Electromyography (EMG)-Based Feature Selection for Detecting Movement Effort in Human-in-the-Loop Optimization of Lower Limb Exoskeletons

applied sciences

MDPI

Article

Electromyography (EMG)-Based Feature Selection for Detecting Movement Effort in Human-in-the-Loop Optimization of Lower Limb Exoskeletons

Martin Grimmer [1,*], Fabian Just [2,3] and Guoping Zhao [4]

[1] Institute of Sport Science, Institute of Automatic Control and Mechatronics, Technical University Darmstadt, 64289 Darmstadt, Germany
[2] Automation Group, Faculty of Electrical Engineering, Chalmers University of Technology, 41296 Gothenburg, Sweden; fabian.just@uni-ulm.de
[3] Institute of Biomedical Engineering, Faculty of Engineering, Computer Science and Psychology, Ulm University, 89081 Ulm, Germany
[4] School of Mechanical Engineering, Southeast University, Nanjing 211189, China; guopingzhao@seu.edu.cn
* Correspondence: martin.grimmer@tu-darmstadt.de; Tel.: +49-06151-16-24134

Abstract

This study identifies electromyography (EMG) features as an alternative to metabolic cost for distinguishing varying levels of movement effort. Data from two experiments was used to analyze the performance of 50 EMG-based features. The first experiment, the Load experiment, involved participants walking with and without carrying loads of 2, 4, and 8 kg, and the second, the Exo experiment, had participants walking with and without varying levels of hip exoskeleton assistance. In the Load experiment, amplitude-based features generally performed well, with Waveform Length (WL) emerging as the top-performing feature achieving a detection rate of 77% when distinguishing between loaded and unloaded conditions in the most challenging 2 kg condition. In contrast, in the Exo experiment, where both increases and decreases in EMG were observed throughout the stride, it failed and mean-based as well as variance-based features performed best and effectively captured fluctuations in muscle activation with a detection rate of up to 71%. This study underscores the importance of selecting EMG features tailored to specific movement tasks and highlights the potential benefits of noise management strategies to improve detection performance for varying levels of movement effort, providing a foundation for EMG-based human-in-the-loop optimization of lower limb exoskeletons.

Keywords: EMG; feature; human; movement; exoskeleton; optimization; electromyography; human-in-the-loop

Academic Editor: Claudio Belvedere

Received: 16 December 2025
Revised: 18 February 2026
Accepted: 25 February 2026
Published: 27 February 2026

1. Introduction

The human lower limb is crucial to performing various daily movement tasks including locomotion. Aging or illness can impose physical and functional constraints on lower limb mobility [1]. Additionally, repetitive tasks or those involving heavy loads, such as lifting or carrying objects, can strain the lower limbs, and result in fatigue and injuries [2]. To overcome mobility limitations and mitigate muscle fatigue, powered wearable assistive technologies, such as lower limb exoskeletons, have been developed [3,4]. Previous work exploring various exoskeleton hardware and control concepts [4,5] clearly supports personalized assistance strategies in order to maximize benefits for users [6]. For example,

Quinlivan et al. employed a single control strategy for all users, which led to inconsistent metabolic cost reductions, ranging from 15% to 36% during walking [7].

To personalize assistance, human-in-the-loop (HITL) optimization has shown considerable promise [8,9]. HITL optimization relies on both a clear optimization objective, such as minimizing walking effort, and a real-time feedback metric that is related to the objective. Thus far, metabolic cost has served as the predominant feedback metric to reduce movement effort. However, while studies have shown significant reductions in walking effort using metabolic cost, this approach is hindered by prolonged optimization periods and its impracticality for everyday scenarios, primarily due to the necessity of wearing a face mask for data acquisition. About two minutes are used for every iteration for an ankle exoskeleton [9], a hip exosuit [10], single and multi-joint exoskeleton configurations [11], a hip exoskeleton [12] and a prosthetic foot with stiffness adaptation [13] during walking, as well as an ankle exoskeleton in running [14]. Most of these approaches used in addition a first-order dynamical model to estimate the metabolic cost within this time window [15].

Considering these limitations, an alternative feedback metric that provides a measure of effort is necessary. One promising candidate is lower limb muscle activity measured via surface electromyography (EMG). While EMG does not measure effort directly, changes in EMG amplitude have been widely used as a practical surrogate for relative changes in muscular demand during locomotion. Further, EMG signals have been increasingly used as real-time feedback in human-in-the-loop (HIL) optimization of lower-limb exoskeletons [12,16,17]. In such closed-loop systems, EMG signals from key lower limb muscles are continuously measured and processed to quantify muscle effort. These measurements are used to construct a cost function, which is then minimized by an optimization algorithm that iteratively adjusts exoskeleton assistance parameters. In this way, the EMG signals directly inform the control system about the human physiological response, completing a closed loop from human effort to device adjustment.

However, the use of EMG as a feedback signal is not without challenges. Due to slow drift, EMG amplitude may change over time independently of movement effort. The drift can arise from multiple physiological and measurement-related factors, including sensor–skin motion early in an experiment [18], sweat accumulation altering electrode impedance [19,20], and increases in skin and muscle temperature during warmup [21–23]. Fatigue has also been proposed as a contributor, although its influence on EMG amplitude remains inconsistent and task dependent [24–26]. Still, previous research has demonstrated the utility of EMG in HITL optimization and has shown notable benefits and quicker optimization compared to methods that rely on metabolic cost [9–11,13,14]. Specifically, Han et al. [16], Xu et al. [17], and Zhao et al. [12] demonstrated that using EMG as feedback substantially accelerates optimization, enhancing efficiency. With only 25–60 s allocated per exploration phase, the number of testable parameter sets per hour can be up to four times higher compared to approaches relying on metabolic cost as the feedback metric. Additionally, Grimmer et al. [27] anticipate further improvements by leveraging combinations of EMG data from multiple strides and various muscles across both limbs. By averaging over a larger dataset, this approach mitigates one of the common challenges associated with EMG, the signal variability, thus paving the way for even more robust optimization processes. In addition to reducing short-term signal variability, EMG-based approaches can also remain effective in the presence of slow signal drift. In earlier work [27], we showed that although EMG amplitude can exhibit substantial drift over extended walking durations, relative comparisons between temporally adjacent conditions remain robust, enabling identification of changes in movement effort.

In Grimmer et al. [27], the mean average value of rectified EMG, also known as mean absolute value (MAV), served as the primary EMG feature for analysis. Han et al. [16],

Xu et al. [17] and Zhao et al. [12] used the root mean square (RMS). However, it is important to acknowledge the vast array of time-, frequency-, and spatial-domain features, as well as their combinations, that are available for EMG analysis [28–31]. According to a review by Toledo et al. [31], mean absolute value (MAV), slope sign change (SSC), zero-crossing (ZC), and wavelength (WL) emerge as the most prevalent EMG features. These features are typically employed in pattern recognition to classify upper- and lower-limb movements for robotic control. In such applications, the features are not required to scale monotonically with muscle activity or walking effort, as the primary objective is movement classification rather than effort minimization. Consequently, changes in EMG feature amplitude may reflect altered coordination strategies or compensatory movements induced by external assistance, rather than a reduction in metabolic effort. This distinction is critical for human-in-the-loop optimization, where the feedback signal must reliably track effort-related changes.

Despite growing interest in EMG-based HITL optimization, it remains largely unexplored which EMG features are best suited to robustly reflect changes in muscle effort across different walking conditions despite inherent EMG signal variability.

The aim of this study is to systematically evaluate and compare a broad set of commonly used EMG features with respect to their ability to detect effort-related changes during walking, using two complementary experimental paradigms: a controlled load manipulation to elicit graded changes in muscle effort, and an exoskeleton-assisted walking experiment that reflects a realistic human-in-the-loop optimization scenario.

In the first experiment (Load), muscle effort was manipulated by having participants walk under an unloaded condition and three loaded conditions, during which they carried dumbbells totaling 2, 4, and 8 kg [27]. EMG features were evaluated based on their ability to detect changes in walking effort during transitions between loaded and unloaded conditions and to differentiate between the three load levels. Particular emphasis was placed on the most challenging 2 kg condition, which represents the smallest change in external load.

We hypothesize that the effectiveness of EMG-based features in distinguishing varying levels of muscle effort differs substantially across features. This effectiveness is quantified using the detection rate, defined as the percentage of observed changes in EMG that align with the expected direction of effort modulation (e.g., increased effort from unloaded to loaded walking). Furthermore, we propose that a key determinant of feature performance is the noise-to-signal-change ratio, defined as the ratio of intra-participant variability to the absolute magnitude of the effort-related signal change.

The second experiment (Exo) examined walking with a hip exoskeleton [12,32,33], where walking assistance was provided under ten different conditions, each featuring varying assistance timings. Additionally, an unassisted condition was included as a baseline for comparison. This exoskeleton experiment offers a valuable opportunity to validate the results concerning feature performance of the Load experiment. This validation specifically pertains to detection rates of changes in walking effort when comparing conditions with and without assistance, as well as the noise-to-signal ratios in a context that closely resembles real-world applications.

2. Materials and Methods

For the analysis of EMG-based features intended to reflect changes in walking effort, experimental conditions must induce systematic variations in effort. Previous studies have shown that walking effort decreases with reduced body weight [34] and increases with additional carried load [35]. These load-induced changes affect not only whole-body metabolic cost but also the activation of individual lower-limb muscles, which can be

observed in corresponding EMG signals [36]. Consequently, experimental paradigms that manipulate external load are well suited for evaluating EMG-based indicators of walking effort.

Many load-modulation experiments apply additional weight using backpacks or weights attached to the lower limbs. For the present study, however, a setup was required that allowed rapid and repeated changes in load within a single experimental session. Miller and Stamford [37] demonstrated that carrying additional weight in the hands increases oxygen consumption, and thus walking effort, by approximately 1.3% per 100 g added. Based on this finding, a previously published study on EMG-based human-in-the-loop optimization [27] employed an experimental design in which participants walked on a treadmill and, guided by a metronome, repeatedly picked up or released a pair of dumbbells. This dataset is particularly well suited for the present analysis, as it includes three distinct load levels of 2 kg, 4 kg, and 8 kg. This allows assessment of whether EMG features systematically increase with increasing load, which is a key requirement for their use in HITL optimization.

Nevertheless, this load-carrying experiment (Load experiment) does not fully capture the EMG adaptations that occur during walking with a lower-limb exoskeleton. To address this limitation, we additionally included a dataset of walking with a hip exoskeleton (Exo experiment) using different assistance timings, which was originally recorded in preparation for a separate HITL optimization study [12]. In this dataset, both lower-limb EMG and metabolic cost of walking were measured. For our analysis, we selected all assistance conditions that resulted on average in a reduction in metabolic cost, assuming that for most muscles this would be accompanied by a corresponding reduction in EMG activity.

2.1. Participant Information

The Load experiment was originally conducted as part of a previous study [27] and included 13 participants (27 $\pm$ 5 years, 1.82 $\pm$ 0.07 m, 81 $\pm$ 11 kg). The Exo experiment was likewise performed in a prior study [12] and involved 11 participants (21 $\pm$ 1 years, 1.75 $\pm$ 0.08 m, 71.7 $\pm$ 8 kg). Both study protocols were approved by the institutional review board of the Technical University of Darmstadt, Germany. All participants provided written informed consent in accordance with the Declaration of Helsinki and reported no gait-related impairments.

2.2. Experimental Setup

In the Load experiment, EMG data was collected from several lower limb muscles for each leg, including the rectus femoris (RF), gluteus maximus (GM), vastus lateralis (VL), tibialis anterior (TA), gastrocnemius lateralis (GAS), and soleus (SOL). In the Exo experiment, EMG data was recorded from the rectus femoris (RF), gluteus maximus (GM), biceps femoris (BF), gastrocnemius lateralis (GAS), and soleus (SOL). Both studies utilized a wireless EMG system (Trigno Avanti, Delsys, Natick, MA, USA). For the Load experiment, the entire acquisition process (1926 Hz) was conducted using this hardware. However, in the Exo experiment, the EMG system was only used for measurement, with data storage synchronized with the exoskeleton data at 1000 Hz on a real-time off-board computer.

EMG electrode placement followed the recommendations of SENIAM (seniam.org). Before sensor placement, the skin was prepared by removing hair and cleaning the area with alcohol. To minimize the risk of sensor displacement due to movement or sweating, sensors were further secured with adhesive non-woven fabric tape (Rudavlies). Each EMG sensor also included a 3D gyroscope (148 Hz). In the Load experiment, the gyroscope data of the EMG sensors attached to both the muscles and vinyl-coated dumbbells were used to identify individual strides and to differentiate between loaded and unloaded walking,

respectively. Both studies were conducted on an instrumented treadmill (ADAL-WR, HEF Tecmachine, Andrezieux-Boutheon, France).

In the Exo experiment, the hip exoskeleton GuroX (3.1 kg) was controlled in real time using an off-board computer running Matlab Simulink xPC (Matlab R2020a, Mathworks, Natick, MA, USA) at 1000 Hz [12,33]. The exoskeleton's extension and flexion torques were governed by a sinusoidal curve with one full cycle corresponding to a single stride. The peak torque amplitudes were set to 10 Nm for extension and 8 Nm for flexion. The peak flexion torque occurred with a fixed delay of 50% of the stride duration following the peak extension torque. The peak extension torque timing varied from 0% to 45% of the stride time in 5% increments, resulting in a total of 10 timing conditions. During the experiments, the peak extension torque timings were randomly assigned. Heel strikes were detected using a vertical ground reaction force (GRF) threshold, and the timing of two consecutive heel strikes of the same limb was used to calculate stride time. The average stride time from the last five strides combined with the elapsed time since the most recent heel strike was used to assess the current stride progress. This information was then used to regulate the exoskeleton's torque amplitude throughout each stride.

2.3. Experimental Protocol

The Load experiment consisted of four walking sessions at 1.3 m/s, each lasting 14 min. After performing the first session without load for warmup, three loaded sessions were performed in a balanced and randomized order where each session required participants to carry two vinyl-coated dumbbells with a total mass of 2, 4, or 8 kg shared equally across both hands. Within each loaded session and at every 40 s, participants were instructed by an acoustic metronome to either pick up or release the pair of dumbbells thus resulting in 11 intervals without loads and 10 intervals with loads over the course of 14 min.

The Exo experiment began with a three-minute standing session used to determine net metabolic cost. Subsequently, a five-minute walking phase without exoskeleton assistance served as a warmup and familiarization period. Then, a four-minute walking session was conducted to gauge metabolic cost without exoskeleton assistance. Ten four-minute walking sessions were then performed with exoskeleton assistance where peak extension and flexion torque timings were varied in a randomized order. All treadmill walking sessions were performed at 1.25 m/s. For all sessions including standing and warmup, participants were wearing the exoskeleton.

2.4. Data Processing

Data from both experiments were exported to Matlab (Mathworks, Natick, MA, USA) for further processing.

2.4.1. Stride Identification

In the Load experiment, strides were identified using the zero-crossing of the shank's angular velocity from the gyroscope data, which occurs near the heel strike, as described in [27]. For the Exo experiment, strides were also identified based on heel strike timing; however, this was achieved using vertical GRFs recorded by two force plates from the treadmill, each covering one leg (left and right). To reduce noise, a zero-lag second-order Butterworth filter with a cutoff frequency of 40 Hz was applied to the GRF data. Force thresholds were then used to detect heel strikes and identify individual strides for each leg [38].

Stride times were calculated from consecutive heel strike timings from the same limb. In both experiments, for each participant and condition, a total of 40 consecutive strides, equally divided between both limbs, were selected for stride time and EMG analyses.

These 40 strides excluded the ten strides immediately prior to and following any change in condition.

2.4.2. Metabolic Cost of Exoskeleton Study

In the Exo experiment, metabolic power was calculated from oxygen consumption and carbon dioxide production using a modified Brockway equation [39]. Metabolic cost was determined for a standing condition (including three minutes of data), a single unassisted walking condition (four minutes), and all assisted walking conditions (four minutes each) using the dynamic model described in [9]. To obtain the net metabolic cost, the average metabolic power from the standing condition was subtracted from each walking condition. Relative changes in net metabolic power between each assisted condition and the unassisted walking condition were then computed following the approach outlined in [40].

2.4.3. EMG-Basic Processing

To eliminate potential EMG offsets, the mean EMG signal was subtracted from the data of each sensor. Subsequently, the EMG data was bandpass filtered using a zero-lag fourth-order Butterworth filter with a frequency range of 40–450 Hz. The 40 Hz cutoff, which is twice the typical value, proved highly effective in reducing EMG drift during prolonged experimental sessions [27]. For calculating the average EMG across a stride, the EMG data was additionally rectified and filtered using a zero-lag second-order Butterworth filter with a 6 Hz cutoff frequency.

Three strategies were employed to remove strides with EMG measurement artifacts [27]. First, phases with significant and prolonged changes in the EMG signal, such as those caused by a loose sensor, were manually excluded. Second, any distinct EMG peaks with amplitudes exceeding three times the average signal amplitude (based on mean peak height across the session), along with the 1 s periods before and after these peaks, were removed. Third, strides with artifacts that exhibited similar EMG amplitudes to expected signals but differed in duration or frequency were identified using the mean absolute value (MAV) of each stride under the same condition. Strides were excluded if their MAV was more than three times greater or less than 0.4 times the average MAV. In total, 10% of the 17,600 analyzed strides in the Exo experiment (11 participants, 7 assisted and 1 unassisted session, 2 limbs, 5 muscles, 20 strides) and 6% of the 374,400 analyzed strides in the Load experiment (13 participants, 3 loaded and 3 unloaded phases, 2 limbs, 6 muscles, 20 strides, 10 transitions, 2 directions of transitions) were excluded.

2.5. EMG Features

The computed EMG features (Table 1) were primarily based on those provided in [29,30]. Both studies listed EMG features commonly used in pattern recognition and motor control applications, as they capture amplitude, frequency, and time–frequency characteristics of the signal, providing a comprehensive representation of muscle activity. In our analysis, we further organized these features by domain for clarity, including a broad range spanning amplitude-, model-, energy-, complexity-, variation-, mean-, and spectral-based features. Features were selected to provide comprehensive coverage of commonly used EMG representations while remaining applicable to cyclic locomotion tasks and suitable for effort-based human-in-the-loop optimization, where features must reflect changes across the entire stride rather than isolated portions of the movement, ensuring that activation changes induced by the exoskeleton are captured throughout the cycle. Equations and Matlab code can be found in the Supplementary Information.

We excluded features that emphasize only specific portions of the movement cycle because they focus on a limited time window (e.g., 25% to 75% of the movement). Examples of these features includes the Absolute Value of Summation in a given window and its Mean

(ASM) [41] and the Modified Mean Absolute Value (MMAV, MMAV2) [28,42]. Similarly, features like the Enhanced Wavelength (EWL) and Enhanced Mean Absolute Value (EMAV) were excluded as they disproportionately increase the contribution of data from 20% to 80% of the movement cycle to the overall feature value [30]. The Multiple Hamming Window (MHW) and Mean Absolute Value Slope (MAVSLP) features were also omitted due to their bias toward specific times within the movement cycle [28]. While these excluded features may be advantageous for pattern recognition by minimizing data with a high noise-to-signal ratio at the beginning and end of non-cyclic movements, they are problematic in repetitive cyclic movements like walking where phases of high muscular activity can occur throughout the entire movement cycle. In particular, changes in activation due to an exoskeleton that can occur in every phase of a stride will not be reflected.

Table 1. List of EMG-based features including identification number (ID), full name, abbreviation, domain, group, and source. Further details of each feature, including any parameters, the equation (Appendix A) and Matlab code, can be found in the Supplementary Information.

ID	Feature	Abbreviation	Domain	Group	Source
1	Average Amplitude Change	AAC	T	Amplitude	[28]
2	Auto-Regressive Model/Coefficients	AR	T	Model	[43–45]
3	Absolute Value of Summation of Square Root	ASS	T	Energy	[41]
4	Cardinality	CARD	T	Complex	[46]
5	Cepstral Coefficients	CC	T	Model	[28,44,47]
6	Coefficient of Variation	COV	T	Variation	[48]
7	Difference Absolute Mean Value	DAMV	T	Amplitude	[49]
8	Difference Absolute Standard Deviation Value	DASDV	T	Variation	[28,49]
9	Difference Variance Value	DVARV	T	Variation	[50]
10	Energy	EN	T	Energy	[51]
11	Frequency Ratio	FR	F	Spectral	[28,52]
12	Integrated EMG	IEMG	T	Amplitude	[28]
13	Interquartile Range	IQR	T	Variation	[53]
14	Kurtosis	KURT	T	Variation	[54]
15	Log Coefficient of Variation	LCOV	T	Variation	[55]
16	Log Detector	LD	T	Energy	[28]
17	Log Difference Absolute Mean Value	LDAMV	T	Amplitude	[50]
18	Log Difference Absolute Standard Deviation	LDASD	T	Variation	[50]
19	Log Teager Kaiser Energy Operator	LTKEO	T	Amplitude	[55]
20	Mean Absolute Deviation	MAD	T	Mean	[53]
21	Mean Absolute Value	MAV	T	Mean	[28,47,51,56,57]
22	Maximum Amplitude	MAX	T	Amplitude	[57]
23	Median Frequency	MDF	F	Spectral	[28,58]
24	Median Absolute Value	MED	T	Mean	by authors
25	Maximum Fractal Length	MFL	T	Complex	[59,60]
26	Modified Median Frequency	MMDF	F	Spectral	[61]
27	Modified Mean Frequency	MMNF	F	Spectral	[61]
28	Mean Energy	MnE	T	Mean	[62,63]
29	Mean Frequency	MNF	F	Mean	[28,58]
30	Mean Power	MNP	F	Mean	[28]
31	Mean Value of Square Root	MSR	T	Mean	[41]
32	Myopulse Percentage Rate	MYOP	T	Amplitude	[64,65]
33	Peak Frequency	PKF	F	Spectral	[28]
34	Power Spectrum Ratio	PSR	F	Energy	[28,66]
35	Root Mean Square	RMS	T	Mean	[28,49,56,57]
36	Standard Deviation	SD	T	Variation	[51,56,57]
37	Sample Entropy	SE	T	Complex	[67,68]
38	Skewness	SKEW	T	Variation	[69]
39	Spectral Moment 1	SM1	F	Spectral	[28]
40	Spectral Moment 2	SM2	F	Spectral	[28]
41	Spectral Moment 3	SM3	F	Spectral	[28]
42	Slope Sign Change	SSC	T	Complex	[28,51,56,70]
43	Temporal Moment	TM	T	Complex	[71]
44	Total Power	TTP	F	Energy	[28]
45	Threshold Zero Crossing	TZC	T	Complex	[72]
46	Variance	VAR	T	Variation	[28,47,56]
47	V-Order	VO	T	Complex	[28,47]
48	Willison Amplitude	WA	T	Amplitude	[28,47,56]
49	Waveform Length	WL	T	Amplitude	[28,47,51,70]
50	Zero Crossing	ZC	T	Complex	[28,47,51,64,70]

We also excluded features referenced by multiple names where the equation or numeric result was that same as that provided by other features. The Energy (EN) [51] feature was selected over the equal Simple Square Integral (SSI) [28]. Furthermore, based on [28], Total Power (TTP) and the Zero Spectral Moment (SM0) result in a similar outcome, and only TTP was analyzed.

The features listed in Table 1 can be organized into subgroups based on their functional roles and performance in specific tasks. Amplitude-based features (WA, AAC, IEMG, WL, DAMV, LTKEO, MYOP, LDAMV, MAX) capture the overall magnitude or strength of the EMG. Mean-based features (MAV, MAD, MED, MNP, MNF, MnE, MSR, RMS) describe the central tendency of the signal also capturing variable muscle activity. Variation-based features (SD, COV, LCOV, DASDV, DVARV, LDASD, KURT, IQR, SKEW, VAR) quantify changes and variability in muscle activity making them suitable for analyzing dynamic or inconsistent movement patterns. Spectral features (SM1, SM2, SM3, MDF, MMDF, MMNF, PKF, FR) represent higher-order frequency-domain properties of the signal and are valuable for tasks with rhythmic components or complex frequency content. Energy-based features (TTP, LD, EN, PSR, ASS) quantify the overall signal power and are useful for assessing the intensity of muscle activation. Complex features (SE, SSC, ZC, TZC, TM, VO, CARD, MFL) reflect advanced properties of the signal, such as nonlinearity or temporal patterns, which may offer unique insights into specific tasks. Model-based features (AR, CC) represent the underlying dynamics of the signal and may be beneficial in tasks requiring an understanding of signal behavior over time.

Feature processing methods, including adaptations, are detailed in the Supplementary Information along with their corresponding equations (Appendix A). Below, we provide two general guidelines on feature usage.

First, several features require threshold values for their determination. Although some studies offer recommendations for these thresholds, using them is critical, as thresholds depend on individual participants and muscles. EMG amplitudes can vary due to factors like tissue thickness between the muscle and sensor, skin temperature, and moisture levels. Therefore, we opted to use normalized signal thresholds based on each participant's EMG sensor data amplitudes. To achieve this, we calculated the median of the rectified EMG for each stride in both the unloaded condition and the condition without exoskeleton assistance. These medians were then averaged across all strides to obtain a basic threshold. Using this basic threshold, thresholds for individual features were manually defined based on visual inspection. For the features WA and MYOP, the basic threshold was used directly. For Zero Crossing (ZC) and Slope Sign Change (SSC), the basic threshold was divided by 10, and for Cardinality (CARD), it was divided by 100.

Second, unlike all other features that produced reasonable outcomes, Maximum Fractal Length (MFL), Log Difference Absolute Standard Deviation (LDASD), Log Teager-Kaiser Energy Operator (LTKEO), and Skewness (SKEW) showed decreasing feature values when transitioning from unloaded to loaded conditions, which is contrary to our expectations. This led to a decrease in detection rate values as the load increased from 2 kg to 8 kg. To address this, for these features, the detection rate was mirrored around the 50% axis.

2.5.1. EMG Outcome Measures and Averaging

EMG features were calculated for each stride in both experiments. We then extracted feature values from specific sets of strides to compute averages for different conditions: loaded vs. unloaded, and with vs. without exoskeleton assistance. These averages were calculated separately for each leg, muscle, participant, and load or assistance condition.

For the Load experiment, feature values from 20 consecutive strides of each leg were extracted from each interval as outlined in [27]. In the Exo experiment, values for

20 consecutive strides were similarly extracted from the beginning of each of the conditions. To account for adaptation to a new condition, values from the first ten strides of each condition preceding these 20 strides were excluded. The average of each 20-stride set was then calculated.

Data were normalized by scaling each feature average for the unloaded conditions in the Load experiment and for the without exoskeleton assistance condition in the Exo experiment to 100%, and all other conditions were expressed relative to this baseline. The intra-participant standard deviation was then calculated for each normalized 20-stride set. Signal-change was determined by subtracting the respective averages of unloaded from loaded conditions and with-assistance from without-assistance conditions. To avoid infinite values when calculating the intra-participant standard deviation to signal-change ratio, signal-change was set to NaN in cases where it was zero. This occurred in 85 out of 468,000 conditions (13 participants, 3 load sessions, 2 limbs, 6 muscles, 20 transitions, 50 features) in the Load experiment and 4 out of 25,410 conditions (11 participants, 7 assistance sessions, 2 limbs, 5 muscles, 1 comparison, 33 features) in the Exo experiment.

Following this, for the Load experiment, the individual participant averages of the 10 transitions for the standard deviation and signal-change were determined. Combined limb averages for standard deviation and signal-change were then calculated for both experiments. Additionally, the detection rate was calculated to indicate the percentage of feature values that changed in the expected direction across all analyzed transitions. For instance, a detection rate of 80% means that 80% of the analyzed changes aligned with expectations: an increase for unloaded to loaded, a decrease for loaded to unloaded, and a decrease for unassisted to assisted conditions. Equal feature outcomes when comparing conditions were considered failed detections, as they did not align with expectations. Eight cases occurred for the Load and one for the Exo experiment. In total, for the Load experiment, 10 loaded-to-unloaded transitions and 10 unloaded-to-loaded transitions were included for each participant and load. Subsequently, participant averages were determined for the signal-change during loaded-to-unloaded and unloaded-to-loaded transitions; participant averages were also determined for the intra-participant standard deviation at each load for the loaded condition. Finally, group averages for each load condition were determined based on the 13 participants for these variables.

The method for determining variables and their group averages in the Exo experiment was comparable. However, when determining the detection rate, multiple assisted conditions were compared to a single unassisted condition for each participant. Based on the metabolic cost reductions (Figure 1), we selected conditions with peak extension torque timings between 10% and 40% as these seven individual samples presumed to reduce muscle activity. These conditions showed the greatest reductions in metabolic cost and were evenly distributed around the maximum metabolic cost reduction of $-12.5 \pm 4.7\%$, which occurred with a peak extension torque timing of 25% and a peak flexion torque timing of 75%.

The absolute value of the feature signal-change was used to determine the intra-participant standard deviation to absolute signal-change ratio. This approach was chosen because individual strides, muscles and participants could all induce both positive and negative signal-change values, potentially resulting in a near-zero average signal-change and thus extreme ratio values. Since the amplitude of change was our primary criterion, using the absolute value was deemed appropriate.

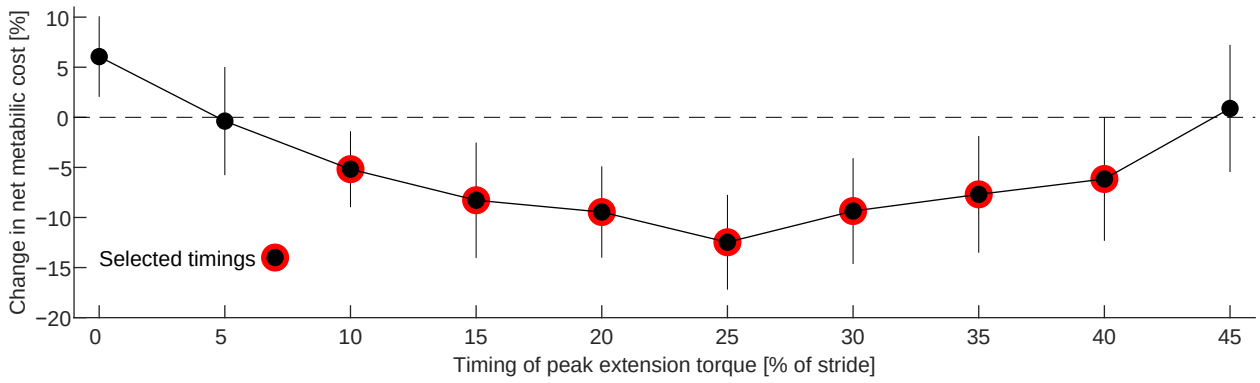

Figure 1. Average change in net metabolic cost of 11 participants for varying the timing of peak hip exoskeleton torque [12] in the Exo experiment. The timings refer to peak extension torque; peak flexion torque occurred at a fixed offset of 50% of the stride. For validating the feature analysis, data from seven timing conditions (red circles) showing decreased net metabolic cost were used, assuming these conditions reduced muscle activity.

2.5.2. Statistics

To compare the performance of individual EMG features, each feature was statistically evaluated against the overall best-performing reference feature at the participant level. In the Load experiment (2 kg condition), Waveform Length (WL) served as the reference feature, whereas in the Exo experiment, Log Difference Absolute Mean Value (LDAMV) was used. For each participant, feature values were first averaged across strides, muscles, limbs, and condition, yielding a single value per feature and participant.

Paired comparisons between each feature and the reference feature were performed using the Wilcoxon signed-rank test, which does not assume normality, to assess whether the distribution of participant-level feature values differed significantly from that of the reference feature.

Because multiple features were tested simultaneously, raw p-values were corrected to account for multiple comparisons within each study. All valid raw p-values for a given experiment were treated as a single family of tests and adjusted using the Benjamini–Hochberg false discovery rate (FDR) procedure. This approach controls the expected proportion of false positives among the tested features within each experiment. Features with FDR-adjusted p-values below 0.05 were considered statistically significant.

3. Results

For the Load experiment participants had similar stride times (about 1.1 ± 0.05 s) for the loaded and unloaded walking conditions when carrying 2, 4 and 8 kg. In contrast, in the Exo experiment, depending on the peak torque timing, the stride times varied between 1.05 ± 0.04 s and 1.11 ± 0.06 s with the without assistance condition lasting 1.07 ± 0.04 s (Table 2).

For the Load experiment, as expected, the normalized EMG of all analyzed muscles increased as the carried load increased (Figure 2). In contrast, in the Exo experiment, the normalized EMG primarily decreased for the GAS and SOL, whereas both increases and decreases were seen for the RF, GM, and BF depending on the peak extension torque timing (Figure 3). The largest EMG fluctuations were observed at the early (0% and 5%) and late (45%) peak extension torque timings. These timings were excluded from further analyses, as they consistently showed either increases or no changes in metabolic cost on average. Consequently, they were unsuitable for evaluating the relationship between EMG reduction and corresponding decreases in metabolic cost.

Table 2. Stride times (±standard deviation) when walking loaded and unloaded for the 2, 4 and 8 kg conditions in the Load experiment, and when walking without and with exoskeleton assistance with different peak extension torque timings (in % of stride time) in the Exo experiment.

Load Experiment											
Condition		Loaded				Unloaded					
		2 kg	4 kg	8 kg		2 kg	4 kg	8 kg			
Stride time [s]		1.1 ± 0.04	1.1 ± 0.05	1.1 ± 0.05		1.09 ± 0.05	1.1 ± 0.05	1.09 ± 0.05			
Exo Experiment											
Condition	without ass.	0%	5%	10%	15%	20%	25%	30%	35%	40%	45%
Stride time [s]	1.07 ± 0.04	1.05 ± 0.04	1.06 ± 0.04	1.06 ± 0.04	1.07 ± 0.04	1.08 ± 0.04	1.09 ± 0.05	1.1 ± 0.04	1.09 ± 0.05	1.11 ± 0.06	1.09 ± 0.05

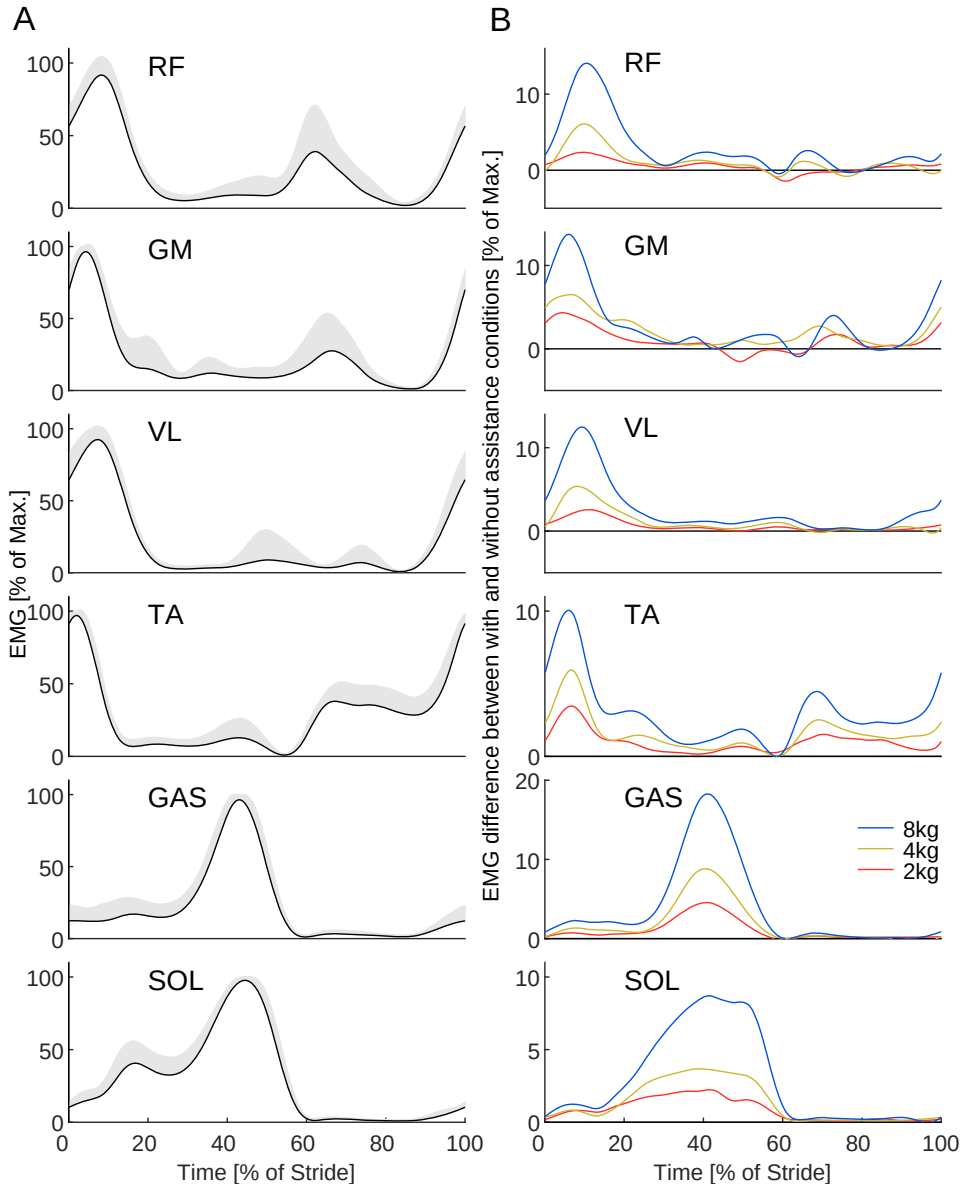

Figure 2. Load experiment: (**A**) Muscle activity for the rectus femoris (RF), gluteus maximus (GM), vastus lateralis (VL), tibialis anterior (TA), gastrocnemius lateralis (GAS), and soleus (SOL) for one stride of walking without load (gray shaded area represents one standard deviation). (**B**) Difference in muscle activity between the loaded condition of 2 kg (red), 4 kg (yellow) and 8 kg (blue).

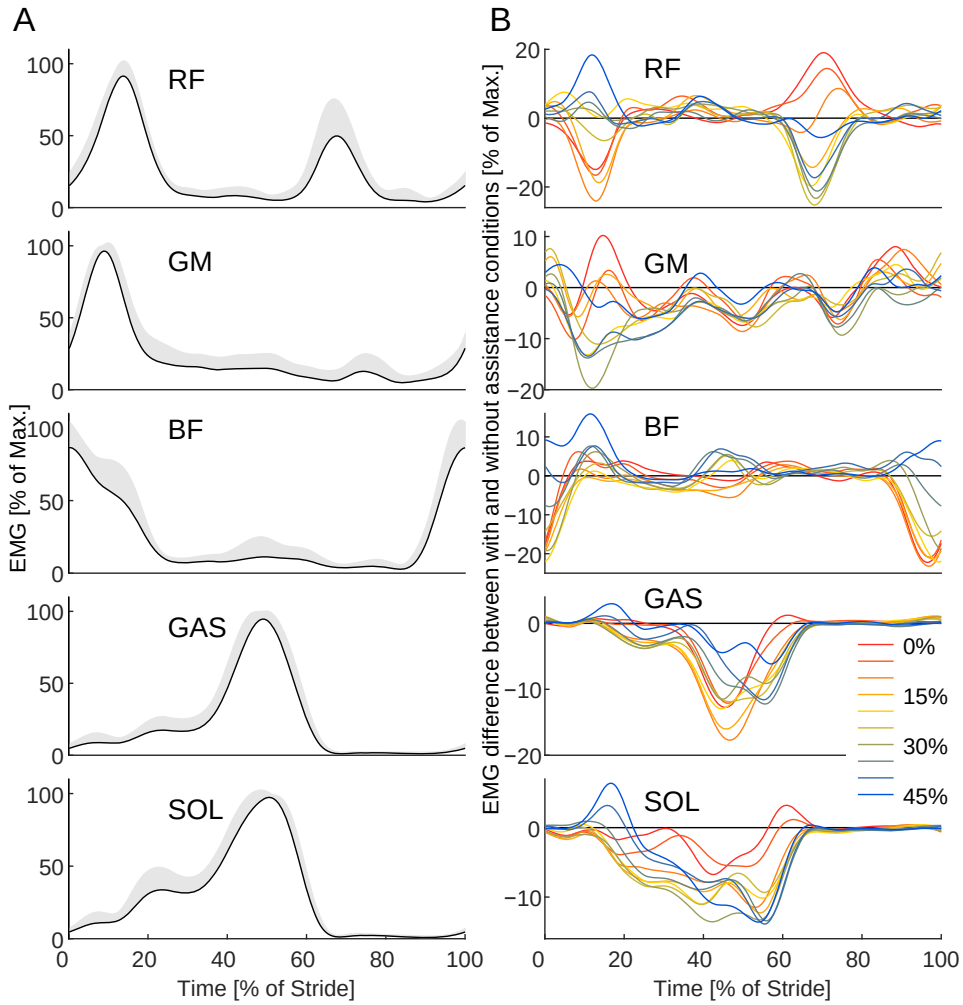

Figure 3. Exo experiment: (**A**) Muscle activity for the rectus femoris (RF), gluteus maximus (GM), biceps femoris (BF), gastrocnemius lateralis (GAS), and soleus (SOL) for on stride of walking without exoskeleton assistance (gray shaded area represents one standard deviation). (**B**) Difference in muscle activity between the condition without exoskeleton assistance and the ten conditions with different peak torque timing with assistance.

Upon analyzing the data from the Load experiment, it became evident that a total of 17 EMG features, representing both time and frequency domains, failed to effectively differentiate between loaded and unloaded conditions (Figure 4). Their detection rates hovered around 50%, akin to random chance. Furthermore, none of these features, nor the signal entropy (SE), exhibited a consistent pattern in detection rates relative to the total carried mass. There was no discernible trend indicating higher detection rates for larger changes in mass as was found for all other features. Therefore, these 17 EMG features were excluded from further analyses.

The greatest detection rates at the most challenging condition for detecting a change between carrying no weight and carrying 2 kg was found for the Waveform Length (WL, 77%). This result was significantly larger compared to all other features (Table 3) but the Absolute Value of Summation of Square Root (ASS), the Sample Entropy (SE) and the Willison Amplitude (WA). On average, for the qualified features, the 2 kg, 4 kg, and 8 kg conditions had detection rates of 70 $\pm$ 4%, 82 $\pm$ 5%, and 92 $\pm$ 5%, respectively.

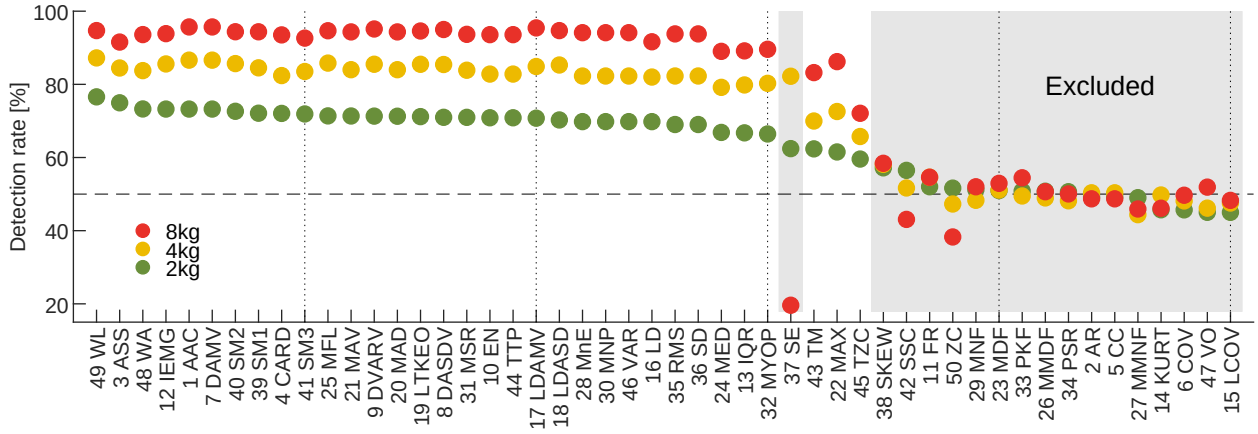

Figure 4. Detection rates for EMG feature changes in the expected direction during transitions between loaded and unloaded conditions in the Load experiment, shown for loads of 2 kg (green), 4 kg (yellow), and 8 kg (red). The data shows the average including all analyzed muscles and participants. A 100% detection rate indicates that all feature changes were identified as expected, while a 50% rate suggests no clear differentiation between loaded and unloaded conditions. Features are ordered by detection rates in the 2 kg condition, which presents the greatest challenge for detecting movement effort changes using EMG. Features highlighted in gray were excluded from further analysis due to their inability to differentiate between load conditions or load levels.

Table 3. FDR-adjusted p-values from paired Wilcoxon signed-rank tests comparing each feature against the reference feature (WL) for the 2 kg condition of the Load experiment. The order is with respect to the order of Figure 4. Significance levels: * $p_{\mathrm{FDR}} < 0.05$, ** $p_{\mathrm{FDR}} < 0.01$.

Feature	p_{FDR}	Feature	p_{FDR}	Feature	p_{FDR}	Feature	p_{FDR}
WL	–	LTKEO	0.002 **	MSR	0.014 *	PKF	0.001 **
ASS	0.455	CARD	0.023 *	LD	0.016 *	MNF	0.001 **
SE	0.945	DASDV	0.012 *	MYOP	0.012 *	MDF	0.001 **
SM1	0.026 *	DVARV	0.010 *	IQR	0.0097 **	AR	0.001 **
SM2	0.025 *	LDASD	0.003 **	MED	0.012 *	CC	0.001 **
IEMG	0.0095 **	MAD	0.012 *	TM	0.005 **	MMNF	0.001 **
WA	0.098	MAV	0.012 *	MAX	0.003 **	MMDF	0.002 **
AAC	0.037 *	MnE	0.0095 **	TZC	0.001 **	COV	0.001 **
DAMV	0.037 *	MNP	0.0095 **	SSC	0.011 *	KURT	0.002 **
SM3	0.02 *	VAR	0.0095 **	SKEW	0.013 *	LCOV	0.001 **
MFL	0.003 **	LDAMV	0.011 *	ZC	0.003 **	VO	0.001 **
EN	0.012 *	RMS	0.0097 **	PSR	0.001 **		
TTP	0.012 *	SD	0.0097 **	FR	0.001 **		

For the Exo experiment, the greatest average detection rates were found for the Log Difference Absolute Mean Value (LDAMV) at 71.1% closely followed by Root Mean Square (RMS), Standard Deviation (SD), Mean Absolute Deviation (MAD) and Mean Absolute Value (MAV) at 70.8%. Also while only including qualified features form the Load experiment in the Exo analyses, several features exhibited detection rates that hovered around random chance (50%) including some of those that performed best during the Load experiment. While the features TM, SM1, LTKEO, IEMG, SM2, MFL, TZC, SM3, WL, ASS, and WA showed significant differences compared to the best performing feature LDAMV before correction, none remained significant after FDR adjustment (Table 4).

The average detection rate of all muscles (MEAN) is defined by the average of individual muscles with the SOL and GAS showing greatest and the RF showing lowest detection rates for the Load and Exo experiments (Figures 5A and 6A).

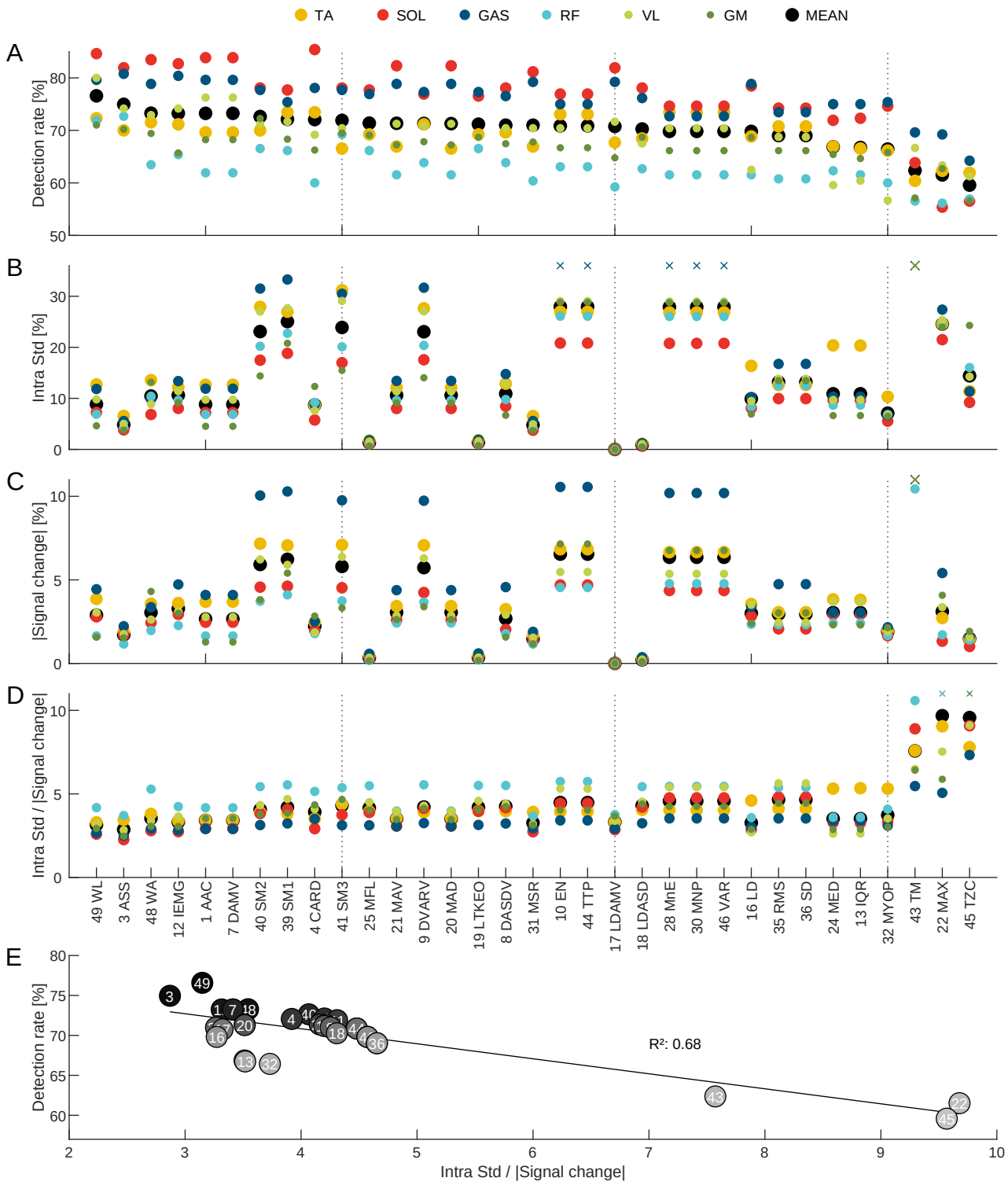

Figure 5. (**A**) Detection rates for EMG feature changes in the expected direction during transitions between loaded and unloaded conditions in the Load experiment for the tibialis (TA), soleus (SOL), gastrocnemius (GAS), rectus femoris (RF), vastus lateralis (VL), gluteus maximus (GM) and the average of these muscles (MEAN). (**B**) Intra-participant standard deviation for individual muscles and their average. (**C**) Absolute signal-changes between conditions with and without loads for individual muscles and their average. (**D**) Noise-to-signal ratio, as represented by intra-participant standard deviation to absolute signal-change, for individual muscles and their average. (**E**) Correlation between the noise-to-signal ratio and detection rate. All subfigures include data from the Load experiment with 2 kg and (**A**–**D**) are ordered based on the detection rate of the MEAN, similar to Figure 4. Data points marked with a cross (×) represent values exceeding the displayed *y*-axis limits.

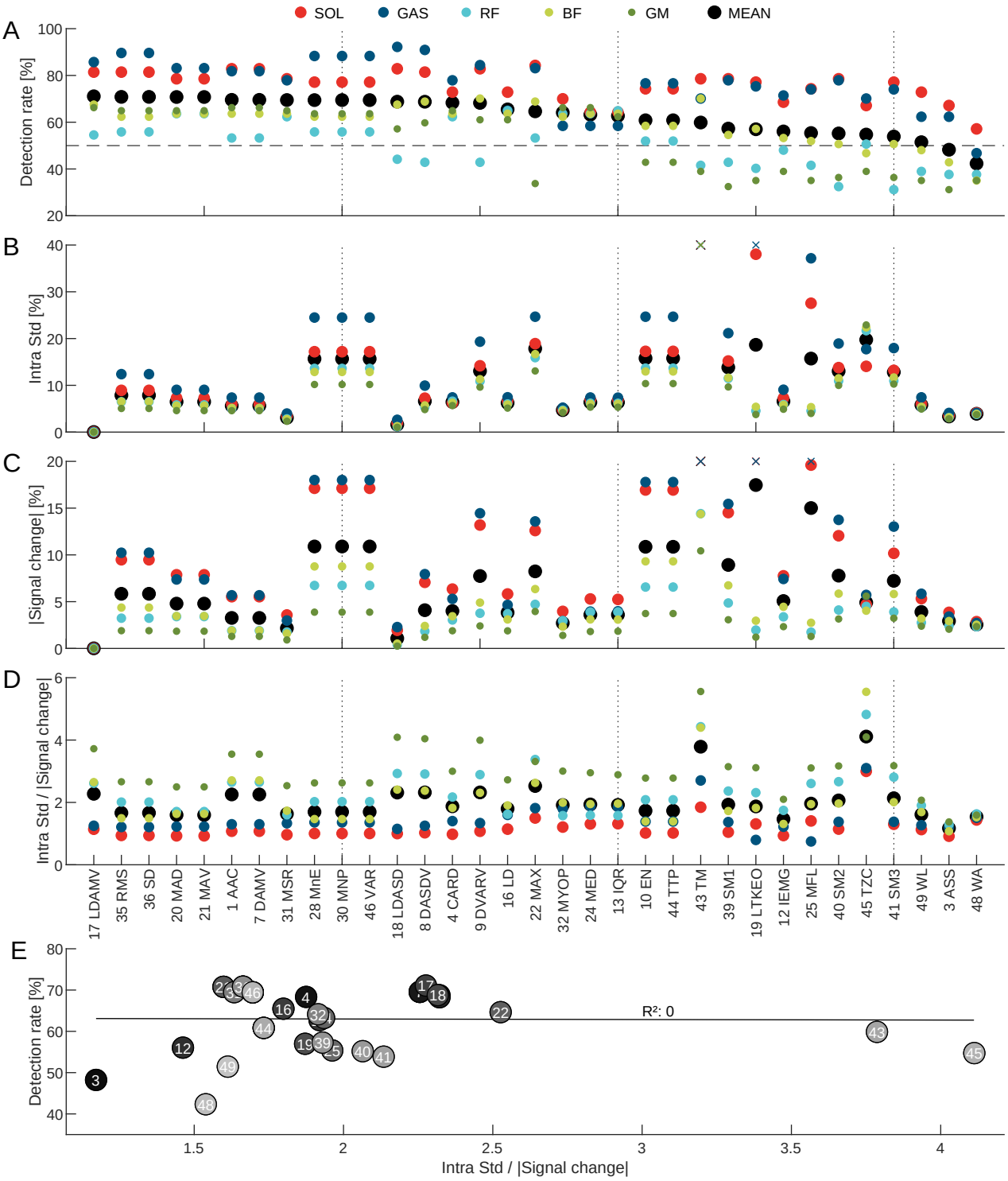

Figure 6. (**A**) Detection rates for EMG feature changes in the expected direction during transitions between loaded and unloaded conditions in the Exo experiment, shown for the soleus (SOL), gastrocnemius (GAS), rectus femoris (RF), biceps femoris (BF), gluteus maximus (GM) and the average of these muscles (MEAN). (**B**) Intra-participant standard deviation for individual muscles and their average. (**C**) Absolute signal-changes between conditions with and without exoskeleton assistance for individual muscles and their average. (**D**) Noise-to-signal ratio, as represented by intra-participant standard deviation and absolute signal-change, for individual muscles and their average. (**E**) Correlation between the noise-to-signal ratio and detection rate. The subfigures (**A**–**D**) are ordered based on the detection rate of the MEAN. Data points marked with a cross (×) represent values exceeding the displayed *y*-axis limits.

Table 4. FDR-adjusted *p*-values from paired Wilcoxon signed-rank tests for the reduced set of 33 features of the Exo experiment when comparing to the best performing feature Log Difference Absolute Mean Value (LDAMV). The order is with respect to the order of Figure 6. While the features TM, SM1, LTEKO, IEMG, SM2, MFL, TZC, SM3, WL, ASS and WA were significant without correction, no features reached significance after FDR correction.

Feature	p_{FDR}	Feature	p_{FDR}	Feature	p_{FDR}	Feature	p_{FDR}
LDAMV	–	DAMV	0.75	MED	0.39	MFL	0.10
MAD	1.0	MSR	0.72	IQR	0.39	TZC	0.10
MAV	1.0	DASDV	0.54	EN	0.15	SM3	0.10
RMS	1.0	LDASD	0.52	TTP	0.15	WL	0.10
SD	1.0	CARD	0.74	TM	0.10	ASS	0.10
MnE	1.0	DVARV	0.61	SM1	0.10	WA	0.10
MNP	1.0	LD	0.33	LTEKO	0.10		
VAR	1.0	MAX	0.15	IEMG	0.10		
AAC	0.75	MYOP	0.61	SM2	0.10		

The intra-participant standard deviation and the absolute signal-change show high variability across features in both experiments. Nonetheless, there exists a strong correlation between the magnitudes of these characteristics. Consequently, when calculating the noise-to-signal-change ratio, the feature averages demonstrate much narrower dispersion, ranging from 2.9 to 9.7 for the Load experiment and 1.3 to 4.1 for the Exo experiment (Figures 5B,C and 6B,C).

In the Load experiment, a strong correlation (R^2 = 0.69) was observed between the detection rate and the noise-to-signal-change ratio. Conversely, no correlation (R^2 = 0.0) was identified between these variables in the Exo experiment (Figures 5D and 6D).

When comparing the detection rates of the Load and Exo experiment, it becomes obvious that there are several features that performed well in one but not the other experiment. In addition, there is a group of fourteen features including CARD (4), DAMV (7), DASDV (8), DVARV (9), LDAMV (17), LDASD (18), MAD (20), MAV (21), MnE (28), MNP (30), MSR (31), RMS (35), SD (36), and VAR (46) that had an average detection rate close or above 70% for both experiments (Figure 7). These include one complex-based, two amplitude-based, five variation-based and six mean-based features.

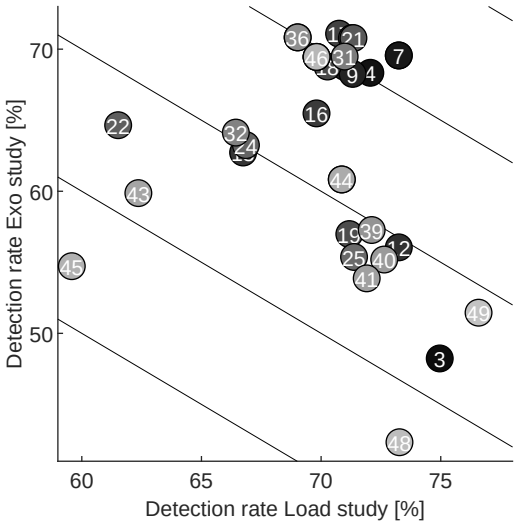

Figure 7. Detection rates for valid features in the Load and Exo experiments. Features in the upper-right corner, with detection rates near or above 70%, represent the most promising candidates. These include CARD (4), DAMV (7), DASDV (8), DVARV (9), LDAMV (17), LDASD (18), MAD (20), MAV (21), MnE (28), MNP (30), MSR (31), RMS (35), SD (36), and VAR (46). Note that RMS (35) is overlapped by SD (36), MAD (20) by MAV (21), and MnE (28) as well as MNP (30) by VAR (46) reflecting their similar performance.

4. Discussion

In this study, data from two experiments was used to identify promising EMG-based features capable of distinguishing varying levels of muscle effort for use in human-in-the-loop optimization. Using the Load experiment, 17 out of 50 analyzed features were unable to differentiate increases in muscle activity associated with increasing load. The remaining features showed a clear trend, with higher detection rates as loads increased. Detection rates for the 2 kg session ranged from 60% to 77%, demonstrating the existence of a significant difference and thus favorable features for this task. We validated these findings with data from the Exo experiment and found that some of the top-performing features in the Load experiment failed in a context more closely resembling real-world exoskeleton applications.

In the Load experiment, Waveform Length (WL) emerged as the best-performing feature with the highest detection rate of 77% for the 2 kg condition. Oskoei et al. were also highlighting the accuracy and stability of WL in EMG analysis [42]. Features such as the Absolute Value of Summation of Square Root (ASS), Willison Amplitude (WA), Integrated EMG (IEMG), and Average Amplitude Change (AAC) also showed high detection rates. These amplitude-based cumulative features (e.g., WL, ASS, WA, IEMG) were most effective due to the consistent increases in muscle activation from carrying loads, which captured gradual changes in muscle effort throughout the stride across all analyzed muscles under the steady, predictable conditions of the Load experiment (Figure 5).

However, in the Exo experiment, these amplitude-based cumulative features like Waveform Length (WL) performed poorly. The dynamic nature of muscle activation in the presence of exoskeleton assistance, where both increases and decreases in EMG were observed depending on assistance timing, challenged these features. Amplitude-based cumulative features such as WL, Absolute Summation of Square Root (ASS), and Willison Amplitude (WA) sum all sample-to-sample changes without accounting for the direction of the fluctuations. Consequently, when muscle activation increases in some phases and decreases in others within a stride, these features accumulate bidirectional fluctuations without reflecting the net muscle effort, effectively "confusing" the signal. For instance, muscles such as the rectus femoris (RF), gluteus maximus (GM), and biceps femoris (BF) showed detection rates well below 50% for WL, indicating that this feature, which performed well in the Load experiment, struggled with the variable, phase-specific activation patterns introduced by the exoskeleton. This suggests that WL and similar features are less suited to scenarios where EMG exhibits phase-dependent bidirectional variations within each stride (Figure 6).

In contrast, mean-based (e.g., MAD, MAV, RMS) and variation-based features (e.g., DASDV, DVARV, LDASD, VAR, SD) performed better during the Exo experiment, effectively capturing the dynamic changes in muscle activation introduced by exoskeleton assistance. These features are more adaptable to the fluctuations in muscle activity and stride-time variability introduced by the exoskeleton's assistance. While amplitude-based cumulative features like Waveform Length (WL) depend on consistent muscle activation, the increased variability in the Exo condition favors features that can capture dynamic changes within a stride. This underscores the importance of context-specific feature selection, especially for tasks involving complex and variable biomechanical interactions, such as those involving exoskeleton assistance (Figure 6).

The best-performing features in the Exo experiment was the Log Difference Absolute Mean Value (LDAMV). Unlike simple amplitude-based cumulative features (e.g., WL), LDAMV and DAMV (both close to 70% detection rate for both studies Figure 7) calculate the mean of absolute differences between consecutive EMG samples, effectively capturing the magnitude of fluctuations without canceling out increases and decreases. This allows

them to reliably reflect dynamic, phase-specific changes in muscle activation within each stride, making them particularly suited to the variable and timing-dependent assistance provided by the exoskeleton.

We acknowledge that no significant differences were observed between the best- and worst-performing features after correcting for multiple comparisons in the Exo experiment. Nevertheless, Figures 5 and 6, as well as Tables 3 and 4, reveal a clear trend: a plateau of high-performing features, including CARD (4), DAMV (7), DASDV (8), DVARV (9), LDAMV (17), LDASD (18), MAD (20), MAV (21), MnE (28), MNP (30), MSR (31), RMS (35), SD (36), and VAR (46) for the combination of Load and Exo experiment. We attribute the absence of significant differences in the Exo experiment primarily to three factors, all of which likely increased data variability. First, there was substantial inter-participant variability in EMG responses to different exoskeleton assistance timings. Second, time-related effects may have contributed, as the reference trial without assistance was performed at the beginning of the study, introducing individual differences in potential EMG signal drift in subsequent conditions as well as individual adaptation or training effects associated with walking with the exoskeleton. Third, the Exo experiment involved fewer comparisons (seven per participant) than the Load experiment (20 per participant), which may have increased the influence of noise and thus limited the ability to detect systematic effects.

The similar performance of several features for the Load and Exo experiment suggests that certain feature types yield comparable outcomes. Therefore, when combining features to differentiate movements or gestures, it may be beneficial to select those that speak to different or distinct EMG qualities to maximize diversity. Toledo et al. showed that it's not necessary to include a large number of features to differentiate movement conditions [31]. We anticipate similar performance outcomes for feature combinations when distinguishing movement effort.

In general, seven out of twelve frequency domain features were dismissed after the initial analysis. Additionally, among the top-ranking features from the Load and Exo datasets, the highest-ranking frequency domain feature ranked only seventh out of all features. This suggests that frequency domain features may be less effective at capturing distinguishing characteristics of movement effort based on EMG data.

So far, RMS has been used in previous EMG-based HITL studies [12,16,17]. Reasons for its selection might include: it provides a direct measure of overall signal amplitude, which correlates with muscle effort; its calculation averages out small fluctuations, offering robustness against random noise and a more stable signal than raw EMG; it is simple to compute and interpret; and its longstanding use in biomechanics and prosthetics research makes it a convenient choice. Our results confirm that RMS seems to be one of the most effective: it ranks among the top 14 features when combining results of the Load and Exo experiments (Figure 7), reliably capturing variations in muscle activation across different conditions.

Our feature analysis highlights that there is no universally best EMG feature, as their effectiveness depends on the task context and how muscle activity changes during the movement. Therefore, feature selection must be carefully adapted to the specific demands of each task. Based on our results, for the application of human-in-the-loop optimization of hip exoskeleton assistance during walking we would currently recommend to explore the use of the top performing features of the Exo experiment that are primarily mean-based and variation-based features (e.g., LDAMV, RMS, SD, MAD, MAV). For unknown conditions with respect to changes in muscle activity, most robust features of both studies could be selected based on Figure 7. The highest detection rates were consistently observed for the gastrocnemius (GAS) and soleus (SOL) muscles across both experiments, suggesting that changes in the movement tasks had a significant impact on the calf muscles. This aligns

with their known biomechanical function during walking: the SOL and GAS are primary contributors to push-off, generating the ankle plantarflexion work that drives forward movement [73]. Consequently, changes in activation due to load carrying or exoskeleton assistance directly reflect variations in mechanical effort, making EMG from these muscles particularly informative. Given that the noise-to-signal-change ratio was also lowest for these muscles, it may be especially beneficial to focus on SOL and GAS within applications of human-in-the-loop optimization with a hip exoskeleton.

We hypothesized that one reason contributing to a high detection rate is the noise-to-signal-change ratio. This assumption was validated for the Load experiment data. However, this was not the case for the Exo experiment despite many of the best-performing features and the muscles with the highest detection rates having the lowest noise-to-signal-change ratios. Several factors could explain this discrepancy. First, features such as Waveform Length (WL), Absolute Value of Summation of Square Root (ASS), and Willison Amplitude (WA), which performed well in the Load experiment, still exhibited low noise-to-signal-change ratios but had very low detection rates in the Exo experiment. This indicates that their inability to detect changes in movement effort for the Exo experiment disrupts the expected relationship. Second, in the Exo experiment only seven timing conditions that resulted in metabolic reductions were included. Consequently, the effective sample size was approximately one third of that in the Load experiment, where 20 transitions per load condition were analyzed for each participant. A larger sample size helps mitigate the influence of signal noise and variability in EMG measurements [27]. Third, the reference condition in the Exo experiment (without assistance) was measured only once at the beginning of the experiment, whereas in the Load experiment (unloaded), it was recorded immediately before and after each 40 s load interval. This reduces the impact of signal drift in the Load experiment, a known issue with EMG data [27]. Finally, although the selected Exo conditions resulted in an average reduction in metabolic cost across participants, individual responses likely varied across muscles and assistance timings. Such participant- and muscle-specific adaptations may further reduce detection rates and obscure the expected relationship between noise-to-signal-change ratio and feature performance. In summary, our results suggest that the noise-to-signal-change ratio is an important factor in determining the effectiveness of an EMG feature for detecting changes in walking effort. However, the nature of individual features also plays a crucial role in regards to handling varying EMG signal qualities, such as during phases of increasing and decreasing activity within a stride.

4.1. Methodological Considerations

In this study, we analyzed a selection of 50 EMG-related features, though we did not conduct a systematic review to identify all potential features used in EMG analysis. It's possible that other features could offer superior performance. Additionally, combining multiple features, exploring alternative feature options, or applying different feature processing methods could improve detection accuracy. For instance, as suggested by [50], exploring the derivative from the EMG rather than the EMG directly may enhance performance. Using rectified EMG could also yield benefits for certain features. However, for individual features, we do not anticipate performance outcomes that significantly exceed the range of those already analyzed.

To better understand the limitations and implications of the selected features, it is important to consider how preprocessing choices influence their behavior. The preprocessing applied to the EMG data, particularly the mean subtraction for offset removal, caused certain features to become numerically similar. For instance, Mean Absolute Value (MAV) and Mean Absolute Deviation (MAD), as well as Average Amplitude Change (AAC) and Difference Absolute Mean Value (DAMV), yielded the same results since these metrics

measure similar properties once the signal mean has been removed. Likewise, Variance (VAR) and Mean Energy (MnE) became nearly identical as subtracting the mean makes VAR effectively measure the average squared signal values, which aligns it with MnE when data is normalized.

These similarities also extend to frequency-domain features, as highlighted by Parseval's theorem, which states that a signal's total energy in the time domain equals its total power in the frequency domain. This theoretical foundation explains why energy- and power-related features, such as Energy (EN) and Total Power (TTP), as well as Mean Power (MNP) and MnE, produced comparable values, since they assess the same signal strength [74]. Recognizing these overlaps is essential for interpreting the results of feature-based analyses and for identifying opportunities to refine feature selection.

In addition, readers should note that while muscle activity generally increased with load in the Load experiment and decreased with exoskeleton assistance in the Exo experiment, this trend was not consistent across all analyzed conditions for every participant. Therefore, detection rates less than 100% are not unusual.

Further, in both experiments the stride times of the reference and the comparison conditions was relatively close. We can imagine that large variation in stride times could benefit features that accumulate signal amplitude over time, such as IEMG, WL, EN, ASS, or WA, since these features scale directly with stride duration.

Finally, as mentioned in the Introduction, EMG signals are inherently sensitive to factors such as electrode movement, perspiration, muscle fatigue, and changes in skin impedance, which may affect feature magnitudes during extended HITL sessions. While our analysis does not explicitly quantify the resilience of individual features to these effects, we mitigated inter-individual and sensor-related variability by avoiding fixed, absolute amplitude thresholds.

Instead, all features requiring thresholds were normalized using participant- and muscle-specific EMG amplitudes derived from reference walking conditions. This normalization strategy reduces sensitivity to absolute mV-level differences across users and sensors, supporting generalization without requiring extensive per-user calibration. Nevertheless, although this approach improves robustness across participants under comparable walking conditions, feature sensitivity to larger changes in gait dynamics, such as walking speed, slope, load magnitude, or assistance level, was not systematically evaluated and remains an important topic for future experimental validation.

4.2. Future Directions and Experimental Validation

While this study identified EMG features with high detection rates for varying muscle effort, these findings are based on retrospective analysis of two experimental datasets. The effectiveness of the identified features for human-in-the-loop (HITL) optimization of hip exoskeleton assistance still needs to be verified experimentally. Future studies should implement the top-performing features in real-time HITL algorithms to determine whether they improve optimization outcomes, including reducing metabolic cost or achieving faster adaptation.

Additional experiments are necessary to evaluate whether these features generalize to other exoskeleton configurations, assistance patterns, or movement tasks, including upper-limb applications or more variable walking conditions, and to systematically assess their robustness to common EMG signal perturbations such as electrode shifts, sweat accumulation, fatigue, and changes in skin impedance, ensuring reliable performance during extended HITL optimization sessions. Such studies could further investigate how EMG features respond to compensatory strategies or co-contraction induced by exoskeleton assistance, for example by combining EMG with joint-level biomechanics or antagonist

muscle analysis. Prospective testing would also allow direct comparison between features under controlled HITL conditions, providing empirical evidence of feature performance and relative effectiveness, which cannot be concluded from the current observational dataset. Identifying muscle synergies can help reduce the number of EMG sensors required. Conversely, combining signals from multiple synergistic muscles may mitigate noise and enhance signal quality, thereby improving the robustness of HITL optimization. While this study focused on healthy participants, the translation of these features to clinical populations with altered neuromuscular control (e.g., stroke, spinal cord injury, cerebral palsy) remains uncertain. Altered muscle recruitment and compensatory strategies may affect feature performance, making the selection of both muscles and features critical for reliably capturing changes in walking effort. Future work should also evaluate the consistency of these features across multiple sessions and days to ensure their reliability for long-term, real-world exoskeleton use.

While EMG-based features provide a direct and sensitive measure of muscle activation, relying on EMG alone as an optimization objective may not fully capture higher-level aspects of locomotor quality. In particular, minimizing muscle activation does not necessarily guarantee preservation of natural gait patterns, balance, or stability, as reductions in EMG amplitude may also arise from altered coordination strategies or redistribution of effort across muscles. Consequently, future HITL implementations may benefit from combining EMG-based objectives with complementary biomechanical or gait-level metrics, such as kinematic symmetry, joint moments, or stability-related measures, to ensure that reductions in muscle effort do not come at the expense of functional or natural locomotion.

From a real-time implementation perspective, computational efficiency remains a relevant consideration. In an initial estimation using a laptop running at up to 2.4 GHz, feature computation times for a single stride were determined in Matlab. RMS was used as a reference feature, as previous studies have demonstrated its effectiveness for EMG-based HITL optimization [12,16,17]. In our implementation, it achieved a computation time of 0.0021 s. When RMS was set to 100%, eight features required 200–400% (e.g., AAC, FR, LCOV, MSR, MED, TTP), seven features required 400–800% (e.g., COV, IQR, KURT, MMDF, SKEW, SM1, SSC), and four features (AR, MDF, MNF, SKEW) exceeded 800% of RMS computation time, with MDF being the most computationally expensive at 0.074 s (approximately 1500%). All other 32 features required less than 200% of the the time. With stride times lasting about 1.1 s, even the more computationally demanding features remain feasible when considering that EMG-based feedback is evaluated over multiple strides to reduce the effects of noise.

Importantly, recent advances in embedded hardware architectures and algorithm deployment techniques substantially reduce traditional constraints related to memory footprint, inference time, and power consumption. For example, Just et al. (2024) [75] demonstrated that neural network quantization strategies combined with modern microcontroller architectures can reduce flash memory requirements by more than an order of magnitude while simultaneously improving inference speed and energy efficiency in self-contained prosthetic systems. Their comparison of state-of-the-art microcontrollers further showed substantial gains in computational performance relative to previous generations, enabling real-time motor intent decoding on resource-constrained wearable hardware. These findings indicate that embedded platforms are rapidly approaching the computational capabilities required for increasingly sophisticated signal processing and machine learning pipelines. Beyond computational complexity, EMG signal artifacts (e.g., caused by sensor movement or transient loss) may pose additional challenges for real-time implementation. Future work could focus on improving signal robustness and developing methods to automatically detect and exclude corrupted data.

Clarifying these limitations ensures that readers understand the scope of the findings. The results presented here provide a foundation for EMG-based feature selection in HITL optimization, but the ultimate performance and comparative efficacy of these features must be demonstrated in future experimental studies.

5. Conclusions

This study utilized data from two experiments to identify promising EMG-based features that are capable of distinguishing different levels of muscle effort during walking. We found that the effectiveness of EMG features depends heavily on the context in which they are applied. While amplitude-based cumulative features performed well under conditions with a continuous increase in muscle load, mean-based and variance-based features performed better with variable dynamic muscle activation introduced by exoskeleton assistance. Therefore, feature selection to detect human effort based on EMG should be tailored to the specific context. While the feature characteristics have a major impact on detection performance, noise management strategies, such as increasing sample size and addressing signal drift, should be incorporated to improve detection quality. Although we analyzed the features in the context of human walking, further investigation is needed to determine whether the best-performing features also excel in other movement tasks and in the upper extremities. We observed that several EMG-based features were unable to differentiate between levels of movement effort and that some features yielded similar results due to their comparable numeric properties. These findings could benefit areas such as movement and gesture recognition. While this study specifically addressed lower-limb exoskeleton assistance, the insights gained here have the potential to enhance the design and control of a broad range of wearable robotics, advancing their usability and effectiveness.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/app16052325/s1, EMG Feature Extraction Toolbox (Extended Version) as Matlab (Natick, MA, USA) code. The Matlab code includes scripts to determine 58 different features based on EMG signals. In addition, it includes an example script (Example_Calculate_Features.m) to run these scripts and determine all features based on EMG from example data of one stride of walking.

Author Contributions: Conceptualization, M.G. and F.J.; methodology, M.G., F.J. and G.Z.; software, M.G., F.J. and G.Z.; validation, M.G. and F.J.; formal analysis, M.G.; investigation, M.G.; resources, M.G.; data curation, M.G. and G.Z.; writing—original draft preparation, M.G.; writing—review and editing, M.G., F.J. and G.Z.; visualization, M.G.; supervision, M.G. and G.Z.; project administration, M.G.; funding acquisition, M.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Technical University of Darmstadt (no. 52700967) and the German Science Foundation (DFG, no. 450821862). G.Z. was supported by the Start-up Research Fund of Southeast University (RF1028624035) and Natural Science Foundation of Jiangsu Province (BK20241313).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of the TU Darmstadt (EK05/2019, 29 March 2019 and EK44/2019, 26 December 2019).

Informed Consent Statement: Informed consent was obtained from all participants involved in the study.

Data Availability Statement: The datasets generated during and/or analyzed during the current study are available for download at the data repository of the TU Darmstadt (TUdatalib) under the titles "Transitions of loaded and unloaded walking dataset" (https://doi.org/10.48328/tudatalib-2068) and "Hip exoskeleton walking with different assistance timings dataset" (https://doi.org/10.48328

/tudatalib-2066). Portions of the MATLAB code used to calculate EMG features were originally provided by Jingwei Too (https://github.com/JingweiToo/EMG-Feature-Extraction-Toolbox (accessed on 28 November 2023) [30,76]). We have extended and modified this code to include additional features. It is available as supplement of this manuscript and in the TU Darmstadt repository under the title "EMG Feature Extraction Toolbox (Extended Version)" (https://doi.org/10.48328/tudatalib-2069, accessed on 28 November 2023).

Acknowledgments: The authors acknowledge C. Gatti for revising the manuscript. During the preparation of this manuscript, the authors used ChatGPT (OpenAI, GPT-5.1) to assist with improving the clarity, grammar, and flow of the English text. The authors reviewed, edited, and take full responsibility for the content of the manuscript. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A

For all formulas, the following unification rules have been applied to ensure that variable names and iteration counters are consistent: N: Total sample count in the time domain. M: Frequency bins. Iteration counters: Standardized to use i for sums over time-domain data, j for sums over frequency-domain data, and p for parameter counters.

Average Amplitude Change (AAC) [28]

$$AAC = \frac{1}{N}\sum_{i=1}^{N-1}|x_{i+1} - x_i| \tag{A1}$$

Auto-Regressive Model/Coefficients (AR) [43–45]

The Auto-Regressive (AR) model is a prediction model that represents each sample of the EMG signal as a linear combination of previous samples, plus a white noise term. It can be expressed as follows:

$$x_i = \sum_{p=1}^{P} a_p x_{i-p} + w_i \tag{A2}$$

where a_p are model parameters, x_{i-p} are previous samples, and w_i is a white noise term. For our analysis, we selected an order P of one.

Absolute Value of Summation of Square Root (ASS) [41]

$$ASS = \left|\sum_{i=1}^{k}(x_i)^{1/2}\right| \tag{A3}$$

where k represents the number of analysis windows, and x_i denotes the data within each window. This feature is computed by taking the square root of each value within a window, summing the resultant values, and then taking the absolute value.

Cardinality (CARD) [46]

Cardinality measures the number of distinct values in a dataset:

$$y_n = \text{sort}(x_n), \quad n = 1, \ldots, N \; CARD = \sum_{n=1}^{N-1}(|y_n - y_{n+1}| > \varepsilon) \tag{A4}$$

where ε is a predefined threshold. Data is first sorted, and distinct values are determined based on the difference exceeding ε.

Cepstral Coefficients (CC) [28,44,47]

Cepstral analysis is the inverse Fourier transform of the logarithm of the power spectrum magnitude. The coefficients can be derived from the AR model:

$$c_1 = -a_1 \tag{A5}$$

$$c_p = -a_p - \sum_{l=1}^{p-1} \left(1 - \frac{l}{p}\right) a_l c_{p-l} \tag{A6}$$

where a_p represents the p-th AR coefficient and c_p represents the p-th Cepstral coefficient.

Coefficient of Variation (COV) [48]

$$COV = \frac{\sigma}{\mu} \tag{A7}$$

where σ is the standard deviation of the rectified EMG, and μ represents the mean absolute value (MAV).

Difference Absolute Mean Value (DAMV) [49]

$$DAMV = \frac{1}{N-1} \sum_{i=1}^{N-1} |x_{i+1} - x_i| \tag{A8}$$

Difference Absolute Standard Deviation Value (DASDV) [28,49]

$$DASDV = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N-1} (x_{i+1} - x_i)^2} \tag{A9}$$

Difference Variance Value (DVARV) [50]

We adapted $DVARV$ from the original formulation to include $N-1$:

$$DVARV = \frac{1}{N-1} \sum_{i=1}^{N-1} (x_{i+1} - x_i)^2 \tag{A10}$$

Energy (EN) [51]

$$EN = \sum_{i=1}^{N} |x_i|^2 \tag{A11}$$

Frequency Ratio (FR) [28,52]

The frequency ratio (FR) is used to distinguish between contraction and relaxation phases of muscle activity by comparing low and high frequency components:

$$FR = \frac{\sum_{j=LLC}^{ULC} P_j}{\sum_{j=LHC}^{UHC} P_j} \tag{A12}$$

where LLC and ULC represent the lower and upper cutoff frequencies for the low-frequency band, while LHC and UHC represent those for the high-frequency band. We adopted cutoff

frequencies of 30–250 Hz for the low band and 250–1000 Hz for the high band, as proposed by Han et al. [52].

Integrated EMG (IEMG) [28]

$$IEMG = \sum_{i=1}^{N} |x_i| \tag{A13}$$

Interquartile Range (IQR) [53]

The interquartile range (IQR) measures the variability between the 75th and 25th percentiles:

$$IQR = Q_3 - Q_1 \tag{A14}$$

where Q_3 is the third quartile and Q_1 is the first quartile.

Kurtosis (KURT) [53,54]

$$KURT = \frac{N \sum_{i=1}^{N} (x_i - \bar{x})^4}{\left(\sum_{i=1}^{N} (x_i - \bar{x})^2 \right)^2} \tag{A15}$$

The MATLAB internal function was used to calculate kurtosis.

Log Coefficient of Variation (LCOV) [55]

$$LCOV = \log(\text{COV}) \tag{A16}$$

Log Detector (LD) [28]

$$LD = \exp\left(\frac{1}{N} \sum_{i=1}^{N} \log(|x_i|) \right) \tag{A17}$$

Log Difference Absolute Mean Value (LDAMV) [50]

$$LDAMV = \log(\text{DAMV}) \tag{A18}$$

Log Difference Absolute Standard Deviation (LDASD) [50]

$$LDASD = \log(\text{DASDV}) \tag{A19}$$

Log Teager Kaiser Energy Operator (LTKEO) [55]

The non-linear TKEO measures the instantaneous energy changes of signals. The LTKEO is calculated as follows:

$$LTKEO = \log\left(\sum_{i=0}^{N-2} x_i^2 - x_{i-1} x_{i+1} \right) \tag{A20}$$

Mean Absolute Deviation (MAD) [53]

$$MAD = \frac{1}{N} \sum_{i=1}^{N} |x_i - \mu| \tag{A21}$$

where μ is the mean value of the signal.

Mean Absolute Value (MAV) [28,47,51,56,57]

$$MAV = \frac{1}{N}\sum_{i=1}^{N}|x_i| \tag{A22}$$

Maximum Amplitude (MAX) [57]

$$MAX = \max(x_i), \quad i = 1, \ldots, N \tag{A23}$$

where the rectified EMG signal was used.

Median Frequency (MDF) [28,58]

$$\sum_{j=1}^{MDF} P_j = \frac{1}{2}\sum_{j=1}^{N} P_j \tag{A24}$$

where P_i is the EMG power spectrum at frequency bin i, and N is length of the frequency bin.

Median Absolute Value (MED) -by authors-

$$MED = \begin{cases} X_{(N/2)} & \text{if } N \text{ is odd} \\ \frac{X_{(N/2)}+X_{(N/2+1)}}{2} & \text{if } N \text{ is even} \end{cases} \tag{A25}$$

The feature was designed based on the idea that reducing the impact of outliers, compared to the MAV, could result in improved detection performance.

Maximum Fractal Length (MFL) [59,60]

$$MFL = \log_{10}\left(\sqrt{\sum_{n=1}^{N-1}(x_{n+1}-x_n)^2}\right) \tag{A26}$$

Modified Median Frequency (MMDF) [61]

$$MMDF = \sum_{i=1}^{MMDF} A_i = \sum_{i=MMDF}^{M} A_i = \frac{1}{2}\sum_{i=1}^{M} A_i \tag{A27}$$

where A_i is the EMG amplitude spectrum at frequency bin i.

Modified Mean Frequency (MMNF) [61]

$$MMNF = \sum_{i=1}^{M} f_i A_i \Big/ \sum_{i=1}^{M} A_i \tag{A28}$$

where f_i is the frequency of spectrum at frequency bin i.

Mean Energy (MnE) [62,63]

$$MnE = \frac{1}{N}\sum_{i=1}^{N} x(i)^2 \tag{A29}$$

Mean Frequency (MNF) [28,58]

The mean frequency (MNF) is an average frequency computed as:

$$MNF = \frac{\sum_{i=1}^{M} f_i P_i}{\sum_{i=1}^{M} P_i} \tag{A30}$$

where M is the length of the frequency bin, f_i is the frequency of bin i, and P_i is the power spectrum at bin i.

Mean Power (MNP) [28]

The mean power (MNP) represents the average power of the EMG power spectrum:

$$MNP = \frac{\sum_{j=1}^{M} P_j}{M} \tag{A31}$$

Mean Value of Square Root (MSR) [41]

The MSR measures the total amount of activity in each analysis window:

$$MSR = \frac{1}{k} \sum_{i=1}^{k} (|x_i|)^{1/2} \tag{A32}$$

where k is the number of windows, and x_i is the data in each window.

Myopulse Percentage Rate (MYOP) [64,65]

The myopulse percentage rate (MYOP) counts the number of times the signal exceeds a predefined threshold:

$$MYOP = \frac{1}{N} \sum_{i=1}^{N} [f(x_i)] \quad \text{with} \quad f(x) = \begin{cases} 1, & \text{if } x \geq \text{threshold} \\ 0, & \text{otherwise} \end{cases} \tag{A33}$$

Frequency (PKF) [28,66]

The peak frequency (PKF) is the frequency at which the maximum power occurs:

$$PKF = \max(P_j), \quad j = 1, \ldots, M \tag{A34}$$

The power spectrum ratio (PSR) compares energy near the peak value of the EMG power spectrum with the total energy:

$$PSR = \frac{\sum_{j=f_0}^{f_0+n} P_j}{\sum_{j=1}^{n} P_j} \tag{A35}$$

where f_0 is the frequency at which the peak power occurs. We used an integral limit of $n = 20$ and energy range from 10 Hz to 450 Hz as suggested in [28].

Root Mean Square (RMS) [28,49,56,57]

The root mean square (RMS) is defined as:

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^{N} x_i^2} \tag{A36}$$

Standard Deviation (SD) [51,56,57]

The standard deviation (SD) is used as a measure of variability:

$$SD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \tag{A37}$$

where μ represents the mean of the rectified EMG signal.

Sample Entropy (SE) [67,68]

Sample entropy (SE) quantifies the complexity of time-series data:

$$x(p) = [x(p+k)]_{k=0}^{m-1}, \quad p = 1, \ldots, n-m+1 \tag{A38}$$

$$SE(x, m, r) = -\ln\left(\frac{A^m(r)}{B^m(r)}\right) \tag{A39}$$

where m is the embedding dimension and r is the tolerance, which was set to $0.25 \times SD$.

Skewness (SKEW) [69]

Skewness measures the asymmetry of the probability distribution:

$$SKEW = \frac{N\sum_{i=1}^{N}(x_i - \bar{x})^3}{\left((N-1)\sum_{i=1}^{N}(x_i - \bar{x})^2\right)^{3/2}} \tag{A40}$$

The according Matlab function was used.

Spectral Moment 1 (SM1) [28]

Spectral moments are statistical features derived from the EMG power spectrum:

$$SM1 = \sum_{j=1}^{M} P_j f_j \tag{A41}$$

Spectral Moment 2 (SM2) [28]

$$SM2 = \sum_{j=1}^{M} P_j f_j^2 \tag{A42}$$

Spectral Moment 3 (SM3) [28]

$$SM3 = \sum_{j=1}^{M} P_j f_j^3 \tag{A43}$$

Slope Sign Change (SSC) [28,51,56,70]

The slope sign change (SSC) is a measure of the frequency of waveform slope changes:

$$SSC = \sum_{i=2}^{N-1} f[(x_i - x_{i-1}) \times (x_i - x_{i+1})] \tag{A44}$$

$$f(x) = \begin{cases} 1, & \text{if } x \geq \text{threshold} \\ 0, & \text{otherwise} \end{cases} \tag{A45}$$

Temporal Moment (TM) [71]

$$TM = \left| \frac{1}{N} \sum_{i=1}^{N} x_i^3 \right| \tag{A46}$$

Here, we used the absolute value of the third temporal moment, as the first and second are similar to MAV and VAR, respectively.

Total Power (TTP) [28]

Total power (TTP) is defined as:

$$TTP = \sum_{j=1}^{M} P_j \tag{A47}$$

Threshold Zero Crossing (TZC) [72]

The threshold zero crossing (TZC) represents the number of times the signal crosses a threshold:

$$TZC = \sum_{i=1}^{N-1} f(x_i, x_{i+1}) \tag{A48}$$

$$f(x_i, x_{i+1}) = \begin{cases} 1, & \text{if } (x_i > T \text{ and } x_{i+1} < T) \text{ or } (x_i < T \text{ and } x_{i+1} > T) \\ 0, & \text{otherwise} \end{cases} \tag{A49}$$

where T is a threshold used to reduce noise in the signal. Instead of using resting state data, we used the window with the lowest five percent of activity found throughout the entire analyzed stride.

Variance (VAR) [28,47,56]

$$VAR = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2 \tag{A50}$$

where $\bar{x}$ is the mean value of x_i.

V-Order (VO) [28,47]

$$VO = \left(\frac{1}{N} \sum_{i=1}^{N} x_i^o \right)^{\frac{1}{o}} \tag{A51}$$

An order o of three was chosen to explore variability, differing from the optimal value of two to avoid duplicating the RMS value.

Willison Amplitude (WA) [28,47,56]

A specific threshold_{WA} is taken for computation of the Willison Amplitude

$$WAMP = \sum_{i=1}^{N-1} f(|x_i - x_{i+1}|) \tag{A52}$$

$$f(x) = \begin{cases} 1, & \text{if } x \geq \text{threshold}_{WA} \\ 0, & \text{otherwise} \end{cases} \tag{A53}$$

Waveform Length (WL) [28,47,51,70]

Waveform length (WL) measures the cumulative length of the waveform over a given interval:

$$WL = \sum_{i=1}^{N-1} |x_{i+1} - x_i| \tag{A54}$$

Zero Crossing (ZC) [28,47,51,64,70]

Zero crossing (ZC) counts the number of times the signal crosses zero, considering a specific chosen threshold$_{ZC}$ to avoid noise:

$$ZC = \sum_{i=1}^{N-1} [\text{sgn}(x_i \times x_{i+1}) \cap |x_i - x_{i+1}| \geq \text{threshold}_{ZC}] \tag{A55}$$

$$\text{sgn}(x) = \begin{cases} 1, & \text{if } x > 0 \\ -1, & \text{if } x < 0 \end{cases} \tag{A56}$$

References

1. Grimmer, M.; Riener, R.; Walsh, C.J.; Seyfarth, A. Mobility related physical and functional losses due to aging and disease-a motivation for lower limb exoskeletons. *J. Neuroeng. Rehabil.* **2019**, *16*, 2. [CrossRef]
2. Cavuoto, L.; Megahed, F. Understanding fatigue: Implications for worker safety. *Prof. Saf.* **2017**, *62*, 16–19.
3. del Carmen Sanchez-Villamañan, M.; Gonzalez-Vargas, J.; Torricelli, D.; Moreno, J.C.; Pons, J.L. Compliant lower limb exoskeletons: A comprehensive review on mechanical design principles. *J. Neuroeng. Rehabil.* **2019**, *16*, 55. [CrossRef]
4. Pinto-Fernandez, D.; Torricelli, D.; del Carmen Sanchez-Villamanan, M.; Aller, F.; Mombaur, K.; Conti, R.; Vitiello, N.; Moreno, J.C.; Pons, J.L. Performance evaluation of lower limb exoskeletons: A systematic review. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2020**, *28*, 1573–1583. [CrossRef] [PubMed]
5. Tucker, M.R.; Olivier, J.; Pagel, A.; Bleuler, H.; Bouri, M.; Lambercy, O.; del R Millán, J.; Riener, R.; Vallery, H.; Gassert, R. Control strategies for active lower extremity prosthetics and orthotics: A review. *J. Neuroeng. Rehabil.* **2015**, *12*, 1. [CrossRef]
6. Sawicki, G.S.; Beck, O.N.; Kang, I.; Young, A.J. The exoskeleton expansion: Improving walking and running economy. *J. Neuroeng. Rehabil.* **2020**, *17*, 25. [CrossRef] [PubMed]
7. Quinlivan, B.T.; Lee, S.; Malcolm, P.; Rossi, D.M.; Grimmer, M.; Siviy, C.; Karavas, N.; Wagner, D.; Asbeck, A.; Galiana, I.; et al. Assistance magnitude versus metabolic cost reductions for a tethered multiarticular soft exosuit. *Sci. Robot.* **2017**, *2*, eaah4416. [CrossRef] [PubMed]
8. Koller, J.R.; Gates, D.H.; Ferris, D.P.; Remy, C.D. 'Body-in-the-loop' optimization of assistive robotic devices: A validation study. In Proceedings of the Robotics: Science and Systems, Ann Arbor, MI, USA, 18–22 June 2016 ; Volume 2016, pp. 1–10. [CrossRef]
9. Zhang, J.; Fiers, P.; Witte, K.A.; Jackson, R.W.; Poggensee, K.L.; Atkeson, C.G.; Collins, S.H. Human-in-the-loop optimization of exoskeleton assistance during walking. *Science* **2017**, *356*, 1280–1284. [CrossRef]
10. Ding, Y.; Kim, M.; Kuindersma, S.; Walsh, C.J. Human-in-the-loop optimization of hip assistance with a soft exosuit during walking. *Sci. Rob.* **2018**, *3*, eaar5438. [CrossRef]
11. Franks, P.W.; Bryan, G.M.; Martin, R.M.; Reyes, R.; Lakmazaheri, A.C.; Collins, S.H. Comparing optimized exoskeleton assistance of the hip, knee, and ankle in single and multi-joint configurations. *Wearable Technol.* **2021**, *2*, e16. [CrossRef]
12. Zhao, G.; Wang, X.; Grimmer, M. Rapid Human-in-the-Loop Optimization for Hip Exoskeleton Assistance During Walking: An EMG-based Approach. In Proceedings of the 12th International Symposium on Adaptive Motion of Animals and Machines (AMAM 2025), Darmstadt, Germany, 8–11 July 2025. [CrossRef]
13. Tankink, T.; Houdijk, H.; Tabucol, J.; Leopaldi, M.; Hijmans, J.M.; Carloni, R. Human-in-the-Loop Optimization of the Stiffness and Alignment of a Prosthetic Foot to Reduce the Metabolic Cost of Walking. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2025**, *33*, 2823–2833. [CrossRef]
14. Witte, K.A.; Fiers, P.; Sheets-Singer, A.L.; Collins, S.H. Improving the energy economy of human running with powered and unpowered ankle exoskeleton assistance. *Sci. Rob.* **2020**, *5*, eaay9108. [CrossRef]
15. Selinger, J.C.; Donelan, J.M. Estimating instantaneous energetic cost during non-steady-state gait. *J. Appl. Physiol.* **2014**, *117*, 1406–1415. [CrossRef]

16. Han, H.; Wang, W.; Zhang, F.; Li, X.; Chen, J.; Han, J.; Zhang, J. Selection of muscle-activity-based cost function in human-in-the-loop optimization of multi-gait ankle exoskeleton assistance. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2021**, *29*, 944–952. [CrossRef] [PubMed]
17. Xu, L.; Liu, X.; Chen, Y.; Yu, L.; Yan, Z.; Yang, C.; Zhou, C.; Yang, W. Reducing the muscle activity of walking using a portable hip exoskeleton based on human-in-the-loop optimization. *Front. Bioeng. Biotechnol.* **2023**, *11*, 1006326. [CrossRef]
18. Day, S. *Important Factors in Surface EMG Measurement*; Bortec Biomedical Ltd Publishers: Calgary, AB, Canada, 2002; pp. 1–17.
19. Abdoli-Eramaki, M.; Damecour, C.; Christenson, J.; Stevenson, J. The effect of perspiration on the sEMG amplitude and power spectrum. *J. Electromyogr. Kinesiol.* **2012**, *22*, 908–913. [CrossRef] [PubMed]
20. Takagi, K.; Ogawa, T.; Terada, E.; Kobayashi, M. Sweating and the electric resistance of the skin. *Acta Neuroveg.* **1962**, *24*, 404–412. [CrossRef]
21. Winkel, J.; Jørgensen, K. Significance of skin temperature changes in surface electromyography. *Eur. J. Appl. Physiol. Occup. Physiol.* **1991**, *63*, 345–348. [CrossRef]
22. Stewart, D.; Macaluso, A.; De Vito, G. The effect of an active warm-up on surface EMG and muscle performance in healthy humans. *Eur. J. Appl. Physiol.* **2003**, *89*, 509–513. [CrossRef]
23. Fröhlich, M.; Ludwig, O.; Zeller, P.; Felder, H. Changes in skin surface temperature after a 10-min warm-up on a bike ergometer. *Int. J. Kin. Sports Sci.* **2015**, *3*, 13–17. [CrossRef]
24. Mizrahi, J.; Voloshin, A.; Russek, D.; Verbitski, O.; Isakov, E. The influence of fatigue on EMG and impact acceleration in running. *Basic Appl. Myol.* **1997**, *7*, 111–118.
25. de Oliveira, C.F.; Soares, D.P.; Bertani, M.C.; Rodrigues, L.J.; Vila-Boas, J. Effects of Fast-Walking on Muscle Activation in Young Adults and Elderly Persons. *J. Nov. Physiother. Rehabil.* **2017**, *1*, 012–019. [CrossRef]
26. Eken, M.M.; Richards, R.; Beckerman, H.; van der Krogt, M.; Gerrits, K.; Rietberg, M.; de Groot, V.; Heine, M. Quantifying muscle fatigue during walking in people with multiple sclerosis. *Clin. Biomech.* **2020**, *72*, 94–101. [CrossRef] [PubMed]
27. Grimmer, M.; Zeiss, J.; Weigand, F.; Zhao, G. Exploring surface electromyography (EMG) as a feedback variable for the human-in-the-loop optimization of lower limb wearable robotics. *Front. Neurorobot.* **2022**, *16*, 948093. [CrossRef]
28. Phinyomark, A.; Phukpattaranont, P.; Limsakul, C. Feature reduction and selection for EMG signal classification. *Expert Syst. Appl.* **2012**, *39*, 7420–7431. [CrossRef]
29. Nazmi, N.; Abdul Rahman, M.A.; Yamamoto, S.I.; Ahmad, S.A.; Zamzuri, H.; Mazlan, S.A. A review of classification techniques of EMG signals during isotonic and isometric contractions. *Sensors* **2016**, *16*, 1304. [CrossRef]
30. Too, J.; Abdullah, A.R.; Saad, N.M. Classification of hand movements based on discrete wavelet transform and enhanced feature extraction. *Int. J. Adv. Comput. Sci. Appl.* **2019**, *10*, 83–89. [CrossRef]
31. Toledo-Pérez, D.C.; Rodríguez-Reséndiz, J.; Gómez-Loenzo, R.A.; Jauregui-Correa, J. Support vector machine-based EMG signal classification techniques: A review. *Appl. Sci.* **2019**, *9*, 4402. [CrossRef]
32. Grimmer, M.; Stasica, M.; Zhao, G. Exoskeleton developments at the Technical University of Darmstadt. In *Proceedings TecPsy 2023*; Technische Universität Darmstadt: Darmstadt, Germany, 2023; p. 82. [CrossRef]
33. Grimmer, M.; Zhao, G. Hip exoskeleton for cycling assistance. *Bioengineering* **2024**, *11*, 683. [CrossRef]
34. Grabowski, A.; Farley, C.T.; Kram, R. Independent metabolic costs of supporting body weight and accelerating body mass during walking. *J. Appl. Physiol.* **2005**, *98*, 579–583. [CrossRef]
35. Browning, R.C.; Modica, J.R.; Kram, R.; Goswami, A. The effects of adding mass to the legs on the energetics and biomechanics of walking. *Med. Sci. Sports Exerc.* **2007**, *39*, 515–525. [CrossRef]
36. Damewood, B.A.P.; Sinkjær, T.; Thompson, A.K. Generation and modification of human locomotor EMG activity when walking faster and carrying additional weight. *Exp. Physiol.* **2025**, *110*, 1316–1335. [CrossRef] [PubMed]
37. Miller, J.F.; Stamford, B.A. Intensity and energy cost of weighted walking vs. running for men and women. *J. Appl. Physiol.* **1987**, *62*, 1497–1501. [CrossRef]
38. Grimmer, M.; Schmidt, K.; Duarte, J.E.; Neuner, L.; Koginov, G.; Riener, R. Stance and swing detection based on the angular velocity of lower limb segments during walking. *Front. Neurorobot.* **2019**, *13*, 57. [CrossRef]
39. Brockway, J. Derivation of formulae used to calculate energy expenditure in man. *Hum. Nutr. Clin. Nutr.* **1987**, *41*, 463–471. [PubMed]
40. Nassour, J.; Zhao, G.; Grimmer, M. Soft pneumatic elbow exoskeleton reduces the muscle activity, metabolic cost and fatigue during holding and carrying of loads. *Sci. Rep.* **2021**, *11*, 12556. [CrossRef]
41. Samuel, O.W.; Zhou, H.; Li, X.; Wang, H.; Zhang, H.; Sangaiah, A.K.; Li, G. Pattern recognition of electromyography signals based on novel time domain features for amputees' limb motion classification. *Comput. Electr. Eng.* **2018**, *67*, 646–655. [CrossRef]
42. Oskoei, M.A.; Hu, H. Support vector machine-based classification scheme for myoelectric control applied to upper limb. *IEEE Trans. Biomed. Eng.* **2008**, *55*, 1956–1965. [CrossRef]
43. Park, S.H.; Lee, S.P. EMG pattern recognition based on artificial intelligence techniques. *IEEE Trans. Rehabil. Eng.* **1998**, *6*, 400–405. [CrossRef]

44. Zecca, M.; Micera, S.; Carrozza, M.C.; Dario, P. Control of multifunctional prosthetic hands by processing the electromyographic signal. *Crit. Rev. Biomed. Eng.* **2002**, *30*, 459–485. [CrossRef]
45. Subasi, A. Classification of EMG signals using combined features and soft computing techniques. *Appl. Soft Comput.* **2012**, *12*, 2188–2198. [CrossRef]
46. Waris, A.; Kamavuako, E.N. Effect of threshold values on the combination of EMG time domain features: Surface versus intramuscular EMG. *Biomed. Signal Process. Control* **2018**, *45*, 267–273. [CrossRef]
47. Tkach, D.; Huang, H.; Kuiken, T.A. Study of stability of time-domain features for electromyographic pattern recognition. *J. Neuroeng. Rehabil.* **2010**, *7*, 21. [CrossRef] [PubMed]
48. Karthick, P.; Ramakrishnan, S. Surface electromyography based muscle fatigue progression analysis using modified B distribution time–frequency features. *Biomed. Signal Process. Control* **2016**, *26*, 42–51. [CrossRef]
49. Kim, K.S.; Choi, H.H.; Moon, C.S.; Mun, C.W. Comparison of k-nearest neighbor, quadratic discriminant and linear discriminant analysis in classification of electromyogram signals based on the wrist-motion directions. *Curr. Appl. Phys.* **2011**, *11*, 740–745. [CrossRef]
50. Phinyomark, A.; Quaine, F.; Charbonnier, S.; Serviere, C.; Tarpin-Bernard, F.; Laurillau, Y. Feature extraction of the first difference of EMG time series for EMG pattern recognition. *Comput. Methods Programs Biomed.* **2014**, *117*, 247–256. [CrossRef]
51. Al-Angari, H.M.; Kanitz, G.; Tarantino, S.; Cipriani, C. Distance and mutual information methods for EMG feature and channel subset selection for classification of hand movements. *Biomed. Signal Process. Control* **2016**, *27*, 24–31. [CrossRef]
52. Han, J.S.; Song, W.K.; Kim, J.S.; Bang, W.C.; Lee, H.; Bien, Z. New EMG pattern recognition based on soft computing techniques and its application to control of a rehabilitation robotic arm. In *Proceedings of the 6th International Conference on Soft Computing, Iizuka, Japan, 1–4 October 2000*; Springer Nature: Berlin/Heidelberg, Germany, 2000 pp. 890–897.
53. Verma, A.R.; Gupta, B. Detecting neuromuscular disorders using EMG signals based on TQWT features. *Augment. Hum. Res.* **2020**, *5*, 8. [CrossRef]
54. Nazarpour, K.; Al-Timemy, A.H.; Bugmann, G.; Jackson, A. A note on the probability distribution function of the surface electromyogram signal. *Brain Res. Bull.* **2013**, *90*, 88–91. [CrossRef]
55. Khushaba, R.N.; Al-Timemy, A.H.; Al-Ani, A.; Al-Jumaily, A. A framework of temporal-spatial descriptors-based feature extraction for improved myoelectric pattern recognition. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2017**, *25*, 1821–1831. [CrossRef]
56. Ahsan, M.R.; Ibrahimy, M.; Khalifa, O. Neural network classifier for hand motion detection from EMG signal. In *Proceedings of the 5th Kuala Lumpur International Conference on Biomedical Engineering 2011: (BIOMED 2011), Kuala Lumpur, Malaysia, 20–23 June 2011*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 536–541. [CrossRef]
57. Daud, W.M.B.W.; Yahya, A.B.; Horng, C.S.; Sulaima, M.F.; Sudirman, R. Features extraction of electromyography signals in time domain on biceps brachii muscle. *Int. J. Model. Optim.* **2013**, *3*, 515. [CrossRef]
58. Kendell, C.; Lemaire, E.D.; Losier, Y.; Wilson, A.; Chan, A.; Hudgins, B. A novel approach to surface electromyography: An exploratory study of electrode-pair selection based on signal characteristics. *J. Neuroeng. Rehabil.* **2012**, *9*, 24. [CrossRef] [PubMed]
59. Arjunan, S.P. Fractal Features of Surface Electromyogram: A New Measure for Low Level Muscle Activation. Ph.D. Thesis, RMIT University, Melbourne, VIC, Australia, 2008.
60. Phinyomark, A.; Phukpattaranont, P.; Limsakul, C. Fractal analysis features for weak and single-channel upper-limb EMG signals. *Expert Syst. Appl.* **2012**, *39*, 11156–11163. [CrossRef]
61. Phinyomark, A.; Limsakul, C.; Phukpattaranont, P. A novel feature extraction for robust EMG pattern recognition. *J. Comput.* **2009**, *1*, 71–80. [CrossRef]
62. D'Alessandro, M.M. The Utility of Intracranial EEG Feature and Channel Synergy for Evaluating the Spatial and Temporal Behavior of Seizure Precursors. Ph.D. Thesis, Georgia Institute of Technology, Atlanta, GA, USA, 2001.
63. Fairley, J.A.; Georgoulas, G.; Mehta, N.A.; Gray, A.G.; Bliwise, D.L. Computer detection approaches for the identification of phasic electromyographic (EMG) activity during human sleep. *Biomed. Signal Process. Control* **2012**, *7*, 606–615. [CrossRef]
64. Philipson, L. The Electromyographic Signal Used for Control of Upper Extremity Prostheses and for Quantification of Motor Blockade During Epidural Anaesthesia. Ph.D. Thesis, Linköping University, Linköping, Sweden, 1987.
65. Fougner, A.L. Proportional Myoelectric Control of a Multifunction Upper-Limb Prosthesis. Master's Thesis, Norwegian University of Science and Technology, Trondheim, Norway, 2007.
66. Qingju, Z.; Zhizeng, L. Wavelet de-noising of electromyography. In *Proceedings of the IEEE/ICMA International Conference on Mechatronics and Automation, Luoyang, China, 25–28 June 2006*; IEEE: Piscataway, NJ, USA, 2006; pp. 1553–1558. [CrossRef]
67. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol.-Heart Circ. Physiol.* **2000**, *278*, H2039–H2049. [CrossRef]
68. Zhang, X.; Zhou, P. Sample entropy analysis of surface EMG for improved muscle activity onset detection against spurious background spikes. *J. Electromyogr. Kinesiol.* **2012**, *22*, 901–907. [CrossRef]
69. Khushaba, R.N.; Al-Ani, A.; Al-Jumaily, A. Orthogonal fuzzy neighborhood discriminant analysis for multifunction myoelectric hand control. *IEEE Trans. Biomed. Eng.* **2010**, *57*, 1410–1419. [CrossRef]

70. Hudgins, B.; Parker, P.; Scott, R.N. A new strategy for multifunction myoelectric control. *IEEE Trans. Biomed. Eng.* **1993**, *40*, 82–94. [CrossRef]
71. Saridis, G.N.; Gootee, T.P. EMG pattern analysis and classification for a prosthetic arm. *IEEE Trans. Biomed. Eng.* **1982**, *BME-29*, 403–412. [CrossRef]
72. Toledo-Perez, D.; Rodriguez-Resendiz, J.; Gomez-Loenzo, R.A. A study of computing zero crossing methods and an improved proposal for EMG signals. *IEEE Access* **2020**, *8*, 8783–8790. [CrossRef]
73. Grimmer, M.; Zeiss, J.; Weigand, F.; Zhao, G. Joint power, joint work and lower limb muscle activity for transitions between level walking and stair ambulation at three inclinations. *PLoS ONE* **2023**, *18*, e0294161. [CrossRef] [PubMed]
74. Moir, T.J. *Rudiments of Signal Processing and Systems*; Springer: Berlin/Heidelberg, Germany, 2022.
75. Just, F.; Ghinami, C.; Zbinden, J.; Ortiz-Catalan, M. Deployment of machine learning algorithms on resource-constrained hardware platforms for prosthetics. *IEEE Access* **2024**, *12*, 40439–40449. [CrossRef]
76. Too, J.; Abdullah, A.R.; Mohd Saad, N.; Tee, W. EMG feature selection and classification using a Pbest-guide binary particle swarm optimization. *Computation* **2019**, *7*, 12. [CrossRef]