

XXVII. DETECTION AND ESTIMATION THEORY*

Academic and Research Staff

Prof. H. L. Van Trees
Prof. D. L. Snyder

Graduate Students

M. E. Austin	L. D. Collins	R. R. Kurth
A. B. Baggeroer	T. J. Cruise	A. P. Tripp, Jr.

RESEARCH OBJECTIVES AND SUMMARY OF RESEARCH

The work of this group may be divided into four major areas.

1. Sonar

The central problem of interest is the development of effective processing techniques for the output of an array with a large number of sensors. Some specific topics of interest in this connection are the following.

(i) A state-variable formulation for the waveform estimation problem when the signal is a sample function from a possibly nonstationary random process that has passed through a dispersive medium before arriving at the array. Some preliminary results have been obtained.¹

(ii) The effects of array velocity on the system performance when operating in a reverberation-limited environment.²

(iii) Iterative techniques to measure the interfering noise field and modify the processor to reduce its effect.

(iv) A hardware-efficient analog transversal equalizer has been designed, and the basic components built and tested. With some additional work, a completed transversal filter will enable us to apply actual channel measurements to its time-variable tap gains, thus making possible laboratory simulations of dispersive channels. Such simulations are expected to provide a convenient means of evaluating channel parameter estimation techniques, and to prove useful in the development of better models of the sonar channel.

2. Communications

a. Digital Systems

Decision-feedback systems offer an effective method for improving digital communications over dispersive channels. A new structure has been derived whose performance should be appreciably better than previous systems. Theoretical work on this decision-feedback scheme will continue; it will be viewed as an integral part of an adaptive receiver whose acquisition and tracking behavior are of interest in equalization of dispersive channels. The analog transversal filter that is being constructed as a channel-measuring and channel-simulation device (see (iv) above) will also prove useful in evaluating the performance of algorithms that have been developed for adjusting the parameters of these adaptive receivers.

*This work was supported by the Joint Services Electronics Programs (U.S. Army, U.S. Navy, and U.S. Air Force) under Contract DA 36-039-AMC-03200(E)).

(XXVII. DETECTION AND ESTIMATION THEORY)

The optimum receiver for the detection of Gaussian signals in Gaussian noise is well known. Except for limiting cases it is difficult to evaluate the error behavior. Work continues on developing performance measures for evaluating the performance, emphasizing techniques that are computationally tractable and, at the same time, give a good measure of the system performance. Both tight upper bounds and computational algorithms have been developed for the probability of error which emphasize the fundamental role of optimum linear systems in detection problems. Future work includes the application of these techniques to the analysis and design of radar, sonar, and communication systems.

The availability of a noiseless feedback channel from receiver-to-transmitter enables a significant increase in performance. By utilizing the continuous feedback signal at the modulator, the behavior system performance at the end of the transmission interval is greatly improved. The feedback link could be used to obtain the same performance over a shorter transmission interval. The actual structure of the system is very flexible and simple.

Noise in the feedback channel degrades the achievable system performance with the simple feedback system. Improvement over the no-feedback system is obtained, but it is not as dramatic as when noiseless feedback is available.

b. Analog Systems

(i) Investigations of the performance of analog modulation systems operating in additive noise channels are essentially completed.³

(ii) When a noiseless feedback link is available from receiver-to-transmitter simple modulation schemes can be developed which achieve the rate-distortion bound. Realizable feedback systems perform very close to the rate-distortion bound.

The effects of additive noise in the feedback link depend on the relative noise levels in the two channels. For relatively small feedback channel noise, the system performance is close to the rate-distortion bound. For large feedback noise the availability of a feedback link does not significantly improve the system performance.

(iii) A new approach has been developed for estimating continuous waveforms in real time. The approach is formulated with continuous Markov processes and use is made of state-variable concepts. The approach has been applied successfully to the problem of estimating continuous stochastic messages transmitted by various linear and nonlinear modulation techniques over continuous random channels; particular emphasis has been given to phase and frequency modulation. An advantage of this approach over alternative schemes is that it leads automatically to physically realizable demodulators that can be readily implemented.⁴

3. Seismic

A substantial portion of our effort in this area is devoted to acquiring an adequate understanding of geophysics in order to formulate meaningful problems. An area of concentration is exploration seismology in land and ocean environments. Some specific problems of interest include array design and effective quantization techniques.

4. Random Process Theory and Application

a. State-Variable and Continuous Markov Process Techniques

(i) In the theory of signal detection and estimation, it is frequently of interest to determine the solutions to a Fredholm integral equation. A state-variable approach to the problem of determining the eigenfunctions and eigenvalues associated with the problem has been formulated.

(XXVII. DETECTION AND ESTIMATION THEORY)

The random process(es) is represented as the output of a linear dynamic system that is described by a state equation. The Fredholm integral equation is reduced to a vector differential equation that is directly related to the state equation of the dynamic system. From this equation, a determinant is found which must vanish in order that an eigenvalue exist. Once the eigenvalue is found, the eigenfunction follows from the transition matrix of the vector differential equation.

The technique is general enough to handle a large class of problems. Constant-parameter (possibly nonstationary) dynamic systems for both scalar and vector processes can be handled in a straightforward and analytic manner. Time-varying systems can also be treated by computational techniques.⁵

(ii) The problem of formulating a state-variable model for random channels encountered in practice is being investigated. A particular class of channels of interest are those exhibiting frequency-selective fading.

(iii) The system identification problem is being studied. Applications include measurement of noise fields, random process statistics, and linear system functions.

b. Detection Techniques

Various extensions of the conventional detection problem to include nonparametric techniques, sequential tests, and adaptive systems are being studied.

H. L. Van Trees

References

1. A. Baggeroer, "MAP State-Variable Estimation in a Dispersive Environment," Internal Memorandum, December 10, 1966 (unpublished).
2. A. P. Tripp, Jr., "Effects of Array Velocity on Sonar Array Performance," S. M. Thesis, Department of Electrical Engineering, M. I. T., 1966.
3. T. L. Rachel, "Optimum F. M. System Performance - An Investigation Employing Digital Simulation Techniques," S. M. Thesis, Department of Electrical Engineering, M. I. T., 1966.
4. D. L. Snyder, "The State-Variable Approach to Continuous Estimation," Ph.D. Thesis, Department of Electrical Engineering, M. I. T., 1966.
5. A. B. Baggeroer, "A State-Variable Technique for the Solution of Fredholm Integral Equation" (submitted to IEEE Transactions on Information Theory).

A. EQUALIZATION OF DISPERSIVE CHANNELS USING DECISION FEEDBACK

1. Introduction

Several authors have recently considered decision feedback as a means of improving digital communication over dispersive channels. Lucky¹ applied decision feedback to the equalization of telephone lines, effectively enabling him to achieve dispersion measurements via the message sequence, rather than having to send additional signals to "sound" the channel before information transmission. Aein and Hancock² have analyzed the performance of a nonlinear decision-feedback receiver, applicable to channels in which the dispersion is restricted to less than two baud durations. Drouilhet and

(XXVII. DETECTION AND ESTIMATION THEORY)

Neissen³ have studied and are constructing a decision-feedback equalizer, differing from that of Lucky in that they use a matched-filter ahead of their equalization filter, and use decision feedback to "subtract out" the effects of the message in order to improve their continuous, "sounding-signal" channel measurement.

The equalizer structure considered by Lucky is the conventional tapped delay line (TDL), which has long been used in correcting telephone lines, while Drouilhet and Niessen are using the MF-TDL structure, arrived at through different approaches and criteria by several persons (for example, Tufts,⁴ George,⁵ and Austin⁶). Both efforts are examples of how decision feedback has been used to facilitate channel measurement and the adjustment of parameters of what we shall refer to henceforth as "conventional" equalizers. We shall show in this report that decision feedback can be used to additional advantage in channel equalization, if one adopts the new equalizer structure developed in the sequel, hereafter referred to as the "decision-feedback" equalizer to distinguish it from the conventional equalizers mentioned above.

We want to consider the problem of determining the structure of a receiver for digital communication over a linear dispersive channel whose equivalent impulse response, $h(t)$, is known. The receiver input is

$$r(t) = \sum_{k=-\infty}^{\infty} \xi_k h(t-kT_b) + n(t), \quad (1)$$

where T_b is the baud duration, and ξ_k contains the information transmitted on the k^{th} baud. Our problem is to decide between the hypothesis H_0 that $\xi_0 = +1$ and hypothesis H_1 that $\xi_0 = -1$. We make the following assumptions.

- (i) The ξ_k are independent.
- (ii) H_0 and H_1 are equally likely.
- (iii) $n(t)$ is white Gaussian noise, $R_n(\tau) = \frac{N_0}{2} \delta(\tau)$.

At this point we would like to derive the optimal receiver structure, assuming only that $\xi_k = +1$ or -1 on each baud. This, however, proves analytically intractable, and therefore we shall adopt an approach leading to a suboptimal receiver structure, which nonetheless exhibits definite advantages over existing conventional equalizer structures. This requires us to make additional assumptions.

- (iv) The ξ_k are $N(0, \sigma)$ for $k > 0$.
- (v) The ξ_k are known for $k < 0$.

We note that assumption (iv) renders our model inaccurate for binary AM or PSK systems in which $\xi_k = +1$ or -1 for all k , while assumption (v) is only valid for decision-feedback equalizers in the absence of decision errors.

Using these assumptions, we shall be led to the desired suboptimal receiver structure

(Fig. XXVII-4); then we shall optimize its parameters for the binary communication problem of interest (Eqs. 5 and 6). Finally, we shall work a simple example demonstrating that our new decision-feedback equalizer is capable of rendering far better performance than the conventional MF-TDL receiver.

2. Structure of the Decision-Feedback Equalizer

We want to determine the optimum receiver structure for the problem and assumptions that have been introduced, but first we introduce some notation and definitions that will prove useful in the derivation that is to follow.

- Definitions: 1. $\underline{\xi}^- = \{\xi_k | k < 0\}$
2. $\underline{\xi}^+ = \{\xi_k | k > 0\}$
3. $a_k = \frac{1}{N_o} \int r(t) h(t - kT_b) dt = \frac{1}{N_o} \int r(t + kT_b) h(t) dt$
4. $b_k = \frac{1}{N_o} \int h(t) h(t - kT_b) dt.$

Under our assumptions, $\underline{\xi}^-$ is known correctly via decision feedback, and with this taken into account, the optimum receiver calculates the likelihood ratio

$$\Lambda = \frac{p[r(t) | \underline{\xi}^-, H_0]}{p[r(t) | \underline{\xi}^-, H_1]} = \frac{\int p[r(t) | \underline{\xi}^-, H_0, \underline{\xi}^+] p(\underline{\xi}^+) d\underline{\xi}^+}{\int p[r(t) | \underline{\xi}^-, H_1, \underline{\xi}^+] p(\underline{\xi}^+) d\underline{\xi}^+}.$$

We shall first consider the numerator. Under assumption (v) we have

$$p[r(t) | \underline{\xi}^-, H_0, \underline{\xi}^+] = K_1 \exp \left[-\frac{1}{N_o} \int \left\{ r(t) - \sum_{k < 0} \xi_k h(t - T_b) - \xi_0 h(t) - \sum_{\ell > 0} \xi_\ell h(t - \ell T_b) \right\}^2 dt \right] \Bigg|_{\xi_0 = 1}$$

and since

$$K_1 \exp \left[-\frac{1}{N_o} \int \left\{ r(t) - \sum_{k < 0} \xi_k h(t - kT_b) \right\}^2 dt \right]$$

(XXVII. DETECTION AND ESTIMATION THEORY)

is independent of $\underline{\xi}^+$ and ξ_o , it will factor out of the integral and cancel with the same term arising in the denominator of Λ . Thus the terms of importance remaining in the integrand of the numerator are

$$\exp \left[\frac{2}{N_o} \int \left\{ r(t) - \sum_{k < 0} \xi_k h(t - kT_b) \right\} \left\{ \xi_o h(t) + \sum_{\ell > 0} \xi_\ell h(t - \ell T_b) \right\} dt \right. \\ \left. - \frac{1}{N_o} \int \left\{ \xi_o h(t) + \sum_{\ell > 0} \xi_\ell h(t - \ell T_b) \right\}^2 dt \right] \Bigg|_{\xi_o = 1}.$$

By applying definitions 3 and 4, this may be written

$$\exp \left[2\xi_o a_o + 2 \sum_{\ell > 0} \xi_\ell a_\ell - 2 \sum_{k < 0} \sum_{\ell > 0} \xi_k \xi_\ell b_{|k - \ell|} - 2\xi_o \sum_{k < 0} \xi_k b_k \right. \\ \left. - \xi_o^2 b_o - 2\xi_o \sum_{\ell > 0} \xi_\ell b_\ell - \sum_{j > 0} \sum_{\ell > 0} \xi_j \xi_\ell b_{|j - \ell|} \right] \Bigg|_{\xi_o = 1}.$$

Now under assumption (iv) we may write

$$p(\underline{\xi}^+) = K_2 \exp \left[-\frac{1}{2\sigma^2} \sum_{\ell > 0} \xi_\ell^2 \right] = K_2 \exp \left[-\sum_{j > 0} \sum_{\ell > 0} \xi_j Q_{j\ell} \xi_\ell \right],$$

where we have defined $Q_{j\ell} = \frac{1}{2\sigma^2} \delta_{j\ell}$. Thus, by factoring out of the integral those terms that are independent of $\underline{\xi}^+$, the numerator of Λ becomes proportional to

$$\exp \left[2\xi_o a_o - 2\xi_o \sum_{k < 0} b_k \xi_k - \xi_o^2 b_o \right] \int \exp \left[-\sum_{j > 0} \sum_{\ell > 0} \xi_j (b_{|j - \ell|} + Q_{j\ell}) \xi_\ell \right. \\ \left. + 2 \sum_{\ell > 0} \xi_\ell \left(a_\ell - \sum_{k < 0} b_{|k - \ell|} \xi_k - \xi_o b_\ell \right) \right] d\underline{\xi}^+ \Bigg|_{\xi_o = 1}.$$

By completing the square in the exponent of the integrand, it is straightforward to show that this numerator of Λ is

$$\exp \left[2\xi_0 a_0 - 2\xi_0 \sum_{k<0} b_k \xi_k - \xi_0^2 b_0 + \sum_{j>0} \sum_{\ell>0} \left(a_j - \sum_{k<0} b_{|j-k|} \xi_k - \xi_0 b_j \right) P_{j\ell} \left(a_\ell - \sum_{m<0} b_{|\ell-m|} \xi_m - \xi_0 b_\ell \right) \right]_{\xi_0=1},$$

where the $P_{j\ell}$ are elements of the matrix \underline{P} defined by $\underline{P} = (\underline{B} + \underline{Q})^{-1}$ (for $j, \ell > 0$), and $Q_{j\ell}$ are as defined above. This same expression evaluated at $\xi_0 = -1$ gives the denominator of Λ . It thus follows that the optimum receiver computes

$$\Lambda = \exp \left[4a_0 - 4 \sum_{k<0} b_k \xi_k - 4 \sum_{\ell>0} b_\ell \left\{ \sum_{j>0} \left(a_j - \sum_{k<0} b_{|k-j|} \xi_k \right) P_{j\ell} \right\} \right],$$

and decides H_0 if $\Lambda \geq 1$, and decides H_1 if $\Lambda < 1$. Equivalently, if one defines

$$g_j \triangleq - \sum_{\ell>0} P_{j\ell} b_\ell \tag{2}$$

$$f_k \triangleq b_k + \sum_{j>0} g_j b_{|j-k|}, \tag{3}$$

then the optimum decision rule may be written

$$\boxed{a_0 + \sum_{j>0} g_j a_j - \sum_{k<0} f_k \xi_k \underset{H_1}{\overset{H_0}{>}} 0} \tag{4}$$

The receiver structure may now be found from this decision rule. From definition 3, it is seen that the sufficient statistics a_j may be generated by using a TDL having taps spaced at the baud duration T_b , as shown in Fig. XXVII-1. Moreover, since the weightings on the a_j may be placed before the integration and multiplication operations, and the multipliers and integrators are common to all taps of Fig. XXVII-1, then clearly we can generate

$$\sum_{j \geq 0} g_j a_j$$

as shown in Fig. XXVII-2, in which we define $g_0 = 1$.

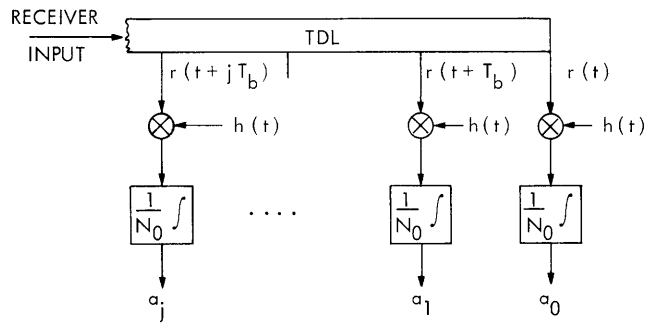


Fig. XXVII-1. Generation of sufficient statistics.

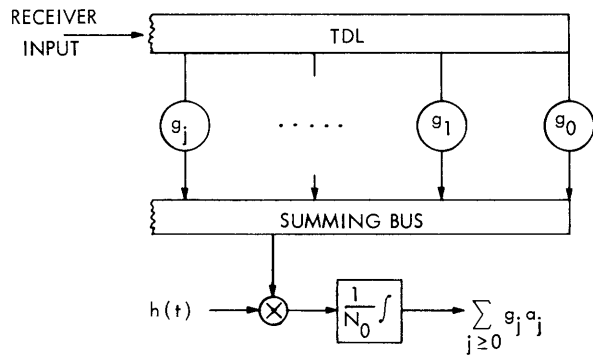


Fig. XXVII-2. Generating the first term of the decision rule.

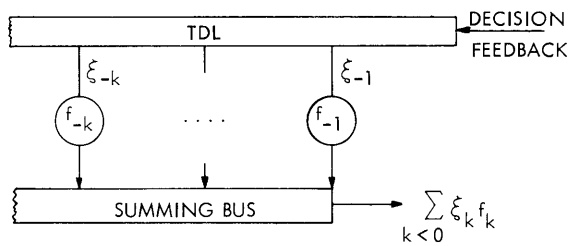


Fig. XXVII-3. Generating the second term of the decision rule.

Similarly, the decision-feedback term may be generated as indicated in Fig. XXVII-3. Noting that the integration need last only over the interval where $h(t)$ is significantly non-zero in Fig. XXVII-2, then we would sample the output at some time, say T , to obtain the desired weighted sum of the a_j . The multiplication-integration procedure is clearly equivalent to a matched filter. Since the MF and TDL are both linear, we may put the MF ahead of the TDL. Also, sampling the TDL output means that one may instead sample the MF output at the baud rate. Thus we have arrived at the final structure of the decision-feedback equalizer shown in Fig. XXVII-4.

We note that this decision-feedback equalizer structure is similar to the conventional MF-TDL equalizer, except that the TDL now only accounts for future bauds, while the feedback TDL accounts for past bauds upon which decisions have been made. Further differences will become apparent in the discussion.

3. Minimum-Output-Sample-Variance Decision-Feedback Equalizer

We now want to adopt the equalizer structure derived above (Fig. XXVII-4) and determine the forward and feedback TDL tap gains that minimize the sum of the signal side-lobes and noise power at the receiver output in the absence of decision errors. We first introduce some additional notation and definitions that will prove useful in the following discussion.

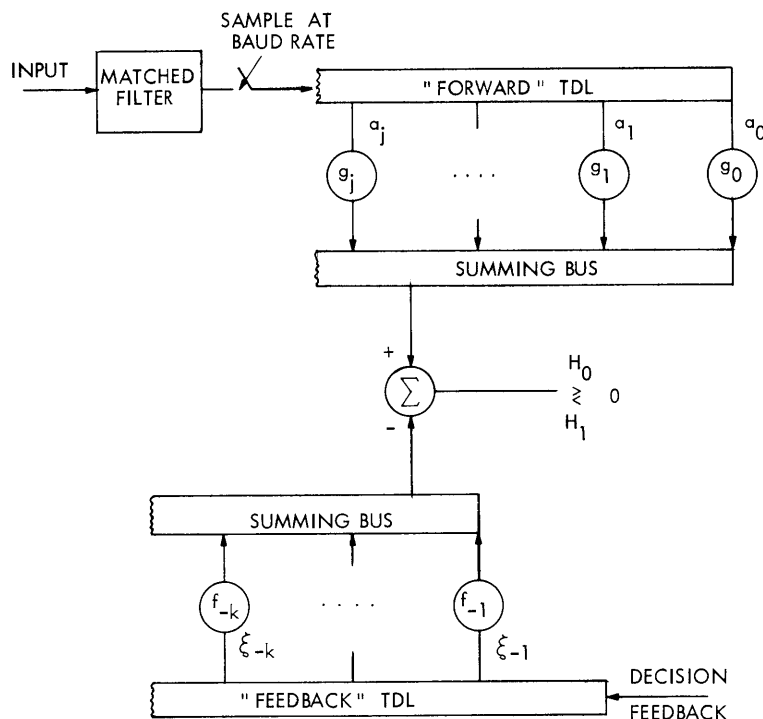


Fig. XXVII-4. Structure of the decision-feedback equalizer.

(XXVII. DETECTION AND ESTIMATION THEORY)

- Definitions:
5. q_ℓ = signal component of the forward-TDL output at the ℓ^{th} sample time, when a single $\xi_0 = +1$ baud is transmitted
 6. $\phi_k = \int h(t) h(t+kT_b) dt$ = sampled channel autocorrelation function at $\tau = kT_b$.
 7. \underline{Y} = matrix with elements $Y_{jk} = \phi_{j-k}$ for $j, k \geq 0$
 8. \underline{X} = matrix with elements $X_{jk} = \sum_{\ell \leq 0} \phi_{j+\ell} \phi_{k+\ell}$ for $j, k \geq 0$
 9. $\underline{\phi}$ = column vector with elements ϕ_i for $i \geq 0$
 10. \underline{g} = column vector with elements g_i for $i \geq 0$
 11. \underline{f} = column vector with elements f_{-i} for $i \geq 1$.

A typical response of the decision-feedback equalizer to a single transmitted baud of $\xi_0 = +1$ is shown in Fig. XXVII-5a. It is always an asymmetrical waveform, having

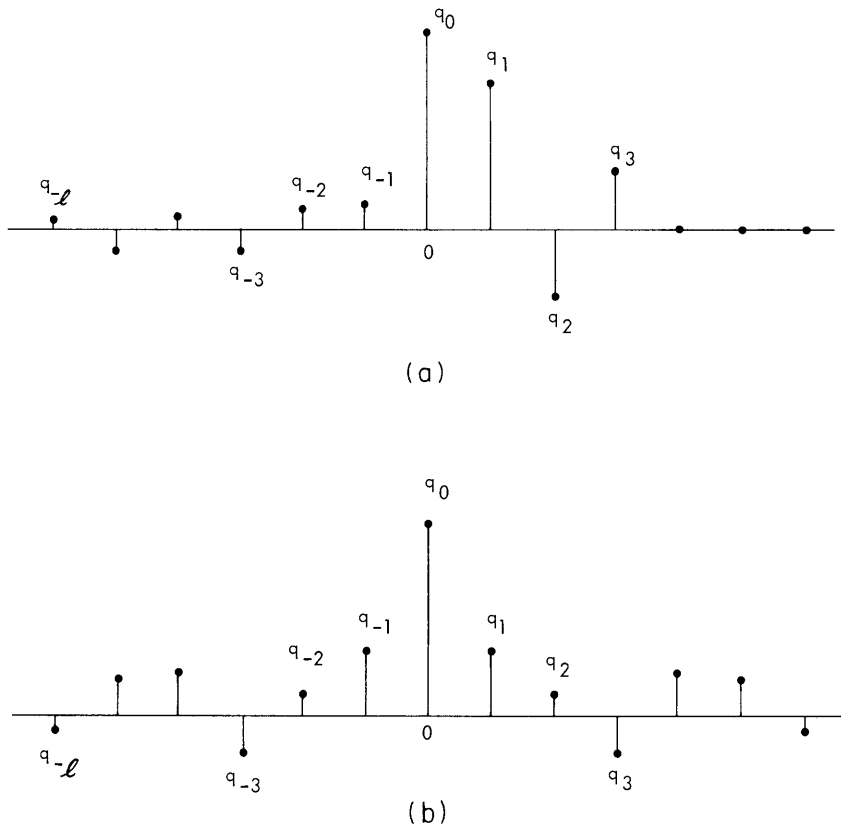


Fig. XXVII-5. Typical responses to a single transmitted $\xi_0 = +1$ baud in the absence of noise: (a) decision-feedback equalizer; and (b) conventional MF-TDL equalizer.

N more samples occurring before the main sample (which is denoted sample number 0) than after it, where the MF output has $2N + 1$ nonzero samples. This is in contrast with the typical output from the conventional MF-TDL equalizer, which is seen in Fig. XXVII-5b to always exhibit symmetry about the main sample.

Before we can proceed to determine the optimum choices of \underline{g} and \underline{f} under our minimum-output-sample-variance criterion, we must first understand the effect of the decision feedback on the output distortion. Consider the signal component out of the forward TDL at the first sample time:

$$\sum_{j \geq 0} g_j [a_{j+1}]_{\text{signal}} = \sum_{j \geq 0} g_j \left[\frac{1}{N_0} \int \left\{ \sum_k \xi_k h(t - kT_b) \right\} \{h(t - (j+1)T_b)\} dt \right].$$

The contribution to this component, which is due to the bauds for which decisions have already been made (that is, on all the ξ_k up to and including ξ_0), is then

$$\sum_{j \geq 0} \sum_{k \leq 0} g_j b_{|j+1-k|} \xi_k.$$

Next, consider the output of the feedback TDL at this same first sample time. With the use of Eq. 3 for the f_k , this becomes

$$\sum_{k < 0} f_k \xi_{k+1} = \sum_{k < 0} \left(b_k + \sum_{j > 0} g_j b_{|j-k|} \right) \xi_{k+1},$$

and if we let $k^* = k + 1$,

$$\begin{aligned} &= \sum_{k^* \leq 0} \left(b_{k^*-1} + \sum_{j > 0} g_j b_{|j+1-k^*|} \right) \xi_{k^*} \\ &= \sum_{j \geq 0} \sum_{k^* \leq 0} g_j b_{|j+1-k^*|} \xi_{k^*}. \end{aligned}$$

Here, we have used $b_{k^*-1} = b_{1-k^*}$ and our earlier definition, $g_0 \triangleq 1$. Thus we see that in the absence of decision errors the feedback-TDL output is exactly the same as the contribution to the forward-TDL output at the first sample time, attributable to past bauds, and hence there is no net contribution to the distortion from those bauds upon which decisions have already been made. We now see the three important advantages that the decision-feedback equalizer enjoys over the conventional equalizer.

(XXVII. DETECTION AND ESTIMATION THEORY)

(i) The conventional equalizer cannot completely eliminate interference due to past bauds, due to noise-enhancement considerations, as well as the practical constraint of a finite TDL length. The decision-feedback equalizer, in the absence of decision errors, completely eliminates the intersymbol interference due to past bauds.

(ii) For the decision-feedback equalizer, the forward-TDL gain-vector \underline{g} may be optimized without consideration of the q_ℓ for $\ell > 0$, since these are eliminated by the decision-feedback, while the conventional equalizer must be designed to simultaneously suppress all of the q_ℓ for $\ell \neq 0$. This additional freedom enables the decision-feedback equalizer to achieve much better suppression of the intersymbol interference due to future bauds.

(iii) Since the intersymbol interference due to past bauds is suppressed through the noiseless feedback-TDL rather than through using additional taps on the forward-TDL as in the conventional equalizer, then clearly the output noise power is significantly reduced.

Each of these advantages contributes to a much better performance for the decision-feedback equalizer compared with that of its conventional counterpart, as we shall show in the example. In view of the conclusions, stated above, it is clear that the output sample variance is given by

$$\sum_{\ell < 0} q_\ell^2 + \text{Output noise variance.}$$

Under the constraint that the main sample be unity, we may include q_0 in the summation to find that

$$\sum_{\ell \leq 0} q_\ell^2 = \sum_{j \geq 0} \sum_{k \geq 0} g_j X_{jk} g_k = \underline{g}^T \underline{X} \underline{g},$$

while the output noise variance can be shown⁷ to be

$$\frac{N_0}{2} \sum_{j \geq 0} \sum_{k \geq 0} g_j Y_{jk} g_k = \frac{N_0}{2} \underline{g}^T \underline{Y} \underline{g}.$$

Thus under the constraint that $q_0 = \underline{g}^T \underline{\phi} = 1$, we want to minimize the quantity

$$J = \underline{g}^T \left[\underline{X} + \frac{N_0}{2} \underline{Y} \right] \underline{g} + \lambda (1 - \underline{g}^T \underline{\phi})$$

over the choice of the forward-TDL gain vector \underline{g} . The unique solution is given by

$$\underline{g} = \frac{\begin{bmatrix} \underline{X} + \frac{N_0}{2} \underline{Y} \end{bmatrix}^{-1} \underline{\phi}}{\underline{\phi}^T \begin{bmatrix} \underline{X} + \frac{N_0}{2} \underline{Y} \end{bmatrix}^{-1} \underline{\phi}} \quad (5)$$

This result appears formally the same as that for the optimum tap-gain vector in the conventional MF-TDL equalizer.⁷ The difference is that for the conventional equalizer the summation of definition 8 is taken over all ℓ , while the vector of definition 9 has elements for all i . Thus \underline{X} matrix for the conventional equalizer becomes symmetrical and Toeplitz, while the $\underline{\phi}$ vector becomes symmetrical, and these properties do not hold for the decision-feedback equalizer.

The proper choice of the \underline{f} vector now follows directly from Eq. 3, except that we have dropped the $1/N_0$ factor common to all terms of Eq. 4 (that is, the a_j contained this factor, while it is not included in the MF output in the present discussion) and thus replace the b_k there by ϕ_k :

$$f_k = \sum_{j \geq 0} g_j \phi_{|j-k|}$$

If we define a matrix $\tilde{\underline{Y}}$ with elements $\tilde{Y}_{jk} = \phi_{j-k}$ for $j \geq 0$ and $k < 0$ (note that it is the range of k which distinguishes this matrix $\tilde{\underline{Y}}$ from the matrix \underline{Y} of definition 7), then the feedback-TDL tap-gain vector may be conveniently written as

$$\underline{f} = \tilde{\underline{Y}} \underline{g} \quad (6)$$

Thus, once the sampled channel autocorrelation function and additive noise level have been specified, one can use Eqs. 5 and 6 to determine the parameters of the minimum-variance decision-feedback equalizer. This is illustrated by the example that follows.

4. Conventional Equalizer versus Decision-Feedback Equalizer: An Example

We shall now work out an example, applying the derived decision-feedback structure to the equalization of the channel whose sampled autocorrelation function is shown in Fig. XXVII-6a. This particular example was chosen, not because it necessarily provides a realistic model of channels of practical importance, but because it illustrates the important advantages that the decision-feedback equalizer has over the conventional MF-TDL equalizer. More complex examples could have been chosen, for they pose no additional difficulties.

We thus want to find the conventional equalizer and decision-feedback equalizer

(XXVII. DETECTION AND ESTIMATION THEORY)

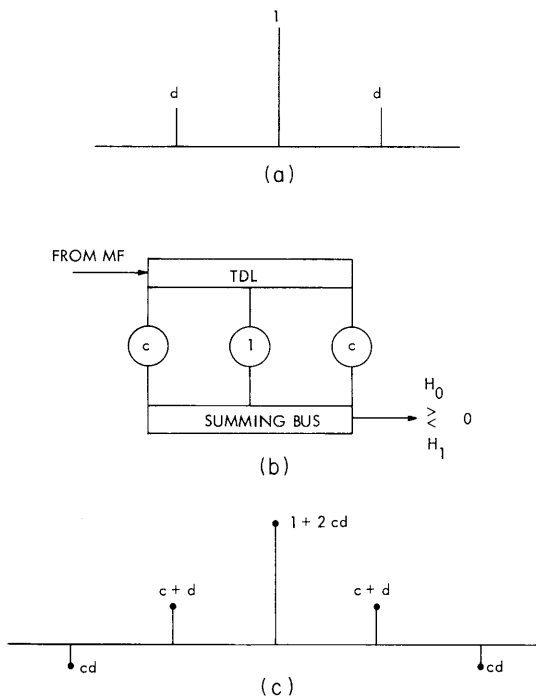


Fig. XXVII-6.

(a) Sampled-channel autocorrelation function. (b) 3-tap conventional equalizer. (c) Output for a single $\xi_0 = +1$ transmitted baud.

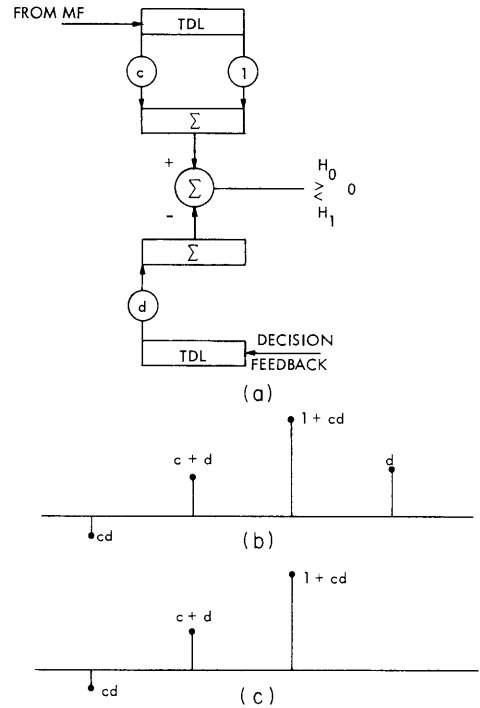


Fig. XXVII-7.

(a) Decision-feedback equalizer. (b) Forward-TDL output. (c) Effective output with decision feedback.

structures as a function of the noise level, $N_0/2$, and the sidelobe level, d , of Fig. XXVII-6a.

Using the 3-tap conventional equalizer shown in Fig. XXVII-6b, one can solve Eq. 5 (with appropriate \underline{X} and $\underline{\phi}$, and normalizing the result so that $g_0 = 1$) to find that

$$c = \frac{2d^3 - d}{1 - d^2 + \frac{N_0}{2}(1-2d^2)} \quad (7)$$

minimizes the output-sample variance, which results in the response to a single transmitted $\xi_0 = +1$ baud shown in Fig. XXVII-6c. The noise present at the output has variance given by

$$\sigma^2 = \frac{N_0}{2} (1+4dc+2c^2). \quad (8)$$

Given σ^2 and an arbitrary set of sidelobes, we can determine the probability of error very efficiently, using an "error tree" algorithm developed by the author.⁷ This

algorithm was applied to the present problem for $d = .4$, $d = .48$, and $d = .50$, with the resulting performance curves for the conventional MF-TDL equalizer shown in Figs. XXVII-8, XXVII-9, and XXVII-10. These curves are discussed further below, when we compare them with the corresponding curves of the decision-feedback equalizer.

For the channel of Fig. XXVII-6a, the decision-feedback equalizer is as shown in Fig. XXVII-7a, where we consider it to be the counterpart of Fig. XXVII-6b, since the delay and tap-gain requirements are the same with both equalizers. We should note at this point that although the forward TDL of the decision-feedback equalizer appears to be half of the conventional-equalizer TDL, this is not the case, in general. In general, if the sampled-channel autocorrelation has M sidelobes, then the feedback TDL requires M taps, with the remaining in the forward TDL. Thus, for example, if one had a conventional equalizer of 55 taps to equalize a channel having 10 autocorrelation sidelobes, then the hardware-equivalent decision-feedback equalizer would have 45 forward-TDL gains and 10 feedback-TDL gains.

Solving Eq. 5 and normalizing so that $g_o = 1$, one finds that

$$c = \frac{d^3 - d}{1 + \frac{N_o}{2}(1-d^2)} \quad (9)$$

minimizes the output-sample variance, and results in the response to a single $\xi_o = +1$ transmitted baud shown in Fig. XXVII-7b. The effective output after decision feedback is shown in Fig. XXVII-7c, where the samples occurring after the main sample have been eliminated. The noise present at the output has variance

$$\sigma^2 = \frac{N_o}{2} (1+2dc+c^2). \quad (10)$$

As discussed further below, the output noise is smaller than that appearing at the output of the conventional equalizer after appropriate normalizations have been made. Note that the parameter c of Eqs. 9 and 10 is numerically different from that of Eqs. 7 and 8.

The performance of the decision-feedback equalizer was determined through digital computer simulations, with the results shown in Figs. XXVII-8, XXVII-9 and XXVII-10. The signal-to-noise ratio in these figures is given by

$$\text{SNR} = 10 \log_{10} \left(\frac{2}{N_o} \right),$$

since we have assumed unit signal energy on each baud. Also, to place the performance curves of the conventional and decision-feedback equalizers in better perspective, Figs. XXVII-8, XXVII-9, and XXVII-10 show the performance curves of an unequalized receiver (matched filter only) and of the ideal receiver (that obtained when transmitting

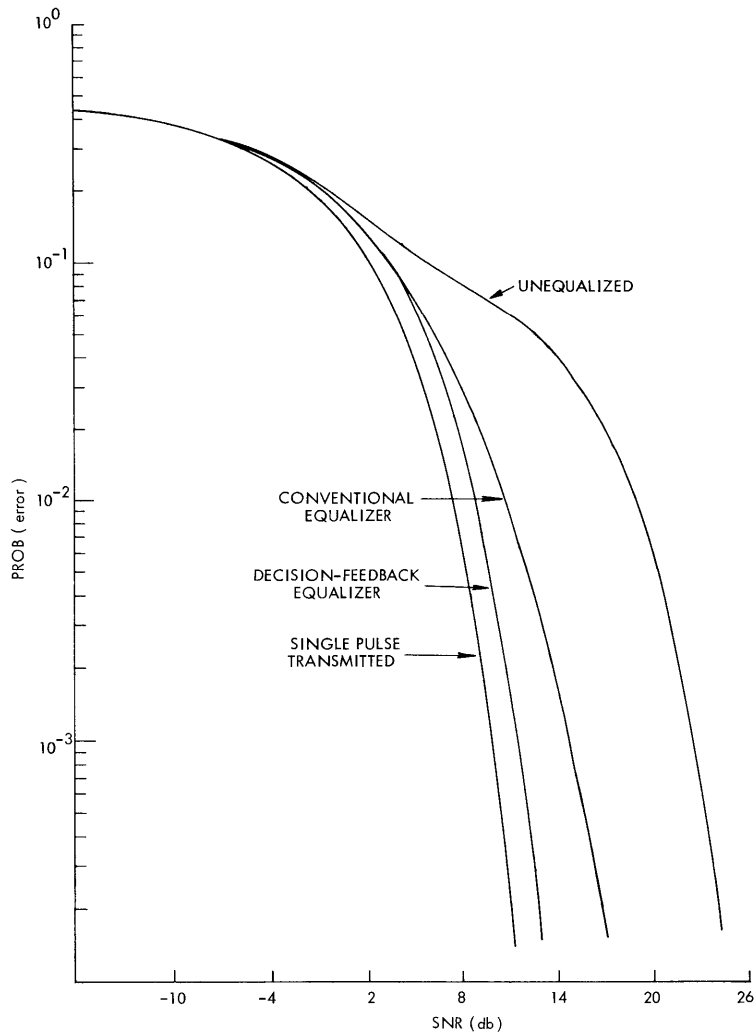


Fig. XXVII-8. Performance for the channel indicated in Fig. XXVII-6a, with $d = .40$.

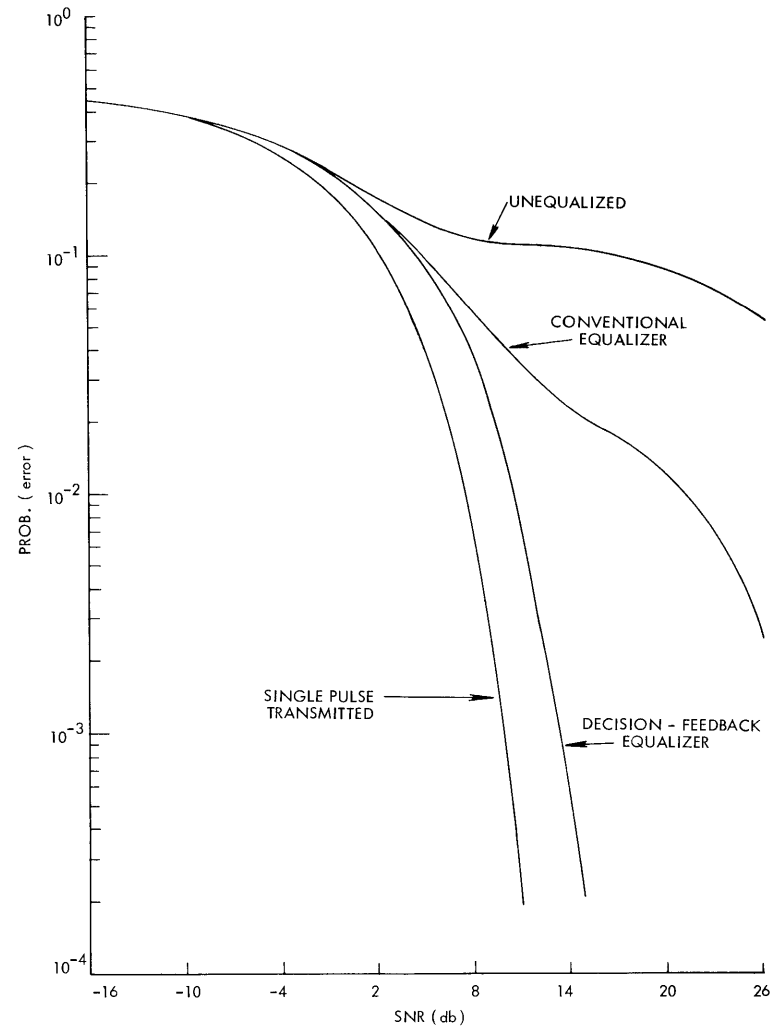


Fig. XXVII-9. Performance for the channel indicated in Fig. XXVII-6a, with $d = .48$.

only a single pulse, where intersymbol interference is no longer a problem).

When the sidelobe energy is approximately one-third that of the main sample (the $d = .40$ case), the decision-feedback equalizer is only approximately 2 db away from the ideal at high SNR, while it is approximately 4 db better than the conventional equalizer.

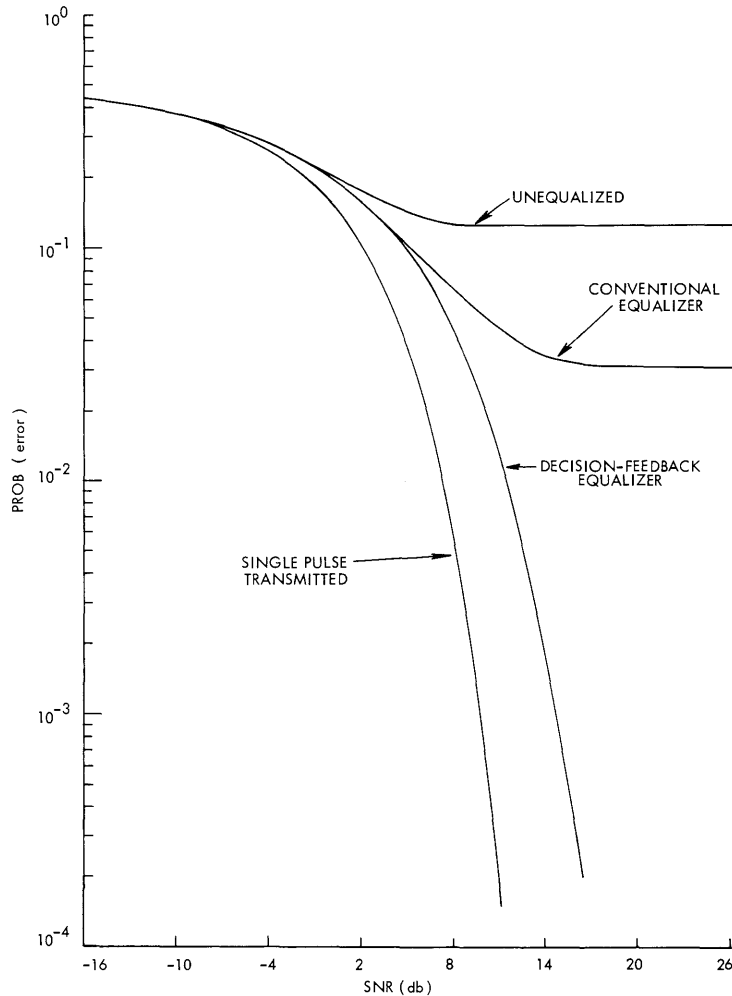


Fig. XXVII-10. Performance for the channel indicated in Fig. XXVII-6a, with $d = .50$.

As the sidelobe energy is increased (Figs. XXVII-9 and XXVII-10), the decision-feedback equalizer becomes 3-5 db off the ideal, and approximately 12 db better than the conventional equalizer at $d = .48$. For $d = .50$, the conventional equalizer is seen in Fig. XXVII-10 to approach a limiting performance with increasing SNR, while the decision-feedback equalizer continues to improve rapidly beyond ~ 5 db. The behavior of the conventional equalizer here is due to the fact that $d = .50$ renders an input distortion

(XXVII. DETECTION AND ESTIMATION THEORY)

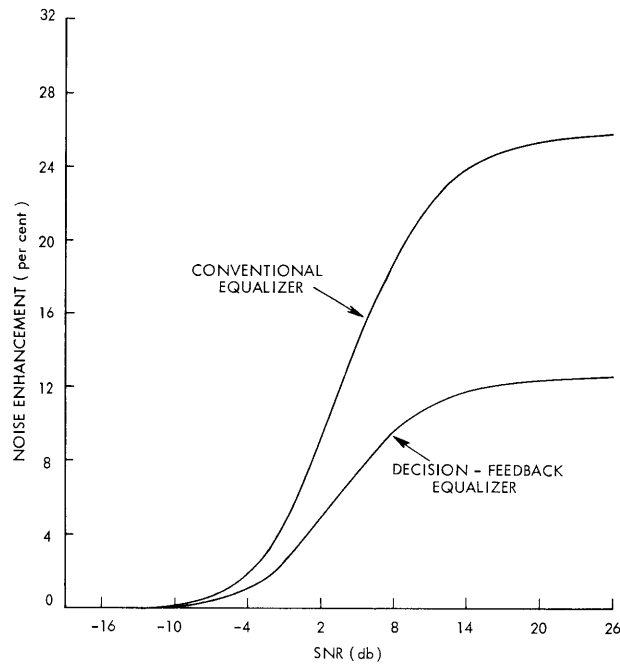


Fig. XXVII-11. Noise enhancement for the channel indicated in Fig. XXVII-6a, with $d = .40$.

of unity (with a sum of the side-lobe magnitudes used as measure; thus the "eye" was closed even in the absence of noise, for those familiar with "eye diagrams"), and a $(2M+1)$ -tap TDL conventional equalizer exhibits a limiting performance of 2^{-2M-3} with increasing SNR.

If one normalizes the main output samples to unity, and correspondingly normalizes the output noise variances, the noise enhancement is considerably more with the conventional equalizer than with the decision-feedback equalizer in this example, as shown in Fig. XXVII-11 for $d = .40$. This illustrates advantage (iii), which we listed for the decision-feedback equalizer.

Note that at low SNR the performance of the two equalizers coincides in Figs. XXVII-8, XXVII-9, and XXVII-10. This is not perhaps what one might have expected from heuristic arguments, which point out that when an error is made by the decision-feedback equalizer the feedback-TDL contribution enhances rather than eliminates the q_1 sample (see Fig. XXVII-7b and 7-c), thereby resulting in this example in an additional equivalent interfering sample whose magnitude exceeds that of the main sample by $|cd|$. This gives a large probability of error on the next decision ($\approx 1/3$ for intermediate SNR, calculated under the assumption of uncorrelated distortion), and thus it appears quite possible that "one bad decision will lead to another," and cause the performance at low SNR to become eventually worse than that of the conventional equalizer.

(XXVII. DETECTION AND ESTIMATION THEORY)

Such behavior was not observed in this example, however, and the decision-feedback equalizer appears at least as good as the conventional equalizer at all SNR.

For more complex channels that we have studied, in which the energy in the side lobes is much greater than in the simple case considered here, the anticipated thresholding effect has been noted,⁷ but only at high conventional equalizer error rates, and the decision-feedback equalizer still exhibits far better performance at all SNR of practical importance.

M. E. Austin

References

1. R. W. Lucky, "Techniques for Adaptive Equalization of Digital Communication Systems," Bell System Tech. J., Vol. XLV, No. 2, pp. 255-286, February 1966.
2. J. M. Aein and J. C. Hancock, "Reducing the Effects of Intersymbol Interference with Correlation Receivers," IEEE Trans., Vol. IT-9, No. 3, pp. 167-175, July 1963.
3. P. Drouilhet and C. W. Niessen, Private discussions, Lincoln Laboratory, M. I. T., February 1966.
4. C. W. Tufts, "Matched Filters and Intersymbol Interference," Technical Report 345, Cruft Laboratory, Harvard University, Cambridge, Massachusetts, July 20, 1961.
5. D. A. George, "Matched Filters for Interfering Signals," IEEE Trans., Vol. IT-11, No. 1, pp. 153-154 (Correspondence), January 1965.
6. M. E. Austin, "Adaptive Signal Processing - Part III," unpublished memoranda, February 1966-October 1966.
7. M. E. Austin, "Adaptive Signal Processing - Part IV," unpublished memoranda, October 1966-January 1967.

B. STATE-VARIABLE ESTIMATION IN THE PRESENCE OF PURE DELAY

1. Introduction

We shall present a state-variable approach for the processing of sonar or seismic array data. In the processing of data from these arrays, one often encounters delayed versions of the same signal, or signals that are correlated by means of a delay, in the same waveform.

Because of the inherent nonrationality of the delay factors, the classical Wiener approach, even for stationary processes, requires finding the impulse response of a filter with an infinite number of poles. When one wishes to use existing state-variable techniques, one requires an infinite dimensional state representation in order to develop the estimation equations. Here, we shall develop a finite set of estimation equations that specify the estimate maximizing the a posteriori probability density of the process.

(XXVII. DETECTION AND ESTIMATION THEORY)

2. Signal Model

Let us now discuss the equations which describe the system of interest (see

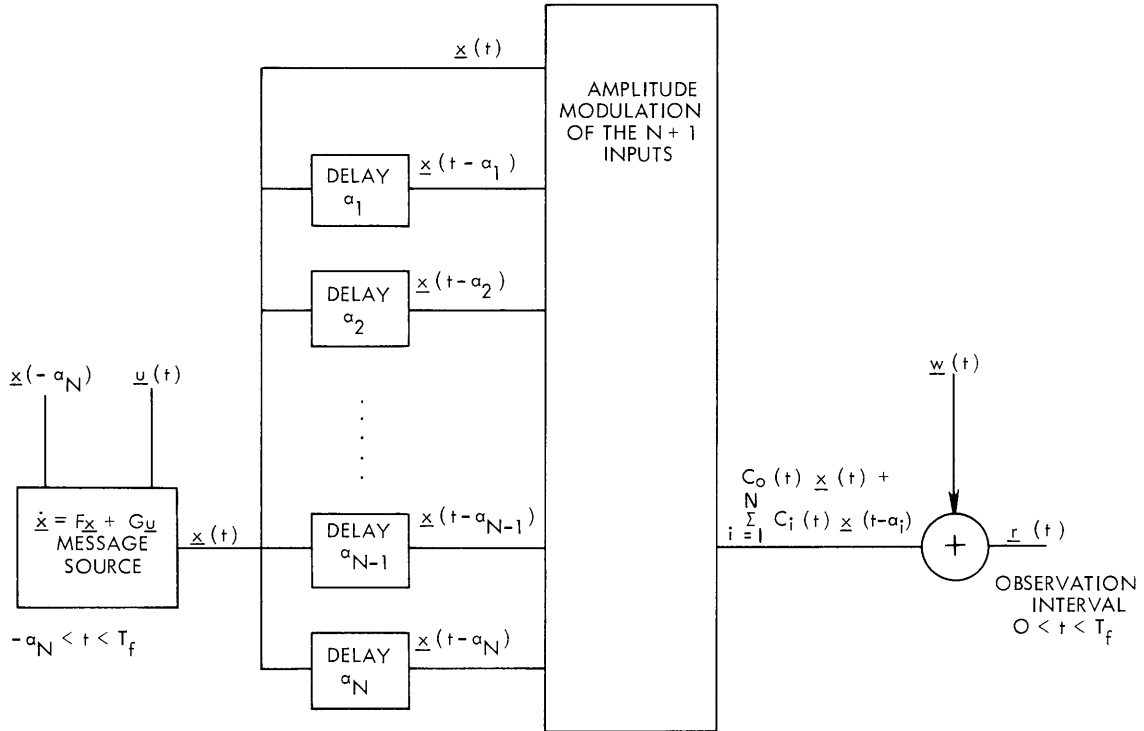


Fig. XXVII-12. System model.

Fig. XXVII-12). We assume that the dynamics of the message process are determined by a linear state equation

$$\frac{d\underline{x}(t)}{dt} = F(t) \underline{x}(t) + G(t) \underline{v}(t) \quad (1)$$

over the time interval $-a_N < t < T_f$, where $F(t)$, $G(t)$ are matrices determining the dynamics of the state equation, and $v(t)$ is a white Gaussian source noise with

$$E[\underline{v}(t)\underline{v}^T(\tau)] = Q\delta(t-\tau). \quad (2)$$

In order to completely specify this random process, we need to make some assumptions about the conditions of the initial state. We assume that the initial state is a Gaussian random vector with

$$E[\underline{x}(-a_N)] = \bar{\underline{x}}(-a_N) \quad (3)$$

$$E\left[(\underline{x}(-a_N) - \bar{\underline{x}}(-a_N))(\underline{x}(-a_N) - \bar{\underline{x}}(-a_N))^T \right] = P(-a_N). \quad (4)$$

We shall now discuss the observation process. We assume that we observe N different delayed versions of an amplitude modulation of the state vector in the presence of a white Gaussian observation noise. Our observation is, therefore,

$$\underline{r}(t) = C_o(t) \underline{x}(t) + \sum_{i=1}^N C_i(t) \underline{x}(t-a_i) + \underline{w}(t), \quad (5)$$

where

$$0 < t < T_f$$

$$0 < a_1 < a_2 < \dots < a_{N-1} < a_N \quad (6)$$

$$E[\underline{w}(t)\underline{w}^T(\tau)] = R\delta(t-\tau). \quad (7)$$

Notice that our observation equation is defined over a different time interval from the state equation. For convenience, let us define

$$\underline{y}_i(t) = \underline{x}(t-a_i). \quad (8)$$

Aside from the delay terms that enter, our assumptions do not differ from the usual ones made in state-variable estimation procedures.

3. Derivation of the MAP Estimation Equations

It can be shown that the problem of maximizing the a posteriori density is equivalent to minimizing the following quadratic functional:

$$\begin{aligned} J(\underline{v}(t), \underline{x}(-a_N)) &= \|\underline{x}(-a_N) - \bar{\underline{x}}(-a_N)\|_{P^{-1}(-a_N)} \\ &+ \int_0^{T_f} \left\| \underline{r}(t) - C_o(t) \underline{x}(t) - \sum_{i=1}^N C_i(t) \underline{y}_i(t) \right\|_{R^{-1}} dt \\ &+ \int_{-a_N}^{T_f} \|\underline{v}(t)\|_{Q^{-1}} dt \end{aligned} \quad (9)$$

Subject to the constraint of the state equation (1) and the delay operations specified by Eq. 6 ($\|\underline{x}\|_A = \underline{x}^T A \underline{x}$).

Let us first consider the state-equation constraint. We can introduce this by using a Lagrangian multiplier. This is done by adding to the quadratic functional the term

$$L_o = \int_{-a_N}^{T_f} \underline{p}^T(t) \left(\frac{d\underline{x}(t)}{dt} - F(t) \underline{x}(t) - G(t) \underline{y}(t) \right) dt. \quad (10)$$

It will be useful to integrate the first term of the integrand by parts. Doing this, we have

$$L_o = \underline{p}^T(T_f) \underline{x}(T_f) - \underline{p}^T(-a_N) \underline{x}(-a_N) \\ - \int_{-a_N}^{T_f} \left\{ \frac{d\underline{p}^T(t)}{dt} \underline{x}(t) + \underline{p}^T(t) F(t) \underline{x}(t) + \underline{p}^T(t) G(t) \underline{y}(t) \right\} dt. \quad (11)$$

When we incorporate the constraints imposed by the delays, we encounter much more difficulty. In order to impose the constraints, we must find a differential equation that the delay operation satisfies. As we would expect from the infinite-state requirement imposed by the delay operation, we cannot find a finite-dimensional ordinary differential equation. It is easy to show, however that the delay operation satisfies the partial differential equation

$$\frac{\partial \phi_i(t, \tau)}{\partial t} + \frac{\partial \phi_i(t, \tau)}{\partial \tau} = 0, \quad (12)$$

where

$$\phi_i(t, 0) = \underline{x}(t).$$

We show this by noting that the general solution to Eq. 12 is

$$\phi_i(t, \tau) = \underline{f}_i(t-\tau). \quad (13)$$

Imposing the boundary condition at $\tau = 0$ yields

$$\phi_i(t, 0) = \underline{f}_i(t) = \underline{x}(t). \quad (14)$$

We see that we now have

$$\phi_i(t, a_i) = \underline{f}_i(t-a_i) = \underline{x}(t-a_i) = \underline{y}_i(t). \quad (15)$$

As a result, we are able to impose the constraint of each delay term by using a Lagrangian multiplier that is a function of two variables, that is, we want to add to the quadratic functional terms of the form

$$L_i = \int_0^{T_f} \int_0^{a_i} \underline{\mu}_i^T(t, \tau) \left(\frac{\partial \phi_i(t, \tau)}{\partial t} + \frac{\partial \phi_i(t, \tau)}{\partial \tau} \right) dt d\tau, \quad (16)$$

in which we have

$$\underline{\phi}_i(t, \tau) = \underline{x}(t-\tau). \quad (17)$$

We again want to perform some integrations by parts. This yields

$$\begin{aligned} L_i = & \int_0^{T_f} \left(\underline{\mu}_i^T(t, a_i) \underline{\phi}_i(t, a_i) - \underline{\mu}_i^T(t, 0) \underline{\phi}_i(t, 0) \right) dt \\ & + \int_0^{a_i} \left(\underline{\mu}_i^T(T_f, \tau) \underline{\phi}_i(T_f, \tau) - \underline{\mu}_i^T(0, \tau) \underline{\phi}_i(0, \tau) \right) d\tau \\ & - \int_0^{T_f} \int_0^{a_i} \left(\frac{\partial \underline{\mu}_i^T(t, \tau)}{\partial t} + \frac{\partial \underline{\mu}_i^T(t, \tau)}{\partial \tau} \right) \underline{\phi}_i(t, \tau) dt d\tau. \end{aligned} \quad (18)$$

When we add the Lagrangian multiplier constraints L_0, L_1, \dots, L_N imposed by the state equation and the delay terms to the quadratic functional described by Eq. 9, we have

$$\begin{aligned} J(\underline{v}(t), \underline{x}(-a_N)) = & \left\| \underline{x}(-a_N) - \bar{\underline{x}}(-a_N) \right\|_{P^{-1}(-a_N)} \\ & + \int_0^{T_f} \left\| \underline{x}(t) - C_0(t) \underline{x}(t) - \sum_{i=1}^N C_i(t) \underline{y}_i(t) \right\|_{R^{-1}} dt \\ & + \int_{-a_N}^{T_f} \left\| \underline{v}(t) \right\|_{Q^{-1}} dt + \underline{p}^T(T_f) \underline{x}(T_f) - \underline{p}^T(T_0) \underline{x}(T_0) \\ & - \int_{-a_N}^{T_f} \left\{ \frac{d\underline{p}^T(t)}{dt} \underline{x}(t) + \underline{p}^T(t) F(t) \underline{x}(t) + \underline{p}^T(t) G(t) \underline{v}(t) \right\} dt \\ & + \sum_{i=1}^N \left[\int_0^{T_f} \left(\underline{\mu}_i^T(t, a_i) \underline{y}_i(t) - \underline{\mu}_i^T(t, 0) \underline{x}(t) \right) dt \right. \\ & + \int_0^{a_i} \left(\underline{\mu}_i^T(T_f, \tau) \underline{x}(T_f - \tau) - \underline{\mu}_i^T(0, \tau) \underline{x}(-\tau) \right) d\tau \\ & \left. - \int_0^{T_f} \int_0^{a_i} \left(\frac{\partial \underline{\mu}_i^T(t, \tau)}{\partial t} + \frac{\partial \underline{\mu}_i^T(t, \tau)}{\partial \tau} \right) \underline{\phi}_i(t, \tau) dt d\tau \right]. \end{aligned} \quad (19)$$

We now are in a position to minimize the quadratic functional by applying variational techniques. We proceed by perturbing $v(t)$ and $x(-a_N)$ from their optimal estimates, that is,

$$\underline{x}(-a_N) = \hat{\underline{x}}(-a_N) + \epsilon \delta \hat{\underline{x}}(-a_N) \quad (20)$$

$$\underline{v}(t) = \hat{\underline{v}}(t) + \epsilon \delta \hat{\underline{v}}(t). \quad (21)$$

The response of the state equation to the perturbed input is

$$\underline{x}(t) = \hat{\underline{x}}(t) + \epsilon \delta \hat{\underline{x}}(t), \quad (22)$$

where $\delta \hat{\underline{x}}(t)$ satisfies the differential equation

$$\frac{d\delta \hat{\underline{x}}(t)}{dt} = F(t) \delta \hat{\underline{x}}(t) + G(t) \delta \underline{v}(t). \quad (23)$$

The variation by the quadratic functional which results from these perturbations is given by

$$\begin{aligned} J(\underline{v}(t), \underline{x}(-a_N)) &= J(\hat{\underline{v}}(t), \hat{\underline{x}}(-a_N)) + \epsilon \left\{ [\underline{x}(-a_N) - \hat{\underline{x}}(-a_N)]^T P^{-1}(-a_N) \delta \underline{x}(-a_N) \right. \\ &\quad + \int_0^{T_f} - \left[\underline{r}(t) - C_o(t) \hat{\underline{x}}(t) - \sum_{i=1}^N C_i(t) \hat{y}_i(t) \right]^T R^{-1} \left[C_o(t) \delta \underline{x}(t) \right. \\ &\quad \left. \left. + \sum_{j=1}^N C_j(t) \delta \hat{y}_j(t) \right] \right\} \\ &\quad + \int_{-a_N}^{T_f} \hat{\underline{v}}(t)^T Q^{-1} \delta \hat{\underline{v}}(t) dt + \underline{p}^T(T_f) \delta \hat{\underline{x}}(T_f) - \underline{p}^T(T_o) \delta \hat{\underline{x}}(-a_N) \\ &\quad - \int_{-a_N}^{T_f} \left\{ \frac{d\underline{p}^T(t)}{dt} \delta \hat{\underline{x}}(t) + \underline{p}^T(t) F(t) \delta \hat{\underline{x}}(t) + \underline{p}^T(t) G(t) \delta \hat{\underline{v}}(t) \right\} dt \\ &\quad + \sum_{i=1}^N \left[\int_0^{T_f} (\underline{\mu}_i^T(t, a_i) \delta \underline{y}_i(t) - \underline{\mu}_i^T(t, 0) \delta \hat{\underline{x}}(t)) dt \right] \\ &\quad + \sum_{i=1}^N \left[\int_0^{a_i} (\underline{\mu}_i^T(T_f, \tau) \delta \underline{x}(T_f - \tau) - \underline{\mu}_i^T(0, \tau) \delta \hat{\underline{x}}(-\tau)) d\tau \right] \\ &\quad \left. + \sum_{i=1}^N \left[\int_0^{T_f} \int_0^{a_i} \frac{\partial \underline{\mu}_i^T(t, \tau)}{\partial t} + \frac{\partial \underline{\mu}_i^T(t, \tau)}{\partial \tau} \delta \hat{\underline{\phi}}_i(t, \tau) dt d\tau \right] \right\} + O(\epsilon^2). \quad (24) \end{aligned}$$

We now want to combine the various variations into common factors. We have

$$\begin{aligned}
J(\underline{y}(t), \underline{x}(-a_N)) &= J(\hat{\underline{y}}(t), \hat{\underline{x}}(-a_N)) + \epsilon \left\{ \left([\hat{\underline{x}}(-a_N) - \bar{\underline{x}}(-a_N)]^T P^{-1}(-a_N) - \underline{p}^T(-a_N) \right) \delta \underline{x}(-a_N) \right. \\
&+ \int_0^{T_f} \left(- \left[\underline{r}(t) - C_o(t) \hat{\underline{x}}(t) - \sum_{j=1}^N C_j(t) \hat{\underline{y}}_j(t) \right]^T R^{-1} C_o(t) \right. \\
&- \left. \left. \left(\frac{d\underline{p}^T(t)}{dt} - \underline{p}^T(t) F(t) - \sum_{i=1}^N \mu_i^T(t, 0) \right) \delta \hat{\underline{x}}(t) \right) dt \right. \\
&+ \int_{-a_N}^{T_f} \left((\hat{\underline{y}}(t)^T Q^{-1} - \underline{p}^T(t) G(t)) \delta \hat{\underline{y}}(t) \right) dt \\
&+ \sum_{i=1}^N \left[\int_0^{T_f} \left(- \left(\underline{r}(t) - C_o(t) \hat{\underline{x}}(t) - \sum_{j=1}^N C_j(t) \hat{\underline{y}}_j(t) \right)^T R^{-1} C_i(t) \right. \right. \\
&+ \left. \left. \mu_i^T(t, a_i) \right) \delta \hat{\underline{y}}_i(t) \right] dt \\
&+ \int_{-a_1}^0 \left(- \frac{d\underline{p}^T(t)}{dt} - \underline{p}^T(t) F(t) - \sum_{j=1}^N \mu_j^T(0, -t) \right) \delta \hat{\underline{x}}(t) dt \\
&+ \sum_{i=1}^{N-1} \left[\int_{-a_{i+1}}^{-a_i} \left(- \frac{d\underline{p}^T(t)}{dt} - \underline{p}^T(t) F(t) - \sum_{j=i+1}^N \mu_j^T(0, -t) \right) \delta \hat{\underline{x}}(t) dt \right] \\
&+ \sum_{j=1}^N \left[\int_0^{a_j} \mu_j^T(T_f, \tau) \delta \hat{\underline{x}}(T_f - \tau) d\tau \right] + \sum_{j=1}^N \left[\int_0^{T_f} \int_0^{a_j} \left(\frac{\partial \mu_j^T(t, \tau)}{\partial t} \right. \right. \\
&\left. \left. + \frac{\partial \mu_j^T(t, \tau)}{\partial \tau} \right) \delta \phi_j(t, \tau) dt d\tau \right]. \quad (25)
\end{aligned}$$

We shall now make a series of arguments to cause the ϵ variation of the functional to vanish. We shall require that the Lagrangian multiplier functions satisfy some equations so that the coefficients of some of the variations vanish.

First we require the delay constraints to have the functional form

$$\mu_i(t, \tau) = \underline{\mu}_{o_i}(t - \tau) \quad (26)$$

so that the last term in the equation vanishes identically ($\underline{\mu}_i$ is the adjoint function for the delay equation). Furthermore, we require that

$$\underline{\mu}_{o_i}(t-\tau) = 0 \quad (27)$$

$$a_i < \tau < T_f \quad (28)$$

and, for $0 < t < T_f$,

$$\underline{\mu}_{o_i}(t-a_i) = C_i^T(t) R^{-1} \left(\underline{r}(t) - C_o(t) \hat{\underline{x}}(t) - \sum_{j=1}^N C_j(t) \underline{y}_j(t) \right). \quad (29)$$

This completes our restrictions of the Lagrangian multipliers for the delay constraints. Notice that the restrictions may be made independently of all the other constraints, including the state-equation constraint.

Now we shall impose some restrictions on the Lagrangian multiplier for the state-equation constraint. First we impose the restriction, $0 < t < T_f$,

$$\frac{d\underline{p}(t)}{dt} = F^T(t) \underline{p}(t) - C_o^T(t) R^{-1} \left(\underline{r}(t) - C_o(t) \hat{\underline{x}}(t) - \sum_{j=1}^N C_j(t) \hat{\underline{y}}_j(t) \right) - \sum_{j=1}^N \underline{\mu}_{o_j}(t). \quad (30)$$

Finally, we impose restrictions for the time before the observation interval. Within this region, for $-a_i < t < 0$, we require

$$\frac{d\underline{p}(t)}{dt} = -F^T(t) \underline{p}(t) - \sum_{j=1}^N \underline{\mu}_{o_j}(t), \quad (31)$$

and, for $-a_{i+1} < t < a_i$, $i = 1, 2, \dots, N-1$,

$$\frac{d\underline{p}(t)}{dt} = -F^T(t) \underline{p}(t) - \sum_{j=i+1}^N \underline{\mu}_{o_j}(t). \quad (32)$$

As a result, we have defined $\underline{p}(t)$ over the entire interval of the process. Therefore we have

$$\begin{aligned} J(\underline{v}(t), \underline{x}(-a_N)) &= J(\hat{\underline{v}}(t), \hat{\underline{x}}(-a_N)) \\ &+ \epsilon \left\{ \left[(\hat{\underline{x}}(-a_N)) - \bar{\underline{x}}(-a_N)^T P^{-1}(-a_N) - \underline{p}^T(-a_N) \right] \delta \hat{\underline{x}}(-a_N) \right. \\ &\left. + \int_{-a_N}^{T_f} (\hat{\underline{v}}(t))^T Q^{-1} - \underline{p}^T(t) G(t) \delta \hat{\underline{v}}(t) + o(\epsilon^2) \right\}. \end{aligned} \quad (33)$$

Because of the optimality of $\hat{\underline{y}}(t)$ and $\hat{\underline{x}}(-a_N)$, we have

$$\underline{p}(-a_N) = P^{-1}(-a_N)((\hat{\underline{x}}(-a_N)) - \bar{\underline{x}}(-a_N)) \quad (34)$$

$$Q^{-1}\hat{\underline{v}}(t) = G^T(t) \underline{p}(t). \quad (35)$$

Equation 34 imposes an initial boundary condition, and Eq. 35 relates the terms $\underline{y}(t)$ and $\underline{p}(t)$. By using Eq. 35, our state equation becomes

$$\frac{d\hat{\underline{x}}(t)}{dt} = F(t) \hat{\underline{x}}(t) + G(t) QG^T(t) \underline{p}(t). \quad (36)$$

We shall now summarize our equations determining the MAP interval estimate of the process $\underline{x}(t)$. For convenience, we shall define

$$a_0 = 0$$

$$\hat{\underline{y}}_0(t) = \hat{\underline{x}}(t).$$

For $-a_N < t < T_f$, we have

$$\frac{d\hat{\underline{x}}(t)}{dt} = F(t) \hat{\underline{x}}(t) + G(t) QG^T(t) \underline{p}(t).$$

By using the definitions stated above, we may write Eqs. 33 and 35 in the same manner. for $-a_{i+1} < t < -a_i$ these equations become

$$\frac{d\underline{p}(t)}{dt} = -F^T(t) \underline{p}(t) - \sum_{j=i+1}^N \mu_{0j}(t),$$

where i may assume the values

$$i = 0, 1, \dots, N-1.$$

For $0 < t < T_f$, we have

$$\frac{d\underline{p}(t)}{dt} = -F^T(t) \underline{p}(t) - C_0(t) R^{-1} \left(\underline{r}(t) - \sum_{j=0}^N C_j(t) \underline{y}_j(t) \right) - \sum_{j=1}^N \mu_{0j}(t).$$

The functions μ_i are defined to be, for $0 < \tau < a_i$,

$$\mu_{0i}(T_f - \tau) = 0,$$

and for $0 < t < T_f$,

$$\mu_{0i}(t - a_i) = C_i^T(t) R^{-1} \left(\underline{r}(t) - \sum_{j=0}^N C_j(t) \hat{\underline{y}}(t) \right).$$

(XXVII. DETECTION AND ESTIMATION THEORY)

The result of our derivation is a set of differential-difference equations with a two-point boundary-value condition. In the absence of any delay terms, the equations are identical to those for the interval estimator with no-memory modulation, that is, amplitude modulation. At the present time, the equation may be solved only for the case of an equally spaced delay, $a_1 = i\Delta a$. We are now working on more general methods of solution, the performance of the estimator in a delay environment, and realizable filters for such an environment.

A. B. Baggeroer

References

1. A. C. Bryson and M. Frazier, "Smoothing for Linear and Nonlinear Dynamic Systems," Proc. Optimum Systems Conference, Wright-Patterson Air Force Base, Ohio, September, 1962.
2. R. E. Kalman and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," ASME Paper 60-JAC, May 31, 1960.
3. M. Athans and P. Falb, Optimal Control (McGraw-Hill Book Company, Inc., New York, 1966).
4. A. Baggeroer, "Maximum A Posteriori Interval Estimation," WESCON/66 Technical Papers, Session 7, Paper 7/3.