An Online Algorithm for Detecting and Tracking the Muscle Contraction Levels with EMG Signals

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

***Abstract*—Objective: Our objective in this work is the estimation of the muscle contraction levels (how much the muscle is contracted), because of its importance in especially gesture recognition, biomechanics, exoskeleton and prosthesis control. Methods: One widely used approach in the literature is the detection of certain waveform patterns in the signal behavior to detect muscle contractions. However, these approaches generally need specific waveforms beforehand and prone to have poor performance under stochastic environments. Another approach is the successive thresholding of certain features of the captured EMG signal. However, most of these features can have erroneous modeling of the EMG signal under low SNR and high interference scenarios (such as PLI). To this end, we propose an algorithm for the detection and tracking of the muscle contraction levels that incorporates methods of convex optimization and expert mixture. Results: For the first time in literature, we introduce an algorithm that has guaranteed performance bounds. Conclusion: Because of its performance bounds, our algorithm is robust under noisy, chaotic or even adversarial environments. Significance: Our work provides a significant tool for use especially in sports science, rehabilitation, medicine and human-machine interface, where the detection and tracking of the muscle contraction levels are of paramount importance.**

I. INTRODUCTION

The analysis of the muscle information with the evaluation and acquisition of the activation potentials through electromyography (EMG) has gained wide popularity [1]. EMG technique is applicable to a wide range of applications including but not limited to sports science and rehabilitation [2]–[10], medical diagnosis [11]–[14], human machine interaction and gesture recognition [15]–[19].

While the EMG technology provides valuable data by itself, its analysis is of paramount importance. Specifically, the amount EMG signals, the detection of contractions and the level of the contractions (i.e., how much the muscle is contracted) are especially useful in gesture recognition, muscle fatigue analysis and muscle power development [19]–[22]. To estimate and track the muscle contraction levels, an unsupervised learning approach is necessary since the contraction times and levels need to be inferred directly from the acquired EMG signals.

There are various approaches in the literature to detect and track the amount of contractions. One school of thought is to detect certain patterns in the biological signals (such as EMG, ECG). There are various QRS detection techniques, which consists of thresholding the signal in a successive manner to capture some specific behavior. These approaches generally embed some sort of transform (such as the derivative) to account for the nonlinearity. In [23], [24], the derivative is approximated with a first order difference on the discrete samples. In [25], the authors model the derivative with the difference of the output signal after a low-pass filter. Another approach is to use a Finite Impulse Response (FIR) filter to model the derivative [26]. All of these methods have nonnegative outputs similar to a full-wave rectified signal, which is than passed through some thresholds for estimation. Moreover, in [27], the authors propose an algorithm that passes the Wavelet Transform of the signal through a matched filter to detect the muscle contractions. Nonetheless, these algorithms are able to recognize the patterns in the signal with a priori knowledge of the specific waveform (consequently they may have poor performance in stochastic environments).

Because of this reason, the general approach has moved away from the detection of patterns in the EMG for detection and tracking of the muscle contractions. It has become popular to use some thresholds on certain features, which are extracted from the EMG signal observations (these metrics/features generally summarize the signal properties). These metrics include Mean Absolute Value (MAV), Variance (VAR), Root Mean Square (RMS), Waveform Length (WL), Zero Crossing (ZC), Slope Sign Change (SSC), Discrete Wavelet Transform (DWT), Wavelet Package (WPT), Mean Frequency (MF), Median Frequency (MDF), Peak Frequency (PF), Mean Power (MP), Total Power (TP), Higuchi's Fractal Dimension (HFD), Detrended Fluctuation Analysis (DFA), Shannon Entropy (SE) [28]–[32]. However, most of these metrics, especially the features related to the signal's spectral behavior can have erroneous modeling of the EMG signal under low SNR and high interference scenarios (such as power line interference, i.e., PLI [33]).

To this end, we model the muscle contractions as Gaussian processes. Since, in noisy environments (especially PLI), parameters can be estimated erroneously, we propose an adaptive algorithm that requires no pre-estimation or calibration in the beginning. Moreover, to account for the noise and interference in the environment, our approach works in an individual sequence manner and has theoretical guarantees on its performance bounds. This way, our algorithm is robust even under chaotic or adversarial environments.

The organization of the paper is as follows. In Section II, we provide some necessary preliminaries about the problem at hand. In Section III, we first provide an algorithm that can

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

H. Gokcesu is with the Department of Electrical and Electronics Engineering, 06800, Bilkent, Ankara, Turkey, (e-mail:hgokcesu@ee.bilkent.edu.tr).

E. Ertan and M. Ergeneci are with Neurocess R&D Labs, Shenzhen, Guangdong, China, (e-mail: ergenecimert@gmail.com, erhertan@gmail.com).

estimate the power level in a single contraction. In Section IV, we propose our main algorithm which can accurately detect the contractions in a sequentially received EMG signal. In Section V, we provide some insights on our performance bounds and finish with some concluding remarks.

II. Preliminaries

In this paper, we are dealing with the problem of detecting and tracking the contraction levels in a sequentially observed EMG signal in real time.

A. EMG Signal Modeling as a Gaussian Process

We formally define our problem by first modeling the EMG signal. An observed EMG signal can be mathematically modeled as the following:

$$x_t = y_t + v_t, \tag{1}$$

where y_t is the Muscle Activation Potential (MAP) and v_t is the noise at time t. Our aim is to track the contraction level in x_t at time t.

We estimate the contraction level of a muscle using its muscle activation potential y_t. However, there is no guarantee that the muscle signals will be present at each observation (there may be times when the muscle is not in contraction). Hence, if there is no contraction at time t, then the observation will simply be given by $x_t = v_t$. We model the muscle signal y_t and the noise v_t as independent zero-mean Gaussian random variables. Consequently, x_t is also a zero-mean Gaussian random variable. Let

$$x_t \in [-1, 1] \tag{2}$$

be the sample observed at time t.

We model x_t as a zero-mean Gaussian random variable with variance σ_t^2, i.e.,

$$x_t \sim \mathcal{N}(0, \sigma_t^2). \tag{3}$$

From (1), we see that σ_t^2 will be at least the noise variance (i.e., the variance of v_t) σ_N^2, i.e.,

$$\sigma_t^2 \geq \sigma_N^2. \tag{4}$$

Moreover, from the range of x_t in (2), we have

$$\sigma_t^2 \leq 1. \tag{5}$$

B. Log-loss Analysis

Our aim is to estimate the variance σ_t^2 at time t (which corresponds to the power of the sEMG signal). Since we model the observations as Gaussian random variables, we adopt a maximum likelihood based approach and maximize the probability of the observations. Based on our variance estimation at time t, let the estimated density function be $f_t(\cdot)$. Thus, the probability of observing x_t in our model will be

$$f_t(x_t) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{x_t^2}{2\sigma_t^2}\right), \tag{6}$$

since we model the observations as in (3). We measure our performance with the log-loss since minimizing log-loss corresponds to the maximization of this likelihood. Thus, at time t, we incur the loss

$$\begin{aligned} l_t(\sigma_t^2) &= -\log(f_t(x_t)), & (7)\\ &= \frac{x_t^2}{2\sigma_t^2} - \frac{1}{2}\log\left(\frac{1}{\sigma_t^2}\right) + \frac{1}{2}\log(2\pi). & (8) \end{aligned}$$

C. Performance Analysis and the Notion of Regret

Suppose we have C contractions in total that start at times $t_{a,c}$ and end at times $t_{b,c}$ for $c \in \{1, 2, \ldots, C\}$. Thus, in total, the number of time segments are upper-bounded by $2C + 1$, which consists of either all contraction samples or all noise samples.

Let t_s for $s \in \{1, 2, \ldots, 2C+1\}$ be the length of such time segment. Based on whether or not a specific time segment corresponds to a contraction, and if it does, the level of that contraction, we have an optimum variance estimation $\sigma_{s,*}^2$.

Let s_t be the segment number of the time index t (hence, $s_1 = 1$ and $s_T = 2C+1$). Then, the log-loss of these optimal estimations are given by

$$l_t(\sigma_{s_t,*}^2) = \frac{x_t^2}{2\sigma_{s_t,*}^2} - \frac{1}{2}\log\left(\frac{1}{\sigma_{s_t,*}^2}\right) + \frac{1}{2}\log(2\pi). \tag{9}$$

Our goal is to minimize the log-loss difference between our estimations σ_t^2 and the optimal estimations $\sigma_{s_t,*}^2$. Thus, we want to minimize the following notion of 'regret'

$$r_t \triangleq l_t(\sigma_t^2) - l_t(\sigma_{s_t,*}^2), \tag{10}$$

or more specifically the cumulative regret for the time horizon T, which is given by

$$\begin{aligned} R_T &\triangleq \sum_{t=1}^{T} r_t, & (11)\\ &= \sum_{t=1}^{T} l_t(\sigma_t^2) - l_t(\sigma_{s_t,*}^2). & (12) \end{aligned}$$

We will analyze the regret results in terms of its complexity with the widely used asymptotic notations.

D. Asymptotic Notations for Complexity Analysis

We first provide two asymptotic notations, which will be useful in our analysis.

1) Big-O Notation: Let T go to infinity, hence, $T \to \infty$. Let $f(T)$ and $g(T)$ be functions of T. If

$$f(T) \leq cg(T), \tag{13}$$

for some constant c, we say that

$$f(T) = O(g(T)). \tag{14}$$

2) Soft-O Notation: Let T go to infinity, hence, $T \to \infty$. Let $f(T)$ and $g(T)$ be functions of T. If

$$f(T) \leq cg(T)\log^k(g(T)), \tag{15}$$

for some constant c and k, we say that

$$f(T) = \tilde{O}(g(T)). \tag{16}$$

III. Estimating a Single Contraction Level

In this section, we first create an algorithm that can accurately estimate a fixed contraction level. Let our variance estimation (which corresponds to the power level) at time t be σ_t^2 and let the optimum fixed variance for the whole time horizon T be σ_*^2. At time t, the incurred losses are

$$l_t(\sigma_t^2) = \frac{1}{2}\frac{x_t^2}{\sigma_t^2} - \frac{1}{2}\log\left(\frac{1}{\sigma_t^2}\right) + \frac{1}{2}\log(2\pi), \tag{17}$$

$$l_t(\sigma_*^2) = \frac{1}{2}\frac{x_t^2}{\sigma_*^2} - \frac{1}{2}\log\left(\frac{1}{\sigma_*^2}\right) + \frac{1}{2}\log(2\pi). \tag{18}$$

We utilize a change of wearables which will make the underlying loss function consistently convex. Our estimation and the optimum estimation are given by

$$\theta_t \triangleq \frac{1}{\sigma_t^2}, \qquad \theta_* \triangleq \frac{1}{\sigma_*^2}. \tag{19}$$

Thus, from (17) and (19), we have the losses

$$l_t(\theta_t) = \frac{1}{2}x_t^2\theta_t - \frac{1}{2}\log(\theta_t) + \frac{1}{2}\log(2\pi), \tag{20}$$

$$l_t(\theta_*) = \frac{1}{2}x_t^2\theta_* - \frac{1}{2}\log(\theta_*) + \frac{1}{2}\log(2\pi). \tag{21}$$

The first and second derivatives of our loss $l_t(\theta_t)$ with respect to the variable θ_t are given by

$$\frac{dl_t(\theta_t)}{d\theta_t} = \frac{1}{2}x_t^2 - \frac{1}{2}\theta_t^{-1}, \tag{22}$$

$$\frac{d^2l_t(\theta_t)}{d\theta_t^2} = \frac{1}{2}\theta_t^{-2}. \tag{23}$$

We use an Online Gradient Descent (OGD) [34]–[36] based approach and update our estimation as follows

$$\theta_{t+1} = \max\left(1, \min\left(M, \theta_t - \eta_t\left(\frac{1}{2}(x_t^2 - \theta_t^{-1})\right)\right)\right) \tag{24}$$

for some $M > 1$. Henceforth, our regret at time t is given by

$$\begin{aligned} r_{t,F} &\triangleq l_t(\theta_t) - l_t(\theta_*), &(25)\\ &= \frac{1}{2}x_t^2\theta_t - \frac{1}{2}\log(\theta_t) - \frac{1}{2}x_t^2\theta_* + \frac{1}{2}\log(\theta_*), &(26)\end{aligned}$$

and the cumulative regret up to time T is

$$\begin{aligned} R_{T,F} &\triangleq \sum_{t=1}^{T} l_t(\theta_t) - l_t(\theta_*), &(27)\\ &= \sum_{t=1}^{T} \frac{1}{2}x_t^2\theta_t - \frac{1}{2}\log(\theta_t) - \frac{1}{2}x_t^2\theta_* + \frac{1}{2}\log(\theta_*). &(28)\end{aligned}$$

In the next theorem, we show our cumulative $R_{T,F}$ against a fixed power level estimation σ_*^2 (or consequently θ_*) for the time horizon T.

Theorem 1. *When the update in* (24) *is done with* $\eta_t = \dfrac{2M^2}{t}$, *we have the following bound*

$$R_{T,F} \leq M^2(\log T + 1), \tag{29}$$

where M is an upper bound on θ_t, T is the time horizon and $R_{T,F} = \sum_{t=1}^{T} l_t(\theta_t) - l_t(\theta_*)$.

Proof. The regret at time t is defined as

$$r_t \triangleq l_t(\theta_t) - l_t(\theta), \tag{30}$$

where $l_t(\theta)$ is as in (20). Assume that for some H, $l_t(\cdot)$ is H-strong convex. Thus, the regret is given by

$$r_t \leq g_t(\theta_t - \theta) - \frac{H}{2}(\theta_t - \theta), \tag{31}$$

where

$$g_t \triangleq \frac{dl_t(\theta_t)}{d\theta_t} = \frac{1}{2}x_t^2 - \frac{1}{2}\theta_t^{-1}. \tag{32}$$

We bound the first term in the right hand side of (31) using the update rule (24). We have

$$|\theta_{t+1} - \theta| \leq |\theta_t - \eta_t g_t - \theta|,$$

since θ_{t+1} is the result of a projection onto the closed convex set $[1, M]$. Hence, we get

$$\begin{aligned}(\theta_{t+1} - \theta)^2 &\leq (\theta_t - \eta_t g_t - \theta)^2,\\ &\leq (\theta_t - \theta)^2 - 2\eta_t g_t(\theta_t - \theta) + \eta_t^2 g_t^2. \end{aligned} \tag{33}$$

Since $\eta_t > 0$ for all t, rearranging (33) results in

$$g_t(\theta_t - \theta) \leq \frac{1}{2\eta_t}\left((\theta_t - \theta)^2 - (\theta_{t+1} - \theta)^2\right) + \frac{1}{2}\eta_t g_t^2. \tag{34}$$

Putting (34) in the right hand side of (31) yields

$$r_t \leq \frac{1}{2\eta_t}\left((\theta_t - \theta)^2 - (\theta_{t+1} - \theta)^2\right) + \frac{1}{2}\eta_t g_t^2 - \frac{H}{2}\|\theta_t - \theta\|^2. \tag{35}$$

Thus, summing (35) from $t = 1$ to T, we have the cumulative regret up to time T, which is given by

$$\begin{aligned} R_{T,F} &\leq \frac{1}{2}\sum_{t=2}^{T}(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}} - H)(\theta_t - \theta)^2\\ &\quad + \frac{1}{2}(\frac{1}{\eta_1} - H)(\theta_t - \theta)^2 - \frac{1}{2\eta_T}(\theta_t - \theta)^2 + \frac{1}{2}\sum_{t=1}^{T}\eta_t g_t^2,\\ &\leq \sum_{t=1}^{T}\frac{g_t^2}{2Ht},\\ &\leq \frac{1}{2H}(\log T + 1), \end{aligned}$$

where we used $\eta_t = (Ht)^{-1}$ and $g_t \leq 1$. From (23) and the fact that $\theta_t \leq M$, the second derivative of the loss function $l_t(\cdot)$ is lower bounded by $0.5M^{-2}$. Thus, we can set $H = 0.5M^{-2}$. Hence, the regret becomes

$$R_{T,F} \leq M^2(\log T + 1), \tag{36}$$

which concludes the proof of the theorem. □

From the definition of the notations in Section II-D, we see that our regret bound on the log-loss against the best fixed estimation is $O(\log(T))$ and $\tilde{O}(T^{\epsilon})$ for all $\epsilon > 0$. By an abuse of notation we will write $\tilde{O}(1)$. In the next section, we will build upon this algorithm to estimate the contraction levels in a sequentially received EMG signal.

IV. Estimating Multiple Contraction Levels

To estimate a varying number of (or multiple) contractions in a sequentially received EMG signal, we run in parallel the algorithm in Section III, which was able to track a fixed variance or power level, and aggregate their results [37]–[39].

Let N_t be the number of parallel running algorithms at time t. Then, we create our aggregated estimation θ_t as

$$\theta_t = \sum_{i=1}^{N_t} w_{i,t}\theta_{i,t}, \tag{37}$$

where $\theta_{i,t}$ is the estimation of the i^{th} parallel running algorithm and the weights $w_{i,t}$ are normalized versions (i.e., the normalized weights) of the weights $\tilde{w}_{i,t}$ of the parallel running algorithms, which are given by

$$w_{i,t} = \frac{\tilde{w}_{i,t}}{\sum_{j=1}^{N_t} \tilde{w}_{j,t}}, \tag{38}$$

so that $w_{i,t}$ creates a probability simplex at time t such that $\sum_{i=1}^{N_t} w_{i,t} = 1$.

We sequentially update the unnormalized weights $\tilde{w}_{i,t}$ with a switching principle as

$$\tilde{w}_{i,t+1} = \frac{t}{t+1}\tilde{w}_{i,t}\exp(-\lambda l_t(\theta_{i,t})) + \frac{1}{t+1}\gamma_t, \tag{39}$$

where λ is the learning rate and γ_t is given by

$$\gamma_t = \frac{1}{N_t}(\sum_{j=1}^{N_t} \tilde{w}_{i,t}\exp(-\lambda l_t(\theta_{i,t}))). \tag{40}$$

In the next theorem, we give an upper bound on the regret of the losses of the aggregated estimation in competition with the losses of an arbitrary expert selection sequence.

Lemma 1. *When the update in* (39) *is used with* $\lambda = 2M^{-2}$*, we have the following upper bound on our cumulative loss, which is*

$$\sum_{t=1}^{T} l_t(\theta_t) \leq -\frac{1}{2}M^2 \log(N_T\gamma_T) \tag{41}$$

where γ_T *is as in* (40).

Proof. Using (22), we see that the square of the derivative of the loss function in (20) is bounded as

$$\left(\frac{dl_t(\theta_t)}{d\theta_t}\right)^2 \leq \frac{1}{4}. \tag{42}$$

Moreover, from (23), we see that the second derivative is lower bounded as

$$\frac{d^2 l_t(\theta_t)}{d\theta_t^2} \geq \frac{1}{2M^2}. \tag{43}$$

Hence, the loss function in (20) is λ-exp-concave for

$$\lambda = \frac{2}{M^2}. \tag{44}$$

Using (37) and the fact that the loss functions $l_t(\cdot)$ are λ-exp-concave, we can write

$$e^{-\lambda l_t(\theta_t)} \geq \sum_{i=1}^{N_t} w_{i,t}e^{-\lambda l_t(\theta_{i,t})}. \tag{45}$$

By taking the logarithm of both sides and dividing by $(-\lambda)$, we acquire the following upper bound on the loss we incur at time t

$$l_t(\theta_t) \leq -\frac{1}{\lambda}\log\left(\sum_{i=1}^{N_t} w_{i,t}e^{-\lambda l_t(\theta_{i,t})}\right). \tag{46}$$

We use the definition in (38) inside the logarithm in (46), which gives

$$\sum_{i=1}^{N_t} w_{i,t}e^{-\lambda l_t(\theta_{i,t})} = \frac{\sum_{i=1}^{N_t} \tilde{w}_{i,t}e^{-\lambda l_t(\theta_{i,t})}}{\sum_{i=1}^{N_t} \tilde{w}_{i,t}}. \tag{47}$$

We substitute in (39) to get

$$\sum_{i=1}^{N_t} w_{i,t}e^{-\lambda l_t(\theta_{i,t})} = \frac{\sum_{i=1}^{N_t} \tilde{w}_{i,t}e^{-\lambda l_t(\theta_{i,t})}}{\sum_{i=1}^{N_t} \frac{t-1}{t}\tilde{w}_{i,t-1}e^{-\lambda l_{t-1}(\theta_{i,t-1})} + \frac{1}{t}\gamma_{t-1}} \tag{48}$$

$$= \frac{\sum_{i=1}^{N_t} \tilde{w}_{i,t}e^{-\lambda l_t(\theta_{i,t})}}{\sum_{j=1}^{N_{t-1}} \tilde{w}_{j,t-1}e^{-\lambda l_{t-1}(\theta_{j,t-1})}} \tag{49}$$

Observe that the denominator of (49) at time t corresponds to the numerator of (49) at time $t-1$. Hence, the sum of (46) from $t=1$ to T is upper bounded by

$$\sum_{t=1}^{T} l_t(\theta_t) \leq -\frac{1}{\lambda}\log\left(\sum_{i=1}^{N_T} \tilde{w}_{i,T}e^{-\lambda l_T(\theta_{i,T})}\right), \tag{50}$$

$$\leq -\frac{1}{\lambda}\log(N_T\gamma_T), \tag{51}$$

and using the fact that $\lambda = 2M^{-2}$ conclude our proof. $\square$

Even though we have an upper bound on the regret in respect to an arbitrary or best algorithm selection, we still need to provide further insight into what this result means. To this end, we next provide an upper bound that encompasses γ_T.

Lemma 2. *We have the following bound in terms of the losses of an arbitrary sequence of experts* $\{i_1, i_2, \ldots, i_T\}$ *(such that* $i_t \in \{1, 2, \ldots, N_t\}$*), which is given by*

$$-\frac{M^2}{2}\log(N_T\gamma_T) \leq \sum_{t=1}^{T} l_t(\theta_{i_t,t}) + \frac{1}{2}\hat{S}M^2\log(N_T T) \tag{52}$$

where $\hat{S} \triangleq 1 + \sum_{t=2}^{T} \mathbb{1}_{i_t \neq i_{t-1}}$.

Proof. We point out that because of the recursive calculation of $\tilde{w}_{i,t}$ coming from all possible algorithm transitions, the sum inside the logarithm in (41) includes all algorithm transition variations possible. Hence, the total incurred loss is actually upper bounded by

$$N_T\gamma_T = (\sum_{i=1}^{N_T} \tilde{w}_{i,T}\exp(-\lambda l_T(\theta_{i,T}))) \tag{53}$$

$$\geq \prod_{t=1}^{T} e^{-\lambda l_t(\theta_{i_t,t})}\left(\prod_{t=2}^{T} P_t(i_{t-1}, i_t)\right), \tag{54}$$

for some index set $\{i_1, i_2, \ldots, i_T\}$ ($i_t \in \{1, \ldots, N_t\}$ for all $t \in \{1, 2, \ldots, T\}$), where

$$P_t(i_{t-1}, i_t) \begin{cases} \frac{t-1}{t} & \text{, if } i_t = i_{t-1} \\ \frac{1}{tN_t} & \text{, if } i_t \neq i_{t-1} \end{cases}, \tag{55}$$

from the equations (39) and (40). Thus, we have

$$N_T \gamma_T \geq \prod_{t=1}^{T} e^{-\lambda l_t(\theta_{i_t,t})} \left(\frac{1}{T} \left(\frac{1}{N_T T} \right)^{\hat{S}-1} \right). \tag{56}$$

where $\hat{S} \triangleq 1 + \sum_{t=2}^{T} \mathbb{1}_{i_t \neq i_{t-1}}$. This gives us

$$-\frac{1}{\lambda} \log (N_T \gamma_T) \leq \sum_{t=1}^{T} l_t(\theta_{i_t,t}) + \frac{\hat{S}}{\lambda} \log(N_T T) \tag{57}$$

and the fact that $\lambda = 2M^{-2}$ conclude our proof. □

In our algorithm, we create our experts such that we start the k^{th} expert at time 2^{k-1} and reset every 2^k time indices. Hence our experts are created (or started) at times $t \in \{1, 2, 4, 8, 16, \ldots\}$ and they reset every $\{2, 4, 8, 16, 32, \ldots\}$ time indices. As an example, the 4^{th} expert is started at time $t = 8$ and it resets at times $t \in \{24, 40, 56, \ldots\}$. Thus, for each $\hat{S}$ time segments, we run the algorithm in Section III, which gives the following regret bound.

Lemma 3. *We have*

$$\sum_{t=1}^{T} l_t(\theta_{i_t,t}) \leq M^2 \hat{S} \log \left(\frac{2T}{\hat{S}} \right) + \sum_{t=1}^{T} l_t(\theta^*_{s_t}), \tag{58}$$

where M is an upper bound on the estimations $\theta_{i,t}$ and $\hat{S}$ is the number of time segments, θ^*_s is the best parameter estimation for s^{th} segment, and s_t is the segment number of the time index t.

Proof. Let $\hat{S}$ be the number of time segments, where each time segment s has length t_s and starts in the time index $t_{a,s}$ and ends in the time index $t_{b,s}$ for $s \in \{1, 2, \ldots, \hat{S}\}$. Let the best expert in each segment s be the $(i_s)^{th}$ expert. Then, we have

$$\sum_{t=1}^{T} l_t(\theta_{i_t,t}) = \sum_{s=1}^{\hat{S}} \sum_{t=t_{a,s}}^{t_{b,s}} l_t(\theta_{i_s,t}). \tag{59}$$

Since we use the algorithm in Section III, using Theorem 1 gives

$$\sum_{t=1}^{T} l_t(\theta_{i_t,t}) \leq \sum_{s=1}^{\hat{S}} \left(M^2 \log(2t_s) + \sum_{t=t_{a,s}}^{t_{b,s}} l_t(\theta^*_s) \right), \tag{60}$$

where θ^*_s is the optimum parameter for the s^{th} segment. From the concavity of $\log(2x)$, we get

$$\sum_{t=1}^{T} l_t(\theta_{i_t,t}) \leq M^2 \hat{S} \log \left(\frac{2T}{\hat{S}} \right) + \sum_{t=1}^{T} l_t(\theta^*_{s_t}), \tag{61}$$

where s_t is the segment number of the time index t and conclude our proof. □

After Lemma 3, we now have an upper bound on the cumulative loss of an arbitrary sequence of expert selections in terms of the losses of the best parameter estimation at time t. With the use of Lemma 1 and Lemma 2 in a successive manner, we can provide an upper bound on the cumulative loss of our parameter estimations in terms of the losses of the best estimations.

Theorem 2. *We have*

$$\sum_{t=1}^{T} l_t(\theta_t) \leq M^2 \hat{S} \log \left(\frac{2T}{\hat{S}} \right) + \frac{1}{2} \hat{S} M^2 \log(N_T T) + \sum_{t=1}^{T} l_t(\theta^*_{s_t}), \tag{62}$$

where M is an upper bound on θ_t, $\hat{S}$ is the number of time segments and N_T is the number of parallel running experts at time T.

Proof. Combining Lemmas 1, 2 and 3, gives us the result and concludes the proof. □

Suppose we have C contractions in total that start at times $t_{a,c}$ and end at times $t_{b,c}$ for $c \in \{1, 2, \ldots, C\}$. Thus, in total, the number of time segments are upper-bounded by $2C + 1$, which consists of either all contraction samples or all noise samples. Let t_s for $s \in \{1, 2, \ldots, 2C + 1\}$ be the length of a such time segment. At any given time inside these intervals, we want to switch to the expert with the greatest reset period to make as little switch as possible. In the worst case (depending on the start of the time segments), we will need to start from the expert with the least reset period and move our way up from there (i.e., switch to the 1^{st} expert then 2^{nd} and then 3^{rd} etc.). Hence, let N_s be the number of switches (number of distinct parallel running algorithms used) during the s^{th} time segment (with length t_s). With this knowledge, we next provide a lemma that gives a bound to the number of switches our algorithm makes ($\hat{S}$) in terms of the number of contractions C.

Lemma 4. *We have the following result*

$$\hat{S} \leq (2C + 1) \log_2 \left(\frac{T}{2C + 1} \right), \tag{63}$$

where C is the number of contractions and $\hat{S}$ is the number of switches our algorithm makes between the parallel running algorithms.

Proof. For the s^{th} time segment with length t_s, let us have

$$t_s = \sum_{s'=1}^{N_s - 1} 2^{s'} + b_s, \tag{64}$$

where $1 \leq b_s \leq 2^{N_s}$ and N_s is the number of switches our algorithm needs to make in the s^{th} segment. We also have

$$1 \leq b_s \leq 2^{N_s} \leq t_s, \qquad N_s \leq \log_2 t_s, \tag{65}$$

from (64). Here, we use the n^{th} expert for a duration of 2^n and the $(N_s)^{th}$ expert for a duration of b_s time. Let $\hat{S}$ be the total number of switches we make between our experts in the time horizon for $\{i_t\}_{t=1}^T$. We have

$$\hat{S} = \sum_{s=1}^{2C+1} N_s, \tag{66}$$

$$\leq \sum_{s=1}^{2C+1} \log_2(t_s), \tag{67}$$

$$\leq (2C+1)\log_2\left(\frac{T}{2C+1}\right), \tag{68}$$

from (65) and the concavity of the logarithm function, which concludes our proof.

Using Theorem 2 and Lemma 4, we next provide a theorem that bounds our cumulative regret for a time horizon T. □

Theorem 3. *We have*

$$R_{T,S} \leq 7M^2(2C+1)\log^2(2T), \tag{69}$$

where M is the upper bound of our estimation, and C is the number of contractions on the time horizon.

Proof. For a length T run of our algorithm, the total number of experts will gradually increase but never exceed $\log_2(T)+1$. Thus, we have

$$N_T \leq \log_2(2T). \tag{70}$$

Thus, using Theorem 2, Lemma 4 and inequality (70) together, we have

$$\begin{aligned}\sum_{t=1}^{T} l_t(\theta_t) \leq & M^2(2C+1)\log_2\left(\frac{T}{2C+1}\right)\log\left(\frac{2T}{2C+1}\right) \\ & + \frac{1}{2}(2C+1)M^2\log_2\left(\frac{T}{2C+1}\right)\log(\log_2(2T)T) \\ & + \sum_{t=1}^{T} l_t(\theta^*_{s_t}).\end{aligned} \tag{71}$$

Since $C \geq 0$ and $\log_2(2T) \leq T$ for $T \in \{1,2,\ldots\}$, (71) becomes

$$\begin{aligned}\sum_{t=1}^{T} l_t(\theta_t) \leq & M^2(2C+1)\log_2(T)\log(2T) \\ & + \frac{(2C+1)M^2}{2}\log_2(T)\log(T^2) + \sum_{t=1}^{T} l_t(\theta^*_{s_t}),\end{aligned} \tag{72}$$

$$\leq \frac{2}{\log(2)}M^2(2C+1)\log^2(2T) + \sum_{t=1}^{T} l_t(\theta^*_{s_t}), \tag{73}$$

$$\leq 7M^2(2C+1)\log^2(2T) + \sum_{t=1}^{T} l_t(\theta^*_{s_t}). \tag{74}$$

Therefore, our cumulative regret up to time T is given by

$$R_{T,S} \triangleq \sum_{t=1}^{T} l_t(\theta_t) - \sum_{t=1}^{T} l_t(\theta^*_{s_t}), \leq 7M^2(2C+1)\log^2(2T), \tag{75}$$

which concludes our proof. □

V. Comparisons

In a sequential EMG signal consisting of C contractions and $2C+1$ time segments, let $t_{a,s}$, $t_{b,s}$ and $\sigma^2_{s,*}$ be the start time, the end time and the power level of the s^{th} time segment respectively. Against the all powerful oracle, which knows $t_{a,s}$, $t_{b,s}$ and $\sigma^2_{s,*}$ for all $s \in \{1,2,\ldots,2C+1\}$, our algorithm can achieve a log-loss regret bound of

$$R_T \leq 7M^2(2C+1)\log^2(2T), \tag{76}$$

with a computational complexity logarithmic in time horizon T. We observe that our regret result is linear in C (the number of contractions) and logarithmic-squared in time horizon T. From the definitions of the asymptotic notations in Section II-D, we observe that our regret is $O(C\log_2(T))$ and $\tilde{O}(CT^\epsilon)$ for all $\epsilon > 0$. By an abuse of notation, we write this as $\tilde{O}(C)$.

We compare the performance of our algorithm with two weaker oracles. The first one knows the contraction times, i.e., $t_{a,s}$ and $t_{b,s}$ but does not know the power level of the contractions, i.e., $\sigma^2_{s,*}$. The second one knows the power levels $\sigma^2_{s,*}$, but does not know the contraction start and end times $t_{a,s}$ and $t_{b,s}$.

A. Oracle with Timing

If the oracle knows the start and end times $t_{a,s}$, $t_{b,s}$ of the time segments $s \in \{1,2,\ldots,2C+1\}$. It can directly estimate the power level $\sigma^2_{s,*}$ in each segment individually. For this purpose, one can use various methods, which includes the maximum likelihood and follow-the-leader based approaches. Such approaches produce $O(\log(T))$ regret for a time horizon T [34], which is minimax optimal. Since in each time segment the algorithm restarts, the cumulative regret will be $O(C\log(T))$ and consequently $\tilde{O}(C)$, which is also minimax optimal. This oracle has a constant time complexity, i.e., $O(1)$.

B. Oracle with Power Level

If the oracle knows power levels $\sigma^2_{s,*}$ of the time segments $s \in \{1,2,\ldots,2C+1\}$, it can use a mixture of experts based approach where each expert uses one of the known variances $\sigma^2_{s,*}$, which will achieve a regret bound of $O(C\log(T))$ [37], and consequently $\tilde{O}(C)$, which is minimax optimal. This oracle will have a time complexity of $O(C)$ since it runs in parallel $2C+1$ number of algorithms.

VI. Conclusion

In this paper, we introduced a completely adaptive contraction detection and tracking algorithm whose performance is similar to the oracles that have the timing or power level information of the contractions in hindsight. Our algorithm has a computational complexity that is only logarithmic in time. Hence, our algorithm is suitable for real time analysis. Moreover, even without any knowledge about the contractions, our guaranteed regret results for any observation sequence, makes our algorithm robust against noisy, chaotic or even adversarial environments. The regrets of both of these oracles (with timing or power level information) are slightly better than us by only a logarithmic in time factor in the Big-O notation and same with us in the Soft-O notation. Hence, our algorithm exhibits near optimal performance rates.

References

[1] G. Robertson, G. Caldwell, J. Hamill, G. Kamen, and S. Whittlesey, *Research methods in biomechanics, 2E*. Human Kinetics, 2013.

[2] Y. Fan and Y. Yin, "Active and progressive exoskeleton rehabilitation using multisource information fusion from emg and force-position epp," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 12, pp. 3314–3321, Dec 2013.

[3] B. J. E. Misgeld, M. Lüken, R. Riener, and S. Leonhardt, "Observer-based human knee stiffness estimation," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 5, pp. 1033–1044, May 2017.

[4] Z. Song, S. Guo, and Y. Fu, "Development of an upper extremity motor function rehabilitation system and an assessment system," *International Journal of Mechatronics and Automation*, vol. 1, no. 1, pp. 19–28, 2011.

[5] X. Navarro-Sune, A. L. Hudson, F. D. V. Fallani, J. Martinerie, A. Witon, P. Pouget, M. Raux, T. Similowski, and M. Chavez, "Riemannian geometry applied to detection of respiratory states from eeg signals: The basis for a brain;ventilator interface," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 5, pp. 1138–1148, May 2017.

[6] K. M. Vamsikrishna, D. P. Dogra, and M. S. Desarkar, "Computer-vision-assisted palm rehabilitation with supervised learning," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 5, pp. 991–1001, May 2016.

[7] M. Sartori, D. G. Llyod, and D. Farina, "Neural data-driven musculoskeletal modeling for personalized neurorehabilitation technologies," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 5, pp. 879–893, May 2016.

[8] 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 Transactions on Biomedical Circuits and Systems*, vol. 12, no. 1, pp. 68–79, Feb 2018.

[9] Y. Bai, D. Do, Q. Ding, J. A. Palacios, Y. Shahriari, M. M. Pelter, N. Boyle, R. Fidler, and X. Hu, "Is the sequence of superalarm triggers more predictive than sequence of the currently utilized patient monitor alarms?" *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 5, pp. 1023–1032, May 2017.

[10] E. Gaeta, G. Cea, M. T. Arredondo, and J. P. Leuteritz, "Amirtem: A functional model for training of aerobic endurance for health improvement," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 11, pp. 3155–3161, Nov 2012.

[11] R. I. Pettigrew, W. J. Heetderks, C. A. Kelley, G. C. Y. Peng, S. H. Krosnick, L. B. Jakeman, K. D. Egan, and M. Marge, "Epidural spinal stimulation to improve bladder, bowel, and sexual function in individuals with spinal cord injuries: A framework for clinical research," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 2, pp. 253–262, Feb 2017.

[12] T. Kamali, R. Boostani, and H. Parsaei, "A multi-classifier approach to muap classification for diagnosis of neuromuscular disorders," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 1, pp. 191–200, 2014.

[13] 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 Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 2, pp. 174–184, 2010.

[14] T. Kamali, R. Boostani, and H. Parsaei, "A multi-classifier approach to muap classification for diagnosis of neuromuscular disorders," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 1, pp. 191–200, 2014.

[15] Y. Oonishi, S. Oh, and Y. Hori, "A new control method for power-assisted wheelchair based on the surface myoelectric signal," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 9, pp. 3191–3196, 2010.

[16] 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 Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 41, no. 6, pp. 1064–1076, 2011.

[17] M. Khezri and M. Jahed, "Real-time intelligent pattern recognition algorithm for surface emg signals," *Biomedical engineering online*, vol. 6, no. 1, p. 1, 2007.

[18] S. Benatti, F. Casamassima, B. Milosevic, E. Farella, P. Schönle, S. Fateh, T. Burger, Q. Huang, and L. Benini, "A versatile embedded platform for emg acquisition and gesture recognition," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 9, no. 5, pp. 620–630, Oct 2015.

[19] R. B. R. Manero, A. Shafti, B. Michael, J. Grewal, J. L. R. Fernández, K. Althoefer, and M. J. Howard, "Wearable embroidered muscle activity sensing device for the human upper leg," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug 2016, pp. 6062–6065.

[20] 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.

[21] A. P. Vinod and C. Y. Da, "An integrated surface emg data acquisition system for sports medicine applications," in *2013 7th International Symposium on Medical Information and Communication Technology (ISMICT)*, March 2013, pp. 98–102.

[22] K. Gokcesu, M. Ergeneci, E. Ertan, A. Z. Alkilani, and P. Kosmas, "An semg-based method to adaptively reject the effect of contraction on spectral analysis for fatigue tracking," in *Proceedings of the 2018 ACM International Symposium on Wearable Computers, UbiComp 2018, Singapore, Singapore, October 8-12, 2018*, 2018, pp. 80–87. [Online]. Available: http://doi.acm.org/10.1145/3267242.3267292

[23] S. Suppappola and Y. Sun, "Nonlinear transforms of ecg signals for digital qrs detection: a quantitative analysis," *IEEE Transactions on Biomedical Engineering*, vol. 41, no. 4, pp. 397–400, April 1994.

[24] B. McCarthy, K. Stephens, C. King, E. Chabot, and Y. Sun, "Graded muscle contractions determined by temporal recruitment," in *2014 40th Annual Northeast Bioengineering Conference (NEBEC)*, April 2014, pp. 1–2.

[25] M. Okada, "A digital filter for the ors complex detection," *IEEE Transactions on Biomedical Engineering*, vol. BME-26, no. 12, pp. 700–703, Dec 1979.

[26] P. S. Hamilton and W. J. Tompkins, "Quantitative investigation of qrs detection rules using the mit/bih arrhythmia database," *IEEE Transactions on Biomedical Engineering*, vol. BME-33, no. 12, pp. 1157–1165, Dec 1986.

[27] A. Merlo, D. Farina, and R. Merletti, "A fast and reliable technique for muscle activity detection from surface emg signals," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 3, pp. 316–323, 2003.

[28] B. Hudgins, P. Parker, and R. N. Scott, "A new strategy for multifunction myoelectric control," *IEEE Transactions on Biomedical Engineering*, vol. 40, no. 1, pp. 82–94, 1993.

[29] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Fractal analysis features for weak and single-channel upper-limb emg signals," *Expert Systems with Applications*, vol. 39, no. 12, pp. 11 156–11 163, 2012.

[30] F. Garavito, J. Gonzalez, J. Cabarcas, D. Chaparro, I. Portocarrero, and A. Vargas, "Emg signal analysis based on fractal dimension for muscle activation detection under exercice protocol," in *2016 XXI Symposium on Signal Processing, Images and Artificial Vision (STSIVA)*, Aug 2016, pp. 1–5.

[31] G. Naik, D. Kumar, and S. Arjunan, "Pattern classification of myo-electrical signal during different maximum voluntary contractions: A study using bss techniques," *Measurement Science Review*, vol. 10, no. 1, pp. 1–6, 2010.

[32] R. Atri, J. S. Marquez, D. Murphy, A. Gorgey, D. Fei, J. Fox, B. Burkhardt, W. Lovegreen, and O. Bai, "Investigation of muscle activity during loaded human gait using signal processing of multi-channel surface emg and imu," in *2016 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, Dec 2016, pp. 1–6.

[33] K. Gokcesu, M. Ergeneci, E. Ertan, and H. Gokcesu, "An adaptive algorithm for online interference cancellation in emg sensors," *IEEE Sensors Journal*, pp. 1–1, 2018.

[34] E. Hazan, A. Agarwal, and S. Kale, "Logarithmic regret algorithms for online convex optimization," *Machine Learning*, vol. 69, no. 2, pp. 169–192, 2007.

[35] K. Gokcesu and S. S. Kozat, "Online density estimation of nonstationary sources using exponential family of distributions," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 9, pp. 4473–4478, Sept 2018.

[36] I. Delibalta, K. Gokcesu, M. Simsek, L. Baruh, and S. S. Kozat, "Online anomaly detection with nested trees," *IEEE Signal Processing Letters*, vol. 23, no. 12, pp. 1867–1871, Dec 2016.

[37] M. Herbster and M. K. Warmuth, "Tracking the best expert," *Machine learning*, vol. 32, no. 2, pp. 151–178, 1998.

[38] K. Gokcesu and S. S. Kozat, "Online anomaly detection with minimax optimal density estimation in nonstationary environments," *IEEE Transactions on Signal Processing*, vol. 66, no. 5, pp. 1213–1227, March 2018.

[39] ——, "An online minimax optimal algorithm for adversarial multiarmed bandit problem," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 11, pp. 5565–5580, Nov 2018.