

Electromyogram (EMG) Signal Analysis: Extraction of a Novel EMG Feature and
Optimal Root Difference of Squares (RDS) Processing in Additive Noise

by

Kiriaki Rajotte

A Thesis

Submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

In partial fulfillment of the requirements for the

Degree of Master of Science

in

Electrical and Computer Engineering

December 2019

APPROVED:

Dr. Edward A. Clancy, Research Advisor

Dr. Xinming Huang, Committee Member

Prof. Stephan Bitar, Committee Member

Abstract

Electromyogram signals generated by human muscles can be measured on the surface of the skin and then processed for use in applications such as prostheses control, kinesiology and diagnostic medicine. Most EMG applications extract an estimate of the EMG amplitude, defined as the time-varying standard deviation of EMG, $EMG\sigma$. To improve the quality of $EMG\sigma$, additional signal processing techniques, such as whitening, noise reduction and additional signal features can be incorporated into the $EMG\sigma$ processing. Implementation of these additional processing techniques improve the quality of the processed signal but at the cost of increased computational complexity and required calibration contractions.

Whitening filters are employed to temporally decorrelate data so that the samples are statistically independent. Different types of whitening filters, linear and adaptive, and their performance have been previously studied in (Clancy and Hogan) and (Clancy and Farry). The linear filter fails at low effort levels and the adaptive filter requires a calibration every time electrodes are removed and reapplied. With the goal of avoiding the disadvantages of the previous whitening filter approaches, the first signal processing technique studied herein developed a universal fixed whitening filter using the ensemble mean of the power spectrum density of EMG recordings from the 64 subjects available in an existing data set. Performance of the EMG to torque model with the universal fixed whitening filter was computed to be 4.8% maximum voluntary contraction (MVC); this is comparable to the 4.84 %MVC error computed for the adaptive whitening filter. The universal fixed whitening filter preserves the performance of the adaptive filter but need not be calibrated for each electrode.

To optimize noise reduction, the second signal processing technique studied derived analytical models using the resting EMG data. The probability density function of the rest contractions was observed to be very close to a Gaussian distribution, showing only a 1.6% difference when compared to a Gaussian distribution. Once the models were developed, they were used to prove that the optimal subtraction of the noise variance is to compute the root of the difference between the signal squared and noise variance (RDS). If this result would lead to a negative value, it must be set to zero; $EMG\sigma$ cannot contain negative components. Once the RDS was proven to be the optimal noise subtraction, it was implemented on 0% MVC and 50%

MVC data. The RDS processing has a considerable impact on lower level contractions (0% MVC), but not on higher level contractions (50% MVC), as expected.

The third signal processing technique involved the creation of a new EMG feature from four individual signal features. Different techniques were used to combine $EMG\sigma$, zero crossings (ZC), slope sign changes (SSC) and waveform length (WL) into a single new EMG feature that would be used in an end application, such as the modeling of torque about the elbow or prosthesis control. The new EMG feature was developed to reduce the variance of the traditional $EMG\sigma$ only feature and to eliminate the need for calibration contractions. Five different methods of combination were attempted, but none of the new EMG features improved performance in EMG to torque model.

Acknowledgements

I am deeply grateful for all the individuals that guided me through the course of this project. Without the support of the following people, the work presented below would not be possible. First, I would like to thank my thesis advisor, Professor Edward A. Clancy for the continuous encouragement and support starting from the day we met. Without his guidance and kindness, I would not have been able to navigate through my academics.

I would also like to thank my colleagues from the Laboratory for Sensory and Physiologic Signal Processing (L(SP)²), Ziling Zhu, Jianan Li, Haopeng Wang and He Wang. My colleagues provided me with continuous support and were always willing to help. The collaborative efforts for this project helped me to improve my technical and software skills.

Finally, I would like to thank my family, Mom, Dad, Johanna and Joseph. You always reminded me that there isn't anything that I can't do if I set my mind to it. You were there through all of it and frequently reminded me that the possibilities are endless.

Table of Contents

<i>Abstract</i>	<i>II</i>
<i>Acknowledgements</i>	<i>IV</i>
<i>Table of Contents</i>	<i>V</i>
<i>Table of Figures</i>	<i>VI</i>
<i>I. Introduction</i>	<i>1</i>
A. Project Statement	1
B. Contributions	1
C. Contents of the Thesis	2
<i>II. Background</i>	<i>4</i>
A. Anatomy and Physiology of Skeletal Muscle	4
B. Characteristics of Surface EMG	7
C. Traditional Electromyography Processing: Time-Varying Standard Deviation of EMG	7
D. EMG to Torque Modeling	10
E. Experimental Data	11
<i>III. Methods</i>	<i>14</i>
A. Raw EMG Processing	14
B. Combination of the Features	21
1) Combination of Uncorrelated Features using Averaging	21
2) Combination of the Features through Averaging	28
3) Combination of Features according to their Statistics	32
4) Combination of Features using Inverse Variance Weights	39
5) Fixed Dynamics of the EMG to Torque Model	44
<i>IV. Root Difference of Squares Processing</i>	<i>52</i>
A. Derivation of the Optional Estimate of $EMG\sigma$ with RDS Processing	52
B. Application of RDS Processing	70
<i>V. Discussion</i>	<i>77</i>
A. Discussion of the New EMG Feature Development	77
B. Discussion of the Applied Root Difference of Squares	78
<i>VI. Fixed Whitening Filter Development</i>	<i>80</i>
<i>VII. Conclusion</i>	<i>83</i>
<i>VIII. References</i>	<i>85</i>

Table of Figures

<i>Figure 1 Structure of Skeletal Muscle (National Cancer Institute)</i>	5
<i>Figure 2 Action Potential and Contraction of Skeletal Muscle (OpenStax)</i>	6
<i>Figure 3 Probability Density of EMG (Clancy and Hogan)</i>	8
<i>Figure 4 Steps Involved in EMGσ</i>	9
<i>Figure 5 Set-up for Data Collection from Previous Experiments (Dai, Bardizbanian and Clancy)</i>	12
<i>Figure 6 Magnitude Response of the 4th Order Butterworth Filter</i>	15
<i>Figure 7 Frequency Response of the IIR Notch Filter at 60 Hz</i>	17
<i>Figure 8 Magnitude Response of the Fixed Whitening Filter</i>	18
<i>Figure 9 Example EMG Standard Deviation Estimate without Smoothing Filter</i>	19
<i>Figure 10 Example Zero Crossing Computation</i>	20
<i>Figure 11 Example Slope Sign Change Computation</i>	21
<i>Figure 12 Initially Proposed Process of Feature Combination</i>	22
<i>Figure 13 Uncorrelated Features Combined by Averaging</i>	27
<i>Figure 14 Procedure used to Average the Features Together</i>	29
<i>Figure 15 Combined Features from Averaging, Unwhitened</i>	29
<i>Figure 16 Combined Features from Averaging, Whitened</i>	30
<i>Figure 17 Torque Model Performance Expressed in %MVC, Unwhitened</i>	31
<i>Figure 18 Torque Model Performance Expressed in %MVC, Whitened</i>	31
<i>Figure 19 Second Feature Combination Process</i>	33
<i>Figure 20 Combined Features from Fixed Weights, Unwhitened</i>	36
<i>Figure 21 Combined Features using Fixed Weights, Whitened</i>	37
<i>Figure 22 Torque Model Performance Expressed in %MVC, Unwhitened</i>	38
<i>Figure 23 Torque Model Performance Expressed in %MVC, Whitened</i>	38
<i>Figure 24 Procedure for Combining the Features using the Inverse Variance Weights</i>	40
<i>Figure 25 Combined Features using Inverse Variance Weights, Unwhitened</i>	41
<i>Figure 26 Combined Features using Inverse Variance Weights, Whitened</i>	42
<i>Figure 27 Torque Model Performance Expressed in %MVC, Unwhitened</i>	43
<i>Figure 28 Torque Model Performance Expressed in %MVC, Whitened</i>	43
<i>Figure 29 Fixed Dynamics Approach</i>	44
<i>Figure 30 FIR Filter Magnitude Responses and Fixed Shape</i>	45
<i>Figure 31 Median Frequency Response of the Four Features</i>	46
<i>Figure 32 Desired and Designed FIR Filter Response for the Feature Pairs</i>	47
<i>Figure 33 Performance of the EMG Torque Model, 0th Order, Universal FIR Filter, Computed Gains, Whitened</i>	48
<i>Figure 34 Mean and Standard Deviation of the Model Gains per Electrode for EMGσ</i>	49
<i>Figure 35 Mean and Standard Deviation of the Model Gains per Electrode for Zero Crossing Feature</i>	49
<i>Figure 36 Mean and Standard Deviation of the Model Gains per Electrode for Slope Sign Changes Feature</i>	50
<i>Figure 37 Mean and Standard Deviation of the Model Gains per Electrode for Waveform Length Feature</i>	50
<i>Figure 38 Performance of the EMG Torque Model, 0th Order, Universal FIR Filter, Universal Gains, Whitened</i>	51
<i>Figure 39 Procedure used to Compare the Effects of Whitening with and without RDS</i>	71
<i>Figure 40 Whitening Comparison without RDS, N = 64 Subjects per Method</i>	72
<i>Figure 41 Whitening Comparison with RDS, N = 64 Subjects per Method</i>	72
<i>Figure 42 Comparison of Data without RDS to Data with RDS, Unwhitened</i>	73
<i>Figure 43 Comparison of Data without RDS to Data with RDS, Fixed Whitened</i>	74
<i>Figure 44 Comparison of Data without RDS to Data with RDS, Adaptive Whitened</i>	74
<i>Figure 45 Results of the Kruskal-Wallis Test (Median, 25th Percentile and 75th Percentile)</i>	76

I. Introduction

A. Project Statement

Like many physical systems, the human body relies on electrical potentials to generate control signals for different biological systems, such as the muscular system. The electrical activity of the muscular system, known as electromyography (EMG), can be measured on the surface of the skin and processed for use in the diagnosis of muscular disorders, kinesiology, ergonomics and prosthesis control (Sornmo and Laguna). At the surface of the skin, the measured signal is susceptible to noise and must be further processed to extract the information content required for the end application. The goal of the work presented in this thesis is to simplify the existing EMG processing techniques as well as to improve performance of the traditional EMG feature when relating EMG to torque about a joint. To achieve this goal, a universal whitening filter was developed, rest contractions were modelled to determine the optimal method of removing additive noise and additional signal features were incorporated into the EMG feature to reduce the EMG-torque error.

B. Contributions

The work presented in this thesis was a collaborative effort between myself, Haopeng Wang and He Wang under the advisement of Professor Edward Clancy. The fundamental goal of this collaborative effort was to develop techniques to improve EMG feature performance in EMG-torque models and to simplify the calibration processing required to compute EMG features. To achieve improvements in performance and simplify the calibration processing required, the following four approaches were studied:

1. Develop and test the use of a universal whitening filter derived from the ensemble average of the individual subject-specific calibrated filters;
2. Study the probability density function and power spectrum of EMG signals at rest to improve the noise rejection of rest contractions;
3. Study the impact of whitening and offset subtraction on 0% maximum voluntary contractions (MVC) and 50% MVC data;

4. Develop and test a novel EMG feature that includes additional signal features (zero crossings, slope sign changes and waveform length) to reduce the variance of the EMG standard deviation only estimate.

My major contributions to the collaborative efforts involved the implementation of the additional signal features into a new EMG feature as well as study of whitening and offset subtraction on the 0% MVC and 50% MVC constant-force, constant-posture data. Multiple approaches to combining four EMG features into a new EMG feature were implemented and tested.

C. Contents of the Thesis

Traditionally, the time varying standard deviation of the EMG signal is used as an estimate of the EMG signal amplitude. In an effort to improve the quality of the processed EMG by decreasing the variance of the EMG σ feature, three additional signal features, zero crossings (ZC), slope sign changes (SSC), and waveform length (WL) were incorporated to create a new EMG feature. Additionally, a fixed whitening filter was developed and implemented in the EMG feature processing and a noise offset subtraction was mathematically modelled and tested.

Previous work that involved the use of features implemented the features in the modelling of EMG to torque or classification analysis. In previous works, the features are individually incorporated into the EMG-torque processor i.e. each was entered as a separate input to the system identification which requires calibration contractions and can increase the computational complexity of the processing. To simplify the computational complexity and attempt to remove the need for calibration contractions, the work completed in this thesis focuses on the combination of features prior to the EMG to torque model. To include the four features in a real time operating system such as a prosthetic, the process of combining four features must be simplified so that the computation time is minimized and no noticeable delay results. Additionally, it is desired to remove the need for contraction data to calibrate each time a combination of the features is implemented.

The contents of this thesis include three traditional-style thesis chapters: a Background, Methods and Discussion section specific to the work that I contributed to the collaborative effort. Citations specific to these three chapters are found in the References section. Additionally, included in this thesis are the author's copies of a conference paper and a journal paper published

based on the results of the collaborative efforts. The published works included in this thesis contain their own references.

II. Background

Communication between the nervous system and skeletal muscles is conducted by electrical impulses that travel from the brain to the muscle, carrying instructions to the muscle or a group of muscles. The study of the electrical activity found in skeletal muscles is called electromyography. The field of electromyography is used in the diagnosis of neuromuscular disorders, control of prostheses and in biomechanics and rehabilitation. Before EMG can be used in a specific application, there is some signal processing that must be completed to take the raw, noisy signal captured on the body to create a useful signal for the end application. The application at hand will dictate the type of processing required to condition the raw EMG.

A. Anatomy and Physiology of Skeletal Muscle

The field of electromyography is constrained to the electrical activity of skeletal muscles found throughout the body. Skeletal muscles are responsible for creating motion, posture and changes in the body's position. The size and shape of skeletal muscle vary depending on where the muscle is in the body, but the general composition remains the same across people. Skeletal muscles are typically divided into and referred to as a muscle group such as the biceps or triceps, but each group can be further divided down to the cellular level. Each individual muscle cell contributes to the contraction of the entire muscle.

A complete muscle can be broken down into three layers. On the smallest layer is the individual muscle cell, which is known as a muscle fiber. The two types of muscle fibers, type I and type II are classified based on the contraction speed of the fiber; type I muscle fibers are slow-twitch and type II are fast-twitch. Slow-twitch muscles require more time to achieve their maximum tension than fast-twitch muscle fibers. The striated or striped appearance in skeletal muscle is created by the laying of the muscle fibers parallel to one another. Multiple muscle fibers grouped together are collectively known as a fascicle. The complete muscle is composed of multiple fascicles. The layered structure of the muscle allows for the relaxation and contraction across an entire muscle or group of muscles. Figure 1 below shows an individual muscle and its structure. In addition to the layers of the muscle, the connective tissue is also labelled in Figure 1. (Fox)

Structure of a Skeletal Muscle

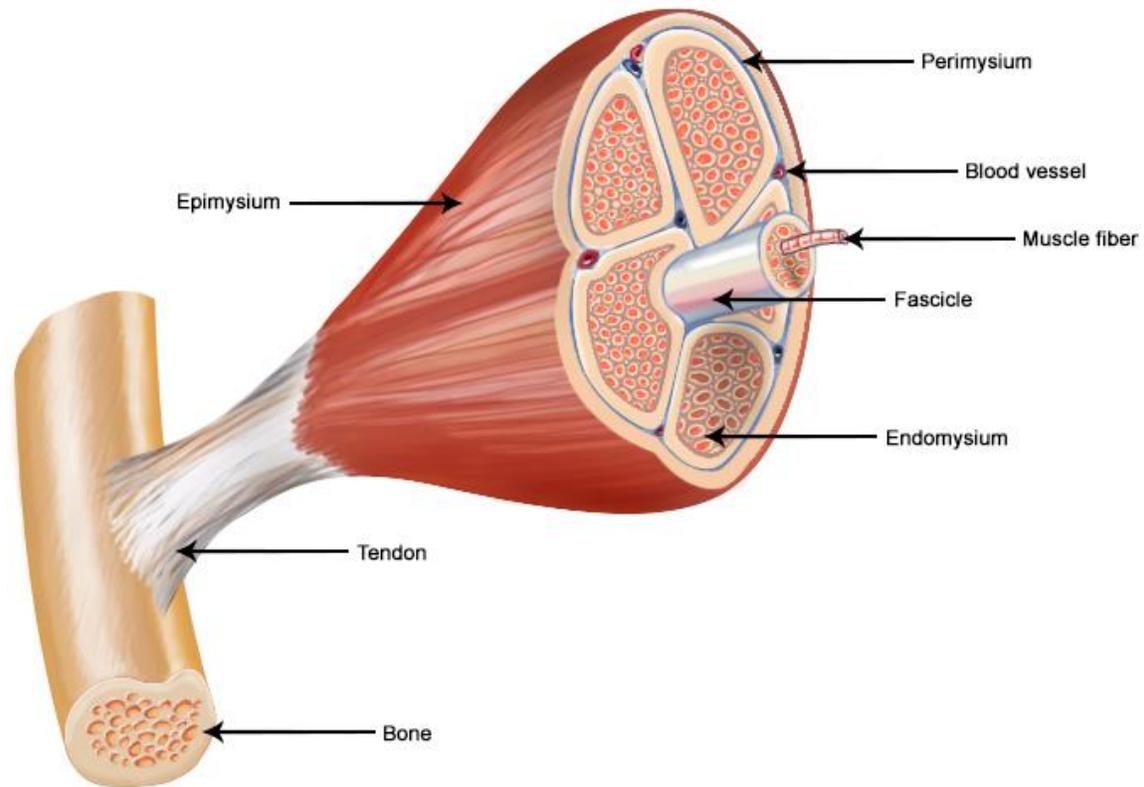


Figure 1 Structure of Skeletal Muscle (National Cancer Institute)

The central and peripheral nervous system control the contraction of muscles by sending an electrical impulse from the brain to the muscle fiber via efferent and motor neurons. Each individual muscle cell contains ions (charged particles) within the cell and outside the cell. At rest, the inside of a cell tends to be negatively charged with respect to the outside of the cell, creating a difference in potential across the membrane. When an impulse travels from the nervous system to the muscle cells (a stimulus), ion channels will open and allow for the movement of ions across the cell membrane. As charge moves across the cell's membrane, the inside of the cell becomes more positively charged with respect to the outside of the cell, resulting in a positive voltage across the cell membrane. Over time, the cell will return to its resting potential. The event in which the cell's potential changes as a response to a stimulus is called an action potential. Action potentials can be observed in muscle cells, cardiac cells and neurons.

In a muscle fiber, an action potential forces the cell to leave its resting potential of roughly -85 mV and move towards roughly 30 mV. Once the initial stimulus is no longer present, the cell will return to its resting potential. The flow of specific ions across the cell's membrane forces a contraction in the muscle. The relationship between the action potential and the resulting contraction is shown below in Figure 2. The top plot shows the change in potential across the cell membrane while the bottom plot shows the resulting contraction as time elapses. There is a short time delay, approximately 20 ms, between the end of the action potential (as the cell returns to its resting potential) and the beginning of a contraction.

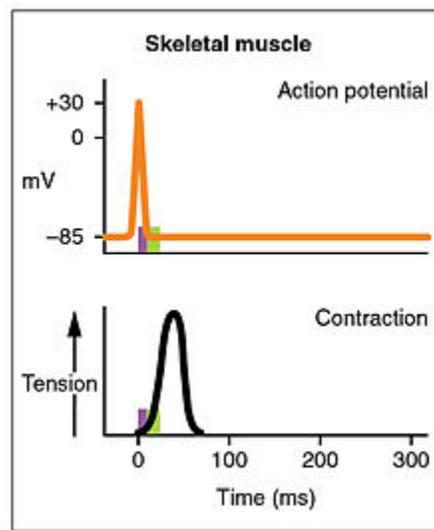


Figure 2 Action Potential and Contraction of Skeletal Muscle (OpenStax)

At the muscle, the motor neuron connects to the muscle fiber, creating the neuromuscular junction. The neuromuscular junction is located approximately half way down the muscle to minimize the action potential's propagation time across the muscle. The motor neuron and its corresponding fibers create a motor unit. Within a motor unit, the temporal and spatial summation of the individual action potentials is called the motor unit action potential (MUAP).

The signal that is measured at the surface of the skin is the sum of the individual MUAPs. The number of MUAPs that are detected depends on the surface area of the electrodes as well as the electrode placement. Additionally, muscle fibers from different motor units may be overlapping which increases the number of MUAPs detected at the skin's surface. At the surface of the skin, the MUAPs are grouped together and cannot be distinguished without further processing.

B. Characteristics of Surface EMG

On the surface of the skin, the measured EMG signal contains activity from multiple motor units, as well as interference from power line, motion artifact, the electrode skin interface, and noise present in the electronics of the analog front end and the data converter. The measured surface EMG signal has an amplitude ranging from 0 mV_{pk-pk} to 10mV_{pk-pk} and a frequency bandwidth ranging from 0 Hz to 2000 Hz, with most of the signal power falling below 600 Hz (De Luca). In the higher frequency range of 600 Hz to 2000 Hz, there is significant noise present and it is challenging to distinguish the signal from the noise. Further signal processing is required to condition the signal measured on the surface of the skin into a signal that can be used in an EMG processor.

C. Traditional Electromyography Processing: Time-Varying Standard Deviation of EMG

The signal measured with surface electrodes is a collection of motor unit firings from a larger area than the needle electrodes. At the surface of the skin, the signal appears to be an amplitude modulated random signal with a probability density that falls between a Laplacian density and a Gaussian Density as shown in Figure 3. The amplitude of the Gaussian (or Laplacian) random variable is modulated by the effort level. As the effort level increases, more motor units are recruited, and the recruitment rate increases which results in an increase in the amplitude of the measured EMG signal.

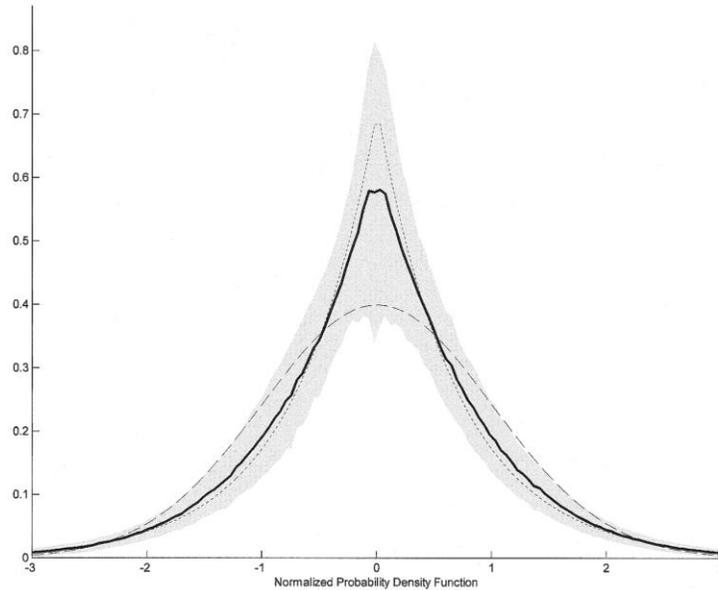


Figure 3 Probability Density of EMG (Clancy and Hogan)

Because the signal can be modeled as either a Laplacian or Gaussian density, the characteristics of these distributions can be used to better model the EMG amplitude and to determine an estimator. To use the raw EMG signal that was measured on the surface of the skin in practical applications, it must be further processed into an EMG feature. The traditional EMG feature is computed as the time varying standard deviation of the measured EMG signal. To develop an EMG processor, the EMG signal is modelled as an amplitude modulated, zero mean random process. Under this assumption, a maximum likelihood estimate of its standard deviations can be developed.



Figure 4 Steps Involved in EMG σ

Figure 4 outlines the steps used to compute an estimate of EMG σ . The first step in processing the raw EMG signal is to filter the signal to remove any unwanted artifacts. To remove motion artifacts, a high pass filter with a cutoff frequency between 10 Hz and 20 Hz is applied to the input signal. To remove interference due to power line, a notch filter is applied at the fundamental, 60 Hz, and its harmonics. Once the raw EMG signal is filtered, an optional whitening filter can be applied to temporally decorrelate the data and produce statistically independent samples. The detection phase rectifies the EMG signal and then a low pass filter can be applied to smooth the data. The final step is to relinearize the data so that the units of the signal are the back to the units of EMG signal. Once the EMG time-varying standard deviation has been computed, it can be used in applications such as prosthesis control or EMG to force modelling. (Clancy, Morin and Merletti)

D. EMG to Torque Modeling

To model the relationship between EMG and joint torque, it is assumed that $EMG\sigma$ can be extracted from the raw EMG and that the torque is a function of $EMG\sigma$. The process of relating EMG to torque can be broken down into three stages, the computation of $EMG\sigma$, training of the model and testing of the model. $EMG\sigma$ is computed from the raw EMG data to produce a signal that can be used in an end application which in this case is a model of torque about the elbow from EMG. For the EMG to torque modelling, a linear least squares regression is computed to model torque about the elbow using EMG measured from the biceps (flexion) and triceps (extension). The training stage of the model is necessary to compute the fit coefficients of the dynamic, linear model.

$$T_{Ext}[n] = f_0 \cdot \hat{s}_F[n] + f_1 \cdot \hat{s}_F[n-1] + \dots + f_L \cdot \hat{s}_F[n-L] - e_0 \cdot \hat{s}_E[n] - e_1 \cdot \hat{s}_E[n-1] - \dots - e_L \cdot \hat{s}_E[n-L] \quad (1)$$

The parameter n represents the sample index and L represents the dynamic model order of the system. The flexion fit coefficients are denoted as f_L and the extension fit coefficients are denoted as e_L . $T_{Ext}[n]$ represents the torque signal values at the n^{th} sample. $\hat{s}_E[n]$ represents the extension $EMG\sigma$ values at n and $\hat{s}_F[n]$ represents the $EMG\sigma$ flexion values at n .

There is no unique solution to solving the dynamic, linear model, so the fit coefficients must be determined using a method that minimizes the least squared error. The least square fit is modelled as:

$$\min \|Ax - b\|^2 \quad (2)$$

where,

$$A = \begin{bmatrix} \hat{s}_F[1] & \dots & -\hat{s}_E[1] \\ \vdots & \ddots & \vdots \\ \hat{s}_F[N] & \dots & -\hat{s}_E[N] \end{bmatrix} \quad x = \begin{bmatrix} f_0 \\ e_0 \end{bmatrix} \quad b = \begin{bmatrix} T_{Ext}[1] \\ T_{Ext}[2] \\ \vdots \\ T_{Ext}[N] \end{bmatrix}$$

To solve for the fit coefficients using the linear least squares fit, the Moore-Penrose pseudo-inverse approach is used. Because the matrix containing the EMG standard deviation estimates for the flexion and extension data is not a square matrix, there is no inverse matrix. The Moore-Penrose pseudo-inverse can be used to find the best fit coefficients for the least squares fit model.

The Moore Penrose pseudoinverse is a unique matrix that can be used to determine a unique solution to a matrix, A . The Moore Penrose pseudoinverse is denoted as A^\dagger . The Moore Penrose Pseudoinverse matrix must meet the following 4 conditions (Penrose):

$$AA^\dagger A = A$$

$$A^\dagger AA^\dagger = A^\dagger$$

$$(AA^\dagger)^T = AA^\dagger$$

$$(A^\dagger A)^T = A^\dagger A$$

Once the fit coefficients have been determined, torque can be estimated using the dynamic model described in equation (1) above. To accurately test the dynamic linear model a second, distinct data set needs to be used. $EMG\sigma$ must also be computed for the second, distinct data set.

Once the torques have been estimated using the linear model, the root mean square error can be calculated to judge the performance of the $EMG\sigma$ -torque model. The root mean square error is calculated by:

$$T_{RMS} = \sqrt{\frac{1}{N} \sum_{n=1}^N \{T_{Actual}[n] - T_{Estimated}[n]\}^2} \quad (3)$$

A second distinct data set needs to be used to gather data about the general performance of the model. Because the data set used in testing is unrelated to the data set used to determine the fit coefficients, the error calculated represents the error found due to the accuracy of the fit coefficients. This method leads to better conclusions about the performance of the model.

E. Experimental Data

The characteristics of the measured signal depend heavily on the type of electrode used to measure the signal. There are two types of electrodes that can be used to acquire an EMG signal, needle electrodes and surface electrodes. Needle electrodes are inserted directly into a muscle through the skin and can record the electrical activity of the adjacent motor units. Although needle electrodes provide a cleaner EMG signal, a physician is required for the insertion of the needle electrodes. Additionally, needle electrodes record a smaller muscle volume than surface electrodes. When modeling EMG to torque, larger muscle volumes lead to better performance. For this reasoning, needle electrodes were not used for collecting any of the data used in this

thesis. For data collection in an academic setting, surface electrodes are simpler to use and do not require the presence of a physician.

The data used in this analysis was collected from 64 subjects during four previous experiments (Dai, Bardizbanian and Clancy) (Liu, Lui and A.) (Clancy and Farry) (E. A. Clancy). Use of the data was approved by WPI's Institutional Review Board and each subject provided written informed consent to participate in the experiments. All subjects were seated and strapped into a custom-built straight back chair and their right shoulder was held at 90°. Once their right arm is strapped into the load cell, the upper arm forms a 90° angle with their forearm. The subject's hand is positioned in a supine orientation so that their thumb is pointed towards the ceiling and their palm is perpendicular to the floor. Figure 5 shows a subject placed in the chair with their right arm configured to collect data from their upper arm.

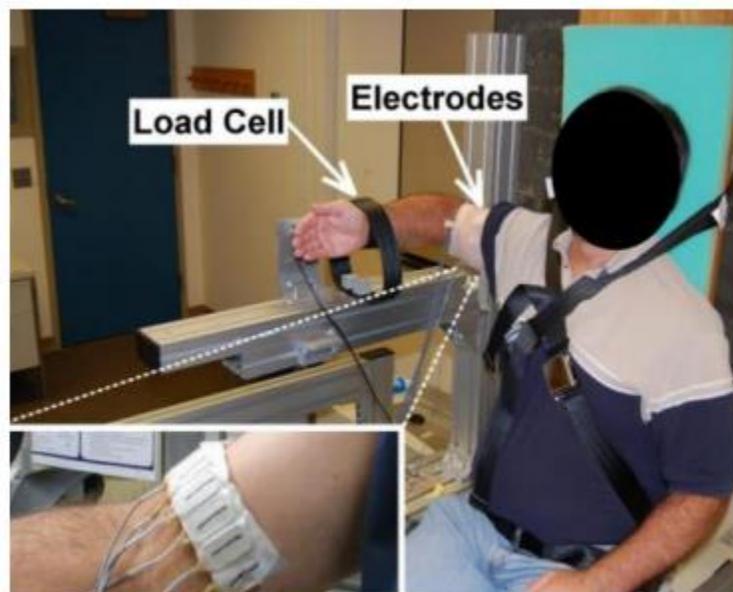


Figure 5 Set-up for Data Collection from Previous Experiments (Dai, Bardizbanian and Clancy)

The data that were collected in the four experiments were gathered from the biceps and triceps muscles using four electrodes (a total of eight electrodes) on the skin above each muscle. Before attaching the electrodes, the test area was cleaned using an alcohol wipe and then coated with a thin layer of electrode gel. The electrodes used in these experiments were custom-built, active, bipolar electrodes designed specifically to measure EMG. The electrode's contacts were 8 mm in diameter, were made from stainless steel, and had a hemispherical shape.

After the skin around the biceps and triceps was cleaned with an alcohol wipe and gelled, four electrode-amplifiers were placed transversely in a row across each muscle, with an approximate spacing of 1.75 cm. The electrodes were placed halfway between the muscle midline and the tendon insertions. An additional reference electrode was placed, after cleaning with an alcohol wipe and gel, between the electrodes on the biceps and triceps. After all nine electrodes were secured using medical tape, an elastic bandage was wrapped around the electrodes for additional support.

The signals measured by the electrodes were passed to an analog front end to remove the common mode voltages, high pass filter the data at 15 Hz with an 8th order Butterworth filter and then low pass filter the data at 1800 Hz with a 4th order Butterworth filter. After being amplified and filtered, the signals were then sampled using a 16-bit analog to digital converter sampling at a rate of 4096 Hz. From these four experiments, data are available for constant force contractions at a 50% effort level (referred to as 50% MVC data), rest signals (referred to as 0% MVC), and active trials where the effort level is varying. A three-minute rest period was included between trials to avoid muscle fatigue over the course of the experiment.

III. Methods

To decrease the variance of the EMG feature traditionally computed using the $EMG\sigma$ only procedure, additional features can be incorporated to create a new feature. The goal of the feature combination is to reduce the variance of the new information while increasing the amount of information contained in the new combined feature. Once the features have been combined, the new feature would ideally eliminate the need for a calibration and simplify some of the traditional EMG processing techniques, such as the EMG to torque modelling.

To use the EMG signal measured at the surface of the skin, the raw EMG is used to calculate the time varying EMG standard deviation, $EMG\sigma$, which is the maximum likelihood estimate of the EMG amplitude. Once the $EMG\sigma$ is computed, it can be used for modelling torque around a joint or generating a control signal for a prosthesis. In an effort to improve on the traditional $EMG\sigma$, additional features of zero crossings, slope sign change, and waveform length are incorporated with the original $EMG\sigma$ to create a new EMG feature using a few different approaches. The process of computing each individual feature and combining the four features are explained below.

A. Raw EMG Processing

Before computing any of the EMG features, the raw EMG data needs to be conditioned. The data must first be normalized so that the scaling across all the data roughly falls between -1 and 1. The normalization stage is required because the data used for this analysis were recorded using different analog to digital converters, so the scaling varies from experiment to experiment. Some of the data were recorded in volts and some were recorded in ADC counts, so normalization is required to scale the raw EMG to the same units for direct comparison across subjects from different experiments. Additionally, the raw EMG data must be conditioned because there are a few sources of interference, such as powerline and motion artifact, that need to be accounted for before computing the four EMG features.

Before any filtering, all of the data recordings used must be normalized so that the data from different subjects and experiments are of the same magnitude. The normalization factor was computed such that the root mean square of the 50% MVC recordings equaled 0.5. The gain applied to the 50% MVC data was also applied to the 0% MVC data so that it is also properly

scaled. To compute the normalization factor, the desired RMS value, 0.5, was divided by the actual RMS value of each 50% MVC channel:

$$\text{Normalization Factor} = \frac{0.5}{\sqrt{\frac{1}{N} \sum_{i=1}^N m[i]^2}},$$

where m is the raw EMG signal and N is the number of samples in the signal. Once the raw EMG has been normalized, it can be filtered.

The first filter that was applied to the normalized EMG data was a high pass filter because motion artifacts and any DC offsets needed to be removed from the data. The digital high pass filter was designed to be a 4th order Butterworth filter with a cut off frequency of 15 Hz. The magnitude response of the filter is shown below in Figure 6 with respect to frequency in Hertz.

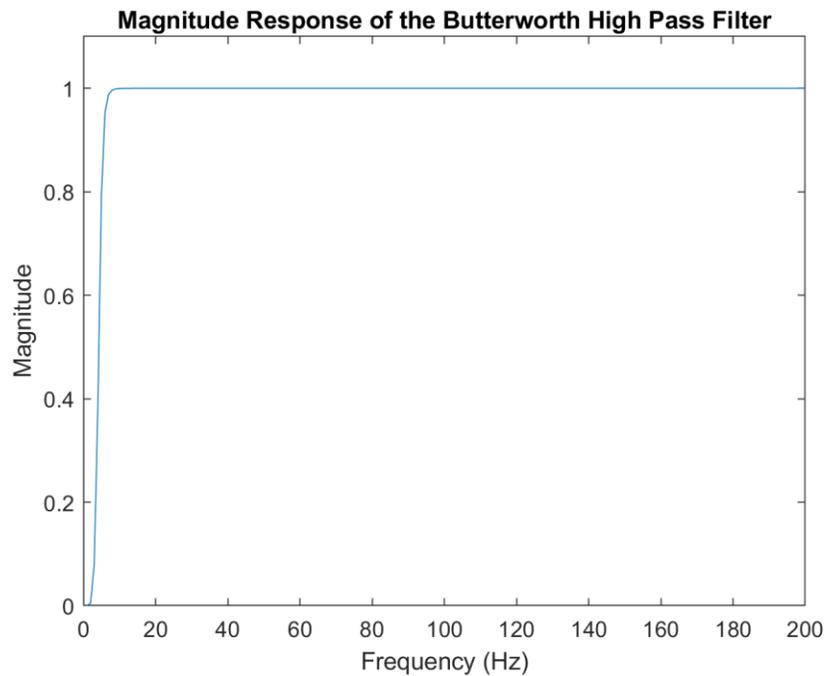


Figure 6 Magnitude Response of the 4th Order Butterworth Filter

After applying the high pass filter, a notch filter is applied to remove powerline interference. Powerline interference typically couples into a measured signal as a result of electrical equipment in the recording environment. Luckily this interference appears only at the fundamental frequency, 60 Hz in the U.S., and its harmonics at 120 Hz, 180 Hz, etc. Because the

location of this interference is known, a narrow notch filter can be placed at each known frequency location to remove the interference. The notch filter that was applied to the data is a second order digital infinite impulse response (IIR) filter. The location of the notch was set at 60 Hz and harmonics of the fundamental. The bandwidth of each notch was designed to be very narrow to avoid the attenuation of the EMG signal found at the neighboring frequencies; the bandwidth of the notch filter for each experiment is presented in Table 1.

Table 1 Notch Filter Locations and Bandwidths by Experiment

Experiment LA		Experiment LB		Experiment ww/wx	
Frequency (Hz)	Bandwidth (Hz)	Frequency (Hz)	Bandwidth (Hz)	Frequency (Hz)	Bandwidth (Hz)
60	0.25	59.97	0.25	59.99	0.25
533.7	0.8	119.94	0.25	419.93	0.25
866.6	0.8	179.91	0.25	659.89	0.8
1031.1	1.2	299.85	0.25	779.87	0.8
1031.5	1.2	359.82	0.25	899.85	0.8
1446.8	1.2	419.79	0.25	1019.83	1.2
1453.8	1.2	479.76	0.25	1139.81	1.2
1638.9	1.5	539.73	0.8	1259.79	1.2
1927.2	1.5	659.67	0.8	1379.77	1.2
		779.61	0.8	1499.75	1.2
		899.55	0.8	1619.73	1.5
		959.52	0.8	1739.71	1.5
		1019.49	1.2	1859.69	1.5
		1139.43	1.2	1979.67	1.5
		1199.4	1.2	1996.5	1.5
		1259.37	1.2		
		1379.31	1.2		
		1439.28	1.2		
		1499.25	1.2		
		1619.19	1.5		
		1739.13	1.5		
		1799.1	1.5		
		1859.07	1.5		
		1979.01	1.5		

Figure 7 shows the plotted magnitude response of the notch filter to remove 60 Hz interference created for subjects from Experiment LA. The other experiments will show a similar shape, but the notch will appear at a different location in frequency and the bandwidth will vary.

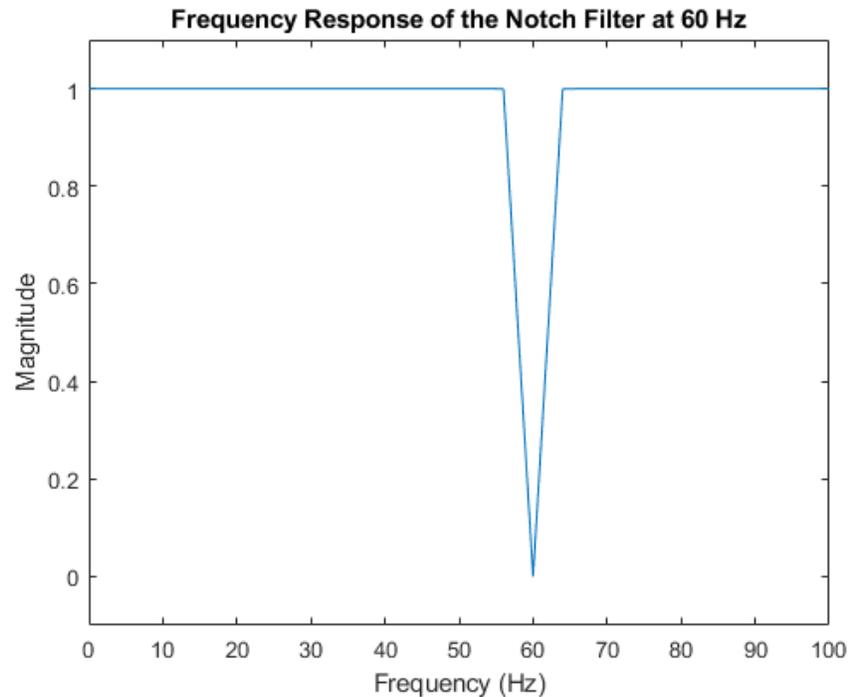


Figure 7 Frequency Response of the IIR Notch Filter at 60 Hz

Once the powerline interference has been removed, an optional whitening filter can be applied to the data. A whitening filter is used to increase the signal's statistical bandwidth and to reduce that EMG's variance. A whitening filter is useful in amplifying in the frequency ranges where the signal has its most power and does not amplify where the signal power is lowest. Application of a whitening filter helps to avoid the amplification of the noise in the higher frequencies, above 600 Hz. The whitening filter used in this analysis is a universal whitening filter that was developed from the ensemble average of the adaptive whitening filter of the individual 50% MVC trials of the subjects. The whitening filter that was applied to the EMG data is a 60th order FIR filter with a band limit of 600 Hz. The magnitude response of this filter can be seen in Figure 8. (Clancy and Farry)

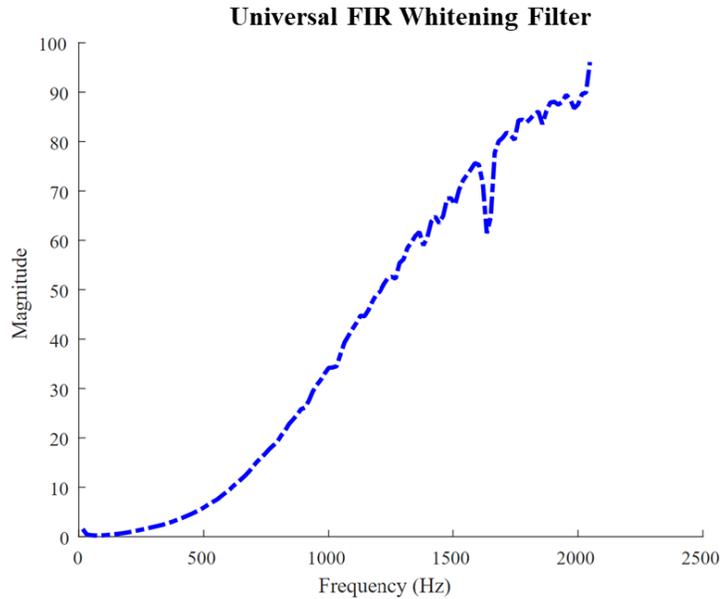


Figure 8 Magnitude Response of the Fixed Whitening Filter

Once the signal has been conditioned to remove powerline interference, DC offsets and motion artifacts, the four features can be computed, and different methods of combination can be tested.

Computing the Features

Once the data have been conditioned, the features can be extracted from the signal. Four features were considered in this analysis: $EMG\sigma$, zero crossings, slope sign changes, and waveform length. These four features are time domain features that can be extracted from the measured EMG and are simple to implement in a system (Kamavuako, Scheme and Englehart) (Hudgins and Parker).

To compute $EMG\sigma$, the absolute value of the normalized, filtered data is computed. To show the process of computing the EMG standard deviation, a sine wave was used to illustrate the computation of the EMG standard deviation features, as well as the additional three features. A sine wave was selected because it is a periodic signal whose parameters (fundamental frequency and amplitude) are simple to control and the locations of zero crossings and slope sign changes are easy to identify. The example sine wave was generated with an amplitude of one and a fundamental frequency of 2 Hz.

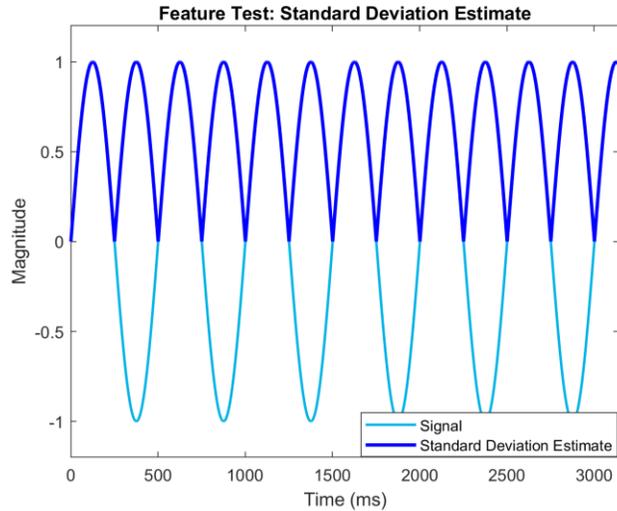


Figure 9 Example EMG Standard Deviation Estimate without Smoothing Filter

The zero crossing feature is a count of every time the signal crosses zero. Zero crossings are a rough estimate of a signal’s frequency and can also be used to calculate the zero crossing rate. To calculate the zero crossing feature, a change in sign, either positive or negative, must be detected between two consecutive samples. The output of the zero crossing computation is a vector of zeros and ones with a length equal to the number of samples in the signal. A zero in the output vector corresponds to no detected zero crossing and a one corresponds to a detected zero crossing.

When computing zero crossings, the sign of the first sample is compared to the second sample. If they have different signs, then a zero crossing may have occurred. To confirm that a zero crossing has been detected, a difference magnitude threshold is included in the calculation of zero crossings. The threshold is set to avoid any mistakenly identified zero crossings that are due to noise in the measured signal. The threshold is set to equal two to three times the root-mean-square (RMS) of the noise recording (Kamavuako, Scheme and Englehart). To determine if a zero crossing has occurred, the difference between the first and second sample is computed and if this difference exceeds the threshold, then a zero crossing has indeed occurred. When a zero crossing has occurred, a count is kept at the corresponding time index. (Hudgins and Parker)

To illustrate the process of computing a zero crossing, a noise-free sine wave was generated and used as the example signal. The example sine wave was generated with an amplitude of one and a fundamental frequency of 2 Hz. To compute the zero crossings in a

noise-free signal, the threshold is set to equal zero. The results of this simulation are shown in Figure 10. Each time that the sine wave crosses through zero, a zero crossing is found and marked by the dark blue stem plot. The dark blue line that is found at y equal to zero shows that no zero crossings have been detected at those locations.

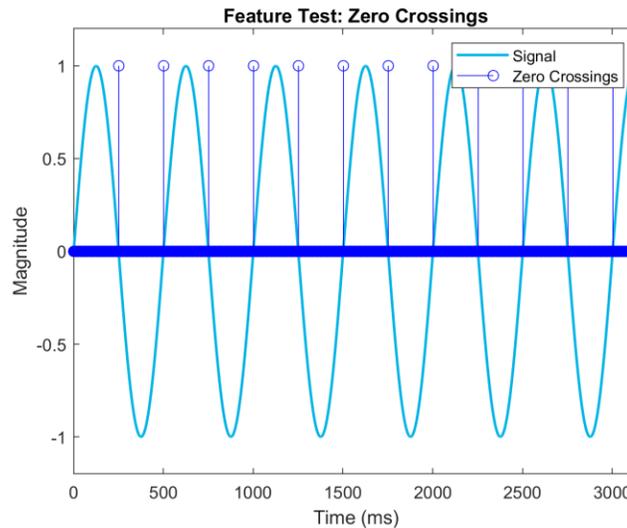


Figure 10 Example Zero Crossing Computation

The third feature that was computed was slope sign changes. Slope sign changes are similar to zero crossings because the computed feature is a vector of zeros and ones. Slope sign changes requires three consecutive samples to determine if the signal's slope has changed sign. To determine if a slope sign change has occurred, the difference is computed between the first and second sample and the second and third sample. If the polarities of the two differences are not equal, then a slope sign change may have occurred. To confirm that a slope sign change has occurred, the absolute value of at least one of the differences must exceed the threshold. The threshold set for the slope sign changes is the same as the threshold set for zero crossings; the threshold is set to be two to three times the root-mean-square of the noise recording. (Hudgins and Parker)

To illustrate the computation of slope sign changes, a noise-free sine wave was generated as a test signal and the slope sign changes were computed. The example sine was generated with an amplitude of one and a fundamental frequency of 2 Hz. Because there is no noise in the generated signal, the threshold is set to equal zero. Figure 11 shows the computed slope sign changes (dark blue stem) as well as the test signal (light blue). For a sine wave, a slope sign

change occurs every time that the sine wave reaches its minimum or maximum amplitude; in this case, a slope sign change occurs each time that the signal is equal to -1 or 1.

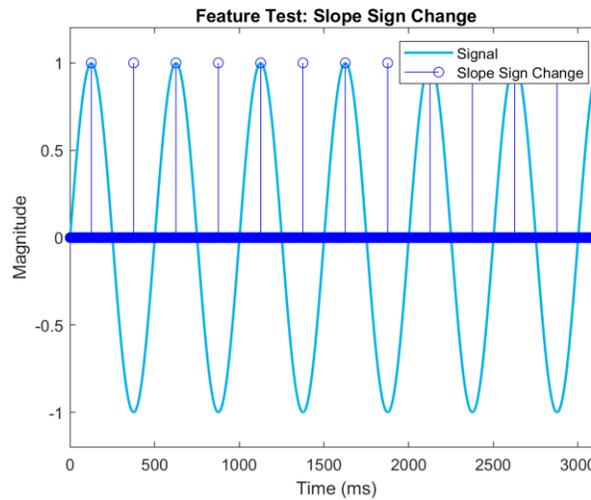


Figure 11 Example Slope Sign Change Computation

The final feature that was considered in this analysis was waveform length. Waveform length is length of the given signal across all samples. To compute the waveform length, the absolute value of the first backward difference is computed over two consecutive samples across the entirety of the measured signal. (Hudgins and Parker)

B. Combination of the Features

After the four features have been computed, they must be combined to form a single feature which will replace the $EMG\sigma$ only feature. Multiple approaches were tested as a procedure to combine the features. These approaches evolved as the behavior of each individual feature was better understood relative to the desired output. The performance of the new EMG feature was compared to the previous $EMG\sigma$ only feature using a linear least squares regression model of EMG to torque. Additional testing included the comparison of computing the features from whitened or unwhitened data. From previous work, the application of a whitening filter on the elbow data shows improvement in the EMG to torque modelling (Clancy and Farry).

1) Combination of Uncorrelated Features using Averaging

The four features are computed from the same data set and some of the information contained within each featured may be redundant. When computing the optimal EMG estimate

of Gaussian distributed samples, the samples must be uncorrelated and statistically independent. If the features are uncorrelated and independent, then a simple combination method is the average of the four features. When averaged together, each feature has an equal contribution to the combined feature.

The initial approach to combining the features included uncorrelating the features using an eigenvalue decomposition. Once the features are computed, their correlations must be removed. Once the correlations are removed, the features can be treated as uncorrelated and independent samples and averaged together to form a new EMG feature. The single feature will replace the EMG σ only feature in the EMG processor. A block diagram of the procedure is shown below in Figure 12.

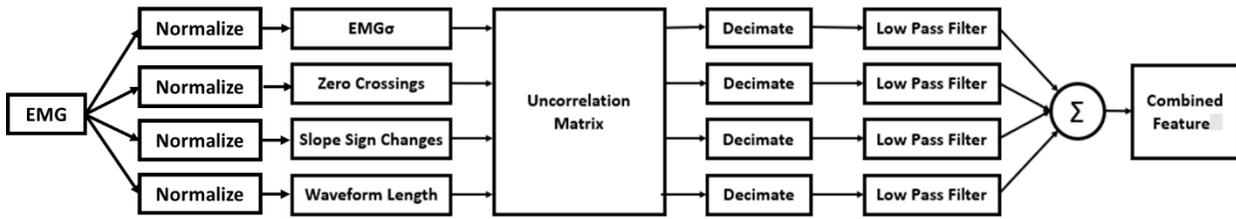


Figure 12 Initially Proposed Process of Feature Combination

Each feature, $f_{feature}$, is treated as a random variable and the four features are arranged into a random vector, \underline{f} .

$$\underline{f} = \begin{bmatrix} f_{EMG\sigma} \\ f_{ZC} \\ f_{SSC} \\ f_{WL} \end{bmatrix}$$

The cross-correlation matrix of the random vector \underline{f} is given as:

$$C_{\underline{f}\underline{f}} = \varepsilon \left[\underline{f}\underline{f}^T \right] = \begin{bmatrix} C_{f_{i1}f_{j1}} & C_{f_{i1}f_{j2}} & C_{f_{i1}f_{j3}} & C_{f_{i1}f_{j4}} \\ C_{f_{i2}f_{j1}} & C_{f_{i2}f_{j2}} & C_{f_{i2}f_{j3}} & C_{f_{i2}f_{j4}} \\ C_{f_{i3}f_{j1}} & C_{f_{i3}f_{j2}} & C_{f_{i3}f_{j3}} & C_{f_{i3}f_{j4}} \\ C_{f_{i4}f_{j1}} & C_{f_{i4}f_{j2}} & C_{f_{i4}f_{j3}} & C_{f_{i4}f_{j4}} \end{bmatrix}$$

The cross-covariance matrix of the random vector \underline{f} is given as:

$$K_{\underline{f}\underline{f}} = \varepsilon \left[(\underline{f} - \underline{\mu}_f)(\underline{f} - \underline{\mu}_f)^T \right] = C_{\underline{f}\underline{f}} - \underline{\mu}_f \underline{\mu}_f^T = \begin{bmatrix} k_{f_{i1}f_{j1}} & k_{f_{i1}f_{j2}} & k_{f_{i1}f_{j3}} & k_{f_{i1}f_{j4}} \\ k_{f_{i2}f_{j1}} & k_{f_{i2}f_{j2}} & k_{f_{i2}f_{j3}} & k_{f_{i2}f_{j4}} \\ k_{f_{i3}f_{j1}} & k_{f_{i3}f_{j2}} & k_{f_{i3}f_{j3}} & k_{f_{i3}f_{j4}} \\ k_{f_{i4}f_{j1}} & k_{f_{i4}f_{j2}} & k_{f_{i4}f_{j3}} & k_{f_{i4}f_{j4}} \end{bmatrix}$$

The cross-correlation coefficients are also useful to compare the relative correlations between the features. The cross-correlation coefficients normalize the correlations so that their value fall between -1 and 1. A cross-correlation coefficient of -1 shows complete negative correlation, a cross-correlation coefficient of 0 shows no correlation and a cross-correlation coefficient of 1 shows complete positive correlation. The cross-correlation coefficient of the random vector \underline{f} is given as:

$$R_{\underline{f}\underline{f}} = \begin{bmatrix} r_{f_{i1}f_{j1}} & r_{f_{i1}f_{j2}} & r_{f_{i1}f_{j3}} & r_{f_{i1}f_{j4}} \\ r_{f_{i2}f_{j1}} & r_{f_{i2}f_{j2}} & r_{f_{i2}f_{j3}} & r_{f_{i2}f_{j4}} \\ r_{f_{i3}f_{j1}} & r_{f_{i3}f_{j2}} & r_{f_{i3}f_{j3}} & r_{f_{i3}f_{j4}} \\ r_{f_{i4}f_{j1}} & r_{f_{i4}f_{j2}} & r_{f_{i4}f_{j3}} & r_{f_{i4}f_{j4}} \end{bmatrix}$$

Although the cross-correlation coefficients are not used in the process of removing the correlations, the cross-correlation coefficients report the relative correlations that exist between the features. Additionally, if the features are successfully uncorrelated, the cross-correlation coefficient matrix should be an identity matrix. The computation of the correlation coefficient matrix is a convenient method to check that the features have been uncorrelated.

To uncorrelate the four features, a linear transformation must be applied to create a new random vector, \underline{y} from the original random vector, \underline{f} .

$$\underline{y} = D\underline{f}$$

The matrix D is created so that the new random vector is composed of uncorrelated random variables. The steps to determine the matrix D that will lead to a set of uncorrelated random variables are shown below.

The cross-covariance matrix of the new random vector, \underline{y} , is equal to:

$$K_{\underline{y}\underline{y}} = DK_{\underline{f}\underline{f}}D^T$$

The cross-covariance matrix is real and symmetric so it has real, linearly independent eigenvectors which can be orthonormal. The eigenvectors, $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_N$, and the corresponding eigenvalues, $\lambda_1, \lambda_2, \dots, \lambda_N$, must satisfy the following:

$$\underline{K}_{ff}\underline{v}_i = \lambda_i\underline{v}_i \quad \text{for } i = 1:N$$

The matrix D^T is a partitioned matrix created from the eigenvectors of the cross-covariance matrix as such:

$$D^T = [\underline{v}_1 | \underline{v}_2 | \dots | \underline{v}_N]$$

If the eigenvectors are contained within the matrix D^T , then the eigenvalue equations for all of the random variables can be written into a single equation using D^T .

$$\underline{K}_{ff}D^T = D^T\underline{\Lambda},$$

where $\underline{\Lambda}$ is equal to

$$\underline{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{bmatrix}$$

Because D^T was created from orthonormal columns, $D^{-1} = D^T$ and the eigenvector equation can be rewritten as:

$$D\underline{K}_{ff}D^T = \underline{\Lambda}$$

The matrix D can then be used to compute a new set of uncorrelated random variables.

$$\underline{y} = D\underline{f}$$

And the covariance of the new, uncorrelated random variables is equal to:

$$\underline{K}_{yy} = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_4 \end{bmatrix}$$

Once the features have been uncorrelated, they must be decimated, and low pass filtered prior to combining. The new features are first decimated individually by an overall factor of 100 which decreases the sampling frequency from 4096 Hz to 40.96 Hz. To achieve a decimation of 100, the signal is decimated in two passes, each with a decimation factor of 10 preceded by a 7th

order infinite impulse response Chebyshev lowpass filter. The first Chebyshev filter has a cut-off frequency of 163.84 Hz and the second Chebyshev filter has a cut-off frequency of 16.384 Hz. The second filter's cut-off frequency is the overall cut-off frequency of the decimation stage.

After decimating each feature, a 2nd order Butterworth lowpass filter with a cut-off frequency of 1 Hz was applied. Once each feature has been decimated and low pass filtered, the features need to be normalized to have a mean equal to the mean of the EMG σ feature. Before combining the features, they must be normalized again to force their means to be on a similar scale.

After uncorrelating, decimating, and low pass filtering, the four features are averaged together to form the new EMG feature. Five different combinations were tested to determine which estimate leads to the best performance in the torque model. The different estimates are computed by averaging all four features, averaging three features (EMG σ , zero crossings, and slope sign changes), and by averaging three two-feature pairs: EMG σ and zero crossings, EMG σ and slope sign changes, and EMG σ and waveform length. The new EMG features are used to train the linear least squares torque model and then the model is applied to estimate the torque of a separate testing data set. The error between the estimated and actual torque is reported as %MVC. If the error calculated using the new EMG feature is lower than the %MVC error computed with the traditional EMG σ only feature, then improvement has been made in the estimation of the EMG amplitude.

Before beginning the process of combining the features using the uncorrelation matrix, the correlations that exist between the features were studied. After computing the features individually, the cross-correlation coefficients of the four features were computed. The cross-correlation coefficients represent the normalized correlations that exist between any two features. Cross-correlation coefficients are normalized to fall within the range of -1 to 1. If the cross correlation is computed to be -1, the features show a highly negative correlation, and if the cross correlation is computed to be +1, the two features are highly correlated in the positive direction. If the cross-correlation coefficient is 0, then the two features are completely uncorrelated. Due to the nature of the features, no negative cross correlation coefficients are expected because all features are expected to be monotonically increasing with respect to the level of force.

For the four features, there are 16 total cross-correlations computed and stored in a four-by-four matrix. The diagonal elements of the cross-correlation coefficient matrix will equal one because a feature is completely correlated with itself. The off-diagonal terms represent the cross-correlation coefficients for each two-feature pair and the matrix is symmetric about the diagonal. The ensemble average of the cross-correlation coefficients, and the standard deviations of the cross-correlation coefficients for the 64 subjects is shown below in Table 2.

Table 2 Ensemble Mean of the Cross-Correlation Coefficients for the Four Features, Across 64 Subjects and 8 Channels

	EMGσ	Zero Crossings	Slope Sign Changes	Waveform Length
EMGσ	1	0.0256±0.0141	0.2498±0.0562	0.6317±0.0177
Zero Crossings	0.0256±0.0141	1	0.0818±0.0547	0.2675±0.0458
Slope Sign Changes	0.2498±0.0562	0.0818±0.0547	1	0.1777±0.0910
Waveform Length	0.6317±0.0177	0.2675±0.0458	0.1777±0.0910	1

Only two features, EMG σ and waveform length are highly correlated, with an ensemble averaged cross-correlation coefficient of 0.63. The remaining cross-correlation coefficients all fall below 0.3, which means that the features are weakly correlated.

After observing the correlations between the four features, the uncorrelation matrix was created for each individual subject using the steps outlined above. Once this matrix was computed, it was applied to the original set of features to create a new set of features. The cross-correlation coefficients of the uncorrelated features are given below in Table 2.

Table 3 Cross Correlation Coefficients computed after the Linear Transform $\underline{y} = D\underline{f}$

	EMGσ	Zero Crossings	Slope Sign Changes	Waveform Length
EMGσ	1	0	0	0
Zero Crossings	0	1	0	0
Slope Sign Changes	0	0	1	0
Waveform Length	0	0	0	1

Multiplying the features by their uncorrelation matrix is a linear transformation of the original four features to a new set of features. The new set of features are derived using information from the original four features, but no longer represent the original four features. The new features were also found to contain negative components. Figure 13 below shows the uncorrelated features created from the 50% MVC data.

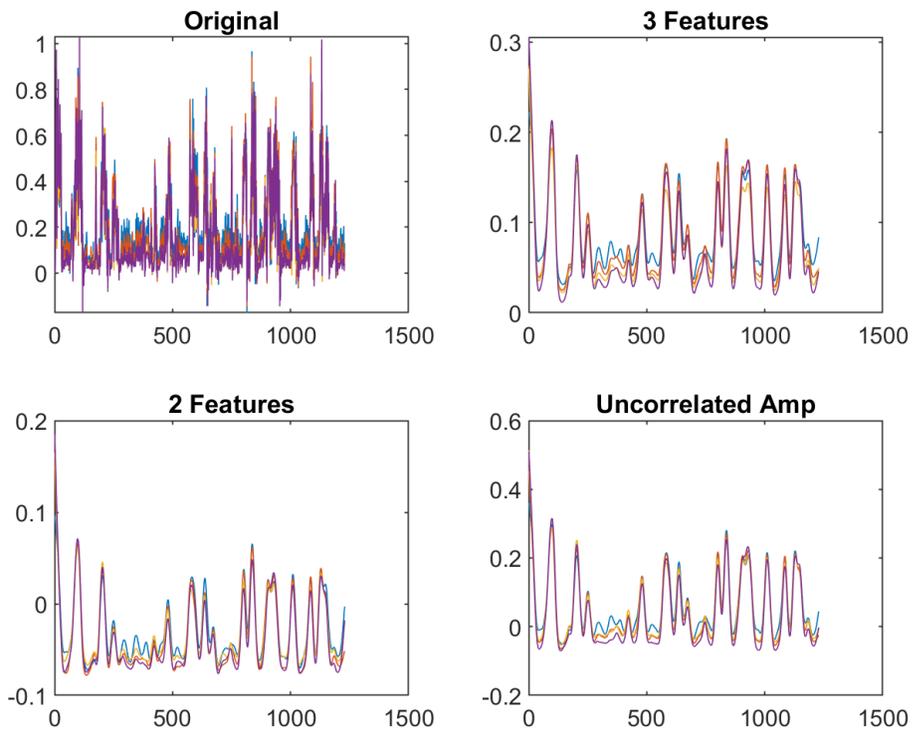


Figure 13 Uncorrelated Features Combined by Averaging

When applying the linear least squares torque model on the dynamic data, the new EMG feature that was derived from the uncorrelated features showed error significantly higher than the baseline error of 4.8% for the whitened data and 5.5% for the unwhitened data. The average %MVC error computed for a model order of 15 for the different uncorrelated feature combinations formed by averaging the uncorrelated features are summarized in Table 4 below.

Table 4 Average %MVC Error of 15th Order EMG to Torque Models

	Whitened	Unwhitened
No Features (Baseline)	4.84%	5.5%
4 Features (Avg.)	17.7%	-
3 Features (Avg.)	18.61%	21.54%
2 Features (Avg.)	19.07%	22.02%
1 Feature	17.35%	21.71%

The uncorrelated EMG feature showed errors higher than the baseline error. Because no improvement was made when uncorrelating the features and the relative correlations between the features are relatively low with the exception of $EMG\sigma$ and waveform length, the next approach to combining the features was to combine them without removing the correlations between the features. The combination of features using different methods of combination showed better performance than the uncorrelated features but did not make any improvements on the baseline error.

2) Combination of the Features through Averaging

If removing the correlations between the features was successful, the optimal estimator for the combined feature would be the average of the four features. If the features are uncorrelated, then they would contribute an equal amount of information to the combined feature. As seen in Table 2, the features are not highly correlated with the exception of $EMG\sigma$ and waveform length, so averaging the features together is still a practical weighing scheme; it is just not the optimal scheme.

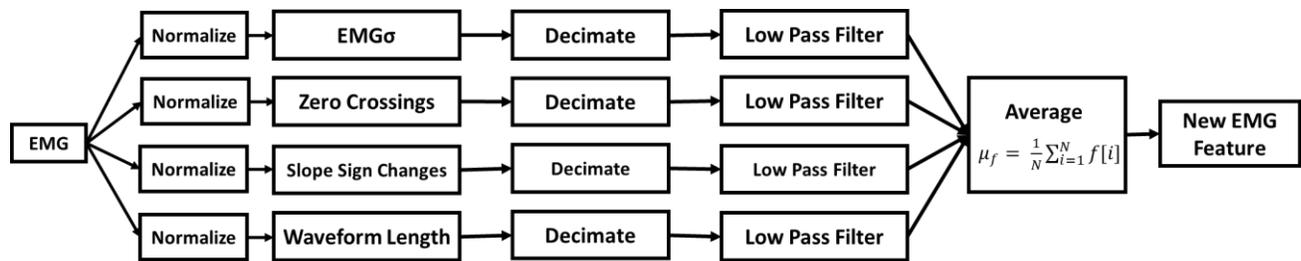


Figure 14 Procedure used to Average the Features Together

When averaged together each individual feature has the same weight in the combined feature, regardless of the feature's actual information content. Although averaging the features together is known to not be the optimal weighing scheme for correlated features, it was implemented because it is a simple method to combine the four features. Figure 15 and Figure 16 show the combined feature created from averaging different feature pair combinations using the 50% MVC data.

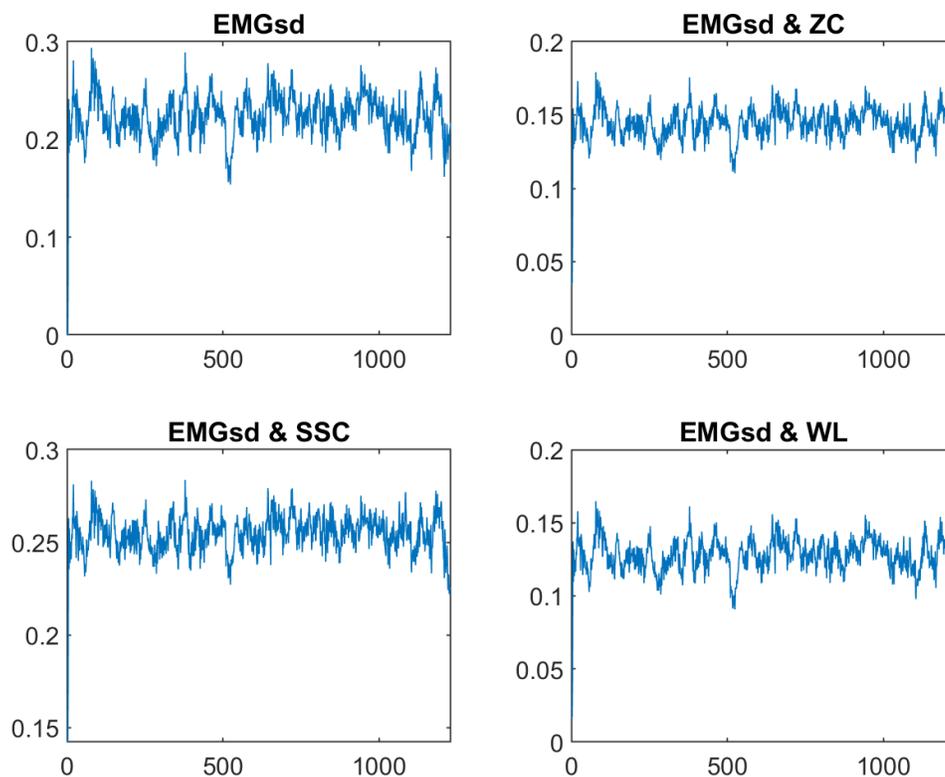


Figure 15 Combined Features from Averaging, Unwhitened

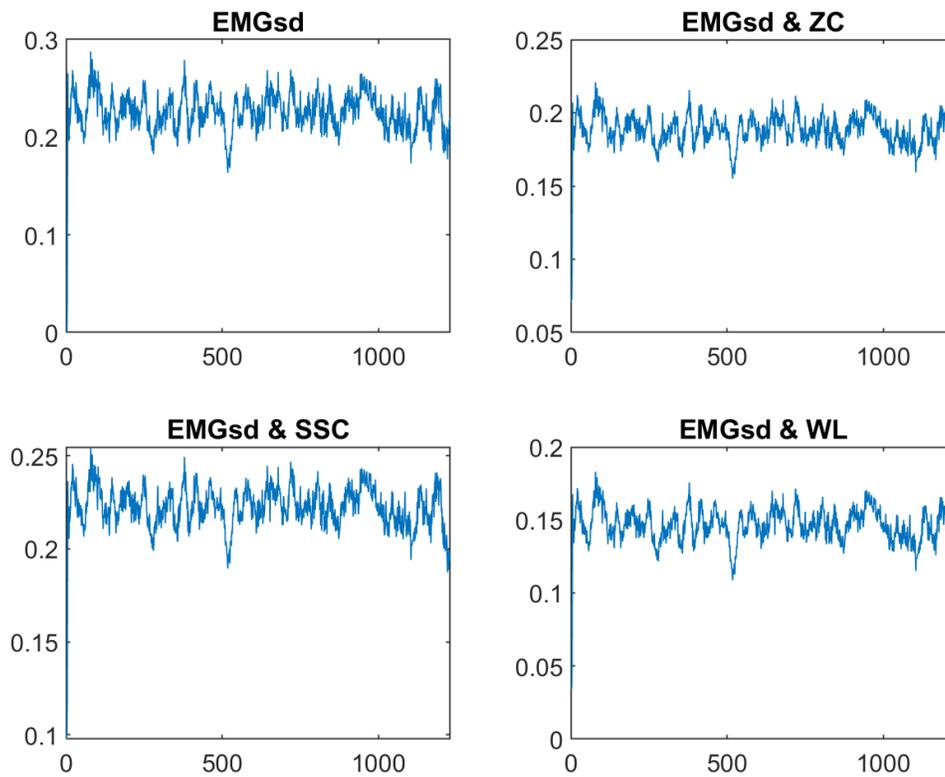


Figure 16 Combined Features from Averaging, Whitened

Once the features have been averaged together to form a new feature, the new feature is then used in the linear least squares EMG to torque model. The performance of EMG-torque model of the new EMG feature created using the dynamic contraction data is shown below in Figure 17 and Figure 18. The mean squared error is expressed in terms of %MVC.

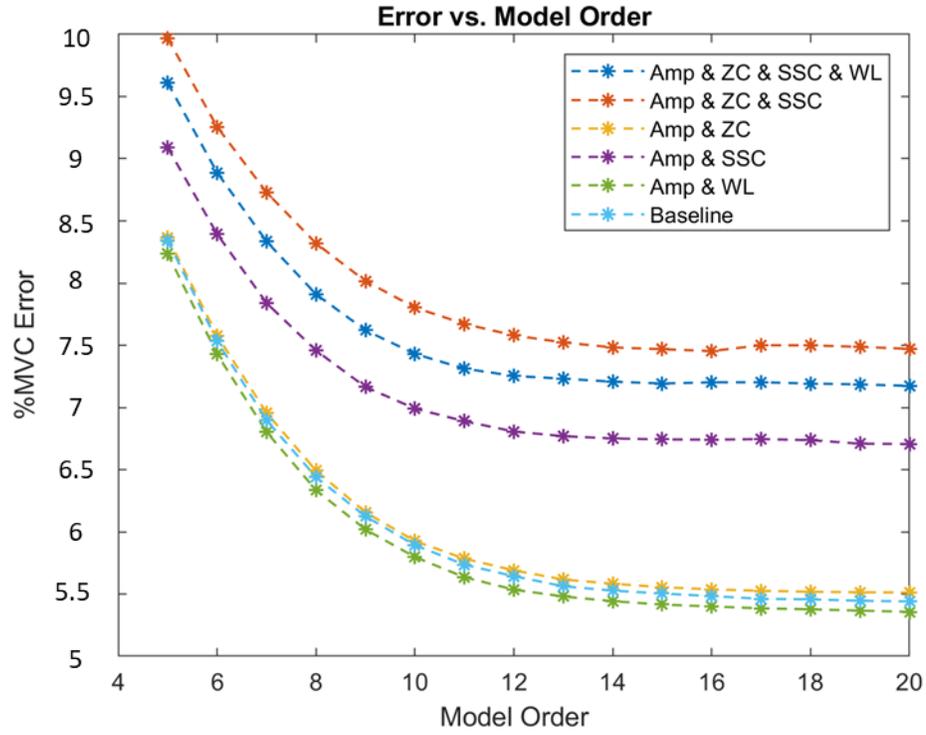


Figure 17 Torque Model Performance Expressed in %MVC, Unwhitened

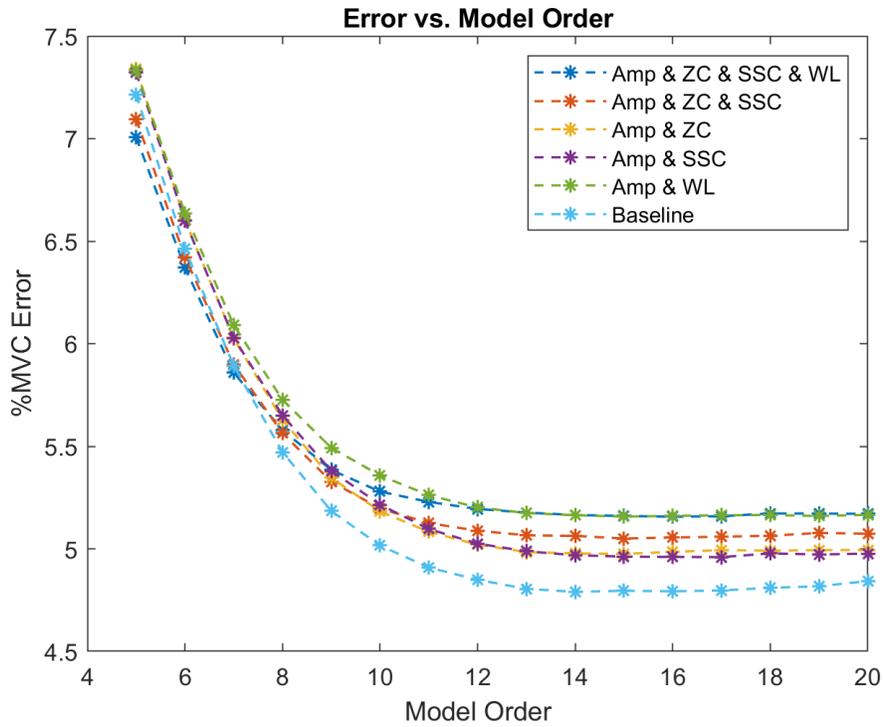


Figure 18 Torque Model Performance Expressed in %MVC, Whitened

For the unwhitened data, the best performance was seen in the EMG σ and waveform length combination. The performance of the EMG σ and waveform length pair tracks very closely with the baseline performance and the EMG σ and zero crossing combination. The improvement in %MVC between the EMG σ and waveform length is approximately 0.1 %MVC.

Whitening the EMG reduces the %MVC error for all of the feature combinations that were considered. For the whitened data, the best performance was seen in the baseline, EMG σ only, feature. The baseline error for the whitened data was computed to be 4.8%. The performance of the new feature combinations that were tested fall at 5 %MVC or greater. From Figure 17 and Figure 18, whitening is shown to decrease the magnitude of the error for all of the features, but does not improve the performance of the combined features relative to the baseline. No combined feature improved EMG to torque performance.

3) Combination of Features according to their Statistics

Another approach that was used to combine the features prior to the EMG to torque model involved deriving weights for each feature based on the features' signal to noise ratio. The means of each feature are normalized so that they are equal, approximately 0.5. As a result, comparing the signal to noise ratio of each feature is a comparison of the variance of the individual feature. The feature with the greatest signal to noise ratio will have the most weight in the combination because it has the least variance. To ensure that the weights sum to one, each individual weight is divided by the sum of the four weights. Once the weights are computed for all electrodes for each subject, a universal weight is derived by averaging the weights across all electrodes and across all subjects for each feature. If computation time is not a concern, the weights could be calculated on a subject to subject basis.

The computation of the features, the decimation stage and the low pass filtering stages are the same as above, but the uncorrelation matrix stage has been removed. Instead of combining four uncorrelated features, the correlated features are combined using weights computed from the signal to noise ratio of each feature. The overview of this combination process is outlined in Figure 19.

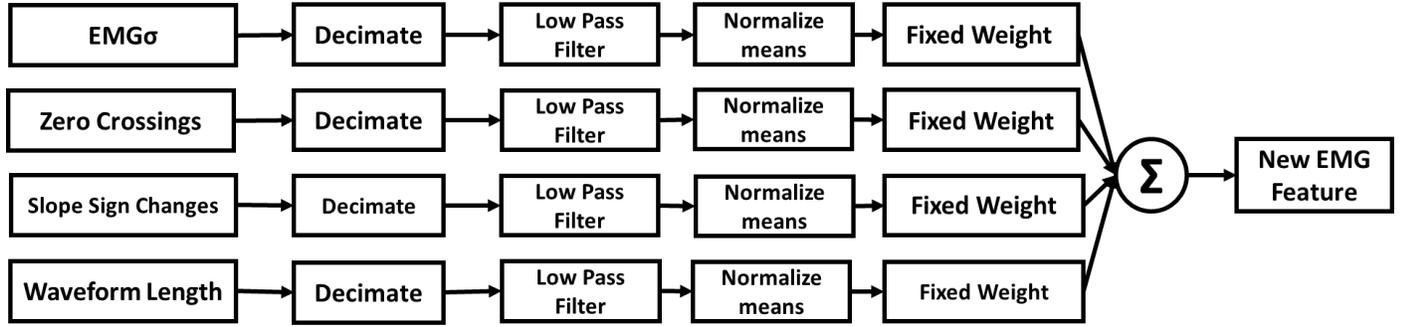


Figure 19 Second Feature Combination Process

The first three stages of Figure 19 are previously described in the Combination of Uncorrelated Features using Averaging Section above. For the combination of four features, the following equations are used to compute the weights for each feature:

$$w_{ZC} = \frac{\frac{4\mu_{ZC}^2}{\sigma_{ZC}^2}}{\frac{4\mu_{ZC}^2}{\sigma_{ZC}^2} + \frac{4\mu_{SSC}^2}{\sigma_{SSC}^2} + \frac{4\mu_{WL}^2}{\sigma_{WL}^2} + \frac{1}{\sigma_{EMG\sigma}^2}}$$

$$w_{SSC} = \frac{\frac{4\mu_{SSC}^2}{\sigma_{SSC}^2}}{\frac{4\mu_{ZC}^2}{\sigma_{ZC}^2} + \frac{4\mu_{SSC}^2}{\sigma_{SSC}^2} + \frac{4\mu_{WL}^2}{\sigma_{WL}^2} + \frac{1}{\sigma_{EMG\sigma}^2}}$$

$$w_{WL} = \frac{\frac{4\mu_{WL}^2}{\sigma_{WL}^2}}{\frac{4\mu_{ZC}^2}{\sigma_{ZC}^2} + \frac{4\mu_{SSC}^2}{\sigma_{SSC}^2} + \frac{4\mu_{WL}^2}{\sigma_{WL}^2} + \frac{1}{\sigma_{EMG\sigma}^2}}$$

$$w_{EMG\sigma} = \frac{\frac{1}{\sigma_{EMG\sigma}^2}}{\frac{4\mu_{ZC}^2}{\sigma_{ZC}^2} + \frac{4\mu_{SSC}^2}{\sigma_{SSC}^2} + \frac{4\mu_{WL}^2}{\sigma_{WL}^2} + \frac{1}{\sigma_{EMG\sigma}^2}}$$

For the three-feature combination of EMGσ, zero crossings, and slope sign changes:

$$w_{ZC} = \frac{\frac{4\mu_{ZC}^2}{\sigma_{ZC}^2}}{\frac{4\mu_{ZC}^2}{\sigma_{ZC}^2} + \frac{4\mu_{SSC}^2}{\sigma_{SSC}^2} + \frac{1}{\sigma_{EMG\sigma}^2}}$$

$$w_{SSC} = \frac{\frac{4\mu_{SSC}^2}{\sigma_{SSC}^2}}{\frac{4\mu_{ZC}^2}{\sigma_{ZC}^2} + \frac{4\mu_{SSC}^2}{\sigma_{SSC}^2} + \frac{1}{\sigma_{EMG\sigma}^2}}$$

$$w_{EMG\sigma} = \frac{\frac{1}{\sigma_{EMG\sigma}^2}}{\frac{4\mu_{ZC}^2}{\sigma_{ZC}^2} + \frac{4\mu_{SSC}^2}{\sigma_{SSC}^2} + \frac{1}{\sigma_{EMG\sigma}^2}}$$

For the two-feature combination of EMG σ and zero crossings:

$$w_{ZC} = \frac{\frac{4\mu_{ZC}^2}{\sigma_{ZC}^2}}{\frac{4\mu_{ZC}^2}{\sigma_{ZC}^2} + \frac{1}{\sigma_{EMG\sigma}^2}}$$

$$w_{EMG\sigma-ZC} = \frac{\frac{1}{\sigma_{EMG\sigma}^2}}{\frac{4\mu_{ZC}^2}{\sigma_{ZC}^2} + \frac{1}{\sigma_{EMG\sigma}^2}}$$

For the two-feature combination of EMG σ and slope sign changes:

$$w_{SSC} = \frac{\frac{4\mu_{SSC}^2}{\sigma_{SSC}^2}}{\frac{4\mu_{SSC}^2}{\sigma_{SSC}^2} + \frac{1}{\sigma_{EMG\sigma}^2}}$$

$$w_{EMG\sigma-SSC} = \frac{\frac{1}{\sigma_{EMG\sigma}^2}}{\frac{4\mu_{SSC}^2}{\sigma_{SSC}^2} + \frac{1}{\sigma_{EMG\sigma}^2}}$$

For the two-feature combination of EMG σ and waveform length:

$$w_{WL} = \frac{\frac{4\mu_{WL}^2}{\sigma_{WL}^2}}{\frac{4\mu_{WL}^2}{\sigma_{WL}^2} + \frac{1}{\sigma_A^2}}$$

$$w_{EMG\sigma-SSC} = \frac{\frac{1}{\sigma_{EMG\sigma}^2}}{\frac{4\mu_{WL}^2}{\sigma_{WL}^2} + \frac{1}{\sigma_{EMG\sigma}^2}}$$

To derive the universal weights that will be applied before combining the features, the individual weights for each feature combination are computed using 50% MVC flexion and extension data for all 64 subjects. Then the 512 weights (8 channels by 64 subjects) for each feature are averaged together to determine the universal weight for each feature. The standard deviations of the weights were computed to confirm that there were no significant differences seen across the weights derived from each subject. The universal weights that were set for each combination are presented in Table 5.

Table 5 Universal Fixed Weights

Combination	EMGσ Weight	ZC Weight	SSC Weight	WL Weight
EMGσ + ZC + SSC + WL	0.2287	0.1201	0.2145	0.4374
EMGσ + ZC + SSC	0.4059	0.2133	0.3808	--
EMGσ + ZC	0.6556	0.3444	--	--
EMGσ + SSC	0.3595	--	0.6405	--
EMGσ + WL	0.3529	--	--	0.6471

After computing the universal weights, they can be applied to their corresponding feature. Then the features can be combined as such:

For the four-feature combination, the new EMG feature is given as:

$$new\ EMG\ Feature = w_{ZC} \cdot ZC + w_{SSC} \cdot SSC + w_{WL} \cdot WL + w_{EMG\sigma} \cdot EMG\sigma$$

For the three-feature combination, the new EMG feature is given as:

$$new\ EMG\ Feature = w_{ZC} \cdot ZC + w_{SSC} \cdot SSC + w_{EMG\sigma} \cdot EMG\sigma$$

For the two-feature combinations, the new EMG feature is given as:

$$\text{new EMG Feature} = w_{ZC} \cdot ZC + w_{EMG\sigma-ZC} \cdot EMG\sigma$$

$$\text{new EMG Feature} = w_{SSC} \cdot SSC + w_{EMG\sigma-SSC} \cdot EMG\sigma$$

$$\text{new EMG Feature} = w_{WL} \cdot WL + w_{EMG\sigma-WL} \cdot EMG\sigma$$

The equations and weights listed above are used to develop the new EMG feature using the dynamic calibration data prior to the linear least squares torque model. The model is trained using the new EMG feature created from one set of dynamic calibration data and then applied to a separate set of dynamic calibration data for testing the model's performance. The following figure shows the average of the traditional time-varying standard deviation, EMG σ feature, and the average of the three 2-feature combinations with and without whitening developed using 50% MVC data.

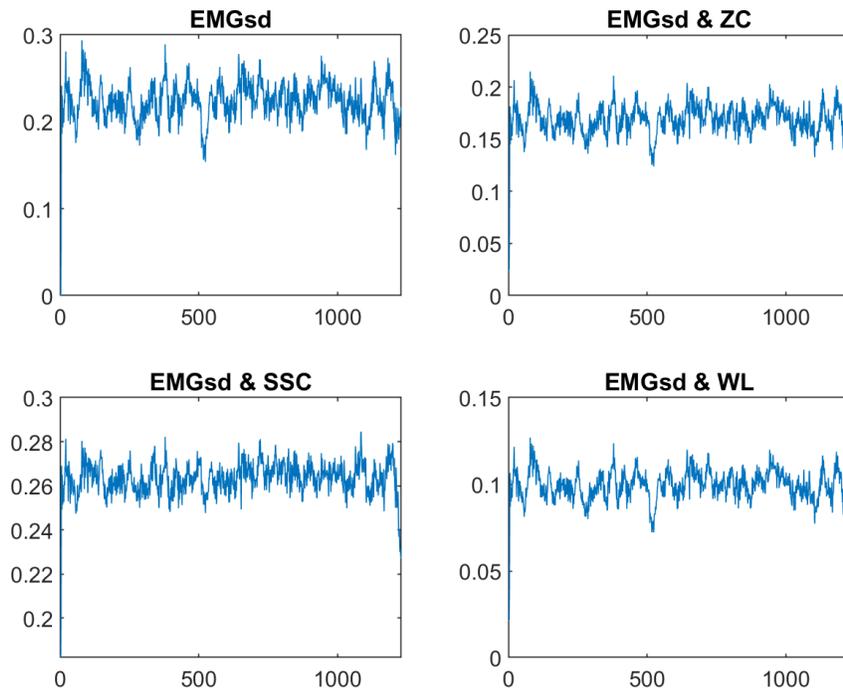


Figure 20 Combined Features from Fixed Weights, Unwhitened

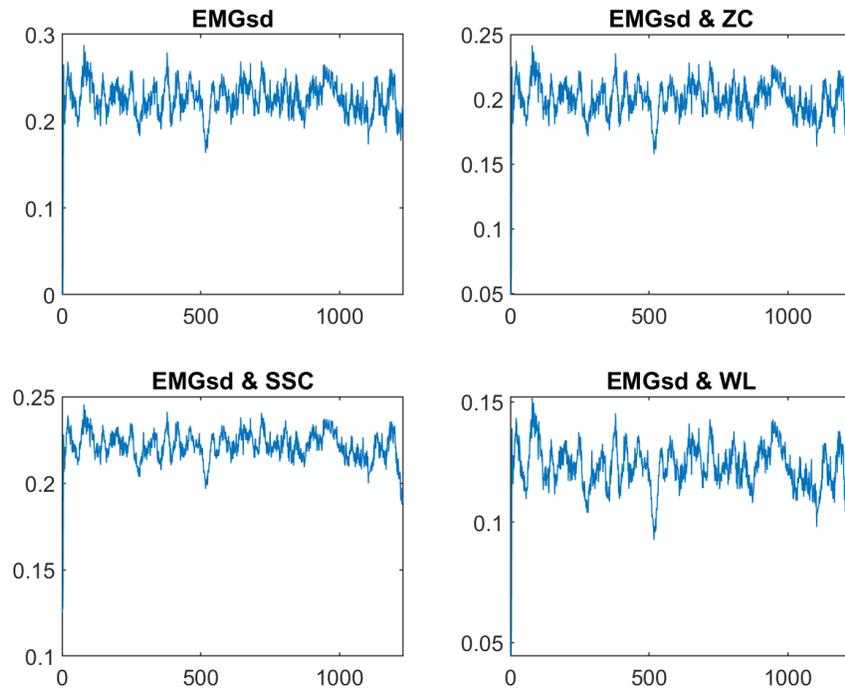


Figure 21 Combined Features using Fixed Weights, Whitened

Once each feature combination has been created from the dynamic calibration data, the new feature can be used to model EMG to torque. The performance is measured in terms of %MVC as the model order is varied from 5 to 20 in increments of 1. Additionally, the impact of whitening was included in the analysis; performance testing was completed for whitened data and unwhitened data. The results of the EMG to torque model for the various model orders is presented below in Figure 22 and Figure 23.

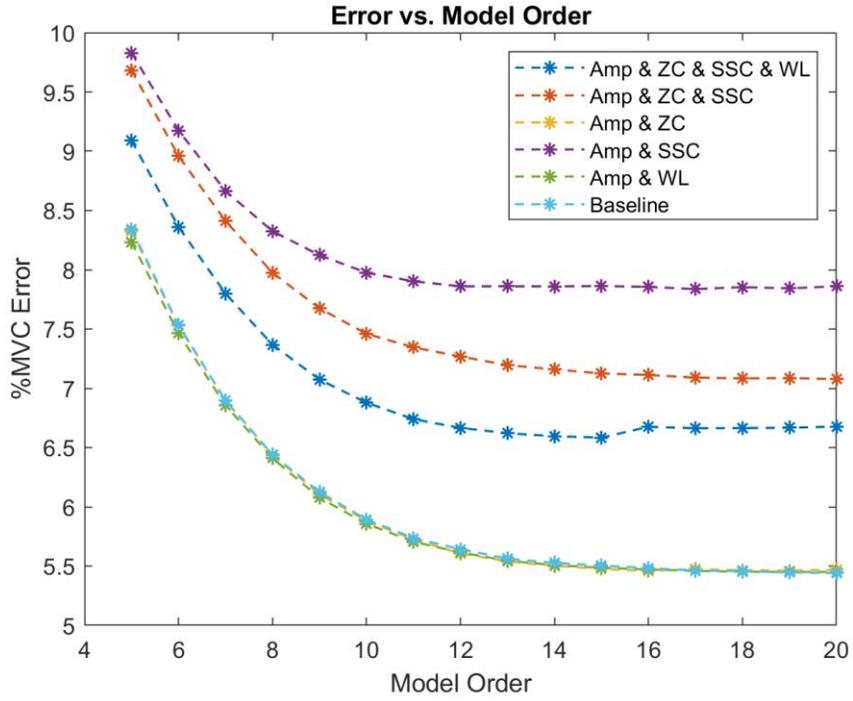


Figure 22 Torque Model Performance Expressed in %MVC, Unwhitened

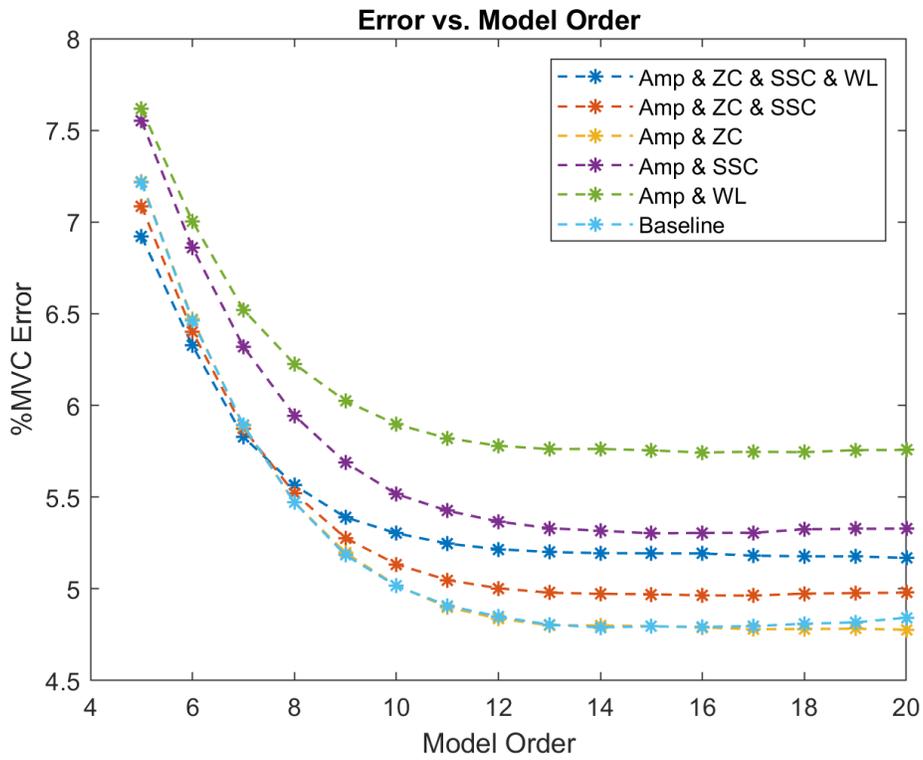


Figure 23 Torque Model Performance Expressed in %MVC, Whitened

The unwhitened data show worse baseline performance, at about 5.5% on average, than the whitened data, about 4.8% on average. (The performance of the whitened data is expected to be better than the unwhitened.) For both the unwhitened and whitened data, none of the feature combinations show any improvement on the performance, but some of the new features exhibit equal performance. For the unwhitened data, the performance of the baseline, zero crossing-EMG σ combination, and the waveform length-EMG σ combination are the same for model orders greater than 10. For the whitened data, the performance of the baseline and the zero crossing-EMG σ combination are the same until the higher model orders (about 17 and greater). The other combinations exhibit worse performance than the baseline.

4) Combination of Features using Inverse Variance Weights

The goal of combining the four features into a new EMG feature is to reduce the variance of the new EMG feature. One method to combine the four features into a single feature that also minimizes the variance of the combined feature is to weigh the features according to the inverse of their variance.

Assuming that the features are uncorrelated and have a non-zero variance, then the weight that will minimize the variance of the combined feature will be the inverse of the variance divided by the sum of the inverses. Induction can be used to prove that the minimum variance of the combined feature is achieved when each individual feature is weighed by the inverse of its variance. The proof is shown for the case of two features. The steps to minimizing the variance of the combined features is as follows:

The generic form of the combined feature is given as:

$$F_{new} = w_1 F_1 + w_2 F_2,$$

where F_{new} is the combined feature, F_1 and F_2 are the individual features and w_1 and w_2 are the weights being applied to each feature. The variance of F_{new} is given as:

$$var(F_{new}) = var(w_1 F_1 + w_2 F_2) = var(w_1 F_1) + var(w_2 F_2) = w_1^2 var(F_1) + w_2^2 var(F_2)$$

The sum of the weights, w_1 and w_2 , is equal to one, so the above equation can be rewritten as:

$$= w_1^2 var(F_1) + (1 - w_1)^2 var(F_2) = w_1^2 (var(F_1) + var(F_2)) - 2w_1 var(F_2) + var(F_2)$$

This equation is minimized when

$$w_1 = \frac{\frac{1}{\text{var}(F_1)}}{\frac{1}{\text{var}(F_1)} + \frac{1}{\text{var}(F_2)}} \text{ and } w_2 = \frac{\frac{1}{\text{var}(F_2)}}{\frac{1}{\text{var}(F_1)} + \frac{1}{\text{var}(F_2)}}$$

Proof of this principle can be found in (Shahar). When combining the four features using the inverse of their variances, the feature with the least variance receives the greatest weight in the combination and the feature with the greatest variance receives the least weight in the combination. To normalize the weights so that they sum to one, each individual inverse variance is divided by the sum of the four inverse variances. Without normalizing the weights such that they sum to one, the weights are large in magnitude, on the order of hundreds or thousands for the features, when the variance is less than one but greater than zero. Without normalization of the weight, an individual weight may enlarge a feature which will dominate the other three features in the combination. To avoid this, the denominator for each weight is the same, and the numerator changes depending on which feature is used. For example, the inverse variance weight for the feature, w_{amp} , is presented below. The weight for the other features is derived by changing the numerator.

$$w_{amp} = \frac{1/\sigma_{amp}^2}{1/\sigma_{amp}^2 + 1/\sigma_{zc}^2 + 1/\sigma_{ssc}^2 + 1/\sigma_{wl}^2}$$

Once each feature has been weighed, the features can be added together to produce the combined feature. Figure 24 shows the process of combining the features to produce the new EMG feature.

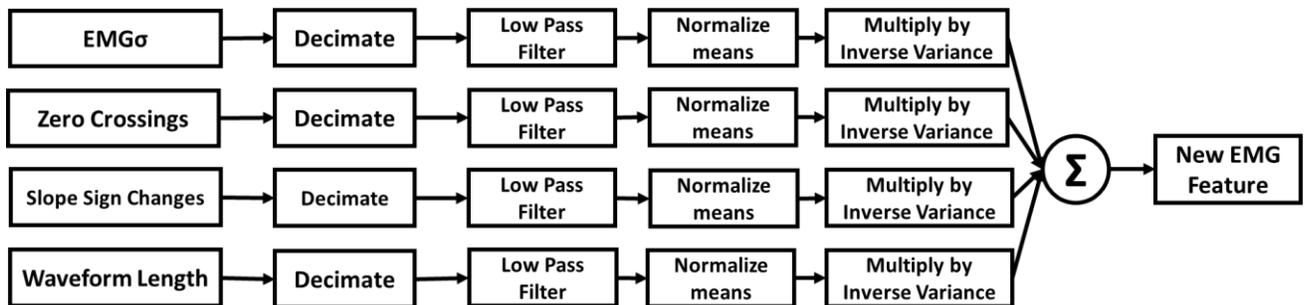


Figure 24 Procedure for Combining the Features using the Inverse Variance Weights

The following figures show the combined feature from the dynamic contraction data using the inverse variance weights computed individually for each subject. Figure 25 and Figure 26 show the combined features for a single subject's dynamic extension data.

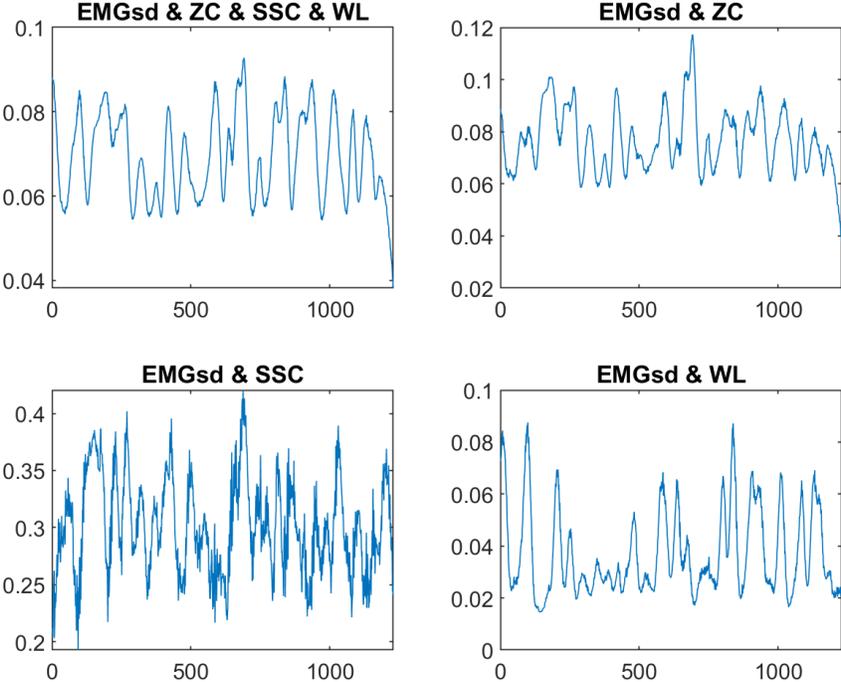


Figure 25 Combined Features using Inverse Variance Weights, Unwhitened

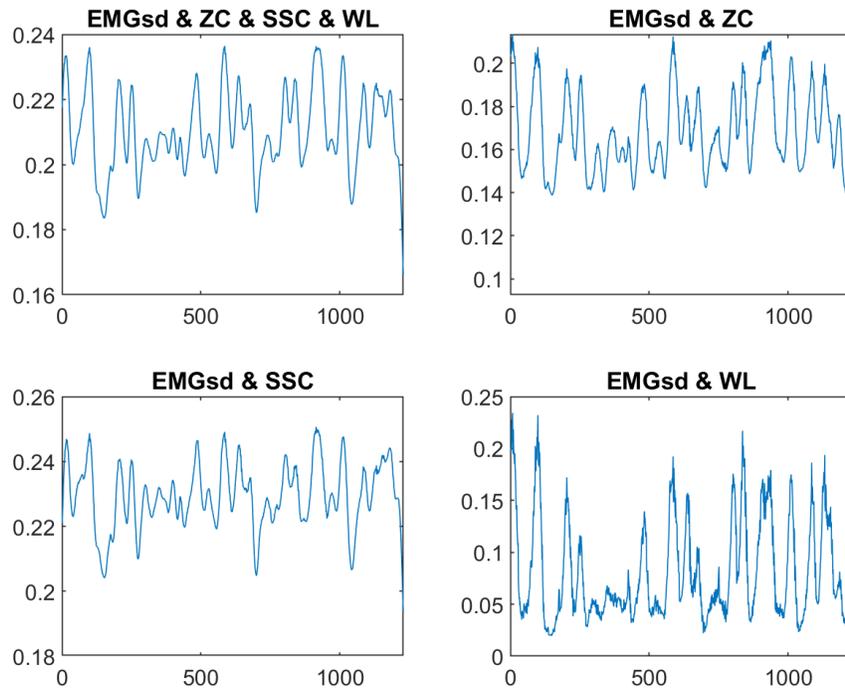


Figure 26 Combined Features using Inverse Variance Weights, Whitened

To test the performance of this combination method, the inverse variance weights were computed and applied to the training and testing dynamic calibration data. The four features were combined prior to the EMG to torque model. When combining the features using this method, no improvement was seen on the baseline performance. Without whitening, the following %MVC error was computed for torque model orders ranging from five to twenty incrementing by one.

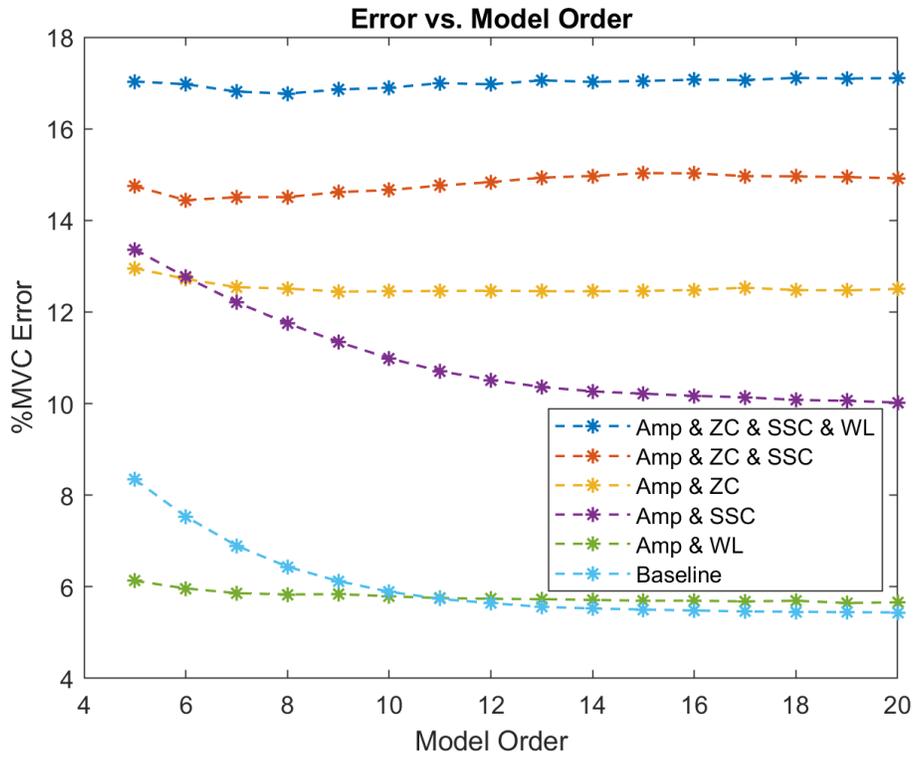


Figure 27 Torque Model Performance Expressed in %MVC, Unwhitened

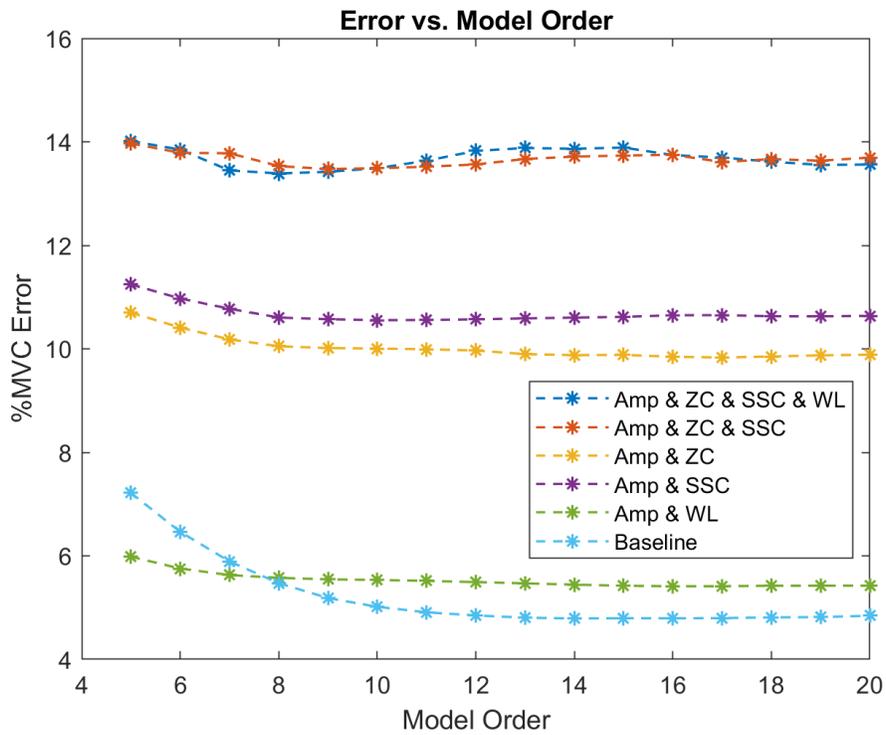


Figure 28 Torque Model Performance Expressed in %MVC, Whitened

5) Fixed Dynamics of the EMG to Torque Model

Previous attempts to combine the features did not show any improvement on the baseline performance. To better understand the mechanics behind the EMG to torque model when each feature is included in the system ID, the coefficients produced by the model were studied as well as the relative gain applied to the model for the dynamic data. A universal filter was developed for each feature using the coefficients for each feature across all subjects and channels. Once the universal filter was created and applied to each feature, the gains applied to each feature are computed using a zeroth order EMG to torque linear least squares. This process is outlined in Figure 29.

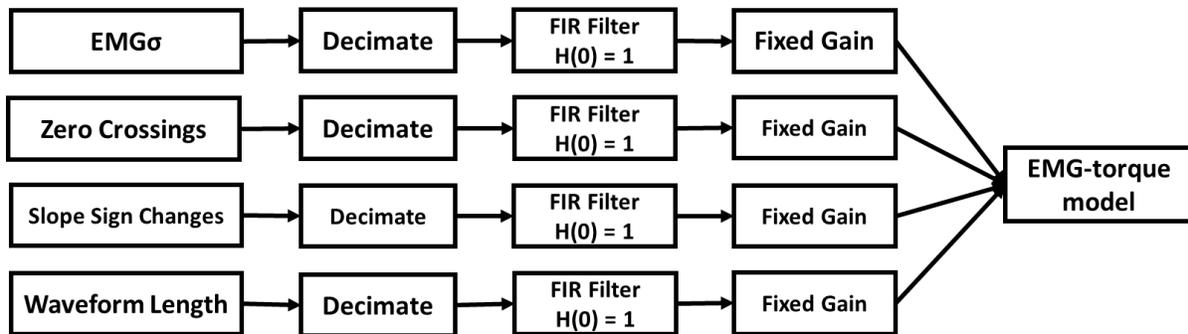


Figure 29 Fixed Dynamics Approach

The coefficients produced by the linear least squares' regression form a FIR filter. For the EMG features, these filters are low pass in nature and have a cut off frequency between 1 Hz to 2 Hz, depending on the feature. For each individual feature, a shimmer plot was created to show the individual magnitude response of the filter for each subject and each channel, for a total of 512 plots after the DC gain has been normalized to 1. A few of the magnitude responses were removed as outliers because their passband gain was significantly higher than the average passband gain. The shimmer plots for each individual feature, after outliers were removed, are shown below in Figure 30. The individual magnitude responses are plotted in grey, the median response is plotted in red and the average response is plotted in light blue.

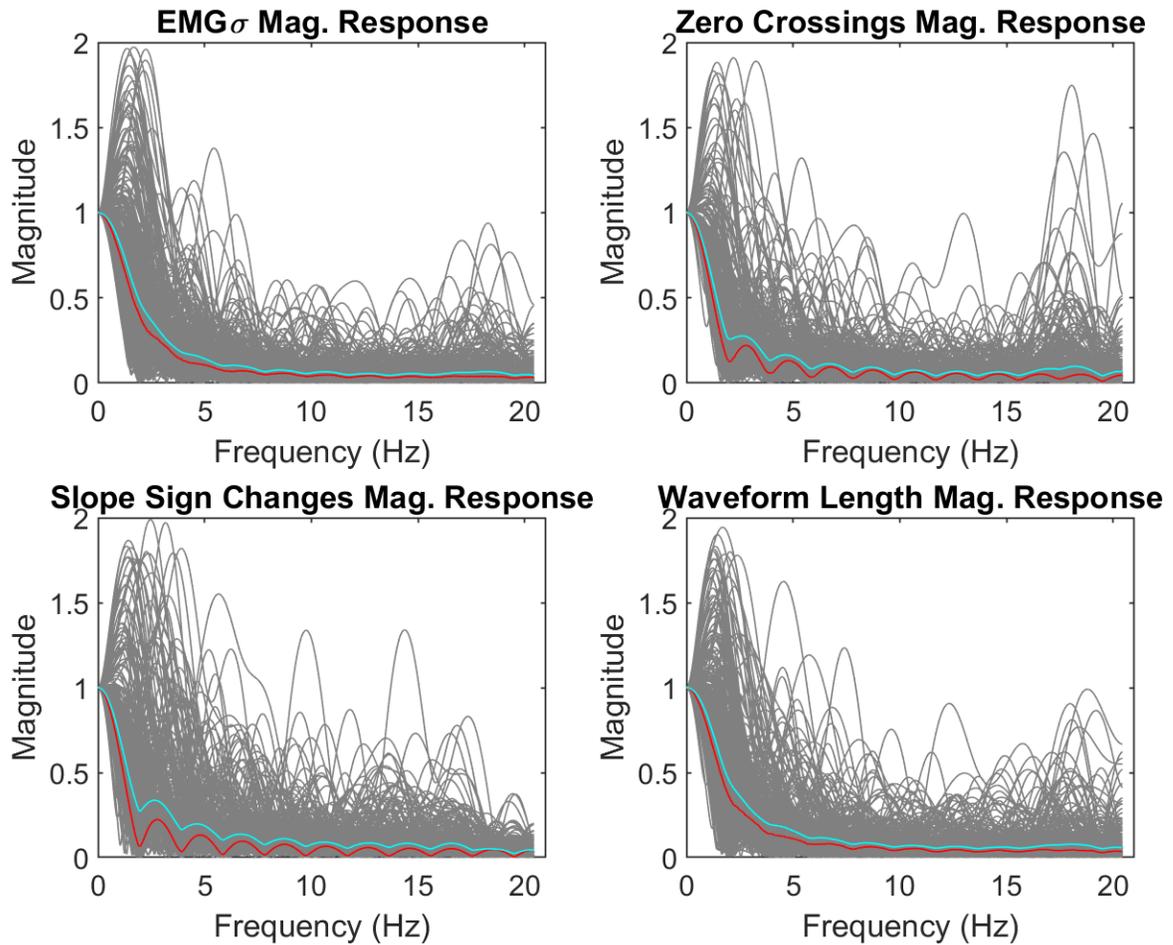


Figure 30 FIR Filter Magnitude Responses and Fixed Shape

To fix the shape of the filter, both the median and average magnitude responses were plotted and considered as a possible magnitude response for a fixed filter. The median filter was selected to be the basis for the fixed filter because it shows a more relaxed passband than the average filter shape. The median filter of the 20th order EMG to torque model was selected as the basis for the universal filter. Different model orders were compared and there were no significant differences seen between the shape of the filters that result from the model of orders ranging between 15 to 20, tested in increments of one. The -3dB point moves closer to 0 Hz as the model order increases, but the change in the -3dB point is less than 0.1 Hz for each increase in model order.

The following figure, Figure 31, shows the universal FIR filters that are generated in the 20th order EMG to torque model for each feature. The response of EMG σ and waveform length

are very close in shape and have a -3dB point of 1.12 Hz. The same behavior is true for the zero crossings and slope sign changes. They have a similar shape and a -3dB point of 0.88 Hz.

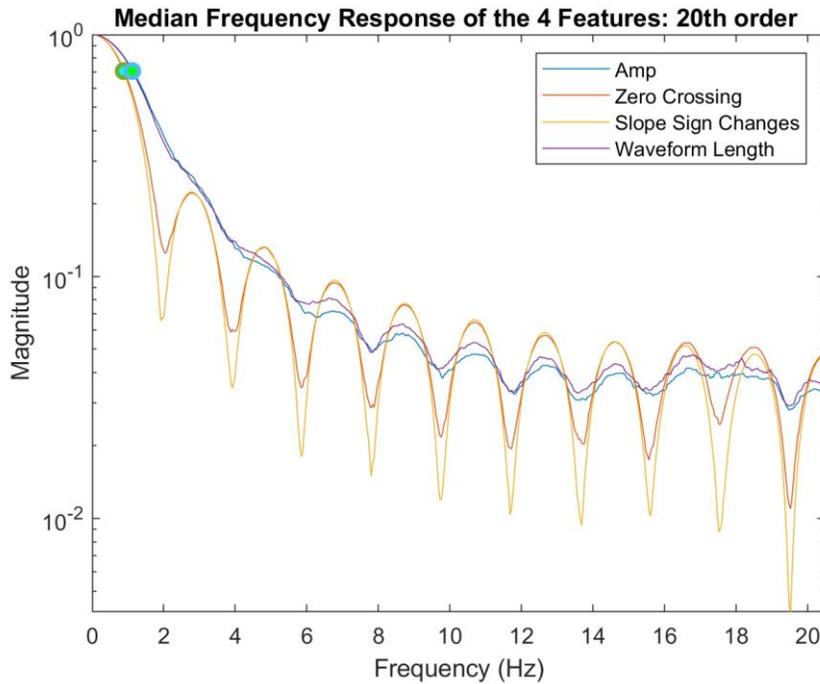


Figure 31 Median Frequency Response of the Four Features

Once the desired response was known, a few filters, IIR and FIR, were designed to match the desired response. The IIR filter for the zero crossings and slope sign changes was designed as 2nd order lowpass Butterworth filter with a cutoff frequency of 0.88 Hz. For the EMG σ and waveform length filter, a 2nd order lowpass Butterworth filter with a cutoff frequency of 1.12 Hz was designed. Neither of the IIR filters were able to match the desired magnitude response perfectly. The roll off of the IIR filters had a different slope than the desired magnitude response and the stopband of the IIR filter was unable to achieve the ripples seen in the FIR filters created in the EMG to torque model, so a FIR filter was designed using the frequency sampling method.

The FIR filters were developed using the frequency sampling method. The frequency sampling method requires the desired magnitude response in the frequency domain and the phase of the filter. The inverse Fourier transform of the desired response is computed, and a Hamming window is applied to compute the filter's b coefficients (for a FIR filter, $a = 1$). The size of the Hamming window is equal to the order of the filter plus one.

For the EMG σ and waveform length, a 120th order FIR filter was designed to match the desired response. The -3dB point is at 0.88 Hz. For the zero crossing and slope sign changes, a 150th order FIR filter was created to match the desired response. The -3dB point of this filter is at 1.02 Hz. The desired and the designed filter responses are shown below in Figure 32.

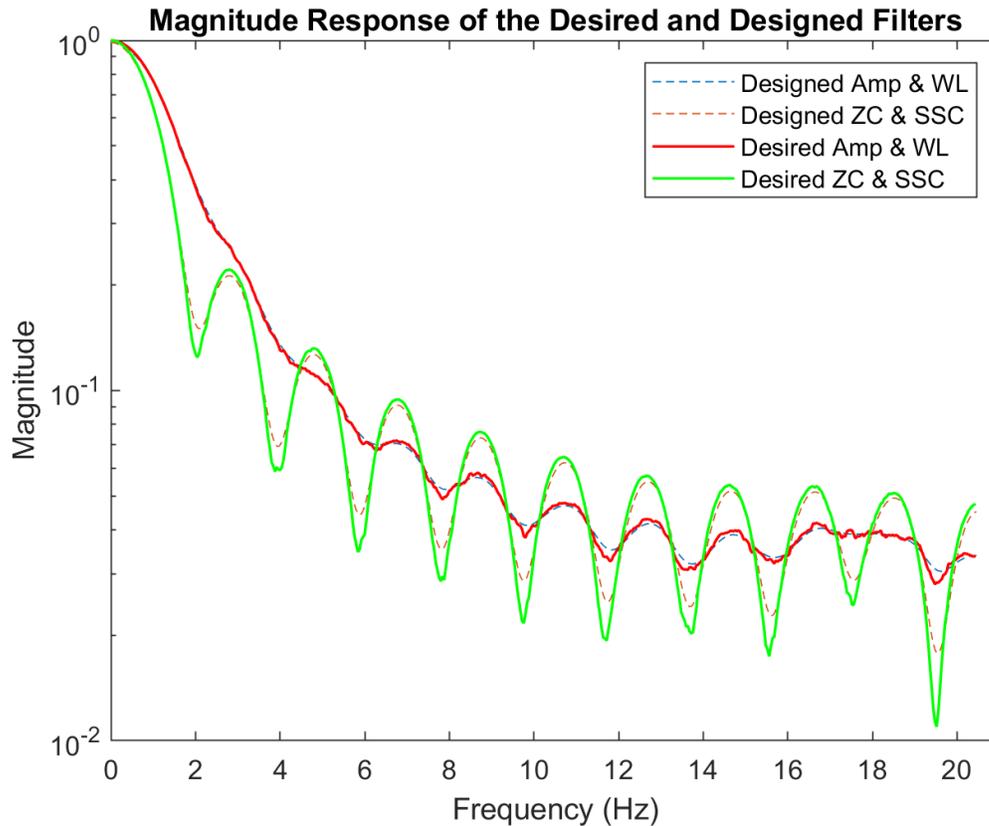


Figure 32 Desired and Designed FIR Filter Response for the Feature Pairs

Once the filter was fixed at the input of the EMG to torque model, the gain applied to each feature was computed by running the linear least squares' regression with a model order set to 0. The gain for each feature for each channel of each subject was computed. Results of the EMG to torque model with the universal filter and gains computed for each feature for each subject are shown in Figure 33. The %MVC error could not be plotted against model order in this case, so it was plotted for each subject. Performance of the zeroth order EMG to torque model with the universal filter was significantly higher, an average of 13.73% for the 64 subjects, than the baseline performance of 4.8% for EMG σ only.

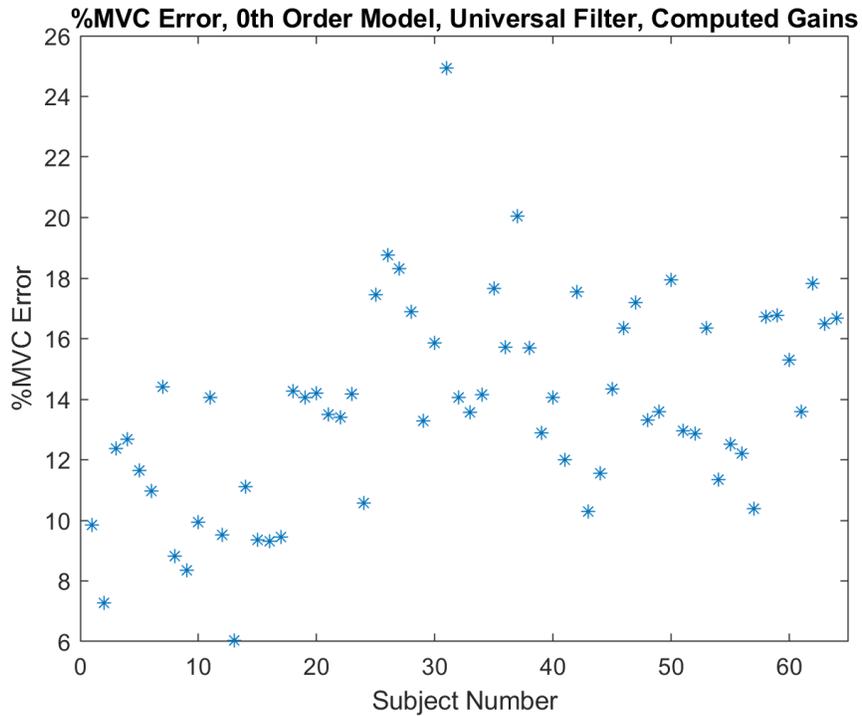


Figure 33 Performance of the EMG Torque Model, 0th Order, Universal FIR Filter, Computed Gains, Whitened

The gains were computed for each feature across all subjects to determine if the gains were relatively consistent, so that they could also be fixed. The distribution of the individual gains was significantly spread across the samples. The following figures show the mean and standard deviations of the 0th order gains for each electrode channel.

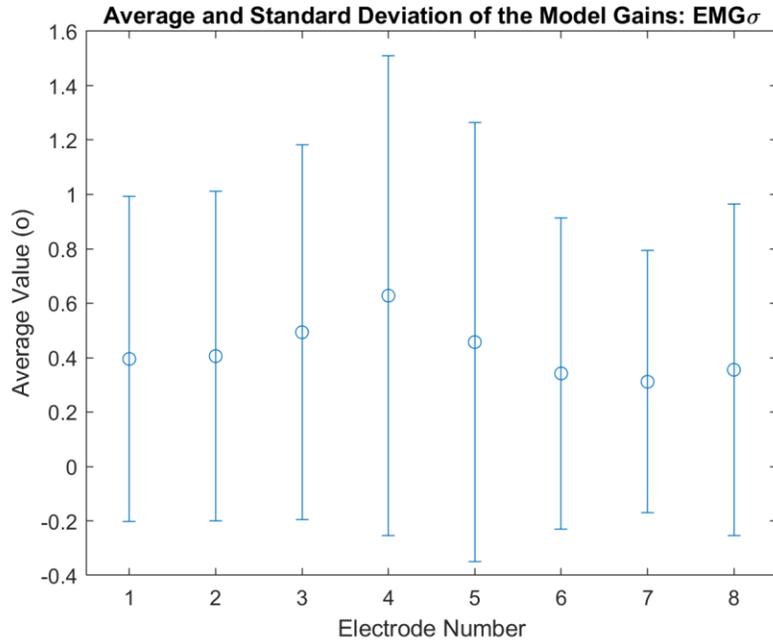


Figure 34 Mean and Standard Deviation of the Model Gains per Electrode for EMG σ

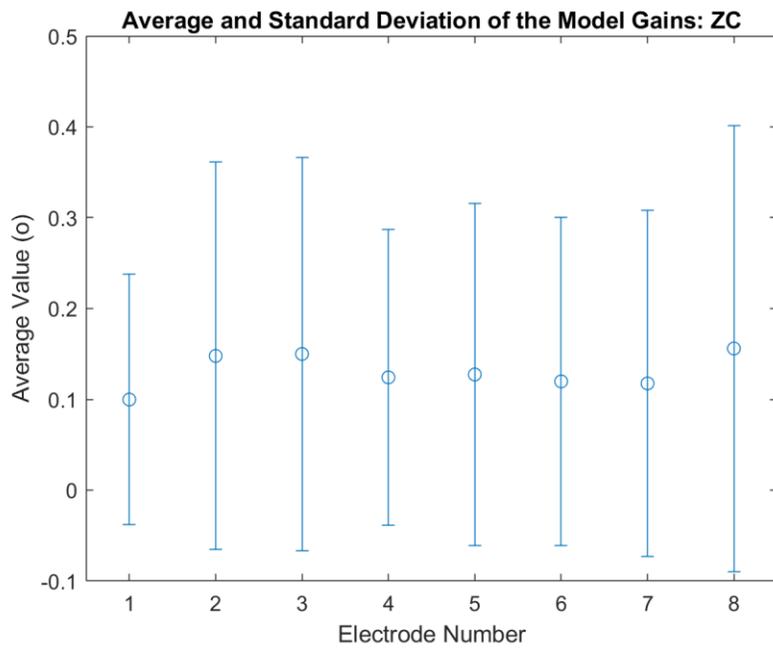


Figure 35 Mean and Standard Deviation of the Model Gains per Electrode for Zero Crossing Feature

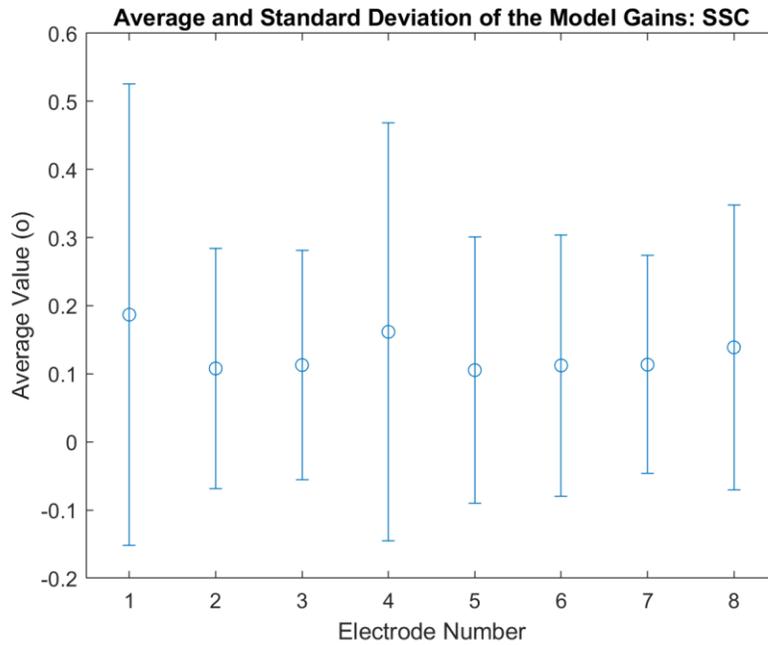


Figure 36 Mean and Standard Deviation of the Model Gains per Electrode for Slope Sign Changes Feature

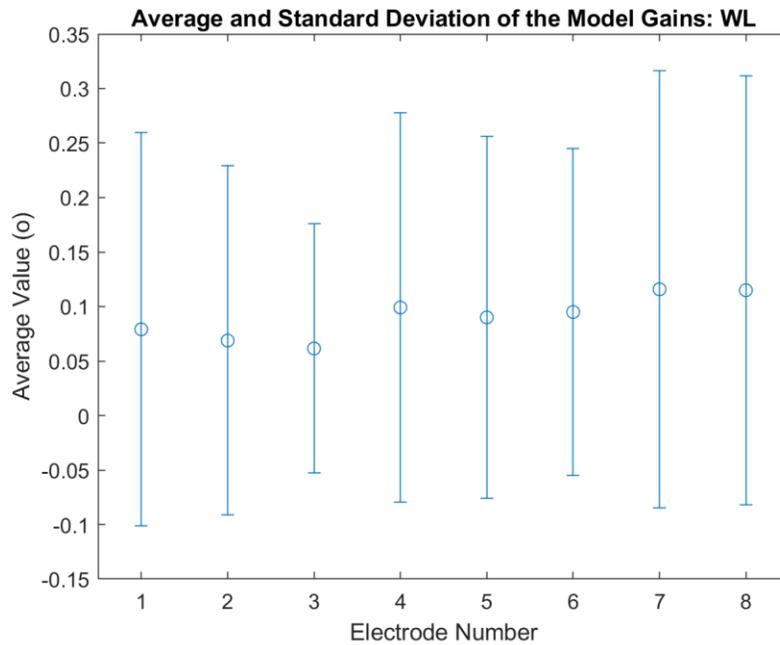


Figure 37 Mean and Standard Deviation of the Model Gains per Electrode for Waveform Length Feature

The average value of the individual gains was used as the universal gain. Performance of the zeroth order EMG to torque model with the universal filter and universal gains for each feature is shown in Figure 38. The %MVC error computed with the universal filter and the

universal gains is higher than the %MVC error seen with only the universal filter and the computed gains (shown in Figure 33).

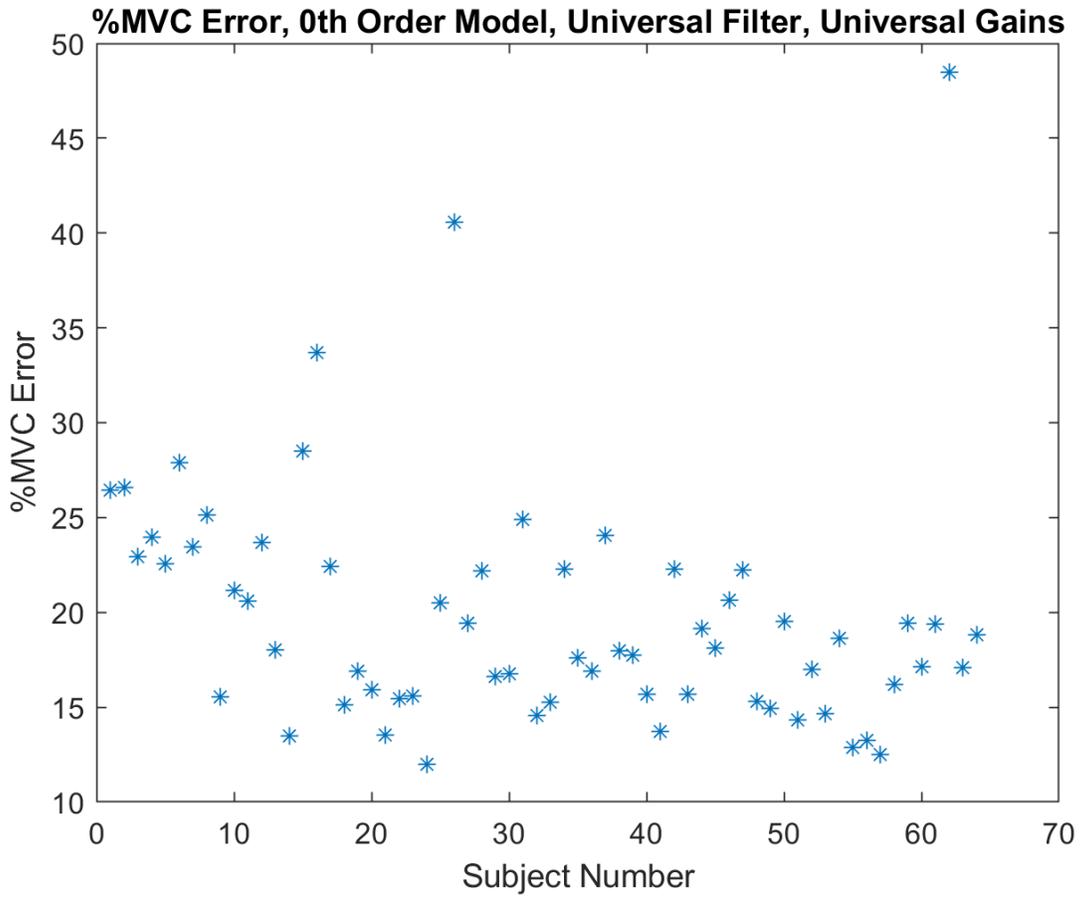


Figure 38 Performance of the EMG Torque Model, 0th Order, Universal FIR Filter, Universal Gains, Whitened

Performance of the model without fixing the dynamics is significantly better than when the dynamics, FIR filter and gains. Breakdown of the EMG to torque model was helpful to show that there is significant variation on a subject to subject basis when combining the four features in the EMG to torque model.

IV. Root Difference of Squares Processing

A. Derivation of the Optional Estimate of $EMG\sigma$ with RDS Processing

This chapter has been published as: Haopeng Wang, Kiriaki J. Rajotte, He Wang, Chenyun Dai, Ziling Zhu, Moinuddin Bhuiyan, Xinming Huang and Edward A. Clancy, “*Optimal Estimation of EMG Standard Deviation ($EMG\sigma$) Requires Noise Subtraction in the Power Domain: Model-Based Derivations and their Implications*,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2019. Available online at:

<http://ieeexplore.ieee.org.ezproxy.wpi.edu/stamp/stamp.jsp?tp=&arnumber=8890663&isnum>

Abstract—Typical electromyogram (EMG) processors estimate EMG signal standard deviation ($EMG\sigma$) via moving average root mean square (RMS) or mean absolute value (MAV) filters, whose outputs are used in force estimation, prosthesis/orthosis control, etc. In the inevitable presence of additive measurement noise, some processors subtract the noise standard deviation from EMG RMS (or MAV). Others compute a root difference of squares (RDS)—subtract the noise variance from the square of EMG RMS (or MAV), all followed by taking the square root. Herein, we model EMG as an amplitude-modulated random process in additive measurement noise. Assuming a Gaussian (or, separately, Laplacian) distribution, we derive analytically that the maximum likelihood estimate of $EMG\sigma$ requires RDS processing. Whenever that subtraction would provide a negative-valued result, we show that $EMG\sigma$ should be set to zero. Our theoretical models further show that during rest, approximately 50% of $EMG\sigma$ estimates are non-zero. This result is problematic when $EMG\sigma$ is used for real-time control, explaining the common use of additional thresholding. We tested our model results experimentally using biceps and triceps EMG from 64 subjects. Experimental results closely followed the Gaussian model. We conclude that EMG processors should use RDS processing and not noise standard deviation subtraction.

Index Terms—Biological system modeling, biomedical signal processing, electromyogram, electromyogram (EMG) amplitude estimation, electromyography, myoelectric signal processing.

INTRODUCTION

THE surface electromyogram (EMG) interference pattern has commonly been processed by the

cascade operations of highpass filtering (to remove DC offsets and attenuate motion artifacts); optional pre-whitening [1-3]; and then taking its moving average root mean square (RMS), moving average mean absolute value (MAV), or by rectifying the signal followed by lowpass filtering. If EMG is modeled as an amplitude-modulated random process, then these schemes estimate its time-varying standard deviation ($EMG\sigma$). For constant-force, non-fatiguing contractions, it has been shown that RMS processing is the optimal estimate of $EMG\sigma$ if the *noise-free* EMG signal is modeled as Gaussian distributed [2, 4-6], and that MAV processing is optimal if the *noise-free* EMG signal is modeled as Laplacian distributed [7]. $EMG\sigma$ has been used to estimate torque [8-13] and mechanical impedance about a joint [14-19], in motor control research [20], and in applications including prosthesis control [21-23], ergonomics [24, 25] and biomechanics [26, 27].

However, EMG is always measured in the presence of additive measurement noise, i.e., noise that exists independent of the level of muscle effort. This noise arises from the measurement apparatus (thermal and active device noise), radiated electromagnetic interference, electrode-to-skin contact resistance [28], unrelated electrophysiological activity, etc. [29]. This noise has an average RMS intensity that is 1.1–4.5% of the RMS EMG at maximum voluntary contraction (MVC) [3, 8, 9, 30-34]. Consequently, the signal to noise ratio (SNR) is low at low contraction levels.

Thus, researchers have proposed alterations to their EMG processors and/or models to include noise. Kaiser and Peterson [1] found that the shape of their whitening filter should be a function of the contraction level, with lower high-frequency gain during low contraction levels. Parker *et al.* [35-37] modeled noise as an additive (white Gaussian) process when *solving* for an optimal multistate EMG classifier, and when *analyzing* (but not *solving*) $EMG\sigma$ estimators. This additive noise model is now common (e.g., [3, 38-40]). Clancy and Farry [3] whitened the raw EMG, then attenuated additive noise using an adaptive Wiener filter. A Wiener filter is the optimal *linear* filter for attenuating additive noise, but is not necessarily the optimal filter overall. Many papers within the ergonomics literature routinely subtract the standard deviation of the background noise from RMS (or MAV) estimates [41]. However, it has been theoretically *argued* [42, 43] that the root difference of squares (RDS) [i.e., subtracting the noise variance from the square of EMG RMS (or MAV), all followed by taking the square root] is the correct approach. An experimental comparison found that RDS processing performs better than standard deviation subtraction [44].

The argument for RDS processing is based on the fact that if the signal and noise are

independent, then their variances add—in *theory*. However, to our knowledge, this proposed processor has not been *derived* (i.e., *solved* for, based on a model) as a statistical estimator in the published literature (although one unpublished preliminary result appears in [45]). Solution via an estimator can demonstrate the optimality (or lack thereof) of a processor and expose its statistical properties. Herein, we provide this derivation, some of its properties and experimental evaluation of the derived optimal results, all for the case of constant-effort contraction.

Mathematical Models of EMG in Additive Noise

Consider an amplitude modulated model of the measured EMG signal, $m[n]$, during constant-effort contraction as [2, 5, 35-37]:

$$m[n] = s \cdot x[n] + v[n], \quad 0 \leq n < N \quad (1)$$

where n is the discrete-time sample index, $s \equiv EMG\sigma$ is the standard deviation (i.e., modulation) of the noise-free EMG, $(s \cdot x[n])$ is the noise-free EMG signal and $v[n]$ is additive noise. Let $x[n]$ be zero mean, unit-variance, wide-sense stationary, correlation-ergodic and have independent samples (i.e., via pre-whitening). Let $v[n]$ be similarly specified, but of variance equal to q^2 and independent of $x[n]$. Let \underline{m} , \underline{x} and \underline{v} be vectors comprised of N successive samples of each respective random variable.

Gaussian Model—EMG σ Estimate [45, 46]

Let both \underline{x} and \underline{v} be jointly Gaussian. Then, \underline{m} is jointly Gaussian with zero mean and covariance matrix: $K_{mm} = \sigma_m^2 I$, where $\sigma_m^2 = s^2 + q^2$ and I is the identity matrix. Thus, the probability density function (PDF) for \underline{m} , given that the standard deviation of the noise-free EMG is $s \equiv EMG\sigma$, is:

$$p_{\underline{m}|s}(\underline{M}|s) = \frac{e^{-\frac{\underline{M}^T K_{mm}^{-1} \underline{M}}{2}}}{(2\pi)^{N/2} |K_{mm}|^{1/2}} = \frac{e^{-\frac{\sum_{n=0}^{N-1} M^2[n]}{2(s^2+q^2)}}}{[2\pi(s^2+q^2)]^{N/2}}, \quad (2)$$

where \underline{M} denotes an instance of the random vector \underline{m} .

The maximum likelihood (ML) estimate of s is the value \hat{s} which maximizes the above PDF. A monotonic transformation of the PDF does not alter the location of the maximum. Thus, taking the natural logarithm yields:

$$\ln[p_{\underline{m}|s}(\underline{M}|\hat{s})] = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\hat{s}^2 + q^2) - \frac{\sum_{n=0}^{N-1} M^2[n]}{2(\hat{s}^2+q^2)}. \quad (3)$$

Differentiating the above with respect to \hat{s} gives:

$$\frac{\partial \ln[p_{m|s}(M|\hat{s})]}{\partial \hat{s}} = -\frac{N}{2} \frac{2\hat{s}}{\hat{s}^2 + q^2} + \frac{\hat{s} \sum_{n=0}^{N-1} M^2[n]}{(\hat{s}^2 + q^2)^2}. \quad (4)$$

Setting this derivative to zero and manipulating leads to a quadratic equation for \hat{s}^2 , the square root of which provides our intermediate result. The quadratic equation has two solutions. But, one of these solutions is not real-valued, so can be eliminated. The retained intermediate result, written as a discrete-time filter, is:

$$\hat{s}[n] = \sqrt{\left(\frac{\sum_{i=0}^{N-1} M^2[n-i]}{N}\right) - q^2}. \quad (5)$$

The parenthesized term within the square root is the mean square value. Hence, the noise correction is made via RDS processing.

The second derivative of (3) with respect to \hat{s} , evaluated at the location of the intermediate result specified by (5) is:

$$\frac{\partial^2 \ln[p_{m|s}(M|\hat{s})]}{\partial \hat{s}^2} = \left[\frac{2N^3}{\left(\sum_{i=0}^{N-1} M^2[n-i]\right)^2} \right] \left[q^2 - \frac{\sum_{i=0}^{N-1} M^2[n-i]}{N} \right]. \quad (6)$$

This second derivative is less than or equal to zero, indicating a local maximum (and *not* a minimum), when $\frac{1}{N} \sum_{n=0}^{N-1} M^2[n-i]$ exceeds the noise variance q^2 . This condition is almost always satisfied during active muscle contraction, but not during low-level contractions or rest. When the condition is not satisfied, maximization with respect to \hat{s} of the PDF occurs at the boundary constraint where $\hat{s} = 0$ [47]. Hence, the complete solution for this ML estimate is:

$$\hat{s}_{\text{RMS}}[n] = \sqrt{\max\left[0, \left(\frac{\sum_{i=0}^{N-1} M^2[n-i]}{N}\right) - g^2 q^2\right]}, \quad (7)$$

where “max” denotes the maximum value operator and the “RMS” subscript emphasizes the use of an RMS processor. Constant scaling factor g has been inserted into this solution, since some applications prefer to artificially inflate the noise threshold. For example, in myoelectric prosthesis control, $g > 1$ helps to insure that the prosthesis is not actuated during rest. For the optimum ML estimate, $g = 1$.

Denote the term in the rounded parenthesis of (7) (i.e., the mean square value of the measured EMG signal) as y . This random variable is Gamma distributed as:

$$p_y(Y) = \frac{Y^{\frac{N}{2}-1} e^{-\frac{Y \cdot N}{2\sigma_m^2}}}{\left(\sigma_m \sqrt{\frac{2}{N}}\right)^N \Gamma\left(\frac{N}{2}\right)} \mu(Y), \quad (8)$$

where $\Gamma(\cdot)$ is the Gamma function and $\mu(\cdot)$ is the step function. Its cumulative density function (CDF) is:

$$P_{y \leq 0}(Y) = 1 - \sum_{k=0}^{\frac{N}{2}-1} \frac{\left(\frac{N}{2}\right)^k Y^k e^{\frac{-Y \cdot N}{2\sigma_m^2}}}{k!} \mu(Y), N \text{ even.} \quad (9)$$

When the muscle is at rest, the true EMG σ is zero ($s = 0$) and the variance of the measured EMG signal is $\sigma_m^2 = q^2$. A fraction of the EMG σ estimates—but not all—will be zero (due to the noise variance subtraction). This probability of estimating a zero value during rest is the CDF of y , evaluated at $Y = g^2 q^2$ (with $s = 0$). This probability, for N even, is:

$$P_{y \leq g^2 q^2, Rest}(Y) = \left[1 - \sum_{k=0}^{\frac{N}{2}-1} \frac{\left(\frac{N}{2}\right)^k g^{2k} e^{\frac{-g^2 N}{2}}}{k!} \right] \mu(Y). \quad (10)$$

Note that this probability is *not* a function of the noise variance and is *only* a function of N and g . Fig. 1 shows this probability as a function of N for four possible values of g . Equation 10 and Fig. 1 show that for $g > 1$, a negative-valued subtraction result within (7) is more likely, producing a higher probability of estimating $\hat{s} = 0$. Conversely, for $g < 1$, a negative-valued subtraction result is less likely, producing a lower probability of estimating $\hat{s} = 0$.

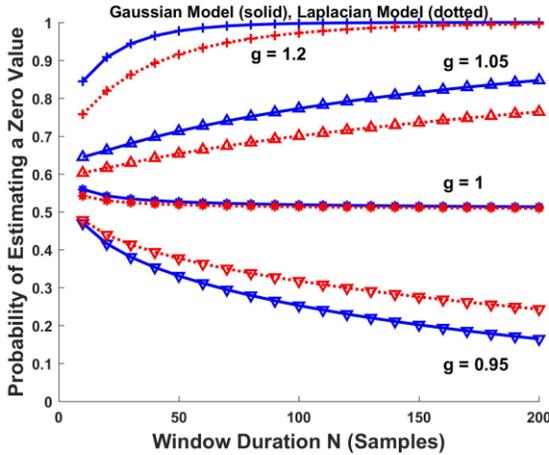


Fig. 1. Probability of estimating a zero EMG σ value during rest for theoretical Gaussian model (moving average RMS processing; solid blue) and Laplacian model (moving average MAV processing; dashed red) as a function of number of independent samples N , for four different noise gain values “ g ”.

Laplacian Model—EMG σ Estimate [7, 45, 46]

MAV processing has been shown to be the ML estimate of EMG σ , if the PDF is Laplacian [7]. So that the additive noise model has a Laplacian PDF, we directly model the measured EMG samples $m[n]$ as being independent and of a Laplacian PDF, without explicit specification of the PDFs of $x[n]$ and $v[n]$. (Note that if $x[n]$ and $v[n]$ are each modeled as Laplacian, then their sum is *not* Laplacian.) Nonetheless, if $x[n]$ and $v[n]$ are assumed independent, then their

variances again add. Thus, the measured EMG again has variance: $s^2 + q^2$, and the PDF for sample $m[n]$ is [48]:

$$p_{m[n]|s}(M[n]|s) = \frac{\sqrt{2}}{2} \cdot \frac{e^{\frac{-\sqrt{2}}{(s^2+q^2)^{1/2}} |M[n]|}}{(s^2+q^2)^{1/2}}. \quad (11)$$

Since the samples of the EMG vector \underline{m} are independent, its joint PDF is the product of the N individual PDFs, which simplifies to:

$$p_{\underline{m}|s}(\underline{M}|s) = \left[\frac{\sqrt{2}}{2 (s^2+q^2)^{1/2}} \right]^N e^{\frac{-\sqrt{2}}{(s^2+q^2)^{1/2}} \sum_{n=0}^{N-1} |M[n]|}. \quad (12)$$

Similar to the Gaussian case above, maximum likelihood estimation of s is found by taking the natural logarithm of the PDF, differentiating with respect to \hat{s} , setting this derivative to zero and solving for \hat{s} . Again, the second derivative proves this intermediate result to, in fact, be a minimum, subject to the same boundary constraint where $\hat{s} = 0$. The complete filter for this ML estimate, again inserting a scaling factor g for the noise, is:

$$\hat{s}_{MAV}[n] = \sqrt{\max \left[0, \left\{ \left(\frac{\sqrt{2}}{N} \sum_{i=0}^{N-1} |M[n-i]| \right)^2 - g^2 q^2 \right\} \right]}. \quad (13)$$

Denote the term in the curly brackets of (13) as w . The PDF for this random variable is:

$$p_w(W) = \frac{e^{-\frac{N\sqrt{W}}{\sigma_m}}}{2} \cdot \left[\sum_{k=0}^{N-1} \left(\left\{ \frac{N}{\sigma_m \sqrt{W}} - \frac{(N-1-k)}{W} \right\} \cdot \prod_{p=1}^{N-1-k} \left\{ \frac{N\sqrt{W}}{\sigma_m p} \right\} \right) \right] \mu(W). \quad (14)$$

Its CDF is:

$$P_{w \leq}(W) = \left\{ 1 - e^{-\frac{N\sqrt{W}}{\sigma_m}} \left[\sum_{k=0}^{N-1} \left(\prod_{p=1}^{N-1-k} \frac{N\sqrt{W}}{\sigma_m p} \right) \right] \right\} \mu(W). \quad (15)$$

The probability of estimating a zero value during rest is the CDF evaluated at $W = g^2 q^2$ (with $s = 0$):

$$P_{w \leq g^2 q^2, Rest}(W) = \left\{ 1 - e^{-Ng} \left[\sum_{k=0}^{N-1} \left(\prod_{p=1}^{N-1-k} \frac{Ng}{p} \right) \right] \right\} \mu(W). \quad (16)$$

Again, the probability of a zero value is only related to N and g . Fig. 1 shows this probability as a function of N for four possible values of g .

Experimental Evaluation of the Models

Experimental Data Set

Data from 64 subjects acquired during four prior experiments with overlapping protocols were used for this study [3, 8, 30, 33]. Re-analysis of these data was exempted from human studies supervision by the WPI Institutional Review Board. Subjects had no known neuromuscular deficits of the right shoulder, arm or hand. In each experiment (see Fig. 1 in [8] for a photograph of the

most recently used experimental apparatus), a subject was seated and secured with seat belts. Their right shoulder was abducted 90°, elbow flexed 90°, and hand supinated perpendicular to the floor. Their wrist was cuffed to a load cell to measure constant-posture elbow torque.

The skin above the triceps and biceps muscles was scrubbed with an alcohol wipe. Gel was applied in the latter two studies. Four bipolar EMG electrode-amplifiers were secured over each of the triceps and biceps muscles, in a tightly-spaced transverse row centered on the muscle midline, midway between the elbow and the midpoint of the upper arm. Each electrode-amplifier had stainless steel, hemispherical contacts of diameter 4 or 8 mm, separated 10 mm edge-to-edge, oriented along the long axis of the muscle. A reference electrode was secured alongside the active electrodes. Each EMG channel had selectable gain, a CMRR ≥ 90 dB at 60 Hz, a 10 or 15 Hz highpass filter (second or fourth order), and a 1800 or 2000 Hz lowpass filter (fourth order). EMG and load cell data were sampled at 4096 Hz at 16-bit resolution. Achieved force was fed back in a real-time display, along with a force target.

After a brief warm-up, separate elbow flexion and extension maximum voluntary contraction (MVC) forces were measured, without the use of force feedback. At least 20–30 minutes had elapsed between the time at which the electrodes were mounted and the completion of these MVC measurements. Then, constant-force 50% MVC extension trials, 50% MVC flexion trials and 0% MVC trials (arm at rest, removed from the wrist cuff) were acquired for 5 s each, using force feedback. (Only one of each type of trial was used in our analysis.) Two or three minutes of rest was provided between trials to avoid cumulative fatigue. Each of the eight, 5-s duration EMG signals from a trial was defined as an “epoch.” Before any further use off-line, each epoch was highpass filtered (15 Hz cut-off, fourth-order Butterworth); IIR notch filtered at 60 Hz and its harmonics (second-order); when selected, adaptively pre-whitened [3, 49]; and bandlimited to 600 Hz [50] (fourth-order Butterworth lowpass). Then the first 500 ms of each epoch was omitted to account for filter start-up transients.

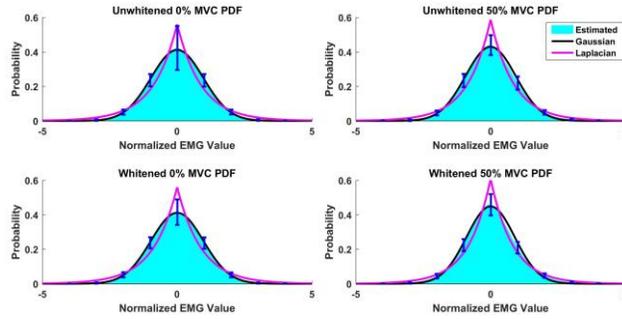


Fig. 2. Top shows ensemble-average PDF estimates of *unwhitened* EMG during 0% MVC (left) and 50% MVC (right), as well as best-fit theoretic Gaussian and Laplacian PDFs. Bottom shows corresponding PDF estimates from *whitened* EMG. $N = 512$ epochs from 64 subjects. Error bars in each plot show ± 1 std. dev. for the ensemble-average estimates.

Evaluating Model Assumptions—EMG PDF

We evaluated the model assumptions related to the first-order PDF of EMG, both at rest and during 50% MVC trials, with and without whitening. During 50% extension trials, only the four epochs from triceps electrodes were examined; during 50% flexion trials, only the four epochs from biceps electrodes were examined. A total of 512 epochs (64 subjects \times 8 electrodes/subject) were available at 0% (rest) and at 50% MVC (combining extension and flexion). Each EMG epoch was normalized to a sample variance of one and a histogram PDF estimate formed (500 bins, equally spaced over the range from -5 to $+5$). The ensemble histogram sample means and standard deviations are shown in Fig. 2.

Best matching between the ensemble vs. theoretic Gaussian/Laplacian PDFs did *not* occur when using theoretic PDFs of unit variances. Thus, the absolute error difference between each ensemble and theoretic PDF was computed for theoretic PDF standard deviations between 0.5 and 2 (increment of 0.01). The minimum area and its corresponding theoretic PDF standard deviation are shown in Table I (see also Fig. 2). In all cases, the data more closely followed the Gaussian model. Kolmogorov-Smirnoff tests between the experimental ensemble PDFs and each of the Gaussian and Laplacian PDFs were not sensitive, finding no statistically significant differences using either the Gaussian model ($p > 0.99$) or the Laplacian model ($p > 0.31$), for the four combinations of effort level (0% MVC, 50% MVC) and whitening. Thus, we computed the absolute area difference between each of the 512 histogram PDF estimates vs. the Gaussian/Laplacian PDFs, finding the best fit standard deviation for each. Paired sign tests (Bonferroni corrected) found the Gaussian PDF to be a better fit ($p < 10^{-6}$) for each of the four

combinations.

Evaluating Estimates of $EMG\sigma$

Historically, quantitative evaluation of constant-effort $EMG\sigma$ has used the ratio of the estimate mean to its standard deviation (the inverse of the coefficient of variation), denoted the SNR. With

TABLE I
ABSOLUTE AREA DIFFERENCES BETWEEN EXPERIMENTAL ENSEMBLE PDFS AND GAUSSIAN/LAPLACIAN PDFS. PARENTHESES LIST STANDARD DEVIATION AT WHICH AREA DIFFERENCE WAS ASSESSED (I.E., STANDARD DEVIATION AT WHICH THE ABSOLUTE ERROR DIFFERENCE BETWEEN EACH ENSEMBLE AND THEORETIC PDF WAS MINIMIZED).

EMG Processing	Gaussian Model		Laplacian Model	
	0% MVC	50% MVC	0% MVC	50% MVC
Unwhite	0.0241 (0.97)	0.0530 (0.93)	0.1981 (1.26)	0.1730 (1.20)
White	0.0188 (0.97)	0.0749 (0.89)	0.2035 (1.26)	0.1532 (1.16)

this definition, variations about the mean of $EMG\sigma$ are considered as “noise.” This definition was convenient, as knowledge of neither the “true” $EMG\sigma$ value nor the $EMG\sigma$ -force relationship was necessary, and the measure is invariant to signal gain. However, that definition is not as indicative of $EMG\sigma$ estimate performance once additive noise is modeled. In particular, the noise can cause the $EMG\sigma$ estimate to incorrectly coalesce about the wrong mean value. In this case, SNR would measure the variation of the processed signal *plus* noise; and not the desired error with respect to the true (noise-free) $EMG\sigma$ —which is more appropriate for this study.

Thus, root mean square error between the true and estimated $EMG\sigma$ value was used as the error measure. However, the true value is not known when assessing with real EMG data. Thus, we pursued an approach similar to [41]. Our available 50% MVC trials assume that muscle effort—and therefore $EMG\sigma$ —is not changing during the contraction. So, we optionally whitened each EMG epoch, then normalized each 0% and, separately, each 50% MVC epoch to have a standard deviation of one. We treated each 50% MVC epoch as the “true” EMG signal and its 0% MVC epoch from the corresponding electrode as noise. We then multiplied each normalized 50% MVC EMG epoch point-by-point by a ramp (1 s zero, 3 s ramping from 0 to 0.1, 1 s at 0.1). To this signal, we added 0.02 times the respective, normalized 0% MVC epoch. This addition gave a SNR of 5, which is representative of measured EMG [3, 8, 9, 30-34]. We then computed the $EMG\sigma$ estimate using a 200 ms duration centered (non-causal) window, only using RMS processing (since the Gaussian model was a much better fit to our data), with and without RDS processing. The root mean square error between the $EMG\sigma$ estimate and the “true” $EMG\sigma$ (i.e., the ramp pattern) was computed at times 1.0, 1.5, ... 4.0 s across the 512 epochs (64 subjects x 8 electrodes per subject).

Fig. 3 shows summary results. Due to non-normality of the data, we computed paired sign tests (separately for each time) between the root mean square error of all six unique paired combinations of the four factors: unwhitened data, whitened data, without RDS processing, and with RDS processing (Bonferroni corrected). Comparing each method with RDS processing to each method without RDS processing (four comparisons) always resulted in significantly lower errors *with* RDS processing for times ≤ 2.5 s ($p < 10^{-5}$), and no differences for times ≥ 3 s ($p > 0.1$). When unwhitened vs. whitened processors were compared *without* RDS processing (one combination), there were no statistical differences ($p > 0.1$), except at 1.5 s ($p = 10^{-4}$)—likely an anomaly. When unwhitened vs. whitened processors were compared *with* RDS processing (one combination), whitening had lower error for times ≤ 1.5 s ($p < 10^{-5}$), and was not significantly different for times ≥ 3.0 s ($p > 0.1$).

Evaluating Probability of a Zero Value at Rest

The theoretical results predict that the probability of estimating a zero value for $EMG\sigma$ during rest is a function of the window length and the noise gain factor “ g ”. We experimentally evaluated this result using the 512 0% MVC epochs. We again limited analysis to RMS processing. We computed the fraction of zero-valued estimates when using RDS processing for all combinations of: unwhitened vs. whitened processing, window length values ranging from $N=2$ –400 ms, and g values of 0.95, 1, 1.05 and 1.2. The sample variance of each rest epoch was computed (after removing a 400 ms startup transient) and used as the noise variance q^2 to compute its respective RMS estimate of $EMG\sigma$.

With this method, the selected window length is misleading for comparison to the theoretical results shown in Fig. 1, because the experimental EMG signal is correlated (i.e., has finite bandwidth). To resolve this conflict, Bendat and Piersol [4, 51] list the number of effective independent samples for a correlated Gaussian process as: $N_{Eff} = 2B_S T$, where B_S is statistical bandwidth (Hz) and T is the window duration (s). Thus, we used the method of [52] to estimate statistical bandwidth from the PSD estimate of each 0% MVC epoch, separately with and without whitening (Welch method, Hamming window, 50% overlap, 614-length DFT). Without whitening

we found the 0% MVC bandwidth to be $B_{S,Unwhite} = 118 \pm 72 \text{ Hz}$, and with whitening we found the 0% MVC bandwidth to be $B_{S,White} = 329 \pm 157 \text{ Hz}$. Fig. 4 plots the fraction of zero values during rest as a function of N_{Eff} and “g”.

Discussion

Maximum Likelihood Estimates of $EMG\sigma$

There has been debate in the literature as to the best way in which to suppress the influence of additive noise when estimating $EMG\sigma$. While RDS processing has been suggested (as well as other approaches), no model-derived optimal solution has been peer-reviewed published. Herein, we analytically derived, using maximum likelihood estimation, that constant-effort EMG, modeled as either a Gaussian or Laplacian random process, requires RDS processing when additive noise is modeled [equations (7) and (13), respectively, with $g = 1$]. Further, our work shows that when the particular instance of the EMG signal is such that RDS processing would result in a negative value within the square root, then $EMG\sigma$ should be estimated as $EMG\sigma = 0$. While these formulae are derived with constant-effort assumptions, existing EMG processors assume a quasi-stationary EMG signal, even during highly dynamic contractions [30, 53-56]. Thus, a moving average window assumes a constant $EMG\sigma$ within that window, but an $EMG\sigma$ that slowly varies between adjacent windows. Hence, these RDS processing results remain valid.

EMG Probability Density Function

It does not appear that the PDF of rest EMG has previously been reported. We found this PDF to closely match the Gaussian PDF.

But, the literature has variously reported the PDF of active EMG as Gaussian or as more peaked near zero than Gaussian (e.g., Laplacian), mostly in small sample size studies. Roesler [57] (sample

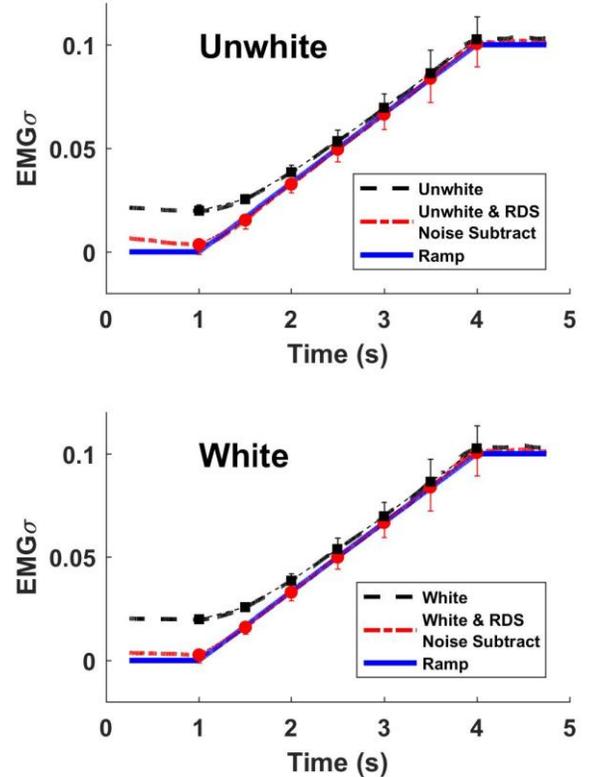


Fig. 3. Top shows ensemble averaged *unwhitened* $EMG\sigma$ estimates along the ramp contraction, with and without RDS processing. Symbols and one-sided error bars show mean and one standard deviation at times 1.0, 1.5, 2.0, ..., 4.0. Bottom shows corresponding results for *whitened* $EMG\sigma$ estimates.

size not listed, perhaps one subject; biceps, triceps and forearm muscles) found the EMG PDF to be precisely Gaussian across a range of isometric contraction levels. Parker *et al.* [35] (sample size not listed, likely one trial reported; intramuscular fine wires within the long head of the biceps brachii) found the EMG PDF to be Gaussian during an ~25% MVC and a just perceptible contraction. Hunter *et al.* [58] (one subject; biceps brachii muscle) found 30% MVC to have a PDF that is more peaked than Gaussian, as did Bilodeau *et al.* [59] for 20% MVCs (16 subjects; biceps brachii and brachioradialis muscles). Nazarpour *et al.* [60] (four subjects; abductor pollicis brevis and flexor carpi radialis muscles) found evidence that the PDF was more peaked (i.e., closer to Laplacian) at low level contractions, but more bell-shaped/Gaussian at higher contraction levels. They postulated that, since more motor unit firings contribute to the EMG during higher contraction levels, the interference signal more closely obeys the central limit theorem—resulting in a more Gaussian shape.

Our own prior work [7] (24 subjects; all distinct from the subjects in the present study) found the PDF from biceps and triceps muscle EMG to be closer to Gaussian than Laplacian, for 10, 25, 50 and 75% constant-force MVCs, using apparatus and methods quite similar to that of the present study. However, this work found that MAV processing produced a higher SNR than RMS processing. A simulation study of constant-effort EMG confirmed that as the EMG PDF is progressively varied from Laplacian to Gaussian, there exists a region wherein the data are more Gaussian in distribution, but MAV processing performs better than RMS.

The present study likely reports the largest sample size to-date. Our EMG exhibited a distribution that closely matched the Gaussian PDF, with a poorer fit to the Laplacian PDF. Since our data were from 50% MVCs (a high contraction level), this result is consistent with the findings of Nazarpour *et al.* [60]. Future comparison to data at lower contraction levels (in which [60] found a more peaked PDF) may be appropriate. The similarity in PDF shapes to our own prior work [7] may be due to the similarity in equipment and use of the identical contraction level. In the end, various factors may influence the EMG PDF, including: electrode shape, size and inter-electrode distance; contraction level; and muscle studied.

EMG σ Estimates

Our root mean square error results from the amplitude-modulated ramp contractions show that noise correction is most important at the lowest contractions levels. RDS processing has the advantage of being progressively less noticeable as effort level increases. For example, once the true EMG σ is four times that of the noise standard deviation, the RDS adjustment is only one sixteenth of the true EMG σ . Once the true EMG σ is five times the noise standard deviation, RDS adjustment is only one 25th the true EMG σ . Etc.

Estimator Performance During Rest

For the ML estimate (c.f., $g = 1$ in Fig. 1 and Fig. 4), we have shown that approximately 50% of EMG σ estimates will be zero, based on either the Gaussian or Laplacian model (excluding unrealistically small N_{Eff} values). Accordingly, nearly half of all EMG σ estimates will be *greater* than zero during rest! In some applications, this result is problematic. For example, the pose of myoelectrically-controlled prostheses, orthoses and exoskeletons would slowly drift at rest, producing an undesired and potentially dangerous action. Thus, we suggest that undesired non-zero EMG σ estimates during rest be eliminated by accentuating the noise standard deviation (i.e., setting $g > 1$). Fig. 1 shows that even modest increases in the gain factor g result in much lower probability of a non-zero value. Indeed, it is common to include threshold subtraction in a prosthesis EMG processor (with zero as the boundary condition), although it is currently applied by subtracting the noise standard deviation from EMG RMS (or MAV) and not via RDS processing [61, 62].

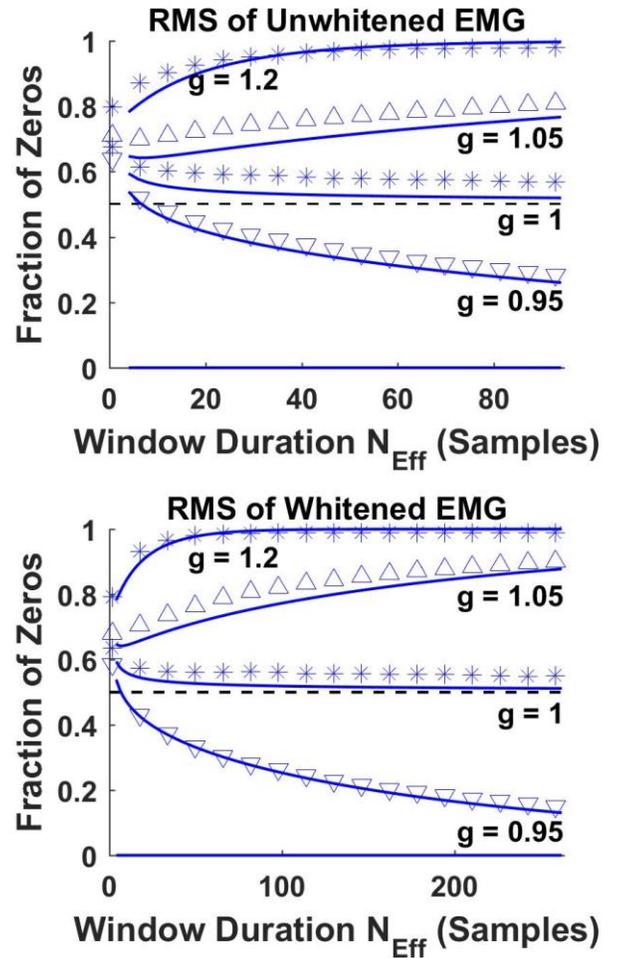


Fig. 4. Symbols show fraction of EMG σ values equal to zero during rest contractions for unwhitened (top) and whitened (bottom) experimental moving average RMS estimates as a function of effective number of samples N_{Eff} , for four different noise gain values " g ". Solid lines show corresponding theoretic probabilities of zero values (same as Fig. 1), for comparison. Dash line show 0.5 probability.

Note that many biomechanics studies in which the subject is active most of the time might *not* want to increase the gain factor “ g ”. Doing so might create a bias in $EMG\sigma$ -force estimates.

Limitations

Our theoretical models assumed independent samples, which are approximated in experimental analysis via whitening. However, since signal and noise have some distinctions in their spectral shape (noise exhibits a lower span of power across frequency [3]), one filter cannot precisely whiten both the noise-free EMG signal and the noise. In particular, whitening filters calibrated to active EMG may contain excessive high frequency gain [45]. Thus, some signal correlation will remain. This dissonance may place practical limits on the bandwidth of whitening filters [50], and might argue for the use of RDS processing in concert with other noise mitigation techniques such as adaptive whitening [3] —in which an adaptive Wiener filter provides lowpass filtering with a progressively lower cutoff at lower $EMG\sigma$ levels.

When evaluating the fraction of zero $EMG\sigma$ values during a rest contraction, we used that same rest contraction to estimate the noise variance (q^2). In practice, q^2 may vary over time; thus, so would the fraction of zero $EMG\sigma$ values during rest. Hence, setting the noise gain factor “ g ” above one might help to mitigate unmeasured changes in q^2 .

Conclusion

Using established stochastic models for EMG in the presence of additive noise, we derived that RDS processing represents the ML estimate of $EMG\sigma$, under both Gaussian and Laplacian PDF assumptions. We concomitantly showed that $EMG\sigma$ should be set to zero whenever RDS processing produces a negative-valued result. Further, we showed that the ML estimate at rest produces zero $EMG\sigma$ estimates only 50% of the time (for all but short-duration smoothing windows). Experimentally, our biceps-triceps EMG data more closely followed a Gaussian PDF than a Laplacian PDF. Our $EMG\sigma$ estimates closely followed theoretical predictions, both during ramp and rest contractions. This work definitively argues that EMG processors should use RDS processing rather than subtracting the noise standard deviation from EMG RMS (or MAV).

References

- [1] E. Kaiser and I. Petersen, "Adaptive Filter for EMG Control Signals," in *The Control of Upper-Extremity Prostheses and Orthoses*, P. Herberts, R. Kadefors, R. Magnusson, and I. Petersen, Eds., Springfield, IL: Charles C. Thomas, 1974, pp. 54–57.
- [2] N. Hogan and R. W. Mann, "Myoelectric signal processing: Optimal estimation applied to electromyography—Part I: Derivation of the optimal myoprocessor," *IEEE Trans. Biomed. Eng.*, vol. 27, no. 7, pp. 382–395, 1980.
- [3] E. A. Clancy and K. A. Farry, "Adaptive whitening of the electromyogram to improve amplitude estimation," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 6, pp. 709–719, 2000.
- [4] N. Hogan and R. W. Mann, "Myoelectric signal processing: Optimal estimation applied to electromyography—Part II: Experimental demonstration of optimal myoprocessor performance," *IEEE Trans. Biomed. Eng.*, vol. 27, no. 7, pp. 396–410, 1980.
- [5] E. A. Clancy and N. Hogan, "Single site electromyograph amplitude estimation," *IEEE Trans. Biomed. Eng.*, vol. 41, no. 2, pp. 159–167, 1994.
- [6] E. A. Clancy and N. Hogan, "Multiple site electromyograph amplitude estimation," *IEEE Trans. Biomed. Eng.*, vol. 42, no. 2, pp. 203–211, 1995.
- [7] E. A. Clancy and N. Hogan, "Probability density of the surface electromyogram and its relation to amplitude detectors," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 6, pp. 730–739, 1999.
- [8] C. Dai, B. Bardizbanian, and E. A. Clancy, "Comparison of constant-posture force-varying EMG-force dynamic models about the elbow," *IEEE Trans. Neural Sys. Rehabil. Eng.*, no. 9, pp. 1529–1538, 2017.
- [9] C. Dai, Z. Zhu, C. Martinez-Luna, T. R. Hunt, T. R. Farrell, and E. A. Clancy, "Two degrees of freedom, dynamic, hand-wrist EMG-force using a minimum number of electrodes," *J. Electromyogr. Kinesiol.*, vol. 47, pp. 10–18, 2019.
- [10] D. Staudenmann, K. Roeleveld, D. F. Stegeman, and J. H. van Dieen, "Methodological aspects of EMG recordings for force estimation—A tutorial and review," *J. Electromyogr. Kinesiol.*, vol. 20, pp. 375–387, 2010.
- [11] J. Hashemi, E. Morin, P. Mousavi, K. Mountjoy, and K. Hashtrudi-Zaad, "EMG-force modeling using parallel cascade identification," *J. Electromyogr. Kinesiol.*, vol. 22, pp. 469–477, 2012.
- [12] J. Hashemi, E. Morin, P. Mousavi, and K. Hashtrudi-Zaad, "Enhanced dynamic EMG-force estimation through calibration and PCI modeling," *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol. 23, no. 1, pp. 41–50, 2015.
- [13] E. P. Doheny, M. M. Lowery, D. P. FitzPatrick, and M. J. O'Malley, "Effect of elbow joint angle on force-EMG relationships in human elbow flexor and extensor muscles," *J. Electromyogr. Kinesiol.*, vol. 18, pp. 760–770, 2008.
- [14] C. Dai, S. Martel, F. Martel, D. Rancourt, and E. A. Clancy, "Single-trial estimation of quasi-static EMG-to-joint-mechanical-impedance relationship over a range of joint torques," *J. Electromyogr. Kinesiol.*, vol. 45, pp. 18–25, 2019.
- [15] C. J. Abul-Haj and N. Hogan, "Functional assessment of control systems for cybernetic elbow prosthesis—Part I: Description of the technique," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 11, pp. 1025–1036, 1990.
- [16] D. Shin, J. Kim, and Y. Koike, "A myokinetic arm model for estimating joint torque and stiffness from EMG signals during maintained posture," *J. Neurophysiol.*, vol. 101, pp. 387–401, 2009.

- [17] T. Kawase, H. Kambara, and Y. Koike, "A power assist device based on joint equilibrium point estimation from EMG signals," *J. Robot. Mechatron.*, vol. 24, no. 1, pp. 205–218, 2012.
- [18] S. Pfeifer, H. Vallery, M. Hardegger, R. Riener, and E. J. Perreault, "Model-based estimation of knee stiffness," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 9, pp. 2604–2612, 2012.
- [19] M. A. Golkar, E. S. Tehrani, and R. E. Kearney, "Linear parametric varying identification of dynamic joint stiffness during time-varying voluntary contractions," *Front. Comput. Neurosci.*, vol. 11, p. 35, 2017.
- [20] D. J. Ostry and A. G. Feldman, "A critical evaluation of the force control hypothesis in motor control," *Exp. Brain Res.*, vol. 153, pp. 275–288, 2003.
- [21] P. Parker, K. Englehart, and B. Hudgins, "Myoelectric signal processing for control of powered limb prostheses," *J. Electromyography. Kinesiol.*, vol. 16, pp. 541–548, 2006.
- [22] D. Farina *et al.*, "The extraction of neural information from the surface EMG for the control of upper-limb prostheses: Emerging avenues and Challenges," *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol. 22, no. 4, pp. 797–809, 2014.
- [23] R. A. Popat *et al.*, "Quantitative assessment of four men using above-elbow prosthetic control," *Arch. Phys. Med. Rehabil.*, vol. 74, pp. 720–729, 1993.
- [24] G. M. Hagg, B. Melin, and R. Kadefors, "Applications in ergonomics," in *Electromyography: Physiology, Engineering, and Noninvasive Applications*, R. Merletti and P. A. Parker, Eds.: IEEE Press/Wiley-Interscience, 2004, pp. 343–363.
- [25] S. Kumar and A. Mital, *Electromyography in Ergonomics*. Briston, PA: Taylor & Francis, 1996.
- [26] C. Disselhorst-Klug, T. Schmitz-Rode, and G. Rau, "Surface electromyography and muscle force: Limits in sEMG-force relationship and new approaches for applications," *Clin. Biomech.*, vol. 24, pp. 225–235, 2009.
- [27] C. A. M. Doorenbosch and J. Harlaar, "A clinically applicable EMG-force model to quantify active stabilization of the knee after a lesion of the anterior cruciate ligament," *Clin. Biomech.*, vol. 18, pp. 142–149, 2003.
- [28] R. Merletti, A. Botter, and U. Barone, "Detection and conditioning of surface EMG signals," in *Surface Electromyography; Physiology, Engineering, and Applications*, R. Merletti and D. Farina, Eds., Hoboken, New Jersey: John Wiley & Sons, Inc., 2016, pp. 54–90.
- [29] E. A. Clancy, E. L. Morin, and R. Merletti, "Sampling, noise-reduction and amplitude estimation issues in surface electromyography," *J. Electromyography. Kinesiol.*, vol. 12, pp. 1–16, 2002.
- [30] E. A. Clancy, "Electromyogram amplitude estimation with adaptive smoothing window length," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 6, pp. 717–729, 1999.
- [31] E. A. Clancy, D. Farina, and R. Merletti, "Cross-comparison of time- and frequency-domain methods for monitoring the myoelectric signal during a cyclic, force-varying, fatiguing hand-grip task," *J. Electromyography. Kinesiol.*, vol. 15, pp. 256–265, 2005.
- [32] P. Liu, L. Liu, F. Martel, D. Rancourt, and E. A. Clancy, "Influence of joint angle on EMG-torque model during constant-posture quasi-constant-torque contractions," *J. Electromyography. Kinesiol.*, vol. 23, pp. 1020–1028, 2013.

- [33] P. Liu, L. Liu, and E. A. Clancy, "Influence of joint angle on EMG-torque model during constant-posture, torque-varying contractions," *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol. 23, no. 6, pp. 1039–1046, 2015.
- [34] E. A. Clancy, C. Martinez-Luna, M. Wartenberg, C. Dai, and T. Farrell, "Two degrees of freedom quasi-static EMG-force at the wrist using a minimum number of electrodes," *J Electromyogr Kinesiol*, vol. 34, pp. 24–36, 2017.
- [35] P. A. Parker, J. A. Stuller, and R. N. Scott, "Signal processing for the multistate myoelectric channel," *Proc. IEEE*, vol. 65, no. 5, pp. 662–674, 1977.
- [36] Y. T. Zhang, P. A. Parker, and R. N. Scott, "Control performance characteristics of myoelectric signal with additive interference," *Med. Biol. Eng. Comput.*, vol. 29, pp. 84–88, 1991.
- [37] P. A. Parker and R. N. Scott, "Myoelectric control of prostheses," *CRC Crit. Reviews Biomed. Eng.*, vol. 13, no. 4, pp. 283–310, 1986.
- [38] P. Bonato, T. D'Alessio, and M. Knaflitz, "A statistical method for the measurement of muscle activation intervals from surface myoelectric signal during gait," *IEEE Trans. Biomed. Eng.*, vol. 45, no. 3, pp. 287–299, 1998.
- [39] R. V. Baratta, M. Solomonow, and B.-H. Zhu, "Methods to reduce the variability of EMG power spectrum estimates," *J. Electromyogr. Kinesiol.*, vol. 8, pp. 279–285, 1998.
- [40] K. Koirala, M. Dasog, P. Liu, and E. A. Clancy, "Using the electromyogram to anticipate torques about the elbow," *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol. 23, no. 3, pp. 396–402, 2015.
- [41] L. Frey Law, C. Krishnan, and K. Avin, "Modeling nonlinear errors in surface electromyography due to baseline noise: A new methodology," *J. Biomech.*, vol. 44, pp. 202–205, 2011.
- [42] R. Ortengren, "Noise and artefacts," in *Electromyography in Ergonomics*, S. Kumar and A. Mital, Eds., London: Taylor & Francis, 1996, pp. 97–107.
- [43] G.-A. Hansson, "Letter to the editor," *J. Biomech.*, vol. 44, p. 1637, 2011.
- [44] L. Frey Law, K. Avin, and C. Krishnan, "Response to letter to the editor," *J. Biomech.*, vol. 44, pp. 1637–1638, 2011.
- [45] E. A. Clancy, "Stochastic Modeling of the Relationship Between the Surface Electromyogram and Muscle Torque," Ph.D Thesis, Massachusetts Institute of Technology, p. 303–352, 449–469, 1991.
- [46] E. A. Clancy, in *Optimal Electromyogram Modeling and Processing During Active Contractions and Rest*: Haopeng Wang M.S. Thesis, Worcester Polytechnic Institute, 2019, pp. 40–58.
- [47] G. B. Thomas Jr., M. D. Weir, J. Hass, and C. Heil, in *Thomas' Calculus* 13th ed., Boston, MA: Pearson, 2014, pp. 185–193.
- [48] A. W. Drake, in *Fundamentals of Applied Probability Theory*, New York: McGraw-Hill Book Company, 1967, pp. 103–107, 273–274.
- [49] P. Prakash, C. A. Salini, J. A. Tranquilli, D. R. Brown, and E. A. Clancy, "Adaptive whitening in electromyogram amplitude estimation for epoch-based applications," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 2, pp. 331–334, 2005.
- [50] M. Dasog, K. Koirala, P. Liu, and E. A. Clancy, "Electromyogram bandwidth requirements when the signal is whitened," *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol. 22, no. 3, pp. 664–670, 2014.

- [51] J. S. Bendat and A. G. Piersol, in *Random Data: Analysis and Measurement Procedures*, New York: John Wiley and Sons, Inc., 1971, pp. 277–281.
- [52] L. Liu, P. Liu, E. A. Clancy, E. Scheme, and K. B. Englehart, "Electromyogram whitening for improved classification accuracy in upper limb prosthesis control," *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol. 21, no. 5, pp. 767–774, 2013.
- [53] R. B. Jerard, T. W. Williams, and C. W. Ohlenbusch, "Practical design of an EMG controlled above elbow prosthesis," in *Proc. 1974 Conf. Eng. Devices Rehabil.*, Boston, MA, 1974, pp. 73–77.
- [54] T. D'Alessio, "Analysis of a digital EMG signal processor in dynamic conditions," *IEEE Trans. Biomed. Eng.*, vol. 32, pp. 78–82, 1985.
- [55] E. Park and S. G. Meek, "Adaptive filter of the electromyographic signal for prosthetic control and force estimation," *IEEE Trans. Biomed. Eng.*, vol. 42, pp. 1048–1052, 1995.
- [56] S. Ranaldi, C. De Marchis, and S. Conforto, "An automatic, adaptive, information-based algorithm for the extraction of the sEMG envelope," *J. Electromyo. Kinesiol.*, vol. 42, pp. 1–9, 2018.
- [57] H. Roesler, "Statistical analysis and evaluation of myoelectric signals for proportional control," in *The Control of Upper-Extremity Prosthesis and Orthoses*, Springfield, IL: Thomas, 1974, pp. 44–53.
- [58] I. W. Hunter, R. E. Kearney, and L. A. Jones, "Estimation of the conduction velocity of muscle action potentials using phase and impulse response function techniques," *Med. Biol. Eng. Comput.*, vol. 25, pp. 121–126, 1987.
- [59] M. Bilodeau, M. Cincera, A. B. Arsenuit, and D. Gravel, "Normality and stationarity of EMG signals of elbow flexor muscles during ramp and step isometric contractions," *J. Electromyo. Kinesiol.*, vol. 7, no. 2, pp. 87–96, 1997.
- [60] K. Nazarpour, A. Al-Timemy, G. Bugmann, and A. Jackson, "A note on the probability distribution function of the surface electromyogram signal," *Brain Res. Bull.*, vol. 90, pp. 88–91, 2013.
- [61] T. W. Williams III, "Practical methods for controlling powered upper-extremity prostheses," *Assist. Technol.*, vol. 2, pp. 3–18, 1990.
- [62] J. M. Hahne, M. A. Schweisfurth, M. Koppe, and D. Farina, "Simultaneous control of multiple functions of bionic hand prostheses: Performance and robustness in end users," *Sci. Robot.*, vol. 3, p. eaat3630, 2018.

B. Application of RDS Processing

The work presented above proves that the maximum likelihood estimate of $EMG\sigma$ includes root difference squared (RDS) processing in the presence of additive noise. To validate the theory presented above in practical applications, six different scenarios were considered to compare the effects of RDS processing with and without whitening. The six scenarios that were considered are:

1. Do not whiten the data, without RDS
2. Do not whiten the data, with RDS
3. Apply the fixed whitening filter, without RDS
4. Apply the fixed whitening filter, with RDS
5. Apply the adaptive whitening filter, without RDS
6. Apply the adaptive whitening filter, with RDS

The work presented in (Wang, Rajotte and Wang) focused on the derivation of the maximum likelihood estimate of $EMG\sigma$ using theoretical models. To study the effects of RDS processing, four distinct data sets were used for each of the 64 subjects (total of 256 data sets): two sets of 0% MVC data and two sets of 50% MVC data.

The raw data are first filtered using the high pass filter shown in Figure 6, then a notch filter is applied to remove power line interference (see Table 1 for locations and bandwidths of the notch filter). The next stage is the whitening stage. There are three possible filters that can be applied at this stage: no filter, an adaptive whitening filter or the universal fixed whitening filter shown in Figure 8. If using the adaptive whitening filter, then two filters are created from the two 50% MVC data, but the filter is applied to the opposite 50% MVC data, respectively. One of the adaptive whitening filters is applied to one of the 0% MVC data and the other whitening filter is applied to the other 0% MVC data. After the whitening stage, the standard deviation of both 0% MVC data is computed as the offset (q) that will be used later in the procedure. After computing the offset, all four data sets are rectified, then smoothed using a window length of 200 ms. To remove the effects of transients, 200 ms of data were removed at the beginning and end of each data set. The final step is the implementation of the RDS processing as:

$$\hat{s}_{MAV} = \sqrt{\max(0, (\sqrt{2}MAV)^2 - q^2)}$$

To cross validate, the offset computed from one 0% MVC was subtracted from the other. Once the optimal estimate was computed for each data set for all subjects, analysis was conducted to compare the effects of RDS processing on low level contractions and the higher, 50% contractions. Figure 39 summarizes the steps.

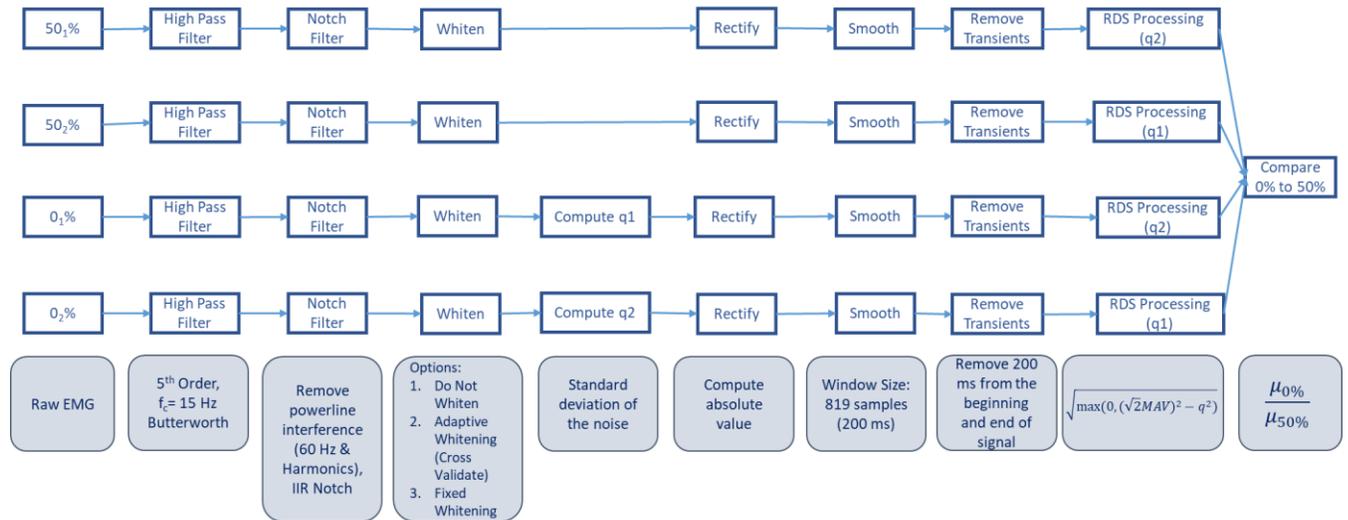


Figure 39 Procedure used to Compare the Effects of Whitening with and without RDS

Once the data were processed using the steps above, the average value of each data set was computed for the six combinations listed above. The average value was computed to compare the magnitude of the data for the different whitening options on the data with and without RDS. To see the impact of whitening on the data, the average value for each data set, with and without RDS was plotted. The following plots show the average value of the 0% MVC data (x-axis) and the average value of the 50% MVC data (y-axis) for the three whitening possibilities.

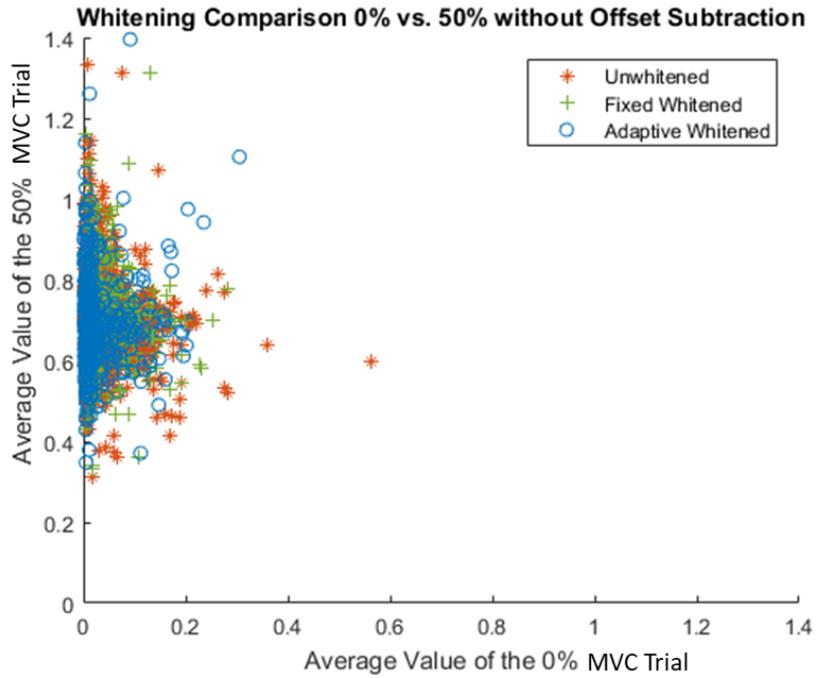


Figure 40 Whitening Comparison without RDS, N = 64 Subjects per Method

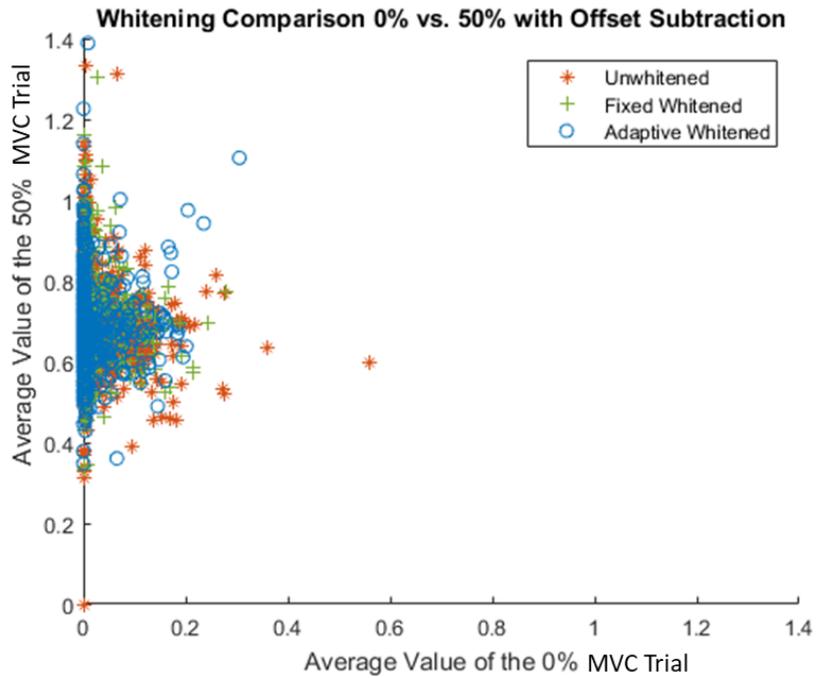


Figure 41 Whitening Comparison with RDS, N = 64 Subjects per Method

Figure 40 and Figure 41 show that the unwhitened 0% MVC data have the greatest magnitude compared to the two types of whitening. The adaptive whitening reduces the 0% MVC signal magnitude on average more than the fixed whitening filter. Both types of whitening reduce the average value of the 0% MVC signal. This behavior is consistent with and without RDS. When RDS is applied, the magnitude of all the data must be less than the data without RDS. For some of the 0% MVCs, they are set to equal 0 in the RDS step, so the points appear on the y-axis.

After observing the behavior of whitening and the RDS subtraction of the 0% MVC and 50% MVC data, the impact of offset subtraction on the 0% and 50% MVC data were studied. The following plots show the data without RDS (x-axis) plotted against the data with RDS (y-axis). The line that runs through each plot, $y = x$, is included as a line of reference. No points can fall above this line, only on or below it.

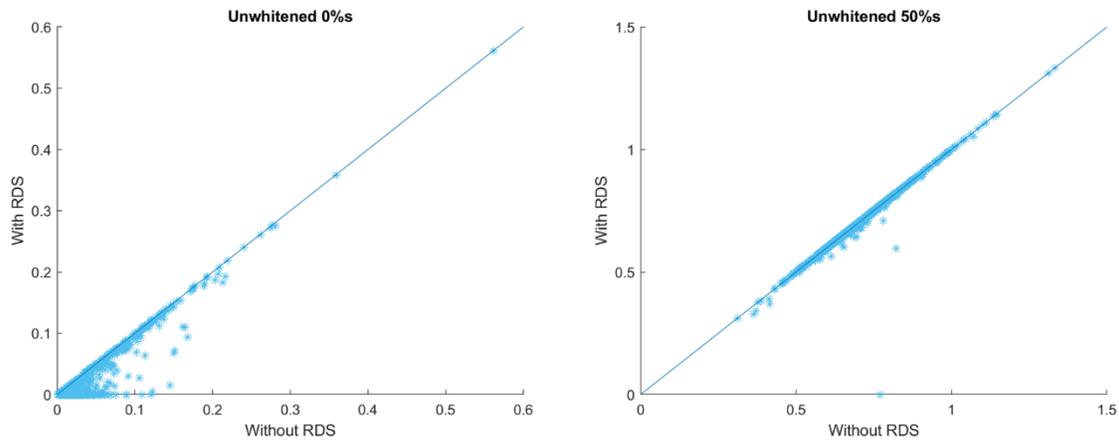


Figure 42 Comparison of Data without RDS to Data with RDS, Unwhitened

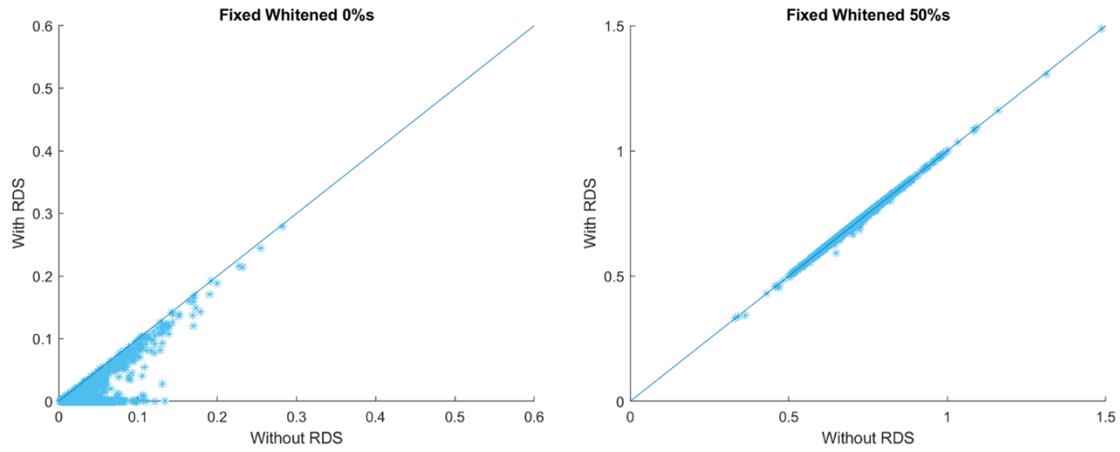


Figure 43 Comparison of Data without RDS to Data with RDS, Fixed Whitened

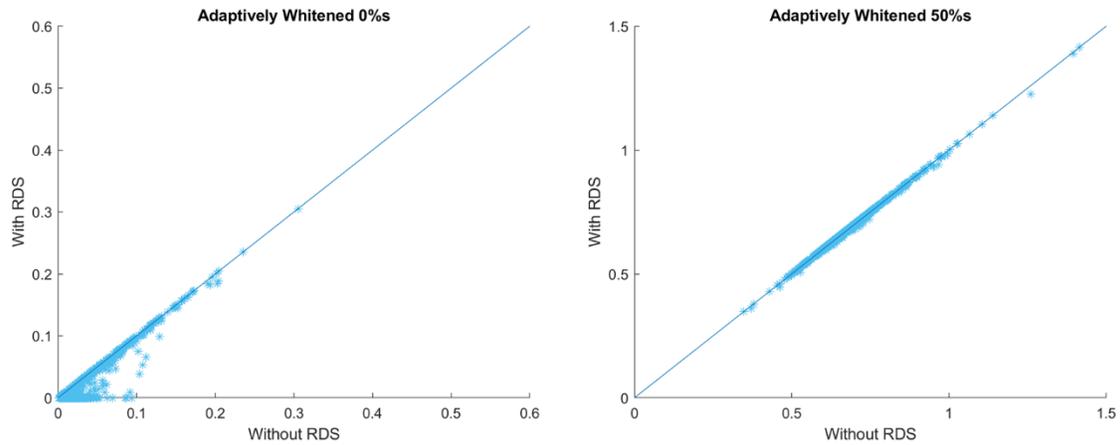


Figure 44 Comparison of Data without RDS to Data with RDS, Adaptive Whitened

Figure 42, Figure 43 and Figure 44 show the impact of RDS on the 0% and 50% MVC data. For the 0% MVC data, all the points fall on or below the line of reference, some even fall on the x-axis. For the 50% MVC data, most of the points remain on the line of reference or slightly below it. These results agreed with the expected outcome of the analysis: RDS would have a greater impact on 0% MVC data than the 50% MVC data.

To compare the effects of RDS with the different whitening options, the average value of each 0% MVC trial was compared to the average value of each 50% MVC trial as a ratio:

$$r = \frac{\mu_{0\%}}{\mu_{50\%}}$$

Ratios were selected to compare the impact of RDS processing because the average value of the 50% MVC trial (denominator) is not expected to change significantly but the average value of

the 0% MVC trial (numerator) is. The ratios of the data with RDS processing must be less than the ratios of the data without RDS processing.

Once the ratios were computed for all 1024 trials, statistical testing was completed to test for significant differences between the six scenarios compared. To test the normality of the ratios ($\frac{\mu_{0\%}}{\mu_{50\%}}$), the Kolmogorov-Smirnov test was used. The Kolmogorov-Smirnov test is a non-parametric test used to determine if data are normally distributed or not. Application of the Kolmogorov-Smirnov test rejected the null hypothesis that the data is normally distributed, so a non-parametric statistical test, not an ANOVA test, must be used to compare the data. (Ghasemi and Zahedisl)

To determine if the ratio have significant differences, the Kruskal-Wallis test was used. The Kruskal-Wallis test is a rank-based, non-parametric test that can be used to compare multiple sets of independent samples. The Kruskal-Wallis test is similar to an ANOVA test but does not assume that the data are normally distributed. It assumes that the samples are random, independent and they share the same distribution. To compare the samples, the Kruskal-Wallis test assigns ranks to the individual samples and compares the ranks rather than directly comparing the raw data. The ranks are assigned by the magnitude of the ratio. For example, the magnitude of the ratios that are equal to zero from the RDS processing are assigned the lowest rank. (Ostertagova, Ostertag and Kovac)

Once the data have been ranked, they are compared to determine if there are significant differences present with a significance level of $p = 0.05$. Bonferroni correction is implemented to reduce the significance level to account for the multiple comparisons being performed. Results of the Kruskal-Wallis test are summarized in Figure 45. In Figure 45, the median value of the ratios is denoted by the square, the 25th percentile of the ratios is denoted by the bar below the square and the 75th percentile is denoted by the bar above the square.

Comparison of the ranked data showed that the data with RDS processing showed significant differences than without RDS processing. The magnitude of the ratios with RDS processing are less than those without RDS processing. When comparing the data with RDS processing, the data whitened with the fixed whitener show significant differences compared to the data whitened with the adaptive whitening filter. The data whitened with the adaptive whitener had a lower ratio than the data whitened with the fixed whitener. When studying the

data without RDS processing, significant differences were seen between the fixed whitened data and the unwhitened and adaptively whitened data.

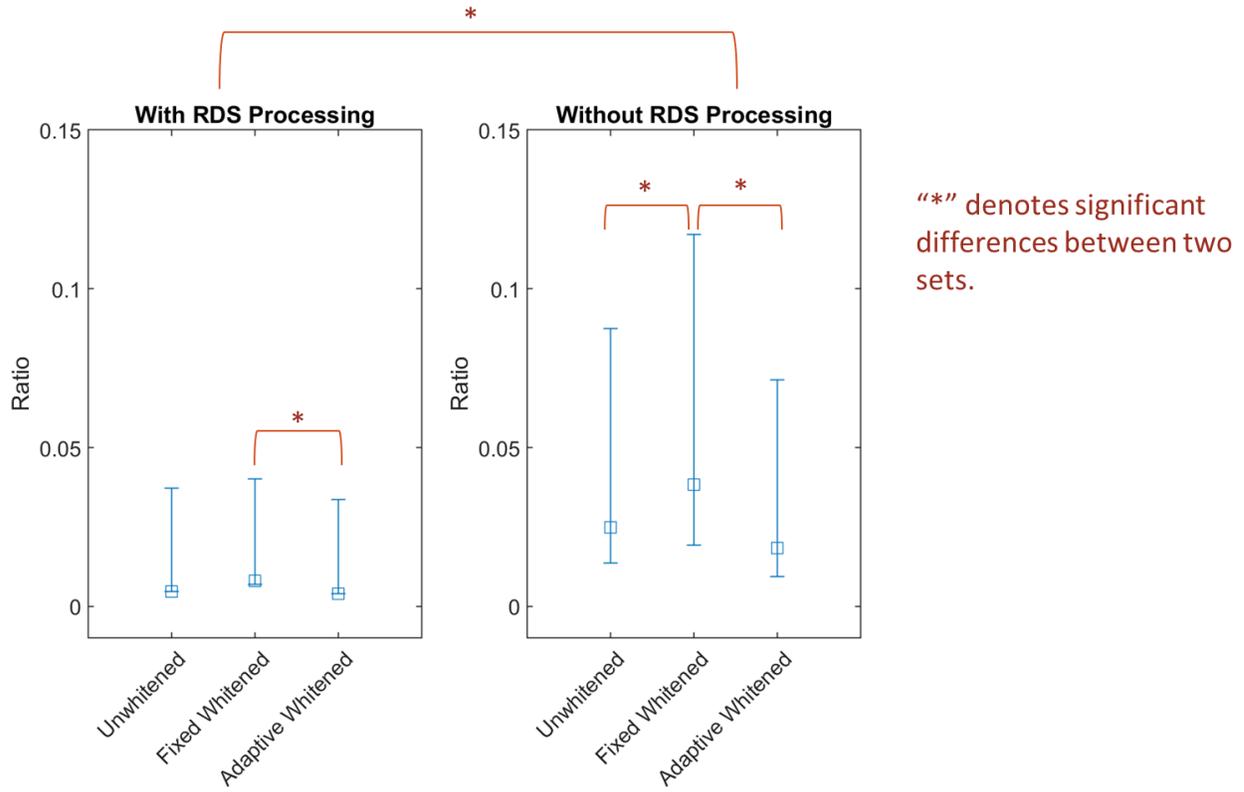


Figure 45 Results of the Kruskal-Wallis Test (Median, 25th Percentile and 75th Percentile)

V. Discussion

A. Discussion of the New EMG Feature Development

The initial proposed approach to completing this project focused on the uncorrelation of the features and then averaging them together to create a new EMG feature that exhibits a decrease in variance on the original EMG feature. The process of removing the correlations between the features involves a linear transformation of the original set of random variables to a new set of random variables. The linear transformation that was used in this work utilized the eigenvalues and eigenvectors of the features. Unfortunately, the new set of random variables showed negative-valued components which is not ideal for the EMG processor.

After the uncorrelation step proved to create negative features, other methods were implemented to combine the four features into a single feature to be used as the EMG feature for the EMG to torque model. The performance of the EMG to torque model, expressed in %MVC, was used to compare the performance of the traditional EMG feature with only $EMG\sigma$ to the performance of the new EMG feature. The performance computed for the different feature combinations (using different methods of combination) never improved on the baseline performance. In the following cases, the performance of the combined pair was equal to the performance of the baseline:

1. Universal Weights, Whitened: Zero Crossings and $EMG\sigma$
2. Universal Weights, Unwhitened: Waveform Length and $EMG\sigma$
3. Averaging, Unwhitened: Waveform Length and $EMG\sigma$, Zero Crossings and $EMG\sigma$

The performance of the unwhitened data in all cases was worse than the performance of the whitened data. The whitened baseline performance of 4.8% was the best performance achieved. The only feature combination that was able to match this performance was the fixed weights method combination of zero crossings and $EMG\sigma$.

After studying the dynamics, FIR filter shape and gains, of the linear least squares model for each subject, it was concluded that the variation between subjects may prevent an implementation of fixed dynamics that result in improved performance with these approaches. Variation between the flexion and extension data may also limit performance improvements. In (Dai, Bardizbanian and Clancy), the only case in which the performance of the four features exceeded the baseline performance was when the EMG estimate was computed using two four-

channel EMG estimates. When the EMG estimate was computed using a single eight channel estimate, the performance of the features was equal to the baseline performance. The two four-channel EMG estimates exhibit worse performance than the single channel estimates, so the 2x4-channel EMG method was not included in this analysis.

Future work can explore other methods of removing correlations from the features so that the uncorrelated features are not negative. The eigenvalue decomposition that was implemented in this work to remove the correlations led to a set of four new features that contained negative components. Principal component analysis and non-negative matrix factorization are other methods that can be used to remove the correlations between the features and may lead to features that do not contain negative components.

Principal component analysis is a data analysis technique used for multivariate data sets to reduce the dimension of the original data set while retaining most of the information content of the individual variables. A linear transformation is applied to the original set of variables to produce a set of uncorrelated variables. The new variables are ordered based on their information content relative to the old set of random variables; the variable with the most information is the first principal component. (Jolliffe)

Non-negative matrix factorization is another data analysis technique that is employed to reduce the dimensionality of the data set. The goal of non-negative matrix factorization is the same as principal component analysis, but it includes the additional constraint of producing a non-negative result. The features are known to be non-negative because of the processing completed prior to the combination and they must remain non-negative after the linear transformation employed to uncorrelated them. Non-negative matrix factorization may lead to uncorrelated features that can be combined to form a new EMG estimate that can be used in the EMG to torque model. (Lee and Seung)

B. Discussion of the Applied Root Difference of Squares

The root difference of squares was seen to have the biggest impact on low effort level contractions, the 0% MVC data used in this analysis, as expected. For the 0% MVC data, more points were set to equal 0 in the estimate than when applied to the 50% MVC. As shown in Figure 42, Figure 43 and Figure 44, the impact of RDS on the 50% data is almost negligible; this is indicated by the majority of the points remaining close to the line of reference.

This analysis was limited to the 0% MVC and 50 %MVC data available. To expand this analysis further, RDS processing can be applied to data of other constant force contraction levels or dynamic data. Application of RDS processing to other types of contractions may provide more insight to the practical applications of the RDS processing. Additional data can be used to determine the point between 0% MVC and 50% MVC where the impact of RDS processing becomes less impactful.

Overall, RDS processing shows the greatest impact at low effort levels and a minimal impact on the higher level, 50% contractions. RDS processing has been modelled to be the optimal maximum likelihood estimate of $EMG\sigma$ and is relatively simple to compute so it is recommended to be included in EMG processors to remove additive noise. At low effort levels, where additive noise will have the biggest impact, RDS processing has the greatest impact.

VI. Fixed Whitening Filter Development

This chapter has been published at the Northeast Bioengineering Conference 2019 as: He Wang, Kiriaki J. Rajotte, Haopeng Wang, Chenyun Dai, Ziling Zhu, Moinuddin Bhuiyan, and Edward A. Clancy, “*Simplified Implementation of Optimized Whitening of the Electromyogram Signal*”.

Introduction: The surface electromyogram (EMG) signal is well modeled as an amplitude modulated, correlated random process. The amplitude modulation, defined as the time-varying standard deviation ($EMG\sigma$) of the signal, is used in various applications as a measure of muscle effort, e.g., EMG-force models, prosthesis control, clinical biomechanics and ergonomics assessment. $EMG\sigma$ can be estimated by rectifying the EMG and then lowpass filtering (cutoff ~ 1 Hz). However, it has long been known that the correlated nature of EMG reduces the statistical efficiency of the $EMG\sigma$ estimate, producing a large variance.

To combat this problem, a whitening filter can be used prior to the rectifier. Whitening removes signal correlation—while preserving signal standard deviation—producing a substantially improved $EMG\sigma$. The advantages of whitening filters have been known since at least 1974 [3]—yet, few researchers use them. A key limitation to widespread use is that most whiteners are “calibrated” to each subject, making them cumbersome to implement.

Since EMG whitening filters have low gain at low frequencies and higher gain at high frequencies, Potvin [4] implemented simple whitening via a fixed, *low-order*, FIR, *highpass* filter that was *not* calibrated to individual subjects. This approach was not compared to the established technique of subject-specific whitening filters.

Our work reported herein describes development of a simplified whitening technique that relies only on EMG magnitude normalization (a measure that is already common). We compare this technique to state-of the art subject-specific whitening.

Experimental Methods: Pre-existing data from 64 subjects [5] were used and did not require human studies supervision per the WPI IRB. Four electrodes over the biceps and four over the triceps muscles were acquired during three trials of 30-s duration, constant-posture, force-varying

elbow contractions in which subjects followed a target displaying a 1 Hz bandlimited, uniform and random process, spanning 50% maximum voluntary contraction (MVC) flexion to 50% MVC extension. Using our existing subject-specific technique to form whitening filters for each electrode (calibrated from additional 5-s rest recordings and constant-effort 50% MVC trials, and limited to 600 Hz in frequency [6,7]), we related $EMG\sigma$ to force. This $EMG\sigma$ -force model used each of the eight $EMG\sigma$ values as inputs, a 15th-order dynamic FIR model per $EMG\sigma$, additionally included the squared value of each $EMG\sigma$ at the 15 time lags (to model the EMG-force non-linearity), and was trained from two trials using least squares. The average \pm std. dev. test error on the distinct third trials was $4.84\pm 1.98\%$ flexion MVC ($\%MVC_F$). This error served as our “baseline” performance.

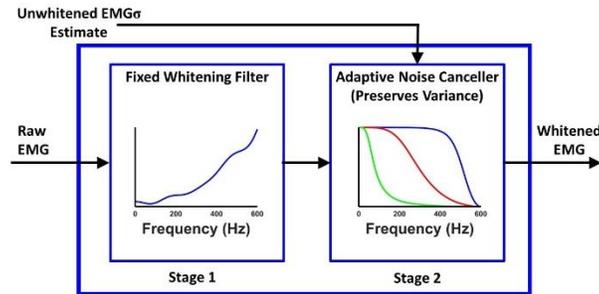


Fig. 1. Two-stage adaptive whitening filter [6].

Analysis Methods and Results: Our whitening filters (Fig. 1) are comprised of a fixed whitening filter followed by an adaptive noise canceller (with variance preservation). The first stage is a fixed linear filter whose magnitude response is the inverse of the square root of the power spectral density (PSD) of the noise-free EMG signal (estimated by subtracting the 0% MVC PSD from the 50% MVC PSD). This filter has low gain at low frequencies and higher gain at high frequencies—the opposite of the spectral content of EMG. The second stage cancels high frequency noise, above the dominant frequency of EMG. This filter is a time-varying lowpass filter, with a cut-off frequency that is lower at lower effort levels. The time adaptation is set via a first-pass unwhitened $EMG\sigma$ estimate. The gain of this stage preserves the overall power of the noise-free signal, so that the full whitening process does not alter $EMG\sigma$.

We contrasted subject-specific whitening filter calibration to “universal” calibration. Each EMG gain was normalized, to account for gain variations between channels. Thereafter, the 0% MVC PSDs and (separately) the 50% MVC PSDs were ensemble-averaged across the 512 calibration recordings (64 subjects x 8 electrodes/subject). The one, ensemble-averaged 0% MVC and the

one, ensemble-averaged 50% MVC were then used to form a single “universal” two-stage whitening filter. This filter was then similarly evaluated on the EMG-force data, producing an average \pm std. dev. test error of 4.80 ± 2.03 %MVC_F—the same as that of subject-specific whiteners.

Conclusions: Our work, combined that of Potvin [4], suggest that the PSD of EMG is sufficiently consistent subject-to-subject that subject-specific calibration of PSDs for EMG whitening may not be necessary (for noise cancellation). Only a gain normalization may be needed per channel. Note that PSD shapes are known to vary with inter-electrode distance [1] and might vary muscle-to-muscle. Also, this set of dynamic contractions may not be particularly sensitive to the magnitude of the noise power, since few of the active-trial contractions were near 0% MVC. (Noise is most impactful at low contraction levels.)

References:

1. Hogan N et al. IEEE TBME. 1980;27:396–410.
2. Clancy EA et al. IEEE TBME. 1995;42:203–211.
3. Kaiser E et al. In “Control of Upper-Extremity Prosthetics & Ortho.,” Charles C. Thomas, 1974:54–57.
4. Potvin et al. J Electromyography Kinesiology. 2004;14:389–399.
5. Dai et al. IEEE TNSRE. 2017;25:1529–1538.
6. Clancy et al. IEEE TBME. 2000;47:709–719.
7. Dasog et al. IEEE TNSRE. 2014;22:664-670.

VII. Conclusion

The goal of the collaborative effort was to develop and test the use of a universal whitening filter from the ensemble average of the individual subject-specific adaptive whitening filters, study the probability density function and power spectrum of rest contractions and develop a new EMG feature that reduces the variance and eliminates the need for calibration contractions. The focus of the work presented in this thesis was the development of a procedure to create a new feature from the combination of four individual features.

Although the combination of the four features did not lead to improvement in the performance of the EMG to torque model, the behavior of the features and their role in the EMG to torque model was better understood. In the EMG to torque model, $EMG\sigma$ and waveform length were observed to have similar responses in the shape of the FIR filter created by their fit coefficients and the same is true for zero crossings and slope sign changes. Although the FIR filter magnitude responses of the feature pairs were relatively consistent, the gains applied to each feature were not. There was significant variation seen in the gains applied to each feature subject-to-subject as well as flexion versus extension. From the work completed in this thesis, it was determined that the combination of the four features using the procedures developed would not eliminate the need for calibration contractions. To account for the subject to subject variation, models that use each feature as a distinct input may lead are likely preferred. Future work may include the exploration of different methods to combine the individual features into a new feature.

The other two goals of the collaborative effort were successful and led to a published conference paper and journal paper, both of which are included above. The universal whitening filter was implemented and reduced the %MVC error of the EMG to torque model from 5.5 %MVC to 4.8 %MVC. The other goal of the collaborative effort involved the modelling of the EMG signals at rest. The modelling of EMG at rest showed that, at rest, EMG more closely resembles a Gaussian distribution than does a 50% MVC contraction. With a better model of the rest contractions, noise rejection of additive noise was proven to be optimized by computing the square root of the variance of the noise subtracted from the squared $EMG\sigma$. From the optimization of the noise rejection of additive noise, analysis was conducted to study the impact of RDS processing on 0% MVC and 50% MVC data. Results of this analysis showed that RDS

processing has the greatest impact on the lower effort level contractions, 0% MVC data, because it sets some values of $EMG\sigma$ to 0. RDS processing does not require significant computation and is recommended to be incorporated into EMG processors to remove additive noise.

VIII. References

- Bhattacharya, Avik, Anasua Sarkarm and Piyali Basak. "Time Domain Multi-Feature Extraction and Classification of Human Hand Movements Using Surface EMG." *2017 International Conference on Advanced Computing and Communication Systems*. Coimbatore: IEEE, 2017. 1-5.
- Clancy, E.A., E.L Morin and R, Merletti. "Sampling, noise-reduction and amplitude estimation issues in surface electromyography." *Journal of Electromyography and Kinesiology* (2002): 1-16.
- Clancy, Edward A. and K. A. Farry. "Adaptive Whitening of the Electromyogram to Improve Amplitude Estimation." *IEEE Transactions on Biomedical Engineering* 47.6 (2000): 709-719.
- Clancy, Edward A. and Neville Hogan. "Multiple Site Electromyograph Amplitude Estimation." *IEEE Transactions on Biomedical Engineering* (1995): 203-211.
- . "Single Site Electromyograph Amplitude Estimate." *IEEE Transactions on Biomedical Engineering* 41.2 (1994): 159-167. PDF.
- Clancy, Edward A. "Electromyogram Amplitude Estimation with Adaptive Smoothing Window Length." *IEEE Transactions on Biomedical Engineering* 46.2 (1999): 717-729.
- Clancy, Edward and Neville Hogan. "Probability Density of the Surface Electromyogram and Its Relation to Amplitude Detectors." *IEEE Transactions on Biomedical Engineering* 46.6 (1999): 730-739.
- Clancy, Edward. "Stochastic Modeling of the Relationship Between the Surface Electromyography and Muscle Torque." *Ph.D. Thesis*. Cambridge: Massachusetts Institute of Technology, 11 January 1991.
- Dai, Chenyun, Berj Bardizbanian and Edward A. Clancy. "Comparison of Constant-Posture Force-Varying EMG-Force Dynamic Models about the Elbow." *IEEE Transactions of Neural Systems and Rehabilitation Engineering* 25.9 (2017): 1529-1538.
- De Luca, Carlo. "Surface Electromyography: Detection and Recording." 2002.
- Fox, Stuart. "Skeletal Muscle." *Human Physiology*. 14th. New York: McGraw-Hill Education, 2016. 360-377.
- Ghasemi, Asghar and Saleh Zahedisl. "Normality Tests for Statistical Analysis: A Guide for Non-Statisticians." *International Journal of Endocrinology and Metabolism* (2012): 486-489. PDF.
- Hogan, Neville and Robert Mann. "Myoelectric Signal Processing: Optimal Estimation Applied to Electromyography - Part I: Derivation of the Optimal Myoprocessor." *IEEE Transactions on Biomedical Engineering* BME-27.No. 7 (1980): 382-395.
- . "Myoelectric Signal Processing: Optimal Estimation Applied to Electromyography - Part II: Experimental Demonstration of Optimal Myoprocessor Performance." *IEEE Transactions on Biomedical Engineering* BME-27.No. 7 (1980): 396-410.
- Hudgins, Bernard and Phillip Parker. "A New Strategy for Multifunctional Myoelectric Control." *IEEE Transactions on Biomedical Engineering* 40.1 (1993): 82-93.
- Jolliffe, I.T. "Principal Component Analysis." New York: Springer, 2002. 1-27.
- Kamavuako, Ernest Nlandu, Erik Justin Scheme and Kevin Brian Englehart. "Determination of Optimum Threshold Values for EMG Time Domain Features; a Multi-dataset Investigation." *Journal of Neural Engineering* (2016): 1-10.
- Lee, Daniel and Sebastian Seung. "Learning the Parts of Objects by Non-Negative Matrix Factorization." *Nature* (1999): 788-791.

- Liu, Lukai, Pu Liu and Edward A. Clancy. "Electromyography Whitening for Improved Classification Accuracy in Upper Limb Prosthesis Control." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 21.5 (2013): 767-774.
- Liu, P., Lukai Lui and Clancy E. A. "Influence of Joint Angle on EMG-Torque Model during Constant-Posture, Torque-Varying Contractions." *IEEE Transactions on Neural Systems Rehabilitation* 23.6 (2015): 1039-1046.
- Merletti, Roberto. "Anatomy and Physiology; Electrophysiology of Muscles and EMG Detection." Electromyography. Merletti, Roberto; Parker, Phillip;. n.d.
- National Cancer Institute. *Structure of Skeletal Muscle*. n.d. <<https://training.seer.cancer.gov/anatomy/muscular/structure.html>>.
- OpenStax. "19.2 Cardiac Muscle and Electrical Activity." University, Rice. *Anatomy and Physiology*. Press Books, 2016. 126. Ebook.
- Ostertagova, Eva, Oskar Ostertag and Jozef. Kovac. "Methodology and Application of the Kruskal-Wallis Test." *Applied Mechanics and Materials*; 611 (2014): 115-120. PDF.
- Penrose, R. "A Generalized Inverse for Matrices." *Cambridge Core* (1954): 406-413.
- Shahar, Doron J. "Minimizing the Variance of a Weighted Average." *Open Journal of Statistics* (2017): 216-224.
- Sornmo, Leif and Pablo Laguna. "The Electromyogram." *Bioelectrical Signal Processing in Cardiac and Neurological Applications*. Boston: Elsevier Academic Press, 2005. 337-360.
- Tipping, Michael and Christopher Bishop. "Probabilistic Principal Component Analysis." *Journal of the Royal Statistical Society* 61.3 (1999): 611-612. <<https://www.jstor.org/stable/2680726>>.
- Wang, Haopeng, et al. "Optimal Estimation of EMG Standard Deviation Requires Noise Subtraction in the Power Domain: Model-Based Derivations and their Implications." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* (2019).