An Adaptive Algorithm for Online Interference Cancellation in EMG Sensors

Kaan Gokcesu, Mert Ergeneci, Erhan Ertan, and Hakan Gokcesu

***Abstract*— One of the biggest issues encountered in the analysis of sensitive electromyography (EMG) sensor data is the power line interference (PLI). Conventional methods in literature either lose valuable sensor data or inadequately decrease the power line noise. Instead of filtering out predetermined frequencies, adaptively estimating the spectrum of PLI can provide better performance. This paper introduces an online adaptive algorithm that removes the power line interference in real time without disturbing the true EMG data. Our method sequentially processes the biomedical signal to properly estimate and remove the PLI component. Through experiments with real EMG data, we compared our method to the five state-of-the-art techniques. Our algorithm outperformed all of them with the highest SNR gain (3.6 dB on average) and with the least disturbance of the true EMG signal (0.0152 dB loss on average). Our method reduces the PLI the most while keeping the valuable sensor data loss at its minimum in comparison to the state-of-the-art. Reducing the noise without disturbing the valuable sensor data provides higher quality signals with decreased interference, which can be better processed and used in biomedical research.**

***Index Terms*— EMG data analysis, PLI cancellation, adaptive filtering, spectrum estimation, online learning.**

I. Introduction

ELECTROMYOGRAPHY (EMG) is a method for evaluating and recording the electrical activity produced by skeletal muscles [1]. Recently, EMG technology has been used in a wide range of applications, e.g., human machine interaction and gesture recognition [2]–[8], rehabilitation [9]–[14], sports science and athlete performance monitoring [15], [16]. A major problem encountered in EMG sensor data measurements is the power line interference (PLI), which can be caused by any AC powered electronic equipment [17]. In general, this interference has a sinusoidal component at 50 Hz or 60 Hz (depending on the geographical region) and may also include its harmonic frequencies [18]. Even though the PLI signal is known to be a sinusoidal at 50/60 Hz, its frequency can vary by ± 2 Hz and its amplitude is dependent on the environment (e.g., the power of electronic components) [19]. Hence, high PLI makes the analysis of an EMG signal difficult since EMG is mostly located in the frequency band of 20-150 Hz [20], [21]. Since the raw EMG signal is generally less than 10 mV [22], its amplification to 0-5 V, amplifies PLI substantially, which causes an interference that can disturb the valuable EMG data.

A widely used analog method to overcome PLI is to use common mode decreasing solutions, e.g., driven-right-leg circuit (DRL) and high common mode rejection ratio instrumentation amplifiers [23]. However, these approaches are insufficient and additional digital methods are required since the analog procedures increase power consumption and cost whilst providing inadequate improvement on the overall signal-to-noise ratio (SNR) [24], [25]. The most common methods to mitigate PLI include the notch filter, the Spectrum Interpolation method (SI), the Laguerre filter, the Regression-Subtraction method (RS) and the Sinusoidal Modeling method (SM) [25], [26]. The notch filters are simply band-stop filters with a narrow stop-band at the specific frequency of PLI (50/60 Hz). The inherent problem with the notch filters lies in the difficulty of adjusting the bandwidth and the center frequency of its stop-band since a too narrow band may lead to a mismatch with the PLI frequency, while a wide band can significantly disturb the valuable EMG data (resulting in decreased signal power) [25], [26]. SI, on the other hand, uses a window smoothing on the spectrum of the EMG data to cancel unnatural components, since it assumes that the spectrum of the acquired EMG signal is the spectrum of true EMG data that is superimposed by a peak at the PLI frequency [26]. Another approach is the Laguerre filter, which is the adaptive and developed version of the notch filter. Although, they provide performance gains against the vanilla notch filters, the inherent problems such as the decrease in signal power and implementation difficulty are still unavoidable [27]. Finally, methods like Regression-Subtraction and Sinusoidal Modelling show substantial performance gains only when the exact PLI frequency is known in advanced [25], [28]. Since those techniques, except Laguerre filter, are not adaptive, the certain frequency of the noise must be given to the algorithm a priori for a sufficient noise rejection. However, in real life EMG data, PLI can change based on the environment, hence, using nonadaptive methods are highly risky.

All in all, these methods are unable to remove PLI adequately and may greedily reject the true EMG data. However, because of the spectral analysis done on the EMG data in most

Manuscript received August 7, 2018; accepted September 21, 2018. Date of publication October 8, 2018; date of current version December 7, 2018. The associate editor coordinating the review of this paper and approving it for publication was Prof. Aime Lay-Ekuakille. *(Corresponding author: Mert Ergeneci.)*

K. Gokcesu is with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: gokcesu@mit.edu).

M. Ergeneci and E. Ertan were with Hercules Biomedical R&D, Ankara 06800, Turkey. They are now with the Research Laboratories, Neurocess, Inc., Shenzhen 518028, China, (e-mail: erhertan@herculesw.com; ergenecimert@herculesw.com).

H. Gokcesu is with the School of IC, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland (e-mail: hakan.gokcesu@epfl.ch).

Digital Object Identifier 10.1109/JSEN.2018.2874724

applications, nearly all of the true EMG data is meaningful and has significance for a complete analysis [21], [29]–[33]. Therefore, the rejection of the true EMG data is undesirable since meaningful EMG data may be lost in the process. To this end, we introduce an efficient, highly adaptive, truly online algorithm, which not only removes PLI sufficiently but also keeps the valuable and meaningful EMG data nearly untouched.

In order to validate and demonstrate our method's advantages, we compared all of the algorithms' signal power, noise power and SNR performances. SNR is a common measurement metric for determining signal acquisition quality for surface EMG (sEMG) sensors. However, SNR, by itself, is not sufficient to properly compare the methods' performances because most techniques disturb the valuable EMG signal and decrease the signal power. Thus, by comparing the algorithms' performances in both signal power and noise power individually, we successfully demonstrate if the noise cancellation procedures cause any loss of the valuable data. Additionally, we demonstrate the average frequency responses of the active and inactive signal frames in order to illustrate the effects of each method on the power spectral density.

II. Problem Description

In this paper, we are dealing with the problem of PLI cancellation in the sEMG signals. A single channel sEMG signal can be mathematically modeled as the following:

$$y[n] = x[n] + w[n] + v[n], \tag{1}$$

where $x[n]$ is the muscle activation potential (clean EMG signal), $w[n]$ is the additive, stationary interference (such as the power line noise) and $v[n]$ is the AWGN noise. Our aim is to produce an estimate $\hat{x}[n]$ that can correctly model $x[n]$.

Generally, in the most prominent real-time EMG applications, the windows (or frames) of the EMG data are processed. Since in real-time systems, the processing of the whole stream is infeasible and a single measurement (sample-by-sample) carries very little information, the applications generally use a sliding window [6], [8], [14], [34]–[36]. To this end, we can define the problem as cleaning the EMG data frames. Let N be the frame length and L be the overlap length (the two adjacent frames have in common a total of L samples). We define the column vector $\boldsymbol{y}_t$ as the frame at time t which is

$$\boldsymbol{y}_t = [y[m+1], y[m+2], \ldots, y[m+N]]^T, \tag{2}$$

where $m = (t-1)(N-L)$. The column vectors $\boldsymbol{x}_t, \boldsymbol{w}_t, \boldsymbol{v}_t$ are similarly defined as the frames at time t of the clean EMG signal, PLI interference and AWGN noise respectively. We also define the Discrete Fourier Transforms (DFT) of these column vectors as $\boldsymbol{Y}_t, \boldsymbol{X}_t, \boldsymbol{W}_t, \boldsymbol{V}_t$ respectively. Hence,

$$\boldsymbol{y}_t = \boldsymbol{x}_t + \boldsymbol{w}_t + \boldsymbol{v}_t, \tag{3}$$

$$\boldsymbol{Y}_t = \boldsymbol{X}_t + \boldsymbol{W}_t + \boldsymbol{V}_t. \tag{4}$$

We denote our estimate of $\boldsymbol{x}_t$ as $\hat{\boldsymbol{x}}_t$ and use the mean squared error (MSE). Hence, our error at t, e_t, is given by

$$e_t = \mathbf{E}[(\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t)^T(\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t)|\boldsymbol{y}_t, \boldsymbol{y}_{t-1}, \ldots, \boldsymbol{y}_1], \tag{5}$$

since the observations $\boldsymbol{y}_t$ up to time t are known observations. We take the gradient of the error e_t with respect to our estimate $\hat{\boldsymbol{x}}_t$ and equate it to the all zero vector to find the optimal estimator (since MSE is convex). Hence,

$$\nabla_{\hat{\boldsymbol{x}}_t} e_t = \mathbf{E}[\boldsymbol{x}_t - \hat{\boldsymbol{x}}_t|\boldsymbol{y}_t, \boldsymbol{y}_{t-1}, \ldots, \boldsymbol{y}_1] = \mathbf{0}. \tag{6}$$

Thus, the optimal estimator of $\boldsymbol{x}_t$ is given by

$$\hat{\boldsymbol{x}}_t = \mathbf{E}[\boldsymbol{x}_t|\boldsymbol{y}_t, \boldsymbol{y}_{t-1}, \ldots, \boldsymbol{y}_1], \tag{7}$$

$$= \mathbf{E}[\boldsymbol{y}_t - \boldsymbol{w}_t - \boldsymbol{v}_t|\boldsymbol{y}_t, \boldsymbol{y}_{t-1}, \ldots, \boldsymbol{y}_1], \tag{8}$$

$$= \boldsymbol{y}_t - \mathbf{E}[\boldsymbol{w}_t|\boldsymbol{y}_t, \boldsymbol{y}_{t-1}, \ldots, \boldsymbol{y}_1] - \mathbf{E}[\boldsymbol{v}_t], \tag{9}$$

$$= \boldsymbol{y}_t - \hat{\boldsymbol{w}}_t, \tag{10}$$

since the expectation is conditioned on $\{\boldsymbol{y}_\tau\}_{\tau=1}^t$, $\boldsymbol{v}_t$ is zero-mean AWGN noise and $\hat{\boldsymbol{w}}_t$ is defined as the estimation of $\boldsymbol{w}_t$ given $\{\boldsymbol{y}_\tau\}_{\tau=1}^t$. Thus, we have reduced the problem of estimating the EMG signal $\boldsymbol{x}_t$ to the problem of estimating the interference $\boldsymbol{w}_t$. In the next section, we provide our algorithm, which can estimate the PLI by using the received frames.

III. Power Line Interference Removal

In this section, we introduce a completely online algorithm called Adaptive Noise Spectrum Estimation and Cancellation (ANSEC) to remove the power line noise in EMG data. Our algorithm works in an incremental manner and updates its parameters continuously. Its computational time and memory complexity are fixed per iteration. Our algorithm works with the frames of the EMG measurements, which are acquired by the sequential windowing of the continuous EMG data stream as given in (3). These frames are used by our adaptive algorithm in proper estimation of the PLI and its removal.

Our aim is to produce the estimate

$$\hat{\boldsymbol{w}}_t = \mathbf{E}[\boldsymbol{w}_t|\boldsymbol{y}_t, \boldsymbol{y}_{t-1}, \ldots, \boldsymbol{y}_1]. \tag{11}$$

We emphasize that the time domain estimation is equivalent to the frequency domain estimation because of the one-to-one mapping. Hence, we can write the estimator as

$$\hat{\boldsymbol{W}}_t = \mathbf{E}[\boldsymbol{W}_t|\boldsymbol{Y}_t, \boldsymbol{Y}_{t-1}, \ldots, \boldsymbol{Y}_1]. \tag{12}$$

Note that if there is no contraction at time t, then the observation frame is simply given by $\boldsymbol{Y}_t = \boldsymbol{W}_t + \boldsymbol{V}_t$. Therefore, we can write the model in a more concise form as

$$\boldsymbol{Y}_t = \alpha_t \boldsymbol{S}_t + \boldsymbol{W}_t + \boldsymbol{V}_t, \tag{13}$$

where $\boldsymbol{S}_t$ is the muscle contraction signal, $\boldsymbol{X}_t = \alpha_t \boldsymbol{S}_t$ and α_t is 1 if there is a contraction and 0 otherwise. Even though α_t are not given in advance, we assume that they are somehow known by us (we will explain later in Section III-B how we can estimate α_t). Thus, we also condition the estimator in (12) to α_t. Hence, our estimator becomes

$$\hat{\boldsymbol{W}}_t = \mathbf{E}[\boldsymbol{W}_t|\boldsymbol{Y}_t, \boldsymbol{Y}_{t-1}, \ldots, \boldsymbol{Y}_1, \alpha_t, \alpha_{t-1}, \ldots, \alpha_1]. \tag{14}$$

We multiply both sides of (13) with $(1-\alpha_t)$, which gives

$$(1-\alpha_t)\boldsymbol{Y}_t = (1-\alpha_t)\alpha_t \boldsymbol{S}_t + (1-\alpha_t)(\boldsymbol{W}_t + \boldsymbol{V}_t), \tag{15}$$

$$= (1-\alpha_t)(\boldsymbol{W}_t + \boldsymbol{V}_t), \tag{16}$$

since $\alpha_t \in \{0, 1\}$ (hence, $(1 - \alpha_t)\alpha_t = 0$). To get rid of $\boldsymbol{S}_t$, we approximate the estimator in (14) as

$$\hat{\boldsymbol{W}}_t = \mathbb{E}[\boldsymbol{W}_t|(1 - \alpha_t)\boldsymbol{Y}_t, (1 - \alpha_{t-1})\boldsymbol{Y}_{t-1}, \ldots, (1 - \alpha_1)\boldsymbol{Y}_1]. \tag{17}$$

We emphasize that $\boldsymbol{W}_t$ is actually included in all $\boldsymbol{Y}_t$. We approximate all of the $\boldsymbol{W}_t$ as the phase shifted versions of the same spectral representations, i.e.,

$$\boldsymbol{W}_t[k] = \boldsymbol{W}_0[k]e^{-jk\sum_{r=1}^{t}\theta_r}, \tag{18}$$

for $k \in \{0, 1, \ldots, N - 1\}$, where θ_r is the phase difference between $\boldsymbol{W}_r$ and $\boldsymbol{W}_{r-1}$. Even though we create the observation frames using a sliding window, we make no assumption about their phase differences (such as equal phase shifting) for the sake of generality. Thus, our algorithm will remain valid for arbitrary samplings of the EMG data stream (e.g., for the sake of power conservation). Moreover, the phase of the PLI component is susceptible to changes because the the contact between sensor and the skin creates a dynamic system that is affected by the human physiology. We denote $\boldsymbol{Z}_\tau$ as the phase shifted version of $\boldsymbol{Y}_\tau$ to $\boldsymbol{Y}_t$, i.e.,

$$\boldsymbol{Z}_\tau[k] = \boldsymbol{Y}_\tau[k]e^{-jk\sum_{r=\tau+1}^{t}\theta_r}. \tag{19}$$

Even though θ_t are not given in advance, we assume that they are somehow known by us (we will explain later in Section III-C how we can estimate θ_t). Thus, we also condition the estimator in (17) to θ_t. Hence,

$$\hat{\boldsymbol{W}}_t = \mathbb{E}[\boldsymbol{W}_t|(1 - \alpha_t)\boldsymbol{Y}_t, (1 - \alpha_{t-1})\boldsymbol{Y}_{t-1}, \ldots, (1 - \alpha_1)\boldsymbol{Y}_1, \theta_t, \theta_{t-1}, \ldots, \theta_1]. \tag{20}$$

Using (19), we approximate (20) as

$$\hat{\boldsymbol{W}}_t = \mathbb{E}[\boldsymbol{W}_t|(1 - \alpha_t)\boldsymbol{Z}_t, (1 - \alpha_{t-1})\boldsymbol{Z}_{t-1}, \ldots, (1 - \alpha_1)\boldsymbol{Z}_1], \tag{21}$$

where all $\boldsymbol{Z}_\tau$ directly includes $\boldsymbol{W}_t$ as an additive component.

To this end, we introduce an algorithm that utilizes three important subroutines. The first subroutine is the noise spectrum estimation, which estimates the PLI spectrum, $\hat{\boldsymbol{W}}_t$. The second subroutine is the silent frame detection, which estimates the contraction labels α_t that are used by the noise spectrum estimation. The third subroutine is the signal synchronization, which estimates the phase differences θ_t that are again used by the noise spectrum estimation. Therefore, before providing our complete algorithm, we first explain these subroutines in detail. We start with the noise spectrum estimation.

A. Noise Spectrum Estimation

In this part, we estimate the frequency spectrum of the interference $\boldsymbol{W}_t$. Due to the conjugate symmetry of the DFT for real signals, we only consider the first half of the transformation. The Minimum Mean Square Error (MMSE) estimator is equivalent to the Maximum Likelihood (ML) estimator for symmetric distributions (such as AWGN in $\boldsymbol{Z}_t$). Hence, (21) becomes

$$\hat{\boldsymbol{W}}_t = \arg\max_{\boldsymbol{W}_t} Pr\{(1-\alpha_t)\boldsymbol{Z}_t, (1-\alpha_{t-1})\boldsymbol{Z}_{t-1}, \ldots, (1-\alpha_1)\boldsymbol{Z}_1\}, \tag{22}$$

where $Pr\{\cdot\}$ is the probability operator. Instead of the joint probability, we greedily use the multiplication of the individual probabilities. Hence, our objective function becomes

$$\hat{\boldsymbol{W}}_t = \arg\max_{\boldsymbol{W}_t} \prod_{\tau=1}^{t} Pr\{(1 - \alpha_\tau)\boldsymbol{Z}_\tau\}, \tag{23}$$

Since, for $\alpha_t = 1$, $(1 - \alpha_t)\boldsymbol{Z}_t$ is 0, we can write (23) as

$$\hat{\boldsymbol{W}}_t = \arg\max_{\boldsymbol{W}_t} \prod_{\tau \in I_t} Pr\{(1 - \alpha_\tau)\boldsymbol{Z}_\tau\}, \tag{24}$$

such that I_t is the time indices where $\alpha_\tau = 0$ up to time t. Hence,

$$\hat{\boldsymbol{W}}_t = \arg\max_{\boldsymbol{W}_t} \prod_{\tau \in I_t} Pr\{\boldsymbol{Z}_\tau\}, \tag{25}$$

since $1 - \alpha_\tau = 1$ if $\tau \in I_t$. $\boldsymbol{Z}_\tau$ is the summation of $\boldsymbol{W}_t$ and a zero-mean AWGN. Thus, (25) is the sample mean, i.e.,

$$\hat{\boldsymbol{W}}_t = \frac{1}{|I_t|}\sum_{\tau \in I_t} \boldsymbol{Z}_t, \tag{26}$$

Using the definitions of $\boldsymbol{W}_t$, $\boldsymbol{Z}_t$, we can write (26) as

$$\hat{\boldsymbol{W}}_t = \begin{cases} \left(1 - \frac{1}{|I_t|}\right)\tilde{\boldsymbol{W}}_{t-1} + \frac{1}{|I_t|}\boldsymbol{Y}_t, & \alpha_t = 0 \\ \tilde{\boldsymbol{W}}_{t-1}, & \alpha_t = 1, \end{cases} \tag{27}$$

where $\tilde{\boldsymbol{W}}_{t-1}[k] = \hat{\boldsymbol{W}}_{t-1}[k]e^{-jk\theta_t}$ (i.e., synchronization).

Thus, to estimate the spectrum in a completely online manner, we start from an initial all-zero estimation and update it using the silent EMG data frames, where α_t and θ_t are given by the subroutines in Section III-B and III-C respectively.

We emphasize that the update given in (27) is completely online and is actually equivalent to the online gradient descent (OGD) update for the square error. Suppose the loss at t is

$$l_t(\tilde{\boldsymbol{W}}_{t-1}) = (\tilde{\boldsymbol{W}}_{t-1} - \boldsymbol{Y}_t)^H(\tilde{\boldsymbol{W}}_{t-1} - \boldsymbol{Y}_t),$$

where $(\cdot)^H$ is the Hermitian (conjugate transpose). Then, the gradient with respect to $\tilde{\boldsymbol{W}}_{t-1}$ is given by

$$\nabla_{\tilde{\boldsymbol{W}}_{t-1}} l_t(\tilde{\boldsymbol{W}}_{t-1}) = 2(\tilde{\boldsymbol{W}}_{t-1} - \boldsymbol{Y}_t). \tag{28}$$

Hence, the OGD update becomes

$$\hat{\boldsymbol{W}}_t = \tilde{\boldsymbol{W}}_{t-1} - 2u_t(\tilde{\boldsymbol{W}}_{t-1} - \boldsymbol{Y}_t). \tag{29}$$

Using the step size $u_t = \frac{1-\alpha_t}{2|I_t|}$ gives the update in (27).

At the beginning of the noise removal, the silent frame detector in Section III-B may have insufficient performance, hence, the noise spectrum estimation may incorrectly use active frames in its estimation. Therefore, to keep the robustness and adaptivity, we can also use the step size

$\mu_t = (1-\alpha_t)/\sqrt{|I_t|}$ instead, which has satisfactory performance for estimation in nonstationary environments [37]. Note that, for (29), we also need to gradually estimate the phase information to create $\tilde{\boldsymbol{W}}_t$ from $\hat{\boldsymbol{W}}_t$ (i.e., we need to estimate the phase shift θ_t). If we do not synchronize the past estimation $\hat{\boldsymbol{W}}_{t-1}$, the spectrum estimator will have unsatisfactory performance because of the linear update in complex space. Thus, we need to properly synchronize the spectra $\hat{\boldsymbol{W}}_{t-1}$ and $\hat{\boldsymbol{W}}_t$ before the update.

B. Silent Frame Detection

The goal of this subroutine is the estimation of the contraction labels α_t. In estimating the interference component $\boldsymbol{W}_t$ on the measured EMG data, we first need to correctly decide which part of the signal should be used in the learning process, i.e., the contraction labels α_t. If the algorithm were to incorrectly use the EMG signal $\boldsymbol{S}_t$ in the estimation of $\boldsymbol{W}_t$, it may remove the valuable EMG signal components in the future. Hence, for each incoming data frame, we first decide whether it contains only the interference $\boldsymbol{W}_t$. These frames with no EMG signal are called silent frames (or inactive frames) since there is no active muscle contraction that would create some sort of EMG signal. To correctly detect these silent frames, we need to use some means of unsupervised learning procedure, since, in most prominent EMG applications, the muscle contraction times are not necessarily provided.

To decide whether an incoming frame is silent or not, we use the energy of that frame. At each time t, we receive the signal frame $\boldsymbol{y}_t$. We denote the energy of $\boldsymbol{y}_t$ by P_t, which is

$$P_t = \boldsymbol{y}_t^T \boldsymbol{y}_t, \tag{30}$$

where $\boldsymbol{y}_t^T$ is the transpose of $\boldsymbol{y}_t$. Since, in our model in (3), the measured EMG data $\boldsymbol{y}_t$ is composed of the superposition of an EMG signal $\boldsymbol{s}_t$, the PLI interference $\boldsymbol{w}_t$ and an AWGN noise $\boldsymbol{v}_t$, the energy P_t of a silent frame should be sufficiently lower than the active frames (contraction). To show this, we calculate the expectation of P_t, which is given by

$$\begin{aligned}\mathbf{E}[P_t] &= \mathbf{E}[\boldsymbol{y}_t^T \boldsymbol{y}_t], &(31)\\ &= \mathbf{E}[(\alpha_t \boldsymbol{s}_t + \boldsymbol{w}_t + \boldsymbol{v}_t)^T (\alpha_t \boldsymbol{s}_t + \boldsymbol{w}_t + \boldsymbol{v}_t)], &(32)\end{aligned}$$

where we used the time domain version of (13). Hence,

$$\begin{aligned}\mathbf{E}[P_t] &= \mathbf{E}[\alpha_t^2 \boldsymbol{s}_t^T \boldsymbol{s}_t + 2\alpha_t \boldsymbol{s}_t^T (\boldsymbol{w}_t + \boldsymbol{v}_t)\\ &\quad \times 2\boldsymbol{w}_t^T \boldsymbol{v}_t + \boldsymbol{w}_t^T \boldsymbol{w}_t + \boldsymbol{v}_t^T \boldsymbol{v}_t]\\ &= \mathbf{E}[\alpha_t^2 \boldsymbol{s}_t^T \boldsymbol{s}_t + \boldsymbol{w}_t^T \boldsymbol{w}_t + \boldsymbol{v}_t^T \boldsymbol{v}_t],\\ &= \alpha_t \mathbf{E}[\boldsymbol{s}_t^T \boldsymbol{s}_t] + \mathbf{E}[\boldsymbol{w}_t^T \boldsymbol{w}_t + \boldsymbol{v}_t^T \boldsymbol{v}_t], &(33)\end{aligned}$$

since $\alpha_t^2 = \alpha_t$ and we assume $\boldsymbol{s}_t$, $\boldsymbol{w}_t$, and $\boldsymbol{v}_t$ are uncorrelated. Note that the right part of (33) exists whether or not there is a contraction in the received frame $\boldsymbol{y}_t$. However, the left part of (33) is dependent on the contraction label $\alpha_t \in \{0, 1\}$. Hence, when there is no contraction, the energy P_t will be lower.

Therefore, to decide whether an incoming frame $\boldsymbol{y}_t$ at time t is silent or not (whether or not it contains no contraction), we compare P_t against a threshold τ_t. Such that

$$\alpha_t = \begin{cases} 0, & P_t \le \tau_t \text{ (inactive)} \\ 1, & P_t > \tau_t \text{ (active).} \end{cases} \tag{34}$$

More generally, this is equivalent to the thresholding of a monotonically increasing transformation $\Phi(P_t)$ such that

$$\alpha_t = \begin{cases} 0, & \Phi(P_t) \le \tau_t \text{ (inactive)} \\ 1, & \Phi(P_t) > \tau_t \text{ (active).} \end{cases} \tag{35}$$

For the algorithm to work satisfactorily, a suitable selection of the threshold τ_t is needed. However, without the knowledge of the energy levels for active and inactive frames, this threshold cannot be selected reliably. Therefore, instead of selecting a fixed threshold, we use a dynamic thresholding scheme and learn it. Starting from an initial value τ_1, we update the threshold with online gradient descent (OGD) [38]. In general, OGD procedure updates the threshold as following

$$\tau_{t+1} = \tau_t - \mu_t \frac{\partial l_t(\tau_t)}{\partial \tau_t}, \tag{36}$$

where $l_t(\cdot)$ is the loss function at time t. OGD algorithm provides logarithmic regret bounds against the best fixed threshold chosen a priori when used with the learning rates $\mu_t = 1/Ht$, where H is a universal lower bound of the second derivatives of the loss functions $l_t(\cdot)$.

Since EMG data acquisition provides an unsupervised learning setting, the loss functions $l_t(\cdot)$ should be chosen in accordance. A suitable choice is the squared-loss against the energy of the received signal such that

$$l_t(\tau_t) = (\tau_t - P_t)^2. \tag{37}$$

The Minimum Mean Square Error (MMSE) solution is

$$\tau^* = \arg\min \sum_{t=1}^{T} (\tau - P_t)^2 = \frac{1}{T} \sum_{t=1}^{T} P_t, \tag{38}$$

which is the batch mean of the energy values P_t. Note that, if there are a lot of active frames in the measured EMG data, this batch mean can be undesirably high, which may cause the algorithm to misclassify active frames as inactive since some contractions may have lower frame energies. Therefore, we need a substantially lower threshold level. Since the silent frames' energies are generally similar and comparably lower to the active frames, we need a threshold value nearer to the minimum received frame energy. Therefore, we use the log transform on the energy values P_t, i.e., $\Phi(P_t) = \log(P_t)$. Thus, the loss function will be given by

$$l_t(\tau_t) = (\tau_t - \log(P_t))^2, \tag{39}$$

which will cause the MMSE solution to be

$$\tau^* = \arg\min \sum_{t=1}^{T} (\tau - \log(P_t))^2 = \frac{1}{T} \sum_{t=1}^{T} \log(P_t). \tag{40}$$

We point out that this MMSE solution τ^* is equivalent to the logarithm of the geometric mean of the energy values P_t. Hence, we will acquire a lower threshold value since geometric mean is lower than the arithmetic mean (AM-GM inequality).

Using the loss function in (39) and the OGD method in (36) with the step size $\mu_t = 1/2t$, our threshold update becomes

$$\tau_{t+1} = \tau_t - (\tau_t - \log(P_t))/t, \tag{41}$$

and our silent frame detection subroutine is given by

$$\alpha_t = \begin{cases} 0, & \log(P_t) \leq \tau_t \text{ (inactive)} \\ 1, & \log(P_t) > \tau_t \text{ (active)}. \end{cases} \tag{42}$$

Up to now, we have explained our silent frame detection, which classifies the incoming frames by its activity (active, inactive) and constructs the contraction labels α_t. Next, we detail our estimation of the phase differences θ_t.

C. Frame Synchronization

In this section, we explain how to synchronize the past estimation $\hat{\boldsymbol{W}}_{t-1}$ to the received signal $\boldsymbol{Y}_t$. The spectra $\boldsymbol{W}_t$ in $\boldsymbol{Y}_t$ and $\hat{\boldsymbol{W}}_{t-1}$ may have different phase values. This phase difference between the observed spectrum ($\boldsymbol{Y}_t$) and estimated noise spectrum ($\hat{\boldsymbol{W}}_{t-1}$) may lead to erroneous learning. Thus, before the update in (29), phases of $\hat{\boldsymbol{W}}_{t-1}$ and $\boldsymbol{Y}_t$ must be synchronized to create $\tilde{\boldsymbol{W}}_{t-1}$. The synchronization of $\hat{\boldsymbol{W}}_{t-1}$ with $\boldsymbol{Y}_t$, requires the multiplication of $\hat{\boldsymbol{W}}_{t-1}$ with a linear phase since this is equivalent to a shift in time domain as

$$\tilde{\boldsymbol{W}}_{t-1}[k] = \hat{\boldsymbol{W}}_{t-1}[k] \exp(-jk\theta_t), \tag{43}$$

where $\tilde{\boldsymbol{W}}_{t-1}[k]$ represents each element of the received frame DFT after synchronization. We set θ_t such that the k_*^{th} element of $\tilde{\boldsymbol{W}}_{t-1}$ will have the same phase as the k_*^{th} element of $\boldsymbol{Y}_t$, i.e., $\Theta(\tilde{\boldsymbol{W}}_{t-1}[k_*]) = \Theta(\boldsymbol{Y}_t[k_*])$, where $\Theta(\cdot)$ is the function that gives the angle of a complex number such that $\Theta(z) = \tan^{-1}(b/a)$, for $z = a + jb$. Therefore, we choose the parameter $\theta_t = (\Theta(\hat{\boldsymbol{W}}_{t-1}[k_*]) - \Theta(\boldsymbol{Y}_t[k_*]))/k_*$.

Naturally, one will be inclined to select the DFT point where the interference most resides (e.g., the point corresponding to the frequency 50 Hz for Power Line Interference). However, without the knowledge of this frequency, we need a more adaptive method to select the sample point k_*. Hence, we select the point k_* in accordance with the point that contains most power most often for the silent frames, i.e., the mode of the sequence of numbers each of which represents the point that contains the most power for each silent frame. We synchronize in accordance with the frequency point that generally contains the most power because selecting a point k_* randomly may lead to erroneous synchronization since the phase information of a frequency point with low power is vulnerable to distortion by white noise. When estimating a nonstationary interference, we can also include a forgetting factor on the number of frames whose power resides the most at the k^{th} frequency component. Next, we provide the complete algorithm, and explain in detail how the subroutines are used in succession to estimate and remove the PLI component in a received data frame.

D. Power Line Interference Removal

In our problem setting, we continuously receive the EMG frames $\boldsymbol{y}_t$ and at the start of each time t, we have our past interference estimation $\hat{\boldsymbol{w}}_{t-1}$. To cancel out the interference $\boldsymbol{w}_t$, we first synchronize $\hat{\boldsymbol{w}}_{t-1}$ (i.e., $\hat{\boldsymbol{W}}_{t-1}$) with $\boldsymbol{y}_t$ (i.e., $\boldsymbol{Y}_t$) and create the synchronized version $\tilde{\boldsymbol{w}}_{t-1}$ (i.e., $\tilde{\boldsymbol{W}}_{t-1}$) as in Sec. III-C. Next, we feed $\boldsymbol{y}_t$ to the silent frame detection explained in Sec. III-B. According to our decision, α_t, we use the synchronized estimate $\tilde{\boldsymbol{W}}_{t-1}$ (and $\boldsymbol{Y}_t$ if $\alpha_t = 0$) to update $\hat{\boldsymbol{W}}_{t-1}$ to $\hat{\boldsymbol{W}}_t$ as in Sec. III-A. Finally, we take the inverse DFT of $\hat{\boldsymbol{W}}_t$ to acquire $\hat{\boldsymbol{w}}_t$ and subtract it from $\boldsymbol{y}_t$ as

$$\hat{\boldsymbol{x}}_t = \boldsymbol{y}_t - \hat{\boldsymbol{w}}_t \tag{44}$$

to create the frame $\hat{\boldsymbol{x}}_t$. The complete algorithm is in Alg. 1.

Algorithm 1 ANSEC

Initialize frame length N and overlap length L.
Set initial thresholds level τ_1.
Set initial noise spectrum estimation $\hat{\boldsymbol{W}}_0$.
Randomly initialize k_* (except $k_* = 0$).
Create all zero maximum power index set M
Initialize $I_0 = \emptyset$
for $t = 1$ **to** ... **do**
  Receive the incoming frame $\boldsymbol{y}_t$
  %Noise Frame Detection
  $P_t = \boldsymbol{y}_t^T \boldsymbol{y}_t$
  if $\log P_t \leq \tau_t$ **then**
    $\alpha_t = 0$ (Silent Frame)
  else
    $\alpha_t = 1$ (Active Frame)
  end if
  $\tau_{t+1} = \tau_t - (\tau_t - \log(P_t))/t$
  %Noise Spectrum Estimation
  Create first half DFT $\boldsymbol{Y}_t$ from $\boldsymbol{y}_t$
  $\theta_t = (\Theta(\hat{\boldsymbol{W}}_{t-1}[k_*]) - \Theta(\boldsymbol{Y}_t[k_*]))/k_*$
  ($\Theta(\cdot)$ is the angle function)
  for $k = 0 : L/2$ **do**
    $\tilde{\boldsymbol{W}}_{t-1}[k] = \hat{\boldsymbol{W}}_{t-1}[k] \exp(-jk\theta_t)$ *%Sync.*
  end for
  if $\alpha_t = 0$ **then**
    $m = \arg\max_{k \neq 0} \|\boldsymbol{X}_t[k]\|$
    $M[m] = M[m] + 1$
    $k_* = \arg\max_k M[k]$
    $I_t = I_{t-1} \cup \{t\}$
  end if
  if $I_t \neq \emptyset$ **then**
    for $k = 0 : L/2$ **do**
      $\hat{\boldsymbol{W}}_t[k] = \tilde{\boldsymbol{W}}_{t-1}[k] - (\tilde{\boldsymbol{W}}_{t-1}[k] - \boldsymbol{Y}_k)(1 - \alpha_t)/|I_t|$
    end for
  end if
  %Noise Subtraction
  Set $\hat{\boldsymbol{w}}_t$ as the inverse DFT of $\hat{\boldsymbol{W}}_t$
  $\hat{\boldsymbol{x}}_t = \boldsymbol{y}_t - \hat{\boldsymbol{w}}_t$
end for

IV. Experiments

In the experiments, we captured sEMG data from biceps brachii, on the line between the medial acromion and the fossa cubit at 1/3 from the fossa cubit [39]. Two and a half minutes

TABLE I
INFORMATION ABOUT GENDER, AGE, HEIGHT AND WEIGHT OF THE SUBJECTS

Subject No	Age	Gender	Height (cm)	Weight (kg)
S1	20	Female	167	61
S2	24	Female	169	67
S3	26	Female	165	58
S4	24	Male	173	72
S5	25	Male	178	68
S6	28	Male	168	70
S7	24	Male	167	72
S8	25	Male	189	66
S9	24	Male	175	78
S10	24	Female	181	60

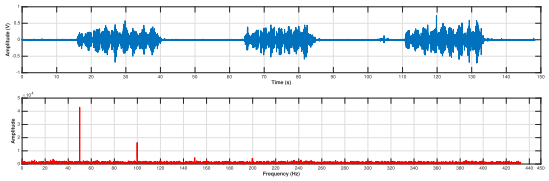

Fig. 1. Plot of captured EMG signal (top) and DFT of the signal's inactive part (bottom) of 4^{th} subject's (S4) raw EMG signal.

of EMG data was collected from ten able bodied subjects, who vary in age, height, gender and weight as shown in Table I.

The subjects realized both dynamic and isometric contractions along with relaxations (no contractions). Each data acquisition is made carefully according to a predetermined procedure. Prior to the data acquisition, the sensor placement locations (biceps brachi) were cleaned with alcohol and the subjects were ensured to be not exhausted. The first six subjects (S1-S6) realized three sets of nine dynamic dumbbell concentration curls (3 kg) with relaxations in between. In order to perform dumbbell concentration curls, the subjects sit on bench, grasp 3 kg dumbbell between feet, place back of upper arm to inner thigh, lean into leg to raise elbow slightly, raise dumbbell to front of shoulder and lower dumbbell until arm is fully extended [40]. S4's (4^{th} subject) plot of captured EMG signal (top plot) and DFT of the signal's inactive part, which corresponds to the spectrum of the PLI interference (bottom plot), are given in Figure 1 for illustration purposes. In the FFT plot in Figure 1, the PLI dominance on the AWGN can be seen clearly. Rest of the subjects (S7-S10) realized random contractions without dumbbell (empty-handed) whilst holding their arms at a 90 degree angle. For this experiment, the contraction start times, durations, and intensities were all left to the discretion of the subjects to collect a dataset that is as diverse and arbitrary as possible. The reason for applying different contraction procedures during data acquisition is to increase the variety of signals gathered and analyze the algorithms' robustness. The data acquisition of subjects S1-S6 is referred as experiment 1 and S7-S10 as experiment 2.

A. Data Collection

The EMG system that is used in the experiments is Biometrics SX230-1000. It is a dry surface EMG sensor with 1000 gain, 10 mm disk shaped stainless steel electrodes with 20 mm

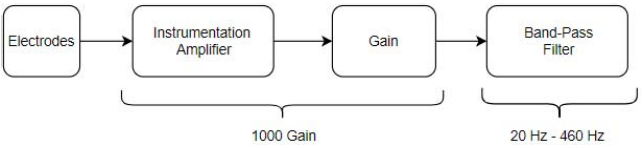

Fig. 2. Scheme of the Biometrics SX230-1000 EMG sensor.

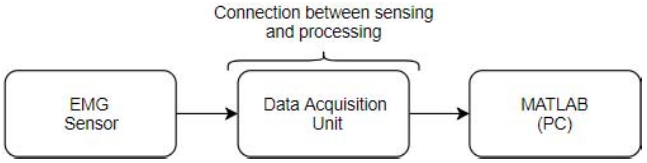

Fig. 3. Overall scheme of the whole data analysis architecture.

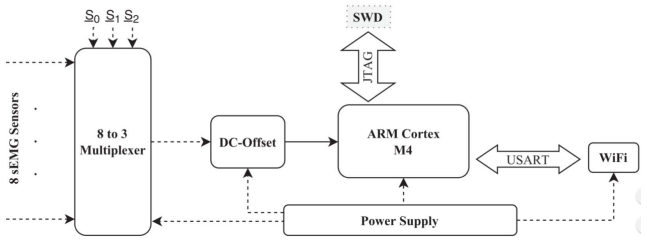

Fig. 4. Scheme of the data acquisition unit between EMG sensor and PC.

inner electrode distance (center to center) [41]. As seen in the Figure 2, the $V+$ and the $V-$ of the electrodes are directly connected to the instrumentation amplifier followed by the gain and band-pass filter blocks. The filter cutoff frequencies are $20Hz$ - $450Hz$ and the overall gain is 1000 [41].

In order to obtain a certain connection between sensing and processing, we made a data acquisition unit. The muscle signals are sensed via Biometrics EMG sensor and the EMG data is processed in Matlab 2016b. The connection between EMG sensor and the MatLab (PC) is realized by the data acquisition unit. In the Figure 3, the overall scheme of the data analysis architecture is visualized.

In the Figure 4, the data acquisition unit scheme is demonstrated. The system is composed of five main stages, which are multiplexer (MUX), DC-Offset, micro controller unit (MCU), power supply and data transmission (WiFi).

The first stage of DAU, i.e., MUX creates the multi-channel architecture. The 8 different sEMG channels enter the MUX where they are transferred to the DC-Offset unit sequentially. Since the sEMG signal band ranges between 20500 Hz, it is required to sample the incoming signals with a sampling rate of minimum 1 kHz, which corresponds to a maximum of 1 ms period. Thus, we decided to change the select pins (i.e., S0, S1, S2) of the MUX with a clock of 8 kHz which would lead to a 1 ms delay (1 kHz sampling frequency) between each sEMG signals' individual samples. In this way, we prevent any data loss and collect 8 different incoming sEMG signals simultaneously. The select pins (S0, S1, S2) of the MUX are controlled by the MCU and the output of the MUX is directly connected to the DC-Offset module where 1.5 V DC-Offset is added to the signal. The DC-offset module

converts the zero-mean signal to a unipolar signal with a range of 03 V to suitably interface with the ADC in the ARM Cortex CPU. A Unipolar ADC cannot convert negative voltage values. Thus, we up-shifted the mean of the EMG signal to 1.5 V, which positions the signal inside the ADC voltage range. After the addition of DC-Offset, the signals are digitalized in the internal ADC of the MCU. We used an ARM cortex-M4 based, stm32f407vgt6, microcontroller. The MCU requires external crystal (8 MHz) and a power supply circuit. It is possible to program the MCU through SWD pins via an external ST-LINKV2 debugger. The sEMG signals coming from the DC-Offset unit are first converted to digital and stored in arrays in accordance with the select of that time instance since each select value corresponds to a different sensor.

The data acquisition unit, which is used to digitize the EMG sensor outputs, has an ADC with a resolution of 12-bits, which is the recommended sensitivity by the SENIAM standards [39]. We use the full-scale range of the ADC, which is $0\ V$ - $3\ V$. Moreover, the accuracy of the EMG sensor is stated as $96dB$ CMRR at $60Hz$ in the product specification sheet [41].

B. Implementation of Traditional PLI Removal Methods

1) *Notch Filter* is designed in MatLab as an IIR, Comb, 17 order notching Filter with a bandwidth of 3 Hz, a fundamental frequency of 50 Hz and other harmonics (i.e., 50 Hz and 100 Hz), which are selected a posteriori to maximize SNR gain.
2) *SI* algorithm has been implemented to smooth the magnitude of the FFT index with maximum magnitude by averaging the magnitudes of its two adjacent indices in each observation frame [26].
3) *Adaptive Laguerre Filter* has been implemented with pole values of 0.5 and learning rate of 6×10^{-5} as instructed in its article [27]. Its order has been set to 50, which is selected a posteriori to maximize SNR gain.
4) *RS* has been implemented in an online manner for 8 harmonics (i.e., 50, 100, . . . , 400 Hz) in each data frame as instructed in its respective article [26].
5) *SM* has been implemented in an online manner for 8 harmonics and up to the second order polynomial of time t (i.e., $1, t, t^2$) in each data frame as instructed in its article [28]. The polynomial order for SM has been selected a posteriori to maximize SNR gain.

C. Signal Waveform Quality Test

In this subsection, the average values of SNR, Signal Power and Noise Power of the captured EMG data for ten subjects are calculated. Even though SNR calculation is commonly used as a performance metric for noise cancellation procedures, Signal Power and Noise Power are also required to be evaluated in order to observe the full effect of noise cancellation methods on the EMG data. A desirable result would be to increase SNR with an insignificant loss of Signal Power, i.e., the increase in SNR should be due to the decrease in Noise Power for optimal noise rejection. In order to compare the performances of all the described methods, we separated the signals by thresholding the RMS values of the received frames. Our algorithm works online but we saved the acquired sEMG signals on MatLab and analyzed the performance of the gathered sEMG data for each method. While making a comparison such as Noise and Signal Power, we first split the output of each method into frames and then taken the RMS of each sEMG frame. Next, by setting a threshold level that is created by naked-eye observation, we classified the original frames into two, which are noise (silent) and signal (contraction) frames. We did not use our adaptive classification method since it was significant to ensure that the active-inactive frames were classified correctly for objective analysis. Thus, checking each frame seperately would ensure that the noise and signal frames are classified correctly.

TABLE II

SIGNAL POWER, NOISE POWER AND SNR COMPARISONS FOR EXPERIMENT 1

	Signal RMS (dB)	Noise RMS (dB)	SNR (dB)
Raw EMG	42.9206	17.3081	25.6125
SI	42.7470	16.4099	26.3371
Laguerre	42.9048	16.0571	26.8477
Notch	42.7065	15.8631	26.8434
RS	42.8469	15.7004	27.1465
SM	42.7028	15.5003	27.2025
ANSEC	**42.9123**	**14.2307**	**28.6816**

TABLE III

SIGNAL POWER, NOISE POWER AND SNR COMPARISONS FOR EXPERIMENT 2

	Signal RMS (dB)	Noise RMS (dB)	SNR (dB)
Raw EMG	41.9179	18.4275	23.4904
SI	41.7851	16.6308	25.1543
Laguerre	41.8771	16.1802	25.6969
Notch	41.7076	15.9846	25.7230
RS	41.8077	15.6169	26.1908
SM	41.6425	15.4269	26.2155
ANSEC	**41.8957**	**13.7441**	**28.1516**

The achievable SNR gain is directly related with the amount of PLI on the EMG signal. Therefore, the SNR gain of these algorithms may substantially differ in accordance with the application and the dataset used, e.g., some experiments in literature achieve SNR gains from 10 to 30 dB [25]–[27]. However, in our experiments, we observed SNR gains less than 5 dB we used a low noise sEMG sensor.

For both of the experiments, i.e., dumbbell (experiment 1) and random (experiment 2) contractions, we also calculated the signal waveform quality comparisons individually and illustrated them in Tables II and III respectively. In Table II, it is seen that ANSEC shows the most desired behaviour, which is the highest increase in SNR value (3.0691 dB), smallest decrease in Signal Power (−0.0082 dB) and the highest decrease in Noise Power (−3.0774 dB). The results show that ANSEC most successfully rejects the PLI without significantly disturbing the valuable EMG data. SM demonstrates the second highest increase in SNR value (1.5900 dB), which is followed by RS (1.5341 dB). However, both RS and SM significantly disturb the true EMG signal and cause loss of data around 50 Hz. In Table II and III, the Signal

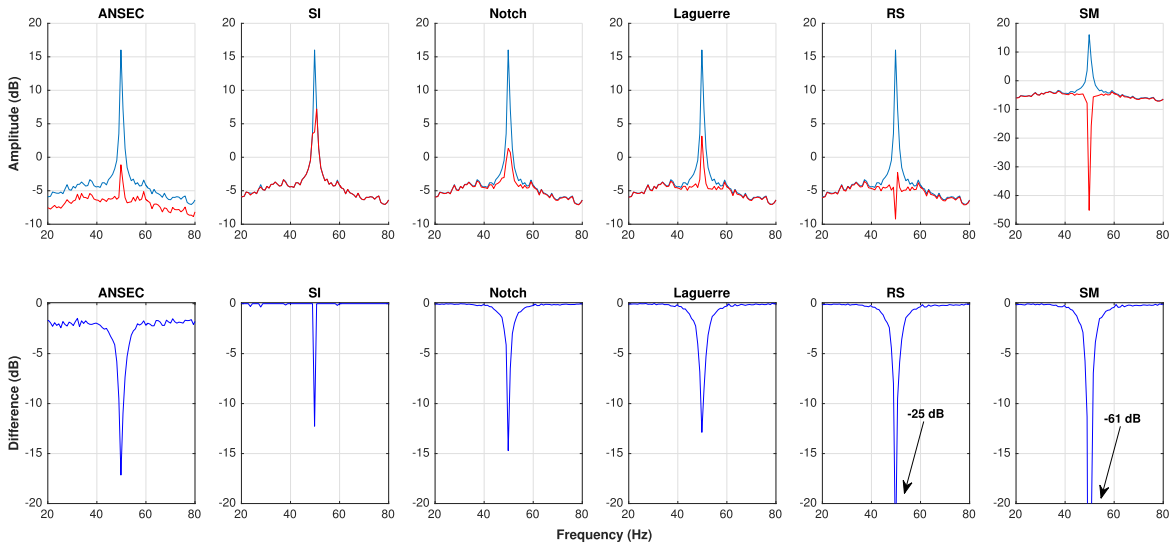

Fig. 5. Comparison of average inactive frequency responses of applied methods (top), display of output-input difference in dB scale of all methods (bottom).

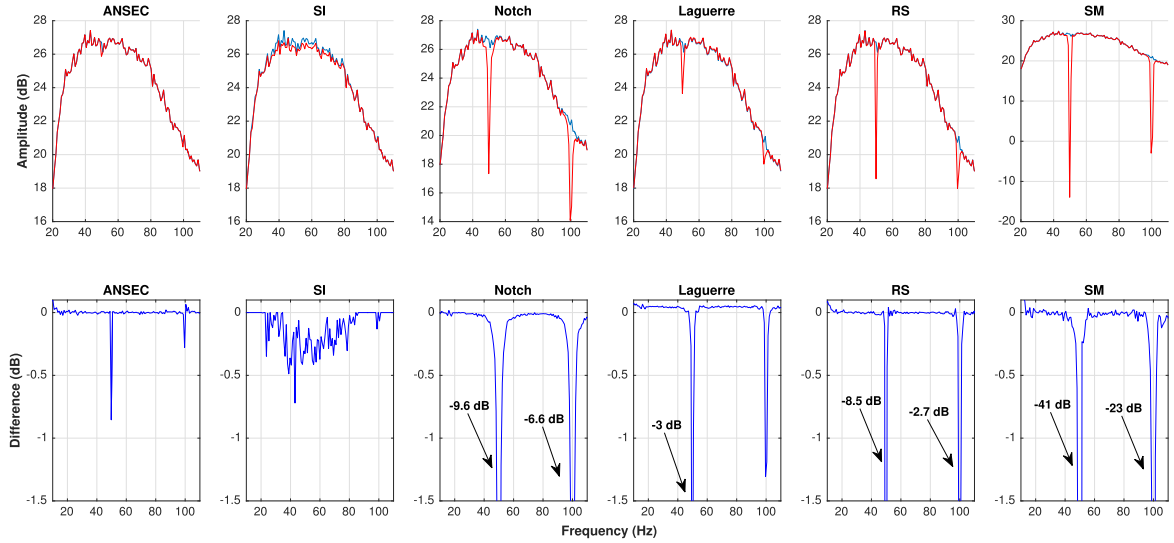

Fig. 6. Comparison of average active frequency responses of applied methods (top), display of output-input difference in dB scale of all methods (bottom).

Power decreases at most 0.2178 dB (Table II) and 0.2754 dB (Table III) by SM. Although, this may seem as a small change, the true EMG data (contractions) is significantly disturbed since the power decrease is only on frequencies around 50 Hz and its harmonics. Considering that EMG signal condenses in a frequency band of 20-150 Hz, the power decrease at 50 and 100 Hz disturbs the valuable data. In Table III, the results resemble Table II but with a slightly better performance. This time ANSEC again demonstrates the best performance by clearly increasing the SNR value most by 4.6613 dB, decreasing the Signal Power least by 0.0222 and decreasing the Noise Power most by 4.6835 dB.

D. Frequency Spectrum Analysis

In this subsection, the frequency response of each method is analyzed in order to evaluate not only the rejection amount at 50 Hz but also the extent of disturbance of true EMG data at PLI frequency. To do that, we first split all EMG data into frames of 1025 samples. The frames were classified as active and inactive as in (34) with an a posteriori selected threshold of $\tau = 3 \times 10^5$. Inactive and active spectral power densities of each subject are calculated. These spectral power densities of the subjects are then averaged in dB scale and illustrated in Figure 5 and 6 respectively. In the following subparts, the average frequency spectra of ANSEC and other five traditional methods were compared. There are two signals in each plot. The blue colored signal demonstrates the behavior of raw EMG, where the red color corresponds to the output of the specific method. In the second row of each figure the output-input difference is displayed in dB scale.

In the first part, the average spectrum of each method is evaluated for inactive frames which include mostly PLI and white noise. In Figure 5, it can be seen that although SM

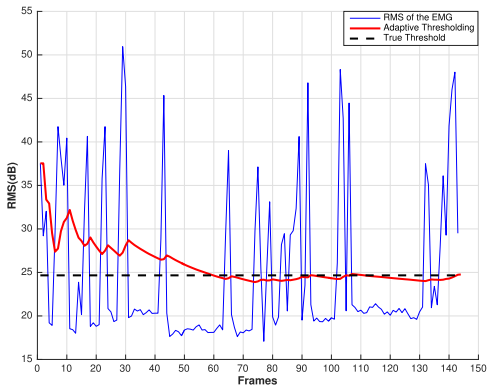

Fig. 7. RMS of S8's EMG data, true threshold and adaptive thresholding.

and RS reject the components at the PLI frequencies the most (approximately 25 dB for RS and 61 dB for SM), ANSEC has better noise cancellation performance since the noise power of the whole frequency band is decreased along with considerable PLI rejection. RS and especially SM unnaturally decrease the power of PLI frequency, which causes serious disturbance in valuable EMG data as will be described in the next subsection. Moreover, SI has the worst performance in decreasing PLI and its subbands. Even though, notch and Laguerre Filters show a similar PLI rejection with ANSEC, they are unsuccessful in significantly rejecting the subbands of PLI.

In the second part, the average spectrum of each method is evaluated for active frames, which include the muscle contractions. Active frames involve valuable EMG data, thus it is paramount to observe less rejection in order to prevent data loss. It can be seen in Figure 6 that notch, RS and SM causes a huge data loss around PLI and its harmonics. Even though SI seems not to bottom at 50 Hz, it decreases the band around 50 Hz, which causes the most data loss in active frames. ANSEC and Laguerre do not disturb the true EMG signal since the rejection amount around 50 Hz and 100 Hz is ignorable (less than 3 dB). All in all, ANSEC disturbs EMG data the least compared to other methods.

E. EMG Classifier Algorithm Validation

In our method, ANSEC, we used an adaptive way for determining active and inactive frames. The method of classification was described in III-B. In this part of the experiments section, we validate our learning algorithm of estimating a threshold value for classification by using the RMS-threshold plot of S8 for illustration purposes. In Figure 7, it is seen that our adaptive threshold converges to the threshold that perfectly splits the regions of contraction and relaxation.

V. Conclusion

We have introduced a new adaptive noise cancellation procedure (ANSEC), which successfully clears out the power line interference without disturbing the valuable sEMG signal. Our algorithm has outperformed all the state-of-the-art PLI removal methods and showed the highest SNR by decreasing the Noise Power whilst leaving the Signal Power of the sEMG signal undisturbed. Other methods significantly decrease the Signal Power disrupt the original EMG information. The main advantage of ANSEC is its adaptive behavior. Other experimented methods have poor adaptation capabilities and do not learn the frequency and amplitude of the PLI noise, which can slightly differ (±2Hz) from the traditional 50 Hz sinusoidal. Hence, the exact frequency of the PLI should be given to most of the competing algorithms for a satisfactory noise rejection. However, ANSEC estimates the exact frequency and amplitude of the PLI and rejects it accordingly. The results show that ANSEC is suitable for sEMG PLI cancellation since it adapts to the changes in PLI and demonstrates the highest noise cancellation without touching the sEMG data around 50 Hz.

References

[1] G. E. Robertson *et al.*, *Research Methods in Biomechanics*, 2nd ed. Champaign, IL, USA: Human Kinetics, 2013.
[2] W. Guo, X. Sheng, H. Liu, and X. Zhu, "Mechanomyography assisted myoeletric sensing for upper-extremity prostheses: A hybrid approach," *IEEE Sensors J.*, vol. 17, no. 10, pp. 3100–3108, May 2017.
[3] Y. Oonishi, S. Oh, and Y. Hori, "A new control method for power-assisted wheelchair based on the surface myoelectric signal," *IEEE Trans. Ind. Electron.*, vol. 57, no. 9, pp. 3191–3196, Sep. 2010.
[4] G. M. Paul, F. Cao, R. Torah, K. Yang, S. Beeby, and J. Tudor, "A smart textile based facial EMG and EOG computer interface," *IEEE Sensors J.*, vol. 14, no. 2, pp. 393–400, Feb. 2014.
[5] Z. Ju, G. Ouyang, M. Wilamowska-Korsak, and H. Liu, "Surface EMG based hand manipulation identification via nonlinear feature extraction and classification," *IEEE Sensors J.*, vol. 13, no. 9, pp. 3302–3311, Sep. 2013.
[6] M. Khezri and M. Jahed, "Real-time intelligent pattern recognition algorithm for surface EMG signals," *BioMedical Eng. OnLine*, vol. 6, p. 45, Dec. 2007.
[7] V. Rantanen *et al.*, "Capacitive measurement of facial activity intensity," *IEEE Sensors J.*, vol. 13, no. 11, pp. 4329–4338, Nov. 2013.
[8] X. Zhang, X. Chen, Y. Li, V. Lantz, K. Wang, and J. Yang, "A framework for hand gesture recognition based on accelerometer and EMG sensors," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 41, no. 6, pp. 1064–1076, Nov. 2011.
[9] D. Brunelli, E. Farella, D. Giovanelli, B. Milosevic, and I. Minakov, "Design considerations for wireless acquisition of multichannel sEMG signals in prosthetic hand control," *IEEE Sensors J.*, vol. 16, no. 23, pp. 8338–8347, Dec. 2016.
[10] T. Torfs, R. F. Yazicioglu, R. Vanspauwen, F. L. Wuyts, and C. V. Hoof, "Wireless vestibular evoked myogenic potentials system," *IEEE Sensors J.*, vol. 8, no. 12, pp. 1941–1947, Dec. 2008.
[11] C. Wong, Z.-Q. Zhang, B. Lo, and G.-Z. Yang, "Wearable sensing for solid biomechanics: A review," *IEEE Sensors J.*, vol. 15, no. 5, pp. 2747–2760, May 2015.
[12] T. Kamali, R. Boostani, and H. Parsaei, "A multi-classifier approach to MUAP classification for diagnosis of neuromuscular disorders," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 1, pp. 191–200, Jan. 2014.
[13] Z. Song, S. Guo, and Y. Fu, "Development of an upper extremity motor function rehabilitation system and an assessment system," *Int. J. Mechatron. Automat.*, vol. 1, no. 1, pp. 19–28, 2011.
[14] S. S. Nair, R. M. French, D. Laroche, and E. Thomas, "The application of machine learning algorithms to the analysis of electromyographic patterns from arthritic patients," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 2, pp. 174–184, Apr. 2010.
[15] M. Nagaraju *et al.*, "Circuit techniques for wireless bioelectrical interfaces," in *Proc. Int. Symp. VLSI Design Autom. Test (VLSI-DAT)*, Apr. 2010, pp. 117–120.
[16] M. Ergeneci, K. Gokcesu, E. Ertan, and P. Kosmas, "An embedded, eight channel, noise canceling, wireless, wearable sEMG data acquisition system with adaptive muscle contraction detection," *IEEE Trans. Biomed. Circuits Syst.*, vol. 12, no. 1, pp. 68–79, Feb. 2018.
[17] M. F. Chimene and R. Pallas-Areny, "A comprehensive model for power line interference in biopotential measurements," *IEEE Trans. Instrum. Meas.*, vol. 49, no. 3, pp. 535–540, Jun. 2000.

[18] C. D. McManus, K.-D. Neubert, and E. Cramer, "Characterization and elimination of AC noise in electrocardiograms: A comparison of digital filtering methods," *Comput. Biomed. Res.*, vol. 26, no. 1, pp. 48–67, Feb. 1993.
[19] R. C. Dugan, S. Santoso, M. F. McGranaghan, and H. W. Beaty, *Electrical Power Systems Quality*. New York, NY, USA: McGraw-Hill, 2003.
[20] R. Merletti and P. Di Torino, "Standards for reporting EMG data," *J. Electromyography Kinesiology*, vol. 9, no. 1, pp. 3–4, 1999.
[21] C. J. De Luca, L. D. Gilmore, M. Kuznetsov, and S. H. Roy, "Filtering the surface EMG signal: Movement artifact and baseline noise contamination," *J. Biomech.*, vol. 43, no. 8, pp. 1573–1579, 2010.
[22] R. M. Rangayyan, *Analysis of Concurrent, Coupled, and Correlated Processes*. Hoboken, NJ, USA: Wiley, 2015, p. 720. [Online]. Available: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7111658
[23] M. A. Haberman and E. M. Spinelli, "A multichannel EEG acquisition scheme based on single ended amplifiers and digital DRL," *IEEE Trans. Biomed. Circuits Syst.*, vol. 6, no. 6, pp. 614–618, Dec. 2012.
[24] J. Xu, R. F. Yazicioglu, B. Grundlehner, P. Harpe, K. A. A. Makinwa, and C. V. Hoof, "A 160μW 8-channel active electrode system for EEG monitoring," *IEEE Trans. Biomed. Circuits Syst.*, vol. 5, no. 6, pp. 555–567, Dec. 2011.
[25] M. Tomasini, S. Benatti, B. Milosevic, E. Farella, and L. Benini, "Power line interference removal for high-quality continuous biosignal monitoring with low-power wearable devices," *IEEE Sensors J.*, vol. 16, vol. 10, pp. 3887–3895, May 2016.
[26] D. T. Mewett, H. Nazeran, and K. J. Reynolds, "Removing power line noise from recorded EMG," in *Proc. 23rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 3, Oct. 2001, pp. 2190–2193.
[27] M. Malboubi, F. Razzazi, and S. M. Aliyari, "Elimination of power line noise from EMG signals using an efficient adaptive Laguerre filter," in *Proc. Int. Conf. Signals Electron. Circuits (ICSES)*, Sep. 2010, pp. 49–52.
[28] M. Zivanovic and M. González-Izal, "Simultaneous powerline interference and baseline wander removal from ECG and EMG signals by sinusoidal modeling," *Med. Eng. Phys.*, vol. 35, no. 6, pp. 1431–1441, Mar. 2013.
[29] J. Wang, L. Tang, and J. E. Bronlund, "Surface EMG signal amplification and filtering," *Int. J. Comput. Appl.*, vol. 82, no. 1, pp. 15–22, Nov. 2013.
[30] U. Imtiaz *et al.*, "Design of a wireless miniature low cost EMG sensor using gold plated dry electrodes for biomechanics research," in *Proc. IEEE Int. Conf. Mechatronics Autom.*, Aug. 2013, pp. 957–962.
[31] D. Scott, "Important factors in surface EMG measurement," Bortec Biomed., Calgary, AB, Canada, Tech. Rep., 2002, pp. 1–17.
[32] A. P. Vinod and C. Y. Da, "An integrated surface EMG data acquisition system for sports medicine applications," in *Proc. 7th Int. Symp. Med. Inf. Commun. Technol. (ISMICT)*, Mar. 2013, pp. 98–102.
[33] M. B. I. Reaz, M. S. Hussain, and F. Mohd-Yasin, "Techniques of EMG signal analysis: Detection, processing, classification and applications," *Biol. Procedures*, vol. 8, no. 1, pp. 11–35, Dec. 2006. [Online]. Available: https://link.springer.com/article/10.1251/bpo115
[34] R. B. R. Manero *et al.*, "Wearable embroidered muscle activity sensing device for the human upper leg," in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2016, pp. 6062–6065.
[35] S. Karlsson, J. Yu, and M. Akay, "Time-frequency analysis of myoelectric signals during dynamic contractions: A comparative study," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 2, pp. 228–238, Feb. 2000.
[36] M. Cifrek, S. Tonković, and V. Medved, "Measurement and analysis of surface myoelectric signals during fatigued cyclic dynamic contractions," *Measurement*, vol. 27, no. 2, pp. 85–92, 2000.
[37] M. Zinkevich, "Online convex programming and generalized infinitesimal gradient ascent," in *Proc. ICML*, T. Fawcett and N. Mishra, Eds., 2003, pp. 928–936.
[38] E. Hazan, A. Agarwal, and S. Kale, "Logarithmic regret algorithms for online convex optimization," *Mach. Learn.*, vol. 69, nos. 2–3, pp. 169–192, 2007.
[39] SENIAM. *Sensor Placement*. Accessed: Dec. 17, 2017. [Online]. Available: http://www.seniam.org/bicepsbrachii.html
[40] ExRx. *Dumbbell*. Accessed: Jan. 4, 2018. [Online]. Available: http://www.exrx.net/WeightExercises/Brachialis/DBConcentrationCurl.html
[41] Biometrics Ltd. *sEMG sensor*. Accessed: Mar. 3, 2018. [Online]. Available: http://www.biometricsltd.com/semg.htm

Kaan Gokcesu received the B.S. and M.S. degrees in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2015 and 2017, respectively. He is currently pursuing the Ph.D. degree in electrical engineering with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, USA. His research interests include sequential learning, adaptive filtering, and biomedical signal processing.

Mert Ergeneci received the Diploma degree in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2016. He is currently a Co-Founder of Neurocess, Inc., Research Labs, Shenzhen, Guangdong, China, focusing on developing a wearable product to make elite athlete performance analysis and injury predictions through sensors. His research interests include biomedical data analysis, adaptive filtering, and wearable electronics.

Erhan Ertan received the Diploma degree in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2015. He is currently a Co-Founder of Neurocess, Inc., Research Labs, Shenzhen, Guangdong, China, focusing on developing a wearable product to make elite athlete performance analysis and injury predictions through sensors. His research interest includes sensor circuitry.

Hakan Gokcesu received the Diploma degree in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 2015, and the M.S. degree from the School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. He is currently with the School of Computer and Communication Sciences, École Polytechnique Fédérale de Lausanne. His research interests include biomedical data analysis, machine learning, and optimization.