

Prediction of Pathological Tremor Signals Using Long Short-Term Memory Neural Networks

Alejandro Pascual-Valdunciel, Víctor Lopo-Martínez, Rafael Sendra-Arranz, Miguel González-Sánchez, Javier Ricardo Pérez-Sánchez, Francisco Grandas, Diego Torricelli, Member IEEE, Juan C. Moreno, Senior Member *IEEE*, Filipe Oliveira Barroso, Member *IEEE*, José L. Pons, Senior Member *IEEE*, Álvaro Gutiérrez, Senior Member *IEEE*

Abstract—Previous implementations of closed-loop peripheral electrical stimulation (PES) strategies have provided evidence about the effect of the stimulation timing on tremor reduction. However, these strategies have used traditional signal processing techniques that only consider phase prediction and might not model the non-stationary behavior of tremor. Here, we tested the use of long short-term memory (LSTM) neural networks to predict tremor signals using kinematic data recorded from Essential Tremor (ET) patients. A dataset comprising wrist flexion-extension data from 12 ET patients was pre-processed to feed the predictors. A total of 180 models resulting from the combination of network (neurons and layers of the LSTM networks, length of the input sequence and prediction horizon) and training parameters (learning rate) were trained, validated and tested. Predicted tremor signals using LSTM-based models presented high correlation values (from 0.709 to 0.998) with the expected values, with a phase delay between the predicted and real signals below 15 ms, which corresponds approximately to 7.5% of a tremor cycle. The prediction horizon was the parameter with a higher impact on the prediction performance. The proposed LSTM-based models were capable of predicting both phase and amplitude of tremor signals outperforming results from previous studies (32-56% decreased phase prediction error compared to the out-of-phase method), which might provide a more robust PES-based closed-loop control applied to PES-based tremor reduction.

Index Terms—machine learning, LSTM, peripheral electrical stimulation, tremor prediction, essential tremor.

This work was funded by the European Union's Horizon 2020 research and innovation programme (Project EXTEND—Bidirectional Hyper-Connected Neural System) under grant agreement No 779982.

A. Pascual-Valdunciel is with the Neural Rehabilitation Group, Cajal Institute, Spanish National Research Council (CSIC), Madrid, Spain, the Department of PM&R, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States, and the E.T.S. Ingenieros de Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain (e-mail: alejandropv@cajal.csic.es).

M. González-Sánchez and J.R. Pérez-Sánchez are with the Movement Disorders Unit, Department of Neurology, Hospital General Universitario Gregorio Marañón, Madrid, Spain (e-mail: miguel_gs_lk@hotmail.com; javineuro@hotmail.com).

F. Grandas is with the Movement Disorders Unit, Department of Neurology, Hospital General Universitario Gregorio Marañón, Madrid,

I. Introduction

Pathological tremors are defined as involuntary rhythmic movements of one or more body parts caused by neural disorders such as Essential Tremor (ET), Parkinson's Disease or cerebellar ataxias, among others [1]. It is not only considered to be the most common motor disorder worldwide, but also to become a disabling condition hindering patients from performing activities of daily living (ADLs) [2]. Particularly, ET is the main neural disorder leading to pathological tremor, and its incidence and severity of symptoms increase with age, contributing to the worldwide aging challenge that society and healthcare systems must face in the next decades [3].

Currently, there is no known cure for ET. The available therapeutic solutions are insufficient and only address the symptoms [4]. Pharmacological treatments are the first line of treatment. However, between 30% and 50% of the ET population is non-responsive to medication or decides to discontinue the treatment due to the adverse effects [5]. Surgical interventions such as Deep Brain Stimulation or high-intensity focused ultrasound are considered for drug-resistant patients with severe tremors [6]. As a counterpart, all surgical approaches require moderate hazardous procedures, which might result in severe side effects, preventing a large number of ET patients from receiving these treatments [7]. Therefore, in

Spain, and the Department of Medicine, Universidad Complutense, Madrid, Spain (e-mail: francisco.grandas@salud.madrid.org).

D. Torricelli, J. C. Moreno and F. O. Barroso are with the Neural Rehabilitation Group, Cajal Institute, Spanish National Research Council (CSIC), Madrid, Spain. (e-mail: diego.torricelli@csic.es; jc.moreno@csic.es; filipe.barroso@cajal.csic.es).

J. L. Pons is with the Legs & Walking AbilityLab, Shirley Ryan AbilityLab, Chicago, IL, United States; the Department of Biomedical Engineering and Mechanical Engineering, McCormick School of Engineering, Northwestern University, Chicago, IL, United States; and the Department of PM&R, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States (e-mail: jpons@sralab.org).

V. Lopo-Martínez, R. Sendra-Arranz and A. Gutiérrez are with E.T.S. Ingenieros de Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain (e-mail: victor.lopo.martinez@alumnos.upm.es; r.sendra@upm.es; a.gutierrez@upm.es).

the last two decades there has been a rising interest in developing alternative solutions to pharmacological and surgical therapies based on mechanical assistive devices or neuroprostheses [8].

Among these emerging techniques, peripheral electrical stimulation (PES) of afferent pathways has shown acute and short-term tremor reduction in several studies with minimum adverse effects [9]. PES of afferent pathways, also called sensory PES, comprises the application of electrical currents below motor threshold to recruit the afferent fibers driving sensory information to the central nervous system. Several studies hypothesize that modulation of afferent fibers disrupts the tremor oscillations, immediately preventing them from reaching the muscles. The first evidence on the effects of PES of afferent pathways on tremor reduction were reported by Dosen et al. [10], who hypothesized that the timely stimulation of afferent Ia fibers could activate reciprocal inhibition circuits at the spinal cord and acutely reduce the supraspinal tremor contribution to the antagonist motor output [11], [12]. In order to synchronize the electrical stimulation with the tremorgenic activity, a closed-loop control algorithm called out-of-phase strategy was applied [10]. This strategy is based on the use of sequential electromyography (EMG) recordings and stimulation windows of a pair of antagonist muscles (*e.g.*, flexor and extensor muscles of the wrist) (Fig. 1(a)). During the recording windows, the EMG signal is demodulated to extract the envelope of the tremor component by applying a band-pass filter (3-12 Hz) to the iterative Hilbert transform of the rectified signal [13]. After that, the peaks of the demodulated signal are identified and, if they are above a certain threshold, it is assumed that tremor is present, and the mean inter-burst interval (MIBI) is computed and used to predict the next tremor bursts. Finally, the stimulation window is enabled, and the electrical pulses are alternately delivered to the antagonist muscles synchronized with the predicted tremor bursts. Other groups have followed a similar approach by synchronizing the stimulation with the predicted physiological tremor activity measured through the EMG or kinematics [14]–[16].

Although tremor frequency is highly variable across pathologies, patients or disease progress, it is usually stable for the same patient along the same disease state [17]. The frequency band of pathological tremor is typically defined between 4 and 12 Hz, therefore tremor can be separated from most voluntary movements by means of frequency-based filtering techniques [18]. In practice, pathological tremors are nonstationary and nonlinear signals, mixed with other voluntary movements, challenging the signal processing paradigm [19]. Current out-of-phase demodulation strategies are limited by the tremor prediction accuracy and the absence of tremor amplitude information for the predictions. A subtle change in tremor phase of tens of milliseconds during the demodulation or stimulation windows can lead to errors in the predicted tremor bursts resulting in relevant de-synchronization issues. In extreme cases, the prediction error can lead to in-phase stimulation of the agonist muscle, eliciting the opposite modulation effect on the nervous system. Furthermore, the intensity of stimulation has to be predetermined, which does not allow any modulation to adapt it to the current tremor amplitude since the tremor amplitude is not considered in the prediction algorithm.

To cope with these issues, Pascual-Valdunciel et al. [20]

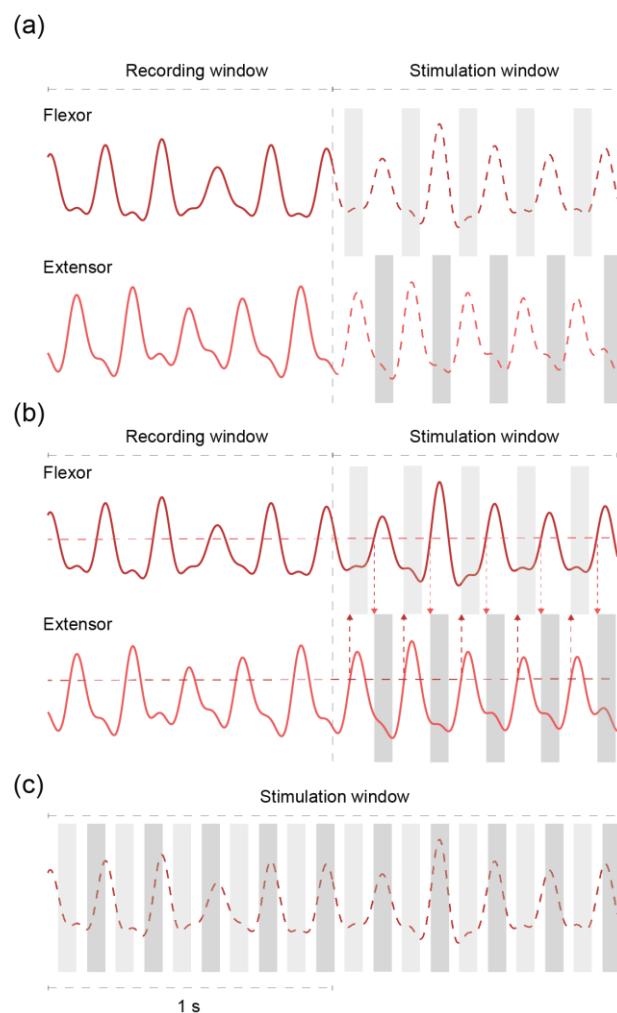


Fig. 1. PES strategies. Filled lines represent recorded signals (kinematics or EMG envelope), while dashed lines represent non-recorded (hypothetical) signals. Note that the displayed non-recorded signals are not predicted or used by the control algorithms and represent the tremor component that the patient would exhibit. Filled light and dark gray rectangles represent stimulation periods applied to the flexor and extensor muscles, respectively (pair of antagonist muscle). (a) Out-of-phase demodulation approach (closed-loop). The envelope of the tremor signal is computed to extract the period and predict the next tremorgenic cycles to synchronize the stimulation bursts. Only the tremor phase is predicted. (b) SATS strategy (closed-loop) uses continuous recordings to detect tremor in real-time and apply stimulation bursts to the antagonist muscle when the measured activity exceeds an adaptive threshold (horizontal dashed lines). No tremor prediction is applied in this strategy. (c) Calibrated open-loop strategy delivers a stimulation pattern locked at the estimated tremor frequency. This strategy does not use tremor data to control the electrical stimulation.

developed the selective and timely stimulation (SATS) strategy, which used a closed-loop algorithm based on consecutive real-time recording of tremor activity and consequent stimulation windows, not including any prediction stage (Fig. 1(b)). The SATS strategy allows quasi-synchronous stimulation and overcomes the prediction errors compared to out-of-phase due to the real-time implementation. However, this strategy presents a main drawback: the tremor activity is estimated by means of computing the root mean square (RMS) values of the EMG signals from short recording windows ($\approx [10, 20]$ ms),

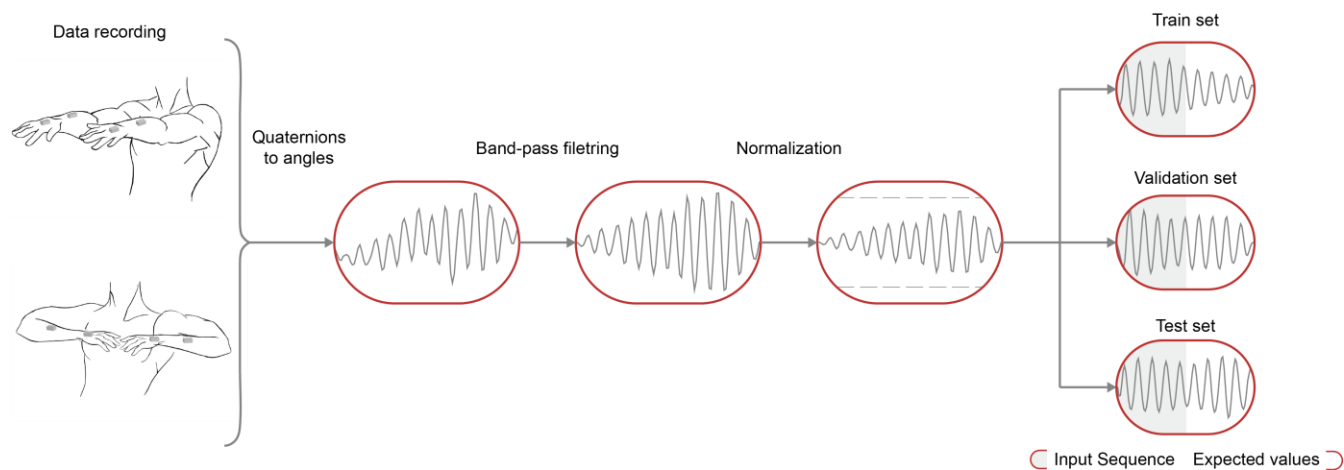


Fig. 2. Schematic of the data recording and pre-processing steps. First, kinematic signals were acquired through IMUs during postural tasks from ET patients (left panel). Then, signals were converted to Euler angles, band-pass filtered in the tremor band and normalized (center panel). Lastly, the signals were split into the training, validation and test sets (right panel).

which implies a limitation of temporal resolution. The physiological activity is always estimated with a delay related to the length of the RMS windows, and therefore the stimulation is always delivered a few milliseconds after the tremor activity is detected.

Previous closed-loop stimulation strategies have been only tested in research scenarios limited to postural tasks [14]–[16], [20], [21]. Outside the clinic environment, a commercial wearable device has been tested in a clinical trial using a calibrated open-loop strategy, which delivers a stimulation pattern locked to the tremor frequency of the patient without considering the tremor phase (Fig. 1(c)) [22]. Though the open-loop strategy allows convenient therapy translation from research to home environments, the stimulation applied is not optimal to specifically modulate the target neural circuits.

An increasing number of neurorehabilitation devices, either based on PES or mechanical assistance, are operated by closed-loop control algorithms synced with the real-time tremor oscillations measured through motion capture systems or electrophysiological signals [23]. In the case of active assistive upper-limb exoskeletons, mechanical actuators counteract tremor by applying opposite forces to the involuntary movements [24]. This application illustrates the importance of accurate estimation of the tremor component in order to synchronize the actuators to only interfere with the involuntary actions and minimize the disturbance of voluntary movements [25]. Different algorithms based on traditional time and frequency signal processing have been proposed for modeling and estimation of tremor, such as the Weighted Fourier Linear Combiner (WFLC), an adaptive tremor estimator based on the sinusoidal model [26]; or Wavelet Adaptive Kalman Estimation (WAKE) frameworks, which combine an adaptive Kalman filter followed by a wavelet transform to estimate tremor in real-time [27]. Recurrent Neural Networks have been used with the same purpose of estimating tremor and isolating it from voluntary movements [28]. Although all these approaches estimate tremor in real-time, there is an absence of real predictive capabilities of the next tremor cycles, therefore, they would not be optimal for triggering synchronized stimulation with physiological activity.

Forthcoming breakthroughs in the medical field related to the management of pathological tremor could be possible by means of integrating machine learning to monitoring systems [29] and tremor reduction interventions [30], [31]. Particularly, artificial neural networks (ANN) could mature into a solution to the tremor prediction paradigm. In the domain of tremor signals forecasting, Ibrahim et al. [32] proposed a hybrid convolutional-multilayer perceptron architecture to predict tremor and voluntary motion using kinematic signals from PD patients. Zanini et al. [33] tested several ANN designs resulting in MLP-encoders and Long Short-Term Memory (LSTM) neural networks as the best architectures to predict the envelope of EMG signals. Both studies were limited by the prediction horizon (PH) applied, which was restricted to 100 ms and 200 ms respectively, a value equivalent up to just one cycle of tremor, if a 5 Hz tremor frequency is assumed. These PHs obtained with both EMG and kinematics data are insufficient if the prediction of several tremor cycles is required, or even the prediction of one tremor cycle for patients exhibiting tremor with a frequency below 5 Hz, since the minimum PH required for estimating the next tremor cycle must be above 200 ms.

Drawing from the previous studies, it has been shown that PES of afferent pathways requires high time synchronization with the tremorgenic activity. Thereby, there is a need of validating a tremor prediction algorithm capable of reliably forecasting more than one cycle of tremor oscillations, consequently allowing to timely trigger the stimulation system with an adaptive intensity to the estimated tremor amplitude and subsequent optimization of the tremor reduction strategy. The purpose of this study was to design and implement a tremor series predictor based on LSTM neural networks. Kinematic signals from ET patients, specifically wrist flexion-extension angle displacement, were used to train, validate and test offline LSTM neural network models with assorted architectures (number of LSTM layers and hidden neurons), training parameters (learning rate), input sequence lengths and PHs. The training results presented here were examined to characterize the best model performance and determine the viability of the LSTM neural networks as tremor predictors compared to the traditional out-of-phase demodulation implementation.

The proposed LSTM predictor architecture seeks to overcome in the following points the results reported by previous studies on the prediction of pathological tremor signals:

- Prediction of the full tremor waveform, including tremor phase and amplitude.
- High precision accuracy (>0.7 correlation).
- PH between 100 ms and 1 s (200 ms is the highest PH explored for other ANNs).
- Lower phase delay prediction compared to traditional methods (out-of-phase).

II. METHODS

A. Participants and data recording

Twelve patients (age range of 62-82 years old) diagnosed with ET were recruited from the Movement Disorder Clinic of Gregorio Marañón Hospital (Madrid, Spain). A neurologist specialized in movement disorders diagnosed and assessed the tremor condition of the patients. The main inclusion criteria comprised the presence of postural tremor of at least one upper-limb, predominantly wrist flexion-extension tremor at the wrist, and no diagnosis of additional neurological or musculoskeletal pathologies. All the participants voluntarily signed the informed consent form to participate in the procedures, which were conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the hospital and the Spanish Agency of Medicines and Medical Devices (AEMPS - record 714/18/EC).

A motion capture system (Tech-MCS, Technaid S.L., Madrid, Spain) comprising four Inertial Measurement Units (IMUs) was used for the acquisition of the kinematic data. Each of the four sensors was located over the dorsal side of the hand and the forearm of both upper limbs. Raw quaternions representing the spatial orientation of the IMUs were sampled at 50 Hz (equivalent to a sampling period of 20 ms) and stored for offline analysis. The presence of tremor while patients are holding a static posture is distinctive for ET, hence, they were asked to hold during 60 seconds any of the following postures

against gravity to elicit tremor: both arms outstretched and pronated; or both arms with the elbows flexed and facing stretched fingers (Fig. 2). Patients were required to attend two experimental recording sessions performed in different weeks, in which a minimum of three trials were recorded for each patient.

B. Data pre-processing

The tremor prediction method here proposed requires several preprocessing steps over the raw kinematic data. Flexion-extension displacement angle was selected as the input signal for the tremor forecasting since the ET patients typically exhibit predominant tremor at this axis. Raw quaternions from the two aligned sensors on the hand and forearm were converted into Euler angles representing the rotation matrix of the wrist joint [20]. Since postural tremor in ET oscillates between 4 and 9 Hz, the tremor signal can be isolated by means of frequency filtering [34]. Though the voluntary components present in the kinematic data were limited due to the postural task execution, a third order Butterworth zero-phase band-pass filter between 4 and 10 Hz was applied on the angle data. LSTMs networks use *tanh* and *sigmoid* activation functions with $[-1, 1]$ and $[0, 1]$ output ranges respectively. Then, all data were normalized between 0 and 1 with the min-max feature scaling to prevent abnormal prediction values [35].

Each of the recording trials was segmented into 2-second recording of the angular variation of the pathological flexion-extension wrist movements. Overall, a total of 7,000 temporal sequences from twelve ET patients were stored and selected to comprise the final dataset. Then, the dataset was randomly divided into 3 subsets: training set, validation set and test set. The train set contains 70% of the data (4900 segments) and is used to train the models; the validation set comprises 15% of the data (1050 segments) and is used to validate the models along the training process to prevent overfitting; and the test set contains 15% of the data (1050 segments) to assess the performance of the models. Fig. 2 summarizes the preprocessing steps applied to the raw kinematic data prior to LSTM neural network training.

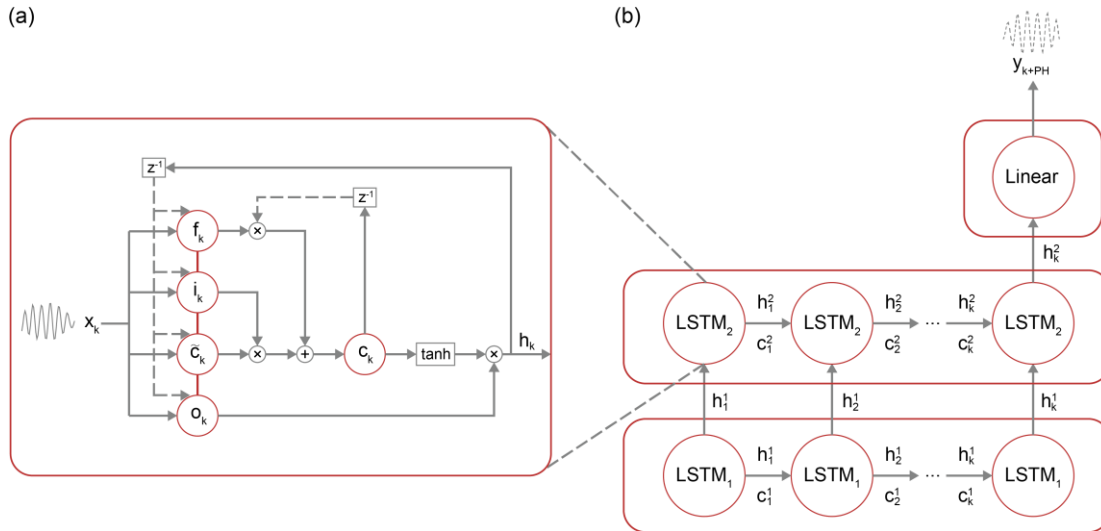


Fig. 3. Predictor architecture diagrams. (a) Computational graph of a LSTM cell. (b) Architecture of the prediction models based on LSTM neural networks. The design illustrated in the figure represents the models comprised of two LSTM layers. Adapted from [30].

C. Network architecture

Feedforward ANNs are not an optimal solution for time series forecasting due to the lack of connections between neurons in the same layer and their structure without interlayer feedback prevent them from having time dependence [36]. RNNs overcome this limitation by introducing recurrent connections among the hidden neurons and producing an output at each time step [37]. The RNNs are trained using the Back-Propagation Through Time (BPTT) algorithm that considers the time dimension, so the loss function of a given time step depends on the previous time step [38]. However, the BPTT algorithm within RNNs presents some drawbacks related to the vanishing and exploding gradients at time steps well before the current one, which prevents the network from learning long-term patterns [39]. These issues are addressed by LSTM neural networks, a particular sort of RNN capable of preserving both short and long-term dependencies through the use of an automatic loop to allow gradients to flow for a long time without vanishing [40].

The operation of a LSTM cell is supported by four structures named gates: the forget gate (f_k), the entry gate (i_k), the status gate ($\tilde{c}_k$) and the exit gate (o_k). Information flows through the different gates encoded in two different states: the cell state (c_k) and the hidden state (h_k). The cell state represents the current memory of the cell, while the hidden state represents the output value of the LSTM network. Equation (1) and Fig. 3 represent the mathematical operations and computational diagram of a LSTM cell. W and U represent the weight matrices for the input and hidden state, and b represent the biases. σ and $\tanh$ represent the sigmoid and hyperbolic tangent activation functions. As a simplified operation account of a LSTM neural network, given an input data sequence of length k , the first data point k_1 enters the first LSTM cell and generates the first hidden and cell states (h_1, c_1) through the four gates. Consequently, the following data points ($k_{2...k}$) are recursively fed into the network until the final hidden state (h_k) is generated, storing the relevant patterns of the sequence.

$$\begin{cases} f_k = \sigma(W_f x_k + U_f h_{k-1} + b_f) \\ i_k = \sigma(W_i x_k + U_i h_{k-1} + b_i) \\ o_k = \sigma(W_o x_k + U_o h_{k-1} + b_o) \\ \tilde{c}_k = \tanh(W_c x_k + U_c h_{k-1} + b_c) \\ c_k = f_k \circ c_{k-1} + \tilde{c}_k \circ i_k \\ h_k = o_k \circ \tanh(c_k) \end{cases} \quad (1)$$

D. Predictor architecture

The formulated forecasting problem requires an angle displacement data sequence that inputs into the neural network, which outputs the predicted angle data sequence subsequent to the input data. Two neural architectures were implemented and tested to elucidate the design impact on the prediction performance. In the first design, the angle data sequence input into one LSTM layer. Here, the time sequence was unfolded and the output corresponding to the hidden state of the first layer at the last time step (h_k^1) was fed into a linear layer composed of one neuron and a sigmoid activation function, which ultimately provided the predicted sequence of the angle displacement. The second design is an extension of the

previous, in which the output sequence h_k^1 from the first LSTM layer was fed into a second LSTM layer. Later, the output h_k^2 from that second layer was used as input of the linear layer to finally compute the forecasted data sequence. Deepening into the architecture, each of the LSTM layers was composed of a set of neurons determining the computational power of the network. In this implementation, three similar numbers of neurons for all the LSTM layers were tested: 20, 35 and 50 neurons, in order to find a trade-off between computational power and cost.

The main goal of this study was to research the feasibility of LSTM neural networks as pathological tremor predictors. Hence, the input and output data sequence lengths are fundamental parameters to be explored to validate a future real application of PES in which the electrical stimulation must be synchronized with the accurately predicted tremor burst. The input sequence (IS) is defined as the window containing the past values of angle displacement used as input to the model. Two IS windows were proposed in this implementation: 600 ms (30 samples) and 1000 ms (50 samples). The output sequence or PH is the time sequence length of the predicted values. In this study, five PH values were explored: 100 ms (5 samples), 200 ms (10 samples), 400 ms (20 samples), 600 ms (30 samples) and 1000ms (50 samples). It is noteworthy that data from the first of the two seconds of the trial were used as IS sequences, while the data from the last second of the trial were used as PH sequences.

Regarding the network training parameters, the Adam optimizer was selected as the algorithm to compute an adaptive update of the weights and biases of the network based on previous LSTM neural network designs used to forecast time series [41]. Furthermore, three learning rate values ($\alpha=0.001$, $\alpha=0.0005$ and $\alpha=0.0001$) are proposed to explore the best training parameters for each of the predictor's implementation, since the learning rate is a parameter implied in backpropagation training algorithm convergence. During the training and validation processes, an early stopping procedure was applied to prevent overfitting by ceasing the training iterations if the loss function value for the validation set stops decreasing [42]. In the aggregate, the combination of all parameters led to 180 models to be trained, validated and tested. The framework used to implement the predictor models is *PyTorch* library from Python [43].

E. Evaluation of prediction accuracy

To assess the accuracy of the predicted sequences provided by the different models in relation to the real or target values a set of metrics and statistical analyses were performed: the mean square error (MSE), which was previously used in the training procedures as the loss function; the root mean square error (RMSE) [28], frequently used to assess forecasting accuracy in other studies [32]; and the Pearson's correlation coefficient ($\rho_{y,\hat{y}}$), which provides information about the linear relationship between the predictions and the expected values. The minimum phase delay or delay between the predicted and the expected values is a fundamental outcome in the application of PES synced with the tremorgenic activity. Hence, the phase delay (PD) of the predictions of the models were compared to the PD values resulting from the out-of-phase prediction approach implemented in [21]. The PD was measured by the following

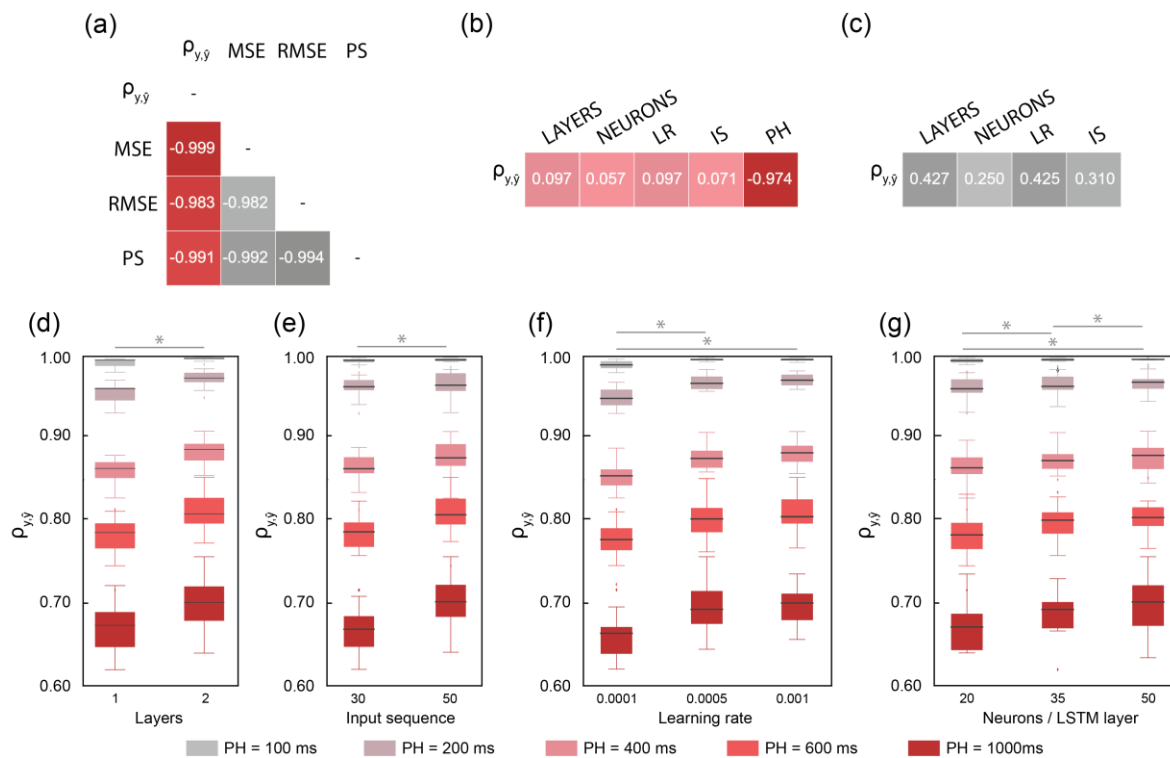


Fig 4. (a) Correlation among the different prediction performance metrics. (b) Correlation between $\rho_{y,\hat{y}}$ and each of the architecture and training parameters of the prediction models. (c) Partial correlation (PH fixed as constant factor) between $\rho_{y,\hat{y}}$ and each of the architecture and training parameters of the prediction models. (d-g) Boxplots representing $\rho_{y,\hat{y}}$ values for all the models grouped by PH and LSTM_{layers} (d), Input Sequence (e), learning rate (f) and LSTM_{neurons} (g). * Represents statistical significance ($p < 0.05$) at the group level.

method: local maxima were detected in both real and predicted signals; then, the time of the maxima from the expected signals were subtracted from the corresponding time maxima from the predicted signals; and finally, the PD average was computed for each trial. A similar procedure was followed to assess the PD results achieved by the out-of-phase prediction algorithm. The out-of-phase algorithm implementation consisted of extracting the envelope of the signal in the tremor frequency band as done in the *B. Pre-processing section*. Then, the local maxima corresponding to individual tremor oscillations were identified, and the period of the signal was estimated by computing the mean inter-burst interval (MIBI). The next tremor bursts were predicted by adding the MIBI to the timing of the last tremor burst detected. Finally, the PD was computed by means of subtracting the tremor time peaks predicted by the out-of-phase approach from the time peaks extracted from the local maxima identified in the expected signals.

In addition to the overall suitability assessment of the model to predict pathological tremor, this study sought to provide insight into how the different network and training parameters impact the prediction outcomes. Since five parameters were explored (LSTM_{layers}, LSTM_{neurons}, PH, IS and LR), after checking normality of the $\rho_{y,\hat{y}}$ results, a 5-way ANOVA test was applied. Afterwards, a post-hoc analysis with Bonferroni correction was applied to check individual differences at the group level.

Regarding the comparison of the PD between the predicted and expected tremor cycles obtained for the LSTM neural networks and out-of-phase methods, the PD values did not

follow a normal distribution and a non-parametric Friedman test was applied to the data. Then, Durbin-Conover post-hoc analysis with Bonferroni correction were applied to test the difference between the LSTM models and out-of-phase approach.

III. RESULTS

In this section, the tremor prediction models trained and validated following the procedures described in the *Methods* section were tested and their performance was analyzed through the metrics previously proposed. Two aims have been identified in this section: identifying the relationship among the different network and training parameters with the models' performance; and assessing the suitability of the models to accurately and reliably predict tremor signals.

Overall, 180 prediction models based on LSTM neural networks were trained, validated and ultimately trained in 1050 tremor sequences, each of them comprising two seconds of wrist angle displacement data windows recorded from twelve ET patients. The model showing the best performance achieved $\rho_{y,\hat{y}} = 0.999$, $MSE = 0.0001^{o^2}$ and $RMSE = 0.01^{o}$ (Table I, first line) while the model showing the worst performance achieved $\rho_{y,\hat{y}} = 0.573$, $MSE = 0.0423^{o^2}$ and $RMSE = 0.197^{o}$.

A linear correlation analysis was performed among the different performance assessment metrics in order to explore the reliability and dependence of each of the metrics. Results showed correlation values more negative than -0.98 for all metrics, proving any of the proposed metrics could serve as

TABLE I
OPTIMAL TREMOR PREDICTION MODELS

Layers	Neurons	Learning Rate	Input Sequence (ms)	Prediction Horizon (ms)	$\rho_{y,\hat{y}}$	MSE ($^{\circ 2}$)	RMSE ($^{\circ}$)	Phase delay LSTM (ms)	Phase delay out-of-phase (ms)
2	35	0.0005	600	100	0.998	0.0002	0.013	-	-
2	35	0.001	600	200	0.977	0.0021	0.039	-	-
2	50	0.001	600	400	0.889	0.0105	0.089	$8 \pm 9^*$	$18 \pm 14^*$
2	50	0.001	600	600	0.824	0.0178	0.119	$10 \pm 7^*$	$20 \pm 13^*$
2	35	0.001	600	1000	0.709	0.0307	0.163	$14 \pm 9^*$	$21 \pm 12^*$
2	50	0.001	1000	100	0.999	0.0001	0.010	-	-
2	50	0.0005	1000	200	0.984	0.0015	0.033	-	-
2	50	0.001	1000	400	0.909	0.0094	0.085	$8 \pm 9^*$	$15 \pm 13^*$
2	50	0.001	1000	600	0.853	0.0156	0.112	$10 \pm 9^*$	$17 \pm 12^*$
2	50	0.0005	1000	1000	0.757	0.0249	0.146	$13 \pm 9^*$	$18 \pm 11^*$

* Represents significant statistical differences between the LSTM and the out-of-phase methods ($p < 0.05$).

indicators of the quality of the predictions (Fig. 4(a)). Since $\rho_{y,\hat{y}}$ provides normalized and easily interpretable information about the fitting between the target and the predicted signals, this metric is preferred to review the results displayed along this section. The 180 models explored resulted from the combination of all the training and network parameters described in the previous section. In order to explore how each of the training and network parameters impact prediction performance, a linear correlation analysis was applied between the $\rho_{y,\hat{y}}$ and each of the parameters. A significant negative correlation (-0.974) was found between PH and $\rho_{y,\hat{y}}$, indicating

better fitness between the predictions provided by the models and the real signals for shorter prediction windows, while the increase of the prediction error raised with longer prediction windows (Fig. 4(b)). No other significant correlation was found for the remaining parameters. However, this analysis included $\rho_{y,\hat{y}}$ values resulting from each of the models combining a set of parameters, thus, inter-variable effects could mask other linear relationships. Then, a partial correlation analysis controlling the independent variable PH was applied. This analysis showed a significant linear relationship between $\rho_{y,\hat{y}}$ and each of the parameters, though the correlation values were lower compared

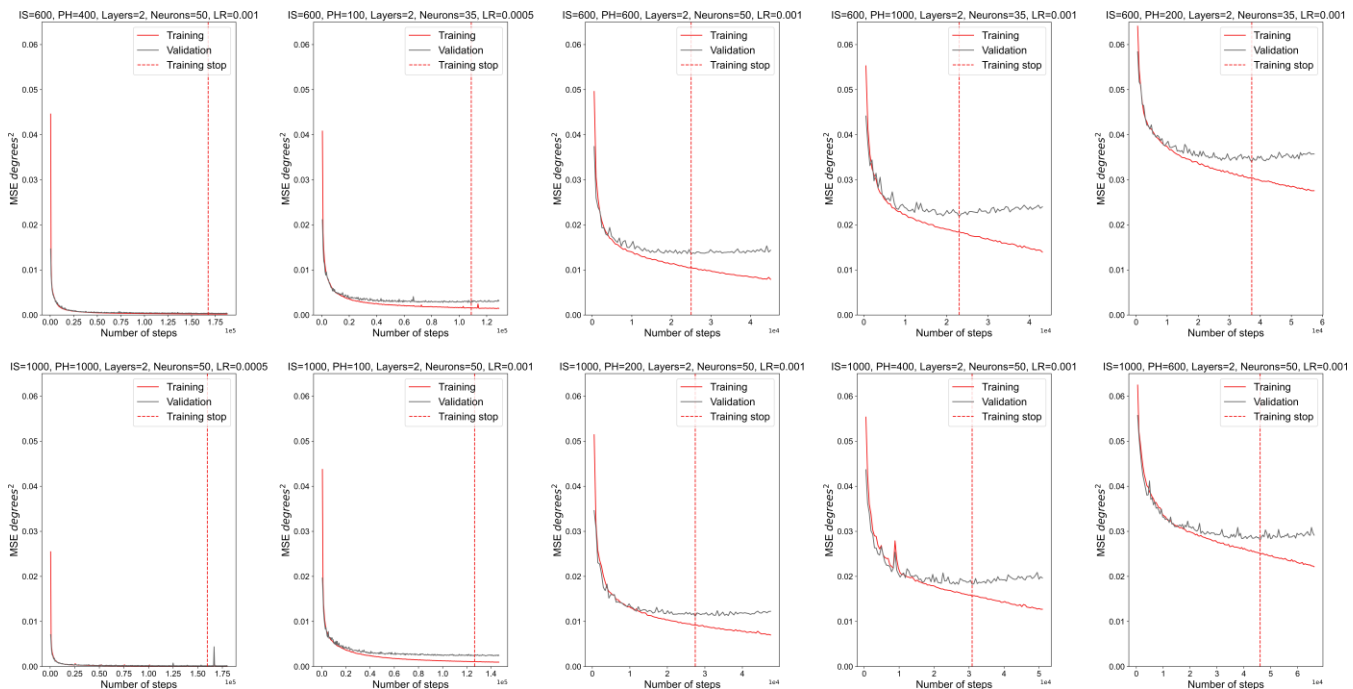


Fig 5. Learning curves for the training and validation datasets corresponding to the optimal LSTM predictors summarized in Table I. In order to prevent overfitting, the final model selected was that one showing the lowest average loss across 5 epochs for the validation test (red vertical dashed lines). The length of the global steps displayed in the figures include all the training steps until the stop function was activated since no decrease in the average validation loss was achieved in the last 40 training epochs.

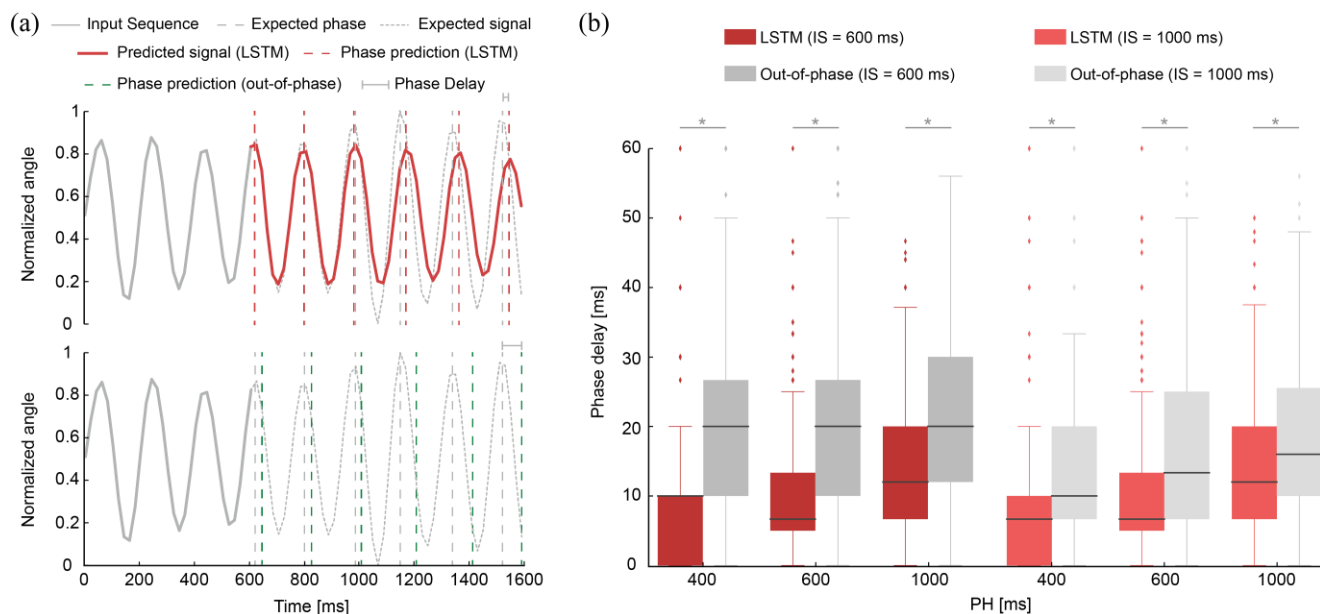


Fig 6. (a) Illustration of the tremor prediction performed by a LSTM neural network model (upper panel, IS = 600 ms, PH = 1000 ms) and the traditional out-of-phase approach (lower panel, IS = 600 ms, PH = 1000 ms). (b) Boxplots representing phase delay (PD) values obtained by the optimal models described in Table I and the out-of-phase method for the same test dataset. Horizontal bars represent median values, error bars represent standard deviations, and the boxes represent the interquartile range. * Represents statistical significance ($p < 0.05$) between groups.

to the interaction between $\rho_{y,\hat{y}}$ and PH (Fig. 4(c)). Based on these results, those prediction models comprising two LSTM layers with 50 neurons in each layer, trained with a learning rate value of 0.001 and 1000 ms of input tremor data (IS) tend to provide the optimal predictions. Figs 4. (d-g) show $\rho_{y,\hat{y}}$ values grouped by PH and each of the remaining parameters ($\text{LSTM}_{\text{layers}}$, $\text{LSTM}_{\text{neurons}}$, LR and IS). In spite of the variability, the linear trend shown in the correlation and partial correlation analyses is observable for all the variables. In addition, the 5-way ANOVA analysis (factors: layers, neurons, LR, IS and PH) showed significant statistical differences for $\rho_{y,\hat{y}}$. When applying the post-hoc tests with corrections for multiple group comparisons, statistically significant differences ($p < 0.001$) in $\rho_{y,\hat{y}}$ were reported within each group ($\text{LSTM}_{\text{layers}}$: 1 vs 2; $\text{LSTM}_{\text{neurons}}$: 20 vs 35 vs 50; IS: 600 ms vs 1000 ms; PH: 100 ms vs 200 ms vs 400 ms vs 600 ms vs 1000 ms). For LR, statistical differences were only reported between LR = 0.0001 and LR = 0.0005 or LR = 0.001.

Table I summarizes the performance metrics and the training parameters for the best models tested, for all combinations of IS and PH. The learning curves for the training and validation datasets can be found in Fig. 5. The linear relationship previously detected between the PH and $\rho_{y,\hat{y}}$ is observable for the optimal models as well. The highest correlation performance metric ($\rho_{y,\hat{y}} = 0.999$) was accomplished by the model predicting 100 ms of tremor signals using 1000 ms as input data, while the optimal model to predict 1000 ms of tremor signals using 1000 ms of input data achieved $\rho_{y,\hat{y}} = 0.756$. Regarding the PD between the predicted and the real signals, all the optimum models performed successfully, providing average phase shift values between 8 ± 9 ms and 13 ± 9 ms. Note that the phase shift metric was not assessed for the models targeting PH = 100 ms and PH = 200 ms, since the fitness for those values was near optimal, and the tremor

frequency was likely too low to exhibit a complete tremor cycles in those short window, and therefore, the first tremor oscillation phase predicted by the out-of-phase method would not fall in that window, and therefore, the peak detection algorithm could not be accurately applied. When comparing the phase lag values provided by the LSTM models predictions with output values of the out-of-phase demodulation method, all the LSTM-based models outperformed the standard demodulation method for all the conditions Fig. 6(a). Friedman and post-hoc tests revealed that the LSTM models achieved lower PD values compared to the out-of-phase demodulation method for all the combinations of IS and PH ($p < 0.05$, Table I, Fig. 6(b)). For instance, the optimal models (IS = 600ms) predicting PH = 400 ms, PH = 600 ms and PH = 1000 ms achieved a relative decrease in phase lag of 56%, 47% and 32% compared to the results provided by the out-of-phase demodulation method. Additionally, for both tremor prediction approaches a linear relationship between the PH and the PD in the predicted signals can be observed, which ultimately depicts the natural time-varying dynamics of pathological tremor.

Fig. 7 illustrates prediction examples for the optimal models summarized in Table I. It is noteworthy the capability of the models to fit the predictions to the real signals for all the PH and IS displayed. The relationship between prediction performance and PH is emphasized once again, since for shorter prediction windows higher accuracy was achieved, being nearly optimal for the 100 ms and 200 ms prediction windows. Though a positive partial correlation was found between the IS and the models performance, this relationship was masked by the strong dependence on the PH, and only for longer PH values the performance improvement is noticeable. For instance, the models predicting 200 ms signals using IS = 600 ms and IS = 1000 ms performed $\rho_{y,\hat{y}} = 0.977$ and $\rho_{y,\hat{y}} = 0.984$, respectively; while the models predicting 1000 ms signals using IS = 600 ms

and IS = 1000 ms performed $\rho_{y,\hat{y}} = 0.709$ and $\rho_{y,\hat{y}} = 0.757$, respectively (Table I). In Fig. 7, it is possible to observe the non-stationary tremor behavior and how this time variability impacts predictions. The models tend to provide better fittings, in both tremor amplitude and phase when the tremor sequence remains stable. Although the non-stationary tremor pattern was even present in short time frames, such as the time windows used in this study, the predictors could capture those amplitude and phase variations through time (Figs. 7 (d, e)). On the

contrary, the out-of-phase tremor prediction method showed limited adaptive capacity to forecast the phase of varying tremor patterns, as it can be observed in the results displayed in Table I and Figs. 7 (e, h).

IV. DISCUSSION

In this study we showed that it is possible to predict wrist flexion-extension tremor signals recorded from ET patients with high accuracy through the use of LSTM neural networks. A set of 180 models was trained, validated and tested with wrist flexion-extension angle data recorded from 12 ET patients. The architecture of the predictors was comprised of one or two layers of LSTM neural networks connected to a linear layer which output the predictions.

The combination of five training and network parameters ($LSTM_{layers}$, $LSTM_{neurons}$, PH, IS and LR) was explored to determine the optimal model features to predict the tremor signals. Overall, the PH was the parameter showing higher impact on the model performance, since it was highly correlated (-0.974) with $\rho_{y,\hat{y}}$, which represents the prediction accuracy of the predicted signals when compared to the real signals. Longer PH led to larger prediction errors, a behavior expected when considering the variability of tremor phase and amplitude throughout time. The additional parameters showed significant positive partial correlation with $\rho_{y,\hat{y}}$. Generally, the models comprised of 2 $LSTM_{layers}$, 50 $LSTM_{neurons}$, input sequence of 1000 ms and trained with a learning rate of 0.001 achieved better predictions. A limited set of network and training parameters ($LSTM_{layers}$, $LSTM_{neurons}$, and learning rate) was heuristically selected to explore their influence in the prediction and to limit the computation time. If these predictors would be implemented in a real-time PES-based system, optimization of the prediction performance, network complexity and computational cost should be considered. Increasing the number of $LSTM_{layers}$ and $LSTM_{neurons}$ might imply an unnecessary increase of the network complexity that would not translate into a significant performance improvement.

Regarding the capabilities of the LSTM-based models to predict tremor signals, the performance results were successful with $\rho_{y,\hat{y}}$ values ranging from 0.709 to 0.998. Shorter prediction windows implied better fitness to the real signals, both in amplitude and tremor phase. When predicting longer tremor sequences, the prediction error increased, though the predictions were highly correlated to the expected values. Previous studies aiming at tremor forecasting using neural networks were limited to short prediction horizons, as illustrated by results provided by Zanini *et al.* [33], where one tremor cycle was predicted based on two seconds of input data; or Ibrahim *et al.* [32], where predictions horizons between 10 ms and 100 ms were provided. On the other hand, algorithms to estimate tremor such as WFLC [26] or WAKE [28] do not present forecasting capabilities and therefore the application of electrical stimuli could not be anticipated to the tremorgenic activity. Compared to these implementations, the PH explored in this study went beyond the 200 ms limitation, which would allow prediction of several tremor cycles from just one input sequence no longer than 600 ms.

The use of modern deep learning structures, such as the transformer networks, could have been considered for this

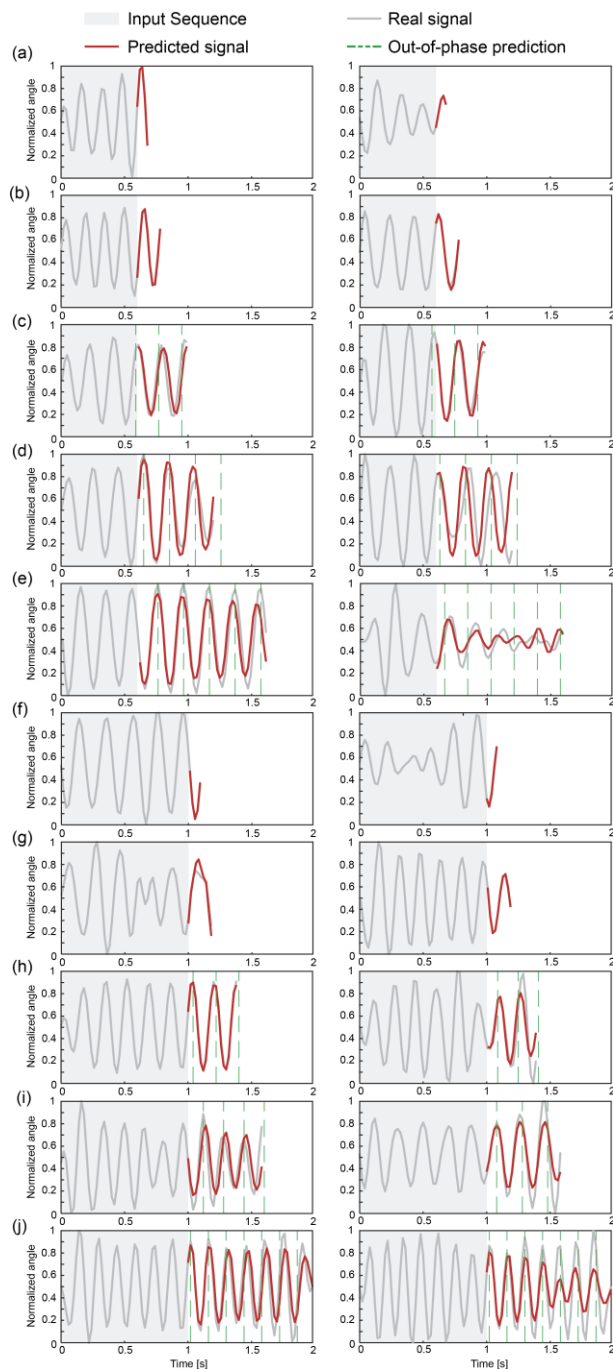


Fig 7. Each row from (a) to (j) illustrates the tremor predictions provided by the models summarized in Table I following the equivalent vertical order. Each column represents a random set of real and predicted signals from the test set. Dashed green lines represent the next tremorgenic bursts predictions provided by the out-of-phase demodulation algorithm.

application. Transformer networks are used in the time series forecasting of widely used datasets and are gaining interest for their performance and reduced training times [44]. Nevertheless, some studies have shown that transformer networks involve a higher number of hyperparameters and require a much more complex fine-tuning, which might not be translated to better outcomes compared to linear or LSTM models in certain applications [45]. Due to these statements, the LSTM networks were selected to solve the tremor forecasting problem in this study, while future work could focus on the comparison of different deep learning algorithms.

Moreover, the characterization of the effect of the input sequence length in the prediction performance is a novel contribution of this study, since prior contributions did not characterize the phase delay of the predicted signals and how this issue could impact the development of closed-loop PES strategies. Results shows that it is possible to predict up to one second of tremor behavior based on 600 ms of recorded data. Regarding the phase delay, the LSTM models outperformed the accuracy in phase prediction compared to the out-of-phase traditional demodulation approach by reducing the phase delay up to 57%. For the training set, statistical analysis showed that the LSTM models achieved more reduced phase delays compared to the out-of-phase method for all the IS and PH combinations.

Although the closed-loop SATS strategy has been proved to reduce acute tremor [20], the tremor detection method suffers from a delay related to the length of the recording window. Performance comparison between SATS and the LSTM models is not straightforward, since the first one is an activity-estimation algorithm while the second one is a predictor. However, the average PD achieved by the LSTM models for the shortest PH assessed here (400 ms) was estimated in 8 ± 9 ms, while additional studies have been estimated that the tremor detection average delay of some SATS strategy implementation is 17 ± 8 ms [46]. Hence, the LSTM models provide higher time accuracy compared to both SATS and out-of-phase approaches. Both traditional strategies, as well as LSTM predictors, have been tested on data recorded during postural tremor. In future studies it would be beneficial to explore the forecasting capabilities in data recorded from kinetic and action tremor, which are tremor manifestations occurring during the execution of activities of daily living.

On the whole, the models were able to capture tremor variability within the input signals and consequently they were capable of adapting the output signals to changes in tremor amplitude or phase. This feature is highly valuable since the models were not trained or tested for single patient data, which can contain patient-specific patterns. Thus, the patient-generalized models overcame the limitation of the study arising from the reduced number of patients' data used to train the models. Predictions far ahead in time led to higher prediction inaccuracies, a fact probably related to the non-stationary tremor behavior. In future studies it would be valuable to characterize the relationship between tremor variability and prediction performance, as well as the development of patient-specific models which could better capture individual tremor patterns and minimize prediction error related to inter-patient variability.

PES of afferent pathways is becoming a promising

alternative to manage pathological tremor [47]. Precise and timely recruitment of the afferent pathways, which might disrupt the tremorgenic input at the central nervous system, is a fundamental feature to optimize neuromodulation mechanisms. Specifically, precision below the order of tens of milliseconds is required to induce neural plasticity [48] as shown by different studies using paired associative stimulation (PAS) protocols. In these interventions PES stimulation is synchronized either with physiological events, such as motor activity [49] or tremor [20], or others forms of central nervous system stimulation such as transcranial magnetic stimulation [50]. The goal of these protocols is to converge different inputs (e.g., the afferent fibers recruited via PES and the tremor neural oscillations) into the same synapse in order to promote plasticity accordingly to Hebbian principles [51]. Hence, there is a need for developing robust closed-loop control algorithms capable of predicting the tremorgenic activity to deliver precise stimulation.

The forecasting models developed and tested in this study overcome tremor prediction performance from previous implementations, achieving correlation values of 0.984, 0.909, 0.853 and 0.757 for prediction of one, two, three and five tremor cycles (assuming 5 Hz tremor). Particularly, not only the phase delay between the predicted and the real signals was improved compared to the out-of-phase demodulation method, but also the models here proposed could estimate the tremor waveform and amplitude. This latter feature opens the possibility of novel PES strategies based on adaptive stimulation intensities in relation to the forecasted tremor amplitude. Future work should explore the implementation and optimization of these prediction models to operate in real-time PES interventions, which should consider the technical framework to ensure accurate and reliable stimulation during ADLs, as well as the possible acute change in tremor dynamics elicited by PES, a phenomenon which has not been explored so far.

Additionally, the EMG signals could be used as control signals instead of kinematics, since the angle displacement presents a characterized electromechanical delay with the muscle activity, and the stimulation delivered would be closer to the physiological activity timing. The LSTM-based predictors could be suitable to be trained with the envelope of the EMG in the tremor band, similarly to the approach proposed by Zanini *et al.* [33]. Interestingly, the results of longer PH achieved here could be extended with EMG data and improve the state-of-the-art results, allowing the synchronization of several cycles of stimulation based on EMG with higher accuracy. On the other hand, closed-loop systems based on kinematic signals might be more suitable to translate the application to the home environment where the quality of the electrophysiological measurements might be compromised by external factors. As an alternative, multimodal concurrent input signals such as kinematics (e.g., acceleration) and EMG from a pair of antagonist muscles would allow to explore different capabilities and performance of the proposed models towards the development of wearable devices or neuroprostheses in out-of-lab environments.

APPENDIX

Source code used to build the LSTM neural network predictors is available in the following repository:

https://github.com/Robolabo/LSTM_tremor_prediction_JBHI.git

ACKNOWLEDGMENT

We would like to thank Cristina Montero Pardo for illustrations from Fig.1. and the patients from Gregorio Marañón Hospital who voluntarily participated in this study.

REFERENCES

- [1] R. C. Helmich, I. Toni, G. Deuschl, and B. R. Bloem, "The pathophysiology of essential tremor and Parkinson's tremor.," *Curr Neurol Neurosci Rep*, vol. 13, no. 9, p. 378, Sep. 2013, doi: 10.1007/s11910-013-0378-8.
- [2] V. Shanker, "Essential tremor: diagnosis and management.," *BMJ*, vol. 366, p. 14485, Aug. 2019, doi: 10.1136/bmj.l4485.
- [3] E. D. Louis *et al.*, "Essential tremor-plus: a controversial new concept," *Lancet Neurol*, vol. 19, no. 3, pp. 266–270, Mar. 2020, doi: 10.1016/S1474-4422(19)30398-9.
- [4] T. A. Zesiewicz *et al.*, "Evidence-based guideline update: treatment of essential tremor: report of the Quality Standards subcommittee of the American Academy of Neurology," *Neurology*, vol. 77, no. 19, pp. 1752–1755, Nov. 2011, doi: 10.1212/WNL.0b013e318236f0fd.
- [5] E. Louis, "Treatment of Essential Tremor: Are there Issues We are Overlooking?," *Front Neurol*, vol. 2, p. 91, 2012, doi: 10.3389/fneur.2011.00091.
- [6] R. F. Dallapiazza *et al.*, "Outcomes from stereotactic surgery for essential tremor.," *J Neurol Neurosurg Psychiatry*, vol. 90, no. 4, pp. 474–482, Apr. 2019, doi: 10.1136/jnnp-2018-318240.
- [7] W. G. Ondo, "Current and Emerging Treatments of Essential Tremor.," *Neurol Clin*, vol. 38, no. 2, pp. 309–323, May 2020, doi: 10.1016/j.ncl.2020.01.002.
- [8] J. S. Lora-Millan, G. Delgado-Oleas, J. Benito-León, and E. Rocon, "A Review on Wearable Technologies for Tremor Suppression.," *Front Neurol*, vol. 12, p. 700600, 2021, doi: 10.3389/fneur.2021.700600.
- [9] A. Pascual-Valdunciel *et al.*, "Peripheral electrical stimulation to reduce pathological tremor: a review.," *J Neuroeng Rehabil*, vol. 18, no. 1, p. 33, Feb. 2021, doi: 10.1186/s12984-021-00811-9.
- [10] S. Dosen *et al.*, "Online tremor suppression using electromyography and low-level electrical stimulation.," *IEEE Trans Neural Syst Rehabil Eng*, vol. 23, no. 3, pp. 385–395, May 2015, doi: 10.1109/TNSRE.2014.2328296.
- [11] J. A. Gallego *et al.*, "Influence of common synaptic input to motor neurons on the neural drive to muscle in essential tremor," *J Neurophysiol*, vol. 113, no. 1, pp. 182–191, Jan. 2015, doi: 10.1152/jn.00531.2014.
- [12] A. Pascual-Valdunciel, F. O. Barroso, S. Muceli, J. Taylor, D. Farina, and J. L. Pons, "Modulation of reciprocal inhibition at the wrist as a neurophysiological correlate of tremor suppression: a pilot healthy subject study.," *Annu Int Conf IEEE Eng Med Biol Soc*, vol. 2019, pp. 6267–6272, Jul. 2019, doi: 10.1109/EMBC.2019.8857018.
- [13] J. L. Dideriksen, F. Gianfelici, L. Z. P. Maneski, and D. Farina, "EMG-based characterization of pathological tremor using the iterated Hilbert transform.," *IEEE Trans Biomed Eng*, vol. 58, no. 10, pp. 2911–2921, Oct. 2011, doi: 10.1109/TBME.2011.2163069.
- [14] J. L. Dideriksen *et al.*, "Electrical Stimulation of Afferent Pathways for the Suppression of Pathological Tremor.," *Front Neurosci*, vol. 11, p. 178, 2017, doi: 10.3389/fnins.2017.00178.
- [15] S. Muceli *et al.*, "A thin-film multichannel electrode for muscle recording and stimulation in neuroprosthetics applications.," *J Neural Eng*, vol. 16, no. 2, p. 26035, Apr. 2019, doi: 10.1088/1741-2552/ab047a.
- [16] J. Kim, T. Wichmann, O. T. Inan, and S. P. Deweerth, "A Wearable System for Attenuating Essential Tremor Based on Peripheral Nerve Stimulation.," *IEEE J Transl Eng Health Med*, vol. 8, p. 2000111, 2020, doi: 10.1109/JTEHM.2020.2985058.
- [17] A. Lenka and J. Jankovic, "Tremor Syndromes: An Updated Review.," *Front Neurol*, vol. 12, p. 684835, 2021, doi: 10.3389/fneur.2021.684835.
- [18] C. A. Vaz and N. V. Thakor, "Adaptive Fourier estimation of time-varying evoked potentials.," *IEEE Trans Biomed Eng*, vol. 36, no. 4, pp. 448–455, Apr. 1989, doi: 10.1109/10.18751.
- [19] J. Timmer, S. Haussler, M. Lauk, and C.-H. Lucking, "Pathological tremors: Deterministic chaos or nonlinear stochastic oscillators?," *Chaos*, vol. 10, no. 1, pp. 278–288, Mar. 2000, doi: 10.1063/1.166494.
- [20] A. Pascual-Valdunciel *et al.*, "Intramuscular Stimulation of Muscle Afferents Attains Prolonged Tremor Reduction in Essential Tremor Patients.," *IEEE Trans Biomed Eng*, vol. 68, no. 6, pp. 1768–1776, Jun. 2021, doi: 10.1109/TBME.2020.3015572.
- [21] S. Dosen *et al.*, "Online tremor suppression using electromyography and low-level electrical stimulation.," *IEEE Trans Neural Syst Rehabil Eng*, vol. 23, no. 3, pp. 385–395, May 2015, doi: 10.1109/TNSRE.2014.2328296.
- [22] S. H. Isaacson *et al.*, "Prospective Home-use Study on Non-invasive Neuromodulation Therapy for Essential Tremor.," *Tremor Other Hyperkinet Mov (N Y)*, vol. 10, p. 29, Aug. 2020, doi: 10.5334/tohm.59.
- [23] J. S. Lora-Millan, G. Delgado-Oleas, J. Benito-León, and E. Rocon, "A Review on Wearable Technologies for Tremor Suppression.," *Front Neurol*, vol. 12, p. 700600, 2021, doi: 10.3389/fneur.2021.700600.
- [24] E. Rocon, J. M. Belda-Lois, A. F. Ruiz, M. Manto, J. C. Moreno, and J. L. Pons, "Design and validation of a rehabilitation robotic exoskeleton for tremor assessment and suppression.," *IEEE Trans Neural Syst Rehabil Eng*, vol. 15, no. 3, pp. 367–378, Sep. 2007, doi: 10.1109/TNSRE.2007.903917.
- [25] V. Castrillo-Fraile, E. C. Peña, J. M. T. Gabriel Y Galán, P. D. Delgado-López, C. Collazo, and E. Cubo,

- “Tremor Control Devices for Essential Tremor: A Systematic Literature Review.” *Tremor Other Hyperkinet Mov (N Y)*, vol. 9, 2019, doi: 10.7916/tohm.v0.688.
- [26] C. N. Riviere, S. G. Reich, and N. V Thakor, “Adaptive Fourier modeling for quantification of tremor.” *J Neurosci Methods*, vol. 74, no. 1, pp. 77–87, Jun. 1997, doi: 10.1016/s0165-0270(97)02263-2.
- [27] S. Shahtalebi, S. F. Atashzar, R. V Patel, and A. Mohammadi, “WAKE: Wavelet decomposition coupled with adaptive Kalman filtering for pathological tremor extraction,” *Biomed Signal Process Control*, vol. 48, pp. 179–188, 2019, doi: 10.1016/j.bspc.2018.10.007.
- [28] S. Shahtalebi, S. F. Atashzar, O. Samotus, R. V Patel, M. S. Jog, and A. Mohammadi, “PHTNet: Characterization and Deep Mining of Involuntary Pathological Hand Tremor using Recurrent Neural Network Models.” *Sci Rep*, vol. 10, no. 1, p. 2195, Feb. 2020, doi: 10.1038/s41598-020-58912-9.
- [29] J. B. Sanderson *et al.*, “Multi-Dimensional, Short-Timescale Quantification of Parkinson’s Disease and Essential Tremor Motor Dysfunction,” *Front Neurol*, vol. 11, Sep. 2020, doi: 10.3389/FNEUR.2020.00886.
- [30] B. Houston, M. Thompson, A. Ko, and H. Chizeck, “A machine-learning approach to volitional control of a closed-loop deep brain stimulation system,” *J Neural Eng*, vol. 16, no. 1, Feb. 2019, doi: 10.1088/1741-2552/AAE67F.
- [31] M. Peralta, P. Jannin, and J. S. H. Baxter, “Machine learning in deep brain stimulation: A systematic review,” *Artif Intell Med*, vol. 122, Dec. 2021, doi: 10.1016/J.ARTMED.2021.102198.
- [32] A. Ibrahim, Y. Zhou, M. E. Jenkins, A. Luisa Trejos, and M. D. Naish, “The Design of a Parkinson’s Tremor Predictor and Estimator Using a Hybrid Convolutional-Multilayer Perceptron Neural Network.” *Annu Int Conf IEEE Eng Med Biol Soc*, vol. 2020, pp. 5996–6000, Jul. 2020, doi: 10.1109/EMBC44109.2020.9176132.
- [33] R. A. Zanini, E. L. Colombini, and M. C. F. de Castro, “Parkinson’s Disease EMG Signal Prediction Using Neural Networks,” in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 2019, pp. 2446–2453. doi: 10.1109/SMC.2019.8914553.
- [34] R. J. Elble and J. McNames, “Using Portable Transducers to Measure Tremor Severity.” *Tremor Other Hyperkinet Mov (N Y)*, vol. 6, p. 375, 2016, doi: 10.7916/D8DR2VCC.
- [35] R. Sendra-Arranz and A. Gutiérrez, “A long short-term memory artificial neural network to predict daily HVAC consumption in buildings,” *Energy Build*, vol. 216, p. 109952, 2020, doi: <https://doi.org/10.1016/j.enbuild.2020.109952>.
- [36] S. Hochreiter, “The Vanishing Gradient Problem during Learning Recurrent Neural Nets and Problem Solutions,” *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, vol. 6, no. 2, pp. 107–116, 1998, doi: 10.1142/S0218488598000094.
- [37] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [38] P. J. Werbos, “Backpropagation through time: what it does and how to do it,” *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1550–1560, 1990, doi: 10.1109/5.58337.
- [39] Y. Bengio, P. Simard, and P. Frasconi, “Learning long-term dependencies with gradient descent is difficult,” *IEEE Trans Neural Netw*, vol. 5, no. 2, pp. 157–166, 1994, doi: 10.1109/72.279181.
- [40] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Comput*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
- [41] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization.” 2017.
- [42] L. Prechelt, “Early Stopping - But When?,” in *Neural Networks: Tricks of the Trade*, G. B. Orr and K.-R. Müller, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998, pp. 55–69. doi: 10.1007/3-540-49430-8_3.
- [43] A. Paszke *et al.*, “Automatic Differentiation in PyTorch,” 2017.
- [44] A. Vaswani *et al.*, “Attention is All you Need,” in *Advances in Neural Information Processing Systems*, 2017, vol. 30. [Online]. Available: <https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf>
- [45] L. and C.-G. M. and L.-R. J. M. Lara-Benítez Pedro and Gallego-Ledesma, “Evaluation of the Transformer Architecture for Univariate Time Series Forecasting,” in *Advances in Artificial Intelligence*, 2021, pp. 106–115.
- [46] A. Pascual-Valdunciel, “Reduction of pathological tremor in Essential Tremor patients through Peripheral Electrical Stimulation of afferent pathways,” Universidad Politécnica de Madrid, Madrid, 2022.
- [47] A. Pascual-Valdunciel, A. Rajagopal, J. Pons, and S. Delp, “Non-Invasive Electrical Stimulation of Peripheral Nerves for the Management of Tremor,” *J Neurol Sci*, vol. accepted, 2022.
- [48] A. Suppa *et al.*, “The associative brain at work: Evidence from paired associative stimulation studies in humans.” *Clin Neurophysiol*, vol. 128, no. 11, pp. 2140–2164, Nov. 2017, doi: 10.1016/j.clinph.2017.08.003.
- [49] J. Ibáñez *et al.*, “Plasticity induced by pairing brain stimulation with motor-related states only targets a subset of cortical neurones,” *Brain Stimul*, vol. 13, no. 2, pp. 464–466, Mar. 2020, doi: 10.1016/J.BRS.2019.12.014.
- [50] M. A. Urbin, R. A. Ozdemir, T. Tazoe, and M. A. Perez, “Spike-timing-dependent plasticity in lower-limb motoneurons after human spinal cord injury.” *J Neurophysiol*, vol. 118, no. 4, pp. 2171–2180, Oct. 2017, doi: 10.1152/jn.00111.2017.
- [51] S. Song, K. D. Miller, and L. F. Abbott, “Competitive Hebbian learning through spike-timing-dependent synaptic plasticity.” *Nat Neurosci*, vol. 3, no. 9, pp. 919–926, Sep. 2000, doi: 10.1038/78829.