

[6] A. Segall, M. H. A. Davis, and T. Kailath, "Nonlinear filtering with counting observations," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 125-134, Mar. 1975.
 [7] B. D. O. Anderson and S. Chirarattananon, "Smoothing as an improvement on filtering: A universal bound," *Electron. Lett.*, vol. 7, no. 18, pp. 524, Sep. 1979.
 [8] B. D. O. Anderson and Doan B. Hoang, "Design procedure for stable sub-optimal fixed lag smoothers," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 949-952, Dec. 1977.
 [9] G. C. Goodwin, Doan B. Hoang, and A. Cantoni, "Application of ARMA models to automatic channel equalization," *Inform. Sci.*, vol. 22, pp. 107-129, 1980.
 [10] J. L. Doob, *Stochastic Processes*. New York, Wiley, 1953.
 [11] P. A. Meyer, "Decomposition of supermartingales: The uniqueness theorem," *Illinois J. Math.*, vol. 7, pp. 1-17, 1963.
 [12] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice Hall, 1979.
 [13] D. Clements and B. D. O. Anderson, "A nonlinear fixed lag smoother for finite state Markov processes," *IEEE Trans. Inform. Theory*, vol. IT-21, no. 4, pp. 446-452 1975.
 [14] M. V. Vaca and S. A. Tretter, "Optimal estimation for discrete time jump processes," *IEEE Trans. Inform. Theory*, vol. IT-24, no. 3, pp. 289-296, 1978.

Improved Maximum-Likelihood Detection and Estimation of Bernoulli-Gaussian Processes

CHONG-YUNG CHI MEMBER, IEEE, AND JERRY M. MENDEL, FELLOW, IEEE

Abstract—When a wavelet to be estimated is not spiky, then a single most likely replacement (SMLR) detector, which is used to detect randomly located impulsive events that have Gaussian-distributed amplitudes, may split a large spike into two smaller ones and may also detect some spikes at wrong locations, although these locations are very close to their true ones. Presented here are two new detection algorithms, namely a single-spike-shift (SSS) detector and an SSS-SMLR detector both of which help correct the SMLR detector's spike-splitting and shifting problem.

I. INTRODUCTION

Kormylo and Mendel [1] have developed a maximum-likelihood detector that produces locally optimal estimates. Their single most likely replacement (SMLR) detector adds or removes an event at any single time point in such a way that the likelihood function always decreases. The SMLR detector has been found to be very useful in reflection seismology, where the problem of pulse overlap can become extremely severe. Although the SMLR detector works well for various wavelets (i.e., channel impulse responses), we have found that it sometimes splits a large spike into smaller ones. Additionally, it sometimes detects spikes at incorrect locations, although these locations are very close to the correct ones. These cases often occur when the wavelet is not spiky (i.e., not broad-band).

In this correspondence, we present new detectors which can help to resolve overlapping nonspiky wavelets, and can increase the accuracy of spike detection.

In Section II we review the background of maximum-likelihood detection and estimation for Bernoulli-Gaussian processes. In

Manuscript received March 15, 1983; revised June 13, 1983. This work was performed at the University of Southern California, Los Angeles, and supported in part under NSF Grant ECS 8200882 and in part under the support of the sponsors of the USC Geo-Signal Processing Program.

C. Y. Chi was with the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089; he is now with the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109.

J. M. Mendel is with the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089-0781.

Section III we derive a single-spike-shift (SSS) detection algorithm. In Section IV we describe an SSS-SMLR detector that is a combination of the SSS and SMLR detectors. Computer simulations are presented in Section V which demonstrate that our SSS detector can help to improve the results obtained from the SMLR detector, and vice versa, and that the SSS-SMLR detector outperforms both the SMLR and SSS detectors.

II. BACKGROUND

As in Kormylo and Mendel [1], we begin with the discrete-time convolutional model

$$z(k) = \sum_{i=1}^k V(k-i)\mu(i) + n(k), \quad (1)$$

where $z(k)$ is observed data, $V(k)$ is the pulse wavelet, $\mu(k)$ is the impulse signal to be estimated, and $n(k)$ is observation noise. We assume that the data is to be processed off-line after N samples of $z(k)$ have been obtained. In (1) wavelet $V(k)$ is modeled as an n th-order autoregressive moving average (ARMA) with transfer function

$$V(z) = \frac{\sum_{i=1}^n \beta_i z^{-i+1}}{1 - \sum_{i=1}^n \alpha_i z^{-i}}. \quad (2)$$

Our detectors are based on Kalman filter/optimal smoother techniques; thus, we also need to express the convolutional model (1) in state-variable format as

$$x(k) = \Phi x(k-1) + \gamma \mu(k)$$

and

$$z(k) = h'x(k) + n(k),$$

where Φ , γ , and h are (known) functions of α_i and β_i (e.g., [2]). As in Kormylo and Mendel [1], the sequence $\mu(k)$ is modeled as a zero mean Bernoulli-Gaussian sequence, one which can be expressed as the product model $\mu(k) = r(k)q(k)$, where $r(k)$ is white Gaussian noise with variance C , and $q(k)$ is a Bernoulli sequence for which

$$\Pr \{q(k)\} = \begin{cases} 1 - \lambda, & q(k) = 0, \\ \lambda, & q(k) = 1. \end{cases}$$

The observation noise $n(k)$ is assumed to be white and Gaussian with variance R , and the sequences $r(k)$ and $q(k)$ are assumed to be independent.

Event detection consists of finding maximum-likelihood estimates $\hat{q}(k)$ of $q(k)$, $k = 1, 2, \dots, N$, and amplitude estimation consists of finding maximum-likelihood estimates $\hat{r}(k)$, of $r(k)$, $k = 1, 2, \dots, N$.

As in Kormylo and Mendel [1], we shall find it convenient during the derivation of our detectors, to express convolutional model (1) in matrix form as $z = V\mu + n$, where $z = \text{col}[z(1), z(2), \dots, z(N)]$,

$$V = \begin{bmatrix} V(0) & 0 & \dots & 0 \\ V(1) & V(0) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ V(N-1) & V(N-2) & \dots & V(0) \end{bmatrix},$$

$$\mu = \text{col}[\mu(1), \mu(2), \dots, \mu(N)],$$

and $n = \text{col}[n(1), n(2), \dots, n(N)]$. From the product model we see that $E\{\mu^2(k)|q(k)\} = Cq^2(k) = Cq(k)$; hence, the conditional covariance matrix for z is given by

$$E\{zz'|q\} \triangleq \Omega_q = VQ_qV' + RI \quad (3)$$

where $\mathbf{q} = \text{col}[q(1), q(2), \dots, q(N)]$, and

$$E\{\boldsymbol{\mu}\boldsymbol{\mu}'|\mathbf{q}\} \triangleq \mathbf{Q}_q = \text{diag}(Cq^2(1), Cq^2(2), \dots, Cq^2(N)) \\ = C \text{diag}(q(1), q(2), \dots, q(N)).$$

The SMLR detector, developed in [1], is an iterative search algorithm that compares the likelihood of a "reference" sequence \mathbf{q}_r to a limited number of different "test" sequences \mathbf{q}_t in each iteration. The SMLR detector was derived by assuming that \mathbf{q}_t differs from \mathbf{q}_r at only one location, so that there are then only N possible test sequences for a given reference sequence. The log-likelihood-ratio decision rule for choosing \mathbf{q}_r and \mathbf{q}_t is given by

$$2 \ln \Lambda_{tr}(k) = 2 \ln \frac{p(\mathbf{z}|\mathbf{q}_t) \Pr(\mathbf{q}_t)}{p(\mathbf{z}|\mathbf{q}_r) \Pr(\mathbf{q}_r)} \\ = \frac{(\mathbf{v}'_k \boldsymbol{\Omega}_r^{-1} \mathbf{z})^2}{C^{-1} [q_t(k) - q_r(k)]^{-1} + \mathbf{v}'_k \boldsymbol{\Omega}_r^{-1} \mathbf{v}_k} \\ - \ln \{1 + \mathbf{v}'_k \boldsymbol{\Omega}_r^{-1} \mathbf{v}_k C [q_t(k) - q_r(k)]\} \\ + 2 [q_t(k) - q_r(k)] \ln \left(\frac{\lambda}{1 - \lambda} \right) \frac{q_t}{q_r} \geq 0, \quad (4)$$

which is (33) in [1], where \mathbf{q}_t is the sequence

$$q_t(i) = \begin{cases} q_r(i), & \text{for all } i \neq k, \\ 1 - q_r(i), & i = k, \end{cases} \quad (5)$$

\mathbf{v}_k is the k th column of matrix \mathbf{V} , and $\boldsymbol{\Omega}_r = \boldsymbol{\Omega}_q|_{q=\mathbf{q}_r}$.

Let k' be associated with the maximum value of $\ln \Lambda(k) \triangleq \ln \Lambda_{tr}(k)$ ($k = 1, 2, \dots, N$). Then the single most likely replacement test sequence is

$$q'_t(i) = \begin{cases} q_r(i), & \text{for all } i \neq k', \\ 1 - q_r(i), & i = k'. \end{cases}$$

It is also true that the log-likelihood function evaluated for \mathbf{q}'_t is at least as large as its value evaluated for \mathbf{q}_r .

As pointed out by Kormylo and Mendel [1], the SMLR search algorithm, initiated by $\mathbf{q}_r = \hat{\mathbf{q}}^{(0)}$, computes N log-likelihood ratios corresponding to N different \mathbf{q}_t sequences. The most likely \mathbf{q}_t sequence is used as the reference sequence $\hat{\mathbf{q}}^{(1)}$ for the next iteration. If, after i iterations we obtain a reference $\mathbf{q}_r = \hat{\mathbf{q}}^{(i)}$ which is more likely than any of the corresponding \mathbf{q}_t sequences, then the search stops and $\hat{\mathbf{q}} = \hat{\mathbf{q}}^{(i)}$ is the final detected event sequence.

After SMLR detection is completed, amplitude estimates of the detected spikes can be obtained by using $\hat{\mathbf{q}}$ in the covariance model of an optimal smoother. Doing this, we obtain $\hat{r}(k|N)$, and subsequently $\hat{\boldsymbol{\mu}}(k|N) = \hat{q}(k)\hat{r}(k|N)$.

III. SINGLE-SPIKE-SHIFT (SSS) DETECTION

The SMLR detector is derived by assuming that \mathbf{q}_t and \mathbf{q}_r differ in just one location. In order to detect a better \mathbf{q} sequence, i.e., one with a higher likelihood function than that obtained by the SMLR detector, we shall derive a detector that computes the likelihood ratio of \mathbf{q}_t and \mathbf{q}_r when \mathbf{q}_t and \mathbf{q}_r differ at exactly two locations.

The total number of different \mathbf{q} sequences obtained from \mathbf{q}_r by changing any two location is $N(N-1)/2$, which can be a very large number. We shall only consider the special case where \mathbf{q}_t and \mathbf{q}_r differ at two consecutive locations. The SSS detector restricts \mathbf{q}_t to be generated from \mathbf{q}_r by shifting only at those k where there is a spike in \mathbf{q}_r . Such spikes are shifted one location forwards or backwards.

Let C_0 be the class of all possible test sequences generated by shifting a spike in \mathbf{q}_r forward or backward one location. We define a "bunch of spikes" as a set of consecutive spikes; thus, a bunch of spikes includes at least two consecutive spikes. We call

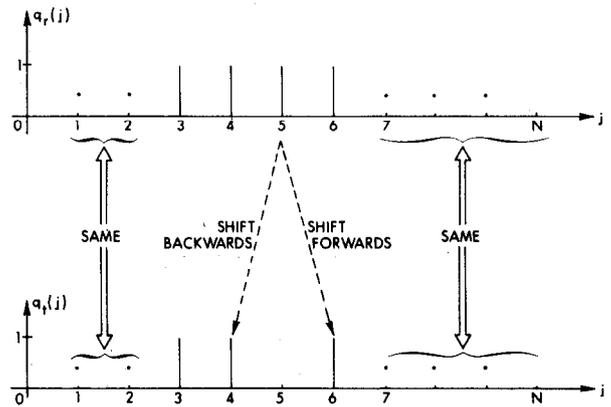


Fig. 1. Example of $\mathbf{q}_r \in C_1$.

a spike at $k = k_i$ an "isolated spike" if $q(k_i) = 1$ and $q(k_i - 1) = q(k_i + 1) = 0$. Assume that \mathbf{q}_r includes L isolated spikes and G bunches of spikes and that the i th bunch has g_i consecutive spikes for $i = 1, 2, \dots, G$. Let the L isolated spikes in \mathbf{q}_r be located at $k = k_1, k_2, \dots, k_L$ and the i th bunch contain spikes located at $k = l_i$ through $l_i + g_i - 1$ for $1 \leq i \leq G$; i.e.,

$$q_r(j) = \begin{cases} 1, & \text{for all } j = k_l, l = 1, 2, \dots, L, \text{ and} \\ & l_i \leq j \leq l_i + g_i - 1, i = 1, 2, \dots, G, \\ 0, & \text{otherwise.} \end{cases}$$

The class C_0 can be expressed as two mutually exclusive subclasses, C_1 and C_2 . Subclass C_1 includes those test sequences in which a change from \mathbf{q}_r to \mathbf{q}_t occurs at just one location,

$$C_1 = \{ \mathbf{q}_t | q_t(j) = q_r(j), \text{ for all } j \neq k, q_t(j) = 1 - q_r(j) \text{ for } j = k, k \in D_1 \}$$

where

$$D_1 = \{ k | l_i \leq k \leq l_i + g_i - 1, 1 \leq i \leq G \}.$$

Subclass C_1 includes all test sequences which are obtained by shifting a spike within any bunch in \mathbf{q}_r one location forwards or backwards. For example, assume that $l_1 = 3$ and $g_1 = 4$, i.e., the first bunch in \mathbf{q}_r is located at $k = 3$ through $k = 6$. Fig. 1 depicts this first bunch of \mathbf{q}_r . A test sequence \mathbf{q}_t obtained by shifting the spike at $k = 5$ one location forwards or backwards is depicted in Fig. 1. Obviously, for $k = 5$, \mathbf{q}_t is a member of C_1 . Note that end points $k = l_i$ and $l_i + g_i - 1$ are included in D_1 when these spikes are shifted into the bunch.

Subclass C_2 includes those test sequences in which a change from \mathbf{q}_r to \mathbf{q}_t occurs at two locations,

$$C_2 = \{ \mathbf{q}_t | q_t(j) = q_r(j) \text{ for all } j \neq k, k + 1, q_t(j) = 1 - q_r(j) \text{ for } j = k, k + 1, k \in D_2 \}$$

where

$$D_2 = \{ k | k = k_l, k_l - 1 \text{ for all } 1 \leq l \leq L; \\ \text{or } k = l_i - 1, l_i + g_i - 1 \text{ for all } 1 \leq i \leq G \}.$$

Note that endpoints $k = l_i - 1$ and $l_i + g_i - 1$ are included in D_2 when these spikes are shifted outside the bunch.

The total number of elements, in C_1 and C_2 , M_1 and M_2 , respectively are

$$M_1 = \sum_{i=1}^G g_i$$

and

$$M_2 = 2L + 2G.$$

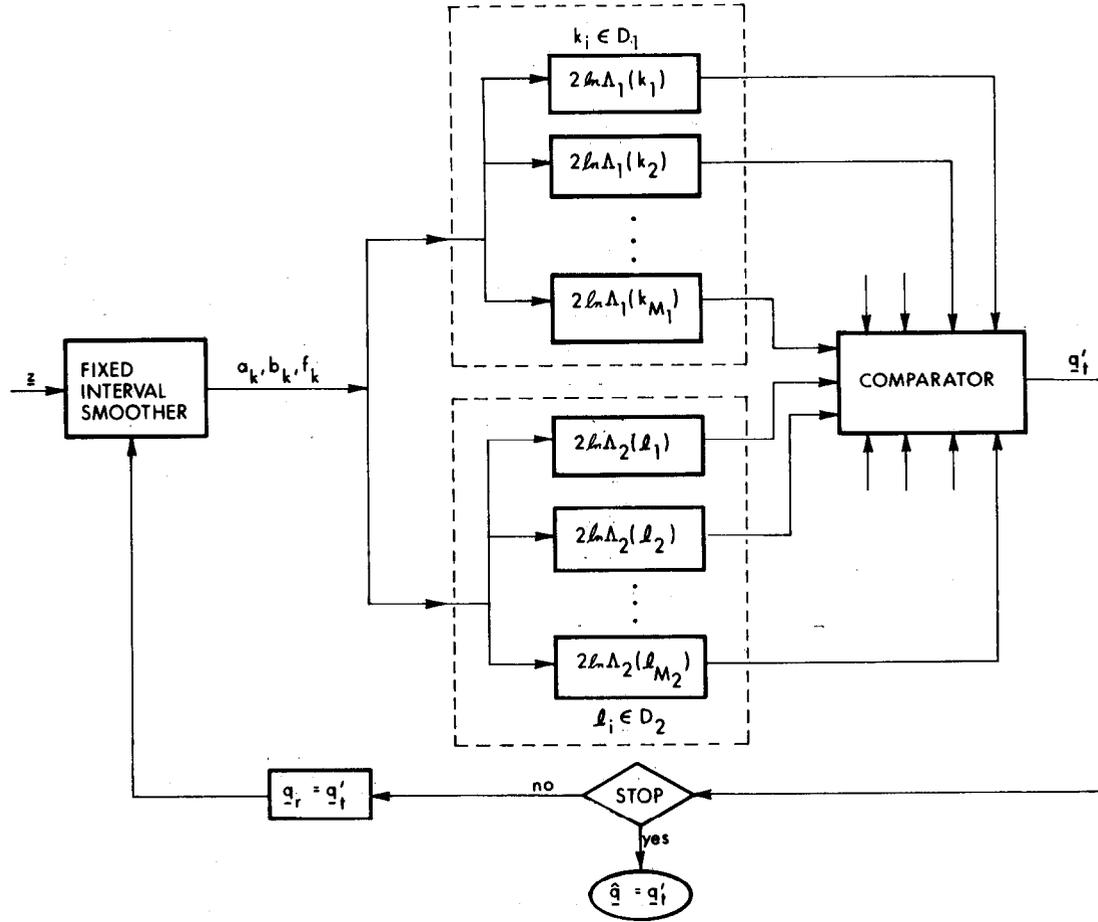


Fig. 2. SSS search algorithm.

For notational simplicity, we let $\ln \Lambda_1(k)$ denote the log-likelihood ratios when $q_t \in C_1$ and $\ln \Lambda_2(k)$ denote the log-likelihood ratios when $q_t \in C_2$. Observe that $\ln \Lambda_1(k)$ can be computed using (4) for all $k \in D_1$. On the other hand, $\ln \Lambda_2(k)$ must be computed using (7) stated in the following theorem for all $k \in D_2$.

Theorem: Assume that q_t and q_r differ at two consecutive locations, i.e.,

$$q_t(i) = \begin{cases} q_r(i), & \text{for all } i \neq k \text{ and } k+1, \\ 1 - q_r(i), & i = k \text{ and } k+1. \end{cases} \quad (6)$$

For this case

$$\begin{aligned} 2 \ln \Lambda_{tr}(k) = & \frac{C}{|A|} [d_k f_k^2 (1 + C d_{k+1} a_{k+1}) \\ & + d_{k+1} f_{k+1}^2 (1 + C d_k a_k) \\ & - 2 C d_k d_{k+1} b_k f_k f_{k+1}] - \ln |A| \\ & + 2(d_k + d_{k+1}) \ln \left(\frac{\lambda}{1 - \lambda} \right) \end{aligned} \quad (7)$$

where

$$\begin{aligned} A = & I + C \begin{bmatrix} v'_k \\ v'_{k+1} \end{bmatrix} \Omega_r^{-1} \begin{bmatrix} v_k & v_{k+1} \end{bmatrix} \begin{bmatrix} d_k & 0 \\ 0 & d_{k+1} \end{bmatrix} \\ = & \begin{bmatrix} 1 + C d_k a_k & C d_{k+1} b_k \\ C d_k b_k & 1 + C d_{k+1} a_{k+1} \end{bmatrix}, \end{aligned} \quad (8)$$

$$f_k \triangleq v'_k \Omega_r^{-1} z,$$

$$a_k \triangleq v'_k \Omega_r^{-1} v_k,$$

and

$$b_k \triangleq v'_k \Omega_r^{-1} v_{k+1}. \quad (11)$$

Quantities f_k , a_k , and b_k can be computed by running the following time-varying backwards state equation which is driven by the innovations process, $\tilde{z}(k|k-1)$ that is obtained from a Kalman filter:

$$r(k|N) = \Phi_b(k) r(k+1|N) + h \eta^{-1}(k) \tilde{z}(k|k-1), \quad (12)$$

where $k = N, N-1, \dots, 1$, $r(N+1|N) = 0$,

$$\Phi_b(k) = [I - K(k)h']\Phi', \quad (13)$$

$K(k)$ is the Kalman gain and $\eta(k)$ is the variance of $\tilde{z}(k|k-1)$. The covariance matrix of $r(k|N)$, $S(k|N)$, satisfies the backwards-recursive equation

$$S(k|N) = \Phi_b(k) S(k+1|N) \Phi_b'(k) + h \eta^{-1}(k) h' \quad (14)$$

where $k = N, N-1, \dots, 1$, and $S(N+1|N) = [0]$. From $r(k|N)$ and $S(k|N)$, we compute f_k , a_k , and b_k as

$$f_k = \gamma' r(k|N) \quad (15)$$

$$a_k = \gamma' S(k|N) \gamma \quad (16)$$

and

$$b_k = \gamma' \Phi_b(k) S(k+1|N) \gamma. \quad (17)$$

The proof of this theorem is given in the Appendix.

Let k' be associated with the maximum value of $\ln \Lambda_1(k)$ ($k \in D_1$) and $\ln \Lambda_2(j)$ ($j \in D_2$). Then the *single-spike-shift test*

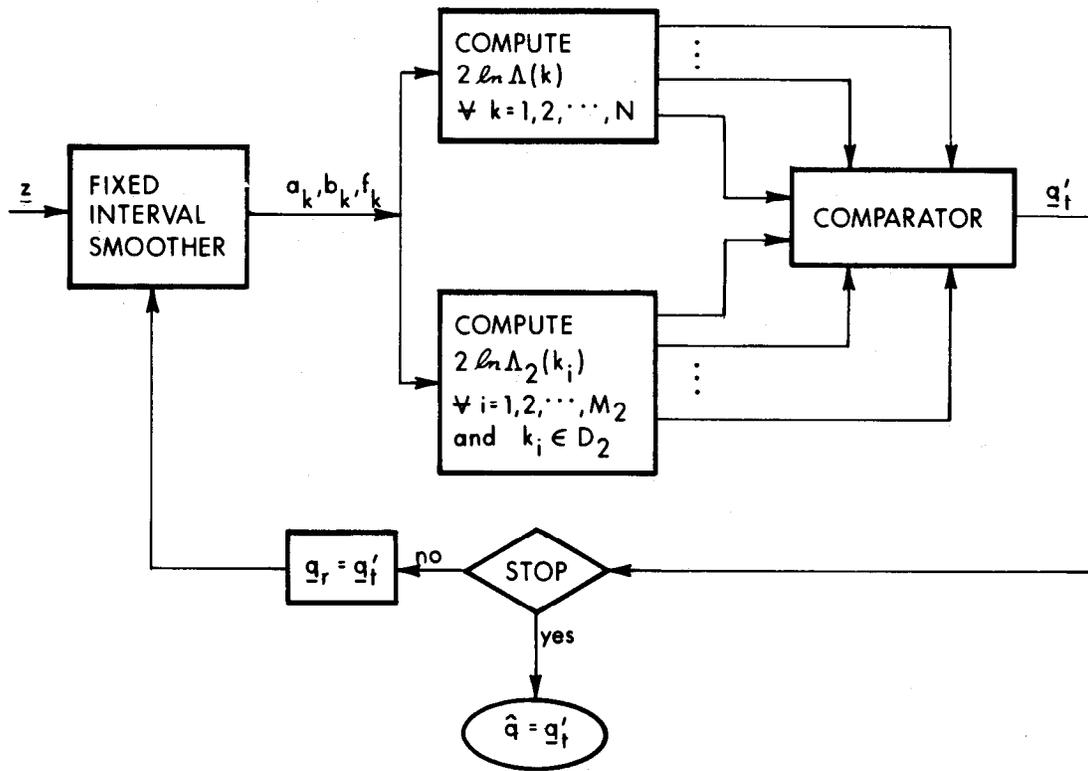


Fig. 3. SSS-SMLR search algorithm.

sequence is

$$q'_i(i) = \begin{cases} q_r(i), & \text{for all } i \neq k', \\ 1 - q_r(i), & i = k', \end{cases} \quad (18)$$

if $k' \in D_1$, or

$$q'_i(i) = \begin{cases} q_r(i), & \text{for all } i \neq k' \text{ and } k' + 1, \\ 1 - q_r(i), & i = k' \text{ and } k' + 1, \end{cases} \quad (19)$$

if $k' \in D_2$. It is also true that the log-likelihood function evaluated for q'_i is at least as large as its value evaluated for q_r .

Beginning with some initial reference $q_r = \hat{q}^{(0)}$, the SSS search algorithm computes M_1 log-likelihood ratios $\ln \Lambda_1(k)$ (for all $k \in D_1$), using (4), and M_2 log-likelihood ratios $\ln \Lambda_2(k)$ (for all $k \in D_2$), using (7) (see Fig. 2). The most likely q_r sequence is used as the reference sequence $\hat{q}^{(1)}$ for the next iteration. If, after i iterations we obtain a reference $q_r = \hat{q}^{(i)}$ which is more likely than any of the corresponding q_r sequences, then the reference sequence no longer changes so the search stops and $\hat{q} = \hat{q}^{(i)}$ is the final detected event sequence. If $\hat{q}^{(0)}$ is chosen to be \hat{q}_{SMLR} , which is obtained from the SMLR detector, then the likelihood function of $\hat{q}^{(i)}$ is surely larger than that of \hat{q}_{SMLR} when $\hat{q}^{(i)} \neq \hat{q}_{\text{SMLR}}$ for all $i \geq 1$.

In summary, the SSS detector is an iterative search algorithm that either shifts a spike forwards or backwards one location per iteration. It is suboptimal in that it may converge to some locally optimal sequence. The keys to this detector are the two expressions which allow us to compute $(M_1 + M_2)$ different log-likelihood ratios using only one optimal smoother (i.e., about two Kalman filters). The numbers M_1 and M_2 usually vary from iteration to iteration.

After SSS detection is completed, amplitude estimation of the detected spikes can be obtained by using \hat{q} as described at the end of Section II.

IV. SSS-SMLR DETECTOR

Once a_k, b_k, f_k are computed, we can calculate the log-likelihood ratios $2 \ln \Lambda(k)$ (for all $k = 1, 2, \dots, N$), $2 \ln \Lambda_1(k_i)$ (for all $i = 1, 2, \dots, M_1$, and $k_i \in D_1$) and $2 \ln \Lambda_2(l_i)$ (for all $i = 1, 2, \dots, M_2$ and $l_i \in D_2$). Doing this leads to another detector which updates a reference sequence q_r by choosing a test sequence that has the largest positive log-likelihood ratio among all computed log-likelihood ratios $2 \ln \Lambda(k)$, $2 \ln \Lambda_1(k)$, and $2 \ln \Lambda_2(k)$. We call this detector an SSS-SMLR detector. Because $\{2 \ln \Lambda_1(k_i), \text{ for all } i = 1, 2, \dots, M_1 \text{ and } k_i \in D_1\}$ are included in $\{2 \ln \Lambda(k), \text{ for all } k = 1, 2, \dots, N\}$, we can drop the calculations of the former terms from the SSS-SMLR detection algorithm.

At each iteration, assume that q'_i is the q sequence that is associated with the largest log-likelihood ratio $2 \ln \Lambda'(k)$. The SSS-SMLR detector either changes a spike location [i.e., q'_i is obtained from (5), and $2 \ln \Lambda'(k) \in \{2 \ln \Lambda(k), k = 1, 2, \dots, N\}$] or shifts a spike one location forwards or backwards [i.e., q'_i is obtained from (6), and $2 \ln \Lambda'(k) \in \{2 \ln \Lambda_2(k_i), i = 1, 2, \dots, M_2, k_i \in D_2\}$] until all computed log-likelihood ratios are less than zero, at which point a locally optimal spike sequence, \hat{q} , has been reached.

Fig. 3 depicts the SSS-SMLR search algorithm. When q'_i is obtained from (5), then the total number of spikes will be either $m'_i = m_r + 1$ or $m'_i = m_r - 1$, whereas when q'_i is obtained from (6), or $m'_i = m_r$. Thus, the total number of spikes at every iteration may or may not be changed.

After SSS-SMLR detection is completed, amplitude estimates of the detected spikes can be obtained by using \hat{q} as described at the end of Section II.

V. COMPUTER SIMULATIONS

We have run many computer simulations and have observed that spike splitting and spike shifting occur often when the

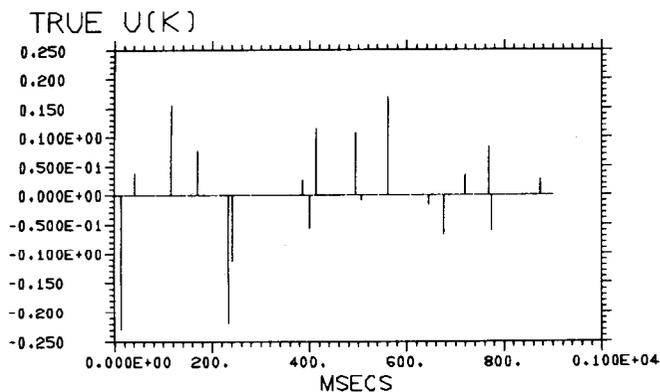


Fig. 4. Bernoulli-Gaussian sequence.

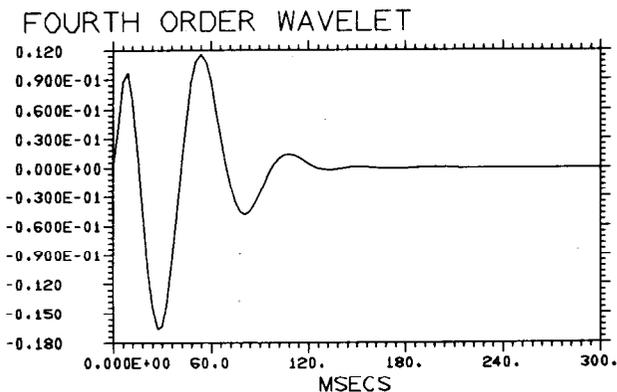


Fig. 5. Fourth-order ARMA wavelet.

SMLR detector is used, and the wavelet is not broad-band, and that both the SSS and SSS-SMLR detectors can help to improve the detection results obtained by the SMLR detector. Our results in this section are representative of all our simulations.

As an example, we consider a synthetic signal designed to fit our modeling assumptions. Using a pseudorandom number generator, we generated the Bernoulli-Gaussian sequence $\mu(k)$ for which $\lambda = 0.05$, $N = 300$, and $m = 18$, shown in Fig. 4. This signal was convolved with a fourth-order ARMA wavelet, shown in Fig. 5, to which white noise is added to produce the synthetic data shown in Fig. 6.

A threshold detector [2], [3] was used to obtain a starting \hat{q} sequence for both the SMLR and SSS-SMLR detectors. We then studied the three schemes depicted in Fig. 7. We denote the output of scheme 1 as $\hat{\mu}_1(k)$, the outputs of scheme 2 as $\hat{\mu}_{21}(k)$ and $\hat{\mu}_{22}(k)$, and the output of scheme 3 as $\hat{\mu}_3(k)$. Note that each output is an estimate of the input $\mu(k)$.

In Figs. 8 through 11, circles depict the true impulse signal $\mu(k)$ and bars depict the outputs $\hat{\mu}_1(k)$, $\hat{\mu}_{21}(k)$, $\hat{\mu}_{22}(k)$, $\hat{\mu}_3(k)$, respectively. Observing $\hat{\mu}_1(k)$ in Fig. 8, we see that the SMLR detector splits the first true spike (i.e., the first circle) into two smaller spikes, detects five true spikes (namely, the 6th, 8th, 9th, 14th, and 16th) but shifts them from their true locations, and gives rise to two false alarms. In Fig. 9, which depicts $\hat{\mu}_{21}(k)$, we see that the SSS detector recovers the first spike, detects the 8th and 9th true spikes, and gives rise to only one false alarm. Only three detected spikes in $\hat{\mu}_{21}(k)$ are shifted from their true locations the 6th, 14th, and 16th. In Fig. 10, which depicts $\hat{\mu}_{22}(k)$, we see that the cascaded SSS and SMLR detectors eliminate false alarms; but the 6th, 14th, and 16th detected spikes still remain

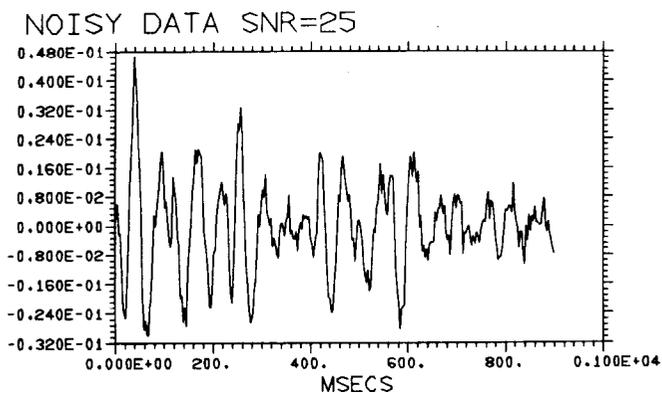


Fig. 6. Synthetic data (signal-to-noise ratio of 25).

shifted. In Fig. 11, which depicts $\hat{\mu}_3(k)$, we see that the SSS-SMLR detector has shifted the 6th spike to its correct location, and only the 14th and 16th detected spikes remain shifted from their true locations.

From these results, we see the SSS detector can improve SMLR detector results, vice versa when they are cascaded together, and the SSS-SMLR works well. Another experiment was performed in which we used $\hat{\mu}_{22}(k)$ as the starting sequence for a second cascade of SSS and SMLR detectors. We were able to improve $\hat{\mu}_{22}(k)$, and, in fact we obtained the same sequence as $\hat{\mu}_3(k)$. Doing this again, we did not improve our results. Apparently, $\hat{\mu}_3(k)$ is a locally optimal estimate of $\mu(k)$ and neither the SSS- or SMLR-detectors can find another \hat{q} sequence such that \hat{q} has a higher likelihood function than \hat{q}_3 associated with $\hat{\mu}_3(k)$.

Finally, observe from Figs. 8 through 11, that better spike location information results in much better estimates of amplitudes.

VI. DISCUSSION AND CONCLUSION

When a wavelet is not spiky then an SMLR detector may split a large spike into two smaller ones and detect some spikes at wrong locations, although the detected locations are very close to their true locations.

We have derived an SSS detector which is based on a maximum-likelihood criterion. Its derivative is very similar to that of the SMLR detector. The SSS detector can help to correct the SMLR detector's spike-splitting and spike shifting problems. All quantities needed to implement the SSS detector can be obtained from one optimal smoother. We also developed an SSS-SMLR detector, which is a combination of SSS and SMLR detectors.

The SSS and SSS-SMLR detection algorithms are iterative. At every iteration they increase the likelihood function $p(z|\hat{q})\Pr(\hat{q})$ until a local maximum of $p(z|\hat{q})\Pr(\hat{q})$ is reached, i.e., they converge to a local maximum of $p(z|\hat{q})\Pr(\hat{q})$; they are, therefore, suboptimal. Our algorithms do guarantee that a local maximum of the likelihood function can be found and that this value will be larger than that obtained by the SMLR detector. Their performance depends on signal-to-noise ratio and bandwidth of the wavelet. Quantitative relationships between their performance and signal-to-noise ratio and bandwidth of the wavelet remain to be developed.

In [8] Kwakernaak computed the likelihood function using results that are familiar from matched filtering. Some approximations were made by him. In this correspondence (and [1]), instead of computing the likelihood function, the SSS-SMLR detector, for example, computes $(N + M_2)$ log-likelihood ratios. To do this, we run an optimal smoother once, but do not make any

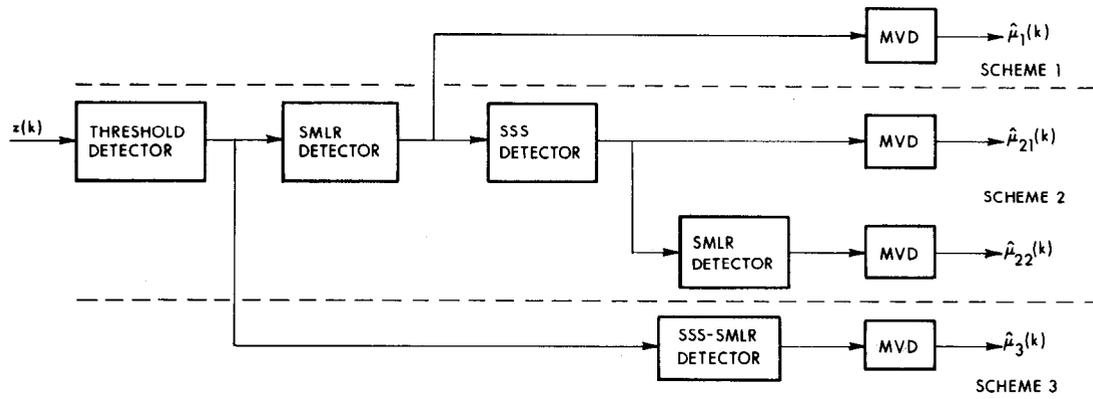


Fig. 7. Simulation schemes.

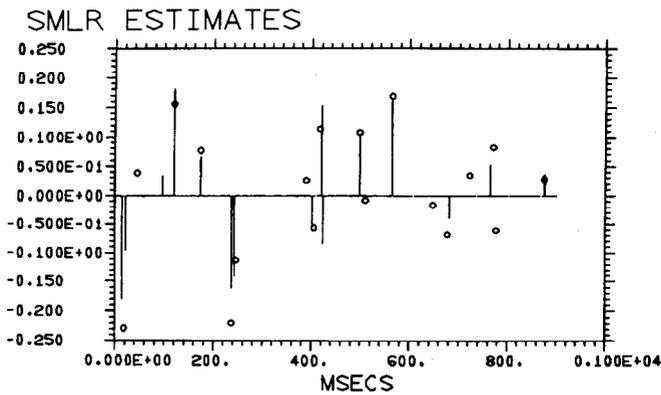


Fig. 8. Output $\hat{\mu}_1(k)$ from scheme 1.

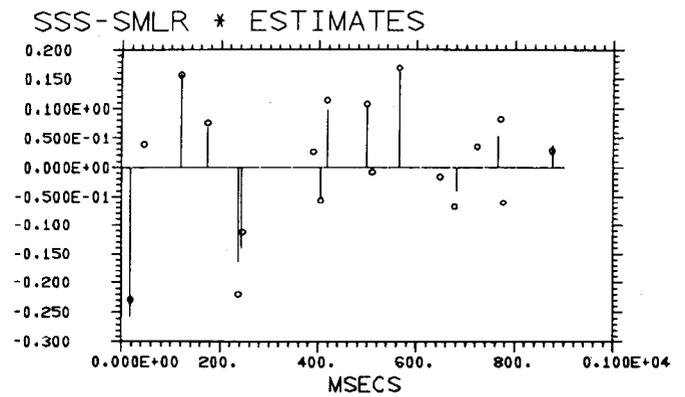


Fig. 10. Output $\hat{\mu}_{22}(k)$ from scheme 2.

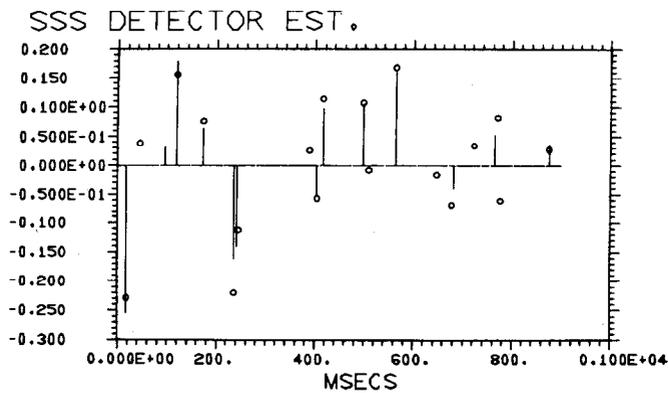


Fig. 9. Output $\hat{\mu}_{21}(k)$ from scheme 2.

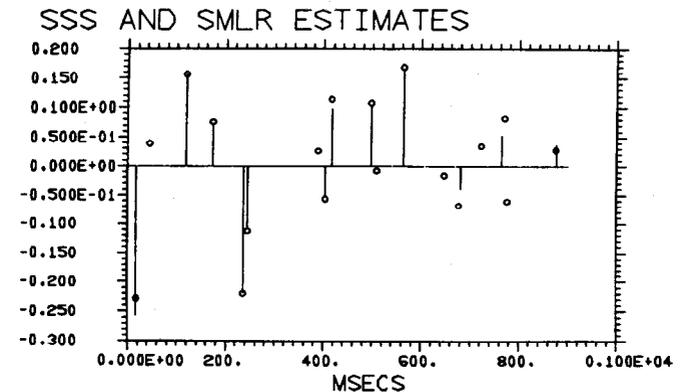


Fig. 11. Output $\hat{\mu}_3(k)$ from scheme 3.

approximations. In other words, we compute $N + M_2$ likelihood functions by running an optimal smoother once instead of using matched filtering approaches $N + M_2$ times. This leads to a great economy in the numerical effort. We then estimate spike amplitudes after detection is completed. The computational load for estimating detected spike amplitudes is equivalent to that for one iteration of detection; thus, amplitude estimation is not computationally burdensome.

Our computer simulations demonstrated that the SSS and SSS-SMLR detectors work quite well; that the SSS detector can help to improve the results obtained from the SMLR detector and vice versa; and, that the SSS-SMLR detector outperforms both the SMLR and SSS detectors.

APPENDIX

Proof of the Theorem

When q is given, then z is Gaussian; hence,

$$p(z|q) = \frac{1}{(2\pi)^{N/2} |\Omega_q|^{1/2}} \exp \left\{ -\frac{1}{2} z' \Omega_q^{-1} z \right\}. \quad (A1)$$

Because the $q(k)$ are independently distributed, we see from (6) that

$$\Pr(q) = \prod_{k=1}^N \Pr(q(k)) = \lambda^m q (1 - \lambda)^{N-m_q} \quad (A2)$$

where m_q is the number of nonzero events in q , i.e.,

$$m_q = \sum_{k=1}^N q(k). \quad (\text{A3})$$

Let U be a unitary matrix such that

$$U'(\mathcal{Q}_t - \mathcal{Q}_r)U = \begin{bmatrix} \mathcal{Q}_{t_1} - \mathcal{Q}_{r_1} & [0] \\ \hline [0] & [0] \end{bmatrix}, \quad (\text{A4})$$

where $\mathcal{Q}_t = \mathcal{Q}_t|_{q=q_t}$, $\mathcal{Q}_r = \mathcal{Q}_r|_{q=q_r}$, and $(\mathcal{Q}_{t_1} - \mathcal{Q}_{r_1})$ is diagonal and invertible. Additionally, let $VU = [V_1; V_2]$.

Kormylo and Mendel [1], [2] proved that the log-likelihood ratio for the test sequence q_t and the reference q_r can be expressed as

$$\begin{aligned} 2 \ln \Lambda_{tr} = & z' \Omega_r^{-1} V_1 (\mathcal{Q}_{t_1} - \mathcal{Q}_{r_1}) \\ & \cdot [I + V_1' \Omega_r^{-1} V_1 (\mathcal{Q}_{t_1} - \mathcal{Q}_{r_1})]^{-1} V_1' \Omega_r^{-1} z \\ & - \ln |I + V_1' \Omega_r^{-1} V_1 (\mathcal{Q}_{t_1} - \mathcal{Q}_{r_1})| \\ & + 2(m_t - m_r) \ln \left(\frac{\lambda}{1 - \lambda} \right), \end{aligned} \quad (\text{A5})$$

where $m_t = m_q|_{q=q_t}$, $m_r = m_q|_{q=q_r}$, $\Omega_t = \Omega_q|_{q=q_t}$, and $\Omega_r = \Omega_q|_{q=q_r}$.

When q_t is defined as in (6), the matrix V_1 in (A5) becomes

$$V_1 = [v_k \quad v_{k+1}]. \quad (\text{A6})$$

Let $d_i \triangleq q_t(i) - q_r(i)$; thus,

$$m_t - m_r = \sum_{i=1}^N q_t(i) - \sum_{i=1}^N q_r(i) = d_k + d_{k+1} \quad (\text{A7})$$

and

$$\begin{aligned} \mathcal{Q}_{t_1} - \mathcal{Q}_{r_1} &= C \begin{bmatrix} q_t(k) & 0 \\ 0 & q_t(k+1) \end{bmatrix} - C \begin{bmatrix} q_r(k) & 0 \\ 0 & q_r(k+1) \end{bmatrix} \\ &= C \begin{bmatrix} d_k & 0 \\ 0 & d_{k+1} \end{bmatrix}. \end{aligned} \quad (\text{A8})$$

The right-hand side of (A5) can now be computed by substituting (A6), (A7), and (A8) into it:

$$\begin{aligned} 2 \ln \Lambda_{tr}(k) = & [d_k f_k \quad d_{k+1} f_{k+1}] A^{-1} \begin{bmatrix} f_k \\ f_{k+1} \end{bmatrix} \\ & - \ln |A| + 2(d_k + d_{k+1}) \ln \left(\frac{\lambda}{1 - \lambda} \right), \end{aligned} \quad (\text{A9})$$

where matrix A and quantities f_k , a_k , and b_k are defined in (8) through (11), respectively. From (8), we see that

$$|A| = (1 + Cd_k a_k)(1 + Cd_{k+1} a_{k+1}) - C^2 d_k d_{k+1} b_k^2 \quad (\text{A10})$$

and

$$A^{-1} = \frac{1}{|A|} \begin{bmatrix} 1 + Cd_{k+1} a_{k+1} & -Cd_{k+1} b_k \\ -Cd_k b_k & 1 + Cd_k a_k \end{bmatrix}. \quad (\text{A11})$$

Substituting (A11) into (A9), we obtain (7).

Next, we demonstrate the truth of (15) through (17). Kormylo and Mendel [1] showed that the minimum-variance estimate of

$\mu(k)$, given z , is

$$\hat{\mu}(k|N) = Cq(k)\gamma'r(k|N), \quad (\text{A12})$$

and,

$$\hat{\mu}(k|N) = Cq(k)v_k' \Omega_q^{-1} z, \quad (\text{A13})$$

$f_k = \gamma'r(k|N)$, and $a_k = \gamma'S(k|N)\gamma$. We therefore only have to demonstrate the truth of (17).

Evaluate $E\{\hat{\mu}(k|N)\hat{\mu}'(k+1|N)\}$ from (A12) and (A13), respectively, as

$$\begin{aligned} E\{\hat{\mu}(k|N)\hat{\mu}'(k+1|N)\} \\ = C^2 q(k)q(k+1)\gamma'E[r(k|N)r'(k+1|N)]\gamma. \end{aligned} \quad (\text{A14})$$

and

$$E\{\hat{\mu}(k|N)\hat{\mu}'(k+1|N)\} = C^2 q(k)q(k+1)b_k. \quad (\text{A15})$$

Compare (A14) and (A15), to see that $b_k = \gamma'E\{r(k|N)r'(k+1|N)\}\gamma$ where, from (12),

$$\begin{aligned} E\{r(k|N)r'(k+1|N)\} \\ = \Phi_b(k)S(k+1|N) + h\eta^{-1}(k)E\{\bar{z}(k|k-1)r'(k+1|N)\}. \end{aligned} \quad (\text{A16})$$

Observe from (12), that $r(k+1|N)$ is a linear combination of $\bar{z}(k+1|k)$, $\bar{z}(k+2|k+1)$, \dots , $\bar{z}(N|N-1)$; thus,

$$\begin{aligned} E\{\bar{z}(k|k-1)r'(k+1|N)\} \\ = E\{\bar{z}(k|k-1)\}E\{r'(k+1|N)\} = \mathbf{0}. \end{aligned} \quad (\text{A17})$$

We conclude that $b_k = \gamma'\Phi_b(k)S(k+1|N)\gamma$.

REFERENCES

- [1] J. Kormylo and J. M. Mendel, "Maximum-likelihood detection and estimation of Bernoulli-Gaussian processes," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 482-488, May, 1982.
- [2] J. M. Mendel, *Optimal Seismic Deconvolution: An Estimation-Based Approach*, New York: Academic, 1983.
- [3] J. Kormylo, "Maximum-likelihood seismic deconvolution," Ph.D. dissertation, Univ. of Southern California, Los Angeles, 1979.
- [4] J. M. Mendel, "Minimum-variance deconvolution," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-19, pp. 161-171, July 1981.
- [5] J. M. Mendel and J. Kormylo, "New fast optimal white-noise estimators for deconvolution," *IEEE Trans. Geosci. Electron.*, vol. GE-15, no. 1, pp. 32-41, Jan. 1977.
- [6] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ: Prentice Hall, 1979.
- [7] J. S. Meditch, *Stochastic Optimal Linear Estimation and Control*. New York: McGraw-Hill, 1969.
- [8] H. Kwakernaak, "Estimation of pulse heights and arrival times," *Automatica*, vol. 16, pp. 367-377, 1980.

Codes for Zero Spectral Density at Zero Frequency

GIANFRANCO L. PIEROBON

Abstract—In pulse-amplitude modulation (PAM) digital transmission systems line encoding is used for shaping the spectrum of the encoded symbol sequence to suit the frequency characteristics of the transmission

Manuscript received May 21, 1982; revised April 4, 1983. This work was supported by the Italian National Council of Research (CNR) under Grant CT81.02374.07.

The author is with the Department of Electrical Engineering, University of Padova, Padova, Italy.