IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 10, 2020, accepted May 22, 2020, date of publication June 8, 2020, date of current version June 16, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3000693

PROLIFIC: A Fast and Robust Profile-Likelihood-Based Muscle Onset Detection in Electromyogram Using Discrete Fibonacci Search

EASTER S. SUVISESHAMUTHU[1,2], (Member, IEEE), DIDIER ALLEXANDRE[1,2], (Member, IEEE), UMBERTO AMATO[3], BIANCAMARIA DELLA VECCHIA[4], AND GUANG H. YUE[1,2], (Member, IEEE)

[1]Center for Mobility and Rehabilitation Engineering Research, Kessler Foundation, West Orange, NJ 07052, USA
[2]Department of Physical Medicine and Rehabilitation, Rutgers New Jersey Medical School, Newark, NJ 07101, USA
[3]Istituto di Scienze Applicate e Sistemi Intelligenti Eduardo Caianiello, Consiglio Nazionale delle Ricerche, 80131 Napoli, Italy
[4]Dipartimento di Matematica, Università degli Studi La Sapienza, 00185 Roma, Italy

Corresponding author: Easter S. Suviseshamuthu (eselvan@kesslerfoundation.org)

The work of Easter S. Suviseshamuthu was supported in part by NIH/National Cancer Institute (NCI) under Grant R01CA18966501A1, and in part by the State of New Jersey Commission on Brain Injury Research under Grant CBIR15MIG004. This work was supported by the European Union's Horizon 2020 Research and Innovation Program under Grant 826589.

ABSTRACT A stochastic scheme, namely, PLM-Lap, has recently been propounded, which relies on the *profile likelihood* (PL) constructed with a *Laplace distribution* for estimating muscle activation onsets (MAOs) in surface electromyographic (sEMG) data. The MAO detection accuracy and robustness of the PLM-Lap have been empirically shown to be better than those of several state-of-the-art approaches. The algorithm designates the data point index associated with the maximum of the PL function as an onset occurrence by regarding every sEMG data point as a candidate onset and hence exhaustively evaluating the objective function. This article concerns an expedient and faster approach premised on the *discrete Fibonacci search* (DFS) to locate the maximum of the discrete PL function. The experimental results support that both the exhaustive and DFS procedures are equivalent in a statistical sense, whereas the latter offers impressive computational savings by a factor of approximately 90. Owing to the speed-up, the accuracy of MAO estimation may further be enhanced by modeling the sEMG data with a set of PL functions, each one built using a suitable probability distribution, and picking the estimate from the best model. Three statistical criteria, i.e., Kolmogorov-Smirnov, Lilliefors, and Anderson-Darling test, for choosing the probability distribution are recommended. A freely downloadable MATLAB package, namely `PROLIFIC`, meant for sEMG onset detection is available on MATLAB File Exchange from the following link: https://www.mathworks.com/matlabcentral/fileexchange/76495-prolific-profile-likelihood-based-on-fibonacci-search.

INDEX TERMS Anderson-Darling test, discrete Fibonacci search, Kolmogorov-Smirnov test, Lilliefors test, muscle activation onset, profile likelihood, surface electromyography.

I. INTRODUCTION

Accurate estimation of muscle activation onsets (MAOs) from the surface electromyographic (sEMG) data is an important research problem, due to its potential implications on the diagnosis of neuromuscular disorders [1], rehabilitation, and sport science [2], [3]. Even though the manual detection of MAOs is feasible for small datasets, it is admittedly laborious and subjective. Therefore, the MAO detection algorithms are preferred to visually pinpointing onsets from the sEMG data. For instance, wavelet transform (WT)-based algorithms are popularized, e.g., [4], since the WT can track the time instant at which the signal undergoes a sudden change. Threshold-based algorithms, on the other hand, are

The associate editor coordinating the review of this manuscript and approving it for publication was Frederico Guimarães.

simple to implement; however, they are too sensitive to signal parameters such as the signal-to-noise ratio (SNR), and hence the choice of a threshold is application-specific [5]. On the contrary, methods relying on statistical principles do not necessitate information concerning the signal properties, except the *a priori* knowledge on the data distribution [6].

Of late, a statistical method has been developed in [7], namely, PLM-Lap, based on the *profile likelihood* (PL) maximization assuming that the underlying data distribution is *Laplace*. The superior performance of PLM-Lap over several state-of-the-art algorithms has been demonstrated by implementing the algorithm with 103 sEMG signals acquired from two muscles of 18 participants (available online from https://github.com/TenanATC/EMG) and 103 simulated signals that were created as explained in [8]. Aside from its capability to accurately find muscle onsets, the PLM-Lap obviates the need for parameter tuning. Nevertheless, the PL maximization is carried out with a naive exhaustive procedure—first by considering each data point as a candidate onset and, then, by evaluating the objective (PL) function to find the data point index that corresponds to the maximum PL value. This exhaustive method requires N function evaluations with N being the total number of sEMG data points.

Toward improving the computational efficiency of PLM-Lap, we introduce the *discrete Fibonacci search* (DFS) to seek the maximum of PL function. Notwithstanding that the PLM-Lap implementation is expedited, the onset detection outcomes from the exhaustive and DFS search are statistically equivalent as shown in Section VI-A. An impressive computational advantage of the latter allowed us to explore several other distribution functions to better approximate sEMG data distributions for computing the PL and, hopefully, enhance the accuracy of the onset estimation. The best fit of distributions is determined via statistical tests. The impact of distribution choice on the MAO error is analyzed in Section VI-B. We direct the readers to [7] and the references therein for a collection of recent works on MAO detection and sEMG distribution modeling.

The key benefits are, therefore, two-pronged: (i) computational savings attained by DFS search and (ii) better (subject-adaptive) sEMG data modeling to improve the MAO detection accuracy.

II. UNDERLYING PRINCIPLE OF PLM-DFS-LAP

Essentially the PLM-Lap approach hypothesizes that the sEMG data preconditioned with the *Teager-Kaiser energy operator* (TKEO) belongs to two disparate distributions—one associated with baseline activities and the other with the muscle activation. Our empirical study in [7] demonstrates that the Laplace distribution best fits the TKEO-conditioned sEMG data. More formally, given the candidate onset index $k = 1, \ldots, N$, the objective function that can be viewed as a measure of "goodness of fit" is evaluated as follows:

$$\mathcal{L}_k(k) = \sum_{i=1}^{k} \log f(x[i]; \hat{\boldsymbol{\theta}}_1(k)) + \sum_{j=k+1}^{N} \log f(x[j]; \hat{\boldsymbol{\theta}}_2(k)) \quad (1)$$

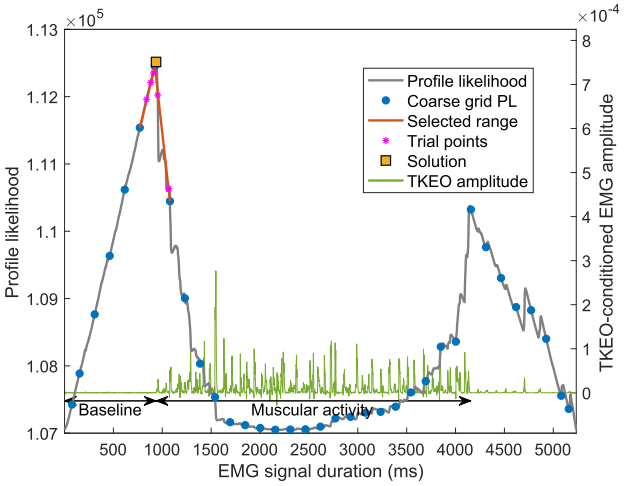

FIGURE 1. The PLM-DFS-Lap is implemented in two stages. First, a coarse search is performed in a uniform grid of finite length that is dependent on the data length (solid blue circles) to estimate a muscle onset, which supposedly lies in a close proximity to the first peak of the PL curve (gray curve). Centered around the data point that corresponds to the rough estimate of the onset, an sEMG data segment of length M (red curve) containing both the baseline and muscle activation is chosen. Second, the DFS is invoked to accurately locate the MAO by evaluating the PL function defined on the interval $[1,\ M]$ in a few selected data points (magenta asterisks) and finding its maximum (solid orange square) in $n \ll M$ evaluations. The respective TKEO-conditioned sEMG data (green) is co-located with the function landscape to show that the first peak of the PL function—MAO detected by the algorithm—coincides with the onset of muscular activity noticeable from the signal.

where $x[i]$ is the i-th data point of the sEMG segment conditioned with the TKEO, $\hat{\boldsymbol{\theta}}_1(\cdot)$ and $\hat{\boldsymbol{\theta}}_2(\cdot)$ are the *maximum likelihood* (ML) estimates (as a function of the argument in the parenthesis) of the parameters that govern the Laplace distributions $f(\cdot;\cdot)$ modeling the baseline and muscle activity, respectively. The downside of the PLM-Lap is that it mandates the evaluation of (1) at all data points to find the maximum of the function, which is presumably the best MAO estimate.

In the pursuit of reducing the number of function evaluations required to maximize (1), a two-stage procedure, namely, PLM-DFS-Lap, as illustrated in Fig. 1 is advocated in this work. Depending on the duration of the sEMG signal, the PL function can have several maxima, each one corresponding to a muscle onset, offset, and random fluctuation in the signal amplitude. As explained in the sequel, the first stage selects the segment of the signal that encompasses an MAO, whereas the second stage detects the MAO within the reduced search interval.

(i) Initially a coarse search with a uniform grid covering the entire interval finds the first peak in the fitness function landscape (PL curve), which is deemed a rough estimate of a muscle onset. In order to precisely locate the MAO, a section of the sEMG data around the onset estimate comprising baseline and muscle activation is selected.

(ii) The DFS is then employed to maximize the unimodal[1] PL function in (1) defined on the interval $[1,\ M]$ with

[1] A function $f(x)$ is said to be unimodal if for some value m it is monotonically increasing for $x \leq m$ and monotonically decreasing for $x \geq m$.

M being a finite integer depending on the duration of the signal.

III. WHY IS THE DFS PREFERRED?

For maximizing (1), one can resort to an optimizer tailored to handle discrete unimodal functions defined on an interval. For instance, a *bounded and discretized Nelder-Mead* (BDNM) algorithm proposed in [9], a *discrete golden section search* (DGSS) introduced in [10], and the DFS described in [11] are viable options. The BDNM is not guaranteed to converge to the optimum of an objective function, and the required number of function evaluations cannot be predicted. The DGSS in [10] is a rudimentary attempt to adopt the *golden section search* (GSS) to seek the optimum of an integer-valued unimodal function. We prefer the DFS algorithm to the BDNM and DGSS for the following rationale.

(i) Desirably, the GSS is linearly convergent [12]. For probing positions of the GSS at only integer sequence indices, it is widely recommended to use its variant, viz., the Fibonacci search. In other words, the DFS is inherently suitable for optimizing integer-valued unimodal functions.

(ii) At every iteration, the DFS reduces the interval that surrounds the optimum such that the width of the "bracketed" interval (in number of samples) will turn out to be a Fibonacci number, defined as:

$$F_0 = 0 \quad F_1 = 1 \quad \text{and} \quad F_{n+1} = F_n + F_{n-1}. \tag{2}$$

This strategy maximizes the amount of interval reduction, while preserving the property of a bracketing algorithm, i.e., the reuse of information concerning function evaluations at intermediate points.

(iii) Unlike other discrete optimizers, in practical instances, the maximum can be found in a very few function evaluations, i.e., $n \ll M$, where n is the largest integer such that $F_n < M$.

IV. DESCRIPTION OF PLM-DFS-LAP ALGORITHM

Given the function interval $[a, b]$, with $a = 1$ and $b = M$, the PLM-DFS-Lap algorithm initially selects two intermediate points, $p = F_{n-1}$ and $q = F_n$, from the Fibonacci sequence in (2), where n is the largest integer such that $F_n < M$. The PL function defined on the interval $[1,\ M]$ is then evaluated at the internal test points, p and q, which are initialized or determined as indicated in Algorithm in Fig. 2. By comparing the function values, $\mathcal{L}_p$ and $\mathcal{L}_q$, one of the subintervals, $[a, p)$ or $(q, b]$, is discarded based on the update rules listed in lines 6–20 in Algorithm in Fig. 2. The interval that brackets the maximum $\hat{\mathcal{L}}$ is recursively reduced until the termination criterion—for instance, $|q - p| = 1$ as suggested in [11]—is satisfied by updating the endpoints and intermediate points. For more details on the DFS search strategy, one may refer to [13].

PLM-DFS Algorithm

input: a TKEO-conditioned sEMG data segment, $x[\iota]$, $\iota = 1, \ldots, M$, comprising baseline activities and the muscle activation

output: sEMG data point index $\widehat{k}$ corresponding to a muscle onset

1: Construct the interval $[a, b]$, with $a = 1$ and $b = M$
2: Compute the largest integer n such that $F_n < M$ ▷ use (2)
3: Initialize intermediate points $p = F_{n-1}$ and $q = F_n$
4: Evaluate $\mathcal{L}_p := \mathcal{L}_p(p)$ and $\mathcal{L}_q := \mathcal{L}_q(q)$ ▷ according to (1)
5: **while** $|q - p| > 1$ **do**
6: **if** $\mathcal{L}_p > \mathcal{L}_q$ **then**
7: $b \leftarrow q$
8: $q \leftarrow p$
9: $\mathcal{L}_q \leftarrow \mathcal{L}_p$
10: $p \leftarrow a + (b - q)$
11: $\mathcal{L}_p \leftarrow \mathcal{L}_p(p)$
12: **else**
13: **if** $\mathcal{L}_p < \mathcal{L}_q$ **then**
14: $a \leftarrow p$
15: $p \leftarrow q$
16: $\mathcal{L}_p \leftarrow \mathcal{L}_q$
17: $q \leftarrow b - (p - a)$
18: $\mathcal{L}_q \leftarrow \mathcal{L}_q(q)$
19: **end if**
20: **end if**
21: **end while**
22: $\widehat{k} \leftarrow p$
23: **return** index $\widehat{k}$ of the data point pertaining to the MAO

FIGURE 2. Description of PLM-DFS algorithm maximizing (1). Depending on the choice of the distribution(s) for modeling the rest and the muscle activity, the objective function stated in (1) is altered. For example, if both distributions in (1) are selected as Laplacian, the algorithm is designated as PLM-DFS-Lap.

V. CHOICE OF DISTRIBUTIONS

The computational relief would facilitate the implementation of PLM-DFS algorithm with other, possible heavy-tailed, distributions that offer a better fit to the data, thereby enhancing the accuracy of MAO detection. The onset would correspond to the data index with the largest PL value from the respective model (if the *a priori* information on the preconditioned sEMG data distribution is unavailable, which is mostly the case). To this end, we will consider models with different distributions and various statistical methods for estimating the best distribution.

A. DISTRIBUTIONS

A wide variety of distributions exist, which can be adapted to several applications, because disparate distributions differ in their shape and support. In the present work, we consider a subset of distributions that are flexible enough to fit the distribution of sEMG signals recorded during the period of

rest or muscle activation. In particular, we also select some distributions linked to extreme events or proven to adapt to phenomena of rupture of materials or occurrence of events that are analogous to the "bursts" of muscle activation, since MAOs may be viewed as disruptive events. In this respect, an important feature of the interesting distributions is their *fat* tails. We remark that some distributions are particular cases of the ones considered in the paper and therefore are not reported.

We denote by $f(x)$ a probability distribution function and $F(x)$ the corresponding cumulative function; moreover, $\boldsymbol{x} \equiv (x_1, \ldots, x_N)$ is a random sample from the distribution. For an easier discussion, we suppose that the sample $\boldsymbol{x}$ has been sorted. We consider the following distributions:

- Gaussian distribution, defined on $]-\infty, +\infty[$:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$
$$F(x) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right],$$

  with estimates

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} x_i,\ \hat{\sigma}^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \hat{\mu}\right)^2 \qquad (3)$$

- Laplace distribution, defined on $]-\infty, +\infty[$:

$$f(x) = \frac{1}{2b}\exp\left(-\frac{|x-\mu|}{b}\right)$$
$$F(x) = \begin{cases} \frac{1}{2}\exp\left(\frac{x-\mu}{b}\right), & \text{if } x \le \mu \\ 1 - \frac{1}{2}\exp\left(-\frac{x-\mu}{b}\right) & \text{if } x > \mu, \end{cases}$$

  with estimates

$$\hat{\mu} = \text{median}(\boldsymbol{x}),\ \hat{b} = \frac{1}{N}\sum_{i=1}^{N}|x_i - \hat{\mu}| \qquad (4)$$

- Cauchy distribution, defined on $]-\infty, +\infty[$:

$$f(x) = \frac{1}{\pi\gamma\left[1 + \left(\frac{x-x_0}{\gamma}\right)^2\right]}$$
$$F(x) = \frac{1}{\pi}\arctan\left(\frac{x-x_0}{\gamma}\right) + \frac{1}{2}.$$

  Estimate of scale (γ) and location (x_0) parameters can be obtained by ML; alternatively, a computationally more efficient way is to estimate them by the following closed formulae [14]:

$$\hat{x}_0 = \frac{1}{0.24N}\sum_{\substack{x_i \ge P_{38\text{th}}(\boldsymbol{x}) \\ x_i \le P_{62\text{th}}(\boldsymbol{x})}} x_i,\ \hat{\gamma} = \text{IQR}(\boldsymbol{x}), \qquad (5)$$

  with $P_{n\text{th}}(\boldsymbol{x})$ and $\text{IQR}(\boldsymbol{x})$ being the n-th percentile and the interquartile range of $\boldsymbol{x}$, respectively.

- Logistic distribution, defined on $]-\infty, +\infty[$:

$$f(x) = \frac{\exp\left(-(x-\mu)/s\right)}{s\left(1 + \exp\left(-(x-\mu)/s\right)\right)^2}$$
$$F(x) = \frac{1}{1 + \exp\left(-(x-\mu)/s\right)}.$$

  The location (μ) and scale (s) parameters are estimated by ML.

- Lognormal distribution, defined on $[0, +\infty[$:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}}\exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right)$$
$$F(x) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{\log x - \mu}{\sigma\sqrt{2}}\right)\right],$$

  with estimates

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N}\log x_i,\ \hat{\sigma}^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(\log x_i - \hat{\mu}\right)^2 \qquad (6)$$

- Weibull distribution, defined on $]0, +\infty[$:

$$f(x) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1}\exp\left(-\left(\frac{x}{\lambda}\right)^k\right)$$
$$F(x) = 1 - \exp\left(-\left(\frac{x}{\lambda}\right)^k\right).$$

  The scale (λ) and shape (k) parameters are estimated by ML.

- Gamma distribution, defined on $]0, +\infty[$:

$$f(x) = \frac{1}{\Gamma(k)\theta^k}x^{k-1}\exp\left(-\frac{x}{\theta}\right)$$
$$F(x) = \frac{1}{\Gamma(k)}\gamma\left(k, \frac{x}{\theta}\right),$$

  with $\gamma(x)$ being the digamma function.
  The shape (k) and scale (θ) parameters can be estimated by ML. However, the closed-form solutions can be obtained by the *method of moments* (MM), which are computationally faster but less efficient, and from the ML of the generalized gamma distribution [15]:

$$\hat{k} = \frac{N\sum_{i=1}^{N}X_i}{N\sum_{i=1}^{N}X_i\log X_i - \sum_{i=1}^{N}\log X_i\sum_{i=1}^{N}X_i}$$
$$\hat{\theta} = \frac{1}{N^2}\left(N\sum_{i=1}^{N}X_i\log X_i - \sum_{i=1}^{N}\log X_i\sum_{i=1}^{N}X_i\right). \qquad (7)$$

  The estimators in (7) are endowed with a proper correction for bias.

- Birnbaum-Saunders distribution, defined on $[0, +\infty[$:

$$f(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{\left(\sqrt{\frac{x}{\beta}} - \sqrt{\frac{\beta}{x}}\right)^2}{2\gamma^2}\right)\left(\frac{\sqrt{\frac{x}{\beta}} + \sqrt{\frac{\beta}{x}}}{2\gamma x}\right)$$
$$F(x) = \Phi\left(\frac{1}{\gamma}\left(\sqrt{\frac{x}{\beta}} - \sqrt{\frac{\beta}{x}}\right)\right),$$

TABLE 1. List of the probability distributions considered for computing the PL, while implementing the Algorithm in Fig. 2. The main features of these distributions—range ($\mathbb{R}$, $\mathbb{R}_0^+$, or $\mathbb{R}^+$), symmetry, and skewness—are indicated. In addition, the method (closed formula or ML) to estimate the parameters and the test(s) suitable for estimating the best distribution, namely, Kolmogorov-Smirnov (KS) and/or Anderson-Darling (AD) and/or Lilliefors (L), are tabulated.

Distribution	Range	Symmetry	Skewness	Estimate	Test
Gaussian	$\mathbb{R}$	Yes	0	Eq. 3	KS, AD, L
Laplace	$\mathbb{R}$	Yes	0	Eq. 4	KS
Cauchy	$\mathbb{R}$	Yes	0	Eq. 5	KS
Logistic	$\mathbb{R}$	Yes	0	ML	KS
Lognormal	$\mathbb{R}_0^+$	No	+	Eq. 6	KS, AD, L
Weibull	$\mathbb{R}^+$	No	+/-	ML	KS, AD, L
Gamma	$\mathbb{R}^+$	No	+	Eq. 7	KS
Birnbaum-Sanders	$\mathbb{R}_0^+$	No	+	ML	KS
Extreme Value	$\mathbb{R}$	No	+	ML	KS, AD, L

with shape (α) and scale (β) parameters estimated by ML.

- Maximum Extreme Value distribution, defined on $]-\infty, +\infty[$:

$$f(x) = \frac{1}{\sigma}\exp\left(\frac{\mu - x}{\sigma}\right)\exp\left(-\exp\left(\frac{\mu - x}{\sigma}\right)\right)$$
$$F(x) = \exp\left(-\exp\left(\frac{\mu - x}{\sigma}\right)\right),$$

with location (μ) and scale (σ) parameters estimated by ML.

Table 1 summarizes the distributions of interest, their key features, means to estimate parameters, and statistical tests employed to select the best distribution. We remark that some distributions are defined on the entire real axis ($\mathbb{R}$), whereas the rest on the positive real axis including zero ($\mathbb{R}_0^+$) or not ($\mathbb{R}^+$). For this reason, the absolute value of TKEO-conditioned sEMG data is considered in the latter case. Moreover, we observe that a few distributions are derived from others through proper transformations; for instance, lognormal and extreme value distributions are obtained by the logarithm of a variable distributed as a Gaussian and Weibull distribution, respectively.

B. SELECTION OF DISTRIBUTIONS

Subsequent to the selection of a set of candidate distributions, a criterion for the choice of the *best*[2] distribution has to be devised. In the sequel, we have explored two possibilities: fixed distributions (i.e., independent of the subject realization) and subject-adaptive distributions.

In principle, one could naturally think of ML itself as a criterion for choosing the best distribution, that is to maximize PL not only with respect to time but also to the distribution among the candidate ones. Note that theoretical arguments tend to discourage this framework, since ML is a parametric quantity intended to estimate a (limited) number of parameters and not to estimate or compare nonparametric distributions themselves. Nevertheless, this approach is empirically investigated in Section VI.

Aside from the ML, we apply some nonparametric tests to compare how well various distributions fit the empirical data. They belong to the category of *minimum distance estimation* (MDE), which could also be adopted for estimating the parameters of a distribution as discussed in Section VII-B.

1) KOLMOGOROV-SMIRNOV TEST WITH ITS LILLIEFORS CORRECTION

Technically it relies on a statistic, D_n, given by the maximum L_1 distance between the empirical distribution function $F_n(x)$ of a signal and the target cumulative distribution $F(x)$:

$$D_n = \sup_x |F_n(x) - F(x)|,$$

estimated from the sample[3] by

$$D_n \approx \max_i \left(F(x_i) - \frac{i-1}{n}, \frac{i}{n} - F(x_i)\right). \tag{8}$$

D_n converges to a known distribution that allows one to infer p-values of the test. The main strength of the Kolmogorov-Smirnov (KS) test is that it is distribution free. However, it assumes independence between F and F_n, in other words, the parameters of distribution F have to be known in advance or by ancillary information and cannot be estimated from the empirical F_n—an assumption that does not hold for our problem. As a consequence, the true (unknown) distribution F approaches to the empirical distribution F_n, and hence p-values are biased upward. See [16] for details. This difficulty will not deter its use in our formulation, since we are interested in merely comparing the p-values among distributions and not inferring significance of fit based on actual p-values. The distribution with the smallest D_n or, equivalently, the largest p-value is hence regarded as the best among the candidate distributions. Nonetheless, we also employ an extension of the KS test due to Lilliefors [17], which is predominantly known as Lilliefors test (L). It relies on the same statistic D_n as the KS test but p-values are instead given by precomputed tables, specific for some selected families of distributions (see [18] for Gaussian, [19] for exponential, and [20] for extreme value). More accurate p-values can be obtained by a *Monte Carlo* strategy. By construction, the best distribution is selected as the one with the largest p-value. Noteworthy is that the KS test (and its Lilliefors correction) are mostly sensible to the middle part of the distribution and not to the tails, because naturally F tends to 0 and 1 at the left and right boundary for all distributions, respectively.

[2] *Best* is meant as minimizing some error indicator of an estimated onset, as specified in Sections V-B1 and V-B2 for the KS, L, and AD tests.

[3] We recall that the sample $\boldsymbol{x}$ has been ordered.

2) ANDERSON-DARLING TEST

The Anderson-Darling test (AD) [21] is based on the L_2 distance between the target and experimental cumulative distribution functions weighted by a function that weights the tails more heavily than the middle of the distributions, in contrast to the KS test. It is given by the statistic

$$A^2 = -N - S$$
$$S = \sum_{i=1}^{N} \frac{2i-1}{N} \left(\log F(x_i) + \log(1 - F(x_{N-i+1}))\right).$$

Similar to the L test, p-values are available from specific tables for selected families of distributions (e.g., [22] for Gaussian, lognormal, and Weibull distributions and [23] for logistic and extreme value distributions). Therefore, in an analogous fashion to the L test, the AD test selects the distribution with the largest p-value as the best one.

VI. EXPERIMENTAL RESULTS

The investigation has two main objectives: (i) to report the speed-up in the run time of PLM-DFS compared to that of the exhaustive-search-based PLM for a specific dataset; and (ii) to choose the best distribution that would serve as a basis for evaluating the PL function.

A. SPEED-UP OF PLM-DFS

The PL function is evaluated in [7] based on the observation that, in general, distributions of preconditioned sEMG data resemble a Laplace distribution. Therefore, in order to readily compare the run time of PLM-DFS with the algorithm proposed in [7], we have adopted the Laplace distribution as a basis for computing the PL. Interestingly, the inferences drawn with this assumption will not be affected by the choice of the distribution function that are studied in Section VI-B. Note that the execution time taken to estimate the onset heavily depends on the distribution, owing to the fact that the method employed for parameter estimation relies on the chosen distribution as shown in Table 1. The run time of the algorithm is proportional to the number of evaluations of the PL, and for each function evaluation, we have to estimate the parameters of the distribution twice—data points to the left (rest) and to the right of the candidate onset (muscle activation). For example, in the case of Gaussian distribution, parameter estimates in (3) have a closed form and incurs $O(N)$ operations; whereas, for the Laplace distribution, the closed-form solution in (4) requires the data to be sorted, which necessitates $O(N \log N)$ operations unless specific algorithms are resorted to, and in any case would result in a higher computational cost than for the Gaussian (see Section VII for more details). When a closed-form solution is not available, generally an iterative procedure has to be implemented to maximize the ML, that is even more expensive from the computational point of view. Nevertheless, irrespective of the procedure to estimate the parameters, the speed-up as a result of PLM-DFS implementation is not contingent on the selected distribution, since the cost of a single evaluation of PL remains the same for both the PLM and PLM-DFS algorithms.

TABLE 2. Mean, SD, median, and IQR of the execution time (in s) of MAO detection algorithms—PLM-Lap and PLM-DFS-Lap—to analyze sEMG data segments that were preconditioned as described in Section VI-A1.

MAO detection	Execution time (in s)			
	Mean	SD	Median	IQR
PLM-Lap	0.717	0.313	0.653	0.064 [0.624; 0.688]
PLM-DFS-Lap	0.008	0.007	0.006	0.004 [0.005; 0.009]

Besides estimating the speed-up, we also intend to infer whether the computational savings due to the DFS-based search would lead to a statistically significant degradation in the MAO detection accuracy.

1) IMPLEMENTATION

The sEMG data segments were first filtered with a second-order Butterworth low-pass (LP) filter having a high cut-off frequency of 60 Hz, and then conditioned with the TKEO operator. Each data segment was fed to the investigated approaches—PLM and PLM-DFS—and the MAO estimates from both algorithms were compared with the *gold standard*. The latter was derived by computing the mean of six (double-blind) annotations made by three researchers after visual inspection of each datum twice with a time gap of one to seven days [8].

2) SPEED-UP

Recall from Section III that the PLM-DFS-Lap (excluding the coarse search) would suffice to perform merely $n \ll M$ function evaluations to seek the maximum of (1). For instance, it requires 18 function evaluations for the sEMG data of duration 2.5 s, which has been sampled at a rate (S_r) of 1 kHz; whereas, the PLM-Lap evaluates the objective 2500 times, given the same data to analyze. Both approaches were implemented with MATLAB R2018a, and executed on a MacBook Pro computer (Intel Core i7-7820HQ 2.9 GHz CPU). The mean, standard deviation (SD), median, and the interquartile range (IQR) of the execution time taken by the PLM-Lap and PLM-DFS-Lap for detecting MAOs over 103 sEMG signals provided in the dataset are recorded in Table 2 and graphically shown in Fig. 3 as jitter-box plots.

Most importantly, the muscle onset estimation is sped up approximately 90 times on average as a desirable consequence of employing the PLM-DFS-Lap. Note that the run time consumed by the faster method takes into account the execution of both the coarse grid search and the DFS for a given sEMG input signal. More details on the computational complexity of the proposed scheme are deferred to Section VII.

3) STATISTICAL ERROR

We have verified whether maximizing (1) with the DFS-based algorithm with a much reduced number of

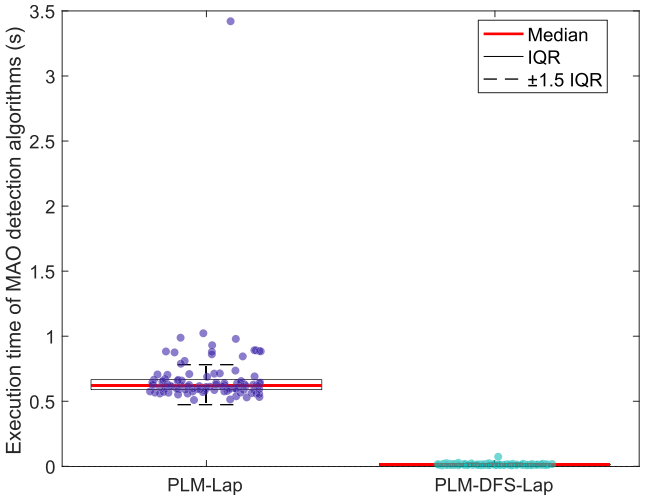

FIGURE 3. Jitter-box plots for the execution time of the onset estimation required by the PLM-Lap and PLM-DFS-Lap, which were generated by overlaying the box-and-whisker plot on the jitter plot.

TABLE 3. Mean, SD, median, and IQR of the signed and absolute MAO detection error (in ms) from PLM-Lap and PLM-DFS-Lap.

	Mean	SD	Median	IQR
Signed error from MAO estimates (in ms)				
PLM-Lap	6	93	−2	30 [−25; 5]
PLM-DFS-Lap	−4	80	−7	40 [−38; 2]
Absolute error from MAO estimates (in ms)				
PLM-Lap	46	81	21	42 [4; 46]
PLM-DFS-Lap	48	63	25	62 [5; 67]

function evaluations would induce a statistically significant degradation in the accuracy of the estimated muscle onset time. The signed error and the absolute error are shown using *jitter-box plots* in Fig. 4(a) and 4(b), respectively. The *jitters* display the error distribution in one dimension without overlapping error values, whereas the *box-and-whisker plot* visually depicts the *bias*, the spread around the median, and the outliers[4] by displaying the five-number summary of the error. Table 3 reports the mean, SD, median, and IQR of the onset error committed by both algorithms in milliseconds. Despite slight variations between the onset time predicted by the two algorithms in certain instances, the differences are not statistically significant per the inference from the *Wilcoxon's signed-rank test* (two-tail)—acceptance of null hypothesis (H_0)—at 1% level of significance as consolidated in Table 4. This implies that the PLM-Lap and PLM-DFS-Lap will yield statistically the same result, while the latter is far superior in the sense of computational efficiency.

Besides the statistical test, the *Bland-Altman plot* in Fig. 5 is meant to graphically compare the two MAO detection algorithms and evaluate the agreement among their outcomes. Since we are supplied with the *gold standard* of a muscle

[4]Whiskers are not included for the absolute error as the underlying distribution is strongly asymmetric.

TABLE 4. Wilcoxon signed-rank test to ascertain that the MAO estimates from PLM-Lap and PLM-DFS-Lap are statistically equivalent. The acceptance of null hypothesis (H_0) ensures this assumption.

Test type	Algorithm pairs	p-value	Hypothesis
Two-tail	PLM-Lap vs. PLM-DFS-Lap	0.042	H_0

onset, which was derived using six double-blind assessments of three sEMG experts (Section VI-A1), it is deemed the best estimate of the true one. Therefore, the *gold standard* onset time (in ms) is plotted on the abscissa, and the absolute error difference between the methods on the ordinate [24]. As can be noticed in Fig. 5, the fixed bias estimated by the median difference is negligible. The *precision* of the detection results is given by median ±1.5 IQR. The 95% *limits of agreement*—mean ±1.96 SD—provide an indication of how far apart the onset detection errors from both methods are more likely to be for most instances. If the differences (mean ±1.96 SD) are not clinically important, the two methods are recommended to be interchangeably used.

B. SELECTION OF DISTRIBUTION

We have investigated the influence of a chosen distribution, which serves as a basis for computing the PL function, on the outcome of PLM-DFS algorithm. The choice of distribution does affect the following: the accuracy of onset estimation and computational time. As discussed in Section VI-A, the type of distribution dictates the parameter estimation approach, thereby altering the run time of the algorithm. However, the speed-up attained by replacing the exhaustive search with DFS in maximizing the PL is independent of the distribution. Therefore, in the remainder of this section, we will focus only on how the accuracy of onset estimation is impacted by the selected distribution.

Several scenarios regarding the choice of the distribution function have been analyzed. First, we try to understand the distinction between the fixed distribution and the subject-adaptive distribution when it comes to the onset estimation accuracy. In the former case, a unique distribution is selected for all the subjects, on the other hand, the choice is subject-specific in the latter. The estimate of the best distribution is obtained with the help of statistical criteria, i.e., ML, KS, AD, and L, described in Section V-B. Furthermore, we also deduced the *oracle* distributions by taking into account the *true* onsets estimated subject-wise by experts. It should be cautioned, though, that the *oracle* can only be treated as a ''reference'' to validate the algorithms; such a virtual distribution is not useful from an operational point of view because it makes use of the *true* onset that is generally unknown. In principle, one can assume that the distributions to the left and to the right of the onset can be different.

First of all, the optimal *oracle* distributions were determined for each subject. To this end, the distributions on either side of the *true* onset were analyzed and the KS method was employed to select the optimal ones. In Table 5, the percentage of occurrence of the distributions across all the

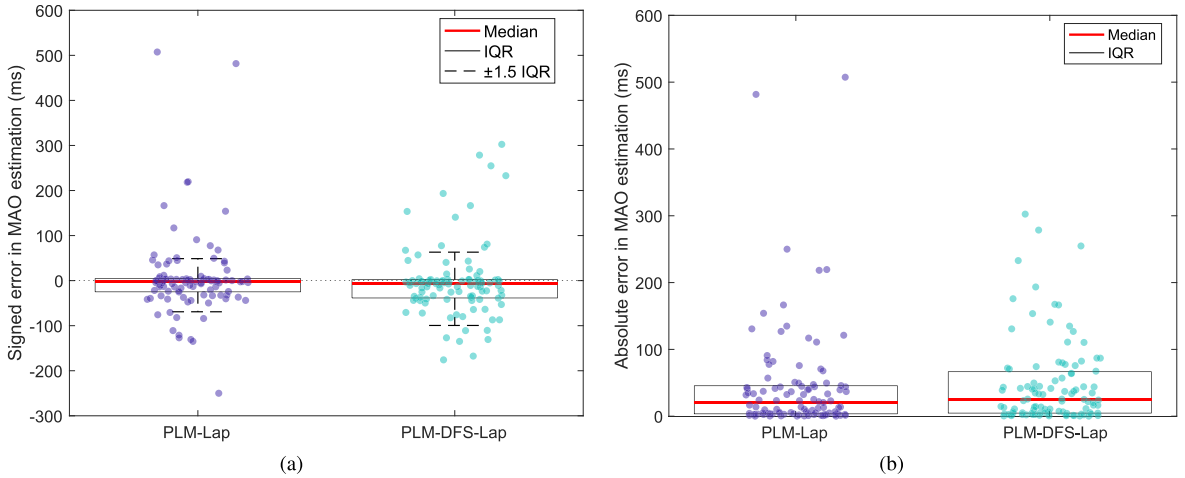

FIGURE 4. Jitter-box plots for the onset estimation error committed by the PLM-Lap and PLM-DFS-Lap, which were generated by overlaying the box-and-whisker plot on the jitter plot. (a) Signed error and (b) absolute error associated with the onsets detected from 103 sEMG signals by both methods. The signed error is useful to qualitatively assess the bias and the spread around the median, while the comparison of the mean or median absolute error renders a straightforward performance evaluation. The whiskers help identify the outliers in the plot with the signed error.

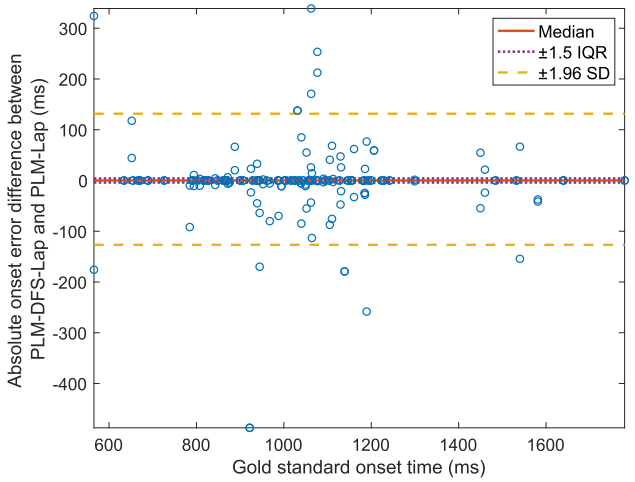

FIGURE 5. The Bland-Altman plot to examine the MAO detection agreement between the PLM-Lap and the PLM-DFS-Lap. Both algorithms were tested with 103 sEMG data segments downloaded from the web link provided in [8]. The absolute MAO error difference between the two approaches in each trial is plotted against the *gold standard* onset time in milliseconds for the respective trial with a blue circle. The median difference is denoted with a red line that represents the bias. The precision of the error differences is indicated using magenta dotted lines, whose value is calculated as median ±1.5 IQR. The orange dashed lines imply the 95% limits of agreement for the differences, and are given by mean ±1.96 SD. Since each trial is compared against itself, the horizontal spread of the error difference between trials remains immaterial.

TABLE 5. The percentage of occurrence of distributions selected by various criteria. Oracle-KS makes use of the *true* onset and the KS criterion. The initial estimate of the onset is obtained by the remaining approaches using PLM-DFS algorithm implemented with Weibull distributions; afterwards, the distributions are chosen based on the KS, L, and AD criteria on either side of the onset.

	Left			Right		
Criterion	Weibull	Lognormal	Gamma	Weibull	Lognormal	Gamma
Oracle-KS*	73.8%	10.7%	14.6%	45.6%	53.4%	1.0%
KS**	70.9%	4.8%	19.4%	59.2%	36.9%	3.9%
L	63.1%	36.9%	–	63.1%	36.9%	–
AD***	71.8%	4.8%	–	63.1%	36.9%	–

Left distribution: *Cauchy 1.0%

Left distribution: **Exponential 4.9%

Left distribution: ***Exponential 23.3%

subjects is recorded in the first row; the distributions that were never selected by any criterion are not included in the table.

In Figs. 6–9, a set of distributions are shown for subject `S01_a1`. Both cumulative (Figs. 6 and 7) and probability distributions (Figs. 8 and 9) are displayed separately for the distributions defined on $]-\infty,+\infty[$ (Figs. 6 and 8) and on $[0,+\infty)$ or $]0,+\infty)$ (Figs. 7 and 9). The upper and lower plots correspond to the TKEO-conditioned sEMG data points that lie on the left and the right side of the *true* onset, respectively. Moreover, the empirical cumulative distributions and histograms are superimposed on the best fit of several distributions listed in Section V-A. Note that a few distributions are omitted from the plots due to lack of convergence during the parameter estimation process (e.g., logistic, Birnbaum-Saunders).

Table 5 (rows 2–4) reports the occurrence of distributions (in %) chosen by KS, L, and AD test applied on the sEMG data, without using the information on the *true* onset. In this case, the onset that divides the sEMG data into the resting state and muscle activation was approximated by PLM-DFS algorithm in Fig. 2, where the objective defined in (1) is formulated with a Weibull distribution. We also modeled the rest and muscle activity by both Gaussian as well as lognormal distributions to have an initial estimate of the onset with the coarse grid; the frequency of optimal distributions obtained with these onset estimates differs from the values recorded in Table 5 (rows 2–4) by only less then 2% (except in two cases $\approx 3\%$ and 4%), and hence not shown for brevity's sake.

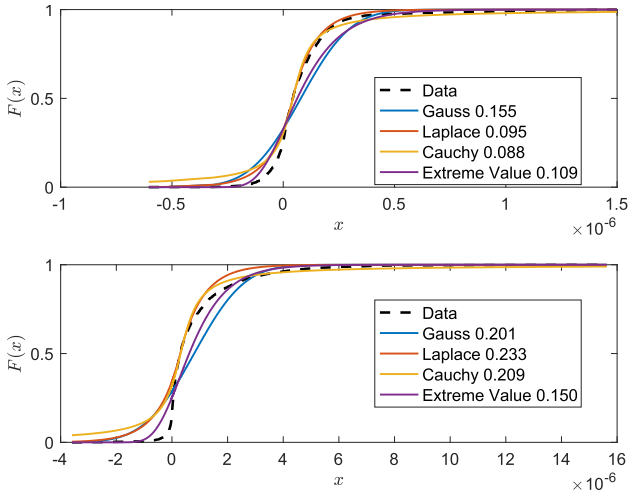

FIGURE 6. Plot of the empirical cumulative distribution function of the TKEO-conditioned sEMG signal (dashed line) and its best fit according to the Gaussian (blue), Laplace (red), Cauchy (orange), and extreme value (purple) distributions defined on $]-\infty, +\infty[$ to the left (upper plot) and to the right (lower plot) of the *true* onset. The legend also reports the statistic of the KS test given in (8). The figure refers to the subject `S01_a1`.

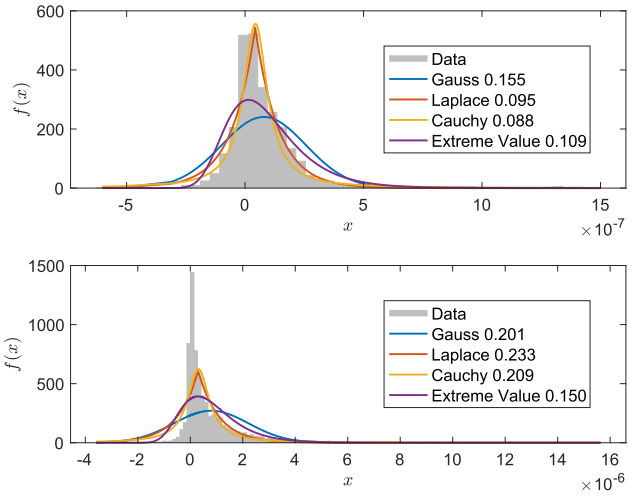

FIGURE 8. Histogram of the TKEO-conditioned sEMG signal (gray) and the corresponding best fit of the distribution functions according to the Gaussian (blue), Laplace (red), Cauchy (orange), and extreme value (purple) distributions defined on $]-\infty, +\infty[$ to the left (upper plot) and to the right (lower plot) of the *true* onset. The legend also reports the statistic of the KS test given in (8). The figure refers to the subject `S01_a1`.

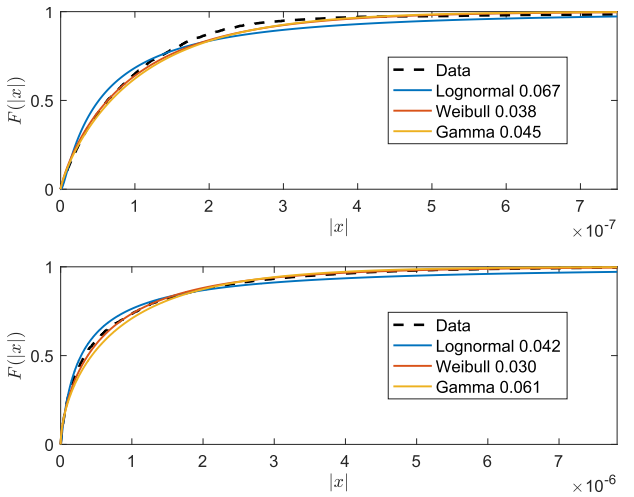

FIGURE 7. Plot of the empirical cumulative distribution function of the absolute TKEO-conditioned sEMG signal (dashed line) and its best fit according to the lognormal (blue), Weibull (red), and gamma (orange) distributions defined on $[0, +\infty[$, $]0, +\infty[$, and $]0, +\infty[$, respectively, to the left (upper plot) and to the right (lower plot) of the *true* onset. The legend also reports the statistic of the KS test given in (8). The figure refers to the subject `S01_a1`.

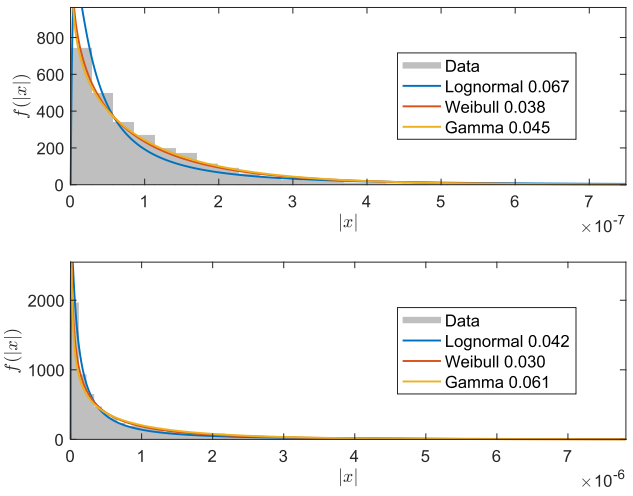

FIGURE 9. Histogram of the absolute TKEO-conditioned sEMG signal (gray) and the respective best fit of the distribution functions according to the lognormal (blue), Weibull (red), and gamma (orange) distributions defined on $[0, +\infty[$, $]0, +\infty[$, and $]0, +\infty[$, respectively, to the left (upper plot) and to the right (lower plot) of the *true* onset. The legend also reports the statistic of the KS test given in (8). The figure refers to the subject `S01_a1`.

Finally, we evaluated the performance of various algorithm scenarios summarized in Table 6 in estimating the onset. We have reported the mean, SD, median, and IQR of the absolute error of onset estimates in Table 7, which would facilitate the user to pick the optimal distribution(s) for his/her application. The onset detection results from PLM-Lap developed in [7] are treated as the reference, wherein the algorithm deploys an exhaustive search for finding the onset index by modeling the sEMG data distribution on either side of each data point as the Laplace function in every subject, thereby evaluating the PL function in all data points. In the PLM-DFS-KS, PLM-DFS-L, and PLM-DFS-AD scenarios, the initial onset estimate was obtained by approximating data distributions with the Weibull function, because this choice consistently yielded better results with respect to other distribution functions.

A few other cases were also tested and yet not reported in Table 7 for the sake of brevity, since they led to a poor estimate of the onset. The scenarios that were excluded from Table 7 due to their inferior outcome are the following:

- Estimating the distribution with the PL function.
- Employing different distributions for the PL to the left and to the right of the onset, meaning that the PL is

TABLE 6. Investigated algorithm scenarios and their design. In the subject-adaptive scenarios, namely, PLM-DFS-KS, PLM-DFS-L, and PLM-DFS-AD, the distribution function $f(x)$ chosen by the statistical tests on the left side of the onset estimate is used to model the distributions of data points that lie on either side of the onset estimate. To determine the suitable $f(x)$ for PLM-DFS-Ora implementation, the MAO errors from non-adaptive scenarios—each one making use of a distribution described in Section V-A—are compared.

Algorithm scenario	Adaptivity of distribution	Distribution/Criterion for distribution selection
PLM-Lap	Fixed	Laplacian
PLM-DFS-Lap	Fixed	Laplacian
PLM-DFS-Wei	Fixed	Weibull
PLM-DFS-Gau	Fixed	Gaussian
PLM-DFS-Log	Fixed	Lognormal
PLM-DFS-KS	Subject-wise	KS test on the left side of the onset
PLM-DFS-L	Subject-wise	L test on the left side of the onset
PLM-DFS-AD	Subject-wise	AD test on the left side of the onset
PLM-DFS-Ora	Subject-wise	Least MAO absolute error among non-adaptive scenarios

TABLE 7. Mean, SD, median, IQR, and maximum of MAO detection absolute error (in ms) from the following scenarios: PLM-Lap, PLM-DFS-Lap, PLM-DFS-Wei, PLM-DFS-Gau, PLM-DFS-Log, PLM-DFS-KS, PLM-DFS-L, PLM-DFS-AD, and PLM-DFS-Ora described in Section VI-B. The first five scenarios are non-adaptive with respect to the choice of the distribution, meaning that the distribution does not depend on the subject. Best error indicators are shown in boldface.

Algorithm scenario	Mean	SD	Median	IQR	Maximum
PLM-Lap	46	81	21	42 [4; 46]	507
PLM-DFS-Lap	48	63	25	62 [**5**; 67]	302
PLM-DFS-Wei	45	55	26	54 [7; 61]	279
PLM-DFS-Gau	**43**	49	30	59 [**5**; 64]	**213**
PLM-DFS-Log	70	120	35	67 [11; 78]	909
PLM-DFS-KS	45	61	**24**	**50** [7; **57**]	431
PLM-DFS-L	48	64	**24**	57 [7; 64]	431
PLM-DFS-AD	48	64	26	57 [7; 64]	431
PLM-DFS-Ora	35	44	20	41 [3; 44]	213

computed using the respective distribution on either side of the onset.

- Scenarios analogous to PLM-DFS-KS, PLM-DFS-L, and PLM-DFS-AD, where the common distribution for computing the PL on both sides of the onset is selected to be the one determined by KS, L, and AD, respectively, as the best distribution on the right side of the onset.

The possible reasons for the less-than-expected performance of these scenarios are discussed in Section VII-C.

We refrain from comparing the performance of PLM-DFS-Lap with the state-of-the-art approaches discussed in [7], because the outcome of PLM-Lap has been shown to have outclassed recent approaches (refer to §III-B (pp. 1286–1288) of [7]). Therefore we restrict the performance comparison to PLM-Lap, which has superseded former methods.

VII. DISCUSSION

A. SPEED-UP

Notice that the PL function in (1) is formulated using the *log-likelihood functions* instead of the *likelihood functions*. The underlying reason is the computational gain, because exponential operations are very expensive, and we can circumvent this difficulty by taking the natural logarithm[5] of the likelihood function constructed with an exponential type of distribution. In our case, the Laplace distribution is otherwise known as the *double exponential distribution*, and it lends itself to the log-likelihood-based formulation. Consequently, the computational time of both the PLM-Lap and PLM-DFS-Lap is reduced by a factor of three.

The exhaustive search for the maximum of the PL function requires N function evaluations, whereas, as shown in p. 4 of [11], the number of iterations of the Fibonacci search is limited to $\log N$ (disregarding the search step via coarse grid). Since the computational cost associated with a function evaluation remains the same in both cases, the theoretical speed-up attained by the proposed method is $N/\log N$, e.g., it is around 1000 for sample sizes of the order of 10000. Nevertheless, we compared the PLM-DFS-Lap with an "optimized version" of PLM-Lap algorithm, which is capable of iteratively restricting the part of the signal that encloses the onset being estimated. This explains the underlying reason for the disparity between the theoretical and practical speed-up (≈ 90 on an average) achieved, despite remarkable speed-up in practice. We refrain from discussing the intricacies of the "improved" PLM-Lap algorithm, since the main focus here is to investigate the PLM-DFS-Lap algorithm that supersedes the former approach. We also note that the computational complexity of a single function evaluation is $O(N)$ at best, when a closed-form solution exists for the estimate of parameters. Worth mentioning is that we also deploy a fast algorithm that returns the sample median for estimating the location parameter of the Laplace distribution, which avoids the full sorting of data points that requires $O(N \log N)$ operations, but instead finds only a fixed number of least values (one half) that incurs $O(N)$ operations [25]. Therefore the computational complexity of the Fibonacci search is limited to $O(N \log N)$. The uniform coarse grid that covers the entire range of the signal in the preliminary step is chosen to have a width $\Delta > 0$ in ms unit. Δ is a trade-off between the exhaustive search ($\Delta \to 0$) and an inconsequential two-point grid comprising only two boundary points of the time interval, where all maxima of PL are lost ($\Delta \to T$, with T being the total duration of the signal). The size of the grid is deduced as

$$N_0 = \left\lceil \frac{T}{\Delta} \right\rceil = \left\lceil \frac{N}{S_r \Delta} \right\rceil = O(N),$$

where S_r is the sampling rate of the acquisition and $\lceil \cdot \rceil$ returns the ceiling of its argument.

Based on the empirical findings with real sEMG data, we advocate the choice of $\Delta = 150$ ms, which offers a good compromise between the onset detection accuracy and the computational load. Also, the empirical study suggests that the PLM-DFS is robust with respect to Δ for Δ ranging

[5]Since the logarithm is a *strictly increasing function*, the logarithm of a function attains its maximum value at the same point as the function itself.

between 100 and 200 ms. Intuitively Δ is related to the temporal dynamics of muscle activation, i.e., slower or faster contractions may require larger or smaller values, respectively.

The choice of N_0 driven by Δ, i.e., electrophysiological conditions, reverts the overall computational complexity (including the coarse grid search) to $O(N^2)$. If one would like to preserve the order as $O(N \log N)$, then it is possible to choose $N_0 = c \log N$, with $c > 0$ being any positive constant.

B. ALTERNATIVE SCHEMES TO FIT AND SELECT DISTRIBUTIONS

To find a distribution that best fits the data, one has to solve two linked sub-problems: (i) estimate relevant parameters of a distribution from the data; (ii) select the *best* distribution according to some data driven criterion. Several methods have been developed in the literature for estimating parameters of a distribution. Most of them fall under the following categories: ML, MM, MDE, and *maximum spacing estimation* (MSE). We mainly rely on the ML, since it has excellent theoretical asymptotic properties and it naturally fits within the PL framework. Note that the PL is premised on the *likelihood*, which is the underlying principle of our methodology. Nevertheless, one can resort to other parameter estimators that may offer computational speed-up. We remark that the ML supersedes the MM as far as theoretical properties are concerned and the MSE turns out to be an approximation of ML as the sample size increases. Notable examples of MDE are KS, AD, and Cramér-von Mises.

Generally speaking, the approaches meant for estimating the parameters of a distribution could also be tailored to select the *best* distribution, by simply comparing the final values of the minimum/maximum criterion that serves as the basis for estimating the parameters. In addition, the information-theoretic criteria such as the classical *Akaike* or *Bayesian information criterion* are advocated in scenarios with an unequal number of parameters across distributions, unlike the case here.

C. CHOICE OF DISTRIBUTIONS

The Oracle-KS scenario in Table 5 implies that different functions render an optimal fit to the distribution of sEMG data lying on the left and the right side of the onset. This is due to intrinsic changes in the sEMG signal during dynamic or sustained isometric contractions. From an electrophysiological perspective, sEMG activity of a muscle is directly related to the number of motor units (MUs) recruited and their firing frequency [26]. At rest or low muscle activity, only smaller MUs are recruited at low firing frequency. The generated motor unit action potentials (MUAPs) are thus small and well separated. During contraction or higher muscle activity, there is an increase in the recruitment of larger MUs at higher frequency, resulting in overlapping MUAPs with larger and sharper spikes [26]. These intrinsic changes lead to modifications of the amplitude and frequency content of the sEMG signal, and are most likely responsible for the observed changes in its distribution profile. Moreover, we observe that the distributions of TKEO-conditioned sEMG pertaining to both rest and muscle activation are skewed (see Fig. 8) and unable to be fitted accurately by any distribution function. This difficulty is mitigated by modeling the absolute value of the TKEO-conditioned sEMG signals; one can visually appreciate a better fit of the distributions in Figs. 7 and 9, besides this claim being confirmed by the KS, L, and AD tests.

Notice that the Oracle-KS scenario in Table 1 predominantly (73.8%) chooses the Weibull distribution to model the resting state sEMG data, i.e., data to the left of the onset. The other distributions selected by the KS test include the gamma (14.6%) and lognormal (10.7%). On the other hand, the distribution functions that offer a better fit to the distribution of the muscle activation data, i.e., sEMG on the right side of the onset, turned out to be very different—the lognormal is chosen more frequently (53.4%) than the Weibull (45.6%). In a data-driven approach (scenarios KS, L, and AD), where we are bereft of or refrain from using ancillary information on the onset, the optimal distributions vary a lot from the Oracle-KS case. The Weibull is the most preferred distribution to model the sEMG data at both sides of the onset, whereas, the lognormal is deemed appropriate for fitting the sEMG data on the right side of the onset in more than one third of the subjects. Furthermore, the occurrence of distributions to the left of the onset determined by the L test differs notably from that predicted by KS and AD. On the contrary, the outcome of L test to the right of the onset is strikingly similar to that of AD. It is also worthwhile to point out that the gamma and exponential distribution are opted quite frequently to the left of the onset by the KS (19.4%) and AD test (23.3%), respectively.

The fundamental difference between the Oracle-KS and the data-driven scenarios lies in the approximation of the initial onset that separates the sEMG data into segments relating to rest (left) and muscle activities (right). In the former approach, the onset is supplied by the experts, while in the latter, it is estimated by the PLM-DFS algorithm implemented with Weibull distributions as discussed in Section VI-B. An inherent pitfall in the data-driven framework is that the initial estimate of the onset would not be as precise as the *true* onset apropos of demarcation between the rest and muscle activation, thereby causing an overlap between the two data segments. By contrast, from a physiological viewpoint, the data points on either side of the *true* onset presumably belong to two disparate distributions, in a sense that either the parameter estimates of both distributions are different or the underlying density functions are dissimilar. Therefore, it is more likely that the data overlap stemming from an inferior onset estimate would adversely interfere with the performance of data-driven scenarios. In addition, the performance of both approaches will be affected due to sEMG data that seldom includes a muscle offset (marking the cessation of muscular activities), since the offset could potentially contribute

to the degradation of ''theoretical'' distributions by mixing sEMG recorded during different conditions. Another source of difficulty is that the muscle activation is not always a well-defined on and off process.

To summarize, adaptive strategies (where the distribution functions are chosen subject-wise) model the distribution of the rectified TKEO-conditioned sEMG more often using the Weibull function and occasionally with the lognormal function. This finding corroborates with the fact that among the non-adaptive scenarios in Table 7, the one relying on the Weibull distribution is reported to have resulted in a reduced MAO detection error in terms of median and IQR range. Nevertheless, the mean error produced by the PLM-DFS-Wei is slightly more than that of PLM-DFS-Gau, since the Gaussian-based scenario handles the extreme instances more effectively as evidenced by the smallest maximum error. Interestingly enough, the scenarios with the Weibull and Gaussian distribution could outperform the Laplacian case, even though the exhaustive-search-based PLM-Lap advocated in [7] is designed with the Laplace distribution. We also observe that the use of lognormal distribution within the non-adaptive framework, despite being recommended by various tests to a certain extent (see Table 5 for percentage of occurrence), has been proven to be ineffective in estimating the onset.

What follows is the inference concerning the adaptive selection of onset by the KS, L, and AD tests. We observe that the KS strategy surpasses other adaptive as well as non-adaptive strategies with regard to the median and IQR of onset error; however, recall that the PLM-DFS-Gau (a non-adaptive strategy) managed to lower the mean error because of its ability to efficiently deal with extreme cases. A possible explanation is that a non-adaptive and less flexible framework, even though less accurate generally than an adaptive one, is less prone to pick one of the local maxima of PL that could result in outliers.

Based on the empirical findings, we conclude that the adaptive choice of distributions by KS is the best possible strategy for estimating the onset.

VIII. CONCLUSION

We have described a discrete-Fibonacci-based search strategy to speed-up the maximization of PL function, so that the muscle onset detection via a robust method, viz., PLM-Lap, can be expedited by an impressive factor[6] (≈ 90). It has also been shown that the computational savings do not degrade the accuracy of the estimated onset time, meaning that the outcomes of PLM-Lap and PLM-DFS-Lap are statistically equivalent.

The speed-up attained due to DFS in the execution of an MAO detection algorithm (as reported in Table 2) enabled us to explore various possibilities to decide the best distribution for constructing the PL function. To this purpose, we have proposed different strategies both non-adaptive (i.e., fixed for all subjects) and adaptive (i.e., subject dependent) based on statistical tests-of-fit, namely, KS, L, and AD. We have empirically verified that the adaptive selection of distributions premised on the KS test could more accurately estimate the onset. As a side note, the non-adaptive use of Weibull function did improve the results obtained in [7] by virtue of relatively better-approximated sEMG distributions with Weibull rather than Laplace function.

The MATLAB source code for a collection of PLM-DFS algorithms described in this article and the distribution choice by KS, L, and AD tests-of-fit, whose acronym is `PROLIFIC` (PROfile LIkelihood based on FIbonaCci search), is publicly available on MATLAB File Exchange at the following link: https://www.mathworks.com/matlabcentral/fileexchange/76495-prolific-profile-likelihood-based-on-fibonacci-search. The software is tailored to the detection of muscle activation onset from sEMG data. However, the same methodology can be applied to other problems, where the data can be split into two regions with different characteristics, because of the following merits: (i) the ability to select the suitable distributions from a wide variety to compute the PL function, as the optimal distributions tend to vary depending on the data at hand and (ii) the impressive computational savings. Furthermore, the MATLAB code could easily be adapted to handle other types of data by non-expert users.

[6]Actually this number is obtained for our dataset, but it could change for others.

ACKNOWLEDGMENT

The authors would like to thank Shyamala Magdalene for proofreading this article and for narration/compilation of the multimedia file. They acknowledge Joan Banks-Smith, Kessler Foundation, for help with the rehabilitation videos. Furthermore, they appreciate the anonymous reviewers and the Associate Editor for their constructive comments that enabled them improve the article.

REFERENCES

[1] C. Frigo, M. Ferrarin, W. Frasson, E. Pavan, and R. Thorsen, ''EMG signals detection and processing for on-line control of functional electrical stimulation,'' *J. Electromyogr. Kinesiol.*, vol. 10, no. 5, pp. 351–360, Oct. 2000.

[2] J. K. Leader III, J. R. Boston, and C. A. Moore, ''A data dependent computer algorithm for the detection of muscle activity onset and offset from EMG recordings,'' *Electroencephalogr. Clin. Neurophysiol./Electromyography Motor Control*, vol. 109, no. 2, pp. 119–123, Apr. 1998.

[3] G. Staude, C. Flachenecker, M. Daumer, and W. Wolf, ''Onset detection in surface electromyographic signals: A systematic comparison of methods,'' *EURASIP J. Adv. Signal Process.*, vol. 2001, no. 2, pp. 67–81, Dec. 2001.

[4] G. Vannozzi, S. Conforto, and T. D'Alessio, ''Automatic detection of surface EMG activation timing using a wavelet transform based method,'' *J. Electromyogr. Kinesiol.*, vol. 20, no. 4, pp. 767–772, Aug. 2010.

[5] K. T. Özgünen, U. Çelik, and S. S. Kurdak, ''Determination of an optimal threshold value for muscle activity detection in EMG analysis,'' *J. Sports Sci. Med.*, vol. 9, no. 4, pp. 620–628, 2010.

[6] J. Drapała, K. Brzostowski, A. Szpala, and A. Rutkowska-Kucharska, "Two stage EMG onset detection method," *Arch. Control Sci.*, vol. 22, no. 4, pp. 427–440, Dec. 2012.

[7] S. E. Selvan, D. Allexandre, U. Amato, and G. H. Yue, "Unsupervised stochastic strategies for robust detection of muscle activation onsets in surface electromyogram," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 6, pp. 1279–1291, Jun. 2018.

[8] M. S. Tenan, A. J. Tweedell, and C. A. Haynes, "Analysis of statistical and standard algorithms for detecting muscle onset with surface electromyography," *PLoS ONE*, vol. 12, no. 5, May 2017, Art. no. e0177312.

[9] E. J. Wyers, M. B. Steer, C. T. Kelley, and P. D. Franzon, "A bounded and discretized Nelder-Mead algorithm suitable for RFIC calibration," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 60, no. 7, pp. 1787–1799, Jul. 2013.

[10] K.-Y. Lian and R.-F. Liu, "A new searching method of splitting threshold values for continuous attribute decision tree problems," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Oct. 2015, pp. 1157–1160.

[11] J. Weatherwax, "An approximation algorithm for computing the mean square error between two high range resolution RADAR profiles," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 11, no. 7, pp. 1–5, Jul. 2005.

[12] L. Kiefer, "Sequential minimax search for a maximum," *Proc. Amer. Math. Soc.*, vol. 4, no. 3, pp. 502–506, Jun. 1953.

[13] D. J. Wilde, *Optimum Seeking Methods*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1964.

[14] T. J. Rothenberg, F. M. Fisher, and C. B. Tilanus, "A note on estimation from a Cauchy sample," *J. Amer. Stat. Assoc.*, vol. 59, no. 306, pp. 460–463, Jun. 1964. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/01621459.1964.10482170

[15] Z.-S. Ye and N. Chen, "Closed-form estimators for the gamma distribution derived from likelihood equations," *Amer. Statistician*, vol. 71, no. 2, pp. 177–181, Apr. 2017, doi: 10.1080/00031305.2016.1209129.

[16] W. J. Conover, *Practical Nonparametric Statistics*, 3rd ed. New York, NY, USA: Wiley, 1998.

[17] H. W. Lilliefors, "On the Kolmogorov-Smirnov test for normality with mean and variance unknown," *J. Amer. Stat. Assoc.*, vol. 62, no. 318, pp. 399–402, Jun. 1967.

[18] G. E. Dallal and L. Wilkinson, "An analytic approximation to the distribution of Lilliefors's test statistic for normality," *Amer. Statistician*, vol. 40, no. 4, pp. 294–296, Nov. 1986.

[19] H. W. Lilliefors, "On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown," *J. Amer. Stat. Assoc.*, vol. 64, no. 325, pp. 387–389, Mar. 1969.

[20] M. Chandra, N. D. Singpurwalla, and M. A. Stephens, "Kolmogorov statistics for tests of fit for the extreme-value and Weibull distributions," *J. Amer. Stat. Assoc.*, vol. 76, no. 375, pp. 729–731, Sep. 1981.

[21] M. A. Stephens, "EDF statistics for goodness of fit and some comparisons," *J. Amer. Stat. Assoc.*, vol. 69, no. 347, pp. 730–737, Sep. 1974.

[22] R. B. D'Agostino and M. A. Stephens, *Goodness-of-Fit Techniques*. New York, NY, USA: Marcel-Dekker, 1986.

[23] H. Shin, Y. Jung, C. Jeong, and J.-H. Heo, "Assessment of modified Anderson-Darling test statistics for the generalized extreme value and generalized logistic distributions," *Stochastic Environ. Res. Risk Assessment*, vol. 26, no. 1, pp. 105–114, Jan. 2012, doi: 10.1007/s00477-011-0463-y.

[24] J. S. Krouwer, "Why Bland-Altman plots should use X, not $(Y + X)/2$ when X is a reference method," *Statist. Med.*, vol. 27, no. 5, pp. 778–780, 2008.

[25] D. R. Musser, "Introspective sorting and selection algorithms," *Software, Pract. Exper.*, vol. 27, no. 8, pp. 983–993, Aug. 1997.

[26] R. Merletti and P. A. Parker, *Electromyography: Physiology, Engineering, and Non-invasive Applications*, vol. 11. Hoboken, NJ, USA: Wiley, 2004.

EASTER S. SUVISESHAMUTHU (Member, IEEE) received the B.E. degree in electronics and communication engineering from the Government College of Engineering, Tirunelveli, India, in 1988, the M.E. degree in applied electronics from Bharathiar University, Coimbatore, India, in 2001, and the Ph.D. degree in multispectral satellite image analysis from the Laboratoire des Sciences de l'Information et des Systèmes, Université de la Méditerranée, Marseille, France, in 2007.

He was a Postdoctoral Fellow with the Consiglio Nazionale delle Ricerche, Naples, Italy, from 2008 to 2010, the Université catholique de Louvain, Louvain-la-Neuve, Belgium, from 2010 to 2014, and Université Joseph Fourier, Grenoble, France, from 2014 to 2015. He has been an Associate Research Scientist with the Center for Mobility and Rehabilitation Engineering Research, Kessler Foundation, West Orange, NJ, USA, since 2015. He has also been a Research Assistant Professor of physical medicine and rehabilitation with the Rutgers New Jersey Medical School, Rutgers University, Newark, NJ, since 2018. His research interests include statistical signal processing, blind source separation, medical imaging, optimization on matrix manifolds, machine learning, bio-signal analysis, and rehabilitation engineering.

DIDIER ALLEXANDRE (Member, IEEE) received the B.S. degree in engineering from the École Nationale Supérieure de Techniques Avancées (ENSTA), Paris, France, in 1996, the M.S. degree in electrical engineering from The University of British Columbia, Vancouver, BC, Canada, in 1996, and the Ph.D. degree in biomedical engineering from Case Western Reserve University, Cleveland, OH, USA, in 2005.

He was an Engineer with General Electric Medical System, Buc, France, from 1997 to 1998. He was a Postdoctoral Fellow with the Cleveland Clinic, Lerner Research Institute, Cleveland, from 2005 to 2012. He has been a Research Scientist with the Center for Mobility and Rehabilitation Engineering Research, Kessler Foundation, West Orange, NJ, USA, since 2012. He has also been an Assistant Professor of physical medicine and rehabilitation with the Rutgers New Jersey Medical School, Rutgers University, Newark, NJ, since 2018. His research interests include investigation of the cortical and neuromuscular mechanisms of movement impairment in stroke, traumatic injury, spinal cord injury, and motor fatigue in cancer and their improvement following novel rehabilitation therapies, using advanced signal processing and neuroimaging.

Dr. Allexandre became a member of the Society for Neuroscience (SFN), in 2010, and the Organization for Human Brain Mapping (OHBM), in 2015.

UMBERTO AMATO received the M.Sc. degree in physics from the University of Napoli Federico II, Naples, Italy, in 1981.

He was a Researcher with Azienda per il Risparmio Energetico, Naples, from 1981 to 1984. He was a Professor with the High Schools, from 1984 to 1986. He was also a Researcher with the Institute for Mathematics Applications, Italian National Research Council, from 1986 to 2006. Since 2006, he has been the Director of Research. He is currently with the Istituto di Scienze Applicate e Sistemi Intelligenti, Naples. His research interests include development of statistical methods for nonparametric regression, classification, and dimension reduction applied to problems arising from remote sensing (analysis of images measured by sensors on-board aircrafts or satellites) and medicine, and virtual metrology (semiconductors).

BIANCAMARIA DELLA VECCHIA received the M.Sc. degree in mathematics (numerical methods for singular integral equations) from the University of Naples, Italy, in 1982.

She was a Researcher with the Consiglio Nazionale delle Ricerche, Naples, from 1983 to 1992. She has been an Associate Professor of numerical analysis with the Università degli Studi La Sapienza, Rome, Italy, since 1992. Her research interests include constructive approximation theory, linear positive operators, rational operators, polynomial interpolation, and numerical methods for CAGD.

GUANG H. YUE (Member, IEEE) received the Ph.D. degree in motor control/exercise science from The University of Iowa, Iowa City, IA, USA, in 1990.

From 1991 to 1993, he was a Postdoctoral Fellow with the Department of Physiology, The University of Arizona, Tucson, AZ, USA. From 1993 to 2011, he was a Faculty Member (Staff Scientist) and the Director of Neural Control Laboratory with the Department of Biomedical Engineering, Cleveland Clinic, Cleveland, OH, USA. Since 2012, he has been the Director of Human Performance and Engineering Research (Center for Mobility and Rehabilitation Engineering Research), West Orange, NJ, USA, and has also been a Professor of physical medicine and rehabilitation with the Rutgers New Jersey Medical School, Rutgers University, Newark, NJ. His research interests include human motor control in health and disease, neuromuscular rehabilitation, and rehabilitation engineering.