Universidade do Minho
Escola de Engenharia

Camilo Augusto Silva Rebocho Vaz

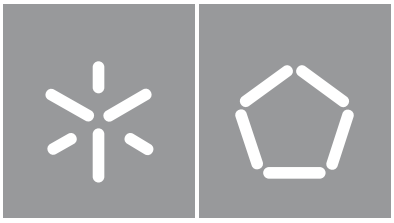
Muscle Synergy Analysis of Lower-Limb Movements

June 2021

Muscle Synergy Analysis of Lower-Limb Movements

Camilo Vaz

UMinho | 2021



Universidade do Minho
Escola de Engenharia

Camilo Augusto Silva Rebocho Vaz

Muscle Synergy Analysis of Lower-Limb Movements

Masters Dissertation
Integrated Masters in Biomedical Engineering
Specialization in Medical Electronics

Dissertation supervised by
Doutora Joana Figueiredo

June 2021

Acknowledgements

All the work carried out during the past year would not have been possbile without the help and support of many people.

First and foremost, I would like to express my deepest gratitude to Prof. Cristina Santos, for welcoming me into the lab and setting up my Erasmus internship, and specially to my advisor Dr. Joana Figueiredo, who helped me immensely every step of the way during the development of this dissertation with inexhaustible patience and understanding.

I am also thankful to Simão, who kindly guided me at every injunction and for his time spent aiding me with every little challenge.

My gratitude also extends to my girlfriend Joana, for supporting me, aiding me and compassionately listening to my venting through this entire process, despite being embroiled in a masters thesis herself.

Finally and most importantly, I would like to thank my mother, who almost didn't make it to see this moment where all her unconditional love and care over the years payed off in a tangible way. This is all thanks to you.

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

Abstract

Neurological disorders and trauma often lead to impaired lower-limb motor coordination. Understanding how muscles combine to produce movement can directly benefit assistive solutions to those afflicted with these impairments. A theory in neuromusculoskeletal research, known as *muscle synergies*, has shown promising results in applications for this field. This hypothesis postulates that the Central Nervous System controls motor tasks through the time-variant combinations of modules (or synergies), each representing the co-activation of a group of muscles. There is, however, no unifying, evidence-based framework to ascertain muscle synergies, as synergy extraction methods vary greatly in the literature. Publications also focus on gait analysis, leaving a knowledge gap when concerning motor tasks important to daily life such as sitting and standing.

The purpose of this dissertation is the development of a robust, evidence-based, task-generic synergy extraction framework unifying the divergent methodologies of this field of study, and to use this framework to study healthy muscle synergies on several activities of daily living: walking, sit-to-stand, stand-to-sit and knee flexion and extension. This was achieved by designing and implementing a cross-validated Non-Negative Matrix Factorization process and applying it to muscle electrical activity data. A preliminary study was undertaken to tune this configuration regarding cross-validating proportions, data structuring prior to factorization and evaluating criteria quantifying accuracy in modularity findings. Muscle synergies results were then investigated for different performing speeds to determine if their structure differed, and for consistency across subjects, to ascertain if a common set of muscle synergies underlay control on all subjects equally. Results revealed that the implemented framework was consistent in its ability to capture modularity ($p < 0.05$). The movements' synergies also did not differ across the studied range of speeds (except one module in Knee Flexion) ($p < 0.05$). Additionally, a common set of muscle synergies was present across several subjects ($p < 0.05$), but shared commonality across every participant was only observed for the walking trials, for which much larger amounts of data were collected.

Overall, the established framework is versatile and applicable for different lower-limb movements; muscle synergies findings for the examined movements may also be used as control references in assistive devices.

Keywords: *Lower-Limb Movement Analysis, Muscle Synergies, Neural Control, Non-Negative Matrix Factorization*

Resumo

As perturbações e traumas neurológicos afetam frequentemente a coordenação motora dos membros inferiores. Uma teoria recente em investigação neuromusculo-esquelética, denominada de *sinergias musculares*, tem demonstrado resultados promissores em soluções de assistência à população afetada por estes distúrbios. Esta teoria propõe que o Sistema Nervoso Central controla as tarefas motoras através de combinações variantes no tempo de módulos (ou sinergias), sendo que cada um representa a co-ativação de um grupo de músculos. No entanto, não existe nenhum processo uniformizante, empiricamente justificado para determinar sinergias musculares, porque os métodos de extração de sinergias variam muito na literatura. Para além disso, as publicações normalmente focam-se em análise da marcha, deixando uma lacuna de conhecimento em tarefas motoras do dia-a-dia, tais como sentar e levantar.

O objetivo desta dissertação é o desenvolvimento de um processo robusto, genérico e empiricamente justificado de extração de sinergias em várias tarefas motoras, unindo as metodologias divergentes neste campo de estudo, e subsequentemente utilizar este processo para estudar sinergias musculares de sujeitos saudáveis em várias atividades do dia-a-dia: marcha, erguer-se de pé partir de uma posição sentada, sentar-se a partir de uma posição de pé e extensão e flexão do joelho. Isto foi alcançado através da implementação de um processo de *cross-validated Non-Negative Matrix Factorization* e subsequente aplicação em dados de atividade elétrica muscular. Um estudo preliminar foi realizado para configurar este processo relativamente às proporções de *cross-validation*, estruturação de dados antes da fatorização e seleção de critério que quantifique o sucesso da representação modular dos dados. Os resultados da extração de sinergias de diferentes velocidades de execução foram depois examinados no sentido de descobrir se este fator influenciava a estrutura dos módulos motores, assim como se semelhanças entre as sinergias de diferentes sujeitos apontavam para um conjunto comum de sinergias musculares subjacente ao controlo do movimento. Os resultados revelaram que o processo implementado foi consistente na sua capacidade de capturar a modularidade nos dados recolhidos ($p < 0.05$). As sinergias de todos os movimentos também não diferiram para toda a gama de velocidades estudada (exceto um módulo na flexão do joelho) ($p < 0.05$). Por fim, um conjunto comum de sinergias musculares esteve presente em vários sujeitos ($p < 0.05$), mas só esteve presente em todos os sujeitos de igual forma para a marcha, para a qual a quantidade de dados recolhida foi muito maior.

Globalmente, o processo implementado é versátil e aplicável a diferentes movimentos dos membros inferiores; os resultados das sinergias musculares para os movimentos examinados podem também ser utilizado como referências de controlo para dispositivos de assistência.

Palavras-chave: *Análise de Movimentos de Membros Inferiores, Controlo Neuronal, Non-Negative Matrix Factorization, Sinergias Musculares*

Contents

List of Figures

List of Tables

Acronyms

ANOVA Analysis of Variance. 59, 60, 65, 67

BF Biceps Femoris. 32, 40, 53, 85, 86, 90–92

BP Band-pass. 12, 37

cosSim Cosine Similarity. 18, 19, 59, 60, 62, 63, 65, 66, 72–80, 82, 88

r^2 Coefficient of Determination. 16–20, 46, 47

CNS Central Nervous System. 9, 10, 26, 27, 30, 80, 96

cvNNMF Cross-Validated Non-negative Matrix Factorization. x, xi, 46–49, 51, 52, 54, 56, 65, 72, 83, 85, 87, 90

DC Direct Current. 37

DOFs Degrees of Freedom. 9

EMG Electromyography. 11, 20, 26–28

FA Factor Analysis. 14, 15, 23

FES Functional Electrical Stimulation. 5, 99

fMRI Functional Magnetic Resonance Imaging. 27

GM Gastrocnemius Medialis. 21, 22, 24, 32, 40, 55, 74, 86, 91–94

HP High-pass. 13, 37

ICA Independent Component Analysis. 14, 15

ICC Intra-Class Correlation. 16

IMU Inertial Measurement Units. 32, 33, 37

LBCI Lower Bound of the Confidence Interval. 54, 56, 57, 83, 85–88, 90, 91, 93, 94

LP Low-pass. 12, 13

MSE Mean Square Error. 42, 44

MUAP Motor Unit Action Potential. 12

MVC Maximum Voluntary Contraction. 13, 33, 37, 40, 41, 91

N_{syn} Minimum Number of Synergies. 46, 47, 49, 50

NNMF Non-negative Matrix Factorization. ix, 14–17, 19, 23, 24, 26–28, 42, 44–47, 54, 63, 83, 95

r Pearson Correlation Coefficient. 18–20, 60, 62, 65–68, 72, 75–79

PCA Principal Component Analysis. 14, 15, 23

RF Rectus Femoris. 21, 24, 25, 32, 40, 48, 57, 85, 86, 89–93

sEMG Surface Electromyography. vii, ix, 10–17, 24–28, 30–33, 36–41, 43–46, 48–54, 74, 85, 87–98

SENIAM Surface Electromyography for Non-Invasive Assessment of Muscles. 11, 32, 33

TA Tibialis Anterior. 21, 24, 25, 27, 32, 40, 55, 57, 74, 85, 86, 89–94

UBCI Upper Bound of the Confidence Interval. 63

VAF Variance Accounted For. x, xi, 16, 43–46, 48–54, 56, 57, 83, 86, 87, 90, 93, 94, 96, 97

Chapter 1

Introduction

This dissertation presents the work developed since March of 2020 in the scope of the fifth year of Integrated Masters in Biomedical Engineering. It has been developed in BiRD LAB (Biomedical Robotic Devices Laboratory) of the Center for MicroElectroMechanical Systems (CMEMs) at the University of Minho. The work presented herein diverges from the one developed during the first semester in the CAMIN team (Control of Artificial Movement & Intuitive Neuroprosthesis), part of the INRIA Sophia Antipolis-Méditerraine Research Group. Work performed at the CAMIN team focused on the development of a classifying framework of shoulder movements as based on inertial measurements with the aim of driving a functional electrical stimulation assistive device for hand-grasping in paraplegic patients.

The work proposed in this dissertation relates to the field of lower-limb movement control analysis and its goal is the development of a framework capable of studying modularity underlying muscles recruitment during lower-limb movements of daily activity. Information gathered through this framework may be used to tune control models of lower limb assistive robotic devices, giving way to quality of life improvements to those affected my motor impairments.

The remainder of this chapter will outline the motivation and problem statement of this dissertation, as well as its goals and contributions.

1.1 Motivation & Problem Statement

One of the major consequences of damage to the nervous system is the development of motor impairments limiting mobility, which in turn are correlated with a reduction in quality of life, propensity for depression and higher likelihood of secondary conditions [1]. Children born with cerebral palsy account

for 1-4 out of 1000 births [2], and one of its main consequences is crouched gait which is a lifelong locomotion impedance. Additionally, around one in four adults globally will have a stroke in their lifetime, with more than 13 million new strokes each year [3]. Of these, more than a third of its survivors will experience significant physical disability hindering their locomotion [4]. Moreover, spinal cord injuries, occurring with an incidence rate of 1 out of 1000 [5], are also responsible for a large share of motor disabilities.

Understanding neural control strategies during muscle recruitment for daily life movements in healthy and pathological subjects may provide useful insights into the development of assistive devices that aim to restore some function following a neurological disability. One such model suggesting a modular nature in neural control has recently been gaining traction in lower limb movement analysis [6]. These recent studies postulate that complex muscle activity for a motor task may be executed through a reduced neural control strategy, involving muscle *synergies*, or the timed co-recruitment of multiple muscles [7]. This hypothesis puts muscle synergies at the center of low-dimensional control as a building block of a modular organization of multiple muscles for movement production.

This understanding of muscle-synergy based control has revealed promising results when applied to rehabilitation engineering. Namely, recent studies have used Functional Electrical Stimulation (FES) with stimulating patterns based on healthy muscle synergies to successfully reduce operating fatigue [8] and to improve long-term rehabilitative outcomes when used therapeutically [9]. Neuromusculoskeletal simulations have also had success making use of these healthy muscle synergies as control inputs in walking [10–12].

This trend reveals a need in studying muscle synergies in diverse movements, specially those associated with daily life activities, in order to gather insights that may accelerate the development of solutions alleviating the impact of motor impairments. However, muscle synergy extraction processes and validating frameworks lack in methodological consistency. The field is divided in choice of extraction algorithm, pre-processing parameters, cross-validating procedure, data structuring decisions and post-extraction validating methods [6, 7, 13]. There is no unifying, evidence-based synergy extraction framework connecting diverging methodologies and applicable to multiple movements and conditions.

Parallel to this challenge, synergy extraction literature for the lower limbs appears to congregate around the same movements such as walking and posture stabilization [6]. Though essential, other movements should be given more attention, specially if they represent important tasks which may be used to develop simpler but still useful task-oriented solutions (e.g. sitting and standing [14]) or to serve as a straightforward basis on which a control framework may be more easily validated (e.g. extending the knee), acting as an understanding stepping stone for more complex movements.

A work tackling both of these major shortcomings in the current muscle synergy extracting paradigm would be useful not only by contributing immediate knowledge to inform decisions in the development of control strategies of assistive devices for the aforementioned tasks, but also by unifying state-of-the-art methodologies on an evidence-based platform that would facilitate the comparison of reported results.

1.2 Goals and Research Questions

The ultimate goal of this dissertation is the design and development of a robust and generic framework capable of extracting synergies from muscle electrical activity data on lower limb movements of daily life (Knee Extension, Knee Flexion, Sit-to-Stand, Stand-to-Sit and Walking) and to validate the findings using statistical inferences, providing a unifying methodology to counter the scattered processes in muscle synergy analysis. This would allow (1) the fast and systematic study of lower-limb muscle synergies, including the minimum number of modules required to explain a given movement and (2) the analysis of significant differences in synergy structure changes across conditions and subjects, providing insights into its usage as a control signal for future work in simulations and assistive devices.

In order to achieve this goal, the following specific objectives were set:

- **Objective 1:** To review relevant literature in order to understand the current paradigm in muscle synergy analysis, the prevailing techniques and their shortcomings and gather the latest understanding in neural modularity of lower-limb movements. This will be addressed in Chapter 2;
- **Objective 2:** To collect electromyographic and kinematic data of healthy subjects during lower limb movements of daily living: sitting, standing, walking and knee extension and flexion; and to organize the collected data into a comprehensive dataset that would facilitate further analysis into these movements. This task is detailed in Chapter 3.
- **Objective 3:** To design and implement a multi-task, generic muscle synergy extraction framework based on robust methods scoured from the literature and to select the minimal number of synergies underlying each movement. The pursuit of this objective is described in Chapter 4;
- **Objective 4:** To validate the framework's robustness and generality through cross-validation methods and statistical analysis of its outputs. Chapter 5 addresses this objective.
- **Objective 5:** To study changes in muscle synergies' structure between subjects and performing speeds for the studied movements. This challenge is approached in chapter 5.

- **Objective 6:** To benchmark the extracted muscle synergy weights and activations by comparison with the literature, when available, and/or to the kinematic variables. This is tackled in Chapter 5.

During the pursuit of these goals, this dissertation aims to answer for the following research questions:

- **RQ 1:** Is there a unifying synergy extraction framework able to capture modularity underlying muscle electrical activity in several lower limb tasks?
- **RQ 2:** What is the minimal number of synergies that can accurately express each of the studied movements' muscle electrical activity at slow speeds?
- **RQ 3:** Are muscle synergies subject and speed dependent for the studied movements?

All of these RQ's will be answered in Chapter 6.

1.3 Main Contributions

With the work done during this dissertation, three main contributions can be mentioned:

- Creation of a comprehensive dataset, gathering lower-limb electromyographic and kinematic data during multiple tasks of daily life to enable further analysis;
- Development of a robust, generic synergy extraction framework and a statistics-based muscle synergy analysis across its runs, performing conditions and participants;
- Reporting of motor modules for the Knee Extension, Knee Flexion, Sit-to-Stand and Stand-to-Sit lower-limb movements, which are poorly studied in the literature.

1.4 Dissertation Structure

This dissertation attempts to bridge gaps in muscle synergy results and methodologies by the development of an independent extraction and validation framework and the recording and reporting of motor module results for a set of lower-limb movements.

Chapter 2 presents a Literature Review for muscle synergy analysis in lower-limb movements. It concludes with a critical analysis of the reviewed literature. Conclusions from this review are then used to support the development of a muscle synergy extraction framework in Chapter 3. Additional input from a small independent study was made to clarify some parameters not explicit in the body of literature in

Chapter 4. Therein, the number of synergies to be extracted in each movement is also determined. Chapter 5 then proceeds to devise, implement and execute null hypothesis tests relating to structural similarity between the gathered results in a hierarchical approach: first examining synergy results across repeated runs of the framework (Section 5.2), then across results from different performing speeds (Section 5.3) and finally between results of each subject (Section 5.4). The dissertation ends in Chapter 6 with a summary of the findings made, and delineates challenges that may be tackled in the future in order to further validate and expand upon the insights described herein.

Chapter 2

Literature Review on Muscle Synergies

In this chapter, the research on muscle synergy extraction and validation will be reviewed. Firstly, an overview of the concept of muscle synergies will be given, its assumptions and evaluation criteria, followed by a more in-depth look at common extraction methodologies and validation frameworks. Subsequently, a review of the current panorama in lower-limb synergies will be performed. The chapter will conclude with a critical analysis of all the examined literature's results, with conclusions drawn in order to best achieve the research goals of this dissertation.

2.1 Muscle Synergy Concept and Principles

A muscle synergy, also known as a muscle motor module, refers to the basic unit of the theorized modular organization of the Central Nervous System (CNS) efferent input during different motor tasks [7, 15]. First proposed by Sherrington in 1910 [16], and supported by evidence gathered by several other studies [17–26]. These researchers hypothesize that the CNS, by recruiting a reduced set of coordinated patterns, controls a small number of muscles rather than coordinating the many thousands of motor units in order to produce movement [15, 27]. In other words, motor behavior results from the linear combination of these synergies.

Different studies point to different motivations and interpretations for the existence of muscle synergies. One preeminent point-of-view states that synergies are a solution to the Degrees of Freedom (DOFs) problem in motor control [28]. The many DOFs afforded by the joints and muscles allows for multiple (i.e., redundant) solutions, conceding flexibility to the nervous system in performing a number of motor tasks (e.g. running, walking, postural stabilization in response to perturbations, etc.). This redundancy poses

a problem to the nervous system: it must choose from a large set of possible solutions because the task requirements are not sufficient to uniquely specify how each muscle and joint must be controlled.

Another interpretation of muscle synergies is that they provide a translation between task level goals (e.g. stabilizing the center of mass) and execution level commands (e.g. activation of individual muscles) [19]. According to this hypothesis, synergies are responsible for selecting muscle groups that when co-activated enable a simplified control of a certain biomechanical feature of the limb (e.g orientation or global limb angle). This understanding makes muscle synergies an integral part of a hierarchical control strategy by structuring complex motor control variables and sensory feedback in a way that allows task-oriented control [29–31].

Additionally, an often-provided explanation for muscle synergies frames it as a primitive solution to motor coordination implemented by phylogenetically "older" neural systems, such as those in the spinal cord [25, 26, 32]. These methods might be re-expressed when the connection between CNS and the spinal cord is damaged (e.g. following a stroke [33, 34]). On the other hand, the CNS may suppress these primitive motor modules by bypassing them in order to express more precisely adaptive behaviors [25, 32, 35].

2.1.1 Synergy Extraction Overview

The process of studying muscle synergies during motor tasks has converged in recent years into a rough common approach: sEMG is measured from a large number of muscles during a complex motor task (or several tasks); secondly, a factorization algorithm is run in order to identify a set of synergies from the resulting dataset; thirdly, a statistical analysis is made to determine whether the recorded sEMG signals can be adequately described as a combination of these synergies; fourthly, a parallel is drawn between the identified synergies and task relevant variables. With this approach, a large set of motor tasks has been suggested to be produced using muscle synergies [6, 7, 36].

Even though these steps are relatively agreed upon, there is no unifying methodology to extract muscle synergies [7]. Many sources of variability are introduced by the choice of factorization algorithm, criteria for dimensionality reduction and data pre-processing, cross-validation and evaluation of quality of reconstruction. These factors present a major obstacle to the successful comparison of results obtained by different methodological paths [37]. Therefore, within the following subsections, an overview of the current techniques and criteria typically used among these studies will be given for each step, with particular focus on the systematic reviews comparing results from the distinct state-of-art methodological approaches.

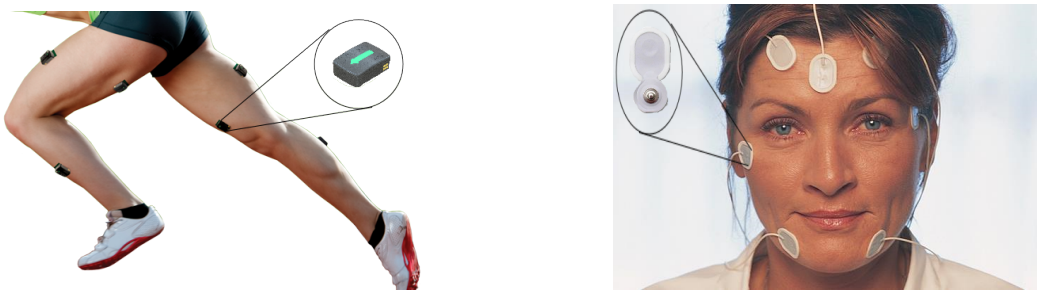
2.1.2 sEMG Data Acquisition

To study muscle coordination, both muscle electrical activity and force patterns of individual muscles during a motor task may be used as a quantitative metric [38]. However, due to the lack of an accurate experimental methodology to estimate muscle force, muscle force modeling approaches are not commonly applied [39]. As such, most modern muscle synergy analysis use common features extracted from electrical muscle activity data to infer this neural modularity. This muscle activity is typically measured using Surface Electromyography (sEMG) data, due to its non-invasive nature and acquisition sensors' increasing availability and affordability [40–42].

Electromyography (EMG) is the extracellular recording of bioelectrical activity generated by muscle fibers [43]. A sEMG electrode typically consists of a pair of poles aligned with the muscle fiber direction. The sEMG electrode is usually made of silver and its area of contact with the skin may be wet (using a silver-chloride gel), or dry in the absence of a conducting medium [44]. Pre-treatment of the instrumented area of skin may be performed to reduce impedance and improve the signal-to-noise ratio. Typical treatments include rubbing with an abrasive paste, cleaning with an alcohol wipe and shaving the hair on the skin [45]. A common issue with sEMG sensors is their sensibility to location, as slight shifts in electrode placement drastically affect the recorded signals, and also to sweat which may naturally occur during a long recording session for demanding motor tasks. Thus, donning and doffing of electrodes should be minimized and carefully performed to be similar to previous configurations [45].

In the muscle synergy analysis literature, most authors settle on commercial bipolar Ag/AgCl electrode configuration [37, 46–52], and choose placement and skin preparation based on guidelines set by the Surface Electromyography for Non-Invasive Assessment of Muscles (Surface Electromyography for Non-Invasive Assessment of Muscles (SENIAM)) project [53]. SENIAM is a concerted action in the Biomedical Health and Research Program of the European Union whose stated goal is to integrate research on sEMG to facilitate exchange of data [53] by standardizing sEMG data collection and processing among researchers. The most common commercial sEMG acquisition systems in synergy research are the Ambu Neuroline 720 01-K/12 [13, 48, 49, 54] (Fig. 2.1b), the Konigsberg Instruments Myopac Wireless/T50 [10, 46, 55, 56], and Delsys Trigno [57–62] (Fig. 2.1a).

Prolonged recording of bioelectrical activity of a muscle allows for a quantitative description of its activity over time. For synergy analysis, an EMG signal should ideally provide a selective recording, uncontaminated by neighboring muscles while faithfully detecting any activity in the target muscle. In practice, this is difficult to achieve simultaneously. Overall detection of the bulk of the muscle can only be achieved with nonselective electrodes, while selective recordings from small muscles can only be achieved with



(a) Wireless Delsys Trigno [63].

(b) Wired Ambu Neuroline 720 [64].

Figure 2.1: sEMG recording systems examples

selective electrodes [43].

2.1.3 Pre-processing of sEMG data

After sEMG acquisition on a dynamic movement, the first step in the synergy extraction process consists in pre-processing the sEMG signal in order to generate linear envelopes describing the activation of each muscle during the motor task. This is typically achieved by performing the following sub-steps [37, 65]:

Band-pass (BP) filtering removes movement artifacts, high-frequency noise, drifts, aliasing effects.

Full wave rectification computes the sEMG envelope. It is additionally useful because it shifts the peak of the spectrum of the Motor Unit Action Potential (MUAP) toward the firing rate frequency, while leaving the firing spectrum unchanged [66].

Low-pass (LP) filtering ensures that no high frequency bursts of activity alters the envelope shape. The lower the cut-off frequency, the smoother the envelope.

Normalization in amplitude assigns equal initial importance to each muscle, allowing non-biased synergy extraction. Normalization in time to a common period allows the direct comparison between movements of different duration.

The choice of parameters for each of these sub-steps affects the validity of the results, as pre-processing directly alters the data available to the factorization algorithm and, consequentially, the number of extracted synergies and the structure of the weighting and activation coefficients themselves [37, 61].

Parameters for the BP filter are relatively common across the synergy extraction literature. For this step, the usual chosen configuration is a 4th order [11, 46, 67] (sometimes 3rd [8, 9]) zero-lag Butterworth

filter. Regarding the bandwidth, several articles point to a LP cut-off frequency from 400 Hz [9, 62] up to 1000 Hz [68] with in-between frequencies of 450 Hz [62]. The High-pass (HP) cut-off frequencies ranges from 10 Hz [68] up to 40 Hz [9], with intermediate values of 20 Hz [8]. On review verified that two bandwidth ranges ([50-500] Hz and [20-500] Hz, 6th order Butterworth) had no significant effect on neither the number of extracted synergies nor on the resulting weighting and activation coefficients on *most* subjects [37].

Full-wave rectification is the process by which negative values in a given signal are transformed into positive, symmetrical values. For sEMG data analysis, this is universally achieved by computing the absolute values of the input data [133].

Prior research has used a wide range of LP filter configurations to smooth the rectified envelope. These range from 1 [70] to 40 [41], including 4 [46], 10 [71–73], 15 [74], 20 [75, 76], 30 [77], and 35 [78] Hz. Systematic studies analyzing a range of cut-off frequencies [37, 65, 79] have reported that reconstruction accuracy of the original signal for a given number of synergies decreased with increasing LP cut-off frequency.

It has been proven than normalization strategies greatly impact the explained variance of the reconstructed sEMG signals [36, 37, 65]; despite this, techniques vary greatly in the literature [48, 80, 81]. Normalization processes can be roughly divided into two categories: normalization by Maximum Voluntary Contraction (MVC), measured before the trial, and by maximum amplitude of the signal during the trial itself. In MVC normalization, subjects are asked to perform a muscle-specific isometric contraction against resistance provided by a researcher. This is typically repeated several times, with a break between each contraction to prevent muscle fatigue. Then, the maximum value obtained from the recorded signal is used to normalize the data for that subject and muscle. Maximum amplitude methods, on the other hand, use the maximum contraction achieved during task performance as the normalization constant. MVC normalization has been found to be best in comparing muscle activity levels and activation patterns between muscles, tasks and individuals in non-synergy related studies [82]. On the other hand, maximum-amplitude during task normalization is reportedly adequate to compare activations between individuals over time, it does not allow comparison between tasks or individuals. In the synergy context, the normalization method modifies the cumulative variance explained by a set of synergies, with MVC normalization leading to higher values of reconstruction accuracy [37].

2.1.4 Muscle Synergy Extraction Methods

Factorization is the most commonly used muscle synergies analysis tool [83]. The goal of this type of algorithms is decomposing acquired sEMG activity patterns into a smaller dimension of time-varying signals, often denoted by H and referred to as muscle activations [80], plus a matrix of weighting coefficients (W) [80], that can be linearly combined to reconstruct the original sEMG signals. Activations are sets of basis functions that represent the time-varying component of the signal. Weights correspond to scalar values that represent activity patterns across all sEMG signals. Collectively, one weight vector and its corresponding activation make up a muscle synergy/motor module [84]. A schematic illustration of the factorization output is represented in Figure 2.2.

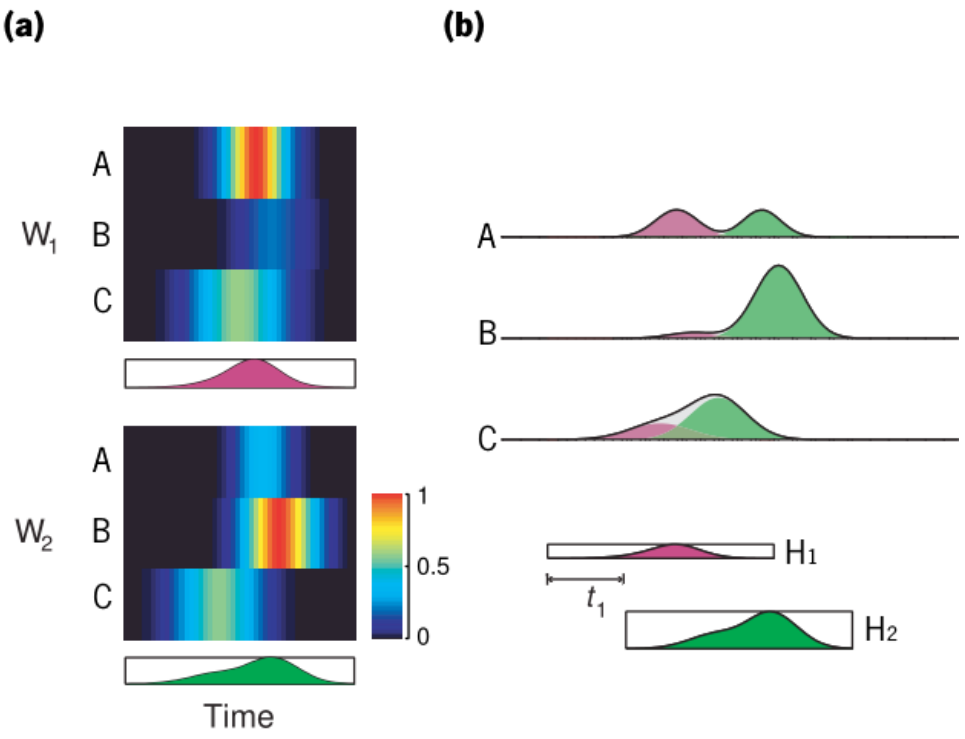


Figure 2.2: Schematic representations for synergies. Two time-varying synergies (W_1 and W_2) are shown in (a). Each synergy specifies a weighting coefficient for one of the three muscles (A-C), indicated by the color of the bar, and a temporal profile for these weightings. An observed pattern of muscle activations is created by scaling each synergy, temporally shifting them (t), then adding them together linearly (b). The purple and green waveforms indicate the contribution from each of the original synergies to the observed response. Adapted from [7].

Commonly used numerical methods to perform decomposition include:Principal Component Analysis (PCA) [57, 85–87], Independent Component Analysis (ICA) [32], Factor Analysis (FA) [37, 88, 89] and NNMF [7, 41, 51, 90–92].

PCA [93] is a linear eigenvalue decomposition technique that consists in transforming the original data vector set into a new mutually uncorrelated set that represents the covariance of the input data. These new vectors are called principal components, and they are arranged according to their variance, effectively remapping the original data into a lower dimensional space. ICA is non-linear blind-source separation technique that identifies the statistically independent sources that can be re-combined to generate a mixed set of signals [32]. FA, similarly to PCA, uses eigenvalue decomposition to produce eigenvectors of the covariance matrix; those with eigenvalues > 1 are considered significant and represent the weighting coefficients, otherwise they're considered noise and are discarded [86]. NNMF creates a parts-based representation of the final signal using only positive, additive components through iterative updates on a set of initial random matrices [94]. This last method stands out from the others for having recently become the most commonly used factorization method for muscle synergy analysis [7, 11, 27, 41, 49, 51, 55, 78, 90–92, 95–97].

A recent study [42], comparing the similarity between the probability distribution of the activations computed by these 4 factorization methods with that of the raw sEMG data during running and walking at different speeds, concluded that NNMF outperformed all other techniques in this metric. Moreover, in the same review, the authors claimed that 62.68% of the reviewed literature between 1999 and 2018 used NNMF, further strengthening the case for universal adoption of this technique as it allows more reliable comparison between results from different authors. Some authors note that NNMF's non-negativity condition is a useful attribute in identifying meaningful synergies, because it prevents the outputs from containing negative activation of the muscles [61], which obeys to the common understanding of muscle contraction. An important observation arising from the variety of factorization algorithms used in muscle synergy literature is that all appear to perform similarly regardless of technique [98], which suggests that the extracted modules are likely not an artifact resulting from the choice of method but rather reflect basic aspects of muscle activations [99].

An important step in an NNMF problem for synergy extraction is the structure of the data that is given as input to the algorithm. The relevant literature appears to usually analyze sEMG signals for a given motor task in the form of single trials, average across trials or concatenated trials, seldom providing a justification for their choice [13]. Early research using NNMF for synergy extraction in lower limbs during walking reported a reconstruction quality of over 90% investigating single step cycles [85, 100], whereas analysis which concatenated trials showed worse results [50, 54, 72, 101]. However, the observed lower reconstruction quality might not be due to intrinsic better performance of single trial analysis, but rather its inability to capture natural step-to-step variability in the sEMG signals [13], and its results may not

translate accurate motor modules underlying a movement.

A recent study explored differences in extracted synergies during treadmill walking when single, averaged or concatenated step-cycle sEMG data were given to the NNMF algorithm [13]; while additionally studying the effect of the number of step-cycles used for concatenation and averaging. They found that, despite the slightly greater observed reconstruction quality in averaging the sEMG signals compared to concatenating, if the resulting synergy weights from the averaged approach were used to reconstruct the concatenated activations, the results were very poor. In addition to this, they observed that the higher the number of concatenated step cycles fed to the NNMF algorithm, the more accurately its weights reconstructed the activations computed by concatenating all available trials. This suggests that a high variance-explained metric might disguise true reconstructing ability in averaging results and that when concatenating trial data prior to factorization, a large amount of data is preferable, though no claims are made regarding the best overall technique [13]. This agrees with the conclusions of the authors of another recent study [62], that concatenating sEMG trials is a more repeatable and robust technique than traditional single-trial or averaging approaches when multiple subjects' data are factorized together. They justify this by noting that when using a concatenated strategy, the NNMF algorithm is forced to find one concatenated matrix of activations, avoiding sparse factorizations from different local minima among the same subjects or among a series of similar trials for the same subject by effectively reducing the dimension of the search space. They show that this leads to a more robust solution as evidenced by their observation that Intra-Class Correlation (ICC) coefficients were generally greater for a concatenated approach (0.72-0.90 for four synergies) versus a single-trial technique (0.63-0.8 for the same number synergies) when comparing synergies across subjects during a weight bearing, force matching task [62].

In order to evaluate the accuracy of the parts-based representation on the original sEMG data, a metric quantifying the performance of the factorization algorithm is needed. This is typically inferred by the reconstruction accuracy of the factorized approximation (i.e. how accurately the product of the factorization represents the original signal), which in turn is usually assessed through Variance Accounted For (VAF), whereby the differences between original muscle contraction patterns and those reconstructed by linearly combining synergy weights and activations are summed and normalized to the amplitude of the original signal. This is typically formulated by $VAF = 1 - SSE/TSS$, where SSE is the sum of the squared residuals and TSS the total sum of squares of each sEMG data point. The similar Coefficient of Determination (r^2) is also used for the same effects but yields high values when the shapes of the signals are well-matched regardless of amplitude [92].

2.1.5 Muscle Synergy Validation

Computed synergies suffer the same problems as supervised learning challenges [102]: a set of motor modules might accurately reconstruct a sEMG signal but may yield low VAF when its weights or activations are multiplied with corresponding factorized parts from another sEMG signal, even if *a priori* the signals seem similar. The cause of this phenomenon is two fold: firstly, the local minima issue may cause two separate factorizations to result in two very different sets of motor modules, even if both are capable of reconstructing the signal; secondly, if the dataset used for decomposition happens to include more variability due to natural inter-subject or inter-trial differences or issues with electrode placement, its factorization result may misrepresent actual neural modularity and generalize poorly for other trials and/or subjects.

Therefore, a validation step is needed, where synergies extracted from a sEMG dataset are correlated to the muscle synergies extracted from an independent data pool [27, 68, 75, 92]. However, a large number of studies do not apply this step. Inspired by the solutions provided by the fields of supervised learning and regression analysis, Cheung *et al.* [68] implemented a cross-validation process in their study of synergies in frogs' hindlimb muscles before and after deafferentation. To estimate the optimal number of synergies during jumping and swimming tasks, they divided the sEMG signals into four equal partitions. Three of the four segments were pooled together and synergies extracted from this dataset. These synergies were then validated by fitting them to the remaining unused partitions (using NNMF with weights held fixed and activations updated across iterations) and the quality of the fit was quantified using the r^2. This procedure was repeated 20 times at each number of synergies, each time with different random initial weights and activations and with different randomly selected data partitions for cross-validation. In addition, they repeated the process on the same partitioned data segments, but randomly and independently shuffling samples for each muscle, yielding a baseline cross-validation r^2. A r^2 greater than this value should indicate that systematic variation in the database is captured by the synergies.

Torres-Oviedo and Ting [41] proposed a similar approach in their study of human responses to postural perturbations but used a 60/40 split across 10 re-runs of the synergy extraction framework. Further authors have applied this simplified form of cross-validation for their studies, re-running the factorization framework, [50] or a combination of both [51], repeated runs and determining weights and activations from independent datasets. Some unilaterally formulate their own approach to validation: Lambert-Shirzad and Van der Loos [61] used k-fold cross validation ($k = 20$) where each fold was validated itself with an 80/20 split.

2.1.6 Similarity

In the muscle motor modules problem space, similarity refers to the degree of agreement between weights and activations from one factorization output and another. Quantifying the similarity between weights means understanding common information between which muscles and to what degree they are recruited; computing inter-activation similarity informs us about the timing and shape of this recruitment. Studying similarity between sets of synergies may be relevant to compare conditions among the same tasks (e.g. different velocities during walking [51], postural balancing with and without stepping [27], squatting before and after induced fatigue [52], etc.), but it may also be useful in inter-trial and inter-subject comparison. In order to properly quantify similarity, a reliable and easily-interpretable choice of metrics is necessary. These metrics may be applied to the weights, activations or both.

One of the first studies attempting to compute similarities between synergy weights made use of Cosine Similarity ($cosSim$) [68]. This measure evaluates similarity by calculating the cosine of the angle between two vectors (i.e. two weighting or activation vectors) normalized to have the same magnitude. Its simplicity and readability has led to it being widely adopted as a similarity measure in the synergy extraction literature [36, 51, 91].

Other researchers have used the Pearson Correlation Coefficient (r) to compute similarities between pairs of both weights and activations [41, 103]. The squared form of this expression is also used for the same effects [92].

Santuz et al. [48] measured both weight and activations' similarity using r^2 and $cosSim$ to evaluate inter-day repeatability of a synergy extraction method. Analyzing performance of both metrics, they found that (1) both were correlated for weights and activations, albeit with different co-domains; (2) r^2 demonstrated a larger evidence-based output for activation analysis. This suggests that $cosSim$ may be adequate to measure agreement between weight's similarity but not for quantifying activations' degree of agreement.

A drawback in using the same characteristic to measure similarity in both weighting and activations coefficients is the fact that information about differences in the $timing$ of curves representing activations is not captured. Furthermore, although r^2 and $cosSim$ inform about the matching of curves in shape and amplitude, they offer no insights into the root cause of a mismatch (i.e. if it is due purely to differences in shape and amplitude, or if one signal is only phase-shifted in relation to the other). A variety of metrics have been used to quantify this characteristic [47, 85], but most recent publications have turned to the measurement of cross-correlation [38, 47, 49, 91, 104].

Frère *et al.* [91] calculated the cross-correlation and used the maximum of the output as a measure of

matching between curves and further defined the timing of activation as the index on which this maximum occurred.

A common goal in similarity studies is to find a subset of motor modules within the extracted pool that are similar to a given reference. Immediately, an issue arises: which value should be defined as the threshold? Here, the relevant literature may be divided into one of two approaches: some arbitrarily choose a value, such as $r^2 > 0.55$ [41], $cosSim > 0.60$ [105], $cosSim \geq 0.75$ [106], $cosSim \geq 0.80$ [13], $cosSim \geq 0.90$ [68], providing no further information of the determination process; others set the value through statistical methods. For instance, some authors set a Pearson Correlation threshold by exploiting the fact that if r is larger than a tabled value, the null hypothesis that the groups are uncorrelated can be rejected, for a given significance level. To provide additional statistical support for this choice, usually a *bootstrapping* study is performed to validate that threshold, by generating a large number of random permutations of the muscle synergies, computing the corresponding r-values and building a distribution with the expected mean and standard deviation. If the pre-selected r is found to belong to a high percentile of the distribution, the correlation is thus empirically proven to represent a higher-than-chance similarity. This statistical approach has the added benefit of allowing the assignment of a confidence value to the similarity findings [91, 92].

2.1.7 Muscle Synergies Sorting

A common problem in studying the similarity between synergies of independent factorizations is that its outputs often require a functional sorting in order to align muscle synergies by role in movement control. An illustration of this challenge may be observed in Fig. 2.3: run 8 has its synergy indices switched compared to the remaining runs and if left unsorted, an analysis of similarity would output misleadingly low values despite the functional role of the synergies being consistent across runs.

A solution developed by Torres-Oviedo and Ting [41] for comparing synergies across different subjects proposed reorganizing the indices by finding the maximum r^2 between repeated subjects. It begins by averaging synergy weights activations across trials for each subject. Then, an initial sorting is performed by grouping muscle synergies based on whether the r^2 between weights and/or activations of an arbitrary reference subject and another participant is greater or lower than a predetermined value. Afterwards, an averaged set of weights and activations is determined from the set which passed this similarity test. Then, with these matrices set as reference and through an iterative process, new r-values are calculated, a new round of grouping is performed and the averaged set updated. The end result is a set of synergies common to all subjects and r values that may serve both as a sorting parameter but also as a measure

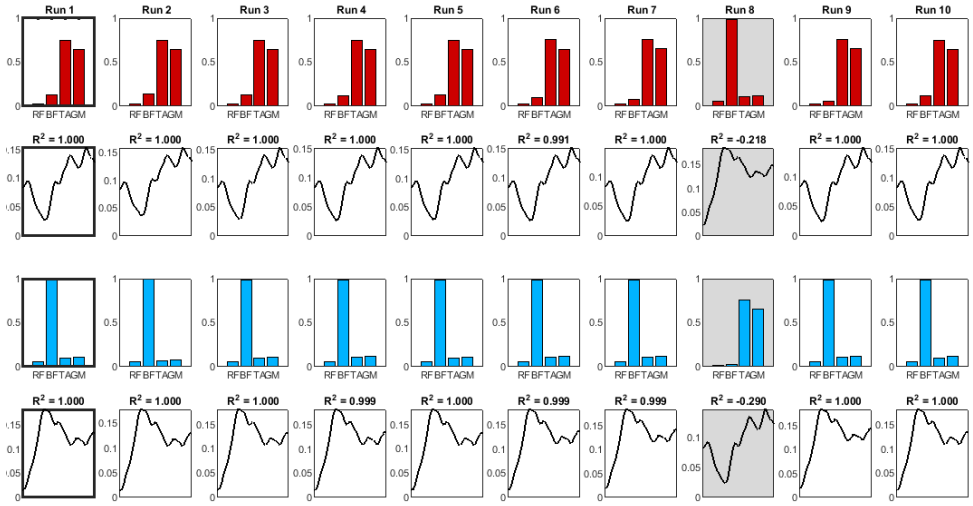


Figure 2.3: Representation of 10 runs of the NNMF algorithm, averaged across trials. Each column represents a different run; each pair of rows a synergy. Activations and weights from run 1 (thick border) were set as a reference to evaluate similarity among the remaining runs. A grey background indicates that the r for the activation of that run is below an arbitrary threshold; possibly signaling a dissimilarity.

to evaluate this generalizability of the computed synergies across subjects. Other studies [51, 107] which implemented the same sorting strategy, used the cosine similarity between synergy weights as a sorting parameter, forgoing a comparison of activations altogether.

2.2 Lower Limb Muscle Synergy Analysis

Over the last two decades, there has been a push for examining EMG signals under a wide range of behavioral conditions in an effort to provide evidence supporting the muscle synergy hypothesis, to deepen understanding on motor impairments, to optimize training in sports rehabilitation and performance and to discover low-dimensional solutions to a motor task in order to model control frameworks [6]. This last field of study is of particular interest to this work as it focuses on task-specific goals and on whether specific complex behaviors can be produced effectively using combinations of motor modules, though conclusions may also be drawn from the other fields.

Throughout this section, a review of recent literature regarding lower limb muscle synergies during motor tasks will be given, with particular focus given to movements relevant for daily life, namely walking and Discrete Motor Tasks. Discrete Motor Tasks, such as sit-to-stand and stand-to-sit are motions important for daily life activity, requiring whole-body coordination and thus commonly impaired by neurological disorders [108, 109]. The most energetically demanding sub-task of these two movements, associated with body ascent and descent, is knee extension and knee flexion, respectively [110]. Despite the fact

that they are indispensable for sitting and standing, among other activities, no studies were found which examined muscle synergies during knee extension and flexion.

2.2.1 Gait

Most lower limb synergy literature focuses on gait analysis as it consists of an highly complex movement involving various sets of muscles being recruited at precise moments in time, while being cyclical and thus easily repeatable and divisible [10, 27, 36, 42, 46, 48, 51, 57, 65, 67, 72, 78, 85, 101, 107, 111–114]. Furthermore, walking and running biomechanics have been extensively studied, with an established general consensus on the composition of gait sub-tasks and how each one influences a stride [115].

The consensus among these early studies is a five-synergy model of human locomotion for walking speed for both level-ground and treadmill walking [10, 57, 85, 87]. The first synergy (composed of activations by the Gluteus Medius, Gluteus Maximus, Vastus Medialis and Rectus Femoris (RF)) corresponds to the biomechanical sub-task of body support, where the body is decelerated in early stance. Synergy 2 (Soleus and Gastrocnemius Medialis (GM)) also provides body support plus forward propulsion in late stance. Motor module 3 (RF and Tibialis Anterior (TA)) acts to decelerate the leg in early and late swing while generating energy to the trunk throughout swing. The fourth synergy (Hamstrings) works by absorbing leg energy (decelerating it) in preparation for foot contact while increasing the leg energy in early stance. Finally, synergy 5 (Iliacus) serves to accelerate the leg forward in pre and early swing by generating energy directly to the leg while simultaneously shifting energy from the trunk to the leg [10, 85–87]. Some studies opt to not record activity of the Iliacus muscles, and therefore only report four synergies, omitting module 5 [10, 46]. Others choose not to record the Iliacus, but attribute Module 5 to a general activation, with no clear bursts, of RF and Gluteus Maximus [13], justifying its relevance as encoding systematic information because they are robust across subjects and especially at higher cadences [74]. One such model is represented in Fig. 2.4.

Ivanenko and colleagues [80] became the first to observe a consistent phase shift in synergy activations with increasing speed, coincident with a kinematic shift in the end of the stance phase. They demonstrated that both the activation peak of all components and the lift-off of the step occurred progressively earlier as treadmill speed was increased. Thus with increasing speed, the duration of stance as a percentage of gait cycle decreases, and the foot is brought through faster during swing to place the heel for the subsequent stance earlier. This phenomenon was observed by several others since [46, 51, 57]. Even though several studies have reported synergy findings for a range of performing speeds (see Table 2.1), none reported

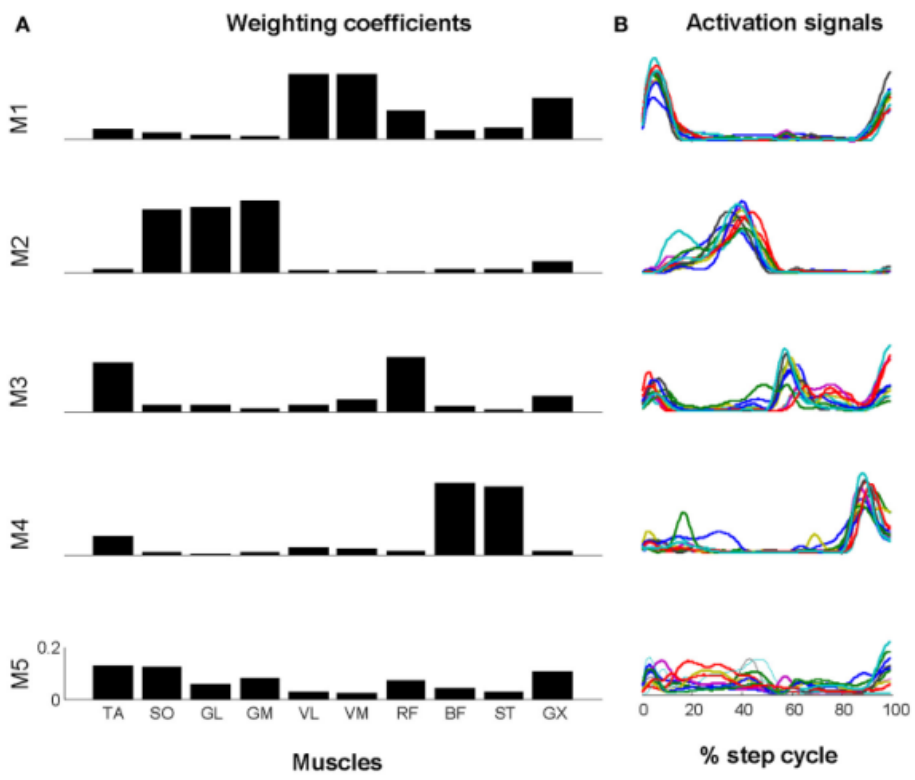


Figure 2.4: Muscle synergies extracted during treadmill walking from twelve healthy subjects. (A) Averaged weights across subjects. (B) Activations for individuals subjects, with each represented by a color. Synergy 5's (M5) existence is controversial; some authors drop it in favor of a reduced number of synergies, others who record a larger set of muscles use its index to represent other motor modules, such as the contribution by Iliacus. Taken from [13]. *TA*: Tibialis Anterior; *SO*: Soleus; *GL*: Gastrocnemius Lateralis; GM: Gastrocnemius Medialis; *VL*: Vastus Lateralis; *VM*: Vastus Medialis; *RF*: Rectus Femoris; *BF*: Biceps Femoris; *ST*: Semitendinosus; *GX*: Gluteus Maximus.

a changing number of motor modules with varying speed, at least for the performed range and for each author's own criteria for determining the minimum number of synergies. The exception was Kibushi *et al.* [51], who claimed a minimum number of synergies of four for the 2.0-4.0 km/h range and five modules for 4.5-8.0 km/h. Additionally, authors who recorded muscle electrical activity from level-ground walking rather than treadmill gait did not report a different number of synergies [86, 87].

Capellini *et al*. [57] also found five muscle synergies in the distinct gaits of walking and running, and further observed that the major differences were concentrated during stance, while walking and running timing was basically the same during the swing phase. In running, the activation peak of the component expressing the posterior calf muscles (such as the GM, Gatrocnemius Latealis, Peroneus Longus and Soleus) occured earlier than the corresponding motor module in walking, whereas other activation timings remained unchanged across gaits. One hypothesis attempts to explain this phenomenon by noting that at higher walking speeds, the ankle extensors are loaded toward the end of stance in an unfavorable portion

of their force-velocity curve, so their force production is limited even as levels of activation increase. By shifting the activation to an earlier phase of stance, the activation is shifted to a more favorable force production range [116].

A modular organization has also been consistently found in impaired walking, though with a lower number of motor modules. This trend has been observed in individuals following neurological injury [60], stroke [46], incomplete spinal cord injury [113] and cerebral palsy [71]. Often, in post-stroke subjects, these impaired synergies resemble merging of specific unimpaired synergies [46].

An important finding following the establishment of a consensus in a four-to-six synergy model of human gait was the successful use of such a strategy (with a few minor modifications regarding the timing of activations) to drive a complex musculoskeletal simulation of the human leg during locomotion. This was firstly proven for a 2D sagittal-plane biomechanical model [10] and later for a 3D model [11, 12] after the inclusion of an additional synergy accounting for mediolateral balance control and contralateral leg swing. These studies theoretically demonstrate that a synergy based control allows for effective locomotion and that in practice a set of motor modules can serve as a basic neural control elements in order to generate the task-specific biomechanical function of walking. Furthermore, it has been recently shown through similar computer simulations that the walking dynamics of unimpaired gait can be reproduced using the lower dimensional control spaces of impaired motor modules such as those present in post-stroke subjects [112].

2.2.2 Sit-to-stand and stand-to-sit movements

Standing from a seated position is essential for human activity as standing upright is a crucial pre-requisite for bipedal walking and a precursor for numerous daily life activities [110]. Understanding the neural control organization of this motor task along with the reciprocal stand-to-sit may provide valuable insights into low-dimensional control of these movement which, in future work, may serve as a basis for targeted strength training or rehabilitative devices such as electrical stimulation.

Muscle synergy research into these motor tasks, however, is scarce. Table 2.2 presents a summary of all the found muscle synergy literature for the sit-to-stand motion. One study examined sit-to-stand from a knee-height stool in four young (22.5 $\pm$ 1.2 mean age) and three older (72.0 $\pm$ 2.0) subjects in two conditions: rising at one's natural speed and rising "as fast as possible" [14]. sEMG data, collected from seven lower-limb muscles ipsilaterally (TA, Soleus, GM, Vastus Lateralis, RF, Semitendinosus and Gluteus Maximus), was correlated with each biomechanical subtask by comparing with kinematic variables. They found three synergies sufficiently explained the variance in the sEMG data (Fig. 2.5): the first, composed

Table 2.1: Reported numbers of synergies for walking in healthy subjects. *LW - Level Ground Walking; TW - Treadmill walking; B - Bilateral placement of electrodes; I - Ipsilateral plcement*

Study	No. Subjects	No. Synergies	No. Measured Muscles	Speed Range	Gait Condition	Factorization Algorithm
[87]	10	5	16 (B)	"brisk walk"	LW	FA
[117]	6	4	7 (I)	0.5*"Normal Speed"-2.0*"Normal Speed"	TW	PCA
[85]	6	5	12-16 (I)	1.00 - 5.00 km/h	TW	PCA
[10]	14	5	8 (I)	4.32 km/h	TW	NNMF
[86]	18	4	16 (I)	106 steps/min	LW	PCA
[46]	20	4	8 (I)	1.08 - 6.48 km/h	TW	NNMF
[57]	8	5	32 (I)	3.00 - 12.00 km/h	TW	PCA
[51]	10	4-5	12 (I)	2.00 - 8.00 km/h	TW	NNMF
[48]	20	4	24 (I)	Preferred speed (5.04 ± 0.72 km/h)	TW	NNMF
[12]	2	5	16 (I)	Preferred speed (4.60 ± 0.9 km/h)	TW	NNMF
[67]	17	4	8 (I)	Preferred speed (4.32 ± 0.68)	TW	NNMF
[13]	12	5	10 (I)	Preferred speed (not reported)	TW	NNMF

mostly of TA activity with additional lighter recruitment of RF and Vastus Lateralis was associated with momentum transfer (by dorsiflexion of the ankle joint) and peaked earliest among the activations at around 50% of the movement time. The peak of this activation also marked the beginning of the activation of the second synergy (encompassing thigh and buttock muscles (RF, Vastus Lateralis, GM, and Semitendinosus), which produced a constant steady activation up until shortly before the end of the movement. This component was correlated with knee extension and hip flexion responsible for the upward movement. Finally, a third synergy, accounting for postural stability, began on the second half of the movement and slowly increased up until the very end of the standing task. This module was mostly composed of Soleus and Gluteus Maximus activity, hinting at a production of knee flexion and ankle plantarflexion acting to decelerate the trunk and stabilize the body.

The timing for the synergy responsible for momentum transfer during forward propulsion occurred noticeably later in the fast condition, while the other two remained relatively synchronous (Fig. 2.5), indicating that in a fast standing scenario, the upward forward movement happened simultaneously with momentum transfer. No differences were found between the older and younger group in both synergy structure and kinematic variables. Similar results were found by other authors [118], recording sEMG activity in six muscles. Another study successfully simulated the forward dynamics of a sit-to-stand movement using a neuro-musculoskeletal model [95], albeit using four synergies because the authors recorded

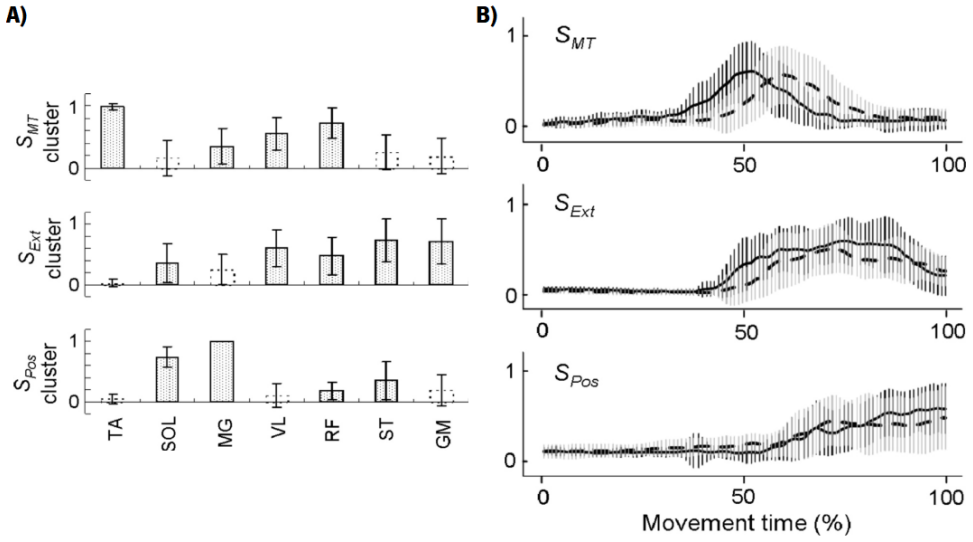


Figure 2.5: Synergy (A) weights and (B) activations during sit-to-stand. In (A), each muscle indicates the average value of activation across subjects with the error bars denoting the standard deviation. Muscles whose average contributions were lower than its standard deviations are represented by empty bars with dashed borders. (B) Solid lines represent average activation across subjects for the comfortable speed condition; dashed lines indicate the "rise as fast as possible condition". Surrounding shaded areas are the standard deviations at each time point. Taken from [14]

Rectus Abdominis activity in order to simulate the flexion of the upper trunk prior to ankle dorsiflexion. The remaining three synergies were consistent with the aforementioned results.

Synergistic relationships for the stand-to-sit movement are hardly studied, with the sparse literature focusing on squatting (weighted or not) rather than natural sitting. In [52], the authors used muscle synergies to understand the task of squatting with an Olympic bar. They concluded that three synergies explained the variance in the movement, as illustrated in Fig. 2.6. The first was marked by Vastus Medialis, RF and TA and dominated during the descending phase, reaching a maximum at approximately the movement's lowest point. The second synergy was composed mostly of Vastus Medialis and Lateralis activity, peaking at the beginning of ascension. A third synergy was attributed to stability and was characterized by plantarflexors (Gastrocnemius Lateralis and Medialis, Soleus) and Gluteus muscles with roughly the same activation throughout the movement.

It is important to note that only half of the squatting movement is similar to a sitting task. This portion of the movement is high in Vastus, RF and TA activity with an activation peak right before sitting, as is illustrated in Fig. 2.6 (Synergies S1 and S2). This may differ from natural sitting if it is assumed that during this movement, muscle activity might abruptly stop long before the corresponding kinematic phase in squatting, as an individual lets their weight descend freely onto the seat.

Table 2.2: Reported numbers of synergies for the sit-to-stand movement in healthy subjects. *B - Bilateral Placement of electrodes; I - Ipsilateral placement*

Study	No. Subjects	No. Synergies	No. Measured Muscles	Seat Height	Factorization Algorithm
[14]	7	3	7 (I)	Equal to Knee Height	NNMF
[118]	3	3	6 (I)	Half of shank height & equal to shank height	NNMF
[97]	11	4	10 (I)	Shank height	NNMF
[119]	12	4	15 (I)	Shank height	NNMF

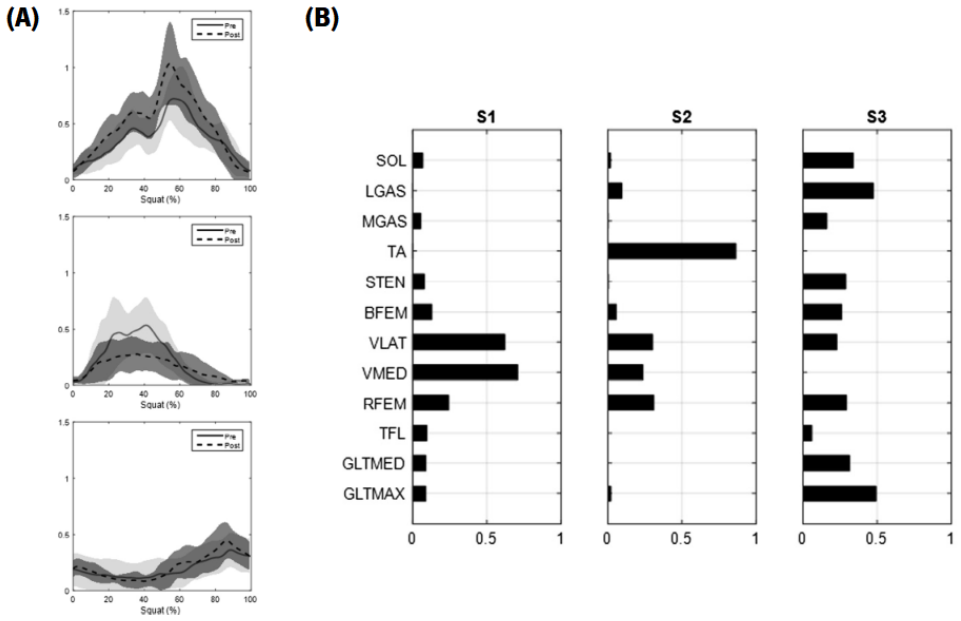


Figure 2.6: Synergy (A) activations and (B) weights during squatting with a weighted bar. In (A), each muscle indicates the value of activation of a concatenated NNMF approach across subjects. S1-S3 correspond to weights of synergy 1 to 3. (B) Activations for Synergies 1, 2 and 3 in a top-down structure. Solid lines represent activations across subjects for the non-fatigued condition (with light gray shaded areas indicating the corresponding standard deviation); dashed lines indicate the activations after fatigue (with dark gray shades denoting standard deviation). Surrounding shaded areas are the standard deviations at each time point. Adapted from [52]

2.3 Critical Analysis

This section aims to critically analyze the collected literature results gathered for the previous sub-chapters. Firstly, critiques of the validity of the motor modularity theory and doubts about its inference from sEMG data will be discussed. Secondly, current methodological choices used for extracting synergies will be examined. Finally, contemporaneous gaps in synergy extraction knowledge will be examined, with

particular attention given to the sparsely studied movements. The chapter ends by summing up the modern synergy extraction paradigm's flaws and how this dissertation plans to address each of them.

Conflicting evidence exists regarding the existence of motor modules as conventionally postulated. Some authors argue that muscle synergies are not encoded in the CNS but rather activated because of task constraints. The recent trend, however, with advances in EMG recording technology and functional resonance imaging data analysis seems to be that synergies are spatiotemporally organized as components in the brain and spinal cord, supporting their neural origin [6]. A recent study, leveraging advances in neural imaging and transcranial stimulation technology, has established a causal relationship between activation of the cerebral cortex region responsible for motor planning and execution (as detected by Functional Magnetic Resonance Imaging (fMRI)) and synergistic voluntary contraction of the Gluteus and pelvic floor muscle as inferred by sEMG recordings [120]. Furthermore, experiments stimulating the spinal cord in animals, including monkeys, frogs, dogs, rabbits and cats result in highly coordinated functional synergies in their musculature [16]. More recent, better designed stimulation studies also support the neural origin hypothesis; for instance, in frogs where the brain was severed from the spine, the simultaneous electrical stimulation of two sites in the spinal cord resulted in endpoint forces being equivalent to the vector summation of each site after individual stimulation [121]. Its controversial origin, however, may be ultimately unimportant for the goals of task-oriented control, such as computational simulation of specific movements, therapeutical insights and functional electrical stimulation pattern generation. This is best evidenced by the fact that neuro-biomechanical models were successful in using synergy based control to simulate walking motion [10–12]. Because this dissertation aims to provide results for future control modeling, muscle synergy theory served our purposes adequately.

The most obvious case open for scrutiny in the conventional synergy extraction pipeline is the usage of sEMG data to ascertain neural control strategies. sEMG data are prone to unreliability due to crosstalk and spatial variability of electrode location both across subjects and trials. Furthermore, previous studies reported a high inter-individual variability of EMG patterns in the TA, during gait [81, 122] and concluded that the control of this muscle is variable across subjects. Nevertheless, despite these drawbacks and barring an invasive method such as intramuscular EMG, it is the best, most readily-available method to study CNS control modularity in motor tasks. Finally, the wide pool of published research using this technique facilitates inter-synergy comparison.

Despite the many reviews and research papers on muscle synergies, the field is limited by methodological inconsistency. Every decision, from choice of pre-processing parameters, of factorization algorithm and of its hyperparameters, of structuring method for data fed to the factorization algorithm, of cross-validating

technique or of evaluating metric for reconstruction quality is often made haphazardly, rarely providing a background in the decision process. Recently, there have been some systematic reviews [6] and studies comparing implications of these choices on final synergy determination [48] that have attempted to unite competing methodological schools on to the same, repeatable, reproducible framework.

Evidence points to NNMF being not only the most reported factorization algorithm to extract synergies, but also the one in which the temporal activation's cumulative distribution function best resembled the input sEMG data; hinting at a more accurate portrayal of neural information present in the raw data [42]. Complementing this, its non-negative nature eases interpretability - a muscle is either positively recruited, implying contraction, or not, and all linearly combine to form the original input waveforms. This justified its selection for this dissertation. However, NNMF is not without its shortfalls. Most notably, being an optimization-based data decomposition technique, it may converge in local minima [123]. Despite this drawback, even local minima can produce results revealing the general "trend of modularity" for a sEMG dataset [124]. This is specially true if it is coupled with a cross-validating method, where several runs' outputs are examined so that the resulting synergies can be investigated for similarity and ranked by structural consistency, the local minima issue can be entirely obviated.

The choice of cross-validating system is particularly divergent in muscle synergy analysis. There exists controversy even regarding what cross-validation should entail in synergy extraction, with some authors making a repeated run analysis with random trial assignment before each run and comparing reconstruction-quality measures [46], and others simply extracting synergies independently and using factors from one synergy output to reconstruct data in the other synergy set [13, 107]. The two techniques have validating ability, and previous studies have leveraged the advantages of both [41, 51, 68], which is what this dissertation will attempt to do as well. However, the existing research is still unclear regarding the parameters this cross-validation system should be configured to, such as data split ratios and reconstruction quality calculation methods; therefore, this dissertation will study these parameters.

The NNMF algorithm in itself may be tweaked, most notoriously by using different iterative update rules, but also by tailoring the NNMF's update rules to model different EMG noise distributions and by configuring the number of maximum iterations, error change tolerance and other convergence criteria. All of these parameters are adequately studied or justified in the body of literature owing to its ubiquity of usage [124], and there are fewer downsides to using the perceived consensus present on published research.

The same can be said for pre-processing decisions such as linear envelope extraction from raw sEMG data and normalization methods. All have been extensively studied and reviews have explored their final

impact specifically on muscle synergies, so the dissertation will lean into these empirical findings [36, 48].

The single upside of the sparsity in methodologies used for muscle synergy extraction is that there is less opposition in trying new techniques and experimenting with different validating models. For example, few studies use statistics to truly leverage multi-participant tests' results to muscle synergy analysis [37, 41, 48, 49]. Analysis of Variance (ANOVA) is particularly useful in drawing parallels between synergies, like comparing modularity between different speed conditions or between healthy and subjects with motor pathologies. If such a statistical framework could be devised, a case could be made for or against muscle synergies using a new perspective employing all variability present in inter-participant data while simultaneously controlling for the population size. Furthermore, assuming similar pre-processing and factorization conditions, multiple studies' results could be aggregated onto the same validating framework in systematic reviews, increasing the statistical power of the inferences. This dissertation will attempt to outline one such framework.

Finally, the lack of human muscle synergy analysis into different movements presents the largest gap in the literature. Published research seems to congregate onto the same few set of movements, most notably: walking and running [13, 41, 51, 85, 107], postural stabilization [24, 37, 67, 92] and grasping [121, 125]. Though these are important movements, additional research is needed into more human daily activities, specially if the goal is to understand neural control to improve solutions given to motor impaired patients. For example, research into standing modularity is repeatable, easily measurable and part of the set of daily activities, but still is very poorly studied in the synergy context. Even simpler sets of movements, that could serve as stepping stones to understanding more complex standing, stabilizing and locomotion synergies, also hardly investigated. This dissertation will also attempt to fill this gap between higher-level complex movements and lower-level simpler motions.

Chapter 3

Synergy Extraction Framework & Dataset Creation

This chapter is divided into two sections; the first presents a high level overview of the muscle synergy analysis process implemented in this dissertation. The second details the experimental protocol used for each studied movement, the data collection system and the data treatment leading to the creation of the dataset for lower-limb sEMG and kinematic variables during multiple tasks of daily life to enable further analysis.

3.1 Synergy Extraction Framework

The methodological framework put in place to reliably and repeatedly obtain the synergies encoding modular CNS information is schematized in Fig. 3.1. This diagram attempts to offer the reader a general overview of the entire process, from the collection of data to the final estimation of common motor modules for a given movement. Each step will be expanded upon in the following chapters. A short description of each block follows:

During Data Collection, sEMG and kinematic data are simultaneously recorded during a given movement and wirelessly transmitted to a computer where they are stored. The next step, Data Processing, has a dual purpose: first, it uses the kinematic data to segment sEMG recordings corresponding to actual contraction; secondly, it extracts from the muscle activity data the time and amplitude normalized envelopes. From this point on, kinematic data is no longer needed for the purposes of muscle synergy analysis; therefore it is stored in a dataset for future reference, along with the sEMG envelopes represented by the

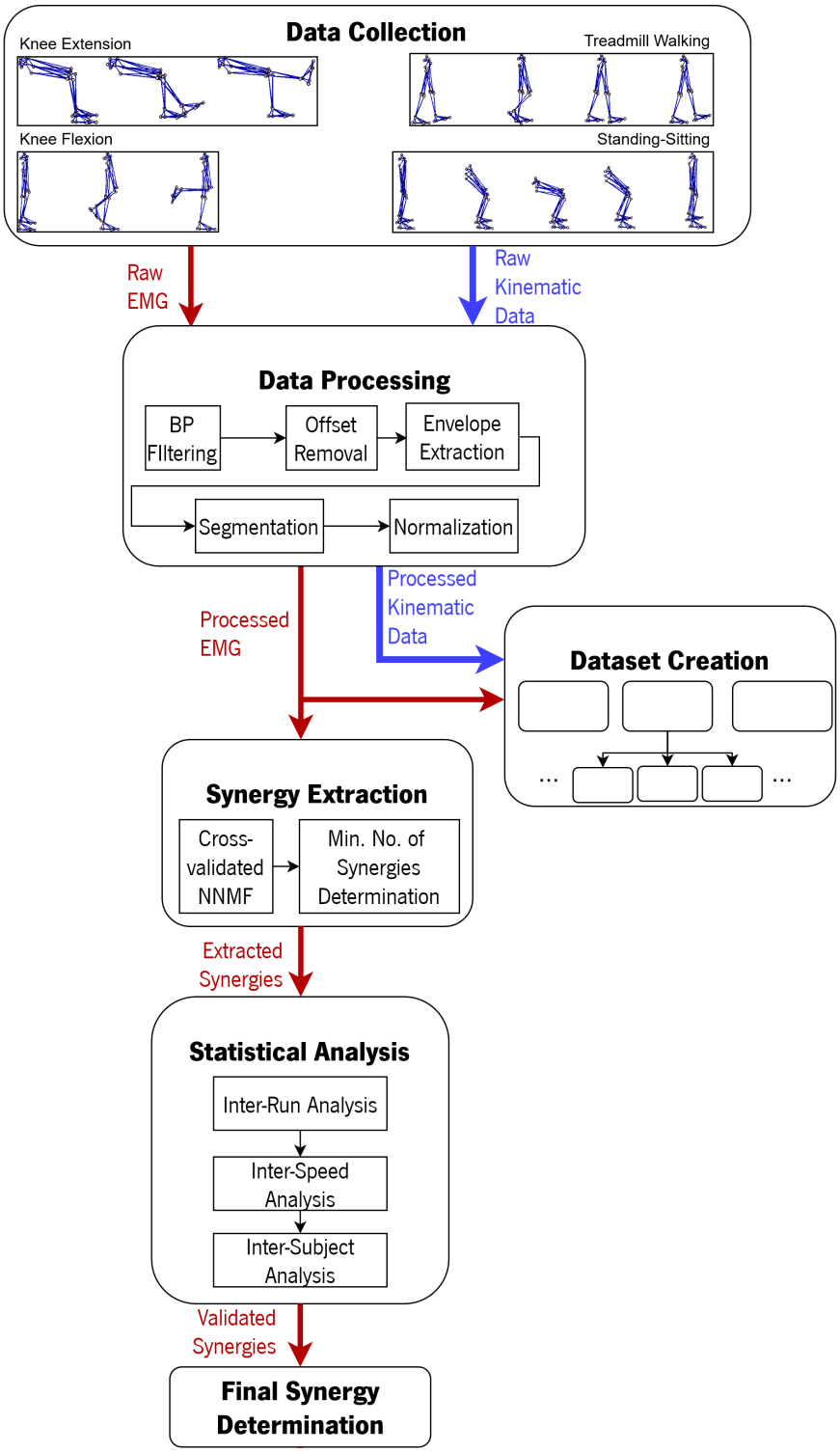


Figure 3.1: Schematic overview of the muscle synergy extraction framework

Dataset Creation block in the diagram. Subsequently, the sEMG envelopes are factorized into muscle synergies in the Synergy Extraction block. Several rounds of extractions are performed, each with random initial conditions in order to minimize the chances of picking a set of synergies from a local minima as evaluated by the cost function. The next block, Statistical Analysis, is tasked with sifting through these

several runs and extracting the most recurrent set across iterations, as well as evaluating the similarity of these motor modules both across movement speeds and subjects. Finally, the information gathered during this statistical process is used to find a common set of synergies through grouping and averaging across subjects.

3.2 Data Collection and Dataset Creation

In order to collect data for synergy analysis in healthy subjects, an experimental protocol was devised. All participants signed a consent form to be part of the study. Subjects' rights were preserved and personal information provided was remained confidential. Data was collected at the University of Minho according to the ethical comission.

3.2.1 Participants

Ten healthy subjects (8 males and 2 females with mean age of 25.2 $\pm$ 2.52 years, mean height of 172.2 $\pm$ 10.6 cm and mean weight of 73.82 $\pm$ 11.09 kg) participated on the experiment trials. None had clinical history or evidence of motor disorders. Fig. 3.2 shows sensor placement on two of the subjects.

3.2.2 Equipment

Each subject was first fitted with an array of sEMG electrodes and Inertial Measurement Units (IMU) sensors (Fig. 3.2). The sEMG electrodes used for data collection were from the Delsys Trigno™ Avanti system (Delsys, Massachussets, USA). After shaving and cleaning the skin on the points of contact with alcohol wipes (according to the recommendations gathered from the literature and presented in section 2.1.2, eight sensors were placed bilaterally (four in each leg) on the *Rectus Femoris* (RF), Biceps Femoris (BF), *Gastrocnemius Medialis* (GM) and *Tibialis Anterior* (TA), following SENIAM recommendations [53]. sEMG data was collected at 2148 Hz.

The four muscles were chosen because their sEMG activity during motor tasks is widely reported in the literature [10, 13, 14, 42, 46, 47, 85, 87, 111, 114, 118, 126, 127] and each may be paired with another in an agonist/antagonist relationship [128], thus maximizing potential modularity findings by minimizing the muscles that are activated simultaneously during a movement.

To measure the joints' angular position, the subjects were instrumented with the Xsens Mtw Awinda™ motion tracking system (Xsens Technologies B.V., Enschede, The Netherlands). This system consists of an array of IMU sensors, each fitted with a tri-axial accelerometer, gyroscope and magnetometer which,

in conjunction with a host device, handle the sampling, buffering, calibration and wireless transmission protocol. These features, combined with a biomechanical model tailored to a subjects anthropometric measurements allows for reliable collection of 3D orientation data. The IMU sensors were placed on the feet, shanks, thighs and lower back of each subject according to the systems' guidelines, allowing the tracking of the ankle, knee and hip joint angles. The sampling frequency was 100 Hz.

(A) (B)

Figure 3.2: (A) sEMG and IMU sensor placement on one of the participants. In order to secure the IMU sensors, straps had to be fixed to the subjects' limbs. Approximate locations of the sEMG sensors are highlighted by dashed green rectangles, while orange represents the IMU sensors' positions. (B) shows the sensor locations of only the sEMG acquisition system, unobstructed by the straps.

3.2.3 Experimental Protocol

Before any movement was performed, a series of tests was done in order to later determine each subjects' Maximum Voluntary Contraction (MVC) for each of the recorded muscles, according to SENIAM recommendations [53]. To this effect, subjects were asked to perform their strongest isometric contraction against a researchers' opposing resistance three times for three seconds, with a five second interval between contractions to prevent muscle fatigue. This was repeated for each muscle with different configurations of subjects' positions depending on the muscle being recorded.

Subsequently, participants started performing the studied motor tasks, by order of execution: Knee Extension, Knee Flexion, Sit-to-Stand, Stand-to-Sit and walking. The performing speed of each motor task was controlled through the use of a metronome, plus the researcher's own verbal guidance.

For Knee Extension, subjects were seated on the edge of a chair and the seat height was adjusted so that the subjects' thighs were parallel to the floor when the feet were flat and resting (Fig. 3.3 (A)). Afterwards, subjects were asked to extend their knee while taking five, three and one seconds to reach full extension in order to simulate a low, medium and high movement speed, respectively. After holding this fully extended position for three seconds, they were allowed to relax. Each speed condition trial was repeated three times and all trials were performed on both the right and left legs, resulting in a total number of trials of 18 (3 trials $\times$ 3 speed conditions $\times$ 2 leg sides).

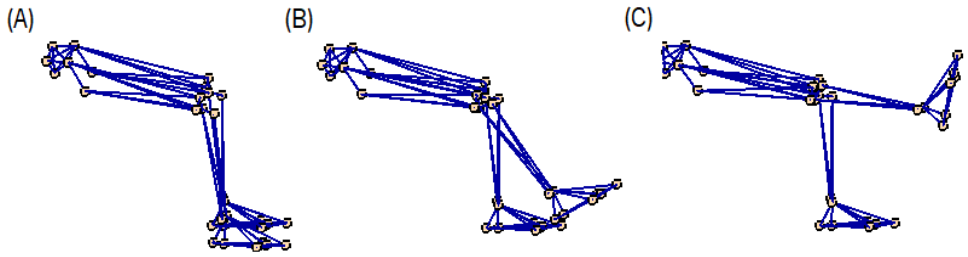


Figure 3.3: Snapshots of a Knee Extension movement as portrayed by a right-side view of a biomechanical model of the lower limbs of an arbitrary subject. (A) starting position; (B) mid-movement; (C) final position

For Knee Flexion, 18 trials were also recorded with the same timings. The subjects started the movement in a standing position (Fig. 3.4 (A)) and were asked to raise one foot off the ground until the shank was parallel with the floor. To maintain balance, subjects were instructed to clasp their hands together and stretch them out. After holding a fully flexed position for three seconds, they were allowed to relax.

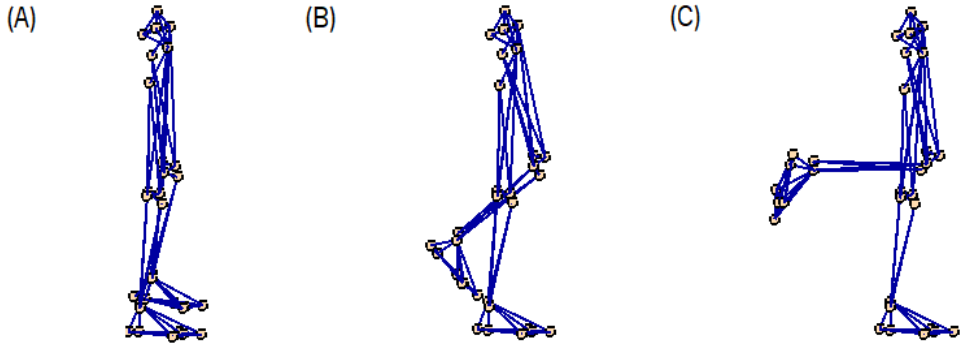


Figure 3.4: Snapshots of a Knee Flexion movement as portrayed by a right-side view of a biomechanical model of the lower limbs of an arbitrary subject. (A) starting position; (B) mid-movement; (C) final position

The sit-to-stand and stand-to-sit movements were included in the same recording session. The starting position consisted of the subject sitting on the edge of a height-adjusted seat with both feet planted on the ground (Fig. 3.5 (A)). On cue, they were asked to stand up in five, three and one seconds for the high, medium and low velocity conditions, respectively. They then held this upright position for five seconds

(Fig. 3.5 (C)). After the five seconds, they sat (Fig. 3.5 (F)), using the same pace they took to rise. This was repeated three times for each speed condition, yielding a total of 9 experimental trials of combined sit-to-stand and stand-to-sit movements for each subject.

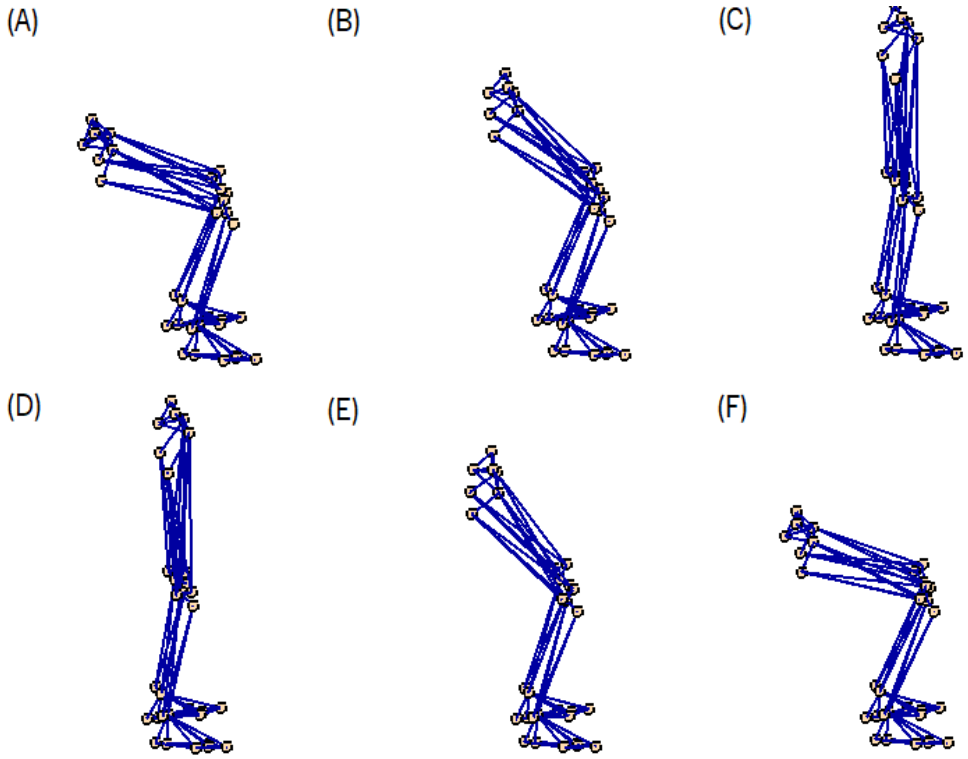


Figure 3.5: Snapshots of a Sit-to-Stand (A-C) and Stand-to-Sit (D-F) movement as portrayed by a right-side view of a biomechanical model of the lower limbs of an arbitrary subject. (A) Sit-to-Stand starting position; (B) Sit-to-Stand mid-movement; (C) and (D) Sit-to-Stand final position and Stand-to-Sit starting position; (E) Stand-to-Sit mid-movement; (F) Stand-to-Sit final position

For acquiring data during walking, subjects were asked to walk into a treadmill operating at 1 km/h, 1.5 km/h and 2 km/h for the fast, medium speed and slow conditions, respectively, and walk as naturally as possible for 1 minute (Fig. 3.6 illustrates a full gait cycle). This was repeated three times for each speed condition, resulting in 9 total gait trials per subject. This range of walking speeds was selected from the low-mid end of ranges reported in the published literature [46, 85]. This was done with future work in control modeling in mind, where data from a finer span of slower velocities is more useful for validating purposes than sparser speeds from a wider range of low to high walking velocities.

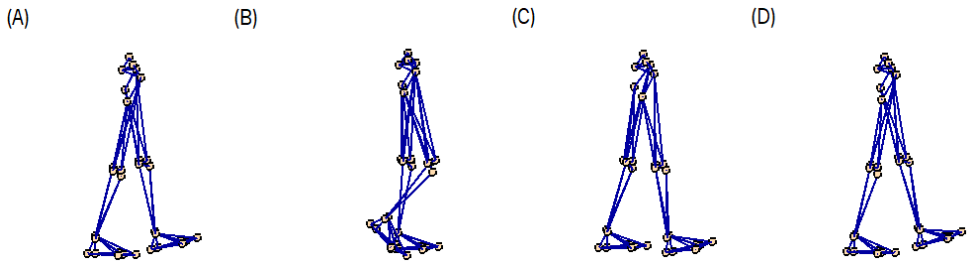


Figure 3.6: Snapshots of the walking movement as portrayed by a left-side view of a biomechanical model of the lower limbs of an arbitrary subject. (A) Initial Contact marking the beginning of stance phase; (B) Mid Stance; (C) Pre-Swing marking the end of Stance phase and beginning of the Swing phase (D) Terminal Swing, marking the end of both the swing phase and the current gait cycle

3.3 Data Pre-processing

After obtaining sensor data, a pre-processing phase was required in order to obtain the sEMG envelopes pertaining to muscle activation during movement. Fig. 3.7 shows the flow of data for this step, divided into the two main subsequent phases: envelope extraction (Fig. 3.7 (A)) and Normalization ((B)).

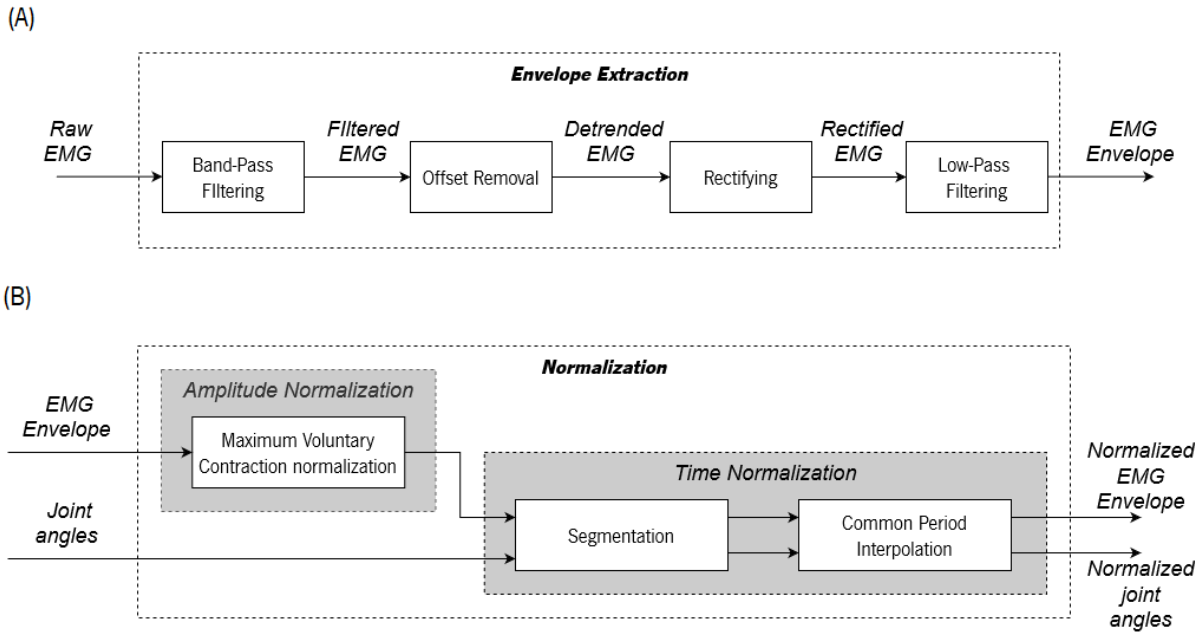


Figure 3.7: Block diagram of the sEMG pre-processing. Data pre-processing may be divided into two sequential phases: (A) Envelope Extraction and (B) Normalization.

Envelope Extraction is responsible for converting high-frequency raw sEMG data into the smooth-curved linear envelope. This is achieved in accordance with the established process in muscle synergy extraction literature: firstly, raw muscle activity data was filtered with a fourth order zero lag BP Butterworth filter with

a lower and upper cut-off frequency of 20 and 400 Hz, respectively [9, 42, 116, 129]. Next, the Direct Current (DC) offset was removed from the filtered data by subtracting its mean [10, 11, 27, 36, 46, 67, 90, 107, 130]. This step may be redundant as the DC offset should have been already removed due to the HP filtering, but it was implemented to be consistent with the literature where it probably plays a role in compensating for incomplete attenuation of the digital filter at 0 Hz. Subsequently, full wave rectification was performed by taking the absolute value of the resulting signal [37]. Finally, the final sEMG envelope was extracted by way of filtering with a zero-lag third order low-pass Butterworth filter, set for a cut-off frequency of 2 Hz. This value was selected based on the conclusions of multiple reviews [48, 65, 79].

The second major phase of pre-processing is the Normalization (Fig. 3.7 (B)), wherein each envelope is normalized in amplitude with respect to the MVC, and in time to a common length of 101 samples. To extract one single MVC for the recordings of multiple contractions, MVC data was first subject to the same processing as shown on Fig. 3.7 (B). The peak of the resulting signal (subject to manual analysis to check for the presence of outlying bursts of amplitude) was defined as full activation of that muscle, and trial data was normalized with respect to this value.

To accomplish time-normalization, sEMG data needs to be segmented. Because sEMG and IMU sensor acquisition were started simultaneously, joint angle data was used to identify the sample range corresponding to knee flexion and extension, and Sit-to-Stand and Stand-to-Sit movements. This process is exemplified in Figure 3.8. The shaded areas, corresponding to periods were knee angle variation was observed, were extracted through a semi-automatic process, were the derivative of the knee angle was calculated, and marked as encoding movement whenever it deviated for more than a few samples from 0. The collected intervals were then subject to manual inspection and *ad-hoc* resegmentation if the periods were found to be improperly segmented.

For the treadmill walking trials, the segmentation was made regarding the gait cycle using the acquisition systems' own markers for heel strike events, and kinematic data spanning from this point up until the next heel strike was classified as belonging to one full gait cycle. The resulting segments may be observed in Fig. 3.9. The first and last three measured gait cycles were excluded from further analysis to allow for natural gait patterns, as the subjects were asked to step into and out of a working treadmill and therefore required time to adjust.

Lastly, the normalized sEMG and kinematic data were adjusted to a common time-basis of 101 points, enabling direct data comparison. Final sEMG envelopes averaged across trials for every movement and speed condition of an arbitrary subject are illustrated in Fig. 3.11. Both resulting signals were then stored in a structured dataset, as explained in section 3.4.

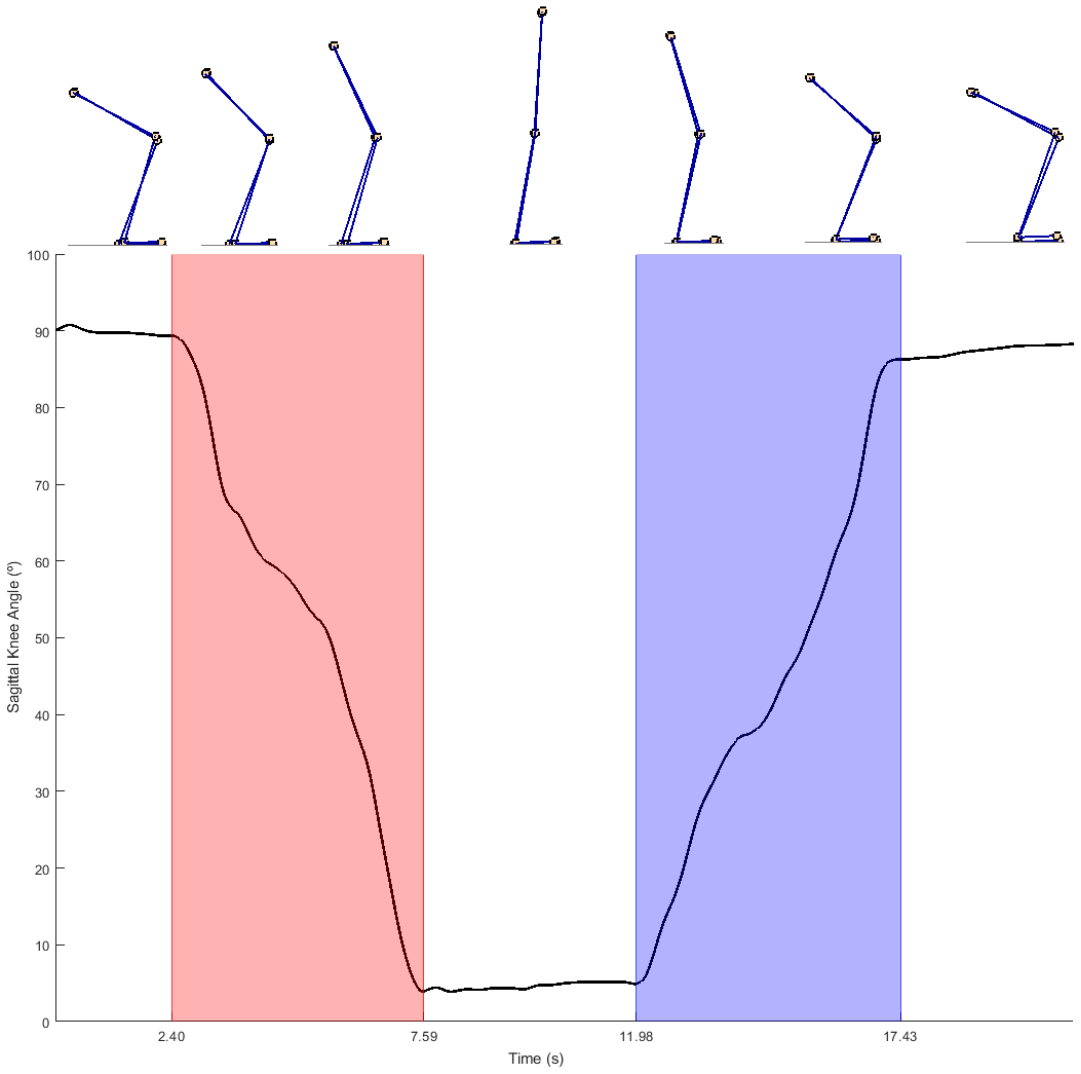


Figure 3.8: Segments classified as sit-to-stand (red shaded area) and stand-to-sit (blue) movement, according to sagittal knee angle (black line) of of an arbitrary trial. Frames of a biomechanical model captured at different timestamps of the same trial are shown for reference. Timestamps on the x-axis were stored for later reference when segmenting sEMG data

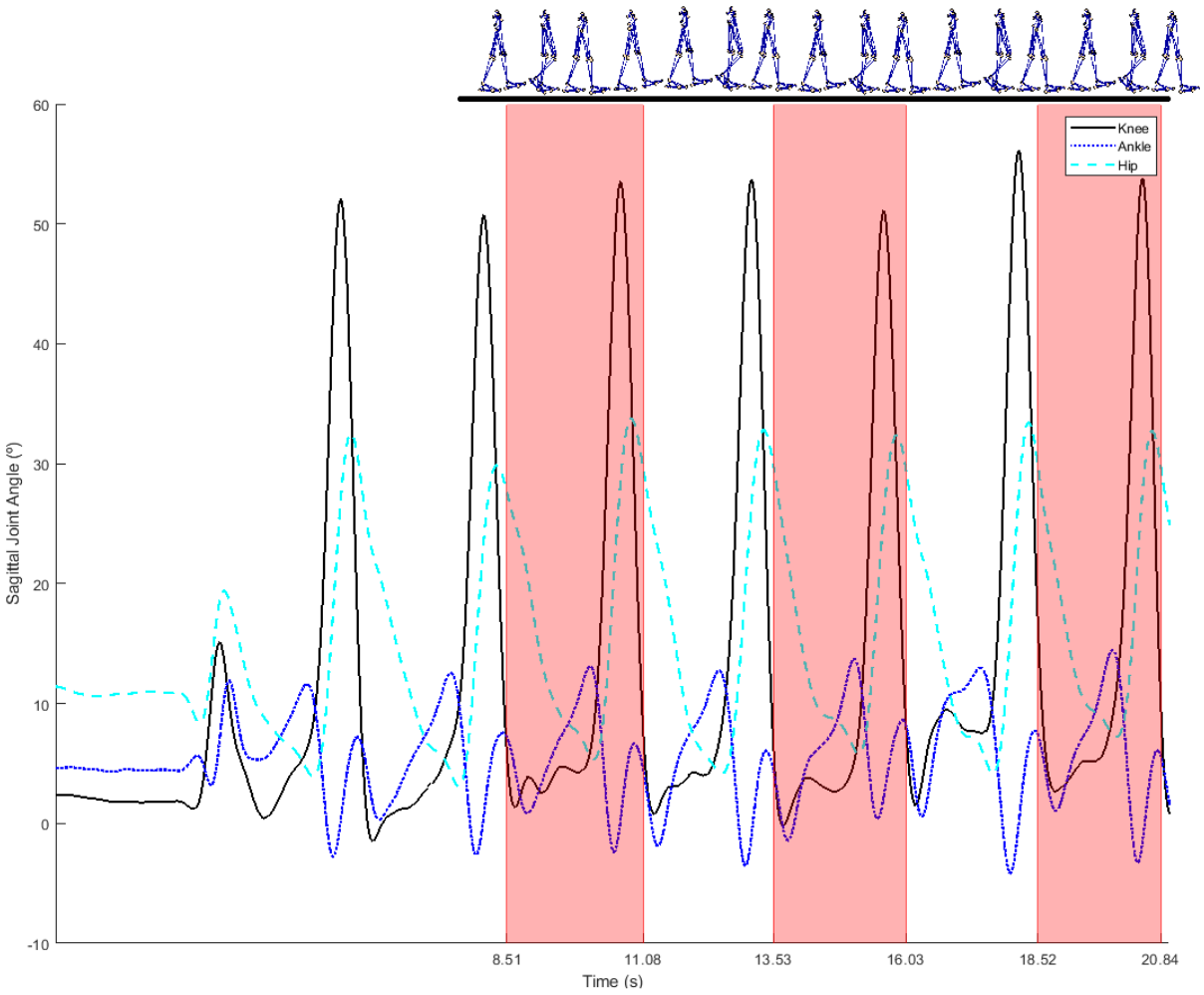


Figure 3.9: Gait cycle segments highlighted as alternating red and white shaded areas of the sagittal angle of the knee (black line), ankle (dark blue line) and hip (dashed light blue line). Note how the first three gait cycles were not segmented to allow the subject to adapt to treadmill walking. Timestamps on the x-axis were stored for later reference when segmenting sEMG data.

3.4 Dataset creation

Raw and processed data was stored in a dataset following the layout illustrated on Fig. 3.10. The dataset is divided into two major subdivisions: "Raw Data", which includes the raw sensor data in the respective acquisition systems' proprietary data format, and "Processed Data" which consists of data in ASCII format stored as **.txt* files. The "Processed Data" field may be further subdivided into three subfolders: "EMG_raw", storing the raw data in ASCII, "EMG_envelope", encompassing the normalized sEMG envelopes extracted using the methods described in section 3.3 (a visualization of these envelopes is present in Fig. 3.11); and "Joint Angles", consisting of the angles of the hip, knee, ankle, and ball of foot joint on the sagittal plane.

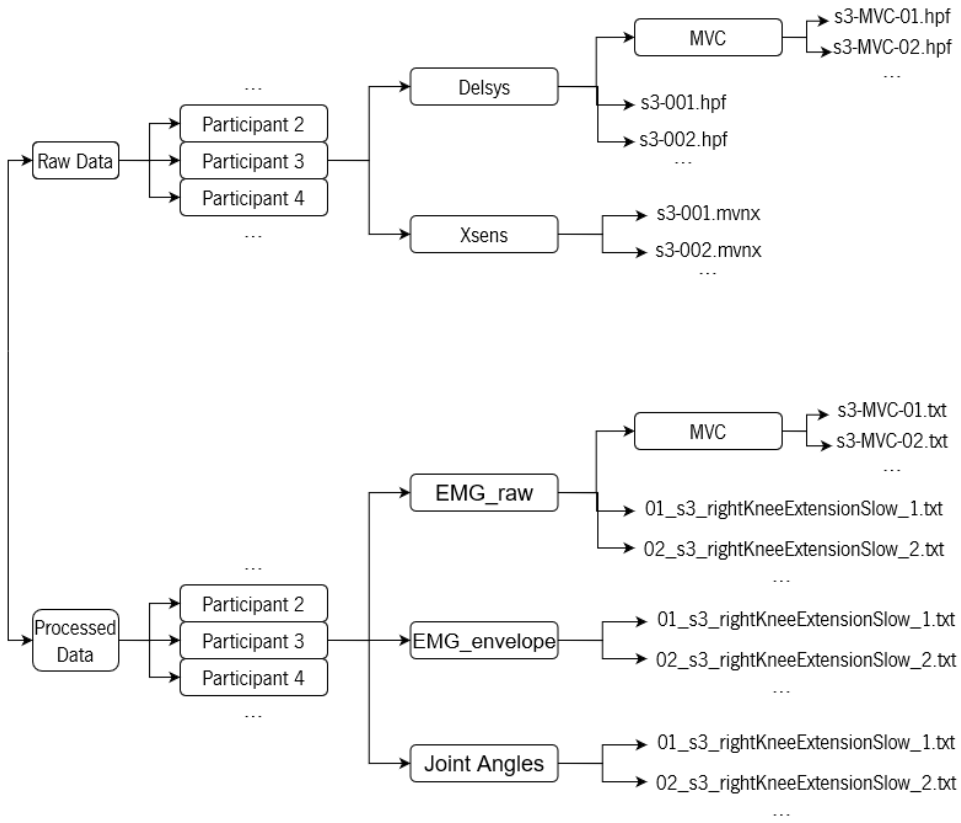


Figure 3.10: Dataset layout

Data was organized on a participant-by-participant basis by assigning each of the ten subjects a corresponding *ParticipantID* folder with all data collected for that specific subject. Each participant had data for each trial saved onto a separate file. For the processed data subdivision, these files were named following a labeling convention intended to make future processing work easier; for example, *01_s3_rightKneeExtensionSlow_1.txt* contains data for the first experiment overall (*01*), for the par-

ticipant with ID 3 (*s3*) during the extension of the right knee in the slow speed condition, (*rightKneeExtensionSlow*) during the first (*_1*) out of a total of three trials where right knee extension was performed.

On the Processed Data branch of the dataset, both the raw sEMG data and envelopes are organized into 9 columns, the first of which encoding the timestamps (in seconds) and the other 8 consisting of sEMG activity (in Volts). Each file is preceded by an 8 line header, describing which muscle was assigned to which data column, as well as other information, namely sampling frequency, measuring unit and total number of data points. The sensor index was consistent across all subjects, with columns 2-5 storing sEMG for right RF, BF, TA and GM respectively and columns 6-9 the analogous muscles on the left leg. This order was also preserved for the MVC trials: files 1-8 inside the MVC subfolders corresponded to columns 2-9 of the actual trial data files.

Joint angle data was organized with the same one-trial-per-file structure, same file extension and same naming convention as the sEMG trials. Data was organized into 9 columns, the first of which encoding the timestamps (in seconds) and the remaining consisting of the sagittal angle of a set of 8 joints. Each file is preceded by an 8 line header, describing which joint was assigned to which data column, as well as other information: sampling frequency, data units and total number of data points. The joint index was consistent across all subjects, with columns 2-5 storing angle data for the right Hip, Knee, Ankle and Ball of Foot (joint between the foot and toe) and columns 6-9 the analogous joints of the left leg.

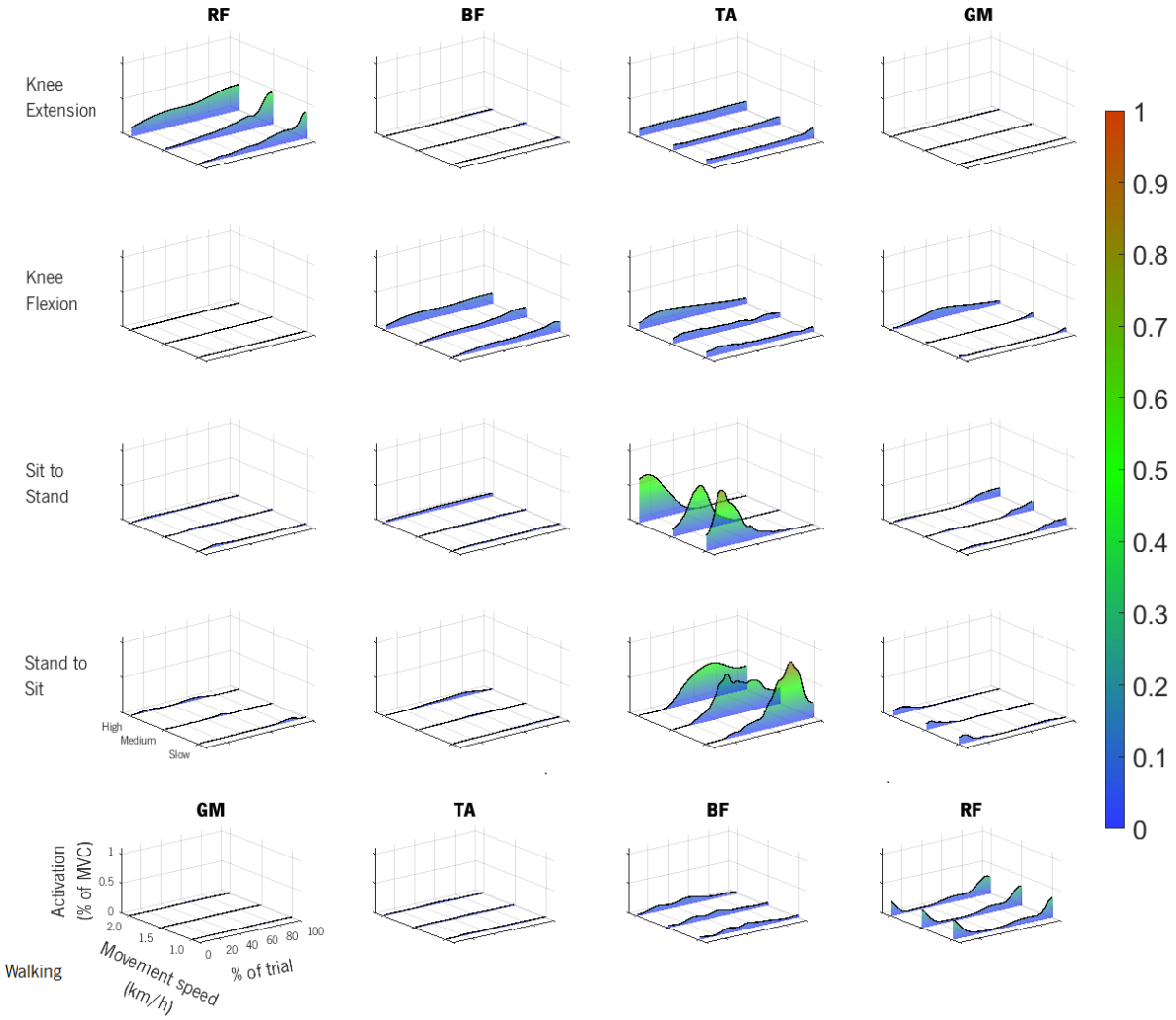


Figure 3.11: Muscle activities expressed as sEMG envelopes in a representative subject (participant 2) for the (A) Discrete Motor Tasks and (B) Walking trials. Each sEMG envelope was normalized by the participant's MVC.

Chapter 4

Multi-task Muscle Synergy Extraction

In this chapter, the methodological framework implemented in this dissertation to allow robust muscle synergy analysis will be explained. To accomplish this logically while justifying the methodological decisions, this chapter will be further subdivided intro three sections. In the first, "Introduction", a brief outline of the selected NNMF and cross-validating framework will be demonstrated, pointing out the areas where there's either a knowledge gap in published literature or a lack of consensus. The "Methods" section delineates the preliminary study taken to address issues raised in the previous section, as well as the criteria used to determine the number of synergies to be extracted for each movement. Finally, the chapter ends with "Results and Discussion" where the final framework is presented with input from the ascertained results, as well as the selected number of synergies to extract from each movement.

4.1 Introduction

Most factorization algorithms used for muscle synergy extraction work by minimizing the mean square error (Mean Square Error (MSE)) between the input matrix and a parts-based approximation reached by the technique [42, 94]. Given a matrix V, each column containing n observations of m, variables, decomposition techniques construct approximate factorizations, as described in Eq. 4.1.

$$V \approx WH = \sum_{i=1}^{k} W_i H_i \tag{4.1}$$

The k columns of W can be regarded as containing a basis that is optimized for the linear approximation of the data in V (that is, W contains some component that crops up again and again in the original n data points). Each column of H is called an encoding and is in one-to-one correspondence with

a data point in W. An enconding consists of the coefficients by which a matrix is represented with a linear combination of basis vectors. The dimensions of the matrix factors W and H are $n \times k$ and $k \times m$, respectively. The rank k of the factorization is generally chosen so that $(n+m)k < nm$ and the product WH can be regarded as a compressed form of the data in V.

In the synergies context, weighting coefficients (W) symbolize the relative importance of a muscle for a given task and activation profiles (H) specify their temporal relevance during the movement [94, 124]. k can be thought of as the number of muscle synergies. Thus, Eq. 4.1 may be rewritten in the problem space of synergy extraction as expressed in Eq. 4.2.

$$M = \sum_{i=1}^{N} W_i H_i + \varepsilon \tag{4.2}$$

Where M is the matrix $N \times t$ of N muscle activation patterns recorded during t time points, W_i is the $N \times k$ matrix of the weighting coefficients, H_i is the $k \times t$ matrix of time varying activation profiles for the i^{th} synergy and ε is the $N \times t$ matrix representing the reconstruction error of each muscle's activation pattern at each time point.

In practical terms, each component W_i represents the contribution of one muscle to synergy i with the implied notion that any one muscle may contribute to multiple synergies by being present in several W's. These so-called *weights* are multiplied by the scalar recruitment coefficient denominated *activations* (H), encoding the time-varying expression of W at each time point.

This, is usually assessed through Variance Accounted For (VAF) [27, 38, 47, 51, 51, 57, 62, 67, 85, 107, 131], whereby muscle contraction patterns are reconstructed by linearly combining synergy weights and activations. Mathematically, this can be expressed as

$$VAF = (1 - \frac{\sum_{j=1}^{t}\sum_{i=1}^{N}(W_i H_i - M)^2}{\sum_{j=1}^{t}\sum_{i=1}^{N}(M)^2}) \times 100\% \tag{4.3}$$

Where a value of VAF = 100% means the original data and the reconstructed approximation are perfectly matched and 0% signifies that the two signals are completely different.

A few studies claim that because VAF uses a statistical approach to quantify reconstruction accuracy, it does not measure actual information on the magnitude on neural information inferred by the sEMG signal. To address this, investigators have recently attempted to compare the shapes of the probability distribution functions between that of the recorded signals and the factorized activations, which reportedly do include sEMG spectral information [42, 132, 133]. Nevertheless, most studies still use VAF as a quantifiable metric of reconstruction accuracy [6, 42, 98], making it the best candidate for drawing parallels with the

literature.

The factorization technique chosen to linearly decompose the sEMG data into muscle synergies was the NNMF MATLAB routine with multiplicative update rules, as described by Lee and Seung [94]. This decision was based on three reasons: firstly, NNMF's widespread adoption in synergy extraction literature may allow direct parallels to be drawn between the results of this dissertation and the published literature; secondly, probabilistic analysis shows that, among the probability distribution of activations (H) derived from different factorization techniques for a given movement, NNMF's is the most similar to the original sEMG signal, suggesting better capturing of neural information present on the muscle activation recordings and hence potentiating more accurate modularity findings [42]; thirdly, the non-negativity constraints of this algorithm facilitates interpretability by serving the intuitive notion of combining parts to form a whole, yielding weights and activations that may only range from 0 (complete non-activation) to 1 (full recruitment) and which may be linearly added to re-obtain the original sEMG signal.

A basic implementation of a synergy extraction framework using NNMF may be comprised of the following steps, schematized in Fig. 4.1. To find the optimal H and W that best represent the modularity on the sEMG data, NNMF starts with two random initial matrices H and W, plus the sEMG envelopes M. It then attempts to search for better approximations by minimizing the Frobenius norm of the mean squared error between the original (M) and the current WH approximation ($\frac{1}{2}||M - WH||_F^2$). This is done by multiplying, at each iterative step, the activations and weights by a scalar proportional to the quality of the approximation (see equations in blocks "Update W" and "Update H" in Fig. 4.1). The process repeats until convergence criteria is met: either a set number of iterations have been performed or the MSE is considered low enough as to accurately represent the original data. Though not part of the factorization, the VAF between the resulting output of the NNMF and the original sEMG data is usually then calculated. If it is found to be high, this may mean that the algorithm successfully converged on a modular representation of the original data.

The convergence settings for the NNMF procedure used for this dissertation were: 500 maximum number iterations for each run, the same as used in the paper introducing the algorithm [94] and a 10^-9 termination tolerance for change in the size of the residual and equal relative change tolerance in the elements of the reconstructed data. These values for tolerance are more conservative than what is usually reported (e.g. 10^-7 in [112], 10^-4 in [62]). This was done so that the NNMF mostly reached convergence by attaining the maximum number of iterations in order to maximize the chances of "sampling" synergy results from the steepest point of different local minima of the objective function.

An issue arises with the previous implementation: it may be misleading to calculate variance-based

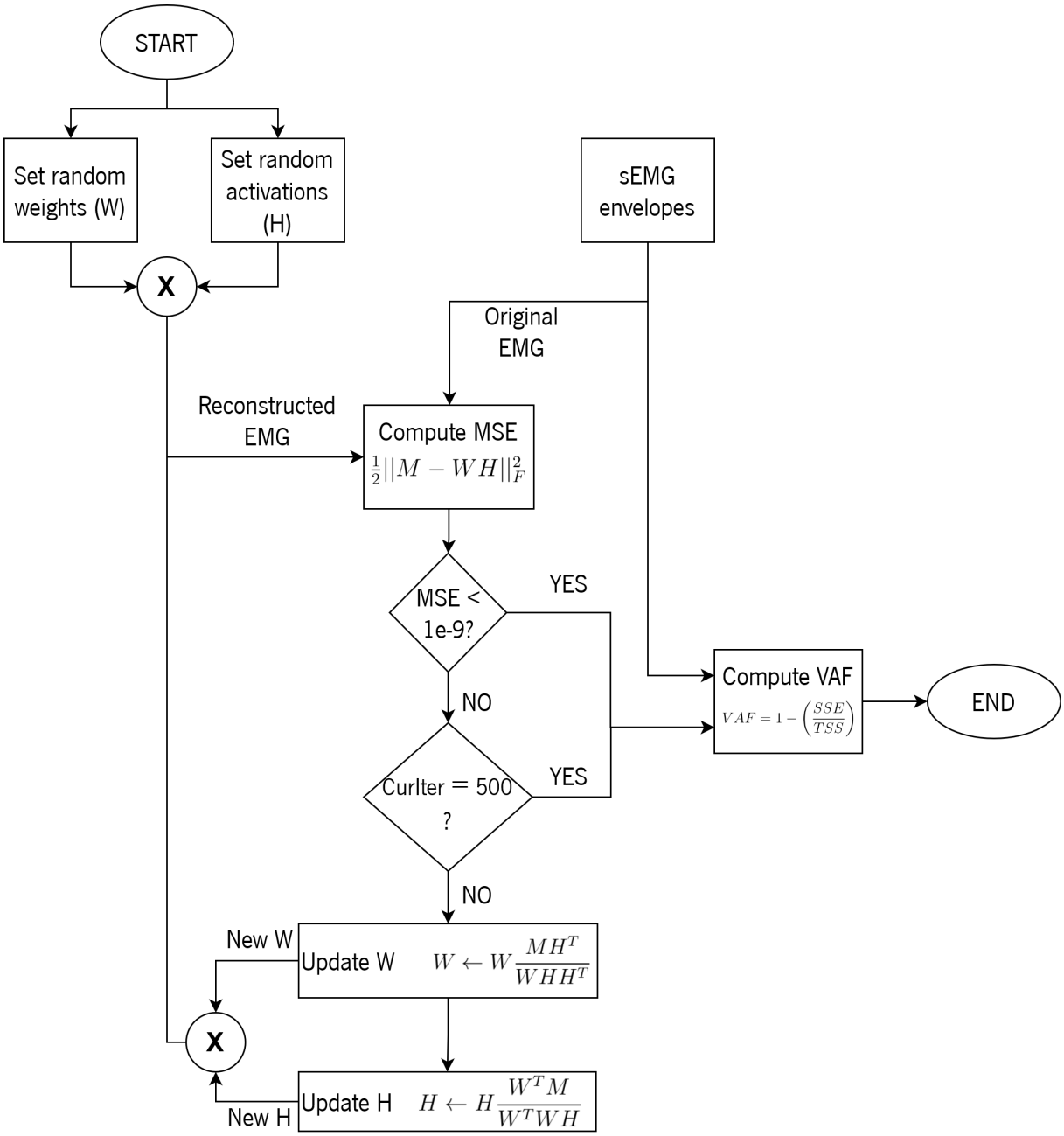


Figure 4.1: Flowchart of a simple NNMF implementation. *MSE*: Mean Squared Error; *SSE*: Sum of Squared Errors; *TSS*: Total Sum of Squares

metrics such as VAF to evaluate the quality of reconstruction on the same data that was used to form the factors, because these results may not be applicable to other slightly different trials.

To address this issue, a cross-validation framework (schematized on Fig. 4.2 (B)) was devised by randomly splitting the sEMG data into two groups (Group 1 and 2) and initially running the NNMF algorithm on data from Group 1. Subsequently, the computed weights were fed to a modified version of the NNMF algorithm where only the activations were allowed to update, with the weights being held fixed. The results were then "crossed" by multiplying the weights derived from the extraction performed on Group 1 by the

activations computed on data from Group 2, when calculating VAF according to original sEMG data in Group 2. The cross-validation process was repeated 10 times, with random partitioning of the data before each of the runs, and different random W and H where applicable. This value of VAF is more accurate in portraying the quality of reconstruction because variability is split between both factors, W and H, acquired on independent datasets and therefore less biased towards trial-specific idiosyncrasies. Furthermore, if conditions in one groups' processing lead to the algorithms' convergence on a local minimum, this will most likely be indicated by a low value of VAF because the chances of the endpoint of NNMF on the parallel dataset being a local minimum with complimenting weights/activations that would lead to an accurate reconstruction are minimal. This approach has been widely reported in synergy extraction literature [6, 68, 91, 92, 107]. This procedure is henceforth referred to as cvNNMF (Cross-Validated Non-Negative Matrix Factorization) to distinguish it from the concept of NNMF as a factorization algorithm indiscriminate in its inputs and outputs, as illustrated in Fig. 4.2 (A).

However, an issue remains relating to how the sEMG data should be split during cross-validation, namely what is the ideal proportion of Group 1 data to Group 2. Some authors use a 60%/40% ratio for data assigned to Group 1 (from which weights are extracted) and Group 2 (with which activations are computed), respectively [41, 51]; others use a 75/25 [68] split, others yet choose 50/50 [107], though none indicates a reason for their choice. This question is represented in Fig. 4.2 (B) as *Q1*.

Furthermore, as this dissertation aims to yield a generic framework, the goal is to include all the variability across all trials. For this purpose, it is needed to verify which of the following options is more accurate: (i) to average, or (ii) to concatenate the trials. This is illustrated by *Q2* in Fig. 4.2 (B). Again, authors differ in their approach, some averaging [41] and others concatenating analogous trials [54, 72] while seldom justifying their choice. Fig. 4.3 illustrates the differences between a concatenating and averaging technique on data from 4 trials of the same muscle.

Finally, should the final VAF for a given cvNNMF run be computed according to original data from Group 1 or Group 2? This corresponds to *Q3* in Fig. 4.2 (B). Because these concerns are rarely addressed in the literature, a short preliminary analysis was performed to evaluate the influence of these parameters on final VAF. The formulation of this study will be examined in section 4.2.

Another step that was made at this point of the analysis was the selection of criteria that defined the Minimum Number of Synergies (N_{syn}) that explained the sEMG data in each of the movements. Previous works have used arbitrary thresholds from 75% to 95% of VAF as criteria [10, 12, 49, 77], and a threshold-based system seems to be the most common among the entirety of research. Other studies have defined the minimum number of synergies by plotting the r^2 as a function of N_{syn}, then fitting decreasing portions

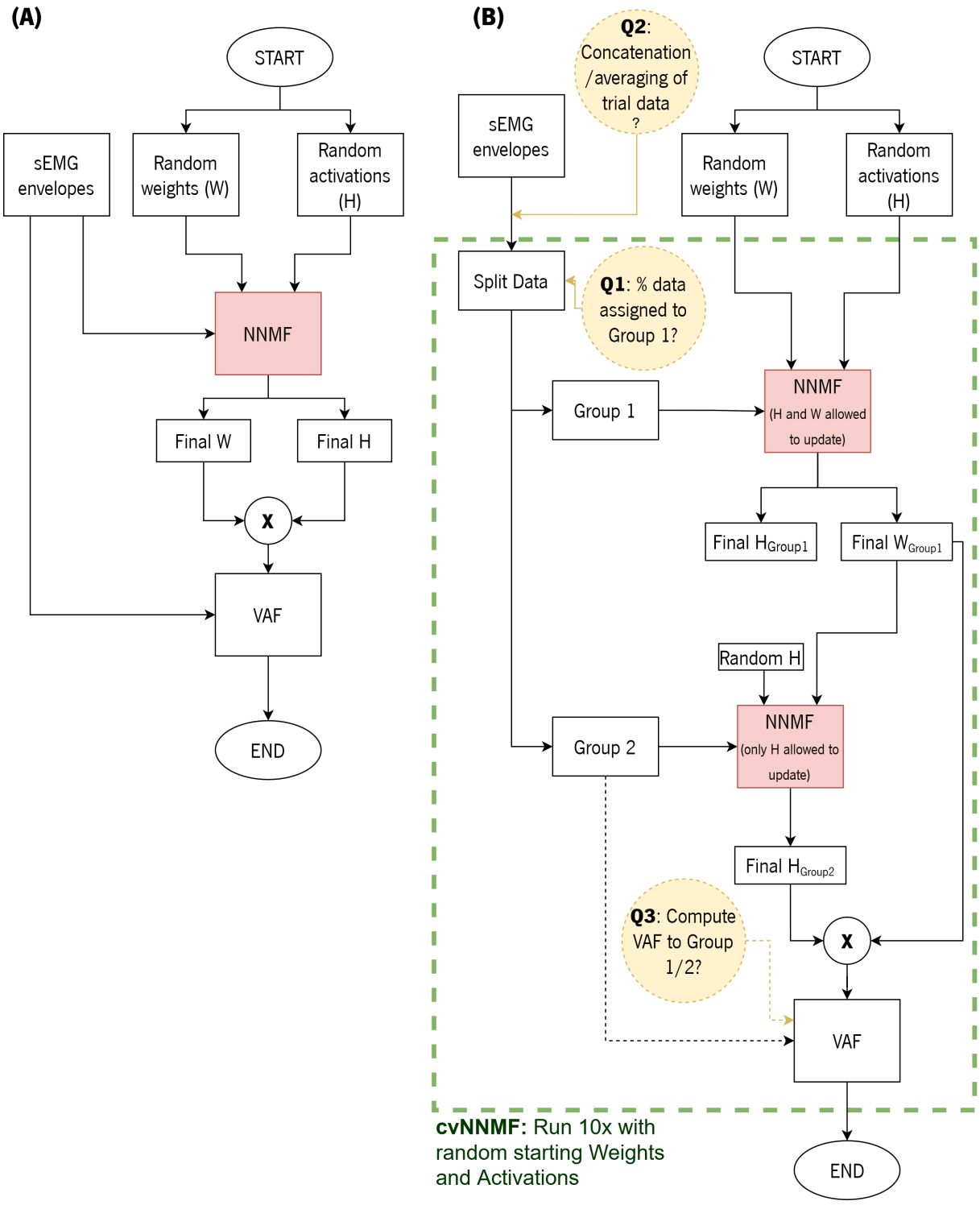


Figure 4.2: Schematic representation of (A) a simple NNMF implementation with no cross-validation and (B) a cross-validated NNMF procedure (cvNNMF) used in this dissertation. Yellow circles highlight points of the process where the relevant literature is unclear on the best configuration.

of the actual r^2 curve to straight lines and choosing N_{syn} as the first point in the curve for which the portion approximates a straight line, as measured by mean squared error to the linear fit [13, 68]. Additional

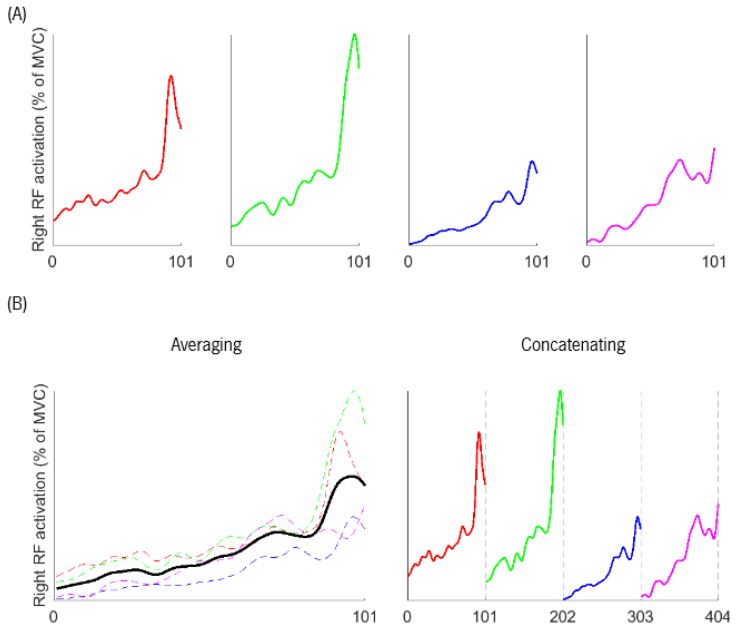


Figure 4.3: Illustration of differences between averaging and concatenating trial data. (A) shows four trials from one muscle (RF), time interpolated to have the same length, intended to be included in a cvNNMF framework. (B) shows the results of both averaging - indicated by the thick black line with dashed colored lines illustrating the contributing trials - and concatenating, where the dashed vertical lines indicate the sample where one trial's data transitions into another

studies have combined threshold-based criteria with cross-validating inference power by constructing a confidence interval with repeated iterations of a cvNNMF, and setting the threshold to the lower bound of that interval for a given significance level - usually $\alpha = 0.05$. This approach is the most robust by capturing results from different local minima of the objective function and from different sets of trial data [14, 46, 51, 75, 78].

4.2 Methods

To explore solutions to questions raised in section 4.1 regarding cross-validation parameters, the framework shown on Fig. 4.2 was ran with all possible permutations of configurations detailed in Table 4.1. That is, each individual parameter in column "Explored Configurations" was held fixed while the rest were tested for a cvNNMF extraction on all available data, and VAF results were recorded. Additionally, all configuration permutations on the cvNNMF framework were ran multiple times for each number of synergies. This span was chosen with the assumption that the number of synergies underlying the sEMG data can not be larger than the number of recorded muscles. Therefore, the final number of configuration tests were 12 for each movement/speed pair of trial data and number of synergies (2 data division ratios

$\times$ 3 VAF calculation possibilities $\times$ 2 data structuring methods).

Table 4.1: Study's questions about cvNNMF parameterization and studied configurations

Question No. (Fig. 4.2 (B))	Description	Explored Configurations
Q1	% data assigned to Group 1/2	50/50 and 60/40
Q2	Data structuring of trial data prior to NNMF	Concatenating and Averaging
Q3	VAF calculation configuration	$W_{G1} * H_{G1}$ vs. EMG_{G2} $W_{G2} * H_{G1}$ vs. EMG_{G1} $W_{G1} * H_{G2}$ vs. EMG_{G2}

This study used data from all subjects for each dataset comprising one motor task and performing speed. It was assumed that having larger amounts of data would lead to more representative results, and that the need for accurate findings in this parameter determination step outweighed the importance of the preservation of subject heterogeneity for statistical analysis, at least regarding this preliminary study. Notably, using data from different subjects for the same synergy extraction iteration has been done in previous studies to find common inter-subject modularity directly [52, 62, 134].

One configuration was impossible to evaluate due to the nature of the VAF parameter. VAF requires equal-length vectors (because it is based on the Pearson Correlation which for any point in the independent variable demands an observation of the dependent variable). This was not the case for the concatenated-60/40 split-$W_{G1} * H_{G1}$ vs. EMG_{G2} configuration, which yielded one vector spanning 60% of the length of the trials on Group 2 and another spanning only 40% of the same length. This is the reason why, on Fig. 4.4 and subsequent colormaps, the cells corresponding to this configuration are blank.

After gathering insights from the study described in the previous section and configuring the final synergy extraction framework to incorporate these changes, the next step involved the determination of the optimal number of synergies (N_{syn}) to represent data from each movement. To select the number of synergies that could best reconstruct the sEMG data, the whole process was ran varying the N_{syn} parameter between 1 and 4 (for the Knee Extension, Knee Flexion, Sit-to-Stand and Stand-to-Sit movements) and 1-8 for the walking trials. The reasoning behind this decision was that the number of motor modules should not be larger than the number of recording electrodes for a given movement, which were 4 for the Discrete Motor Tasks (even though 8 sensors were recorded but the movements either involved only the right or left side (Knee Extension and Flexion) or were bilaterally symmetrical (Sit-to-Stand and Stand-to-Sit) and 8 for the gait analysis (which is bilaterally asymmetrical and therefore data from all 8 sensors are relevant). A number of synergies was considered sufficient in reconstructing the original data if the lower bound of the 95% Confidence Interval of the VAF constructed from the collective VAF values results of the 10 runs across

the 10 subjects was superior to 90%, following suggestions from several publications [14, 46, 51, 75, 78]. If N_{syn} is shown to differ across speeds, the number of synergies for that movement will be selected as being the *largest number across speeds*.

4.3 Results & Discussion

This section contains results gathered for both preliminary studies and its discussion. In section 4.3.1, findings from the parameter study will be presented and conclusions drawn. Based on these conclusions, the final synergy extraction framework that will be used throughout the dissertation is described. With the framework established, it is possible to proceed onto section 4.3.2 where the number of synergies to represent the data is chosen.

4.3.1 Evaluation of Framework Parameters

Results for the preliminary study of framework parameters (summarized in Table 4.1) spanned too many movement-speed combinations to be represented in its entirety in this section. Therefore, a selection of a couple of movement-speed pairs was made. Results for the Knee Extension trials at slow speed are represented on Fig. 4.4 and those of the walking trials at slow speed may be observed on Fig. 4.5.

By analyzing Fig. 4.4, the first impression that may be taken is that all values are relatively high and similar among each other, except the ones yielded by the concatenated data structuring method on a 50/50 division when the product of the factors derived from one group are used to evaluate reconstruction accuracy on the other groups (i.e., the 4 light blue cells on the lower left corner). This is because, effectively, two different sEMG signals are being compared, one reconstructed from a factorization on a set of trials and the other being raw data from a completely different set of trials, which cannot be expected to be similar, even if the movement being performed is the same and the trials are time normalized. This phenomenon is portrayed in Fig. 4.6. The reason why this happens only for this set of concatenation configurations is because in the other approaches, the activations originate in the same dataset they're being compared against, and therefore the timings of both vectors are aligned ($W_{G2} * H_{\mathbf{G1}}$ vs. $EMG_{\mathbf{G1}}$ and $W_{G1} * H_{\mathbf{G2}}$ vs. $EMG_{\mathbf{G2}}$). VAF penalizes much more slightly mistimed or dissimilar activations than slightly different weighting vectors, because the former when reconstructed produces entirely different shaped curves while the latter produces similar curves with slightly mismatched amplitudes. This is the intended effect because in a motor task, the timing of onset of the muscles' activations is much more important for the production of joint torque and task-specific movement than the slight under or over-

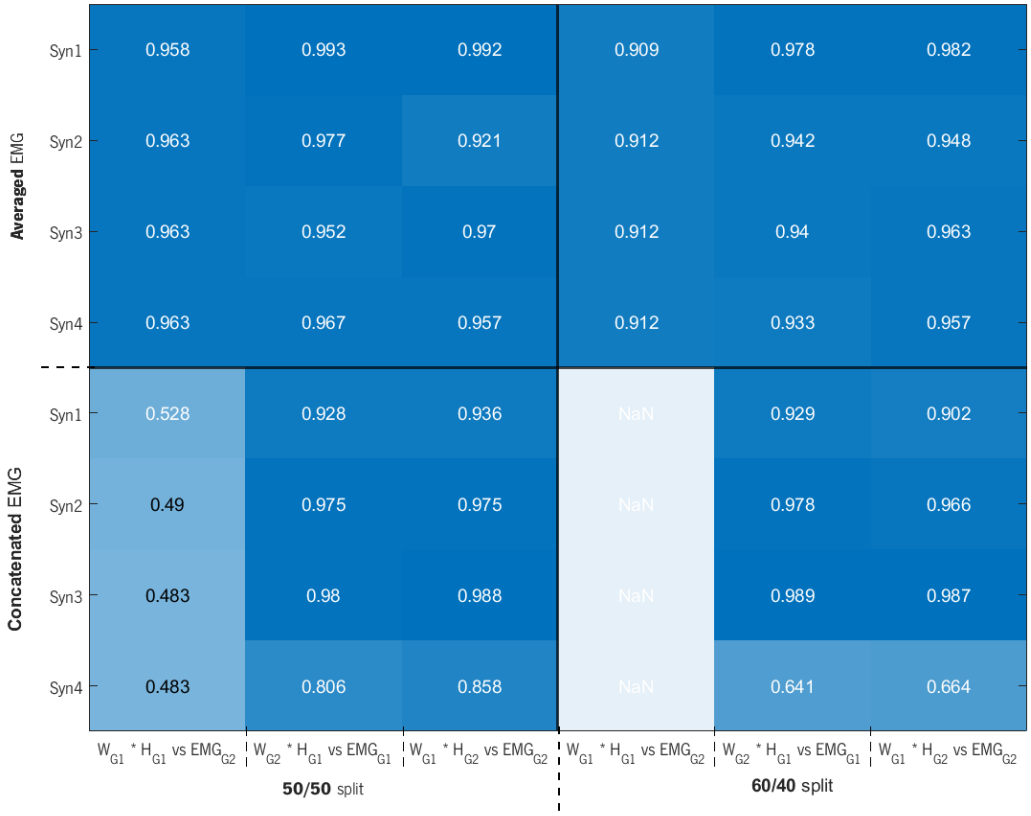


Figure 4.4: Lower bounds for the 95% Confidence Interval of the VAF of 10 runs of cvNNMF on sEMG envelope data from 10 subjects during Knee Extension at slow speed and across two data structuring methods (Y-axis: averaging trials, concatenating trials) and two extraction/validation ratios (X-axis: 60%/40%, 50%/50%). Each data structuring method is further subdivided into rows representing the number of synergies (Syn1-Syn4) the cvNNMF algorithm was configured to extract. Each of the two main divisions of the X-axis is further subdivided into three methods of validation ($W_{G1} * H_{G1}$ vs. EMG_{G2}: using weights from the first group (50% or 60% parcel) combined with activations of the second group (50% or 40% parcel, respectively) to measure the reconstruction quality against the original sEMG data from the second group. Results from Concatenated/60-40 split/$W_{G1} * H_{G1}$ vs. EMG_{G2} are missing because the lengths of the reconstructed and original sEMG signals are different and therefore impossible to evaluate according to the VAF criteria.

recruitment of intervening muscles, but has the side effect of restricting the possible VAF benchmarks when using a concatenated approach.

The second inference that may be extracted from both Figures 4.4 and 4.5 is that the values of VAF for averaged structuring are generally higher than those of the concatenated structuring: this may be taken to mean that the averaged approach results in better reconstructions. However, upon closer inspection, one can observe that as the number of synergies increases, the VAF using the concatenated approach rises

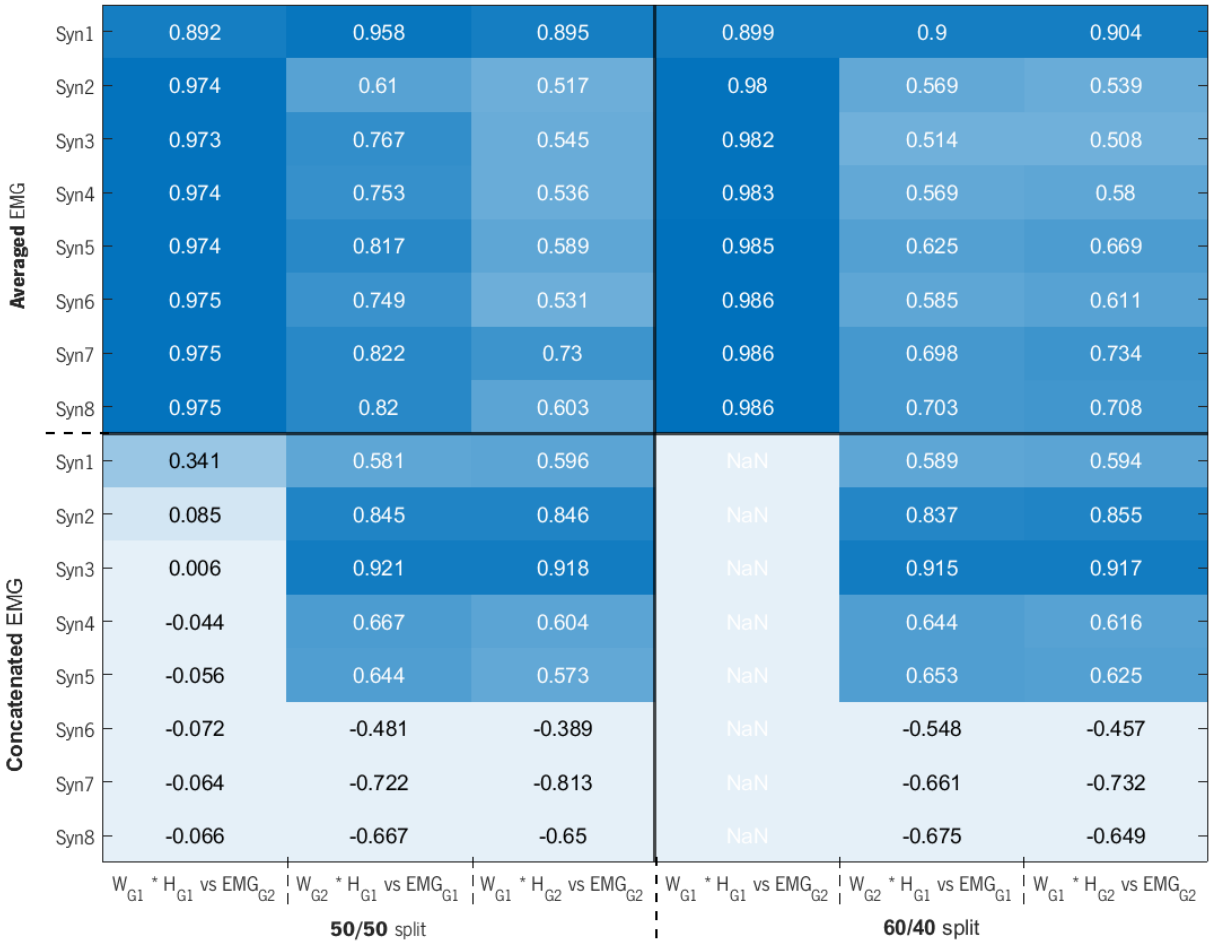


Figure 4.5: Lower bounds for the 95% Confidence Interval of the VAF of 10 runs of cvNNMF on sEMG envelope data from 10 subjects during walking at slow speed and across two data structuring methods (Y-axis: averaging trials, concatenating trials) and two extraction/validation ratios (X-axis: 60%/40%, 50%/50%).

from Syn1 to Syn2, maintains the same value for Syn3, then decreases for Synergy 4, regardless of data splitting proportion; while results from the averaging approach are consistently high (>0.9) independent of number of synergies. This may mean one of two things: either the averaging technique is simply better at reconstructing the sEMG envelopes and 1 synergy is already the optimal number of motor modules, or the concatenating technique is accounting for variability not present in the averaging approach and therefore portraying reconstruction quality more accurately when yielding the highest values of VAF for 2 and 3 synergies. Results from the walking trials at slow speed (Fig. 4.5), with many more trials (350 trials vs. 30), may lead to more robust conclusions regarding the most appropriate structuring method. The findings of averaging structuring state that VAF is only large (>0.9) for one synergy using the averaging technique, and that for synergies 2-8 it is considerably lower than 0.9. This finding is in contrast to results acquired from the concatenating technique, where VAF rises steadily up until three synergies, attaining values higher than 0.9. This finding supports the hypothesis that the concatenated approach portrays

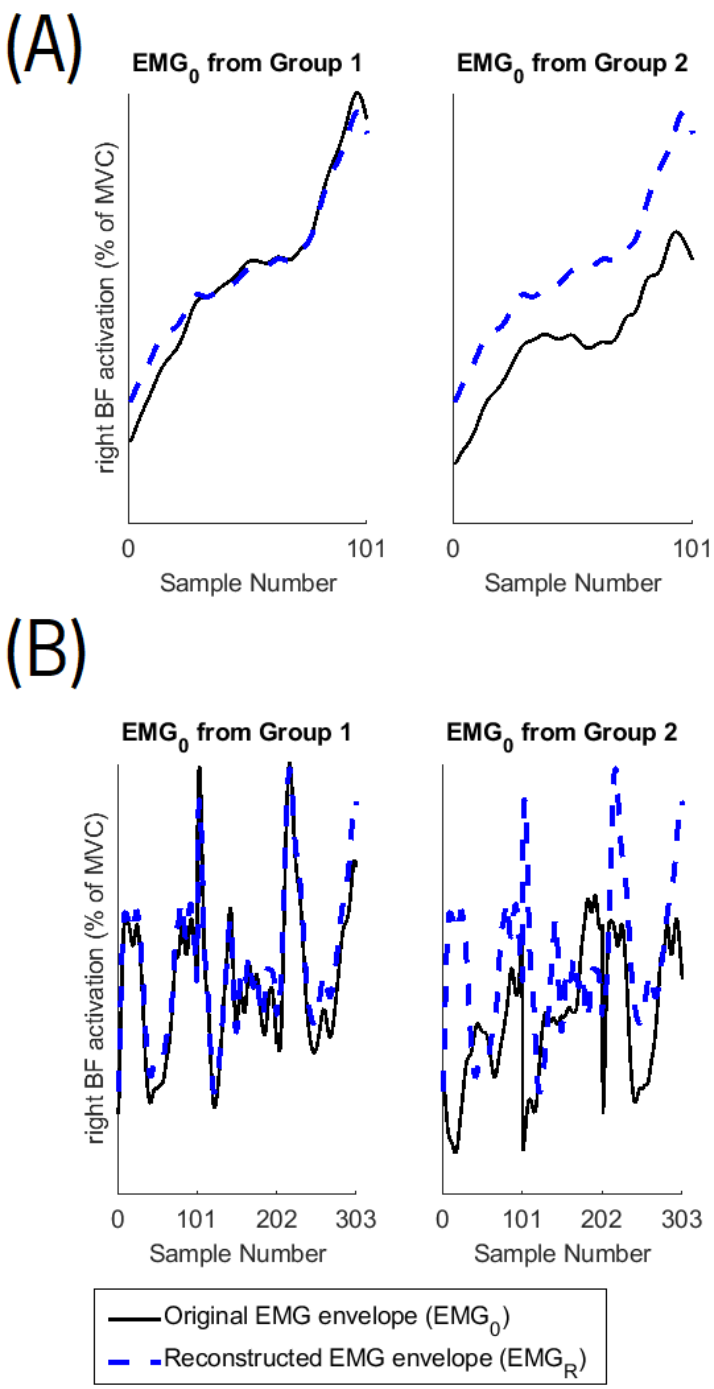


Figure 4.6: Illustration of two different VAF evaluation methods. The dark thick lines represent the original data fed to the synergy extraction framework (EMG_0), while blue dashed lines show the sEMG reconstructed from the resulting synergies using (A) an averaged approach and (B) a concatenated technique, for three normalized trials of the BF during Knee Flexion. For the figures on the left side, the benchmark data was taken from Group 1, and the right's, from Group 2. The clear mismatch in shape in the bottom right figure explains the poor VAF results from evaluating reconstruction accuracy with activations derived from a different group's data.

variability on the input data better than by averaging, because averaging tends to "smooth" individual trials' idiosyncrasies to the point that important information may be lost. Concatenating, on the other hand, forces the cvNNMF algorithm to incorporate this sporadic-but-important variability, yielding motor modules more reliable across trials, subjects and performing conditions and better representative of the modularity theoretically present in sEMG signals of a task-oriented movement.

Therefore, the chosen methodology for structuring the data was to concatenate it prior to factorization. Among the results portrayed in Fig. 4.4 and Fig. 4.5, the differences between data division ratios seem negligible, with no method outperforming the other consistently. The same may be observed for the group of data chosen as reference ($W_1 * H_2$ vs. EMG_2 or vice-versa). Therefore, if no method appears to be empirically preferable, the decision made was to split the variability equally between the two factors (W_1 and H_1) by using 50% of the data to form Group 1 and an equal percentage to create Group 2, thereby maximizing the trials available for each of the two sequential NNMF runs in a cvNNMF framework. This made the choice of VAF benchmark even more irrelevant as both groups have equal number of randomly-assigned trials. Therefore, the dataset to act as the VAF benchmark was arbitrarily set to Group 1 ($W_1 * H_2$ vs. EMG_2). Conclusions drawn from the previous study were used to build the final synergy extraction framework, schematized in Fig. 4.7.

Q1 (Table 4.1) posed the question of what was the ideal data partition proportion for Cross-Validation. The preliminary study revealed that no method consistently reported higher VAF's, therefore the ratio was set at 50%/50% for Group 1 and Group 2 data, in order to split the variability equally among both groups; hence "Randomized 50/50 data partition" in the Data Structuring block of Fig. 4.7. Addressing Q2, dealing with the best method of structuring trials data prior to NNMF, it was decided to follow a concatenated approach. This conditioned the answer to Q3 ("What is the best VAF calculation configuration?"), as activations had to be timed to the benchmarking sEMG signal (i.e. $W_1 * H_1$ vs. EMG_2 could not convey reconstructing ability accurately). Of the remaining two options for computing of VAF, both were indifferent considering the data split proportion was set to 50/50; therefore, the final VAF benchmark was set to $W_1 * H_2$ vs. EMG_2, hence the origin of inputs that result in EMG_R on Fig. 4.7.

4.3.2 Minimum Number of Synergies

Fig. 4.8 shows the results of the evolution of VAF with a varying number of synergies. Table 4.2 demonstrates the minimum number of motor modules that fulfilled the criteria and final number of synergies that was taken from this analysis. For the Knee Extension movement, one synergy was sufficient to achieve a Lower Bound of the Confidence Interval (LBCI) superior to 90 across all speed conditions. This

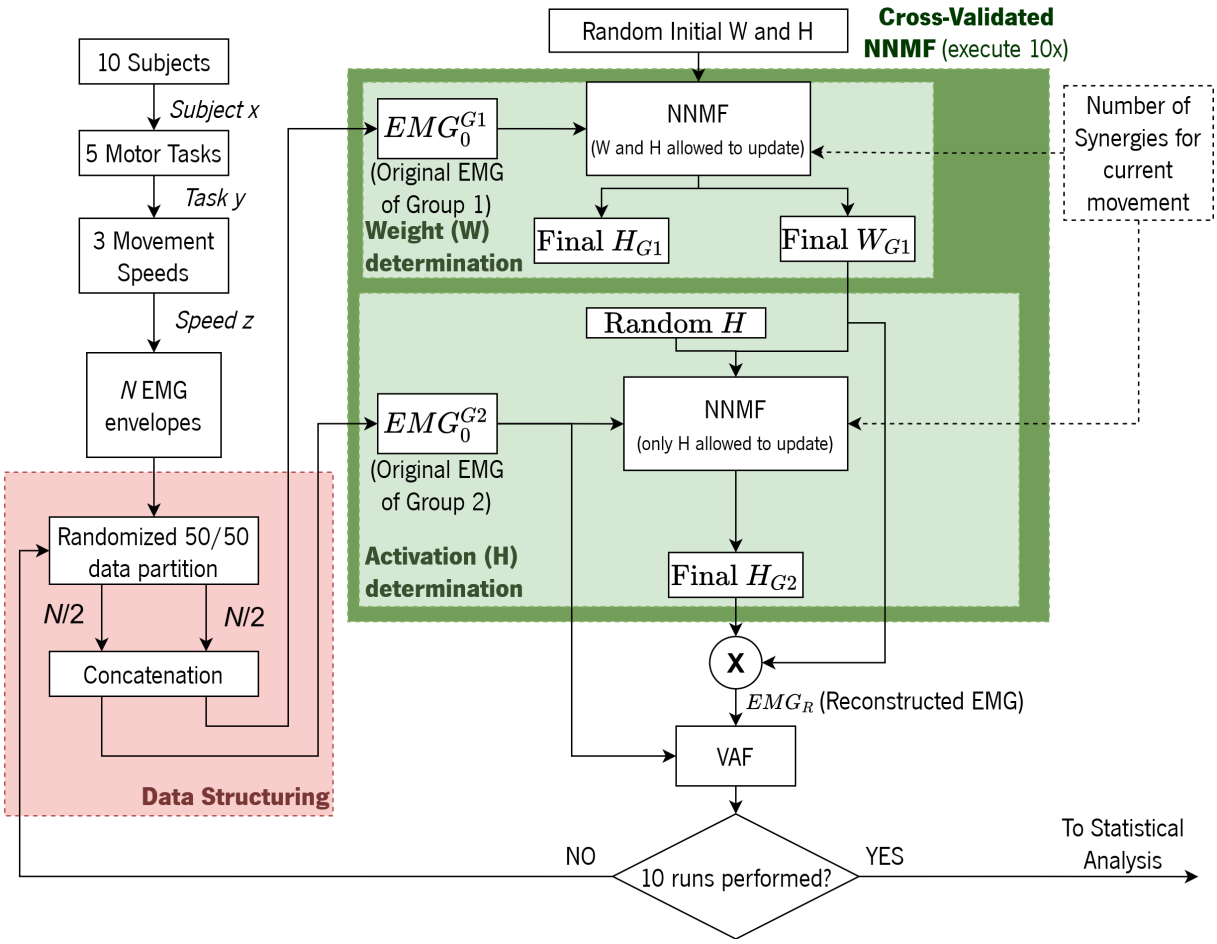


Figure 4.7: Schematic of the Final Multi-Task Synergy Extraction Framework

was expected since extending the knee is a somewhat simple movement requiring the recruitment of only the quadriceps. The same would have been expected of Knee Flexion, requiring the contraction of the Hamstrings to elevate the shank; however, empirical tests revealed that the number of synergies was 2 for all velocities. This finding was observed across almost all subjects. We hypothesize that this may be due to the experimental procedure: subjects were not instructed on how to move their foot during flexion of the knee, in order to allow the natural performance of the movement. Consequently, the second synergy might correspond to activation of the dorsiflexors (of which we measured the TA) or of the plantarflexors (of which we measured the GM), recruited in order to maintain the ankle static throughout the movement by responding to inertial changes on the foot. This hypothesis will be examined when correlating the synergies with the kinematic variables in section 5.4.3. For the previous two movements, literature in the synergy extraction space is lacking and therefore cannot be used to support or refute any hypothesis.

Regarding Sit-to-Stand task, this dissertation found 2 synergies explaining 90% of the variance of the data across 10 subjects each with 95% confidence. In the literature, authors have repeatedly found 3 motor modules underlying standing while measuring 6-7 muscles [14, 95, 118]. This does not directly

corroborate the results found for this study but does support the possibility that several motor modules are needed to encode sitting-to-standing.

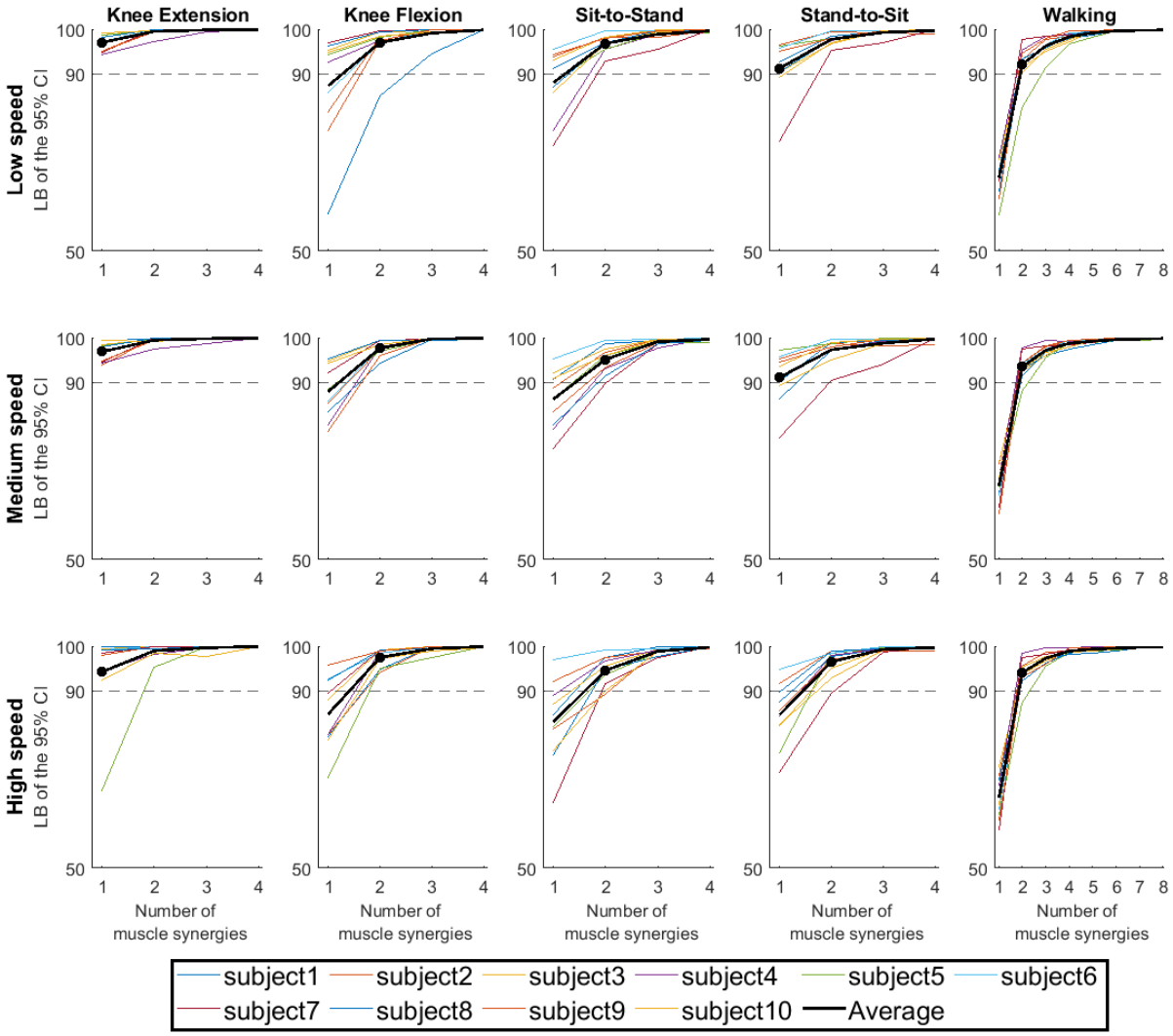


Figure 4.8: Values of the Lower Bound of the 95% Confidence Interval (LBCI) of the VAF of 10 runs of cvNNMF as a function of the number of synergies across five lower-limb movements (columns) and three speed conditions (rows). For each movement-speed pair, the Lower Bounds of the 95% CI across the 10 participants (thin colored lines) were averaged (thick black line) and the minimum number of synergies that equaled or surpassed 90 (indicated by horizontal dashes) was highlighted (black filled circle). This number was consistent across speeds, except for the Stand-to-Sit movement

Findings from Stand-to-Sit were the only ones that were not consistent across speeds. The slow and medium speed data yielded one motor module explaining the vast majority of the variance across subjects, but in the fast condition, this number was two. There is no research into synergy extraction for unweighted sitting to look for hints of the cause of this phenomenon. There is, however, one work that has studied the motor modules in weighted squatting from which rough parallels may be drawn [52]. In this study, three

Table 4.2: Minimum number of synergies fulfilling the VAF threshold criteria and final number of synergies chosen for each movement

Movement	Speed Condition	Minimum number of synergies with average LBCI>90	Final number of synergies
Knee Extension	Slow	1	1
	Medium	1	
	High	1	
Knee Flexion	Slow	2	2
	Medium	2	
	High	2	
Sit-to-Stand	Slow	2	2
	Medium	2	
	High	2	
Stand-To-Sit	Slow	1	2
	Medium	1	
	High	2	
Gait	Slow	2	2
	Medium	2	
	High	2	

synergies were found in the combined sit-to-stand and stand-to-sit movement, of which the activation of two seemed to be more prevalent in the period corresponding to the descending portion of a squat (Fig. 2.6 (A) up until 60% of normalized squat period). These were modules involving a heavy TA recruitment during the descent, peaking just before the lowest point of the participants' center of mass (i.e. when the thighs were parallel to the floor). The second synergy, on the other hand, was characterized by Vastus and RF expression, rising steadily from the beginning and peaking at the lowest point of the subjects' center of mass. These findings are hardly applicable to our analysis, but are useful in that they reject an unidimensional synergy model. Furthermore, as the faster sitting motion requires more energy, more torque being produced at the joints and greater postural stabilization effort [110], this may be taken to explain why the results differed for the fast condition: a concentrated recruitment of TA during the descent may not be as necessary to stabilize the center of mass in slow and unweighted sitting than during fast and weighted motion. It may not warrant another motor module, or its expression may be absorbed into other synergies without significantly affecting the task. Regardless, to allow for later inter-velocity analysis, it was decided to keep the number of synergies constant across speeds and hence a number of synergies equal to two was chosen for all Stand-to-Sit trials. This fulfilled the criteria of the LBCI being larger than 90 on all speeds simultaneously.

Finally, for the walking experiment, two synergies were found underlying movement in all speed conditions. This is at odds with insights from published research which point to 4-5 synergies. However, these

studies measure anywhere from 8 to 16 ipsilateral muscles [13, 46, 51, 85, 87, 100] whereas only four were recorded during our data acquisition protocol.

Chapter 5

Statistical Analysis of Extracted Synergies

In this chapter, the synergies extracted using the previously established framework will be examined and validated through statistical methods.

Section 5.1 explains the methodological assumptions and gives an overview of the statistical tools used on the rest of this chapter. To allow the drawing of statistical conclusions through ANOVA, synergies will be examined on an intra-subject basis (i.e. subjects' motor modules will be examined independently of one another). This will be done hierarchically from a ground-up approach, first investigating similarities and differences across the 10 runs for each subject (section 5.2 - Inter-Run Analysis), then relationships between results from distinct speed conditions (section 5.3 - Inter-Speed Analysis) and finally similarities across results from different subjects (section 5.4 - Inter-Subject Analysis) in an attempt to find a set of motor modules common to all subjects.

5.1 Methods for Statistical Analysis

Based on the literature findings, similarity between a pair of synergies was measured with three metrics, given their versatility and interpretability. These three metrics and their role in the problem context are illustrated in Fig. 5.1. The first data characteristic, Cosine Similarity ($cosSim$) was used to investigate agreement between weights (W). For two weighting vectors $\vec{W}_1$ and $\vec{W}_2$ this can be expressed as described in Eq. 5.1:

$$cos(\theta) = \frac{\vec{W}_1 \cdot \vec{W}_2}{\left\|\vec{W}_1\right\| \left\|\vec{W}_2\right\|} \tag{5.1}$$

Thus, two vector are maximally similar if they're parallel ($cos(\theta) = 1$) and maximally dissimilar if

they're orthogonal ($cos(\theta) = 0$).

The second metric, the Pearson Correlation (r), measured the likeness of the waveforms of the activations (H). For a data point $H1_i$ in an activation curve $H1$, and point $H2_i$ in curve $H2$, with corresponding means $\bar{H1}$ and $\bar{H2}$, r may be stated as in Eq. 5.2.

$$r = \frac{\sum_{i=1}^{n}(H1_i - \bar{H1})(H1_i - \bar{H2})}{\sqrt{\sum_{i=1}^{n}(H1_i - \bar{H1})^2(H2_i - \bar{H2})^2}} \tag{5.2}$$

Therefore $r = 1$ if the curves are completely matched in shape and timing.

Finally, the third consisted of the lag time at the maximum of the cross-correlation output between two activation vectors (*lag*), normalized by the length of the curves, and quantified the difference in the timing of two activations (i.e. the magnitude of the phase shift between two waveforms plus the direction of the shift indicated by the sign - negative meaning the reference curve is delayed in relation the other and positive the inverse; that the reference curve is anticipated). Following the same nomenclature used for the previous equation, cross correlation function may be expressed by Eq. 5.3

$$R_k = \frac{\sum_{i=1}^{n-k}(H1_i - \overline{H1})(H2_{i+k} - \overline{H2})}{\sqrt{\sum_{i=1}^{n}(H1_i - \overline{H1})\sum_{i=1}^{n}(H2_i - \overline{H2})}} \tag{5.3}$$

where k is an integer encoding the lag index. The *lag* metric is simply the k for which R_k is maximized, normalized to the length of the curves n.

With synergy similarity quantified in type and magnitude, it was possible to assess if the synergies were different across conditions through one-way ANOVA which attempts to find significant differences across group means (groups being the synergy extraction runs or synergies in different speed conditions). A Shapiro-Wilk test of normality was applied and revealed that the computed similarities were not normally distributed, regardless of movement or speed condition ($p < 0.05$). Therefore, the Friedman test was chosen as this is the non-parametric option for ANOVA across group means. However, this test only informs about the *existence* of different means between groups, but not the groups responsible for these means. Therefore, a *post-hoc* paired test, namely the Wilcoxon signed rank test with Bonferroni adjustment was selected [37, 67]. Finally, the significance level was set to $\alpha = 0.05$. All statistical tests were ran on SPSS version 26.0 (IBM Corp.; Armonk, NY, USA).

The use of information conveyed by the results of these statistical analysis was structured in a hierarchical manner. Firstly, if statistical tests operating on *cosSim* of weights to a reference revealed these weights to be dissimilar, the synergy was marked as outright dissimilar, regardless of results on the remaining metrics. This followed the reasoning that if the weighting vectors were found to be different,

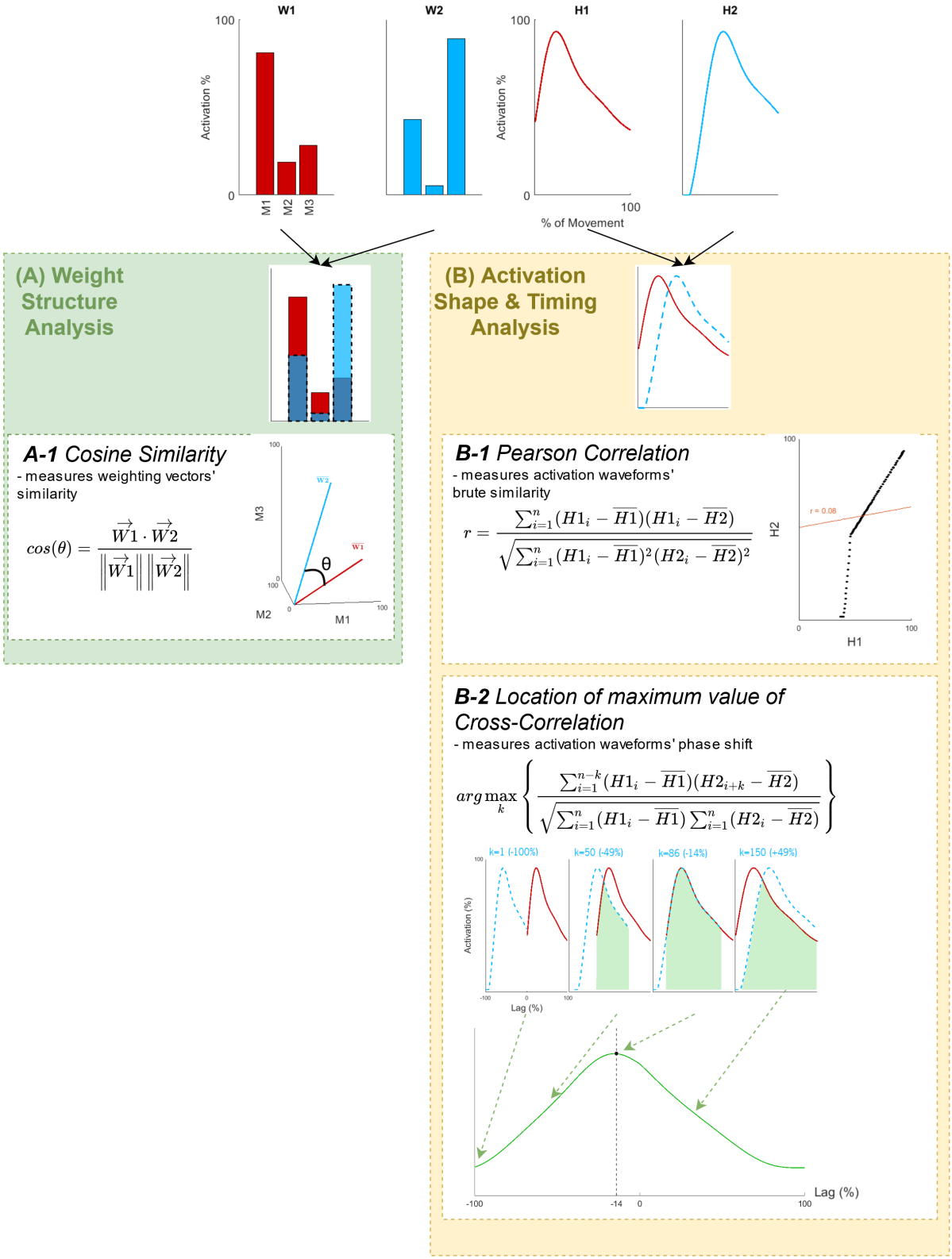


Figure 5.1: Illustration of the extraction of similarity metrics between two synergy weighting vectors (W1 and W2) and their corresponding activations (H1 and H2).

the muscles being recruited were different and the corresponding activation curves were also different, or if found to be similar would be so by coincidence as they would be encoding temporal activation of a different set of muscles. This decision process is represented in Fig. 5.3 (B) in the block labeled "Weight Structure Analysis". On the other hand, if there were no differences among means of $cosSim$ across subjects, the next step would be to investigate activations: first in morphology, then in magnitude of phase shift. This was done by running the statistical framework on computed r's and in the case of rejection of the null hypothesis, marking the entire synergy as "similar with dissimilar activation curve" (Fig. 5.3 (B): "Activation Shape Analysis"). Subsequently, if both weight structure and activation shapes were deemed not significantly dissimilar, the same protocol would be ran again, this time on the lag metric and the synergy would be considered as "Similar with lagged activations" (Fig. 5.3 (B): "Phase Shift Analysis"). Finally, failing to reject any of the null hypothesis would result in a "Similar Synergy" classification.

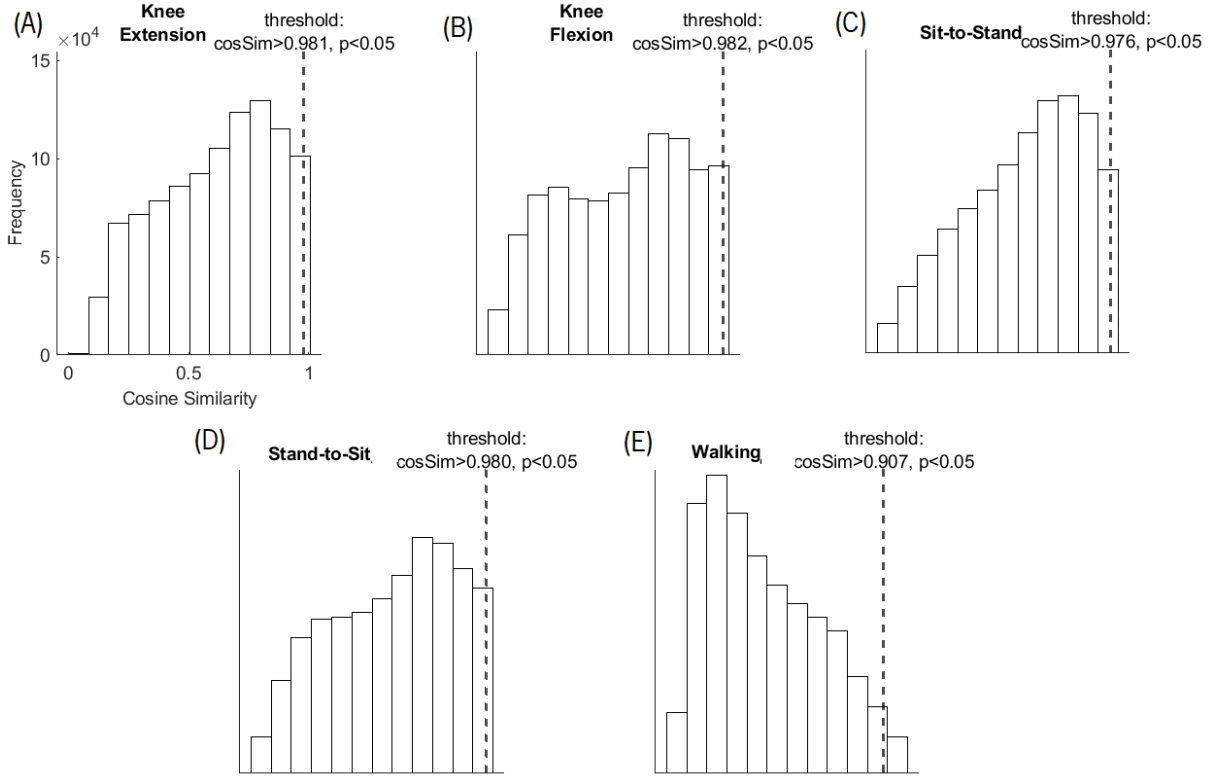


Figure 5.2: Illustration of the criteria used to classify muscle synergies' similarity for the inter-subject study. Two sets of 1000 random muscle synergies were extracted from the available synergies and their pairwise cosine similarities computed to generate a distribution of cosine similarities expected by chance. A pair of synergies were considered "similar" if $cosSim$>threshold, corresponding to the Upper Bound of the 95% bootstrapped Confidence Interval (dashed vertical line)

The statistical analysis illustrated in Fig. 5.3 (B) was used for the inter-run and inter-speed study. However, for the inter-subject study, this statistical analysis framework is not applicable since the inde-

Table 5.1: Upper Bounds of the 95% Confidence Intervals (Upper Bound of the Confidence Interval (UBCI)) of the bootstrapped distributions for each movement

Movement	Knee Extension	Knee Flexion	Sit-to-Stand	Stand-to-Sit	Gait
(UBCI)	0.9814	0.9812	0.9764	0.9802	0.9084

pendent variable is not consistent across observations (i.e. each subjects synergies' similarity to all other participants). Therefore, for the inter-subject study, *bootstrapping* was used to apply statistical confidence to synergy similarity analysis. Bootstrapping is a non-parametric approach to statistical inference that substitutes computation for distributional assumptions and asymptotic results. It works by imitating sampling from the population by treating the available samples as if they were the population; that is, by randomly selecting a set of observations repeatedly, computing the desired metric on the resulting dataset and building a confidence interval with the results [135].

Hence, this technique was used in order to empirically determine a *cosSim* threshold that would ascertain if a pair of synergy weights were similar or dissimilar, for a significance level of $\alpha = 0.05$. In practical terms, this was achieved by pooling all extracted synergies for a given movement, spanning all NNMF runs over all subjects across all speed conditions, and randomly sampling n weights (n being the number of muscles) k times (k being the number of *replications*), generating k random synergies. This process is repeated on the same dataset. Then, the Cosine Similarity is computed between all possible pairs of random synergies from the two sets ($k \times k$ pairs in total). These values are then sorted, yielding a distribution of Cosine Similarities expected by chance. Finally, the 95% Confidence Interval is estimated by selecting the 2.5 and 97.5 percentiles of the resulting distribution, and the upper bound is set as the threshold cosine similarity (each movements' distribution and corresponding upper percentile is demonstrated on Fig. 5.2). Therefore, weight pairs with a cosine similarity above this value were considered more similar than expected by chance, and those on or below were categorized as dissimilar ($p < 0.05$). Results for each movement are summed up in Table 5.1. This distribution-based thresholding approach and the number of replications k was set to 1000 following previously published literature [91, 131].

Fig. 5.3 (A) shows a schematic representation of this three-step statistical analysis.

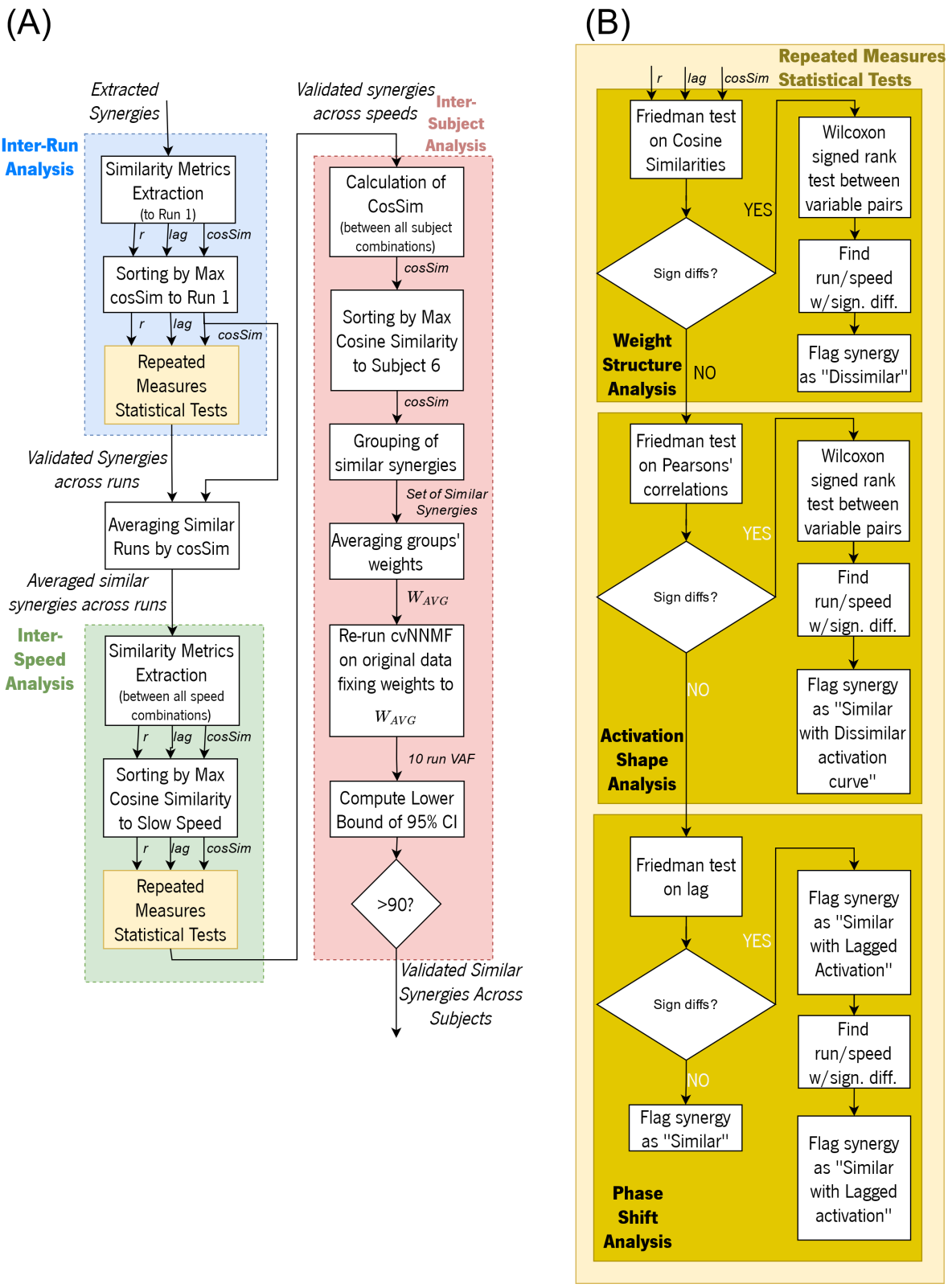


Figure 5.3: Schematic representation of the Statistical Analysis framework used to validate synergy extraction outputs. (A) shows the flow of data for the entire process from output of the synergy extraction framework up until the determination of common synergies across subjects. (B) represents a flow chart of the "Repeated Measure Statistical Tests" block on (A), demonstrating the hierarchical architecture of classifying synergies based on different measures of similarity between their weights or activations.

5.2 Inter-Run Analysis

For the first phase of the study, the 10 runs of the cvNNMF for the 10 subjects were subject to an ANOVA, aiming to determine if the output synergies from these runs were significantly different from each other. After the statistical tests, synergies dissimilar to the reference, as determined by the distribution-derived similarity threshold, were discarded from future analysis in order to determine a set of motor modules that would represent the structural consensus of the synergy extraction runs.

5.2.1 Methods

The first step consisted of computing similarity metrics for each of the 10 synergy extraction runs. To accomplish this, an arbitrary reference had to be set to act as a common baseline to which other runs could be compared to. Run 1 was chosen for all the studied datasets. As explained in Chapter 2, a functional sorting of the synergies is required if weight and activation comparison results are to represent functional differences rather than quantifying agreement between synergies responsible for other muscle contractions. Therefore, a sorting step was implemented, where synergies were sorted by their weights' similarity to the reference run, as measured by $cosSim$. If reordering occurred for a given run, new metrics were computed to reflect it.

By this point, each of the 10 subjects had 9 variables encoding similarity of run 2-10 to the reference run 1. This could be stored in a 10 x 9 matrix which could be examined for differences across means of each run through the framework established in the previous section (Fig. 5.3 (B)). These resulted in three matrices for each movement, speed and synergy index combination.

5.2.2 Results

The resulting test statistics for the Friedman test (Chi-square distribution - $\tilde{\chi}^2$) and p-values are summed up in Table 5.2. No significant differences were recorded for cosine similarity to run 1 on any movement, speed or synergy ($p < 0.05$). For the study of the waveforms similarity (r to Run 1), three sets of synergy activations revealed significant differences: two from synergies extracted from the sit-to-stand movement (slow, synergy index 1, $p = 0.007$; and fast, synergy index 2, $p = 0.026$) and another set from synergies computed on data from the stand-to-sit movement (fast, synergy index 1, $p = 0.041$). After performing the Wilcoxon signed rank test with a Bonferroni adjusted significance level on all pairs of Run similarities, no significant differences were reported for any of the pairs on the examined similarity matrices.

Table 5.2: Inter-Run Friedman tests results. Chi-square ($\tilde{\chi}^2$) distributions and corresponding p-values are reported for each movement-speed-synergy combination. Bold values indicate Friedman tests where the null hypothesis - that synergy extraction runs' mean similarity to a reference run were the same for every subject - was rejected ($p < 0.05$). *cosSim: Cosine Similarity; r: Pearson's correlation coefficient; W: Synergy weighting vector; H: Synergy temporal activation.*

		cosSim of W's		r of H's		*lag* of H's	
Movement (Speed)	Synergy Index	$\tilde{\chi}^2$	p	$\tilde{\chi}^2$	p	$\tilde{\chi}^2$	p
Knee Extension (Slow)	1	7.280	0.507	9.001	0.342	0.000	1.000
Knee Extension (Medium)		7.980	0.435	8.849	0.355	8.000	0.433
Knee Extension (Fast)		4.514	0.808	8.668	0.371	8.000	0.433
Knee Flexion (Slow)	1	6.533	0.588	6.187	0.626	10.314	0.244
Knee Flexion (Medium)		13.920	0.084	13.813	0.087	5.730	0.677
Knee Flexion (Fast)		6.827	0.555	11.333	0.184	11.767	0.162
Knee Flexion (Slow)	2	10.027	0.263	7.387	0.496	4.235	0.835
Knee Flexion (Medium)		3.973	0.860	2.453	0.964	9.112	0.333
Knee Flexion (Fast)		6.107	0.635	12.293	0.139	4.780	0.781
Sit-to-Stand (Slow)	1	6.533	0.588	**20.987**	**0.007**	8.798	0.360
Sit-to-Stand (Medium)		10.373	0.240	3.413	0.906	7.427	0.491
Sit-to-Stand (Fast)		8.373	0.398	15.147	0.056	4.920	0.766
Sit-to-Stand (Slow)	2	8.267	0.408	4.987	0.759	5.916	0.657
Sit-to-Stand (Medium)		5.093	0.748	4.373	0.822	6.980	0.539
Sit-to-Stand (Fast)		8.693	0.369	**17.387**	**0.026**	8.212	0.413
Stand-to-Sit (Slow)	1	5.360	0.718	3.627	0.889	13.616	0.092
Stand-to-Sit (Medium)		7.093	0.527	8.747	0.364	7.349	0.499
Stand-to-Sit (Fast)		11.547	0.173	**16.107**	**0.041**	5.971	0.650
Stand-to-Sit (Slow)	2	5.360	0.718	5.147	0.742	6.331	0.610
Stand-to-Sit (Medium)		10.720	0.218	8.267	0.408	5.328	0.722
Stand-to-Sit (Fast)		6.693	0.570	1.680	0.989	5.454	0.708
Gait (Slow)	1	4.480	0.811	5.040	0.753	9.963	0.268
Gait (Medium)		7.760	0.457	5.893	0.659	6.174	0.628
Gait (Fast)		11.067	0.198	6.347	0.608	7.111	0.525
Gait (Slow)	2	10.800	0.213	6.480	0.594	11.307	0.185
Gait (Medium)		9.360	0.313	7.947	0.439	8.000	0.433
Gait (Fast)		13.973	0.082	12.880	0.116	9.989	0.266

5.2.3 Discussion of results

The absence of significant differences across runs for the weight structure analysis (columns labeled as "*cosSim* of W's" in Table 5.2) indicates that similarity to the reference run was consistent regardless of cross-validated iteration. This can be interpreted as meaning that muscle recruitment was consistent regardless of synergy extraction run for every subject. Regarding the activation curve shape as measured

by r ("r of H's" in Table 5.2), three synergies revealed significant differences across runs. *Post-hoc* analysis, however, indicated that the activation's waveform was not different across synergy extraction runs.

It is nonetheless important to examine the three runs that yielded differences in the r of H's. Firstly, the activations for the 10 runs of Sit-to-Stand performed on a slow pace are shown in Fig. 5.4, along with the corresponding r-values to Run 1. By analyzing Fig. 5.4 (A) it is evident that waveforms in these runs do not seem to be different from the remaining iterations. Upon closer inspection of the same runs in Fig. 5.4 (B), the potential source of sparse r's becomes clearer: the shape of the activation of the reference run in subject 3 and 10 is different from most of the remaining runs'. As observed in Fig. 5.4 (A), both of these subjects scored low on Run 2, and represent the lowest average similarity from all the participants (0.80 ± 0.15 and 0.85 ± 0.09 for participants 3 and 10, respectively). This dual source of inconsistency plus natural variability present in other subjects' iterations (like subject 1 which scored an r of 0.69 on Run 2) may have led to the rejection of the null hypothesis of equality in run similarity means.

A similar phenomenon resulted in rejection of the null hypothesis for synergy 1 of the fast sit-to-stand, as is demonstrated by Fig. 5.5. Activation of run 1 on several subjects' results were structurally different from the remaining runs. Visually, this is most obvious for subjects 3, 9 and 10 (Fig. 5.5 (B)), supported by the low collective average similarity (0.66 ± 0.35, 0.38 ± 0.62, 0.73 ± 0.41, respectively, shown in (A)). The choice of reference run contributed to incorrect signaling by the analysis of variance of a dissimilar synergy across runs. In fact, most runs were similar among each other as can be seen in Fig.5.5 (B).

Fig. 5.6 illustrates the results for the last synergy in which the null hypothesis was rejected (fast stand-to-sit). Results for this synergy were notorious because the p-value was the largest ($p = 0.041$) of the three examined modules. Significant differences reported in Table 5.2 may be attributed partly to reference selection bias, and to the Friedman's ANOVA tendency to find nonexistent significant differences with increasing number of groups. There exists literature supporting our conclusions regarding the shortfalls of ANOVA followed by a pair-wise *post-hoc* test: omnibus tests such as the Friedman's are designed to preserve type I errors (signaling differences between groups where there were none) in order to avoid unnecessary *post-hoc* testing, but also to minimize the occurrence of not signaling paired differences where they existed significantly, a much more damaging finding to oversee [136].

Lastly, the statistical analysis revealed the activation timing was consistent across runs, as evidenced by the failure to reject the null hypothesis in all synergies for the lag similarity metric analysis (*lag* of H's pair of columns in Table 5.2) ($p < 0.05$). It is therefore demonstrated that synergies were similar in muscle recruitment, temporal activation and temporal timing ($p < 0.05$).

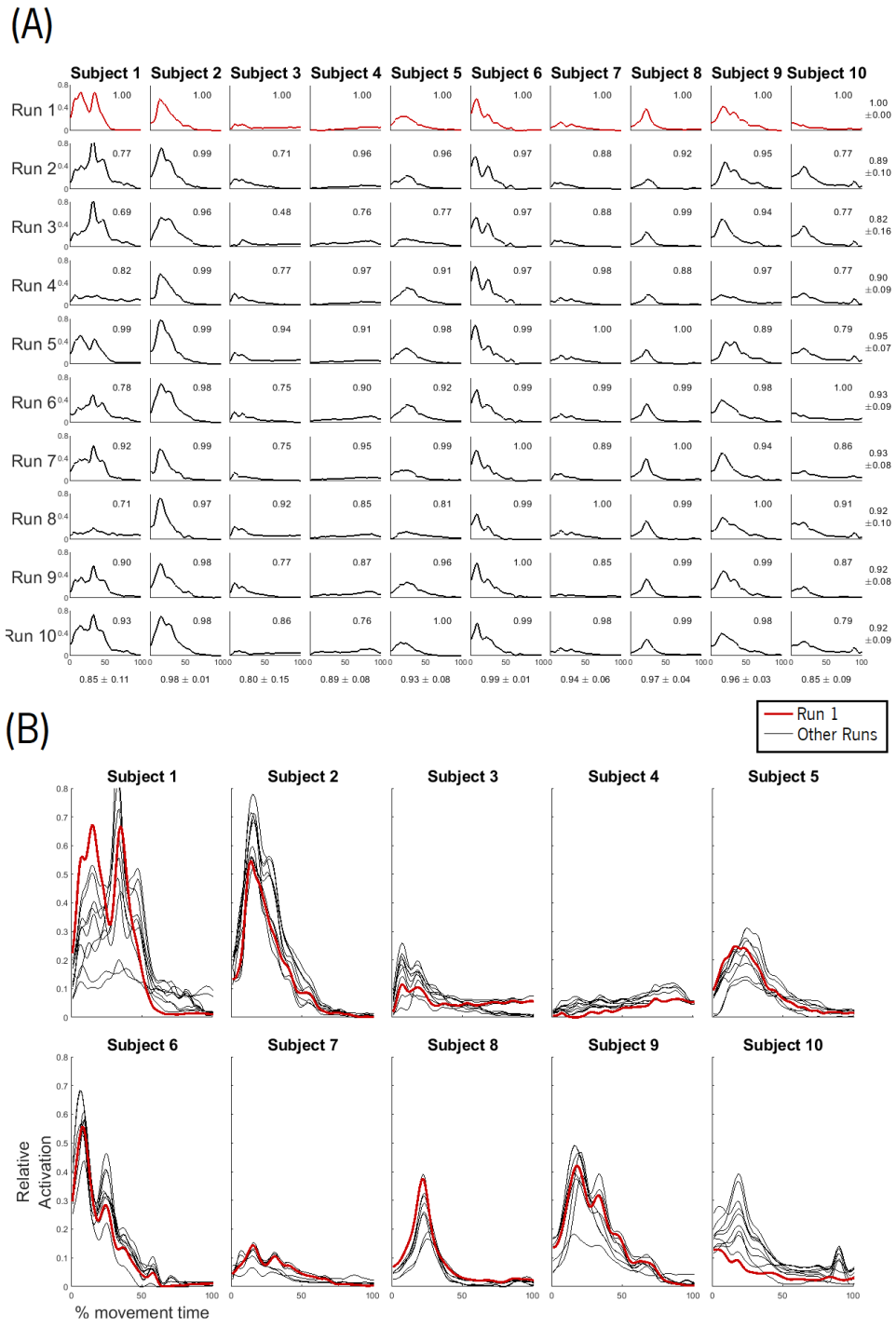


Figure 5.4: Activations of synergy 2 for runs on data from slow sit-to-stand movement. Run 1 (red line) was set as a benchmark with which other runs' similarity was evaluated with, as assessed by the Pearson Correlation coefficient (r) (inset value in (A)). Mean similarity $\pm$ standard deviations are also indicated for each subject across runs and for each run across subjects to aid understanding of the source of significant differences. (A) shows activations of each run separately and corresponding r; (B) shows the same waveforms grouped by subject and overlaid on top of each other.

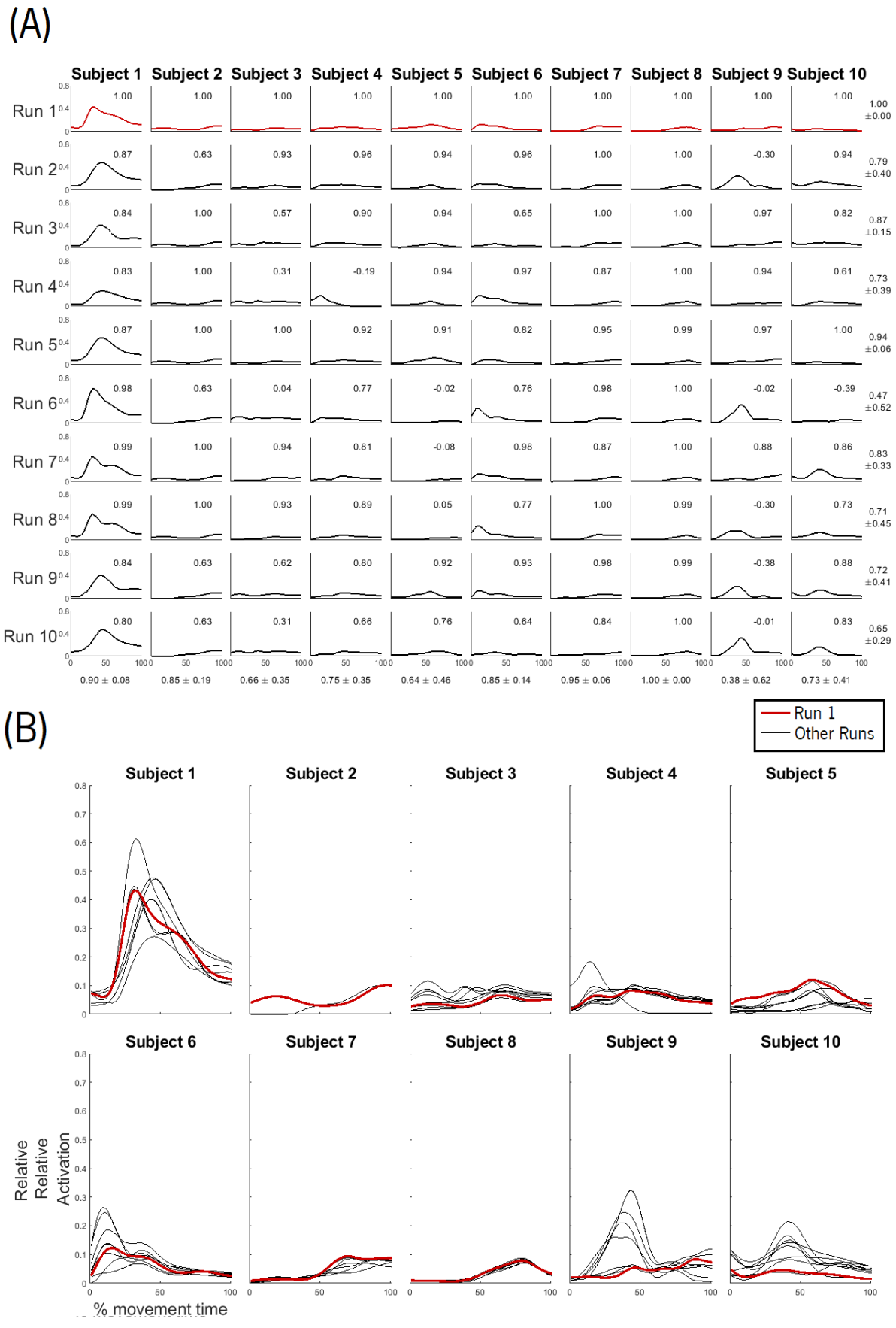


Figure 5.5: Activations of synergy 1 for runs on data from fast Sit-to-Stand movement.

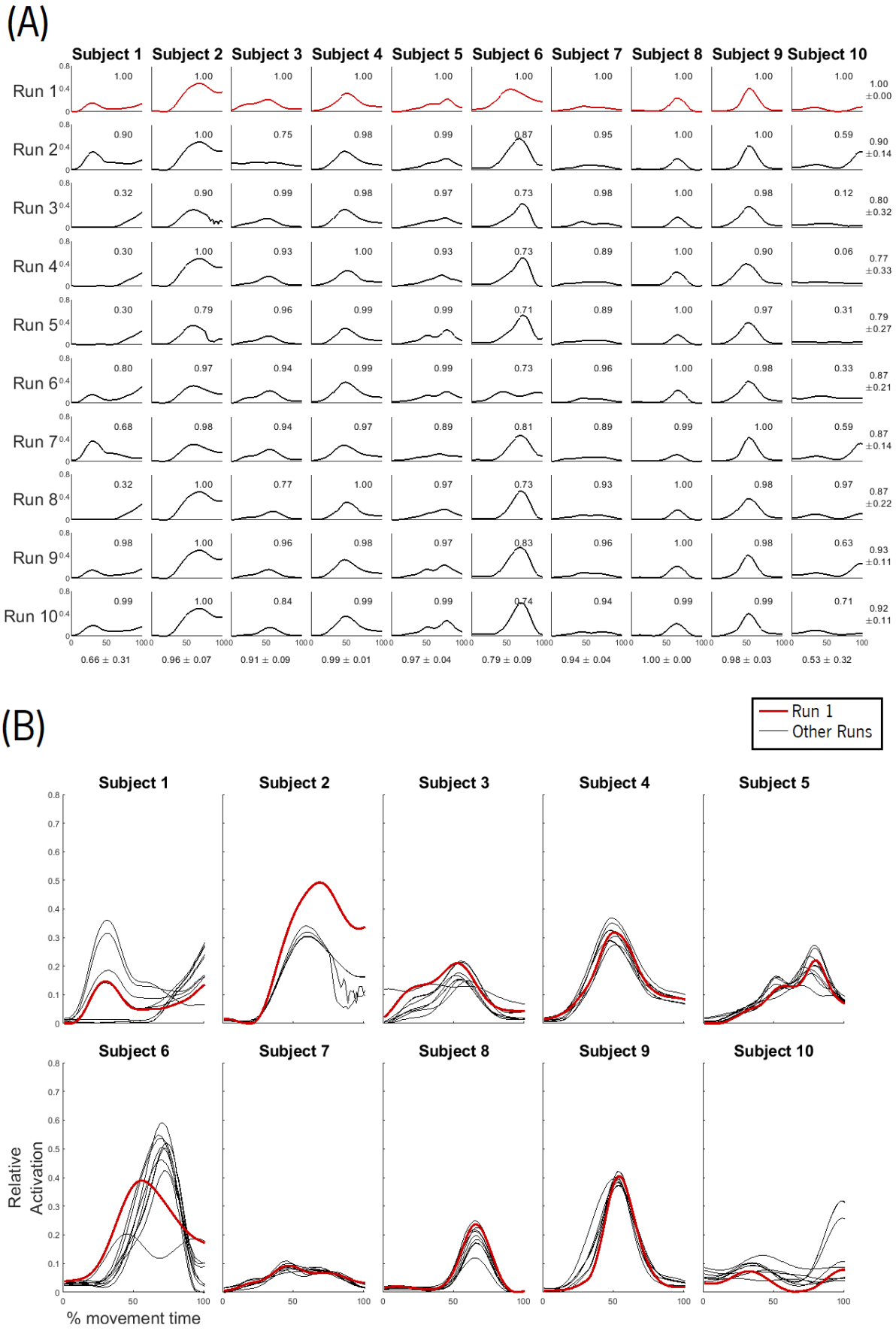


Figure 5.6: Activations of synergy 1 for runs on data from fast Stand-to-Sit movement.

5.3 Inter-Speed Analysis

For the second phase of similarity analysis, subjects' synergy outputs were studied across speed conditions in order to determine if there were significant differences in motor modules composition when varying performing speed.

5.3.1 Methods

The first methodological challenge that arose was how to best represent each subjects' synergies for each speed condition from the 10 run output that characterized each movement's dataset. Averaging all 10 runs indiscriminately would result in a synergy set that included synergies from different points of the objective function, which would fail to represent the *most-commonly derived* motor modules of the cvNNMF across the 10 runs. It was thus decided to only include runs of the structural consensus, by eliminating any synergy run whose weights' similarity to the reference were lower than the pre-computed threshold for $\alpha = 0.05$ (Table 5.1) before averaging.

The next step was to compute the similarity metrics on each speeds' synergy results. It was decided to determine similarity metrics from all variable combinations (i.e. Speed 1 vs Speed 2, Speed 1 vs Speed 3 and Speed 1 vs Speed 3 - a total of 3 similarity tests). Therefore, by this point, each of the 10 subjects had 3 variables representing similarity for all paired combinations of speeds. This could be stored in a 10×3 matrix, each cell encoding similarity for a pair of synergies, where the statistical tests were performed. Because there were three similarity metrics (*cosSim*, r and *lag*), there were three such matrices for each movement and synergy index combination.

Similarly to the Inter-Run Analysis, a functional sorting step was necessary to "align" synergy indexes so as to allow comparison between weights and activations responsible for similar functions, and to avoid low similarity results due to mismatched synergies rather than actual structural or timing differences. The arbitrary reference motor modules were set to each subjects' slow speed condition synergies, and both remaining synergy sets (medium and high speed) were sorted by maximum *cosSim* to this benchmark. After sorting, the computed metrics were ready to be evaluated through Repeated Measures Statistical tests as laid out on Fig. (5.3 (B)).

5.3.2 Results

Test statistics ($\tilde{\chi}^2$ distributions) and p-values (p) for the Friedman tests' analysis of variance are shown in Table 5.3 for all three metrics. Three synergy weights across two movements were revealed to contain

Table 5.3: Inter-Speed Friedman tests results. Chi-square ($\tilde{\chi}^2$) distributions and corresponding p-values are reported for each movement-synergy pair. Bold values indicate Friedman tests where the null hypothesis was rejected ($p < 0.05$)

		cosSim of W's		r of H's		*lag* of H's	
Movement	Synergy Index	$\tilde{\chi}^2$	p	$\tilde{\chi}^2$	p	$\tilde{\chi}^2$	p
Knee Extension	1	1.4	0.497	4.2	0.122	2	0.368
Knee Flexion	1	1.8	0.407	5	0.082	5.636	0.06
	2	**7.4**	**0.025**	2.4	0.301	2	0.368
Sit-to-Stand	1	1.4	0.497	1.8	0.407	0.359	0.836
	2	2.6	0.273	2.4	0.301	0	1
Stand-to-sit	1	1.4	0.497	2.6	0.273	2.526	0.283
	2	2.4	0.301	1.4	0.497	3.677	0.159
Gait	1	**12.8**	**0.002**	2.4	0.301	5.733	0.057
	2	**10.4**	**0.006**	3.2	0.202	1.714	0.424

Table 5.4: Wilcoxon Paired Signed-Ranks test results for all speed combinations of weights' *cosSim* from selected synergies where similarity across speeds was statistically different ($p < 0.05$). Test-statistic Z and corresponding Bonferroni-adjusted p-values are reported for each speed combination. The Bonferroni adjustment was applied by multiplying the unadjusted p-value by the number of comparisons. Bold values indicate tests where the null hypothesis was rejected ($p < 0.05$). *Speed 1:* Slow Speed; *Speed 2:* Medium Speed; *Speed 3:* Fast Speed

		Speed1/Speed2 vs. Speed1/Speed3		Speed1/Speed3 vs. Speed2/Speed3		Speed1/Speed2 vs. Speed2/Speed3	
Movement	Synergy Index	Z	p	Z	p	Z	p
Knee Flexion	2	-2.293	0.059	**-2.497**	**0.029**	-0.663	1.000
Gait	1	-1.784	0.252	**-2.395**	**0.042**	-2.09	0.111
Gait	2	-1.274	0.696	-2.09	0.111	-2.191	0.081

significant differences between speeds ($p < 0.05$). These were the second synergies of the Knee Flexion movement and both synergies of the walking trials. All three were therefore selected for *post-hoc* study to investigate which speeds were responsible for these differences. Results of this analysis are summed up in Table 5.4. The second synergy of the walking motor modules had no significant pairwise differences ($p < 0.05$) despite being flagged by the Friedman test. For the remaining two (synergy 2 of Knee Flexion and synergy 1 of the Gait set), significant differences were reported on both for the pair Slow Speed/Fast Speed vs. Medium Speed/Fast Speed ($p < 0.05$), indicating that the motor modules corresponding to the fast movement were dissimilar to the ones from the remaining speeds. Spatial and temporal structure for these two synergy sets, along with synergies from the other speeds may be consulted in Fig. 5.7 (for

the Knee Flexion trials) and in Fig. 5.8 (walking trials). The remaining synergy sets are present in Fig. 5.9 (Knee Extension), 5.10 (Sit-to-Stand), 5.11 (Stand-to-Sit).

5.3.3 Discussion of results

Most movements had no significant differences across speeds in all three measured metrics of weight similarity, activation shape and phase shift, indicating that the same set of synergies could represent muscle activations for all of the tested velocities.

The pairs that did reject the Wilcoxon test's null hypothesis (highlighted in bold in Table 5.4) were the same for both synergies: Speed 1/Speed 3 vs. Speed 2/Speed 3 ($p = 0.029$ for synergy 2 of Knee Flexion, $p = 0.042$ for synergy 1 of walking synergies). These findings suggest that the set of synergies for speed 3 (fast movement) were "dissimilar" set, because the common element for the pair of pairwise similarities (Speed 1/Speed 3 vs. Speed2/Speed3) was speed 3; in other words, it was speed 3 the one where mean similarity was significantly different to both speed 1 and speed 2.

Focusing on the Knee Flexion synergies, it is clear that the dissimilarity for the fast speed condition lies with the recruitment of GM during the flexion of the knee (as evidenced by the red-blue bar plots in Fig. 5.7). Since the participants were not instructed to minimize their foot movement during knee flexion, some subjects not only used their TA to dorsiflex the ankle, but also unilaterally activated their GM to plantarflex the ankle and maintain the foot perpendicular to the shank. Although this requires lower contraction for the slower 5-second (slow condition) and 3-second (medium speed condition) movements, a rapid and intense contraction of the GM is required to maintain an immobile foot when fully flexing the knee in just one second, as was the case for the fast condition. This results in an increment in that muscle's relative activation for a synergy's weights in order to account for the increased sEMG amplitudes, maximizing its dissimilarity to the other "softer" speed condition trials. This hypothesis is supported by the muscle activations shown in Fig. 3.11 (A): GM sEMG envelope activity is notoriously higher for the "Fast" speed condition in Knee Flexion. Therefore, it was determined that, for the present dataset, synergies extracted on Knee Flexion performed rapidly were significantly different that those extracted in slower speed conditions ($p < 0.05$).

Regarding synergy 1 of the walking modules for the fast condition (highlighted with a gray background in Fig. 5.8), a more in depth analysis at the dissimilar synergy reveals that the lowest $cosSim$ to the slow speed outputs was 0.95, recorded for subject 3. In fact, mean similarity of "fast" to "slow" synergies was 0.991 ± 0.016, compared to 0.995 ± 0.010 for "medium" to "slow" modules. In other words, despite the signaling of significant differences, the similarity across speeds is consistent. The disagreement for

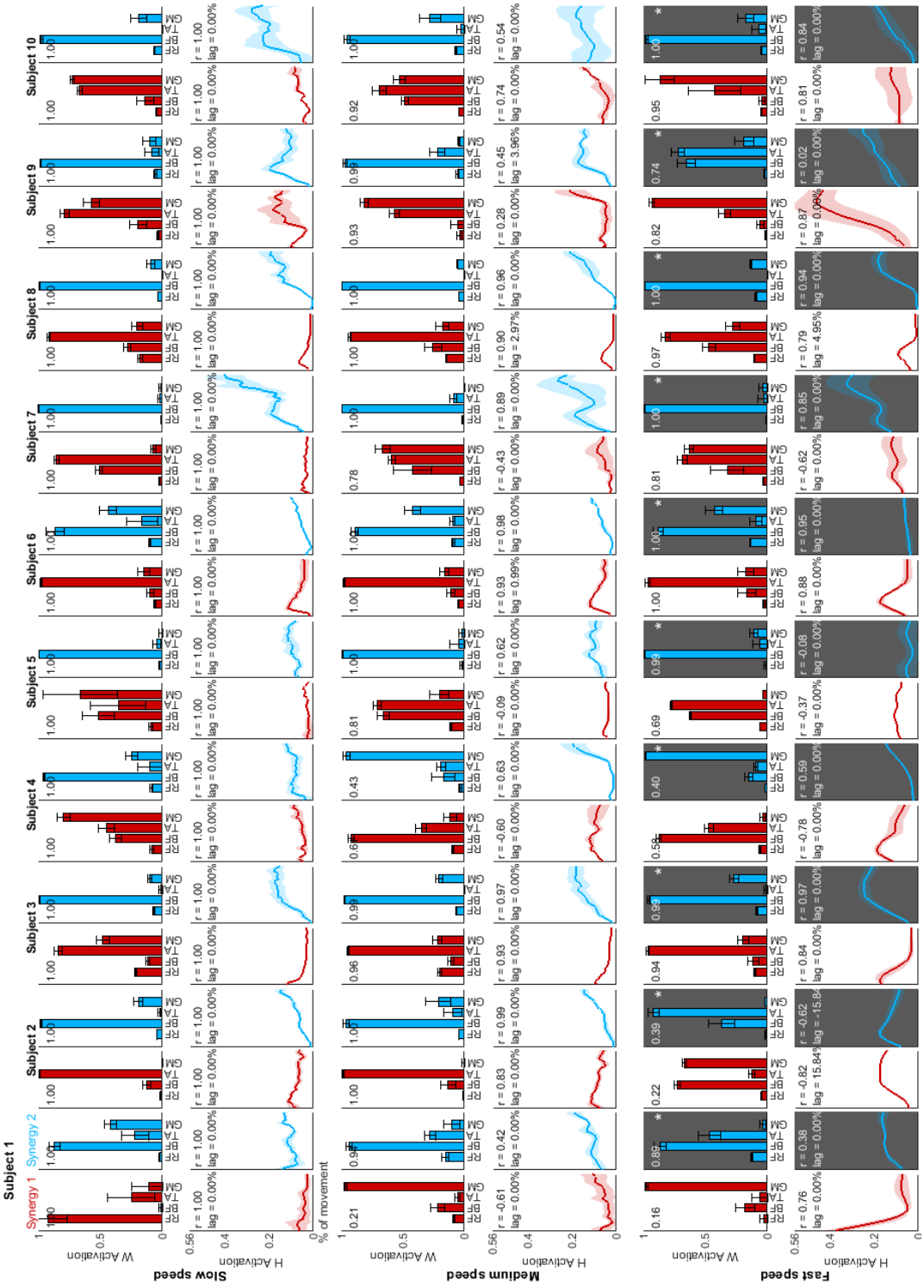


Figure 5.7: Knee Flexion synergies. Each pair of rows illustrates motor modules of a movement performed at a given speed for all 10 participants, and each pair of columns (aligned by color) the resulting synergies for all three speed conditions for one subject. Because synergies were averaged across similar runs, the colored bar plots and error bars represent means and standard deviations of the resulting weights, respectively, and the value on the top left shows the $cosSim$ to the corresponding synergy weights for a slow movement of the same subject. The colored thick lines and surrounding shaded areas illustrate the activation curve means and standard deviations, respectively, with the two inset values showing the Pearson correlation (r) and lag between that waveform and the corresponding curve for the slow condition of the same subject. A gray background represents motor modules that were found to be significantly different from synergies of the remaining speed conditions ($p < 0.05$)

.

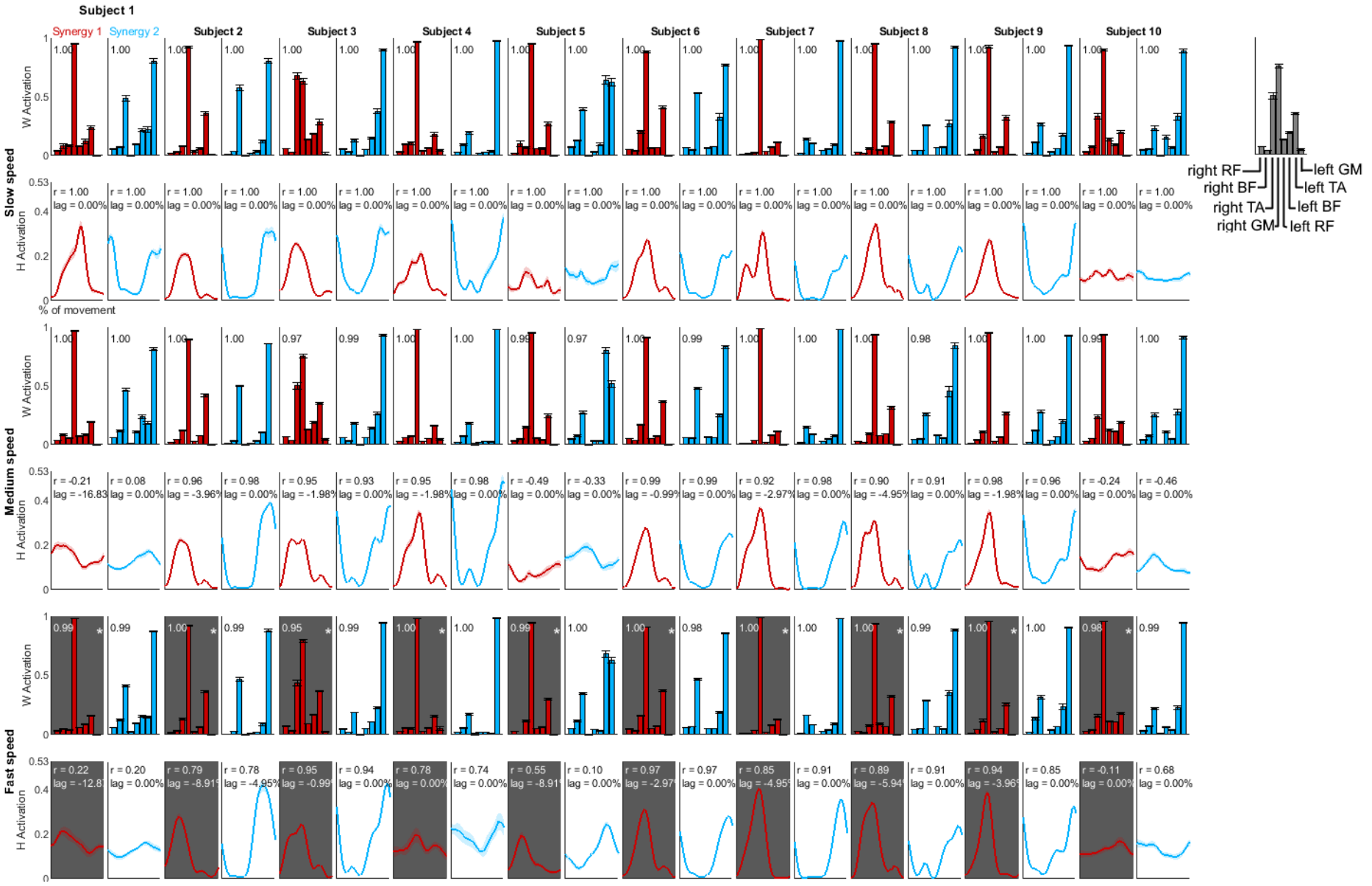


Figure 5.8: Walking synergies. Each pair of rows illustrates motor modules of a movement performed at a given speed for all 10 participants, and each pair of columns (aligned by color) the resulting synergies for all three speed conditions for one subject. Because synergies were averaged across similar runs, the colored bar plots and error bars represent means and standard deviations of the resulting weights, respectively, and the value on the top left shows the *cosSim* to the corresponding synergy weights for a slow movement of the same subject. The colored thick lines and surrounding shaded areas illustrate the activation curve means and standard deviations, respectively, with the two inset values showing the Pearson correlation (r) and lag between that waveform and the corresponding curve for the slow condition of the same subject. A dark-gray background represents motor modules that were found to be significantly different from synergies of the remaining speed conditions ($p < 0.05$)

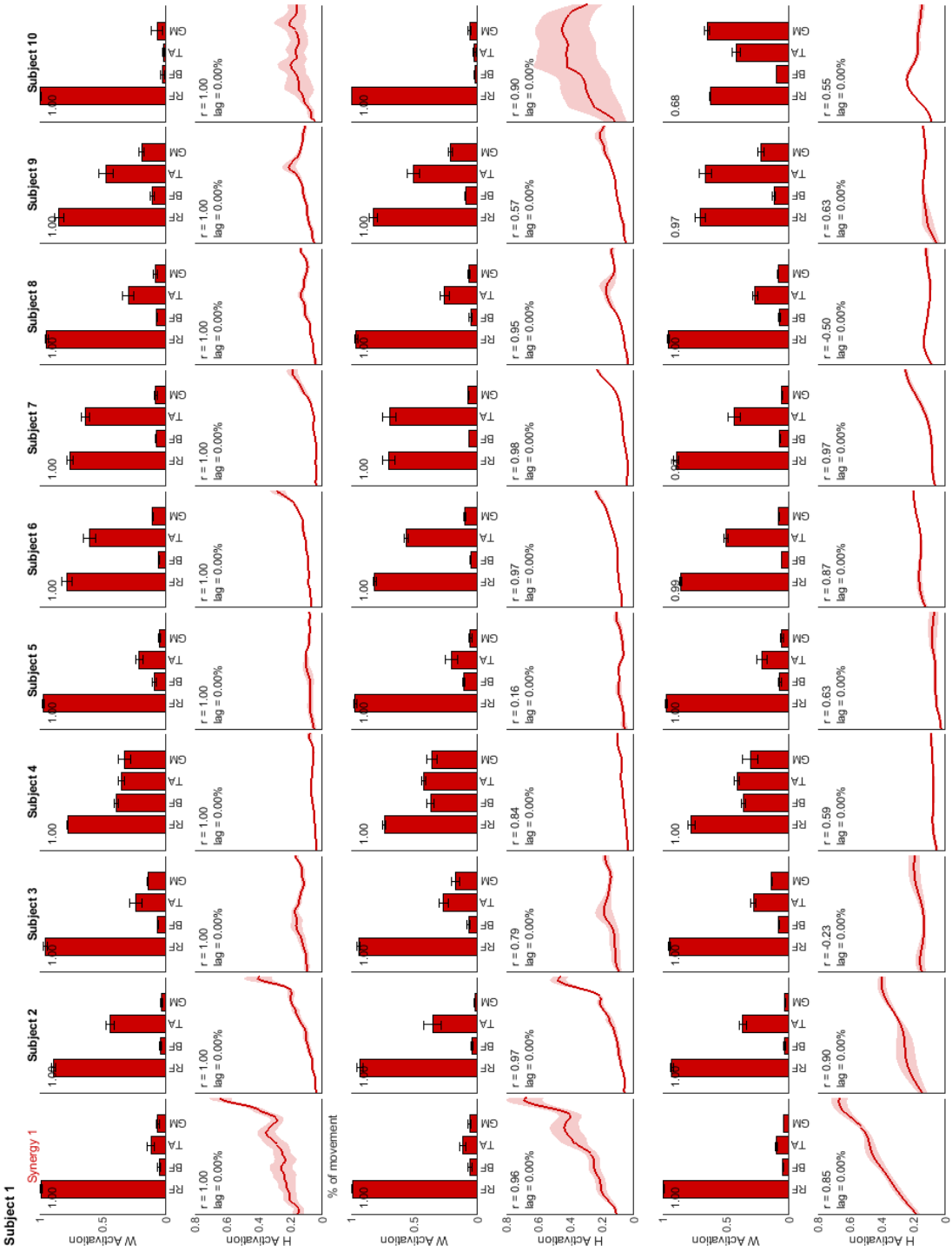


Figure 5.9: Knee Extension synergies. Each row illustrates motor modules of a movement performed at a given speed for all 10 participants, and each column the resulting synergies for all three speed conditions for one subject. Because synergies were averaged across similar runs, the colored bar plots and error bars represent means and standard deviations of the resulting weights, respectively, and the value on the top left shows the $cosSim$ to the corresponding synergy weights for a slow movement of the same subject. The colored thick lines and surrounding shaded areas illustrate the activation curve means and standard deviations, respectively, with the two inset values showing the Pearson correlation (r) and lag between that waveform and the corresponding curve for the slow condition of the same subject.

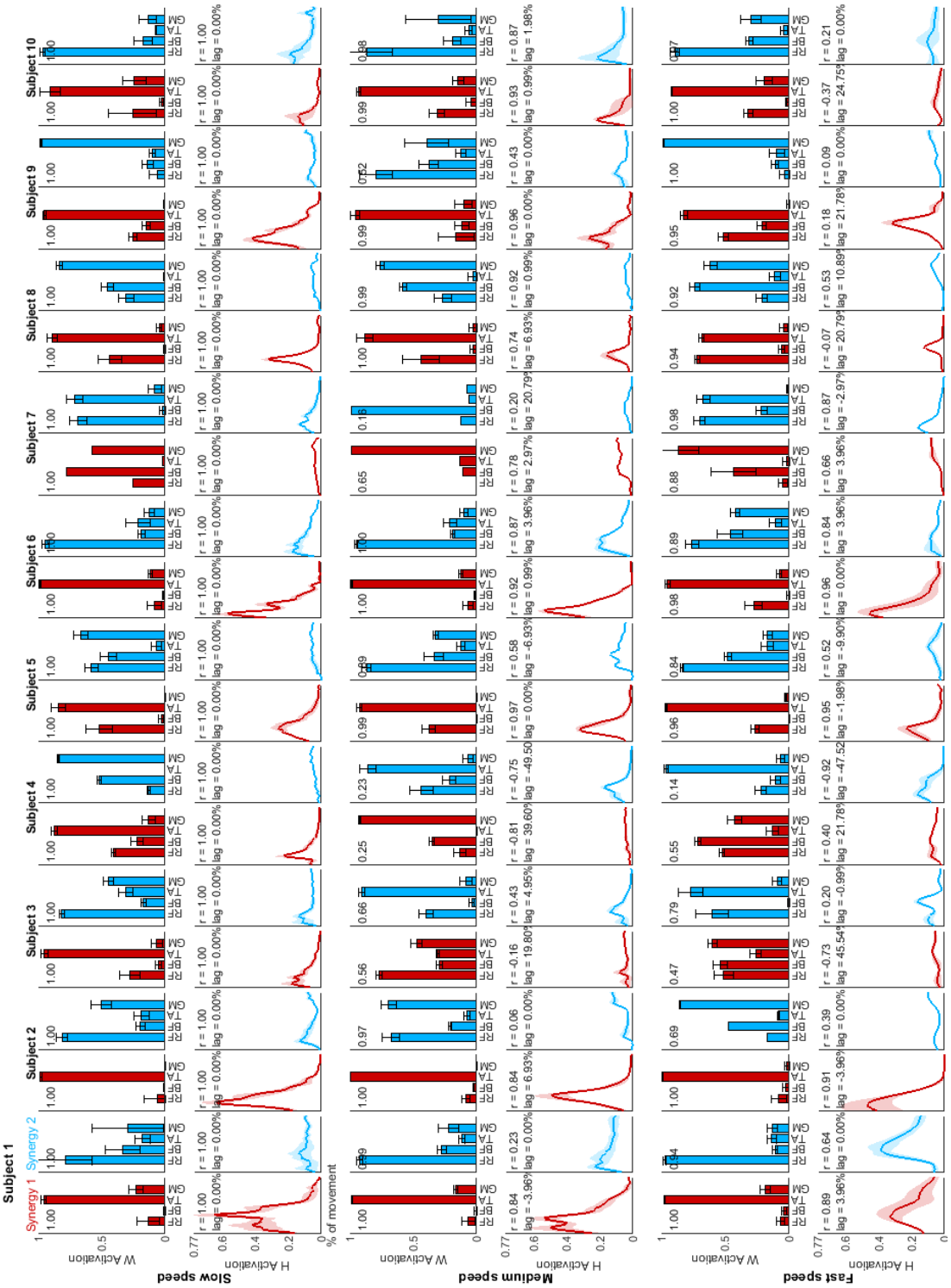


Figure 5.10: Sit-to-Stand synergies. Each pair of rows illustrates motor modules of a movement performed at a given speed for all 10 participants, and each pair of columns (aligned by color) the resulting synergies for all three speed conditions for one subject. Because synergies were averaged across similar runs, the colored bar plots and error bars represent means and standard deviations of the resulting weights, respectively, and the value on the top left shows the $cosSim$ to the corresponding synergy weights for a slow movement of the same subject. The colored thick lines and surrounding shaded areas illustrate the activation curve means and standard deviations, respectively, with the two inset values showing the Pearson correlation (r) and lag between that waveform and the corresponding curve for the slow condition of the same subject.

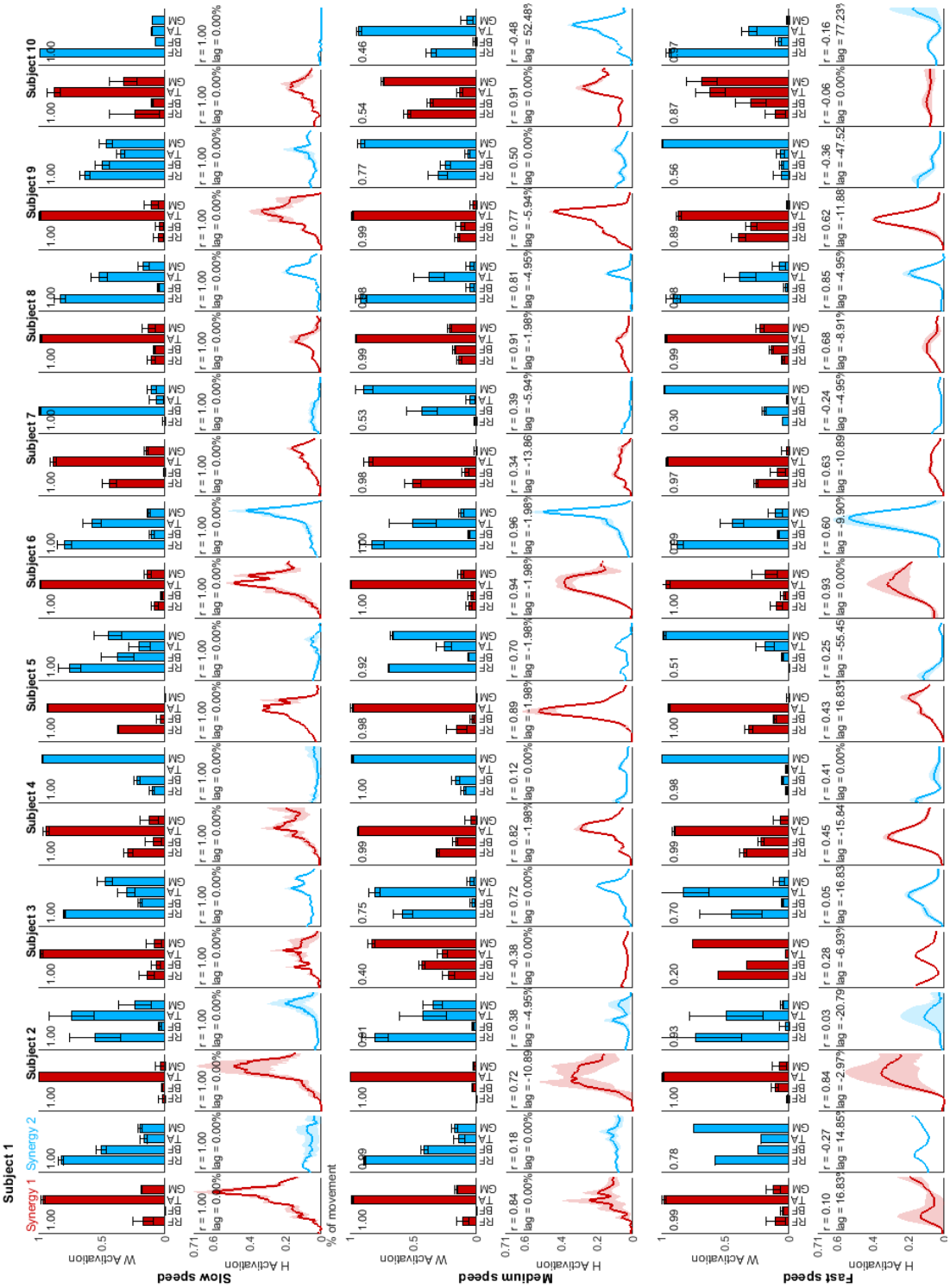


Figure 5.11: Stand-to-Sit synergies. Each pair of rows illustrates motor modules of a movement performed at a given speed for all 10 participants, and each pair of columns (aligned by color) the resulting synergies for all three speed conditions for one subject. Because synergies were averaged across similar runs, the colored bar plots and error bars represent means and standard deviations of the resulting weights, respectively, and the value on the top left shows the $cosSim$ to the corresponding synergy weights for a slow movement of the same subject. The colored thick lines and surrounding shaded areas illustrate the activation curve means and standard deviations, respectively, with the two inset values showing the Pearson correlation (r) and lag between that waveform and the corresponding curve for the slow condition of the same subject.

this particular synergy arises from the fact that synergy structure is so consistently similar across speeds for all participants that subject 3's subtly different synergy structure leads to a small change in mean $cosSim$ across subjects that is sufficient to trigger the rejection the null hypothesis that synergy weights are equal regardless of speed. To circumvent this "false positive" in the problem context, the subject's results should have been removed from the statistical analysis through an outlier detection mechanism.

Previously reported literature examining synergy composition for different walking speeds in healthy participants claimed that the synergies responsible for foot clearance in the mid-swing phase were temporally shifted, with swing initiation occurring later and leg deceleration earlier with decreasing treadmill speed, from 1.0 up to 8.0 km/h. However, that was not the synergy that was found in our analysis. For the dorsiflexor-heavy synergy found in those studies, analogous to the module represented in Fig. 5.8 of this dissertation and thought to correspond to forward propulsion, the activation was invariant both in spatial composition and timing regardless of speed condition [46, 51, 85]. This would indicate that there should be no differences between weights and/or activations across speeds, supporting the hypothesis that the significant differences reported in Table 5.4 for the walking trials were due to very high homogeneity in similarity across speeds, possibly due to the short range of walking speeds tested (1.0-2.0 km/h), not sufficiently different to elicit a different neural control strategy.

In sum, the literature suggests that the statistical framework's findings of significant differences across speeds for synergy 2 of the gait movement are not expected, specially for the small range of low speeds and for the structure of the extracted motor modules. It is then hypothesized that the differences were due to subject 3's outlying weight similarity results originating in an unknown factor, possibly electrode displacement.

5.4 Inter-Subject Analysis

In this section, synergy extraction outputs between subjects will be examined in an attempt to find common synergies between subjects, which would hint that efferent modularity of the CNS exists and is consistent across healthy participants for the studied movements. Repeated measures statistical tests used for the inter-run and inter-speed study are not applicable for this analysis, and hence a different approach will be used to find common motor modules among subjects through statistical inferences.

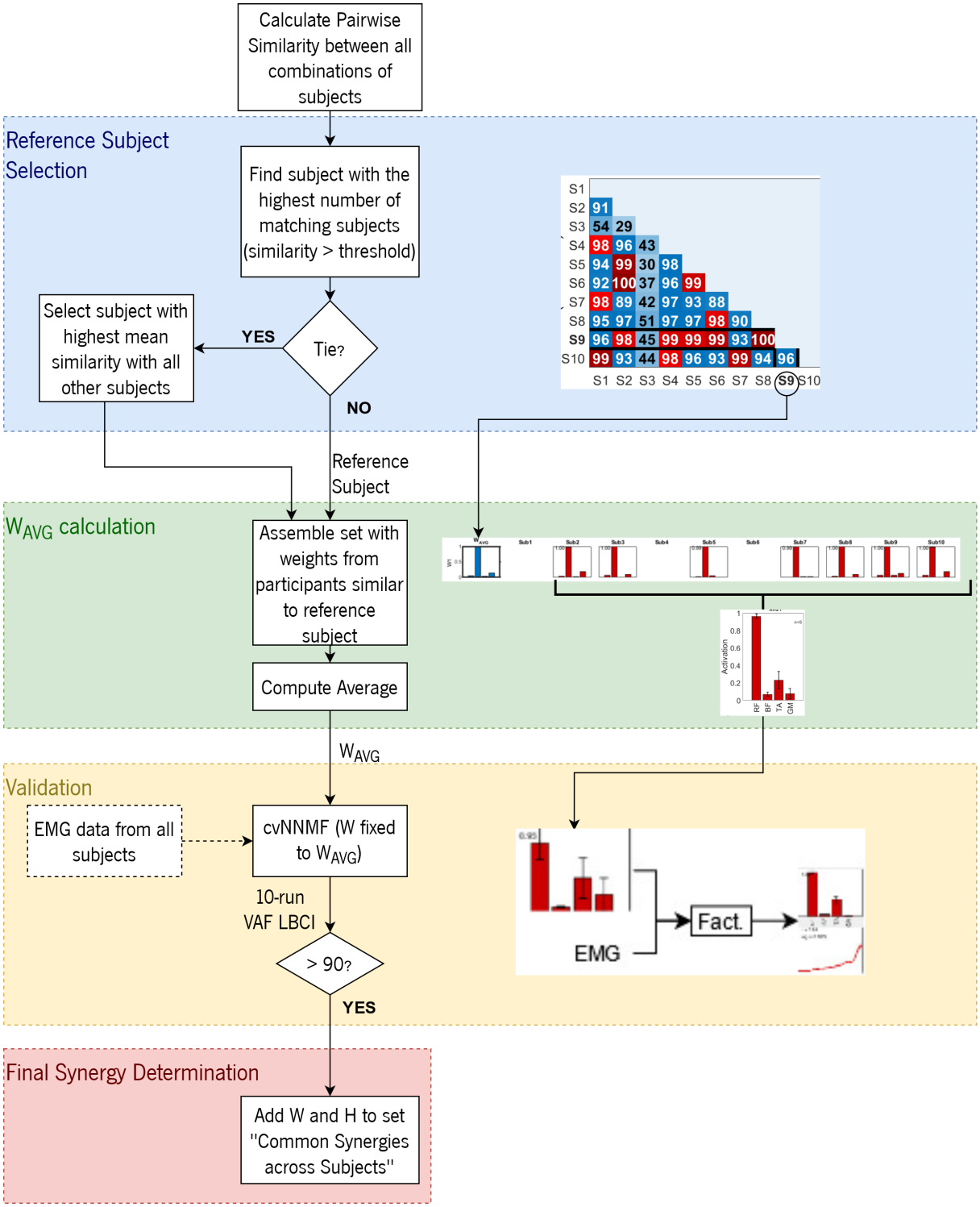


Figure 5.12: Flowchart for the process of determination of most common synergies across subjects.

5.4.1 Methods

The first step made to allow inter-subject comparison was to analyze the weights across subjects so as to investigate if the motor modules' structure were similar across participants. Fig. 5.12 summarizes the structure of this process. This analysis was performed on synergies extracted from medium speed data, with the notion that results for this speed would generalize best for the "neighboring" fast and slow movement speeds.

A functional sorting step was implemented to align motor modules across subjects. Subject 6 was chosen as an arbitrary reference to which the remaining subjects synergies were sorted by maximum cosine similarity. If reordering occurred, the *cosSim* would be recalculated to reflect the correct correspondence.

Because the goal of this analysis was to determine which synergies were most similar across subjects, a pairwise similarity analysis of all available subjects was required. Following the previous assumptions made for the Inter-Run and Inter-Speed study, that structurally different synergies as measured by the *cosSim* of their weights were outright dissimilar because they encoded the recruitment of a different set of muscles, inter-subject study was made exclusively using *cosSim* on subjects' weights. To accomplish this, in each synergy index, the *cosSim* between weights from every subject and every other subject (i.e. Subject 1 vs Subject 2, Subject 1 vs. Subject 3, ..., Subject 2 vs. Subject 3, Subject 2 vs. Subject 4, ..., Subject 9 vs Subject 10) was calculated, resulting in 45 pairwise similarities. The mean value of *cosSim* of each subject with all other participants was also calculated and stored for later reference.

Using the *cosSim* threshold criteria, it was possible to determine participant pairs that had similar weighting vectors with 95% confidence ($p < 0.05$). For each synergy index on each movement, the subject who had the most similarity matches with other subjects (i.e. *cosSim* higher than the thresholds outlined in Table 5.1) was chosen as the *reference* for that synergy. In case of a tie with another participant, the one with highest mean *cosSim* with all other subjects' weights was selected. This process is illustrated in the blue block of Fig. 5.12 - "Reference Subject Calculation".

The need for a "reference participant" arises from the fact that there were no motor modules similar across all subjects, and therefore it was necessary to determine the synergies *most similar* across subjects. The choice of reference, therefore, was based on *a priori* knowledge of which participant would yield the largest set of modules from among the available ones.

Subsequently, using this participants' synergy for a given index as "reference", a group of synergies was assembled following the criteria "weight similarity higher than that expected by chance" ($p < 0.05$). This process resulted in groups composed of weighting vectors sourced from a number of participants ranging from 2 to 10. Afterwards, these groups were averaged in order to obtain one weighting vector

representing the "most similar weights" across subjects - W_{AVG}. This step corresponds to the green box in Fig. 5.12 - "W_{AVG} Calculation".

This analysis is similar to the inter-subject analysis methodology observed in synergy extraction literature [41, 107, 131]; however, this procedure deviates in two ways: first, by not including all participants in the W_{AVG} calculation. This was done because, as has been shown in previous sections, synergy structure differed among subjects. Therefore, it was decided to select a group from within all participants that were already similar among themselves in order to proceed with the inter-subject study. This was the cause of the second alteration to the common inter-subject analysis: the reference subject, rather than being chosen arbitrarily, was selected by examining the similarities between all combinations of participants synergies, with criteria derived from confidence intervals extracted from a distribution of similarity observations.

To confirm the supposition that this process of determining W_{AVG} resulted in synergy weights representative of muscle recruitment of a common inter-subject modular control strategy, a final validation step was implemented, highlighted in yellow in Fig. 5.12. Here, using the original data acquired for each participant, the synergy extraction framework (as depicted in Fig. 4.7 of section 4.3) was ran, but with weights in the NNMF being held fixed to W_{AVG}. As such, only H was allowed to vary: it was initialized with random values and iteratively updated until convergence. The resulting VAF indicates to what degree W_{AVG} was accurate in portraying modular underlying control for that particular subject. In order to quantify overall generalizability of the computed W_{AVG} across participants, a 95% Confidence Interval was built with VAF's calculated from the 10 cvNNMF runs for each subject. By comparing the LBCI for the participants who contributed to the formation of W_{AVG} (in other words, those whose subject-specific weights were classified as "most similar" between themselves) with the LBCI of those who did not, it was possible to ascertain to what extent a neural control strategy was common across subjects.

Additionally, activations from those subjects who met the LBCI>90 criteria following this validation step were used to examine synergy temporal activation for comparison with the kinematic variables. This set of activations in conjunction with the participant-invariant W_{AVG} was considered as the set of final "most common synergies across subjects".

5.4.2 Results

Fig. 5.13 shows the pairwise cosine similarities between every combination of subjects' synergy weights, after sorting. Pairs which demonstrated a higher similarity than that expected by chance ($p < 0.05$) are marked in red. For each synergy index, the reference participant is highlighted in bold. Grouping

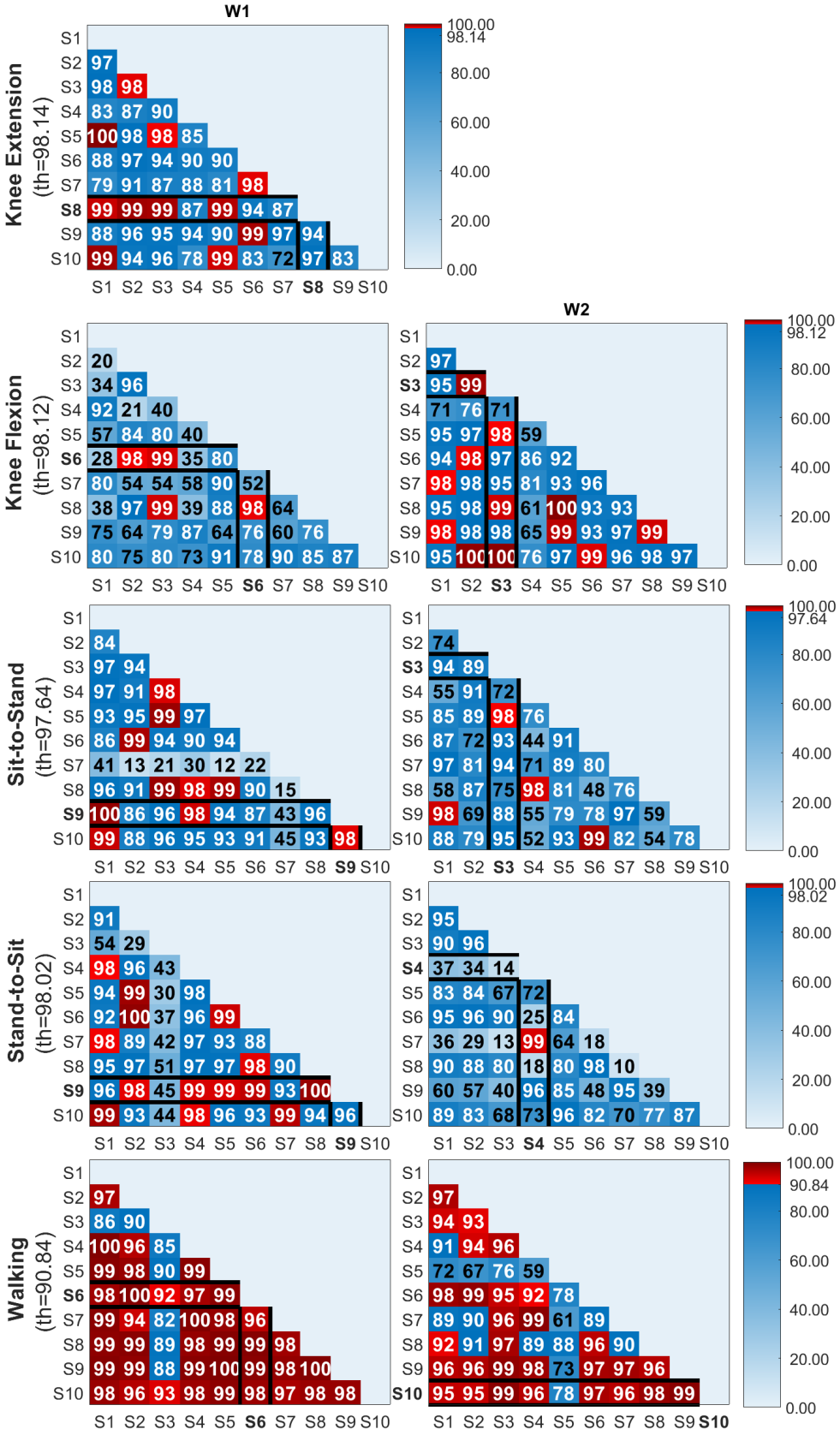


Figure 5.13: Cosine similarities of medium speed synergy weights for all pairwise combinations of participants. Red cells highlight pairs more similar than what's expected by chance ($p < 0.05$). The "most matched" subject, per the criteria explained in section 5.4.1, is highlighted in bold.

by similarity to the reference subject yielded a different number of subjects contributing their weights to W_{AVG} determination, depending on the reference participant's "number of matching subjects". This value is equal to the number of red cells in the highlighted rows/columns of Fig. 5.13 and is also indicated by "n" in Fig. 5.14 (A). This figure also illustrates the corresponding W_{AVG} - the synergy weights averaged across the group of most similar subjects - as bar plots and respective standard deviations as error bars. Fig. 5.14 (B) is a visualization of the reconstruction accuracy on each participant's original data when using these weights. Subjects that were part of the group that resulted in that movement's W_{AVG} are marked by an asterisk. In order to determine if reconstruction quality was higher for this category of participants, average 10-run LBCI was computed for two sets of subjects: those contributing to the W_{AVG} computation ("In" bar in Fig. 5.15) and those not contributing ("Out" bar in Fig. 5.15). For reference, the LBCI for the original cvNNMF framework allowing both factors to freely vary is also denoted in the red bar ("Original"). Finally, 5.16 illustrates the average activations H_1 and H_2 produced by the cvNNMF by fixing weights to W_{AVG} when reconstruction accuracy for a participant across 10 runs was adequate according to the initial criteria (LBCI >90). Synchronized sagittal joint angles for the relevant lower limbs are also plotted to allow comparison with kinematic variables. The remainder of this section summarizes the results for the inter-subject study in a synergy-by-synergy basis, outlining the structure of the common motor modules found for each synergy for all five movements, the number of participants who contributed to its determination and the ability of these synergies to extend to all subjects, as measured by the reconstruction accuracy found during the validation step.

Knee Extension's single averaged synergy was sourced from 5 participants out of a total of 10, meaning that half of the subjects had different motor modules ($p < 0.05$) from the common modular structure as determined by the present analysis. The five contributing subjects were averaged into the vector depicted in Fig. 5.14 (A), encoding a dominating recruitment of the RF (0.96 $\pm$ 0.03 relative activation) with additional lighter selection of the TA (0.23 $\pm$ 0.10). The validation step (for which subject-by-subject results are shown in Fig. 5.14 (B)) revealed a low reconstruction accuracy in data of subjects present for the W_{AVG} calculation and those absent. Only subject 6 of the non-contributing group showed a LBCI > 80%. Averaging results for both groups separately (depicted by bar plots in Fig. 5.15) showed that W_{AVG} contributors had a 25% superior LBCI compared to non-contributors.

The first synergy of the Knee Flexion (W_{AVG1} plot in 5.14 (A)) originated from a set of four subjects, and appeared to be composed almost exclusively of TA recruitment (0.96 $\pm$ 0.02). W_{AVG2}, on the other hand, originated in a consensus formed by five participants and was virtually unanimous in its expression of BF (0.98 $\pm$ 0.02). Therefore, both weighting vectors W_{AVG1} and W_{AVG2} encoded the activation

of one muscle each: TA and BF, respectively. Their combined reconstructing ability (Fig. 5.14 (B)) on the original sEMG data was superior to 90 in 6 subjects, including two that had not contributed their factorization outputs to W_{AVG} determination. As denoted in the respective bar plots of 5.15, LBCI was 91.8 $\pm$ 5.6 % for contributing participants vs. 71.5 $\pm$ 16.2 % for the remainder, a discrepancy of 20.3. The high standard deviation for the "Out" group suggests that results for this analysis was very sparse; that is, the set of W_{AVG} generalized well for muscle recruitment on some of the non-contributing participants and very poorly for others.

Sit-to-Stand's synergy of index 1 (W_{AVG1} in Fig. 5.14 (A)), sourced from four participants, consisted mostly of TA recruitment (0.74 $\pm$ 0.10) with appreciable contribution from the RF muscle (0.39 $\pm$ 0.04). The second synergy, W_{AVG2}, originated from outputs of two participants, and corresponded to RF expression (0.79 $\pm$ 0.01) followed by the GM and BF, both roughly equal in mean magnitude at half that of the RF's (0.40 $\pm$ 0.01 and 0.38 $\pm$ 0.08, respectively). The reconstructing accuracy of this pair of synergies yielded three subjects with LBCI >90, of which two were absent from the computing of W_{AVG}, as illustrated by the three orange cells in Fig. 5.14 (B). Averaging LBCI's across the six contributing subjects revealed a value of 85.2 $\pm$ 7.9 %, 14.5 higher than the analogous characteristic on the set of subjects not involved in W_{AVG} (70.7 $\pm$ 22.9). These discrepancy may be visually inspected in Fig. 5.15.

Like Sit-to-Stand, results for Stand-to-Sit also showed high agreement between subjects of the first synergy, but not of the second. As evidenced by Fig. 5.14 (A), Stand-to-Sit's W_{AVG1}, averaged from six participants, showed a heavy recruitment of TA (0.98 $\pm$ 0.04), with the next most prominent muscle being the RF with 0.15 $\pm$ 0.09 relative expression. On the other hand, W_{AVG2}'s spatial composition was derived from 2 participants and consisted in dominating GM expression (0.93 $\pm$ 0.02) with additional recruitment of the BF (0.28 $\pm$ 0.09). Reconstructing on original data while fixing the weights to this set, showed that four participants met the criteria of having LBCI > 90. One of these (Subject 9) did not cede its weights to W_{AVG} computation (Fig. 5.14). Mean LBCI's for the two group of participants, contributing and non-contributing, yielded a value of 85.4 $\pm$ 14.66 and 43.8 $\pm$ 45.0, respectively, showing a discrepancy of 41.6 between the means. The "Out" group in Fig. 5.15 was notorious for having its standard deviation higher than the mean. This can be pinpointed to the fact that subject 10 had its lower bound of the confidence interval built from the 10-run VAF results equal to -8, indicating very low and very sparse VAF's across runs.

Lastly, inter-subject analysis on walking synergies found all ten subjects sourcing W_{AVG1}, and nine involved in the computing of W_{AVG2}. Spatial structure was symmetrical, as demonstrated in Fig. 5.14 (A): W_{AVG1} corresponded to right side GM recruitment (0.90 $\pm$ 0.06) and W_{AVG2} to the left GM (0.93

$\pm$ 0.07), both sharing additional expression of the TA on the opposite side in the order of 0.15 relative activation. Validating on the original walking dataset revealed that six subjects had a LBCI > 90. This was the only movement for which all subjects contributed to the process of determining W_{AVG}, meaning that all the extracted muscle synergies, except one, were similar regardless of participant ($p < 0.05$). This is why, unlike the remaining movements, it was impossible to divide the participants into two groups by their presence in the group from which W_{AVG} was calculated, as all met the criteria to be included in W_{AVG}; hence the absence of a bar plot in "Out" in Fig. 5.15. This marks a contrast that reappeared regularly for the inter-subject study, between high similarity results for walking synergies and lower measurements for the Discrete Motor Tasks.

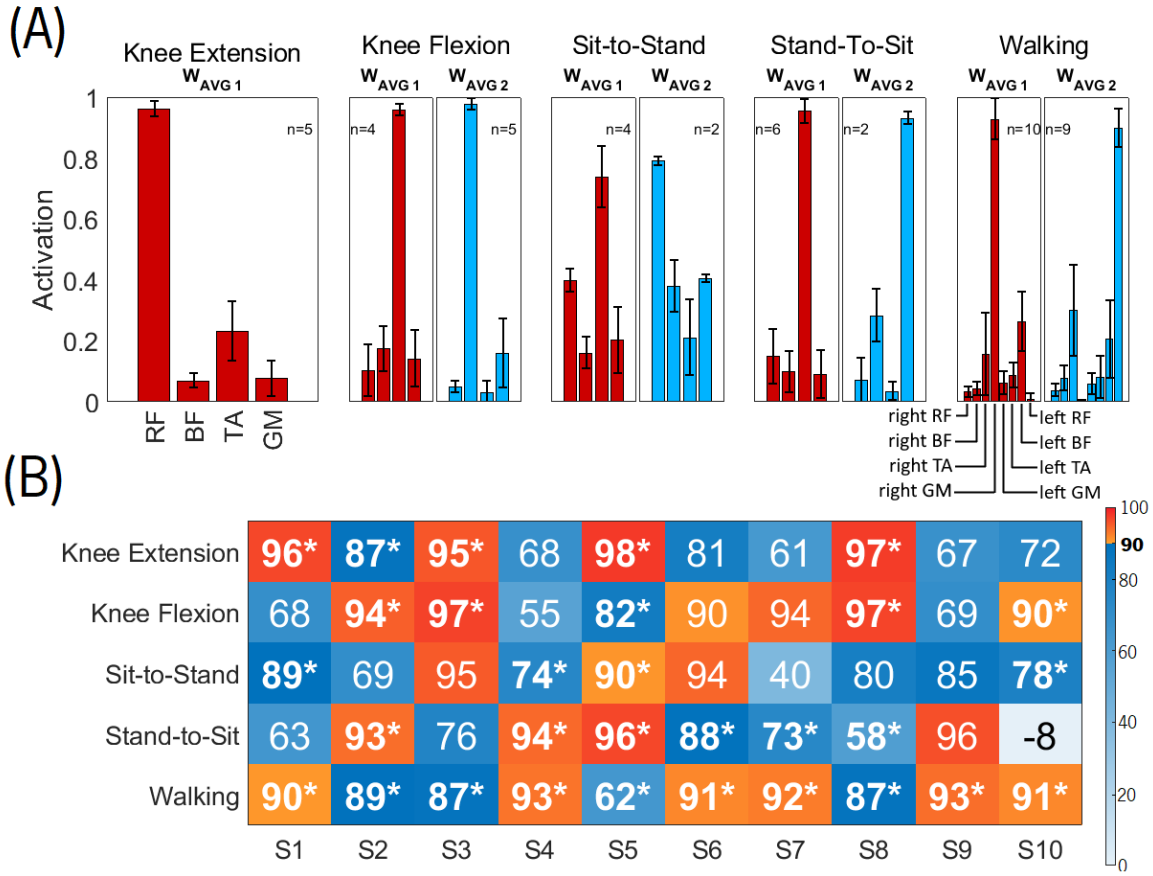


Figure 5.14: Illustration of W_{AVG} and variance results for the re-running of cvNNMF on the original data fixing weights to W_{AVG}. (A) depicts W_{AVG} - the averaged weighting vectors across similar weights - as bar plots and each muscle's respective standard deviation as a black error bar. These were the weights fed on each of the 10 runs to the cvNNMF algorithm, yielding VAF values with which a 95% Confidence Interval was built. The Lower Bounds of these intervals are shown in (B). Values larger than 90 (our original criteria to determine synergy adequateness in reconstructing sEMG data) are highlighted in orange. If a subject's own weights were used for the determination of that movement's W_{AVG}, it is denoted with an asterisk next to its LBCI.

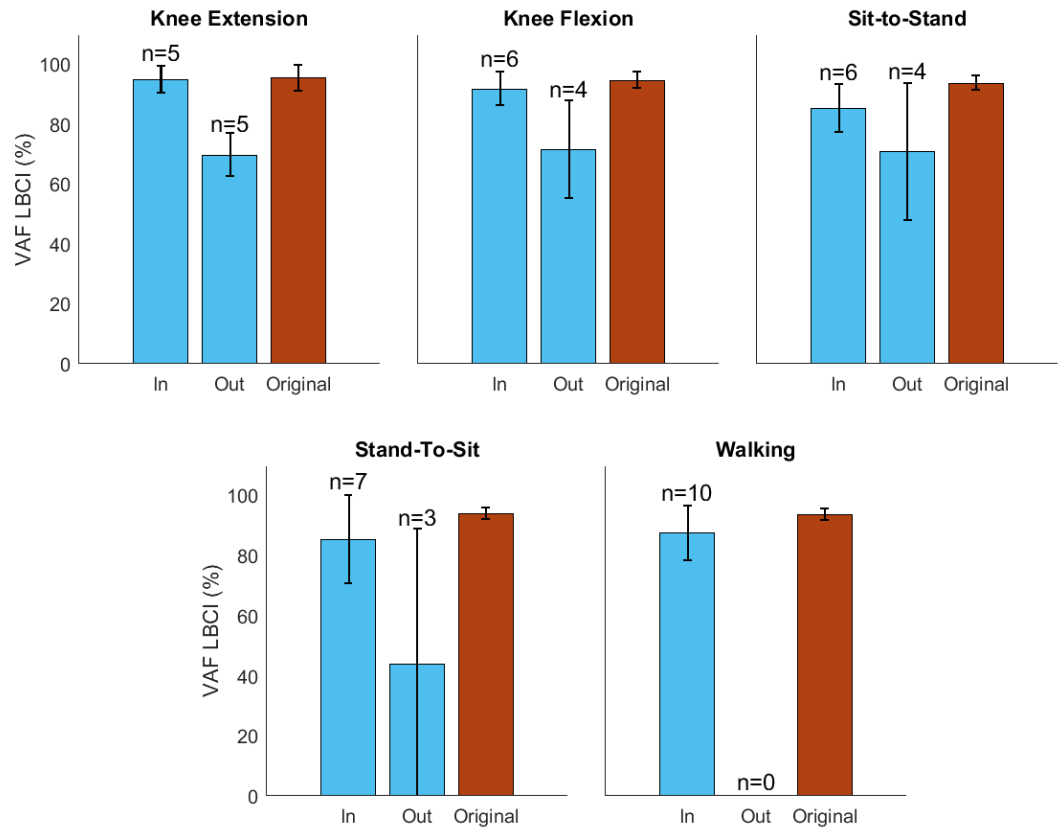


Figure 5.15: Validation Results. Blue bar plots denote average Lower Bounds of the 95% Confidence Interval built from 10 runs of the synergy extraction framework, but constraining weights to be equal W_{AVG}. Results were divided by category of participants: those that contributed to the calculation of W_{AVG} (i.e. those whose weights were more common across subjects) are indicated by the bar "In"; the remaining participants are included in "Out". "n" encodes the number of subjects in each category. The red bar plot "Original" is the LBCI for the original synergy extraction framework and is calculated for all subjects. The black error bars at each bar plot illustrate the standard deviation.

5.4.3 Discussion of Results

For the inter-subject analysis, it was attempted to find common modularity underlying sEMG data for five lower limb movements across subjects. Results from pairwise *cosSim* between all subject permutations (shown in confusion matrix from in Fig. 5.13) appeared to indicate that motor modules were different across subjects for most movements, with the notable exception of the reported results for the walking synergies. This may be explained by the difference in the amounts of data fed to the synergy extraction framework. For instance, the number of concatenated trials used to compute the motor modules for subject 5 in walking at medium speed was 74, while in sit-to-stand the total available trials were only

3. This phenomenon was consistent for all other participants. A previous study examining different gait sEMG envelopes structuring methods concluded that for a concatenated approach, like the one we used for this dissertation, the lower the number of concatenated gait cycles, the poorer the quality of reconstruction [13]. It is therefore reasonable to assume that the smaller data pool from which the Discrete Motor Tasks' synergies were extracted is responsible for overall lower inter-subject similarity and higher variability between synergy extraction outputs.

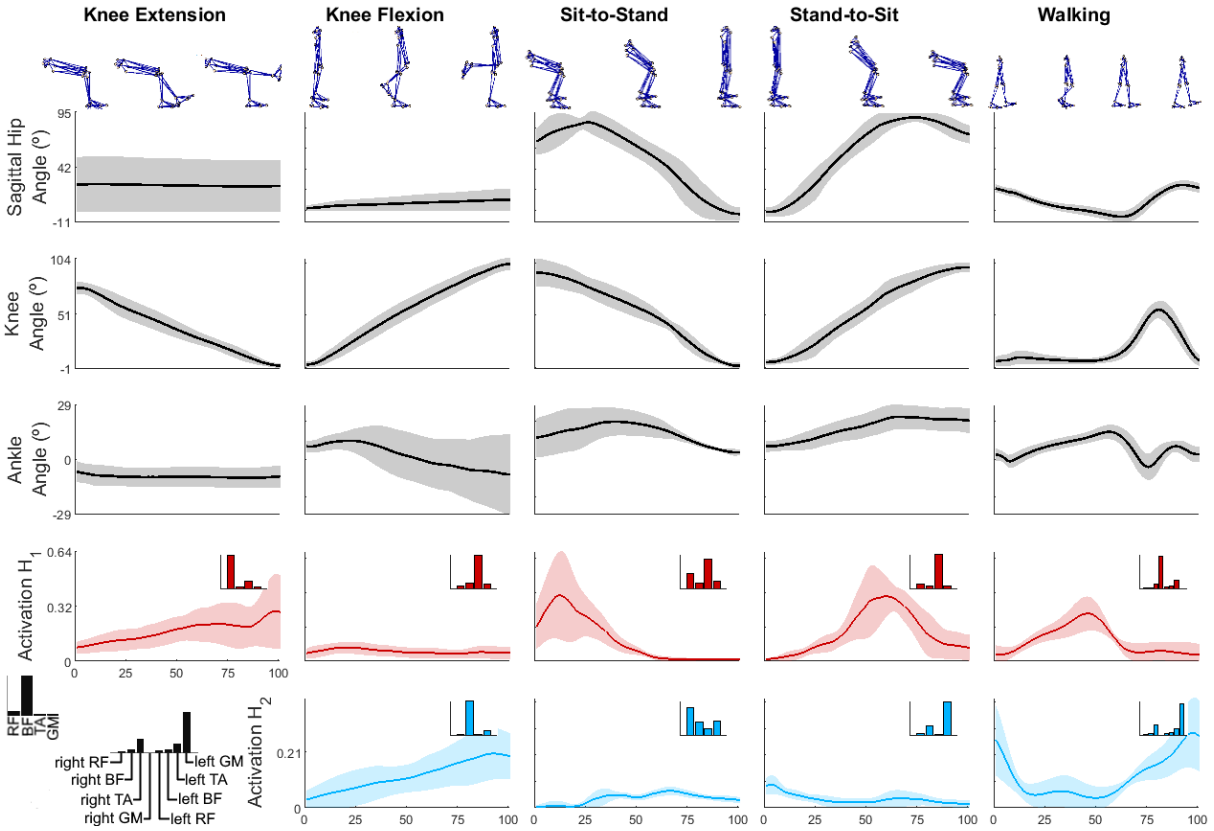


Figure 5.16: Final sets of synergies most common across similar subjects. The first three rows illustrate the mean sagittal hip, knee and ankle joint angles, with gray shaded areas denoting standard deviations at each normalized time point. The last two rows illustrate mean synergy activation with standard deviation represented by the colored shaded area. The corresponding weights W_{AVG} are represented by the inset bar plot.

By examining Fig. 5.14 (A), it is observable that the Knee Extension single synergy consisted in the dominating recruitment of the RF (0.96 $\pm$ 0.03 relative activation) with additional lighter activation of the TA (0.23 $\pm$ 0.10). This might be taken to mean that the participants mostly recruited their quadriceps to produce sufficient torque at the knee joint to elevate the shank, and used their dorsiflexors to maintain the ankle flexed when approaching full extension. This is supported by the Activation of H_1 of the validated synergies in Fig. 5.16: it rises steadily from near non-expression at the very beginning of the movement to 0.3 relative activation, with a short burst in activity in the last 10%. Up until this burst, activation rises in

a proportionally inverse relation to sagittal knee joint angle, corroborating the hypothesis that the synergy encodes quadriceps contraction responsible for shank elevation. Raw sEMG activations (Fig. 3.11 (A) of section 3.3) demonstrates clearly a very similar burst in RF expression for the analogous period in medium speed trials. It is therefore inferrable that this burst is responsible for the "locking" of the knee joint immediately after attainment of full extension and may be responsible for continuing static extension of the shank after the studied period. The fact that the 5 subjects sourced to compute Knee Extension's W_{AVG} all obtained high LBCI (> 87) seems to indicate neural control as ascertained by sEMG factorization did not extend to other subjects for this movement. However, one subject (participant 6) showed a VAF LBCI of 81%; though short of the pre-determined criteria of 90, it is still relatively high value.

Knee Flexion was structurally dominated by the activation of two muscles, one responsible for knee joint torque production and shank displacement, and the other actuating on the ankle joint to maintain it static throughout the movement. As is illustrated in Fig. 5.14 (A), each muscle's activity was clearly divided into two synergies. Analyzing the sagittal ankle angle in Fig. 5.16 provides a possible explanation: the standard deviation for this joint's angle was higher compared to the other movements', and it increased with movement time, which indicates that the participants had varying ankle poses which mostly differed towards the end of the movement. This was due the fact that some subjects chose to dorsiflex the ankle while others avoided it. This introduced large variability in the VAF results: participants who did not dorsiflex the ankle angle would not have needed two synergies to capture the variability of the sEMG data, while does that did, required a second synergy to explain TA expression. This is the reason why Knee Flexion results divided sEMG data in two synergies: one "principal" responsible for shank movement and another "secondary" causing ankle immobility.

Parallels to the kinematic data support the theorized functional roles of each motor module: the first synergy was composed almost exclusively of TA recruitment and was temporally constant across movement time (Fig. 5.16 "Activation H_1"), with a very subtle peak at around 20% of execution, potentially responsible for raising the foot off the ground when initiating movement. The remaining synergy was temporally expressed in a much higher proportion and consisted overwhelmingly of BF activation, rising proportionally to knee angle and thus most likely responsible for joint actuation and shank movement.

Validation LBCIs of VAF values (Fig. 5.14 (B)) for these movements demonstrated a phenomenon that was consistent for the results of the remaining movements: they showed that (i) not every subject who contributed to W_{AVG} attained the highest LBCIs and (ii) some non-contributing participants achieved very high LBCI's. The former observation indicates that in some participants the cvNNMF converged on a set of factors that poorly represented variance on the original sEMG dataset despite the same weights

resulting in good reconstruction on a cross-validated synergy extraction framework on the exact same data. A methodological quirk may be at the root of this phenomenon: a subject "contributing to W_{AVG}" could have contributed by lending *at least* one of its weight vectors ($W1$ or $W2$) to W_{AVG} determination, but not necessarily both, and thus a subject could have had the weighting vector $W1$ similar across subjects, but have $W2$ be completely different from the W_{AVG2} and thus the pair of synergies that reconstructed well on its data was not the same tested for validation, yielding poor results.

The latter observation, that non-contributing subjects nevertheless achieved a high reconstruction metric, may indicate that common modularity underlying neural commands may exist and may be inferred by the sEMG data. Specifically, for Knee Flexion, 6 subjects attained >90 LBCI with the same weighting vectors.

For the Sit-to-Stand task, W_{AVG1}, sourced from four participants, consisted mostly of TA recruitment with appreciable contribution from the RF muscle as evidenced by the bar plots in Fig. 5.14. This weighting vector acts mostly during momentum transfer from sitting to standing (hinted by the presence of TA activity, dorsiflexing the ankle and propelling the body forward) and subsequent rising of the center of mass (caused by the quadriceps of which RF is a member), as has already been discussed in Chapter 4. This hypothesis is corroborated upon inspection of mean activation across participants after validation ("Activation H_1" on Fig. 5.16): there is a clear and noticeable dominance of W_1 expression for the first half of the movement. This early peak is to be expected if it is considered that the TA's relative recruitment in W_{AVG1} is twice that of RF, which means that, on average, at the time of H_1 peak, the dorsiflexors were contracted at double the recruitment potential of the quadriceps as measured by the MVC. From this observation, it may be inferred that the most likely scenario is that the TA is acting on the ankle, dorsiflexing it, propelling the body forward and upward from a rest state and ultimately inducing momentum transfer from sitting to a standing position, while the simultaneously contracting RF is acting to extend the knee providing the upward movement. There is no obvious relationship with the kinematic variables in Fig. 5.16, but the sagittal hip angle peaking right after the peak in H_1 may support these findings because, in a natural standing movement, the hip joint attains its lowest angle to the thighs at the moment of separation from the seat, as the trunk generates upper-body momentum during lift-off [110].

For W_{AVG2} of sit-to-stand, the muscle most relatively activated was RF, followed by GM and BF, both roughly equal in magnitude at half that of the RF's. It may seem counter-intuitive that the agonist-antagonist pair of RF/BF are recruited simultaneously in this synergy, but in fact this phenomenon has been previously documented in human biomechanics where it is known as Lombard's paradox: the fact that, when rising to stand from a sitting position, both the hamstrings and quadriceps are contracted,

despite them being antagonists to each other [137]. This is justified as being caused by the bi-articular nature of the RF and BF: the RF acting over the hip joint has a smaller hip moment arm than the BF; however, the RF moment arm is greater over the knee than the hamstrings' knee moment, meaning that simultaneous contraction of both will result in hip and knee extension, allowing standing [138]. Therefore, if sEMG activity is recorded, both muscles will appear to be contracting simultaneously.

Sit-to-stand also marks the first movement thus far for which synergy extraction results have been reported in the literature. W_{AVG1} found in this dissertation is very similar to the "Momentum Transfer" synergy extracted in [14, 118] (Fig. 2.6 in section 2.2.2), both dominated by TA and RF activity. Likewise, W_{AVG2} seems to be a merger of the remaining two "Extension" and "Posture Stabilization" motor modules, incorporating both quadriceps (of the "Extension" synergy) and GM activity ("Posture Stabilization") plus some expression of the BF, a muscle of the hamstrings which the authors did not measure [14, 118]. Temporal activation in H_1, as illustrated in Fig. 5.16, is also analogous, though the same can not be said of H_2.

Results for the synergy extraction framework on stand-to-sit revealed that W_{AVG1}, averaged across 6 similar subjects, was composed almost exclusively of TA activation. Here it is expected that the TA acts as a stabilizer during the late stages of sitting, acting on the ankle joint to balance the body during descent. W_{AVG2}'s spatial composition, on the other hand, was derived from 2 participants and consisted in dominating GM expression with additional recruitment of the BF. It is presumed that, much like the TA in standing, the GM acts on the ankle joint to initiate movement in sitting, this time by plantarflexing the ankle to begin descent, at which point the thigh muscles, including the BF, "take over" by allowing controlled whole-body descent up until seat contact.

The corresponding validated synergy activations, illustrated in Fig. 5.16, confirm this hypothesis: H_2 peaks at the beginning of the movement, before dropping quickly to a near-zero baseline at around 40% trial time. This peak is coincident with initiation of descent as evidenced by the sagittal knee angle, supporting the assignment of the role of sitting propulsor to W_2. Kinematic analysis also confirms the theorized function of synergy 1, because H_1's peak at 60% movement time coincides with both a decrease in slope of the knee angle, suggesting a reduction in movement speed, and with peak hip flexion, indicating the alignment of the upper limbs with the vertical axis, orthogonal to the seat, marking a phase transition between rapid sitting and a more controlled, "seat-approaching" controlled descent, typical of natural unassisted sitting motion.

There were fewer benchmark studies to evaluate the validity of final W_{AVG} of the stand-to-sit movement. As was discussed in section 4.3.2, the only synergy extraction results reported are for weighted

squatting [52] - i.e. combined sit-to-stand and stand-to-sit while holding a weightlifting bar on the shoulders. Nevertheless, there are some parallels that may be drawn: from the three synergies found underlying sEMG data acquired from 12 muscles in that study, two of them were temporally dominant in the descending portion of the squatting movement (Fig. 2.5 (A) up until 60% normalized squat period). Of these, one was similar to W_{AVG1} both in weighting vectors and activation curve, by encoding high TA activity coupled with weak quadriceps recruitment peaking at the movements' lowest point. The remaining reported synergy was dominated by heavier quadriceps expression and finds no correspondence with our GM-heavy W_{AVG2}, possibly because the GM's role in weighted squatting is absorbed into the other synergies owing to its diminished relative contraction compared with those muscles responsible for knee flexion (i.e., RF, *VMED* and *VLAT* in Fig. 2.5).

Findings from the determination of LBCI's of VAF results for the previous two movements of Sit-to-Stand and Stand-to-Sit seem to indicate that some contributing subjects exhibited low reconstruction ability and some that were not present in the set with which W_{AVG} was computed showed high VAF's, hinting at common inter-subject underlying modularity for these movements.

Lastly, the bilateral study of walking sEMG data found 2 symmetrical synergy weights: W_{AVG1} encoded right GM recruitment and W_{AVG2} left GM, and both shared additional expression of the TA on the opposite side in the order of 0.15 relative activation. There is direct evidence for the existence of this motor module in the synergy extraction literature [13, 46, 85] at a similar speed, where GM presence is usually accompanied by other plantarflexors', such as the Soleus [46], and the Gastrocnemius Lateralis [13]. The validated activations ($H1$ and $H2$ in Fig. 5.16) also match both in shape and phase to the reported results. Because it is mostly active in late stance, at around 40-60% of movement time, it is typically associated with the biomechanical task of forward propulsion, where plantarflexion of the ankle joint produces a ground reaction force which acts to accelerate the body center of mass, propelling the body forward. This is simultaneous with contralateral weight acceptance, because as one side propels the body forward and swings the leg, the other foot rotates over the heel until flat so as to stabilize in anticipation of the swinging leg. Few recent studies, however, examine walking muscle synergies bilaterally, so the role played by the contralateral TA during swing initiation, weighted at about 15% of maximum expression, can only be explained by correspondence with temporal activation spikes of the TA-heavy module reported in the literature's synergies. Though the TA is notorious for being recruited by several synergies during walking [85], there exists a category of synergies where TA heavy expression is omnipresent (one of which is illustrated in Fig. 2.4, M3, of section 2.2.1 of the Literature Review on Muscle Synergies) where it is usually paired with RF and whose temporal activation peaks just before ipsilateral heel strike

[13, 46, 85]. The researchers point out, during heel strike of the foot of the opposing leg, the TA is likely responsible for ankle dorsiflexion in anticipation of upcoming leg loading/weight acceptance during early stance. It is simultaneous with forward propulsion - as one ankle plantarflexes (GM), the other dorsiflexes almost symmetrically (TA) - which is why the activation is uniphasic for both synergies in our results. Furthermore, the kinematic properties support this, as evidenced by a small peak in plantarflexion (the local maximum attained for ankle angle in Fig. 5.16 at around 85% gait cycle time) being simultaneous with the midpoint of temporal activation of synergy 2, which encodes ipsilateral TA recruitment. This clearly correlates contralateral swing initiation with ipsilateral preparation for leg loading. Finally, a 1995 study that did examine bilateral sEMG signals, found a positive correlation between GM and contralateral TA in both factors responsible for left and right forward propulsion [87].

For the walking trials, the gap between the Lower Bounds the Confidence Intervals for the "In" and "Out" groups of participants in Fig. 5.15 could not be calculated, because all subjects met the criteria to be included in W_{AVG} (either through W_1, W_2 or both), hence why there is an absence of a bar plot in "Out" in Fig. 5.15. This demonstrates the high similarity between weights across subjects and helps explain the homogeneous and high LBCI's of VAF values shown in Fig. 5.14 (B). Specifically, LBCI for all was larger than 87%, except for subject 5, who had a very different spatial recruitment of muscles characterized by activation of the TA in addition to the GM in W_2, as is illustrated in Fig. 5.8. This dissertation therefore postulates that the high values of reconstructing ability for the gait results resulted from the large amounts of data available for each factorization run. Moreover, these findings support the hypothesis that a common neural control strategy underlies motor patterns in treadmill walking of healthy subjects, since the same set of muscle weightings was capable of reconstructing the sEMG signal with demonstrable, repeatable, accuracy in all but one of the participants.

Chapter 6

Conclusions

In this dissertation, muscle synergies from five lower-limb movements were extracted and examined using a methodological framework leveraging multi-participant data to make statistical inferences with the ultimate goal of finding common modularity underlying neuromuscular control in healthy subjects. Findings from this work may be used in the future to (i) aid the development of task-oriented control models for the studied movements using the determined muscle synergies; (ii) use the developed synergy extraction framework uniting the scattered methods from the state of the art techniques and (iii) assign statistical confidence to muscle synergy comparisons.

A review performed on the relevant literature in chapter 2 revealed that a modular neural control strategy might underly muscle recruitment during motor tasks and may be ascertained from sEMG data. These muscle synergies have had promising results in several applications, namely control modeling for assistive devices. However, the field was shown to be inconsistent in its methodologies of extracting synergies and lacking in reporting of muscle synergies for lower limb movements of daily life, such as standing and sitting and knee extension and flexion. Additionally, the same review revealed some consensual steps in muscle synergy extraction process, namely that NNMF is the most accurate and widespread factorization algorithm for the purposes of extracting synergies, that a cross-validation step is important in order to circumvent the local minima issue, and that sEMG data should be pre-processed using specific configurations of techniques in order to maximize its potential in extracting useful motor modularity data (Objective 1).

To address the gap in variety of lower-limb movements in the muscle synergy literature, in Chapter 3, a data acquisition protocol was devised aiming to record muscle activities during these seldom studied lower-limb tasks, while also recording data for walking movements (Objective 2). Moreover, all movements

were performed at three different speed conditions, with the expectation that studying data from different speeds would provide further insights into the modular organization of of the CNS input.

Aiming to accomplish Objective 3, in Chapter 4, a synergy extraction framework equipped with a cross-validating system was designed and implemented according to the perceived methodological consensus in the State-of-the-Art. To fine-tune this framework, three parameters left unclear by the literature were studied to investigate the impact on final reconstructing ability: data structuring method prior to factorization, data partitioning ratio for cross-validation and benchmark choice for computing VAF when quantifying reconstructing ability. Results from this study indicated that concatenating trials, partitioning data into two equal-sized groups prior to factorization and using weights from one group coupled with activations from the second to measure reconstruction accuracy against the original signal of the second group were the most accurate configurations to infer neural control strategies from muscle electrical activity data.

Still within the scope of Objective 3, the configured synergy extraction framework was used to determine the minimum number of muscle synergies that adequately explained variance in the participant's sEMG data and therefore underlied each movement's control. The final number of motor modules was one for the Knee Extension movement, and two for the remaining movements, regardless of performing speed.

Chapter 5 examined muscle synergies from a hierarchical approach. It may be concluded with 95% confidence that the synergies did not differ across repeated factorization runs (Objective 4). Furthermore, it may be stated that most of the extracted muscle synergies for the five movements do not differ across speed conditions ($p < 0.05$). Only one module for the Knee Flexion performed in one second (fast condition) showed significant differences from slower movement executions, addressing Objective 5.

Moreover, the inter-subject analysis was designed to reveal the *most common* synergies across subjects despite not all participants reporting similar modules (Objective 5). It determined that for the walking data, almost all of the studied participants converged onto the same similar set of synergies, and though this was not true for the Discrete Motor Tasks, the *most common* inter-subject set was successful in reconstructing the sEMG signals in participants that did not have their synergies determined to be part of that set. Both of these observations point to an inherent modularity in muscle activity data indicating a common control strategy underlying these tasks. The fact that the Discrete Motor Tasks required this extra validation step to study this hypothesis was born of necessity by (i) the small number of sEMG trials available for each subject on each movement (excluding walking) which made results less repeatable by being more sensible to individual trial's variability and (ii) the small number of recorded muscles, making the low-dimensional nature of the input data more likely to produce similar results by chance, elevating

the similarity threshold and cutting off many similar synergies from being classified as such.

From the criteria available for evaluation, such as parallels to kinematic data and to the sparse literature's results, this method was successful in representing common neuromuscular organization despite dataset hindrances. Only one of the determined synergies across subjects (synergy 2 of the Stand-to-Sit movement) was not explained either by available research or through relationship with the kinematic variables (Objective 6).

As far as the literature review appears to indicate, this is the first time a unifying, evidence-based muscle synergy extraction framework and hierarchical study of its outputs leveraging statistical methods is devised and implemented. The processes described in this dissertation could be particularly useful for future research examining movements where there is little available data because sEMG data collection is difficult, tedious and/or tiring.

The research questions posed at the beginning of this dissertation were answered as follows:

RQ 1: Is there a unifying synergy extraction framework able to capture modularity underlying muscle electrical activity in several lower limb tasks?

The literature review indicated that there exists some methodological consensus regarding the sequence of steps to follow in muscle synergy extraction, but methods entailing each of those steps and its parameters are very divergent across publications. In Chapter 4, an evidence-based framework uniting these techniques was implemented and tested on lower-limb sEMG data. In short, this process consists of randomly splitting the available data into equal amounts, concatenating it prior to input into a cross-validated Non-Negative Matrix Factorization procedure, where weights are computed on one group and activations on the other, and where factorization is repeated ten times, including the random partitioning of the data. The outputs from these ten iterations should accurately convey the modularity present in the dataset, inferring the control strategy underlying the movement.

RQ 2: What is the minimal number of synergies that can accurately express each of the studied movements' muscle electrical activity at slow speeds?

This RQ was tackled in Chapter 4. The ascertained number of motor modules, defined as the minimum number of synergies for which the mean Lower Bounds of a 95% Confidence Interval built with VAF values from 10 cross-validated synergy extraction runs across subjects, was one synergy for Knee Extension and two for Knee Flexion, Sit-to-Stand, Stand-to-Sit and Gait data.

RQ 3: Are muscle synergies' structure subject and speed dependent for the studied movements?

This RQ was answered in Chapter 5. Statistical analysis revealed that the extracted muscle syner-

gies for all of the five studied movements did not differ significantly across runs ($p < 0.05$), indicating that a modular representation of the sEMG data existed and was repeatedly and reliably inferred by this framework. Furthermore, the muscle synergies' morphology was invariant across executing velocities for all movements except Knee Flexion, which when performed at a fast speed, (i.e. in 1 second from immobilization up until full flexion) revealed significant differences in its recruiting muscles ($p < 0.05$). Some participants exhibited a personalized controlling strategy for a given lower-limb movement, but the *most common* synergies across subjects were successful in determining a set of motor modules extendable to multiple participants, whose number increased when amount of available data for extracting synergies was larger; hinting at a set of latent underlying muscle synergies controlling each movement.

6.1 Future Work

Future work may include the experimental protocol's optimization by increasing sample size and extending the recording to more trials for Discrete Motor tasks, additional muscles and a broader range of performing speeds. The measuring of a larger set of muscles would be particularly useful because it would confirm whether or not the reporting of only two synergies for the walking data in this dissertation was due to the fact that only four ipsilateral muscles were measured, when the body of literature claims four to six motor modules underlies gait control. If a number of synergies within this range were to be found, it would also help to corroborate results on the Discrete Motor Tasks, for which the literature is scarcer; this is in addition to the direct benefit the recording of additional muscles has on muscle synergy extraction findings. A larger number of trials recorded for each subject would contribute to validate and expand the findings of this dissertation, as the small pool of data available for each subject is thought to have been at the root of the significant dissimilarity among some of the tests' results and poor generalizability for the Discrete Motor Tasks. Additionally, if Knee Flexion trials are replicated, the instructions given to the participants should specify foot positioning in order to avoid the inconsistency in synergy structure found in this work. Greater care in general should be taken to avoid unforeseen activations of muscles not specifically associated with a given movement.

The repeated measures statistical tests developed herein could be particularly interesting in validating synergies extracted from participants with motor impairments. This would allow a very useful assignment of statistical confidence to findings of an healthy vs. pathological muscle synergy structure. Moreover, due to the statistical nature of the framework, it benefits directly from the aggregation of data from additional participants, because the larger the number of subjects and/or trials, the larger the statistical power of

the null hypothesis tests and the more the distribution from which the similarity thresholds are derived approach the reality of the population. Future work could therefore contribute to the findings of this dissertation by simply recording additional subjects performing the same movements. Different additional lower-limb movements, however, would also be very useful for clinical insights and future control models, specially considering the lack of daily life activities' results in the muscle synergy literature.

Lastly, the present findings in muscle synergies could serve many applications. In the domain of control modeling, for instance, it may provide muscle activation patterns for the driving of musculoskeletal models or as a reference for the generation of stimulating patterns for FES assistive devices. In a more clinical domain, these results could also have useful applications, such as complementing a biomechanical analysis or drawing parallels with muscle synergies in subjects with impaired motor coordination, in order to study specific pathologies' characteristics.

Bibliography

[1] G. V. Ostir, I. M. Berges, Y.-F. Kuo, J. S. Goodwin, S. R. Fisher, and J. M. Guralnik, "Mobility Activity and Its Value as a Prognostic Indicator of Survival in Hospitalized Older Adults," *Journal of the American Geriatrics Society*, vol. 61, no. 4, pp. 551–557, 2013. [Online]. Available: https://agsjournals.onlinelibrary.wiley.com/doi/abs/10.1111/jgs.12170

[2] K. Van Naarden Braun, N. Doernberg, L. Schieve, D. Christensen, A. Goodman, and M. Yeargin-Allsopp, "Birth Prevalence of Cerebral Palsy: A Population-Based Study," *Pediatrics*, vol. 137, no. 1, p. e20152872, jan 2016. [Online]. Available: http://pediatrics.aappublications.org/content/137/1/e20152872.abstract

[3] M. P. Lindsay, B. Norrving, R. L. Sacco, M. Brainin, W. Hacke, S. Martins, J. Pandian, and V. Feigin, "World Stroke Organization (WSO): Global Stroke Fact Sheet 2019," *International Journal of Stroke*, vol. 14, no. 8, pp. 806–817, 2019. [Online]. Available: https://doi.org/10.1177/1747493019881353

[4] S. S. Virani, A. Alonso, H. J. Aparicio, E. J. Benjamin, M. S. Bittencourt, C. W. Callaway, A. P. Carson, A. M. Chamberlain, S. Cheng, F. N. Delling, M. S. V. Elkind, K. R. Evenson, J. F. Ferguson, D. K. Gupta, S. S. Khan, B. M. Kissela, K. L. Knutson, C. D. Lee, T. T. Lewis, J. Liu, M. S. Loop, P. L. Lutsey, J. Ma, J. Mackey, S. S. Martin, D. B. Matchar, M. E. Mussolino, S. D. Navaneethan, A. M. Perak, G. A. Roth, Z. Samad, G. M. Satou, E. B. Schroeder, S. H. Shah, C. M. Shay, A. Stokes, L. B. VanWagner, N.-Y. Wang, C. W. Tsao, and N. null, "Heart Disease and Stroke Statistics - 2021 Update," *Circulation*, vol. 143, no. 8, pp. 254–e743, 2021.

[5] A. S. Burns, R. J. Marino, A. E. Flanders, and H. Flett, "Chapter 3 - Clinical diagnosis and prognosis following spinal cord injury," in *Spinal Cord Injury*, ser. Handbook of Clinical Neurology, J. Verhaagen and J. W. McDonald, Eds. Elsevier, 2012, vol. 109, pp. 47–62. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780444521378000036

[6] R. E. Singh, K. Iqbal, G. White, and T. E. Hutchinson, "A systematic review on muscle synergies: From building blocks of motor behavior to a neurorehabilitation tool," *Applied Bionics and Biomechanics*, vol. 2018, 2018.

[7] M. C. Tresch and A. Jarc, "The case for and against muscle synergies," *Current Opinion in Neurobiology*, vol. 19, no. 6, pp. 601–607, 2009.

[8] E. Ambrosini, M. Parati, E. Peri, C. De Marchis, C. Nava, A. Pedrocchi, G. Ferriero, and S. Ferrante, "Changes in leg cycling muscle synergies after training augmented by functional electrical stimulation in subacute stroke survivors: a pilot study," *Journal of NeuroEngineering and Rehabilitation*, vol. 17, no. 1, pp. 1–14, 2020.

[9] S. Ferrante, N. C. Bejarano, E. Ambrosini, A. Nardone, A. M. Turcato, M. Monticone, G. Ferrigno, and A. Pedrocchi, "A personalized multi-channel FES controller based on muscle synergies to support gait rehabilitation after stroke," *Frontiers in Neuroscience*, vol. 10, no. SEP, 2016.

[10] R. R. Neptune, D. J. Clark, and S. A. Kautz, "Modular control of human walking: A simulation study," *Journal of Biomechanics*, vol. 42, no. 9, pp. 1282–1287, 2009.

[11] J. L. Allen, S. A. Kautz, and R. R. Neptune, "The influence of merged muscle excitation modules on post-stroke hemiparetic walking performance," *Clinical Biomechanics*, vol. 28, no. 6, pp. 697–704, 2013. [Online]. Available: http://dx.doi.org/10.1016/j.clinbiomech.2013.06.003

[12] M. Sartori, L. Gizzi, D. G. Lloyd, and D. Farina, "A musculoskeletal model of human locomotion driven by a low dimensional set of impulsive excitation primitives," *Frontiers in Computational Neuroscience*, vol. 7, no. JUN, pp. 1–22, 2013.

[13] A. S. Oliveira, L. Gizzi, D. Farina, and U. G. Kersting, "Motor modules of human locomotion: Influence of EMG averaging, concatenation, and number of step cycles," *Frontiers in Human Neuroscience*, vol. 8, no. MAY, pp. 1–9, 2014.

[14] H. Hanawa, K. Kubota, T. Kokubun, T. Marumo, F. Hoshi, A. Kobayashi, and N. Kanemura, "Muscle synergies underlying sit-to-stand tasks in elderly people and their relationship with kinetic characteristics," *Journal of Electromyography and Kinesiology*, vol. 37, no. August, pp. 15–20, 2017. [Online]. Available: http://dx.doi.org/10.1016/j.jelekin.2017.08.004

[15] M. Ghislieri, V. Agostini, and M. Knaflitz, "How to Improve Robustness in Muscle Synergy Extraction," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 1525–1528, 2019.

[16] C. S. Sherrington, "Flexion-reflex of the limb, crossed extension-reflex, and reflex stepping and standing," *The Journal of Physiology*, vol. 40, no. 1-2, pp. 28–121, 1910.

[17] E. Bizzi, A. D'Avella, P. Saltiel, and M. Tresch, "Modular organization of spinal motor systems," *Neuroscientist*, vol. 8, no. 5, pp. 437–442, 2002.

[18] R. Gentner and J. Classen, "Modular Organization of Finger Movements by the Human Central Nervous System," *Neuron*, vol. 52, no. 4, pp. 731–742, 2006.

[19] S. Grillner, "Neurobiological bases of rhythmic motor acts in vertebrates," *Science*, vol. 228, no. 4696, pp. 143–149, 1985. [Online]. Available: https://science.sciencemag.org/content/228/4696/143

[20] W. A. Lee, "Neuromotor synergies as a basis for coordinated intentional action," *Journal of Motor Behavior*, vol. 16, no. 2, pp. 135–170, 1984.

[21] J. M. Macpherson, "How flexible are muscle synergies?" *Motor control: concepts and issues*, vol. 14, pp. 33–47, 1991.

[22] D. A. McCrea and I. A. Rybak, "Organization of mammalian locomotor rhythm and pattern generation," *Brain Research Reviews*, vol. 57, no. 1, pp. 134–146, 2008.

[23] L. H. Ting, "Dimensional reduction in sensorimotor systems: a framework for understanding muscle coordination of posture," *Progress in Brain Research*, vol. 165, pp. 299–321, 2007.

[24] L. H. Ting and J. L. McKay, "Neuromechanics of muscle synergies for posture and movement," *Current Opinion in Neurobiology*, vol. 17, no. 6, pp. 622–628, 2007.

[25] M. C. Tresch, P. Saltiel, A. D'Avella, and E. Bizzi, "Coordination and localization in spinal motor systems," *Brain Research Reviews*, vol. 40, no. 1-3, pp. 66–79, 2002.

[26] S. Giszter, V. Patil, and C. Hart, "Primitives, premotor drives, and pattern generation: a combined computational and neuroethological perspective," *Progress in Brain Research*, vol. 165, pp. 323–346, 2007.

[27] S. A. Chvatal, G. Torres-Oviedo, S. A. Safavynia, and L. H. Ting, "Common muscle synergies for control of center of mass and force in nonstepping and stepping postural behaviors," *Journal of Neurophysiology*, vol. 106, no. 2, pp. 999–1015, 2011.

[28] N. Bernstein, "The Coordination And Regulation Of Movements," New York, New York, USA, p. 191, 1967. [Online]. Available: https://archive.org/details/bernsteinthecoordinationandregulationofmovements/page/n4/mode/2up

[29] H. Barlow, "Optimal Unsupervised Motor Learning for Dimensionality Reduction of Nonlinear Control Systems," *Neural Computation*, vol. 1, no. 3, pp. 295–311, 1989.

[30] M. Chhabra and R. A. Jacobs, "Properties of synergies arising from a theory of optimal motor behavior," *Neural Computation*, vol. 18, no. 10, pp. 2320–2342, 2006.

[31] T. D. Sanger, "Optimal Unsupervised Motor Learning for Dimensionality Reduction of Nonlinear Control Systems," *IEEE Transactions on Neural Networks*, vol. 5, no. 6, pp. 965–973, 1994.

[32] C. B. Hart and S. F. Giszter, "Modular premotor drives and unit bursts as primitives for frog motor behaviors," *Journal of Neuroscience*, vol. 24, no. 22, pp. 5269–5282, 2004.

[33] L. Dipietro, H. I. Krebs, S. E. Fasoli, B. T. Volpe, J. Stein, C. Bever, and N. Hogan, "Changing motor synergies in chronic stroke," *Journal of Neurophysiology*, vol. 98, no. 2, pp. 757–768, 2007.

[34] T. H. Cruz and Y. Y. Dhaher, "Evidence of abnormal lower-limb torque coupling after stroke: An isometric study," *Stroke*, vol. 39, no. 1, pp. 139–147, 2008.

[35] T. Drew, J. Kalaska, and N. Krouchev, "Muscle synergies during locomotion in the cat: A model for motor cortex control," *Journal of Physiology*, vol. 586, no. 5, pp. 1239–1245, 2008.

[36] C. L. Banks, M. M. Pai, T. E. McGuirk, B. J. Fregly, and C. Patten, "Methodological choices in muscle synergy analysis impact differentiation of physiological characteristics following stroke," *Frontiers in Computational Neuroscience*, vol. 11, no. August, pp. 1–12, 2017.

[37] P. Kieliba, P. Tropea, E. Pirondini, M. Coscia, S. Micera, and F. Artoni, "How are Muscle Synergies Affected by Electromyography Pre-Processing?" *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 4, pp. 882–893, 2018.

[38] F. Hug, "Can muscle coordination be precisely studied by surface electromyography?" *Journal of Electromyography and Kinesiology*, vol. 21, no. 1, pp. 1–12, 2011.

[39] A. Erdemir, S. McLean, W. Herzog, and A. J. van den Bogert, "Model-based estimation of muscle forces exerted during movements," *Clinical Biomechanics*, vol. 22, no. 2, pp. 131–154, 2007.

[40] L. Ting and S. Chvatal, *Decomposing Muscle Activity in Motor Tasks: Methods and Interpretation*, 2010, pp. 102–138.

[41] G. Torres-Oviedo and L. H. Ting, "Muscle synergies characterizing human postural responses," *Journal of Neurophysiology*, vol. 98, no. 4, pp. 2144–2156, 2007.

[42] M. F. Rabbi, C. Pizzolato, D. G. Lloyd, C. P. Carty, D. Devaprakash, and L. E. Diamond, "Non-negative matrix factorisation is the most appropriate method for extraction of muscle synergies in walking and running," *Scientific Reports*, vol. 10, no. 1, pp. 1–11, 2020.

[43] "Chapter 5 - Measurement of pelvic floor muscle function and strength, and pelvic organ prolapse," in *Evidence-Based Physical Therapy for the Pelvic Floor (Second Edition)*, second edi ed., K. Bø, B. Berghmans, S. Mørkved, and M. Van Kampen, Eds. Churchill Livingstone, 2015, pp. 43–109. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780702044434000054

[44] A. G. S. Rayo, L. H. H. Gómez, A. T. V. Sánchez, J. A. B. Fernández, J. A. F. Campos, G. U. Calderón, V. M. S. Rayo, and A. E. F. Peñaloza, *Design and Manufacturing of a Dry Electrode for EMG Signals Recording with Microneedles*. Cham: Springer International Publishing, 2018, pp. 259–267. [Online]. Available: https://doi.org/10.1007/978-3-319-59590-0{_}22

[45] G. Piervirgili, F. Petracca, and R. Merletti, "A new method to assess skin treatments for lowering the impedance and noise of individual gelled Ag{\textendash}{AgCl} electrodes," *Physiological Measurement*, vol. 35, no. 10, pp. 2101–2118, sep 2014. [Online]. Available: https://doi.org/10.1088/0967-3334/35/10/2101

[46] D. J. Clark, L. H. Ting, F. E. Zajac, R. R. Neptune, and S. A. Kautz, "Merging of healthy motor modules predicts reduced locomotor performance and muscle coordination complexity post-stroke," *Journal of Neurophysiology*, vol. 103, no. 2, pp. 844–857, 2010.

[47] K. Nishida, S. Hagio, B. Kibushi, T. Moritani, and M. Kouzaki, "Comparison of muscle synergies for running between different foot strike patterns," *PLoS ONE*, vol. 12, no. 2, pp. 1–19, 2017.

[48] A. Santuz, A. Ekizos, L. Janshen, V. Baltzopoulos, and A. Arampatzis, "On the Methodological Implications of Extracting Muscle Synergies from Human Locomotion," *International Journal of Neural Systems*, vol. 27, no. 5, pp. 1–15, 2017.

[49] M. Kristiansen, A. Samani, P. Madeleine, and E. A. Hansen, *Effects of 5 weeks of bench press training on muscle synergies: A randomized controlled study*, 2016, vol. 30, no. 7.

[50] A. S. Oliveira, P. B. Silva, M. E. Lund, U. G. Kersting, and D. Farina, "Fast changes in direction during human locomotion are executed by impulsive activation of motor modules," *Neuroscience*, vol. 228, pp. 283–293, 2013. [Online]. Available: http://dx.doi.org/10.1016/j.neuroscience.2012.10.027

[51] B. Kibushi, S. Hagio, T. Moritani, and M. Kouzaki, "Speed-Dependent Modulation of Muscle Activity Based on Muscle Synergies during Treadmill Walking," *Frontiers in Human Neuroscience*, vol. 12, no. January, pp. 1–13, jan 2018. [Online]. Available: http://journal.frontiersin.org/article/10.3389/fnhum.2018.00004/full

[52] K. B. Smale, M. S. Shourijeh, and D. L. Benoit, "Use of muscle synergies and wavelet transforms to identify fatigue during squatting," *Journal of Electromyography and Kinesiology*, vol. 28, pp. 158–166, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.jelekin.2016.04.008

[53] H. J. Hermens, B. Freriks, C. Disselhorst-Klug, and G. Rau, "Development of recommendations for SEMG sensors and sensor placement procedures," *Journal of Electromyography and Kinesiology*, vol. 10, no. 5, pp. 361–374, 2000. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1050641100000274

[54] A. S. Oliveira, P. B. Silva, M. E. Lund, L. Gizzi, D. Farina, and U. G. Kersting, "Effects of Perturbations to Balance on Neuromechanics of Fast Changes in Direction during Locomotion," *PLoS ONE*, vol. 8, no. 3, pp. 1–13, 2013.

[55] J. L. Allen and R. R. Neptune, "Three-dimensional modular control of human walking," *Journal of Biomechanics*, vol. 45, no. 12, pp. 2157–2163, 2012. [Online]. Available: http://dx.doi.org/10.1016/j.jbiomech.2012.05.037

[56] S. Dorel, A. Couturier, and F. Hug, "Intra-session repeatability of lower limb muscles activation pattern during pedaling," *Journal of Electromyography and Kinesiology*, vol. 18, no. 5, pp. 857–865, 2008.

[57] G. Cappellini, Y. P. Ivanenko, R. E. Poppele, and F. Lacquaniti, "Motor patterns in human walking and running," *Journal of Neurophysiology*, vol. 95, no. 6, pp. 3426–3437, 2006.

[58] D. J. Berger and A. D'Avella, "Effective force control by muscle synergies," *Frontiers in Computational Neuroscience*, vol. 8, no. 1 APR, pp. 1–13, 2014.

[59] S. Israely, G. Leisman, C. C. Machluf, and E. Carmeli, "Muscle synergies control during hand-reaching tasks in multiple directions post-stroke," *Frontiers in Computational Neuroscience*, vol. 12, no. February, 2018.

[60] Y. P. Ivanenko, G. Cappellini, I. A. Solopova, A. A. Grishin, M. J. MacLellan, R. E. Poppele, and F. Lacquaniti, "Plasticity and modular control of locomotor patterns in neurological disorders with motor deficits," *Frontiers in Computational Neuroscience*, vol. 7, no. SEP, pp. 1–11, 2013.

[61] N. Lambert-Shirzad and H. F. Machiel Van Der Loos, "On identifying kinematic and muscle synergies: A comparison of matrix factorization methods using experimental data from the healthy population," *Journal of Neurophysiology*, vol. 117, no. 1, pp. 290–302, 2017.

[62] M. S. Shourijeh, T. E. Flaxman, and D. L. Benoit, "An approach for improving repeatability and reliability of non-negative matrix factorization for muscle synergy analysis," *Journal of Electromyography and Kinesiology*, vol. 26, pp. 36–43, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.jelekin.2015.12.001

[63] "Trigno™ EMG & Additional Sensors," may 2020. [Online]. Available: https://delsys.com/trigno/sensors/

[64] "Ambu® Neuroline 720." [Online]. Available: https://www.ambu.com/neurology/intraoperative-monitoring/product/ambu-neuroline-720

[65] B. R. Shuman, M. H. Schwartz, and K. M. Steele, "Electromyography data processing impacts muscle synergies during gait for unimpaired children and children with cerebral palsy," *Frontiers in Computational Neuroscience*, vol. 11, no. June, pp. 1–9, 2017.

[66] L. J. Myers, M. Lowery, M. O'Malley, C. L. Vaughan, C. Heneghan, A. St. Clair Gibson, Y. X. Harley, and R. Sreenivasan, "Rectification and non-linear pre-processing of EMG signals for cortico-muscular analysis," *Journal of Neuroscience Methods*, vol. 124, no. 2, pp. 157–165, 2003.

[67] R. L. Routson, S. A. Kautz, and R. R. Neptune, "Modular organization across changing task demands in healthy and poststroke gait," *Physiological Reports*, vol. 2, no. 6, pp. 1–14, 2014.

[68] V. C. Cheung, A. D'Avella, M. C. Tresch, and E. Bizzi, "Central and sensory contributions to the activation and organization of muscle synergies during natural motor behaviors," *Journal of Neuroscience*, vol. 25, no. 27, pp. 6419–6434, 2005.

[69] D. Farina, N. Jiang, H. Rehbaum, A. Holobar, B. Graimann, H. Dietl, and O. C. Aszmann, "The extraction of neural information from the surface EMG for the control of upper-limb prostheses: Emerging avenues and challenges," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 4, pp. 797–809, 2014.

[70] S. Muceli, A. T. Boye, A. D'Avella, and D. Farina, "Identifying Representative Synergy Matrices for Describing Muscular Activation Patterns During Multidirectional Reaching in the Horizontal Plane," *Journal of Neurophysiology*, vol. 103, no. 3, pp. 1532–1542, 2010. [Online]. Available: https://doi.org/10.1152/jn.00559.2009

[71] K. M. Steele, A. Rozumalski, and M. H. Schwartz, "Muscle synergies and complexity of neuromuscular control during gait in cerebral palsy," *Developmental Medicine & Child Neurology*, vol. 57, no. 12, pp. 1176–1182, 2015. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/dmcn.12826

[72] L. Gizzi, J. F. Nielsen, F. Felici, Y. P. Ivanenko, and D. Farina, "Impulses of activation but not motor modules are preserved in the locomotion of subacute stroke patients." *Journal of neurophysiology*, vol. 106, no. 1, pp. 202–210, jul 2011.

[73] D. Torricelli, F. Barroso, M. Coscia, C. Alessandro, F. Lunardini, E. Bravo Esteban, and A. D'Avella, *Muscle Synergies in Clinical Practice: Theoretical and Practical Implications*. Cham: Springer International Publishing, 2016, pp. 251–272. [Online]. Available: https://doi.org/10.1007/978-3-319-24901-8{_}10

[74] V. Monaco, A. Ghionzoli, and S. Micera, "Age-related modifications of muscle synergies and spinal cord activity during locomotion," *Journal of Neurophysiology*, vol. 104, no. 4, pp. 2092–2102, 2010.

[75] V. C. Cheung, L. Piron, M. Agostini, S. Silvoni, A. Turolla, and E. Bizzi, "Stability of muscle synergies

for voluntary actions after cortical stroke in humans," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 46, pp. 19 563–19 568, 2009.

[76] V. C. K. Cheung, A. D'Avella, and E. Bizzi, "Adjustments of Motor Pattern for Load Compensation Via Modulated Activations of Muscle Synergies During Natural Behaviors," *Journal of Neurophysiology*, vol. 101, no. 3, pp. 1235–1257, 2009. [Online]. Available: https://doi.org/10.1152/jn.01387.2007

[77] G. Torres-Oviedo, J. M. Macpherson, and L. H. Ting, "Muscle synergy organization is robust across a variety of postural perturbations," *Journal of Neurophysiology*, vol. 96, no. 3, pp. 1530–1546, 2006.

[78] A. Sawers, J. L. Allen, and L. H. Ting, "Long-term training modifies the modular structure and organization of walking balance control," *Journal of Neurophysiology*, vol. 114, no. 6, pp. 3359–3373, 2015. [Online]. Available: https://doi.org/10.1152/jn.00758.2015

[79] M. V. D. Krogt, L. Oudenhoven, A. Buizer, A. Dallmeijer, N. Dominici, and J. Harlaar, "The effect of EMG processing choices on muscle synergies before and after BoNT-A treatment in cerebral palsy," *Gait & Posture*, vol. 49, p. 31, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.gaitpost.2016.07.095

[80] Y. P. Ivanenko, R. Grasso, M. Zago, M. Molinari, G. Scivoletto, V. Castellano, V. Macellari, and F. Lacquaniti, "Temporal Components of the Motor Patterns Expressed by the Human Spinal Cord Reflect Foot Kinematics," *Journal of Neurophysiology*, vol. 90, no. 5, pp. 3555–3565, 2003.

[81] F. Hug, J. M. Drouet, Y. Champoux, A. Couturier, and S. Dorel, "Interindividual variability of electromyographic patterns and pedal force profiles in trained cyclists," *European Journal of Applied Physiology*, vol. 104, no. 4, pp. 667–678, 2008.

[82] M. Halaki and K. Ginn, "Normalization of EMG Signals: To Normalize or Not to Normalize and What to Normalize to?, Computational Intelligence in Electromyography Analysis - A Perspective on Current Applications and Future Challenges," *Computational Intelligence in Electromyography Analysis - A Perspective on Current Applications and Future Challenges*, p. 448, 2012. [Online]. Available: http://dx.doi.org/10.5772/49957{%}0Ahttp://www.intechopen.com/books/computational-intelligence-in-electromyography-analysis-a-perspective-on-current-applications-and-future-challenges/normalization-of-emg-signals-to-normalize-or-not-to-normalize-and-what-to

[83] D. Rimini, V. Agostini, and M. Knaflitz, "Evaluation of Muscle Synergies Stability in Human Locomotion :," pp. 1–5, 2017.

[84] L. Ting and S. Chvatal, *Decomposing Muscle Activity in Motor Tasks: Methods and Interpretation*, 2010, pp. 102–138.

[85] Y. P. Ivanenko, R. E. Poppele, and F. Lacquaniti, "Five basic muscle activation patterns account for muscle activity during human locomotion," *Journal of Physiology*, vol. 556, no. 1, pp. 267–282, 2004.

[86] B. L. Davis and C. L. Vaughan, "Phasic behavior of EMG signals during gait: Use of multivariate statistics," *Journal of Electromyography and Kinesiology*, vol. 3, no. 1, pp. 51–60, 1993.

[87] K. S. Olree and C. L. Vaughan, "Fundamental patterns of bilateral muscle activity in human locomotion," *Biological Cybernetics*, vol. 73, no. 5, pp. 409–414, 1995.

[88] A. Saito, K. Watanabe, and H. Akima, "Coordination among thigh muscles including the vastus intermedius and adductor magnus at different cycling intensities," *Human Movement Science*, vol. 40, pp. 14–23, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167945714002061

[89] L. A. Merkle, C. S. Layne, J. J. Bloomberg, and J. J. Zhang, "Using factor analysis to identify neuromuscular synergies during treadmill walking," *Journal of Neuroscience Methods*, vol. 82, no. 2, pp. 207–214, 1998. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0165027098000545

[90] S. Hagio and M. Kouzaki, "The flexible recruitment of muscle synergies depends on the required force-generating capability," *Journal of Neurophysiology*, vol. 112, no. 2, pp. 316–327, 2014.

[91] J. Frère and F. Hug, "Between-subject variability of muscle synergies during a complex motor skill," *Frontiers in Computational Neuroscience*, vol. 6, no. DEC, pp. 1–13, 2012.

[92] S. A. Safavynia and L. H. Ting, "Task-level feedback can explain temporal recruitment of spatially fixed muscle synergies throughout postural perturbations," *Journal of Neurophysiology*, vol. 107, no. 1, pp. 159–177, 2012.

[93] T. Chau, "A review of analytical techniques for gait data. Part 1: Fuzzy, statistical and fractal methods," *Gait & posture*, vol. 13, pp. 49–66, 2001.

[94] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization." *Nature*, vol. 401, no. 6755, pp. 788–791, oct 1999.

[95] Q. An, Y. Ishikawa, S. Aoi, T. Funato, H. Oka, H. Yamakawa, A. Yamashita, and H. Asama, "Analysis of muscle synergy contribution on human standing-up motion using a neuro-musculoskeletal model," *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 2015-June, no. June, pp. 5885–5890, 2015.

[96] R. Cheng and J. W. Burdick, "Extraction of Muscle Synergies in Spinal Cord Injured Patients," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, vol. 2018-July, pp. 2623–2626, 2018.

[97] N. Yang, Q. An, H. Yamakawa, Y. Tamura, A. Yamashita, and H. Asama, "Muscle synergy structure using different strategies in human standing-up motion," *Advanced Robotics*, vol. 31, no. 1-2, pp. 40–54, 2017.

[98] M. C. Tresch, V. C. Cheung, and A. D'Avella, "Matrix factorization algorithms for the identification of muscle synergies: Evaluation on simulated and experimental data sets," *Journal of Neurophysiology*, vol. 95, no. 4, pp. 2199–2212, 2006.

[99] E. Bizzi and V. C. Cheung, "The neural origin of muscle synergies," *Frontiers in Computational Neuroscience*, vol. 7, no. April, pp. 1–6, 2013.

[100] Y. P. Ivanenko, G. Cappellini, N. Dominici, R. E. Poppele, and F. Lacquaniti, "Coordination of locomotion with voluntary movements in humans," *Journal of Neuroscience*, vol. 25, no. 31, pp. 7238–7253, 2005.

[101] A. S. C. Oliveira, L. Gizzi, U. G. Kersting, and D. Farina, "Modular organization of balance control following perturbations during walking," *Journal of Neurophysiology*, vol. 108, no. 7, pp. 1895–1906, 2012.

[102] M. W. Browne, "Cross-Validation Methods," *Journal of Mathematical Psychology*, vol. 44, no. 1, pp. 108–132, 2000. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0022249699912798

[103] N. Dominici, Y. P. Ivanenko, G. Cappellini, A. D'Avella, V. Mondì, M. Cicchese, A. Fabiano, T. Silei, A. Di Paolo, C. Giannini, R. E. Poppele, and F. Lacquaniti, "Locomotor primitives in newborn babies and their development," *Science*, vol. 334, no. 6058, pp. 997–999, 2011.

[104] B. Pan, Y. Sun, B. Xie, Z. Huang, J. Wu, J. Hou, Y. Liu, Z. Huang, and Z. Zhang, "Alterations of muscle synergies during voluntary arm reaching movement in subacute stroke survivors at different levels of impairment," *Frontiers in Computational Neuroscience*, vol. 12, no. August, pp. 1–11, 2018.

[105] A. D'Avella, P. Saltiel, and E. Bizzi, "Combinations of muscle synergies in the construction of a natural motor behavior," *Nature Neuroscience*, vol. 6, no. 3, pp. 300–308, 2003.

[106] V. C. Cheung, A. Turolla, M. Agostini, S. Silvoni, C. Bennis, P. Kasi, S. Paganoni, P. Bonato, and E. Bizzi, "Muscle synergy patterns as physiological markers of motor cortical damage," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, no. 36, pp. 14 652–14 656, 2012.

[107] S. Hagio, M. Fukuda, and M. Kouzaki, "Identification of muscle synergies associated with gait transition in humans," *Frontiers in Human Neuroscience*, vol. 9, no. FEB, pp. 1–12, 2015.

[108] A. C. H. Geurts, M. de Haart, I. J. W. van Nes, and J. Duysens, "A review of standing balance recovery from stroke," *Gait & Posture*, vol. 22, no. 3, pp. 267–281, 2005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0966636204002012

[109] P.-T. Cheng, M.-Y. Liaw, M.-K. Wong, F.-T. Tang, M.-Y. Lee, and P.-S. Lin, "The sit-to-stand movement in stroke patients and its correlation with falling," *Archives of Physical Medicine and Rehabilitation*, vol. 79, no. 9, pp. 1043–1046, 1998. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S000399939890168X

[110] M. Schenkman, R. A. Berger, P. O. Riley, R. W. Mann, and W. A. Hodge, "Whole-Body Movements During Rising to Standing from Sitting," *Physical Therapy*, vol. 70, no. 10, pp. 638–648, oct 1990. [Online]. Available: https://doi.org/10.1093/ptj/70.10.638

[111] A. R. Den Otter, A. C. Geurts, T. Mulder, and J. Duysens, "Speed related changes in muscle activity from normal to very slow walking speeds," *Gait and Posture*, vol. 19, no. 3, pp. 270–278, 2004.

[112] N. Mehrabi, M. H. Schwartz, and K. M. Steele, "Can altered muscle synergies control unimpaired gait?" *Journal of Biomechanics*, vol. 90, pp. 84–91, 2019. [Online]. Available: https://doi.org/10.1016/j.jbiomech.2019.04.038

[113] E. J. Fox, N. J. Tester, S. A. Kautz, D. R. Howland, D. J. Clark, C. Garvan, and A. L. Behrman, "Modular control of varied locomotor tasks in children with incomplete spinal cord injuries," *Journal of Neurophysiology*, vol. 110, no. 6, pp. 1415–1425, 2013.

[114] A. Santuz, A. Ekizos, L. Janshen, F. Mersmann, S. Bohm, V. Baltzopoulos, and A. Arampatzis, "Modular control of human movement during running: An open access data set," *Frontiers in Physiology*, vol. 9, no. OCT, pp. 1–11, 2018.

[115] F. E. Zajac, R. R. Neptune, and S. A. Kautz, "Biomechanics and muscle coordination of human walking: part II: lessons from dynamical simulations and clinical implications." *Gait & posture*, vol. 17, no. 1, pp. 1–17, feb 2003.

[116] R. R. Neptune and K. Sasaki, "Ankle plantar flexor force production is an important determinant of the preferred walk-to-run transition speed," *Journal of Experimental Biology*, vol. 208, no. 5, pp. 799–808, 2005.

[117] A. E. Patla, "Some characteristics of EMG patterns during locomotion: implications for the locomotor control process." *Journal of motor behavior*, vol. 17, no. 4, pp. 443–461, dec 1985.

[118] Q. An, Y. Ishikawa, J. Nakagawa, H. Oka, H. Yamakawa, A. Yamashita, and H. Asama, "Muscle synergy analysis of human standing-up motion with different chair heights and different motion speeds," *Proceedings - 2013 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2013*, pp. 3579–3584, 2013.

[119] N. Yang, Q. An, H. Kogami, H. Yamakawa, Y. Tamura, K. Takahashi, M. Kinomoto, H. Yamasaki, M. Itkonen, F. Shibata-Alnajjar, S. Shimoda, N. Hattori, T. Fujii, H. Otomune, I. Miyai, A. Yamashita, and H. Asama, "Temporal Features of Muscle Synergies in Sit-to-Stand Motion Reflect the Motor Impairment of Post-Stroke Patients," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 10, pp. 2118–2127, 2019.

[120] S. Asavasopon, M. Rana, D. J. Kirages, M. S. Yani, B. E. Fisher, D. H. Hwang, E. B. Lohman, L. S. Berk, and J. J. Kutch, "Cortical activation associated with muscle synergies of the human male pelvic floor," *Journal of Neuroscience*, vol. 34, no. 41, pp. 13 811–13 818, 2014.

[121] F. A. Mussa-Ivaldi and E. Bizzi, "Motor learning through the combination of primitives," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 355, no. 1404, pp. 1755–1769, 2000.

[122] L. Guidetti, G. Rivellini, and F. Figura, "EMG patterns during running: Intra-and inter-individual variability," *Journal of Electromyography and Kinesiology*, vol. 6, no. 1, pp. 37–48, 1996.

[123] D. D. Lee and H. S. Seung, "44565," vol. 401, no. October 1999, pp. 788–791, 2000.

[124] M. W. Berry, M. Browne, A. N. Langville, V. P. Pauca, and R. J. Plemmons, "Algorithms and applications for approximate nonnegative matrix factorization," *Computational Statistics and Data Analysis*, vol. 52, no. 1, pp. 155–173, 2007.

[125] T. Takei, J. Confais, S. Tomatsu, T. Oya, and K. Seki, "Neural basis for hand muscle synergies in the primate spinal cord," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 114, no. 32, pp. 8643–8648, 2017.

[126] M. J. MacLellan, Y. P. Ivanenko, F. Massaad, S. M. Bruijn, J. Duysens, and F. Lacquaniti, "Muscle activation patterns are bilaterally linked during split-belt treadmill walking in humans," *Journal of Neurophysiology*, vol. 111, no. 8, pp. 1541–1552, 2014.

[127] M. M. Khemlani, J. H. Carr, and W. J. Crosbie, "Muscle synergies and joint linkages in sit-to-stand under two initial foot positions," *Clinical Biomechanics*, vol. 14, no. 4, pp. 236–246, 1999.

[128] P. Calmels and P. Minaire, "A review of the role of the agonist/antagonist muscle pairs ratio in rehabilitation," *Disability and Rehabilitation*, vol. 17, no. 6, pp. 265–276, 1995. [Online]. Available: https://doi.org/10.3109/09638289509166646

[129] F. Rossi, P. M. Ros, R. M. Rosales, and D. Demarchi, "Embedded bio-mimetic system for functional electrical stimulation controlled by event-driven sEMG," *Sensors (Switzerland)*, vol. 20, no. 5, 2020.

[130] L. H. Ting and J. M. Macpherson, "A limited set of muscle synergies for force control during a postural task," *Journal of Neurophysiology*, vol. 93, no. 1, pp. 609–613, 2005.

[131] J. Roh, W. Z. Rymer, and R. F. Beer, "Robustness of muscle synergies underlying three-dimensional force generation at the hand in healthy humans," *Journal of Neurophysiology*, vol. 107, no. 8, pp. 2123–2142, 2012. [Online]. Available: https://doi.org/10.1152/jn.00173.2011

[132] M. Kaufman, U. Zurcher, and P. S. Sung, "Entropy of electromyography time series," *Physica A: Statistical Mechanics and its Applications*, vol. 386, no. 2, pp. 698–707, 2007.

[133] D. Farina, R. Merletti, and R. M. Enoka, "The extraction of neural strategies from the surface EMG: An update," *Journal of Applied Physiology*, vol. 117, no. 11, pp. 1215–1230, 2014.

[134] T. E. Flaxman, M. S. Shourijeh, K. B. Smale, T. Alkjær, E. B. Simonsen, M. R. Krogsgaard, and D. L. Benoit, "Functional muscle synergies to support the knee against moment specific loads while weight bearing," *Journal of Electromyography and Kinesiology*, vol. 56, no. November 2020, p. 102506, 2021. [Online]. Available: https://doi.org/10.1016/j.jelekin.2020.102506

[135] J. Fox, *Applied Regression Analysis and Generalized Linear Models*, 3rd ed. SAGE Publications, 2015. [Online]. Available: https://books.google.pt/books?id=cjB3BwAAQBAJ

[136] T. Chen, M. Xu, J. Tu, H. Wang, and X. Niu, "Relationship between Omnibus and Post-hoc Tests: An Investigation of performance of the F test in ANOVA," *Shanghai Archives of Psychiatry*, vol. 30, no. 1, pp. 60–64, 2018.

[137] W. P. Lombard and F. M. Abbott, "THE MECHANICAL EFFECTS PRODUCED BY THE CONTRACTION OF INDIVIDUAL MUSCLES OF THE THIGH OF THE FROG," *American Journal of Physiology-Legacy Content*, vol. 20, no. 1, pp. 1–60, 1907. [Online]. Available: https://doi.org/10.1152/ajplegacy.1907.20.1.1

[138] J. G. Andrews, "The functional roles of the hamstrings and quadriceps during cycling: Lombard's Paradox revisited," *Journal of Biomechanics*, vol. 20, no. 6, pp. 565–575, 1987. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0021929087902788