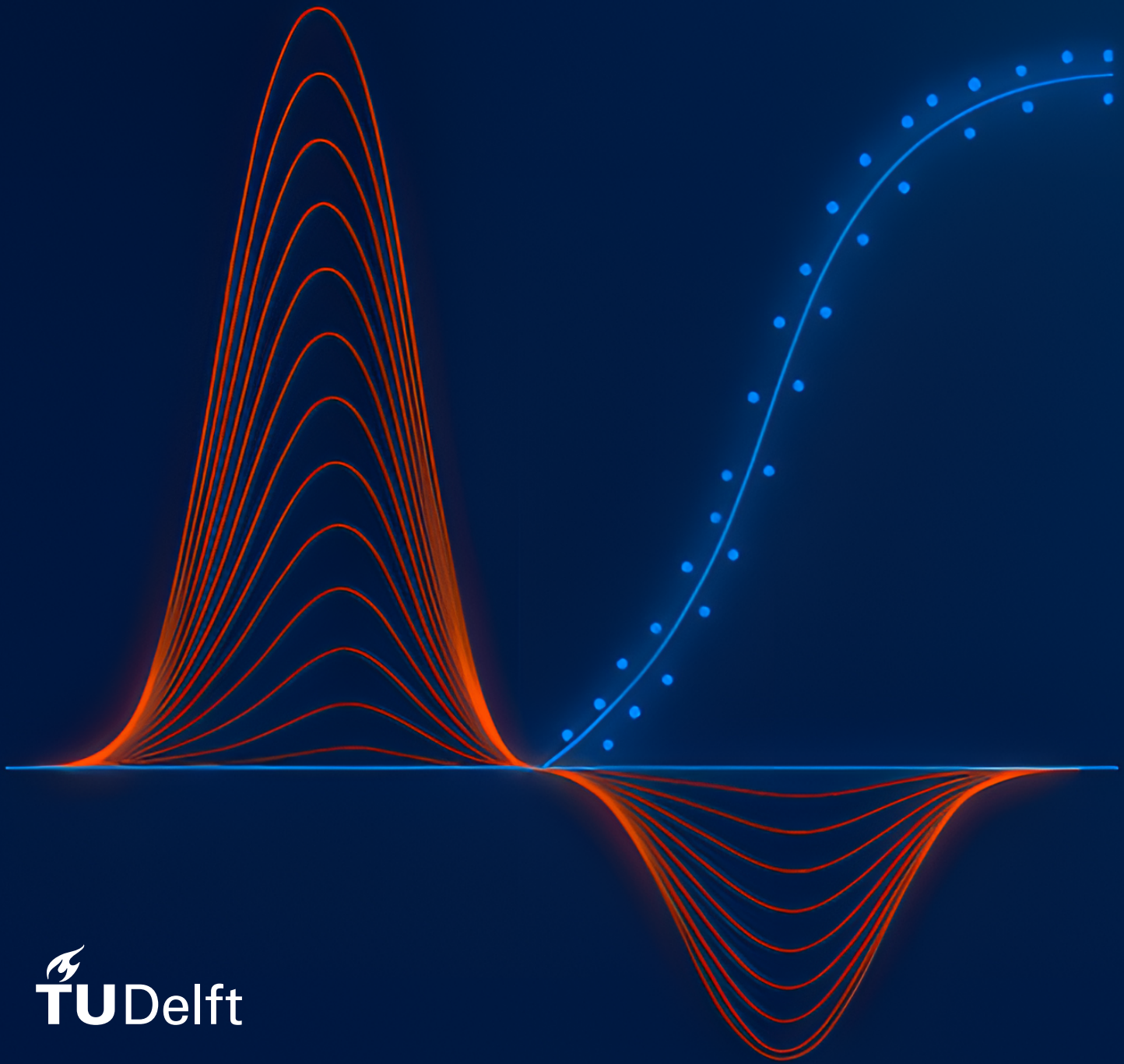
Wavolution: a Novel Algorithm for Motor Unit Number Estimation

Developing a new algorithm for motor unit number estimation from compound muscle action potential scans

P.L. Ottenhoff



TU Delft

Wavolution: a Novel Algorithm for Motor Unit Number Estimation

Developing a new algorithm for motor unit number estimation from compound muscle action potential scans

by

P.L. Ottenhoff

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on September 29th, 2025 at 16:00.

Student number:	4675843	
Project duration:	February 3rd, 2025 – September 29th, 2025	
Thesis committee:	Dr. ir. W. Mugge,	TU Delft, supervisor
	Prof. dr. ir. A. C. Schouten,	TU Delft
	Dr. ir. D. J. L. Stikvoort García	UMC Utrecht, daily supervisor

This thesis is confidential and cannot be made public until September 29th, 2025.

TU Delft

Preface

This thesis marks one of the final steps in completing my Master's in Biomedical Engineering at Technical University (TU) of Delft. The project was carried out from December 2024 to September 2025 and focused on identifying the strengths, limitations and feasibility for clinical trials of current motor unit number estimation (MUNE) methods, and developing a new algorithm to estimate the number of motor units from compound muscle action potential (CMAP) scans with the aim to improve disease monitoring in patients with amyotrophic lateral sclerosis (ALS).

This topic is very close to me. I chose to work on a project that involved ALS because someone I care about was affected by the disease. First of all, I am grateful that I have gotten the chance to work on this project, for which I want to thank Dr. ir. W. Mugge and Dr. ir. D. J. L. Stikvoort García in special. I am very thankful to Dr. ir. W. Mugge for his clear feedback and guidance in areas outside my expertise throughout the process. Secondly, I want to thank Dr. ir. D. J. L. Stikvoort García for all his practical input and ideas, feedback, shared knowledge and experience, and for his time to endure all my questions – it really helped me shape the direction of my thesis. I would also like to thank Prof. dr. ir. A. C. Schouten for his role in the committee, as well as the people at the UMC Utrecht for their help, data, and feedback I received during the project, with special thanks to Dr. ir. Boudewijn Sleutjes. Last but not least, I would like to thank all my friends and family for their support over the past months.

During the development of the novel program: "Wavolution", ChatGPT 5.0 Thinking was used. This Artificial Intelligence (AI) model was used to assist in writing the Python and JavaScript code that serves the application, as well as writing the Python code that was used to develop the algorithm and results present in this report. The use of AI was not a replacement for my own work, but rather a supportive tool to help accelerate programming tasks. During this project, I have gained a lot of experience in formulating clear prompts and carefully refining instructions were essential in obtaining relevant outputs for such tasks. Finally, I would like to point out that any errors or inaccuracies remain my responsibility.

P.L. Ottenhoff
Utrecht, September 2025

Contents

Abstract

Introduction

Compound muscle action potential (CMAP) scans are detailed stimulus-response curves that provide information on the activation of motor units (MUs) and can be used to provide a motor unit number estimate (MUNE). As loss of MUs is the hallmark of amyotrophic lateral sclerosis (ALS), accurate MUNE will be a valuable biomarker for monitoring disease progression in patients with ALS. MScanFit, the current gold standard in CMAP scan-based MUNE, is efficient and widely used, but underestimates MUNE in muscles with a high number of MUs.

Methods

The current study introduces Wavolution, a waveform-based algorithm that estimates MUNE from CMAP scans. Whereas MScanFit assigns solely numerical values to MU properties in its fitting procedure, Wavolution assigns single MU action potentials (SMUAPs) to MUs to account for physiological phenomena such as phase cancellation and temporal dispersion present in CMAP recordings. Furthermore, Wavolution has a highly efficient way of simulating CMAP scans by matrix multiplication. The fitting procedure of simulated CMAP scans onto a target CMAP scan consists of two stages: (1) an initialization phase that generates preliminary MU pools based on amplitude and threshold density distributions of the input scan, and (2) an optimization phase, in which a genetic-like algorithm iteratively updates the simulated MU pools to minimize the error between simulated and target CMAP scans.

Results

Wavolution was validated on 920 simulated CMAP scans, ranging from 5 to 150 MUs with each MU number containing various noise levels ranging from 1.0 to 100.0 µV. Wavolution achieved significantly lower percentual MUNE discrepancy compared to MScanFit (13.2% vs. 25.5%) and had a significantly reduced computation time (67.8 vs. 92.9 seconds per scan). Wavolution achieved higher accuracy in the mean amplitude size of the MUs in the estimated MU pool than MScanFit when MU counts are high and reproduced reductions in maximum CMAP amplitude caused by phase cancellation and temporal dispersion, although it tended to overestimated their magnitude.

Discussion

Wavolution provides a physiologically transparent and computationally efficient approach to CMAP scan–based MUNE. By modeling SMUAP waveforms, it accounts for phase cancellation effects and reduces the ceiling effects of MScanFit in muscles with high MU counts, where muscles are still largely intact. A Windows application was developed to enable bulk processing, reducing operator workload in large-scale ALS trials. Future research should focus on extending the algorithm to make full use of the information available in CMAP recordings, validating Wavolution on experimental CMAP scans, and improving noise estimation. With further development, Wavolution could become a valuable tool for disease monitoring in ALS clinical trials.

1

Introduction

Amyotrophic lateral sclerosis (ALS) is a multisystem neurodegenerative disorder characterized by the degeneration of central and peripheral motor neurons, causing progressive weakness, paralysis and eventually death [1, 2]. With the exception of Riluzole, no effective disease-modifying treatments have been approved for the general population of ALS patients [3, 4]. For patients with the rare SOD1 mutation, however, the recent approval of Tofersen, an antisense oligonucleotide, has proved that slowing ALS is possible. Reliable biomarkers are critical to demonstrate slowing of the disease [5, 6]. However, the heterogeneous presentation and a lack of reliable biomarkers for the disease prohibit early diagnosis and challenge accurate monitoring of disease progression, especially when muscles are largely intact [7–9]. With current biomarkers, the sample size required to obtain adequately powered results for drug efficacy is considerable (> 200 participants) and even further amplified by missing data, dropouts and death [10]. Biomarkers that are sensitive throughout the entire disease duration observed in clinical trials will, therefore, be integral to the validation of new treatments for patients with ALS.

As loss of motor units (MUs) is the hallmark of ALS, quantifying the remaining number of MUs is a sensible measure of the disease's progression. Although this actual number cannot be directly measured, it can be approximated using various motor unit number estimation (MUNE) methods [11]. Compound muscle action potential (CMAP) scan-based MUNE is increasingly being included in ALS trials as outcome measure due to its ease of application, reproducibility, and non-invasiveness [12–17]. The CMAP scan, colloquially referred to as the "muscle scan" or "MScan", is a highly detailed stimulus-response curve that represents the relationship between stimulus intensity and CMAP size. As MUs in a muscle have distinct activation thresholds, the number of firing MUs in a muscle decreases with a decreasing stimulus current. A CMAP scan records the response size of a muscle for supramaximal to subthreshold stimulus currents, providing a detailed stimulus-response curve of the entire muscle [18].

MScanFit is a program for MUNE that functions by means of minimizing the discrepancy between a CMAP scan of an estimated MU pool and a target CMAP scan [12]. Due to its high processing efficiency and reproducibility, it is the current gold standard in CMAP scan-based MUNE (Appendix A). However, results are influenced by initial user settings of the program and in muscles with a large number of MUs, MScanFit lacks precision and underestimates MUNE [12]. As a result, this method encounters ceiling effects in early disease stages when muscles are largely intact. In large-scale clinical trials, additionally, the total operator time becomes substantial, as multiple muscles of all participants are measured over multiple visits. Every fitted CMAP scan has to be evaluated manually before and after performing the fitting procedure. A reduction in computation time and manual operator time with bulk processing can therefore offer large benefits in large-scale clinical trials, making it more practical for routine application.

Therefore, the aim of the current research was to develop a new algorithm that: (1) is less sensitive to ceiling effects than MScanFit MUNE and (2) has a reduced computation time. This algorithm, termed "Wavolution", contains a computationally efficient but extensive model for generating realistic MU pools and simulating CMAP scans, as well as an algorithm for the fitting procedure of estimated MU pools, which consists of two stages: an initialization stage and an optimization stage. The performance and accuracy on performing MUNE from CMAP scans of Wavolution were evaluated by comparing key metrics, including the percentual MUNE discrepancy and computation time, to MScanFit.

2

Methods

2.1. Simulating CMAP scans

A CMAP scan represents the amplitude size of individual CMAP recordings over a range of stimulus intensities, where the CMAP recordings consist of superimposed waveforms of SMUAPs. Wavolution assigns an actual recorded SMUAP to each MU in an estimated MU pool to generate a CMAP scan. Five distinct normalized SMUAP waveforms ($A_{\text{max}} = 1$ µV) varying in shape were extracted from ALS patient recordings of the abductor pollicis brevis (APB) muscle (Appendix B) to serve as fundamental waveforms for MU pools (Figure 2.1A). To generate realistic MU pools, transformation were applied to these fundamental waveforms' amplitude (A), phase (φ), distal motor latency (δ), activation threshold (t), and relative spread (RS) of the activation threshold (ρ). As such, parameters A, φ, and δ describe the shape and size of the simulated SMUAP (Figure 2.1B). This is an extension of the MScanFit model that assigns only scalar amplitude value to each MU rather than actual waveforms [12]. Parameters t and ρ describe the probability of firing of a MU, which is defined by a cumulative Gaussian probability function with mean $\mu = t$ and standard deviation $\sigma = \frac{\rho \cdot t}{100}$ [19] (Figure 2.1C). As MUs exhibit "all-or-none" firing behavior, this probabilistic firing results in alternation effects in the responses of individual MUs, of which $> 99\%$ is captured within $\mu \pm 3\sigma$ (Figure 2.1D). When these regions of alternation of MUs overlap, it is hard to determine if CMAP amplitudes stem from alternation effects or from contribution of newly recruited MUs. Furthermore, SMUAPs typically have an initial positive phase (upward) in their bi-phasic waveform (SMUAP 1 in Figure 2.1B). However, in cases of advanced reinnervation, inverted SMUAPs (*e.g.* an initial negative phase (downward) in the bi-phasic waveform) can be recorded (SMUAP 2 in Figure 2.1B) [20]. As CMAP recordings are superpositions imposed by individual waveforms, opposing phases of SMUAPs in CMAP recordings exacerbate phase cancellation, where the positive and negative phases of SMUAPs interfere and cancel each other out [21]. These phase cancellation effects can decrease the maximum CMAP amplitude up to 40% and can lead to a decrease in CMAP amplitude with a newly recruited MU (Figure 2.1B and D) [21]. Similarly, the contribution to CMAP amplitude by the actual amplitudes of individual SMUAP waveforms can be masked by temporal dispersion, which is caused by distal motor latency and shape of the SMUAP waveform (Appendix A). The effects of temporal dispersion and phase cancellation subsequently impact the estimation of MU pool sizes and MU properties.

Wavolution stores various transformed waveforms of SMUAPs of MU pools in a matrix, M_{SMUAP}, with each row denoting a different MU and each column the amplitude of the SMUAP at each timepoint (Appendix B). The activation thresholds (t) and relative spreads of the activation thresholds (ρ) of the MUs are stored in in a firing matrix M_{firing} of size $N \times S$; for every MU N, firing values 0 (no activation) and 1 (activation) are randomly drawn from the distributions in Figure 2.1C for each stimulus intensity S (Appendix B):

$$M_{N\times S} = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & \cdots \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{2.1}$$

Multiplying the matrix of transformed SMUAPs (one per MU) by the firing matrix results in a matrix

of accumulated SMUAPs for every stimulus intensity:

$$M^{T}_{\text{SMUAP}(N\times T)} \times M_{\text{firing }(N\times S)} = M_{\text{CMAP }(T\times S)} \tag{2.2}$$

Figure 2.1D shows a small region of a CMAP scan, with the accumulated SMUAP waveforms of the MUs presented in Figure 2.1B. For every stimulus intensity, the maximum amplitude (baseline-peak) of the CMAP was extracted and baseline noise was added (Figure 2.1D and E). As such, Wavolution simulates realistic CMAP scans, taking into account interference between SMUAP waveforms. The use of matrix multiplication is highly computationally efficient – the generation of 1000 CMAP scans takes approximately 6.5 seconds on a personal computer (Windows 11, Intel i7).

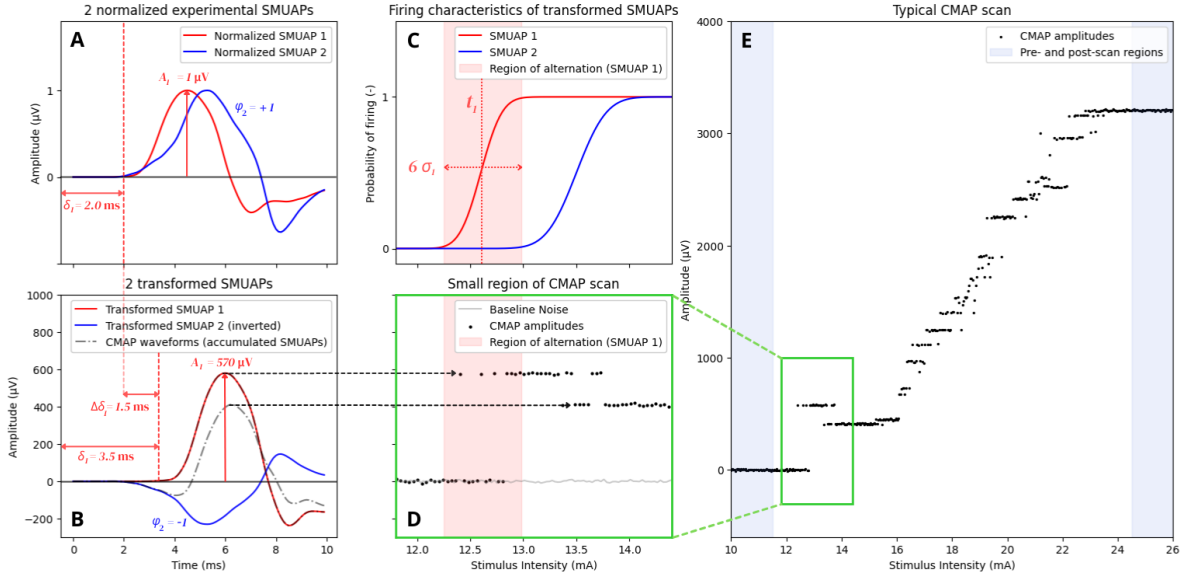


Figure 2.1: Generation of a MU pool and simulation of the CMAP scan using single SMUAP waveforms: (A) 2 distinct normalized ($A_{\text{max}} = 1$ µV) SMUAP waveforms extracted from ALS patient recordings of the abductor pollicis brevis (APB) muscle. (B) Transformation are applied to the normalized waveforms, changing amplitude (A), distal motor latency (δ), and phase (φ). (C) The activation threshold and relative spread of the MUs describe the firing behavior of each MU for the stimulus range. (D) CMAP amplitudes are extracted from the accumulated CMAP waveforms of B. As the second SMUAP is inverted (*e.g.* an initial negative (downward) phase in the bi-phasic waveform), there is a decrease in CMAP amplitude. Baseline noise ($\sigma = 5$ µV) is added to simulate realistic CMAP scans. (E) A typical CMAP scan that shows CMAP amplitudes over a range of stimulus intensities, where the first and last data points, called the pre- and post-scan regions, are used for noise estimation.

2.2. Novel algorithm for MUNE from CMAP scans: "Wavolution"

The novel algorithm of Wavolution for the fitting procedure consists of two main stages: (1) the initialization stage, which assesses a target CMAP scan to obtain model parameters and generate preliminary estimated MU pools; and (2) the optimization stage, which makes use of a evolutionary-like algorithm. A flowchart of the general fitting process is shown in Figure 2.2. Wavolution aims to optimize the fitness of multiple MU pools such that the resulting estimated CMAP scans match the target CMAP scan.

2.2.1. Error score

To quantify the error between an estimated and target CMAP scan, a combined error function of 4 error metrics was created. The error metrics quantify the absolute error, global trend of the stimulus-response curve, amplitude densities and densities of the differences in amplitude. Consequently, the error function captures local and global differences between the estimated and target CMAP scan. The error function quantified as follows:

$$E_{\text{total}} = (w_1 \cdot E_{\text{abs, mean}}) + (w_2 \cdot E_{\text{amp, PDF}}) + (w_3 \cdot E_{\text{fit, mean}}) + (w_4 \cdot E_{\text{thresh, PDF}}) \tag{2.3}$$

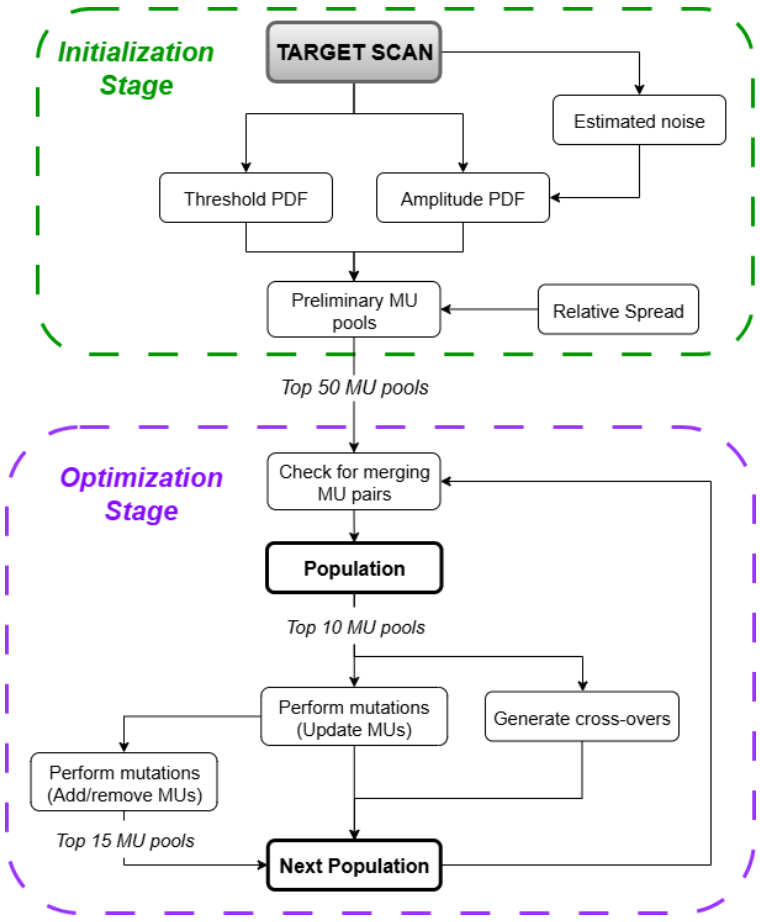


Figure 2.2: Flowchart of Wavolution, including the initialization stage in **green** and the optimization stage in **purple**. In the initialization stage, preliminary estimated MU pools are generated based on an amplitude and threshold density of the target scan. These preliminary estimated MU pools are sorted based on their error score, where the top 50 MU pools are selected as the starting population for the evolutionary algorithm. In the optimization stage, this starting population of MU pools is optimized for multiple generations to find the MU pool with the lowest error score. First, MU pools are checked for possible merging MU pairs based on their amplitudes or thresholds. Consequently, mutations and cross-overs are performed to the 10 best MU pools of a population and stored as the next population: mutations are performed by updating a MU pool based on its amplitude density and cross-overs are generated by combining two MU pools. The mutated MU pools are mutated again, where MUs are added and removed to generate 100 new MU pools. The top 15 MU pools of these 100 are added to the next population as well, after which the iterative process starts again.

where w_i are the weights in the error function. The weights of the error function were chosen such that each error score contributes an approximately equal proportion to the total error and were kept fixed throughout the optimization process for all target scans (Appendix C). $E_{\text{abs, mean}}$ is the absolute mean error between the estimated CMAP scan and the target scan (Figure 2.3A) and is used to reflect the absolute error of all data points in the CMAP scan. $E_{\text{amp, PDF}}$ is the difference in area between amplitude probability density function of the estimated and target CMAP scans, where large distances and height differences are amplified (Figure 2.3B). The amplitude probability density function is used to analyze the amount of data points at each unique amplitude of the CMAP scan and is not affected by alternation effects. $E_{\text{fit, mean}}$ is the mean error of the smoothed fit through the CMAP amplitudes of the estimated and target CMAP scans (Figure 2.3C) and is used to reflect the global trend of the CMAP scan. $E_{\text{thresh, PDF}}$ is the difference in area between threshold probability density function of the estimated and target CMAP scans (Figure 2.3D). The threshold probability density function is used to analyze the locations of steps over the whole stimulus range of the CMAP scan and is not affected by alternation effects.

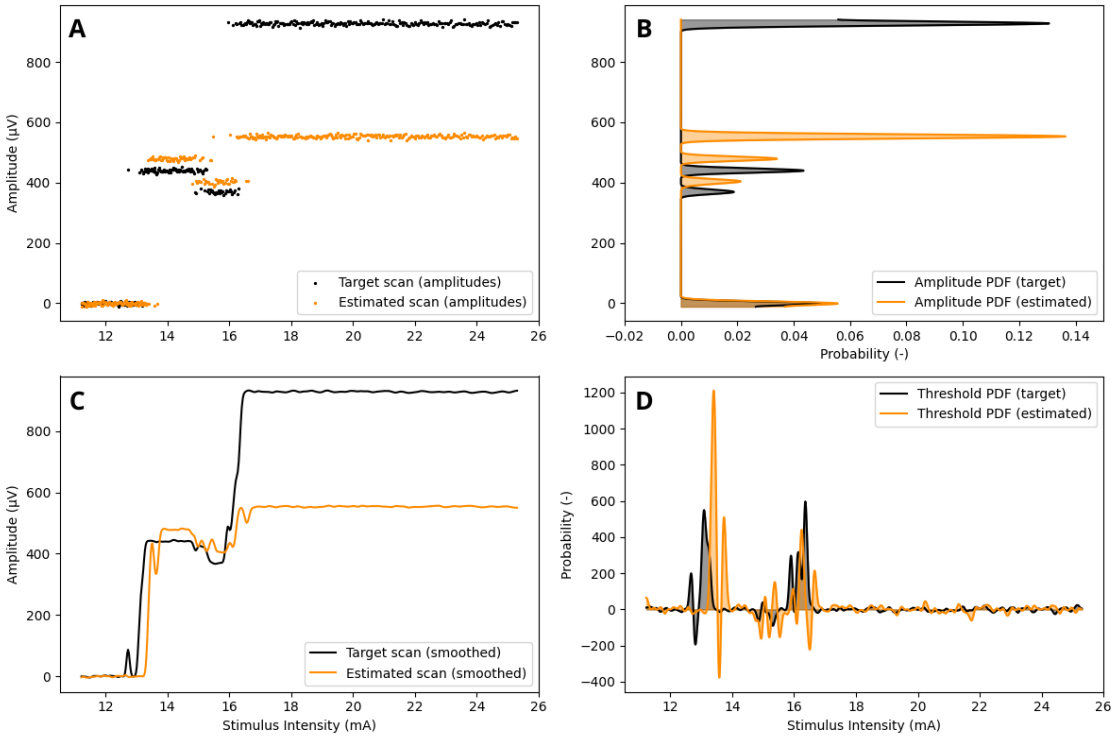


Figure 2.3: The separate error score metrics for an estimated and target scan. (A) The absolute amplitudes are used for the mean absolute error. (B) The difference between the target and estimated amplitude PDF is used for a summed error score. (C) A smoothed trend line is used for the mean smoothed fitting error. (D) The difference between the target and estimated threshold PDF is used for a summed error score. The error score for this estimated MU pool is relatively high, as there is a large amplitude difference from 16 mA, leading to a high absolute error score (A), a high amplitude PDF error score (B) and a high smoothed error score (C).

2.2.2. Initialization stage

A CMAP scan is typically recorded with pre- and post-scan sweeps at constant supramaximal and subthreshold stimulus levels to assess the variability of the responses and the baseline noise [12]. Wavolution first estimates the noise level of the target scan as the standard deviation of the pre- and post-scan regions (Figure 2.1E) – the first 10 and last 10 data points of the target scan. These 20 data points are only a small region of the target scan, which can lead to an underestimation of the true noise. To account for uncertainty and underestimation of the actual noise, the noise estimate was multiplied by a range of factors between 1.0 and 2.5. For every scaled estimated noise value, preliminary estimated MU pools were generated and the associated noise value was stored, as this value is used for generating additional noise to estimated CMAP scans (Figure 2.1D and E), as well as for generating an amplitude density distribution of a CMAP scan.

Subsequently, an amplitude and threshold density distribution are generated of the target scan. The amplitude and threshold probability density functions are generated by Gaussian smoothing of a histogram of the amplitude values and stimulus values of amplitude differences of a CMAP scan. The spread of the smoothing of the amplitude and threshold histograms is determined by the estimated noise and threshold spread (Appendix C). Unique contributing amplitude values of a CMAP scan are extracted by taking the peaks of the amplitude density distribution. The threshold peaks in the threshold density distribution are assigned to a unique amplitude peak and are used to generate the threshold values of the SMUAPs. To ensure coverage of the entire potential range of MUNE values, Wavolution generates preliminary estimated MU pools with maximum MU counts up to 200. If $N_{\text{peaks}} < 200$, N_{peaks} is maximum MU count for the preliminary MU pools. For a MU pool of N MUs, the N largest peaks in the amplitude density distribution are selected. As these peaks have corresponding stimulus intensity values, we estimated a threshold value for each unique amplitude. Subsequently, N SMUAPs are generated and assigned threshold and amplitude values, sorted by increasing stimulus intensity. An

SMUAP is assigned a phase value ($\varphi = +1$ or -1) based on if there is an increase or decrease in CMAP amplitude with respect to the previous CMAP amplitude peak, such that the SMUAP waveform either increases or decrease the amplitude of the CMAP waveform. Then, we assigned a random value from a Gaussian distribution based on Sleutjes *et al.* [19] ($(\mu, \sigma) = (1.65, 0.43)$) for the relative spread of the activation threshold to each MU in the MU pool. Finally, all preliminary MU pools are ranked based on their error score and the best 50 are kept as the starting population in the optimization stage.

2.2.3. Optimization stage

The optimization stage of Wavolution contains a evolutionary-like algorithm that optimizes a population of MU pools for multiple generations. The individual MU pools in a population are evaluated based on their average error score, where the individuals with the lowest average error score of a generation are called the best individuals. Figure 2.2 shows the sequence of steps in the evolutionary process, which are discussed in more detail below.

Checking for merging MU pairs

First, every MU pool in the population is checked for the possible presence of MU pairs that should be ’merged’. A MU pair is merged if: one MU in the pair is too small (< 5 μV); or the difference in activation thresholds of the MU pair is below the step size of the decremental stimuli, which is 0.2% of the stimulus intensity at that point ($|t_1 - t_2| < 0.2\%\ t_1$) (Appendix C).

Performing mutations (updating MUs)

Then, the whole population of multiple MU pools is ranked based on their error score according to Equation 2.3. The 10 fittest MU pools of generation X are mutated: for every MU pool, MUs are removed and added based on the amplitude density distribution, which indicates the errors in unique amplitudes between the target and estimated CMAP scan (Figure 2.4). A density peak is considered an erroneous peak if it has a minimum distance of 5 μV to another peak, as this is subjectively chosen to be the minimum size that a MU can be distinguished from noise, or if the estimated threshold distance is too large ($D_{\text{peaks}} > D_{\text{threshold}}$ (Appendix C)). An estimated MU pool is updated in one iteration such that the estimated amplitude PDF matches the target amplitude PDF, where the amount of MUs can drastically increase or decrease: first, all MUs at the locations of the erroneous peaks in the estimated scan are removed, after which MUs are added at the locations of the erroneous peaks of the target scan. These 10 mutated MU pools are added to the population of generation $X + 1$.

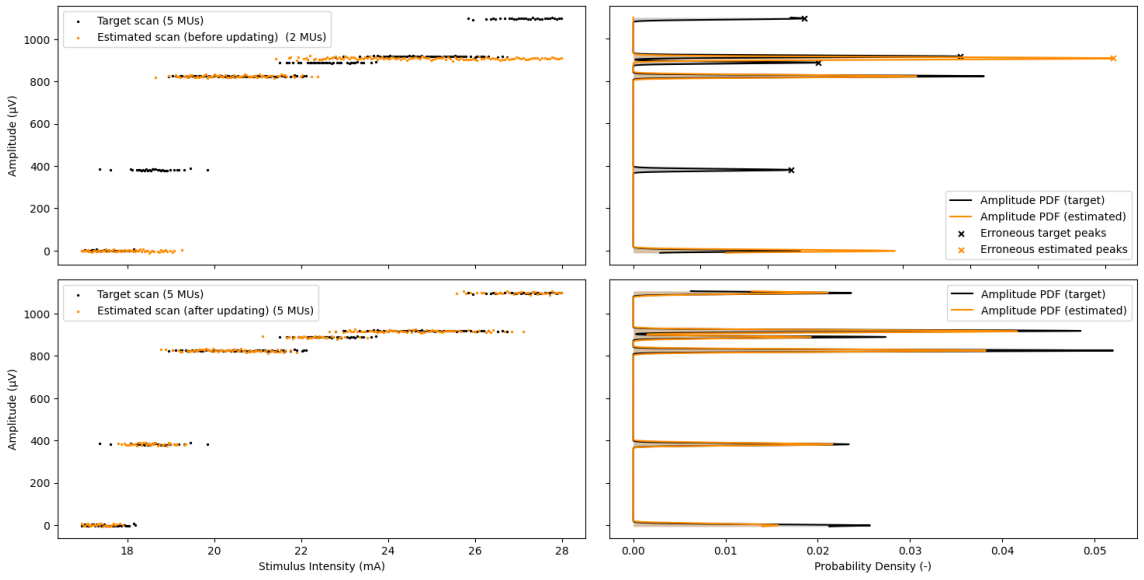


Figure 2.4: An example of updating MUs based on erroneous peaks in the amplitude density distributions of the target scan and an estimated scan. All erroneous peaks are updated in one iteration: all MUs at the erroneous peaks of the estimated CMAP scan are removed (orange peaks) and new MUs are added with properties corresponding to the erroneous peaks of the target scan (black peaks).

Performing mutations (adding/removing MUs)
Consequently, $1-5$ MUs are added and removed to each of the 10 mutated MU pools based on the location of the largest absolute error between the target and estimated CMAP scan. This results in 10 additional mutated MU pools, with a total of 100. The MU merging procedure is applied to these 100 mutated MU pools after which the MU pools are ranked based on their average error score. The fittest 15 MU pools are added to the population of generation $X+1$.

Generating cross-overs
Finally, the 10 fittest MU pools of generation X are used to generate cross-overs: a cross-over MU pool is generated by taking one half of MUs of MU pool A and the second half of MUs of MU pool B (Appendix C). With pairwise combinations of 10 unique parent MU pools, this leads to a total of $\binom{10}{2} = 45$ new cross-over MU pools, which are added to the population of generation '$X+1$' as well and the iterative process starts again ($X = X+1$). All estimated MU pools of each generations are stored. After 5 generations, the best 15 MU pools of all generations each generate 10 scans and are ranked by their average error score. The MU pool with the lowest average error score is decided to be the optimal MU pool.

2.3. Validation of Wavolution

2.3.1. Training and validation set

Using the five distinct SMUAP waveforms that were extracted from ALS patient recordings of the abductor pollicis brevis (APB) muscle (Appendix B), a training and validation set was built. In total, 16 different MU numbers ($[5, 10, 20, 30, 40, \ldots, 130, 140, 150]$) were tested. For each MU number, 10 different MU pools were generated, leading to 160 distinct MU pools. As MUs have probabilistic firing characteristics, two scans were generated from each MU pool, a 'test' and 'retest' scan. To assess the robustness of Wavolution at various noise levels, 5 distinct levels of increasing baseline noise were added to each scan. Thus, for every MU pool, 10 scans were simulated (2 (test/retest) $\times$ 5 (noise levels)), resulting in $160 \times 10 = 1600$ scans. This dataset was split up into a training (40%) set to empirically determine hyperparameters and develop the model, and validation (60%) set to assess the performance of Wavolution.

2.3.2. Validation of accuracy and performance

We compared the accuracy and performance of Wavolution to MScanFit, with the main outcome measure being the percentual MUNE discrepancy $\left(100\% \cdot \left(\frac{\text{MUNE} - N_{\text{MU, true}}}{N_{\text{MU, true}}}\right)\right)$. This assessment was performed at all noise levels included in the validation set. Similarly, the computation times of both algorithms were evaluated. To assess whether Wavolution accurately captures the individual amplitudes of MUs in the MU pool, we compared the absolute error in mean SMUAP size of Wavolution and MScanFit. Finally, to assess whether Wavolution accurately captures temporal dispersion and phase cancellation effects of the MU pool, we compared the amount of reduction in maximum CMAP amplitude due to temporal dispersion and phase cancellation of the estimated MU pool to that of the target MU pool.

2.4. Statistical Analysis

Differences in the validation metrics of the accuracy and performance between MScanFit and Wavolution were assessed at each noise level using a Mann-Whitney U-test with paired samples. To control for false positives, p-values were adjusted using the Benjamini-Hochberg method. Adjusted $p < 0.05$ were considered significant.

3

Results

3.1. Single scan accuracy and performance of Wavolution

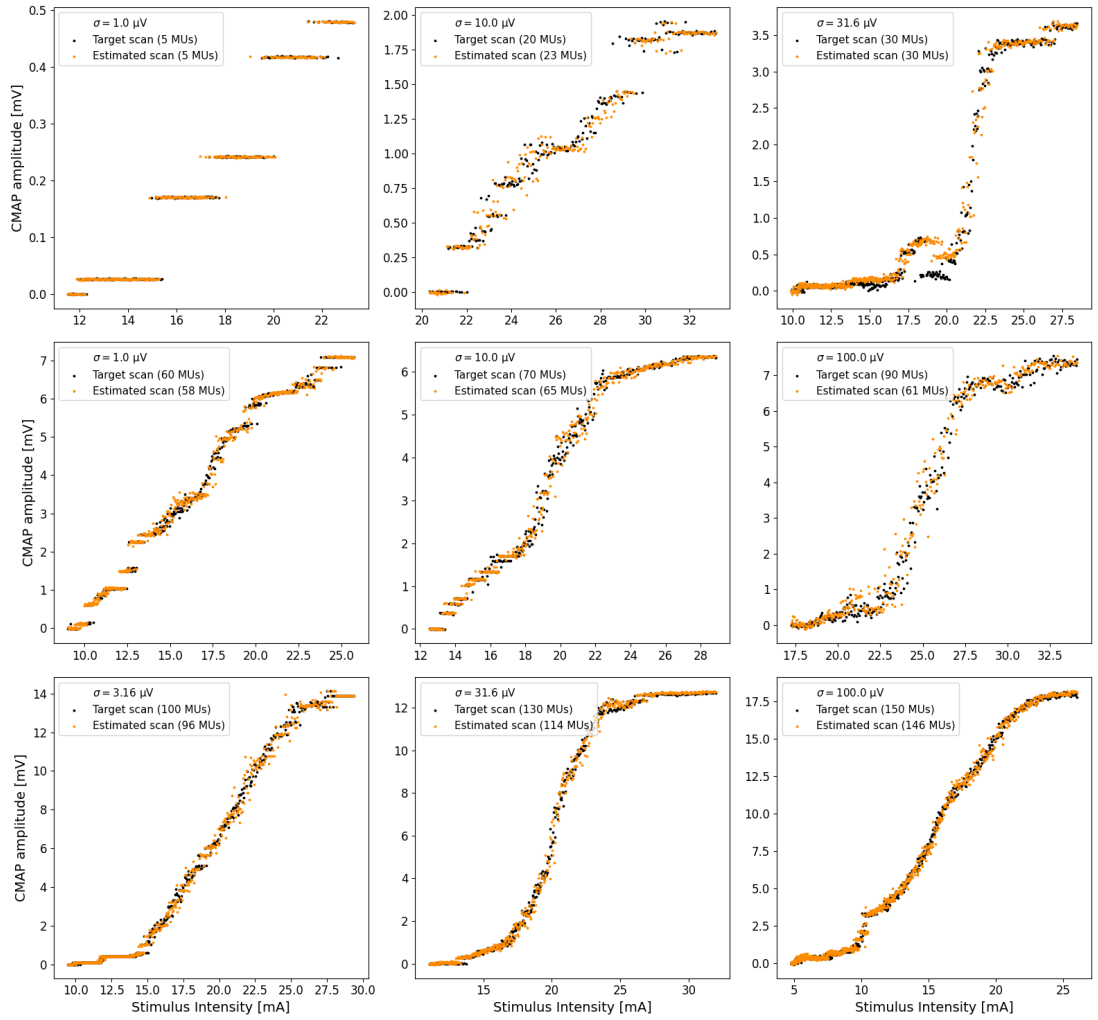


Figure 3.1: Nine examples of target and estimated CMAP scans by Wavolution chosen by their varying MU count and complexity. The top row depicts three scans with low MU numbers ($N_{\text{MU}} < 50$) with increasing complexity from left to right. Similarly, the middle and bottom row contain three scans with medium ($50 \leq N_{\text{MU}} < 100$) and high ($N_{\text{MU}} \geq 100$) MU numbers, respectively.

In Figure 3.1, nine fitted CMAP scans are shown over their target scan to evaluate Wavolution's ability to fit MU pool properties and perform MUNE. Wavolution was able to capture the global CMAP amplitude curve and the overlap between the estimated and target CMAP scan shows the model's ability to replicate the MU firing behavior. The top middle figure shows Wavolution's ability to include inverted SMUAPs as there is a decrease in CMAP amplitude in both the target and estimated CMAP scans at $31 - 32$ mA stimulus intensity. However, Wavolution had difficulty to fit the top right target scan, containing a high noise level and a large inverted SMUAP (at around 18 mA). The exclusion of such a large inverted SMUAP leads to an underestimation of the true amount of MUs in the target scan. For other noise levels of the same target scan, the CMAP scan of the estimated MU pool was fitted more accurately.

3.2. Comparison with MScanFit

3.2.1. Single CMAP scan

In Figure 3.2, fitted CMAP scans by both MScanFit and Wavolution are shown over their target scan. Both algorithms are able to capture the global CMAP amplitude curve and match CMAP amplitudes. The computation time of MScanFit was 73.2 seconds, whereas Wavolution performed the fitting procedure in 28.0 seconds. Although MScanFit MUNE has a higher accuracy, MScanFit is not able to fit a decrease in CMAP amplitude caused by phase cancellation effects as it is not able to include SMUAPs with a negative amplitude contribution. Wavolution, on the other hand, has the ability to include inverted SMUAPs and can thus include a decrease CMAP amplitude, which results in a more accurate fit in the higher stimulus intensity region of the scan (> 29 mA).

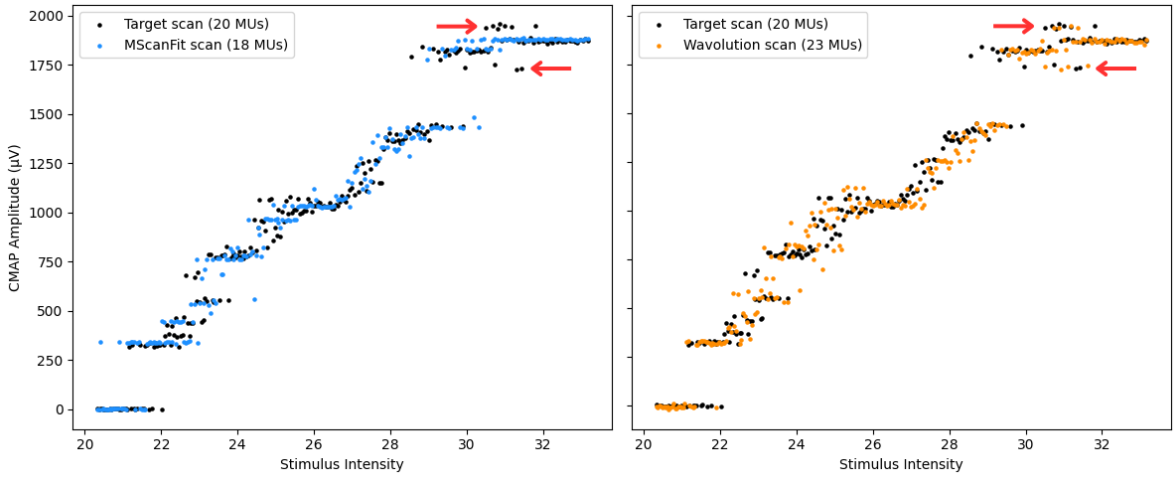


Figure 3.2: An example of target and estimated CMAP scans by MScanFit and Wavolution. At the locations of the red arrows, Wavolution is able to decrease CMAP amplitude by including inverted SMUAPs, whereas MScanFit assumes a monotonically increasing CMAP amplitude.

3.2.2. Accuracy

Taken over the entire range of MU numbers, the MUNE accuracy of Wavolution outperformed MScanFit at all noise levels, except for the low MU range at noise levels of $\sigma_4 = 31.6$ μV and $\sigma_5 = 100.0$ μV (Table I). A summary of the accuracies at a noise level of $\sigma_2 = 3.16$ μV is presented in Figure 3.3). At this noise level, the decrease in percentual MUNE discrepancy of Wavolution compared to MScanFit was 5.0%, 7.1%, and 21.6% at low ($N_{\text{MU}} < 50$), medium ($50 \leq N_{\text{MU}} < 100$) and high ($N_{\text{MU}} \geq 100$) MU numbers, respectively; over the full range the decrease in percentual MUNE discrepancy was 12.5% ($p < 0.001$). The largest increase in accuracy was reached at high MU numbers $N_{\text{MU}} \geq 100$, which was as expected. For all noise levels, similar results were found (Table I), except for the low MU range of noise levels of $\sigma_4 = 31.6$ μV and $\sigma_5 = 100.0$ μV, where no decrease in percentual MUNE discrepancy was found.

Table I: Average percentual MUNE discrepancy (E_{MUNE} (%)) for different MU number ranges and noise levels.

Noise Level	$E_{\text{MUNE, low}}$ (%)		$E_{\text{MUNE, medium}}$ (%)		$E_{\text{MUNE, high}}$ (%)		$E_{\text{MUNE, total}}$ (%)	
	Wavolution	MScanFit	Wavolution	MScanFit	Wavolution	MScanFit	Wavolution	MScanFit
$\sigma_1 = 1.0$ µV	5.8	10.0	9.4	20.0	14.3	32.7	10.0	24.0
$\sigma_2 = 3.16$ µV	5.0	10.0	9.3	17.1	12.5	33.6	9.1	21.7
$\sigma_3 = 10.0$ µV	10.0	15.0	10.0	18.2	14.5	32.3	11.0	22.9
$\sigma_4 = 31.6$ µV	21.3	17.5	13.1	22.1	18.2	33.0	17.9	25.0
$\sigma_5 = 100.0$ µV	20.0	20.0	12.9	32.3	18.6	38.8	17.9	34.0
Mean	12.4	14.5	11.0	21.9	15.6	34.1	13.2	25.5

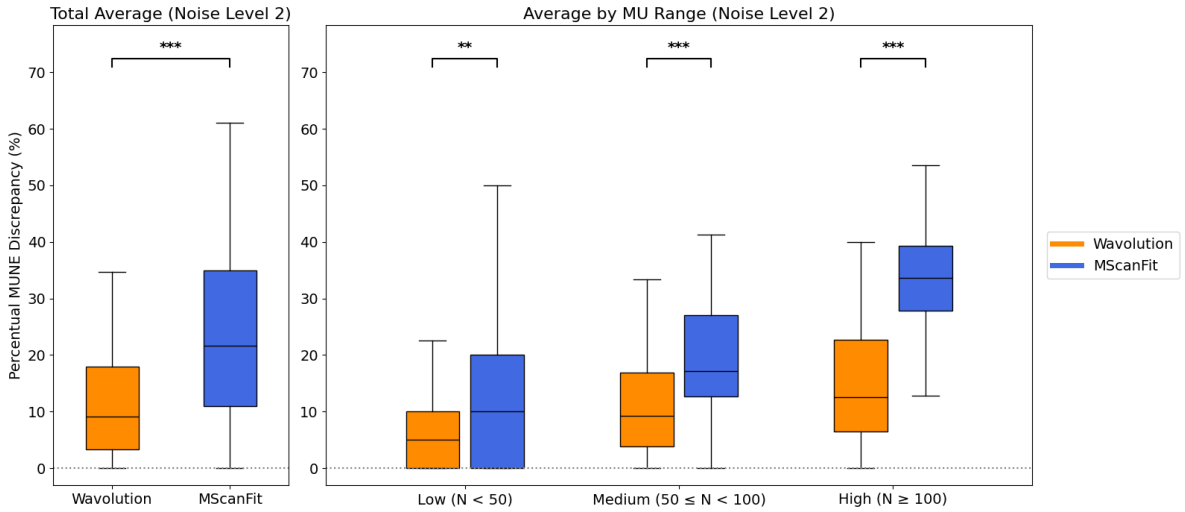


Figure 3.3: Average percentual MUNE discrepancy of Wavolution and MScanFit for a noise level of $\sigma_2 = 3.16$ µV. (A) shows the average percentual MUNE discrepancy for a total over 192 scans. (B) shows the average percentual MUNE discrepancy for low (60 scans), medium (60 scans) and high (72 scans) MU range. **Adjusted $p < 0.01$; ***adjusted $p < 0.001$.

3.2.3. Computation time

Taken over the entire range of MU numbers, the computation time of Wavolution was lower than that of MScanFit at all noise levels (Table II). A summary of the computation times at a noise level of $\sigma_2 = 3.16$ µV is presented in Figure 3.4. At this noise level, the decrease in computation time of Wavolution with respect to MScanFit was 57.1 s, 27.9 s, and 24.1 s at low, medium and high MU numbers, respectively; over the full range the decrease in computation time was 31.5 s ($p < 0.001$). For all other noise levels, similar results were found (Table II), except for the high MU range at a noise level of $\sigma_5 = 100.0$ µV, where there was no significant decrease in computation time. Results for the other noise levels ($\sigma_1 = 1.0$ µV, $\sigma_3 = 10.0$ µV, $\sigma_4 = 31.6$ µV, and $\sigma_5 = 100$ µV) are included in Appendix D.

Table II: Average computation time (T_{comp} (s)) for different MU number ranges and noise levels.

Noise Level	$T_{\text{comp, low}}$ (s)		$T_{\text{comp, medium}}$ (s)		$T_{\text{comp, high}}$ (s)		$T_{\text{comp, total}}$ (s)	
	Wavolution	MScanFit	Wavolution	MScanFit	Wavolution	MScanFit	Wavolution	MScanFit
$\sigma_1 = 1.0$ µV	36.1	89.4	64.4	89.1	76.4	101.4	62.9	94.5
$\sigma_2 = 3.16$ µV	35.9	93.0	64.2	92.1	78.5	102.6	64.2	95.7
$\sigma_3 = 10.0$ µV	35.5	87.6	64.7	93.0	75.1	103.5	63.1	95.7
$\sigma_4 = 31.6$ µV	39.2	81.9	70.0	85.5	83.9	101.1	67.3	90.6
$\sigma_5 = 100.0$ µV	46.4	76.5	84.2	89.1	99.4	100.8	81.4	87.9
Mean	38.6	85.7	69.5	89.8	82.7	101.9	67.8	92.9

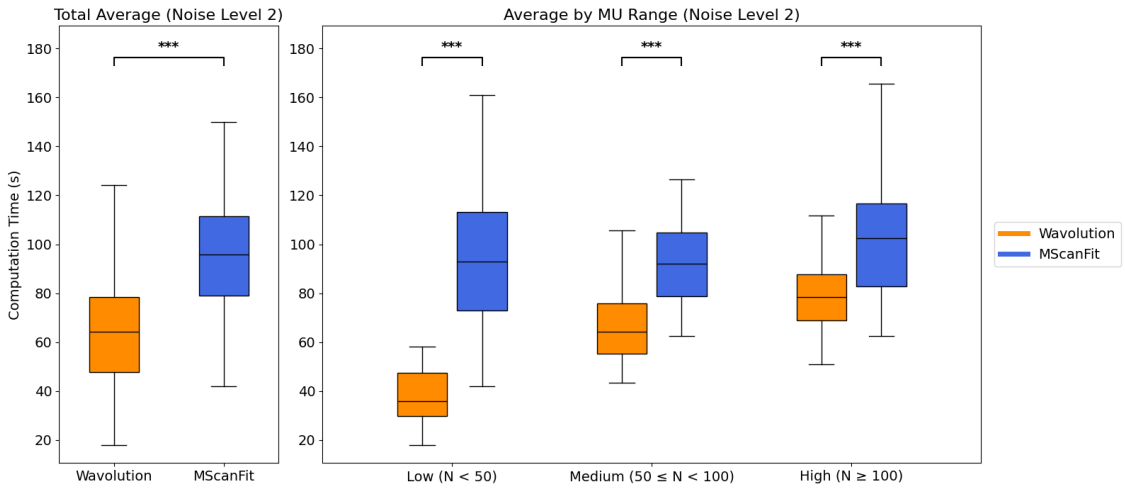


Figure 3.4: Average computation time of Wavolution and MScanFit for a noise level of $\sigma_2 = 3.16$ µV. (A) shows the average computation time for a total over 192 scans. (B) shows the average computation time for low (60 scans), medium (60 scans) and high (72 scans) MU range. ***Adjusted $p < 0.001$.

3.2.4. MU pool properties

Mean SMUAP size

Table III: Mean SMUAP size discrepancy ($E_{\text{SMUAP, mean}}$ (µV)) for different MU number ranges and noise levels.

Noise Level	$E_{\text{SMUAP, low}}$ (µV)		$E_{\text{SMUAP, mid}}$ (µV)		$E_{\text{SMUAP, high}}$ (µV)		$E_{\text{SMUAP, total}}$ (µV)	
	Wavolution	MScanFit	Wavolution	MScanFit	Wavolution	MScanFit	Wavolution	MScanFit
$\sigma_1 = 1.0$ µV	15.5	22.6	23.2	18.5	24.7	36.1	21.4	25.7
$\sigma_2 = 3.16$ µV	16.0	20.2	23.5	11.9	24.2	39.0	21.7	24.4
$\sigma_3 = 10.0$ µV	17.2	30.6	27.3	18.2	25.5	32.3	24.2	28.3
$\sigma_4 = 31.6$ µV	40.9	35.2	32.5	21.8	23.3	39.4	31.3	29.0
$\sigma_5 = 100.0$ µV	44.5	39.0	29.7	33.7	20.8	46.9	29.7	42.5
Mean	26.8	29.5	27.3	20.8	23.7	38.8	25.7	30.0

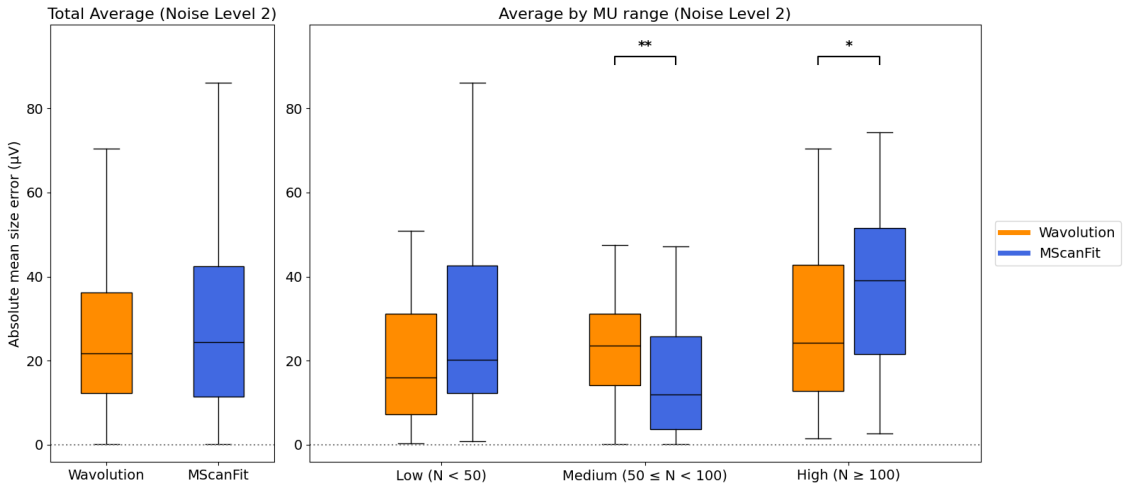


Figure 3.5: Average mean SMUAP size error in µV of Wavolution and MScanFit for a noise level of $\sigma_2 = 3.16$ µV. (A) shows the average mean SMUAP size error for a total over 192 scans. (B) shows the average mean SMUAP size error for low (60 scans), medium (60 scans) and high (72 scans) MU range. *Adjusted $p < 0.05$; **adjusted $p < 0.01$.

A summary of the mean SMUAP size errors of Wavolution and MScanFit is shown in Table II. A summary of the mean SMUAP size errors at a noise level of $\sigma_2 = 3.16$ µV is presented in Figure 3.5. Especially at the high MU range, the error in mean SMUAP size is significantly smaller for Wavolution than for MScanFit, whereas for the medium MU range, MScanFit obtained lower errors. For all other noise levels, similar results were found (Table III), except for the low MU range at noise levels of $\sigma_4 = 31.6$ µV and $\sigma_5 = 100.0$ µV, and the medium MU range at a noise level of $\sigma_5 = 100.0$ µV. Results for the other noise levels ($\sigma_1 = 1.0$ µV, $\sigma_3 = 10.0$ µV, $\sigma_4 = 31.6$ µV, and $\sigma_5 = 100$ µV) are included in Appendix D.

Reduction in maximum CMAP amplitude

As a result of temporal dispersion and phase cancellation effects, the maximum CMAP amplitude of a CMAP scan is generally lower than the summed absolute amplitude of all contributing individual SMUAPs in the MU pool that is used to generate a CMAP scan. This reduction in maximum CMAP amplitude was quantified as:

$$1 - \frac{\mathrm{CMAP}_{\max}}{\sum_i |A_{\mathrm{MU}\,i}|} \tag{3.1}$$

Figure 3.6 shows the absolute difference in this reduction between estimated and target MU pools at a noise level of $\sigma_2 = 3.16$ µV. Ideally, the difference is 0%, meaning the estimated MU pool perfectly captures phase cancellation and temporal dispersion.

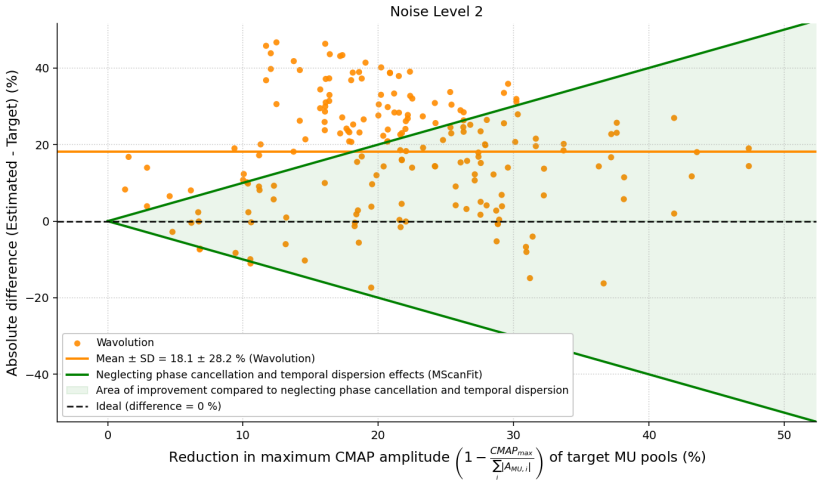


Figure 3.6: The absolute difference in reduction in maximum CMAP amplitude between the estimated and target MU pools for a noise level of $\sigma_2 = 3.16$ µV. Ideally, the absolute difference should be 0%. Wavolution generally overestimates the effects of temporal dispersion and phase cancellation, while MScanFit completely neglects the presence of temporal dispersion and phase cancellation.

As MScanFit does not account for phase cancellation and temporal dispersion effects, it always underestimates these effects when they are present in CMAP scans. Wavolution, however, has the ability to quantify these effects in the estimated MU pools. Especially when interference between SMUAPs is strong ($> 25\%$ reduction in maximum CMAP amplitude), Wavolution more accurately captures these effects in the estimated MU pool, which indicates that it more accurately estimates the sizes of individual SMUAPs, particularly those with negative phase contributions. This is consistent with Figure 3.5, where Wavolution more accurately estimates the mean SMUAP size in larger MU pools, which naturally contain more phase cancellation and temporal dispersion effects due to the higher amount of interference between SMUAP waveforms. When the reduction in maximum CMAP amplitude in the target MU pools is smaller ($< 25\%$), however, Wavolution overestimates the effects of phase cancellation and temporal dispersion, which is also reflected in Figure 3.5. For a noise level of $\sigma_2 = 3.16$ µV, Wavolution systematically overestimates the reduction in maximum CMAP amplitude with 18.1 ± 28.2 % (mean + SD), indicating that Wavolution exaggerates the effects of temporal dispersion and phase cancellation in the estimated MU pools. However, when the effects of phase cancellation and temporal dispersion are larger, this overestimation outperforms MScanFit, which cannot capture these effects at all.

4

Discussion

In this study, we introduced a novel algorithm that performs MUNE on CMAP scans. Wavolution outperforms MScanFit in terms of accuracy and computational efficiency, with a reduction of 12.3% in overall percentual MUNE discrepancy and 25.1 s in overall computation time over 920 CMAP scans. Furthermore, Wavolution makes use of actual SMUAP waveforms and is able to simulate inverted SMUAP waveforms, which can contribute to a higher MUNE accuracy in CMAP scans with high MU counts, suggesting that the algorithm has improved utilization of the available information.

4.1. Key findings

In comparison to the reference outcome, MScanFit, Wavolution has substantially higher MUNE accuracy, particularly in scans with a larger number of MUs. Current CMAP scan-based MUNE methods do not account for the presence of inverted SMUAPs in CMAP responses, assuming all MUs contribute positively to the CMAP amplitude. The incorporation of inverted SMUAPs in Wavolution more closely reflects real physiological behavior and can be an explanation for the underestimation of MScanFit in CMAP scans with larger number of MUs. Furthermore, while MScanFit relies on a preliminary estimate of the MU count for its fitting process, Wavolution covers the possibility that every CMAP scan can contain up to 200 MUs, reducing potential operator and estimation bias. A Windows application was developed with the capability to bulk process CMAP scans, where manual operation time is minimized (Appendix E). Together with the significantly reduced computation time, Wavolution can decrease the total operation time for bulk processing of CMAP scans in large-scale ALS clinical trials.

4.2. Impact of using SMUAP waveforms

To the best of our knowledge, current CMAP scan-based MUNE methods do not incorporate SMUAPs in the fitting process. Instead, they assign numerical values to MU properties, such as amplitude, threshold, and relative spread. Although computationally efficient, this approach limits these methods' ability to fully capture complex physiological effects such as phase cancellation and temporal dispersion, both of which strongly influence the shape and amplitude of the CMAP. Also an AI–based approach to CMAP scan–based MUNE is, at present, unlikely to be adopted in clinical practice. AI models generally behave as "black box" models, which limit the transparency between raw CMAP recordings and MUNE. Additionally, AI models rely on simulated training data, which can introduce modeling assumptions and can neglect phenomena present in experimental scans, which can limit interpretability and clinical trust. A Convolutional Neural Network approach to MUNE from CMAP scans was proposed by Junjun *et al.* [22]. While the method demonstrated high accuracy and short execution time, the approach is not able to capture waveform-level interactions such as phase cancellation and temporal dispersion, both of which strongly affect CMAP morphology.

Wavolution addresses these limitations by modeling the CMAP scan as successive, superimposed SMUAP waveforms, combining them into a CMAP using highly efficient matrix multiplications. As such, Wavolution allows for complex interactions between MUs without adding much computational cost. Additionally, the use of SMUAP waveforms makes Wavolution physiologically transparent: it enables quality control, such as inspection of individual simulated CMAP waveforms, and facilitates reasoning

about phenomena such as baseline drift, temporal dispersion, and phase cancellation. Furthermore, waveform-based modeling allows future iterations of the algorithm to incorporate other SMUAP properties, such as area, duration, distal motor latency, and shape.

Currently, Wavolution uses only five fundamental SMUAP waveforms of the APB muscle extracted from ALS patient recordings in its database for simulating CMAP scans. The same fundamental waveforms have also been used to simulate the target CMAP scans, which could in theory introduce bias in favor of Wavolution's performance. In practice, however, this effect is expected to be small, as each MU is assigned a random waveform which is extensively transformed in distal motor latency, temporal scaling, amplitude and phase. These transformations ensure that the use of the same fundamental SMUAP yields different interactions between the waveforms. As a result, target SMUAP sizes can be under- or overestimated, even when the compound CMAP amplitude is matched with high accuracy. Due to temporal dispersion and phase cancellation, two SMUAPs with different properties may produce similar contributions to CMAP amplitude (Appendix B). Moreover, Wavolution systematically overestimates the reduction in maximum CMAP amplitude and the proportion of inverted SMUAPs in the target MU pools. This effect may be partly explained by alternation, where thresholds of successive MUs of the target MU pool are estimated incorrectly. In such cases, two small target SMUAPs can be estimated as one larger SMUAP followed by a small inverted SMUAP, incorrectly increasing the proportion of inverted SMUAPs in the estimated MU pools. However, for all CMAP scan-based MUNE methods, such ambiguities are inherent when making use of solely baseline-to-peak CMAP amplitudes, as these only represent limited information present in the CMAP recordings.

For future research, Wavolution could be extended to exploit more information contained in CMAP waveforms rather than only the baseline-to-peak CMAP amplitudes. Successive CMAP recordings can be subtracted to obtain isolated increments due to the contribution of newly recruited MUs. These increments can then be characterized by area, shape, duration, distal motor latency and phase structure and be used to mitigate ambiguities arising from temporal dispersion, phase cancellation, and alternation. Other transformations such as time shifting, and time warping could be used to fit CMAP waveforms rather than the CMAP amplitudes. Together with physiological constraints as bounded SMUAP duration, realistic distal motor latencies and limited phase counts, controlled reshaping of simulated SMUAPs can result in a more accurate estimation of individual MU properties. AI methods could assist Wavolution in this reshaping of simulated SMUAPs by optimizing the SMUAP parameters to match CMAP waveforms. The use of AI should, however, be limited to this optimization process only to preserve the traceability of individual SMUAP contributions and to ensure the transparency of Wavolution's fitting process. Furthermore, Wavolution's dataset of fundamental SMUAPs can easily be expanded with more diverse SMUAPs, as well as with SMUAPs of other muscles, provided that isolated SMUAP recordings are available. Using more diverse fundamental SMUAPs can minimize the amount of transformations required to fit SMUAPs to CMAP waveforms present in a target CMAP scan. Furthermore, expanding the dataset of SMUAPs beyond solely SMUAPs of the APB muscle enables Wavolution to be applied to CMAP scans of different muscles. For instance, simulating scans of the tibialis anterior (TA) would only require the inclusion of isolated, fundamental SMUAPs recorded from that muscle.

4.3. Feasibility for clinical trials

A CMAP scan–based motor unit number estimate (MUNE) is an attractive outcome measure for clinical trials. It directly reflects lower motor neuron loss which precedes the loss of clinical function, thereby increasing sensitivity to treatment effects [23, 24]. The Windows-based application of Wavolution makes this approach globally accessible, requiring only stimulus–response recordings that can be obtained with relatively simple EMG hardware. This broad compatibility facilitates application in both research and clinical trial settings worldwide.

Wavolution is significantly faster than MScanFit due to: (1) its efficient generation of CMAP waveforms due to matrix multiplication; and (2) the use of modern computing strategies, such as parallel processing. This significant improvement in computation time can be used to further improve the accuracy of the model, but also decreases operator workload. MScanFit requires manual preparation of numerous individual CMAP scans by selecting the pre- and post-scan limits in separate files. Wavolution´s feature to complete all manual preparation at the start of a session for all CMAP scans reduces the total operator time and minimizes practical hurdles of using CMAP scan-based MUNE as an outcome measure in large-scale clinical trials.

At present, Wavolution has only been validated using simulated CMAP scan data. While these simulations were designed to mimic real physiological responses as closely as possible, they do not fully capture the variability and complexity present in human EMG recordings. Factors such as unpredictable noise, baseline drift, or electrode placement differences could affect performance when applied to real data – common limitations for every CMAP scan-based MUNE technique. Therefore, one of the key next steps is to validate the algorithm using recordings from human subjects. As MUNE provides only an estimate rather than an absolute ground truth, it is difficult to determine whether one method is better than another. However, Wavolution demonstrates higher accuracy in CMAP scans with larger MU populations, making it more sensitive in disease stages where other approaches typically lose reliability. This enhanced sensitivity could extend the clinical utility of CMAP scan–based MUNE and improve the statistical power in clinical trials. With further development and validation, Wavolution has the potential to become a valuable tool in the early monitoring stages of patients in ALS clinical trials.

4.4. Signal quality

Research on the effects of baseline noise on MUNE values is limited, even though baseline noise is a critical parameter in all CMAP scan-based MUNE methods [12, 25, 26]. Whereas MScanFit explicitly estimates the noise level directly from the CMAP scan, other CMAP scan-based MUNE methods instead rely on various predefined or assumed values for the noise present in the recordings. The noise level is often used to set the minimum size for MUs. In cases of high noise levels, this can mask the presence of small MUs and in turn affect MUNE values. This effect is stronger in CMAP recordings with high MU counts, as there is a strong inverse relationship between MUNE and mean SMUAP size [27]. Reducing noise in CMAP recordings could therefore improve the accuracy of CMAP scan-based MUNE methods, especially in early stages of disease progression, where muscles are still largely intact.

The noise estimate of Wavolution in target CMAP scans is not always accurate, as it relies on the pre- and post-scan limits, which only provide partial information of the actual noise present in the whole scan. A deviation in noise estimation can introduce high discrepancies between MUNE and the true amount of MUs present in a scan, as the estimated noise level plays a significant role in various functions of the algorithm, such as the amplitude density function. If the noise level is *underestimated*, low-amplitude responses are more likely to be interpreted as contributions of small MUs leading to an overestimation of MUNE values. Conversely, if the noise level is *overestimated*, low-amplitude changes are masked by spread in the amplitude density function, which can lead to underestimation of the MU count. As such, high accuracy in noise estimation is critical for the performance of Wavolution.

As MUNE accuracy decreases with increasing noise levels, an improvement to Wavolution (and other CMAP scan-based MUNE methods) would be to pre-process the target scan before it is analyzed for the fitting process by detrending and denoising the individual CMAP recordings. Detrending is useful when the EMG recordings contain baseline drift, which 'masks' the actual amplitude of the CMAP measurements (Figure F.1). Detrending 'removes' this baseline drift from each measurement such that the CMAP scans represent the actual EMG recordings from baseline to peak. Secondly, an EMG signal of a CMAP recording to a stimulus can be seen as a discrete, one-dimensional wavelet signal that contains noise. By using wavelet decomposition, coefficients that separate the signal into low- and high-frequency bands can remove signal noise (Appendix F). Therefore, wavelet denoising can decrease the overall baseline noise in CMAP recordings and thus improve the accuracy of CMAP scan-based MUNE methods. Future iterations of Wavolution could include a pre-processing module.

4.5. Conclusion

The main goal of this study was to improve accuracy and computation time compared to MScanFit. The proposed Wavolution algorithm achieves both aims: it lowers overall percentual MUNE discrepancy by 12.3% and reduces average runtime by 25.1 s per scan relative to MScanFit. Wavolution achieves these improvements by combining a waveform-based representation of CMAP responses with efficient population-based optimization. By superimposing SMUAPs, including inverted SMUAPs to represent temporal dispersion and phase cancellation, a higher MUNE accuracy is reached, especially at higher MU counts, where current MUNE methods typically lose sensitivity. Together with efficient, population-based optimization that updates all units per generation without relying on a preliminary estimate, Wavolution is able to search for the best solution in a large range of possible MUNE values, introducing robustness against local minima and ceiling effects. Wavolution is compatible with

standard EMG hardware and usable as a Windows application, making it feasible for worldwide clinical trial use. As such, Wavolution can be used to produce disease progression biomarkers for multi-center ALS studies, where sensitivity to early motor neuron loss, time-efficiency, and accessibility are critical.

References

[1] M. A. van Es *et al.*, "Amyotrophic lateral sclerosis," *The Lancet*, vol. 390, no. 10107, pp. 2084–2098, Nov. 2017. DOI: 10.1016/S0140-6736(17)31287-4.

[2] O. Hardiman *et al.*, "Amyotrophic lateral sclerosis," *Nature Reviews Disease Primers*, vol. 3, no. 1, Oct. 2017. DOI: 10.1038/nrdp.2017.71.

[3] H. Mitsumoto, B. R. Brooks, and V. Silani, "Clinical trials in amyotrophic lateral sclerosis: Why so many negative trials and how can trials be improved?" *The Lancet Neurology*, vol. 13, no. 11, pp. 1127–1138, Nov. 2014. DOI: 10.1016/S1474-4422(14)70129-2.

[4] R. G. Miller, J. D. Mitchell, and D. H. Moore, "Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND)," *Cochrane Database of Systematic Reviews*, 2012. DOI: 10.1002/14651858.CD001447.pub3.

[5] T. M. Miller *et al.*, "Trial of Antisense Oligonucleotide Tofersen for SOD1 ALS," *New England Journal of Medicine*, vol. 387, no. 12, pp. 1099–1110, Sep. 2022. DOI: 10.1056/NEJMoa2204705.

[6] C. Shin-Yi Lin *et al.*, "Neurophysiological and imaging biomarkers of lower motor neuron dysfunction in motor neuron diseases/amyotrophic lateral sclerosis: IFCN handbook chapter.," *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, vol. 162, pp. 91–120, Jun. 2024. DOI: 10.1016/j.clinph.2024.03.015.

[7] R. P. Van Eijk *et al.*, "An old friend who has overstayed their welcome: The ALSFRS-R total score as primary endpoint for ALS clinical trials," *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, vol. 22, no. 3-4, pp. 300–307, Apr. 2021. DOI: 10.1080/21678421.2021.1879865.

[8] C. N. Fournier, "Considerations for Amyotrophic Lateral Sclerosis (ALS) Clinical Trial Design," *Neurotherapeutics*, vol. 19, no. 4, pp. 1180–1192, Jul. 2022. DOI: 10.1007/s13311-022-01271-2.

[9] D. J. L. Stikvoort García, B. Sleutjes, L. J. van Schelven, H. S. Goedee, and L. H. van den Berg, "Diagnostic accuracy of nerve excitability and compound muscle action potential scan derived biomarkers in amyotrophic lateral sclerosis," *European Journal of Neurology*, vol. 30, no. 10, pp. 3068–3078, 2023. DOI: 10.1111/ene.15954.

[10] C. Wong *et al.*, "Clinical trials in amyotrophic lateral sclerosis: A systematic review and perspective," *Brain Communications*, vol. 3, no. 4, fcab242, Oct. 2021. DOI: 10.1093/braincomms/fcab242.

[11] J. R. Daube, "Estimating the Number of Motor Units in a Muscle," *Journal of Clinical Neurophysiology*, vol. 12, no. 6, p. 585, Nov. 1995.

[12] H. Bostock, "Estimating motor unit numbers from a CMAP scan," *Muscle & Nerve*, vol. 53, no. 6, pp. 889–896, 2016. DOI: 10.1002/mus.24945.

[13] A. B. Jacobsen, H. Bostock, and H. Tankisi, "Following disease progression in motor neuron disorders with 3 motor unit number estimation methods.," *Muscle & nerve*, vol. 59, no. 1, pp. 82–87, Jan. 2019. DOI: 10.1002/mus.26304.

[14] B. T. H. M. Sleutjes *et al.*, "Assessment of motor unit loss in patients with spinal muscular atrophy.," *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, vol. 131, no. 6, pp. 1280–1286, Jun. 2020. DOI: 10.1016/j.clinph.2020.01.018.

[15] R. S. Kristensen *et al.*, "MScanFit motor unit number estimation (MScan) and muscle velocity recovery cycle recordings in amyotrophic lateral sclerosis patients.," *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, vol. 130, no. 8, pp. 1280–1288, Aug. 2019. DOI: 10.1016/j.clinph.2019.04.713.

[16] S. Willemse *et al.*, “Lithium carbonate in amyotrophic lateral sclerosis patients homozygous for the C-allele at SNP rs12608932 in UNC13A: Protocol for a confirmatory, randomized, group-sequential, event-driven, double-blind, placebo-controlled trial,” *TRIALS*, vol. 23, no. 1, Dec. 2022. DOI: 10.1186/s13063-022-06906-5.

[17] argenx, “A Phase 2a, Double-Blinded, Randomized, Placebo-Controlled, and Active-Treatment Extension Study to Assess the Safety, Tolerability, Efficacy, Pharmacokinetics, and Immunogenicity of ARGX-119 in Participants with Amyotrophic Lateral Sclerosis,” clinicaltrials.gov, Clinical trial registration NCT06441682, Mar. 2025.

[18] J. H. Blok, A. Ruitenberg, E. M. Maathuis, and G. H. Visser, “The electrophysiological muscle scan,” *Muscle & Nerve*, vol. 36, no. 4, pp. 436–446, 2007. DOI: 10.1002/mus.20838.

[19] B. T. H. M. Sleutjes *et al.*, “Impact of stimulus duration on motor unit thresholds and alternation in compound muscle action potential scans.,” *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, vol. 132, no. 2, pp. 323–331, Feb. 2021. DOI: 10.1016/j.clinph.2020.10.026.

[20] M. de Carvalho, P. Barkhaus, S. Nandedkar, and M. Swash, “Motor unit number estimation (MUNE): Where are we now?” *CLINICAL NEUROPHYSIOLOGY*, vol. 129, no. 8, pp. 1507–1516, Aug. 2018. DOI: 10.1016/j.clinph.2018.04.748.

[21] B. T. H. M. Sleutjes, D. J. L. Stikvoort García, P. A. van Doorn, H. S. Goedee, and L. H. van den Berg, “Simulating progressive motor neuron degeneration and collateral reinnervation in motor neuron diseases using a dynamic muscle model based on human single motor unit recordings.,” *Journal of neural engineering*, vol. 20, no. 5, Oct. 2023. DOI: 10.1088/1741-2552/acfe9d.

[22] C. Junjun, Z. Li, L. Wu, Z. Lu, M. Chen, and P. Zhou, “Motor unit number estimation based on convolutional neural network,” *Journal of Neural Engineering*, 2025. DOI: 10.1088/1741-2552/ae01da.

[23] V. Vacchiano *et al.*, “Motor unit number estimation via MScanFit MUNE in spinal muscular atrophy,” *Muscle and Nerve*, vol. 70, no. 1, pp. 71–81, 2024. DOI: 10.1002/mus.28091.

[24] D. M. Sørensen *et al.*, “Estimating motor unit numbers from a CMAP scan: Repeatability study on three muscles at 15 centres.,” *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, vol. 151, pp. 92–99, Jul. 2023. DOI: 10.1016/j.clinph.2023.04.008.

[25] R. Henderson, P. Ridall, N. Hutchinson, A. Pettitt, and P. Mccombe, “Bayesian statistical MUNE method,” *Muscle & nerve*, vol. 36, pp. 206–13, Aug. 2007. DOI: 10.1002/mus.20805.

[26] S. D. Nandedkar, P. E. Barkhaus, and E. V. Stålberg, “Analysis of the compound muscle action potential scan: Step index (STEPIX) and amplitude index (AMPIX).,” *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, vol. 139, pp. 119–127, Jul. 2022. DOI: 10.1016/j.clinph.2022.04.011.

[27] C. S. Klein, H. Liu, and Y. Xiong, “Estimation of the number of motor units in the human extensor digitorum brevis using MScanFit.,” *PloS one*, vol. 19, no. 4, e0302214, 2024. DOI: 10.1371/journal.pone.0302214.

[28] P. Masrori and P. Van Damme, “Amyotrophic lateral sclerosis: A clinical review,” *European Journal of Neurology*, vol. 27, no. 10, pp. 1918–1929, 2020. DOI: 10.1111/ene.14393.

[29] R. Chia, A. Chiò, and B. J. Traynor, “Novel genes associated with amyotrophic lateral sclerosis: Diagnostic and clinical implications,” *The Lancet Neurology*, vol. 17, no. 1, pp. 94–102, 2018. DOI: https://doi.org/10.1016/S1474-4422(17)30401-5.

[30] C. Maurel *et al.*, “Causative Genes in Amyotrophic Lateral Sclerosis and Protein Degradation Pathways: A Link to Neurodegeneration,” *Molecular Neurobiology*, vol. 55, no. 8, pp. 6480–6499, Aug. 2018. DOI: 10.1007/s12035-017-0856-0.

[31] B. Brooks, R. Miller, M. Swash, and T. Munsat, “El escorial revisited: Revised criteria for the diagnosis of amyotrophic lateral sclerosis,” *Amyotrophic Lateral Sclerosis and Other Motor Neuron Disorders*, vol. 1, pp. 293–299, 2000. DOI: 10.1080/146608200300079536.

[32] M. de Carvalho *et al.*, “Electrodiagnostic criteria for diagnosis of als,” *Clinical Neurophysiology*, vol. 119, no. 3, pp. 497–503, 2008. DOI: 10.1016/j.clinph.2007.09.143.

[33] B. Johnsen, “Diagnostic criteria for amyotrophic lateral sclerosis from el escorial to gold coast,” *Clinical Neurophysiology*, vol. 131, no. 8, pp. 1962–1963, 2020. DOI: https://doi.org/10.1016/j.clinph.2020.04.012.

[34] D. Yuan, S. Jiang, and R. Xu, “Clinical features and progress in diagnosis and treatment of amyotrophic lateral sclerosis,” *Annals of Medicine*, vol. 56, no. 1, Dec. 2024. DOI: 10.1080/07853890.2024.2399962.

[35] J. C. Davies *et al.*, “Limited value of serum neurofilament light chain in diagnosing amyotrophic lateral sclerosis,” *Brain Communications*, vol. 5, no. 3, fcad163, Jun. 2023. DOI: 10.1093/braincomms/fcad163.

[36] R. P. Van Eijk *et al.*, “Refining eligibility criteria for amyotrophic lateral sclerosis clinical trials,” *Neurology*, vol. 92, no. 5, Jan. 2019. DOI: 10.1212/WNL.0000000000006855.

[37] N. J. Thakore, B. R. Lapin, H. Mitsumoto, and P. R. O.-A. A. C. T. Consortium, “Early initiation of riluzole may improve absolute survival in amyotrophic lateral sclerosis,” *Muscle & Nerve*, vol. 66, no. 6, pp. 702–708, 2022. DOI: https://doi.org/10.1002/mus.27724.

[38] F. Franchignoni, J. Mandrioli, A. Giordano, and S. Ferro, “A further rasch study confirms that alsfrs-r does not conform to fundamental measurement requirements,” *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, vol. 16, no. 5-6, pp. 331–337, 2015. DOI: 10.3109/21678421.2015.1026829.

[39] R. S. Bedlack *et al.*, “How common are als plateaus and reversals?” *Neurology*, vol. 86, no. 9, pp. 808–812, 2016. DOI: 10.1212/WNL.0000000000002251.

[40] B. T. Sleutjes *et al.*, “Advancing disease monitoring of amyotrophic lateral sclerosis with the compound muscle action potential scan,” *Clinical Neurophysiology*, vol. 132, no. 12, pp. 3152–3159, Dec. 2021. DOI: 10.1016/j.clinph.2021.09.014.

[41] J. Larkindale *et al.*, “Cost of illness for neuromuscular diseases in the united states,” *Muscle & Nerve*, vol. 49, no. 3, pp. 431–438, 2014. DOI: https://doi.org/10.1002/mus.23942.

[42] A. J. McComas, P. R. W. Fawcett, M. J. Campbell, and R. E. P. Sica, “Electrophysiological estimation of the number of motor units within a human muscle,” *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 34, no. 2, pp. 121–131, 1971. DOI: 10.1136/jnnp.34.2.121.

[43] W. F. Brown and H. S. Milner-Brown, “Some electrical properties of motor units and their effects on the methods of estimating motor unit numbers.,” *Journal of Neurology, Neurosurgery, and Psychiatry*, vol. 39, no. 3, pp. 249–257, Mar. 1976.

[44] M. J. Strong, W. F. Brown, A. J. Hudson, and R. Snow, “Motor unit estimates in the biceps-brachialis in amyotrophic lateral sclerosis,” *Muscle & Nerve*, vol. 11, no. 5, pp. 415–422, 1988. DOI: https://doi.org/10.1002/mus.880110502.

[45] D. W. Stashuk, T. J. Doherty, A. Kassam, and W. F. Brown, “Motor unit number estimates based on the automated analysis of f-responses,” *Muscle & Nerve*, vol. 17, no. 8, pp. 881–890, 1994. DOI: https://doi.org/10.1002/mus.880170807.

[46] C. Neuwirth *et al.*, “Motor unit number index (munix): A novel neurophysiological marker for neuromuscular disorders; test–retest reliability in healthy volunteers,” *Clinical Neurophysiology*, vol. 122, no. 9, pp. 1867–1872, 2011. DOI: https://doi.org/10.1016/j.clinph.2011.02.017.

[47] J. P. van Dijk, J. H. Blok, B. G. Lapatki, I. N. van Schaik, M. J. Zwarts, and D. F. Stegeman, “Motor unit number estimation using high-density surface electromyography,” *Clinical Neurophysiology*, vol. 119, no. 1, pp. 33–42, 2008. DOI: https://doi.org/10.1016/j.clinph.2007.09.133.

[48] B. T. H. M. Sleutjes *et al.*, “CMAP scan discontinuities: Automated detection and relation to motor unit loss.,” *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, vol. 125, no. 2, pp. 388–395, Feb. 2014. DOI: 10.1016/j.clinph.2013.07.016.

[49] H. Bostock, A. B. Jacobsen, and H. Tankisi, “Motor unit number index and compound muscle action potential amplitude,” *Clinical Neurophysiology*, vol. 130, no. 9, pp. 1734–1740, Sep. 2019. DOI: 10.1016/j.clinph.2019.05.031.

[50] E. M. Maathuis, J. Drenthen, G. H. Visser, and J. H. Blok, “Reproducibility of the CMAP scan,” *Journal of Electromyography and Kinesiology*, vol. 21, no. 3, pp. 433–437, Jun. 2011. DOI: 10.1016/j.jelekin.2010.11.007.

[51] P. Barkhaus, S. Nandedkar, M. de Carvalho, M. Swash, and E. Stålberg, “Revisiting the compound muscle action potential (CMAP),” *CLINICAL NEUROPHYSIOLOGY PRACTICE*, vol. 9, pp. 176–200, 2024. DOI: 10.1016/j.cnp.2024.04.002.

[52] C. He, J. Xing, J. Li, Q. Yang, and R. Wang, “A New Wavelet Threshold Determination Method Considering Interscale Correlation in Signal Denoising,” *Mathematical Problems in Engineering*, vol. 2015, no. 1, p. 280 251, 2015. DOI: 10.1155/2015/280251.

[53] D. Donoho, “De-noising by soft-thresholding,” *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, May 1995. DOI: 10.1109/18.382009.

A

Background Information

Amyotrophic lateral sclerosis (ALS) is a multisystem neurodegenerative disorder characterized by the degeneration of central and peripheral motor neurons, causing progressive weakness, paralysis and eventually death [1, 2]. ALS is clinically and genetically heterogeneous, with various phenotypes and disease trajectories in terms of symptom spreading and progression rates. In most cases, the disease ultimately ends in respiratory failure with a median survival of approximately 3 years [28]. ALS can be classified as familial ALS (10% of the cases), which suggests an autosomal dominant genetic inheritance, or sporadic ALS (90% of the cases), which has no affected family members. In addition to the multiple underlying pathophysiological mechanisms, which are not fully understood, more than 40 ALS-related causative genes have been identified [29, 30]. This clinical and genetic heterogeneity creates various challenges in early diagnosis of the disease, as well as accurate tracking of the disease's progression in the early stages of the disease.

Diagnostic criteria of ALS have been revised multiple times over the years, from the revised El Escorial diagnostic criteria (rEEC) and Awaji criteria (AC), to the 'Gold Coast' criteria (GCC) [31–33]. While varying in specific criteria, the diagnostic process has consistently encompassed three fundamental steps: (1) a documented history that demonstrates progressive motor impairment after a period of normal motor function; (2) the presence of both upper and lower motor neuron signs in at least one body region, or lower motor neuron dysfunction in at least two body regions; (3) thorough investigations must be conducted to rule out any other potential disease process [33]. The clinical and genetic heterogeneity of ALS causes delays in the diagnosis, as accurately diagnosing ALS at the early stage of the disease remains a challenge [34]. Removing diagnostic doubt at the suspicion of the disease would be a major advance, allowing earlier and possibly more inclusive enrollment to clinical trials in ALS [35]. Inclusion in clinical trials is still a challenge, as the majority of patients with ALS is excluded from trial participation at diagnosis [36]. At the same time, early inclusion in clinical trials is crucial for the efficacy of the treatment, as early initiation of riluzole may improve absolute survival in ALS [37].

To examine whether a new drug has disease modifying properties, clinical trials need to identify a slowing in the rate of functional decline followed by a prolongation in survival. The progression and progression rate of the disease is most commonly monitored using the ALS functional rating revised scores (ALSFRS-R), which is a questionnaire that evaluates the severity of ALS, including respiratory function. However, there are various limitations with the ALSFRS-R as an outcome measure. A one-point change may indicate either a minor or significant functional shift, depending on the specific question and item, and thus, it does not serve as a consistent or measurable unit of functional change across the scale [38]. Furthermore, a study on the ALSFRS-R using data from a pooled clinical trial database found that 25% and 16% of patients in the placebo group exhibited no change in their ALSFRS-R score over six months and 12 months, respectively [39]. To summarize, the ALSFRS-R score has significant implications for clinical trials evaluating treatments designed to slow disease progression by a reduced ability and sensitivity to detect treatment effects and early alterations in disease progression rate [8, 40]. In addition to this, the inability to determine drug efficacy in a single, definitive and timely trial may lead to re-evaluation of the same investigational medicinal products in subsequent trials, which is inefficient in time, cost and patient resource [10]. ALS is an expensive disease, with annual costs per-patient exceeding $70{,}000$ dollars [41]. This calls for the need of a marker that is sensitive in the

early monitoring process in clinical trials, as treatment effects may have greater benefit if treatment is initiated earlier.

As loss of motor units (MUs) is the hallmark of ALS, obtaining estimates of the number of living MUs in a muscle is a sensible measure of the disease's progression. It would be ideal to have an actual measure of the number of functional MUs, but current methods are not able to provide this number [11]. Over the past decades, various methods have emerged for estimating this number based on electrophysiological methods – so-called motor unit number estimation (MUNE) methods. MUNE was first performed via manual incremented stimulation. Stimulation of a peripheral nerve evokes an electrical response, the muscle action potential, of the muscle (or group of muscles) that can be recorded. It is possible to recruit successive single MUs by carefully grading the strength of an electrical stimulus, leading to single MU action potential (SMUAP) amplitude; the response of the total population of units, also called the compound muscle action potential (CMAP), is then evoked by a maximal stimulus to the nerve [42]. The final estimate of the functional number of MUs was then defined as the CMAP amplitude divided by the average SMUAP. Since then, many methods have been introduced for MUNE, including multiple point stimulation (MPS) MUNE, the spike-triggered averaging technique, the F-response Method, MUNIX, High Density Surface EMG (HDSEMG) and CMAP-scan analysis methods, including Bayesian statistics, D50, and MScanFit [12, 25, 43–48].

MUNIX and CMAP-scan analysis methods are both convenient and automated, making them most suitable for inclusion in larger clinical trials. MUNIX, however, highly correlates with CMAP amplitude and is less sensitive for detecting early MU loss compared to CMAP-scanning methods, limiting the feasibility for clinical trials, especially in the early stages of ALS [49]. In addition, the method uses voluntary activation for the MUs instead of electrical stimulation. There is no consensus on which provides a better estimation of the response size. However, factors such as fatigue can introduce challenges when voluntary activation is required. Electrical stimulation is often less dependent on the participant, especially in severe disease cases. This reduces variability between operators and participants, making it more suitable for inclusion in clinical trials. Earlier MUNE methods, including the incremental and F-response method, are less suitable for clinical use, as improvements in reliability have been made over the years. Techniques that require skilled and trained operators with expertise, such as MPS, HD-sEMG and invasive methods, are less suitable for larger clinical trials due to their complex recording protocols and required hardware that is not everywhere accessible. Methods that require voluntary activation are less suitable for clinical trials as this can be challenging for patients with severe symptoms of ALS. Compared to MUNIX and statistical methods, CMAP-scan analysis methods are currently the most suitable, as they analyze the entire MU pool, are efficient, automated, and non-invasive. Furthermore, the method does not require highly skilled operators and can be performed on standard EMG equipment, making it a feasible method for clinical use. Compared to other MUNE methods, CMAP-scan analysis is a non-invasive, highly-reproducible, and less labor intensive way to estimate MUs controlling a muscle [40, 50], making it a promising technique for clinical application. However, as complexity increases with the number of MUs, the accuracy of the estimates reduce when muscles are largely intact. Therefore, especially in the early stages of ALS, accurate estimation of MUs is challenging, even though crucial for effective monitoring of the disease and treatment effects. Key challenges for accurate estimation of MUs include alternation of MU firing, phase cancellation effects, and baseline noise.

MScanFit, the most widely used CMAP-scanning method, has high efficiency in its processing, taking a few minutes to provide an estimate [12]. It covers amplitude, threshold, and relative spread of the threshold in the simulated scans, increasing the physiological reliability, but neglects waveform shapes, which can provide additional information. It has been evaluated with experimental data by comparing MUNE values of patients and healthy controls, represented by the Area under the Curve of a Receiver-operating characteristic curve (ROC). MScanFit showed very high AUC values when discriminating patients and healthy controls [15, 23]. Also for a longitudinal study, MScanFit MUNE showed to be the most promising candidate biomarker in terms of the required sample size for clinical trials [40]. As such, MScanFit is the current 'gold' standard as a MUNE technique. However, in these studies, baseline MUNE was already relatively low, suggesting that the population was already affected by MU loss. When MU numbers are high, MScanFit lacks precision and underestimates MUNE [12]. Given the ceiling effects of MScanFit, it is unclear whether these results are equally promising for clinical trials where the population is less affected by MU loss.

B

Simulating CMAP Scans

B.1. SMUAPs from ALS patient recordings

Five distinct SMUAPs were extracted from ALS patient recordings of the abductor pollis brevis (APB). These SMUAPs originated from MUs with a clear separated threshold and were the only MUs that were firing at certain stimulus intensities in the CMAP scan. As such, clear isolated SMUAPs could be extracted from a single CMAP recording of the total CMAP scan and could be used to serve as fundamental waveforms for the simulation of CMAP scans. These five distinct SMUAP waveforms are shown below in Figure B.1.

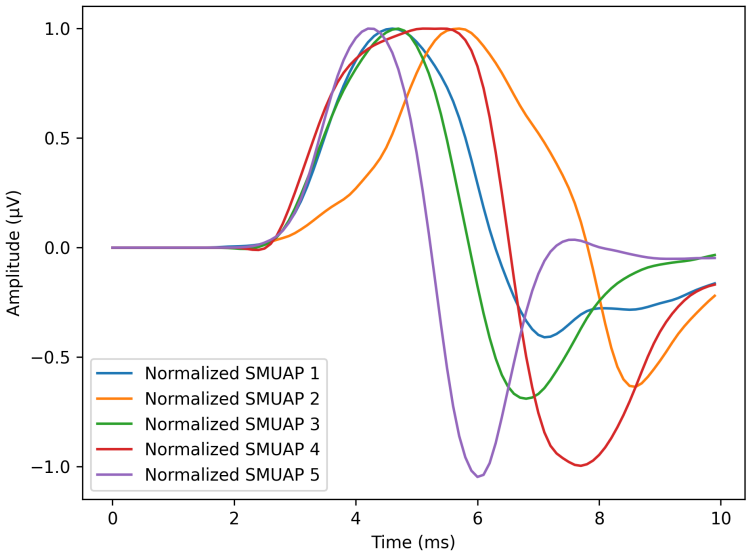
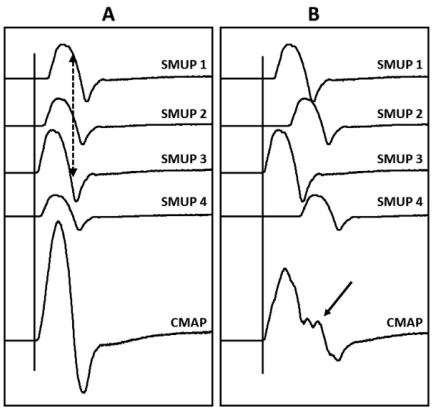


Figure B.1: Five distinct SMUAPs extracted from ALS patient recordings of the abductor pollis brevis (APB).

B.2. Temporal dispersion and phase cancellation

SMUAPs have various other properties, such as shape, duration, and distal motor latency, assigning actual SMUAPs to MUs replicates real world scenario's more closely. CMAP measurements contain superpositions of individual SMUAP waveforms (Figure B.2). The actual amplitudes of individual SMUAP waveforms can be masked by temporal dispersion, which is caused by distal motor latency and shape of the waveform. Additionally, CMAP recordings typically exhibit an initial negative deflection in their bi-phasic waveform. However, in cases of advanced reinnervation, CMAPs may demonstrate an initial positive deflection, according to Carvalho *et al.* [20]. This induces effects of phase cancellation, which can reach up to 40% decrease in maximum CMAP amplitude [21]. This subsequently impacts MUNE values estimated from CMAP scans.



C

SMUP	Amplitude (μV)	Area (μV x ms)	Duration (ms)
1	100	210	3.25
2	80	168	3.25
3	125	263	3.25
4	60	126	3.25
Total	365	767	

D

Measurement	CMAP (A)	CMAP (B)
Latency (ms)	3	4
Amplitude (μV)	339	210
Area (μV x ms)	729	594
Duration (ms)	3.65	6

Figure B.2: Individual MU potential and compound muscle action potential (CMAP) measurements are shown schematically in A and B. The individual SMUP measurements are shown in top table in C and the CMAP measurements in D. Due to temporal dispersion and phase cancellation, the total amplitudes and areas of the CMAP measurement are smaller than the sum of the individual SMUP measurements [51].

B.3. Simulating a CMAP scan with matrix multiplication

A MU pool of N MUs consists of $4N$ variables: N activation thresholds (t), N values for the relative spread of the threshold (ρ), N SMUAP waveforms of time-length T (with N amplitudes A), and N phase values (φ). These variables are used to generate two matrices: an SMUAP matrix ($N \times T$), consisting of the waveform amplitude values over time of each MUs, and a firing matrix ($N \times S$), consisting of the firing behavior of every MU at each stimulus intensity value S_i.

The SMUAP matrix contains the SMUAP waveform of every MU in the MU pool, with each row denoting the waveforms' phase and amplitude $\varphi \cdot A_i$ at time point T_i. An example of such a matrix is shown below, where each row and column represents an SMUAP and time point in the top figures of Figure B.3 .

$$M_{N\times T} = \begin{bmatrix} 0 & 0 & 4.3 & 15.3 & 174.3 & 211.4 & 51.2 & \cdots \\ 0 & 0 & 2.1 & -62.5 & -77.4 & -71.9 & -7.3 & \cdots \\ 0 & 0 & -3.4 & 86.9 & 192.6 & 203.2 & 141.6 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{B.1}$$

The activation threshold and relative spread of the activation threshold determine the firing behavior of MUs at each stimulus intensity. Figure B.4 shows the firing probability of a MU depending on the activation threshold (μ) and relative spread of the activation threshold, which is used to determine the standard deviation, σ, of the all-or-none response [19]:

$$\sigma = \frac{\mu \cdot \mathrm{RS}}{100} \tag{B.2}$$

where μ is the mean activation threshold and RS is the value of the relative spread of the activation threshold.

The probability of firing is determined by the accumulated probability density, indicated by the sigmoid curve in Figure B.4. By randomly sampling 0 (no activation) and 1 (activation) from this probability density function at each stimulus intensity value S_i, a firing matrix of size $N \times S$ is obtained for N MUs:

$$M_{N\times S} = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & \cdots \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{B.3}$$

Multiplying matrix of SMUAP values of MUs by a firing matrix results in a matrix of accumulated SMUAPs for every stimulus intensity represented in each column (shown in the left bottom figure of Figure B.3):

$$M^T_{\text{SMUAP }(N\times T)} \times M_{\text{Firing }(N\times S)} = M_{\text{CMAP }(T\times S)} \tag{B.4}$$

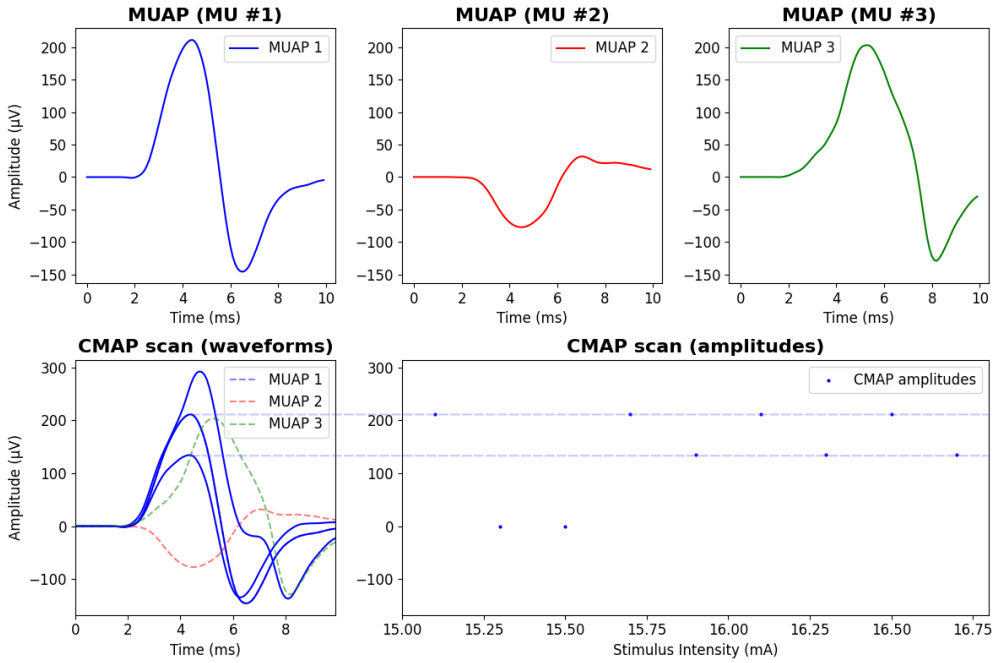


Figure B.3: Part of a CMAP scan with the three MUs presented in the top figures. The left bottom figure shows the accumulated SMUAPs, resulting in the accumulated amplitudes in the right bottom figure.

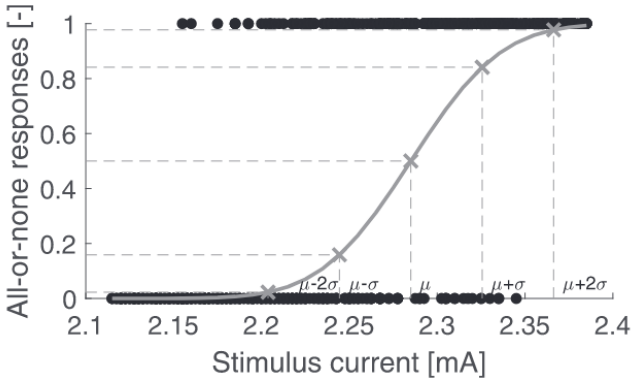


Figure B.4: The firing probability of a MU with the mean activation threshold, μ, and standard deviation, σ, depending on the stimulus intensity [19].

$$M_{T\times S} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 4.3 & 0 & 0 & 4.3 & 6.4 & 4.3 & 6.4 & 4.3 & 6.4 & \dots \\ 15.3 & 0 & 0 & 15.3 & -47.2 & 15.3 & -47.2 & 15.3 & -47.2 & \dots \\ 174.3 & 0 & 0 & 174.3 & 96.9 & 174.3 & 96.9 & 174.3 & 96.9 & \dots \\ 211.4 & 0 & 0 & 211.4 & 139.5 & 211.4 & 139.5 & 211.4 & 139.5 & \dots \\ 51.2 & 0 & 0 & 51.2 & 43.9 & 51.2 & 43.9 & 51.2 & 43.9 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{B.5}$$

For every stimulus intensity, the maximum amplitude (baseline-peak) of the CMAP waveform (indicated in blue) is extracted, leading to a CMAP scan of amplitude values (shown in the right bottom figure of Figure B.3). This matrix multiplication very high compuational efficiency: generating 1000 CMAP scans takes approximately 6.5 seconds.

C

Optimization Stage

C.1. Discrepancy score

The weights in the discrepancy score (Equation 2.3) are chosen such that they contribute approximately equal;y to the total error score. The weights are kept fixed throughout the optimization process. These weights are:

$$\begin{aligned} w_1 &: \quad 750 \\ w_2 &: \quad 2 \\ w_3 &: \quad 0.1 \\ w_4 &: \quad 1e5 \end{aligned}$$

C.2. Values for amplitude and threshold density distributions

The differences in amplitude and threshold density distributions between the simulated and target scans are used in the error score. The density distributions are generated by smoothing histograms of the occurrence of amplitudes and locations of amplitude differences, as this was > 100 times faster than directly generating a probability density function.

The histograms ($N_{\text{bins}} = 10 \times N_{\text{stimuli}}$) are smoothed using Gaussian smoothing. The spread for the amplitude density distribution is $0.5 \cdot \max(A_{\text{min}},\ A_{\text{min}} \cdot (1 + \log(\text{estimated noise}))$, where $A_{\text{min}} = 5$ μV, as this is considered to be the minimum amplitude. Due to a higher spread of data points with increasing noise, the estimated noise level is included as a multiplier. The spread for the the threshold density distribution is calculated according to the stimulus spread based on the average estimated value of the relative spread [19]. Each stimulus intensity has a width of relative spread:

$$w_i = 2 \cdot \frac{RS}{100} \cdot SI_i \tag{C.1}$$

where $RS = 1.65$. Consequently, the amount of stimulus samples within this range is computed for every stimulus intensity:

$$S_i = \sum \left[SI_i - \frac{w_i}{2}, SI_i + \frac{w_i}{2} \right] \tag{C.2}$$

The average of this width over the quantiles (25th, 50th, and 75th percentiles) is the threshold spread:

$$T_s = \frac{Q_1(S) + Q_2(S) + Q_3(S)}{3} \tag{C.3}$$

Additionally, the MUs of the top 10 MU pools of each population are updated according to the amplitude density distribution. The amplitude density distribution indicates where errors between the simulated scan and the target scan occur. Based on erroneous peaks in the amplitude distribution, simulated MUs are removed and added. For the identification of erroneous peaks, peaks of the density distributions of the simulated and target scan need to have a certain distance between them. The minimum distance in amplitude between a simulated and target peak is 5 μV, as this is considered to

be the minimum amplitude for a MU. Secondly, an amplitude peak is considered an erroneous peak if the mean stimulus intensity of simulated and target amplitude peaks has a minimum distance in stimulus intensity. This minimum distance is calculated with the threshold spread:

$$D = \frac{SI_{\text{max}} - SI_{\text{min}}}{n/T_s} \tag{C.4}$$

where SI_{min} and SI_{max} are the minimum and maximum stimulus intensity value respectively, n is the amount of stimulus points, and T_s is the threshold spread.

C.3. Additional figures of fitting procedure

The figures below correspond to important steps in the fitting procedure. Figure C.1 shows the various ways of possible merging of MU pairs based on their amplitudes or thresholds. Figure C.2 shows an example of the generation of a cross-over MU pool, based on two separate MU pools.

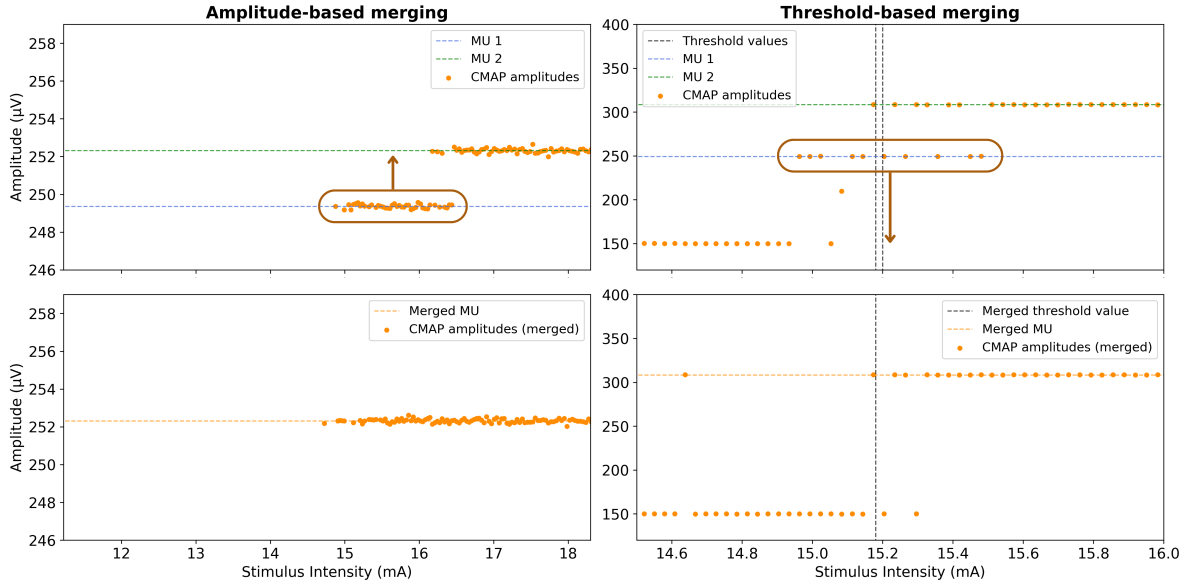


Figure C.1: Merging of MUs occurs if a MU amplitude is too small, or if the activation threshold of two adjacent MUs are smaller than the decremental size of stimuli.

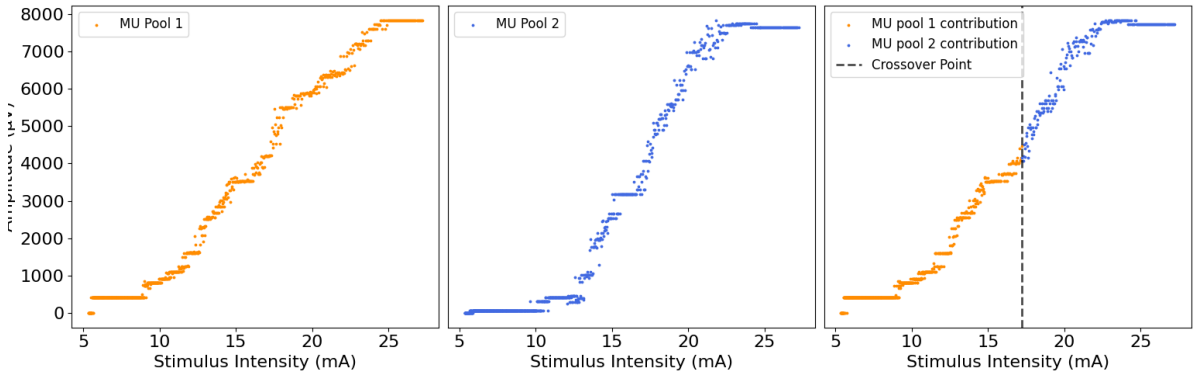


Figure C.2: An example of a cross-over MU pool consisting of the lowest threshold MUs of MU pool 1 and the highest threshold MUs of MU pool 2.

C.4. Alternatively explored optimization methods

A hybrid fitting model was tested, making use of the evolutionary characteristics in combination with global optimization methods. Multiple derivative-free optimization algorithms have been explored, such as Differential Evolution (DE), Bayesian optimization, Covariance Matrix Adaptation Evolution Strategy (CMA-ES) and Particle Swarm Optimization (PSO). The latter optimization method was found most suitable in terms of computation time and convergence.

Particle Swarm Optimization uses multiple particles in the parameter landscape to investigate search directions and minima. These particles have a cognitive component (e.g. self-exploratory component), a social component (e.g. a component that 'draws' them towards better performing particles), and a momentum component (which determines the 'velocity' and 'overshoot' at which a particle moves). By adjusting these parameters, particles may be more sensitive to individual exploration or to faster convergence at the risk of ending up in a local minimum. In the particular case of CMAP scan fitting, however, the parameter space grows quickly with increasing MU numbers as it scales with $M \times N$ for N MUs with M parameters. Additionally, due to the probabilistic firing of MUs, the error score of each scan of the same MU pool varies. This creates a very 'rugged' landscape for the parameter space, requiring a lot of particles and iterations to avoid local minima. It was found that these optimization methods therefore did not add enough value in terms of error minimization compared to the computation time needed for the algorithm. Therefore, it was decided not to use this optimization in the evolutionary process of FitWave.

D

Additional Results

D.1. MUNE accuracy

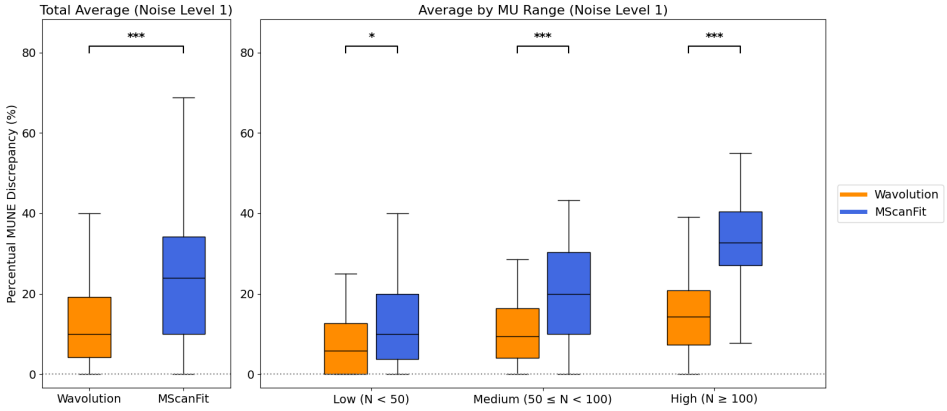


Figure D.1: Average percentual MUNE discrepancy of FitWave and MScanFit for a noise level of $\sigma_1 = 1.0$ µV. (A) shows the average percentual MUNE discrepancy for a total over 192 scans. (B) shows the average percentual MUNE discrepancy for low (60 scans), medium (60 scans) and high (72 scans) MU range. *Adjusted $p < 0.05$; ***adjusted $p < 0.001$.

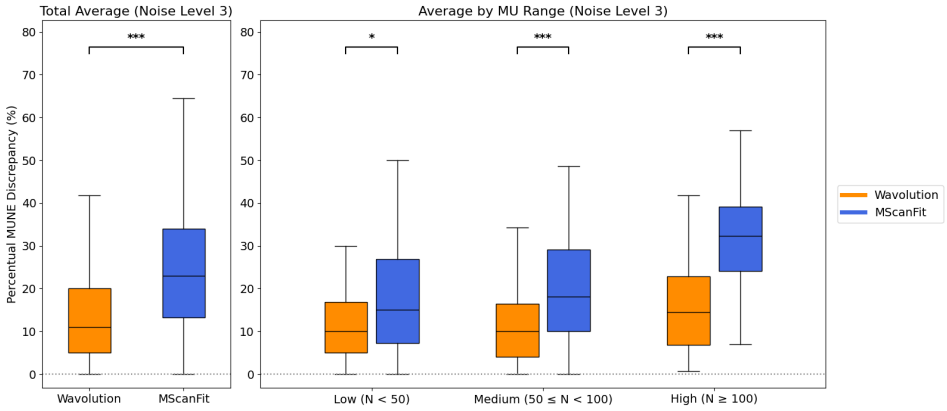


Figure D.2: Average percentual MUNE discrepancy of FitWave and MScanFit for a noise level of $\sigma_3 = 10.0$ µV. (A) shows the average percentual MUNE discrepancy for a total over 192 scans. (B) shows the average percentual MUNE discrepancy for low (60 scans), medium (60 scans) and high (72 scans) MU range. *Adjusted $p < 0.05$; ***adjusted $p < 0.001$.

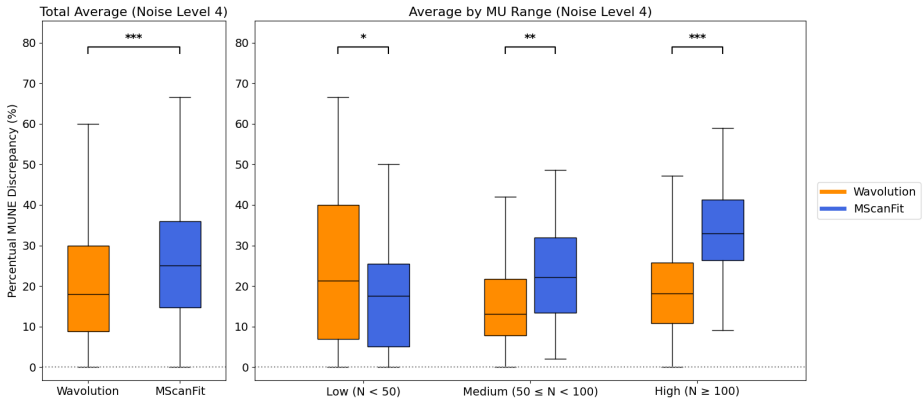


Figure D.3: Average percentual MUNE discrepancy of FitWave and MScanFit for a noise level of $\sigma_4 = 31.6$ µV. (A) shows the average percentual MUNE discrepancy for a total over 192 scans. (B) shows the average percentual MUNE discrepancy for low (60 scans), medium (60 scans) and high (72 scans) MU range. *Adjusted $p < 0.05$; **adjusted $p < 0.01$; ***adjusted $p < 0.001$.

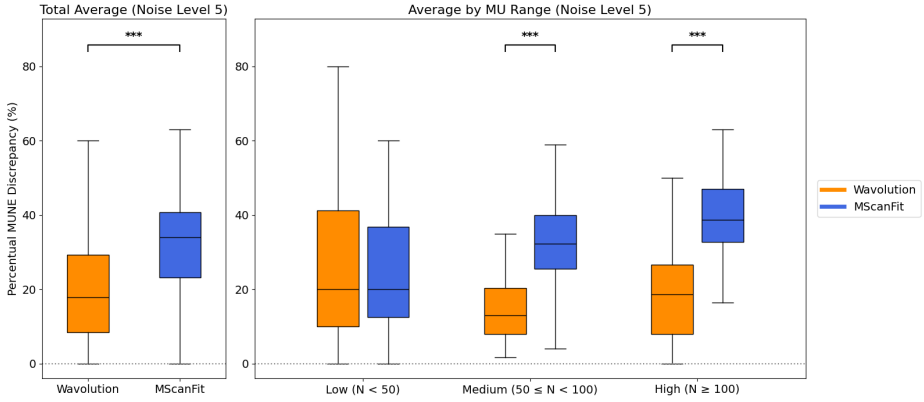


Figure D.4: Average percentual MUNE discrepancy of FitWave and MScanFit for a noise level of $\sigma_5 = 100.0$ µV. (A) shows the average percentual MUNE discrepancy for a total over 192 scans. (B) shows the average percentual MUNE discrepancy for low (60 scans), medium (60 scans) and high (72 scans) MU range. ***Adjusted $p < 0.001$.

D.2. Computation time

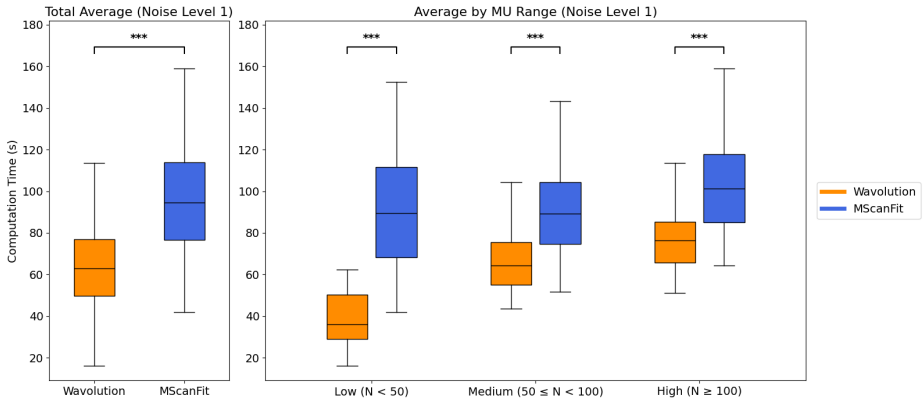


Figure D.5: Average computation time of FitWave and MScanFit for a noise level of $\sigma_1 = 1.0$ µV. (A) shows the average computation time for a total over 192 scans. (B) shows the average computation time for low (60 scans), medium (60 scans) and high (72 scans) MU range. ***Adjusted $p < 0.001$.

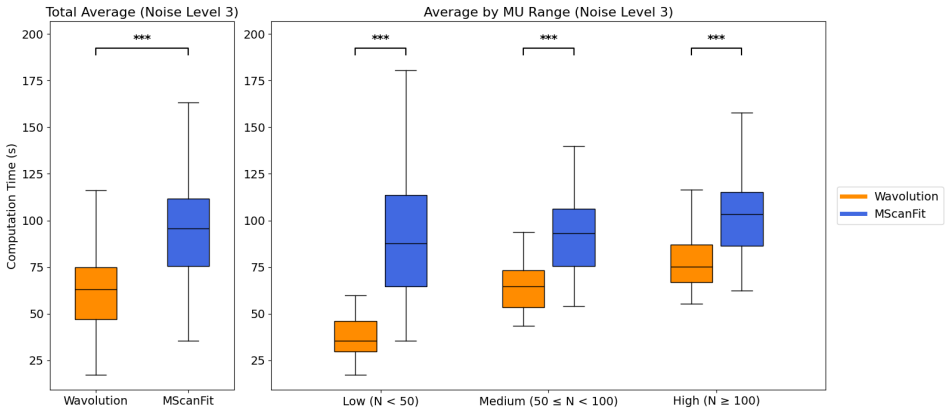


Figure D.6: Average computation time of FitWave and MScanFit for a noise level of $\sigma_3 = 10.0$ µV. (A) shows the average computation time for a total over 192 scans. (B) shows the average computation time for low (60 scans), medium (60 scans) and high (72 scans) MU range. ***Adjusted $p < 0.001$.

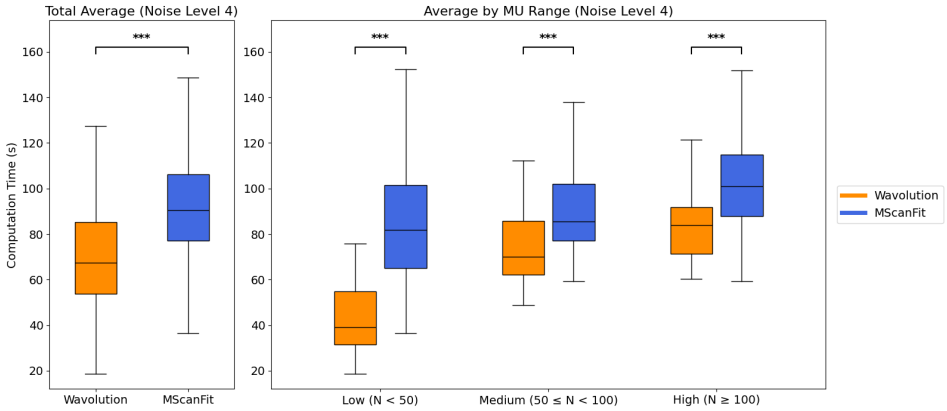


Figure D.7: Average computation time of FitWave and MScanFit for a noise level of $\sigma_4 = 31.6$ µV. (A) shows the average computation time for a total over 192 scans. (B) shows the average computation time for low (60 scans), medium (60 scans) and high (72 scans) MU range. ***Adjusted $p < 0.001$.

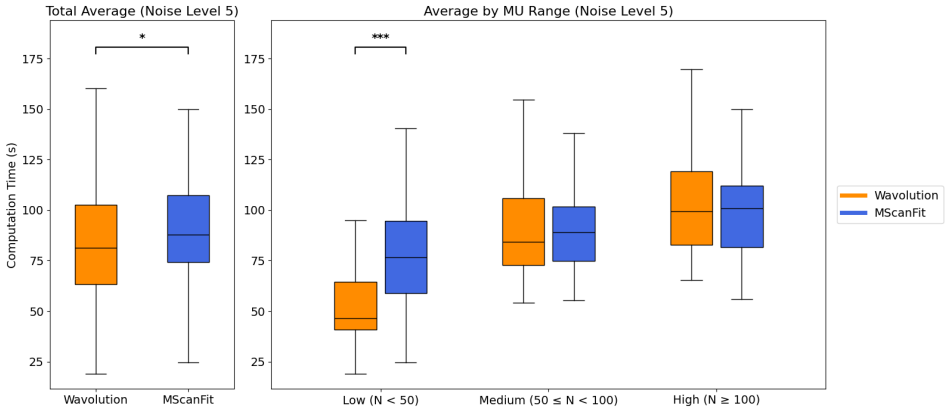


Figure D.8: Average computation time of FitWave and MScanFit for a noise level of $\sigma_5 = 100.0$ µV. (A) shows the average computation time for a total over 192 scans. (B) shows the average computation time for low (60 scans), medium (60 scans) and high (72 scans) MU range. *Adjusted $p < 0.05$; ***adjusted $p < 0.001$.

D.3. MU pool properties

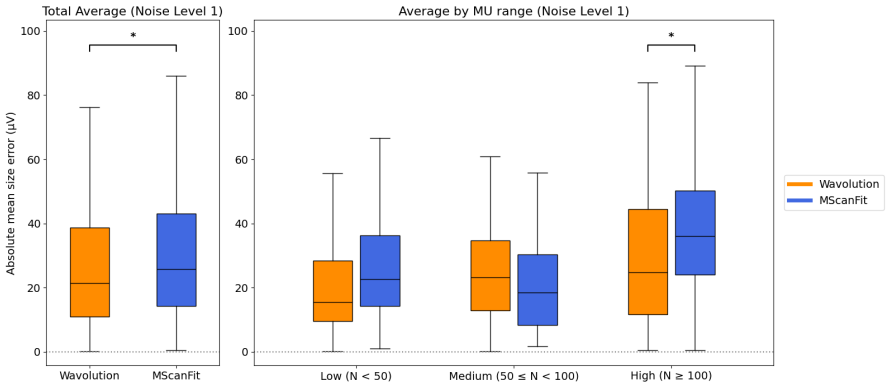


Figure D.9: Average mean SMUAP size error in µV of FitWave and MScanFit for a noise level of $\sigma_1 = 1.0$ µV. (A) shows the average mean SMUAP size error for a total over 192 scans. (B) shows the average mean SMUAP size error for low (60 scans), medium (60 scans) and high (72 scans) MU range. *Adjusted $p < 0.05$; **adjusted $p < 0.01$.

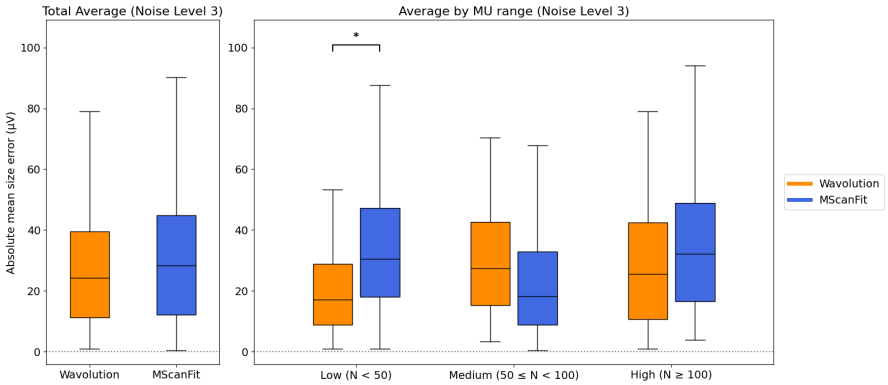


Figure D.10: Average mean SMUAP size error in µV of FitWave and MScanFit for a noise level of $\sigma_3 = 10.0$ µV. (A) shows the average mean SMUAP size error for a total over 192 scans. (B) shows the average mean SMUAP size error for low (60 scans), medium (60 scans) and high (72 scans) MU range. *Adjusted $p < 0.05$.

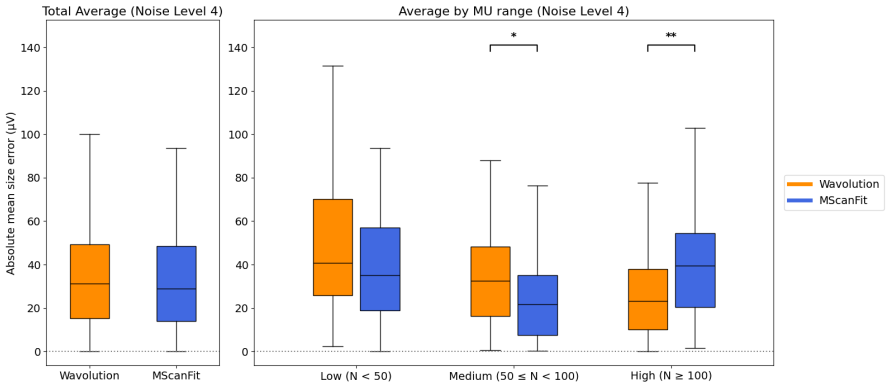


Figure D.11: Average mean SMUAP size error in µV of FitWave and MScanFit for a noise level of $\sigma_4 = 31.6$ µV. (A) shows the average mean SMUAP size error for a total over 192 scans. (B) shows the average mean SMUAP size error for low (60 scans), medium (60 scans) and high (72 scans) MU range. **Adjusted $p < 0.01$.

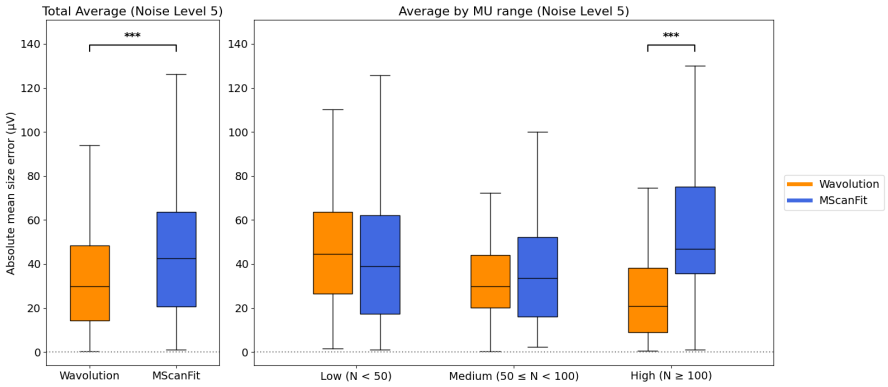


Figure D.12: Average mean SMUAP size error in µV of FitWave and MScanFit for a noise level of $\sigma_5 = 100.0$ µV. (A) shows the average mean SMUAP size error for a total over 192 scans. (B) shows the average mean SMUAP size error for low (60 scans), medium (60 scans) and high (72 scans) MU range. ***Adjusted $p < 0.001$.

D.4. Reduction in maximum CMAP amplitude

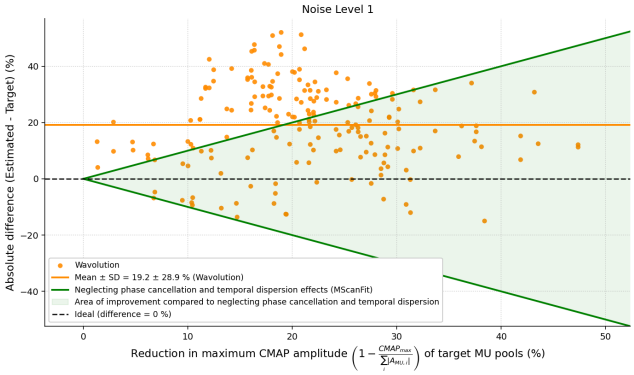


Figure D.13: The reduction in maximum CMAP amplitude in estimated and target MU pools for a noise level of $\sigma_1 = 1.0$ µV.

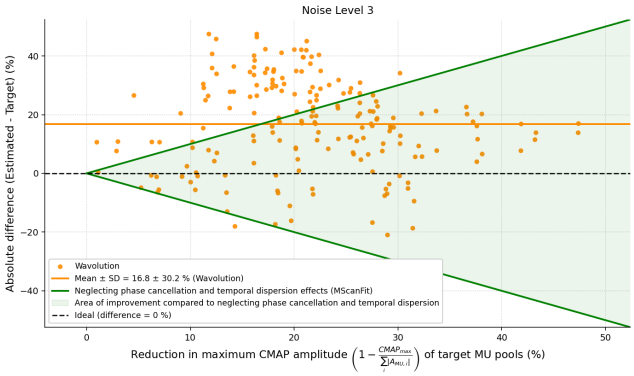


Figure D.14: The reduction in maximum CMAP amplitude in estimated and target MU pools for a noise level of $\sigma_3 = 10.0$ µV.

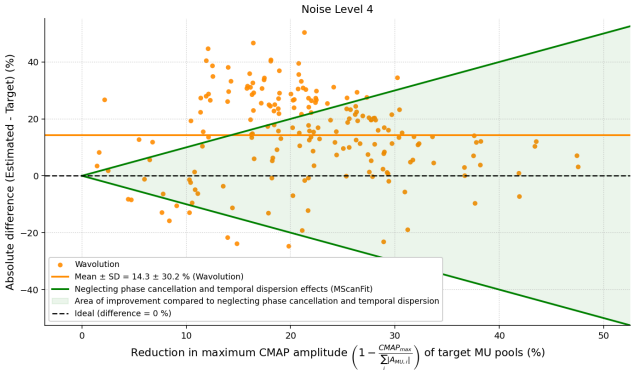


Figure D.15: The reduction in maximum CMAP amplitude in estimated and target MU pools for a noise level of $\sigma_4 = 31.6$ µV.

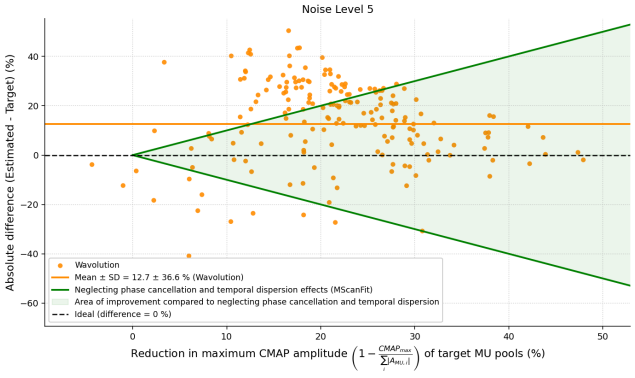


Figure D.16: The reduction in maximum CMAP amplitude in estimated and target MU pools for a noise level of $\sigma_5 = 100.0$ µV.

E

Wavolution Application

In order to make sure that Wavolution is accessible everywhere/worldwide, a Windows application was developed. This appendix chapter contains descriptive information on how to use this application. To ensure that Wavolution can be used for bulk processing, the application has not only the option to fit one CMAP scan, but also the option to upload a .zip file containing multiple target scans for the analysis of multiple scans.

E.1. Uploading target scan(s)

In Figure E.1, the upload screen is shown. Clicking the button 'Upload' opens the local file explorer, from which a target CMAP scan can be uploaded. The accepted formats are .xlsx, .xls, .mem (MScan-Fit QtracP software file format) and .csv. For bulk processing, a .zip file can be uploaded. The .zip file should contain target CMAP scans in similar format as the target CMAP scans used for single processing. By clicking the 'Continue' button, the uploaded files can be used for selecting pre- and post-scan limits.

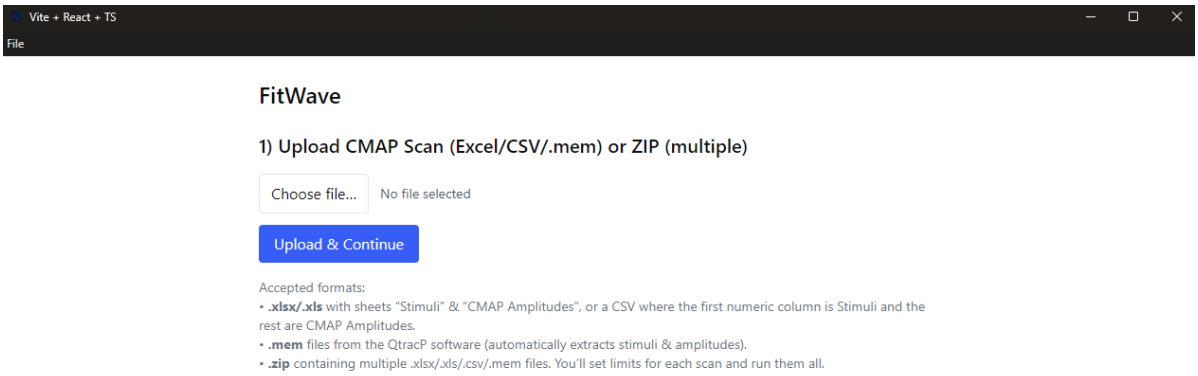


Figure E.1: The upload page of the Wavolution application.

E.2. Selecting pre- and post-scan limits

After uploading the target scan(s), the pre- and post-scan limits have to be set for estimating the baseline noise in the target scan(s). For a single scan, this page is shown in Figure E.2. By selecting the ”Set pre-scan limits” button, the pre-scan limits can be selected by clicking anywhere on the target scan twice. This selects the two boundaries in which the pre-scan region falls. In a similar way, the post-scan region can be selected. By clicking and holding a region, it can be ’shifted’ horizontally. When clicking and holding one of the two boundaries of a region, the region can be enlarged or decreased by sliding horizontally. Additionally, the number of generations for the genetic-like process in Wavolution can be selected (default is 5). By clicking the ’Run Wavolution’ button, the fitting process is started. In the case of bulk processing, however, the fitting process can only be started when all pre- and post-scan limits are selected. The user can switch between scans in the .zip file by clicking on the ’Next scan’ or ’Previous scan’ buttons. After all pre- and post-scan regions in the bulk file are selected, the button ’Run Wavolution’ can be selected.

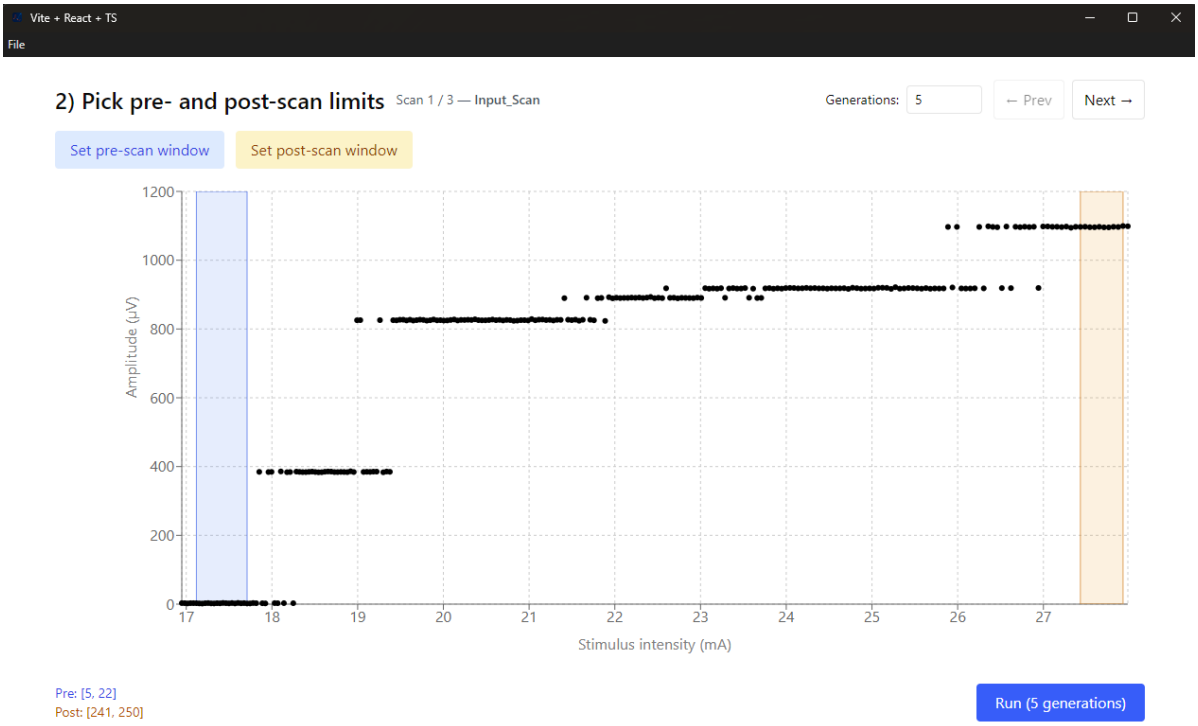


Figure E.2: The page of the Wavolution application where pre- and post-scan limits are selected. For bulk processing, the user can switch between scans with the buttons in the top right of the application.

E.3. Wavolution’s fitting process

In the next screen, shown in Figure E.3, the fitting process of Wavolution can be seen. At the top, the elapsed time is shown, including a blue progress bar and the current generation. The left plot shows a simulated CMAP scan of the best current MU pool of all generations. The right top plot shows the mean + SD MUNE values of the population for each generation. The right bottom plot shows the error score of the best MU pool of each generation. In the case of bulk processing, another progress bar is shown all the way at the top, indicating how many scans have been processed and how many are left for processing. When the fitting process for all uploaded scans has been completed, the user can click the button ’Export results’ in the right bottom of the screen.

Figure E.3: The live fitting page of the Wavolution application that shows the current best MU pool and the population and error over the generations. Once all scans have been processed, the results can be exported by clicking the 'Export Results' button in the bottom right.

E.4. Exported results

The results are exported to a folder of the user's choice, where the folder is named:

{SCAN_ID}*_Wavolution_Results_*{DD/MM/YYYY_HH/MM}

In the case of bulk processing, multiple folders are stored in one folder named:

Wavolution_Results_*{DD/MM/YYYY_HH/MM}*

For each scan that is processed, Wavolution creates five files:

- ***{SCAN_ID}*_MU_properties.xlsx**
  This file contains one sheet of all the numerical MU properties (amplitude A, relative spread ρ, activation threshold t, and phase φ), as well as one sheet with the actual SMUAP signals of every simulated MU in the estimated MU pool.
- ***{SCAN_ID}*_CMAP_Scan.xlsx**
  This file contains one sheet with all the stimulus intensity values, as well as one sheet with the actual CMAP signal values for every stimulus intensity value.
- ***{SCAN_ID}*_ScanResults.xlsx**
  This file contains one sheet with an overview of the scan results, including MUNE, runtime, mean MU amplitude, largest MU amplitude, smallest MU amplitude, and mean relative spread.
- ***{SCAN_ID}*_CMAP_Scan.png**
  This file shows a figure of the CMAP scan described in the file above.
- ***{SCAN_ID}*_TotalResultsOverview.png**
  This file shows an overview of the CMAP scan and fitting process, including the CMAP scan, the smoothed line, the absolute and smoothed error, and the mean + SD MUNE plot and the error score plot of Wavolution's fitting process.

F

Pre-processing of CMAP recordings

F.1. Detrending

Baseline drift can lead to atypical CMAP scans of the actual muscle responses, which in turn has an effect on the MUNE value. Detrending individual CMAP recordings 'removes' this baseline drift from each measurement such that the CMAP scans represent the actual EMG recordings from baseline to peak. Detrending occurs based on the 'trend' that is present in the separate CMAP recordings. The baseline of these measurements is estimated by a spline that is fitted through the first and last data points of the measurement. By subtracting this spline from the baseline, the difference is removed, resulting in an amplitude measurement from 0 to V_{max} (baseline-peak). An example of baseline correction for a recorded EMG signal is shown in Figure F.1. Baseline correction based on low frequency removal was also tested, but spline-based correction resulted in more accurate baseline correction.

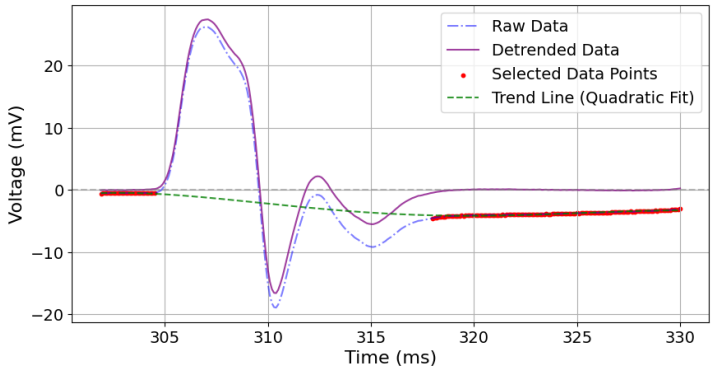


Figure F.1: Baseline drift correction for an EMG signal recorded from the abductor pollis brevis.

F.2. Denoising

One-dimensional signals that contain noise can be seen as the original signal with additional noise through linear superposition:

$$X(k) = S(k) + E(k) \tag{F.1}$$

where $X(k)$ is the noisy signal, $S(k)$ is the original signal, and $E(k)$ is white Gaussian noise, subject to $N(0, \sigma^2)$ distribution [52].

An EMG recording of a CMAP response to a stimulus can be seen as a discrete, one-dimensional wavelet signal that contains noise. Especially at lower stimulus levels, where the responses have lower magnitudes, the noise affects the measurements of amplitude and area of the recording, which in turn affects the CMAP-scan. Therefore, wavelet denoising can be performed based on wavelet decomposition. An input signal is decomposed into wavelet coefficients using a discrete wavelet transform:

```
coeffs = pywt.wavedec(signal, wavelet=wavelet, level=level)
```

This returns a list of coefficients separating the signal into different frequency bands, where `coeffs[0]` contains the approximation coefficients (e.g., the low-frequency components), and the higher levels contain the detail coefficients (e.g., the high-frequency components). Large amplitude wavelet coefficients may be produced by the useful signal, and the small amplitude is likely to represent the noise [52].

There are many ways to estimate the noise level based on these coefficients. The universal thresholding method introduced by Donoho [53] is widely used in wavelet-based denoising. It automatically estimates the noise level in the wavelet coefficients, making it suitable when noise levels are unknown or variable. It is defined as:

$$T = \sigma\sqrt{2\log N} \tag{F.2}$$

where σ is the average variance of the noise and N is the signal length. The average variance, σ, is calculated using the median estimate method:

$$\sigma = \frac{\text{median}(|W_{1,k}|)}{0.6745} \tag{F.3}$$

where $W_{1,k}$ represent all the wavelet coefficients in level 1 [52].

Based on this threshold, hard thresholding is performed, where wavelet coefficients smaller than threshold T are removed. The performance of wavelet denoising depends on the type of wavelet used for decomposition and the level of wavelet decomposition. In order to investigate which type of wavelet and level provide the most accurate denoising, wavelet denoising was performed for all types of discrete wavelets and levels 2 to 6 (the denoising parameters). A simulated set of clean and noisy signals was used, where the noisy signals were denoised. The resulting denoised signals were compared to the clean signals in terms of root-mean-squared error (RMSE), where the lowest RMSE values corresponded to the most accurate denoising parameters.

As there are 127 discrete wavelets in the wavelet decomposition program, first, a preliminary analysis was done, where the RMSE values of all levels were summed for the discrete wavelets. From this, the 10 discrete wavelets with the lowest total RMSE values were selected. These wavelets were tested at levels 2 to 6 for single MU action potentials (MUAPs) of CMAP recordings. For simulated CMAP recordings of 5, 40, 80, and 120 MUs, responses at 4 different stimulus intensities were analyzed, leading to a total of 16 different CMAP recordings. An example of the denoising process of one of these recordings is shown in Figure F.2.

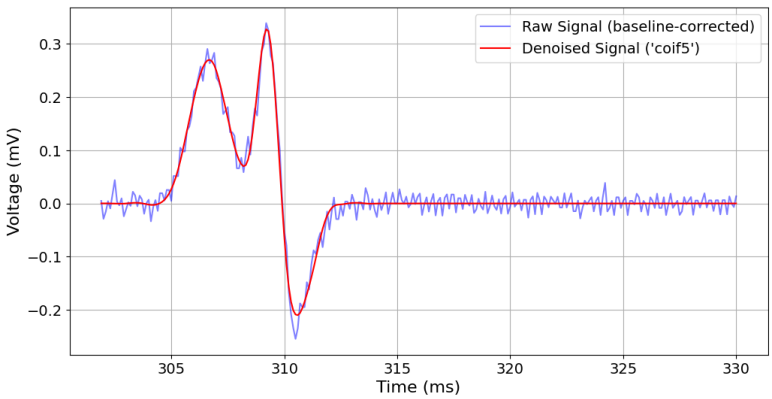


Figure F.2: Wavelet-based denoising for an EMG signal recorded from the abductor pollis brevis.

For all the recordings, the RMSE was summed for each level of every wavelet. The wavelet and corresponding level with the lowest total RMSE was found to be the `'coif5'` wavelet with level 3. Therefore, these denoising parameters were used in the wavelet denoising process of CMAP recordings. One should keep in mind that the recordings used in this analysis of wavelet families and levels are based on the abductor pollicis brevis (APB) muscle. Therefore, when analyzing other muscle responses in CMAP recordings, it could be that other wavelet families or levels lead to a more accurate denoised signal.