Corresponding phase and residual plots have been generated.

After applying a narrow band-pass filter to the measured EEG oscillations, the resulting x_t, θ_t can be described by

$$x_t = X_t \cos(\omega_0 t + \theta_t) + W_t = s_t + W_t$$

such that θ_t is same as $r(t)$ and $\omega_0 = \frac{2\pi f_c}{f_s}$. The noise or the background EEG is modeled by W_t . It is assumed that the noise is a Gaussian random process $W_t \sim N(0, \sigma^2)$. The IP θ_t ($\theta_t \in [-\pi, \pi)$) and the envelope X_t are slowly varying functions of t . Hence, it is expected that the changes between θ_t and θ_{t+1} to be small. In Sameni and Seraj (2017), it has been shown as the IE becomes small (near zero), the IP will contain sudden changes or jitters (see Fig. 2.15) that can be falsely correlated to phase-resets induced from an event.

We can describe the effect of spurious phase variation as shown in Fig. 2.16. Here are four samples of an analytic form of a signal. The distance between the first two samples $|a_1 - a_2|$ and the second

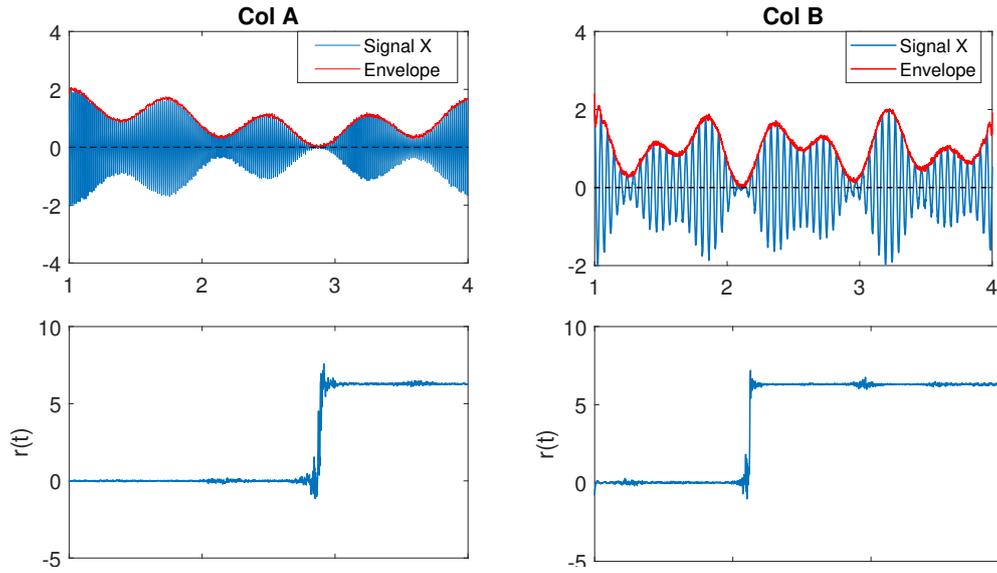


Figure 2.15: Col A, *first plot*: An example of an amplitude modulated signal with a mean frequency $f_c = 50\text{Hz}$. Red line shows the IE. The second plot shows $r(t)$. As the IE of the signal approaches to zero, the phase variations increase, best seen as an abrupt change. This is clearly evident in comparison to the other envelope values throughout the signal at other time samples. Col B, *first plot*: The same explanation as in Col A, however with a lower mean frequency signal $f_c = 15\text{Hz}$.

two samples $|b_1 - b_2|$ are the same. However, the envelope, which is the distance of the individual samples to the origin is different. Considering the pair a_1, a_2 , such that the estimated phase is $a_2 = x_t$ and the true value of phase is $a_1 = x_{t+1}$, the angle between these two, given the magnitude of the envelope, is about 13° . However, if the same samples with the same distance are spotted with low envelopes as in (b_1, b_2) , then the phase difference is approximately 116° . Another interpretation is that a_1 corresponds to the sample without any noise, and a_2 is the location of the sample after getting perturbed with noise. In this case, the difference in phase is not large if they have larger envelopes.

This effect has been studied in Sameni and Seraj (2017); Chavez et al. (2006); Rudrauf et al. (2006) where the effect of spurious phase jumps have been described in terms of calculation of the phase. The IP is computed using the arctan operator $(\tan^{-1}(\frac{u(t)}{x(t)}))$. Minor changes due to noise or background EEG variations to the real and imaginary parts of the analytic signal (which are narrow-bandpassed) can lead to significant changes to the computation of phase as the numerator and denominator values are small (see Rudrauf et al. (2006) and the references therein). Hence, the aim of this study is to estimate the *true phase* using the IE and IP and have a more reliable measure of the IP such that the spurious phase variations due to a low envelope will be removed.

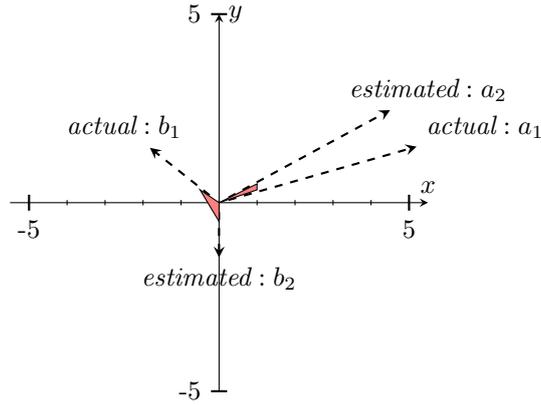


Figure 2.16: An example of a sudden phase change between b_1 and b_2 due to the low envelope in comparison to a_1 and a_2 . The distance of (a_1, a_2) and (b_1, b_2) are approximately the same, however their envelopes are different.

2.4.2 Modeling the Variations Between Instantaneous Phase and Amplitude

The analytic form of the signal $s_t \in \mathbb{R}$ is denoted as $\hat{s}_t = s_t + i\mathcal{H}(s_t)$, $\hat{s}_t \in \mathbb{C}$. We assume that the measurements are corrupted by Gaussian noise with mean 0 and variance α . The noise of the measurements is denoted by W_t . It's analytical form \widehat{W}_t is modeled using a symmetric complex Gaussian distribution $\widehat{W}_t \sim \mathcal{CN}(0, \alpha I)$. This can be shown as follows: as stated, the bandpassed signal is modeled as

$$x_t = X_t \cos(\omega_0 t + \theta_t) + W_t = s_t + W_t.$$

It's HT is denoted by

$$\begin{aligned} \mathcal{H}(x_t) &= \mathcal{H}(X_t \cos(\omega_0 t + \theta_t) + W_t) \\ &= X_t \sin(\omega_0 t + \theta_t) + \mathcal{H}(W_t) \end{aligned}$$

using the linearity of the HT. Here $\mathcal{H}(W_t)$ is also a Gaussian distributed random variable (see Chapter 8 of Davenport et al. (1958) and Sameni and Seraj (2017)) with mean 0 and variance α . The analytic form of the signal x_t using the HT is then

$$\begin{aligned} \hat{x}_t &= x_t + i\mathcal{H}(x_t) \\ &= X_t \cos(\omega_0 t + \theta_t) + W_t + i[X_t \sin(\omega_0 t + \theta_t) + \mathcal{H}(W_t)] \\ &= X_t [\cos(\omega_0 t + \theta_t) + i \sin(\omega_0 t + \theta_t)] + [W_t + i\mathcal{H}(W_t)] \\ &= X_t e^{i(\omega_0 t + \theta_t)} + [W_t + i\mathcal{H}(W_t)] \end{aligned} \tag{2.28}$$

where the analytic form of W_t is modeled by a complex Gaussian distribution $W_t + i\mathcal{H}(W_t) \sim$

$\mathcal{CN}(0, I\alpha)$ with a zero mean and independent real and imaginary part with σ^2 as the variance⁵. The resulting complex Gaussian distribution is a symmetric distribution with non-diagonal elements set to zero.

The analytic form of the measurements are hence modeled as $\hat{x}_t = \hat{s}_t + \widehat{W}_t$, where $\hat{s}_t \in \mathbb{C}$ is the analytic form of the signal without noise. We work directly with the complex analytic signals, as we can later obtain the denoised version of the IP to simplify our model. With our assumption of s_t , the observations can be modeled using a complex Gaussian distribution $\hat{x}_t \sim \mathcal{CN}(\hat{s}_t, \alpha)$, where \hat{s}_t is the mean and α is the variance. The signal s_t is filtered with a narrow bandpass filter centered around f_c prior to computing the analytic signal \hat{x}_t (see Eq.2.28). Thus we can assume that $\hat{x}_t = A_t e^{i(\omega_0 t + \theta_t)} + \widehat{W}_t$, using the linearity of HT. Here, $\hat{s}_t = A_t e^{i(\omega_0 t + \theta_t)}$ is the denoised analytical signal that we want to recover. The ratio of successive values of \hat{s}_t is

$$\begin{aligned} \hat{s}_{t+1}/\hat{s}_t &= (A_{t+1}/A_t) e^{i(\omega_0(t+1) + \theta_{t+1} - \omega_0 t - \theta_t)} \\ &= (A_{t+1}/A_t) e^{i(\omega_0 + (\theta_{t+1} - \theta_t))} \end{aligned} \quad (2.29)$$

$$\approx e^{i\omega_0} \quad (2.30)$$

The results are obtained based on the assumption that at a narrow frequency, A_t and θ_t are varying slowly compared to ω_0 . We can therefore express the phase modulations over time using our assumptions as $\hat{s}_{t+1} \approx e^{i\omega_0} \hat{s}_t$, and model $\hat{s}_{t+1} = e^{i\omega_0} \hat{s}_t + \hat{\eta}_t$, where $\hat{\eta}_t \sim \mathcal{CN}(0, \sigma)$. This means that the analytic signal at time $t+1$ is obtained using the phase at time t multiplied by a small factor of $e^{i\omega_0}$ with some additive noise. The additive noise η correspond to the simplifications that have been applied for obtaining Eq.(2.30). We assume that the additive noise follows a Gaussian distribution. The proposed model for the phase guarantees that the changes in phase of a narrow-bandpassed signal are rather slow and gradual over time. The model has a structure similar to a HMM, as in Fig. 2.12, in which the states s_t and ultimately phase values that we want to estimate are the analytic form of the signal.

2.4.2.1 Model Derivation

We describe the model derivation in two parts of forward and backward pass. As described previously, the model assumes the following:

$$\begin{aligned} \hat{s}_{t+1} &= e^{i\omega_0} \hat{s}_t + \hat{\eta}_t \text{ with } \hat{\eta}_t \sim \mathcal{CN}(0, \sigma) \\ \hat{x}_t &= \hat{s}_t + \widehat{W}_t \text{ with } \widehat{W}_t \sim \mathcal{CN}(0, \alpha) \end{aligned} \quad (2.31)$$

⁵For simplicity of notation, we write $\mathcal{CN}(a, b)$ instead of $\mathcal{CN}(a, Ib)$ in the rest of the paper.

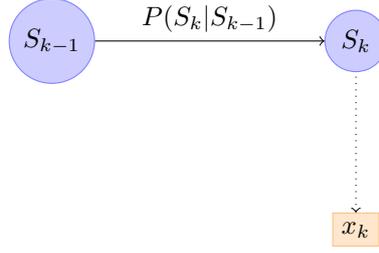


Figure 2.17: An illustration of part of a Bayesian Network corresponding to forward-passing of information.

with the following assumptions about the distributions of states and the data observations:

$$\begin{aligned}
 P(\hat{s}_1) &\sim \mathcal{CN}(\mu_1, p_1) \\
 P(\hat{s}_{t+1} | \hat{s}_t) &\sim \mathcal{CN}(e^{i\omega_0} s_t, \sigma) \\
 P(\hat{x}_t | \hat{s}_t) &\sim \mathcal{N}(\hat{s}_t, \alpha)
 \end{aligned} \tag{2.32}$$

We can write the above expressions as a two dimensional real linear Gaussian state-space model by separating the real and imaginary parts. In the following we take \hat{s}_t, \hat{x}_t to be two dimensional vectors of real numbers consisting of the real and imaginary parts of the underlying complex number. With this convention we have:

$$\begin{aligned}
 P(\hat{s}_1) &\sim \mathcal{N} \left(\begin{bmatrix} \Re(\mu_1) \\ \Im(\mu_1) \end{bmatrix}, \begin{bmatrix} p_1 & 0 \\ 0 & p_1 \end{bmatrix} \right) \\
 P(\hat{s}_{t+1} | \hat{s}_t) &\sim \mathcal{N} \left(\begin{bmatrix} \cos(\omega_0) & -\sin(\omega_0) \\ \sin(\omega_0) & \cos(\omega_0) \end{bmatrix} \hat{s}_t, \begin{bmatrix} \sigma & 0 \\ 0 & \sigma \end{bmatrix} \right) \\
 P(\hat{x}_t | \hat{s}_t) &\sim \mathcal{N} \left(\hat{s}_t, \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} \right)
 \end{aligned} \tag{2.33}$$

In this form we can directly apply the Kalman smoother that has been described in Briers et al. (2010). This consists of a forward and backward pass that we explain in the following two sections.

The Forward Pass To estimate the state (an estimation of the analytic form of the signal), we have to derive the following posterior distribution:

$$\begin{aligned}
P(\hat{s}_t|\hat{x}_{1:t}) &= \frac{P(\hat{x}_t|\hat{s}_t, \hat{x}_{1:t-1}) P(\hat{s}_t, \hat{x}_{1:t-1})}{P(\hat{x}_{1:t})} \\
&= P(\hat{x}_t|\hat{s}_t) \frac{P(\hat{s}_t|\hat{x}_{1:t-1})}{P(\hat{x}_t|\hat{x}_{1:t-1})}
\end{aligned}$$

by using the first order Markov property that the current data at time t is independent from the past given the state at time \hat{s}_t . Simplifying the normalization factor we can write:

$$\begin{aligned}
P(\hat{s}_t|\hat{x}_{1:t}) &\propto P(\hat{x}_t|\hat{s}_t) P(\hat{s}_t|\hat{x}_{1:t-1}) \\
&\propto P(\hat{x}_t|\hat{s}_t) \underbrace{\int P(\hat{s}_t, \hat{s}_{t-1}|\hat{x}_{1:t-1}) d\hat{s}_{t-1}}_{\text{Marginalizing over } \hat{s}_{t-1}}
\end{aligned}$$

after expanding the inner bracket, we obtain:

$$P(\hat{s}_t|\hat{x}_{1:t}) \propto P(\hat{x}_t|\hat{s}_t) \int P(\hat{s}_t|\hat{s}_{t-1}) P(\hat{s}_{t-1}|\hat{x}_{1:t-1}) d\hat{s}_{t-1} \quad (2.34)$$

Eq.2.34 can be realized as in Fig. 2.17. Fig. 2.17 shows a part of the first order Markov model. We can compute the distribution of $P(\hat{s}_t | \hat{x}_{1:t})$ recursively starting from $P(\hat{s}_0)$. It turns out that in our model the distribution of $P(\hat{s}_t | \hat{x}_{1:t})$ is always Gaussian so it suffices to compute its mean and covariance matrix. Writing μ_t, P_t for the mean and covariance matrix of $P(\hat{s}_t | \hat{x}_{1:t})$ we have the following equations from Briers et al. (2010).

$$\begin{aligned}
P'_t &= BP_t B^T + Q \\
K_t &= P'_t (P'_t + R)^{-1} \\
\mu_{t+1} &= B\mu_t + K_t(\hat{x}_t - B\mu_t) \\
P_{t+1} &= P'_t - K_t P'_t{}^T
\end{aligned}$$

Furthermore, it turns out that P_t is always diagonal of the form $p_t I$. This allows us to simplify the equations above. To see this, assume that $P_t = p_t I$ and plug this expression into the equations

above.

$$\begin{aligned}
P'_t &= B(p_t I) B^T + Q \\
&= p_t B B^T + \sigma I = (p_t + \sigma) I \\
K_t &= (p_t + \sigma) I ((p_t + \sigma) I + \alpha I)^{-1} \\
&= \frac{p_t + \sigma}{p_t + \sigma + \alpha} I \\
\mu_t &= B\mu_t + \frac{p_t + \sigma}{p_t + \sigma + \alpha} (\hat{X}_t - B\mu_t) \\
P_{t+1} &= \left(p_t + \sigma - \frac{(p_t + \sigma)^2}{p_t + \sigma + \alpha} \right) I
\end{aligned}$$

The *kalman gain factor* $\left(K_t = \frac{p_t + \sigma}{p_t + \sigma + \alpha} I \right)$ is an expression for determining how reliable the measurement at time $t + 1$ is compared to the estimated state (signal) based on the level of noise in the data. If data has a small level of SNR (high noise), the algorithm will rely more on the estimated value than the measurement. Therefore, a realistic and good estimate of SNR can improve the performance of the KS.

The Backward Pass Given $P(\hat{s}_t | \hat{x}_{1:t})$ we compute $P(\hat{s}_t | \hat{x}_{1:T})$ recursively starting from $P(\hat{s}_T | \hat{x}_{1:T})$ and moving backwards. In general we have

$$P(\hat{s}_t | \hat{x}_{1:T}) = P(\hat{s}_t | \hat{x}_{1:t}) \int \frac{P(\hat{s}_{t+1} | \hat{x}_{1:T}) P(\hat{s}_{t+1} | \hat{s}_t)}{P(\hat{s}_{t+1} | \hat{x}_{1:t})} d\hat{s}_{t+1} \quad (2.35)$$

and hence we can compute $P(\hat{s}_t | \hat{x}_{1:T})$ given $P(\hat{s}_{t+1} | \hat{x}_{1:T})$. Again, we find that $P(\hat{s}_t | \hat{x}_{1:T})$ is always a Gaussian distributed random variable. Writing $\bar{\mu}_t, \bar{P}_t$ for the mean and covariance matrix of $P(\hat{s}_t | \hat{x}_{1:T})$ we have the following equations from (Briers et al., 2010).

$$\begin{aligned}
\Gamma_t &= P_t B^T (P'_t)^{-1} \\
\bar{\mu}_t &= \mu_t + \Gamma_t (\bar{\mu}_{t+1} - B\mu_t) \\
\bar{P}_t &= P_t + \Gamma_t (\bar{P}_{t+1} - P'_t) \Gamma_t^T
\end{aligned}$$

As in the forward pass this simplifies in our model, since we only deal with symmetric covariance matrices of the form $\bar{p}_t I$ and since $B^T = B^{-1}$. In the following we assume that $\bar{P}_{t+1} = \bar{p}_{t+1} I$. We

have

$$\begin{aligned}\Gamma_t &= p_t I B^T ((p_t + \sigma) I)^{-1} = \frac{p_t}{p_t + \sigma} B^T \\ \bar{\mu}_t &= \mu_t + \frac{p_t}{p_t + \sigma} B^T (\bar{\mu}_{t+1} - B \mu_t) \\ &= \mu_t + \frac{p_t}{p_t + \sigma} (B^T \bar{\mu}_{t+1} - \mu_t) \\ \bar{P}_t &= p_t I + \Gamma_t (\bar{p}_{t+1} - p_t - \sigma) \Gamma_t^T \\ &= \left(p_t + \frac{p_t^2 (\bar{p}_{t+1} - p_t - \sigma)}{(p_t + \sigma)^2} \right) I\end{aligned}$$

The steps for applying a KS are shown in the algorithm below:

Input:

α : Variance of measurement noise

σ : Variance of signal $< \left(\frac{f_c}{f_s} \right)^2$

Output:

$\mu_{1:T}$: Estimated signal

$V_{1:T}$: Estimated variance

Forward Pass

$p_{1:T} = 0$ Initial variance estimate

$\mu_1 = \hat{x}_1$

$p_1 = \alpha$

for $t = 2$ to T **do**

$$\mu_t = e^{i\omega_0} \mu_{t-1} + \frac{p_{t-1} + \sigma}{p_{t-1} + \alpha + \sigma} (\hat{x}_t - e^{i\omega_0} \mu_{t-1})$$

$$p_t = \frac{\alpha(p_{t-1} + \sigma)}{p_{t-1} + \sigma + \alpha}$$

end for

Backward Pass

$V_T = p_T$

for $t = T - 1$ to 1 **do**

$$\mu_t = \left(1 - \frac{p_t}{p_t + \sigma} \right) \mu_t + \frac{p_t}{p_t + \sigma} (e^{-i\omega_0} \mu_{t+1})$$

$$V_t = p_t + \left(\frac{p_t}{p_t + \sigma} \right)^2 (V_{t+1} - p_t - \sigma)$$

end for

Algorithm 1: Algorithmic representation of a Kalman smoother (The forward and backward pass)

2.4.3 Neural correlates of Selective Attention & Ongoing EEG

The mapping between the segmented ERP and ongoing EEG measurement is a debatable topic that we extensively discussed in the first chapter. Regarding the generation of ERP and spontaneous EEG activities, we considered exclusively the role of the phase-reorganization to explore the neural

signature of selective attention in different experimental paradigms. We are interested in exploring the possibility of mapping the features obtained from the N1-P2 neural correlate of averaged ERPs with ongoing EEG measurements. However, this requires sophisticated methods to understand the activity of the ongoing EEG signals and validation of results with averaged ERP results. In various studies, the IP was used to assess the attentional effort related to the listening tasks (Strauss et al., 2008a, 2010; Bernarding et al., 2013b; Morteza pouraghdam et al., 2017). In particular it was shown that listening effort induced by task difficulty could be quantified using the IP of ongoing EEG measurements in various subject groups and listening tasks.

In this study, we analyze the effect of selective attention using ongoing EEG signals and compare the objective measures with the results obtained from the segmented and averaged ERPs. By doing so, we explore the possibility of translating and mapping the N1-P2 effect in ongoing oscillations and validate the results with ERP analysis. The details regarding the data is explained in the next section.

2.4.4 Experimental Setting

The experimental setting in this study is based on a former experimental paradigm of (Hillyard et al., 1973), in which two auditory tone burst sequences with different frequency contents are presented dichotically. Each sequence has standard tone bursts and deviant bursts that has a slightly different frequency and an appearance probability of 10%. The bursts have a duration of 50 ms and an amplitude window with a linear rise and fall time of 10 ms and plateau of 30 ms. The sequence presented to the right ear consisted of tone bursts of 1500 Hz and deviants of 1560 Hz, whereas those to the left ear have tone bursts of 800 Hz and deviants of 840 Hz. The ISIs between bursts have a randomized duration oscillating between 250-1250 ms. The order between standard and deviants was also randomized. The resultant auditory stimulation signals have a total of 560 bursts each.

Ten normal-hearing subjects (mean age: 26.6 ± 6.24 years, 6m/4f, right-handed) participated in the study. Normal-hearing thresholds are controlled by pure tone audiogram signals preceding the measurements. After a detailed explanation of the procedure, the subjects signed an information consent form. Next, Ag/AgCl electrodes (Schwarzer GmbH, Germany) were attached at the right (M2) and left (M1) mastoids, at the vertex (Cz) and upper forehead (Fpz) for the common reference and ground, respectively. Electrode impedances are below 5 k Ω throughout the experiment. Subsequently, the subjects were instructed to sit on a chair and avoid moving as much as possible. Next, headphones were placed on the subjects and the auditory thresholds (0 dB SL) of the 4 different tone bursts were collected.

Each subject was instructed to follow one out of the three following conditions: attend to the stimulation on the left ear and count the deviants (A-attL), ignore any stimulation and instead read a magazine (B-unatt R/L), and attend to the stimulation on the right ear and count the deviants (C-attR). Auditory stimulation was presented at the 50 dB sensation level (SL) and calibrated using

peak equivalent measures (pe SPL), see (Corona-Strauss and Strauss, 2016; European Committee for Standardization, 2007; Richter and Fedtke, 2005). Two presentation orders as in Hillyard et al. (1973) were used (ABCCBA or CBAABC). The latest resulted in a total of six conditions, meaning a total experiment duration of around 90 min, including electrode placement. The experiment length varied depending on the duration of the breaks between conditions needed for each individual subject.

Measurement Setup All stimulation and corresponding trigger files were created with a scientific computing software (Mathworks Inc., USA) using a sampling frequency of 44.1 kHz. Via a programmable attenuator headphone buffer (gPAH, gTec, Austria) that also controlled the intensity level, the stimulation files were delivered to the subjects through circumaural headphones (HDA-200, Sennheiser, Germany). The trigger channels were connected to a trigger conditioner box (g.Trigbox, gTec, Austria) that adjusted the signals for further acquisition using a biosignal amplifier. The biosignal amplifier (gUSBamp, gTec, Austria), controlled by computer, recorded the EEG activity (sampling frequency 512 Hz) and the trigger signals. The computer controlled also the presentation of the auditory stimulation by means of a sound card (Scarlett 2i4, Focusrite). All dataset were filtered (bandpass filter 1-70 Hz) and segmented accordingly after artifact removal using an amplitude threshold of (50 mV).

2.5 Laminar Auditory Activities and its Relation to Auditory Evoked Responses

Up until this point, we have analyzed the neural correlates of auditory selective attention using EEG at different auditory stimulation and experimental paradigms. We investigated different decoding techniques for analyzing the effect of the N1-P2 component of late AERs. The AER measurements are obtained from repetition of pure-tone stimulation (for studying the effect of long-term habituation) to the study of ongoing EEG measurements in presence of more complicated combination of auditory stimuli and longer duration.

All the previous studies in this work are based on neural activities on the surface of the scalp. However, the EEG signals are produced as a result of different neural processing from different sources projected through various spatial depths in the cortex. In Khodai (2014), it is reported that a deeper study of the neural activities of individual layer specific provide a better and more accurate perspective on the problem as there are *alteration of neural activities in different layers* of the cortical structure of auditory cortex. Therefore, it is important to understand how phase regularities translate across spatio-temporal scales of neural processing. In addition, not all hearing pathologies can be necessarily explained in terms of disorders in the peripheral system. In cases where peripheral regions are not the cause of hearing pathology, a deeper view on the auditory

pathway (i.e, examples include hearing hallucinations and tinnitus) is required, as damage to any of the components can lead to disorders in higher level processing in auditory cortex (Butler and Lomber, 2013; Khodai, 2014). In Butler and Lomber (2013), the functional connectivity between different cortical layers in congenitally deaf animals and hearing animals have been shown, in which the differences can be observed in the neural activity and connectivity at different layers.

Invasive electrophysiological measurements in many experimental settings are costly and hazardous. In the case of humans, there are many methodological limitations that constrain the level of the experimental procedures, however, generalizations from animal auditory system models to the human auditory system can be applied as the subcortical auditory system between mammalian animals has been highly conserved (Butler and Lomber, 2013; Glendenning and Masterton, 1998). Therefore, in this study, *we aim to understand the activity of the N1-P2 component of auditory ERP with relation to the cortical neural activities of auditory cortex in rats.* The ERP components equivalent to human N1 and P2 components have been studied previously in different species and has been shown that it has a shorter latency than of humans (Knight et al., 1985; Iwanami et al., 1994). It has been illustrated that the equivalent N1 component in humans occur approximately 50ms post-stimulus in rats (Knight et al., 1985) with the first early positive component appearing approximately at 30ms (P30)(Iwanami et al., 1994). This is followed by a second positive peak at about 80ms (P80) (Iwanami et al., 1994; Knight et al., 1985), similar to the P2 component in humans. In Iwanami et al. (1994), ERPs very similar to the human has been obtained in a two-tone discrimination oddball task in two conditions of resting and performing, at the surface and depth of the rat auditory cortex.

In Szymanski et al. (2011), the phase locking effects of local field potentials (LFP), measured at different cortical layers in rats with varying sensory stimulation have been fully analyzed. It has been shown that the phase of low frequency LFP follows a naturalistic stimuli, representing the fluctuating excitability of the particular neural population (Szymanski et al., 2011).

There has been studies showing a relation between the non-invasive AERs and invasive measurements at deeper layers of the cortex for different experimental paradigms (Celesia, 1976; Iwanami et al., 1994; Knight et al., 1985). The neural phase information of sensory cortical LFPs and EEGs resets in response to the stimuli onset at particular times (Kayser et al., 2009; Szymanski et al., 2011; Schyns et al., 2011). However the relation of AERs to the *functionality of laminar circuit mechanisms in presence of a salient stimulus* is not fully understood. In previous studies of (Kayser et al., 2009), it was shown that the phase information of laminar LFPs are aligned with respect to complex, naturalistic sensory stimuli at low frequency domains. Similar studies have been carried out by Szymanski et al. (2011).

In this study, we aim to understand the mapping of the thalamo-cortical activities onto AERs, as it will facilitate the measuring procedures and its application on human subjects. We employ *information theoretic* approaches to understand the relation between AERs and LFPs of cortical layers, using the phase information of neural responses at different spatio-temporal domains. We

attempt to describe the findings with of the laminar cortical processing and its relation to the evoked responses.

We use the LFPs of a penetrated multi-electrode array at ten different depths in the primary auditory cortex of rats. The experimental setup aims to reproduce the effect of sensory auditory adaptation, in which a tone burst is repeatedly applied to the hearing system. A broad-band transient acoustic stimulation during the experiment is used as our experimental paradigm. The penetrating electrodes are in columnar alignment with respect to the potential changes in the non-invasive measurable ERPs. The main incentive of the study is to understand the laminar projections onto the measured AERs using the IP of measured neural responses.

In the following sections, we describe the anatomy of thalamo-cortical interactions and main characteristics of the cytoarchitecture layers of A1 of a rat. We then describe the experimental setting and methodology.

2.5.1 The Cortical Laminar Organization in A1

The laminar cytoarchitecture of the auditory cortex, similar to other sensory cortices, consists of six layers (with the exception of the archicortex and paleocortex) that are commonly denoted by Roman letters. Every layer consists of different neuronal types responsible for a specific functional connectivity and task (Winer, 2011; Khodai, 2014). The layers are divided into three groups: the supra-granular (layers I and II), granular-layer (III and IV) and infra-granular layer (V and VI) (see (Winer, 2011)).

Layer IV, also known as the internal granular layer, consists of *smooth stellate (star-like)* and *spiny stellate* neurons, which are categorized as granule cells. The neurons in layer IV receive a high a number of afferents through the white matter⁶ and other cortical and subcortical regions (Linden and Schreiner, 2003). Layer IV is usually known as the main thalamorecipient⁷ layer (in case of primary auditory cortex, layer III is also involved as a receiver) as it receives projections from auditory pathway (i.e., specific thalamic nuclei). Layer III is dense in cell density and consists mostly of medium to large pyramidal cells, some of which that have apical dendrites extending up to layer I, and others, extending into deeper layers of the cortex.

The information that have been projected to layers III and IV run to the supragranular layers (II/III) (Winer, 2011). The supragranular layers project to other cortical layers by lateral spread of axons, which creates a patch processing patterns, as well to infragranular layers (V/VI) (Linden and Schreiner, 2003).

Layer V, also referred to as the internal pyramidal layer, is known for having medium to large pyramidal cells. Dendrites of large pyramidal cells in this layer can extend up to layer I, whereas

⁶white matter is primarily known as the main pathway between different cortical regions, remote cortical areas, as well as interaction between cortical and subcortical areas.

⁷The receiving afferents to the cortical layers are from different thalamus divisions. For more information in this regard see (Lee and Imaizumi, 2013; Khodai, 2014).

axons extend to white matter and are a major source of efferent fibers. Layer V and VI are considered to be the main contributors for sending the feedback information to other cortical and subcortical areas via white matter (Linden and Schreiner, 2003).

2.5.2 Experimental Setting

We examine the laminar activity of the auditory cortex in adult Sprague-Dawley rats of both sexes. The animals' body weight range from 450g to 850g. All experiments are conducted in accordance with the German animal protection act (5/2006) and were approved by an ethics committee.

The rats were anesthetized by the administration of isoflurane (3 vol.%) and oxygen (2l per min). Once the animal was non-sentient, the anaesthesia was replaced with Ketamine (Ketavet, 100mg/ml, 1-1.2 ml per 100g bodyweight and Xylazine (Rompun 2%, 4.5 μ l per 100g bodyweight) in a physiologic saline solution and administered via intraperitoneal injection.

The animals were placed in a fixed position in a tailor-made stereotactic frame for acoustic stimulation. After removal of scalp and muscular tissue, two micro screws (Mondeal, LOCK-DRIVE, titanium, 1.5x4mm), rostral to Bregma and caudal to Lambda are inserted for the epidural acquisition of auditory evoked responses. A trepanation, 1.4mm in diameter, was made 5.25mm posterior to Bregma and 5.25mm lateral to the midline for the insertion of the penetrating multitrode (Thomas Recording Multitrode Type I). An electrode 3.5mm posterior to Bregma and 3.2mm lateral to midline was inserted into the Thalamic Nuclei (TN) (Swanson, 1992). See Fig. 2.18 for an illustration of electrode positioning and surgical preparation.

We recorded extracellular potentials at 8 different depths, spaced at intervals of 125 μ m for the electrodes 1-7 and 650-700 μ m to the electrode top. The AERs were generated by a click-tone of 100 μ s duration and ISI of 1.5s (E.A.R.Tone 3A). Each recording session lasted 90s. Data was collected using a g.Tec USBamp amplifier; sampling frequency was set to 4.8 kHz. For the processing of single-trial responses we additionally acquired a stimulus locked trigger signal using the g.Tec TRIGbox. In this study we focus on AER-related electrode and laminar cortical processing.

2.5.3 Mutual Information of IP between Laminar Cortex

The main goals of this study are *obtaining meaningful information regarding projections of cortical layers' activities onto the measured AERs using the phase information of signals*; *decoding the delay process of information transfer between different layers as well as AERs using phase information*. We consider studying the level of information shared between neural activities of different cortical layers with AERs. For this means, we use the concept of mutual information. The mutual information (MI) between two random variables X and Y denoted as $I(X, Y)$ quantifies the degree of dependence between X and Y . The mutual information between X and Y is defined in terms of marginal entropies and the joint entropy $H(X, Y)$

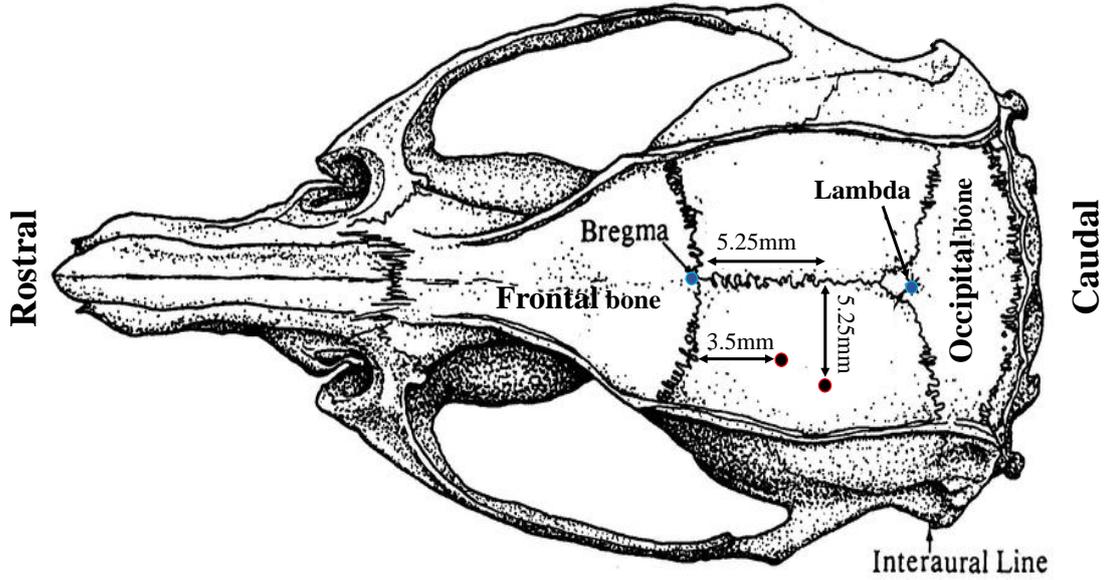


Figure 2.18: Two micro screws were inserted rostral to bregma and caudal to lambda for epidural AER recording (Blue outlines). A stereotactic burr hole was made 5.25mm posterior to Bregma and 5.25mm lateral to mid-line for the insertion of the penetrating multitrode (red outlines). Background illustration from George and Charles (2007).

$$I(X, Y) = H(X) + H(Y) - H(X, Y). \quad (2.36)$$

In our case, we want to compute the MI of the phase of segmented data between different electrodes at different frequencies. To compute the degree of MI ($I(P, P')$) between two given phase matrices P, P' , we first approximate the underlying probability density functions with normalized histograms $hist(P) = (\frac{m_1}{n}, \frac{m_2}{n}, \dots, \frac{m_k}{n})$ and $hist(P') = (\frac{m'_1}{n}, \frac{m'_2}{n}, \dots, \frac{m'_k}{n})$. Here, n is the number of data points in P and P' (we assume they have the same size), and m_i, m'_i is the number of data points from P and P' respectively that fall in the i -th bin of the histogram. We use $k = 5$ bins to evenly divide the range $[-\pi, \pi)$ (the method is similar to Russakoff et al. (2004)). Similarly, we compute a joint histogram

$$hist(P, P') = \left(\frac{m_{1,1}}{n}, \dots, \frac{m_{k,1}}{n}, \frac{m_{1,2}}{n}, \dots, \frac{m_{k,2}}{n}, \frac{m_{1,k}}{n}, \dots, \frac{m_{k,k}}{n} \right) \quad (2.37)$$

to approximate the joint probability density function of X, Y . Here $m_{i,j}$ counts how often we have a measurement from P fall into bin i with the corresponding measurement from P' simultaneously falling into bin j .

Hence, the MI between \mathbf{P} and \mathbf{P}' is approximated by

$$\begin{aligned}
I(\mathbf{P}, \mathbf{P}') &\approx H(\text{hist}(\mathbf{P})) + H(\text{hist}(\mathbf{P}')) - H(\text{hist}(\mathbf{P}, \mathbf{P}')) \\
&= -\sum_{i=1}^k \frac{m_i}{n} \log\left(\frac{m_i}{n}\right) - \sum_{j=1}^k \frac{m'_j}{n} \log\left(\frac{m'_j}{n}\right) \\
&\quad + \sum_{i=1}^k \sum_{j=1}^k \frac{m_{i,j}}{n} \log\left(\frac{m_{i,j}}{n}\right). \tag{2.38}
\end{aligned}$$

Image Registration We use a registration technique based on MI to detect time delays between different channels and/or frequencies. Given two phase matrices \mathbf{P} , \mathbf{P}' , we determine the time delay d between the matrices that maximizes their mutual information. The delay d can be either positive or negative, depending on whether \mathbf{P} is ahead of \mathbf{P}' or vice versa. We distinguish between two cases:

1. $d \geq 0$, that is the signal in \mathbf{P}' is delayed (or in sync) compared to \mathbf{P} . In this case, we use the phase data corresponding to the first 500 ms of each trial from \mathbf{P} , and then determine $d \geq 0$ such that the mutual information between this section of \mathbf{P} and the data from \mathbf{P}' that corresponds to the time window $[d, d + 500\text{ms}]$ in \mathbf{P}' is maximized. We choose a window of 500 ms in order to narrow the time interval to that contains the most relevant phase response.
2. $d < 0$, that is the signal in \mathbf{P}' is ahead of \mathbf{P} . We handle this case in the same way as case 1, with the roles of \mathbf{P} and \mathbf{P}' reversed. The delay d' of \mathbf{P} compared to \mathbf{P}' determined in this way is then simply negated: $d = -d'$.

Since we do not know ahead of time whether case 1 or 2 applies to a given pair of phase matrices, we always compute both and pick the one with the larger maximized mutual information. The maximum MI between different cortical layers and AERs for different frequencies is in Chapter 3, Section 3.6.2.

Chapter 3

Results

In this section, the results regarding the studies in chapter 2 are fully described. For each section, we will cite the corresponding study regarding the methodology and the measured data from chapter 2. Further discussions about the results are followed in chapter 4. The following section corresponds to the study described in section 2.3.2. *Results of this study has been published in MortezaPouraghdam et al. (2015a).*

3.1 Long-term Habituation Decoding Using Von-Mises Model

As described previously in section 2.3.2-data processing part, the phase information of all individual ERPs is extracted and visualized in a matrix \mathbf{P} (see section 2.3.1). To observe the effect of habituation, we analyze the data in a cross-sectional (column-wise) manner, that is the phase variations over different trials at a specific time instance (See Fig. 2.7 as an example). For a fixed time of interest, we segment the data into groups of Q samples with a mutual overlap of q samples between the groups. The benefit of using an overlap lies in obtaining smoother and more detailed results, similar to a moving average.

Fig. 3.1 shows the distribution of the phase data $\{\theta_{k,50}\}_k$ at 97 ms (sample 50) for four segments. The phase data are plotted as points on the unit circle for habituation process data in the top row (A) and for the non-habituation data in the bottom row (B). The resultant vector \mathbf{R} divided by the number $n = 200$ of samples for each group is the red vector. As described in section 2.3.1 the direction of the red vector is the circular mean of the data. The length of \mathbf{R} corresponds to the concentration of the data around the mean; the concentration is higher if the red vector is larger.

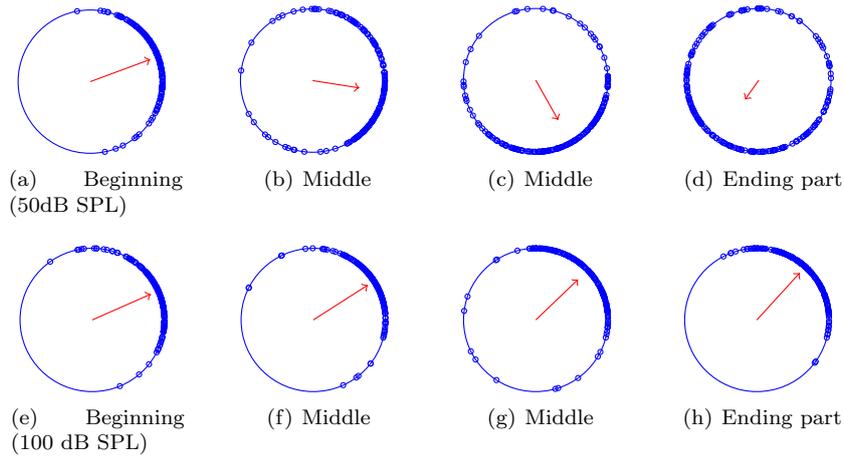


Figure 3.1: We show the dynamic of phase data over trials at a specific time of $t = 97ms$ for one human subject. The segment sizes are $Q = 200, q = 0$. Plots (1 to 4) in row A and row B correspond to 50dB SPL and 100dB SPL data respectively. The resultant vector \mathbf{R} divided by the number of segment samples is shown in red.

In the case of non-habituation data (B), the phase information is densely clustered around the mean and remains constant throughout consecutive segments. In the habituation process (A) the mean direction varies between the segments and the degree of phase cluster is low at the end of the experiment.

To model the variations of IP, we initially use simulated data of the neurofunctional model as described in section 2.3.2. We generate this data based on a normal distribution of physiological parameters. The applied inter-individual variance ensures a general validity of the simulation and subsequent data processing. The parameter range describes inter-individual variability in the sensory pathway and is based on physiological limits taken from Robinson et al. (204). Ten different data sets is generated. For each data set we consider the time between 97ms and 127ms, i.e., an interval which approximately covers the N1 wave. More precisely we use the corresponding samples $\{\theta_{k,b_j}\}_k$ at six different time samples $b_j := 50 + 3j, j = 0, \dots, 5$. The segments have a chosen length of $Q = 200$ and an overlap of $q = 195$ so that there are 120 segments at a fixed time b_j . The choice of a high overlap is to show a smoother and more detailed results of phase-cluster modulations. For each segment we fit a von Mises model. Fig. 3.2 shows the estimated concentrations κ of the von Mises model averaged over the six points of time $b_j, j = 0, \dots, 5$ and the 10 data sets. The decreasing effect of the concentration for the habituation process and a stable and high phase-lock for non-habituation is clearly evident.

Next we analyze the phase of measured ERPs in 10 subjects for the same time course from 97ms to 127ms. We segment the trials in the same manner as for the simulated data ($Q = 200, q = 195$). Fig. 3.3 gives the averaged estimated concentrations κ for 120 segments averaged over the six points of time b_j and the 10 subjects.

The behavior of the estimated concentration parameter can be also illustrated in the Shannon entropy

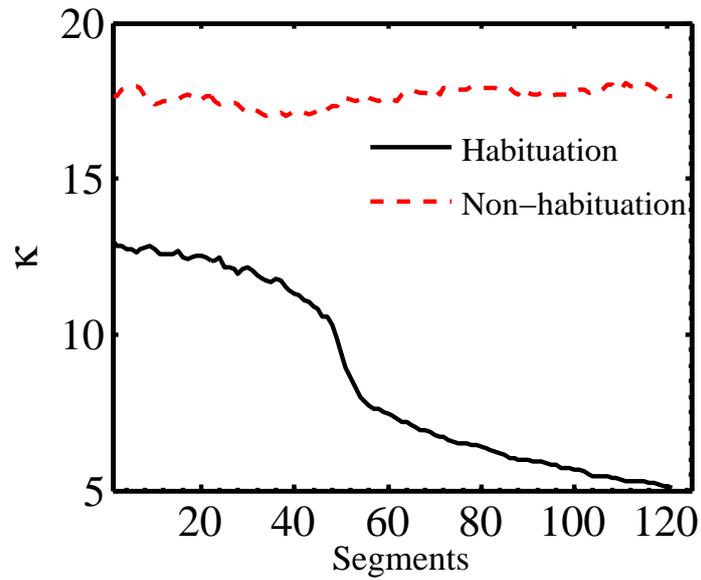


Figure 3.2: Average of the estimated von Mises concentration κ in the course of habituation and non-habituation for simulated data by the neurofunctional model. The concentration for the habituation data decreases drastically between the segments 40 to 70 but remains on the same level for non-habituation data.

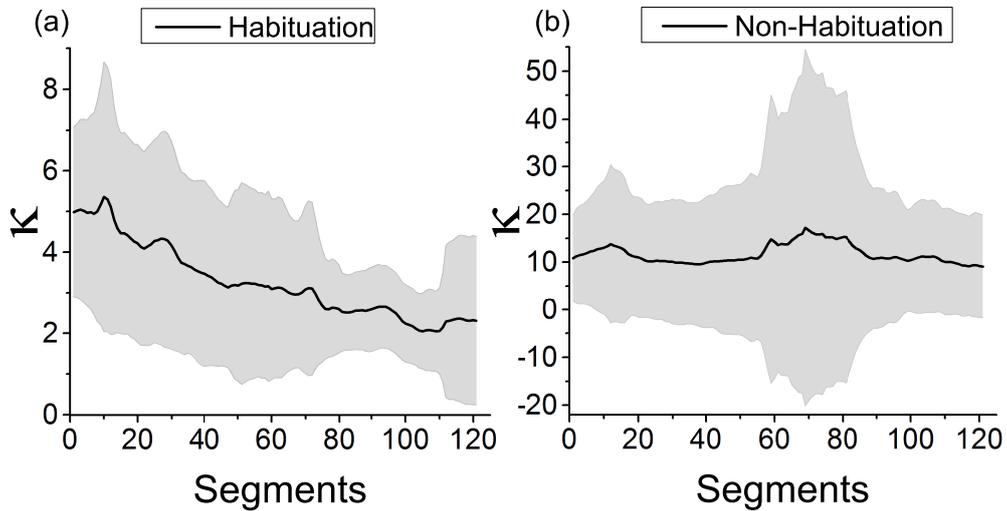


Figure 3.3: Averages of the estimated von Mises concentration κ in the course of habituation and non-habituation for measured human data. The shaded area shows the standard deviation. The decreasing effect of the concentration for the habituation process in contrast to the non-habituation process is clearly visible.

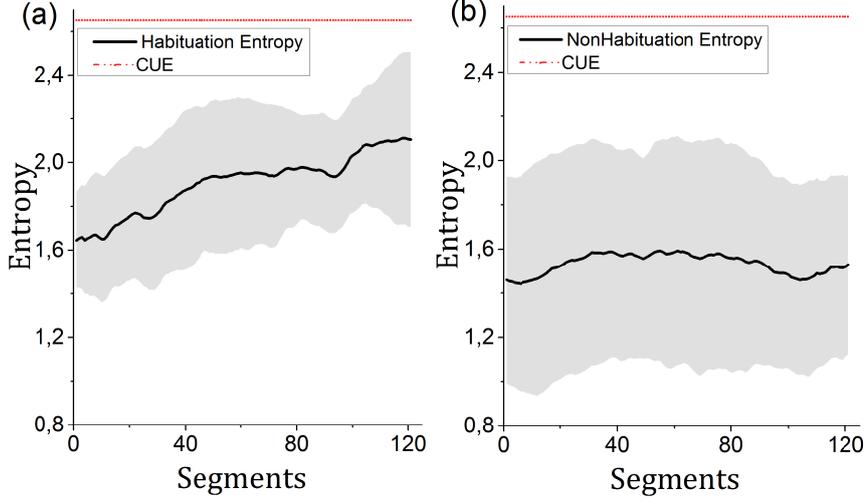


Figure 3.4: The plot shows the averages entropy of phase samples in consecutive segments ($Q = 200, q = 195$) for all measured data subjects. The CUE corresponds to the maximum entropy of continuous uniformly distributed data ($-\log_2(\frac{1}{2\pi})$) and is shown for comparison. The entropy of the habituation data increases to the entropy of uniformly distributed data, something that is not the case for non-habituation data. The shaded area shows the standard deviation.

of the segmented samples. Fig. 3.4 shows the entropy of the samples in the consecutive segments (averaged again over the six values for b_j and the 10 subjects). As the concentration decreases, the amount of entropy increases, approaching a level of entropy corresponding to a uniform distribution. This is consistent with the fact that the phase of the data becomes more uniformly distributed. On the other hand, an increase in the concentration parameter is reflected in a decrease of the entropy value, indicating a high density of neuronal synchronization in non-habituation responses.

So far we have dealt with the IP for the NLM filtered data. In Table 3.1 we show the average of the estimated concentrations κ_{s,b_j} per segment and time

$$\bar{\kappa} := \sum_{s=1}^{120} \sum_{j=0}^5 \kappa_{s,b_j}$$

for NLM and non-NLM filtered habituation data and non-habituation data for each of the 10 subjects together with their variances. In agreement with our previous findings, we have again the effect that the concentration is in general higher for non-habituation data. We see further that the results obtained from the NLM filtered data reflect a higher degree of phase cluster in comparison to the non-NLM filtered data. To measure the reliability of using a von Mises model in our computations, we evaluate the confidence interval for the estimated concentration parameters. Let $[\delta_{s,b_j}^L, \delta_{s,b_j}^U]$, $s = 1, \dots, 120$, $j = 0, \dots, 5$ denote the calculated confidence interval of concentration κ . Superscripts U and L stand for the upper and lower bound of the confidence interval. We average the size of the confidence interval for all segments at each time by computing $I_{b_j}^\kappa = \frac{1}{120} \sum_{s=1}^{120} |\delta_{s,b_j}^U - \delta_{s,b_j}^L|$. In Fig. 3.5 we show the $I_{b_j}^\kappa$, $j = 0, \dots, 5$ for the 10 subjects. The confidence interval was computed

S	$\bar{\kappa}_H$	$\bar{\kappa}_{NH}$	$\bar{\kappa}_H^{NLM}$	$\bar{\kappa}_{NH}^{NLM}$
1	0.47±0.12	0.68±0.09	2.09±0.88	4.14±0.73
2	0.96±0.15	1.76±0.13	5.98± 2.52	13.5±2.04
3	0.56±0.13	0.66±0.05	2.81±1.24	3.06±0.54
4	0.39±0.12	1.41±0.30	1.23±0.42	7.62±4.69
5	0.76±0.17	2.37±0.23	4.49±2.16	59.9±26.2
6	0.83± 0.12	0.50±0.15	3.36±0.56	1.67±0.43
7	0.78±0.19	0.88±0.11	2.90±0.76	3.43±0.96
8	0.29±0.13	0.55±0.10	1.04±0.38	1.93±0.32
9	0.70±0.05	0.74±0.21	3.56±0.48	6.40±3.96
10	0.51±0.09	1.13±0.10	2.61±0.92	9.11±1.84

Table 3.1: Average concentration of the 6 time intervals for each of the 10 subjects for NLM filtered habituation data $\bar{\kappa}_H^{NLM}$, NLM filtered non-habituation data $\bar{\kappa}_{NH}^{NLM}$, non-filtered habituation data $\bar{\kappa}_H$, and non-filtered non-habituation data $\bar{\kappa}_{NH}$ is shown. The concentration for NLM filtered data is higher than that of the non-NLM filtered data, reflecting a better phase clustering.

according to Fisher (1995). The findings regarding this study are discussed in chapter 4.

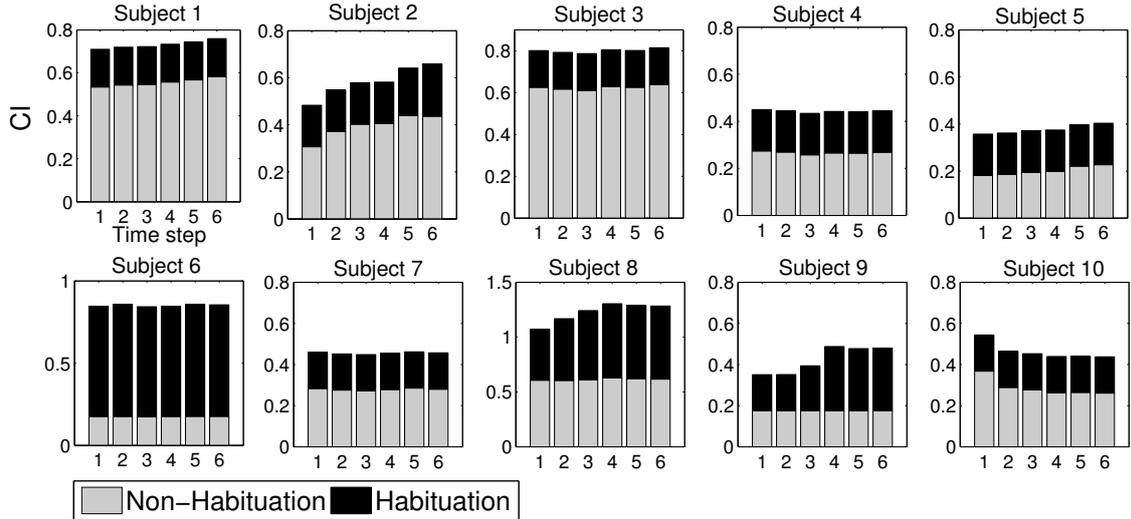


Figure 3.5: The average confidence interval size of the estimated concentrations over different times b_j , $j = 1, \dots, 6$ is presented for all subjects.

3.2 Abrupt and Continuous Detection of Temporal

Changes in ERPs in Long-term Habituation Process

Results regarding this study have been published in MortezaPouraghdam et al. (2014). In section 3.1, we tried to discriminate between the two processes of habituation and non-habituation by fitting a von-Mises model over different segment sizes. However, one of the drawbacks of the previous approach is selecting an appropriate length of Q and overlap samples q , as different values can yield different results in regard to capturing the long-term habituation dynamics. If the window

sizes Q are very large and coarse, the dynamic of habituation may not be captured. If they are very fine, the data can be over-fit. The same window size Q may also not be optimal for all subjects. The effects of selecting a coarse and fine window size are further discussed in chapter 4.

As a continuation to the previous study, we aim to investigate the possibility of detecting a more precise range at which habituation takes place. We expect the non-habituation state data to stay similar over time. As described in section 3.1 and the studies conducted in MortezaPouraghdam et al. (2015a), the diffusion of data samples at the N1 component of ERs over trials can be used as a signature for detecting the habituation state. In this study, we aim to use a change point approach to detect the sample point at which a significant diffusion in data concentration occurs. The method is an abrupt change point model combined with a Viterbi algorithm to examine the time range at which a habituation process occurs (see chapter 2, section 2.3.3 for more details about the method).

3.2.1 Synthetic Data Evaluation

We first evaluated the model on synthetic data with varying abrupt change points and on synthetic data with a distribution similar to the habituation. Fig. 3.6 shows two examples of different synthetic time series with abrupt changes in terms of the mean/concentration of a von-Mises distribution. The domain of data is in $[-\pi, \pi)$ and is generated using a rejection sampling approach, after selecting the number of change points. In Fig. 3.6-(A) which consists of three change points, the mean of the data after every change point randomly varies, but the concentration of data remains constant. In another illustration of a time-series with abrupt changes, when a change point occurs, the corresponding concentration parameter of data changes. This example can be compared to the real ERs over trials in case of habituation.

We generated synthetic data with different number of change points from $cp = 1, \dots, 5$. For every change point amount, 100 batches of data sets were generated randomly. The hazard rate selected for detecting change points is based on a geometric probability distribution. Let p be the probability of a change point, then the probability of having change points at different lengths is given by $P(g = x) = (1 - p)^{x-1}p$. The hazard rate function is computed as follows:

$$\begin{aligned}
 H(x) &= \frac{P(g = x)}{\sum_{t=x}^{\infty} P(g = t)} \\
 &= \frac{(1 - p)^{x-1}p}{\sum_{t=x}^{\infty} (1 - p)^{t-1}p} \\
 &= \frac{1}{\sum_{t=0}^{\infty} (1 - p)^t} \\
 &= p.
 \end{aligned} \tag{3.1}$$

We use the hazard rate as $H = \frac{m}{T}$ where m is the number of change points and T is the length of the synthetic data. The length of the synthetic time series is $T = 1000$ samples. To obtain a measure of

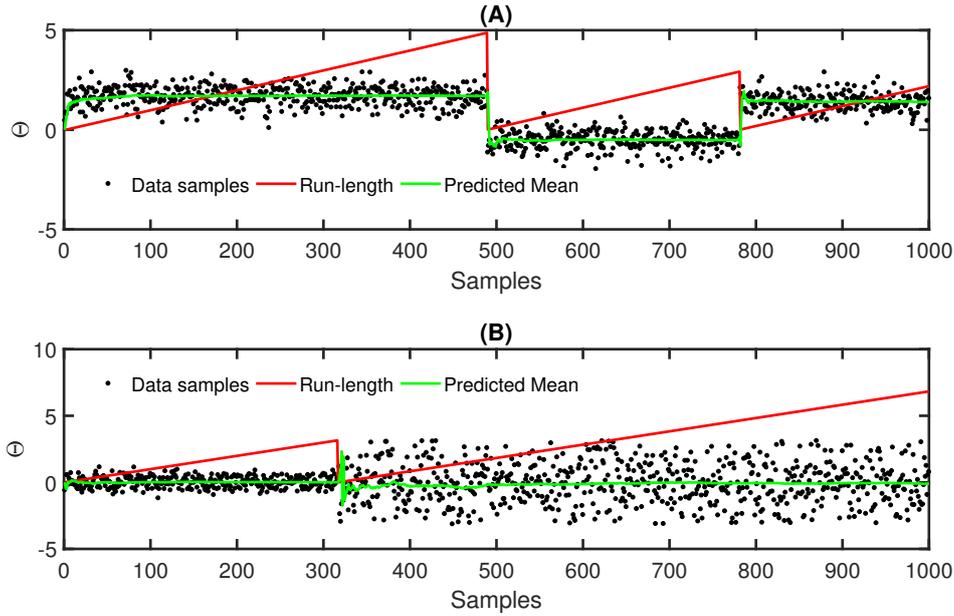


Figure 3.6: Detection of change points for simulated circular data. (A) The randomly generated data is illustrated as black circles for a better illustration of the abrupt change points. The changes are defined in terms of abrupt changes in the mean. The red line shows the variations of the run-state (or state transitions) over time. The green line is the predicted mean value. (B) Same explanation as for (A), however the abrupt change is defined in terms of the variation of concentration parameter.

how well the change points are estimated, the average difference between the estimated run-length and the actual state transitions based on the actual change points is computed. Fig. 3.7 shows one such example.

The same procedure as in Fig. 3.7 is repeated for different numbers of change points up to $m = 5$. The average results for 100 different sets of data have been shown in Fig. 3.8. As the length of the signals is $T = 1000$ samples, in worst case scenario where none of the detected change points are valid, the average difference between two run-lengths will be 500. The results in Fig. 3.8 suggest that the average difference is relatively small.

3.2.2 Evaluation of CPs on the Measurements

The method was applied on measured N1 of ERs (θ_{k,b_j}) at two different auditory stimulations of 50dB SPL and 100dB SPL (See experimental-settings part of section 2.3.2 and section 2.3.3.3 for details of the experiment). It is assumed that the habituation for a soft and comfortable stimulus will occur faster in comparison to an aversive stimulus. Hence, a change point for a soft stimulus will be detected earlier time comparison to the aversive stimulus. Therefore, run-lengths obtained from an aversive stimulus will be larger than for a soft stimulus. The changes in the run-length that would occur at the end of aversive stimuli have been discarded as it is difficult to assert if they are due to habituation or noise. As we are interested predominantly in change points detected due

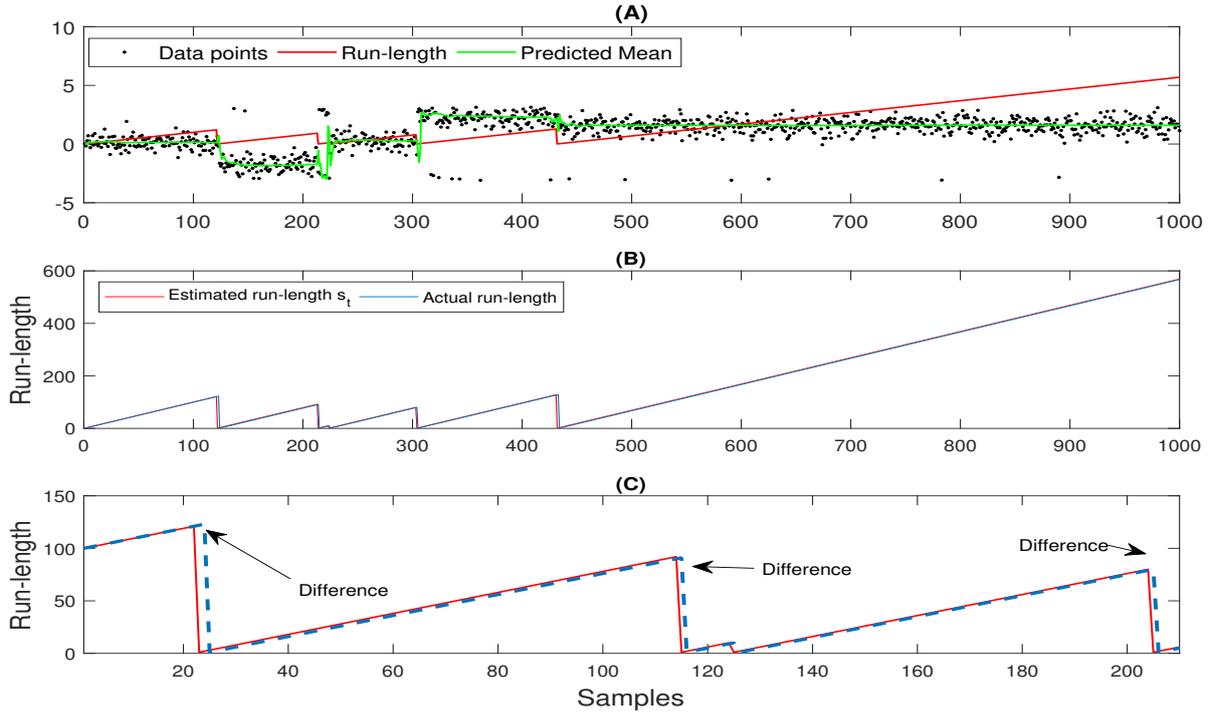


Figure 3.7: (A): An example of a time series with four abrupt changes. (B): The overlap of the estimated run-length and the actual run-length. The average of the difference between the two run-lengths is used as a measure to determine how well the run-length has been estimated. (C) This shows a small zoomed-in section of the plot in (B). The small delays in detecting the change points have been marked.

to changes in the von Mises concentration parameter rather than the mean, we center the data to have a mean of zero. The run-length variation for few subjects over trials at a particular time has been plotted in Fig. 3.9. For an aversive 100dB SPL stimulus, hardly any changes due to the high synchronization of phase information can be detected. However in 50dB stimulus, the changes in diffusion of phase modulations are more tractable and occur at earlier trials in comparison to the 100dB stimulus. There may remain fluctuations at later times due to the sensitization effect.

The subjects presented in Fig. 3.9 are among subjects with a best fit of run-length in terms of distinguishing between habituation and non-habituation stimuli.

In Fig. 3.10 we summarize some of the statistical properties of the estimated run-lengths variation for each subjects at different times. Fig. 3.10(A) shows the average maximum value of the run-length for each specific time for all of the subjects. The index values t_i correspond to data over trials at times $t_1 = 97ms$, $t_2 = 107ms$, $t_3 = 117ms$ and $t_4 = 123ms$. The average maximum run-length is consistently larger for 100dB SPL than for the 50dB SPL, indicating a lower variation in run-length transition. The average maximum run-length value over all times for specific subjects has been shown in plot B. There were statistically significant differences of maximum run-length values between soft and aversive stimulations as determined by a one-way ANOVA test ($F(1, 12.23)$, $p = 0.0026$) using a 5% significance test. In Fig. 3.10-(C) the average run-length for individual subjects has been

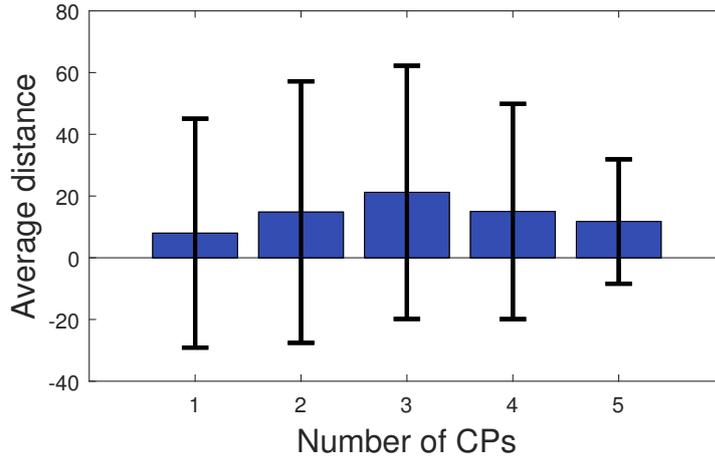


Figure 3.8: The average distance between the actual change points and the optimal run-length path.

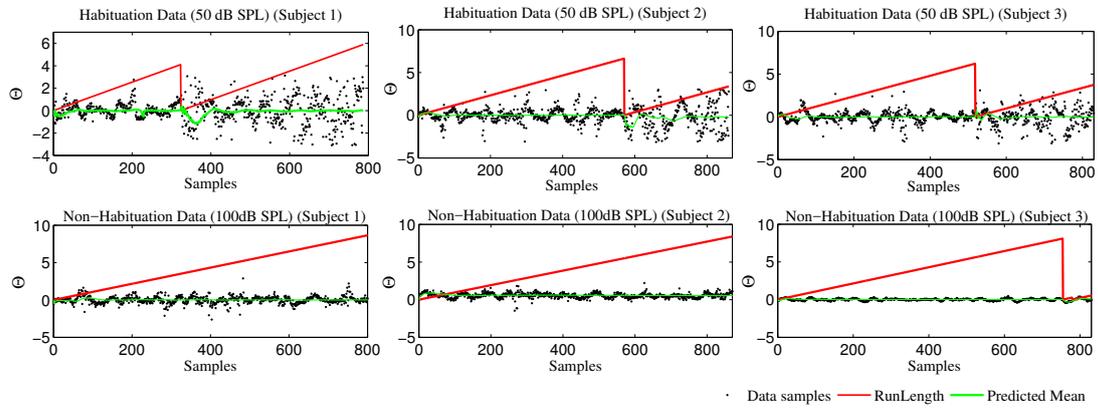


Figure 3.9: The top row shows the results of the change point algorithm for 50dB SPL data for three subjects. The corresponding 100dB SPL for individual subject is plotted on the second row.

plotted. The mean difference between the two stimulations is significant as $F(1, 13.36), p = 0.0018$ in an ANOVA analysis test.

Setting of Hazard Rate For Measured Data As we were interested in detecting change points in terms of concentration corresponding either to the occurrence or absence of habituation process, the hazard rate had to be adjust accordingly. We therefore empirically enforced a hazard rate of $H = 9.8e - 12$ for our study. As in the case of measured signals, noise artifacts are present and the hazard rate has to be set to a small value to avoid noise-induced fluctuations in the run-length behavior. The small choice of the hazard rate indicates that the detected change point has been significant in terms of underlying generative model such that it could not be ignored by the model. In addition, one of the reasons for enforcing a small hazard rate is that we are interested to detected one change point at most such that it can be related to the occurrence of long-term habituation.

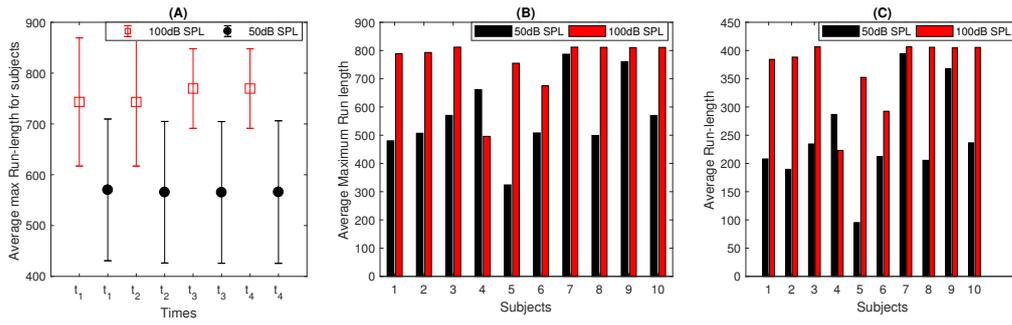


Figure 3.10: (A): The average maximum run-length value over all subjects at different times for two different stimulations. (B): Average maximum run-length for different times for individual subjects. (C): The average run-length at different times for individual subjects.

3.3 Detecting the Gradual changes of long-term habituation using a State-Space Model

Results regarding this study have been published in MortezaPouraghdam et al. (2016). One of the limitations of studying abrupt change points is that it will be difficult to determine the status of habituation for a range of auditory stimulations that may not differ significantly with regards to loudness perception and comfort level. A single transition in run-length may not be accurate in understanding the occurrence of habituation in the temporal domain. Multiple rapid transitions make it difficult to draw concrete conclusions or differentiate the habituation state from the noisy epochs. In addition, the choice of an optimal hazard rate to describe meaningful changes in the signal can be a challenging task, as controlling the number of change points and avoiding noisy detections can lead to small hazard rate values. An off-line evaluation of run-length for classifications of more complicated stimuli can be also challenging.

To overcome some of these limitations, we considered a *gradual change point detection* method instead of an abrupt change point proposal. This is also more consistent with the definition of long-term habituation, by which changes in the level of attention-binding are defined as a gradual process (See section 1.3.3). Therefore, instead of considering a binary change point paradigm at every time instance (i.e., $s_t = 0$ or $s_t = s_{t-1} + 1$), we considered a state-space model such that gradual changes can be captured. The model description is described in section 2.3.4.

The state-space model designed in section 2.3.4 is applied on four different auditory stimulation. The auditory stimuli are 60dB SPL, 70dB SPL, 80dB SPL and 90dB SPL. As the difference between stimuli are not as large as the difference between 50dB SPL and 100dB SPL, detecting the long-term habituation is also more difficult. In Fig. 3.11, the phase distribution of data Θ at different parts of the experiment for all stimuli is illustrated. Our goal is to detect the long-term habituation by tracking the gradual changes occurring in the phase data Θ_t .

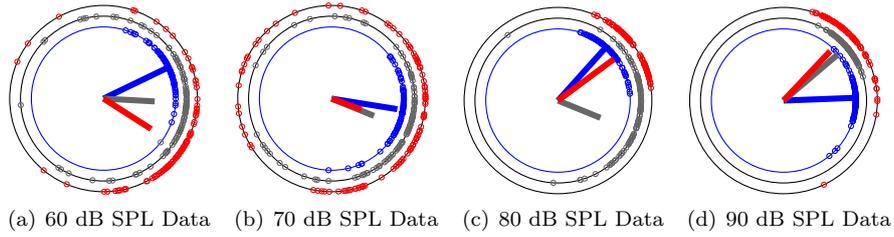


Figure 3.11: An example of phase distribution over trials at a specific time $\Theta_t, t = 97ms$ for a subject at four different stimuli levels of 60 dB SPL, 70 dB SPL, 80 dB SPL and 90dB SPL has been shown. The phase data Θ_t has been divided into three different segments corresponding to the beginning (B), middle (M) and end (E) of the experiment. Segment size is $G = 150$. Blue, gray and red data circles correspond to the B, M and E of the experiment respectively. The lines are the resultant mean vector divided by the number of segment samples. As phase data spread more around the circle, the length of the vector mean decreases.

3.3.1 Initial Conditions

One problem of the forward method as described in Chapter 2, section 2.3.3.2 is that the results can be heavily influenced by the initialization at $t = 1$. Depending on how the state likelihoods for the first observation are initialized, state likelihoods at subsequent times can be far from the ideal (see Fig. 3.12 for the effect of non-adequate initialization in data). A similar effect is observed if there is a sudden change in the underlying states at a later time in the data or if the initial data points are outliers (noise). Because the forward method does not *look ahead* of the current time, it adapts to changes in the data only with a delay. We add a backward pass to compensate for these effects, and to remove any influence of the *direction of time* on the results.

We also mitigate the initialization problem by using the result of the forward pass at t_N to initialize the backward pass, and vice-versa. Unless the data distribution drastically changes close to the end of the data t_N , the forward method usually converges sufficiently at that point, thus providing a very stable initialization for the backward pass. After the backward pass is done, we run the forward pass for a second time, using the result at t_1 from the backward pass for its initialization. The results of the second forward pass together with the results of the backward pass are used in analysis. The results of the first, randomly initialized forward pass are discarded in the final result to remove the influence of the original random initialization.

3.3.2 Optimization of Prior Parameters

In this section we report on the setting of the models' prior parameters K and σ^2 and other settings that were used for analyzing IP of ERPs. The ERPs correspond to nineteen subjects at different auditory stimulations of 60dB SPL, 70dB SPL, 80dB SPL and 90dB SPL. The phase information over trials is analyzed at seven distinct times of Θ_t ($t = 97ms$ to $t = 197ms$ with step sizes of three,

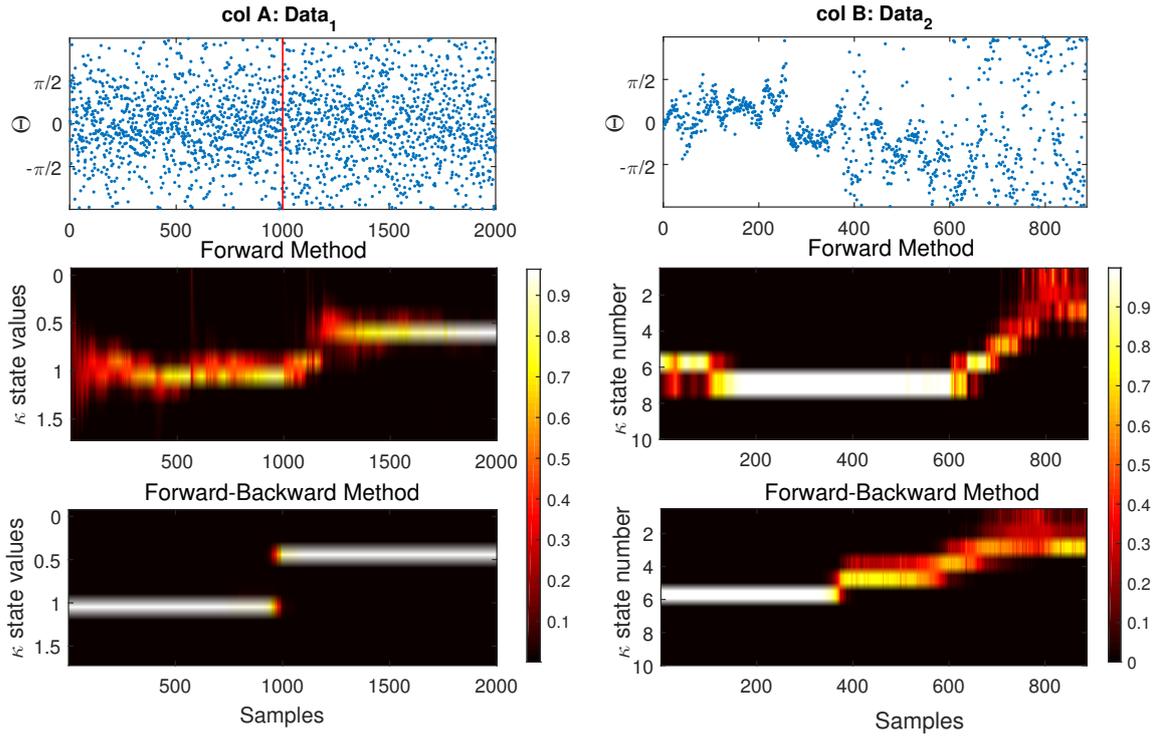


Figure 3.12: We show the effect of the forward and forward-backward approach for two different data examples. Column A: shows an artificial data set generated using a rejection sampling approach with an exact time change at sample 1000. The estimated concentration for the forward and forward-backward approach has been shown respectively. Column B: shows the estimated concentration for the forward method (1-B) and (2-B) the phase information of habituation data. The effect of the initialization on the results of the forward method is clearly visible. With regards to detecting change in phase modulations, the forward-backward approach provides more accurate timing information than the forward approach. The change occurs at sample 400.

or more specifically for all trials at samples of $M = 44$ to $M = 62$) for the phase matrix $P \in \mathbb{R}^{N \times M}$. The number of discrete states for R_μ and R_κ is $m = 20$ with an upper bound of $\ell = 63$ for the concentration parameter. We discretize the concentration parameter κ using a logarithmic scale.

The state transition model described in section 2.3.4.1 has two free parameters, K and σ^2 that determine the speed with which μ and κ change over time respectively. If we fit the parameters such that the likelihood of the observations $p(\Theta_{1:T})$ is maximized, we overfit the model to a specific dataset. The main problem with this approach is that it does not help us discriminate between different groups of stimuli reliably as possible. In addition, no prior information regarding the decaying behavior of the habituation process is used. We propose to set the free parameters K and σ^2 such that the resulting fits are robust under noise, allowing to reliably predict the stimulus level that the subject has been exposed to. The optimization criteria for the prior parameters is defined such that the ratio of the variance between different groups of stimulus levels to the variance of within groups of our novelty measure is maximized as defined in Eq.3.2. Here the novelty measure is defined as the ‘normalized’ expectation value $E'(\kappa_t)$ and reflects the level of change in the κ parameter rather than its absolute value.

$$\arg \max_{K, \sigma^2} \left(\rho = \frac{\text{between group variance of } E'(\kappa_t)}{\text{average within group variance of } E'(\kappa_t)} \right) \quad (3.2)$$

The procedure for computing the normalized expected value $E'(\kappa_t)$ of concentration κ is as follows: For a given observation Θ , we first obtain $E(\kappa_t) = \sum_{\kappa_j \in \kappa} \{ \kappa_j p(\kappa_t = \kappa_j | \Theta, \sigma^2, K) \} \forall t = 1 \dots N$ by performing the forward-backward passes. Since our primary interest lies in the change of κ over trials and not in its absolute value, we apply a normalization as such: Given a sequence of $E(\kappa_t)$ over time t for a given subject and stimulus, we compute the average of the last 50 samples. We then divide all values $E(\kappa_t)$ by the computed average: $E'(\kappa_t) = \frac{E(\kappa_t)}{\sum_{i=N-50}^N E(\kappa_j)}$. Our choice in using the last 50 samples for the normalization is based on the assumption that any habituation-related change in the values of κ_t will have happened before that point.

For the optimization itself, we consider a finite set of possible parameters $K \in R_{K_{prior}}$ for $R_{K_{prior}} = \{K_1, K_2, \dots, K_r\}$ and $\sigma^2 \in R_{\sigma^2} = \{\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2\}$. This is based on prior information about the habituation process in which the changes are defined to occur steadily in time. Thereby they are set such that the corresponding probability distributions for the mean and concentration parameters are peaky (small width), indicating the transition in mean and concentration states cannot happen abruptly. Given the range of possible values for K and σ^2 as $R_{K_{prior}} = \{K_1, K_2, \dots, K_r\}$ and $R_{\sigma^2} = \{\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2\}$ respectively, we obtain r^2 different configuration prior pairs (K_i, σ_j^2) , $i, j \in 1, 2, \dots, r$. For each of these configurations, we first perform the forward-backward procedure and then compute the ρ criteria. Finally we select the parameters that lead to the highest ρ value.

Before describing the results of the ERP measurements, we explain the results of the following steps: 1- Validating the model on artificial circular data. 2- Reporting on the results of model transition parameters K and σ^2 for experimental data. 3- Tracking data distribution over trials. 4- Reporting

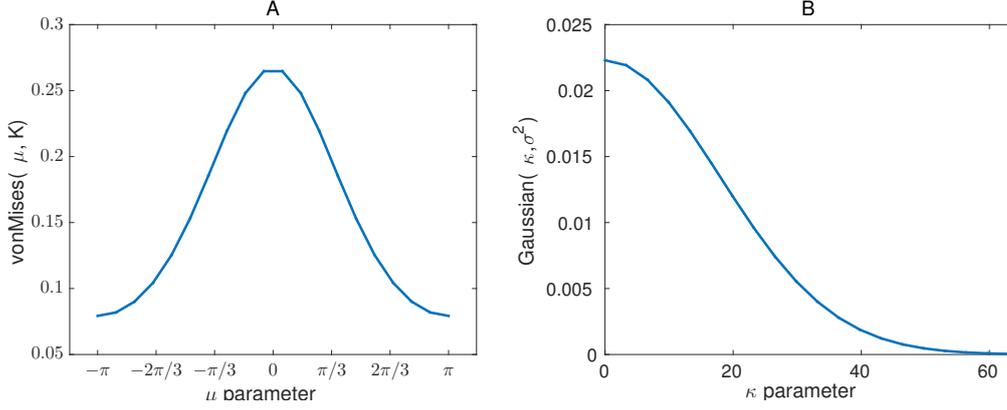


Figure 3.13: The plots show the probability distribution for the mean and concentration states with chosen prior parameters K and σ^2 . The plot (A) corresponds to the pdf (von Mises distribution) of the mean state transition model with the estimated concentration of $K = 0.6071$. The plot (B) corresponds to the pdf (Gaussian distribution) of the concentration state transition model with the estimated variance of $\sigma^2 = 320$.

average concentration for different auditory stimuli.

3.3.3 Model Validation on Artificial Circular Data

We first validate the forward-backward method on artificial data generated from a von Mises distribution with different concentration parameters κ . The artificial data is generated using a rejection sampling method $\Theta = [\Theta^{(1)}, \Theta^{(2)}, \dots, \Theta^{(J)}]$, such that $\Theta^{(i)} \sim \text{vonMises}(\mu_i, \kappa_i) \forall i = 1, \dots, J$ and with Θ of length L . As we are not interested in the mean changes of data, we keep the mean parameters constant for all data segments $\Theta^{(i)}$.

In Fig. 3.14 we show two artificial data series that are composed of segments sampled from von Mises distributions with different κ parameters. The transition times between the distributions are at $t_1 = 1000$ and $t_2 = 2000$ for plot (a) and (b) and $t_1 = 1000$, $t_2 = 1500$ and $t_3 = 2500$ for plot (c). After applying our Bayesian model, we compute the mean square error (MSE) between the estimated concentration and the actual concentration used for generating the data. We also apply a maximum-likelihood approach to estimate the Von-Mises parameters using a moving window of different lengths (the ML estimation is described in section 2.3.2). The MSE of our method for the data presented in Fig. 3.14-(a,b,c) are 0.0033, 0.014 and 0.033 respectively. We generated 50 artificial circular data sets with the same concentration and transition points for each of the examples (a,b,c) in Fig. 3.14 and evaluated the average of MSE for the estimated concentration parameters. The MSE are 0.041 ± 0.04 , 0.085 ± 0.064 and 0.010 ± 0.020 respectively.

In addition, we tested the accuracy of our model and the ML approach in Fig. 3.14-(c) under additive noise, to assess the robustness of the methods. The noise was generated from a normal distribution with different variances $\tilde{\Theta} = \Theta + N(0, \sigma^2)$ for $\sigma^2 \in \{0.01, 0.02, \dots, 0.05\}$. The resulting signal was wrapped to the range $[0, 2\pi)$. For the windowing approach, we use 4 different sizes $G = 50, 100, 200, 400$ with 98% overlap. The corresponding MSE results are shown in Fig. 3.15-(a).

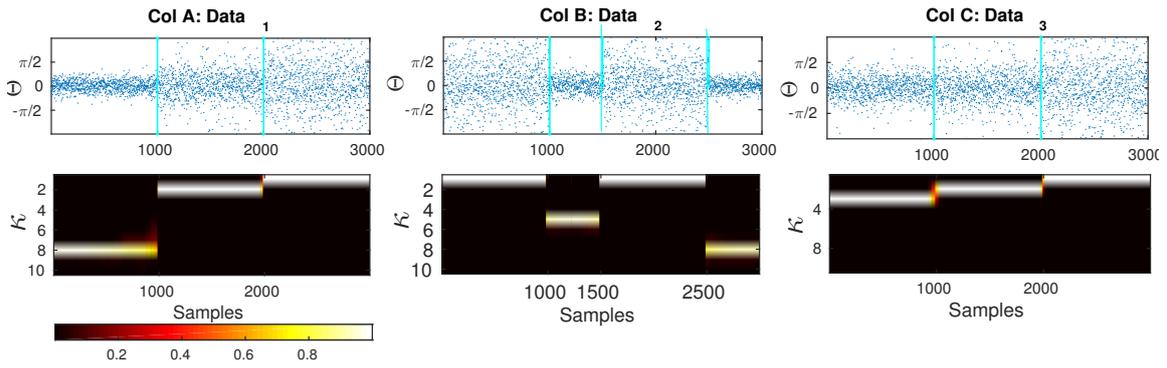


Figure 3.14: Results of applying the forward-backward Bayesian model on three different data sets. (a): Artificial circular data that is generated from 3 different Von-mises distributions with concentration parameters of $\kappa_1 = 8$, $\kappa_2 = 2$ and $\kappa_3 = 1$. The transitions are at samples $t_1 = 1000$ and $t_2 = 2000$. (b): Circular data generated with dispersion values of $\kappa_1 = 1$, $\kappa_2 = 5$, $\kappa_3 = 1$, $\kappa_4 = 8$ with changes at samples $t_1 = 1000$, $t_2 = 1500$ and $t_3 = 2500$. (c): Artificial data that is generated from $\kappa_1 = 3$, $\kappa_2 = 2$ and $\kappa_3 = 1$. The transitions are at samples $t_1 = 1000$ and $t_2 = 2000$.

We used a prior of $\sigma^2 = 0.08$ for Fig. 3.14-(a). The small value of σ is due to the small range of concentration values in this example. Note that the transition is from $\kappa \rightarrow \kappa + 1$ in a single time-step. The transition probability of this event with the given σ is fairly low (about 0.0002), avoiding overfitting.

Furthermore, we tested the model on data with random change points and random concentration parameters. We generated 50 data-sets Θ , each with two random change points at different concentrations $\kappa_1 \in [6, 10]$, $\kappa_2 \in [3, 5]$ and $\kappa_3 \in [1, 2]$. In each case we generated a series of $L = 3000$ sample points. We then computed the average MSE for both the windowing / ML approach and for our forward-backward approach. See Fig. 3.15-(b). The same prior parameter as in part (a) was used for this part.

Note that for the windowing approach with window size w , an input of L samples only results in $L - w$ parameter estimations. In order to make the MSE values comparable, we extend the input signal on both ends with $\frac{w}{2}$ additional samples from the first distribution, and the same number of additional samples from the final distribution.

In a third example in Fig. 3.15-(c) we generated data sets with a higher concentrations, i.e., $\kappa_1 = 9$, $\kappa_2 = 1$ and $\kappa_3 = 9$ at transition times $t_1 = 1000$ and $t_2 = 2000$. We used the same prior parameter $\sigma^2 = 0.6071$ for artificial data as was obtained with experimental data (In section 3.3.4 we report on the results of the prior parameters for the experimental data). The model is able to track the data distribution at different noise levels. This example was used to show that given a similar range of concentration values of the data samples, the prior parameters from the experimental data can be used to track the distribution of artificial data-sets.

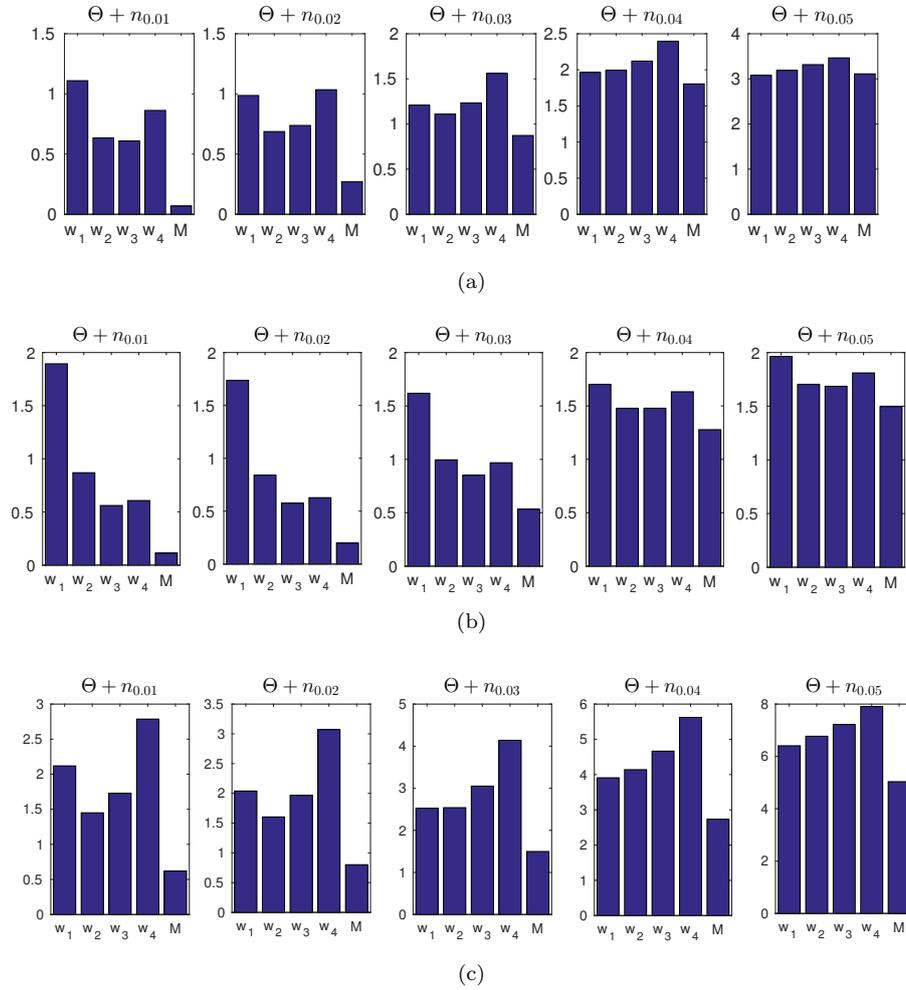


Figure 3.15: (a): The average mean square error (MSE) of estimated concentration for 50 artificial data-sets. The generated data consists of concentration parameters of $\kappa_1 = 3$, $\kappa_2 = 2$ and $\kappa_3 = 1$ with change points at times $t = 1000$ and $t = 2000$. The label M refers to the results obtained from our Bayesian forward-backward model. The labels w_1 to w_4 refer to different window sizes of 50, 100, 200 and 400 samples. (b): The MSE of 50 artificially generated data-sets with random transition times in concentration, and random κ values. (c): The MSE for 50 artificially generated data-sets with transition times at $t = 1000$ and $t = 2000$ with the concentrations of $\kappa_1 = 9$, $\kappa_2 = 1$ and $\kappa_3 = 9$. We used the prior parameter $\sigma^2 = 0.6071$ for $p(\kappa_t | \kappa_{t-1}) \sim N(\mu, \sigma^2)$ for our model.

3.3.4 Determining the Model Parameters for Experimental Data

The prior parameters obtained for the transition models are set to $K = 0.6071$ and $\sigma^2 = 320$. The criteria for optimizing σ^2 and K is described in section 3.3 and is based on Eq. 3.2. In Fig. 3.13 we show the state transition probability distribution for the selected mean and concentration parameters K and σ^2 respectively. These parameters affect the speed of changes in the transition models and consequently, how abruptly the changes are determined. As an example, as the variance parameter σ^2 in $N_{\mu, \sigma^2}(x)$ increases and the concentration parameter K in $VM_{\mu, K}(x)$ decreases, the state transitions can occur more abruptly. This leads to a more abrupt and sudden detection of changes in the signal's phase distribution. In the extreme case of a uniform state transition probability at any time, all subsequent states would have the same likelihood, independently of the current state. As a consequence, the μ_t state would simply track the signal itself, while leaving κ_t at a constant high level. In our case, the chosen parameters K and σ^2 consider the prior information of habituation into the model.

3.3.5 Tracking Data Distribution over Trials

In Fig. 3.16 we show the analysis of the Bayesian change point algorithm on a data set in which habituation occurs. The data corresponds to 60dB SPL which has in general a comfortable loudness perception. The likelihoods of the set of states in time (sample times) allow us to track the temporal changes in the mean and concentration of the phase data. In this study we are particularly interested in the changes in concentration over trials as an indicator of the degree of attention allocated to the sensory stimulus. The corresponding marginal distribution of concentration parameter over the samples is shown in part B. Every state number (here it is between 1 and 20) corresponds to a different concentration values distributed logarithmically between [0,63]. In C the likelihood of the discrete state space at four sample times corresponds to samples from the beginning, middle and end of the experiment. The concentration intensity drops significantly at the end of the experiment in relation to its initial value at the beginning of the experiment.

The effect of habituation between different subjects is variable. For additional clarity, in Fig. 3.17, we illustrate the expected value of concentration at different stimuli intensities along with the individual temporal changes in concentration for three different subjects. The results of three different subjects in Fig. 3.17 show the change process in the concentration states of phase data over all trials at a specific time $t = 97ms$ at different stimulus levels. The marginal likelihoods of the concentration parameter illustrate the detected changes in the phase synchronization. The results of 60 dB SPL in subject 1-(A,B) present a gradual diffusion in phase synchronization, whereas in 70dB SPL (C,D), the degree of phase cluster on average is low and contains many cyclic changes. For 80dB SPL and 90dB SPL (E-F) and (G-H) respectively, the phase clustering remains relatively high throughout the experiment before undergoing significant state transitions. This is particularly evident in the case

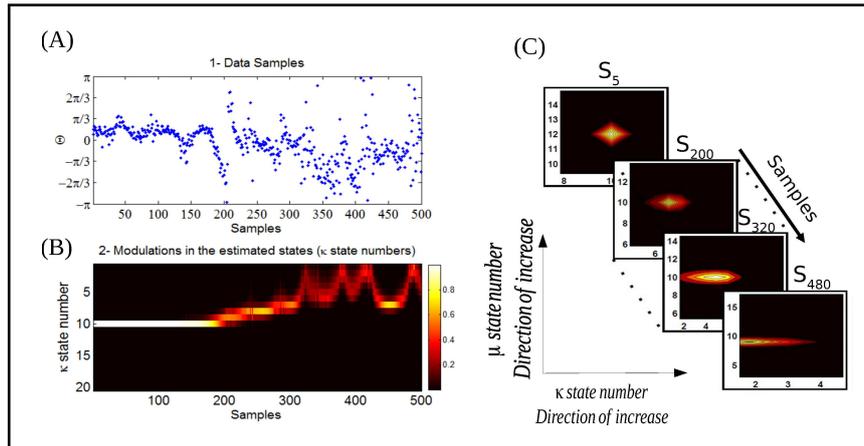


Figure 3.16: (A): The data corresponds to the phase information of auditory ERPs over trials for a long term habituation. (B) The marginal likelihood of the concentration parameter over trials has been shown. (C) The individual likelihood values of the discrete state space at four different samples has been shown. The first state-space likelihood corresponds to the fifth sample (beginning of the experiment) up to the sample 480 at the end of the experiment. The estimated concentration at the beginning of the experiment is high (reflecting a high binding of attention) and decreases significantly towards the end of the experiment (lower binding of attention).

of 90dB SPL in which a significant change occurs around the sample 510. The averaged results of the concentration at different times suggest that the level of phase diffusion for higher stimuli levels such as 80dB SPL and 90dB SPL is relatively lower. This indicates a stronger attention-binding to the stimulus because of the subjective unpleasant perception of the stimulus.

The same explanation applies to the other two subjects in Fig. 3.17. We explain the findings as follows: The detected change in the concentration states of 60dB SPL and 70dB SPL in subject 2-(A,B) and (C,D) tends to fluctuate rapidly between lower-concentration states. The average expected value of the concentration at these two stimuli are very low and contain little structural synchronization. The diffused phase information throughout the experiment indicates a low attention-binding to the stimulus. In the case of 80dB SPL phase samples are uniformly distributed for the first 410ms and then followed by a higher degree synchronization, that is closely reflected in the fitted concentration states. In the last subject in Fig. 3.17, a same type of explanation regarding phase modulations can be applied.

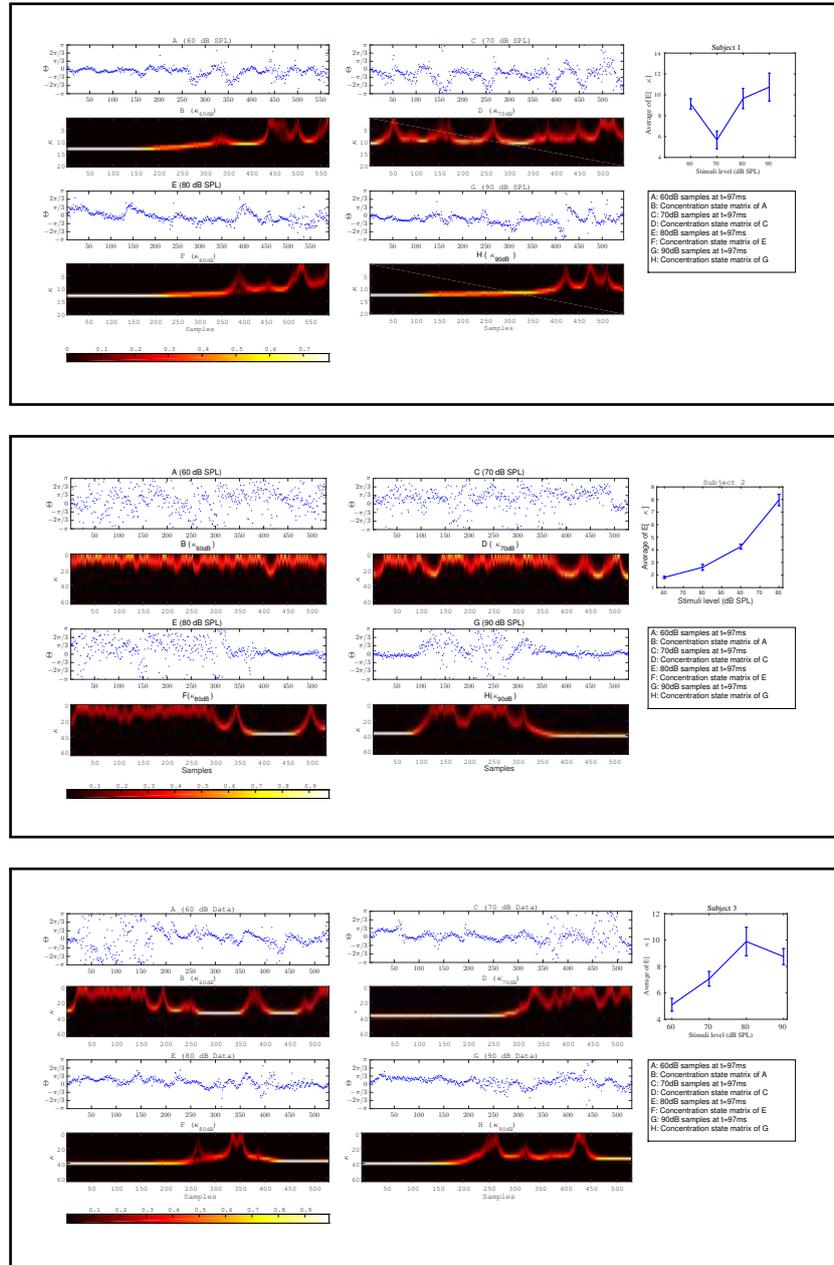


Figure 3.17: The plots in box [1] and [2] and [3] correspond to the results of three different subjects. Plots (A) to (H) for every box correspond to the forward-backward results of the concentration likelihood over trials at different stimuli. Part I is the expected value of concentration for data at different times t . Based on the magnitude of the expected value of the concentration parameter at different times as shown in part I for every subject, we are able to distinguish between different stimuli.

3.3.6 Average Concentration For Different Auditory Stimulations

The objective of this study is to evaluate the degree of habituation using the IP of ERPs induced at the different stimuli of 60dB SPL, 70dB SPL, 80dB SPL and 90dB SPL. To do this, we use a Bayesian model to track the changes in the underlying concentration in the instantaneous phase information of ERPs. In addition we use verbal responses from participants about the loudness of different stimuli. This knowledge was used to validate the conclusion that the relation between the objective measure and the loudness scale at different stimuli.

Despite the high variability among subjects with regard to changes in the concentration states (see Fig. 3.17), the average concentration results in Fig. 3.18 suggests that *as the loudness level increases, it is highly probable that the degree of phase synchronization increases as well*. To validate the obtained results, we compare the averaged concentration against the average verbal responses of the participants regarding their perception of the sounds. As shown in Fig. 3.18 (b), as stimuli level increases, the intensity of the loudness perception increases as well. This is consistent with the studies conducted by Hood and Poole (1966) and Stephens and Anderson (1971), that the stimuli between 90dB SPL to 100dB SPL is considered uncomfortably loud in normal hearing subjects. Furthermore we applied a one way ANOVA test over different stimulus levels across all subjects and found out that we can reliably distinguish between the 60dB SPL and 90dB SPL using the average of the expected concentration at a significance level of 5% with $p = 0.0101 (F = 7.37)$.

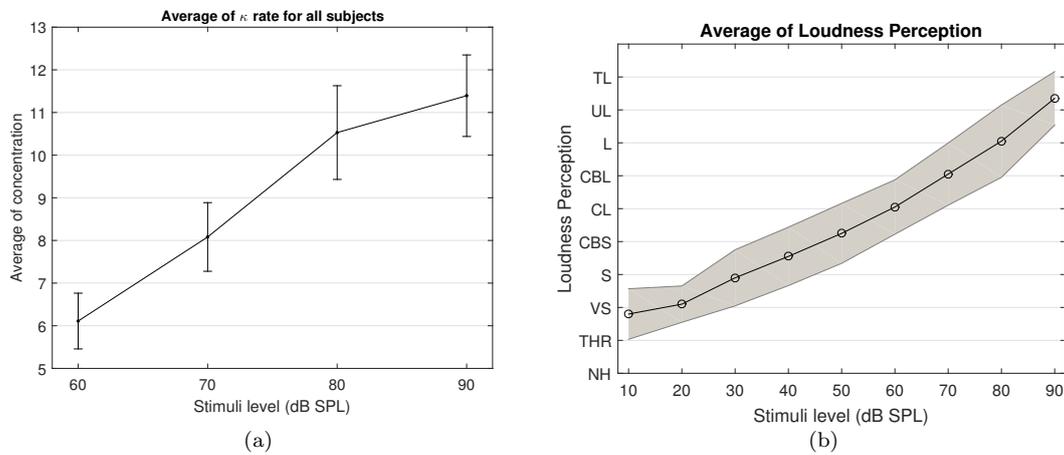


Figure 3.18: (a): The average expected concentration over all subjects at different stimuli levels. (b) The average loudness perception for all subjects at different stimuli. The loudness perception y-axis is defined as NH: Not heard, THR: Threshold, VS: very soft, S: Soft, CBS: comfortable but soft, CL: comfortable loud, L: loud, UL: upper level, TL: too loud.

In Chapter 4, we further discuss the effectiveness of the algorithm by applying it on different data. Other statistical tests are conducted to better understand classification between different stimulation groups. Major limitations in this study are also discussed fully in Chapter 4.

3.4 Removal of Spurious Phase Variations in Ongoing EEG Signals

The initial results of this section has been published in MortezaPouraghdam et al. (2018); MortezaPouraghdam and Strauss (2017). The application part of the study (*Mapping the Effect of Selective Attention between Ongoing EEG Activities and Averaged ERPs by means of Instantaneous Phase Information, 2019*) is in preparation for submission to the journal of *IEEE transactions on Neural Systems and Rehabilitation Engineering*. The study corresponds to the study in section 2.4.

3.4.1 Validation of Model on Synthetic Data

In Fig. 3.19, we show an example of an amplitude modulated signal with a phase reset at the time step $t = 2.18s$. The signal has been generated according to

$$X(t) = \begin{cases} [\cos(\omega_0 t)^2 + \epsilon] \cos(\omega_1 t) & \text{if } t < u \\ [\cos(\omega_0 t)^2 + \epsilon] \cos(\omega_1 t + \pi) & \text{otherwise} \end{cases}. \quad (3.3)$$

The signal generated after the time step u is shifted for π , hence the time step u is the actual change point. The parameters ω_0 and ω_1 have different frequencies, such that one of them corresponds to the generation of a lower amplitude signal. The shift after the time step u will lead to a phase reset that is related to the signal and we are interested to track this change over time. The other phase jitters which we aim to remove correspond to the the low envelope of the signal. Fig. 3.19(b) shows the estimated phase before and after applying the KS. Before applying KS, there are two phase jitters due to the low envelope and the actual shift in the signal. After applying the KS, the jitter corresponding to the low envelope has been diminished. In addition, the standard deviation of the estimated phase is an indicator for the degree of reliability of the phase jitter. A low standard deviation in the estimated signal phase indicates that it is less likely that the phase variation is due to the low envelope, whereas a high standard deviation indicates a higher likelihood that the phase variation has been generated due to noise. The details of the standard deviation estimation has been given in Appendix.G. In this example, the standard deviation from the sample $t = 300$ increases, indicating a high uncertainty due to the artificial phase reset.

To assess the accuracy of the model, we generated 500 synthetic data-series with different levels of added noise and random shifts in the signal. For every number of change points n , we randomly shift the signal at n random time steps u between $\frac{\pi}{8}$ and $\frac{7\pi}{8}$. This generates n instances at which the signal has been shifted, and the other phase variations correspond to either the noise or low envelope of the signal (see Fig. 3.19).

We apply a KS to remove the noisy variations in the signal. To get an estimate of how reliably

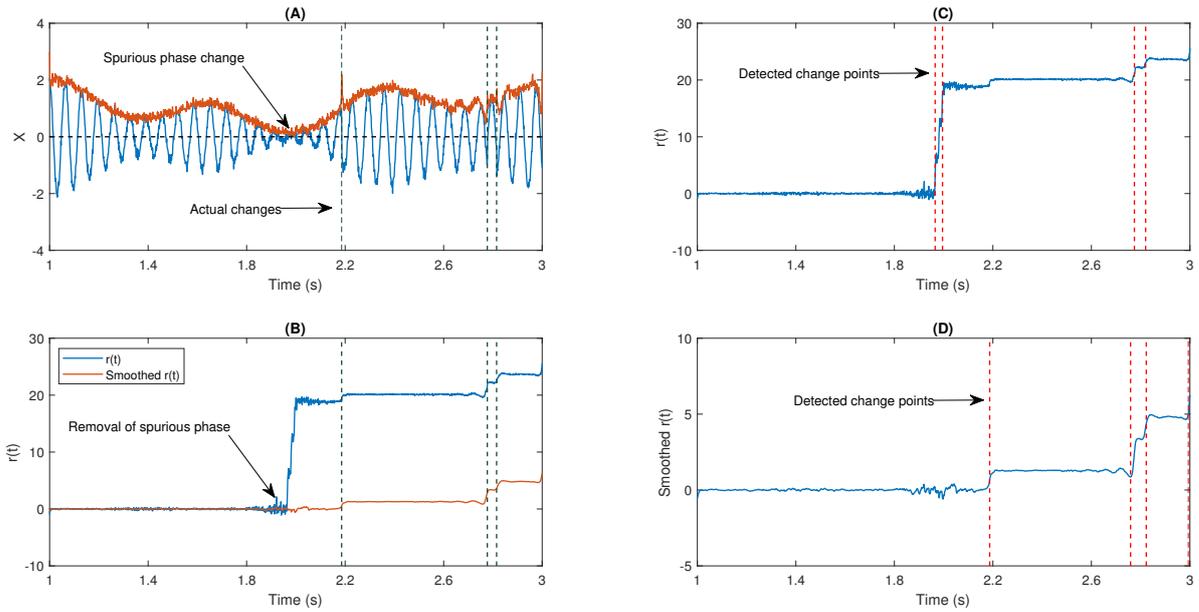


Figure 3.19: (A): An example of an oscillatory signal which contains a low envelope. The signal length is about 2s. The vertical lines show the actual shifts in the signal which are random. (B): The phase after removing the line corresponding to the center frequency before and after applying a Kalman smoother (KS). (C) The detected changes in $r(t)$ without any smoothing have been plotted. (D) The detected changes in $r(t)$ after applying KS.

the change points have been removed, we apply a change point algorithm to detect the time steps at which the mean of the signal has been significantly altered¹. The detected time instances are recorded as the estimated significant change points before and after applying KS (See Fig. 3.19). If the *difference of the estimated change points and the actual random change points* are less than 10 samples, the estimated change point is assumed to be correct (*true positive, TP*). However, if the difference is larger than the determined threshold, the data point has been *falsely recognized* as a change point (*and is referred to false positive, FP*). If the change point algorithm fails to detect the actual change point, then the point is referred to as *false negative, FN*.

In Fig. 3.20, the average number of false positives (fp) and false negatives (fn) for different number of change points have been plotted. For every number of change points n , we generated 2000 batches of data with different levels of noise and reported the average number of FN and FP. *After applying KS, the average number of FPs is significantly reduced.* However, the case with no filtering yields very unstable results as the noise level increases. The average false negatives is however lower for the case that we apply no smoothing compared to results after KS. This is mainly due to the fact that more random changes are detected in the pre-smoothing condition. Therefore, as many indices will be assigned correctly as change points, satisfying the minimum distance criteria. In the case of smoothing, the overlap of a change point and low envelope can cause an increase in the number of

¹The function *findchangepts* in MATLAB using the *mean* returns the time instances at which the mean of the data changes abruptly.

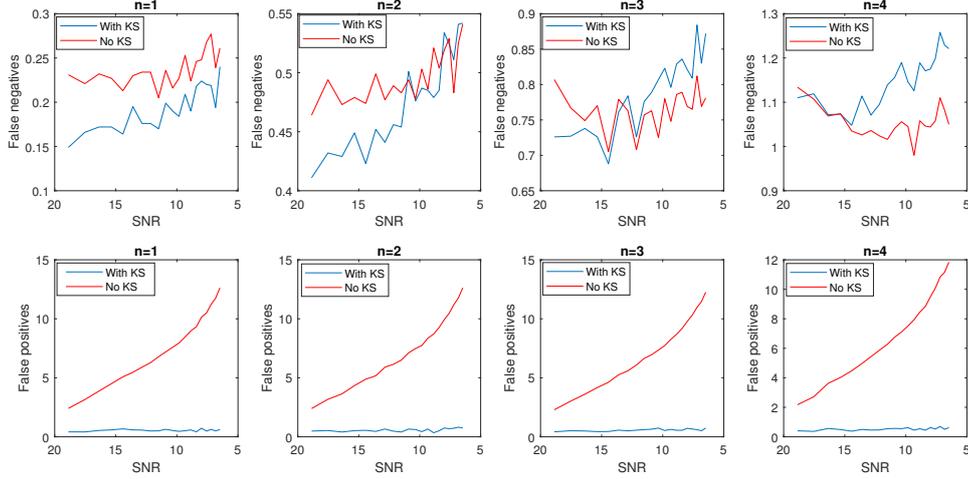


Figure 3.20: Average number of false positives and false negatives for 2000 batch of synthetic time-series signals with length of 1500 samples. The number of actual changes in the signal are indicated as n and results are shown for an increasing SNR.

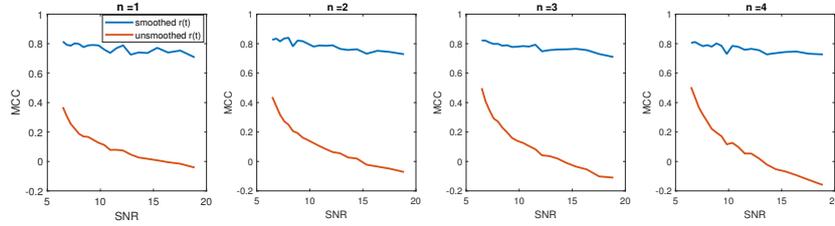


Figure 3.21: The average Matthews correlation coefficient for different number of SNRs and different number of change points.

detected false negatives.

Using the measured rates of false negatives and positives, we computed the *Matthews correlation coefficient*, *MCC*. MCC takes the number of false/true positive/negatives and returns a correlation coefficient between the observed and predicted binary classifier. It is computed as

$$MCC = \frac{TP.TN - FP.FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}. \quad (3.4)$$

A coefficient of +1 indicates a perfect prediction (i.e., in our case a perfect detection of change points at the correct indices), zero indicates no better than random assignment of change points and a coefficient of -1 means a complete disagreement between the predicted change points and the actual ones. Fig. 3.21 shows the average MCC for different SNRs for varying change points n . The MCC significantly decreases as the SNR increases (and noise level decreases), indicating a random assignment of change points. This is also consistent with the average results of falsely detected change points. As the number of falsely detected change points increases, we have a more random assignment of change points. This is however not the case for the post-KS condition. The higher average of MCC indicates a significant improvement in the accuracy of detected change points.

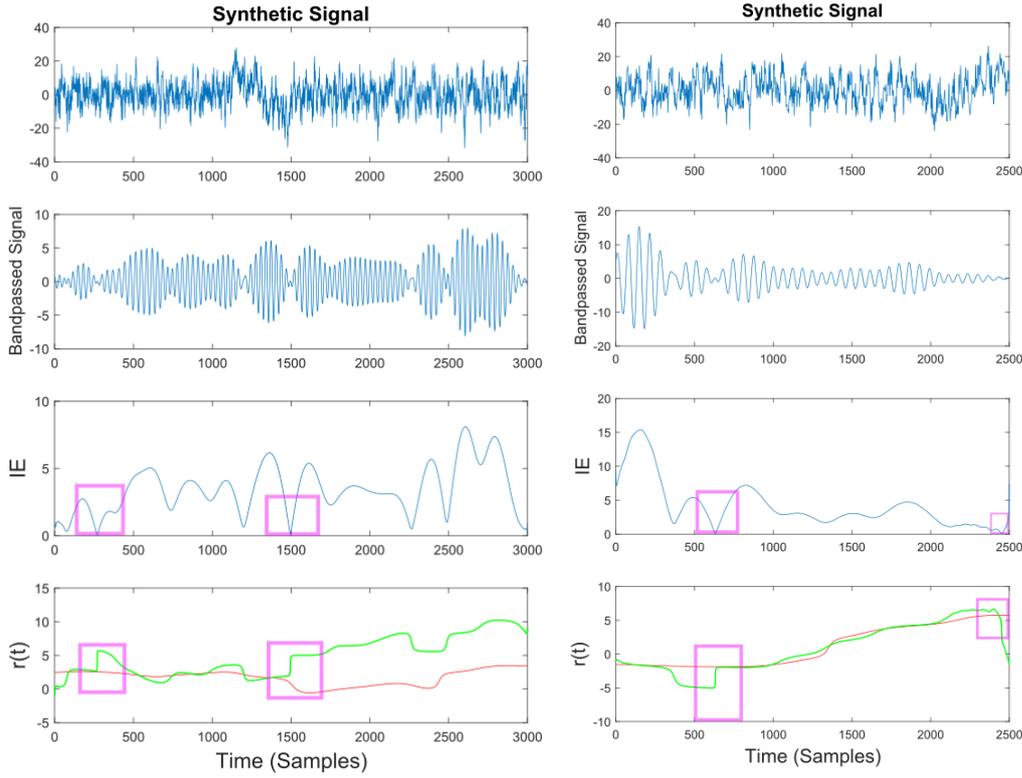


Figure 3.22: (A): Synthetic EEG signal generated at a $f_s=250\text{Hz}$ with a SNR of 0.11 (high noise variance) bandpassed at $f_c=7.6\text{Hz}$. Different noise amplitudes have been used for the two examples. The instantaneous envelope (IE) has been plotted with red boxes showing the regions with a low envelope below 0.2. In the last plot the resulting $r(t)$ has been plotted for the smoothed (red) and non-smoothed (green) signal. The red regions show the effect of the smoothing on the regions with a low IE. (B): Same description as in (A) applies to (B) with a SNR of 0.13.

Applying KS on Additional Synthetic Data

In this section we generate additional synthetic signals for testing the effect of spurious phase variations. The synthetic signals are generated as a superposition of sinusoidal signals and noise with different amplitudes ². The signals has been narrow-bandpassed using a FIR filter to a center frequency of 7.6Hz. The goal is to analyze the effect of the KS on removing the spurious phase variation for regions where the envelope is low. In this case, we considered the IP of signal corresponding to envelopes below 0.2 to be noise, and therefore need to be removed.

In Fig. 3.22 and Fig. 3.23 we show different examples of synthetic EEG signals where at some instances the corresponding envelope of the band-passed signal approaches zero. The regions corresponding to the low envelope has been illustrated in a red box. We show how applying the method with proper set of parameters can remove the spurious changes in the IP of the signal. The main incentive is to remove the spurious variations in $r(t)$. As we are considering the phase information

²The code for generating the signals has been mainly used from <https://github.com/pchrapka/phasereset> with slight variation

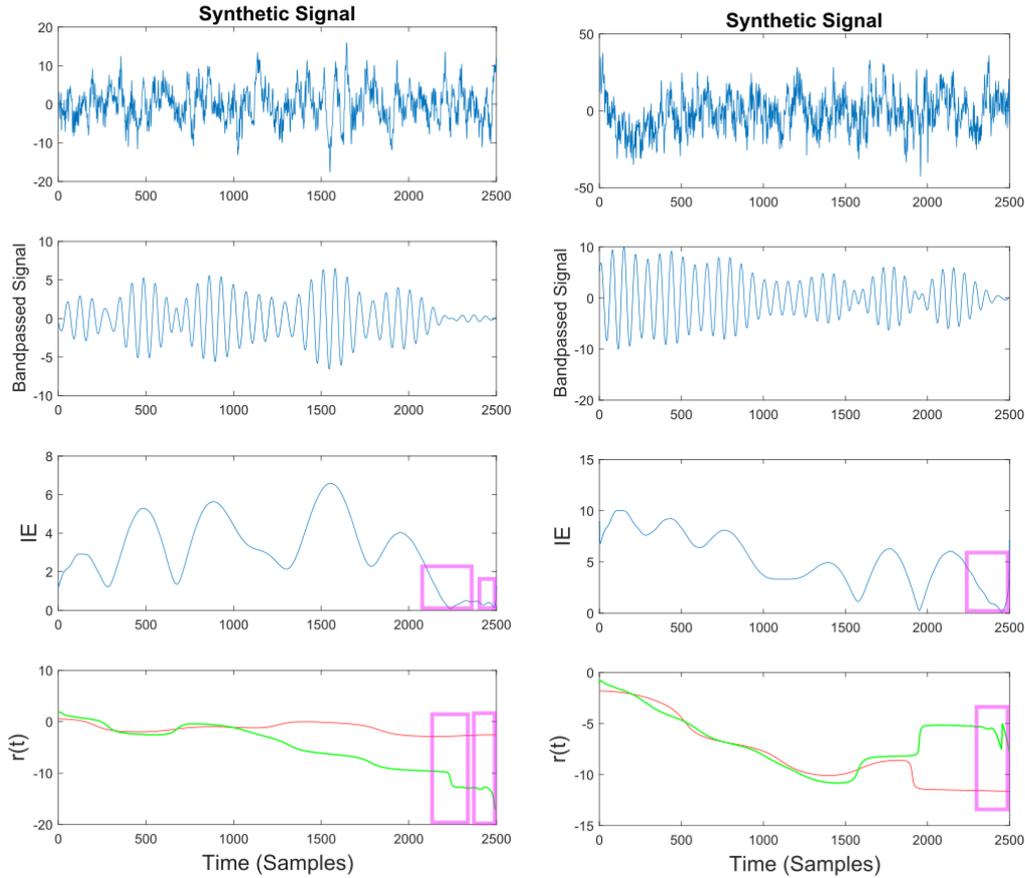


Figure 3.23: (A): Synthetic EEG signal generated at a $f_s=250\text{Hz}$ with a SNR of 0.43 bandpassed at $f_c=7.6\text{Hz}$. The instantaneous envelope (IE) has been plotted with red boxes showing the regions with a low envelope below 0.2. In the last plot the resulting $r(t)$ has been plotted for the smoothed (red) and non-smoothed (green) signal. The red regions show the effect of the smoothing on the regions with a low IE. (B): Same description as in (A) applies to (B) with a SNR of 0.047.

in a narrow-bandpass, we are required to filter the data accordingly. In chapter 4, section 4.4.1 we describe in more detail on the choices of parameter setting.

3.5 Analysis of IP of Ongoing EEG Measurements

In section 3.4.1, we describe the validation of the proposed model on synthetic data. In the following sections, we study the effect of selective attention on ongoing EEG oscillations as well as how the proposed approach can be incorporated for a more reliable study of measured data.

3.5.1 Filtering procedure & IP Extraction

The measured electrical activities from the right and left mastoids are bandpass-filtered between 1-70Hz. As described in section 2.2.1, to obtain the IP, there are two main steps. First, the signal is

narrow-bandpassed and secondly, the IP of the narrow-bandpassed signal is estimated. We compute the IP using the analytic form of the signal with the Hilbert transformation. To have a meaningful interpretation of the IP, the signal is narrow-bandpassed around a certain center frequency of interest. A zero-phase forward backward IIR elliptic filter is used to narrow-bandpass the signal. As noted in Seraj and Sameni (2017), an IIR filter requires a much lower order than a FIR filter and a zero-phase forward backward ensures a zero-phase distortion.

In Sameni and Seraj (2017); Seraj and Sameni (2017), it is shown that slight variations to the filter parameters in filtering process can lead to variations in IP and IF response. The reasons for phase slipping effect has been thoroughly discussed in section 2.4.1 and Seraj and Sameni (2017); Sameni and Seraj (2017). Therefore, a robust estimation method that estimates the IP from the average ensemble of infinitesimal perturbations to the filter parameters is presented by Seraj and Sameni (2017).

To estimate the IP, we apply a narrow-bandpass filter with slight variations in the frequency range for $M = 100$ iterations. At every iteration, filter parameters are as follows: the filter order is 6, the reduction in the stop-band is 50dB and the ripple-passband is 0.01 (the changes are very small that the effect is irrelevant for study of most physiological effects). The other set of parameters is selected based on quantifying the level of selective attention in attended and non-attended cases which is described in section 3.5.3. These parameters are the frequency and filter's bandwidth. The set of center-frequencies that have been tested are $f_c = (6.8, 7, 7.4, 7.6)$ Hz and the set of bandwidths are $bw = (0.1, 0.25, 0.4, 0.5)$. At every iteration i , the IP is computed for different parameter settings f_c , bw and KS factors. The KS factor is explained in the next section.

3.5.2 Setting of KS factors

The main parameters in KS are σ and α and *determine the degree to which we can rely on the measurements*. The lower the factor, the lesser the trust we put on the measurements; therefore the effect of the model prediction in computing the phase is higher. In other words, the factor is an indicator of the noise level in the data (SNR).

The effect of various KS factors is shown in Fig. 3.24 in a segment of EEG data. The signal is first band-passed at a center frequency of 7Hz and plotted in Fig. 3.24(b) along with the envelope of the signal. The envelope approaches near zero at some time samples, resulting in abrupt change points in the residual that are plotted in (c). The gray lines are the residuals corresponding to filters with different parameters that are slightly perturbed. In most cases, the residuals contain many abrupt jumps due to low envelope values. The black line in Fig. 3.24 (c) is the residual after applying KS. To have a better view of how KS removes the spurious phase variations, the IF is plotted in (d,e,f). In (e) the IF of one of filters has been plotted. The comparison between the IF after applying KS and before KS has been shown in (f).

The way we assign the parameters α and σ for measurements is as follows: We set $\hat{\mathbf{X}}_t$ to be the

average over all M filters as described in section 3.5.1. We estimate α using the variance of $\hat{\mathbf{X}}_t$ over different filter parameters. The parameter σ is estimated as β times the variance of $\hat{\mathbf{X}}_t - e^{i\omega_0} \hat{\mathbf{X}}_{t-1}$ where β is a free parameter that controls the degree we rely on the measurements over the predictions. Similar to section 3.5.1, we search for a range of β values that would maximize the difference between the attended and non-attended channels in our study. The β parameter range that we choose is between $1e^{-4}$ and 100 and equally spaced into 150 values.

3.5.3 Quantification of Attentional Effort In Experimental Setting

After applying KS, we quantify the degree of attentional effort using a measure obtained from Corona-Strauss and Strauss (2017). The measure is

$$\Lambda \propto 1 - e^{-TR^2} \quad (3.5)$$

where R is the resultant mean vector length of IP samples and T is the total number of samples in the signal. The measure maps the resultant vector length R to an exponential function that is bounded between 0 and 1. A high value of Λ indicates a higher attentional effort. The objective is to determine the level of attentional effort between the attending and non-attending conditions using the ongoing EEG.

3.5.4 Optimization of Filter and KS Parameters

We therefore apply the measure in Eq.3.5 for the extracted IP information of ongoing EEG signals with different parameter settings of bw , f_c and β . We select the settings that maximally distinguish between the attending and non-attending ongoing EEG signals for the cases of (attL_{ipsi} vs. IgnL_{ipsi}) and (attR_{ipsi} vs. IgnR_{ipsi}). The setting attL_{ipsi} corresponds to the electrode located on the same side as the stimulus arrival. This is compared against IgnL_{ipsi}, the case of non-attendance. For consistency, we measure ongoing activities from the same side as in attL_{ipsi}. Similar explanation is given for comparing attR_{ipsi} versus IgnR_{ipsi}, measures taken from the right side. In the following algorithm (algorithm 2), the main steps of optimization process are written. The set of parameters that maximized the difference between the attending and non-attending cases are included in Table.3.2.

In Fig. 3.25, the averaged Λ for different signal lengths is shown. The results show a higher Λ in the case of attending condition compared to the non-attending case. More explanations regarding the effect of selective attention and the mapping of segmented ERP onto ongoing EEG is described in chapter 4.

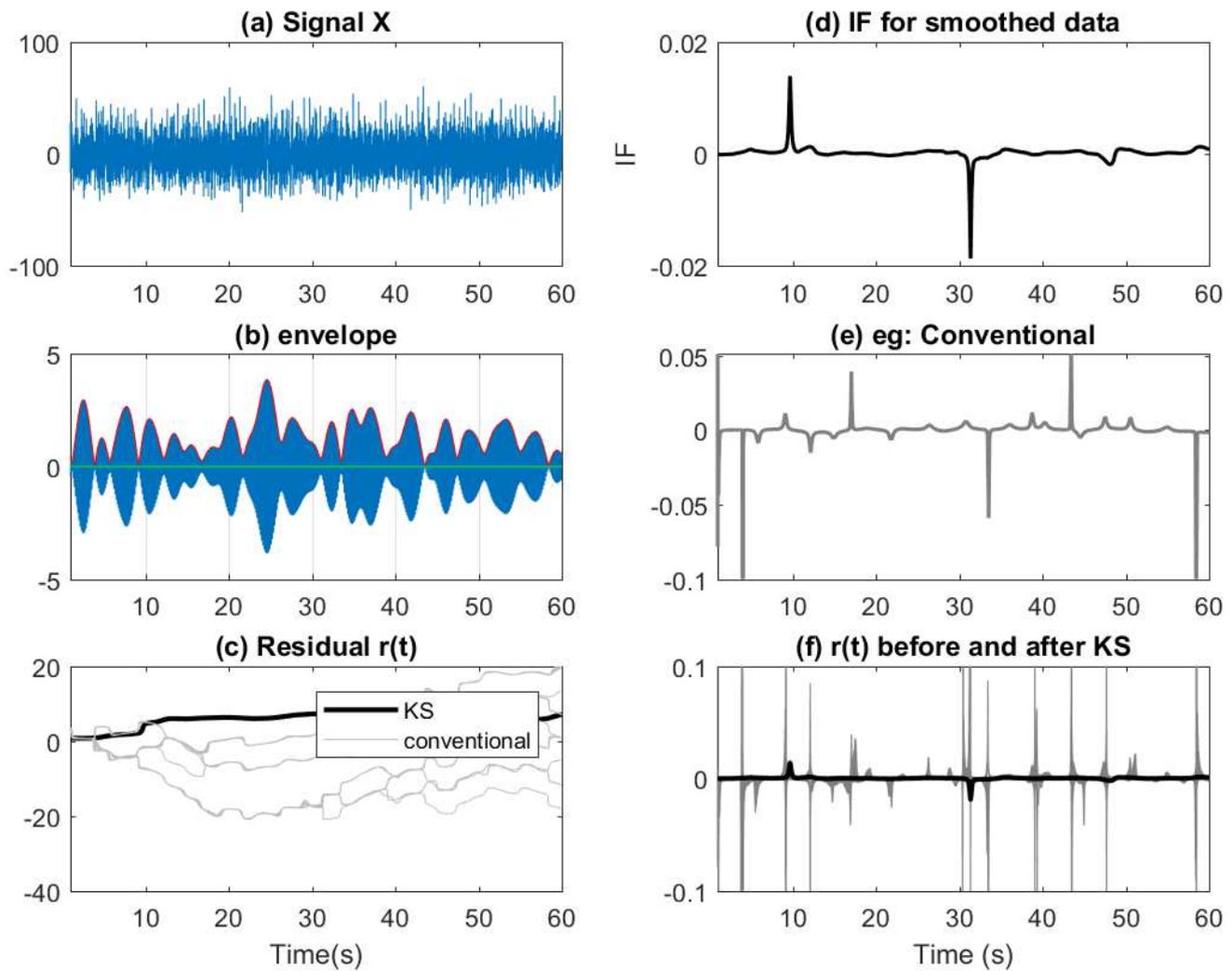


Figure 3.24: (a): A 60 seconds segment of measured EEG data. (b): Data has been narrow band-passed with a center frequency of $7Hz$. The envelope of the analytic band-passed signal has been plotted in red. (c): The residual of IP before and after applying KS has been shown. The term conventional corresponds to the filtered signals without any KS processing. As described previously, for every filter, a slight perturbation has been applied to the data. (d): The IF of the smoothed data. (e,f): Comparing the IF after and before applying KS. The changes in regard to low envelopes in (b) can be observed as fewer abrupt changes after applying KS.

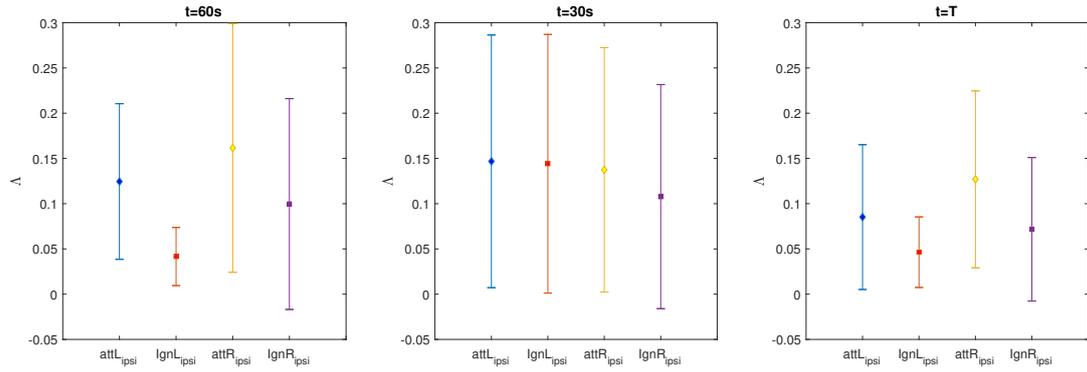


Figure 3.25: Average value of Λ over all subjects for different lengths. The terms $attL/attR_{ipsi}$ correspond to the same side as the stimulus. The conditions $IgnR/IgnL_{ipsi}$ correspond to non-attending case.

Input:

$\mathbf{X} \in \mathbb{R}^{1 \times T}$: Time series to be analyzed

M : Total number of filters

$\{filterorder, rp, sb\}$: Filter parameters such as filter order, ripple passband and stop-band.

bw : Filter bandwidth

f_c : center frequency

β : Kalman smoother factor

Output:

$\mu_{1:T}$: Estimated signal

$V_{1:T}$: Estimated variance

for $j = 1$ to M **do**

$x_j = filter(\mathbf{X}, filterorder, rp, sb, bw, f_c, \beta)$

$y_j = HT(x_j)$, Apply Hilbert Transform

end for

$\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_M]$, $\mathbf{Y} \in \mathbb{R}^{M \times T}$

$\hat{\mathbf{y}} = mean(\mathbf{Y})$, $\hat{\mathbf{y}} \in \mathbb{R}^{1 \times T}$

$\boldsymbol{\mu} = Kalman - Smoother(\hat{\mathbf{y}})$

$\Lambda(\boldsymbol{\mu})$, Compute the Selective Attention Measure

Algorithm 2: The main procedures for applying the Kalman smoother algorithm. The Kalman-Smoother corresponds to the function defined in Alg.1, *filter* corresponds to any suitable filtering procedure given the filter parameters.

	$f_c(Hz)$	bw	β
attL_ipsi, IgnL_ipsi	7	0.4	69.28
attR_ipsi, IgnR_ipsi	7	0.4	67.07

Table 3.2: The set of parameter values that maximize the differences between the attending and non-attending cases.

3.6 Relating AER to Laminar Phase Dynamics

As described in Chapter 2, section 2.5, we are interested in understanding the mapping between thalamo-cortical layers with the measured AERs. A better understanding of the projections of the cortical neural activities onto the late auditory ERPs can enhance our understanding of the underlying neural circuitries responsible for the generation of AERs, as well as experimental utility. **The results of this study have been published in Mortezapouraghdam et al. (2015b).**

3.6.1 Data Processing

To remove unwanted frequencies, we apply a bandpass FIR filter with 500 taps to the recorded raw LFP data for all electrode measurements $e_i, i = 1, \dots, 10$. We apply a bandpass filter of 2 Hz - 100 Hz for our further analysis. In our case, we applied additional filtering to remove frequencies of 50 Hz and its harmonics using a notch filter. This was necessary because our measurement setup introduced high noise levels at this specific frequency.

For more efficient processing, we subsequently down-sampled the data to two times the high cut-off frequency of the bandpass filter (200 Hz). Note that no information is lost in this step, according to the Nyquist-Shannon sampling theorem. We then segment the signal into trials of 1.5 seconds post-stimulus with respect to the trigger signal. Given the new sampling frequency of 200Hz, we obtain $T = 300$ samples per trial. For a given subject, we denote the trials as $\mathbf{q}_k^T \in \mathbb{R}^T$ for every trial $k = 1, \dots, N$. We represent the $N \times T$ ERP image as $\mathbf{Q} = (\mathbf{q}_1, \dots, \mathbf{q}_N)^T$ where $\mathbf{q}_k \in \mathbb{R}^T$ is the band-passed filtered post-stimulus trial k that appear as a row in the matrix.

In Fig. 3.26 we show the average over all trials at every electrode for different cortical layers as well as the averaged auditory ERPs. We show the activities for the first 300ms. The blue region correspond to the regions where the peak latency occurs about 35ms ~ 40ms post-stimulus for the measurements at the TN up to ALR. The first two post-stimulus peaks for averaged ALRs are observed at 20ms (positive voltage) and 45ms (negative voltage) post-stimulus.

To compute the IP of the band-filtered trials $\mathbf{q}_k \in \mathbb{R}^T, k = 1, \dots, N$ we employ a continuous wavelet transform. The wavelet coefficients are computed as $\omega_{k,b} := \langle \mathbf{q}_k, \psi_b^a \rangle_{L^2}$, $b = 1, \dots, T$, where $\psi \in L^2(\mathbb{R})$ is the wavelet function with $a \in \mathbb{R}$, $a \neq 0$ being the scale parameter. The wavelet scales are selected such that we adequately sample from the full signal frequency band. Specifically, we selected 40 scales such that the corresponding wavelet pseudo-frequencies lie in the boundaries of $[f_1, f_2]$, where f_1 and f_2 are slightly below and above the bandpass filter cut-offs of 2Hz and 100Hz respectively. The 40 scales are spaced logarithmically within these bounds.

By applying the wavelet transformation (see chapter 2, section 2.2.3) over the range of different scales and discretized translations $b_m (m = 1, 2, \dots, T)$, we obtain the complex entries of $\omega_{k,b} = \text{Re}(\omega_{k,b}) + i\text{Im}(\omega_{k,b}) = |\omega_{k,b}| \exp(ip_{k,b})$ for every scale a . The phase information for every electrode e_i at different scales (corresponding to frequencies between f_1 and f_2) is denoted as a $N \times T$ matrix

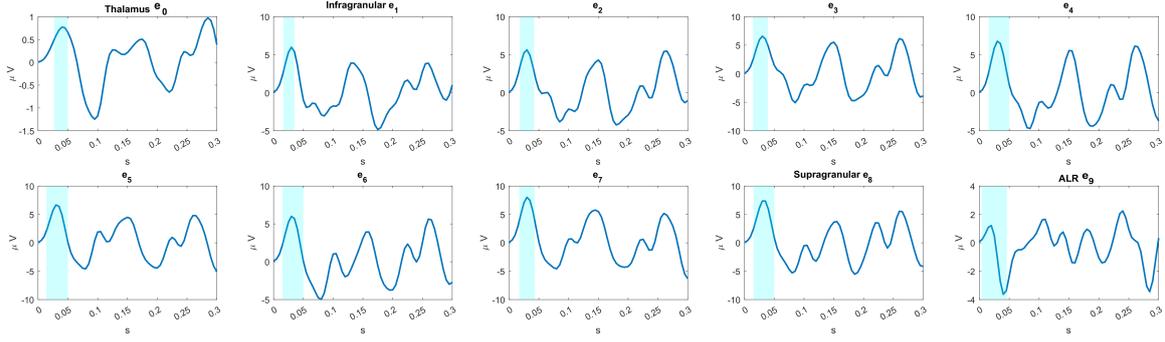


Figure 3.26: The average results over all post-stimulus recordings for electrodes from the thalamic nuclei (TN) to the averaged ALR. The electrodes e_1 to e_8 correspond to the electrodes inserted from the infra-granular layer up to the supra-granular layer. The first peak latency for recordings TN to e_8 are at $35ms \sim 40ms$ post-stimulus. The first two peaks of ALR occur at $20ms$ and $45ms$ post-stimulus. The results are smoothed with a Gaussian filter of size 5.

$\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N)^T$, where \mathbf{p}_k is a row vector corresponding to trial k . Furthermore, we applied a variational denoising phase algorithm for every two dimensional phase matrix (Bergmann et al., 2014). For more information see (Bergmann et al., 2014).

3.6.2 Interaction Between Cortical Layers and AERs Using MI

In Fig. 3.27 we show the MI between cortical layers (excluding e_0 and e_9) at different frequency domains for one of the measurements. The frequency ranges are divided as follow: the delta band corresponds to the range of frequencies below 4 Hz. The theta band correspond to the range of frequencies between 4 and 7 Hz. The alpha frequency is between 7.5 and 12Hz. The beta band is divided into the lower, middle and higher rhythms, corresponding to 12.5 - 16, 16.5 - 20 and 16.5 to 20 Hz respectively. Frequencies above 30Hz are considered in the gamma band. The averaged results in Fig. 3.28 over all measurements at seven distinct frequency domains are also shown. The average results over the nine subjects indicate the existence of a high level information sharing at lower frequency domains (i.e., ≤ 7.5 Hz). Results are consistent with previously reported studies of Kayser et al. (2009) at lower frequency domains. We also examined the MI between electrode pairs that are located next to each other $I(\mathbf{P}_{e_i}, \mathbf{P}_{e_{i+1}})$. Fig. 3.30 shows the maximum MI between electrode pairs at every frequency domain. The MI was computed at different delays between pairs of electrodes and the maximum existent MI has been used for illustration. As the location of electrodes are not specifically determined, we refer to the results with respect to likely infra/supra-granular regions. We speculate on the results as follows: The pair of electrodes at the supragranular layer are expected to show a high level of MI when computed to one another due to the flow of information from the granular layer and thalamic input. This is mainly the case for layer II and III. The thalamic input to layer VI and projection from supragranular layer to infragranular layers V and VI leads to a high information sharing at this region.

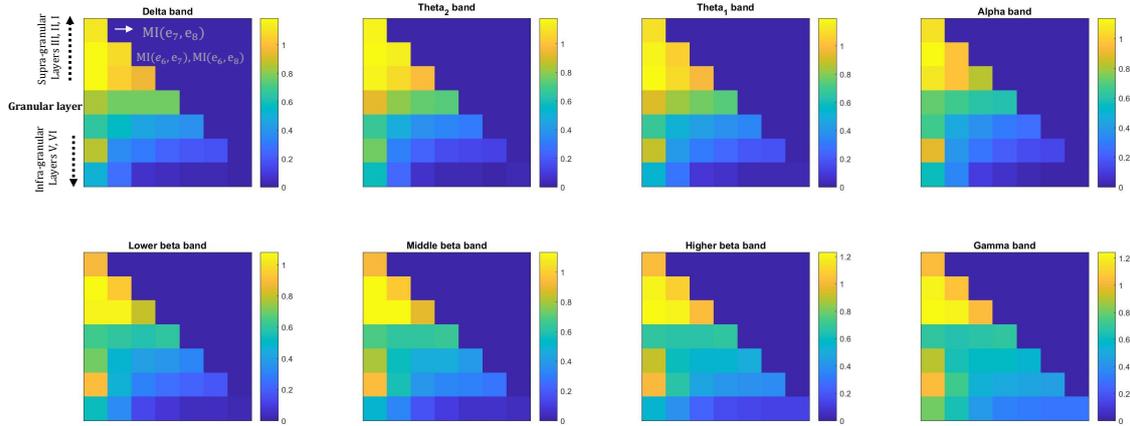


Figure 3.27: The MI between pair of electrodes from the supra-granular layers (i.e., e_8 , e_7) to infragranular layers (i.e., e_1) at different frequency domains. The frequencies are delta (≤ 4), theta (≥ 4 & ≤ 7), alpha (≥ 7.5 & ≤ 12), lower beta (≥ 12.5 & ≤ 16), mid-beta (≥ 16.5 & ≤ 20) and gamma (≥ 30 Hz).

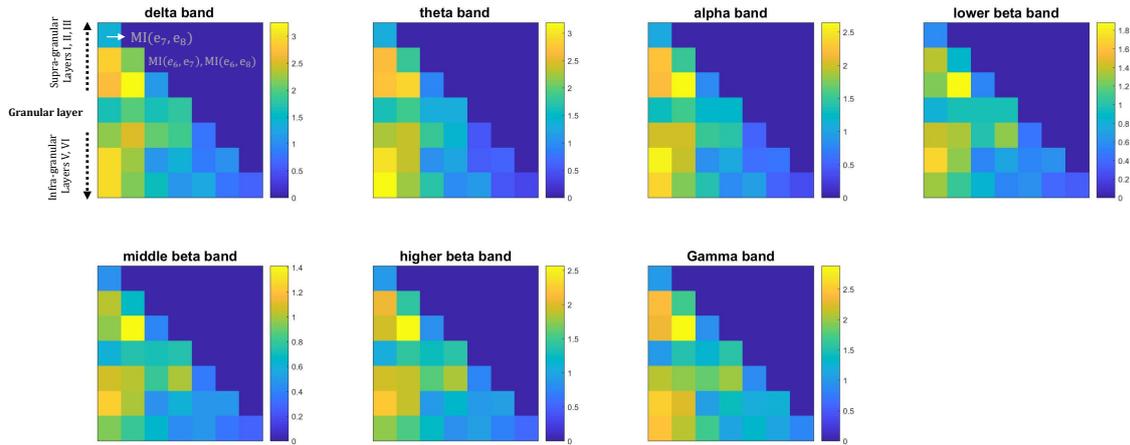


Figure 3.28: The average MI between electrode pairs at different frequency domains. Same explanation applies as in Fig.3.27.

To investigate the relation between the AERs and other cortical layers, we computed the average mutual information values over nine subjects at different frequency bands. Fig. 3.29 shows the MI among different electrodes. The pattern clearly suggests a significant level of dependency between the granular layer, supragranular layer with the measured AERs at different frequency domains. The distribution of MI between AERs and the granular layer appears to be highly correlated and repeated across different frequency domains (Delta, Theta rhythms).

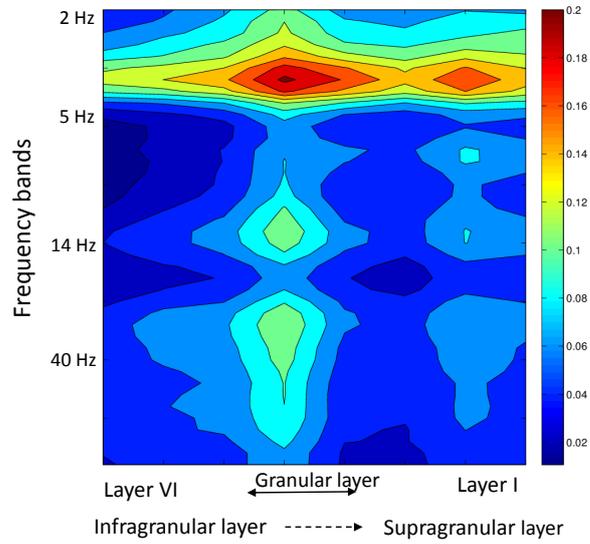


Figure 3.29: The average MI of the nine subjects between the outward-electrode (AER) and the cortical supra/intra-granular layers.

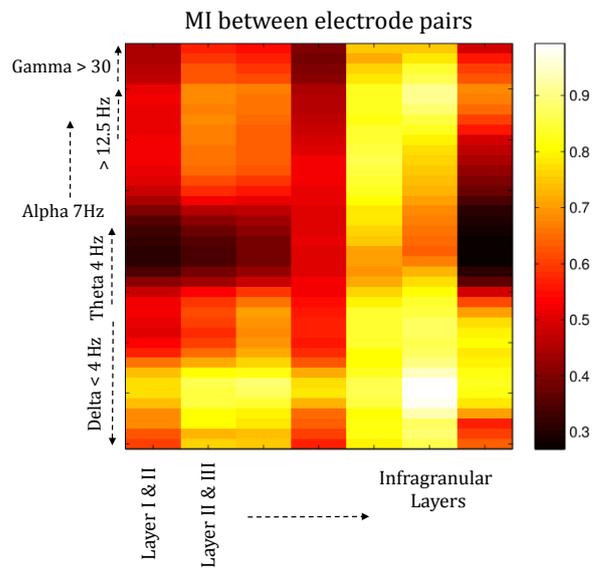


Figure 3.30: The average MI between subsequent pairs of electrodes at different frequency domains for nine subjects.

Chapter 4

Discussions & Limitations

In this section, we discuss the results, limitations and future works regarding the studies presented in chapter 3, section 3.1.

4.1 Tracking Changes in Long-term Habituation Over Trials

In the study presented in section 3.1, the time-course behavior in the concentration of simulated data (see Fig. 3.2) is related to the reduction of the CA1 integration time window (as in the physiological data of (Vinogradova, 2001)) and is a function of the number of subsequent stimuli. It is simulated using weakly coupled oscillators; see (Haab et al., 2011) for details. Thereafter, the temporal reduction of the integration time window, and thus the reduction of excitatory influence to the thalamus, is fitted with a sigmoidal function. We use this sigmoidal function in thalamocortical interaction model to simulate decreasing excitatory effects to the thalamocortical (G1) gain ((Haab et al., 2011)). No actual ERP data is used for fitting the limbic influence onto the thalamocortical circuitry. As Vinogradova (2001) invasively measured the spiking activity in the hippocampus of the cat, we observe a different slope behavior in the non-invasively recorded human habituation data, with the consequence of different time scales for small- and large-scale habituation. This result of two coexisting habituation timescales is consistent with Mutschler et al. (2010), hinting towards a hierarchical organized network of stimulus processing. The simulated habituation time-course bears a greater similarity to the reduction of the CA1 integration time window, as we largely neglect individual noise effects, such as, e.g., the voluntary shift of attention away from the target stimulus. The effect that we observe for habituation and non-habituation is consistent with the attention-load hypothesis by Lavie (1995, 2000). Our model is able to predict experimental trends in ERPs of saliently aversive and neutral stimuli. Furthermore, a parallel study examines the applicability

of the models predictions on medical conditions involving pathologic attentional binding, such as tinnitus.

The averaged estimated concentration results for the experimental data in Fig. 3.3 resulting from a 50dB(SPL) stimuli tend to be low and have a decreasing trend. The same effect is observed for the simulated data in Fig. 3.2. This is in contrast to 100dB(SPL) responses where the estimated concentration remains at a high level. To test the relation between changes in the estimated concentration value over trials for measured habituation and non-habituation data, we conduct a two-way ANOVA test. The chosen factors in our ANOVA study are the segments and conditions of the experiment (habituation and non-habituation). We want to understand the following cases: (a)- Does the distinction between habituation (50dB SPL) and non-habituation (100dB SPL) influence how the concentration between data segments changes? (b)- Do different segments have significantly different concentration values?

We compute the concentration values for each segment by fitting a von Mises model to both the 50dB and the 100dB case of all 10 subjects. We use a segment size of $Q = 200$ trials with $q = 0$ giving four segments. Since we are interested in *how the concentration changes over trials rather than in its absolute value*, we normalize each segment's concentration by dividing it by the concentration of the last segment for each given subject and habituation/non-habituation case independently. Since the last segment now always has a relative concentration of 1, we discard it for the ANOVA test, leaving three segments to consider. The two-way ANOVA test yields the following results: The first null hypothesis that the mean relative concentrations are equal for 50 dB and 100 dB is highly unlikely to explain the observed data at $p = 1.9 \times 10^{-5}$ ($F = 21.98$). The second null hypothesis that mean relative concentrations are equal across the first three segments has a very low likelihood as well with $p = 0.027$ ($F = 3.83$). At a significance level of 5%, we can confidently reject both of these hypotheses. Finally, the null hypothesis that there is no interaction between the two factors is also unlikely to explain the data, though at a weaker $p = 0.051$ ($F = 3.14$). The results of the two-way ANOVA support the findings in Fig. 3.3.

The confidence interval of the estimated parameters contains information about the precision of the fitted parameters. It could be related to how accurate the data variations are fitted to the model. The large CI in the case of habituation can be explained in terms of changes in the data modality over the trials. The phase information at the early segments is highly dense around the mean, having a pdf of Gaussian shape. A von Mises model can model such data accurately as it is similar to a Gaussian pdf and shares its unimodality property, whereas in the consequent segments, as phase diffuses, the von Mises model is no longer able to fit the data accurately.

The higher variations in CI of the non-habituation data can be interpreted in Fig. 3.3. Although the estimated concentration is significantly higher and its trend can be clearly distinguished from the habituation process, it does not follow a constant trend. From segments 60 to 80, the average concentration and the standard deviation of the estimated parameters slightly increase. There can

be numerous reasons that lead to such variations in the concentration parameter. One of them is the individual's variable response to an aversive stimulus. In the case of bimodal data, one single model may not completely explain the dynamics of the data .

4.1.1 Limitations:

As discussed previously, in the case of habituation, as the phase begins to diffuse in later segments, the data becomes more spread out on a unit circle. The distribution of the data has more heavy-tails. It has been previously reported that in cases of heavy-tailed or peaky pdfs, a wrapped Cauchy model can often fit data more accurately than a von Mises model (Vo and Oraintara, 2010). The improvement of results using a wrapped Cauchy model was not investigated in this study.

We investigate the multi-modality of the data by fitting a mixture of two von-Mises models. Under the assumption that the data $\Theta = \{\theta_1, \theta_2, \dots, \theta_T\}$ is i.i.d, we assume that the data comes from a *mixture of two von-Mises distributions*, each given by

$$f(\theta, \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp\{\kappa \cos(\theta - \mu)\}, \quad (4.1)$$

with $0 \leq \theta \leq 2\pi, 0 \leq \kappa, 0 \leq \mu \leq 2\pi$. We assume that with a probability of p_1 , a data point was drawn from the distribution f_1 , and with a probability of p_2 from f_2 . The likelihood of data samples is hence given by $L = \prod_{i=1}^T p_1 f_1(\theta_i) + p_2 f_2(\theta_i)$, where $0 \leq p_j \leq 1$ and $p_1 + p_2 = 1$.

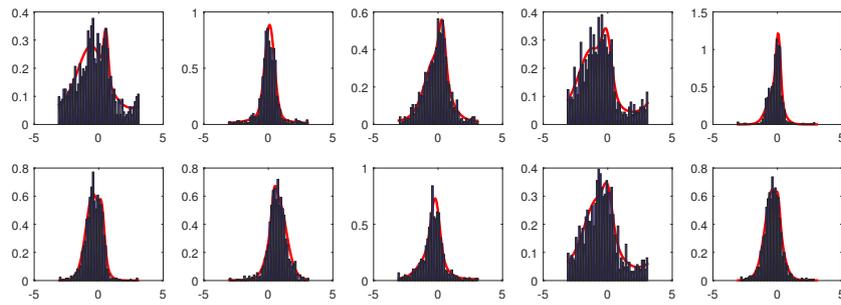
The mixture of two von-Mises distributions can be extended up to any arbitrary number of underlying von-Mises distributions. Given the mixture of two von-Mises distributions, we want to maximize the log of the likelihood function with respect to the parameters μ and κ . The log-likelihood function is given by:

$$l = \sum_{i=1}^T \log[p_1 f_1(\theta_i) + p_2 f_2(\theta_i)] \quad (4.2)$$

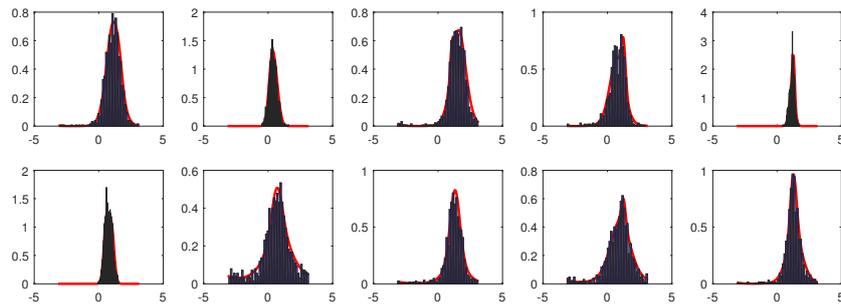
The optimized parameters are taken from the study in Banerjee et al. (2005), where the Expectation Maximization approach is used to estimate the parameters μ and κ . We therefore use the optimized μ and κ from (Banerjee et al., 2005) to study the fit of two von-Mises distributions to the data. In Fig. 4.1 we plot the resulting fitted model for all subjects.

The results obtained in Fig. 4.1 suggest that the changes in modalities between the soft and aversive stimuli are different. This can better be shown in the case of a soft stimulus of 50dB SPL, where data distribution undergoes more changes. In Fig. 4.2 we plot the changes in the IP information over different segments. Different colors correspond to the samples assigned to either f_1 or f_2 .

The observations regarding the model fittings in Fig. 4.1 and Fig. 4.2 illustrate that the changes in the aversive stimulus are more or less unimodal. However, there can be changes over the trials during the stimulus exposure. The changes are however more relevant and observable in the case of



(a) 50dB SPL



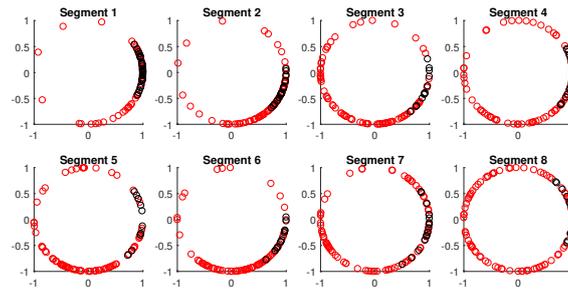
(b) 100dB SPL

Figure 4.1: In (a) and (b) we show the fitted mixture of two von-Mises distributions for all subjects.

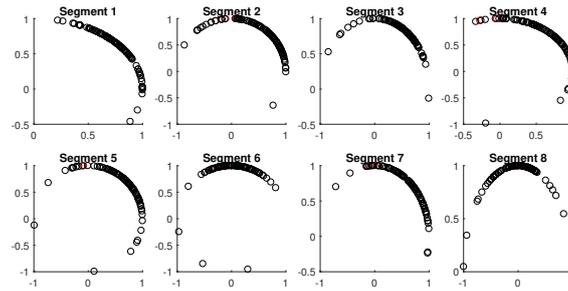
a soft stimulus. A better fit of a better fitting model to the data should be further investigated in related future studies.

Moreover, it must be noted that the neurofunctional model used for data-generation is a simplified model for describing attention-regulation. The physiological parameters used for the simulations are randomly taken from a set of physiological variables, such as, e.g., axon length and axonal velocity. The modeling results thus match the experimental group behavior but can hardly match single ERP dynamics, as those parameters are unidentifiable in the individual patient. In addition, due to unspecified background activities in the data, the outcomes in Fig. 3.3 are subject to more noise than the simulated data as in Fig. 3.2.

Another limitation of the current approach is determining the optimal window size for observing meaningful changes in the long-term habituation process. More precisely, the choice of the window size can be important in understanding the behavior of data in long-term. When the window size is too small, we end up over-fitting the data leading to high fluctuations in the concentration parameter in the time-course of data. A coarse window-size leads to a loss of time precision, makes it impossible to detect the exact time at which a change in the underlying data distribution occurs. This has been demonstrated in Fig. 4.3. It shows an example of habituation data and how incorrect windowing choice can misrepresent data behavior. In general, the method lacks the ability to detect gradual changes in the habituation process with a high temporal precision. In addition, it does not



(a) 50dB SPL



(b) 100dB SPL

Figure 4.2: In (a) and (b) we show the changes in the IP of assigned data samples for one subject fitted with a mixture of two von-Mises distributions. The two different colors, black and red, correspond to data samples assigned to models f_1 or f_2 .

incorporate our prior knowledge of long-habituation process in attentional circuits into the model that makes the results more robust against noise. This is the main motivation for applying different approaches which takes into account our prior knowledge of habituation and make the model more robust and precise in terms of detecting significant phase changes.

Furthermore, one of the main areas that is considered for future studies is the enhancement of results by studying the distortion effects of 2D filtering to the phase information. More specifically, the analysis of how noise-models translate from time-domain to phase can help us in developing sophisticated denoising procedures to reveal the systematic effect of habituation and non-habituation more reliably.

4.2 From the Abrupt CPM to Continuous Detection of Changes

Because of the limitations in our primary study, we applied a different method with higher temporal precision for detecting the changes in the signal. We aim to learn about measures that rely on the neural correlates of selective-attention and distinguish between the two states of long-term habituation and its absence with a high temporal-resolution. Therefore, the problem was modeled

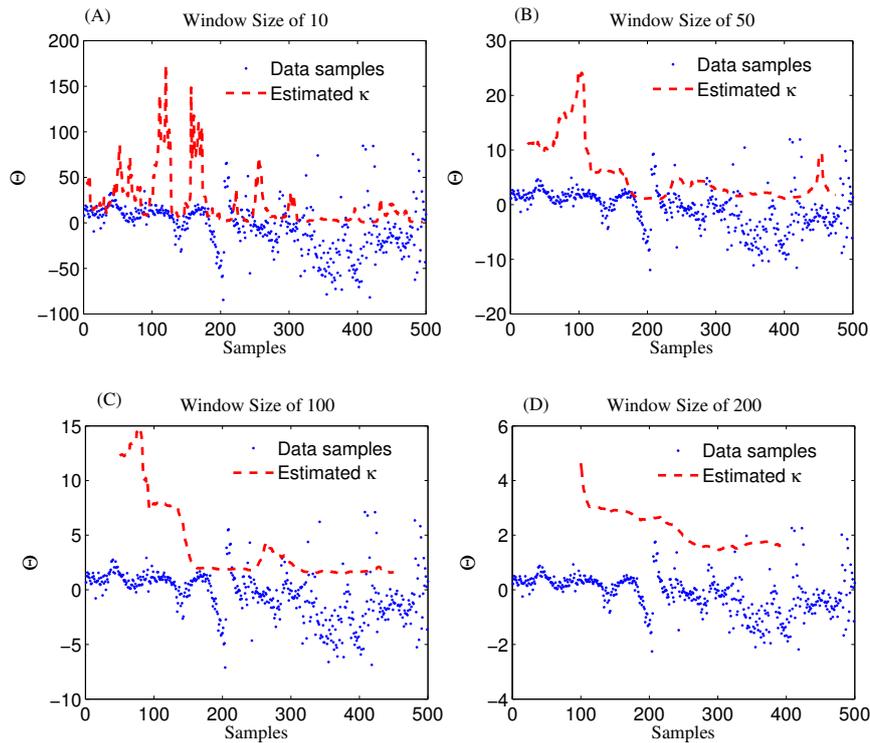


Figure 4.3: The subplots (A), (B), (C), (D) present the maximum likelihood estimate of the concentration over different window sizes. As illustrated, large window sizes provide stable results, but exact temporal information is lost and changes are difficult to detect precisely. On the other hand, a small window size tends to over-fit the data, and is sensitive to noise. In (F) we show the results of our forward-backward Bayesian model for data (E). The plot shows the likelihood of different concentration parameters in an underlying von Mises model at each sample. Note that at the later samples the likelihood of low phase concentration increases.

as a binary setting and the two main states are the presence and absence of a habituation over the experimental trials.

The abrupt change point models that were presented in Adams and MacKay (2007); Paquet (2007) were adapted for the detection of habituation in the data. The results of an abrupt change point model in section 2.3.3.2 fitted to the measurements corresponding to neural correlates of 50dB SPL and 100dB SPL stimulations were presented in Fig. 3.2.2. The main motive is to create a model that is able to detect one significant change point (i.e., a point) such that the data's statistics after this point significantly change. There are, however, complications in training a universal model for detecting such abrupt changes. One of the problems is determining the expected time at which habituation occurs. As the stimuli differences between 50dB SPL and 100dB SPL in terms of the level of hearing-comfort are large, we can expect an earlier habituation occurrence in 50dB SPL than in 100dB SPL. In 100dB SPL, the aversive intensity prohibits an early habituation whereas in 50dB SPL, we expect to observe habituation after the first few hundred trials, hence, for a 50 dB SPL a significant change is expected in the IP distribution (see Fig. 4.4).

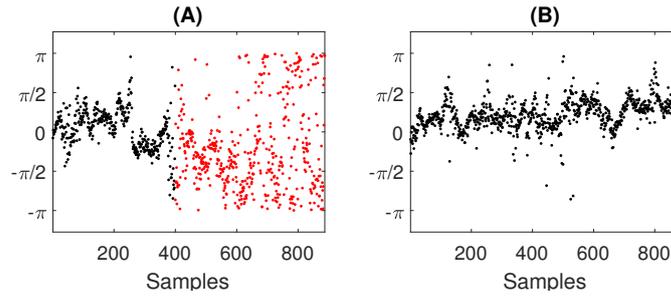


Figure 4.4: In the ACP approach, we search for a sample time that the underlying generative model is significantly changed. In this example (A) at about sample 400, data illustrates the effect of habituation. In the second data example (B), there is no significant change point.

The measurements suggest that the timing of long-term habituation is a highly variable factor, particularly in the 50dB SPL neural activities. We could obtain several change points in the run-length, and as they might not be near each other, it is difficult to determine an actual valid trial number that represents the presence of habituation. Therefore, we enforced the model to contain only one significant change point to increase the distinction between the two different stimuli in term of extracted features by applying a small hazard rate. The hazard rate of $H = 9.8e^{-12}$ is chosen empirically and the maximum and averaged run-lengths are selected to analyze the results (See Fig. 3.10 and the corresponding ANOVA test results).

The changes in the run-length are extracted using a Viterbi algorithm that computes the most likely state sequence of run-lengths. The details on the algorithm has been explained in section 2.3.3. Considering the most likely state sequences in the aversive stimulus, the run-length is on average higher for aversive stimulus (100dB SPL) than the soft stimulus (50dB SPL). This indicates the presence of a change point at earlier trials in the case of 50dB SPL than in 100dB SPL. The run-length changes more rapidly if a small hazard rate is applied.

4.2.1 Limitation With Abrupt Definition of Changes

One of the main difficulties with the ACP model is determining a meaningful time range in which habituation occurs. Although the model detects time at which concentration of data significantly changes, this partially depends on the enforcement of a small hazard rate. In addition, the discrimination between the 50dB SPL and 100dB SPL tones based on the run-length feature is not difficult. With a 100dB SPL stimulus, we expect to have an increasing run-length over the samples as the level of attentional binding stays the same throughout the experiment. With 50dB SPL, it is expected to have more variations in the run-length parameter. Exploring the run-length differences between stimulations that are not significantly distinguishable in terms of loudness (as an example 50dB SPL vs. 60dB SPL) is a more challenging task. This was detected for the two cases of 60 and 70dB SPL. It becomes difficult to interpret the changes in the run-length sequence based on different choices of the hazard rate. As more stimuli are added, setting the model parameters become more

difficult. In addition, when the expected number of changes is unknown, it is difficult to optimally detect all meaningful change points as the success of model heavily depends on a proper choice of prior change point (Adams and MacKay, 2007).

4.3 Continuous Definition of Long-term Habituation and Tracking The Changes

Due to some of the limitations discussed in section 4.2.1, we adopted a different approach to detect the changes in neural correlates of selective attention at different stimuli. As described in section 1.3.3, the drift in attention in long-term habituation is a *gradual process*. Therefore, instead of using an abrupt change point model, we use a gradual change point model that is able to capture the slow changes in long-term habituation. This way, we aim to classify between different stimuli that are not significantly distinguishable in terms of loudness. We propose a Bayesian change point model as described in section 2.3.4 for tracking the *overall gradual changes* in the neural correlates of selective attention. The objective of this study is to evaluate the degree of habituation effect using the IP of ERPs induced at different stimuli: of 60dB SPL, 70dB SPL, 80dB SPL and 90dB SPL. We use a Bayesian model to track the changes in the underlying concentration parameter of IP. In addition, we use the verbal responses of the participants about the loudness of different stimuli. This knowledge is used to validate the conclusion about the relation between the objective measure and the loudness scale at different stimuli (results are included in section 3.3.5).

Despite the high variability among subjects with regard to changes in the concentration states (see Fig. 3.17), the average concentration results in Fig. 3.18 suggest that *as the loudness level increases, it is highly probable that the degree of phase synchronization increases as well*. To validate the obtained results, we compare the averaged concentration level against the average of the verbal responses from the participants. As shown in Fig. 3.18 (b), as the stimuli level increases, the intensity of the loudness perception increases as well. This is consistent with the studies conducted by Hood and Poole (1966) and Stephens and Anderson (1971), that stimuli between 90dB SPL to 100dB SPL is considered uncomfortably loud in normal hearing subjects. Furthermore we applied a one way ANOVA test over different stimulus levels across all subjects to reliably distinguish between the 60dB SPL and 90dB SPL using the average of the expected concentration at a significance level of 5% with $p = 0.0101 (F = 7.37)$.

To test the effectiveness of the algorithm on experimental data and the significance of the results between 60dB SPL and 90dB SPL, we applied an additional test as follows: we generated a signal that consists of two parts, the first part corresponds to the data samples from the first half of the 60dB SPL and the second part of the signal contains samples from the second half of the 90dB SPL data. We applied the proposed forward-backward Bayesian model with the same empirical prior parameters as in 3.3.4 to check if we are able to detect the artificial change point between two

stimuli. In Fig. 4.5, we show a few examples of our observations. Throughout all observations, the model is able to track the distribution of concentration changes.

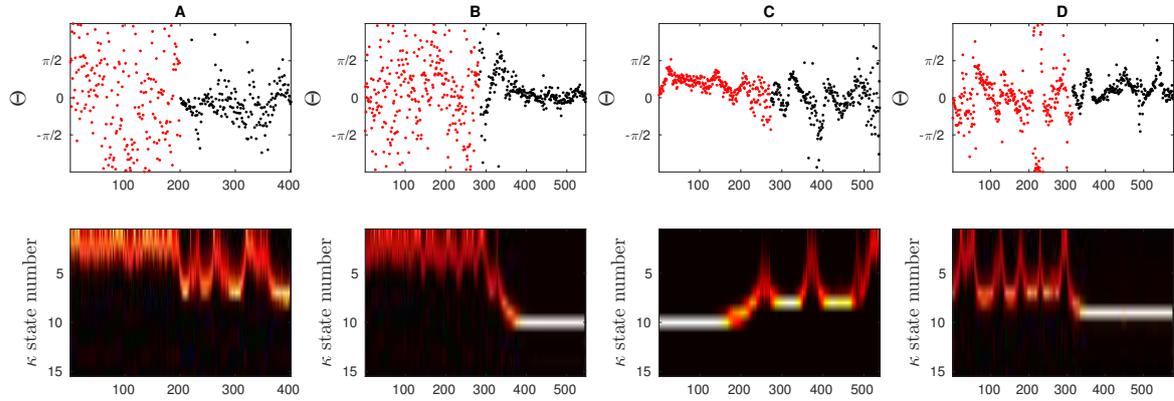


Figure 4.5: To show the effectiveness of the algorithm in significantly differentiating between 60dB SPL and 90dB SPL, we constructed a signal that is composed of two parts. The first half of the signal (in red) corresponds to the first half of a signal at 60dB SPL, and the other half (in black) corresponds to the second half of a signal at 90dB SPL. We run the proposed algorithm with the empirical prior parameters and check if the model is able to track the distribution of data in terms of the concentration.

In some of the tests it is more difficult to detect the artificial transition. This is mainly because the data in 60dB SPL and 90dB SPL can behave the same at different time intervals. If the samples in the first half of the 60dB SPL have lower concentration and the second half in 90dB SPL is in a similar state, then the transition may not be evident. The same argument holds when different halves of the signals have high concentration values. However, given the results and the additional test, we can confidently report on the significant difference between 60dB SPL and 90dB SPL.

4.3.1 The Rate of Change of the concentration Parameter

To test whether the stimulus level has an impact on the *degree of change* in the concentration over trials, we conducted a second ANOVA test as follows: We adjusted the subject specific factor in the concentration values by dividing the expected values of the concentration by the average of the expected value of the concentration of the last 50 samples (the resulting values are denoted as \mathbf{W} and see the corresponding algorithm in Appendix. F). The meaning of the resulting values can be understood as follows: In case of no change at all, all values in \mathbf{W} will be exactly 1. If the concentration increases for a subject at a given stimulus level throughout the experiment, values at the beginning of the experiment will be less than 1. A decreasing level of concentration corresponds to values greater than 1. We next compute the average of the expected concentration for every subject at different stimulus levels. This shows the *relative changes in the phase concentration over trials with respect to the last 50 samples*. A one-way ANOVA test yields the following results: The null hypothesis that the relative changes of the expected concentration is independent of the stimulus

level can be rejected at a significance level of 5% with $p = 0.0178 (F = 3.58)$.

4.3.2 Limitations

Despite the success in objectively differentiating between different stimuli by using the IP of the ERPs, and determine a significant difference between the rate of change in the concentration parameter among different groups, the variability in the dynamics of the IP among subjects is very large. Therefore, it is difficult to draw additional general conclusions, such as determining a general time at which habituation may occur for different subjects, or to conclude a unique uncomfortable sound level threshold among all subjects. Our results are also constrained by amount of data.

To demonstrate habituation variability, we describe the habituation effect for the 60dB SPL and 70dB SPL for subjects 1 and 2 in Fig. 3.17. The corresponding results for subject 2 show that the phase information is uniformly distributed throughout the experiment. As a result, the estimated concentration fluctuates rapidly between lower concentration states. However the habituation process in subject 1 is visible as a continuous decay in the phase concentration. The change in concentration is more significant in the first subject than the second one. However, based on the average expected concentration, we can conclude that in both cases, a weak level of attention has been allocated to the stimulus due to a low concentration value. This effect is mainly due to the variances of neural responses over the number of stimulus presentations. The same acoustic stimulus presented to a subject never elicits identical neural responses across a series of presentation. While this effect holds for a single subject (intra-individual variability), the effect is even more pronounced across a group of subjects (inter-individual variability) due to physiological variations in the neural architecture. Moreover our experimental paradigm cannot force the participants to ignore a presented stimulus. The information carried by the stimulus is not limited to its intensity. Associative processes in the individual subject can alter the subjective impact of a stimulus (see Busse et al. (2009)) and lead to a variation of time periods in which the subjects voluntarily pay attention to the presented stimulus. At this stage, with the current data and analysis tools, we believe it is not possible to describe a definite habituation onset in time which is solely determined by stimulus intensity.

Future Work There are several promising directions for future work:

- We can try to join similar repetitive state transitions into a single state and show only the state transitions with significant differences. This could be achieved by applying a discrete two state Hidden Markov Model on the results of the marginal value of the concentration parameter.
- To adapt the scale of the states with respect to the overall average of the expected concentration value (i.e., different resolutions). Applying such a mechanism will help us to track the behavior of a habituation in cases that a weak attention level has been allocated to the stimulus.

Additionally, one other area that can improve our results is denoising the phase information more

effectively as a pre-processing step. In this study the phase information was extracted after denoising the ERPs using a NLM algorithm. We suggest utilizing denoising methods directly on the phase data to reduce the noise level in the phase information.

4.4 Tracking the level of selective attention in ongoing EEG

In previous studies, we developed/employed methods for analyzing the ERs induced from repetitions of sound stimuli (pure tone beeps). Despite limitations in some methods or lack of data, we were able to mainly distinguish between the absence or presence of long-term habituation and explore different methodologies possible for analyzing IP. The signals that were processed are all segmented and the IP over trials at different times are considered for further study. However, in many studies, the stimulus is continuous or of duration and hence, it is required to analyze the ongoing activity of EEG signals without segmentation.

One of the advantages of studying the ongoing EEG signals as opposed to segmentation is avoiding spurious phases jumps or *phase-slipping*. It has been shown in Freeman et al. (2003); Sameni and Seraj (2017) that segmentation may introduce additional spurious variations due to the discontinuity that is introduced in the signal. In addition, the study of the phase is not reliable in phase-slipping epochs and there is no best way of segmenting the data. Therefore, to quantify the level of selective attention in ongoing EEG signals in this study, we consider one of the effects of spurious phase variation in signals as described thoroughly in section 2.4.1. We first describe some of the general strategies for setting the parameters α and σ .

4.4.1 Setting of the α & σ Parameters

As described previously in chapter 2, section 2.4.2, setting of α and σ are important since they are used for computing the Kalman gain factor. The factor determines how much do we rely on the measurements than the model. In Fig. 4.6, we show an example of the model dynamics with low noise in comparison to the measurements that have been produced based on Eq. 2.31 with constant noise over time. The model is having a more smooth path and avoids crossing the zero, whereas in the measurements, due to the noise and small envelope (approaching to zero), we can have jitters or abrupt variations in the IP or IF.

In the examples provided in Fig. 3.23, Fig. 3.23 and the ones corresponding to the synthetic sine waves, we have access to the original signal without noise. This is equivalent to having prior knowledge of the distribution or shape of the signal for real scenarios. However, in most cases, we don't have access to the actual distribution of the signal and hence a proper *estimation* of the parameters is required. In this section we give suggestions for setting the parameters of the KS.

As described in Eq. 2.31, the additive noise of the signal and measurement are defined as $\hat{\eta}_t \sim \mathcal{N}(0, \sigma)$ and $\hat{W}_t \sim \mathcal{N}(0, \alpha)$. By reordering $\hat{s}_{t+1} + e^{i\omega_0} \hat{s}_t + \hat{\eta}_t$, we have $\hat{\eta}_t = \hat{s}_{t+1} - e^{i\omega_0} \hat{s}_t$. Therefore the

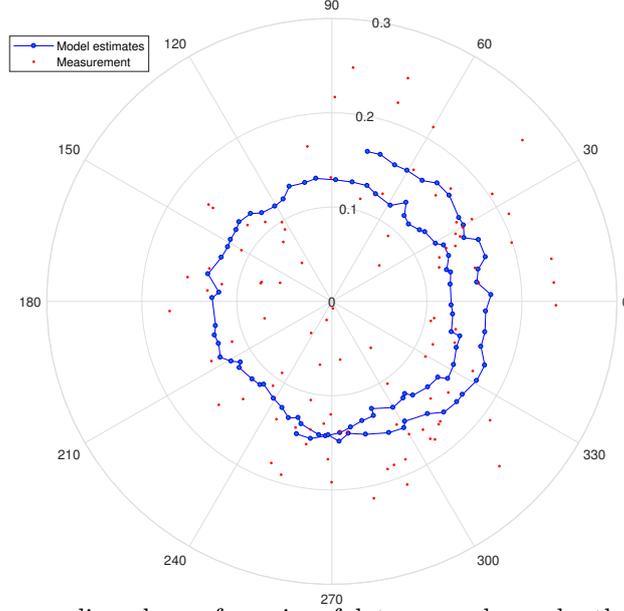


Figure 4.6: The corresponding phase of a series of data examples under the proposed model $\hat{s}_{t+1} = \hat{s}_t e^{i\omega_0} + \hat{\eta}_t$ has been shown in blue. The distance of the samples to the center is the envelope. The red dots show the resulting measurement ($\hat{x}_t = \hat{s}_t + \hat{W}_t$). Depending on the degree of smoothing, we can remove the areas at which the envelope of the measurements approach zero.

variance of noise of the signal σ can be estimated as the variance of the difference equation of the analytic signal. In case of real measurements, as we don't have access to the actual signal, we can estimate the variance of the noise as follows:

$$\begin{aligned}
 \hat{x}_{t+1} - e^{i\omega_0} \hat{x}_t &= \hat{s}_{t+1} + W_{t+1} - e^{i\omega_0} (\hat{s}_t + W_t) \\
 &= \hat{s}_{t+1} + W_{t+1} - e^{i\omega_0} \hat{s}_t - e^{i\omega_0} W_t \\
 &= \underbrace{\hat{s}_{t+1} - e^{i\omega_0} \hat{s}_t}_{\eta_t} + (W_{t+1} - e^{i\omega_0} W_t)
 \end{aligned} \tag{4.3}$$

As stated earlier W_{t+1} and W_t are samples from a Gaussian distribution with a variance of α , $\mathcal{N}(0, \alpha I)$. Due to symmetry of the Gaussian distributions, the sum of two normal distribution is a Gaussian with the summation of the means and the variances. Therefore, in our case, the resulting distribution will be $\mathcal{N}(0, 2\alpha I)$. By taking the variance of both sides in Eq. 4.3, we have

$$\begin{aligned}
 \text{var}(\hat{x}_{t+1} - e^{i\omega_0} \hat{x}_t) &= \text{var}(\eta_t) + \text{var}(W_{t+1} - e^{i\omega_0} W_t) \\
 &= \text{var}(\eta) + 2\alpha
 \end{aligned}$$

Hence the variance of noise of the signal can be estimated as $\text{var}(\eta_t) = \text{var}(\hat{x}_{t+1} - e^{i\omega_0} \hat{x}_t) - 2\alpha$. Setting $\epsilon = \text{var}(\hat{x}_{t+1} - e^{i\omega_0} \hat{x}_t)$, the optimal σ is in the range $\sigma \in [\max(\epsilon - 2\alpha, 0), \epsilon]$ where ϵ is the

upper-bound of our estimation. The input for EEG measurements is \hat{X}_t that is the average over all M filters. Another possibility for controlling the level of smoothing is to optimize over $\beta var(\eta_t)$ where β is a free parameter that controls the degree we rely on the measurements over the predictions. In the next section we emphasize more on the role of the β .

4.4.2 Setting of the β Parameter

One of the main parameters for modeling the phase as described in section 2.4 is the factor β which determines to what extent the estimated phase is reliable. This effect has been shown in Fig. 4.7 in terms of changes in IF. Different instances of IF have been illustrated in plots (c,d,e) for different choices of the predictability factor β . A low β indicates a small SNR ratio and therefore the predicted phase is more reliable than the measurement. In the extreme case of a zero factor as in plot (c), the IF will be a straight line. In the case where we rely more on measurements than predicted data, the estimated IF will contain more changes in IF, including those regarding the low envelope (as in (e)). A proper choice can smooth the unwanted changes in IP and preserve the relevant changes.

In our study, we set the factor value and the filter parameters based on the discrimination ability between the attending and non-attending conditions using Λ . Despite the possibility of other methods for setting β such as applying an expectation maximization approach, they are not always the ideal approaches. In most applications we are not searching for a perfect model of the data as it can lead to over-fitting. Instead, we search for measures that are able to differentiate between different classes of cognitive activities.

Results corresponding to seminal dichotic tone detection as described in section 2.4.4 are shown in Fig. 3.25. A high average Λ in the attending case indicates a higher effort compared to the non-attending case. This effect has been shown in many former studies before (Corona-Strauss and Strauss, 2017; Bernarding et al., 2013b; Mortezapouraghdam et al., 2017).

4.4.3 Different Choices of β and Λ

To further explore the effect of changing β , we run a 7 minutes of the recorded ongoing EEG with different choices of β . We keep the parameters such as the filter's bandwidth ($bw = 0.4$), frequency ($f_c = 7$) and $\delta = 0.1$ fixed (they are set to values that were found to be relevant to distinguish between the attending and non-attending cases). The β parameter used for estimating IP has been linearly spaced between $1.0e^{-4}$ and 100. In Fig. 4.7, we plot the corresponding extracted IF for different β values in increasing order where the spikes are related to variations in the IP. The small values of β rely heavily on the model's prediction whereas the higher β values correspond to a stronger contribution of data measurements, as well as the model's prediction for estimating IP. In the case where $\beta = 0$, the resulting IF is zero since the estimated phase will be a sine wave and its unwrapped phase is a straight line. However as β increases, there will be more jitters in the computed IF as the model relies less on its own predictions for estimating the phase information.

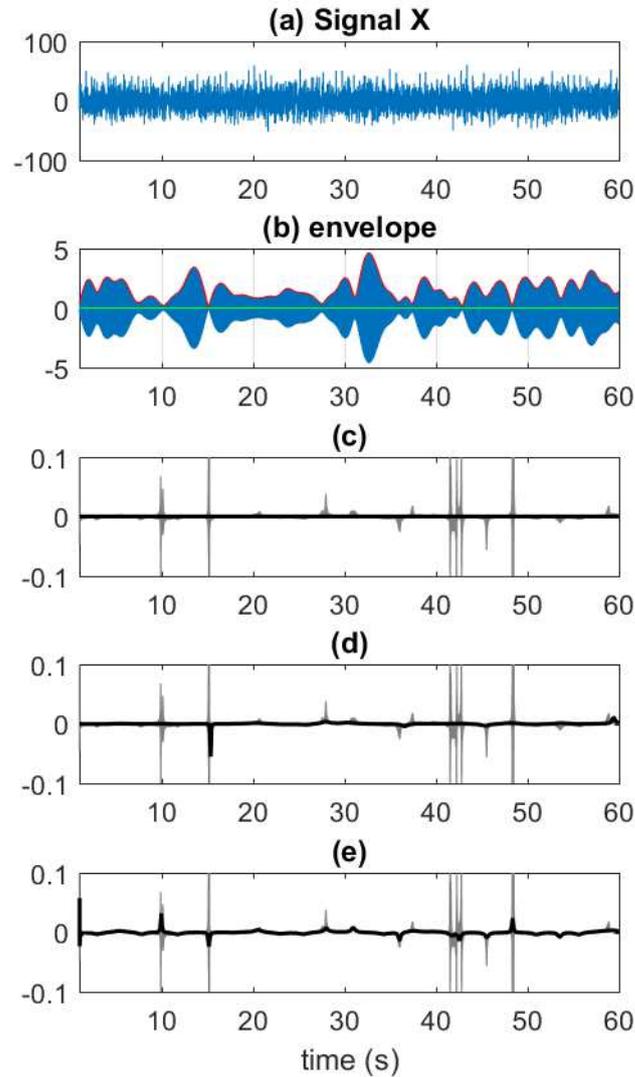


Figure 4.7: Different choices of β . (a) A 60-second segment of EEG measurements. (b) The narrow band-passed signal along with its envelope. (c) The extreme case of zero factor yields no changes in the IF. The gray lines correspond to IF without any KS preprocessing. (d) IF for $\beta = 0.005$. (e) IF resulting from a higher β factor ≈ 100 .

We compute Λ for different choices of β as shown in Fig. 4.8(b). However, in this case, we use a simpler version of the Λ measure, where we don't divide the resultant mean length by the length of the signal T . This is mainly for numerical reasons and consistency.

Different choices of β can lead to different Λ values due to different variations and localizations of the phase samples. However, there is no linear relationship between β and Λ . One of the reasons for a low Λ when β gets higher, can be explained by the addition of different noise factors that can lead to random resets in different parts of the signal. As phase jitters become more spread out, detecting directed concentration using Λ vanishes. However, a low Λ for small choices of β is more certain as predictions are heavily based on the model.

To show a broader view, we compute Λ for different choices of β parameters in Fig. 4.9 (using the

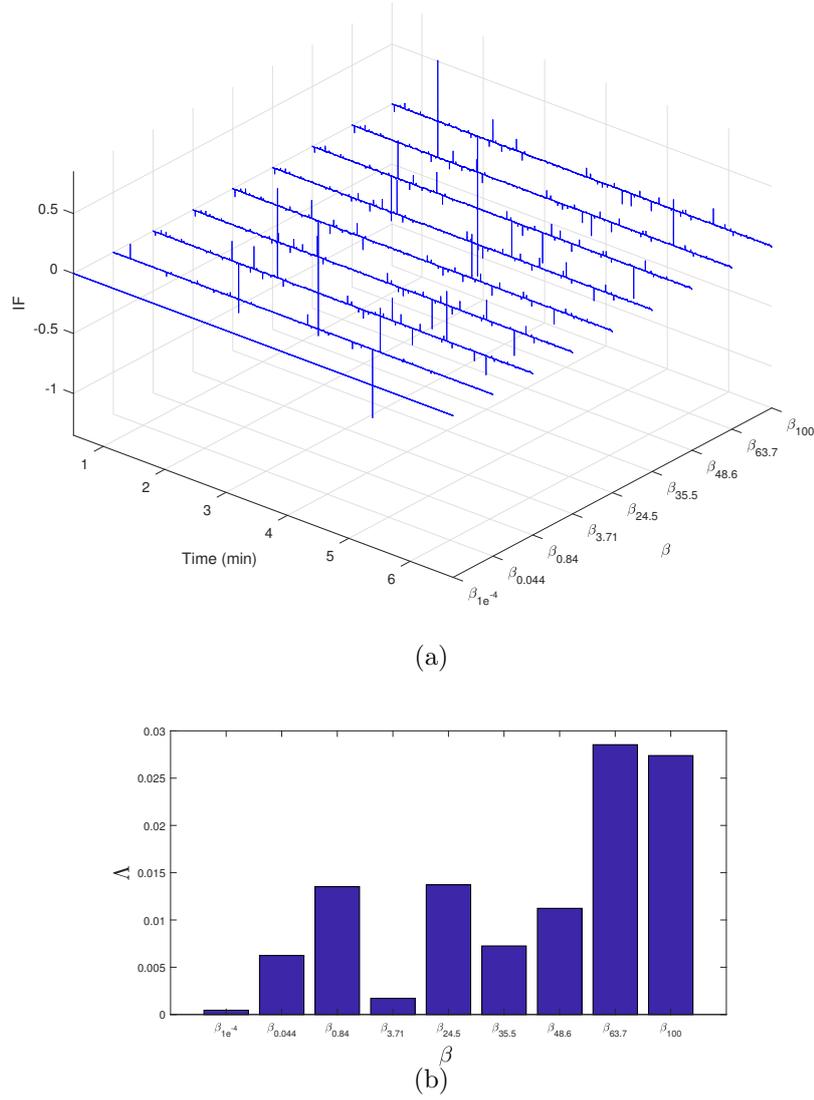


Figure 4.8: (a) Different estimations of IF for β values. (b) The corresponding Λ for different β values.

same set of parameters as before). We show the variations in attended and un-attended cases for two subjects. As β increases, the measure Λ converges to the Λ value corresponding to data with no smoothing effect, as the influence of the model in estimating the phase decreases. Therefore, after a certain threshold, it is expected that Λ converges and shows no changes to further variations in the β parameter. As stated previously, in the beginning the model is more susceptible to changes, as we rely more on the model estimates. Based on where the phase resets are located in the signal, the Λ measure can vary for different β values. We expect that other changes in the Λ measure are due to numerical inaccuracies.

The results of the current study are consistent with the conclusions presented in Corona-Strauss and Strauss (2017). However, one of the main focuses in this study, after removing spurious phase variations using KS is to measure the effect of selective attention using all *temporal resolution of the*

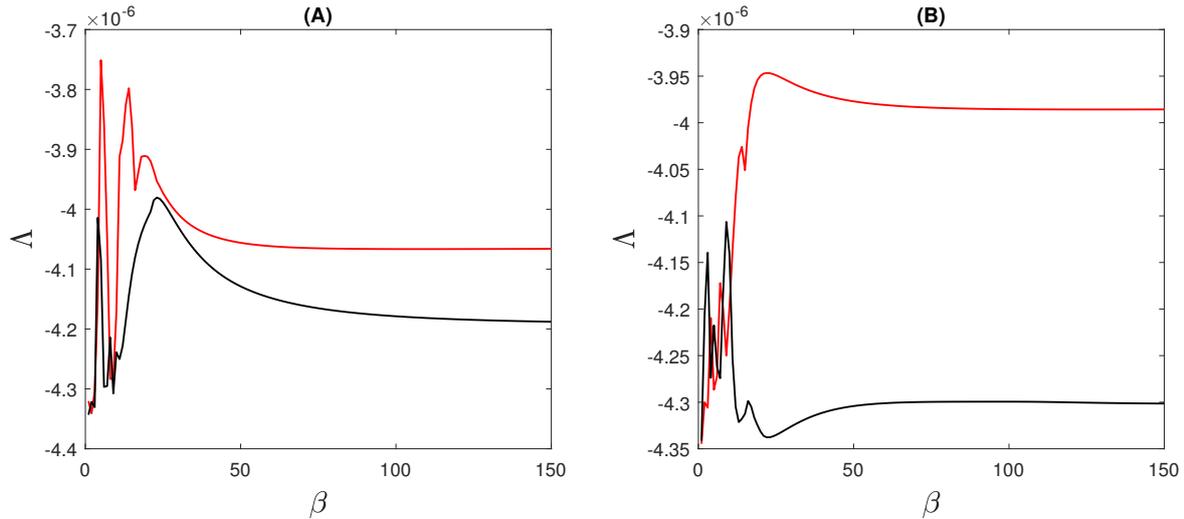


Figure 4.9: An example of changes to Λ for different choices of β factor in applying the KS model. The red and black lines correspond to the attending and non-attending conditions respectively.

ongoing ALR activity. The study in Corona-Strauss and Strauss (2017) uses only the first 60 seconds of the experiment for a meaningful contrast between the attending and non-attending conditions. Although using the complete ongoing EEG, the statistics between attending and non-attending are not significant, the average of Λ over all subjects between the two different groups are shown to differ.

To compare the results with Corona-Strauss and Strauss (2017), we also computed Λ for the first 60 seconds using the proposed method. A one-way ANOVA test between attending (*attL_{ipsi}*) and non-attending (ignoring) conditions (*IgnL_{ipsi}*) is applied to distinguish the effect of selective attention in the two processes. We used the computed Λ values of all subjects for the ignoring and attending conditions. The test statistics results yield a p-value of $p = 0.0465, F(4.57, 1)$. The average Λ results for different lengths of the first 60 and 30 seconds have been also included in Fig. 3.25.

4.4.4 Limitations

One of the major achievements in this study is distinguishing between the two processes of attending and non-attending using neural correlates of selective attention with ongoing EEG activities. We mapped the results of the study of segmented ERs with the study of ongoing EEG signals. In the averaging procedure of ERPs (segmentation of ALR activity), the noisy trials can be removed to improve the SNR and study the coherence between different trials. However, when using the complete ongoing EEG, removing noise is more difficult as none of the latter techniques are easily applicable. We therefore require methods that enable us to de-noise the ongoing signal while preserving the continuity of the signal, as segmentations can introduce artificial phase jumps.

In this study, we removed only one of the abrupt jumps in IP that is not dependent on physiological effects. This was achieved by analyzing the signal in a narrow frequency range, in order to obtain

meaningful phase oscillations. However, the overlap possibility of abrupt variations in IP due to a low envelope and physiologically induced changes has to be carefully investigated.

In future studies, the measures/techniques that are a better indicator of selective attention in terms of the level of phase clustering should be investigated. The current common measures for measuring phase clustering are biased and depend on the number of samples. Examples include the inter-trial-coherence (ITC) measure that suffers from sample size bias as well as number of trials in the experiment. The measure's output is between zero (random phase alignment) and one (perfect phase alignment). A modified version of the ITC, namely the Rayleigh Z-measure has been shown to slightly enhance the results as the sample size increases in artificially generate data ((Chou and Hsu, 2018)). The level of synchronization¹ is given by $PS = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \cos(\theta_i - \theta_j)$. One of the disadvantages of the measure Λ used in this study, is that as the number of samples increases, Λ becomes very small such that for a long temporal recording, the significance between two different processes is no longer reflected in the measure. In addition, it can suffer from numerical inaccuracies as the length of the samples increases. In the future, we will try additional approaches to measure either, continually or discretely, how the clustering of phase samples around the circle varies.

Optimization of Model Parameters Using Expectation Maximization The other main improvement is to find a better strategy to optimize the current parameters in our model. We require function to optimize the parameters such that the selected parameters are a good fit for the data and enable us to discriminate between the two processes. We consider optimizing the model parameters σ and α using *expectation maximization* (EM), where the log of the data likelihood under α and σ parameters is maximized. To avoid blind over-fitting of data with no regard to the classification aspect of the problem, we additionally introduce a cost function that measures the ratio of the variance between the two groups divided by the variations within groups. We believe that by finding a measure that is more numerically consistent, we are able to find a better set of parameters for a reliable study of the ongoing EEG activities.

4.5 The Mappings Between Laminae Activities and AERs

The average MI between the AER activities with different cortical layers as in Fig. 3.29, is a high MI across different frequencies for the mid allocated electrodes. Although, we cannot say with certainty at what layer the electrodes are allocated, a consistent high MI across the middle ranged electrodes that is related approximately to the granular layer is observed. This pattern is also observable for layer I. The redundancy of information across the granular layer and layer I can be explained due to cortical laminar processing. The main flow of information initiated from the TN is to the granular and layer III of the cortex. The thalamic input information targeted as these layers enter the lower

¹The phase angles can be expressed in terms of unit vectors on a complex unit-length circle (i.e., $z_i = e^{i\theta_i} = \cos \theta_i + i \sin \theta_i$). The measure computes the level of phase alignment in terms of average of distances between any pair of samples. The term $\frac{2}{n(n-1)}$ is the number of sample pairs that are evaluated.

supragranular layers which in turn project back to the infragranular layers. Thereby the results across these layers appear to be redundant as a result of the information flow of laminar cortical processing.

The individual result as illustrated in Fig. 3.27 suggest that there is a higher amount of information sharing between the mid-range electrodes, as it is the main receiver of thalamic input. We regard the lack of information in the deeper layers to be related to distortion of data and noisy measurements, specifically the electrode located at the TN. As it showed no significant amount of information with the other layers, it was excluded from the analysis. To more accurately observe the interaction between the cortical laminae with AER, we require more information on the specific allocation of the electrodes. In the current study, there are eight contacts in the laminar cortex, which can lead to the positioning of more than one electrode in a single layer. Therefore, the redundant information and the mixture of the recordings make it difficult to draw any concrete conclusions regarding the measured activities.

In the future studies, our goal is to employ measures that would be a better indicator of the information transfer within different layers and to the corresponding AER electrode. MI is a symmetric measure that is not able to determine the direction of interactions between cortical contacts and their relationship with the AER. In the current study, we investigated the MI between different layers with different delays d . However, no significant results explaining a meaningful or interpretable interaction was found. Measures such as transfer entropy and granger causality (GC) may help in obtaining additional information. In one of the studies in progress, we are using GC to model the activities between different electrodes before and after onset of stimulation. Using the fitted auto-regressive model characteristics such as the model order and delay, over different sequences of pre/post stimulation patterns, we are looking for information regarding the changes between different cortical layers. In addition the quality of the data measurements as well as the exact positioning of the electrodes in the laminae must be improved.

Chapter 5

Conclusions

5.1 Conclusion

We presented new approaches for measuring and tracking the changes in the dynamics of attentional binding and long-term habituation process. The EEG recordings are analyzed using different objective measures for decoding the underlying neural signature of attention over different trials. We particularly use the instantaneous phase information of the N1-P2 component of the the ERPs as it has been shown to be a reliable indicator of the endogenous effect of attentional binding. The phase information is highly clustered around a unit circle when there is a high level of attentional binding to the stimulus. On the contrary, as the attentional binding reduces, the phase information is more uniformly spread.

In the first set of studies, we model the distribution of IP information using von-Mises probability density functions. We apply this method on two distinct auditory stimulations (soft vs. aversive) to determine their likelihood to habituate. The 50dB SPL is a soft stimulus and has a comfortable loudness level, whereas 100dB SPL is considered aversive (uncomfortably loud) by the participants. We analyze the N1-P2 complex in the corresponding measured ERs by fitting a von-Mises model. The ERs over the trials are segmented with overlap and a von-Mises pdf is fitted for each segment. The estimated concentration parameters are then used to decode the level of attentional binding over time. Using a proper overlap window, we can clearly distinguish between the changes in a habituation and non-habituation process. In the case of soft stimulation, the concentration parameter at the beginning of the segments is high and gradually decreases. The same pattern is observed in other measures such as the changes in the entropy values. However, the changes in the concentration parameter of an aversive stimulus remain high with less variations throughout the experiment. The

rate of changes in the estimated concentration parameter between the soft and aversive stimulus has been shown to significantly differ ($p < 0.05$). The results are cross-validated with a neurofunctional model that simulates ERPs and the results are consistent.

The proposed method, however, relies heavily on a proper window size and the number of overlapping samples to obtain the meaningful variations in the phase information. In addition, we are required to fit a model to many data segments. We therefore improved the temporal resolution of detecting changes in concentration by applying a Bayesian change point model. The model is used for detecting the time instances at which a significant change in the underlying parameters of the generative model occurs. Run-length, a feature indicating the presence of change points is used for discriminating between the habituation and non-habituation process. As no change point is expected in case of non-habituation, the average run-length is higher than the habituation process as no resetting in the run-length parameter would be detected. In the case of a soft stimulus, there are more resets in the run-length parameter. The computed average run-length over all subjects as well as individual subjects is higher for an aversive stimulus in comparison to a soft stimulus.

Despite detecting the change points with high temporal resolution, the abrupt change point detection algorithm however has drawbacks in terms of tracking the variations in long-term habituation. Because the long-term habituation is defined as a gradual process, an abrupt change point model may not reflect the gradual changes clearly in the signal. In addition, a more accurate estimate of the timing of habituation may be difficult to obtain. We therefore applied a Bayesian change point model that is able to capture the gradual changes in the concentration parameter. The method is able to track the gradual changes in the IP information at different sound stimulation levels. However, no unique timing was obtained to indicate the occurrence of habituation for different stimulation levels as the variability between subjects is high. We computed the average of the expected value of the concentration parameter over different sample points. The results indicate as the sound intensity level increases, the average of the concentration parameter over all samples is more likely to increase as well.

We extended our studies on the neural modulations of the selective attention by employing a longer time-interval stimulation. For this aim, we used the measured neural activity of a dichotic listening experiment. No segmentation is applied on the recorded EEG activities. We investigated the effect of spurious phase variation in ongoing signal by applying a variant of a Kalman smoother model to remove the unwanted phase jitters. The proposed approach is first tested on synthetic generated data with spurious phase jumps due to a low envelope. In addition, random phase shifts were assigned in the signal as the main change points. The proposed model is able to significantly reduce the number of false positives in detecting the change points. After validating the model with synthetic data, we applied the model on measured ongoing EEG signals for removing spurious phase variations and objectively computed the level of selective attention for the two conditions of attending and ignoring a stimulus. The average objective measure for the case of attending is higher than the ignoring

condition, reflecting a higher attentional binding to the stimulus than ignoring case. In the future studies, we aim to improve significantly the optimization of model parameters.

Despite the success in developing and applying different methods in decoding the underlying neural signature of selective attention and long-term habituation, we can investigate the effects deeper in terms of the projection of AERs to the laminar phase dynamics in-vitro. For this aim, we assess the mutual information between the AERs with the individual cortical layers at different frequencies in rats. It is shown that the highest information sharing is between the AER and the granular layers as they are considered to be the main thalamo-recipients. Redundant information sharing occurs between layer I and AERs and is explained by cortical laminar processing. No further results regarding the delay process between different cortical layers and its relation to AER could be obtained at this stage. In future studies, we will investigate asymmetric measures such as transfer entropy and granger causality to obtain a more detailed information on the direction of information flow between different layers and the projection of cortical neural activities onto the AERs. We aim to use the pre-stimulus data to understand the possible delay, model order using the granger causality, and compare the results with the post-stimulus estimated parameters. We believe that over a long period of stimulation, we are able to have a better understanding of the changes and possible delays between recorded AERs and cortical layers.

Bibliography

Ryan Prescott Adams and David JC MacKay. Bayesian online changepoint detection. *arXiv preprint arXiv:0710.3742*, 2007.

Peter W Alberti. The anatomy and physiology of the ear and hearing. *Occupational exposure to noise: Evaluation, prevention, and control*, pages 53–62, 2001.

Galka Andreas. *Topics in nonlinear time series analysis, with implications for EEG analysis*, volume 14. World Scientific, 2000.

Claudio Babiloni, Fabio Babiloni, Filippo Carducci, Febo Cincotti, Fabrizio Rosciarelli, Lars Arendt-Nielsen, Andrew CN Chen, and Paolo Maria Rossini. Human brain oscillatory activity phase-locked to painful electrical stimulations: A multi-channel EEG study. *Human brain mapping*, 15(2):112–123, 2002.

Arindam Banerjee, Inderjit S Dhillon, Joydeep Ghosh, and Suvrit Sra. Clustering on the unit hypersphere using von mises-fisher distributions. *Journal of Machine Learning Research*, 6(Sep):1345–1382, 2005.

Daniel Barry and John A Hartigan. Product partition models for change point problems. *The Annals of Statistics*, pages 260–279, 1992.

Dwight W Batteau. The role of the pinna in human localization. *Proceedings of the Royal Society of London B: Biological Sciences*, 168(1011):158–180, 1967.

Georgvon Békésy. Concerning the pleasures of observing, and the mechanics of the inner ear. *Nobel Lecture, December*, 11, 1961.

Daniel Bendor and Xiaoqin Wang. The neuronal representation of pitch in primate auditory cortex. *Nature*, 436(7054):1161–1165, 2005.

- Gadi Benshalom and Edward L White. Quantification of thalamocortical synapses with spiny stellate neurons in layer iv of mouse somatosensory cortex. *J Comp Neurol*, 253:303–314, 1986.
- Hans Berger. Über das elektrenkephalogramm des menschen. *Archiv für psychiatrie und nervenkrankheiten*, 87(1):527–570, 1929.
- Ronny Bergmann, Friederike Laus, Gabriele Steidl, and Andreas Weinmann. Second order differences of cyclic data and applications in variational denoising. *SIAM Journal on Imaging Sciences*, 7(4): 2916–2953, 2014.
- Corinna Bernarding, Ronny Hannemann, David Herrmann, Daniel J Strauss, and Farah I Corona-Strauss. Extraction of listening effort correlates in the oscillatory EEG activity: Investigation of different hearing aid configurations. In *Neural Engineering (NER), 2013 6th International IEEE/EMBS Conference on*, pages 1258–1261, 2013a.
- Corinna Bernarding, Daniel J Strauss, Ronny Hannemann, Harald Seidler, and Farah I Corona-Strauss. Neural correlates of listening effort related factors: influence of age and hearing impairment. *Brain research bulletin*, 91:21–30, 2013b.
- Corinna Bernarding, Farah I Corona-Strauss, Ronny Hannemann, and Daniel J Strauss. Objective assessment of listening effort: Effects of an increased task demand. In *Conf Proc IEEE Eng Med Biol Soc*, pages 3684–3687, 2016.
- Christoph Börgers and Nancy J Kopell. Gamma oscillations and stimulus selection. *Neural computation*, 20(2):383–414, 2008.
- ME Brandt. Visual and auditory evoked phase resetting of the alpha EEG. *International journal of psychophysiology*, 26(1):285–298, 1997.
- Albert S Bregman. *Auditory scene analysis: The perceptual organization of sound*. MIT press, 1994.
- Mark Briers, Arnaud Doucet, and Simon Maskell. Smoothing algorithms for state–space models. *Annals of the Institute of Statistical Mathematics*, 62(1):61–89, 2010.
- Donald E Broadbent. *Perception and Communication*. Pergamon, London, UK, 1958.
- Randy M Bruno and Bert Sakmann. Cortex is driven by weak but synchronously active thalamocortical synapses. *Science*, 312(5780):1622–1627, 2006.
- Andreas Bruns. Fourier-, Hilbert-, and wavelet based signal analysis: are they really different approaches? *J. of Neuroscience Methods*, 137:321–332, 2004.
- Niko A Busch, Julien Dubois, and Rufin VanRullen. The phase of ongoing EEG oscillations predicts visual perception. *The Journal of neuroscience*, 29(24):7869–7876, 2009.

- Michael Busse, Lars Haab, Mai Mariam, Christoph Krick, Tina Weis, Wolfgang Reith, and Daniel J Strauss. Assessment of aversive stimuli dependent attentional binding by the n170 vep component. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pages 3975–3978. IEEE, 2009.
- Blake E Butler and Stephen G Lomber. Functional and structural changes throughout the auditory system following congenital and early-onset deafness: implications for hearing restoration. *Frontiers in systems neuroscience*, 7:92, 2013.
- Robert A Butler, Wolf-Dieter Keidel, and M Spreng. An investigation of the human cortical evoked potential under conditions of monaural and binaural stimulation. *Acta Otolaryngologica*, 68(1-6): 317–326, 1969.
- Thomas J Carew and Eric R Kandel. Acquisition and retention of long-term habituation in aplysia: correlation of behavioral and cellular processes. *Science*, 182(4117):1158–1160, 1973.
- Sven G Carlsson and Soly I Erlandsson. Habituation and tinnitus: an experimental study. *Journal of psychosomatic research*, 35(4):509–514, 1991.
- Gastone G Celesia. Organization of auditory cortical areas in man. *Brain*, 99(3):403–414, 1976.
- Chandramouli Chandrasekaran, Hjalmar K Turesson, Charles H Brown, and Asif A Ghazanfar. The influence of natural scene dynamics on auditory cortical activity. *The Journal of Neuroscience*, 30(42):13919–13931, 2010.
- M Chavez, Michel Besserve, C Adam, and Jacques Martinerie. Towards a proper estimation of phase synchronization from time series. *Journal of neuroscience methods*, 154(1-2):149–160, 2006.
- Edward Colin Cherry. Some experiments on the recognition of speech, with one and with two ears. *The Journal of the acoustical society of America*, 25(5):975–979, 1953.
- Edward Colin Cherry. Some further experiments upon the recognition of speech with one two ears. *The Journal of the acoustical society of America*, 26(5):554–559, 1954.
- Elizabeth P Chou and Shen-Mou Hsu. Cosine similarity as a sample size-free measure to quantify phase clustering within a single neurophysiological signal. *Journal of neuroscience methods*, 295: 111–120, 2018.
- Donna Coch, Lisa D Sanders, and Helen J Neville. An event-related potential study of selective auditory attention in children and adults. *Journal of cognitive neuroscience*, 17(4):605–622, 2005.
- Andrew RA Conway, Nelson Cowan, and Michael F Bunting. The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic bulletin & review*, 8(2): 331–335, 2001.

- Farah I Corona-Strauss and Daniel J Strauss. Robust extraction of the N1-effect in dichotic listening using hardy space mappings of auditory late single trials. In *Conf Proc IEEE Eng Med Biol Soc*, volume 2016:1, pages 3418–3421, 2016.
- Farah I Corona-Strauss and Daniel J Strauss. Circular organization of the instantaneous phase in ERPs and the ongoing EEG due to selective attention. In *Neural Engineering (NER), 2017 8th International IEEE/EMBS Conference on*, pages 625–628. IEEE, 2017.
- Farah I Corona-Strauss, Corinna Bernarding, Matthias Latzel, and Daniel J Strauss. Syllable evoked auditory late responses: Effects of noise onsets and noise types. In *Neural Engineering (NER), 2011 5th International IEEE/EMBS Conference on*, pages 140–143. IEEE, 2011.
- George H Crampton and Wallace J Schwam. Effects of arousal reaction on nystagmus habituation in the cat. *American Journal of Physiology-Legacy Content*, 200(1):29–33, 1961.
- F Crick. Function of the thalamic reticular nucleus: The searchlight hypothesis. *Proc. Natl. Acad. Sc. USA*, 81:4586–4590, 1984.
- Wilbur B Davenport, William L Root, et al. *An introduction to the theory of random signals and noise*, volume 159. McGraw-Hill New York, 1958.
- Michael Domjan. *The principles of learning and behavior*. Cengage Learning, 2014.
- European Committee for Standardization. Electroacoustics- audiometric equipment. part 3: Test signals of short duration. Technical Report, The European Standard. EN 60645-3:2007., 2007.
- Michael W Eysenck and Mark T Keane. *Cognitive psychology: A student's handbook*. Taylor & Francis, 2000.
- Nicholas I Fisher. *Statistical Analysis of Circular Data*. Cambridge University Press, 1995.
- Antonio Elia Forte, Octave Etard, and Tobias Reichenbach. The human auditory brainstem response to running speech reveals a subcortical mechanism for selective attention. *eLife*, 6, 2017.
- Walter J Freeman, Brian C Burke, and Mark D Holmes. Aperiodic phase re-setting in scalp EEG of beta–gamma oscillations by state transitions at alpha–theta rates. *Human brain mapping*, 19(4):248–272, 2003.
- Pascal Fries. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in cognitive sciences*, 9(10):474–480, 2005.
- Lluís Fuentemilla, Josep Marco-Pallarés, and Carles Grau. Modulation of spectral power and of phase resetting of EEG contributes differentially to the generation of auditory event-related potentials. *Neuroimage*, 30(3):909–916, 2006.
- Stanley A Gelfand. *Essentials of audiology*, 2001.

- Paxinos George and Watson Charles. *The rat brain in stereotaxic coordinates*. 2007.
- Zoubin Ghahramani. An introduction to hidden markov models and bayesian networks. *International journal of pattern recognition and artificial intelligence*, 15(01):9–42, 2001.
- KK Glendenning and RB Masterton. Comparative morphometry of mammalian central auditory systems: variation in nuclei and form of the ascending system. *Brain, behavior and evolution*, 51(2):59–89, 1998.
- E Bruce Goldstein and James Brockmole. *Sensation and perception*. Cengage Learning, 2016.
- Timothy D Griffiths and Jason D Warren. The planum temporale as a computational hub. *Trends in neurosciences*, 25(7):348–353, 2002.
- Lars Haab, Elisabeth Wallhauser-Franke, Carlos Trenado, and Daniel J Strauss. Modeling limbic influences on habituation deficits in chronic tinnitus aurium. In *Conf Proc IEEE Engineering in Medicine and Biology Society*, pages 4234–4237, Minneapolis, MN, USA, 2009.
- Lars Haab, Carlos Trenado, Mai Mariam, and Daniel J Strauss. Neurofunctional model of large-scale correlates of selective attention governed by stimulus-novelty. *Cognitive Neurodynamics*, 5:103–111, 2011.
- James Wilbur Hall. *Handbook of auditory evoked responses*. Allyn & Bacon, 1992.
- Todd C Handy. *Event-related potentials: A methods handbook*. MIT press, 2005.
- Jonathan C Hansen and Steven A Hillyard. Endogeneous brain potentials associated with selective auditory attention. *Electroencephalography and clinical neurophysiology*, 49(3):277–290, 1980.
- J.D. Harris. Habituation response decrement in the intact organism. *Psychological Bulletin*, 40(6):385, 1943.
- Robert D Hawkins, Eric R Kandel, and Steven A Siegelbaum. Learning to modulate transmitter release: themes and variations in synaptic plasticity. *Annual review of neuroscience*, 16(1):625–665, 1993.
- Steven A Hillyard, Robert F Hink, Vincent L Schwent, and Terence W Picton. Electrical signs of selective attention in the human brain. *Science*, 182(108):177–180, 1973.
- Steven A Hillyard, Terence W Picton, D Regan, et al. Sensation, perception and attention: Analysis using ERPs. In *Event-related brain potentials in man*, pages 223–321. Academic Press New York, 1978.
- Steven A Hillyard, Edward K Vogel, and Steven J Luck. Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 353(1373):1257–1270, 1998.

- JD Hood and JP Poole. Tolerable limit of loudness: its clinical and physiological significance. *The Journal of the Acoustical Society of America*, 40(1):47–53, 1966.
- Akira Iwanami, Toshikazu Shinba, Michihisa Sumi, Nobuyuki Ozawa, and Ken-ichi Yamamoto. Event-related potentials during an auditory discrimination task in rats. *Neuroscience research*, 21(1):103–106, 1994.
- William James. *The principles of psychology*, 1890.
- Ben H Jansen, Gopal Agarwal, Anant Hegde, and Nashaat N Boutros. Phase synchronization of the ongoing EEG and auditory EP generation. *Clinical Neurophysiology*, 114(1):79–85, 2003.
- Daniel Kahneman. *Attention and effort*, volume 1063. Prentice-Hall Englewood Cliffs, NJ, 1973.
- Christoph Kayser, Marcelo A Montemurro, Nikos K Logothetis, and Stefano Panzeri. Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. *Neuron*, 61(4):597–608, 2009.
- John Kent. Some probabilistic properties of Bessel functions. *The Annals of Probability*, pages 760–770, 1978.
- Shyam M Khanna and Debra GB Leonard. Basilar membrane tuning in the cat cochlea. *Science*, 215:305–306, 1982.
- Tansi Jamshed Khodai. *Functional laminar architecture of rat primary auditory cortex following acoustic trauma*. PhD thesis, University of Strathclyde, 2014.
- Nelson Yuan-Sheng Kiang. Discharge patterns of single fibers in the cat’s auditory nerve. Technical report, DTIC Document, 1965.
- Akihisa Kimura, Isao Yokoi, Hiroki Imbe, Tomohiro Donishi, and Yoshiaki Kaneoke. Auditory thalamic reticular nucleus of the rat: anatomical nodes for modulation of auditory and cross-modal sensory processing in the loop connectivity between the cortex and thalamus. *J Comp Neurol.*, 520:1457–1480, 2012.
- Wolfgang Klimesch, Manuel Schabus, Michael Doppelmayr, Walter Gruber, and Paul Sauseng. Evoked oscillations and early components of event-related potentials: an analysis. *International Journal of Bifurcation and Chaos*, 14(02):705–718, 2004a.
- Wolfgang Klimesch, Bärbel Schack, Manuel Schabus, Michael Doppelmayr, Walter Gruber, and Paul Sauseng. Phase-locked alpha and theta oscillations generate the p1–n1 complex and are related to memory performance. *Cognitive Brain Research*, 19(3):302–316, 2004b.
- Wolfgang Klimesch, Paul Sauseng, Simon Hanslmayr, Walter Gruber, and Roman Freunberger. Event-related phase reorganization may explain evoked neural dynamics. *Neuroscience & Biobehavioral Reviews*, 31(7):1003–1016, 2007.

- Robert T Knight, Simon Brailowsky, Donatella Scabini, and Gregory V Simpson. Surface auditory evoked potentials in the unrestrained rat: component definition. *Electroencephalography and clinical neurophysiology*, 61(5):430–439, 1985.
- David C Knill and Alexandre Pouget. The Bayesian brain: the role of uncertainty in neural coding and computation. *TRENDS in Neurosciences*, 27(12):712–719, 2004.
- David JM Kraemer, C Neil Macrae, Adam E Green, and William M Kelley. Musical imagery: sound of silence activates auditory cortex. *Nature*, 434(7030):158–158, 2005.
- Jean-Philippe Lachaux, Eugenio Rodriguez, Jacques Martinerie, Francisco J Varela, et al. Measuring phase synchrony in brain signals. *Human brain mapping*, 8(4):194–208, 1999.
- Joel Lachter, Kenneth I Forster, and Eric Ruthruff. Forty-five years after broadbent (1958): still no identification without attention. *Psychological review*, 111(4):880, 2004.
- Paul. Langevin. Magnetisme et theorie des electrons. *Ann. Chim. Phys.*, 5:71–127, 1994.
- Nilli Lavie. Perceptual load as a necessary condition for selective attention. *Journal Of Experimental Psychology Human Perception And Performance*, 21:451–468, 1995.
- Nilli Lavie. *Selective attention and cognitive control: Dissociating attentional functions through different types of load*. MIT–Press, Cambridge, MA, USA, 2000.
- Nilli Lavie and Yehoshua Tsal. Perceptual load as a major determinant of the locus of selection in visual attention. *Attention, Perception, & Psychophysics*, 56(2):183–197, 1994.
- Charles C Lee and Kazuo Imaizumi. Functional convergence of thalamic and intrinsic projections to cortical layers 4 and 6. *Neurophysiology*, 45(5-6):396–406, 2013.
- Jennifer F Linden and Christoph E Schreiner. Columnar transformations in auditory cortex? a comparison to visual and somatosensory cortices. *Cerebral Cortex*, 13(1):83–89, 2003.
- Yin F Low and Daniel J. Strauss. A performance study of the wavelet–phase stability (wps) in auditory selective attention. *Brain Res Bull*, 86:110–117, 2011.
- Joachim Lübke, Veronica Egger, Bert Sakmann, and Dirk Feldmeyer. Columnar organization of dendrites and axons of single and synaptically coupled excitatory spiny neurons in layer 4 of the rat barrel cortex. *Journal of Neuroscience*, 20(14):5300–5311, 2000.
- Steven J Luck. *An introduction to the event-related potential technique*. MIT press, 2014.
- Huan Luo and David Poeppel. Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, 54(6):1001–1010, 2007.

- Scott Makeig, Marissa Westerfield, T-P Jung, S Enghoff, Jeanne Townsend, Eric Courchesne, and Terrence J Sejnowski. Dynamic brain sources of visual evoked responses. *Science*, 295(5555): 690–694, 2002.
- Kanti V Mardia. Statistics of directional data. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 349–393, 1975.
- Kanti V Mardia and Peter E Jupp. *Directional statistics*, volume 494. John Wiley & Sons, 2009.
- Mai Mariam, Wolfgang Delb, Farah I Corona-Strauss, Marc Bloching, and Daniel J Strauss. Comparing the habituation of the late auditory potentials to loud and soft sounds. *J Physiol Measurement*, 30:141–153, 2009.
- Mai Mariam, Wolfgang Delb, Bernhard Schick, and Daniel J Strauss. Objective loudness scaling by late auditory evoked potentials : Kernel based novelty detection approach. *Art Intell Medicine*, 55:185–195, 2012.
- Shigeaki Matsuoka. Theta rhythms: state of consciousness. *Brain Topography*, 3(1):203–208, 1990.
- Andrew R Mayer, Deborah Harrington, John C Adair, and Roland Lee. The neural networks underlying endogenous auditory covert orienting and reorienting. *Neuroimage*, 30(3):938–949, 2006.
- Kerry McAlonan, Verity J Brown, and Eric M Bowman. Thalamic reticular nucleus activation reflects attentional gating during classical conditioning. *J Neurosci.*, 20:8897–8901, 2000.
- Neville Moray. Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly journal of experimental psychology*, 11(1):56–60, 1959.
- Zeinab Morteza pouraghdam and Daniel J Strauss. Removal of spurious phase variations in oscillatory signals. pages 2209–2212, 2017.
- Zeinab Morteza pouraghdam, Lars Haab, Gabriele Steidl, and Daniel J Strauss. Detection of change points in phase data: A bayesian analysis of habituation processes. In *Conf Proc IEEE Eng Med Biol Soc*, volume 2014:1, pages 1014–1017, 2014.
- Zeinab Morteza pouraghdam, Lars Haab, Farah I Corona-Strauss, Gabriele Steidl, and Daniel J Strauss. Assessment of Long-Term Habituation Correlates in Event-Related Potentials Using a von Mises Model. *IEEE Trans Neural Syst Rehabil Eng*, 23:363–373, 2015a.
- Zeinab Morteza pouraghdam, Lars Haab, Karsten Schwerdtfeger, and Daniel J Strauss. Relating auditory evoked responses to the laminar phase dynamics in rats using mutual information. In *Neural Engineering (NER), 2015 7th International IEEE/EMBS Conference on*, pages 952–955. IEEE, 2015b.

- Zeinab Mortezapouraghdam, Robert C Wilson, Lars Schwabe, and Daniel J Strauss. Bayesian Modeling of the Dynamics of Phase Modulations and their Application to Auditory Event Related Potentials at Different Loudness Scales. *Frontiers in computational neuroscience*, 10, 2016.
- Zeinab Mortezapouraghdam, Corinna Bernarding, and Daniel J Strauss. Objective assessment of perceived effort in listening by employing EEG feature. In *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE*, pages 2908–2911. IEEE, 2017.
- Zeinab Mortezapouraghdam, Farah I Corona-Strauss, Kazutaka Takahashi, and Daniel J Strauss. Reducing the effect of spurious phase variations in neural oscillatory signals. *Frontiers in Computational Neuroscience*, 12:1–16, 2018. ISSN 1662-5188. doi: 10.3389/fncom.2018.00082. URL <https://www.frontiersin.org/article/10.3389/fncom.2018.00082>.
- Frank E Musiek, Jane A Baran, Jennifer B Shinn, Linda Guenette, Elena Zaidan, and Jeffrey Weihing. Central deafness: An audiological case study. *International journal of audiology*, 46(8): 433–441, 2007.
- Isabella Mutschler, Birgit Wieckhorst, Oliver Speck, A. Schulze-Bonhage, Jú Hennig, Erich Seifritz, and Tonio Ball. Time scales of auditory habituation in the amygdala and cerebral cortex. *Cerebral Cortex*, 20:2531–2539, 2010.
- S Shyamla Narayan, Andrei N Temchin, Alberto Recio, and Mario A Ruggero. Frequency tuning of basilar membrane and auditory nerve fibers in the same cochleae. *Science*, 282(5395):1882–1884, 1998.
- Arne Öhman and Malcolm Lader. Selective attention and “habituation” of the auditory averaged evoked response in humans. *Physiology & behavior*, 8(1):79–85, 1972.
- Alan V Oppenheim, W Schafer Ronald, and RB John. Discrete-time signal processing. *New Jersey, Printice Hall Inc*, 1989.
- Ulrich Paquet. Empirical Bayesian change point detection. *Graphical Models*, 1995:1–20, 2007.
- Will D Penny, Stephan J Kiebel, James M Kilner, and Mick D Rugg. Event-related brain dynamics. *Trends in neurosciences*, 25(8):387–389, 2002.
- Gert Pfurtscheller and Fernando L Da Silva. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical neurophysiology*, 110(11):1842–1857, 1999.
- Dale Purves, George J Augustine, and David Fitzpatrick. *Neurosciences*. De Boeck Supérieur, 2015.
- Lawrence Rabiner and Biing-Hwang Juang. An introduction to hidden markov models. *ASSP Magazine, IEEE*, 3(1):4–16, 1986.

- Catharine H Rankin, Thomas Abrams, Robert J Barry, Seema Bhatnagar, David F Clayton, John Colombo, Gianluca Coppola, Mark A Geyer, David L Glanzman, Stephen Marsland, et al. Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. *Neurobiology of learning and memory*, 92(2):135–138, 2009.
- Aparna Rao, Yang Zhang, and Sharon Miller. Selective listening of concurrent auditory stimuli: an event-related potential study. *Hearing research*, 268(1):123–132, 2010.
- Joseph P Rauschecker. Neural encoding and retrieval of sound sequences. *Ann N Y Acad Sci*, 1060:125–135, 2005.
- Joseph P Rauschecker, Amber M Leaver, and Mark Mühlau. Tuning out the noise: Limbic–auditory interactions in tinnitus. *Neuron*, pages 819–826, 2010.
- Ronald A Rensink. Attention: Change blindness and inattention blindness. 2009.
- Utz Richter and Thomas. Fedtke. Reference zero for the calibration of audiometric equipment using clicks as test signals. 44:478–487, 2005.
- Peter A Robinson, C J Rennie, Donald L Rowe, S C O’Connor, and E Gordon. Multiscale brain modelling. *Philosophical Transactions of the Royal Society of London*, B 360:1043–1050, 2005.
- Peter A Robinson, C J Rennie, Donald L Rowe, and S C O’Connor. Estimation of multiscale neurophysiologic parameters by electroencephalographic means. *Hum Brain Mapp*, 23:53–72, 204.
- Timm Rosburg, Peter Trautner, Nashaat N Boutros, Oleg A Korzyukov, Carlo Schaller, Christian Erich Elger, and Martin Kurthen. Habituation of auditory evoked potentials in intracranial and extracranial recordings. *Psychophysiology*, 43(2):137–144, 2006.
- Stuart Rosen. Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 336(1278):367–373, 1992.
- David Rudrauf, Abdel Douiri, Christopher Kovach, Jean-Philippe Lachaux, Diego Cosmelli, Mario Chavez, Claude Adam, Bernard Renault, Jacques Martinerie, and Michel Le Van Quyen. Frequency flows and the time-frequency dynamics of multivariate phase synchronization in brain signals. *Neuroimage*, 31(1):209–227, 2006.
- Daniel B Russakoff, Carlo Tomasi, Torsten Rohlfing, and Calvin R Maurer. Image similarity using mutual information of regions. In *European Conference on Computer Vision*, pages 596–607. Springer, 2004.
- Reza Sameni and Esmaeil Seraj. A robust statistical framework for instantaneous electroencephalogram phase and frequency estimation and analysis. *Physiological measurement*, 38(12):2141, 2017.

- Naoyuki Sato and Yoko Yamaguchi. Theta synchronization networks emerge during human object-place memory encoding. *Neuroreport*, 18:419–424, 2007.
- Paul Sauseng, Wolfgang Klimesch, Walter Gruber, Simon Hanslmayr, Roman Freunberger, and Michael Doppelmayr. Are event-related potential components generated by phase resetting of brain oscillations? a critical discussion. *Neuroscience*, 146(4):1435–1444, 2007.
- B McA Sayers, HA Beagley, and WR Henshall. The mechanism of auditory evoked EEG responses. *Nature*, 1974.
- Alfons Schnitzler and Joachim Gross. Normal and pathological oscillatory communication in the brain. *Nature reviews neuroscience*, 6(4):285, 2005.
- Philippe G Schyns, Gregor Thut, and Joachim Gross. Cracking the code of oscillatory activity. *PLoS biology*, 9(5):e1001064, 2011.
- Esmail Seraj. Cerebral synchrony assessment: A general review on cerebral signals' synchronization estimation concepts and methods. *arXiv preprint arXiv:1612.04295*, 2016.
- Esmail Seraj and Reza Sameni. Robust electroencephalogram phase estimation with applications in brain-computer interface systems. *Physiological measurement*, 38(3):501, 2017.
- Ankoor S Shah, Steven L Bressler, Kevin H Knuth, Mingzhou Ding, Ashesh D Mehta, Istvan Ulbert, and Charles E Schroeder. Neural dynamics and the fundamental mechanisms of event-related brain potentials. *Cerebral cortex*, 14(5):476–483, 2004.
- Barbara G Shinn-Cunningham. Object-based auditory and visual attention. *Trends in cognitive sciences*, 12(5):182–186, 2008.
- Daniel J Simons and Christopher F Chabris. Gorillas in our midst: Sustained inattentive blindness for dynamic events. *Perception*, 28(9):1059–1074, 1999.
- Charles Spence and Valerio Santangelo. Auditory attention. *Oxford Handbook of Auditory Science: Hearing*, 3:249, 2010.
- Mallat Stéphane. *A Wavelet Tour of Signal Processing*. Academic Press, 1999.
- SD Stephens and CM Anderson. Experimental studies on the uncomfortable loudness level. *Journal of Speech, Language, and Hearing Research*, 14(2):262–270, 1971.
- Robert J Sternberg and Karin Sternberg. *Cognitive psychology*. Nelson Education, 2016.
- Stanley S Stevens. The relation of pitch to intensity. *The Journal of the Acoustical Society of America*, 6(3):150–154, 1935.
- Gilbert Strang and Truong Nguyen. *Wavelets and filter banks*. SIAM, 1996.

- Daniel J Strauss and Alexander L Francis. Toward a taxonomic model of attention in effortful listening. *Cogn, Affect Behav Neurosci*, pages 1–17, 2017. doi: 10.3758/s13415-017-0513-0.
- Daniel J Strauss, Wolfgang Delb, Roberto D’Amelio, and Peter Falkai. Neural synchronization stability in the tinnitus decompensation. In *Proceedings of the 2st Int. IEEE EMBS Conference on Neural Engineering*, pages 186–189, Arlington, VA, USA, 2005.
- Daniel J Strauss, Farah I Corona-Strauss, and Matthias Froehlich. Objective estimation of the listening effort: Towards a neuropsychological and neurophysical model. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 1777–1780. IEEE, 2008a.
- Daniel J Strauss, Wolfgang Delb, Roberto D’Amelio, Yin F Low, and Peter Falkai. Objective quantification of the tinnitus decompensation by synchronization measures of auditory evoked single sweeps. *IEEE Trans Neural Syst Rehabil Eng.*, pages 74–81, 2008b.
- Daniel J Strauss, Farah I Corona-Strauss, Carlos Trenado, Corinna Bernarding, Wolfgang Reith, Matthias Latzel, and Matthias Froehlich. Electrophysiological correlates of listening effort: neurodynamical modeling and measurement. *Cognitive neurodynamics*, 4(2):119–131, 2010.
- Daniel J Strauss, Tanja Teuber, Gabriele Steidl, and Farah I Corona-Strauss. Exploiting the self-similarity in ERP images by nonlocal means for single-trial denoising. *IEEE Trans Neural Syst Rehabil Eng.*, 21(4):576–583, 2013.
- Shravani Sur and VK Sinha. Event-related potential: An overview. *Industrial psychiatry journal*, 18(1):70, 2009.
- Larry W Swanson. Brain maps: structure of the rat brain, 1992.
- Francois D Szymanski, Neil C Rabinowitz, Cesare Magri, Stefano Panzeri, and Jan WH Schnupp. The laminar and temporal structure of stimulus information in the phase of field potentials of auditory cortex. *The Journal of Neuroscience*, 31(44):15787–15801, 2011.
- Richard F Thompson. The neurobiology of learning and memory. *Science*, 233(4767):941–947, 1986.
- Richard F Thompson. Habituation: a history. *Neurobiology of learning and memory*, 92(2):127, 2009.
- Richard F Thompson and William A Spencer. Habituation: a model phenomenon for the study of neuronal substrates of behavior. *Psychological review*, 73(1):16, 1966.
- Anne M Treisman. Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, 12(4):242–248, 1960.
- Carlos Trenado, Lars Haab, and Daniel J Strauss. Corticothalamic feedback dynamics for neural correlates of auditory selective attention. *IEEE Trans Neural Syst Rehabil Eng.*, 17:46–52, 2009.

- Geoffrey Underwood. Moray vs. the rest: The effects of extended shadowing practice. *The Quarterly journal of experimental psychology*, 26(3):368–372, 1974.
- Olga S Vinogradova. Hippocampus as comparator: role of the two input and two output systems of the hippocampus in selection and registration of information. *Hippocampus*, 11:578–598, 2001.
- An Vo and Soontorn Orintara. A study of relative phase in complex wavelet domain: Property, statistics and applications in texture image retrieval and segmentation. *Signal Processing: Image Communication*, 25(1):28–46, 2010.
- Richard. von Mises. Über die "Ganzzahligkeit" der Atomgewichte und verwandte Fragen. *Physikal. Z.*, 19:490–500, 1918.
- Johan Magnus von Wright, K Andersson, and Ulla Stenman. *Generalization of conditioned GSRs in dichotic listening*. In P. M. A. Rabbitt and S. Dornic(Eds), *Attention and Performance V*. New York: Academic Press, 1973.
- Verena Walpurger, Gabriele Hebing-Lennartz, Heide Denecke, and Reinhard Pietrowsky. Habituation deficit in auditory event-related potentials in tinnitus complainers. *Hearing research*, 181(1): 57–64, 2003.
- Jamie Ward. *The student's guide to cognitive neuroscience*. Psychology Press, 2015.
- Elizabeth M Wenzel, Marianne Arruda, Doris J Kistler, and Frederic L Wightman. Localization using nonindividualized head-related transfer functions. *The Journal of the Acoustical Society of America*, 94(1):111–123, 1993.
- Lisa E Williams, Jennifer Urbano Blackford, Andrew Luksik, Isabel Gauthier, and Stephan Heckers. Reduced habituation in patients with schizophrenia. *Schizophrenia research*, 151(1):124–132, 2013.
- Robert C Wilson, Matthew R Nassar, and Joshua I Gold. Bayesian online learning of the hazard rate in change-point problems. *Neural computation*, 22(9):2452–2476, 2010.
- Jeffery A Winer. A profile of auditory forebrain connections and circuits. In *The auditory cortex*, pages 41–74. Springer, 2011.
- Noelle Wood and Nelson Cowan. The cocktail party phenomenon revisited: how frequent are attention shifts to one's name in an irrelevant auditory channel? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(1):255, 1995.
- Geoffrey F Woodman. A brief introduction to the use of event-related potentials in studies of perception and attention. *Attention, Perception, & Psychophysics*, 72(8):2031–2046, 2010.
- Yoko Yamaguchi, Naoyuki Sato, Hiraoki Wagatsuma, Zhihua Wu, Colin Molter, and Yoshito Aota. A unified view of theta-phase coding in the entorhinal-hippocampal system. *Current Opinion in Neurobiology*, 17:197–204, 2007.

Nick Yeung, Rafal Bogacz, Clay B Holroyd, and Jonathan D Cohen. Detection of synchronized oscillations in the electroencephalogram: an evaluation of methods. *Psychophysiology*, 41(6): 822–832, 2004.

Nick Yeung, Rafal Bogacz, Clay B Holroyd, Sander Nieuwenhuis, and Jonathan D Cohen. Theta phase resetting and the error-related negativity. *Psychophysiology*, 44(1):39–49, 2007.

Basilis Zikopoulos and Helen. Barbas. Prefrontal projections to the thalamic reticular nucleus form a unique circuit for attentional mechanisms. *J Neurosci*, 26:7348–7361, 2006.

Basilis Zikopoulos and Helen. Barbas. Pathways for emotions and attention converge on the thalamic reticular nucleus in primates. *Journal of Neuroscience*, 32:5338–5350, 2012.

Appendices

A Cauchy Principal value

Definition A.1. Let f be a real-valued function that is defined on $[a, b] \setminus \{\xi\}$. If f is not defined near a point ξ in the interval $[a, b]$, the integral of f over $[a, b]$ does not always exist. The *improper integral* of f can be defined by summing the two integrals

$$\lim_{\epsilon \rightarrow 0^+} \int_a^{\xi-\epsilon} f(x) dx$$
$$\lim_{\epsilon \rightarrow 0^+} \int_{\xi+\epsilon}^b f(x) dx$$

If the two integrals don't exist, we can use the *symmetric limit*, known as the Cauchy Principal Value (PV) as follows:

$$\lim_{\epsilon \rightarrow 0^+} \left(\int_a^{\xi-\epsilon} f(x) dx + \int_{\xi+\epsilon}^b f(x) dx \right). \quad (1)$$

We write $PV \int_a^b f(x) dx$ for the integral of f between a and b .

B Fourier Transform of Complex Wavelet Function

The fourier transform of ψ , the sixth derivative of a complex Gaussian function, is

$$\begin{aligned}
\mathcal{F}(\psi) &= \mathcal{F}\left(c\left(e^{it}e^{-t^2}\right)^{(6)}\right) \\
&= c\mathcal{F}\left(c\left(e^{it}e^{-t^2}\right)^{(6)}\right) \\
&= c(2\pi i\omega)^6 \mathcal{F}\left(e^{it}e^{-t^2}\right)
\end{aligned} \tag{2}$$

This is based on the differentiation property of the Fourier transform which states that if a function f is a differentiable with f and f' is integrable, its Fourier transform is given by $\mathcal{F}f'(\omega) = 2\pi i\omega\mathcal{F}(f)$. More generally, the Fourier transform of n -th derivative $f^{(n)}$ is given by

$$\mathcal{F}f^{(n)}(\omega) = (2\pi i\omega)^n \mathcal{F}(f).$$

The Fourier transform of a function of the form $f(t)e^{iat}$ is $\mathcal{F}f\left(\omega - \frac{a}{2\pi}\right)$, that is, the Fourier transform of f , which is shifted in the frequency domain. In Eq. 2 we have $f(t) = e^{-t^2}$ and $e^{iat} = e^{it}$ with $a=1$. Therefore, substituting its Fourier transform in the Eq. 2 we obtain

$$\mathcal{F}(\psi) = c(2\pi i\omega)^6 \mathcal{F}\left(e^{-t^2}\right)\left(\omega - \frac{1}{2\pi}\right)$$

where the Fourier transform of $\mathcal{F}(e^{-t^2})$ is $\sqrt{\pi}e^{-(\pi\omega)^2}$. In general, the Fourier transform of a function $f = e^{-\alpha t^2}$ is $\mathcal{F}(f) = \sqrt{\frac{\pi}{\alpha}}e^{-\frac{(\pi\omega)^2}{\alpha}}$. However, because of the shift in the frequency, ω is replaced by $\omega - \frac{1}{2\pi}$. Therefore we have :

$$\begin{aligned}
\mathcal{F}(\psi) &= c(2\pi i\omega)^6 \sqrt{\pi}e^{-(\pi(\omega - \frac{1}{2\pi}))^2} \\
&= -c64\pi^6 \omega^6 \sqrt{\pi}e^{-(\pi\omega - \frac{1}{2})^2}
\end{aligned} \tag{3}$$

The factor c is chosen such that $\|F\psi\|_2 = \|\psi\|_2 = 1$. Summarizing we have:

$$\mathcal{F}(\psi) \propto -\omega^6 e^{-(\pi\omega - \frac{1}{2})^2}.$$

C Non-local Means Method

The details of the NLM method have been explained in Strauss et al. (2013). Here we briefly report on the parameters and the main idea behind the technique. ERP images contain a high degree of self-similarity over individual trials. This is mainly because of the N1-P2 components of ERPs that are evoked after the onset of stimulus. We exploit the characteristic correlations of ERP features across the rows in ERP images. Let $\mathbf{S} \in \mathbb{R}^{N \times M}$ represent the matrix of ERP trials. There are N

trials, each with a length of M samples. We want to replace every sample $s_i, i = 1, \dots, J$ ($J = NM$) in the matrix with a denoised version based on the structural similarities in the signals. Every pixel s_i and its neighborhood is compared to the other ERP patches. For every comparison, a weight coefficient $\eta_{i,j}$ is assigned to the center sample s_i based on the degree of the similarity of the patches. The sample s_i is replaced with the denoised sample q_i , which is computed as:

$$q_i = \frac{1}{\gamma} \sum_{j=1}^J \eta_{i,j} s_j \quad (4)$$

where $\gamma_i = \sum_{j=1}^J \eta_{i,j}$. We denote the image patches with center samples s_i and s_j as s_{i+I} and s_{j+I} respectively. The weights $\eta_{i,j}$ are computed based on the similarity between the patches:

$$\eta_{i,j} = \exp \left(-\frac{1}{\lambda} \sum_{k \in I} \varphi_{\sigma,k} |s_{i+I} - s_{j+I}|^2 \right) \quad (5)$$

where $\varphi_{\sigma,k}$ denotes a sampled version of a 2-D Gaussian kernel with standard deviation of σ . The parameter σ determines the degree of influence of neighboring patches on the weight and the parameter $\lambda > 0$ controls the amount of denoising. For more details see Strauss et al. (2013). The parameters that are used for this study are a fixed 1×11 similarity patch, an asymmetric Gaussian with $\sigma = (1.0, 5.0)^T$ and $\lambda = 1000$ being the denoising parameter.

D Derivation of Forward Pass of Abrupt CPM

Defining $\alpha_i^t = P(\Theta_{1:t}, s_t = i)$, we have:

$$\begin{aligned} \alpha_i^t = P(\Theta_{1:t}, s_t = i) &= P(\theta_1, \dots, \theta_{t-1}, \theta_t, s_t = i) \\ &= P(\theta_t | \Theta_{1:t-1}, s_t = i) P(\Theta_{1:t-1}, s_t = i) \\ &= P(\theta_t | \Theta_{1:t-1}, s_t = i) \sum_{j=1}^{t-1} P(\Theta_{1:t-1}, s_t = i, s_{t-1} = j) \\ &= P(\theta_t | \Theta_{1:t-1}, s_t = i) \sum_{j=1}^{t-1} P(s_t = i | s_{t-1} = j) P(s_{t-1} = j, \Theta_{1:t-1}) \\ &= P(\theta_t | \theta_{1:t-1}, s_t = i) \sum_{j=1}^{t-1} P(s_t = i | s_{t-1} = j) \alpha_j^{t-1} \end{aligned}$$

where $P(\theta_t | \theta_{1:t-1}, s_t)$ is the emission distribution, $P(s_t = i | s_{t-1} = j)$ is the transition distribution and α_j^{t-1} is the recursive part.

E Derivation of the Posterior Distribution Over the States in CPM

$$\begin{aligned}
P(\mu_t, \kappa_t | \theta_{1:N}) &= \frac{P(\mu_t, \kappa_t, \theta_{1:N})}{P(\theta_{1:N})} \\
&\propto P(\mu_t, \kappa_t, \theta_{1:t}, \theta_{t+1}, \theta_{t+2:N}) \\
&\propto P(\theta_{t+1} | \theta_{1:t}, \theta_{t+2:N}, \mu_t, \kappa_t) P(\theta_{1:t}, \theta_{t+2:N}, \mu_t, \kappa_t) \\
&\propto P(\theta_{t+1} | \mu_t, \kappa_t) P(\mu_t, \kappa_t | \theta_{1:t}, \theta_{t+2:N}) \\
&\propto P(\theta_{t+1} | \mu_t, \kappa_t) P(\mu_t, \kappa_t | \theta_{1:t})
\end{aligned}$$

F Derivation of Weights

The procedure for ρ optimization:

```

* The parameters of the algorithm are set as follows: *
* numStimuli = 4, referring to the different stimuli levels. *
* numSub = 20, referring to the subjects used in this study. *
* N, the number of trials per subject/stimulus level. *
* Matrix Initializations:  $\mathbf{W} \in \mathbb{R}^{\text{numSub} \times N}$ ,  $\mathbf{A} \in \mathbb{R}^{\text{numStimuli} \times N}$ ,  $\mathbf{V} \in \mathbb{R}^{\text{numStimuli} \times N}$ .
* Vector Initializations:  $\mathbf{v}' \in \mathbb{R}^{\text{numStimuli} \times 1}$ ,  $\tilde{\mathbf{v}} \in \mathbb{R}^{N \times 1}$ .
for each candidate configuration  $(\sigma^2, K)$  do
  for  $g=1:\text{numStimuli}$  do
    for  $i=1:\text{numSubjects}$  with data set  $\Theta_{i,g}$  do
      for  $t=1:N$  do
        Compute  $p(\kappa_t, \mu_t | \Theta_{i,g}, \sigma^2, K)$  using the forward-backward algorithm.
        Marginalize over  $\mu_t$ :  $p(\kappa_t | \Theta_{i,g}, \sigma^2, K)$ ;
        Compute the expected value  $E[\kappa_t] = \sum_{\kappa_j \in \kappa} \kappa_j p(\kappa_t = \kappa_j | \Theta_{i,g}, \sigma^2, K)$ ;

        *Store the result in the matrix  $\mathbf{W}_g$ .*
         $\mathbf{W}_g(i, t) = E[\kappa_t]$ ;
      end

      *Adjust for subject-specific factor in concentration values:*
      * divide  $\kappa$  by mean  $\kappa$  of the last 50 data points*
       $\mathbf{W}_g(i, :) = \mathbf{W}_g(i, :) / \text{avg}(\mathbf{W}_g(i, (N - 50) : N))$ ;
    end

    *Compute the group mean and in-group variance at each time  $t$ *
    for  $t=1:N$  do
      *Store the group mean at time  $t$ , averaging over all subjects.*
       $\mathbf{A}(g, t) = \text{avg}(\mathbf{W}_g(:, t))$ ;
      *Store the in-group variance at time  $t$ .*
       $\mathbf{V}(g, t) = \text{variance}(\mathbf{W}_g(:, t))$ ;
    end
    Average over in-group variances at different  $t$ :  $\mathbf{v}'(g) = \text{avg}(\mathbf{V}(g, :))$ .
  end

  *Compute between-groups variance, at each time  $t$ *
  for  $t=1:N$  do
     $\tilde{\mathbf{v}}(t) = \text{variance}_g(\mathbf{A}(:, t))$ 
  end
  Average over between-group variances at different  $t$ :  $\tilde{\mathbf{v}}' = \text{avg}(\tilde{\mathbf{v}})$ 

  *Compute the optimization criteria*
   $\rho(\sigma^2, K) = \frac{\tilde{\mathbf{v}}'}{\text{avg}_g(\mathbf{v}'(g))}$ ;
end

```

Algorithm 3: The algorithm used for computing the optimization criteria ρ for each set of prior parameter candidates (K_i, σ_i^2) .

G Derivation of Probability Density Functions

We assume that λ has a complex gaussian distribution with mean μ and standard deviation σ . The probability density function of λ is

$$P(\lambda) = \frac{1}{2\pi\sigma^2} e^{-\frac{|\lambda-\mu|^2}{2\sigma^2}}$$

In cartesian coordinates, for $\lambda = x + iy$ we have

$$P(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x-\Re\{\mu\})^2 + (y-\Im\{\mu\})^2}{2\sigma^2}}$$

From this we obtain the probability density function for λ in polar coordinates, where $\lambda = Ae^{i\theta}$. It is convenient to express μ in polar coordinates as well. In the remainder, let $\mu = Be^{i\phi}$.

$$\begin{aligned} P(A, \theta) &= AP(\Re\{Ae^{i\theta}\}, \Im\{Ae^{i\theta}\}) \\ &= \frac{A}{2\pi\sigma^2} e^{-\frac{|Ae^{i\theta} - Be^{i\phi}|^2}{2\sigma^2}} \\ &= \frac{A}{2\pi\sigma^2} e^{-\frac{A^2 - 2\Re\{Ae^{i\theta}Be^{-i\phi}\} + B^2}{2\sigma^2}} \\ &= \frac{A}{2\pi\sigma^2} e^{-\frac{A^2 - 2AB\cos(\theta-\phi) + B^2}{2\sigma^2}} \end{aligned}$$

By Bayes's rule we have $P(\theta | A) = \frac{p(A, \theta)}{P(A)}$. To derive $P(A)$, we marginalize over θ , that is

$$\begin{aligned} P(A) &= \int_0^{2\pi} p(A, \theta) d\theta \\ &= \int_0^{2\pi} \frac{A}{2\pi\sigma^2} e^{-\frac{A^2+B^2}{2\sigma^2}} e^{\frac{AB}{\sigma^2}\cos(\theta-\phi)} d\theta \\ &= \frac{A}{2\pi\sigma^2} e^{-\frac{A^2+B^2}{2\sigma^2}} \int e^{\frac{AB}{\sigma^2}\cos(\theta-\phi)} d\theta \\ &= \frac{A}{2\pi\sigma^2} e^{-\frac{A^2+B^2}{2\sigma^2}} \int e^{\frac{AB}{\sigma^2}\cos(\theta)} d\theta \\ &= \frac{A}{2\pi\sigma^2} e^{-\frac{A^2+B^2}{2\sigma^2}} 2\pi I_0\left(\frac{AB}{\sigma^2}\right) \\ &= \frac{A}{\sigma^2} e^{-\frac{A^2+B^2}{2\sigma^2}} I_0\left(\frac{AB}{\sigma^2}\right) \end{aligned}$$

The derived pdf is also known as the *Rice distribution* and I_0 is the *modified Bessel function* of order zero. Using Bayes rule we obtain

$$P(\theta | A) = \frac{1}{2\pi} e^{\frac{AB}{\sigma^2}\cos(\theta-\phi)} / I_0\left(\frac{AB}{\sigma^2}\right)$$

In particular, $P(\theta | A)$ is a von-Mises distribution with concentration $\kappa = \frac{AB}{\sigma^2}$. The mean resultant vector is

$$\begin{aligned}
 E[e^{i\theta} | A] &= \int_0^{2\pi} e^{i\theta} \frac{1}{2\pi I_0(\kappa)} e^{\kappa \cos(\theta-\phi)} d\theta \\
 &= \frac{1}{2\pi I_0(\kappa)} \int_0^{2\pi} e^{\kappa \cos(\theta-\phi)+i\theta} d\theta \\
 &= \frac{1}{2\pi I_0(\kappa)} e^{i\phi} \int_0^{2\pi} e^{\kappa \cos(\theta)+i\theta} d\theta \\
 &= \frac{1}{2\pi I_0(\kappa)} e^{i\phi} 2\pi I_1(\kappa) \\
 &= \frac{I_1(\kappa)}{I_0(\kappa)} e^{i\phi}
 \end{aligned}$$

Using the law of total expectation this allows us to determine the mean resultant vector of θ .

$$\begin{aligned}
 E[e^{i\theta}] &= E[E[e^{i\theta} | A]] \\
 &= E \left[\int P(e^{i\theta}|A) e^{i\theta} dA \right] \\
 &= e^{i\theta} E_A \left[I_1 \left(\frac{AB}{\sigma^2} \right) / I_0 \left(\frac{AB}{\sigma^2} \right) \right] \\
 &= \left[\int_0^\infty \frac{A}{\sigma^2} e^{-\frac{A^2+B^2}{2\sigma^2}} I_0 \left(\frac{AB}{\sigma^2} \right) I_1 \left(\frac{AB}{\sigma^2} \right) / I_0 \left(\frac{AB}{\sigma^2} \right) dA \right] e^{i\theta} \\
 &= \left[\int_0^\infty \frac{A}{\sigma} e^{-\frac{1}{2} \left(\frac{A}{\sigma} \right)^2} e^{-\frac{B^2}{2\sigma^2}} I_1 \left(\frac{A}{\sigma} \frac{B}{\sigma} \right) dA \right] e^{i\theta} \\
 &= \left[\int_0^\infty A e^{-\frac{A^2}{2}} I_1 \left(A \frac{B}{\sigma} \right) dA \right] (e^{i\theta}) \left(-e^{-\frac{B^2}{2\sigma^2}} \right) \\
 &= \left[g e^{-g^2} \sqrt{\frac{\pi}{2}} (I_0(g^2) + I_1(g^2)) \right] e^{i\phi} \tag{6}
 \end{aligned}$$

With $g = \frac{B}{2\sigma}$.

Thus we obtain a closed form for the circular standard deviation S of θ .

$$\begin{aligned} S &= \sqrt{-2 \ln(|E[e^{i\theta}]|)} \\ &= \sqrt{2g^2 - 2 \ln(g) - \ln(\pi/2) - 2 \ln(I_0(g^2) + I_1(g^2))} \end{aligned}$$

which has values between 0 and infinity.

List of Publications

Journal Publications

Z. Morteza pouraghdam, F. I. Corona-Strauss, K. Takahashi, D. J. Strauss.

Reducing the Effect of Spurious Phase Variations in Neural Oscillatory Signals, *Frontiers in Computational Neuroscience*, Vol. 12, 2018.

Z. Morteza pouraghdam, R. C. Wilson, L. Schwabe, D.J. Strauss.

Bayesian Modeling of the Dynamics of Phase Modulations and their Application to Auditory Event Related Potentials at Different Loudness Scales, *Frontiers in computational neuroscience*, Vol. 10, 2016.

Z. Morteza pouraghdam, L. Haab, F. I. Corona-Strauss, G. Steidl, D. J. Strauss.

Assessment of Long-Term Habituation Correlates in Event-Related Potentials Using a von Mises Model, *IEEE Trans Neural Syst Rehabil Eng*, pp. 363-373, Vol. 23, 2015.

Z. Morteza pouraghdam, F. I. Corona-Strauss, D. J. Strauss.

* Mapping the Effect of Selective Attention between Ongoing EEG Activities and Averaged ERPs by means of Instantaneous Phase Information, *In preparation for submission to IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2019.

Conference Publications

Z. Morteza pouraghdam, D. J. Strauss, C. Bernarding.

Objective Assessment of Perceived Effort in Listening by Employing EEG Features, *In 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2908-2911, 2017.

Z. Morteza pouraghdam, D. J. Strauss.

Removal of Spurious Phase Resets in Oscillatory Signals, *In 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2209-2212, 2017.

Z. Morteza pouraghdam, L. Haab, K. Schwerdtfeger, and D. J. Strauss.

Relating Auditory Evoked Responses to the Laminar Phase Dynamics in Rats Using Mutual Information, *In 7th Annual International Conference of the IEEE Neural Engineering*, pp. 952-955, 2015.

Z. Morteza pouraghdam, L. Haab, G. Steidl, D. J. Strauss.

Detection of Change Points in Phase Data: A Bayesian Analysis of Habituation Processes, *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1014-1017, 2014.

L. Haab, **Z. Morteza pouraghdam**, D. J. Strauss.

Modeling prediction of a generalized habituation deficit in decompensated tinnitus sufferers, *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5691-5694, 2014.

Danksagung

I would like to extend my sincere gratitude to my advisor Prof. Daniel Strauss for guiding and mentoring me over the years. Because of him, I was able to work in the exciting interdisciplinary area of computational neuroscience and achieve invaluable experiences at different topics in the field. I also thank him for providing me with the opportunities to interact with many different researchers. Throughout my collaboration with Princeton University, I learned a lot from Dr. Robert Wilson, especially in the area of change point modeling using Bayesian models. Under his skilled supervision, I was able to publish my second paper on the Bayesian modeling of instantaneous phase data. I thank him wholeheartedly for providing me with a robust learning experience during my short visit. Special thanks to Prof. Steidl for all of her support and patience during my time at Kaiserslautern University. Because of her guidance, I was able to improve my writing style and publish my first paper.

For their help during my graduate school years and beyond, I thank Prof. Klakow, Dr. Rahil Mahdian Toroghi and Dr. Friedlich Faubel. I particularly would like to thank Dr. Rahil Mahdian for his time and patience in addressing my questions and problems. I am truly grateful for the generous amount of time he spent helping mold me into a better researcher.

I would like to extend my sincere thanks to Prof. Lars Schwabe. I am thankful for all the time he has dedicated to my scholarship during and after graduate school. I am also grateful that I was able to collaborate with him on my second paper; working with him is always an invaluable learning experience.

Thank you to Dr. Farah Corona-Strauss, Dr. Lars Haab, Dr. Corinna Bernarding and Dr. Takahashi for all of their feedback on my thesis and the different projects I participated in; I learned a lot from their expertise.

I also appreciate the time that I spent at the graduate school and systems neuroscience & neurotechnology unit (SNNU), in which I got to know some of my best friends, Narsis Salafzoon and

Daniel Mewes. Last but not least, I thank my family for their unconditional support and constant encouragement throughout the years. In particular, I would like to thank my husband, Steven, for everything.