PURDUE UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING
ELECTRONIC SYSTEMS RESEARCH LABORATORY

SEMI-ANNUAL REPORT OF RESEARCH

PERFORMED UNDER GRANT NsG—553

JULY 1, 1965 through DECEMBER 31, 1965

LAFAYETTE, INDIANA

FOREWORD

This report summarizes work carried out at the Electronic Systems Research Laboratory of Purdue University in the Communication Sciences area during the period July 1, 1965 through December 31, 1965. The research reported herein was supported in full or in part under NASA Grant NsG-553.

C D McGillem

C. D. McGillem, Director
Electronic Systems
Research Laboratory

TABLE OF CONTENTS

	Page
FOREWORD	i
PROJECT PERSONNEL	iii
SECTION I - LEARNING SYSTEMS	1
A. Nonsupervision and Parameter Conditional Mixture	1
B. Cognitive Signal Processing	7
C. Adaptive Detection without Syncronization	13
SECTION II - ADAPTIVE SYSTEMS	19
A. Probability of Error for Correlators with Noisy Reference Signals	19
B. Purdue Channel Simulator	24
SECTION III - SIGNAL DESIGN	26
A. Experimental Research on Communication Systems Subject to Intersymbol Interference	26
B. Optimum Waveforms and Receivers for Channels with Memory	29
C. Signal Selection for Telemetry Channels	37
SECTION IV - PHASE LOCK LOOP STUDIES	42

PROJECT PERSONNEL

Professorial Staff

D. R. Anderson

J. C. Hancock

J. C. Lindenlaub

E. A. Patrick

R. E. Totty

P. A. Wintz

Graduate Instructors and Assistants

C. C. Bailey

W. D. Gregg

E. J. Luecke

R. A. Markley

S. E. Nykanen

E. A. Quincy

G. R. Redinbo

T. L. Stewart

J. J. Uhran

I. LEARNING SYSTEMS

A. NONSUPERVISION AND PARAMETER-CONDITIONAL MIXTURES

J. C. Hancock

E. A. Patrick

1. Mixtures and Parameter Conditional Mixture

In the last Semi-Annual report we defined a parameter-conditional mixture. The type of the mixture depends upon the a priori knowledge used in its construction. By approaching nonsupervisory problems through first defining mixtures, we are able to define precisely different nonsupervisory problems and the a priori knowledge they utilize. For example, the mixture defined in Pt. 4 p. 3 can be used to apply histogram concepts to the nonsupervisory problem. The mixture defined in Pt. 4 p. 2 corresponds to the "classical" nonsupervisory problem. The mixture defined in Pt. 5 p. 4 arises when more than one class can be active on the same sample.

A mixture results when a vector X can be partitioned W ways, $\pi_1, \pi_2, \ldots, \pi_W$. If, for example, $X = \{X_s\}_{n-v+1}^{n}$ with a single pattern class active causing each X_s, there are $W = M^v$ ways the pattern classes could be active to cause X. If, as another example, $X = X_s$ with a single pattern class active causing X_s, there are W = M ways the pattern classes could be active to cause X. Since the partitions are mutually exclusive and exhaustive,

$$F(X) = \sum_{r=1}^{W} F(X \mid \pi_r) \, P(\pi_r) \tag{1}$$

where F(X) is called the mixture c.d.f., $F(X \mid \pi_r)$ the rth partition-conditional c.d.f., and $P(\pi_r)$ the rth mixing parameter.

When we speak of a family of gaussian c.d.f.'s or a family of multinomial c.d.f.'s, we have in mind the nature of the parameters which characterize the family. It is, therefore, appropriate to define a parameter-conditional mixture c.d.f. $F(X \mid B)$ constructed using the family $\{F(X \mid \pi_r, B_r)\}$ of rth partition, parameter-conditional c.d.f.'s. To do this, define

$$B = B_1 U B_2 U \ldots, U B_M U B_{M+1} \tag{2}$$

$$B_{M+1} = \{P(\pi_r)\}_1^W \tag{3}$$

Since (X, π_1), $(X,\pi_2),\ldots,(X,\pi_W)$ are mutually exclusive and exhaustive events,

$$F(X|B) = \sum_{r=1}^{W} F(X,\pi_r|B)$$

$$= \sum_{r=1}^{W} F(X|\pi_r,B)\, P(\pi_r|B) \tag{4}$$

Now, the rth partition-conditional c.d.f. is characterized by B_r,

$$F(X|\pi_r,B) = F(X|\pi_r,B_r) \tag{5}$$

and since B contains $P(\pi_r)$,

$$P(\pi_r|B) = P(\pi_r) \tag{6}$$

Thus, (4) becomes

$$F(X|B) = \sum_{r=1}^{W} F(X|\pi_r,B_r)\, P(\pi_r) \tag{7}$$

If we are given $F(X)$, W, and the family $\{F(X|\pi_r, B_r\}$, then when can B be uniquely found? Or, put another way, given $F(X)$, when does $F(X) = F(X|B)$ have a unique solution for B? The answer is that B can be uniquely found when the class of parameter-conditional mixtures is identifiable, sufficient conditions for which are given in Ref. 1.

In the following sections we proceed to relate Eq. 7 to nonsupervisory problems arising in practice.

2. $X = X_s$ with Single Class Active

Let $X = X_s$ with one of M pattern classes possibly active. Then $W = M$ and Eq. (7) becomes

$$F(X_s|B) = \sum_{r=1}^{M} F(X_s|\omega_i,\, B_i)\, P_i \tag{8}$$

This parameter-conditional mixture, Eq. (8), arises when samples X_1, $X_2,\ldots,$ X_n are parameter-conditionally independent.

3. v Samples Parameter-Conditionally Dependent

Let $X = \{X_s\}_{n-v+1}^{n}$ with a single pattern class active causing each sample X_s. Then $W = M^v$. Equation (7) becomes

$$F(\{X_s\}_{n-v+1}^{n}|B) = \sum_{r=1}^{M^v} F(\{X_s\}_{n-v+1}^{n}|\pi_r, B_r)\ P(\pi_r) \tag{9}$$

A mixture of this form arises when making a decision on sample X_n if X_n, given π_r and B_r, is statistically dependent on the previous (v-1) samples. The distribution function of X_n, conditioned on $\{X_s\}_{n-v+1}^{n-1}$ and B, can be expressed as

$$F(X_n|B,\{X_s\}_{n-v+1}^{n-1}) = \frac{\sum_{r=1}^{M^v} F(\{X_s\}_{n-v+1}^{n}|B_r^v, \pi_r^v)\ P(\pi_r^v)}{\sum_{r=1}^{M^v} F(\{X_s\}_{n-v+1}^{n-1}|B_r^{v-1}, \pi_r^{v-1}) P(\pi_r^{v-1})} \tag{10}$$

where π_r^v denotes the rth partition for samples $X_{n-v+1},\ldots,X_n$, and π_r^{v-1} denotes the rth partition for samples $X_{n-v+1},\ldots,X_{n-1}$.

Thus, when the v samples are statistically dependent, a priori knowledge must include the family $\{F(\{X_s\}_{n-v+1}^{n}|B_r,\pi_r)\}$ of multidimensional rth partition, parameter-conditional c.d.f.'s, the dimension of each member increasing as v increases. Furthermore, the number of terms in this mixture grows as v increases.

4. $X = X_s = \{X_{s_k}\}_1^v$ with Single Pattern Class Active

Let $X = X_s = X_{s_1}, X_{s_2}, \ldots, X_{s_v}$ with class ω_i active for all v samples. The parameter-conditional mixture c.d.f. $F(X_s|B)$ is

$$F(X_s|B) = \sum_{i=1}^{M} F(\{X_{s_k}\}_1^v|\omega_i, B_i)\ P_i \tag{11}$$

This mixture does not grow with increasing v as did the previous mixture because the statistically dependent samples are supervised. The a priori knowledge used to construct this mixture is knowledge of M, the family, and the fact that $X_s = \{X_{s_k}\}_1^v$ with one pattern class active for all samples.

We find in Ref. 1 that this type of mixture arises when applying the histogram concept to nonsupervisory problems. By taking v samples at the sth observation with pattern class ω_i active, the class of mixtures may be identifiable whereas it would not be with only one sample taken.

5. $X = X_s$ with Interclass Interference

Let $X = X_s$ with any number of M classes possibly active causing X_s, a situation we will call interclass interference. The a priori knowledge also includes knowledge of M, the family, and that a class ω_i is active on the sth sample with probability P_i. Since a class ω_i is either active or not for each sample X_s, there are 2^M mutually exclusive and exhaustive ways that the sth sample can occur. Thus the parameter-conditional mixture c.d.f. $F(X_s|B)$ is

$$F(X_s|B) = \sum_{r=1}^{2^M} F(X_s|\pi_r, B_r)\, P(\pi_r) \tag{12}$$

6. Two Possible Sets of Mixing Parameters

Let $X = X_s$ and a single class ω_i active for X_s. The a priori knowledge includes knowledge that M = 2, the family is known, and that there are two possible sets of mixing parameters defined as follows:

It is known that either P_1 or $(1-P_1)$ is equal to P; $P_1 = P$ with probability Q; and $(1-P_1) = P$ with probability (1-Q). Since the events $P_1 = P$ and $(1-P_1) = P$ are mutually exclusive (assume $P \neq \frac{1}{2}$), the parameter conditional mixture c.d.f. is

$$F(X_s|B) = Q[PF(X_s|\omega_1,B_1) + (1-P_1)F(X_s|\omega_2,B_2)] +$$
$$(1-Q)[(1-P)F(X_s|\omega_1,B_1) + PF(X_s|\omega_2,B_2)] \quad (13)$$

where

$$B = (Q, P, B_1, B_2) \quad (14)$$

Define

$$F_1(X_s|B) = PF(X_s|\omega_1,B_1) + (1-P)F(X_s|\omega_2,B_2)$$
$$F_2(X_s|B) = (1-P)F(X_s|\omega_1,B_1) + PF(X_s|\omega_2,B_2)$$

Equation (14) then simplifies to

$$F(X_s|B) = QF_1(X_s|B) + (1-Q)F_2(X_s|B) \quad (15)$$

As the problem is formulated, Q is either 1 or 0 since only one of the two sets of mixing parameters is active at a given time. Thus, (15) is a parameter-conditional mixture with one zero mixing parameter. The sufficient conditions given in Ref. 1 require all mixing parameters to be greater than zero but less than one. We, therefore, cannot conclude sufficient conditions for identifiability in this present problem. On the other hand, the fact that one of the mixing parameters has value P is a priori knowledge, and thus should not impose greater constraints on the class of resulting parameter-conditional mixtures for identifiability. This shows the need for a study of identifiability when a mixture has one or more mixing parameters of value zero, and corresponds to the nonsupervisory problem with an unknown number of pattern classes M.

7. Given a Set of Families

Consider now a situation where there are R possible families, $\mathcal{F}_j = \{F^j(X_s|\omega_i,B_i)\}$, $j = 1,2,\ldots,R$. This might correspond to a problem where the class-conditional c.d.f. depends upon some parameter, for example phase, which changes from sample to sample, and takes on R possible values. Or, it might correspond to a problem where the noise statistics change from sample to sample, being represented by one of R possible c.d.f.'s. We will now assume that the samples are classified but that the families are not. That is, let $X = X_s$ with ω_i

known active causing X_s, and the jth family active with probability Q_j. Then

$$F(X_s) = \sum_{j=1}^{R} Q_j F^j(X_s|\omega_i), \ \omega_i \text{ known} \tag{16}$$

Thus the probability distribution of X_s is given by a mixture c.d.f. even though the samples are classified. In this case, the families active in causing the samples are unclassified.

REFERENCES

1. Patrick, E. A., "Learning Probability Spaces for Classification and Recognition of Patterns with or without Supervision," Ph.D. Thesis, Purdue University, November, 1965.

B. COGNITIVE SIGNAL PROCESSING

J. C. Hancock

W. D. Gregg

N66 20099

1. Summary

This report is a summary of the research on unsupervised signal pattern recognition under the above title performed since the second semi-annual report[1] with the following progress to date:

(1) A time series analog of the multivariate "distance statistic" discussed previously[1] has been developed for successively testing the (i+1)st vector observation against the ith vector observation for statistical similarity of class parameters. A computer simulation of a noisy signal pattern format subjected to the "separation test" demonstrates quite reliable results for signal to noise ratios (defined below) down to -4.5db.

(2) An asymptotic form of the Bayes decision boundary as constructed from unsupervised or unclassified vector samples has been developed. The effect of no preclassification or supervision appears as a bias in the asymptotic error probability; and for signal-to-noise ratios less than 0 db, the asymptotic error probability for the model considered is on the order of two to four times greater than that of the known parameter Bayes matched filter.

2. A Priori Mode

The a priori mode consists of the separation of the observed vectors into two classes on the basis of statistical similarity and combination to provide a priori estimates, $\underline{\theta}_{1o}$, $\underline{\theta}_{2o}$, Θ_o of the class parameters. A time series analog (of Eq. (1)[1]) calls for a linear combination of the elements of $\underline{z}_i$, $\underline{z}_{i+1}$, to be tested against a threshold for similarity or dis-similarity about the mean. Maximization of the distance function with respect to $\underline{h}$ by the Schwartz inequality for probability has

$$\underset{\underline{h}}{\text{Max}}\ d\left[\left(\underline{h}_i'\ \underline{z}_i\right)\left(\underline{h}_{i+1}'\underline{z}_{i+1}\right)\right] = \underset{\underline{h}}{\text{Max}} \left\{\frac{\overline{\lambda_A(\underline{d})}}{\overline{\lambda_A^2(d)}}\right\}_{H_o} - \left\{\frac{\overline{\lambda_A(\underline{d})}}{\overline{\lambda_A^2(\underline{d})}}\right\}_{H_1} \tag{1}$$

and for the anti-polar case, $\underline{\theta}_1 = -\underline{\theta}_2$,

$$\underline{h}_i = \underline{z}_{i+1},\ \underline{h}_{i+1} = \underline{z}_i \tag{2}$$

yielding a minimax decision rule

$$\lambda_{A_i} = (\underline{z}_i,\ \underline{z}_{i+1})^2 \begin{cases} \geq 0 & H_o \\ < 0 & H_1 \end{cases} \tag{3}$$

The hypotheses are: (a) H_o, that the ith and (i+1)th vector observations are both from ω_1 or ω_2, and (b) H_1, that one vector observation is from ω_1 while the other one is from ω_2. For the anti-polar case, (see Fig. 1b, Ref. 1) the corresponding computer-simulated sampled noisy signal pattern for a SNR of 3db is illustrated in Fig. 1e.* SNR for white noise is defined as $||\underline{\theta}||^2/\sigma^2$. For equal class covariances, the time slot (vector) samples can differ only in the pulse shapes (mean vectors $\underline{\theta}_1$, $\underline{\theta}_2$). The a priori estimates of the elements of the vectors, $\underline{\theta}_1$, $\underline{\theta}_2$, as a function of the number of time slots observed are illustrated in Figs. 1a, b, c, and d* for one computer simulation experiment. The parameter values are extracted from the noisy signal pattern (mixture) without any pre-classification or supervision and without any a priori knowledge of the signal parameters. Various computer simulation experiments are in progress in order to establish and confirm bounds on dynamic performance.

3. Classification Mode

The development of an asymptotic form of the unsupervised Bayes decision boundary proceeds as follows. The parametric Bayes formulation of the classification mode against an a priori distribution $P(\{\theta\})$, with parameters $\underline{\theta}_{1o}$, $\underline{\theta}_{2o}$, Θ_o provided by an

*Refer to page 12 of this report for part a, b, c, d and e of Fig. 1.

a priori mode, yields the decision rule

$$(\lambda(\underline{z}_{N+1}|\underline{Z}) - \lambda_o) \underset{<\,0}{\geq 0} \tag{4}$$

where for equal risk

$$\lambda(\underline{z}_{N+1}|\underline{Z}) = \frac{\int_{R\{\theta\}} P(\underline{z}_{N+1}|\omega_1,\{\theta\})\; P(\underline{Z}|\{\theta\})\; P(\{\theta\})\; d\{\theta\}}{\int_{R\{\theta\}} P(\underline{z}_{N+1}|\omega_2,\{\theta\})\; P(\underline{Z}|\{\theta\})\; P(\{\theta\})\; d\{\theta\}} \tag{5}$$

and

$$\lambda_o = {}^{p_2}/p_1;\ \{\theta\} = (\underline{\theta}_1,\underline{\theta}_2,\Theta);\ \Theta \overset{\Delta}{=} \text{Noise Covariance Matrix} \tag{6}$$

For the multivariate gaussian case, assuming no pre-classification of $\underline{Z}$, the formal expression[2] for $P(\underline{Z}|\{\theta\})$ can be factored into

$$P(\underline{Z}|\{\theta\}) = p_2^N\, P(\underline{Z}|\{\theta\}^{(2)}) \prod_{i=1}^{N} \left[1 + \lambda_o^{-1}\, P(\underline{z}_i|\{\theta\}^{(1)})/P(\underline{z}_i|\{\theta\}^{(2)})\right] \tag{7}$$

The bi-polar and off-on signal pattern models can be treated simultaneously by introducing the a priori information about the populations as

$$\underline{\theta}_2 = k\,\underline{\theta}_1;\ P(\{\theta\}) = P(\Theta|\underline{\theta}_1,\underline{\theta}_2)\; P(\underline{\theta}_1|\underline{\theta}_2)\; \delta(\underline{\theta}_2 - k\,\underline{\theta}_1) \tag{8}$$

The substitution of (7) and (8) into the numerator and denominator of (6), followed by a factoring and completion of the square with respect to the $\underline{\theta}_i$ of the exponents in the integrand and integration, yields for (5) a ratio of a series of Pearson Type VII terms, e.g.

$$P(\underline{z}_{N+1}|\omega_i,\underline{Z}) \sim \left(1 + \frac{{}^{(o)}Q_{iN}}{N}\right)^{-\frac{N}{2}} + \lambda_o^{-1}\left(1 + \frac{{}^{(1)}Q_{iN}}{N}\right)^{-\frac{N}{2}} + \ldots + \lambda_o^{-N}\left(1 + \frac{{}^{(1+\ldots+N)}Q_{iN}}{N}\right)^{-\frac{N}{2}}, \quad i = 1,2 \tag{9}$$

The elements ${}^{(\xi)}Q_{iN}$ are quadratic forms composed of combinations of the observed vector samples. The decision boundaries for the bi-polar and off-on cases are obtained by taking $\lim_k \lambda(\underline{z}_{N+1}|\underline{Z})$ as $k \to -1$ and 0 respectively. A subsequent application of the law of large numbers to the sequences of observations in each term of (9) has

$$\lim_{N \to \infty} \left(1 + \frac{{}^{(\xi)}Q_{iN}}{N}\right)^{-\frac{N}{2}} \longrightarrow e^{-\frac{{}^{(\xi)}Q_{iL}}{2}} \qquad (10)$$

where the ${}^{(\xi)}Q_{iL}$ are quadratic forms with parameters comprising the asymptotic values of the sequences of observations in the ${}^{(\xi)}Q_{iN}$ terms. Thus in

$$P \left\{ \lim_{N \to \infty} \lambda(\underline{z}_{N+1} | \underline{Z}) = \lambda_{\infty} \right\} = 1 \qquad (11)$$

for the bi-polar case, k = -1, letting $\lambda_L = \ln \lambda_{\infty}$, the asymptotic expression for the "unsupervised decision boundary" is

$$\hat{\beta}_L(\underline{z}) = \left[2 \; (p_1 - p_2) \; \sum\nolimits_M^{-1} \underline{\theta}_1 \right]' \left[\underline{z} - \underline{o} \right] + \lambda_o \qquad (12)$$

which has the same form as the known parameter decision boundary

$$\beta(\underline{z}) = \left[2 \sum\nolimits^{-1} \underline{\theta}_1 \right]' \left[\underline{z} - \underline{o} \right] + \lambda_o \qquad (13)$$

Since $\sum$ is the true noise covariance matrix and $\sum_M$ is the asymptotic form of the matrix in ${}^{(\xi)}Q_{iN}$, the discriminant function coefficients differ in (12) and (13). For n = 2 (two samples per time slot),

$$\sum\nolimits_M = \begin{bmatrix} \{\sigma^2 + 6p_1p_2\theta_{21}^2\} & \{\rho \; \sigma^2 + 4p_1p_2\theta_{21}\theta_{22}\} \\ \{\rho \; \sigma^2 + 4p_1p_2\theta_{21}\theta_{22}\} & \{\sigma^2 + 6p_1p_2\theta_{22}^2\} \end{bmatrix} \qquad (14)$$

The presence of (14) in (12) tends to "color" white noise and not exactly whiten "colored" noise with the effect of biasing the probability of classification error away from that of (13). This difference in probability of classification error can be expressed as

$$\Delta \, P_e = p_1 \int_{\hat{h}_1}^{h_1} f(x)dx + p_2 \int_{\hat{h}_2}^{h_2} f(x)dx \qquad (15)$$

where f(x) is the gaussian P.D.F. For square bi-polar pulses of amplitude c in white noise,

$$h = \frac{1}{2\sqrt{2}} \left\{ \frac{\text{Ln}(p_2/p_1)}{(c/\sigma)} \pm 4 \right\} (c/\sigma) \tag{16}$$

$$\hat{h} = \frac{(p_1-p_2)^{-1}}{2\sqrt{2}} \left\{ \frac{\text{Ln}(p_2/p_1)\ [(\sigma/c)^4 + 12p_1p_2\ (\sigma/c)^2 + 20(p_1p_2)^2]}{[(\sigma/c)^2 + 2p_1p_2]} \pm \sqrt{2} \right\} (c/\sigma)$$

with the subscripts 1 and 2 in (15) associated with the + and - signs in (16) respectively. An illustration of (15) is shown in Fig. 2 as the class probability weighted sum of the shaded areas. Both $|h|$ and $|\hat{h}|$ increase with SNR and thus

$$\lim_{SNR \to \infty} |h_i| - |\hat{h}_i| \to 0; \ SNR \sim (c/\sigma)^2; \ i = 1,2 \tag{17}$$

with the effect that the probability of error for the asymptotic unsupervised decision boundary approaches that for the known parameter decision boundary as SNR becomes large. A comparison for a particular case appears in Table I.

TABLE I

($p_1 = 7/8$, $p_2 = 1/8$, Square Anti-polar pulses)

SNR(db)	P_e(Known Parameter) Case	P_e(Asymptotic Unsupervised) Case
-4.5	.36	.48
-3.0	.30	.45
-1.5	.23	.41
0.0	.15	.36
1.5	.09	.33
3.0	.05	.29

REFERENCES

1. Hancock, J.C. and Gregg, W.D., "Cognitive Signal Processing", Purdue University School of Electrical Engineering, Second Semi-Annual Research Summary, Jan.-July 1965, pp. 15-22, Section 5.1.

2. Rao, C.R. Advanced Statistical Methods in Biometric Research, J. Wiley and Sons, New York, N.Y., 1952.

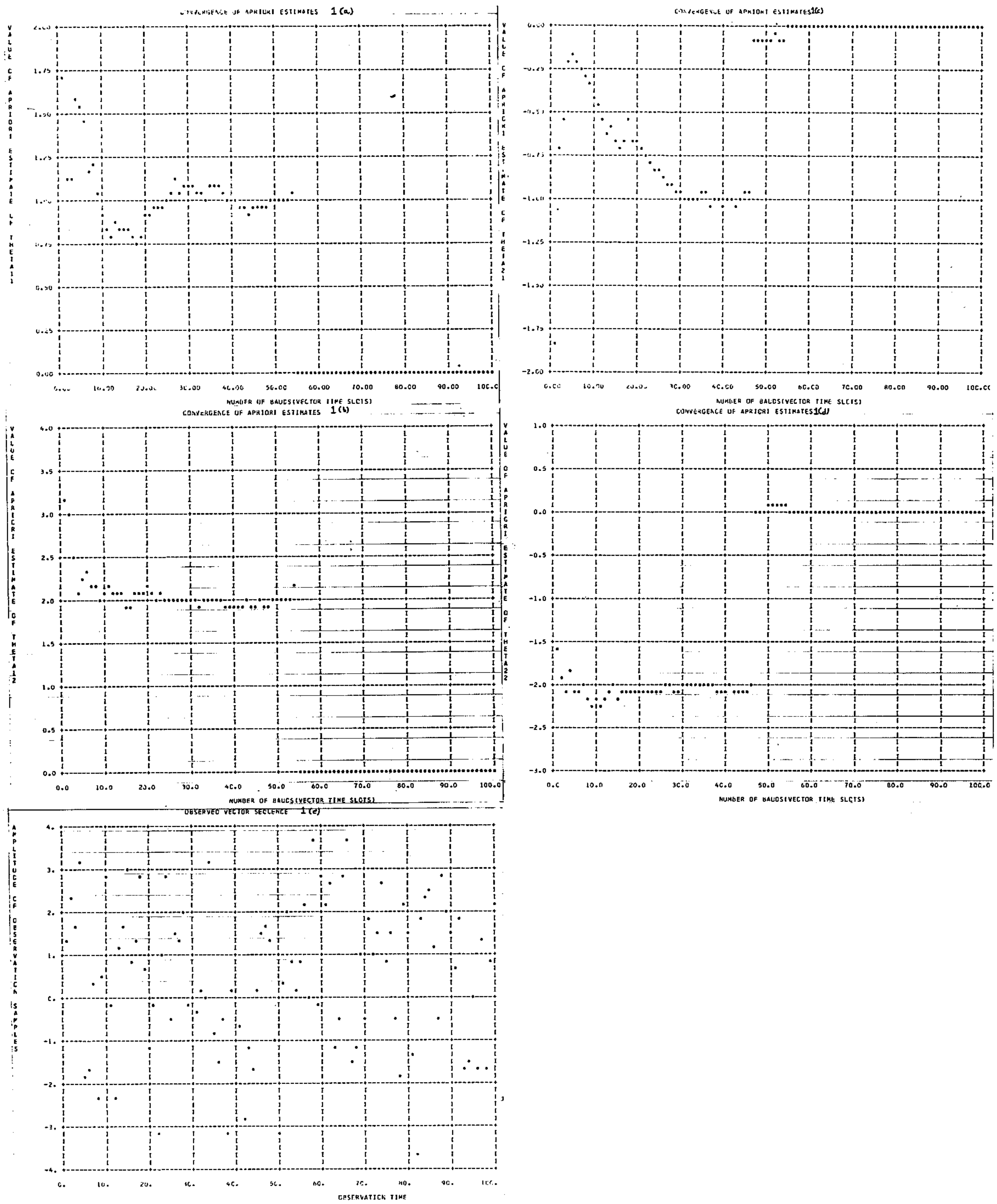


Fig. 1

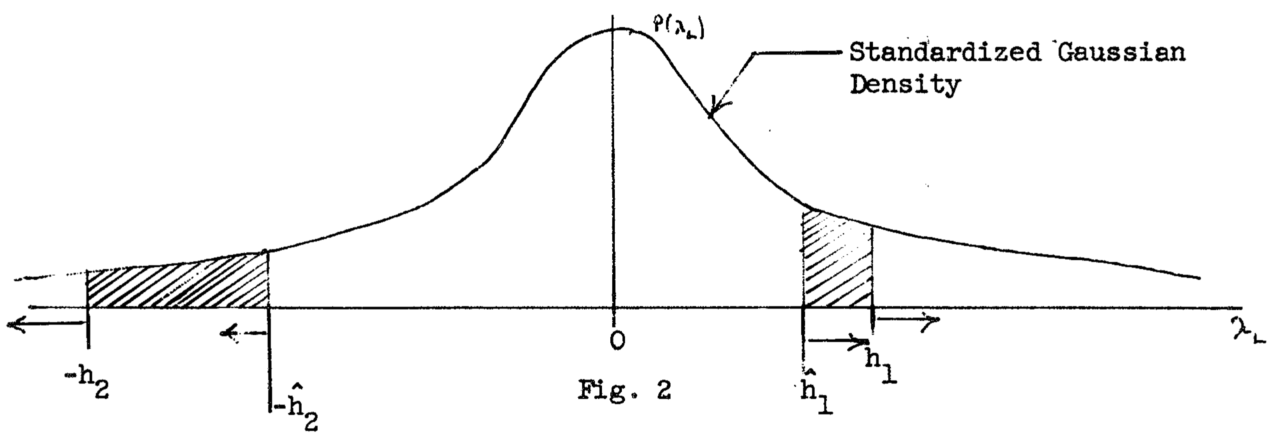

Standardized Gaussian Density
$\rho(\lambda_L)$
0
λ_L
$-h_2$
$-\hat{h}_2$
$\hat{h}_1$
h_1

Fig. 2

C. ADAPTIVE DETECTION WITHOUT SYNCHRONIZATION

J. C. Hancock

T. L. Stewart

1. Introduction

The detection theory area of the communication sciences has been concerned with the classification and recognition of signals in the presence of noise. In the general M-ary detection problem, the signal sequence is constructed from a family of M possible waveforms. The detection problem then consists of determining, in each signal interval, which of the M possible waveforms was transmitted. Disturbances to the transmitted signal which prevent correct identification of the signal are of two general types:[1]

1) Additive disturbances characterized by an additive noise term at the receiver.

2) Multiplicative disturbances which directly alter the transmitted waveshape.

The receiver structure for detecting deterministic signals in gaussian noise has been extensively studied.[2,3,4] These techniques are heavily dependent on the availability of complete a priori information on the transmitted signals and on the effects of the disturbances. Generally, the receiver must know the waveshape of all possible transmitted signals and their exact arrival times. Recent work has been concentrated in the areas of supervised and unsupervised detection methods.[5,6,7] In these cases, incomplete knowledge of the signal waveshapes is assumed and the receiver must adapt to the actual transmitted waveshape.

A particularly troublesome problem in the area of adaptive detection, and also in the conventional detection case where the signal parameters are completely known, is that of locating the arrival time of the transmitted sequence. Lack of synchronization gives poor or else completely meaningless performance in the cases mentioned previously. Uncertainty in determining the signal baud can occur in several ways. An unknown time delay in the channel results in a synchronization error. Special techniques for obtaining synchronization have been developed, i.e., sending synchronization pulses or else synchronizing codes; however, there is still uncertainty due to the presence of noise in the received data.

This study is devoted to the adaptive detection problem, with the lack of time synchronization as a fundamental assumption. Of major concern is the establishment of a receiver structure that will simultaneously adapt to the unknown signal parameters and unknown synchronization. It is also assumed that no auxiliary synchronization symbols are transmitted. Since implementation of the more sophisticated detection techniques in use today necessarily involves digital processing, it will be assumed that the data presented to the receiver is in digital form. The model for the system is shown in Fig. 1.

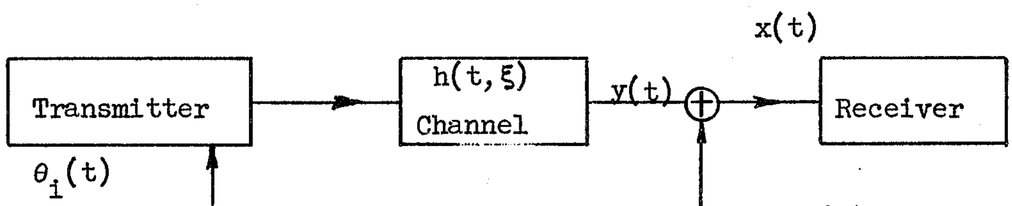


Fig. 1

Every T seconds, the transmitter selects one of a finite set of waveshape $\underline{\theta_i}(t)$, $1 \leq i \leq M$. The channel alters the transmitted waveshape such that $y(t) = \int_{-\infty}^{\infty} h(t,\xi)\, \theta_i(\xi)\, d\xi$. Noise, n(t), is added, and the resultant signal X(t) is presented to the receiver. The receiver must now decide which of the M possible signals was sent.

2. Approach

The receiver is presented with a sampled time sequence, $x(t_1)$, $x(t_2)$, $x(t_3)$,..... The duration or baud length is known at the receiver; however, the receiver does not know the instant of time, t_i, that the signal sequence begins. Assume that the signal is characterized by N time samples. Index the first N samples by T_1, T_2,...,T_N. It is clear that for the conditions assumed, one of these points, T_i, is the correct reference time for the sequence. The time series will be partitioned into bauds of length N and the symbol $\underline{X}_k$ will designate the k^{th} baud. The symbol $\underline{Z}_k$ will be used to designate the observation sequence $\underline{X}_1$, $\underline{X}_2$,...,$\underline{X}_k$. The problem is to pick which of the possible T_i is the correct reference time. This will be done by choosing the T_i that maximizes the a posteriori probability

$$P(T = T_1/\underline{Z_k})$$
$$\vdots$$
$$P(T = T_N/\underline{Z_k})$$

After the k^{th} baud is received, it can be shown that

$$P(T = T_i/\underline{Z_k}) = P(T = T_i/\underline{X_k},\underline{Z_{k-1}}) = \frac{P(\underline{X_k}/T = T_i,\underline{Z_{k-1}})\,P(T = T_i/\underline{Z_{k-1}})}{\sum_{i=1}^{N} P(\underline{X_k}/T = T_i)\,P(T = T_i/Z_{k-1})}$$

Therefore, $P(T = T_i/\underline{X_k},\underline{Z_{k-1}})$, can be computed from $P(T = T_i/\underline{Z_{k-1}})$ if $P(\underline{X_k}/T = T_i,\underline{Z_{k-1}})$ can be determined

$$P(\underline{X_k}/T = T_i,\underline{Z_{k-1}}) = \sum_{j=1}^{M} P(\underline{X_k}/T = T_i,\underline{Z_{k-1}},\underline{\theta_j})\,P(\underline{\theta_j}/T = T_i,\underline{Z_{k-1}})$$

If the signal vectors, $\underline{\theta_j}$, are known, then $P(\underline{X_k}/T = T_i,\underline{Z_{k-1}},\underline{\theta_j})$ and $P(\underline{\theta_j}/T=T_i,\underline{Z_{k-1}})$ can be determined. If the vectors $\underline{\theta_j}$ are unknown, then an iterative scheme must be used to determine these functions. Iterative solutions to this problem have been developed.[5,6,7]

Thus, an iterative approach exists for the adaptive detection problem without synchronization. The solution is optimum in the sense that at each decision instant, the reference time is chosen that minimizes the risk. In addition, the proposed method provides a means of synchronization for the conventional detection problem without using any additional synchronization signals. Figures 2, 3, 4, and 5*illustrate computer simulations of the proposed technique for various assumptions on the signal and noise.

*All shown on page 17 and 18 of this report.

REFERENCES

1. Sherman, H., "Channel Characterization," Ch. 5 of Lectures on Communication Theory, ed. E.J. Baghdady, McGraw-Hill Book Co., New York, 1961.

2. Woodword, P.M. Probability and Information Theory with Applications to Radar, Pergamon Press, London, England, 1953.

3. Price, R. "Optimum Detection of Random Signals in Noise," IRE Trans. on Information Theory, vol. IT-2, Dec., 1956.

4. Helstrom, C., Statistical Theory of Signal Detection, Pergamon Press, London England, 1960.

5. Braverman, D., "Learning Filters for Pattern Recognition," Trans. IRE, IT-8, July, 1962.

6. Daly, R.F., "Adaptive Binary Detectors," Stanford Electronics Labs., TR. No. 2003-2, June, 1961.

7. Fralick, S.C., "The Synthesis of Machines which Learn without a Teacher," Stanford Electronics Labs., TR. No. 6103-8, April, 1964.

Binary Case, Known Signal; P(T) - Average of 20 Experiments

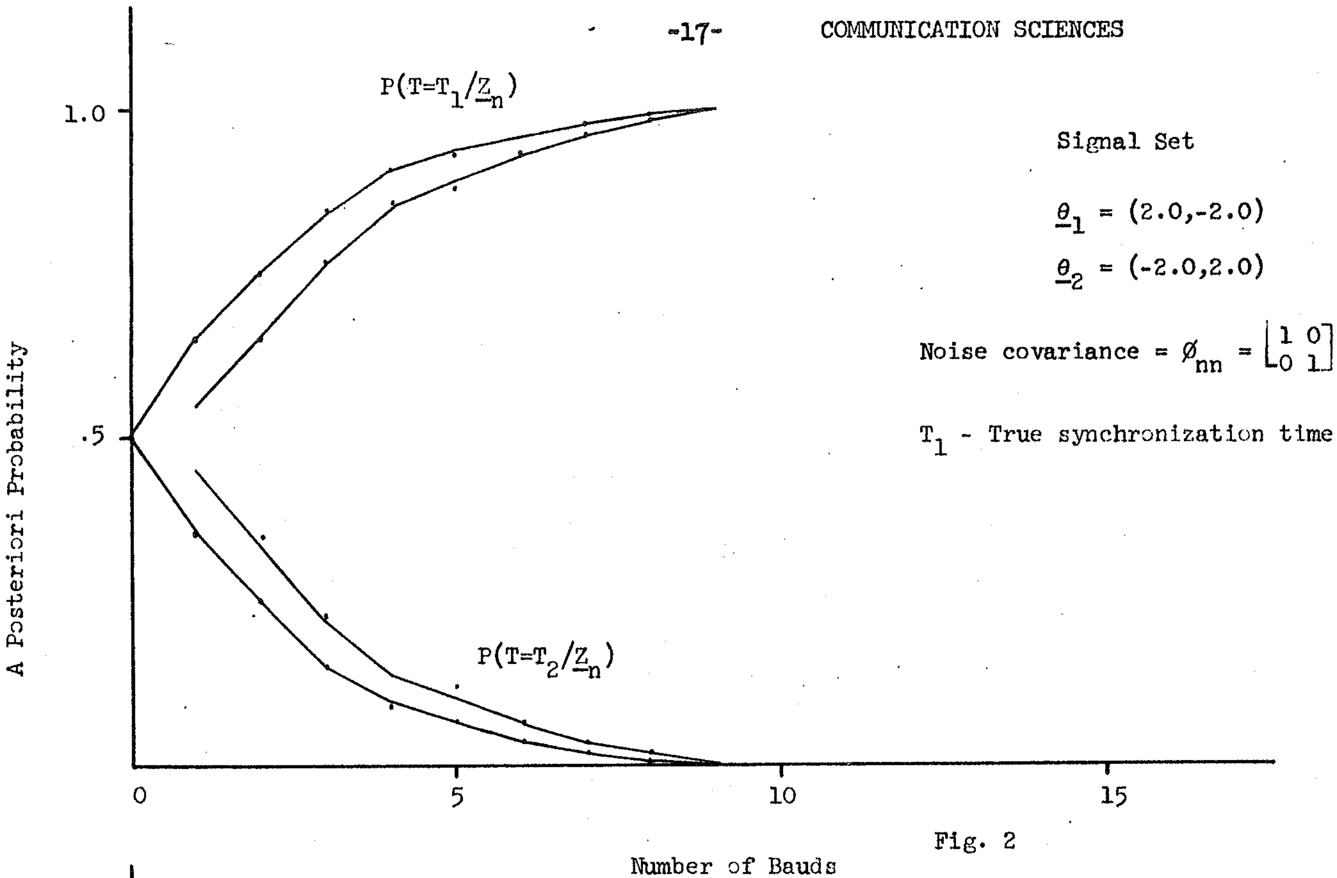


Fig. 2

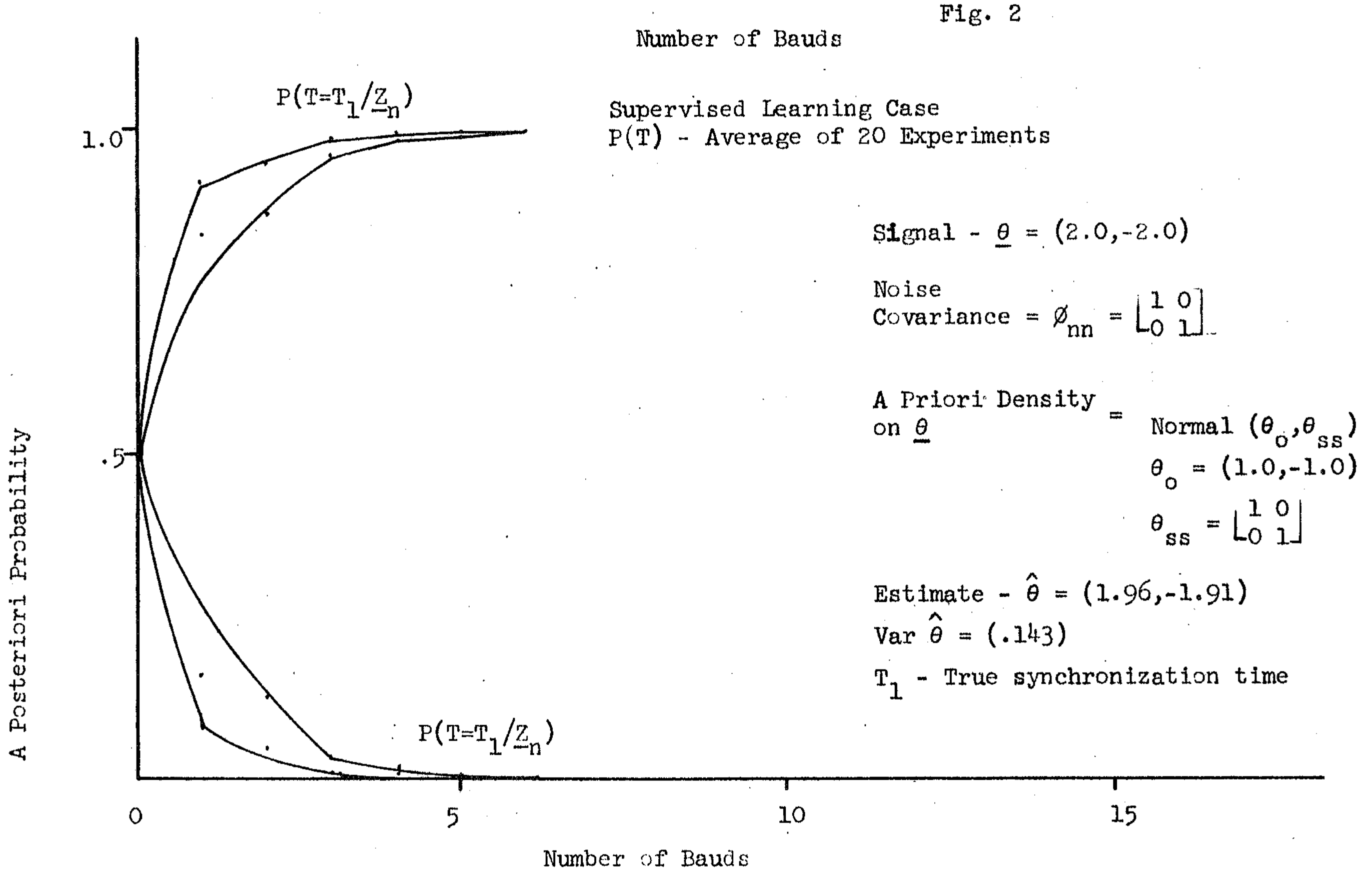


Fig. 3

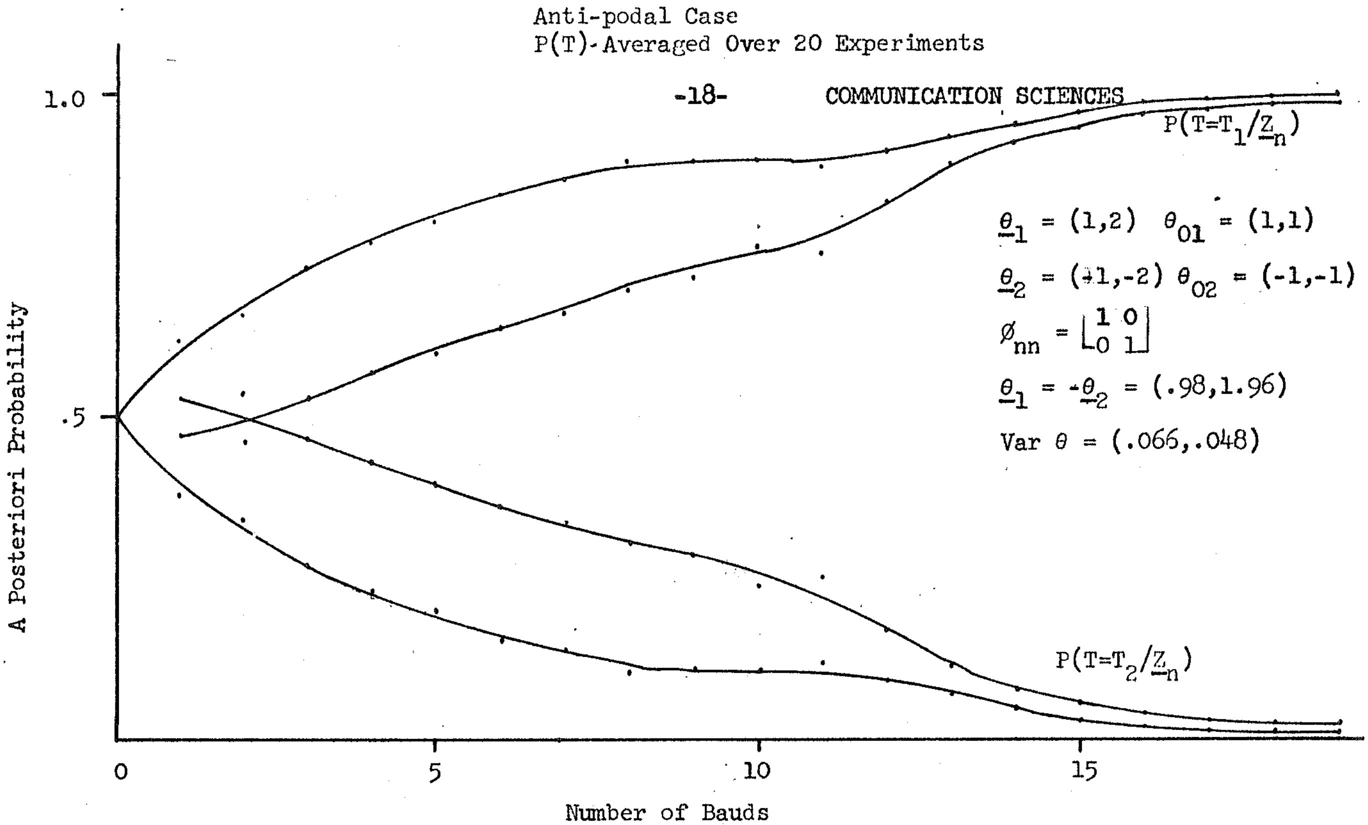

Anti-podal Case
P(T)-Averaged Over 20 Experiments
1.0
.5
A Posteriori Probability
P(T=T1/Zn)
P(T=T2/Zn)
θ1 = (1,2) θ01 = (1,1)
θ2 = (-1,-2) θ02 = (-1,-1)
Ønn = [1 0; 0 1]
θ1 = -θ2 = (.98,1.96)
Var θ = (.066,.048)
0
5
10
15
Number of Bauds

Fig. 4

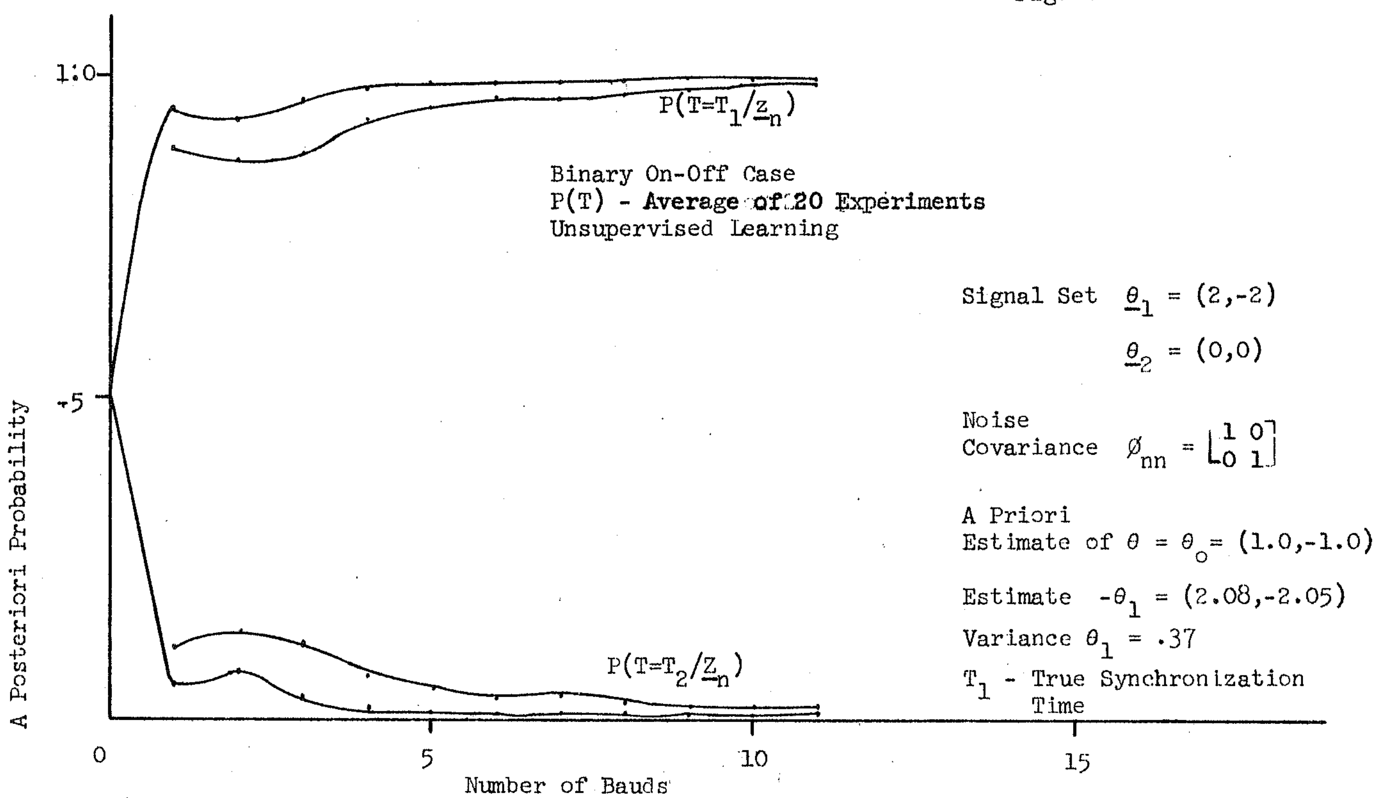

1.0
.5
A Posteriori Probability
P(T=T1/Zn)
Binary On-Off Case
P(T) - Average of 20 Experiments
Unsupervised Learning
Signal Set θ1 = (2,-2)
θ2 = (0,0)
Noise
Covariance Ønn = [1 0; 0 1]
A Priori
Estimate of θ = θo = (1.0,-1.0)
Estimate -θ1 = (2.08,-2.05)
Variance θ1 = .37
T1 - True Synchronization
Time
P(T=T2/Zn)
0
5
10
15
Number of Bauds

Fig. 5

II. ADAPTIVE SYSTEMS

Most of the results concerning the transmission of information via digital communication systems are based on certain assumptions concerning the channel environment. In particular, the channel statistics are usually assumed known so that the problem of finding optimum detector structures can be solved, at least in principle, by the methods of statistical inference. When adequate a priori information concerning the channel environment is not available, however, the standard decision theory formulation of the detection problem is not applicable. Under these circumstances, it is reasonable to consider receivers capable of extracting knowledge of the channel conditions directly from the information bearing signals as they emerge from the channel. Systems that attempt to measure the unknown properties of the channel, and use these measurements in the processing of the received data, are sometimes termed "adaptive." Note, however, that it is not really necessary to learn the properties of the channel characteristics. Since the purpose of the binary receiver is to make decisions as to which of two signals is present in the received data, the receiver need only learn how to distinguish between the signals - i.e., the distinguishing features of the two types of received data. Both theoretical and experimental investigations of this adaptive receiver approach to the problem of communicating through unknown channels are presently in progress.

A. PROBABILITY OF ERROR FOR CORRELATORS WITH NOISY REFERENCE SIGNALS

P. A. Wintz

Correlation detection is an optimum strategy for detecting signals in noise for various combinations of criteria of goodness and assumptions concerning the observed data. A simple correlator is shown in Fig. 1. The input waveform $x(t)$ consists of a deterministic signal $s(t)$ of T seconds duration plus a random perturbation n(t) called noise. The reference waveform y(t) also consists of a deterministic signal r(t) of T seconds duration plus a random component m(t) also called noise. Ordinarily, both n(t) and m(t) are assumed to be gaussian random processes. The output of the integrator z(t) is sampled at t = T, and the statistic z(T) used in the decision process. Hence, the average performance of the correlation detector depends on the statistics of the variate z(T). An attempt to determine the probability law governing z(T)

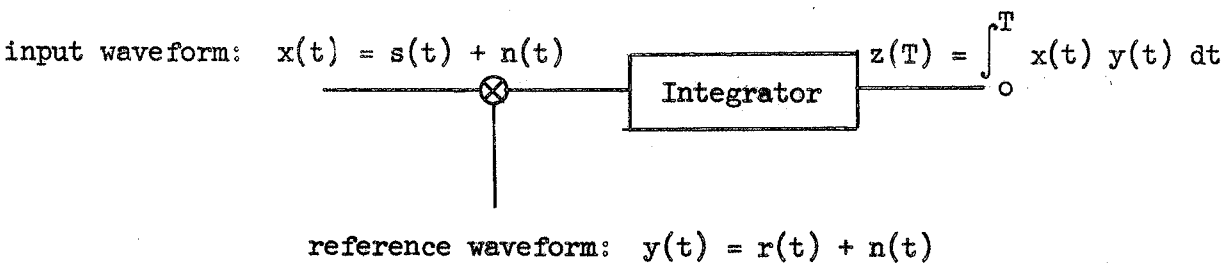


Fig. 1.

directly from

$$z(T) = \int_0^T [s(t) + n(t)]\,[r(t) + m(t)]\,dt$$

leads to severe mathematical difficulties. However, by employing an appropriate set of orthonormal basis functions $\{\psi_i(t)\}$, the required waveforms can be approximated by finite dimensional vectors which, in turn, can be used to write

$$z(T) \doteq z = \int_0^T \sum_{j=1}^{2N} x_j \psi_j(t)\,dt$$

$$= \sum_{j=1}^{2N} x_j y_j = \sum_{j=1}^{2N} (s_j + n_j)(r_j + m_j)$$

The problem of determining the probability law governing z(T) and/or z has already been investigated for some special cases. However, no results for the case of unequal signal-to-noise ratios in the input and reference channels have been reported in the published literature. The statistics of $z(T)$ and/or z for this case are important since, in most practical situations, it is possible to obtain a reference waveform of considerably higher signal-to-noise ratio than the input waveform. Another important case, not previously considered, is that of non-identical signals in the two channels, i.e., $r(t) \neq s(t)$. This situation may arise due to imperfect time synchronization, e.g., $r(t) = s(t-\tau)$, or when it is necessary to estimate $s(t)$ from previously received data and an unbiased estimate is not available. The characteristic function, probability density function, and the probability distribution function for the random variable z have been computed for the general case, subject only to the following restrictions: both signal vectors have finite energy, i.e., $\sum_{j=1}^{2N} s_j^2 < \infty$, $\sum_{j=1}^{2N} r_j^2 < \infty$; the random variables n_j, $j=1,\ldots,2N$ form a set of 2N mutually independent gaussian random variables of identical variances, i.e., $\overline{(n_j-\overline{n_j})(n_k-\overline{n_k})} = \sigma_n^2\,\delta_{jk}$; the random variables m_j form a set of 2N mutually independent gaussian random variables of identical variances, i.e., $\overline{(m_j-\overline{m_j})(m_k-\overline{m_k})} = \sigma_m^2\,\delta_{jk}$. Note, however, that no loss of generality results from assuming that both the n_j and m_j have zero mean and unit variance. For an unbiased correlation detector the probability of a decision error is simply the probability that the correlator output does not exceed zero, i.e.,

$$P_E = \text{Prob}\left[\sum_{j=1}^{2N}(s_j+n_j)(r_j+m_j) < 0\right]$$

It has been shown that

$$P_E = \frac{1}{2} - \frac{\exp\left[-E_a\right]}{2^{2N}} \sum_{k=0}^{\infty} \frac{(E_a/4)^k}{k!} \sum_{j=0}^{\infty} \frac{(E_a/2)^{j+1}}{(j+1)!} \sum_{n=0}^{j} \binom{j+2k+2N+n}{k+N+n} \binom{j}{n} \left(-\frac{1}{2}\right)^n$$

where

$$E_a = \frac{1}{2}\sum_{j=1}^{2N}(s_j^2 + r_j^2) = \text{average of the energies of } s(t) \text{ and } r(t)$$

$$E_c = \sum_{j=1}^{2N} s_j r_j = \text{cross energy of } s(t) \text{ and } r(t).$$

Note that performance of the correlator does not depend on the waveshapes of the signal components in the two channels. All waveform pairs s(t), r(t) having the same average and cross energies perform equally well. It is also interesting that the correlator output does not depend on the two signal energies $\sum s_j^2$ and $\sum r_j^2$, but only on their average.

The probability of error P_E is easily evaluated for various values of N, E_c, and E_a on a digital computer. Fig. 2 illustrates the dependence of P_E on N, E_a, and E_c for selected values of these parameters. Next, consider the case of identical signal components, but different signal-to-noise ratios in the input and reference channels. For this case

$$z = \sum_{j=1}^{2N} (s_i + n_i)(\beta s_i + m_i)$$

If we let E represent the signal energy, i.e.,

$$E = \sum_{j=1}^{2N} s_j^2$$

then

$$E_a = \frac{E}{2}(1 + \beta^2);$$

$$E_c = \beta E$$

In Fig. 3, P_E is plotted vs. E for selected values of β and N.

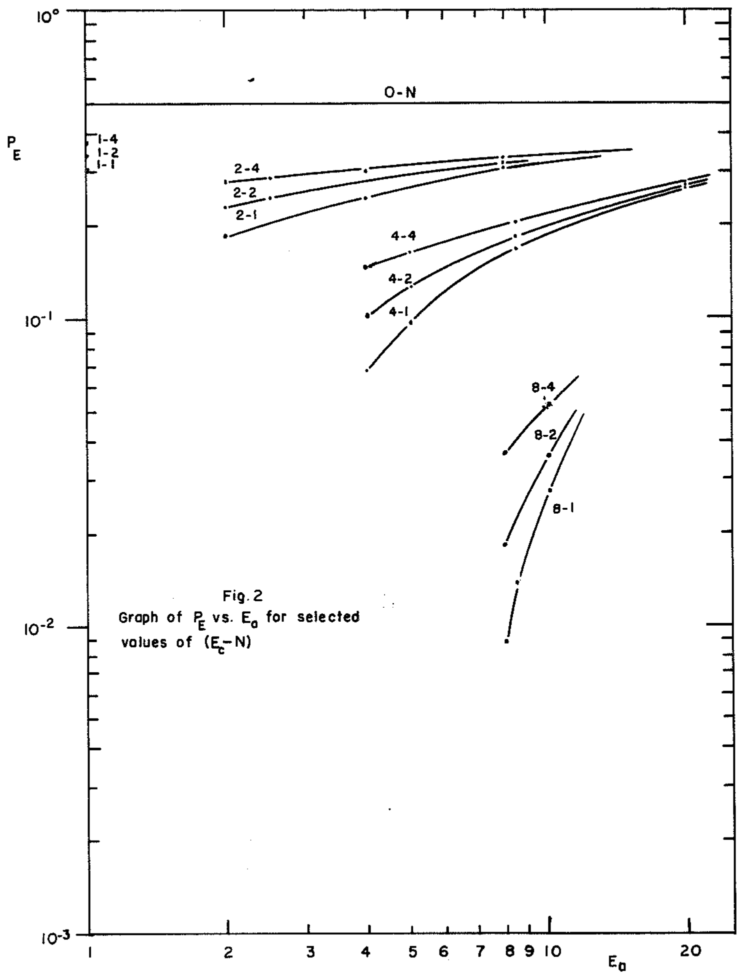


Fig. 2
Graph of P_E vs. E_a for selected values of (E_c-N)

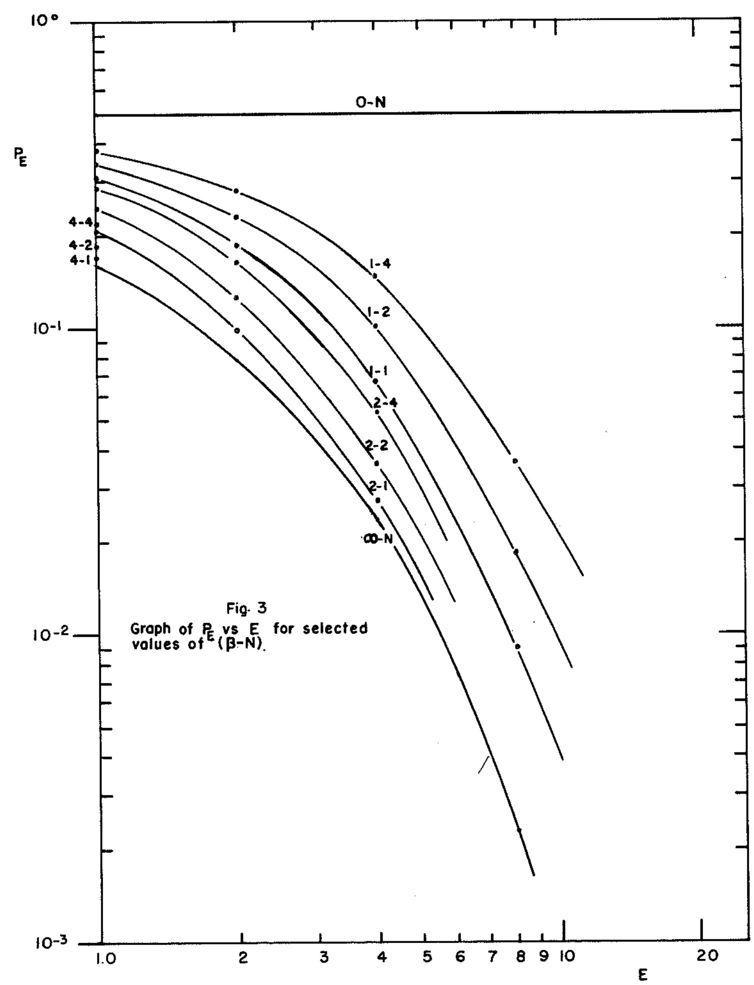


Fig. 3
Graph of P_E vs E for selected values of $(\beta-N)$.

B. PURDUE CHANNEL SIMULATOR

P. A. Wintz

R. A. Markley

This project is concerned with a water tank simulator of randomly time-varying communication channels. Hydrophones are used to connect the transmitter and receiver to the "channel" just as antennas are used to couple transmitters and receivers to atmospheric channels. The advantage of the water tank simulator is due to the effective scaling of path lengths, antenna beam widths, etc. by a factor that allows the channel to be contained within the laboratory. Hence, experiments under controlled laboratory conditions are possible. For example, the ratio of the velocity of electromagnetic radiation in free space to the velocity of sound in water is approximately 2×10^5. Therefore, distances, wavelengths, delays, etc. are scaled by this factor. We find that kilometers are scaled to centimeters and transducers (antennas) of a few centimeters diameter produce beam widths of a few degrees at megacycle frequencies. Although the attenuation of sound in water increases with frequency, the path lengths are short enough that signaling at megacycle frequencies requires a transmitter power of approximately 100 milliwatts. The tank, incidentally, is lined with sound-absorbing material so that reflections from the sides of the tank are not a problem.

By adjusting various parameters, the channel characteristics can be adjusted to fit the user's needs. Placing the transmitter and receiver transducers facing each other at two points in the tank results in an arrangement that produces a strong specular component. The transit time can be adjusted by changing the distance between the transducers. Scatter components are easily introduced by releasing air from nozzles at the bottom of the tank. The scatter component introduced by the air-bubble-water interfaces can be controlled by regulating the air nozzles. For example, the strength of the scatter component depends on the number of bubbles, while its spectrum depends on the size of the bubbles. The specular component can be removed completely by a slightly different arrangement of the transducers. For example, a scatter path can be simulated by placing the transducers side by side at one end of the tank with the air nozzles in front of them. The spread of the multipath structure (range of delays) can be adjusted by changing the distances to the closest and fartherest nozzles. This arrangement also allows for one or more specular components (fixed paths of different lengths) to be introduced by inserting one or more reflectors at selected distances from the transducers. The strength of the specular component relative to the scatter component can be adjusted by adjusting the size of the fixed reflector relative to the number of bubbles. Finally, doppler shifts can be introduced by moving the medium relative to the transducers. Products of maximum doppler shift (cycles/sec) times maximum delay (seconds) on the order of 5 can be obtained.

Peripheral equipment already constructed includes a transmitter, receiver, and a device for measuring the statistics of the channel output. The receiver contains a linear envelope detector that can be switched either in or out. The "statistics" circuit employs two Schmidt triggers with slightly different firing levels followed by a logic circuit. The logic circuit determines the fraction of time the waveform spends in the window between the two firing levels and presents this information as a reading on a counter. By changing the DC level of the input signal a fraction of time, histogram of the signal amplitude (or envelope amplitude) is easily constructed.

The Purdue Channel simulator is available for use by any interested person. It is expected that some users will attempt to simulate real-life channels.

Other users will use the simulator as a convenient source of perturbed signals to test adaptive receivers designed to operate efficiently for a variety of channel environments.

III. SIGNAL DESIGN

A. EXPERIMENTAL RESEARCH ON COMMUNICATION SYSTEMS SUBJECT TO INTERSYMBOL INTERFERENCE

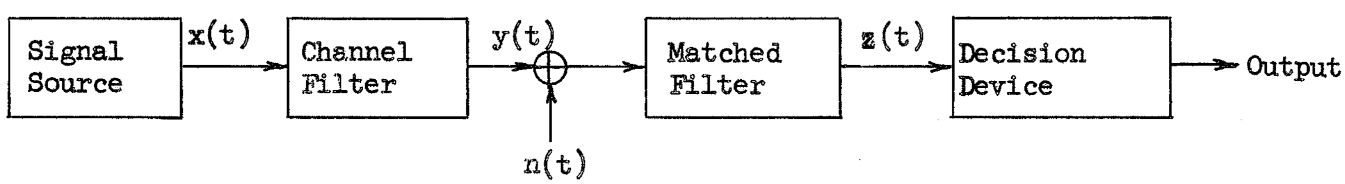
J. C. Lindenlaub

C. C. Bailey

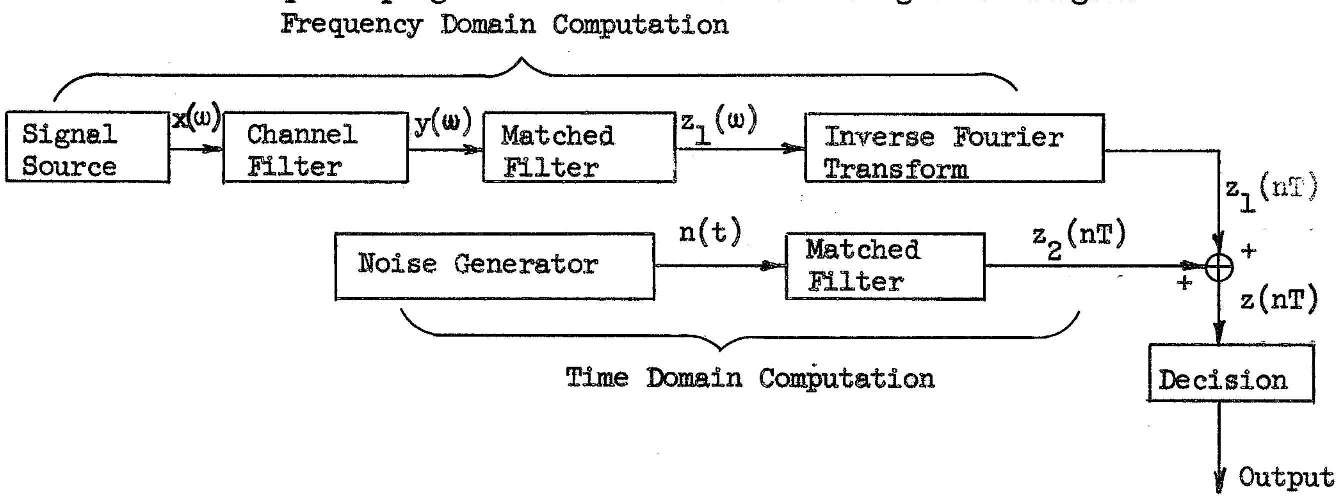
In order to experimentally determine the effects of intersymbol interference phenomena on the performance of digital communication systems, two experimental models of binary digital communication systems subject to intersymbol interference are being developed. The two models are 1) a laboratory model and 2) a computer simulation model. The purpose of these models is to investigate the degradation in probability of error in matched-filter type digital communication systems when controlled amounts and types of intersymbol interference causing distortion are introduced. Methods of reducing the effects of such distortion can also be investigated with these systems.

Computer Model

The digital computer simulation model is designed to simulate the following system:



The program includes a counter to detect and count the detection errors made by the system. The actual computer program is based on the following block diagram.

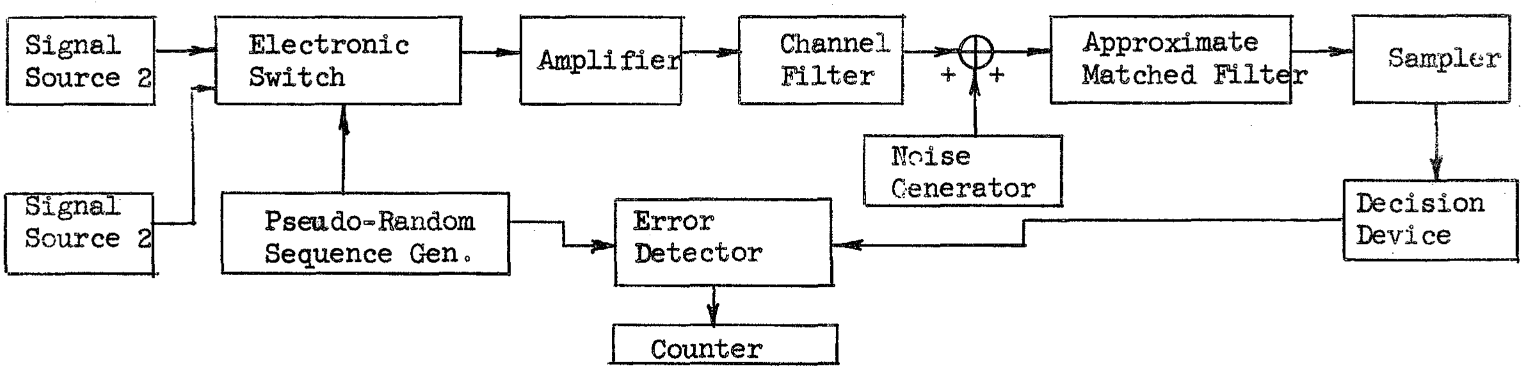


The computations made on the signal as it passes through the system are done with frequency domain specifications--the magnitude and phase functions associated with the signal's Fourier transform. This was done since it was desired to specify channel filters which may not be representable as finite lumped-parameter or simple distributed-parameter circuits. Thus the spectrum of the signal portion of z(t), $z_1(t)$, is computed and transformed into samples of the appropriate time-domain signal component at the input of the decision device. The noise process is most easily simulated by the use of time samples derived from pseudo-random number generation on the computer. This is followed by convolution of the time samples with the impulse response of the matched filter to produce the noise component of the input to the decision device.

This computer simulation model is presently being used in an investigation of binary two-phase synchronous systems and binary unipolar AM systems. In each case a raised cosine spectrum for the transmitted signal is used and an ideal matched filter for this receiver is employed at the receiver.

Laboratory Model

The laboratory model communication system was designed according to the following block diagram.



This model was built with transistorized electronic circuitry, and operates with a center frequency of 455 kc. The system presently operates with unipolar AM square pulse modulation. It can be modified to operate as an FSK or PSK system. The signal spectrum is shaped at the transmitting amplifier by a single-tuned RLC circuit. The approximate matched filter consists of a single-tuned RLC circuit followed by an envelope detector. The error detector makes logical comparisons between the output of the pseudo-random sequence generator and the output of the decision device, and produces two sets of pulses for input to the counter, one set indicating false alarm errors and the other indicating false dismissal errors.

Both of the communication system models described above allow probability of error performance to be determined for a wide range of channel distortion filters and modulation schemes for varying signal-to-noise ratios. This enables the study of the effects of intersymbol interference under the influence of controlled amounts of signal distortion and noise power. The digital computer model utilizes a more ideal representation of the signal transmission and detection processes, and allows somewhat more freedom in the specification of the nature of the channel distortion. The laboratory model utilizes a more practical representation of the signal transmission and detection processing found in existing systems.

B. OPTIMUM WAVEFORMS AND RECEIVERS FOR CHANNELS WITH MEMORY

J. C. Hancock

E. A. Quincy

1. Statement of the Problem

The specific problem considered in this research is the optimization of entire binary communication systems, i.e. joint optimization of the transmitted pulse waveforms and the receiver when the channel response is time-invariant and known. Also, the channel is assumed to exhibit sufficient memory such that intersymbol interference results at the receiver. The criterion of optimality considered is minimum average probability of detection error. For recent literature pertinent to this problem, see Refs. 1,2,3, and 4.

Figure 1 shows a model of the binary communication system considered in this research. The additive noise is assumed stationary, gaussian with zero mean and covariance $\phi_n = NI$ such that the noise samples are assumed to be statistically independent with variance (or noise power) N. Also, the received signal is assumed to be representable by a finite sum of weighted basis functions such that the weighting coefficients form a K-dimensional vector denoted by a bar beneath an upper-case letter.

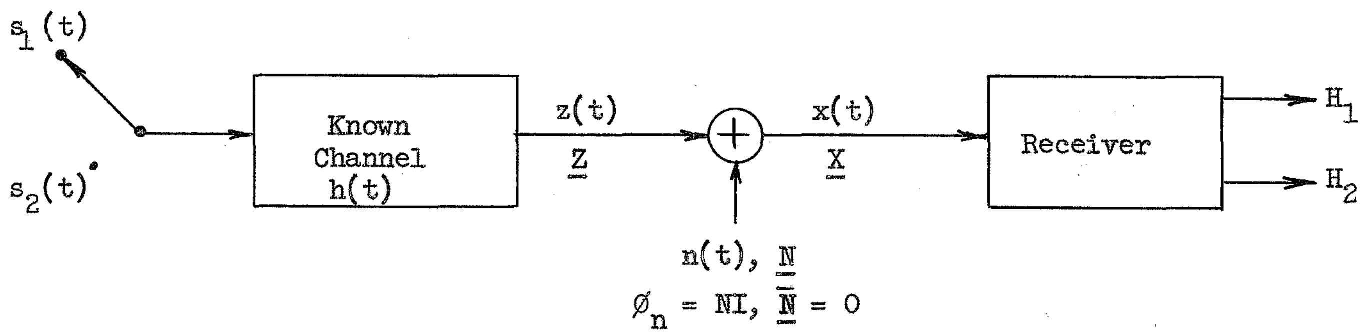


Figure 1

Binary Communication System Model

2. Assumptions

The following assumptions were employed in this research:

1) $\underline{N}$ is zero-mean gaussian with independent samples.

2) z(t) is representable by K numbers, $\underline{Z}$

3) h(t) is time-invariant, known and exhibits M bauds of memory; i.e., an input

pulse of duration T is stretched to (M+1)T at the output.

4) Signaling rate is 1/T.

5) Receiver synchronized to transmitter.

6) Receiver has zero-memory with observation period [0, (M+1)T].

3. Approach

Ideal Approach

An ideal approach to this problem is to derive the Bayes receiver from the Maximum Likelihood Ratio, $\lambda(x)$, for arbitrary transmitted signal waveforms, $s_1(t)$ and $s_2(t)$, of length T and an arbitrary time-variant channel with impulse response h(t). Then, the average probability of error P_e would be derived for this receiver. The resulting functional, P_e, would depend on the signal parameters: energies $\underline{E}$, all possible cross-correlations $\underline{\rho}$ of the two desired signal waveforms, with all possible combinations of received waveforms of overlapping pulses. P_e would also depend on the noise power N and the channel impulse response h(t). However, these quantities would be fixed for any given problem. Hence, we can denote the probability of error by

$$P_e(s_1, s_2, \underline{E}, \underline{\rho}) = \int_R f(\underline{u}) d\underline{u} \tag{1}$$

With an explicit expression available for P_e, it would then be a matter of applying variational techniques to minimize P_e with respect to the transmitted waveforms, with possible physical constraints applied, i.e.,

$$\min_{s_1, s_2} \; P_e(s_1, s_2, \underline{E}, \underline{\rho}) - \lambda_1 f(s_1, s_2) - \lambda_2 g(s_1, s_2) \tag{2}$$

with

$$f(s_1, s_2) = f_o \tag{3}$$

$$g(s_1, s_2) = g_o \tag{4}$$

This would produce the optimum waveforms for a specified channel.

Approach Used

The approach used in this research differed from the ideal above since an explicit for P_e in terms of signal parameters could not be obtained. However, the integral in (1) was numerically integrated; and by plotting a family of curves, the effect of the pertinent signal parameters on P_e was obtained. The two significant signal parameters obtained for bipolar signals and M=1 were E_o and E_{ht};i.e., the energy out of the channel

$$E_o = \int_o^{2T} z^2(t)\, dt \tag{5}$$

and the head-tail cross-correlation energy

$$E_{ht} = \int_o^{T} z(t)\, z(t + T)dt \tag{6}$$

The family of curves showed that Min P_e was equivalent to $\underset{s}{\text{Max}}\, E_o$, with constraints on the energy into the channel and on the head-tail cross-correlation energy; i.e.,

$$E_1 = \int_o^{T} s^2(t)dt = \text{constant} \tag{7}$$

$$E_{ht} = \text{constant} \tag{8}$$

Hence, the optimum waveform was obtained from the following expression.

$$\underset{s}{\text{Max}}\, E_o - \lambda_1, E_1 - \lambda_2 E_{ht} \tag{9}$$

Bayes Receiver

The generalized, zero-memory, non-linear Bayes receiver (maximum likelihood ratio) was derived for M bauds of channel memory and shown to be

$$\lambda(\underline{x}) = \frac{P(\underline{X}|\underline{S}_1)}{P(\underline{X}|S_2)} \quad : \quad \frac{P_2}{P_1} \tag{10}$$

and

$$(\underline{x}) = \frac{\sum_{j=1}^{r} P_{1j} a_{1j} e^{\frac{1}{N}(\underline{x}^T \underline{Z}_{1j})}}{\sum_{j=1}^{r} P_{2j} a_{2j} e^{\frac{1}{N}(\underline{x}^T \underline{Z}_{2j})}} \quad : \quad \frac{P_2}{P_1} \tag{11}$$

or the equivalent modified likelihood ratio is

$$\lambda(\underline{x}) = \ln \left[\sum_{j=1}^{r} P_{1j} a_{1j} e^{\frac{1}{N}(\underline{x}^T \underline{Z}_{1j})} \right]$$

$$- \ln \left[\sum_{j=1}^{r} P_{2j} a_{2j} e^{\frac{1}{N}(\underline{x}^T \underline{Z}_{2j})} \right]$$

$$- \ln \left(\frac{P_2}{P_1} \right) \; : \; 0 \qquad (12)$$

where

$$r = 2^{2M} \qquad (13)$$

$$a_{ij} = e^{-\frac{1}{2N}(\underline{Z}_{ij}^T \underline{Z}_{ij})}, \quad i = 1,2; \; j = 1,2,\ldots,r \qquad (14)$$

5. Average Probability of Error for Bayes Receiver

General Case

The average probability of error, P_e, for the general receiver described in Eq. (12) was formulated and shown to be

$$P_e = P\left[\lambda < \frac{P_2}{P_1} \mid \underline{s}_1\right] P_1 + P\left[\lambda > \frac{P_2}{P_1} \mid \underline{s}_2\right] P_2 \qquad (15)$$

$$= \sum_{\ell=1}^{r} \left[P_{1\ell} P_1 \iint_{R_{1\ell}} \cdots \int P(\underline{w}^{1\ell}) \prod_{j=1}^{2r} dw_j^{1\ell} \right.$$

$$\left. + P_{2\ell} P_2 \iint \cdots \int P(\underline{w}^{2\ell}) \prod_{j=1}^{2r} dw_j^{2\ell} \right] \qquad (16)$$

where $P(\underline{w}^{k\ell})$ are gaussian multi-variate densities and the regions of integration are:

$$R_{1\ell} = \left\{ w_1^{1\ell}, w_2^{1\ell}, \ldots, w_{2r}^{1\ell} : w_{2r}^{1\ell} > w_{2r}^{1\ell^*} = \ell n \left[\frac{1}{kb_{2r}} \sum_{j=1}^{r} b_j e^{w_j^{1\ell}} - \frac{1}{b_{2r}} \sum_{j=r+1}^{2r-1} b_j e^{w_j^{1\ell}} \right] \right\} \qquad (17)$$

$$R_{2\ell} = \left\{ w_1^{2\ell}, w_2^{2\ell}, \ldots, w_{2r}^{2\ell} : w_{2r}^{2\ell^*} < w_{2r}^{2\ell} = \ell n \left[\frac{1}{kb_{2r}} \sum_{j=1}^{r} b_j e^{w_j^{2\ell}} - \frac{1}{b_{2r}} \sum_{j=r+1}^{2r-1} b_j w_j^{2\ell} \right] \right\} \qquad (18)$$

Special Case: Adjacent Band Overlap (M=1), Equi-Probable, Bipolar Signals

For this special case, P_e in (16) reduces to

$$P_e = \frac{1}{2} - \frac{1}{8}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\bar{\sigma}^{ij}y_i y_j}}{2\pi\sqrt{\sigma^{33}\,|\,\phi_W\,|}} \sum_{\ell=1}^{4} \operatorname{erf}\left(-\sqrt{\frac{\sigma^{33}}{2}}\, z_{1\ell}^{*}\right) dy_2 dy_1 \tag{19}$$

where

$$z_{1\ell}^{*} = f(y_1, y_2) \text{ (nonlinear)} \tag{20}$$

The expression in (19) was numerically integrated on the computer to yield

$$P_e = f\left(\frac{E_o}{N}, \rho_{ht}\right) \tag{21}$$

where

$$E_o = \int_o^{2T} z^2(t)dt \tag{22}$$

$$\rho_{ht} = \frac{\int_o^T z(t)\; z(t+T)dt}{E_o} \tag{23}$$

The sketch in Fig. 2 shows the dependence of P_e on these parameters.

1. No intersymbol interference
2. $\rho_{ht} = 0.15$
3. $\rho_{ht} = 0.2$
4. $\rho_{ht} = 0.3$

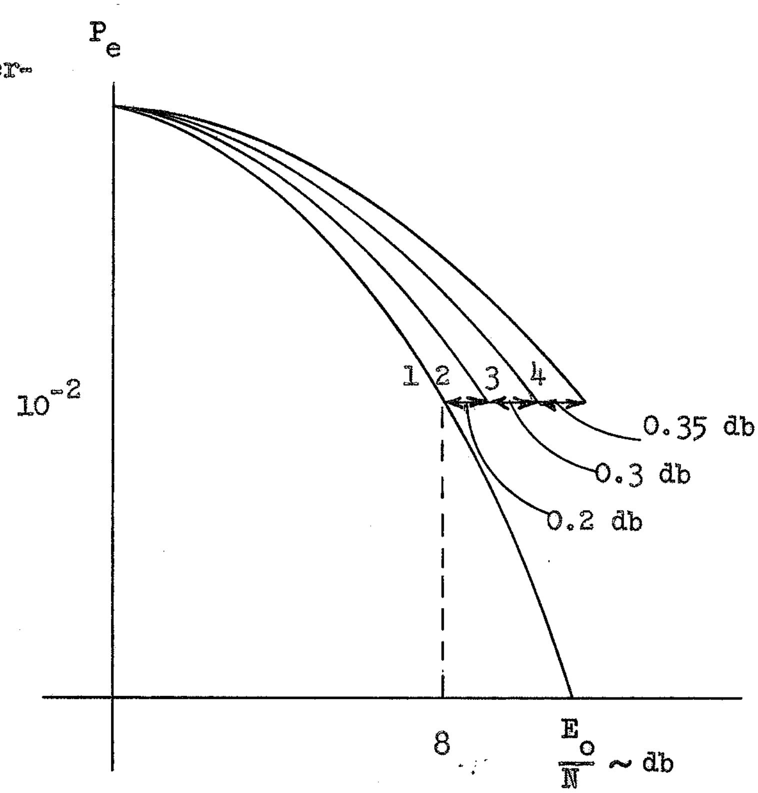


Fig. 2

From Fig. 2, we note that the nonlinear Bayes receiver is not very sensitive to changes in ρ_{ht} and its performance is very close to that of no intersymbol interference. Also, we note that for a fixed $\frac{E_o}{N}$, P_e is a monotonic increasing function of ρ_{ht}; and for a fixed ρ_{ht} and N, P_e is a monotonic decreasing function of E_o. Hence, we would like to maximize E_o and minimize ρ_{ht} for a fixed energy into the channel. Since, E_o is a function of ρ_{ht} we conclude that we should maximize E_o for a fixed energy input and fixed ρ_{ht}.

6. Signal Design

Problem

In the previous section we concluded that the signal design problem is to

$$\text{Msx } I(s) = E_o - \lambda_1 E_1 - \lambda_2 E_{ht} \tag{24}$$

where

$$E_o = \int_{-T}^{3T} z^2(t)dt \tag{25}$$

$$E_1 = \int_{-T}^{T} s^2(t)\, dt = \text{constant} \tag{26}$$

$$E_{ht} = \int_{-T}^{T} z(t)\, z(t+2T)dt = \text{constant} \tag{27}$$

and λ_1, λ_2 are Lagrangian multipliers.

The signal is transmitted on (-T,T) and observed on (-T, 3T) for M = 1 in order to take advantage of the symmetry. Eq. 24 can be rewritten as

$$\text{Max } I(s) = \int_{-T}^{T} \left[\int_{-T}^{T} s(t)\, s(\tau)\, K(t,\tau)\, d\tau - \lambda_1\, s^2(t) \right] dt \tag{28}$$

where

$$K(t,\tau) = \int_{-\infty}^{\infty} |H(f)|^2\, e^{j\omega(t-\tau)}\, df$$

$$\frac{-\lambda_2}{2} \int_{-T}^{T} \left[H(x-\tau)\, h(x+2T - t) + h(x-t)\, h(x+2t-\tau) \right] dx \tag{29}$$

and h(t) is the channel impulse response. H(f) is the Fourier transform of h(t).

General Solution

By applying a first variation to s(t) in (28), the solution for the optimum s(t) was obtained in terms of the following Fredholm integral equation of the second kind:

$$s(t) = \lambda \int_{-T}^{T} s(\tau)\, K(t,\tau) d\tau \tag{30}$$

where

$$\lambda = \frac{1}{\lambda_1} \tag{31}$$

and

$$K(t,\tau) = K(\tau,t) \tag{32}$$

RC - Lowpass Channel Solution

For this particular channel

$$h(t) = \begin{cases} \alpha = e^{-\alpha t} & , \quad t \geq 0 \quad (33a) \\ 0 & , \quad t \leq 0 \quad (33b) \end{cases}$$

where

$$\alpha = \frac{1}{RC} \tag{34}$$

Substituting (33) into (29) and (29) into (30), an expression was obtained which can be differentiated twice with respect to t to obtain

$$s''(t) + \alpha^2 (\lambda - 1)\, s(t) = 0 \tag{35}$$

Hence, the form of the solution for the RC channel is

$$s(t) = A\, e^{s_1 t} + A^* e^{s_1^* t} , \quad -T \leq t \leq T \tag{36}$$

where

$$s_1 = j\, \alpha \sqrt{\lambda - 1} \tag{37}$$

$$s_1^* = -s \tag{38}$$

and A and τ are determined by substituting (36) back into the integral (30).

REFERENCES

1. Quincy, E.A., 1st Semi-Annual Research Summary, "Digital Communication Systems Optimization for Channels with Memory", School of Electrical Engineering, Purdue University, July-December 1964, page 33, Communication Sciences, Section 4.

2. Tufts, D.W. and Aron, J.R., "Intersymbol Interference and Estimation of Digital Message Sequences", International Conf. on Microwaves, Circuit Theory and Information Theory, Summaries of Papers, Part 3, Information Theory, pp. 55-56, Tokyo, Japan, September 1964.

3. Tufts, D.W., "Nyquist's Problem - The Joint Optimization of Transmitter and Receiver in Pulse Amplitude Modulation," Proc. of IEEE; March 1965, pp: 248-259.

4. Tufts, D.W., and Shnidman, D.A., "Effect of Correlated Data and Colored Noise on Error Probability and Mean Square Error in PAM Data Transmission," IEEE Int'l. Conv. Record, Part 7; March 1965, pp. 227-237.

N66 20105

C. SIGNAL SELECTION FOR TELEMETRY CHANNELS

R. E. Totty

In the telemetry problem, a waveform $x(t)$ is to be transmitted in such a fashion that the received waveform $x_*(t)$ is a scaled replica of $x(t)$. It is assumed here that a digital transmission scheme is used. That is, the waveform $x(t)$ is "sampled"; the samples are quantized; the quantized samples are associated one-to-one with a set of transmission waveform; and the receiver decides which one of these possible waveforms was transmitted. The received "samples" are then used to produce the reconstructed waveform $x_*(t)$. There are three sources of error. First, there is the initial approximation error, denoted by ϵ_a^2, which is the error incurred by describing the waveform by a finite number of samples. A quantization error, denoted by ϵ_q^2, is incurred by quantizing the samples. Another source of error is the error due to incorrect decisions made at the receiver, which we call channel error, and is denoted by ϵ_*^2. It would be convenient, as it is sometimes erroneously assumed to be, if the total error, ϵ^2, were represented by

$$\epsilon^2 = \epsilon_a^2 + \epsilon_q^2 + \epsilon_*^2 \tag{1}$$

In this report, we show conditions under which Eq. (1) is valid for a particular error criterion. If $x(t)$ is the original waveform and $x_*(t)$ the reconstructed waveform, we define ϵ^2 to be the mean-integral-square error between $x(t)$ and $x_*(t)$, i.e.,

$$\epsilon^2 = E\left\{ \int_I [x(t) - x_*(t)]^2 dt \right\} , \tag{2}$$

where E denotes expectation and I is time interval over which $x(t)$ is defined. We assume that $x(t)$ is a sample function from a random process with a countable basis. The "samples" here are taken to be the generalized Fourier coefficients $a_1, \ldots, a_N$ of the expansion $\hat{x}(t) = \sum_{i=1}^{N} a_i \phi_i(t)$, where the $\phi_i(t)$ are an arbitrary set of orthonormal (over I) basis functions. The approximation error is then given by

$$\epsilon_a^2 = E\left\{ \int_I [x(t) - \sum^{N} a_i \phi_i(t)]^2 \, dt \right\} . \tag{3}$$

Now, if the a_i are chosen to minimize ϵ_a^2, $a_k = \int_I x(t)\phi_k(t)dt$; and for this best set of a_i,

$$\int_I [x(t) - \Sigma a_i\phi_i(t)]\,[\Sigma b_i\phi_i(t)]dt = 0 \tag{4}$$

for any choice of real numbers b_i.[1] We assume in the sequel that the approximation is done optimally. Denoting the received samples by a_i^*, the reconstructed waveform $x_*(t)$ has the form

$$x_*(t) = \sum_{i=1}^{N} a_i^*\phi_i(t)\ . \tag{5}$$

Now

$$\epsilon^2 = E\{\int_I [x(t) - x_*(t)]^2\, dt\} \tag{6}$$

$$= E\{\int_I [x(t) - \Sigma^N a_i^*\phi_i(t)]^2\, dt\} \tag{7}$$

Adding and subtracting the term $\hat{x} = \Sigma^N a_i\phi_i(t)$ inside the brackets in Eq. (7), we have

$$\epsilon^2 = E\{\int_I [x(t) - \Sigma^N a_i\phi_i(t) + \Sigma^N a_i\phi_i(t) - \Sigma^N a_i^*\phi_i(t)]^2\, dt\} \tag{8}$$

$$= E\{\int_I [x(t) - \Sigma^N a_i\phi_i(t)]^2 dt + \int_I [\Sigma a_i\phi_i(t) - \Sigma a_i^*\phi_i(t)]^2 dt$$

$$+ \int_I [x(t) - \Sigma a_i\phi_i(t)]\,[\Sigma(a_i - a_i^*)\phi(t)]dt\}\ . \tag{9}$$

However, from Eq. (4), the last term in Eq. (9) is zero. Then, using the orthonormality of the $\phi_i(t)$, we have

$$\epsilon^2 = \epsilon_a^2 + E\{\Sigma^N (a_i - a_i^*)^2\} \tag{10}$$

We denote by $\hat{a}_i$ the quantized version of a_i. Then

$$\epsilon_q^2 = E\{\int_I [\Sigma^N a_i\phi_i(t) - \Sigma^N \hat{a}_i\phi_i(t)]^2 dt$$

$$= E\{\Sigma^N (a_i - \hat{a}_i)^2\} \tag{11}$$

In the last term in Eq. (10), we add and subtract the term $\hat{a}_i$ so that

$$E\{\sum^N (a_i - a_i^*)^2\} = E\{\sum^N (a_i - \hat{a}_i + \hat{a}_i - a_i^*)^2\} \tag{12}$$

$$= E\{\sum^N (a_i - \hat{a}_i)^2\} + E\{\sum^N (\hat{a}_i - a_i)^2\} + 2E\{\sum^N (a_i - \hat{a}_i)(\hat{a}_i - a_i^*)\}. \tag{13}$$

Now, a_i^* differs from $\hat{a}_i$ only due to incorrect decisions of the receiver, so that

$$\epsilon_*^2 = E\{ \int_I [\sum^N \hat{a}_i \phi_i(t) - \sum^N a_i^* \phi_i(t)]^2 dt\} \tag{14}$$

$$= E\{\sum^N (\hat{a}_i - a_i^*)^2 \tag{15}$$

Equation (10) then may be written as

$$\epsilon^2 = \epsilon_a^2 + \epsilon_q^2 + \epsilon_*^2 + 2E\{\sum^N (a_i - \hat{a}_i)(\hat{a}_i - a_i^*)\} \tag{16}$$

Now, if the last term in Eq. (16) were zero, the total error would indeed be the sum of the approximation error, the quantization error, and the channel error.

We now concentrate on a single term

$$E\{(a_k - \hat{a}_k)(\hat{a}_k - a_k^*)\} \tag{17}$$

and, for simplicity, drop the subscripts. To compute Eq. (17), we need the joint density functions $p_1(a,\hat{a})$, $p_2(a,a^*)$, $p_3(\hat{a},a^*)$. The quantization scheme used here is the usual one where the number x_i is transmitted (or rather a waveform corresponding to x_i) whenever $z_i < a \leq z_{i+1}$, so that

$$P[\hat{a} = x_j] = P[z_i < a \leq z_{j+1}] = \int_{z_i}^{z_{j+1}} p(a)da \qquad \text{where } p(\ )$$

is the probability density function of the sample a. Note that the type of transmission, noise characteristics, and receiver structure are completely arbitrary. The effect of all these factors is described by the conditional probability matrix $[P_{ij}]$ where the element P_{ij} is given by

$$P_{ij} = \text{Prob } [a^* = x_j/\hat{a} = x_i] \tag{18}$$

and is the probability that a transmitted point x_i is interpreted at the receiver as x_j.

For ease of notation, we will denote by p_j the junction which is equal to p(a) for $z_j \leq a \leq z_{j+1}$ and is zero elsewhere, and call

$$\overline{p_j} = \int_{z_j}^{z_{j+1}} p(a)da. \tag{19}$$

We find after some straight-forward calculations that the various probability density functions are given by

$$p_1(a,\hat{a}) = \sum_i \delta(\hat{a}-x_i)p_i \tag{20}$$

$$p_2(a,a*) = \sum_i \delta(a*-x_i) \sum_j p_j P_{ji} \tag{21}$$

$$p_3(\hat{a},a*) = \sum_i \sum_j \overline{p_i} P_{ij} \delta(\hat{a}-x_i)\ \delta(a*-x_j) \tag{22}$$

$$p_4(\hat{a}) = \sum_i \overline{p_i}\ \delta(\hat{a}-x_i) \tag{23}$$

Equation (17) may be expanded (omitting subscripts) to yield

$$E[(a-\hat{a})(\hat{a}-a*)] = E[a\hat{a}] - E[\hat{a}]^2 + E[\hat{a}a*] - E[aa*] \tag{24}$$

Denoting

$\int_{z_j}^{z_{j+1}} ap(a)da$ by $\overline{ap_j}$, we find that

$$E[a\hat{a}] = \sum_i x_i \overline{ap_i} \tag{25}$$

$$E[\hat{a}]^2 = \sum x_i^2 \overline{p_i} \tag{26}$$

$$E[\hat{a}a*] = \sum_i \sum_j x_i x_j \overline{p_i} P_{ij} \tag{27}$$

$$E[aa*] = \sum_i x_i \sum_j \overline{ap_j} P_{ji} \quad . \tag{28}$$

Now it can be shown[2], that the optimum (minimizes $E[a-\hat{a}]^2$) choice of the x_i is obtained when

$$x_j = \frac{\int_{z_j}^{z_{j+1}} ap(a)da}{\int_{z_j}^{z_{j+1}} p(a)da} = \frac{\overline{ap_j}}{p_j} \tag{29}$$

Substituting $x_i\overline{p_i}$ for $\overline{ap_i}$ in Eq. (25) yields Eq. (26) so that $E[a\hat{a}] - E[\hat{a}]^2 = 0$ whenever the x_i are chosen optimally. Also, substituting Eq. (29) into Eq. (28) yields Eq. (27). Hence, $E[\hat{a}a^*] - E[aa^*] = 0$ whenever the quantization is done optimally, so that

$$\epsilon^2 = \epsilon_a{}^2 + \epsilon_q{}^2 + \epsilon_*{}^2 \qquad (30)$$

whenever both the approximation and the quantization are optimally performed. No assumption regarding the independence of the "quantization noise" and the "channel noise" is made nor required. Equation (30) neither implies this independence nor (what is more important) is implied by this assumption.

Results concerning tradeoffs between approximation, quantization, and channel error, together with waveform selection and finite-time coding, will be presented at a later date.

REFERENCES

1. N. I. Acheiser, Theory of Approximation, Ungar Publishing Co., New York, 1956.

2. J. Max, "Quantizing for Minimum Distortion," IRE Trans, Vol. IT-6, pp 7-12.

IV. PHASE LOCK LOOP STUDIES

J. C. Lindenlaub

J. J. Uhran

The overall purpose of this investigation is to study the threshold properties of a phase lock loop using various phase detectors. So far, the investigation has been restricted to the first order case. The results are being extended to higher order systems to determine if the hierarchy of improvements still results.

The second semi-annual report showed the relative improvements that could be obtained by changing the form of the phase detector for a first order phase lock loop. Still many aspects of the threshold problem remain unresolved. Therefore, as an aid to continuing the study, an experimental phase-lock loop system has been built and is being used to observe the threshold properties of the systems considered. Results from the first-order system using a sine comparator, shown schematically in Fig. 1, will be reported here. The addition of an analog divider to this system forms a tanlock receiver, Fig. 2.

Static and dynamic tests made on the components and the entire 1st order system of Fig. 1 indicate that the physical plant conforms to the theoretical model, Fig. 3, in the regions of interest.

The phase detector has a static characteristic very close to the desired cosine function, Fig. 4. The VCO has a large adjustable gain factor, β, linear over 5% of the center frequency 450kc. A typical gain curve of the VCO, combined with a large resistive divider, m, is shown in Fig. 5. The combined gain--mβ--is 2.4kc/volt.

When the system of Fig. 1 is synchronized to an unmodulated carrier, the d.c. control voltage, e_o, should be proportional to the offset frequency $\Delta\omega$ as in Eq. (1), while the

$$e_o = A \sin \phi = \frac{\Delta\omega}{m\beta} \tag{1}$$

maximum offset frequency, $\Delta \omega_{max}$, is proportional to the input signal level A as in Eq. (2). Figures (6) and (7) are typical plots

$$\Delta \omega_{max} = A \beta m \tag{2}$$

of these equations for our system. Both m and β were fixed for the above situation.

A thorough investigation of the system has been undertaken since it has been shown to conform to our model. Some of the preliminary conclusions are summarized in the following three sections. It should be noted that both the carrier and modulating signal are sinusoidal while the interference is band-limited, white, gaussian noise.

Preliminary Conclusions

A. Frequency Modulated Carrier--No Noise

(1) Provided the signal rate is within the loop bandwidth, signal breakup or distortion is caused by the frequency deviation, not the signal rate. The maximum deviation is approximately equal to the open loop gain when $\Delta \omega = 0$. If $\Delta \omega \neq 0$, the maximum deviation is proportionately reduced.

(2) The rms signal output will increase linearly with deviation until breakup occurs. At that point it will begin to decrease if deviation is increased as shown in Fig. 8. This effect will also appear at threshold in the presence of noise.

(3) The plot of the output signal vs. frequency indicates the shape to be the same as expected from a small signal linear analysis, with the 3db point occurring at the open loop gain.

B. Fixed Carrier--White Noise

(1) Within the equivalent noise bandwidth the output spectrum is parabolic for high carrier to noise ratio (C/N).

(2) The total output power varies linearly with C/N in db.

(3) Not evident in the total output power is an additional noise component with a spectral width approximately equal to the inverse of the sync time. This component appears at threshold and is due to the system randomly losing sync and causing

spikes in the output. When the noise is measured only in the information band, this component has a significant effect.

C. Frequency Modulated Carrier--White Noise

(1) The standard demodulation curves are shown in Fig. 9 for two cases of modulation index. They are similar to the standard discriminator curves. It should be noted that the threshold is due to two simultaneous effects: (a) a large increase in noise due to the sync spikes, and (b) a sharp decrease in signal power as discussed above.

A model of the phase-lock loop is being formulated to more adequately explain the threshold effects. It should be noted that for the analog signal case, the output information is practically useless below threshold unless it is extremely redundant. However, in the case of digital signals, where a probability of error criterion is used, operation might be quite useful below threshold for data rates less than the loop bandwidth.

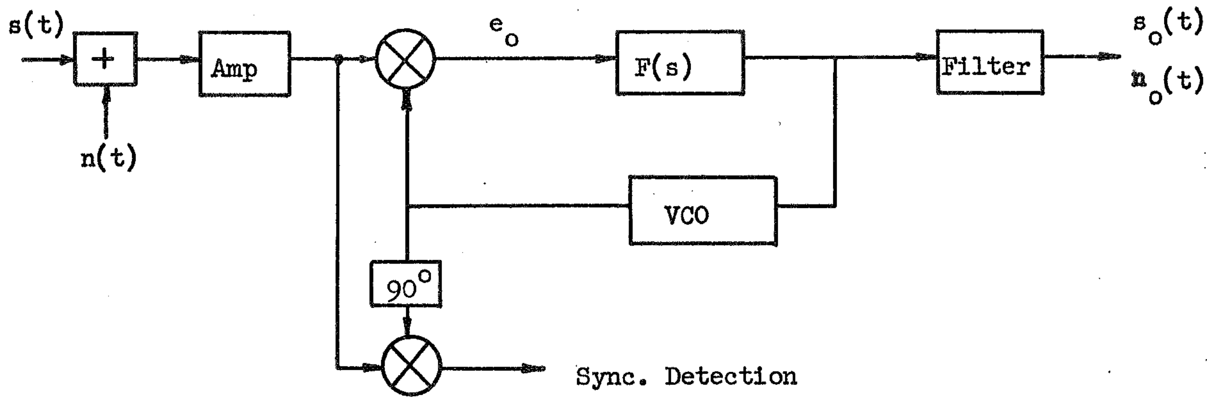


Fig. 1

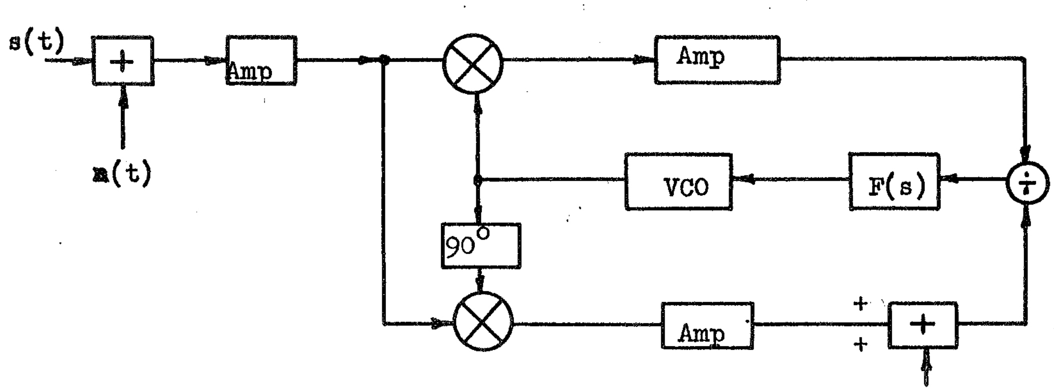


Fig. 2

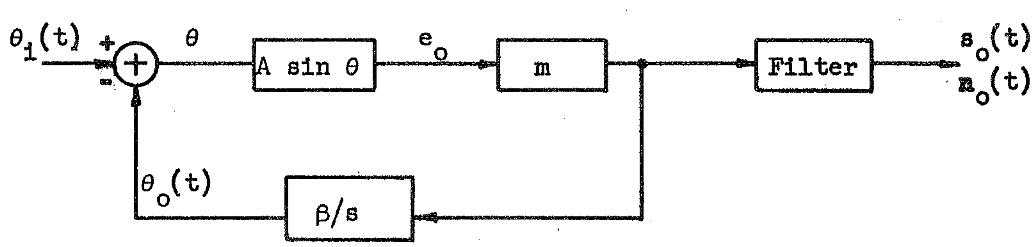
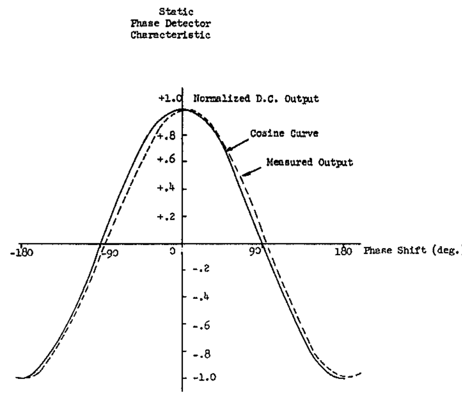


Fig. 3



Note: Shift due to a small unbalanced voltage which is cancelled out.

Fig. 4

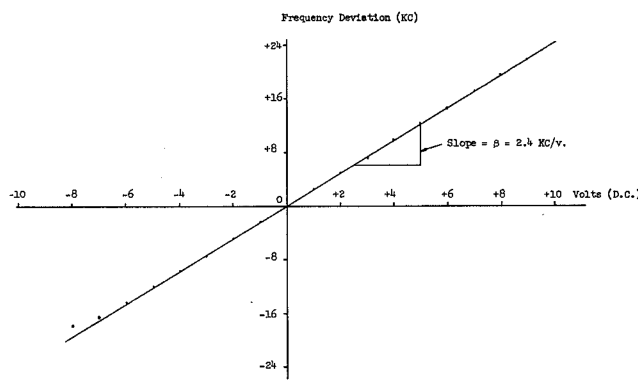


Fig. 5

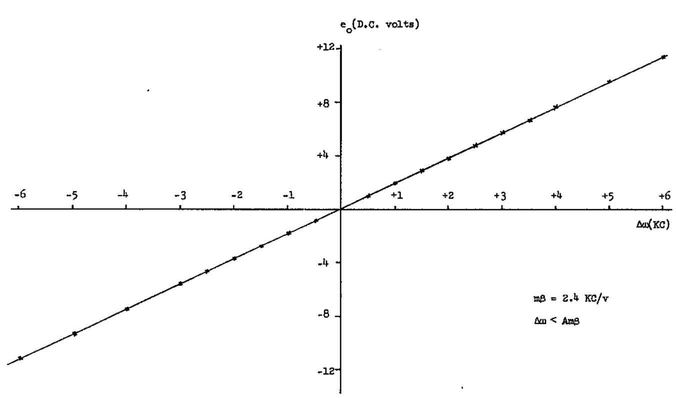


Fig. 6

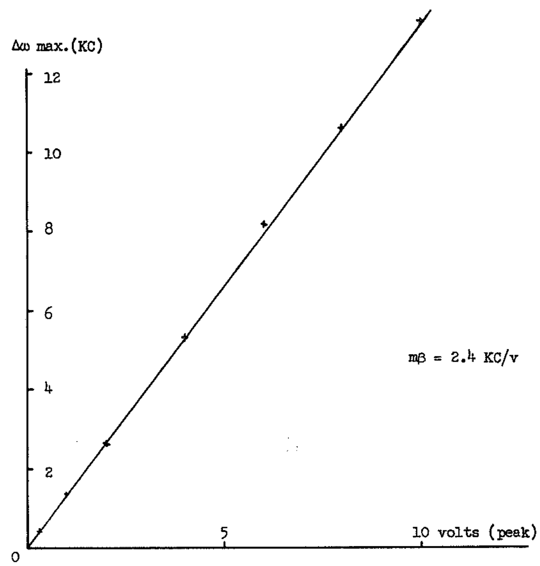


Fig. 7

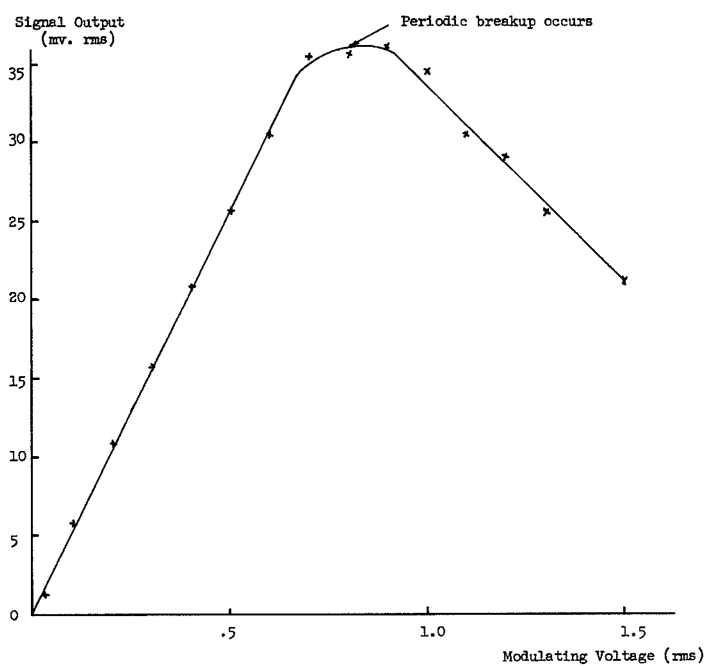


Fig. 8

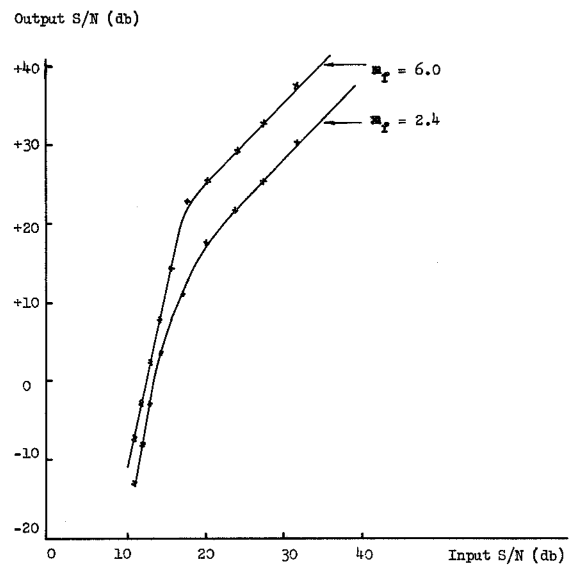


Fig. 9