QATAR UNIVERSITY

COLLEGE OF ENGINEERING

DESIGN OF A SIMULATOR FOR NEONATAL MULTICHANNEL EEG: APPLICATION TO TIME-FREQUENCY APPROACHES FOR AUTOMATIC ARTIFACT REMOVAL AND SEIZURE DETECTION

BY

MOHAMMAD FATHI AL-SA'D

A Thesis Submitted to the Faculty of

Engineering

in Partial Fulfilment

of the Requirements

for the Degree of

Master of Science

January 2016

COMMITTEE PAGE

The members of the Committee approve the thesis of Mohamed Fathi Al-Sa'd defended on 20th of January 2016.

_____

Professor Boualem Boashash
Thesis Supervisor

_____

Professor Abdeldjalil Aïssa-El-Bey
Committee Member

_____

Professor Serkan Kiranyaz
Committee Member

_____

Doctor Uvais Qidwai
Committee Member

Approved:

_____

Dr Rashid Al-Ammari, Dean, College of Engineering

ABSTRACT

The electroencephalogram (EEG) is used to noninvasively monitor brain activities; it is the most utilized tool to detect abnormalities such as seizures.

In recent studies, detection of neonatal EEG seizures has been automated to assist neurophysiologists in diagnosing EEG as manual detection is time consuming and subjective; however it still lacks the necessary robustness that is required for clinical implementation. Moreover, as EEG is intended to record the cerebral activities, extra-cerebral activities external to the brain are also recorded; these are called “artifacts” and can seriously degrade the accuracy of seizure detection.

Seizures are one of the most common neurologic problems managed by hospitals occurring in 0.1%-0.5% livebirths. Neonates with seizures are at higher risk for mortality and are reported to be 55-70 times more likely to have severe cerebral-palsy. Therefore, early and accurate detection of neonatal seizures is important to prevent long-term neurological damage.

Several attempts in modelling the neonatal EEG and artifacts have been done, but most did not consider the multichannel case. Furthermore, these models were used to test artifact or seizure detection separately, but not together. This study aims to design synthetic models that generate clean or corrupted multichannel EEG to test the accuracy of available artifact and seizure detection algorithms in a controlled environment.

In this thesis, synthetic neonatal EEG model is constructed by using; single-channel EEG simulators, head model, 21-electrodes, and propagation equations, to produce clean multichannel EEG. Furthermore, neonatal EEG artifact model is designed using synthetic signals to corrupt EEG waveforms. After that, an automated EEG artifact

detection and removal system is designed in both time and time-frequency domains. Artifact detection is optimised and removal performance is evaluated. Finally, an automated seizure detection technique is developed, utilising fused and extended multichannel features along a cross-validated SVM classifier.

Results show that the synthetic EEG model mimics real neonatal EEG with 0.62 average correlation, and corrupted-EEG can degrade seizure detection average accuracy from 100% to 70.9%. They also show that using artifact detection and removal enhances the average accuracy to 89.6%, and utilising the extended features enhances it to 97.4% and strengthened its robustness.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Firstly, I would like to express my sincere gratitude to my supervisor Professor Boualem Boashash, for his excellent guidance, caring, patience, and support throughout my master's thesis. Your feedbacks had sharpened my knowledge, and our scientific debates had broadened my mind.

Beside my supervisor, I would like to thank Dr Tamer Khattab, Dr Andrei Sleptchenko, Dr Nabeel Khan, Dr Sadiq Ali, and Dr Samir Ouelha for their enormous support and help.

My sincere thanks also goes to the Department of Electrical Engineering at Qatar University for their support throughout my thesis and master studies.

Last but not the least, I would like to thank my family and my fiancée for their continuous support and encouragement. Your presence enriches my life, and helps me to move forward.

DEDICATION

I dedicate this work to my parents

Fathi and Asmahan Al-Sad and to my lovely fiancée

Wafa Saad

CHAPTER 1: INTRODUCTION

1.1 Background and Motivation

Neonatal seizures are one of the most common and crucial signs of acute neonatal encephalopathy. They represent a major risk of death or morbidity, and may contribute to an adverse neurodevelopmental outcomes. Seizures occur as reported in population-based studies, in 1 to 5 per 1000 live births, increasing to 11.1% for preterm neonates and to 13.5% for infants with a birth weight less than 2500 grams. Furthermore, neonates with seizures are reported to be 55 to 70 times more likely to have severe cerebral palsy [1] [2] [3] [4] [5].

Clinical signs of seizures may not be present in as many as 85% of neonates. In addition, treating these seizure events with antiepileptic drugs may further suppress the clinical signs while electrographic seizures persist. Consequently, electroencephalogram (*EEG*) monitoring is essential to accurately diagnose neonatal seizure events and to evaluate the efficacy of treatment [6].

Few staff in Neonatal Intensive Care Unit (*NICU*) receive sufficient training to interpret EEG patterns. Therefore, automated neonatal seizure detectors were introduced in the NICU to assist neurophysiologists in diagnosing EEG, as manual detection is time consuming and subjective. Current seizure detection methods still lack the necessary robustness that is required for clinical implementation, due the variability of EEG recordings in time and among patients [7] [8]. Such variability is inevitable, as the morphology of background and seizure patterns change significantly among different patients, and is greatly affected by the conceptual age (*CA*) of the patient. Moreover, as EEG is intended to record the cerebral activities, extra-cerebral activities external to the

brain are also recorded; these are called “artifacts”. Artifacts are undesired signals originating from different sources that can introduce changes in the EEG recordings. They can mimic abnormal EEG patterns, thus deforming the EEG information content. EEG artifacts such as cardiac activity, ocular movements, eye blinks and muscular activity are among the most common types of artifacts and can have durations ranging from seconds to hours [6] [9] [10] [11].

Neonatal seizure detection accuracy and robustness can be enhanced by testing and validating the system using large, annotated, clinical databases [12]. Unfortunately, obtaining such databases requires intense effort, and their protection policies restrain independent scientists from comparing their results on a common ground basis. Therefore, modelling neonatal EEG is needed to provide a common ground database. Such models can assist the development of new features for the automated seizure detection systems, and can facilitate comparisons of neonatal seizure detection algorithms. Furthermore, it can be used for calibrating new EEG machines, and can serve as an alternative when patient’s confidentiality and ethical laws prohibit the use of real EEG [13].

1.2 Project Characteristics

1.2.1 Definition and Objectives

The project is defined as the design of a simulator for neonatal multichannel EEG, and its application as a test bed for automatic artifact detection and removal for abnormality detection systems. The project investigates whether neonatal seizure detection rates can be enhanced using; time and time-frequency artifact detection and removal techniques, and time-frequency features based on channels alikeness. Thus, the objective of this

thesis is to design synthetic models that generate clean or corrupted neonatal multichannel EEG to test the accuracy of available artifact and seizure detection algorithms in a controlled environment.

1.2.2 Scope

The project scope is defined as designing the neonatal multichannel EEG model by assessing validated: simulators, neonatal head model, and multichannel propagation scheme. Furthermore, the automatic artifact detection and removal techniques use time and time-frequency signal processing methods. Such methods are preferred more than frequency domain techniques, because; they require standard real-valued algorithms rather than complex-valued [14], they do not omit the non-stationary behaviour of EEG [15], and additionally time-frequency methods produce unique signatures for different signals and are not limited to non-Gaussian signals [15] [16]. Finally, the automatic abnormality detection system utilise: time-frequency feature extraction methodology, features fusion and selection techniques, and a machine learning algorithm.

1.2.3 General Overview

The general overview of the project stages is depicted in Figure 1.1. Synthetic neonatal multichannel EEG waveforms are produced using the developed model along with synthetic multichannel artifacts. Both signals are added to produce contaminated version of the clean multichannel EEG. After that, multichannel time-frequency representations are computed using Spatial Time-Frequency Distributions (*STFD*) and plugged into the automatic artifact detection and removal system along with the time domain contaminated version. Time and time-frequency artifact detection and removal

algorithms are computed and the results are finally plugged into the automatic EEG abnormality detection to recognise seizure events.

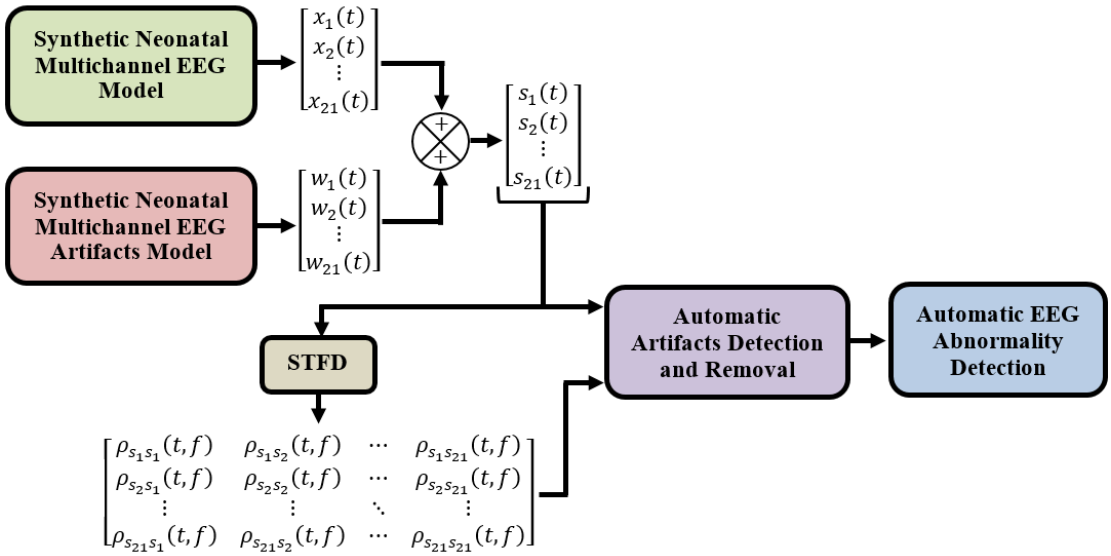

Figure 1.1: Schematic overview of the project stages

1.3 Contributions

The most important contributions of this thesis to the neonatal EEG modelling and detection are five folds.

1. Modelling synthetic neonatal multichannel EEG: We developed a model that mimics real life neonatal multichannel EEG by using: single-channel EEG simulators, head model, 21 electrodes, and propagation equations, and which is controlled by user inputs to alter the properties of produced patterns. The synthetic neonatal multichannel EEG has been validated against real neonatal multichannel EEG using Spearman's correlation. Validation results show 0.58065 and 0.57796 mean correlations when comparing synthetic and real background epochs, and 0.65107 and 0.66254 when comparing seizure epochs in the frequency and time-frequency domains respectively.

2. Modelling multichannel EEG propagation: We developed a novel neonatal EEG propagation model by utilising the radiation transport equation, which describes the dispersion and decay in light intensity due absorption, scattering, and reflection when propagating through biological tissues. The developed model is simple, descriptive, and require low computational power when compared to current modelling techniques. The model has been validated by comparing its seizure-to-background ratio distribution against a validated reference using Probability-Probability plots. The resultant normal line show an R^2 value of 0.99999036 when fitting the reference, thus proving the goodness of the model.
3. Modelling synthetic neonatal multichannel EEG artifacts: We developed a model that contaminates the synthetic clean EEG, and which is also controlled by user inputs to alter the properties of produced artifacts. The produced EEG artifacts effectively supress the EEG information content and lower the seizure detection average accuracies from 100% to; 99.5%, 98.2%, 93.0 %, 82.8%, and 70.9% when using signal-to-artifact ratios of; 0, -7.2206, -13.2412, and -16.763, and -19.262 dB respectively, thus mimicking real life contamination.
4. Automated artifacts detection and removal: We automated and extended time domain Blind Source Separation (*BSS*) algorithms into the joint time-frequency and statistically optimised their detection performances. They effectively detect contaminated EEG with mean detection rates of 97% and 90.4% in the time and time-frequency domains respectively. Furthermore, they maintained their detection accuracies when using signal-to-artifact ratios of; 0, -7.2206, -13.2412, and -16.763, and -19.262 dB, thus proving their robustness to contamination level. Moreover, seizure detection mean accuracy have increased

from 70.9% to 89.6% when removing artifacts from corrupted EEG with -19.262 dB SAR using time domain Fast-ICA, thus enhancing the overall system performance (by 18.7% increment).

5. Extended multichannel features: We introduced extended time-frequency multichannel features that are based on information alikeness between EEG channels. By utilising the extended features, seizure detection average accuracies have increased from 93.0% to 97.8%, 82.8% to 97.6%, and from 70.9% to 97.4% when using corrupted EEG with SAR values of -13.2412, -16.763, and -19.262 dB respectively (4.8%, 14.8% and 26.5% increments). This proves the effectiveness of the extended features in enhancing the seizure detection performance.

Additional contributions are:

1. We developed an EEG corruption evaluation methodology using various statistical approaches. This methodology quantifies the artifact masking effect on the EEG information contents, thus assessing the EEG contamination level.
2. We introduced a classification accuracy profiling to illustrate the effect of feature subset size on the accuracy of the automated seizure detection system. This provides a benchmark to choose a number of feasible features that maximises the system accuracy and robustness.

1.4 Thesis Organisation

Chapter 2: Theoretical Background and Literature Review

Chapter 2 provides a review of the fundamentals needed to proceed in this thesis along with the literature reviews. The chapter starts with an explanation of the anatomical

structure of the neonatal head and its current modelling techniques. After that, fundamentals of EEG signals and its normal and abnormal behaviours are explained along with their modelling techniques. Furthermore, methods and techniques for processing the EEG signals are reviewed to understand their adequacy and limitations. Finally, it presents a review on blind source separation methods and fundamentals of pattern recognition to summarise the edge of knowledge in; detecting and removing EEG artifacts, and detecting neonatal seizures.

Chapter 3: Modelling Approaches and Techniques

Chapter 3 provides in depth explanations for the modelling approaches and techniques used in this thesis along with all the procedures undertaken. First, it illustrates an analysis for real neonatal EEG data, then it explains the synthetic neonatal multichannel EEG modelling technique that is based on: time-frequency synthetic simulators, a head model, and a novel EEG propagation scheme. Furthermore, it explains the EEG propagation model validation procedure using Probability-Probability plots. Finally, it explains the construction of the synthetic neonatal multichannel EEG artifacts model that is based on using simulated signals to contaminate the clean EEG.

Chapter 4: Application to Time-Frequency Algorithms Enhancement for Abnormality Detection

Chapter 4 provides in depth explanations for the abnormality detection time-frequency enhancement techniques used in this thesis along with all the procedures undertaken. First, it explains the construction of the automatic artifact detection and removal system using BSS in time and time-frequency domains to filter the contaminated EEG patterns.

After that, an optimisation procedure for the automated artifact detection technique is presented along with an evaluation for the artifact removal quality to quantify the system performance. Furthermore, it demonstrates the construction of the automated EEG abnormality detection system, embedding fused and extended multichannel EEG features to detect neonatal seizures. Finally, validation of the automated EEG abnormality detection performance using cross-validation techniques is explained.

Chapter 5: Results and Discussions

Results and outcomes of the methods and techniques used in chapter 3 and 4 are presented in this chapter, along with their discussions and interpretations. First, the synthetic neonatal multichannel EEG and artifacts models results are illustrated to present the clean and corrupted multichannel waveforms. After that, results of the automated artifact detection and removal system using BSS in the time and time-frequency domains along with the automated EEG abnormality detection system, are presented and discussed to assess the system performance enhancement.

Chapter 6: Conclusions and Future Perspectives

The last chapter summarises the important findings of this thesis and tests the validity of the null hypothesis of this work. It also presents several suggestions and perspectives for future research and development.

CHAPTER 2: THEORETICAL BACKGROUND AND LITERATURE REVIEW

In this chapter, a review of the fundamentals needed to proceed in this thesis are presented along with the literature reviews. First, the anatomy of the neonatal head and its modelling techniques are discussed to understand the complexity of the Electroencephalogram source environment. After that, the EEG fundamentals and its modelling techniques are discussed to create a baseline in discriminating normal and abnormal signal behaviours. Furthermore, a review on the neonatal EEG processing methods is presented to illustrate the adequacy of such techniques along with their limitations. Finally, blind source separation methods and fundamentals of pattern recognition are reviewed to summarise the current techniques in; detecting and removing EEG artifacts, and detecting neonatal seizures.

2.1 The Neonatal Head

2.1.1 Anatomy of the Neonatal Head

The neonatal head shares the same main structures with adults. It consists of a scalp, skull, cerebrospinal fluid (*CSF*), and a brain (Figure 2.1). The scalp is the soft tissue envelope of the cranial vault containing the skull, CSF, and the brain, while the skull is the bony structure that mainly forms the head and protects inner soft tissues. The CSF is a clear, colourless body fluid found in the brain and spine, which provides basic mechanical and immunological protection to the brain. Finally, the brain is the main organ of the human nervous system [17].

The brain consists of two symmetric hemispheres and can be divided into cerebrum, brain stem and cerebellum. The cortex, or upper portion of the brain, is composed of

four lobes namely; frontal, parietal, occipital and temporal lobes (Figure 2.1). Each of these lobes are subdivided into many areas that control different functions. The cortex is about 1.5-4.0 mm thick and consists of nerve cells, which are darker in colour and therefore called grey matter. The larger inner part of the brain consists only of axons of the nerve cells and is lighter in colour, therefore called the white matter [17].

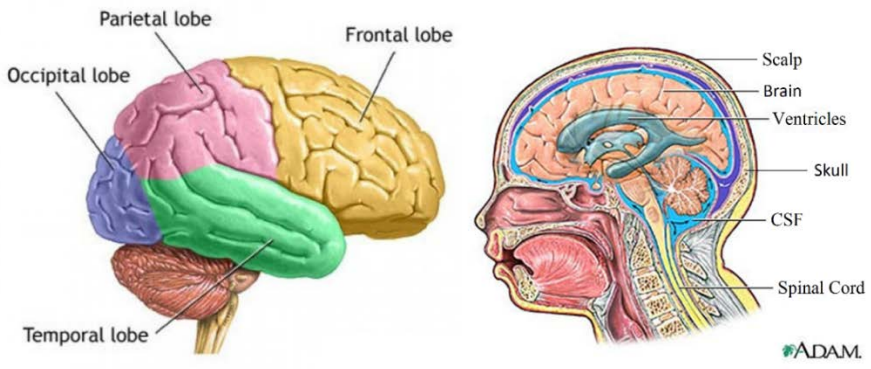

Figure 2.1: Brain lobes and head structures [Redrawn from www.md-health.com]

The brain is made of about 10^{10} nerve cells called neurons. Neurons consist of a cell body, dendrites, nucleus, an axon, and synaptic terminals (Figure 2.2). Dendrites extend from the cell body and share with it the function of receiving information from synaptic connections from adjoining neurons. Information is then transmitted to the next neuron through the axon and the synaptic terminals [17].

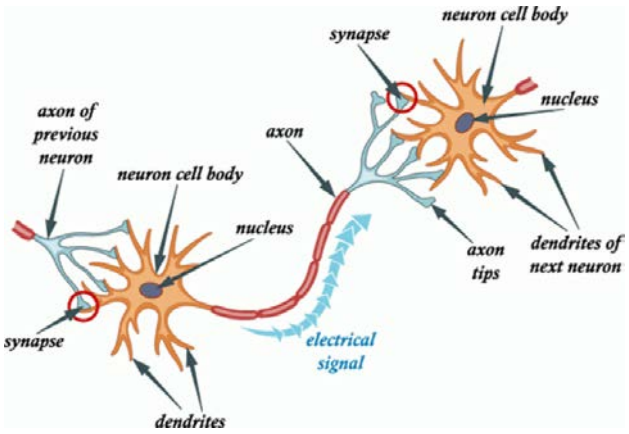

Figure 2.2: Neuron structure and information transmission [Taken from www.pixgood.com]

2.1.2 Neonatal Head Modelling

Spherical and realistic neonatal head models had been successfully used in different areas such as; solving EEG forward and inverse problems, and validation of optical tomographic images. Several spherical head models had been proposed in the literature such as; homogeneous spherical model, three sphere model, four sphere model, isotropic multi-sphere model, anisotropic multi-sphere model, and five layer model.

Despotovi et al. had proposed using the four sphere model in [18] and [19] to mimic the neonatal head by dividing the head structure into four concentric spheres; scalp, skull, CSF, and the brain. They had used this model in [18] to present a pipeline for modelling realistic volume conductor models for the neonatal head, and they addressed the challenges and presented their solutions. They had also used this model in [19] to develop an integrated method for ictal EEG dipole source localization based on a realistic head model, and had investigated the utility of EEG source imaging in neonates with postasphyxial seizures.

Another adaption for the four sphere model was used by *Vatta et al.* in [20]. They compared different spherical and realistic head modelling techniques to estimate the EEG forward solution from current dipole sources, distributed on a standard cortical space, and reconstructed from Montreal Neurological Institute Magnetic Resonance Imaging (*MRI*) data. They presented computer simulations for three different four-shell head models, two with realistic geometries; either surface-based (*BEM*) or volume-based (*FDM*), and the corresponding sensor-fitted spherical-shaped model.

Sadleir et al. had also adapted the four sphere model in [21]. They had used the model to estimate the sensitivity of three 16-electrode and impedance measurement configurations to bleeding in premature neonates.

Ala et al. had proposed using the isotropic multi-sphere model in [22]. They built a multi-sphere particle method in order to estimate the solution of the Poisson's equation with Nuemann boundary conditions describing the neuronal brain activity.

Hayde et al. had proposed adapting the anisotropic multi-sphere model in [23]. Their study examined the effects that partial volume errors in CSF segmentations had upon the electromagnetic source localization bioelectric model. They introduced a new approach for using estimates of partial volume fractions in the construction of patient specific bioelectric models.

Gupta et al. had used the five sphere model in [24]. They studied the effect of taking white matter and grey matter into account while solving the forward EEG problem. Their work shows that five sphere head models, comprising scalp, skull, CSF, grey matter and white matter, are more accurate in finding the forward EEG solution in comparison to four sphere models.

Brigadoi et al. had also adapted the five sphere model in [25]. They presented a 4D neonatal head model (for each week from 29 to 44 weeks post-menstrual age) that included a multi-layered tissue mask identifying; CSF, grey matter, white matter, cerebellum and brainstem. They developed an open source package which can be applied by users of near-infrared spectroscopy and diffuse optical tomography. It can be used to optimise probe locations and image reconstruction, register data to cortical locations, and ultimately improve the accuracy and interpretation of diffuse optical techniques in neonatal populations.

2.2 Neonatal Electroencephalogram (EEG)

2.2.1 Electrical Activities and Scalp EEG

The sources of the electrical activity in the brain measurable with EEG are cortical. Each synapse behaves like a voltage source driving current in a small loop. Electric potentials around individual cells are negligible, and not recordable at the scalp surface, except for the fact that the pyramidal cells are all aligned perpendicular to the surface of the cortex. Therefore, if activities were synchronous, the voltage fields produced by individual cells accumulate to produce a relatively larger potential that is recordable at the scalp. Such spatiotemporal synchronisation of neural networks results in net polarisation of extended brain regions, which may be transient, slow, or oscillatory [26].

A widely used standard for placing and denoting EEG electrodes on the scalp is the 10-20 system [27] [28] [29] [30]. It is an internationally recognized method that defines the location of the scalp electrodes, and it provides a uniform coverage of the entire scalp. The system is based on the relationship between the location of an electrode and the underlying area of cerebral cortex. The numbers 10 and 20 correspond to the fact that the distances between adjacent electrodes are either 10% or 20% of the total front-back or right-left distance of the scalp, which defines 21 electrode positions. Each electrode has: a main letter to identify the lobe, a number to identify the hemisphere, and some electrodes has subscripted letter used for accurate localization (Figure 2.1 and Table 2.1) [31].

Four anatomical landmarks are used to position the electrodes. The first landmark is the nasion which is the point between the forehead and the nose. The second landmark is the inion which is the lowest point of the skull from the back of the head, and lastly the pre-auricular points anterior to the ear (Figure 2.3). The exact electrode placement

procedure is discussed extensively in [31] and [32]. The electrodes positions in x, y, z, azimuth and elevation/altitude are summarized in Table 2.2 [33].

Table 2.1: Electrodes letters/labels according to the brain lobes and hemispheres

Main Letter	Main Location	Number	Hemisphere	Subscripted Letter	Specific Location
F	Frontal Lobe	**Even**	Right	**Z**	Midline
T	Temporal Lobe	**Odd**	Left		
C	Scalp Centre				
P	Parietal Lobe				
O	Occipital Lobe				

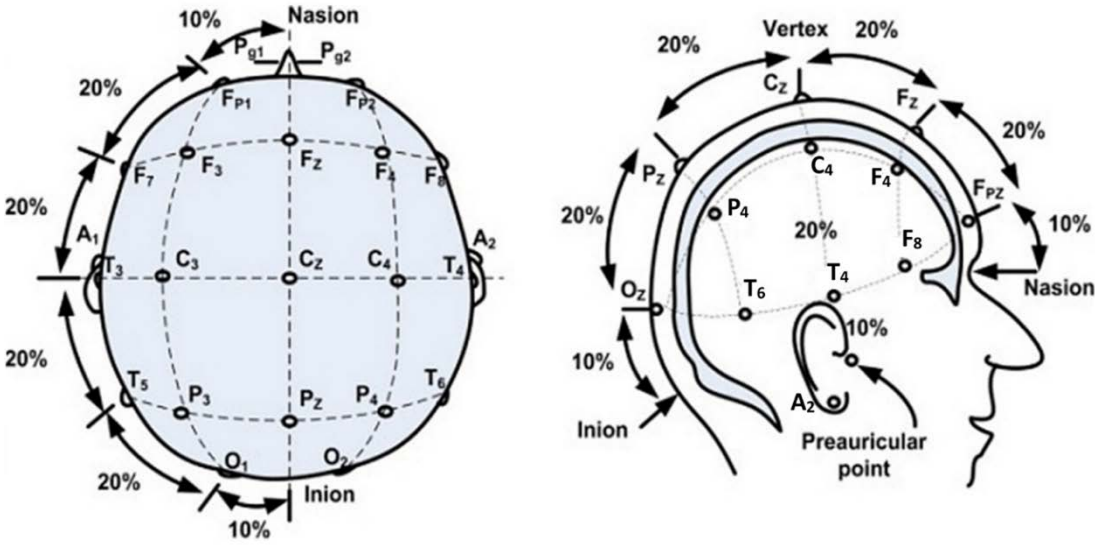

Figure 2.3: The International 10-20 standard for EEG electrodes placement (Redrawn from [34])

Table 2.2: 21 EEG electrodes positions in catesian and spherical coordinates. Cartesian coordinates are caluclated with respect to a head radius of 5.95 cm. Spherical angles measuring references are; azimuth angle is between +X and +Y, and elevation is between the electrode vector and the x-y plane. The computed values can be validated by checking the results in [33], but note that transforming to a common angle reference is needed

Electrode Label	Cartesian Coordinates			Spherical Coordinates (Degrees)	
	X	Y	Z	Azimuth	Elevation
F_z	-3.50	0.00	4.81	180	54
C_z	0.00	0.00	5.95	180	90
P_z	3.50	0.00	4.81	360	54
F_{p1}	-5.38	-1.75	1.84	198	18
F_3	-3.45	-2.73	4.00	218.36	42.3
F_7	-3.33	-4.58	1.84	234	18
C_3	0.00	-3.50	4.81	270	54
T_3	0.00	-5.66	1.84	270	18
P_3	3.45	-2.73	4.00	321.64	42.3
T_5	3.33	-4.58	1.84	306	18
O_1	5.38	-1.75	1.84	342	18
F_{p2}	-5.38	1.75	1.84	162	18
F_4	-3.45	2.73	4.00	141.64	42.3
F_8	-3.33	4.58	1.84	126	18
C_4	0.00	3.50	4.81	90	54
T_4	0.00	5.66	1.84	90	18
P_4	3.45	2.73	4.00	38.36	42.3
T_6	3.33	4.58	1.84	54	18
O_2	5.38	1.75	1.84	18	18
F_{pz}	-5.66	0.00	1.84	180	18
O_z	5.66	0.00	1.84	360	18

2.2.2 Normal and Abnormal Neonatal EEG

The most common way of analysing EEG is by visual inspection in the time domain and by estimating the signal spectral contents. Normally, EEG is visualised in time windows ranging from 10 to 15 seconds and showing all channels simultaneously. The frequency distribution is important when characterising the EEG pattern. Different oscillations are consequences of different underlying mechanisms. The EEG is commonly divided into four sub-bands as the following (Figure 2.4):

- Delta: $0.5 \leq f < 4$ Hz (Occur in infants and during deep sleep or anaesthesia).
- Theta: $4 \leq f < 8$ Hz (High in children below 13 years old during drowsiness or light sleep).
- Alpha: $8 \leq f < 13$ Hz (Occur in walking and resting state. High when eyes are closed. Sinusoidal in shape).
- Beta: $13 \leq f < 30$ Hz (The lower frequencies are present during mental activity, and higher frequencies are associated with tension and intense).

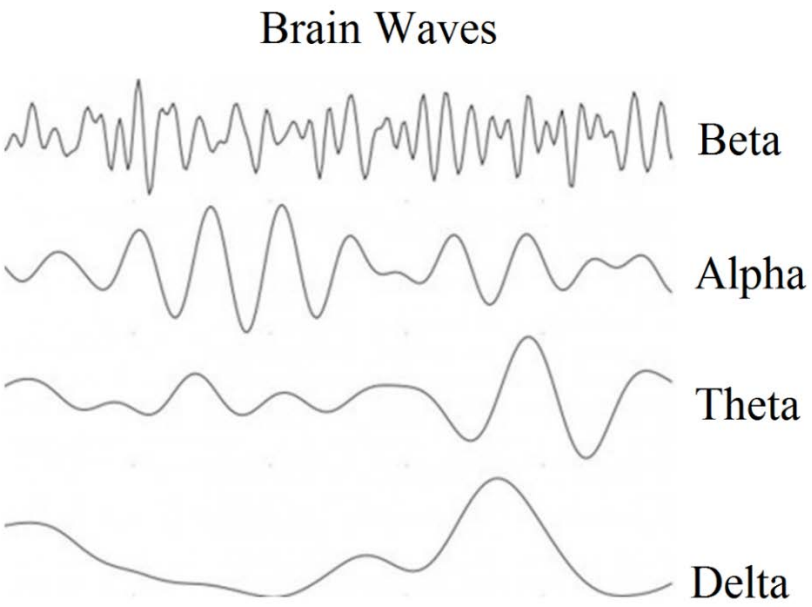

Figure 2.4: Example of Different brain waves (Redrawn from www.hubpages.com)

Classification of neonatal EEG into normal and abnormal is a difficult task as neurologists need to be aware of the exact age of the neonate. Vast maturational changes happen in the 25 to 48 week period of Conception Age (*CA*). Normal patterns at one age can be abnormal only a few weeks later and vice versa [35].

In extreme prematurity, normal EEG patters are typically discontinues, with bursts of continuous cerebral activity separated by intervals of relative quiescence and lower amplitude. This discontinuity improves with ages and the inter-burst interval becomes much shorter and higher in amplitude as the baby approaches full term. Between 30 to 32 weeks CA, Quite Sleep (*QS*) activity consists of a *tracé discontinue* patterns, where periods of cerebral activity are separated by isoelectric periods of quiescence with voltages less than 25μV. By 35 to 36 weeks CA, QS changes behaviour into *tracé alternant* pattern, where cerebral activities are maintained above 25μV, but with cycles between higher-amplitude bursts and more quiescent periods. Finally, by 40 to 44 weeks CA, the EEG background becomes continuous in both wake and sleep cycles [35].

EEG abnormalities can be detected by examining the background EEG voltage and frequency. Examples of abnormalities in the background are: discontinuous pattern (Figure 2.5), burst suppression pattern (Figure 2.6), continuous low voltage EEG (less than 10 μV), and isoelectric EEG (less than 2 μV) [35].

Seizures are one of the most important EEG abnormality that needs to be detected. They can be grouped into the following categories: clonic, tonic, and myoclonic. Due to the incomplete myelination, infants cannot generate generalised tonic-clonic seizures♣, but

♣ Generalized tonic-clonic seizure is one type of seizure that involves the entire body [35].

they can have multifocal seizures that can appear as generalised to the inexperienced examiner. They can also have generalised epileptic spams that are hypothesised to be more subcortically driven [35]. Generally, seizures are characterised by sudden changes in frequency in the measured EEG (limited by the alpha band < 13 Hz [13] [35] [36]). These changes are imposed on the background EEG as repetitive spikes or oscillatory type of signals (Figure 2.7). Moreover, the morphology of epileptic seizures slightly change from one type to another as they may appear in different frequency ranges.

Figure 2.5: Tracé discontinue in a neonate with moderately severe HIE (Redrawn from [37]). Note that discontinuous EEG in neonates with HIE is equivalent to the tracé discontinue in premature.

Figure 2.6: Burst suppression pattern in severe birth asphyxia (Redrawn from [37])

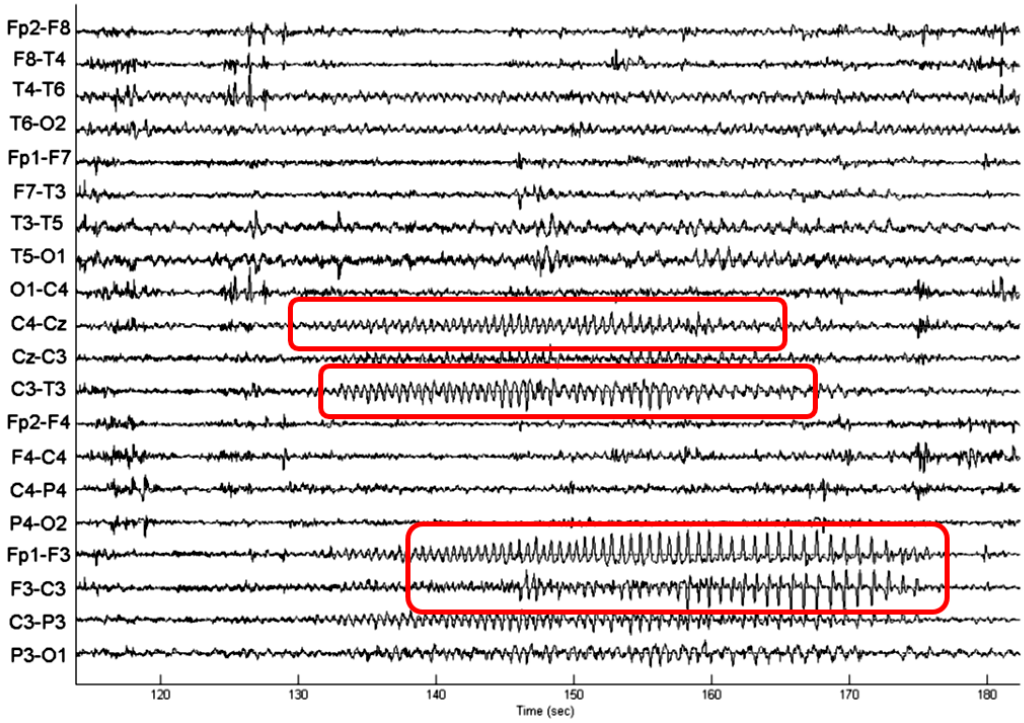

Figure 2.7: Example of a typical neonatal seizure (Redrawn from [37])

2.2.3 Neonatal Single-Channel EEG Modelling

Several attempts and techniques to model single-channel neonatal EEG has been found in the literature. The first attempt (to the best of the author's knowledge) at modelling the neonatal EEG for seizure detection was performed by Roessgen *et al*. in [38] "*The Roessgen Model*". The model was constructed by modifying the adult EEG model proposed by Lopes da Silva et al. in [39]. The modification was done by adding a stationary sawtooth input to Lopes da Silva's model to simulate neonatal seizures. The produced synthetic waveforms had coloured spectrum but lacked the broad nonstationary behaviour that is known to be present in neonatal EEG.

Another attempt to model the neonatal EEG was performed by Celka and Colditz [40] "*The Celka Model*". Their model consisted of two different wiener filters to simulate seizure and background patterns. The model included a nonstationary sawtooth input

with a piecewise linear instantaneous frequency (*IF*) law to comprise the nonstationary behaviour of the neonatal EEG seizure patterns. This model offered two major improvements in modelling the neonatal EEG; the inclusion of the nonlinear mapping function, and the inclusion of a nonstationary input based on the analysis of the seizure waveform. However, the model was not capable of simulating the various morphologies of the seizure waveform due to a lack of variability in the mapping function. Furthermore, this model could not be implemented as the range and distribution of possible inputs and model parameters used are still not known.

Rankine *et al.* had presented a different approach to the neonatal EEG modelling problem in [41] "*The Rankine Model*". It is based on the fact that the seizure waveform can be modelled as a multicomponent signal with a piecewise linear IF law [42], and that background EEG can be modelled as a random signal with a time varying spectrum [43]. The statistical properties of the model parameters (ranges and the distributions) were al known. Consequently the model was capable of simulating the neonatal EEG patterns. However, due the large number of parameters; the model failed to account for dependence between them, which meant that many seizure patterns generated by the model were not mimicking real neonatal EEG seizure.

An improvement to the background EEG waveform produced in the Rankine model was proposed in [44] "*The FBM Model*". This model was based on using a band limited fractional Brownian process with time-varying Hurst exponent to model the neonatal background EEG. This approach had provided a more detailed definition of the neonatal background EEG when compared to the previous models. The advantages of this approach were; simple simulation of continues neonatal background EEG with variable

spectral characteristics, and that development of features for analysing the neonatal background EEG is inherited in the model.

The latest attempt (to the best of the author's knowledge) in neonatal EEG modelling was proposed by *Stevenson et al.* in [13]. The model was based on a nonlinear dynamic system; the Duffing oscillator *"The Duffing Model"*. The Duffing oscillator was driven by a nonstationary impulse train to simulate neonatal seizure EEG and by white Gaussian noise to simulate neonatal background EEG. The advantages of this model were; reduction in the required number of parameters, and more accurate/realistic, life–like EEG when compared with the previous models.

2.2.4 Neonatal Multichannel EEG Modelling

Modelling the neonatal multichannel EEG is done by adapting the Radiation Transport Equation (*RTE*) and by considering a mean seizure propagation speed. The RTE describes the dispersion and decay in light intensity due absorption, scattering, and reflection when propagating through biological tissues. Despite the fact that, RTE uses light and EEG is an electrical measurement; both signals are fundamentally electromagnetic, thus they obey the same fundamental laws of physics (absorption, scattering, reflection, and the inverse square law [45]). Furthermore, applying the RTE on a validated neonatal head model, would represent the EEG propagation through the different head structures, thus EEG power will decay and disperse according to the head tissue's optical properties.

The RTE is used extensively in Optical Imaging, Near Infrared Spectroscopy (*NIRS*), Diffuse Optical Tomography (*DOT*), Single Photon Emission Computed Tomography (*SPECT*), and to study light absorption and scattering in different

biological tissues. *Jacques* had used the RTE in [46] to review reported tissue optical properties. The author concluded that using a generic tissue can adequately mimic any real tissue, and it has the advantage of generating smoothly predictable spectra for absorption and scattering.

Tuchin had also adapted the RTE in [47] to discuss both (strongly) multiple scattering tissues, such as: skin, brain tissues, and vessel walls, and (weakly) scattering high-transparent tissues, such as eye tissues (cornea and lens).

Dehaes et al. had investigated the effect of the neonatal fontanel by predicting the photon propagation using a probabilistic Monte Carlo approach in [48] to improve NIRS measurements. They had created two anatomical neonatal head models from computed tomography and magnetic resonance images. The first model is a realistic head including the fontanel tissue, while the second model replaces the fontanel by skull tissue. Their results suggested that the fontanel should be taken into account in quantification of NIRS responses to avoid errors.

Jager et al. had investigated the performance of a neural network in the derivation of the absorption coefficient of the brain in [49]. This was done by using simulated non-invasive time-resolved reflectance measurements on the head. They had determined the absorption coefficient of the brain with an RMS error of less than 6% from reflectance data at a single distance calculated by diffusion theory.

Another adaptation of the RTE was done by *Brigadoi et al.* in [25]. They had used absorption coefficients, reduced scattering coefficients, and refractive indices from *Dehaes et al.* to model the photon transport using Mote Carlo approach, Finite Element Method, and a Voxel-based Monte Carlo method.

The neonatal seizure propagation speed has no record in the literature (to the best of the author's knowledge). However adult seizure propagation speed ranges are studied, but it is still poorly understood why it varies over several orders of magnitude. Several attempts and techniques had been done to understand this behaviour. *Trevelyan et al.* had proposed in [50] that the speed of propagation is set by the extent of the recruitment steps, which in turn is set by how successfully the feedforward inhibitory restraint contains the excitatory drive. Thus, a single mechanism could account for the wide range of propagation velocities of epileptiform events observed in vitro and in vivo (0.1 - 100 mm/s). *Trevelyan et al.* had also suggested in [51] that the after-discharges propagate with relative uniform speed and are independent of the speed of the ictal waveform (20 - 100 mm/s). Another attempt was done by *Asakawa et al.* in [52]. They had evaluated the effect of different anxiety states on information processing as measured by an EEG using emotional stimuli on a smartphone. They showed how anxiety level can change the EEG propagation speed.

2.2.5 Artifacts in Neonatal EEG

Artifacts that corrupt neonatal EEG measurement can be categorised into three distinct sources: environmental artifacts, experimental errors, and systemic physiological artifacts.

2.2.5.1 Environmental Artifacts

Environmental artifacts generally originate from the mains power leads that surround the body. It can be seen in the form of 50/60 Hz sinusoid and can additionally arise from electromagnetic interference. Instrumentation artifact is a type of environmental artifact which originates from within the circuit components and can be observed in the

form of; thermal noise, shot noise, or 1/*f* noise. These types of artifacts can easily be suppressed by means of linear filters. This is due to either the narrow frequency band of the artifact such as the 50/60 Hz power signal or because the different frequency band these artifacts occupy when compared to the neonatal EEG. Although white noise covers all frequencies, filtering the undesired high-frequency components removes the majority of the noise, thereby improving the signal-to-noise ratio (*SNR*) [9].

2.2.5.2 Experimental Errors

Experimental errors are seen as uncontrolled/unwanted variations in the experimental setup. They can be reduced by proper planning, but it is almost impossible to eliminate completely. This inability to completely supress such errors is in most part due to; human errors during the experimental setup, or subject motion during data acquisition. Examples of such errors are: incorrect procedural setup, and poor application or placement of the EEG electrodes which can create discrepancies in the measured signal. Neonate's motion can create large amount of experimental errors and it is commonly the most detrimental for EEG recordings. Neonate's motion alters the position of the electrodes on the scalp and can cause differences in the distance between recording electrodes and the scalp, thus a change in the electrical coupling causing signal distortion. Moreover, changes in the conduction volume between the electrode and the EEG signal source due neonate's motion induces potential changes at the recording sites. Movement artifacts can be reduced by correct preparation of the interface surface by abrasion and cleansing of the skin prior to the adhesion of the electrodes [9].

2.2.5.3 Physiological Artifacts

Physiological processes in the body other than EEG can contribute to the measured signal. The major contributors of such artifacts are: eye movement related artifacts,

Electrocardiogram signals (*ECG*), and muscle tension signals measured using Electromyography (*EMG*).

The eye movement related artifacts have the highest supressing effect on EEG. Alternations in the eye position (due rapid eye movement sleep *REMS*, blinking, or eye lid movements) change the resting potential of the retina. These electrical changes can be measured using an Electrooculogram (*EOG*) and can propagate over the scalp to be imposed on the EEG recordings. This type of artifact can seriously degrade the utility of EEG analysis, as its amplitude can be greater than the EEG.

The cardiac signals can also cause artifacts on the EEG readings. The signal originating from the heart has a relative high amplitude that can be picked up at different locations on the patient's body. Furthermore, the expansion and contraction of the blood vessels caused by the heart beatings may introduce pulsation artifacts on the EEG recordings. These artifacts are characterised by frequency ranges of around 0.5–40 Hz and 1.2 Hz, respectively, making them difficult to remove from EEG waveforms, as they are overlaid on the EEG spectrum [9].

2.2.6 Neonatal EEG Artifacts Modelling

Several attempts and techniques had been done to understand the characteristics of the neonatal EEG artifacts. *Uriguen et al.* had reviewed the background knowledge on the characteristics of the EEG activities along with the EEG artifacts in [10]. They had categorised EEG artifacts into physiological and non-physiological waveforms, and then discussed only physiological artifacts (ocular, muscle, cardiac, perspiration, and tongue artifacts) as the other type can be reduced by proper measurement techniques.

Another attempt in reviewing the characteristics of EEG artifacts was done by *Sweeney et al.* in [9]. In this publication EEG artifacts had been categorised into: environmental artifacts, experimental errors, and physiological artifacts. They had extensively discussed the physiological signals that most likely to be recorded in a home environment, and documented the artifacts which occurred most frequently that had the largest degrading effect on the EEG signals.

Vos et al. had described and evaluated algorithms using Independent Component Analysis (*ICA*) for the automatic removal of ECG, pulsation and respiration artifacts in neonatal EEG in [53]. They concluded that these artifacts had similar morphologies as seizures, and were characterised by a high degree of repetitiveness. Because of this, these artifacts were the main cause of false positives in seizure detection algorithms.

Janardhan et al. had discussed ECG artifact's influence on the analysis and interpretation of the EEG signals in [54]. They had modelled the cardiac artifact by combining two signals to approximate the main characteristics of the artifact. The cardiac artifact consisted of a 1 Hz spike train signal simulating corrupted QRS complexes of the ECG, and a 2 Hz sine wave that corresponded to the pulsation artifact of an electrode close to a blood vessel.

Matic et al. had also discussed the ECG artifact's influence on the analysis and interpretation of the EEG signals in [55]. They had modelled the cardiac artifact by combining two signals to approximate the main characteristics of the artifact. The cardiac artifact consisted of a 2.5 Hz spike train signal simulating corrupted QRS complexes of the ECG having an amplitude of 6, and a 2 Hz sine wave that corresponded to the pulsation artifact of an electrode close to a blood vessel having an amplitude of 1.5.

Khalifa et al. had discussed the contamination of EEG data by different sources external to the brain in [56]. They had found out that pulse artifacts having a frequency near 1.2 Hz occur when an electrode is placed in the vicinity of a blood vessel. This artifact can appear as a sharp spike or as a smooth wave. They had also found out that an electrode artifact is generally characterized by behaviour like short time, high amplitude (*STHA*) jump due to impedance alteration, and confinement to a single electrode. Electrode artifacts resembled inter-ictal spikes and at times could mimic an ictal pattern.

Zima et al. had discussed STHA artifacts on long-term neonatal EEG recordings, and their robust removal in [57]. They described the STHA artifact to be mainly caused by movement activity, and had an adverse effect on automatic processing of long-term sleep recordings.

Brotchie et al. had also discussed the STHA EEG artifact characteristics in [58]. They had constructed a model that simulates STHA EEG artifacts using heavy-tailed noise. This type of noise was generated using a Levy stable symmetrical stochastic process and had a maximum amplitude 15 times greater than the normalised EEG.

2.3 Processing of Neonatal EEG Signals

Neonatal EEG is a nonstationary signal where its spectral contents change with time. Classical processing domains such as time and frequency cannot take into consideration the time varying spectrum, thus they cannot represent the EEG information adequately. Consequently, a Time-Frequency Representation (*TFR*) is used to represent the time varying spectral information of nonstationary signals. It represents both time (t) and frequency (f) in a joint distribution called the Time-Frequency Distribution (*TFD*) and denoted by $\rho(t, f)$. More details can be found in [1] chapter 1

pages 31-64. Time-frequency signal processing is studied and discussed extensively in Appendix A page 140, along with a novel approximation for the compact support kernel time-frequency formulation.

Several recent publications used time-frequency processing techniques in different applications. For instance, *Castaño-Candamil et al.* had introduced a novel method for source analysis of EEG recordings based on a physiologically-motivated source representation in [59]. They had used Time-Frequency Mixed-Norm Estimates along with Sparse Basis Field Expansions to extract EEG sources.

Another adaptation for time-frequency processing techniques was done by *Siddiqui et al.* in [60]. They had used time frequency processing to analyse the power spectrum density of EEG signals to diagnosis Rapid Eye Movement Behaviour Disorder (*REMBD*). Their results suggested that normal people and REMBD patients can be differentiated using this methodology.

Yan et al. had employed the Stockwell transform to obtain the time–frequency representation of the EEG signals, and then calculated the power spectral density in the time-frequency plane to characterize the behaviour of EEG recordings in [61]. This was done to propose a novel seizure detection method based on Stockwell transform for intracranial long-term EEG data. *Samiee et al.* had used an approach based on an adaptive and localized time-frequency representation of EEG signals by means of rational functions to classify epileptic seizures in [62]. They also implemented a novel feature extraction technique based on rational discrete short-time Fourier transform (*DSTFT*) that outperformed others in terms of classification accuracy.

Lastly, *Boashash et al.* had used time-frequency processing techniques mostly for neonatal seizure detection applications as in [27], [28], [63] and [64]. In [27] they had

proposed a methodology to define new time-frequency features to detect changes in nonstationary signals and applied it on a neonatal seizure detection system. In [28] they had designed a novel time-frequency matched filter to solve the problem of detecting non-stationary signals in the presence of additive noise, and then applied the design to detect neonatal seizures using multichannel EEG signals. In [63] they had reviewed the recent advances in the field of time-frequency signal processing with focus on introducing image feature information using pattern recognition techniques. Finally, in [64] they presented a methodical approach to improve quadratic time-frequency distribution (*QTFD*) methods by designing adaptive time-frequency kernels for diagnosis applications. They had illustrated their approach by using EEG, heart rate variability (*HRV*), and pathological speech signals.

2.4 Blind Source Separation

2.4.1 The Cocktail Party Problem

Blind Source Separation (*BSS*) refers to the methods that allow the separation of a set of source signals from a set of mixed signals without information about the nature of the source signals. This problem is called the cocktail party problem.

The neonatal EEG represents a linear mixture of numerous sources of neural and non-neural activities (EEG and artifacts). Because EEG is acquired with a number of electrodes on different locations on the scalp, each electrode will measure a weighted sum of these different sources namely the electrode observation. This mixture of sources can be expressed using Equation 2.1, where; Y is the matrix of acquired signals on the ith electrode (Equation 2.2), S is the matrix of source signals (Equation 2.3), and

M is the mixing matrix where its coefficients determine the contribution of each source signal in the observation (Equation 2.4).

Equation 2.1 describes the standard linear statistical model in the time domain (Two Dimensional Signals [Channel, Time]). This equation can be extended to include three dimensional signals such as images, or in this particular case the auto-TFDs of the STFD matrix. Equation 2.5 describes the time-frequency extension of the standard linear statistical model, where; Q is the matrix of the mixed auto-TFDs on the ith electrode, and T is the vectorisation of the auto-TFDs from the STFD matrix. Vectorisation converts the auto-TFDs matrices into vectors having n number of rows and (time samples x frequency samples) number of columns. Such transformation is valid as the fundamental property of the linear mixing models states that transformation of the mixed data does not in any way change the mixing model, and it is essential as time-domain BSS techniques can be used in the time-frequency domain [65].

The EEG and artifacts mixture can be separated by estimating an un-mixing matrix $\widehat{W}$, where "ideally" it is the inverse of the mixing matrix M. Equation 2.6 and Equation 2.7 describe the source separation in the time and the time-frequency domain respectively. The objective of BSS is to find or estimate an un-mixing matrix that separates the observed mixture into source signals and Independent Component Analysis (*ICA*) is a widely used tool to solve this problem (Figure 2.8).

$$Y = M \,.\, S \qquad \text{Equation 2.1}$$

$$Y = [y_1(t),\ y_2(t), \dots, y_n(t)]^T \qquad \text{Equation 2.2}$$

$$S = [s_1(t),\ s_2(t), \dots, s_n(t)]^T \qquad \text{Equation 2.3}$$

$$M = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{nn} \end{bmatrix}$$ Equation 2.4

$$Q = M \,.\, T = M \,.\, vec\left(diag(\boldsymbol{\rho}(t,f))\right)$$ Equation 2.5

$$\hat{S} = \widehat{W} \,.Y = \widehat{M}^{-1} \,.Y$$ Equation 2.6

$$\hat{T} = \widehat{W} \,.Q = \widehat{M}^{-1} \,.Q$$ Equation 2.7

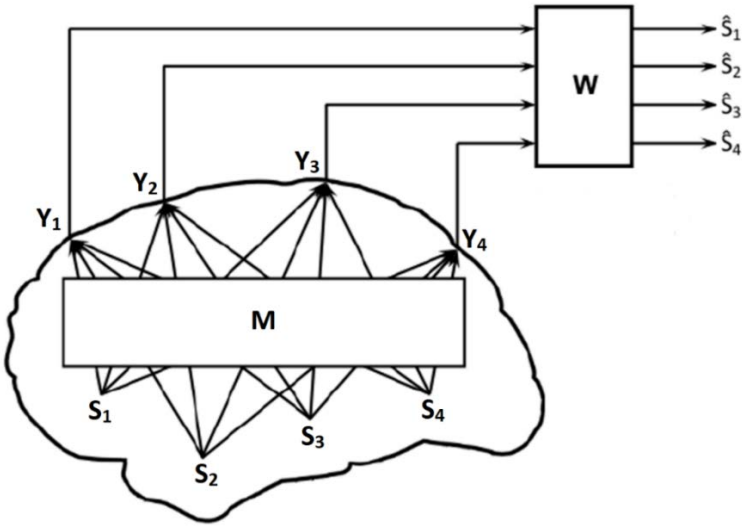

Figure 2.8: Linear mixing and de-mixing models (Redrawn from [10])

2.4.2 Independent Component Analysis

Since both M and S are unknown, it is generally impossible to determine them without imposing additional constraints. Therefore, ICA imposes several assumptions about the sources in order to obtain a unique decomposition which are:

- M is an unknown invertible mixing matrix.
- The source signals are mutually independent random variables.
- The source signals are non-Gaussian signals (This assumption can be relaxed as only one source signal can be Gaussian at a time).

- The number of source signals are always less than or equal the number of electrodes/observations.

Furthermore, several choices need to be made as several BSS algorithms exist and it is necessary to decide how many components should be retrieved. More details on BSS algorithms can be found in Appendix B, page 164.

Several attempts had been done to analyse the different EEG artifact detection and removal techniques in the time domain. *Uriguen et al.* had reviewed artifact removal algorithms used to extract the main sources of interference encountered in the EEG, specifically ocular, muscular and cardiac artifacts in [10]. With that review they had concluded that, without prior knowledge of the recorded EEG signal or the contaminants, the safest approach is to correct the measured EEG using ICA, precisely an algorithm based on second order statistics such as second order blind identification (*SOBI*). They had also concluded that correlation measures and normalised root mean square error (*NRMSE*) are the most prevalent metrics to verify the effectiveness of a noise/artifact removal methods.

Another attempts in reviewing the artifact removal algorithms were done by *Sweeney et al.* and *Adali et al.* in [9] and [66] respectively. *Sweeney et al.* provided a detailed analysis of the current artifact removal techniques and evaluated the advantages and disadvantages of each of the proposed artifact detection and removal techniques. *Adali et al.* had presented ICA and then its generalization to multiple data sets Independent Vector Analysis (*IVA*). They had addressed how various methods fall under this umbrella and had given examples of performance for a few sample algorithms in medical image analysis.

Sarfraz et al. had proposed the application of ICA to ECG signal pre-processing and then compared the performances of two major types of ICAs namely Infomax and Fast-ICA in ECG signal de-noising in [67]. They concluded that Fast ICA outperformed Infomax algorithm and can effectively improve the ECG recognition in the presence of non-trivial artifacts.

Vos et al. had described and evaluated artifact removal algorithms using ICA for the automatic removal of ECG, pulsation and respiration artifacts in neonatal EEG in [53]. They concluded that these techniques reduced the number of false positive detections and are beneficial in long term EEG seizure monitoring in the presence of disturbing biological artifacts.

Janardhan et al. had quantified which artifact separation algorithm would be most effective when used for contaminated EEG in [54]. They had evaluated the influence of noise and artifacts on the performance of Fixed Point ICA (*FpICA*), SOBI, Joint Approximation Diagonalization of Eigen matrices (*JADE*), and Algorithm for Multiple Unknown Source Extraction (*AMUSE*) algorithms. They had used spearman correlation as detection/comparison criterion, and concluded that SOBI is the best for sinusoidal artifacts removal, while FpICA is the best for spiky artifacts removal.

Another comparison of common ICA methods was done by *Matic et al.* and *Miljkovic et al.* in [55] and [68] respectively. They compared the performance of Robust-ICA, SOBI, JADE, and Blind Source Separation-Canonical Correlation Analysis (*BSS-CCA*) algorithms when applied to real-life neonatal EEG data with added artifacts. They both had used spearman correlation as detection/comparison criterion, and concluded that SOBI and BSS-CCA are the best algorithms for sinusoidal artifacts removal, while Robust ICA was the best for spiky artifacts removal.

Vanderperren et al. had also used ICA algorithms to remove Ballistocardiogram (*BCG*) artifacts from EEG recordings inside MR scanner in [69]. They had used JADE, SOBI, Fast-ICA, and Infomax ICA algorithms to remove artifacts and adapted the spearman correlation as the artifact component selection criteria.

On the other hand, many attempts had been done to analyse the different EEG artifact removal techniques in the time-frequency domain. *Sardouie et al.* had used the BSS-CCA method to extract the time-frequency signature of ictal sources in [11]. Moreover, they proposed two time-frequency based semi-BSS approaches, namely the Time-Frequency-Generalized Eigen Value Decomposition (*TF-GEVD*) and the Time-Frequency-Denoising Source Separation (*TF-DSS*), for the denoising of ictal signals based on the extracted time-frequency signatures.

Zhao et al. had proposed a hybrid de-noising method combining Discrete Wavelet Transformation (*DWT*) and an Adaptive Predictor Filter (*APF*) to remove Ocular Artifacts (*OA*) from the EEG signals in [70]. Their results outperformed existing techniques such as; Wavelet Packet Transform (*WPT*), ICA, DWT and Adaptive Noise Cancellation (*ANC*). *Muthukumaraswamy* had reviewed the spectral, spatial, and temporal characteristics of muscle artifacts along with the techniques to suppress such artifacts in EEG in [71]. The author had concluded that none of the artifact detection and removal techniques was able to guarantee that the analysed data are artifacts free.

Daly et al. had proposed to place an accelerometer on cerebral palsy patient's head along with ICA to separate head movement's artifacts from EEG signals in [72]. Their results illustrated that the approach significantly reduces the influence of head movement artifacts for Brain-Computer Interface (*BCI*) applications.

Hamaneh et al. had proposed an automated algorithm for removal of EKG artifact from EEG data using ICA and Continuous Wavelet Transformation (*CWT*) in [73]. The method applies CWT to the ICs to detect if an EKG artifact component exist using different wavelet based features. *Hyvarinen et al.* had proposed to apply ICA on Short-Time Fourier Transforms (*STFT*) of EEG/MEG signals in [65] to find more "interesting" sources than with time-domain ICA, and to sort the obtained components more meaningfully. They had used Fast-ICA, SOBI, and JADE algorithms and applied them on a concatenated short-time Fourier transform of the EEG signal.

Zachariah et al. had designed a system for removal of artifacts from multichannel EEG signals based on ICA and using signal rhythmic components in [74]. Wavelet decomposition was used as a pre-processing step because different artifacts were overlapped with different EEG rhythms. Using wavelet decomposition had increased the redundancy and the rejection of the suitable wavelet artifact related components and decreased the probability of rejecting useful information from the EEG signals.

He-ping et al. had proposed a BSS algorithm by combining the characteristic of time-frequency analysis (*TFA*) and BSS in [15]. Their simulations showed that this algorithm not only suppresses cross term interference but also enhanced time-frequency resolutions. It also provided an effective technology for fault diagnosis of mechanical equipment. *Ramkumar et al.* had developed a data driven method to spatiotemporally and spectrally characterize the dynamics of the brain oscillations for resting state magnetoencephalography (*MEG*) data in [75], and called it envelope Spatial Fourier Independent Component Analysis (*eSFICA*). They had compared their method using simulated data set against 5 other variants of ICA and found that eSFICA performed on

par with its temporal variant eTFICA, and better than other ICA variants in characterizing dynamics at time scales of the order of minutes.

2.4.3 Number of Source Signals Estimation

In general, it is not possible to know how many sources are present in the time domain or in the time-frequency domain EEG. However, this number is known to affect the quality of the BSS solution. When more sources are estimated than there really are, BSS algorithms tend to 'overfit', which can lead to estimation errors [53]. An accepted method to define the number of active sources is based on Principle Component Analysis (*PCA*) decomposition or Singular Value Decomposition (*SVD*) of the multichannel EEG waveform [53]. The number of dominant eigenvalues provides an estimate for the number of active sources in the EEG. More specifically, a percentage of the total sorted eigenvalues can be defined, explaining the variance. The number of required eigenvalues to satisfy the percentage is chosen as the number of active sources in the multichannel EEG waveform. In this way, the number of sources to be extracted is automatically adapted to the complexity of the EEG and the number of EEG channels available.

2.5 Fundamentals of Pattern Recognition

2.5.1 Feature Extraction

Feature extraction is a key stage in pattern recognition, classification of abnormalities, and in automated decision making algorithms such as machine learning. Features concerning the statistical and spectral properties are traditionally used to obtain specific details about the underlying signal. Such statistical features include:

mean, variance, skewness and kurtosis, coefficient of variation, inter-quartile range, and entropy. On the other hand, spectral features include: spectral flux, spectral centroid, spectral roll-off, and spectral flatness. In the case of EEG, such features would fail to give sufficient relevant discriminatory information as the spectral contents are changing with time. Consequently, time-frequency extension of these features is required to increase the relevant information content of these features.

Several authors have extended neonatal EEG features into the time-frequency domain such as: *Hidalgo-Muñoz et al.* in [29], *Balakrishnan et al.* in [30], *Nagaraj et al.* in [36], *Boashash et al.* in [27], [63], [64], [76], and [77], *Al-Fahoum et al.* in [78], *Bruser et al.* in [79], *Fu et al.* in [80], *Jenke et al.* in [81], and *Kumari et al.* in [82]. From all of these publications, the time-frequency features that illustrated good performance in characterising the abnormalities in neonatal EEG are categorised as: statistical features, frequency extended features, instantaneous frequency features, and sub-bands energy features. More details on the time-frequency features along with their equations and interpretations are discussed in Appendix C, page 173.

2.5.2 Feature Fusion: Sum Basis

Feature fusion, also known as early integration or feature-level combining, is a common framework for the fusion of different feature sets. It combines features derived from the different single-channel EEG into a more global feature set to be used for classification of multichannel EEG, thus, it requires only one classifier.

The Multichannel Feature Fusion Sum Basis (*MFFSB*) scheme is based on summing the extracted features from all single-channels [30]. Equation 2.8 expresses the extracted d features from M channels, while Equation 2.9 expresses the fusion process

to produce the global features set. Recent publications concluded that this methodology outperforms other fusion techniques such as decision fusion [28] [30].

$$\mathbf{J} = \begin{bmatrix} J_{1,1} & J_{1,2} & \cdots & J_{1,d} \\ J_{2,1} & J_{2,2} & \cdots & J_{2,d} \\ \vdots & \vdots & \ddots & \vdots \\ J_{M,1} & J_{M,2} & \cdots & J_{M,d} \end{bmatrix} \quad \text{Equation 2.8}$$

$$\boldsymbol{\Psi}_i = \sum_{i=1}^{M} J_{i,j} \quad \text{Equation 2.9}$$

2.5.3 Feature Selection

"Feature selection is a process of selecting an optimal subset from the original set of extracted features that is both relevant and non-redundant" [30]. A large number of features containing redundant and/or irrelevant features can reduce the classifier performance. Thus, this stage is crucial to maintain the highest possible classification performance with the minimum number of features. Furthermore, feature selection has to be utilised even when using ensemble of classifiers. This is essential because utilising the full feature set (containing irrelevant features) in every used classifier, ensures an ensemble performance that is at least higher than any of the individual classifiers, but their individual performance will not be maximised. Consequently, the overall performance can still be higher when utilising a feature selection algorithm

Existing feature selection algorithms generally belong to the following categories: wrappers, filters, and hybrid. The filter method is independent of any classification algorithm, as it orders (ranks) features using specific statistical metrics. Such methodologies are time efficient, but produce low accuracy. Some popular filter methods are F-score criterion, mutual information and information gain. On the other hand, the wrapper method depends on the specific classification algorithm, thus it

produces high accuracy, but it requires high computational power. Finally, the hybrid method attempts to take advantage of the filter and wrapper techniques by employing their opposing strengths [83].

Several authors have used different feature selection algorithms. *Balakrishnan et al.* had used the filter-wrapper based feature selection approach to select an optimal subset of EEG features from a larger set extracted in [30]. *Bruser et al.* have applied the filter approach using a second-order extension of the first-order utility (*FOU*) method to rank the available features for atrial fibrillation detection in cardiac vibration signals in [79]. *Liu et al.* had proposed a new feature selection method based on Fisher criterion and genetic optimization (called *FIG*) to tackle the CT imaging signs of lung diseases recognition problem in [84]. *Jenke et al.* have selected features to recognise emotions from EEG using; ReliefF, Min-Redundancy-Max-Relevance (*mRMR*), and Effect-Size (*ES*) based algorithms in [81]. Lastly, *Adam et al.* have selected the best features to detect EEG peaks using two algorithms namely; Particle Swarm Optimization (*PSO*) and Random Asynchronous Particle Swarm Optimization (*RAPSO*) in [85].

2.5.3.1 Fisher's Score

The Fisher's score is a filter based feature selection algorithm. It finds a subset of features such that, the distances between data points in different classes are as large as possible, while the distances between data points in the same class are as small as possible. In particular, the input feature vector $\Psi \in \mathbb{R}^{d \times n}$ is reduced to $\mathbf{Z} \in \mathbb{R}^{m \times n}$ by computing the Fisher score as expressed by Equation 2.10, where $\mu_{i,k}$ and $\sigma_{i,k}$ are the mean and standard deviation of k-th class corresponding to the i-th feature. μ_i and σ_i

denote the mean and standard deviation of the whole data set corresponding to the i-th feature. Finally, after computing the Fisher score for each feature, the top m ranked features with large scores are selected. This feature selection method is suboptimal as the score of each feature is computed independently. More in depth details on Fisher's Score can be found in [86], while the Matlab code for the Fisher's score algorithm can be found in [87].

$$F(\Psi_i) = \frac{\sum_{k=1}^{c} n_k \left(\mu_{i,k} - \mu_i\right)^2}{(\sigma_i)^2} \qquad \text{Equation 2.10}$$

$$(\sigma_i)^2 = \sum_{k=1}^{c} n_k \left(\sigma_{i,k}\right)^2 \qquad \text{Equation 2.11}$$

2.5.4 Classification

Classification is the problem of identifying to which set of categories a new observation belongs to on the basis of a training set whose category membership is known. Different classification algorithms exist such as: k-Nearest Neighbour (*k-NN*), Linear Discriminant Analysis (*LDA*), Naive Bayes, Decision Trees, Artificial Neural Networks (*ANN*), and Support Vector Machine (*SVM*) [30] [79] [80] [82] [88].

The classifier detection performance can be evaluated in terms of sensitivity, specificity, and balanced accuracy (more details on detection performance metrics can be found in Appendix D, page 177). Unfortunately, these classifier evaluation metrics are not robust as they depend on the selected training and testing sets, thus producing unreliable results. Cross-validation techniques are used to overcome this problem and to produce robust results independent from the selection of training and testing sets.

2.5.4.1 Support Vector Machine

SVM is a supervised learning method that generates input-output mapping functions from a set of labelled training data. It produces a model, based on the training data to classify new test data. SVM is capable of classifying data separated by linear and non-linear boundaries by mapping the problem to a higher dimensional space using different kernel functions (Figure 2.9). It is also capable of classifying overlapping and non-separable data by assigning a penalty for input data that fall on the wrong side of the hyper planes.

For a two-class problem (i.e. normal and abnormal EEG), consider a given training set $\{x_k, y_k\}_{k=1}^{N}$ with input data $x_k \in \mathbb{R}^n$ and output data $y_k \in \mathbb{R}$ with class labels $y_k \in \{-1, 1\}$, a nonlinear SVM classifier can be expressed using Equation 2.12, where α_k are Lagrange multipliers, and $K(x, x_k)$ is a kernel function. The Radius Basis Function (*RBF*) is widely used as SVM kernel and it can be expressed using Equation 2.13, where σ controls the width of the RFB kernel.

$$y(x) = sign\left[\sum_{k=1}^{N} \alpha_k y_k K(x, x_k) + b\right] \quad \text{Equation 2.12}$$

$$K(x, x_k) = \exp\left(\frac{-\|x - x_k\|^2}{2\sigma^2}\right) \quad \text{Equation 2.13}$$

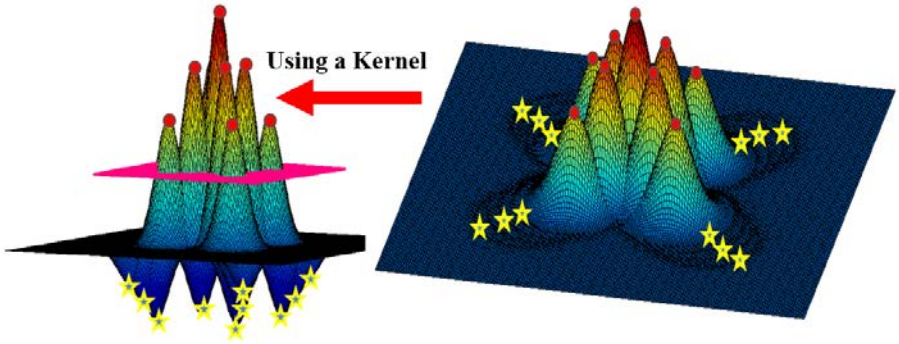

Figure 2.9: SVM kernel used to separate mixed data in a higher dimension (This figure is for illustration purposes)

2.5.4.2 Cross-Validation

Several cross-validation techniques were used in the literature to validate the SVM detection performance, and were categorised into: Exhaustive, and Non-exhaustive cross-validation techniques [80]. Exhaustive cross-validation methods learn and test on all possible combinations to divide the original sample into training and validation sets. On the other hand, Non-exhaustive cross-validation methods do not consider all the combinations but only portions.

Leave-one-out cross-validation (*LOOCV*) is an exhaustive method that provides an unbiased estimate of the generalisation error for the SVM classifier. It takes one instance from the original sample of size N_{seg} as the validation data, and leaves the rest observations $N_{seg} - 1$ as the training data. This process is repeated N_{seg} times until each data sample is considered as validation, thus covering all possible combinations (Figure 2.10) [30] [89] [90] [91].

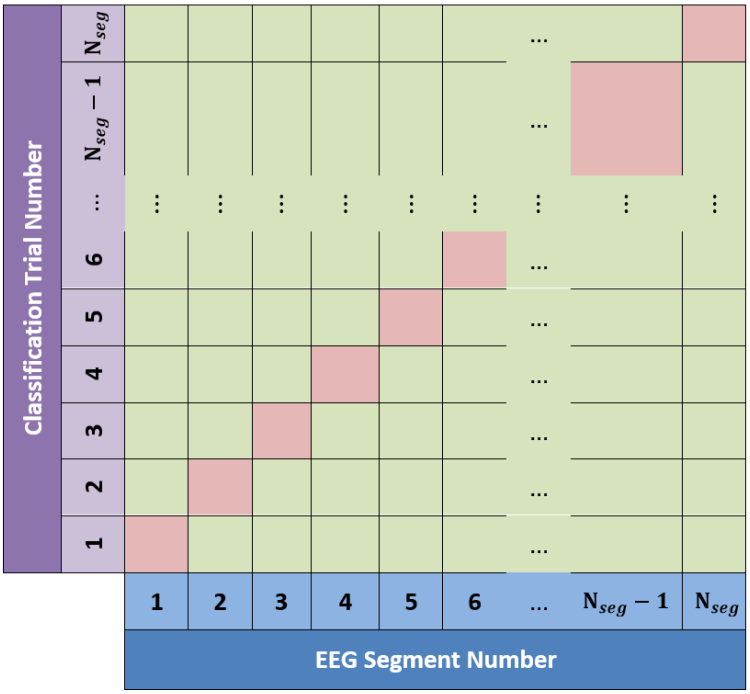

Figure 2.10: Leave-one-out cross-validation process. Training segments are highlighted in green, while validation segments are highlighted in red

CHAPTER 3: MODELLING APPROACHES AND TECHNIQUES

In this chapter, the modelling approaches and techniques used in this thesis are explained in depth along with all the procedures undertaken. First, real neonatal multichannel EEG data are analysed to understand their properties that must exist in a synthetic model. After that, the synthetic neonatal multichannel modelling technique based on synthetic simulators, a head model, and a novel EEG propagation scheme is explained. The utilised synthetic simulators are based on time-frequency modelling of the neonatal EEG background and seizure epochs. Moreover, the multichannel model relationships with previous publications are presented. Furthermore, the EEG propagation scheme produce three different propagation models, thus a model validation procedure is discussed and explained to choose the best model. After that, the synthetic neonatal multichannel EEG is validated against real neonatal multichannel EEG data to quantify its effectiveness. Finally, the synthetic neonatal multichannel EEG artifacts modelling methodology based on simulated signals is presented to contaminate the clean EEG, thus masking its information content.

3.1 Data Analysis

Multichannel continuous EEG recordings were collected from 7 neonatal patients at the NICU of the Royal Brisbane and Women's Hospital, Brisbane, Australia, and were transferred to professor Boashash, lead PI at Qatar University of the National Priorities Research Program (*NPRP*) grant no. 6-885-2-364. Data were recorded using the MEDELEC Profile System (Medelec, Oxford Instruments, UK) that utilises 12 Ag/AgCL electrodes placed according to the 10-20 standard, thus constructing a 20

channel bipolar scheme (more details on the 10-20 standard can be found in Section 2.2.1, page 13, but note that electrode F_{PZ} is not utilised). Acquired data are: filtered using an analog bandpass filter with cut-off frequencies of 0.5Hz and 70Hz, sampled at 256Hz, and then down-sampled to 32Hz. Acquired data contains background and seizure epochs that are marked by a paediatric neurologist from the Royal Children's Hospital, thus providing a method for validation. Figure 3.1 and Figure 3.2 illustrate a single-channel sample of the acquired database depicting background and seizure epochs respectively. Background epochs are characterised by: nonstationary temporal behaviours, decaying power spectrums, and random time-frequency signatures. On the other hand, seizure patterns are characterised by: oscillatory and evolving temporal behaviours, spiky power spectrums, and deterministic time-frequency signatures [1]. These properties will be the basis in constructing the multichannel neonatal EEG model.

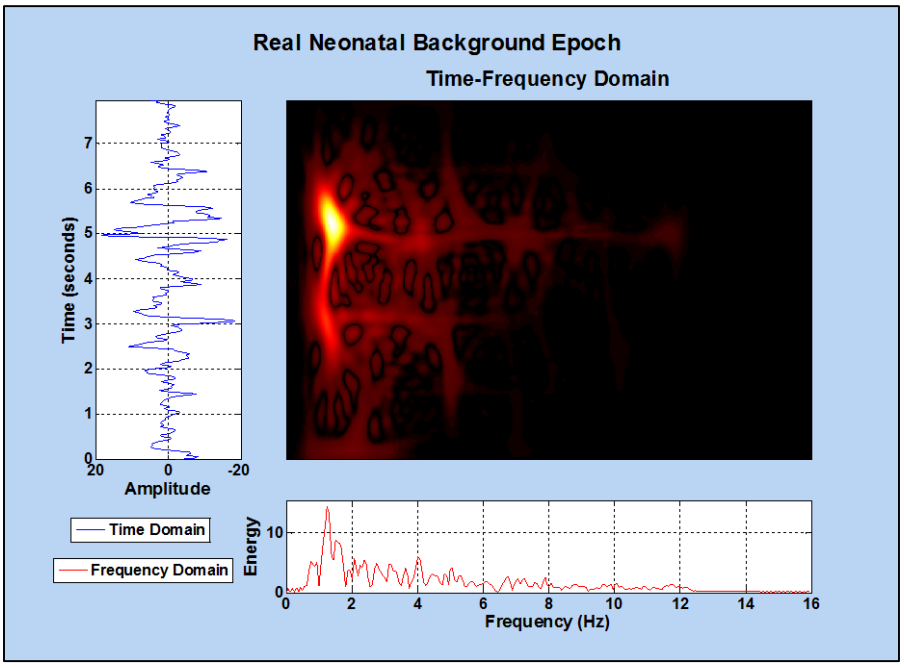

Figure 3.1: Real neonatal background EEG. TFD parameters (EMBD, N = 256, Fs = 32 Hz, FFT_N = 512, lag window = 91, alpha = 0.08, beta = 0.9, time resolution = 1). Time and frequency steps are (0.03125, 0.0626) respectively

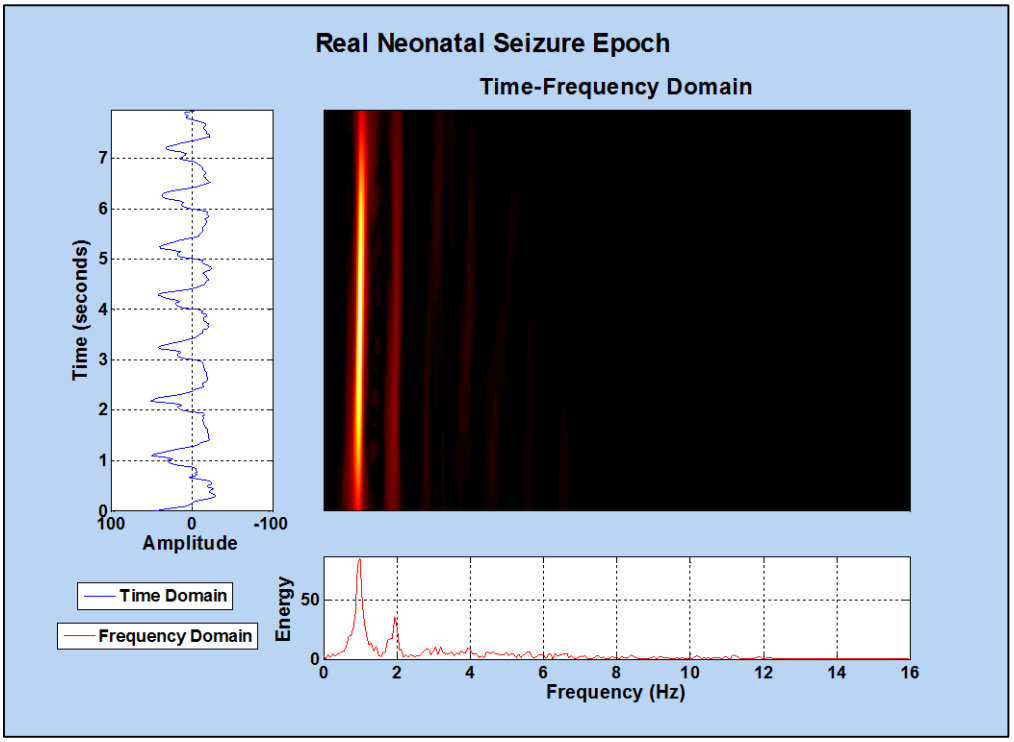

Figure 3.2: Real neonatal seizure EEG. TFD parameters (EMBD, N = 256, Fs = 32 Hz, FFT_N = 512, lag window = 255, alpha = 0.01, beta = 0.9, time resolution = 1). Time and frequency steps are (0.03125, 0.0196) respectively

3.2 Synthetic Neonatal Multichannel EEG Model

3.2.1 Neonatal Head Model

The neonatal head model is constructed by utilising the four sphere approach used by *Sadleir et al.* in [21]. This *Sadleir et al.* approach is adapted as authors supply the dimensions of all spheres. These dimensions were chosen according to data gathered from the inspection of archived neonatal MRI models, which makes them more reliable in mimicking the true neonatal head (more details of the Sadleir four sphere head model can be found in Appendix E, page 179). Furthermore, 21 EEG electrodes are placed on the scalp surface of the four sphere head model according to the international 10-20 system (Section 2.2.1, page 13). The electrodes placement on the four sphere neonatal head model is illustrated in Figure 3.3 by modelling the electrodes as small spheres

(relative to the four sphere head model). It shows; the right and left hemisphere's electrodes in blue and in red respectively, the midline electrodes are in black, and finally the front and back electrodes are in green.

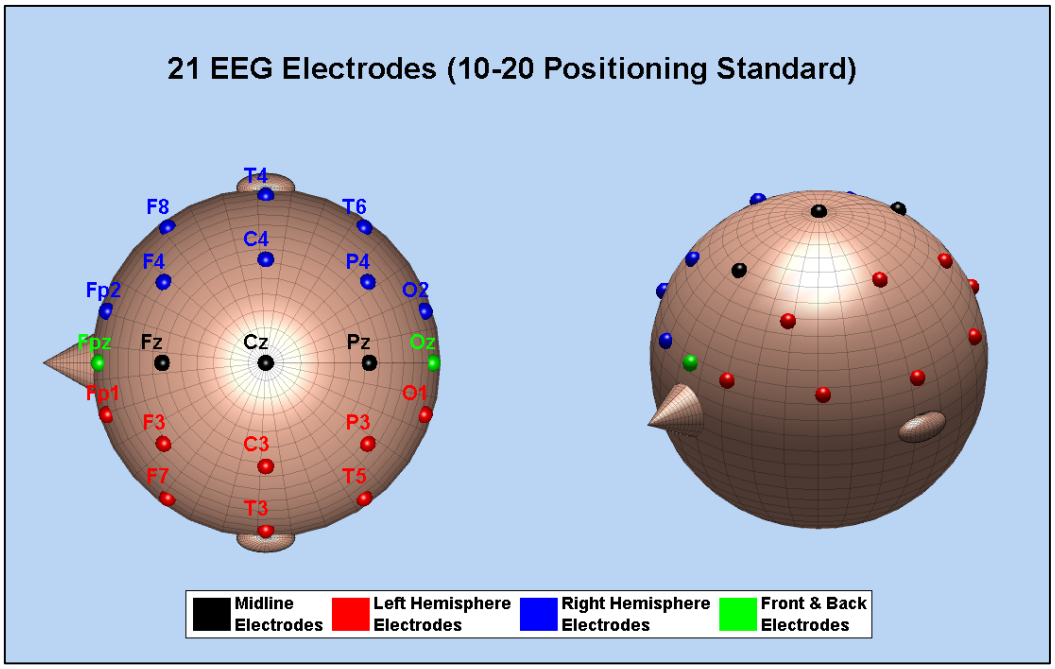

Figure 3.3: 21 EEG electrodes placement on the four sphere neonatal head model. The plot on the left shows the 21 electrodes positions according to the 10-20 standard, while the plot on the right shows the placement of the 21 electrodes on the neonatal head model. The electrodes are modelled by spheres having a radius of 0.3 cm constructed from a 32x32 structure

3.2.2 Neonatal Single-Channel EEG Model

The neonatal single-channel EEG model is constructed using the Rankine model [41] that was developed by Nathan Stevenson and Luke Rankine. This model is the latest that is available as an open source to the public and it can be downloaded from [92] (more details on the Rankine model can be found in Appendix F, page 181). The model consists of two simulators; the first simulator models the neonatal background EEG (Figure 3.4), while the second simulator models the neonatal seizure EEG (Figure 3.5). Both simulators produce good approximations as they were validated by

correlating their results with real neonatal EEG epochs (0.817 for background and 0.901 for seizure in the time-frequency domain).

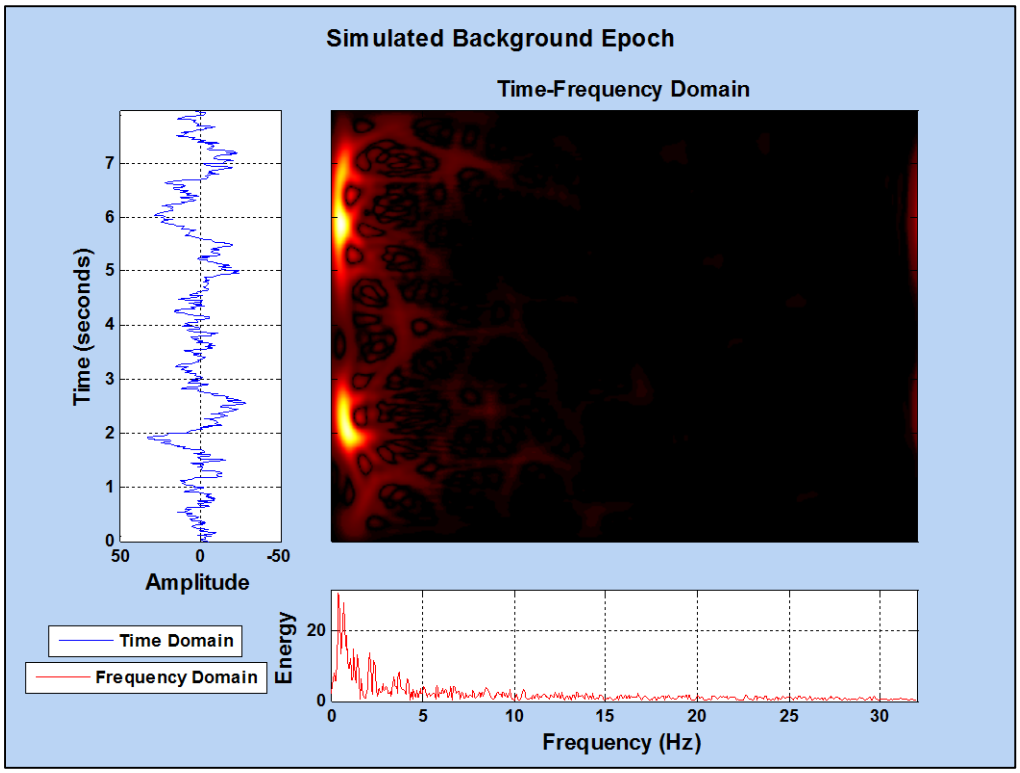

Figure 3.4: Simulated neonatal background EEG. TFD parameters (EMBD, N = 512, Fs = 64 Hz, FFT_N = 1024, lag window = 91, alpha = 0.08, beta = 0.9, time resolution = 1). Time and frequency steps are (0.0156, 0.0626) respectively

The multiple harmonics of the neonatal seizure illustrate the multipath reception of LFM signal that propagated through a nonlinear frequency shifting material (Figure 3.5). The fundamental harmonic has the highest power, while each of the next harmonics has a power less than its previous one (power is decaying as we move across the harmonic number). Furthermore, the fourth harmonic has the highest frequency deviation and distortion, while each of the previous harmonics has a frequency deviation and distortion less than its next one (frequency deviation and distortion is increasing as we move across the harmonic number). The fundamental harmonic (having the highest power and least frequency distortion) represents the seizure event

that had to propagate the least distance to reach the electrode, while others represent scattered versions of the fundamental harmonic; thus the Rankine single-channel EEG model takes into consideration the multipath reception of neonatal EEG.

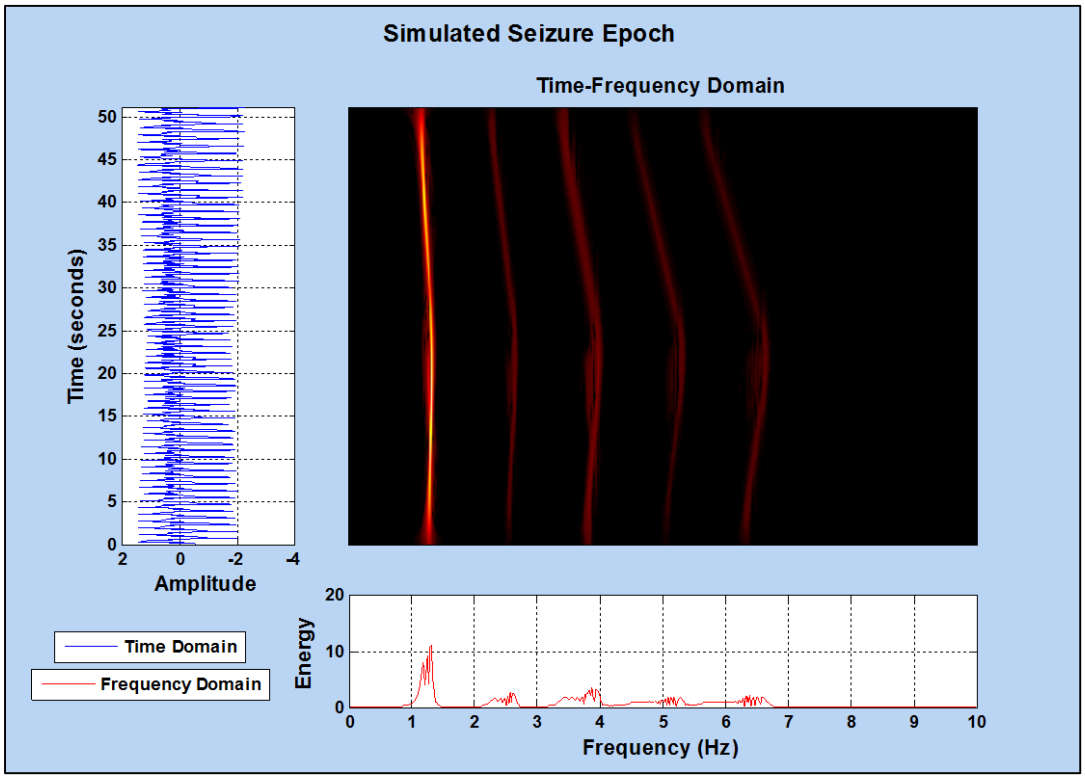

Figure 3.5: Simulated neonatal seizure EEG. TFD parameters (EMBD, N = 1024, Fs = 20 Hz, FFT_N = 1024, lag window = 511, alpha = 0.01, beta = 0.9, time resolution = 1). Time and frequency steps are (0.05, 0.0196) respectively

3.2.3 Neonatal EEG Propagation Model

EEG propagation describes how EEG signals propagate through the different brain tissues that can attenuate, scatter the propagation path, and/or modulate the signal frequency. One problem arise in constructing the propagation model is calculating the electrode potentials for a given source (usually a current dipole). This problem is called the Forward Problem, and it is the reciprocal of the Inverse Problem which finds the dipole parameters that best represent measured potentials at the scalp electrodes.

The Forward Problem had been solved using many different methods such as: combining both EEG and MRI to localize seizures in 3D space as in [19], using Particle Numerical Models as in [22], using Electromagnetic Source Localisation (*ESL*) as in [23], using analytical solutions for the forward problem as in [24], and using Boundary Element Method (*BEM*), Finite-Element Method (*FEM*) and Finite Difference Method (*FDM*) as in [20] [93] [94] [95].

All these methodologies have common pitfalls such as:

- Rapid and extensive need for high computational power [20].
- Deep knowledge in specialised areas such as: numerical 3D meshing, optimisation techniques, electromagnetics, and Monte Carlo simulations.
- The need of specialised costly software, because commercial ones do not solve 3D numerical problems such as MATLAB.
- The need of accurate measurement/estimation of the neonatal different brain structure conductivities [18].
- The need of accurate placement for EEG electrodes [19].
- The need of accurate head modelling, especially in including or excluding the neonatal fontanel [24] [94] [95].

Since we are using the neonatal single-channel EEG simulators from Section 3.2.2, the forward problem is simplified to become a matter of assigning relative amplitudes and delays to synthetic signals that take into account the multipath reception. Such assigning shall be done with respect to the location of the EEG event in time and space in a way that, the relative amplitudes and delays will be influenced by the depth and properties of the brain structure the signal would propagate through. After that, these synthetic signals would appear on the 21 EEG channels on the neonatal four sphere head model.

The neonatal EEG propagation scheme is constructed in a novel way by combining a space-propagation and a time-propagation models together. The space-propagation model is constructed by modifying the Radiation Transport Equation (*RTE*) that was adapted by *Jacques* in [46], by *Tuchin* in [47], and by *Dehaes et al.* in [48] and by applying it to the neonatal four sphere head model. The equation describes the decay in light intensity (photons power per unit area) due absorption, scattering, and reflection when propagating through biological tissues. The model maps relative amplitudes to each EEG channel which depends on the medium the signal propagate through and the signal path length. On the other hand, the time-propagation model is constructed by considering an adult mean seizure propagation speed and the signal total path length, as the neonatal seizure propagation speed has no record in the literature (to the best of the author's knowledge). The model maps relative delays to each EEG channel depending on the signal total path length. The simplification of considering the mean speed is valid as the seizure propagation speed vary over several orders of magnitude with no clear explanation [50] [51] [52].

Combining both space and time propagation models creates the neonatal EEG propagation model which assigns relative amplitudes and delays to the synthetic seizure signals that would appear on the 21 EEG channels on the neonatal four sphere head model. Such model would have tremendous advantages over the previous methodologies (Combining EEG and MRI, ESL, Particle Numerical Models, BEM, FEM, and FDM) which are:

- Simple descriptive model that does not acquire high level mathematics.
- Calculations acquire relatively low computational power when compared to the previous methodologies.

- Tissue optical properties are more stable and well established when compared to their electrical conductivities. This is true as tissue optical properties had been extensively studied for lots of medical imaging modalities such as Optical Imaging, Near Infrared Spectroscopy (*NIRS*), Diffuse Optical Tomography (*DOT*), and Single Photon Emission Computed Tomography (*SPECT*).

Assumptions for constructing the neonatal EEG propagation model are stated as follows:

- The electrodes and seizure event locations are well known.
- The neonatal four sphere head model has homogenous optical properties, such that each sphere optical properties do not change over time nor space.
- Seizure source location is represented by a point source.
- Background EEG is a stochastic process that will appear on all channels, and will have a normalized amplitude. Thus it does not need to have a propagation model.
- Seizure patterns are not stochastic; consequently the same pattern will appear on all channels in the case of a seizure simulation but with different amplitudes and delays that are proportional to the event generation location in time and space. Thus, it is needed to develop a propagation model for this signal.

3.2.3.1 The Space-Propagation Model

The model describes the seizure signal decay in the neonatal four sphere head model (Equation 3.1). I_i is the signal intensity estimated at electrode i. $R_{j,j+1}$ is the coefficient of Fresnel reflection at the normal beam incidence between regions j and $j+1$. I_o is the initial intensity of the signal. $\mu_{t,j}$ is the extinction coefficient for region j (summation of absorption and scattering coefficients). $D_{i,j}$ is the signal path length in

region j to reach electrode i. $r_i, \theta_i,$ and ϕ_i define electrode i position in spherical coordinates with respect to the seizure event location, and lastly M is the number of head regions and N is the number of electrodes [46] [47] [48].

$$I_i = \left[1 - R_{j,j+1}\right] I_o \exp\left[-\sum_{j=1}^{M} \mu_{t,j}\, D_{i,j}(r_i, \theta_i, \phi_i)\right] \quad , \quad i \in \mathbb{N} | 1 \leq i \leq N \qquad \text{Equation 3.1}$$

The optical properties (absorption coefficients, scattering coefficients, anisotropy factors and refractive indices for every sphere in the Sadlier head model) selected to simulate the EEG propagation are summarised in Table 3.1 and were taken from *Dehaes et al.* in [48]. The same values were used by *Brigadoi et al.* in [25], while some of them were also used by older publications as in [96], [97] and [98]. Table 3.1 values exclude the coefficient of Fresnel reflection from the RTE, and if we assume the initial intensity of the signal to be unity (as we are looking for relative amplitudes) Equation 3.1 is modified into the matrix form and expressed by Equation 3.2, Equation 3.3, and Equation 3.4. By substituting the values and variables from Equation 3.2 and Equation 3.4 into Equation 3.3, the signal path lengths become the only unknowns (Equation 3.5). The complete process for solving the RTE can be found in Appendix G, page 185.

Table 3.1: Optical properties of the neonatal four sphere head model. Absorption coefficient μ_a, scattering coefficient μ_s, anisotropy factor g, refractive index n, region number j. Taken from [48]

j	Tissue Type	$\mu_{a,j}\ (cm^{-1})$	$\mu_{s,j}\ (cm^{-1})$	g_j	n_j
1	**Brain**	0.425	50	0.9	1.3
2	**CSF**	0.041	3.2	0.9	1.3
3	**Skull**	0.16	160	0.9	1.3
4	**Scalp**	0.18	190	0.9	1.3

$$\begin{bmatrix}\mu_{t,1}\\ \mu_{t,2}\\ \mu_{t,3}\\ \mu_{t,4}\end{bmatrix} = \begin{bmatrix}\mu_{a,1}+\mu_{s,1}(1-g_1)\\ \mu_{a,2}+\mu_{s,2}(1-g_2)\\ \mu_{a,3}+\mu_{s,3}(1-g_3)\\ \mu_{a,4}+\mu_{s,4}(1-g_4)\end{bmatrix} = \begin{bmatrix}0.425+5\\ 0.041+0.32\\ 0.16+16\\ 0.18+19\end{bmatrix} = \begin{bmatrix}5.425\\ 0.361\\ 16.16\\ 19.18\end{bmatrix} \quad \text{Equation 3.2}$$

$$\begin{bmatrix}A_1\\ A_2\\ \vdots\\ A_{20}\\ A_{21}\end{bmatrix} \propto \begin{bmatrix}\sqrt{I_1}\\ \sqrt{I_2}\\ \vdots\\ \sqrt{I_{20}}\\ \sqrt{I_{21}}\end{bmatrix} \xrightarrow{yields} \begin{bmatrix}A_1\\ A_2\\ \vdots\\ A_{20}\\ A_{21}\end{bmatrix} = \begin{bmatrix}\sqrt{I_1}\\ \sqrt{I_2}\\ \vdots\\ \sqrt{I_{20}}\\ \sqrt{I_{21}}\end{bmatrix} \quad \text{Equation 3.3}$$

$$\begin{bmatrix}I_1\\ I_2\\ \vdots\\ I_{20}\\ I_{21}\end{bmatrix} = \begin{bmatrix}\exp\left(-\mu_{t,1}\,D_{1,1}-\mu_{t,2}\,D_{1,2}-\mu_{t,3}\,D_{1,3}-\mu_{t,4}\,D_{1,4}\right)\\ \exp\left(-\mu_{t,1}\,D_{2,1}-\mu_{t,2}\,D_{2,2}-\mu_{t,3}\,D_{2,3}-\mu_{t,4}\,D_{2,4}\right)\\ \vdots\\ \exp\left(-\mu_{t,1}\,D_{20,1}-\mu_{t,2}\,D_{20,2}-\mu_{t,3}\,D_{20,3}-\mu_{t,4}\,D_{20,4}\right)\\ \exp\left(-\mu_{t,1}\,D_{21,1}-\mu_{t,2}\,D_{21,2}-\mu_{t,3}\,D_{21,3}-\mu_{t,4}\,D_{21,4}\right)\end{bmatrix} \quad \text{Equation 3.4}$$

$$\begin{bmatrix}A_1\\ A_2\\ \vdots\\ A_{20}\\ A_{21}\end{bmatrix} = \begin{matrix}\sqrt{\exp\left(-5.425\,D_{1,1}-0.361\,D_{1,2}-16.16\,D_{1,3}-19.18\,D_{1,4}\right)}\\ \sqrt{\exp\left(-5.425\,D_{2,1}-0.361\,D_{2,2}-16.16\,D_{2,3}-19.18\,D_{2,4}\right)}\\ \vdots\\ \sqrt{\exp\left(-5.425\,D_{20,1}-0.361\,D_{20,2}-16.16\,D_{20,3}-19.18\,D_{20,4}\right)}\\ \sqrt{\exp\left(-5.425\,D_{21,1}-0.361\,D_{21,2}-16.16\,D_{21,3}-19.18\,D_{21,4}\right)}\end{matrix} \quad \text{Equation 3.5}$$

3.2.3.2 The Time-Propagation Model

The model is constructed by considering the mean seizure propagation speed of an adult and the signal total path length l_i. The neonatal seizure propagation speed has no record in the literature (to the best of the author's knowledge). However adult seizure propagation speed ranges are studied, but it is still poorly understood why it varies over several orders of magnitude, thus the mean value is considered [50] [51] [52].

The time delay t_i of the signal reaching electrode i is calculated by dividing the total path length by the seizure event propagation speed v (Equation 3.6). The signal total

path length is calculated by summing all the signal pathways in all the M regions reaching electrode i. The seizure propagation speed is considered to be 6 cm/s as it is the mean value of the seizure speed uniform distribution (20 - 100 mm/s) [51].

$$t_i = \frac{l_i}{v} = \frac{\sum_{j=1}^{M} D_{i,j}}{v} = \frac{1}{6}\sum_{j=1}^{M} D_{i,j} \quad , \quad i \in \mathbb{N} | 1 \leq i \leq N \qquad \text{Equation 3.6}$$

Finding the signal time delay on all the EEG electrodes on the neonatal four sphere head model is done by extending Equation 3.6 into the matrix form (Equation 3.7), where $M = 4$ and $N = 21$. The relative time delays Δt_i are calculated by subtracting the inferior of the signal time delays vector, making the signal with the largest amplitude (smallest total path length) having a relative time delay of zero (Equation 3.8).

$$\boldsymbol{T} = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_{20} \\ t_{21} \end{bmatrix} = \frac{1}{6}\begin{bmatrix} D_{1,1} + D_{1,2} + D_{1,3} + D_{1,4} \\ D_{2,1} + D_{2,2} + D_{2,3} + D_{2,4} \\ \vdots \\ D_{20,1} + D_{20,2} + D_{20,3} + D_{20,4} \\ D_{21,1} + D_{21,2} + D_{21,3} + D_{21,4} \end{bmatrix} \qquad \text{Equation 3.7}$$

$$\Delta \boldsymbol{T} = \boldsymbol{T} - Inf(\boldsymbol{T}) \qquad \text{Equation 3.8}$$

This concludes the construction of the neonatal EEG propagation model, as the relative amplitudes and delays were calculated for all the 21 EEG channels using the space and time propagation models. The neonatal EEG propagation model has two tuning parameters, which are the absorption and scattering coefficients. Omitting the scattering coefficient creates a pure optical absorbing head model (Model A), while omitting the absorption coefficient creates a pure optical scattering head model (Model B). Using both coefficients creates a third head model that is absorbing and scattering (Model C). Models A, B, and C will be all used in the validation process to choose the best of them that fits published criteria.

3.2.4 Neonatal Multichannel EEG Model

The neonatal multichannel EEG model is intended to: create a system that controls what type of EEG patterns should appear on every channel (seizure or background), the duration of these patterns, and the number of segments in the whole EEG signal. The model should have a number of tuning parameters, making it suitable for different applications such as: testing of algorithms, generation of synthetic neonatal EEG, validation of seizure localising algorithms, and validation of EEG classification results. The model would also generate a flag or a mask that tells where seizures are located on the whole EEG signal. This is very important to make the system suitable for validation purposes.

The neonatal multichannel EEG model has a number of assumptions listed below *(note that some of these assumptions are similar to those in Section 3.2.3 but are mentioned here to emphasize on their need and usage).*

- Seizure source location is represented by a point source that is generated using a uniform distribution, which is localised in the northern/upper hemisphere of the brain model.
- Background EEG is a stochastic process that will appear on all electrodes/channels all the time with a normalized amplitude.
- Seizure patterns are not stochastic, consequently the same pattern will appear on all channels in the case of a seizure simulation but with different amplitudes and delays that are determined by the EEG propagation model.
- The beginning of a seizure is determined by the beginning of the first seizure pattern appearing on any of the channels.

- The end of a seizure is determined by the end of the last seizure pattern appearing on any of the channels.
- The beginning and ending of a background EEG is determined by the absence of a seizure pattern on all the EEG channels simultaneously.

3.2.4.1 Multichannel EEG Epochs Arrangement

The multichannel EEG epochs arrangement produces a mask that will be the baseline for plugging in different types of EEG epochs. The mask or the flag illustrates all the properties that a multichannel waveforms must have. This production is controlled by user inputs such as: the number of segments that the single-channel waveform consist of N_{seg}, the time duration of the waveform segment T_{seg}, the sampling frequency F_s of the waveform, the number of electrodes n_{ch}, and the EEG operation mode m_o.

The number of segments N_{seg} and their duration T_{seg} determine the time duration T_{ch} of a single-channel, which can be described by Equation 3.9. The sampling frequency F_s and the time duration T_{ch} of a single-channel determine the number of samples generated per channel N_{ch}, which can be described by Equation 3.10. The total number of samples generated for the 21 EEG electrodes N_w can be determined by using Equation 3.11. These basic equations are essential in understanding the memory requirements to simulate such multichannel waveforms.

$$T_{ch} = N_{seg}T_{seg} \quad , \quad N_{seg} \in \mathbb{N} \ \& \ T_{seg} \in \mathbb{N} \qquad \text{Equation 3.9}$$

$$N_{ch} = F_sT_{ch} \quad , \quad F_s \in \mathbb{N} \ \& \ T_{ch} \in \mathbb{N} \qquad \text{Equation 3.10}$$

$$N_w = n_{ch}N_{ch} = 21\left(F_s\, N_{seg}\, T_{seg}\right) \ , \quad F_s \in \mathbb{N} \ \& \ N_{seg} \in \mathbb{N} \ \& \ T_{seg} \in \mathbb{N} \qquad \text{Equation 3.11}$$

The EEG mask is initially filled up with zeros indicating only background epochs. Depending on the EEG operation mode m_o the values of the mask change as follows:

- EEG Operation Mode 1:

In this operation mode different background epochs will be generated for every segment in every channel. This multi-plugging is done to ensure the stochasticity of the EEG background waveform. In this operation mode the mask values will remain zeros. Figure 3.6 illustrates the construction of the multichannel EEG waveforms having six segments using mode 1. In this case all segments are different background epochs and are indicated by the green colour (Seg 1.1 is different from Seg. 2.1).

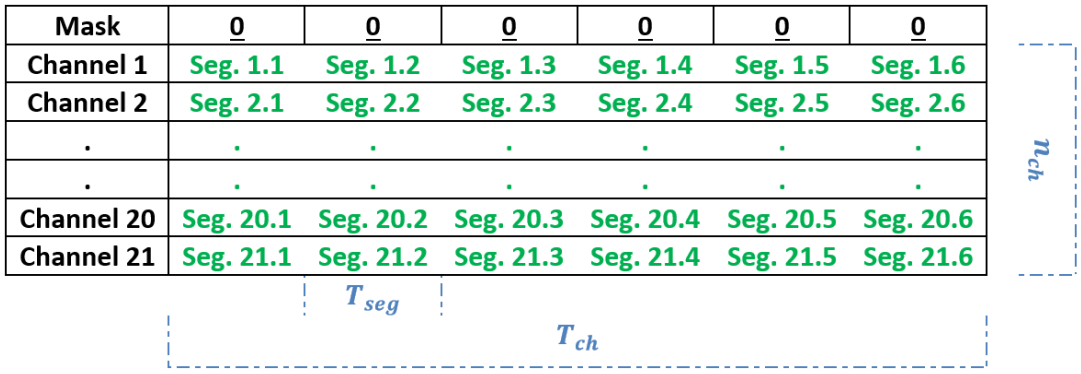

Figure 3.6: EEG operation mode 1. All segments are different background epochs through all channels

- EEG Operation Mode 2:

In this operation mode different seizure epochs will be generated for every segment, but the same seizure behaviour will exist through the different channels. This is done as seizure epochs are not stochastic in nature. In this operation mode the mask values will be filled up with ones. Figure 3.7 illustrates the construction of the multichannel EEG waveforms having six segments using mode 2. In this case all segments are seizure epochs and are indicated by the red colour (Seg 1.1 behaviour is different from Seg. 1.2, but not different from Seg. 2.1).

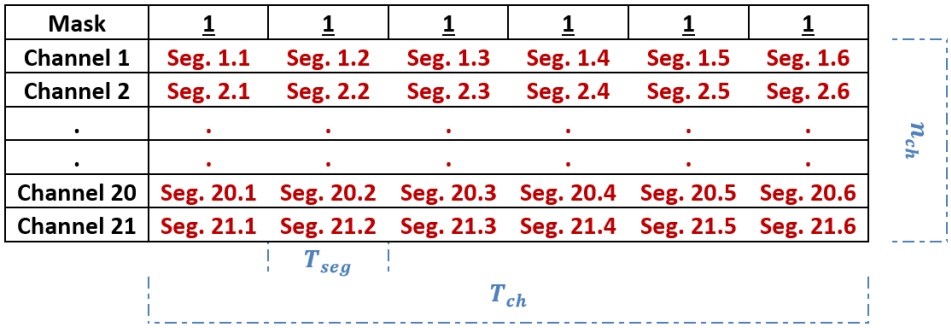

Figure 3.7: EEG operation mode 2. All segments are different seizure epochs but the same through all channels

- EEG Operation Mode 3:

In this operation mode both modes 1 and 2 logics are combined to produce an alternating epoch type signal in every channel. Different background epochs will be generated for every segment that is masked as background in every channel, while different seizure epochs will be generated for every segment that is masked as seizure, but the same seizure behaviour will exist through the different channels. In this operation mode the mask values will be alternating between zero and one. Figure 3.8 illustrates the construction of the multichannel EEG waveforms having six segments using mode 3. In this case all seizure epochs are indicated by the red colour, while background epochs are indicated by the green colour (Seg 1.1 is different from Seg. 2.1, while Seg 1.2 behaviour is different from Seg. 1.4, but not different from Seg. 2.2).

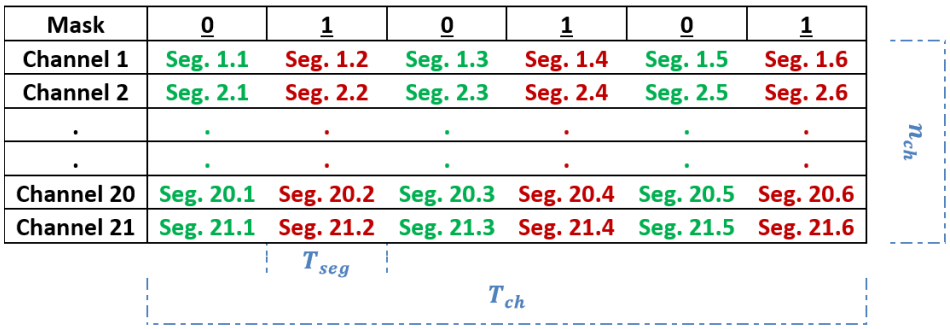

Figure 3.8: EEG operation mode 3. All segments are alternating between background and seizure epochs

3.2.4.2 Multichannel EEG Epochs Generation

The multichannel EEG epochs generation produces neonatal EEG segments that are either background or seizure epochs. This production of segments is controlled by user inputs such as: the number of segments that the single-channel waveform consist of N_{seg}, the time duration of the waveform segment T_{seg}, the sampling frequency F_s of the waveform, the number of electrodes n_{ch}, the EEG operation mode m_o, and the propagation model m_p. One problem arises immediately concerning the sampling frequency F_s, as the neonatal EEG background and seizure simulators in Section 3.2.2 use different sampling frequencies which are 64 and 20 Hz respectively. Seizure epochs are oversampled from 20 Hz to 32 Hz. While background epochs are down-sampled from 64 Hz to 32 Hz satisfying the sampling theorem and solving the problem of different sampling rates. This modification in the sampling frequency of the simulators is valid as most of the background and seizure EEG energies are located below 12 Hz (do not exceed the alpha band 8-12 Hz) [13].

3.2.4.2.a Background EEG Generation:

All EEG waveforms contain background epochs, as all the abnormalities are imposed on the background epoch and might be demolishing it. Background EEG epochs are stochastic in nature and to respect this property different background epochs are generated for every segment in every channel by re-exciting the simulator independently using the new sampling rate of 32 Hz.

Every generated background EEG segment is normalised with respect to its amplitude independently from other segments. This is done so no segments would overcome others by amplitude, and because background EEG information is not embedded in its amplitude rather than in its frequency and entropy.

The total number of background epochs needed in the multichannel waveform depends on the EEG operation mode m_o, the number of segments N_{seg}, and the number of channels $n_{ch} = 21$, as for an EEG operation mode $m_o = 1$, $21\ N_{seg}$ segments are needed, and for an EEG operation mode $m_o = 3$ either $\frac{21\ N_{seg}}{2}$ or $\frac{21\ N_{seg}}{2} + 1$ are needed. This depends whether N_{seg} is even or odd. In the case of an EEG operation mode of 1, all needed background segments are stacked in a $21xN_{ch}$ matrix, which represents the neonatal multichannel background EEG signal (Figure 3.6).

3.2.4.2.b Seizure EEG Generation:

The EEG seizure simulator is excited once to supply an epoch for a specific segment, and its amplitude is normalised. The normalised seizure epoch is supplied to the EEG propagation model to provide delayed/attenuated copies of the signal, and then passed through all 21 channels.

The EEG propagation model requires a specific source location. Seizure source location is represented as a point source that is generated using a three-dimensional uniform distribution. This is done to simulate different seizure source locations for every generated segment. The locations of the seizure event sources are confined within the northern/upper hemisphere of the brain model. This is done to mimic realistic cases as seizures only happen in the brain not in the whole sphere (Figure 3.9). The three-dimensional uniform distribution has an azimuth angle changing from 0 to 2π, an elevation angle that is changing from 0 to π/2, and a radius changing from 0 to 4.75 cm. The maximum radius of the event generation is limited to 4.75 cm which is 0.1 cm less than the brain sphere radius. This is done to ensure that seizures are generated within the brain sphere.

The total number of seizure epochs needed in the multichannel waveform depends on the EEG operation mode m_o, the number of segments N_{seg}, and the number of channels $n_{ch} = 21$, as for an EEG operation mode $m_o = 2$, $21\ N_{seg}$ segments are needed, and for an EEG operation mode $m_o = 3$, either $\frac{21\ N_{seg}}{2}$ or $\frac{21\ N_{seg}}{2} - 1$ segments are needed. This depends whether N_{seg} is even or odd. In the case of an operation mode of 2, all needed seizure segments are stacked in a $21xN_{ch}$ matrix and then summed with the background $21xN_{ch}$ matrix. The resultant matrix represents the neonatal multichannel seizure EEG signal (Figure 3.7). In the case of an operation mode of 3, all needed seizure segments are stacked in $\left[\frac{21N_{ch}}{2}\right]$ or $\left[\frac{21\ N_{seg}}{2} - 1\right]$ matrix and then summed with the background $\left[\frac{21N_{ch}}{2}\right]$ or $\left[\frac{21\ N_{seg}}{2} + 1\right]$ matrix. The resultant matrix represents the alternating multichannel neonatal EEG signal (Figure 3.8).

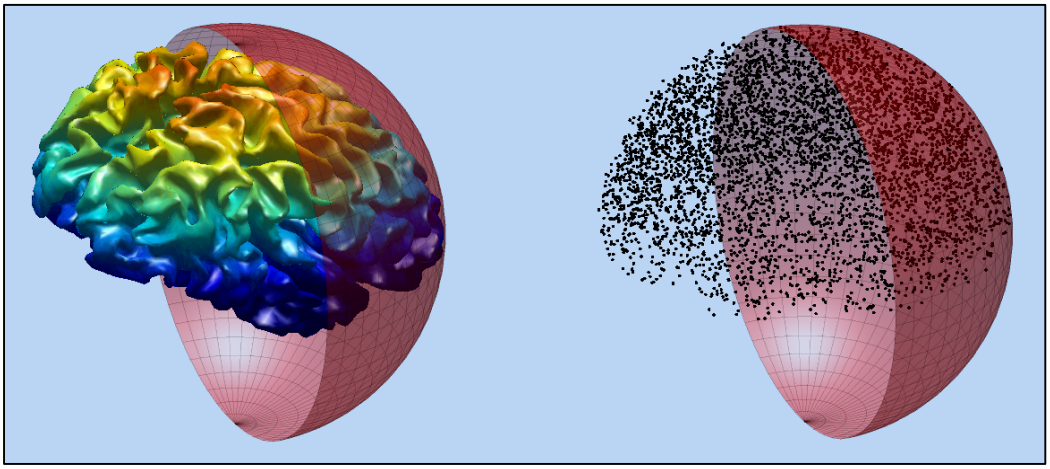

Figure 3.9: Seizure source three dimensional uniformly distributed locations that are localised in the northern-upper hemisphere. 5000 locations are generated in this figure

3.2.5 Relationships with Previous Models

A literature survey for the neonatal multichannel EEG modelling was conducted first in 2014 and later in 2016 (using IEEE database and Google Scholar). There was no

evidence of any modelling attempt for the neonatal multichannel EEG (to the best of the author's knowledge), but two publications were found relevant.

Shen et al. had proposed modelling the adult multichannel EEG by using a new local spatiotemporal prediction method based on support vector machines in [99]. The model estimates the dynamics of the system using time delay embedding according to the Takens embedding theorem. Their training data (taken from [100]) were strictly stationary epochs, and their model had failed to simulate abnormal EEG epochs. Furthermore, *Cui et al.* had proposed modelling the multichannel EEG using a multi-kinetics and multichannel coupled Neural Mass Model (*NMM*) in [101]. Their model was able to simulate background and seizure epochs of Sprague-Dawley rats, but failed to produce nonstationary epochs. Their technique in modelling the multichannel EEG is based on using multiple NMMs to mimic the brain functions, which are based on coupled cortical areas.

Both multichannel modelling techniques are related to the neonatal multichannel EEG model in this thesis, as they all take into consideration the relationships between the EEG channels along with the multipath reception of an EEG event.

3.2.6 EEG Propagation Model Validation

The validation process of the EEG propagation model is done by comparing the model generated Seizure-to-Background Ratio (*SBR*) against a reference using Probability-Probability (*P-P*) plots (more details on Probability-Probability plots can be found in Appendix I, page 211).

Different SBRs had been suggested and used in many publications. *Mesbah et al.* had used an SBR value of 5 dB in [102] and 10 dB in [103], while *Stevenson et al.* and

Celka et al. had used a uniform random SBR ranging from 10 dB to 30 dB in [13] and [40] respectively. *Stevenson et al.* had also used a uniform random SBR, but ranging from 10 dB to 20 dB in [42]. Furthermore, different range was used by *Rankine et al.* in [104] starting from -2.5 dB to 6 dB.

In this thesis, the SBR reference is chosen to have a normal distribution with a mean value of 15 and variance of 25 to cover most of the published ranges. Such mean and variance have; 68% confidence interval ranging from 10 to 20 dB, 90% confidence interval ranging from 5 to 25 dB, and a 95% confidence interval ranging from 0 to 30 dB. The neonatal EEG propagation model has two tuning parameters that produced three different propagation models (Model A, Model B, and Model C), hence all models must be validated against the reference SBR.

The model validation process consists of the following steps:

1. Generate 50 minutes of multichannel background EEG (200 segments per channel, 15 seconds each, 96,000 samples per channel, and 2,016,000 samples in total).
2. Generate 200 seizure source locations.
3. Generate 50 minutes of multichannel seizure EEG by utilising the 200 source locations and by using models A, B, and C.
4. Calculate the power of every segment in the background multichannel signal (This produces a 21x200 background power matrix).
5. Calculate the power of every segment in models A, B, and C seizure multichannel signals (This produces a 21x200 power matrix for every model).
6. Calculate the SBR for all models by dividing each value in the models power matrices by every value in the background power matrix. This produces 17,640,000 SBR values for every model.

7. Compare the models SBR values with the SBR reference using P-P plots.
8. Calculate the model normal line goodness of fit metric R^2 when fitting the reference normal line.
9. Choose the model that creates the largest R^2 value.

3.2.7 Neonatal Multichannel EEG Model Validation

The validation process of the neonatal multichannel EEG model is done by quantifying the alikeness between the synthetic multichannel generated waveforms and the real multichannel neonatal EEG (described in Section 3.1, page 43). EEG alikeness is quantified using Spearman's correlation, which is computed on segment per segment basis and averaged across all channels, thus producing a correlation distribution (more details on correlation can be found in Appendix H, page 188). The real EEG data are comprised of 20 channels, consequently channel F_{pz} will be omitted from the synthetic model (only for the validation process) as it is the furthest electrode from any seizure event (Figure 2.3 and Figure 3.3). The EEG model validation process consists of the following steps:

1. Repeat the first 2 steps of the EEG propagation validation process, but by generating 100 segments per channel, 8 seconds each.
2. Extract clean and annotated background and seizure segments (8 seconds each) from the real neonatal EEG dataset (this will be done for all 7 patients).
3. Calculate and average the absolute Spearman's correlation between the synthetic and real multichannel waveforms for every segment, thus producing a correlation distribution (e.g. segment no. 1 of synthetic multichannel background versus segment no. 20 of real multichannel background).

4. Calculate the mean of the background and seizure EEG correlation distributions.
5. Repeat steps 3 and 4, but by transforming the data into the frequency and time-frequency domains.

This process will produce six correlation distributions, 3 of them are for the background EEG comparison (time, frequency, and time-frequency comparisons), and the others are for the seizure EEG comparison.

3.3 Synthetic Neonatal Multichannel EEG Artifacts Model

3.3.1 Neonatal EEG Artifacts Model

The neonatal EEG artifacts model is constructed by combining different synthetic EEG artifacts to corrupt the synthetic neonatal multichannel EEG waveforms. Blood Vessel Pulsation (*BVP*), Electrocardiogram Spikes (*ECGS*), and Short-Time High-Amplitude (*STHA*) artifacts are chosen to be the contamination signals. These artifacts are physiological, and this type of signals is chosen for modelling because of its nature of occurrence, as neonatal seizure and background EEG spectrums overlap with these signals. BVP, ECGS, and STHA physiological artifacts are chosen specifically because of their unique behaviour in mimicking EEG seizure patterns, and this contributes to high false rates in automatic seizure detection systems [9] [10] [53] (more details on EEG artifacts can be found in Section 2.2.5, page 23 and in Section 2.2.6, page 25).

3.3.1.1 Blood Vessel Pulsation

The BVP artifact corresponds to the pulsation effect of an electrode close to a blood vessel. It is a continuous oscillatory Gaussian type of artifact, highly correlated in time, but uncorrelated to the other activities in the EEG signal [53] [54]. The BVP artifact is

modelled as a sine wave having a Gaussian probability distribution, and a frequency close to the heart rate of neonates [55].

The model is constructed by using Equation 3.12, where f_{BVP} is the pulsation frequency, and $\mathcal{N}_g(t, \mu_1, \sigma_1)$ is a Gaussian noise having a mean of μ_1 and a standard deviation of σ_1 (Equation 3.13). The Gaussian noise is added to the sine wave to force the probability distribution of the BVP artifact in taking a bell-shape, as sine waves are not Gaussian signals in nature and this would contradict the characteristics of the BVP artifact. This addition is also important to simulate the noisy nature of the BVP artifact. The BVP artifact $BVP(t)$ is then normalised to have a maximum amplitude of one.

The pulsation frequency f_{BVP} of the sine wave is chosen as 2 Hz because both *Janardhan et al.* in [54] and *Matic et al.* in [55] agreed on this value as a valid frequency in modelling the BVP artifact. The Gaussian noise $\mathcal{N}_g(t, \mu_1, \sigma_1)$ has a mean value of zero ($\mu_1 = 0$) and a standard deviation of 0.5 ($\sigma_1 = 0.5$).

$$BVP(t) = \sin(2\pi f_{BVP} t) + \mathcal{N}_g(t, \mu_1, \sigma_1) \quad \text{Equation 3.12}$$

$$\mathcal{N}_g(t, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \quad \text{Equation 3.13}$$

3.3.1.2 Electrocardiogram Spikes

The ECGS artifact corresponds to the corrupted QRS complexes of the ECG picked up by the EEG electrodes when placed in the vicinity of a blood vessel. This artifact appears as a sharp spike and the occurrence of these spikes is not strictly periodical but correlated with the neonatal heart rate [54] [55]. The ECGS artifact is modelled as a spike train signal having a Gaussian probability distribution [56].

The model is constructed by using Equation 3.14, where f_e is the frequency of the spike train, and $\mathcal{N}_g(t, \mu_2, \sigma_2)$ is a Gaussian noise having a mean of μ_2 and a standard deviation

of σ_2 (Equation 3.13). The Gaussian noise is added to the spike train to force the probability distribution of the ECGS artifact in taking a bell-shape as spikes are not Gaussian signals in nature and this would contradict the characteristics of the ECGS artifact. This addition is also important to simulate the noisy nature of the ECGS artifact. The ECGS artifact $ECGS(t)$ is then normalised to have a maximum amplitude of one. *Janardhan et al. 2015*, *Matic et al. 2009*, and *Khalifa et al. 2010* had all disagreed on the frequency of the spike train, as it was 1 Hz in [54], 2.5 Hz in [55], and 1.2 Hz in [56]. In this thesis, the spike train frequency f_e is chosen as 1 Hz because *Janardhan et al.* in [54] is the most recent publication and it is the closest to *Khalifa et al. 2010* in [56] where they dealt with neonatal EEG artifacts. The Gaussian noise $\mathcal{N}_g(t, \mu_2, \sigma_2)$ has a mean value of zero ($\mu_2 = 0$) and a standard deviation of 0.1 ($\sigma_2 = 0.1$).

$$ECGS(t) = \begin{Bmatrix} 1 \;\; ; \;\; t = \dfrac{n}{f_e} \\ 0 \;\; ; \;\; t \neq \dfrac{n}{f_e} \end{Bmatrix} \;\; ; \;\; n \in N^+ \Bigg\} + \mathcal{N}_g(t, \mu_2, \sigma_2) \qquad \text{Equation 3.14}$$

3.3.1.3 Short-Time High-Amplitude

The STHA artifact corresponds to the burst suppression in the neonatal brain and can also correspond to movement activities or electrode artifacts [56] [57]. It is an EEG artifact that is characterized by periods of high-voltage electrical activity alternating with periods of no activity in the brain. The STHA EEG artifact is modelled using a heavy-tailed noise [58].

The model is constructed as a type of noise that exhibits impulsive characteristics. The probability density function (PDF) of this impulsive noise is modelled by a heavy-tailed distribution, and Levy stable symmetrical stochastic process is used to simulate such

noise [58]. The Levy stable symmetrical stochastic process is defined by Equation 3.15, which is the characteristic function of the process $\varphi(t)$ [105]. α is the stability index, β is the skewness parameter, γ is the scale parameter, and δ is the location parameter. Their varying ranges are $0 < \alpha \leq 2, -1 \leq \beta \leq 1, \gamma > 0, and\ \delta \in R$. It is worthy of noting that the well-known Gaussian and Cauchy distributions are special cases of this distribution and they exist when the stability index α is 2 or 1 respectively [105]. The STHA artifact is then normalised to have a maximum amplitude of one. The characteristic function parameters are chosen from *Liang et al.* in [105]. The stability index α and the skewness parameter β have values of 1.4 and 0.8 respectively producing a Levy distribution. The scale parameter γ is unity so it does not affect the distribution amplitude, and lastly the location parameter δ has a value of zero to centralise the levy PDF.

$$\varphi(t) = \begin{cases} exp\left(j\delta t - \gamma^{\alpha}|t|^{\alpha}\left[1 - j\beta \tan\left(\frac{\pi\alpha}{2}\right) sign(t)\right]\right) & ;\ \alpha \neq 1 \\ exp\left(j\delta t - \gamma|t|\left[1 + j\beta\left(\frac{2}{\pi}\right) sign(t) \ln|t|\right]\right) & ;\ \alpha = 1 \end{cases} \quad \text{Equation 3.15}$$

3.3.2 Multichannel Neonatal EEG Artifacts Model

The multichannel neonatal EEG artifacts model is intended to create a system that controls: the duration of patterns, the amplitude of patterns, and the number of artifact segments in the whole EEG signal. The model should have a number of tuning parameters, making it suitable for different applications such as: testing of artifact detection algorithms, generation of synthetic neonatal EEG artifacts, and validation of artifact removal algorithms. The model would also generate an artifact flag or a mask that tells where EEG artifacts are located on the whole EEG signal. This is very important to make the system suitable for validation purposes.

The neonatal multichannel EEG artifacts model assumptions are summarised as:

- EEG Artifacts change behaviour on EEG segment basis (the model will be re-excited for every segment), such that different artifact behaviours would appear across a single channel.
- EEG Artifacts do not change behaviour on EEG channel basis (the same generated artifact segment will be replicated on all channels), such that the same artifact behaviour would appear on all channels for the same EEG segment.
- The beginning of an artifact is determined by the beginning of the first artifact pattern appearing on any of the channels.
- The end of an artifact is determined by the end of the last artifact pattern appearing on any of the channels.
- The beginning and ending of a clean EEG segment is determined by the absence of an artifact pattern on all the EEG channels simultaneously.

3.3.2.1 Multichannel EEG Artifact Epochs Arrangement

The EEG artifact epochs arrangement produces a mask that forms the baseline for plugging in EEG artifact epochs. The mask (or flag) illustrates all the properties that a multichannel artifact waveforms must have. This production is controlled by user inputs such as: the number of segments that the single-channel artifact waveform consist of N_{a_seg}, the time duration of the artifact waveform segment T_{a_seg}, the sampling frequency F_s of the artifact waveform, the number of electrodes n_{ch}, and the artifact operation mode m_{ao}. The number of artifact segments N_{a_seg} and their duration T_{a_seg} determine the time duration T_{a_ch} of a single channel, which can be described by Equation 3.16. The sampling frequency F_s and the time duration T_{a_ch} of a single channel determine the

number of artifact samples generated per channel N_{a_ch}, which can be described by Equation 3.17. The total number of artifact samples generated for the 21 EEG electrodes N_{a_w} can be determined by using Equation 3.18.

$$T_{a\_ch} = N_{a\_seg} T_{a\_seg} \quad , \quad N_{a\_seg} \in \mathbb{N} \ \& \ T_{a\_seg} \in \mathbb{N} \qquad \text{Equation 3.16}$$

$$N_{a\_ch} = F_s T_{a\_ch} \quad , \quad F_s \in \mathbb{N} \ \& \ T_{a\_ch} \in \mathbb{N} \qquad \text{Equation 3.17}$$

$$N_{a\_w} = n_{ch} N_{a\_ch} = 21\left(F_s \, N_{a\_seg} \, T_{a\_seg}\right) \ , \ F_s \in \mathbb{N} \ \& \ N_{a\_seg} \in \mathbb{N} \ \& \ T_{a\_seg} \in \mathbb{N} \qquad \text{Equation 3.18}$$

The artifact mask is initially filled up with zeros indicating that no artifact epochs exist. Depending on the artifact operation mode m_{a_o} the values of the artifact mask change.

- Artifact operation Mode 1:

In this artifact operation mode, no artifact epochs will be generated. This operation is essential in generating clean EEG signals that can be used for EEG corruption evaluation purposes, and in artifact removal validation.

- Artifact operation Mode 2:

In this artifact operation mode, different artifact epochs will be generated for every segment, but the same artifact behaviour will exist through the different channels. This multi-plugging is done to ensure different artifact situation in every segment to cover all possible artifact behaviour scenarios. In this operation mode, the artifact mask values will be filled up with ones. Figure 3.10 illustrates the construction of the multichannel artifact waveforms having six segments and using an artifact operation mode of 2. In this case all segments are artifact epochs and are indicated by the grey colour (Seg. 1 behaviour is different from Seg. 2).

- Artifact operation Mode 3:

In this artifact operation mode, different artifact epochs will be generated only for the segments that exist on the first half of the signal length, but the same artifact behaviour will exist through the different channels. The first half of the artifact mask values are filled up with ones while the rest is filled up with zeros *(Note that if the number of segments was odd, artifact operation mode 3 will floor the artifact segments number).* Figure 3.11 illustrates the construction of the multichannel artifact waveforms having six segments and using an artifact operation mode of 3. In this case half of the segments are artifact epochs and are indicated by the grey colour (Seg. 1 behaviour is different from Seg. 2).

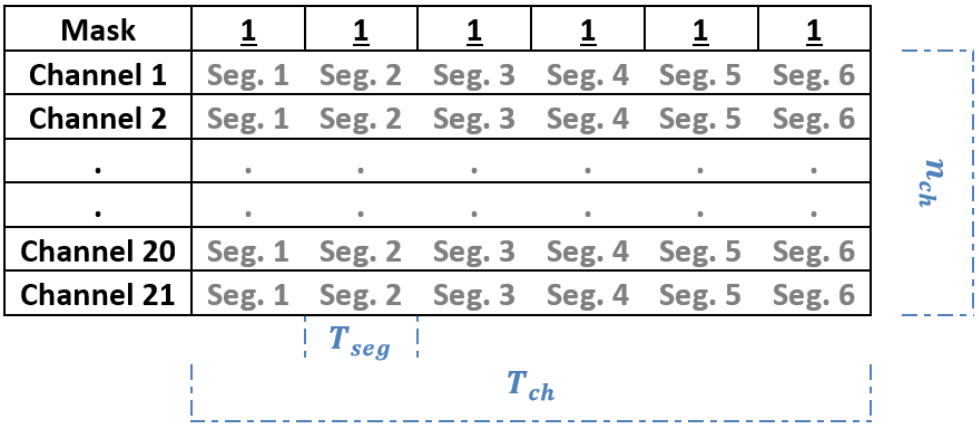

Figure 3.10: Artifact operation mode 2. All segments are different artifact epochs but the same through all channels

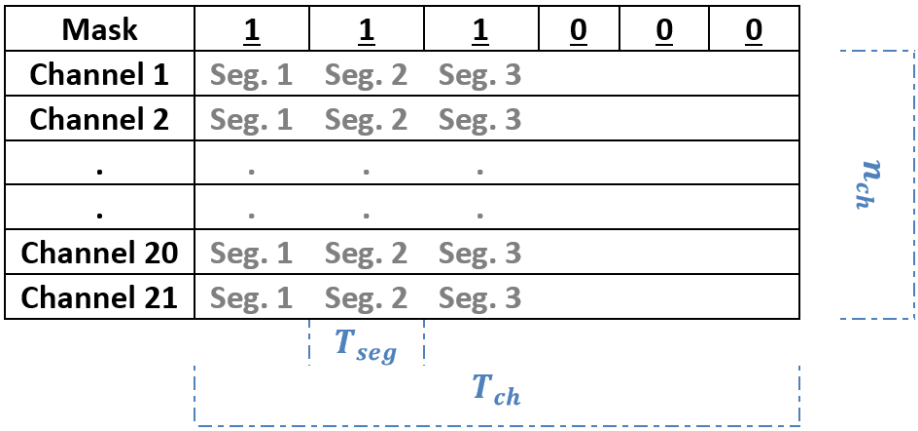

Figure 3.11: Artifact operation mode 3. The first half segments are different artifact epochs but the same through all channels

3.3.2.2 Multichannel EEG Artifacts Epochs Generation

The EEG artifacts epoch's generation produces neonatal EEG artifact segments that are combined from the BVP, ECGS, and the STHA signals by using the artifact mask. This production of artifact epochs is controlled by user inputs such as: the number of segments that the single-channel artifact waveform consist of N_{a_seg}, the time duration of the artifact waveform segment T_{a_seg}, the sampling frequency F_s of the waveform, the number of electrodes n_{ch}, and the artifact operation mode m_{a_o}.

All EEG artifacts; BVP, ECGS, and STHA have normalised amplitudes. The combined EEG artifacts $w(t)$ can be constructed using Equation 3.19, where ρ_1, ρ_2, and ρ_3 are the contamination factors for the $BVP(t)$, $ECGS(t)$, and the $STHA(t)$ respectively. The contamination factors values for the BVP and the ECGS artifacts were taken from *Matic et al.* in [55] as 1.5 and 6 respectively, while the contamination factor value of the STHA artifact was taken from *Brotchie et al.* in [58] as 15.

The total number of the EEG artifacts epochs needed in the multichannel waveform depends on the artifact operation mode m_{a_o}, the number of artifact segments N_{a_seg}, and the number of channels $n_{ch} = 21$, as for an artifact operation mode $m_{a\_o} = 2, 21$ N_{a_seg} segments are needed, and for an artifact operation mode $m_{a\_o} = 3$, either $\frac{21\ N_{a\_seg}}{2}$ or $\frac{21\ N_{a\_seg}}{2} - 1$ are needed. This depends whether N_{a_seg} is even or odd. In the case of an operation mode of 2, all needed artifact segments are stacked in a $21xN_{a\_ch}$ matrix, which represents the neonatal multichannel EEG artifact signal $w(t)$ (Figure 3.10)

$$w(t) = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{bmatrix} [BVP(t) \quad ECGS(t) \quad STHA(t)] \qquad \text{Equation 3.19}$$

Finally, the neonatal multichannel corrupted EEG waveforms are constructed using Equation 3.20, where; $s(t)$ is the neonatal multichannel corrupted EEG waveform, $x(t)$ is the neonatal multichannel clean EEG waveform, $w(t)$ is the neonatal multichannel EEG artifact waveform, and α is a factor that alters the signal-to-artifact ratio (*SAR*). The corrupted multichannel signal is simply constructed by adding up the clean multichannel EEG signal with the multichannel EEG artifact signal. The generation of $s(t)$ can be tuned by changing the EEG operation mode m_o, the artifact operation mode m_{a_o}, and the SAR. This creates lots of possibilities in generating multichannel EEG waveforms; it can be e.g. only background, only seizure, or alternating between the two; and each of these possibilities can either be clean, or have artifacts across all segments, or on the first half of the EEG waveform with different SARs.

Evaluation of EEG information masking due artifacts is studied and discussed extensively in Appendix H, page 188. This study is beyond the scope of this thesis; hence it will not be discussed here.

$$s(t) = x(t) + \alpha\, w(t)$$ Equation 3.20

CHAPTER 4: APPLICATION TO TIME-FREQUENCY ALGORITHMS ENHANCEMENT FOR ABNORMALITY DETECTION

In this chapter, abnormality detection time-frequency enhancement techniques used in this thesis are explained in depth along with all the procedures undertaken. First, an automated artifact detection and removal system utilising BSS in time and time-frequency domains is constructed to filter the contaminated EEG patterns. Time and time-frequency BSS techniques are utilised rather than frequency domain methods because they require standard real-valued algorithms rather than complex-valued, and additionally time-frequency methods produce unique signatures and are not limited to non-Gaussian signals. After that, an optimisation procedure for the automated artifact detection technique is discussed and an evaluation method for assessing the quality of filtered EEG is presented to quantify the system performance. Furthermore, an automated EEG abnormality detection system is constructed embedding fused and extended multichannel EEG features to detect neonatal seizures. This stage considers clean, corrupted, and BSS filtered EEG as input patterns to quantify the differences in detection accuracy for a range of signal-to-artifact ratios. Finally, validation of the automated EEG abnormality detection performance is discussed and explained using cross-validation techniques in order to produce reliable results.

4.1 Automatic Artifacts Detection and Removal

The automatic artifacts detection and removal stage is intended to create a system that: separates the EEG and artifact source signals (x_i and w_i) from the corrupted mixture y_i by estimating an un-mixing matrix $\widehat{W}$ using Blind Source Separation (*BSS*)

in the time and time-frequency domains. Furthermore, it detects and determines the artifact components by processing the estimated source signals to remove its contribution from the mixture, in order to filter the corrupted multichannel EEG signal (more details of BSS can be found in Section 2.4.2, page 31 and Appendix B, page 164).

In this thesis, estimation of the number of sources is done by choosing the percentage of the total eigenvalues explaining the variance as one (more details on estimation of the number of sources can be found in Section 2.4.3, page 36). This means that the number of estimated independent components (source signals) will be equal to the number of EEG channels (21 components). This is done as the number of EEG channels is relatively low when compared to dense EEG electrode configurations (62 channels are used in [69]), so further reduction in dimensionality might mean loss of information.

4.1.1 Neonatal EEG Artifact Detection

Once BSS has been performed, it is crucial to carefully identify and select the artifact related component for removal purposes. Selecting the artifact-related component must be automated and should not rely on user inputs. Furthermore, BSS methods have limitations in the sense that an estimated signal $\hat{s}_i$ cannot determine the variance of the source signal s_i. That is, there exists an infinite number of factors α_i as described by Equation 3.1. Moreover, α_i can be always chosen in a way to create a unit variance signal, but this still leaves the ambiguity of the sign. This implies that artifact component detection must not rely on the amplitude of the components rather than its shape or the morphology.

$$\hat{s}_i = \frac{1}{\alpha_i} s_i$$ Equation 3.1

The most widely reported EEG artifact selection criterion is based on the amount of correlation between the resulting independent components and an artifact reference signal [69]. The Spearman's Correlation Coefficient (*SCC*) proves to be a good choice to compare the artifact reference signal with the independent components because it depends on the relative shape of the signal [10] [54] [55] [68] [69].

An artifact reference signal is generated independently from the corrupted multichannel EEG waveform containing; BVP, ECGS, and STHA signals, and having different contamination factors from the multichannel model. The 21 estimated source signals are correlated with this reference signal using the SCC. If the maximum correlation output was above a certain threshold (C_{thr}), this source signal is considered an EEG artifact, and an artifact prediction mask (or flag) will contain ones for this segment indicating the detection of an EEG artifact. If the maximum correlation output was below the chosen threshold, this segment will be considered as clean EEG and no further processing will be applied and the artifact prediction mask will contain zeros for this segment. If multiple source signals produce Spearman's correlations above the correlation threshold C_{thr}, the source signal with the maximum correlation output will be considered as the EEG artifact, while the other is not.

This correlation process between the artifact reference signal and the independent components will be repeated for all the BSS algorithms (Fast-ICA [106], Robust-ICA [107], BSS-CCA [108], SOBI [109] [110], and JADE [111]) and for all EEG segments covering the whole neonatal multichannel EEG waveform in time and time-frequency

domains. Figure 4.1 illustrates the correlation process between the artifact reference signal and the independent components/source signals.

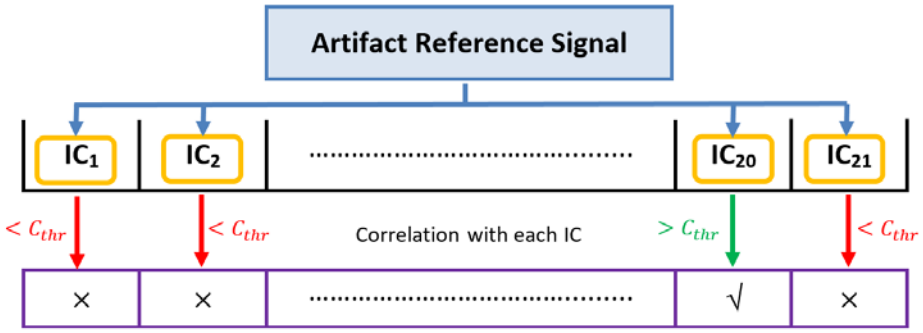

Figure 4.1: Artifact detection procedure based on Spearman's correlation coefficient. IC is the Independent Component, C_{thr} is the correlation threshold, × means no artifact detected, and √ means artifact is detected

4.1.2 Artifact Detection Optimisation

The artifact detection technique needs to be optimised, as a low correlation threshold would result in detecting more artifacts than there really exist, while a high correlation threshold would result in detecting less artifacts than there really exist. The artifact detection optimisation is done by maximising the alikeness between the artifact mask, and the time and time-frequency prediction masks. Such optimisation scheme must be applied to all BSS algorithms independently as they process data differently even though they are based on the same BSS knowledge. Artifact detection performance is quantified in terms of sensitivity, specificity, and balanced accuracy (more details on detection performance metrics can be found in Appendix D, page 177). Artifact detection optimisation aims to find a correlation threshold that would maximise both sensitivity and specificity for every BSS algorithm in the time and time-frequency domains. This is a multivariable optimisation problem, but it reduces to a single variable by maximising the balanced accuracy, as it will then maximise both sensitivity and specificity and will reduce the optimisation problem complexity. Such maximisation

will not produce a deterministic answer, because maximising the balanced accuracy can be done using a range of correlation thresholds. Consequently, artifact detection optimisation becomes a statistical optimisation problem that would produce a random variable instead of a deterministic answer. The artifact detection statistical optimisation process consists of the following steps:

1. Generate 25 minutes (21 channels, 100 segments each channel, 15 seconds each segment) of clean multichannel EEG using an EEG operation mode of 3 (alternating between background and seizure epochs).
2. Generate 25 minutes (21 channels, 100 segments each channel, 15 seconds each segment) of multichannel EEG artifact using an artifact operation mode of 3 (artifacts only exist on the first half of the waveform).
3. Add the generated multichannel artifacts to the clean EEG from step 1, then segment the resultant waveform to reduce the needed computational time and memory.
4. Estimate the independent components using Fast-ICA, Robust-ICA, BSS-CCA, JADE, and SOBI algorithms for every segment in the corrupted multichannel EEG.
5. Initiate a correlation threshold of 0.
6. Apply the artifact detection technique on all the BSS components for all the corrupted multichannel EEG segments using the defined correlation threshold.
7. Generate prediction masks for all BSS algorithms.
8. Calculate sensitivity, specificity, and balanced accuracy for all BSS algorithms by comparing the prediction masks with the true artifact mask.
9. Update the correlation threshold value by adding a correlation threshold step of 0.001.

10. Repeat steps 6-11 until correlation threshold reaches unity (This will create a loop of 1001 iteration).
11. Repeat steps 1-11 hundred times to produce the statistical optimisation random variable for every BSS algorithm.

This extensive operation will produce 100 curves for sensitivity, specificity, and the balanced accuracy for every BSS algorithm (namely performance curves), where every curve consists of 1001 samples. This is done to ensure the integrity of the results and to establish a robust random variable for the correlation threshold of every BSS algorithm. Note that, the same steps are followed to optimise the artifact detection for the time-frequency BSS. The only difference is that an extra step will be added after step 4, which transforms the segmented corrupted multichannel EEG into the time-frequency domain. Moreover, steps 1-11 could not be repeated hundred times due to limited memory and processing power, as each step produced approximately 10 GB of data. Instead, it was repeated for three times using a correlation threshold range of 0-0.3 and a step of 0.003, as preliminary results showed that maximum values would fall within this range.

4.1.3 Neonatal EEG Artifact Removal

After detecting the artifact components in the corrupt multichannel EEG, the next step becomes removing the contribution of these components. By expanding the matrix notations in the BSS model, we can see that each column in the estimated mixing matrix corresponds to contributions from a specific source signal estimate (Equation 3.2). Following the detection example in Figure 4.1 where the 20^{th} independent component is the detected artifact signal, the 20^{th} column of the estimated mixing matrix holds the

contributions of that artifact. By removing the 20th column of the estimated mixing matrix, the electrode observations become cleaner versions of the corrupted multichannel EEG waveform (Equation 3.3). This process will be repeated in time and time-frequency domains for all the BSS algorithms and for all the EEG segments that were marked or flagged as corrupted.

$$\begin{bmatrix} \hat{y}_1 \\ y_2 \\ \vdots \\ \hat{y}_{20} \\ \hat{y}_{21} \end{bmatrix} = \begin{bmatrix} \hat{m}_{1,1} & \hat{m}_{1,2} & \cdots & \hat{m}_{1,20} & \hat{m}_{1,21} \\ \hat{m}_{2,1} & \hat{m}_{2,2} & \cdots & \hat{m}_{2,20} & \hat{m}_{2,21} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \hat{m}_{20,1} & \hat{m}_{20,2} & \cdots & \hat{m}_{20,20} & \hat{m}_{20,21} \\ \hat{m}_{21,1} & \hat{m}_{21,2} & \cdots & \hat{m}_{21,20} & \hat{m}_{21,21} \end{bmatrix} \begin{bmatrix} \hat{s}_1 \\ \hat{s}_2 \\ \vdots \\ \hat{s}_{20} \\ \hat{s}_{21} \end{bmatrix}$$

Equation 3.2

$$\begin{bmatrix} \hat{y}_1 \\ y_2 \\ \vdots \\ \hat{y}_{20} \\ \hat{y}_{21} \end{bmatrix} = \begin{bmatrix} \hat{m}_{1,1} & \hat{m}_{1,2} & \cdots & 0 & \hat{m}_{1,21} \\ \hat{m}_{2,1} & \hat{m}_{2,2} & \cdots & 0 & \hat{m}_{2,21} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \hat{m}_{20,1} & \hat{m}_{20,2} & \cdots & 0 & \hat{m}_{20,21} \\ \hat{m}_{21,1} & \hat{m}_{21,2} & \cdots & 0 & \hat{m}_{21,21} \end{bmatrix} \begin{bmatrix} \hat{s}_1 \\ \hat{s}_2 \\ \vdots \\ \hat{s}_{20} \\ \hat{s}_{21} \end{bmatrix}$$

Equation 3.3

4.1.4 Artifact Removal Evaluation

The output of the artifact removal scheme in the time and time-frequency domains must be evaluated to quantify its performance. The performance evaluation is done by quantifying the alikeness between the filtered EEG (artifacts removed) and the clean multichannel EEG waveform in the time and time-frequency domains. This is done by using the Normalised Root Mean Squared Error (*NRMSE*) and the Pearson's Correlation Coefficient (*PCC*) metrics (more details on performance metrics can be found in Appendix D, page 177). The artifact removal performance evaluation process consists of the following steps:

6. Repeat the first 4 steps of the artifact detection optimisation, but with an artifact operation mode of 2, and segment both clean and corrupted waveforms.

7. Identify and remove artifact components using the optimised artifact detection technique.
8. Estimate the clean multichannel EEG segments for all BSS algorithms.
9. Calculate time-frequency NRMSE and PCC for every estimated segment.

4.2 Automatic EEG Abnormality Detection

The automatic EEG abnormality detection stage is intended to create a system that detects neonatal EEG seizures, by utilising fused and selected multichannel time-frequency features along with a statistical classifier.

4.2.1 Feature Extraction and Fusion

Single-channel feature extraction is done by extending the time or frequency domain features into the joint time-frequency domain (more details on feature extraction and the time-frequency features equations can be found in Section 2.5.1, page 36 and Appendix C, page 173). In this thesis, time-frequency features are categorised as: statistical, spectral, instantaneous frequency, and sub-bands energy features. These features are utilised as they have shown good discrimination power for the neonatal EEG [27] [30] [36] [63] [78] [80] [81]. After extracting the time-frequency features $\mathbf{J}$ from every single channel $\mathbf{Z}$, these features are fused together forming a global feature set $\mathbf{\Psi}$ using the Multichannel Feature Fusion Sum Basis (*MFFSB*) technique [30] (more details on feature fusion can be found in Section 2.5.2, page 37). The feature extraction and fusion process for the 21 EEG channels and for the 16 time-frequency features is depicted in Figure 4.2. This process will be applied on the clean, corrupted, and BSS estimated multichannel EEG waveforms to be further processed.

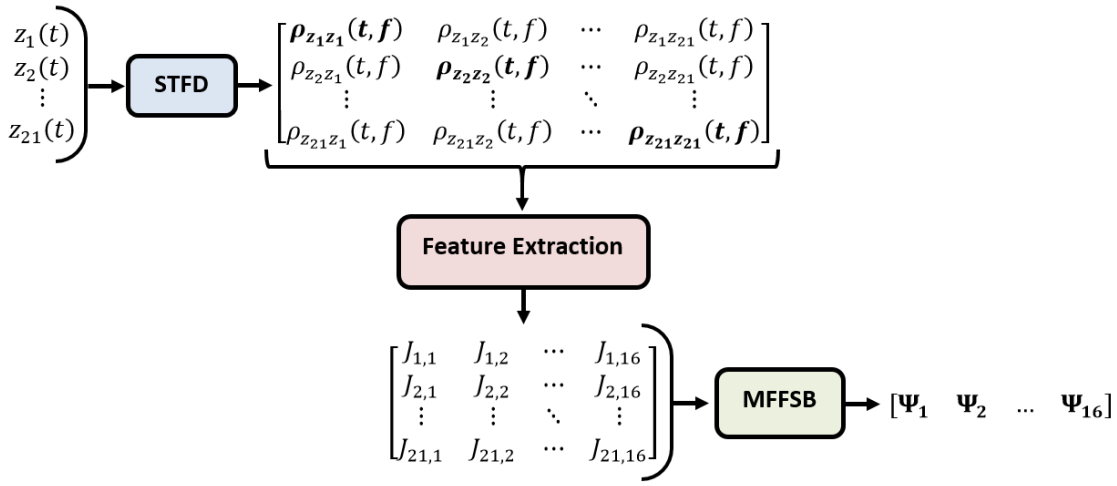

Figure 4.2: Feature extraction and fusion process. STFD is the Spatial Time-Frequency Distribution, and MFFSB is the Multichannel Feature Fusion Sum Basis. Note that only auto-TFDs (*bolded in the STFD matrix*) will be used for the feature extraction stage

4.2.2 Feature Selection and Classification

Once time-frequency features had been extracted and fused, an optimal subset from these features is selected on the basis of relevance and non-redundancy to be fed into a statistical classifier.

The feature selection process is constructed by using the filter approach, and more specifically by computing Fisher's score for each feature (more details on feature selection can be found in Section 2.5.3, page 38). Features are ordered according to their Fisher's scores and the top ranked m features are selected. This process is only applied on the clean EEG features and not on the corrupted or estimated EEG features. This is done to quantify; the possible reduction in classification accuracy due EEG corruption, and the possible increment in accuracy due EEG filtering. The selected subsets $\boldsymbol{\psi}$ are then fed simultaneously into nonlinear SVM classifier to discriminate the multichannel EEG segments into background or seizure patterns. The nonlinear SVM

classifier is constructed using an RBF kernel with σ of 1 (more details on classification and SVM can be found in Section 2.5.4, page 40).

Since the Fisher's score method is suboptimal, an iterative feature selection and classification process must be utilised. This is done by increasing the top ranked feature size m from 1 up to the total number of features and feeding each to the SVM classifier (Note that all features will be selected when $i = 16$). Such iterative process would produce an accuracy profile for the SVM classifier, illustrating its accuracy relation with the size of selected features (Figure 4.3).

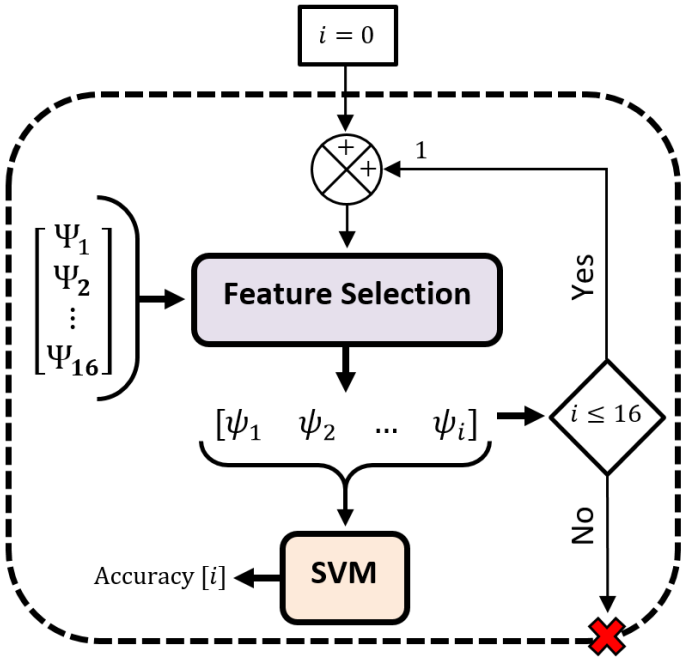

Figure 4.3: Feature selection and classification iterative process (The maximum number of selected features is 16)

4.2.3 Defining Extended Multichannel EEG Features

It was established in Sections 2.2.2 and 3.1, pages 16 and 43, that background EEG is a stochastic process, which produces independent random behaviours across all EEG channels. On the other hand, seizure events are not stochastic and their channel behaviour depends on neighbouring channels. Consequently, features that describe the

multichannel information alikeness can be used to discriminate background from seizure events. Information alikeness can be quantified in terms of correlation, or joint entropy. In this thesis, correlation is adapted as the information alikeness metric.

Pearson's correlation equation can be extended to be used with time-frequency distributions. The extension is expressed in Equation 3.4 and Equation 3.5, where Q is the number of channels, and $\mu_{z_i z_i}$ is the mean value of the auto-TFD $\rho_{z_i z_i}[n,k]$. The Pearson's correlation forms a symmetric matrix **R** containing; multichannel auto-correlations on the diagonal, and multichannel cross-correlations off the diagonal (Equation 3.6). The auto-correlations do not contain information concerning the multichannel information alikeness, thus they must be removed. Furthermore, since **R** is symmetric, only the lower or upper triangle of the matrix has to be computed in order to minimise the computation requirements (the upper triangle of the matrix **R** is highlighted in red). The extracted upper or lower triangle is then vectorised into a one-dimensional array r of length P (Equation 3.7). This array contains the pure information concerning the multichannel alikeness, thus the statistical features described in Appendix C in page 173 can be used to create extended multichannel features.

The extended multichannel EEG features are expressed in Equation 3.8, Equation 3.9, Equation 3.10, Equation 3.11, and Equation 3.12. The feature selection and classification iterative process described in Section 4.2.2, has to be repeated twice considering only the multichannel fused features and then considering the combination of the extended multichannel features and the fused multichannel features (Figure 4.4). The maximum number of selected features in this case is not 21, but has to be maintained as 16. This is crucial, so a comparison between the two classification profiles can be made.

$$R_{ij} = \frac{\sum_{k=1}^{M}\sum_{n=1}^{N}\left(\rho_{z_i z_i}[n,k] - \mu_{z_i z_i}\right)\left(\rho_{z_j z_j}[n,k] - \mu_{z_j z_j}\right)}{\sqrt{\sum_{k=1}^{M}\sum_{n=1}^{N}\left(\rho_{z_i z_i}[n,k] - \mu_{z_i z_i}\right)^2}\sqrt{\sum_{k=1}^{M}\sum_{n=1}^{N}\left(\rho_{z_j z_j}[n,k] - \mu_{z_j z_j}\right)^2}}$$ Equation 3.4

$i \in [1, Q]\ \ j \in [1, Q]$

$$\mu_{zz} = \frac{1}{NM}\sum_{k=1}^{M}\sum_{n=1}^{N}\rho_{zz}[n,k]$$ Equation 3.5

$$\mathbf{R} = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1Q} \\ R_{21} & R_{22} & \cdots & R_{2Q} \\ \vdots & \vdots & \ddots & \vdots \\ R_{Q1} & R_{Q2} & \cdots & R_{QQ} \end{bmatrix}$$ Equation 3.6

$$r_{1\times P} = R_{ij}\ ,\qquad j > i\ ,\qquad P = \frac{Q^2 - Q}{2}$$ Equation 3.7

- Mean:

$$M_1 = \ \mu = \frac{1}{P}\sum_{k=1}^{P} r_k$$ Equation 3.8

- Variance:

$$M_2 = \ \sigma^2 = \frac{1}{P}\sum_{k=1}^{P}(\mu - r_k)^2$$ Equation 3.9

- Skewness:

$$M_3 = \frac{1}{(P-1)\sigma^3}\sum_{k=1}^{P}(r_k - \mu)^3$$ Equation 3.10

- Kurtosis:

$$M_4 = \frac{1}{(P-1)\sigma^4}\sum_{k=1}^{P}(r_k - \mu)^4$$ Equation 3.11

- Coefficient of Variation:

$$M_5 = \frac{\sigma}{\mu}$$ Equation 3.12

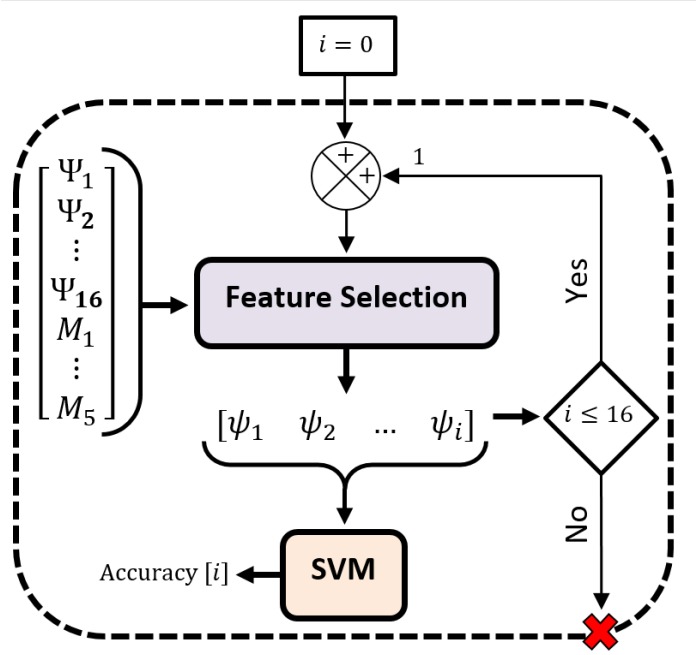

Figure 4.4: Feature selection and classification iterative process including the extended multichannel features (The maximum number of selected features is 16)

4.2.4 Abnormalities Detection Evaluation

The abnormality detection evaluation procedure produces an accuracy profile for the Leave-one-out cross-validated (*LOOCV*) SVM classifier (more details on Leave-one-out cross-validation can be found in Section 2.5.4.2, page 42). The SVM kernel and parameters are kept constant during the process to quantify the relative change in performance when utilising; the artifact detection and removal stage, and the extended multichannel features. The process consists of the following steps:

1. Repeat the first 3 steps of the artifact removal evaluation (Section 4.1.4, page 80).
2. Extract time-frequency features from every single channel in every segment for the clean, corrupted, and estimated multichannel EEG waveforms.
3. Fuse the extracted features using MFFSB to form the *fused multichannel time-frequency features*.

4. Extract the *extended multichannel time-frequency features* from every segment for the clean, corrupted, and estimated multichannel EEG waveforms, and then combine them with the *fused multichannel time-frequency features*.
5. Compute Fisher scores for every fused and combined multichannel features of the clean EEG, and sort them in descending order.
6. Apply the feature selection and classification iterative process on the clean EEG fused and combined ordered features.
7. Cross validate the SVM classifier using LOOCV in every iteration to produce a validated accuracy profile for the *fused and combined multichannel time-frequency features*.

Summing up, the synthetic neonatal multichannel EEG model (constructed in Chapter 3) is used to produce clean normal and abnormal EEG patterns, thus establishing the control setup in this thesis. Furthermore, the experimental setup is established by studying the effect of four independent variables on the performance of the neonatal seizure detection system. The independent variables are summarised as:

- Artifact relative power: the synthetic neonatal multichannel artifacts model (Chapter 3) is used to produce corrupted EEG with different signal-to-artifact ratios.
- Artifact detection and removal: Each of the corrupted multichannel waveforms is pre-processed by the optimised automated time and time-frequency artifact detection and removal systems.
- Extended features: proposed features are extracted from the clean, corrupted, and pre-processed multichannel waveforms.
- Number of utilised features: performance curves of the neonatal seizure detection system are generated with respect to the number of utilised features.

CHAPTER 5: RESULTS AND DISCUSSIONS

In this chapter, the synthetic neonatal multichannel EEG results are illustrated to present the clean multichannel waveform. The model results consist of the propagation model validation outcomes so as to choose the best model that fits the reference distribution along with its results. Moreover, validation results for the synthetic neonatal multichannel EEG model are presented to quantify its effectiveness in mimicking real neonatal EEG. After that, the synthetic neonatal multichannel EEG artifacts model results are illustrated in terms of how they contaminate the clean waveform. Results include the synthetic EEG artifacts outcomes, the multichannel artifact waveforms, and the contaminated neonatal multichannel EEG. Furthermore, results of the automated artifact detection and removal system using BSS in the time and time-frequency domains are presented in terms of how they filter the contaminated EEG signals. Artifact detection and removal performances are evaluated and discussed with respect to different signal-to-artifact ratios to quantify the quality of the system. Finally, the automated EEG abnormality detection system results are presented to assess the enhancements in seizure detection rates when utilising; the artifact detection and removal system, and the extended multichannel features.

5.1 Synthetic Neonatal Multichannel EEG Model Outputs

5.1.1 EEG Propagation Model Validation Results

EEG propagation models were validated against the reference SBR range. All models output signals were amplified to produce an approximated mean SBR of 15 so they can be compared with the reference distribution. Model A output was amplified by 11.33 to produce a mean SBR of 14.8, while models B and C were amplified by

400e6 and 600e6 respectively to produce mean SBRs of 14.83 and 14.26. Models B and C amplification factors are relatively large when compared to model A, but they are essential as scattering coefficients were large enough to attenuate the seizure signal picked up by the scalp electrodes. Figure 5.1 illustrates the output SBR distributions of all models against the reference distribution. Model A has the highest likelihood to be chosen as "The EEG Propagation Model", because the P-P plot emphasises on the fact that Model A (pure optical absorber) is the best candidate to be the propagation model as it is the best in fitting the reference distribution behaviour when calculating the total probability (Figure 5.2). Moreover, Model A normal line has an R^2 value of 0.99999036 when fitting the reference normal line, thus proving its goodness of fit. Further simulations and testing will be carried out using only Model A. P-P plot with linear SBR axis is depicted in the supplementary material shown in Appendix J.1, page 212.

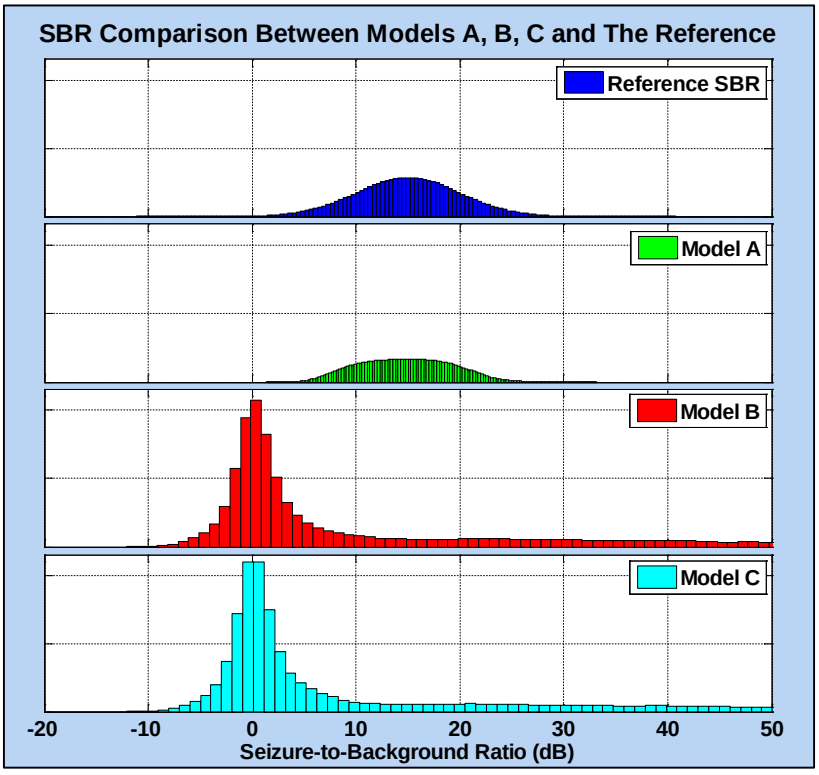

Figure 5.1: SBR distributions of the reference and models A, B, and C outputs. Note that amplitudes of all plots have the same range

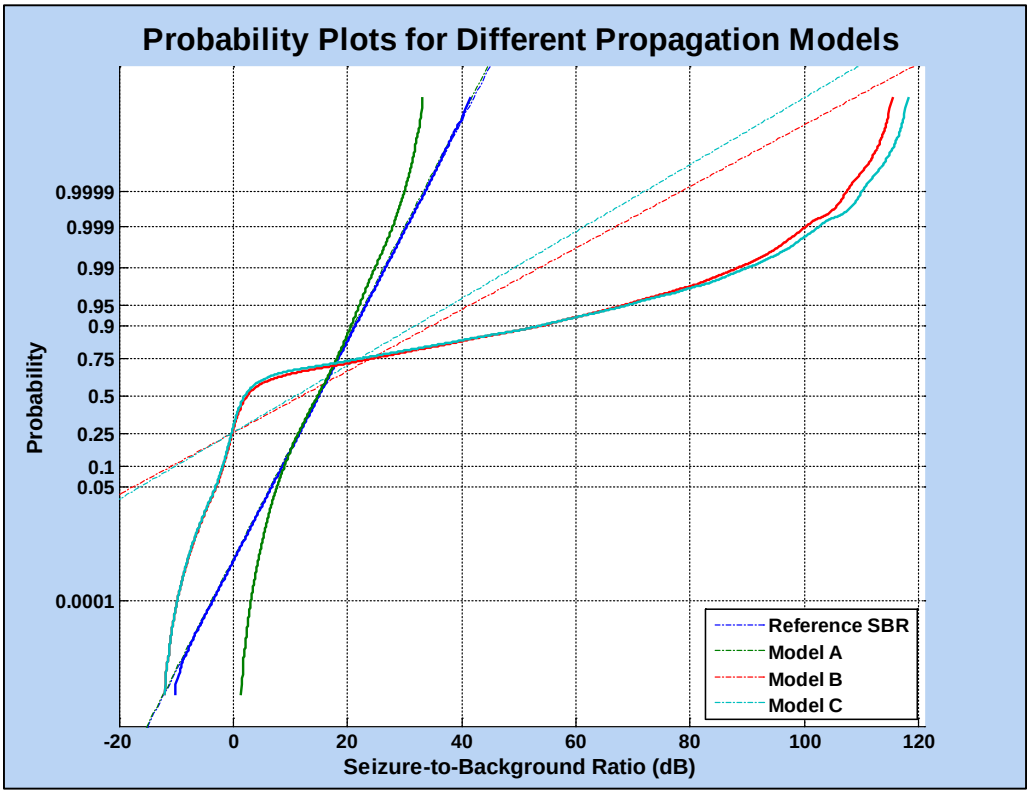

Figure 5.2: P-P plot of models A, B, and C output (Logarithmic) SBR against the reference distribution

5.1.2 Neonatal EEG Propagation Model Output

The EEG propagation model simulations are carried out using only Model A with a seizure source location defined by a radius 4.5, an azimuth angle 180 degrees, and an elevation angle of 45 degrees.

5.1.2.1 The Space-Propagation Model

Figure 5.3 illustrates the seizure event dispersion in the neonatal four sphere head model reaching the scalp electrodes. Closer electrodes have relatively higher reception power than further ones. The space model output illustrates the change in the seizure relative amplitude with respect to the electrodes location (Figure 5.4). Note that Figure 5.4 does not show the time delay in every channel as it only illustrates the change in the relative amplitude, so all seizure patterns start and end together.

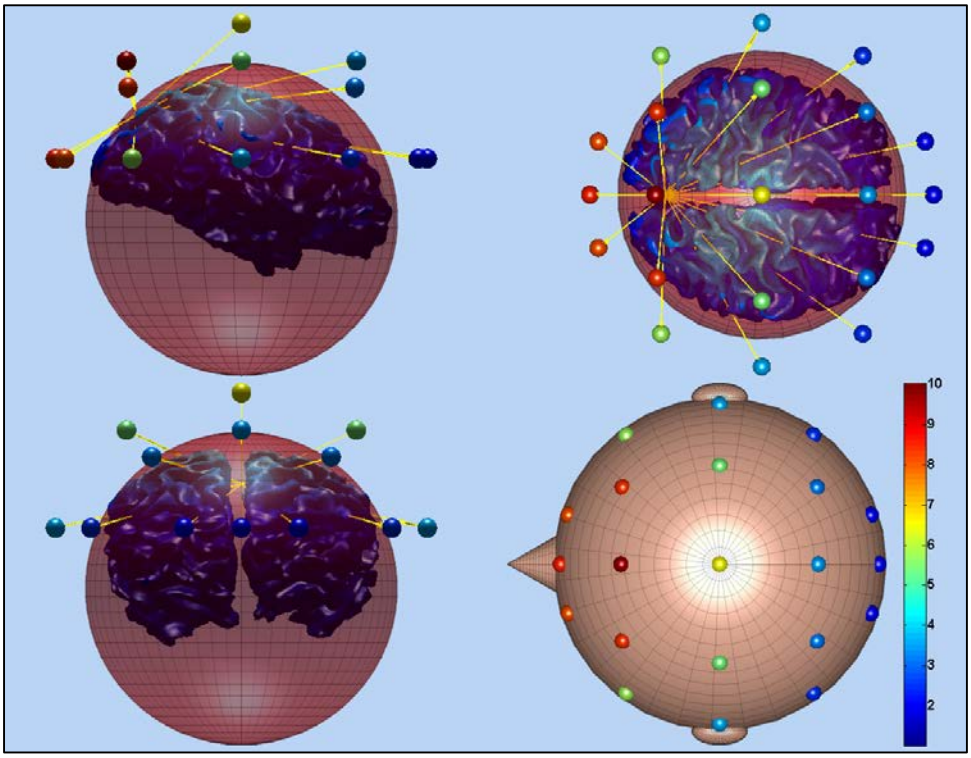

Figure 5.3: The seizure event dispersion in the neonatal four sphere head model

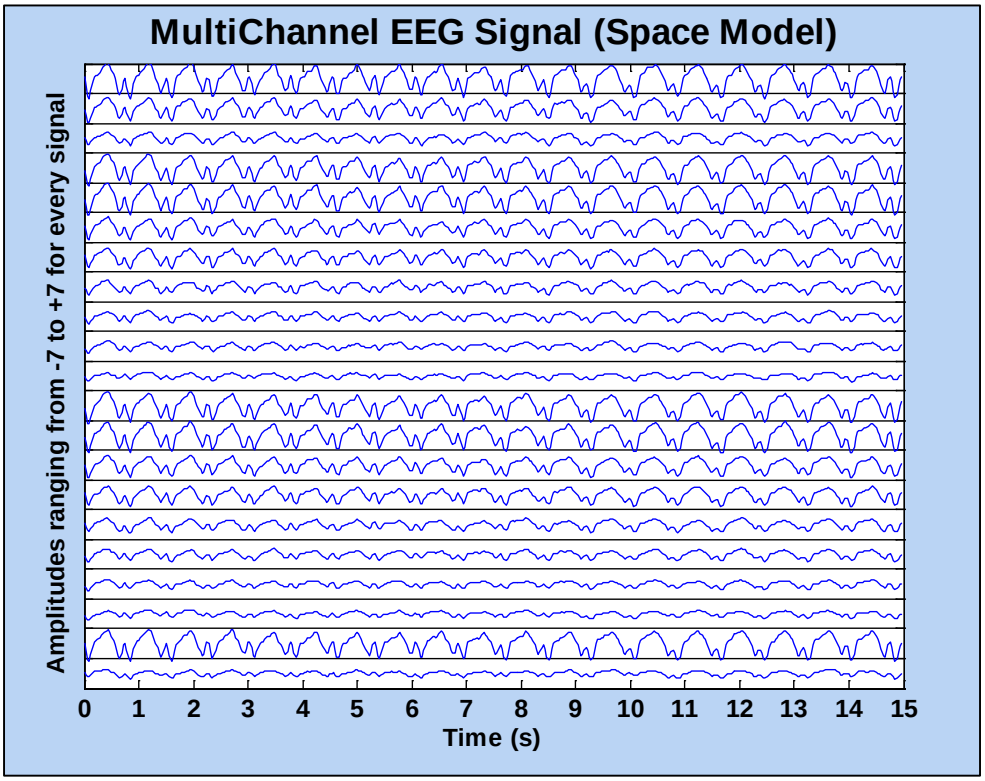

Figure 5.4: The multichannel EEG space-propagation model output

5.1.2.2 The Time-Propagation Model

Figure 5.5 illustrates every channel delay in receiving the seizure pattern. Delays depend on the signal total path length, as farther electrodes result in relatively larger time delays. Note that all channel amplitudes are normalised to illustrate the differences in the reception relative time between the channels. The beginning and ending of the seizure pattern in every channel are depicted using green and red vertical lines respectively.

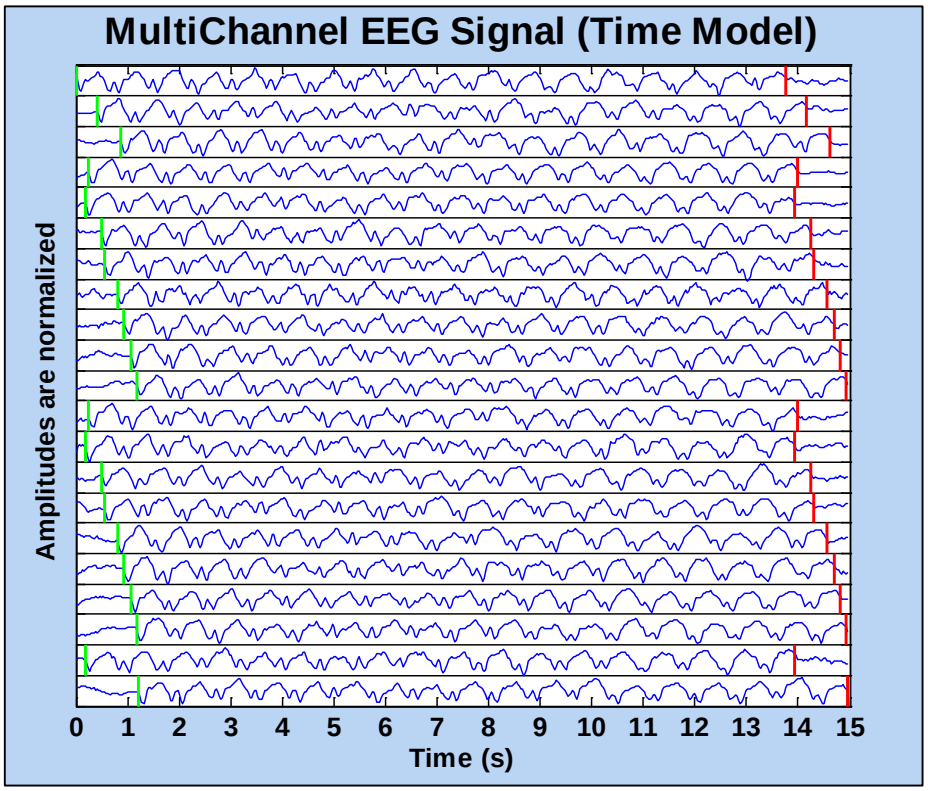

Figure 5.5: The multichannel EEG time-propagation model output (Green lines represent the beginning of the seizure pattern, while red lines represent the end of the seizure pattern)

5.1.2.3 The Time-Space Propagation Model

Table 5.1 and Figure 5.6 illustrate the output when combining the time and space models into a time-space propagation model. Results show the relative amplitudes and delays the seizure pattern would experience on every channel placed on the neonatal

four sphere head model. Electrode F_z has the highest relative amplitude and the lowest relative delay, as it is the closest electrode to the seizure source location, while in contrast O_z has the lowest relative amplitude and the highest relative delay. Note that the relative amplitudes and delays have some sort of symmetry across the channels. This is a consequence of the perfect spheres representing the neonatal head regions, as such shapes produce symmetry in modelling results.

Table 5.1: The multichannel EEG time-space propagation model output

i	Electrode Label	Relative Amplitude (V)	Relative Delay (s)
1	F_z	9.689199655	0
2	C_z	6.547759961	0.40625
3	P_z	3.391266637	0.875
4	F_{p1}	8.124688322	0.25
5	F_3	8.430654536	0.1875
6	F_7	5.729977323	0.5
7	C_3	5.413214954	0.5625
8	T_3	3.602627909	0.8125
9	P_3	3.118237446	0.9375
10	T_5	2.54271105	1.0625
11	O_1	2.118622305	1.1875
12	F_{p2}	8.124688322	0.25
13	F_4	8.430654536	0.1875
14	F_8	5.729977323	0.5
15	C_4	5.413214954	0.5625
16	T_4	3.602627909	0.8125
17	P_4	3.118237446	0.9375
18	T_6	2.54271105	1.0625
19	O_2	2.118622305	1.1875
20	F_{pz}	8.486164054	0.1875
21	O_z	2.070078352	1.21875

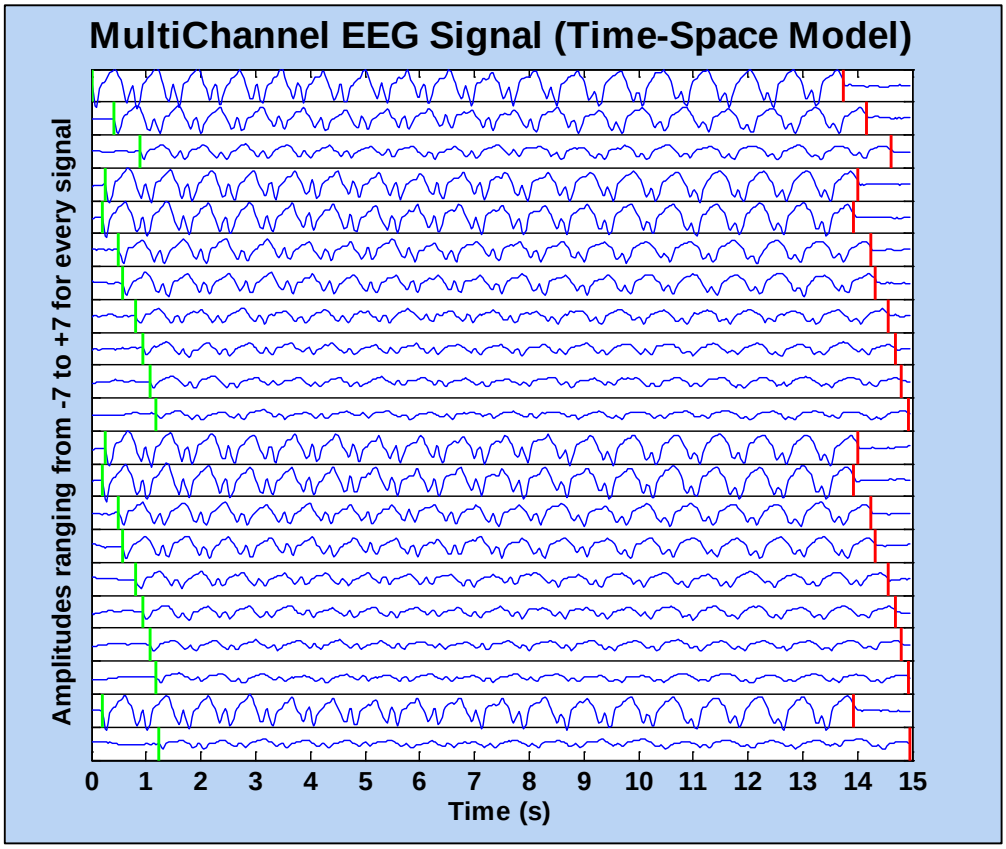

Figure 5.6: The multichannel EEG time-space propagation model output (Green lines represent the beginning of the seizure pattern, while red lines represent the end of the seizure pattern)

5.1.3 Neonatal Multichannel EEG Model Output

Figure 5.7 depicts the output of the synthetic neonatal multichannel EEG model. It shows the multichannel waveform by using an EEG operation mode of 3 (alternating between background and seizure epochs). Every channel is constructed by 8 segments, where 4 of them are background epochs and the others are seizure epochs. Every segment has a duration of 15 seconds making the total waveform to have a duration of 120 seconds. Seizure patterns are highlighted with light red background, while background patterns are not.

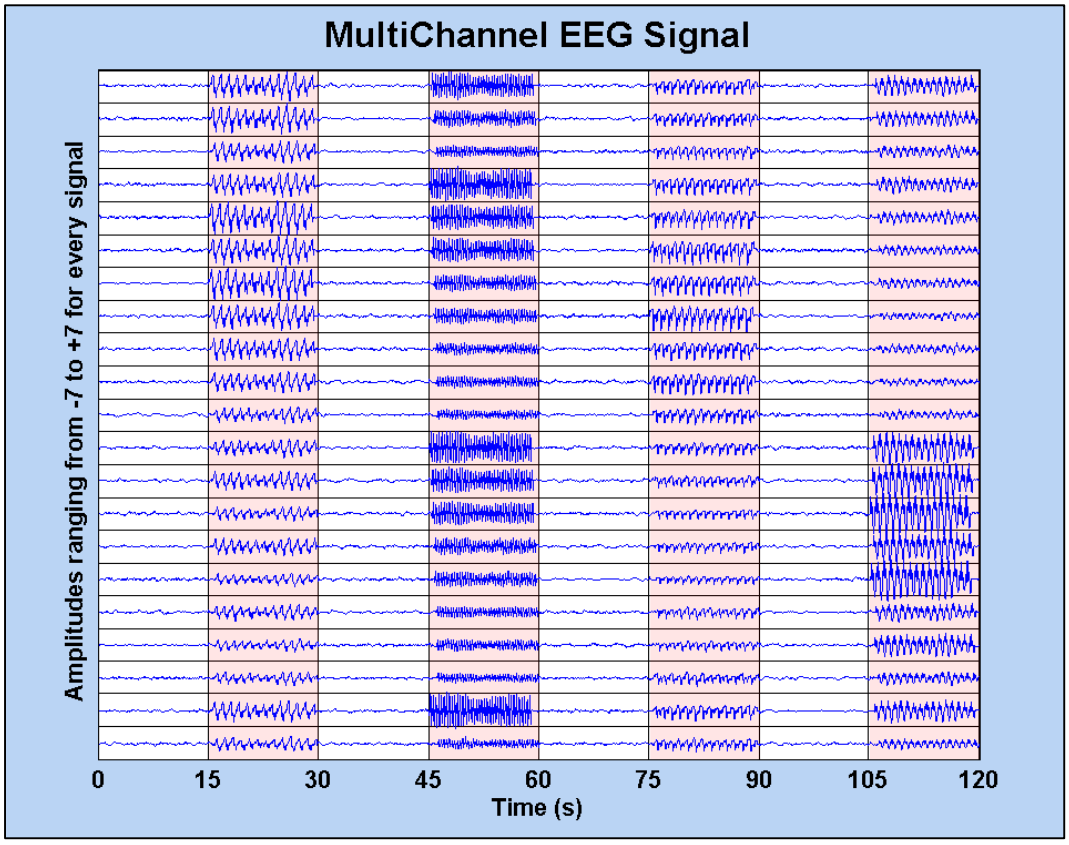

Figure 5.7: The neonatal multichannel EEG model output (Seizure segments are highlighted with light red background)

5.1.4 Neonatal Multichannel EEG Model Validation Results

Validation of the neonatal multichannel EEG model was done by comparing the synthetic multichannel generated waveforms and the real multichannel neonatal EEG (described in Section 3.1, page 43) using Spearman's correlation. A selection of 201 background and 80 seizure annotated segments were extracted from the real neonatal EEG database, and then correlated with 100 synthetic multichannel background and seizure segments.

Figure 5.8 illustrates the correlation distributions for the background and seizure EEG comparisons. Time, frequency, and time-frequency correlation distributions are depicted in blue, red, and green respectively. Comparisons for the background and

seizure segments in the time domain, illustrate negligible mean correlations of 0.1584 and 0.32923 respectively. However, the "EEG Corruption Evaluation" study in Appendix H, page 188 showed that background and seizure segments are weakly correlated among themselves in the time domain (Table A.4). On the other hand, comparisons in the frequency and time-frequency domains illustrate high mean correlations of 0.58065 and 0.57796 for the background EEG comparison, and 0.65107 and 0.66254 for the seizure EEG comparison respectively. These high correlation results illustrate the effectiveness of the synthetic neonatal multichannel EEG model in mimicking the morphology of real neonatal EEG in the frequency and the time-frequency domains.

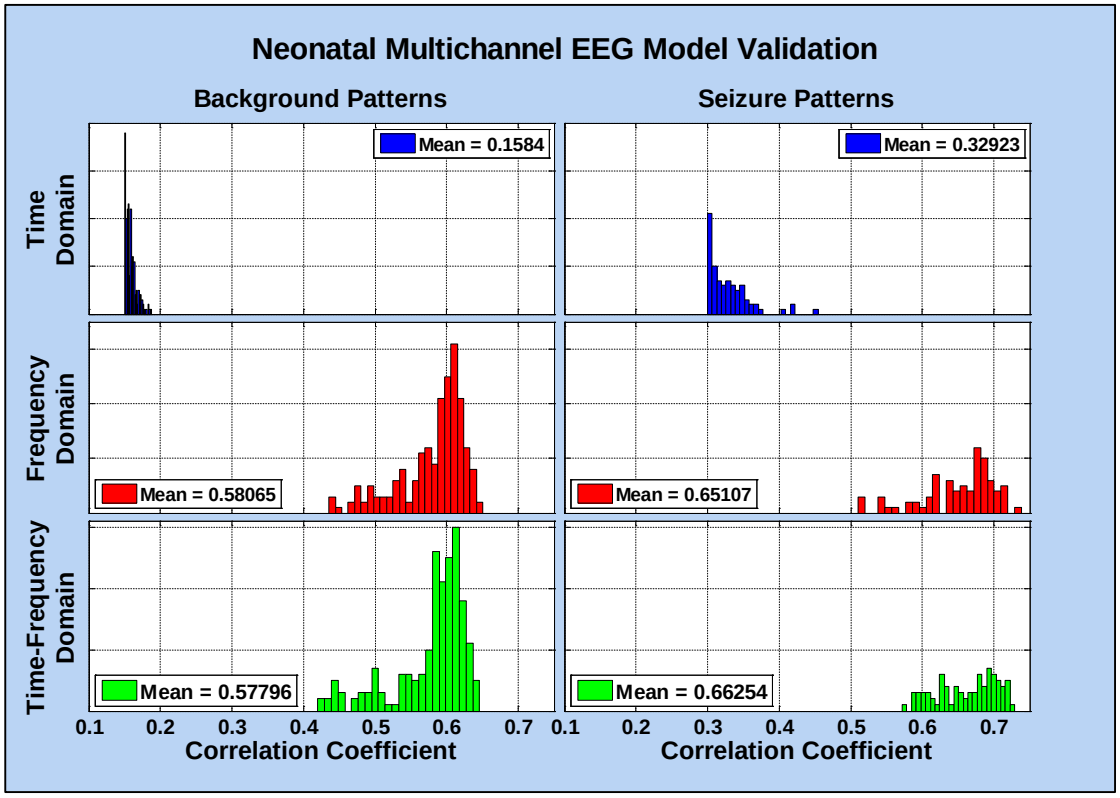

Figure 5.8: The neonatal multichannel EEG model validation results

5.2 Synthetic Neonatal Multichannel EEG Artifacts Model Outputs

5.2.1 Neonatal EEG Artifacts Model Output

The neonatal EEG artifacts model produces single-channel EEG artifacts, consisting of BVP, ECGS, and STHA synthetic signals. All produced segments have a duration of 15 seconds and a sampling frequency of 32 Hz.

5.2.1.1 Blood Vessel Pulsation

Figure 5.9 illustrates the behaviour of the synthetic BVP artifact when Gaussian noise is added. It contains the simulated artifact in time, frequency, and time-frequency domains, along with the time domain histogram. Furthermore, it shows the Gaussian behaviour and the noisy nature of the BVP artifact when picked up by EEG electrodes. The initial non-Gaussian simulated BVP artifact is attached in the supplementary material shown in Appendix J.2, page 213.

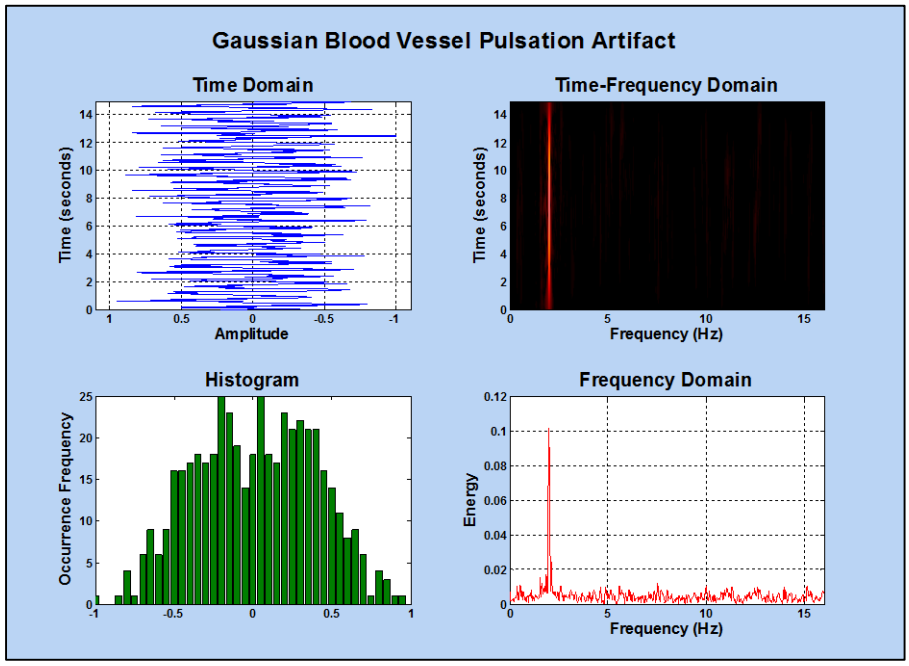

Figure 5.9: Simulated Gaussian Blood Vessel Artifact. TFD parameters (MB, N = 480, Fs = 32 Hz, FFT_N = 1024, Lag Window = 479, Alpha = 0.01, time resolution = 1). Time and frequency steps are (0.0313, 0.0313) respectively

5.2.1.2 Electrocardiogram Spikes

Figure 5.10 illustrates the behaviour of the synthetic ECGS artifact when Gaussian noise is added. It contains the simulated artifact in time, frequency, and time-frequency domains, along with the time signal histogram. Moreover, it shows the Gaussian behaviour and the noisy nature of the ECGS artifact when picked up by EEG electrodes. The initial non-Gaussian simulated ECGS artifact is attached in the supplementary material shown in Appendix J.2, page 213.

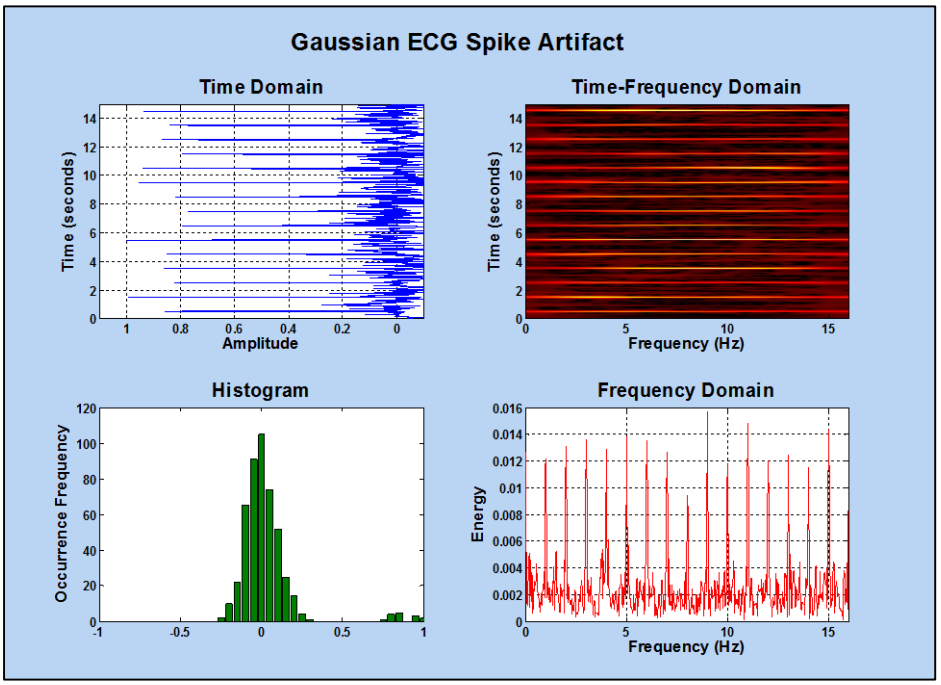

Figure 5.10: Simulated Gaussian ECG Spike artifact. Simulated ECG Spike artifact. TFD parameters (EMB, N = 480, Fs = 32 Hz, FFT_N = 1024, Lag Window = 479, time resolution = 1, Alpha = 0.99, Beta = 0.01). Time and frequency steps are (0.0313, 0.0313) respectively

5.2.1.3 Short-Time High-Amplitude

The simulated STHA artifact is illustrated in Figure 5.11. It contains the synthetic artifact in time, frequency, and time-frequency domains, along with the time signal histogram. The histogram shows the heavy-tailed distribution of such artifact, which

fits a Levy Distribution. LSD toolbox has been used to simulate Levy stable symmetrical stochastic process. More details on the toolbox can be found in [105].

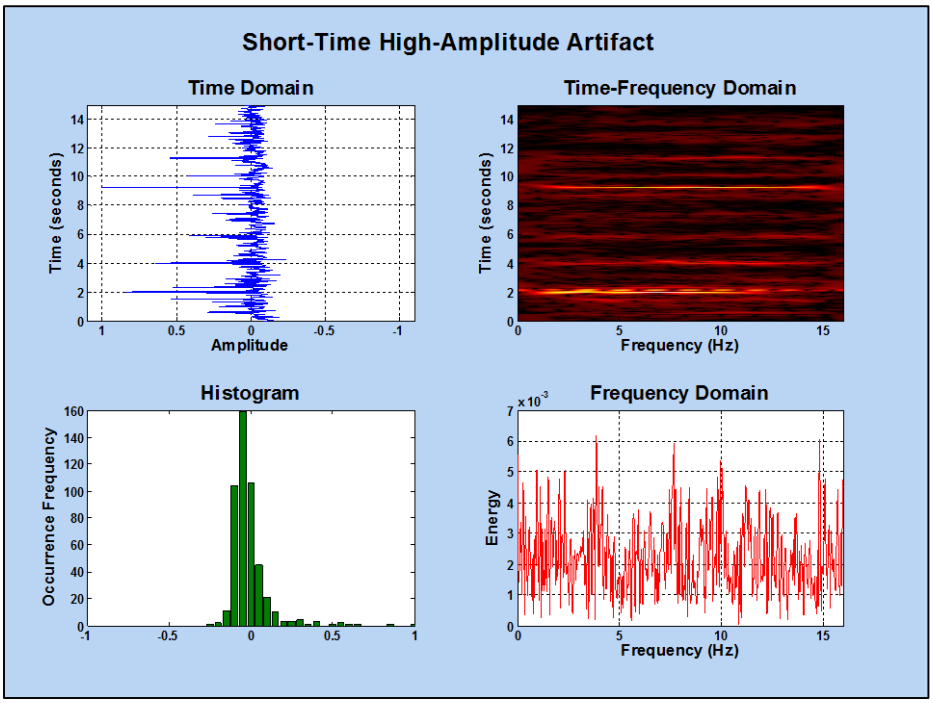

Figure 5.11: Simulated STHA artifact. TFD parameters (MB, N = 480, Fs = 32 Hz, FFT_N = 1024, Lag Window = 11, Alpha = 0.99, time resolution = 1). Time and frequency steps are (0.0313, 0.0313) respectively

5.2.2 Neonatal Multichannel EEG Artifacts Model Output

The neonatal multichannel EEG artifacts model is capable of producing different multichannel artifact patterns to contaminate the clean neonatal multichannel EEG.

5.2.2.1 Multichannel Clean EEG

Figure 5.12 illustrates the neonatal multichannel clean EEG. It was produced using an EEG operation mode of 3. Every channel is constructed from 6 segments, where 3 of them are background epochs and the other half are seizure epochs. Every segment has a duration of 15 seconds making the total waveform to have a duration of 90 seconds. Seizure patterns are highlighted with light red background, while background patterns are not. Figure 5.13 illustrates the first channel of the multichannel clean EEG output

in time, frequency, and time-frequency domains. A diagnosis mask is attached to illustrate the alternation between background and seizure epochs.

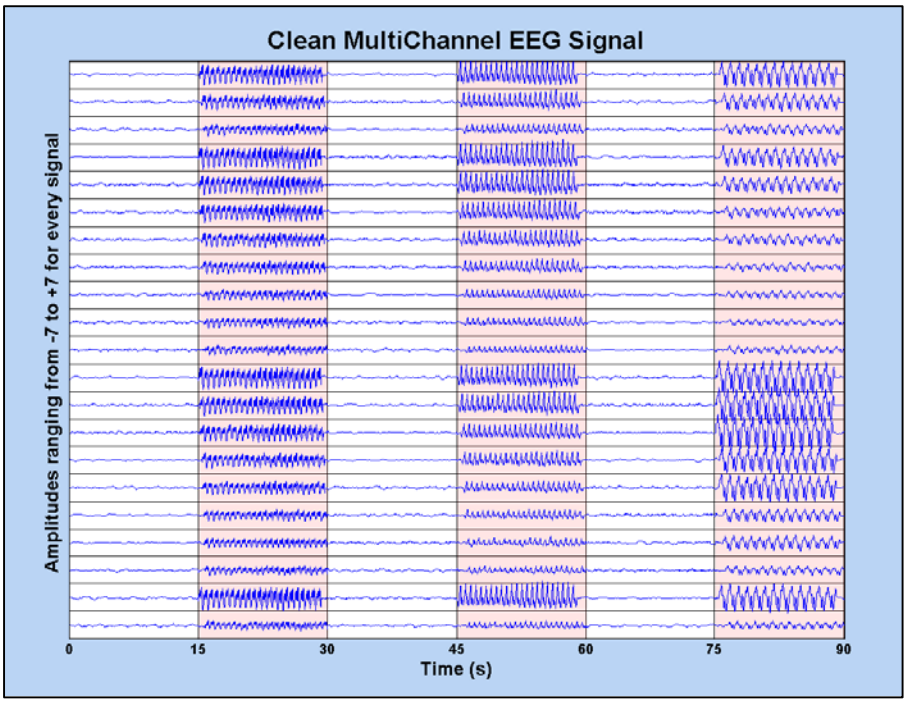

Figure 5.12: The neonatal multichannel clean EEG (EEG operation mode of 3). Seizure segments are highlighted with light red background

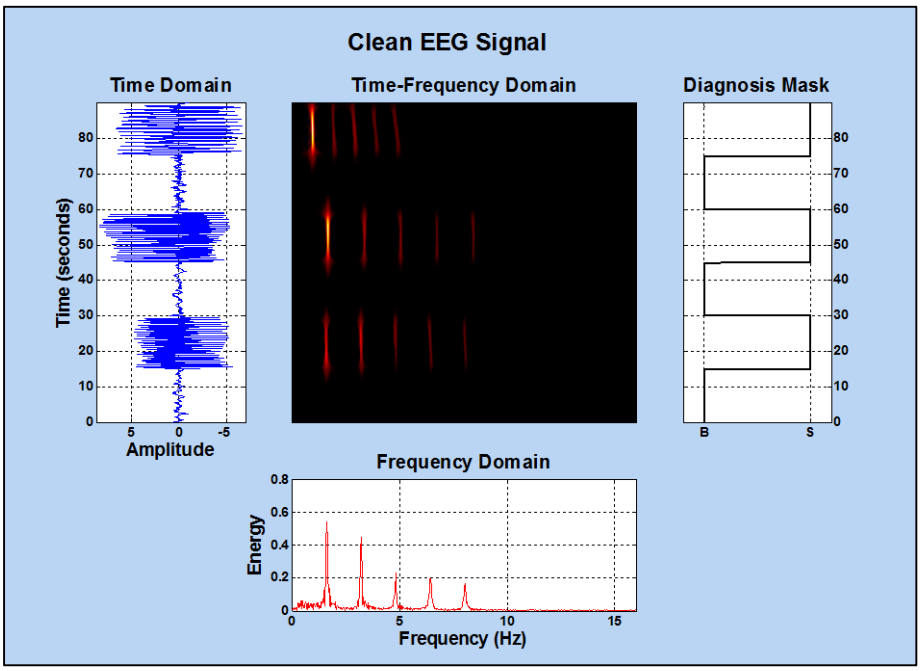

Figure 5.13: The neonatal multichannel clean EEG (First Channel). TFD parameters (EMB, N = 2880, Fs = 32 Hz, FFT_N = 1024, Lag Window = 2879, Alpha = 0.01, Beta = 0.9, time resolution = 1). Time and frequency steps are (0.0313, 0.0313) respectively. S: Seizure, B: Background

5.2.2.2 Multichannel EEG Artifact

Figure 5.14 illustrates the neonatal multichannel EEG artifact. It was produced using an artifact operation mode of 3 (artifacts exist on the first half of the signal length). Every channel is constructed by 6 segments, where the first 3 of them are artifacts, while the rest are empty. Every segment has a duration of 15 seconds making the total waveform to have a duration of 90 seconds. Artifacts patterns are highlighted with light grey background, while clean patterns are not. Figure 5.15 illustrates the first channel of the multichannel EEG artifact output in time, frequency, and time-frequency domains. An artifact mask is attached to illustrate the location of the artifact patterns in the waveform.

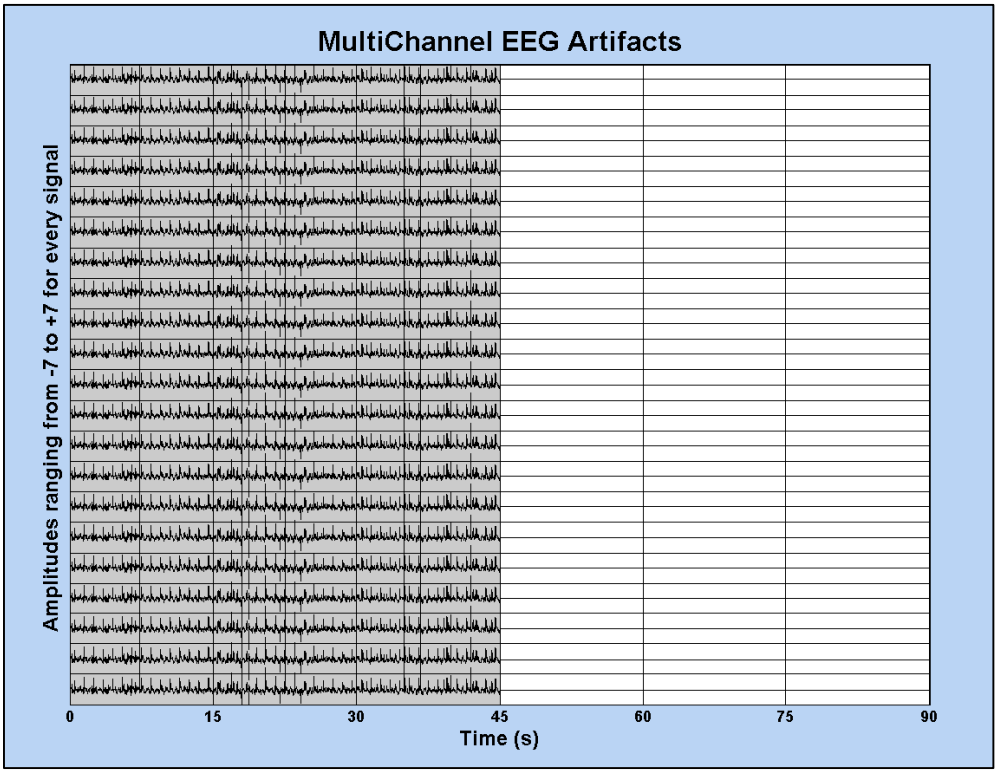

Figure 5.14: The neonatal multichannel EEG Artifact (Artifact operation mode of 3)

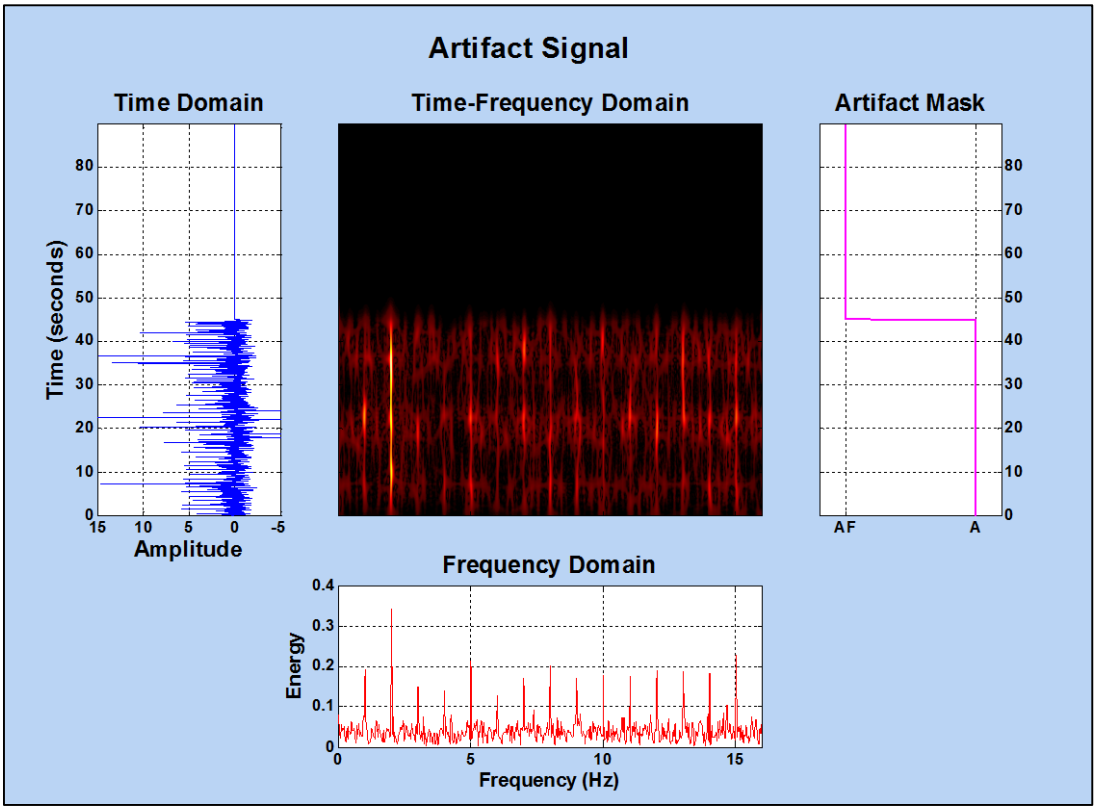

Figure 5.15: The neonatal multichannel EEG Artifact (First Channel). TFD parameters (EMB, N = 2880, Fs = 32 Hz, FFT_N = 1024, Lag Window = 2879, Alpha = 0.01, Beta = 0.9, time resolution = 1). Time and frequency steps are (0.0313, 0.0313) respectively. A: Artifact, AF: Artifact Free

5.2.2.3 Multichannel Corrupted EEG

Figure 5.16 depicts the neonatal multichannel corrupted EEG. It was produced using an EEG operation mode of 3 and an artifact operation mode of 3. Every channel is constructed from 6 segments, where 3 of them are background, and 3 of them are seizure epochs. 3 artifact segments are imposed on the first 3 segments, while the rest are clean. Every segment has a duration of 15 seconds making the total waveform to have a duration of 90 seconds. Figure 5.17 illustrates the first channel of the multichannel corrupted EEG output in time, frequency, and time-frequency domains. Artifact and diagnosis masks are attached to illustrate the location of the artifact and seizure patterns in the waveform.

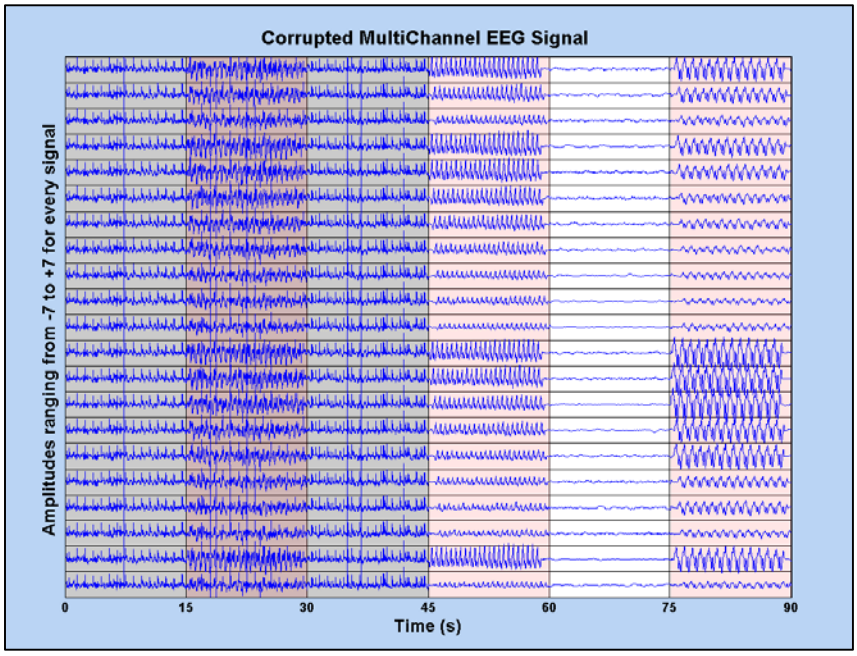

Figure 5.16: The neonatal multichannel corrupted EEG (EEG operation mode of 3, Artifact operation mode of 3). Seizure segments are highlighted with light red background and artifact segments are highlighted with light grey background

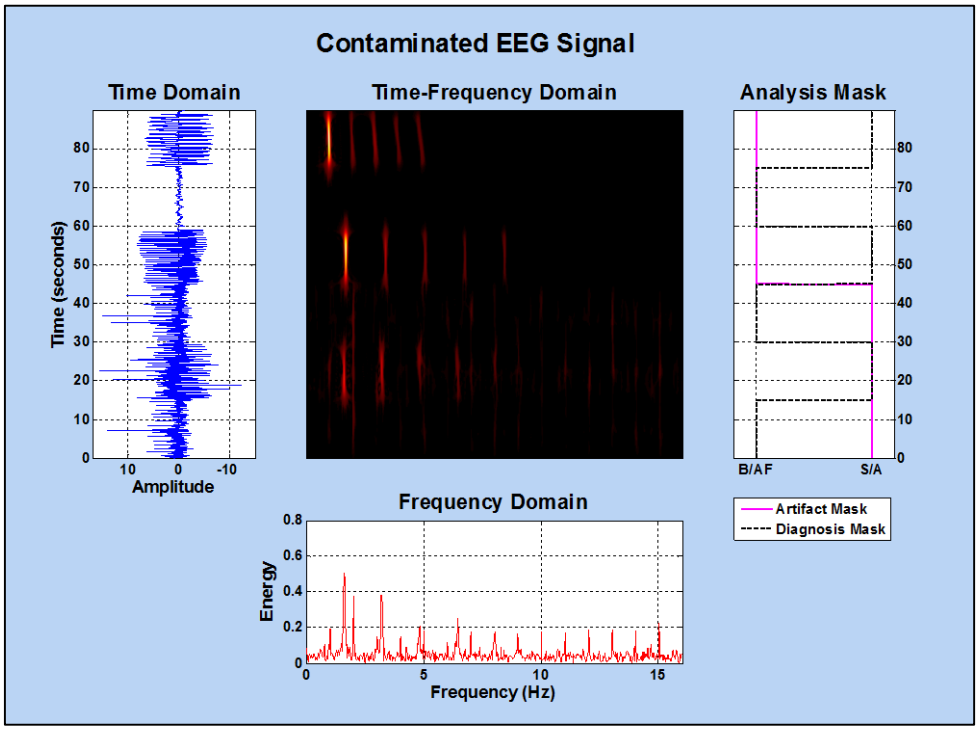

Figure 5.17: The neonatal multichannel corrupted EEG (First Channel). TFD parameters (EMB, N = 2880, Fs = 32 Hz, FFT_N = 1024, Lag Window = 2879, Alpha = 0.01, Beta = 0.9, time resolution = 1). Time and frequency steps are (0.0313, 0.0313) respectively. A: Artifact, AF: Artifact Free, B: Background, S: Seizure

5.3 Automatic Artifacts Detection and Removal

The automatic artifacts detection and removal system output illustrate the detection and removal performance results for the time and time-frequency BSS algorithms using different signal-to-artifact ratios (*SAR*). Corrupted neonatal multichannel EEG signals are generated using an EEG and artifact operation modes of 3 and 2 to test the time and time-frequency artifact detectors. Five versions of the corrupted signals are created by altering the artifact factor to be 0.4355, 1, 2, 3, and 4, thus producing corrupted waveforms with SAR values of 0, -7.2206, -13.2412, -16.763, and -19.262 dB respectively. This is done to assess the robustness and accuracy of the artifact detection and removal system when applied to different EEG contamination levels.

5.3.1 Artifact Detection Performance

Artifact detection performances for the time and time-frequency BSS algorithms are maximised by statistically optimising the correlation thresholds between the ICs and the artifact reference signal (the complete detection optimisation results are attached in the supplementary material shown in Appendix J.3, page 214 because of their quantity).

The time and time-frequency domain BSS artifact detection performance results are illustrated in Figure 5.18. Results show that time domain JADE has the highest detection rate when compared to the other examined time domain BSS algorithms. It illustrates maintained 97% accuracy, 94% sensitivity, and 100% specificity through the different SAR levels. This makes time domain JADE the most robust time domain algorithm with respect to its detection capabilities. Fast-ICA and BSS-CCA algorithms show a slight decrement in their detection rate (maximum differences in accuracies are

3% and 1% respectively). They show 95% and 91.6% mean accuracies, 91.6% and 85.2% mean sensitivities, and lastly 98.4% and 98% mean specificities. Furthermore, SOBI shows an increment in its detection rate due the prominence of the artifact signal (maximum difference in accuracy is 4%). It illustrates 92.2% mean accuracy, 86.4% mean sensitivity, and 98% mean specificity. Lastly, Robust-ICA illustrates a maintained 95% accuracy, 92% sensitivity, and 98% specificity through the different SAR levels.

On the other hand, time-frequency domain artifact detection results show that none of the BSS algorithms maintained its detection rate through the different SAR levels. Time-frequency SOBI produces the highest detection rate when compared to the other examined time-frequency BSS algorithms. It illustrates 90.4% mean accuracy, 86.8% mean sensitivity, and 94% mean specificity through the different SAR levels (maximum difference in accuracy is 3%). This makes SOBI the most robust time-frequency domain algorithm with respect to its detection capabilities. Other BSS algorithms show an overall decrement in their detection rates. Fast-ICA, Robust-ICA and BSS-CCA algorithms show 88%, 86.8%, and 86.6% mean accuracies, 82.4%, 82%, and 81.6% mean sensitivities, and lastly 93.6%, 91.6% and 91.6% mean specificities (maximum differences in accuracies are 3%, 2%, and 3% respectively). Furthermore, JADE algorithm has the least detection rate consisting of: 85.8% mean accuracy, 81.6% mean sensitivity, and 90% mean specificity (maximum difference in accuracy is 5%).

Summing up, all time and time-frequency artifact detection algorithms can be considered robust through the different SAR levels. This is valid as all algorithms have maintained a maximum difference in their detection accuracies below 5%, thus no prior knowledge on the SAR level is required for such algorithms. In addition, time domain

artifact detection algorithms produce higher overall detection performance when compared to the time-frequency BSS, but SOBI is the exception. It maintains the same level of performance when used in the time and time-frequency domains, thus making it reliable and independent from the domain of analysis.

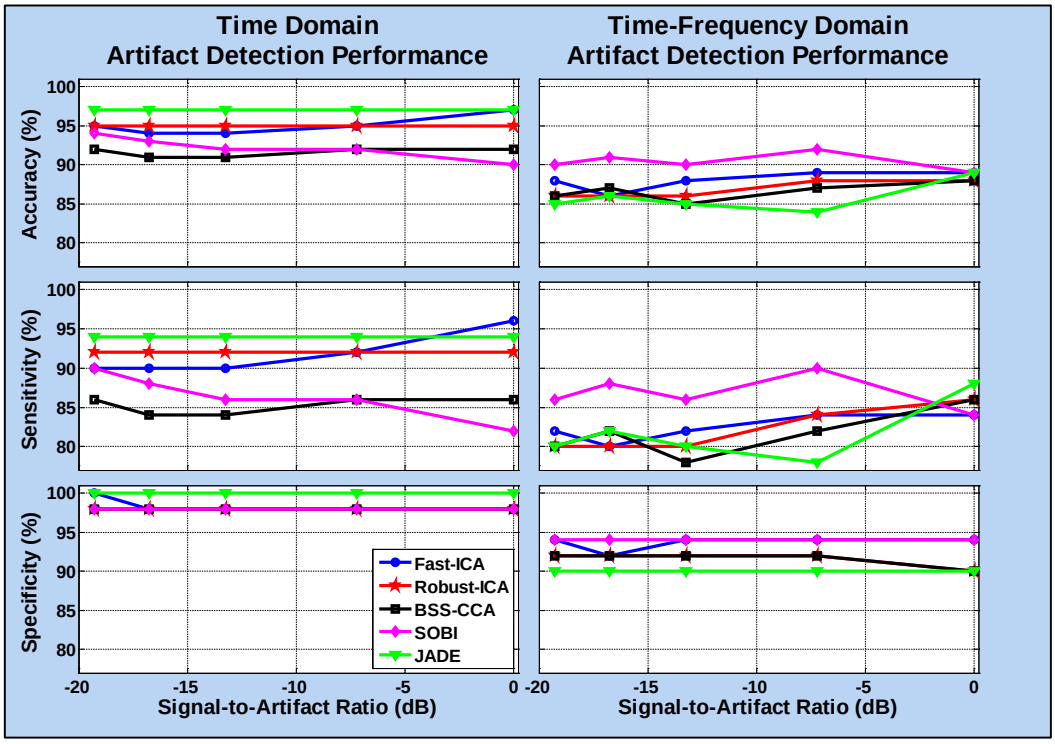

Figure 5.18: Time and time-frequency artifact detection performance assessment through different signal-to-artifact ratios

5.3.2 Artifact Removal Performance

Artifact removal performance is evaluated using the NRMSE and PCC metrics in the time-frequency domain. This evaluation procedure is done for all the time and time-frequency BSS algorithms.

The time and time-frequency domain BSS artifact removal evaluation results are illustrated in Figure 5.19. Results show that time domain SOBI algorithm has the highest removal performance when compared to the other examined time domain BSS algorithms. It illustrates 51.3% mean NRMSE and 79.5% mean PCC through the

different SAR levels. This makes it the best time domain algorithm with respect to its artifact removal capabilities *(*maximum differences in NRMSE and PCC are 42.8% and 26.9% respectively). On the other hand, time domain Robust-ICA illustrates the lowest removal performance when compared to the other examined time domain BSS algorithms. It illustrates 56.8% mean NRMSE and 74.2% mean PCC through the different SAR levels. This makes it the worst time domain algorithm with respect to its artifact removal capabilities (maximum differences in NRMSE and PCC are 47.9% and 29.7% respectively).

Furthermore, time-frequency JADE algorithm shows the highest removal performance when compared to the other examined time-frequency BSS algorithms. It illustrates 70.3% mean NRMSE and 52.3% mean PCC through the different SAR levels. This makes it the best time-frequency domain algorithm with respect to its artifact removal capabilities (maximum differences in NRMSE and PCC are 44.8% and 44.6% respectively). On the other hand, time-frequency BSS-CCA and SOBI algorithms illustrate the lowest removal performance when compared to the other examined time-frequency BSS algorithms. They illustrate 76% and 77.6% mean NRMSEs, and 47.6% and 49.6% mean PCCs through the different SAR levels respectively (maximum differences in NRMSEs are 45.6% and 35.5% respectively, and maximum differences in PCCs are 49% and 46.2 respectively).

NRMSE and PCC values for the time-frequency BSS have been degraded when compared to the time domain BSS results. Transforming time-domain BSS estimated signals to the time-frequency domain created cross terms between; EEG components, artifact components, and between EEG and artifact components. Removing artifacts in the time domain minimised the existence of artifacts in the EEG time domain signal,

thus the time-frequency transformation would contain the EEG information patterns and the EEG components cross terms. On the other hand, removing artifacts in the time-frequency domain minimised the existence of artifacts patterns, but cross terms between artifact components, and between EEG and artifact components still exist. This created a noticeable difference between the clean signal and the estimation, consequently producing high NRMSE and low PCC values.

Summing up, all time and time-frequency artifact removal algorithms cannot be considered robust through the different SAR levels, as their performance metrics were depending on the contamination level, thus prior knowledge of the SAR level is essential for these algorithms. In addition, time domain artifact removal algorithms produce higher overall removal performance when compared to the time-frequency BSS due artifact cross-terms existence.

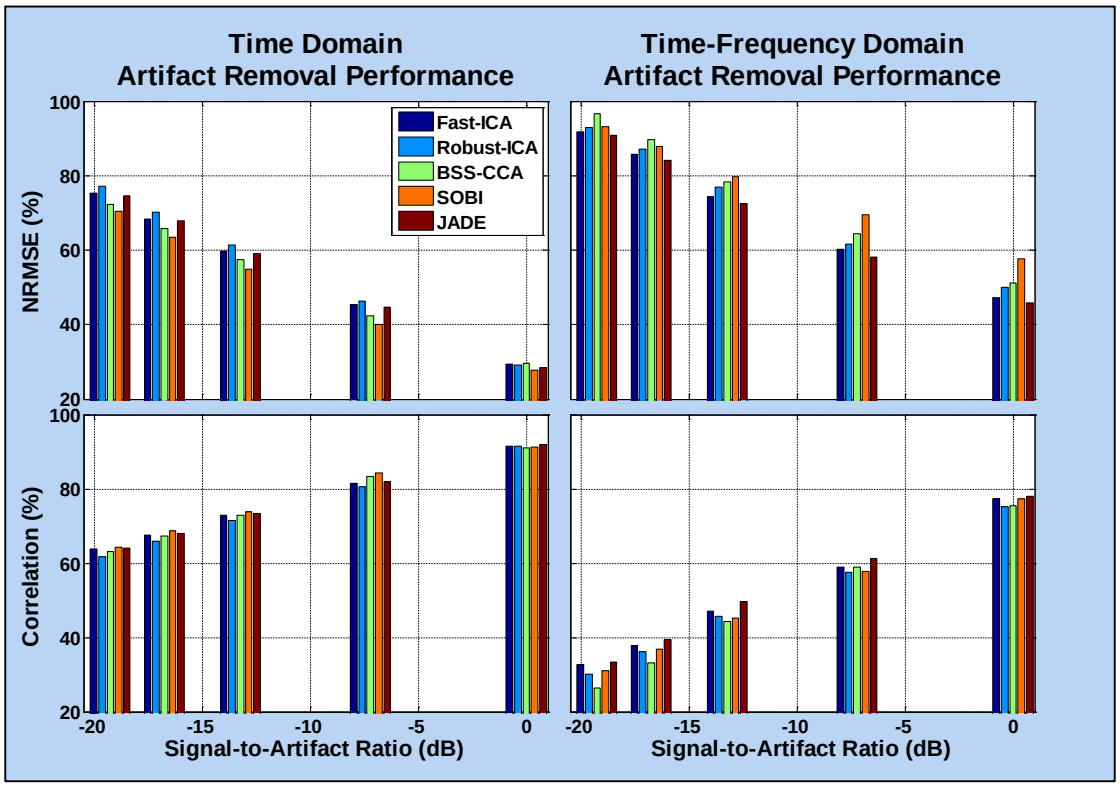

Figure 5.19: Time and time-frequency artifact removal performance assessment through different signal-to-artifact ratios

5.4 Automatic EEG Abnormality Detection

The automatic EEG abnormality detection system output show the feature selection results when using the fused multichannel features, and when using the combination of the fused and extended multichannel features. Furthermore, it demonstrates the performance results of the SVM classifier when using the fused multichannel features, and the combination of the fused and extended multichannel features for the clean, corrupted, and filtered EEG epochs through different SAR levels.

5.4.1 Feature Selection Results

The features selection process is based on calculating the Fisher's score for every feature in the clean EEG waveform independently from the others. Then, ordering them according to their Fisher's scores to select a top ranked subset. Table 5.2 illustrates the ordered fused multichannel time-frequency features and their Fisher's scores. It shows that the instantaneous frequency features and spectral flatness are the worst features to be used in the EEG abnormality detection system, and that Shannon entropy, normalised Renyi entropy, and the TFD mean features are the best. On the other hand, Table 5.3 illustrates the ordered combination of the extended and fused multichannel time-frequency features and their Fisher's scores. It also shows that the instantaneous frequency features and spectral flatness features are the worst to be used in the EEG abnormality detection system, but it demonstrates that the mean time-frequency correlation between the EEG channels would be the best feature to be used along with the Shannon and normalised Renyi entropies. Such change in the best feature to be used could improve the EEG abnormality detection results. Feature selection complete tables are attached in the supplementary material shown in Appendix J.4, page 222.

Table 5.2: Fused multichannel time-frequency features ordered according to their Fisher's scores

#	Feature Label	Fisher Score
12	$\boldsymbol{F_7}$	10.89603343
9	$\boldsymbol{F_6}$	8.408607824
1	$\boldsymbol{T_1}$	6.802327383
⋮	⋮	⋮
10	$\boldsymbol{F_5}$	0
13	$\boldsymbol{IF_1}$	0
14	$\boldsymbol{IF_2}$	0

Table 5.3: Combined extended and fused multichannel time-frequency features ordered according to their Fisher's scores

#	Feature Label	Fisher Score
17	$\boldsymbol{M_1}$	13.10198119
12	$\boldsymbol{F_7}$	10.89603343
9	$\boldsymbol{F_6}$	8.408607824
⋮	⋮	⋮
10	$\boldsymbol{F_5}$	0
13	$\boldsymbol{IF_1}$	0
14	$\boldsymbol{IF_2}$	0

5.4.2 Abnormality Detection Results

The abnormality detection process is based on adapting the feature selection and classification iterative process, along with an SVM classifier that will be cross-validated in every iteration using the LOOCV method. This process will be repeated twice; first using the fused multichannel time-frequency features, and then using the combined extended and fused multichannel time-frequency features. Abnormality detection results will be discussed with respect to accuracy and robustness, where abnormality detection robustness is defined as the maximum accuracy delivered by utilising the maximum number of features.

The classification accuracy profiles when using the fused and the combined extended and fused time-frequency multichannel features are illustrated in Figure 5.20, while averages of the classification accuracy profiles are illustrated in Table 5.4. Firstly, it can be seen that the system accuracy profiles when using corrupted EEG has been improved by utilising the extended multichannel features. The average detection accuracies have increased from 93.0% to 97.8%, 82.8% to 97.6%, and from 70.9% to 97.4% when using corrupted EEG with SAR of -13.2412, -16.763, and -19.262 dB respectively. This proves that the overall system accuracy can be improved by using the combined features with average addition of 4.8%, 14.8% and 26.5%.

Moreover, the maximum accuracies when using corrupted EEG with SAR of: 0, -7.2206, -13.2412, and -16.763 dB are: 100%, 100%, 99%, and 90% respectively by utilising 8 fused features. Furthermore, 77% is the maximum accuracy when using corrupted EEG with SAR of -19.2618 dB and by utilising 7 fused features. On the other hand, maximum accuracies when using corrupted EEG with all different SARs have reached 100% by utilising 9 combined multichannel features. This proves that the overall system robustness and accuracy are enhanced when using the extended features as the number of feasible features has increased from 7 or 8 to 9 and the least accuracy has increased by 23%.

Classification accuracy profiles for the time and time-frequency filtered EEG using the fused and the combined extended and fused time-frequency multichannel features are attached in Appendix J.5, page 224.

Results show that time domain artifact detection and removal improves the abnormality detection average accuracies when utilising the fused multichannel features. They have increased by approximately 3.5%, 11.8%, and 16% when applying Robust-ICA on the

corrupted EEG, and by 2.8%, 10.3%, and 18.7% when applying Fast-ICA on the corrupted EEG with SAR of -13.2412, -16.763, and -19.262 dB respectively. On the other hand, average accuracies have degraded when applying time domain artifact detection and removal on the corrupted EEG and utilising the extended multichannel features. They have decreased by minimum of 2.3%, 1.9%, 2.5%, and 3.2% when applying SOBI, JADE, Robust-ICA, and Fast-ICA on the corrupted EEG with SAR of -7.22, -13.2412, -16.763, and -19.262 dB respectively.

Moreover, time domain artifact detection and removal does not improve the abnormality detection robustness, because maximum accuracies were achieved using the same number of features as the corrupted EEG. In addition, results show that different algorithms must be used to maintain the same number of utilised features on different SAR levels. SOBI achieves the maximum accuracy at 0 and -7.22 dB SAR when utilising 9 and 8 features respectively, while Robust-ICA achieves the maximum accuracy at -13.2412 and -16.763 dB SAR when utilising 8 features, and finally Fast-ICA achieves the maximum accuracy at -19.262 dB SAR when utilising 8 features. This wide range of used algorithms due different SAR levels prohibits the usage of time domain artifact detection and removal techniques, as adequate algorithms cannot be chosen without prior knowledge on the contamination level.

Time-frequency domain artifact detection and removal results illustrate that abnormality detection average accuracies can be improved when utilising the fused multichannel features. They have approximately increased by 1.9%, 4.6%, and 3.4% when applying JADE on the corrupted EEG, and by 1.3%, 6.7%, and 3% when applying Fast-ICA on the corrupted EEG with SAR of -13.2412, -16.763, and -19.262 dB respectively.

In addition, abnormality detection average accuracies has also improved when utilising the extended multichannel features. They have approximately increased by 0.13% and 0.3% when applying Fast-ICA on the corrupted EEG with SAR of -13.2412 and -19.262 dB. Furthermore, JADE has increased the average accuracy by 0.4% and 0.1% when applied on the corrupted EEG with SAR of -16.763 and -19.262 dB.

Moreover, time-frequency artifact detection and removal does not improve the abnormality detection robustness, because maximum accuracies were achieved using the same number of features as the corrupted EEG. In addition, results show that only the time-frequency JADE algorithm can be used to maintain the same number of features on all the SAR levels. This makes time-frequency artifact detection and removal preferable as they do not require prior knowledge on the SAR level in contrast with time domain techniques.

Summing up, abnormality detection average accuracies can be improved when using corrupted EEG with SARs of -13.2412, -16.763, and -19.262 dB. Utilising the fused multichannel features along with time domain Fast-ICA increases the average accuracies from; 93.0%, 82.8%, and 70.9% to; 95.8%, 93.1%, and 89.6% respectively (*2.8%, 10.3%, and 18.7% increments*). On the other hand, using time-frequency JADE increases the average accuracies to; 94.9%, 87.4%, and 74.3% respectively (*1.9%, 4.6%, and 3.4% increments*). Furthermore, by only utilising the extended multichannel features average accuracies have increased from; 93.0%, 82.8%, and 70.9% to; 97.8%, 97.6%, and 97.4% (*4.8%, 14.8% and 26.5% increments*) when using corrupted EEG with SAR of -13.2412, -16.763, and -19.262 dB respectively. Adding to that, using time-frequency JADE slightly improves these accuracies to be 97.7%, 98.0%, and 97.5% (*-0.1%, 0.4% and 0.1% increments*), while using time domain techniques

degrade them. Moreover, both time and time-frequency artifact detection and removal techniques do not improve the abnormality detection robustness, because most maximum accuracies were achieved using the same number of features as the corrupted EEG. In addition, results show that only the time-frequency JADE algorithm can be used to maintain the same number of features on all the SAR levels, thus time-frequency artifact detection and removal techniques are preferable.

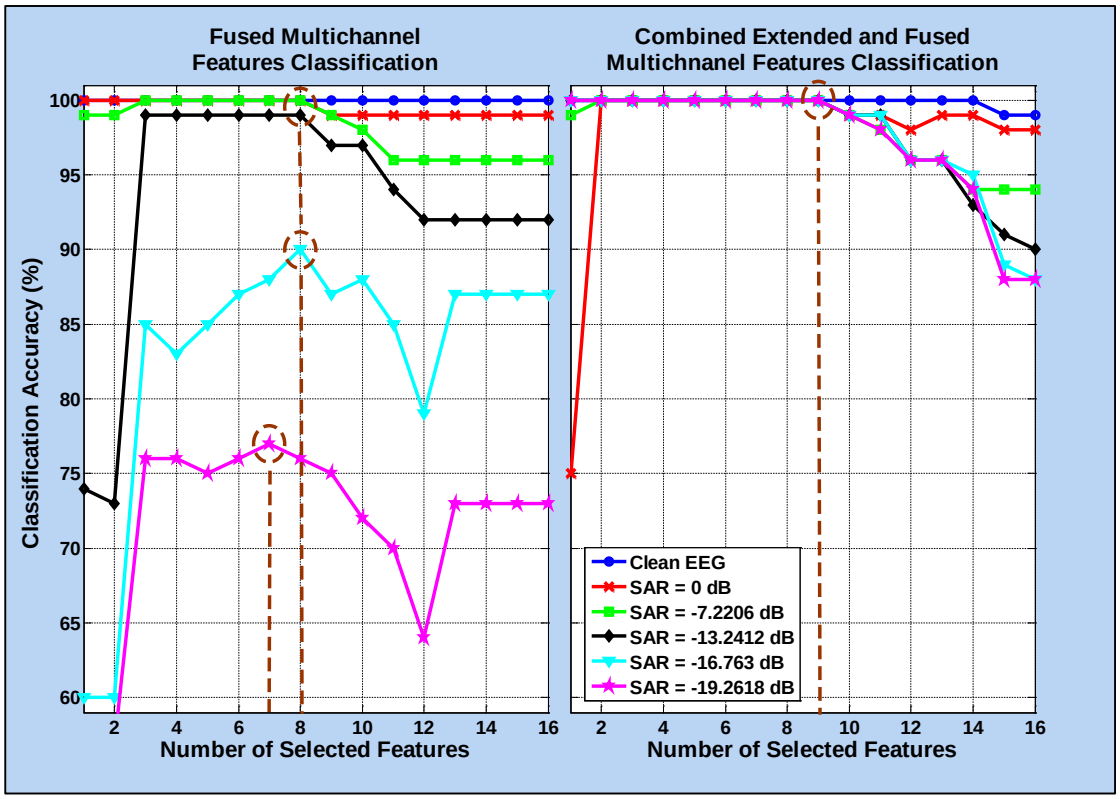

Figure 5.20: EEG abnormality detection accuracy profiles when using the fused and the combined extended and fused time-frequency multichannel features for the clean and corrupted EEG. Maximum accuracies delivered by utilising the maximum number of features are contoured in dark brown circles

Table 5.4: EEG abnormality detection average performances when using the fused and the combined extended and fused time-frequency multichannel features for the clean and corrupted EEG. *min* is the minimum value, *max* is the maximum value, μ is the mean value

	Fused Multichannel Time-Frequency Features								
	Balanced Accuracy (%)			**Sensitivity (%)**			**Specificity (%)**		
	min	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*
Clean EEG	100	**100**	100	100	**100**	100	100	**100**	100
SAR ≈ 0 dB	99	**99.5**	100	100	**100**	100	98	**99**	100
SAR ≈ -7.22 dB	96	**98.1875**	100	98	**99.75**	100	92	**96.625**	100
SAR ≈ -13.241 dB	73	**93.0625**	99	60	**93.5**	98	84	**92.625**	100
SAR ≈ -16.763 dB	60	**82.8125**	90	40	**82**	90	66	**83.625**	90
SAR ≈ -19.262 dB	49	**70.9375**	77	40	**71.25**	80	50	**70.625**	76
	Combined Extended and Fused Multichannel Time-Frequency Features								
Clean EEG	99	**99.875**	100	100	**100**	100	98	**99.75**	100
SAR ≈ 0 dB	75	**97.8125**	100	66	**97.875**	100	84	**97.75**	100
SAR ≈ -7.22 dB	94	**98.125**	100	98	**99.875**	100	88	**96.375**	100
SAR ≈ -13.241 dB	90	**97.75**	100	100	**100**	100	80	**95.5**	100
SAR ≈ -16.763 dB	88	**97.625**	100	100	**100**	100	76	**95.25**	100
SAR ≈ -19.262 dB	88	**97.4375**	100	94	**99.25**	100	76	**95.625**	100

CHAPTER 6: CONCLUSIONS AND FUTURE PERSPECTIVES

The synthetic neonatal multichannel EEG was constructed by using: a neonatal single channel EEG model, a neonatal head model, 21 EEG electrodes, and an EEG propagation model. After that, the generated seizure-to-background ratios were validated against published references using Probability-Probability (*P-P*) plots. Furthermore, the synthetic neonatal multichannel model was validated against real neonatal EEG data using Spearman's correlation.

The neonatal single-channel EEG used the Rankine model [41] to produce background and seizure epochs. Secondly, the neonatal head model used Sadleir four sphere approach [21], which divided the head into four concentric spheres namely; scalp, skull, cerebrospinal fluid, and the brain. Furthermore, 21 EEG electrodes were placed on the scalp surface of the neonatal head model according to the international 10-20 system [27] [28] [29] [30]. Moreover, the EEG propagation model was constructed by modifying the radiation transport equation to describe the EEG propagation through the head model. It had three tuning options to alter the head optical properties making them; pure optical absorbing, pure optical scattering, or optical absorbing and scattering♣.

The P-P validation results illustrated that pure optical absorbing medium is the best candidate to be used further in the system, as its normal line had an R^2 value of 0.999990 when fitting the reference normal line. After that, the synthetic neonatal

♣ Despite the fact that, RTE uses light and EEG is an electrical measurement; both signals are fundamentally electromagnetic, thus they obey the same fundamental laws of physics (absorption, scattering, reflection, and the inverse square law [45]). Furthermore, applying the RTE on a validated neonatal head model, would represent the EEG propagation through the different head structures, thus EEG power will decay and disperse according to the head tissue's optical properties.

multichannel EEG was assembled and user options were created to alter: the type, duration, and the number of EEG patterns appearing on every channel. Finally, the model validation results showed the model effectiveness in mimicking the morphology of real neonatal EEG, as it illustrated 0.58065 and 0.57796 mean correlations for the background EEG comparison and 0.65107 and 0.66254 for the seizure EEG comparison in the frequency and time-frequency domains respectively

Capabilities of this model can be upgraded by utilising several options such as: different background and seizure simulators, different electrode placement standards such as the 10-20, 10-10, and 5-10 systems, and an option to include or exclude the neonatal fontanel in the neonatal head model. These additions can tailor the model according to the user specific needs, and can broaden its validation and simulation capabilities.

The synthetic neonatal multichannel EEG artifacts model was constructed by combining three synthetic physiological signals to corrupt the synthetic neonatal EEG. After that, corrupted EEG was generated with different signal-to-artifact ratios (*SAR*) to assess the system performance through different contamination levels.

The model used three synthetic physiological signals namely; blood vessel pulsation, electrocardiogram spikes, and short-time high-amplitude signals to produce EEG artifacts. Furthermore, an EEG multichannel artifact system was created with user inputs to alter the: location, duration, amplitude, and the number of artifact segments in the EEG signal. After that, five versions of corrupted EEG were created with signal-to-artifact ratios (*SAR*) of: 0, -7.2206, -13.2412, -16.763, and -19.262 dB respectively, to assess the system robustness and accuracy for all the different SAR levels.

Contamination potentials of the model can be upgraded and enhanced by expanding the synthetic templates to include other types of EEG artifacts, such as eye blinking and

body movement artifacts. Furthermore, randomness in choosing and placing the synthetic artifacts on different channels can be introduced, thus making the model more suitable in testing artifact detection and removal algorithms.

The automatic artifacts detection and removal system was designed to process corrupted EEG signals in the time (*t*) and time-frequency (*t-f*) domains using BSS algorithms. Furthermore, the artifact detection technique was statistically optimised and the artifact removal was evaluated. After that, artifact detection and removal performances were tested using corrupted EEG with different SAR levels in order to assess its accuracy and robustness to contamination levels.

The system utilised Fast-ICA, Robust-ICA, BSS-CCA, SOBI, and JADE algorithms to estimate the source signals in the t and t-f domains. After that, artifacts were detected by comparing an optimised correlation threshold with the correlation between the estimated source signals and a reference signal. Finally, artifact removal performance was evaluated by computing Normalised Root Mean Square Error (*NRMSE*) and Pearson's Correlation Coefficient (*PCC*) metrics in the t-f domain between the filtered and original clean epochs.

Artifact detection results illustrated that, all t and t-f artifact detection and removal algorithms were robust with respect to their detection capabilities for all the different SAR levels (maintained 5% maximum difference), but not with respect to their removal capabilities (NRMSE and PCC were degrading when increasing the SAR level). JADE algorithm had produced the highest and the most robust detection performance when utilised in the t domain (97% accuracy, 94% sensitivity, and 100% specificity), and the highest mean removal performance when utilised in the t-f domain (70.3% NRMSE and 52.3% PCC). On the other hand, SOBI algorithm had produced the highest and the

most robust mean detection performance when utilised in the t-f domain (90.4% accuracy, 86.8% sensitivity, and 94% specificity), and the highest mean removal performance when utilised in the t domain (51.3% NRMSE and 79.5% PCC).

The automatic EEG abnormality detection was constructed using: t-f multichannel features, extended multichannel features, multichannel feature fusion scheme, feature selection algorithm, and a cross validated Support Vector Machine (*SVM*) classifier

Multichannel t-f features were extracted from every channel, then fused using the Multichannel Feature Fusion Sum Basis technique. Furthermore, extended t-f multichannel features were introduced based on EEG multichannel information alikeness, and then combined with the fused features. After that, Fisher's scores were computed for both feature sets of the clean EEG waveform, and then ordered according to their scores to select a top ranked subset. Finally, the feature subset was plugged into the feature selection and classification iterative process to produce classification accuracy profiles using cross-validated SVM classifier.

Feature selection results illustrated that Shannon and normalised Renyi entropies were the best features to be used in the EEG abnormality detection system when utilising the fused multichannel feature set (Fisher's scores are 10.9 and 8.4 respectively). Furthermore, the average t-f correlation among EEG channels became the best feature to be used when utilising the combined multichannel feature set (Fisher's score 13.1). Abnormality detection results demonstrated that, using corrupted EEG with SARs of -13.2412, -16.763, and -19.262 dB along with utilising the fused feature set, had degraded the average accuracies of the classifier from 100% to 93.0%, 82.8%, and 70.9% respectively. However, they were increased by using the time domain Fast-ICA algorithm (95.8%, 93.1%, and 89.6% respectively), and by using the t-f JADE

algorithm (94.9%, 87.4%, and 74.3% respectively). Furthermore, by only utilising the extended multichannel features the average accuracies were increased to 97.8%, 97.6%, and 97.4% respectively. Adding to that, using the t-f JADE along with the extended multichannel features had slightly improved these accuracies to be 97.7%, 98.0%, and 97.5%, but using time domain artifact detection and removal degraded them. Furthermore, the maximum accuracies when using corrupted EEG with SAR of; 0, -7.2206, -13.2412, and -16.763 dB were; 100%, 100%, 99%, and 90% respectively by utilising 8 fused features, and 77% for an SAR of -19.2618 dB by utilising 7 fused features. However, they were increased to 100% by utilising 9 combined multichannel features. This indicates that the robustness and accuracy of the system had improved when using the extended features as the number of feasible features increased from 7 or 8 to 9, and the least maximum accuracy had increased by 23%.

Improvements to the automatic EEG abnormality detection system can be done by utilising: different multichannel fusion schemes, such as symmetrical uncertainty, different feature selection algorithms, such as mRMR, adaptive feature selection based on patients and/or data, and regularized tree, and different classifiers, such as: naive Bayes, K-nearest neighbour, Extended Nearest Neighbour, and Convolutional Neural Networks. These added options can help determine which feature selection algorithm and classifier result in the maximum detection performance.

BIBLIOGRAPHY

[1] B. Boashash, Time-Frequency Signal Analysis and Processing: A Comprehensive Review, Second ed., Elsevier, 2015.

[2] N. S. Abend and C. J. Wusthoff, "Neonatal Seizures and Status Epilepticus," *J Clin Neurophysiol,* vol. 29, no. 5, pp. 441-448, 2013.

[3] F. Pisani, C. Spagnoli, E. Pavlidis, C. Facini, G. M. Kouamou Ntonfo, G. Ferrari and R. Raheli, "Real-time automated detection of clonic seizures in newborns," *Clinical neurophysiology,* vol. 125, no. 8, pp. 1533-40, 2014.

[4] L. G. Van Rooij, L. Hellström-Westas and L. S. De Vries, "Treatment of neonatal seizures," *Seminars in Fetal and Neonatal Medicine,* vol. 18, no. 4, pp. 209-215, 2013.

[5] C. Uria-Avellanal, N. Marlow and J. M. Rennie, "Outcome following neonatal seizures," *Seminars in fetal & neonatal medicine,* vol. 18, no. 4, pp. 224-32, 2013.

[6] E. M. Thomas, A. Temko, W. P. Marnane, G. B. Boylan and G. Lightbody, "Discriminative and generative classification techniques applied to automated neonatal seizure detection," *IEEE Journal of Biomedical and Health Informatics,* vol. 17, no. 2, pp. 297-304, 2013.

[7] G. B. Boylan, N. J. Stevenson and S. Vanhatalo, "Monitoring neonatal seizures," *Seminars in Fetal and Neonatal Medicine,* vol. 18, no. 4, pp. 202-208, 2013.

[8] L. Orosco, A. G. Correa and E. Laciar, “Review: A Survey of Performance and Techniques for Automatic Epilepsy Detection,” *Journal of Medical and Biological Engineering*, vol. 33, no. 6, pp. 526-537, 2013.

[9] K. T. Sweeney, T. E. Ward and S. F. McLoone, “Artifact Removal in Physiological Signals—Practices and Possibilities,” *IEEE Transactions on Information Technology In Biomedicine*, vol. 16, no. 3, pp. 488-500, 2012.

[10] J. A. Uriguen and B. G. Zapirain, “EEG artifact removal - State-of-the-art and guidelines,” *Journal of Neural Engineering*, vol. 12, no. 3, pp. 1741-2560, 2015.

[11] S. H. Sardouie, M. B. Shamsollahi, L. Albera and I. Merlet, “Denoising of Ictal EEG Data Using Semi-Blind Source Separation Methods Based on Time-Frequency Priors,” *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 3, pp. 839-847, 2015.

[12] A. Schulze-Bonhage, H. Feldwisch-Drentrup and M. Ihle, “The role of high-quality EEG databases in the improvement and assessment of seizure prediction methods,” *Epilepsy and Behavior*, vol. 22, no. SUPPL. 1, pp. 88-93, 2011.

[13] N. Stevenson, M. Mesbah, G. Boylan, P. Colditz and B. Boashash, “A nonlinear model of newborn EEG with nonstationary inputs,” *Annals of Biomedical Engineering*, vol. 38, no. 9, pp. 3010-3021, 2010.

[14] J. Anemüller, T. J. Sejnowski and S. Makeig, “Complex independent component analysis of frequency-domain,” *Neural Networks*, vol. 16, no. 9, pp. 1311-1323, 2003.

[15] S. He-ping, C. Ji-hua and L. Xiao, “Blind Source Separation For Non-Stationary Signal Based On Time-frequency Analysis,” in *Fourth International Conference on Intelligent Networks and Intelligent Systems*, Kuming, China, 2011.

[16] S. H. Sardouie, M. B. Shamsollahi, L. Albera and I. Merlet, “Denoising of Ictal EEG Data Using Semi-Blind Source Separation Methods Based on Time-Frequency Priors,” *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 3, pp. 839-847, 2015.

[17] K. R. Miller and J. Levine, Biology, New Jersey: Pearson Prentice Hall, 2006.

[18] I. Despotovic, E. Vansteenkiste and W. Philips, “A realistic volume conductor model of the neonatal head: methods, challenges and applications,” in *35th Annual International Conference of the IEEE EMBS*, Osaka, Japan, 2013.

[19] I. Despotovic, P. J. Cherian, M. D. Vos, H. Hallez, W. Deburchgraeve, P. Govaert, M. Lequin, G. H. Visser, R. M. Swarte, E. Vansteenkiste, S. V. Huffel and W. Philips, “Relationship of EEG Sources of Neonatal Seizures to Acute Perinatal Brain Lesions Seen on MRI: A Pilot Study,” *Human Brain Mapping*, vol. 10, pp. 2402 - 2417, 2013.

[20] F. Vatta, F. Meneghini, F. Esposito, S. Mininel and F. D. Salle, “Realistic and Spherical Head Modeling for EEG Forward Problem Solution: A Comparative Cortex-Based Analysis,” *Computational Intelligence and Neuroscience*, vol. 2010, 2010.

[21] R. J. Sadleir and T. Tang, “Electrode configurations for detection of intraventricular haemorrhage in the premature neonate,” *Physiol Meas.*, vol. 30, no. 1, pp. 63 - 79, 2010.

[22] G. Ala and E. Francomano, “A Multi-Sphere Particle Numerical Model for Non-Invasive Investigations of Neuronal Human Brain Activity,” *Progress In Electromagnetics Research Letters,* vol. 36, pp. 143 - 153, 2013.

[23] D. E. Hyde, F. H. Duffy and S. K. Warfield, “Anisotropic Partial Volume CSF Modeling for EEG Source Localization,” *Neuroimage.,* vol. 62, no. 3, pp. 2161 - 2170, 2012.

[24] N. Gupta and S. Devi, “EEG Forward Problem Solution for a Multi-Shell Head Model,” *International Journal of Recent Technology and Engineering (IJRTE),* vol. 2, no. 1, pp. 2277 - 3878, 2013.

[25] S. Brigadoi, P. Aljabar, M. Kuklisova-Murgasova, S. R. Arridge and R. J. Cooper, “A 4D neonatal head model for diffuse optical imaging of pre-term to term infants,” *NeuroImage,* vol. 100, pp. 385 - 394, 2014.

[26] T. Banaschewski and D. Brandeis, “Annotation: What electrical brain activity tells us about brain function that other techniques cannot tell us – a child psychiatric perspective,” *Journal of Child Psychology and Psychiatry,* vol. 48, no. 5, pp. 415 - 435, 2007.

[27] B. Boashash, G. Azemi and N. Ali Khan, “Principles of time–frequency feature extraction for change detection in non-stationary signals: Applications to newborn EEG abnormality detection,” *Pattern Recognition,* vol. 48, no. 3, pp. 616-627, 2015.

[28] B. Boashash and G. Azemi, “A review of time-frequency matched filter design with application to seizure detection in multichannel newborn EEG,” *Digital Signal Processing: A Review Journal,* vol. 28, pp. 28-38, 2014.

[29] a. R. Hidalgo-Muñoz, M. M. López, a. T. Pereira, I. M. Santos and a. M. Tomé, "Spectral turbulence measuring as feature extraction method from EEG on affective computing," *Biomedical Signal Processing and Control*, vol. 8, pp. 945-950, 2013.

[30] M. Balakrishnan, P. Colditz and B. Boashash, "A Multi-Channel Fusion Based Newborn Seizure Detection," *J. Biomedical Science and Engineering*, pp. 533-545, 2014.

[31] "10/20 System Positioning Manual," Trans Cranial Technologies, Hong Kong, 2012.

[32] S. Sanei and J. Chambers, EEG Signal Processing, John Wiley & Sons, Ltd, 2007.

[33] K. B. Böcker, J. A. Van Avermaete and M. M. Van den Berg-Lenssen, "The International 10-20 System Revisited: Cartesian and Spherical Co-ordinates," *Brain Topography*, vol. 6, no. 3, pp. 231 - 235, 1994.

[34] H. Martinez-Alfaro, "Integrating Neural Signal and Embedded System for Controlling Small Motor," in *Advances in Mechatronics*, InTech, 2011, pp. 31 - 42.

[35] G. M. Galloway, Clinical Neurophysiology in Pediatrics: A Practical Approach to Neurodiagnostic Testing and Management, Springer Publishing Compan, 2015.

[36] S. B. Nagaraj, N. J. Stevenson, W. P. Marnane, G. B. Boylan and G. Lightbody, "Neonatal Seizure Detection Using Atomic Decomposition With a Novel

Dictionary," *IEEE Transactions on Biomedical Engineering,* vol. 61, no. 11, pp. 2724-2732, 2014.

[37] P. J. Cherian, R. M. Swarte and G. H. Visser, "Technical standards for recording and interpretation of neonatal electroencephalogram in clinical practice," *Ann Indian Acad Neurol,* vol. 12, no. 1, pp. 58-70, 2009.

[38] M. Roessgen, A. M. Zoubir and B. Boashash, "Seizure Detection of Newborn EEG Using a Model-Based Approach," *IEEE Transactions On Biomedical Engineering,* vol. 45, no. 6, pp. 673 - 685, 1998.

[39] F. H. Lopes da Silva, A. Hocks, H. Smits and L. H. Zetterberg, "Model of Brain Rhythmic Activity The Alpha-Rhythm of the Thalamus," *Kybernetik,* vol. 15, pp. 27 - 37, 1974.

[40] P. Celka and P. Colditz, "Nonlinear Nonstationary Wiener Model of Infant EEG Seizures," *IEEE Transactions On Biomedical Engineering,* vol. 49, no. 6, pp. 556 - 564, 2002.

[41] L. Rankine, N. Stevenson, M. Mesbah and B. Boashash, "A Nonstationary Model of Newborn EEG," *IEEE Transactions On Biomedical Engineering,* vol. 54, no. 1, pp. 19 - 28, 2007.

[42] N. Stevenson, L. Rankine, M. Mesbah and B. Boashash, "Newborn EEG Seizure Simulation Using Time–Frequency Signal Synthesis," in *APRS Workshop on Digital Image Computing: Pattern Recognition and Imaging for Medical Applications,* Griffith University, Southbank, Brisbane, Australia, 2005.

[43] L. Rankine, H. Hassanpour, M. Mesbah and B. Boashash, “Newborn EEG simulation from nonlinear analysis,” in *The Eight International Symposium on Signal Processing and its Applications*, Sydney, Australia, 2005.

[44] N. Stevenson, L. Rankine, M. Mesbah and B. Boashash, “Modelling Newborn EEG Background Using A Time-Varying Fractional Brownian Process,” in *15th European Signal Processing Conference*, Poznan, Poland, 2007.

[45] F. T. Ulaby , E. Michielssen and U. Ravaioli, Fundamentals of Applied Electromagnetics, Sixth ed., Prentice Hall, 2010.

[46] S. L. Jacques, “Optical properties of biological tissues: a review,” *IOP Publishing : Phys. Med. Biol.*, vol. 58, pp. 37 - 61, 2013.

[47] V. V. Tuchin, “Light scattering study of tissues,” *Physics - Uspekhi*, vol. 40, no. 5, pp. 495 - 515, 1997.

[48] M. Dehaes, K. Kazemi, M. Pelegrini-Issac, R. Grebe, H. Benali and F. Wallois, “Quantitative Effect of the Neonatal Fontanel on Synthetic Near Infrared Spectroscopy Measurements,” *Human Brain Mapping*, vol. 34, pp. 878 - 889, 2013.

[49] M. Jager and A. Kienle, “Non-invasive determination of the absorption coefficient of the brain from time-resolved reflectance using a neural network,” *Phys Med Biol.*, vol. 56, no. 11, pp. 39 - 44, 2011.

[50] A. J. Trevelyan, D. Sussillo and R. Yuste, “Feedforward Inhibition Contributes to the Control of Epileptiform Propagation Speed,” *The Journal of Neuroscience*, vol. 27, no. 13, pp. 3383 - 3387, 2007.

[51] A. J. Trevelyan, T. Baldeweg, W. v. Drongelen, R. Yuste and M. Whittington, "The Source of Afterdischarge Activity in Neocortical Tonic–Clonic Epilepsy," *The Journal of Neuroscience,* vol. 27, no. 49, pp. 13513 - 13519, 2007.

[52] A. Tetsuya, M. Ayumi, H. Takuto, U. Tatsuya, T. Masato and M.-M. Yuko, "Comparison of EEG propagation speeds under emotional stimuli on smartphone between the different anxiety states," *Frontiers in Human Neuroscience,* vol. 8, pp. 1662 - 5161, 2014.

[53] M. D. Vos, W. Deburchgraeve, P. Cherian, V. Matic, R. Swarte, P. Govaert, G. Visser and S. V. Huffel, "Automated artifact removal as preprocessing refines neonatal seizure detection," *Clinical Neurophysiology,* vol. 122, no. 12, p. 2345–2354, 2011.

[54] A. Janardhan and K. K. Rao, "Application of Signal Separation Algorithms for Artifact Removal from EEG Signals," *International Journal of Modern Communication Technologies & Research,* vol. 3, no. 1, 2015.

[55] V. Matic, W. Deburchgraeve and S. Van Huffel, "Comparison of ICA Algorithms for ECG Artifact Removal from EEG Signals," *IEEE-EMBS Benelux Chapter Symposium,* 2009.

[56] M. Khalifa, M. Mesbah, B. Boashash and P. Colditz, "Influence of EEG Artifacts on Detecting Neonatal Seizure," in *International Conference on Information Science, Signal Processing and their Applications,* Kuala Lumpur, Malaysia, 2010.

[57] M. Zima, P. Tichavsky, K. Paul and V. Krajca, "Robust removal of short-duration artifacts in long neonatal EEG recordings using wavelet-enhanced ICA

and adaptive combining of tentative reconstructions," *Physiological Measurement*, vol. 33, no. 8, 2012.

[58] J. Brotchie, L. Rankine, M. Mesbah, P. Colditz and B. Boashash, "Robust Time-Frequency Analysis of Newborn EEG Seizure Corrupted by Impulsive Artefacts," in *Annual International Conference of the IEEE EMBS*, Lyon, France, 2007.

[59] S. Castaño-Candamil, J. Höhne, J.-D. Martínez-Vargas, X.-W. An, G. Castellanos-Domínguez and S. Haufe, "Solving the EEG inverse problem based on space–time–frequency structured sparsity constraints," *NeuroImage*, vol. 118, pp. 598-612, 2015.

[60] M. M. Siddiqui, G. Srivastava and S. H. Saeed, "Detection of rapid eye movement behaviour disorder using short time frequency analysis of PSD approach applied on EEG signal (ROC-LOC).," *Biomedical Research*, vol. 26, no. 3, pp. 587-593, 2015.

[61] A. Yan, W. Zhou, Q. Yuan, S. Yuan, Q. Wu, . X. Zhao and J. Wang, "Automatic seizure detection using Stockwell transform and boosting algorithm for long-term EEG," *Epilepsy & Behavior*, vol. 45, pp. 8-14, 2015.

[62] K. Samiee, P. Kovacs and M. Gabbouj, "Epileptic Seizure Classification of EEG Time-Series Using Rational Discrete Short-Time Fourier Transform," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 2, pp. 541-552, 2015.

[63] B. Boashash, N. A. Khan and T. Ben-Jabeur, "Time-frequency features for pattern recognition using high-resolution TFDs: A tutorial review," *Digital Signal Processing*, vol. 40, pp. 1-30, 2015.

[64] B. Boashash, G. Azemi and J. M. O'Toole, “Time-Frequency Processing of Nonstationary Signals: Advanced TFD Design to Aid Diagnosis with Highlights from Medical Applications,” *IEEE Signal Processing Magazine,* vol. 30, no. November , pp. 108-119, 2013.

[65] A. Hyvarinen, P. Ramkumar, L. Parkkonen and R. Hari, “Independent component analysis of short-time Fourier transforms for spontaneous EEG/MEG analysis,” *NeuroImage,* vol. 49, pp. 257 - 271, 2010.

[66] T. Adali, M. Anderson and F. Geng-Shen, “Diversity in Independent Component and Vector Analyses: Identifiability, algorithms, and applications in medical imaging,” *Signal Processing Magazine,* vol. 31, no. 3, pp. 18 - 33, 2014.

[67] M. Sarfraz, F. F. Li and A. A. Khan, “Independent Component Analysis Methods to Improve Electrocardiogram Patterns Recognition in the Presence of Non-Trivial Artifacts,” *Journal of Medical and Bioengineering,* vol. 4, no. 3, pp. 221 - 226, 2015.

[68] N. Miljkovic, V. Matic, S. V. Huffel and M. B. Popovic, “Independent Component Analysis (ICA) Methods for Neonatal EEG Artifact Extraction: Sensitivity to Variation of Artifact Properties,” in *10th Symposium on Nueral Network Applications in Electrical Engineering (NEUREL-2010),* Serbia, 2010.

[69] K. Vanderperren, M. D. Vos, J. R. Ramautar, N. Novitskiy, M. Mennes, S. Assecondi, B. Vanrumste, P. Stiers, B. R. Van den Bergh, J. Wagemans, L. Lagae, S. Sunaert and S. V. Huffel, “Removal of BCG artifacts from EEG recordings inside the MR scanner: A comparison of methodological and validation-related aspects,” *NeuroImage,* vol. 50, pp. 920 - 934, 2010.

[70] Q. Zhao, B. Hu, Y. Shi, Y. Li, P. Moore, M. Sun and H. Peng, “Automatic identification and removal of ocular artifacts in EEG - Improved adaptive predictor filtering for portable applications,” *IEEE Transactions on Nanobioscience,* vol. 13, no. 2, pp. 109-117, 2014.

[71] S. D. Muthukumaraswamy, “High-frequency brain activity and muscle artifacts in MEG/EEG: a review and recommendations,” *Frontiers in human neuroscience,* vol. 7, no. April, p. Article 138, 2013.

[72] I. Daly, M. Billinger, R. Scherer and G. Müller-Putz, “On the automated removal of artifacts related to head movement from the EEG,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 21, no. 3, pp. 427-434, 2013.

[73] M. B. Hamaneh, N. Chitravas, K. Kaiboriboon, S. D. Lhatoo and K. A. Loparo, “Automated Removal of EKG Artifact From EEG Data Using Independent Component Analysis and Continuous Wavelet Transformation,” *Biomedical Engineering,* vol. 61, no. 6, pp. 1634 - 1641, 2014.

[74] A. Zachariah, J. Jai and G. Titus, “Automatic EEG Artifact Removal by Independent Component Analysis Using Critical EEG Rhythms,” in *International Conference on Control Communication and Computing (ICCC),* Thiruvananthapuram, India, 2013.

[75] P. Ramkumar, L. Parkkonen and A. Hyvarinen, “Group-level spatial independent component analysis of Fourier envelopes of resting-state MEG data,” *NeuroImage,* vol. 86, pp. 480 - 491, 2014.

[76] B. Boashash and L. Boubchir, “On the Selection of Time-Frequency Features for Improving the Detection and Classification of Newborn EEG Seizure Signals and Other Abnormalities,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics),* no. 4, pp. 634-643, 2012.

[77] B. Boashash, L. Boubchir and G. Azemi, “Time-frequency signal and image processing of non-stationary signals with application to the classification of newborn EEG abnormalities,” *IEEE International Symposium on Signal Processing and Information Technology,* vol. 2, pp. 120-129, 2011.

[78] A. S. Al-Fahoum and A. a. Al-Fraihat, “Methods of EEG signal features extraction using linear analysis in frequency and time-frequency domains,” *ISRN neuroscience,* vol. 2014, 2014.

[79] C. Bruser, J. Diesel, M. D. H. Zink, S. Winter, P. Schauerte and S. Leonhardt, “Automatic Detection of Atrial Fibrillation in Cardiac Vibration Signals,” *IEEE journal of biomedical and health informatics,* vol. 17, no. 1, pp. 1-10, 2013.

[80] K. Fu, J. Qu, Y. Chai and Y. Dong, “Classification of seizure based on the time-frequency image of EEG signals using HHT and SVM,” *Biomedical Signal Processing and Control,* vol. 13, pp. 15-22, 2014.

[81] R. Jenke, A. Peer and M. Buss, “Feature Extraction and Selection for Emotion Recognition from EEG,” *Affective Computing, IEEE Transactions on Computing,* vol. 5, no. 3, pp. 327-339, 2014.

[82] R. .. S. Kumari and J. J. Prabin, “Seizure Detection in EEG Using Time Frequency Anlysis and SVM,” in *International Conference on Emerging Trends in Electrical and Computer Technology (ICETECT)*, 2011.

[83] S. Kashef and H. Nezamabadi-pour, “A new feature selection algorithm based on binary ant colony optimization,” in *Conference on Information and Knowledge Technology (IKT) 5th*, Shiraz, 2013.

[84] X. B. Liu, L. Ma, L. Song, Y. F. Zhao, X. M. Zhao and C. W. Zhou, “Recognizing Common CT Imaging Signs of Lung Diseases Through a New Feature Selection Method Based on Fisher Criterion and Genetic Optimization,” *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 2, pp. 635-647, 2015.

[85] A. Adam, M. I. Shapiai, M. Z. M. Tumari, M. S. Mohamad and M. Mubin, “Feature selection and classifier parameters estimation for EEG signals peak detection using particle swarm optimization,” *The Scientific World Journal*, 2014.

[86] Q. Gu, Z. Li and J. Han, “Generalized Fisher Score for Feature Selection,” in *Proceedings of the Twenty-Seventh Conference on Uncertainty in Artificial Intelligence*, Barcelona, Spain, 2012.

[87] “Feature Selection Algorithm,” arizona state university, [Online]. Available: http://featureselection.asu.edu/software.php. [Accessed November 2015].

[88] A. Subasi and M. Ismail Gursoy, “EEG signal classification using PCA, ICA, LDA and support vector machines,” *Expert Systems with Applications*, vol. 37, no. 12, pp. 8659-8666, 2010.

[89] F. Yan, Q. Shang, S. Xia, Q. Wang and P. Ma, “Topological study on the toxicity of ionic liquids on Vibrio fischeri by the quantitative structure–activity relationship method,” *Journal of Hazardous Materials,* vol. 268, pp. 410-415, 2015.

[90] X. Wang, Y. Jiao, T. Tang, H. Wang and Z. Lu, “Altered regional homogeneity patterns in adults with attention-deficit hyperactivity disorder,” *European Journal of Radiology,* vol. 82, no. 9, pp. 1552-1557, 2013.

[91] H. Cai, P. Ruan, M. Ng and T. Akutsu, “Feature weight estimation for gene selection: a local hyperlinear learning approach,” *BMC bioinformatics,* vol. 15, pp. 70-83, 2014.

[92] “Newborn EEG Simulation,” The University of Queensland, [Online]. Available: http://www.som.uq.edu.au. [Accessed 10 March 2015].

[93] P. Gargiulo, P. Belfiore, E. Friðgeirsson, S. Vanhatalo and C. Ramon, “The effect of fontanel on scalp EEG potentials in the neonate,” *Clinical Neurophysiology,* 2014.

[94] B. Lanfer, M. Scherg, M. Dannhauer, T. Knösche, M. Burger and C. Wolters, “Influences of skull segmentation inaccuracies on EEG source analysis,” *NeuroImage,* vol. 62, pp. 418 - 431, 2012.

[95] S. Lew, D. D. Sliva, M.-s. Choe, P. E. Grant, Y. Okada, C. H. Wolters and M. S. Hämäläinen, “Effects of sutures and fontanels on MEG and EEG source analysis in a realistic infant head model,” *NeuroImage,* vol. 76, pp. 282 - 293, 2013.

[96] A. Custo and D. A. Boas, “Comparison of Diffusion and Transport in human head,” *Biomedical Topical Meeting, OSA Technical Digest,* 2004.

[97] J. Heiskala, T. Neuvonen, P. E. Grant and I. Nissilä, “Significance of tissue anisotropy in optical tomography of the infant brain,” *Applied Optics,* vol. 46, no. 10, p. 16331640, 2007.

[98] J. Heiskala, M. Pollari, M. Metsäranta, P. E. Grant and I. Nissilä, “Probabilistic atlas can improve reconstruction from optical imaging of the neonatal,” *Optics Express,* vol. 17, no. 17, p. 1497714992, 2009.

[99] M. Shen, L. Lin, J. Chen and C. Q. Chang, “A Prediction Approach for Multichannel EEG Signals Modeling Using Local Wavelet SVM,” *IEEE Transactions on Instrumentation and Measurement,* vol. 59, no. 5, pp. 1485-1492, 2010.

[100] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David and C. E. Elger, “Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state,” *Physical Review E,* vol. 64, no. 6, pp. 1-8, 2001.

[101] D. Cui, X. Li, X. Ji and L. Liu, “Multi-channel neural mass modelling and analyzing,” *Science China Information Sciences,* vol. 54, no. 6, pp. 1283-1292, 2011.

[102] M. Mesbah, J. M. O. Toole, P. B. Colditz and B. Boashash, “Instantaneous frequency based newborn EEG seizure characterisation,” *Advances in Signal Processing,* vol. 143, no. 1, 2012.

[103] M. Mesbah and B. Boashash, “Performance Comparison of Seizure Detection Methods Using EEG of Newborns for Implementation of A DSP Subsystem,” in *IEEE International Conference in Acoustics, Speech, and Signal Processing*, Orlando, Florida, 2002.

[104] L. Rankine, M. Mesbah and B. Boashash, “A matching pursuit-based signal complexity measure for the analysis of newborn EEG,” *Med Bio Eng Comput*, vol. 45, pp. 251 - 260, 2007.

[105] Y. Liang and W. Chen, “A survey on computing Le´vy stable distributions and a new MATLAB toolbox,” *ELSEVIER*, vol. 93, no. 1, pp. 242-251, 2012.

[106] A. Hyvarinen and E. Oja, “Independent Component Analysis: Algorithms and Applications,” *Neural Networks*, vol. 13, pp. 411 - 430, 2000.

[107] V. Zarzoso and P. Comon, “Robust Independent Component Analysis by Iterative Maximization of the Kurtosis Contrast With Algebraic Optimal Step Size,” *Neural Networks*, vol. 21, no. 2, pp. 248 - 261, 2010.

[108] N. M. Correa, T. Adali, Y.-O. Li and V. D. Calhoun, “Canonical Correlation Analysis for Data Fusion and Group Inferences: Examining applications of medical imaging data,” *IEEE Signal Process Mag.*, vol. 27, no. 4, pp. 39 - 50, 2010.

[109] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso and E. Moulines, “A Blind Source Separation Technique Using Second-Order Statistics,” *IEEE Transactions On Signal Processing*, vol. 45, no. 2, pp. 434 - 444, 1997.

[110] A. Yeredor, “Blind Separation of Gaussian Sources via Second-Order Statistics with Asymptotically Optimal Weighting,” *IEEE Signal Processing Letters,* vol. 7, no. 7, pp. 197 - 200, 2000.

[111] K. Nordhausen, J.-F. Cardoso, J. Miettinen, H. Oja, E. Ollila and S. Taskinen, “Blind Source Separation Methods Based on Joint Diagonalization and Some BSS Performance Criteria,” 2015.

[112] B. Boashash, Modern Signal Processing, Brisbane, Queensland: Queenslan University of Technology, 1999.

[113] M. Geissbuehler and T. Lasser, “How to display data by color schemes compatible with red-green color perception deficiencies,” *Optics Express,* vol. 21, no. 8, pp. 9862-9874, 2013.

[114] B. Boashash and T. Ben-Jabeur, “Design of a High-Resolution Separable-Kernel Quadratic TFD for Improving Newborn Health Outcomes Using Fetal Movement Detection,” in *The 11th International Conference on Information Sciences, Signal Processing and their Applications*, 2012.

[115] M.-T. Shih, F. Doctor, S.-Z. Fan, K.-K. Jen and J.-S. Shieh, “Instantaneous 3D EEG Signal Analysis Based on Empirical Mode Decomposition and the Hilbert–Huang Transform Applied to Depth of Anaesthesia,” *Entropy,* vol. 17, no. 3, pp. 928-949, 2015.

[116] M. Wacker and H. Witte, “Time-frequency Techniques in Biomedical Signal Analysis,” *Methods of Information in Medicine,* vol. 52, no. 4, pp. 279-296, 2013.

[117] M. Abed, A. Belouchrani, M. Cheriet and B. Boashash, “Compact Support Kernels Based Time-Frequency Distributions: Performance Evaluation,” in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, 2011.

[118] M. Abed, A. Belouchrani and M. Cheriet, “Time-Frequency Distributions Based on Compact Support Kernels : Properties and Performance Evaluation,” *IEEE Transactions on Signal Processing*, vol. 60, no. 6, pp. 2814-2827, 2012.

[119] H. Gavert, J. Hurri, J. Sarela and A. Hyvarinen, “Independent Component Analysis (ICA) and Blind Source Separation (BSS),” Alto University, [Online]. Available: http://research.ics.aalto.fi/ica/newindex.shtml. [Accessed 1 October 2015].

[120] V. Z. Gascón-Pelegrí, “RobustICA algorithm for independent component analysis,” Laboratoire d'Informatique, Signaux et Systèmes de Sophia Antipolis, [Online]. Available: http://www.i3s.unice.fr/~zarzoso/robustica.html. [Accessed 1 October 2015].

[121] M. Borga, “Canonical Correlation Analysis,” LINKOPINGS UNIVERSITET, [Online]. Available: http://www.imt.liu.se/people/magnus/cca/. [Accessed 1 October 2015].

[122] A. Turnip, “Comparison of ICA-Based JADE and SOBI Methods EOG Artifacts Removal,” *Journal of Medical and Bioengineering*, vol. 4, no. 6, pp. 436-440, 2015.

[123] A. Cichocki, “ACSOBIRO,” Swartz Center for Computational Neuroscience, [Online]. Available: http://sccn.ucsd.edu/eeglab/allfunctions/acsobiro.html. [Accessed 1 October 2015].

[124] J. F. Cardoso, “Blind Source Separation and Independent Component Analysis,” TELECOM ParisTech, [Online]. Available: http://perso.telecom-paristech.fr/~cardoso/Algo/Jade/jadeR.m. [Accessed 1 October 2015].

[125] D. M. Powers, “Evaluation Evaluation a Monte Carlo study,” Cornell University Library, 2015.

[126] M. Mukaka, “Statistics Corner: A guide to appropriate use of Correlation coefficient in medical research,” *Malawi Medical Journal,* vol. 24, no. 3, pp. 69-71, 2012.

[127] R. M. Warner, Applied Statistics From Bivariate Through Multivariate Techniques, 2 ed., SAGE, 2013.

[128] D. C. Montgomery and G. C. Runger, Applied Statistics and Probability for Engineers, John Wiley & Sons, 2011.

[129] A. Ghasemi and S. Zahediasl, “Normality Tests for Statistical Analysis: A Guide for Non-Statisticians,” *Int J Endocrinol Metab.,* vol. 10, no. 2, pp. 486-489, 2012.

[130] K. B. Bocker, J. A. Van Avermaete and M. M. Van den Berg-Lenssen, “The International 10-20 System Revisited: Cartesian and Spherical Co-ordinates,” *Brain Topography,* vol. 6, no. 3, pp. 231 - 235, 1994.

Appendix A: Fundamentals of Non-Stationary Signal Processing

The material presented in this appendix are used to explain the fundamentals of neonatal EEG time-frequency processing and analysis, which are needed for Section 2.3, page 27.

A.1 Stationary and Nonstationary Processes

A random process $X(t)$ is said to be stationary to the Nth order if for any $t_1, t_2, \dots, t_N$, $and\ x_1, x_2, \dots, x_N$ Equation A.1 holds, where the probability is P, F_x is the probability density function, and t_0 is an arbitrary real constant [112].

$$F_x(x_1, \dots, x_N; t_1, \dots, t_N) = P(\{x(t_1) \le x_1, \dots, x(t_N) \le x_N\})$$
$$= P(\{x(t_1 + t_0) \le x_1, \dots, x(t_N + t_0) \le x_N\}) = F_x(x_1, \dots, x_N; t_1 + t_0, \dots, t_N + t_0) \quad \text{Equation A.1}$$

This process is said to be strictly stationary if it is stationary to the infinite order ($N \to \infty$). This implies that all moments are constants and equal through the random process, thus a process cannot be proven to be strictly stationary as infinite moments have to be calculated. Wide-sense stationary processes are stationary to the second order, thus the expectation of the process (first moment) is constant and the autocorrelation function R_{XX} (second moment) does not depend on time (Equation A.2 and Equation A.3) [112].

$$\mathrm{E}[X(t)] = \int_{-\infty}^{\infty} x\, f_x(x; t)\, dx = c \quad \text{Equation A.2}$$

$$R_{XX}(t_1, t_2) = R_{XX}(\tau) = \mathrm{E}[X(t_1).X(t_1 + \tau)]\,, \qquad \tau = |t_2 - t_1| \quad \text{Equation A.3}$$

Most real life signals are nonstationary for instance: ocean waves, atmospheric turbulence, economic time-series data, EEG, and much simpler example is a linear frequency modulated (*LFM*) signal. All of these examples share that both or at least one

of the first two moments changes with time. However, it is still possible to consider such signals as piece wise stationary for measurement and analysis purposes. This can be done by segmenting the nonstationary signal into N wide-sense stationary segments, thus stationary analysis tools can be used. This is very important as the Wiener-Khintchine theorem still holds for such small segments. Consequently, a time-varying power spectrum $S_X(t_i, f)$ for nonstationary signals can be obtained from the Fourier transform of the time-varying autocorrelation function $R_{X_i}(t, \tau)$. Thus creating a time-frequency representation (TFR) [1] [112].

$$S_X(t_i, f) = \mathcal{F}\{R_{X_i}(t, \tau)\} = \int_{-\infty}^{\infty} R_{X_i}(t, \tau)\, e^{-j2\pi f\tau}\, d\tau, \qquad i \in [1, N] \quad \text{Equation A.4}$$

A.2 Wigner-Ville Distribution

The Wigner Distribution (*WD*) $\mathcal{W}_s(t, f)$ was discovered by Eugene Wigner in 1932 in a quantum mechanical concept. It is a Quadratic Time-Frequency Distribution (*QTFD*) based on computing the Fourier transform of a signal kernel $K_s(t, \tau)$ which represents the Instantaneous Autocorrelation Function (*IAF*) of a signal $s(t)$ using Equation A.5 and Equation A.6 [1].

$$K_s(t, \tau) = s\left(t + \frac{\tau}{2}\right)\, s^*\left(t - \frac{\tau}{2}\right) \quad \text{Equation A.5}$$

$$\rho_s(t, f) = \int_{-\infty}^{\infty} s\left(t + \frac{\tau}{2}\right)\, s^*\left(t - \frac{\tau}{2}\right)\, e^{-j2\pi f\tau}\, d\tau = \mathcal{W}_s(t, f) \quad \text{Equation A.6}$$

The Wigner-Ville Distribution (*WVD*) $W_z(t, f)$ is an extension of the WD. It modifies the signal kernel to adapt the analytic associate $z(t)$ of $s(t)$. The analytic association is a realistic version of the real signal, where this association has no negative frequencies. This concept mimics real life signals where negative frequencies have no meaning as a consequence from the basic definition of frequency (number oscillations per unit time),

this quantity cannot be negative [1]. The analytic association can be calculated using Equation A.7 and Equation A.8, where $y(t)$ is the Hilbert transform of $s(t)$.

$$z(t) = s(t) + jy(t) \quad \text{Equation A.7}$$

$$y(t) = \mathcal{H}\{s(t)\} \quad \text{Equation A.8}$$

The signal kernel of the analytic associate is calculated using Equation A.9 and is used to compute the WVD using Equation A.10.

$$K_z(t,\tau) = z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right) \quad \text{Equation A.9}$$

$$\rho_z(t,f) = \int_{-\infty}^{\infty} z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f\tau}\, d\tau = W_z(t,f) \quad \text{Equation A.10}$$

Figure A.1 illustrates a real signal $s_1(t)$ and its analytic associate $z_1(t)$ in the frequency and time-frequency domains using WD and WVD. $s_1(t)$ is a LFM signal and can be described using Equation A.11. It is clear that the analytic associate has no negative frequencies, but only mimics the positive portion of the real signal frequency representation (Equation A.12). The WD shows two LFM signals on the positive and negative frequency axes respectively. On the other hand, the WVD only shows only one LFM on the positive frequency axis which again mimics real life signals [1].

$$s_1(t) = \cos(2\pi f_o t + \pi\alpha t^2) \quad \text{Equation A.11}$$

$$z(t) = s_1(t) + j\mathcal{H}[s_1(t)] = e^{\,j[2\pi f_o t + \pi\alpha t^2]} \quad \text{Equation A.12}$$

Noisy terms are added on the centre of the WD. These noisy additions are called inner artifacts and they act as embedded noise on the TFR. Figure A.2 illustrate the WVD of a multicomponent signal consisting of a LFM and a Quadratic Frequency Modulated signal (*QFM*). The TFR consists of many artifacts that did not exist in the original signal. These artifacts are contoured with red and yellow blubs. The real information of the signal are contoured in green and are called the auto-terms of the TFR where they

describe the instantaneous frequency (*IF*) law of the multicomponent signal. Artifacts that are between two components are called outer-terms or cross-terms (contoured in red) and are generated because of the nonlinear nature of TFDs. On the other hand, artifacts that are generated within the signal are called inner-terms (contoured in yellow) and are generated because of the nonlinear IF law of the signal [1].

Inner and outer-terms alternate in sign as we move normal to the IF law in the t-f plane; this is a characteristic feature of the inner and outer terms and will be used in cross-terms minimisation. Note that MATLAB default linear colour-maps are not adequate for the human eye to see small changes in colour, thus a different linear colour-map is used and illustrated in Figure A.3. This colour-map will be used further in this thesis. More details on colour-maps and human eye perception can be found in [113].

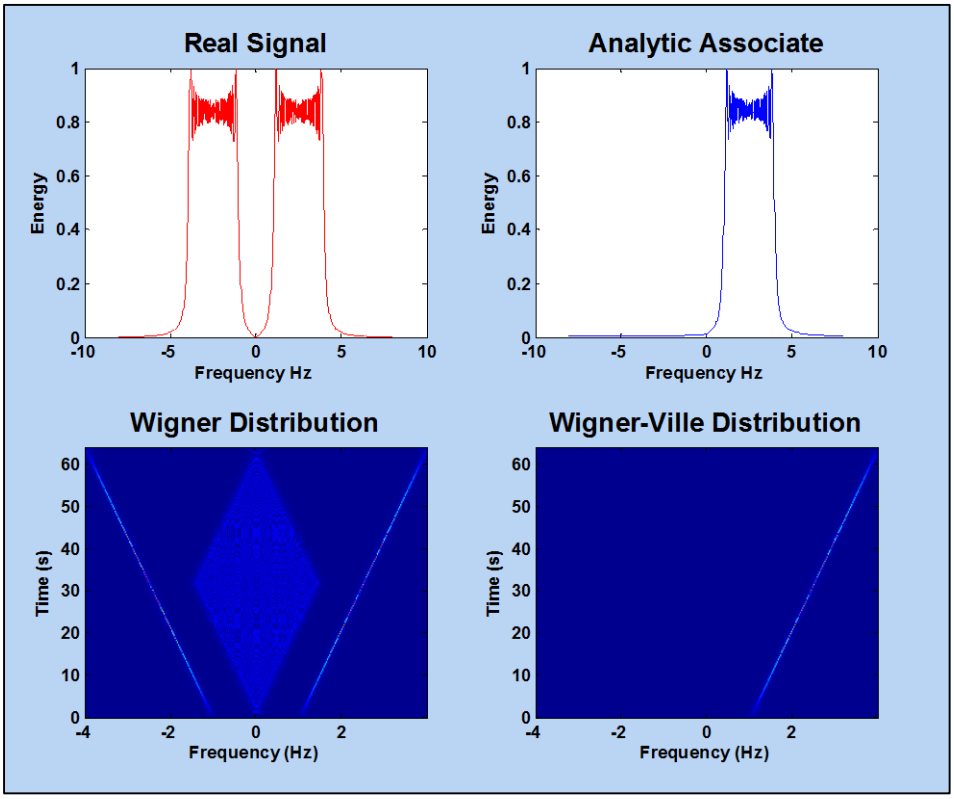

Figure A.1: Real LFM and its analytic associate frequency and time-frequency representations using WD and WVD. The starting frequency of the LFM is 1 Hz and ends at 4 Hz. The signal duration is 64 seconds and the sampling frequency is 16 Hz

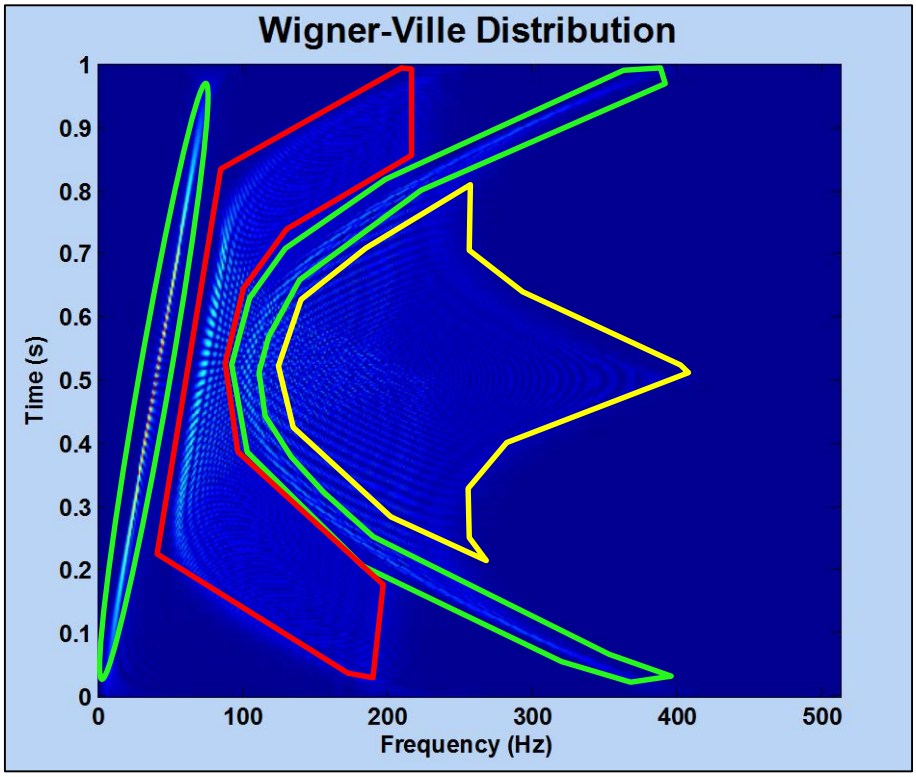

Figure A.2: WVD of a multicomponent signal consisting of a LFM and a QFM. The LFM frequency starts at 1 and finishes at 40 Hz, while the QFM frequency starts at 400 goes to 100 Hz at ends at 400 Hz. The signal duration is 1 second using a sampling frequency of 512 Hz

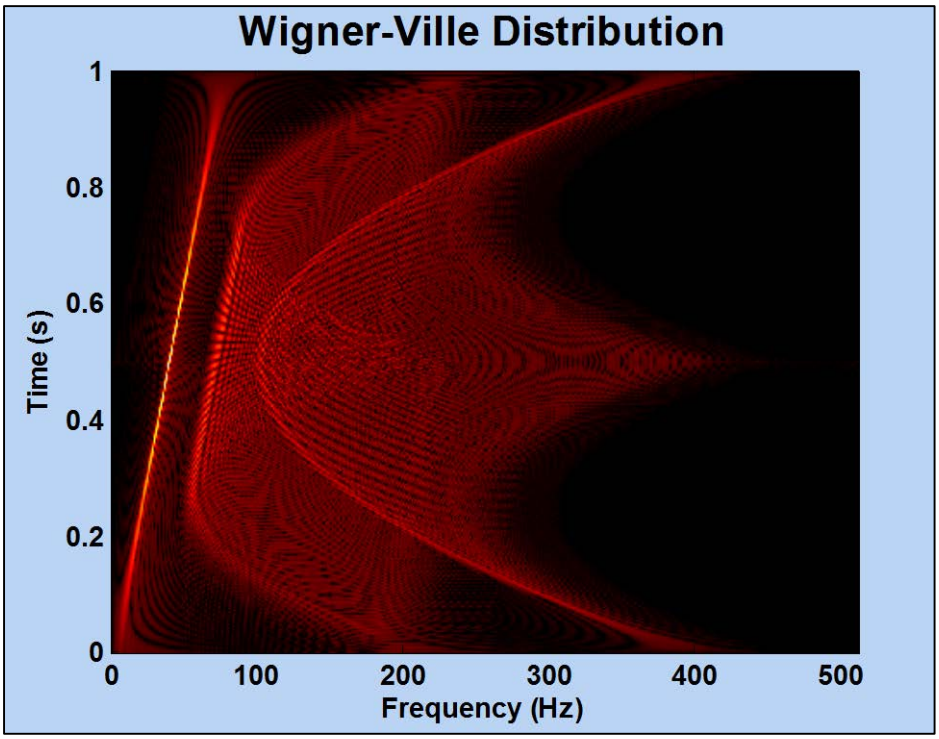

Figure A.3: WVD of a multicomponent signal consisting of a LFM and a QFM. This is the same as Figure A.2 but with different colour map

A.3 Formulation of Quadratic Time-Frequency Distributions

The WVD defined in Equation A.10 gives ideal concentration for mono-component LFM signals, but it produces undesired cross-terms for non-linear frequency modulated or multicomponent signals (Figure A.1 and Figure A.3). Cross-terms can be reduced by convolving the WVD with a two dimensional time-frequency kernel $\gamma(t, f)$, which can be described using Equation A.13.

Equation A.13 also represents the general time-frequency formulation of QTFDs. The 2D smoothing reduces cross-terms contribution, but it blurs the auto-terms as well. Consequently, kernels must be designed to achieve the best trade-off between cross-terms minimisation and auto-terms resolution [1].

$$\rho(t, f) = \gamma(t, f) *_t *_f W_z(t, f) \qquad \text{Equation A.13}$$

Computation of QTFDs based on Equation A.13 requires one Fourier transform and two convolutions along the time and frequency axes. In addition to this general expression, QTFDs can be expressed starting from a time-lag formulation, doppler-lag formulation or a doppler-frequency formulation [1].

Equation A.13 can be described in the time-lag formulation by using Equation A.14. The convolution operation along the frequency domain axis is replaced with multiplication along the lag axis, where $G(t, \tau)$ is called the time-lag kernel of the TFD. This formulation requires only one convolution with respect to time and one Fourier transform from lag domain to frequency domain [1].

Another formulation can be done in the doppler-frequency domain as in Equation A.15. In this case the time convolution is replaced by a multiplication in the Doppler axis, where $\mathcal{G}(v, f)$ is called the doppler-frequency kernel of the TFD and $k_z(v, f)$ is referred to as the Spectral Autocorrelation Function (*SAF*).

Lastly, Equation A.13 can be also described in the doppler-lag formulation using Equation A.16. The time convolution of Equation A.14 is replaced by multiplication in the Doppler domain, where $A_z(v,\tau)$ represents the ambiguity function and $g(v,\tau)$ is the doppler-lag kernel (Equation A.17 and Equation A.18). This formulation is widely used for designing high resolution TFDs as it allows entering the filter specifications directly in the formulation which is equivalent to designing a 1D filter in the frequency domain. The relationship between all formulations is depicted in Figure A.4 [1].

$$\rho(t,f) = \int G(t,\tau) *_t K_z(t,\tau)\, e^{-j2\pi f\tau}\, d\tau = \int R_z(t,\tau)\, e^{-j2\pi f\tau}\, d\tau \qquad \text{Equation A.14}$$

$$\rho(t,f) = \int \mathcal{G}(v,f) *_f k_z(v,f)\, e^{j2\pi vt}\, dv = \int r(v,f)\, e^{j2\pi vt}\, dv \qquad \text{Equation A.15}$$

$$\rho(t,f) = \iint g(v,\tau)\, A_z(v,\tau) e^{j2\pi(vt-f\tau)}\, d\tau dv = \iint \mathcal{A}_z(v,\tau)\, e^{j2\pi(vt-f\tau)}\, d\tau dv \qquad \text{Equation A.16}$$

$$A_z(v,\tau) = \int K_z(t,\tau)\, e^{-j2\pi tv}\, dt \qquad \text{Equation A.17}$$

$$g(v,\tau) = \int G(t,\tau)\, e^{-j2\pi tv}\, dt \qquad \text{Equation A.18}$$

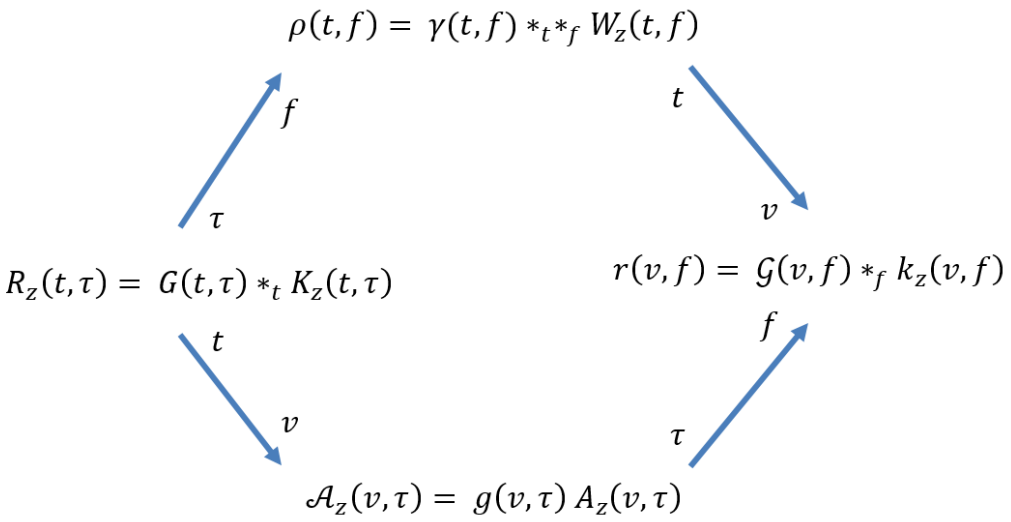

Figure A.4: TFD formulation relationships in several domains (Redrawn from [1] and [63])

As stated before inner and outer-terms sign alternate as we move normal to the IF law in the t-f plane, and their alternation frequency is higher than the auto-terms. Consequently, their contribution in the doppler-lag domain would exist far away from the origin where low alternating components (auto-terms) resign. Figure A.5 illustrates the ambiguity domain of the multicomponent signal in Figure A.3. The contribution of auto-terms is contoured in green while the rest is contributions from inner and outer-terms. Designing high resolution TFDs becomes a matter of designing a 2D filter in the time-frequency domain that is able to filter auto-terms from inner and outer-terms. This can be achieved by designing a 2D filter in the ambiguity domain conserving the shape of the auto-terms contribution [1].

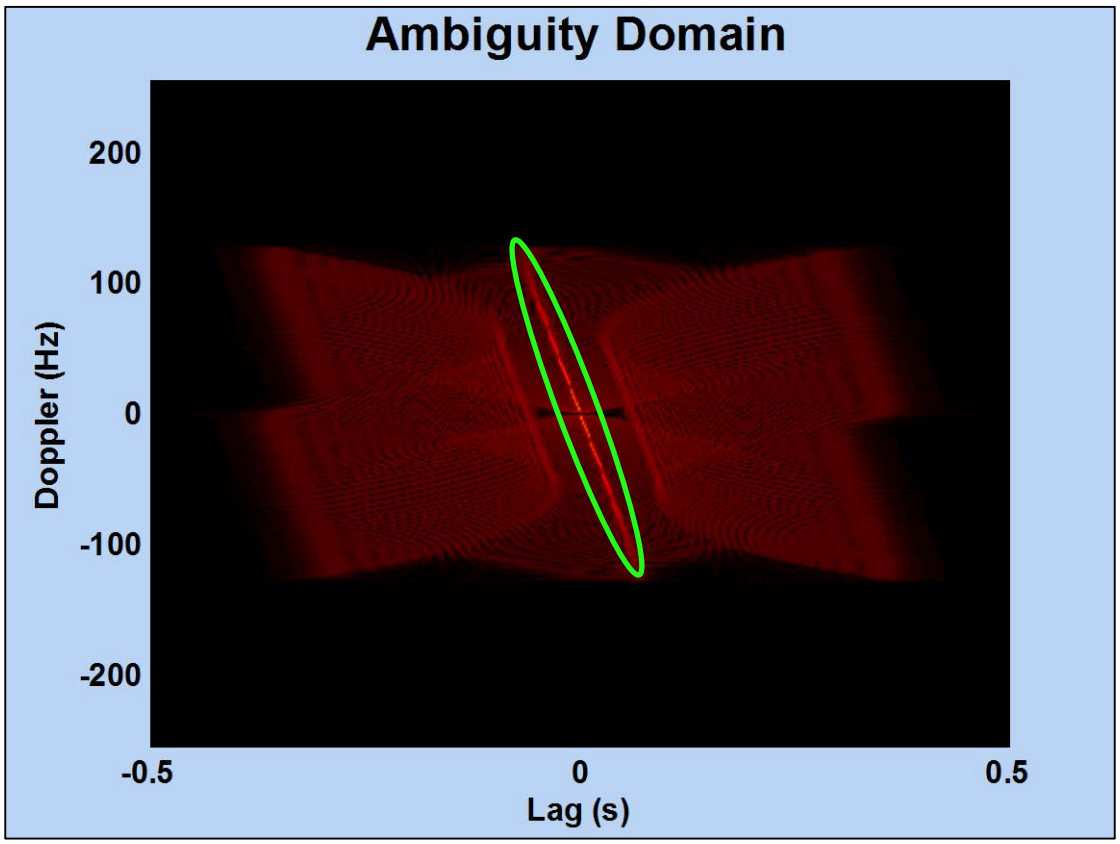

Figure A.5: Doppler-lag or Ambiguity domain of the signal depicted in Figure A.3

A.4 High Resolution Quadratic Time-Frequency Distributions

High resolution QTFDs can be computed using separable kernels, where $\gamma(t,f)$ can be expressed using Equation A.19. The meaning and significance of separable kernels is that QTFDs are defined and designed by smoothing the WVD in t and then in f. The shape and size of $g_1(t)$ or $G_2(f)$ defines the smoothing along the time or frequency axis respectively. The relationship between all formulations is depicted in Figure A.6.

$$\gamma(t,f) = g_1(t)G_2(f) \qquad \text{Equation A.19}$$

Examples of high resolution TFDs are: the B-Distribution (*BD*), Modified B-Distribution (*MBD*), Extended Modified B-Distribution (*EMBD*), Compact Support Kernel (*CSK*), and Extended Compact Support Kernel (*ECK*) [1] [27] [63] [77] [114].

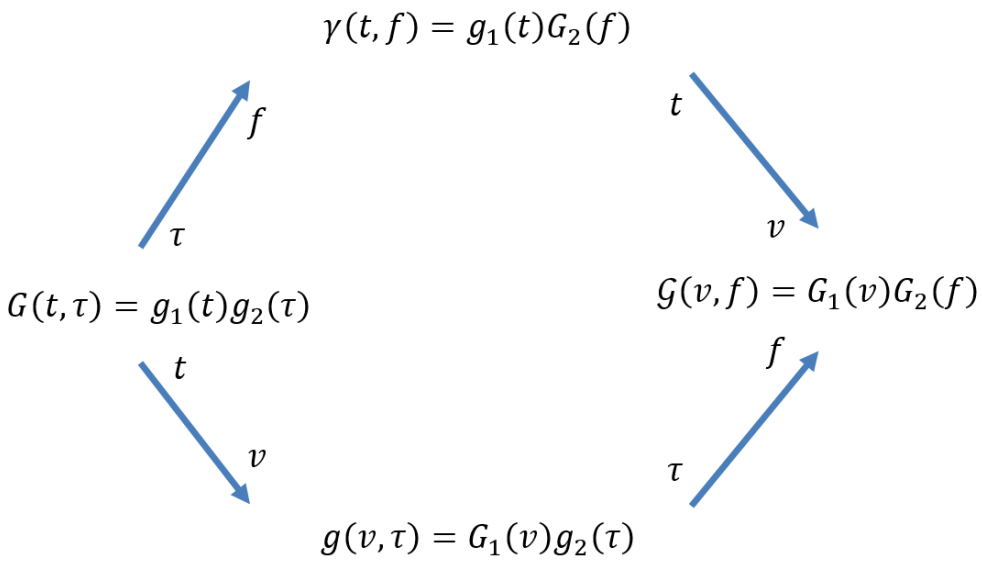

Figure A.6: Separable kernels formulation relationships in several domains (Redrawn from [1] and [63])

A.4.1 B-Distribution

The BD kernel expression in the ambiguity domain is depicted in Equation A.20, where τ, v, and β are bounded to ensure that the BD contains a low-pass filter. It illustrates the product between a low-pass filter $G_1(v)$ and high-pass filter $g_2(\tau)$. The BD produces high resolution TFD for small values of β. The disadvantage of the BD is that a zero at the origin appears as a consequence of the high pass filtering along the lag axis. This results in degradation of the resolution for specific types of signals [63].

$$g(v,\tau) = G_1(v)g_2(\tau) = |\tau|^{\beta} \frac{|\Gamma(\beta + j\pi v)|^2}{2^{1-2\beta}\Gamma(2\beta)} \qquad \text{Equation A.20}$$

$$|v| \leq 0.5, |\tau| \leq 0.5, 0 \leq \beta \leq 1$$

A.4.2 Modified B-Distribution

The MBD design overcomes some of the problems illustrated in the BD. Its kernel expression in the ambiguity domain is depicted in Equation A.21. The kernel of the MBD includes only the Doppler window of the BD kernels so as to avoid the distortion caused by high pass filtering along the lag-axis. As a consequence the MBD gives the highest energy concentration only for signals whose IF does not vary rapidly with time, such as EEG seizure signals and LFMs [1] [63].

$$g(v,\tau) = G_1(v) = \frac{|\Gamma(\beta + j\pi v)|^2}{\Gamma^2(\beta)}, \qquad |v| \leq 0.5, 0 \leq \beta \leq 1 \qquad \text{Equation A.21}$$

A.4.3 Extended Modified B-Distribution

The EMBD was designed as an improvement to the MBD. It extends the MBD by applying its kernel filter along both lag and Doppler axes as depicted in Equation A.22. The lengths of the lag and Doppler windows are controlled independently by α and β respectively. The extra degree of freedom in the formulation of the EMBD allows to

independently adjust the lengths of the windows along both lag and Doppler axes. This advantage makes it a useful tool in analysing real life signals such as; EEG signals with seizure and EEG spike signals [1] [63].

$$g(v,\tau) = G_1(v)g_2(\tau) = \frac{|\Gamma(\beta + j\pi v)|^2}{\Gamma^2(\beta)} \frac{|\Gamma(\alpha + j\pi\tau)|^2}{\Gamma^2(\alpha)}, \qquad \text{Equation A.22}$$

$$-0.5 \le v \le 0.5, \quad -0.5 \le \tau \le 0.5, \quad 0 \le \beta \le 1, \quad 0 \le \alpha \le 1$$

A.4.4 Compact Support Kernel

The CSK is designed to vanish outside a given range in the ambiguity domain. Consequently, it does not have infinite length and does not require rectangular window truncation which may cause loss of information. TFDs utilising such compact kernels have been shown to outperform other kernel-based methods in terms of their ability to supress cross-terms and to retain auto-terms resolution. Such high-resolution performance is achieved by the kernel finite length, and its flexibility in adjusting both the shape and the size independently. The CSK expression in the ambiguity domain is depicted in Equation A.23, Equation A.24, and Equation A.25. These equations show that the shape and size of both Doppler and lag windows are determined by the parameters c and D respectively. However, it restricts the Doppler and lag windows to be of same length, so that smoothing along time or frequency cannot be adjusted independently.

$$g(v,\tau) = G_1(v)g_2(\tau) = \begin{Bmatrix} e^{2c}e^{\frac{cD^2}{v^2-D^2}+\frac{cD^2}{\tau^2-D^2}}, & v^2 < D^2 \text{ and } \tau^2 < D^2 \\ 0 \quad , & otherwise \end{Bmatrix} \qquad \text{Equation A.23}$$

$$G_1(v) = \begin{Bmatrix} e^c e^{\frac{cD^2}{v^2-D^2}}, & v^2 < D^2 \\ 0 \quad , & otherwise \end{Bmatrix} \qquad \text{Equation A.24}$$

$$g_2(\tau) = \left\{ \begin{array}{cc} e^c e^{\frac{cD^2}{\tau^2 - D^2}}, & \tau^2 < D^2 \\ 0 \quad , & otherwise \end{array} \right\}$$ Equation A.25

A.4.5 Extended Compact Support Kernel

The CSK shape and size adjusting pitfalls limits its application to the analysis of signals whose energy is homogeneously distributed in the ambiguity domain. Consequently, the CSK TFD cannot deal optimally with signals like EEG seizure signals, whose IF is almost parallel to the time axis, or EEG spike signals, whose IF is almost parallel to the frequency axis [1] [63].

Fortunately, the CSK can be extended by modifying the formulation of Doppler and lag windows such that their lengths can be adjusted independently as expressed in Equation A.26, Equation A.27, and Equation A.28. These equations show that the shape and size of both Doppler and lag windows are determined independently by the parameters c, D, and E respectively, which leads to the formulation of the Extended Compact Support Kernel (*ECK*). The TFD defined by the ECK kernel may be referred to as ECK TFD, but this abbreviation is further shortened to Compact Kernel Distribution (*CKD*) for simplicity [1] [63].

$$g(v,\tau) = \left\{ \begin{array}{cc} e^{2c} e^{\frac{cD^2}{v^2 - D^2} + \frac{cE^2}{\tau^2 - E^2}}, & if \ |v| < D \ and \ |\tau| < E \\ 0 \quad , & otherwise \end{array} \right\}$$ Equation A.26

$$G_1(v) = \left\{ \begin{array}{cc} e^c e^{\frac{cD^2}{v^2 - D^2}}, & if \ |v| < D \ a \\ 0 \quad , & otherwise \end{array} \right\}$$ Equation A.27

$$g_2(\tau) = \left\{ \begin{array}{cc} e^c e^{\frac{cE^2}{\tau^2 - E^2}}, & if \ |\tau| < E \\ 0 \quad , & otherwise \end{array} \right\}$$ Equation A.28

The ECK time-frequency formulation can be derived using the saddle-point integration method. Starting from Equation A.29, Equation A.30, Equation A.31, and Equation A.32 we can see that the time-frequency formulation requires a Fourier transform from lag to frequency and an inverse Fourier transform from Doppler to time (Equation A.33). Thus, formulations of $g_1(t)$ and $G_2(f)$ are the only unknowns and can be derived independently (Equation A.34 and Equation A.35).

$$g(\nu,\tau) = \lambda\, G_1(\nu)\, g_2(\tau), \qquad if\ \ |\nu| < D\ \ and\ \ |\tau| < E \qquad \text{Equation A.29}$$

$$\lambda = \exp(2c) \qquad \text{Equation A.30}$$

$$G_1(\nu) = \exp\left[\frac{cD^2}{\nu^2 - D^2}\right], \qquad if\ \ |\nu| < D \qquad \text{Equation A.31}$$

$$g_2(\tau) = \exp\left[\frac{cE^2}{\tau^2 - E^2}\right], \qquad if\ |\tau| < E \qquad \text{Equation A.32}$$

$$\gamma(t,f) = \lambda\, g_1(t)\, G_2(f) \qquad \text{Equation A.33}$$

$$g_1(t) = \mathcal{F}^{-1}_{\nu\to t}\{G_1(\nu)\} \qquad \text{Equation A.34}$$

$$G_2(f) = \mathcal{F}_{\tau\to f}\{g_2(\tau)\} \qquad \text{Equation A.35}$$

The derivation starts by modifying Equation A.31 into the form of Equation A.36, thus changing the Fourier transform expression into Equation A.37, where $\phi(\nu) = \exp(j2\pi t\nu)$.

$$G_1(\nu) = \exp\left[\frac{-cD^2}{D^2 - \nu^2}\right] = \exp\left[\frac{-c}{1 - \left(\frac{\nu}{D}\right)^2}\right], \qquad if\ \ |\nu| < D \qquad \text{Equation A.36}$$

$$g_1(t) = \mathcal{F}^{-1}_{\nu\to t}\{G_1(\nu)\} = \langle G_1(\nu), \phi(\nu)\rangle = \int_{-1}^{1} G_1(\nu)\, \phi(\nu)\ d\nu \qquad \text{Equation A.37}$$

By letting $v = \psi D \;\; , \;\; dv = \; Dd\psi \;\; , \;\; \psi \in (-1, 1)$ Equation A.36 expression changes into Equation A.38, and Equation A.37 changes into Equation A.39, where $\phi(\psi) = \exp(j2\pi t\psi D) = \; \exp(k\psi)$.

$$G_1(\psi) = \; \exp\left[\frac{-c}{1-\psi^2}\right], \qquad if \;\; |\psi| < 1 \qquad \text{Equation A.38}$$

$$g_1(t) = \mathcal{F}_{v\to t}^{-1}\{G_1(\psi)\} = \langle G_1(\psi), \phi(\psi)\rangle = D\int_{-1}^{1} G_1(\psi)\;\phi(\psi)\;\; d\psi \qquad \text{Equation A.39}$$

Since $G_1(\psi)$ is an even function Equation A.39 can be expressed as;

$$g_1(t) = 2D\;\text{Re}\left\{\int_{0}^{1} G_1(\psi)\;\phi(\psi)\;\; d\psi\right\} \qquad \text{Equation A.40}$$

The inner term of the integration can be expressed by Equation A.41, where $\alpha(\psi) = k\psi - \frac{c}{1-\psi^2}$, which changes the form of Equation A.40 into Equation A.42.

$$G_1(\psi)\;\phi(\psi) = \exp[\alpha(\psi)] \qquad \text{Equation A.41}$$

$$g_1(t) = 2D\;\text{Re}\left\{\int_{0}^{1} \exp[\alpha(\psi)]\;\; d\psi\right\} \qquad \text{Equation A.42}$$

By letting $\psi = 1 - x \;\; , \;\; d\psi = -dx \;\; , \;\; x \in (1, 0)$ $\alpha(\psi)$ changes into $\alpha(x)$ which is expressed by Equation A.43, and $g_1(t)$ changes into the form expressed by Equation A.44.

$$\alpha(x) = k - kx - \frac{c}{1-(1-x)^2} = k - kx - \frac{c}{(2-x)x} \qquad \text{Equation A.43}$$

$$g_1(t) = 2D\;\text{Re}\left\{\int_{1}^{0} -\exp[\alpha(x)]\;\; dx\right\} = 2D\;\text{Re}\left\{\int_{0}^{1} \exp[\alpha(x)]\;\; dx\right\} \qquad \text{Equation A.44}$$

Now, by using rational regression $\frac{c}{(2-x)x}$ can be approximated as $\frac{cP_1x+cP_2}{x+Q_1}$, where the mean values of P_1, P_2, and Q_1 are 0.35524, 0.5, and 0 respectively, thus modifying Equation A.43 expression into Equation A.45 (Figure A.7). The goodness of fit for this rational regression is 0.9999926, which proves the feasibility of this approximation.

$$\alpha(x) \approx k - kx - \frac{c}{2x} - 0.35524c$$ Equation A.45

By using Taylor Series approximation up to the second order derivative, $\alpha(x)$ can be approximated as $\hat{\alpha}(x) = \alpha(x_0) + \alpha'(x_0)(x - x_0) + \frac{1}{2}\alpha''(x_0)(x - x_0)^2$, thus changing Equation A.44 into Equation A.46. Choosing x_0 as the saddle point makes $\alpha'(x_0) = 0$.

$$g_1(t) = 2D \text{ Re}\left\{\int_0^1 \exp[\hat{\alpha}(x)]\ dx\right\}$$ Equation A.46

$\hat{\alpha}(x)$ can be expressed as $\hat{\alpha}(x) = \alpha(x_0) + \frac{1}{2}\alpha''(x_0)(x - x_0)^2$, where $\alpha'(x) = -k + \frac{c}{2(x)^2}$, $\alpha''(x) = \frac{-c}{x^3}$, and $\alpha'(x_0) = -k + \frac{c}{2(x_0)^2} = 0$, thus $x_0 = \pm\sqrt{\frac{c}{2k}}$

Since x is changing from 0 to 1, $x_0 = \sqrt{\frac{c}{2k}}$ making $\alpha(x_0) = k - \sqrt{2ck} - 0.35524c$ and $\alpha''(x_0) = \frac{-c}{\left(\sqrt{\frac{c}{2k}}\right)^3} = -\sqrt{\frac{(2k)^3}{c}}$, thus;

$$\hat{\alpha}(x) = k - \sqrt{2ck} - 0.35524c - 2\sqrt{\frac{k^3}{c}}\left(x - \sqrt{\frac{c}{2k}}\right)^2$$ Equation A.47

By replacing the variables of Equation A.47 with the dummy variables: $\gamma = k - \sqrt{2ck} - 0.35524c$, $\sigma = 2\sqrt{\frac{k^3}{c}}$, $\mu = \sqrt{\frac{c}{2k}}$, and $k = j2\pi tD$, it can be expressed as $\hat{\alpha}(x) = \gamma - \sigma(x - \mu)^2$, thus changing Equation A.46 into the form of Equation A.48.

$$g_1(t) = 2D \text{ Re}\left\{\exp(\gamma)\int_0^1 \exp[-\sigma(x - \mu)^2]\ dx\right\}$$ Equation A.48

By knowing that the integration of $\exp[-a(x)^2]$ or $\exp[-\sigma(x - \mu)^2]$ is equal to $\sqrt{\frac{\pi}{a}}$ for $a > 0$, $g_1(t)$ can be solved as; $\boldsymbol{g_1(t) = 2D \text{ Re}\left\{\exp(\gamma)\sqrt{\frac{\pi}{\sigma}}\right\}}$

Using the same analogy and steps of $g_1(t)$, derivation $G_2(f)$ can be expressed as;

$$\boldsymbol{G_2(f) = 2E\ \mathrm{Re}\left\{\exp(\zeta)\sqrt{\frac{\pi}{\varepsilon}}\right\}},\quad \text{where}\quad \zeta = \eta - \sqrt{2c\eta} - 0.35524c,\ \ \varepsilon = -2\sqrt{\frac{\eta^3}{c}}\ ,$$

and $\eta = -j2\pi fE$, thus expanding Equation A.33 into Equation A.49.

$$\boldsymbol{\gamma(t,f) = 4DE\left[\mathrm{Re}\left\{\exp(\gamma)\sqrt{\frac{\pi}{\sigma}}\right\}\mathrm{Re}\left\{\exp(\zeta)\sqrt{\frac{\pi}{\varepsilon}}\right\}\right]} \qquad \text{Equation A.49}$$

One problem arises immediately in this approximation, is the infinite value at the centre of the time-frequency approximation (t = 0 and f = 0). In this work, the centre value is simply assumed to be zero, which disturbs the approximation quality and calls for further work.

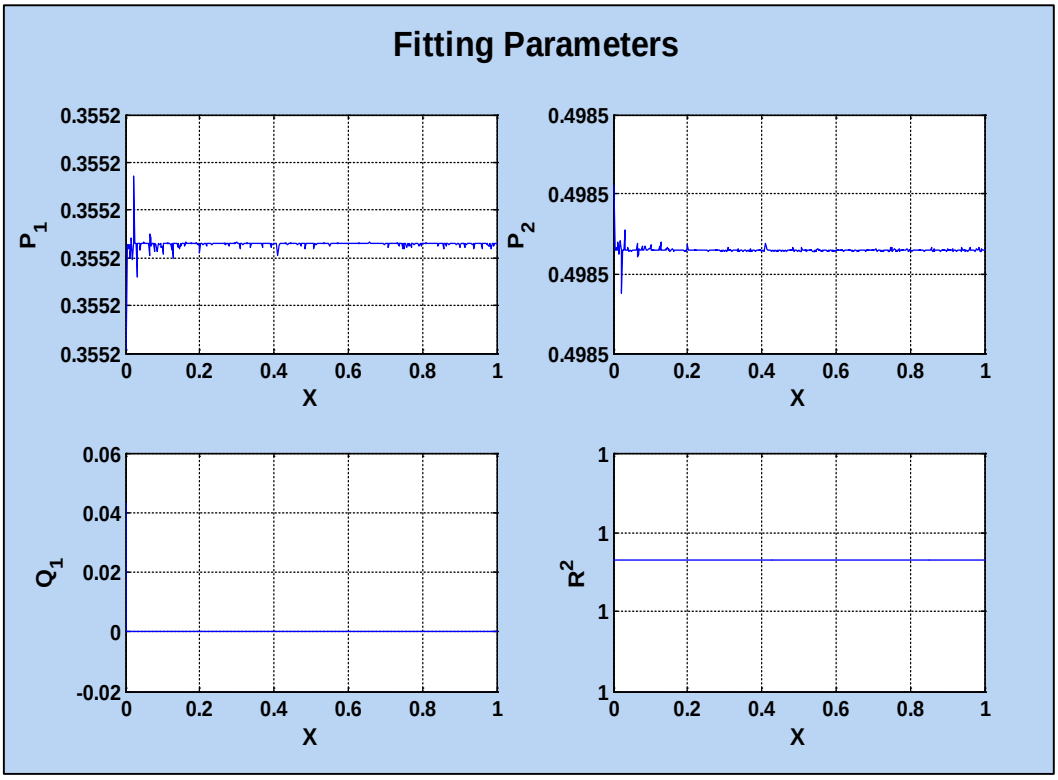

Figure A.7: Rational regression computed parameters

Figure A.8, Figure A.9, Figure A.10, Figure A.11, Figure A.12, and Figure A.13 illustrate the approximation outcomes along with the real CKD outputs in the time-

frequency and doppler-lag domains for different C, D and E parameters. Note that, high alternations are imposed on the time-frequency approximation, but the general morphology of the real CKD is followed.

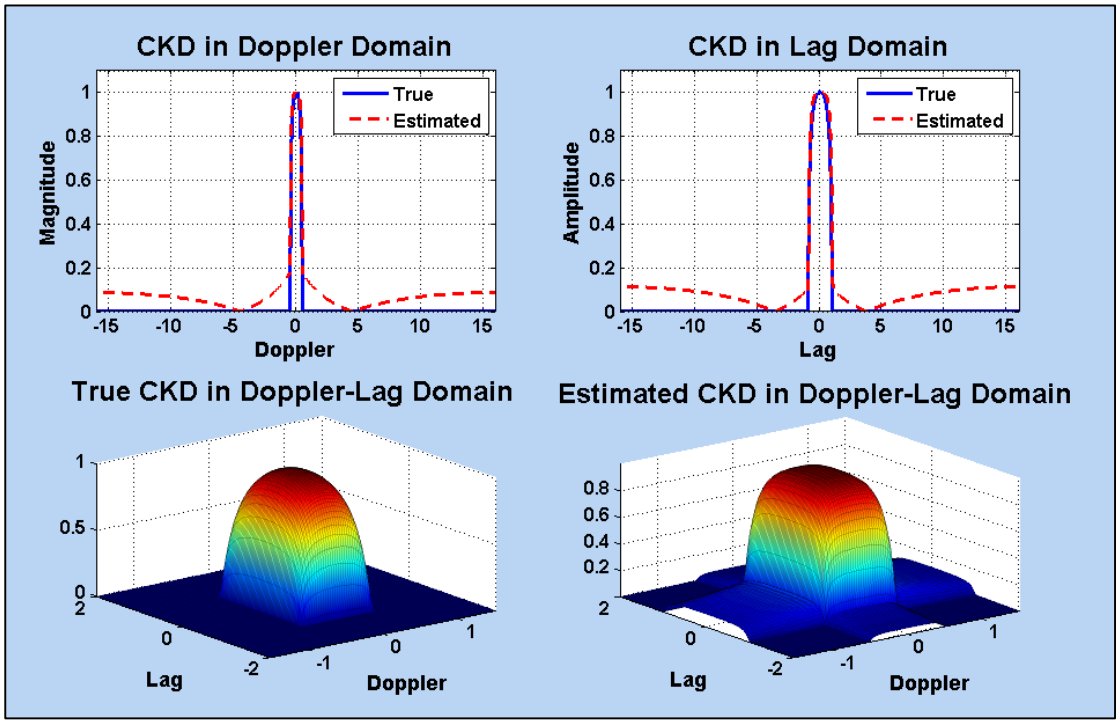

Figure A.8: CKD approximation resultant doppler-lag domain. C = 0.2, D = 0.5, and E = 1. Sampling frequency is 32 Hz, and time duration is 32 seconds

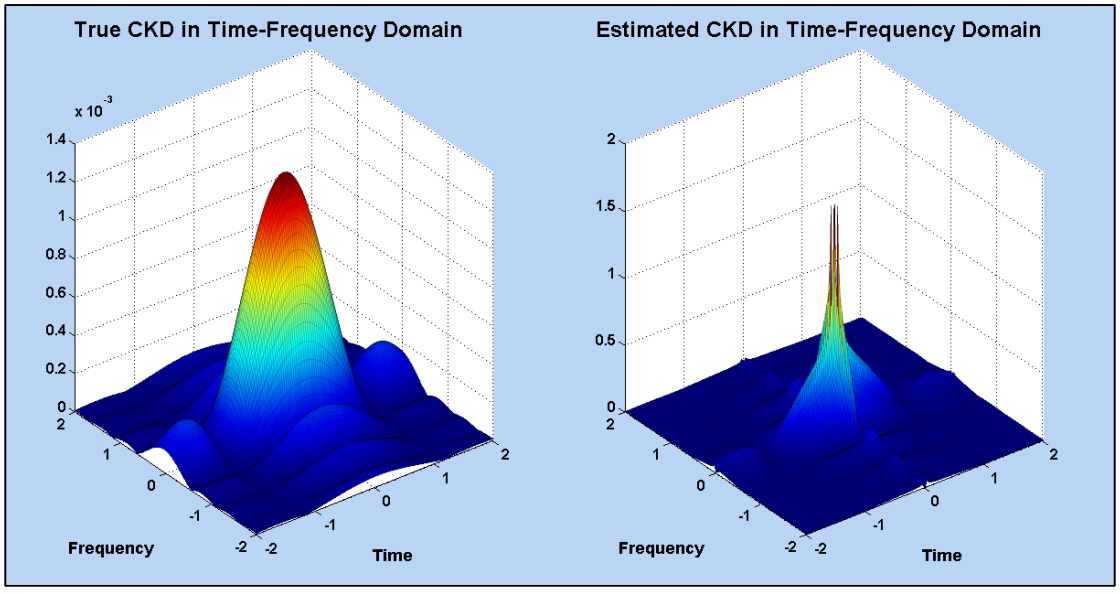

Figure A.9: CKD time-frequency approximation. C = 0.2, D = 0.5, and E = 1. Sampling frequency is 32 Hz, and time duration is 32 seconds

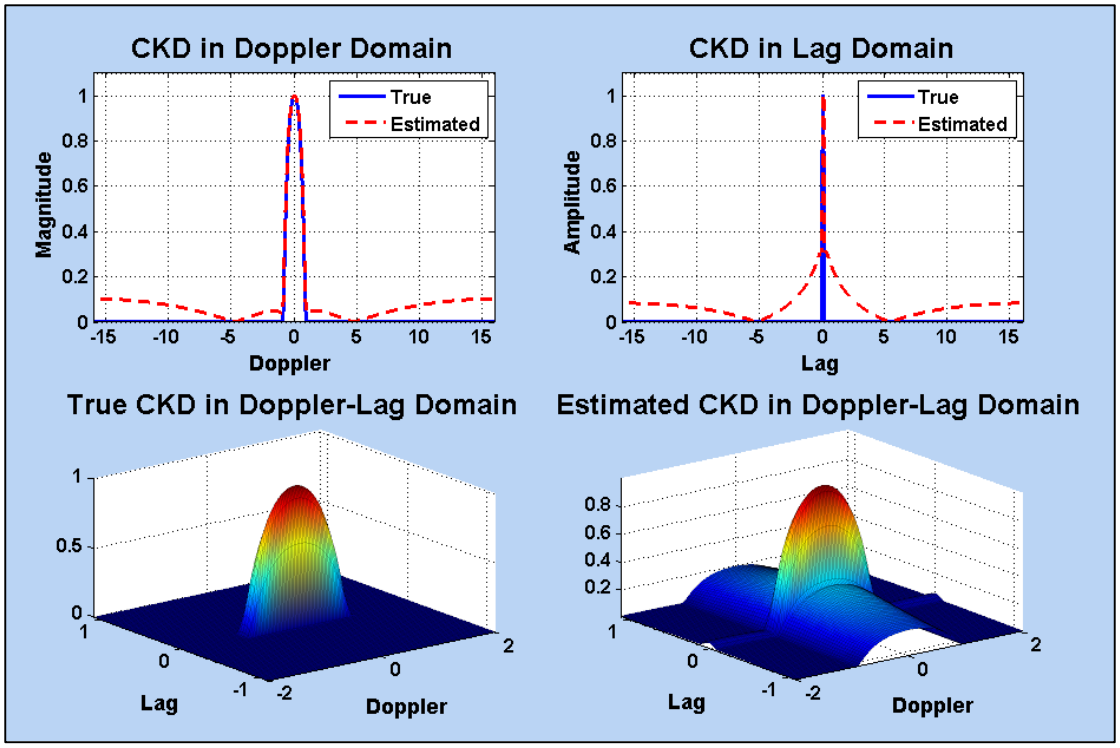

Figure A.10: CKD approximation resultant doppler-lag domain. C = 0.8, D = 1, and E = 0.1. Sampling frequency is 32 Hz, and time duration is 32 seconds

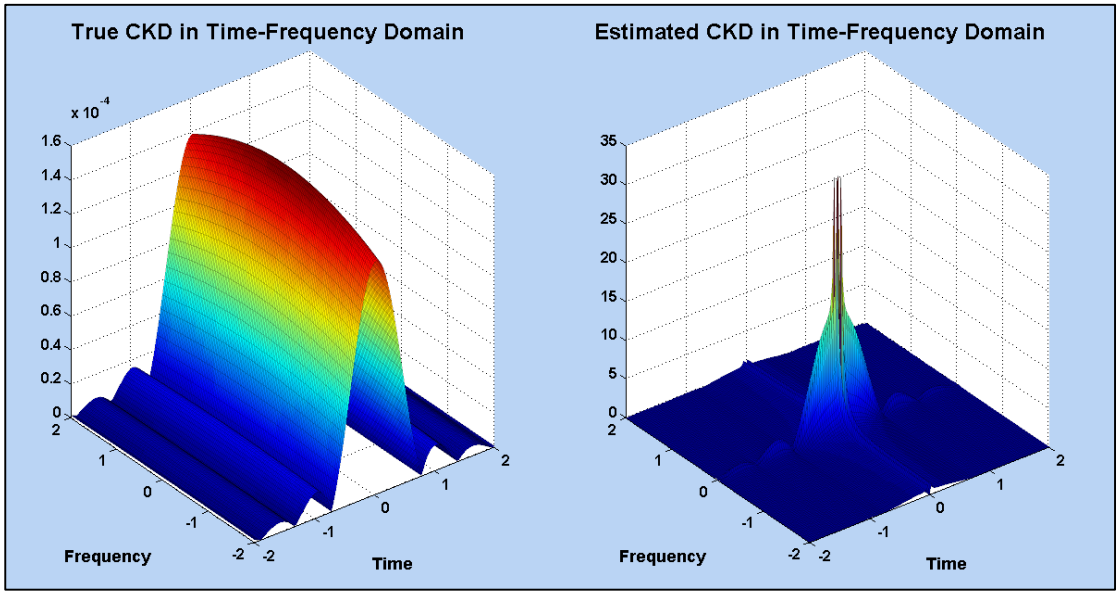

Figure A.11: CKD time-frequency approximation. C = 0.8, D = 1, and E = 0.1. Sampling frequency is 32 Hz, and time duration is 32 seconds

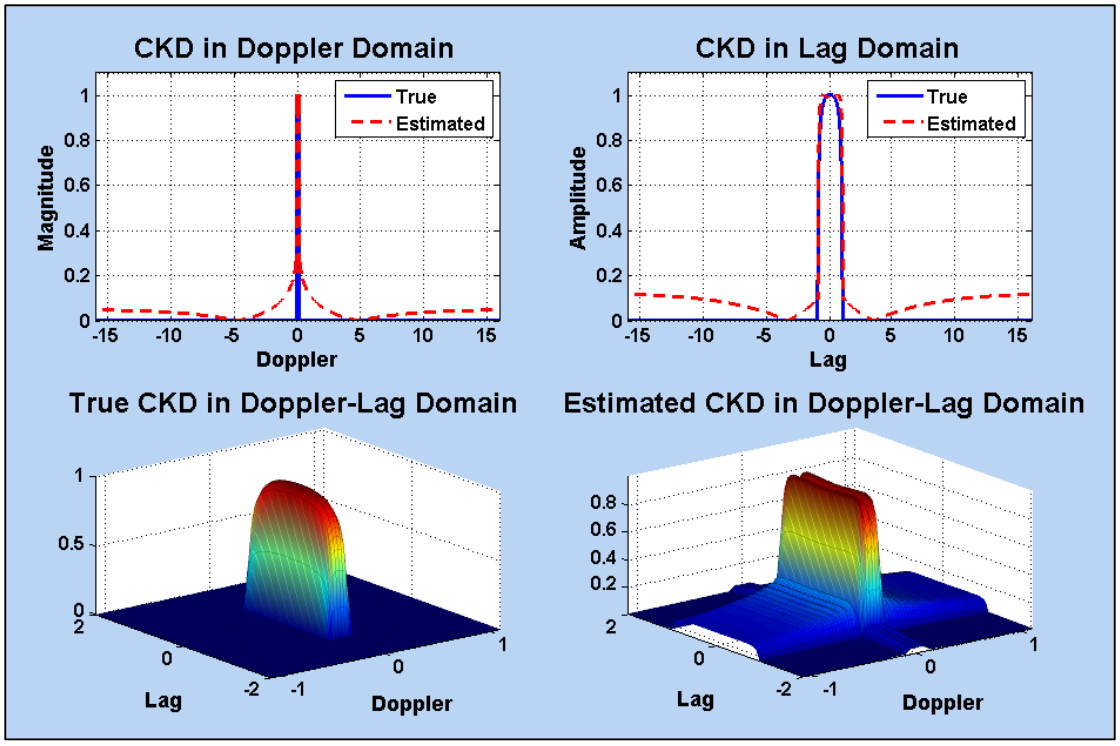

Figure A.12: CKD approximation resultant doppler-lag domain. C = 0.1, D = 0.1, and E = 1. Sampling frequency is 32 Hz, and time duration is 32 seconds

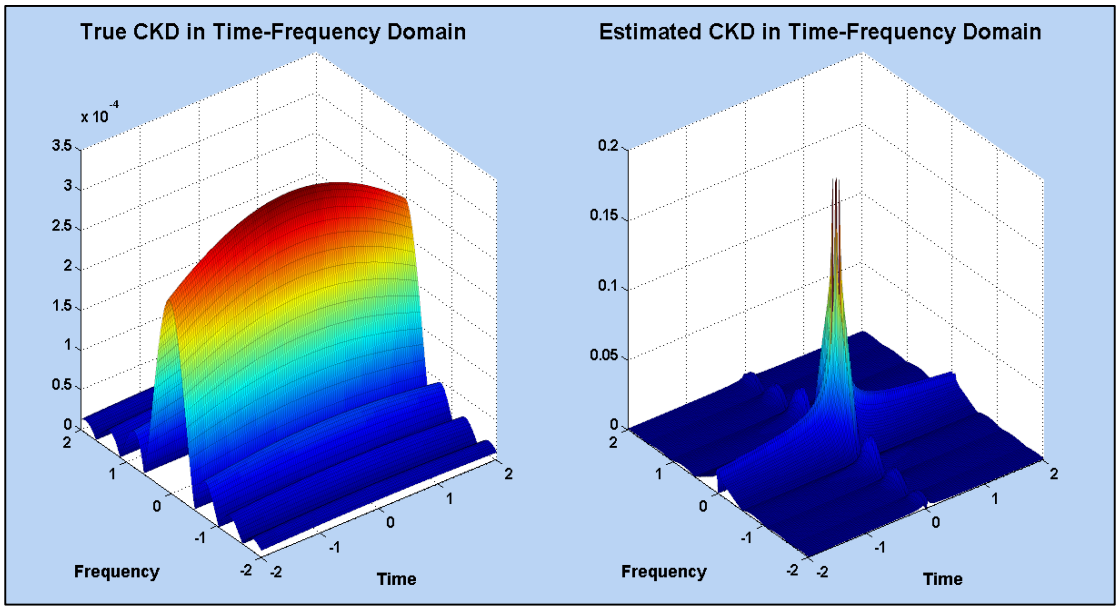

Figure A.13: CKD time-frequency approximation. C = 0.1, D = 0.1, and E = 1. Sampling frequency is 32 Hz, and time duration is 32 seconds

Figure A.14 illustrates the Normalised Root Mean Square Error (*NRMSE*) distribution of the approximated ECK doppler-lag domain transformation using different C, D, and E values. Regions that show NRMSE values below 50% are contoured in black. It can be seen that the approximation is most feasible when *D* and *E* parameters are below 1, but the feasibility range increases when C approaches 1. However, the error distributions show large values, which is due to the high alternations in the time-frequency domain. This again calls for further work to develop a more practical approximation.

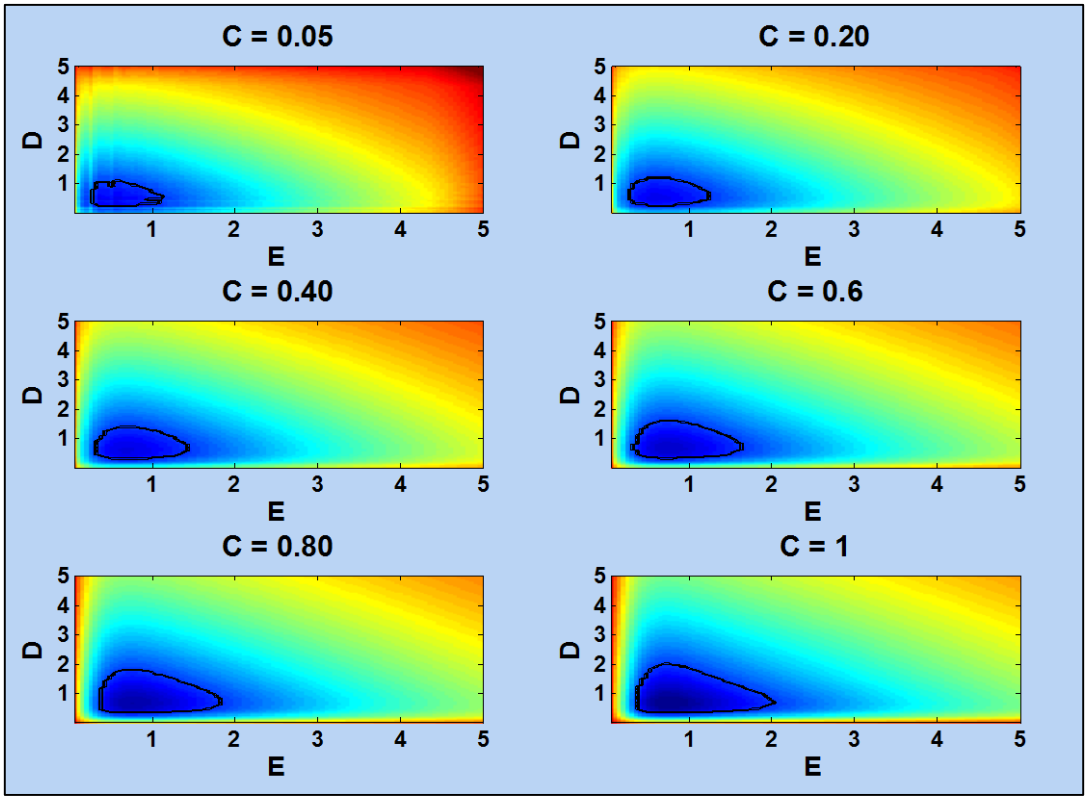

Figure A.14: CKD approximation resultant doppler-lag domain NRMSE distribution. Plot colours change from blue to red reflecting 40% to 100% NRMSE percentages. Errors that are less than 50% are contoured in black. Sampling frequency is 16 Hz, and time duration is 16 seconds

A.4.6 Performance Assessment

Figure A.15 and Figure A.16 illustrate the WVD, BD, MBD, EMBD, CSK, and ECK of the multicomponent signal depicted in Figure A.3, along with the ambiguity domains of each result. It can be clearly seen that the EMBD and the ECK give the highest energy concentration for the QFM while others failed to do so. It also supresses most of the cross-terms as depicted in the TFR and in the ambiguity domain. The ECK formulation was not available at the beginning of this thesis, thus the EMBD will be used further in this work as it produces adequate resolution.

Details of CSK, ECK, and other high resolution TFDs such as smoothed WVD, S-Method, Stockwell transform, Hilbert-Huang Transform, and Adaptive Kernel TFD can be found in [27], [61], [63], [80], [114], [115], [116], [117], and [118].

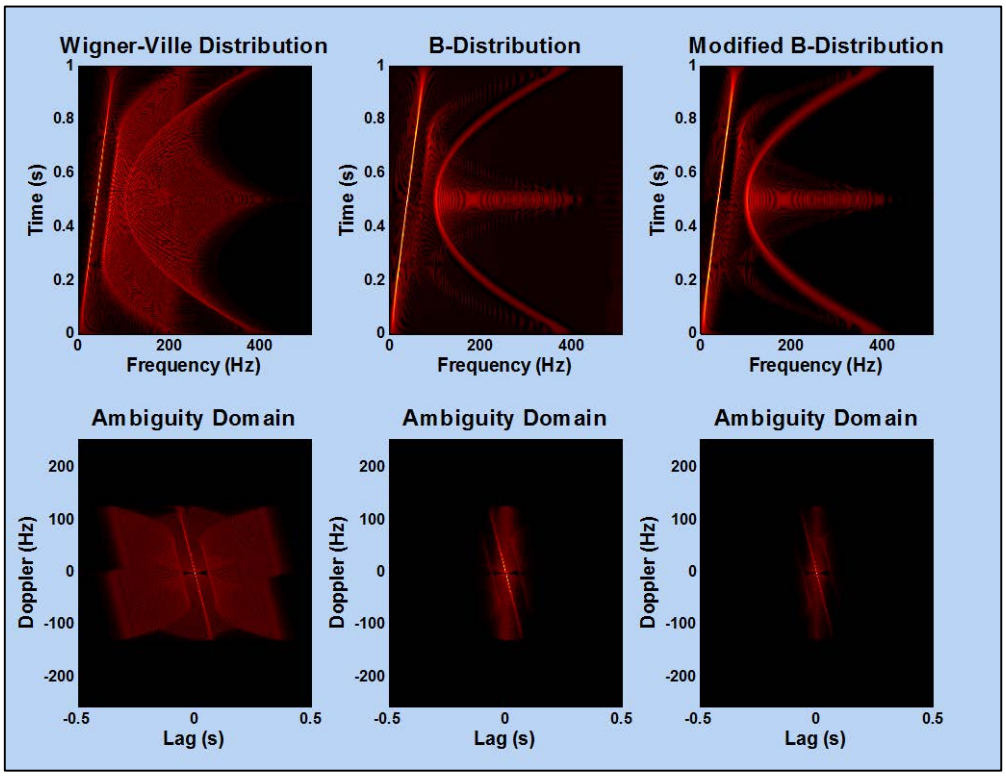

Figure A.15: WVD, BD, and MBD results of a multicomponent signal consisting of a LFM and a QFM. BD parameters are β = 0.08, and MBD parameters are β = 0.08

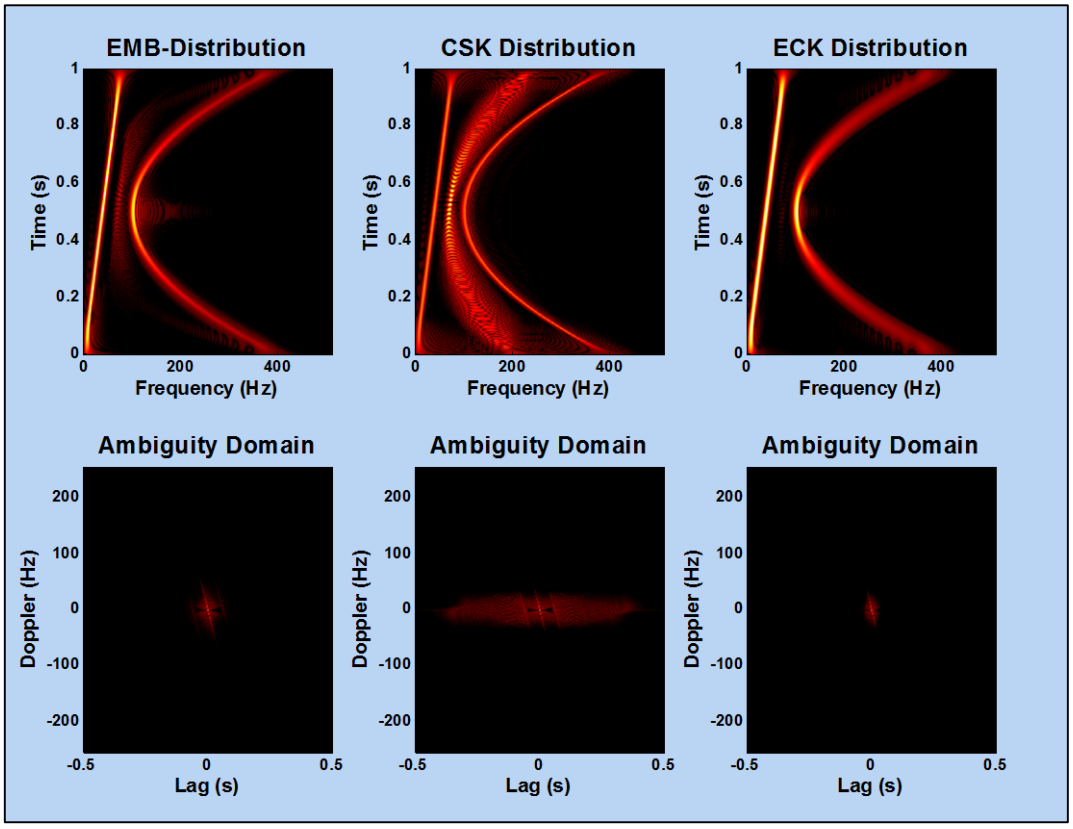

Figure A.16: EMBD, CSK, and ECK results of a multicomponent signal consisting of a LFM and a QFM. EMBD parameters are: $\alpha = 0.08$ and $\beta = 0.3$, CSK parameters are C = 4 and D = 50, and ECK parameters are C = 4, D = 50, and E = 0.05

A.5 Extension of Single-Channel QTFDs to Multichannel QTFDs

EEG is recorded using multiple electrodes (sensors) for measurement reliability and events localisation. The multisensory scheme of EEG recording can take multiple cases such as: Single Input Single Output (*SISO*), Multiple Input Single Output (*MISO*), Single Input Multiple Output (*SIMO*), and Multiple Input Multiple Output (*MIMO*). Figure A.17 illustrates all of these cases where S_i are different source signals. The QTFDs for a single sensor (SISO case) must be extended to consider the multisensory scheme to form multi-sensor QTFDs [1].

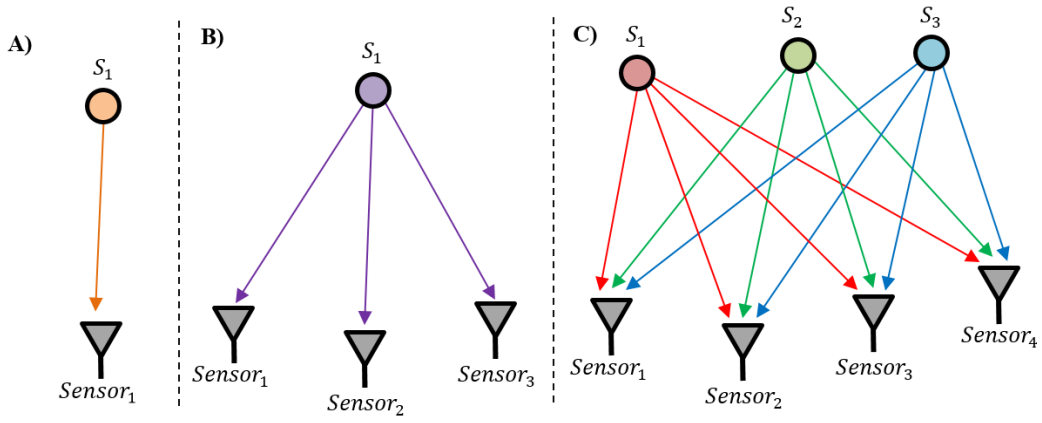

Figure A.17: EEG recording multisensory schemes. A illustrates SISO, B illustrates SIMO, and C illustrates MIMO (Redrawn from [1])

Excluding the SISO case, the sensor received signals and their analytic associates would form vectors as expressed by Equation A.50 and Equation A.51. The class of quadratic Spatial TFDS (*STFDs*) is then defined by Equation A.52, where the time convolution operator is applied to each entry of the spatial instantaneous correlation matrix $\boldsymbol{K}_{zz}(t,\tau)$ as defined by Equation A.53. The spatial instantaneous correlation matrix contains the instantaneous auto-correlation functions (diagonal terms) and cross-correlation functions (off-diagonal terms). Furthermore, in the STFD matrix $\boldsymbol{\rho}_{zz}(t,f)$ the diagonal terms are called auto-TFDs (Equation A.54) and the off-diagonal terms are called cross-TFDs (Equation A.55). Consequently the multichannel or spatial time-frequency analysis contains both auto-TFDs and cross-TFDs. Moreover, there are two types of cross-terms in multichannel TFDs. The first type is cross-terms embedded within the auto-TFDs due its nonlinear nature, and the second type of cross-terms is associated with cross-TFDs when having different source signals as they interact together forming artifacts. Figure A.18 illustrates the STFD 3-dimensional matrix that holds the auto-TFDs of the example signal depicted in Figure A.3 [1].

$$\boldsymbol{s}(t) = [s_1(t),\ s_2(t), \dots, s_M(t)]^T$$ Equation A.50

$$\boldsymbol{z}(t) = [z_1(t),\ z_2(t), \dots, z_M(t)]^T$$ Equation A.51

$$\boldsymbol{\rho}_{zz}(t,f) = \mathcal{F}_{\tau \to f}\{\, G(t,\tau) *_t \boldsymbol{K}_{zz}(t,\tau)\} =$$

$$\begin{bmatrix} \rho_{z_1z_1}(t,f) & \rho_{z_1z_2}(t,f) & \cdots & \rho_{z_1z_M}(t,f) \\ \rho_{z_2z_1}(t,f) & \rho_{z_2z_2}(t,f) & \cdots & \rho_{z_2z_M}(t,f) \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{z_Mz_1}(t,f) & \rho_{z_Mz_2}(t,f) & \cdots & \rho_{z_Mz_M}(t,f) \end{bmatrix}$$ Equation A.52

$$\boldsymbol{K}_{zz}(t,\tau) = \boldsymbol{z}\left(t+\frac{\tau}{2}\right)\boldsymbol{z}^{\boldsymbol{H}}\left(t+\frac{\tau}{2}\right) =$$

$$\begin{bmatrix} K_{z_1z_1}(t,f) & K_{z_1z_2}(t,f) & \cdots & K_{z_1z_M}(t,f) \\ K_{z_2z_1}(t,f) & K_{z_2z_2}(t,f) & \cdots & K_{z_2z_M}(t,f) \\ \vdots & \vdots & \ddots & \vdots \\ K_{z_Mz_1}(t,f) & K_{z_Mz_2}(t,f) & \cdots & K_{z_Mz_M}(t,f) \end{bmatrix}$$ Equation A.53

$$\rho_{z_iz_i}(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(t-u,\tau)\, z_i\left(u+\frac{\tau}{2}\right) z_i^*\left(u-\frac{\tau}{2}\right) e^{-j2\pi\tau f}\, du d\tau$$ Equation A.54

$$\rho_{z_iz_j}(t,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(t-u,\tau)\, z_i\left(u+\frac{\tau}{2}\right) z_j^*\left(u-\frac{\tau}{2}\right) e^{-j2\pi\tau f}\, du d\tau$$ Equation A.55

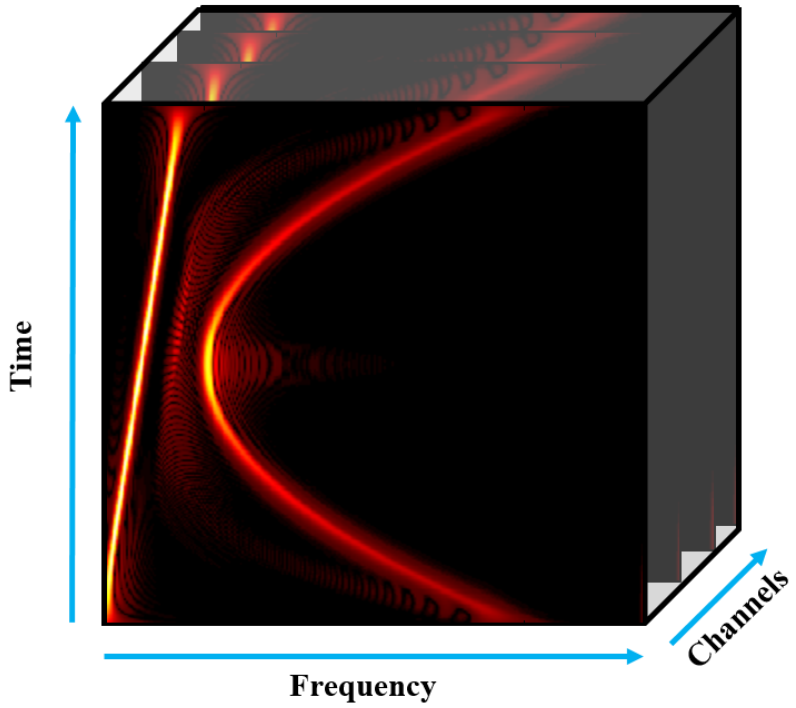

Figure A.18: Auto-TFDs 3-dimensional matrix of the example signal in Figure A.3

Appendix B: Blind Source Separation

The material presented in this appendix are used to explain BSS algorithms in depth along with the extension of time domain BSS to the time-frequency domain. These materials are needed to for Section 2.4.2 in page 31, and Section 4.1 in page 74.

B.1 BSS Algorithms

B.1.1 Fast-ICA

Fast-ICA is described as a computationally efficient method for performing the estimation of ICA [69]. It uses fixed-point iteration scheme that has been proven to be 10-100 times faster than conventional gradient descent methods. Fast-ICA extracts the sources one-by-one and uses the normalized kurtosis as its criterion function that every extracted vector must maximise (Equation A.56). More in depth details on Fast-ICA can be found in [106]. The Matlab code for the Fast-ICA algorithm was taken from [119].

$$K(w) = \frac{E\{|\hat{s}|^4\} - 2E^2\{|\hat{s}|^2\} - |E\{s^2\}|^2}{E^2\{|\hat{s}|^2\}} \qquad \text{Equation A.56}$$

B.1.2 Robust-ICA

Robust-ICA algorithm represents a simple modification on the Fast-ICA criterion function. The search for maxima's is done by using a search direction g. The search direction is the gradient of the criterion function. The Robust-ICA criterion function and its search direction are described using Equation A.57 and Equation A.58. Robust-ICA as its name is more robust than Fast-ICA and has a very high convergence speed [55]. More in depth details on Robust-ICA can be found in [107]. The Matlab code for the Robust-ICA algorithm was taken from [120].

$$\mu_{OPT} = \arg\max_{\mu} K(w + \mu g)$$ Equation A.57

$$g = \nabla_w K(w)$$ Equation A.58

B.1.3 Blind Source Separation-Canonical Correlation Analysis

BSS-CCA solves the BSS problem by forcing the sources to be mutually uncorrelated and maximally correlated with a predefined function. The predefined function z is the delayed version of the signal $y(t)$, or $y(t, f)$ and it is described using Equation A.59.

$$z(t) = y(t - 1)$$ Equation A.59

Canonical correlation analysis obtains two sets of basis vectors, one for $y(t)$ and the other for $z(t)$, such that the correlations between the projections of the variables onto these basis vectors are mutually maximised [55]. Linear combinations of components in y and z are described using Equation A.60 and Equation A.61.

$$y = W_y^T . Y$$ Equation A.60

$$z = W_z^T . Z$$ Equation A.61

BSS-CCA finds the vectors W_y and W_z that maximise the correlation ρ between the y and z by solving the following maximisation problem:

$$max_{W_y, W_z}[\rho(y, z)] = \frac{E\{yz\}}{\sqrt{E\{y^2\}.E\{z^2\}}} = \frac{W_y^T C_{yz} W_z}{\sqrt{(W_y^T C_{yy} W_y).(W_z^T C_{zz} W_z)}}$$ Equation A.62

The solution of the maximisation problem is computed by setting the derivatives of Equation A.62 to zero. Which results in the following two eigenvalue problems.

$$solution: \begin{cases} {C_{yy}}^{-1} C_{yz} {C_{zz}}^{-1} C_{zy} \widehat{W}_y = \rho^2 \widehat{W}_y \\ {C_{zz}}^{-1} C_{zy} {C_{yy}}^{-1} C_{yz} \widehat{W}_z = \rho^2 \widehat{W}_z \end{cases}$$ Equation A.63

After solving Equation A.63, we can calculate the estimates of the sources using Equation A.64. More in depth details on BSS-CCA can be found in [108]. The Matlab code for the BSS-CCA algorithm was taken from [121].

$$\hat{s}_i = \widehat{W}_{y_i}{}^T y \qquad \text{Equation A.64}$$

B.1.4 Second Order Blind Identification

SOBI uses second order statistics to decompose the measurements and it is employed in numerous EEG studies. Furthermore, SOBI exploits non-stationarity, which is typical for EEG data and SOBI is proved to be among the best methods for separating EEG data [69] [122].

It is the most appropriate algorithm for sources that are individually correlated in time, but mutually uncorrelated [54] [55]. It is based on a joint diagonalization of correlation matrices. Mathematically, this means that for all time lags τ the source correlation matrices are diagonal. Equation A.65 describes the correlation of the observation, where R_s represents the correlation matrix of the source signals and M is the mixing matrix. Considering that this equation holds for all values of τ, the mixing matrix M is the one that jointly diagonalises all the correlation matrices. More in depth details on SOBI can be found in [109] and [110]. The Matlab code for the SOBI algorithm was taken from [123].

$$R_y(\tau) = E\{y(y)\, y(y+\tau)^T\} = M\, R_s(\tau)\, M^T \qquad \text{Equation A.65}$$

B.1.5 Joint Approximate Diagonalization of Eigen Matrices

JADE is one of the earliest BSS implementations and has a procedure based on higher-order statistics. JADE proved to be effective for removing artifacts from EEG

data [69] [122]. This approach exploits the fourth order moments in order to separate the source signals from mixed signals.

First, the whitening matrix P and the signal Z are estimated using Equation A.66. Secondly, the cumulants of the whitened mixtures $\hat{Q}_i^Z$ are computed. After that, an estimate of the unitary matrix R is obtained by maximizing the criteria $\lambda_i V_i$ by means of the joint diagonalization. If $\lambda_i V_i$ cannot be exactly jointly diagonalised, the maximisation of the criteria defines a joint approximate diagonalization. An orthogonal contrast is optimised by finding the rotation matrix R such that the cumulant matrices are as diagonal as possible (Equation A.67). Finally, the mixing matrix is estimated using Equation A.68 and the independent components are estimated by using Equation A.69. More in depth details on JADE can be found in [111]. The Matlab code for the JADE algorithm was taken from [124].

$$Z = P\,Y \qquad \text{Equation A.66}$$

$$R = \arg\min_R \sum_i Off\left(R^T \hat{Q}_i^Z R\right) \qquad \text{Equation A.67}$$

$$\widehat{M} = R\,P^{-1} \qquad \text{Equation A.68}$$

$$\hat{s}_i = \widehat{M}^{-1}\,y_i = \widehat{W} y_i \qquad \text{Equation A.69}$$

B.2 BSS Examples

Example signals are used to test the validity of the artifact detection and removal methodology along the time-frequency extension of BSS using two and three dimensional signals.

Firstly, source signals are generated independently from each other and then mixed linearly using a random mixing matrix. Figure A.19 illustrate three source signals that

are generated in the time domain which are: a pulse train, a Linear Frequency Modulated signal (LFM), and an Amplitude Modulated signal (AM). These signals are linear mixed to form 3 channel waveforms. Figure A.20 demonstrate the time-frequency transformation of the time-domain signals. These three dimensional signals are linearly mixed to form 3 channel three dimensional waveforms. Finally, Figure A.21 illustrate three source images which are for; Albert Einstein, Leonhard Euler, and Alan Turing. These images are linearly mixed to form a 3 channels.

Secondly, source signals/images are estimated using different ICA algorithms. Figure A.22, Figure A.23, and Figure A.24 illustrate the source estimations for the time, time-frequency, and image examples. The sign ambiguity in the source estimation can be clearly seen in Figure A.24 as few image colours are inverted.

Lastly, source signal/image number two is considered to be the real information while others are considered to be artifacts. Figure A.25, Figure A.26, and Figure A.27 illustrate the output of the artifact detection and removal methodologies.

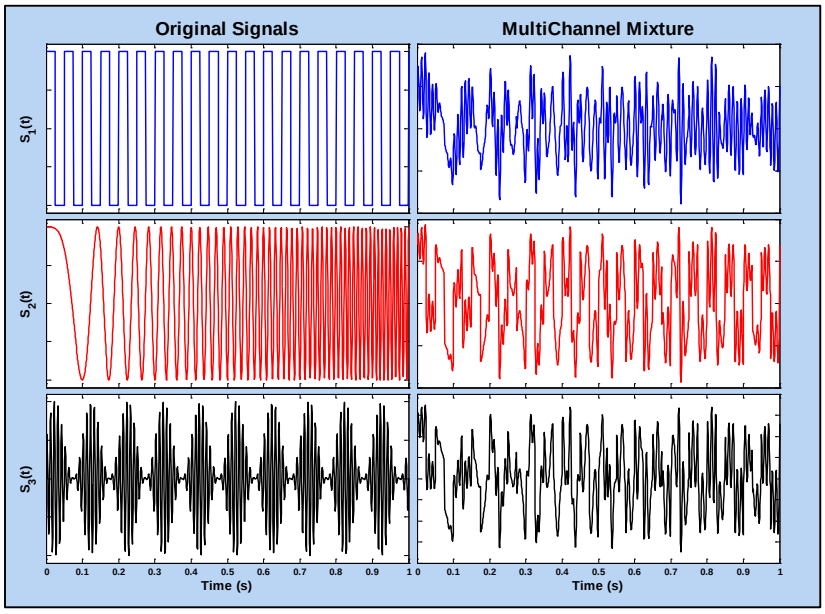

Figure A.19: Time domain example. Source signals are on the left, while multichannel mixtures are on the right

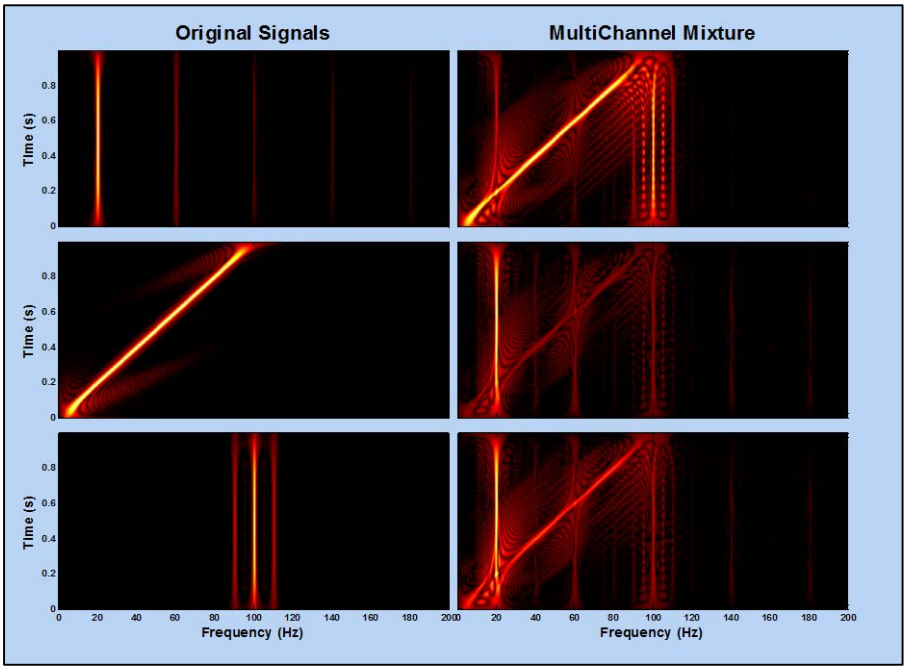

Figure A.20: Time-Frequency domain example. Source signals are on the left, while multichannel mixtures are on the right

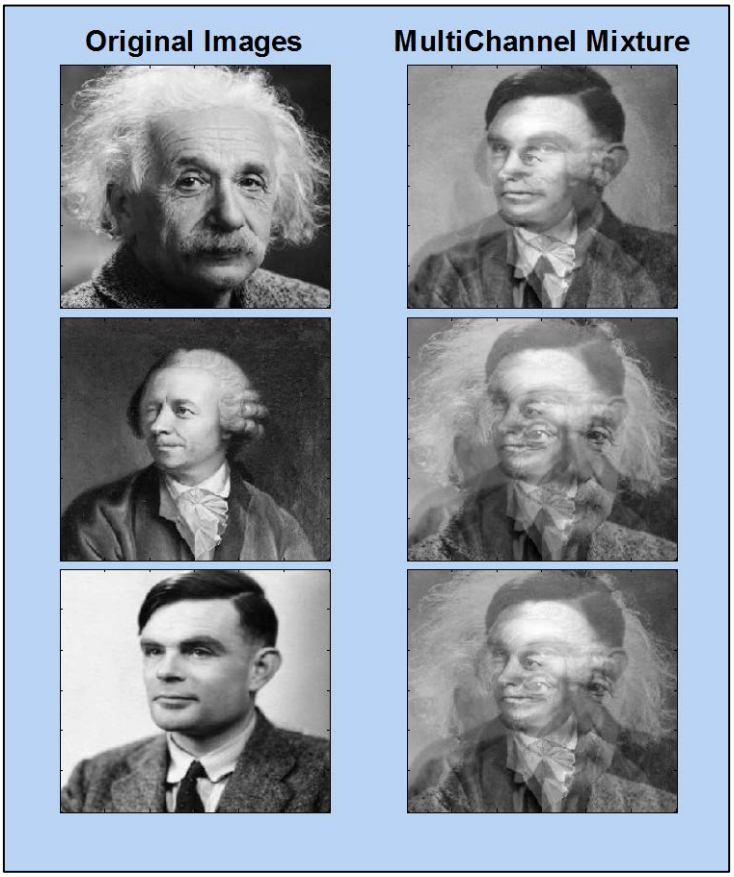

Figure A.21: Image example. Source signals are on the left, while multichannel mixtures are on the right

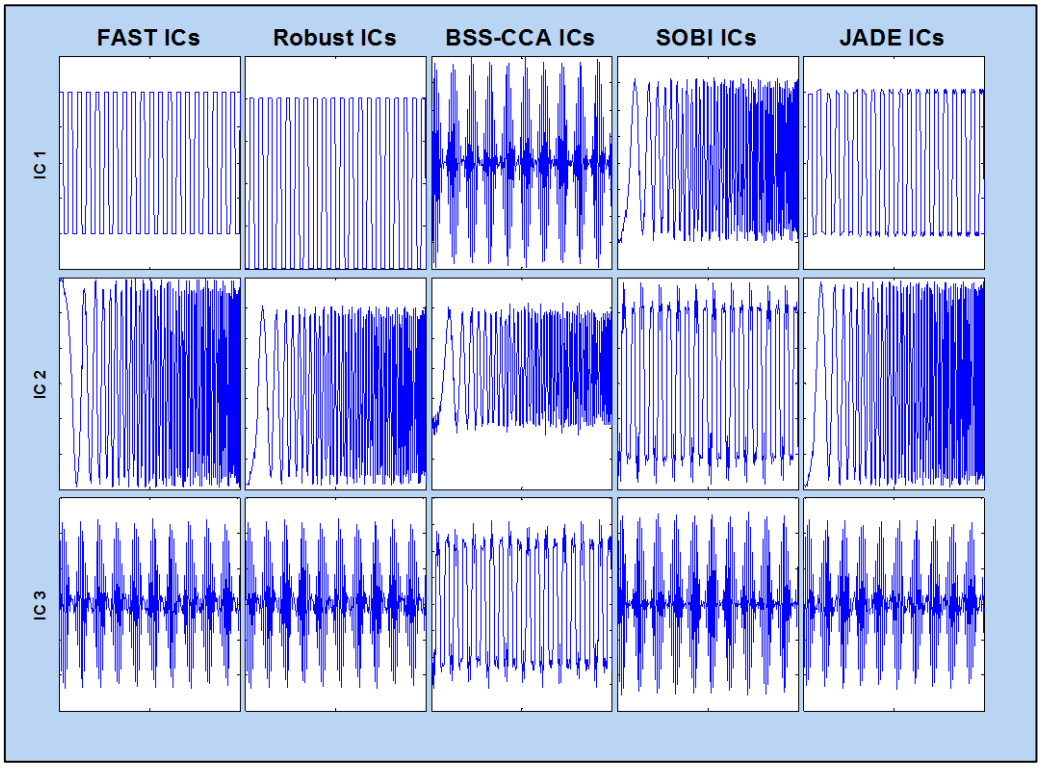

Figure A.22: Time domain example. Two dimensional source signals estimation

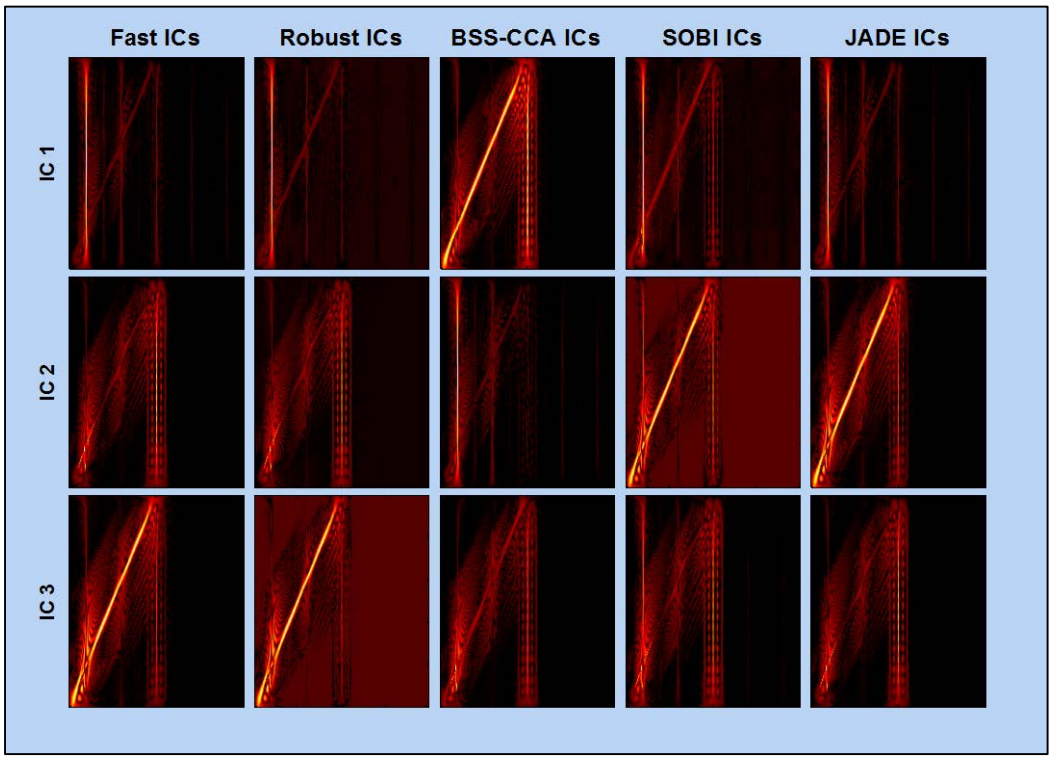

Figure A.23: Time-Frequency domain example. Three dimensional source signals estimation

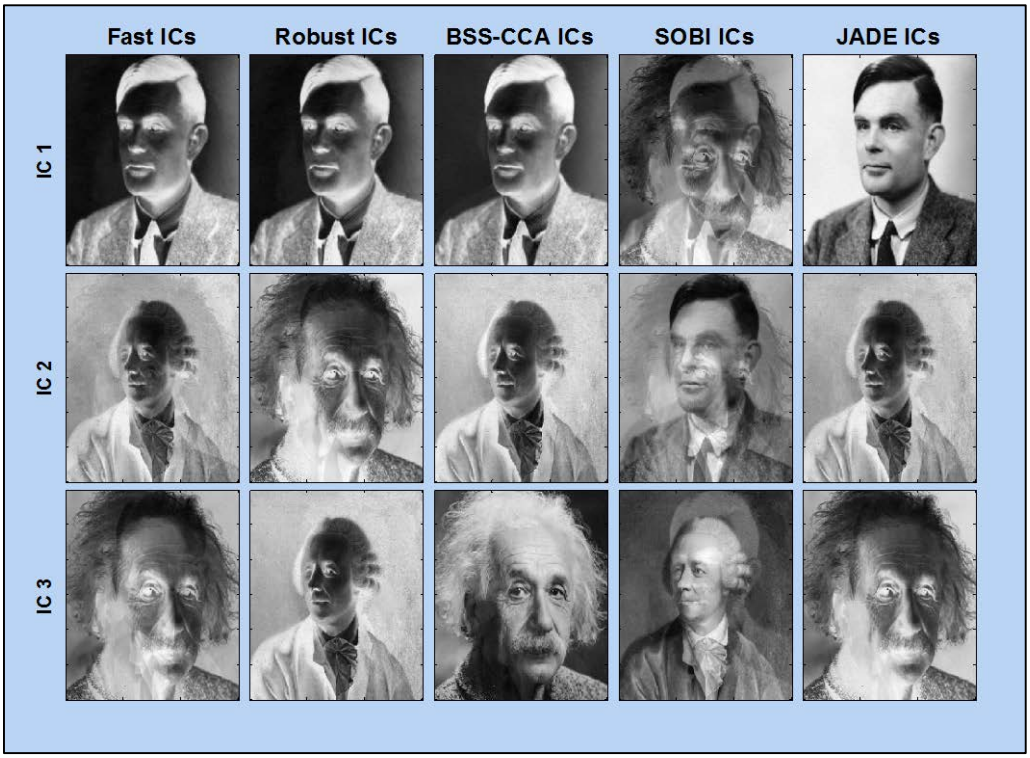

Figure A.24: Image example. Three dimensional source images estimation

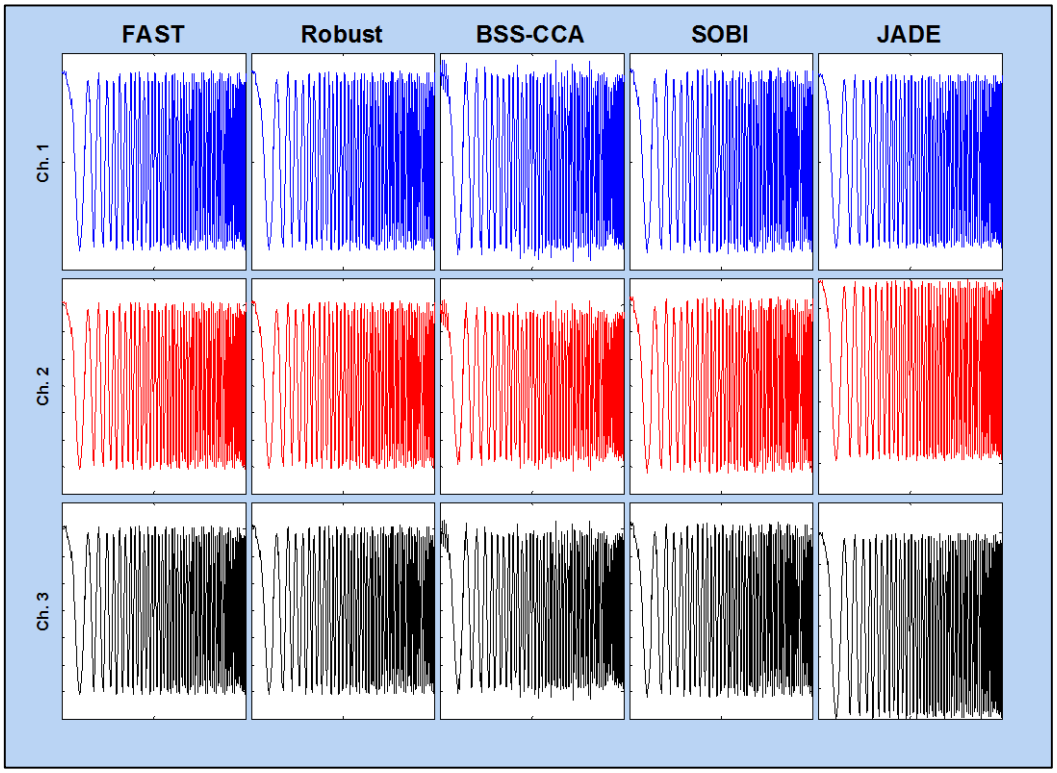

Figure A.25: Time domain example. Artifact detection and removal results

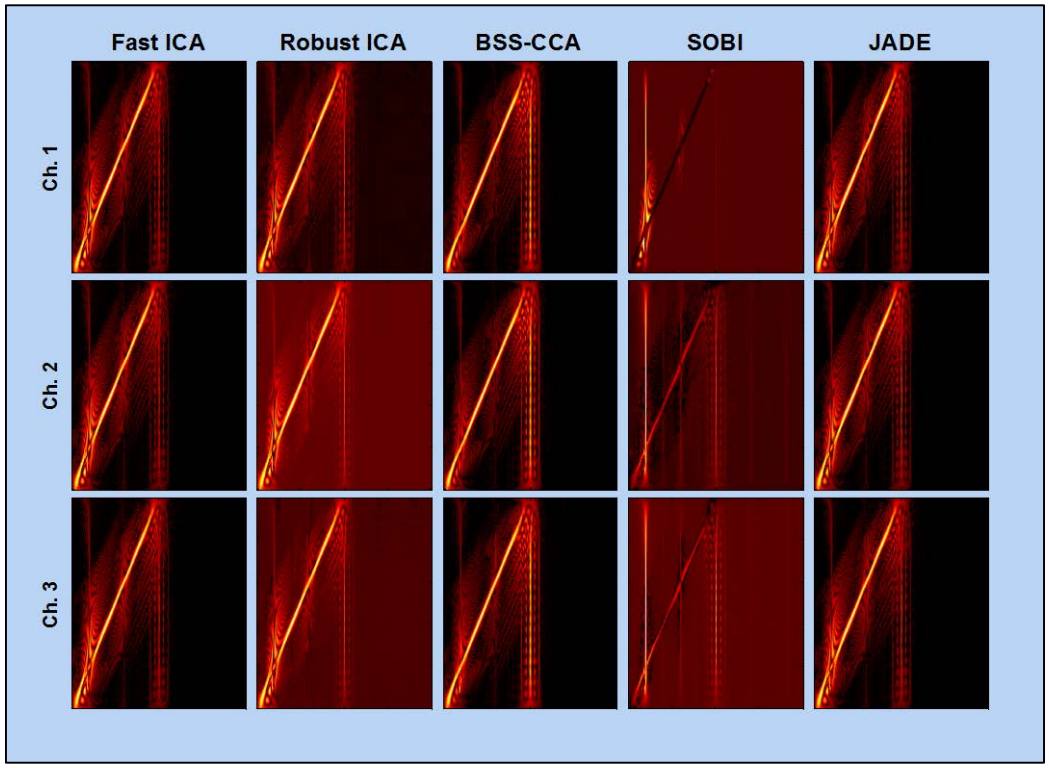

Figure A.26: Time-Frequency domain example. Artifact detection and removal results

Figure A.27: Image example. Artifact detection and removal results

Appendix C: Time-Frequency Features

The material presented in this appendix are used to introduce the time-frequency features that are used in this thesis along with their equations and interpretations. These equations are used in Section 2.5.1 page 36, Section 4.2.1 page 81, and in Section 4.2.3 page 83.

C.1 Statistical Features

Assuming that abnormal and normal EEG have different probability distributions, the following features would result in high discrimination (Equation A.70, Equation A.71, Equation A.72, Equation A.73, and Equation A.74) [27] [30] [36] [63] [78] [80] [81].

- Mean:

$$T_1 = \mu = \frac{1}{NM}\sum_{k=1}^{M}\sum_{n=1}^{N}\rho[n,k] \qquad \text{Equation A.70}$$

- Variance:

$$T_2 = \sigma^2 = \frac{1}{NM}\sum_{k=1}^{M}\sum_{n=1}^{N}(\mu - \rho[n,k])^2 \qquad \text{Equation A.71}$$

- Skewness:

$$T_3 = \frac{1}{(NM-1)\sigma^3}\sum_{k=1}^{M}\sum_{n=1}^{N}(\rho[n,k] - \mu)^3 \qquad \text{Equation A.72}$$

- Kurtosis:

$$T_4 = \frac{1}{(NM-1)\sigma^4}\sum_{k=1}^{M}\sum_{n=1}^{N}(\rho[n,k] - \mu)^4 \qquad \text{Equation A.73}$$

- Coefficient of Variation:

$$T_5 = \frac{\sigma}{\mu} \qquad \text{Equation A.74}$$

C.2 Spectral Extended Features

Assuming that abnormal and normal EEG have different spectral information, the following features would result in high discrimination (Equation A.75, Equation A.76, Equation A.77, Equation A.78, Equation A.79, Equation A.80, Equation A.81, and Equation A.82) [27] [30] [36] [63] [78] [80] [81].

- Spectral Flux: It measures the rate of change of the spectral content of a signal with time (Equation A.75). This measurement has to be done along the time axis ($l = 0, p = 1$), the frequency axis ($l = 1, p = 0$), and along the diagonal axis ($l = 1, p = 1$) to measure any sudden change in all directions (Equation A.76, Equation A.77, and Equation A.78).

$$Flux = \sum_{n=1}^{N-l} \sum_{k=1}^{M-p} \rho[n + l, k + p] - \rho[n, k] \quad \text{Equation A.75}$$

$$F_1 = \sum_{n=1}^{N} \sum_{k=1}^{M-1} \rho[n, k + 1] - \rho[n, k] \quad \text{Equation A.76}$$

$$F_2 = \sum_{n=1}^{N-1} \sum_{k=1}^{M} \rho[n + 1, k] - \rho[n, k] \quad \text{Equation A.77}$$

$$F_3 = \sum_{n=1}^{N-1} \sum_{k=1}^{M-1} \rho[n + 1, k + 1] - \rho[n, k] \quad \text{Equation A.78}$$

- Energy Concentration: It determines the sparsity of the signal energy in the time-frequency domain.

$$F_4 = \left(\sum_{n=1}^{N} \sum_{k=1}^{M} \sqrt{|\rho[n, k]|} \right)^2 \quad \text{Equation A.79}$$

- Spectral Flatness: It measures the level of uniformity of the energy distribution in the time-frequency domain. It is computed as the ratio of the geometric mean of a TFD divided by its arithmetic mean.

$$F_5 = MN \frac{\prod_{n=1}^{N} \prod_{k=1}^{M} \rho[n,k]}{\sum_{k=1}^{M} \sum_{n=1}^{N} \rho[n,k]} \quad \text{Equation A.80}$$

- Normalised Renyi Entropy: It measures the randomness in the distribution of signal energy in the time-frequency domain.

$$F_6 = -\frac{1}{2} \log_2 \left(\sum_{n=1}^{N} \sum_{k=1}^{M} \left(\frac{\rho[n,k]}{\sum_{k=1}^{M} \sum_{n=1}^{N} \rho[n,k]} \right)^3 \right) \quad \text{Equation A.81}$$

- Shannon Entropy: It quantifies the amount of uncertainty in the distribution of the signal energy in the time-frequency domain.

$$F_7 = -\sum_{k=1}^{M} \sum_{n=1}^{N} \left(\frac{\rho[n,k]}{\sum_{k=1}^{M} \sum_{n=1}^{N} \rho[n,k]} \log_2 \left(\frac{\rho[n,k]}{\sum_{k=1}^{M} \sum_{n=1}^{N} \rho[n,k]} \right) \right) \quad \text{Equation A.82}$$

C.3 Instantaneous Frequency Features

Assuming that abnormal and normal EEG have different IF law (Equation A.83), the IF statistics would result in high discrimination (Equation A.84 and Equation A.85) [27] [30] [36] [63] [78] [80] [81].

$$f[n] = \frac{F_s}{2M} \frac{\sum_{k=1}^{M} k\rho[n,k]}{\sum_{k=1}^{M} \rho[n,k]} \quad \text{Equation A.83}$$

$$IF_1 = \frac{1}{N} \sum_{n=1}^{N} f[n] \quad \text{Equation A.84}$$

$$IF_2 = \Delta f[n] = sup(f[n]) - Inf(f[n])$$ Equation A.85

C.4 Sub-bands Energy Features

Assuming that abnormal and normal EEG exist on different frequency bands, the following features would result in high discrimination (Equation A.86 and Equation A.87) [27] [30] [36] [63] [78] [80] [81].

$$E_1 = \sum_{n=1}^{N} \sum_{k=1}^{M_\delta} \rho[n,k]$$ Equation A.86

$$E_2 = \sum_{n=1}^{N} \sum_{k=M_\delta}^{2M_\delta} \rho[n,k]$$ Equation A.87

$$M_\delta = 10\frac{M}{F_s}$$ Equation A.88

Appendix D: Performance Metrics

The material presented in this appendix are used to explain how to quantify the performance of any detection or removal (filtering) algorithm. This material is used in Section 4.1 page 74 to quantify artifact detection and removal performance, and in Section 2.5.4 page 40 to quantify the abnormality detection performance.

D.1 Detection Performance

Accuracy, sensitivity, specificity, and balanced accuracy are performance metrics that are generally used to quantify any detection algorithm or detection system performance. Accuracy measures the level of measurement that yields true and consistent results, while sensitivity (also called true positive rate) measures the proportion of positives that are correctly identified (target signals exist and detected), and lastly specificity (also called true negative rate) measures the proportion of negatives that are correctly identified (target signals do not exist and not detected). Balanced accuracy is a trade-off between sensitivity and specificity as it is their average, while another term can be also used namely “Informedness” where it re-normalises the balanced accuracy to have a range from zero to one [125]. Accuracy, sensitivity, specificity, balanced accuracy, and informedness are described in Equation A.89, Equation A.90, Equation A.91, Equation A.92, and Equation A.93, where TP is true positive, TN is true negative, FP is false positive, and FN is false negative. The explanations of these terminologies are as follows: True positive: target signal exist and correctly identified, False positive: target signal do not exist but identified, True negative: target signal do not exist and not identified, False negative: target signal exist but not identified.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad \text{Equation A.89}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad \text{Equation A.90}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad \text{Equation A.91}$$

$$\text{Balanced Accuracy} = \frac{\text{Sensitivity} + \text{Specificity}}{2} \quad \text{Equation A.92}$$

$$\text{Informedness} = 2 \times \text{Balanced Accuracy} - 1 \quad \text{Equation A.93}$$

D.2 Error Evaluation

The Normalised Root Mean Square Error (*NRMSE*) quantifies the relative differences or errors between an estimate and a real value [10]. The NRMSE can be calculated for a one-dimensional segment i of a signal or static using Equation A.94. $\hat{x}$ is the estimated signal, x is the real or reference signal, N is the total number of samples within a segment, and N_{seg} is the total number of segments. This definition can be extended to include N-dimensional signals and statics by using tensor notations. Equation A.95 describes the NRMSE computation for two-dimensional segments of size $N \times M$.

$$\text{NRMSE}_i^{\mathbb{R}} = \sqrt{\frac{\sum_{n=1}^{N} |x(n) - \hat{x}(n)|^2}{\sum_{n=1}^{N} |x(n)|^2}}, \qquad i \in \left[1, N_{seg}\right] \quad \text{Equation A.94}$$

$$\text{NRMSE}_i^{\mathbb{R}^2} = \sqrt{\frac{\sum_{m=1}^{M} \sum_{n=1}^{N} |x(n,m) - \hat{x}(n,m)|^2}{\sum_{m=1}^{M} \sum_{n=1}^{N} |x(n,m)|^2}}, \qquad i \in [1, N_{seg}] \quad \text{Equation A.95}$$

Appendix E: The Sadleir Four Sphere Head Model

The Sadleir four sphere head model is used in this thesis to simulate the neonatal head, which is utilised in Section 3.2.1 page 45. It divides the neonatal head into four concentric spheres (structures) namely: scalp, skull, cerebrospinal fluid (*CSF*), and the brain [21]. The dimensions of the spheres are summarized in Table A.1. A basic head model is constructed to reference the viewing perspective of front, back, and side views of the neonatal head. This is done by creating; a cone representing the nose to illustrate the front and back views, and two small ellipsoids representing the ears to illustrate the side views (Figure A.28 and Figure A.29).

Table A.1: The neonatal head four sphere model dimensions. Taken from [21]

	Radius (cm)	Thickness (cm)	Volume (cm^3)
Head	**5.95**	**---**	**882.3472**
Scalp	**---**	**0.29**	**122.8295**
Skull	**---**	**0.6**	**216.8423**
CSF	**---**	**0.3**	**90.9137**
Brain	**4.76**	**---**	**451.7618**

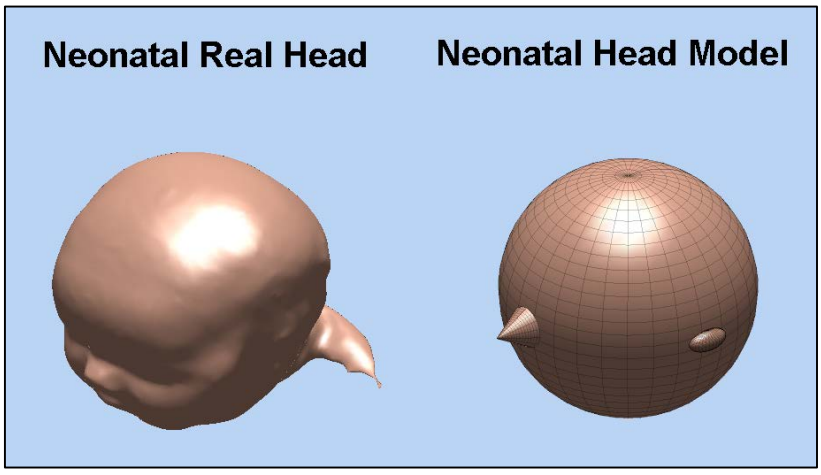

Figure A.28: Basic neonatal head model. The head structure size is 32x32

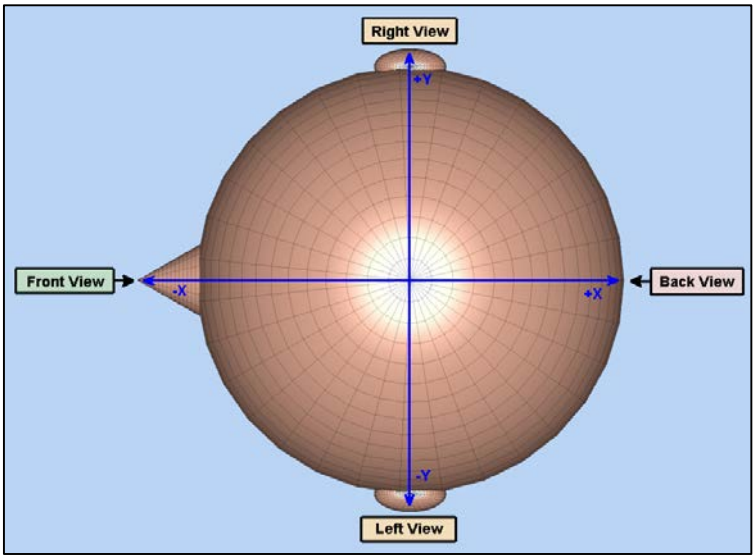

Figure A.29: Basic neonatal head model to reference viewing perspective. The head structure size is 32x32

The neonatal four sphere head model is illustrated in 3D and in 2D top view cross sections (Figure A.30). It shows the different head regions along with their dimensions in different colours; the scalp region is illustrated in beige, the skull region is in green, the CSF region is in light blue, and finally the brain region is in red.

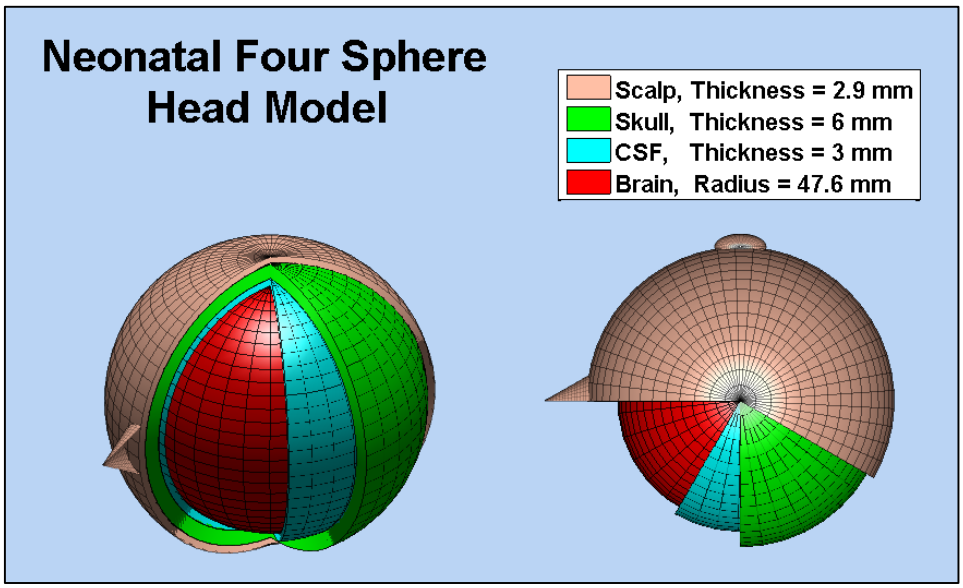

Figure A.30: Neonatal four sphere head model cross sections in 3D and in 2D top view, every sphere is constructed using a 32x32 structure

Appendix F: The Rankine EEG Model

The material presented in this appendix are used to explain the single-channel EEG model used in this thesis in depth. This model is utilised in Section 3.2.2, page 46.

F.1 Neonatal EEG Background Simulator

The neonatal background EEG simulator is based on the nonlinear analysis of real neonatal EEG signals using Fractal Dimension (*FD*) estimation. The synthetic neonatal EEG signal is derived from the relationship between FD and the spectrum power law index to produce a synthetic epoch [41] [43].

The power spectrum of the neonatal background EEG approximately follows a power law of the form expressed in Equation A.96, where c is a constant, f is frequency and γ is the power law exponent [41].

$$S(f) \approx \frac{c}{|f|^{\gamma}} \qquad \text{Equation A.96}$$

The nonstationary behaviour of the neonatal background EEG is modelled using a time-varying power law exponent γ_n, therefore the modified model can be described using Equation A.97, where $S_n(f)$ is the power spectrum associated with nth epoch of duration T [41] [43].

$$S_n(f) = \frac{c}{|f|^{\gamma_n}} \qquad \text{Equation A.97}$$

This model assumes that the signal is quasi-stationary in every duration, which means that γ_n will be constant for the duration of an epoch, but will vary on an epoch by epoch basis. The power spectrum $S_n(f)$ can be expressed using Equation A.98, where $X_n(f)$ is the FT of the nth epoch $x_n(t)$ (Equation A.99), and $\theta_n(f)$ is the phase spectrum which is assumed to be a realisation of a random process [41] [43].

$$S_n(f) = \frac{c}{|f|^{\gamma_n}} = X_n(f)X_n^*(f)$$ Equation A.98

$$X_n(f) = \frac{\sqrt{c}}{|f|^{\gamma_n/2}} e^{j\theta_n(f)}$$ Equation A.99

Synthesis of $x_n(t)$ is then done by taking the IFT of Equation A.99. The modelled epoch $x_n(t)$ has a power spectrum with a smooth power law. However the power spectra of the real neonatal EEG background exhibits random fluctuations around the power law. Therefore, 15 sub-epochs with the same power law exponent, but different random phase spectra $\theta_n(f)$ are created then added together. The constructive and destructive interference of the sub-epochs result in fluctuations around the desired power law; mimicking the power spectrum of real neonatal EEG background [41].

The power law exponent is estimated using its linear relationship with the Fractal Dimension FD (Equation A.100). Studies have shown that the analysis of the estimated γ_n values shows random fluctuations following a Beta distribution with parameters α and β equal to 7.82 and 7.44 respectively [41]. The phase spectrum $\theta_n(f)$ is estimated to be a random process with uniform distribution having a range of $[0,\ 2\pi]$ [41] [43].

$$FD = \frac{5 - \gamma_n}{2}$$ Equation A.100

F.2 Neonatal EEG Seizure Simulator

The neonatal EEG seizure simulator is based on the well-known time-frequency nonstationary signal model (Equation A.101). This model accounts for all the significant time-frequency characteristics of neonatal EEG seizure including: multiple components or harmonics, piece-wise linear instantaneous frequency laws and harmonic's amplitude modulation. A time-frequency model for real nonstationary

signals with multiple components can be expressed using Equation A.101, where $a_k(t)$, $f_k(\tau)$, and θ_k are the amplitude modulation, time-varying IF function, and initial phase for the kth signal component, respectively [41] [42].

$$s(t) = \sum_{k=1}^{K} a_k(t) cos\left(2\pi \int_0^t f_k(\tau) d\tau + \theta_k\right) \quad \text{Equation A.101}$$

It can be seen from Equation A.101 that estimates for the functions $a_k(t)$, $f_k(\tau)$, θ_k and the number of harmonics K are required for the simulation of neonatal EEG seizure.

The IF function $f_k(t)$ is modelled as a piecewise linear function, where the general form of a piecewise LFM function $f(t)$, with M pieces is given by Equation A.102. The start frequency of the LFM is given by f_{st}, $\xi = [\xi_1, \xi_2, \dots, \xi_M]$ are the gradients in Hz/sec, $B = [B_1 = 0, B_2, B_3, \dots, B_M, B_{M+1} = N]$ are the turning points in seconds, N is the discreet length of the seizure and C_m is the alignment intercept that ensures continuity [41] [42].

$$f(t) = \sum_{m=1}^{M} F_m(\xi_m, C_m; t)\, rect\left(\frac{t - 0.5(B_{m+1} - B_m)}{B_{m+1} - B_m}\right) \quad \text{Equation A.102}$$

$$F_m(\xi_m, C_m, t) = \xi_m t + C_m \quad \text{Equation A.103}$$

$$C_m = \begin{cases} f_{st} & ; m = 1 \\ F_{m-1}(\xi_{m-1}, C_{m-1}, B_m) - \xi_m B_m & ; m \geq 2 \end{cases} \quad \text{Equation A.104}$$

The multiple harmonics of the neonatal EEG seizure are related to the fundamental frequency, therefore the IF for each harmonic can be derived from the fundamental using Equation A.105 [41] [42].

$$f_k(t) = k\, f_1(t) \quad \text{Equation A.105}$$

The amplitude modulation function $a_k(t)$ of each harmonic is parameterized by a gain factor R_k, normalized variation V_n, and number of turning points P (Equation A.106).

$$a_k(t) = \Phi(R_k, V_k, P, t)$$ Equation A.106

The gain factor R_k is referred to as the harmonic ratio, which gives an indication that $R_1 = 1$. The component amplitude modulation function $a_k(q)$ is determined from a cubic spline interpolation of P randomly assigned turning points with amplitudes using Equation A.107, where the mean of $V_n = 0.33$ and the locations of the turning points are found using Equation A.108 [41] [42].

$$a_k(q) = R_k(0.67 + V_n)$$ Equation A.107

$$q = \frac{N(p + X)}{P}$$ Equation A.108

$p = [0, 1, \dots, P - 1]$ and X is a stationary random process uniformly distributed between 0 and 1. The boundary conditions of the cubic spline fit are set to have a derivative of zero. Finally, the initial phase value θ_k is assumed to be a random variable with a stationary uniform distribution on $[-\pi, \pi)$.

The complexity of this model is reduced by setting the number of harmonics K to 5, setting the number of pieces M in $f_k(\tau)$ to 3, assuming that B is a stationary uniformly distributed random process ranging across the epoch, and limiting the number of turning points P in $a_k(t)$ to a maximum of 8. Distribution estimates for all neonatal EEG seizure model parameters are given in Table A.2 [41] [42].

Table A.2: Distribution estimates for the neonatal EEG seizure model parameters. B is a continuous beta distribution, B* is a discreet beta distribution and L-N is a Log Normal Distribution. Taken from [41]

	P	R_2	R_3	R_4	R_5	V_n	ξ	f_{st}
Distribution	**B***	**B**	**B**	**B**	**B**	**B**	**B**	**L-N**
α	1.8	1.7	1.5	1.9	1.4	3.9	69.1	-0.17
β	3.0	3.2	4.1	3.6	1.2	8.0	69.8	0.55
Minimum	1	0.2	0.2	0.2	0.2	0	-0.06	0.425
Maximum	8	1.2	1.0	0.6	0.4	1	0.06	∞

Appendix G: Solving the Radiation Transport Equation

The material presented in this appendix are used to explain in details how to solve the Radiation Transport Equation (*RTE*). The solution of the RTE is utilised in Section 3.2.3.1, page 51.

Two concentric biological disks are illustrated in Figure A.31 to describe the RTE solution for the 2-dimensional case. Equation A.109 describes the signal intensity decay, where I_i is the signal intensity estimated at electrode i. $R_{j,j+1}$ is the coefficient of Fresnel reflection at the normal beam incidence between the disk regions j and $j+1$ (Equation A.110). n_j is the refractive index of disk j. I_o is the initial intensity of the signal. $\mu_{t,j}$, $\mu_{a,j}$, and $\mu_{s,j}$ are the extinction, absorption, and scattering coefficients respectively for the region j (Equation A.111). g_j is the anisotropy factor of disk j. $D_{i,j}$ is the signal path length in region j to reach electrode i. r_i and θ_i defines electrode i position in polar coordinates with respect to the seizure event location (O in Figure A.31), and lastly M is the number of disk regions and N is the number of electrodes. Equation A.112 illustrates the relationship between the electrode relative amplitude A_i and its signal intensity. Since relative amplitudes need to be computed the relationship becomes equal, as the multiplication factors add no relative information.

$$I_i = \left[1 - R_{j,j+1}\right] I_o \exp\left[-\sum_{j=1}^{M} \mu_{t,j}\, D_{i,j}(r_i, \theta_i)\right] \quad , \quad i \in \mathbb{N} | 1 \leq i \leq N \qquad \text{Equation A.109}$$

$$R_{j,j+1} = \left[\frac{(n_j - n_{j+1})}{(n_j + n_{j+1})}\right]^2 \quad , \quad j \in \mathbb{N} | 1 \leq j \leq M \qquad \text{Equation A.110}$$

$$\mu_{t,j} = \mu_{a,j} + \mu_{s,j}(1 - g_j) \qquad \text{Equation A.111}$$

$$A_i \propto \sqrt{I_i} \gg A_i = \sqrt{I_i}$$ Equation A.112

Solving the intensity decay problem at electrode number 1, make Equation A.109, Equation A.110, and Equation A.111 in the form of;

$$I_1 = \left[1 - R_{1,2}\right] I_o \exp\left[-\mu_{t,1} D_{1,1}(r_1, \theta_1) - \mu_{t,2} D_{1,2}(r_1, \theta_1)\right]$$

$$R_{1,2} = \left[\frac{(n_1 - n_2)}{(n_1 + n_2)}\right]^2 \quad , \quad \mu_{t,1} = \mu_{a,1} + \mu_{s,1}(1 - g_1) \quad , \quad \mu_{t,2} = \mu_{a,2} + \mu_{s,2}(1 - g_2)$$

All of these variables are known, except for $D_{1,1}$ and $D_{1,2}$, which are the propagation path lengths through regions 1 and 2. Finding $D_{1,1}$ and $D_{1,2}$ relies on knowing the exact location of the electrode (E_1) and the seizure event location (O in Figure A.31). The solution of this problem is explained in Figure A.31 and consists of the following steps:

a) Create a vector starting from the seizure event location (O) and ending at the electrode location (E_1). Call it $\overrightarrow{OE_1}$, and then calculate its magnitude and angle (r_1, and θ_1).

b) Find the two intersection points between the border of region 1 (which is a circle) and the infinite line L, where $\overrightarrow{OE_1} \subseteq L$.

c) Consider the intersection point (G_1) that creates a vector $\overrightarrow{OG_1}$ parallel to $\overrightarrow{OE_1}$, and then calculate its magnitude and angle. This point is called a ghost electrode where it mimics an electrode attached on region 1 and simplifies the problem to become a mono-region problem, where $D_{1,1}$ is equal to $\left\|\overrightarrow{OG_1}\right\|$.

d) Find the two intersection points between the outer border of region 2 (which is a circle) and the infinite line L.

e) Consider the intersection point (E_1) that creates a vector parallel to $\overrightarrow{OE_1}$, and then calculate its magnitude and angle. $D_{1,2}$ is equal to $\left\|\overrightarrow{OE_1}\right\| - \left\|\overrightarrow{OG_1}\right\|$.

The same steps can be followed to solve the signal path lengths in M regions for N electrodes. This is done by repeating the line-and-circle intersection method NxM times. After finding the electrode signal intensity, the relative amplitude can be calculated using Equation A.112.

The 2-dimensional RTE solution can be extended to the 3-dimensional case by replacing the line-and-circle intersection method with a line-and-sphere intersection method, which considers concentric spheres instead of disks.

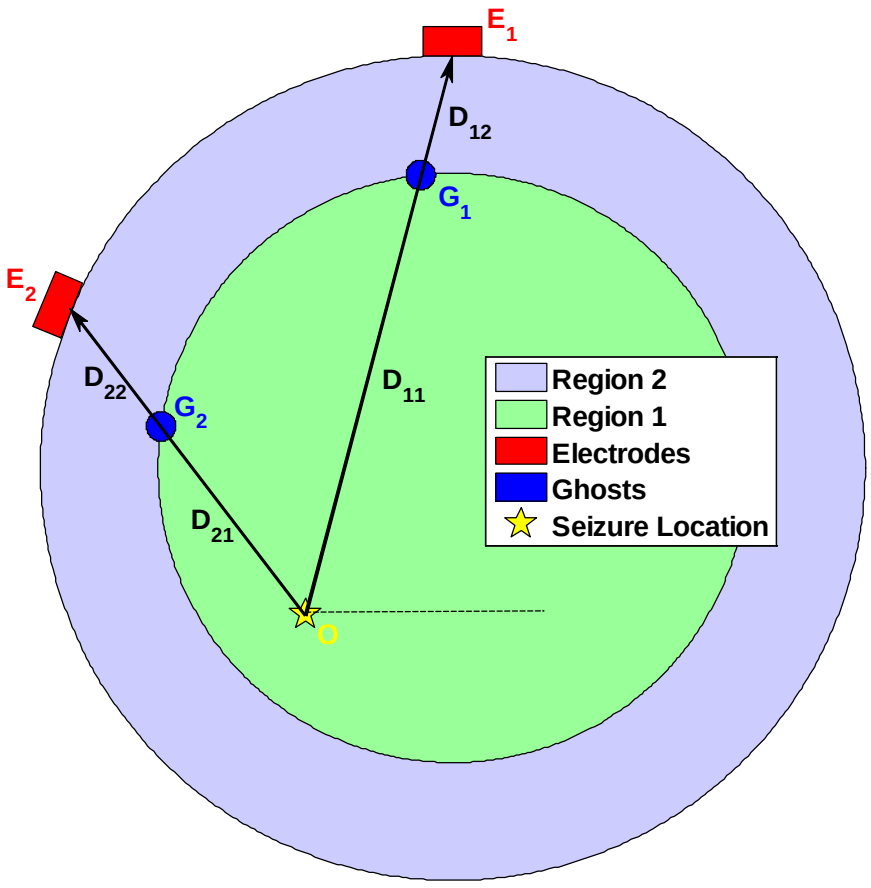

Figure A.31: Solving the radiation transport equation for two concentric disks. O is the seizure event location. E_1 and E_2 are the electrodes locations, while G_1 and G_2 are the ghost electrodes locations. D_{11} is the signal path in region 1 reaching electrode 1, D_{12} is the signal path in region 2 reaching electrode 1, D_{21} is the signal path in region 1 reaching electrode 2, and D_{22} is the signal path in region 2 reaching electrode 2

Appendix H: EEG Corruption Evaluation

The material presented in this appendix are used to evaluate EEG information masking due artifacts. This study is considered as outside the scope of this thesis; hence it is discussed here. This material was referred to in Section 3.3, page 65.

The neonatal EEG corruption evaluation process is intended to quantify the relationship between corrupted background and corrupted seizure EEG epochs in time, frequency and time-frequency domains. This relationship shall illustrate relative common behaviour between the different corrupted EEG epochs. This is caused by the dominance of artifact signals leading to unreliable interpretation/detection of background/seizure segments, consequently mimicking real neonatal EEG epoch corruption situations. The common behaviour is relative as it represents the increase in behaviour alikeness between corrupted background and seizure epochs when compared to what their relationship used to be before the artifact corruption. The quantification of the masking effect on EEG information by artifacts is done using different techniques such as: Pearson's Correlation Coefficient (*PCC*), Two Samples t-Test, and one-way Analysis of Variance (*ANOVA*).

H.1 Testing Signals Generation

The first step in evaluation is generating the signals under test. These signals must have high number of samples covering most of the possible outcomes to produce statistically valuable/reproducible results. The testing signals generation procedure consists of the following steps:

1. Generate 50 minutes (21 channels, 200 segments each channel, 15 seconds each segment) of clean multichannel EEG background and seizure signals.

2. Another 50 minutes (21 channels, 200 segments each channel, 15 seconds each segment) of multichannel EEG artifact is generated using the Neonatal Multichannel EEG Artifacts Model.
3. Corrupted multichannel background EEG is created by adding up the generated multichannel artifacts with the clean background EEG from step 1.
4. Corrupted multichannel seizure EEG is created by adding up the generated multichannel artifacts with the clean seizure EEG from step 1.
5. The first channel of each of the following signals (the clean multichannel background EEG, the clean multichannel seizure EEG, the corrupted multichannel background EEG, and the corrupted multichannel seizure EEG) will be extracted and used in the evaluation processes (200 segments per signal).
6. DFT is computed for each of the extracted signals in step 5.
7. Time-Frequency transformation is computed using the EMBD for each of the extracted signals in step 5.

The first channel was only extracted in step 5 instead of vectorising the multichannel matrix into one vector, because the first channel signals along with their frequency and time-frequency transformations take lots of memory and computation time (2.8 GB to be exact), so it is statistically enough to consider the first channel as the extracted time domain samples are more than 90,000. The extracted single channel signals form four groups of signals under test. The first group contains 200 only clean background epochs, while the second group contains 200 only clean seizure epochs. The third group contains 200 only corrupted background epochs, and the fourth group contains 200 only corrupted seizure epochs.

Evaluation of EEG corruption is done by comparing and testing the alikeness for different combinations of the testing groups. The testing group's combinations fall also under a likelihood umbrella, namely the self-likelihood, and the mutual-likelihood umbrellas.

The self-likelihood umbrella contains the following four testing group's combinations: clean background versus clean background, clean background versus corrupted background, clean seizure versus clean seizure, and clean seizure versus corrupted seizure. The self-likelihood umbrella quantifies the behaviour alikeness between clean signals with themselves, and between clean signals with their corrupted versions to represent the common relative behaviour or the change in the common behaviour. This umbrella should produce negligible results as the nature of background and seizure epochs are totally different from their corrupted versions.

The mutual-likelihood umbrella contains the following three major testing group's combinations: clean background versus clean seizure, clean background versus corrupted seizure, and corrupted background versus corrupted seizure. The mutual-likelihood umbrella quantifies the behaviour alikeness between; clean signals, between corrupted signals, and between corrupted background and clean seizure to represent the likelihood for mistakes in the analysis or interpretation of the EEG patterns. This umbrella should produce significant results, because artifacts dominate the overall behaviour of the corrupted EEG signal.

H.2 Evaluation Techniques

The second step in evaluation is processing the testing signals by different techniques. PCC, Two Samples t-Test, and one-way ANOVA techniques are used to

process the signal under test for every likelihood umbrella in the time, frequency, and time-frequency domains.

H.2.1 Pearson's Correlation Coefficient

Correlation is "a method of assessing a possible two-way linear association between two continuous variables/signals" [126]. It is measured by a statistic called the correlation coefficient, which measures the strength of the linear association between variables under test. It is a dimensionless quantity that can vary from -1 (perfect negative correlation) through 0 (no correlation) to +1 (perfect positive correlation). Zero correlation indicates that there is not any linear relationship between two testing signals, and a correlation of -1 or +1 indicates a perfect linear relationship. The stronger the linear correlation, the closer the correlation coefficient comes to ±1. If the coefficient is a positive number, the variables are directly related, and if it is a negative number, the variables are inversely related. Table A.3 illustrates the linear correlation coefficient ranges with their scientific interpretations [126] [127].

Table A.3: Rule of thumb for interpreting the size of a correlation coefficient

Correlation Value	Interpretation
0.90 to 1.00 (-0.90 to -1.00)	Very high positive (negative) correlation
0.70 to 0.90 (-0.70 to -0.90)	High positive (negative) correlation
0.50 to 0.70 (-0.50 to -0.70)	Moderate positive (negative) correlation
0.30 to 0.50 (-0.30 to -0.50)	Low positive (negative) correlation
0.00 to 0.30 (0.00 to -0.30)	Negligible correlation

There are two main types of correlation coefficients namely: Pearson's product moment Correlation Coefficient (*PCC*) and Spearman's rank Correlation Coefficient (*SCC*). The PCC is denoted as ρ for a population and as r for a sample statistic. It is used when

both variables are normally distributed or generally coming from the same distribution. The PCC between variables x and y can be calculated using Equation A.113. The SCC is denoted as ρ_s for a population and as r_s for a sample statistic. It is appropriate when one or both variables are skewed or ordinal and is robust when extreme values are present [126]. Equation A.114 describes the SCC between variables x and y, where d_i is the difference in ranks for x and y.

$$r = \frac{cov(X,Y)}{\sigma_Y \sigma_Y} = \frac{\sum_{i=1}^{N}(X_i - \bar{X})\,(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{N}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{N}(Y_i - \bar{Y})^2}} \qquad \text{Equation A.113}$$

$$r_s = 1 - \frac{6\sum_{i=1}^{n} {d_i}^2}{n(n^2-1)} \qquad \text{Equation A.114}$$

The PCC evaluation is done by correlating every segment in one testing group with every single segment in the other testing group (Figure A.32). This extensive process is done per segment to respect the stochastic nature of EEG epochs where every segment is considered to be wide-sense stationary which can be realized as a nonlinear system with different initial conditions.

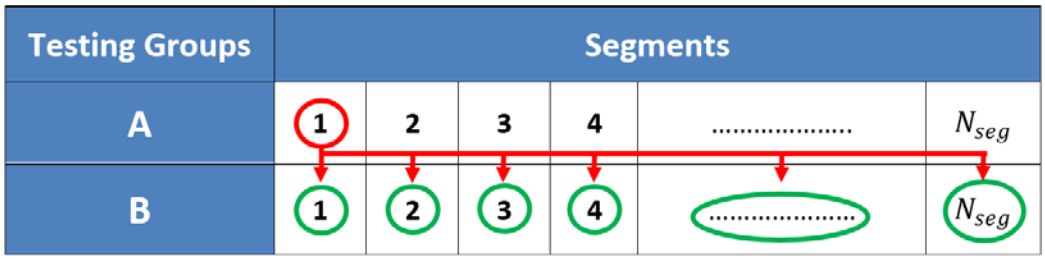

Figure A.32: The PCC evaluation procedure. Segment 1 in group A is correlated with every segment in group B. This process is repeated for every segment in group A producing a correlation distribution

PCC evaluation of the neonatal EEG corruption is done by repeating the correlation process for the time, frequency, and time-frequency representations for the different combinations of testing groups in the likelihood umbrellas. The total number of tested

segments N_{LCC} in the PCC evaluation procedure can be described using Equation A.115, where $N_{test\ group}$ represent the number of testing group's combination, N_{domain} is the number of domains, and N_{seg} is the number of segments per signal.

$$N_{LCC} = N_{test\ group} \times N_{domain} \times {N_{seg}}^2 = 7 \times 3 \times 200^2 = 840{,}000 \quad \text{Equation A.115}$$

The output of this extensive evaluation process can be interpreted as a correlation distribution where the mean of this distribution reflects the linear correlation coefficient between the different testing signals; in every group combination; in every likelihood umbrella.

The correlation distributions can be divided into three types: positive, negative, and mixed distributions. The positive correlation distribution is a distribution that totally lies in the positive correlation axis (Only positive correlation coefficients), while the negative correlation distribution is a distribution that lies totally in the negative correlation axis (Only negative correlation coefficients). The mixed distribution is a distribution that is combined from positive and negative distributions (positive and negative correlation coefficients). This happens mostly in stochastic time domain correlations as the relationship can alternate between positive and negative, but producing discontinuity at zero, which leads to consider this distribution as a combination of positive and negative distributions. In such cases positive and negative means are calculated with their respective probabilities resulting in two linear correlation coefficients for the mixed distributions.

H.2.2 Two Samples t-Test

The two samples t-test allows evaluation for the mean difference between two populations using data from two separate samples. This method is used in situations where no prior knowledge on either of the two populations under test. In particular, the

population means and standard deviations are all unknown, because the population variances are not known, and these values must be estimated from the sample data [128].

This evaluation technique relies on the condition that populations under test have to be normally distributed. This condition results in producing a random variable t that has a t-distribution with v degrees of freedom (Equation A.116 and Equation A.117). Where $\overline{x_1}$ and $\overline{x_2}$ are the means of the two samples, Δ is the hypothesized difference between population means, S_1 and S_2 are the standard deviations of the two samples, and n_1 and n_2 are the sizes of the two samples.

$$t = \frac{\overline{x_1} - \overline{x_2} - \Delta}{\sqrt{\frac{{S_1}^2}{n_1} + \frac{{S_2}^2}{n_2}}} \qquad \text{Equation A.116}$$

$$v = \frac{\left(\frac{{S_1}^2}{n_1} + \frac{{S_2}^2}{n_2}\right)^2}{\frac{\left({S_1}^2/n_1\right)^2}{n_1 - 1} + \frac{\left({S_2}^2/n_2\right)^2}{n_2 - 1}} \approx n_1 + n_2 - 2 \qquad \text{Equation A.117}$$

The Two Samples t-Test evaluation procedure is done by firstly calculating the means of the N_{seg} segments. This will produce a total number of N_{seg} means for every testing group (Figure A.33). These means are considered as a sampling distribution which follows a normal distribution as the number of segments increases (here we have 200) *"Central Limit Theorem"* [127].

The random variable t is calculated to test the zero mean difference hypothesis (Null Hypothesis $H_0: \Delta = 0$, alternative Hypothesis $H_1: \Delta \neq 0$) between; the different testing group's sampling distributions for every testing group's combination in every likelihood umbrella. The degrees of freedom (Equation A.117) for large number of

samples can be approximated to be $n_1 + n_2 - 2$, and because we have 200 segments this approximation is valid. Finally, after calculating the random variable t, error probabilities when rejecting the null hypothesis are calculated by accessing the t-table with $\alpha = 0.05$. These calculations are done in time, frequency, and time-frequency domains to compare the zero mean hypothesis for every testing group's combinations in every likelihood umbrella.

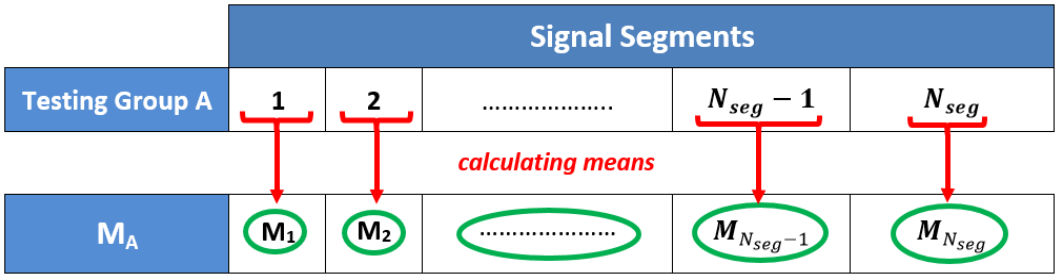

Figure A.33: The Two Samples t-Test evaluation procedure. Mean values are calculated for every testing group by calculating the mean value of every segment producing a sampling distribution that follows a normal distribution

H.2.3 One-Way Analysis of Variance

ANOVA is "a generalization of the t-Test; it compares means on a quantitative variable across any number of groups" [127]. This technique reduces the type I error conducted in the two samples t-Test because it performs a single omnibus test (F-test) that examines all the comparisons in the study as a set. The F-test in the one-way ANOVA, tests the null hypothesis that the means of all k populations are equal (Equation A.118).

$$H_0: \mu_1 = \mu_2 = \cdots = \mu_k \qquad \text{Equation A.118}$$

It produces a ratio (F-ratio) that determines the statistical significance of the results (Equation A.119). The F-ratio is distributed according to F-distribution with $k - 1$ and $N - k$ degrees of freedom. Where k is the number of groups, N is the total number

of samples, n_i is the number of samples for group i, $\tilde{\mu}_i$ is the estimated mean for group i, μ is the overall mean, n_k is the number of samples for group k, and x_{ij} is the element j for group i.

$$F_{ratio} = \frac{MS_{between}}{MS_{within}} = \frac{SS_{between}/(k-1)}{SS_{within}/(N-k)} = \frac{(N-k)\sum_{i=1}^{k} n_i(\tilde{\mu}_i - \mu)^2}{(k-1)\sum_{i=1}^{k}\sum_{j=1}^{n_k}(x_{ij} - \mu_i)^2} \quad \text{Equation A.119}$$

The ANOVA evaluation procedure is done by firstly calculating the means of the N_{seg} segments, exactly the same as the Two Samples t-Test evaluation procedure (Figure A.33). The F ratio is calculated to test the equal means hypothesis (Equation A.118) between all testing group's sampling distributions. After calculating the F ratio, error probabilities when rejecting the null hypothesis are calculated by accessing the F-table with $\alpha = 0.05$. These calculations are done in time, frequency, and time-frequency domains.

H.3 EEG Corruption Evaluation Results

H.3.1 Pearson's Correlation Coefficient

The PCC evaluation process for the time, frequency, and time-frequency representations for the two likelihood umbrellas is done by collecting statistically significant correlation coefficients with $\alpha = 0.05$ and representing their distribution. Distribution's positive and negative means with their corresponding probabilities are calculated to reflect the linear correlation coefficient between the testing groups. Initial results showed negligible correlation results for all testing groups in the mutual-likelihood umbrella, thus the BVP and ECGS contamination coefficients are doubled to become 3 and 12 respectively increasing the EEG corruption (Figure A.34, Figure A.35, and Table A.4).

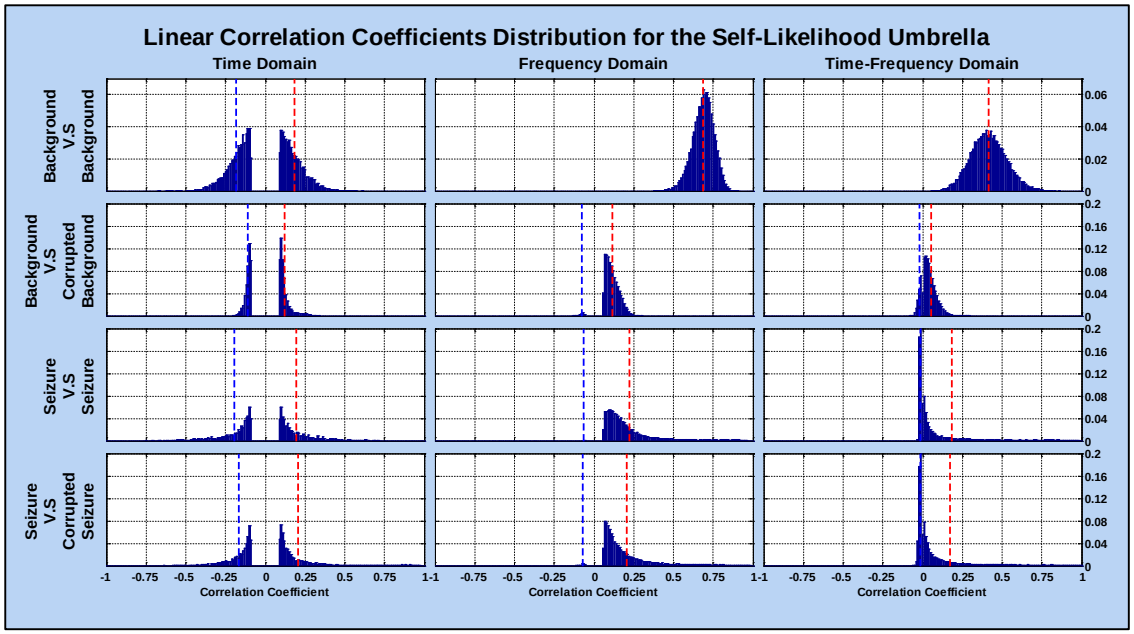

Figure A.34: Correlation distributions for the self-likelihood umbrella. Red dotted lines are positive means, while blue dotted lines are negative means. TFD parameters (EMB, N = 480, Fs = 32 Hz, FFT_N = 1024, Lag Window = 479, Alpha = 0.01, Beta = 0.9, time resolution = 1)

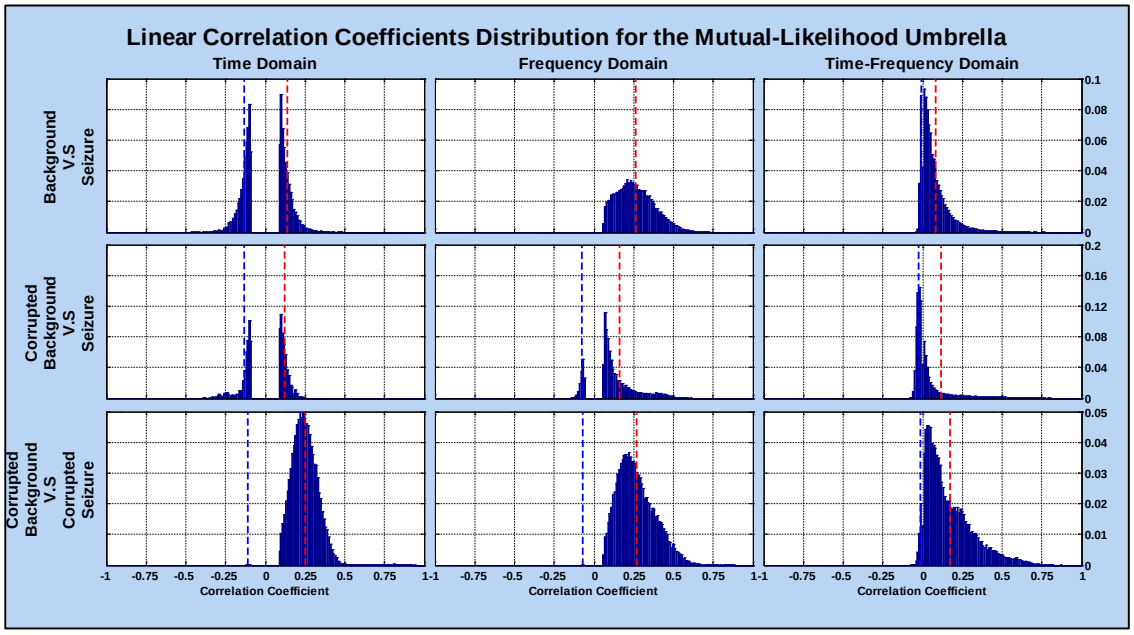

Figure A.35: Correlation distributions for the mutual-likelihood umbrella. Red dotted lines are positive means, while blue dotted lines are negative means. TFD parameters (EMB, N = 480, Fs = 32 Hz, FFT_N = 1024, Lag Window = 479, Alpha = 0.01, Beta = 0.9, time resolution = 1)

Table A.4: Detailed results of the PCC evaluation. r^+**:** positive correlation coefficient. r^-**:** negative correlation coefficient. p_r: probability. ***t*:** time. *f*: frequency. ***tf*:** time-frequency. X no value

Likelihood Umbrella	Testing Groups	Correlation Coefficients / Correlation Distributions Means					
		r_t^+ (p_r)	r_t^- (p_r)	r_f^+ (p_r)	r_f^- (p_r)	r_{tf}^+ (p_r)	r_{tf}^- (p_r)
Self-Likelihood	**Background V.S Background**	0.1820 (0.4977)	**-0.1817 (0.5023)**	**0.6866 (1)**	X (0)	**0.4137 (1)**	X (0)
	Background V.S Corrupted	**0.1200 (0.5445)**	-0.1102 (0.4555)	**0.1163 (0.9847)**	-0.0783 (0.0153)	**0.0514 (0.8324)**	-0.0185 (0.1676)
	Background Seizure V.S Seizure	0.1972 (0.4976)	**-0.1943 (0.5024)**	**0.2208 (0.9999)**	-0.0644 (0.0001)	**0.1819 (0.5313)**	-0.0129 (0.4687)
	Seizure versus Corrupted Seizure	**0.2049 (0.5223)**	-0.1684 (0.4777)	**0.2045 (0.9915)**	-0.0726 (0.0085)	**0.1728 (0.5706)**	-0.0156 (0.4294)
Mutual Likelihood	**Background V.S Seizure Corrupted**	**0.1363 (0.5130)**	-0.1342 (0.4870)	**0.2604 (1)**	X (0)	**0.0840 (0.8770)**	-0.0111 (0.1230)
	Background V.S Seizure Corrupted	**0.1203 (0.5397)**	-0.1315 (0.4603)	**0.1588 (0.8517)**	-0.0775 (0.1483)	**0.1139 (0.4535)**	-0.0252 (0.5465)
	Background V.S Corrupted Seizure	**0.2492 (0.9999)**	-0.1079 (0.0001)	**0.2662 (0.9999)**	-0.0682 (0.0001)	**0.1726 (0.9643)**	-0.0141 (0.0357)

Figure A.36 and Figure A.37 illustrate the linear correlation coefficients distributions for the self-likelihood and the mutual-likelihood umbrellas in the time,

frequency, and time-frequency domains. Positive and negative means are illustrated by dotted red and blue vertical lines respectively. In Figure A.36 the first row of plots (subplots 1-3) show the correlation distributions for the clean background versus clean background testing group combination, while the second row of plots (subplots 4-6) show the correlation distributions for the clean background versus corrupted background testing group combination. All means (positive and negative) of the second row in time, frequency, and time-frequency domains are below ±0.15 and less than the first row (especially for frequency and time-frequency domains). These low mean values of row two are interpreted as; clean background epochs are poorly correlated with corrupted backgrounds, and do not share much behaviour with its corrupted versions. The difference in the correlation coefficients means between row one and two is interpreted as; clean background epochs are more correlated with themselves than their corrupted versions.

The third row of plots (subplots 7-9) in Figure A.36 show the correlation distributions for the clean seizure versus clean seizure testing group combination, while the fourth row of plots show the correlation distributions for the clean seizure versus corrupted seizure testing group combination. All means (positive and negative) of the third and fourth rows in time, frequency, and time-frequency domains are below ±0.25 and similar in values. These low mean values are interpreted as; clean seizure epochs are poorly correlated with corrupted seizures, and do not share much behaviour with its corrupted versions. The similarity in the correlation coefficients means between row three and four is interpreted as; clean seizure epochs are poorly correlated with themselves and with their corrupted versions.

In Figure A.37 the first row of plots (subplots 1-3) show the correlation distributions for the clean background versus clean seizure testing group combination, the second row of plots (subplots 4-6) show the correlation distributions for the corrupted background versus clean seizure testing group combination, and the third row of plots (subplots 7-9) show the correlation distributions for the corrupted background versus corrupted seizure testing group combination. All means (positive and negative) of the first and second rows in time, frequency, and time-frequency domains are below ±0.3. These low mean values are interpreted as; clean background epochs are poorly correlated with clean seizures, and corrupted background are also poorly correlated with clean seizures. All means of the third row in time, frequency, and time-frequency domains are above +0.4. These relatively high mean values indicate high relative correlation between corrupted background epochs and corrupted seizure epochs, consequently resulting in difficulty when differentiating between the two of them.

Table A.5 illustrate detailed results, showing positive and negative means for every correlation distribution and their corresponding probabilities. The detailed results of Table A.5 match Figure A.36 and Figure A.37 outcomes. It shows relatively high correlation between corrupted background and corrupted seizure epochs, which proves the difficulty in distinguishing between those. Other correlations show negligible results (below ±0.3) which means possible identification between these groups. Since the initial amplitudes of the BVP and ECGS did not result in adequate EEG corruption while the doubled amplitude did, consequently the new doubled amplitudes will used through the rest of the results, and will be considered as the default values unless otherwise mentioned.

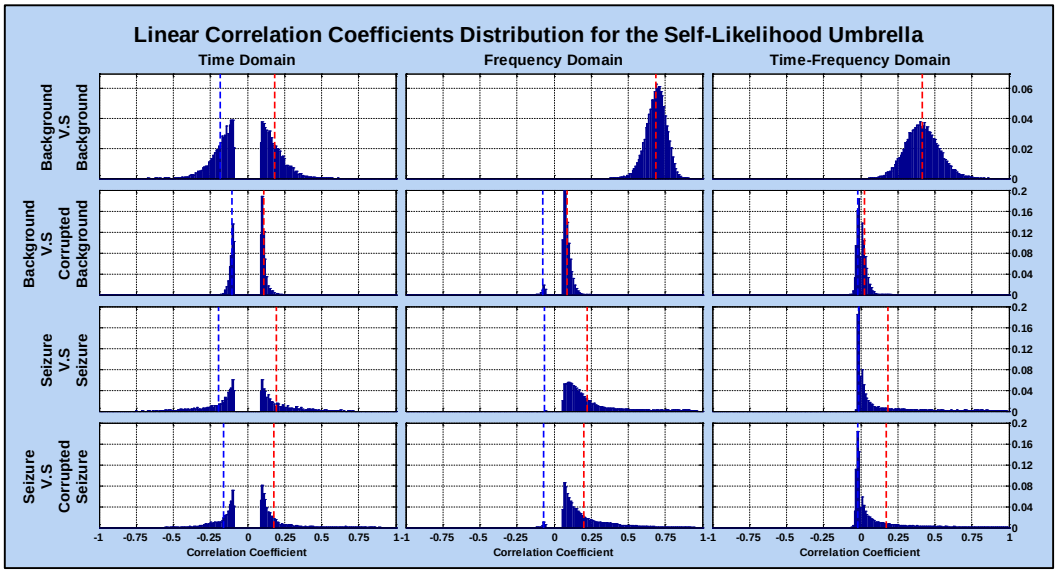

Figure A.36: Correlation distributions for the self-likelihood umbrella (BVP and ECGS amplitudes are doubled). Red dotted lines are positive means, while blue dotted lines are negative means. TFD parameters (EMB, N = 480, Fs = 32 Hz, FFT_N = 1024, Lag Window = 479, Alpha = 0.01, Beta = 0.9, time resolution = 1)

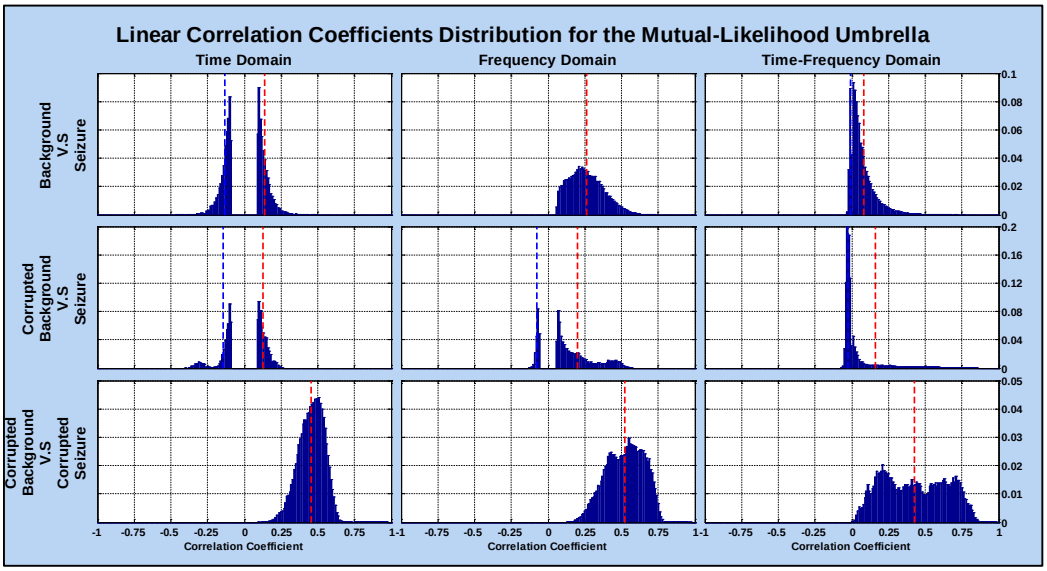

Figure A.37: Correlation distributions for the mutual-likelihood umbrella (BVP and ECGS amplitudes are doubled). Red dotted lines are positive means, while blue dotted lines are negative means. TFD parameters (EMB, N = 480, Fs = 32 Hz, FFT_N = 1024, Lag Window = 479, Alpha = 0.01, Beta = 0.9, time resolution = 1)

Table A.5: Detailed results of the PCC evaluation (BVP and ECGS amplitudes are doubled). r^+**:** positive correlation coefficient. r^-**:** negative correlation coefficient. p_r: probability. *t***:** time. *f*: frequency. *tf***:** time-frequency. X no value

Likelihood Umbrella	Testing Groups	Correlation Coefficients / Correlation Distributions Means					
		r_t^+ (p_r)	r_t^- (p_r)	r_f^+ (p_r)	r_f^- (p_r)	r_{tf}^+ (p_r)	r_{tf}^- (p_r)
Self-Likelihood	**Background V.S Background**	0.1820 (0.4977)	**-0.1817 (0.5023)**	**0.6866 (1)**	X (0)	**0.4137 (1)**	X (0)
	Background V.S Corrupted	**0.1101 (0.5798)**	-0.1071 (0.4202)	**0.0882 (0.9562)**	-0.0741 (0.0438)	**0.0273 (0.5208)**	-0.0185 (0.4792)
	Background Seizure V.S Seizure	0.1972 (0.4976)	**-0.1943 (0.5024)**	**0.2208 (0.9999)**	-0.0644 (0.0001)	**0.1819 (0.5313)**	-0.0129 (0.4687)
	Seizure versus Corrupted Seizure	**0.1802 (0.5482)**	-0.1601 (0.4518)	**0.1990 (0.9791)**	-0.0719 (0.0209)	**0.1746 (0.5246)**	-0.0200 (0.4754)
Mutual Likelihood	**Background V.S Seizure Corrupted**	**0.1363 (0.5130)**	-0.1342 (0.4870)	**0.2604 (1)**	X (0)	**0.0840 (0.8770)**	-0.0111 (0.1230)
	Background V.S Seizure Corrupted	**0.1296 (0.5538)**	-0.1433 (0.4462)	**0.1984 (0.7930)**	-0.0737 (0.2070)	0.1628 (0.3278)	**-0.0254 (0.6722)**
	Background V.S Corrupted Seizure	**0.4571 (1)**	X (0)	**0.5242 (1)**	X (0)	**0.4284 (1)**	X (0)

H.3.2 Two Samples t-Test

The Two Samples t-Test evaluation process for the time, frequency, and time-frequency representations for the different testing groups is done by calculating the mean value for every testing group's segments, then generating the mean sampling distributions for every testing group. Sampling distributions absolute mean differences are calculated to test the null hypothesis of equal means along with the error

probabilities when rejecting the null hypothesizes at 5% significance for every group combination.

Figure A.38 illustrates the sampling distributions for every testing group's in time, frequency, and time-frequency domains. Mean values are illustrated by red dotted vertical lines. The first row (subplots 1-3) show the sampling distributions for the first testing group (Background) in time, frequency, and time-frequency domains. The second row (subplots 4-6) show the sampling distributions for the second testing group (Seizure) in time, frequency, and time-frequency domains. The third row (subplots 7-9) show the sampling distributions for the third testing group (Corrupted Background) in time, frequency, and time-frequency domains, and finally the fourth row (subplots 10-12) show the sampling distributions for the fourth testing group (Corrupted Seizure) in time, frequency, and time-frequency domains.

Table A.6 illustrate the absolute mean differences for the two likelihood umbrellas, along with the error probabilities when rejecting the null hypothesis. For the self-likelihood umbrella, the clean background versus corrupted background, and the clean seizure versus corrupted seizure show negligible error probabilities when rejecting the null hypothesis of equal means for all; time, frequency, and time-frequency domains. This suggests that clean background epochs and corrupted background epochs are relatively easy to be identified in the time, frequency, and time-frequency domains, as well as clean seizure epochs and corrupted seizure epochs. For the mutual-likelihood umbrella, the corrupted background versus clean seizure show negligible error probabilities when rejecting the null hypothesis of equal means for all; time, frequency, and time-frequency domains. This suggests that corrupted background epochs and clean seizure epochs are relatively easy to be identified in the time, frequency, and time-

frequency domains. The clean background versus clean seizure, and the corrupted background versus corrupted seizure show negligible error probabilities when rejecting the null hypothesis of equal means for the frequency and time-frequency domains. On the other hand, they show 12.9% and 80.3% error probability for the time domain comparison respectively. This suggests that clean background and clean seizure epochs relatively difficult to be identified in the time domain, as well as corrupted background and corrupted seizure epochs. This proves the difficulty in distinguishing between corrupted background and corrupted seizure epochs in the time domain.

Table A.6: Detailed results of the Two Samples t-Test evaluation (BVP and ECGS amplitudes are doubled). μ_t, μ_f, and μ_{tf} are the time, frequency, and time-frequency domains sampling distribution means, p: error probability when rejecting the null hypothesis

Likelihood Umbrella	Testing Groups	Mean Difference		
		$\lvert\mu_{At} - \mu_{Bt}\rvert$ $(\boldsymbol{p})$	$\lvert\mu_{Af} - \mu_{Bf}\rvert$ $(\boldsymbol{p})$	$\lvert\mu_{Atf} - \mu_{Btf}\rvert$ $(\boldsymbol{p})$
Self-Likelihood	**Background V.S Background**	0 (1)	0 (1)	0 (1)
	Background V.S Corrupted Background	0.09584 (9.46e-28 ≈ 0)	0.027594 (5.11e-193 ≈ 0)	547.1867 (8.55e-106 ≈ 0)
	Seizure V.S Seizure	0 (1)	0 (1)	0 (1)
	Seizure versus Corrupted Seizure	0.09584 (1.89e-25 ≈ 0)	0.023424 (5.89e-113 ≈ 0)	553.3039 (3.64e-11 ≈ 0)
Mutual Likelihood	**Background V.S Seizure Corrupted**	**0.002901 (0.129239)**	0.012691 (3.74e-97 ≈ 0)	902.6766 (5.46e-45 ≈ 0)
	Background V.S Seizure Corrupted	0.092938 (1.97e-25 ≈ 0)	0.014903 (1.00e-73 ≈ 0)	355.4898 (4.08e-09 ≈ 0)
	Background V.S Corrupted Seizure	**0.002901 (0.803469)**	0.008522 (4.32e-26 ≈ 0)	908.7938 (6.74e-40 ≈ 0)

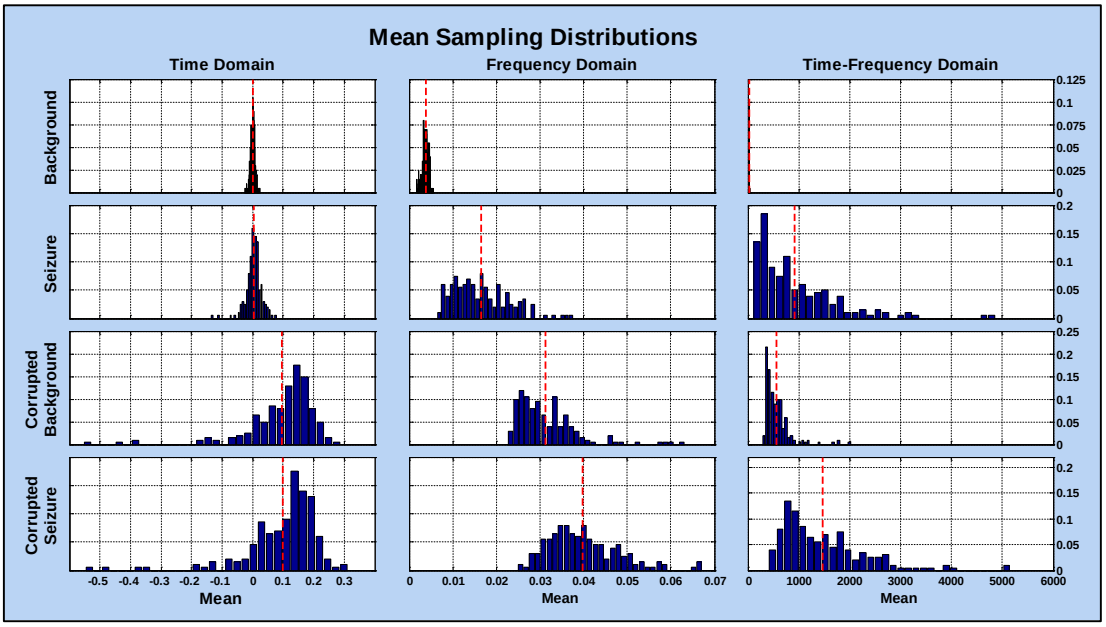

Figure A.38: Mean sampling distributions for every testing group (BVP and ECGS amplitudes are doubled). Red dotted lines are sampling distributions means. TFD parameters (EMB, N = 480, Fs = 32 Hz, FFT_N = 1024, Lag Window = 479, Alpha = 0.01, Beta = 0.9, time resolution = 1)

H.3.3 One-way Analysis of Variance

The ANOVA evaluation process for the time, frequency, and time-frequency representations for all testing groups is done by calculating the mean values for every testing group segments, then generating the mean sampling distributions for every testing group. Sampling distributions mean differences are calculated to test the null hypothesis of equal means along with the 95% confidence interval upper and lower limits. F-ratios and error probabilities when rejecting the null hypothesizes are calculated at 5% significance for all group combinations in all domains.

Table A.7, Table A.8, and Table A.9 illustrate negligible zero error probabilities when rejecting the null hypothesizes of equal means in all domains. This is due to the differences between the mean values when considering corrupted epochs (Figure A.39, Figure A.40, and Figure A.41). The F-ratios suggest that the testing groups are most

discriminated in the frequency domain, followed by the time-frequency domain, and finally by the time domain (Large F-ratios suggest more distance between means "discrimination/identification").

Table A.7: ANOVA evaluation results in time domain (ANOVA Table). *SS*: Sum of squared deviations, *df*: degrees of freedom, *F*: F-ratio, $p > F$: error probability when rejecting the null hypothesis

Source	SS	df	MS	F	$p > F$
Time Domain	1.8387	3	0.6129	87.9468	3.7170e-49
Error	5.5474	796	0.0070		
Total	7.3861	799			

Table A.8: ANOVA evaluation results in frequency domain (ANOVA Table). *SS*: Sum of squared deviations, *df*: degrees of freedom, *F*: F-ratio, $p > F$: error probability when rejecting the null hypothesis

Source	SS	df	MS	F	$p > F$
Frequency Domain	0.1535	3	0.0512	1338.6	2.0971e-310
Error	0.0304	796	3.8227e-05		
Total	0.1839	799			

Table A.9: ANOVA evaluation results in time-frequency domain (ANOVA Table). *SS*: Sum of squared deviations, *df*: degrees of freedom, *F*: F-ratio, $p > F$: error probability when rejecting the null hypothesis

Source	SS	df	MS	F	$p > F$
TF Domain	2.2463e+08	3	7.4876000	216.1568	1.5098e-102
Error	2.7573e+08	796	3.46400		
Total	5.0036e+08	799			

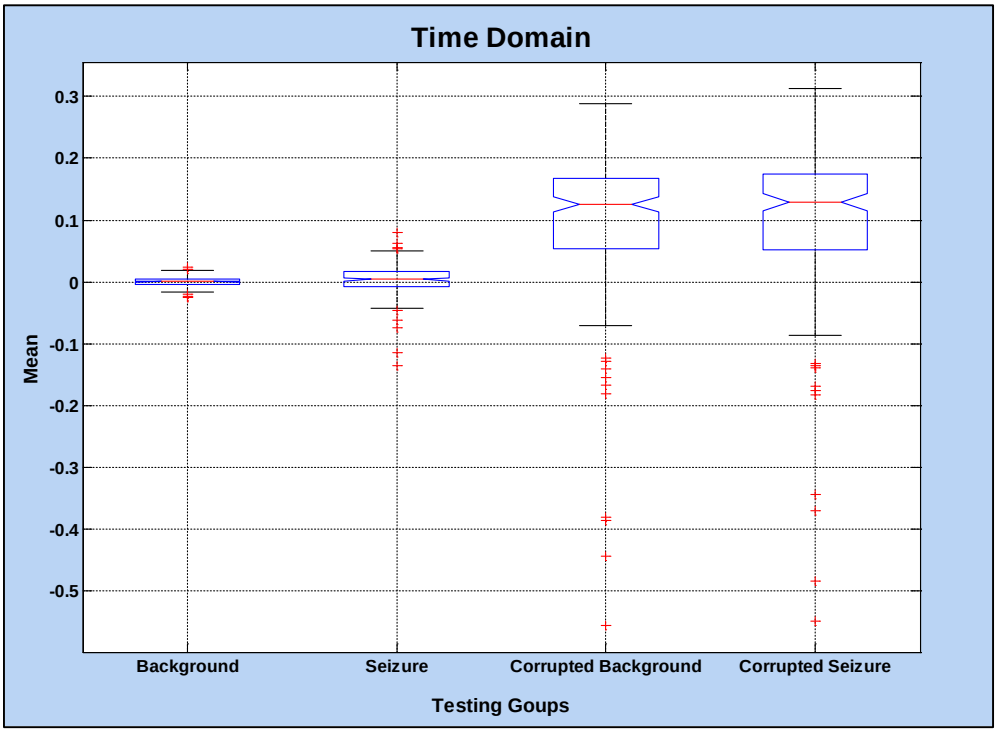

Figure A.39: ANOVA evaluation results in time domain (Boxplot)

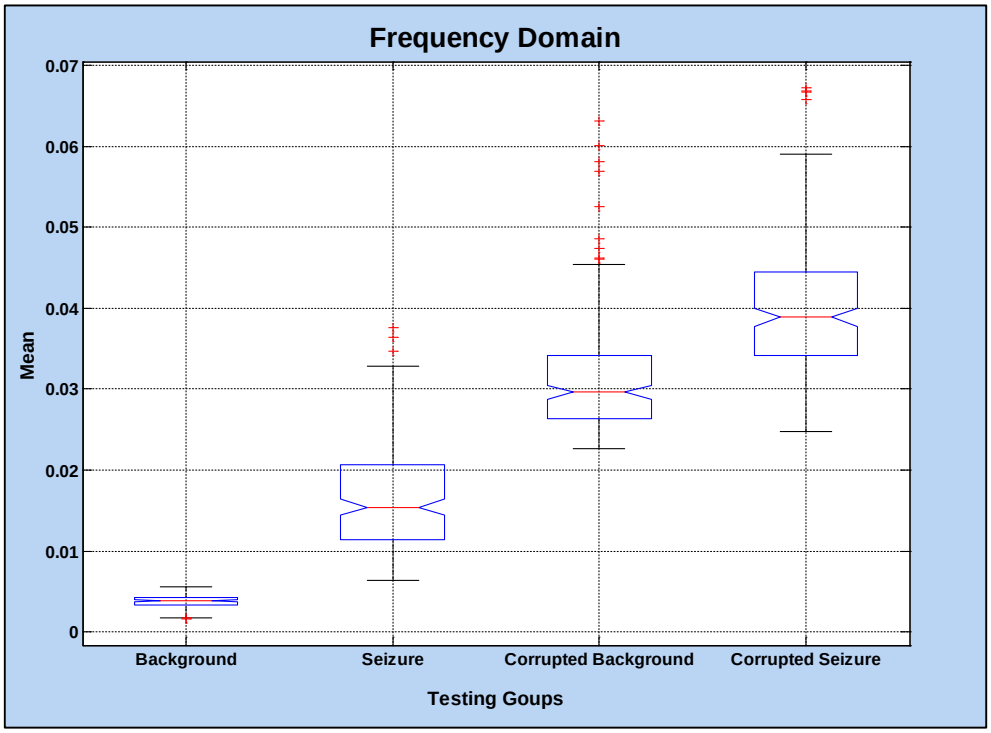

Figure A.40: ANOVA evaluation results in frequency domain (Boxplot)

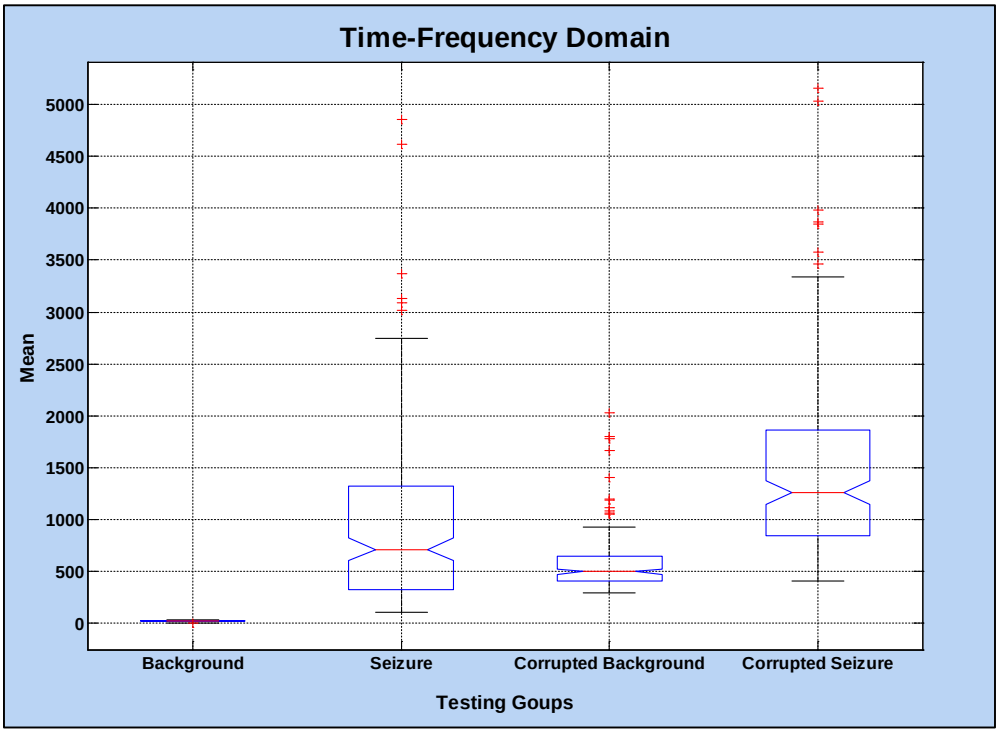

Figure A.41: ANOVA evaluation results in time-frequency domain (Boxplot)

Table A.10, Table A.11, and Table A.12 show the Tukey's Honest Significant Difference (Tukey's HSD) results, which is used to identify the means that are significantly different from each other.

Table A.10 illustrate the Tukey's HSD results for all the testing groups in the time domain. The 95% confidence intervals for the clean background versus clean seizure and the corrupted background versus corrupted seizure group combinations include the zero mean difference. This suggests that background versus seizure epochs, and corrupted background versus corrupted seizure epochs are relatively difficult to identify in the time domain. These results can be visualized in Figure A.39.

Table A.11 illustrate the Tukey's HSD results for all the testing groups in the frequency domain. The 95% confidence for all testing groups do not include the zero mean different. On the other hand the corrupted background versus corrupted seizure testing

group 95% confidence interval values are relatively smaller when compared to the other testing groups. This suggests that corrupted background and corrupted seizure epochs are the hardest to distinguish between them in the frequency domain. These results can be visualized in Figure A.40.

Table A.12 illustrate the Tukey's HSD results for all the testing groups in the time-frequency domain. The 95% confidence intervals for all combination groups do not include the zero mean difference. These results can be visualized in Figure A.41.

Table A.10: Tukey's HSD results in time domain, [L] 95% confidence interval lower bound, [U] 95% confidence interval upper bound, μ_t: time domain sampling distribution mean. B: Background, S: Seizure, CB: Corrupted Background, CS: Corrupted Seizure

Likelihood Umbrella	Testing Groups	Means Difference		
		$[L]$	$\mu_{At} - \mu_{Bt}$	$[U]$
Self-Likelihood	**B versus CB**	-0.11729	-0.09584	-0.07439
	S versus CS	-0.11729	-0.09584	-0.07439
Mutual Likelihood	**B versus S**	**-0.02435**	**-0.0029**	**0.018545**
	CB versus S	-0.11438	-0.09294	-0.07149
	CB versus CS	**-0.02435**	**-0.0029**	**0.018545**

Table A.11: Tukey's HSD results in frequency domain, [L] 95% confidence interval lower bound, [U] 95% confidence interval upper bound, μ_f: frequency domain sampling distribution mean

Likelihood Umbrella	Testing Groups	Means Difference		
		$[L]$	$\mu_{At} - \mu_{Bt}$	$[U]$
Self-Likelihood	**B versus CB**	-0.02918	-0.02759	-0.02601
	S versus CS	-0.02501	-0.02342	-0.02184
Mutual Likelihood	**B versus S**	-0.01428	-0.01269	-0.0111
	CB versus S	-0.01649	-0.0149	-0.01331
	CB versus CS	**-0.01011**	**-0.00852**	**-0.00693**

Table A.12: Tukey's HSD results in time-frequency domain, [L] 95% confidence interval lower bound, [U] 95% confidence interval upper bound, μ_{tf}: time-frequency domain sampling distribution mean

Likelihood Umbrella	Testing Groups	Means Difference		
		$[L]$	$\mu_{At} - \mu_{Bt}$	$[U]$
Self-Likelihood	**B versus CB**	-698.388	-547.187	-395.985
	S versus CS	-704.505	-553.304	-402.103
Mutual Likelihood	**B versus S**	-1053.88	-902.677	-751.475
	CB versus S	204.2886	355.4898	506.6911
	CB versus CS	-1060	-908.794	-757.592

Appendix I: Probability-Probability Plots

The material presented in this appendix are used to explain probability-probability plots in details as it was used to validate the EEG propagation model in Section 3.2.6, page 62.

Probability-Probability (*P-P*) plots are graphical techniques for assessing the similarities between a distribution of data against a given distribution such as the normal or Weibull distributions. P–P plots can be used as graphical adjunct to test the fit of probability distributions. It plots the cumulative distribution function (*CDF*) of a variable against the CDF of the reference distribution. Data under-analysis are ranked and sorted, and the corresponding z-score is calculated for each rank using Equation A.120, where S is the standard deviation. This is the expected value that data should have in a normal distribution. The computed scores are then converted to z-scores and the actual z-scores are plotted against the expected z-scores. If the data are normally distributed, the result would be a straight diagonal line [129].

This methodology can be extended to test the alikeness of N distributions, as P-P plots can be plotted for two processes and if the processes have the same distribution (does not have to be normal) their semi-diagonal lines would be on each other.

$$Z = \frac{X - \bar{X}}{S} \qquad \text{Equation A.120}$$

Appendix J: Supplementary Material

The material presented in this appendix include all supplementary results that do not contribute to the main research question, or have been moved due its large quantity. These results are referred to in Chapter 5, page 88 (Results and Discussions).

J.1 EEG Propagation Model Validation

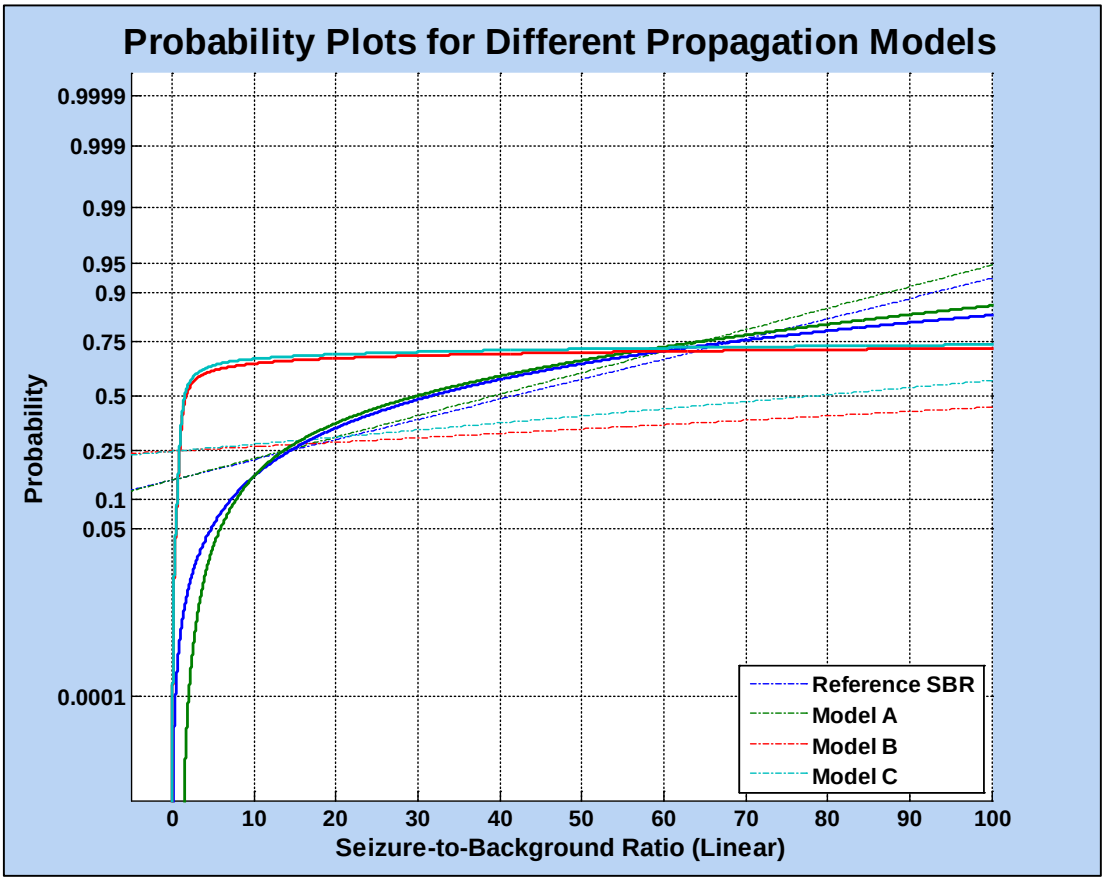

Figure A.42: P-P plot of models A, B, and C output (Linear) SBR against the reference distribution

J.2 Neonatal EEG Artifacts Model

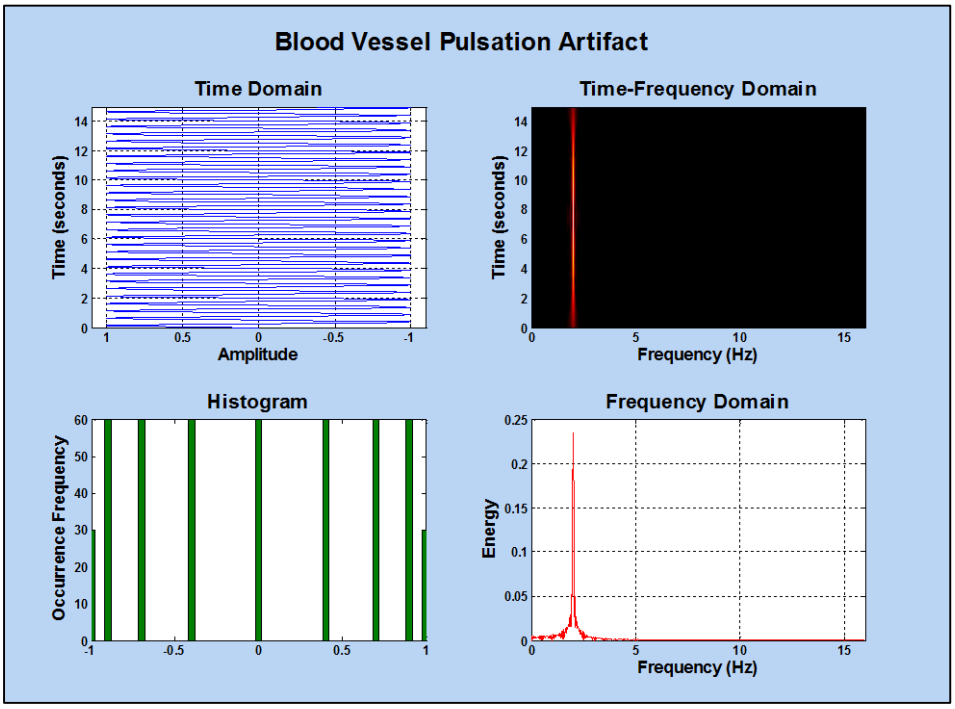

Figure A.43: Simulated blood vessel pulsation artifact. TFD parameters (MB, N = 480, Fs = 32 Hz, FFT_N = 1024, Lag Window = 479, Alpha = 0.01, time resolution = 1). Time and frequency steps are (0.0313, 0.0313) respectively

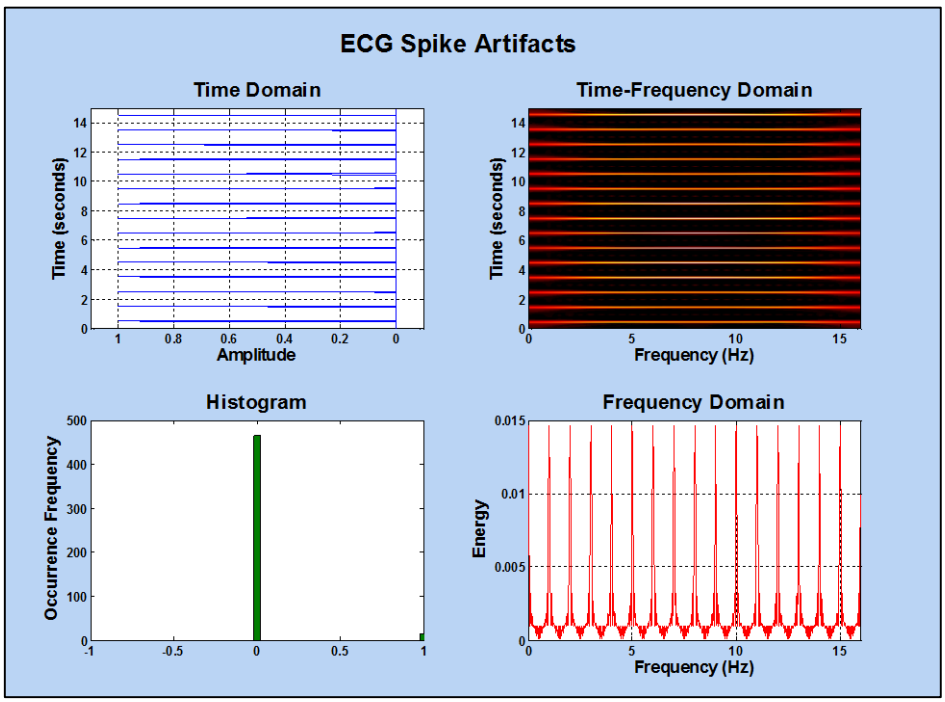

Figure A.44: Simulated ECG Spike artifact. TFD parameters (EMB, N = 480, Fs = 32 Hz, FFT_N = 1024, Lag Window = 479, time resolution = 1, Alpha = 0.99, Beta = 0.01). Time and frequency steps are (0.0313, 0.0313) respectively

J.3 Artifact Detection Optimisation

Detection optimisation maximises the detection balanced accuracy by spanning all possible correlation thresholds. This procedure is done for the time and time-frequency BSS algorithms.

Time and time-frequency BSS optimisation results consist of 100,100 and 303 samples for each performance curve respectively, thus average curves are drawn and denoted by a thick black curve (Figure A.45-Figure A.49 and Figure A.51-Figure A.55). The detection thresholds that maximises the balanced accuracy for all BSS algorithms are extracted, normal distributions are fitted and mean values and standard deviations are calculated along with their 95% confidence intervals (Figure A.50, Table A.13, Figure A.56, and Table A.14).

J.3.1 Time Domain Artifact Detection

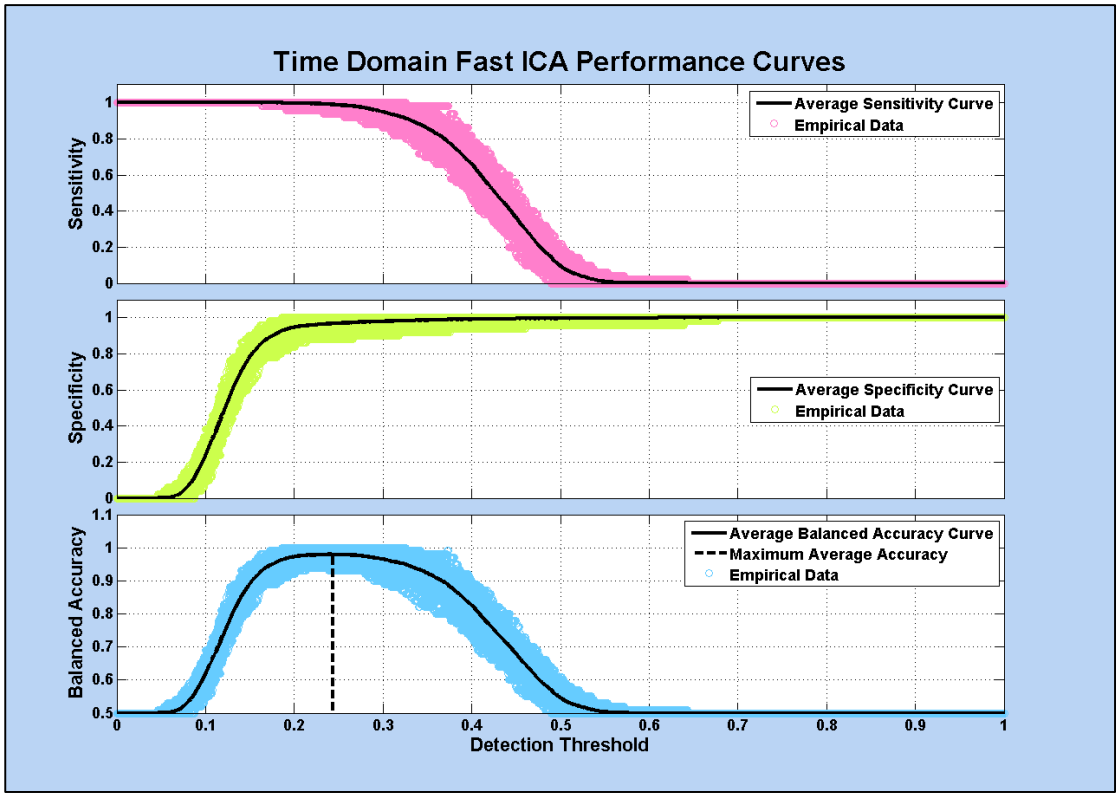

Figure A.45: Time domain Fast ICA detection performance curves

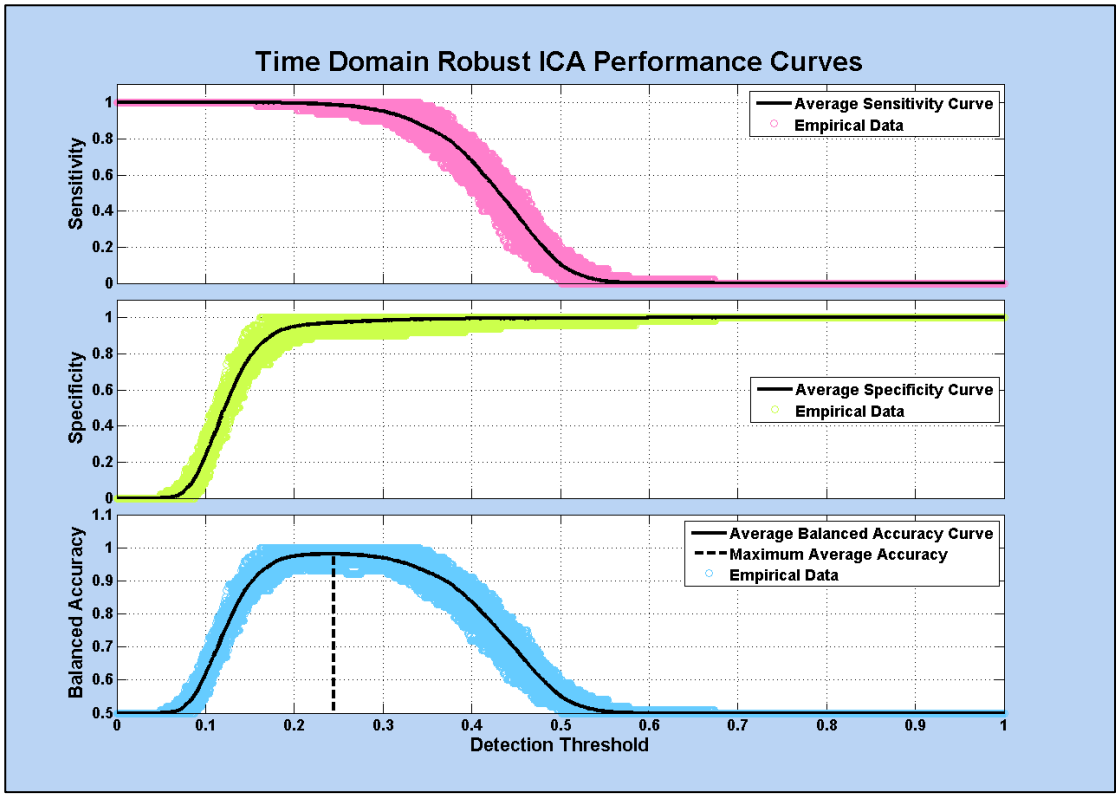

Figure A.46: Time domain Robust ICA detection performance curves

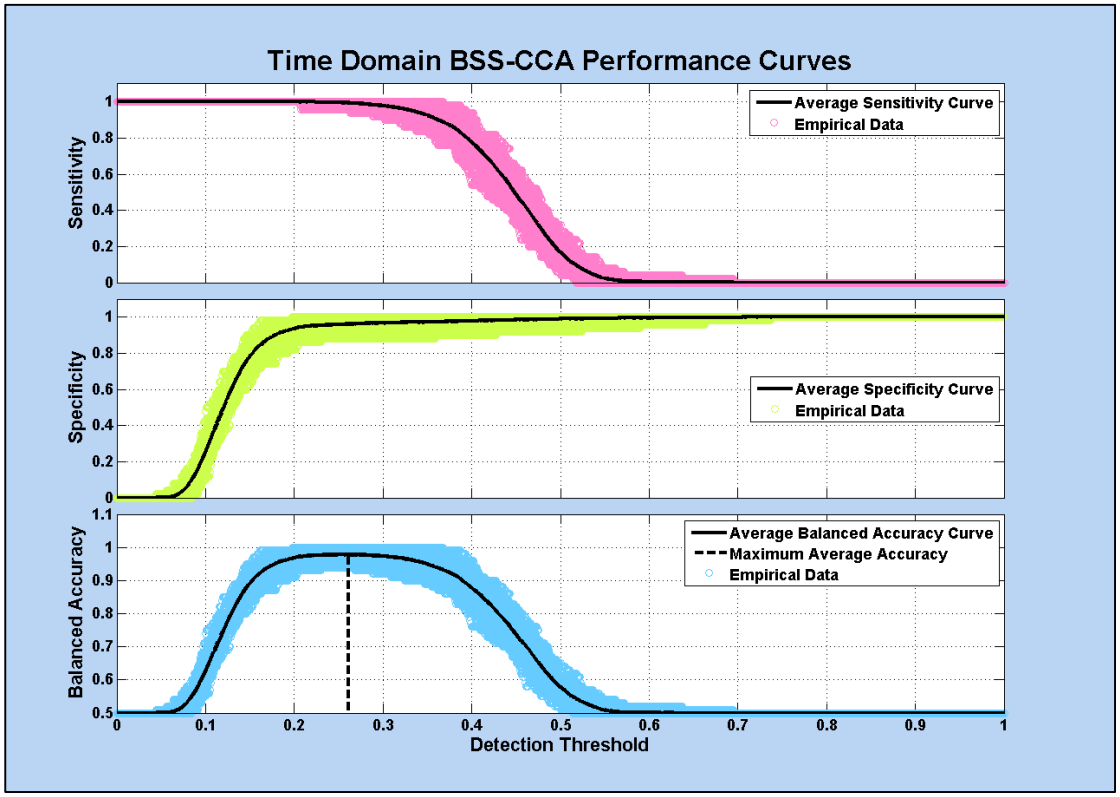

Figure A.47: Time domain BSS-CCA detection performance curves

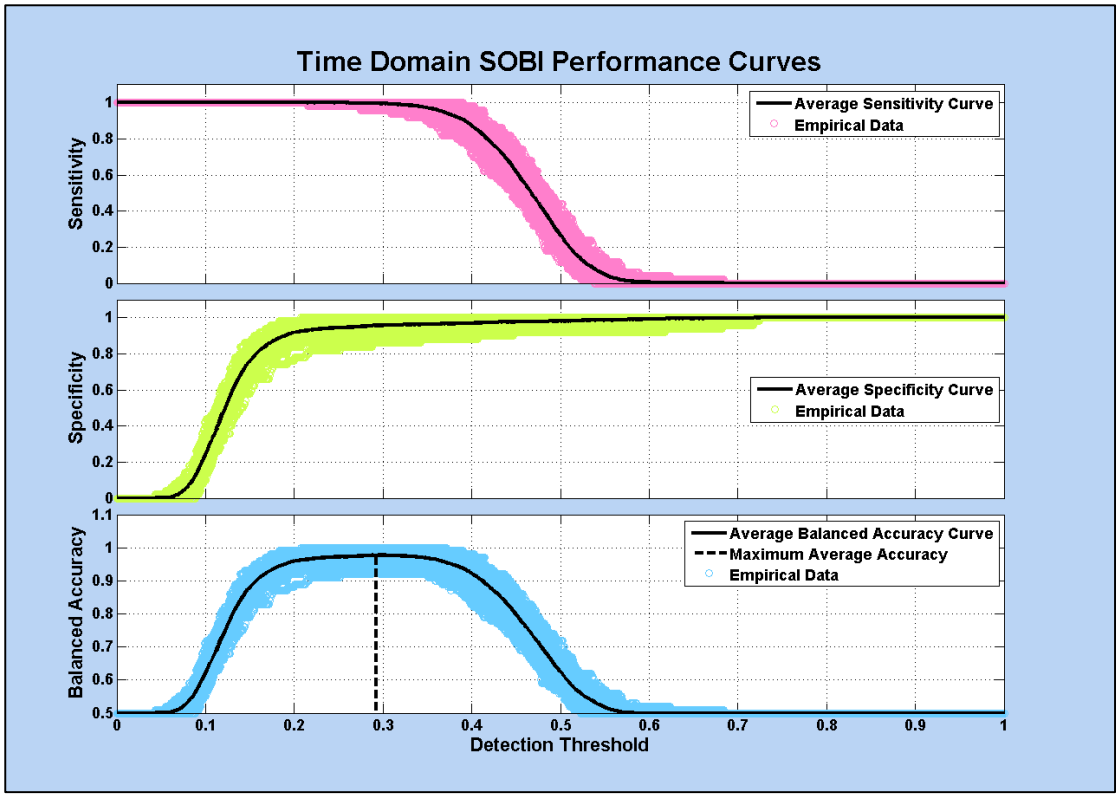

Figure A.48: Time domain SOBI detection performance curves

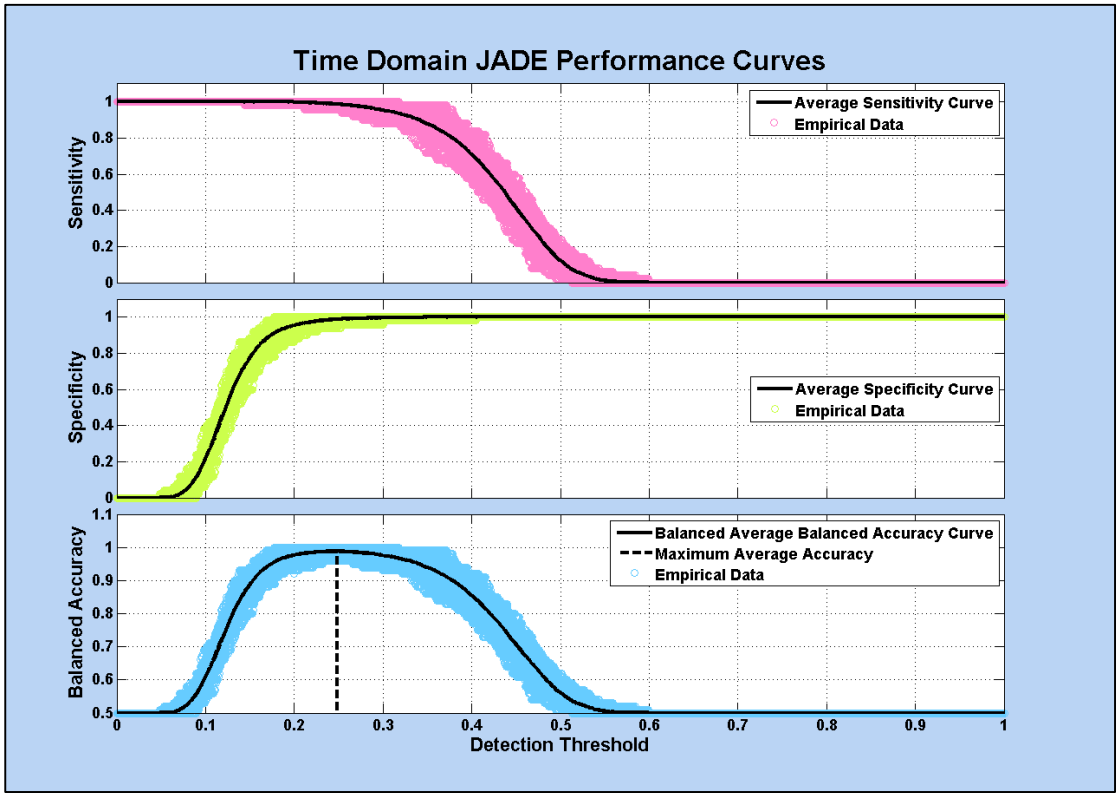

Figure A.49: Time domain JADE detection performance curves

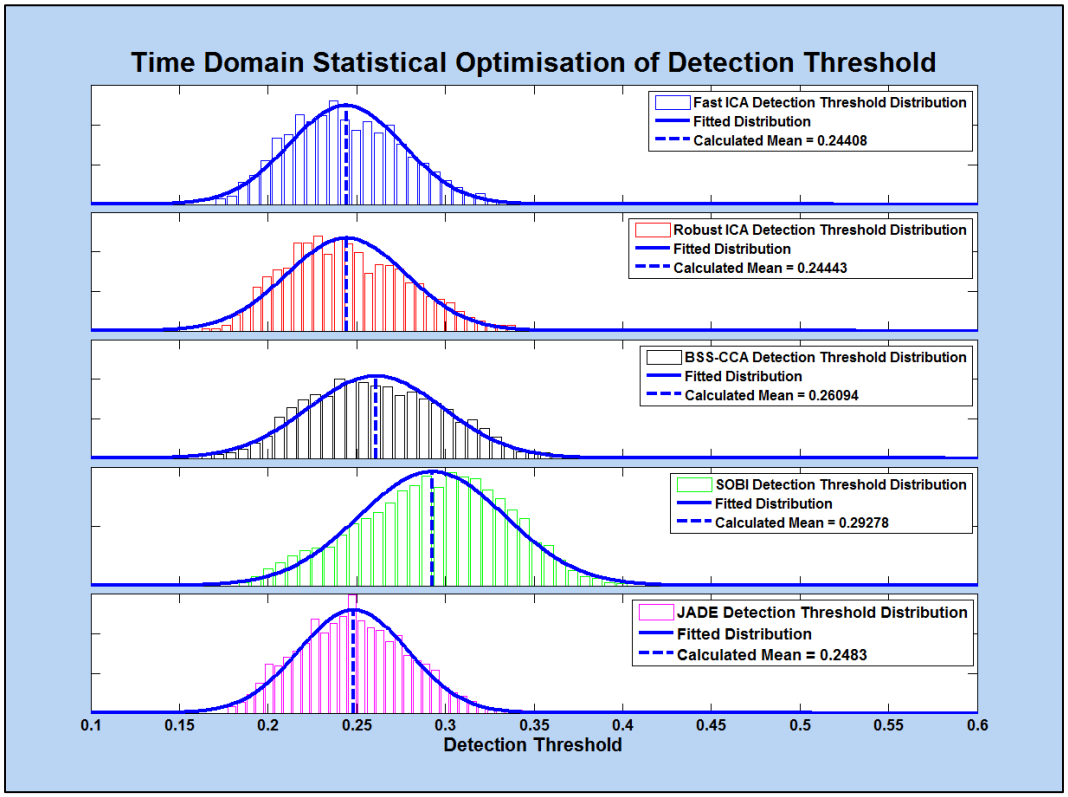

Figure A.50: Time domain optimised correlation thresholds for all BSS algorithms

Table A.13: Time domain artifact detection optimisation results. μ_L is the mean 95% confidence interval lower bound, μ_U is the mean 95% confidence interval upper bound, μ is the detection threshold mean, σ_L is the standard deviation 95% confidence interval lower bound, σ_U is the standard deviation 95% confidence interval upper bound, σ is the detection threshold standard deviation

BSS Algorithms	**Time Domain Optimised Detection Threshold Parameters**					
	μ_L	$\boldsymbol{\mu}$	μ_U	σ_L	$\boldsymbol{\sigma}$	σ_U
Fast-ICA	0.243130	**0.244078**	0.245026	0.031509	**0.032165**	0.032849
Robust-ICA	0.243432	**0.244430**	0.245429	0.033351	**0.034042**	0.034763
BSS-CCA	0.259972	**0.260944**	0.261916	0.037681	**0.038356**	0.039055
SOBI	0.291865	**0.292784**	0.293704	0.040910	**0.041549**	0.042210
JADE	0.247396	**0.248297**	0.249198	0.030231	**0.030856**	0.031506

J.3.2 Time-Frequency Domain Artifact Detection

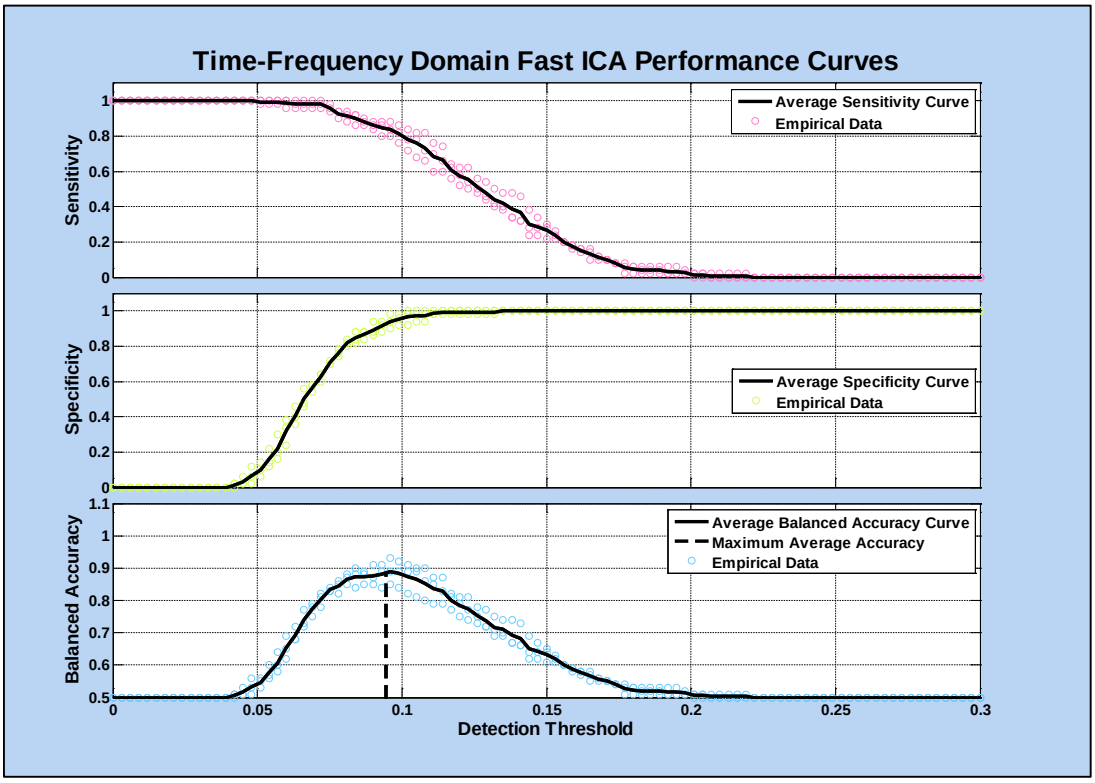

Figure A.51: Time-Frequency domain Fast ICA detection performance curves

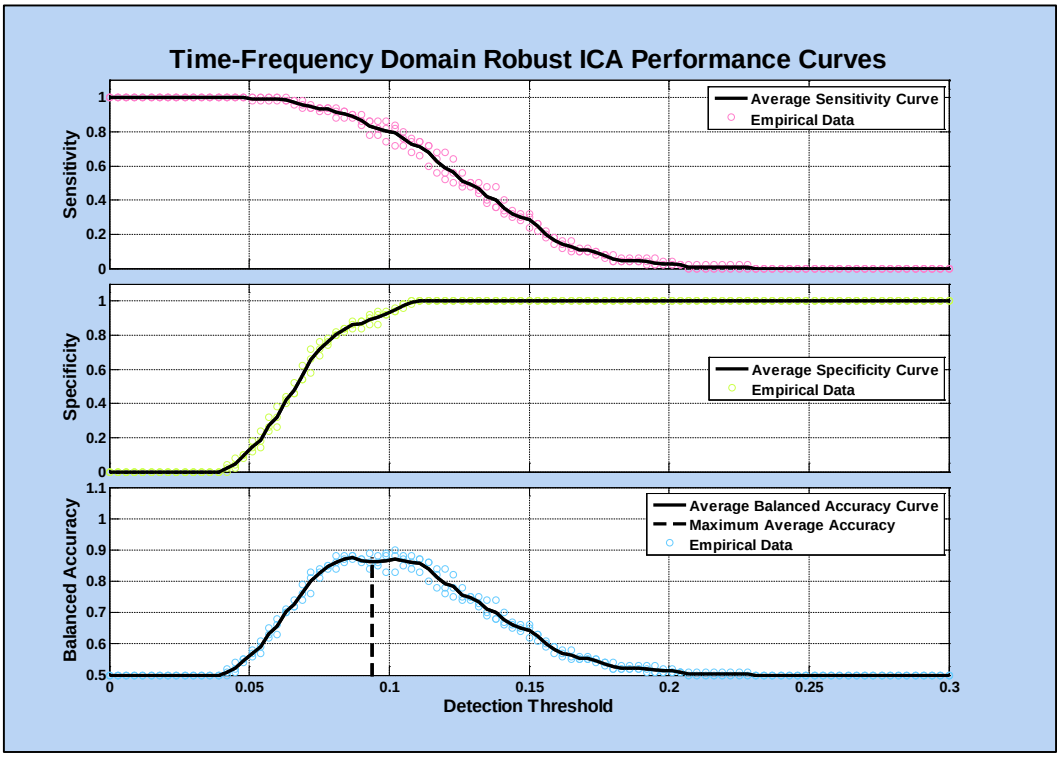

Figure A.52: Time-Frequency domain Robust ICA detection performance curves

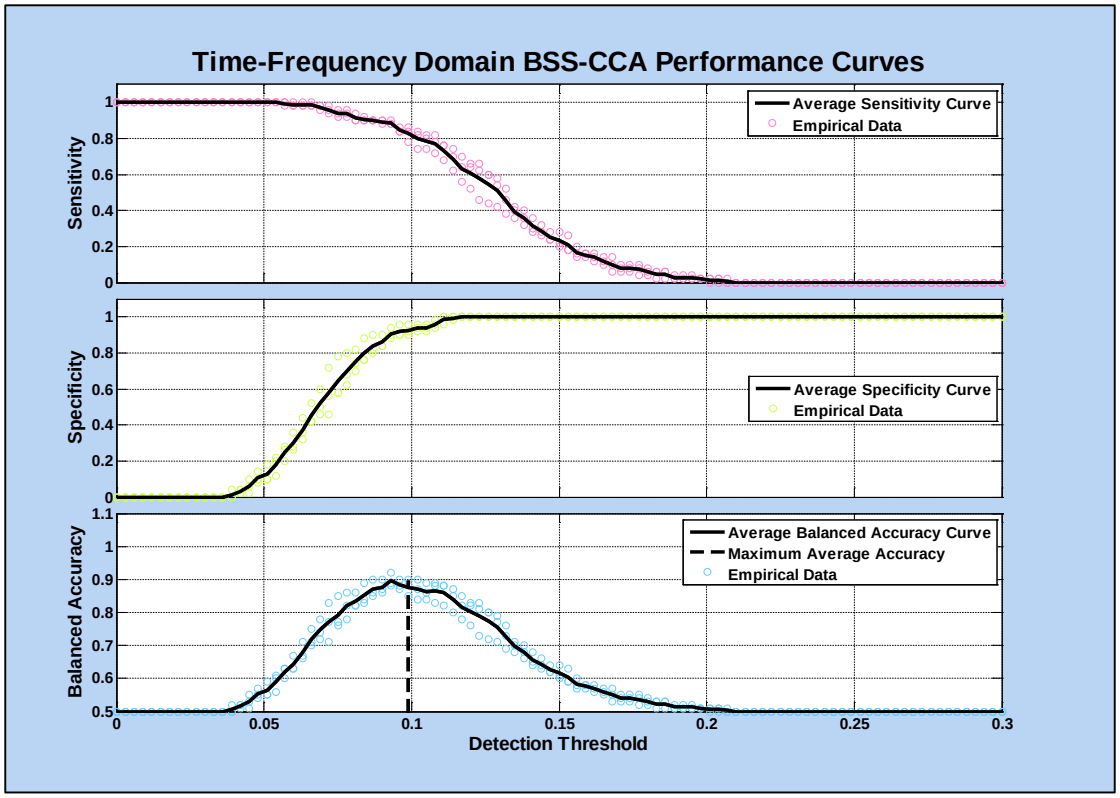

Figure A.53: Time-Frequency domain BSS-CCA detection performance curves

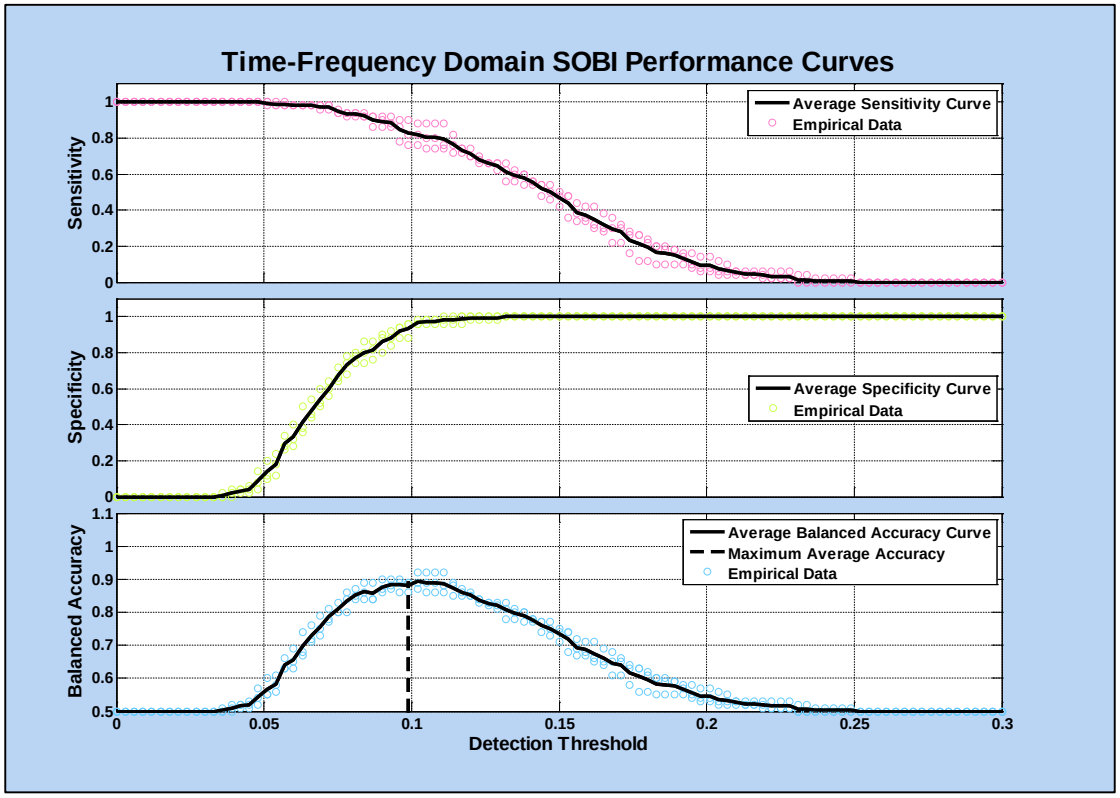

Figure A.54: Time-Frequency domain SOBI detection performance curves

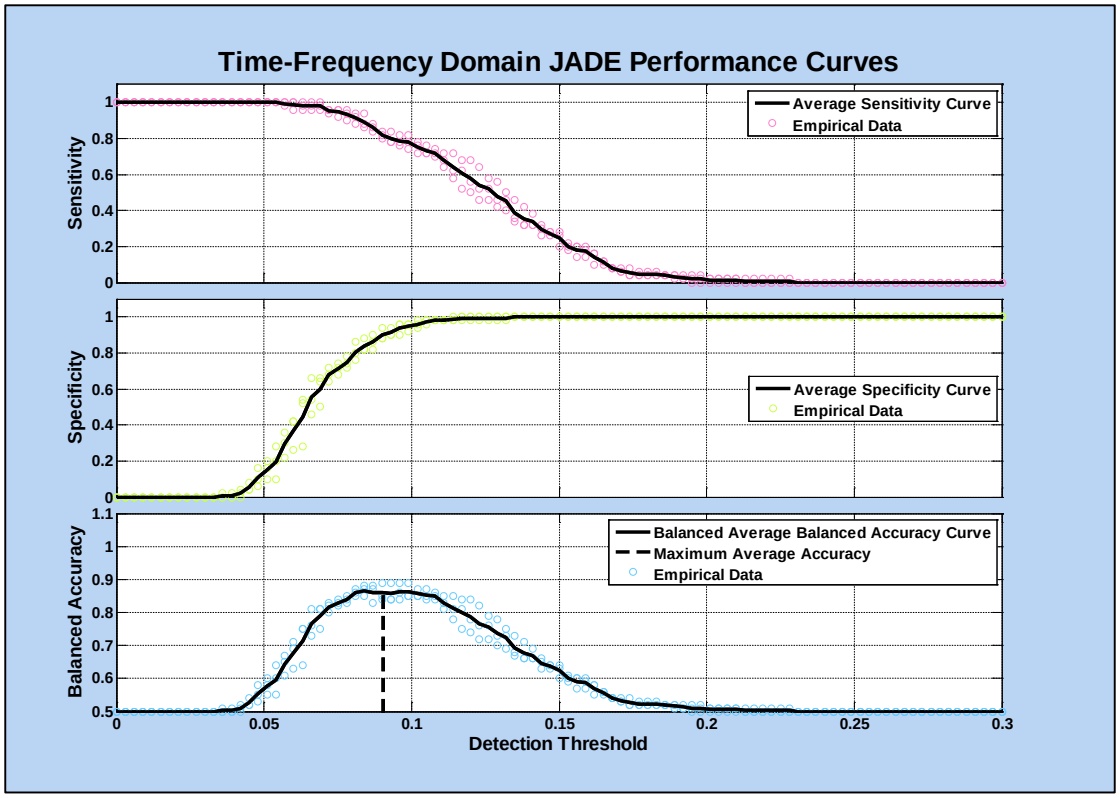

Figure A.55: Time-Frequency domain JADE detection performance curves

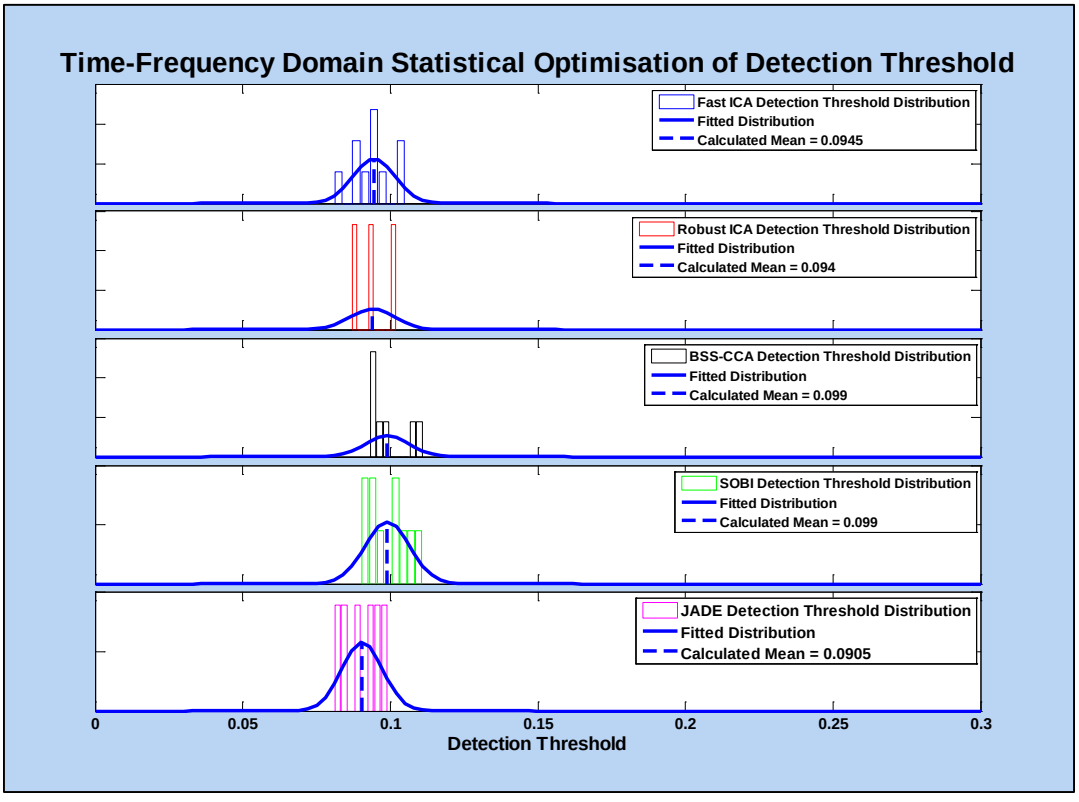

Figure A.56: Time-Frequency domain optimised correlation thresholds for all BSS algorithms

Table A.14: Time-Frequency domain artifact detection optimisation results. μ_L is the mean 95% confidence interval lower bound, μ_U is the mean 95% confidence interval upper bound, μ is the detection threshold mean, σ_L is the standard deviation 95% confidence interval lower bound, σ_U is the standard deviation 95% confidence interval upper bound, σ is the detection threshold standard deviation

BSS Algorithms	**Time-Frequency Domain Optimised Detection Threshold Parameters**					
	μ_L	$\boldsymbol{\mu}$	μ_U	σ_L	$\boldsymbol{\sigma}$	σ_U
Fast-ICA	0.089416	**0.0945**	0.099584	0.004888	**0.007106**	0.012973
Robust-ICA	0.075245	**0.094**	0.112755	0.003931	**0.00755**	0.047449
BSS-CCA	0.092018	**0.099**	0.105982	0.004865	**0.00755**	0.016625
SOBI	0.093552	**0.099**	0.104448	0.005238	**0.007616**	0.013903
JADE	0.083207	**0.0905**	0.097793	0.004338	**0.00695**	0.017045

J.4 Feature Selection

Table A.15: Fused multichannel time-frequency features ordered according to their Fisher's scores

#	Feature Label	Fisher Score
12	$\boldsymbol{F_7}$	10.89603343
9	$\boldsymbol{F_6}$	8.408607824
1	$\boldsymbol{T_1}$	6.802327383
11	$\boldsymbol{T_5}$	5.799446158
5	$\boldsymbol{F_4}$	5.637856925
3	$\boldsymbol{T_3}$	4.005305596
4	$\boldsymbol{T_4}$	2.911641887
2	$\boldsymbol{T_2}$	1.683835312
6	$\boldsymbol{F_1}$	0.790666701
8	$\boldsymbol{F_3}$	0.416335709
7	$\boldsymbol{F_2}$	0.364105607
16	$\boldsymbol{E_2}$	0.212894762
15	$\boldsymbol{E_1}$	0.105505172
10	$\boldsymbol{F_5}$	0
13	$\boldsymbol{IF_1}$	0
14	$\boldsymbol{IF_2}$	0

Table A.16: Combined extended and fused multichannel time-frequency features ordered according to their Fisher's scores

#	Feature Label	Fisher Score
17	$\boldsymbol{M_1}$	13.10198119
12	$\boldsymbol{F_7}$	10.89603343
9	$\boldsymbol{F_6}$	8.408607824
1	$\boldsymbol{T_1}$	6.802327383
11	$\boldsymbol{T_5}$	5.799446158
5	$\boldsymbol{F_4}$	5.637856925
3	$\boldsymbol{T_3}$	4.005305596
4	$\boldsymbol{T_4}$	2.911641887
2	$\boldsymbol{T_2}$	1.683835312
21	$\boldsymbol{F_1}$	0.790666701
6	$\boldsymbol{F_3}$	0.416335709
8	$\boldsymbol{F_2}$	0.364105607
7	$\boldsymbol{M_5}$	0.344553663
16	$\boldsymbol{E_2}$	0.212894762
19	$\boldsymbol{M_3}$	0.193284894
20	$\boldsymbol{E_4}$	0.105505172
15	$\boldsymbol{M_4}$	0.039815504
18	$\boldsymbol{M_2}$	0.03667936
10	$\boldsymbol{F_5}$	0
13	$\boldsymbol{IF_1}$	0
14	$\boldsymbol{IF_2}$	0

J.5 Abnormality Detection

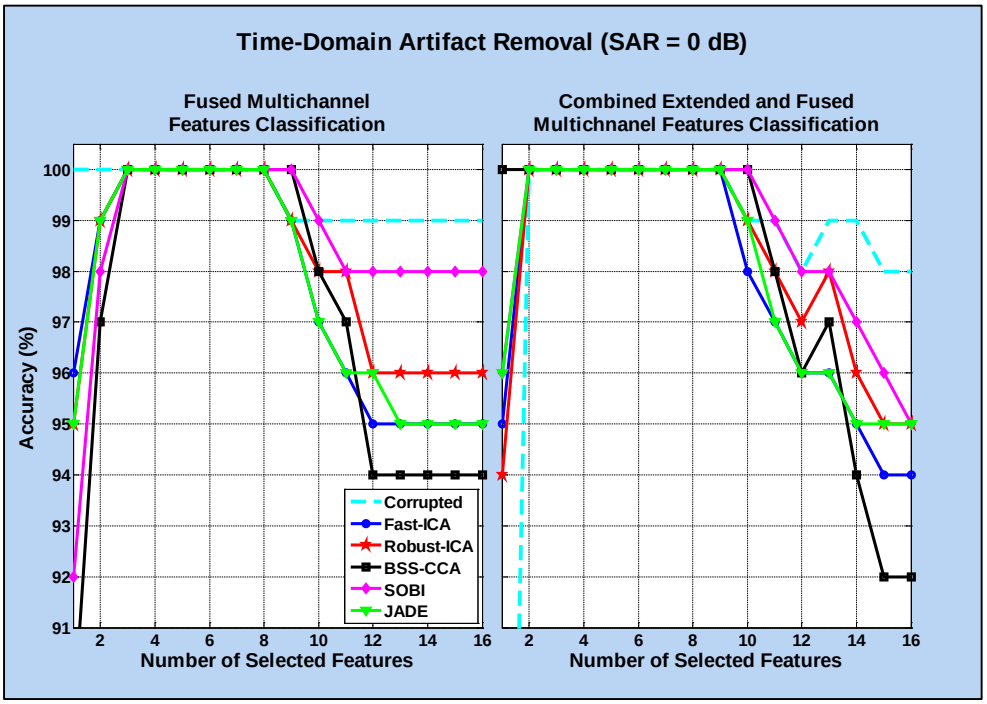

Figure A.57: Time domain filtered EEG abnormality detection accuracy profiles when using the fused and the combined extended and fused time-frequency multichannel features. SAR level is equal to 0 dB

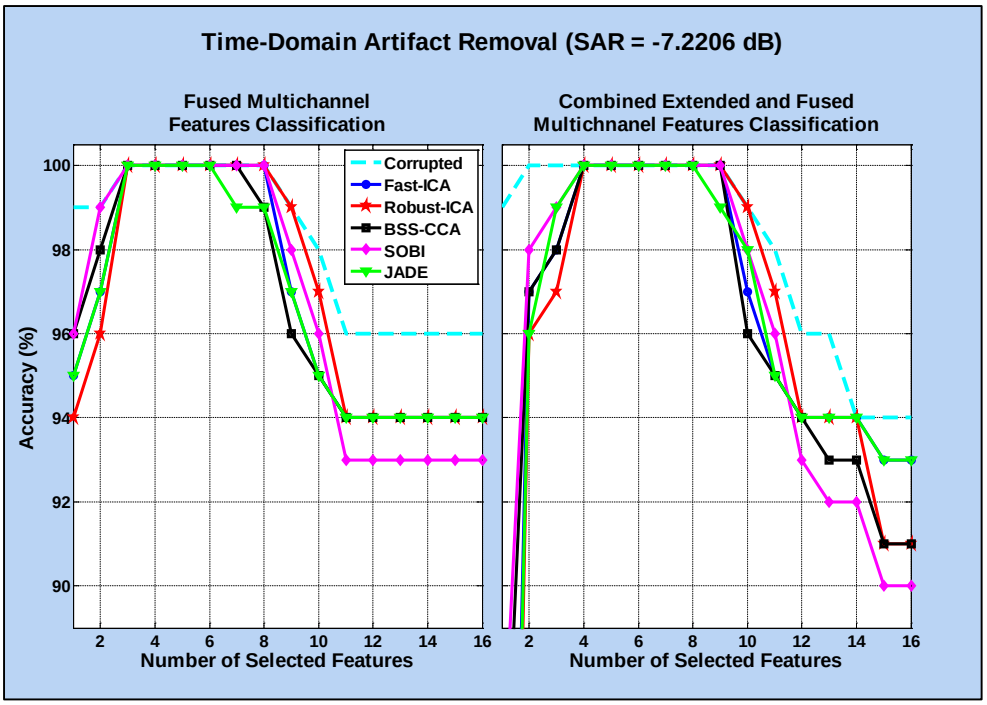

Figure A.58: Time domain filtered EEG abnormality detection accuracy profiles when using the fused and the combined extended and fused time-frequency multichannel features. SAR level is equal to -7.2206 dB

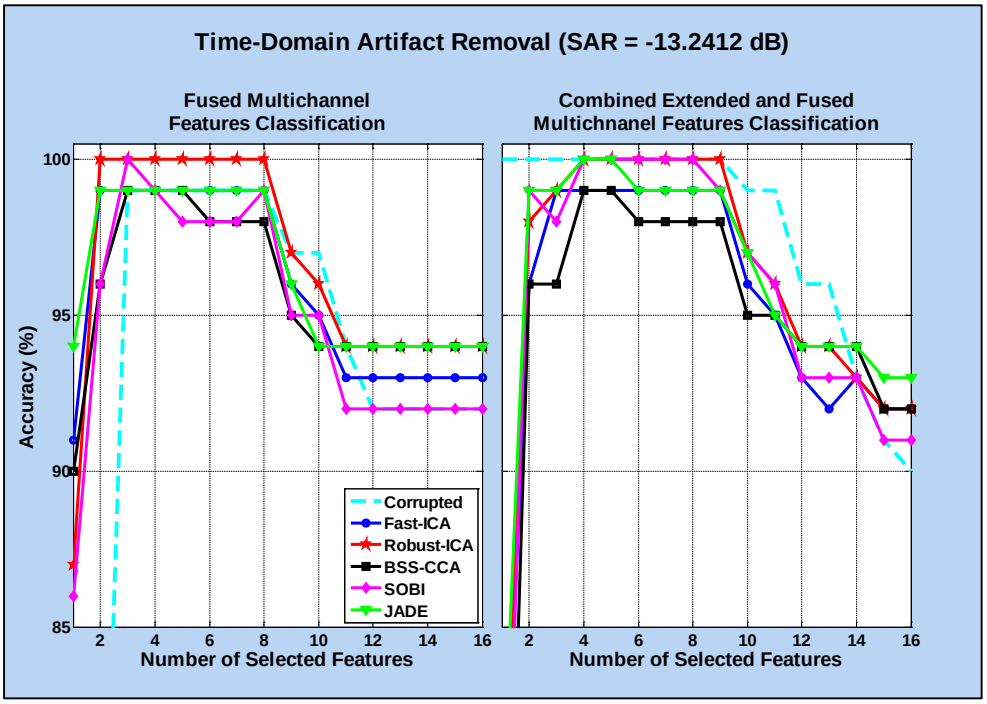

Figure A.59: Time domain filtered EEG abnormality detection accuracy profiles when using the fused and the combined extended and fused time-frequency multichannel features. SAR level is equal to -13.2412 dB

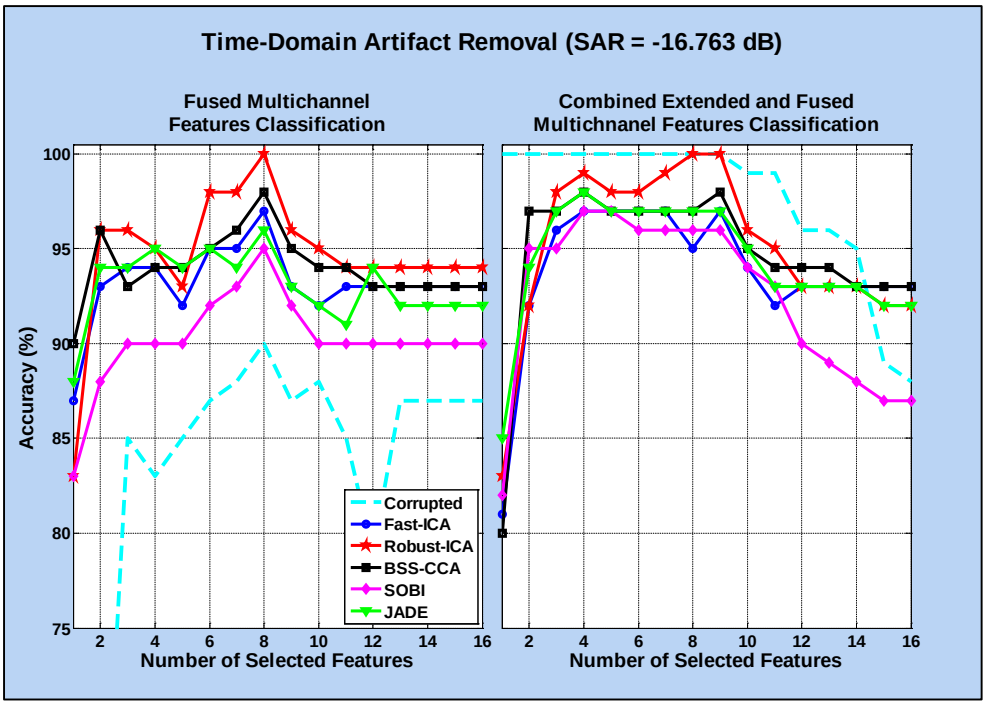

Figure A.60: Time domain filtered EEG abnormality detection accuracy profiles when using the fused and the combined extended and fused time-frequency multichannel features. SAR level is equal to -16.763 dB

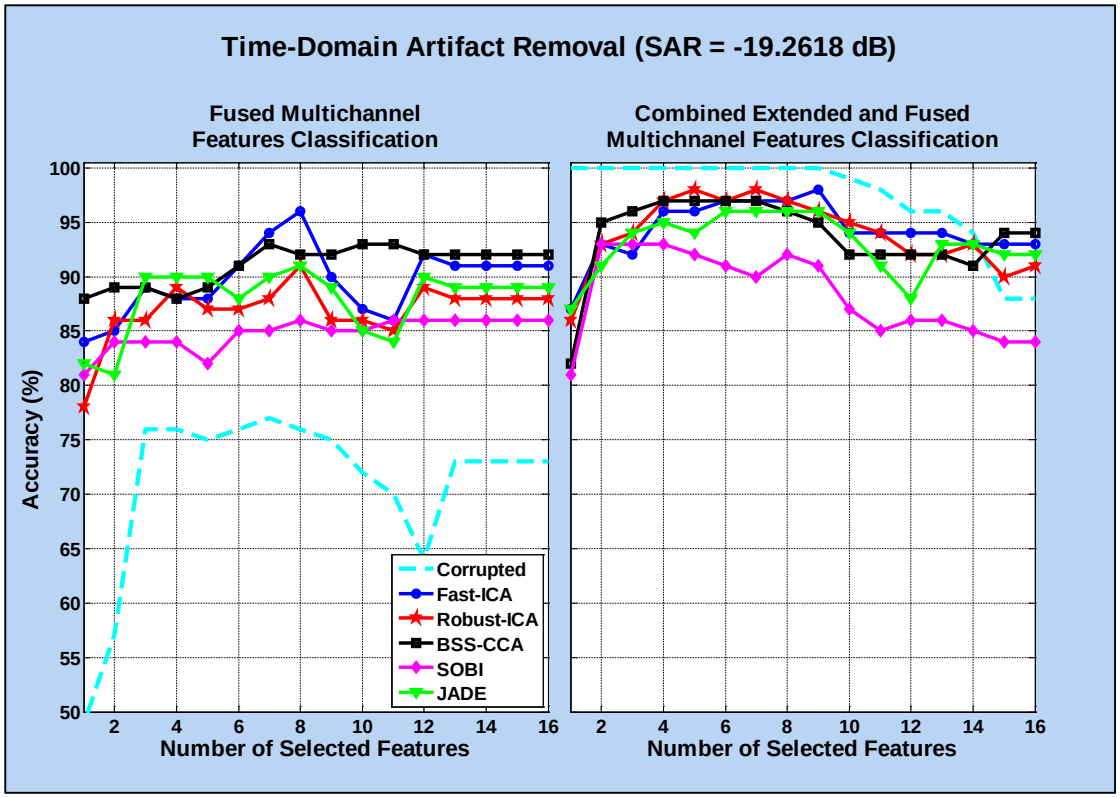

Figure A.61: Time domain filtered EEG abnormality detection accuracy profiles when using the fused and the combined extended and fused time-frequency multichannel features. SAR level is equal to -19.262 dB

Table A.17: Time domain filtered EEG abnormality detection average performances when using the fused and the combined extended and fused time-frequency multichannel features. SAR level is equal to 0 dB. *min* is the minimum value, *max* is the maximum value, μ is the mean value

	Fused Multichannel Time-Frequency Features								
	Balanced Accuracy (%)			**Sensitivity (%)**			**Specificity (%)**		
	min	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*
SAR ≈ 0 dB	99	**99.5**	100	100	**100**	100	98	**99**	100
Fast − ICA $_t$	95	**97.625**	100	92	**99.375**	100	90	**95.875**	100
Robust − ICA $_t$	95	**98.0625**	100	90	**99.25**	100	92	**96.875**	100
BSS − CCA $_t$	89	**96.9375**	100	78	**98.25**	100	88	**95.625**	100
SOBI $_t$	92	**98.5625**	100	84	**98.75**	100	96	**98.375**	100
JADE $_t$	95	**97.625**	100	90	**99.25**	100	90	**96**	100
	Combined Extended and Fused Multichannel Time-Frequency Features								
SAR ≈ 0 dB	75	**97.8125**	100	66	**97.875**	100	84	**97.75**	100
Fast − ICA $_t$	94	**97.8125**	100	98	**99.875**	100	88	**95.75**	100
Robust − ICA $_t$	94	**98.25**	100	94	**99.625**	100	90	**96.875**	100
BSS − CCA $_t$	92	**98.0625**	100	100	**100**	100	84	**96.125**	100
SOBI $_t$	95	**98.6875**	100	94	**99.625**	100	90	**97.75**	100
JADE $_t$	95	**98.0625**	100	98	**99.875**	100	90	**96.25**	100

Table A.18: Time domain filtered EEG abnormality detection average performances when using the fused and the combined extended and fused time-frequency multichannel features. SAR level is equal to -7.2206 dB. *min* is the minimum value, *max* is the maximum value, μ is the mean value

	Fused Multichannel Time-Frequency Features								
	Balanced Accuracy (%)			**Sensitivity (%)**			**Specificity (%)**		
	min	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*
SAR ≈ -7.22 dB	96	**98.1875**	100	98	**99.75**	100	92	**96.625**	100
Fast − ICA $_t$	94	**96.75**	100	90	**99**	100	88	**94.5**	100
Robust − ICA $_t$	94	**96.875**	100	88	**98.875**	100	88	**94.875**	100
BSS − CCA $_t$	94	**96.75**	100	92	**99.25**	100	88	**94.25**	100
SOBI $_t$	93	**96.6875**	100	92	**99.375**	100	86	**94**	100
JADE $_t$	94	**96.625**	100	90	**99**	100	88	**94.25**	100
	Combined Extended and Fused Multichannel Time-Frequency Features								
SAR ≈ -7.22 dB	94	**98.125**	100	98	**99.875**	100	88	**96.375**	100
Fast − ICA $_t$	69	**95.25**	100	68	**97.375**	100	70	**93.125**	100
Robust − ICA $_t$	66	**94.9375**	100	64	**97.125**	100	68	**92.75**	100
BSS − CCA $_t$	83	**95.6875**	100	86	**98.875**	100	80	**92.5**	100
SOBI $_t$	85	**95.8125**	100	90	**99.25**	100	80	**92.375**	100
JADE $_t$	70	**95.3125**	100	68	**97.5**	100	72	**93.125**	100

Table A.19: Time domain filtered EEG abnormality detection average performances when using the fused and the combined extended and fused time-frequency multichannel features. SAR level is equal to -13.241 dB. *min* is the minimum value, *max* is the maximum value, μ is the mean value

	Fused Multichannel Time-Frequency Features								
	Balanced Accuracy (%)			**Sensitivity (%)**			**Specificity (%)**		
	min	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*
SAR ≈ -13.241 dB	73	**93.0625**	99	60	**93.5**	98	84	**92.625**	100
Fast − ICA $_t$	91	**95.8125**	99	82	**97.625**	100	86	**94**	100
Robust − ICA $_t$	87	**96.5**	100	86	**99.125**	100	88	**93.875**	100
BSS − CCA $_t$	90	**95.625**	99	90	**99.125**	100	88	**92.125**	98
SOBI $_t$	86	**94.75**	100	88	**97.375**	100	84	**92.125**	100
JADE $_t$	94	**96.3125**	99	88	**99**	100	88	**93.625**	100
	Combined Extended and Fused Multichannel Time-Frequency Features								
SAR ≈ -13.241 dB	90	**97.75**	100	100	**100**	100	80	**95.5**	100
Fast − ICA $_t$	77	**94.9375**	99	76	**96.75**	100	78	**93.125**	100
Robust − ICA $_t$	77	**95.75**	100	76	**98.375**	100	78	**93.125**	100
BSS − CCA $_t$	69	**94.1875**	99	64	**97**	100	74	**91.375**	98
SOBI $_t$	73	**95.1875**	100	66	**97.625**	100	80	**92.75**	100
JADE $_t$	79	**95.8125**	100	80	**98.625**	100	78	**93**	100

Table A.20: Time domain filtered EEG abnormality detection average performances when using the fused and the combined extended and fused time-frequency multichannel features. SAR level is equal to -16.763 dB. *min* is the minimum value, *max* is the maximum value, μ is the mean value

	Fused Multichannel Time-Frequency Features								
	Balanced Accuracy (%)			**Sensitivity (%)**			**Specificity (%)**		
	min	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*
SAR ≈ -16.763 dB	60	**82.8125**	90	40	**82**	90	66	**83.625**	90
Fast − ICA $_t$	87	**93.125**	97	78	**95.375**	100	86	**90.875**	96
Robust − ICA $_t$	83	**94.625**	100	78	**96.875**	100	88	**92.375**	100
BSS − CCA $_t$	90	**94**	98	90	**97.25**	100	88	**90.75**	96
SOBI $_t$	83	**90.1875**	95	88	**95.5**	98	78	**84.875**	94
JADE $_t$	88	**93**	96	78	**94.5**	100	86	**91.5**	98
	Combined Extended and Fused Multichannel Time-Frequency Features								
SAR ≈ -16.763 dB	88	**97.625**	100	100	**100**	100	76	**95.25**	100
Fast − ICA $_t$	81	**93.75**	97	80	**96.875**	100	82	**90.625**	96
Robust − ICA $_t$	83	**95.0625**	100	82	**97.875**	100	84	**92.25**	100
BSS − CCA $_t$	80	**94.625**	98	70	**96.875**	100	86	**92.375**	98
SOBI $_t$	82	**92.375**	97	74	**95.875**	98	76	**88.875**	96
JADE $_t$	85	**94.375**	98	86	**97.5**	100	84	**91.25**	96

Table A.21: Time domain filtered EEG abnormality detection average performances when using the fused and the combined extended and fused time-frequency multichannel features. SAR level is equal to -19.262 dB. *min* is the minimum value, *max* is the maximum value, μ is the mean value

	Fused Multichannel Time-Frequency Features								
	Balanced Accuracy (%)			**Sensitivity (%)**			**Specificity (%)**		
	min	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*
SAR ≈ -19.262 dB	49	**70.9375**	77	40	**71.25**	80	50	**70.625**	76
Fast − ICA $_t$	84	**89.625**	96	74	**90.375**	98	86	**88.875**	96
Robust − ICA $_t$	78	**86.875**	91	72	**89**	96	80	**84.75**	92
BSS − CCA $_t$	88	**91.0625**	93	90	**94.25**	98	86	**87.875**	92
SOBI $_t$	81	**84.8125**	86	86	**91**	94	76	**78.625**	82
JADE $_t$	81	**87.875**	91	72	**88.25**	96	82	**87.5**	94
	Combined Extended and Fused Multichannel Time-Frequency Features								
SAR ≈ -19.262 dB	88	**97.4375**	100	94	**99.25**	100	76	**95.625**	100
Fast − ICA $_t$	87	**94.25**	98	86	**95.625**	100	88	**92.875**	98
Robust − ICA $_t$	86	**93.9375**	98	84	**96.875**	100	82	**91**	98
BSS − CCA $_t$	82	**93.6875**	97	72	**96**	100	86	**91.375**	98
SOBI $_t$	81	**88.3125**	93	72	**92.5**	96	72	**84.125**	92
JADE $_t$	87	**93**	96	88	**95.75**	100	86	**90.25**	94

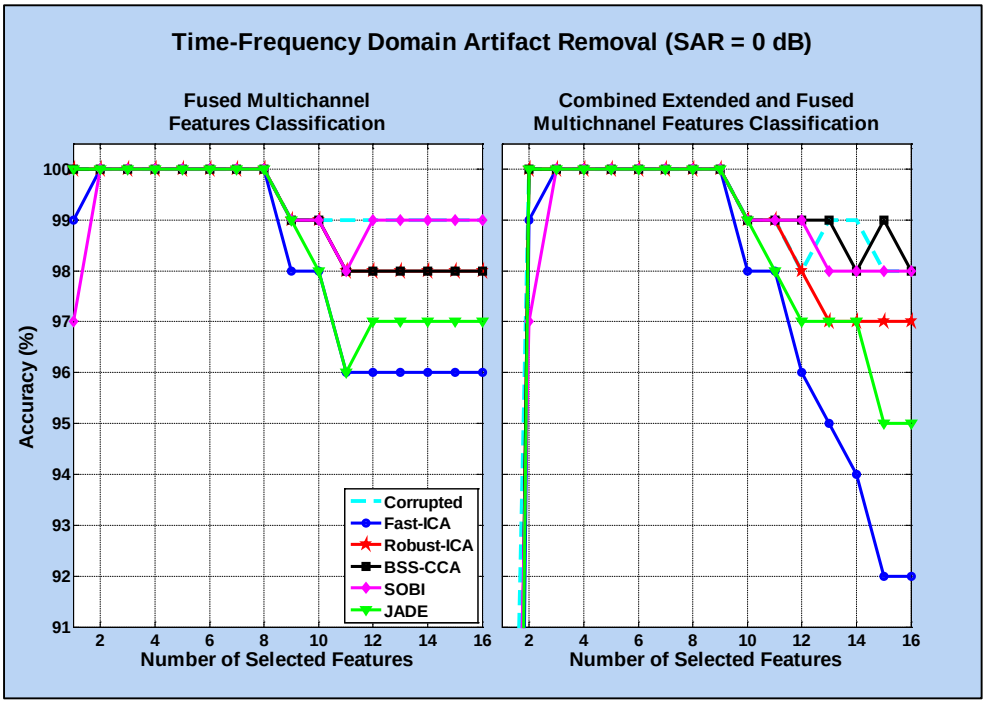

Figure A.62: Time-frequency domain filtered EEG abnormality detection accuracy profiles when using the fused and the combined extended and fused time-frequency multichannel features. SAR level is equal to 0 dB

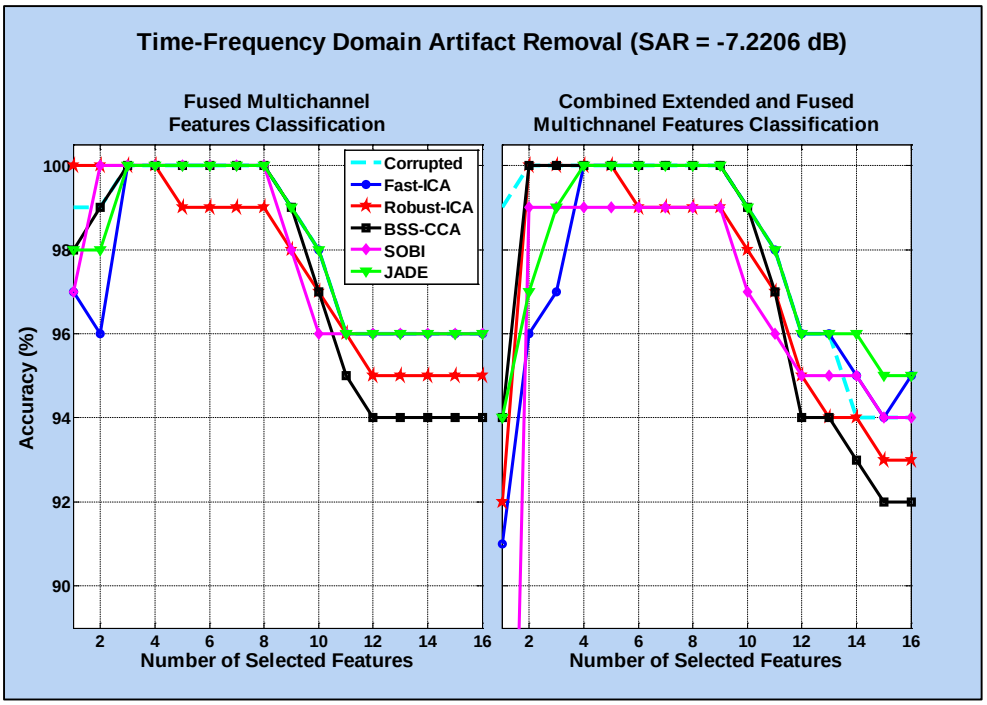

Figure A.63: Time-frequency domain filtered EEG abnormality detection accuracy profiles when using the fused and the combined extended and fused time-frequency multichannel features. SAR level is equal to -7.2206 dB

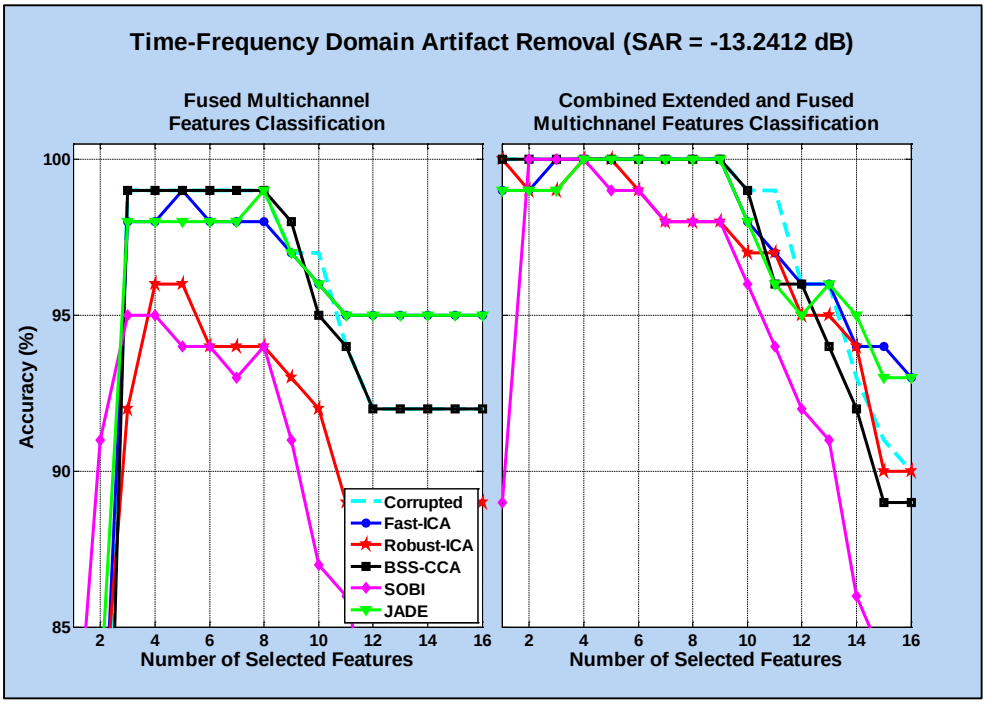

Figure A.64: Time-frequency domain filtered EEG abnormality detection accuracy profiles when using the fused and the combined extended and fused time-frequency multichannel features. SAR level is equal to -13.2412 dB

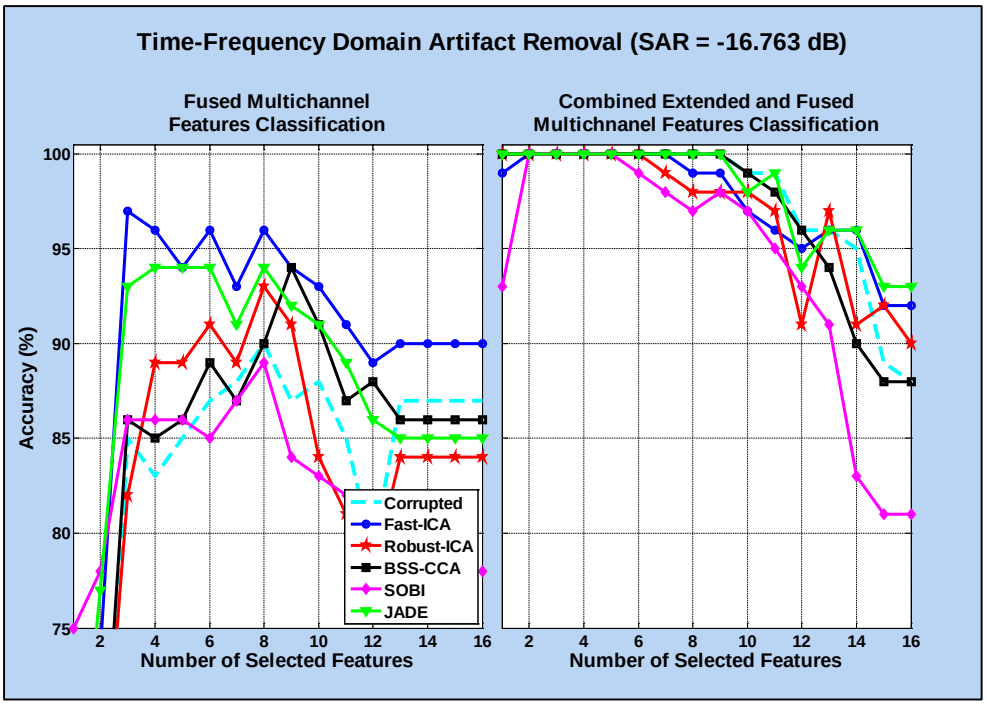

Figure A.65: Time-frequency domain filtered EEG abnormality detection accuracy profiles when using the fused and the combined extended and fused time-frequency multichannel features. SAR level is equal to -16.763 dB

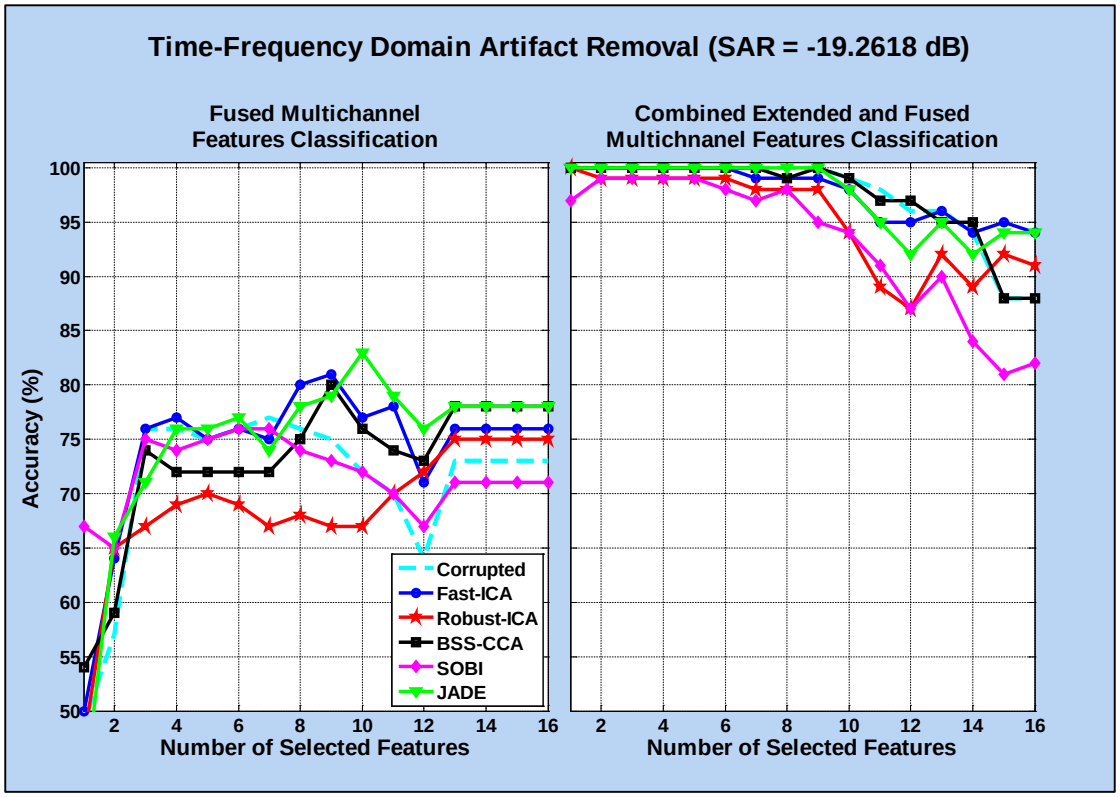

Figure A.66: Time-frequency domain filtered EEG abnormality detection accuracy profiles when using the fused and the combined extended and fused time-frequency multichannel features. SAR level is equal to -19.262 dB

Table A.22: Time-frequency domain filtered EEG abnormality detection average performances when using the fused and the combined extended and fused time-frequency multichannel features. SAR level is equal to 0 dB. *min* is the minimum value, *max* is the maximum value, μ is the mean value

	Fused Multichannel Time-Frequency Features								
	Balanced Accuracy (%)			**Sensitivity (%)**			**Specificity (%)**		
	min	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*
SAR ≈ 0 dB	99	**99.5**	100	100	**100**	100	98	**99**	100
$\mathbf{Fast-ICA}_{tf}$	96	**98.1875**	100	98	**99.875**	100	92	**96.5**	100
$\mathbf{Robust-ICA}_{tf}$	98	**99.125**	100	100	**100**	100	96	**98.25**	100
$\mathbf{BSS-CCA}_{tf}$	98	**99.125**	100	100	**100**	100	96	**98.25**	100
$\mathbf{SOBI}_{tf}$	97	**99.25**	100	94	**99.625**	100	96	**98.875**	100
$\mathbf{JADE}_{tf}$	96	**98.625**	100	100	**100**	100	92	**97.25**	100
	Combined Extended and Fused Multichannel Time-Frequency Features								
SAR ≈ 0 dB	75	**97.8125**	100	66	**97.875**	100	84	**97.75**	100
$\mathbf{Fast-ICA}_{tf}$	50	**94.625**	100	80	**98.625**	100	20	**90.625**	100
$\mathbf{Robust-ICA}_{tf}$	59	**96.4375**	100	70	**98.125**	100	48	**94.75**	100
$\mathbf{BSS-CCA}_{tf}$	61	**97**	100	54	**97.125**	100	68	**96.875**	100
$\mathbf{SOBI}_{tf}$	75	**97.5625**	100	84	**98.625**	100	66	**96.5**	100
$\mathbf{JADE}_{tf}$	55	**95.8125**	100	66	**97.875**	100	44	**93.75**	100

Table A.23: Time-frequency domain filtered EEG abnormality detection average performances when using the fused and the combined extended and fused time-frequency multichannel features. SAR level is equal to -7.2206 dB. *min* is the minimum value, *max* is the maximum value, μ is the mean value

	Fused Multichannel Time-Frequency Features								
	Balanced Accuracy (%)			**Sensitivity (%)**			**Specificity (%)**		
	min	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*
SAR ≈ -7.22 dB	96	**98.1875**	100	98	**99.75**	100	92	**96.625**	100
$\mathbf{Fast-ICA}_{tf}$	96	**97.875**	100	94	**99.25**	100	92	**96.5**	100
$\mathbf{Robust-ICA}_{tf}$	95	**97.625**	100	100	**100**	100	90	**95.25**	100
$\mathbf{BSS-CCA}_{tf}$	94	**97.375**	100	96	**99.625**	100	88	**95.125**	100
$\mathbf{SOBI}_{tf}$	96	**97.9375**	100	94	**99.625**	100	92	**96.25**	100
$\mathbf{JADE}_{tf}$	96	**98.0625**	100	96	**99.5**	100	92	**96.625**	100
	Combined Extended and Fused Multichannel Time-Frequency Features								
SAR ≈ -7.22 dB	94	**98.125**	100	98	**99.875**	100	88	**96.375**	100
$\mathbf{Fast-ICA}_{tf}$	91	**97.3125**	100	92	**98.875**	100	88	**95.75**	100
$\mathbf{Robust-ICA}_{tf}$	92	**97**	100	92	**99.5**	100	86	**94.5**	100
$\mathbf{BSS-CCA}_{tf}$	92	**97.1875**	100	96	**99.75**	100	84	**94.625**	100
$\mathbf{SOBI}_{tf}$	71	**95.5625**	99	70	**98.125**	100	72	**93**	98
$\mathbf{JADE}_{tf}$	94	**97.8125**	100	96	**99.5**	100	90	**96.125**	100

Table A.24: Time-frequency domain filtered EEG abnormality detection average performances when using the fused and the combined extended and fused time-frequency multichannel features. SAR level is equal to -13.2412 dB. *min* is the minimum value, *max* is the maximum value, μ is the mean value

	Fused Multichannel Time-Frequency Features								
	Balanced Accuracy (%)			**Sensitivity (%)**			**Specificity (%)**		
	min	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*
SAR ≈ -13.241 dB	73	**93.0625**	99	60	**93.5**	98	84	**92.625**	100
$\mathbf{Fast-ICA}_{tf}$	79	**94.375**	99	72	**94**	98	86	**94.75**	100
$\mathbf{Robust-ICA}_{tf}$	81	**90.5**	96	74	**85.625**	94	86	**95.375**	98
$\mathbf{BSS-CCA}_{tf}$	70	**92.8125**	99	62	**93.625**	98	78	**92**	100
$\mathbf{SOBI}_{tf}$	80	**88.375**	95	84	**93.125**	94	70	**83.625**	96
$\mathbf{JADE}_{tf}$	83	**94.9375**	99	72	**94.25**	98	92	**95.625**	100
	Combined Extended and Fused Multichannel Time-Frequency Features								
SAR ≈ -13.241 dB	90	**97.75**	100	100	**100**	100	80	**95.5**	100
$\mathbf{Fast-ICA}_{tf}$	93	**97.875**	100	100	**100**	100	86	**95.75**	100
$\mathbf{Robust-ICA}_{tf}$	90	**96.8125**	100	98	**98.875**	100	82	**94.75**	100
$\mathbf{BSS-CCA}_{tf}$	89	**97.1875**	100	100	**100**	100	78	**94.375**	100
$\mathbf{SOBI}_{tf}$	84	**94.25**	100	84	**98.375**	100	68	**90.125**	100
$\mathbf{JADE}_{tf}$	93	**97.6875**	100	98	**99.75**	100	86	**95.625**	100

Table A.25: Time-frequency domain filtered EEG abnormality detection average performances when using the fused and the combined extended and fused time-frequency multichannel features. SAR level is equal to -16.763 dB. *min* is the minimum value, *max* is the maximum value, μ is the mean value

	Fused Multichannel Time-Frequency Features								
	Balanced Accuracy (%)			**Sensitivity (%)**			**Specificity (%)**		
	min	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*
SAR ≈ -16.763 dB	60	**82.8125**	90	40	**82**	90	66	**83.625**	90
$\mathbf{Fast-ICA}_{tf}$	59	**89.5**	97	52	**88.5**	96	66	**90.5**	98
$\mathbf{Robust-ICA}_{tf}$	56	**82.6875**	93	52	**77**	90	58	**88.375**	96
$\mathbf{BSS-CCA}_{tf}$	64	**85**	94	48	**83.375**	96	74	**86.625**	92
$\mathbf{SOBI}_{tf}$	75	**82.125**	89	82	**91.875**	96	58	**72.375**	88
$\mathbf{JADE}_{tf}$	64	**87.4375**	94	54	**82.375**	94	74	**92.5**	96
	Combined Extended and Fused Multichannel Time-Frequency Features								
SAR ≈ -16.763 dB	88	**97.625**	100	100	**100**	100	76	**95.25**	100
$\mathbf{Fast-ICA}_{tf}$	92	**97.5625**	100	98	**99.375**	100	84	**95.75**	100
$\mathbf{Robust-ICA}_{tf}$	90	**96.9375**	100	82	**96.625**	100	80	**97.25**	100
$\mathbf{BSS-CCA}_{tf}$	88	**97.0625**	100	100	**100**	100	76	**94.125**	100
$\mathbf{SOBI}_{tf}$	81	**94.125**	100	90	**99.25**	100	62	**89**	100
$\mathbf{JADE}_{tf}$	93	**98.0625**	100	96	**99.75**	100	86	**96.375**	100

Table A.26: Time-frequency domain filtered EEG abnormality detection average performances when using the fused and the combined extended and fused time-frequency multichannel features. SAR level is equal to -19.262 dB. *min* is the minimum value, *max* is the maximum value, μ is the mean value

	Fused Multichannel Time-Frequency Features								
	Balanced Accuracy (%)			**Sensitivity (%)**			**Specificity (%)**		
	min	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*	*min*	$\boldsymbol{\mu}$	*max*
SAR ≈ -19.262 dB	49	**70.9375**	77	40	**71.25**	80	50	**70.625**	76
$\mathbf{Fast-ICA}_{tf}$	50	**74**	81	52	**74**	84	48	**74**	82
$\mathbf{Robust-ICA}_{tf}$	47	**68.625**	75	60	**67.125**	78	34	**70.125**	90
$\mathbf{BSS-CCA}_{tf}$	54	**72.8125**	80	24	**70.125**	80	52	**75.5**	86
$\mathbf{SOBI}_{tf}$	65	**71.75**	76	70	**81**	86	50	**62.5**	70
$\mathbf{JADE}_{tf}$	42	**74.3125**	83	44	**73.25**	82	40	**75.375**	84
	Combined Extended and Fused Multichannel Time-Frequency Features								
SAR ≈ -19.262 dB	88	**97.4375**	100	94	**99.25**	100	76	**95.625**	100
$\mathbf{Fast-ICA}_{tf}$	94	**97.75**	100	92	**98.375**	100	88	**97.125**	100
$\mathbf{Robust-ICA}_{tf}$	87	**95.1875**	100	78	**91.25**	100	96	**99.125**	100
$\mathbf{BSS-CCA}_{tf}$	88	**97.375**	100	96	**99.25**	100	76	**95.5**	100
$\mathbf{SOBI}_{tf}$	81	**93.125**	99	92	**98.5**	100	62	**87.75**	100
$\mathbf{JADE}_{tf}$	92	**97.5**	100	86	**97.625**	100	88	**97.375**	100


תודה
Dankie
Gracias
Спасибо
شكراً
Merci
Takk
Köszönjük
Terima kasih
Grazie
Dziękujemy
Dėkojame
Ďakujeme
Vielen Dank
Paldies
Kiitos
Täname teid
谢谢
Thank You
Tak
感謝您
Obrigado
Teşekkür Ederiz
감사합니다
Σας ευχαριστούμε
ขอบคุณ
Bedankt
Děkujeme vám
ありがとうございます
Tack