

Removal of Myoelectric Signal Disturbances in Prosthetics Control with PCA and ICA

An Experimental Signal Processing Study
Master's thesis in Biomedical Engineering

JOHANNA LUNDIN

DEPARTMENT OF ELECTRICAL ENGINEERING
CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2022
www.chalmers.se

DEPARTMENT OF ELECTRICAL ENGINEERING
CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2022
www.chalmers.se

Master Thesis

Removal of Myoelectric Signal Disturbances in Prosthetics Control with PCA and ICA *An Experimental Signal Processing Study*

Johanna Lundin

February 2nd, 2022



CHALMERS
UNIVERSITY OF TECHNOLOGY

Removal of Myoelectric Signal Disturbances in Prosthetics Control with PCA and ICA
An Experimental Signal Processing Study

JOHANNA LUNDIN
© JOHANNA LUNDIN, 2022.

Supervisor:
Jan Zbinden, Department of Electrical Engineering

Examiner:
Max Ortiz Catalan, Department of Electrical Engineering

Master's Thesis 2022
Center for Bionics and Pain Research
Department of Electrical Engineering
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: A Venn diagram describing several factors that affect the quality of EMG signal during acquisition.

Gothenburg, Sweden 2022

Acknowledgments

I would like to express my gratitude to Professor Max Ortiz Catalan for the possibility to conduct my master thesis project at the Center for Bionics and Pain Research. Since the beginning, I have been included in the team and surrounded by competent and supporting researchers. So, thanks to all of you at the Center for Bionics and Pain Research for the warm welcome and for introducing me to the research field.

I would like to express an extra and special thank you to my supervisor Jan Zbinden who has contributed a lot with both his expertise and time during the entire project. Without your guidance, the experience would not have been as pleasant as it finally was. Another thanks to Dr Eric Earley for contributing with your knowledge and time.

Finally, a big thank you for the discussions and emotional support given by my fellow students Julia Molin, Selena van der Horst, Cajsa Hjohlman, and Sofia Cvetkovic Destouni!

Johanna Lundin, Gothenburg, February 2022

Abstract

The most common technique to control active prostheses is via electromyography (EMG), where myoelectric signals from voluntary muscle contractions are recorded and interpreted as different movements. One longstanding challenge in EMG prosthetic control is crosstalk. Crosstalk in EMG can be explained as the part of an EMG signal recorded over a muscle, but that has been produced by another muscle. The appearance of crosstalk makes decoding of the intended prosthetic movement more difficult and the resulting control less reliable which in turn affects the performance of the prosthesis. This project investigates how principal component analysis (PCA) and independent component analysis (ICA) affect the EMG signals acquired from implanted electrodes of a participant with a trans-humeral amputation. Classification with a multilayer perceptron showed that the application of PCA over offline featured EMG signals could increase the accuracy for pattern recognition from 87.7% to 99.9%. Meanwhile, the application of ICA increased the accuracy from 87.7% to 93.6%. One modified offline method with PCA which also could be adapted for online usage was tested and resulted in an accuracy of 93.9%. These results were corroborated on a dataset of a participant with trans-radial amputation where an offline PCA lead to an accuracy improvement from 74.9% to 94.9% and the PCA algorithm adapted for online use resulted in an achieved accuracy of 83.6%. If this improved performance could be achieved online these findings could in the long term improve prosthetics control and yield better life quality for people suffering from limb loss.

Table of Contents

Abbreviations	viii
2 Introduction.....	1
2.1 Aim and motivation behind the project.....	2
3 Theory	3
3.1 Electromyography and electrodes	3
3.2 Muscle and nerve supplies in the arm.....	4
3.3 Crosstalk.....	4
3.4 Signal Processing	6
3.4.1 Principal Component Analysis.....	6
3.4.2 Individual Component Analysis	7
4 Method	10
4.1 Literature Review	10
4.2 Experimental Signal Processing	10
4.2.1 Participants	10
4.2.2 Datasets	11
4.2.3 Experiment 1	12
4.2.4 Experiment 2	13
4.2.5 Experiment 3	14
4.2.6 Experiment 4	15
4.2.7 Experiment 5	15
4.2.8 Statistical analysis.....	15
5 Results	16
5.1 Experiment 1.....	16
5.1.1 PCA over EMG signals.....	16
5.1.2 ICA over EMG signals	17
5.2 Experiment 2.....	18
5.2.1 PCA over features extracted from EMG signals	18
5.2.2 ICA over features extracted from EMG signals	19
5.3 Experiment 3.....	21
5.4 Experiment 4.....	24
5.5 Experiment 5.....	27
6 Discussion	30
6.1 Interpretations of the results	30
6.2 Limitations within the project.....	32

6.3	Future work.....	32
7	Conclusion.....	34
	References.....	35
	Appendix.....	38

Abbreviations

ALC	Artificial Limb Controller
CBPR	Center for Bionics and Pain Research
EMG	Electromyography
ICA	Individual Component Analysis
kNN	k-Nearest Neighbor
LDA	Linear Discriminant Analysis
MAV	Mean Absolut Value
MLP	Multilayer Perceptron
MUAP	Muscle Unit Action Potential
PCA	Principal Component Analysis
RMS	Root Mean Square
RPNI	Regenerative Peripheral Nerve Interfaces
SLPCH	Slope changes
SVD	Single Value Decomposition
TMR	Targeted Muscle Reinnervation
TWL	Time Wavelength
ZC	Zero Crossing

1 Introduction

Upper limb amputation affects daily living to great extent. The amputation itself implies loss of function, which forces the people with amputation to adapt their living and interactions with the environment to new circumstances. Easy tasks such as lifting a cup of coffee or performing a handshake are no longer trivial.

Prosthetics can provide great help for those people suffering from limb loss, nevertheless, it is not rare that amputees to different grades abandon their devices [1]. Common reasons behind this can be related to the discomfort of the prosthesis [2], limited functionality [3], [4], or limitations of the prosthetic control [5], to mention a few examples. Looking closer into prosthetic control there exist several factors which challenge the performance, where the movement predictability of the device is one.

Electromyography (EMG) is the most common technique to control active prostheses. The technique is based on the idea of recording myoelectrical signals, evoked by voluntary muscle contractions, and interpretation of these signals as different movements. One phenomenon which can occur during recording of EMG signals, and which aggravates the decoding into different movements is the appearance of crosstalk. Crosstalk observed in EMG can be described as the part of an EMG signal recorded over a muscle, but that has been produced by another muscle [6]. Unfortunately, there do not yet exist any ideal suggestions on how to identify, quantify or reduce muscular crosstalk [7]. In addition, most research is performed on surface EMG, leaving the whole field of implantable EMG electrodes (see Figure 1 for an example application of implantable electrodes within prosthetics) yet to be explored.

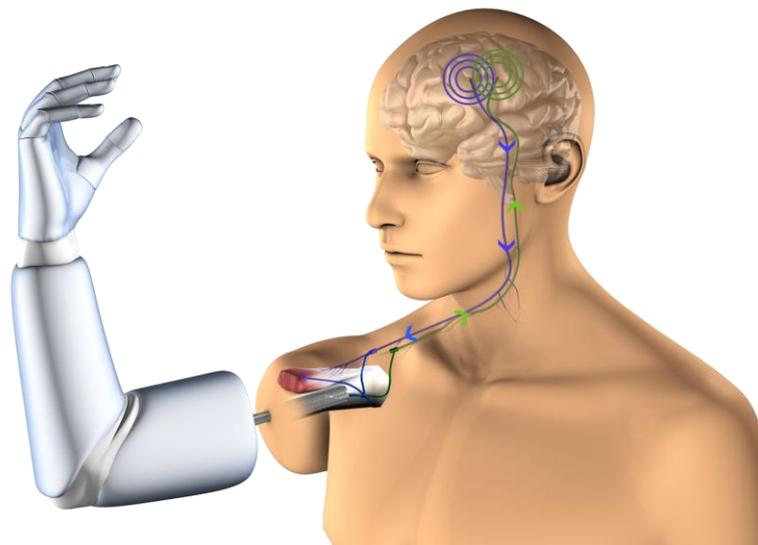


Figure 1 An osseointegrated human-machine gateway [8], allowing for bidirectional communication between the human body and the prosthesis. To control the prosthesis, the user's movement intent (blue signal) is transported via the nerves to the muscles, where the bio-signal generated by the muscles is measured. An algorithm decodes the measured bio-signal and moves the prosthesis accordingly.

Two interesting algorithms used for reduction of crosstalk are Principal Components Analysis (PCA) and Independent Component Analysis (ICA). Where PCA is an algorithm used to compress information [9], while ICA is an algorithm used to separate information [10]. The two algorithms can be used one by one, or in combinations with each other. If used together it is not rare to use PCA as a pre-processing step before ICA, as it reduces the dimensionality of data [11].

1.1 Aim and motivation behind the project

This project aims to investigate if crosstalk in signals obtained from implantable EMG electrodes can be reduced to improve prosthetic control performance. The performance of the two algorithms PCA and ICA applied on EMG signals recorded from two different research participants will be evaluated.

The motivation behind the project is to contribute to the research within the prosthetics field. Reduction of crosstalk leads to improved volition decoding and this in turn results in better prosthetic control. In a bigger context, this could improve the life quality for people with upper limb loss, today and in the future.

2 Theory

To understand how upper limb EMG prosthetics work, knowledge in the interdisciplinary field of the human body and electronics is beneficial. This section provides a background on EMG, anatomy, crosstalk, and signal processing, relevant for the field of prosthetics.

2.1 Electromyography and electrodes

The human body has three different types of muscle tissue: skeletal, cardiac, and smooth. The skeletal muscle tissue is the one used to produce and stabilize different movements from voluntary contractions, such as the movement of an arm. These movements can be studied and recorded with the help of EMG, which measures electrical changes within the muscular tissue. [12]

The electrical activity within muscular tissue appears due to activation of individual motor unit action potentials (MUAPs). The MUAPs propagate through specific neurons, called motoneurons, on-demand from the brain or spinal cord and to their targeting muscle fibers. The summation of several MUAPs recorded in the targeting muscle together builds up to an EMG signals, as pictured in Figure 2. [13]

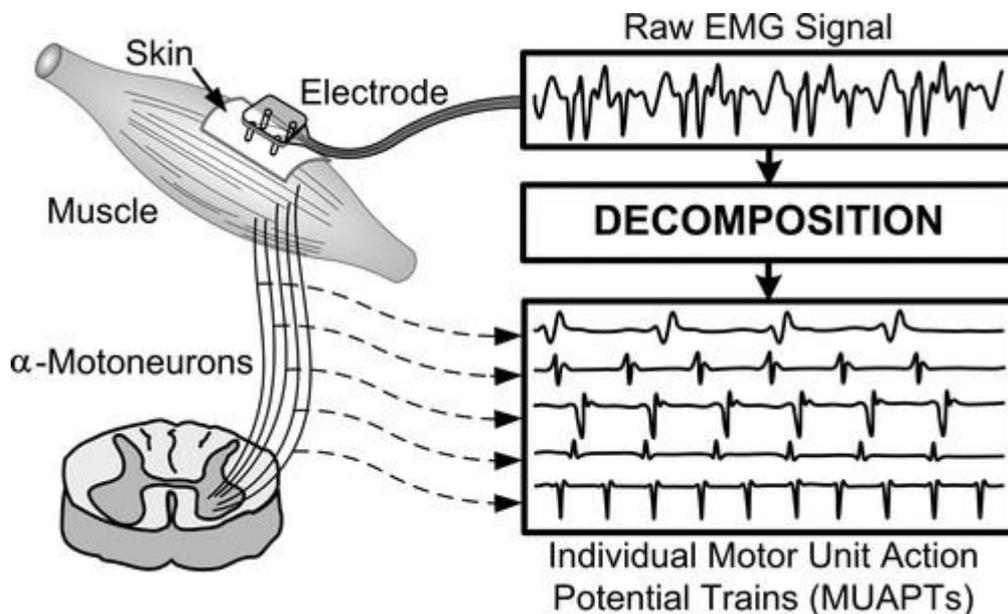


Figure 2 Decomposition of surface EMG signals into individual motor unit action potentials [14]

To record electrical activity within a muscle several options exist. Commonly used solutions are non-invasive surface EMG where electrodes are placed directly on the skin above the muscle of interest, or needle electrodes temporarily inserted through the skin [13]. In addition to surface and needle EMG, there are also solutions where the electrodes are permanently implanted in the body. These electrodes can be either extra-muscular or intra-muscular [15]. An extra-muscular electrode, here also

referred to as an epimysial electrode, is sewn directly onto the epimysium [15], the outer layer of tissue surrounding the muscle within the body [12]. Meanwhile, intramuscular electrodes are fully or partially implanted through the muscle [15]. At the Center for Bionics and Pain Research, the latter options are used, intra-, and extra-muscular electrodes.

2.2 Muscle and nerve supplies in the arm

The arm can be divided into the upper arm, forearm, wrist, and hand, controlled by different muscles and nerve supplies. In general, muscles in the upper arm are fewer and bigger compared to muscles in the lower arm and hand which are several, smaller, and more specifically in the movements that they control. Muscles can be placed deep into the body or superficial, close to the skin. The muscles of the arm are mainly supplied by the Median, Radial, and Ulnar nerves, which all branch out more and more the closer to the hand they come. [12]

When limb loss occurs, nerve and muscle sources are lost or damaged, resulting in loss of function. The limb itself can often be replaced with a prosthesis, but the loss and damage to muscles and nerves remain. Two techniques used to restore function and increase the number of muscle sources for prosthetics control are targeted muscle reinnervation (TMR) [16], and regenerative peripheral nerve interfaces (RPNI) [17], [18]. In TMR a free ending of a cut nerve is relocated to a new but native muscle target which it can reinnervate. RPNI does the same but with the difference that the muscle target no longer is native, instead free muscle graft is denervated and then wrapped around the free nerve ending. Usage of TMR and RPNI in prosthetics control makes it possible to increase the number of myoelectrical sources in the limb which in the best case can increase the number of controllable joints of a prosthesis.

2.3 Crosstalk

According to Mezzarane and Kohn, *“The crosstalk phenomenon consists in recording the volume-conducted electromyographic activity of muscles other than that under study.”* [19]. This is an effect commonly observed in surface EMG where the surface electrodes do not allow specifying the target precisely, especially if the muscle is not superficial [6]. Hence, it is possible that electrical activity from a volume bigger than that one of interest can contribute to the captured signals [6].

Figure 3 visualizes a diagram of several factors affecting the quality of the acquired EMG signals from a recording. Crosstalk is one out of several factors, and should not be mixed up with the terms co-activation and co-contraction which often are used interchangeably and describe muscle pairs that work together to create and stabilize movements [20]. Muscle synergies is also a term that can create confusion. Muscle synergies is a theory where movements are explained to be controlled by specific groups of muscles [21]. Comparing these latter terms with crosstalk, a major

difference is that in crosstalk the measured “contributions” are un-desired, meanwhile, the electrical activity is expected and needed in co-activation and muscle synergies.

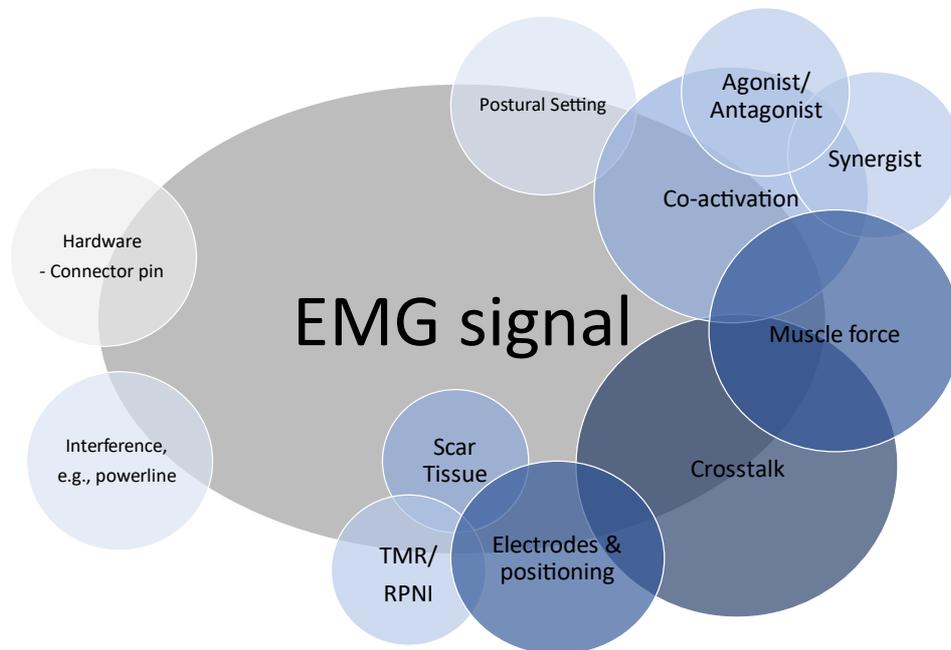


Figure 3 Venn diagram visualizing crosstalk as one out of several factors that effects the quality of an EMG signal

Even though the crosstalk phenomenon and its appearance are well known there are no unambiguous explanations why and how it appears, how to identify it, or how to reduce it. Talib et al. have compiled a summary of several methods that previously have been used to handle crosstalk, including identification, quantification, and reduction [7]. One method for identification of crosstalk is through simultaneous measurements with needle and surface EMG [22], [23]. The hypothesis is that the differences between the signals will include crosstalk, as the needle is expected to measure more local activity in comparison to the surface electrodes which will capture activity from a bigger volume. Another suggestion is to isolate the muscle of activity, meanwhile measuring the activity of surrounding muscles which are expected to be quiet and then compare those [24].

Proceeding to the step about how to quantify the crosstalk, three different indices, amplitude- [25], power- [26], and cross-correlation-based index [27], are presented. The methods compare amplitudes, power of a signal, and waveforms of the signals, from measurements with and without crosstalk. The methods are widely discussed, and unfortunately, all have their disadvantages. They might be hard to apply in practice or have the drawback that even if the indices indicate a difference, it cannot be decided whether it is due to crosstalk or co-activation.

In the final step of reduction, Talib et al. mention electrode geometry and interelectrode distance [28], [29], electrode configuration [25], and filtering [30], as

examples. One method is to use bipolar electrodes which filter away differences between the measuring points, a difference which could include crosstalk. Although bipolar electrodes have demonstrated successful results for crosstalk removal, it is not always a suitable option. One example is when using implanted EMG electrodes where there might be reasons to minimize the number of wires needed, meanwhile maximizing the sites of measuring. Then bipolar electrodes have the disadvantage of requiring more cables compared to monopolar electrodes for the same amount of measuring sites [31]. Here filtering might be a better solution or a good complement as it does not affect the number of physical electrodes needed. Examples of filtering methods that previously have been tested for crosstalk or similar challenges, with either interesting or predominant positive outcomes, are ICA and PCA [10], [31], [32].

2.4 Signal Processing

In this project, several signal processing methods for crosstalk reduction have been developed and tested, all based on the two algorithms PCA and ICA. PCA is an algorithm for compressing data while ICA is for separating data, each algorithm and its motivation for this specific project is described below.

2.4.1 Principal Component Analysis

PCA is an algorithm that can be used to decrease the number of dimensions within a correlated dataset [9]. The dimensionality reduction is performed through a projection of the data onto another coordinate system [9]. The new coordinate system is built up of orthogonal components maximizing the variance of the initial dataset [9]. During transformation, the components are ranked in decreasing order with the component maximizing the variance first [9]. Thus, one can safely reduce the number of dimensions by disregarding components, starting from the last.

Assuming that crosstalk is a disturbance appearing as parts of muscular signals with low variance (at one or several recording sites), then there exists a possibility that the PCA algorithm would extract the crosstalk as one or several principal components. Removal of these “crosstalk components” makes it theoretically possible to reconstruct new signals free from crosstalk. Figure 4 describes the process where muscular activity interferes with each other and is being captured by EMG electrodes. PCA transforms the observed signals into new extracted sources, where the important signal content is conserved separated from minor disturbances, such as crosstalk.

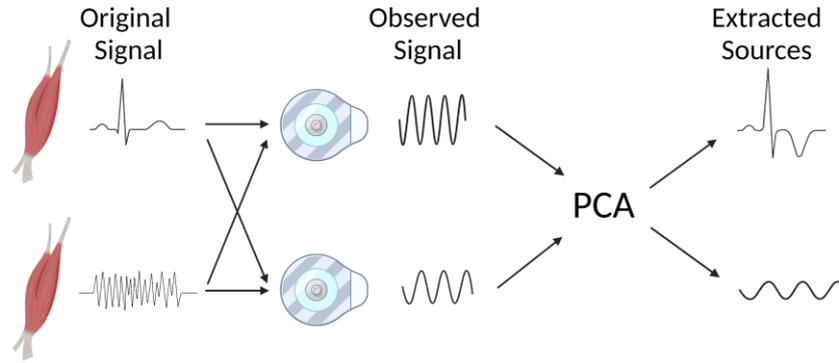


Figure 4 PCA diagram, showing how the original sources are mixed up before being observed. Followed by PCA transformation to get the extracted sources. Figure created with BioRender.com.

Two common implementations of PCA are through eigenvalue decomposition of the data covariance matrix, and through singular values decomposition (SVD) over the centered data [32]. In this project the latter has been used, mathematically explained by the following steps [32]:

1. Centering of the data
2. Compute the empirical covariance matrix
3. Obtain eigenvalues and eigenvectors of the covariance matrix
4. Sort the eigenvalues in descending variance order
5. Project data onto the transformation matrix

Defining the raw EMG signals as, Z , the reconstructed signals, Z_{pca_r} , can be achieved through the equation:

$$Z_{pca_r} = USV' + \mu.$$

Here U , S , and V , respectively are the left singular vector, the singular values, and the right singular vector computed through SVD [9], over the signals Z centered by its mean μ . To eliminate undesired principal components the matrices U , S , and V , can be truncated at values between one and the total number of signals in Z [9]. When referring to a matrix U later in this report this is the U being referred to.

2.4.2 Individual Component Analysis

ICA is a statistical signal processing technique, initially developed to solve problems related to the cocktail-party problem [33]. The cocktail-party problem describes situations where two or more sources simultaneously emit signals (voices from guests at a party) that mixed up arrives at the receiver (another guest). Despite the signals arriving mixed up, the receiver can separate the signals and understand which source emitted each signal. ICA is an algorithm trying to describe and solve this challenge mathematically.

Applying the cocktail-party problem on the case of EMG signals with crosstalk (Figure 5), the muscles could be considered the sources, the electrodes as the receiver, and the ICA the part of the receiver trying to separate the signals.

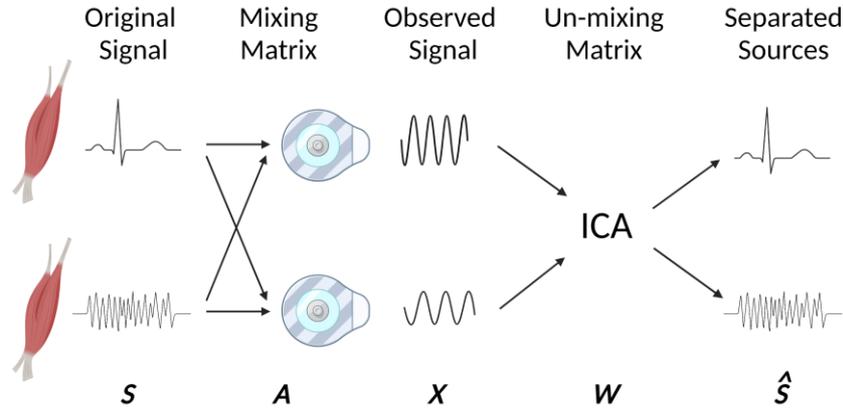


Figure 5 ICA diagram where S are the sources, X are the correlated recorded signals, \hat{S} the separated signals. A is the matrix mixing the signals and W is the matrix un-mixing them. Figure created with BioRender.com.

If it is true that crosstalk appears from independent sources, it is theoretically possible to remove these and reconstruct signals free from crosstalk. An important difference for ICA compared to PCA is that the components from ICA are unsorted [33]. Hence, it is up to the user of ICA to decide whether any component is more “important or correct” than another.

In this project, the fast ICA algorithm has been deployed on data preprocessed with centering and whitening. The algorithm is commonly used thanks to its fast convergence, easy implementation, and satisfying performance [34]. The general idea is to find the best approximated un-mixing matrix W , such that the measured signals X , can be transformed into a format \hat{S} , mimicking the initial sources S [35]. Using equations, the measured signals X can be formulated as,

$$X = AS,$$

where S are the original signals, and A is the matrix describing how these original signals influence each other before being observed by the electrodes. The reconstructed signals can be formulated as,

$$\hat{S} = WX.$$

Here W is a matrix approximated to maximize the non-Gaussianity through iteration of a function g , measuring the Negentropy [33]. The steps followed are presented below [33]:

1. Divide a random initial vector w by its norm
2. Let $w^+ = E\{xg(w^T x)\} - E\{g'(w^T x)\}w$
3. Let $w = \frac{w^+}{\|w^+\|}$
4. Go back to step 2 until w converge

After approximation of W the separated sources \hat{S} , can be studied, and desired data reconstructed.

3 Method

The method is divided into one initial theoretical part, and one following practical part. The theoretical part consists of a literature review, meanwhile, the practical part includes an experimental signal processing study of recorded EMG signals from prosthetic users.

3.1 Literature Review

The literature review was performed through the reading of scientific articles and books. The study did mainly focus on the current state of the art within the field of active prosthetics, crosstalk in muscles, and the muscular system of the upper arm. Articles and books were found through the databases Google Scholar, Scopus, and the Chalmers University of Technology Library. Different combinations of the following keywords were used: prosthetics, EMG, epimysial, intramuscular, implanted, electrodes, crosstalk, co-activation, co-contraction, muscle synergy, far-field potential, anatomy, PCA, and ICA.

The outcome of the literature review was used as an introduction to the research field, and as a design tool for the study of the following experimental work.

3.2 Experimental Signal Processing

In the experimental signal processing part of the project data acquisition and five different experimental set-ups were performed. Statistical analysis was used to interpret the results.

3.2.1 Participants

Two participants took part in the study, one with a trans-humeral amputation (Participant A) and one with a trans-radial amputation (Participant B). Before the experiment started both participants signed an informed consent.

3.2.1.1 Data acquisition

Both Participant A and Participant B have had muscular electrodes for prosthetics control surgically implanted in their upper or lower arm. After surgery, both participants have taken part in several follow-ups at the Center for Bionics and Pain Research. During these occasions, the performance of the EMG signals and their possibility to be decoded into different hand and finger movements have been studied and recorded.

Recordings have been made through the open-source platform BioPatRec [36] and with an Artificial Limb Controller (ALC) [37]. Where BioPatRec is a research tool

focused on myoelectric pattern recognition, and the ALC is an embedded system developed for the control of prosthetic devices. The ALC sampled data at 500 Hz and with 16-bit resolution. In addition, a high-pass filter at 20 Hz and a notch filter of 50 Hz was applied in real-time.

Each follow-up consisted of two recording sessions, one for gross movements:

- Open/close hand (2 movements)
- Flex/extend hand (2 movements)
- Pronation/supination of the forearm (2 movements)
- Flex/extend of the elbow (2 movements – only executed by Participant A)

And one recording session for single finger movements:

- Flex/extend of each finger independently (10 movements)

When recording the data, the participant was asked to sit down comfortably on a chair in front of a computer. The computer screen instructed the participant to conduct one movement at a time, with 70% of the maximum voluntary contraction. Each movement was repeated three times for three seconds and with three seconds of rest in between.

3.2.2 Datasets

The recordings resulted in two different datasets, Dataset A and Dataset B (corresponding to Patient A and Patient B), presented in Table 1.

Table 1 Dataset information

Designation	Dataset A	Dataset B
Gender	Male	Female
Age of participant [years]	54	49
Amputation	Trans-humeral	Trans-radial
Time frame	09-01-2019 – 19-05-2021	22-10-2019 – 28-04-2021
Number of sessions	9	7

Dataset A has been used in Experiment 1-3, and Dataset B has been used in Experiment 4. Both participants had 12 monopolar muscular electrodes implanted each.

3.2.3 Experiment 1

In Experiment 1 the movement classification accuracy of EMG signals, and the EMG signals processed with PCA, respectively ICA, were calculated and compared. Three different inputs were used for the PCA and ICA algorithm, the raw format of the EMG signals, min-max normalization, and Z-score normalization of the EMG signals. Normalization was applied to facilitate visual comparison of the signals before and after PCA and ICA and to detect possible differences in the performance of the algorithms. Figure 6 visualizes the process.

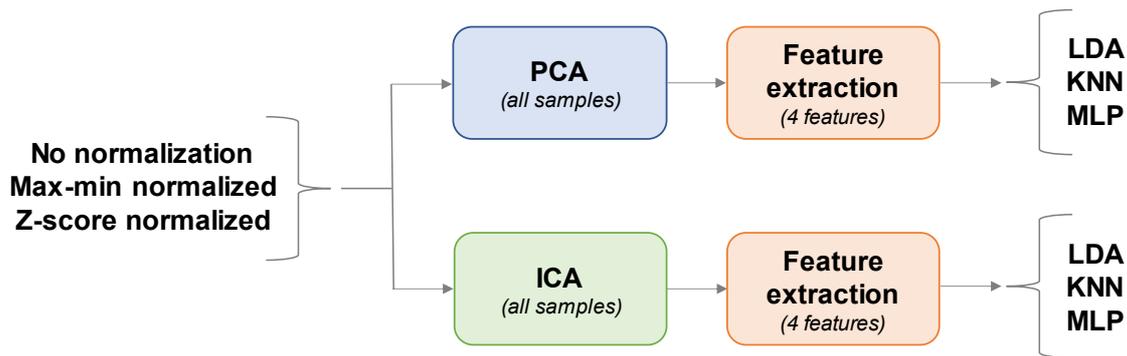


Figure 6 Block diagram describing the process in Experiment 1

Both PCA and ICA were performed on a 2D matrix format of the EMG signals, constituted of 12 channels times all their 18 or 16 movements followed by each other (including rest), resulting in the matrix dimension $[12 \text{ channels} \times 162000 \text{ samples}]$ or $[12 \text{ channels} \times 144000 \text{ samples}]$. A format resulting in a maximum truncation value of 12 for the PCA and the possibility of 12 independent components for ICA.

In the case of PCA, all truncation values between one and 12 were used, resulting in 12 different reconstructions of the same input data. Here the truncation reveals the number of principal components that are being kept. The ICA was used similarly to the PCA. Although, with the difference that instead of reconstructing the signals for different truncation values the reconstructions were now performed with one out of all twelve components removed, one at a time. This resulted in 12 different reconstructions of the ICA data.

From the reconstructed PCA and ICA signals, the four features mean absolute value (MAV), waveform length (TWL) in the time domain, zero crossing (ZC), and slope changes (SLPCH) were extracted [38]. The featured data were then classified with Linear Discriminant Analysis (LDA), k-nearest neighbor (KNN), and a multilayer perceptron (MLP), in the pattern recognition function in BioPatRec. Each classification was performed with ten repetitions to increase the reliability of the results, and the achieved accuracies were then used as a metric to evaluate the results.

3.2.4 Experiment 2

In Experiment 2 the features were now extracted before applying PCA or ICA, instead of afterward. Usage of all four features MAV, TWL, ZC, and SLPCH, allowed for up to 48 different truncation values in PCA, or 48 individual components for ICA. Z-score normalized data was used and features were extracted from windows of 200 ms, without any overlap between the windows.

In the case of PCA, two different methods were tested, see Figure 7. The first one included PCA applied over all four features MAV, TWL, ZC, and SLPCH. This was followed by classification over all four features, and all different subsets of three, two, and one feature, one at a time. In the second case, the PCA and classification were performed only over two features at a time.

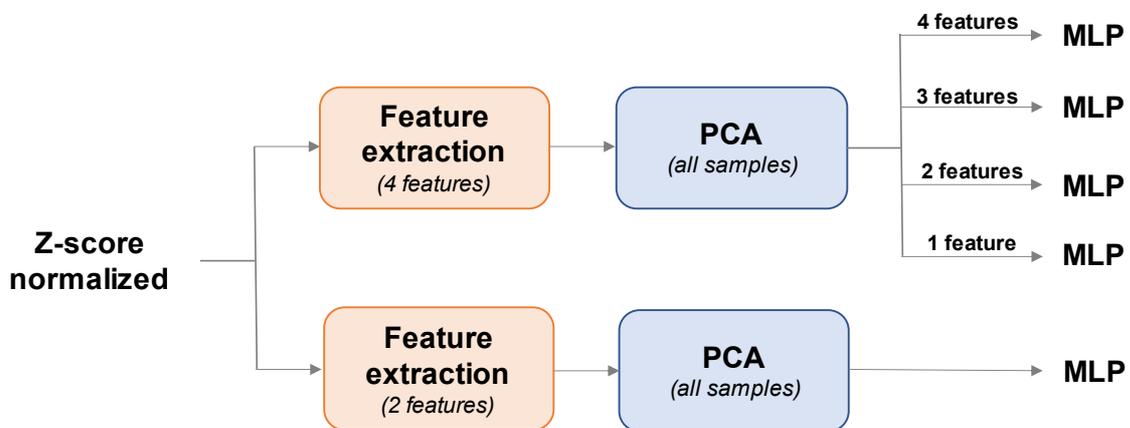


Figure 7 Block diagram describing the process for PCA in Experiment 2

Proceeding with ICA, it was tested to apply ICA directly on the extracted data, and over the reconstructed PCA data over all four features. A visual study of the ICA components did not reveal any obvious components to remove, hence a threshold was used to remove components with a specific magnitude, either bigger or smaller than the maximal peak-to-peak value among all 48 components. The process is visualized in Figure 8.

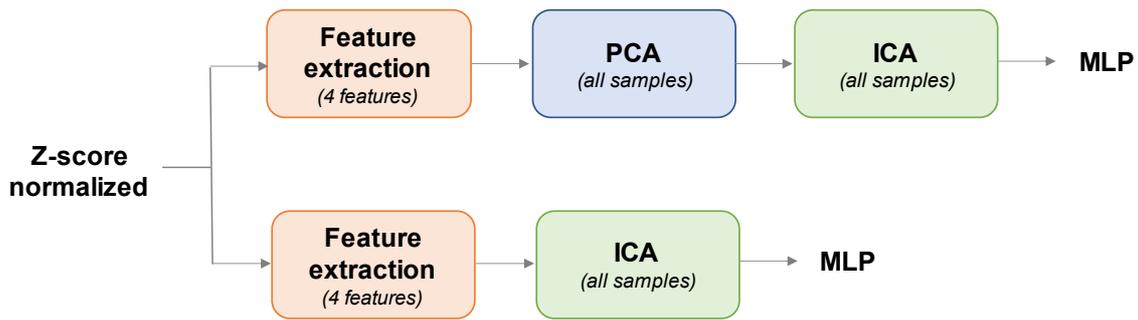


Figure 8 Block diagram describing the process for ICA in Experiment 2

3.2.5 Experiment 3

In Experiment 3 PCA was applied on extracted features, but with the difference from Experiment 2 that these windows for feature extraction now did overlap.

First, the case of the biggest possible overlap was tested, a step of one sample between each extraction was used, which in time corresponded to 2 ms. This resulted in a new dataset with the same number of samples but of featured data instead of the “raw data”. From this new dataset new windows of 200 ms was extracted and PCA was performed within each window, see the upper process in Figure 9. Different conditions were used and U was either taken from the PCA calculation within each window or substituted with an U calculated over the entire signal from the same session, or another session. The results were then compared with the case of no PCA:

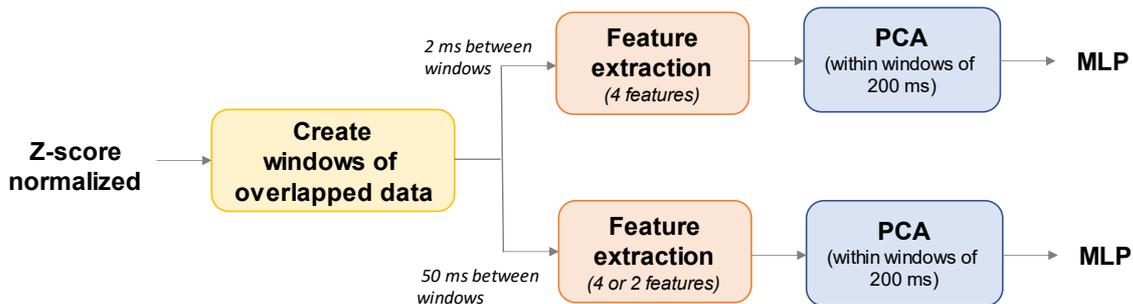


Figure 9 Block diagram describing the process in Experiment 2

In the second case, a smaller overlap was used, only an overlap of 25 samples corresponding to a feature extraction each 50 ms. See the lower process in Figure 9. Again, features were extracted and PCA applied within a window corresponding to 200 ms. It was tested to use extraction with all four features, and extractions only with the features TWL and SLPCH. This time the U used were taken from another

session than that other study, this to simulate an “online test”. Once again, the results were compared with the case of no PCA.

3.2.6 Experiment 4

In Experiment 4 selected cases with significantly improved accuracies from Experiment 2-3 were repeated but with Dataset B instead of the previously used Dataset A. This resulted in the repetition of the PCA part of Experiment 2 and the entire Experiment 3.

3.2.7 Experiment 5

In Experiment 5 the signals and different U before and after PCA for selected cases with significantly improved accuracies once again were displayed and compared. This includes PCA over features extracted from the entire recording without overlap, PCA after feature extraction each 2 ms, and extraction each 50 ms. Also, U matrices for the same experiments and cases were compared. An important difference from previous experiments is the fact that the accuracy no longer is used as a metric. Instead, an amplitude-based root mean square (RMS) index was used to compare signals before and after PCA, implemented accordingly to the following equation,

$$index = \frac{RMS(signal\ after\ PCA)}{RMS(signal\ before\ PCA)}.$$

For comparison of U the summation of the absolute values for all samples corresponding to each feature MAV, TWL, ZC, and SLPCH, were compared with the content for all features together.

3.2.8 Statistical analysis

To compare results two different methods have been used. Series of paired-sampled t-tests and one-way ANOVA, both with a Bonferroni correction. Before the performance of t-test and ANOVA all data was confirmed to be standard normally distributed with help of a One-sample Kolmogorov-Smirnov test.

The paired-sampled t-test tested against null hypothesis, that the compared data, X and Y, had normally distributed data with equal means and equal but unknown variances. Using this in a series the Y was kept constant while changing the data samples in variable X. One-way ANOVA was used similarly, but tested if there was a common mean between several groups of data, and not only between X and Y.

The significance level was set to 5%, meaning that a rejected test result corresponded to a Bonferroni-corrected $p < 0.05$. A rejected test value indicates a significant difference between the compared variables.

4 Results

Experiments 1-3 present results from the usage of Dataset A, and Experiment 4 presents results from the usage of Dataset B. In Experiment 5 both Datasets A and B are studied.

All figures are based on Z-score normalization (of session 9 for Dataset A and session 7 for Dataset B) and pattern recognition with MLP if nothing else is mentioned.

4.1 Experiment 1

In experiment 1 PCA and ICA were performed over raw data, max-min normalized data, and Z-score normalized EMG data of 12 signals. Pattern recognition was executed with ten repetitions and with three different algorithms, LDA, KNN, and MLP.

4.1.1 PCA over EMG signals

Figure 10 visualizes the results from applying PCA over max-min normalized data classified by MLP.

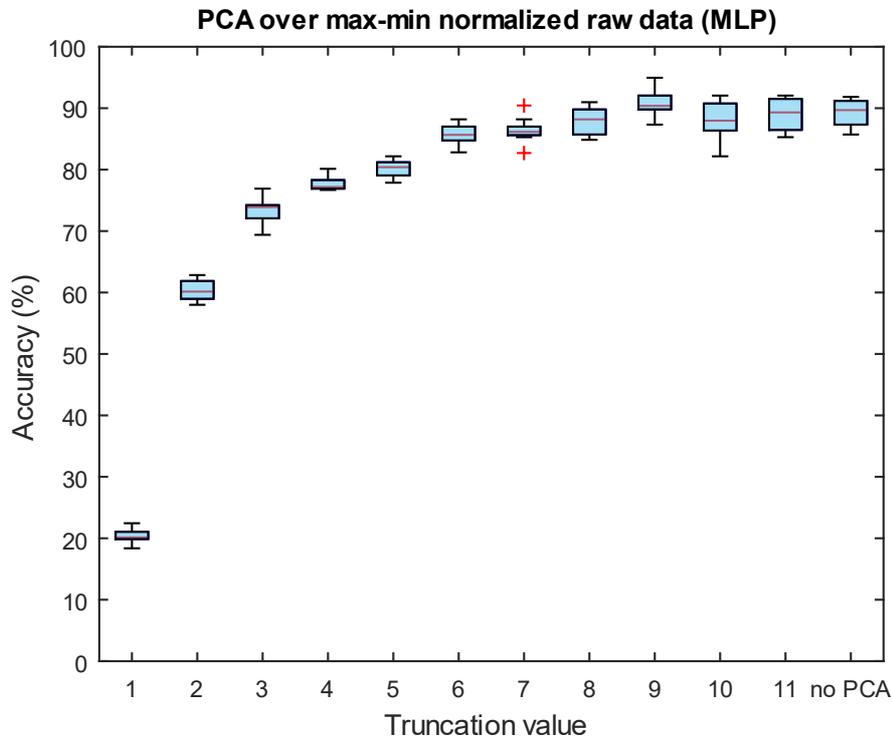


Figure 10 PCA over max-min normalized raw data and pattern recognition with MLP

According to t-test results, there was no significant difference between the case of no PCA and truncation values 7-11 ($p \geq 0.118$). Among the truncation values 1-6, there

was a significant difference ($p \leq 0.0129$) but none had an accuracy higher than the case of no PCA. No significant difference was found in either the LDA or kNN case as well (see Figure 24 in Appendix). This indicates that the PCA method in Experiment 1 did not improve the performance of the pattern recognition.

4.1.2 ICA over EMG signals

Figure 11 displays ICA over raw data and pattern recognition with an LDA classifier.

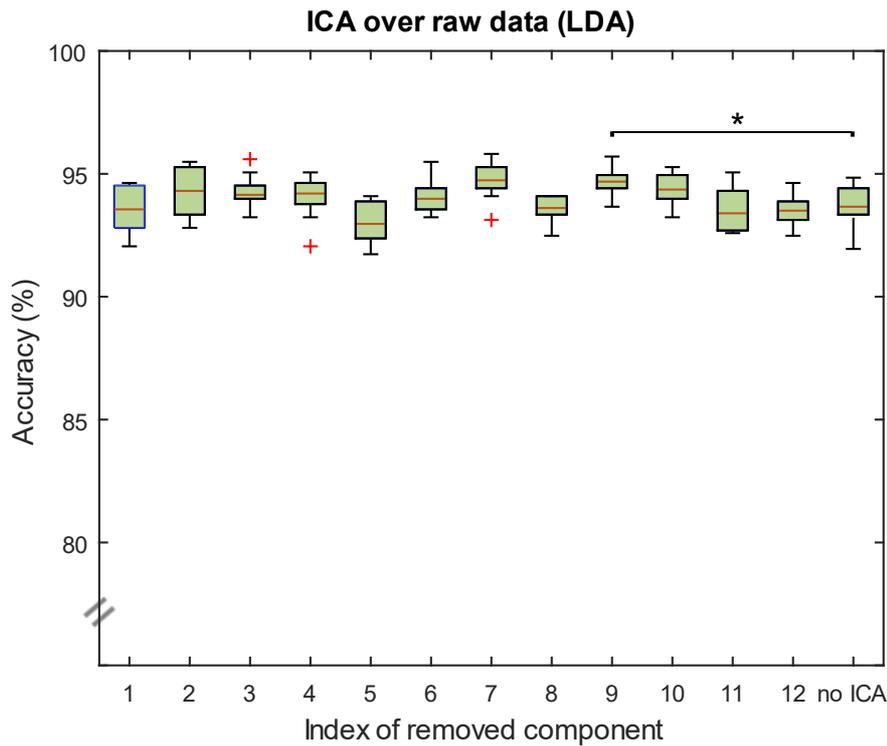


Figure 11 ICA over raw data and pattern recognition with an LDA classifier. The star indicated that the removal of component 9 yielded significantly better results than the case of "no ICA"

The star indicates that the pattern recognition performed better after removal of ICA component 9, compared with the case of no ICA ($p = 0.0347$). No other cases were significantly different than the case of no ICA ($p \geq 0.0509$), according to performed t-tests. Neither any of the cases for MLP or kNN indicated a significant improved performance (see Figure 25 in Appendix).

4.2 Experiment 2

In Experiment 2, PCA was performed over featured data, and ICA over PCA transformed data.

4.2.1 PCA over features extracted from EMG signals

Figure 12 visualizes the pattern recognition accuracy from applying PCA over 48 different truncation values (12 signals times four features) and all four features MAV, TWL, ZC, and SLPCH (red dashed line). This result is compared with the case of pattern recognition directly over featured data, without PCA (blue continuous line).

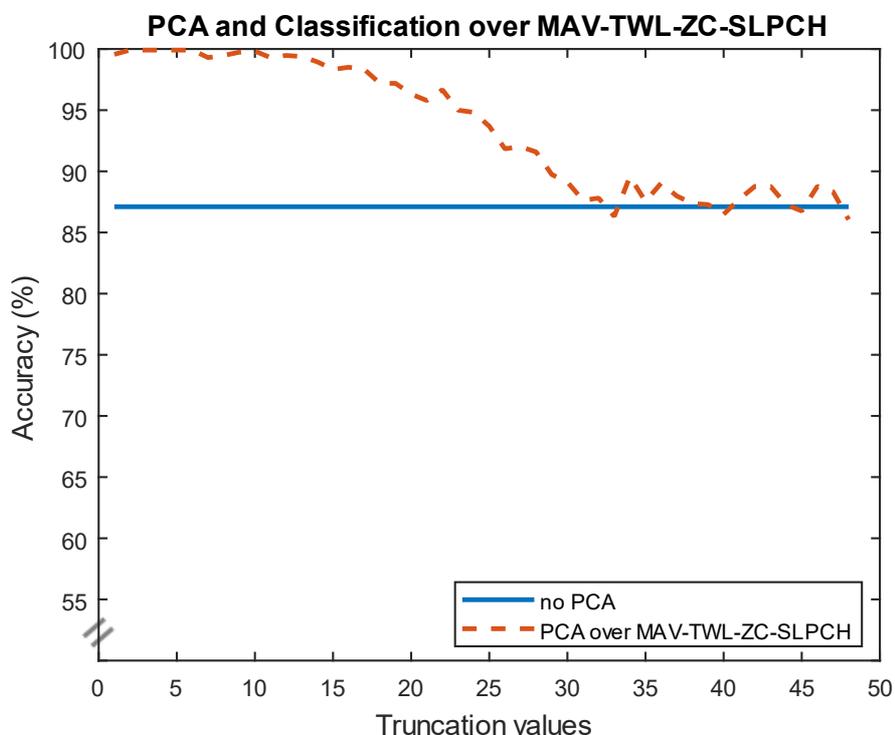


Figure 12 Red dashed line displays PCA performed with truncation values 1-48 over the four features MAV, TWL, ZC, and SLPCH. The blue continuous line shows the accuracy for classification of the featured data without applied PCA

As seen in the figure the red line goes towards the blue line with increasing truncation value. This is in alignment with the paired-sampled t-test between the case of no PCA and each truncation value. The t-tests showed that the accuracy results from truncation values 1-22 ($p \leq 0.023$) are significantly higher than the performance without any PCA applied. The achieved accuracy for no PCA was 87.7%, meanwhile, the best performance, 99.9%, was achieved for PCA with truncation values 2-6. This corresponds to a total improvement of 13.9%.

Figure 13 shows PCA performed over the four features and different subsets of the features, indicated for each subplot. Red dashed lines are PCA over all four features but classification only over two features, yellow dotted lines are instead PCA and classification over the same two features. The blue continuous lines are classification over indicated two features without the usage of PCA. Having two features instead of four decreases the number of possible truncation values from 48 to 24, hence the differences in length between different graphs within each plot.

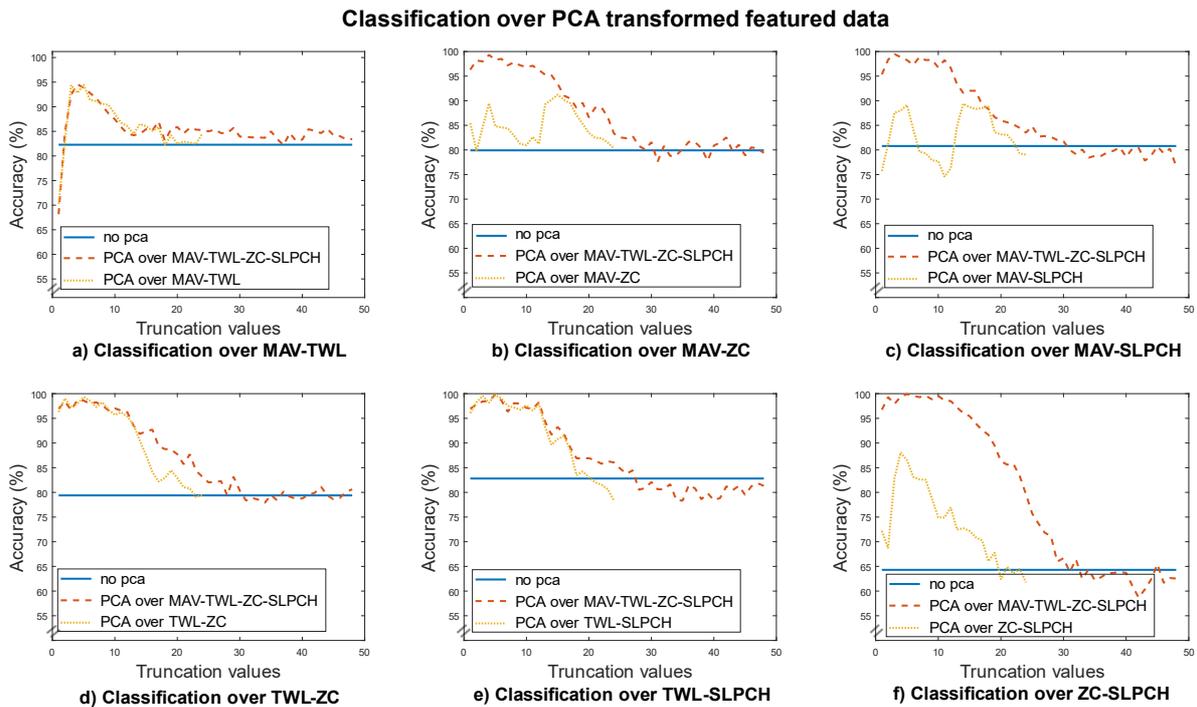


Figure 13 PCA performed over all four features MAV, TWL, ZC, and SLPCH, and different subsets of features. Red dashed lines are PCA over all four features but classification only over those indicated for each figure. Yellow dotted lines are PCA and classification over indicated features. Blue continuous lines are classification over indicated features without PCA

Like in Figure 12, a trend of decreasing accuracy for higher truncation values is also visible in Figure 13. The best accuracy was achieved for PCA over all four features, but classification only over ZC and SLPCH (Figure 13f). Although, PCA and classification over TWL and ZC (Figure 13d) and PCA and classification over TWL and SLPCH (Figure 13e) did achieve high accuracies for lower truncation values as well. Exact numbers for these and more cases can be found in Figure 26 and Figure 27 in Appendix.

4.2.2 ICA over features extracted from EMG signals

In Figure 14 ICA was performed over all four features MAV, TWL, ZC, and SLPCH (dotted yellow line), and over PCA transformed data over the same features (dashed red line). Removal of ICA components were performed with the help of a threshold of

60% of the maximum peak-to-peak value among all components. The results were compared with the case of neither PCA nor ICA (continuous blue line).

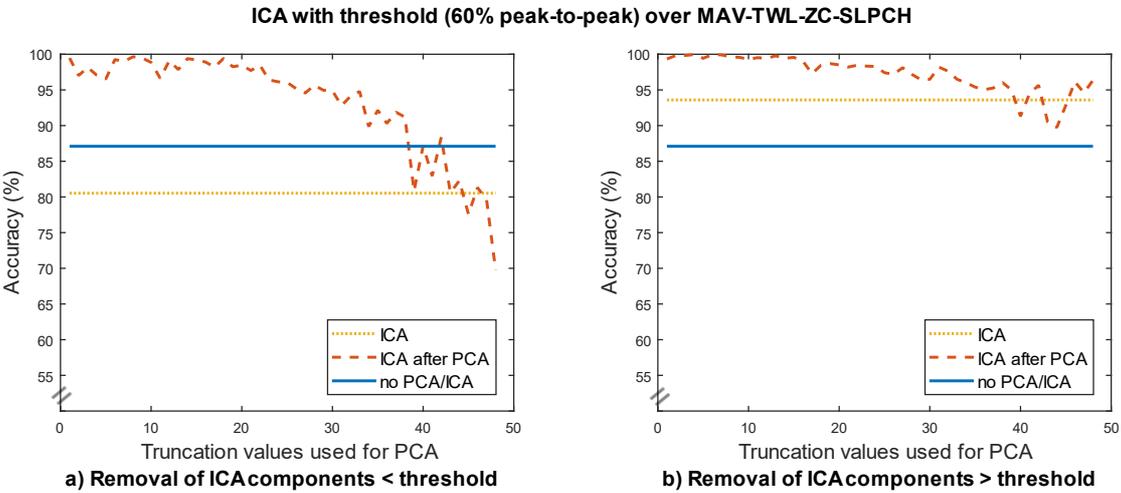


Figure 14 ICA performed over all four features MAV, TWL, ZC, and SLPCH (yellow dotted line), and over PCA transformed data over the same features (dashed red line). ICA components were removed with a threshold either smaller (case a) or bigger (case b) than 60% of the maximal peak-to-peak value

Figure 14a) shows how the removal of smaller components decreased the performance for ICA compared with no PCA or ICA. Removal of bigger components, on the other hand, seems to increase the performance, see Figure 14b). The red dashed graphs show that ICA after PCA of lower truncation values yields accuracies in the range 95%-100%. Although, the removal of bigger components results in more stable results meanwhile the results from the removal of smaller components fluctuate more.

4.3 Experiment 3

In Experiment 3, PCA was performed on featured data within windows of 200 ms.

Figure 15 shows the results for feature extraction each 2 ms and the cases:

- No PCA (feature extraction from overlapped data)
- PCA performed within a window but with U from Experiment 2
- PCA within a window and U calculated for each window

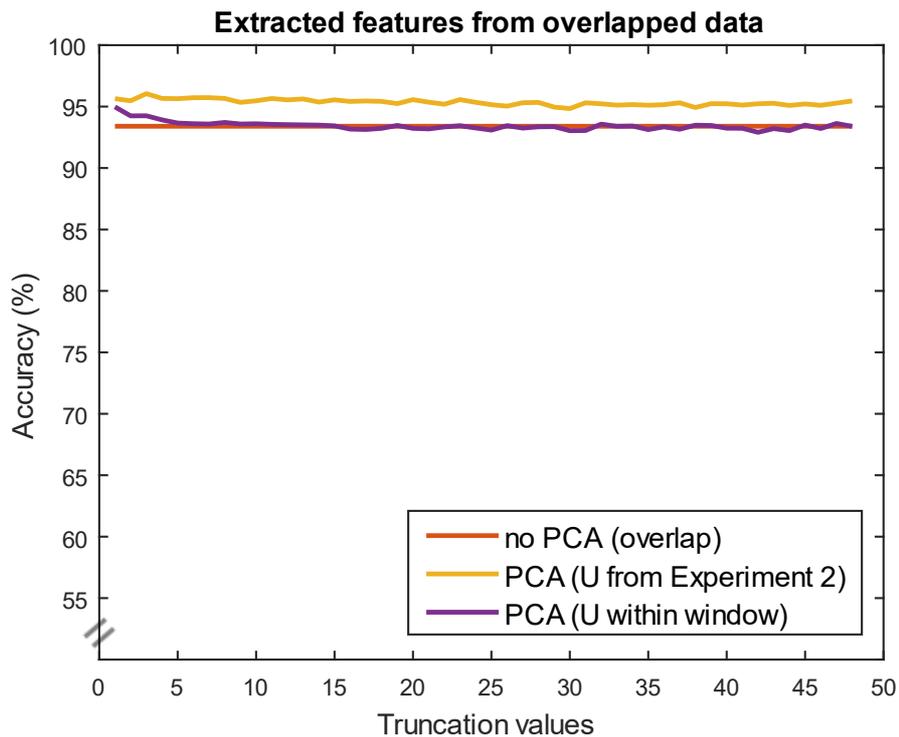


Figure 15 PCA over features extracted from overlapped data

The figure reveals that the truncation value has a minor impact on the accuracy. Although, higher accuracies around 95.4% were achieved when adding PCA on the overlapped data with the U calculated from each truncation value in experiment 2 (Figure 12), compared with 93.4% for overlapped data without PCA.

Figure 16 visualizes a boxplot for overlapped data with feature extraction each 2 ms (same case as above), and PCA with truncation value 5. The different boxplots display the results for:

- a) Entire PCA within windows
- b) PCA within windows, but U substituted with U from Experiment 2
- c) Entire PCA within windows but over a subset of session 9A
- d) PCA within windows over a subset of session 9A, but U substituted with an U calculated over the entire other subset of 9A
- e) PCA within windows but U substituted with U calculated over entire session 8A
- f) No PCA with overlap

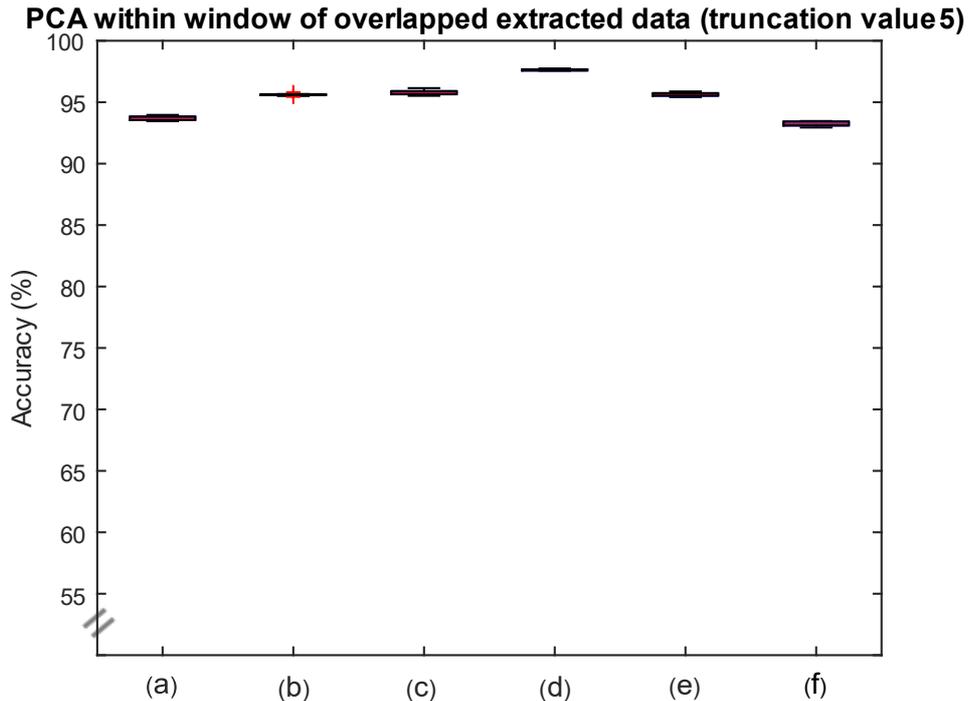


Figure 16 Boxplot for PCA with truncation value 5 and overlapped windows (2 ms between each feature extraction). a) Entire PCA within windows, b) PCA within windows, but U substituted with U from Experiment 2, c) Entire PCA within windows but over a subset of session 9A, d) PCA within windows over a subset of session 9A, but U substituted with an U calculated over the entire other subset of 9A, e) PCA within windows but U substituted with U calculated over entire session 8A, f) No PCA with overlap

A one-way ANOVA revealed significant differences between the different cases ($F(5,54) = 1160$, $p = 1.24E-53$). Repeated U from a subset of unseen data (d) performed best (97.6%), with an improvement of 4.72% over no PCA with overlap. Pairwise t-tests indicated this treatment as significantly better than all other cases ($p \leq 3.49E-16$).

To simulate a set-up that could be used for online testing, PCA was applied over four and two features, and the feature extraction rate decreased from 2 ms to 50 ms. A U was calculated from session 8A and a truncation value of four was used. Figure 17 shows the results, with the following boxes:

- a) Entire PCA within windows (4 features)
- b) PCA within windows but U substituted with an U calculated over entire session 8A (4 features)
- c) no PCA (4 features)
- d) Entire PCA within windows (TWL & SLPCH)
- e) PCA within windows but U substituted with an U calculated over entire session 8A (TWL & SLPCH)
- f) no PCA (TWL & SLPCH)

PCA within window of overlapped extracted data (truncation value 4)

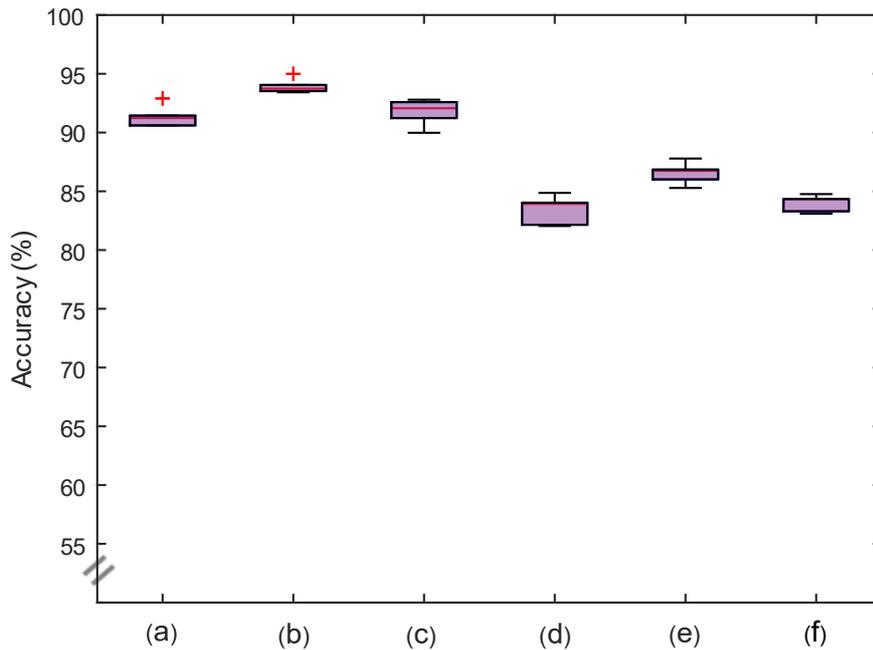


Figure 17 Boxplot for PCA with truncation value 4 and overlapped windows (50 ms between each feature extraction). a) Entire PCA within windows (4 features), b) PCA within windows but U substituted with an U calculated over entire session 8A (4 features), c) no PCA (4 features), d) Entire PCA within windows (TWL & SLPCH), e) PCA within windows but U substituted with an U calculated over entire session 8A (TWL & SLPCH), f) no PCA (TWL & SLPCH)

The figure shows that PCA over four features achieved higher accuracies than PCA over two features. The highest accuracy was achieved for case (b) with an accuracy of 93.9%. An improvement of 2.40%, compared to the case of no PCA over all four features (c) with an accuracy of 91.7%. One-way ANOVA revealed statistically significant differences between the cases ($F(5,54) = 239$, $p = 1.62E-35$), and pairwise t-tests indicated case (b) as statistically significantly different from all other cases ($p \leq 1.39E-04$).

4.4 Experiment 4

In Experiment 4 selected experiments with statistically significant improved accuracies for Dataset A were repeated, but now with Dataset B.

Figure 18 visualizes PCA over the four features MAV, TWL, ZC, and SLPCH and all their possible truncation values (dashed red line), compared with the case of no PCA (continuous blue line).

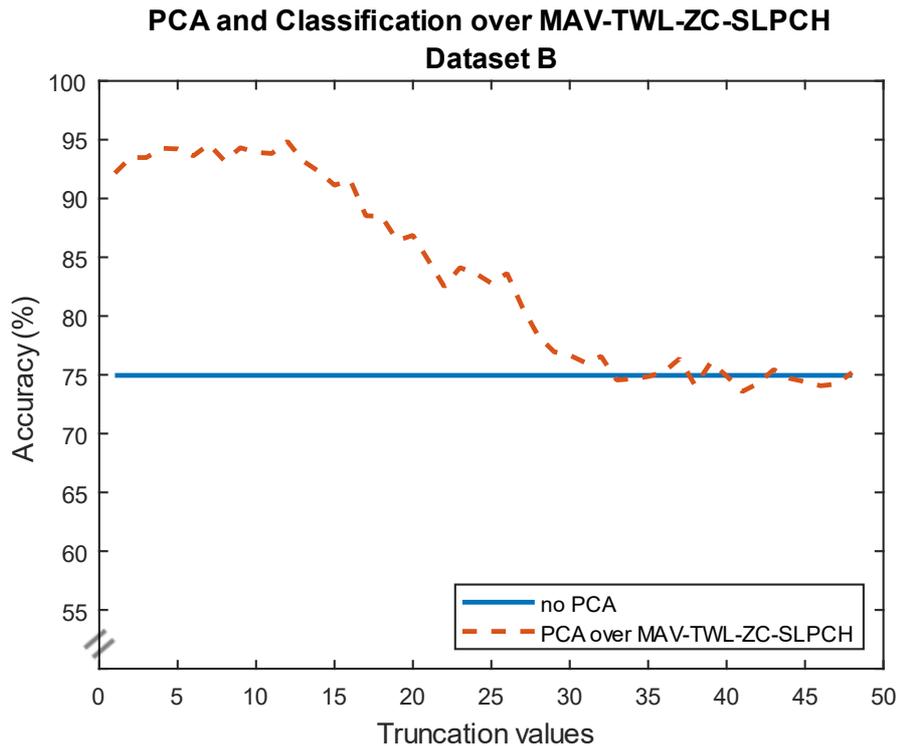


Figure 18 Red dashed line displays PCA performed with truncation values 1-48 over the four features MAV, TWL, ZC and SLPCH for Dataset B. Blue continuous line shows the accuracy for classification of the featured data without applied PCA

A series of paired-sampled t-tests indicated that truncation values 1-27 were significantly better than the case of no PCA ($p \leq 0.00317$). The highest accuracy of 94.9% was achieved for truncation value 12, an improvement of 26.7% compared with 74.9% for the case of no PCA.

The test with overlapped data and extraction each 2 ms was also repeated. The results can be seen in Figure 19 with the following boxes:

- a) Entire PCA within windows
- b) PCA within windows, but U substituted with U from Figure 18
- c) Entire PCA within windows but over a subset of session 9A
- d) PCA within windows over a subset of session 7B, but U substituted with an U calculated over the entire other subset of 7B
- e) PCA within windows but U substituted with U calculated over entire 6B
- f) No PCA with overlap

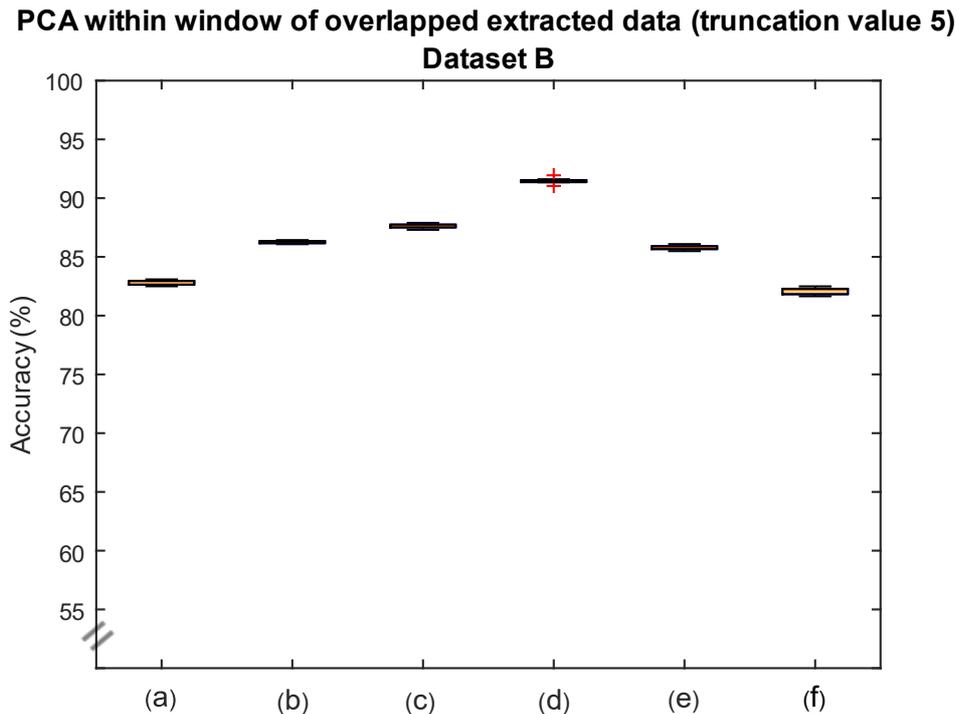


Figure 19 Boxplot for PCA with truncation value 5 and overlapped windows (2 ms between each feature extraction) for Dataset B. a) Entire PCA within windows, b) PCA within windows, but U substituted with U from Figure 18, c) Entire PCA within windows but over a subset of session 7B, d) PCA within windows over a subset of session 9A, but U substituted with an U calculated over the entire other subset of 7B, e) PCA within windows but U substituted with U calculated over entire session 6B, f) No PCA with overlap

In similarity to the same test performed with Dataset A (Figure 16) the cases with PCA over all four features did perform better than the case of no PCA with overlap (f). One-way ANOVA revealed significant differences between the different cases ($F(5,54) = 2510$, $p = 1.34E-62$). Once again, the highest accuracy was achieved for U from a subset of unseen data, 91.4% compared with 82.1% in the case of no PCA with overlap. A pairwise t-test indicated this case (d) as statistically significant different from all the other groups ($p \leq 2.04E-18$).

Proceeding with the last test where feature extraction was performed each 50 ms, the results can be seen in Figure 20 with the following boxes:

- a) Entire PCA within windows (4 features)
- b) PCA within windows but U substituted with an U calculated over entire session 6B (4 features)
- c) no PCA (4 features)
- d) Entire PCA within windows (TWL & SLPCH)
- e) PCA within windows but U substituted with an U calculated over entire session 6B (TWL & SLPCH)
- f) no PCA (TWL & SLPCH)

**PCA within window of overlapped extracted data (truncation value 4)
Dataset B**

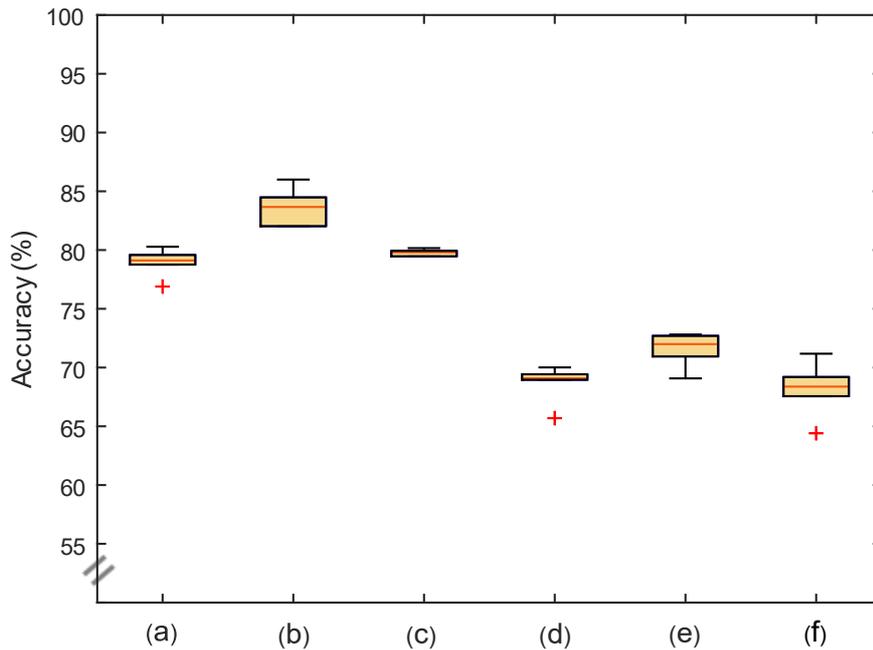


Figure 20 Boxplot for PCA with truncation value 4 and overlapped windows (50 ms between each feature extraction) for Dataset B. a) Entire PCA within windows (4 features), b) PCA within windows but U substituted with an U calculated over entire session 6B (4 features), c) no PCA (4 features), d) Entire PCA within windows (TWL & SLPCH), e) PCA within windows but U substituted with an U calculated over entire session 6B (TWL & SLPCH), f) no PCA (TWL & SLPCH)

Once more the results follow the same distribution as for Dataset A (compare with Figure 17). The best results, 83.6%, were achieved for case (b), PCA over all four features and with U calculated over entire session 6B. An increase of 4.76% compared to the case of no PCA over all four features (c) and 79.8%. One-way ANOVA indicated statistically significant differences between the cases ($F(5,54) = 179$, $p = 2.40E-32$), and pairwise t-tests comparing (b) to all the other cases did confirm a significantly improvement ($p \leq 3.01E-05$).

4.5 Experiment 5

In Experiment 5 the EMG signals before and after application of PCA were visually compared. An RMS amplitude-based index was used as a metric to support the visual comparison.

Figure 21 shows the results for Dataset A and the signal from one implanted electrode for the cases:

- (i) PCA reconstruction with no overlap
- (ii) PCA with unseen U from a subset of data and feature extraction each 2 ms
- (iii) PCA with U calculated over entire session 8A and feature extraction each 50 ms

The yellow line in the top graph shows the Z-score normalized data, the blue lines the extracted data for each feature, and the red lines the PCA reconstructions for each feature.

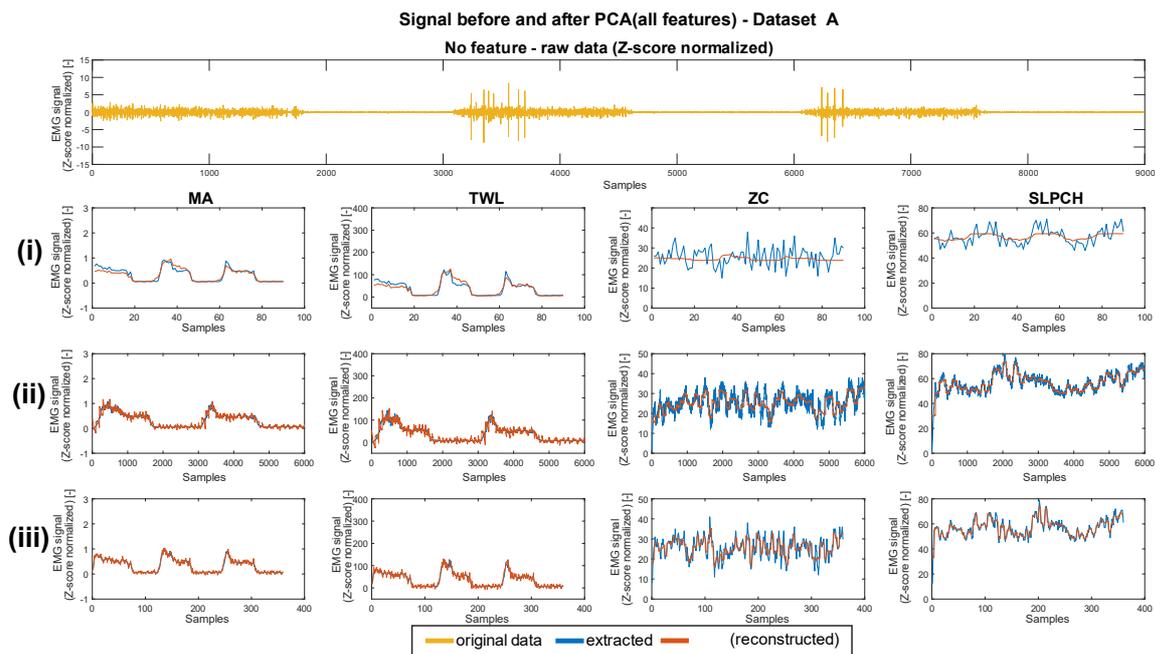


Figure 21 Signals before and after PCA for Dataset A. The top graph shows the Z-score normalized data meanwhile the graphs below show the featured data MAV, TWL, ZC, and SLPCH. Case (i) display PCA reconstruction with no overlap, (ii) displays PCA with unseen U from a subset of data with extraction each 2 ms, and (iii) PCA with U calculated over entire session 8A and feature extraction each 50 ms

Visually no clear differences between the cases (i), (ii), and (iii), are identified, although case (i) has a slightly different appearance for ZC and SLPCH compared with (ii) and (iii). The RMS metric used supported the visual observation and pointed out ZC for case (i) as the signal with biggest amplitude decrease when comparing the signal before and after PCA, a decrease with 2.90%. All cases had slightly decreased

RMS ratios, except from MAV and TWL for case (ii) and (iii) which had amplitudes of the same magnitude before and after application of PCA. (Table 2 in Appendix presents the RMS ratio for all cases.)

The corresponding cases but for Dataset B and one of its signals, are presented in Figure 22.

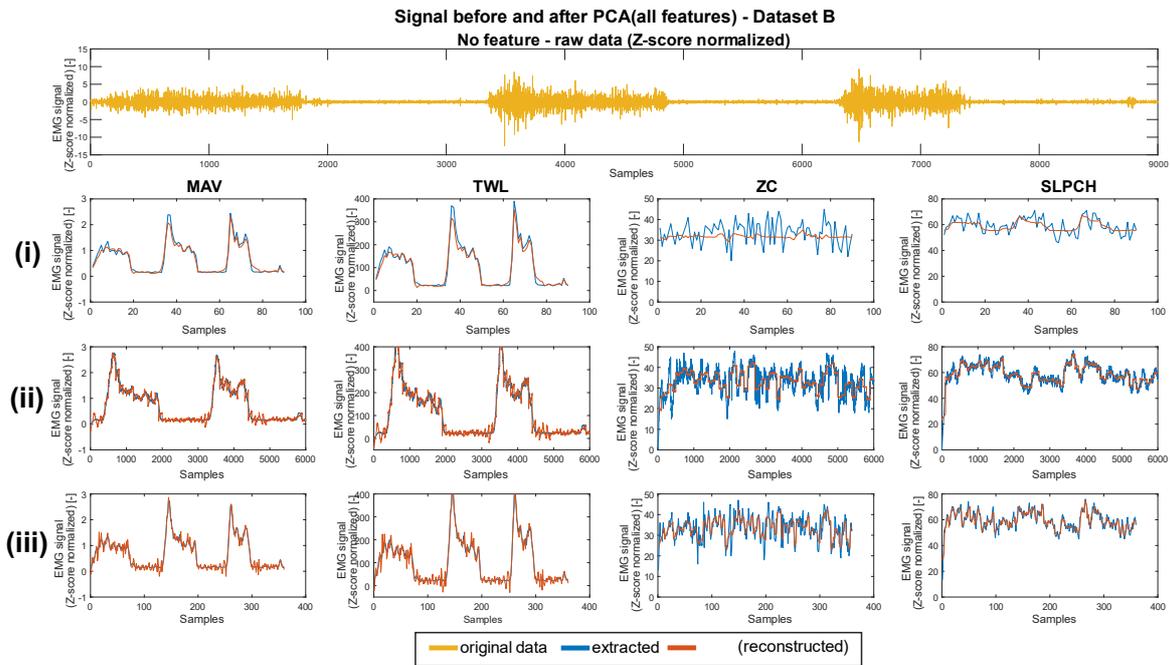


Figure 22 Signals before and after PCA for Dataset B. The top graph shows the Z-score normalized data meanwhile the graphs below show the featured data MAV, TWL, ZC, and SLPCH. Case (i) display PCA reconstruction with no overlap, (ii) displays PCA with unseen U from a subset of data with extraction each 2 ms, and (iii) PCA with U calculated over entire session 6B and feature extraction each 50 ms

In similarity to Dataset A, there are no clear differences between the cases (i), (ii), and (iii). Neither comparison between the two different datasets highlights any outstanding differences. The RMS index did also in this case indicate ZC for case (i) as the signal with biggest amplitude decrease when comparing the signal before and after PCA, a decrease with 2.40%. (Table 3 in Appendix shows the RMS results for all different cases in Dataset B.)

In Figure 23 the same cases are studied once again, but instead of showing the reconstructed signals, the graphs now visualize the U used in the PCA. The figure includes U both from Dataset A and Dataset B.

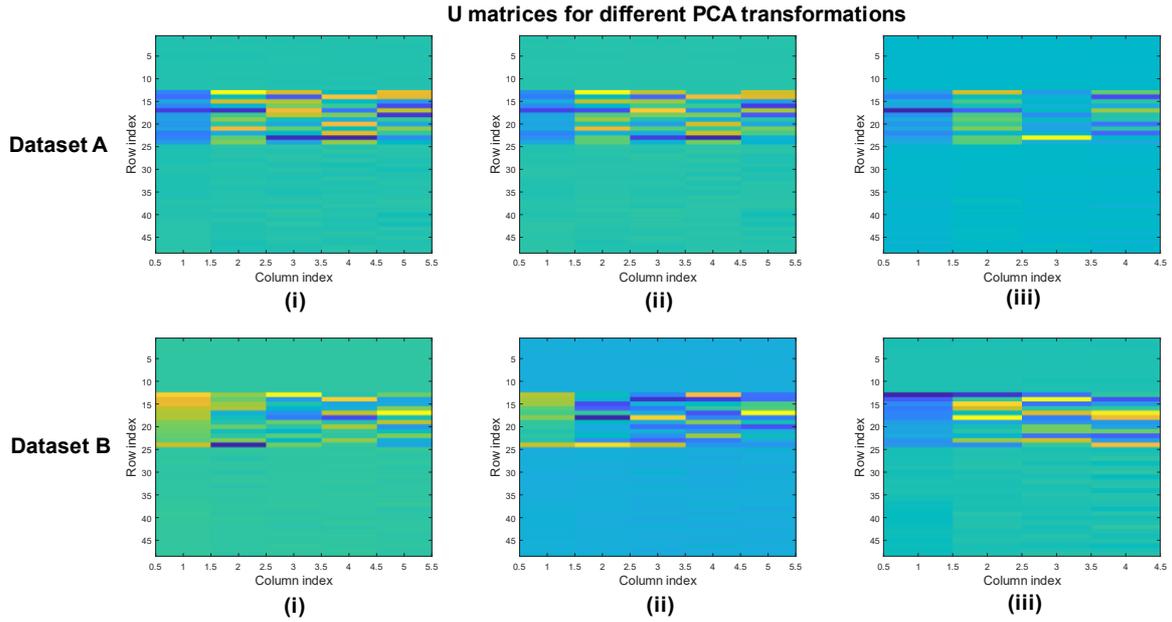


Figure 23 U matrices for different transformations. Top row for Dataset A and lower row for Dataset B. Case (i) display U from PCA with no overlap, (ii) displays U from a subset of data with extraction each 2 ms, and (iii) U for extraction each 50 ms

Studying the different datasets and their cases it can be concluded that all cases have similar characteristics. Row indices 1-12 and 25-48 corresponds to MAV, ZC, and SLPCH, and have values of similar magnitudes and with small changes. Row indices 13-24 on the other hand, describing the TWL feature appear somewhat chaotic in comparison. Studying the absolute values of U , at least 89.2% of the signal content were constituted of the TWL features and less than 0.700% of MAV, for each individual case. (Table 4 in Appendix show the ratios of signal content for each feature.)

5 Discussion

This project aimed to investigate and evaluate if PCA and ICA applied on EMG signals recorded from implanted electrodes could improve prosthetic control performance. In this section, it is discussed what the results point to, limitations within the performed project, and what practical actions could or should take place in the future.

5.1 Interpretations of the results

Several different experiments were designed and performed, and it was shown that movement predictability accuracy could be improved for both PCA and ICA under certain circumstances. Most interesting were the findings that increased accuracy could be achieved after PCA over four features extracted each 50 ms, compared to the same cases but without applied PCA. These results were especially interesting as they were a consequence of the experimental approach most promising for online usage. This is due to its application of PCA within small windows of 200 ms, and its decrease in feature extraction rate compared to the other experimental approaches. As feature extraction slows down the signal processing it can be expected that a too high feature extraction rate (such as 2 ms) will decrease the speed of the prosthetic control, affecting the user experience and functionality of the prosthesis.

Other interesting findings were the ones achieving the highest accuracies. Those came from the PCA approach in Experiment 2 where PCA was applied over all samples from the four features MAV, TWL, ZC, and SLPCH at once. Unfortunately, the method is non-feasible for online usage as PCA was applied over the entire session, corresponding to 18 s times the number of performed movement (16 or 18), which is a huge time window in context of prosthetic control. Despite, the approach being unrealistic for online usage it demonstrated that PCA could improve the accuracy which motivated the work yielding the results achieved for decreased feature extraction rate.

The achieved results support the hypothesis that PCA is a promising tool for the reduction of crosstalk, which also agrees with previously conducted research [31]. Although, the achieved results are hard to directly compare with earlier performed research. This as the main metric for this project has been accuracy, while most studies reviewed have used other types of indices. Nevertheless, the RMS metric used in Experiment 5 indicated small decreases in amplitude when comparing the signal before and after crosstalk removal. This is in accordance with other results where removal of crosstalk has been shown to decrease the total signal content [25], [39], [40]. The decreases identified in this project were $\leq 2.90\%$, quite small results in comparison with other research where crosstalk could make up to 30% of the total signal content. Possible reasons for the big differences are that the other studies

reviewed used other muscles, and surface EMG where the crosstalk is expected to be bigger in comparison to the usage of implantable electrodes.

Going back to Experiment 1 and Experiment 2 one major learning is that the performance of ICA and PCA increased significantly when applied over featured data instead of the raw data. These improvements include PCA and ICA over both four and two features. Unfortunately, the same trend for two features could not be seen when extracting features for each 50 ms (compare Figure 13 with Figure 17 and Figure 20). The fewer features needed the less time to perform feature extraction which in turn accelerates the signal processing speed.

Overall, the results achieved in this project were more promising for PCA than ICA. This can be connected to the fact that the PCA algorithm itself ranked the most important components, making it easy to exclude the components of least importance. Meanwhile, an own created condition had to be set for the removal of the ICA components. A challenging problem as there were only small differences in the component's appearance. Hence, the only condition tested was for different thresholds where components of certain sizes were removed. It is not unreasonable to believe that a better method for component removal could be developed. Although, improved accuracy was achieved when using ICA which supports the hypothesis and previous research that ICA can be used to discriminate crosstalk [10].

Another interesting observation was the fact that the truncation values for PCA had a big impact on the achieved accuracies in Experiment 1 and Experiment 2. Experiment 3 on the other hand, gave no reason to believe that the truncation was of big importance. Rather seemed like the big improvement was due to feature extraction from overlapped data. An explanation for this could be that when overlapped data was used, there was also an overlap between the data in the training and test sets due to the current code implementation. Hence, the MLP was trained on a subset of data which later was used when computing the accuracies. Hence the results from this project and their meaning should be interpreted carefully as they are not yet tested online. Nevertheless, the positive results indicate that it is an interesting approach that could be worth trying online. To enable online testing the code developed must be adapted for the continuous gathering of data and implemented on the ALC, used for prosthetic control.

A bit surprising was the outcome that no major differences between the signals from different electrodes, after PCA or ICA, were identified. Hence, the visualization of only one signal for each dataset in Experiment 5. Before the execution of the experiments, it was expected to see visual differences between the signals from different electrodes. This is due to different surgical techniques and muscle sources, which implies different conditions for the appearance and recording of myoelectric activity [41]. Another outcome worth mentioning is the similarities between the U 's visualized in Experiment 5, both within and between each dataset. Although, an

outstanding appearance for the TWL features, this was nothing visually observable when viewing the signals after the application of PCA.

5.2 Limitations within the project

One of the major limitations within this project is the fact that no index has been used which can confirm whether the accuracy changes, before and after PCA and ICA, are due to the removal of crosstalk or the removal of other undesired data. Therefore, it was decided to mainly measure the performance of the algorithms through the accuracy of pattern recognition. This as the literature review did not reveal any index as entirely trustable or suitable for this specific project. Improved pattern recognition, on the other hand, will probably always be a positive result no matter what type of information has been removed. Hence, no results from either the accuracy or RMS metric entirely confirm that removal of crosstalk is the actual cause of improved pattern recognition in this project. If not derived from the removal of crosstalk the improvements could maybe be derived from the removal of other types of disturbances in the recorded signals, such as noise or artifacts. Or, only by the fact that the PCA algorithm compresses information. This compression maybe focuses the EMG signals to their most important characteristics which makes the decoding into different movements easier.

Another thought to have in mind is that this project does not aim to compare different classifiers, such as KNN, LDA, MLP, nor different normalizations of the raw data. The purpose of the classifier choice was to enable the comparison of the data before and after signal processing. In similarity, normalization was used to enable visual interpretation of the data. Although, the choices made may have influenced the results, which should be kept in mind.

The participants chosen for this study have different types of amputations, surgical techniques for electrode implantation, and different muscles conserved within the arm. Although, this is not a full representation of the variations within the group of active prosthesis users. In addition, the experiments so far have only been conducted for a certain set of movements recorded in a laboratory environment. The laboratory environment contributes with better conditions than those given in daily living. Therefore, the results from this project and what it can bring for the future should be interpreted carefully and not applied to the general prosthesis user and in their everyday living.

5.3 Future work

As there is a need for improved prosthetics techniques, the findings in this project have several aspects which could be interesting to explore in future work. One of the most interesting findings is to confirm whether the approach of feature extracting

each 50 ms can improve the accuracy during online testing. This as a positive online test can implicate the possibility for implementation in real-life situations.

It would also be interesting to compare different U 's. Considering their similarities shown in Experiment 5 it is not entirely out of the scope, that it would be possible to manually create an artificial U optimized for PCA over four features.

Another point of view interesting to explore is the potential of these experiments in usage adapted for daily living and training for prosthetic control. This as the quality of the signals captured by the implanted electrodes can vary over time. It can be expected to have more stable signals a certain time after surgery, compared with the period following directly after surgery. If there exists an U such as pattern recognition that can be performed with positive results, despite having clear signals, this could perhaps reduce the time needed for prosthetics training. Or, decrease the time between surgery and when prosthetics training can begin.

There are also certain things related to the PCA algorithm directly worth investigating deeper. The first example is the truncation values. In this project, several truncation values have been tested, despite the knowledge that literature exists stating recommended truncation values to use [9], [42]. These recommendations are based on the amount of variance to keep through the transformation, common ratios to conserve are 90% or 99% of the variance. An interesting approach might be to explore how the truncation values yield the best results in this report corresponding to the "recommended" truncation values. Although, it could be assumed that quite high variance has been conserved in Experiment 2-4 (corresponding to the recommendations), based on the visualization of the signals before and after PCA in Experiment 5, as no major differences were observed. The second experiment which could be interesting is to not only create an artificial U , but also to create an artificial US to compare potential differences.

To summarize the project has shown that both PCA and ICA have the potential to improve prosthetic control. Although, the findings of improved accuracy after applying PCA within windows over extracted features are the most interesting. It is recommended to investigate that approach deeper with the motivation that it potentially could improve the performance of EMG prosthetics, and with that also increase the life quality for those people that today or in the future are or will be suffering from limb loss.

6 Conclusion

To summarize, there exists a need for improved prosthesis performance to avoid prosthetic abandonment in the future. The new technique shows promising results but there are still several factors needed to be improved, such as the control. Crosstalk and other undesired content in measured EMG signals aggravate decoding of the muscles' pattern, resulting in non-optimal pattern recognition.

In this project, signal processing through PCA and ICA showed improved results when applied over featured data, both for Participant A and Participant B. The best results were achieved when PCA was applied over the four features MAV, TWL, ZC, and SLPCH, for an entire session of 18 s, and pattern recognition was calculated over windows of 200 ms. The accuracy for Participant A increased from 87.7% to 99.9%, and for Participant B from 74.9% to 94.9%. In the test aiming to simulate a scenario more realistic to real-time usage, the features were extracted each 50 ms and PCA applied within windows of 200 ms. Participant A achieved an accuracy of 93.9% and Participant B an accuracy of 83.6%.

The results do not reveal whether these improvements are consequences of the removal of crosstalk specifically. Although, these are improvements significant enough to make it interesting to explore how the same method would perform online, and not only over offline data. Regardless of what the signal processing removes, these results could in the long-term yield better life quality for people suffering from limb loss, if the approach also yields improved results online.

References

- [1] S. M. Engdahl, B. P. Christie, B. Kelly, A. Davis, C. A. Chestek, and D. H. Gates, "Surveying the interest of individuals with upper limb loss in novel prosthetic control techniques," *J. Neuroeng. Rehabil.*, vol. 12, no. 1, pp. 1–11, 2015, doi: 10.1186/s12984-015-0044-2.
- [2] E. Biddiss and T. Chau, "Upper-limb prosthetics: Critical factors in device abandonment," *Am. J. Phys. Med. Rehabil.*, vol. 86, no. 12, pp. 977–987, 2007, doi: 10.1097/PHM.0b013e3181587f6c.
- [3] M. A. LeBlanc, "Innovation and improvement of body-powered arm prostheses: A first step," *Clin. Prosthetics Orthot.*, vol. 9, no. 1, pp. 13–16, 1985, [Online]. Available: http://oandplibrary.com/cpo/pdf/1985_01_013.pdf.
- [4] P. J. Kyberd *et al.*, "Survey of Upper-Extremity Prosthesis Users in Sweden and the United Kingdom," *JPO J Prosthetics Orthot*, vol. 19, no. 2, pp. 55–62, 2007.
- [5] M. Espinosa and D. Nathan-Roberts, "Understanding Prosthetic Abandonment," *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, vol. 63, no. 1, pp. 1644–1648, Nov. 2019, doi: 10.1177/1071181319631508.
- [6] L. Mesin, "Crosstalk in surface electromyogram: literature review and some insights," *Phys. Eng. Sci. Med.*, vol. 43, no. 2, pp. 481–492, 2020, doi: 10.1007/s13246-020-00868-1.
- [7] I. Talib, K. Sundaraj, C. K. Lam, J. Hussain, and M. A. Ali, "A review on crosstalk in myographic signals," *Eur. J. Appl. Physiol.*, vol. 119, no. 1, pp. 9–28, 2019, doi: 10.1007/s00421-018-3994-9.
- [8] M. Ortiz-Catalan, B. Hakansson, and R. Branemark, "An osseointegrated human-machine gateway for long-term sensory feedback and motor control of artificial limbs," *Sci. Transl. Med.*, vol. 6, no. 257, Oct. 2014, doi: 10.1126/scitranslmed.3008933.
- [9] S. L. Brunton and K. J. Nathan, *Data-driven science and engineering: machine learning, dynamical systems, and control*, 1st ed. Cambridge University Press, 2019.
- [10] Y. Kim, S. Stapornchaisit, M. Miyakoshi, N. Yoshimura, and Y. Koike, "The Effect of ICA and Non-negative Matrix Factorization Analysis for EMG Signals Recorded From Multi-Channel EMG Sensors," *Front. Neurosci.*, vol. 14, no. December, pp. 1–10, 2020, doi: 10.3389/fnins.2020.600804.
- [11] A. Delorme and S. Makeig, "EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J. Neurosci. Methods*, vol. 134, no. 1, pp. 9–21, 2004, doi: 10.1016/j.jneumeth.2003.10.009.
- [12] G. J. Tortora and B. Derrickson, *Introduction to the Human Body*, Ninth Edit. John Wiley & Sons, 2012.
- [13] M. B. I. Reaz, M. S. Hussain, and F. Mohd-Yasin, "Techniques of EMG signal analysis: Detection, processing, classification and applications," *Biol. Proced. Online*, vol. 8, no. 1, pp. 11–35, 2006, doi: 10.1251/bpo115.
- [14] C. J. De Luca, A. Adam, R. Wotiz, L. D. Gilmore, and S. H. Nawab, "Decomposition of surface EMG signals," *J. Neurophysiol.*, vol. 96, no. 3, pp. 1646–1657, 2006, doi: 10.1152/jn.00009.2006.
- [15] M. Ortiz-Catalan, R. Brånemark, B. Håkansson, and J. Delbeke, "On the viability of implantable electrodes for the natural control of artificial limbs:

- Review and discussion," *Biomed. Eng. Online*, vol. 11, 2012, doi: 10.1186/1475-925X-11-33.
- [16] T. A. Kuiken *et al.*, "Targeted muscle reinnervation for real-time myoelectric control of multifunction artificial arms," *JAMA - J. Am. Med. Assoc.*, vol. 301, no. 6, pp. 619–628, 2009, doi: 10.1001/jama.2009.116.
- [17] M. G. Urbanek, Z. Baghmanli, J. D. Moon, K. B. Sugg, B. Nicholas, and P. S. Cederna, "Quantification of Regenerative Peripheral Nerve Interface Signal Transmission," *Suppl. to Plast. Reconstr. Surg.*, vol. 130, no. 5, pp. 55–56.
- [18] M. G. Urbanek *et al.*, "Development of a Regenerative Peripheral Nerve Interface for Control of a Neuroprosthetic Limb," *Biomed Res. Int.*, vol. 2016, 2016, doi: 10.1155/2016/5726730.
- [19] R. A. Mezzarane and A. F. Kohn, "A method to estimate EMG crosstalk between two muscles based on the silent period following an H-reflex," *Med. Eng. Phys.*, vol. 31, no. 10, pp. 1331–1336, 2009, doi: 10.1016/j.medengphy.2009.09.005.
- [20] M. E. Busse, C. M. Wiles, and R. W. M. van Deursen, "Muscle co-activation in neurological conditions," *Phys. Ther. Rev.*, vol. 10, no. 4, pp. 247–253, 2005, doi: 10.1179/108331905X78915.
- [21] T. Wojtara, F. Alnajjar, S. Shimoda, and H. Kimura, "Muscle synergy stability and human balance maintenance," *J. Neuroeng. Rehabil.*, vol. 11, no. 1, pp. 1–9, 2014, doi: 10.1186/1743-0003-11-129.
- [22] S. Shahid, J. Walker, G. M. Lyons, C. A. Byrne, and A. V. Nene, "Application of higher order statistics techniques to EMG signals to characterize the motor unit action potential," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 7, pp. 1195–1209, 2005, doi: 10.1109/TBME.2005.847525.
- [23] A. Nene, C. Byrne, and H. Hermens, "Is rectus femoris really a part of quadriceps? Assessment of rectus femoris function during gait in able-bodied adults," *Gait Posture*, vol. 20, no. 1, pp. 1–13, 2004, doi: 10.1016/S0966-6362(03)00074-2.
- [24] D. A. Winter, A. J. Fuglevand, and S. E. Archer, "Crosstalk in surface electromyography: Theoretical and practical estimates," *J. Electromyogr. Kinesiol.*, vol. 4, no. 1, pp. 15–26, 1994, doi: 10.1016/1050-6411(94)90023-X.
- [25] T. J. Koh and M. D. Grabner, "Cross talk in surface electromyograms of human hamstring muscles," *J. Orthop. Res.*, vol. 10, no. 5, pp. 701–709, 1992, doi: 10.1002/jor.1100100512.
- [26] D. Farina, M.-F. Lucas, and C. Doncarli, "Optimized Wavelets for Blind Separation of Nonstationary Surface Myoelectric Signals," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 1, pp. 78–86, 2008, doi: 10.1109/TBME.2007.897844.
- [27] M. M. Lowery, N. S. Stoykov, and T. A. Kuiken, "A simulation study to examine the use of cross-correlation as an estimate of surface EMG cross talk," *J. Appl. Physiol.*, vol. 94, no. 4, pp. 1324–1334, 2003, doi: 10.1152/jappphysiol.00698.2002.
- [28] P. Zipp, "Effect of electrode geometry on the selectivity of myoelectric recordings with surface electrodes," *Eur. J. Appl. Physiol. Occup. Physiol.*, vol. 50, no. 1, pp. 35–40, 1982, doi: 10.1007/BF00952242.
- [29] D. A. Winter, A. J. Fuglevand, and S. E. Archer, "Crosstalk in surface electromyography: Theoretical and practical estimates," *J. Electromyogr. Kinesiol.*, vol. 4, no. 1, pp. 15–26, 1994, doi: [https://doi.org/10.1016/1050-6411\(94\)90023-X](https://doi.org/10.1016/1050-6411(94)90023-X).

- [30] L. Mesin, "Optimal spatio-temporal filter for the reduction of crosstalk in surface electromyogram," *J. Neural Eng.*, vol. 15, no. 1, p. 16013, Jan. 2018, doi: 10.1088/1741-2552/aa8f03.
- [31] A. Matran-Fernandez, E. Mastinu, R. Poli, M. Ortiz-Catalan, and L. Citi, "Crosstalk Reduction in Epimysial EMG Recordings from Transhumeral Amputees with Principal Component Analysis," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, vol. 2018-July, pp. 2124–2127, 2018, doi: 10.1109/EMBC.2018.8512645.
- [32] G. R. Naik, S. E. Selvan, M. Gobbo, A. Acharyya, and H. T. Nguyen, "Principal Component Analysis Applied to Surface Electromyography: A Comprehensive Review," *IEEE Access*, vol. 4, pp. 4025–4037, 2016, doi: 10.1109/ACCESS.2016.2593013.
- [33] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural Networks*, vol. 13, pp. 411–430, 2000, doi: 10.7819/rbgn.v19i63.1905.
- [34] R. H. Chowdhury, M. B. I. Reaz, M. A. Bin Mohd Ali, A. A. A. Bakar, K. Chellappan, and T. G. Chang, "Surface electromyography signal processing and classification techniques," *Sensors (Switzerland)*, vol. 13, no. 9, pp. 12431–12466, 2013, doi: 10.3390/s130912431.
- [35] G. R. Naik and D. K. Kumar, "Estimation of independent and dependent components of non-invasive EMG using fast ICA: Validation in recognising complex gestures," *Comput. Methods Biomech. Biomed. Engin.*, vol. 14, no. 12, pp. 1105–1111, 2011, doi: 10.1080/10255842.2010.515211.
- [36] M. Ortiz-Catalan, R. Brånemark, and B. Håkansson, "BioPatRec: A modular research platform for the control of artificial limbs based on pattern recognition algorithms," *Source Code Biol. Med.*, vol. 8, pp. 1–18, 2013, doi: 10.1186/1751-0473-8-11.
- [37] E. Mastinu, P. Doguet, Y. Botquin, B. Hakansson, and M. Ortiz-Catalan, "Embedded System for Prosthetic Control Using Implanted Neuromuscular Interfaces Accessed Via an Osseointegrated Implant," *IEEE Trans. Biomed. Circuits Syst.*, vol. 11, no. 4, pp. 867–877, 2017, doi: 10.1109/TBCAS.2017.2694710.
- [38] B. Hudgins, P. Parker, and R. N. Scott, "A New Strategy for Multifunction Myoelectric Control," *IEEE Trans. Biomed. Eng.*, vol. 40, no. 1, pp. 82–94, 1993, doi: 10.1109/10.204774.
- [39] J. P. van Vugt and J. G. van Dijk, "A convenient method to reduce crosstalk in surface EMG. Cobb Award-winning article, 2001.," *Clin. Neurophysiol.*, vol. 112, no. 4, pp. 583–92, 2001, [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S1388245701004825%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/11275529>.
- [40] C. J. De Luca and R. Merletti, "Surface myoelectric signal cross-talk among muscles of the leg," *Electroencephalogr. Clin. Neurophysiol.*, vol. 69, no. 6, pp. 568–575, 1988, doi: [https://doi.org/10.1016/0013-4694\(88\)90169-1](https://doi.org/10.1016/0013-4694(88)90169-1).
- [41] A. L. Lowe and N. V. Thakor, "Cut wires: The Electrophysiology of Regenerated Tissue," *Bioelectron. Med.*, vol. 7, no. 1, pp. 1–15, 2021, doi: 10.1186/s42234-021-00062-y.
- [42] M. Gavish and D. L. Donoho, "The optimal hard threshold for singular values is $4/\sqrt{3}$," *IEEE Trans. Inf. Theory*, vol. 60, no. 8, pp. 5040–5053, 2014, doi: 10.1109/TIT.2014.2323359.

Appendix

Figure 24 displays classification with LDA, KNN, and MLP of PCA over raw, max-min normalized, and Z-score normalized data.

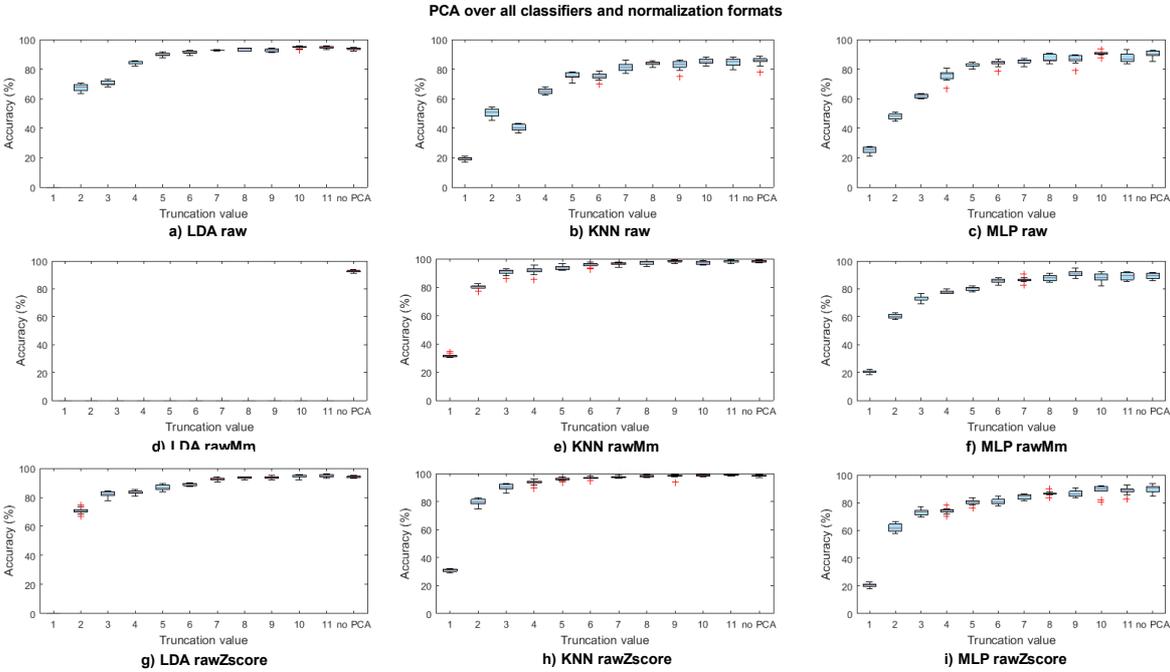


Figure 24 PCA over different data formats and classified by different algorithms (Experiment 1)

Figure 25 display classification with LDA, KNN and MLP of ICA over raw, max-min normalized, and Z-score normalized data.

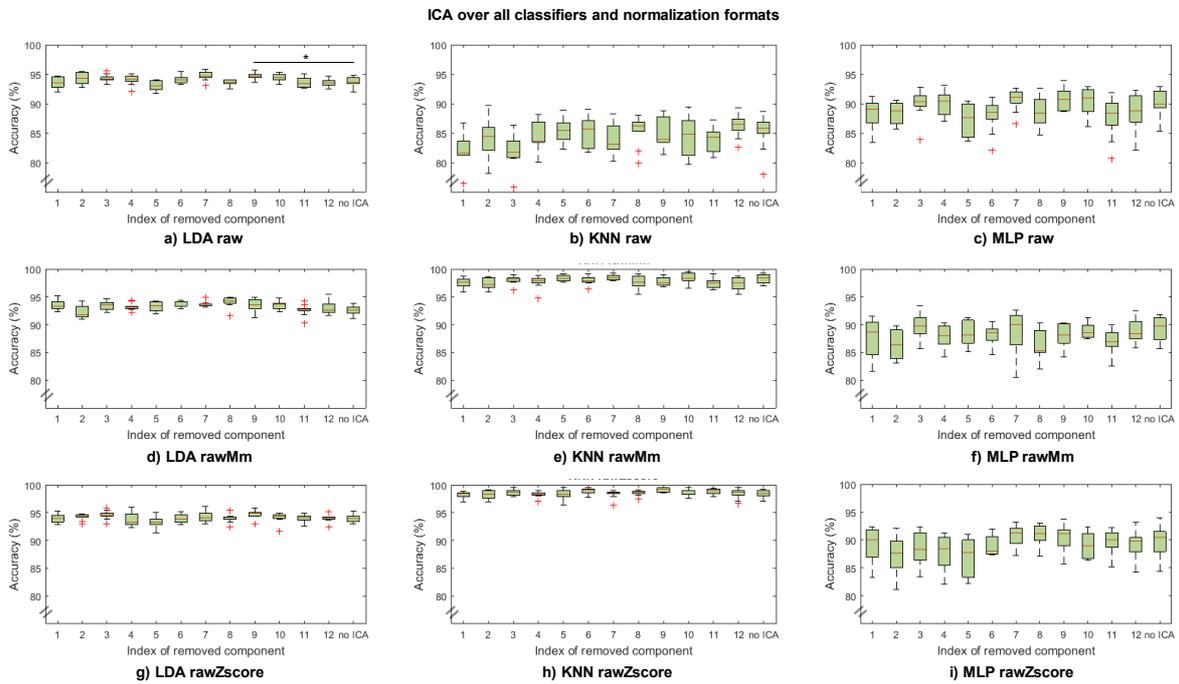


Figure 25 ICA over different data formats and classified by different algorithms (Experiment 1)

Figure 26 shows accuracies for PCA performed over four features, and accuracies calculated over the different subsets of the features.

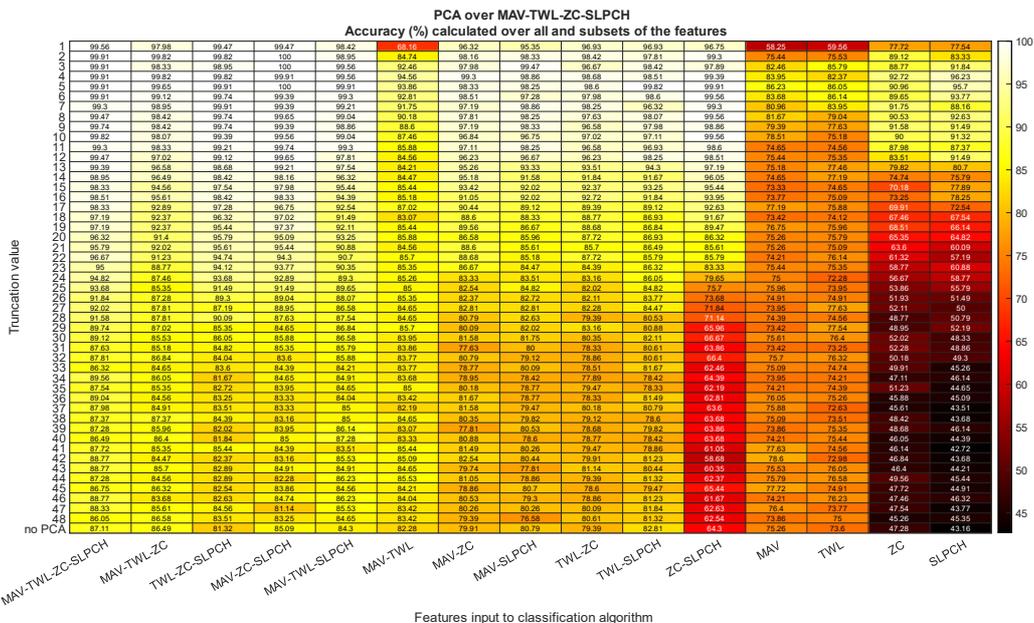


Figure 26 Accuracies calculated over different subsets of features from PCA transformed data (Experiment 2)

Figure 27 displays PCA performed, and accuracy calculated over two features at a time.

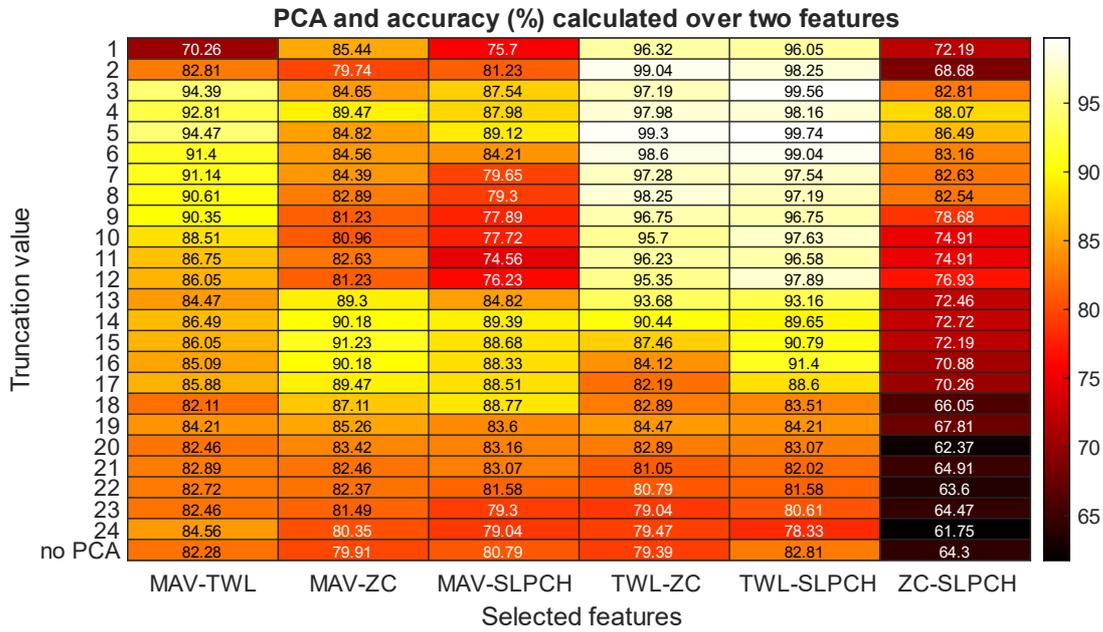


Figure 27 PCA and accuracies calculated over two features (Experiment 2)

Table 2 presents the RMS ratios before and after PCA for the cases of Dataset A presented in Experiment 4.

Table 2 Root mean square amplitude-based ratio for signal after PCA divided with the signal before PCA (%) (Experiment 4)

Dataset A	MAV [%]	TWL [%]	ZC [%]	SLPCH [%]
(i)	98.9	99.1	97.1	99.4
(ii)	100	100	99.0	99.9
(iii)	100	100	99.1	99.9

Table 3 presents the RMS ratios before and after PCA for the cases of Dataset B presented in Experiment 4.

Table 3 Root mean square amplitude-based ratio for signal after PCA divided with the signal before PCA (%) (Experiment 4)

Dataset B	MAV [%]	TWL [%]	ZC [%]	SLPCH [%]
(i)	99.0	99.1	97.6	99.9
(ii)	99.8	99.7	99.3	99.9
(iii)	101	100	99.4	99.9

Table 4 presents the ratios of signal content for each feature within U , for both Dataset A and Dataset B.

Table 4 The ratios of U content from each feature, dataset and case (%) (Experiment 5)

Dataset	Feature	(i) [%]	(ii) [%]	(iii) [%]
A	MAV	0.700	0.680	0.670
	TWL	92.8	92.3	89.2
	ZC	2.88	3.26	3.13
	SLPCH	3.55	3.80	6.99
B	MAV	0.610	0.600	0.640
	TWL	9.29	92.4	90.1
	ZC	2.80	0.311	3.96
	SLPCH	3.71	3.88	5.33