Use of Neurophysiological Measures to Estimate and Influence Change in Motor Function

by

Nader Riahi

M.A.Sc., The University of British Columbia, Vancouver, Canada, 1987

B.Sc., University of Wales, Cardiff, UK, 1984

Thesis Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in the

School of Engineering Science

Faculty of Applied Sciences

SIMON FRASER UNIVERSITY

Spring 2023

Declaration of Committee

Name:	**Nader Riahi**
Degree:	**Doctor of Philosophy**
Title:	**Use of Neurophysiological Measures to Estimate and Influence Change in Motor Function**
Committee:	**Chair:** **Shervin Jannesar** Lecturer, Engineering Science
	Carlo Menon Supervisor Professor, Engineering Science
	Ryan D'Arcy Committee Member Professor, Engineering Science
	Jie Liang Committee Member Professor, Engineering Science
	Teresa Cheung Examiner Assistant Professor, Engineering Science
	Domenico Formica External Examiner Professor, Engineering Newcastle University

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University

or has conducted the research

c. as a co-investigator, collaborator, or research assistant in a research project approved in advance.

A copy of the approval letter has been filed with the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, British Columbia, Canada

Update Spring 2016

Abstract

Frequent assessment of motor function would contribute towards personalization of therapeutic activities for stroke survivors. This study aimed to propose a pragmatic motor assessment tool that could facilitate frequent assessments towards recovery of lost function during chronic phase. We focused our investigation on the use of EEG-based resting state functional connectivity (FC) for this purpose.

The first phase investigated the suitability of FC for accurate estimation of motor impairment in stroke survivors. We selected phase synchronization as a measure of FC and Fugl-Meyer assessment for upper extremities (FMU) as a measure of impairment. We used projection algorithms to develop models for estimating FMU from FC. We showed cross-validated R^2 of 0.97 and an RMS error of 1.9 on FMU scale. The proposed method showed promise as a practical tool for frequent assessments of motor function.

There is diminished incentive for frequent assessments, however, if FC could only estimate large changes in function, as is the case with FMU. The second phase of our study investigated the ability of the proposed algorithms in estimating small incremental changes in motor function. Using objective measures of motor skill in healthy participants, we showed 98% accuracy in estimating small longitudinal changes in skill, thereby improving the incentive for frequent assessments.

The high estimation accuracy of our proposed method presented an opportunity to assess bidirectional interactions between FC and motor function. The third phase of our study investigated the prospects of inducing a change in motor skill by influencing FC. We selected mental imagery as the mechanism to influence FC and provided real-time neurofeedback on the selected connectivity channels to guide the mental imagery. We showed over 20% improvement in motor skill of a healthy individual through neurofeedback training alone.

Our proposed method showed promise in facilitating an individualized approach towards improvement of motor function that could complement the conventional therapeutic activities. It also showed potential for providing an accurate assessment of motor impairment, while addressing the challenges associated with the availability and expertise of examiners. It promotes frequent assessments of motor function and personalization of therapeutic activities for stroke survivors in chronic phase.

Keywords: Motor function; Electroencephalography; Neurofeedback; Mental Imagery; Functional connectivity; Partial least squares correlation and regression

Dedication

I dedicate this study to all stroke survivors that continue to endure challenging living conditions. I hope our research can contribute towards further recovery of their lost function.

Acknowledgements

My sincere thanks to Dr. Menon for his unwavering commitment and technical advice that kept me engaged in this research. I am especially grateful for his support throughout the challenging pandemic years. I thank Dr. D'Arcy for his help with the neuroscience aspects of my research and his insightful suggestions on academic reporting and publications. I would like to acknowledge Dr. Rawicz for his recommendation to consider mental imagery for the last phase of my research, which played a fundamental role in setting the direction of this study. I am grateful for Dr. Levy's expeditious dealings with all the administrative aspects of the research and her feedback on the manuscripts. My special thanks to Dr. Tehrani and Dr. Ahmadizadeh for the time spent on the content of the manuscripts and our long discussions towards clarification and flow of technical information in the reports. My cordial thanks to all members of the MENRVA group for their help in easing me back into the student life and my daughter Leila for her patient attempts in correcting my broken English in academic writing. I would like to end with expressing my most heartfelt gratitude to my wife Janice for her relentless reassurance to not be afraid of leaving a thirty-year career to pursue this lifelong ambition. She was my anchor during this humbling experience with her devoted encouragement that helped me through many moments of doubt. This work would not have been possible without her loving support.

Table of Contents

List of Tables

List of Figures

List of Acronyms

-tiveCH	Functional connectivity channels with negative regression coefficients
+tiveCH	Functional connectivity channels with positive regression coefficients
DL	Deep Learning
EEG	Electroencephalograph
ERD	Event Related Desynchronization
ERS	Event Related Synchronization
FC	Functional Connectivity
fMRI	functional Magnetic Resonance Imaging
FMU	Fugl-Meyer upper extremity motor score
M1	Primary Motor Cortex
MI	Mental Imagery
NF	Neurofeedback
NFT	Neurofeedback Training
NIPALS	Non-linear Iterative Partial Least Squares
PLI	Phase Lag Index
PLS	Partial Least Squares
PLSC	Partial Least Squares Correlation
PLSR	Partial Least Squares Regression
PT	Physical Training
R^2	Coefficient of Determination
RMSE	Root Mean Squared Error
rsFC	resting state Functional Connectivity
SD	Standard Deviation
SMR	Sensorimotor Rhythms
SVD	Singular Value Decomposition

Glossary

Bootstrap Resampling	Sampling with replacement from the original observations to calculate robustness of a feature within a data set.
Ceiling Effect	The condition where the output has reached its maximum value and cannot increase with any further changes in the input.
Channel	Refers to EEG electrode-pairs in conjunction with functional connectivity analysis.
Floor Effect	The condition where the output has reached its minimum value and cannot be reduced with any further changes in the input.
Functional Connectivity	Temporal correlation or synchronization between neurophysiological activities at different locations in the brain.
Imaginary Coherence	This is the same as spectral coherence but keeping only the imaginary part of the complex cross spectral density.
Instantaneous Functional Connectivity	Measures of functional connectivity from real time EEG signals.
Mental Imagery	Mental imagination of objects, places, attributes or activities, without any accompanying physical movement or engagement of muscles.
Morlet Wavelet	A sinusoidal waveform that is modulated by a Gaussian envelope. It is used as a bandpass filter, especially for non-stationary signals.
Neurofeedback	A form of biofeedback, where an individual's neurophysiological activities are measured in real time and presented back to the individual through audio, visual, or haptic feedback. It is used to help individuals with volitional control of their neurophysiological activities through mental imagery.
Permutation Test	Used to develop a distribution for a test statistic by shuffling the dependant variable with resect to independent variables of all observations.
Phase Clustering	magnitude of average phase angle differences between two EEG electrodes over a selected period.
Please Lag Index	average of sign of imaginary part of cross spectral density. There is no power component in this algorithm.
Predictors	Input variables in a regression model

Rehabilitation	Application of interventional activities for the purpose of generating improvement in function.
Spectral Coherence	cross spectral density that is normalized by auto spectral densities from two EEG electrodes.
Spontaneous Functional Connectivity	Measures of functional connectivity from resting state EEG data.

Executive Summary

Overarching goal of our study was to propose a method that had the potential to deliver an objective measure of motor impairment, facilitate frequent assessments, and provide a mechanism to guide mental imagery towards improving motor function.

Based on the results from prior studies, we selected neurophysiological measures for this purpose but aimed to remove the conventional constraint of selecting the motor areas as the focus of our analysis. We opted to use brain functional connectivity measures due to their global representation of interactions between different brain areas. For pragmatic reasons, we chose EEG-based resting state functional connectivity (rsFC) and used the peak phase synchronization index as the neurophysiological measure to estimate motor impairment. To the best of our knowledge, peak resting-state synchronization index had not been previously used for this purpose.

Our decision to include global brain interactions, both in location and frequencies, introduced other challenges associated with disproportionate dimensionality of the features (brain activities) and behavior (motor function), especially for small number of observations (participants). Presence of large number of features has a negative impact on the prediction performance of the regression models. This compelled us to use methodologies that were less sensitive to these conditions. We opted to use algorithms based on Projection to Latent Structures, also known as Partial Least Squares (PLS), to develop regression models for estimating motor impairment. We used a combination of PLS-Correlation and iterative PLS-Regression for dimensionality reduction and elimination of unrelated features from the regression models. Again, to the best of our knowledge, this combinational approach had not been previously used for feature selection and generalization of PLS regression models.

We selected FMU as a measure of motor impairment in stroke survivors. The first phase of our research investigated the performance of the proposed PLS method in using rsFC to estimate motor impairment in ten stroke survivors. Performance of the method was assessed with respect to its accuracy in estimating FMU scores. Cross-validation resulted in R^2 of 0.97 and a root-mean-squared-error of 1.9 (~ 2.9%) on FMU scale, which was less than the minimum detectable change (3 points or ~ 4.5%). We argued that the proposed method eliminated the need for execution of physical tasks during assessments,

was cost effective, portable, and simple to use. The proposed method showed potential in providing an objective measure of motor impairment, while addressing the challenges associated with the availability of trained examiners to carry out the assessments. The latter was deemed to be conducive towards more frequent assessment of motor function.

Accurate and frequent assessment of motor function would facilitate continual adjustment of rehabilitation strategies and personalization of therapeutic activities that could result in further recovery of lost function. However, there is very little incentive for frequent assessments if they can only measure large changes in motor function, which might take a long time to achieve. FMU is an example of such assessments, where a point change in FMU score would correspond to a large change in function. Since the performance of our proposed method was evaluated by its accuracy in estimating FMU scores, we needed to extend our analysis to include an evaluation of the performance in estimating small incremental changes in function. Small longitudinal changes in motor skill of healthy individuals through physical training was considered to be a good candidate for this purpose. We designed computer-based tracing tasks for skill training and used the spatial error in tracing as an objective measure of skill. We collected resting-state EEG data throughout the skill training program and developed regression models using our PLS-method and rsFC. Longitudinal data from seven healthy participants yielded an average accuracy of 98% (standard deviation of 1.2%) in estimating tracing errors. Although these results do not necessarily extend to stroke survivors, they were nevertheless indicative of the capability of the proposed method for estimating small incremental changes in motor function. We considered this as a foundational work that was necessary before future longitudinal research with other populations. The results also revealed the potential use of this method to study the impact of targeted brain stimulation on motor function through activation of individualized relevant FC networks.

Mental imagery (MI) is gaining attention as a strategy towards endogenous brain stimulation for improving motor skill. Neurofeedback (NF) is commonly used to guide MI in order to activate the relevant brain networks. Prior studies had primarily focused on generalized models that described the relationship between motor function and neurophysiological activities. These models were mostly based on SMR band power or FC with motor areas. Involvement of different brain networks and synchronization frequencies across subjects, along with differences due to age or existing capabilities, motivated us to deviate from the generalized approach and focus on individualized method

towards NF training. Our research on skill training with healthy participants showed promising results on developing individualized models for estimating small incremental changes in motor skill from rsFC. We used the models for identifying contributing EEG channels and band frequencies towards changes in skill. The identified channels were then used as individualized target candidates for NF training.

Consistent with the previous phase, we developed digital tracing tasks for physical skill training and used the spatial error in tracing to gauge the change in skill of a healthy participant. Selection of healthy participants allowed us to monitor small changes in motor skill and FC from physical and NF training. We applied our proposed method to identify the channels with the largest margin for increasing FC as the candidates for NF training. We showed over 20% reduction in tracing error through neurofeedback training alone, without any additional physical training. We also showed retention of improvement in skill for several days after the completion of neurofeedback training. We are not aware of any prior study that used an individualized selection of channels and frequencies for application of NF to guide MI towards improving motor skill.

In summary, we proposed a method based on peak-rsFC between networks that were not constrained to interactions with motor areas or any specific synchronization frequencies. We used PLS algorithms for both selecting relevant features (channels and frequencies) and developing generalized models for estimating motor impairment. We also showed that the method exhibited good performance in estimating small longitudinal changes in motor skill from individualized models for healthy participants. In the final phase of our research and through a case study, we demonstrated that the method had potential in identifying individualized channels that could be used for NF training to guide mental imagery towards incremental improvements in motor skill.

This study was limited to the basic introduction of a new method and foundational demonstration of concept. The future direction of this study can be divided into two categories: 1) Development of a prototype for the generalized model to be used for estimating motor impairment. This will require the collection of 2-minutes resting state EEG data along with a physical appraisal of motor impairment from as many stroke survivors as feasible. We recommend this to be initially limited to individuals in chronic phase; 2) Development of an NF prototype that could initially be used with motivated healthy participants that are actively involved in personal skill improvement training. A

successful demonstration of the prototype with healthy participants could then be extended to stroke survivors in chronic phase that might be actively working towards regaining specific skills.

Chapter 1.

Introduction

Material in this chapter is extracted, reproduced, and modified with permission from the following papers:

N. Riahi, V. A. Vakorin, C. Menon, “Estimating Fugl-Meyer Upper Extremity Motor Score from Functional-Connectivity Measures,” *IEEE transactions on neural systems and rehabilitation engineering*, vol. 28, no. 4, pp. 860-868, Apr. 2020.

N. Riahi, R. D’Arcy, C. Menon, “A Method for Estimating Longitudinal Change in Motor Skill from Individualized Functional-Connectivity Measures,” *Sensors* 2022, DOI: 10.3390/s22249857.

N. Riahi, W. Ruth, R. D’Arcy, C. Menon, “A Method for Using Neurofeedback to Guide Mental Imagery for Improving Motor Skill,” *IEEE transactions on neural systems and rehabilitation engineering* 2022, DOI: 10.1109/TNSRE.2022.3218514.

1.1. Chapter Overview

This chapter explains our motivation, research questions and the resulting objectives. It concludes with the potential contributions towards stroke rehabilitation.

1.2. Background and Motivation

A large percentage of stroke survivors experience motor impairment in upper extremities that affects their activities of daily living [1]. An estimated 30% of the survivors continue to experience functional inadequacies throughout the chronic phase with psychological and financial impact to both the stroke survivors and their caregivers [2]. There is limited institutional and governmental support for continuing rehabilitation during chronic phase [3]. This has partially been based on the perceived diminished capacity for

improvement, or plateauing, despite existing evidence for further recovery of function from application of novel rehabilitation protocols. Continual adjustment of rehabilitation strategy and therapeutic activities is conducive towards breaking out of this plateau [4]. Accurate and frequent assessment of motor function would contribute towards development of rehabilitation strategies and personalization of therapeutic activities [5]. Frequency of assessments, however, is hampered by the availability of trained examiners. Cost effective assessment tools and complementary therapeutic methodologies may help in this regard.

Fugl-Meyer assessment is an accepted method for evaluation and numerical representation of motor function after stroke [6]. It is used to obtain a baseline assessment, as well as monitor and quantify longitudinal changes in motor function during rehabilitation [7], [8]. Given proper training of the examiners and following standardized approach, the minimum detectable change for Fugl-Meyer upper extremity motor score (FMU) can be reduced to approximately 3 points, which is just under 5% of maximum scale for FMU assessment [9]. This is exacerbated by the subjectivity of the evaluation, when terms such as “performed partly” or “limited range” are used to quantify behavior and can cause variability in the final assessment scores. Furthermore, every point change in FMU score corresponds to large changes in function and as such may not contribute towards assessment of small incremental improvements in execution of motor tasks. This would further discourage frequent assessments since achieving a minimum detectable change in FMU score might take a long time to accomplish. Finding physiological measures that can objectively and accurately estimate small changes in motor function may help reduce dependency on the skill level and experience of personnel that are tasked with carrying out the assessment. Selection of physiological measures, however, should not be solely based on the technical challenges associated with how well they can estimate behavior, but also on their ability to address pragmatic issues associated with reduced capacity of stroke survivors as well as cost and duration of assessment [10]. Addressing the pragmatic challenges might help facilitate more frequent assessments, which in turn could be conducive towards the development of personalized therapeutic activities and rehabilitation strategies. Pertaining to technical considerations and specifically the accuracy of assessment, use of physiological measures must result in comparable or better accuracy than conventional protocols used for evaluating motor function and require minimal expertise to carryout the measurements.

Although stroke can cause local structural damage, the dysfunction is generally extended to remote sections of the brain that are functionally connected to the damaged areas [11]. The ability to perform a broad analysis of network interaction should therefore be considered when selecting a physiological measure. Given our focus on brain networks, it seems appropriate to concentrate on neurophysiological measures to quantify the network interactions. Brain functional connectivity (FC) has shown potential for use as such neurophysiological measures, especially due to their ability to analyze global network interactions [12], [13]. Taking into account the aforementioned pragmatic challenges associated with mobility and feasibility, there is increased interest in using resting state FC measures through electroencephalography (EEG) for this purpose [14]. FC is generally evaluated in terms of correlation in activities of different brain areas [15]. Empirical research with EEG has shown that engagement and communication between brain regions are facilitated through neural synchronization at different frequencies [16], [17] and that the synchronization parameters can be used as metrics to quantify FC between associated brain areas [18]. The high temporal resolution of EEG systems is favourable towards accurate measurement of synchronization between different regions of the brain and is why EEG systems are commonly used in studies where coherence is selected as a measure of FC. Considering the potential for quantifying motor related FC from resting state measurements and the suitability of EEG systems in measuring FC through coherence, along with its relative low cost and portability, motivates the use of EEG-based resting-state FC (rsFC) as the neurophysiological measure of choice for estimating motor function.

Given that FMU is a standard tool for the assessment of motor impairment, it would be reasonable to evaluate the performance of methodologies based on rsFC against that of FMU. Provided that the use of rsFC can result in comparable accuracy, it could then be promoted as an alternative pragmatic assessment tool for the appraisal of motor function during rehabilitation. However, while methodologies based on rsFC might show good performance in estimating FMU, their accuracy in estimating motor function could only be evaluated relative to the accuracy of FMU. Assessment of function based on FMU is somewhat subjective in nature and generally quantifies fairly large changes in motor function [9]. To promote the use of these new methodologies for frequent assessment of function, they need to show potential for evaluating small changes in motor function that are not feasible with FMU assessment. To gain a better understanding of the estimation

accuracy of the methods based on rsFC, their performance needs to be evaluated against a more objective measure of behavior that can be quantified at smaller incremental change as compared with conventional assessment protocols. Specific to motor function, acquisition of a skill by healthy individuals through physical training is a good candidate for this purpose. The ability of the algorithms to accurately estimate modest longitudinal changes in motor skill may prove useful in addressing the technical requirements of the new assessment tool. It can also motivate using the tool at shorter intervals, thereby increasing the assessment frequency and enabling small adjustments in rehabilitation strategy accordingly. This can improve the prospects of accomplishing further improvements in motor function during chronic phase [4].

Pertaining to diminished potential for improvement during chronic phase and the need for continual adjustment of rehabilitation strategy, there is increased interest in brain stimulation as a complementary approach along with physical therapy [19]. Specific to motor function and use of FC as a neurophysiological measure, the enhanced computing power of current electronic devices has facilitated the real-time processing of coherence and the subsequent analysis of instantaneous FC [20]. This has presented an opportunity to not only monitor changes in FC for the purpose of assessment, but also investigate non-physical activities that might influence FC for the purpose of inducing change in motor function. The effect of these influencing activities on related behavior such as motor function may then be examined and if favorable, integrated into relevant training programs for improvement of function [21]. Brain FC may be influenced through external stimulation [22], [23] or endogenous stimulating activities. The latter is the concept behind operant conditioning through neurofeedback (NF), where individuals try to self-regulate and control the desired brain activation through mental imagery (MI) in a closed-loop process [24], [25], [26]. The low implementation risk associated with MI as a mechanism to influence behavior motivates its use as a complementary therapeutic activity to improve motor function during chronic phase.

1.2.1. Summary

In summary, our motivation for this research includes the following:

- Use of neurophysiological measures can reduce the possible subjectivity of motor assessments, remove the potential complications associated with language barriers, and reduce the need for examiners' subject matter expertise.
- Selecting resting state FC as the neurophysiological measure of choice eliminates the need for execution of physical tasks as part of the assessment protocol. This can address the challenges associated with reduced capacity and mobility of stroke survivors as well as duration of assessment.
- EEG modality is well suited for high resolution evaluation of FC when using coherence as a measure of FC. EEG systems can be operated with minimal training and at a capital cost that could potentially be recovered after a limited number of assessments in a clinical setting or at the point of care.
- Development of an easy-to-use motor assessment protocol based on the above considerations could facilitate more frequent appraisal of motor function, which is conducive towards the development of personalized intervention strategy and adjustments of therapeutic activities at shorter intervals.
- Utilization of EEG based FC as predictors of motor function enables the real-time application of neurofeedback to guide targeted MI that could influence the pertinent FC towards improvement of motor function. Targeted MI could be used as a complementary approach to expedite and potentially enhance the efficacy of established intervention and rehabilitation activities during chronic phase.

1.3. Research Questions and Objectives

In this study, we aimed to propose a method that could be used for more frequent assessment of motor impairments in stroke survivors, thereby allowing for adjustment of rehabilitation strategy towards extending the improvement in function. We also aspired to use the proposed method in facilitating a complementary therapeutic activity for improving motor function, particularly during chronic phase. For pragmatic reasons, we focused on investigating the potential use of EEG-based FC for assessment, monitoring, and influencing motor function. As such, our study had to address three research questions:

1. Is EEG-based rsFC a suitable neurophysiological measure to accurately estimate motor impairment in stroke survivors?

2. Can rsFC be used to estimate small incremental changes in motor function?
3. Can we induce a change in motor function by influencing individualized FC channels through MI?

Our goal was to investigate signal processing and analysis algorithms that could be used to identify suitable measures of FC towards accurate estimation of motor function. We opted to use synchronization at different frequencies as a measure of FC. The signal processing algorithms had to deal with non-stationary nature of EEG signals when extracting frequency specific features. There was also the complication associated with large data sets where the number of features (FC measures) far exceeded the number of observations (participants). In summary, the algorithms had to deal with non-stationarity of the EEG signals, potential collinearity within the feature set, and presence of large number of unrelated features that could compromise the performance of prediction models. This led to our first research objective to explore the performance of a combination of processing algorithms for estimating motor impairment.

Objective-1: Propose a method for estimating FMU from rsFC measures.

Performance of the proposed method could be quantified by its accuracy in estimating FMU. This, however, was a subjective assessment of impairment that measured relatively large changes in motor function. To gain a better understanding of the estimation accuracy of the method, its performance needed to be evaluated against a more objective measure of motor function that could be quantified at smaller incremental changes as compared with FMU.

Objective-2: Test the performance of the proposed method in estimating small incremental changes in motor function.

If the proposed method proved suitable for estimating small incremental changes in motor function, then it could be used for monitoring the short-term longitudinal interactions between FC and function. Our focus this far was on quantifying the change in FC as a consequence of change in motor function. The next step was to investigate the reverse relationship and assess the impact of changing FC towards inducing incremental change in motor function. We opted to use MI as a non-invasive and cost-effective approach towards influencing FC. To address the effect of MI on relevant FC, and its subsequent impact on motor function, we needed to develop a real-time processing algorithm that

could monitor instantaneous measures of FC. To evaluate the applicability of our proposed method for this purpose, we aimed to use the same processing modules from objective-2 to quantify the FC measures that were selected for estimating the change in motor function. This allowed us to monitor the impact of different MI by providing real-time feedback to guide volitional control of the pertinent individualized FC measures.

Objective-3: Investigate the prospects of influencing the individualized FC measures through MI for improving motor function.

1.4. Thesis Structure

This section describes the outline of the thesis.

Chapter 2 covers the prior work related to the use of FC for baseline assessment of motor function, patient stratification, and prediction of longitudinal change in motor skill. It continues with a review of prior studies on the application of neurofeedback to influence motor function. Common algorithms for signal processing, feature extraction, and analysis are introduced. The chapter concludes with a list of equipment and materials used in this study.

Chapter 3 details the specifics of the methodology used towards selection of features for estimating motor impairment in stroke survivors. Results pertaining to the performance of the methodology along with some limitations of the approach are presented.

Chapter 4 presents the performance of the proposed methodology in estimating small longitudinal change in motor skill. Focus is on individualized approach as a precursor towards application of neurofeedback for skill improvement.

Chapter 5 builds on the results from estimating change in motor skill. It details a process for selection of pertinent features that could be used for neurofeedback. A method for utilization of neurofeedback to guide mental imagery as an endogenous brain stimulation is then presented. The chapter concludes with the results of applying the proposed individualized method for improving motor skill in a healthy participant.

Chapter 6 lists the potential approach for implementation of proposed methodology as well as future work to address the limitations and areas for improvement.

1.5. Contributions

We proposed a method that could be used for both assessment and intervention, and consequently contribute towards continual adjustment of rehabilitation strategies and facilitate complementary therapeutic activity for improving motor function.

As an assessment tool, it could provide an objective measure of motor function that limits the dependence on the availability and expertise of trained examiners. It is easy to use, and economical to implement. It could improve the feasibility of carrying out more frequent assessments, which should help manage the therapeutic activities and overall rehabilitation strategy. Resting state analysis is a preferential approach due to its least dependence on the physical abilities of individuals with stroke. Furthermore, use of neurophysiological measures can reduce the possible subjectivity of assessments and remove the potential complications associated with language barriers and cognitive abilities. When considering these factors, then our proposed method (if proven consistent through further research with a larger number of participants) would provide considerable incentive for clinical use. The EEG system can be operated with minimal training and at a capital cost that could potentially be recovered after a limited number of assessments in a clinical setting or at the point of care. Depending on the accuracy and resolution of these neurophysiological measures in estimating motor function, they may also act as early indicators of the impact of therapeutic activities before any discernable change in motor function, thereby facilitating the opportunity to finetune and personalize rehabilitation strategies at shorter intervals.

As an intervention tool, it could assist in guiding individuals with their mental imagery that would stimulate relevant networks of the brain to improve motor function. It would contribute towards a complementary mechanism to expedite and potentially enhance the efficacy of physical therapy. Collecting resting state EEG data before (or after) the NF sessions may also facilitate a progressive identification of newly available contributing brain areas as motor function improves. A continuous development of

alternative prediction models as new data samples become available may help extend further improvements in behavior. This is especially relevant to stroke survivors in chronic phase with historically low expectations for additional recovery of lost function.

Chapter 2.

Literature Review, Materials and Method

Material in this chapter is extracted, reproduced, and modified with permission from the following papers:

N. Riahi, V. A. Vakorin, C. Menon, “Estimating Fugl-Meyer Upper Extremity Motor Score from Functional-Connectivity Measures,” *IEEE transactions on neural systems and rehabilitation engineering*, vol. 28, no. 4, pp. 860-868, Apr. 2020.

N. Riahi, R. D’Arcy, C. Menon, “A Method for Estimating Longitudinal Change in Motor Skill from Individualized Functional-Connectivity Measures,” *Sensors* 2022, DOI: 10.3390/s22249857.

N. Riahi, W. Ruth, R. D’Arcy, C. Menon, “A Method for Using Neurofeedback to Guide Mental Imagery for Improving Motor Skill,” *IEEE transactions on neural systems and rehabilitation engineering* 2022, DOI: 10.1109/TNSRE.2022.3218514.

2.1. Chapter Overview

This chapter covers the prior work related to the use of brain functional connectivity for baseline assessment of motor function, patient stratification, and prediction of longitudinal change in motor skill. It continues with a review of prior studies on the application of neurofeedback to influence motor function. Common algorithms used for signal processing, feature extraction, and analysis are then introduced. The chapter concludes with a list of equipment and materials used in this study.

2.2. Prior Work

Prior studies with stroke survivors, using functional Magnetic Resonance Imaging (fMRI), showed neural network reorganization and formation of compensatory functional

networks between different areas of the brain to recover the lost motor function [7], [8], [11]. Their results indicated correlation between resting state FC measures and motor function, concluding that resting state FC could be a relevant neurophysiological measure to represent the extent of impairment after stroke. This was based on prior observations that lack of physical activity associated with resting state does not imply a silent brain. There is evidence of resting state network activities between motor regions of the brain while a person remains motionless but awake [27].

Application of FC analysis has not been limited to understanding the current state of individuals' motor impairment. Correlation between baseline FC measures and future improvements in motor function has been proposed for patient stratification [12]. Longitudinal studies have also investigated the correlation between change in FC measures and ongoing improvement in motor function after stroke [13]. Interestingly, the authors showed the relationship between changes in FC and treatment-induced improvement in motor function was contingent on the baseline status. For patients with more sever baseline impairment, increases in FC between ipsilesional and contralesional areas were beneficial towards achieving better treatment-induced motor gains. Whereas the reverse relation (decrease in FC) was associated with larger gain in motor function for patients with less sever baseline impairment. Use of FC as neurophysiological measures for assessment, monitoring, and prediction of change in motor function has therefore gained attention for developing and finetuning rehabilitation strategies.

The modality, however, is not constrained to fMRI. Given the challenges associated with mobility of stroke survivors and operational cost of modalities, there is increased interest in using resting state FC measures through EEG for this purpose [14]. EEG-based FC measures were shown to correlate with motor status after stroke [28]. The authors focused on the correlation between global connectivity measures and behavioral assessments such as FMU. Estimates of correlation with FMU was evaluated from the FC with the primary motor cortex as the seed location. Individual connectivity measures between the seed network and all other locations around the brain were then assessed to generate an overall map of interacting networks [8], [28], [29]. This approach has also been extended to produce a single functional connectivity index reflecting the spatial average value of functional connectivity between all inter or intra-hemispheric brain regions [30]. In one study, connectivity with ipsilesional primary motor cortex (M1) accounted for an impressive 78% of variance in motor impairment of stroke survivors [28].

The authors used coherence at specific canonical frequency bands as a measure of FC. Partial Least Square (PLS) algorithms were applied to estimate motor function from FC measures. However, addition of other measures such as corticospinal tract injury into the model were recommended to increase performance in estimating impairment. The need to include other measures than connectivity alone in explaining the variance with motor function may have been due to the exclusion of other relevant connectivity measures between areas that did not include M1. The authors also restricted connectivity analysis to beta-bands which further constrained the contribution of other FC measures towards correlation with motor function.

Although FC related studies showed good results in estimating motor function, the accuracy of the algorithms was evaluated against subjective measures that quantify relatively large changes in motor function [9]. To gain a better understanding of the estimation accuracy of these algorithms, their performance needed to be evaluated against a more objective measure of behavior that could be quantified at smaller increments as compared with conventional assessment protocols. Incremental improvement in motor skill by healthy individuals through physical training was considered to be a good candidate for this purpose.

Longitudinal motor learning studies with fMRI showed that some areas of the brain exhibit a transient change in FC while other areas showed a more lasting change towards consolidation and long-term retention [31]. The authors observed an initial increase in resting state FC in right postcentral gyrus and supramarginal gyrus during the early stages of skill acquisition, reflecting the integration of sensorimotor and visuospatial attention. This was followed by a decrease in FC in the same areas later in the motor learning process, indicating a reduced recruitment of these areas. They also observed a consistent increase in FC within the left supramarginal gyrus throughout the motor skill learning, which they interpreted as evidence of persistent FC change towards retention of learned skill. In another motor skill training study, the reduction in FC was interpreted as an indication that some networks may have become less important or conversely, developed specialized neuronal circuits for efficient performance of motor tasks [32]. These persistent changes in FC were shown to have potential for quantifying, and subsequently estimating the change in motor skill.

Investigations of relationship between change in FC and motor skill has not been constrained to fMRI modality. Prior work using EEG modality showed correlation between motor learning and FC in specific frequency bands [33] [34]. Several studies focused on the ability of global configuration of FC to predict skill acquisition in healthy participants. In a study with 3D visuomotor learning using movement smoothness as a measure of skill, the authors showed that the acquired skill could be predicted by the strength of alpha-band FC before the motor training tasks [35]. In a similar study with mirror drawing tasks, the authors concluded that the resting state alpha-band FC could be implicated in providing the optimal resources necessary for performing an upcoming task and that the FC measures could predict how well the skill could be learnt [36]. Analysis of FC measures were not limited to alpha-band alone. In a study with healthy participant the authors predicted subsequent degree of motor adaptation from resting state beta-band FC between primary motor cortex and anterior prefrontal cortex [37]. There was also variability in the areas of brain for which connectivity measures were evaluated. Prior work using PLS analysis showed the ability to predict skill acquisition from beta-band FC between primary motor cortex and left parietal cortex [38]. The need for broader analysis of network interactions that were not limited to primary motor cortex was further reinforced through a motor adaptation study that showed involvement of parieto-occipital and fronto-parietal cortical component [39]. These studies point towards manifestation of wide-ranging frequencies and brain areas towards correlation between FC and change in motor skill.

Specific to motor function, analysis of FC as a neurophysiological measure has been focused on predicting or monitoring the changes in function as a consequence of physical training or therapeutic activities. There is, however, an increased interest in investigating the impact of influencing these neurophysiological measures through brain stimulation to induce change in function. Brain stimulation can be through external sources [40], [41], or endogenous stimulation through MI [42], [43], [44]. The latter is generally guided through NF in the form of visual, audio, or tactile feedback to help individuals select the appropriate MI that impacts the neurophysiological measures of interest.

Selection of neurophysiological measures for the purpose of NF is not limited to connectivity measures. In a review paper by Jeunet *et al.* [45], the authors listed several studies on volitional self-regulation of sensory-motor rhythms (SMR) through NF training to trigger neuroplasticity. Results showed promise towards acquisition of motor skill. SMR covers a broad range of frequencies within alpha and beta bands that are selected based

on the research objectives. In a case study with a stroke survivor in chronic phase, increasing the alpha-band power over CZ electrode was shown to improve motor function in terms of gait speed [46]. The authors speculated that improvement in function was due to NF training alone, since there was no change to the exercise routine and the individual was many years into chronic phase. In a controlled study to improve shooting performance, the authors provided NF on SMR (12–15 Hz) over C3, CZ, and C4 areas with the goal of enhancing power through MI [47]. The sham group received instruction to increase alpha power over T3 and reduce power over T4. The control group did not receive NF training. Shooting performance was not improved in the sham or control group. Application of MI using NF has also been extended towards indirect control of motor skill. In a recent study by Sidhu *et al.* [48] on dual-task performance, the authors showed improvement in performance of both single and dual-task motor activity by decreasing the alpha-power over the supplementary motor area. They reasoned that movements are automized through repetition but can be compromised through illness or external constraints. These can result in conscious control of motor activities that reduces movement proficiency with adverse effects on dual-task performance. They argued that NF resulted in improvement in cognitive performance, which in turn reduced conscious control of motor tasks, thereby improving dual-task performance.

Regarding the use of NF to influence FC or SMR spectral power, investigations have been expanded to observations of auxiliary impact of NF on other parts of the brain that were not the target of NF. Experiments with healthy participants indicated a progressive reduction in FC within associative areas that were not the focus of the NF training [49]. Participants were asked to perform MI of grasping with their right hand. They were provided feedback on the spectral power over C3 and CP3 by adjusting the position of a moving visual target. The authors showed that learning to control the target position was accompanied with a progressive reduction in functional integration of the associative areas including fronto-occipital, parieto-occipital, fronto-central, and bilateral temporal. They concluded that the changes in FC were associated with the increasing automaticity and were the ancillary effect of NF training on the spectral power over the motor areas.

Impact of regulating spectral power in motor areas through NF is not limited to changes in FC of associative areas. In a recent study using bi-modal EEG-fMRI experiment with stroke survivors in chronic phase, the authors showed a decrease in inter-hemispheric connectivity between premotor and primary motor areas and a reduction in

ipsilesional self-inhibitory connections [50]. They argued that the impact of NF training that target localized motor areas should be broadened to assess FC changes throughout the brain. Other studies using fMRI modality have shown that motor imageries result in overlapping activation of areas including primary motor, premotor, supplementary motor, and parietal areas that include sensorimotor and posterior parietal lobe [19]. It has therefore been argued that analysis of network interactions might be a more wholistic approach for NF implementation [42], [51]. Allaman *et al.* [52] took this one step further, proposing that the performance in motor sequence tasks were directly related to the strength of spontaneous FC and less related to classical task induced activations such as Event Related De/Synchronization (ERD/ERS). The authors concluded that the latter was more of a compensatory mechanism for lower global connectivity states which enabled "low-performing subjects to accomplish a task despite unfavorable neural states". In a related study with both healthy participants [53] and stroke survivors [54] the authors showed that an NF protocol based on increasing the global alpha-band FC of the primary motor cortex resulted in improvement in motor function. These are significant findings and encouraging results in favor of using NF to facilitate an endogenous self stimulation of brain towards improving motor skill. Restriction of analysis to *a priori* selection of specific frequency band and brain areas, such as alpha-band and motor cortex, is conducive towards generalization and subsequent standardization of approach, but might be limiting in terms of performance [55], [56], [57].

2.3. Methodologies

2.3.1. Assessment of Motor Function

Assessment of motor impairment for stroke survivors were carried out by trained examiners as part of another study [58] and shared with us to carry out this research.

Pertaining to motor learning and acquisition of motor skill by healthy individuals that could be quantified at small incremental change, we used a computer-based tracing task and evaluated the spatial error in tracing as an objective measure of skill. Python 3.7 was used to create our experimental track patterns on a computer screen. We opted for elliptical trajectories instead of straight lines [36] to increase the degree of difficulty for

tracing tasks. The track was constructed from eight quarter-ellipses that were arranged to form a four-section curved-pattern as shown in Figure 2.1(a). In this figure, the green dot represents the starting point for the placement of the mouse pointer, before tracing the corresponding track towards the red dot. Selection of the active track section was controlled through the program. We adjusted the pointer speed so that the distance between opposing tips of the track pattern corresponded to approximately 35 cm of mouse travel across the torso (x-axis: left to right) and 25 cm away from the torso (y-axis: top to bottom). The rationale was to promote large enough physical movements to engage multiple arm joints without the need to move the torso. We also disabled the driver option that allowed a non-linear relation between the pointer's speed and the mouse acceleration. The objective was to maintain relational consistency between the mouse and pointer coordinate systems. The horizontal area in physical space was thus represented by pixels on the computer screen, and time was measured through the computer's real-time clock. Total position-error while tracing a track section was determined by the area (in pixels) between the actual pointer trajectory (Trace) and the desired track path (Track) as shown in Figure 2.1(b).

Tracing Performance

Position-error between the participants' tracing trajectory and the intended track pattern was selected as one of the performance indicators. Time taken to trace each section (trial) was selected as another indicator. Total position-error during a trial was determined by the area between the trace and track trajectories (Figure 2.1(b)). To discourage participants' attempt to reduce position-error by tracing slower, we used each trial time as a penalizing (multiplication) factor to inflate the respective position-error. Conversely, to discourage participants' attempt to reduce trial time by moving the mouse too quickly to stop at the destination vertex, we accumulated positional offsets from the track endpoint until the pointer came to rest at the destination vertex.

We used both accumulated position-error as well as the accumulated product of position-error and its corresponding tracing time from each trial, as two separate indictors of tracing performance. Improvement in motor skill through training reduces the magnitude of one or both these measures. Participants were updated on their tracing performance after each training session. Position-error was measured in units of pixels-squared,

representing the area between the track and trace trajectories, and converted to spatial units of squared-centimetres (cm^2) based on an estimated coverage of 0.25 mm^2 per pixel. The product of position-error and its associated tracing time was measured in units of cm^2seconds.

Each training session consisted of 90 tracing trials, generating 90 intermediate performance values. We used two different approaches to produce a measure of tracing performance for each session: First, a single value corresponding to the median of all 90 trials (single-median option), and second, the median of the first 30 trials as a measure of performance before training, and the median of the last 30 trials as a measure of performance after training (dual-median option). The two measures allowed for separate evaluation of consolidated and short-term learning.

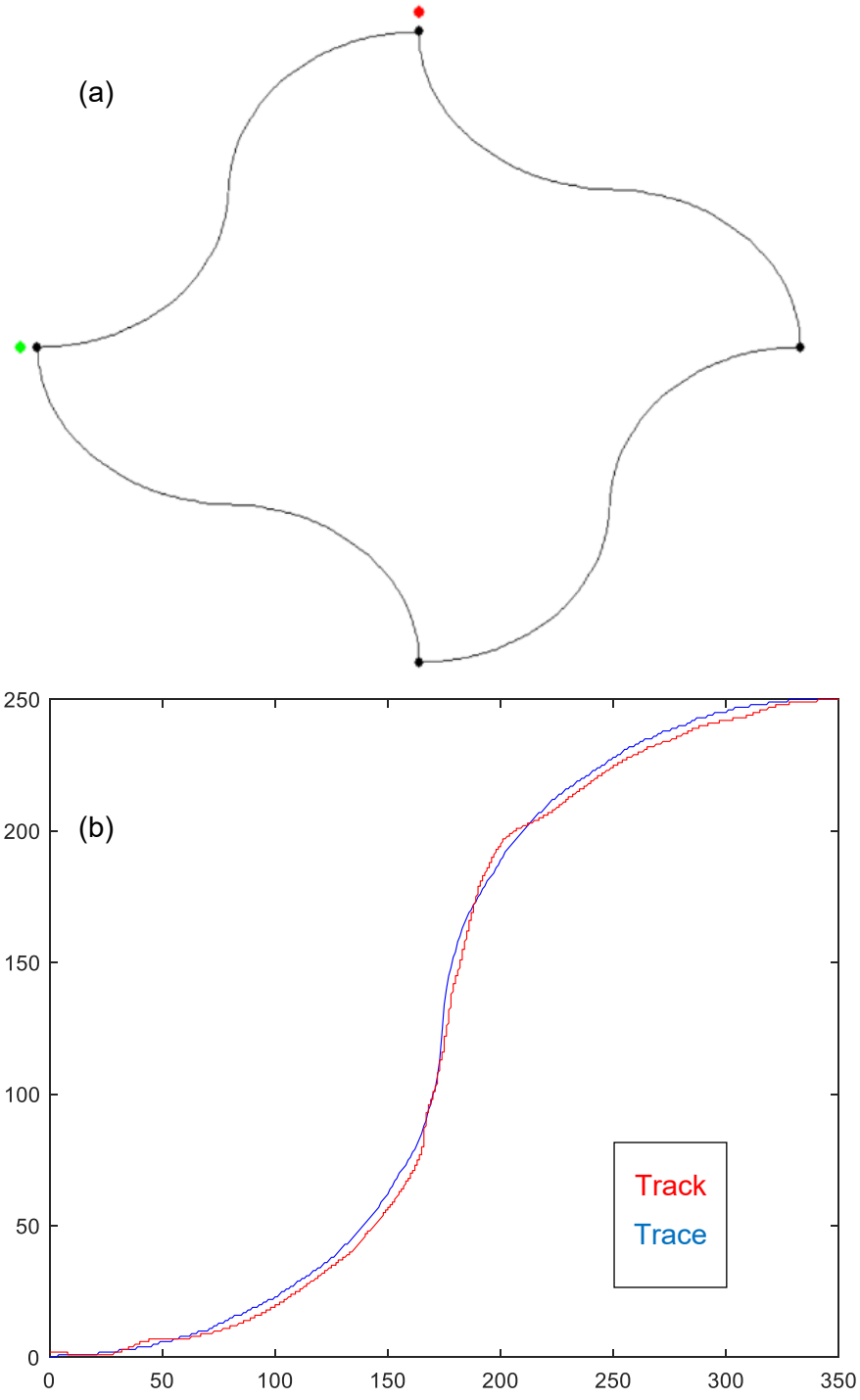

Figure 2.1 **(a) General shape of the track pattern to be traced by participants. (b) An example of the tracing trajectory over a single-track section. The position-error is the total area between the trace (red) and track (blue) pattern. Axes for (b) are in units of screen pixels.**

2.3.2. Band Separation

We used complex Morlet wavelets for bandpass filtering of EEG signals from each electrode. Morlet wavelet is a sinusoid wave within a Gaussian envelope. The center frequency of the wavelet is determined by the frequency of the sinusoid and the bandwidth is determined by the standard deviation of the Gaussian envelope [59]. For a given sinusoid frequency, increasing the number of sinusoid cycles within the wavelet increases the standard deviation of the Gaussian envelope, resulting in a narrower frequency bandwidth of the bandpass filter. The wavelet filter characteristics can therefore be adjusted by the center frequency and the number of cycles of the sinusoid within the wavelet [60]. The actual filtering process can be implemented by multiplying the Fourier Transform of the wavelet and EEG data from each electrode and then applying Inverse Fourier Transform to convert back to time domain. The resulting frequency-specific amplitude and phase information of the filtered EEG data can be used to compute instantaneous coherence between activities at different electrodes [60].

Morlet Wavelets

This section uses material from [59] and [62].

Fourier decomposition is an established mechanism by which spectral information within a signal is analyzed and reported on. The basic premise behind Fourier transform is stationarity, which assumes the constituent components of the signal maintain their frequency structure over the transform time. This is not necessarily the case for many sources including biological signals. To remedy this, the signal may be analyzed in smaller portions, during which, signal stationarity can be assumed. One common approach is the application of wavelet transforms where the window kernel is imbedded in the decomposition as part of the transform and changes as a function of the frequency. Phase values for wavelet decomposition are estimates of the instantaneous angles at each time point.

Wavelet transforms have been found to be particularly useful for analyzing aperiodic, intermittent, transient and localized discontinuities such as edges or temporary appearance of unrelated signals. It can simultaneously examine data in both time and

frequency. Wavelet transform uses localized wavelike functions known as wavelets. It can be manipulated through translation (time) and dilation (stretched or squeezed). It effectively quantifies local matching of the wavelet with the input signal at a particular dilation scale. There are many published wavelets that one can choose from with the selection criteria depending on the nature of input signal and ultimate objective of analysis. But irrespective of the selection, wavelets generally start with a mother wavelet with unit scale and no translation. The mother wavelet is the basis for all other derived wavelets through translation and dilation. This study focuses on gaussian-based wavelets.

Amongst basic features of wavelets and their transforms is that Fourier transform of the wavelet has a finite energy and does not contain a zero-frequency. Combination of these two features constrains the wavelets at both ends of their spectrum and makes them functionally behave as bandpass filters. For applications where the input signal has an inherent oscillatory nature, then Morlet wavelets may be a good choice for decomposition and analysis of the data. Morlet wavelet is a single sinusoid that has been tapered by a gaussian kernel (window). Frequency of the sinusoid determines the peak (or center) frequency of the wavelet. For the mother wavelet, this is referred to as the characteristic frequency. These wavelets can localize transient oscillatory components of the input signal that are in the same frequency range as the peak frequency of the wavelet. We will see later that the peak frequency of a Morlet wavelet is proportional to the degree of dilation (scale) of that wavelet which re-establishes the relation between the perturbation period and scale of the wavelet. Representing the sinusoid in its analytic form ($e^{i2\pi ft}$) will introduce the imaginary components into the transform and provides an opportunity to extract an estimate of the instantaneous phase of the transformed signal. A basic restriction for the wavelet construction is that the peak frequency of a wavelet should not have a period that is longer than the total length of the input signal. For example, one cannot analyze activities at lower frequencies than 1 Hz for a signal that is only one second long. In practice one should allow for several cycles (> 3) of the peak frequency within the length of the input signal. It should also not violate the Nyquist limit as determined by the sample rate.

As mentioned in previous paragraphs, Morlet wavelet uses a gaussian kernel. It is helpful to formulate the relationship between the kernel's standard-deviation and its frequency spectrum to establish the behavior in terms of filter properties. This is why the length of a kernel is generally assessed in terms of its standard-deviation, even though

the actual window length is a multiple of the standard-deviation. Note that in practice the kernel is weighted by some proportion of the standard-deviation to create a unit energy kernel.

Frequency spectrum of a function can be obtained through its Fourier transform as shown below

$$\boldsymbol{H(f) = \int h(t) e^{-i2\pi ft} dt} \tag{2.1}$$

We can assign **h(t)** to be a gaussian envelope in the form of

$$\boldsymbol{h(t) = C_{\sigma_t} e^{-\frac{t^2}{2{\sigma_t}^2}}}$$

with C as the normalization factor (in reference for this study it is $\frac{1}{\sqrt{\sigma}\pi^{1/4}}$ and $\boldsymbol{\sigma_t}$ the standard-deviation of the kernel in time domain.

The frequency spectrum ***H(f)*** will also be a gaussian envelope [61] of the form

$$\boldsymbol{H(f) = K_{\sigma_f} e^{-\frac{f^2}{2{\sigma_f}^2}}}$$

with K as the normalization factor in terms of $\boldsymbol{\sigma_f}$. But more importantly, in the derivation by Derpanis [61], replacing $\boldsymbol{\chi}$ with $\boldsymbol{t}$ and $\boldsymbol{\omega}$ with $\boldsymbol{2\pi f}$, the relationship between the standard-deviation in time and frequency domain is derived to be

$$\boldsymbol{\sigma_f = \frac{1}{2\pi\sigma_t}} \tag{2.2}$$

This highlights the inverse relation between the standard-deviations in time ($\boldsymbol{\sigma_t}$) and frequency ($\boldsymbol{\sigma_f}$) domain. The longer the time domain window the narrower the bandwidth of the filter and vise versa.

The mother wavelet for the Morlet transform is constructed by a complex sinusoid that is windowed by a gaussian kernel of unit standard-deviation (scale = 1) and no translation, as shown below

$$\boldsymbol{\varphi(t) = e^{i2\pi f_0 t} . e^{\frac{-t^2}{2}}}$$

The frequency of the sinusoid in the mother wavelet describes the peak frequency (f_0) of that wavelet, which is also the center frequency of the wavelet spectrum. Following the relationship in Equation-2.2, the frequency domain spectrum also has a gaussian form with the bandwidth inversely proportional to the dilation scale ($a = \sigma_t$) and centered at f_0. It is important to note that f_0 determines the number of cycles in the mother wavelet and by definition, remains constant for all derived wavelets at different scales. This is one of the reasons that f_0 is referred to as the characteristic frequency of the Morlet wavelet. As the gaussian envelope is dilated ($a > 1$), the enveloped sinusoid is stretched while keeping the number of cycles constant, which effectively reduces the center frequency of the resulting wavelet. Conversely, contracting the gaussian envelope increases the center frequency of the resulting wavelet. This can be seen by replacing the time-index (t) with $(t-b)/a$ to account for the dilation (a) and translation (b) of the wavelet.

$$\varphi_{a,b}(t) = \varphi\left(\frac{t-b}{a}\right) = e^{i2\pi f_0[(t-b)/a]}.e^{\frac{-[(t-b)/a]^2}{2}} \tag{2.3}$$

We can now proceed with the wavelet transform in the same way as the Fourier transform, keeping in mind that we have to use the conjugate of the complex basis function $\psi^*_{a,b}(t)$. The Morlet transform is therefore given by

$$T(a,b) = C_a \int x(t).e^{\frac{-(t-b)^2}{2a^2}} e^{-i2\pi(f_0/a)(t-b)}\, dt \tag{2.4}$$

Equation-2.4 shows the relationship between the dilated wavelet center-frequency f in the analytic form in terms of the characteristic frequency f_0 and the wavelet scale a ($f = f_0/a$).

There is one more point to consider here and that is the frequency spectrum of the Morlet wavelet itself. This is important as it shows the relationship between the bandwidth of the effective filter and the standard-deviation of the gaussian kernel (scale a of the wavelet). It also describes the dependency of the phase of the Morlet transform on the translation b across the input signal. Following the Fourier transform as formulated in Equation-2.1,

$$\emptyset_{a,b}(f) = \int \varphi_{a,b}(t) e^{-i2\pi ft} dt = \int \varphi\left(\frac{t-b}{a}\right) e^{-i2\pi ft} dt$$

Making the substitution $\tau = (t-b)/a$ and therefore $dt = a.d\tau$

$$\emptyset_{a,b}(f) = \int \varphi(\tau) e^{-i2\pi f(a\tau+b)} a d\tau$$

Bringing out the ***b*** term will result in

$$\emptyset_{a,b}(f) = a.e^{-i(2\pi f)b} \int \varphi(\tau) e^{-i2\pi(af)\tau} d\tau \quad (2.5)$$

Noting the rescaled frequency (***a.f***) in the Fourier integral and the phase term ***b*** in the analytic sinusoid term outside the Fourier integral.

Using Fourier transform, it can be seen that the frequency response of a gaussian envelope is in the shape of another gaussian, with its bandwidth inversely proportional to the standard-deviation of the kernel in time domain (Equation-2.2). The impact of changing the standard deviation of the gaussian kernel is the smearing in time (larger standard deviation) and frequency (smaller standard deviation) domain. The smearing changes as a function of decomposition frequency due to its relationship with the scale of the wavelet (Equation-2.4).The decomposition frequencies (***f***) are constrained by the number of cycles (f_0) in the mother wavelet and change with the size of the gaussian kernel through the relationship $f = f_0 / a$. As a side note, it may be confusing why the literature uses the scale ***a*** instead of the usual σ as the descriptor for standard-deviation of the gaussian kernel. This is because there are other wavelets that are not described in terms of a gaussian kernel. Yet the concept of wavelet dilation still applies with the degree of dilation represented by the scale ***a*** of the wavelet.

There are two observations to be made from the spectrum analysis of the Morlet wavelet as described by Equation-2.5; First, the peak frequency of the Morlet wavelet is shifted down in frequency with increasing scale of the wavelet. This can be seen from the term (***af***) in the analytical sinusoid inside the Fourier integral. Second, the phase of the Morlet transform is centered on the translation as can be seen from the ***b*** term in the analytical sinusoid outside the Fourier integral. In the Morlet transform the decomposition sinusoids are windowed with a gaussian first and then collectively used to transform the input signal. The decomposition sinusoids are therefore centered at (***b***) and express the instantaneous phase angles at each time point.

To summarize, as a Morlet wavelet is dilated, its frequency spectrum shifts to the lower frequencies with smaller bandwidth and as it is contracted, its frequency spectrum

shifts to the higher frequencies with larger bandwidth. This is because of the relationship between the peak frequency and scale ($f=f_0/a$) for Morlet wavelet and the general relationship between the scale/size of the gaussian and the frequency bandwidth ($\sigma_f=1/2\pi\sigma_t$). Noting that the standard deviation of the gaussian determines the tradeoff between temporal and frequency precision, and the property of Morlet wavelet that the size of the gaussian at a particular frequency is determined by the number of cycles in the mother wavelet, one can rephrase the relationship to state that the number of cycles in a wavelet controls the tradeoff between temporal and frequency precision in the Morlet transform. The kernel size of the derived wavelets increases with the increasing number of cycles in the mother wavelet. Larger number of cycles results in a better frequency precision and smaller number of cycles results in better temporal precision. The temporal precision should not be confused with temporal resolution as the latter is entirely determined by the sampling rate. The fact that temporal precision of Morlet transform increases with frequency (at a constant number of cycles) is also conducive towards generating much cleaner transform plots around the nonstationary periods of the input signal. The frequency resolution in Morlet transform is related to the number of scales used for the transform and can be selected based on the amount of desired spectrum overlap in the scale-space.

2.3.3. Functional Connectivity Measures

Execution of motor tasks involves coordinated activities of broad areas of the brain [11], [63], which is generally referred to as functionally connected group of neural networks [15]. Empirical research with electroencephalography (EEG) has shown that engagement and communication between brain regions are facilitated through neuronal synchronization at different frequencies [16], [17], and that the synchronization parameters can be used as metrics to quantify FC between associated brain areas [18]. One such parameter is the phase of neuronal oscillations at similar frequencies, where the consistency of phase differences over a typical duration of hundreds of milliseconds can be considered as an indication of synchronized activities [64]. It is therefore reasonable to concentrate on frequency domain processes that can estimate the phase of these oscillatory behaviours. The dynamic nature of these synchronized activities dictates a

localized approach towards computation of phase. We have tried to facilitate this through the use of wavelet filters as described in the previous section.

A popular approach for quantifying the phase synchronization between oscillatory signals is the measurement of coherence. Strictly speaking, the term coherence is commonly used to indicate the use of both amplitude and phase [62]. For example, spectral coherence refers to calculation of instantaneous phase differences that are modulated by corresponding signal amplitude. Taking power out of the equation and focusing only on phase-differences is commonly referred to as synchronization. For the sake of simplicity, we use the term coherence to refer to both cases and qualify whether amplitude is included or not through the specific names for each of the processing algorithms that are considered.

Coherence can be viewed as a measure of variance explained between two signals $s_1(\omega,t)$ and $s_2(\omega,t)$ at a specific frequency ω over a localized period t and is quantified by the ratio of cross spectral density and auto spectral density (Equation-2.6). Its values are bounded between 0 and 1 [18].

$$Coh_{s_1 s_2}(\omega) = \frac{\left|\frac{1}{n}\sum_{t=1}^{n} |s_1(\omega,t)|\ |s_2(\omega,t)|\ e^{i\Delta\varphi(\omega,t)}\right|}{\sqrt{(\frac{1}{n}\sum_{t=1}^{n}|s_1(\omega,t)|^2)(\frac{1}{n}\sum_{t=1}^{n}|s_2(\omega,t)|^2)}} \quad (2.6)$$

$\Delta\varphi(\omega,t)$ is the phase difference between s_1 and s_2 at frequency ω and time point *t*. There are different versions of Equation-2.6 used to quantify coherence between signals from two EEG electrodes. Some of these extensions exclude the signal amplitude and focus on the phase component, arguing that the amplitude is a confounding factor towards evaluation of phase synchronization. Although advocates of coherence as a measure of FC argue that random phase differences would most likely result in much smaller coherence, even in the presence of good amplitude correlation. Similarly, strong phase synchronization (consistent phase difference) would result in a large coherence despite small amplitude correlation. There is further assertion that larger signal amplitude would generally provide a better signal to noise ratio, which results in better quality of phase calculation and quantification of phase differences.

We computed five different measures of coherence/synchronization for all unique combinations of electrode pairs and frequency points. These were used to produce non-directional functional connectivity measures between different brain regions.

Spectral Coherence

Coherence measures suffer from a common problem with EEG modality related to volume conduction. A signal from an electrode is a superposition of different current sources, some of which are due to field spread through volume conductors (brain, skull, and scalp) [65]. These are instantaneous components that appear in almost perfect synchrony at different electrodes. As such, they result in significant coherence values that may not be a true reflection of FC between the areas represented by those electrodes. One way to reduce the impact of volume conduction is to apply spatial band-pass filters that suppress low frequency variations across the scalp electrodes [66]. This would effectively increase topographical selectivity and localize the signal activities to areas close to each electrode. Of course, a drawback with this approach is that it equally reduces the impact of synchronization from widely distributed sources as well as deep sources, which generally appear as spatially broad activities. We used a Surface Laplacian filter based on spherical spline method to reduce the impact of volume conduction prior to coherence analysis [67]. This is a second order spatial derivative on the surface tangent to local scalp. Coherence is then evaluated by the ratio of cross spectral density normalized by auto spectral densities from each electrode as presented by Equation-2.6 [60].

Phase Clustering

Amplitude normalizing the components in Equation-2.6, will remove the effect of amplitude and evaluates coherence in terms of phase differences only [68]. Vector addition of these resulting phase differences over the time period of concern is referred to as phase clustering or phase-locking-value represented by Equation-2.7.

$$\boldsymbol{PC}_{s_1 s_2}(\boldsymbol{\omega}) = \left| \frac{\mathbf{1}}{\boldsymbol{n}} \sum_{\boldsymbol{t}=\mathbf{1}}^{\boldsymbol{n}} \boldsymbol{e}^{i\Delta\boldsymbol{\varphi}(\boldsymbol{\omega},\boldsymbol{t})} \right| \tag{2.7}$$

$\Delta\boldsymbol{\varphi}(\boldsymbol{\omega},\boldsymbol{t})$ is the phase difference between $\boldsymbol{s_1}$ and $\boldsymbol{s_2}$ at frequency $\boldsymbol{\omega}$ and time point $\boldsymbol{t}$.

Use of phase clustering is motivated by the claim that it is more reflective of true synchronization than coherence, where the latter is confounded by the measures of

amplitude. Phase clustering method is also impacted by volume conduction as is the case with coherence. Use of Surface Laplacian filters can help in this regard, with the same advantages and drawbacks as highlighted for coherence measures.

Imaginary Part of Coherence

Projection of the analytical representation of coherence onto the imaginary axis, or simply taking the imaginary part of the complex coherence, has gained traction for evaluation of FC [69]. It is motivated by the assumption that volume conduction is instantaneous and as such primarily concentrated around phase difference of 0° or 180°. This is represented by the real axis of the complex representation of coherence and taking only the imaginary part would effectively remove any synchronization measures associated with volume conduction. Rewriting Equation-2.6 in its cartesian form, we will have

$$\boldsymbol{Coh}_{s_1 s_2}(\boldsymbol{\omega}) \;=\; \frac{\left|\frac{1}{n}\sum_{t=1}^{n} \; |s_1(\omega,t)| \; |s_2(\omega,t)| \; \left(cos\left(\Delta\varphi(\omega,t)\right) + \; i \; sin\left(\Delta\varphi(\omega,t)\right)\right)\right|}{\sqrt{\left(\frac{1}{n}\sum_{t=1}^{n}|s_1(\omega,t)|^2\right)\left(\frac{1}{n}\sum_{t=1}^{n}|s_2(\omega,t)|^2\right)}}$$

The imaginary part of the coherence can then be defined as

$$\boldsymbol{IC}_{s_1 s_2}(\boldsymbol{\omega}) \;=\; \frac{\left|\frac{1}{n}\sum_{t=1}^{n} \; |s_1(\omega,t)| \; |s_2(\omega,t)| \;\; sin\left(\Delta\varphi(\omega,t)\right)\right|}{\sqrt{\left(\frac{1}{n}\sum_{t=1}^{n}|s_1(\omega,t)|^2\right)\left(\frac{1}{n}\sum_{t=1}^{n}|s_2(\omega,t)|^2\right)}} \tag{2.8}$$

A point to note about Equation-2.8 is its dependence on the average amplitude of the signals. It has the effect of reducing the relative amplitude of the imaginary part in the presence of strong volume conduction, and as such reducing the signal to noise ratio of the resulting IC and the overall increase in susceptibility to system noise [70].

Phase Lag Index (PLI)

To reduce susceptibility to noise and indirect impact of volume conduction, one can use the sign of the imaginary component, without the amplitude modulation. [70].

$$\boldsymbol{PLI}_{s_1 s_2}(\boldsymbol{\omega}) \;=\; \frac{1}{n}\sum_{t=1}^{n} \; \boldsymbol{sign}\left(sin\left(\Delta\varphi(\omega,t)\right)\right) \tag{2.9}$$

Note that PLI removes the magnitude of phase difference from the analysis, which effectively allows for large variance in phase difference during the analysis period. This may be viewed contrary to the assumption that FC is based on strong phase synchronization between different neuronal population.

Weighted Phase Lag Index

To compensate for potential discontinuity that can arise at phase differences around the 0° or π° from small perturbations when using PLI, the sign of the imaginary component can be weighed by the amplitude of the imaginary component [71]. This is referred to as wPLI. Defining the imaginary part of the cross-spectrum as ***Im(ω,t)***, we can rewrite PLI equation as:

$$\boldsymbol{Im_{s_1s_2}(\omega,t)} \;=\; \boldsymbol{|s_1(\omega,t)|}\;\boldsymbol{|s_2(\omega,t)|}\quad \boldsymbol{sin\big(\Delta\varphi(\omega,t)\big)}$$

$$\boldsymbol{wPLI_{s_1s_2}(\omega)} \;=\; \frac{\frac{1}{n}\sum_{t=1}^{n}\left(\boldsymbol{|Im_{s_1s_2}(\omega,t)|}\;\boldsymbol{sign\big(Im_{s_1s_2}(\omega,t)\big)}\right)}{\frac{1}{n}\sum_{t=1}^{n}\boldsymbol{|Im_{s_1s_2}(\omega,t)|}} \qquad (2.10)$$

The difference between $\boldsymbol{wPLI_{s_1s_2}(\omega)}$ and $\boldsymbol{IC_{s_1s_2}(\omega)}$ is primarily in the denominator where normalization by the amplitude of the cross spectrum is not carried out, which removes the potential influence of volume conduction. Addition of the amplitude of the imaginary component in the numerator, however, biases the FC measures towards larger phase differences.

2.3.4. Partial Least Squares Algorithm

Unless specifically stated, details of information presented in this section can be found in [72] and [73].

We applied PLS analysis to examine the relationship between EEG functional connectivity and FMU. PLS is a multivariate statistical approach [74] that is well suited for analyzing large data sets such as electrophysiological activities, especially when the number of features, *i.e.,* connectivity measures at different frequencies, are much larger

than the number of observations (participants). PLS-Correlation (PLSC) and PLS-Regression (PLSR) are two categories of PLS analysis. They both utilize singular value decomposition (SVD) to decompose the covariance of brain-activities such as functional connectivity measures, and behavior such as motor impairment (FMU) into a set of latent variables (LV) at lower dimensions.

The goal of PLSC is to analyze the percentage of variance in observations that is explained by the corresponding latent variables. When input variables (brain activities and behavior) are mean centered and normalized, then SVD can be used to analyze correlation between the input variables at a lower dimension as determined by the rank of the covariance matrix. Given that we are interested in single behavior variable (FMU), then the latent space for PLSC is made up of only one column over all observations. PLSC can therefore be used to study the manifestation of functional connectivity across specific frequencies and electrode-pairs [72], [73], [74], [75] to identify the contributing channels and frequencies towards correlation between FMU and functional connectivity indices.

Permutation and bootstrap resampling are two computational methods of statistical inference for significance and robustness of contributing electrode-pairs [73] at specific frequencies. We applied one ‘global’ permutation-based test to evaluate the significance of overall correlations in PLSC and one ‘local’ bootstrap-based test to evaluate the robustness of contribution of individual functional connectivity measures. For permutation test, a new singular value sample is obtained by randomly reordering the FMU measures while leaving the original order of the connectivity measures. We generated a sampling distribution of singular values under null hypothesis through (500 to 1000) iterations. We then computed the *p-value* of the original singular value based on its location within the generated distribution. To obtain a measure of the robustness of contribution of connectivity measures, we used bootstrap-resampling to compute the standard error of the saliences associated with each connectivity measure. This involved sampling with replacement of the covariance matrix through 500 to 1000 iterations. The bootstrap ratio value (original saliences divided by the standard error of the corresponding bootstrap distributions) represents the robustness of contribution of individual EEG features to the overall correlations between functional connectivity and FMU and is equivalent to z-scores. [74]. In our study, we used the terms bootstrap ratio values and z-scores interchangeably.

PLSR is an asymmetric approach that can be used to predict dependent-variables such as FMU from independent-variables such as connectivity measures. PLSR is well suited for noisy independent variables with multicollinearity. It employs Non-linear Iterative Partial Least Squares (NIPALS) technique to simultaneously extract latent vectors with maximal covariance from the dependent and independent variables [76]. The process is repeated iteratively until fully deflating the input variables. The latent variables are used to compute the regression coefficients. Overfitting is addressed by extracting the least number of latent vectors that account for most of the variance. The prediction quality is evaluated through cross-validation techniques such as leave-one-out. The quality first increases with increasing number of latent variables, but then starts to decrease as it approaches towards overfitting [76]. One method to avoid overfitting is to stop adding latent variables as soon as the quality starts to decrease.

Overview of PLS

PLS finds the latent variables, sometimes referred to as factors, that explain large portions of the variation in the input data. Latent variables are identified through Singular Value Decomposition (SVD) for PLSC and Non-linear Iterative Partial Least Squares (NIPALS) for PLSR.

To carryout PLS analysis, we arrange the brain activities in a matrix of independent variables with observations (or participants) as the rows and features (functional connectivity measures) as the columns, and similarly the dependant variables with observations as rows and behaviour measures (FMU or skill) as the columns. The relationship between the columns of independent variables and columns of dependant variables can then be quantified through the dot product of respective columns. If the columns are centered (subtract the column mean), then the dot product produces the covariance between these columns. If the columns are further normalized (divided by the sum of the squared entries of each column), then the dot product quantifies the correlation between the features and behaviour columns. The correlation matrix R is then formulated as:

$$\boldsymbol{R = X^T Y} \tag{2.11}$$

where $\boldsymbol{X^T}$ is the transpose of the brain activities and $\boldsymbol{Y}$ the behavior measures. Given $\boldsymbol{n}$ observations, $\boldsymbol{i}$ brain activity measures, and $\boldsymbol{j}$ behaviour measures, then $\boldsymbol{X}$ is an $(\boldsymbol{n \times i})$ matrix of brain activities and $\boldsymbol{Y}$ an $\boldsymbol{(n \times j)}$ matrix of behaviours. The singular value decomposition of the correlation matrix is then given by:

$$\boldsymbol{R = W \Delta C^T} \tag{2.12}$$

$\boldsymbol{W}$ and $\boldsymbol{C}$ are the matrices of saliences for the brain activities and behaviour measures respectively. The number of salience (columns of $\boldsymbol{W}$ and $\boldsymbol{C}$) are determined by the rank of matrix $\boldsymbol{R}$. The latent variables (or scores) are quantified by the projection of the brain activities and behaviour measures ($\boldsymbol{X}$ and $\boldsymbol{Y}$) onto their respective saliences. Each latent variable is a linear combination of the respective input variables from $\boldsymbol{X}$ and $\boldsymbol{Y}$. The advantage of working in the latent space is the resulting reduction in the dimensionality of the input data.

So far, we have explained the conventional approach of data organization and decomposition that is common between PLSC and PLSR. This is where the two methods diverge in how the projections are utilized for the purpose of determining the largest amount of information common to both $\boldsymbol{X}$ and $\boldsymbol{Y}$ (PLSC) or predicting $\boldsymbol{Y}$ from $\boldsymbol{X}$ (PLSR).

PLSC

As a brief recap, given that both $\boldsymbol{X}$ and $\boldsymbol{Y}$ are centred and normalized, the correlation matrix R is computed from $\boldsymbol{X}$ and $\boldsymbol{Y}$ using Equation-2.11. Singular value decomposition of $\boldsymbol{R}$ produces the matrices of saliences of $\boldsymbol{X}$ and $\boldsymbol{Y}$ represented by $\boldsymbol{W}$ and C respectively (Equation-2.12). Latent variables are then generated by the projection of the columns of $\boldsymbol{X}$ and $\boldsymbol{Y}$ onto their respective saliences as linear combination of the original entries. They can be used to explain the correlation between the input matrices by quantifying the latent vectors that exhibit the maximum covariance between the input data. This is symbolically represented by the following expressions [73]. The lower-case variables are the column vectors of the corresponding matrices represented by the upper-case variables.

$$\boldsymbol{L_X = X\,W} \quad \text{and} \quad \boldsymbol{L_Y = Y\,C} \tag{2.13}$$

Column vectors of $\boldsymbol{L_X}$ and $\boldsymbol{L_Y}$ represent the latent vectors of $\boldsymbol{X}$ and $\boldsymbol{Y}$ respectively.

$$\boldsymbol{L_{X,l} = X\, W_l} \quad \text{and} \quad \boldsymbol{L_{Y,l} = Y\, C_l}$$

$$\boldsymbol{L_{X,l}^T\, L_{Y,l'} = 0} \text{ for } \boldsymbol{l \neq l'}$$

$$\boldsymbol{w_l^T\, w_l = c_l\, c_l^T = 1}$$

$$\boldsymbol{L_{X,l}^T\, L_{Y,l} = \Delta_l}$$

The covariance is presented by the corresponding singular value from the diagonal matrix $\boldsymbol{\Delta}$ (Equation-2.12) and are presented in a descending order of magnitude. Note that in our case, the behaviour matrix has a single column representing the FMU (or measure of motor skill), which limits the rank of the covariance matrix to 1, resulting in a single latent vector for the brain scores. The correlation can then be computed from $\boldsymbol{Y.L_X}$ that is scaled by the number of observations.

Computational methods such as permutation test are utilized to build a probability distribution which can then be used to check for significance of the latent variables in evaluating the correlation. A permutation sample can be generated by random re-ordering of the observations in $\boldsymbol{X}$ while leaving the order of the observations in $\boldsymbol{Y}$ unaltered. The fixed effect model of PLSC is then recalculated with the new permutation sample to generate new singular values. Repeating the process for several times (between 500 to 1000) will generate the probability distribution, which can then be used to test for statistical significance of the original singular value. The robustness of the entries $\boldsymbol{i}$ of salience vectors w_l can then be tested for reliable contribution towards correlation. This is done by dividing the individual elements $\boldsymbol{i}$ by their standard error. The standard error is calculated through bootstrap resampling. The process involves random resampling with replacement (between 500 to 1000) of the original observations of $\boldsymbol{X}$ and $\boldsymbol{Y}$. This is similar in nature to evaluation of z-score for each element.

Score = $w_{i,l} / \hat{\sigma}(w_{i,l})$ where $\hat{\sigma}(w_{i,l})$ is the standard error of the element $\boldsymbol{i}$ in salience $\boldsymbol{w_l}$. Elements with scores larger than 3 are considered stable contributors towards correlation.

PLSR

PLS regression is an iterative process where the latent vectors associated with the largest singular values are utilized to construct estimates of ***X*** and ***Y*** that are then used to deflate the original matrices. The process is repeated until full deflation of the input matrices or stopped after a limited number of latent variables based on other criteria related to performance of the prediction. The basic assumption is that the observations can be reconstructed from smaller number of latent variables.

The process starts the same as PLSC, where matrices of the saliences of ***X*** and ***Y*** (***W*** and ***C***) are computed from the covariance matrix $\boldsymbol{R=X^TY}$ (Equation-2.11 and 2.12). The brain and behaviour scores ($\boldsymbol{L_X}$ and $\boldsymbol{L_Y}$) are then computed from the projection of the ***X*** and ***Y*** onto their respective saliences (Equation-2.13). This is where the two processes part directions where for regression analysis, the first latent vectors associated with the largest singular value are used to reconstruct an estimate of the input matrices. The estimates are then used to deflate the original matrices.

$$\boldsymbol{t_l = X\, w_l} \quad \text{and} \quad \boldsymbol{u_l = Y\, c_l} \tag{2.14}$$

$\boldsymbol{t_l}$ *and* $\boldsymbol{u_l}$ will be the first columns of the iteratively built latent vectors (***T*** and ***U***) that are computed from the first columns of ***W*** and ***C*** saliences. Loading of the input data onto the first latent variable is then computed through the projection of ***X*** onto its respective latent vector. These matrices of first latent vector and loadings are then used to compute the first estimate of the original matrix.

$$\boldsymbol{p_l = X^T\, t_l} \tag{2.15}$$

$$\boldsymbol{\widehat{X}_l = t_l\ p_l^T} \tag{2.16}$$

$$\boldsymbol{\widehat{Y}_l = u_l\ c_l^T} \tag{2.17}$$

We then make the following relationship between the latent vectors of the dependant and independent variables ($\boldsymbol{u_l}$ *and* $\boldsymbol{t_l}$ respectively) that can be used to compute the regression coefficients to estimate ***Y*** from ***X***.

$$\boldsymbol{u_l = b_l\, t_l\ +\ e} \tag{2.18}$$

Where $\boldsymbol{e}$ is the residual which is ignored during the iterative process of computing the regression coefficients for each of the latent variables extracted at each iteration. We use the estimated values $\widehat{\boldsymbol{X}}_{\boldsymbol{l}}$ and $\widehat{\boldsymbol{Y}}_{\boldsymbol{l}}$ to deflate the original matrices and then repeat the process (Equation-2.11 and 2.12) to compute the remaining latent vectors starting with a new covariance matrix and singular value decomposition. Equations-2.16, 2.17 and 2.18 can be used to relate ***X*** and ***Y*** as follows:

$$\widehat{\boldsymbol{Y}}_{\boldsymbol{l}} = \boldsymbol{u}_{\boldsymbol{l}}\ \boldsymbol{c}_{\boldsymbol{l}}^{\boldsymbol{T}} = \boldsymbol{b}_{\boldsymbol{l}}\,\boldsymbol{t}_{\boldsymbol{l}}\,\boldsymbol{c}_{\boldsymbol{l}}^{\boldsymbol{T}} \tag{2.19}$$

Equation-2.19 shows the relationship between the dependant variables and the latent vectors $\boldsymbol{t}_{\boldsymbol{l}}$ of the independent variables. Using Equation-2.16, we can rewrite the estimate of ***Y*** as [73]

$$\widehat{\boldsymbol{Y}} = \boldsymbol{X}\,\boldsymbol{B}_{\boldsymbol{reg}} \quad \text{where } \boldsymbol{B}_{\boldsymbol{reg}} = \boldsymbol{P}^{\boldsymbol{T}+}\ \boldsymbol{B}\ \boldsymbol{C}^{\boldsymbol{T}} \tag{2.20}$$

$\boldsymbol{P}^{\boldsymbol{T}+}$ is the Moore-Penrose pseudo-inverse of $\boldsymbol{P}^{\boldsymbol{T}}$, ***C*** is the loading matrix built from the column matrices $\boldsymbol{c}_{\boldsymbol{l}}$, and ***B*** is the diagonal matrix constructed from $\boldsymbol{b}_{\boldsymbol{l}}$s. The obvious question with Equation-2.20 is the number of latent vectors ($\boldsymbol{l}$) to be used for constructing the matrices ***P***, ***C*** and ***B***. In PLSC, we used bootstrap resampling to estimate a z-score for each feature of the independent variables in the latent space. This method can still be applied in PLSR analysis to limit the number of input features to those that consistently contribute towards estimation of the dependant variable and generalization of the regression coefficients. Another method that can help towards generalization, is the use of Residual Estimated Sum of Squares (RESS), along with cross validation through leave-one-out technique [76]. The idea is to sequentially leave one observation out of the regression analysis, compute the regression coefficients and use them to estimate the left-out observation. The residual errors ($\left\|Y - \hat{Y}\right\|^2$) can then be summed to get an estimate of the regression performance.

Full deflation of the ***X*** matrix through iterative singular value decomposition and computation of latent vectors is a fixed model approach that might result in the best regression model for the given set of observations. This is useful when the given set of observations represent the complete (or majority of) experimental conditions. In general, however, the set of observation is a sample of much larger population that is being used as a random representative. In this case, using the full set of latent vectors will generally

result in overfitting and poor performance of estimation from new observations. One approach to avoid overfitting is to use leave-one-out technique at each iteration of adding a latent variable and assessing the resulting RESS values. The idea is to keep adding latent vectors as long as the RESS values continue to decrease. The idea being that the value of RESS will turn around when addition of latent variables steers towards fixed model. It is assumed that the lower the number of latent variables, the more generalized the regression model, but with potentially lower estimation performance.

2.3.5. Model Generation Flowchart

Error! Reference source not found. shows the processing steps involved in the g eneration of regression models for predicting motor function from functional connectivity measures. The processing steps are common for all three phases of this study.

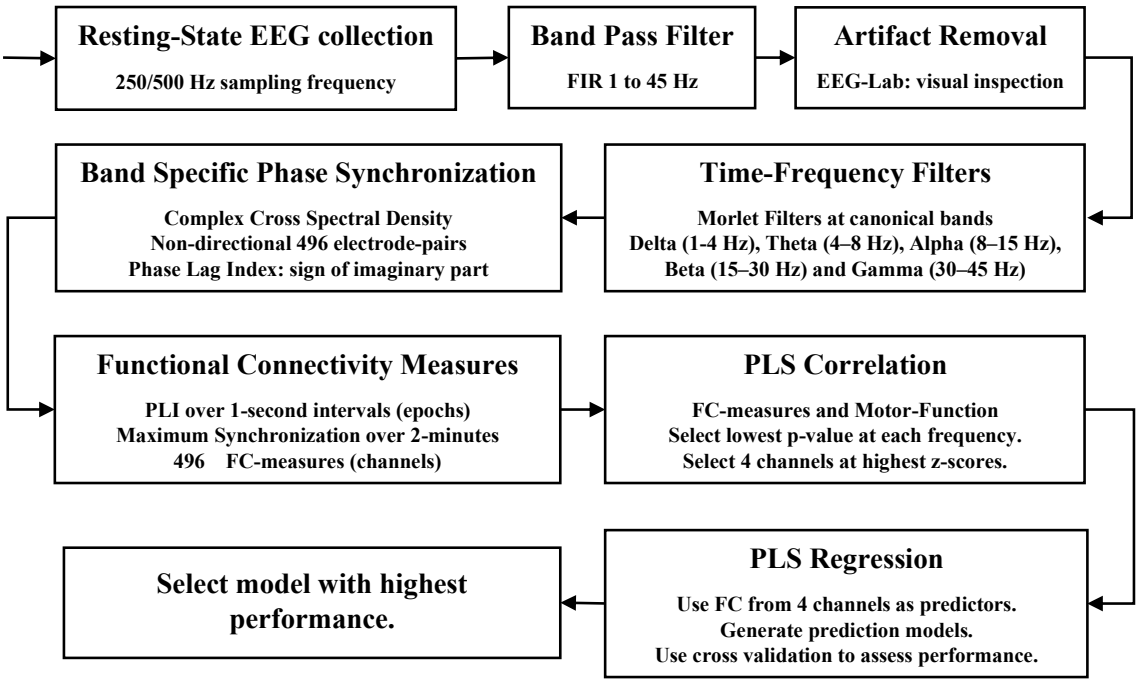

Figure 2.2 Flowchart representation of the processing steps from EEG data collection to generation of prediction model for estimating motor function.

2.4. Materials and Equipment

We used a 32-electrode gel-based EEG cap (g.Nautilus, g.tec medical engineering, Austria) with a standard 10-20 montage, operating at 500 samples per second, referenced to right earlobe and common ground at midpoint between FPZ and FZ. A single size EEG cap was used for all participants (Figure 2.3). Cap placement was based on centering the CZ electrode on the mid-point between left and right Pre-Auricular, Nasion and Inion locations. Conductive gel was applied under each electrode to reduce the contact impedance below 25 K-Ohm. All collected EEG data was imported into MATLAB-7.8.0 (MathWorks Inc) for further processing. EEG-Lab V14.1.2 [77] was used to visually inspect the EEG signals for the presence and removal of artifacts. We did not use independent component analysis (ICA) to remove these artifacts. The decision was made as a compromise towards allowing more processing time for the computation of functional connectivity measures. This was particularly important for the third phase of our study with respect to real-time analysis of EEG data during neurofeedback training.

For data collection during neurofeedback study we used a 32-electrode dry-EEG cap (g.SAHARA, g.tec medical engineering, Austria) with a standard 10-20 montage, operating at a sampling rate of 250 Hz for data acquisition (Figure 2.4). The reference electrode was placed on the right mastoid and the ground electrode on the left mastoid. Our rationale behind switching to dry caps during this phase of our study was purely for pragmatic considerations with respect to setup time associated with the gel-based EEG caps. The concern was that the long setup time could result in participant fatigue before the start of NFT program, which could negatively influence the participant's focus and consequently the effectiveness of mental imagery.

Implementation of all EEG preprocessing algorithms were based on excerpts from the on-line MATLAB scripts from Analyzing Neural Time Series Data [62]. We used the on-line MATLAB toolbox by McIntosh [74] for PLS processing.

Figure 2.3 **EEG gel-based electrode cap and acquisition device. g.tec medical engineering, GmbH. Brain-Computer Interfaces & Neurotechnology (www.gtec.at).**

Figure 2.4 **EEG dry electrodecCap. g.tec medical engineering, GmbH. Brain-Computer Interfaces & Neurotechnology (www.gtec.at).**

Chapter 3.

Estimating Motor Impairment from Functional Connectivity Measures

Material in this chapter is extracted, reproduced, and modified with permission from the following paper:

N. Riahi, V. A. Vakorin, C. Menon, “Estimating Fugl-Meyer Upper Extremity Motor Score from Functional-Connectivity Measures,” *IEEE transactions on neural systems and rehabilitation engineering*, vol. 28, no. 4, pp. 860-868, Apr. 2020.

3.1. Chapter Overview

Our goal for this part of the study was to propose a method that could provide an objective measure of motor impairment, while addressing the challenges associated with the availability of trained examiners to carry out the assessments. The latter was deemed to be conducive towards more frequent assessment of motor function, thereby enabling the adjustment of rehabilitation strategies and personalization of therapeutic activities. For pragmatic reasons, we focused on the use of EEG-based functional connectivity for this purpose. We aimed to answer the research question on whether resting state functional connectivity was a suitable neurophysiological measure to accurately estimate motor impairment in stroke survivors. Our objective was to propose a method based on measures of functional connectivity for estimating motor impairment that was evaluated through Fugl-Meyer assessment. A favorable result would further motivate the appraisal of the performance of the proposed method based on a more objective measures of motor function than Fugle-Meyer, and our goal for the second phase of the study.

This chapter explains the details of the methodology used towards selection and processing of neurophysiological features. Results pertaining to the performance of the methodology along with some limitations of the approach are presented.

3.2. Introduction

Fugl-Meyer assessment is an accepted method of evaluating motor function for people with stroke. A challenge associated with this assessment is the availability of trained examiners to carry out the evaluation. Neurophysiological measures show promise in addressing the above impediment. Our study investigated the potential of using resting state electroencephalographic (EEG) functional connectivity as a Neurophysiological measure for estimating Fugl-Meyer upper extremity motor score (FMU) in people with chronic stroke.

Prior work has shown connectivity with ipsilesional primary motor cortex (M1) to account for 78% of variance in motor impairment [28], addition of other measures such as corticospinal tract injury into the model were recommended to increase R^2 to 93%. The need to include other measures than connectivity alone in explaining the variance with motor function may have been due to the exclusion of other relevant connectivity measures between areas that do not include M1. We therefore propose a different approach that eliminates the *a priori* restriction on connectivity measures being tied to any specific seed location or any particular electrophysiological frequency band. We suggest using coherence as a measure of functional connectivity index between brain regions and use statistical analysis of these indices to identify the contributing brain regions that correlate with FMU at different frequencies. We further propose applying the connectivity indices from the identified regions and frequencies as regressors for estimating FMU.

To this end, we collected resting state EEG data from 10 individuals with stroke. Functional connectivity was evaluated through five different processing algorithms and quantified in terms of both average as well as maximum-coherence between EEG electrodes at 15 frequencies from 1 to 45 Hz. We applied a multi-variate Partial Least Squares (PLS) Correlation analysis to simultaneously identify specific connectivity channels (EEG electrode pairs) and frequencies that robustly correlated with FMU. We then applied PLS-Regression to the identified channels and frequencies to generate a set of coefficients for estimating the FMU. Participants were randomly assigned to a training-set of eight and a test-set of two. Cross-validation with leave-one-out approach on the training-set, using Phase-Lag-Index processing algorithm, resulted in an R^2 of 0.97 and a least-square linear fit slope of 1 for predicted versus actual FMU, with a root-mean-square error of 1.9 on FMU scale. Application of regression coefficients to the connectivity

measures from the test-set resulted in predicted FMU of 47 and 38 versus actual scores of 46 and 39, respectively. Our results demonstrated that the evaluation of neural correlates of FMU shows promise in addressing the challenges associated with the availability of trained examiners to carry out the assessments.

3.3. Method

3.3.1. Participants

Ten stroke survivors volunteered for this research. EEG recordings were collected as part of other studies [58] and shared with us to carry out this research. Table 3.1 shows the demographics of the participants. Based on the date of participation, the first 8 participants were arbitrarily allocated to a training-set and the last two participants to a test-set. Stroke participants were all in chronic phase with motor deficits specifically caused by their stroke. As part of the inclusion criteria, the participants were required to understand English and be able to communicate with experimenters. They were also required to have a minimum Montreal Cognitive Assessment Score of 23. Exclusion criteria included other neurological conditions besides stroke (i.e., Parkinson's disease, multiple sclerosis etc.), inability to move most-affected upper extremity, known musculoskeletal injury or conditions that affected the upper extremities and known history of epilepsy or seizures. The Research Ethics Board of Simon Fraser University approved the protocol for this study, and all participants signed informed written consent.

3.3.2. Protocol

EEG data was collected on the same day and immediately after carrying out Fugl-Meyer upper extremity motor assessment for each stroke participant. EEG recordings were carried out while participants were sitting comfortably upright with feet flat on the floor, calm, still and quiet, but awake. Participants were asked to keep their eyes closed while 2 minutes of resting state EEG data was collected. Participants were informed of the start of EEG recording.

Table 3.1 **Demographic data for all participants. SP represent stroke participant. MoCA is the Montreal Cognitive Assessment Score. FMU is the Fugl-Meyer upper extremity motor score for the affected hand of stroke participants. SP5 was with aphasia. Although he was only able to get a score of 15 in MoCA, he was evaluated by a professional physical therapist, who confirmed his ability to give consent and follow the instructions in the protocol of this study. Therefore, he was also included in this study.**

Participant ID	Gender	Age	Affected hand	Years after stroke	MoCA	FMU
SP1	F	62	R	5	26	41
SP2	M	64	L	6	24	47
SP3	M	80	R	11	23	39
SP4	F	50	R	6	23	46
SP5	M	39	R	11	15	18
SP6	M	64	L	3	27	45
SP7	M	75	R	1	25	49
SP8	M	64	L	0.5	23	24
SP9	M	66	L	8	24	46
SP10	F	78	R	1	23	39

3.3.3. EEG Data Pre-processing

Electrode data for stroke participants with left hand impairment were left-right flipped prior to any processing and analysis. We applied finite impulse response filter (1 to 45 Hz) to raw electrode data in both forward and reverse direction to obtain a zero-phase band-passed EEG data. The filtered data was then visually inspected in EEG-Lab V14.1.2 for the presence of noise due to muscular activities. The first 4 seconds of all recordings were excluded to allow for a more stable resting state phase, reducing the total duration of EEG data per participant to 116 seconds. Surface Laplacian filter was then applied to reduce the impact of volume conduction [66]. Surface Laplacian is a spatial bandpass filter that reduces the effects of broad spatial activities and improves topographical localization. We used the spherical spline method for this purpose [67].

3.3.4. Band Separation and Filtering

We used complex Morlet wavelets for bandpass filtering of EEG signals from each electrode. Morlet wavelet is a sinusoid wave within a Gaussian envelope. The center

frequency of the wavelet is determined by the frequency of the sinusoid and the bandwidth is determined by the standard deviation of the Gaussian envelope. The actual filtering process can be implemented by multiplying the Fourier Transform of the wavelet and EEG data from each electrode and then applying Inverse Fourier Transform to convert back to time domain. The resulting frequency-specific amplitude and phase information of the filtered EEG data can be used to compute instantaneous coherence between activities at different electrodes (see 2.3.2).

We considered five canonical frequency bands defined as Delta (1-4 Hz), Theta (4–8 Hz), Alpha (8–15 Hz), Beta (15–30 Hz) and Gamma (30–45 Hz) in EEG research. We further divided each band into low, medium and high sub-bands, resulting in 15 individual center-frequencies and 15 specific Morlet wavelets for bandpass filtering. The number of cycles were increased by 1 for each band, from 4 at Delta to 8 at Gamma band, resulting in individual bandpass filters that were approximately a bandwidth apart at consecutive center frequencies between 1 to 45 Hz. With 116-seconds trial at 500 Hz sampling rate and 15 frequencies, the filtering process generated 870,000 complex samples for each of the 32 electrodes.

3.3.5. Connectivity Measures Through Coherence

We used five different coherence algorithms to produce non-directional functional connectivity measures between different brain regions. Here, we assumed that each electrode represented the electrophysiological activities of a specific brain region. Implementation of all algorithms were based on excerpts from the on-line MATLAB scripts from Analyzing Neural Time Series Data [62]. Processing algorithms included:

- Spectral coherence represented by the cross spectral density normalized by auto spectral densities from each electrode [66], [60].
- Phase clustering represented by the magnitude of average phase angle differences between two electrodes over a selected period [68].
- Imaginary part of coherence is the same as spectral coherence but keeping only the imaginary part of the complex cross spectral density. It further minimizes the impact of volume conduction [69].

- Phase Lag Index (PLI), represented by the average of sign of imaginary part of cross spectral density. There is no power component in this algorithm [70].
- Weighted Phase Lag Index is PLI that is weighted by the instantaneous distance of the phase differences from the real axis [71].

For each frequency band, we calculated non-directional coherence measures for all possible electrode pairings, resulting in 496 (32*31/2) connectivity channels that represented the complete set of electrode-pairs for a 32-electrode EEG cap. Note that in our study we use the terms EEG electrode pairings, connectivity channels, or simply channels interchangeably. We partitioned the filtered EEG data from each participant into 1-second non-overlapping epochs before applying the above coherence algorithms at each frequency. We generated a single average-coherence value over the 500 samples in each epoch, resulting in 116 coherence measures for each channel at each frequency. One-second epoch was not long enough for Delta band coherence calculations. We therefore excluded Delta band from connectivity analysis at 1-second epochs. We did, however, repeat the analysis for 0.5, 2, and 4-second epochs and included the Delta band for 4-second epoch only.

We considered two separate processing steps to further reduce the 116 averaged-coherence measures to a single connectivity index for each channel and at each frequency for the complete 2-minute resting state trial. The first was to compute a single average value of the 116 coherence measures. The second was to extract the maximum value from all 116 measures. Thus, for each method, we generated 12 frequency-specific (Theta to Gamma) connectivity indices for each of the 496 channels, resulting in a total of 5,952 connectivity indices for each participant. We repeated the connectivity-index computation with epochs of 0.5, 2, and 4 seconds to investigate the impact of epoch length on connectivity-index.

3.3.6. PLS Analysis

Overview

We applied PLS analysis to examine the relationship between EEG functional connectivity and FMU. PLSC was used to study the manifestation of functional connectivity across specific frequencies and electrode-pairs (channels) [73]-[75] to identify the

contributing channels and frequencies towards correlation between FMU and functional connectivity indices. We used PLSR to predict FMU (dependent-variables) from functional connectivity measures (independent-variables) [76]. The latent variables were used to compute the regression coefficients. Overfitting was addressed by extracting the least number of latent vectors that accounted for most of the variance. The prediction quality was evaluated through coefficient of determination (R^2) and root-mean-squared-error (RMSE) that were evaluated from cross-validated leave-one-out approach. To avoid overfitting, we stopped adding latent variables as soon as the quality started to decrease.

We used the on-line MATLAB toolbox by McIntosh [74] for PLSC processing.

Workflow of analysis and effects of noise and multi-collinearity

Our feature space is multi-dimensional with potential multi-collinearity, for example, due to spatially close channels. We therefore started with a simulation process to assess the impact of multi-collinearity and noise on the performance of PLSR without any restrictions on the least-contributing channels. To achieve this, we constructed a feature matrix of artificial coherence measures: 8 rows (equivalent to 8 participants) and 60 columns (60 EEG features). To introduce collinearity in the feature space, we filled the first 10 columns with different linear combinations of the participants' actual FMU. We then applied noise by adding uniformly distributed random numbers to all 60 columns. Starting with only the first 10 columns (8x10 feature matrix), we calculated the R^2 and slope of linear fit (fit-slope) between predicted and actual FMU through leave-one-out approach. We repeated the process for progressively larger numbers of random feature columns and increased the magnitude of added noise. We then carried out 500 iterations of each evaluation to generate a box plot of the R^2 and fit-slopes.

The connectivity channels and frequencies that contributed towards correlation between FMU and functional connectivity indices of stroke participants were identified through PLSC analysis [75]. It is worth mentioning again that connectivity channels with respect to seed electrode or connecting electrodes were not restricted to motor cortex. We started with z-scores > 2.33 and leave-one-out approach on the training set, to find the most common set of contributing channels and frequencies. We repeated the analysis for each of the four epoch lengths (0.5, 1, 2 and 4-seconds).

We carried out PLSR on connectivity data from the stroke participants in our training set, using only the most contributing connectivity channels and frequencies that were identified through PLSC analysis. To reduce overfitting, we stopped adding latent vectors when the quality of prediction from leave-one-out approach began to decrease [76]. Similar to our simulation stage, we assessed prediction performance of the PLSR by evaluating the R^2 and fit-slope of predicted vs actual FMU through the leave-one-out approach. We also used RMSE to further qualify the fit-slope measures. Regression coefficients generated from the training set were then used to evaluate the FMU for the 2 stroke participants in our test-set. To further assess the impact of additional participants on the regression performance, we repeated the cross-validation using leave-one-out approach with all 10 participants as well as iteratively swapping the two participants in the test-set.

We used the on-line MATLAB toolbox from Abdi [76] for our PLSR analysis.

3.4. Results

3.4.1. Simulations: effects of noise and multi-collinearity

Figure 3.1 shows the effect of added noise and selection of features on the coefficient of determination (R^2) in a model used to simulate the situation wherein we manipulate the number of features selected for predicting FMU from EEG coherence. Median values of R^2 changed by less than 0.01 between 20% and 200% added noise, irrespective of percentage of random features. Percentage of added random features however, had a larger impact on the regression performance with over 10% drop in the median values of R^2 from no added random features to 80% random features.

Figure 3.2 shows the regression performance in terms of fit-slope between the predicted and actual FMU. Prediction performance as measured by fit-slope decreased by approximately 50% from no added random features to 80% random features. Impact of increased levels of added noise was similar to that of R^2 analysis.

3.4.2. FMU Correlation

We applied PLSC analysis to assess the correlation between the connectivity indices and FMU at different frequencies. Note that two measures of coherence (one for the average and one for the maximum across EEG segments) were tested with five algorithms for computing coherence itself. Using the average-coherence measure as the connectivity index for each channel at each frequency did not produce statistically significant results ($p > 0.05$) for any of the five coherence processing algorithms. At the same time, using the maximum-coherence values led to statistically significant ($p < 0.04$) correlation between FMU and connectivity indices in the Alpha frequency band.

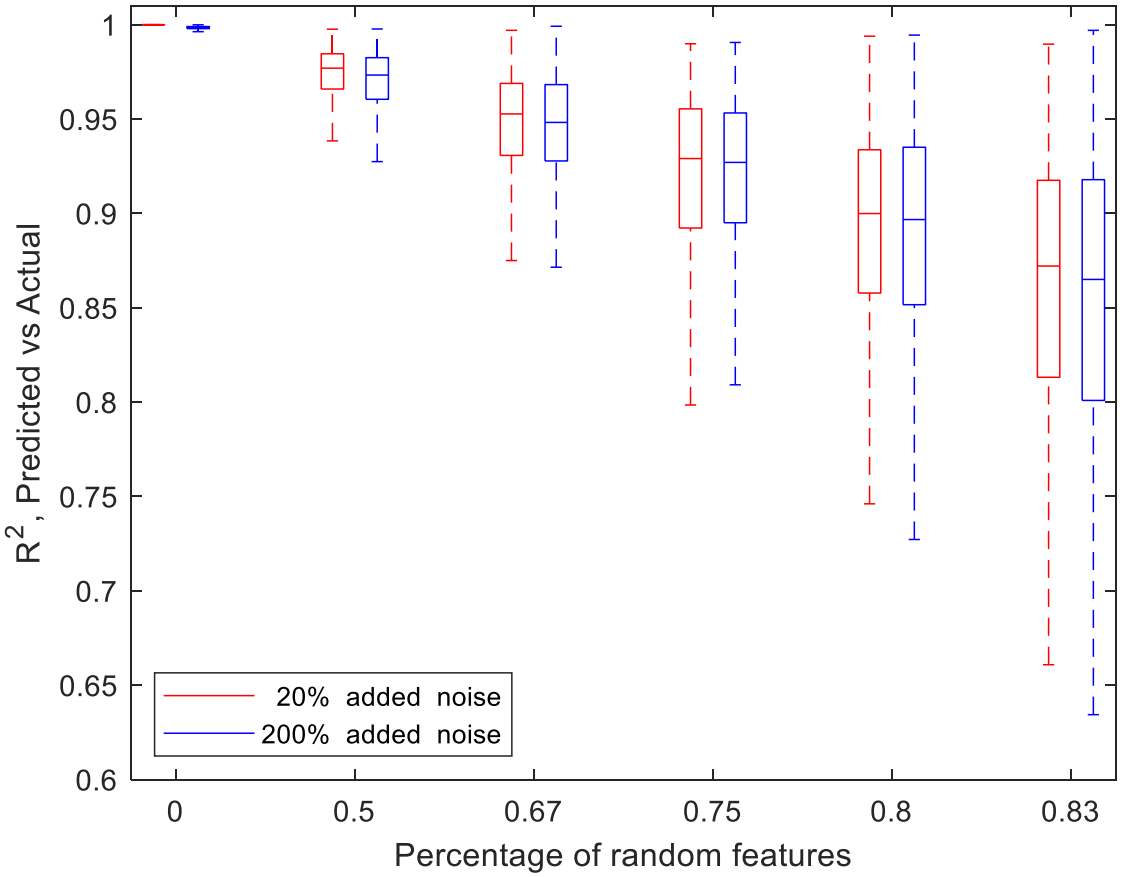

Figure 3.1 **Degradation in regression performance for simulated data, measured through R^2, for increasing levels of added noise at 20% and 200%. Percentage of random features represents the ratio of random features over total number of features. Random features had larger impact on R^2 than the level of noise. Results were generated from 500 iterations.**

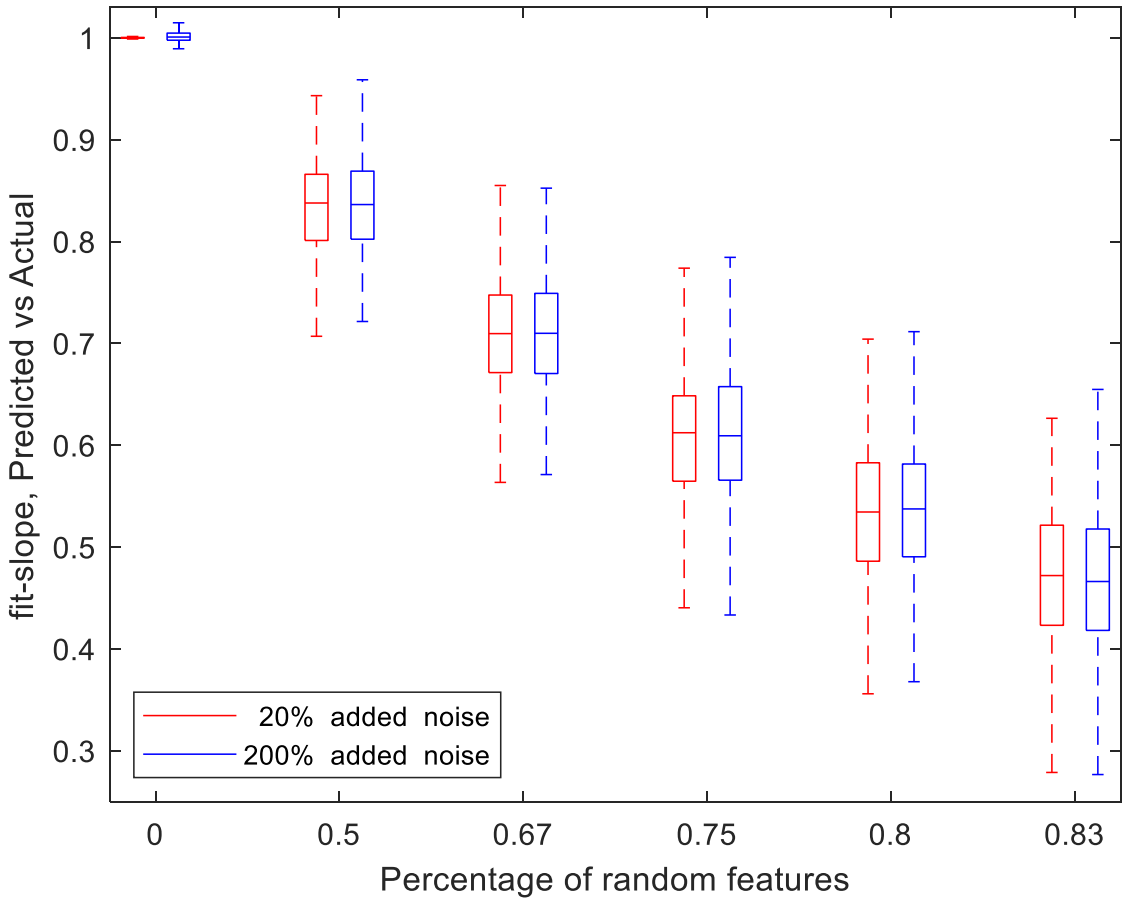

Figure 3.2 **Degradation in regression performance for simulated data, measured through fit-slope, for increasing levels of added noise at 20% and 200%. Percentage of random features represents the ratio of random features over total number of features. Random features had larger impact on fit-slope than the level of added noise. Results were generated from 500 iterations.**

Figure 3.3 shows the results of PLSC analysis using PLI processing algorithm and a z-score threshold of 2.33. There were more than 50 functionally connected channels at Alpha-band with z-scores greater than 2.33 that exhibited positive correlations with FMU. Using only the Theta, Beta, or Gamma band in PLSC did not produce a statistically significant correlation ($p > 0.05$).

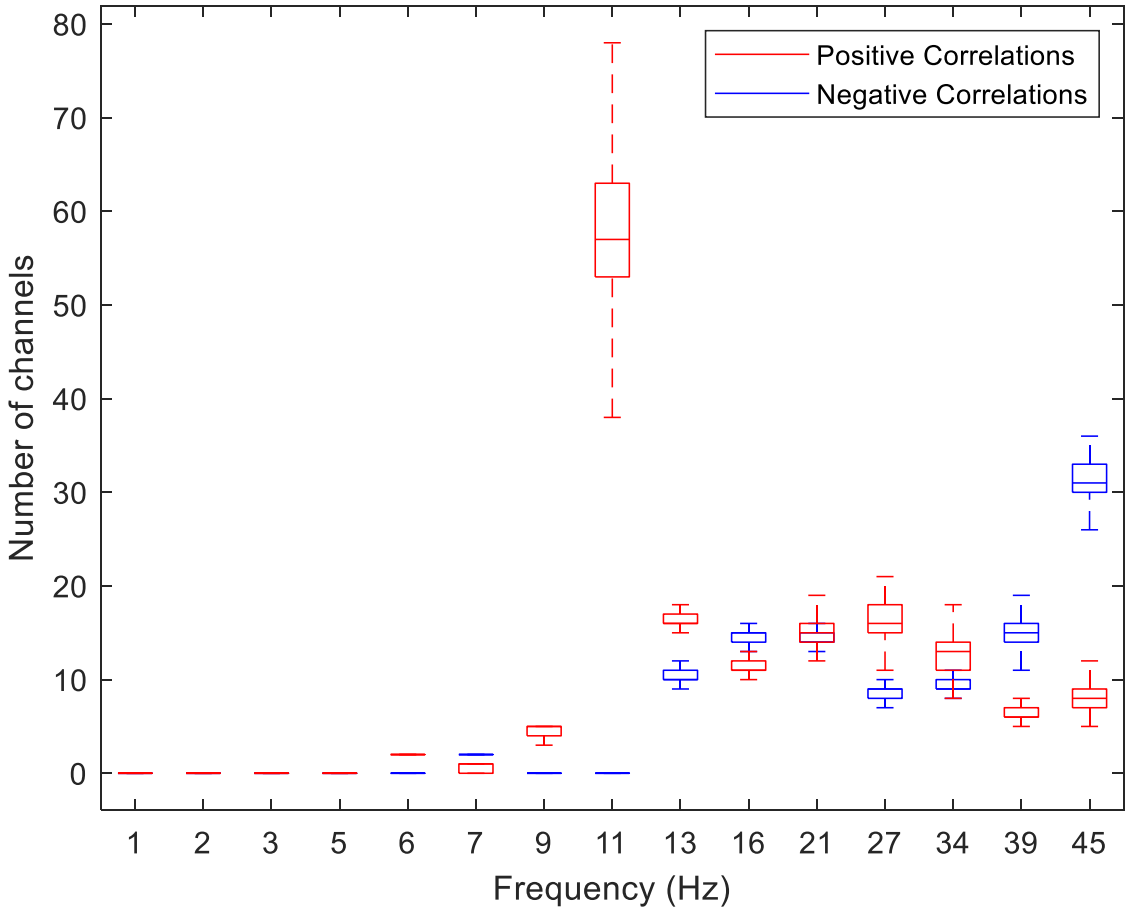

Figure 3.3 **Correlation analysis with 500 iterations showed statistically significant ($p < 0.04$) positive correlation between FMU and over 50 channels at 11 Hz. Number of channels (y-axis) represent the electrode-pairs with synchronized activities at each frequency (x-axis) and z-score > 2.33.**

Correlation analysis was carried out with 1-second epochs. There was no statistically significant correlation with 0.5, 2, and 4-second epochs for either the average or maximum-coherence approach ($p > 0.05$). We repeated the correlation analysis using leave-one-out approach and z-score threshold of 2.33 to look for the most common contributing channels at each frequency. We found channels F7-F3, FP2-F7, F8-C4 and FC2-CZ at medium Alpha frequency (11 Hz) to be the most robust (z-score > 4).

3.4.3. FMU Regression

Training-set

We used the four identified channels from our PLSC analysis and medium Alpha frequency (11 Hz), as regressors for our PLSR analysis. The leave-one-out approach provided an estimate of FMU for each stroke participant by using the regression coefficients produced from the remaining stroke participants.

Figure 3.4 shows the linear fit of predicted versus actual FMU. PLI processing algorithm produced the best prediction performance indicated by the combination of fit-slope and R^2 that were closest to 1 along with the lowest RMS-error, as seen from Table 3.2. The RMS-error of 1.9 is in Fugl-Meyer upper extremity scale and is calculated using the differences between actual and predicted FMU for each stroke participants within the training set. With an R^2 of 0.97, the statistical power analysis at α=0.05, 4 predictors and a desired power level of 0.8 indicated a minimum sample size of less than 8. This was consistent with the number of participants in our training-set.

Table 3.2 FMU prediction results from PLSR through cross-validation using leave-one-out approach on the training-set. Regression coefficients were computed for the four connectivity channels FP2-F7, F7-F3, F8-C4 and FC2-CZ at medium Alpha frequency (11 Hz). Fit-Slope represents the slope of least-square linear fit between the predicted and actual FMU with the corresponding RMS-errors. The latter is in Fugl-Meyer scale for upper extremity (maximum 66). Predicted data in bold-text are the best results and were obtained from PLI processing algorithm and maximum-coherence as a measure of connectivity index.

Participants		SP1	SP2	SP3	SP4	SP5	SP6	SP7	SP8			
Actual FMU		41	47	39	46	18	45	49	24			
Predicted FMU	Processing									Fit-Slope	RMS-error	R^2
	Spectral coherence	27.7	41.9	41.7	62.6	35.7	40.2	30.6	2.4	0.904	14.3	0.26
	Imaginary part of coherence	66.1	37.2	63.9	48.7	-3.9	30.2	39.0	39.3	1.038	17.3	0.30
	Phase clustering	12.1	39.8	58.9	44.0	35.6	29.5	45.0	20.0	0.882	15.3	0.08
	Phase Lag Index (PLI)	**40.8**	**47.2**	**38.5**	**47.5**	**15.8**	**45.8**	**46.7**	**27.8**	**1.002**	**1.9**	**0.97**
	Weighted PLI	45.1	46.3	42.6	46.6	-24.7	46.1	45.7	20.8	0.949	15.3	0.83

The last step in regression analysis was to generate a set of regression coefficients for the four identified channels (F7-F3, FP2-F7, F8-C4 and FC2-CZ), using all stroke

participants in the training set. The corresponding regression coefficients for each of these channels were 88, 80, -28 and -5 respectively.

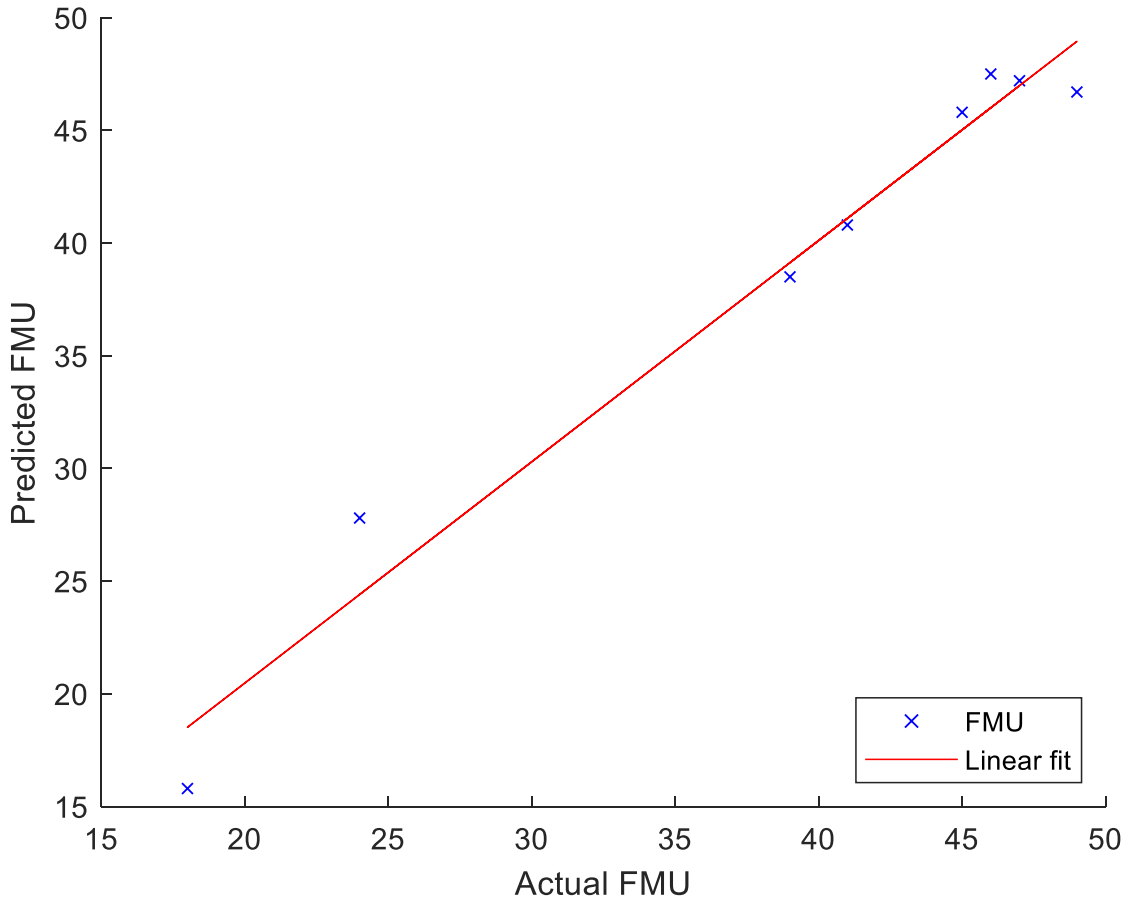

Figure 3.4 **Linear fit of predicted versus actual FMU from leave-one-out approach on the training-set. Predictions were obtained from PLSR analysis, using PLI processing and maximum-coherence measures as connectivity index for channels FP2-F7, F7-F3, F8-C4 and FC2-CZ at Alpha band (11 Hz).**

Test-set

PLSR analysis on the training-set using the PLI processing algorithm and the four identified connectivity channels (F7-F3, FP2-F7, F8-C4 and FC2-CZ) at medium Alpha frequency (11 Hz) generated a set of regression coefficients that were then used to predict the FMU of the stroke participants in the test-set. We applied the regression coefficients

to the connectivity indices of SP9 and SP10 from the same four connectivity channels and at the same medium Alpha band frequency, resulting in predicted FMU of 47 and 38 compared to the actual FMU of 46 and 39 respectively. The prediction error of ±1 point on Fugl-Meyer scale for upper extremity was below the RMS-error of 1.9 obtained from cross-validation on the training-set.

Full participants

Cross-validation with the same four channels (F7-F3, FP2-F7, F8-C4 and FC2-CZ) using all 10 participants resulted in an R^2 of 0.91 and a fit-slope of 0.991. Using the regression coefficients from these four channels to estimate the FMU resulted in an RMS-error of 2.1 on the Fugl-Meyer scale. Repeating the PLSC analysis with the 10 participants resulted in a reduced number of contributing channels that included F7-F3, FP2-F7, F8-C4. Using only these three channels to assess the regression performance through cross-validation with all 10 participants resulted in a similar R^2 of 0.91 and a fit-slope of 0.991. Using the regression coefficients from these three channels to estimate the FMU resulted in an RMS-error of 2.0 on the Fugl-Meyer scale. Further limiting the channels to F7-F3 and FP2-F7 resulted in better prediction at the lower extremes of motor scores (prediction of 17.7 for an actual FMU of 18,). Iteratively swapping the two participants in the test-set resulted in a similar R^2 of 0.91 and a fit-slope of 0.99. It is noteworthy that a statistical power analysis with a reduced R^2 of 0.91 resulted in a similar minimum sample size of 8 for a desired power level of 0.8 with 4 predictors.

3.5. Discussion

The main objective of this study was to investigate the suitability of resting state functional connectivity measures as biomarkers for predicting FMU. Using the primary motor cortex as the seed location for connectivity analysis, Wu et al [28] showed good prediction performance from ipsilesional M1–Premotor connectivity measures within Beta band. For our investigation however, we proposed carrying out a global consideration of functional connectivity that was not restricted to motor areas of the brain as the seed location. Nor did we want to limit our search to Beta band as the primary frequency of

interest. We selected PLSR for predicting FMU based on its performance in the presence of multicollinearity and noise [76]. Through simulation, we showed that PLSR was sensitive to large numbers of random features and that its performance degraded with increasing numbers of unrelated functional connectivity measures. This was an inherent challenge with our suggestion of using global functional connectivity measures that were not restricted to *a priori* selection of seed location and frequency bands. We showed that PLSC was a suitable approach [74] to identify the contributing channels and frequencies that correlated with FMU, thereby allowing us to exclude the unrelated features that were unfavorable to the performance of regression. Using only the four identified contributing channels and frequencies for our PLSR analysis on the training-set, we found cross-validated prediction performance with R^2 of 0.97, and linear fit-slope of 1 between the predicted and actual FMU, and an RMS-error of 1.9 on Fugl-Mayer scale. The RMS-error was below the minimum detectable change of 3 for Fugl-Meyer upper extremity motor assessment [9]. Using all 10 participants resulted in a slight reduction in cross-validated R^2 to 0.91 and a fit-slope to 0.99 but also a reduction in the number of contributing channels to 3. Using the regression coefficients from these three channels to estimate the FMU resulted in an RMS-error of 2.0. Iteratively swapping the two participants in the test-set resulted in similar fit-slope of 0.99 and R^2 of 0.91. We argue that our proposed approach has merit in not only improving the prediction performance but is also informative in identifying brain regions that could have been overlooked through *a priori* restriction of seed locations.

There are two impacting factors at play when using PLSC to identify the contributing brain regions and frequencies; one is the selection of features that might corelate with FMU and the second is the robustness of the correlating features in explaining the covariance. Our study was focused on resting state functional connectivity index as the feature of interest. The index, however, can be computed in different ways. We implemented a two-stage computation of index. First, we calculated the average coherence in each epoch. For the second stage, we implemented two different approaches; one was to average over all epochs [38] and the other was to use the maximum-coherence from all epochs. We are not aware of any current publications that use this second method. Permutation test on the results of PLSC analysis showed statistically significant correlation ($p < 0.04$) for Alpha band with the maximum-coherence method under our experimental setup of 1-second non-overlapping epochs. In this study,

we did not investigate the reasoning behind significant correlation at 1-second epochs only. We speculate that coherence occurs in bursts of certain duration and increasing the epoch size would reduce the maximum-coherence through the process of averaging within each epoch. Smaller epochs have the opposite effect, by reducing the signal to noise ratio and potentially more spurious maximum-coherence. Our interest was to find robust synchronization between regions and not necessarily long duration of synchronization. Permutation test with 0.5, 2 and 4-second epochs did not produce statistically significant results ($p > 0.05$).

Robustness of contributing features in PLSC analysis can be examined through bootstrapping and the resulting z-score for each channel at specific frequencies [75]. Using leave-one-out approach during PLSC analysis and inspecting the most common channels and frequencies along with the corresponding z-scores can help select the predictors for regression analysis. Using this approach, we found that a z-score ≥ 4 identified the most robust contributing channels and frequencies. Considering our consistent results from the experimental approach, there may be merit in further research on an analytical method of choosing a threshold for the z-scores to find the corresponding contributing channels from the training-set.

We found F7-F3, FP2-F7, F8-C4, FC2-CZ and the central frequency of the alpha-band as the most robust contributing factor towards a correlation between the connectivity indices and FMU, under our experimental conditions. The corresponding regression coefficients for each of these channels were 88, 80, -28 and -5 respectively. We can make several observations about the channels and corresponding coefficients. The first observation is the strong positive influence of increased connectivity index between F7 and both F3 and FP2 towards higher FMU. This is interesting when considering the general approach of constraining the seed region for connectivity measures to primary motor cortex [28]. If we had followed that initial constraint, we would have missed these two channels as predictors of FMU. Furthermore, assuming the ipsilesional primary motor cortex to be represented by C3, then contribution of connectivity with this motor area at Alpha band under our experimental conditions was less relevant to FMU prediction than the connectivity between prefrontal and frontal areas. The second observation is the interhemispheric nature of FP2-F7 channel. Higher level of interhemispheric functional connectivity in the prefrontal areas seems to contribute to higher FMU. This relationship appears to agree with published studies [12], [78], although not specific to primary motor

cortex. The third observation is the negative regression coefficient associated with higher contralesional functional connectivity with motor areas (F8-C4). At first glance, this seems confusing when considering that PLSC analysis did not produce robust negative correlation between F8-C4 channel and FMU. One explanation is that the covariance of connectivity index from this channel with FMU exhibits positive correlation, meaning that connectivity index varies in unison with FMU, but that its influence as a regressor is more inhibitory towards the other channels and regulates the extent of increase in FMU due to increase connectivity index between F7, F3 and FP2. One last observation is the positive influence of the ipsilesional functional connectivity in the frontal area (F7-F3) towards higher FMU. It will be interesting to understand the importance of this observation from a neuroscientific point of view. It is noteworthy that the topmost robust contributing channels (F7-F3, FP2-F7, F8-C4) were consistently identified through PLSC analysis when increasing the number of training set to 10 participants.

With respect to clinical relevance and applicability, identification and elimination of least-contributing channels from FMU prediction has a favorable impact on cost and setup time of the EEG system. We estimate that an upper extremity motor function assessment could be completed in less than 15 minutes when using a 7-electrode EEG system to collect 2 minutes of resting state data. Moreover, resting state analysis is a preferential strategy due to its least dependence on the physical abilities of individuals with stroke, especially when suffering from severe motor impairment [14]. Furthermore, use of biomarkers can reduce the possible subjectivity of assessments, remove the potential complications associated with language barriers, and minimize the requirement for examiners' subject matter expertise. When considering these factors, then our proposed approach (if proved consistent through further research with a larger number of participants) would provide significant incentives for clinical use. The EEG system can be operated with minimal training and at a capital cost that could potentially be recovered after a limited number of assessments in a clinical setting or at the point of care. A limitation of this study is that participants were all in the chronic phase of stroke. Further research with individuals in acute and sub-acute phases is necessary to explore the implications on the contributing channels or regression coefficients.

Our study was a proof of concept with limited number of participants and no information on the extent of structural damage to the brain. The only available information on the stroke participants were related to the affected hand. The latter was used to swap

the EEG measures between the two brain hemispheres. Lack of information on the extent and location of structural damage could play a biasing role with respect to potential subgroups within the stroke participants. FMU measures had a mean of 38.6, median of 40, minimum 18 and maximum 49 on FMU scale. A training set with more stroke participants at the extreme ends (FMU < 25 and FMU > 50) would be beneficial towards a more informed statistical analysis. There are different views of the extent of floor and ceiling effect at these extreme ends. Previous work has shown floor and ceiling effect when using FMU for stratification or to predict outcome [9]. In this study, using the maximum-coherence as a measure of connectivity index does have the potential to introduce ceiling effect and needs to undergo further research with more participants at the upper ends of FMU (> 50). If further work at the upper end of motor scores point to a ceiling effect, then separate regression models could be applied for the upper extremes of FMU. Our analysis also revealed larger RMS-error at the lower extremes of FMU, especially when iteratively swapping the participants in the test-set. The error at the lower extremes between actual and predicted FMU was reduced (actual=18, predicted=17.7) when we applied a different regression model that only used the first two contributing channels, namely F7-F3 and FP2-F7. These results may indicate the need for using different models for the lower extremes, middle and upper extremes of FMU.

This study did not investigate the minimum detectable change that is achievable through Partial Least Squares approach for predicting FMU. This is an important topic for future research on this approach. If minimum detectable change with PLS approach is smaller than that of Fugl-Meyer assessment, then the former will have the potential for use in monitoring the efficacy of rehabilitation at shorter intervals and personalizing the techniques as need be. This is a plausible prospect, since a measure of motor deficit that is independent of physical motion may provide a higher resolution of assessment and deliver intermediate and non-visually-observable appraisal of motor function.

To summarize, we investigated the potential of applying resting state functional connectivity measures as biomarkers for estimating Fugl-Meyer upper extremity motor scores. We selected maximum-coherence as a measure of connectivity index and employed Partial Least Squares Correlation to identify the electrode pairs and frequencies that correlated with motor scores. We then used Partial Least Squares Regression to generate coefficients for predicting the motor scores. Regression coefficients were generated from 8 stroke participants in our training-set and applied to 2 stroke participants

in our test-set for predicting their motor score. Application of regression coefficients to the connectivity indices from the test-set resulted in predicted scores of 47 and 38 versus actual Fugl-Meyer upper extremity motor scores of 46 and 39, respectively.

Chapter 4.

Estimating Longitudinal Change in Motor Skill from Individualized Functional Connectivity Measures

Material in this chapter is extracted, reproduced, and modified with permission from the following papers:

N. Riahi, R. D'Arcy, C. Menon, "A Method for Estimating Longitudinal Change in Motor Skill from Individualized Functional-Connectivity Measures," *Sensors* 2022, DOI: 10.3390/s22249857.

4.1. Chapter Overview

In the first phase of our study, we introduced a method based on PLS analysis of rsFC for estimating motor impairment in stroke survivors. We argued that addressing the challenges associated with the availability of trained examiners, ease of use, and the accuracy of estimation could be conducive towards frequent assessments of motor impairment and personalization of therapeutic activities. However, there is little incentive for frequent assessments if they can only measure large changes in motor function, which might take a long time to achieve. FMU is an example of such assessments, where a point change in FMU score would correspond to a large change in function. Performance of our proposed method was quantified based on its accuracy in estimating FMU. To gain a better understanding of the estimation accuracy of the method, its performance needed to be evaluated against a more objective measure of motor function that could be quantified at small incremental changes as compared with FMU. The second phase of our study aimed to answer the research question on whether rsFC could be used to estimate small incremental changes in motor function. Our focus was on individualized approach as a precursor towards application of neurofeedback for skill improvement, and the last phase of our research.

This chapter presents the design of an objective motor skill training and assessment program, along with the performance of Individualized models in estimating the resulting longitudinal change in motor skill.

4.2. Introduction

Pragmatic, objective, and accurate motor assessment tools could facilitate more frequent appraisal of longitudinal change in motor function and subsequent development of personalized therapeutic strategies. Brain functional connectivity (FC) has shown promise as an objective neurophysiological measure for this purpose. Longitudinal motor learning studies with healthy participants using functional magnetic resonance imaging showed that some areas of the brain exhibited a transient change in FC while other areas showed a more lasting change towards consolidation and long-term retention [31], [32], [79]. This persistent change in FC has the potential for quantifying, and subsequently estimating the change in motor skill. Prior work using EEG modality showed correlation between rsFC and motor learning [33], with several studies focusing on the ability of generalized global configuration of FC to estimate skill acquisition in healthy participants [34], [35], [36], [37], [38], [39]. The authors showed interaction between a wide range of brain networks at different synchronization frequencies for motor learning. Involvement of different networks and frequencies along with differences across subjects due to age or existing capabilities, motivates an individualized approach towards evaluation of correlating FC measures. We advocate the use of EEG-based resting-state FC (rsFC) measures to address the pragmatic requirements. Pertaining to appraisal of accuracy, we suggest using the acquisition of motor skill by healthy individuals that could be quantified at small incremental change. Computer-based tracing tasks are a good candidate in this regard when using the spatial error in tracing as an objective measure of skill.

In this study we explored a longitudinal motor skill training program involving a computer-based tracing task, in which healthy participants used a computer mouse with their non-dominant hand to trace a predetermined and non-trivial pattern on a computer screen. The aim was to improve motor skill by attempting to reduce tracing error over a period of six to eight training sessions spread over as many days. Tracing error was used as an objective measure of behavioral performance that could be quantified at small incremental change. Resting state EEG data were collected before and after each training session and used for evaluation of rsFC. We then investigated the accuracy of an individualized PLS method in estimating longitudinal change in tracing error from changes in rsFC. Longitudinal data from participants yielded an average accuracy of 98.2% (standard deviation of 1.2%) in estimating tracing error. The results show potential for an

accurate individualized motor assessment tool that reduces the dependence on the expertise and availability of trained examiners, thereby facilitating more frequent appraisal of function and development of personalized training programs

4.3. Method

4.3.1. Study Design

Workflow

Figure 4.1 shows the experimental workflow for data collection and analysis. Tracing errors during physical training were used to appraise the change in acquired motor skill.

Setup

We used a 32-electrode gel-based EEG cap (g.Nautilus, g.tec medical engineering, Austria) operating at a sampling rate of 250 Hz, and OpenVibe v2.2 for data acquisition and storage. The reference electrode was placed on the right earlobe with the ground electrode located midway between FZ and FPZ. We applied enough conductive gel to maintain contact impedance below 30 K-Ohm. Recorded EEG data were imported into MATLAB-7.8.0 (MathWorks Inc) for signal processing and analysis.

Python 3.7 was used to create our experimental track patterns on a computer screen. We opted for elliptical trajectories instead of straight lines [36] to increase the degree of difficulty for tracing tasks. The track was constructed from eight quarter-ellipses that were arranged to form a four-section curved-pattern as shown in the inset (top-left corner) of Figure 4.2. In this figure, the green dot represents the starting point for the placement of the mouse pointer, before tracing the corresponding track towards the red dot. Selection of the active track section was controlled through the program as explained later in the protocol. Total position-error while tracing a track section was determined by the area (in pixels) between the actual pointer trajectory (Trace) and the desired track path (Track) as shown in Figure 4.2.

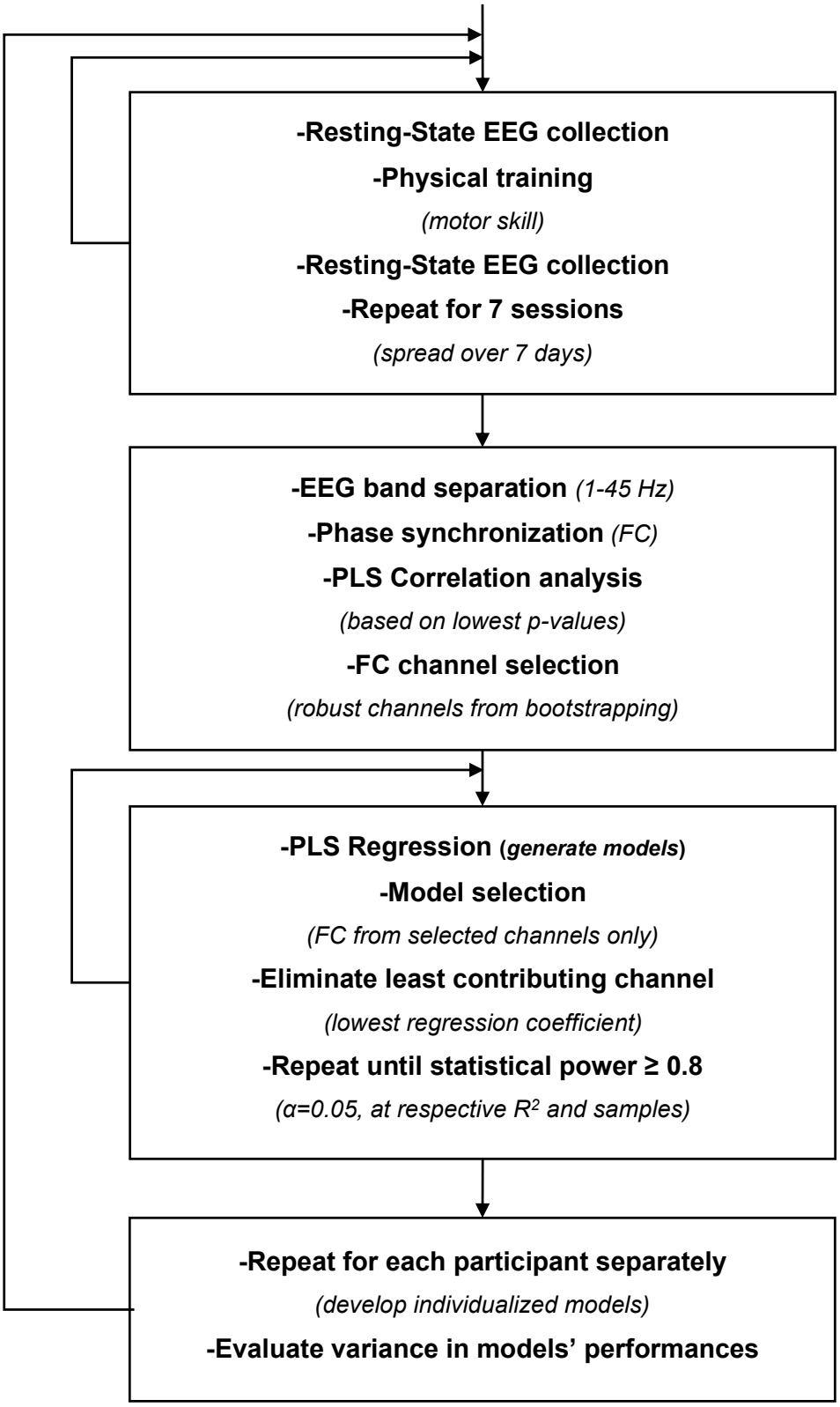

Figure 4.1 Experimental workflow for data collection and analysis.

Participants

Seven healthy right-handed participants, HP1 through HP7, (Mean Age=38.7, SD=21.5, 3 females) volunteered for the research study. Participants had no known neurological conditions, artifact inducing implants, or physical condition that would exclude them from the study. The Research Ethics Board of Simon Fraser University approved the protocol for this study, and all participants signed informed written consent forms.

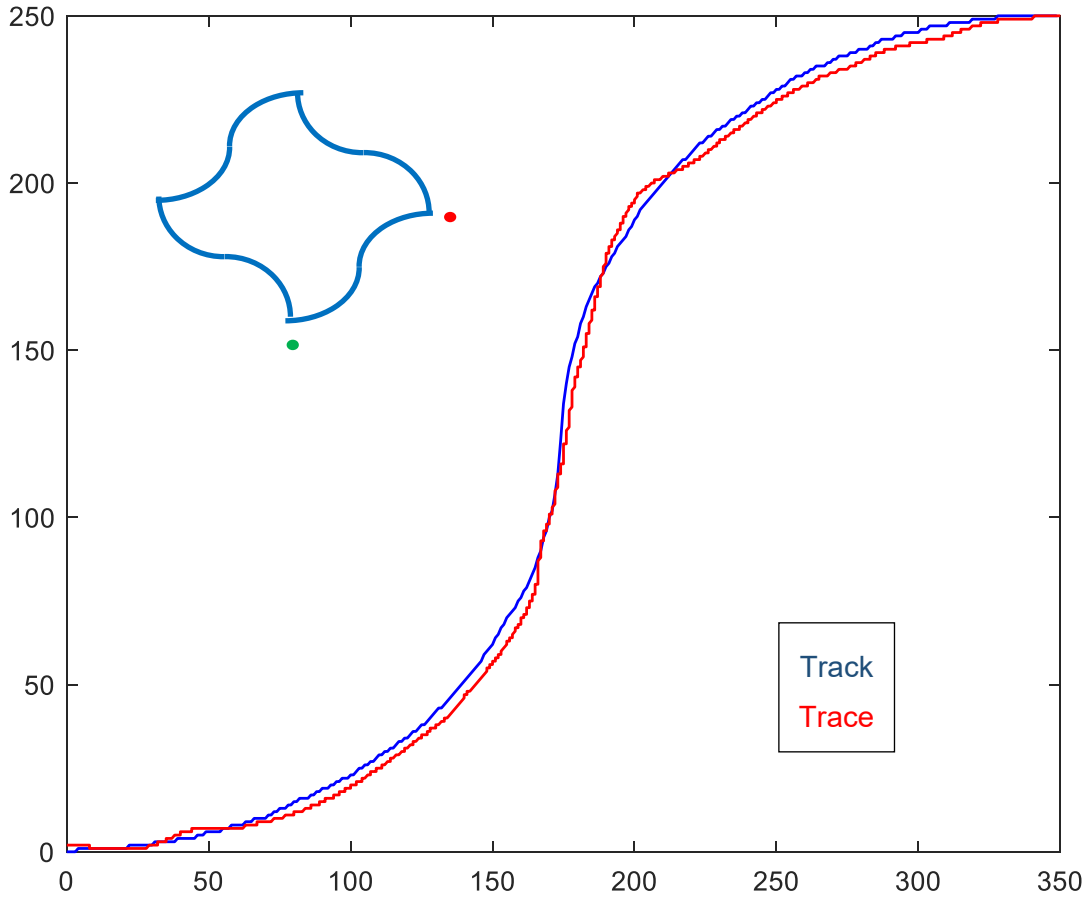

Figure 4.2 **Participant's tracing trajectory over a track section. The position-error is the total area between the Trace (red) and Track (blue) section. The inset at the top-left corner shows the complete track pattern. The green and red dots identify the active track to be traced. Axes are in units of screen pixels.**

Protocol

Every participant completed a longitudinal experiment that included 7 sessions, limited to a single session per day over two weeks. Each session consisted of four phases. There were no breaks in between or during each phase unless the participant specifically asked for one due to fatigue.

- Phase-1: The pre-training 5-minutes resting state EEG data collection. Participants were asked to sit comfortably upright with their feet flat on the floor, still and quiet with their eyes closed, but awake. Participants were notified of the start of EEG recording.
- Phase-2: an 8-trial test, with each trial including a tracing task with the right hand (dominant hand). Data from Phase-2 were collected but not used in this study.
- Phase-3: 90-trial training, with each trial including a tracing task with the left hand (non-dominant hand). These trials were over a randomly selected section of track and direction of tracing. Participants were asked to move the mouse on the table using only their arm and not their torso. Participants were prompted to move the mouse-pointer to the vertex identified by the green dot and instructed to trace the track towards the vertex with the red dot (Figure 4.2). They were asked to trace quickly and accurately, without compromising one for the other. To complete each trial, participants had to keep the pointer at the destination vertex for one second. This would penalize performance indicators when moving too fast to stop at the destination vertex.
- Phase-4: post-training 5-minutes resting state EEG data collection.

Each session lasted between 70 to 90 minutes depending on EEG setup time and the participants' tracing speed during Phase-3.

4.3.2. Tracing Performance

Position-error between the participants' tracing trajectory and the intended track pattern was selected as one of the performance indicators. Time taken to trace each section (trial) was selected as another indicator. Total position-error during a trial was determined by the area between the trace and track trajectories (Figure 4.2). To discourage participants' attempt to reduce position-error by tracing slower, we used each

trial time as a penalizing (multiplication) factor to inflate the respective position-error. Conversely, to discourage participants' attempt to reduce trial time by moving the mouse too quickly to stop at the destination vertex, we accumulated positional offsets from the track endpoint until the pointer came to rest at the destination vertex.

We used both accumulated position-error as well as the accumulated product of position-error and its corresponding tracing time from each trial, as two separate indictors of tracing performance. Changes in motor skill between sessions were reflected in variations in the magnitude of one or both of these measures. Selection of two metrics was our attempt in addressing the differences in participants that were more focused on tracing error rather than the tracing speed (or vice versa). We speculate that these may involve different interacting brain regions. Participants were updated on their tracing performance after each training session. Position-error was measured in units of pixels-squared, representing the area between the track and trace trajectories, and converted to spatial units of squared-centimetres (cm^2) based on an estimated coverage of 0.25 mm^2 per pixel. The product of position-error and its associated tracing time was measured in units of cm^2seconds.

Each training session consisted of 90 tracing trials, generating 90 intermediate performance values. We used two different approaches to produce a measure of tracing performance for each session: First, a single value corresponding to the median of all 90 trials (single-median option), and second, the median of the first 30 trials as a measure of performance before training, and the median of the last 30 trials as a measure of performance after training (dual-median option). The two measures allowed for separate evaluation of consolidated and short-term learning.

4.3.3. EEG Data Processing and PLS Analysis

Longitudinal EEG data from each participant were processed independently of the other participants, making this an individualized assessment rather than a cross-sectional or inter-participant analysis. We used an EEG sampling frequency of 250 Hz and a frontend finite impulse response bandpass filter of 1-45 Hz. We defined five canonical frequency bands specified as Delta (1-4 Hz), Theta (4–8 Hz), Alpha (8–15 Hz), Beta (15–

30 Hz) and Gamma (30–45 Hz) for our second stage filters and band separation. We further divided each band into low, medium, and high sub-bands, resulting in 15 distinct center-frequencies and 15 different Morlet wavelets for bandpass filtering [62]. Center-frequencies were approximately one bandwidth apart. We used coherence as a measure of functional connectivity and evaluated the instantaneous coherence through five different algorithms, namely Phase-Clustering, Spectral Coherence, Imaginary part of Coherence, Phase Lag Index (PLI) and weighted PLI [62]. Coherence was evaluated for every combination of electrode-pairs (496 channels) at each of the 15 center frequencies. Resulting coherence measures from each algorithm were then separately averaged over 1-second non-overlapping epochs, generating a total of 300 samples for each channel and at each frequency. We used only a 2-minute section starting at an offset of 30 seconds from the beginning of EEG data for PLS analysis. We used peak-detection over the 2-minute window to generate a single connectivity index for each channel. Both single-median and dual-median behavioral measures were used for correlation analysis with the connectivity indices. We selected the median of all 90 trials for each session as the behavior data (tracing performance) associated with that session for the single-median option. These were used for correlation analysis with connectivity indices from both the pre- and post-training EEG data separately. For the dual-median option, correlation analysis was carried out between the median of the first 30 trials and the connectivity indices from the pre-training EEG, and the median of the last 30 trials with that of post-training EEG.

PLS Correlation (PLSC) was used to identify the most robust channels and frequencies that contributed towards correlation between the connectivity indices and the selected measure of tracing performance [73]. We selected the options with the lowest *p*-values for this purpose. We repeated the process 50 times for each participant and selected the channels that were consistently present in over 80% of the repetitions and at the same level of robustness. We used permutation for quantifying the *p*-values and bootstrap to test for robustness. The resulting connectivity indices for the identified channels and frequencies were then used in PLS-Regression (PLSR) to produce a model for estimating the corresponding tracing performance. The estimation accuracy was examined through a cross validated leave-one-out approach over the seven training sessions and quantified by the root-mean-square-error (RMSE) in estimation. The estimation error was presented as a percentage of the average tracing error obtained from

the selected single- or dual-median option. The ratio of estimation RMSE over average tracing error allowed for comparison between estimation accuracy from different tracing performance measures obtained through position or position-time error. Data from participants with multiple combinations of contributing channels and frequencies were passed through iterative PLSR analysis. The goal was to find the specific combination that resulted in the smallest RMSE in estimating the tracing performance with the lowest number of contributing channels that resulted in a statistical power of 0.8 at $\alpha=0.05$ for the number of available samples for each participant.

4.4. Results

4.4.1. Data Collection

We carried out 53 training sessions, collected over 500 minutes of EEG data and 4700 tracing trials from seven participants. Data from 4 sessions were discarded due to external noise and technical recording issues. Except for HP1 and HP2, all other participants had 7 training sessions with 70 minutes of EEG data and 630 tracing performance measurements for each participant. Participant HP2 had to stop after 6 sessions and participant HP1 volunteered for 8 sessions.

4.4.2. Tracing Performance

Figure 4.3 shows the results of longitudinal training for participant HP1 with position error (blue bars) and position-time error (red bars) from all 90 tracing trials. The bullseye represents the median value. The results for the first and last 30 tracing trials were similar in nature and were omitted for the sake of clarity and space. Note that a change in one metric without a corresponding change in the other, could be used to identify and isolate the active metric. For example, a participant that focuses primarily on the position error may achieve a change in this measure of performance by slowing the speed of tracing. This could potentially appear as very little overall change in position-time measures. The opposite situation would correspond to a participant that focuses primarily on speed with

less attention to positional error. These activities may involve different groups of interacting brain areas, resulting in different models for predicting the tracing performance. Figure 4.4 shows the longitudinal training results for the remaining participants HP2 to HP7 but limited to the final selection of the position or position-time errors after PLSR analysis (Table 4.1).

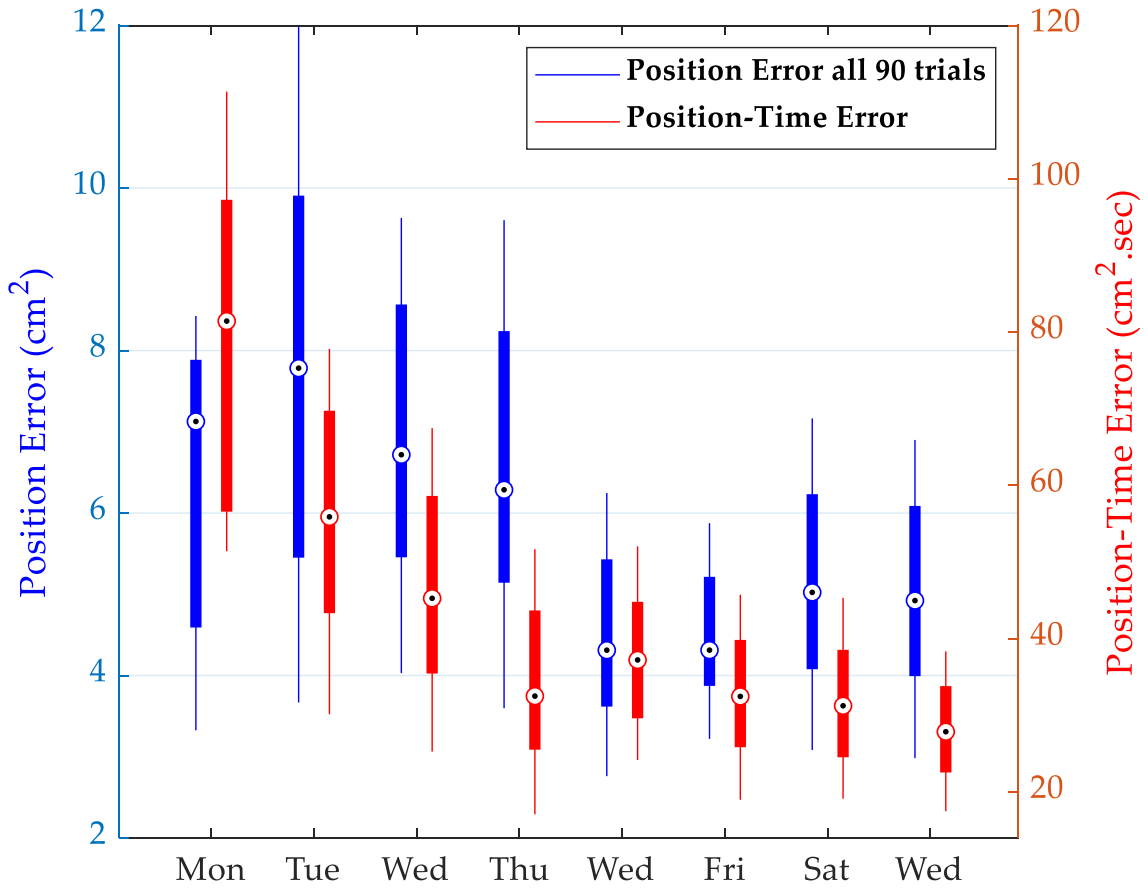

Figure 4.3 **Longitudinal tracing performance for HP1 during 8 sessions of physical training program with 90 trials in each session. Performance is in terms of position error (blue bars) and product of position error and time (red bars). The bullseye indicates the median value. Corresponding results from the first and last 30 trials were similar in nature and were excluded for the sake of clarity.**

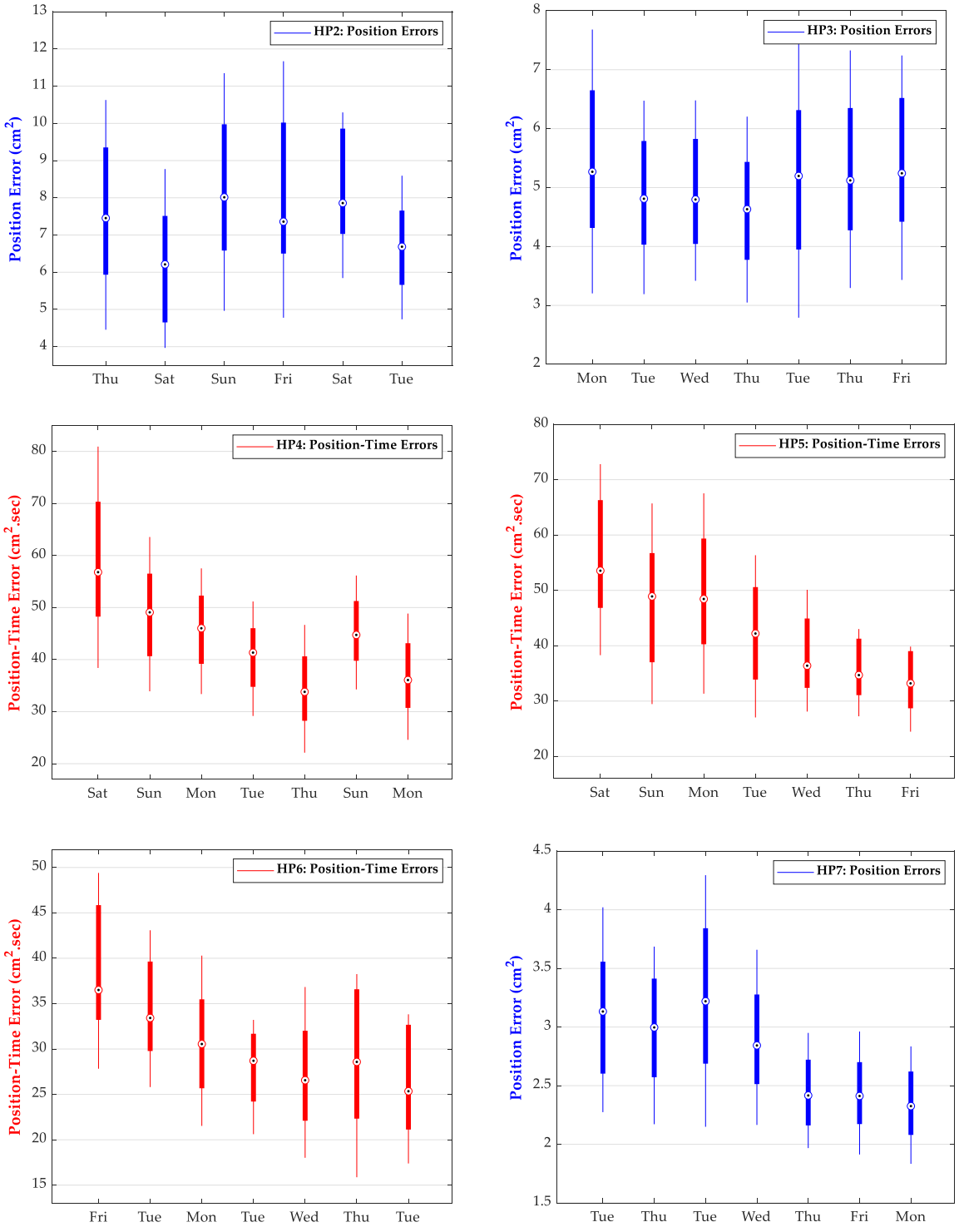

Figure 4.4 **Longitudinal tracing performance in terms of Position Error (blue bars) or Position-Time Error (red bars) for participants HP2 to HP7. Selection of performance measure is based on the final PLS analysis as shown in Table 4.1.**

4.4.3. PLS Analysis

Synchronization values using the PLI algorithm, averaged over 1-second non-overlapping epochs, and peak-detected across a 2-minute EEG interval at a start-offset of 30-seconds, resulted in strong correlation (p < 0.05 uncorrected) between the tracing performance and connectivity indices. For most participants, PLSC analysis generated multiple combinations of promising correlation between rsFC extracted from pre- or post-training EEG data and different tracing performance measures. We used bootstrapping to identify the most robust channels and center frequencies that contributed towards the correlation with tracing performance. PLSR analysis was constrained to these channels and frequencies. We generated a regression model for each combination and applied cross-validated leave-one-out approach to evaluate the estimation RMSE, which was subsequently used to represent the estimation accuracy. We then selected the model with the lowest RMSE as the best performing estimator of the change in motor skill for that participant. RMSE was presented as a percentage of the average tracing error for the respective participant. This allowed for comparison between estimation accuracy from different tracing performance measures obtained through position or position-time error.

Table 4.1 PLS analysis of the rsFC from pre- or post-training EEG data. PLSC was used to identify the contributing channels and frequency bands that correlated with the tracing performance at p-val < 0.05 (uncorrected). Selection of channels were based on robustness of contribution, evaluated through bootstrapping. PLSR was used to generate the estimation model. The number of channels were further reduced iteratively to obtain a statistical power of 0.8 at α=0.05 for the estimation model. RMSE is calculated through cross-validated leave-one-out approach for each model.

Participant	EEG	Freq. Band	Tracing Performance	Channels	R^2	RMSE (%)
HP1	Pre-training	Beta-High (27 Hz)	Single-median Pos-time	5	0.988	4.29
HP2	Post-training	Alpha-High (13 Hz)	Dual-median Pos-only	3	0.987	1.06
HP3	Pre-training	Beta-Med (21 Hz)	Single-median Pos-only	2	0.868	1.81
HP4	Post-training	Beta-Med (21 Hz)	Single-median Pos-time	4	0.996	1.15
HP5	Pre-training	Beta-Low (16 Hz)	Dual-median Pos-time	4	0.997	0.98
HP6	Pre-training	Beta-Low (16 Hz)	Dual-median Pos-time	4	0.984	1.60
HP7	Pre-training	Beta-High (27 Hz)	Single-median Pos-only	3	0.955	2.98

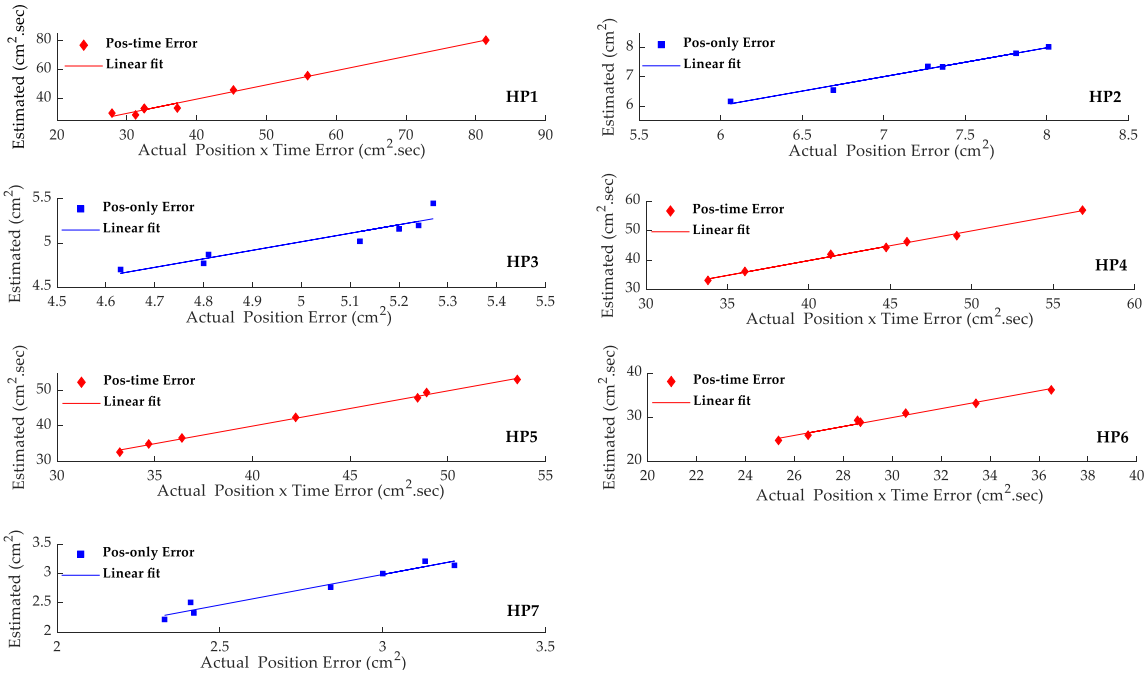

Figure 4.5 **Linear fit of the estimated versus actual tracing performance from cross-validated leave-one-out approach for each participant. Estimations were obtained from PLSR analysis, using PLI processing and peak-coherence measures as connectivity index (rsFC).**

Regression coefficients represent the contribution of corresponding channels (predictors) at the identified frequency bands towards estimating the tracing performance (behavior) from rsFC indices. The small number of training sessions (samples) for each participant has a negative impact on the statistical power of our analysis. We therefore aimed to reduce the number of predictors to counter the effects of our small sample size. The goal was to reduce the channel count to the maximum number of predictors that resulted in a statistical power of greater than 0.8 at α=0.05, and the R^2 and the number of samples (sessions) for the respective participant. To achieve this, we iteratively removed the channel with the lowest regression coefficient, based on the argument that these channels would have less impact on the overall estimation accuracy compared to the channels with the higher regression coefficients. Table 4.1 shows the channel count and the corresponding R^2 that resulted in a statistical power of greater than 0.8. Figure 4.5 shows the results of the cross-validated leave-one-out approach in graphical form. The linear-fit in each graph was generated by the line-of-best-fit using least-squares method. We set the y-intercept to 0 under the assumption that the line should pass through the

origin. Mean slope of the Linear-fit was at 0.997. Removing this constraint did not introduce a change in mean slope of the Linear-fit.

4.5. Discussion

The primary objective in this study was to investigate the accuracy of a proposed individualized method for estimating modest incremental changes in motor skill from objective neurophysiological measures. The methodology was expected to address the technical and pragmatic challenges associated with appraisal of motor function to facilitate more frequent assessments. rsFC was selected as the objective neurophysiological measure to allow for the evaluation of wide-ranging network interactions while eliminating the need for execution of physical tasks during assessments. We focused on EEG systems as the measurement modality of choice due to their portability, relative low cost, and requiring a minimal level of the examiners' expertise and time. These were critical design requirements for a tool that could potentially facilitate more frequent individualized assessments of motor function and subsequent development of personalized intervention strategies. We opted to use computer-based tracing tasks to evaluate spatial error in tracing as an objective measure of motor skill that could be quantified in small incremental changes. PLSC analysis was applied to limit the number of contributing EEG channels for estimating change in skills, followed by PLSR analysis to build a model for estimation. We further reduced the number of predictors (channels) in our estimation model to maintain a statistical power of ≥ 0.8. The proposed approach resulted in an average estimation accuracy of 98.2% with a standard deviation of 1.2%. At this level of accuracy in estimating objective measures of behavior, the proposed method may have the potential to provide intermediate valuations of motor function from subjective measures of behavior that inherently have larger margins of error [28], [55], [9].

PLSC analysis resulted in multiple combinations of behavioral and neurophysiological measures for each participant. The identified channels (electrode-pairs) for each participant were more consistent for rsFC at similar frequency bands and using the same tracing performance option (Pos-only or Pos-time). The latter may indicate the involvement of different brain areas when combining the precision of tracing (Pos-only error) with corresponding speed of tracing (Pos-time error) as a measure of change in

motor skill. Similarly, synchronization at different frequency bands may indicate an alternative group of neural networks associated with different aspects of skill improvements [16, 80, 81, 64]. However, our assessment was focused on evaluating the estimation accuracy of the proposed method rather than identifying the contributing brain areas. Source localization algorithms may have to be applied to better understand the relationship between synchronization frequency and contributing brain areas [16, 82]. Prior studies with healthy participants had shown rsFC as a predictor of skill acquisition and the extent of global connectivity as indicators of future motor performance [36, 38]. This is consistent with our results with per-training rsFC as predictors of tracing errors. We expected that training related changes in rsFC would consolidate several hours after the training session [63, 83] and would therefore relate to the next pre-training tracing performance. We can see this from the results in Table 4.1 for five of the participants. The relationship with post-training performance (HP2 and HP4) might be worth further investigation in future studies with larger number of participants.

Concerning the measure of tracing performance, we selected the first and last 30 trials to quantify the pre- and post-training skill levels. This was done based on a cursory examination of the variance in tracing performance over the course of all training sessions. Our analysis revealed larger variation in tracing performance for smaller number of trials (< 30) during the earlier sessions than later in the training program. We used the ratio of the standard deviation of tracing performance over the corresponding median of those trials (SD-Ratio) for this analysis. This indicated that during earlier sessions when tracing skill was less developed, we needed more trials to get a reasonable estimate of tracing performance. But as training progressed towards later sessions, the participants could trace more consistently, thereby requiring smaller number of trials. The change in SD-Ratio between 30 and a lower number of trials during these later sessions were relatively small, and as such, we opted to use the first 30 trials as a more stable measure to evaluate the pre-training tracing performance throughout the training program. To maintain consistency, we used the last 30 trials to measure the tracing performance for post-training. It should be noted that in some cases, we observed a reduction in tracing performance during the last 30 trials despite having practiced over the prior 60 trials. This may reflect a confounding factor related to fatigue. A more comprehensive assessment of performance variation over different trial-count selections for pre- and post-training may

reveal additional information and is warranted for future studies specific to motor skill improvement.

Our objective for this study did not necessitate a trendline for changes in motor skill, nor a statistically significant change in the acquired skill throughout the physical training sessions. Our aim was to investigate the presence of a relationship between rsFC and motor skill (measured through median tracing performance) and whether the relationship could be captured with a regression model to accurately estimate the individuals' tracing performance from their respective rsFC. We also did not carry out any cross-participant analysis of motor performance, as our objective was to develop an individualized assessment method and not a generalized model to investigate commonality between participants. We speculated that the ability to acquire motor skill was highly individualized and expected to see large differences between participants with respect to the extent and rate of change in skill. The results in Figure 4.4 are in line with this speculation and support our motivation for development of individualized models for estimating motor skill.

Inherent in our analysis is the dependence on passage of time. The stable changes in functional connectivity have been attributed to the development of specialized neural circuits for fast and efficient execution of tasks. These changes are not necessarily all associated with an increase in recruitment but also the opposite, indicating that some networks may have become either more efficient or less important for the respective motor skill acquisition [32, 13]. We can make similar interpretations about the sign of regression coefficients in our study to indicate an increase or decrease in network recruitment. The longitudinal impact of contributing channels is expected to change as motor skill improves over time. Having a simple and cost-effective tool might facilitate an opportunity to monitor and track the individualized time dependent changes in contributing channels by examining the change in their respective regression coefficients. We expect these changes to be gradual and suggest that new individualized models could be developed based on a rolling reassessment of the rsFC as new longitudinal samples become available. This will also allow for the progressive exclusion of older samples and the associated channels that are less contributing towards the regression model, thereby maintaining the statistical power of the regression analysis. It may also be possible to expedite this temporal change in the regression models by influencing the synchronization

levels of the contributing channels through brain stimulation, be it external or through endogenous techniques such as mental imagery and real-time neurofeedback.

To summarize, we investigated the accuracy of an individualized PLS processing technique for estimating an objective measure of change in motor skill. We used resting-state EEG to evaluate functional-connectivity indices and position errors from computer-based tracing tasks as a measure of motor skill. We carried out a longitudinal motor skill training program in which seven right-handed healthy participants used a computer mouse with their non-dominant left hand to trace a pattern on a computer screen. Each participant went through six to eight training sessions spread over as many days. We used PLSC to identify the contributing channels specific to each participant and PLSR to develop an individualized model for estimating the longitudinal change in motor skill of the respective participant. Using leave-one-out cross validation technique, we observed an average root-mean-square estimation error of 1.8% corresponding to an average estimation accuracy of 98.2% (standard deviation 1.2%). Considering the pragmatic advantages of using EEG-based resting-state functional connectivity measures for estimating longitudinal change in motor skill, the proposed method shows potential towards an objective and accurate motor assessment tool.

Chapter 5.

Using Neurofeedback to Guide Mental Imagery for Improving Motor Skill

Material in this chapter is extracted, reproduced, and modified with permission from the following papers:

N. Riahi, W. Ruth, R. D'Arcy, C. Menon, "A Method for Using Neurofeedback to Guide Mental Imagery for Improving Motor Skill," *IEEE transactions on neural systems and rehabilitation engineering* 2022, DOI: 10.1109/TNSRE.2022.3218514.

5.1. Chapter Overview

The second phase of our study established the performance of the proposed method in estimating small longitudinal change in motor skill. We argued that the accuracy of the method showed promise for frequent assessment of motor function and personalization of rehabilitation strategy. It also showed potential for monitoring short-term incremental interactions between changes in FC and motor skill. Our focus thus far was on quantifying the change in FC as a consequence of change in motor function. The next step was to consider the reverse relationship and assess the impact of changing FC towards inducing a change in motor function. This was inline with our research aspiration for using this method to facilitate a complementary therapeutic activity and our objective to Investigate the prospects of influencing the individualized FC measures through MI for improving motor function.

This chapter details a process for the selection of individualized FC channels and real-time monitoring of instantaneous FC measures. A method for utilization of neurofeedback to guide mental imagery as an endogenous brain stimulation is then presented. The chapter concludes with the results from applying the proposed individualized method for improving motor skill.

5.2. Introduction

Mental imagery (MI) is gaining attention as a strategy towards endogenous brain stimulation for improving motor skill. Neurofeedback (NF) is commonly used to guide MI in order to activate the relevant brain networks.

Specific to motor function, prior work showed change in motor skill by influencing regional brain activities [45], [46], [47], [48], [49], [50]. In these studies, the focus was to regulate the sensorimotor rhythms (SMR) through MI, based on the hypothesis that modulating a specific band power over motor areas would influence the related motor behavior. NF on the instantaneous strength of SMR was used to inform the participants about the impact of their MI. However, studies using functional magnetic resonance imaging have shown that MI can result in overlapping activation of different brain areas including primary motor, premotor, supplementary motor, and parietal areas that include the sensorimotor and posterior parietal lobe [19]. It has therefore been argued that analysis of network interactions might be a more wholistic approach for NF implementation [42], [51], [52]. In a study with both healthy participants [53] and stroke survivors [54], the authors showed that an NF protocol based on increasing the global alpha-band FC with the primary motor cortex resulted in improvement in motor function. These are significant findings and encouraging results in favor of using NF to facilitate an endogenous self stimulation of brain towards improving motor skill. The targeted frequency band and brain networks for FC analysis is not limited to alpha-band or motor areas. Prior study with healthy participants showed interaction between networks at both alpha and beta-band that included FC with prefrontal cortex [57] for motor learning. Involvement of different brain networks and synchronization frequencies along with differences across subjects due to age or existing capabilities, motivates an individualized approach towards NF training [55], [56]. We therefore investigated an NF training method based on individualized FC analysis between different brain networks at multiple frequencies.

To this end we used a longitudinal physical motor skill training (PT) program involving a computer-based tracing task. We then investigated an individualized EEG-based method for NF through broad consideration of interactions between different brain networks. We selected the change in brain functional connectivity (FC) as an objective neurophysiological measure of change in motor skill during a longitudinal PT program.

Digital tracing tasks were developed for skill training and the spatial error in tracing was used to gauge the change in skill.

Spontaneous FC was estimated from coherence measures [18] using resting state EEG data collected as part of the physical training program. We used partial least squares algorithms [38], [73] to find the most robust contributing networks (EEG electrode-pairs or channels) towards correlation between the resting state FC and the acquired motor skill. We used the network with the largest margin for increasing FC as the candidate for NF training while experimenting with MI during a neurofeedback training program. The participant was informed of the changes in instantaneous FC through real-time audio feedback to help guide volitional control of the connectivity indices with the goal of improving motor skill without execution of additional physical training. We showed over 20% reduction in tracing error through neurofeedback training alone, without any additional physical training. We also showed retention of improvement in skill for several days after the completion of neurofeedback training. Our proposed methodology shows promise for a highly individualized approach towards improvement in motor skill. Given that EEG is an accessible health and wellness technology, such a method could provide a practical complementary option towards personalized therapeutic strategies to improve motor function.

5.3. Method

5.3.1. Study Design

Workflow

Figure 5.1 shows the experimental workflow for data collection and analysis. The three motor skill assessments after the end of physical training program were carried out to appraise any change in skill without physical training. Similarly, the four assessments after the end of neurofeedback training were carried out to test for retention of improvement in motor skill. We balanced the requirements of quality data collection with the participant engagement and optimal performance. The latter was an important condition with respect to NF, as task demands have direct impact on the success of MI (e.g., attention/distraction). The different repetition numbers during successive stages of

this study were selected to address practical training requirement that balanced the trade-off between optimal learning and fatigue.

Setup

We used a 32-electrode dry-EEG cap (g.SAHARA, g.tec medical engineering, Austria) operating at a sampling rate of 250 Hz for data acquisition. The reference electrode was placed on the right mastoid and the ground electrode on the left mastoid. Our rationale behind switching to dry caps during this phase of our study was purely for pragmatic considerations with respect to setup time associated with the gel-based EEG caps. The concern was that the long setup time could result in participant fatigue before the start of NFT program, which could negatively influence the participant's focus and consequently the effectiveness of mental imagery.

Python 3.7 was used to create elliptical track patterns on a computer screen as depicted in the inset of Figure 5.2. Participant was asked to use the computer mouse to trace the track section between the vertices identified by the green and red dots, respectively. The active track was selected randomly as explained later in the protocol. Distance between opposing tips of the track pattern was arranged to be approximately 35 cm of mouse travel across the torso (x-axis: left to right) and 25 cm away from the torso (y-axis: top to bottom). The rationale was to deliver large physical movements that engaged multiple arm joints without the need to involve the torso. Motor skill was measured in terms of position error between the track and the participant's tracing trajectory, as well as the time taken to complete the corresponding tracing task. Position error for each tracing task was quantified as the area, in pixels, between the active track and tracing trajectory as shown in Figure 5.2.

Participant

A healthy right-handed 61-year-old female volunteered in order to test the methodological feasibility for this study. The participant had no known neurological or physically limiting conditions, and no implants. The Research Ethics Board of Simon

Fraser University approved the protocol for this study, and the participant signed an informed written consent form.

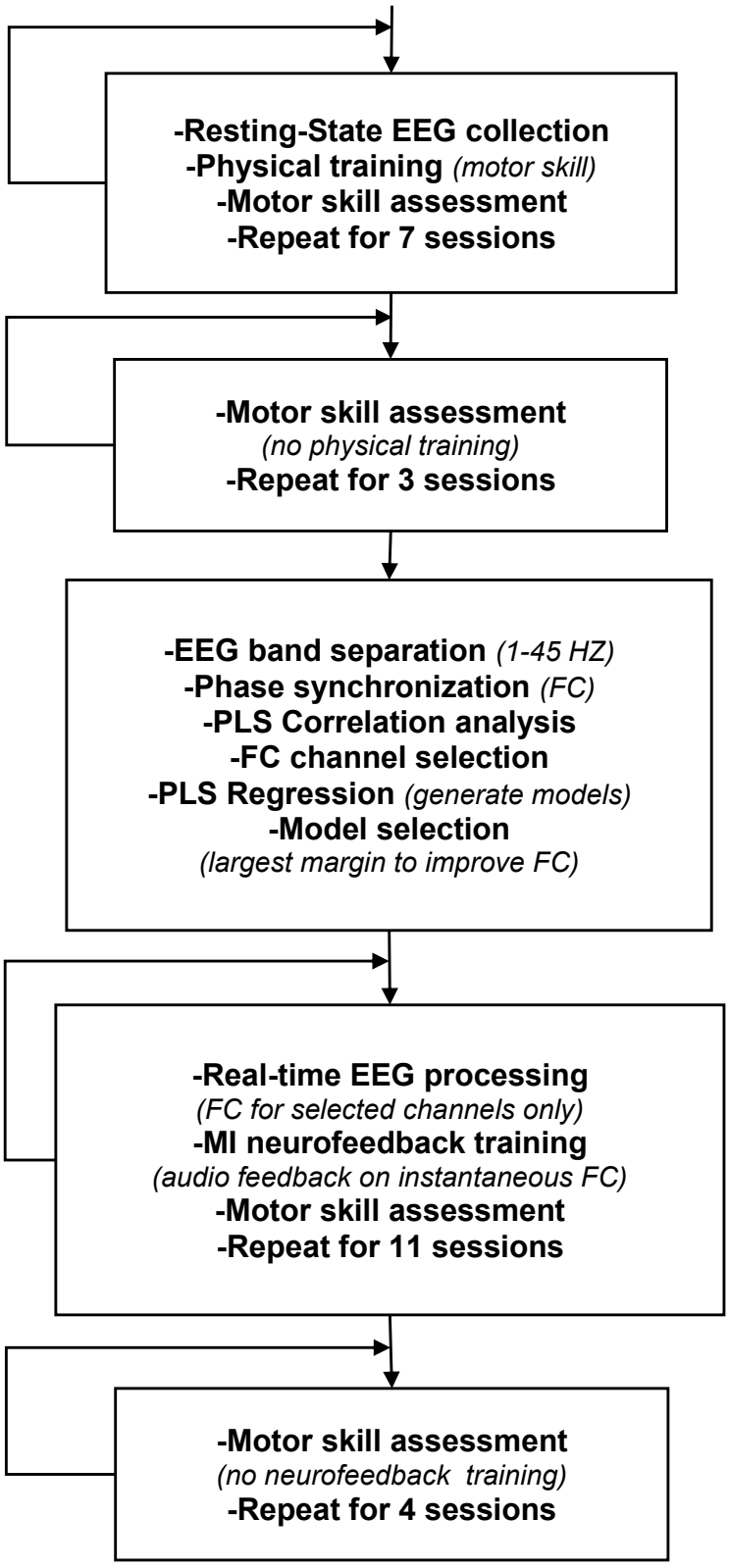

Figure 5.1 Experimental workflow for data collection and analysis.

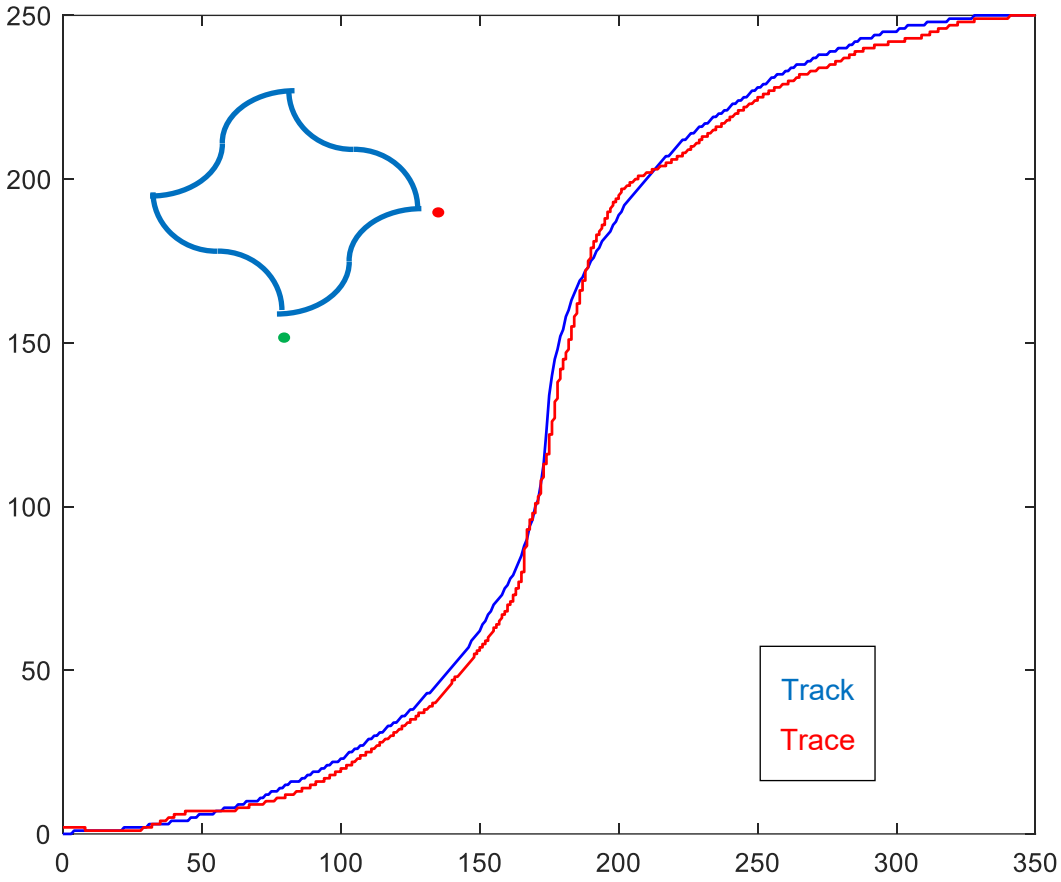

Figure 5.2 **Participant's tracing trajectory over a track section. The position-error is the total area between the trace (red) and track (blue) section. The inset at the top-left corner shows the complete track pattern. The green and red dots identify the active track to be traced. Axes are in units of screen pixels.**

Protocol

The participant completed a longitudinal PT program that included 7 training sessions, limited to a single session per day. Each session consisted of 8 tracing trials with the right hand followed by 90 tracing trials with the left hand. A trial was defined as a tracing task over a track section, where the specific track and direction of tracing was randomly selected by the program. The rationale behind using the left (non dominant) hand for tracing was to induce motor learning effect [79] thereby increasing the potential

for larger changes in FC during a short PT program. Data for the right hand were collected but not used in this study. We collected 5 minutes of pre- and post-PT resting state EEG data in each session. Each session lasted approximately 50 minutes.

The protocol for assessment of motor skill was like that of PT but limited to only 30 tracing trials for the left hand. We carried out three motor assessments over three consecutive days after the end of PT program to obtain a baseline for skill level before the start of the neurofeedback training (NFT) program. Three assessments were the minimum number of samples needed to determine a trend and measurements' deviation around that trend. Our goal was to minimize the number of sessions before NFT to reduce the potential for fatigue in our participant. We also carried out four motor assessments over four days after the completion of the NFT program to test for retention of the acquired skill. We had planned to obtain more than the minimum samples after the NFT program, but the participant could only partake in 4.

NFT started after the completion of the PT program. EEG data were acquired at 1-second intervals and processed in real-time to calculate the instantaneous FC measures and generate the subsequent audio feedback. Volume of the audio feedback was proportional to the magnitude of the FC measures. Each NFT session lasted 5 minutes. Depending on the participants level of fatigue, we carried out at least two sessions, and at most, five sessions in each day of the NFT program. There were a few minutes break between each session to document the participant's recollection of MI in the preceding session and the strategy, if any, for the next session. We also collected 5 minutes of resting-state EEG data before the start, and after the end of all NFT sessions in each day.

5.3.2. Measure of Motor Skill

We evaluated two measures of performance indicators, namely the accumulated position error during each tracing task, and the product of the accumulated position error and the time taken to complete the corresponding track (position-time error). The latter brought the tracing speed into consideration. Position error was calculated from the total area between the track and tracing trajectory (Figure 5.2) and converted to spatial units of

squared centimeter (cm^2) based on an approximate conversion factor of 0.25 mm^2 per square pixel. Position-time error was in units of cm^2seconds.

We generated three separate measures of motor skill for each session. 1) A single value corresponding to the median error of all 90 tracing trials, referred to as single-median option. 2) Median error of the first 30 trials to represent the tracing performance before PT, and 3) that of the last 30 trials for performance after PT, referred to as dual-median option. The process was carried out for both position and position-time errors. These were used to evaluate four distinct FC corelates of motor skill as explained in the signal processing section.

For assessments of motor skill associated with NFT, a single set of 30 tracing trials was carried out before the NFT sessions and another 30 tracing trials after the completion of all NFT sessions on that day. Motor skill before and after NFT was quantified by the median of the respective tracing errors. We also used 30 tracing trials for each skill assessment after the completion of PT and NFT programs.

5.3.3. Signal Processing and Analysis

EEG preprocessing

Recorded resting state EEG data were imported into MATLAB-7.8.0 (MathWorks Inc) for signal processing and analysis during the PT program. We applied a frontend bandpass filter of 1-45 Hz using a finite impulse response digital filter. The filtered data was then visually inspected in EEG-Lab V14.1.2 to select a 2-minute continuous section with minimal amount of interference from muscular activities. The frequency range was further divided into five canonical bands of Delta (1-4 Hz), Theta (4–8 Hz), Alpha (8–15 Hz), Beta (15–30 Hz) and Gamma (30–45 Hz), each with three additional sub-bands of low, medium, and high frequencies. This was the design criteria for the 15 distinct Morlet-wavelet filters with center frequencies that were approximately one bandwidth apart [62]. We used coherence between EEG electrode pairs (channels) at each of the 15 frequencies as a measure of functional connectivity at that frequency and evaluated the instantaneous coherence through Phase Lag Index (PLI) algorithms [62].

Motor skill correlates

With the 32 electrode EEG, coherence had to be evaluated for each of the 496 non-directional channels. The individual coherence samples from each channel were averaged over a one-second non-overlapping epoch, resulting in 120 coherence measures in each of the 15 frequency bands. We used the maximum coherence in each band as an index of FC at the corresponding center frequency, resulting in an array of 7,440 indices for PLS analysis [55]. We used single-median tracing performance for correlation analysis between FC indices from both the pre- and post-PT EEG data separately. For the dual-median option, we used the pre-PT tracing performance (median of the first 30 trials) for correlation analysis with FC indices from pre-PT EEG data, and post-PT tracing performance (median of the last 30 trials) for correlation analysis with FC indices from post-PT EEG data.

We screened for promising channels and frequencies by selecting those with strong correlation ($p < 0.05$) with each of the single- and dual-median tracing performances using PLS correlation (PLSC) analysis [73]. We carried out 50 iterations of the correlation analysis to identify the channels that were present in at least 80% of the repetitions, and at the same level of robustness. We used bootstrapping to quantify robustness [73]. We then used PLS-Regression (PLSR) to generate a model for estimating the tracing performance from the FC indices of the identified channels. The error in estimation was quantified by root-mean-square-error (RMSE) using leave-one-out cross-validation over the seven PT sessions. The RMSE was presented as a percentage of the average tracing error obtained from each of the single- and dual-median options. This allowed for comparison between models for position or position-time tracing performances.

Neurofeedback training program

EEG data during NFT program were acquired at one-second intervals and processed in real-time using C++ programming under Microsoft Visual Studio 2019 environment. Coherence was only evaluated for the channels and frequency bands that were identified through PLSR analysis during the PT program. The instantaneous

coherence measures were averaged over the one-second epoch for each channel and subtracted from a baseline threshold separately. The threshold was selected from the minimum FC index at each of the channels in the regression model. The resulting data were then used to adjust the volume of an audio feedback. The participant was advised to finetune their metal imagery with the goal of increasing the volume of the audio feedback.

5.3.4. Participant

A healthy right-handed 61-year-old female volunteered in order to test the methodological feasibility for this study. The participant had no known neurological or physically limiting conditions, and no implants. The Research Ethics Board of Simon Fraser University approved the protocol for this study, and the participant signed an informed written consent form.

5.4. Results

5.4.1. Physical Training

The participant completed seven tracing sessions as part of the PT program. Tracing was done with the left hand and repeated for 90 trials over randomly selected track sections and directions of movement. Figure 5.3 shows the resulting position error (blue bars) and position-time error (red bars) for the first 30 tracing trials during the seven sessions of the PT program. The results for the last 30 and all 90 tracing trials were similar in nature and were omitted for the sake of clarity and space. Note that the position-time error has two degrees of freedom, where a change in its value without a corresponding change in position error is also an indication of change in skill due to the tracing speed. The red asterisk indicates significant ($p < 0.05$) change in position-time performance over the PT program. Statistical analysis of the change in motor skill is explained later in this section.

5.4.2. PLSC Analysis

FC indices from 496 EEG channels at 15 center frequencies for each of the pre- and post-PT data were separately used in correlation analysis with each of the position (Pos-only) and position-time (Pos-time) assessments of motor skill. Each skill assessment was quantified through a single- or dual-median approach. The latter was represented by the median of the first and last 30 tracing trials and were used in conjunction with the pre- and post-PT EEG data, respectively. This resulted in four separate PLSC analysis for each of the pre- and post-PT EEG data. Table 5.1 shows the results of the PLSC analysis that generated promising correlations ($p < 0.05$ uncorrected). We used bootstrapping to identify the most robust channels and center frequencies that contributed towards the correlation with tracing performance [55]. PLSR analysis was constrained to these channels and frequencies as described in the next section.

5.4.3. PLSR Analysis

We generated a separate regression model for each of the entries in Table 5.1. Predictors of the models were limited to the most robust channels and frequencies that were identified from PLSC analysis. Regression coefficients in each model represented the contribution from the identified channels (predictors) at the specific frequency sub-band towards valuation of the tracing performance (behavior) from FC indices (objective measures). We used leave-one-out cross-validation to quantify the RMSE in estimating tracing performance. Estimation error was represented as the ratio of the RMSE and the average of all 7 tracing performances. The small number of behavior samples had an unfavorable impact on the statistical power of our regression analysis. To counter the effects of the small sample size, we reduced the number of predictors to achieve a statistical power of 0.8 at $\alpha=0.05$ and R^2 of the respective regression model. This was done by removing the predictor with the lowest regression coefficient and generating a new model through iterative PLSR analysis. Our rationale for this approach was the argument that these channels would have less impact on the overall estimation error compared to those with higher regression coefficients. PLSR columns in Table 1 show the results of this iterative process with the final number of channel counts and corresponding RMSE ratio. The statistical power for Option-D was well above 0.8, but we could only

achieve a power of 0.78 for Option-B as we reduced the number of channels to 4 and less. The R^2 for options A and C were too low and became worse for lower number of channels, resulting in a best-case power of 0.57 and lower. We therefore excluded options A and C and only focused on Option-B and D for neurofeedback training.

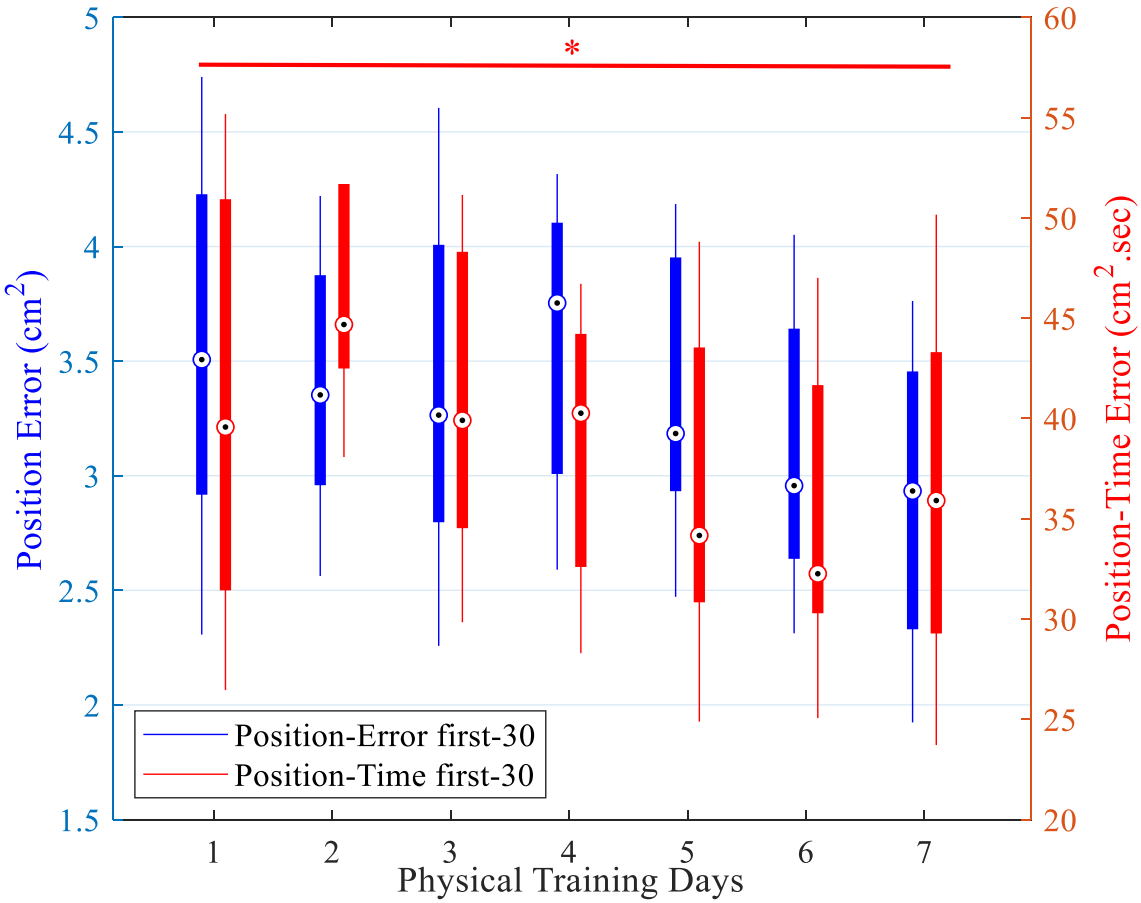

Figure 5.3 **Longitudinal tracing performance during physical training in terms of position error (blue bars), and product of position error and time (red bars). The bullseye indicates the median value. Corresponding results from all 90 and last 30 trials were similar in nature and were excluded for the sake of clarity. (*) indicates significant ($p < 0.05$) change in tracing performance.**

Table 5.1 **Results of PLS analysis for FC indices from EEG data collected before (pre-PT) and after (post-PT) physical training. Input data consisted of 496 EEG channels at 15 center frequencies, using Single- or Dual-median assessments for position (Pos-only) or position-time (Pos-time) tracing errors. Only four combinations (Option-A to D) met our screening criterion for strong correlation ($p < 0.05$ uncorrected), shown in descending order of coefficients (Corr. Coeff.). Channels represent the final number of predictors used in the regression model from PLSR analysis. R^2 is the coefficient of determination. Model performance is determined by the ratio of RMSE over the average tracing errors from all 7 PT sessions.**

Input				PLSC Results		PLSR Results		
Option	EEG	Freq. Band (center)	Tracing Performance	*p-value*	Corr. Coeff.	Channels	R^2	RMSE (%)
A	Pre-PT	Beta-Med (21 Hz)	Single-median Pos-only	0.008	0.999	4	0.879	3.04
B	Pre-PT	Beta-Low (16 Hz)	Dual-median Pos-time	0.019	0.999	4	0.945	2.51
C	Post-PT	Beta-Low (16 Hz)	Single-median Pos-only	0.001	0.998	4	0.908	2.81
D	Post-PT	Beta-Low (16 Hz)	Dual-median Pos-only	0.020	0.998	4	0.974	2.13

5.4.4. Neurofeedback Training

The regression models for estimating tracing error had both positive and negative coefficients. This meant that an increasing FC in channels with negative coefficients resulted in a reduction in tracing error, or an improvement in motor skill. Increasing FC in channels with positive coefficients had the opposite effect. We only provided feedback on FC of channels that were conducive towards a reduction in tracing error. The volume of the audio feedback was proportional to the increase in FC above a baseline threshold. We selected the baseline threshold to be the smallest FC index for the corresponding channel during the PT program. The indices of the channels with negative regression coefficients for Options B and D are shown in Figure 5.4(a) and (b), respectively. Options A and C were not used for neurofeedback, due to their low statistical power of the regression model. Bearing in mind that the maximum value of an FC index is limited to 1, we looked for the regression model that provided the largest potential for increasing FC through neurofeedback. Option-B was the only model that showed an opportunity for the

participant to increase the FC indices of both contributing channels (B1 and B2) by over 0.12 (~ 14% of the last reading). Option-D was limited to 0.05 and only available for one contributing channel (D2).

We carried out three assessments of tracing performance between the end of PT and start of the NFT program. Motor skill assessments were also done for each of the eleven NFT days. After the completion of the NFT program, the participant went through an additional four tracing assessments to examine the short-term retention of acquired motor skill. Figure 5.5 shows the results of all the motor skill assessments in terms of the tracing error specific to Option-B in Table 5.1. Impact of guided MI on the resting state FC of selected channels during the NFT program is shown in Figure 5.6. FC indices are from resting state EEG data collected before the NFT sessions. The corresponding FC indices during the PT program are also included in Figure 5.6 (dashed lines) to provide a reference for relative changes in individual FC indices.

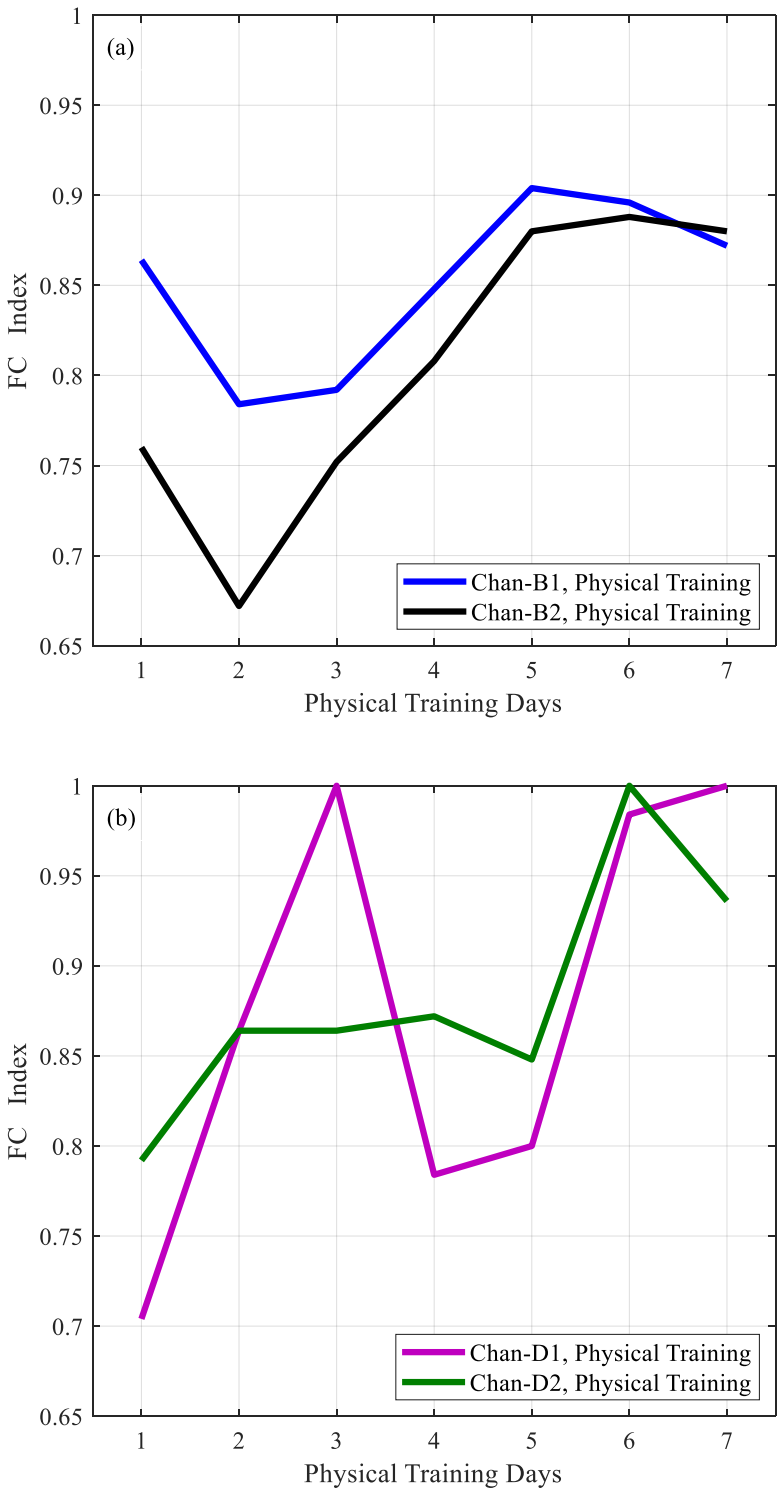

Figure 5.4 **FC indices for (a) EEG data from Option-B in Table 1, and (b) those of Option-D. Only the two channels with negative regression coefficients are shown. These channels were selected for neurofeedback training. Option-B provided a larger potential margin for increasing FC indices as compared with option-D.**

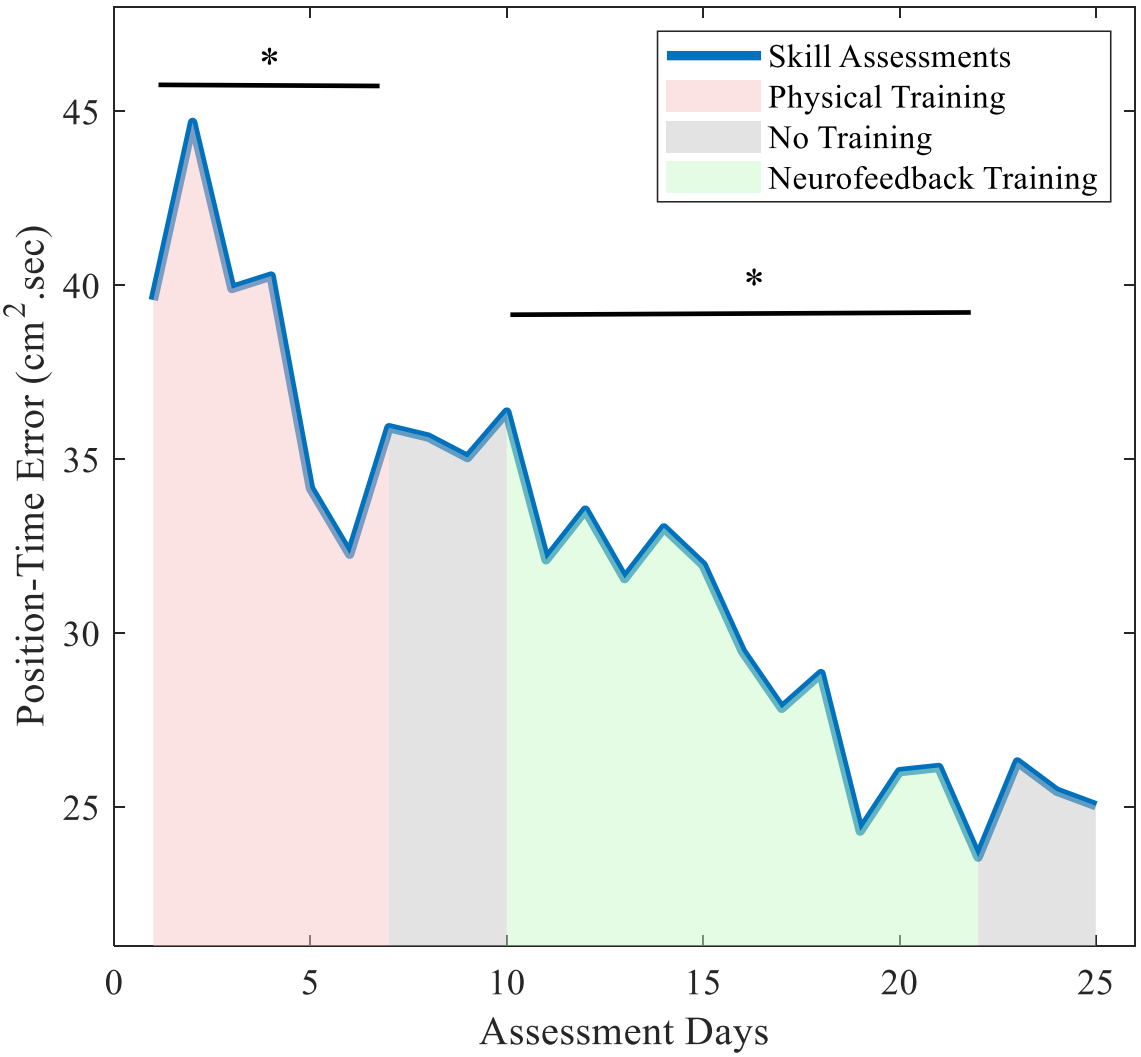

Figure 5.5 **Motor assessments for Option-B with Dual-median Position-time selection. (*) indicates significant ($p < 0.05$) change in tracing performance.**

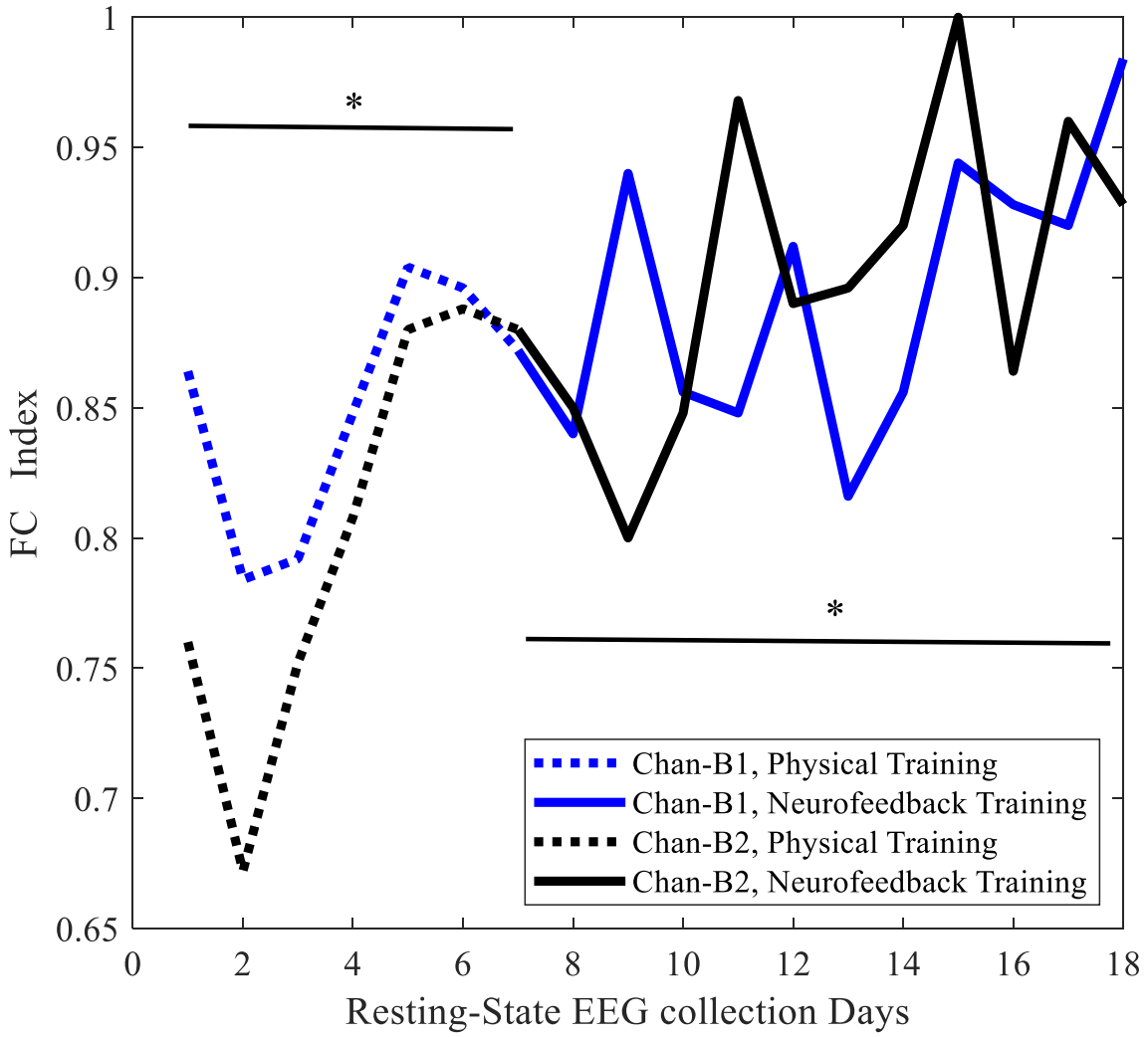

Figure 5.6 **Resting state FC indices in channels from Option-B. Processing was done on pre-training EEG data for both PT and NFT program. B1 and B2 are the channels with negative regression coefficients. Data from both PT (dashed lines) and NFT (solid lines) are included for comparison. (*) indicates significant (p < 0.05) change in FC indices.**

5.4.5. Statistical Analysis

We began our analysis by testing for the presence of a significant ($p < 0.05$) change in motor skill during each of the PT and NFT programs. We also investigated the manifestation of discernable structure in the way motor skill changed during the NFT program. Statistical analysis was done using version 4.2.1 of R (R Core Team, 2022).

We fit a simple linear regression model to the motor skill assessments during the PT program. The slope of this regression model was significantly different from 0 (p = 0.048) and indicated a decrease in mean tracing errors across PT sessions. Next, we fit a linear regression model to the motor assessments scores during the NFT program, which also showed a significant ($p = 8.5 \times 10^{-5}$) decrease in mean tracing errors. We also considered higher-order polynomial regression models using standard information criteria [84] and found that a linear fit was most appropriate.

We probed the motor assessment scores during the NFT program for a possible substantive change in the relationship between motor skill and NFT sessions. This was done to explore the presence of change-points in the regression models, indicating different phases in the way NFT influenced motor skill. Processing was done by performing a structural change analysis using the 'strucchange' package in R [85]. The three F-statistic based tests gave inconsistent conclusions (p-values: sup$F$ = 0.064, ave$F$ = 0.073, exp$F$ = 0.04). Repeating this analysis using quadratic regression models gave similar results. We therefore concluded that there was insufficient evidence to suggest a structural change during the NFT program.

Using a two-sample T-test, we found a significant ($p = 7.19 \times 10^{-7}$) difference between the mean tracing performance in motor skill assessments before the start of the NFT program to that of post NFT assessments (Figure 5.5). The 95% confidence interval reported that mean performance was lower by between 9.3 and 12.3 cm^2sec after the completion of the NFT program. To analyze the impact of the NFT on resting state FC, we used an exact permutation test to check if the mean FC index was increased in the channels that were the focus of MI (Figure 5.6). Computation was done using the 'EnvStats' package in R [86]. A preliminary investigation did not find evidence of strong heteroscedasticity, and permutation tests found a significant increase in peak resting state FC of both channels (p = B1: 0.047, B2: 0.0016 with 31,824 permutations).

5.5. Discussion

Prior research has characterized the relationship between MI, NF, and motor skill. Motor learning is associated with repetitive activation of relevant neural networks [87], with

similar structures being involved for real and imagined actions [88]. This is the rationale for using MI as an adjunct for acquisition of skill [89] and NF to guide the MI for activation of the relevant networks [90]. Our last objective in this study was to investigate the causal effects of changing FC on motor skill based on individualized selection of brain networks and synchronization frequencies. We built on the results from our objective-1 in this study that showed encouraging performance in estimating motor function from connectivity measures based on global consideration of brain areas. To this end, we chose a longitudinal motor skill training program for an initial identification of EEG channels that exhibited strong correlation between their respective FC and the change in motor skill. The identified channels were then specifically targeted to induce a change in their connectivity indices. We selected MI as a non-invasive and pragmatic technique to influence FC and chose neurofeedback to guide the MI that had the highest impact on the FC indices of the identified channels. Our results showed over 20% improvement ($p = 7.19 \times 10^{-7}$) in motor skill at the end of the NFT program as compared with the skill assessment after PT. The results also indicated a retention of the motor skill that was measured through assessments spread over five days after the completion of the NFT program. This, however, needs to be better evaluated in future studies.

A cursory look at the tracing performance during the NFT program (Figure 5.5) seems to indicate an initial learning period of about five days that did not result in a major improvement in motor skill. Additional NFT sessions showed a gradual decrease in tracing error over the subsequent NFT days. But statistical analysis of this data pointed towards insufficient evidence for a structural change in the regression model during the NFT program. However, all three F-statistic tests exhibited *p*-values near our significance threshold ($p = 0.05$), suggesting that future research with a larger cohort may identify a more complicated relationship than a single linear model. Similar learning periods were observed in other neurofeedback studies, but the authors did not consider the observations conclusive [89]. Such a structural change in how motor performance improves would provide consequential information for future adoption of NFT as a complementary therapeutic activity. Knowing that there may be a learning period during which a major change in skill is not to be expected can help limit the discouraging psychological effects from lack of progress.

Our proposed approach is contingent on an initial physical training program that triggers measurable change in motor function to enable the identification of contributing

channels and synchronization frequencies. This may appear as an impediment towards the application of our methodology for individuals with limited ability to partake in the physical training. In such cases, we believe that the alternative approach of neurofeedback on generalized areas of interest, such as global connectivity with motor areas, may be an appropriate strategy for inducing an initial level of improvement [54]. The latter would still require the collection of EEG data and assessment of motor function, which could then be used for subsequent PLS analysis and identification of relevant channels. A shift in strategy from a generalized to an individualized approach can then follow these initial intervention steps with the potential to expedite or possibly increase further improvements in function [56]. Concerning the clinical application of our proposed method, resting state EEG data can be collected without major hindrance during the initial therapeutic activities as the individuals improve their motor function. Individualized contributing channels and frequencies can then be identified and used for NF training as a complementary activity to the ongoing physical therapy.

In our design of the experiment, we used the error in tracing as a measure of motor skill. This represents an inverse relationship between the regression model and behavior, meaning that a decrease in tracing error signifies an improvement in skill. Consequently, increasing the FC in channels with negative regression coefficients resulted in a better tracing performance. This was equivalent in effect to decreasing FC in channels with positive regression coefficient. We could not, however, propose a mechanism to guide specific MI that would reduce the peak FC in a particular channel. We therefore focused on providing neurofeedback on channels with negative regression coefficients to help the participant identify the MI that resulted in an increase in FC. The model represented by Option-D (Table 5.1) had the lowest RMSE in estimating the tracing performance and was therefore the preferred model for NFT. But an examination of the maximum resting state FC in the channels with negative regression coefficients (Figure 5.4(b)) showed a small margin for increasing FC. The latter represents a ceiling effect for our models and the reasoning behind the selection of our second candidate model, despite its lower performance in estimating skill (higher RMSE ratio). Option-B provided a better opportunity for inducing a larger range of change in FC (Figure 5.4(a)). This ceiling effect might be a limiting factor for individuals without multiple options for predictive models. Future work in this area should investigate the proportion of participants that exhibit this limitation.

The positive regression coefficients in our model may imply a gradual development of specialized networks for efficient execution of motor tasks, thereby requiring a decreasing recruitment of those networks to perform the respective tasks [32], [13]. This represents an inherent dependence of our approach on the passage of time, where alternative networks start to contribute more towards further improvements in skill. A continuous development of alternative regression models as new data samples become available may be conducive towards indirect circumvention of the ceiling effect, by introducing new contributing channels with potentially larger margins for increasing FC. This dynamic evolution of predictive models may help extend further improvements in skill. It is important to clarify that we are not arguing for targeted brain stimulation through neurofeedback as a replacement for physical therapy or training. We instead propose our approach as a complementary activity that could potentially enhance and finetune the efficacy of established intervention strategies. This study shows promise for individualized targeted endogenous brain stimulation that is optimized through the use of neurofeedback. We are hopeful that our results would encourage a more comprehensive study of this methodology with a sufficiently larger number of participants and under controlled experimental conditions. The objective would be to investigate the potential of individualized stimulation in expediting or possibly increasing the effect of MI, as compared with the conventional approach of mental practice that focuses on the SMR activities or generalized connectivity measures with motor areas only.

To summarize, we investigated the effect of individualized approach towards improving motor skill through influencing FC in EEG channels that were not constrained to the motor areas. We selected coherence at specific synchronization frequencies between electrode pairs (channels) as a measure of FC and used peak resting state FC as an objective measure for estimating change in motor skill through physical training. We applied PLSC analysis to detect the contributing channels and PLSR to generate a model for estimating change in motor skill. We then used MI as an endogenous brain stimulation mechanism to influence the FC in the contributing channels of the regression model. We provided real time feedback on the instantaneous FC in the identified channels to guide the MI in a way that was conducive towards improvement in skill. We showed over 20% improvement in motor skill through neurofeedback training alone, without any additional physical training. We also showed retention of improvement in skill for several days after the completion of NF training.

Chapter 6.

Concluding Remarks

Material in this chapter is extracted, reproduced, and modified with permission from the following papers:

N. Riahi, V. A. Vakorin, C. Menon, “Estimating Fugl-Meyer Upper Extremity Motor Score from Functional-Connectivity Measures,” *IEEE transactions on neural systems and rehabilitation engineering*, vol. 28, no. 4, pp. 860-868, Apr. 2020.

N. Riahi, R. D’Arcy, C. Menon, “A Method for Estimating Longitudinal Change in Motor Skill from Individualized Functional-Connectivity Measures,” *Sensors* 2022, DOI: 10.3390/s22249857.

N. Riahi, W. Ruth, R. D’Arcy, C. Menon, “A Method for Using Neurofeedback to Guide Mental Imagery for Improving Motor Skill,” *IEEE transactions on neural systems and rehabilitation engineering* 2022, DOI: 10.1109/TNSRE.2022.3218514.

6.1. Chapter Overview

Chapter 6 summarizes the research goals and the proposed method in achieving the resulting objectives. It presents the potential approach for implementation of the proposed methodology as well as future work to address the limitations and areas for improvement.

6.2. Summary

A large percentage of stroke survivors endure functional inadequacies throughout the chronic phase with psychological and financial impact to both the stroke survivors and their caregivers [2]. There is evidence that continual adjustment of rehabilitation strategy and therapeutic activities is conducive towards further recovery of function during chronic phase [4]. Accurate and frequent assessment of motor function would contribute towards

modification of rehabilitation strategies and personalization of therapeutic activities [5]. Frequency of assessments, however, is hampered by the availability and expertise of trained examiners. This motivated our investigation for a cost-effective assessment tool and a complementary therapeutic activity that could help with further recovery of function during chronic phase.

For pragmatic reasons, we focused on investigating the potential use of EEG-based resting state FC for assessment, monitoring, and influencing motor function. As such, our study aimed to address three research questions:

1. Is EEG-based rsFC a suitable neurophysiological measure to accurately estimate motor impairment in stroke survivors,
2. can rsFC be used to estimate small incremental changes in motor function, and
3. can we induce a change in motor function by influencing individualized FC channels through MI?

and the subsequent objectives:

1. Propose a method for estimating FMU from rsFC measures.
2. Test the performance of the proposed method in estimating small incremental changes in motor function.
3. Investigate the prospects of influencing the individualized FC measures through MI for improving motor function.

The first phase of our research attempted to answer the question on suitability of rsFC for accurate estimation of motor impairment in stroke survivors. In chapter 3 we described a method for estimating FMU from rsFC measures. We selected spectral coherence as a measure of FC and applied PLS algorithms to identify contributing connectivity channels and generated models for estimating FMU. Cross-validation resulted in an R^2 of 0.97 and a root-mean-square error of 1.9 on FMU scale. We argued that the ease of use and the accuracy of estimation reduced the dependency on the availability of trained examiners and was conducive towards frequent assessments of motor impairment and personalization of therapeutic activities.

A limitation of the resulting method from the first phase of our research was that its accuracy was evaluated against FMU, which measures relatively large changes in motor

function. Gaining large improvements in function could take a long time to achieve, which weakens the incentive for more frequent assessments. The second phase of our research aimed to address the question on whether rsFC could be used to estimate small incremental changes in motor function. In chapter 4 we described an objective measure of motor skill that could be quantified at small longitudinal changes and tested the performance of the proposed method in estimating incremental changes in motor skill of healthy participants. Our rationale for using healthy participants was that incremental changes in motor skill of healthy individuals were assumed to be much smaller than functional improvements associated with a point change in FMU. We presented a computer-based tracing task for motor skill training and used the spatial error in tracing as an objective measure of motor function. The results yielded an average accuracy of 98% (standard deviation of 1.2%) in estimating tracing error. Our focus was on individualized approach as a precursor towards application of neurofeedback for skill improvement and the last phase of our research.

The ability to estimate small longitudinal changes in motor skill at short intervals paved the way for an investigation into bidirectional interactions between skill and FC. Our focus in the previous two phases was on quantifying the change in FC as a consequence of change in motor function. Our goal for the third phase of this study was to examine the reverse relationship and assess the impact of changes in FC on motor function. This was inline with our research aspiration for using this method to facilitate a complementary therapeutic activity and our objective to investigate the prospects of influencing the individualized FC measures through MI for improving motor function. In chapter 5 we described a process for the selection of individualized FC channels and real-time monitoring of instantaneous FC measures in those channels. Digital tracing tasks were used for skill training and assessment. PLS algorithms identified the most robust contributing channels towards correlation between the resting state FC and the acquired motor skill. A method for utilization of NF to guide MI as an endogenous brain stimulation was presented. We showed over 20% reduction in tracing error through NF training alone, without any additional physical training.

Our proposed method shows promise in facilitating an individualized approach towards improvement of motor function that could complement the conventional therapeutic activities. It also has the potential for providing an accurate assessment of

motor impairment while addressing the challenges associated with the availability of trained examiners.

6.3. Discussion

We focused on FMU as a widely accepted assessment tool and proposed EEG-based rsFC as an objective neurophysiological measure for estimating FMU. We used PLS algorithms to limit the number of EEG channels for evaluation of rsFC and the subsequent estimation of motor impairment. Our goal was to propose a methodology that not only resulted in an objective and accurate estimation of motor function, but also addressed the pragmatic requirements associated with cost, portability, and ease of use while requiring minimal operator expertise.

We argued that our proposed approach using EEG-based rsFC could meet the pragmatic requirements, when considering the following:

- EEG systems
    - require minimal amount of training to setup and operate,
    - are relatively inexpensive and can be purchased with lower number of electrodes, thereby making them more cost effective, and
    - are portable, which addresses some of the mobility challenges associated with stroke survivors.
- Reducing the number of EEG electrodes for data collection can decrease the setup time, which combined with resting state analysis, can reduce the overall assessment time to less than 20 minutes, and potentially 15 minutes when using Dry-EEG caps.
- Use of resting state data for estimating motor function can further improve the applicability of the proposed method when considering the stroke survivors' potential challenges in
    - performing physical tasks that may be required for motor assessment, and
    - comprehension of instructions.

In summary, and specific to clinical relevance and applicability, our proposed method may motivate more frequent assessments of motor function, thereby facilitating a

continual adjustment of rehabilitation strategy and therapeutic activities at shorter intervals. Considering these factors and a capital cost that could potentially be recovered after a limited number of assessments in a clinical setting, or at the point of care, the advantages of the proposed method may provide significant incentive for adoption and use.

Another factor that influences the frequency of motor assessments is the resolution of assessment scores in quantifying the change in motor function. Given proper training of the examiners and following standardized approach, the minimum detectable change for FMU can be reduced to approximately 3 points, which is just under 5% of maximum scale for FMU assessment [9]. This is exacerbated by the subjectivity of the evaluation, which can cause variability in the final assessment scores [91]. Furthermore, every point change in FMU score corresponds to large changes in function and as such may not contribute towards assessment of small incremental improvements in execution of motor tasks [92]. This would further discourage frequent assessments since achieving a minimum detectable change in FMU score might take a long time to accomplish. The accuracy of our proposed method was evaluated by its ability to estimate FMU. It showed good performance with an RMSE of about 1.9 on FMU scale. The measure of accuracy for the method, however, is still subjective in nature as it is estimated with respect to a subjective assessment. We therefore proposed appraising the accuracy against a more objective measure of motor function that could be evaluated at modest incremental changes. We opted to use computer-based tracing tasks to evaluate spatial error in tracing. This was an objective measure of individualized longitudinal change in motor skill that could be quantified at small increments for healthy participants. The proposed method resulted in an average estimation accuracy of 98.2% (SD 1.2%). Incremental changes in motor skill of healthy individuals were assumed to be much smaller than functional improvements associated with a point change in FMU. We consequently argued that at this level of accuracy in estimating objective measures of behaviour, the proposed method might have the potential to provide intermediate valuations of change in motor function that are less than the minimum detectable change by FMU assessments. It may therefore justify more frequent assessment with the expectation of detecting changes in assessment scores before any discernible change in motor function. This might therefore present a potential for finetuning therapeutic activities at shorter intervals. Specific to clinical relevance and applicability, the time/resources needed for the initial collection of data

towards developing the individualized models may seem impractical. But upon closer look at the sequence of events during a therapy session, one could realize minimal impact to the flow of activities. The generalized model could initially be used to produce a baseline assessment of impairment for stratification, development of rehabilitation strategy, and selection of therapeutic activities. The only additional activity during the subsequent therapy sessions is the collection of 2-minutes resting state EEG data, which could be completed in 5 to 10 minutes when using dry caps. The physiotherapist provides an assessment of change in function that would be used along with the EEG data to produce the individualized model after 7 therapy sessions. The model could subsequently be used for more frequent objective assessments that would help the therapist to adjust and finetune the therapeutic activities, or potentially augment them with the application of NFT as a complementary approach along with the prescribed physical therapy.

Inherent in our analysis of motor skill improvement is the dependence on passage of time. The stable changes in functional connectivity have been attributed to the development of specialized neural circuits for fast and efficient execution of tasks. These changes are not necessarily all associated with an increase in recruitment but also the opposite, indicating that some networks may have become either more efficient or less important for the respective motor skill acquisition [13, 32]. We can make similar interpretations about the sign of the regression coefficients from PLSR models to indicate an increase or decrease in network recruitment. The longitudinal impact of the contributing channels is expected to change as motor skill improves over time. Having a simple and cost-effective tool could facilitate an opportunity to monitor and track the time dependent changes in contributing channels. We expect these changes to be gradual and suggest that new individualized models could be developed based on a rolling reassessment of the FC as new samples become available. This longitudinal view of individualized assessments and continual adjustments of regression models might be conducive towards extending the improvement in motor skill through application of mental imagery (MI) and real-time neurofeedback.

Specific to the use of MI as a complementary therapeutic activity, our objective was to investigate the causal effects of changing FC on motor skill based on individualized selection of brain networks and synchronization frequencies. We used a longitudinal physical training (PT) program for the initial identification of EEG channels that exhibited strong correlation between their respective FC and the change in motor skill. The identified

channels were then specifically targeted to induce a change in their connectivity indices. We selected MI as a non-invasive and pragmatic technique to influence FC and chose neurofeedback to guide the MI that had the highest impact on the FC indices of the identified channels. The approach, however, is contingent on an initial physical training program that triggers measurable change in motor function to enable the identification of contributing channels. This may appear as an impediment towards the application of the methodology for individuals with limited ability to partake in the physical training. In such cases, it is clearly better to stay with the traditional approach of applying well established conventional rehabilitation strategies. However, we recommend collection of resting state EEG data during these initial therapeutic activities without major hindrance on the prescribed rehabilitation strategy. The proposed method can then be used to identify Individualized contributing channels that correlate with the improvement in function as a consequence of the initial intervention activities. These FC channels can then be employed for NF training to provide a complementary therapeutic activity along with the ongoing physical therapy. Our study was not based on any *a priori* selection of specific brain areas, which allowed us to use the algorithms to identify the contributing channels at the sensor level. For applications where hypothesis is based on influencing specific brain areas for the purpose of improving function, then source level functional connectivity analysis would be preferred and deemed necessary to facilitate stimulation of the correct brain networks.

Our approach generated multiple potential models, some of which exhibited a small margin for increasing FC. This represents a ceiling effect for the respective models. A continuous development of alternative regression models as new data samples become available may be conducive towards indirect circumvention of this ceiling effect, by introducing new contributing channels with potentially larger margins for increasing FC. This dynamic evolution of predictive models may help extend further improvements in skill. It is important to clarify, however, that we are not arguing for targeted brain stimulation through neurofeedback as a replacement for physical therapy or training. We instead propose our approach as a complementary activity that could potentially enhance and finetune the efficacy of established intervention strategies.

6.4. Limitations

A major limitation of our study is the small number of observations for PLS analysis. Although data from 10 stroke survivors showed promising results in estimating FMU, the reliability of the regression model is a concern for PLS analysis with less than 30 observations [72]. This is further exacerbated by the distribution of FMU scores that covered a broad range of impairments from a minimum score of 18 to the maximum of 49 on FMU scale. Prior publications have shown second degree polynomial (with an inverted U-shape) relationship between the baseline FMU score and treatment gain [9]. This may be indicative of the need to develop separate models for baseline FMU scores of less than 25 and greater than 50 [9]. The number of participants in our study did not allow for separation into different groups. Our study is therefore inconclusive with respect to the estimation performance of these group-specific models. An obvious dilemma with the use of different models is that the selection of a specific model depends on the score it is trying to estimate. We speculate that the generic model described in this study might be an adequate starting point to provide a high-level estimate of motor function that could guide the selection criteria for the more appropriate baseline dependant model.

The limitation associated with the number of observations can be extended to the second phase of our research with respect to appraisal of estimation accuracy against objective measures of motor function. Our study with stroke survivors had identified a minimum of 3 connectivity channels for estimating FMU. Given the expected R^2 values and the required statistical power of 0.8 with 3 predictors (channels), we proposed 7 observations as the minimum number of sessions required for estimating motor skill. To increase the reliability of our regression models, we needed more sessions per participant. This, however, proved very challenging under the environmental conditions at the time. The impact of lower number of observations per participant may have manifested itself through higher *p-values* from permutation test. Although all *p-values* were less than 0.05, only three were at 0.01 and remaining four were closer to 0.04. This may not have been a problem for FMU analysis, but it is a concern for motor skill study because of multiple comparison. Our study was exploratory in nature and as such, we considered calculating the tracing error from either 30 trials (dual median option) or 90 trials (single median option). This is a problem for the participants with *p-values* of 0.04, since the corrected significance threshold is now at 0.025, which makes the appraisal of estimation accuracy

from those participants less reliable. The reasoning can be further extended to include the measure of skill (position-only or position-time), which would further reduce the significance threshold to 0.0125. But this is debatable since the two measures of skill might involve different networks in the brain, which effectively separates the independent variables. Although these higher thresholds of significance could influence our conclusions related to the statistical analysis of the contributing channels, it did not have a material impact on our conclusions with respect to the outcome of selecting those channels for the purpose of NFT. If the identified channels were random, then influencing their FC measures would most likely not have induced a significant change in motor skill. Nevertheless, excluding the results from four of the participants reduces our trust in the average estimation accuracy across the participants. We believe higher number of observations per participant, as well as larger number of participants for this analysis would have been informative towards reliability of our conclusions. Interestingly enough, focusing only on the results from the three participants with *p-values* of 0.01, increases the average estimation accuracy to 98.5% (from 98.2%) and reduces the standard deviation to 0.4% (from 1.2%). This is speculatively encouraging as it could point towards an expectation to have higher estimation accuracy with more reliable (lower *p-values*) individual measurements. However, deciding the required number of observations remains a challenge due to dynamic dependence of regression models on the passage of time.

In our design of experiment for an objective assessment of motor function, we used the error in tracing as a measure of motor skill. This represents an inverse relationship between the regression model and behavior, meaning that a decrease in tracing error signifies an improvement in skill. Consequently, increasing FC in channels with negative regression coefficients would result in a better tracing performance. Conversely, decreasing FC in channels with positive regression coefficients would also result in improvement in performance. We could not, however, propose a mechanism to guide specific MI that would reduce the peak FC in a particular channel. Being able to devise a process by which the peak FC could be manipulated in a descending trajectory would considerably increase the dynamic range of NF guided MI. This seems intuitively challenging, since not recruiting a network during a prescribed NF training session might not be equivalent to intentionally using that network at a lower capacity or more efficiently (lower FC index). Being restricted to only using the channels with negative regression

coefficients can limit the potential application of our proposed method, since it is not guaranteed that such channels with enough margin for increasing peak FC can be identified. Specific to the limitation of this study with respect to the efficacy of targeted MI through NF for improving motor function, we did not have the opportunity to design a control condition into our experiment. This would have strengthened our claim against a potential placebo effect. We did, however, take steps to mitigate this through multiple assessments of skill without any accompanying physical or neurofeedback training. Our analysis of change in skill during a physical training program (phase-2 of our study) did not show significant change in skill during a 30-trial portion of training. It also had reasonably less variance compared with 15 or 8-trial assessments. We therefore limited our skill assessment sessions to 30 trials. We carried out assessment of skill between PT and NFT programs, and after the completion of NFT program, without any accompanying physical training. These assessments did not present any significant change in skill [Figure 5.5], which was an indication of retention of acquired skill, be it after PT or NFT. We therefore concluded that the change in skill during NFT was primarily due to the application of MI on the selected EEG channels. However, further research under strict experimental conditions with the inclusion of a control group/condition that would receive NFT from unrelated channels is warranted and recommended for future work.

6.5. Future Work

We identified the number of observations (number of stroke survivors or number of skill improvement sessions) as a major limitation of this study. Assuming that the starting point for the recommended future work is the repeat of analysis with larger number of observations, we can list a few related investigations that could be informative towards further development of the proposed methodology.

6.5.1. Regression Models

There are two distinct characteristics of observations that need to be considered when generating the regression models:

- Observations that represent static measures of motor function across different stroke survivors, where each observation corresponds to a separate individual, and

- Observations that represent longitudinal measures of motor function for one stroke survivor, where each observation is a temporal snapshot of impairment from the same individual.

The former represents a cross-participant analysis with no temporal component, whereas the latter represents an individualized longitudinal analysis that may not apply to other participants.

Cross Participant Analysis

In the case of FMU, we carried out cross participant analysis with no temporal dependence between observations. Grouping of participants could be based on the population characteristic that is determined by the baseline appraisal of motor function. This poses no obvious constraint on the number of observations, meaning that more observations would result in a better representation of the populations, and presumably a more reliable conclusion of analysis. Here, the limitation is related to the selection of boundary between different populations and the resulting criteria for generation of different estimation models. We can use the baseline FMU scores as the '*separation-threshold*', where the expected level of improvement is the deciding criteria for grouping of population subsets [9]. Prior studies point towards FMU scores of 25 and 50 as two intermediate thresholds, resulting in three separate populations and our subsequent suggestion for three separate regression models. Here we hypothesize that the separation of population based on the baseline FMU scores would result in better performing regression models and is warranted for further investigation with larger number of participants and wide range of impairment.

Another approach for the selection of boundary between different populations is an empirical search for separation-thresholds. Baseline FMU measures could be sequentially assigned to prospectus separation-threshold in an ascending order. The process would be iterative, where separate models are generated for the observations on each side of the threshold. We speculate that the performance of the model corresponding

to the lower FMU scores will initially increase with the progressive inclusion of observations at higher baseline FMU scores. The score at which the performance of the model starts to plateau, or decrease, would represent the lower score threshold. The process can be repeated towards discovery of subsequent thresholds at higher FMU scores. The approach can be constrained to selection of two thresholds corresponding to the lower and upper separation boundaries, much the same as the recommendations based on the margin for potential improvement in function [9]. The selection of thresholds can also be further finetuned to include multiple separation-thresholds. We believe this to be a well worth study, not only to improve the performance of the proposed assessment method, but also informative towards the relationship between potential gain in function and baseline assessment of impairment.

On a related note, our study did not have any information on the details of structural damage to the brain. The only available information on the stroke survivors were related to the affected hand, which was used to swap the EEG measures between the two brain hemispheres. There is merit in extending the grouping of stroke survivors to include the location and extent of damage to the brain. The resulting models may highlight potential correlation between the contributing FC channels and the structural damage.

It is worth revisiting the dilemma associated with the selection of an appropriate model for assessment of impairment, where selection of the model would be based on the score it is designed to estimate. There are two perspectives related to this dilemma; 1) using the proposed method as an assessment tool in a clinical setting for patient stratification, and 2) using the proposed method to facilitate more frequent assessments towards development of individualized rehabilitation strategy.

1) Development of a generalized model that covers a broad range of FMU scores, be it at a lower accuracy, may address this problem. Having a large number of observations can help gain a better understanding of the validity of this approach. It is worth noting that our analysis with 10 stroke survivors generated a better estimation of the lower FMU scores, when we attempted to generalize the model by reducing the number of predictors (contributing channels) to the two channels with the highest regression coefficients. This is encouraging results with respect to the development of a generalized model that could be used to provide an initial estimate of the FMU score

(at a lower accuracy) that could then be used to select a more appropriate model to generate the final (and presumably better) estimate of motor impairment.

2) We can handle this perspective in a similar manner, whereby the final estimate of the impairment is evaluated in a two-step approach. The first step can be either through a generalized model (as in case-1 above), or through conventional FMU assessment, which will help select the appropriate model that could then be used as frequently as needed to personalize the therapeutic activities. Note that the two-step approach does not actually impact the length of each assessment. The models were presumably generated during the research phase with large number of participants and are simply applied to the same two-minutes of resting state EEG data collected from the stroke survivor.

Longitudinal Analysis

As part of developing a model for estimating the change in motor skill, each participant was asked to go through a longitudinal physical training program (tracing tasks). Objective measures of change in skill (tracing error) were used to build a regression model based on their correlation with changes in FC. We speculate that the sign of regression coefficients is indicative of the temporal change in functional recruitment and contribution of networks towards the change in skill. It is therefore plausible to assume that contribution of some networks diminishes throughout the training program. It is also plausible to assume further training introduces new contributing networks towards improvement of skill. This presents an inherent temporal variability in the regression models, meaning that some predictors can be taken out of the model while new predictors enter the equation. It therefore seems reasonable to expect the need to generate new regression models at some intervals during the progression of change in motor skill. The limits for these predictors are the maximum FC for negative regression coefficients (ceiling) and minimum FC for positive regression coefficients (floor). We speculate that once the predictors reach their limits of FC measures, the estimation accuracy of the corresponding model starts to decrease with any further improvements in skill. This introduces a conflicting requirement on the number of observations for the appraisal of the estimation accuracy. Generalization of approach towards the selection of optimal number of observations would be a topic that warrants future experimental study and investigation.

Complications with the number of new observations, however, is not necessarily problematic when using the regression models for the purpose of channel selection in neurofeedback. The available margin for increasing FC (ceiling effect) limits the number of new observations that can be used with the existing model before having to initiate the generation of a new model. The challenge in this case is the decision on the number of older observations to exclude before generating a new model as new samples become available. An interesting question is how the PLSC algorithms might deal with this temporal change in the contributing channels. The ceiling effect reduces the contribution of a channel towards change in motor function. The behaviour of this channel would therefore appear transitory when considering the addition of new observations without excluding the older samples. The transitory behaviour may reduce the robustness (z-score) of these channels' contribution towards correlation with change in motor function. We evaluate the z-scores through bootstrap resampling and exclude channels at lower z-scores before using the remaining channels to generate the regression models. Investigating the validity of this speculated behaviour of PLSC, through simulation, can be very informative towards the development of algorithms for the rolling generation of new models without having to exclude any of the collected samples.

The motivation for a rolling generation of new regression models is to extend the longitudinal lifecycle of the neurofeedback approach as a complementary therapeutic activity towards improvement of motor function. Generation of new models is an indirect circumvention of ceiling effect due to the limited margin for increasing FC. Our analysis, however, was only focused on channels with negative regression coefficients (-tiveCH), primarily because of our inability to propose an algorithm for using the FC from channels with positive coefficients (+tiveCH). In this study, we did not examine the auxiliary effect of MI on the FC of +tiveCH while reaching the maximum FC in -tiveCH. An auxiliary reduction in the FC of +tiveCH would effectively increase the overall dynamic range of the model by the available margin for the +tiveCH (addition of floor effect), thereby extending the impact of targeted mental imagery beyond the ceiling effect. Study of the auxiliary effects of MI on the channels that were not directly targeted by NF can be informative in this regard. It is noteworthy that the results may also indicate a general limit for the use of neurofeedback as a complementary approach towards improvement of motor function. This is plausible, as it is intuitively unreasonable to expect a continual and unlimited amount of improvement in function through neurofeedback.

Related to the prospects of longitudinal improvements in motor function from the application of guided mental imagery, is the management of expectations. We have already touched on factors such as ceiling effect, and availability of models with enough margin for increasing FC. But there is another factor related to the rate of improvement in function during the course of NFT. A cursory look at the tracing performance during the NFT program (Figure 5.5) seems to indicate an initial learning period of about five days that did not result in a major improvement in motor skill. Additional NFT sessions showed a gradual decrease in tracing error over the subsequent NFT days. But statistical analysis of this data pointed towards insufficient evidence for a structural change in the rate of change in skill during the NFT program. However, the test statistics exhibited *p*-values near our significance threshold ($p = 0.05$), suggesting that future research with larger number of participants may identify a more complicated relationship than a single linear model. Such a structural change in how motor performance might improve, would provide consequential information for future adoption of NFT as a complementary therapeutic activity. Knowing that there may be a learning period during which a major change in skill is not to be expected can help limit the discouraging psychological effects from lack of progress. It is noteworthy that similar learning periods were observed in other neurofeedback studies, but inline with our study, the authors did not consider the observations conclusive [89].

6.5.2. Hybrid Approach from Generalized to Individualized

We already touched on the combined use of generalized and individualized models to assess motor impairment for the purpose of stratification. We can extend this approach to the implementation of NFT for intervention and therapeutic activities. As mentioned before, the neurofeedback approach is contingent on an initial physical training program that triggers measurable change in motor function to facilitate the identification of contributing channels and synchronization frequencies. This may appear as an impediment towards the application of the methodology for individuals with limited ability to partake in the physical training. In such cases, we believe that the alternative approach of neurofeedback on generalized areas of interest may be an appropriate strategy for inducing an initial level of improvement [54]. Continual collection of resting state EEG and conventional assessment of motor function could then be used for subsequent PLS

analysis and identification of relevant channels for generation of individualized estimation models. A shift in strategy from a generalized to an individualized approach can then follow these initial intervention steps with the potential to expedite or possibly increase further improvements in function [56]. Our study was a proof of concept that used only one participant, which makes the hybrid use of generalized and individualized strategy a speculation on our part. Investigation of the validity of such approach with a large number of stroke survivors will be informative towards the clinical application of our proposed method.

6.5.3. Miscellaneous investigations

In this study we used phase synchronization between different neural populations as a measure of functional connectivity between the respective networks. We partitioned the filtered EEG data from each participant into 1-second non-overlapping epochs before applying PLI algorithm to quantify synchronization. We averaged the instantaneous phase synchronization measures over each epoch and selected the peak value as a representative of overall synchronization (FC index). The regression model for estimating FMU was based on this selection of 1-second epochs. We repeated the procedure for different selection of epochs at 0.5-, 2-, and 4-seconds durations, none of which resulted in statistically significant correlation between the FC indices and FMU. We did not investigate the reasoning behind significant correlation at 1-second epochs only. We speculate that coherence occurs in bursts of certain duration and increasing the epoch size would reduce the maximum-coherence through the process of averaging within each epoch. Smaller epochs have the opposite effect, resulting in a reduction of the signal to noise ratio and potentially more spurious maximum-coherence. Temporal distribution of instantaneous synchronization measures prior to averaging may be informative towards understanding the relationship between the length of the epochs and the final connectivity index. This may also provide information on any existing relationship between the bursts of synchronization in the identified FC channels and whether there is an underlying dynamic temporal relationship between the activities of these channels. It is worth mentioning again that we used Morlet filters for computation of synchronization as a measure of Functional Connectivity, which is inherently a non-directional analysis. For

applications requiring directional analysis (e.g., Effective Connectivity) other wavelet options such as Dual Tree Complex Wavelet Transform [93] may be more appropriate.

We did not investigate the minimum detectable change that is achievable through PLS approach for predicting FMU. This is an important topic for future research on this approach and its efficacy as an assessment tool for stroke survivors. Our expectation is that a measure of motor deficit that is independent of physical tasks may provide a higher resolution of assessment and deliver intermediate appraisal of motor function that is not considered to be clinically insignificant. Although the ultimate goal is to induce clinically significant change in function, the availability of reliable intermediate appraisal would be conducive towards personalization of rehabilitation strategy. A follow up study using a new design of experiment with objective measures of behavior that might include time, range, or accuracy of movement with stroke survivors might be informative in this regard.

In closing, the large number of EEG channels and temporal samples makes neural networks an attractive candidate for EEG signal processing and analysis. Our choice of Functional Connectivity as the attribute of interest, increases the number of input features by more than an order of magnitude. This along with a variety of potential processing algorithms, makes a future investigation with deep learning (DL) approach a worthwhile study. Prior study using multilayer convolutional neural network with preprocessed EEG data (spectral and phase information) from a tapping task showed promising results in predicting FMU scores in 14 stroke survivors [58]. But the study did not venture into the interpretation of extracted features in the network. These features would be informative towards selection of stimulation activities for improving motor function, which was the focus of our study. At the time of this report, majority of DL studies had focused on applications that were specific to classification of EEG data for Brain Computer Interface, not prediction of function [94]. Use of DL for different stages of EEG processing may be a viable approach, focusing on improvement of signal to noise ratio, artifact removal, and feature extraction [95]. End-to-end processing with DL to cover artifact removal, signal preprocessing, feature extraction, and classification is also a desirable prospect where one could primarily concentrate on a single optimization model [96]. There are, however, pitfalls with this end-to-end approach. Without an understanding of the behavior of the model and what DL has learnt and used for making the respective classifications, we would not know whether artifacts were removed or incorporated into the optimization model. Interpretability would be an important aspect of future studies [97], [98], [99] for

applications such as the one described in this study, where the learnt features need to be used for development of strategies towards improving motor function.

References

[1] M. Y. C. Pang et al., “A community-based group upper extremity exercise program improves motor function and performance of functional activities in chronic stroke: a randomized controlled trial,” Archives of physical medicine and rehabilitation, vol. 87, no. 1, pp. 1–9, 2006.

[2] R. W. Teasell et al., “Rethinking the Continuum of Stroke Rehabilitation,” Archives of physical medicine and rehabilitation, vol. 95, no. 4, pp. 595–596, 2014.

[3] R. W. Teasell et al., “Time to Rethink Long-Term Rehabilitation Management of Stroke Patients,” Topics in Stroke Rehabilitation, vol. 19, no. 6, pp. 457-462, 2012.

[4] S. J. Page et al., “Reconsidering the motor recovery plateau in stroke rehabilitation,” Archives of physical medicine and rehabilitation, vol. 85, no. 8, pp. 1377-1381, 2004.

[5] C. M. Stinear, “Prediction of motor recovery after stroke: advances in biomarkers,” *Lancet neurology*, vol. 16, no. 10, pp. 826-836, Oct. 2017.

[6] A. R. Fugl-Meyer, L. Jääskö, I. Leyman, S. Olsson, S. Steglind, “The post-stroke hemiplegic patient. 1. A method for evaluation of physical performance,” Scand. J. Rehabil. Med. 7, pp. 13–31, 1975

[7] C. H. Park et al., “Longitudinal changes of resting-state functional connectivity during motor recovery after stroke,” Stroke, vol. 42, pp. 1357–1362, May 2011.

[8] Y. T. Fan et al., “Neuroplastic changes in resting-state functional connectivity after stroke rehabilitation,” Frontiers in Human Neuroscience, vol. 9, Article 546, Oct. 2015.

[9] J. See et al., “A standardized approach to the Fugl-Meyer assessment and its implications for clinical trials,” Neurorehabilitation Neural Repair, vol. 27, no. 8, pp. 732–741, Oct 2013.

[10] M. H. Milot, S. C. Cramer, “Biomarkers of recovery after stroke,” Current Opinion in Neurology, vol. 21, no. 6, pp. 654–659, Dec. 2008.

[11] A. R. Carter, G. L. Shulman, M. Corbetta, “Why use a connectivity-based approach to study stroke and recovery of function,” NeuroImage, vol. 62, no. 4, pp. 2271–2280, Oct. 2012.

[12] C. Grefkes, G. R. Fink, “Connectivity-based approaches in stroke and recovery of function,” The Lancet Neurology, vol. 13, no. 2, pp. 206–216, Feb. 2014.

[13] E. Burke Quinlan et al., “Biomarkers of Rehabilitation Therapy Vary according to Stroke Severity,” Neural Plasticity, vol. 2018, p. 8, Mar. 2018.

[14] L. A. Boyd et al., “Biomarkers of stroke recovery: consensus-based core recommendations from the stroke recovery and rehabilitation roundtable,” International Journal of Stroke, vol. 12, no. 5, pp. 480–493, Jul. 2017.

[15] K. J. Friston, “Functional and Effective Connectivity: A Review,” Brain Connectivity, vol. 1, no. 1, Jun. 2011, DOI: 10.1089/brain.2011.0008.

[16] M. Marino et al., “Neuronal dynamics enable the functional differentiation of resting state networks in the human brain,” Human brain mapping, vol. 40, no. 5, pp.1445-1457, Apr. 2019.

[17] P. Fries, “A mechanism for cognitive dynamics: neuronal communication through neuronal coherence,” Trends in cognitive sciences, vol. 9, no. 10, pp. 474-480, Aug. 2005.

[18] A. M. Bastos, J. M. Schoffelen, “A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls,” Frontiers in systems neuroscience, vol. 9, Article 175, Jan. 2016.

[19] E. S. Claflin et al., “Emerging Treatments for Motor Rehabilitation After Stroke,” Neurohospitalist, vol. 5, no. 2, pp. 77-88, Apr. 2015.

[20] J. García-Prieto et al., “Efficient Computation of Functional Brain Networks: toward Real-Time Functional Connectivity,” Frontiers in neuroinformatics, vol. 11, p.8, Feb. 2017.

[21] J. A. Micoulaud-Franchi et al., “Towards a Pragmatic Approach to a Psychophysiological Unit of Analysis for Mental and Brain Disorders: An EEG-Copeia for Neurofeedback,” Applied psychophysiology and biofeedback, vol. 44, no. 3, pp. 151-172, May 2019.

[22] J. P. Koch, F. Hummel, “Toward precision medicine: tailoring interventional strategies based on noninvasive brain stimulation for motor recovery after stroke,” Current opinion in neurology, vol. 30, no. 4, pp. 388-397, Aug. 2017.

[23] Z. Frehlick et al., “Human translingual neurostimulation alters resting brain activity in high-density EEG,” Journal of neuroengineering and rehabilitation, vol. 16, no. 1, p.60, May 2019.

[24] R. Sitaram et al., “Closed-loop brain training: the science of neurofeedback,” Nat. Rev. Neurosci., vol. 18, no. 2, Feb. 2017.

[25] T. Renton et al., “Neurofeedback as a form of cognitive rehabilitation therapy following stroke: A systematic review,” PloS one, vol. 12, no. 5, p.e0177290, May 2017.

[26] B. Zoefel et al., “Neurofeedback training of the upper alpha frequency band in EEG improves cognitive performance,” NeuroImage, vol. 54, no. 2, pp.1427-1431, Jan. 2011.

[27] M. P. Van Den Heuvel, H. E Hulshoff Pol, “Exploring the brain network: a review on resting-state fMRI functional connectivity,” European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology, vol. 20, no. 8, pp. 519-534, Aug. 2010.

[28] J. Wu et al., “Connectivity measures are robust biomarkers of cortical function and plasticity after stroke,” Brain, vol. 138, no. 8, pp. 2359–2369, Aug. 2015.

[29] P. Nicolo et al., “Coherent neural oscillations predict future motor and language improvement after stroke”, Brain, vol. 138, no. 10, pp. 3048-3060, Oct. 2015.

[30] T. Kawano et al., “Large-Scale Phase Synchrony Reflects Clinical Status After Stroke: An EEG Study”, Neurorehabilitation and Neural Repair, vol. 31, no. 6, pp. 561-570, Jun. 2017.

[31] L. Ma et al., “Changes occur in resting state network of motor system during 4 weeks of motor skill learning,” NeuroImage, vol. 58, no. 1, pp. 226-233, Jun 2011.

[32] T. Wiestler, J. Diedrichsen, “Skill learning strengthens cortical representations of motor sequences,” eLife, vol. 2, p. e00801, Jul. 2013.

[33] G. Guggisberg et al., “Two Intrinsic Coupling Types for Resting-State Integration in the Human Brain,” Brain topography, vol. 28 no. 2, pp. 318-329, Sep. 2014.

[34] R. Sigala, “The role of alpha-rhythm states in perceptual learning: insights from experiments and computational models,” Frontiers in computational neuroscience, vol. 8, p. 36, Apr. 2014.

[35] O. Ozdenizci et al., “Neural Signatures of Motor Skill in the Resting Brain,” IEEE International Conference on Systems, Man and Cybernetics (SMC), DOI: 10.1109/SMC.2019.8914252, Oct. 2019.

[36] L. Manuel et al., “Resting-state connectivity predicts visuo-motor skill learning,” NeuroImage, vol.176, pp. 446–453, May 2018.

[37] I. Faiman, S. Pizzamiglio, D. L. Turner, “Resting-state functional connectivity predicts the ability to adapt arm reaching in a robot-mediated force field,” NeuroImage, vol.v174, pp. 494-503, Jul. 2018.

[38] J. Wu, R. Srinivasan, A. Kaur, S. Cramer, “Resting-state cortical connectivity predicts motor skill acquisition,” NeuroImage, vol. 91, pp. 84–90, May 2014.

[39] O. Ozdenizci et al., “Electroencephalographic identifiers of motor adaptation learning,” Journal of neural engineering, vol.14, no. 4, p. 046027, Jun. 2017.

[40] P. Koch, F. Hummel, “Toward precision medicine: tailoring interventional strategies based on noninvasive brain stimulation for motor recovery after stroke,” Current opinion in neurology, vol. 30, no. 4, pp. 388-397, Aug. 2017.

[41] Z. Frehlick et al., “Human translingual neurostimulation alters resting brain activity in high-density EEG,” Journal of neuroengineering and rehabilitation, vol. 16, no. 1, p.60, May 2019.

[42] R. Sitaram et al., “Closed-loop brain training: the science of neurofeedback,” Nat. Rev. Neurosci., vol. 18, no. 2, Feb. 2017.

[43] T. Renton et al., “Neurofeedback as a form of cognitive rehabilitation therapy following stroke: A systematic review,” PloS one, vol. 12, no. 5, p.e0177290, May 2017.

[44] B. Zoefel et al., “Neurofeedback training of the upper alpha frequency band in EEG improves cognitive performance,” NeuroImage, vol. 54, no. 2, pp.1427-1431, Jan. 2011.

[45] C. Jeunet et al., “Using Recent BCI Literature to Deepen our Understanding of Clinical Neurofeedback: A Short Review,” Neuroscience, vol. 378, May 2018.

[46] W. Nan et al., “Neurofeedback Training for Cognitive and Motor Function Rehabilitation in Chronic Stroke: Two Case Reports,” Frontiers in neurology, vol. 10, p.800, Jul. 2019.

[47] A. Gong et al., “Efficacy, Trainability, and Neuroplasticity of SMR vs. Alpha Rhythm Shooting Performance Neurofeedback Training,” Frontiers in human neuroscience, vol. 14, p.94, Mar. 2020.

[48] A. Sidhu, A. Cooke, “Electroencephalographic neurofeedback training can decrease conscious motor control and increase single and dual-task psychomotor performance,” Experimental brain research, vol. 239, no. 1, pp. 301-313, Nov. 2021.

[49] M. Corsi et al., “Functional disconnection of associative cortical areas predicts performance during BCI training,” NeuroImage, vol. 209, p.116500, Apr. 2020.

[50] L. Giulia et al., “The impact of neurofeedback on effective connectivity networks in chronic stroke patients: an exploratory study,” Journal of neural engineering, vol. 18, no. 5, p.56052, Oct. 2021.

[51] J. Gonzalez-Astudillo et al., “Network-based brain-computer interfaces: principles and applications,” Journal of neural engineering, vol. 18, no. 1, p.11001, Jan. 2021.

[52] L. Allaman et al., “Spontaneous Network Coupling Enables Efficient Task Performance without Local Task-Induced Activations,” The Journal of neuroscience, vol. 40, no. 50, pp.9663-9675, Dec. 2020.

[53] A. Mottaz et al., “Neurofeedback training of alpha-band coherence enhances motor performance,” Clinical neurophysiology, vol. 126, no. 9, pp.1754-1760, 2014.

[54] A. Mottaz et al., “Modulating functional connectivity after stroke with neurofeedback: Effect on motor deficits in a controlled cross-over study,” NeuroImage clinical, vol. 20, pp.336-346, 2018.

[55] N. Riahi, V. A. Vakorin, C. Menon, “Estimating Fugl-Meyer Upper Extremity Motor Score from Functional-Connectivity Measures,” IEEE transactions on neural systems and rehabilitation engineering., vol. 28, no. 4, pp. 860-868, Apr. 2020.

[56] S. Enriquez Geppert, R. J. Huster, C. Herrmann, “EEG-neurofeedback as a tool to modulate cognition and behaviour: a review tutorial,” Frontiers in human neuroscience, vol. 11, pp. 1-19, Feb. 2017.

[57] M. Vukelić, A. Gharabaghi, “Self-regulation of circumscribed brain activity modulates spatially selective and frequency specific connectivity of distributed resting state networks,” Frontiers in behavioral neuroscience, vol.9, p. 181, Jul. 2015.

[58] X. Zhang, R. D’Arcy, C. Menon, “Scoring upper-extremity motor function from eeg with artificial neural networks: a preliminary study,” Journal of Neural Engineering, vol.16, no. 3, p.036013, Apr. 2019.

[59] P. S. Addison, The illustrated wavelet transform handbook, introductory theory and applications in science, engineering, medicine and finance. CRC Press, Florida: Boca Raton, Second Edition, 2017, ch. 2.

[60] J. P. Lachaux, et al., “Estimating the time-course of coherence between single-trial brain signals: an introduction to wavelet coherence,” Clinical Neurophysiology, vol. 32, no. 3, pp. 157–174, Jun. 2002.

[61] K. G. Derpanis, “Fourier Transform of the Gaussian,” 2005. http://www.cse.yorku.ca/~kosta/CompVis_Notes/fourier_transform_Gaussian.pdf

[62] M. X. Cohen, *Analyzing neural time series data: theory and practice*. MIT Press, Massachusetts: Cambridge, 2014.

[63] J. W. Krakauer et al., “Motor Learning,” Comprehensive Physiology, vol. 9, no. 2, pp. 613–663. Wiley. https://doi.org/10.1002/cphy.c170043, Mar. 2019.

[64] F. Varela et al., “The brainweb: Phase synchronization and large-scale integration,” Nature reviews. Neuroscience, vol. 2, no. 4, pp. 229-239, Apr. 2001.

[65] P. L. Nunez, R. Srinivasan, *Electric Fields of the Brain: The neurophysics of EEG*, New York: Oxford University Press, 2006.

[66] R. Srinivasan, W. R. Winter, J. Ding, P. L. Nunez, “EEG and MEG coherence: measures of functional connectivity at distinct spatial scales of neocortical dynamics,” Journal of Neuroscience Methods, vol. 166, no. 1, pp. 41–52, Oct. 2007.

[67] F. Perrin, J. Pernier, O. Bertrand, J. F. Echallier, “Spherical splines for scalp potential and current density mapping,” Electroencephalography and clinical neurophysiology, vol.72, no. 2, pp. 184-187, Feb. 1989.

[68] J. P. Lachaux, E. Rodriguez, J. Martinerie, F. J. Varela, “Measuring phase synchrony in brain signals,” Human Brain Mapping, vol. *8*, no. 4, pp. 194–208, 1999.

[69] G. Nolte et al., “Identifying true brain interaction from EEG data using the imaginary part of coherency,” Clinical Neurophysiology, vol. 115, no. 10, pp. 2292–2307, Oct. 2004.

[70] C. J. Stam, G. Nolte, A. Daffertshofer, “Phase lag index: Assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources,” Human Brain Mapping, vol. 28, no. 11, pp. 1178–1193, Nov. 2007.

[71] M. Vinck, M. van Wingerden, T. Womelsdorf, P. Fries, C. M. A. Pennartz, C. M. A., “The pairwise phase consistency: a bias-free measure of rhythmic neuronal synchronization,” NeuroImage, vol. 51, no. 1, pp. 112–122. May 2010.

[72] P. van Roon, J. Zakizadeh, S. Chartier, “Partial Least Squares tutorial for analyzing neuroimaging data,” Tutorials in Quantitative Methods for Psychology, vol. 10, no. 2, pp. 200-215, Sep. 2014.

[73] A. Krishnan, L. J. Williams, A. R. McIntosh, H. Abdi, “Partial Least Squares (PLS) methods for neuroimaging: a tutorial and review,” NeuroImage, vol. 56, no. 2, pp. 455–475, May 2011.

[74] A. R. McIntosh, N. Lobaugh, “Partial least squares analysis of neuroimaging data: Applications and advances,” NeuroImage, vol. 23, Supp. 1, pp. S250–S263, 2004.

[75] V. A. Vakorin et al., “Detecting Mild Traumatic Brain Injury Using Resting State Magnetoencephalographic Connectivity,” PLoS Computational Biology, vol. 12, no. 12, p.e1004914, Dec. 2016.

[76] H. Abdi, “Partial least squares regression and projection on latent structure regression (PLS Regression),” Wiley Interdisciplinary Reviews: Computational Statistics, vol. 2, no. 1, pp. 97–106, Jan. 2010.

[77] A. Delorme, S. Makeig, “EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis,” Journal of neuroscience methods, vol. 134, no. 1, pp. 9–21, Mar. 2004.

[78] A. Baldassarre et al. “Dissociated functional connectivity profiles for motor and attention deficits in acute right-hemisphere stroke,” Brain vol. 139, no. 7, pp. 2024–2038, Jul. 2016.

[79] T. O. Frizzell et al., “white matter neuroplasticity: motor learning activates the internal capsule and reduces hemodynamic response variability,” Front. Hum. Neurosci., vol.14, p.509258, Oct 2020.

[80] P. Fries, “Rhythms for Cognition: Communication through Coherence,” Neuron *(Cambridge, Mass.)*, vol. 88 no. 1, pp. 220-235, Oct. 2015.

[81] G. Rabiller et al., “Perturbation of Brain Oscillations after Ischemic Stroke: A Potential Biomarker for Post-Stroke Function and Therapy,” International journal of molecular sciences, vol. 16, no. 10, pp. 25605-25640, Oct. 2015.

[82] M. Hassan, F. Wendling, “Electroencephalography Source Connectivity: Aiming for High Resolution of Brain Networks in Time and Space,” IEEE signal processing magazine, vol. 35, no. 3, pp. 81-96, May 2018.

[83] J. Doyon, V. Penhune, L. G. Ungerleider, “Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning,” Neuropsychologia, vol. 41, no. 3, pp. 252-262, 2003.

[84] G. Claeskens, N. L. Hjort, Model Selection and Model Averaging. Cambridge: Cambridge University Press, 2008, ch. 4.

[85] A. Zeileis et al., “strucchange: An R Package for Testing for Structural Change in Linear Regression Models,” Journal of statistical software, vol. 7, no. 1, pp.1-38, Jan. 2002.

[86] S. P. Millard, EnvStats: An R Package for Environmental Statistics. New York, NY: Springer New York, 2013, ch. 7.

[87] R. J. Nudo et al., “Use-dependent alterations of movement representations in primary motor cortex of adult squirrel monkeys,” The Journal of neuroscience, vol. 16, no. 2, p. 785-807, 1996.

[88] A. J. Szameitat et al., “Cortical activation during executed, imagined, observed, and passive wrist movements in healthy volunteers and stroke patients,” NeuroImage, vol. 62, no. 1, p. 266-280, 2012.

[89] S. Boe et al., “Laterality of brain activity during motor imagery is modulated by the provision of source level neurofeedback,” NeuroImage, vol. 101, p. 159-167, 2014.

[90] S. Darvishi et al., “Proprioceptive Feedback Facilitates Motor Imagery-Related Operant Learning of Sensorimotor β-Band Modulation,” Frontiers in neuroscience, vol. 11, p. 60, 2017.

[91] T. Platz et al., “Reliability and validity of arm function assessment with standardized guidelines for the Fugl-Meyer Test, Action Research Arm Test and Box and Block Test: a multicentre study,” Clinical rehabilitation, vol. 19, no. 4, pp.404-411, 2005.

[92] S. Hiragami et al., “Minimal clinically important difference for the Fugl-Meyer assessment of the upper extremity in convalescent stroke patients with moderate to severe hemiparesis,” Journal of physical therapy science, vol. 31, no. 11, pp.917-921, Aug. 2019.

[93] N. Kingsbury, “Complex Wavelets for Shift Invariant Analysis and Filtering of Signals,” Applied and Computational Harmonic Analysis, vol. 10, no. 3, pp 234-253, May 2001.

[94] Y. Roy et al., “Deep learning-based electroencephalography analysis: a systematic review,” Journal of neural engineering, vol. 16, no. 5, p. 51001, Aug. 2019.

[95] A. Essa, H. Kotte, “Brain Signals Analysis Based Deep Learning Methods: Recent advances in the study of non-invasive brain signals,” arXiv.org 2021, https://doi.org/10.48550/arxiv.2201.04229.

[96] S. Nakagome et al., (2022). Deep Learning Methods for EEG Neural Classification. In: Thakor, N.V. (eds) Handbook of Neuroengineering. Springer, Singapore. https://doi.org/10.1007/978-981-15-2848-4_78-1.

[97] N. Mathur et al., “Deep learning helps EEG signals predict different stages of visual processing in the human brain,” Biomedical signal processing and control, vol.70, p.102996, July 2021.

[98] R. Selvaraju et al., “GradCAM: Visual explanations from deep networks via gradient-based localization,” Proceedings of the IEEE International Conference on Computer Vision, pp. 618–626, 2017.

[99] C. Molnar “Interpretable Machine Learning. A Guide for Making Black Box Models Explainable,” https://christophm.github.io/interpretable-ml-book/. leanpub.com 2019.