Phantom movement training without classifier performance feedback improves mobilization ability while maintaining EMG pattern classification
Olivier Rossel, Manon Chateaux, Nathanaël Jarrassé, Fabien Vérité, Amélie Touillet, Caroline Nicol, Jean Paysant, Jozina B. De Graaf

To cite this version:
Olivier Rossel, Manon Chateaux, Nathanaël Jarrassé, Fabien Vérité, Amélie Touillet, et al.. Phantom movement training without classifier performance feedback improves mobilization ability while maintaining EMG pattern classification. IEEE Transactions on Medical Robotics and Bionics, 2023, 5 (1), pp.133-142. 10.1109/TMRB.2023.3237918 . hal-04137971

HAL Id: hal-04137971
https://inria.hal.science/hal-04137971
Submitted on 22 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
Phantom movement training without classifier performance feedback improves mobilization ability while maintaining EMG pattern classification

Olivier Rossel, Manon Chateaux, Nathanaël Jarrassé, Fabien Vérité, Amélie Touillet, Caroline Nicol, Jean Paysant, and Jozina B. De Graaf.

Abstract—Voluntary phantom movements are systematically associated with muscle contractions in the residual limb. These latter are specific to the type of movement and can be classified by pattern recognition algorithms. However, phantom mobility generates fatigue that could impact classification metrics. This study explored whether daily phantom movement training at home with no other feedback than inherent somatosensory information can impact the classification success rate. Kinematics and muscle activity were compared between before and after a two-month home training in six major upper limb amputees. Surface EMG patterns were classified to quantify a potential change in the features space with training. Our results showed that this type of training induces faster, smoother, and richer phantom mobility. However, classification metrics did not change with training. When including the new types of movements achievable after training, accuracy did not decrease, indicating that muscle activation patterns associated with these movements were sufficiently different not to interfere with the already existing movement classes. Thus, although phantom training with only somatosensory feedback increases the overall phantom movement capacity, it does not increase the classification success rate. Yet, it is possible that paired with other forms of feedback, phantom training could improve this success rate.

Index Terms— Amputation, EMG-classification, Feedback, Phantom movement training

I. INTRODUCTION

Most upper limb myoelectric prostheses currently available on the market involve non-intuitive sequential control of only a few degrees of freedom (DoF). They also entail a demanding learning period and the still poor ergonomics at the end leads 44% of users to give up using them [1], [2]. Thus, many research teams are working on the challenging issue of improving the performance of myoelectric control by increasing the number of DoF controllable in an intuitive way. Most current methods involve pattern recognition (PR) algorithms. PR relies on the hypothesis that each motor action is characterized by a repeatable muscle activity pattern, which can be defined by a set of elementary characteristics (amplitude, frequency, etc.). PR algorithms are classifiers that must be trained by presenting EMG data labeled according to the associated movement. Although successful in research studies and starting to be popular in clinical context (e.g., Coapt, Ottobock), PR techniques are rarely applied to above-
elbow amputees without targeted muscle reinnervation, and most importantly, their currently limited robustness remains a major obstacle to everyday use by amputees [3]. Misclassification of myoelectric patterns can come from the user, for example from variability in muscular contractions because of effort modulation [4] or from motor intention variations or muscle fatigue [5], but also from numerous external factors such as the changes in the residual limb posture or in the placement of electrodes [6]. Also, the necessary consistency of muscular patterns is difficult to obtain since, depending on the situation, muscles can be contracted in numerous different ways to perform a similar gesture [7]. Moreover, current PR control requires amputees to perform explicit muscle contractions instead of gestures, since the limb is absent. As it is known that one only has limited awareness of one’s muscle contractions [8], [9], such instruction can lead to low reproducibility.

Voluntary phantom mobility is used in research as a promising way to control prostheses via PR [10]–[14]. This phenomenon is defined as the ability to send a motor command to one’s phantom limb that will in turn elicit the somatosensory sensation of a movement via systematically associated contractions in residual muscles [12]–[17]. Phantom mobility, and more specifically phantom hand movement (PHM), is a widespread phenomenon irrespective of the level of amputation. A recent epidemiological study [18] showed that nearly 80% of below-elbow and 70% of above-elbow amputees described the natural ability to perform voluntary PHM. For the latter group, it is important to note that the contractions associated with PHM are measured in the muscles of the upper arm, whereas these were not involved in the movements of the intact hand before amputation. This could be partly explained by a reorganization of the peripheral nervous system. Indeed, results obtained from amputated monkeys strongly suggest that after amputation, axons of spinal motoneurons are able to sprout again and get involved in mechanisms of muscle fibers recapture [14], [15]. Thus, contractions observed in the residual muscles could be the direct consequence of the motor command sent to the limb that was amputated. The spinal motoneurons that previously targeted muscles involved in intact hand movements would have retargeted muscle fibers in these more proximal muscles. The sensation of movement in the phantom limb would in turn be the somatosensory feedback caused by these muscle contractions, also associated with movements of skin and other tissues (see [16] for a convincing illustration of this statement).

As each contraction pattern is specific to a given type of phantom movement, surface EMG signals are particularly informative about the type of PHM being performed [13], [21]. PR has been used to identify the type of performed movement, in above-elbow amputees in particular, which eventually allows intuitive control of myoelectric prostheses with many DoF [13], [21]. Thus, phantom-based prosthesis control is a promising non-invasive technique that might improve both ergonomics and performance of myoelectric prostheses. Yet, although phantom movements are natural, they generate fatigue that sometimes makes PHM difficult to perform [22]. This leads to a lack of endurance and potentially influences EMG signals by generating unreproducible and insufficiently specific EMG patterns [23]. This could be an obstacle to consistently high classification rates and thus to the use of PHM for prosthesis control.

Several recent studies have thus been focusing on movement training to produce more distinct EMG patterns between movements [24]–[26]. What would happen if training were associated with phantom-based prosthesis control? We hypothesize that phantom movement training might not only decrease the associated fatigue and thus facilitate the execution of phantom movements, but also improve phantom-based PR performance by increasing the amount of information contained in the EMG signals that is usable by the classifier. Phantom movement training has mostly been investigated to treat phantom pain (through virtual reality using PR, or mirroring) [24], [25]. Thus, the evolution of phantom movement characteristics (through analyses of kinematic or EMG data) is often overlooked. Thanks to an instrumented glove that allows to collect kinematic data through mimicking of PHM with the intact hand in real time, Garbarini and colleagues [26] showed that after a short period of phantom finger-tapping training the execution speed was increased. This study included only one amputee with phantom mobility. One of our recent studies [18] took interest in 5 participants who had undergone an above-elbow amputation a few years earlier. They participated in a two-month protocol of daily PHM training. With training, it appeared that PHM became faster but also more enduring, as shown by the increasing number of repetitions a given phantom movement was executed without blocking, reported as an “impossibility to move despite strong effort” [22]. Moreover, new types of PHM became achievable with training. No evolution was seen when no training was performed [18]. This suggests that training phantom movements improves some aspects of phantom mobility, at least in above-elbow amputees. However, because no EMG recording was reported in these studies, it is not known whether the associated EMG signals, and therefore PR performance, were altered by such training.

A few studies have focused on combining training and PR based on EMG signals. But mostly below-elbow amputees were included in these studies [10], [27]–[31]. Moreover, the types of training that were chosen required laboratory set-ups which limits the number of amputees that could benefit from it in a clinical context. Finally, the...
feedback provided during training was based on the performance of the PR algorithm, which allowed the participants to adapt their movements to increase this performance by modifying the information contained in the EMG signals (e.g. [30]). Such training might have resulted in movement intentions that did not match the prosthesis movements. Although for current prosthesis use, amputees learn matching instructed muscle contractions with device output, the interest of PHM is that it needs no learning to obtain natural and intuitive control (i.e., without mismatch). This is ergonomic and contributes to the embodiment of the prosthesis (e.g. [32]).

To avoid the previously cited disadvantages caused by feedback on PR performance, we chose to consider another type of training: a two-month daily PHM training at home with no other feedback than natural somatosensory information. The aim of the present study was to explore whether improvements in kinematics following home training [18] were accompanied by an increase in phantom-based PR performance. For six amputees (of whom 5 above-elbow), we compared the PHM kinematics and the muscle activity on the residual limb between before and after training. We classified EMG patterns to evaluate the change in the amount of information usable by the classifier. The results of this study may have consequences for both the amputee rehabilitation process and the control of myoelectric prostheses.

II. METHODS

A. Participants

Six male participants (P1 to P6, 24 to 75 years old) with an amputation of their dominant upper limb were included in the study. Four of them were above-elbow amputees (P1-P4), P5 was a below-elbow amputee, and P6 had a scapulothoracic disarticulation. For 3 of the participants some kinematic data had already been shown in [18]. The inclusion criteria required that the amputation dated back to at least one year before the first experiment. Participants also had to be able to generate PHM and mimic in real time with the intact hand (for more details see [22]). They all declared motivation to train daily at home by repeating all the phantom movements that they could perform. Amputees who experienced phantom limb pain while performing PHM were excluded. The time since amputation varied from 1 to 34 years. Three of the 6 amputees used a myoelectric prosthesis daily, controlled by the classical threshold-based method (on/off and co-contractions strategy). The demographics of the participants are summarized in Table 1. This study was conducted in accordance with the World Medical Association Declaration of Helsinki. Informed consent was obtained from all participants and the study received approval from the Clinical Research Ethics Committee (CERES No 2016-57).

Table 1. Demographic data concerning the six participants, all males. TH=transhumeral, TR=transradial, and SD=scapulothoracic disarticulation. R=Right, L=Left. For all participants, the amputated arm was the dominant one before amputation. Fx = flexion and extension of finger x (with 1=thumb... 5=little finger), Fxy =fingers x and y together, P=Pinch opening/closing, H=Hand opening/closing, Wfl=wrist flexion/extension, Winc=wrist ulnar/radial inclination. The new movements that became achievable during the training period are marked in brackets in the last colon.

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Delay (years)</th>
<th>Level</th>
<th>Amp side</th>
<th>My o</th>
<th>PHM</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>42</td>
<td>TH</td>
<td>R</td>
<td>Yes</td>
<td>F5, H (F1, F2, Wfl)</td>
</tr>
<tr>
<td>P2</td>
<td>49</td>
<td>3</td>
<td>TH</td>
<td>L</td>
<td>No</td>
</tr>
<tr>
<td>P3</td>
<td>24</td>
<td>2</td>
<td>TH</td>
<td>L</td>
<td>Yes</td>
</tr>
<tr>
<td>P4</td>
<td>58</td>
<td>34</td>
<td>TH</td>
<td>L</td>
<td>No</td>
</tr>
<tr>
<td>P5</td>
<td>40</td>
<td>1</td>
<td>TR</td>
<td>L</td>
<td>Yes</td>
</tr>
<tr>
<td>P6</td>
<td>57</td>
<td>6</td>
<td>SD</td>
<td>R</td>
<td>No</td>
</tr>
</tbody>
</table>

B. Protocol

All participants performed two recording sessions, one before and one after the two-month home training period. Each recording session started with a short questioning to determine the different types of PHM the participant could perform and the self-assessed difficulty in performing them. The tested movements were actual movements that involved a change of hand posture: individual finger movements, closing/opening of the hand and pinch, and global hand movements around the wrist (flexion/extension and inclination). Movements were recorded from easiest (less effortful) to most difficult. For each movement, the participant was instructed to simultaneously mimic the kinematics of the PHM (amplitude, velocity; for more details see [22]) with his intact hand, so that EMG activity was recorded on the residual upper limb while kinematics was measured from the contralateral hand. Each type of PHM was performed in series of up to 10 repetitions of a complete cycle (i.e., flexion/extension or closing/opening) at a comfortable speed chosen by the participant himself. Special care was taken to protect participants from fatigue. If the participant could not perform 10 cycles in a row (e.g., movement blocking, residual limb tremor), the recording was stopped and the participant rested for a few minutes before recordings were resumed for either the same or another type of movement. This was repeated until a total of 10 cycles was reached.
for each movement type. Each session lasted approximately 60 minutes and was video recorded. Recordings took place in the presence of a medical doctor.

Between these two recording sessions, the participants were asked to find a moment each day to train their PHM. They were instructed to execute as many cycles as possible (up to 15) without blocking. If they felt that they were able to perform a new type of PHM (i.e., one they were previously unable to execute), they had to integrate this movement into their daily training. Contrary to recording sessions before and after training, the participants were specifically instructed not to mimic the PHM with their intact hand. They were also instructed to stop immediately if the training induced pain. During the training period, a weekly phone interview was organized to evaluate compliance, tolerance, and observance. Each training session was reported to take at most 15 minutes.

C. Recordings

The kinematics of the mimicked phantom finger movements were recorded on the intact hand using a Cyberglove® II, at a frequency of 100 Hz. Details on the calibration of the Cyberglove have been reported previously [12], [22]. Global hand movements around the wrist were recorded with an inertial measurement unit (IMU, nine degree-of-freedom Sensor Stick from Sparkfun ®) that was attached to the dorsal surface of the glove. A sensor fusion algorithm computed IMU orientations based on the data from the accelerometer, gyroscope, and magnetometer. This allowed a reconstruction of the hand orientations in space and the range of the participants’ wrist movements (flexion/extension, pronation/supination, ulnar/radial inclination).

EMG activity was recorded at a frequency of 1 kHz with an ANT-Neuro® Eego-sports system with shielded cables. EMG signals were recorded in a continuous way during the whole experimental session and were synchronized with the kinematic recordings thanks to a dedicated push button used at the start of each series of movement cycles. The placement of surface EMG electrodes (Ambu© BlueSensor Ag/AgCl snap bipolar electrodes, 2 cm inter-electrode distance) was individualized for each amputee based on palpation of the active muscle volumes during PHM. They were positioned on various sites of the residual limb as shown in Fig. 1. In the case of above-elbow amputees, our present analysis focused on the EMG signals recorded by electrodes placed distally to the deltoid muscles (i.e., on the sole biceps and triceps brachii muscles). As a function of the length of the residual limb, 8 to 11 pairs of electrodes were placed. For the participant with a scapulothoracic disarticulation, 12 electrode pairs were placed around the shoulder and all of them were considered. For the below-elbow amputee, 11 electrode pairs on the residual forearm were selected. Electrode positions were photographed so that they could be reproduced in the second recording session after training.

III. ANALYSES

Since flexion and extension (or opening and closing) are antagonistic movements, and thus produced by different muscle volumes, it was necessary to distinguish them in the EMG signals. Therefore, the entire kinematic and EMG analysis was based on movement half-cycles. In the following, the ”number of different types of PHM” refers to all the different movements that the participant could perform, considering flexion and extension (or opening and closing) as two types of movements. Furthermore, in most analyses, the comparison of variables for initially existing movements between before (called ”Pre”) and after (called ”Post”) was made on all 10 performed cycles whether they were executed in one sequence or in different series. The new movement types that appeared during the training period were not considered in this Pre/Post comparison since they did not undergo the exact same duration of training. However, to analyze whether these new PHM changed classification results, they were added in a separate analysis (called ”PostNew”) and compared to Pre and Post.
Electrode placements on the residual limb of the 6 participants. Besides for P6, two views allow visualization of almost all electrode pairs. For P1 to P4 with an above-elbow amputation, only the electrodes placed on the biceps and triceps were selected for further analyses. The ones on the deltoid, pectoralis and dorsal muscles were excluded. For P5 with a below-elbow amputation, all 11 electrodes placed on the residual forearm were selected. For P6 with a scapulothoracic amputation, all 12 electrode pairs were selected.

A. **Kinematics**

First, we counted the number of different types of PHM as well as the maximum number of cycles per type of PHM that each participant was able to perform without blocking. The latter measure is an indicator of susceptibility to fatigue. Then, for each participant and each type of phantom movement, we selected the specific joints involved and the kinematics was calculated by averaging the angular displacements over all the selected joints. The angular velocity was calculated as the time derivative of the recorded angles and was filtered with a third-order lowpass Butterworth filter at 2 Hz. The beginning and end of each half-cycle was identified by a change in movement direction visible in the velocity sign (see Fig. 2). For each participant, each type of phantom movement and each half-cycle, movement amplitude was computed as the difference between maximum and minimum angular displacements. Half-cycle duration and peak velocity were also extracted. To compare PHM kinematics between Pre and Post, different variables were computed and explained here-under.

Relationship between amplitude and peak velocity

For unconstrained cyclic movements, amplitude and velocity varied between cycles and between movement types, even in intact movements. Yet, what remained stable is the relationship between movement amplitude and peak velocity, the latter increasing linearly with increasing amplitude [33]. To investigate whether PHM training influences this relationship, the slope $\beta$ of the linear relation between peak velocity and amplitude was compared between Pre and Post for all movement types combined.

**Relationship between amplitude and peak velocity**

For unconstrained cyclic movements, amplitude and velocity vary between cycles and between movement types, even in intact movements. Yet, what remains stable is the relationship between movement amplitude and peak velocity, the latter increasing linearly with increasing amplitude [25]. To investigate whether PHM training influences this relationship, the slope $\beta$ of the significant linear relationship between peak velocity and amplitude was compared between Pre and Post for all movement types combined.

---

**Fig. 1.** Electrode placements on the residual limb of the 6 participants. Besides for P6, two views allow visualization of almost all electrode pairs. For P1 to P4 with an above-elbow amputation, only the electrodes placed on the biceps and triceps were selected for further analyses. The ones on the deltoid, pectoralis and dorsal muscles were excluded. For P5 with a below-elbow amputation, all 11 electrodes placed on the residual forearm were selected. For P6 with a scapulothoracic amputation, all 12 electrode pairs were selected.

**Fig. 2.** Illustration of the computation of the reproducibility indices. (a) Angle, angular velocity (upper part), EMG of the Triceps Brachial (TB) and the Biceps Brachial (BB) (lower part) for 3.5 cycles of finger flexion and extension. Separation of half-cycles was done according to the sign of the velocity signal, allowing extraction of the amplitude and peak velocity for each type of movement. (b) Temporal normalization on 100 samples per half cycle. Thin traces indicate the temporal normalized signal for all executed repetitions (i.e., mostly 10), bold traces the average over all repetitions. For the sake of...
clearness, standard deviations are only represented for 5 different samples. The reproducibility indices (RI-Kin and RI-EMG) calculated for this example are reported for each half-cycle above the obtained curves.

Reproducibility Index (RI-Kin)

As the ability to reproduce an intact limb movement increases with training, the same is expected for phantom movements. Therefore, we defined a Reproducibility Index (RI) to quantify the ability to reproduce the kinematics over repetitions of a given movement. Before calculating this index, each type of movement (i.e., each half-cycle) was normalized to a time-base of 100%, containing 100 samples. This allowed us, for each type of movement, to average the velocity signals over repetitions (bold traces in Fig. 2b). For each type t of PHM, RI is computed as follows:

\[
RI^m = 1 - \frac{1}{100} \sum_{t=1}^{100} CV_t^m
\]

where \( CV_t^m \) is the coefficient of variation over repetitions for velocity values at “time” t. RI is close to 1 if \( CV_t^m \) averaged over t is close to 0 (i.e., standard deviations are small compared to their respective mean velocity values). Some values of RI are reported for illustration in Fig. 2b above the velocity signal. For each participant, the global RI-Kin was then computed as the RI averaged over all movement types.

Fluidity index (FI-Kin)

Before training, PHM were reported to be difficult to produce and to repeat: participants often perceived their cyclic phantom movements slowing down and sometimes eventually blocking [22]. As with intact limb movements, the ability to produce smooth movements should increase with training, which has been suggested by our previous results [18]. Smooth cyclic movements are characterized by an acceleration and deceleration phase that always precede a change in movement direction, whereas numerous acceleration/deceleration phases without any change of direction indicate ‘non-smooth’ movements. Thus, we propose the calculation of a Fluidity Index (FI) based on the count of deceleration phases. For each movement type m, FI is computed as follows:

\[
FI^m = \frac{1}{R} \sum_{r=1}^{R} \frac{2}{1 + \frac{D_r^m}{R}}
\]

Where \( D_r^m \) is the number of deceleration phases for movement type m at repetition r. R is the total number of repetitions (in general 10). If the PHM is perfectly fluid, there is only one deceleration phase per repetition: \( D_r^m = 1 \) so \( FI^m = 1 \). Figure 3 illustrates the method of FI calculation performed on the velocity signal before and after training. For each participant, the global FI-Kin was then computed as the FI averaged over all movement types.

---

**Fig. 3.** Calculation of the fluidity index, here for 2 cycles of little finger flexion and extension (a) before (Pre) and (b) after (Post) training for one participant. Upper trace: angle as a function of time. Second trace: derivative of the angle with in red, for illustration, the deceleration count. Third and fourth traces: rectified EMG signal from the BB, respectively TB, muscle, with in orange the RMS envelope.
B. EMG patterns

Two types of EMG signal analyses were performed. One was based on the temporal evolution of EMG amplitude and the other consisted in classifying EMG patterns (i.e., pattern recognition, PR). Both types of analyses started with filtering the recorded EMG signals with a third-order bandpass Butterworth filter between 10 Hz and 400 Hz. To distinguish EMG activity associated to PHM half-cycles, the beginning and end were identified with the help of the kinematic data (see Section IIIA Kinematics). Two variables were computed to explore the temporal evolution of EMG-amplitude: the EMG fluidity index (FI-EMG) and the EMG reproducibility index (RI-EMG) (see below). They were both computed before (Pre) and after (Post) training, for each type of movement. For the classification of EMG patterns, we used a linear discriminant analysis (LDA) classifier which is commonly used in pattern recognition algorithms (e.g., [13], [21], [30], [34]). Using the above-mentioned half-cycles, each movement type was assigned to a different class. For each class, the following features were computed from the EMG: the root mean square (RMS envelope), the first 4 autoregressive coefficients, the zero crossings, the wavelength and the sample entropy [13], [21]. Because of the extended duration of sample entropy computation, a 512-ms sliding analysis window with a 128-ms overlap between successive windows was used. To assess the possible change caused by phantom training, we used three commonly used metrics (e.g., [9], [10], [35]): Classification Accuracy (CA), Mean Euclidian Distances between classes (MED), and the Inter-Class Distance Nearest Neighbor (IDNN). These metrics were computed on movement types that already existed before training for both Pre and Post, as well as for all movement types including the new ones that had appeared during the training period (PostNew). All analyses were performed using MATLAB software (The MathWorks, Natick, MA).

Reproducibility index (RI-EMG)

The EMG reproducibility index was computed from the EMG RMS envelope of each channel in the same way as that used for kinematics (see Fig. 2). Thus, EMG signals of each type of movement (i.e., each half-cycle) and each channel was normalized to a time-base of 100% such that the signals could be averaged over repetitions. RI_{mc} - EMG is defined as the reproducibility index of movement type m at EMG channel c. In Figure 2a, the two lower traces show EMG bursts during two cycles of finger movement measured on two distinct channel pairs on the Triceps Brachii (TB) and on the Biceps Brachii (BB). For each participant, RI_{mc} represents the average of RI_{mc} over all EMG channels and the global RI-EMG was computed as the average of RI_{mc} over all movement types.

Fluidity index (FI-EMG)

While FI-Kin was calculated on the velocity signal, FI-EMG was calculated based on the RMS envelope. The number of deflections in the RMS envelope, given by the number of times the acceleration was zero, was automatically determined for each half-cycle. First, for each PHM type m of series of cyclic movements, the number of deflections for each repetition r and each EMG channel c was extracted and represented by \( N_{r mc} \). FI_{mc} was defined as the fluidity index of movement type m for EMG channel c. For each participant, FI_{mc} represents the average of FI_{mc} over all EMG channels and the global FI-EMG was computed as the average of FI_{mc} over all movement types. Fig. 3 shows an example for two EMG channels.

Classification Accuracy (CA)

The stability and specificity of the movement commands was assessed by the participant’s ability to produce classifiable patterns. To avoid problems such as overfitting or selection bias, classification accuracy (CA) resulted from a 5-fold cross-validation [34], [36]. Each class was divided into 5 segments of 2 full repetitions containing the same number of classes. Four segments were used for training and the obtained decoder was applied to the remaining segment. This was done for an exhaustive number of combinations. From this, the final confusion matrix was extracted where each row represents an actual class while each column represents a predicted class. The true positive rate for the class \( m \) is CA_{mc}, which is the diagonal of the confusion matrix. Then CA was calculated as the total true positive rate calculated over all movement types, higher CA indicating better recognition of the movement classes. The maximum score of this metric is 1.

Mean Euclidian Distance (MED)

Euclidian Distance is a metric used to assess changes in the distribution of class clusters in high dimensional space. First, the Euclidian Distances between each feature centroid was computed [10], [35]. Then MED_{mc} was
defined as the Mean Euclidian inter-class Distance between centroid for the movement \(m\) and all the other classes. Finally, for each participant, the mean value was extracted to construct the MED.

**Inter-class Distance Nearest Neighbor (IDNN)**

The separability of EMG patterns over different movements was quantified by the Inter-class Distance Nearest Neighbor (IDNN) as defined by [9]. It relies on the Mahalanobis distance that is currently used in pattern recognition because it measures the distance between two classes while considering the variance of each class in the space between the centroids. It therefore included the size and shape of class distributions. First, the inter-class distance between classes \(m_i\) and \(m_j\) is computed as the reduced sum of Mahalanobis distances in feature space from classes \(m_i\) and \(m_j\). Then \(IDNNm\) was defined as the minimum of the inter-class distances between the class \(m_i\) and all the other classes. Finally, for each participant, the \(IDNN\) was computed as the value of \(IDNNm\) averaged over all types of movements.

### C. Statistics

Statistical testing was done for each kinematic and EMG variable. The kinematic and EMG variables averaged over all movement types were not normally distributed over subjects. Thus, for group results, the median values and their corresponding 1st and 3rd quartiles were reported. The Wilcoxon paired signed-rank test was used to compare results between Pre and Post. Concerning the relationship between movement amplitude and its peak velocity, a Pearson correlation test was performed. For classification metrics, a Friedman test was performed to compare the 3 conditions Pre, Post and PostNew. For all tests, the significance level was set to 0.05.

### IV. RESULTS

All 6 participants completed the training period and declared no sensation of pain in either the residual or the phantom limb during or following phantom movement executions. During our weekly phone interview, they all confirmed their daily implication in the training.

#### A. Kinematics

All kinematic Pre and Post data are reported in Figure 4. Overall, 5 out of the 6 participants performed more types of PHM after training than before training (group comparison p<0.03, Fig. 4a; see also Table 1). New movements had appeared 1 to 2 weeks after the beginning of training, depending on the participant. Finger movements were the ones that appeared most often. A new global hand movement around the wrist appeared in P1. For P6, movements that involved several fingers as a block before training (f2345), became individualized finger movements after training.

The number of cycles that could be performed in a row (i.e., without blocking) is reported in Figure 4b. Over all participants, the median (1st-3rd quartile) number of cycles before training was 6.3 (4.9-9.7) per movement and significantly increased to 10.0 (9.4-10.0) after training (p<0.03). The participants reached the number of 10 repetitions without blocking for 8 out of 10 new movements that had appeared during training. The mean duration of movements is shown for each type of movement in the left-hand part of Figure 4c, and as an average over all types of movement in the right-hand part of the same figure. Over all participants, the median duration dropped from 3.9 s (2.8-5.4 s) before to 1.8 s (1.0-2.7 s) after training (p<0.03). Durations for the new movements were not significantly different.

Figure 5 shows an example of the amplitude/peak-velocity relationship for one participant: a significant positive linear relationship exists both in Pre and Post training but with a higher slope in Post. Thus, for a given amplitude, the peak velocity was higher after training. The group analysis confirmed these observations as it shows a significant slope increase between Pre and Post from a median value of 0.7 (0.7- 0.9) to 1.2 (0.9-1.5) (p<0.03) (Figure 4d).

Although the Reproducibility Index (RI-Kin) in Fig.4e increased for many movements, globally it did not reach significance (p>0.05) mainly because of P3 whose averaged value decreased after training. The Fluidity Index (FI-Kin) increased from 0.7 (0.4-0.9) before training to 0.9 (0.8-1.5) after training (p<0.03) (Fig. 4f).
Fig. 4. Kinematic performances for participants P1 to P6 before (Pre) and after a 2-month home training period (Post), as well as the ones averaged over all movement types for the 6 participants (right side of the figure). (a) Number of different types of PHM. (b) Number of cycles before blocking. Note that results for different movement types are sometimes superimposed. (c) Movement duration. (d) Slope $\beta$ of amplitude/peak velocity relationship. For each participant, regressions were significant ($p<0.001$). (e) Reproducibility index (RI-Kin). (f) Fluidity index (FI-Kin). The lines connect the results for a given PHM obtained during Pre and Post sessions. * = $p<0.03$.

Interestingly, participant P5 – a below-elbow amputee – already had a higher number of different PHM and cycle repetitions than others before training. Yet, he still made progress in movement speed – which was shown by the decrease in cycle duration and the increase in $\beta$ value – and his movements became more fluid.

B. **EMG pattern**
The reproducibility index (RI-EMG) is shown in Figure 6a for each movement type and each participant for the Pre and Post conditions. Although all participants showed increased reproducibility after training for some movements, the group comparison showed no significant difference between the Pre and Post conditions (p>0.05) because one participant (P3) had a lower average value over movements after training (Fig. 6a). On the other hand, group comparison on FI-EMG, represented in Figure 6b, significantly increased after training from 0.2 (0.1-0.3) to 0.4 (0.2-0.7) (p<0.05). The results of classification analyses are shown in Figure 6c to 6e for all participants. Over all participants, the individual CA values varied between 0.6 and 1.0 before training and between 0.4 and 1.0 after training. Group comparison showed no significant difference in CA, MED and IDNN values between Pre and Post sessions, even when the types of PHM that had appeared during training (PostNew) were included.

**V. DISCUSSION**

The aim of this study was to examine whether daily PHM training at home for two months impacts EMG PR performance in the same way as it impacts the overall phantom mobilization ability, measured from the kinematic
data. Participants trained all their phantom movements daily with no other feedback than somatosensory information coming from the associated contractions in residual muscles. Our results showed a significant improvement in their overall phantom mobilization ability after training. This was expressed through the increased number of PHM as well as the higher speed, fluidity and, for all participants but P3, reproducibility of movements. Note that the number of cycles that could be performed in a row also increased after training, showing that susceptibility to fatigue decreased. However, although EMG bursts also became significantly smoother, none of the EMG metrics used for classification nor the reproducibility index (RI-EMG) improved significantly.

Our previous results showed a decrease in cycle duration after training [18] suggesting a faster phantom movement execution. This is confirmed and extended in the present study as the slope between amplitude and peak velocity increased with training: for a given amplitude, the peak velocity of movements became higher after training. This can be related to the fact that PHM became more fluid with training (i.e., executed in one shot without being interrupted by decelerations or stops) such that the peak velocity could reach a higher value. Interestingly, P6, who underwent a scapulothoracic disarticulation, showed kinematic characteristics close to those of above-elbow and below-elbow amputees, both before and after training. This needs yet to be confirmed through a real comparison between balanced groups of different levels of amputation. However, as far as we know, the kinematics of phantom movements had never been studied before in amputees with such high level of amputation.

Overall, our kinematic results show that phantom mobility improved with training, which should somehow correlate with modifications in EMG data. Indeed, as stated in the introduction section, the sensation of voluntary phantom limb movement is based on somatosensory feedback coming from movements of the residual muscles, skin, and other tissues caused by the associated muscle contractions [16], [19], [20]. Our results are consistent with this idea as, first, shorter movement durations come with shorter EMG bursts in the residual muscles after training (see time axis of Fig. 3 for an illustration). Second, adding the new types of PHM that the participants discovered during training as independent movements (PostNew) did not significantly decrease the overall classification accuracy. Adding movement classes to the classification would have decreased its success rate if the EMG patterns associated with the new movements had resembled the ones of the already existing classes. Therefore, our results suggest that these new types of PHM are associated to EMG patterns that are distinct from the ones of already existing movements. This is consistent with the fact that the participants reported to perceive these new types of PHM as undoubtedly different from the others. Thus, training of phantom mobility is indeed reflected in some aspects of the EMG signals.

Still, classification metrics did not significantly improve with training, meaning that there was no increase in the amount of usable information contained in the EMG data. This lack of increase in classification accuracy might be caused by the specificity of our method. First, the restricted number of EMG electrodes could limit our chances to observe topological changes in muscle contraction patterns. Second, surface EMG recording might not be able to detect deeper muscle activity. If deeper muscle contractions contribute to the somatosensory feedback giving rise to the sensation of phantom movements, we might miss their contribution to the recorded EMG signal and thus lower the success rate of the classifier. Third, PR is based on a selection of specific features that inherently leads to a possible loss of information contained in the EMG signals.

One might have expected the appearance of new movement types to have negatively influenced CA values, which would have partly explained our results. However, we believe that this is not the case here as the new movements appeared early in the training session (after 1 to 2 weeks). In addition, we have not seen any systematic relationship between CA outcomes and the appearance of new movements. For example, P2 didn’t discover new movements during training while for some movements, CA slightly decreased after training. Conversely, P3 had 3 new movements and its associated CA values mostly increased when comparing Pre and Post but also Pre and PostNew. Finally, if a confusion had occurred because of new movements, the overall interclass distances (IDNN) would have decreased for all participants but P2, which is not the case. So, overall, we believe that the new movements did not have any influence on CA.

It is important to note that before training, some CA values were already close or equal to one. One could think of a ceiling effect that would explain why training does not allow to significantly increase CA. However, by looking more closely at the data, it can be noticed that these high values only concern some types of movements for some participants. Even for P2 and P5, who display the highest values, some are still not higher than 95% which lets some room for improvement. In addition, if there were a ceiling effect, these CA values would have all remained constant with training which is not the case here as some high pre-training CA values in fact diminished after training. This decrease probably reflects the effect of confounding variables such as stress, climate change
and general fatigue, that have long been known to influence phantom sensations in general (and thus mobility) [37]. It is possible that their effect is strong enough to overcome the effect of our 2-month daily home training.

The present study is conducted without a control group. However, given that without training, the kinematic aspects of phantom mobility remained stable [18], the possible use of a control group in the present study would not have been useful. Indeed, even an increase in kinematic performance was not accompanied by an improved classification, so it seems unlikely that in a control group the classification would have evolved differently.

The type of training we chose was very different from the ones used in studies that show an increase in classification accuracy. Indeed, these studies used online visual feedback during training, that either came from an actual prosthesis [27] or from a virtual one [10], [28], [29], both inherently providing information on PR performance. If the movement of the prosthesis initiated by the decoder does not correspond to the requested movement, the user can adjust his motor command to better copy the input expected by the machine. Such movement adjustments induce modifications in the information contained in the EMG signals by changing the feature space and increasing inter-class distances. De Montalivet and colleagues [30] even used a pattern-similarity graphical biofeedback which gave a precise measure of the space-distance between the different EMG patterns specific to a movement type. They encouraged participants to increase this distance by finding a slightly different way of executing movements. Therefore, the most likely reason for the lack of increase in classification rate in the present study seems to be the absence of PR performance feedback during training. This hypothesis is supported by results of a study by He and colleagues [35] who, although focussing on static posture maintenance in mostly intact participants, did not find any improvement of classification accuracy after training without feedback on PR performance (constant “within-day error rates”). Irrespective of the exact technology or protocol used, additional feedback on PR performance could increase EMG classification accuracy of phantom movements.

The present study confirmed our previous results ([12], [13], [21]) that phantom mobility is classifiable, even before training, with overall classification success rates that are comparable to values found in the literature, using other methods in a clinical setting where amputees trained the PR controller using static muscle contractions, sometimes with additional instructions such as varying EMG amplitudes [31], [35]. If so, is phantom training at home useful? We believe it is. First, endurance of the participants is clearly enhanced by this type of training. This shows that regular use of phantom movements induces less fatigue than might have been expected from reports of amputees who had not trained their phantom movements (e.g.,[16], [22], [38], [39]). Training allows for a high level of endurance early in the rehabilitation process, i.e., even before the actual use of a prosthesis. Second, the EMG bursts became more fluid and often more reproducible. In real-life conditions, prosthesis users need the delay between their motor command and the action of the prosthesis to be as short as possible. Thus, the prosthesis must recognize a phantom movement from the start of its execution to imitate it immediately. The fact that the EMG has become more fluid and, for most of the participants, more reproducible, could prove to be an advantage when classification is performed in real time. Third, training increased the number of different types of phantom movements without penalizing classification performance. This potentially allows an increase in the number of DoF while keeping the prosthetic control intuitive. For these reasons, we encourage considering phantom mobility by training phantom movements as early as possible in the rehabilitation process.

Concerning prosthesis control, several steps must now be made before phantom mobility can be transferred to the clinical context. The most important step is the development of prostheses with independently controllable finger, wrist and elbow joints that are not based on predefined grips. Also, the number of EMG detection points in the prostheses should be increased and their positioning on the residual limb should be flexible (e.g., they should not be aligned as in the Coapt system). These steps will allow phantom-based control to be tested during actual prosthesis wear. Then, to refine the control of the prosthesis, the kinematics of a given phantom movement would have the be detected from the associated EMG pattern when amputees are using the prosthesis in daily life.

REFERENCES


