Identification of Nonlinear Biological Systems Using Laguerre Expansions of Kernels

Vasilis Z. Marmarelis

Biomedical and Electrical Engineering, University of Southern California, Los Angeles, CA

Abstract — Identification of nonlinear dynamic systems using the Volterra-Wiener approach requires the estimation of system kernels from input-output data. A kernel estimation technique, originally proposed by Wiener (1958) and recently studied by Ogura (1986), employs Laguerre expansions of the kernels and estimates the unknown expansion coefficients via time-averaging of covariance samples. This paper presents another implementation of the technique which utilizes least-squares fitting instead of covariance time-averaging and provides for the proper selection of the intrinsic Laguerre parameter "α". Results from simulation examples demonstrate that this implementation can yield accurate kernel estimates up to 3rd-order from short input-output data records. Furthermore, it is shown that this implementation remains effective in the presence of noise and when the spectral characteristics of the input signal deviate somewhat from the theoretical requirements of whiteness. The computational requirements and the overall performance of this technique compare favorably to existing methods, especially in cases where the system kernels can be represented with a relatively small number of Laguerre basis functions.

Keywords — Nonlinear systems, Kernel estimation, Laguerre expansions.

INTRODUCTION

The problem of identification and modeling of nonlinear biological systems from experimental input-output (stimulus-response) data has been attracting increasing attention in recent years. The need arises in the study of important nonlinearities inherent in physiological function. For instance, the nervous system relies on essential nonlinear operations (such as, generation of action potentials, stimulus compression, nonlinear inhibition or facilitation, etc.) to accomplish a great diversity of tasks; the cardiovascular, renal, and respiratory systems rely on nonlinear mechanisms for homeodynamic control and autoregulation; to mention a few celebrated nonlinearities in physiological systems. Among the possible methodological approaches, the methods based on the Volterra-Wiener theory of nonlinear systems have gained considerable popularity in recent years. A number of important applications and theoretical studies have established the strengths and limitations of this approach (for partial review see [1,6,7,8,10,11]).

For systems that can be tested experimentally with broadband stimuli and can be represented adequately by low-order nonlinearities, this approach appears particularly efficacious. Among the advantages are: (a) the canonical representation of nonlinearities with kernel functions which provide rigorous and predictive models of the input-output functional relation; (b) the ability to obtain these models accurately under noisy experimental conditions; (c) the validity of these models over the entire frequency range of the system operation; (d) the fact that these models do not rely on prior postulates and can be extended to nonstationary cases; (e) the ability to study nonlinear systems with multiple inputs and multiple outputs; (f) the experimental efficiency of this approach, resulting from the rich information content of the employed test stimulus; and (g) the inclusion of linear dynamics as a first-order component of the obtained results. Among the current disadvantages of this approach are: (a) the practical limitations of modeling high-order nonlinearities, except in certain special cases; (b) the more sophisticated experimental apparatus required for the necessary broadband stimulus, and the associated data collection and processing requirements; and (c) the physiological interpretation of high-order (2nd and 3rd) kernels. We must note that these limitations have been receding gradually with the advent of new methodological and technological advances.

This paper addresses the important practical issue of how to obtain accurate results (kernels) in the presence of noise from short experimental data-records. Notable advances on the same issue have been recently reported

(3,4) and constitute, along with the method presented herein, powerful tools in resolving one of the critical problems in the application of the Volterra-Wiener approach. The advocated approach employs the orthonormal basis of discrete-time Laguerre functions to expand the kernels and reduce the number of unknown parameters that need be estimated, relative to the traditional discrete-time representation of the kernels. This approach was first suggested by Wiener in his pivotal monograph (16) and was first implemented by Watanabe and Stark (15). More recently, Ogura (9) provided an insightful treatment of this approach and its actual application to real data. Ogura's method utilizes time-averages of covariance samples (between the system output and the respective Laguerre filter outputs) to obtain estimates of the unknown expansion coefficients. This requires strict whiteness of the system input. The implementation proposed in this paper utilizes, instead, least-squares estimation of the unknown expansion coefficients. This leads to increased estimation accuracy in the presence of noise and reduces the requirements in terms of length of experimental data-records. It also allows accurate kernel estimation when the experimental stimulus deviates from the theoretical requirement of whiteness, as long as the stimulus remains sufficiently broadband and all significant kernels are included in the estimated model. Caution must be exercised in the case of non-white stimuli, because biases in the obtained kernel estimates may be caused by lack of orthogonality of possible high order functionals not included in the estimated model. Furthermore, the implementation presented herein provides for proper selection of the intrinsic Laguerre parameter "α", an important practical issue that was not addressed in previous work.

The purpose of this paper is to draw attention to this promising approach and provide sufficient technical detail and illustrative examples in order to allow its expanded use by interested investigators. This may address the confining issues of limited experimental data and/or non-white broadband experimental stimuli encountered in certain biomedical studies. The methodological approach is presented in the Discrete-Time Laguerre Expansion Method section and some illustrative examples from computer simulations are given in the Illustrative Examples of Kernel Estimation section. Application of this method to actual experimental data is currently under way in a variety of biomedical systems (neural, renal, respiratory).

DISCRETE-TIME LAGUERRE EXPANSION METHOD

The Volterra-Wiener approach to nonlinear system identification utilizes a functional expansion of the system response signal in terms of the stimulus signal. Details of this approach and its various extensions can be

found in numerous publications (e.g. [1,6,7,10,11]). The background necessary for the specific objectives of this paper is summarized below.

For the general nonlinear time-invariant continuous system, the response signal $y(t)$ can be expressed in terms of a series of functionals (Volterra series) that represent multiple convolutions of the stimulus signal $x(t)$:

$$y(t) = \sum_{n=0}^{\infty} \int_0^{\infty} \ldots \int_0^{\infty} k_n(\tau_1, \ldots, \tau_n) x(t - \tau_1)$$
$$\ldots x(t - \tau_n) d\tau_1 \ldots d\tau_n . \quad (1)$$

The multiple convolution integrals of the Volterra series involve high-order kernel functions $\{k_n(\tau_1, \ldots, \tau_n)\}$ which constitute the descriptors of the system nonlinear dynamics (Volterra kernels). Consequently, the system identification task is to obtain estimates of these kernels from stimulus-response data. These kernel functions are symmetric with respect to their arguments, i.e., attain the same value for any permutation of given $(\tau_1, \ldots, \tau_n)$ values.

In order to facilitate the estimation of the system kernels, Wiener suggested the orthogonalization of the Volterra functionals for a Gaussian white noise (GWN) stimulus. The functional terms of the Wiener series are constructed on the basis of a Gram-Schmidt orthogonalization procedure requiring that the covariance between any two Wiener functionals $\{G_n(t)\}$ be zero. The resulting Wiener series expansion takes the form:

$$y(t) = \sum_{n=0}^{\infty} G_n[h_n; x(t'), t' \le t]$$

$$= \sum_{n=0}^{\infty} \sum_{m=0}^{[n/2]} \frac{(-1)^m n! P^m}{(n - 2m)! m! 2^m}$$

$$\times \int_0^{\infty} \ldots \int_0^{\infty} h_n(\tau_1, \ldots, \tau_{n-2m},$$

$$\lambda_1, \lambda_1, \ldots, \lambda_m, \lambda_m)$$

$$\times x(t - \tau_1) \ldots x(t - \tau_{n-2m})$$

$$\times d\tau_1 \ldots d\tau_{n-2m} d\lambda_1 \ldots d\lambda_m \quad (2)$$

where $[n/2]$ is the integer part of $n/2$ and P is the power level of the GWN stimulus. The set of Wiener kernels $\{h_n\}$ is, in general, different from the set of Volterra kernels $\{k_n\}$. Specific relations, however, exist between the two sets of kernels (6).

Clearly, the Wiener kernels depend on the GWN stimulus power level whereas the Volterra kernels are independent of any stimulus characteristics. This is due to the fact that the Wiener kernels are associated with an orthog-

onal functional expansion (when the input is GWN of some power level P) whereas the Volterra kernels are associated with an analytic functional expansion that does not depend on any input characteristics. Both sets of kernels can be used to predict the system response to any given stimulus, provided these sets are complete.

The orthogonality of the Wiener series allows the estimation of Wiener kernels from input-output data in the general case, as described below. Since the orthogonality of the Wiener terms is independent of the specific kernel functions involved, "instrumental" Wiener functionals can be used to isolate each term in a series and subsequently obtain the corresponding kernel. For instance, if a mth order instrumental functional $Q_m[q_m;x(t'),t' \leq t]$ is used then:

$$
\begin{aligned}
E[y(t)Q_m(t)] &= \sum_{n=0}^{\infty} E[G_n(t)Q_m(t)] \\
&= E[G_m(t)Q_m(t)] \\
&= m!P^m \int_0^{\infty} \ldots \int_0^{\infty} h_m(\tau_1, \ldots, \tau_m) \\
&\quad \times q_m(\tau_1, \ldots, \tau_m) \, d\tau_1, \ldots, d\tau_m \quad (3)
\end{aligned}
$$

where the "instrumental" kernel $q_m(\tau_1, \ldots, \tau_m)$ can be arbitrarily chosen. Since q_m is known, the left-hand-side of Eq. 3 can be evaluated from input-output data. In order to evaluate the unknown kernel $h_m(\tau_1, \ldots, \tau_m)$ on the right-hand-side, judicious choice of the instrumental kernel q_m is necessary to facilitate the solution of Eq. 3. Wiener suggested the use of any suitable multidimensional orthonormal basis for this purpose, and the Laguerre basis in particular for reasons explained below. So, if $\{b_j(\tau)\}$ is a complete orthonormal (CON) basis over the range of the system memory μ, then instrumental kernels of the form:

$$
q_m(\tau_1, \ldots, \tau_m; j_1, \ldots, j_m) = b_{j_1}(\tau_1) \ldots b_{j_m}(\tau_m) \quad (4)
$$

can be used to obtain the expansion coefficients $\{c_{j_1, \ldots, j_m}\}$ of the unknown kernel over the specified CON basis as:

$$
c_{j_1, \ldots, j_m} = \frac{1}{m!P^m} E[y(t)Q_m(t;j_1, \ldots, j_m)] \quad (5)
$$

where

$$
h_m(\tau_1, \ldots, \tau_m) = \sum_{j_1} \ldots \sum_{j_m} c_{j_1, \ldots, j_m} b_{j_1}(\tau_1) \ldots b_{j_m}(\tau_m) .
$$

$$(6)$$

Note that in this case:

$$
\begin{aligned}
&Q_m(t;j_1, \ldots, j_m) \\
&= \sum_{k=0}^{[m/2]} \frac{(-1)^k P^k m!}{(m-2k)!k!2^k} \\
&\quad \times v_{j_1}(t) \ldots v_{j_{m-2k}}(t) \delta_{j_{m-2k+1}, j_{m-2k+2}} \cdots \delta_{j_{m-1}, j_m} \quad (7)
\end{aligned}
$$

where

$$
v_j(t) = \int_0^{\mu} b_j(\tau)x(t-\tau) \, d\tau \quad (8)
$$

and δ denotes the Kronecker delta. The importance of this observation lies in the fact that the functions $v_j(t)$ are independent Gaussian random processes with zero mean and variance P, and the instrumental functionals $\{Q_m\}$ can be seen as orthogonal (Hermite) polynomials in the variables $(v_1, \ldots, v_m)$, when $x(t)$ is GWN. For the general causal system, Wiener proposed the use of Laguerre functions for the CON set $\{b_j(\tau)\}$ since they are defined over the interval $\tau \in [0, \infty)$ and can be generated easily in analog form—a fashionable mode at that time—by a simple RC ladder network. Furthermore, they have a "built-in" exponential which makes them suitable for kernels of exponential asymptotic values, as in most physical systems.

Lee and Schetzen (5) proposed a different implementation of Wiener's original ideas for kernel estimation that has been widely used because of its relative simplicity. The Lee and Schetzen technique is based on the observation that the product of m time-shifted versions of the GWN input can be written in the form of the leading term of an instrumental functional Q_m using a product of delta functions for the instrumental kernel. They were able to show that use of this instrumental functional allows kernel estimation through input-output crosscorrelation:

$$
\begin{aligned}
&h_m(\tau_1, \ldots, \tau_m) \\
&= \frac{1}{m!P^m} E[y_m(t)x(t-\tau_1) \ldots x(t-\tau_m)] \quad (9)
\end{aligned}
$$

where $y_m(t)$ represents the mth order response residual

$$
y_m(t) = y(t) - \sum_{n=0}^{m-1} G_n(t) . \quad (10)
$$

The simplicity and elegance of the crosscorrelation technique led to many applications, particularly in the area of physiological systems (for partial review, see [6,7,8]).

The use of the crosscorrelation technique revealed certain practical limitations of this approach. The most important ones are the following: (a) the stringent requirement of a band-limited white-noise stimulus (that covers the entire system bandwidth); and (b) the considerable estimation variance of the kernels (especially of order higher than first) obtained from finite length input-output data-records. The latter limitation is an inevitable consequence of the stochastic nature of the employed stimulus and the fact that the crosscorrelation estimates converge to the true values at a rate proportional to the square-root of the record length. This limitation is also incumbent on the Ogura implementation of the Laguerre expansion approach, since the unknown expansion coefficients are estimated by time-averaging computation of the covariance shown in Eq. 5. Thus, long data-records are often required to obtain estimates of satisfactory accuracy, resulting also in increased computational burden. This has prompted the use of specially designed stimuli, such as pseudo-random m-sequences (13) and sum-of-sinusoids of incommensurate frequencies (14), as well as stochastic approximation methods (2). These efforts have yielded improved results in certain cases but still present the user with a set of confining trade-offs. The most effective solution to date has been offered by the exact orthogonalization method of Korenberg (4), which overcomes the problem of estimation variance by orthogonalizing the discrete-time form of the functional series with respect to the particular stimulus data-record used in the experiment. This approach also alleviates the need for strict whiteness of the stimulus, even though broad-band spectral characteristics are still required and caution must be exercised in interpreting the obtained kernel estimates, since they may deviate from the Wiener set if the stimulus is non-white and higher order system nonlinearities exist above and beyond the ones represented by the estimated kernels.

The implementation of the Laguerre expansion approach presented in this paper exhibits strengths and weaknesses akin to Korenberg's method. It overcomes the problem of estimation variance by performing a least-squares fit of the actual data and relaxes the whiteness requirements of the experimental stimulus (with the caveat mentioned above). The main difference between the two methods is that the representation of the kernels in Laguerre expansion form, instead of discrete-time form, may yield more compact kernel representations which may result in higher estimation accuracy and reduced computational burden, especially in high-order kernel estimation. This is illustrated in the next section, where accurate estimates of 2nd- and 3rd-order kernels from short data-records are presented. We must note that the relative efficacy of these two methods will depend, in general, on the specific kernel characteristics.

The proposed method is cast in discrete-time. The kernels are expanded on the orthonormal set of discrete-time Laguerre functions as indicated in Eq. 6, where the τ's assume discrete non-negative values and $\{b_j(\tau)\}$ denotes the jth order discrete-time orthonormal Laguerre function:

$$b_j(\tau) = \alpha^{(\tau-j)/2}(1 - \alpha)^{1/2}$$
$$\times \sum_{k=0}^{j} (-1)^k \binom{\tau}{k} \binom{j}{k} \alpha^{j-k}(1 - \alpha)^k \quad (\tau \geq 0)$$

$$(11)$$

where α is the discrete-time Laguerre parameter ($0 < \alpha < 1$) which determines the rate of exponential asymptotic decline of these functions. The key variables $\{v_j(t)\}$ are the discrete-time convolutions of the discrete-time stimulus $x(t)$ with $\{b_j(\tau)\}$:

$$v_j(t) = T \sum_{\tau=0}^{M} b_j(\tau)x(t - \tau) \tag{12}$$

where T is the sampling interval, M is the kernel memory extent in number of lags ($\mu = MT$), and t assumes discrete-time values ($t = 0,1,\ldots,N$); where $R = NT$ is the record length.

The sought kernel expansion coefficients $\{c_{j_1,\ldots,j_m}\}$ can be estimated by time-averaging implementation of the ensemble averaging operation indicated in Eq. 5, if $x(t)$ is GWN (as suggested previously by Wiener and Ogura). However, as we indicated previously, more accurate estimates result from a least-squares regression of the discrete-time response data $y(t)$ on the computed functional terms $\{Q_m(t;j_1,\ldots,j_m)\}$. This also alleviates the strict requirement of stimulus whiteness, if all significant terms are included in the model. These computations would be rather burdensome, if it were not for the following two key observations: (a) the Laguerre expansion of kernels typically results in significant reduction of the number of estimated coefficients, as illustrated in the following section; and (b) the computation of the variables $\{v_j(t)\}$ can be accelerated significantly by use of the recursive relation:

$$v_j(t) = \sqrt{\alpha}v_j(t - 1) + \sqrt{\alpha}v_{j-1}(t) - v_{j-1}(t - 1) \tag{13}$$

which is due to the particular form of the discrete-time Laguerre functions. Computation of this recursive relation must be initialized by the following recursive equation that yields $v_0(t)$ for given stimulus $x(t)$:

$$v_0(t) = \sqrt{\alpha}v_0(t - 1) + T\sqrt{1 - \alpha}x(t) . \tag{14}$$

These computations can be preformed rather fast, for $t = 0, 1, \ldots, N$ and $j = 0, 1, \ldots, L - 1$; where L is the total number of Laguerre functions used in the kernel expansion. The choice of the Laguerre parameter α is rather critical in achieving efficient kernel expansions (and, consequently, fast and accurate kernel estimation) and its judicious selection is discussed in the following section.

ILLUSTRATIVE EXAMPLES OF KERNEL ESTIMATION

In this section, the use of discrete-time Laguerre expansions for kernel estimation is illustrated and certain important properties of this method are discussed.

We begin by illustrating certain basic properties of the discrete-time Laguerre functions (LF). The first 5 LFs for $\alpha = 0.2$ are shown in Fig. 1 (T = 1). We note that the number of zero-crossings (roots) of each LF equals its order. Furthermore, the higher the order the longer the significant values of a LF spread over time, and the time separation between zero-crossings increases. This is further illustrated in Fig. 2, where the LFs of order 4, 8, 12, and 16 are shown for $\alpha = 0.2$. An interesting illustration of

the first 50 LFs for $\alpha = 0.2$ is given in Fig. 3, plotted as a square matrix over 50 time untis (0–49 lag) in 3-D perspective (top display) and contour plot (bottom display). We note the symmetry of this matrix and the fact that higher order LFs are increasingly "delayed" in their significant values. This observation implies that kernels with a pure delay may require higher order LFs for their accurate representation. Nonetheless, their accurate estimation is still possible and the number of required LFs can be potentially reduced by use of "associated Laguerre functions" of appropriate order (9).

In the frequency domain, the FFT magnitude of all LFs for given α is identical, as illustrated in Fig. 4 for the 5 LFs of Fig. 1. The FFT phase, however, of these LFs is different, as illustrated in Fig. 5. We note that the nth order LF exhibits a maximum phase shift of $(n\pi)$, with the odd-order ones commencing at π radians and the even-order ones commencing at 0 radians (at zero frequency). Thus, the minimum-phase LF is always for order zero.

In order to examine the effect of α on the form of the LFs, we show in Fig. 6 the 4th order LF for $\alpha = 0.1, 0.2$, and 0.4. We observe that increasing α results in longer spread of significant values and zero-crossings. Thus ker-
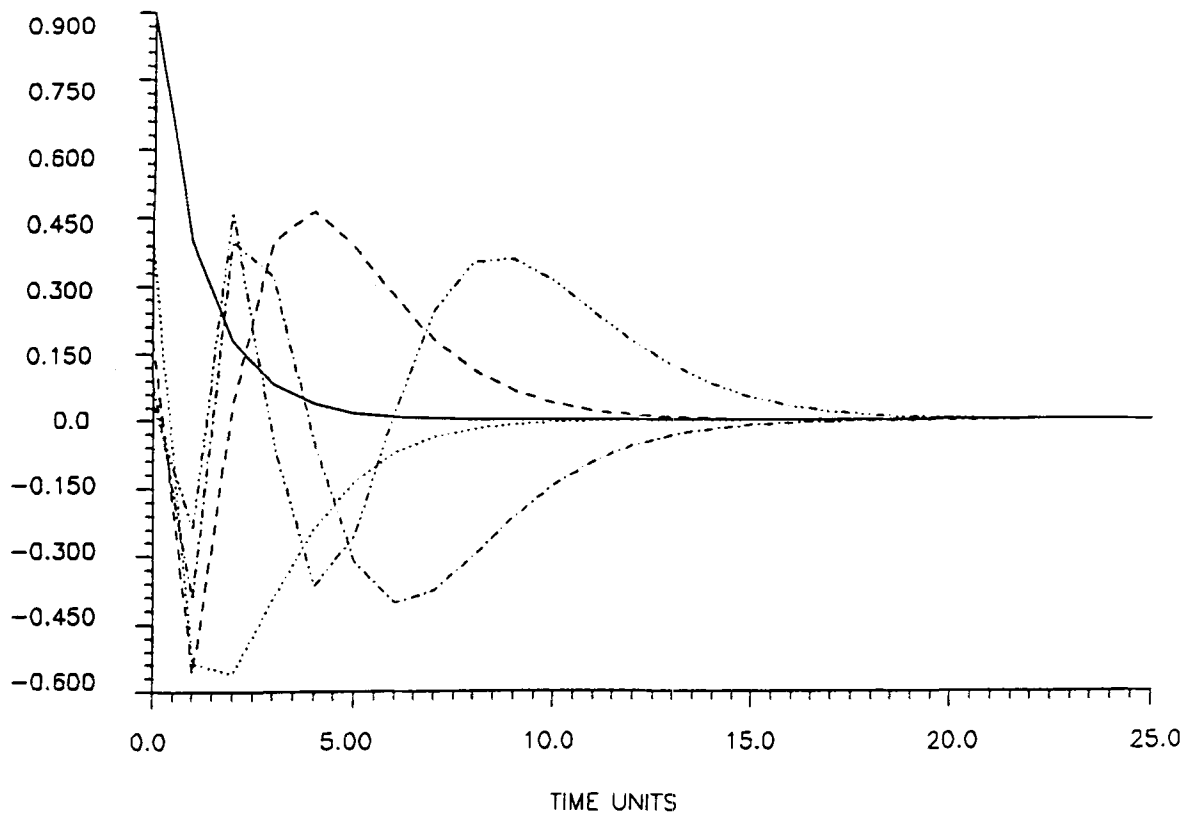

FIGURE 1. Discrete-time Laguerre functions (LFs) of order 0 (solid), 1 (dotted), 2 (dashed), 3 (dot-dash), 4 (dot-dot-dot-dash) for $\alpha = 0.2$, plotted over the first 25 lags.
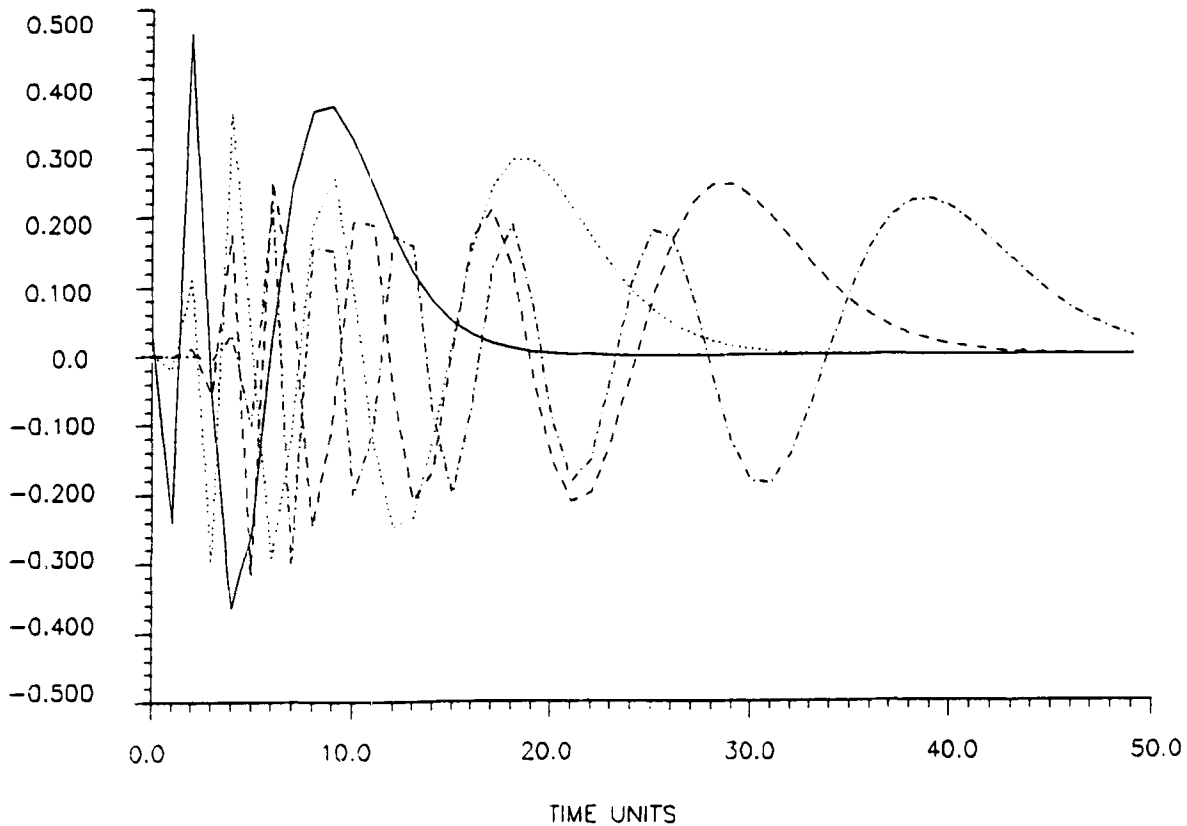
FIGURE 2. Discrete-time Laguerre functions (LFs) of order 4 (solid), 8 (dotted), 12 (dashed), 16 (dot-dash) for $\alpha = 0.2$, plotted over the first 25 lags.

nels with longer memory may require a larger α for efficient representation. In the frequency domain, the FFT magnitudes and phases of the LFs of Fig. 6 are shown in Figs. 7 and 8, respectively. We observe that for larger α, the lower frequencies are emphasized more (Fig. 7) and the phase lags faster (Fig. 8) although the total phase shift is the same (4 π for all 4th order LFs). A more complete illustration of the effect of α is given in Fig. 9, where the matrix of the first 50 LFs for $\alpha = 0.1$ is plotted in the same fashion as in Fig. 3 (for $\alpha = 0.2$). It is clear from these figures that increasing α increases the separation of the zero crossings (ripple in 3-D perspective plots) and broadens the "fan" formation evident in the contour plots. These observations can guide us in selecting the proper value of α for given kernel memory (M) and number of LFs (L). Specifically, we select the value of α for which the kernel memory is covered by significant values of LFs while, at the same time, the LF values are diminished for $\tau \gg M$ (in order to secure their orthogonality). In other words, for given values (M, L), α must be chosen so that the point (M, L) in the contour plane be near the edge of the "fan" formation but outside this "fan."

We now demonstrate the use of LF expansions for kernel estimation. Consider first a second-order nonlinear system with the 1st- and 2nd-order kernels shown in Figs. 10 and 11, respectively. This system is simulated for a GWN stimulus of 512 data points, and the kernels are estimated using both the Lee-Schetzen crosscorrelation (LSC) technique and the proposed Laguerre expansion of kernels (LEK) technique. The 1st-order kernel estimates are shown in Fig. 12 along with the exact kernel (solid line), where the LEK estimate is plotted with dotted line (almost exactly superimposed on the solid line) and the LSC estimate is plotted with dashed line. The 2nd-order kernel estimates are shown in Fig. 13 for LEK and Fig. 14 for LSC. The superiority of the LEK estimates is evident.

These LEK estimates were obtained by use of 10 LFs and $\alpha = 0.1$. The required computing time on a VAX Workstation was 6 s. The obtained Laguerre coefficients (LC) in this case are shown in Figs. 15 and 16 for 1st- and 2nd-order kernels, respectively. Note that the 2nd-order LCs are plotted as a complete (symmetric) matrix, even though only the entries of one triangular region are esti-
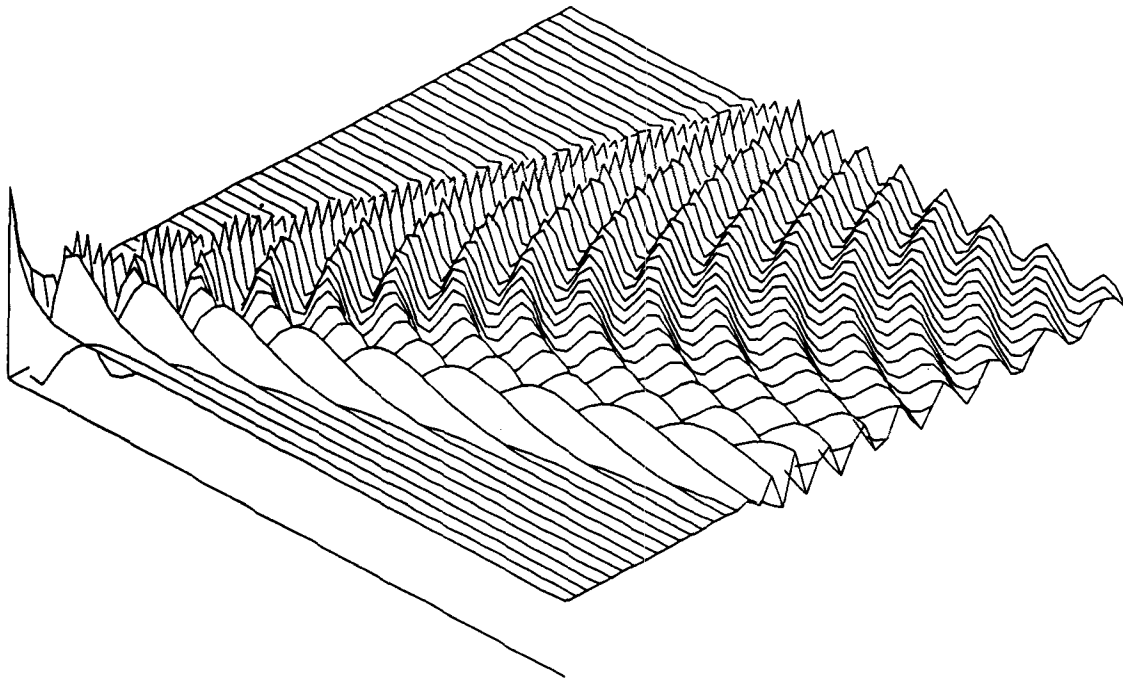
```
x-min=    0.0000     y-min=    0.0000     z-min=  -0.5600E+00
x-max=   49.0000     y-max=   49.0000     z-max=   0.8944E+00
```

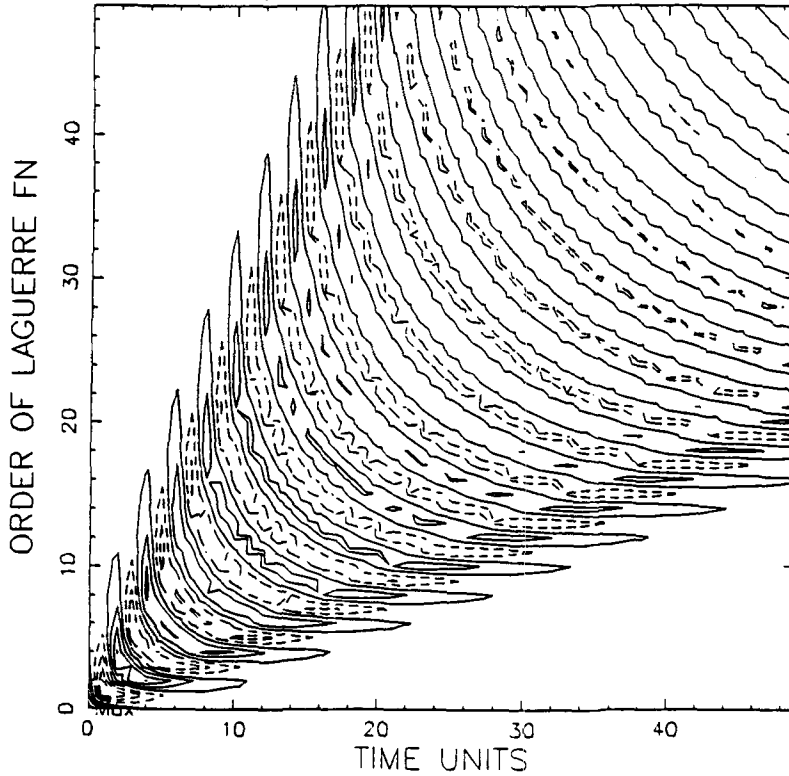| Data file: LM2	Cont.Inc.	Cont.lines	Data-set	
	0.1454	0.8944	0.8944	Max
L-FNS MATRIX (A=0.2)		-0.5600	-0.5600	Min

FIGURE 3. The first 50 LFs for $\alpha = 0.2$, plotted from 0 to 49 lags in 3-D perspective plot (top display) and contour plot (bottom display).
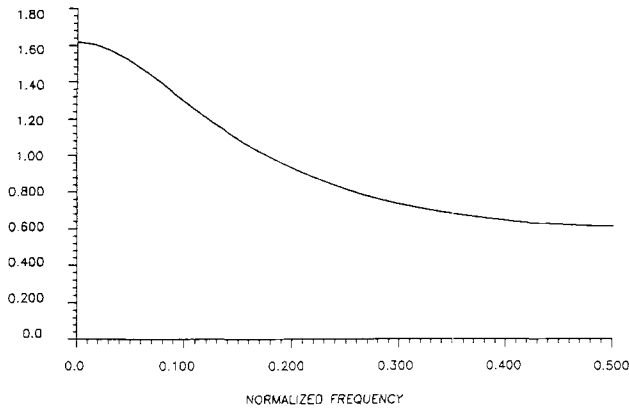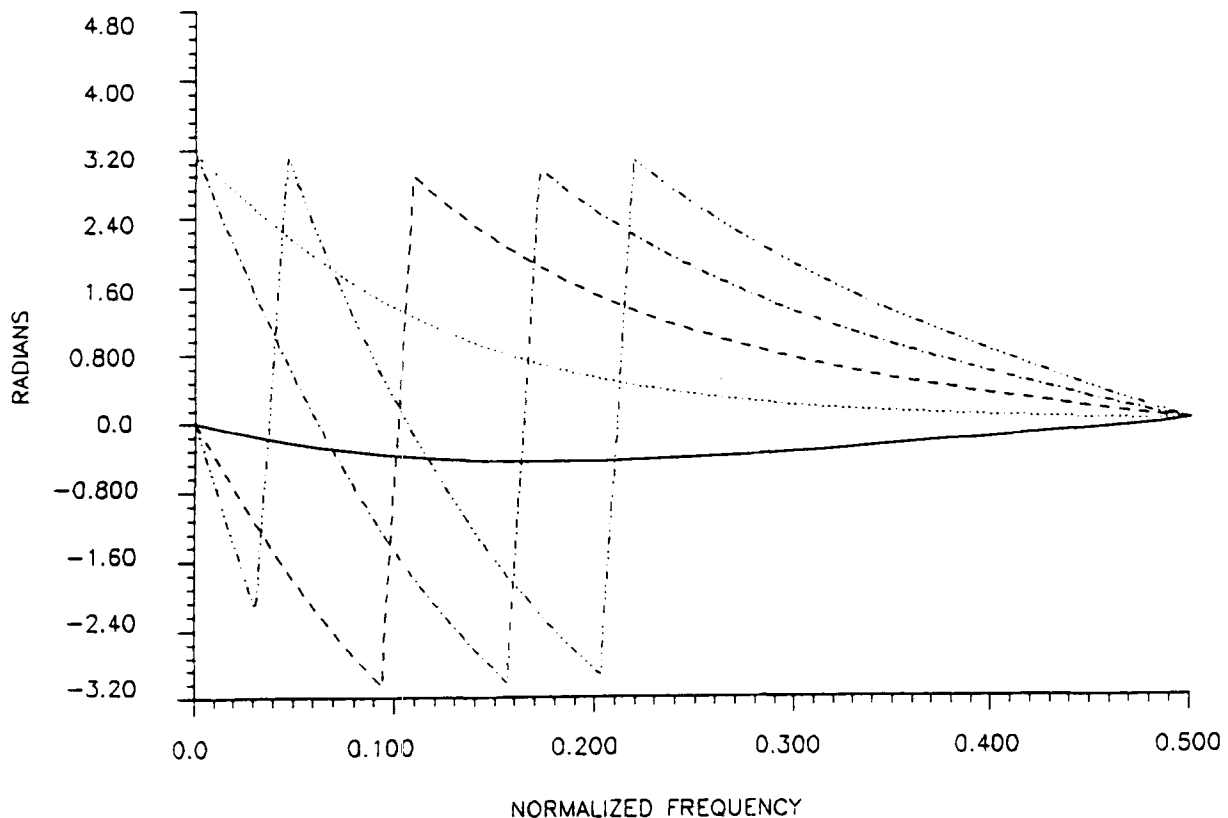
FIGURE 4. FFT magnitude of the first 5 (orders 0 to 4) LFs (shown in Fig. 1) for $\alpha = 0.2$, plotted up to the normalized Nyquist frequency of 0.5.

of these kernels can also be accomplished with $L = 8$ (due to the very small values of the coefficients for the 9th and 10th LFs), which demonstrates the significant reduction in kernel representation accomplished by the Laguerre expansion.

It must be noted that the number L of required Laguerre functions for accurate representation of the kernels of a given system critically affects the computational burden associated with this method. In general, the total number of resulting expansion coefficients for a system of order K is: $(L + K)!/(L!K!)$. Note that this number includes all kernels up to order K and takes into account the well-known symmetries in high-order kernels. The required L in a given application can be practically determined by estimating the 1st-order kernel for a large L, and then inspect the resulting coefficient estimates to select the minimum number that corresponds to significant values. For instance, in the example of Fig. 15, estimation of coefficients beyond the plotted ones would yield negligible values. The same reasoning can be applied to high-order kernels (e.g., the 2nd-order coefficients shown in Fig. 16). Selection of the required maximum kernel order K is based on the adequacy of the output prediction

mated. Thus, in general, the number of LCs estimated using "L" LFs for kernel expansion is: $[L(L + 1)/2]$ for the 2nd-order kernel and L for the 1st-order kernel. It is evident from Figs. 15 and 16 that an adequate estimation

NORMALIZED FREQUENCY

FIGURE 5. FFT phase of the first 5 (orders 0 to 4) LFs (shown in Fig. 1) for $\alpha = 0.2$, plotted up to the normalized Nyquist frequency of 0.5.
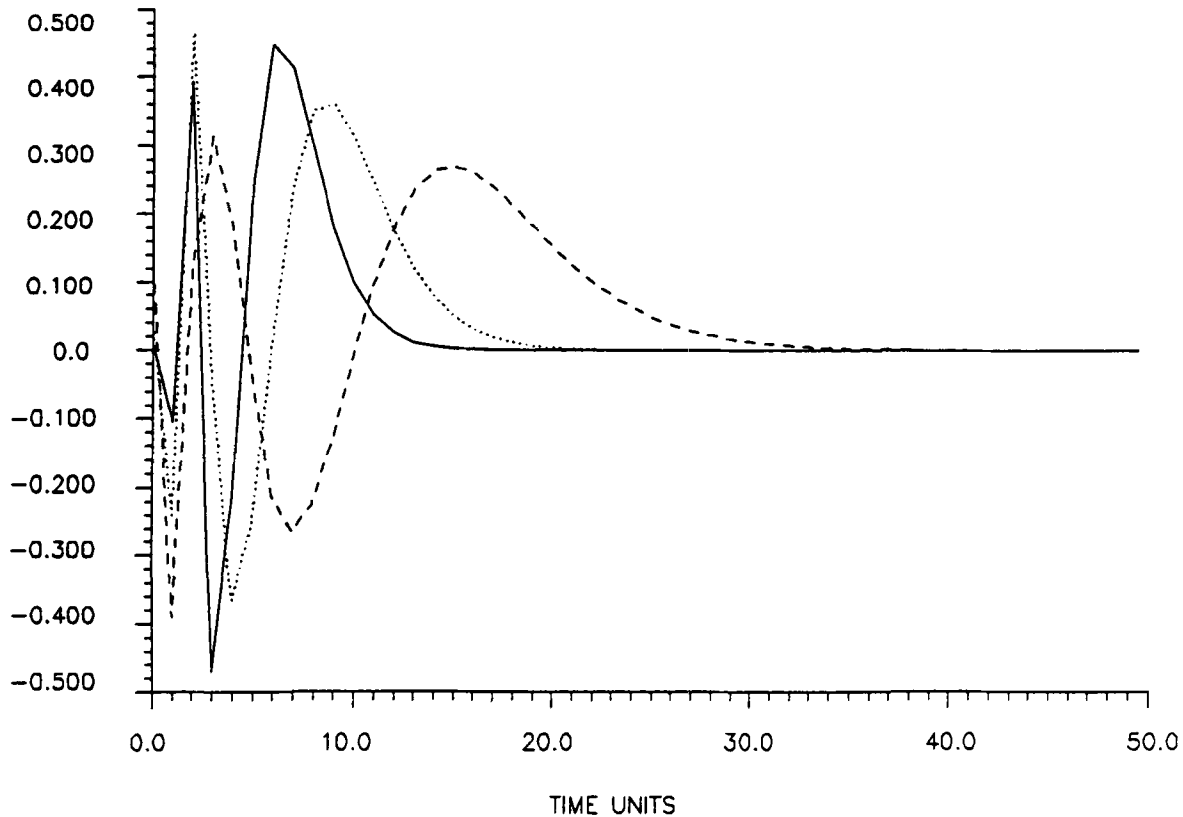
FIGURE 6. The 4th-order LFs for $\alpha = 0.1$ (solid), 0.2 (dotted), 0.4 (dashed), plotted from 0 to 49 lag.

accuracy of the model of a given order, as in all previous applications of the Volterra-Wiener approach.

We now examine the effect of noise on the obtained kernel estimates by adding independent GWN to the response signal for a signal-to-noise ratio of 10 dB. The 1st-
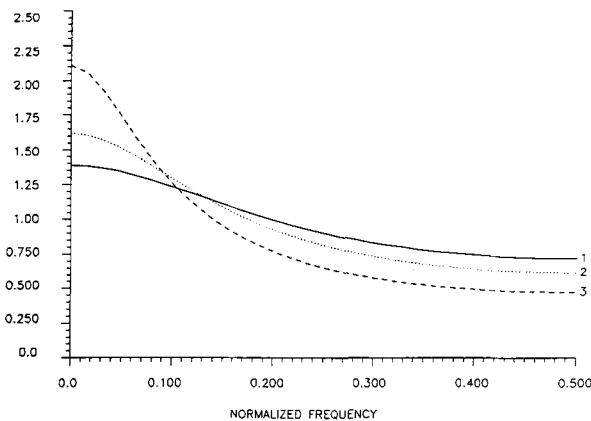

FIGURE 7. FFT magnitude of 4th-order LFs shown in Fig. 6. Trace 1: $\alpha = 0.1$; Trace 2: $\alpha = 0.2$; Trace 3: $\alpha = 0.4$.

order kernel estimates obtained by the LEK and LSC techniques are shown in Fig. 17 in the following manner: LEK in solid line and LSC in dotted line. The corresponding 2nd-order kernel estimates are shown in Fig. 18 for LEK and Fig. 19 for LSC. Comparison with the exact kernels shown in Figs. 10–11 indicates the superiority of the LEK technique under these conditions. Note that the LEK estimates in this demonstration were computed with $L = 8$, which resulted also in shorter computing time, i.e., 4 s. The fact that kernel estimates of this quality can be obtained from such short records of data (512 data points), even in cases with significant noise (SNR = 10 dB), can have important implications in actual applications of the Volterra-Wiener approach.

Another important issue in actual applications is the effect of higher order nonlinearities on the obtained lower order kernel estimates. Since most applications limit themselves to the first two kernels, the presence of higher-order (>2) terms acts as a source of "correlated noise" which is dependent on the input. To illustrate this effect, we add 3rd- and 4th-order nonlinearities to the previous system and recompute the kernel estimates. Note that the simulated system is a simple cascade of a linear filter followed
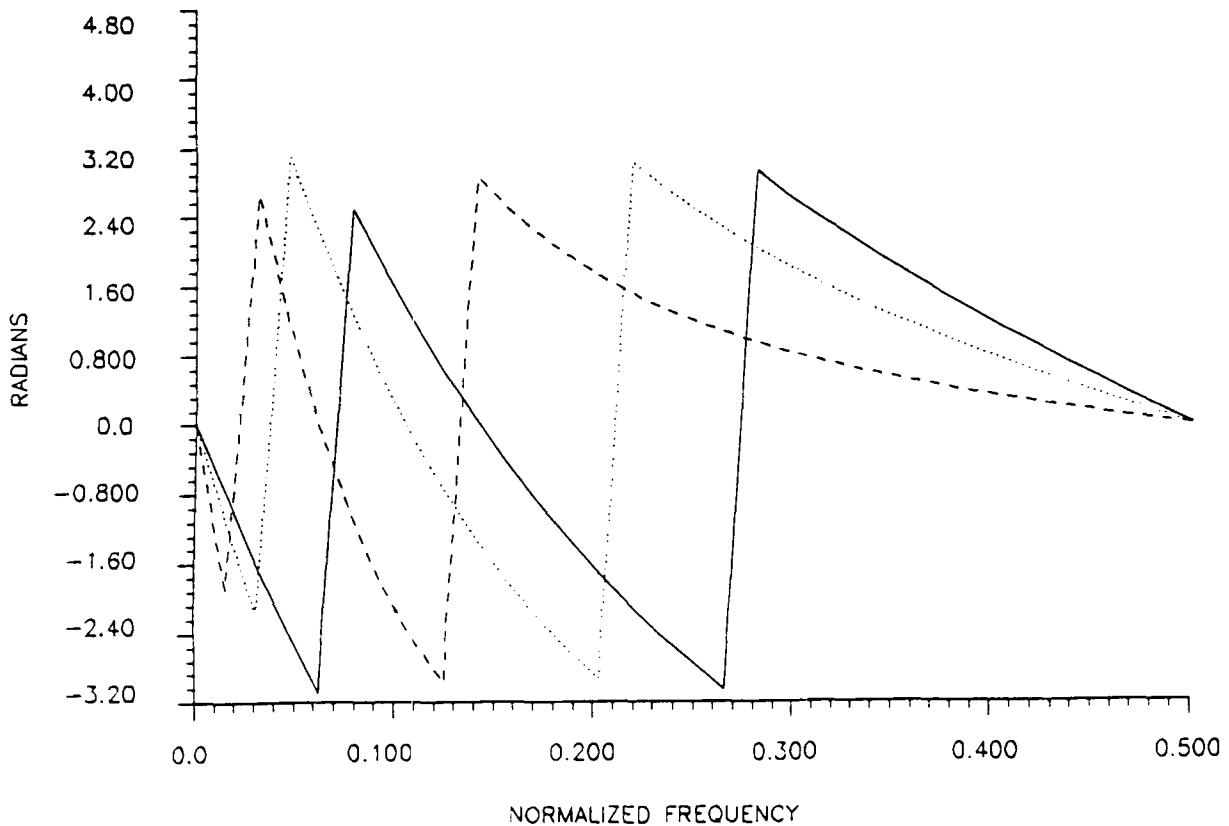
FIGURE 8. FFT phase of 4th-order LFs shown in Fig. 6. Trace 1: $\alpha = 0.1$; Trace 2: $\alpha = 0.2$; Trace 3: $\alpha = 0.4$ (line drawing style corresponds to Figs. 6 and 7).

by a static nonlinearity of the form: $y = z + z^2$ (for the 2nd-order system) and $y = z + z^2 + z^3/3 + z^4/4$ (for the 4th-order system) where $z(t)$ is the output of the linear filter. Thus the exact 1st- and 2nd-order Wiener kernels of this 4th-order system do not change in waveform but only in scale. The obtained 1st-order kernels are shown in Fig. 20: the exact kernel is plotted with solid line, the LEK estimate with dotted line, and the LSC estimate with dashed line. The LEK estimate is better than its LSC counterpart, but it exhibits certain deviations from the exact kernel due to the presence of the 3rd- and 4th-order terms. The exact 2nd-order kernel of the 4th-order system has the same form as in Fig. 11 but scaled by a factor of 2.23. The 2nd-order kernel estimates obtained via LEK and LSC techniques are shown in Figs. 21 and 22, respectively. Clearly, the LEK estimate is better than its LSC counterpart, and closely resembles the exact kernel in form and size.

Another important attribute of the proposed technique is its ability to yield accurate kernel estimates even when the stimulus deviates from the theoretically required white noise. This is critically important in experimental studies

where a white-noise stimulus cannot be practically secured. As an illustrative example, consider the previous 2nd-order system being simulated for a non-white (broadband) stimulus that has the spectrum shown in Fig. 23. The 1st-order kernel estimates obtained via LEK and LSC are shown in Fig. 24, along with the exact kernel. The LEK estimate (dotted line) is closely superimposed on the exact kernel (solid line), while the LSC estimate shows the effects of the non-white stimulus in terms of estimation bias (in addition to the anticipated estimation variance). The 2nd-order kernel estimates obtained via LEK and LSC are shown in Figs. 25 and 26, respectively. Clearly, the LEK estimates are far superior, and they are not affected by the non-whiteness of the employed stimulus. Note, however, that these LEK estimates will be affected by the non-white stimulus in the presence of higher order nonlinearities. This is illustrated by simulating the previous 4th-order system with the non-white stimulus. The obtained 1st-order kernel estimates via LEK and LSC are shown in Fig. 27, along with the exact kernel. The LEK estimate is affected more than before, when a white stimulus was used for the 4th-order system (see Fig. 20), but,
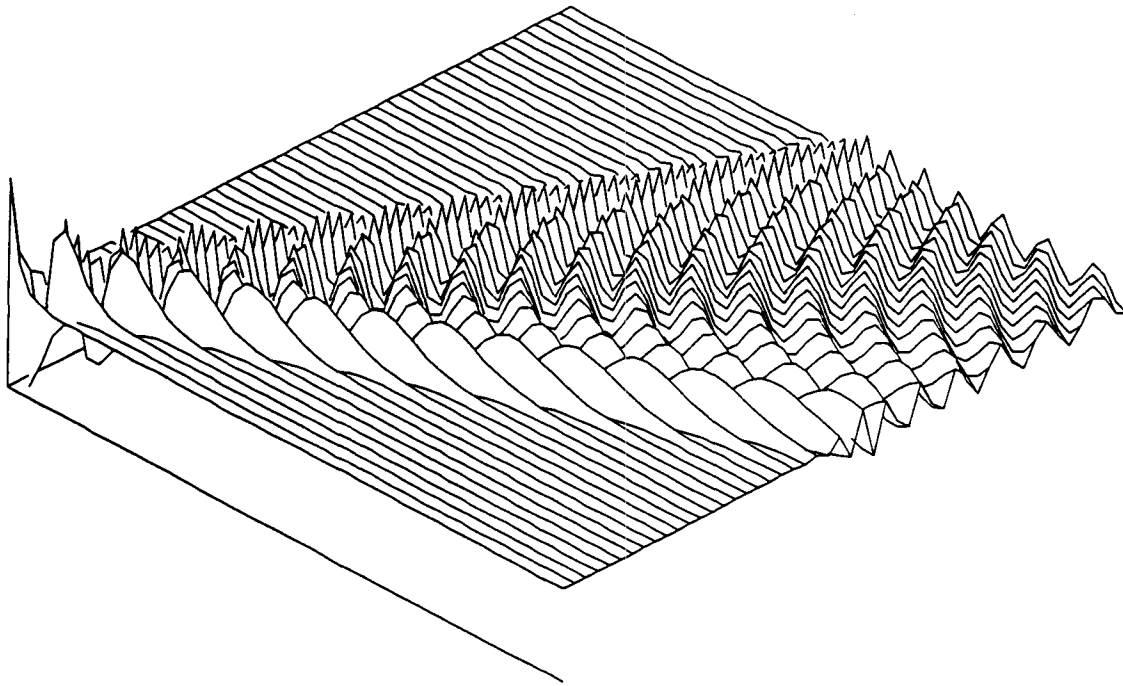
```
x-min=    0.0000      y-min=    0.0000      z-min= -0.7589E+00
x-max=   49.0000      v-max=   49.0000      z-max=  0.9487E+00
```

Data file: LM1   Cont.Inc.   Cont.lines Data-set
                            0.1708   0.9487   0.9487 Max
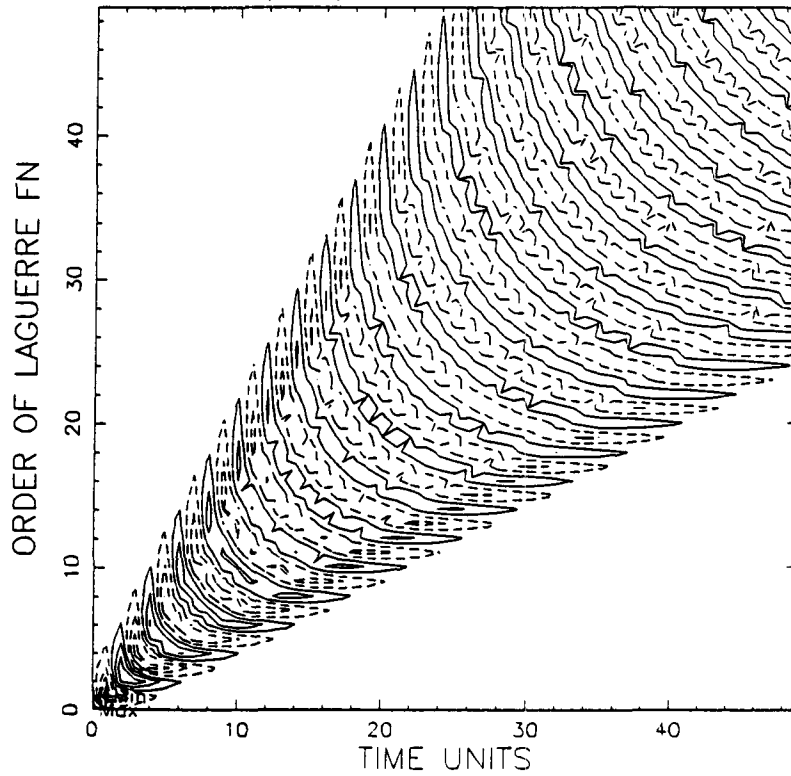L-FNS MATRIX (A=0.1)   -0.7589   -0.7589 Min

FIGURE 9. The first 50 LFs for α = 0.1, plotted from 0 to 49 lags in 3-D perspective plot (top display) and contour plot (bottom display).
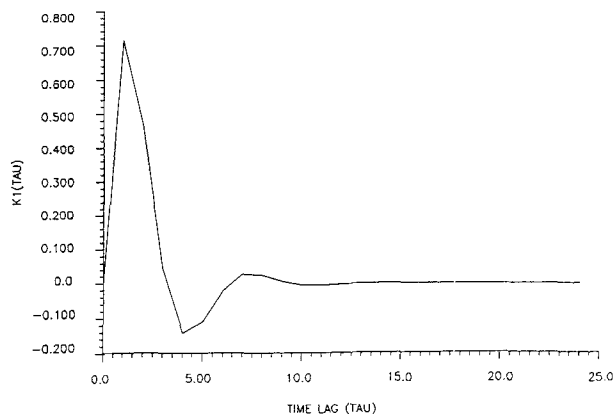
FIGURE 10. Exact 1st-order kernel of simulated system.

still, much less than its LSC counterpart. The same is true for the 2nd-order kernel estimates shown in Fig. 28 for LEK and Fig. 29 for LSC. It is interesting to note that the form of the LSC 2nd-order estimates for the non-white stimulus (Figs. 26 and 29) is not affected much by the presence of higher order terms, even though they are rather poor in both cases.

These results demonstrate the fact that the proposed LEK technique yields very accurate kernel estimates from short experimental data-records, even for non-white (but broadband) stimuli, when there are no significant higher order nonlinearities (higher than the estimated ones). However, these kernel estimates may be seriously affected when the experimental stimulus is non-white and significant higher order nonlinearities exist. This is due to the fact that the corresponding Wiener functionals are no longer orthogonal when the stimulus is non-white, resulting in certain estimation bias owing to "projections" from these higher order terms. Of course, this problem is alleviated when all significant nonlinearities (kernels) are
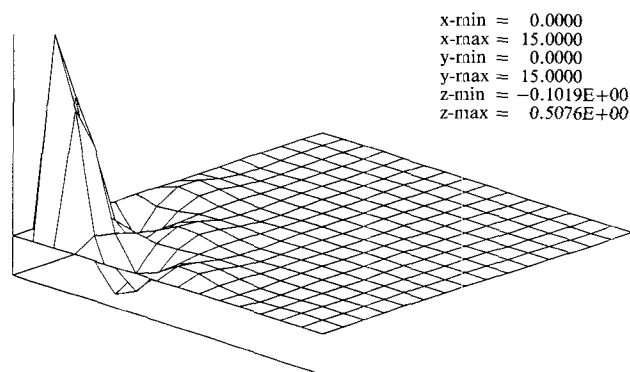
estimated. For this reason, the Laguerre expansion technique was extended to 3rd-order kernel estimation, as illustrated below.

The accurate estimation of 3rd-order kernels from short data-records is made possible by the ability of the LEK technique to reduce (in most cases) the number of required parameters for kernel representation. Although this is not guaranteed in every possible situation, the structure of the Laguerre functions (i.e., exponentially weighted polynomials) makes it likely for most physical or physiological systems whose kernels usually exhibit asymptotically exponential structure. As an illustrative example, consider the previously simulated system extending to the 3rd-order nonlinearity receiving a GWN input (1,024 data points). The resulting 3rd-order kernel estimate via the LEK technique was rather accurate. This is demonstrated in Fig. 30, where the 3rd-order kernel estimate for $\tau_3 = 4$ is shown as a 3-D function (note that visualization of 3rd-order kernels requires taking 2-D "cuts" for specific values of τ_3). Comparison with the exact 3rd-order kernel "cut" at $\tau_3 = 4$, shown in Fig. 31, indicates the efficacy of the LEK technique. Results of similar quality were obtained for all other values of τ_3. To demonstrate the superiority of the proposed technique over the traditional crosscorrelation technique, the corresponding LSC estimate (for $\tau_3 = 4$) is shown in Fig. 32. The ability of the proposed technique to make the estimation of 3rd-order kernels practical and accurate from short data-records creates a host of exciting possibilities for the actual identification and analysis of 3rd-order nonlinearities in biological systems with odd-order (e.g., sigmoid-type) nonlinearities that are rather common in biology and have gained increasing prominence in neuroscience in recent years.

These results illustrate the efficacy of the proposed kernel estimation technique and demonstrate some of its important properties. Similar results have been obtained from numerous other simulations. The technique is currently applied to the analysis of actual experimental data.

CONCLUSIONS

The use of discrete-time Laguerre expansions for kernel estimation via least-squares fitting offers certain practical advantages over existing techniques in applications of the Volterra-Wiener approach. These advantages result from the ability of Laguerre expansions to yield compact kernel representations in many cases of practical interest. Specifically, accurate kernel estimates (up to 3rd-order) can be obtained from short experimental data-records, even if the experimental stimulus deviates from white noise (provided that it remains sufficiently broadband and all significant kernels are included in the estimated model). Furthermore, the proposed technique is


```
x-min  =    0.0000
x-max  =   15.0000
y-min  =    0.0000
y-max  =   15.0000
z-min  =   -0.1019E+00
z-max  =    0.5076E+00
```

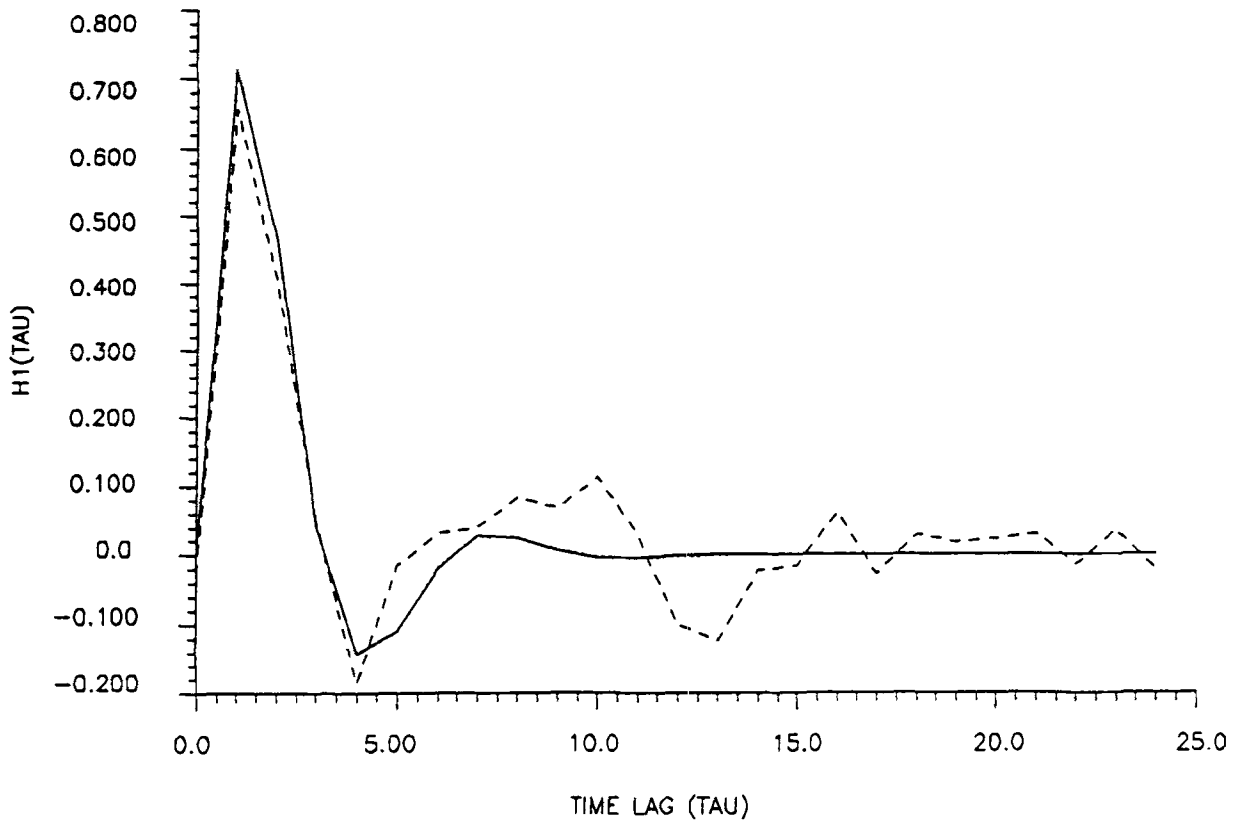**FIGURE 11. Exact 2nd-order kernel of simulated system.**

FIGURE 12. 1st-order kernel estimates obtained via LEK: Laguerre expansion of kernels (dotted) and LSC: Lee-Schetzen crosscorrelation (dashed). The exact kernel is plotted with solid line and it is nearly superimposed on the LEK estimate.

remarkably robust in the presence of data-contaminating noise and does not require long computing time.

The efficacy of the proposed technique and its aforementioned advantages have been demonstrated (up to 3rd-order kernel estimation) with computer simulations. It is

hoped that use of this technique will enhance and broaden the applicability of the Volterra-Wiener approach in cases where only short data-records (possibly noisy) are available and/or the experimental stimulus deviates from white noise.

x-min =   0.0000
x-max = 15.0000
y-min =   0.0000
y-max = 15.0000
z-min = −0.1013E+00
z-max =   0.5075E+00

x-min =   0.0000
x-max = 15.0000
y-min =   0.0000
y-max = 15.0000
z-min = −0.1464E+00
z-max =   0.5421E+00
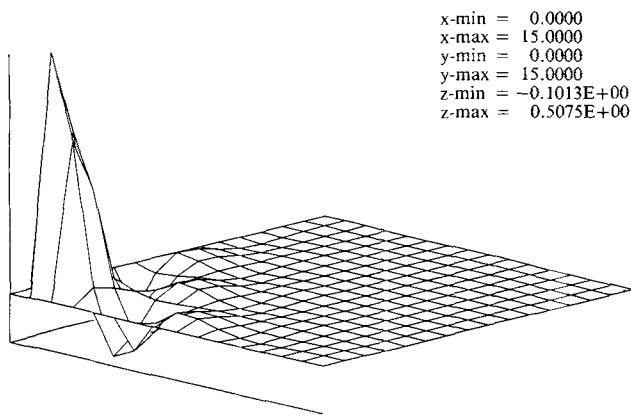
FIGURE 13. 2nd-order kernel estimated obtained via LEK technique.
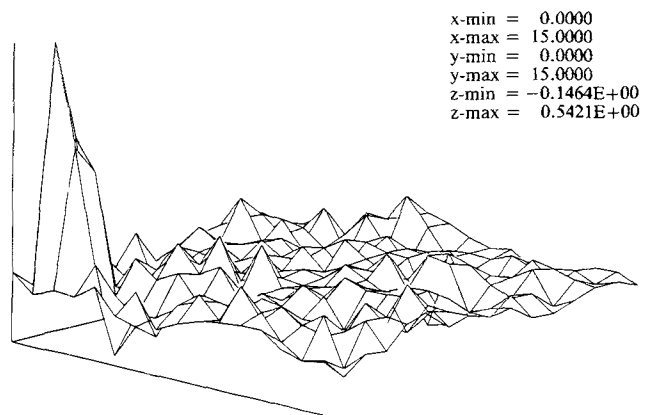
FIGURE 14. 2nd-order kernel estimated obtained via LSC technique.
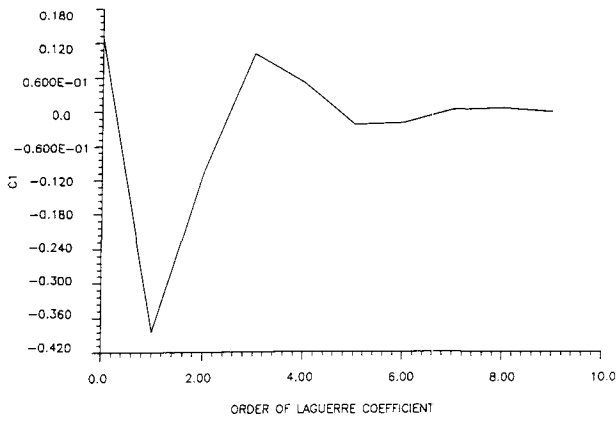
FIGURE 15. Estimates of the 10 (order 0 to 9) Laguerre expansion coefficients for the 1st-order kernel.
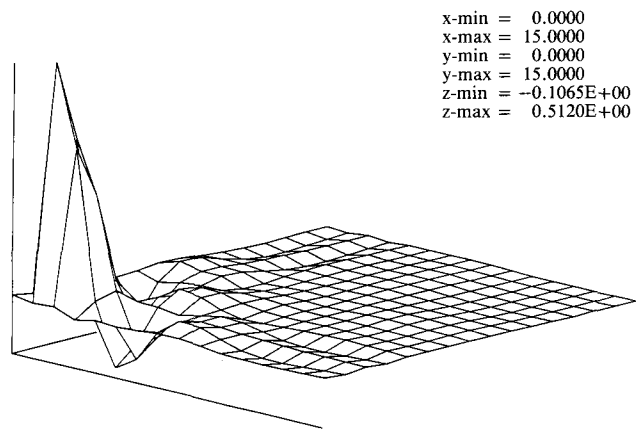
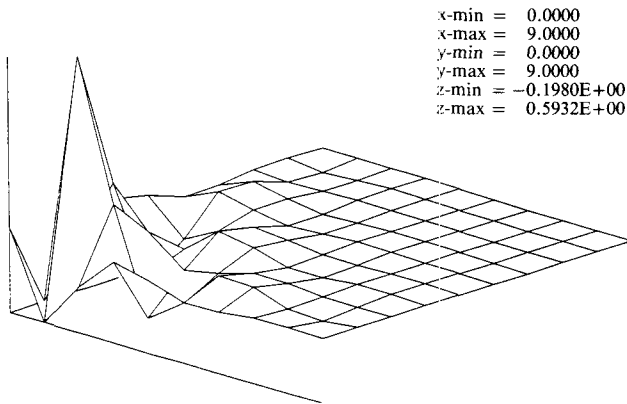FIGURE 18. 2nd-order kernel estimate via LEK for noisy data.

FIGURE 16. Estimates of the Laguerre expansion coefficients (order 0 to 9 in each dimension) for the 2nd-order kernel, plotted as a symmetric 2-D array (total number of distinct coefficients is 55).
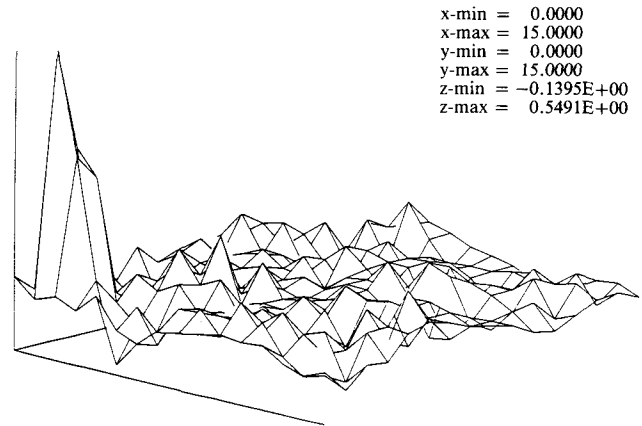

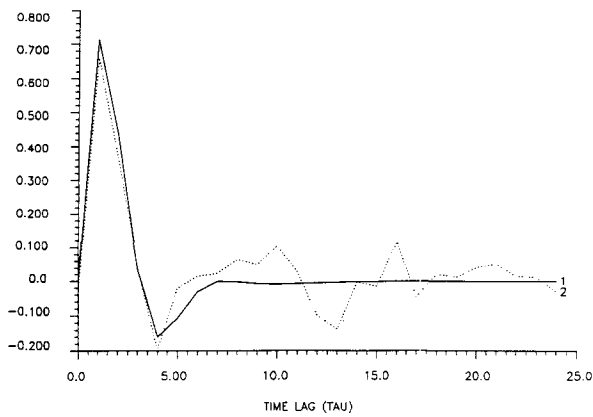FIGURE 19. 2nd-order kernel estimate via LSC for noisy data.

FIGURE 17. 1st-order kernel estimates via LEK (solid) and LSC (dotted) for noisy data (SNR = 10 dB).
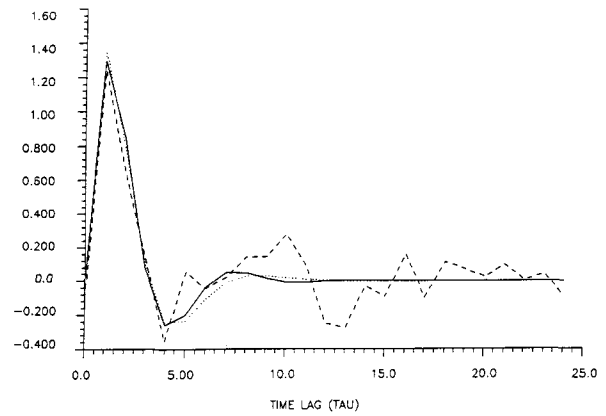
FIGURE 20. 1st-order kernel estimates for 4th-order simulated system obtained via LEK (dotted) and LSC (dashed). The exact kernel is plotted with solid line.

x-min = 0.0000
x-max = 15.0000
y-min = 0.0000
y-max = 15.0000
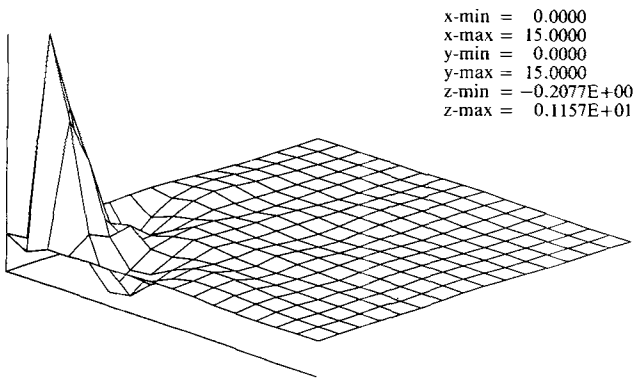z-min = −0.2077E+00
z-max = 0.1157E+01

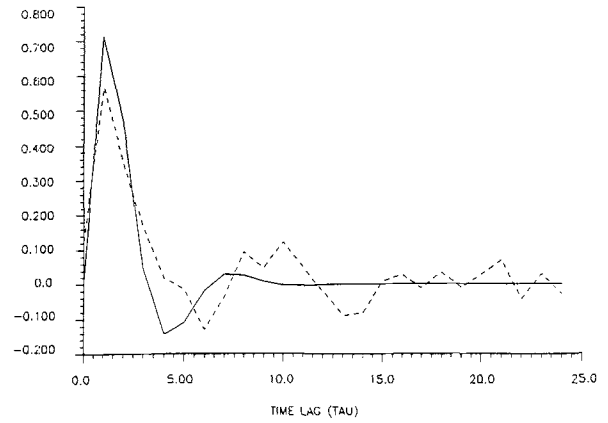FIGURE 21. 2nd-order kernel estimate for 4th-order simulated system obtained via LEK.

FIGURE 24. 1st-order kernel estimates for non-white stimulus obtained via LEK (dotted) and LSC (dashed). The exact kernel in solid line is nearly superimposed on the LEK estimate.

x-min = 0.0000
x-max = 15.0000
y-min = 0.0000
y-max = 15.0000
z-min = −0.3564E+00
z-max = 0.1223E+01

x-min = 0.0000
x-max = 15.0000
y-min = 0.0000
y-max = 15.0000
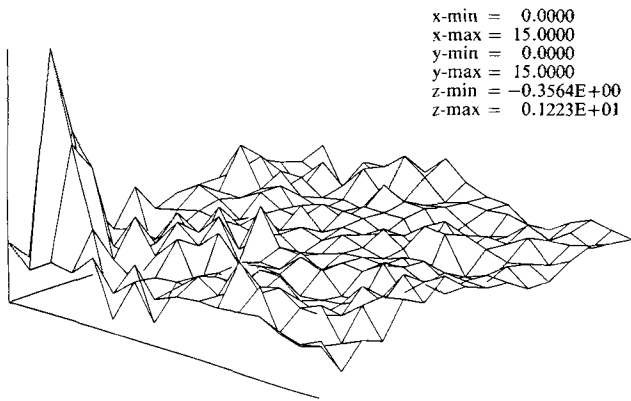z-min = −0.1021E+00
z-max = 0.5080E+00

FIGURE 22. 2nd-order kernel estimate for the 4th-order simulated system obtained via LSC.
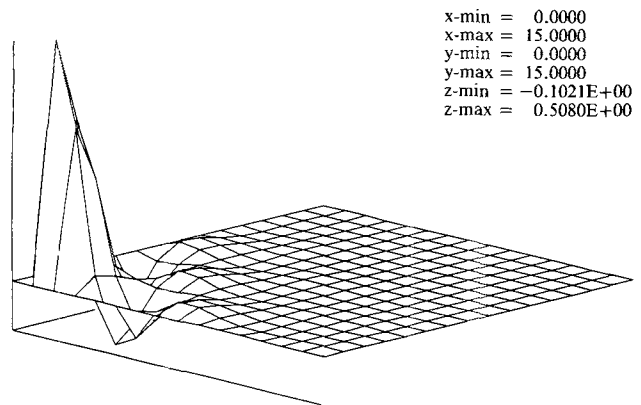
FIGURE 25. 2nd-order kernel estimate for non-white stimulus obtained via LEK.

x-min = 0.0000
x-max = 15.0000
y-min = 0.0000
y-max = 15.0000
z-min = −0.1876E+00
z-max = 0.6445E+00
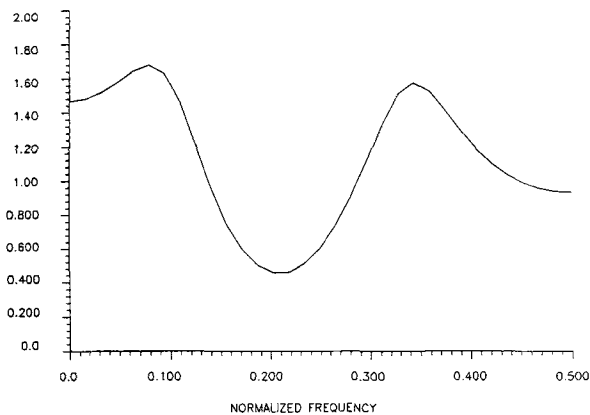
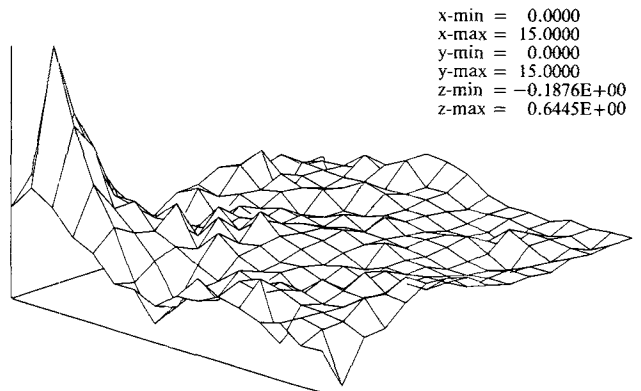FIGURE 23. Spectrum of non-white stimulus used in simulation.

FIGURE 26. 2nd-order kernel estimate for non-white stimulus obtained via LSC.
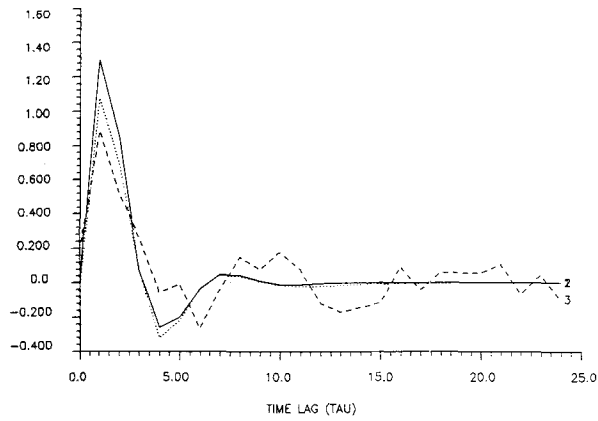
FIGURE 27. 1st-order kernel estimates for non-white stimulus and 4th-order system obtained via LEK (dotted) and LSC (dashed). The exact kernel is plotted with solid line.
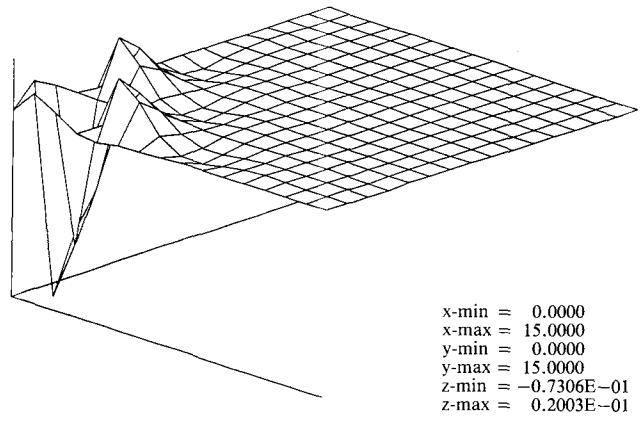

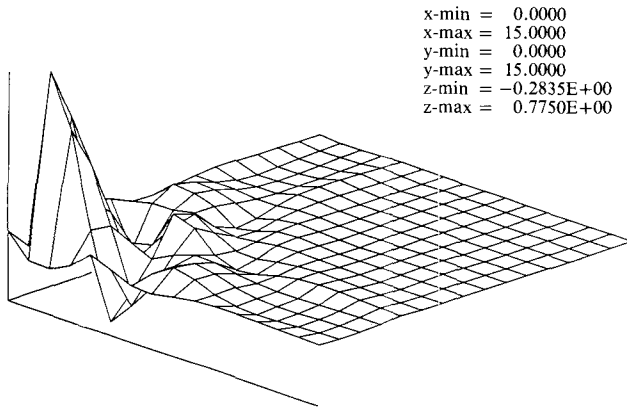FIGURE 30. 3rd-order kernel estimate (2-D "cut" at $\tau_3 = 4$) obtained via LEK.

x-min =   0.0000
x-max = 15.0000
y-min =   0.0000
y-max = 15.0000
z-min = $-0.7306E-01$
z-max =   0.2003E$-$01

x-min =   0.0000
x-max = 15.0000
y-min =   0.0000
y-max = 15.0000
z-min = $-0.2835E+00$
z-max =   0.7750E+00

FIGURE 28. 2nd-order kernel estimate for non-white stimulus and 4th-order system obtained via LEK.

x-min =   0.0000
x-max = 15.0000
y-min =   0.0000
y-max = 15.0000
z-min = $-0.7259E-01$
z-max =   0.1457E$-$01
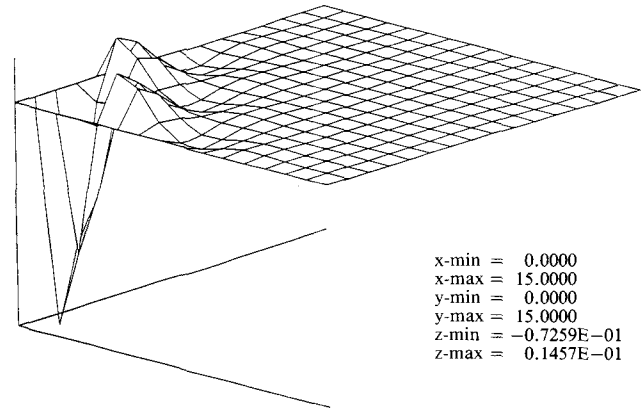
FIGURE 31. Exact 3rd-order kernel "cut" at $\tau_3 = 4$.

x-min =   0.0000
x-max = 15.0000
y-min =   0.0000
y-max = 15.0000
z-min = $-0.3677E+00$
z-max =   0.9611E+00
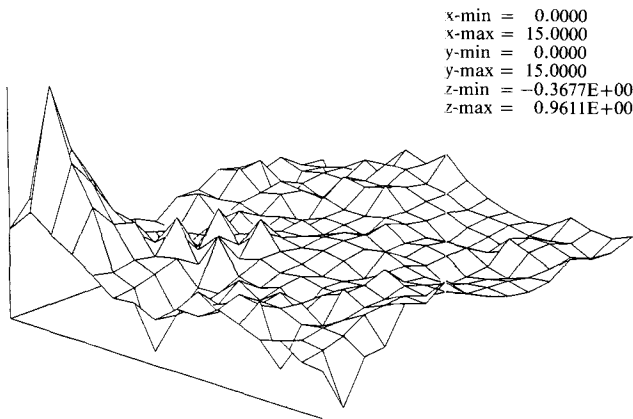
FIGURE 29. 2nd-order kernel estimate for non-white stimulus and 4th-order system obtained via LSC.

x-min =   0.0000
x-max = 15.0000
y-min =   0.0000
y-max = 15.0000
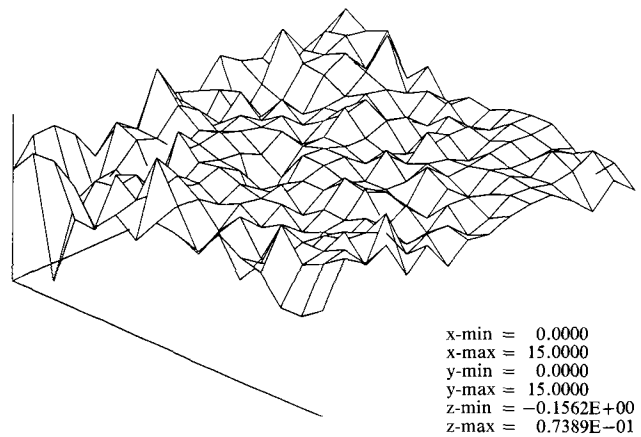z-min = $-0.1562E+00$
z-max =   0.7389E$-$01

FIGURE 32. 3rd-order kernel estimate (2-D "cut" at $\tau_3 = 4$) obtained via LSC.

REFERENCES

1. Barrett, J.F. The use of functionals in the analysis of non-linear physical systems. J. Electron. Control. 15:567–615; 1963.
2. Goussard, Y. Wiener kernel estimation: A comparison of cross-correlation and stochastic approximation methods. In: Advanced methods of physiological system modeling: Vol. I. Los Angeles, California: USC Biomedical Simulations Resource; 1987: pp. 289–302.
3. Korenberg, M.J. Functional expansions, parallel cascades and nonlinear difference equations. In: Marmarelis, V.Z., ed. Advanced methods of physiological system modeling: Volume I. Los Angeles: USC Biomedical Simulations Resource; 1987: pp. 221–240.
4. Korenberg, M.J. Identifying nonlinear difference equation and functional expansion representations: The fast orthogonal algorithm. Ann. Biomed. Eng. 16:123–142; 1988.
5. Lee, Y.W.; Schetzen, M. Measurement of the Wiener kernels of a nonlinear system by crosscorrelation. Int. J. Contr. 2:237–254; 1965.
6. Marmarelis, P.Z.; Marmarelis, V.Z. Analysis of physiological systems: The white-noise approach. New York: Plenum; 1978.
7. Marmarelis, V.Z., ed. Advanced methods of physiological system modeling: Vol. I. Los Angeles: USC Biomedical Simulations Resource; 1987.
8. Marmarelis, V.Z., ed. Advanced methods of physiological system modeling: Vol. II. New York: Plenum; 1989.
9. Ogura, H. Estimation of Wiener kernels of a nonlinear system and a fast algorithm using digital Laguerre filters. 15th NIBB Conference, Okazaki, Japan; 1985: pp. 14–62.
10. Rugh, W.J. Nonlinear system theory: The Volterra/Wiener approach. Baltimore: Johns Hopkins University Press; 1981.
11. Schetzen, M. The Volterra and Wiener theories of nonlinear systems. New York: Wiley; 1980.
12. Sun, H.H.; Shi, J.H. New algorithm for Korenberg-Billings model of nonlinear system identification. In: Advanced methods of physiological system modeling: Vol. II. New York: Plenum Publishing; 1989: pp. 179–200.
13. Sutter, E.E. A practical nonstochastic approach to nonlinear time-domain analysis. In: Advanced methods of physiological system modeling: Vol. I. Los Angeles, California: USC Biomedical Simulations Resource; 1987: pp. 303–315.
14. Victor, J.D. Nonlinear systems analysis: Comparison of white noise and sum of sinusoids in a biological system. Proc. Nat. Acad. Sci. 76:996–998; 1979.
15. Watanabe, A.; Stark, L. Kernel method for nonlinear analysis: Identification of a biological control system. Math. Biosci. 27:99–108; 1975.
16. Wiener, N. Nonlinear problems in random theory. New York: Wiley; 1958.