Adaptive Fourier Estimation of Time-Varying Evoked Potentials

CHRISTOPHER A. VAZ, STUDENT MEMBER, IEEE AND NITISH V. THAKOR, MEMBER, IEEE

Abstract—**An estimation procedure for dealing with time-varying evoked potentials is presented here. The evoked response is modeled as a dynamic Fourier series and the Fourier coefficients are estimated adaptively by the least mean square algorithm. Approximate expressions have been developed for the estimation error and time constant of adaptation. A procedure for optimizing the estimator performance is also presented. The effectiveness of the estimator in dealing with simulated as well as actual evoked responses is demonstrated.**

I. INTRODUCTION

EVOKED potentials (EP's) are electrical responses of the central nervous system to sensory stimuli applied in a controlled manner. Signal-to-noise ratios (SNR's) may be as low as 0 to -10 dB. Several methods have been used so far to try to improve the SNR. Some good overviews of EP estimation tools are to be found in [1]. The purpose of this paper is to devise an efficient means of estimating EP's that are subject to pathological changes. Early detection of such changes is crucial in such environments as a neurological critical care unit or an operating room [2]. For example, McPherson *et al.* [13] have reported transient changes in EP's during surgery with certain anesthetic agents (see also Section IV). For this purpose, the EP is modeled as being essentially invariant unless or until a change in physiological state occurs. In other words, it is *partly deterministic, partly stochastic.*

Most EP estimation methods can be broadly classified into two categories. The first category consists of techniques that regard the EP as a set of samples and attempt to estimate each sample individually instead of estimating the EP as a coherent waveform. Of all these, the most widely used is signal averaging. Variations of the common summing averager include weighted averaging (giving larger weights to recent data), latency-corrected averaging (adjusting for shifts), and selective averaging (considering only relatively high SNR waveforms). Also included are certain adaptive averaging schemes such as the time-sequenced filter [3], [4] which is an adaptation of the least-mean square (LMS) algorithm [5] for recurrent nonstationary signals. All these methods use a cyclostationary model of the signal which takes advantage of its recurrence property. These could be useful in tracking changes, but do so slowly because the estimates are updated just once every trial.

The second category consists of filters that use the spectral properties of the band-limited signal. This information is only available *a posteriori* and is generally obtained from an averaged waveform. *A posteriori* Wiener filtering (APWF) was one of the earliest techniques to be used for EP processing [6]. There has been a controversy over APWF [1] with its rather unrealistic model of the EP as a stationary random process. To overcome this limitation, time-varying filters were developed both in the frequency domain [7] as well as in the time-domain [8]. These exhibit superior performance. Their ability to follow transients is unpredictable because that depends on how accurately and how quickly the *a posteriori* estimated statistics of signal and noise reflect the transient. The mean-square error (MSE) is a fixed function of these statistics and cannot be controlled (except by averaging prior to filtering). These filters are useful for studying trial-to-trial variability in normal subjects or in improving the SNR of an averaged waveform.

An efficient estimator would exploit both the recurrence property (as in the first category methods) and the spectral properties (as in the second category methods). An interesting technique mentioned in the literature uses a two-dimensional discrete Fourier transform [9]. Its concept of spectrum variation across trials is similar to ours and it is potentially useful for following transients. But its performance has not been analytically described. Consequently, proper guidelines for the choice of parameters is lacking.

A Fourier series model for the EP is presented here. The model parameters are estimated by an adaptive process, enabling the estimator to efficiently track clinically significant alterations in the EP. Performance and design of the estimator are also discussed.

II. THE FOURIER LINEAR COMBINER

The signal model regards the noise-free evoked responses as comprising a quasi-periodic signal of the form

$$s_{k+T} \sim s_k \tag{1}$$

where k is the discrete time index and T is the fixed recurring interval or number of samples per trial. The relation is not a perfect equality, representing the partially stochastic nature of the signal. Such a signal can be rep-

Manuscript received March 21, 1988; revised October 7, 1988. This work was supported by Grant NS 24282, a Research Career Development Award 5K04HL01509-03 from the NIH, and a Presidential Young Investigator Award from the NSF to N. V. Thakor.

The authors are with the Department of Biomedical Engineering, The Johns Hopkins Medical School, Baltimore, MD 21205.

IEEE Log Number 8825629.

0018-9294/89/0400-0448$01.00 © 1989 IEEE

resented by a discrete dynamic Fourier series [10], of the form

$$s_k = \sum_{r=1}^{T/2-1} a_r \sin(2\pi rk/T) + b_r \cos(2\pi rk/T). \quad (2)$$

Note that ac-coupled amplification is implied in the omission of the dc term. This representation is illustrated in Fig. 1.

The series is dynamic because the Fourier coefficients in general are time-varying. An adaptive algorithm is used to estimate the coefficients. The adaptive Fourier linear combiner (FLC) is shown schematically in Fig. 2. The uncorrupted EP is denoted by s and the background EEG noise by v. d is the observed data. The output $\hat{d}$ is the estimate of the signal while e_k is the error output. The input vector $\boldsymbol{x}_k$ consists of the sinusoidal components in (2). The weight vector $\boldsymbol{w}_k$ is updated at every iteration according to the LMS algorithm of Widrow and Hoff [5]

$$\boldsymbol{w}_{k+1} = \boldsymbol{w}_k + \mu(-\partial e_k^2/\partial \boldsymbol{w}_k) = \boldsymbol{w}_k + 2\mu \boldsymbol{x}_k e_k \quad (3)$$

where μ is a user-chosen step-size parameter which determines the gain of the algorithm. For those who are unfamiliar with the LMS algorithm, a brief description is provided in Appendix A. For a more detailed discussion on convergence properties of the LMS algorithm, the reader is referred to [5] and [11].

Although the elements of the input vector $\boldsymbol{x}_k$ are deterministic functions they can always be modeled as stochastic, with moments computed over any one segment of length T. Since the functions are orthogonal, each having a mean power of $1/2$, the autocorrelation matrix of $\boldsymbol{x}$ is diagonal, i.e.,

$$\boldsymbol{R} = E[\boldsymbol{x}_k \boldsymbol{x}_k^T] = (1/2)\boldsymbol{I}. \quad (4)$$

Since $\boldsymbol{x}$ is a vector of dimension 2M (see Fig. 2), $\boldsymbol{R}$ will have 2M eigenvalues in general. In the absence of noise, the MSE $\epsilon_k = E[e_k^2]$ converges exponentially to zero, provided μ lies within certain bounds (see Appendix A). Corresponding to each eigenvalue of $\boldsymbol{R}$ is a distinct mode of the system and a time constant of convergence for each mode. The theoretical time constant is inversely proportional to the eigenvalue for that mode. In the FLC, each parameter (Fourier coefficient) being estimated corresponds to a separate mode because of orthogonality. In this case, all the eigenvalues of $\boldsymbol{R}$ are identical and equal to $1/2$, because the eigenvalues of a diagonal matrix are merely the diagonal elements themselves. According to [5], the ideal time constant for the relaxation of ϵ_k is given by $\tau_\epsilon = 1/4\mu\alpha$ where α is the corresponding eigenvalue. In the case of the FLC,

$$\tau_\epsilon = 1/(2\mu). \quad (5)$$

The actual time constant would be somewhat greater than this ideal value due to the inherent approximations in the LMS algorithm. The step size cannot be made arbitrarily large as this would lead to instability in the system [5], [2.]

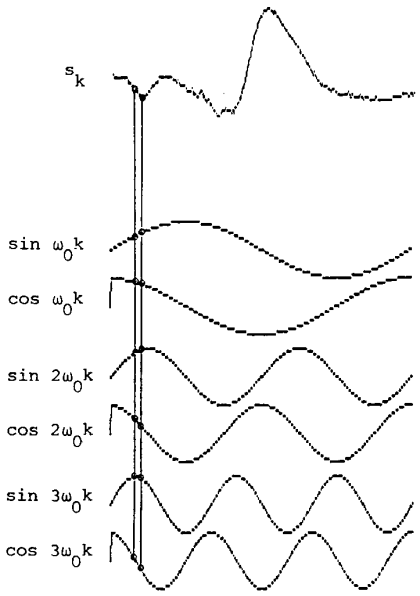


Fig. 1. Illustration of the Fourier linear combiner. The base frequency $\omega_0 = 2\pi/T$.

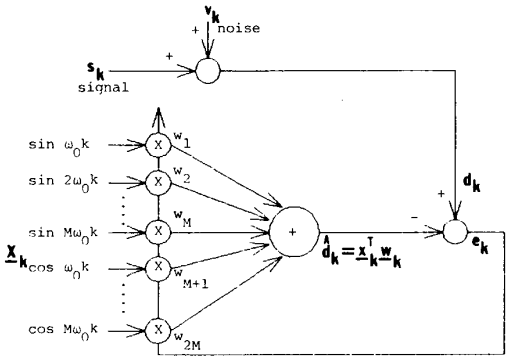


Fig. 2. The adaptive FLC estimator.

III. The Mean Square Error Function

A. Deriving the Error in Estimation

The series in (2) can usually be truncated to just a few terms without appreciable distortion because of the limited bandwidth of EP's. In fact, truncation may even reduce the estimation error. This is because the LMS algorithm, even after reaching steady state, produces an error over and above the minimum Wiener MSE when noise is present. This is an unavoidable consequence of the adaptation process [5]. Let the series in (2) be truncated to $r = M$. M is the *order* of the model. This section will be devoted to the performance analysis of the FLC estimator. An expression for the error of estimation will be derived. This will be the basis for a judicious choice of the model order and the step size.

The following definitions will be used:

$$\omega_0 = 2\pi/T.$$

The input basis vector:

$$\boldsymbol{x}_k = [x_{1,k} \quad x_{2,k} \cdots x_{2M,k}]^T \quad \text{where}$$

$$x_{r,k} = \begin{cases} \sin(r\omega_0 k), & r = 1, 2, \cdots, M \\ \cos[(r - M)\omega_0 k], & r = M + 1, \\ & M + 2, \cdots, 2M. \end{cases} \quad (6)$$

The solution vector of Fourier coefficients (stationary case):

$$\boldsymbol{w}^* = [a_1 \quad a_2 \cdots a_M \quad b_1 \quad b_2 \cdots b_M]^T.$$

The truncated Fourier series represents the approximated signal:

$$s_k' = \boldsymbol{x}_k^T \boldsymbol{w}^* = \sum_{r=1}^{M} a_r \sin(r\omega_0 k) + b_r \cos(r\omega_0 k). \quad (7)$$

Assuming stationarity for the moment, we are faced with the system identification problem

$$d_k = \boldsymbol{x}_k^T \boldsymbol{w}^* + z_k$$

where $\boldsymbol{w}^*$ is the unknown parameter vector defined above; $\boldsymbol{x}_k^T \boldsymbol{w}^*$ is the (approximated) unknown signal s_k'. The input d_k is the observed data while the noise v_k is unobservable; z_k is the sum of the noise and the residual signal, i.e.,

$$z_k = v_k + (s_k - s_k').$$

Our analysis is based upon the following assumptions:

1) $2\mu << 1$;

2) the noise v_k is zero-mean, stationary and uncorrelated with the signal as well as with every component of the Fourier series;

3) the autocovariance function of v is nonzero over only a finite interval, i.e.,

$$K_v(i) = E[v_k v_{k+i}] = 0, \qquad \text{for } i > p$$

where p is a positive integer;

4) the approximation residue $s - s'$ has a power that is small compared to the noise variance $K_v(0)$; as a result, it makes a negligible contribution to the steady-state error generated by the adaptive process (see below).

Based on these assumptions, it is possible to derive an approximate expression for the MSE of estimation in terms of the model order and the step size (the user-chosen quantities). The reader should bear in mind that the MSE here refers to the *steady-state error* after the statistical moments of the weight vector have converged. Implied in assumption 1) is the fact that μ is small enough for convergence to occur (refer again to [5] and [11]). Due to the fact that an orthogonal basis vector is being used, the MSE, $\xi = E[(s - \hat{d})^2]$ is minimized for $\boldsymbol{w} = \boldsymbol{w}^*$. Let $f(M)$ denote the approximation error due to truncation of the Fourier series to M terms. Then

$$f(M) = E[(s_k - s_k')^2] = \sum_{i=M+1}^{T/2-1} (a_i^2 + b_i^2)/2. \quad (8)$$

The steady-state MSE then consists of two components, i.e.,

$$\xi = \lim_{k\to\infty} E[(s_k - \hat{d}_k)^2] = \epsilon' + f(M) \quad (9a)$$

where

$$\epsilon' = \lim_{k\to\infty} E[(s_k' - \hat{d}_k)^2]. \quad (9b)$$

ϵ' is the steady-state error due to the adaptation process alone. Note that (9) is valid because of assumption 4), which neglects the correlation between the approximation residue and the estimator output $\hat{d}$. Under assumptions 1)–4) listed above, the following expression for ϵ' has been derived (see Appendix B for the derivation),

$$\epsilon' \sim \mu\left[2\sum_{i=1}^{p} (1 - \mu)^i K_v(i) \cdot \sum_{j=1}^{M} \cos(j\omega_0 i) + K_v(0)M\right]. \quad (10)$$

The total MSE can be calculated by substituting (10) into (9). $f(M)$, the error in approximation, decreases in a monotonic fashion with the model order. From (8)–(10), it is clear that ξ is a complex function of the model order M. For low values of M, the approximation error given by (8) is high; for higher values of M, it is quite possible that ϵ' may be large. It follows that there should be an optimal model order that minimizes ξ. Alternatively, for a fixed ξ, there should be an optimal order that maximizes μ.

B. Performance Evaluation

A simulation was carried out to test the accuracy of the approximation in (10). The signal was the null signal so that there was no approximation error present. The noise was generated by a first-order autoregressive process

$$v_k = av_{k-1} + n_k \quad (11)$$

where a is the AR parameter (determining the amount of color) and n_k is white noise with variance adjusted to give $K_v(0) = 1000$. A first-order estimator was used with $\mu = 0.01$ and $T = 100$. The FLC estimator was run on the data for different values of a. Each run consisted of 100 periods and the MSE was calculated by averaging over the last 90 periods. The theoretical MSE was also calculated using (10). It can be seen that the theoretical and experimental MSE plots in Fig. 3 are in close agreement. Note also how performance is degraded as the noise becomes increasingly colored.

Another simulation was carried out in which a triangular waveform (inset in Fig. 4) was used as the signal and noise was superimposed on it according to (11) with $a = 0.7$. The theoretical and observed MSE values for FLC estimation were computed as a function of the model order and are compared in Fig. 4. Theoretical values were calculated using (8)–(10). Notice that, for both plots, the estimation error is minimized by choosing an order of three. For lower orders, the MSE is largely due to the

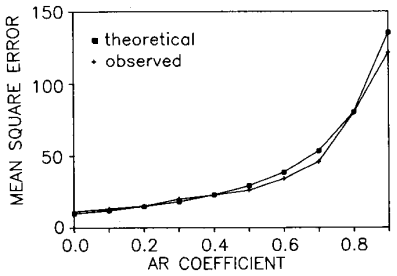


Fig. 3. A plot of FLC estimation error as a function of noise color [AR coefficient in (11)] in order to test the validity of (10).

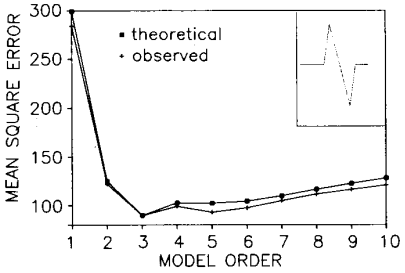


Fig. 4. A plot of FLC estimation error as a function of model order using a simulated signal (shown in inset) and simulated noise. The SNR is −4.1 dB. The step size is 0.004.

high approximation error [see (8)]; for higher orders the approximation error dies down but the error due to adaptation, ϵ' is the major contributor [see (10)].

In our studies, we have used the following procedure to determine a suitable step size for operating the FLC.

a) In general most of the signal energy lies within a limited bandwidth (this is illustrated later in Section V-B). Based on ensemble averages of trial runs or earlier recordings of the same kind of EP, a model order $M_{\max}$ is chosen which covers at least 99 percent of the signal energy. It is assumed that, for all practical purposes, approximation error is absent for $M = M_{\max}$.

b) A desired estimation error is chosen.

c) The estimator is initially run with order $M_{\max}$ and μ chosen by solving (10) with ϵ' as the desired MSE.

d) After the weights have reasonably settled (two or three time constants), (8)–(10) are used to solve for μ with ξ as the desired MSE. This is done for each order up to $M_{\max}$ and the optimal model order is the one that gives the highest μ. The estimator is then run with this optimal order and the corresponding step size.

In like manner, the above procedure can be modified so that, given a desired μ, one can arrive at the model order that minimizes ξ.

IV. Results

A. Simulated Data

A simulation was carried our using the waveform in Fig. 5(a) (in dark ink) as the known signal (derived from a previous EP recording) and independently derived physiological noise. The noise was obtained from a human volunteer by recording data from the scalp (position $C3$) without stimulation, in such a manner as to pick up electrical activity in the somatosensory cortex. The reference electrode was placed on the forehead. Fig. 5 illustrates the effect of varying the model order and the step size. In Fig. 5(a), typical FLC outputs corresponding to two different model orders are overlaid (in light ink) over the original waveform. The model order is four in Fig. 5(a) i) and nine in Fig. 5(a) ii). The theoretical MSE [as given by (9a)] and the time constant of MSE relaxation, τ_ϵ [given by (5)] are practically the same in the two cases. However in Fig. 5(a) i), it is apparent that the major contribution to the MSE is the approximation error, $f(M)$, resulting in a poorer reconstruction of the signal. On the other hand, the MSE in Fig. 5(a) ii) is largely due to ϵ', the deviations caused by the adaptive process itself. Fig. 5(b) illustrates

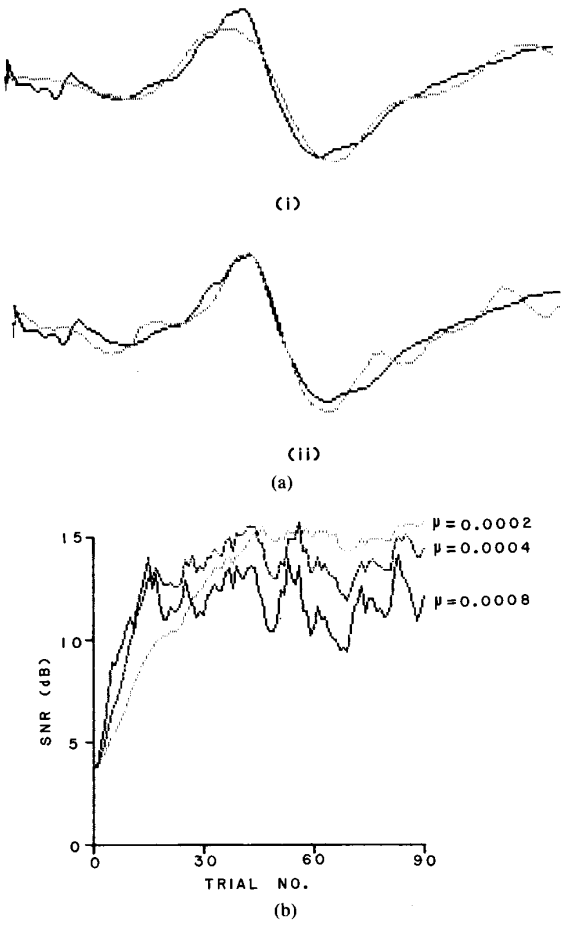


Fig. 5. Results of a simulation using a known signal and physiological noise. EP segments are 40 ms long. (a) Effect of varying model order. Typical FLC outputs are shown in light ink overlaid on the original signal for i) a model order of 4, and ii) for an order of 9. The theoretical MSE is identical in the two cases while the step sizes do not significantly differ [9.94×10^{-4} for i) and 9.36×10^{-4} for ii)]. (b) Effect of varying μ. Trend plots of the SNR of the estimates are shown for three different values of μ. The SNR here is defined as the ratio of the mean signal power to the MSE of the estimate.

how the choice of the step-size μ affects the estimator performance. A larger step size leads to faster adaptation [as one would expect from (5)], but a lower steady-state SNR.

The same data were used to simulate an EP transient. At a certain point in the simulation, the shape of the signal was abruptly changed. This was accomplished by altering the phase spectrum of the original signal. The SNR of the estimates is plotted in Fig. 6, and the underlying signals are also displayed at the top. In the figure, three different estimators are compared: a moving window averager (MWA), an exponentially weighted averager (EWA), and the FLC estimator. Note how each estimator responds to the change in the signal. In order to provide a fair comparison of tracking capability, all three estimators were designed to yield the same steady-state estimation error. For this estimation error, (8)–(10) were used to choose the FLC parameters (i.e., step size and model order), based on known signal and noise information so as to minimize τ_ϵ.

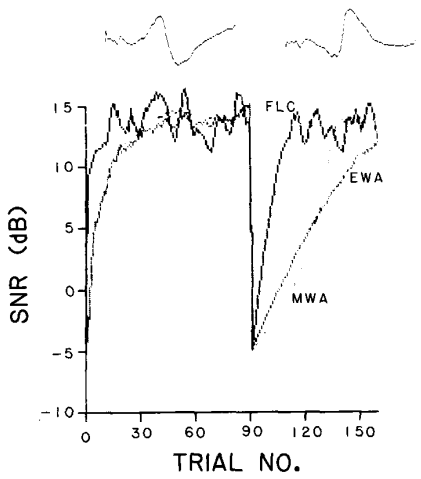


Fig. 6. Simulation of an EP transient with SNR trend plots displayed. At the 90th trial, the underlying signal was abruptly changed; the simulated signals are shown at the top of the figure. For the FLC, $M = 7$ and $\mu = 5.82 \times 10^{-4}$.

B. Physiological Data

In order to test the performance of the FLC on actual EP's, cortical somatosensory evoked potentials (SEP's) were recorded from a normal adult. The subject's left median nerve was stimulated transcutaneously by delivering current pulses of 200 μs duration through a pair of surface electrodes on the wrist. The recording electrode was at $C4$ and the reference at F_z. Initially, the stimulating current was set at 6.5 mA which was above the motor threshold, sufficient to cause a distinct thenar twitch. Pulses were delivered at a rate of 3.7 Hz. In the middle of the experiment, the current was abruptly lowered to 3.5 mA which was above the sensory threshold but below the motor threshold. The three estimators that were used on the simulated data were also used here; their performance is comparatively evaluated in Fig. 7. Instead of SEP amplitude, the root mean square amplitude of the estimated signal w.r.t. baseline (which is a more integrated measure of the entire evoked response) is plotted in Fig. 7. Also shown, for comparison, are the root mean square amplitude levels of independently computed averages corresponding to the two stimulus intensities. Once again the parameters of the three estimators were chosen to yield the same steady-state MSE. FLC parameters were chosen to minimize the time constant of MSE relaxation, based on signal and noise estimates from the first 100 trials. It is clear that the FLC surpasses the other methods in following the change in response elicited by the change in stimulus.

In order to demonstrate the importance of EP monitoring in critical situations, SEP's were recorded from patients undergoing surgery. The effect of the anesthetic induction agent etomidate in significantly enhancing the cortical SEP has been recorded in the literature [13]. This effect had been observed by simple averaging at certain intervals. We decided to record SEP's continuously before and after administration of etomidate (0.2 mg/kg IV bolus) and carry out an FLC analysis of the data off-line.

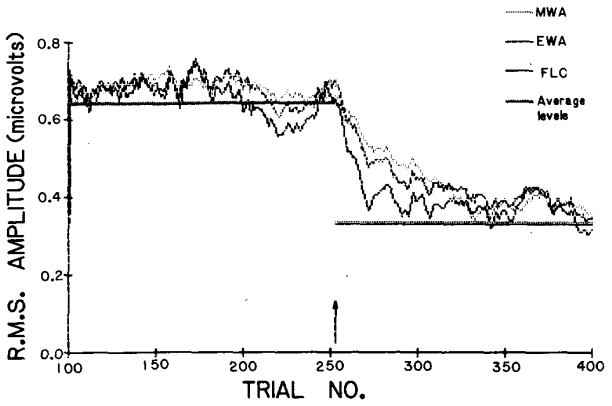


Fig. 7. Results of cortical SEP estimation from recordings on a normal human volunteer. Each thick straight line represents an rms amplitude level of an average of 250 trials. The point of transition is indicated by an arrow. In order to "zoom in" on the transient change, the estimates for the first 100 trials are not shown. For FLC estimation, $M = 6$ and $\mu = 1.41 \times 10^{-4}$.

The results are shown in the form of an amplitude trend plot in Fig. 8(a). Amplitude was measured as the difference between $N20$ and $P23$. Stimulation rate was 5.9/s. The time of drug administration is indicated by an arrow. Some sample estimated waveforms are also shown. Note the dramatic rise in amplitude after delivery of the drug and the subsequent return to normalcy. For the FLC, M_{max} was chosen as 30. The signal and noise spectra during the control period are plotted on a logarithmic scale in Fig. 8(b) (the signal spectrum is based on an ensemble average). The dotted line indicates the cutoff point corresponding to $M_{max} = 30$. It is apparent that this accounts for almost all the signal energy (well over 99 percent as a matter of fact). In this case, the optimal order was also determined as 30, according to the procedure outlined in Section III-B. The estimator was designed for a root-mean-square error of 0.3 μV.

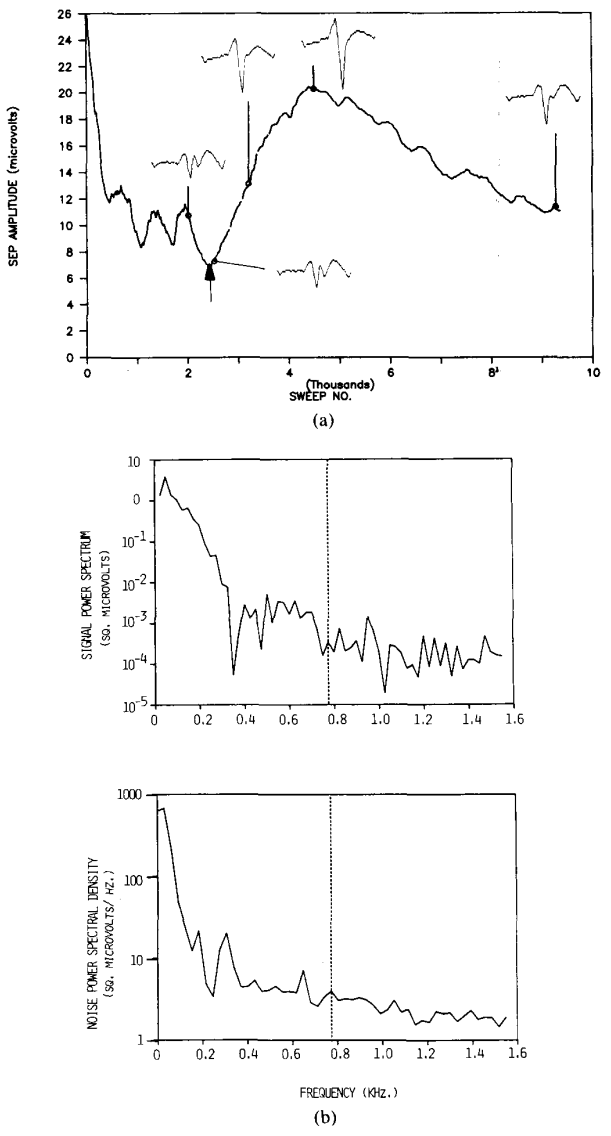


Fig. 8. Effect of etomidate on the cortical SEP. (a) Amplitude trend plot, generated by Fourier estimation, compressed for convenience of presentation. Arrow indicates moment of drug administration. The initial transient is due to the learning period of the algorithm. (b) Signal and noise spectral plots based on data recorded during the control period. Note the logarithmic scale. The dotted line corresponds to $M_{\max} = 30$.

V. Discussion and Conclusion

In the area of critical care monitoring, an adaptive process is essential in order to track changes in EP morphology; such changes could quite well be indicative of some insult to the central nervous system (such as trauma or oxygen deprivation, see [2]). Fourier modeling is highly appropriate since it takes into account both the periodic nature of the response as well as its band-limited nature (the latter is accounted for by truncating the series). Many linear filters used for EP's model the EP as an autoregressive (AR) or autoregressive moving average (ARMA) process [3], [8], [14]. The problem with implementing such a model adaptively is that the AR/ARMA parameters may be highly nonstationary even when the EP is invariant [7]. This is an unnecessary burden on the adaptive process. Furthermore, in these models, each point in the EP waveform is a linear combination of preceding points which are unknown due to noisy data. In contrast, the FLC models the EP as a linear combination of known sinusoidal functions.

There are limitations in the implementation of the FLC on EPs that need to be recognized and worked upon. Highly colored noise can cause a drastic drop in rate of convergence for a given steady-state accuracy. Further development of the algorithm will include a study of the effects of prewhitening. Another way of dealing with the problem is to reduce the model order by finding a way to select those sinusoidal functions that maximally represent the signal. The model does not take into account random fluctuations of the EP that have no pathological origin, although it may be possible to model these as another component of the additive noise.

Nonetheless, the FLC estimation technique does offer very significant advantages over other methods for detecting EP transients. It is not only possible to specify the estimation error but also to control it (recognizing, of course, the tradeoff between accuracy and speed of convergence). The model does not make ad hoc assumptions about the EP (such as it being an AR or ARMA process). Unlike most filters, there is zero phase shift. The Fourier coefficients are stationary as long as the EP is invariant. This makes it possible to use a simple gradient-search adaptive algorithm such as LMS; adaptation is required only when the evoked response is altered, as it should be. Convergence of the estimates is exponential and there is no problem of eigenvalue spread as in most applications of the LMS algorithm. When the eigenvalues of $\boldsymbol{R}$ are unequal, the smallest eigenvalue determines the rate of convergence since the time constant for each mode of the system is inversely proportional to the corresponding eigenvalue. This problem is obviated in the FLC because the eigenvalues are all equal and known, so that the speed of adaptation is uniform and a function solely of the step size. The input vector consists of known functions which can be precomputed and do not require additional data acquisition. The parametric representation of the EP makes the estimator ideal for data compression and storage of large banks of EP data. Furthermore, with some modifications, the estimator can also be used for other kinds of periodic signals. It is even capable of eliminating periodic interference if the period is distinct from that of the desired signal (as, for example, in pneumograms contaminated by cardiogenic interference).

Appendix A

Refer once more to Fig. 2. We have the system identification problem

$$d_k = \boldsymbol{x}_k^T \boldsymbol{w}^* + v_k \qquad \text{(A1)}$$

where $\boldsymbol{w}^*$ is the unknown weight vector to be estimated;

$s_k = \boldsymbol{x}_k^T \boldsymbol{w}^*$ is the unknown signal. The noise v_k is also unobservable while d_k is the actual observed data (signal + noise). $\boldsymbol{x}_k$ is also known. The estimate at instant k

$$\hat{d}_k = \boldsymbol{x}_k^T \boldsymbol{w}_k. \tag{A2}$$

It is assumed that d_k, $\boldsymbol{x}_k$, and e_k are stationary in the wide sense. The MSE, $\epsilon_k = E[e_k^2]$ is a quadratic function of the weight vector $\boldsymbol{w}$ at any instant k. To examine the MSE as a function of the weights, it is necessary to drop the subscript k from $\boldsymbol{w}$ and ϵ.

$$\begin{aligned} e_k &= d_k - \boldsymbol{w}^T \boldsymbol{x}_k \\ e_k^2 &= d_k^2 + \boldsymbol{w}^T \boldsymbol{x}_k \boldsymbol{x}_k^T \boldsymbol{w} - 2 d_k \boldsymbol{x}_k^T \boldsymbol{w} \\ \epsilon &= E[e_k^2] = E[d_k^2] + \boldsymbol{w}^T \boldsymbol{R} \boldsymbol{w} - 2E[d_k \boldsymbol{x}_k^T] \boldsymbol{w} \end{aligned} \tag{A3}$$

where $\boldsymbol{R} = E[\boldsymbol{x}_k \boldsymbol{x}_k^T]$ [see (4)]. Taking the gradient of ϵ,

$$\partial \epsilon / \partial \boldsymbol{w} = 2\boldsymbol{R}\boldsymbol{w} - 2E[d_k \boldsymbol{x}_k].$$

It is apparent that the gradient is minimized for $\boldsymbol{w} = \boldsymbol{R}^{-1} E[d_k \boldsymbol{x}_k]$, i.e., for $\boldsymbol{w} = \boldsymbol{w}^*$ since $E[d_k \boldsymbol{x}_k] = \boldsymbol{R}\boldsymbol{w}^*$ due to decorrelation of signal and noise. Estimation can be done iteratively by using the instantaneous squared error e_k^2 itself as an estimate of ϵ_k:

$$\partial (e_k^2) / \partial \boldsymbol{w} = 2e_k (\partial e_k / \partial \boldsymbol{w}) = -2e_k \boldsymbol{x}_k. \tag{A4}$$

One can now use a steepest-descent type of gradient search algorithm to arrive at the minimum MSE:

$$\boldsymbol{w}_{k+1} = \boldsymbol{w}_k - \mu \partial (e_k^2) / \partial \boldsymbol{w} = \boldsymbol{w}_k + 2\mu e_k \boldsymbol{x}_k \tag{A5}$$

where μ is a gain constant determining the speed of adaptation. Equation (A5) is the well-known LMS algorithm. The authors of [5] state that theoretically the algorithm converges if

$$0 < \mu < 1/\alpha_{\max} \tag{A6}$$

where $\alpha_{\max}$ is the largest eigenvalue of $\boldsymbol{R}$. However, this result is based on certain ideal conditions which are not satisfied for the LMS algorithm. A more practical bound is also mentioned in [5], namely

$$0 < \mu < 1/tr[\boldsymbol{R}]. \tag{A7}$$

More rigorous proofs of convergence are found in [11].

Appendix B

The objective is to derive the approximation (10). Define

$$\tilde{\boldsymbol{w}}_k = \boldsymbol{w}_k - \boldsymbol{w}^* \tag{B1}$$

i.e., the weight error vector. There exists a recursion equation for $\tilde{\boldsymbol{w}}$ (see [12]),

$$\tilde{\boldsymbol{w}}_{k+1} = (\boldsymbol{I} - 2\mu \boldsymbol{x}_k \boldsymbol{x}_k^T) \tilde{\boldsymbol{w}}_k + 2\mu \boldsymbol{x}_k v_k. \tag{B2}$$

Strictly speaking, z_k should be used in place of v_k in (B2). Assumption 4) in Section III-A justifies this approximation. Now define the covariance matrix of the weight error vector as

$$\boldsymbol{\Gamma}_k = E[\tilde{\boldsymbol{w}}_k \tilde{\boldsymbol{w}}_k^T]. \tag{B3}$$

From (9b), we know that

$$\begin{aligned} \epsilon' &= \lim_{k\to\infty} E\left[(s_k' - \hat{d}_k)^2\right] \\ &= \lim_{k\to\infty} E\left[(\boldsymbol{w}^{*T} \boldsymbol{x}_k - \boldsymbol{w}_k^T \boldsymbol{x}_k)^2\right] \\ &= \lim_{k\to\infty} E\left[(\tilde{\boldsymbol{w}}_k^T \boldsymbol{x}_k)^2\right] = \lim_{k\to\infty} E\left[(\boldsymbol{x}_k^T \tilde{\boldsymbol{w}}_k)^2\right] \end{aligned} \tag{B4}$$

where $s_k' = \boldsymbol{w}^{*T} \boldsymbol{x}_k$ has been defined in (7). If the expectation is viewed as an ensemble average, then (B4) can be rewritten as

$$\begin{aligned} \epsilon' &= \lim_{k\to\infty} E\left[E\{\boldsymbol{x}^T \tilde{\boldsymbol{w}}_k \tilde{\boldsymbol{w}}_k^T \boldsymbol{x}\} \,\middle|\, \boldsymbol{x} = \boldsymbol{x}_k\right] \\ &= E\left[\boldsymbol{x}^T (\lim_{k\to\infty} \boldsymbol{\Gamma}_k) \boldsymbol{x} \,\middle|\, \boldsymbol{x} = \boldsymbol{x}_k\right] \\ &= E[\boldsymbol{x}_k^T \boldsymbol{\Gamma}_\infty \boldsymbol{x}_k] = E[tr(\boldsymbol{\Gamma}_\infty \boldsymbol{x}_k \boldsymbol{x}_k^T)] = tr[\boldsymbol{\Gamma}_\infty \boldsymbol{R}] \end{aligned} \tag{B5}$$

where $\boldsymbol{\Gamma}_\infty = \lim_{k\to\infty} \boldsymbol{\Gamma}_k$. It will subsequently be clear that this limit exists.

In general, it is difficult to come up with a tractable solution of the difference in (B2) that is an explicit function of μ. However, for $2\mu \ll 1$ and for $\boldsymbol{x}_k$ defined as in (6), (B2) can be approximated, for large k, by the exponential regression (refer to [15])

$$\tilde{\boldsymbol{w}}_k = \beta \tilde{\boldsymbol{w}}_{k-1} + 2\mu \boldsymbol{x}_{k-1} v_{k-1} \tag{B6}$$

where $\beta = 1 - \mu$. Solving,

$$\tilde{\boldsymbol{w}}_k = 2\mu \sum_{i=0}^{k-1} \beta^i \boldsymbol{x}_{k-1-i} v_{k-1-i}. \tag{B7}$$

From (B3) and (B7),

$$\begin{aligned} \boldsymbol{\Gamma}_k = (2\mu)^2 \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} \beta^{i+j} E[v_{k-1-i} v_{k-1-j}] \\ \cdot E[\boldsymbol{x}_{k-1-i} \boldsymbol{x}_{k-1-j}^T]. \end{aligned} \tag{B8}$$

A little algebraic work will show that

$$E[\boldsymbol{x}_{k-1-i} \boldsymbol{x}_{k-1-j}^T] + E[\boldsymbol{x}_{k-1-j} \boldsymbol{x}_{k-1-i}^T] = \boldsymbol{R}_{|i-j|}$$

where

$$\begin{aligned} \boldsymbol{R}_n = \operatorname{diag} \big[&\cos(\omega_0 n) \cos(2\omega_0 n) \cdots \cos(M\omega_0 n) \\ &\cos(\omega_0 n) \cdots \cos(M\omega_0 n)\big] \end{aligned} \tag{B9}$$

for $n > 0$. Rearranging (B8),

$$\begin{aligned} \boldsymbol{\Gamma}_k = (2\mu)^2 \big[&K_v(0) \boldsymbol{R}(\beta^0 + \beta^2 + \cdots + \beta^{2k-2}) \\ &+ K_v(1) \boldsymbol{R}_1(\beta^1 + \beta^3 + \cdots + \beta^{2k-3}) \\ &+ \cdots + K_v(n) \boldsymbol{R}_n(\beta^n + \cdots + \beta^{2k-2-n}) \\ &+ \cdots + K_v(k-1) \boldsymbol{R}_{k-1} \beta^{k-1} \end{aligned}$$

where $K_v(i) = E[v_k v_{k+i}]$. The limit value of the error covariance matrix

$$\begin{aligned} \boldsymbol{\Gamma}_\infty = \left[(2\mu)^2/(1-\beta^2)\right] \big(&K_v(0) \boldsymbol{R} + \beta K_v(1) \boldsymbol{R}_1 \\ &+ \cdots + \beta^p K_v(p) \boldsymbol{R}_p\big) \end{aligned} \tag{B10}$$

which exists since $0 < \beta < 1$. Remembering that $\beta = 1 - \mu$, $\mu << 1$, (B10) can be reduced to

$$\mathbf{\Gamma}_\infty \sim 2\mu \sum_{i=1}^{p} \beta^i K_v(i)\mathbf{R}_i + 2\mu K_v(0)\mathbf{R}. \quad \text{(B11)}$$

From (4) and (B5),

$$\epsilon' = \lim_{k\to\infty} E\left[tr\left(\mathbf{\Gamma}_\infty \mathbf{R}\right)\right] = tr\left(\mathbf{\Gamma}_\infty\right)/2$$

$$\sim \mu\left[\sum_{i=1}^{p} \beta^i K_v(i) \sum_{j=1}^{M} 2\cos\left(j\omega_0 i\right) + K_v(0)M\right]$$

by (B11), (B9)

$$= \mu\left[2\sum_{i=1}^{p} (1-\mu)^i K_v(i) \sum_{j=1}^{M} \cos\left(j\omega_0 i\right) + K_v(0)M\right]$$

which is the same as (10).

ACKNOWLEDGMENT

The authors are indebted to Dr. R. McPherson, Dr. G. Empting-Koschorke, and M. Petr of The Johns Hopkins Hospital for their valuable assistance in collecting data necessary for this paper.

REFERENCES

[1] C. D. McGillem, J. I. Aunon, and D. G. Childers, "Signal processing in evoked potential research: Applications of filtering and pattern recognition," in *Critical Reviews in Bioengineering*. Cleveland, OH: CRC, Oct. 1981, pp. 225–265.

[2] T. Tsubokawa and R. Ramsay, "Evoked responses: Use in neurological intensive care" in *Intensive Care for Neurological Trauma and Disease*. B. A. Green, L. F. Marshall, T. J. Gallagher, Eds. New York: Academic, 1982, pp. 201–216.

[3] N. V. Thakor, "Improved SNR in evoked potentials by adaptive filtering," in *Proc. VII Ann. Conf. IEEE/EMBS*, Chicago, IL, 1985, pp. 112–113.

[4] E. R. Ferrara and B. Widrow, "The time-sequenced adaptive filter," *IEEE Trans. Circuits Syst.*, vol. CAS-28, pp. 519–523, 1981.

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

[6] D. O. Walter, "*A posteriori* 'Wiener filtering' of average evoked responses," *Electroencephalogr. clin. Neurophysiol.*, vol. 27, pp. 61–70, 1969.

[7] J. P. C. de Weerd, "*A posteriori* time-varying filtering of average evoked potentials," *Biol. Cybern.*, vol. 41, pp. 211–222, 1981.

[8] K. Yu and C. D. McGillem, "Optimum filters for estimating evoked potential waveforms," *IEEE Trans. Biomed. Eng.*, vol. BME-24, pp. 232–241, 1977.

[9] J. A. Sgro, R. G. Emerson, and T. A. Pedley, "Real-time reconstruction of evoked potentials using a new two-dimensional filter method," *Electroencephalogr. clin. Neurophysiol.*, vol. 62, pp. 372–380, 1985.

[10] C. A. Vaz, I. N. Bankman, and N. V. Thakor, "Evoked potential estimation using a Fourier series model," in *Proc. IX Ann. Conf. IEEE Eng. Med. Biol. Soc.*, Boston, MA, 1987, pp. 600–601.

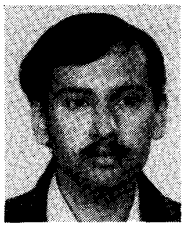
[11] R. R. Bitmead, "Convergence in distribution of LMS-type adaptive parameter estimates," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 54–60, 1983.

[12] 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, 1985.

[13] R. W. McPherson, B. Sell, and R. J. Traystman, "Effects of thiopental, fentanyl and etomidate on upper extremity somatosensory evoked potentials in humans," *Anesthesiology*, vol. 65, pp. 584–589, 1986.

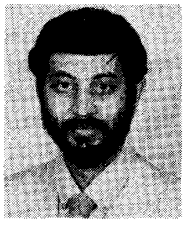
[14] H. J. Heinze and H. Künkel, "ARMA—Filtering of evoked potentials," *Meth. Inform. Med.*, vol. 23, pp. 29–36, 1984.

[15] J. E. Mason, E. W. Bai, L-C. Fu, M. Bodson, and S. S. Sastry, "Analysis of adaptive identifiers in the presence of unmodelled dynamics: Averaging and tuned parameters," in *Proc. 26th IEEE Conf. Decision and Control*, Los Angeles, CA, 1987, pp. 360–365.



Christopher Vaz (S'83) 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.



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

For two years he worked as an Electronics Engineer for Philips India. Between 1981 and 1983 he served on the faculty of Northwestern University, and is currently an Associate Professor in the Biomedical Engineering Department of The Johns Hopkins Medical School, Baltimore, MD. His current research interests are cardiovascular and neurosensory instrumentation, signal processing, and microcomputer and VLSI applications. Currently, he teaches courses on biomedical instrumentation and computer applications in medicine. He also serves as a consultant to medical instrumentation industries in these areas.

Dr. Thakor is a recent recipient of two research awards: a Research Career Development Award from the National Institutes of Health and a Presidential Young Investigator Award from the National Science Foundation.