UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

Electromyographic Artifact Removal from Electroencephalography Using Simulations and Spatially Constrained, Optimised, Wavelet Enhanced Independent Component Analysis

Gabriel Devin Eisenberg

Johannesburg, 2019

A dissertation submitted to the Faculty of Engineering and the Built Environment, University of Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science in Engineering.

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

Declaration

I declare that this dissertation is my own unaided work. It is being submitted to the degree of Master of Science in Engineering to the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination to any other University.

.....

(Signature of Candidate)

……. day of …… year …….

Abstract

Electroencephalography (EEG) is used to measure electrical activity produced in the brain through multiple electrodes placed on the scalp. EEG has clinical, research, and brain-computer interface (BCI) device applications. EEG is susceptible to artifacts and noise, resulting in a diminished signal to noise ratio (SNR). A particularly challenging type of artifact is produced by muscles of the face and scalp, called electromyographic (EMG) artifacts.

This study aimed to develop an EMG artifact removal algorithm for EEG. An algorithm called spatially constrained, wavelet enhanced independent component analysis (SCWEICA) was adapted to enable EMG artifact removal. It was then improved by optimising the wavelet denoising parameters (WDPs).

The new optimised SCWEICA (O-SCWEICA) algorithm and two EMG artifact removal algorithms were used to remove EMG artifacts from various datasets of simulated 31-channel EMG artifact contaminated EEG (EAC-EEG). The datasets varied in duration and artifact topographies. The performance of the three algorithms were then compared.

O-SCWEICA was shown to be robust to changes in EMG artifact topography and well suited for offline EMG artifact removal from longer duration EAC-EEG recordings. On average, O-SCWEICA took 0.56 s to process the longest EAC-EEG recordings, those of 200 s, while maintaining high mutual information (MI) of 2.40 on average across the electrodes (AV). The worst case (WC) MI was 1.76, where WC refers to the channel with the largest EMG artifact. Compared to U-SCWEICA, O-SCWEICA was 1.14% and 12.3% more accurate in terms of AV and WC MI, respectively.

Acknowledgements

I would like to thank the two most significant people who have assisted and guided me in the completion of this study: Mr Abdul-Khaaliq Mohamed and Prof Vered Aharonson. Their patience, guidance, and support in this study, my personal life, and employment has been invaluable and is greatly appreciated. With their influence, my problem-solving ability and critical thinking have improved tremendously.

I would also like to thank Mr Abdul-Khaaliq Mohamed for obtaining funding for me through his grant with the National Research Foundation (NRF). In turn, I would like to acknowledge the financial assistance of the NRF.

The completion of this study would not have been possible without the help of academics who have worked in this field. I would like to thank Dr Muhammed Akhtar for providing me with the SCFastICA algorithm. I thank Dr Hugh Nolan and Dr Robert Whelan for advising me on EMG artifact simulation and for providing me with the algorithm to do so.

Many thanks to Jeannine Davids, the secretary in the School of Electrical and Information Engineering. She is kind, friendly, and has been incredibly supportive of me in the last few years.

The support of my family and close friends is greatly appreciated. Their understanding, support, and encouragement throughout this long process, even while completing anatomy, is appreciated more than I can possibly express. Thank you to my parents, Simon and Melanie, and my sisters, Micaela and Ariella. Thank you to Kyle Fyvie, Denette Holloway, and Samantha Dos Santos for being there for me throughout the highs and lows of this work. In particular, I would like to thank Kyle for allowing me to bounce ideas off him and assisting me in finding solutions to the numerous problems I encountered, from ideation to solution. I would like to thank Isabel Tollman for selflessly offering to edit my dissertation. Her keen eye for editing has transformed this document.

Table of Contents

Declaration i

Abstract ii

Acknowledgements iii

Table of Contents iv

List of Figures ix

List of Tables xii

Glossary of Terms and Abbreviations xiv

Chapter 1 – Introduction 1

1.1 Problem Outline 1

1.2 Dissertation Structure 5

Chapter 2 – Background 6

2.1 Electroencephalography (EEG) 6

2.1.1 EEG Characteristics 6

2.1.2 Fundamentals of Scalp EEG Modelling 7

2.1.3 Artifacts in EEG 8

2.2 Electromyographic (EMG) Artifacts in EEG 9

2.3 Chapter Summary 10

Chapter 3 – Literature Review 11

3.1 Independent Component Analysis (ICA) 11

3.1.1 Cocktail Party Problem 12

3.1.2 Overview of ICA Model and Algorithm 12

3.1.3 Standard ICA (STD-ICA) for Artifact Removal 16
3.1.4 Spatially Constrained ICA (SCICA)..... 17
3.1.5 Summary..... 20
3.2 Wavelet Denoising..... 21
3.2.1 Wavelet Transform (WT) 21
3.2.2 Discrete Wavelet Transform (DWT) 23
3.2.3 Stationary Wavelet Transform (SWT)..... 24
3.2.4 Wavelet Denoising Methodology and Wavelet Denoising Parameters (WDPs)..... 25
3.2.5 Summary..... 28
3.3 Removal of EMG Artifacts from EEG..... 29
3.3.1 Non-BSS Based EMG Artifact Removal..... 29
3.3.2 BSS Based EMG Artifact Removal..... 29
3.3.3 Wavelet Enhanced ICA for EMG Artifact Removal 31
3.3.4 Summary..... 34
3.4 WDP Selection and Optimisation 34
3.4.1 WDPs for Artifact Component Denoising..... 34
3.4.2 WDP Optimisation for Temporal Signals..... 35
3.4.3 Summary..... 36
3.5 EEG and EMG Artifact Simulations..... 37
3.5.1 EEG Simulation 37
3.5.2 EMG Artifact Simulation..... 40
3.5.3 Summary..... 42
3.6 Performance Measures 42
3.6.1 Normalised Mean Square Error (NMSE) 42

3.6.2 Mutual Information (MI) 43
3.6.3 Absolute Maximum Cross-Correlation (AMCC) 43
3.6.4 Computational Efficiency 44
3.6.5 Relative Difference 44
3.7 Chapter Summary 45
Chapter 4 – Problem Specification **46**
4.1 Research Question (RQ) 46
4.2 Research Areas (RAs) 47
4.2.1 Research Area 1 (RA1) 47
4.2.2 Research Area 2 (RA2) 47
4.2.3 Research Area 3 (RA3) 47
4.3 Hypotheses 48
4.4 Expected Outcomes 48
Chapter 5 – Methodology **49**
5.1 Generation of Simulated Data 50
5.1.1 Ground Truth EEG Simulation 51
5.1.2 EMG Artifact Simulation and Their Addition to 'Clean' EEG 54
5.1.3 Validation of Simulations 64
5.1.4 Simulated Datasets to Address the RAs 65
5.2 Implementation of SCWEICA for U-SCWEICA and O-SCWEICA 70
5.2.1 Step-by-Step Overview of the SCWEICA Algorithm 72
5.2.2 Spatially Constrained ICA (SCICA) 72
5.2.3 Wavelet Denoising of the EMG Artifact Component 74

5.2.4 EEG Reconstruction 75

5.3 Implementation of RA1 – Selection of Decomposition Level for U-SCWEICA 75

5.3.1 Accuracy Measures 76

5.3.2 Methodology 76

5.4 Implementation of RA2 – Optimisation of WDPs for O-SCWEICA 77

5.4.1 Explored WDPs for Optimisation 78

5.4.2 WDP Optimisation Methodology 79

5.5 Implementation of RA3 – Experimental Setup for Performance Comparison of EMG Artifact Removal Algorithms 80

5.5.1 Adaptation of Standard ICA Denoising Algorithm for Automated Denoising and Testing 80

5.5.2 Testing of EMG Artifact Removal Algorithms 80

5.5.3 Hardware and Software 83

5.6 Chapter Summary 83

Chapter 6 – Results and Discussion **85**

6.1 RA1 – Decomposition Level Selection for U-SCWEICA 85

6.2 RA2 – WDP Optimisation for O-SCWEICA 86

6.2.1 Determination of Optimal WDPs for O-SCWEICA 90

6.3 RA3 – Performance Comparison of EMG Artifact Removal Algorithms 91

6.3.1 Performance Comparison Results 91

6.3.2 Performance Comparison Discussion 98

6.4 Answering the Research Question 104

6.5 Chapter Summary 105

Chapter 7 – Conclusion **107**

7.1 Summary 107

7.2 Main Findings 109

7.3 Recommendations for Future Work 109

References **112**

List of Figures

Figure 2.1: An example of EOG and EMG artifacts in a selection of EEG channels (adapted from [2]). 9

Figure 2.2: Muscles of the face and scalp approximately overlaid with an EEG electrode layout. **(a)** The muscles of the face and scalp with electrodes approximately overlaid as white circles (adapted from [30]). **(b)** An electrode layout over a subject's head (adapted from [31])..... 10

Figure 3.1: Examples of mother wavelets. (a) Haar (haar), (b) Daubechies order 4 (db4), (c) Coiflet order 1 (coif1) and (d) Symlet order 9 (sym9)..... 22

Figure 3.2: Recursive filtering in the application of the DWT for three decomposition levels (adapted from [41]) 24

Figure 3.3: Bandwidths of three decomposition levels with the DWT (adapted from [42])..... 24

Figure 3.4: Cascading of artifact removal techniques. Adapted from [9]..... 31

Figure 3.5: Abstracted wavelet denoising of artifact independent component. **(a)** Artifact component with neural 'noise'. **(b)** 'Denoised' artifact component. **(c)** Extracted neural signal which appears to be 'noise' within the component. 32

Figure 3.6: High level SCWEICA block diagram (adapted from [14])..... 33

Figure 4.1: Development of RAs 47

Figure 5.1: High-level block diagram of the methodology used in this study..... 49

Figure 5.2: A sample of simulated, 'clean' EEG 52

Figure 5.3: A 64 EEG electrode layout according to the international 10-20 system, with the 31 simulated electrodes filled in grey (adapted from [60])..... 52

Figure 5.4: The spectra of **(a)** typical, observed EEG used to simulate 'clean' EEG signals and **(b)** simulated EEG from channel FC4..... 53

Figure 5.5: Electrode layout and EMG weight mapping process 55

Figure 5.6: Mapping of a **(a)** 128-electrode layout (the six additional EOG and mastoid electrodes used in [54] are not shown) to a **(b)** 31-electrode layout (adapted from [61] and [60], respectively) 55

Figure 5.7: Mapping of EMG weights. **(a)** EMG weights for the 134-electrode layout. **(b)** EMG weights for the 31-electrode layout. **(c)** Spatially distributed EMG weights for the 31-electrode layout. 60

Figure 5.8: Primary electrodes (filled in grey) affected by: **(a)** right-side temporalis muscle contraction, **(b)** left-side temporalis muscle contraction and **(c)** both left-side and right-side temporalis muscle contraction. Adapted from [60]. 62

Figure 5.9: Spatially distributed EMG artifacts 63

Figure 5.10: Simulated EAC-EEG..... 63

Figure 5.11: Comparison of the spectrum of channel FT8 **(a)** before and **(b)** after the addition of EMG artifacts. **(c)** Spectrum of isolated EMG artifact in channel FT8. 64

Figure 5.12: An example of the windows (marked by dashed lines) used for the Towards RT Datasets. The signal is the EAC-EEG from channel FT8, where the EMG artifact occurred from approximately 0.4 to 2.4 s. The windowed sections across all channels in Windows 1, 2, and 3 were included in the Towards RT Datasets 1, 2, and 3, respectively. 69

Figure 5.13: SCWEICA block diagram (adapted from Akhtar *et al*. [14])..... 70

Figure 5.14: Signals at each stage of SCWEICA algorithm. **(a)** EAC-EEG. **(b)** Components from SCICA. Component 1 corresponds to the EMG artifact and component 2 corresponds to neural signals. **(c)** Wavelet denoising of component 1

and zeroing of component 2. **(d)** Spatially distributed EMG artifact. **(e)** Denoised EEG after subtraction of spatially distributed EMG artifact from EAC-EEG. **(f)** Ground truth EEG which appears to be indistinguishable from the denoised EEG. 71

Figure 5.15: Block diagram of WDP selection methodology 77

Figure 5.16: EMG artifact removal and performance quantification..... 83

Figure 6.1: Average AMCC results for the testing of O-SCWEICA for the **(a)** AV and **(b)** WC cases. Higher AMCC implies better denoising. 93

Figure 6.2: Average MI results for the testing of O-SCWEICA for the **(a)** AV and **(b)** WC cases. Higher MI implies better denoising..... 94

Figure 6.3: Average NMSE results for the testing of O-SCWEICA for the **(a)** AV and **(b)** WC cases. Lower NMSE implies better denoising. 95

Figure 6.4: Average denoising duration results for **(a)** STD-ICA and **(b)** U-SCWEICA and O-SCWEICA. The results are shown separately owing to the vast differences in magnitude – they would not be adequately displayed in the same figure. 96

Figure 6.5: O-SCWEICA denoising duration over multiple iterations with MATLAB optimisation enabled 103

List of Tables

Table 5.1: Distances and distance errors between the target electrodes in the 31-electrode layout and the source electrodes in the 134-electrode layout 57

Table 5.2: Datasets for WDP selection and performance comparison..... 66

Table 5.3: Signals in the WDP dataset and datasets for performance comparison of EMG artifact removal algorithms 66

Table 5.4: Selection of discrete mother wavelets offered by MATLAB, used in the optimisation problem. 79

Table 5.5: Summary of tests..... 81

Table 6.1: Average NMI, AMCC, shift λ, and NMSE for wavelet decomposition levels one to ten for the combination of the Haar wavelet, Global threshold determination method and soft thresholding..... 86

Table 6.2: Best 20 WDP combinations in terms of average NMI 87

Table 6.3: Best 20 WDP combinations in terms of average AMCC 88

Table 6.4: Best 20 WDP combinations in terms of average NMSE 89

Table 6.5: A comparison of the average hard and soft thresholding in denoising EAC-EEG in terms of the AV and WC MI, AMCC and NMSE..... 90

Table 6.6: Absolute and relative difference between AV and WC results for hard and soft thresholding with reference to hard thresholding..... 91

Table 6.7: Average relative differences in accuracy between O-SCWEICA and U-SCWEICA. Positive AMCC and MI results imply better O-SCWEICA accuracy. Negative NMSE results imply better O-SCWEICA accuracy..... 97

Table 6.8: Average absolute and relative differences in efficiency between O-SCWEICA and U-SCWEICA. Negative values imply better O-SCWEICA efficiency. 97

Table 6.9: Comparison between literature and the AV denoising performance of O-SCWEICA when denoising the Akhtar Comparison Dataset 104

Glossary of Terms and Abbreviations

Abbreviation/Term	Description
A_c	Constrained column/s of mixing matrix **A** corresponding to the artifact topography.
AMCC	Absolute, maximum cross-correlation
AMICA	adaptive mixture of independent analysers
AV	Average for all channels/electrodes
BCI	Brain-computer interface
BSS	Blind source separation
CCA	Canonical correlation analysis
DWT	Discrete wavelet transform
EAC-EEG	Electromyographic artifact contaminated electroencephalogram
ECG	Electrocardiogram/ Electrocardiographic
EEG	Electroencephalography/ Electroencephalographic
EMG	Electromyography/ Electromyographic
EMG weights	Amplitudes of a typical muscle contraction signals at each EEG electrode
EOG	Electrooculogram/ Electrooculographic
Epoch	EEG recordings are typically divided into a number of equisized samples called epochs. Epochs are used in repetitive EEG experiments.
ERPs	Event related potentials
Frame	The number of points in an epoch
FT7	Frontotemporal electrode 7. The electrode most affected by EMG artifacts produced by the left temporalis muscle.
FT8	Frontotemporal electrode 8. The electrode most affected by EMG artifacts produced by the right temporalis muscle.
Ground truth	A known reference for accurate, comparative purposes

Abbreviation/Term	Description
ICA	Independent component analysis
InfoMax	Information maximisation
MAD	Mean absolute deviation
MI	Mutual information
Muscle map	The distribution of EMG weights of a muscle contraction across EEG electrodes
NMI	Normalised mutual information
NMSE	Normalised mean square error
Offline	Refers to recorded EEG signals
O-SCWEICA	Spatially constrained, optimised, wavelet enhanced independent component analysis
PCA	Principal component analysis
RA/s	Research area/s
RA1	Research area 1
RA2	Research area 2
RA3	Research area 3
RQ	Research question
RT	Real-time
SCFastICA	Spatially constrained FastICA
SCICA	Spatially constrained independent component analysis
SCWEICA	Spatially constrained, wavelet enhanced independent component analysis
SNR	Signal to noise ratio
SOBI	Second order blind identification
Spatially distributed artifact	The result of multiplication of a mixing matrix and an artifact component. The artifact is distributed across EEG electrodes by the weightings in the mixing matrix.
STD-ICA	Standard ICA denoising
SWT	Stationary wavelet transform

Abbreviation/Term	Description
U-SCWEICA	Spatially constrained, unoptimised, wavelet enhanced independent component analysis
WC	Worst-case measure for electrodes affected by the largest source of EMG artifacts
WDP/s	Wavelet denoising parameter/s
WT	Wavelet transform

Chapter 1 – Introduction

This chapter introduces the problem that serves as a foundation for this study. Section 1.1 details the problem, previous solutions and hypothesises a new solution to the problem. The structure of the dissertation is presented in Section 1.2.

1.1 Problem Outline

Electroencephalography (EEG) measures the electrical activity of the brain through multiple electrodes placed on the scalp [1]. EEG is a non-invasive, portable, and cost effective tool for brain research, patient monitoring, and diagnosis of neurological disorders [2]. EEG can also be used to harness electrical potentials from the brain for applications such as brain-computer interface (BCI) controlled devices [2]. BCI devices, such as EEG controlled prosthetic hands, can greatly improve the quality of life of amputees [3].

The measurements of EEG electrodes are susceptible to artifacts and noise as the brain is not the only detectable source of electrical activity [3]. These sources of artifacts and noise are either physiological (such as muscle interference, eye blinks, or pulse artifacts via arteries in the brain) or non-physiological (such as line noise, or movement of electrodes or wires) [4]. These phenomena cause a low signal to noise ratio (SNR) in EEG recordings [3], [5]. The low SNR can cause unintentional activity of a BCI device as artifacts can be misinterpreted as a neurological signal [2], [6]. Furthermore, if the amplitude of the artifact is large enough, the desired neurological signals can be further obscured within regions of artifacts [6]–[8].

Muscle, or electromyographic (EMG), artifacts are particularly challenging to remove as they overlap with EEG in both amplitude and frequency [1], [7], [8]. Simple signal processing methods such as linear filtering are therefore not viable to separate the EEG signal from EMG artifacts [9]. This is further complicated as

EMG artifacts vary per muscle and per person [7]. EMG artifacts are different for males and females, vary between individuals, and change according to the level and direction of contraction of the muscle [7]. As such, no reference signal is known *a-priori* [9]. Reference signals are required for common artifact removal techniques, including regression and adaptive filtering [9].

While the EMG artifacts can be measured alongside or within the EEG recording to obtain reference signals representing EMG artifacts, there are many muscles that would need to be measured with electrodes. This limits the number of electrodes that can be used for EEG recording. Furthermore, within the EMG artifacts, EEG appears as noise [9] and as such, simple subtraction of EMG artifacts from EEG would lead to a loss of EEG signals. Extensive research has been conducted on the removal of EMG artifacts from EEG signals and has shown that no single artifact removal technique can effectively suppress all EMG artifacts [4], [7], [9]. The need for a better EMG artifact removal algorithm, that yields an EEG signal with a higher SNR, is essential for research, clinical, and BCI applications.

A reliable EMG artifact removal algorithm necessitates an understanding of the underlying system that produces the signals of interest. Multichannel EEG is considered to be a combination of signals from the brain and artifact sources, like EMG artifacts [9]. The separation of these signals in such a system can be modelled as a blind source separation (BSS) problem [9], [10]. BSS problems require estimating the original sources of measured signals, with little or no knowledge of the sources and mixing process that produced the measured data [11]. BSS problems can be solved by several techniques such as independent component analysis (ICA), canonical correlation analysis (CCA), and principal component analysis (PCA) [9].

ICA has been implemented using various algorithms in several studies that showed relatively successful removal of EMG artifacts from EEG [6], [8], [12]. Additionally, ICA yielded greater success than PCA and has been recommended over CCA for EMG artifact removal from EEG [9]. ICA separates a combination

of measured signals into the underlying sources (called components) that generated the series of measurements [13]. These components can be processed to obtain a denoised representation of the original, measured signals [13]. In standard ICA (STD-ICA) artifact removal, artifact components are manually identified and zeroed before reconstruction takes place [8], [14].

STD-ICA is limited as it cannot be fully automated because it requires visual inspection to determine artifact components [2]. There are two additional problems with STD-ICA. Firstly, the STD-ICA implementation is computationally complex which hinders real-time (RT) implementation [14]. Secondly, as neural signals can 'leak into' artifact components, they will be lost upon removal of these components [2], [6], [8], [14]. The loss of these residual signals leads to inaccurate EEG reconstruction [2].

These issues – long processing time and leaking of neural signals into artifact components – were addressed by Akhtar *et al.* [14] with spatially constrained, wavelet enhanced ICA (SCWEICA). In SCWEICA, the prior, known spatial information about an artifact (such as specific EMG artifact topographies for the given EEG electrode layout) can be used to constrain the ICA algorithm [14], [15]. This greatly reduces processing time [14], [15]. The application of spatial constraints to the ICA algorithm is called spatially constrained ICA (SCICA) [14]–[16]. Wavelet denoising is then performed on the identified artifact component/s that were produced by SCICA [8], [14]. This enables the removal of neural signals that have leaked into the artifact component/s for accurate reconstruction of EEG signals [8], [14].

It is hypothesised that SCWEICA is a good candidate for EMG artifact removal based on the properties of the wavelet transform (WT) to separate signals and the spectral properties of both EEG and EMG artifacts [7], [9], [17]. SCWEICA has been used for successful removal of electrooculographic (EOG) artifacts, which occur owing to eye movement and blinking [14]. The algorithm has not yet been tested in the more challenging task of removing EMG artifacts [14]. Hence, this study explores how accurately SCWEICA can remove EMG artifacts from EEG.

By optimising the wavelet denoising parameters (WDPs) in SCWEICA, the accuracy of EMG artifact removal from EEG can be improved [18]. WDP optimisation has not been performed in previous studies that applied wavelet denoising to artifact components [14], [19], [20]. This study also explores how the optimisation of WDPs will affect the accuracy of SCWEICA in the application of EMG artifact removal.

It is hypothesised that the reduced computation time offered by SCWEICA [14] may have potential for RT applications. This is a vital trait, as most EEG applications and devices require EMG artifact removal in RT to provide an immediate response. Therefore, the algorithm duration is noted to ascertain its efficiency.

Simulations provide a convenient platform for an extensive artifact removal study, to assess the performance and RT viability of artifact removal algorithms [9]. Additionally, simulations allow for ground truth evaluation of such algorithms as the quality of the signal before and after artifact removal can be quantified [9]. Ground truth refers to using a known reference for accurate comparative purposes. However, simulations are a preliminary step to determining the final outcome of an algorithm [6], [9]. This must be done by applying it to real EEG data [6], [9]. This study employed simulated EEG data that was accurate enough to represent EEG-like signals that have been contaminated with EMG artifacts. The EMG artifact removal algorithm was developed and tested on this data. Testing on real EEG data was not attempted as it was out of the scope of this study.

This study optimised the WDPs in the SCWEICA algorithm (termed O-SCWEICA) to remove EMG artifacts from EEG and was compared to STD-ICA and unoptimised SCWEICA (termed U-SCWEICA) in terms of accuracy and efficiency, using simulations.

1.2 Dissertation Structure

The dissertation is structured as follows:

- Chapter 2 details background information on EEG signals and EMG artifacts.
- Chapter 3 presents the literature review of ICA and SCICA, wavelet denoising and methods for WDP optimisation, methods used in EMG artifact removal, the modelling and simulation of EEG and EMG artifacts, and performance measures to assess the accuracy and efficiency of artifact removal.
- Chapter 4 describes the research question (RQ), research areas (RAs), hypotheses, and expected outcomes of this study.
- Chapter 5 details the proposed methodology for the study, which includes the generation of simulations of 'clean' EEG and EMG artifact contaminated EEG (EAC-EEG), the implementation of both U-SCWEICA and O-SCWEICA for EMG artifact removal from EEG, and the testing procedures used in the study.
- Chapter 6 presents the results of tests performed in the study and discusses them. The RQ is subsequently answered.
- Chapter 7 concludes the dissertation with a summary of the work performed, main findings, and future recommendations.

Chapter 2 – Background

This chapter outlines background information for the literature review in Chapter 3, and for the remainder of the study. Background information is provided for EEG characteristics, modelling, and artifacts in Section 2.1, and for EMG artifacts in EEG in Section 2.2.

2.1 Electroencephalography (EEG)

EEG is used to measure superficial voltages generated in the cortex of the brain through electrodes placed on the scalp, often over hair [9], [21]. The voltages measured by EEG electrodes are produced by neurons in the brain and are a sum of the active neurons (sources), as well as other sources of voltage (artifacts) in the head [9], [22]. The summation of voltages is superposed as a function of the distance and orientation of sources, as well as the resistivity of tissues in the head [22]. This can be modelled using a number of approaches which are discussed later in this section. EEG characteristics and artifacts in EEG recordings are also discussed.

2.1.1 EEG Characteristics

EEG can be characterised in terms of temporal, spectral, and rhythmic features, as well as specific event related potentials (ERPs) called phasic peaks. These characteristics are common across individuals and are used to diagnose clinical conditions (such as epilepsy) and normal behaviour [9].

Temporal Characteristics of EEG

EEG voltages are generally in the range of 10 – 100 μV_{p-p} [23]. In adults, typical voltages range between 10 – 50 μV_{p-p} [23]. The frequency of these signals span between 0 – 100 Hz, which can be subdivided into various bandwidths [9].

Bandwidths of Rhythmic EEG Activity

There are several typical bandwidths of EEG activity that have been discovered through frequency and wavelet analysis [24]. These are delta (~ 0.5 – 4 Hz), theta (4 – 7 Hz), alpha (8 – 13 Hz), beta (14 – 30 Hz), and gamma (30 – 100 Hz) [9]. The definition of each bandwidth range varies in literature [2], [9], [24], [25]. This rhythmic EEG activity is often considered to be background EEG [9].

Phasic Peaks

Phasic peaks are a class of ERPs produced after an event and are typically bursts of transient signals [24]. Phasic peaks occur as a series of positive and negative peaks at fixed times after an event [24]. The amplitudes of these signals are small and attenuate over time [9], [24].

According to Yeung *et* al. [24], it can be assumed that phasic peaks are uncorrelated with background EEG signals. This allows for the signals to be simulated separately and combined for analytic purposes. In turn, background EEG is considered noise relative to phasic peaks [24]. However, this terminology does not imply that other EEG signals are random or meaningless, but simply refers to the uncorrelation between the two groups of signals [26]. This distinction is necessary to describe Yeung EEG Simulation in Section 3.5.1.

2.1.2 Fundamentals of Scalp EEG Modelling

There are several fundamental models used to approximate the generation of EEG signals that are measured on the scalp. These include the dipole model of neuronal populations, the forward model/problem and the spherical head shell model. Each topic is summarised in the following subsections. For a more detailed explanation of these topics, refer to [27], [28].

Dipole Model

Dipole models are used to represent the electrical activity of individual neurons or groups of neurons [27], [28]. The electrical activity is modelled as a current dipole (with an associated electric field) which can be measured at the surface of the

head using EEG electrodes [27]. Dipoles are directional and can be modelled to project within three dimensions at a particular location and voltage [27]. One or more dipoles may be used in a given model [27].

Forward Model/Problem

The forward problem is phrased as follows [28]: how can an electrical field be calculated from a known arrangement of neuronal sources? The forward model consists of using a source of electrical activity in the head (such as dipole models) and calculating the voltages measured at electrodes on the scalp [27]. This can be achieved using Maxwell's equations which are expressed in detail in [27], [28].

Spherical Head Shell Model

One of the methods to solve the forward problem is to use the spherical head shell model [27]. This model can contain multiple layers that represent parts of the head (such as the cerebrospinal fluid, skull and scalp) which have individual conductivities and thicknesses [27], [28]. These layers are modelled as concentric layers around an inner space – the brain [28]. The layers separate the dipoles in the centre of the head, modelled as a sphere, and the electrodes on the scalp [28]. The layers attenuate the EEG signals measured on the scalp compared to those on the surface of the brain [28].

2.1.3 Artifacts in EEG

The most common types of physiological artifacts in EEG are: EOG, EMG, and electrocardiogram (ECG) artifacts [4], [9]. ECG artifacts are produced by pulsations of blood vessels in the head [4], [9]. An example of EOG and EMG artifacts in EEG is shown in Figure 2.1. EOG artifacts are seen in channels F3, Fz, and F4, while EMG artifacts are primarily seen in channels P3 and P4 [2]. However, the EMG artifacts have affected most of the channels [2]. EOG artifacts appear as relatively periodic, peak-like signals that can be larger than 100 μV [4]. EMG artifacts typically appear as high frequency and high amplitude signals [7].

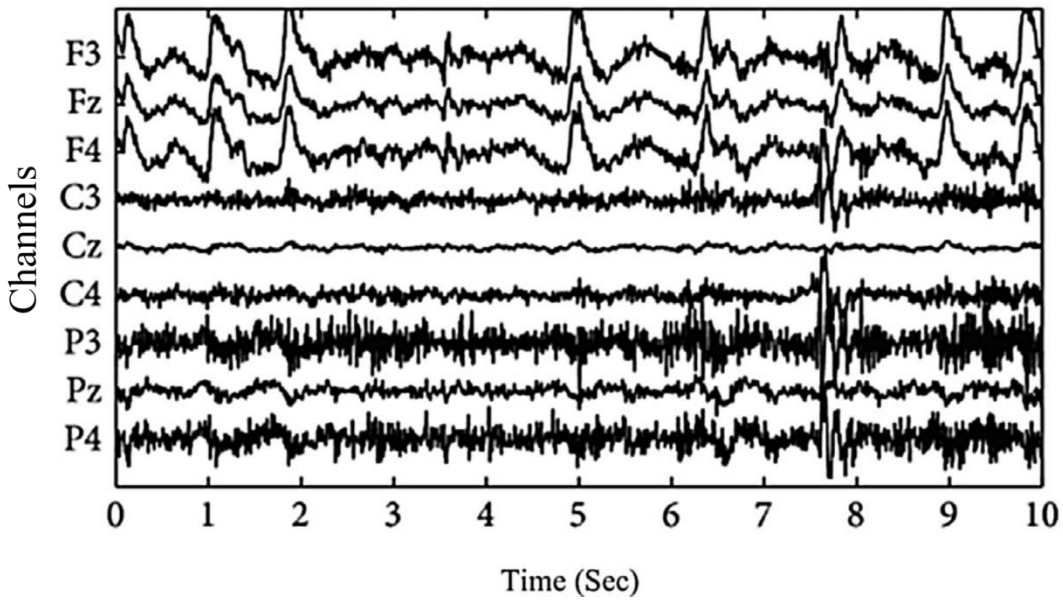
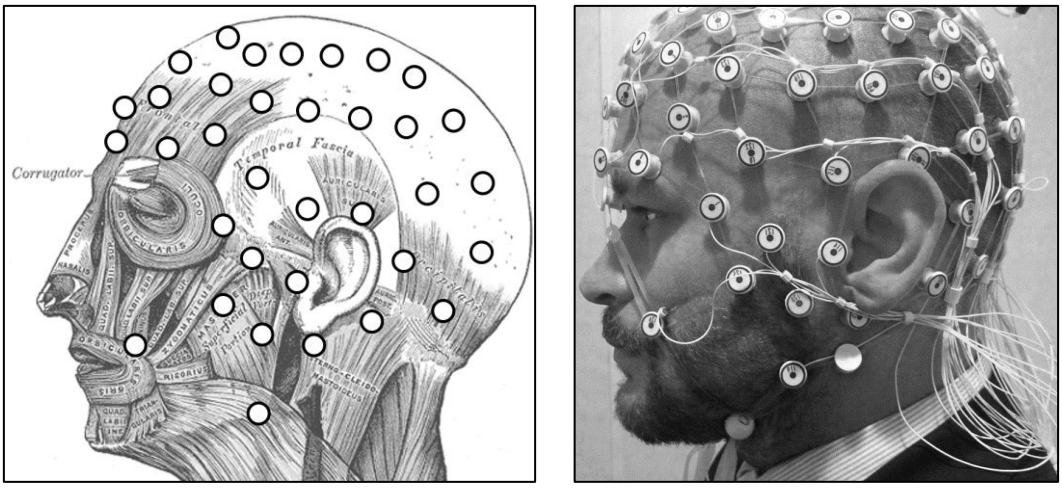


Figure 2.1: An example of EOG and EMG artifacts in a selection of EEG channels (adapted from [2]).

An extensive description of EEG artifacts, as well as techniques to remove them, is provided in [2], [4], [9].

2.2 Electromyographic (EMG) Artifacts in EEG

EMG artifacts in EEG recordings are generated by muscles in proximity to the electrodes placed on the scalp when measuring neural activity [29]. When these muscles contract, relatively large voltages are measured, which obscures the EEG signal measured by the EEG electrodes [7]. A comparison of the muscles of the face and scalp, and an EEG electrode layout is presented in Figure 2.2. From Figure 2.2, it is clear that the electrodes will detect significant muscle activity either during muscle contraction or during muscle twitches (sporadic contractions).



(a) **(b)**

Figure 2.2: Muscles of the face and scalp approximately overlaid with an EEG electrode layout. **(a)** The muscles of the face and scalp with electrodes approximately overlaid as white circles (adapted from [30]). **(b)** An electrode layout over a subject's head (adapted from [31]).

The reported properties of striated/voluntary muscle potentials measured by superficial electrodes include the bandwidth of EMG signals, EMG amplitude range and peak frequency of signals. The bandwidth of EMG signals range from 20 – 300 Hz, with most power in the lower end of the spectrum [7], [29]. EMG amplitudes range from 1 – 100 µV in temporal regions of the head [7]. When EEG signals are measured in the proximity of the temporalis and frontalis muscles (of the scalp), different frequency ranges are observed for each muscle [29]. The spectrum of the temporalis EMG signals show a peak in the range of 35 – 80 Hz (and a small peak at 20 Hz), whereas the frontalis EMG spectrum ranges from 45 – 60 Hz [29]. These signals attenuate over the remainder of the head [29]. This is particularly relevant in the removal of EMG artifacts from EEG recordings.

2.3 Chapter Summary

Background information has been provided for EEG characteristics and modelling, and EMG artifacts. This information is required to define the methods in this study, as well as in the literature review.

Chapter 3 – Literature Review

Relevant literature for the dissertation is discussed in this chapter. It is structured to initially address how the methods used in this study have been used in literature. Thereafter, the gap in literature is identified, which is followed by simulation techniques and performance measures to address it. The topics that are discussed in this chapter occur as follows:

- Section 3.1 – independent component analysis (ICA), standard ICA (STD-ICA) denoising and the application of spatial constraints to the ICA algorithm, i.e. spatially constrained ICA (SCICA).
- Section 3.2 – wavelet denoising and the parameters associated with it (i.e. the WDPs).
- Section 3.3 – a selection of EMG artifact removal techniques, which make use of ICA, SCICA and wavelet denoising. The gap in literature is identified in this section.
- Section 3.4 – WDP optimisation methods that have been applied to temporal signals. Such methods may be used to assist in addressing the gap in literature.
- Section 3.5 – EEG and EMG artifact simulation techniques.
- Section 3.6 – performance measures to assess the accuracy and efficiency of EMG artifact removal.

3.1 Independent Component Analysis (ICA)

Section 3.1.1 presents an analogy known as the ‘cocktail party problem’ to describe blind source separation (BSS) problems and how the signals in this study represent a BSS problem. ICA is a BSS algorithm that can be used to process such signals. An overview of the ICA algorithm is provided in Section 3.1.2. Finally, STD-ICA denoising and SCICA are described in Sections 3.1.3 and 3.1.4, respectively, for use in the methodology in Chapter 5.

3.1.1 Cocktail Party Problem

During a cocktail party, many conversations occur at the same time [32]. If multiple microphones are placed in the room, several conversations can be recorded simultaneously [32]. With enough recorded conversations it may be challenging to determine what individuals are saying in each recording [32]. Therefore, the problem involves separating the mixture of measurements to obtain each individual's speech signal [32].

The 'cocktail party problem' is an example where BSS can be employed [32]. By definition, BSS problems occur when multiple sensors measure a significant mixture of signals, generated by several sources, which need to be estimated [13]. This can be done without specific knowledge of the source signals prior to mixing [32].

EEG recordings, comprised of multiple measurements taken from electrodes placed on the scalp, include a mixture of neural signals and artifacts [9]. Such artifacts include EMG and electrooculographic (EOG) signals [9]. By separating the recordings into sources, artifacts may be processed and removed from the recording [14]. To perform effective BSS for EEG artifact removal, ICA may be employed [9].

3.1.2 Overview of ICA Model and Algorithm

ICA is an unsupervised, multivariate data analysis algorithm that solves BSS problems [33]. ICA utilises statistical methods to separate independent sources that have been collected in a multichannel fashion [13]. According to Akhtar *et al*. [14], ICA separates a mixture of measurements $\boldsymbol{x}(t)$ into:

1. $\boldsymbol{s}(t)$ – the sources/components that generated the mixed signals.
2. $\boldsymbol{A}$ – the mixing matrix which mixes the components to produce the given mixed signals.

In its most basic and linear form, this is modelled according to Equation 3.1 [14]:

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{v}(t) \tag{3.1}$$

Where M sensors yield $\boldsymbol{x}(t) = [x_1(t), x_2(t), \dots, x_M(t)]^T$ measurements, which is a linear mixture of the desired N sources $\boldsymbol{s}(t) = [s_1(t), s(t), \dots, s_N(t)]^T$ that have been mixed according to a mixing matrix, $\boldsymbol{A}$, of size M x N [14]. Additive noise sources are described by $\boldsymbol{v}(t) = [v_1(t), v_2(t), \dots, v_M(t)]^T$ [14]. Equation 3.1 can be expressed in matrix form as: **X** = **AS** + **V**. For simplicity, the matrix representation is used in the remainder of this section.

In this study, **X** refers to the EEG signals measured by electrodes [14]. An artifact-free version of **X** is desired for this study. Methods to achieve this are described in Section 3.3 and require estimation of **S** and **A**. **S** represents the neural signals and artifact sources (such as EMG artifacts), called components [14]. **A** is the mixing matrix that describes how the sources in **S** combine to form **X**. **V** is additive noise (such as line noise) [14].

According to James and Hesse [13], to apply ICA to EEG signals, it must be assumed that:

1. The mixing of signals is linear and instantaneous.
2. The number of sources **S** is less than or equal to the number of sensor channels (electrodes).
3. The mixing matrix **A** is stationary – i.e. the electrodes do not move during measurement.
4. The sources, **S**, are mutually independent and have non-Gaussian distributions [33].

The general ICA algorithm will be described. This is followed by a brief discussion of variations of the algorithm and limitations of ICA.

General ICA Algorithm

According to Hyvärinen [33], ICA is typically divided into two stages:

1. Pre-whitening of the measured data to orthogonalise the measured signals in **X** so that the signals are decorrelated, have less redundancy, and have unit variance [34].
2. ICA estimation.

For simplicity, additive noise **V** – in Equation 3.1 – is assumed to be zero. The whitening procedure maintains the ICA model (in the form of **X** = **AS**) and is shown in Equation 3.2 [33]:

$$\boldsymbol{H} = \boldsymbol{QX} = \boldsymbol{QAS} = \widetilde{\boldsymbol{A}}\boldsymbol{S} \tag{3.2}$$

The measured signals, **X**, are whitened by a linear transformation using matrix **Q** to produce matrix **H**, an orthogonal set of signals [33]. This can be done using an algorithm such as principal component analysis (PCA) [33].

The transformed mixing matrix, $\widetilde{\boldsymbol{A}}$, is now orthogonal [33]. This is useful as the number of parameters/elements that need to be estimated in **A** are reduced [33], [34]. Assuming that **A** is an n x n matrix with n^2 elements, $\widetilde{\boldsymbol{A}}$ will contain $n(n-1)/2$ elements after whitening [33], [34]. Typically, numerical optimisation of orthogonal matrices is faster and more stable than numerical optimisation of non-orthogonal matrices [33]. $\widetilde{\boldsymbol{A}}$ can now be estimated by maximising an objective function related to the non-Gaussianity of the components [33].

The two most common optimisation approaches are maximum-likelihood estimation and the minimisation of the mutual information (MI) between components, which yield the same objective function [33]. MI is discussed further in Section 3.6.2. In short, MI measures the dependence of one signal to another, and can therefore be used to estimate components that are as independent as possible [33].

The objective function, $L(\boldsymbol{W})$, used in these algorithms is described in Equation 3.3 [33]:

$$L(\boldsymbol{W}) = \sum_{i=1}^{M} \sum_{t=1}^{N} G_i \left(\boldsymbol{w}_i^T \boldsymbol{h}(t) \right) \tag{3.3}$$

Where:

- **W** is the inverse of $\widetilde{\boldsymbol{A}}$ (where **S** = **WX**), with rows $\boldsymbol{w}_i^T$. It must be assumed that **A** is invertible [34].
- G_i is the logarithm of the probability density function of each component $s_i(t)$ of **S** and is estimated by $\boldsymbol{w}_i^T \boldsymbol{h}(t)$.
- $\boldsymbol{h}(t)$ is the whitened version of the observed points $\boldsymbol{x}(t)$.

G_i is often approximated with a smoothed version of the negative absolute function, and is shown in Equation 3.4 [33]:

$$G(u) = -\log(\cosh(u)) \tag{3.4}$$

$L(\boldsymbol{W})$ is then maximised while maintaining orthogonality between the rows of **W** [33]. Once **W** is determined, **A** can be obtained by finding the inverse of **W** and **S** can be acquired by multiplying **W** by the measured signals **X**.

$L(\boldsymbol{W})$ can be optimised by a variety of optimisation methods [33], which constitutes the variations of the ICA algorithm.

Variations of the ICA Algorithm

The general ICA algorithm has been adapted into various algorithms such as FastICA, extended information maximisation (InfoMax), second order blind identification (SOBI), and adaptive mixture of independent analysers (AMICA) [9], [13]. The use of each algorithm depends on the given application. Hyvärinen *et al.* [34] discuss such algorithms and provide justification for their selection. FastICA is briefly described for its applicability in the methodology and for its use in SCICA, as detailed in Section 3.1.4.

FastICA is an efficient, fixed-point iteration algorithm that maximises non-Gaussianity [34]. FastICA is more efficient than gradient based methods, which are commonly used in optimisation [34]. FastICA convergences rapidly and, unlike gradient based algorithms, has no learning rate and no other adjustable parameters [34]. This means that FastICA is easy to use and reliable [34]. For a complete description of FastICA, refer to the original study by Hyvärinen [35].

Following the application of FastICA (or another variant of the ICA algorithm) on the mixed signals, the mixing matrix and the components are obtained. As with any algorithm, ICA has limitations.

Limitations of ICA

According to Hyvärinen [33], ICA is limited in the following ways:

1. The signs and scaling (amplitudes) of components are not determined.
2. The ordering of components is not determined and is not necessarily constant when ICA is applied to the same set of signals.

The first limitation is of minimal relevance as the signs and scaling of components are accounted for by the mixing matrix. The second limitation may be relevant to automated processing if a specific order of components is required. However, if the components need to be manually identified (as in STD-ICA), the order of components is irrelevant.

3.1.3 Standard ICA (STD-ICA) for Artifact Removal

STD-ICA is the typical approach for artifact removal in EEG applications [14]. In STD-ICA, unconstrained ICA (described in the following paragraph), such as FastICA, is applied to EEG data, **X**, to obtain the mixing matrix, **A**, and components, **S** [14]. The artifact components in **S** are then visually identified and zeroed [14]. Thereafter, reconstruction occurs by multiplying **A** and **S** to obtain the cleaned EEG data [14].

Unconstrained refers to using the general ICA algorithm which is 'blind' to the properties of the measured system, as in BSS. This is contrasted to constrained ICA where properties of the source signals are incorporated into the algorithm [9]. The constraints may be temporal or spatial [9].

STD-ICA is limited by the need for manual identification of artifact components. STD-ICA is typically slower than approaches that can be automated and therefore reduced computation time is required [14]. One such algorithm that achieves this is SCICA.

3.1.4 Spatially Constrained ICA (SCICA)

SCICA is an adaptation of the ICA algorithm that utilises prior knowledge about the spatial topography of the system [15]. According to Hesse and James [16] and Akhtar *et al.* [14], SCICA has three advantages over unconstrained ICA:

1. Reliable ICA decomposition.
2. Reduced ICA computation time.
3. Known order of constrained sources which is useful for automation of the algorithm and in turn, for automated noise removal.

SCICA occurs in two stages which are elaborated on in the below subsections. Firstly, the first *n* columns of mixing matrix **A** (in Equation 3.1) are predefined with the expected spatial topography of *n* relatively known sources [16]. These columns (termed constrained columns) are then subject to constraints (termed spatial constraints) throughout the iterations of the ICA algorithm [16]. Secondly, ICA decomposition occurs in which the constrained columns undergo an adjusted algorithm, while unconstrained columns undergo unconstrained ICA decomposition [16].

The below sections describe: the spatial constraints that are applied to the mixing matrix **A** and the determination of spatial topographies of interest; adjustments to the FastICA algorithm for SCICA; and the applicability of SCICA to EEG artifact removal.

Implementation of Spatial Constraints

To use SCICA, the assumptions of independence and non-Gaussianity in Section 3.1.2 must be relaxed for sources corresponding to constrained columns [15]. Such sources are called constrained sources [16] and are assumed to be independent of unconstrained sources while not necessarily being mutually independent of them [15]. The columns subject to constraints are defined in Equation 3.5 [15]:

$$\boldsymbol{A} = [\widehat{\boldsymbol{A}}_c, \boldsymbol{A}_u] \tag{3.5}$$

$\widehat{\boldsymbol{A}}_{\boldsymbol{c}}$ is the resulting constrained columns of **A** from the initial, predefined columns $\boldsymbol{A}_c$. $\boldsymbol{A}_u$ is the remaining unconstrained columns of **A** and contains initial guesses to ensure that the columns in **A** are independent [15], [16]. The initial guesses for the constrained columns, $\boldsymbol{A}_c$, can be derived from a pre-determined set of sensor projections [16]. In this study, the sensor projections correspond to the projection of EMG artifacts across EEG electrodes. The sensor projections are referred to as the reference or constraint topographies [16].

These topographies can be obtained using a mathematical model of the system or by a manual selection of sources [16], [36]. Manual selection requires applying an unconstrained BSS method (such as FastICA) to a previous set of data [16], [36].

Manual selection is the easier of the two approaches and is done by selecting the mixing vector/s in matrix **A** corresponding to the component/s containing the source/s of interest [36]. The mixing vector/s give an approximation of the spatial distribution of the source/s [36]. To use this approach, it must be assumed that knowledge of one or several mixing vectors is available and that the spatial, prior knowledge about the source/s is reliable [36]. In this study, the sources of interest are those corresponding to EMG artifacts.

Throughout SCICA iterations, $\widehat{\boldsymbol{A}}_{\boldsymbol{c}}$ will diverge from $\boldsymbol{A}_c$ depending on the implemented type of constraint [15]. Depending on the level of uncertainty of accuracy of the reference/constrained topography, different applications of

constraints can be implemented [16]. Categories include hard (high certainty), soft (intermediate certainty), and weak (low certainty) constraints [16]. According to Hesse and James [16], columns that are subject to:

- Hard constraints – remain fixed.
- Soft constraints – allow for small deviation around the initial guess.
- Weak constraints – are used for an initial guess, after which ICA proceeds without constraint.

Soft constraints are generally preferred as they allow for a limited amount of angular divergence from the reference/constraint topography [16]. This restricts the amount of deviation allowed for the constrained columns in the mixing matrix as the SCICA algorithm iterates [16].

Implementation of SCICA with FastICA

SCICA can be implemented for BSS algorithms that either do or do not require pre-whitening [16]. For example, FastICA requires pre-whitening whereas the original InfoMax algorithm does not [16]. As FastICA is used in the methodology, SCICA is described for this algorithm. The methodology for other algorithms is described in detail by Hesse and James [16].

For statistically independent sources, the orthogonal mixing matrix, **H**, in Equation 3.2 is orthonormal [16]. Orthonormality must be preserved after each iteration of the algorithm by a method such as Gram-Schmidt orthonormalisation [16]. In SCICA, **H** is represented by Equation 3.6:

$$\boldsymbol{H} = [\widehat{\boldsymbol{H}}_c, \boldsymbol{H}_u] = \boldsymbol{Q}\boldsymbol{A} = \boldsymbol{Q}[\widehat{\boldsymbol{A}}_c, \boldsymbol{A}_u] \tag{3.6}$$

As **H** is required to be orthonormal and $\mathbf{A}_c$ is specified, $\mathbf{H}_c$ will not necessarily be orthonormal [16]. Therefore, provided that all constrained columns are subject to the same type of constraint (hard or soft constraints), Hesse and James [16] propose the following approach for each iteration of the FastICA algorithm:

1. Apply FastICA to unconstrained columns $\widehat{\boldsymbol{H}}_u$.
2. Set $\widehat{\boldsymbol{H}} = [\widehat{\boldsymbol{H}}_u, \widehat{\boldsymbol{H}}_c]$ and apply Gram-Schmidt orthonormalisation, which changes $\widehat{\boldsymbol{H}}_c$.
3. Apply the spatial constraints to $\widehat{\boldsymbol{H}}_c$ within the now orthonormalised space.
4. Set $\widehat{\boldsymbol{H}} = [\widehat{\boldsymbol{H}}_c, \widehat{\boldsymbol{H}}_u]$ and apply Gram-Schmidt orthonormalisation, which changes $\widehat{\boldsymbol{H}}_u$.

The application of orthonormalisation to $\widehat{\boldsymbol{H}}$ and spatial constraints to $\widehat{\boldsymbol{H}}_c$ is sufficient to enforce changes based on the changes applied to the unconstrained columns [16].

Application of SCICA to Artifact Removal from EEG

In EEG denoising, artifact sources are of interest and can be targeted with SCICA for stable and efficient ICA decomposition [14], [16]. An unconstrained BSS algorithm such as FastICA can be used to obtain the reference topographies for artifact sources, which can then be used as initial guesses in the constrained columns $\mathbf{A}_c$ [16], [36]. Thereafter, SCICA need only be performed until the artifact components have converged [14]. In doing so, computation time is reduced – not all components need to be calculated in each iteration of ICA [14], [15]. Finally, processing of the artifact components may occur.

3.1.5 Summary

ICA is an unsupervised BSS method that can be used to separate a mixture of signals into the sources/components that generated the mixed signals and a mixing matrix which mixes the components to produce the given mixed signals.

The ICA algorithm has been adapted into several variations including FastICA, InfoMax, SOBI, and AMICA. These algorithms may be used to denoise a mixture of signals by STD-ICA. In STD-ICA, artifact components are visually identified, zeroed, and multiplied by the mixing matrix, resulting in denoised signals.

ICA can be spatially constrained by utilising the topographies of signals of interest – such as artifacts – for reliable, stable, accurate, and efficient

decomposition. SCICA may be used to automate denoising algorithms and has applications in artifact removal from EEG.

3.2 Wavelet Denoising

Wavelet denoising is a powerful, versatile technique which has been used extensively in EEG denoising [9]. Various aspects of wavelet denoising are discussed in this section.

Wavelet denoising makes use of the wavelet transform (WT) which models a given, measured signal, $x(t)$, as a sum of a desired signal, $s(t)$, and additive noise, $v(t)$ [37]. The WT can be applied to continuous and discrete signals to localise events in both time and frequency at various, pre-defined resolution levels [9], [17]. The stationary WT (SWT) is an adaptation of the discrete WT (DWT) that has been indicated as suitable for denoising applications [14].

The WT, DWT and SWT are described in Sections 3.2.1, 3.2.2 and 3.2.3 respectively. Finally, Section 3.2.4 discusses the wavelet denoising methodology that has been used in literature and that was employed in this study. The parameters that affect the success of wavelet denoising are detailed.

3.2.1 Wavelet Transform (WT)

The WT provides good time-frequency presentation and denoising capabilities [9], [17]. The WT has been used extensively in biomedical signal analysis [9], [17]. As most biological signals (such as EEG and EMG) are non-stationary, the WT is preferable to the Fourier transform for multi-resolution analysis [17], [38]. The continuous WT transforms a time signal $x(t)$ according to Equation 3.7 [17]:

$$X(a,b) = \int_{-\infty}^{\infty} x(t)\psi_{a,b}^{*}(t)dt \qquad (3.7)$$

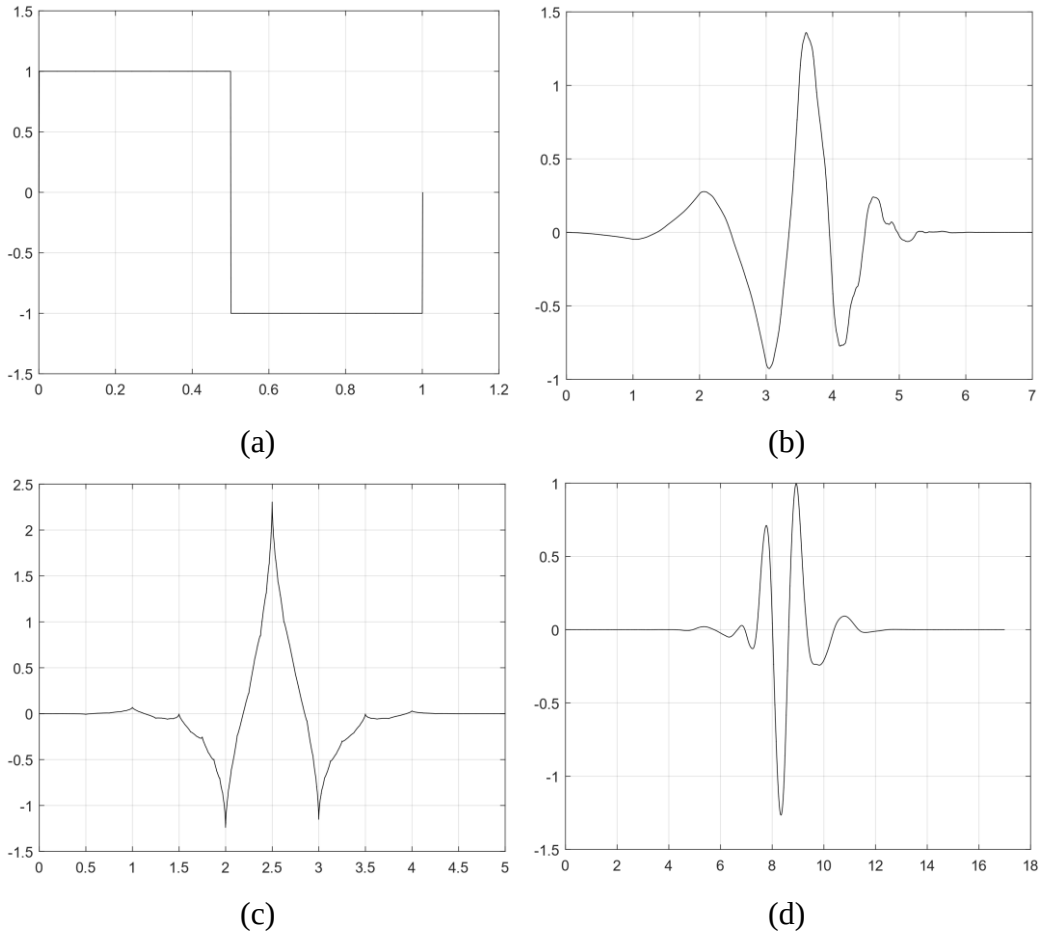


Figure 3.1: Examples of mother wavelets. (a) Haar (haar), (b) Daubechies order 4 (db4), (c) Coiflet order 1 (coif1) and (d) Symlet order 9 (sym9)

Where $\psi_{a,b}^{*}$ is the complex conjugate of the wavelet function, $\psi_{a,b}$, a is the scaling factor, and b is the time shift factor [17]. Factors a and b are real numbers [17]. The wavelet function (or "mother wavelet" [20]), $\psi_{a,b}(t)$, is defined in Equation 3.8 [17]:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right) \tag{3.8}$$

When a is large, the wavelet function is stretched [17]. This is useful for the analysis of low frequency components [17]. When a is small, the wavelet function is compressed, which is useful for the analysis of high frequency components [17]. $\psi(t)$ can be of a variety of wavelet families such as the Haar, Daubechies, Coiflet, or Symlet wavelets [38], which are illustrated in Figure 3.1. Each wavelet

family can further be described by the number of vanishing moments, also known as the order of the wavelet [39]. This allows for the generation of different wavelets within a wavelet family – based on their governing equations – to yield such wavelets as Daubechies order 1, 2, 3, and more.

3.2.2 Discrete Wavelet Transform (DWT)

The WT can be applied to both continuous and discrete signals [17]. The DWT is acquired by discretising time, as well as factors a and b [17]. Factors a and b are discretised according to Equation 3.9 [19]:

$$a = 2^j \quad , \quad b = k2^j \tag{3.9}$$

Where j is the decomposition level and k is the temporal localisation [19]. k contributes to the extent to which the mother wavelet is shifted in time as indicated in Equation 3.9. A discrete signal $x(n)$ can therefore be represented as a series of wavelet coefficients (W) for a desired signal $s(n)$ and additive noise $v(n)$ in Equation 3.10 [19]:

$$W_x(a,b) = W_s(a,b) + W_v(a,b) \tag{3.10}$$

The DWT can be modelled as a convolution between the discrete, measured signal and two filters [20]. The filters are a low pass filter $h(n)$ and a high pass filter $g(n)$ [20]. The filters allow for an efficient, recursive approach to analyse a discrete signal, $x(n)$, at different frequency bands [20]. Initially, the bandwidth of these filters occurs at half the bandwidth of $x(n)$ [38].

The output of each filter is N total wavelet coefficients, resulting in a total of $2N$ samples, where N is the number of samples in $x(n)$ [40]. Therefore, every second sample of the filtered outputs is discarded to maintain the total number of samples of the input signal [40]. This is called down sampling [40]. After filtering and down sampling, the output of the low pass filter is called the approximation and the output of the high pass filter is called the detail [38]. This is one decomposition – the first decomposition level.

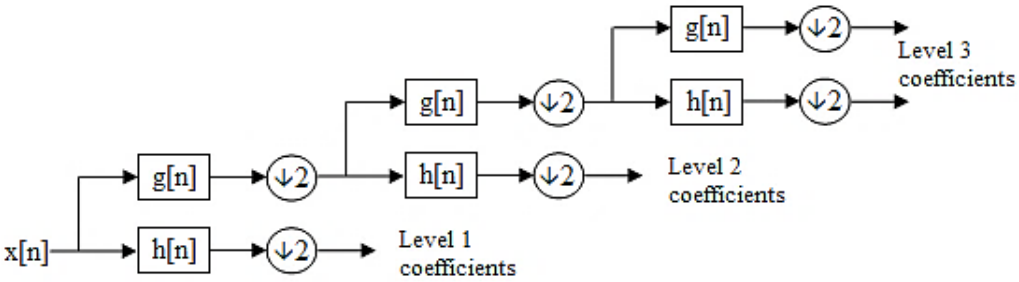


Figure 3.2: Recursive filtering in the application of the DWT for three decomposition levels (adapted from [41])

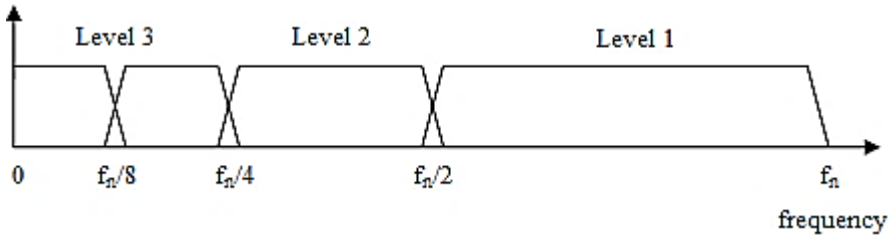


Figure 3.3: Bandwidths of three decomposition levels with the DWT (adapted from [42])

The filtering process can be applied recursively to the approximation of each decomposition level [17]. This is shown for three decomposition levels in Figure 3.2. In Figure 3.2, down sampling is represented by ↓2 [40]. At each decomposition level, the bandwidth of the applied filters is reduced by half according to Figure 3.3, where f_n is the bandwidth of the input signal [40].

3.2.3 Stationary Wavelet Transform (SWT)

The DWT lacks translation-invariance [14]. This means that the coefficients of a time-delayed signal are not correspondingly shifted [43]. Translation-invariance is important in signal processing applications when the arrival time of a signal is unknown [43]. The SWT is an adaptation of the DWT which achieves translation-invariance [14], [43].

The SWT does not down sample like the DWT. Instead, it up samples the filter coefficients [44]. This means that a zero is inserted between each element that the filter acts on [44]. The up sampling of filter coefficients is done by a factor of 2^{j-1} on the j^{th} level of decomposition [14].

It has been shown that wavelet denoising is improved with translation-invariance [45] and as such, the SWT provides good denoising properties [14]. The SWT is also well suited to detect change detection, which is useful for the analysis of signals such as EEG and EMG [14].

3.2.4 Wavelet Denoising Methodology and Wavelet Denoising Parameters (WDPs)

According to Urigüen and Garcia-Zapirain [9] and Akhtar *et al*. [14], wavelet denoising occurs as follows:

1. Determine a threshold for a noisy signal using a threshold determination method. The threshold should separate wavelet coefficients associated with noise (typically lower amplitude) from those associated with the signal (typically larger amplitude).
2. Apply the WT with a chosen mother wavelet to a certain decomposition level. The chosen decomposition level should isolate the signal of interest in time and frequency.
3. Threshold the detail coefficients at the determined threshold using either hard or soft thresholding. This zeroes wavelet coefficients below the threshold. Wavelet coefficients above the threshold are maintained at their value with hard thresholding and are reduced with soft thresholding.
4. Reconstruct the signal with the filtered coefficients using the inverse WT. As a result, the signal is denoised.

From the above methodology it is clear that wavelet denoising is influenced by several parameters. These are termed WDPs. WDPs include the mother wavelet and decomposition level, discussed previously in Sections 3.2.1 and 3.2.2 respectively. Other parameters include the threshold determination method and thresholding type. This section focuses on the role of these parameters in the wavelet denoising methodology.

Mother Wavelet

For improved wavelet denoising, the mother wavelet should approximately match the signal that needs to be denoised [18]. This is not an intuitive choice and optimisation methods can assist with this decision, as described in Section 3.4.

Decomposition Level

As shown in Figures 3.2 and 3.3, the decomposition level describes the number of times that filtering is performed in the DWT or SWT methodology. As such, an appropriate decomposition level is required to localise the desired signal [14]. For too little decomposition, the signal will be poorly localised [14]. For too much decomposition, the signal will be spread across frequency bands, and thresholding the detail coefficients would smear the signal in time [14].

Threshold Determination Method

A threshold should be determined such that noise and signal wavelet coefficients can be adequately distinguished to perform wavelet denoising. There are typically four thresholding methods that are used in wavelet denoising: Sqtwolog, Rigrsure, Heursure, and Minimaxi [46], [47]. There is a discrepancy between Sqtwolog in literature and MATLAB implementation. As MATLAB was used in the methodology, a fifth threshold determination method is defined as Global. The literature in Section 3.3.3 uses the Global threshold determination method and therefore, a distinction is made.

According to Valencia *et al*. [47], Sqtwolog uses the universal method to determine a threshold for a signal according to Equation 3.11:

$$th_j = \omega_j \sqrt{2 \log N_j} \tag{3.11}$$

Where th_j is the threshold, ω_j is the mean absolute deviation (MAD) of the signal, and N_j is the length of the noisy signal [47]. MAD is an estimation of the noise level [37], and is defined in Equation 3.12 [47]:

$$\omega_j = \frac{\text{median}(|\omega|)}{0.6745} \tag{3.12}$$

Where ω is the wavelet coefficient at scale j [47]. In the MATLAB R2015b implementation of the Sqtwolog threshold determination method, Equation 3.11 is not multiplied by ω_j [48]. The Global threshold determination method (using a different MATLAB function) follows Equation 3.11. As MATLAB R2015b was used in the study, the distinction in naming conventions is used in the remainder of the dissertation.

The Rigrsure threshold determination method is defined in Equation 3.13 [47]:

$$th_j = \sigma_j \sqrt{\omega_b} \tag{3.13}$$

Where σ_j is the standard deviation of the noisy signal and ω_b is the b^{th} squared wavelet coefficient chosen at minimal risk from the vector $W = [\omega_1, \omega_2, \dots, \omega_N]$ [46], [47]. W contains square wavelet coefficients that have been arranged in ascending order [46], [47].

The Heursure threshold determination method is a heuristic combination of Global and Rigrsure but is not described mathematically in the reviewed literature [46], [47]. The combination is useful because if the signal to noise ratio (SNR) of the signal is low then Global allows for more appropriate threshold determination than Rigrsure [47].

The Minimaxi threshold determination method is based on the minimax principle in statistics [46]. A fixed threshold is used for minimax performance for root mean square error against an ideal procedure [46], [47]. This is defined in Equation 3.14 where N is the number of samples:

$$th_j = \begin{cases} \sigma\left(0.3936 + 0.10829\, log_2 N\right) & N > 32 \\ 0 & N < 32 \end{cases} \tag{3.14}$$

Thresholding Type

Thresholding can either be hard or soft. The performance of each type depends on the threshold determination method [47]. Hard and soft thresholding are detailed below.

In hard thresholding, wavelet coefficients are zeroed if they are less than the given threshold [47]. Otherwise, they are maintained at their original value according to Equation 3.15 [47]:

$$f(n) = \begin{Bmatrix} W(n) & W(n) > th_j \\ 0 & Otherwise \end{Bmatrix} \tag{3.15}$$

Where $W(n)$ is the wavelet coefficient. In soft thresholding, coefficients are reduced in magnitude if they are greater than the threshold but are zeroed otherwise, according to Equation 3.16 [47]:

$$f(n) = \begin{Bmatrix} sgn\big(W(n)\big) \times \big(W(n) - th_j\big) & |W(n)| \geq th_j \\ 0 & Otherwise \end{Bmatrix} \tag{3.16}$$

Where *sgn*() is defined in Equation 3.17 [47]:

$$sgn(x) = \begin{Bmatrix} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{Bmatrix} \tag{3.17}$$

3.2.5 Summary

Wavelet denoising is a powerful technique with many applications in EEG denoising. The SWT has advantages over the DWT for this purpose. Wavelet denoising is influenced by the choice of WDPs such as the mother wavelet, decomposition level, threshold determination method, and thresholding type. The combination of WDPs will impact the accuracy of wavelet denoising.

3.3 Removal of EMG Artifacts from EEG

A variety of algorithms have been proposed in literature to remove EMG artifacts from EEG. This review addresses different approaches in literature that have been applied to remove EMG artifacts from EEG. This includes:

- Section 3.3.1 – approaches that are not specific to the BSS problem.
- Section 3.3.2 – techniques that are suited to the BSS problem.
- Section 3.3.3 – the WT in combination with ICA to solve the BSS problem more accurately.

3.3.1 Non-BSS Based EMG Artifact Removal

Yong *et al.* [3] and Goncharova *et al.* [29] used spatial filtering (including monopolar, bipolar, Laplacian, and common average reference montages) to remove EMG artifacts from EEG. Both studies reported that facial EMG contamination is significant across all EEG electrodes and that spatial filtering is ineffective in EMG artifact removal [3], [29]. This implies that EMG artifact removal techniques must take all EEG electrodes into account and not only those over, or adjacent to, individual scalp or facial muscles.

Other EMG artifact removal techniques have been used such as adaptive filtering in combination with neural networks [49], but this is rare in literature. In general, BSS approaches are favoured as the problem itself is BSS based.

3.3.2 BSS Based EMG Artifact Removal

Techniques that can solve the BSS problem are generally favoured as a reference EMG waveform is not available [7], [9]. Such techniques include PCA, canonical correlation analysis (CCA), and ICA [9]. Standard PCA is inadequate for the removal of artifacts in EEG, as it models sources as orthogonal [9]. It is more appropriate to model sources as independent, which is done in ICA [9]. Urigüen and Garcia-Zapirain [9] reported that ICA is comparable to CCA for EMG artifact removal, and recommended ICA over CCA for this purpose.

Safieddine *et al.* [11] contrasted and compared the use of ICA, CCA, empirical mode decomposition, and the DWT, each in isolation, to remove EMG artifacts from EEG. Each algorithm was applied to both simulated and real 32-channel EEG data [11]. Simulations were performed for one and two sources of EMG artifacts with varying SNRs [11]. For very low EMG SNR (-30 dB), empirical mode decomposition outperformed the other algorithms [11]. For higher EMG SNRs (-25 to -5 dB), ICA and CCA performed better [11]. The DWT performed poorly at very low EMG SNR but improved in performance at higher EMG SNRs (approximately -20 to -15 dB) [11]. Safieddine *et al.* [11] reported that the performance of ICA and CCA might be improved at very low EMG SNR by utilising additional electrodes.

Taheri *et al.* [50] showed that a number of ICA methods including SOBI and FastICA, and CCA are accurate and efficient in the removal of EMG artifacts from EEG. Leutheuse *et al.* [51] showed that AMICA outperforms InfoMax in the removal of EMG artifacts generated by head and neck movements.

As discussed, EMG artifacts can be removed from EEG with relative success by using variations of ICA. Typically, these approaches are not used in isolation and can be cascaded from multiple channel to single channel techniques as depicted in Figure 3.4 [9]. Figure 3.4 illustrates a BSS algorithm applied to EEG data which generates components. Each component is processed by a single channel based algorithm before the reconstruction process occurs [9]. Spatially constrained, wavelet enhanced ICA (SCWEICA) is an example of such a method and is described in Section 3.3.3. An alternative cascade was implemented by Chen *et al.* [1]. The cascade used a single channel algorithm (ensemble empirical mode decomposition) to generate a multichannel signal which was processed by a multichannel algorithm (CCA) [1]. Thereafter, reconstruction occurred [1], [9].

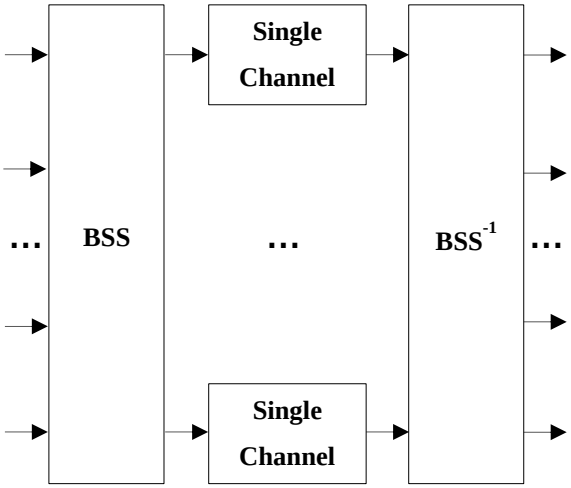


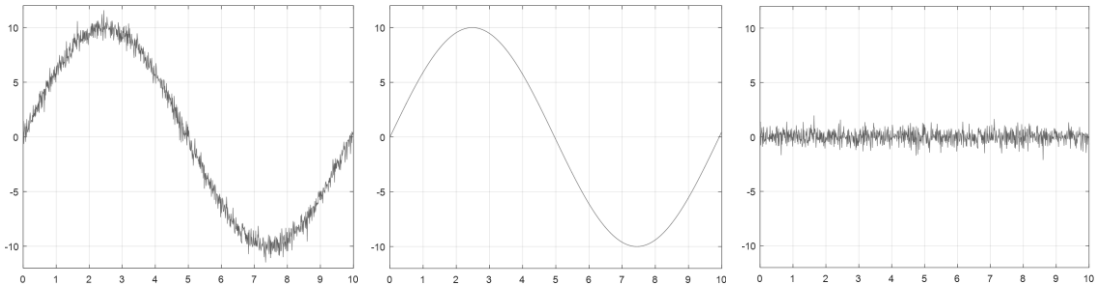
Figure 3.4: Cascading of artifact removal techniques. Adapted from [9].

Multiple denoising techniques may be used in cascade for more accurate denoising. Therefore, to achieve better denoising, a multichannel BSS method may be combined with a single channel method. This combination of techniques is more accurate than applying the methods individually.

3.3.3 Wavelet Enhanced ICA for EMG Artifact Removal

Castellanos and Makarov [19] and Akhtar *et al.* [14] described the use of the WT to remove neural signals that leak into artifact components. The artifact components are output by ICA. Neural signals may leak into artifact components because of a limited number of electrodes [19]. The limited number of electrodes constrains the maximum number of sources that can be approximated by ICA [19]. This is a consequence of assuming that the number of sources is less than or equal to the number of electrodes (see Section 3.1.2).

To address this problem, the artifact component can be modelled as the sum of a large amplitude artifact signal and a low amplitude neural signal [19], shown abstractly in Figure 3.5. During periods where artifacts are not generated, a neural signal remains in the artifact components [19]. Therefore, the artifact components cannot be set to zero (as is done in STD-ICA), because relevant EEG information will be lost [19]. This results in less accurate denoising and therefore, the neural signal should be retained.



(a) **(b)** **(c)**

Figure 3.5: Abstracted wavelet denoising of artifact independent component. **(a)** Artifact component with neural 'noise'. **(b)** 'Denoised' artifact component. **(c)** Extracted neural signal which appears to be 'noise' within the component.

Castellanos and Makarov [19] used the DWT to remove neural signals in the artifact components to ensure that only the artifact is subtracted in EEG reconstruction [19]. Therefore, the neural signal is preserved [19] which allows for more accurate artifact removal. This method was called wavelet enhanced ICA [19]. Akhtar *et al*. [14] stated that Castellanos and Makarov [19] were the first to use wavelet enhanced ICA for the removal of EOG artifacts.

Another approach, similar to Castellanos and Makarov [19], used an automatic, wavelet-enhanced ICA algorithm to remove multiple artifacts, including EMG artifacts [20]. According to Mammone *et al.* [20], this artifact removal algorithm was more successful than wavelet enhanced ICA. Mammone *et al*. [20] used the Daubechies order 4 wavelet (db4) to match the spike-like nature of EEG after assessing different orders of Daubechies wavelets. EMG artifacts were suppressed and removed in both simulated and real settings [20]. While it was stated that the EMG artifacts were successfully removed from EEG, the EMG artifact removal was not quantified.

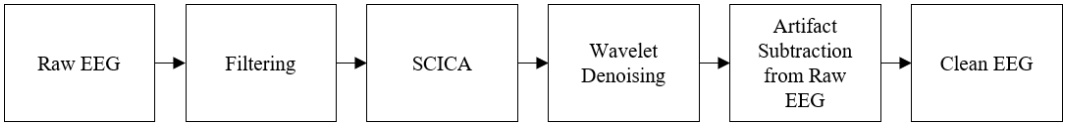


Figure 3.6: High level SCWEICA block diagram (adapted from [14])

SCWEICA was described by Akhtar *et al.* [14]. It is composed of three parts as per its name, which are described in Sections 3.1.2, 3.1.4, and 3.2:

1. Applying spatial constraints to an ICA algorithm (such as FastICA) by using the prior, known, spatial information about the artifacts to identify artifact components [14].
2. Wavelet denoising (using the SWT and Haar wavelet) applied to isolate the artifact from the leaked EEG signal in artifact components (generated by SCICA) [14].

The high level SCWEICA denoising process is depicted in a block diagram in Figure 3.6. After wavelet denoising of artifact components, in which the artifacts are isolated from leaked neural signals, the artifacts are spatially distributed [14]. This is done by multiplying the mixing matrix by the denoised artifact components [14]. Finally, reconstruction occurs whereby the spatially distributed artifacts are subtracted from filtered EEG using matrix subtraction [14].

Akhtar *et al*. [14] compared SCWEICA to STD-ICA and wavelet enhanced ICA. SCWEICA achieved better EOG artifact removal than these algorithms from both simulated and real EEG [14]. For six channels of simulated EOG artifact contaminated EEG, SCWEICA yielded 9.02% normalised mean squared error (NMSE) and MI of 1.34 [14]. NMSE and MI are detailed in Section 3.6. When utilised on real EEG data (for two subjects) SCWEICA yielded an average performance of: 2.16% NMSE and 1.71 MI over 1.05 s denoising duration [14]. With improved hardware and software, the denoising duration could potentially be improved to fall within acceptable times for real-time (RT) use.

Akhtar *et al*. [14] used SCWEICA to remove EOG artifacts from EEG. It is probable that the SCWEICA algorithm can be extended to remove EMG artifacts,

owing to the ability of the WT to denoise EMG artifacts [20]. However, from the noted differences in EOG and EMG (as detailed in Section 2.1.3), it is likely that the selection of WDPs should differ per artifact removal application for more accurate results.

3.3.4 Summary

BSS methods are the most suitable to EMG artifact removal from EEG. ICA has been shown to provide accurate EMG artifact removal and may be improved when used in conjunction with wavelet denoising. By spatially constraining the ICA algorithm to obtain SCWEICA, more efficient and accurate artifact removal from EEG is possible. SCWEICA has been used for EOG artifact removal and is hypothesised to be applicable to EMG artifact removal. However, this may require a different selection of WDPs.

3.4 WDP Selection and Optimisation

The combination of WDPs that should be used in wavelet denoising depends on the particular denoising application [18], [46], [47]. An objective method, such as optimisation, can be used to determine the desirable combination of WDPs per application. WDP optimisation has been performed for a variety of temporal electrophysiological signals such as EEG, EMG, and electrocardiogram (ECG) [18]. However, to the author's knowledge, optimisation has not been performed for the wavelet denoising of artifact components that were produced by ICA.

Section 3.4.1 discusses the justification given in previous studies for WDP selection to denoise artifact components. Section 3.4.2 describes temporal WDP optimisation approaches that have been performed in literature.

3.4.1 WDPs for Artifact Component Denoising

This section describes the justification WDP selection in studies that have applied wavelet denoising to artifact components but did not report optimisation of WDPs [14], [19], [20]. These studies are summarised below:

- Castellanos and Makarov [19] used the Global threshold determination method without specifying the other WDPs used in the study.
- Mammone *et al.* [20] used the Daubechies order 4 (db4) wavelet at the third level of decomposition. The third decomposition level was chosen to approximately match four bandwidths of EEG [20]. The study explored additional, unspecified Daubechies wavelets [20]. Thresholding details were not disclosed.
- Akhtar *et al*. [14] used the Haar wavelet at the fifth level of decomposition. Thresholding was performed using the Global threshold determination method and soft thresholding [14]. Reasons for these choices were justified with literature and were empirical in the case of the decomposition level. However, it is unclear if optimisation of the WDPs occurred.

3.4.2 WDP Optimisation for Temporal Signals

WDP optimisation approaches that have been applied to temporal signals are described in this section. These optimisation methods may be adapted for WDP selection in wavelet denoising of EMG artifact components generated by ICA in the SCWEICA algorithm. By optimising the WDPs in SCWEICA, the neural 'noise' [14] can likely be better removed from the EMG artifact component/s. This will likely allow for more accurate EEG reconstruction and overall denoising.

Rafiee *et al*. [18] suggested using the Daubechies order 44 wavelet (db44) with the WT to process various electrophysiological signals. Such signals include EEG over three midline sites corresponding to the visual cortex, EMG over the forearm, and vaginal pulse amplitude [18]. The correlation coefficient was used to judge how well mother wavelet functions matched the average investigated signals [18]. Refer to the study for details of how the average signals were determined. Rafiee *et al*. [18] investigated 324 continuous and discrete individual mother wavelets from wavelet families at the fourth level decomposition. Wavelet families included the discrete Haar, Daubechies, Coiflet, discrete Meyers,

biorthogonal, and Symlet families [18]. Wavelet denoising was not performed in this study.

Al-Qazzaz *et al.* [46] assessed a variety of WDPs to determine the best mother wavelet and threshold determination method. The investigation was done using soft thresholding at the fourth level of decomposition for EEG denoising [46]. 45 discrete mother wavelets were used, including Daubechies (db1 – db20), Symlet (sym1 – sym20), and Coiflet (coif1 – coif5) [46]. Four threshold determination methods were investigated, including Global (in an implementation without using a logarithmic calculation, as defined in Equation 3.11), Rigrsure, Heursure, and Minimaxi [46]. Denoising performance was assessed with cross correlation, mean square error, SNR, and peak SNR [46]. The results determined that Symlet order 9 (sym9) and the Rigrsure threshold determination method yielded the best performance [46].

Valencia *et al*. [47] compared the wavelet denoising performance of four threshold methods – Global, Rigrsure, Heursure, and Minimaxi – for both hard and soft thresholding. This was done on a single EEG channel with the Haar wavelet at decomposition levels one to ten [47]. White Gaussian noise was added to the EEG signal before denoising [47]. Performance was measured using root mean square error and the correlation coefficient [47]. The study found that the fourth and fifth decomposition levels were optimal for hard and soft thresholding, respectively [47]. Furthermore, the study found that Rigrsure was the best threshold determination method [47].

3.4.3 Summary

Studies that performed wavelet denoising of artifact components selected WDPs without reported optimisation of the WDPs. Such optimisation was performed in studies of temporal EEG wavelet denoising and may be adapted for wavelet denoising of EMG artifact components in SCWEICA. This will likely allow for better isolation of the EMG artifact from leaked neural signals and, in turn, result in improved EMG artifact removal.

3.5 EEG and EMG Artifact Simulations

Simulation is a useful method commonly employed to evaluate the performance of an algorithm [9], [11]. This is particularly relevant to artifact removal from real EEG data as an artifact-free, referential EEG signal cannot be recorded [2], [9]. Therefore, simulations can be used for ground truth evaluation of an EEG denoising algorithm [2], [9]. EEG and EMG artifact simulation methods are described in Sections 3.5.1 and 3.5.2, respectively.

3.5.1 EEG Simulation

EEG simulation can be performed with methods of varying complexity and quality [9]. Such methods include use of the spherical head shell model and inferred simulations from real EEG data.

Spherical Head Shell Model

Makeig *et al*. [52] used a three-shell spherical head model with four dipoles and six scalp electrodes. This approach is limited by the number of electrodes and would need to be expanded to a greater number of electrodes that correspond to typical brain-computer interface (BCI) based EEG analyses [22]. Typically, 32 or 64 electrodes are used in EEG [22].

A free to use EEG simulation package called BESA Simulator [53] was used by Nolan *et al*. [54]. The simulation package allows for dipole sources to be added at desired locations in the spherical head shell model with individual time-frequency conditions [53].

A drawback of the BESA Simulator is that the number of necessary dipoles can be in the order of hundreds, making this process time consuming and tedious. Nolan *et al*. [54] used 200 dipoles in their simulations. Dipoles could be determined by computing the inverse problem, where sources are estimated from measured EEG [28]. However, this requires real EEG data which was out of the scope of this study.

Wavelet Reconstruction

A realistic EEG simulation approach, called Wavelet Reconstruction, is possible by using the DWT to decompose real EEG data within the typical EEG frequency bandwidths [55]. These bandwidths are discussed in Section 2.1.1. A random selection of the wavelet coefficients can then be used to simulate realistic EEG data [55].

Bridwell *et al*. [55] used the DWT on recorded EEG data with a biorthogonal mother wavelet (bior3.9) over five decomposition levels [55]. The decomposition levels were then used to approximate the delta, theta, alpha, and beta EEG bandwidths in terms of wavelet coefficients [55]. The coefficients of each bandwidth were then distributed using a logistic distribution, which has heavier tails than a normal distribution [55]. EEG signals were simulated by selecting random values from each distribution, reconstructing time signals from the coefficients, and adding the time signals together [55]. This was done in the following methodology:

1. The EEG electrode topography was simulated by assigning locations of the electrodes as a series of current sources and sinks [55].
2. Four pairs of electrodes were defined as sources and sinks. One electrode in each pair outputs the time signal for each bandwidth [55].
3. The signals in the remaining electrodes were determined by spherical interpolation using EEGLAB [55].

The MATLAB code for this algorithm is freely available. However, real EEG data is required for Wavelet Reconstruction and each simulation is specific to the individual whose EEG data was recorded. As the use of real EEG data was out of the scope of this study, wavelet reconstruction was not viable.

Yeung EEG Simulation

Yeung *et al.* [24] simulated 31-channel EEG-like data, representative of uncorrelated background EEG noise. The description of uncorrelated background EEG noise is discussed with Phasic Peaks in Section 2.1.1, according to the event

related potential (ERP) processing methodology. The simulation algorithm requires three parameters: sampling frequency, length of the frames and number of epochs. Therefore, prior to the discussion of the simulation method, epochs and frames must be described.

Recorded EEG data is typically divided into a number of epochs (length of time or samples). Epochs are commonly used for repetitive experiments so that the repetitions can be compared to one another [24]. Frames are the number of samples in an epoch.

Yeung *et al.* [24] simulated 31-channel EEG data by summing 50 sinusoids of randomly selected frequencies and phases for each epoch. Simulation was performed separately for each channel [24]. The frequencies of the sinusoids were in the range of 0.1 to 125 Hz and the phases varied from 0 to 2π [24]. The maximum amplitude of the signal, at 0.1 Hz, was set to 20 μV [24]. The remaining amplitudes at each frequency were estimated from the spectrum of an observed EEG [24]. Thereafter, phasic peaks were added to the random sinusoids using a single dipole model, as per the standard EEG forward model [24]. As random parameters are utilised, each EEG simulation will be different.

The same simulation methodology was used again by Yeung *et al.* in a 2007 study [26] and other studies including, but not limited to, [1] and [14]. This method is henceforth referred to as Yeung EEG Simulation.

Yeung EEG Simulation is a simple model used to replicate the primary features of EEG but is not completely accurate [24]. While this approach yields an accurate simulation of the spectral characteristics of EEG, it may not fully capture the synchronisation across multiple channels [9].

The MATLAB code for Yeung EEG Simulation has been made freely available for reuse. It is simple to implement and adapt to this study. Yeung EEG Simulation does not require real EEG data, which falls within the scope of the study. Therefore, this method is superior to Spherical Head Shell Models and Wavelet Reconstruction despite its limitations.

3.5.2 EMG Artifact Simulation

The simulation of EMG artifacts in EEG can be accomplished using a number of methods. Such methods include extracting EMG artifacts from EEG recordings [11] and bandpass filtering random noise to represent specific muscle activity [1], [54], [56]. Both methods produce EMG artifacts that can be added to EEG signals.

Extraction of EMG Artifacts from Real EEG

Safieddine *et al.* [11] examined real EEG data and extracted 50 epochs containing EMG interference. The signals were normalised with reference to the channel with greatest power, and added to simulated EEG signals [11]. It is unclear how this addition was performed. However, it is likely that the addition was performed on a channel-by-channel basis, as the extracted epochs encompass all channels in the time period.

An alternative EMG artifact extraction approach involves instructing a person to contract and relax muscles during an EEG recording [8]. This method simplifies the task of searching for EMG artifacts in EEG recordings. Thereafter, the EMG artifacts can be extracted and added to EEG signals.

The above approaches assume that EMG artifacts are absent when the subject is at rest, and that neurological activity is constant throughout rest and during contraction periods [8]. These two assumptions are inconsistent with the nature of EEG and EMG artifacts [8]. Therefore, the addition of these signals is not truly representative of EMG artifacts in EEG for the purposes of simulation. Furthermore, these approaches require real EEG data which was out of the scope of this study.

Bandpass Filtering Random Noise

Delorme *et al*. [56] modelled EMG artifacts in EEG by bandpass filtering random noise between 20 and 60 Hz with a finite impulse response filter. The noise characteristics were unspecified. The filtered noise was multiplied by a

32-channel muscle map, representative of a temporalis muscle contraction [56]. Temporalis muscle contraction occurs when the jaw is clenched [29]. A muscle map refers to the distribution of a muscle contraction across scalp electrodes – called EMG weights – which were isolated by ICA [56]. The muscle map had high gains at temporal electrodes and near zero gain at other electrodes [56]. This EMG simulation method was also used by Mammone *et al*. [20].

The above method by Delorme *et al*. [56] was adapted for 134-channel EEG by Nolan *et al*. [54]. According to this method, and from the simulation code provided by Nolan *et al*. [54], random noise was bandpass filtered between 20 and 60 Hz using a finite impulse response filter and added to simulated EEG signals on an epoch-by-epoch basis [54]. For each epoch there was a 35% chance that a contraction would be simulated [54]. If a contraction occurred, the length of each simulated EMG contraction occupied a random portion of the epoch [54]. The start and end of the contraction within an epoch, and the amplitude factor, were randomly chosen [54]. The amplitude factor was multiplied by the EMG weights at each electrode [54]. The spatial distribution of the contractions was performed by multiplying the contractions with a 134-channel scalp muscle map, representative of right-side temporalis muscle contraction [54].

Chen *et al.* [1] implemented a similar method whereby random noise was filtered between 20 and 60 Hz. However, a different spatial distribution method was used. The amplitudes of these signals were multiplied by a 19 x 19 mixing matrix [1], as shown in Equation 3.1. Each column of the mixing matrix contained between five and eight non-zero entries to represent the fact that electrodes can detect EMG signals from nearby muscles as the signals attenuate [1].

The method by Chen *et al.* [1] is less representative of spatially distributed EMG artifacts than using a muscle map, as implemented by Delorme *et al*. [56] and Nolan *et al*. [54]. Therefore, for more realistic EMG noise simulation, a muscle map should be used similarly to Delorme *et al*. [56] and Nolan *et al*. [54].

3.5.3 Summary

A variety of EEG and EMG simulation methods have been discussed which may be used to simulate EMG artifact contaminated EEG (EAC-EEG). The best combination of methods is Yeung EEG Simulation and EMG artifact simulation with muscle maps. Both approaches do not require real EEG to generate simulations but used real EEG to produce the algorithms. As such, these methods approximate real EAC-EEG which will allow for assessment of SCWEICA in EMG artifact removal.

3.6 Performance Measures

To assess the performance of an EEG artifact removal algorithm, it is necessary to quantitatively measure its accuracy and efficiency. Accuracy measures include NMSE, MI, and absolute, maximum cross-correlation (AMCC). These measures are discussed in Sections 3.6.1, 3.6.2, and 3.6.3, respectively. A commonly used efficiency measure is duration, which is described in Section 3.6.4. Additionally, it is useful to contrast the performance EEG artifact removal algorithms against each other using relative difference. Relative difference is described in Section 3.6.5.

3.6.1 Normalised Mean Square Error (NMSE)

NMSE is the most prevalent verification measure for artifact removal [9]. It measures the deviation between reference and denoised signals [9]. NMSE is defined in Equation 3.18 [14]:

$$NMSE = E\left\{\frac{\sum_{i=0}^{N}[z_1(i) - z_2(i)]^2}{\sum_{i=0}^{N}[z_1(i)]^2}\right\} \tag{3.18}$$

Where $z_1(i)$ is the artifact free, reference signal, $z_2(i)$ is the cleaned signal after artifact removal, N is the total number of samples, i is the corresponding sample in the signals and $E\{\cdot\}$ is the mathematical expectation [14]. The lower the NMSE, the more closely the cleaned signal represents the reference signal.

NMSE does not fully quantify the dynamics of EEG, as it compares the deviation of signals on a point-to-point basis [14]. Therefore, the cleaned signal must also be quantified with MI [14].

3.6.2 Mutual Information (MI)

MI is a measure of the mutual dependence between two random variables z_1 and z_2 [14], [33]. MI is defined in Equation 3.19 [14]:

$$MI = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(z_1, z_2) log\left(\frac{f(z_1, z_2)}{f(z_1)f(z_2)}\right) dz_1 dz_2 \tag{3.19}$$

Where $f(z_1, z_2)$ is the joint probability distribution function of z_1 and z_2, and $f(z_1)$ and $f(z_2)$ are the marginal probability distribution functions of z_1 and z_2, respectively [14]. MI is non-negative and zero if z_1 and z_2 are independent [14]. The larger the MI, the more related the two variables [14]. When applied to the reference and cleaned EEG signals, it can be determined how closely the cleaned signal represents the artifact-free reference signal [14].

Normalised Mutual Information (NMI)

Normalised MI (NMI) is a version of MI that ranges from 0 to 1 [57]. NMI is calculated by dividing Equation 3.19 by the minimum entropy of the two signals, z_1 and z_2[57]. z_1 and z_2 are modelled as random variables [57]. When comparing two signals, NMI may be preferable to MI as MI is does not have an upper limit.

3.6.3 Absolute Maximum Cross-Correlation (AMCC)

Cross-correlation measures the similarity of two time series functions of equal lengths that have been delayed from one another by a constant delay, λ [58]. Cross-correlation is defined in Equation 3.20 [58]:

$$c_{z_1 z_2} = \frac{1}{N} \frac{\sum_{i=0}^{N-\lambda}(z_1(i) - E\{z_1\})(z_2(i+\lambda) - E\{z_2\})}{\sqrt{\sum_{i=0}^{N-\lambda}(z_1(i) - E\{z_1\})^2 \sum_{i=0}^{N-\lambda}(z_2(i) - E\{z_2\})^2}} \tag{3.20}$$

Where $\lambda \epsilon \{1-N, 2-N, \ldots, N-2, N-1\}$ and $c_{z_1 z_2}$ ranges from -1 to 1 [58]. AMCC is the absolute, maximum value in the cross-correlation signal.

If the AMCC is one, the signals are perfectly correlated at shift λ; if the AMCC is close to zero, there is no correlation between the signals [58]. The AMCC is therefore the largest value determined by cross correlation and if greater than zero, it implies that there is a possible causal relationship between z_1 and z_2 [58].

3.6.4 Computational Efficiency

As the computational complexity of a denoising algorithm is essential for its real-time (RT) applicability, it is necessary to quantify the speed of the algorithm [9]. The time taken for an algorithm to complete (converge) depends on the hardware and software on which the process is run, and on the complexity of the algorithm itself. By measuring the duration of an algorithm, the potential RT use of the algorithm can be determined.

3.6.5 Relative Difference

To compare the performance of an artifact removal algorithm to another, it is useful to compare each algorithm using the same measure using relative difference. This can be done using Equation 3.21:

$$Relative\ Difference\ (\%) = \frac{x-y}{x} \times 100 \qquad (3.21)$$

The interpretation of relative difference depends on the performance measure that generated the results. A larger positive value in terms of AMCC and MI indicates that the denoising algorithm that produced x performs better than the algorithm that produced y. A larger, negative value in terms of NMSE and denoising duration indicates that the denoising algorithm that produced x performs better than the algorithm that produced y.

3.7 Chapter Summary

EMG artifacts are challenging to remove from EEG and may hinder the interpretation of EEG signals. The literature review shows the viability of exploring SCWEICA for EMG artifact removal from EAC-EEG – which has not been performed before. SCWEICA can be adapted for accurate EMG artifact removal by optimising the WDPs. This distinguishes an optimised algorithm (denoted O-SCWEICA) from its unoptimised equivalent (denoted U-SCWEICA). The performance of O-SCWEICA can be quantified by using simulations for ground truth assessment, and a variety of accuracy and efficiency performance measures.

Chapter 4 – Problem Specification

The research question (RQ) and research areas (RAs) for this study are presented in Sections 4.1 and 4.2, respectively. The hypotheses for the study are outlined in Section 4.3, and Section 4.4 describes the expected outcomes of the study.

4.1 Research Question (RQ)

This study aims to answer the following RQ: to what extent can spatially constrained, optimised, wavelet enhanced FastICA remove EMG artifacts from EEG, both accurately and efficiently in time, for later real-time adaptation?

The new EMG artifact removal algorithm explored in this study, optimised SCWEICA (O-SCWEICA), needs to be compared to existing algorithms to assess its relative performance. The existing algorithms are previous EEG artifact removal methods used in literature: SCWEICA and STD-ICA.

In the RQ, accuracy refers to the comparison between ground truth, clean EEG and EEG after EMG artifact removal. The comparison can indicate how closely artifact removal reconstructs the ground truth, clean EEG. Accuracy was quantified in terms of AMCC, MI, and NMSE. In the RQ, efficiency pertains to the computation time of the algorithm. The efficiency measure may indicate the capability of O-SCWEICA to remove EMG artifacts from EEG in real-time (RT).

The four performance measures – AMCC, MI, NMSE, and denoising duration – were used to compare O-SCWEICA to existing EEG artifact removal algorithms, to answer the RQ. The comparison was made using relative difference to easily understand and identify differences in performance.

4.2 Research Areas (RAs)

To answer the RQ, three RAs were identified. The RAs are shown in Figure 4.1 and are each described in Sections 4.2.1, 4.2.2, and 4.2.3.

4.2.1 Research Area 1 (RA1)

RA1 involves the selection of an appropriate decomposition level to ensure that SCWEICA can remove EMG artifacts from EEG. This change to SCWEICA is termed unoptimised SCWEICA (U-SCWEICA). From the description of SCWEICA, Akhtar *et al*. [14] allowed for selection of the decomposition level to isolate a particular artifact. The other WDPs – Haar mother wavelet, Global threshold determination method, and soft thresholding – were fixed.

4.2.2 Research Area 2 (RA2)

RA2 requires the optimisation of WDPs in the SCWEICA algorithm for optimal EMG artifact isolation in the ICA artifact component. This adaptation of the SCWEICA algorithm is termed O-SCWEICA.

4.2.3 Research Area 3 (RA3)

RA3 comprises the performance comparison of EMG artifact removal algorithms applied to EMG artifact contaminated EEG (EAC-EEG) simulations. The algorithms are: O-SCWEICA, U-SCWEICA, and STD-ICA. The performance of the EMG artifact removal algorithms are quantified in terms of accuracy and efficiency, allowing for relative performance comparison.

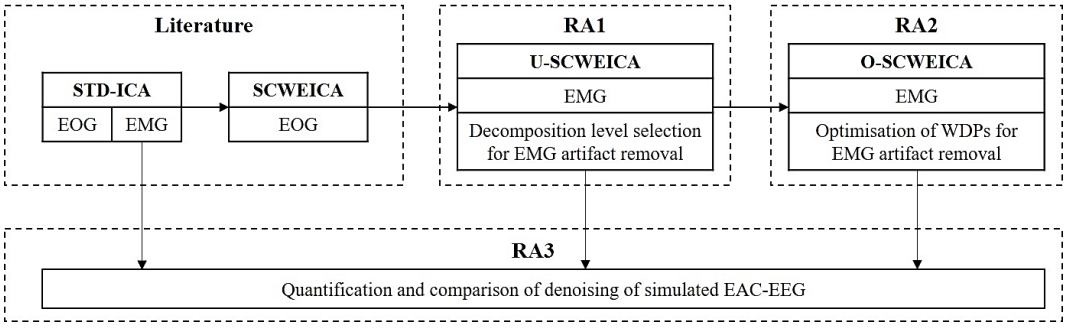


Figure 4.1: Development of RAs

4.3 Hypotheses

It is hypothesised that:

1. The optimisation of WDPs will allow for more accurate EMG artifact removal from EEG when using O-SCWEICA, as compared to U-SCWEICA and STD-ICA.
2. O-SCWEICA can produce more accurate EMG artifact removal (compared to existing algorithms) when applied to relatively short windows of sampled data. Hence, in future O-SCWEICA may be adapted for RT EMG artifact removal.

4.4 Expected Outcomes

The expected outcomes for this research include:

1. A WDP optimisation methodology for EMG artifact removal from EEG, with reference to the components generated by ICA. This has not been done in literature according to the author's knowledge.
2. The first application and quantification of the performance of SCWEICA (unoptimised or optimised) for the removal of EMG artifact from EEG.
3. An indication of whether or not O-SCWEICA is a viable algorithm for RT EEG artifact removal applications.

Chapter 5 – Methodology

This chapter details the methodology developed to implement the research areas (RAs) and to answer the research question (RQ). The RAs necessitate a framework to implement and test the algorithms. Methods for the simulation of ground truth signals are first described. Thereafter, the development of EMG artifact removal algorithms pertaining to each RA are detailed, followed by the experimental procedures for performance analysis. The high-level block diagram for the methodology is presented in Figure 5.1, and is referred to extensively throughout the chapter.

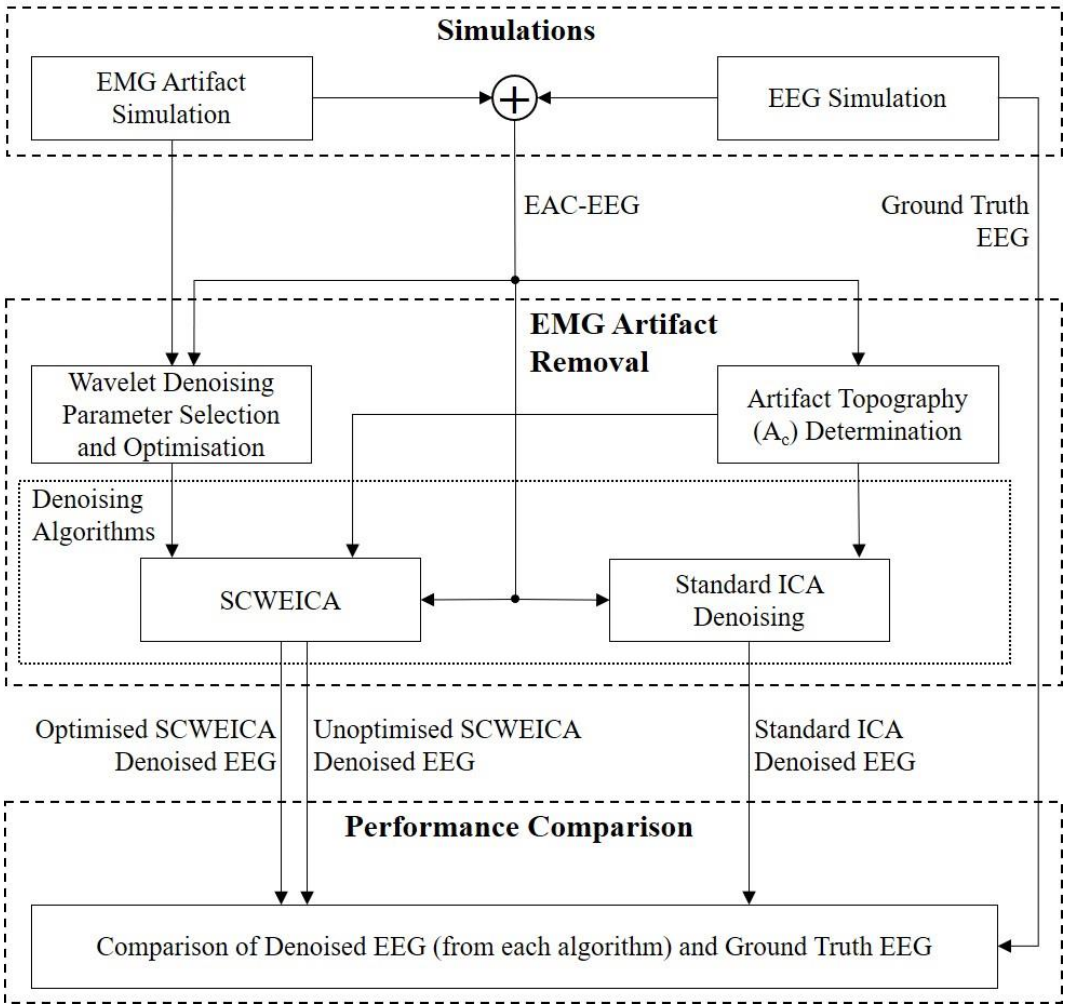


Figure 5.1: High-level block diagram of the methodology used in this study

Section 5.1 describes the simulation of ground truth EEG, EMG artifacts, and their addition to form EMG artifact contaminated EEG (EAC-EEG). A key component of the EMG artifact generation was mapping the electrode layouts from 134 electrodes (as per the original EMG simulation code by Nolan *et al*. [54]) to 31 electrodes (as per Yeung EEG Simulation). Mapping was based on the shortest straight-line distance between the electrodes in each layout, and allowed for the addition of the EEG and EMG artifact signals. Thereafter, the generation of datasets for testing is described.

The implementation of the SCWEICA algorithm, is foundational to RA1 and RA2, and is described in Section 5.2. The implementation for RA1 is described in Section 5.3, where SCWEICA was adapted for EMG artifact removal from EAC-EEG. This was done by selecting the appropriate decomposition level while maintaining the other WDPs, as described by Akhtar *et al*. [14]. Section 5.4 details the implementation of RA2, where the WDPs in SCWEICA were optimised for EMG artifact removal from EAC-EEG.

Section 5.5 outlines the experimental procedure, defined in RA3, to compare the performance of O-SCWEICA against U-SCWEICA and STD-ICA. The experimental procedure is discussed in terms of performance measures and datasets that were employed.

5.1 Generation of Simulated Data

This section describes the “Simulations” block in Figure 5.1, which is divided into four subsections as follows:

- Section 5.1.1 outlines the simulation of the ground truth EEG signals.
- Section 5.1.2 details the simulation of EMG artifacts and the addition of the EMG artifacts to the EEG signals. This results in EAC-EEG, shown in Figure 5.1.
- Section 5.1.3 presents validation of the simulations.

- Section 5.1.4 details the generation of datasets from the simulated signals. The datasets were used for performance assessment of the denoising algorithms, as shown in the Performance Comparison block in Figure 5.1.

5.1.1 Ground Truth EEG Simulation

Yeung EEG Simulation [24], [26], as described in Section 3.5.1, was used in this study to generate EEG-like signals. The MATLAB code for the simulation algorithm was obtained from one of the authors of [24], [26].

The simulated, 'clean', 31-channel EEG serves as a ground truth measure for comparisons after EMG artifact removal from EAC-EEG, as shown in Figure 5.1. This set of signals should represent an ideal signal, after perfect filtering removed all artifacts and noise from real EEG, for a person with no abnormalities in their EEG signals. An example of a 'clean' EEG simulation is presented in Figure 5.2.

The simulation in Figure 5.2 was run for 20 seconds at a sampling frequency of 250 Hz, with 500 frames and 10 epochs, for a total 5000 samples per channel [24], [26]. Each epoch corresponds to two seconds to ensure that EMG simulation would be likely to occur, as discussed in Section 5.1.2, and would be representative of sustained muscle contraction rather than muscle twitches.

The 31 channels used for the simulation are shown topographically over the head in Figure 5.3, and are filled in grey. In Figure 5.3, "nasion" refers to an intersection of three bones just above the bridge of the nose and "inion" refers to the tip of the external occipital protuberance at the back of the head [59]. These points serve as references in analysis of the skull and are used as guidelines for electrode placement [59].

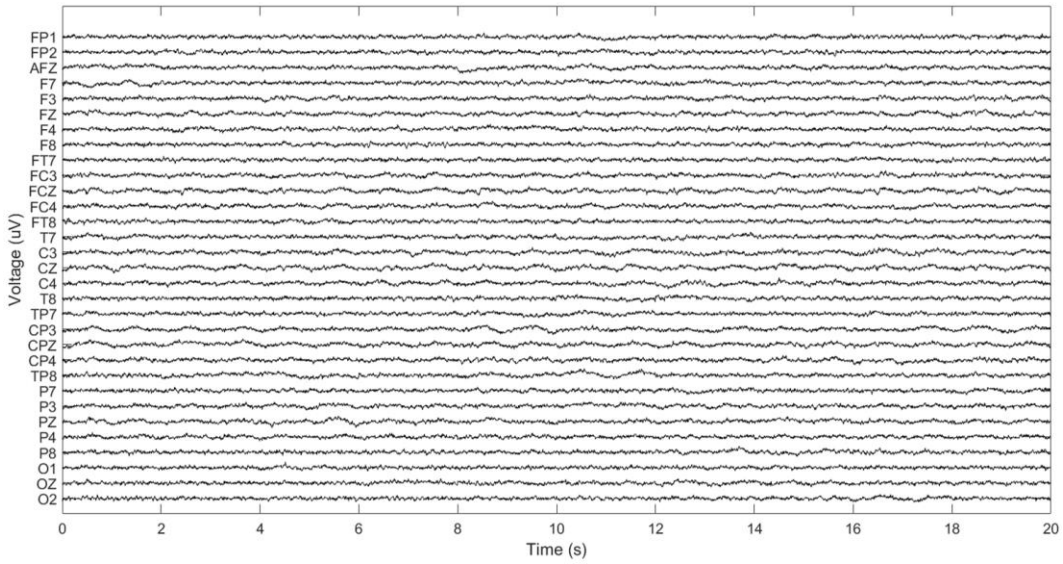


Figure 5.2: A sample of simulated, ‘clean’ EEG

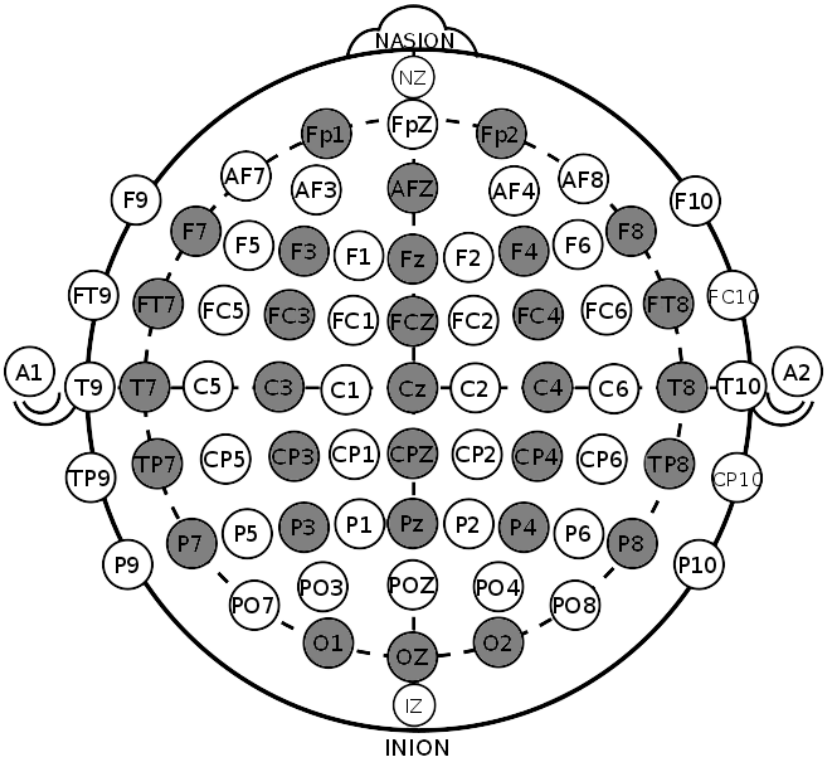
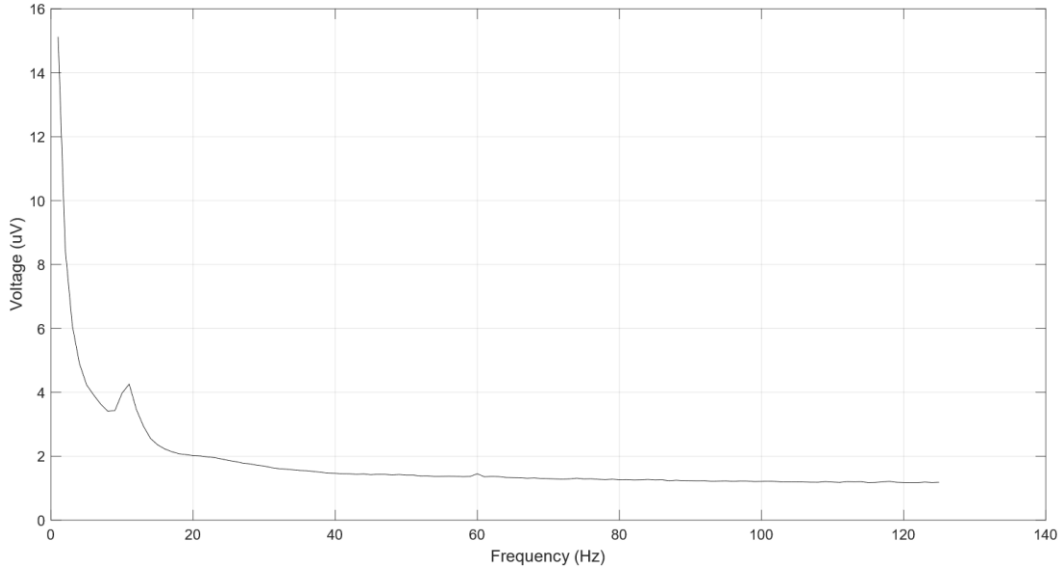
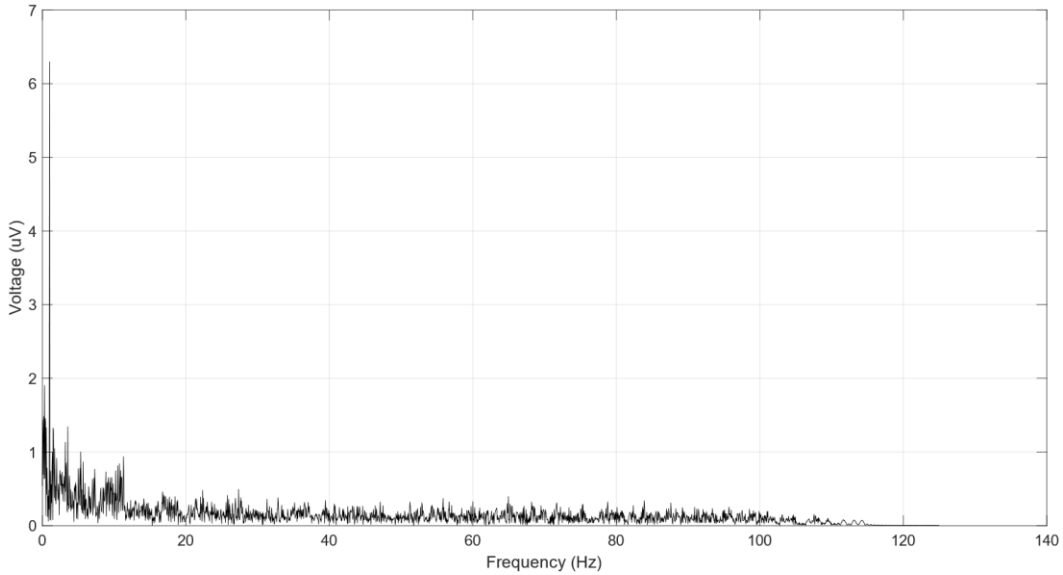


Figure 5.3: A 64 EEG electrode layout according to the international 10-20 system, with the 31 simulated electrodes filled in grey (adapted from [60])



(a)



(b)

Figure 5.4: The spectra of **(a)** typical, observed EEG used to simulate 'clean' EEG signals and **(b)** simulated EEG from channel FC4.

The spectrum of a typical, observed EEG (from the MATLAB code provided by [24], [26]) is shown in Figure 5.4a. Yeung *et al*. [24], [26] derived this spectrum from empirical observations of EEG data in an EEG study of 16 subjects. EEG simulation results in a similar spectrum to that of Figure 5.4a and is shown in Figure 5.4b for a randomly selected channel (FC4 in this case).

According to Yeung *et al.* [24], the simulation method was intended to replicate primary, observable features of empirically acquired EEG data from their experiments. One of the main drawbacks of this method is that the same frequency distribution, in Figure 5.4a, was used across all electrodes and is not electrode specific [24]. Similarly to Yeung *et al.* [24], this study is intended to perform EMG artifact removal and does require a reliable and accurate EEG simulation method to validate the artifact removal algorithm. Consequently, the drawbacks of the method can be disregarded for the purposes of this study.

5.1.2 EMG Artifact Simulation and Their Addition to 'Clean' EEG

This section details the EMG artifact simulation method used in the study to obtain EAC-EEG, indicated in Figure 5.1. EMG artifact simulation was based on Nolan *et al*. [54] and Delorme *et al*. [56], using code provided by Nolan *et al*. [54]. The simulation code was adapted for use in this study. The details of this approach is described in Section 3.5.2.

A key part in implementing the EMG artifact simulation algorithm was mapping the electrode layouts from a 134-electrode topography to a 31-electrode topography. This was required as the EMG simulation code provided by Nolan *et al*. [54] needed to be adapted for use in the 31-electrode layout used in EEG simulation, shown in Figure 5.3.

This mapping enabled adaptation of the 134-channel EMG weights, provided with the EMG simulation code, to represent 31-channel EMG weights. EMG weights are discussed in Section 3.5.2. The EMG weights were then spatially distributed (to represent electric field attenuation) to simulate a spatially distributed EMG artifact from the right temporalis muscle across the 31 electrodes. Mapping the electrode layouts and EMG weights is summarised in Figure 5.5. The EMG weights can then be used to simulate left, right, or both temporalis muscle contractions for the generation of unilateral or bilateral EMG artifacts. Finally, the spatially distributed EMG artifacts were added to the 'clean' EEG signals, resulting in EAC-EEG.

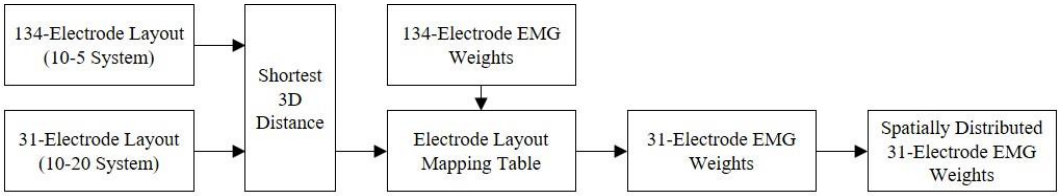
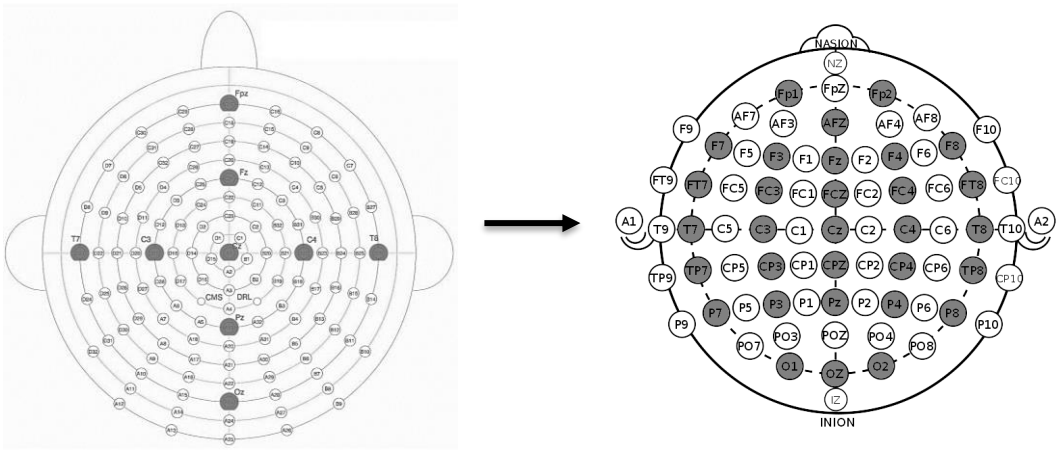


Figure 5.5: Electrode layout and EMG weight mapping process



(a) **(b)**

Figure 5.6: Mapping of a **(a)** 128-electrode layout (the six additional EOG and mastoid electrodes used in [54] are not shown) to a **(b)** 31-electrode layout (adapted from [61] and [60], respectively)

Mapping of Electrode Layouts

The EMG weights provided with the EMG artifact simulation code was for 134 electrodes according to the 10-5 system [54]. This study used 31 electrodes according to the 10-20 system, as per Yeung EEG Simulation [24]. The 10-20 and 10-5 systems are standardised methods for EEG electrode placement, which utilise relative distances based on cranial features [62]. The 10-20 system is used for less dense EEG layouts, such as 21 or 31 electrodes, and the 10-5 system is used for more dense EEG layouts (typically for more than 300 electrodes) [62]. Therefore, to simulate EMG artifacts over a 31-electrode topography, the 134-electrode layout needed to be mapped to a 31-electrode layout. This is shown in Figure 5.6. The mapping procedure is an original algorithm and was specifically developed for this study to add EMG artifacts to EEG signals.

The 134-electrode layout was in three dimensional Euclidean coordinates, however, the 31-electrode layout was in polar coordinates. EEGLAB was used to determine the appropriate XYZ coordinates for the 31-electrode layout [63]. As per EEGLAB, each axis in the XYZ domain ranges from -1 to 1 [63]. Therefore, the distances used in this domain were normalised – relative – distances of the electrodes used in practise.

Relative distances are used as the sizes of peoples' heads vary and the placement of electrodes is based on physical features of the head, primarily the distance from the nasion to inion [64]. The nasion and inion are illustrated in Figure 5.6b. Electrodes are placed at 20% and 10% of the distance from the nasion to inion as per the 10-20 system [64]. Additionally, a reference electrode is required and is placed at the vertex of the head – typically electrode C_z in the 10-20 system – for consistency between subjects.

The electrodes of the two topographies were overlaid in the XYZ domain and the nearest electrodes (based on shortest three-dimensional Euclidean straight-line distance) were selected for mapping between the layouts. The error in mapping electrodes was calculated as a percentage of the diameter of the head. As per the EEGLAB coordinate system, the diameter of the head is 2 units. Therefore, the error was calculated according to Equation 5.1:

$$Error\ (\%) = \frac{distance\ between\ electrodes}{diameter\ of\ the\ head} \times 100 \qquad (5.1)$$

The calculation assumes that the electrodes are represented as points located at the centre of mass of each electrode. Hence, the distance between electrodes is calculated from the centre of one electrode to the centre of another. As such, the spatial characteristics of the electrodes did not need to be considered.

Table 5.1 shows the nearest electrodes between the two electrode layouts, the three-dimensional Euclidean straight-line between them, and the error in the electrode mapping. This constitutes the electrode mapping table shown in Figure 5.5.

Table 5.1: Distances and distance errors between the target electrodes in the 31-electrode layout and the source electrodes in the 134-electrode layout

Target Electrode	Source Electrode	Straight Line Distance	Distance Error (%)
FP1	C28	0.0238	1.19
FP2	C15	0.0238	1.19
AFz	C20	0.0175	0.875
F7	D6	0.0896	4.48
F3	C25	0.1568	7.84
Fz	C21	0.0678	3.39
F4	C12	0.1568	7.84
F8	C6	0.0896	4.48
FT7	D9	0.1194	5.97
FC3	D3	0.1387	6.94
FCz	C23	0.0070	0.35
FC4	C3	0.1387	6.94
FT8	B28	0.1194	5.97
T7	D21	0.0618	3.09
C3	D18	0.0513	2.57
Cz	A1	0	0
C4	B21	0.0513	2.57
T8	B24	0.0618	3.09
TP7	D26	0.0639	3.20
CP3	D28	0.1232	6.16
CPz	A3	0.0339	1.70
CP4	B18	0.1232	6.16
TP8	B16	0.0639	3.20
P7	D30	0.0572	2.86
P3	A6	0.1034	5.17
Pz	A4	0.0513	2.57
P4	B3	0.1034	5.17
P8	B12	0.0572	2.86
O1	A16	0.0381	1.91
Oz	A22	0.0980	4.90
O2	A29	0.0381	1.91

Table 5.1 shows that the maximum error between electrodes was 7.84% for electrodes F3 and C25, and electrodes F4 and C12. These electrodes do not contain large EMG weights, as shown in Figure 5.7b. Therefore, the error is tolerable. As these EMG weights are relatively small, it is assumed that they will not contribute greatly to the simulated EMG artifact.

The electrode mapping table in Table 5.1 allowed for mapping from the 134-electrode layout to the 31-electrode layout. In turn, this enabled the approximation of the muscle maps from one layout to another. This was done using the EMG weights in the provided EMG simulation code, as shown in Figure 5.5.

Determination of 31-Channel, Spatially Distributed EMG Weights

Using the electrode mapping table, the EMG weights of the 134-electrode muscle map (Figure 5.7a) were converted to an equivalent 31-electrode muscle map (Figure 5.7b). In Figure 5.7a, the labels between the start of each set of electrodes, A to D, are not listed for the purposes of visibility. Each set of electrodes, A to D, contain 32 electrodes, whereas set EXG has six electrodes – the additional EOG and mastoid electrodes used in [54].

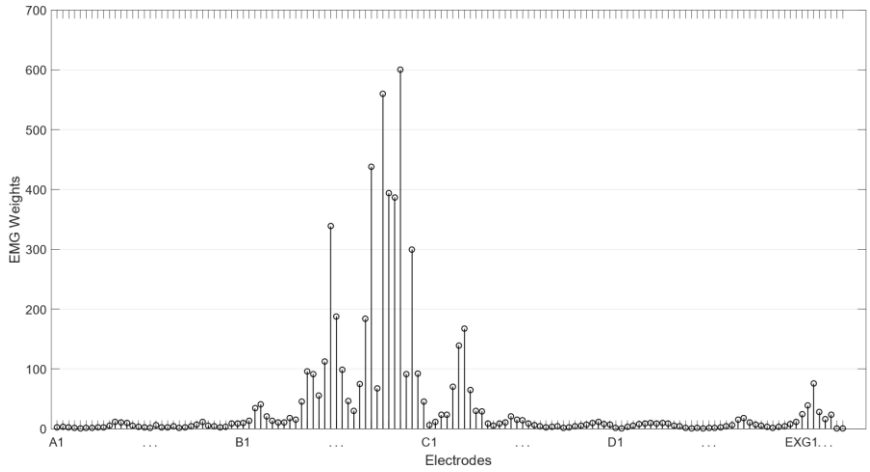
For example, the EMG weight for electrode FP1 in the 31-electrode layout was set as the EMG weight for electrode C28 in the 134-electrode layout. The largest EMG weight in Figure 5.7b is in electrode FT8 which is representative of a right temporalis muscle contraction [65]. This corresponds with Delorme *et al*. [56] and is used as validation of the mapping process.

The mapping of EMG weights from the 134-electrode layout (Figure 5.7a) to the 31-electrode layout (Figure 5.7b) may lose EMG weight information. This is acceptable because when a sparser electrode layout is used rather than a denser electrode layout, there will inevitably be a loss of information. As such, this is realistic for the use of a 31-electrode layout.

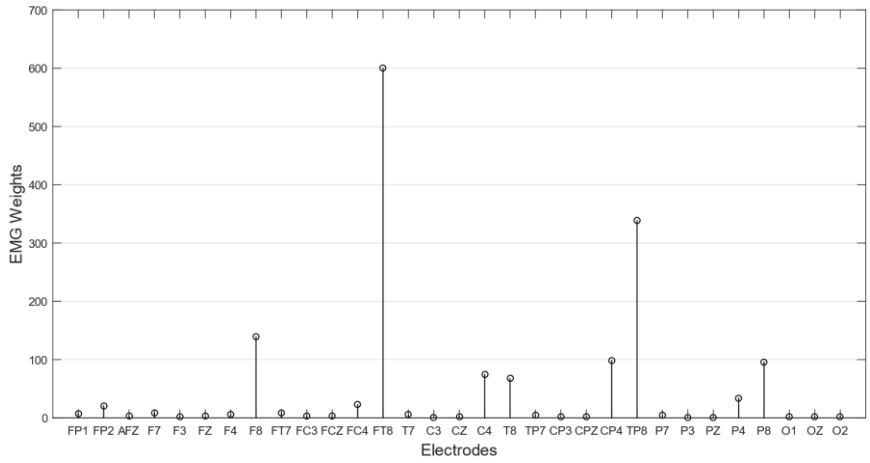
The EMG weights for the 31-electrode layout, shown in Figure 5.7b, were then spatially distributed over the electrode layout, resulting in Figure 5.7c. The spatial distribution follows the code provided by Nolan *et al*. [54] and was performed according to Equation 5.2:

$$W_s = \left|\sqrt{W}\right| \times \sqrt{\max(W)} \tag{5.2}$$

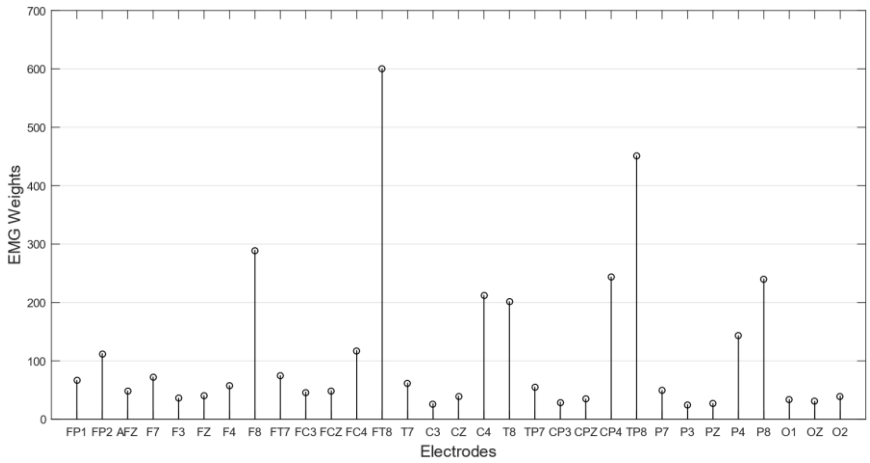
Where W is the vector of EMG weights and W_s is the vector of spatially distributed EMG weights. Spatial distribution allows for realistic EMG artifact simulation. The spatially distributed EMG artifact can then be added to the 31-channel, 'clean' EEG (described in Section 5.1.1), by matrix addition to obtain EAC-EEG.



(a)



(b)



(c)

Figure 5.7: Mapping of EMG weights. **(a)** EMG weights for the 134-electrode layout. **(b)** EMG weights for the 31-electrode layout. **(c)** Spatially distributed EMG weights for the 31-electrode layout.

Side-Based Muscle Contraction Simulation

The muscle map in the EMG simulation algorithm was representative of a right-side temporalis muscle contraction [54]. The algorithm was further adapted to allow for simulation of a left-side temporalis muscle contraction and for the simulated contraction of both temporalis muscles. This resulted in unilateral or bilateral EMG artifacts. The simultaneous contraction of both muscles is the most likely combination of contractions, as the muscles act together to clench the jaw.

The primary electrodes in the 31-electrode layout affected by right-side temporalis muscle contraction are shown in Figure 5.8a. This figure illustrates the primary affected electrodes before spatial distribution of EMG weights. This configuration can be projected for left-side temporalis muscle contraction or for the contraction of both temporalis muscles, before the spatial distribution of EMG weights.

For left-side, unilateral temporalis muscle contraction, the midline EMG weights were maintained, while the left and right sides were swapped appropriately. For a bilateral temporalis muscle contraction, the right-side EMG weights were copied to the left side and the midline was maintained. The primary affected electrodes for the left-side and both sides are shown in Figures 5.8b and 5.8c, respectively. Once the EMG weights are spatially distributed, they affect all 31 electrodes as shown in Figure 5.7c (for right-side temporalis muscle contraction).

If left-side unilateral or bilateral EMG artifacts were simulated, the EMG weights were varied on the left side within a user-defined percentage. For example, if a -5% variation is desired, the left-side EMG weights will be 5% smaller than the original values. The right-side EMG weights were kept constant with the original EMG weights.

The left-side EMG weights were varied as it is unlikely that muscles on either side of the body will have the same contraction power; even if they vary slightly [66]. This may occur owing to hypertrophy or atrophy of one of the temporalis muscles [67], [68].

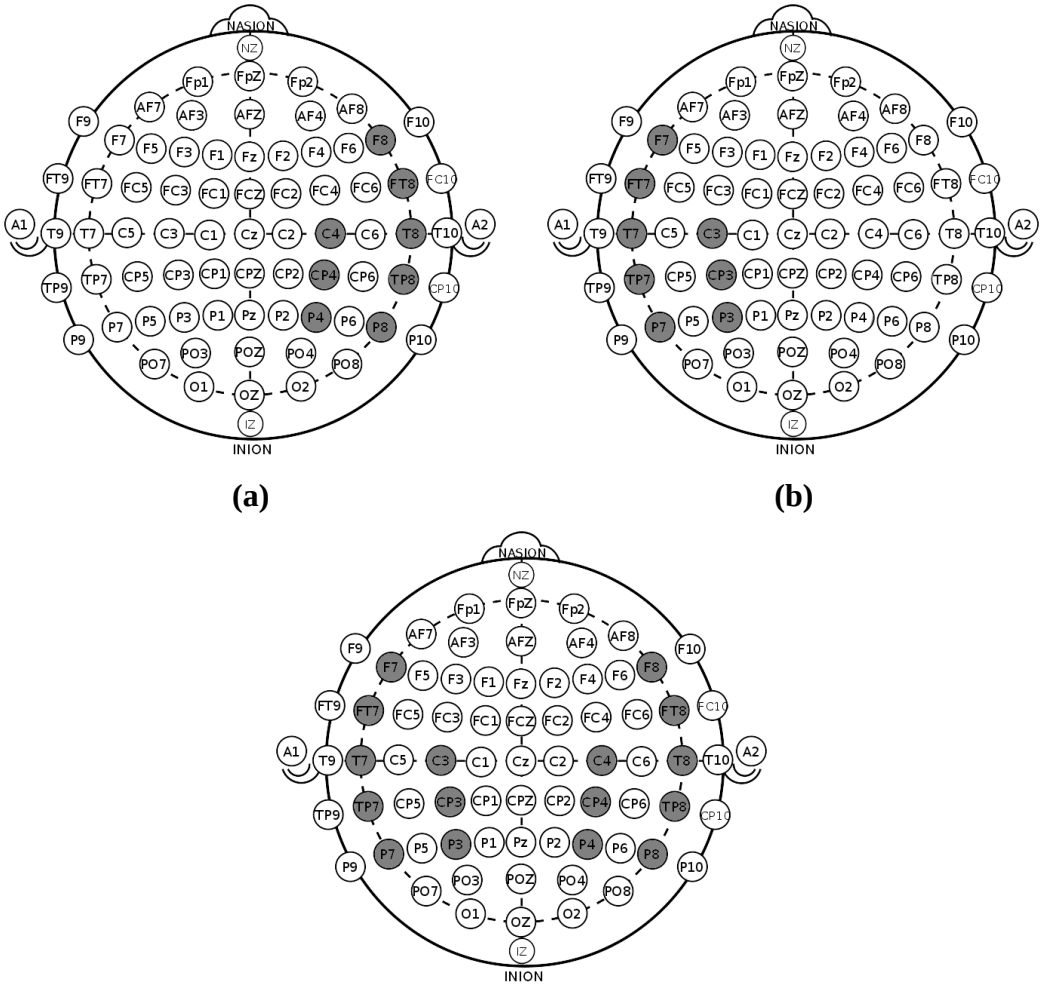


Figure 5.8: Primary electrodes (filled in grey) affected by: **(a)** right-side temporalis muscle contraction, **(b)** left-side temporalis muscle contraction and **(c)** both left-side and right-side temporalis muscle contraction. Adapted from [60].

EMG Artifact Contaminated EEG (EAC-EEG) Simulation

This section describes the simulation of unilateral right-side temporalis muscle contraction. The simulation of left-side or right-side unilateral temporalis muscle contractions will primarily affect different electrodes, shown in Figures 5.8a and 5.8b, respectively, but are fundamentally similar. The same applies to the simulation of bilateral temporalis muscle contraction. An example of simulated EMG artifacts for 31 electrodes is shown in Figure 5.9. The EMG artifact was then added to 'clean' EEG to form EAC-EEG, as shown in Figure 5.10.

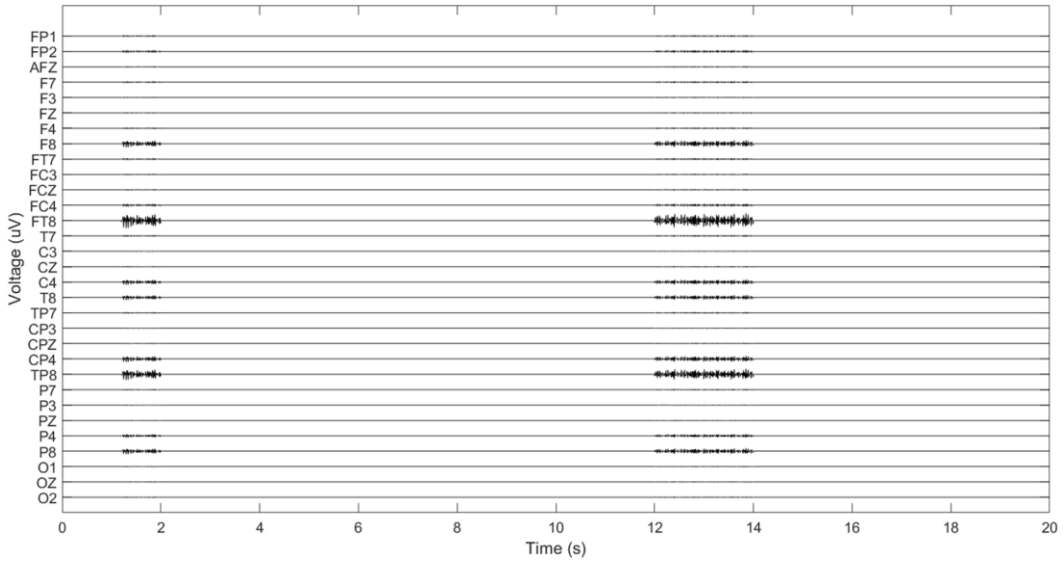


Figure 5.9: Spatially distributed EMG artifacts

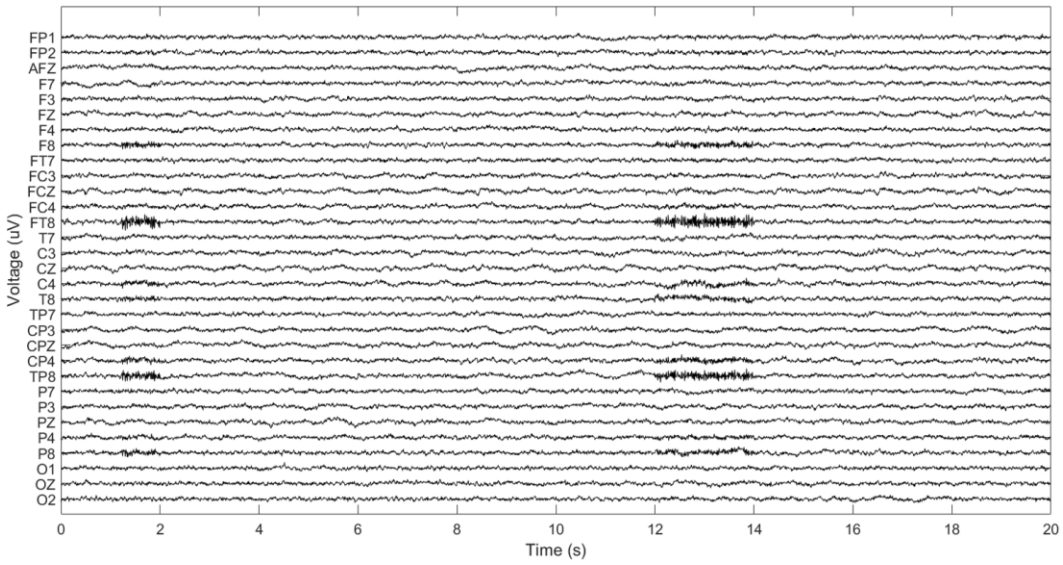


Figure 5.10: Simulated EAC-EEG

A comparison of the spectrum of channel FT8 – the channel with the largest EMG artifact from right-side temporalis muscle contraction – before and after the addition of EMG artifacts is shown in Figures 5.11a and 5.11b, respectively. The isolated EMG artifact in channel FT8 is displayed in Figure 5.11c. There is a distinct increase in power in the 20 to 60 Hz range from Figure 5.11a to Figure 5.11b owing to the addition of EMG artifacts.

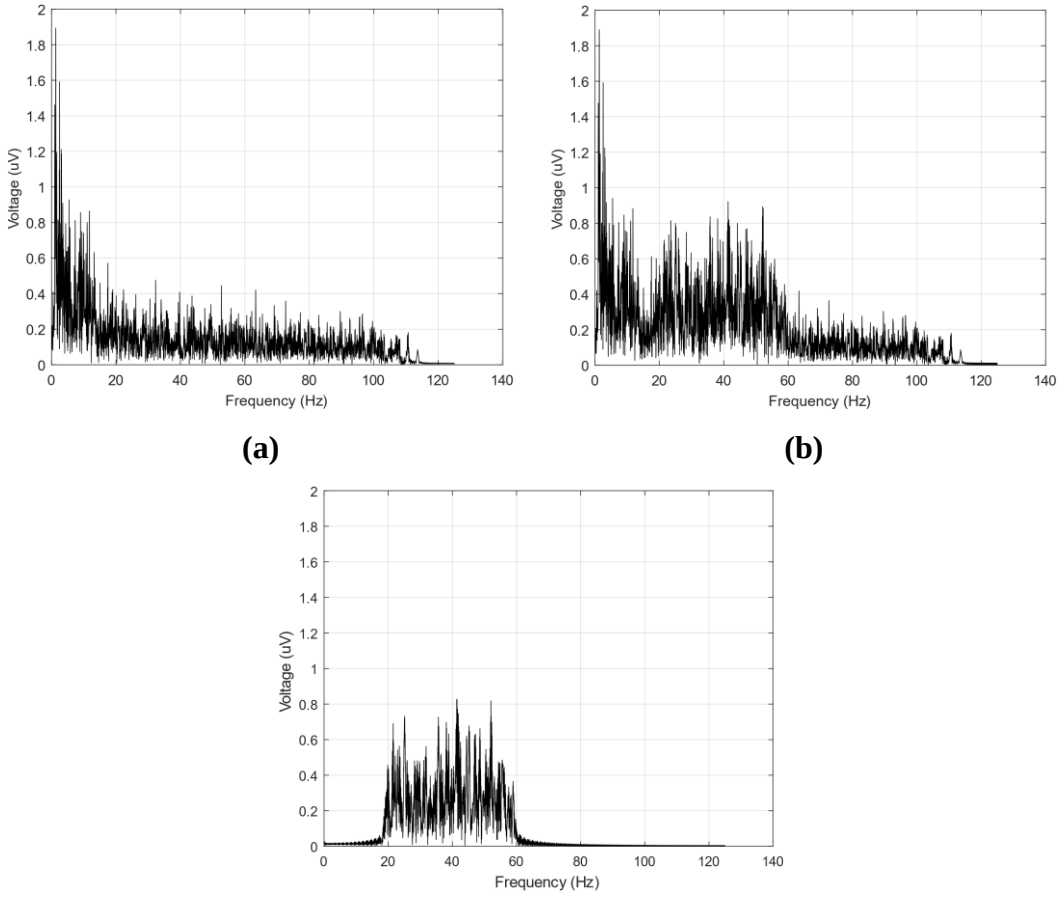


Figure 5.11: Comparison of the spectrum of channel FT8 **(a)** before and **(b)** after the addition of EMG artifacts. **(c)** Spectrum of isolated EMG artifact in channel FT8.

5.1.3 Validation of Simulations

The simulations in this study are assumed to replicate normal EEG signals. Normal EEG is defined as EEG in the absence of identifiable abnormalities, such as signals that are indicative of epilepsy and other disorders [23], [69]. It is difficult to validate the simulations quantitatively because, in the absence of identifiable abnormalities, no feature or characteristic of EEG makes it normal [23]. Therefore, the simulations were validated by methodologies of the studies that introduced the simulation techniques. In this study, simulations were generated utilising two methods:

1. EEG simulation by Yeung *et al*. [24], [26].
2. EMG artifact simulation by Nolan *et al*. [54] and Delorme *et al*. [56].

Yeung *et al*. [24], [26] utilised real EEG from their experiments to generate an average EEG spectrum, which was used in the simulation methodology. The spectrum of simulated EEG in this study is comparable to the spectrum of that by Yeung *et al*. [24], [26], as shown in Figure 5.4. Deviations in the spectrum are a result of the random nature of the simulation method, as detailed in Section 3.5.1.

Nolan *et al*. [54] and Delorme *et al*. [56] derived a temporalis muscle map for EMG artifact simulation from real 134-channel EEG using ICA [56]. The bandwidth of random noise used to represent a temporalis muscle contraction was representative of the indicated muscle [56]. The muscle map was then reduced to 31 channels in this study.

The simulated EMG artifact was subsequently added to the EEG simulation, as illustrated in Figure 5.10. Changes in the spectrum of the simulated signals were observed as expected and are shown in Figure 5.11. Therefore, there is confidence that simulations in this study approximate both real, 'clean' EEG and real EAC-EEG.

5.1.4 Simulated Datasets to Address the RAs

To implement the RAs described in Section 4.1, simulated EAC-EEG was required. Therefore, various datasets were simulated and divided into two groups:

1. A dataset for WDP selection for U-SCWEICA (RA1) and O-SCWEICA (RA2).
2. Datasets for EMG artifact removal denoising performance comparison between O-SCWEICA, U-SCWEICA, and STD-ICA (RA3). These datasets varied in both signal duration and artifact topography to thoroughly assess denoising performance.

The datasets consist of 20 sets of simulated data. Each set comprises a reference signal/s and a signal/s to be cleaned. The datasets are described in Table 5.2 and elaborated on in the subsequent sections. The number of datasets were chosen to

show the effects of artifact removal offered by each algorithm. The datasets were used in the experimental procedures in RA3, which is discussed in Section 5.5.

Table 5.3 shows the signals in the WDP selection dataset and datasets generated for algorithm performance comparison. Unless otherwise specified, simulations were performed for 20 seconds at 250 Hz with 500 frames and 10 epochs – the required parameters in EAC-EEG simulation – for a unilateral right-side EMG artifact with 0% left-side variation.

Table 5.2: Datasets for WDP selection and performance comparison

RA/s Involved	**Dataset Category**	**Dataset Name**	**Number of Datasets**
RA1 and RA2	WDP Selection	WDP	1
RA3	Comparative Datasets	Reference	1
		MATLAB Optimisation	1
		Akhtar Comparison	1
	Duration-Varying Datasets	Towards Real-Time	3
		Offline	5
	Artifact Topography Datasets	Left-Side	7
		Both-Sides	7

Table 5.3: Signals in the WDP dataset and datasets for performance comparison of EMG artifact removal algorithms

Datasets	**Reference Signal/s**	**Signal/s to be Denoised**
WDP	EMG signal before addition to ‘clean’ EEG	EMG artifact component generated by FastICA
Performance comparison of EMG artifact removal algorithms	‘Clean’ EEG	EAC-EEG

Dataset for WDP Selection

The process of obtaining each set of signals in this dataset, denoted the "WDP Dataset", is summarised below:

- EAC-EEG simulation.
- The spatially distributed EMG artifact (before addition to 'clean' EEG) from one channel was saved. Any channel could be chosen since the signal was normalised in the optimisation process, as discussed in Section 5.3.2.
- EAC-EEG was processed with unconstrained FastICA and the resulting components were visually inspected. The artifact component corresponding to the EMG artifact was identified and saved.

Datasets for EMG Artifact Removal Performance Comparison

Three datasets, denoted "Comparative Datasets", were used to compare artifact removal performance. The datasets are the Reference Dataset, MATLAB Optimisation Dataset, and Akhtar Comparison Dataset.

The Reference Dataset was generated to enable performance comparison of all other denoised datasets within the study. The EAC-EEG and corresponding 'clean' EEG were saved for each set of simulations.

As detailed in Section 6.3.2, MATLAB optimises functions with each repeated iteration and, as such, testing numerous datasets concurrently causes the duration of each iteration to decrease (on average). This leads to an inaccurate representation of the efficiency of the algorithms in the initial versus later denoising iterations. To illustrate this, the Reference Dataset was processed a second time but with the optimisation option on. The new dataset was named the MATLAB Optimisation Dataset. Testing of the MATLAB Optimisation Dataset is the only scenario where the optimisation option was enabled. For all other tests, where duration is vital, this option was disabled.

The Akhtar Comparison Dataset was generated to compare the results of this study to Akhtar *et al*. [14]. This dataset was simulated to resemble the simulations

used by Akhtar *et al*. [14]. Therefore, the sampling frequency was set to 200 Hz. 10 epochs and 400 frames were used to correspond with the sampling frequency for 20 seconds of simulated data.

Duration-Varying Datasets for EMG Artifact Removal Performance Comparison

The Duration-Varying Datasets were used to show the impact of varying signal lengths on the performance of each algorithm. The Duration-Varying Datasets were comprised of the Towards Real-Time (RT) and Offline Datasets. The Towards RT Datasets resembled windowed EEG data that would be used in RT denoising. The Offline Datasets resembled recorded EEG data which requires artifact removal. Three datasets were generated for the Towards RT Datasets, and five datasets were generated for the Offline Datasets.

The Towards RT Datasets were generated by extracting windows from the Reference Dataset. Each of the EAC-EEG simulations from the Reference dataset were visually inspected and EMG contractions were identified. Windows of increasing duration were taken from the same point to encompass increasing portions of the EMG artifact (or of samples after the EMG artifact) over all channels. This is illustrated in Figure 5.12, for channel FT8, where the EMG signal is largest. Following the approach in Figure 5.12, three datasets were generated: Towards RT Dataset 1, 2, and 3. These correspond to windows of 1, 2, and 3 seconds, respectively.

The Offline Datasets were simulated with increasing simulation durations. Five Offline Datasets were generated for durations of 40, 80, 120, 160, and 200 seconds. The datasets were named according to the duration of the simulation; “Offline Dataset X” is the dataset for which “X” is the duration. For example, Offline Dataset 40 contains simulations of 40 seconds.

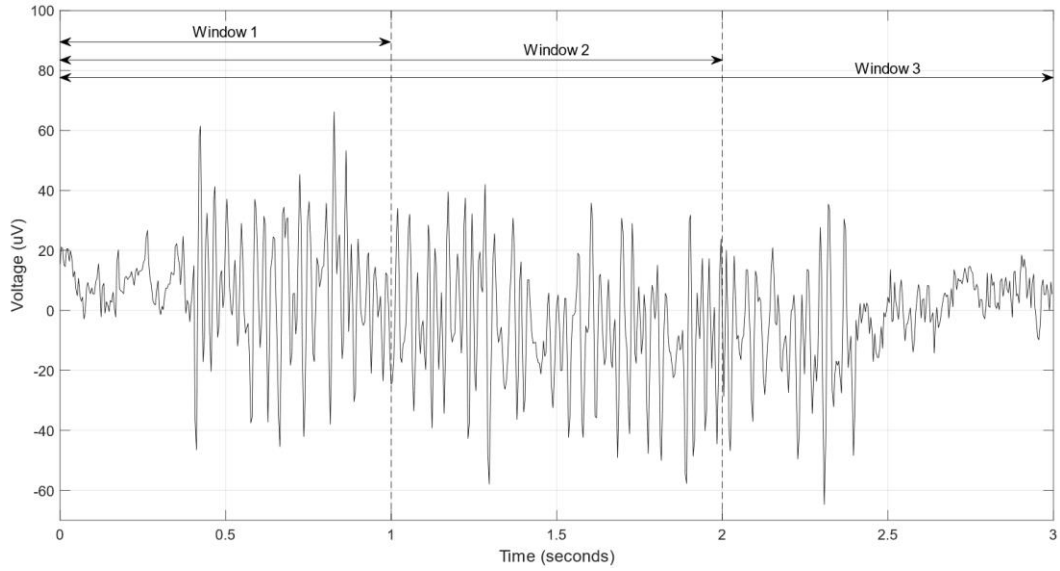


Figure 5.12: An example of the windows (marked by dashed lines) used for the Towards RT Datasets. The signal is the EAC-EEG from channel FT8, where the EMG artifact occurred from approximately 0.4 to 2.4 s. The windowed sections across all channels in Windows 1, 2, and 3 were included in the Towards RT Datasets 1, 2, and 3, respectively.

Artifact Topography Datasets for EMG Artifact Removal Performance Comparison

The left-side EMG weights were adjusted in each of the Artifact Topography Datasets, as described in Section 5.1.2 and shown in Figure 5.8. Seven datasets were generated for bilateral temporalis muscle contractions (termed Both-Sides Datasets) and for unilateral left-side contractions (termed Left-Side Datasets), respectively.

The Artifact Topography Datasets were simulated with varying left-side EMG weights of -90%, -50%, -10%, 0%, 10%, 50%, and 90%. This resulted in seven datasets for each Artifact Topography Dataset which were named according to the left-side EMG weight variation. For example, the Both-Sides -90% Dataset contains simulations with left-side EMG weights set to 90% less than the original weights.

5.2 Implementation of SCWEICA for U-SCWEICA and O-SCWEICA

The SCWEICA implementation to remove EMG artifacts from simulated EAC-EEG is outlined in this section. U-SCWEICA and O-SCWEICA differ in the selection of WDPs but are identical in all other aspects of the SCWEICA implementation. The development of U-SCWEICA for RA1 and O-SCWEICA for RA2 are discussed in Sections 5.3 and 5.4, respectively.

A block diagram of the SCWEICA algorithm, as implemented by Akhtar *et al*. [14] and applied to EAC-EEG is shown in Figure 5.13. Figure 5.14 shows the signals at each step of the SCWEICA algorithm. The figures that required wavelet denoising were generated using the optimal WDPs from RA2 in Section 6.2.1.

Figures 5.13 and 5.14 depict the artifact removal process to obtain denoised EEG in several key steps. The steps are summarised in Section 5.2.1. Thereafter, Sections 5.2.2, 5.2.3 and 5.2.4 describe the implementations of SCICA, wavelet denoising and EEG reconstruction, respectively.

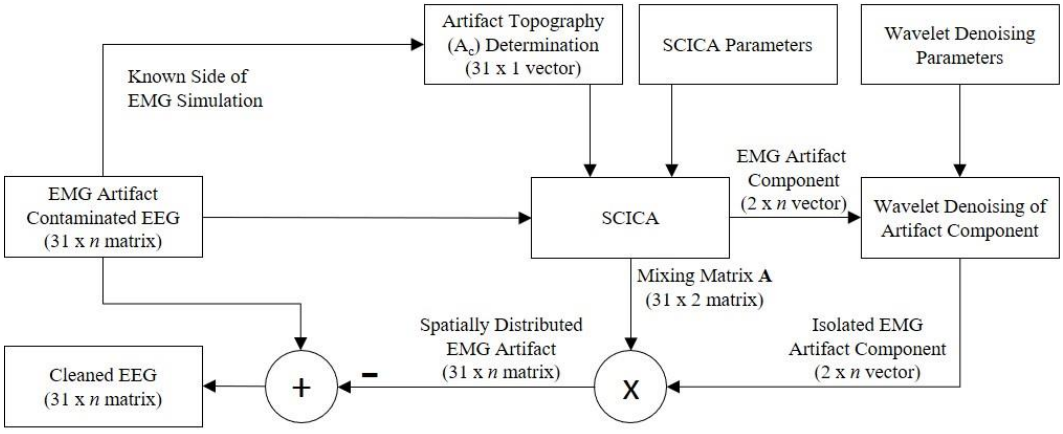


Figure 5.13: SCWEICA block diagram (adapted from Akhtar *et al*. [14])

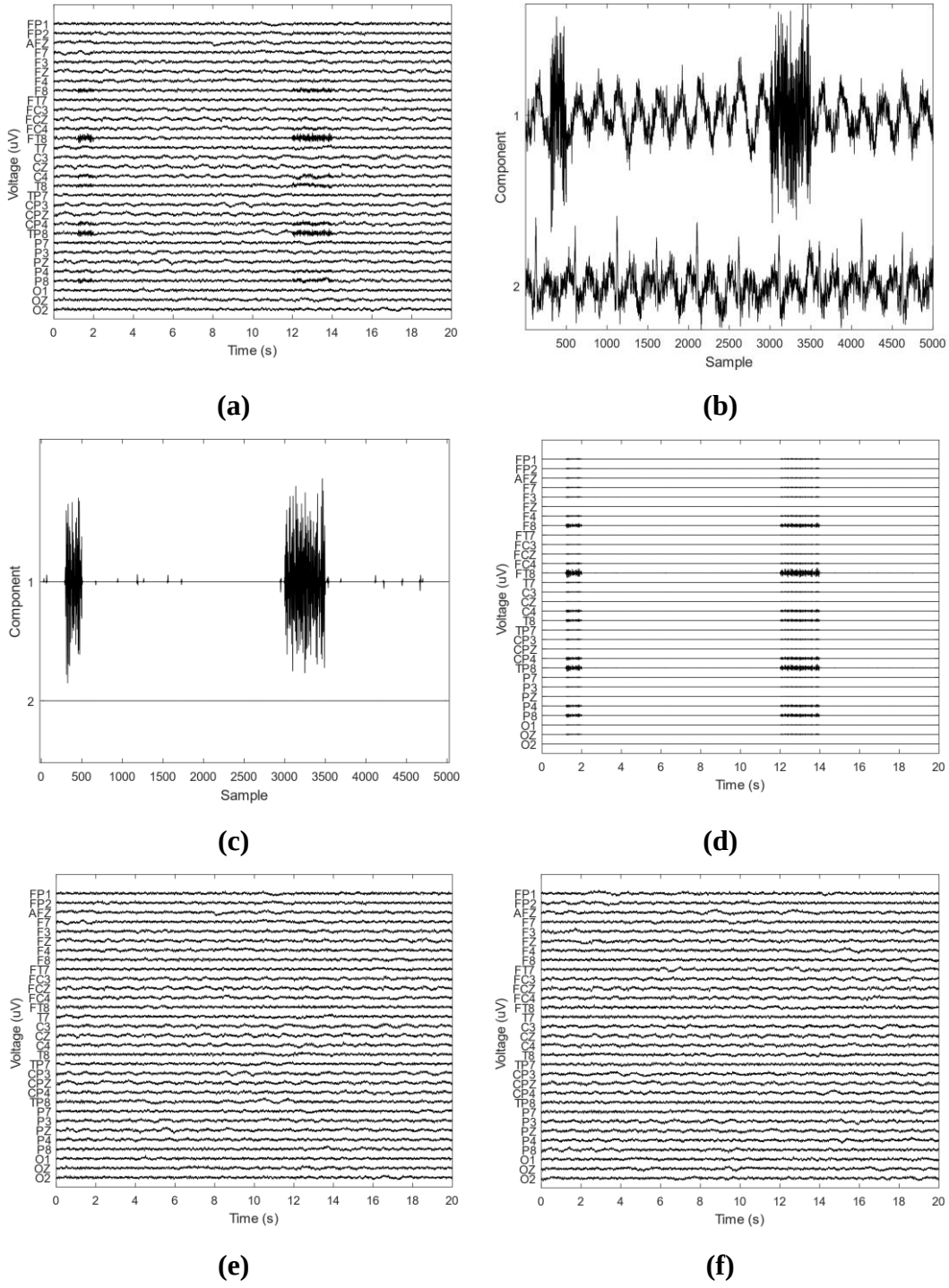


Figure 5.14: Signals at each stage of SCWEICA algorithm. **(a)** EAC-EEG. **(b)** Components from SCICA. Component 1 corresponds to the EMG artifact and component 2 corresponds to neural signals. **(c)** Wavelet denoising of component 1 and zeroing of component 2. **(d)** Spatially distributed EMG artifact. **(e)** Denoised EEG after subtraction of spatially distributed EMG artifact from EAC-EEG. **(f)** Ground truth EEG which appears to be indistinguishable from the denoised EEG.

5.2.1 Step-by-Step Overview of the SCWEICA Algorithm

Firstly, SCICA is applied to EAC-EEG (Figure 5.14a) to obtain the relevant artifact component corresponding to the simulated EMG artifact (Figure 5.14b). The need for two components is detailed in Section 5.2.2. SCICA uses the appropriate artifact topography depending on the side of the simulated EMG artifact, as described in Section 5.2.2. The EMG artifact component in Figure 5.14b contains neural signals that have leaked into it, as discussed in Section 3.2.3.

Secondly, wavelet denoising is required to remove the leaked EEG signals from the EMG artifact component to isolate the artifact, as described in Section 5.2.3. Figure 5.14c illustrates the output of wavelet denoising, where the output EMG artifact resembles the simulated EMG artifact before spatial distribution. The neural signals in the second component are zeroed.

The choice of WDPs affects the accuracy of EMG artifact isolation in Figure 5.14c. The WDPs must be optimised for optimal isolation of the EMG artifact from leaked neural signals in the artifact component, as detailed in Section 5.2.3. The optimisation of WDPs distinguishes O-SCWEICA in this study from SCWEICA, as developed by Akhtar *et al*. [14].

Finally, EEG reconstruction is performed to obtain cleaned EEG. This is described in Section 5.2.4. EEG reconstruction occurs by spatially distributing the isolated EMG artifact (Figure 5.14d) and subtracting it from EAC-EEG, using matrix subtraction (Figure 5.14e) [14]. The denoised EEG in Figure 5.14e should be similar to the ground truth, simulated EEG in Figure 5.14f.

5.2.2 Spatially Constrained ICA (SCICA)

SCICA was used in SCWEICA to efficiently isolate the simulated EMG artifact using prior, known, spatial information about the artifact topography. This section details spatially constraining the ICA algorithm, the chosen parameters for SCICA, and how the artifact topography was determined depending on the side of the simulated EMG artifact.

Choice of ICA Algorithm

According to Hesse and James [16], spatial constraints can be applied to a variety of ICA algorithms. This includes but is not limited to Infomax and FastICA. Spatially constrained FastICA (SCFastICA) was chosen for this study because spatial constraints have been successfully applied to this algorithm [15], [16]. Furthermore, FastICA is reliable, efficient and easy to use as described in Section 3.1.2. The SCFastICA algorithm was provided by Akhtar *et al.* [14].

While another ICA algorithm may be better suited for EMG artifact removal from EEG, such as SOBI or AMICA [9], this was out of the scope of this study. This possible improvement is addressed in the future recommendations in Section 7.3.

SCICA Parameters

SCICA was implemented with soft spatial constraints within an angular threshold of 1° [15], [16]. This was done as the initial A_c is expected to diverge, given the random nature of the simulations. The maximum number of iterations in estimating a component was set to 1000 and the stopping criterion for component estimation was set to 10^{-8} [14]. The SCFastICA algorithm was provided in an encrypted form, owing to intellectual property constraints [70]. However, from the description of the SCICA implementation in [14], the remainder of the parameters that were used were the default parameters for FastICA [71].

The SCICA algorithm, provided by Akhtar *et al.* [14], requires at least one constrained component and one unconstrained component. Therefore, as one artifact component was determined from the artifact topography in A_c (description to follow), SCICA generated two components. This implies that the mixing matrix produced by SCICA will be 31 rows by 2 columns. As shown in Equation 3.5, the first column is the constrained column (representative of the artifact topography) and the second column is the unconstrained column.

As the second component was not an artifact component, it was unnecessary and was therefore zeroed. Only the isolated EMG component was used to produce the spatially distributed EMG artifact shown in Figure 5.14d.

Side Based A_c Selection

As discussed in Section 5.1.2, a simulated EMG artifact can be generated for three cases:

1. Right-side unilateral temporalis muscle contraction
2. Left-side unilateral temporalis muscle contraction
3. Both left-side and right-side (bilateral) temporalis muscle contractions

Each case of artifact generation produced a different EMG artifact topography. As per Section 3.1.4, to obtain an artifact topography (A_c) for SCICA, unconstrained ICA must first be used. As such, the unconstrained FastICA algorithm was applied to a random EAC-EEG simulation for each case. The resulting components were inspected for the component that corresponded to the EMG artifact and leaked EEG signal. The corresponding column in the **A** matrix was then saved as the side-appropriate A_c. For example, if the seventh component contained the artifact, then the seventh column of the **A** matrix corresponded to the artifact topography. Each A_c was then used in the EMG artifact removal process, depending on the origin of artifact in the simulations.

5.2.3 Wavelet Denoising of the EMG Artifact Component

Wavelet denoising was used to isolate the EMG artifact component from neural signals that leaked into the component (Figure 5.14c). This section describes the WT that was selected, and briefly discusses the different parameters that affect the performance of wavelet denoising.

Choice of Wavelet Transform Algorithm

In this study, the stationary wavelet transform (SWT) was used for its superior denoising properties compared to the discrete wavelet transform (DWT) [14]. Refer to Section 3.2.3 for advantages of the SWT over the DWT.

Selection of WDPs

As detailed in Section 3.2.4 and 3.4, there are a number of parameters that may affect the wavelet denoising process. These include the mother wavelet, decomposition level, threshold determination method, and the thresholding type. The selection of these parameters distinguishes U-SCWEICA from O-SCWEICA and are elaborated on in Sections 5.3 and 5.4, respectively.

5.2.4 EEG Reconstruction

EEG reconstruction was performed by subtracting the spatially distributed EMG artifact from EAC-EEG, as shown in Figure 5.13. EEG reconstruction entails multiplying the mixing matrix (generated by SCICA) by the isolated EMG artifact (after wavelet denoising), resulting in a spatially distributed EMG artifact. Thereafter, the spatially distributed EMG artifact was subtracted (using matrix subtraction) from EAC-EEG to produce cleaned EEG.

5.3 Implementation of RA1 – Selection of Decomposition Level for U-SCWEICA

RA1 requires that SCWEICA be adapted for EMG artifact removal, thus producing U-SCWEICA. This was done by selecting the decomposition level that produced the best isolation of the EMG artifact from neural noise in the artifact component. The other WDPs were set according to Akhtar *et al*. [14], because the method in [14] allowed for leeway in the selection of the decomposition level to localise the artifact of interest. The remaining WDPs were specified as follows:

- Mother wavelet: Haar.
- Threshold determination method: Global.
- Thresholding type: soft.

To select the decomposition level, wavelet denoising was performed on the WDP Dataset, with the above WDPs, while the decomposition level was increased from one to ten. The accuracy of wavelet denoising was quantified in terms of average

normalised mutual information (NMI), absolute, maximum cross-correlation (AMCC), shift λ and normalised mean square error (NMSE). The accuracy measures are described in Section 5.3.1, and the methodology to implement RA1 is outlined in Section 5.3.2.

5.3.1 Accuracy Measures

NMI, AMCC, shift λ, and NMSE were used to evaluate the accuracy of estimation of the known, simulated, EMG artifact signal. This was compared to the EMG artifact component after wavelet denoising. The resulting denoised signal should be:

- High in both NMI and AMCC.
- Low in NMSE.
- Have a shift $\lambda = 0$ (such that translation invariance is ensured).

MATLAB does not provide an NMI function. As such, a reliable MATLAB function that can calculate NMI was desired. A “free-to-use” function was written by Dr Jason Palmer, which can also calculate mutual information (MI). The function is available at http://sccn.ucsd.edu/~jason/minfo.m, and has been used in peer-reviewed literature [14].

5.3.2 Methodology

For each set in the WDP Dataset, the artifact component was denoised with the Haar wavelet, Global threshold determination method, and soft thresholding, while the decomposition level was increased from one to ten. The denoising performance was quantified in terms of NMI, AMCC, shift λ, and NMSE. This process is summarised in Figure 5.15.

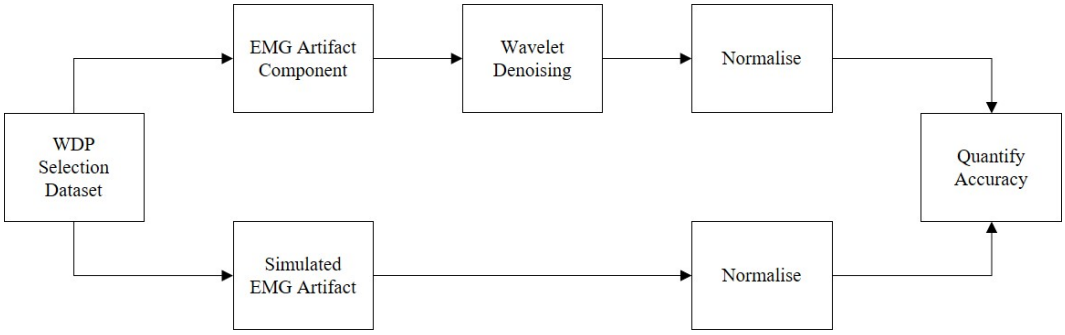


Figure 5.15: Block diagram of WDP selection methodology

For consistent comparison of the denoised EMG artifact component and ground truth simulated EMG signal, each signal was normalised with reference to its respective absolute maximum value. Normalisation was necessary as the amplitude of the simulated EMG signal was not necessarily comparable to that of the denoised EMG artifact component. This is because the spatial distribution of the EMG artifact across channels resulted in different amplitudes per channel.

The results for each measure were averaged over all 20 simulations in the dataset. Therefore, the average denoising representation for each decomposition level was obtained. The best performing decomposition level was then identified and selected.

5.4 Implementation of RA2 – Optimisation of WDPs for O-SCWEICA

As reviewed in Section 3.4.1, previous studies that applied wavelet denoising to artifact components [14], [19], [20] did not attempt to optimise the selection of WDPs. Generally, the choice of parameters in these studies was qualitative. As the aim of denoising is to achieve the best possible separation of signals, optimal WDPs are required to isolate the EMG signal in the EMG artifact component resulting from SCICA. The optimised WDPs were used in the wavelet denoising process in O-SCWEICA.

The various WDPs that were optimised for O-SCWEICA are presented in Section 5.4.1, and the optimisation methodology is detailed in Section 5.4.2. The optimisation methodology is an adaptation of that in Section 5.3.

5.4.1 Explored WDPs for Optimisation

The WDPs that were explored for optimisation include:

- Various discrete mother wavelets that were tested for use with the SWT, as summarised in Table 5.4. The selection of mother wavelets included higher order Daubechies [18] and Symlet [46] wavelets.
- Decomposition levels one to ten. By the tenth decomposition level, as per Figure 3.3, the signal has been decomposed to the level of $\frac{f_n}{1024}$ to $\frac{f_n}{512}$. Here, f_n is the sampling frequency, set to 250 Hz. This translates to 0.24 to 0.49 Hz – a vast over-estimation of the EMG artifact. In turn, the high decomposition level will likely result in poor denoising as the EMG artifact will be smeared in time [14]. Therefore, the optimal decomposition level will likely be much lower than ten.
- Five threshold determination methods offered by MATLAB were tested. These are Global, Heursure, Rigrsure, Sqtwolog, and Minimaxi. The methods are described in Section 3.2.4.
- The thresholding type can be either hard or soft.

The combination of these WDPs yielded 5000 individual combinations for the optimisation problem. The combinations were made from 50 mother wavelets, ten decomposition levels, five threshold determination methods, and two thresholding methods.

Table 5.4: Selection of discrete mother wavelets offered by MATLAB, used in the optimisation problem.

Wavelet Family	Mother Wavelets
Daubechies	haar, db2, db3, db4, db5, db6, db7, db8, db9, db10, db11, db12, db15, db20, db25, db30, db35, db40, db44, db45
Meyers	dmey
Biorthogonal	bior1.1, bior1.3, bior1.5, bior2.2, bior2.4, bior2.6, bior2.8, bior3.1, bior3.3, bior3.5, bior3.7, bior3.9, bior4.4, bior5.5, bior6.8
Symlet	sym2, sym3, sym4, sym5, sym6, sym7, sym8, sym9, sym10
Coiflet	coif1, coif2, coif3, coif4, coif5

5.4.2 WDP Optimisation Methodology

The methodology used to determine the optimal WDPs for O-SCWEICA is an adaptation of that used in Section 5.3. In this case, all WDPs (mother wavelet, decomposition level, threshold determination method, and threshold type) were optimised. Additionally, to ensure translation invariance, any WDP combination with an average shift not equal to zero was discarded.

If all three performance measures (NMI, AMCC, and NMSE) agreed on the optimal combination of WDPs, then the combination was used in O-SCWEICA. If the performance measures disagreed on the optimal combination of WDPs then the SCWEICA algorithm was used to determine which combination performed best. To determine the best WDP combination, SCWEICA was applied to the EAC-EEG signals from the WDP Dataset using each combination of WDPs. The performance was quantified in terms of AMCC, MI, and NMSE, as per Section 5.5.2. The combination of WDPs that produced the most accurate EMG artifact removal was then used in O-SCWEICA.

5.5 Implementation of RA3 – Experimental Setup for Performance Comparison of EMG Artifact Removal Algorithms

This section refers to the "Performance Comparison" block in Figure 5.1. The experimental setup for the testing of O-SCWEICA compared to U-SCWEICA and STD-ICA for EMG artifact removal from EAC-EEG is presented.

Firstly, STD-ICA is adapted for automated testing in Section 5.5.1. Thereafter, the performance measures, in terms of accuracy and efficiency, and testing methodology are described in Section 5.5.2. The hardware and software that was used in the study are detailed in Section 5.5.3.

5.5.1 Adaptation of Standard ICA Denoising Algorithm for Automated Denoising and Testing

The implementation of STD-ICA differed from literature because SCFastICA was used as opposed to unconstrained FastICA. SCFastICA was run for 31 components to mimic unconstrained FastICA, ensuring that the EMG artifact would always be in the first component, so that the denoising process could be automated. Otherwise, visual inspection and user input would be required, which could lead to potential human error in selecting the appropriate component.

5.5.2 Testing of EMG Artifact Removal Algorithms

This section describes Tests 1 to 7 in Table 5.5, where the EMG artifact removal performance of O-SCWEICA is compared to U-SCWEICA and STD-ICA. The datasets in Table 5.5 are described in Section 5.1.4. Chapter 6 presents and discusses the results of the tests.

Table 5.5: Summary of tests

Test Number	Description	Dataset
1	Referential comparison for all other tests in this study	Reference
2	Effect of MATLAB's optimisation on EMG artifact removal for each algorithm	MATLAB Optimisation
3	Literature comparison to Akhtar *et al*. [14]	Akhtar Comparison
4	RT EMG artifact removal viability assessment of O-SCWEICA	Towards RT
5	Viability of O-SCWEICA for EMG artifact removal from long duration EEG recordings	Offline
6	Effect of amplitude-varying unilateral left-side EMG artifact on the performance of O-SCWEICA	Left-Side
7	Effect of amplitude-varying bilateral EMG artifacts on the performance of O-SCWEICA	Both-Sides

Accuracy Measures

AMCC, MI, and NMSE were used to assess the accuracy of the three EMG artifact removal algorithms. Each of these measures quantified the similarity of the 'cleaned' EEG to the ground truth EEG. The measures were averaged across the simulations in each dataset to assess the average performance of each algorithm.

Better EMG artifact removal is manifested in larger MI and AMCC (the latter should tend to 1), and smaller NMSE (0% for the best performance). NMSE and MI were used by Akhtar *et al*. [14] allowing for fair comparison between these studies.

AMCC, MI, and NMSE were quantified in two ways. The first was the average across all channels (AV), where the given measure was quantified for each channel and then averaged. The second was the worst-case (WC) measure. The WC was calculated for the channel with the greatest EMG weight and therefore the largest EMG artifact amplitude. This distinction was made as the AV results can be misleading in describing the success of overall EMG artifact removal as, depending on the EMG simulation, many channels will have little interference. After the denoising process, the AV results can skew the perceived performance.

If a unilateral right-side temporalis muscle contraction was simulated, the WC channel is channel FT8. However, if a unilateral left-side temporalis muscle contraction was simulated, the WC channel is FT7. For the simulation of bilateral temporalis muscle contractions, both channels FT7 and FT8 were used for the WC measure, to ensure consistent comparison.

Efficiency Measure

Efficiency of the three artifact removal algorithms was assessed based on the duration of each algorithm. The average denoising duration was calculated per dataset, per algorithm. The number of floating-point operations for each algorithm should also be computed but this was out of the scope of this study. Instead, it is presented as future work in Section 7.3.

Testing

Each of the denoising algorithms (STD-ICA, U-SCWEICA, and O-SCWEICA) was applied to each of the datasets in the Comparative, Duration-Varying, and Artifact Topography Datasets, detailed in Table 5.5. Each denoised simulation within each dataset was quantified in terms of AMCC, MI, and NMSE (for both AV and WC), and compared to the ground truth EEG. The denoising duration was also noted. The testing process is shown in Figure 5.16. The average of each performance measure for each algorithm in each dataset was determined.

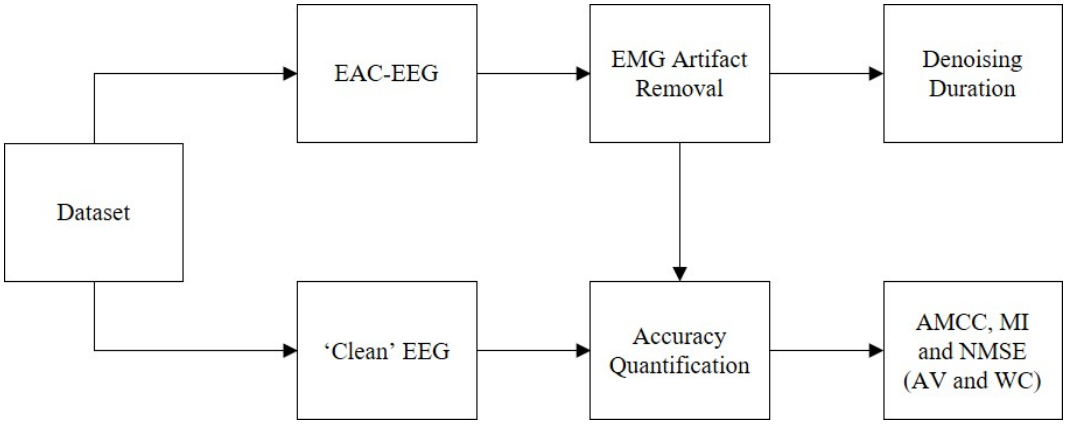


Figure 5.16: EMG artifact removal and performance quantification

5.5.3 Hardware and Software

The hardware used in this study was a Windows 10 laptop with an Intel Core i5-7200U 2.5 GHz processor, 8 GB DDR4 RAM, and 1 TB hard drive. The simulations and denoising algorithms were implemented in MATLAB R2015b.

5.6 Chapter Summary

The simulations and datasets used in the study were described. Clean, ground truth 31-channel EEG and EMG artifacts were simulated and added to form EAC-EEG. 26 datasets of these signals were generated. One dataset was used to adapt the SCWEICA algorithm for EMG artifact removal. The remaining datasets varied in simulation duration and artifact topographies which were used for performance assessment of O-SCWEICA against U-SCWEICA and STD-ICA.

The SCWEICA algorithm implementation was outlined in terms of the various components of the algorithm: SCICA, wavelet denoising and EEG reconstruction. SCICA was applied to EAC-EEG to determine the EMG artifact component and mixing matrix. Wavelet denoising was used to remove the neural signals that leaked into the EMG artifact component. The mixing matrix was then multiplied by the EMG artifact component to spatially distribute the artifact across the electrodes. Thereafter, EEG reconstruction was performed by subtracting the spatially distributed EMG artifact from EAC-EEG to obtain cleaned EEG.

Two variants of the SCWEICA algorithm were produced: U-SCWEICA and O-SCWEICA. The best decomposition level was selected for U-SCWEICA to adapt the SCWEICA algorithm for EMG artifact removal. The WDPs were optimised in O-SCWEICA to best isolate the EMG signal from the neural signals in the EMG artifact component. This should allow for more accurate EMG artifact removal compared to the other algorithms explored in the study.

A testing paradigm was detailed for the testing of O-SCWEICA against STD-ICA and U-SCWEICA. The average performance of the algorithms was quantified in terms of accuracy (AMCC, MI, and NMSE) and efficiency (denoising duration) when removing EMG artifacts from various EAC-EEG datasets.

Chapter 6 – Results and Discussion

This chapter highlights the trends observed in the results, as well as noteworthy exceptions, main observations, and comparisons to literature. This chapter is structured to address each of the RAs, which together answer the RQ.

RA1, RA2, and RA3 are addressed in Sections 6.1, 6.2, and 6.3, respectively. In Section 6.1, wavelet denoising parameters (WDPs) were selected to adapt SCWEICA for EMG artifact removal, producing U-SCWEICA. In Section 6.2, the optimal WDPs for EMG artifact removal were chosen for SCWEICA, resulting in O-SCWEICA. The EMG artifact removal performance of O-SCWEICA, U-SCWEICA, and STD-ICA were compared in Section 6.3. The RQ is answered in Section 6.4.

6.1 RA1 – Decomposition Level Selection for U-SCWEICA

As detailed in Section 5.3, the best decomposition level was selected for U-SCWEICA while all other WDPs were fixed according to Akhtar *et al*. [14]. The WDPs were the Haar wavelet, Global threshold determination method, and soft thresholding [14].

Table 6.1 shows the decomposition level selection results using the pre-selected WDPs for U-SCWEICA. The results indicate that the third decomposition level performed best in terms of NMI, AMCC, shift λ, and NMSE. The selection of the third decomposition level was expected. The bandwidth of the simulated EMG artifact is 20 – 60 Hz and for 250 Hz sampling frequency (from Figure 3.3), the third decomposition level encompasses 31.3 – 62.5 Hz. Therefore, the WDPs used in U-SCWEICA are:

- Mother wavelet: Haar.
- Decomposition level: three.
- Threshold determination method: Global.
- Thresholding type: soft.

Table 6.1: Average NMI, AMCC, shift λ, and NMSE for wavelet decomposition levels one to ten for the combination of the Haar wavelet, Global threshold determination method and soft thresholding

Decomposition Level	NMI	AMCC	Shift λ	NMSE (%)
1	0.275	0.557	0.20	92.9
2	0.335	0.718	0.05	49.2
3	0.359	0.835	0	31.3
4	0.260	0.799	0	44.8
5	0.188	0.743	0	75.8
6	0.136	0.705	0	83.1
7	0.110	0.671	0	88.6
8	0.099	0.644	0	92.4
9	0.097	0.636	0	94.0
10	0.099	0.638	0	93.0

6.2 RA2 – WDP Optimisation for O-SCWEICA

As described in Section 5.4.1, 5000 combinations of WDPs were used to denoise each of the EMG artifact components in the WDP Dataset. Thereafter, the results were averaged over the simulations in the dataset to determine the average performance of each WDP combination. Combinations were discarded if the average shift λ was not equal to zero to ensure translation invariance. The average results of the 20 best performing WDP combinations are shown for NMI, AMCC, and NMSE in Tables 6.2, 6.3, and 6.4, respectively.

Table 6.2: Best 20 WDP combinations in terms of average NMI

Mother Wavelet	Threshold Determination Method	Threshold Type	Decomposition Level	NMI
sym9	Global	soft	3	0.437
sym4	Global	soft	2	0.432
sym5	Global	soft	3	0.429
sym6	Global	soft	2	0.424
sym4	Global	hard	2	0.419
sym8	Global	soft	2	0.416
sym6	Global	hard	2	0.413
sym10	Global	soft	2	0.411
sym8	Global	hard	2	0.408
sym10	Global	hard	2	0.404
sym9	Global	hard	3	0.394
dmey	Global	soft	3	0.387
sym5	Global	hard	3	0.384
bior6.8	Global	soft	3	0.377
bior4.4	Global	soft	3	0.376
coif1	Global	soft	3	0.365
bior1.1	Global	soft	3	0.359
haar	Global	soft	3	0.359
bior1.3	Global	soft	3	0.358
bior1.5	Global	soft	3	0.357

Table 6.3: Best 20 WDP combinations in terms of average AMCC

Mother Wavelet	Threshold Determination Method	Threshold Type	Decomposition Level	AMCC
sym9	Global	hard	3	0.921
sym5	Global	hard	3	0.914
sym9	Global	soft	3	0.896
sym9	Sqtwolog	soft	3	0.893
sym5	Global	soft	3	0.890
sym5	Sqtwolog	soft	3	0.889
sym4	Global	hard	2	0.889
bior3.9	Global	soft	4	0.883
sym6	Global	hard	2	0.883
bior3.7	Global	soft	4	0.881
sym4	Sqtwolog	soft	2	0.877
bior3.5	Global	soft	4	0.877
sym8	Global	hard	2	0.876
db35	Global	soft	4	0.874
bior3.9	Global	hard	4	0.874
db30	Global	soft	4	0.874
bior3.7	Global	hard	4	0.873
db40	Global	soft	4	0.873
sym9	Minimaxi	soft	3	0.873
db25	Global	soft	4	0.873

Table 6.4: Best 20 WDP combinations in terms of average NMSE

Mother Wavelet	**Threshold Determination Method**	**Threshold Type**	**Decomposition Level**	**NMSE (%)**
sym9	Global	hard	3	15.225
sym5	Global	hard	3	16.484
sym9	Global	soft	3	21.322
sym5	Global	soft	3	22.733
sym9	Sqtwolog	soft	3	23.180
sym5	Sqtwolog	soft	3	24.124
bior3.9	Global	soft	4	24.147
bior3.7	Global	soft	4	24.578
haar	Global	hard	3	25.177
bior1.1	Global	hard	3	25.177
bior3.5	Global	soft	4	25.199
bior1.3	Global	hard	3	25.637
bior1.5	Global	hard	3	25.806
bior3.3	Global	soft	4	26.845
bior1.5	Global	soft	3	27.724
bior2.8	Global	hard	3	28.571
coif1	Global	hard	3	28.800
bior1.3	Global	soft	3	28.868
bior2.4	Global	hard	3	28.906
bior2.6	Global	hard	3	29.096

6.2.1 Determination of Optimal WDPs for O-SCWEICA

For all performance measures in Tables 6.2, 6.3, and 6.4, the best mother wavelet was Symlet order 9, using the Global threshold determination method at the third decomposition level. The selection of Symlet order 9 corresponds with Al-Qazzaz *et al.* [46]. The results differed in the use of the soft or hard threshold type. Hard thresholding performed best in terms of AMCC and NMSE, while soft thresholding performed best in terms of NMI.

To address the disagreement, the aforementioned WDPs were used in SCWEICA to clean the corresponding EAC-EEG signals from the WDP Dataset, as detailed in Section 5.4.2. The cleaned EEG was compared to the corresponding 'clean', ground truth EEG in terms of MI, AMCC, and NMSE. The average results are shown in Table 6.5 for the average across all channels (AV) and worst case (WC) results – the channel with the largest EMG artifact. WC refers to channel FT8 as the EMG contractions in the WDP Dataset were simulated for the right-side temporalis muscle.

Table 6.5: A comparison of the average hard and soft thresholding in denoising EAC-EEG in terms of the AV and WC MI, AMCC and NMSE

Performance Measures	**Threshold Type**			
	Hard		**Soft**	
	AV	**WC**	**AV**	**WC**
MI	2.243	1.721	2.236	1.451
AMCC	0.997	0.969	0.993	0.915
NMSE (%)	0.669	6.40	1.58	19.7

Table 6.6: Absolute and relative difference between AV and WC results for hard and soft thresholding with reference to hard thresholding

Performance Measures	Absolute Difference		Relative Difference (%)	
	AV	WC	AV	WC
MI	0.007	0.27	0.31	15.7
AMCC	0.004	0.054	0.40	5.57
NMSE (%)	-0.911	-13.3	-136	-207

Table 6.6 indicates that the absolute and relative differences (see Equation 3.21 in Section 3.6.5) in the thresholding methods are close to negligible. However, hard thresholding yields better AV and WC results in all metrics compared to soft thresholding. Therefore, the optimal WDPs used in O-SCWEICA are:

- Mother wavelet: Symlet Order 9 (sym9).
- Decomposition level: three.
- Threshold determination method: Global.
  Thresholding type: hard.

6.3 RA3 – Performance Comparison of EMG Artifact Removal Algorithms

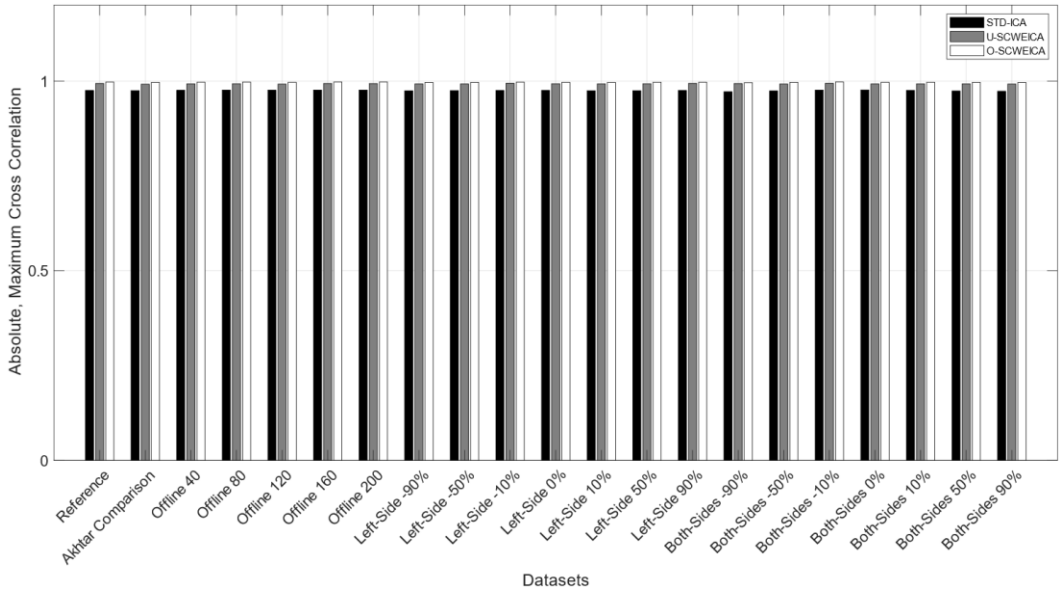
This section compares the EMG artifact removal performance of O-SCWEICA, U-SCWEICA, and STD-ICA. The results of the tests in Table 5.5 are presented and discussed. All results were compared to the denoising results of the Reference Dataset. Henceforth, these results are referred to as the Reference Results.

6.3.1 Performance Comparison Results

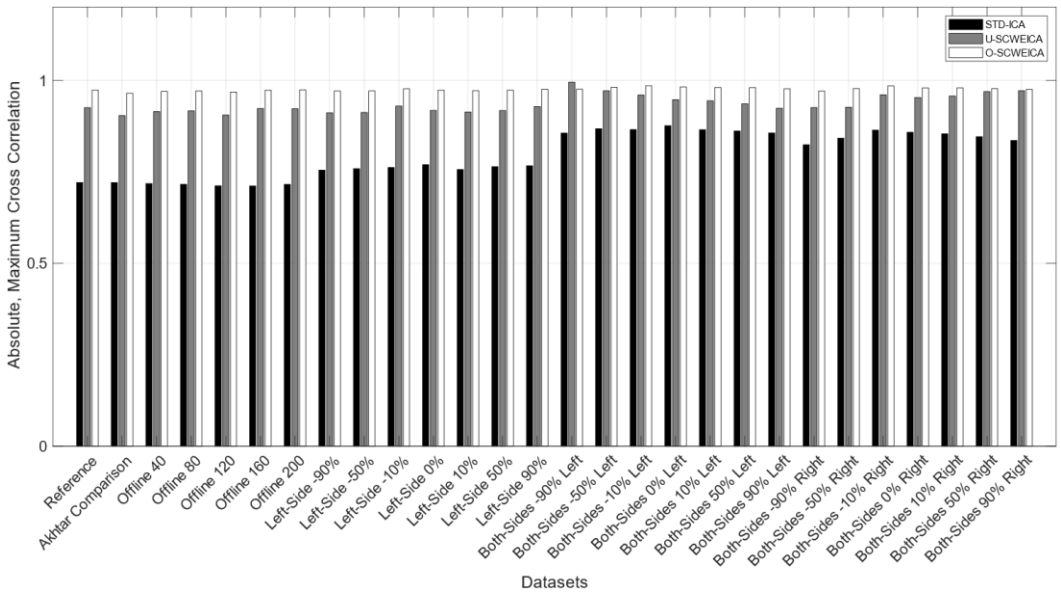
The denoising results in terms of AMCC, MI, NMSE, and denoising duration are shown in Figures 6.1, 6.2, 6.3 and 6.4 respectively. Note that:

- The Both-Sides Dataset produced results for the left WC channel (FT7) and the right WC channel (FT8). The other datasets were characterised for either the right or left WC channel.
- The Towards RT results (Test 4 in Table 5.5) were omitted from the results figures. Compared to the Reference Results, the Towards RT denoising accuracies were poor and the efficiencies were not greatly improved.
- The MATLAB Optimisation results (Test 2 in Table 5.5) were omitted from the accuracy results as they were identical to the Reference Results. This is because the respective datasets were identical.

The average relative differences between O-SCWEICA and U-SCWEICA for denoising accuracy and efficiency are shown in Tables 6.7 and 6.8, respectively. These calculations were taken with reference to the O-SCWEICA results (see Equation 3.21 in Section 3.6.5). STD-ICA was not compared to the other algorithms as its performance was vastly inferior across all metrics.

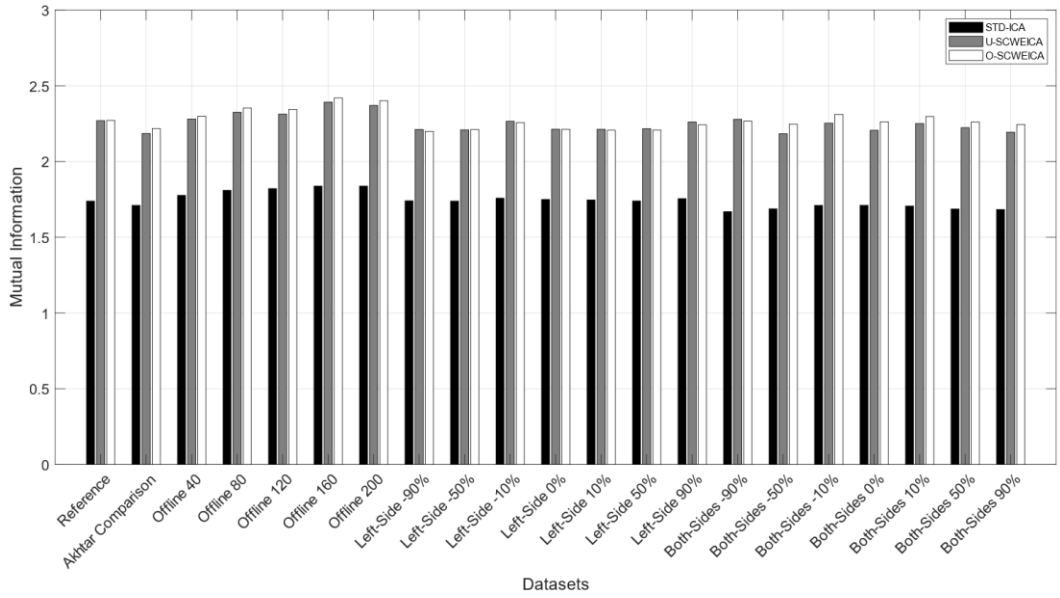


(a)

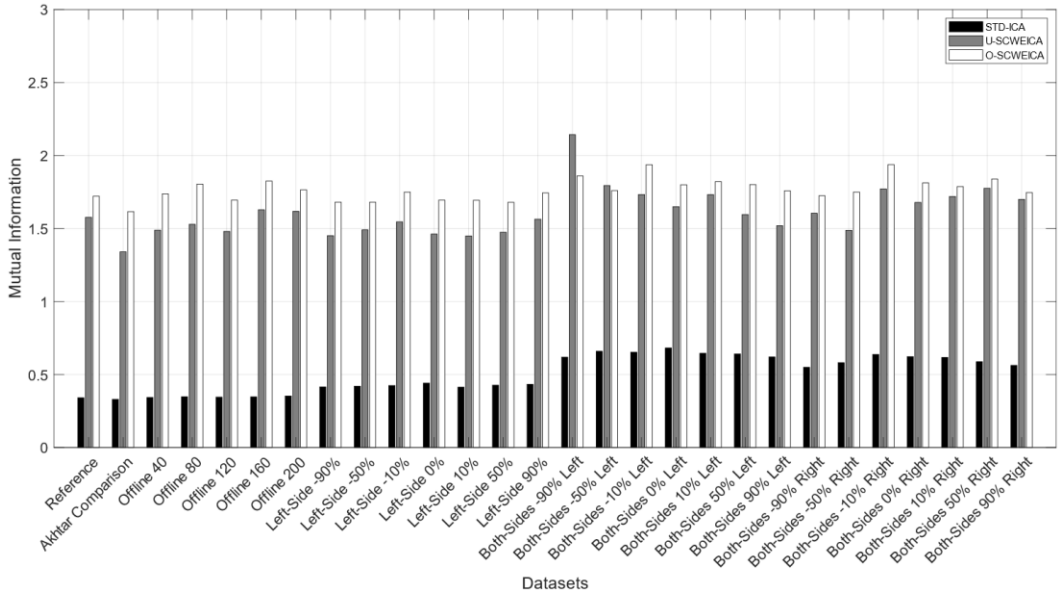


(b)

Figure 6.1: Average AMCC results for the testing of O-SCWEICA for the **(a)** AV and **(b)** WC cases. Higher AMCC implies better denoising.

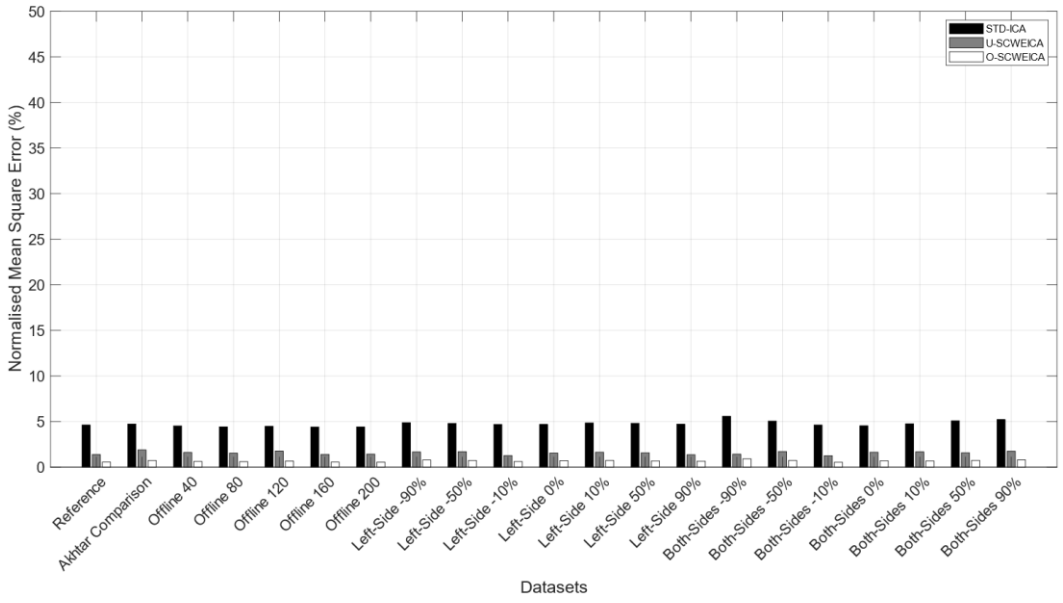


(a)

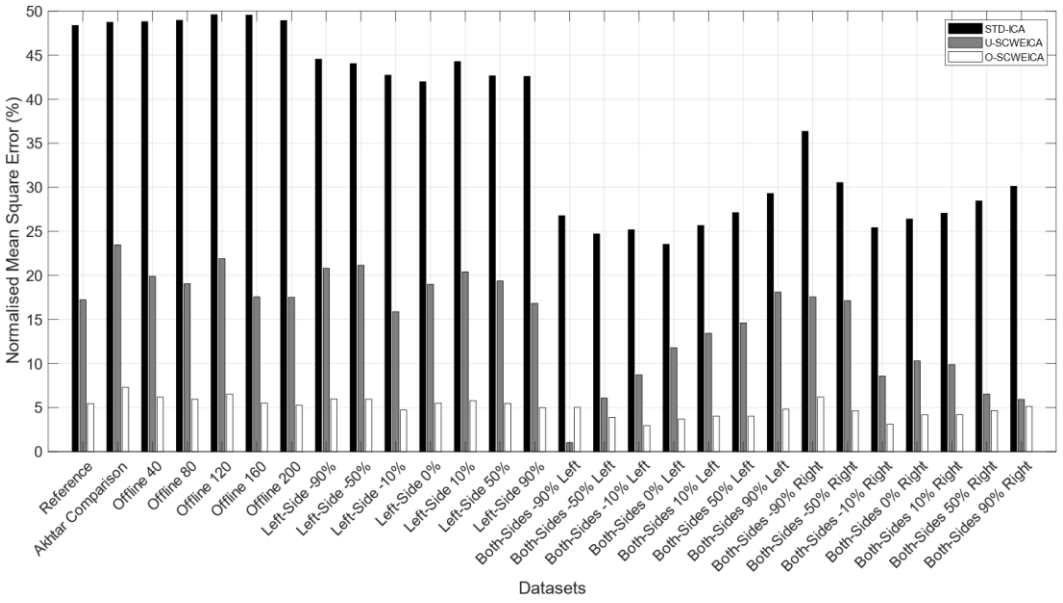


(b)

Figure 6.2: Average MI results for the testing of O-SCWEICA for the **(a)** AV and **(b)** WC cases. Higher MI implies better denoising.

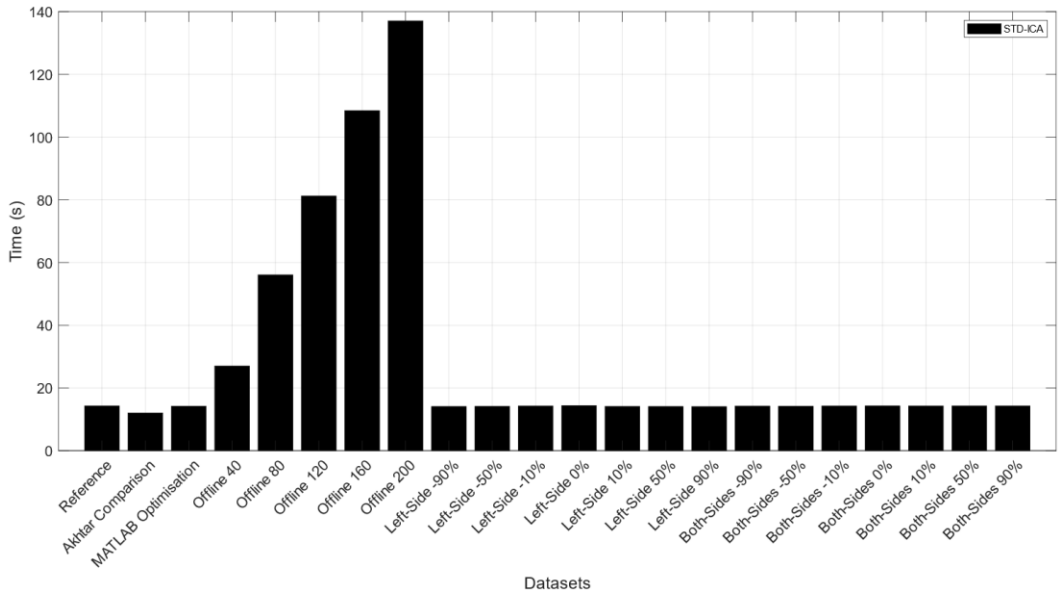


(a)

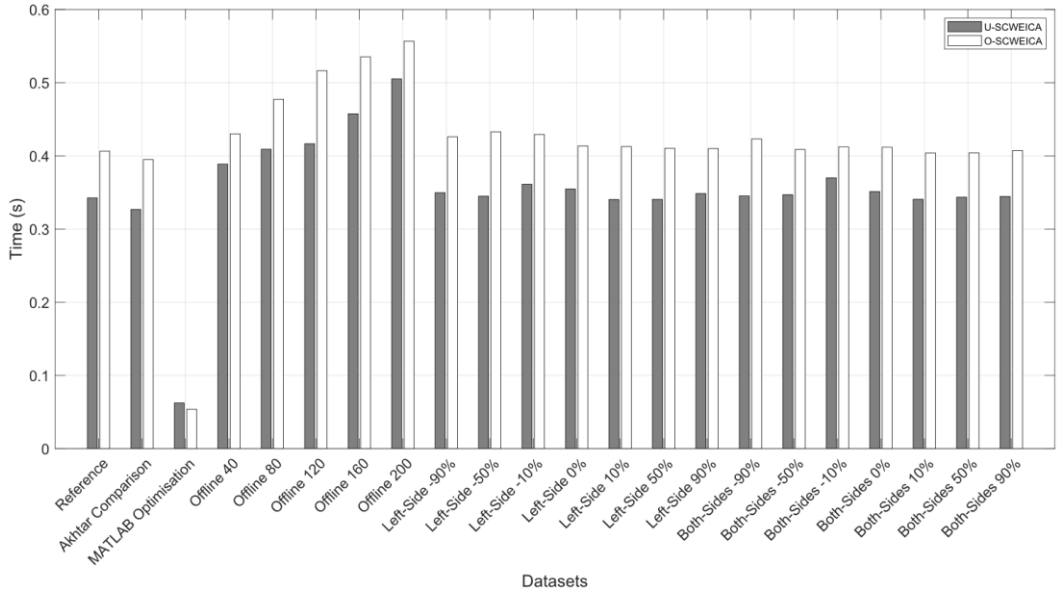


(b)

Figure 6.3: Average NMSE results for the testing of O-SCWEICA for the **(a)** AV and **(b)** WC cases. Lower NMSE implies better denoising.



(a)



(b)

Figure 6.4: Average denoising duration results for **(a)** STD-ICA and **(b)** U-SCWEICA and O-SCWEICA. The results are shown separately owing to the vast differences in magnitude – they would not be adequately displayed in the same figure.

Table 6.7: Average relative differences in accuracy between O-SCWEICA and U-SCWEICA. Positive AMCC and MI results imply better O-SCWEICA accuracy. Negative NMSE results imply better O-SCWEICA accuracy.

Test	Relative Difference (%)					
	AMCC		MI		NMSE (%)	
	AV	WC	AV	WC	AV	WC
1	0.36	4.92	0.09	8.40	-145	-216
3	0.49	6.35	1.46	17.1	-158	-222
5	0.42	5.65	1.14	12.3	-158	-226
6	0.36	5.61	-0.34	12.5	-118	-247
7: Right Side	0.36	2.63	1.89	6.84	-119	-138
7: Left Side	-	2.70	-	4.51	-	-166

Table 6.8: Average absolute and relative differences in efficiency between O-SCWEICA and U-SCWEICA. Negative values imply better O-SCWEICA efficiency.

Test	Absolute Difference in Duration (s)	Relative Difference in Duration (%)
1	0.064	15.68
2	-0.009	-16.09
3	0.068	17.24
5	0.068	13.41
6	0.071	16.83
7	0.061	14.93

6.3.2 Performance Comparison Discussion

The denoising results discussion is divided into general trends, notable exceptions and main observations. This is followed by comparisons to literature. Together, the results and discussion assisted in answering the RQ in Section 6.4.

General Trends

For all results, across all performance measures, STD-ICA performed worse than U-SCWEICA and O-SCWEICA. This corresponds with Akhtar *et al*. [14] where STD-ICA performed worse than SCWEICA for EOG artifact removal. Hence, the focus of the discussion is on differences in performance between U-SCWEICA and O-SCWEICA.

For the majority of results, O-SCWEICA was more accurate than U-SCWEICA, likely owing to the optimisation of WDPs. This is clearly seen in Table 6.7, with few exceptions. U-SCWEICA performed slightly more efficiently than O-SCWEICA, except for Test 2 in Table 5.5 where in which O-SCWEICA was 0.009 s faster This is likely owing to the fast convergence of the Haar wavelet [14]. On average across all datasets, there was 0.06 s difference in denoising duration between the two algorithms.

From the general trends seen in Section 6.3.1 it is clear that O-SCWEICA is preferable to U-SCWEICA and STD-ICA for more accurate EMG artifact removal. The small loss of efficiency by O-SCWEICA compared to U-SCWEICA is tolerable.

Notable Exceptions

U-SCWEICA was more accurate than O-SCWEICA in two cases:

1. The Left Side AV MI results (Test 6 in Table 5.5) in Table 6.6 indicate negligible better accuracy for U-SCWEICA compared to O-SCWEICA. It is unclear why this discrepancy occurred.
2. U-SCWEICA performed more accurately than O-SCWEICA in denoising the Both-Sides -90% Dataset for all the left WC performance measures. It

is hypothesised that the low amplitude EMG artifact in this channel may be of similar amplitude to the neural signals. Therefore, threshold determination may not sufficiently distinguish the EMG artifact from neural signals. Applying soft thresholding to the remaining wavelet coefficients may lead to less subtraction of neural signals than hard thresholding in EEG reconstruction. In turn, greater accuracy may be possible with soft thresholding.

Main Observations

This section details the main observations from the denoising results and the implications that they have for the use of O-SCWEICA in the removal of EMG artifacts from EEG. The observations are described from major to minor importance and assist in answering the RQ.

Offline Denoising Viability

Test 5 in Table 5.5 showed that O-SCWEICA is more accurate than U-SCWEICA in EMG artifact removal from long duration EAC-EEG signals. The Offline Dataset denoising results and Reference Results are similar for AMCC and NSME. Except for Offline Dataset 120, all Offline Datasets performed better than the Reference Results in terms of MI.

Compared to the Reference Results, artifact removal from the Offline Datasets takes between 0.02 and 0.15 s longer, as expected with more samples. However, the increase in denoising duration is small compared to the vast increase in EEG signal duration. When the longest dataset (Offline Dataset 200) was processed with O-SCWEICA, the average denoising duration was 0.56 s and the accuracy was generally better than the other Offline datasets. For Offline Dataset 200, the AV AMCC, MI, and NMSE results were 0.997, 2.40, and 0.54%, respectively. The WC results were 0.974, 1.76, and 5.26% for AMCC, MI, and NMSE, respectively. These results are indicative of satisfactory EMG artifact removal that closely resembles the ‘clean’, ground truth EEG.

The high accuracies imply that even with an increased number of samples per channel, when using O-SCWEICA to remove EMG artifacts from EAC-EEG:

- Accuracy is maintained and potentially improved.
- Denoising duration decreases slightly but is minimal compared to the increase in EEG recording duration.

Therefore, O-SCWEICA can be used for more accurate and efficient removal of EMG artifacts from large EAC-EEG recordings compared to U-SCWEICA and STD-ICA.

Adaptation to Changes in Artifact Topography

Test 6 in Table 5.5 showed that O-SCWEICA performs similarly to the Reference Results for unilateral EMG artifacts. Therefore, whether the EMG artifact was generated from the left-side or right-side, regardless of changes in amplitude, there was minimal effect on the accuracy or efficiency of O-SCWEICA.

Test 7 in Table 5.5 showed generally improved accuracy and maintained efficiency of O-SCWEICA when applied to the Both-Sides Datasets. Compared to the WC Reference Results, all three algorithms performed better in processing the majority of the Both-Sides Datasets. However, the AV performance was approximately the same for each of the three algorithms compared to the Reference Results. It is unclear why this is the case but may occur owing to the averaging and smoothing out of notable phenomena.

There is an observable trend that for U-SCWEICA in all WC accuracy measures but is most notable for NMSE in Figure 6.3b. For the left WC, as the variation in left-side EMG weight increases from -90% to 90%, the accuracy of U-SCWEICA decreases. For the right WC, as the variation in left-side EMG weight increases from -90% to 90%, the accuracy tends to increase.

As discussed, the observed trend for the left WC may be related to the improved performance of soft thresholding for lower amplitude EMG artifacts relative to the

EEG signals. Therefore, as the EMG weight and EMG artifact increase in amplitude, the artifact removal accuracy decreases. It is unclear why the observed trend in the right WC occurs. It was expected that the right WC accuracy results remain relatively constant because the right-side EMG weights were not varied in the simulations for this dataset. Therefore, the results imply that for bilateral EMG artifacts, U-SCWEICA achieves better denoising on one side of the head, but not on the other. As U-SCWEICA is not the focus of this study, further explanations were not pursued.

For Tests 6 and 7 in Table 5.5, O-SCWEICA was relatively stable. For the AV accuracy results, O-SCWEICA had standard deviations of 0.0004, 0.034, and 0.089% for AMCC, MI, and NMSE, respectively. For the WC accuracy results, O-SCWEICA had standard deviations of 0.004, 0.076, and 0.920% for AMCC, MI, and NMSE respectively. Additionally, from Table 6.7, O-SCWEICA performs better on average than U-SCWEICA for all WC accuracy measures. Therefore, O-SCWEICA is the best performing algorithm regardless of changes in the relative EMG amplitudes of the left-side and right-side temporalis muscles.

Suitability for RT EMG Artifact Removal

Test 4 in Table 5.5 showed that O-SCWEICA is unsuitable for RT artifact removal – the accuracies were low for all Towards RT Datasets. While all accuracy measures improved minimally with increased window size, the results were far worse than the Reference Results. For example, the O-SCWEICA AV MI results for the Towards RT 1 Dataset were 46.8% worse than the Reference Results.

The different window sizes in the Towards RT Datasets did not affect the denoising duration. The simulations in the Reference Dataset encompassed 20 s and were processed, on average, in 0.41 s. Comparatively, a 1 s window in the Towards RT 1 Dataset was processed, on average, in 0.39 s.

None of the Towards RT datasets allow for a good representation of expected RT sampling and processing because all windows were 1 s or greater. If a smaller

window size is required for RT processing, it is assumed that the accuracy of O-SCWEICA would be further reduced as less samples are available for the algorithm. Therefore, O-SCWEICA is not viable for RT EMG artifact removal in the methodology employed in this study.

Efficiency Improvement

Test 2 in Table 5.5 showed that the denoising duration of O-SCWEICA – when applied to the MATLAB Optimisation Dataset – was 0.37 s shorter (on average) than the other datasets. This is illustrated in Figure 6.4b.

The MATLAB efficiency option has a drawback: the denoising durations are not necessarily consistent across iterations. Figure 6.5 depicts this drawback. The first O-SCWEICA iteration takes approximately as long as the average denoising duration for the Reference Results with the efficiency option disabled. Thereafter, the optimisation clearly reduces the denoising duration.

The use of the efficiency option did not seem to affect the STD-ICA algorithm as much as the SCWEICA algorithms, as seen in Figure 6.4a. On average, with the efficiency option enabled, STD-ICA took 14.2 s compared to 14.3 s for the Reference Results, with the efficiency option disabled.

For the shortest denoising durations when using O-SCWEICA in MATLAB, the efficiency option should be enabled. However, the algorithm may need to be run multiple times to see an improvement in efficiency.

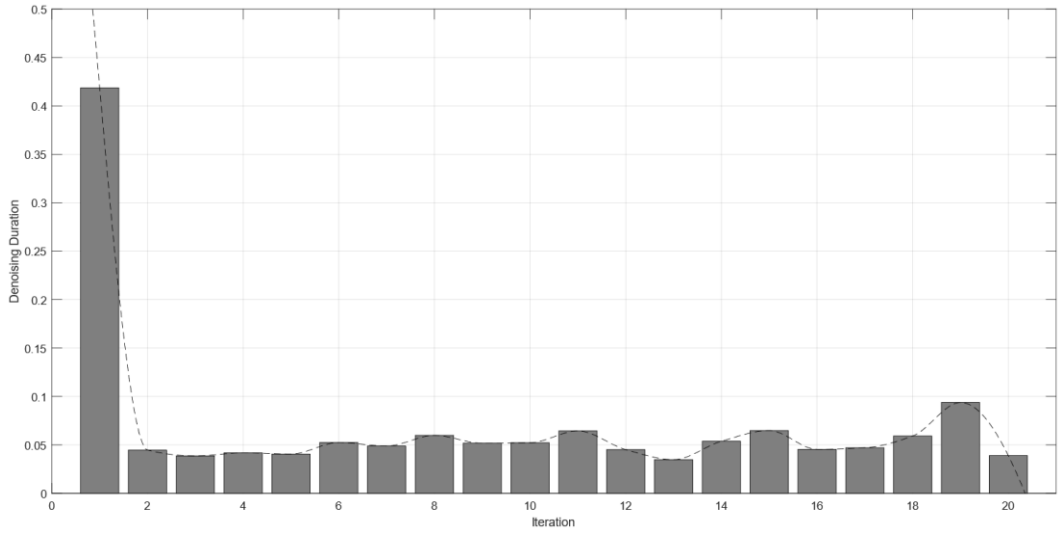


Figure 6.5: O-SCWEICA denoising duration over multiple iterations with MATLAB optimisation enabled

Comparison to Literature

Test 3 in Table 5.5 allows for comparison to literature. The Akhtar Comparison Dataset was processed with O-SCWEICA and the results were compared to those of Akhtar *et al.* [14]. Comparisons were made to the average MI and NMSE results in their study for both the simulations and human subjects. The comparison between these studies (using the AV results of the Akhtar Comparison Dataset) is shown in Table 6.9.

The improvement in accuracy is likely attributed to the optimisation of the WDPs, as detailed in Section 6.2. The improvement in efficiency is likely owing to the better hardware and software used in the current study, which was conducted six years later than Akhtar *et al.* [14].

There are differences between these studies and as such, the comparison was not comprehensive. Akhtar *et al*. [14] used six electrodes in the simulations and 19 electrodes for human subjects, while this study used 31 electrodes for simulations. AMCC was not used by Akhtar *et al*. [14] and hence, could not be used for comparison. Additionally, the implementation of SCWEICA by Akhtar *et al*. [14]

aimed to remove EOG artifacts while this study aimed to remove EMG artifacts. The differences in these types of artifacts necessitated different WDPs.

Table 6.9: Comparison between literature and the AV denoising performance of O-SCWEICA when denoising the Akhtar Comparison Dataset

	SCWEICA Implementation by Akhtar *et al*. [14]			**O-SCWEICA**
	Simulations	**Subject 1**	**Subject 2**	
NMSE (%)	9.02	1.72	2.71	0.73
MI	1.34	1.77	1.66	2.22
Duration (s)	-	1.07	1.02	0.40

6.4 Answering the Research Question

The RQ in Section 4.1 is: to what extent can spatially constrained, optimised, wavelet enhanced FastICA remove EMG artifacts from EEG, both accurately and efficiently in time, for later real-time adaptation?

RA1 was completed in Section 6.1. SCWEICA was adapted for EMG artifact removal by selecting the third decomposition level, while all other WDPs were set according to Akhtar *et al*. [14]. This produced U-SCWEICA.

RA2 was completed in Section 6.2. 5000 combinations of WDPs were optimised for SCWEICA to remove EMG artifacts from EAC-EEG. The resulting, optimal WDPs were: Symlet order 9 (sym9) mother wavelet, decomposition level three, Global threshold determination method and hard thresholding. This resulted in O-SCWEICA.

RA3 was completed in Section 6.3. The results indicate that for the majority of tests, regardless of the artifact topography or duration of EEG recording, O-SCWEICA outperforms all other EMG artifact removal algorithm in this study. However, O-SCWEICA is unsuitable for RT EMG artifact removal in the

methodology employed in this study. Nevertheless, O-SCWEICA is efficient in removing EMG artifacts from large, recorded EAC-EEG datasets as the denoising duration increases slightly for a large increase in EEG recording duration. On average, O-SCWEICA took 0.56 s to process the dataset with the longest duration (200 s). Furthermore, the accuracy of O-SCWEICA appears to improve with an increased number of samples. This is likely owing to more samples being available for the SCICA algorithm. O-SCWEICA was shown to yield better accuracy and efficiency compared to literature.

Therefore, the RQ can be answered: O-SCWEICA is not suitable for RT EMG artifact removal but is accurate and efficient in removing EMG artifacts from large, recorded EAC-EEG datasets. On average, O-SCWEICA removed EMG artifacts from 200 s EAC-EEG simulations in 0.56 s, with AV results of 2.40 MI and 0.54% NMSE. The performance of O-SCWEICA was largely unaffected by changes in the EMG artifact topography. Therefore, O-SCWEICA is robust for EMG artifact removal from EAC-EEG.

6.5 Chapter Summary

This chapter answered the RQ. O-SCWEICA – the new algorithm for the removal of EMG artifacts from EAC-EEG – required the optimisation of the WDPs in components generated by unconstrained FastICA. The optimal WDPs were Symlet order 9, at decomposition level three, using the Global threshold determination method and hard thresholding.

O-SCWEICA, U-SCWEICA, and STD-ICA were tested in EMG artifact removal from EAC-EEG datasets that varied in duration and artifact topographies. The results of each denoising algorithm were compared in terms of accuracy and efficiency. The results showed that O-SCWEICA and U-SCWEICA outperform STD-ICA in both accuracy and efficiency. U-SCWEICA was slightly more efficient than O-SCWEICA but the difference was minor, varying on average between 0.071 seconds faster and 0.009 s slower for the various datasets. O-SCWEICA performed more accurately than U-SCWEICA in almost all cases

owing to the optimisation of WDPs. Furthermore, O-SCWEICA demonstrated more accurate and efficient artifact removal compared to Akhtar *et al*. [14].

The RQ was subsequently answered and all RAs were addressed. In summary, O-SCWEICA is the best performing, robust algorithm used in this study and can likely be used for accurate and efficient EMG artifact removal of large EAC-EEG recordings. However, in the current methodology O-SCWEICA is unsuitable for RT EMG artifact removal.

Chapter 7 – Conclusion

Section 7.1 summarises the study in terms of objectives, methods and results. Thereafter, the main findings of the study are discussed in Section 7.2. Finally, Section 7.3 provides recommendations for future work. Terms are redefined in this chapter for ease of reading.

7.1 Summary

This study aimed to develop an accurate and efficient electromyographic (EMG) artifact removal algorithm for electroencephalography (EEG). Such an algorithm is necessary for research, clinical, and brain-computer interface (BCI) applications, where accurate and efficient EMG artifact removal is vital. EMG artifact removal is challenging – the EMG artifacts overlap with EEG in both amplitude and frequency.

Various methods have been implemented to remove EMG artifacts from EEG. The most successful approaches have combined multichannel methods, such as independent component analysis (ICA), with single channel methods, such as wavelet denoising. Spatially constrained, wavelet enhanced ICA (SCWEICA) was used to remove electrooculographic (EOG) artifacts from EEG with relative success but had not been extended to EMG artifact removal.

SCWEICA was adapted for EMG artifact removal by selecting an appropriate decomposition level for wavelet denoising. The other wavelet denoising parameters (WDPS) were the Haar mother wavelet, Global threshold determination method, and soft thresholding. The altered algorithm was called unoptimised SCWEICA (U-SCWEICA).

SCWEICA was then optimised for EMG artifact removal by optimising the WDPs. 5000 combinations of WDPs were explored. The optimal combination of WDPs were the Symlet order 9 mother wavelet at the third decomposition level, with the Global threshold determination method, and hard thresholding. These

WDPs were used with SCWEICA for EMG artifact removal and termed optimised SCWEICA (O-SCWEICA).

Both O-SCWEICA and U-SCWEICA were compared against the standard ICA (STD-ICA) denoising method using simulations. Simulations allow for ground truth assessment of an algorithm's performance. Therefore, 31-channel EMG artifact contaminated EEG (EAC-EEG) datasets were simulated. The datasets varied in both signal duration and artifact topography. The variations allowed for generalised performance assessment. Performance was analysed in terms of accuracy and efficiency. Accuracy was quantified using absolute maximum cross-correlation (AMCC), mutual information (MI), and normalised mean square error (NMSE). Efficiency was assessed using denoising duration – the time taken for the algorithm to output the cleaned EEG signals. Accuracies were quantified as an average across all electrodes (AV) and for the worst case (WC) channel – the channel with the largest EMG artifact.

Both O-SCWEICA and U-SCWEICA were shown to be more accurate and efficient than STD-ICA for all datasets. For the majority of results, O-SCWEICA was more accurate than U-SCWEICA, while U-SCWEICA was more efficient than O-SCWEICA.

The results imply that O-SCWEICA is not feasible for real-time (RT) EMG artifact removal but is robust to changes in artifact topography. O-SCWEICA was shown to be well suited for offline EMG artifact removal from longer duration EAC-EEG recordings. On average, O-SCWEICA took 0.56 s to process the longest EAC-EEG recordings, 200 s, while maintaining high MI of 2.40 AV and 1.76 WC. Compared to U-SCWEICA, O-SCWEICA was 1.14% and 12.3% more accurate in terms of AV and WC MI, respectively.

7.2 Main Findings

The main findings from this study are listed below, in order of significance:

- O-SCWEICA was shown to be accurate and efficient for EMG artifact removal from EAC-EEG, regardless of the size of the recording, or changes in the EMG artifact topography. O-SCWEICA is therefore a robust EMG artifact removal algorithm for EAC-EEG.
- O-SCWEICA is well suited for EMG artifact removal from large, EAC-EEG recordings. Recorded EEG is traditionally denoised with STD-ICA, which was shown to be far less accurate and efficient than O-SCWEICA. The use of O-SCWEICA instead of STD-ICA in denoising recorded EEG could increase the response rate for diagnostics of neurological disorders.
- The optimisation of WDPs in O-SCWEICA improved the accuracy of EMG artifact removal. The informed selection of WDPs allowed for optimal removal of neural signals from EMG artifact components. To the author's knowledge, this is the first study that performed optimisation of WDPs for independent components.
- Spatially constrained FastICA (SCFastICA) has shown to perform well with EMG artifacts. To the author's knowledge, this has not been tested before. Furthermore, for the first time, SCWEICA was used to remove EMG artifacts from EEG.

7.3 Recommendations for Future Work

There are a number of possible areas for future work. A few areas are discussed.

The optimised WDPs from this study should not necessarily be used in future studies, purely because they were determined by an optimisation process. Rather, future studies should use the optimisation process outlined in Section 5.4 to determine the best combination of WDPs for that study.

This study used simulations to represent EAC-EEG recordings. Simulations are inherently limited by the assumptions of their models. As such, O-SCWEICA should be tested on real EEG which likely requires an adaptation of O-SCWEICA to remove signals such as EOG and electrocardiographic (ECG) artifacts. The adaptation would require new calculations of the reference artifact topographies for SCFastICA. Different optimal WDPs would probably be required for each type of artifact. The application of different WDPs to different components may be possible with an algorithm that can identify different artifacts. Such an algorithm has been implemented by Mammone *et al*. [20].

This study used fixed reference topographies in SCFastICA depending on the three types of artifact topographies that were generated. Since the results for different artifact topographies were similar, the use of fixed topographies did not seem to affect the performance of O-SCWEICA. However, this might vary in other artifact topographies, especially when multiple sources of artifacts are present. Hence, different artifact topographies should be explored for EMG artifact removal.

This study used short time, separate windows to determine if O-SCWEICA could be used for RT artifact removal. The RT applicability of the algorithm could potentially be improved by implementing it over multiple, overlapping windows. RT artifact removal may be possible by increasing the sampling frequency so that more samples are available to the SCFastICA algorithm. However, an increased sampling frequency could be wasteful in terms of memory and might increase the denoising duration, hindering the aim of RT denoising.

As discussed in Section 5.2.2, different ICA algorithms – such as second order blind identification (SOBI) – may be better suited to the removal of EMG artifacts from EEG. Future studies may apply spatial constraints to various ICA algorithms in the O-SCWEICA methodology and contrast the impact that each ICA algorithm has on EMG artifact removal. The efficiency of each algorithm may affect potential RT implementation. As such, the number of floating-point

operations for each algorithm should also be noted as an additional efficiency measure.

ICA performance could potentially be improved at very low signal to noise ratio (SNR) by utilising additional electrodes [11]. Therefore, the simulations could be expanded for different and more numerous electrode layouts. The increased number of electrodes has the potential to improve accuracy but will likely decrease efficiency of O-SCWEICA. With larger electrode layouts of 128 or 256 electrodes, STD-ICA would probably be far less efficient compared to the results for 31 electrodes. Therefore, O-SCWEICA should be tested when applied to denser electrode layouts to determine if a substantial decrease in denoising duration is possible compared to STD-ICA.

References

[1] X. Chen, A. Liu, J. Chiang, Z. J. Wang, M. J. McKeown, and R. K. Ward, "Removing Muscle Artifacts From EEG Data: Multichannel or Single-Channel Techniques?," *IEEE Sensors Journal*, vol. 16, no. 7, pp. 1986–1997, Apr. 2016.

[2] M. K. Islam, A. Rastegarnia, and Z. Yang, "Methods for artifact detection and removal from scalp EEG: A review," *Neurophysiologie Clinique/Clinical Neurophysiology*, vol. 46, no. 4–5, pp. 287–305, Nov. 2016.

[3] X. Yong, R. K. Ward, and G. E. Birch, "Facial EMG contamination of EEG signals: Characteristics and effects of spatial filtering," in *2008 3rd International Symposium on Communications, Control and Signal Processing*, 2008, pp. 729–734.

[4] F. A. Rahman, M. F. Othman, and N. A. Shaharuddin, "A review on the current state of artifact removal methods for electroencephalogram signals," in *2015 10th Asian Control Conference (ASCC)*, 2015, pp. 1–6.

[5] D. M. Goldenholz *et al.*, "Mapping the Signal-To-Noise-Ratios of Cortical Sources in Magnetoencephalography and Electroencephalography," *Human Brain Mapping*, vol. 30, no. 4, pp. 1077–1086, Apr. 2009.

[6] B. W. McMenamin, A. J. Shackman, L. L. Greischar, and R. J. Davidson, "Electromyogenic Artifacts and Electroencephalographic Inferences Revisited," *Neuroimage*, vol. 54, no. 1, pp. 4–9, Jan. 2011.

[7] S. D. Muthukumaraswamy, "High-frequency brain activity and muscle artifacts in MEG/EEG: a review and recommendations," *Frontiers in Human Neuroscience*, vol. 7, Apr. 2013.

[8] A. J. Shackman, B. W. McMenamin, H. A. Slagter, J. S. Maxwell, L. L. Greischar, and R. J. Davidson, "Electromyogenic Artifacts and Electroencephalographic Inferences," *Brain Topography*, vol. 22, no. 1, pp. 7–12, Jun. 2009.

[9] J. A. Urigüen and B. Garcia-Zapirain, “EEG artifact removal—state-of-the-art and guidelines,” *Journal of Neural Engineering*, vol. 12, no. 3, p. 031001, 2015.

[10] T.-P. Jung *et al.*, “Removing electroencephalographic artifacts by blind source separation,” *Psychophysiology*, vol. 37, no. 2, pp. 163–178, 2000.

[11] D. Safieddine *et al.*, “Removal of muscle artifact from EEG data: comparison between stochastic (ICA and CCA) and deterministic (EMD and wavelet-based) approaches,” *EURASIP Journal on Advances in Signal Processing*, vol. 2012, no. 1, p. 127, Dec. 2012.

[12] M. Crespo-Garcia, M. Atienza, and J. L. Cantero, “Muscle Artifact Removal from Human Sleep EEG by Using Independent Component Analysis,” *Annals of Biomedical Engineering*, vol. 36, no. 3, pp. 467–475, Mar. 2008.

[13] C. J. James and C. W. Hesse, “Independent Component Analysis for Biomedical Signals,” *Physiological Measurement*, vol. 26, no. 1, p. R15, 2005.

[14] M. T. Akhtar, W. Mitsuhashi, and C. J. James, “Employing spatially constrained ICA and wavelet denoising, for automatic removal of artifacts from multichannel EEG data,” *Signal Processing*, vol. 92, no. 2, pp. 401–416, Feb. 2012.

[15] C. W. Hesse and C. J. James, “The FastICA algorithm with spatial constraints,” *IEEE Signal Processing Letters*, vol. 12, no. 11, pp. 792–795, Nov. 2005.

[16] C. W. Hesse and C. J. James, “On Semi-Blind Source Separation Using Spatial Constraints With Applications in EEG Analysis,” *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 12, pp. 2525–2534, Dec. 2006.

[17] M. Akay, “Wavelets in biomedical engineering,” *Annals of Biomedical Engineering*, vol. 23, no. 5, pp. 531–542, Sep. 1995.

[18] J. Rafiee, M. A. Rafiee, N. Prause, and M. P. Schoen, “Wavelet basis functions in biomedical signal processing,” *Expert Systems with Applications*, vol. 38, no. 5, pp. 6190–6201, May 2011.

[19] N. P. Castellanos and V. A. Makarov, “Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component

analysis," *Journal of Neuroscience Methods*, vol. 158, no. 2, pp. 300–312, Dec. 2006.

[20] N. Mammone, F. L. Foresta, and F. C. Morabito, "Automatic Artifact Rejection From Multichannel Scalp EEG by Wavelet ICA," *IEEE Sensors Journal*, vol. 12, no. 3, pp. 533–542, Mar. 2012.

[21] S. Bulárka and A. Gontean, "EEG Pattern Recognition Techniques Review," in *2015 IEEE 21st International Symposium for Design and Technology in Electronic Packaging (SIITME)*, 2015, pp. 273–276.

[22] J. Song *et al.*, "EEG source localization: Sensor density and head surface coverage," *Journal of Neuroscience Methods*, vol. 256, pp. 9–21, Dec. 2015.

[23] B. S. Chang, D. L. Schomer, and E. Niedermeyer, "Normal EEG and Sleep: Adults and Elderly," in *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, Sixth., Philadelphia: Lippincott Williams & Wilkins, 2011, pp. 183–214.

[24] N. Yeung, R. Bogacz, C. B. Holroyd, and J. D. Cohen, "Detection of synchronized oscillations in the electroencephalogram: An evaluation of methods," *Psychophysiology*, vol. 41, no. 6, pp. 822–832, Nov. 2004.

[25] F. Amzica and F. H. Lopes Da Silva, "Cellular Substrates of Brain Rhythyms," in *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, Sixth., Philadelphia: Lippincott Williams & Wilkins, 2011, pp. 33–63.

[26] N. Yeung, R. Bogacz, C. B. Holroyd, S. Nieuwenhuis, and J. D. Cohen, "Theta phase resetting and the error-related negativity," *Psychophysiology*, vol. 44, no. 1, pp. 39–49, Jan. 2007.

[27] H. Hallez *et al.*, "Review on solving the forward problem in EEG source analysis," *Journal of NeuroEngineering and Rehabilitation*, vol. 4, p. 46, Nov. 2007.

[28] W. Wadman J. and F. H. Lopes Da Silva, "Biophysical Aspects of EEG and MEG Generation," in *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, Seventh., New York, NY: Oxford University Press, 2018, pp. 89–103.

[29] I. I. Goncharova, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, “EMG contamination of EEG: spectral and topographical characteristics,” *Clinical Neurophysiology*, vol. 114, no. 9, pp. 1580–1593, Sep. 2003.

[30] H. V. Carter, *Muscles of the Head, Face, and Neck, (plate 378), Anatomy of the Human Body (Henry Gray), Sourced from Wikimedia Commons*. 1918.

[31] P. Kallioinen, *EEG Recording, Sourced from Wikimedia Commons*. 2012.

[32] Y. Ainhoren, S. Engelberg, and S. Friedman, “The cocktail party problem [instrumentation notes],” *IEEE Instrumentation Measurement Magazine*, vol. 11, no. 3, pp. 44–48, Jun. 2008.

[33] A. Hyvärinen, “Independent component analysis: recent advances,” *Philos Trans A Math Phys Eng Sci*, vol. 371, no. 1984, Feb. 2013.

[34] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. New York: John Wiley & Sons, Inc., 2001.

[35] A. Hyvärinen, “Fast and robust fixed-point algorithms for independent component analysis,” *IEEE Transactions on Neural Networks*, vol. 10, no. 3, pp. 626–634, May 1999.

[36] M. De Vos, L. De Lathauwer, and S. Van Huffel, “Spatially constrained ICA algorithm with an application in EEG processing,” *Signal Processing*, vol. 91, no. 8, pp. 1963–1972, Aug. 2011.

[37] G. P. Nason, “Choice of the threshold parameter in wavelet function estimation,” *Wavelets and Statistics*, vol. 103, pp. 261–280, 1995.

[38] S. S. Patil and M. K. Pawar, “Quality advancement of EEG by wavelet denoising for biomedical analysis,” in *2012 International Conference on Communication, Information Computing Technology (ICCICT)*, 2012, pp. 1–6.

[39] A. Graps, “An Introduction to Wavelets,” *IEEE Computational Science and Engineering*, vol. 2, no. 2, pp. 50–61, 1995.

[40] V. J. Samar, A. Bopardikar, R. Rao, and K. Swartz, “Wavelet Analysis of Neuroelectric Waveforms: A Conceptual Tutorial,” *Brain and Language*, vol. 66, no. 1, pp. 7–60, Jan. 1999.

[41] John TE Slade, *English: Discrete wavelet transform - a 3 level filter bank.* 2005.

[42] John TE Slade, *English: Frequency domain representation of the DWT (discrete wavelet transform).* 2005.

[43] J. C. Pesquet, H. Krim, and H. Carfantan, "Time-invariant orthonormal wavelet representations," *IEEE Transactions on Signal Processing*, vol. 44, no. 8, pp. 1964–1970, Aug. 1996.

[44] G. P. Nason and B. W. Silverman, "The Stationary Wavelet Transform and some Statistical Applications," in *Wavelets and Statistics*, vol. 103, A. Antoniadis and G. Oppenheim, Eds. New York, NY: Springer New York, 1995, pp. 281–299.

[45] R. R. Coifman and D. L. Donoho, "Translation-Invariant De-Noising," in *Wavelets and Statistics*, vol. 103, A. Antoniadis and G. Oppenheim, Eds. New York, NY: Springer New York, 1995, pp. 125–150.

[46] N. K. Al-Qazzaz, S. Ali, S. A. Ahmad, M. S. Islam, and M. I. Ariff, "Selection of mother wavelets thresholding methods in denoising multi-channel EEG signals during working memory task," in *2014 IEEE Conference on Biomedical Engineering and Sciences (IECBES)*, 2014, pp. 214–219.

[47] D. Valencia, D. Orejuela, J. Salazar, and J. Valencia, "Comparison analysis between rigrsure, sqtwolog, heursure and minimaxi techniques using hard and soft thresholding methods," in *2016 XXI Symposium on Signal Processing, Images and Artificial Vision (STSIVA)*, 2016, pp. 1–5.

[48] "Threshold selection for de-noising - MATLAB thselect." [Online]. Available: https://www.mathworks.com/help/wavelet/ref/thselect.html. [Accessed: 18-May-2018].

[49] A. Jafarifarmand and M. A. Badamchizadeh, "Artifacts removal in EEG signal using a new neural network enhanced adaptive filter," *Neurocomputing*, vol. 103, pp. 222–231, Mar. 2013.

[50] N. Taheri *et al.*, "Feasibility of blind source separation methods for the denoising of dense-array EEG," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015, pp. 4773–4776.

[51] H. Leutheuser, F. Gabsteiger, F. Hebenstreit, P. Reis, M. Lochmann, and B. Eskofier, “Comparison of the AMICA and the InfoMax algorithm for the reduction of electromyogenic artifacts in EEG data,” in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2013, pp. 6804–6807.

[52] S. Makeig, T. P. Jung, D. Ghahremani, and T. J. Sejnowski, “Independent Component Analysis of Simulated ERP Data,” Institute for Neural Computation, Technical, 2000.

[53] “BESA® | Brain Electrical Source Analysis: BESA Simulator > Overview,” *BESA® | Brain Electrical Source Analysis*, 11-Jul-2017. [Online]. Available: https://www.besa.de/products/besa-simulator/besa-simulator-overview/.

[54] H. Nolan, R. Whelan, and R. B. Reilly, “FASTER: Fully Automated Statistical Thresholding for EEG artifact Rejection,” *Journal of Neuroscience Methods*, vol. 192, no. 1, pp. 152–162, Sep. 2010.

[55] D. A. Bridwell, S. Rachakonda, R. F. Silva, G. D. Pearlson, and V. D. Calhoun, “Spatiospectral Decomposition of Multi-subject EEG: Evaluating Blind Source Separation Algorithms on Real and Realistic Simulated Data,” *Brain Topogr*, pp. 1–15, Feb. 2016.

[56] A. Delorme, T. Sejnowski, and S. Makeig, “Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis,” *NeuroImage*, vol. 34, no. 4, pp. 1443–1449, Feb. 2007.

[57] P. A. Estevez, M. Tesmer, C. A. Perez, and J. M. Zurada, “Normalized Mutual Information Feature Selection,” *IEEE Transactions on Neural Networks*, vol. 20, no. 2, pp. 189–201, Feb. 2009.

[58] K. Christian, “Data-Driven Methods for Fault Localization,” in *Data-Driven Methods for Fault Localization in Process Technology*, KIT Scientific Publishing, 2013, pp. 41–98.

[59] K. L. Moore, A. F. Dalley, and A. M. R. Agur, “Head,” in *Clinically Oriented Anatomy*, Seven., Lippincott Williams & Wilkins, 2013, pp. 820–980.

[60] B. C. Oxley, *International 10-20 EEG placement system showing Modified Combinatorial Nomenclature*. 2017.

[61] Biosemi, “128 Electrode Headcap,” 18-Apr-2006. [Online]. Available: https://www.biosemi.com/pics/cap_128_layout_medium.jpg. [Accessed: 19-May-2018].

[62] V. Jurcak, D. Tsuzuki, and I. Dan, “10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems,” *NeuroImage*, vol. 34, no. 4, pp. 1600–1611, Feb. 2007.

[63] “A03: Importing Channel Locations - SCCN.” [Online]. Available: https://sccn.ucsd.edu/wiki/A03:_Importing_Channel_Locations. [Accessed: 15-Mar-2018].

[64] D. J. Chong, G. L. Sahlem, and C. W. Bazil, “Introduction to Electroencephalography,” in *Review of Sleep Medicine (Second Edition)*, T. J. Barkoukis and A. Y. Avidan, Eds. Philadelphia: Butterworth-Heinemann, 2007, pp. 105–141.

[65] M. J. Fu, J. J. Daly, and M. C. Cavusoglu, “A Detection Scheme for Frontalis and Temporalis Muscle EMG Contamination of EEG Data,” in *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, 2006, pp. 4514–4518.

[66] B. Campillo, C. Martín, J.-C. Palma, A.-D. Fuentes, and J.-A. Alarcón, “Electromyographic activity of the jaw muscles and mandibular kinematics in young adults with theoretically ideal dental occlusion: Reference values,” *Medicina Oral, Patologia Oral Y Cirugia Bucal*, vol. 22, no. 3, pp. e383–e391, May 2017.

[67] A. Serrat, J. M. García-Cantera, and L. M. Redondo, “Isolated unilateral temporalis muscle hypertrophy: A case report,” *International Journal of Oral and Maxillofacial Surgery*, vol. 27, no. 2, pp. 92–93, Apr. 1998.

[68] S. W. Hwang, M. M. Abozed, A. J. Antoniou, A. M. Malek, and C. B. Heilman, “Postoperative Temporalis Muscle Atrophy and the Use of Electrocautery: A Volumetric MRI Comparison,” *Skull Base*, vol. 20, no. 5, pp. 321–326, Sep. 2010.

[69] M. G. Chez, M. Chang, V. Krasne, C. Coughlan, M. Kominsky, and A. Schwartz, “Frequency of epileptiform EEG abnormalities in a sequential

screening of autistic patients with no known clinical epilepsy from 1996 to 2005," *Epilepsy & Behavior*, vol. 8, no. 1, pp. 267–271, Feb. 2006.

[70] "Protect Your Source Code - MATLAB & Simulink." [Online]. Available: https://www.mathworks.com/help/matlab/matlab_prog/protect-your-source-code.html. [Accessed: 03-Apr-2018].

[71] A. Hyvärinen and P. Pajunen, "Nonlinear independent component analysis: Existence and uniqueness results," *Neural Networks*, vol. 12, no. 3, pp. 429–439, Apr. 1999.

[72] F. Gabsteiger, H. Leutheuser, P. Reis, M. Lochmann, and B. M. Eskofier, "ICA-based reduction of electromyogenic artifacts in EEG data: Comparison with and without EMG data," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014, pp. 3861–3864.