Influence of artifacts on movement intention decoding from EEG activity in severely paralyzed stroke patients

Eduardo López-Larraz[1,*], Carlos Bibián[1], Niels Birbaumer[1,2], Ander Ramos-Murguialday[1,3]

Abstract— **Brain-machine interfaces (BMI) can be used to control robotic and prosthetic devices for rehabilitation of motor disorders, such as stroke. The calibration of these BMI systems is of paramount importance in order to establish a precise contingent link between the brain activity related to movement intention and the peripheral feedback. However, electroencephalographic (EEG) activity, commonly used to build non-invasive BMIs, can be easily contaminated by artifacts of electrical or physiological origin. The way these interferences can affect the performance of movement intention decoders has not been deeply studied, especially when dealing with severely paralyzed patients, which often generate more artifacts by compensatory movements. This paper evaluates the effects of removing artifacts from the data used to train a BMI decoder on a dataset of 28 severely paralyzed stroke patients. We show that cleaning the training datasets reduces the global BMI performance for decoding attempts of movement. Further, we demonstrate that this performance drop especially affects the test trials contaminated by artifacts (i.e., trials that might not reflect cortical activity but noise), but not the clean test trials (i.e., trials representing correct cortical activity). This paper underlines the importance of cleaning the datasets used to train BMI systems to improve their efficacy for decoding movement intention and maximize their neurorehabilitative potential.**

I. Introduction

Non-invasive brain-machine interfaces (BMI) constitute a feasible and relatively cheap way to control external devices for communication and motor restoration using brain activity only [1]. Thanks to this technology, patients with severe motor deficits have managed to control robotic or prosthetic systems and used them to facilitate the movement of their paralyzed arms or legs [2], [3], [4]. Recent studies have evidenced that the contingent association between the brain signatures of movement and proprioceptive peripheral feedback promotes neural plasticity [5]. Further, rehabilitative interventions based on this methodology have shown that functional recovery can be achieved after several weeks or months of training both in SCI and stroke patients [6], [7].

The electroencephalogram (EEG) constitutes an especially relevant tool for decoding motor commands from the brain, as has been demonstrated in the aforementioned studies. By placing electrodes on the surface of the scalp, in the regions above the motor cortex, it allows the real-time acquisition of the cortical activity generated in response to movement intentions. Paralyzed patients such as stroke and SCI suffer a cortical reorganization that generally makes them produce a pathologic neural activation during their movement intentions [8], [9], [10], although it is still possible to decode different movements from their EEG activity [11], [12], [13], [14]. However, one of the main limitations of EEG technology is its low signal to noise ratio, and its easiness to get contaminated by interferences of electrical or physiological origin: i.e., artifacts.

Different techniques have been proposed to remove artifacts from EEG datasets used to train motion intention decoders [15]. We can divide those techniques into two groups: the ones that filter (or clean) the artifacts, preserving the amount of data; and the ones that detect the artifactual trials and discard them. The first group includes methods like linear regression or independent component analysis (ICA) to reconstruct EEG data after removing artifacts–such as eye movements [16], [17]–or median filters to remove external electric/magnetic contaminations [18]. The second type is generally based on the use of statistical distributions to eliminate those segments of signal that exceed a certain threshold on a given parameter (e.g., amplitude, power on a frequency band, etc) [19]. Despite many studies and reviews have addressed the issue of how to minimize the artifacts for EEG-based BMIs, they generally evaluate the success of the methods on the resulting signals only. However, the influence that removing artifacts has for the control of BMIs that decode movement commands is rarely studied, more so when considering data recorded with paralyzed patients, who generally produce higher number of artifacts due to compensatory activity.

This paper evaluates the effect of removing artifacts from EEG datasets used to calibrate a BMI decoder of movement intentions. The study is performed with a clinically relevant dataset that includes EEG recordings of 28 stroke patients while they move their healthy hand, or try to move their paretic hand. An automatized procedure is utilized to eliminate the three main sources of artifacts on EEG datasets: eye movements, motion artifacts, and muscular contamination. The performances of a BMI decoder implemented in a pseudo-online manner are compared after the application or not of the artifact removal procedure to the training dataset.

II. Methods

A. Dataset

Twenty eight chronic stroke patients (16 male, age 47.9±13.2, time since stroke 64.9±59.4 months) were involved in this study. From the 28 patients, 14 had the stroke in the left hemisphere and 14 in the right one. Regarding the typology of lesion, 14 of them presented mixed and 14 had

[1]Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Silcher Str. 5, 72076, Tübingen, Germany.
[2]Wyss Center for Bio and Neuroengeneering, Geneva, Switzerland.
[3]TECNALIA Health Technologies, San Sebastian, Spain.
[*]Correspondence: eduardo.lopez-larraz@uni-tuebingen.de

subcortical lesions. All the patients suffered hand paralysis resulting in no residual finger extension in the paretic arm, and no other psychiatric or neurological condition apart from the stroke. More details about the inclusion criteria, exclusion criteria, clinical and demographic data can be seen elsewhere [6]. The experiments were conducted at the University of Tübingen, Germany. The experimental procedure was approved by the ethics committee of the Faculty of Medicine of the University of Tübingen, and all the patients provided written informed consent before participation.

The patients were invited to attend to one experimental session, in which their electroencephalographic (EEG) activity was recorded while they opened/closed their healthy hand or attempted to open/close their paralyzed hand. Each patient performed between 4 and 6 blocks, each of which included, in a random order, 17 trials moving the healthy hand, and 17 trials attempting to move the paralyzed one. The trials included audiovisual cues to instruct the patients about the three phases: rest (5 sec), motor execution/attempt (5 secs), inter-trial interval (random duration between 3-4 sec). In the motor execution/attempt interval, the patients were asked to open and close the indicated hand at their own comfortable pace.

EEG activity was recorded with a 16-electrodes Acticap system (BrainProducts GmbH, Germany) from Fp1, Fp2, F3, Fz, F4, T7, C3, Cz, C4, T8, CP3, CP4, P3, Pz, P4, and Oz locations, with the ground in AFz and reference in FCz (according to the international 10/10 system). Vertical and horizontal electrooculographic (EOG) derivations were also recorded to capture eye movements. Both signals were sampled synchronously at 500 Hz.

B. BMI Design

Despite the dataset studied in this paper was analyzed offline, all the procedures were performed simulating an online scenario. Sliding windows, causal filters, and auto-regressive models were used in order to obtain realistic results that can be replicated in a real-time setting.

A block-based *N*-fold cross validation procedure was applied, where *N* corresponded to the number of blocks recorded for each patient. All the procedures were applied separately for the healthy and the paretic arms. In each fold, the data of the corresponding test block was kept apart, and the training dataset was constructed with the rest of the blocks. All the specific procedures involving data-distribution parameters were performed on the training data only (e.g., artifact removal or feature normalization). Subsequently, when applicable, the parameters computed on the basis of the training data were used on the test trials (e.g., normalization and EOG regression).

C. Artifact removal

The artifact removal method was applied only to the data used for training the BMI. This simulates a scenario in which, prior to starting a BMI therapeutic session, we have recorded some data for calibration and we have to decide if we perform an artifact rejection method to the data or not.

The method to remove the artifacts is designed to eliminate the three main sources of contamination in EEG data: *(i)* eye movements; *(ii)* motion artifacts; and *(iii)* muscular artifacts. Ocular artifacts are due to the electrical currents generated by blinks or vertical and lateral movements of the eyes. Motion artifacts are low-frequency oscillations that are due to head or whole-body movements, and cause an increase in signal power that can be easily confused with cortical activity [20]. Muscular artifacts are high-frequency interferences due to the activation of cranial or neck muscles, which lead to high-power in beta and gamma frequency bands.

1) Removal of ocular artifacts: A linear regression was used to remove the components of vertical and horizontal eye movements affecting the EEG channels, as proposed in [16]. Here is assumed that the recorded signal is a linear combination between the actual EEG activity and the contamination coming from the eyes:

$$X_{c,t} = EEG_{c,t} + b_{c,d} \cdot EOG_{d,t} \tag{1}$$

where $X_{c,t}$ is the acquired signal in c channels during t time samples; $EEG_{c,t}$ corresponds to the clean EEG activity; $EOG_{d,t}$ are the d EOG derivations; and b are the coefficients that represent the degree of contamination of each EEG channel by the EOG activity. The coefficients b can be estimated as the product between the auto-covariance matrix of the EOG derivations and the cross-covariance between the EEG and the EOG [16]. The estimation can be done from a training segment of data, and used to clean such training segment and new unseen data. The clean signal is calculated as:

$$EEG = X - b \cdot EOG \tag{2}$$

With this method, the influence of EOG activity is minimized but the amount of data is preserved.

2) Removal of motion and muscular artifacts: Motion and muscular artifacts cannot be so easily removed by using linear regression as eye movements, and more complex techniques have been proposed to try to minimize their influence [15]. Despite independent component analysis (ICA) has been used to reduce these types of contamination, there is evidence showing that it can introduce artificial variance to the signals and bias the classification performance [21]. Therefore, we used a variance-based method to completely remove the trials contaminated by head/body motions (low frequency) or muscular activations (high frequency). First, the power in delta ([0.1-4] Hz) and gamma ([30-48] Hz) frequency bands was calculated for the rest and movement intervals of each trial. The mean and standard deviation (std) of the values calculated during the rest intervals were used to define the first rejection threshold, set as 3 stds above the mean delta and gamma power. All the trials exceeding this threshold either in the rest or in the movement intervals were marked for rejection. Subsequently, we recalculated the mean and std of the non-rejected trials and also rejected all those that exceeded the newly calculated threshold (i.e., 3 stds above the new mean) during rest or movement intervals in any of the bands.

The removal of artifacts was done considering only the channels that were used by the BMI (i.e., channels for feature extraction and for re-referencing, see Subsection II-D below). All the thresholds calculated on each fold for each channel to remove the artifacts were saved for an additional analysis in which we separated the test data between clean and contaminated trials to evaluate if they were classified differently.

D. Feature extraction and training

Only the electrodes placed on the contralateral centro-parietal cortex with respect to the moved (or attempted) limb were considered: i.e., C3, CP3, and P3 when considering the right arm, and C4, CP4, and P4 when considering the left arm. The signals were bandpass filtered between 0.1 and 48 Hz with a 4-th order Butterworth filter. The selected electrodes were re-referenced using Laplacian derivations (for the electrodes whose 4 closest neighbors were not available, we considered the 2 or 3 available neighbors only).

The signals were divided into 7-second trials, from -3 to +4 seconds, with 0 being the presentation of the cue that instructed the patients to move. We extracted features from one-second time windows to model the *rest* and the *movement* classes. The windows of the *rest* class were extracted from the time interval [-2, 0] s, and the windows of the *movement* class were extracted from the interval [1, 3] s, with a sliding step of 0.25 s.

For each one-second EEG window we computed the average power in the alpha ([7-13] Hz) and the beta ([14-30] Hz) bands of the 3 contralateral channels (i.e., feature vectors corresponding to each window included six values). To compute the power spectral density, we used an order-20 autoregressive model based on the Burg algorithm [22]. All the feature vectors belonging to the *rest* and to the *movement* classes were normalized to have zero mean and unit variance. Subsequently, these vectors were fed to an SVM classifier with a radial basis function (RBF) kernel to build the model that best discriminates between both classes.

E. Classification

The evaluation of the test trials was done in a pseudo-online manner, simulating a real-time use of the BMI. The features were extracted as explained in Section II-D, and the parameters computed on the basis of the training dataset were used to normalize the test features and to compute the regression for EOG correction (when this method was applied). The trials were evaluated with the one-second sliding window from -3 to +4 s (notice that the first output was generated at t = -2 s), with the classifier providing an output every 20 ms.

We plotted the average outputs of the decoder to analyze its behavior during the rest and the movement periods. The three main metrics to quantify the BMI performance were: true positives (TP), false positives (FP) and average accuracy (ACC). The TP measure the success of the classifier during the movement period, which was considered as the interval [1, 4] s. The FP measure the errors of the classifier during the rest interval [-2, 0] s. The average accuracy was computed as the mean between the TP and the true negatives (i.e., 1-FP), and was used as a threshold-independent metric to deal with coupled TP and FP differences that are due to an offset in the classifier output, but do not translate into a better average performance.

F. Statistical analysis

Two analyses were performed to quantify the impact of artifacts on BMI performance. First, we compared the two calibration approaches (i.e., training with all the trials without removing artifacts, or training after applying the artifact removal method) for classifying the same test data. Secondly, we divided the test data into clean and contaminated trials, and evaluated those two datasets with both calibration methods. We used non-parametric statistical tests and false discovery rate (FDR) correction for multiple comparisons, and considered as statistically significant those p-values smaller than 0.05 after correction.

The influence of the artifact removal applied to the training was evaluated separately for each hand by comparing the TP, FP and ACC (two tailed Wilcoxon signed-rank test, FDR corrected for the three comparisons). For the comparison between classification performances of clean and contaminated test datasets we compared only the ACC values, due to simplicity reasons. We evaluated, given a fixed test dataset (e.g., clean or contaminated test trials), the influence of the two calibration methods; and given a fixed calibration method (e.g., without or with artifact removal), the influence of the test dataset. For the results of each hand, two tailed Wilcoxon signed-rank tests were used with FDR correction for four comparisons (i.e., two training procedures and two test datasets).

III. Results

A. Influence of artifact removal applied to the training data

The average number of trials available for training the BMI on each fold was 59.2 ± 12.3 both for the healthy and for the paretic arm. The artifact rejection method eliminated, on average, $28.6 \pm 10.0\%$ of the trials corresponding to the healthy arm and $39.5 \pm 20.7\%$ for the paretic arm.

Figure 1 shows the average performance achieved for the healthy hand (left) and the paretic hand (right) when applying or not the artifact removal method to the dataset used to train the BMI decoder. For the healthy hand, both the percentage of TP and FP were significantly lower when removing artifacts from the training data ($p = 0.0002$ and $p = 0.0004$, respectively). However, this behavior resembles an offset in the probability outputs of the classifier (see Fig. 1-left), a problem that can be solved online by adjusting the classifier thresholds. In fact, the average accuracy was not significantly different between both methods ($p > 0.05$). For the attempt of movement of the paretic hand, the TP were significantly lower when applying the artifact removal ($p = 0.0002$), but the FP were not significantly different ($p > 0.05$), which resulted in a significantly lower accuracy also ($p < 0.042$).

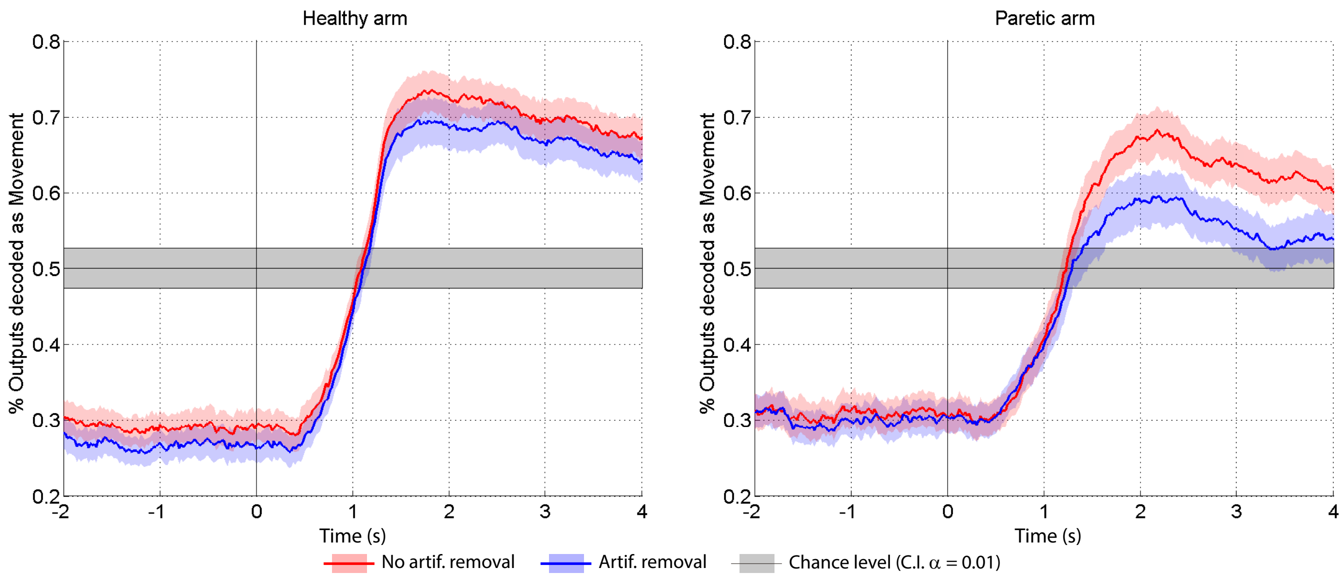

Fig. 1. Average decoding accuracies for the healthy (left) and the paretic (right) arms for all the test trials. On each panel, the lines represent the average of all the patients, and the shades indicate the standard error of the mean. The red color corresponds to the results of the classifier trained without removing artifacts, while the blue color represents the results of the classifier trained after removing artifacts. The shaded gray area indicates the confidence interval of the chance level ($\alpha = 0.01$), computed on the basis of all the test trials, according to [23].

B. *Influence of artifacts in testing*

The analysis presented in section III-A corresponds to a simulation in which different datasets are used to train the BMI and used pseudo-online to classify all the available test trials, as it would be done in real-time. For the current analysis, we separated the test data between clean and contaminated trials. The thresholds to define a trial as contaminated were obtained from the training dataset used on each fold (see Section II-C.2). For the healthy arm, $40.8 \pm 11.4\%$ of the test trials were marked as contaminated, while for the paretic arm the percentage of contaminated trials was $49.1 \pm 19.9\%$.

Figure 2 shows the decoding results of the classifiers, trained without or with artifact removal, applied to test datasets consisting of clean or contaminated trials. Table I shows the accuracies for all the possible combinations of training (without or with artifact removal) and test (clean or contaminated trials). Here, we used statistical tests to analyze two situations: *(i)* the influence of the training dataset (i.e., without or with artifact removal) on a fixed test dataset (e.g., in clean trials or in contaminated trials); and *(ii)* the influence of the test data (i.e., clean trials or contaminated trials) for a fixed training dataset (e.g., trained without or trained with artifact removal).

For the healthy hand, the comparison between both training procedures revealed that, for the clean trials, there were no significant differences between both procedures (comparison of both solid lines in Figure 2-left, $p > 0.05$), while for the contaminated trials, the classifier trained without removing artifacts provided significantly higher accuracies (comparison of both dotted lines in Figure 2-left, $p = 0.015$). The comparison between the two test datasets for a fixed training procedure showed that the clean trials were classified significantly better than the contaminated trials, both for the decoder trained without artifact removal (comparison of solid and dotted red lines in Figure 2-left, $p = 0.0005$), and for the decoder trained with artifact removal (comparison of solid and dotted blue lines in Figure 2-left, $p = 0.0003$).

When analyzing the decoding of movement attempt of the paralyzed hand, the clean trials were again similarly classified by both training procedures (comparison of both solid lines in Figure 2-right, $p > 0.05$), and the contaminated trials were significantly better decoded by the classifier trained without removing artifacts (comparison of both dotted lines in Figure 2-right, $p = 0.049$). The comparison between the two test datasets for a fixed training procedure showed that, for the decoder trained without artifact removal, there were no significant differences between clean and contaminated trials (comparison of solid and dotted red lines in Figure 2-right, $p > 0.05$), while for the decoder trained with artifact removal, the clean trials were classified significantly better (comparison of solid and dotted blue lines in Figure 2-right, $p < 0.047$).

IV. Discussion and Conclusions

When designing a neurorehabilitation intervention based on brain-machine interfaces (BMI), several methodological and experimental components have to be carefully designed. In this paper we focused on the stage when we have to calibrate the BMI with examples of brain activity of rest and movement (execution or attempt), and we have to decide if we clean these training examples to remove potential artifacts that might bias the decoding performance. We studied the effects of artifacts contaminating EEG activity on the decoding of movement and movement intention in severely paralyzed stroke patients. The major outcome of this work

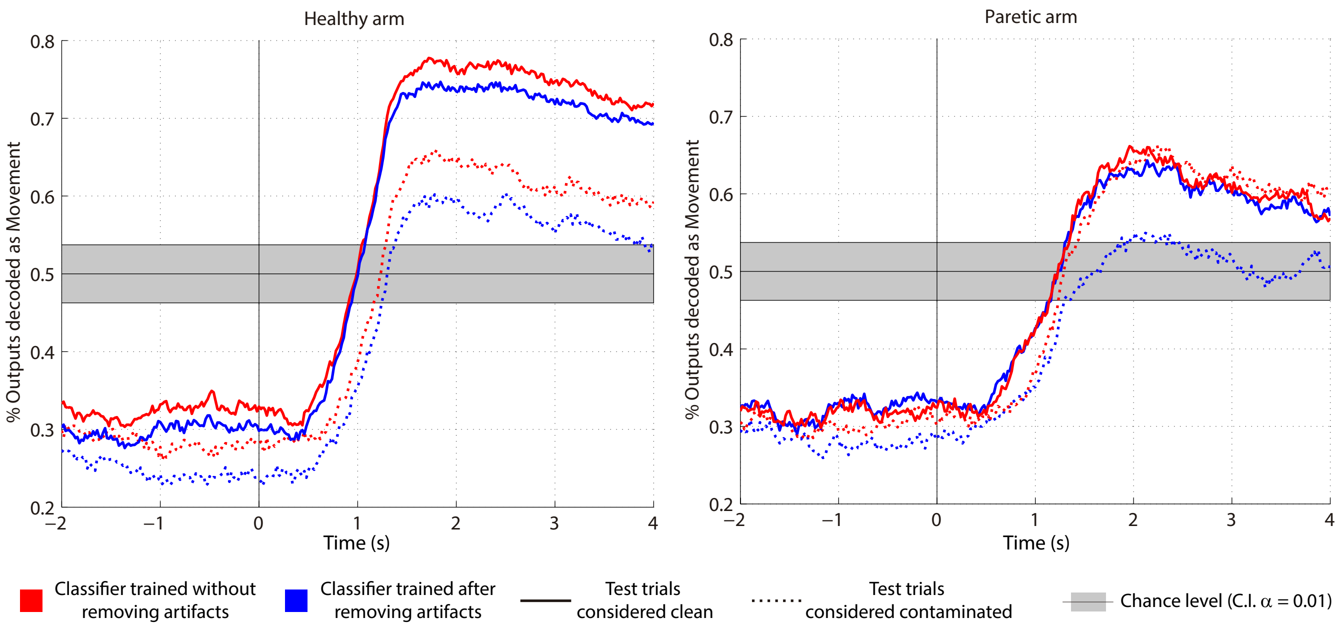

Fig. 2. Decoding accuracies separated for clean and contaminated trials for the healthy (left) and paretic (right) arms. The red color corresponds to the results of the classifiers trained without removing artifacts, while the blue color represents the results of the classifiers trained after removing artifacts. Solid and dotted lines correspond to clean and contaminated trials, respectively. The shaded gray area indicates the confidence interval of the chance level ($\alpha = 0.01$), computed on the basis of the minimum number of test trials for all the cases, according to [23].

is an empirical demonstration of how the EEG artifacts can affect the performance of brain-machine interfaces (BMI) from the training and the testing perspectives, which may have a negative impact on the use of neuro-robotic and neuro-prosthetic systems for stroke rehabilitation.

We observed that eliminating artifacts from the training datasets leads to significantly lower accuracy values when decoding the attempt of movements of the paretic arm, although it does not affect the decoding of movements of the healthy arm. This has two major practical implications for the development and improvement of clinical BMIs. Firstly, the fact that the results for the healthy and paretic arm are different implies that some conclusions extracted from preliminary studies with healthy subjects might not be directly applicable to clinical populations. This underlines the need of investigating how to optimize BMI systems with the potential end-users of the technology, and not only with healthy participants [24]. Secondly, even though training with cleaner EEG signals provided lower accuracy values for the decoding of motion intentions, this does not necessarily mean that the efficacy of the BMI is lower. Indeed, the higher performances achieved when having contaminated signals in the training dataset might not reflect the ability of the system to decode cortical activation, but the artifacts present in the test data. This would cause that the BMI reinforces abnormal patterns of activity, potentially inducing maladaptive plasticity [25].

In this line, we studied the test datasets, separating them between clean and contaminated trials. The relevant result here was again a difference between the healthy and paretic hands. For the healthy hand, the clean trials were always classified better than the contaminated ones (regardless of having or not artifacts within the training data). This might reflect that the artifacts occurring during a healthy movement are not so frequent, and more importantly, they are of smaller magnitude and do not influence much how the classifier models the *rest* and *movement* classes. Conversely, for the paretic arm, only having a clean training dataset led to significantly better accuracies for the clean than for the contaminated trials (which are more likely to represent non-cortical activity). A possible explanation for this could be the extra effort required to try to elicit paretic movements, which generates more and bigger artifacts (head and eye movements, or neck and cranial muscle contractions). The main consequence of this is that if we train the BMI decoder using signals that include artifacts, we will have a biased increase in performance that does not reflect the accuracy of the system to decode correct cortical activation.

Despite the promising results shown so far to induce functional recovery with BMI [6], there is still a great margin of improvement for this technology before it can be included in the portfolio of standard treatments that therapists can use for motor rehabilitation [26]. The results presented in this work provide relevant insights for the development of clinical BMIs for stroke patients. Currently, there is a general trend in the community aiming at the enhancement of performance of movement intention decoders, using advanced machine learning and signal processing strategies. However, BMI-triggered rehabilitation therapies rely on Hebbian learning, assuming that the association between the brain activation and the feedback will promote neuroplasticity. Therefore, it is reasonable to think that not only it is important to achieve

TABLE I

ACCURACIES FOR ALL THE COMBINATIONS OF TRAINING AND TEST DATASETS.

	Healthy arm		**Paretic arm**	
	Training dataset without artif. removal	**Training dataset with artif. removal**	**Training dataset without artif. removal**	**Training dataset with artif. removal**
Clean trials	71.76 ± 11.74	71.90 ± 11.57	64.59 ± 10.62	63.92 ± 11.88
Contaminated trials	67.37 ± 10.56	66.50 ± 10.57	64.90 ± 11.83	61.03 ± 10.59

high performances, but to have precise and accurate decoding of the relevant cortical activation, which would improve the quality (more than the quantity) of the feedback. As shown in this paper, the use of artifact-free data to train the BMI decoders can be very relevant to improve the contingent link between brain and muscles, and to intensify the long term rehabilitative effects of this technology.

ACKNOWLEDGMENT

This study was funded by the Baden-Württemberg Stiftung (GRUENS ROB-1), the Deutsche Forschungsgemeinschaft (DFG, Koselleck), the Fortüne-Program of the University of Tübingen (2422-0-0), and the Bundesministerium für Bildung und Forschung BMBF MOTORBIC (FKZ 13GW0053) and AMORSA (FKZ 16SV7754).

REFERENCES

[1] U. Chaudhary, N. Birbaumer, and A. Ramos-Murguialday, "Brain-computer interfaces for communication and rehabilitation," *Nature Reviews Neurology*, vol. 12, no. 9, pp. 513–525, 2016.

[2] G. Pfurtscheller, C. Guger, G. Müller, G. Krausz, and C. Neuper, "Brain oscillations control hand orthosis in a tetraplegic," *Neuroscience Letters*, vol. 292, no. 3, pp. 211–214, 2000.

[3] G. Pfurtscheller, G. R. Müller, J. Pfurtscheller, H. J. Gerner, and R. Rupp, "'Thought' - Control of functional electrical stimulation to restore hand grasp in a patient with tetraplegia," *Neuroscience Letters*, vol. 351, no. 1, pp. 33–36, 2003.

[4] E. López-Larraz, F. Trincado-Alonso, V. Rajasekaran, A. J. Del-Ama, J. Aranda, J. Minguez, Á. Gil-Agudo, and L. Montesano, "Control of an ambulatory exoskeleton with a brain-machine interface for spinal cord injury gait rehabilitation," *Frontiers in Neuroscience*, vol. 10, no. 359, 2016.

[5] N. Mrachacz-Kersting, N. Jiang, A. J. T. Stevenson, I. K. Niazi, V. Kostic, A. Pavlovic, S. Radovanovic, M. Djuric-Jovicic, F. Agosta, K. Dremstrup, and D. Farina, "Efficient neuroplasticity induction in chronic stroke patients by an associative brain-computer interface." *Journal of neurophysiology*, vol. 115, no. 3, pp. 1410–1421, 2016.

[6] A. Ramos-Murguialday, D. Broetz, M. Rea, L. Läer, O. Yilmaz, F. L. Brasil, G. Liberati, M. R. Curado, E. Garcia-Cossio, A. Vyziotis, W. Cho, M. Agostini, E. Soares, S. Soekadar, A. Caria, L. G. Cohen, and N. Birbaumer, "Brain-machine interface in chronic stroke rehabilitation: a controlled study." *Annals of neurology*, vol. 74, no. 1, pp. 100–108, 2013.

[7] A. R. C. Donati, S. Shokur, E. Morya, D. S. F. Campos, R. C. Moioli, C. M. Gitti, P. B. Augusto, S. Tripodi, C. G. Pires, G. A. Pereira, F. L. Brasil, S. Gallo, A. A. Lin, A. K. Takigami, M. A. Aratanha, S. Joshi, H. Bleuler, G. Cheng, A. Rudolph, and M. A. L. Nicolelis, "Long-Term Training with a Brain-Machine Interface-Based Gait Protocol Induces Partial Neurological Recovery in Paraplegic Patients," *Scientific Reports*, vol. 6, p. 30383, 2016.

[8] V. Kaiser, I. Daly, F. Pichiorri, D. Mattia, G. R. Müller-Putz, and C. Neuper, "Relationship between electrical brain responses to motor imagery and motor impairment in stroke," *Stroke*, vol. 43, pp. 2735–2740, 2012.

[9] C. Tangwiriyasakul, R. Verhagen, W. L. C. Rutten, and M. J. A. M. van Putten, "Temporal evolution of event-related desynchronization in acute stroke: a pilot study." *Clinical neurophysiology*, vol. 125, no. 6, pp. 1112–1120, 2014.

[10] E. López-Larraz, L. Montesano, Á. Gil-Agudo, J. Minguez, and A. Oliviero, "Evolution of EEG motor rhythms after spinal cord injury: a longitudinal study," *PLoS ONE*, vol. 10, no. 7, p. e0131759, 2015.

[11] E. López-Larraz, J. M. Antelis, L. Montesano, Á. Gil-Agudo, and J. Minguez, "Continuous decoding of motor attempt and motor imagery from EEG activity in spinal cord injury patients." in *34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, San Diego, 2012, pp. 1798–1801.

[12] J. Ibáñez, J. I. Serrano, M. D. del Castillo, E. Monge-Pereira, F. Molina-Rueda, I. Alguacil-Diego, and J. L. Pons, "Detection of the onset of upper-limb movements based on the combined analysis of changes in the sensorimotor rhythms and slow cortical potentials." *Journal of neural engineering*, vol. 11, no. 5, p. 056009, 2014.

[13] J. M. Antelis, L. Montesano, A. Ramos, N. Birbaumer, and J. Minguez, "Decoding Upper Limb Movement Attempt from EEG Measurements of the Contralesional Motor Cortex in Chronic Stroke Patients." *IEEE transactions on biomedical engineering*, vol. 64, pp. 99–111, 2017.

[14] E. López-Larraz, A. M. Ray, T. C. Figueiredo, C. Bibián, N. Birbaumer, and A. Ramos-Murguialday, "Stroke lesion location influences the decoding of movement intention from EEG," in *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017.

[15] J. A. Urigüen and B. Garcia-Zapirain, "EEG artifact removalstate-of-the-art and guidelines," *Journal of Neural Engineering*, vol. 12, 2015.

[16] A. Schlögl, C. Keinrath, D. Zimmermann, R. Scherer, R. Leeb, and G. Pfurtscheller, "A fully automated correction method of EOG artifacts in EEG recordings," *Clinical Neurophysiology*, vol. 118, no. 1, pp. 98–104, 2007.

[17] M. A. Klados, C. L. Papadelis, and P. D. Bamidis, "REG-ICA: A new hybrid method for EOG artifact rejection," in *Final Program and Abstract Book - 9th International Conference on Information Technology and Applications in Biomedicine, ITAB 2009*, 2009.

[18] A. Insausti-Delgado, E. López-Larraz, C. Bibián, Y. Nishimura, N. Birbaumer, and A. Ramos-Murguialday, "Influence of trans-spinal magnetic stimulation in electrophysiological recordings for closed-loop rehabilitative systems," in *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2017.

[19] H. Nolan, R. Whelan, and R. B. Reilly, "FASTER: Fully Automated Statistical Thresholding for EEG artifact Rejection," *Journal of Neuroscience Methods*, vol. 192, no. 1, pp. 152–162, 2010.

[20] T. Castermans, M. Duvinage, G. Cheron, and T. Dutoit, "About the cortical origin of the low-delta and high-gamma rhythms observed in EEG signals during treadmill walking," *Neuroscience Letters*, 2014.

[21] M. B. Pontifex, K. L. Gwizdala, A. C. Parks, M. Billinger, and C. Brunner, "Variability of ICA decomposition may impact EEG signals when used to remove eyeblink artifacts," *Psychophysiology*, 2016.

[22] J. P. Burg, "Maximum Entropy Spectral Analysis," in *Proceedings of the 37th Annual International SEG Meeting*, 1975.

[23] G. R. Müller-Putz, R. Scherer, C. Brunner, R. Leeb, and G. Pfurtscheller, "Better than random? A closer look on BCI results," *International Journal of Bioelectromagnetism*, 2008.

[24] E. López-Larraz, F. Trincado-Alonso, and L. Montesano, "Brain-machine interfaces for motor rehabilitation: is recalibration important?" in *14th International Conference on Rehabilitation Robotics (ICORR)*, Singapore, 2015, pp. 223–228.

[25] K. A. Moxon, A. Oliviero, J. Aguilar, and G. Foffani, "Cortical reorganization after spinal cord injury: Always for good?" *Neuroscience*, vol. 283, pp. 78–94, 2014.

[26] G. Asín Prieto, R. Cano-de-la Cuerda, E. López-Larraz, J. Metrot, M. Molinari, and L. E. H. van Dokkum, "Emerging perspectives in stroke rehabilitation," in *Emerging Therapies in Neurorehabilitation*. Berlin: Springer Berlin Heidelberg, 2014, vol. 1, pp. 3–21.