Signal Processing 114 (2015) 34–44

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Legendre polynomial modeling of time-varying delay applied to surface EMG signals—Derivation of the appropriate time-dependent CRBs

Abdelbassit Boualem, Meryem Jabloun *, Philippe Ravier, Olivier Buttelli

Univ. Orleans, PRISME, EA 4229, F45072, Orleans, France

ARTICLE INFO

Article history:
Received 12 June 2014
Received in revised form
9 October 2014
Accepted 1 February 2015
Available online 20 February 2015

Keywords:
Surface electromyography
Conduction velocity
Time-varying delay
Maximum likelihood estimator
Cramer–Rao lower bound
Simulated annealing

ABSTRACT

In surface electromyography (sEMG), the assessment of the muscle fiber conduction velocity (CV) can be performed by estimating the time delay between sEMG signals recorded by means of electrodes. During the dynamic contractions that often occur in daily life, the sEMG signals are not stationary which, induces time-varying delays (TVD). In the present study, the problem of TVD estimation is considered using a new parametric method. First, we propose a polynomial modeling of the TVD. The model is expressed on an orthonormal polynomial basis. Then, the TVD model parameters are estimated by using a maximum likelihood (ML) strategy solved by a stochastic optimization technique, called simulated annealing (SA).

The orthogonality property allows parameter decoupling, which helps enhance the estimation accuracy. The SA algorithm helps escape local minima, thereby keeping the optimality property benefit of the ML estimator. We also derive two appropriate Cramer–Rao lower bounds (CRB) for the estimated TVD model parameters and for the TVD waveforms. Simulation results show that the estimation of both the model parameters and the TVD function is unbiased and that the mean square error (MSE) obtained is close to the derived CRBs. A comparison with nonparametric approaches of TVD estimation is also presented and shows the superiority of the method proposed.

1. Introduction

The estimation of time-varying delay (TVD) is of great interest in many research fields [1–7]. This task is however a real challenge in the field of bioelectrical signals, especially for surface electromyography (sEMG) signals [4,8–11], since the TVD between two or more EMG signals is tightly related to the muscle fiber conduction velocity (MFCV). MFCV is an interesting indicator of muscle activity. Muscle is composed of various muscle fibers organized in functional units known as motor units (MU), i.e. one motoneuron innervating a group of muscle fibers [12]. The MFCV reflects the global propagation velocity of the motor unit action potential (MUAP) generated by the activated MU. It provides information about neuromuscular control including the features and working conditions of the MU. The MFCV is therefore an useful tool for studying fatigue and neuromuscular diseases [9,13]. Moreover, the MFCV variation reflects the modification of the type of the motor unit being active during the muscle contraction. This gives indications about the motor unit recruitment strategies [14,15].

Several classical methods for TVD estimation assume stationary sEMG signals and a constant time delay within

* Corresponding author.
E-mail addresses: abdelbassit.boualem@univ-orleans.fr (A. Boualem), meryem.jabloun@univ-orleans.fr (M. Jabloun), philippe.ravier@univ-orleans.fr (P. Ravier), olivier.buttelli@univ-orleans.fr (O. Buttelli).

http://dx.doi.org/10.1016/j.sigpro.2015.02.011
0165-1684/

the analysis time window [7,16]. However the non-constant force levels during fast dynamic contractions induce physiological modifications in the recruited muscle fiber features and consequently a time-varying MFCV. Therefore the sEMG signals should be assumed nonstationary and the time delays nonconstant [10].

Few published methods address TVD estimation [10,11,17]. In [10], the authors assumed a TVD during bursts of EMG activity under dynamic contractions. They proposed to divide each burst into segments which are short enough for the time delay to be considered locally constant. The maximum likelihood estimation (MLE) of the local time delays was achieved using an iterative Newton-gradient technique applied to sliding windows. However, this optimization technique does not guarantee convergence toward a global extremum, which adversely affects the estimation. Moreover, when applied to real data this method is highly sensitive to the noise level.

In [18], the time delay was modeled by a finite impulse response filter. The filter coefficients were estimated by using an iterative recursive least squares algorithm. The proposed algorithm is able to track fractional delay changes using a very short time window. However, the performance of the algorithm is poorer when the analyzed signals are not white. To improve the robustness of this algorithm, the authors in [11] proposed decimation and whitening preprocessing.

In [17], three methods were investigated based on time-frequency or time-scale representations to follow time delay changes. At each instant of these representations, the phase difference between two EMG recorded channels was approximately a linear curve with respect to frequency. The curve slope was proportional to the time delay. The performance of these methods is highly dependent on the data power spectrum density and the analysis window parametrization.

In the present paper, we propose to model the TVD as a polynomial function varying over time. This model is a non-discontinuous time function that takes into account the physiological modifications during dynamic contraction. Indeed, abrupt variations in the MFCV cannot occur due to physiological considerations related to the genesis mechanism of the action potentials. Moreover, thanks to a TVD polynomial model, the number of parameters to be estimated is reduced: only (polynomial degree $+1$) parameters are required. Additionally, we propose to define the TVD polynomial model using an orthonormal polynomial basis. The orthogonality property allows decoupled estimation parameters and thus helps reduce the estimation error. The estimation of the TVD polynomial model parameters was carried out using a maximum likelihood procedure (ML) solved by a stochastic optimization method, simulated annealing. This stochastic optimization technique has the ability to avoid local extrema which helps keep the optimality and efficiency properties of the ML procedure.

In [19], we derived time-dependent CRBs appropriate to polynomial TVD functions. In the present paper, we generalize the CRB derivation to nonpolynomial TVD functions. Monte Carlo simulations of sEMG signals are carried out to demonstrate the robustness of the proposed approach. The obtained variances and biases of the estimated TVD functions are then compared to the derived CRBs.

The paper is organized as follows. Section 2 presents the sEMG acquisition model and the synthetic sEMG generating model. Section 3 details the TVD estimation method that we propose. The appropriate CRBs are derived in Section 4. Section 5 shows the estimation results and a comparison to the CRBs. Finally, Section 6 concludes and presents some perspectives for work in progress.

2. sEMG modeling

In this section, we first present an analytical model for two-channel sEMG acquired signals, and then a generating model of synthetic sEMG signals. This model is helpful for statistical performance studies.

2.1. Time-acquisition model

We consider a simplified model of the sEMG signal propagation with no scaling and amplitude distortion factors, as these factors can be neglected for small inter-electrode distances (smaller than 5–7 mm). Let us consider a sEMG signal $s(n)$ that propagates along the muscle fiber. An analytical discrete model of two sEMG signals recorded by two spatially separated sensors is given by [11,17,19]

$$\begin{aligned} x_1(n) &= s(n)+w_1(n) \\ x_2(n) &= s(n-\theta(n))+w_2(n) \quad n=0,\ldots,N-1 \end{aligned} \tag{1}$$

where N is the sample size and $\theta(n)$ is the TVD. w_1 and w_2 are additive white Gaussian noises (AWGN). The noises are assumed to be mutually independent and identically distributed with zero mean and equal variance σ^2. $\theta(n)$ is linked to the CV which is the commonly used indicator for physiological studies through

$$\mathrm{CV}(n)=\frac{F_s\Delta e}{\theta(n)} \tag{2}$$

where F_s is the sampling frequency and Δe is the inter-electrode distance.

An efficient CV estimation needs an accurate estimation of the TVD $\theta(n)$ which is the purpose of the present paper. However, improving the TVD estimation is a real challenge in the sEMG field given the specific features of sEMG. In addition to their non-stationarity, sEMG signals have neither narrow nor wide spectral bands as will be detailed in the following section.

2.2. Synthetic sEMG generating model

The generating model of synthetic sEMG signals must reach real-like sEMG spectrum features by providing synthetic sEMG signals whose spectra are neither narrow nor wide band [4,20]. Such a spectrum can be modeled by [20]

$$\mathrm{PSD}(f)=\frac{Kf_l^2f^2}{(f^2+f_l^2)(f^2+f_h^2)^2} \tag{3}$$

where f_l and f_h are two frequency parameters that reflect changes on the shape of the power spectral density PSD(f) [4].

Therefore, synthetic sEMG signals can be generated using stochastic white Gaussian processes filtered through the square root of Eq. (3) [4,8,20]. A delayed version $s(n-\theta(n))$

of $s(n)$ is obtained using a cardinal sine interpolation [18]

$$s(n-\theta(n)) = \sum_{m=-M}^{M-1} \mathrm{sinc}(m-\theta(n))s(n-m). \quad (4)$$

3. Proposed TVD estimation method

This section focuses on the sEMG TVD modeling and estimation. First we propose to model the TVD as a polynomial time function. Then the model parameter estimation is achieved using the likelihood procedure solved by a stochastic optimization technique.

3.1. TVD polynomial function modeling

In [11,16,17], nonparametric TVD estimation was addressed. These nonparametric approaches lead to a high-dimension optimization problem which is intractable especially for the ML procedure. To reduce the dimensionality of the problem, the authors focus on local estimation using sliding window analysis. However, this still requires estimating the same number of parameters as the data sample size. In the present paper, we reduce the dimension of the estimation problem by considering a parametric approach. Since the TVD is a continuous time-varying function due to physiological considerations [21], $\theta(n)$ (1) can be approximated by a polynomial function according to the Weierstrass theorem [22]:

$$\theta(n) \simeq \theta^{(d)}(n) = \sum_{i=0}^{d} \Theta_i P_i(n) \quad (5)$$

where $\theta^{(d)}(n)$ is the TVD polynomial approximation of degree d. $\{P_i(n)\}_{i=0,\ldots,d}$ is a set of discrete polynomial basis functions of degree i. The model parameters to be estimated are

$$\underline{\Theta}^{(d)} = [\Theta_0, \ldots, \Theta_d] \quad (6)$$

and these parameter values are constrained by the physiological ranges of CV(n) and its kinetics (first derivative) [23]. Thanks to (5), only $d+1$ parameters have to be estimated. We propose to use an orthonormal polynomial basis since orthogonality allows parameter uncoupling while normalization unifies the range order of these parameters. In this study, the well-known Legendre polynomials are used to create the orthonormal polynomial basis [24].

3.2. Likelihood function

We propose to estimate the TVD using an ML estimator in order to keep its attractive asymptotic properties such as consistency, normality and efficiency [25]. First, let us consider the vector $\underline{\theta} = [\theta(0), \ldots, \theta(N-1)]^T$ where the TVD $\theta(n)$ is introduced in (1).

The ML estimator of $\underline{\theta}$, denoted by $\hat{\underline{\theta}}$ is the vector that maximizes the likelihood function $L(\underline{x}_1, \underline{x}_2, \underline{s}, \underline{\theta}, \sigma)$

$$\hat{\underline{\theta}} = \arg\max_{\underline{\theta}} L(\underline{x}_1, \underline{x}_2, \underline{s}, \underline{\theta}, \sigma) \quad (7)$$

where $\underline{x}_1$, $\underline{x}_2$ and $\underline{s}$ are the vectors $\underline{x}_1 = [x_1(0), \ldots, x_1(N-1)]^T$, $\underline{x}_2 = [x_2(0), \ldots, x_2(N-1)]^T$ and $\underline{s} = [s(0), \ldots, s(N-1)]^T$, respectively. σ is the noise standard deviation (1). $\{.\}^T$ denotes the transpose operator.

Combining the assumptions of mutually independent and Gaussian noises ($w_1(n)$ and $w_2(n)$ of (1)), the likelihood function is written as

$$L\left(\underline{x}_1, \underline{x}_2, \underline{s}, \underline{\theta}, \sigma\right) = \frac{1}{(2\pi\sigma^2)^N} \exp\left(\frac{-1}{2\sigma^2} \sum_{n=0}^{N-1} \left([x_1(n)-s(n)]^2 + [x_2(n)-s(n-\theta(n))]^2\right)\right). \quad (8)$$

Since $s(n)$ is unknown, it can be estimated by maximizing the log-likelihood function with respect to $s(n)$:

$$\hat{s}\left(n-\theta(n)\right) = \tfrac{1}{2}\left[x_1\left(n-\theta(n)\right) + x_2(n)\right]. \quad (9)$$

By replacing the estimation of $s(n)$ (9) in (8), the maximization of the likelihood function corresponds to the minimization of the least squares function $\ell(\underline{\theta})$:

$$\ell(\underline{\theta}) = \sum_{n=0}^{N-1} (x_1(n-\theta(n)) - x_2(n))^2. \quad (10)$$

Second, by taking into account the approximation of $\theta(n)$ by $\theta^{(d)}(n)$ (5), we can write the least squares function as

$$\ell(\underline{\Theta}^{(d)}) = \sum_{n=0}^{N-1} (x_1(n-\theta^{(d)}(n)) - x_2(n))^2. \quad (11)$$

Thus the estimator $\hat{\underline{\Theta}}^{(d)}$ of the vector $\underline{\Theta}^{(d)}$ is given by

$$\hat{\underline{\Theta}}^{(d)} = \arg\min_{\underline{\Theta}^{(d)}} \ell(\underline{\Theta}^{(d)}). \quad (12)$$

3.3. Stochastic optimization

The minimization of $\ell(\underline{\Theta}^{(d)})$ should be numerically obtained using an optimal approach. This is a multidimensional continuous optimization problem with constraints due to CV physiological limits [21,23]. Since $\ell(\underline{\Theta}^{(d)})$ is a function with local minima classical optimization techniques such as the gradient descent algorithm do not guarantee the convergence to the global minimum and they converge only if the initial guess is already somewhat close to the final solution [26]. A stochastic optimization technique helps converge towards the global minimum. We propose to use the simulated annealing (SA) technique [27] to handle this problem. SA allows uphill moves, which potentially avoids being stuck at local minima. We propose the SA algorithm steps summarized in Table 1. We emphasize that the proposed SA algorithm is working with a random initialization step. This makes sure the algorithm is not to trapped in a local minimum.

To sum up, the originality of the TVD estimation addressed in the present paper is twofold. The first novel feature is the use of an orthonormal polynomial modeling of the TVD. The second is the use of the SA algorithm to solve the ML procedure. In this way, first we ensure a continuous time function model of the TVD that meets physiological constraints and limits on the CV. Second we reduce the dimension of the estimation problem. Lastly,

Table 1
SA algorithm for the minimization of the least squares function $\ell(\underline{\Theta}^{(d)})$ (11). The specification of the SA algorithm parameters is detailed in Appendix B.

1. Initialize $\underline{\Theta}^{(d)} = \underline{\Theta}^{(d)^{(0)}}$ and $\mathcal{T} = \mathcal{T}_0$
   $\underline{\Theta}^{(d)^{(0)}}$ is randomly chosen.
   $\mathcal{T}$ (Temperature) is a global parameter control.
2. Repeat
   - Repeat
     i. Generate a new candidate solution
     $\underline{\Theta}^{(d)^{(new)}} = \underline{\Theta}^{(d)} + \underline{\delta}$
     where $\underline{\delta}$ is an agitation vector;
     ii. Calculate the acceptance probability function
     $P(\underline{\Theta}^{(d)}, \underline{\Theta}^{(d)^{(new)}}, \mathcal{T})$;
     iii. If $P(\underline{\Theta}^{(d)}, \underline{\Theta}^{(d)^{(new)}}, \mathcal{T}) > rand()$
     then $\underline{\Theta}^{(d)} \leftarrow \underline{\Theta}^{(d)^{(new)}}$;
     until end of the temperature level.
   - If $\ell(\underline{\Theta}^{(d)}) < \ell(\underline{\Theta}^{(d)^{(best)}})$ then $\underline{\Theta}^{(d)^{(best)}} \leftarrow \underline{\Theta}^{(d)}$
     ($\underline{\Theta}^{(d)^{(best)}}$ is the best found solution).
     Gradually decrease the temperature $\mathcal{T}$.
   until the stopping criterion has been checked.
3. Return $\underline{\Theta}^{(d)^{(best)}}$.

we preserve optimality of the estimation by avoiding local minima and keeping the asymptotic property of the ML estimator. To ascertain the performance of the proposed method we derive the statistical CRB in the next section.

4. Cramer–Rao lower bound

In [19], we computed the appropriate CRB expressions for a canonical polynomial model of the TVD and for two sEMG recorded channels. The CRB were time-dependent functions and we emphasized their relationship to the classical CRB calculated for a constant time delay [19]. In the present paper, first we extend the CRB calculation to the Legendre polynomial modeling of the TVD and for two sEMG recorded channels. Second, we derive the appropriate CRB for any TVD waveform.

4.1. CRB for the TVD Legendre polynomial model

Let us derive the CRB for the parameters of the TVD Legendre polynomial model $\underline{\Theta}^{(d)}$ (6). Therefore, we have to determine the Fisher information matrix (FIM), denoted by F. It is a $(d+1)\times(d+1)$ matrix and its elements are as follows (see Appendix A for details):

$$F_{i,j} = \frac{1}{\sigma^2}\sum_{n=0}^{N-1}\left[(s'(n-\theta(n)))^2 P_i(n) P_j(n)\right] \tag{13}$$

where s' is the derivative of the signal s with respect to time and $P_i(n)$ is the Legendre polynomial function of degree i. If $\hat{\Theta}_i$ is an unbiased estimator of the parameter Θ_i then the CRB asserts

$$\mathrm{var}(\hat{\Theta}_i) \geq \mathrm{CRB}(\Theta_i) = [F^{-1}]_{i,i} \tag{14}$$

The computation of the FIM (1) is made possible through the availability of both the signal and the TVD law. However we should ensure that the FIM is well conditioned to obtain reliable bounds.

Next, we derive the CRB for the TVD Legendre polynomial model $\theta^{(d)}(n)$ (5). This CRB is a time-dependent function and it is obtained using the formula transformation described in [19,28]. If $\hat{\theta}^{(d)}(n)$ is an unbiased estimator of $\theta^{(d)}(n)$, then

$$\mathrm{var}(\hat{\theta}^{(d)}(n)) \geq \mathrm{CRB}(\theta^{(d)}(n)) = \underline{h}_n^T F \underline{h}_n \tag{15}$$

where $\underline{h}_n = [P_0(n), \ldots, P_d(n)]^T$ since $\partial\theta^{(d)}(n)/\partial\Theta_i = P_i(n)$.

We note that this derived CRB (15) is sensitive to the under/over estimation of the degree of the TVD polynomial model (as shown in [19] for a canonical TVD polynomial model). Indeed, the higher the number of model parameters to be estimated, the higher the mean square errors.

4.2. CRB for a general analytic TVD waveform

Let us assume a nonconstant TVD $\theta(n)$, then the CRB is written as (see calculation details in Appendix D):

$$\mathrm{CRB}(\theta(n)) = \sigma^2\left(\frac{\theta'(n)}{1-\theta'(n)}\right)^2 \frac{1}{s'(n-\theta(n))^2} \tag{16}$$

If $\hat{\theta}(n)$ is an unbiased estimator of $\theta(n)$, then

$$\mathrm{var}(\hat{\theta}(n)) \geq \mathrm{CRB}(\theta(n)) \tag{17}$$

We note that this derived CRB($\theta(n)$) (17) is also a time-dependent function. Its computation is made possible through the availability of both the derivative of the signal and the derivative of the TVD law with respect to time. As one can see, this CRB (17) tends to be 0 when the TVD tends to be a constant, which means a highly optimistic and noninformative bound. The denominator in Eq. (17) cannot vanish. Indeed, this corresponds to the TVD case $\theta(n) = n + \text{constant}$, which means that the assumed acquisition model (1) is meaningless.

5. Simulation results and discussion

In this section, the statistical study of the robustness of the proposed approach is addressed. First, the expressions of the statistical tools useful for this study are reviewed . Then, the Monte Carlo procedure useful for generating EMG signal trials is detailed. Next, the results are discussed and compared to the CRB.

5.1. Statistical tools

To ascertain the statistical performance of the approach proposed, the obtained results are expressed by means of classical normalized bias and mean square errors (MSE). The normalized bias is

$$\mathrm{Bias}\left(\theta^{(d)}(n)\right) = \frac{|E[\hat{\theta}^{(d)}(n)] - \theta^{(d)}(n)|}{\theta^{(d)}(n)} \tag{18}$$

where $\hat{\theta}^{(d)}(n)$ (5) is the estimator of $\theta^{(d)}(n)$. The symbol $E[.]$ denotes the expectation operator and it is practically calculated based on the Monte Carlo simulated data.

The MSE is defined by

$$\mathrm{MSE}(\theta^{(d)}(n)) = E[(\theta^{(d)}(n) - \hat{\theta}^{(d)}(n))^2]. \quad (19)$$

We also introduce two additional indicators in order to gain deeper insight into the impact of the TVD polynomial modeling. The first indicator is the deterministic mismatch error between the true TVD and the proposed polynomial model:

$$Er(n) = 100 \times \frac{|\theta(n) - \theta^{(d)}(n)|}{|\theta(n)|}. \quad (20)$$

The second indicator is the mean of the instantaneous root mean square error (RMSE) of $\theta(n)$ defined by

$$\mathrm{RMSE}_{TVD} = \frac{1}{N}\sum_{n=0}^{N-1} \sqrt{E[(\theta(n) - \hat{\theta}^{(d)}(n))^2]}. \quad (21)$$

We note that a similar indicator can be obtained for CV (n):

$$\mathrm{RMSE}_{CV} = \frac{1}{N}\sum_{n=0}^{N-1} \sqrt{E[(\mathrm{CV}(n) - \widehat{\mathrm{CV}}^{(d)}(n))^2]}, \quad (22)$$

where $\widehat{\mathrm{CV}}^{(d)}(n)$ is the estimation of $\mathrm{CV}(n)$ determined by substituting $\theta(n)$ to $\hat{\theta}^{(d)}(n)$ in (2).

5.2. Monte Carlo simulation

150 independent trials of synthetic sEMG signals were generated using the procedure described in Section 5.5 with fixed parameters $f_l = 60$ Hz, $f_h = 120$ Hz and $M = 30$. Two values of the signal to noise ratio (SNR), 10 dB and 20 dB, were considered. The sampling frequency was $F_s = 1024$ Hz and the inter-electrode distance $\Delta e = 10$ mm. Two time evolution models of CV introduced in [11,17] were considered, the unit of which is expressed in meter per second. The first CV model is a monotonic sigmoid function which presents fast and slow changes

$$CV_{sigmoid}(n) = 2 + \frac{1}{1 + e^{-40/6(n/F_s - 0.5)}}\ \mathrm{m\ s^{-1}}. \quad (23)$$

The second CV model is a sinusoidal function with positive and negative gradients

$$CV_{sinusoid}(n) = 4 + 2\ \sin\left(2\pi\frac{n}{F_s}\right)\quad \mathrm{m\ s^{-1}}. \quad (24)$$

Of course, the TVD $\theta(n)$ corresponding to each CV model is determined using (2). Both CV evolution models (23) and (24) satisfy CV physiological constraints when assuming a classical acquisition procedure [21,23]:

- the range of $\mathrm{CV}(n)$ values should fall within the interval $[2\ \mathrm{m\ s^{-1}}, 8\ \mathrm{m\ s^{-1}}]$,
- $\mathrm{CV}(n)$ derivative values should not exceed $2.5\ \mathrm{m\ s^{-2}}$.

Note that both CV models (23) and (24) are interesting to study the impact of the polynomial model degree on the TVD estimation accuracy. They do not have the same number of the sign changes in the TVD derivative. Since the polynomial model degree is an increasing function of the latter number, these two CV models need distinct TVD polynomial model degrees to reach comparable fitting errors.

In the following, we only consider $\theta^{(3)}(n)$ and $\theta^{(7)}(n)$ the approximations of $\theta(n)$ using two model orders $d=3$ and $d=7$ respectively (see (5)). For the sigmoidal CV model (23), the curves of $\theta(n)$, $\theta^{(3)}(n)$ and $\theta^{(7)}(n)$ are superimposed in Fig. 1(c) while their corresponding CV are superimposed in Fig. 1(a). Similarly, Fig. 1(b) and (d) shows the TVD and CV curves of the sinusoidal CV model (24).

The deterministic mismatch errors between the true TVD and proposed polynomial models of order 3 and 7 are superimposed in Fig. 1(e) and (f) for the sigmoidal and sinusoidal CV models respectively. As can be seen, a low order is actually enough to reach a TVD approximation with a systematic error less than ($< 12\%$) in the case of the sigmoidal CV model, whereas a higher order is required to reach the same systematic error in the case of the sinusoidal CV model. We also note that, as theoretically expected, these model errors decrease with increasing d.

5.3. Maximum likelihood estimation errors

This section concerns the study of the noise impact on the estimation procedure. Estimation results obtained by Monte Carlo simulations for the sigmoidal and sinusoidal CV models are shown in Figs. 2 and 3, respectively.

The biases of the ML estimator of $\theta^{(d)}$ are displayed in Figs. 2(a–b) and 3(a–b) for model orders of 3 and 7 respectively. Figs. 2(c–d) and 3(c–d) show the variances of the ML estimator of $\theta^{(d)}$ for both model orders, respectively.

According to Figs. 2(a–b) and 3(a–b), the bias can be neglected since it is smaller than 3.5%. Therefore, the ML estimator of $\theta^{(d)}(n)$ is almost unbiased. This makes it possible to compare the estimator variances with the appropriate derived CRB (14). From Figs. 2(c–d) and 3(c–d), one can first notice that the CRB is highly dependent on the dimension d and SNR value. A difference of at least 3 dB is observed between the cases $d=3$ and $d=7$, with $d=3$ exhibiting the lowest CRB. Second, both the CRB and estimator variance are high on the edges of the analysis window compared to the middle. This is a known problem in time-varying parameter estimation [19,28].

Finally, the estimator variance is approximately 3 dB and 5 dB higher than the appropriate CRB at SNR 20 dB and 10 dB, respectively. This difference can be further reduced by selecting a better stopping criterion. However the more selective the criterion, the longer the execution time will be.

Tables 2 and 3 summarize the RMSE_{TVD} (21) and RMSE_{CV} (22) indicator values. Since both indicators are defined as time-averaging errors, they help get a global performance comparison between the estimators. For both indicators, the best results are obtained whatever the SNR value with the model order $d=7$, the mismatch errors being lower in this case compared to the case $d=3$. Furthermore, when comparing the sinusoidal model results with sigmoidal ones, the indicator values are multiplied by at least 4 in the case $d=3$ for RMSE_{TVD}.

As a conclusion, the more the model order d is increased, the greater the reduction in systematic errors. However, increasing the model order means that the complexity of the optimization problem increases and hence convergence

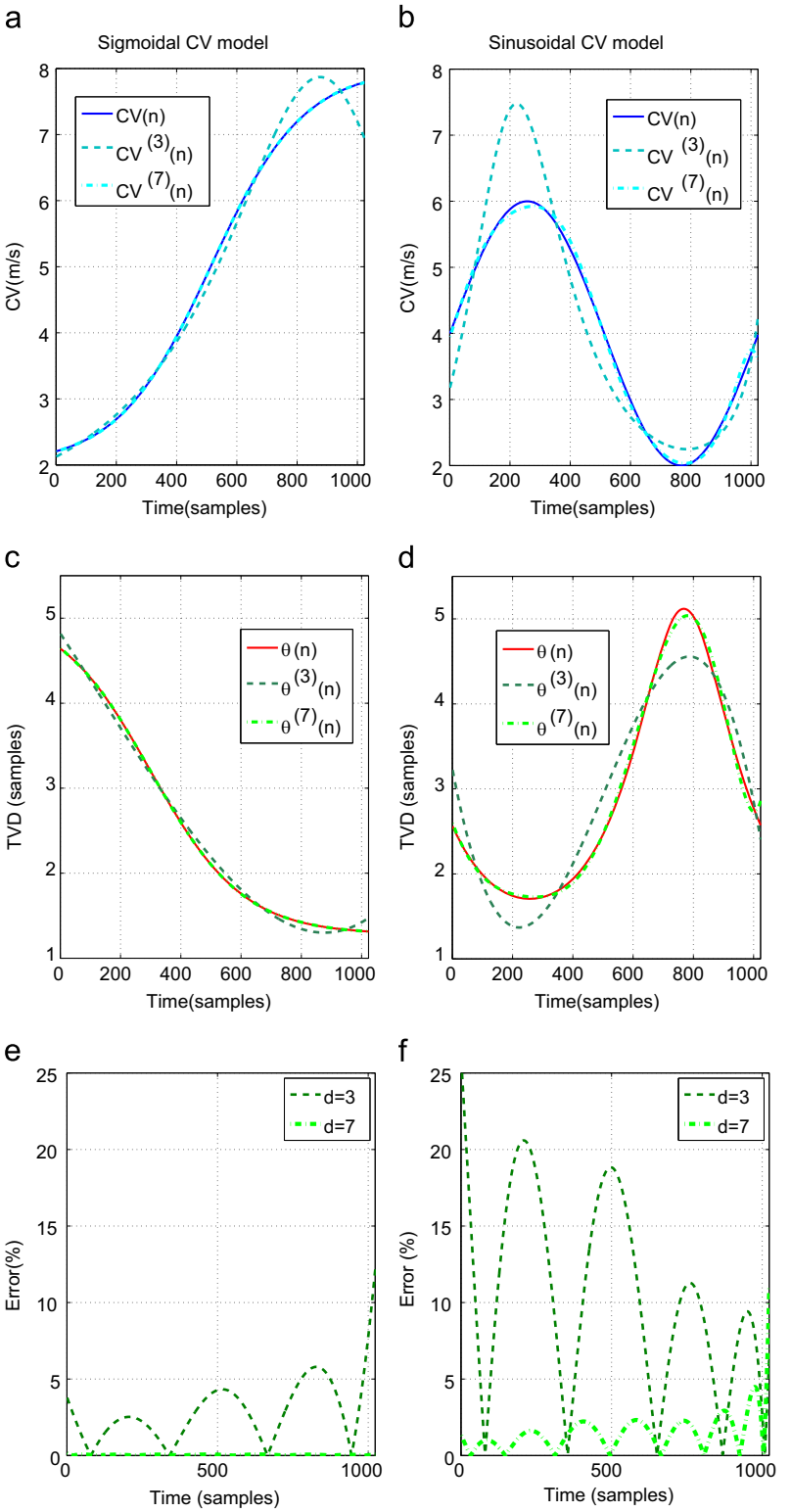

Fig. 1. Polynomial model degree impact on the TVD polynomial models. (a) and (b) Curves of CV(n) corresponding to the sigmoidal and sinusoidal CV models respectively. (c) and (d) Curves of the TVD $\theta(n)$ and TVD polynomial approximations $\theta^{(3)}(n)$ and $\theta^{(7)}(n)$ corresponding to the sigmoidal and sinusoidal CV models respectively. (e) Mismatch errors on the TVD for $d=3$ and $d=7$ corresponding to the sigmoidal CV model. (f) Mismatch errors on the TVD for $d=3$ and $d=7$ corresponding to the sinusoidal CV model. Note that the curves of the CV approximations CV $^{(3)}(n)$ and CV $^{(7)}(n)$ are obtained by substituting $\theta^{(3)}(n)$ and $\theta^{(7)}(n)$ for $\theta(n)$ in (2), respectively.

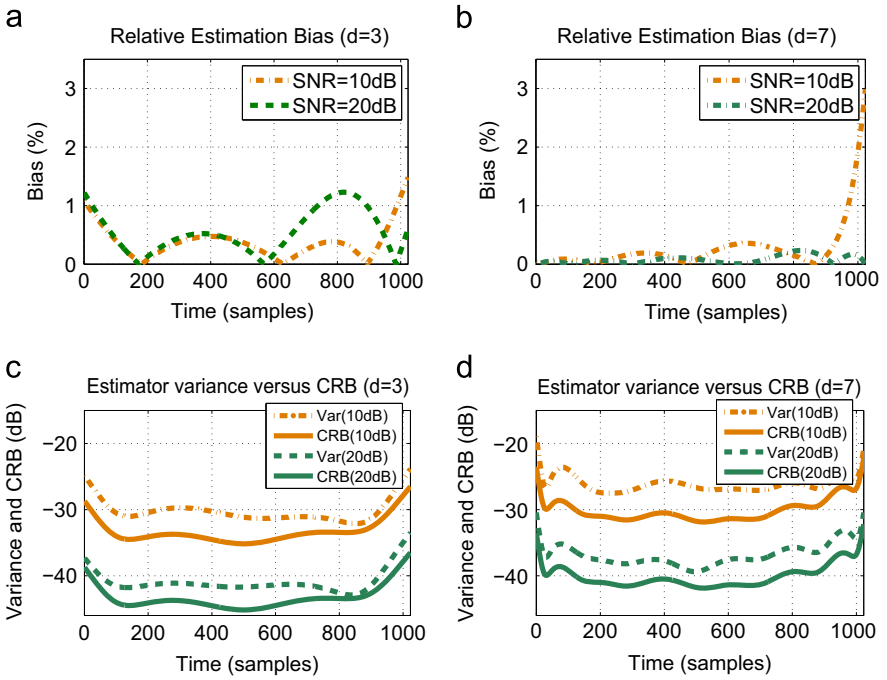

Fig. 2. Estimation results for the sigmoidal CV model at SNR = 10 dB and SNR = 20 dB: (a–b) Bias in percent of the ML estimator of $\theta^{(d)}(n)$ with (left column) $d=3$ and (right column) $d=7$. (c–d) Estimator variance and time-varying CRB (15) with (left column) $d=3$ and (right column) $d=7$.

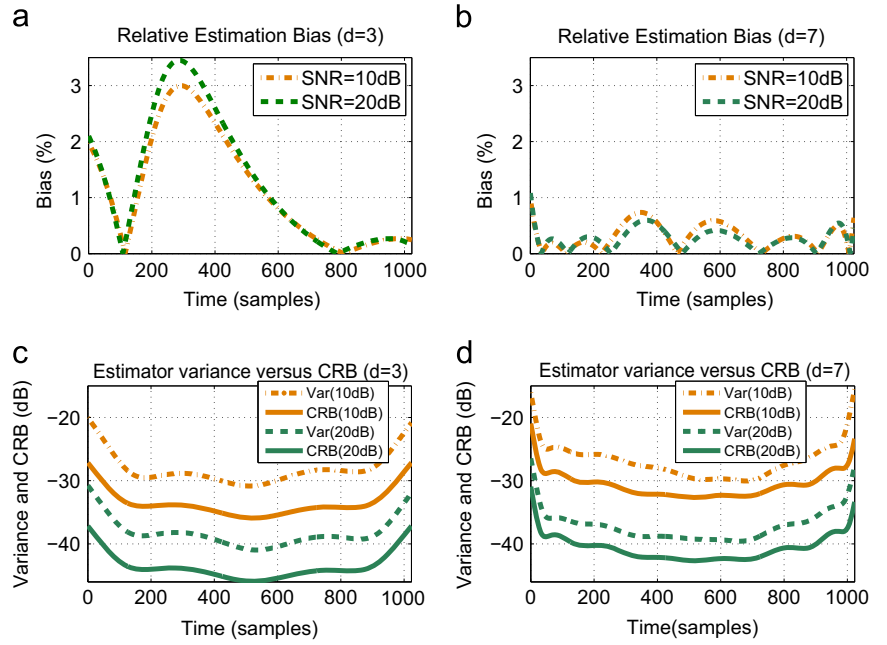

Fig. 3. Estimation results for the sinusoidal CV model at SNR = 10 dB and SNR = 20 dB: (a–b) Bias in percent of the ML estimator of $\theta^{(d)}(n)$ with (left column) $d=3$ and (right column) $d=7$. (c–d) Estimator variance and time-varying CRB (15) with $d=3$ (left column) and $d=7$ (right column).

problems may occur. Solving the problem therefore requires a longer execution time to reach a reasonable solution. It is necessary to make a compromise between the polynomial order and estimation errors.

Fig. 4(a) and (b) shows the obtained least squares errors $\ell(\hat{\underline{\boldsymbol{\Theta}}}^{(d)})$ (11) as a function of the model order for both the simulated sigmoidal and sinusoidal CV models at an SNR of 20 dB. As can be seen, the improvement is significant up to a given model order ($d=5$ and $d=7$ for sigmoidal and sinusoidal CV models, respectively). Above this level, however, no further improvement in execution time is significant.

5.4. Empirical strategy for selecting the model order

This section describes an empirical method to select a model order that satisfies the compromise between reasonable execution time and reduced estimation errors:

Step 0 : Estimate the TVD model parameters $\underline{\hat{\Theta}}^{(d)} = [\hat{\Theta}_0, \ldots, \hat{\Theta}_{d-1}]$ only once using a high model order, for example $d=10$. Then, evaluate the least squares error function $\ell(\underline{\hat{\Theta}}^{(d)})$ (11).

Step 1 : Remove the highest order estimated parameter $\hat{\Theta}_{d-1}$ from the vector $\underline{\hat{\Theta}}^{(d)}$. Decrease the value of d by 1. Reevaluate $\ell(\underline{\hat{\Theta}}^{(d)})$ (11).

Step 2 : Restart Step 1 until $d=1$.

The optimal model order is then selected when the value of $\ell(\underline{\hat{\Theta}}^{(d)})$ no longer evolves significantly. This selected model order can then be considered for data collected under the same protocol. This empirical method is entirely justified by the orthogonality property of the Legendre polynomial basis.

Table 2
RMSE on the TVD and CV estimation for the example of simulated sigmoidal CV function.

	SNR=10 dB		SNR=20 dB	
	$d=3$	$d=7$	$d=3$	$d=7$
$RMSE_{TVD}$ (samples)	0.072	0.050	0.064	0.015
$RMSE_{CV}$ (m/s)	0.202	0.140	0.189	0.044

Table 3
RMSE on the TVD and CV estimation for the example of simulated sinusoidal CV function.

	SNR=10 dB		SNR=20 dB	
	$d=3$	$d=7$	$d=3$	$d=7$
$RMSE_{TVD}$ (samples)	0.287	0.071	0.283	0.049
$RMSE_{CV}$ (m/s)	0.505	0.108	0.500	0.060

5.5. Comparison with a nonparametric method

To illustrate the potential of the proposed method, we present a comparison with the Fourier phase coherence estimator (FPCE) proposed in [17] and applied to sEMG signals. The method proposed is a parametric approach, the only required parameters being the polynomial model order and stopping criterion, whereas the FPCE method is a nonparametric method requiring prior parameters such as the shape and duration of the sliding window and the frequency range of interest. Fig. 5 shows the MSE obtained using these two methods for the sigmoidal CV model at SNR 20 dB. As can be seen, the lowest MSE is obtained using the proposed approach; a gain of 10 dB on average is observed.

6. Conclusions

The estimation of the time-varying delay between two-channel sEMG signals was addressed using a new parametric approach. This approach consists of a polynomial modeling of the TVD expressed on an orthonormal polynomial basis. The maximum likelihood estimation of the parameter model was carried out using the simulated annealing algorithm. The proposed approach provides unbiased TVD estimators with a MSE 3 dB higher than the CRB we derived. It also outperforms the FPCE method [17] by exhibiting a 10 dB performance improvement. This improvement is mainly due first to the parameter decoupling stemming from the basis orthogonality property and second to the ability of the SA algorithm to avoid local minima. In practice, better estimation accuracy can be obtained by increasing the TVD model order. However this increases the execution time needed to reach the expected performance, so a compromise should be made. We propose a manual method to select the optimal model order. In the future, we aim at taking into account the force level and fatigue impact on the sEMG PSD and addressing

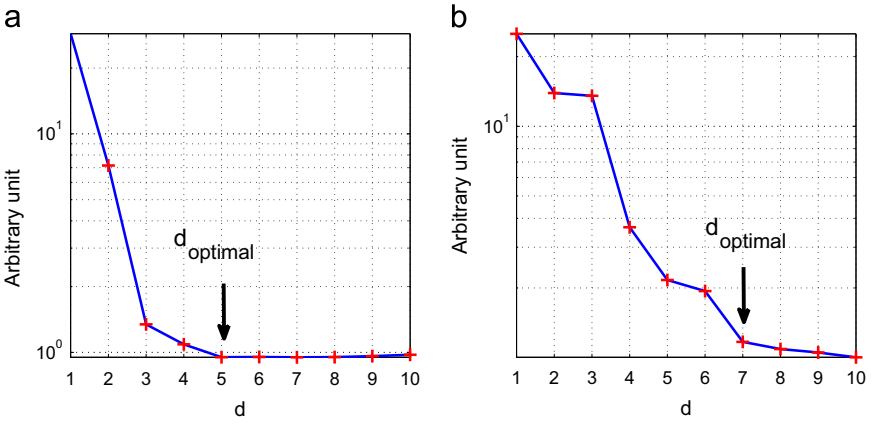

Fig. 4. Evolution of the least squares errors $\ell(\underline{\hat{\Theta}}^{(d)})$ (11) as a function of the model order for the (a) simulated sigmoidal CV model and (b) simulated sinusoidal CV model (SNR=20 dB).

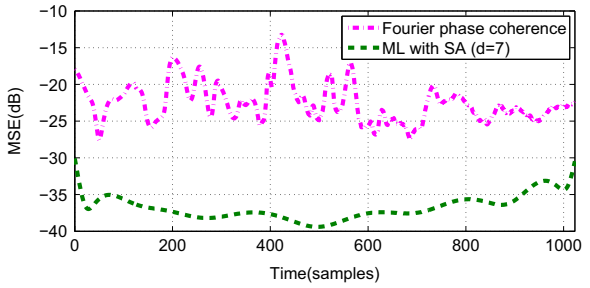

Fig. 5. Comparison between Fourier phase coherence estimator [17] (with a window duration of 0.25 s) and the proposed TVD estimator (with $d=7$) for the sigmoidal CV model at SNR 20 dB.

the robustness of the proposed approach. Better promising optimization techniques [29] will be investigated .

Appendix A

In Section 3.1, an orthonormal base is considered since the base orthogonality serves to decouple the model parameters. To verify this decoupling, we numerically calculate the normalized covariance matrix M of the model parameters with $d=3$ applied on the TVD appropriate to the sinusoidal CV (Fig. 1(b)). Each element of the matrix M is numerically calculated from the parameter estimates with 150 Monte Carlo simulations at an SNR of 20 dB

$$M_{i,j}=\frac{\mathrm{cov}(\hat{\Theta}_i,\hat{\Theta}_j)}{\mathrm{var}(\hat{\Theta}_0)} \quad i,j=0,\ldots,3. \tag{A.1}$$

The result obtained is as follows:

$$M=\begin{bmatrix} 1.0000 & -0.1320 & 0.0536 & -0.0380 \\ -0.1320 & 1.1019 & -0.0134 & 0.0342 \\ 0.0536 & -0.0134 & 0.9320 & -0.1491 \\ -0.0380 & 0.0342 & -0.1491 & 1.2970 \end{bmatrix}. \tag{A.2}$$

The diagonal elements are almost identical, so the model parameters have the same variance. The other elements that represent the coupling between the model parameters are almost zero.

Appendix B

In order to apply the simulated annealing method, one must specify the following parameters [27,30–34]:

1. The candidate generator procedure: A new candidate solution is selected in the neighborhood of the current solution.

$$\underline{\Theta}^{(d)(new)}=\underline{\Theta}^{(d)}+\underline{\delta} \tag{B.1}$$

where $\underline{\delta}$ is an agitation vector generated from Gaussian distribution $\mathcal{N}(0,\rho^2)$. The choice of the same variance ρ^2 for all model parameters is justified by the uncoupling property as verified in Appendix A.

2. The acceptance probability function: We use the Metropolis–Hasting criterion [35] defined by

$$P\left(\underline{\Theta}^{(d)},\underline{\Theta}^{(d)(new)},\mathcal{T}\right)=\exp\left(-\frac{\ell(\underline{\Theta}^{(d)(new)})-\ell(\underline{\Theta}^{(d)})}{\mathcal{T}}\right) \tag{B.2}$$

3. The initial temperature:

$$\mathcal{T}_0=\frac{\ell(\underline{\Theta}^{(d)(0)})}{A} \tag{B.3}$$

where A is a fixed parameter ($A=0.75$).

4. Cooling diagram: The temperature is reduced slowly by marking bearings of sufficient length.

$$\mathcal{T}_k=\mathcal{T}_{k-1}*C \tag{B.4}$$

with $0<C<1$, ($C=0.95$).

5. The stopping criterion:

The program will be stopped when there is no evolution of the current solution on D consecutive temperature levels ($D=10$).

Appendix C

The FIM elements are computed using the well-known formula given by [25]

$$F_{i,j}=-E\left[\frac{\partial^2 \ln(L(\underline{x}_1,\underline{x}_2,\underline{s},\underline{\Theta}^{(d)},\sigma))}{\partial\Theta_i\partial\Theta_j}\right] \tag{C.1}$$

By taking into account the approximation of $\theta(n)$ by $\theta^{(d)}(n)$ (5) in the likelihood function: (8), we have

$$\frac{\partial \ln(L(\underline{x}_1,\underline{x}_2,\underline{s},\underline{\Theta}^{(d)},\sigma))}{\partial\Theta_i}=\frac{1}{\sigma^2}\sum_{n=0}^{N-1}\left[\left(x_2(n)-s\left(n-\theta^{(d)}(n)\right)\right)\times\frac{\partial s(n-\theta^{(d)}(n))}{\partial\Theta_i}\right] \tag{C.2}$$

By making $f=n-\theta^{(d)}(n)$, we have

$$\frac{\partial s(n-\theta^{(d)}(n))}{\partial\Theta_i}=\frac{\partial s(f)}{\partial f}\frac{\partial f}{\partial\Theta_i} \tag{C.3}$$

From (5), it results: $\frac{\partial f}{\partial\Theta_i}=-P_i(n)$ and we have

$$\frac{\partial \ln(L(\underline{x}_1,\underline{x}_2,\underline{s},\underline{\Theta}^{(d)},\sigma))}{\partial\Theta_i}=\frac{-1}{\sigma^2}\sum_{n=0}^{N-1}\left[\left(x_2(n)-s\left(n-\theta^{(d)}(n)\right)\right)\times s'(n-\theta^{(d)}(n))P_i(n)\right] \tag{C.4}$$

Hence,

$$\frac{\partial^2 \ln(L(\underline{x}_1,\underline{x}_2,\underline{s},\underline{\Theta}^{(d)},\sigma))}{\partial\Theta_i\partial\Theta_j}=\frac{-1}{\sigma^2}\sum_{n=0}^{N-1}P_i(n)P_j(n)\left[(s'(n-\theta^{(d)}(n)))^2-(x_2(n)-s(n-\theta^{(d)}(n)))s''\times(n-\theta^{(d)}(n))\right] \tag{C.5}$$

From (1), we can approximate $E[x_2(n)-s(n-\theta^{(d)}(n))]=0$ and we have finally

$$F_{i,j}=\frac{1}{\sigma^2}\sum_{n=0}^{N-1}\left[(s'(n-\theta(n)))^2P_i(n)P_j(n)\right] \tag{C.6}$$

Appendix D

The Fisher information $F(\theta(k))$ is written using the classical formula [25]:

$$F(\theta(k))=-E\left[\frac{\partial^2\ln(L(\underline{x}_1,\underline{x}_2,\underline{s},\underline{\theta},\sigma))}{\partial\theta(k)^2}\right], \tag{D.1}$$

where k denotes the discrete time here for the sake of clarity of the proof. The first derivation of the log-likelihood function (8) with respect to $\theta(k)$ is expressed as

$$\begin{aligned}\frac{\partial\ln(L(\underline{x}_1,\underline{x}_2,\underline{s},\underline{\theta},\sigma))}{\partial\theta(k)}&=\frac{-1}{2\sigma^2}\sum_{n=0}^{N-1}\left[\frac{\partial(x_1(n)-s(n))^2}{\partial\theta(k)}\right.\\&\left.+\frac{\partial(x_2(n)-s(n-\theta(n)))^2}{\partial\theta(k)}\right],\\&=\frac{-1}{\sigma^2}\frac{\partial(x_2(k)-s(k-\theta(k)))}{\partial\theta(k)}(x_2(k)-s\\&\times(k-\theta(k))),\\&=\frac{1}{\sigma^2}\frac{\partial s(k-\theta(k))}{\partial\theta(k)}(x_2(k)-s(k-\theta(k))).\end{aligned}$$

The second derivation of the log-likelihood function (8) with respect to $\theta(k)$ is then given by

$$\begin{aligned}\frac{\partial^2\ln(L(\underline{x}_1,\underline{x}_2,\underline{s},\underline{\theta},\sigma))}{\partial\theta(k)^2}&=\frac{1}{\sigma^2}\left[\frac{\partial^2s(k-\theta(k))}{\partial\theta(k)^2}(x_2(k)-s(k-\theta(k)))\right.\\&\left.+\frac{\partial(s(k-\theta(k)))}{\partial\theta(k)}\frac{\partial(x_2(k)-s(k-\theta(k)))}{\partial\theta(k)}\right].\end{aligned}$$

From (1), we have $E[x_2(k)-s(k-\theta(k))]=0$ and therefore

$$E\left[\frac{\partial^2\ln(L(\underline{x}_1,\underline{x}_2,\underline{s},\underline{\theta},\sigma))}{\partial\theta(k)^2}\right]=\frac{-1}{\sigma^2}E\left[\left(\frac{\partial(s(k-\theta(k)))}{\partial\theta(k)}\right)^2\right]. \tag{D.2}$$

Under the assumptions of regularity and no abrupt changes in the TVD function (i.e $\partial\theta(k)/\partial k=\theta'(k)$ is finite), we can obtain for the case $\theta'(k)\neq 0$

$$\begin{aligned}\frac{\partial(s(k-\theta(k)))}{\partial\theta(k)}&=\frac{\partial(s(k-\theta(k)))}{\partial k}\frac{\partial k}{\partial\theta(k)},\\&=(1-\theta'(k))s'(k-\theta(k))\frac{\partial k}{\partial\theta(k)},\\&=\frac{(1-\theta'(k))}{\theta'(k)}s'(k-\theta(k)).\end{aligned} \tag{D.3}$$

The CRB of $\theta(k)$ obtained under the conditions cited above is then

$$\mathrm{CRB}(\theta(k))=F(\theta(k))^{-1}=\sigma^2\left(\frac{\theta'(k)}{(1-\theta'(k))}\right)^2\frac{1}{(s'(k-\theta(k)))^2}. \tag{D.4}$$

As one can notice from (D.4), the CRB of $\theta(k)$ has a finite expression ($\mathrm{CRB}(\theta(k))=0$) for the limiting case $\theta'(k)=0$. The term $(1-\theta'(k))$ in the denominator of (D.4) cannot vanish since the acquisition model (1) that we assume is meaningless if $\theta(k)=k+\text{constant}$.

References

[1] Y. Chan, J. Riley, J. Plant, A parameter estimation approach to time-delay estimation and signal detection, IEEE Trans. Acoust. Speech Signal Process. 28 (1) (1980) 8–16.
[2] G. Carter, Time delay estimation for passive sonar signal processing, IEEE Trans. Acoust. Speech Signal Process. 29 (3) (1981) 463–470.
[3] B. Harris, I. Gath, G. Rondouin, C. Feuerstein, On time delay estimation of epileptic EEG, IEEE Trans. Biomed. Eng. 41 (9) (1994) 820–829.
[4] D. Farina, R. Merletti, et al., Comparison of algorithms for estimation of EMG variables during voluntary isometric contractions, J. Electromyogr. Kinesiol. 10 (5) (2000) 337–349.
[5] M. Jakubowski, R.J. Fonck, C. Fenzi, G.R. McKee, Wavelet-based time-delay estimation for time-resolved turbulent flow analysis, Rev. Sci. Instrum. 72 (1) (2001) 996–999.
[6] B.C. Kirkwood, Acoustic Source Localization Using Time-Delay Estimation (Ph.D. thesis), Technical University of Denmark, 2003.
[7] T. Müller, M. Lauk, M. Reinhard, A. Hetzel, C. Lücking, J. Timmer, Estimation of delay times in biological systems, Ann. Biomed. Eng. 31 (2003) 1423–1439.
[8] D. Farina, R. Merletti, A novel approach for precise simulation of the EMG signal detected by surface electrodes, IEEE Trans. Biomed. Eng. 48 (6) (2001) 637–646.
[9] D. Farina, L. Arendt-Nielsen, R. Merletti, T. Graven-Nielsen, Assessment of single motor unit conduction velocity during sustained contractions of the tibialis anterior muscle with advanced spike triggered averaging, J. Neurosci. Methods 115 (1) (2002) 1–12.
[10] D. Farina, M. Pozzo, E. Merlo, A. Bottin, R. Merletti, Assessment of average muscle fiber conduction velocity from surface EMG signals during fatiguing dynamic contractions, IEEE Trans. Biomed. Eng. 51 (8) (2004) 1383–1393.
[11] F. Leclerc, P. Ravier, D. Farina, J.C. Jouanin, O. Buttelli, Time-varying delay estimation with application to electromyography, in: Proceedings of EUSIPCO, Lausanne, 2008.
[12] E.G.T. Liddell, C.S. Sherrington, Recruitment and some other factors of reflex inhibition, Proc. R. Soc. Lond. 97 (1925) 488–518.
[13] D. Farina, M. Gazzoni, R. Merletti, Assessment of low back muscle fatigue by surface EMG signal analysis: methodological aspects, J. Electromyogr. Kinesiol. 13 (4) (2003) 319–332.
[14] D. Farina, R. Merletti, R.M. Enoka, The extraction of neural strategies from the surface EMG, J. Appl. Physiol. 96 (2004) 1486–1495.
[15] D. Farina, A. Macaluso, R. Ferguson, G. De Vito, Effect of power, pedal rate, and force on average muscle fiber conduction velocity during cycling, J. Appl. Physiol. 97 (2004) 2035–2041.
[16] D. Farina, R. Merletti, Methods for estimating muscle fibre conduction velocity from surface electromyographic signals, Med. Biol. Eng. Comput. 42 (4) (2004) 432–445.
[17] F. Leclerc, P. Ravier, O. Buttelli, J. Jouanin, Comparison of three time varying delay estimators with application to electromyography, in: Proceedings of EUSIPCO, Poznan, 2007.
[18] Y. Chan, J. Riley, J. Plant, Modeling of time delay and its application to estimation of nonstationary delays, IEEE Trans. Acoust. Speech Signal Process. 29 (3) (1981) 577–581.
[19] M. Jabloun, P. Ravier, O. Buttelli, Cramer–Rao lower bounds for estimating the time-varying delay of surface EMG signals, in: IEEE Statistical Signal Processing Workshop (SSP), Nice, 2011.
[20] E. Shwedyk, R. Balasubramanian, R. Scott, A nonstationary model for the electromyogram, IEEE Trans. Biomed. Eng. (1977) 417–424.
[21] E. Schulte, O. Miltner, E. Junker, G. Rau, C.D. Klug, Upper trapezius muscle conduction velocity during fatigue in subjects with and without work related muscular disorders: a noninvasive high spatial resolution approach, Eur. J. Appl. Physiol. 96 (2006) 194–202.
[22] A. Pinkus, Weierstrass and approximation theory, J. Approx. Theory 107 (2000) 1–66.
[23] F. Leclerc, Développement d'outils non-stationnaires pour la mesure de délais variables appliqués aux signaux bioélectriques (Ph.D. thesis), Orleans University, 2008.
[24] M. Abramowitz, I. Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Dove, New York, 1972.
[25] S.M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory, vol. 1, Prentice Hall PTR, 1998.
[26] G.T. Luu, P. Ravier, O. Buttelli, Comparison of maximum likelihood and time frequency approaches for time varying delay estimation in

the case of electromyography signals, in: Biomedical Engineering International Conference BMEiCON, 2012.
[27] S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, Optimization by simulated annealing, Science 220 (1983) 671–680.
[28] B. Friedlander, J. Francos, Estimation of amplitude and phase parameters of multicomponent signals, IEEE Trans. Signal Process. 43 (4) (1995) 917–926.
[29] P. Indranil, D. Saptarshi, Design of hybrid regrouping PSO-GA based sub-optimal networked control system with random packet losses, Memet. Comput. 5 (2013) 141–153.
[30] L. Pibouleau, S. Domenech, A. Davin, C. Azzaro-Pantel, Expérimentations numériques sur les variantes et paramètres de la méthode du recuit simulé, Chem. Eng. J. 105 (2005) 117–130.
[31] P. Siarry, G. Dreyfus, La méthode du recuit simulé: théorie et applications, IDSET, 1988.
[32] P. Siarry, G. Berthiau, F. Durbin, J. Haussy, Enhanced simulated annealing for globally minimizing functions of many continuous variables, ACM Trans. Math. Softw. 23 (1997) 209–228.
[33] J. Dréo, A. Pétrowski, P. Siarry, E. Taillard, Metaheuristics for hard optimization: methods and case studies, SPI Publisher Services, Springer-Verlag, Berlin, Heidelberg, 2006.
[34] D. Henderson, S.H. Jacobson, A.W. Johnson, The theory and practice of simulated annealing, Handbook of Metaheuristics: International Series in Operations Research & Management Science 57 (2003) 287–319.
[35] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, E. Teller, Equation of state calculations by fast computing machines, J. Chem. Phys. 21 (1953) 1087.