Steady-State MSE Convergence of LMS Adaptive Filters with Deterministic Reference Inputs with Applications to Biomedical Signals

Salvador Olmos and Pablo Laguna

Abstract—In this paper, we analyze the steady-state mean square error (MSE) convergence of the LMS algorithm when deterministic functions are used as reference inputs. A particular adaptive linear combiner is presented where the reference inputs are any set of orthogonal basis functions—the adaptive orthogonal linear combiner (AOLC). Several authors have applied this structure always considering in the analysis a time-average behavior over one signal occurrence. In this paper, we make a more precise analysis using the deterministic nature of the reference inputs and their time-variant correlation matrix. Two different situations are considered in the analysis: orthogonal complete expansions and incomplete expansions. The steady-state misadjustment is calculated using two different procedures with equivalent results: the classical one (analyzing the transient behavior of the MSE) and as the residual noise at the output of the equivalent time-variant transfer function of the system. The latter procedure allows a very simple formalism being valid for colored noise as well. The derived expressions for steady-state misadjustment are contrasted with experimental results in electrocardiographic (ECG) signals, giving exact concordance for any value of the step size.

Index Terms—Biomedical signal, deterministic input, LMS adaptive filters, steady-state analysis.

I. INTRODUCTION

THE MOST well-studied bioelectrical signals are the event-related signals that are time locked to a stimulus. The stimulus can be external, as in visual, auditory, or electrical in the case of evoked potentials, or internal, as in electrocardiograms, (ECG's). For internal stimuli, a time-reference point can be defined from every signal occurrence, for example, the QRS fiducial point for ECG. The repetitive signals are often contaminated by noise from several sources. In general, an event-related signal can be considered to be a stochastic process that can be decomposed into a periodic deterministic signal that is time locked to a stimulus and an additive stationary noise uncorrelated with the signal. Several signal processing techniques are used to recover the signal hidden in the noise. The adaptive signal processing technique appears to be appropriate for such situations [1]–[3].

The LMS algorithm [4] is undoubtedly the most popular algorithm for adaptive signal processing. The popularity of the LMS algorithm is to a large extent due to its computational simplicity. Furthermore, it is generally felt that its behavior is quite simple to understand [4], [5], and the algorithm appears to be very robust.

The most common applications of the LMS algorithm (noise canceling, prediction, identification systems, etc.) use random reference inputs. As a consequence, the majority of authors have analyzed the properties of the LMS algorithm for random inputs. Several authors analyzed the MSE convergence of the LMS algorithm for Gaussian random inputs under the *independence assumption* [5]–[9]. This assumption, although clearly violated in many applications, simplifies the analysis significantly. The discrepancies between theoretical results based on this assumption and the true algorithm behavior was investigated in [10] and found to be relatively small. A more realistic assumption (less strong) has also been used by several authors [11], [12], where statistically dependent reference inputs are considered.

Much less work has been done with deterministic reference inputs. Some of the applications are related to adaptive noise cancellers of sinusoidal interferences [13]–[17], where a deterministic periodic waveform can be used because the disturbance period is known *a priori* or can be estimated from noise source measurements. The behavior of the LMS algorithm for sinusoidal references is slightly different than when the inputs are random and is denoted as a non-Wiener solution of the LMS algorithm [13], [14], [16]. In all these works, the structure of the adaptive filter was a transversal filter.

In the field of biomedical signals, several applications of the multiple-input adaptive linear combiner (ALC) [4] have been proposed, where several deterministic functions are used as reference inputs [18]–[23]. Very little accurate work has been addressed to the MSE convergence analysis of the LMS algorithm with deterministic reference inputs. Most of the authors normally use all the basis functions (the number of basis functions p is the same as the signal duration: N samples). However, many applications need a reduced number of coefficients (e.g., data compression [24], monitoring, detection, and analysis of pathologies like ischemia in ECG [25] and hypoxia in evoked potentials [18]). Two different situations will be considered in this paper: complete expansions $(p = N)$ and incomplete expansions $(p < N)$. In Section II, we introduce the *adaptive orthogonal linear combiner*, which is a generalization of previous applications, whose reference inputs are the basis functions of any orthogonal transform.

In addition, all the authors analyzed the convergence using a time-average over a signal occurrence. Recently, Barros *et*

al. [26] presented a simple MSE convergence analysis of the LMS algorithm when exponential functions were used as reference inputs. They applied the classical expressions of MSE convergence based on the independence assumption derived for random input reference signals [27]. In Section III, we show a more precise analysis of the MSE convergence considering the deterministic reference input signal nature and their time-variant correlation matrix.

In [19], [28], and [29], it was shown that the LMS algorithm with periodic impulsive reference inputs is equivalent to a linear time-invariant filter, whose transfer function is a comb filter. In addition, the misadjustment was interpreted in [29] as the residual noise that passes through the filter. In Section IV, we generalize the same result for any complete orthogonal transform (not only the identity transform formed by impulse functions) obtaining the same misadjustment result as in [29]. Moreover, we recently showed that when a reduced number of functions is used in the ALC, the adaptive filter is equivalent to a linear time-variant periodic filter [30], [31]. The misadjustment interpretation as the residual noise through the linear filter can also be applied for incomplete expansions obtaining the same results as with the classical time-domain analysis but in a more elegant and direct fashion.

Finally, simulation results with ECG signals from the QTDB database [32] corroborate that the derived equations for the steady-state *excess MSE* give exact results even for high values of the step-size μ.

II. ADAPTIVE ORTHOGONAL LINER COMBINER

The ALC [4] with the LMS algorithm has been applied to the analysis of ECG signals [3], [4], [19], [22], [33], evoked potentials [18], [34]–[36], and impedance cardiography signals [21]. It makes use of the recurring features of the signal. In this work, we denote the *adaptive orthogonal linear combiner* (AOLC) filter as a particular form of the ALC whose reference inputs are the basis functions of any orthogonal transform. Several authors have analyzed special cases of this structure using as reference inputs impulse functions [15], [19], [28], [29], Walsh functions [18], cosine functions [35], [36], exponential functions [20], [26], Hermite functions [22], and KLT functions [23], [25], [30]. In this paper, we generalize all these configurations to any orthogonal transform whose basis functions at time instant k are denoted as $\mathbf{X}[k] = (X_0[k]X_1[k]\cdots X_{p-1}[k])^T$, where $X_i[k]$ is the value of the ith basis function at time instant k, and p is the number of functions used in the modeling.

The structure of the adaptive filter is shown in Fig. 1. The primary input $d[k] = s[k] + n[k]$ consists of concatenated signal occurrences (composed of the deterministic part $s[k]$ and the noise part $n[k]$). The noise $n[k]$ is a wide-sense stationary stochastic process, whereas $s[k]$ is the biomedical signal under study after an N-sample segmentation defined around the stimulus instant. In the steady-state analysis of the algorithm, we assume that $s[k]$ is periodic. In practice, $s[k]$ will be time variant, and the algorithm will try to track the signal changes in a finite adaptation time. A first approximation analysis is to consider that the adaptive algorithm has infinite time to adapt its
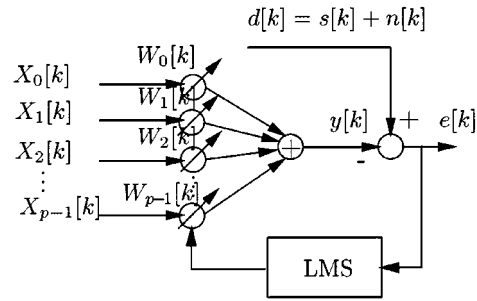

Fig. 1.   Adaptive linear combiner with orthogonal basis functions as reference inputs (AOLC).

weights. The adaptive system dynamically estimates the amount of each reference input $X_i[k]$ present in the input signal $d[k]$. For the analysis, we will consider that the basis functions $\mathbf{X}[k]$ are periodic, i.e., $\mathbf{X}[k + N] = \mathbf{X}[k]$ for all k. The number of basis functions $p \leq N$ will be variable. The filter output $y[k] = \mathbf{W}^T[k]\mathbf{X}[k]$ with $\mathbf{W}^T[k] = (W_0[k]W_1[k]\cdots W_{p-1}[k])$ recovers the deterministic part of $d[k]$ correlated with the reference inputs, whereas the uncorrelated noise is attenuated. The LMS algorithm tries to minimize the mean-square value of the error signal $e[k]$.

In the next section, we will show that when complete expansions are considered $p = N$, the weight vector converges to the optimum Wiener solution $\mathbf{W}_{\text{opt}}$, which is the projection of the deterministic signal $s[n]$ onto the space generated by $\{X_i[k]\}$. The estimation error at the optimum Wiener solution $e_{\text{opt}}[k]$ can be decomposed into two terms

$$
\begin{aligned}
e_{\text{opt}}[k] &= d[k] - \mathbf{W}_{\text{opt}}^T\mathbf{X}[k] \\
&= s[k] + n[k] - \mathbf{W}_{\text{opt}}^T\mathbf{X}[k] \\
&= c[k] + n[k].
\end{aligned} \tag{1}
$$

The first component $c[k] = s[k] - \mathbf{W}_{\text{opt}}^T\mathbf{X}[k]$ represents the estimation error due to the truncation of the orthogonal expansion. If we assume that the deterministic part of the input occurrence-concatenated signal $s[k]$ remains constant over all occurrences, then both $s[k]$ and $c[k]$ are periodic. The second term $n[k]$ is the noise present in the observed signal $d[k]$. Moreover, the components $c[k]$ and $n[k]$ are mutually independent. In the AOLC, the reference inputs are deterministic and statistically independent from the noise $n[k]$. No independence assumptions are needed in this case.

The MSE performance of the AOLC filter is analyzed here using two different ways with equivalent results. First, in Section III, we use the classical analysis, i.e., the transient analysis of the MSE for zero-mean white noise, and second, we evaluate the steady-state MSE as the residual noise at the output of the system using its equivalent transfer function in Section IV.

III. TRANSIENT ANALYSIS OF THE MSE

The solution to the finite difference equation of the LMS algorithm

$$
\mathbf{W}[k + 1] = \mathbf{W}[k] + 2\mu e[k]\mathbf{X}[k] \tag{2}
$$

is given by [14], [37]

$$\mathbf{W}[k] = \mathbf{F}_{k-1,0}\mathbf{W}[0] + 2\mu \sum_{m=0}^{k-1} d[m]\mathbf{F}_{k-1,m+1}\mathbf{X}[m] \quad (3)$$

where

$$\mathbf{F}_{j,m} = \begin{cases} \prod_{k=m}^{j} (\mathbf{I} - 2\mu\mathbf{X}[k]\mathbf{X}^T[k]), & j \geq m \\ \mathbf{I}, & j < m. \end{cases} \quad (4)$$

The first term $\mathbf{F}_{k-1,0}\mathbf{W}[0]$ is a transient, and it will be null at steady-state if low values of the step-size μ are used ($0 < \mu < 1$) because $\lim_{k\to\infty} \|\mathbf{F}_{k-1,0}\| = \mathbf{0}$. Alternatively, we may assume that $\mathbf{W}[0] = \mathbf{0}$.

The convergence analysis of the LMS algorithm for the AOLC is a bit simpler because both $\mathbf{X}[k]$ and $\mathbf{F}_{j,m}$ are deterministic as well as periodic. We consider two different situations in the analysis: complete expansions and incomplete expansions.

A. Complete Expansions

When complete orthonormal expansions are used, the basis function matrix

$$\mathbf{M} = (\mathbf{X}[0] \quad \mathbf{X}[1] \quad \cdots \quad \mathbf{X}[N-1]) \quad (5)$$

is square and unitary [38], i.e.,

$$\mathbf{MM}^T = \mathbf{I}_N = \mathbf{M}^T\mathbf{M}. \quad (6)$$

The first equality is equivalent to the orthogonality property of the basis functions over the time index $\sum_{k=0}^{N-1} X_j[k]X_m[k] = \delta_{jm}$, where δ_{jm} is the Kronecker delta function. The second equality implies a second kind of orthogonality involving different time-index vectors over the basis index, i.e., $\mathbf{X}^T[j]\mathbf{X}[m] = \sum_{i=0}^{N-1} X_i[j]X_i[m] = \delta_{jm}$. In this case, the time-variant transition matrix products $\mathbf{F}_{j,m}$ are greatly simplified, and it is easy to demonstrate that

$$\mathbf{F}_{j,m} = \mathbf{I} - 2\mu \sum_{k=m}^{j} \mathbf{X}[k]\mathbf{X}^T[k] \qquad j \geq m. \quad (7)$$

If we consider the transition matrix product of a complete signal occurrence, we have $\mathbf{F}_{N-1,0} = (1 - 2\mu)\mathbf{I}$, and the product $\mathbf{F}_{N-1,m+1}\mathbf{X}[m]$ reduces to $\mathbf{X}[m]$. Hence, the weight vector at time N from (3) reduces to

$$\mathbf{W}[N] = (1 - 2\mu)\mathbf{W}[0] + 2\mu \sum_{m=0}^{N-1} d[m]\mathbf{X}[m]. \quad (8)$$

We can iteratively apply (8) to an integer number i of signal occurrences giving

$$\mathbf{W}[iN] = (1 - 2\mu)^i\mathbf{W}[0]$$
$$+ 2\mu \sum_{l=0}^{i-1} (1 - 2\mu)^l \sum_{m=0}^{N-1} d[lN + m]\mathbf{X}[m]. \quad (9)$$

The first term is a transient if $0 < \mu < 1$ and assuming that the deterministic part $s[k]$ is periodic, the steady-state mean weight vector will be

$$E\{\lim_{i\to\infty} \mathbf{W}[iN]\} = \lim_{i\to\infty} \left\{ 2\mu \sum_{l=0}^{i-1} (1 - 2\mu)^l \right.$$
$$\left. \cdot \sum_{m=0}^{N-1} E\{d[lN + m]\}\mathbf{X}[m] \right\}$$
$$= \sum_{m=0}^{N-1} s[m]\mathbf{X}[m] = \mathbf{W}_{\text{opt}} \quad (10)$$

if zero-mean noise is assumed. Thus, the weight vector of the AOLC for complete expansions converges to the optimum Wiener solution, i.e., the clean signal projection onto the transformed domain defined by the basis functions $\{\mathbf{X}[k]\}$. Equivalently, the steady-state weight vector is an unbiased estimate of $\mathbf{W}_{\text{opt}}$. The weight error vector $\mathbf{V}[k] = \mathbf{W}[k] - \mathbf{W}_{\text{opt}}$ at any occurrence time instant j of the ith occurrence can be written using (3) as

$$\mathbf{V}[iN + j] = \mathbf{F}_{iN+j-1,j} \mathbf{V}[j]$$
$$+ 2\mu \sum_{m=j}^{iN+j-1} e_{\text{opt}}[m]\mathbf{F}_{iN+j-1,m+1}\mathbf{X}[m]$$
$$0 \leq j \leq N - 1. \quad (11)$$

In the complete expansion case, $\mathbf{V}[k]$ is due to the noise $n[k]$ because the truncation error $c[k]$ is null, and $e_{\text{opt}}[k] = n[k]$.

The minimum error can be calculated from (1) as $\xi_{\min} = E\{e_{\text{opt}}^2[k]\} = \sigma_n^2$. The total mean square error will be the sum $\xi[k] = E\{e^2[k]\} = \xi_{\min} + \xi^{\text{ex}}[k]$, where the *excess MSE* $\xi^{\text{ex}}[k]$ can be written [27] as

$$\xi^{\text{ex}}[k] = E\{\mathbf{X}^T[k]\mathbf{V}[k]\mathbf{V}^T[k]\mathbf{X}[k]\}$$
$$- 2E\{e_{\text{opt}}[k]\mathbf{X}^T[k]\mathbf{V}[k]\}. \quad (12)$$

The last term of (12) can be decomposed using (1), as

$$2E\{e_{\text{opt}}[k]\mathbf{X}^T[k]\mathbf{V}[k]\}$$
$$= 2c[k]\mathbf{X}^T[k]E\{\mathbf{V}[k]\} + 2\mathbf{X}^T[k]E\{n[k]\mathbf{V}[k]\}. \quad (13)$$

When complete expansions are considered, $c[k] = 0$. In addition, the term $E\{n[k]\mathbf{V}[k]\}$ is also null if zero-mean white noise is assumed. Hence

$$\xi^{\text{ex}}[k] = \mathbf{X}^T[k]E\{\mathbf{V}[k]\mathbf{V}^T[k]\}\mathbf{X}[k]. \quad (14)$$

The weight error correlation matrix $E\{\mathbf{V}[k]\mathbf{V}^T[k]\}$ at the end of the Nth occurrence can be calculated from (11) as

$$
\begin{aligned}
E\{\mathbf{V}&[iN]\mathbf{V}^T[iN]\}\\
&= (1-2\mu)^{2i} E\{\mathbf{V}[0]\mathbf{V}^T[0]\}\\
&\quad + 2\mu(1-2\mu)^i \sum_{l=0}^{i-1}(1-2\mu)^l\\
&\quad \cdot \sum_{m=0}^{N-1} \mathbf{X}[m] E\{n[lN+m]\mathbf{V}^T[0]\}\\
&\quad + 2\mu(1-2\mu)^i \sum_{l=0}^{i-1}(1-2\mu)^l\\
&\quad \cdot \sum_{m=0}^{N-1} E\{n[lN+m]\mathbf{V}[0]\}\mathbf{X}^T[m]\\
&\quad + 4\mu^2 \sum_{l_1=0}^{i-1}\sum_{l_2=0}^{i-1}(1-2\mu)^{l_1+l_2}\\
&\quad \cdot \sum_{m_1=0}^{N-1}\sum_{m_2=0}^{N-1} E\{n[l_1N+m_1]\\
&\qquad n\cdot[l_2N+m_2]\}\mathbf{X}[m_1]\mathbf{X}^T[m_2].
\end{aligned} \tag{15}
$$

The second and third terms are null if zero-mean white noise is assumed. Therefore, in that case

$$
\begin{aligned}
E\{\mathbf{V}&[iN]\mathbf{V}^T[iN]\}\\
&= (1-2\mu)^{2i} E\{\mathbf{V}[0]\mathbf{V}^T[0]\}\\
&\quad + 4\mu^2\sigma_n^2\mathbf{I}\sum_{l=0}^{i-1}(1-2\mu)^{2l}
\end{aligned} \tag{16}
$$

since $\sum_{m=0}^{N-1}\mathbf{X}[m]\mathbf{X}^T[m]=\mathbf{I}$. Finally, the steady-state excess MSE can be written using (14) and (16) as

$$
\lim_{i\to\infty}\xi^{\mathrm{ex}}[iN] = \frac{4\mu^2\sigma_n^2}{1-(1-2\mu)^2} = \xi_{\min}\frac{\mu}{1-\mu}. \tag{17}
$$

If we want to calculate $\xi^{\mathrm{ex}}[k]$ at time instants different from the end-occurrence time ($k=iN+j, j\neq 0$), we can apply an equivalent recursive relation to (16) for $j\neq 0$, obtaining the same steady-state excess MSE value as in (17). The normalized steady-state misadjustment is

$$
M[\infty] = \frac{\mu}{1-\mu}. \tag{18}
$$

The same result was obtained in [19], [29] for periodic impulse functions, where the simplicity of the basis functions allowed an easy estimation of the misadjustment. The expression obtained now is valid for any complete orthogonal transform. In contrast, the expression[1] in [26, eq. (34)] gives a steady-state misadjustment of

$$
M = \frac{N\mu\lambda}{1-\mu\lambda(N+1)} = \frac{\mu}{1-\mu-\mu/N} \tag{19}
$$

[1]The definition of the step-size μ in [26] was twice the value here.

which is very close to the exact result, especially for very low values of μ, as was the case in [26]. The difference arises from the fact that Barros *et al.* used a recursive equation for the weight error correlation matrix [26, eq. (7)] that was derived in [27] for random reference input signals by applying the Gaussian moment factoring theorem. However, the reference input signals considered in [26] were deterministic (exponential functions).

B. Incomplete Expansions

When a reduced number of basis functions is used in the AOLC $p < N$, we can conceptually analyze the MSE convergence in a similar way as in the last section. The only difference is that now, the analytical expressions are more complex. In addition, the truncation error $c[k]$ is no null.

When an incomplete set of basis functions is considered, the complete basis function matrix $\mathbf{M}$ can be partitioned as

$$
\mathbf{M} = \begin{bmatrix}\mathbf{M}_1\\\mathbf{M}_2\end{bmatrix} \tag{20}
$$

$\mathbf{M}_1$ being the $(p\times N)$ matrix formed by the selected basis functions and $\mathbf{M}_2$ the $((N-p)\times N)$ complementary matrix. Applying (20) to the first equality in (6), we obtain

$$
\begin{aligned}
\mathbf{M}_1\mathbf{M}_1^T &= \mathbf{I}_p, &\quad \mathbf{M}_2\mathbf{M}_1^T &= \mathbf{0}\\
\mathbf{M}_1\mathbf{M}_2^T &= \mathbf{0}, &\quad \mathbf{M}_2\mathbf{M}_2^T &= \mathbf{I}_{N-p}
\end{aligned} \tag{21}
$$

and from the second equality in (6)

$$
\mathbf{M}_1^T\mathbf{M}_1 + \mathbf{M}_2^T\mathbf{M}_2 = \mathbf{I}_N. \tag{22}
$$

From (21), we corroborate that the orthonormality over the time index is also true for a reduced number of orthogonal functions, i.e., $\mathbf{M}_1\mathbf{M}_1^T = \mathbf{I}_p$. In contrast, the orthonormality over the basis index is lost in general ($\mathbf{M}_1^T\mathbf{M}_1 \neq \mathbf{I}_N$) because at least one term of $\mathbf{M}_2^T\mathbf{M}_2$ is nonzero. In summary, when incomplete expansions are used $\mathbf{X}^T[j]\mathbf{X}[m] = r_{jm} \neq \delta_{jm}$ and the transition matrix products $\mathbf{F}_{j,m}$ in (4) have a more complex description. The steady-state weight error vector can be calculated from (11) by taking the limit as $i\to\infty$. The first term will converge to the null matrix for small values of μ because

$$
\lim_{i\to\infty}\mathbf{F}_{iN+j-1,j} = \lim_{i\to\infty}(\mathbf{F}_{N+j-1,j})^i = \mathbf{0}. \tag{23}
$$

The difference is that two different driving terms must be considered now because $e_{\mathrm{opt}}[k]$ has two components (1). Hence

$$
\begin{aligned}
\lim_{i\to\infty}&\mathbf{V}[iN+j]\\
&= 2\mu\sum_{l=0}^{i-1}(\mathbf{F}_{N+j-1,j})^l \sum_{m=j}^{N+j-1} c[m]\mathbf{F}_{N+j-1,m+1}\mathbf{X}[m]\\
&\quad + 2\mu\sum_{l=0}^{i-1}(\mathbf{F}_{N+j-1,j})^l\\
&\quad \cdot \sum_{m=j}^{N+j-1} n[lN+m]\mathbf{F}_{N+j-1,m+1}\mathbf{X}[m].
\end{aligned} \tag{24}
$$

Now, $\mathbf{V}[k]$ is originated by two different sources: the truncation error $c[k]$ and the noise $n[k]$. Applying the expected value, we obtain

$$
\begin{aligned}
E\{ &\lim_{i\to\infty} \mathbf{V}[iN+j]\} \\
&= 2\mu(\mathbf{I} - \mathbf{F}_{N+j-1,j})^{-1} \\
&\quad \cdot \sum_{m=j}^{N+j-1} c[m]\mathbf{F}_{N+j-1,m+1}\mathbf{X}_m \neq 0.
\end{aligned} \quad (25)
$$

The steady-state weight vector of incomplete expansions is a biased estimate of $\mathbf{W}_{\mathrm{opt}}$, and the bias is different at different occurrence-time instants j. The bias is originated by $c[k]$ and can be made small using a high number of basis functions or using a transform that packs the signal energy in a low number of basis functions. The bias depends on the step-size μ in a complex way because $\mathbf{F}_{i,j}$ also depends on μ.

The minimum error is now time-variant

$$
\begin{aligned}
\xi_{\min}[k] &= E\{e_{\mathrm{opt}}^2[k]\} \\
&= c^2[k] + \sigma_n^2.
\end{aligned} \quad (26)
$$

The excess MSE can be calculated using (12) and (13). The weight error correlation matrix can be calculated by multiplying (11) and applying the expected value

$$
\begin{aligned}
E\{\mathbf{V}[iN&+j]\mathbf{V}^T[iN+j]\} \\
&= \mathbf{F}_{iN+j-1,j}E\{\mathbf{V}[j]\mathbf{V}^T[j]\}\mathbf{F}_{iN+j-1,j}^T \\
&\quad + 2\mu\mathbf{F}_{iN+j-1,j}\sum_{m=j}^{iN+j-1}E\{\mathbf{V}[j]e_{\mathrm{opt}}[m]\} \\
&\quad \cdot \mathbf{X}^T[m]\mathbf{F}_{iN+j-1,m+1}^T \\
&\quad + 2\mu\sum_{m=j}^{iN+j-1}\mathbf{F}_{iN+j-1,m+1}\mathbf{X}[m] \\
&\quad \cdot E\{e_{\mathrm{opt}}[m]\mathbf{V}^T[j]\}\mathbf{F}_{iN+j-1,j}^T \\
&\quad + 4\mu^2\sum_{m=j}^{iN+j-1}\sum_{l=j}^{iN+j-1}E\{e_{\mathrm{opt}}[m]e_{\mathrm{opt}}[l]\} \\
&\quad \cdot \mathbf{F}_{iN+j-1,m+1}\mathbf{X}[m]\mathbf{X}^T[l]\mathbf{F}_{iN+j-1,l+1}^T.
\end{aligned} \quad (27)
$$

To calculate the steady-state value, we take the limit as $i \to \infty$. All the first three terms are transient (null at steady-state) if the step-size μ is selected to accomplish (23). As a consequence, the steady-state weight error correlation matrix can be written as

$$
\begin{aligned}
\lim_{i\to\infty} &E\{\mathbf{V}[iN+j]\mathbf{V}^T[iN+j]\} \\
&= 4\mu^2\sum_{m=j}^{iN+j-1}\sum_{l=j}^{iN+j-1}E\{e_{\mathrm{opt}}[m]e_{\mathrm{opt}}[l]\} \\
&\quad \cdot \mathbf{F}_{iN+j-1,m+1}\mathbf{X}[m]\mathbf{X}^T[l]\mathbf{F}_{iN+j-1,l+1}^T
\end{aligned} \quad (28)
$$

where

$$
\begin{aligned}
E\{e_{\mathrm{opt}}[m]e_{\mathrm{opt}}[l]\} &= E\{(c[m]+n[m])(c[l]+n[l])\} \\
&= c[m]c[l] + \sigma_n^2\delta[m-l]
\end{aligned} \quad (29)
$$

if $n[k]$ is assumed to be a zero-mean white noise with variance σ_n^2. In this case

$$
\begin{aligned}
\lim_{i\to\infty} &E\{\mathbf{V}[iN+j]\mathbf{V}^T[iN+j]\} \\
&= 4\mu^2\sum_{m=j}^{iN+j-1}\sum_{l=j}^{iN+j-1}c[m]c[l] \\
&\quad \cdot \mathbf{F}_{iN+j-1,m+1}\mathbf{X}[m]\mathbf{X}^T[l]\mathbf{F}_{iN+j-1,l+1}^T \\
&\quad + 4\mu^2\sigma_n^2\sum_{m=j}^{iN+j-1}\mathbf{F}_{iN+j-1,m+1}\mathbf{X}[m]\mathbf{X}^T[m] \\
&\quad \cdot \mathbf{F}_{iN+j-1,m+1}^T.
\end{aligned} \quad (30)
$$

The first term is due to the truncation error $c[k]$ (deterministic and periodic), and the second term is generated by the presence of noise $n[k]$. Applying the periodicity of the basis functions, (30) can be written as

$$
\begin{aligned}
\lim_{i\to\infty} &E\{\mathbf{V}[iN+j]\mathbf{V}^T[iN+j]\} \\
&= 4\mu^2\mathbf{B}_j\left(\sum_{m=j}^{N+j-1}\sum_{l=j}^{N+j-1}c[m]c[l] \right. \\
&\quad \left. \cdot \mathbf{F}_{N+j-1,m+1}\mathbf{X}[m]\mathbf{X}^T[l]\mathbf{F}_{N+j-1,l+1}^T\right)\mathbf{B}_j^T \\
&\quad + 4\mu^2\sigma_n^2\sum_{s=0}^{\infty}(\mathbf{F}_{N+j-1,j})^s\mathbf{Q}_j(\mathbf{F}_{N+j-1,j}^T)^s
\end{aligned} \quad (31)
$$

where

$$
\mathbf{B}_j = (\mathbf{I} - \mathbf{F}_{N+j-1,j})^{-1} \quad (32)
$$

and

$$
\mathbf{Q}_j = \sum_{m=j}^{N+j-1}\mathbf{F}_{N+j-1,m+1}\mathbf{X}[m]\mathbf{X}^T[m]\mathbf{F}_{N+j-1,m+1}^T. \quad (33)
$$

We do not have a closed form of the sum of the second term series in (31)

$$
\sum_{s=0}^{\infty}(\mathbf{F}_{N+j-1,j})^s\mathbf{Q}_j(\mathbf{F}_{N+j-1,j}^T)^s = \mathbf{H}_j \quad (34)
$$

but its convergence is guaranteed because of (23). A numerical approximation of the sum of the series $\mathbf{H}_j$ can be obtained by truncating the series to a finite number of additive terms. Finally, the steady-state weight error correlation matrix can be written as

$$
\lim_{i\to\infty} E\{\mathbf{V}[iN+j]\mathbf{V}^T[iN+j]\} = 4\mu^2\mathbf{D}_j + 4\mu^2\sigma_n^2\mathbf{H}_j \quad (35)
$$

where

$$\mathbf{D}_j = \mathbf{B}_j \left(\sum_{m=j}^{N+j-1} \sum_{l=j}^{N+j-1} c[m]c[l] \right.$$
$$\left. \cdot \mathbf{F}_{N+j-1,m+1}\mathbf{X}[m]\mathbf{X}^T[l]\mathbf{F}_{N+j-1,l+1}^T \right) \mathbf{B}_j^T. \quad (36)$$

The weight error correlation matrix is composed of two different terms: the first one due to the truncation error and the second term due to the presence of the noise.

Finally, the steady-state excess MSE at time instant $k = iN + j, i \rightarrow \infty$ can be calculated using (12), (13), (25), and (35) as

$$\lim_{i\rightarrow\infty} \xi^{\mathrm{ex}}[iN+j]$$
$$= 4\mu^2\mathbf{X}^T[j]\mathbf{D}_j\mathbf{X}[j] + 4\mu^2\sigma_n^2\mathbf{X}^T[j]\mathbf{H}_j\mathbf{X}[j]$$
$$- 4\mu c[j]\mathbf{X}^T[j]\mathbf{B}_j \sum_{m=j}^{N+j-1} c[m]\mathbf{F}_{N+j-1,m+1}\mathbf{X}[m]. \quad (37)$$

Remember that this expression is only valid for white noise. Furthermore, the steady-state excess MSE of incomplete expansions is different at different time indexes j. The last equation reduces to the complete expansions case (17) doing $c[k] = 0$ and considering complete basis functions $\mathbf{X}[k]$.

We can distinguish three different terms in (37): the first one $4\mu^2\mathbf{X}^T[j]\mathbf{D}_j\mathbf{X}[j]$ is the power of the biased estimation of $\mathbf{V}[k]$ originated by $c[k]$; the second term $4\mu^2\sigma_n^2\mathbf{X}^T[j]\mathbf{H}_j\mathbf{X}[j]$ is the variance introduced by the noise $n[k]$, and the third term is the cross-term of the truncation error $c[k]$ and $\mathbf{V}[k]$. A new term could be obtained from the interaction between $n[k]$ and $\mathbf{V}[k]$, but it was null in (37) because white noise was assumed in the analysis. There is no interaction between the truncation error $c[k]$ and noise $n[k]$ because they come from two different physical sources. The first two terms in (37) are related to the first term in (12) because $\mathbf{V}[k]$ has two independent sources: $c[k]$ and $n[k]$. The third and the "missing" terms in (37) are related to the second term in (12).

IV. Transfer Function Interpretation

A. Complete Expansions

Several authors have analyzed the ALC when periodic impulse functions are used as reference inputs (a complete expansion), showing its equivalence to a linear time-invariant filter [19], [28], [29], [39]. The impulse response $h[k]$ of the system is an impulse train with exponentially decreasing amplitude depending on the value of the step-size μ [29, eq. (21)]. The frequency response $H(e^{j\omega})$ is a comb filter whose lobe-width is proportional to μ

$$H(e^{j\omega}) = \frac{Y(e^{j\omega})}{D(e^{j\omega})} = \frac{2\mu e^{-j\omega N}}{1 + (2\mu - 1)e^{-j\omega N}}. \quad (38)$$

We have recently shown [30], [31] that any complete orthogonal transform (and not only the identity transform made of im-

pulse functions) obtains the same time-invariant transfer function (38).

In [29], the misadjustment was interpreted as the residual noise that passes through the equivalent transfer function. As a consequence, the steady-state excess MSE for complete expansions can be obtained as the integral value

$$\xi^{\mathrm{ex}}[\infty] = \frac{1}{2\pi} \int_{-\pi}^{\pi} |H(e^{j\omega})|^2 S_n(e^{j\omega}) \, d\omega \quad (39)$$

where $S_n(e^{j\omega})$ is the noise power spectral density. In the case of white noise with variance σ_n^2, the integral can be easily evaluated using Parseval's relation as

$$\xi^{\mathrm{ex}}[\infty] = \sigma_n^2 \sum_{k=-\infty}^{\infty} |h[k]|^2$$
$$= \sigma_n^2 4\mu^2 \sum_{l=0}^{\infty} (1 - 2\mu)^{2l} = \sigma_n^2 \frac{\mu}{1 - \mu}. \quad (40)$$

The excess MSE is the same as in (17). One major advantage of this misadjustment interpretation is that it can be easily calculated for other noise spectral density functions by evaluating the integral (39).

B. Incomplete Expansions

When a reduced number of basis functions are used in the AOLC, the system is equivalent to a linear time-variant periodic filter [30], [31]. Hence, the steady-state excess MSE will also be time variant, as it was shown in (37). The time-variant impulse response $h[n, k]$ is introduced in Appendix A.

If the primary input signal is decomposed as $d[k] = b[k] + c[k] + n[k]$, where

$b[k]$ deterministic component over the signal subspace spanned by the p basis functions;
$c[k]$ truncation error;
$n[k]$ observed noise;

the excess MSE can be calculated using (26) as

$$\xi^{\mathrm{ex}}[k] = \xi[k] - \xi_{\min}[k]$$
$$= E\{(b[k] + c[k] + n[k] - y[k])^2\}$$
$$- (c^2[k] + E\{n^2[k]\}). \quad (41)$$

The AOLC system is linear, and therefore, the output signal $y[k]$ can be decomposed into three terms $y[k] = y_b[k] + y_c[k] + y_n[k]$, where each term is the output corresponding to one of the input terms $b[k]$, $c[k]$, and $n[k]$, respectively. The steady-state output $y_b[k]$ in the absence of noise will be $y_b[k] = b[k]$ because the corresponding input signal $b[k]$ is completely represented by the p basis functions. Applying the zero-mean noise assumption to (41) and after some simple algebraic manipulations, we get

$$\xi^{\mathrm{ex}}[k] = y_c^2[k] + E\{y_n^2[k]\} - 2c[k]y_c[k] - 2E\{n[k]y_n[k]\} \quad (42)$$

where the cross-term products between $n[k]$ and the deterministic signals $b[k]$, $c[k]$, $y_b[k]$, and $y_c[k]$ are null. In the case of
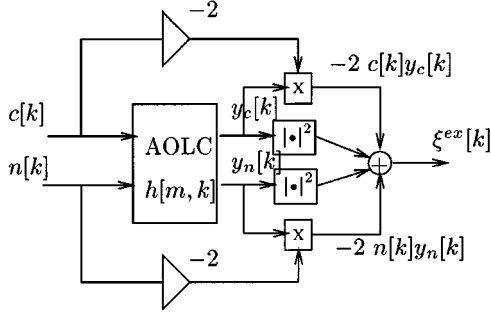
Fig. 2. Schematic block diagram of steady-state $\xi^{\text{ex}}[k]$ calculation according to the transform function interpretation (42).

white noise, the term $E\{n[k]y_n[k]\}$ is also null. Fig. 2 illustrates (42) with a block diagram. The inputs to the linear time-variant system AOLC are both error sources: the truncation error $c[k]$ and the noise $n[k]$. Four different outputs are obtained. The first term $y_c^2[k]$ is the power of the response to $c[k]$. This term is the same than $4\mu^2\mathbf{X}^T[j]\mathbf{D}_j\mathbf{X}[j]$ in (37). The second term $E\{y_n^2[k]\}$ is the noise variance at the output, which is the same as $4\mu^2\sigma_n^2\mathbf{X}^T[j]\mathbf{H}_j\mathbf{X}[j]$ in (37). The cross terms $2c[k]y_c[k]$ and $2E\{n[k]y_n[k]\}$ also have a equivalence term in (37). We can conclude that there is a direct relationship between each term of (37) and (42).

We need to calculate the output signal power of a time-variant periodic system. In Appendix A, we derive expressions for the response of a time-variant filter to deterministic and random input signals. The term $y_c^2[k]$ in (42) can be evaluated using (58) as

$$\lim_{i\to\infty} y_c^2[iN+j] = \left(\sum_{m=-\infty}^{\infty} c[m]f[m,j]\right)^2$$
$$0 \le j \le N-1 \tag{43}$$

where $f[m,k]$ is defined in Appendix A as the response at time k to the input signal $\delta[k-m]$. The second term $E\{y_n^2[k]\}$ can be easily calculated at steady-state using (64) in Appendix A as

$$\lim_{i\to\infty} E\{y_n^2[iN+j]\}$$
$$= \lim_{i\to\infty} r_{y_n}[iN+j, iN+j]$$
$$= \sum_{m=-\infty}^{\infty} r_n[m] \sum_{m_1=-\infty}^{\infty} h[m_1,j]h[m_1+m,j] \tag{44}$$

where $r_n[m]$ is the input noise autocorrelation function, and $h[m,j]$ is the impulse response of the AOLC system at time instant j of the occurrence. In the special case of a white noise random input signal $r_n[m] = \sigma_n^2\delta[m]$, the output noise energy is simplified to

$$\lim_{i\to\infty} r_{y_n}[iN+j, iN+j] = \sigma_n^2 \sum_{m_1=-\infty}^{\infty} h^2[m_1,j]. \tag{45}$$

Finally, the third term of $\xi^{\text{ex}}[k]$, $-2c[k]y_c[k]$ can be easily calculated as

$$\lim_{i\to\infty} -2c[iN+j]y_c[iN+j]$$
$$= -2c[j]y_c[j]$$
$$= -2c[j] \sum_{m=-\infty}^{\infty} c[m]f[m,j]. \tag{46}$$

Summing all three terms, the steady-state excess MSE is

$$\lim_{i\to\infty} \xi^{\text{ex}}[iN+j]$$
$$= \left(\sum_{m=-\infty}^{\infty} c[m]f[m,j]\right)^2$$
$$+ \sigma_n^2 \sum_{m=-\infty}^{\infty} h^2[m,j] - 2c[j] \sum_{m=-\infty}^{\infty} c[m]f[m,j]. \tag{47}$$

This expression can be more easily evaluated than (37) because the sequence $f[m,j]$ can be obtained running the filter with an input impulse at time instant j, and the instantaneous impulse response $h[m,j]$ is directly related to $f[m,j]$ by (55) of Appendix A.

The two excess MSE expressions (37) and (47), corresponding to the two different approaches, are term-by-term equivalent, even though their appearance is very different. When the truncation error $c[k]$ is small and very low values of step-size μ are used, we demonstrate in Appendix B that both equations can be greatly simplified, showing analytically its equivalence for this particular case.

C. Excess MSE for Colored Noise

Biomedical signals are always embedded in physiological noise generated by contiguous physiological systems. For example, muscle electrical activity, motion artifacts, and baseline wandering are often also recorded in ECG signals. Therefore, nonwhite noise should also be considered in the convergence analysis of the AOLC. Most random processes with a continuous power spectrum density can be generated as the output of a causal linear filter driven by white noise [40]. This white noise-driven model is called the *innovations representation* of the random process.

1) Complete Expansions: Let $g[k]$ be the impulse response of the linear filter of the innovations representation of the colored noise $n[k]$, and let $w[k]$ the white input noise. The impulse response can be normalized in order to get a unity energy filter $\sum_{k=0}^{\infty} |g[k]|^2 = 1$; therefore, noise power information will be in the white input noise variance σ_w^2. The excess MSE can be calculated using (39) as

$$\xi^{\text{ex}}[\infty] = \sigma_w^2 \frac{1}{2\pi} \int_{-\pi}^{\pi} |H(e^{j\omega})|^2 |G(e^{j\omega})|^2 \, d\omega \tag{48}$$

which can be also easily evaluated using the Parseval's relation as

$$\xi^{\text{ex}}[\infty] = \sigma_w^2 \sum_{k=-\infty}^{\infty} |h[k] * g[k]|^2. \tag{49}$$

The convolution $y[k] = h[k] * g[k]$ can be written as

$$y[k] = 2\mu \sum_{i=1}^{\infty} (1 - 2\mu)^{i-1} g[k - iN]. \tag{50}$$

For many biomedical signals, such as evoked potentials or ECG signals, we can assume that the noise autocorrelation function is shorter than N, which is the occurrence length. From (49) and (50) and assuming that the length of $g[k]$ is shorter than N, we get

$$\xi^{\text{ex}}[\infty] \simeq \sigma_w^2 (2\mu)^2 \sum_{k=0}^{\infty} \sum_{i=1}^{\infty} (1 - 2\mu)^{2(i-1)} g^2[k - iN]$$

$$\simeq 4\mu^2 \sigma_w^2 \sum_{i=1}^{\infty} (1 - 2\mu)^{2(i-1)} \sum_{k=0}^{N-1} g^2[k - iN]$$

$$\simeq \frac{\mu}{1 - \mu} \sigma_w^2 = \frac{\mu}{1 - \mu} \xi_{\min}. \tag{51}$$

The excess MSE for complete expansions with colored input noise is the same than for white noise.

2) Incomplete Expansions: When a reduced number of functions are used in the AOLC, we have shown that the excess MSE can be easily evaluated using (42). The last term $2E\{n[k]y_n[k]\}$ is due to the interaction between the present noise $n[k]$ and past noise samples and is evaluated as

$$2E\{n[k]y_n[k]\}$$

$$= 2E\left\{n[k] \sum_{m=-\infty}^{\infty} h[m, k]n[k - m]\right\}$$

$$= 2 \sum_{m=-\infty}^{\infty} h[m, k]E\{n[k]n[k - m]\}$$

$$= 2 \sum_{m=-\infty}^{\infty} h[m, k]r_n[m]. \tag{52}$$

Finally, evaluating the steady-state excess MSE at time instant $k = iN + j$ with $i \to \infty$, we get

$$\lim_{i \to \infty} \xi^{\text{ex}}[iN + j]$$

$$= \left(\sum_{m=-\infty}^{\infty} c[m]f[m, j]\right)^2 + \sum_{m=-\infty}^{\infty} r_n[m]$$

$$\cdot \sum_{m_1=-\infty}^{\infty} h[m_1, j]h[m + m_1, j]$$

$$- 2c[j] \sum_{m=-\infty}^{\infty} c[m]f[m, j] - 2 \sum_{m=-\infty}^{\infty} h[m, j]r_n[m]. \tag{53}$$
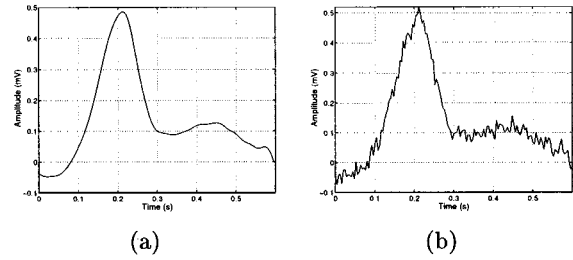

Fig. 3. Original and simulated noisy observed signal occurrence with SNR = 20 dB. (a) Clean signal $s[k]$. (b) Observed noisy signal $d[k]$.
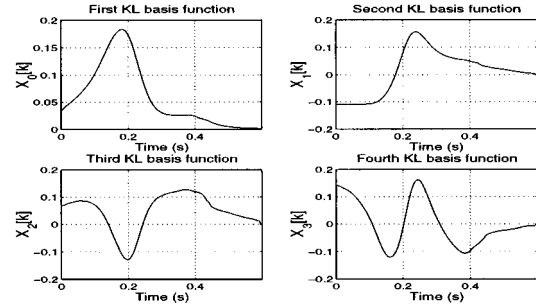

Fig. 4. First four KL basis functions of the ST-T complex.

In order to get an equivalent result using the time-domain transient analysis of $\xi^{\text{ex}}[k]$, more complex equations should be analyzed. When white noise is considered, $r_n[m] = \sigma_n^2 \delta[m]$, (53) is equivalent to (47) because the impulse responses accomplish $h[m, j] = 0; m \leq 0$.

V. RESULTS

In this section, we experimentally verify the validity of the derived equations so far. As a first step, we made a simulation to evaluate the steady-state performance for stationary signals. A signal was synthesized as a sequence of records $d[k]$. Each one consisted of a selected invariant ST-T complex of one normal heartbeat from record 103 of the QTDB database [32] and additive Gaussian white noise, $n[k]$, with a value of SNR = 20 dB. We show in Fig. 3 the selected clean heartbeat and a signal occurrence when the simulated noise is added.

We selected the optimal Karhunen–Loéve transform [40] as an example of one commonly used orthogonal transform, but any orthogonal transform could be used. The basis functions were estimated using a training set of signals from several databases [25] in order to adapt the basis functions to a large population of ECG morphologies. These basis functions are optimum in the sense that they represent the highest percentage energy of the training set ensemble with the minimum number of functions [40]. We show in Fig. 4 the first four KL basis functions for the ST-T complex of the ECG.

The AOLC filter was applied to the simulated signal, and the results of steady-state excess MSE were compared with derived equations in previous sections. As a first step, we study the MSE convergence of incomplete expansions without simulated noise. In this case, the weight error vector trajectory $\mathbf{V}[k]$ is completely deterministic, and we will expect exact results from the analysis, setting $n[k] = 0$. In this simplified case, the error
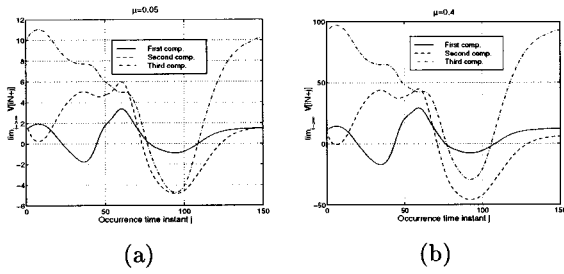
Fig. 5. Theoretical and experimental steady-state weight error vector $\lim_{i \to \infty} \mathbf{V}[iN + j]$ for two different values of $\mu = 0.05$ and 0.4 when $p = 3$ basis functions are used. (a) $\mu = 0.05$. (b) $\mu = 0.4$.

Fig. 6. Impact of the biased estimation of the steady-state weight vector on the reconstructions. (a) $\mu = 0.05$. (b) $\mu = 0.4$.

signal will be $e[k] = c[k] - \mathbf{V}^T[k]\mathbf{X}[k]$. Thus, the MSE can be written as

$$
\begin{aligned}
\xi[k] &= (c[k] - \mathbf{V}^T[k]\mathbf{X}[k])^2 \\
&= c^2[k] + \mathbf{V}^T[k]\mathbf{X}[k]\mathbf{X}^T[k]\mathbf{V}[k] - 2c[k]\mathbf{V}^T[k]\mathbf{X}[k].
\end{aligned}
\tag{54}
$$

In this case, the steady-state weight error vector is only due to the truncation error $c[k]$ because there is no noise in the simulated signal. These are the best conditions for the AOLC to estimate the signal $s[k]$. However, the steady-state weight vector will be biased with respect to the optimum weight vector. We show in Fig. 5 the values of the steady-state weight error vector with two different values of the step-size $\mu = 0.05$ and 0.4 when only $p = 3$ basis functions are used in the AOLC. Theoretical results are calculated using (25). Experimental results are obtained by running the AOLC filter. Results of the the weight vector are shown after $i = 100$ signal occurrences where all transients have died. Experimental and theoretical results are completely overlaid. It is clearly seen that all the $p = 3$ components of the weight vector are a biased estimate of the optimum weight vector, and the steady-state bias is different for every occurrence instant j. Moreover, the bias is larger for higher values of μ following an approximated linear relation with μ, as predicted in (25).

In order to illustrate the impact of the biased estimation of the steady-state weight vector on the reconstructed signal, we show in Fig. 6 the output signal $y[iN + j] = \mathbf{W}^T[iN + j]\mathbf{X}[j]$ using $p = 3$ basis functions after all transients have died. The difference between the reconstructed signals obtained with the biased weight vector and the optimum weight vector are almost invisible, especially for low values of the step-size μ. The impact of the biased estimation of the steady-state weight vector on the reconstructed signal is very low in terms of signal deformation.

The steady-state excess MSE for $p = 3$ basis functions are shown in Fig. 7 for two different values of the step-size ($\mu = 0.05$ and 0.4). We show the theoretical values of $\xi^{\text{ex}}[k]$ [the last two terms in (54)] and experimental values for two different values of μ. Experimental results are shown after $i = 100$ occurrences. Again, both results (theoretical and experimental) are completely overprinted. Moreover, theoretical results are also calculated as the sum of the first and third terms of both (37) and (42), showing their equivalence. It is corroborated that the steady-state excess MSE is different for every time instant j of the occurrence with higher values when the step-size is larger, although no noise is present in the input signal. Moreover, the
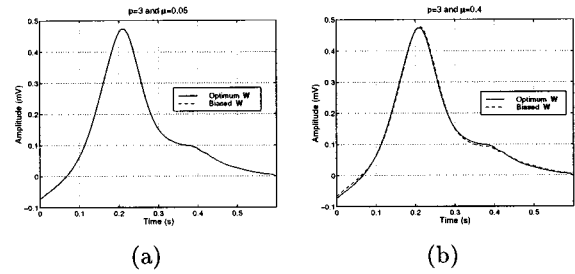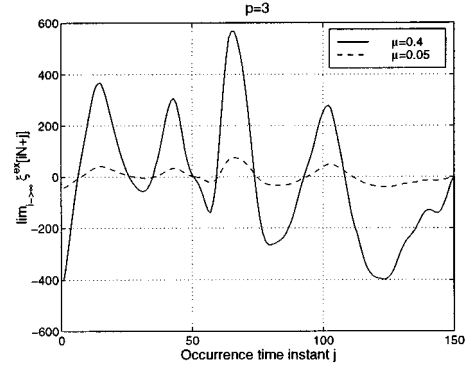

Fig. 7. Theoretical and experimental values of the steady-state $\xi^{\text{ex}}[k]$ with $p = 3$ KL basis functions without noise.

steady-state excess MSE can be negative because of the truncation error.

The second step of verification is to consider the presence of noise in the observed signals. In Fig. 8, we show the steady-state excess MSE using derived expressions (37) and (47) and experimental measures averaging 10 000 runs of the filter with simulated noisy signals of SNR $= 20$ dB after $i = 100$ occurrences. We show three different values of the number of functions p. The mean values of the experimental results of steady-state excess MSE are overprinted on the theoretical values for both cases: small and large values of the step-size μ. The value of steady-state $\xi^{\text{ex}}[k]$ decreases for high values of the number p of basis functions and for low values of the step-size μ. When p is low, e.g., $p = 3$, the truncation error $c[k]$ is much more important than noise, and Fig. 8(a) is very similar to Fig. 7. When p is low, the steady-state $\xi^{\text{ex}}[k]$ can be negative, as can be seen from Fig. 8(a) and (b). However, the total MSE $\xi[k] = \xi_{\min}[k] + \xi^{ex}[k]$ is always a semidefinite positive quantity. When p is high, the truncation error $c[k]$ is very small, and $\xi^{\text{ex}}[k]$ is positive [see Fig. 8(c)].

If we want to reduce the value of the total output signal error $\xi[k]$, we can use a higher value of the number p of functions and/or select a lower value of the step-size μ. A question that arises now is which of both actions will be more efficient in order to reduce the total error $\xi[k]$. To answer this equation, we can use the three-term decomposition of the steady-state $\xi^{\text{ex}}[k]$ in (37) and (47) and see which term is more important for a given condition of the input signal, the SNR of the contaminating noise, etc. Moreover, $\xi_{\min}[k] = c^2[k] + E\{n^2[k]\}$ only depends on the number of functions p. When the truncation error is more important than the noise, the number of functions should
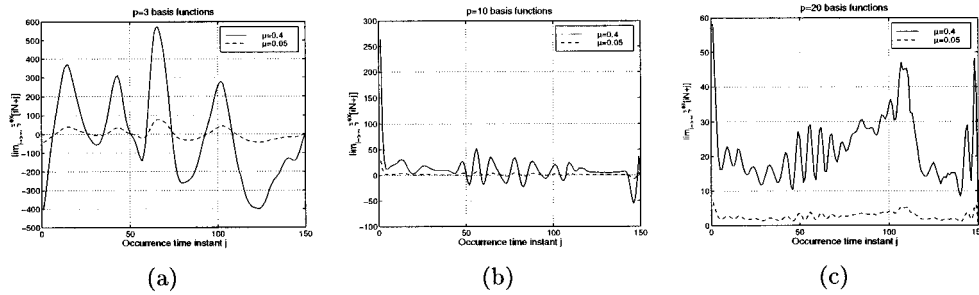
Fig. 8.   Comparison of theoretical and experimental values of steady-state ξ^{ex} using $p = 3$, 10, and 20 KL basis functions with SNR $= 20$ dB. Note that vertical scales are different. (a) $p = 3$. (b) $p = 10$. (c) $p = 20$.
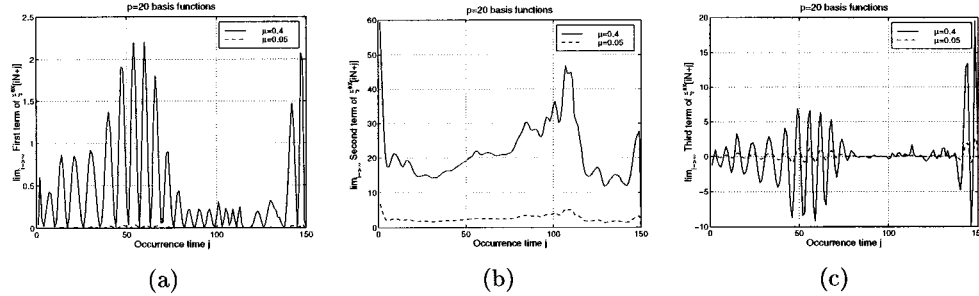

Fig. 9.   Decomposition of steady-state ξ^{ex} using $p = 20$ KL basis functions with SNR $= 20$ dB. Note that vertical scales are different. (a) First term. (b) Second term. (c) Third term.

be preferably increased. For example, from the conditions $p = 3$ and $\mu = 0.4$ in Fig. 8(a), it is more efficient to increase the number of functions to $p = 10$ with $\mu = 0.4$ than to decrease the step size to $\mu = 0.05$ with $p = 3$ (without evaluating the reduction in ξ_{min} due to the increasing number of basis functions). However, when the truncation error is small compared with the noise, a decrease of the step size is the more appropriate choice. For example, we show in Fig. 9 the decomposition of $\xi^{\mathrm{ex}}[k]$ into the three terms for $p = 20$ basis functions with low and high values of μ and SNR $= 20$ dB. The noise term shown in Fig. 9(b) is more important than truncation error, and therefore, a decrease of the step size is more efficient.

For clinical applications, we can be interested in specific areas of the repetitive signal that are located at equal time instants j inside the occurrence, e.g., ST elevation for ischemia detection, QRS amplitude, etc. With our analysis, we can evaluate the steady-state MSE for every time of the signal occurrence instead of a mean value, as had been analyzed in previous works.

VI. CONCLUSIONS

In this paper, we analyzed the steady-state MSE convergence of the LMS algorithm using the adaptive orthogonal linear combiner (AOLC), where the reference inputs were any set of orthogonal functions. The deterministic and periodic properties of the reference inputs allowed an exact steady-state analysis of the LMS algorithm. The primary input was a deterministic and periodic signal contaminated by stationary noise.

Two alternative formulations of the problem were used: First, we used a time-domain formulation based on the solution of the discrete-time recursive equation for the evolution of the weight vector. The second formulation is based on a transfer domain

approach, where the misadjustment could be interpreted as the residual noise power that passes through the equivalent transfer function of the system. This interpretation allowed an easy calculation of the excess MSE for the case of colored input noise.

In addition, the analysis was performed in two different situations: complete expansions and incomplete expansions. The steady-state misadjustment expressions for complete expansions were in concordance with exact results previously obtained when periodic impulses were used as reference inputs [19], [29]. The same result of steady-state misadjustment is now generalized to any complete orthogonal transform. Some important differences are obtained when only a reduced number of functions $p < N$ are used in the expansion. The first one is that the weight vector converges to a biased estimate of the optimum Wiener solution. The bias is due to the truncation error. Moreover, the value of steady-state misadjustment is different for every occurrence time instant.

The decomposition of the steady-state excess MSE into three different terms (for the case of white noise) gives a useful criteria for selecting the more appropriate parameters (number of basis functions p and the step-size μ) that define the AOLC system. When the first term is higher than the others, the number of basis functions should be increased because the truncation error is more important than the noise present in the signal. On the contrary, if the second term is higher, it means that lower values of the step-size μ should be used in order to reduce the amount of noise.

Experimental results with electrocardiographic signals show that derived expressions give exact results of steady-state excess MSE for any value of the step-size μ. Many previous published results were close to the exact solution given here because they only considered low values of the step-size μ in their analysis, and in that case, the results are approximately equivalent.

APPENDIX A

In this Appendix, we calculate the response of deterministic and random inputs to a linear time-variant periodic filter. Many digital signal processing textbooks analyze the response of a linear time-invariant system to stationary random input signals [39], [40], showing that if the random input signal is stationary in the wide sense, the output is also stationary in the wide sense. Moreover, the power spectrum of the output is the product of the input spectrum and the modulus squared frequency response of the system. However, when incomplete expansions are used, the equivalent transfer function of the AOLC is linear time-variant and periodic with impulse response $h[m,k]$. A closed form for the instantaneous impulse responses of the AOLC is not known in general. However, the instantaneous impulse response can be related to the output of the filter in response to an impulse function. Let $f[j,k]$ be the output of the AOLC at instant k when the input impulse was located at sample j. This signal will be causal $f[j,k] = 0; k \le j$. The impulse responses of the system $\{h[m,k]; k = 0, 1, \cdots, N-1\}$, where the first index denotes the impulse response waveform and the second index is the time instant when the impulse response is valid, can be written as $h[m,k] = f[k-m,k]$ because $f[j,k]$ can be expressed as the linear convolution

$$f[j,k] = \sum_{m=-\infty}^{\infty} h[m,k]x[k-m]$$
$$= \sum_{m=-\infty}^{\infty} h[m,k]\delta[k-j-m] = h[k-j,k]. \quad (55)$$

Therefore, $h[m,k] = 0; m \le 0$. The output $f[j,k]$ can be easily obtained running the AOLC filter. We calculate the response of such a time-variant system to deterministic and random input signals.

Let $d[k] = s[k] + n[k]$ denote the input signal composed of a periodic deterministic component $s[k]$ and a wide-sense stationary zero-mean random signal $n[k]$, with autocorrelation function $r_n[q] = E\{n[k]n[k+q]\}$. The output signal $y[k]$ can be also decomposed into two different components because the system is linear: $y[k] = y_s[k] + y_n[k]$, where $y_s[k]$ is the response of the system to the deterministic component $s[k]$ and $y_n[k]$ the response to $n[k]$.

The deterministic component can be directly obtained applying the linear convolution

$$y_s[k] = \sum_{m=-\infty}^{\infty} s[m]h[k-m,k] = \sum_{m=-\infty}^{\infty} s[m]f[m,k]. \quad (56)$$

The output $y_s[k]$ will be periodic because $s[k] = s[k+N]$, and $f[m,k] = f[m+N,k+N]$

$$y_s[k+N] = \sum_{m=-\infty}^{\infty} s[m]f[m,k+N]$$
$$= \sum_{m=-\infty}^{\infty} s[m-N]f[m-N,k] = y_s[k]. \quad (57)$$

The output signal power will also be periodic

$$y_s^2[iN+j] = \left(\sum_{m=-\infty}^{\infty} s[m]f[m,j]\right)^2 \quad \forall i. \quad (58)$$

The response to the stationary random input signal $y_n[k]$ will be random with expected value

$$E\{y_n[k]\} = \sum_{m=-\infty}^{\infty} E\{n[m]\}h[k-m,k] \quad (59)$$

and $E\{y_n[k]\} = 0$ if zero-mean noise is assumed. The autocorrelation of $y_n[k]$ will be

$$E\{y_n[k]y_n[k+q]\} = E\left\{\sum_{m_1=-\infty}^{\infty} h[m_1,k]n[k-m_1]\right.$$
$$\left. \cdot \sum_{m_2=-\infty}^{\infty} h[m_2,k+q]n[k+q-m_2]\right\} \quad (60)$$

which can be written as

$$E\{y_n[k]y_n[k+q]\}$$
$$= \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} h[m_1,k]\, h[m_2,k+q]$$
$$\cdot E\{n[k-m_1]n[k+q-m_2]\}. \quad (61)$$

The random input signal is stationary, and hence

$$E\{y_n[k]y_n[k+q]\}$$
$$= \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} h[m_1,k]$$
$$h \cdot [m_2,k+q]r_n[q+m_1-m_2]. \quad (62)$$

The output random signal is not wide-sense stationary because the autocorrelation function depends on the absolute time instant k, due to the time-varying impulse response of the system. The expression (62) can be simplified making the change of variables $m = m_2 - m_1$ obtaining

$$E\{y_n[k]y_n[k+q]\} = \sum_{m=-\infty}^{\infty} r_n[q-m]$$
$$\cdot \sum_{m_1=-\infty}^{\infty} h[m_1,k]h[m_1+m,k+q]. \quad (63)$$

For the misadjustment evaluation, we are interested in the steady-state residual noise power of the output signal. This value

can be obtained from (63) by setting the autocorrelation lag $q = 0$ at time instant $k = iN + j$ and taking the limit $i \to \infty$ as

$$
\lim_{i \to \infty} r_{y_n}[iN + j, iN + j]
$$

$$
= \sum_{m=-\infty}^{\infty} r_n[-m] \sum_{m_1=-\infty}^{\infty} h[m_1, j] h[m_1 + m, j]
$$

$$
= \sum_{m=-\infty}^{\infty} r_n[m] \sum_{m_1=-\infty}^{\infty} h[m_1, j] h[m_1 + m, j]. \quad (64)
$$

In the special case of a white-noise random input signal, $r_n[q] = \sigma_n^2 \delta[q]$, and the output signal power is simplified to

$$
\lim_{i \to \infty} r_{y_n}[iN + j, iN + j] = \sigma_n^2 \sum_{m_1=-\infty}^{\infty} h^2[m_1, j]. \quad (65)
$$

APPENDIX B

When the truncation error $c[k]$ is very small (due to high number of basis functions p) and very low values of the step-size μ are used, the analytical expressions that describe the excess MSE [(37) and (47)] can be greatly simplified. In this Appendix, we theoretically show that for this particular case both expressions are equivalent.

The transition matrices product over a complete signal occurrence $\mathbf{F}_{N+j-1,j}$ can be expanded as an Nth-degree polynomial of μ. If very low values of the step size are used and quadratic and higher order terms on μ can be neglected, $\mathbf{F}_{N+j-1,j}$ in (1) can be approximated as $\mathbf{F}_{N+j-1,j} \simeq (1 - 2\mu)\mathbf{I}$, which is the same result as in the complete expansion case. Using this approximation in (34), we can write

$$
\mathbf{H}_j \simeq \frac{1}{4\mu - 4\mu^2} \mathbf{Q}_j. \quad (66)
$$

If the truncation error can be considered null ($c[k] = 0$), the steady-state excess MSE will only be composed of the term

$$
\lim_{i \to \infty} \xi^{\text{ex}}[iN + j] = 4\mu^2 \sigma_n^2 \mathbf{X}^T[j] \mathbf{H}_j \mathbf{X}[j]
$$

$$
\simeq \frac{\mu \sigma_n^2}{1 - \mu} \mathbf{X}^T[j] \mathbf{Q}_j \mathbf{X}[j] \quad (67)
$$

where $\mathbf{Q}_j$ was defined in (33). Using (33) and neglecting quadratic terms on μ, we can write

$$
\lim_{i \to \infty} \xi^{\text{ex}}[iN + j]
$$

$$
\simeq \frac{\mu}{1 - \mu} \sigma_n^2 \sum_{m=j}^{N+j-1} \mathbf{X}^T[j] \mathbf{X}[m] \mathbf{X}^T[m] \mathbf{X}[j]
$$

$$
\simeq \frac{\mu}{1 - \mu} \sigma_n^2 \sum_{m=j}^{N+j-1} r_{jm}^2 = \frac{\mu}{1 - \mu} \sigma_n^2 \sum_{m=0}^{N-1} r_{jm}^2. \quad (68)
$$

For the second expression of steady-state excess MSE (47), it is shown in [30] and [31] that when low values of the step-size μ are used, the response to an impulse function at sample j

is approximately equal to the convolution of the time-invariant impulse response corresponding to the complete expansion (38) and the inner product values of the reference vectors $r_{jm} = \mathbf{X}^T[j]\mathbf{X}[m]$. Let $f[j, k]$ be the output of the system at instant k when the impulse is located at sample j, writing

$$
f[j, k] \simeq \begin{cases} 0, & k \leq j \\ r_{jk} 2\mu (1 - 2\mu)^u, & k > j, u = \lfloor \frac{k}{N} \rfloor. \end{cases} \quad (69)
$$

The steady-state excess MSE can be calculated by setting $c[k] = 0$ in (47) and then taking the limit

$$
\lim_{i \to \infty} \xi^{\text{ex}}[iN + j] = \sigma_n^2 \sum_{m=-\infty}^{\infty} h^2[m, j]. \quad (70)
$$

If we apply (69) and (55) to the last equation, we finally obtain

$$
\lim_{i \to \infty} \xi^{\text{ex}}[iN + j]
$$

$$
\simeq \sigma_n^2 \sum_{i=0}^{\infty} (2\mu)^2 (1 - 2\mu)^{2i} \sum_{m=j+1}^{N+j-1} r_{jm}^2
$$

$$
= \frac{\mu \sigma_n^2}{1 - \mu} \sum_{m=0}^{N-1} r_{jm}^2. \quad (71)
$$

In conclusion, we have theoretically shown that both expressions (37) and (47) for the steady-state excess MSE in the particular case of very low values of the step-size and low truncation error give the same approximation. However, experimental results show their equivalence for any value of the step-size μ and the number of functions p.

REFERENCES

[1] E. Ferrara and B. Widrow, "Fetal electrocardiogram enhancement by time-sequenced adaptive filtering," *IEEE Trans. Biomed. Eng.*, vol. BME-29, pp. 458–459, June 1982.

[2] Z. Yi-Sheng and N. V. Thakor, "P-wave detection by an adaptive QRS-T cancellation technique," in *Comput. Cardiology*. New York: IEEE Comput. Soc. Press, 1987, pp. 249–252.

[3] N. V. Thakor and Z. Yi-Sheng, "Applications of adaptive filtering to ECG analysis: Noise cancelation and arrhythmia detection," *IEEE Trans. Biomed. Eng.*, vol. 38, pp. 785–794, Aug. 1991.

[4] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.

[5] B. Widrow, J. McCool, M. G. Larimore, and R. Johnson, "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, pp. 1151–1162, Aug. 1976.

[6] L. Horowitz and K. D. Senne, "Performance advantage of complex LMS for controlling narrow-band adaptive arrays," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 722–736, Mar. 1981.

[7] B. Fisher and N. J. Bershad, "The complex LMS adaptive algorithm—Transient weight mean and covariance with application to the ALE," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 34–44, Feb. 1983.

[8] W. Gardner, "Learning characteristics of stochastic-gradient-descent algorithms: A general study, analysis and critique," *Signal Process.*, vol. 6, pp. 113–133, Apr. 1984.

[9] A. Feuer and E. Weinstein, "Convergence analysis of LMS filters with uncorrelated Gaussian data," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 222–230, Feb. 1985.

[10] J. E. Mazo, "On the independence theory of equalizer convergence," *Bell Syst. Tech. J.*, vol. 58, pp. 963–993, 1979.

[11] D. C. Farden, "Stochastic approximation with correlated data," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 105–113, 1981.

[12] O. Macchi, *Adaptive Processing: The Least Mean Square Approach with Applications in Transmission*. New York: Wiley, 1995.

[13] J. R. Glover, "Adaptive noise cancelling applied to sinusoidal interferences," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-25, pp. 484–491, Dec. 1977.

[14] M. J. Shensa, "Non-Wiener solutions for the adaptive canceller with a noisy reference," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, pp. 468–473, Aug. 1980.

[15] S. J. Elliott and P. Darlington, "Adaptive cancellation of periodic, synchronously sampled interference," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 715–717, June 1985.

[16] N. J. Bershad and P. L. Feintuch, "Non-Wiener solutions for the LMS altorithm—A time domain approach," *IEEE Trans. Signal Processing*, vol. 43, pp. 1273–1275, May 1995.

[17] N. Bershad and J. Bermudez, "Sinusoidal interference rejection analysis of an LMS adaptive feedforward controller with anoisy peridic interference," *IEEE Trans. Signal Processing*, vol. 46, pp. 1298–1313, May 1998.

[18] N. V. Thakor, X. Guo, C. Vaz, P. Laguna, R. Jané, P. Caminal, H. Rix, and D. Hanley, "Orthonormal (Fourier and Walsh) models of time-varying evoked potentials in neurological injury," *IEEE Trans. Biomed. Eng.*, vol. 40, pp. 213–221, Mar. 1993.

[19] P. Laguna, R. Jané, O. Meste, P. W. Poon, P. Caminal, H. Rix, and N. V. Thakor, "Adaptive filter for event-related bioelectric signals using an impulse correlated reference input: Comparison with signal averaging techniques," *IEEE Trans. Biomed. Eng.*, vol. 39, pp. 1032–1044, Oct. 1992.

[20] C. Vax, X. Kong, and N. V. Thakor, "An adaptive estimation of periodic signals using a Fourier linear combiner," *IEEE Trans. Signal Processing*, vol. 42, pp. 1–10, Jan. 1994.

[21] A. K. Barros and M. Y. Yasuda, "Filtering noncorrelated noise in impedance cardiography," *IEEE Trans. Biomed. Eng.*, vol. 42, pp. 324–327, Mar. 1995.

[22] P. Laguna, R. Jané, S. Olmos, N. V. Thakor, H. Rix, and P. Caminal, "Adaptive estimation of QRS complex by the Hermite model for classification and ectopic beat detection," *Med. Biol. Eng. Comput.*, vol. 34, pp. 58–68, 1996.

[23] J. García, S. Olmos, G. Moody, R. Mark, and P. Laguna, "Adaptive estimation of Karhunen-Loève series applied to the study of ischemic ECG records," in *Comput. Cardiol.*: IEEE Comput. Soc., 1996, pp. 249–252.

[24] S. Olmos, M. Millán, J. García, and P. Laguna, "ECG data compression with the Karhunen-Loève transform," in *Comput. Cardiology*. New York: IEEE Comput. Soc. Press, 1996, pp. 253–256.

[25] P. Laguna, G. Moody, R. Jané, P. Caminal, and R. Mark, "Karhunen-Loève transform as a tool to analyze the ST-segment," *J. Electrocardiol.*, vol. 28, pp. 41–49, 1996.

[26] A. K. Barros and N. Ohnishi, "MSE behavior of biomedical event-related filters," *IEEE Trans. Biomed. Eng.*, vol. 44, pp. 848–855, Sept. 1997.

[27] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[28] P. Clarkson and P. White, "Simplified analysis of the LMS adaptive filter using a transfer function approximation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 987–993, July 1987.

[29] P. Laguna, R. Jané, E. Masgrau, and P. Caminal, "The adaptive linear combiner with a periodic-impulse reference input as a linear comb filter," *Signal Process.*, vol. 48, no. 3, pp. 193–203, 1996.

[30] S. Olmos, J. Garcia, R. Jané, and P. Laguna, "Truncated orthogonal expansions of recurrent signals: Equivalence to a periodic time-variant filter," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 1998, pp. 1709–1712.

[31] ——, "Truncated orthogonal expansions of recurrent signals: Equivalence to a linear time-variant periodic filter," *IEEE Trans. Signal Processing*, vol. 47, pp. 3164–3172, Nov. 1999.

[32] P. Laguna, R. Mark, A. Goldberger, and G. Moody, "A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG," in *Comput. Cardiology*. New York: IEEE Comput. Soc. Press, 1997.

[33] B. Widrow, J. Glover, J. McCool, J. Kaunitz, C. S. Williams, R. Hearn, J. Zeidler, E. Dong, Jr., and R. Goodlin, "Adaptive noise cancelling: Principles and applications," *Proc. IEEE*, vol. 63, pp. 1692–1716, Dec. 1975.

[34] N. V. Thakor, "Adaptive filtering of evoked potentials," *IEEE Trans. Biomed. Eng.*, vol. BME-34, pp. 6–12, 1987.

[35] P. Poon and N. V. Thakor, "Adaptive Fourier modeling of evoked potentials: Application to cerebral focal ischemia," in *Proc. 12th Annu. Conf. IEEE EMBS*, 1990, pp. 885–886.

[36] C. Vaz and N. V. Thakor, "Adaptive Fourier estimation of time-varying evoked potentials," *IEEE Trans. Biomed. Eng.*, vol. 36, pp. 448–455, Apr. 1989.

[37] D. G. Luemberger, *Introduction to Dynamic Systems: Theory, Models and Applications*. New York: Wiley, 1979.

[38] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[39] P. M. Clarkson, *Optimal and Adaptive Signal Processing*. Boca Raton, FL: CRC, 1993.

[40] C. W. Therrien, *Discrete Random Signals and Statistical Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1992.

Salvador Olmos was born in Valencia, Spain, in 1969. He received the Ind. Eng. and the Ph.D. degrees from the Polytechnic University of Catalonia, Barcelona, Spain, in 1993 and 1998, respectively.

Since 1994, he has been an Assistant Professor with the Department of Electronic Engineering and Communications, University of Zaragoza, Zaragoza, Spain. His professional research interests are in data compression and signal processing of biomedical signals.

Pablo Laguna was born in Jaca, Huesca, Spain, in 1962. He received the M.S. degree in physics and the Ph.D. degree from the University of Zaragoza (UZ), Zaragoza, Spain, in 1985 and 1990, respectively. His Ph.D. thesis was developed at the Biomedical Engineering Division of the Institute of Cybernetics (IC), Polytechnic University of Catalonia (UPC), Barcelona, Spain.

He is currently an Associate Professor of Signal Processing and Communications with the Department of Electronics Engineering and Communications, Centro Politécnico Superior, UZ. From 1987 to 1992, he was an Assistant Professor with the Department of Control Engineering, UPC, and a Researcher at the Biomedical Engineering Division, IC. His professional research interests are in signal processing, in particular applied to biomedical applications.