An Adaptive Estimation of Periodic Signals Using a Fourier Linear Combiner

Christopher Vaz, Xuan Kong, *Member, IEEE*, and Nitish Thakor, *Senior Member, IEEE*

***Abstract*—Here, we present an adaptive algorithm for estimating from noisy observations, periodic signals of known period subject to transient disturbances. The estimator is based on the LMS algorithm and works by tracking the Fourier coefficients of the data. The estimator is analyzed for convergence, noise misadjustment and lag misadjustment for signals with both time invariant and time variant parameters. The analysis is greatly facilitated by a change of variable that results in a time invariant difference equation. At sufficiently small values of the LMS step size, the system is shown to exhibit decoupling with each Fourier component converging independently and uniformly. Detection of rapid transients in data with low signal to noise ratio can be improved by using larger step sizes for more prominent components of the estimated signal. An application of the Fourier estimator to estimation of brain evoked responses is included.**

I. Introduction

NUMEROUS applications involve the recording of periodic signals in background noise. In the biomedical field, several signals recorded during patient monitoring are periodic; examples are neurologically evoked potentials, cardiograms, and pneumograms. It is quite common that many biological signals are also driven to be periodic by the application of external stimuli. Examples are the paced (electrical stimulation by a pacemaker device) heart rhythm (the electrocardiogram) and the evoked brain response (response of the brain to periodic external stimulation such as electrical shock or auditory clicks or visual flashes). Therefore, study of periodic signals is broadly important. Very often, transient changes in these signals carry important information about the system. When very little *a priori* information about the signal is available, an adaptive estimator usually offers the best solution.

One of the most widely used adaptive algorithms is the least mean square (LMS) algorithm, developed by Widrow and Hoff [12]. The algorithm is simple and computationally efficient, and has been successfully applied in a wide variety of situations in the past decades. When used for the purpose of system identification, the algorithm attempts to estimate a possibly time-varying linear system whose weight or parameter vector $\underline{w}_k^o$ obeys

$$y_k = \underline{x}_k^{*\prime} \underline{w}_k \underline{w}_k^o + v_k \tag{1.1}$$

where the prime denotes transposition and the asterisk denotes complex conjugation, and v_k is the inherent noise in the system, assumed to be uncorrelated with $\underline{x}_k$. LMS is the steepest descent gradient search type iterative algorithm, with the weight estimate denoted by $\underline{w}_k$. It attempts to minimize the square of the error

$$e_k = y_k - \underline{x}_k^{*\prime} \underline{w}_k \tag{1.2}$$

by means of the recursion [17]

$$\underline{w}_{k+1} = \underline{w}_k + \mu \underline{x}_k e_k \tag{1.3}$$

where μ is a positive real number, known as the step size, that determines the size of the adjustment for each iteration.

There has been a good deal of work on the subject which has contributed to an analytical description of the properties of LMS. Such analyses have dealt with the issues of convergence, misadjustment due to noise, and misadjustment due to parameter variation (known as lag misadjustment). Widrow and his colleagues [14], [15] have provided a heuristic analysis of the LMS algorithm, discussing convergence properties, misadjustment error and learning characteristics. Bitmead [3] has rigorously derived conditions for convergence in distribution of the LMS estimates; if those conditions are satisfied, convergence of (1.3) to a stationary solution is exponential. Feuer and Weinstein [4] have analyzed convergence and misadjustment in the case of uncorrelated Gaussian data, a case that is of limited practical significance. However, Macchi and Eweda [6] and Shi and Kozin [8] have derived conditions for convergence for the more general case of dependent or colored data; Bitmead and Anderson [2] and Solo [9] have looked at both convergence and misadjustment.

There are instances which call for the analysis of the LMS algorithms with periodic regression vector $\underline{x}_k$ [1], [10], [16]. Many of the more general results mentioned above cannot be applied, since they are based on the assumption that $\underline{x}_k$ in (1.1) is not totally deterministic; even if they could be modified to accommodate a deterministic $\underline{x}_k$, there are certain simplifications and important insights that can be gained by starting out with the periodicity assumption. Widrow *et al.* [16] discussed the relationship between the DFT of the input s_k and the weight vector $\underline{w}_k$ of the LMS algorithm (1.3) with

$$\underline{x}_k = \frac{1}{\sqrt{N}}[1 \quad e^{-j\omega_0 k} \cdots e^{-j\omega_0 (N-1)}]', \qquad \omega_0 = 2\pi/N \tag{1.4}$$

Manuscript received June 7, 1990; revised January 12, 1993. The associate editor coordinating the review of this paper and approving it for publication was Prof. John J. Shynk. This work was supported in part by the National Institutes of Health under Grant NS24282.

C. Vaz and N. Thakor are with the Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD 21205.

X. Kong is with the Department of Electrical Engineering, Northern Illinois University, DeKalb, IL 60115.

IEEE Log Number 9213267.

and

$$e_k = s_k - x_w^{*\prime}\underline{w}_k. \tag{1.5}$$

It was found that for properly selected step size μ, the weight vector $\underline{w}_k|_{k=lN}$ is proportional to the DFT of input $\{s_m\}_{m=lN-N}^{lN-1}$. Because of this, the algorithm (1.3) and (1.4) was referred to as the LMS spectrum analyzer. In [1], the spectrum analyzer was generalized to cover the case where the components of the regression vector $\underline{x}_k$ are not evenly spaced in the frequency domain.

If we have a periodic signal s_k with a period of T sampling intervals, then it can be represented by a Fourier series, i.e.,

$$s_k = \sum_{n=\pm 1}^{\pm N} w_n^o e^{j\omega_0 nk} = x_k^{*\prime}\underline{w}^o, \qquad \omega_0 = 2\pi/T \tag{1.6}$$

where

$$\underline{x}_k = [e^{j\omega_0 Nk} \quad e^{j\omega_0(N-1)k} \cdots e^{j\omega_0 k} \quad e^{-j\omega_0 k} \cdots e^{-j\omega_0 Nk}]' \tag{1.7}$$

$$\underline{w}^o = [w_{-N}^o \quad w_{-N-1}^o \cdots w_{-1}^o \quad w_1^o \cdots w_N^o]. \tag{1.8}$$

It is assumed that the signal s_k can be described by no more than N harmonics (N is called the model order) and the maximum model order is $T/2$ (for even T). The choice of model order is dictated by the bandwidth of the signal s_k. With $2N + 1 = T$, analysis of the adaptive algorithm can be simplified with the argument similar to the one in [16]. In our application [10], however, $2N + 1 < T$ holds; therefore the analysis presented in this paper becomes essential. The "dc" term is omitted since "ac" coupling is assumed. If a "dc" component is present, it can always be removed prior to processing and added later.

With regression vector (1.7), the Fourier series w_n^o of the signal s_k can be adaptively estimated using (1.3). An adaptive system shown in Fig. 1 has been used by Vaz and Thakor [10] to estimate periodic signal s_k from noisy observations $y_k = s_k + v_k$. The noise v_k is assumed to be independent of the signal s_k and it is stationary with covariance $K_v(i) = E[v_k v_{k+i}]$ and power spectral density $S_v(\omega)$. Thus, the LMS algorithm (1.3) with $\underline{x}_k$ given in (1.7) and

$$e_k = \underline{x}_k^{*\prime}\underline{w}^o + v_k - x_k^{*\prime}\underline{w}_k \tag{1.9}$$

yields adaptive estimates of the Fourier series $\underline{w}^o$. The LMS spectrum analyzer was renamed as the Fourier linear combiner (FLC) in [10], since the signal is represented by a linear combination of complex exponentials.

In this paper, the adaptive FLC for estimating periodic signal in the presence of noise is studied. It is worth noting that in our current analysis the period of the signal is assumed to be known, since this is the case in the applications we are concerned with. The analytical results have been used successfully to guide design and implementation of the adaptive algorithms in several biomedical engineering applications. In Section II, a transformation is introduced to facilitate the analysis of the adaptive algorithm. Next, in Section III, the noise misadjustment analysis is carried out for the case of the

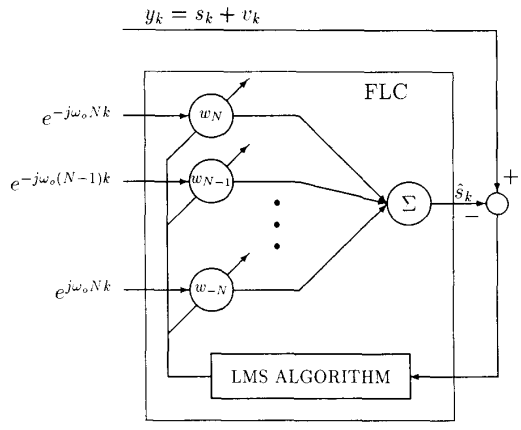


Fig. 1. Block diagram of the Fourier linear combiner (also defined as the spectrum analyzer in [1].)

periodic signal with the time invariant Fourier coefficients. The decoupling phenomenon of the FLC with small adaptive gain μ is investigated and a closed form for the noise misadjustment is derived.

In many situations the signal characteristics vary with time and, in fact, such variability is one of the most important reasons for using adaptive algorithms. A model of such a periodic signal, with its Fourier coefficients changing from period to period (the period itself is known and fixed), is proposed in Section IV. An expression for the lag misadjustment, defined as the extra amount of mean square parameter estimation error due to the time varying nature of the signal characteristics, is obtained. This is the first analysis of the adaptive FLC with time-varying signal characteristics. A nonuniform FLC is proposed and analyzed in Section V. Also, a formula for calculating the noise misadjustment is derived. Based on this formula, the nonuniform FLC can be properly set up to achieve the desired misadjustment level. Simulation results and discussions are presented in Section VI. Finally, a brief summary is given in Section VII.

II. The Time- Invariant Transformation

In order to analyze the performance of the FLC, we shall look at the weight error vector. We assume stationarity, i.e., perfectly periodic s_k and stationary v_k. Then, the solution weight vector is $\underline{w}^o$ satisfying

$$s_k = \underline{x}_k^{*\prime}\underline{w}^o \tag{2.1}$$

where $\underline{x}_k$ and $\underline{w}^o$ are given by (1.7) and (1.8), respectively. To compute statistical moments, we introduce cyclic moments, which are moments averaged over any cycle or period consisting of T consecutive samples. Such cyclic moments will be denoted by "E_{cyc}," instead of "E." The relationship between the two operators may be expressed as

$$E_{\text{cyc}} f(\cdot) \equiv \frac{1}{T}\sum_{i=k+1}^{k+T} Ef(\cdot, i), \qquad k \text{ being any integer.} \tag{2.2}$$

Our goal is to quantify the total steady state misadjustment $\mathcal{E}$ after all the transients have died down. This misadjustment may be defined as the mean square error of the estimate $\hat{s}_k = \underline{x}_k^{*\prime}\underline{w}_k$ with reference to the underlying periodic signal $s_k = \underline{x}_k^{*\prime}\underline{w}^o$, i.e.

$$E = \lim_{k\to\infty} E_{\text{cyc}}[(\underline{x}_k^{*\prime}\underline{w}^o - x_k^{*\prime}\underline{w}_k)^2]. \tag{2.3}$$

Let us see how $\mathcal{E}$ can be expressed in terms of the second order statistics of parameter estimation error. From (1.3) and (1.9)

$$\begin{aligned}\underline{w}_{k+1} &= \underline{w}_k + \mu\underline{x}_k(\underline{x}_k^{*\prime}\underline{w}^o + v_k - \underline{x}_k^{*\prime}\underline{w}_k)\\ \Rightarrow\quad \underline{w}_{k+1} - \underline{w}^o &= \underline{w}_k - \underline{w}^o + \mu\underline{x}_k v_k\\ &\quad - \mu\underline{x}_k\underline{x}_k^{*\prime}(\underline{w}_k - \underline{w}^o).\end{aligned}$$

Introducing the instantaneous weight error vector $\underline{\tilde{w}}_k = \underline{w}_k - \underline{w}^o$, we have the error system

$$\underline{\tilde{w}}_{k+1} = (I - \mu\underline{x}_k\underline{x}_k^{*\prime})\underline{\tilde{w}}_k + \mu\underline{x}_k v_k. \tag{2.4}$$

Since $\underline{v}_k$ is independent of $\underline{x}_k$, and both are zero-mean, it is clear from (2.4) that $E[\underline{\tilde{w}}_k] = 0$ in steady state. This observation ensures that the estimator is unbiased. Following the discussions in [15, p. 102], we can easily see that $\mu < 2/\text{tr}\,(R_{xx}) = 1/N$ is a sufficient condition to guarantee the convergence of the algorithm in mean (where R_{xx} is the covariance matrix of $\underline{x}_k$).

System (2.4) has a time-varying transition matrix, i.e., $(I - \mu\underline{x}_k\underline{x}_k^{*\prime})$, which is hard to manipulate when looking at second-order statistics. However, one can take advantage of the periodicity of $\underline{x}_k$ and of the transition matrix, to transform (2.4) into a time-invariant system by means of a simple change of variables. This transformation was first presented in Annex 1 of [1] and in [13].

Let us define a $2N \times 2N$ diagonal matrix D with diagonal elements

$$\begin{aligned}D(q,q) &= d_q = e^{j\omega_0 q},\\ q &= -N, -(N-1), \cdots, -1, 1, \cdots, N\end{aligned} \tag{2.5}$$

and

$$\underline{x}_0 = \overbrace{[1 \quad 1 \cdots 1]'}^{2N \text{ times}} \tag{2.6}$$

then it is easy to see that the following relations hold

$$\underline{x}_k = D^{-k}\underline{x}_0, \quad \text{and} \quad \underline{x}_k^* = D^k\underline{x}_0. \tag{2.7}$$

Premultiplying both sides of (2.4) by D^{k+1}, we have

$$\begin{aligned}D^{k+1}\underline{\tilde{w}}_{k+1} &= D^{k+1}\underline{\tilde{w}}_k - \mu D^{k+1}D^{-k}\underline{x}_0\underline{x}_0'D^k\underline{\tilde{w}}_k\\ &\quad + \mu D^{k+1}D^{-k}\underline{x}_0 v_k\\ &= D(I - \mu\underline{x}_0\underline{x}_0')D^k\underline{\tilde{w}}_k + \mu D\underline{x}_0 v_k.\end{aligned} \tag{2.8}$$

Defining

$$\underline{\tilde{f}}_k = D^k\underline{\tilde{w}}_k, \quad \text{and} \quad H = D(I - \mu\underline{x}_0\underline{x}_0') \tag{2.9}$$

(2.8) can be rewritten as

$$\underline{\tilde{f}}_{k+1} = H\underline{\tilde{f}}_k + \mu D\underline{x}_0 v_k. \tag{2.10}$$

Note that system (2.10) is a time-invariant system as desired.

III. Noise Misadjustment Analysis

A. Calculation of Noise Misadjustment and the Convergence Criterion

If we assume that the true parameter vector is time invariant (2.1), then the steady state mean square error (MSE) is synonymous with the noise misadjustment, M. In other words,

$$M = \lim_{k\to\infty} E_{\text{cyc}}[(s_k - \hat{s}_k)^2] \tag{3.1}$$

i.e., $\mathcal{E}$ can be replaced by M in (2.3) provided s_k is defined as in (2.1).

We are interested in the error term $s_k - \hat{s}_k$:

$$s_k - \hat{s}_k = -\underline{x}_k^{*\prime}\underline{\tilde{w}}_k = -(D^k\underline{x}_k)^{*\prime}\underline{\tilde{f}}_k. \tag{3.2}$$

From (2.7) we have $D^k\underline{x}_k = \underline{x}_0 = [1 \quad 1 \cdots 1]'$, thus

$$s_k - \hat{s}_k = -\underline{x}_0'\underline{\tilde{f}}_k. \tag{3.3}$$

It follows from (2.3) and (3.3) that the quantity we are seeking is

$$\mathcal{E} = M = \lim_{k\to\infty} E_{\text{cyc}}[\underline{x}_0'\underline{\tilde{f}}_k\underline{\tilde{f}}_k^{*\prime}\underline{x}_0]. \tag{3.4}$$

In Appendix A, it is shown that

$$\mathcal{E} = \frac{\mu^2}{2\pi}\int_{-\pi}^{\pi} S_v(\omega)H(\omega)d\omega \tag{3.5}$$

where $S_v(\omega)$ is the power spectral density of the noise v_k and $H(\omega)$ is defined as follows:

$$H(\omega) \equiv ||\underline{x}_0'(H - Ie^{j\omega})^{-1}D\underline{x}_0||^2. \tag{3.6}$$

A valid concern is whether the second-order moment (3.4) converges as $k \to \infty$. The analysis in Appendix A shows that limit (3.4) exists provided $\lim_{k\to\infty} H^k = 0$ (i.e., the null matrix). Of course, this is intuitively obvious from (2.10). In other words, the eigenvalues of H must be of absolute value less than unity in order for the MSE to converge.

B. Frequency Response of the Noise Misadjustment Function

Examination of frequency response (3.5) reveals some interesting properties. In Appendix B, a closed form expression for the function $\mu^2 H(\omega)$ is obtained for the case $N = 1$. It is noted that the function $\mu^2 H(\omega)$ has unit amplitude at frequencies $\omega = \pm\omega_0$ no matter how small the adaptive gain μ is. This function, when modulated by $S_v(\omega)$, constitutes the integrand in (3.5). Fig. 2 plots $\mu^2 H(\mu)$ as a function of normalized frequency ω for two different μ. Note, the two sharp peaks at $\omega = \pm\omega_0 = \pm 2\pi/128$. In general, for an estimator of order N, there will be $2N$ peaks at $\pm\omega_0, \pm 2\omega_0, \cdots, \pm N\omega_0$.

Thus, we see that the overwhelming contribution to the noise misadjustment function is from frequencies in the neighborhood of $\pm\omega_0, \pm 2\omega_0, \cdots, \pm N\omega_0$ for an estimator of order N. The smaller the value of μ is, the more narrow this neighborhood will be. In Appendix B, we have obtained a formula relating the bandwidth (BW) of the peaks (the frequency range within which the amplitude is no less than 50

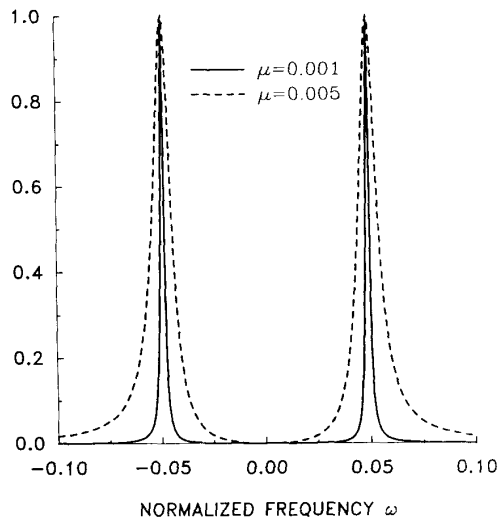


Fig. 2. Plot of $\mu^2 H(\omega)$. This function may be viewed as the "signal factor" in the integrand of (3.5), $S_v(\omega)$ being the "noise factor." When multiplied by $S_v(\omega)/2\pi$. the area under the curve will be the steady state MSE. N = 1, T = 128.

TABLE I
TUNING FILTER BANDWIDTH AS A FUNCTION OF THE ADAPTIVE GAIN

μ	Bandwidth	
	From (B3) Actual Value	From (3.7) Predicted Value
0.001	0.002	0.002
0.002	0.004	0.004
0.005	0.010	0.010
0.01	0.020	0.021
0.02	0.041	0.048

percent of the maximum) and the adaptive gain μ for $N = 1$ as follows:

$$\mathrm{BW} = \frac{2\sin^2 \omega_0}{|\mu^2 \cos^2 \omega_0 - \sin^2 \omega_0|}\mu. \tag{3.7}$$

Note, that when $\mu \ll 1$, (3.7) can be approximated by

$$\mathrm{BW} \approx 2\mu \tag{3.8}$$

which is in agreement with the result in [1]. Table I shows that the bandwidth given by (3.7) are in very good agreement with the actual values obtained from (B3). Though the above formula is obtained for the case $N = 1$, it is expected to be a good approximation for $N > 1$ if BW $\ll \omega_0$, since the bandpass filter centered at one harmonic contributes very little to the filter output near any other harmonics.

From (3.5), one can see that the integrand would then be negligible at frequencies that are not very close to any of the harmonics of the fundamental frequency ω_0. Thus, the noise within these bands virtually has no effect on the system performance. One obvious advantage of this result is the rejection of periodic interference. For instance, if $T = 40$ msec., $\omega_0 = 25$ Hz. Thus, any periodic noise with spectral peaks sufficiently far away from multiples of 25 Hz (such as 60 Hz power line interference) would not corrupt the FLC estimates at small values of μ. One may say that the algorithm resembles a tuning filter, tuned to the fundamental frequency and its harmonics. Transfer function analysis of the algorithm in [1] shows that this is indeed the case.

C. The Decoupled Linear Approximation

Another interesting aspect of the tuning phenomenon is the link between overlapping of adjacent peaks and the degree of coupling of the system as expressed by the vector differential equation (2.10). The larger the value of μ is, the greater is the overlap between adjacent peaks as shown in Fig. 2, indicating a stronger coupling. Such a coupling can be seen intuitively from the definition of H (2.9). As μ becomes smaller, H approaches the form of a diagonal matrix ($\approx (1-\mu)D$, with error $||H - (1-\mu)D|| = O(\mu)$). As a result, the system equation (2.10) becomes more and more decoupled. The decoupled approximation is quite useful for processing data of low SNR by the adaptive FLC with small μ. We show in Appendix C, that the decoupling approximation results in an error of $O(\mu)$ for $\underline{\tilde{f}}_k$. Thus, this approximation will introduce an error in misadjustment calculation of size $O(\mu^{3/2})$ at most.

The decoupled system can be written as

$$\underline{\tilde{f}}_{k+1} = (1-\mu)D\underline{\tilde{f}}_k + \mu D\underline{x}_0 v_k. \tag{3.9}$$

Let $\underline{a}_{k,q}$ denote the qth element of any vector function $\underline{a}_k$. Then, (3.9) can be broken down into $2N$ independent scalar difference equations

$$f_{k+1,q} = (1-\mu)d_q f_{k,q} + \mu d_q v_k, \quad q = \pm 1, \cdots, \pm N. \tag{3.10}$$

Summing up (3.10) from 1 to k and neglecting the exponentially decaying term $(1-\mu)^k d_q f_{0,q}$, we get

$$\begin{aligned} f_{k,q} &= \mu \sum_{i=0}^{k-1} (1-\mu)^{k-i-1} d_q^{k-1-i} d_q v_i \\ &= \mu \sum_{i=0}^{k-1} (1-\mu)^i d_q^i d_q v_{k-1-i}. \end{aligned} \tag{3.11}$$

The misadjustment M of the decoupled system is still given by (3.4). To calculate M, we need the quantity $\underline{x}_0' \underline{\tilde{f}}_k$, which can be easily obtained based on the observation that $\underline{x}_0' = [1 \quad 1 \cdots 1]$. Thus,

$$\underline{x}_0' \underline{\tilde{f}}_k = \mu \sum_{q=\pm 1}^{\pm N} \sum_{i=0}^{k-1} (1-\mu)^i d_q^i d_q v_{k-1-i}. \tag{3.12}$$

The assumption that the noise autocovariance of lags greater than p is negligible reduces (3.4) to (see Appendix C)

$$M = \frac{\mu}{2}\left(K_v(0)\kappa_T(0) + 2\sum_{m=1}^{p} K_v(m)\kappa_T(m) \right) + O(\mu)^2 \tag{3.13}$$

where $K_v(i) \equiv E[v_k v_{k+i}]$ is the noise autocovariance function and $\kappa_T(m)$ is defined as

$$\kappa_T(m) = \frac{1}{T}\sum_{i=0}^{T-1}(1-\mu)^{2i}\sum_{p=\pm1}^{\pm N}\sum_{q=\pm1}^{\pm N}\exp[j\omega_0(i+1)p] \exp[-j\omega_0(i+m+1)q]. \quad (3.14)$$

Equation (3.13) confirms the statement we have made earlier about the second order statistics being $O(\mu)$. Ignoring terms of μ^2 and higher leaves us with the decoupled linear approximation. It is of the same form as the approximate expression for the misadjustment in [10] which was derived using averaging approximations. For small step sizes, (3.13) provides a useful means of choosing the step size in order to satisfy prescribed misadjustment specifications. A specified misadjustment value may function as a "tolerance" of the estimates. Of course, availability of $K_v(i)$ needs to be investigated. In many (certainly in biomedical engineering) applications, controlled experiments can be performed for estimating the noise statistics [10].

It is also apparent that, in the absence of noise, (3.9) is the equation for a simple exponential decay. The time constant of the decay, i.e., of the convergence of the weights, is

$$\tau_w = -1/\log(1-\mu) \approx 1/\mu, \qquad \text{for } \mu \ll 1. \quad (3.15)$$

According to (2.3), the misadjustment is a quadratic function of the weights. So, for small μ, the time constant for the misadjustment is approximately half that of the mean time constant for weights

$$\tau_M \approx \tau_w/2 \approx 0.5\mu^{-1}, \qquad \text{for } \mu \ll 1 \quad (3.16)$$

(since all weights have the same time constant τ_w). A significant implication of this result is that the convergence tends to be uniform with a single time constant. Thus, owing to the fact that $\underline{x}_k$ is an orthonormal data-independent reference set, the chief shortcoming of LMS (i.e., disparity of time constants of convergence) has been overcome [7].

IV. Lag Misadjustment Analysis

When the signal to be estimated is perfectly periodic, the only misadjustment is caused by the background noise. However, if the signal deviates from stationary representation (2.1), so that there is a time-varying solution, there is an additional misadjustment known as the lag misadjustment. An analysis of this situation necessitates the formulation of a model for the parameter variation.

A. The Punctuated Perturbation Model

We formulate a simple variation model where each cycle (period) of the underlying signal is considered to be represented by a single set of Fourier coefficients. However, this set of Fourier parameters varies from one cycle to the next in a random fashion. This model has been found to work well for evoked response studies [11], where each cycle consists of a response of the central nervous system to an applied peripheral stimulus. It is especially suited to situations where each cycle is considered to be an independent phenomenon. We call this parameter variation model the punctuated perturbation model (abbreviated as PPM). In this scheme, denote

$$k = nT + p \quad (4.1)$$

where k is the time index and T is the fixed and known cycle length in conformity with the notation used thus far. The integers n (quotient of k/T) and p (remainder of k/T) denote the cycle index ($n \geq 0$) and latency index ($0 \leq p \leq T-1$), respectively. In other words, the time k corresponds to the pth point in the nth cycle. According to the PPM scheme then, under steady state conditions,

$$s_k = \underline{x}_k^{*\prime}\underline{w}_n^o. \quad (4.2)$$

Equation (4.2) may be compared to (2.1) in which $\underline{w}_n^o$ is simply $\underline{w}^o$, a constant parameter vector. The perturbation is "punctuated" because the parameter vector is constant for an entire cycle and then jumps to another point in the parameter space at the beginning of the next cycle. The following conditions are assumed to hold:

i) $\underline{w}_n^o$ is a stationary stochastic vector process.
ii) $E[\underline{w}_n^o] = \overline{\underline{w}}^o$.
iii) $\underline{w}_n^o$ is an independent, identically distributed (i.i.d.) random vector process with $E[(\underline{w}_l^o - \overline{\underline{w}}^o)(\underline{w}_m^o - \overline{\underline{w}}^o)^{*\prime}] = \underline{0}$ for $l \neq m$, where $\underline{0}$ is the null matrix.

Let us now define

$$\underline{\delta}_n = \underline{w}_{n+1}^o - \underline{w}_n^o \quad (4.3)$$

$\underline{\delta}_n$ refers to the parameter "jump" from one cycle to the next. Obviously, $\underline{\delta}_n$ would have nonzero autocorrelation only for lags of 0 and 1. Denote the variance of $\underline{\delta}_n$

$$\underline{\Lambda} = E[\underline{\delta}_n\underline{\delta}_n^{*\prime}] = 2E[(\underline{w}_n^o - \overline{\underline{w}}^o)(\underline{w}_n^o - \overline{\underline{w}}^o)^{*\prime}] \quad (4.4)$$

so that

$$E[\underline{\delta}_n\underline{\delta}_{n-1}^{*\prime}] = -E[(\underline{w}_n^o - \overline{\underline{w}}^o)(\underline{w}_n^o - \overline{\underline{w}}^o)^{*\prime}] = \tfrac{1}{2}\underline{\Lambda}. \quad (4.5)$$

B. Calculation of Lag Misadjustment

The same time invariant transformation that was used in Section II can be used here provided we redefine the weight error vector as

$$\underline{\tilde{w}}_k = \underline{w}_k - \underline{w}_n^o. \quad (4.6)$$

For $k = 0, 1, 2, \cdots, T-2$,

$$\underline{w}_{k+1} - \underline{w}_0^o = \underline{w}_k - \underline{w}_0^o + \mu\underline{x}_k v_k - \mu\underline{x}_k\underline{x}_k^{*\prime}(\underline{w}_k - \underline{w}_0^o)$$
$$\Rightarrow \quad \underline{\tilde{w}}_{k+1} = (I - \mu\underline{x}_k\underline{x}_k^{*\prime})\underline{\tilde{w}}_k + \mu\underline{x}_k v_k \quad (4.7)$$

which can be transformed into a time invariant equation in $\underline{\tilde{f}} = D^k\underline{\tilde{w}}_k$, namely

$$\underline{\tilde{f}}_{k+1} = H\underline{\tilde{f}}_k \pm \mu D\underline{x}_0 v_k. \quad (4.8)$$

This is the same as (2.10) with the constant matrix H defined in (2.9). However, for $k = T-1$,

$$\underline{w}_T - \underline{w}_1^o = \underline{w}_{T-1} - \underline{w}_0^o + \mu\underline{x}_{T-1}v_{T-1} - \mu\underline{x}_{T-1}\underline{x}_{T-1}^{*\prime}(\underline{w}_{T-1} - \underline{w}_0^o) - (\underline{w}_1^o - \underline{w}_0^o)$$
$$\Rightarrow \quad \underline{\tilde{w}}_T = (I - \mu\underline{x}_{T-1}\underline{x}_{T-1}^{*\prime})\underline{\tilde{w}}_{T-1} + \mu\underline{x}_{T-1}v_{T-1} - \underline{\delta}_0 \quad (4.9)$$

which can be transformed into (note that $D^T = I$ by (2.5))

$$\begin{aligned}\underline{\tilde{f}}_T &= H\underline{\tilde{f}}_{T-1} + \mu D\underline{x}_0 v_{T-1} - \underline{\delta}_0 \\ &= \mu \sum_{i=0}^{T-1} H^{T-1-i} D\underline{x}_0 v_i - \underline{\delta}_0 + H^T \underline{\tilde{f}}_0. \end{aligned} \quad (4.10)$$

Extending the argument to $k = nT + p(0 \le p \le T)$, we have

$$\begin{aligned}\underline{\tilde{f}}_k &= \mu \sum_{i=0}^{k-1} H^{k-1-i} D\underline{x}_0 v_i \\ &\quad - \sum_{j=0}^{n-1} H^{(n-1-j)T+p}\underline{\delta}_j + H^k \underline{\tilde{f}})0. \end{aligned} \quad (4.11)$$

Since $H^k \underline{\tilde{f}}_0$ vanishes at an exponential rate, it can be ignored in the steady state (i.e., k sufficiently large) analysis. So, for large k, we have

$$\begin{aligned}\underline{\tilde{f}}_k &= \mu \sum_{i=0}^{k-1} H_{k-1-i} D\underline{x}_0 v_i \\ &\quad - \sum_{j=0}^{n-1} H^{(n-1-j)T+p}\underline{\delta}_j = \underline{\tilde{f}}_k^1 + \underline{\tilde{f}}_k^2. \end{aligned} \quad (4.12)$$

The steady state MSE can now be obtained using (3.4). Note that the expression for $\underline{\tilde{f}}_k$ in (4.12) consists of two components. The first is due to the additive noise alone and is identical to (A.2) in Appendix A. The second is due purely to the parameter variation. Since the two are independent of each other, the expectation of the cross-product terms will vanish and so the MSE is simply the sum of the *noise misadjustment* term which is given in (3.5), and the *lag misadjustment*, which is the contribution of the parameter variation to the MSE. The total MSE is

$$\begin{aligned}\mathcal{E} &= \lim_{k\to\infty} E_{\text{cyc}}[||\underline{x}_0' \underline{\tilde{f}}_k||^2] \\ &= \lim_{k\to\infty} E_{\text{cyc}}||\underline{x}_0' \underline{\tilde{f}}_k^1||^2 + \lim_{k\to\infty} E_{\text{cyc}}||\underline{x}_0' \underline{\tilde{f}}_k^2||^2 \\ &= M\,(\text{as before}) + L \end{aligned} \quad (4.13)$$

where M denotes the noise misadjustment and L the excess lag. From (4.12),

$$\begin{aligned}L &= \lim_{n\to\infty} E_{\text{cyc}}\left[\underline{x}_0'\left(\sum_{r=0}^{n-1} H^{(n-1-r)T+p}\underline{\delta}_r \right.\right. \\ &\quad \left.\left. \cdot \sum_{s=0}^{n-1} \underline{\delta}_s^{*\prime}(H^{(n-1-s)T+p})^{*\prime}\right)\underline{x}_0\right] \\ &= \underline{x}_0' H^p \lim_{n\to\infty} E_{\text{cyc}}\left[\sum_{r=0}^{n-1}\sum_{s=0}^{n-1} H^{(n-1-r)T}\underline{\delta}_r \right. \\ &\quad \left. \cdot \underline{\delta}_s^{*\prime}(H^{(n-1-s)T})^{*\prime}\right](H^p)^{*\prime}\underline{x}_0. \end{aligned} \quad (4.14)$$

In Appendix D, it is shown that

$$L = \underline{x}_0' H^p A (H^p)^{*\prime} \underline{x}_0 \quad (4.15)$$

where

$$A = S - \tfrac{1}{2}H^T S - \tfrac{1}{2}S(H^T)^{*\prime}$$

with S being the solution of the discrete time Lyapunov equation

$$S - H^T S (H^T)^{*\prime} = \underline{\Lambda}$$

To further explore the relationship between L and μ, we consider the case when μ is small so that

$$H \approx (1-\mu)D. \quad (4.16)$$

With detailed steps in Appendix D, we can obtain the following expression for the excess lag at $k = nT + p$

$$L \approx N(1-\mu)^{2p}\sigma_\delta^2 \quad (4.17)$$

where σ_δ^2 is the variance of the identically distributed independent weight vector components.

Equation (4.17) conforms with one's intuition about the adaptive estimation of the time variant parameters. Since the parameters only change at the beginning of each cycle, the longer the adaptive process is undertaken within that cycle (indexed by p), the better the quality of the estimation will be and thus the smaller the excess lag will be. As in the case of adaptive estimation of the time invariant parameters, the larger the adaptive gain μ is, the faster the adaptation will be and thus the smaller the estimation error (excess lag here) will be. Finally, the excess lag is proportional to the variance of the parameter variation; the larger the "jump" of the true weight vector from one cycle to the other is, the bigger the excess lag (estimation error) becomes.

V. The Nonuniform Fourier Linear Combiner

In Section III, we had demonstrated how the FLC estimator can be approximated by a decoupled system for sufficiently low values of the step size. In the event of a transient, the convergence of each frequency component of the system is virtually independent of that of the others. Now it is almost always the case that some components of the signal are "stronger" than others in the sense that they make a larger contribution to the signal energy. It makes sense therefore that the stronger components should be made to converge faster and the weaker ones slower in such a manner as to maintain the steady state MSE at a constant level. This would enable an abrupt transient to be picked up more rapidly without sacrificing the steady state accuracy. This strategy of variable convergence leads us to introduce the nonuniform Fourier linear combiner (NFLC). The price that must be paid is the longer time needed to attain steady state, owing to the relatively slow convergence of the "weak" components.

A. Performance Analysis of the NFLC Estimator

The FLC estimator is modified by replacing the LMS algorithm (1.3) with the nonuniform LMS (NULMS) algorithm

$$\underline{w}_{k+1} = \underline{w}_k + \underline{\mu}\underline{x}_k e_k \quad (5.1)$$

where $\underline{\mu}$ is a diagonal matrix with $\underline{\mu}(q,q) = \mu_q (q = \pm 1, \pm 2, \cdots, \pm N)$. It is convenient to define another diagonal matrix U as

$$\underline{\mu} = \mu U \quad (5.2)$$

where μ is the step size for uniform convergence and $u_q = U(q,q)$ is a number that is greater than unity for faster converging components and less than unity for slower ones. The NULMS is similar to an algorithm that was developed by Harris *et al.* [5], who presented some empirical results.

The time invariant transformation holds for the NFLC too, but now the time invariant transition matrix is

$$H = D(I - \underline{\mu x}_0 \underline{x}_0'). \tag{5.3}$$

Then, the recursion equation of (2.10) becomes

$$\begin{aligned} \underline{\tilde{f}}_{k+1} &= H\underline{\tilde{f}}_k + \underline{\mu} D\underline{x}_0 v_k \\ &= H\underline{\tilde{f}}_k + \mu U D\underline{x}_0 v_k. \end{aligned} \tag{5.4}$$

In fact (5.4) describes the more general case whereas (2.10) is the special case of uniform convergence with $U = I$. The analysis for the steady state MSE follows along exactly the same lines and needs no repetition. The corresponding expression for the MSE is

$$\mathcal{E} = \frac{\mu^2}{2\pi} \int_{-\pi}^{\pi} S_v(\omega) H(\omega)\, d\omega \tag{5.5}$$

where

$$H(\omega) = ||\underline{x}_0'(H - Ie^{j\omega})^{-1} U D \underline{x}_0||^2. \tag{5.6}$$

Note that when $U = I$, (5.5) returns to (3.5), the MSE of the uniform convergence case.

The expression for the lag misadjustment is the same as that found in (4.15). The only difference, of course, is that H is now defined as in (5.3).

Of a greater practical interest is the decoupled linear approximation for the NFLC, especially a formula like (3.13), based on which we can calculate the step sizes μ_q to achieve a given steady state misadjustment M. To calculate M, the following heuristic argument is proposed. First, we show that in the case of the uniform FLC, the linear approximation of (3.13) can be obtained by computing

$$M = \lim_{k\to\infty} \sum_{q=\pm 1}^{\pm N} E_{\text{cyc}} |\tilde{f}_{k,q}|^2 \tag{5.7}$$

where $\tilde{f}_{k,q}$ is the solution to the univariate difference equation

$$\tilde{f}_{k+1,q} = (1-\mu)\tilde{f}_{k,q} + \mu x_{k,q} v_k, \qquad x_{k,q} = \sqrt{2N} e^{jq\omega_q k} \tag{5.8}$$

and ω_q is chosen so that

$$\begin{aligned} &\cos\left((i-j)\omega_q q\right) \\ &\quad = \frac{1}{N} cos(\omega_0(j+1)q) \sum_{p=1}^{N} \cos\left(\omega_0(i+1)p\right). \end{aligned} \tag{5.9}$$

This is equivalent to [see (C.8) for definition of $\kappa(i,j)$]

$$\sum_{q=\pm 1}^{\pm N} x_{k-1-i,q} x_{k-1-j,q}^* = \kappa(i, j-i). \tag{5.10}$$

Note the similarity between (5.8) and (3.10). Expanding (5.8) and neglecting the exponentially decaying term $(1-\mu)^k f_{0,q}$,

$$f_{k,q} = \mu \sum_{i=0}^{k-1} (1-\mu)^i x_{k-1-i,q} v_{k-1-i}. \tag{5.11}$$

From (5.7) and (5.10)

$$\begin{aligned} M &= \lim_{k\to\infty} \mu^2 \sum_{q=\pm 1}^{\pm N} E_{\text{cyc}} \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} (1-\mu)^i x_{k-1-i,q} \\ &\quad \cdot v_{k-1-i} (1-\mu)^j x_{k-1-j,q}^* v_{k-1-j}^* \\ &= \lim_{k\to\infty} \mu^2 \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} (1-\mu)^{i+j} K_v(i-j) \\ &\quad \cdot \sum_{q=\pm 1}^{\pm N} x_{k-1-i,q} x_{k-1-j,q}^* \qquad (5.12) \\ &= \mu^2 \lim_{k\to\infty} \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} (1-\mu)^{i+j} K_v(i-j)\kappa(i, j-i). \end{aligned}$$

(5.13)

Repeating the argument of Appendix C, we get (3.13).

Next, we can proceed as follows: Substituting $x_{k,q}$ in (5.8) into (5.12), we get

$$\begin{aligned} M &= 4\mu^2 N \sum_{q=1}^{N} \lim_{k\to\infty} \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} (1-\mu)^{i+j} K_v(i-j) \\ &\quad \cdot (q\omega_q(i-j)) \\ &= 4\mu^2 N \sum_{q=1}^{N} \lim_{k\to\infty} \left[K_v(0) \sum_{i=0}^{k-1} (1-\mu)^{2i} \right. \\ &\quad \left. + 2\sum_{m=1}^{p} K_v(m) \cos(m\omega_q q) \sum_{i=0}^{k-m-1} (1-\mu)^{2i+m} \right] \\ &= 2\mu N \sum_{q=1}^{N} \left(K_v(0) + 2\sum_{m=1}^{p} K_v(m) \cos(m\omega_q q) \right) \\ &\quad + O(\mu^2). \end{aligned} \tag{5.14}$$

The whole point of this argument is that the linear approximation enables one to decompose the misadjustment into N components, each of which can be obtained by solving a scalar difference equation of form (5.8). Nonuniformity is now introduced merely by replacing μ in (5.8) by μ_q. The linear approximation now becomes

$$\begin{aligned} M &\approx 2N \sum_{q=1}^{N} \mu_q \left(K_v(0) + 2\sum_{m=1}^{p} K_v(m) \cos(m\omega_q q) \right) \\ &\quad + O(\mu^2). \end{aligned} \tag{5.15}$$

We now see that (3.13) is but a special case of (5.15). In the former, every μ_q is set to μ, the uniform step size. It is easy to solve the linear part of (5.15), since the ratios of the step sizes to one another can be determined from the relative estimated strengths of the spectral components of the signal. The "strength" of a component may be defined differently in different applications. It could be defined as the mean energy

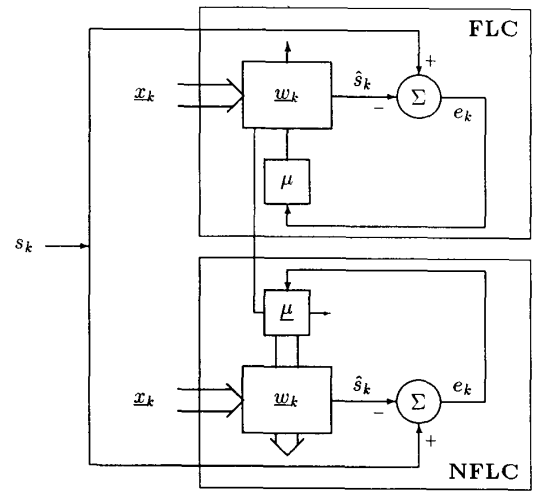


Fig. 3. Block diagram showing the NFLC implementation. A uniform FLC is required to provide a *posteriori* information about the signal spectrum in order to update the step size matrix of the NFLC.

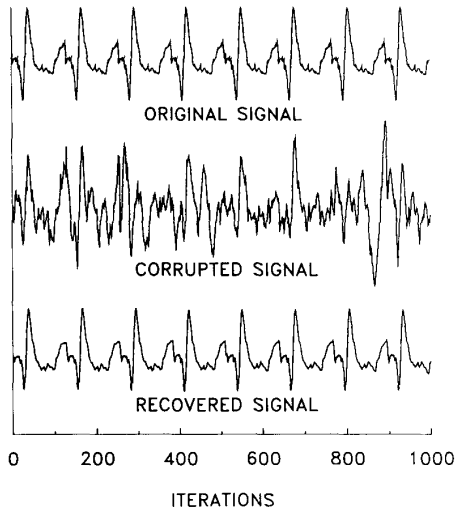


Fig. 4. Application of the adaptive FLC technique to estimate brain evoked responses of an anesthetized cat. The responses correspond to successive stimuli delivered at the rate of 5.9 Hz. The top trace shows eight consecutive 40 ms raw responses concatenated together. The center trace is the noisy response obtained during physiological experiment (SNR = -7.10 dB). The bottom trace displays the results of adaptive FLC (SNR = 40.09 dB). These responses have been extracted from the middle of a data set, when the estimator has already attained steady state.

or its square root or any other function of the spectrum. Thus, in the absence of any parameter variations, one can solve (5.15) for the step sizes to achieve the desired steady state misadjustment with desired convergence properties.

B. Practical Implementation of the NFLC

The scheme for implementing the NFLC estimator is illustrated in Fig. 3. Notice that the configuration includes both the uniform and the nonuniform FLC estimators. The reason for including the uniform FLC is to provide a reliable estimate of the strengths of the spectral components of the signal. The weight vector of the NFLC should not be used because the estimates of the relatively weaker components adapt more slowly. Furthermore, if the signal becomes very small or disappears altogether, using the NFLC weight vector to adjust the step sizes could easily lead to instability. The NFLC performance is the most efficient when there is a sharp drop (or rise) in the amplitude of the underlying signal. The performance of the NFLC is somewhat compromised when there is a drastic and abrupt change in signal spectrum. This is due to that formerly weak components, which have a large time constant of adaptation, may become more prominent following the transient.

The weight estimates of the uniform FLC give us the relative strength of the spectrum components of the signal s_k. Thus, the ratio between $\mu_q[u_q$ in (5.2)] can be determined. Substituting μ_q in (5.15) with μu_q, one can obtain easily μ (therefore μ_q) for any desired M given, if the estimates of $K_v(i)$ are available. This completes the design of the NFLC algorithm.

VI. Discussion

The Fourier linear combiner is particularly well-suited for estimating periodic or recurrent signals with low SNR. Such signals may include stimulus-response type data, where successive responses to applied stimuli are concatenated together to form a periodic signal. An example of such a situation in clinical monitoring is to estimate brain evoked responses to successively applied stimuli. The noise consists mainly of background brain activity not related to the stimulus, and often has no periodic component in the bandwidth of recording. Fig. 4 illustrates the effectiveness of the adaptive FLC. The top trace consists of eight successive "ideal" brain responses to electrical pulses delivered to a peripheral nerve of an anesthetized cat, each 40 msec in length. They are the result of concatenation of the processed evoked potentials, which in turn are obtained through synchronized averaging. The center trace is the corrupted response (with background EEG noise) obtained directly from the experiment (SNR = -7.10 dB), the bottom trace consists of the FLC estimates of the evoked potential responses. The FLC recovers the periodic signal successfully and the SNR of this recovered signal is 40.09 dB. The order of the adaptive FLC is set to $N = 32$ to cover $> 99\%$ of the signal's energy. The adaptive gain is set to $\mu = 0.0007$.

Fig. 5 demonstrates how the weights themselves (i.e., the estimated Fourier coefficients) can provide useful spectral information. Data obtained by our clinical collaborators during physiological investigations of the brain function were analyzed. An anesthetized animal was made to inhale first 100% oxygen, next 10% oxygen gas mixture (which resulted in hypoxia). The time trends of the mean energy of the fundamental, second, third, and fourth harmonics are shown in Fig. 5. While this hypoxic insult resulted in brain death signified by a flattening of the evoked response, each harmonic responds differently to the disturbance. While the third and fourth harmonics decline steadily a few minutes after the disturbance is initiated, there is an enhancement of the fundamental and the second harmonic before their declination.

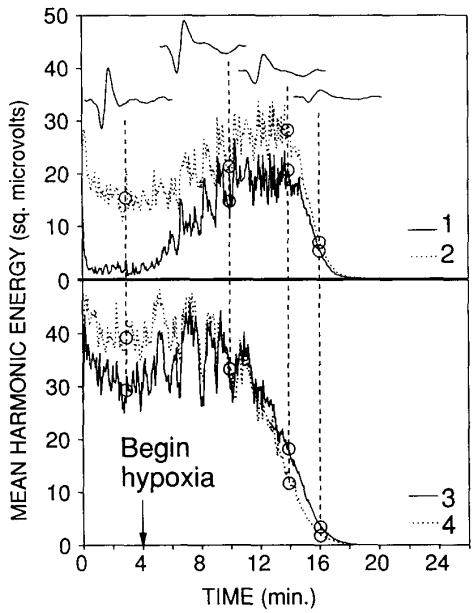


Fig. 5. FLC-estimated harmonic trends in an experiment. An anesthetized cat is made hypoxic. At the point in time indicated, the gas inspired by the animal is changed from 100% oxygen to 10% oxygen, 90% nitrogen. Time trends are shown for the mean energy in the first four components of the Fourier series, fundamental and second harmonic at the top, third and fourth harmonics at the bottom. Sample estimated response waveforms are also shown: the first corresponds to the control period, and the other three correspond to various stages in brain function deterioration.

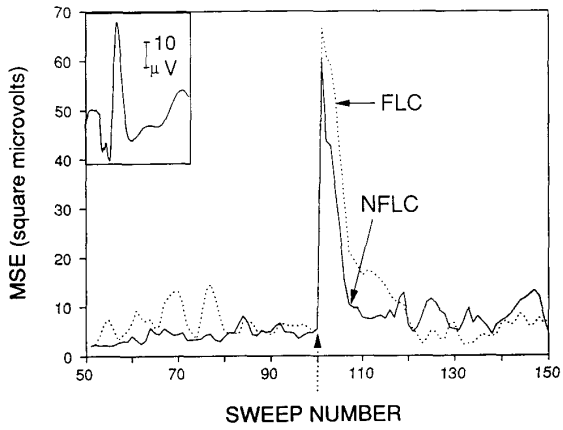


Fig. 6. Comparison of the NFLC and the FLC algorithms for detecting abrupt changes of the signal parameters. MSE trend plots for FLC and NFLC with simulated data. The signal used is shown at the top-left corner. At the point indicated by the arrow, the signal amplitude is abruptly reduced by 80%. Initial SNR is −3 dB.

TABLE II
COMPARISON OF THE MSE FROM THEORETICAL ANALYSIS AND COMPUTER SIMULATION. THE MSE IS NORMALIZED BY THE SIGNAL POWER.

μ	0.01	0.001	0.0001	0.00001
order of FLC covers 99 % of signal energy				
theory	0.2349	0.0430	0.0058	0.0006
simulation	0.3000	0.0526	0.0136	0.0120
order of FLC covers 100 % of signal energy				
theory	0.1690	0.0127	0.0015	0.0002
simulation	0.1771	0.0132	0.0015	0.0002

The enhancement of the fundamental component is particularly striking, and is clinically significant [11]. The adaptive Fourier linear combiner has its parameters as follows: $\mu = 0.007$ and $N = 32$.

The computational complexity of the FLC algorithm (as measured by the number of real multiplications) is 4N per input sample. The NFLC algorithm (coupled with a FLC as shown in Fig. 3) has computational complexity of $6N$ per input sample. This 50% increase in computation complexity is justified by the fact that the performance of the NFLC algorithm for detecting abrupt changes in parameter variation is better than the FLC algorithm alone, as evidenced in Fig. 6. (At the point indicated by the arrow, the signal amplitude is reduced by 80%.) As one can see, the MSE of the NFLC decreases faster than the FLC algorithm does. This fast reduction in the MSE signifies a faster tracking of the abrupt changes by the NFLC algorithm (due to the faster adaptation of the stronger components). As stated in Section V, the NFLC algorithm takes longer to reach the steady state (due to the slower adaptation of the weaker components). At steady state (not shown in Fig. 6), both algorithms have the same misadjustment as designed.

The analysis of misadjustment that has been presented here should be useful for assessing the effectiveness of the FLC in any given situation, and in many cases, even designing the estimator (i.e., choosing the appropriate step size) to attain desired performance characteristics. Table II summarizes the MSE (normalized by the signal power) obtained from simulation and the theoretical analysis (3.13). The true signal used is the same as the one described in Fig. 4, the additive noise is an AR (2) process with SNR = 0 dB. The model order of the FLC is 32 and $p = 5$ in (3.13). We see that the predicted MSE at steady state by (3.13) is in good agreement with the simulation results. The discrepancy between them can be attributed to that in theory perfect modeling of the desired signal by FLC is assumed and in simulation the order of FLC is chosen to cover 99% of the signal energy. The bottom part of the table shows the result where the signal is generated artificially with only first eight harmonics and the adaptive FLC order is also eight to allow perfect modeling. We observe a better agreement of the theoretical and simulation results.

An interesting aspect of the FLC is the decoupling that occurs when the step size becomes small. In this case, the relationship between noise misadjustment and step size is virtually linear. Furthermore, the transition matrix gets more diagonalized and the system tends to converge uniformly. This result is also apparent when one considers the "autocorrelation" matrix defined by the cyclic moment $E_{\text{cyc}}[\underline{x}\underline{x}^{*\prime}]$, which is none other than the identity matrix. According to [15], the time constants of convergence of the various modes of the system depend on the eigenvalues of this correlation matrix, and on the step size. Since the correlation matrix is the identity matrix, all eigenvalues will be unity, and one would expect

uniform convergence, thus overcoming one major drawback of LMS. However, since we are considering cyclic moments, it follows that this would be true for time constants that are larger than one period of the signal. As we saw earlier, one could take advantage of the decoupling phenomenon to set up an estimator with a selective nonuniform convergence, based on a posteriori estimate of the signal spectrum (i.e., the Fourier coefficients).

Our analyses are based on the assumption that the period of the signal is known, it would be interesting to extend our analysis to cover the case of simultaneously estimating the period of the signal as well. We hope to address this problem in future.

VII. Summary

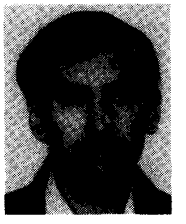
An LMS-based Fourier series estimator for periodic data with both time invariant and time varying parameters has been studied in this paper. The analysis includes a derivation of the steady-state noise misadjustment as well as lag misadjustment for a particular kind of parameter variation model. System decoupling and uniform convergence occur when the LMS step size is small. However, the selective nonuniformity in the step size, based on a posteriori estimate of the Fourier coefficients, can greatly enhance the detection of abrupt transients.

Acknowledgment

The authors gratefully acknowledge the assistance given by D. Hanley, M.D., of the Neurology Department of the Johns Hopkins Hospital in furnishing data to demonstrate application of this work. They are also grateful to one anonymous reviewer for suggesting a simpler time independent transformation.

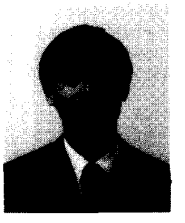
References

[1] P. Baudrenghien, "The adaptive spectrum analyzer," Ph.D. dissertation, Stanford Univ., Stanford, CA, 1984.
[2] R. R. Bitmead and B. D. O. Anderson, "Performance of adaptive estimation algorithms in dependent random environments," *IEEE Trans. Automat. Contr.*, vol. AC–25, pp. 788–794, 1980.
[3] R. R. Bitmead, "Convergence in distribution of LMS-type adaptive parameter estimates," *IEEE Trans. Automat. Contr.*, vol. AC–28, pp. 54–60, 1983.
[4] A. Feuer and E. Weinstein, "Convergence analysis of LMS filters with uncorrelated Gaussian data," *IEEE Trans. Acoust. Speech Signal Processing*, vol. ASSP-33, pp. 222–229, 1985.
[5] R. W. Harris, D. M. Chabries, and F. A. Bishop, "A variable step (VS) adaptive filter algorithm," *IEEE Trans. Acoust. Speech Signal Processing*, vol. ASSP-34, pp. 309–316, 1986.
[6] O. Macchi and E. Eweda, "Second-order convergence analysis of stochastic adaptive linear filtering," *IEEE Trans. Automat. Contr.*, vol. AC–28, pp. 76–85, 1983.
[7] S. Narayan and A. Peterson, "Frequency domain least-mean-square algorithm," *Proc. IEEE*, vol. 69, pp. 124–126, 1981.
[8] K. H. Shi and F. Kozin, "On almost sure convergence of adaptive algorithms," *IEEE Trans. Automat. Contr.*, vol. AC-31, pp. 471–474, 1986.
[9] V. Solo, "The limiting behavior of LMS," *IEEE Trans. Acoust. Speech Signal Processing*, vol. 37, pp. 1909–1922, 1989.
[10] C. A. Vaz and N. V. Thakor, "Adaptive fourier estimation of time-varying evoked potentials," *IEEE Trans. Biomed. Eng.*, vol. 36, pp. 448–455, 1989.
[11] C. A. Vaz, "Adaptive fourier estimation of time-varying evoked potentials," Ph.D. dissertation, Johns Hopkins Univ., Baltimore, MD, 1990.
[12] B. Widrow and M. E. Hoff, Jr., "Adaptive switching circuits," in *IRE WESCON Conv. Rec.*, 1960, pp. 96–104, pt. 4.
[13] B. Widrow *et al.*, "Adaptive noise canceling: Principles and applications," *Proc. IEEE*, vol. 63, pp. 1692–1716, 1975.
[14] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, Jr., "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, pp. 1151–1162, 1976.
[15] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.
[16] B. Widrow, P. Baudrenghien, M. Vetterli, and P. Titchener, "Fundamental relations between the LMS algorithm and the DFT," *IEEE Trans. Circuits Syst.*, vol. CAS-34, pp. 814–820, 1987.
[17] B. Widrow, J. McCool, and M. Ball, "The complex LMS algorithm," *Proc. IEEE*, vol. 63, pp. 719–720, 1975.



Christopher Vaz was born in Pune, India in 1962. He received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Bombay.

He is currently working on his doctoral dissertation in biomedical engineering at The Johns Hopkins University in Baltimore, MD. His research interests include signal processing and its applications to biosignal analysis.



Xuan Kong (M'91) received the B.S. degree from Sichuan University in 1984, the M.S. degree from University of Manitoba, Winnepeg, Canada, in 1986, and the M.S.E. and the Ph.D. degrees from Johns Hopkins University, Baltimore, MD, in 1989 and 1991, respectively, all in electrical engineering.

He is currently an Assistant Professor in the Department of Electrical Enginering at Northern Illinois University. His research interests include adaptive signal processing, biomedical signal processing, morphological image processing, and audio quality evaluation in telecommunication systems. He is the co-author (with V. Solo) of *Adaptive Signal Processing Algorithms: Stability and Performance* (Prentice-Hall, 1994).

Dr. Kong is a member of INNS.

Nitish V. Thakor (S'78–M'81–SM'89) received the bachelor's degree in electrical engineering from the Indian Institute of Technology, Bombay, India, in 1974, and the Ph.D. degree in electrical and computer engineering from the University of Wisconsin, Madison, in 1981.

He previously served on the faculty of the Northwestern University, and is currently on the faculty of the Biomedical Engineering Department of the Johns Hopkins School of Medicine, Baltimore, MD. He also serves as an Associate Editor for the IEEE Transactions on Biomedical Engineering, Journal of Ambulatory Monitoring, and Journal of Biological Systems, and is also the editor of an upcoming text-book on Biomedical Signal Processing: Theory and Applications. He is engaged in international collaborative research and educational activities supported in part by a Fulbright Award and NATO. His research interests are cardiovascular and neurological signal processing, instrumentation, and large-scale computer applications. He currently teaches courses on related subjects and serves as a consultant to industries in these specialties.

Dr. Thakor is a previous recipient of a Research Career Development Award from the National Institutes of Health and a Presidential Young Investigator Award from the National Science Foundation.