

AT&T  
BELL LABORATORIES

December 1984  
Vol. 63 No. 10 Part 1

# TECHNICAL JOURNAL

A JOURNAL OF THE AT&T COMPANIES

Wire-Tap Channel

Buffer Behaviour

Impulse Response

Decision Demodulation of PCM Signals

Microwave Radio Transmission

## EDITORIAL COMMITTEE

A. A. PENZIAS,<sup>1</sup> *Committee Chairman*

M. M. BUCHNER, JR.<sup>1</sup>

R. C. FLETCHER<sup>1</sup>

J. S. NOWAK<sup>1</sup>

R. P. CLAGETT<sup>2</sup>

D. HIRSCH<sup>4</sup>

B. B. OLIVER<sup>5</sup>

R. P. CREAN<sup>2</sup>

S. HORING<sup>1</sup>

J. W. TIMKO<sup>3</sup>

B. R. DARNALL<sup>1</sup>

R. A. KELLEY<sup>1</sup>

V. A. VYSSOTSKY<sup>1</sup>

B. P. DONOHUE, III<sup>3</sup>

J. F. MARTIN<sup>2</sup>

<sup>1</sup>AT&T Bell Laboratories

<sup>2</sup>AT&T Technologies

<sup>3</sup>AT&T Information Systems

<sup>4</sup>AT&T Consumer Products

<sup>5</sup>AT&T Communications

## EDITORIAL STAFF

B. G. KING, *Editor*

L. S. GOLLER, *Assistant Editor*

P. WHEELER, *Managing Editor*

A. M. SHARTS, *Assistant Editor*

B. G. GRUBER, *Circulation*

AT&T BELL LABORATORIES TECHNICAL JOURNAL (ISSN0005-8580) is published ten times each year by AT&T, 550 Madison Avenue, New York, NY 10022; C. L. Brown, Chairman of the Board; T. O. Davis, Secretary. The Computing Science and Systems section and the special issues are included as they become available. Subscriptions: United States—1 year \$35; 2 years \$63; 3 years \$84; foreign—1 year \$45; 2 years \$73; 3 years \$94. Payment for foreign subscriptions or single copies must be made in United States funds, or by check drawn on a United States bank and made payable to the Technical Journal and sent to AT&T Bell Laboratories, Circulation Dept., Room 1E335, 101 J. F. Kennedy Pky, Short Hills, NJ 07078.

Single copies of material from this issue of the Journal may be reproduced for personal, noncommercial use. Permission to make multiple copies must be obtained from the Editor.

Comments on the technical content of any article or brief are welcome. These and other editorial inquiries should be addressed to the Editor, AT&T Bell Laboratories Technical Journal, Room 1H321, 101 J. F. Kennedy Pky, Short Hills, NJ 07078. Comments and inquiries, whether or not published, shall not be regarded as confidential or otherwise restricted in use and will become the property of AT&T. Comments selected for publication may be edited for brevity, subject to author approval.

Printed in U.S.A. Second-class postage paid at Short Hills, NJ 07078 and additional mailing offices. Postmaster: Send address changes to the AT&T Bell Laboratories Technical Journal, Room 1E335, 101 J. F. Kennedy Pky, Short Hills, NJ 07078.

Copyright © 1984 AT&T.

AT&T Bell Laboratories

# Technical Journal

---

VOL. 63

DECEMBER 1984

NO. 10, PART 1

---

Copyright © 1984 AT&T. Printed in U.S.A.

<b>Wire-Tap Channel II</b>	2135
L. H. Ozarow and A. D. Wyner	
<b>Stationary and Cyclostationary Finite Buffer Behaviour Computation via Levinson's Method</b>	2159
M. H. Ackroyd	
<b>Distortion Analysis on Measuring the Impulse Response of a System Using a Crosscorrelation Method</b>	2171
N. Benvenuto	
<b>Soft Decision Demodulation to Reduce the Effect of Transmission Errors in Logarithmic PCM Transmitted Over Rayleigh Fading Channels</b>	2193
W. C. Wong, R. Steele, and C.-E. Sundberg	
<b>Linear Equalization Theory in Digital Data Transmission Over Dually Polarized Fading Radio Channels</b>	2215
N. Amitay and J. Salz	
PAPERS BY AT&T BELL LABORATORIES AUTHORS	2261
CONTENTS, JANUARY ISSUE	2271



## Wire-Tap Channel II

By L. H. OZAROW and A. D. WYNER\*

(Manuscript received April 2, 1984)

Consider the following situation.  $K$  data bits are to be encoded into  $N > K$  bits and transmitted over a noiseless channel. An intruder can observe a subset of his choice of size  $\mu < N$ . The encoder is to be designed to maximize the intruder's uncertainty about the data given his  $\mu$  intercepted channel bits, subject to the condition that the intended receiver can recover the  $K$  data bits perfectly from the  $N$  channel bits. The optimal trade-offs among the parameters  $K$ ,  $N$ , and  $\mu$  and the intruder's uncertainty  $H$  ( $H$  is the "conditional entropy" of the data given the  $\mu$  intercepted channel bits) were found. In particular, it was shown that for  $\mu = N - K$ , a system exists with  $H \approx K - 1$ . Thus, for example, when  $N = 2K$  and  $\mu = K$ , it is possible to encode the  $K$  data bits into  $2K$  channel bits, so that by looking at any  $K$  channel bits, the intruder obtains no more than one bit of the data.

### I. INTRODUCTION

In this paper we study a communication system in which an unauthorized intruder is able to intercept a subset of the transmitted symbols, and it is desired to maximize the intruder's uncertainty about the data without the use of an encryption key (either public or private).

Specifically, the encoder associates with the  $K$ -bit binary data sequence  $\mathbf{S}^K$  an  $N$ -bit binary "transmitted" sequence  $\mathbf{X}^N$ , where  $N > K$ . It is required that a decoder can correctly obtain  $\mathbf{S}^K$  with high probability by examining  $\mathbf{X}^N$ . The intruder can examine a subset of his choice of size  $\mu$  of the  $N$  positions in  $\mathbf{X}^N$ , and the system designer's task is to make the intruder's equivocation (uncertainty) about the

---

\* Authors are employees of AT&T Bell Laboratories.

---

Copyright © 1984 AT&T. Photo reproduction for noncommercial use is permitted without payment of royalty provided that each reproduction is done without alteration and that the Journal reference and copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free by computer-based and other information-service systems without further permission. Permission to reproduce or republish any other portion of this paper must be obtained from the Editor.

data as large as possible. The encoder is allowed to introduce randomness into the transformation  $\mathbf{S}^K \rightarrow \mathbf{X}^N$ , but we make the assumption that the decoder and the intruder must share any information about the encoding and the randomness. This assumption precludes the use of “key” cryptography, where the decoder has the exclusive possession of certain information.

As an example, suppose that  $K = 1$ ,  $N = 2$ , and  $\mu = 1$ . Let the data bit be  $S$ , and let  $\xi$  be a uniform binary random variable that is independent of  $S$ . Let  $\mathbf{X}^2 = (\xi, \xi \oplus S)$ , where  $\oplus$  denotes modulo 2 addition. If the intruder looks at either coordinate of  $\mathbf{X}^2$ , he gains no information about  $S$ , so that the system has perfect secrecy. The decoder, however, can obtain  $S$  by adding (modulo 2) the two components of  $\mathbf{X}^2$ .

Our problem is to replicate this type of performance with large  $K$ ,  $N$ , and  $\mu$ . In fact, we assume that  $K \approx RN$  and  $\mu \approx \alpha N$ , where  $R$  and  $\alpha$  are held fixed and  $N$  becomes large. Roughly speaking, we show that perfect secrecy is attainable provided that  $\mu$  is not too large, specifically  $\mu \leq N - K$  or  $\alpha \leq 1 - R$ . In Section II we give a precise statement and discussion of our problem and results, leaving the proofs for Sections III through V.

This problem is similar to the wire-tap channel problem studied in Ref. 1. A special case of the problem studied there allows an intruder to examine a subset of the encoder symbols that is chosen at random by nature. In the present problem, the system designer must make the system secure against a more powerful intruder who can select which subset to examine.

## II. FORMAL STATEMENT OF THE PROBLEM AND RESULTS

In this section we give a precise statement of our problem and state the results.

First a word about notation. Let  $\mathcal{U}$  be an arbitrary finite set. Denote its cardinality by  $|\mathcal{U}|$ . Consider  $\mathcal{U}^N$ , the set of  $N$ -vectors with components in  $\mathcal{U}$ . The members of  $\mathcal{U}^N$  will be written as

$$\mathbf{u}^N = (u_1, u_2, \dots, u_N),$$

where subscripted letters denote components and boldface superscripted letters denote vectors. A similar convention applies to random vectors, which are denoted by uppercase letters. When the dimension of a vector is clear from the context, we omit the superscript. Finally, for random variables  $X$ ,  $Y$ ,  $Z$ , etc., the notation  $H(X)$ ,  $H(X|Y)$ ,  $I(X; Y)$ , etc., denotes the standard information theoretic quantities as defined, for example, in Ref. 2.

We now turn to the description of the communication system.

1. The *source* output is a sequence  $\{S_k\}_1^\infty$ , where the  $S_k$  are independent and identically distributed (i.i.d.) binary random variables with uniform distribution.

2. The *encoder* with parameters  $(K, N)$  is a channel with input alphabet  $\{0, 1\}^K$ , output alphabet  $\{0, 1\}^N$ , and transition probability  $q_E(\mathbf{x}^N | \mathbf{s}^K)$ . Let  $\mathbf{S}^K$  and  $\mathbf{X}^N$  be the input and output, respectively, of the encoder.

3. The *decoder* is a mapping

$$f_D: \{0, 1\}^N \rightarrow \{0, 1\}^K.$$

Let  $\hat{\mathbf{S}} = (\hat{S}_1, \hat{S}_2, \dots, \hat{S}_K) = f_D(\mathbf{X}^N)$ . The *error rate* is

$$P_e = \frac{1}{K} \sum_{k=1}^N \Pr\{S_k \neq \hat{S}_k\}.$$

4. An intruder with parameter  $\mu \leq N$  picks a subset  $\mathcal{S} \subseteq \{1, 2, \dots, N\}$ , such that  $|\mathcal{S}| = \mu$  and is allowed to observe  $X_n, n \in \mathcal{S}$ . Let  $\mathbf{Z}^N = (Z_1, \dots, Z_N)$ , defined by

$$Z_n = \begin{cases} X_n, & n \in \mathcal{S}, \\ ?, & n \notin \mathcal{S}, \end{cases}$$

denote the intruder's information. The system designer seeks to maximize the equivocation

$$\Delta \triangleq \min_{\mathcal{S}: |\mathcal{S}|=\mu} H(\mathbf{S}^K | \mathbf{Z}^N).$$

Thus, the designer is assured that no matter what subset  $\mathcal{S}$  the intruder chooses, the intruder's remaining uncertainty about the source vector is at least  $\Delta$ . When  $\Delta = K$ , the intruder obtains no information about the source, and the system has attained perfect secrecy.

In this paper we study the trade-offs between  $K, N, \Delta$ , and  $P_e$ . As we shall see, it will be useful to consider the normalized qualities  $K/N, \mu/N$ , and  $\Delta/K$ . Thus,  $K/N$  is the rate of the encoder equal to the number of data bits per encoded bit,  $\mu/N$  is the fraction of the encoded bits that the intruder is able to observe, and  $\Delta/K$  is the normalized entropy.

The intruder who observes  $\mathbf{Z}^N$  can reconstruct the data sequence  $\mathbf{S}^K$  with a per-bit-error probability of, say,  $P'_e$ . It follows from Fano's inequality that  $h(P'_e) \geq \Delta/K$ , where  $h(\cdot)$  is the binary entropy function defined below eq. (2). Thus,  $\Delta/K \approx 1$  implies that  $P'_e \approx 1/2$ , which is essentially perfect secrecy.

We will say that the triple  $(R, \alpha, \delta)$  is *achievable* if for all  $\epsilon > 0$  and all integers  $N_0 > 0$ , there exists an encoder/decoder with parameters  $N \geq N_0, K \geq (R - \epsilon)N, \mu \geq (\alpha - \epsilon)N, \Delta \geq (\delta - \epsilon)N$ , and  $P_e \leq \epsilon$ . We

will show in the sequel that  $(R, \alpha, \delta)$  is achievable for  $0 \leq R, \alpha \leq 1$ , and

$$0 \leq \delta \leq \begin{cases} 1, & 0 \leq \alpha \leq 1 - R, \\ \frac{(1 - \alpha)}{R}, & 1 - R \leq \alpha \leq 1. \end{cases} \quad (1)$$

A graph of the achievable  $(\alpha, \delta)$  pairs for fixed  $R$  is given in Fig. 1.

The following theorem, a proof of which is given in Section III, is a "converse" result that gives a necessary condition on achievable codes.

*Theorem 1: If there exists a code with parameters  $(K, N, \Delta, P_e)$ , then*

$$\Delta \leq \begin{cases} K, & 0 \leq \mu \leq N - K, \\ N - \mu + Kh(P_e), & N - K \leq \mu \leq N, \end{cases} \quad (2)$$

where  $h(\lambda) = -\lambda \log \lambda - (1 - \lambda) \log(1 - \lambda)$  is the binary entropy function.

Now if  $(R, \alpha, \delta)$  is achievable, for arbitrary  $\epsilon > 0$ , there must be an encoder/decoder with parameters  $N, K \geq (R - \epsilon)N, \mu \geq (\alpha - \epsilon)N, \Delta \geq (\delta - \epsilon)N, P_e \leq \epsilon$ . Applying Theorem 1 to this code yields

$$\delta \leq \begin{cases} 1, \\ \frac{(1 - \alpha)}{R} + O(\epsilon) + h(\epsilon), & 1 - R \leq \alpha + O(\epsilon) \leq 1, \end{cases},$$

which is (1) as  $\epsilon \rightarrow 0$ . Thus conditions (1) are necessary for a triple to be achievable. Theorem 2, which is also proved in Section III, implies that  $(R, \alpha, \delta)$  is achievable if (1) is satisfied.

*Theorem 2: Let  $1 - R < \alpha < 1$ . Then, for all  $\epsilon > 0, N_0 \geq 1$ , there exist an  $N \geq N_0$  and an encoder/decoder with parameters  $K = RN, \mu = \alpha N, \Delta/K \geq [(1 - \alpha)/R] - \epsilon$ , and  $P_e = 0$ .*

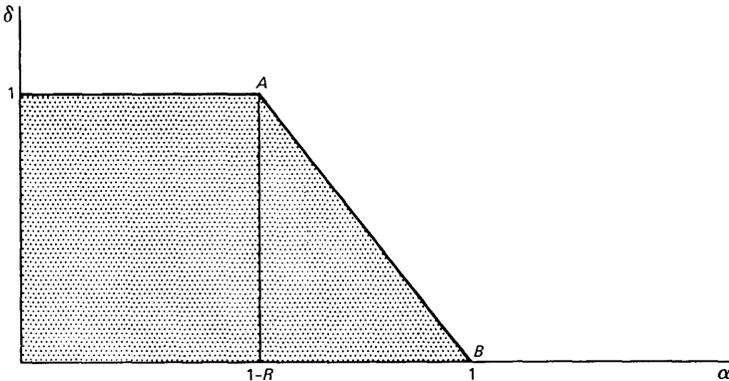


Fig. 1—Achievable  $(\alpha, \delta)$  for fixed  $R$ .

The idea behind the proof of Theorem 2 is the following. Partition the set  $\{0, 1\}^N$  into  $2^K$  subsets  $\{A_m\}_i^{2^K}$  with equal cardinality—i.e.,  $|A_m| = 2^{N-K}$ . The  $2^K$  possible values of  $\mathbf{S}^K$  can be put in 1 to 1 correspondence with these subsets. When  $\mathbf{S}^K$  corresponds to  $A_m$  ( $1 \leq m \leq 2^K$ ), the encoder output is uniformly distributed on  $A_m$ . Since the  $\{A_m\}$  are disjoint, the decoder can recover  $\mathbf{S}^K$  perfectly and  $P_e = 0$ . We show (by random coding) that there exists a partition satisfying Theorem 2.

A convenient way to partition  $\{0, 1\}^N$  is to let the sets  $\{A_m\}$  be the cosets of a group code  $G$  with  $N - K$  information symbols (so that  $G$  has  $2^K$  cosets). Theorem 3, which is proved in Section IV, asserts that, in fact, we can do quite well with codes of this type.

*Theorem 3: If the triple  $(R, \alpha, \delta)$  satisfies (1), then it is achievable using an encoder/decoder derived from a group code.*

The following simple lemma allows us to establish the achievability of all triples on the straight line of Fig. 1 connecting points  $A$  and  $B$  by proving only the achievability of point  $A$ .

*Lemma 1: Suppose that we are given an encoder/decoder  $f_E, f_D$  with parameters  $N, K$ , and  $P_e$ . Suppose that there are two intruders that have parameters  $\mu = \mu_1, \mu_2$  and  $\Delta = \Delta_1, \Delta_2$ , respectively. Then, if  $\mu_2 \geq \mu_1$ ,*

$$\Delta_2 \geq \Delta_1 - (\mu_2 - \mu_1). \quad (3)$$

*Remark:* Inequality (3) can be rewritten as

$$(\Delta_2/K) \geq (\Delta_1/K) - \left( \frac{\mu_2/N - \mu_1/N}{K/N} \right),$$

from which we conclude that  $(R, \alpha_1, \delta_1)$  is achievable implies that  $(R, \alpha_2, \delta_2)$  is achievable where  $\alpha_2 \geq \alpha_1$  and

$$\delta_2 = \delta_1 - \left( \frac{\alpha_2 - \alpha_1}{R} \right).$$

In particular, if  $\alpha_1 = 1 - R, \delta_1 = 1$ , then

$$\delta_2 = (1 - \alpha_2)/R.$$

*Proof of Lemma 1:* Let  $\mathcal{S}_1 \subseteq \mathcal{S}_2 \subseteq \{1, 2, \dots, N\}$ , where  $|\mathcal{S}_1| = \mu_1, |\mathcal{S}_2| = \mu_2$ . Let  $\mathbf{Z}_i^N$  correspond to  $\mathcal{S}_i (i = 1, 2)$ , i.e.,  $\mathbf{Z}_i = (Z_{i1}, \dots, Z_{iN})$  and

$$Z_{ij} = \begin{cases} X_j, & j \in \mathcal{S}_i \\ ?, & j \notin \mathcal{S}_i \end{cases}$$

Then,

$$\begin{aligned} H(\mathbf{S}^K | \mathbf{Z}_2) - H(\mathbf{S}^K | \mathbf{Z}_1) &= H(\mathbf{S}^K | \mathbf{Z}_2, \mathbf{Z}_1) - H(\mathbf{S}^K | \mathbf{Z}_1) \\ &= -I(\mathbf{S}^K; \mathbf{Z}_2 | \mathbf{Z}_1) \geq -H(\mathbf{Z}_2 | \mathbf{Z}_1) \geq -(\mu_2 - \mu_1), \end{aligned}$$

where the first equality follows from  $\mathcal{S}_1 \subseteq \mathcal{S}_2$ . Thus

$$\begin{aligned} H(\mathbf{S}^K | \mathbf{Z}_2) &\geq H(\mathbf{S}^K | \mathbf{Z}_1) - (\mu_2 - \mu_1) \\ &\geq \Delta_1 - (\mu_2 - \mu_1) \end{aligned} \quad (4)$$

from the definition of  $\Delta$ . Minimizing (4) over all  $\mathcal{S}_2$ , with  $|\mathcal{S}_2| = \mu_2$ , yields (3) and the lemma.  $\square$

Finally, we state a theorem that is a rather surprising strengthening of Theorem 2. Its proof is given in Section V.

*Theorem 4: For arbitrary  $K, N$  ( $1 \leq K \leq N$ ), and  $\mu = N - K$ , there exists an encoder/decoder with  $P_e = 0$  and*

$$\Delta \geq K - 1 - \frac{2.23}{\sqrt[4]{N}}.$$

### III. PROOF OF THEOREMS 1 AND 2

Assume that  $\mathbf{S}^K, \mathbf{X}^N, \mathbf{Z}^N$ , and  $\hat{\mathbf{S}}$  correspond to a source/encoder/decoder as defined in Section II, with parameters  $K, N, \Delta$ , and  $P_e$ . Then, making repeated use of the identity  $H(U, V) = H(U) + H(V|U)$ , we obtain

$$\begin{aligned} \Delta &= H(\mathbf{S}^K | \mathbf{Z}^N) = H(\mathbf{S}, \mathbf{Z}) - H(\mathbf{Z}) \\ &= H(\mathbf{S}, \mathbf{X}, \mathbf{Z}) - H(\mathbf{X} | \mathbf{S}, \mathbf{Z}) - H(\mathbf{Z}) \\ &= H(\mathbf{S} | \mathbf{X}, \mathbf{Z}) + H(\mathbf{X}, \mathbf{Z}) - H(\mathbf{X} | \mathbf{S}, \mathbf{Z}) - H(\mathbf{Z}) \\ &= H(\mathbf{S} | \mathbf{X}, \mathbf{Z}) + H(\mathbf{X} | \mathbf{Z}) - H(\mathbf{X} | \mathbf{S}, \mathbf{Z}). \end{aligned} \quad (5)$$

Now

$$\begin{aligned} H(\mathbf{S} | \mathbf{X}, \mathbf{Z}) &= H(\mathbf{S} | \mathbf{X}, \mathbf{Z}, \hat{\mathbf{S}}) \leq H(\mathbf{S} | \hat{\mathbf{S}}) \\ &\leq Kh(P_e), \end{aligned}$$

where the last inequality follows from Fano's inequality (see Ref. 2). Also, since  $H(\mathbf{X} | \mathbf{Z})$  is the entropy of those  $N - \mu$  coordinates of  $\mathbf{X}$  not specified by  $\mathbf{Z}$ , we have  $H(\mathbf{X} | \mathbf{Z}) \leq N - \mu$ . Finally, noting that  $H(\mathbf{X} | \mathbf{S}, \mathbf{Z}) \geq 0$ , we have from (5)

$$\Delta \leq N - \mu + Kh(P_e),$$

which is Theorem 1.

We now give a proof of Theorem 2, which proceeds along the lines suggested in Section II. Let  $K, N$  be given, and let  $\{A_m\}$ ,  $1 \leq m \leq 2^K$ , be a partition of  $\{0, 1\}^N$  into subsets  $A_m \subseteq \{0, 1\}^N$  such that  $|A_m| \equiv 2^{N-K}$ . As in Section II, the partition defines a code: to encode message  $m$  ( $1 \leq m \leq 2^K$ ), we let  $\mathbf{X}^N$  be a randomly chosen vector in  $A_m$ . Since the  $A_m$  are disjoint,  $P_e = 0$  and  $H(\mathbf{S} | \mathbf{X}, \mathbf{Z}) = 0$ . Further, since the  $2^K$  messages are equally likely and  $|A_m| \equiv 2^{N-K}$ ,  $\mathbf{X}$  is uniformly distributed on  $\{0, 1\}^N$ , so that its coordinates are i.i.d. uniform binary random

variables. Thus  $H(\mathbf{X}^N | \mathbf{Z}^N) = N - \mu$ . We conclude from (5) that for this encoder

$$\Delta = N - \mu - H(\mathbf{X}^N | \mathbf{S}^K, \mathbf{Z}^N). \quad (6)$$

Now let  $\mathbf{z} \in \{0, 1, ?\}^N$  be a possible value for the intruder's information, and let  $\mathbf{x} \in \{0, 1\}^N$ . We say that  $\mathbf{z}$  is "consistent" with  $\mathbf{x}$  if  $\mathbf{z}$  can be obtained from  $\mathbf{x}$  by changing a subset of the coordinates of  $\mathbf{x}$  to '?'s. Next, let  $L \geq 1$  be an integer to be chosen later. We say that a partition  $\{A_m\}$  is "good" if for all  $m (1 \leq m \leq 2^K)$  and all  $\mathbf{z} \in \{0, 1, ?\}^N$  with exactly  $N - \mu$  '?'s,

$$\text{card}\{\mathbf{x} \in A_m: \mathbf{z} \text{ is consistent with } \mathbf{x}\} < L.$$

If our encoder corresponds to a good partition for some  $L$ , then

$$H(\mathbf{X}^N | \mathbf{S}^K, \mathbf{Z}^N) < \log L,$$

and (6) yields

$$\Delta \geq N - \mu - \log L. \quad (7)$$

At the end of this section we will prove the following proposition about the existence of good partitions. This will lead us directly to Theorem 2.

*Lemma 2: Let  $K, N$ , and  $\mu$  be such that*

$$N - \mu - K < 0. \quad (8)$$

*Then, there exists a good partition (with parameters  $K, N$ , and  $\mu$ ) provided*

$$L > \frac{2N + K + 2 \log e}{K + \mu - N}. \quad (9)$$

Now let  $R, \alpha, \epsilon$ , and  $N_0$  be given as in the hypothesis of Theorem 2. Then, using  $2 \log e \leq 3$ , we write for  $N \geq 1$ ,

$$\frac{N + K + 2 \log e}{K + \mu - N} \leq \frac{1 + R + 3}{\alpha - (1 - R)} \triangleq B < \infty.$$

Thus, there exists a good partition with  $L \leq B + 1$ , and we conclude from (7) that there exists a code with  $\Delta/K \geq [(1 - \alpha)/R] - \{[\log(B + 1)]/(RN)\}$ . If we choose  $N \geq N_0, \epsilon R/\log(B + 1)$ , then the existence of this code establishes Theorem 2. It remains to prove Lemma 2.

*Proof of Lemma 2:* Let  $\{A_m\}$ ,  $1 \leq m \leq 2^K$ , be a partition of  $\{0, 1\}^N$ , where  $|A_m| \equiv 2^{N-K}$ . Let  $\Psi(A_1, \dots, A_{2^K}) = 0$  or 1, depending on whether  $\{A_m\}$  is good. We write

$$\Psi(A_1, \dots, A_{2^K}) \leq \sum_{m=1}^{2^K} \sum_{\mathbf{z}} \phi(A_m, \mathbf{z}), \quad (10)$$

where the inner sum is taken over all  $\mathbf{z} \in \{0, 1, \dots\}^N$  with exactly  $N - \mu$  1's, and  $\phi(A_m, \mathbf{z}) = 1$  if

$$\text{card}\{\mathbf{x} \in A_m: \mathbf{z} \text{ is consistent with } \mathbf{x}\} \geq L,$$

and  $\phi(A_m, \mathbf{z}) = 0$  otherwise.

We now choose the partition at random with uniform distribution on the set of partitions of  $\{0, 1\}^N$  into  $2^K$  classes of equal size. The expectation  $E\Psi$  satisfies

$$E\Psi \leq \sum_m \sum_{\mathbf{z}} E\Phi(A_m, \mathbf{z}). \tag{11}$$

The expectation in the right member of (11) is taken, as indicated, with  $\mathbf{z}$  held fixed. Let us define the following quantities:

$$\begin{aligned} Q(\mathbf{z}) &= \{\mathbf{x} \subseteq \{0, 1\}^N: \mathbf{x} \text{ is consistent with } \mathbf{z}\}, \\ n_1 &= |Q(\mathbf{z})| = 2^{N-\mu}, \\ n &= |\{0, 1\}^N| = 2^N, \\ r &= |A_m| = 2^{N-K}. \end{aligned} \tag{12}$$

We now compute  $E\Phi(A_m, \mathbf{z})$ . The  $r$  members of  $A_m$  are chosen at random from  $\{0, 1\}^N$  (without replacement). The probability that exactly  $t$  members of  $A_m$  belong to  $Q(\mathbf{z})$  is

$$\frac{\binom{n_1}{t} \binom{n - n_1}{r - t}}{\binom{n}{r}} \triangleq \pi_t.$$

To see this, observe that there are  $\binom{n}{r}$  ways to choose the set  $A_m$ . The  $t$  members of  $A_m$  that belong to  $Q(\mathbf{z})$  can be chosen in  $\binom{n_1}{t}$  ways, and the remaining  $(r - t)$  members of  $A_m$  can be chosen from the complement of  $Q(\mathbf{z})$  in  $\binom{n - n_1}{r - t}$  ways.

Now

$$\pi_t = \frac{\binom{n_1}{t} \binom{n - n_1}{r - t}}{\binom{n}{r}} \leq \frac{\binom{n_1}{t} \binom{n}{r - t}}{\binom{n}{r}}.$$

Also, using  $\binom{n_1}{t} \leq n_1^t/t!$ , and

$$\begin{aligned}
\frac{\binom{n}{r-t}}{\binom{n}{r}} &= \frac{n!}{(n-r+t)!(r-t)!} \cdot \frac{r!(n-r)!}{n!} \\
&= \frac{r(r-1)(r-2)\dots(r-t+1)}{(n+t-r)(n+t-r-1)\dots(n-r+1)} \\
&\leq \frac{r^t}{(n-r)^t} = \frac{(r/n)^t}{(1-r/n)^t},
\end{aligned}$$

we have

$$\pi_t \leq \frac{(n_1 r/n)^t}{t!(1-r/n)^t}.$$

Thus

$$E\Phi(A_m, \mathbf{z}) = \sum_{t=L}^{2^{N-K}} \pi_t \leq \sum_{t=L}^{\infty} \frac{(n_1 r/n)^t}{t!(1-r/n)^t}.$$

Using (12), we have  $(n_1 r/n) = 2^{N-\mu-K}$ ,  $(1-r/n) \geq 1/2$ , so that

$$\begin{aligned}
E\Phi(A_m, \mathbf{z}) &\leq \sum_{t=L}^{\infty} 2^{(N-\mu-K)t} \frac{2^t}{t!} \\
&\leq 2^{(N-\mu-K)L} \sum_{t=0}^{\infty} \frac{2^t}{t!} = 2^{(N-\mu-K)L} e^2.
\end{aligned}$$

Substituting into (11) yields

$$\begin{aligned}
E\Psi &\leq \sum_m \sum_{\mathbf{z}} 2^{(N-\mu-K)L+2\log e} \\
&\leq 2^{(N-\mu-K)L+2\log e+K+2N}.
\end{aligned}$$

If  $L$  satisfies (9), then  $E\Psi < 1$ . Since  $\Psi$  is integer valued, there must exist a particular partition, say  $\{A_m^*\}$  such that  $\Psi(A_1^*, \dots, A_{2^K}^*) = 0$ . This is our good partition.  $\square$

#### IV. GROUP CODES AND THEOREM 3

In Sections II and III, we discussed how to construct encoder/decoders based on a partition  $\{A_m\}$  of  $\{0, 1\}^N$ . In this section we consider the special case where the partition  $\{A_m\}$  is defined by a group code and its cosets.

Let  $H$  be a  $K \times N$  parity-check matrix, which we assume has rank  $K$ . Let the partition  $\{A_m\}$ ,  $1 \leq m \leq 2^K$ , be the code defined by  $H$  and its cosets. Thus  $|A_m| \equiv 2^{N-K}$ , for  $1 \leq m \leq 2^K$ . To encode message  $\mathbf{s} = (s_1, \dots, s_K)$ , the encoder makes a random selection of one of the  $2^{N-K}$  members of the  $A_m$  corresponding to  $\mathbf{s}$ . This is equivalent to letting  $\mathbf{X}^N$  be a random choice from the  $2^{N-K}$  solutions of

$$HX^\dagger = \mathbf{s}^\dagger, \quad (13)$$

where  $\dagger$  denotes matrix transpose. Note that since  $\mathbf{S}$  is uniformly distributed on  $\{0, 1\}^K$ ,  $\mathbf{X}^N$  is uniformly distributed on  $\{0, 1\}^N$ , and its coordinates  $X_1, X_2, \dots, X_N$  are i.i.d. uniform binary random variables.

The decoder observes  $\mathbf{X}^N$  and computes  $H\mathbf{X}^\dagger$ , which is the message. Thus  $P_e = 0$ . We now show how to compute  $\Delta$  in terms of certain distance-like properties of the parity-check matrix.

*Definition:* Let  $\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_N$  be the columns of  $H$  ( $\mathbf{C}_n$  is a  $K$ -vector). Let  $\mathcal{S} \subseteq \{1, 2, \dots, N\}$  and define  $D(\mathcal{S})$  to be the dimension of the subspace spanned by  $\{\mathbf{C}_n\}$ ,  $n \in \mathcal{S}$ . For a given  $K \times N$  parity-check matrix  $H$ , define for  $0 \leq \mu \leq N$ ,

$$D^*(\mu) = \min_{|\mathcal{S}|=N-\mu} D(\mathcal{S}). \quad (14)$$

We now state Lemma 3.

*Lemma 3:* Let  $D^*(\mu)$  correspond to the  $K \times N$  parity-check matrix  $H$ . Let  $w, w'$  be the minimum weight of the code and dual code, respectively, defined by  $H$ . Then, (a) for  $N - w + 1 \leq \mu \leq N$ ,  $D^*(\mu) = N - \mu$ ; (b) for  $0 \leq \mu \leq w' - 1$ ,  $D^*(\mu) = K$ .

*Proof of Lemma 3:* Assertion (a) follows immediately on observing that all sets of  $w - 1$  columns of  $H$  are linearly independent. Thus  $D(\mathcal{S}) = |\mathcal{S}|$ , for  $|\mathcal{S}| \leq w - 1$ . If  $N - w + 1 \leq \mu \leq N$ , then  $N - \mu \leq w - 1$ , so that

$$D^*(\mu) = \min_{|\mathcal{S}|=N-\mu} D(\mathcal{S}) = N - \mu,$$

which is assertion (a).

Now assertion (b) states that all submatrices  $\hat{H} = (\mathbf{C}_{i_1}, \mathbf{C}_{i_2}, \dots, \mathbf{C}_{i_q})$  of  $H$  have rank  $K$  when  $q \geq N - w' + 1$ . To establish this assertion, assume that  $\text{rank } \hat{H} < K$ . Then there exists a set of linear row manipulations which transform  $\hat{H}$  into a matrix with a row of zeros. These same row manipulations will transform  $H$  into a matrix for which a row has weight  $\leq N - q$ . Since the dual code is the row space of  $H$ ,  $N - q \geq w'$  or  $q \leq N - w'$ , establishing assertion (b).  $\square$

We now give Lemma 4.

*Lemma 4:* When an encoder/decoder is constructed to correspond to the parity-check matrix  $H$ , then

$$\Delta = D^*(\mu). \quad (15)$$

*Proof of Lemma 4:* Let  $\mathbf{S}, \mathbf{X}, \mathbf{Z}$  correspond to an encoder/decoder with parameters  $K, N, \Delta$  and ( $P_e = 0$ ), derived, as discussed above, from a parity-check matrix  $H = (\mathbf{C}_1, \dots, \mathbf{C}_N)$ . Since  $P_e = 0$  and  $\mathbf{X}^N$  is uniformly distributed on  $\{0, 1\}^N$ , eq. (6) applies. Thus, Lemma 4 will be established when we show that

$$H(\mathbf{X}^N | \mathbf{S}^K, \mathbf{Z}^N) = N - \mu - D^*(\mu). \quad (16)$$

Now suppose that  $\mathbf{S}^K = \mathbf{s}$  and  $\mathbf{Z}^N = \mathbf{z}$ . Without loss of generality, assume that the last  $\mu$  coordinates of  $\mathbf{z}$  are copies of the corresponding coordinates of  $\mathbf{X}$ . Thus, given  $\mathbf{S}^K = \mathbf{s}$ ,  $\mathbf{Z}^N = \mathbf{z}$ , the remaining unknown coordinates of  $\mathbf{X}$  are precisely the solutions for  $x_1, \dots, x_{N-\mu}$  of

$$\sum_{n=1}^{N-\mu} \mathbf{C}_n x_n = \mathbf{s}^\dagger + \sum_{n=N-\mu+1}^N \mathbf{C}_n \mathbf{x}_n \triangleq \alpha. \quad (17)$$

Since the number of solutions is  $N - \mu - \text{rank}(\mathbf{C}_1, \dots, \mathbf{C}_{N-\mu})$ , and, given  $\mathbf{S} = \mathbf{s}$ ,  $\mathbf{Z} = \mathbf{z}$ , all these solutions are equally likely, then (16) follows; hence the lemma.  $\square$

Before continuing with the proof of Theorem 3, we digress to apply Lemma 4 in an example. Let  $K = 4$ ,  $N = 8$ , and construct an encoder/decoder using the self-dual Hamming code with block length 8 and four information digits and four check digits. Then,  $w = w' = 4$ , so that

$$\Delta = D^*(\mu) = \begin{cases} 4 = K, & 0 \leq \mu \leq 3, \\ 3, & \mu = 4, \\ N - \mu, & 5 \leq \mu \leq 8. \end{cases}$$

Thus, the encoder/decoder is optimal for all  $\mu$  except  $\mu = 4$ , when  $\Delta$  is only one bit less than ideal.

We will establish Theorem 3 via a random code argument. Towards this end, we establish the following lemmas.

*Lemma 5:* Let  $1 \leq m \leq n$ , and let the  $m \times n$  matrix  $A$  over  $GF(2)$  be chosen at random with uniform distribution on the set of  $2^{mn}$  binary  $m \times n$  matrices. Then, for  $1 \leq L \leq m$ ,

$$\Pr\{\text{rank } A < m - L\} \leq 2^{-(L+1)(n-m)+n}.$$

*Proof of Lemma 5:* Let us choose the  $n$  columns of  $A$  sequentially and independently. Let  $d(j)$  be the dimension of the linear space spanned by the first  $j$  columns. Suppose that  $d(j) = k \leq m$ . With probability  $2^{k-m}$ ,  $d(j+1) = k$ ; and with probability  $(1 - 2^{k-m})$ ,  $d(j+1) = k+1$ . This sequential choice of the columns is modeled by the Markov chain of Fig. 2.

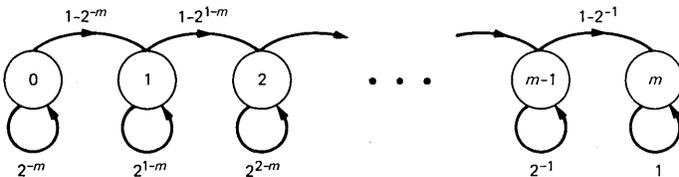


Fig. 2—Markov chain used in proof of Lemma 5.

Begin at state 0. With each choice of a column, advance one state if and only if this choice increases the dimension of the space spanned by the columns chosen so far. The rank of the matrix  $A$  is  $d(n)$  and is equal to the state at which we find ourselves after all  $n$  columns are chosen. Let  $\Gamma(k)$  denote the set of paths  $\pi$  that start at state 0 and terminate at state  $k$  ( $0 \leq k \leq m$ ). Then

$$\Pr\{\text{rank } A < m - L\} = \sum_{k=0}^{m-L-1} \sum_{\pi \in \Gamma(k)} \Pr\{\pi\}. \quad (18)$$

Now let the path  $\pi \in \Gamma(k)$ . This path contains exactly  $n - k$  self-loops, each of which has probability  $\leq 2^{-m+k}$ . Thus, for  $\pi \in \Gamma(k)$ ,

$$\Pr\{\pi\} \leq 2^{-(m+k)(n-k)}.$$

Also, since  $|\Gamma(k)| = \binom{n}{k}$ , eq. (18) yields

$$\Pr\{\text{rank } A < m - L\} \leq \sum_{k=0}^{m-L-1} \binom{n}{k} 2^{-(m-k)(n-k)}.$$

Since the exponent is nondecreasing in  $k$  ( $k \leq m \leq n$ ), we have

$$\begin{aligned} \Pr\{\text{rank } A < m - L\} &\leq \sum_{k=0}^{m-L-1} \binom{n}{k} 2^{-(L+1)(n-m+L+1)} \\ &\leq 2^n 2^{-(L+1)(n-m)}, \end{aligned}$$

which is Lemma 5.  $\square$

*Lemma 6:* Let  $1 \leq m \leq n$ , and let the  $m \times n$  matrix  $A$  over  $GF(2)$  be chosen at random with uniform distribution on the set of  $2^{mn}$  binary  $m \times n$  matrices. Then

$$\begin{aligned} \Pr\{\text{rank } A = m\} &= \prod_{j=0}^{m-1} (1 - 2^{j-n}) \\ &\geq \exp \left\{ \frac{-m2^{m-1-n}}{1 - 2^{m-1-n}} \right\} \geq \left( 1 - \frac{m2^{m-1-n}}{1 - 2^{m-1-n}} \right). \end{aligned}$$

*Proof of Lemma 6:* Choose the rows of  $A$  sequentially. As in the proof of Lemma 5, the probability that the dimension of the space spanned by the first  $j$  rows is equal to  $j$  is

$$(1 - 2^{-n})(1 - 2^{-n+1}) \dots (1 - 2^{-n+j-1}).$$

The rest of the lemma follows from  $\ln(1 - u) \geq -u/(1 - u)$  and  $e^{-u} \geq 1 - u$ .

We now turn to Theorem 3. Let  $R > 0$  be given and held fixed. We will show that  $\delta = 1$  and  $\alpha = 1 - R$  is achievable, and the remainder of the theorem will follow from Lemma 1. Let  $\epsilon > 0$  be arbitrary. We

will show that there exists an encoder/decoder with parameters  $N$ ,  $K = RN$ ,  $\mu = (1 - R - \epsilon)N$ , and  $\Delta \geq K - L$ , provided that

$$L \geq 3/\epsilon. \quad (19)$$

We proceed as follows. Let  $H$  be a  $K \times N$  parity-check matrix, and let  $L$  satisfy (19). Let  $D^*(\mu)$  correspond to  $H$ , and define

$$\Psi(H) = \begin{cases} 1, & D^*(\mu) < K - L \text{ or } \text{rank}(H) < K, \\ 0, & \text{otherwise.} \end{cases} \quad (20)$$

We must show that there exists an  $H$  with  $\Phi(H) = 0$ . We can write

$$\Psi(H) \leq \sum_{\substack{\mathcal{S} \subseteq \{1, \dots, N\} \\ |\mathcal{S}| = \mu}} \Phi(H, \mathcal{S}) + \Phi_0(H), \quad (21a)$$

where

$$\Phi_0(H) = \begin{cases} 1, & \text{rank}(H) < K, \\ 0, & \text{otherwise,} \end{cases} \quad (21b)$$

and

$$\Phi(H, \mathcal{S}) = \begin{cases} 1, & D(\mathcal{S}) < K - L, \\ 0, & \text{otherwise.} \end{cases} \quad (21c)$$

If we choose  $H = (\mathbf{C}_1, \dots, \mathbf{C}_N)$  at random with uniform distribution on the set of  $2^{K \cdot N}$  binary  $K \times N$  matrices, then (21) yields

$$E\Psi(H) \leq \sum_{|\mathcal{S}| = \mu} E\Phi(H, \mathcal{S}) + E\Phi_0(H). \quad (22)$$

Let  $\mathcal{S}$ , with  $|\mathcal{S}| = \mu$ , be arbitrary, and let  $A = (\mathbf{C}_{i_1} \mathbf{C}_{i_2} \dots \mathbf{C}_{i_{N-\mu}})$ , where  $\mathcal{S} = \{i_1, \dots, i_{N-\mu}\}$ . Then,  $\Phi(H, \mathcal{S}) = 1$  if and only if  $\text{rank } A < K - L$ , and  $E\Phi(H, \mathcal{S}) = \Pr\{\text{rank } A < K - L\}$ . We can apply Lemma 5 with  $n = N - \mu$ ,  $m = K$ , to obtain

$$E\Phi(H, \mathcal{S}) \leq 2^{-(L+1)(N-\mu-K)+(N-\mu)}. \quad (23)$$

Similarly, we can apply Lemma 6 with  $A = H$ ,  $n = N$ , and  $m = K$ , to obtain

$$E\Phi_0(H) \leq \frac{K2^{K-N-1}}{1 - 2^{K-N-1}} \leq \frac{K2^{K-N}}{1 - 2^{K-N}}. \quad (24)$$

Since there are no more than  $2^N$  subsets  $\mathcal{S}$ , (22) through (24) yield (using  $N - \mu - K = \epsilon N$ ,  $K = RN$ )

$$\begin{aligned} E\Psi(H) &\leq 2^{-(L+1)(N-\mu-K)+(N-\mu)+N} + \frac{K2^{K-N}}{1 - 2^{K-N}} \\ &\leq 2^{-L\epsilon N + 2N} + \frac{RN2^{-(1-R)N}}{1 - 2^{-(1-R)N}}. \end{aligned} \quad (25)$$

Since  $L$  satisfies (19), the first term in the right member of (25) is less than  $1/2$ . Furthermore, for  $N$  sufficiently large, the second term in (25) is also less than  $1/2$ . Thus

$$E\Psi(H) < 1.$$

Since  $\Psi(\cdot)$  is an integer-valued function, there must exist a  $K \times N$  matrix  $H_0$  such that  $\Psi(H_0) = 0$ , so that  $\text{rank } H_0 = K$  and for the corresponding encoder/decoder,  $\Delta = D^*(\mu) \geq K - L$ , which is what we set out to prove. Thus, we have shown that for arbitrary  $R > 0$ , the triples  $(R, \alpha, \delta)$ , where  $\alpha \leq 1 - R$ ,  $\delta \leq 1$ , are achievable, completing the proof of Theorem 3.

## V. PROOF OF THEOREM 4

We restate Theorem 4 here:

*Theorem 4: For all  $K, N$  for which  $1 \leq K \leq N$ , and  $\mu = N - K$  there exists an encoder/decoder pair for which error-free decoding is possible, and*

$$\Delta \geq K - 1 - \frac{2.23}{\sqrt[4]{N}}.$$

In particular, as  $K$  and  $N$  grow at a fixed ratio (i.e.,  $K = RN$  as  $N \rightarrow \infty$ ),  $\Delta$  can be made as close as we like to  $K - 1$ .

We shall use a random coding argument to prove Theorem 4. A code for this problem is a partition of all  $2^N$  sequences into  $2^K$  message "bins," each of size  $2^{N-K}$ . Each message is transmitted by randomly choosing an element of the bin corresponding to the desired message, with all elements chosen equally likely. The ensemble of codes is the set of all partitions, chosen equally likely.

We shall obtain a lower bound on  $\Delta$  as a function of the code selected and then show that the average (over the ensemble of codes) of  $\Delta$  for any subset of  $\mu$  bits selected is greater than  $K - 1$ . We then obtain the variance (again, over the ensemble of codes) of the bound. Using the variance, we can invoke Chebyshev's inequality to bound the probability that the code yields  $\Delta < K - 1 - \epsilon$  for any positive  $\epsilon$ . Since the probability distribution over the code ensemble is uniform, this bound is the fraction of the number of all possible codes that fail to provide acceptable secrecy for the  $\mu$  bits selected. Since there are a limited number of sets of  $\mu$  bits  $\left[ \binom{N}{\mu} \right]$ , to be precise, the total number of codes that fail to provide acceptable secrecy for *any* choice of  $\mu$  bits can be bounded. The condition that the total number of such codes must be less than the number of codes in the ensemble, which guarantees the existence of a good code, yields the theorem.

While the proof as just outlined is conceptually simple, the mechanics (primarily the variance computation) are complicated, and are outlined in the Appendix.

We proceed: Let  $\mathcal{S}$  be an index set (of size  $N - K$ ), which is a subset of  $\{1, 2, \dots, N\}$ . Let  $\mathbf{Z}$  be the observed vector resulting from observations of the positions of transmitted code word  $\mathbf{X}$  with indices in  $\mathcal{S}$ , and let  $\mathbf{S}$  be the message (we omit superscripts). Now

$$\Delta \triangleq H(\mathbf{S} | \mathbf{Z}) = H(\mathbf{S}) - H(\mathbf{Z}) + H(\mathbf{Z} | \mathbf{S}) \quad (26)$$

follows from applying the identities

$$H(A, B) = H(B, A) = H(B | A) + H(A).$$

By the construction of the code, all values of  $\mathbf{X}$  are equally likely, so that all values of  $\mathbf{Z}$  are as well, and (26) becomes

$$\Delta = K - (N - K) + H(\mathbf{Z} | \mathbf{S}).$$

Let  $n(\mathbf{z}, \mathbf{s})$  be the number of code words corresponding to message  $\mathbf{s}$  that are consistent with  $\mathbf{z}$ , for each  $\mathbf{s}$  and  $\mathbf{z}$ . Then,

$$\Pr[\mathbf{Z} = \mathbf{z} | \mathbf{S} = \mathbf{s}] = \frac{n(\mathbf{z}, \mathbf{s})}{2^{N-K}},$$

since the total number of code words corresponding to  $\mathbf{s}$  is  $2^{N-K}$ . Therefore,

$$\begin{aligned} H(\mathbf{Z} | \mathbf{S}) &= - \sum_{\mathbf{s}} \sum_{\mathbf{z}} p(\mathbf{z}, \mathbf{s}) \log_2 p(\mathbf{z} | \mathbf{s}) \\ &= - \sum_{\mathbf{s}} \sum_{\mathbf{z}} \frac{n(\mathbf{z}, \mathbf{s})}{2^{N-K}} p(\mathbf{s}) \log_2 \frac{n(\mathbf{z}, \mathbf{s})}{2^{N-K}} \\ &= \frac{1}{2^N} \sum_{\mathbf{s}} \sum_{\mathbf{z}} n(\mathbf{z}, \mathbf{s}) [N - K - \log_2 n(\mathbf{z}, \mathbf{s})]. \end{aligned}$$

Clearly,

$$\begin{aligned} \sum_{\mathbf{z}} n(\mathbf{z}, \mathbf{s}) &= 2^{N-K} \sum_{\mathbf{z}} P(\mathbf{Z} = \mathbf{z} | \mathbf{S} = \mathbf{s}) \\ &= 2^{N-K}, \end{aligned}$$

and the number of  $\mathbf{s}$  is  $2^K$ , so

$$H(\mathbf{Z} | \mathbf{S}) = N - K - \frac{1}{2^N} \sum_{\mathbf{s}} \sum_{\mathbf{z}} n(\mathbf{z}, \mathbf{s}) \log_2 n(\mathbf{z}, \mathbf{s}),$$

yielding

$$\Delta = K - \frac{1}{2^N} \sum_{\mathbf{s}} \sum_{\mathbf{z}} n(\mathbf{z}, \mathbf{s}) \log_2 n(\mathbf{z}, \mathbf{s}).$$

We will lower bound  $\Delta$  by using

$$\begin{aligned} \log_2 x &= \frac{1}{\ln 2} \ln x \\ &= \frac{1}{\ln 2} \left( \ln \frac{x}{\alpha} + \ln \alpha \right) \\ &\leq \frac{1}{\ln 2} \left( \frac{x}{\alpha} - 1 + \ln \alpha \right), \end{aligned}$$

which is true for any positive  $\alpha$ . Therefore,

$$\Delta \geq \Delta_b = K - \frac{1}{2^N \ln 2} \sum_{\mathbf{s}} \sum_{\mathbf{z}} n(\mathbf{z}, \mathbf{s}) \left[ \frac{n(\mathbf{z}, \mathbf{s})}{\alpha} - 1 + \ln \alpha \right].$$

Since we have just shown that

$$\sum_{\mathbf{s}} \sum_{\mathbf{z}} n(\mathbf{z}, \mathbf{s}) = 2^N,$$

then

$$\Delta_b = K + \frac{1}{\ln 2} - \log_2 \alpha - \frac{1}{2^N \ln 2} \sum_{\mathbf{s}} \sum_{\mathbf{z}} \frac{n^2(\mathbf{z}, \mathbf{s})}{\alpha}.$$

Interpreted as a function of the randomly selected code,  $\Delta_b$  is a random variable. The expectation  $\overline{n^2(\mathbf{z}, \mathbf{s})}$  is constant for all  $\mathbf{z}$  and  $\mathbf{s}$  by the symmetry of the code selection, and

$$\overline{\Delta_b} = K + \frac{1}{\ln 2} - \log_2 \alpha - \frac{1}{\ln 2} \frac{\overline{n^2(\mathbf{z}, \mathbf{s})}}{\alpha}.$$

This bound is maximized by using  $\alpha = \overline{n^2(\mathbf{z}, \mathbf{s})}$ , which yields

$$\overline{\Delta_b} = K - \log_2 \overline{n^2(\mathbf{z}, \mathbf{s})}.$$

The variance of  $\Delta_b$  can be written as

$$\text{Var}(\Delta_b) = E[(\Delta_b - \overline{\Delta_b})^2] = \frac{1}{2^{2N} \ln^2 2} \left[ \sum_{\mathbf{s}} \sum_{\mathbf{z}} \frac{n^2(\mathbf{z}, \mathbf{s}) - \overline{n^2}}{n^2} \right]^2. \quad (27)$$

Denote the four generic  $(\mathbf{z}, \mathbf{s})$  pairs by (1, 1), (1, 2), (2, 1), and (2, 2). Also let

$$\begin{aligned} A &= 2^N, \\ B &= 2^K, \end{aligned}$$

and

$$C = 2^{N-K}.$$

Equation (27) can be written as

$$\begin{aligned} \text{Var}(\Delta_b) = & \frac{1}{A \ln^2 2} \left\{ \left[ \frac{n^2(1, 1) - (\bar{n}^2)^2}{(\bar{n}^2)^2} \right] \right. \\ & + (C - 1) \left[ \frac{n^2(1, 1)n^2(2, 1) - (\bar{n}^2)^2}{(\bar{n}^2)^2} \right] \\ & + (B - 1) \left[ \frac{n^2(1, 1)n^2(1, 2) - (\bar{n}^2)^2}{(\bar{n}^2)^2} \right] \\ & \left. + (B - 1)(C - 1) \left[ \frac{n^2(1, 1)n^2(2, 2) - (\bar{n}^2)^2}{(\bar{n}^2)^2} \right] \right\}, \end{aligned}$$

or

$$\text{Var}(\Delta_b) = \frac{1}{(\bar{n}^2)^2 A \ln^2 2} [M_1 + M_2 + M_3 + M_4]. \quad (28)$$

Using the bounds on  $M_1, M_2, M_3,$  and  $M_4$  given in the Appendix gives

$$\begin{aligned} \text{Var}(\Delta_b) & \leq \frac{1}{A \ln^2 2} \frac{1}{(1 + T_1)^2} \\ & \quad \cdot [5T_1 + 5T_1^2 + T_1^3 - 2T_1 - 4T_1^2 - 2T_1^3 + T_1 + 4T_1^2] \\ & = \frac{1}{A \ln^2 2} \frac{1}{(1 + T_1)^2} [8T_1 + 5T_1^2 - T_1^3], \end{aligned}$$

where  $T_1$  (defined in the Appendix) is less than or equal to one. The right-hand side is maximized (over the allowable range of  $T_1$ ) by  $T_1 = 1$ , so that

$$\begin{aligned} \text{Var}(\Delta_b) & \leq \frac{3}{A \ln^2 2} \cong \frac{6.244}{A} \\ & = \frac{6.244}{2^N}. \end{aligned}$$

The mean of  $\Delta_b$  is

$$\bar{\Delta}_b = K - \log_2 \bar{n}^2.$$

Since (see the Appendix)

$$\bar{n}^2 = 1 + \frac{(B - 1)(C - 1)}{A - 1} \leq 2,$$

then

$$\bar{\Delta}_b \geq K - 1.$$

Chebyshev's inequality states

$$\Pr[\Delta_b \leq \bar{\Delta}_b - \epsilon] \leq \frac{\text{Var}(\Delta_b)}{\epsilon^2}.$$

Since  $\Delta \geq \Delta_b$  and  $\bar{\Delta}_b \geq K - 1$ , we have

$$\begin{aligned} \Pr [\Delta \leq K - 1 - \epsilon] &\leq \frac{\text{Var}(\Delta_b)}{\epsilon^2} \\ &\leq \frac{6.244}{\epsilon^2 2^N}. \end{aligned} \quad (29)$$

If we define

$C$  = set of all codes,  
 $C(\mathcal{S})$  = set of codes for which  $H(\mathbf{S} | \mathbf{Z}) \leq K - 1 - \epsilon$ ,

and

$|A|$  = number of elements of  $A$ ,

then, since all codes are equally likely, (29) is equivalent to

$$\frac{|C(\mathcal{S})|}{|C|} \leq \frac{6.244}{\epsilon^2 2^N}.$$

Now the set of codes for which  $H(\mathbf{S} | \mathbf{Z}) \leq K - 1 - \epsilon$  for some  $\mathcal{S}$  is  $\bigcup_{\mathcal{S}} C(\mathcal{S})$ , and

$$|\bigcup_{\mathcal{S}} C(\mathcal{S})| \leq \sum_{\mathcal{S}} |C(\mathcal{S})|.$$

Therefore,

$$\begin{aligned} \frac{|\bigcup_{\mathcal{S}} C(\mathcal{S})|}{|C|} &\leq \sum_{\mathcal{S}} \frac{|C(\mathcal{S})|}{|C|} \\ &\leq (\# \mathcal{S}'\text{s}) \frac{6.244}{\epsilon^2 2^N}. \end{aligned}$$

The number of possible  $\mathcal{S}$ 's is the number of subsets of the set  $\{1, \dots, N\}$  of size  $N - K$ , given by  $\binom{N}{N - K}$ . Therefore,

$$\frac{|\bigcup_{\mathcal{S}} C(\mathcal{S})|}{|C|} \leq \binom{N}{N - K} \frac{6.244}{\epsilon^2 2^N}.$$

If  $R = K/N$ , then

$$\binom{N}{N - K} \leq \frac{1}{\sqrt{2\pi R(1 - R)N}} 2^{Nh(R)},$$

where  $h(\cdot)$  is the binary entropy function and

$$\frac{|\bigcup_{\mathcal{S}} C(\mathcal{S})|}{|C|} \leq \frac{6.244}{\epsilon^2 \sqrt{2\pi R(1 - R)N}} 2^{N(h(R)-1)}. \quad (30)$$

As long as the left-hand side of (30) is strictly less than 1, at least one code on the ensemble falls outside  $\bigcup_{\mathcal{S}} C(\mathcal{S})$ , that is, provides

$\Delta > K - 1 - \epsilon$  for all  $\mathcal{L}$ . This is guaranteed as long as

$$\epsilon^2 \geq \frac{6.244}{\sqrt{2\pi R(1-R)}} \frac{2^{N(h(R)-1)}}{\sqrt{N}} \frac{2.49}{\sqrt{R(1-R)}} \frac{2^{N(h(R)-1)}}{\sqrt{N}}.$$

Therefore, the existence of codes for which

$$\Delta \geq K - 1 - \frac{\sqrt{2.49}}{\sqrt[4]{R(1-R)N}} 2^{\frac{N}{2}[h(R)-1]}$$

is guaranteed.

For  $R = 1/2$ , the worst case, this becomes

$$\Delta \geq K - 1 - \frac{2.23}{\sqrt[4]{N}},$$

which is Theorem 4.

## REFERENCES

1. A. D. Wyner, "The Wire-Tap Channel," B.S.T.J., 54, No. 8 (October 1975), pp. 1355-87.
2. R. G. Gallager, *Information Theory and Reliable Communication*, New York: Wiley, 1968.

## APPENDIX

### *Statistics of Codes Used in Section V*

Here we present the expectations necessary to obtain the variance of  $\Delta_b$ , which is defined in Section V.

Let  $n(1, 1)$  and  $n(2, 1)$  [resp.  $n(1, 2)$  and  $n(2, 2)$ ] be the numbers of code words in the bin assigned to message 1 (resp. 2), which are consistent with distinct sequences  $\mathbf{z}_1$  and  $\mathbf{z}_2$ . Recall that we have previously defined:

$$\begin{aligned} A &\triangleq 2^N \\ B &\triangleq 2^K \\ C &\triangleq 2^{N-K}. \end{aligned}$$

Under the assumption that all partitions of the  $2^N$  code word sequences into  $2^K$  message bins are chosen with equal likelihood, then the probability of the four-tuple  $[n(1, 1), n(2, 1), n(1, 2), n(2, 2)]$  is the number of codes with those values divided by the total number of codes.

This last ratio can be reduced to the ratio between (a) a numerator consisting of the number of ways that the bins (each of size  $C$ ) corresponding to messages one and two can be assembled to include exactly  $n(1, 1)$  and  $n(1, 2)$  code words consistent with  $\mathbf{z}_1$  (from a

candidate pool of  $B$  elements) and exactly  $n(2, 1)$  and  $n(2, 2)$  code words consistent with  $\mathbf{z}_2$  (from a disjoint pool of  $B$  elements); and (b) a denominator consisting of the total number of ways the two bins can be assembled.

The numerator is

$$N = \begin{bmatrix} B \\ n(1, 1), n(1, 2) \end{bmatrix} \begin{bmatrix} B \\ n(2, 1), n(2, 2) \end{bmatrix} \cdot \begin{bmatrix} A - 2B \\ C - n(1, 1) - n(2, 1), C - n(1, 2) - n(2, 2) \end{bmatrix},$$

which is the product of the number of ways the elements consistent with  $\mathbf{z}_1$  can be drawn, times the number of ways the elements consistent with  $\mathbf{z}_2$  can be drawn, times the number of ways the elements consistent with neither can be drawn.

The denominator is just

$$D = \begin{pmatrix} A \\ C, C \end{pmatrix},$$

i.e., the number of ways that two arbitrary sets of size  $C$  can be drawn. The notation we have used is the standard trinomial coefficient:

$$\begin{pmatrix} X \\ Y, Z \end{pmatrix} \triangleq \frac{X!}{Y!Z!(X - Y - Z)!},$$

defined to be zero when any of  $Y, Z$  or  $Y + Z$  are greater than  $X$ , or when any of  $X, Y$ , or  $Z$  are negative. Defining, for the sake of notational compactness,

$$\begin{aligned} \alpha &\triangleq n(1, 1) \\ \beta &\triangleq n(1, 2) \\ \gamma &\triangleq n(2, 1) \\ \delta &\triangleq n(2, 2), \end{aligned}$$

then

$$P(\alpha, \beta, \gamma, \delta) = \frac{\begin{pmatrix} B \\ \alpha, \beta \end{pmatrix} \begin{pmatrix} B \\ \gamma, \delta \end{pmatrix} \begin{pmatrix} A - 2B \\ C - \alpha - \gamma, C - \beta - \delta \end{pmatrix}}{\begin{pmatrix} A \\ C, C \end{pmatrix}}.$$

The quantities needed for the analysis of Section V are  $\overline{\alpha^2}$ ,  $\overline{\alpha^4}$ ,  $\overline{\alpha^2\beta^2}$ ,  $\overline{\alpha^2\gamma^2}$ , and  $\overline{\alpha^2\gamma^2}$ . The actual calculations are involved and unenlightening. Everything can be evaluated using the relationships

$$\sum_i \begin{pmatrix} X \\ i \end{pmatrix} \begin{pmatrix} Y \\ Z - i \end{pmatrix} = \begin{pmatrix} X + Y \\ Z \end{pmatrix}$$

and

$$\begin{pmatrix} X \\ Y, Z \end{pmatrix} Y = X \begin{pmatrix} X - 1 \\ Y - 1, Z \end{pmatrix}.$$

If we define the auxiliary quantities:

$$T_j = \frac{(B - j)(C - j)}{A - j},$$

and note since  $BC = A$ , that  $T_0 = 1$  and  $T_{j+1} < T_j$  so long as  $j$  is less than the smaller of  $B$  and  $C$ , then

$$\overline{\alpha^2} = 1 + T_1$$

and

$$\overline{\alpha^4} = 1 + 7T_1 + 6T_1T_2 + T_1T_2T_3.$$

Also,

$$\overline{\alpha^2\beta^2} = \frac{A - C}{A - 1} \left[ 1 + \frac{2(C - 1)(B - 2)}{A - 2} + \frac{(C - 1)^2(B - 2)(B - 3)}{(A - 2)(A - 3)} \right],$$

$$\overline{\alpha^2\gamma^2} = \frac{A - B}{A - 1} \left[ 1 + \frac{2(B - 1)(C - 2)}{A - 2} + \frac{(B - 1)^2(C - 2)(C - 3)}{(A - 2)(A - 3)} \right],$$

and

$$\overline{\alpha^2\delta^2} = \frac{A}{A - 1} \left[ 1 + \frac{2(B - 1)(C - 1)}{A - 2} + \frac{(B - 1)^2(C - 1)^2}{(A - 2)(A - 3)} \right].$$

Referring back to the quantities needed in (25), we have

$$\begin{aligned} M_1 &= \overline{\alpha^4} - (\overline{\alpha^2})^2 \\ &= 1 + 7T_1 + 6T_1T_2 + T_1T_2T_3 - (1 + T_1)^2 \\ &= 5T_1 + 6T_1T_2 + T_1T_2T_3 - T_1^2. \end{aligned}$$

Since  $T_3 < T_2 < T_1$ , then

$$M_1 < 5T_1 + 5T_1^2 + T_1^3.$$

Next

$$\begin{aligned} M_2 &= (C - 1) \overline{[n^2(1, 1)n^2(2, 1) - (n^2)^2]} \\ &= (C - 1) \overline{[\alpha^2\gamma^2 - (\overline{\alpha^2})^2]} \\ (C - 1) &\left\{ \frac{A - B}{A - 1} \left[ 1 + \frac{2(B - 1)(C - 2)}{A - 2} \right. \right. \\ &\quad \left. \left. + \frac{(B - 1)^2(C - 2)(C - 3)}{(A - 2)(A - 3)} \right] - (1 + T_1)^2 \right\}. \end{aligned}$$

Now

$$\frac{1}{T_1} \frac{(B-1)(C-2)}{A-2} = \frac{C-2}{C-1} \frac{A-1}{A-2} = \frac{AC-A-C+2-A}{AC-A-C+2-C} < 1,$$

since  $A > C$ . Similarly,  $[(B-1)(C-3)]/(A-3) < T_1$ , so

$$\begin{aligned} M_2 &\leq (C-1) \left[ \frac{A-B}{A-1} (1+2T_1+T_1^2) - (1+T_1)^2 \right] \\ &= (C-1) \left[ (1+T_1)^2 \left( \frac{A-B}{A-1} - 1 \right) \right] \\ &= -\frac{(B-1)(C-1)}{A-1} (1+T_1)^2 \\ &= -T_1(1+T_1)^2. \end{aligned}$$

Since  $\alpha^2\gamma^2$  is just  $\alpha^2\beta^2$  with the roles of  $B$  and  $C$  reversed,  $M_2$  and  $M_3$  are “dual” to each other in this sense. Therefore,  $M_3$  satisfies the same bound as  $M_2$ :

$$M_3 \leq -T_1(1+T_1)^2.$$

The last term needed is

$$\begin{aligned} M_4 &= (B-1)(C-1) [\overline{n^2(1,1)n^2(2,2)} - (\overline{n^2})^2] \\ &= (B-1)(C-1) \left\{ \frac{A}{A-1} \left[ 1 + \frac{2(B-1)(C-1)}{A-2} \right. \right. \\ &\quad \left. \left. + \frac{(B-1)^2(C-1)^2}{(A-2)(A-3)} \right] - (1+T_1)^2 \right\} \\ &= (B-1)(C-1) \left[ \frac{A}{A-1} + 2T_1 \frac{A}{A-2} \right. \\ &\quad \left. + \frac{A(A-1)}{(A-2)(A-3)} T_1^2 - (1+T_1)^2 \right] \\ &= (B-1)(C-1) \left[ \frac{1}{A-1} + 2T_1 \frac{2}{A-2} \right. \\ &\quad \left. + T_1^2 \left( \frac{A(A-1)}{(A-2)(A-3)} - 1 \right) \right] \\ &= (B-1)(C-1) \left[ \frac{1}{A-1} + 2T_1 \frac{2}{A-2} + T_1^2 \frac{4A-6}{(A-2)(A-3)} \right] \\ &= T_1 + 4T_1 \frac{(B-1)(C-1)}{A-2} \\ &\quad + T_1^2 \frac{4A-6}{(A-2)(A-3)} (B-1)(C-1). \end{aligned}$$

All nontrivial cases of interest have  $K$  and  $N-K \geq 2$ , so that  $B$  and  $C$  are each  $\geq 4$ , which implies that

$$(B - 1)(C - 1) = BC - B - C + 1 \leq A - 7,$$

so that

$$\frac{(B - 1)(C - 1)}{A - 3} < 1.$$

Therefore,

$$\begin{aligned} M_4 &\leq T_1 + 4T_1 \frac{A - 7}{A - 2} + T_2^2 \frac{4A - 6}{A - 2} \\ &\leq T_1 + 4T_1 - \frac{20}{A - 2} T_1 + T_1^2 \frac{4A - 6}{A - 2}. \end{aligned}$$

Since  $T_1 < 1$ , then  $T_1^2 < T_1$ , so

$$\begin{aligned} M_4 &\leq 5T_1 + T_1^2 \left( \frac{4A - 6}{A - 2} - \frac{20}{A - 2} \right) \\ &= 5T_1 + T_1^2 \frac{4A - 26}{A - 2} \\ &< 5T_1 + 4T_1^2. \end{aligned}$$

#### AUTHORS

**Lawrence H. Ozarow**, B.S. (Electrical Engineering), 1970, Columbia University; S.M. and Ph.D. (Electrical Engineering), 1973 and 1979, respectively, The Massachusetts Institute of Technology; CNR, Inc., 1973–1976; AT&T Bell Laboratories, 1979—. From 1973 to 1976 Mr. Ozarow was with CNR, Inc., Needham, Massachusetts, working in modulation and coding for high-frequency channels. From 1976 to 1978 he was a research assistant at the MIT Lincoln Laboratories, in the area of packet switched satellite communication. Since 1979 he has been with the Mathematics and Statistics Research Center at AT&T Bell Laboratories working in the areas of Information and Communication Theory. In 1982 he was at AT&T Bell Laboratories Transmission Systems Division.

**Aaron D. Wyner**, B.S., 1960, Queens College; B.S.E.E., M.S., Ph.D., 1960, 1961, and 1963, respectively, Columbia University; AT&T Bell Laboratories, 1963—. Mr. Wyner has been doing research in various aspects of information and communication theory and related mathematical problems. He is presently Head of the Communications Analysis Research department. He spent the year 1969–70 visiting the Department of Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel, and the Faculty of Electrical Engineering, at Technion, Haifa, Israel, on a Guggenheim Foundation Fellowship. He has also been a full- and part-time faculty member at Columbia University, Princeton University, and the Polytechnic Institute of Brooklyn. He is Chairman of the Metropolitan New York Chapter of the IEEE Information Theory Group, Associate Editor of the Group's Transactions, and co-chairperson of two international symposia. President, IEEE Information Theory Group, 1976; Fellow, IEEE Member, AAAS, Tau Beta Pi, Eta Kappa Nu, Sigma Xi. Since September 1983 he has been Editor-in-Chief of the IEEE Transactions on Information Theory.



## Stationary and Cyclostationary Finite Buffer Behaviour Computation via Levinson's Method

By M. H. ACKROYD\*

(Manuscript received November 8, 1983)

A model for a finite buffer is described. The equilibrium equations for the number in the buffer have Toeplitz form. This enables Levinson's method, which is widely used in the computation of optimal linear filters, to be used to compute efficiently the distribution of the number in the buffer, for both stationary and cyclostationary systems. The method can also be used to compute distributions in other, mathematically similar, queueing problems. These include the computation of the waiting time distribution in discrete time G/G/1 queues.

### I. INTRODUCTION

Buffer analysis is important in the design of data communication systems in which messages are queued whilst awaiting transmission. In particular, it is important to be able to choose the capacity of a buffer so that the probability of overflow is acceptably small. Methods for computing the distribution of the number in a finite buffer continue, therefore, to interest telecommunications engineers.

Various authors have dealt with computing the distribution of the number in a buffer (see, for example, Refs. 1 through 7). These references deal with models where the arrival statistics are constant, except for Ref. 7, which discusses the case where the distributions of successive arrivals at the buffer vary cyclically. This case arises when the arrivals are messages from cyclically polled sources, with differing

---

\* AT&T Bell Laboratories.

---

Copyright © 1984 AT&T. Photo reproduction for noncommercial use is permitted without payment of royalty provided that each reproduction is done without alteration and that the Journal reference and copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free by computer-based and other information-service systems without further permission. Permission to reproduce or republish any other portion of this paper must be obtained from the Editor.

characteristics. In such cases, the equilibrium distribution of the number in the buffer varies cyclically, and the behaviour is termed *cyclostationary*. The present paper deals with both stationary and cyclostationary systems.

Dor<sup>1</sup> and Chu<sup>2</sup> studied systems where a single character is removed from the buffer at a time. In such cases the distribution of the number in the buffer satisfies a linear recurrence relation, from which the distribution can be computed directly. Chu<sup>3</sup> and Rudin<sup>4</sup> dealt with the distribution of the number in the buffer in systems where up to  $N$  characters can be transmitted at a time. Chu used Gaussian elimination for the direct solution of the equilibrium equations for buffer sizes of up to 50 or so. Rudin avoided the direct solution of the equilibrium equations by dealing with the transient case. By starting with a known distribution, such as the one corresponding to an empty system, the successive distributions can be computed iteratively. The equilibrium distribution can be obtained by continuing the iterations until convergence, to within an appropriate criterion, is obtained. Bagchi and Templeton<sup>5</sup> also applied this approach to the treatment of finite buffer problems. The method is applicable to stationary, to cyclostationary, and, indeed, to other nonstationary problems. Each iteration involves the computation of a discrete convolution. (Later in the paper the method is referred to as *iterated convolutions*.) For systems with a high utilization factor and large buffer size, the total volume of computation can be immense—even when the convolutions are computed efficiently via a fast Fourier transform algorithm. Except for this disadvantage, experience shows that the method is a satisfactory way of obtaining the equilibrium distribution.

Leihs and Kobayashi<sup>6</sup> avoided the difficulties of exact computation by obtaining bounds on buffer overflow probabilities. They exploited the similarity between the fundamental equation for the number in the buffer and the fundamental equation for the waiting time in a G/G/1 queue. From known bounds on the G/G/1 waiting-time distribution, they derived bounds on the overflow probability in finite buffer systems.

It seems that the solution of the equilibrium equations for buffer sizes such as one hundred to one thousand or more has widely been considered impractical. In this paper, it is shown that the matrix of the equilibrium equations, for one finite buffer model, has Toeplitz form, i.e., the elements on any given diagonal are all equal. This enables Levinson's method,<sup>8,9</sup> which is widely used to solve Toeplitz equations in the design of optimal linear filters, to be used to compute, economically, the stationary distribution of the number in the buffer. To solve for  $K$  unknowns, Levinson's method requires computer time proportional to  $K^2$ , rather than  $K^3$ , as needed by general methods of

solving linear equations. The paper also shows that, in the cyclostationary case, the matrix of the equilibrium equations has block Toeplitz form, i.e., when partitioned, the submatrices on any given diagonal are all equal. This enables the extension of Levinson's method to the block Toeplitz case, which is used in multichannel linear filter design,<sup>9,10</sup> to be used to compute, economically, the cyclostationary distribution of the number in the buffer.

One commonly used model for a finite buffer has equilibrium equations for the distribution of the number in the buffer that do not have Toeplitz form. However, it is shown that the system can be represented as a cyclostationary one, with period two, so that it becomes possible to use Levinson's method to solve its equilibrium equations efficiently. Finally, the paper points out that Levinson's method can supply numerical solutions to other queueing problems whose form is mathematically similar to the one dealt with here. These problems include the calculation of the waiting time distribution in discrete time G/G/1 queues and cyclostationary queueing problems arising in the performance analysis of clocked schedules for real-time software.

## II. FORMULATION—STATIONARY CASE

Time is assumed to be divided into a sequence of intervals. The number of units in the buffer just prior to the end of each interval is the quantity of prime interest. The units may be bits, characters, messages, or other entities, depending on the application. During the  $n$ th interval, there may be both arrivals and departures. The net potential number of units arriving in the interval, a number that can be positive or negative, is denoted by  $X_n$ . The number in the buffer is constrained to the range  $[0, K]$ , so that the number in the buffer at the end of the  $n$ th interval is given by

$$Y_n = \min[K, (Y_{n-1} + X_n)^+], \quad (1)$$

where  $(z)^+$  denotes  $\max(0, z)$ .

There are various physical interpretations of this formulation. As one example, suppose that  $V_n$  units arrive during the  $n$ th interval and, at the end of the interval, as many units as permitted by a packet size  $W_n$  are removed from the buffer for transmission.  $X_n$  in (1) then represents the difference  $V_n - W_n$ . With this interpretation, the number in the buffer during the interval can exceed  $K$ . If the limit on the number in the buffer were imposed throughout the interval, rather than just at its end, the system would be correctly described not by (1) but by

$$Y'_n = \{\min[(Y'_{n-1} + V_n), K] - W_n\}^+. \quad (2)$$

The behaviour of this system can differ somewhat from that of (1),

such as when a large value of  $V_n$  is accompanied by a similar large value of  $W_n$ . The computation of the distribution of the number in the buffer for a system described by (2) is discussed in Section VI.

Another interpretation of (1) is applicable when units arrive at the buffer in batches at purely random instants in continuous time and when, also at purely random instants, units are removed, in batches, for transmission. With this interpretation,  $X_n$ , if positive, represents the number of units arriving in a group. If  $X_n$  is negative, it represents the number of units offered transmission at an instant.

The successive  $X_n$  are assumed to be independent and, in the stationary case, identically distributed, with distribution  $\{a_k\}$ , i.e.,

$$a_k = \text{probability}[X_n = k].$$

The cumulative distribution of  $Y_n$ , with the system in equilibrium, is  $\{q_k\}$ , i.e.,

$$q_k = \text{probability}[Y_n \leq k].$$

The cumulative distribution of  $Y_{n-1} + X_n$  is, because of the independence of  $X_n$  and  $Y_{n-1}$ , the discrete convolution of  $\{a_k\}$  and  $\{q_k\}$ . Because of this and because  $Y_n$  is constrained to the range  $[0, K]$ , the cumulative distribution of the number in the buffer can be expressed in terms of itself as

$$q_k = \begin{cases} 0, & k < 0 \\ \sum_{m=-\infty}^{\infty} q_m a_{k-m}, & 0 \leq k < K \\ 1, & k \geq K. \end{cases}$$

This leads to the following system of equations for the  $K$  unknown points in the distribution  $\{q_k\}$ :

$$\underbrace{\begin{bmatrix} d_0 & d_1 & \cdots & d_{K-1} \\ d_{-1} & d_0 & \cdots & d_{K-2} \\ \vdots & \vdots & \ddots & \vdots \\ d_{-(K-1)} & d_{-(K-2)} & \cdots & d_0 \end{bmatrix}}_{\mathbf{D}_K} \begin{bmatrix} q_0 \\ q_1 \\ \vdots \\ q_{K-1} \end{bmatrix} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{K-1} \end{bmatrix}, \quad (3)$$

where

$$d_0 = 1 - a_0$$

$$d_k = -a_{-k}, \quad k = \pm 1, \pm 2, \dots, \pm(K-1)$$

$$b_k = \sum_{m=K-k}^{\infty} a_{-m}, \quad k = 0, 1, \dots, K-1.$$

The equations in (3) could be solved by the use of a general-purpose routine for the solution of linear equations. However, for large  $K$ , the computer time and storage needed would be impractical. But the matrix in (3) is a Toeplitz matrix, i.e., one whose elements on any given diagonal are all equal. This enables Levinson's method to be used for the economical solution for the unknown probabilities.

### III. LEVINSON'S METHOD

When used to solve a system of Toeplitz equations with  $K$  unknowns, Levinson's method requires computer time proportional to  $K^2$ , rather than the  $K^3$  required by general methods for the solution of linear equations. In the process of solving (3) for a buffer size  $K$ , it also yields the distributions for buffer sizes 2, 3,  $\dots$ ,  $K - 1$ . This is an advantage when the purpose of the computation is to determine a suitable buffer size for a given application.

In signal processing applications, the elements of the matrix are normally autocorrelation coefficients, for which  $d_k = d_{-k}$ , a condition that normally does not apply in the buffer problem. Computer programs based on Levinson's method for signal processing applications are normally written to exploit the symmetry of the autocorrelation coefficients. Such programs, therefore, require modification before they can be used for the finite buffer problem.

Levinson's method, in a form suitable for the stationary finite buffer problem, is outlined as follows. At the  $p$ th stage of the recursion, the solution vector\*  $\mathbf{q}(p) = [q_0(p), q_1(p), \dots, q_{p-1}(p)]^T$  has previously been obtained, together with two auxiliary solution vectors,  $\mathbf{x}(p)$  and  $\mathbf{y}(p)$ , which satisfy

$$\begin{bmatrix} x_0(p), x_1(p), \dots, x_{p-1}(p) \\ y_{p-1}(p), y_{p-2}(p), \dots, y_0(p) \end{bmatrix} \mathbf{D}_p = \begin{bmatrix} u_x(p), 0, \dots, 0, 0 \\ 0, 0, \dots, 0, u_y(p) \end{bmatrix}, \quad (4)$$

where  $x_0(p) = y_0(p) = 1$ . The  $p$ th stage of the recursion involves two steps:

Step 1. Compute  $\mathbf{x}(p + 1)$ ,  $\mathbf{y}(p + 1)$  from  $\mathbf{x}(p)$ ,  $\mathbf{y}(p)$ .

Step 2. Compute  $\mathbf{q}(p + 1)$  from  $\mathbf{q}(p)$  and  $\mathbf{y}(p + 1)$ .

$\mathbf{x}(p + 1)$  and  $\mathbf{y}(p + 1)$  are formed as linear combinations of  $\mathbf{x}(p)$  and  $\mathbf{y}(p)$ :

$$\begin{aligned} & [x_0(p + 1), x_1(p + 1), \dots, x_{p-1}(p + 1), x_p(p + 1)] \\ & = [x_0(p), x_1(p), \dots, x_{p-1}(p), 0] \\ & \quad + c_x(p + 1)[0, y_{p-1}(p), \dots, y_1(p), y_0(p)] \end{aligned} \quad (5a)$$

\* The notation is altered here to denote explicitly the buffer size  $p$  to which the solution corresponds.

$$\begin{aligned}
& [y_p(p+1), y_{p-1}(p+1), \dots, y_1(p+1), y_0(p+1)] \\
& = c_y(p+1)[x_0(p), x_1(p), \dots, x_{p-1}(p), 0] \\
& \quad + [0, y_{p-1}(p), \dots, y_1(p), y_0(p)]. \tag{5b}
\end{aligned}$$

The coefficients  $c_x(p+1)$  and  $c_y(p+1)$  are chosen so that the computed  $\mathbf{x}(p+1)$  and  $\mathbf{y}(p+1)$  will satisfy an equation like (4). For this,

$$c_x(p+1) = -v_x(p)/u_x(p), \quad c_y(p+1) = -v_y(p)/u_x(p), \tag{6}$$

where

$$\begin{aligned}
v_x(p) &= \sum_{k=0}^{p-1} x_k(p) d_{k-p} \\
v_y(p) &= \sum_{k=0}^{p-1} y_k(p) d_{p-k} \tag{7}
\end{aligned}$$

and where

$$\begin{aligned}
u_x(p) &= \sum_{k=0}^{p-1} x_k(p) d_k = u_x(p-1) + x_{p-1}(p) d_{p-1} \\
u_y(p) &= \sum_{k=0}^{p-1} y_k(p) d_{-k} = u_y(p-1) + y_{p-1}(p) d_{1-p}. \tag{8}
\end{aligned}$$

The first step in the  $p$ th stage of the recursion consists of evaluating, in appropriate order, (5) through (8), so that  $\mathbf{x}(p+1)$  and  $\mathbf{y}(p+1)$  are obtained.

The solution vector  $\mathbf{q}(p+1)$  is obtained as a linear combination of  $\mathbf{q}(p)$  and  $\mathbf{y}(p+1)$ :

$$\begin{aligned}
[q_0(p+1), \dots, q_{p-1}(p+1), q_p(p+1)] &= [q_0(p), \dots, q_{p-1}(p), 0] \\
& \quad + c_q(p+1)[y_p(p+1), \dots, y_1(p+1), y_0(p+1)], \tag{9}
\end{aligned}$$

in which the coefficient  $c_q(p+1)$  is chosen so that  $\mathbf{q}(p+1)$  is a correct solution. For this,

$$c_q(p+1) = [b_p - \gamma(p)]/u_p(p+1), \tag{10}$$

where

$$\gamma(p) = \sum_{k=0}^{p-1} q_k(p) d_{k-p}. \tag{11}$$

The evaluation of (9) through (11), in proper order, completes the  $p$ th stage of the recursions. The recursions are started with  $x_0(1) = y_0(1) = 1$ ,  $v_x(1) = v_y(1) = d_0$  and  $q_0(1) = b_0/d_0$ .

These formulas can be programmed simply; a set of Fortran

subroutines to compute the stationary finite buffer solution via Levinson's method used a total of 49 lines of code.

#### IV. NUMERICAL EXPERIENCE

Accumulation of computational error is a possibility with recursive computations. However, the Levinson method is known to be numerically stable,<sup>11</sup> and practical experience shows that it can give adequate accuracy for buffer analysis.

The accuracy of results obtained by the use of Levinson's method has been checked by comparison with results obtained by iterated convolutions and with known analytical solutions. For example, when  $a_1 = 1 - \alpha$ ,  $a_{-1} = \alpha$ ,  $a_k = 0$  otherwise, the distribution of the number in the buffer is a truncated geometric distribution, with parameter  $(1 - \alpha)/\alpha$ . The computations were performed in double precision arithmetic on an IBM computer, in which numbers are represented to about 17 decimal digits and where chopping arithmetic (rather than rounding) is used. The combination of large buffer size and  $(1 - \alpha)/\alpha$  close to unity is least conducive to accuracy. However, even with a buffer size of 1900 and  $(1 - \alpha)/\alpha = 0.99$ , results accurate to seven and eight decimal places were obtained. A CPU time of 26 seconds on an IBM 3081/K was needed to produce the solution for a buffer size of one thousand.

By way of contrast, 126 seconds of execution time was required for a 5K VIC-20\* computer, programmed in BASIC, to give the solution for a buffer size of 48. This was the largest size permitted by the small memory of this small machine. The VIC-20 computer gave an accuracy of eight decimal places on the various examples that were tried. Levinson's method thus makes practical the solution of the equilibrium equations, for moderate buffer sizes, even on small personal computers.

#### V. THE CYCLOSTATIONARY CASE

The previous sections discussed stationary systems in which the distributions of  $X_n$  and  $Y_n$  are independent of  $n$ . However, in some situations the distribution of  $X_n$  varies cyclically with  $n$ . An example is where the arrivals come from a set of data sources, having differing characteristics, which are polled cyclically. This section discusses systems described by the same fundamental equation as before, (1), but where the distribution of  $X_n$  varies cyclically. This distribution is denoted by

$$a_{kn} = \text{probability}[X_n = k].$$

---

\* Registered trademark of Commodore Electronics Ltd.

With the system in equilibrium, the distribution of the number in the buffer will also vary cyclically. The cumulative distribution of the number in the buffer at the end of the  $n$ th interval is denoted by

$$q_{kn} = \text{probability}[Y_n \leq k].$$

Because the distributions vary cyclically, it is sufficient to consider just one period, e.g.,  $n = 0, 1, \dots, N - 1$ .

With the system in cyclic equilibrium, the distribution of the number in the buffer at the end of the  $(n + 1)$ th interval can be expressed in terms of the distribution at the end of the  $n$ th,  $n$  being computed modulo  $N$ .

$$q_{k,n+1} = \begin{cases} 0, & k < 0 \\ \sum_{m=-\infty}^{\infty} q_{mn} a_{k-m,n+1}, & 0 \leq k < K \\ 1, & k \geq K. \end{cases}$$

This leads to a set of equations, for the  $KN$  unknown probabilities, whose matrix has block Toeplitz form; i.e., when partitioned into  $N \times N$  submatrices, the submatrices on any given diagonal are all equal. The equations can be written

$$\begin{bmatrix} \mathbf{d}_0 & \mathbf{d}_1 & \cdots & \mathbf{d}_{K-1} \\ \mathbf{d}_{-1} & \mathbf{d}_0 & \cdots & \mathbf{d}_{K-2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{d}_{-(K-1)} & \mathbf{d}_{-(K-2)} & \cdots & \mathbf{d}_0 \end{bmatrix} \begin{bmatrix} \mathbf{q}_0 \\ \mathbf{q}_1 \\ \vdots \\ \mathbf{q}_{K-1} \end{bmatrix} = \begin{bmatrix} \mathbf{b}_0 \\ \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_{K-1} \end{bmatrix},$$

where the submatrices and subvectors are given by

$$\mathbf{q}_k = \begin{bmatrix} q_{k,0} \\ q_{k,1} \\ \vdots \\ q_{k,N-1} \end{bmatrix} \quad \mathbf{b}_k = \begin{bmatrix} \sum_{m=K-k}^{\infty} a_{-m,0} \\ \sum_{m=K-k}^{\infty} a_{-m,1} \\ \vdots \\ \sum_{m=K-k}^{\infty} a_{-m,N-1} \end{bmatrix},$$

$$\mathbf{d}_k = \begin{bmatrix} \delta_k & 0 & \cdots & 0 & -a_{-k,0} \\ -a_{-k,1} & \delta_k & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \delta_k & 0 \\ 0 & 0 & \cdots & -a_{-k,N-1} & \delta_k \end{bmatrix},$$

where

$$\delta_k = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0. \end{cases}$$

Block Toeplitz equations occur in the formulation of multichannel optimal linear filters. Levinson's method was extended to the block Toeplitz case for computing the coefficients in such filters.<sup>9,10</sup> For block Toeplitz systems the method follows the outline given in Section III, but with operations on scalar elements replaced by the corresponding operations on subvectors and submatrices and with the matrices corresponding to the  $c$ 's in (6) computed via the solution of linear equations. Programs for computing the coefficients of optimal multichannel linear filters require modification before they can be used for the cyclostationary buffer problem. This is because they are normally written to exploit the symmetry  $\mathbf{d}_k = \mathbf{d}_k^T$  that exists in the multichannel filtering problem but not, generally, in the cyclostationary buffer problem.

## VI. CYCLOSTATIONARY REPRESENTATION OF STATIONARY QUEUES

The description of a finite buffer by (2) corresponds more closely to reality, in many cases, than its description by (1). As mentioned before, with (2), the number waiting never exceeds the buffer size. However, the equilibrium equations for the distribution of  $Y'_n$  in (2) do not have Toeplitz form, and so Levinson's method cannot be applied directly to their solution. However, Levinson's method can be applied when it is recognized that a stationary system described by (2) can be represented by a cyclostationary system, described by (1), having a period of two.

$X_n$ , the net arrival size in (1), is identified, alternately, with the arrival size and the service size in (2). That is,

$$X_n = \begin{cases} V_m, & n = 2m \\ W_m, & n = 2m + 1. \end{cases}$$

The distribution of  $X_n$  alternates between the stationary distributions

of  $V_m$  and  $W_m$ . The equilibrium distribution of  $Y_n$  varies cyclically and, for even  $n$ , is identical to the stationary distribution of  $Y'_m$ . Levinson's method can thus be used to obtain the distribution  $Y'_m$  by solving the block Toeplitz equations corresponding to (2) and retaining just the distribution for even-indexed  $Y_n$ .

In most cases, the behaviour of a system described by (1) is very similar to one described by (2). They differ appreciably only in cases where there is a significant probability of buffer overflow due to a large value of  $V_m$ , which is accompanied by a similar value of  $W_m$ . As an example of a case where the behaviour of the systems differs appreciably, results were computed for a buffer size of five where  $V_n$  and  $W_n$  are both geometrically distributed, with parameter 0.5. The computed distributions are as follows:

$k$	$q_k$ , system described by (1)	$q_k$ , system described by (2)
0	0.2500	0.1429
1	0.3750	0.2857
2	0.5000	0.4286
3	0.6250	0.5714
4	0.7500	0.7143
5	1	1

## VII. OTHER APPLICATIONS

It is often necessary in telecommunication system analysis to calculate the distribution of the waiting time in a queueing system. A stationary G/G/1 queue—i.e., where the distributions of interarrival times and service times are arbitrary—often forms a suitable model of the system to be analyzed. The fundamental equation for  $T_n$ , the waiting time of the  $n$ th customer in a G/G/1 queue, is

$$T_n = (T_{n-1} + U_{n-1})^+, \quad (12)$$

where  $U_{n-1}$  is the difference between the service time for the  $(n - 1)$ th customer and the interarrival time between the  $n$ th and the  $(n - 1)$ th customer. When  $K$  in (1) is sufficiently large that the probability of  $Y_{n-1} + X_n$  exceeding  $K$  is negligible, (1) has, effectively, the same form as (2). A method for computing the distribution of the number in a system described by (1) can, therefore, also be used to compute the distribution of waiting times in a G/G/1 queue with discrete service and interarrival times. Some error is incurred by the use of finite  $K$  but, by choosing  $K$  sufficiently large, this error can be made small compared with arithmetic error.

Many switching systems use a processor whose time is shared among different tasks. The tasks have various requirements in terms of the

delay that can be tolerated between the demand for the execution of a task and the completion of its execution. A commonly used scheme for controlling the execution of such tasks is a clocked schedule.<sup>12</sup> The tasks have different priorities, and the schedule specifies which combination of tasks is to be executed in each interval between clock instants. Fredericks<sup>12</sup> has shown that the resulting multiclass priority queueing system is equivalent to a set of single input G/G/1 queueing systems. In each of these queues the distribution of the quantity corresponding to  $U_n$  in (12) varies periodically. Fredericks approximated the periodically varying queue by a stationary queue. This gives accurate results for long delays but gives no information about the short time variation of delays that results from the varying nature of the system. Because of the similarity between (1) and (12), Levinson's method for solving block Toeplitz systems permits the computation of the cyclically varying distribution of the waiting time in a discrete G/G/1 queue in which the distribution of  $U_n$  varies periodically with  $n$ . The method thus provides a means for the analysis of short-term delays in clocked schedule systems, as well as longer delays.

### VIII. CONCLUSION

Levinson's method provides a practical means for computing the equilibrium distribution of the number in a finite buffer, by explicit solution of the equilibrium equations. On a mainframe computer, the method is practical even with quite large buffer sizes. With moderate buffer sizes, it becomes feasible to do the computations even on a small personal computer. The method also provides a practical means for computing the waiting-time distribution in discrete time G/G/1 queues.

### IX. ACKNOWLEDGMENT

I am grateful to B. S. Gotz for carefully reading an earlier version of this paper.

### REFERENCES

1. N. M. Dor, "Guide to the Length of Buffer Storage Required for Random (Poisson) Input and Constant Output Rates," *IEEE Trans. Elect. Comput.*, *EC-16* (October 1967), pp. 683-4.
2. W. W. Chu, "Buffer Behaviour for Batch Poisson Arrivals and Single Constant Output," *IEEE Trans. Commun. Technol.*, *COM-18* (October 1970), pp. 613-8.
3. W. W. Chu, "Buffer Behaviour for Poisson Arrivals and Multiple Synchronous Constant Outputs," *IEEE Trans. Comput.*, *C-19* (June 1970), pp. 530-4.
4. H. Rudin, Jr., "Performance of Simple Multiplexer-Concentrators for Data Communication," *IEEE Trans. Commun. Technol.*, *COM-19* (April 1971), pp. 178-87.
5. T. P. Bagchi and J. G. C. Templeton, *Numerical Methods in Markov Chains and Bulk Queues*, New York: Springer-Verlag, 1972.
6. K. Geihs and H. Kobayashi, "Bounds on Buffer Overflow Probabilities in Communication Systems," *Performance Eval.*, *2* (October 1982), pp. 149-60.

7. R. R. Anderson, G. J. Foschini, and B. Gopinath, "A Queueing Model for a Hybrid Data Multiplexer," *B.S.T.J.*, 58 (February 1979), pp. 279-300.
8. N. Levinson, "The Weiner RMS (Root Mean Square) Error Criterion in Filter Design and Prediction," *J. Math. Phys.*, 25 (1947), pp. 261-78.
9. E. A. Robinson, *Multichannel Time Series Analysis with Digital Computer Programs*, San Francisco: Holden-Day, 1967.
10. R. A. Wiggins and E. A. Robinson, "Recursive Solution to the Multichannel Filtering Problem," *J. Geophys. Res.*, 70 (April 1965), pp. 1885-91.
11. G. Cybenko, "The Numerical Stability of the Levinson-Durbin Algorithm for Toeplitz Systems of Equations," *SIAM J. Sci. Statist. Comput.*, 1 (September 1980), pp. 303-19.
12. A. A. Fredericks, "Analysis of a Class of Schedules with Real Time Applications," *Performance of Computer Systems*, M. Arrato, A. Butrimenko, and E. Gelenbe, Eds., Amsterdam, New York, and Oxford: North-Holland, 1979, pp. 201-16.

## AUTHOR

**Martin H. Ackroyd**, B.Sc. (Electronic and Electrical Engineering), 1966, Birmingham University, U.K.; Ph.D. (Electronic and Electrical Engineering), 1970, Loughborough University of Technology, U.K.; Loughborough University of Technology, U.K., 1969-1973; Central Research Laboratories of EMI Limited, 1973-1975; Aston University, U.K., 1975-1982; AT&T Bell Laboratories, 1982-1984; Essex University, U.K., 1984—. At Loughborough University of Technology, Mr. Ackroyd was a Lecturer in Signal Processing. At Central Research Laboratories of EMI Limited, he was the Head of Cognitive Systems Group. Mr. Ackroyd was a Reader in Telecommunications at Aston University. At AT&T Bell Laboratories, he was a Member of Technical Staff in the Performance Analysis department of the Network Planning division. Mr. Ackroyd is now the Professor of Telecommunication and Information Systems in the Department of Electrical Engineering Science at Essex University. His research interests include the numerical solution of problems in information system performance engineering. Previously he has done research in other areas, including digital signal and image processing. Senior member, IEEE.

# Distortion Analysis on Measuring the Impulse Response of a System Using a Crosscorrelation Method

By N. BENVENUTO\*

(Manuscript received June 14, 1984)

In this paper we analyze the characteristics of a crosscorrelation method for measuring the impulse response of an unknown system. This analysis may also be of some interest in the area of synchronization of pseudonoise systems. In the first part of the paper a closed-form expression for the main component of the probing pulse is derived. This result may be used to quickly estimate the effect of various system parameters on the measurement's accuracy. In the second part results are presented for two system configurations.

## I. INTRODUCTION

Among the various methods for measuring the impulse response of a linear time-invariant system, a simple and effective method is to excite the system with a white noise-like input and crosscorrelate the output with the input. Actually, instead of a random white noise, a pseudonoise waveform may be used as an input.<sup>1</sup> Such measurement systems are designed to produce the same effect as exciting the unknown system with a probing pulse having a flat spectrum over the entire frequency band of the unknown system.

For these measurement systems the actual implementation is of fundamental importance, and various schemes have been proposed depending upon the system to be measured; among others we recall

---

\* AT&T Bell Laboratories.

---

Copyright © 1984 AT&T. Photo reproduction for noncommercial use is permitted without payment of royalty provided that each reproduction is done without alteration and that the Journal reference and copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free by computer-based and other information-service systems without further permission. Permission to reproduce or republish any other portion of this paper must be obtained from the Editor.

Refs. 2 and 3 (see also the extensive bibliography in Ref. 1). This paper presents an analysis of an implementation introduced by D. Cox,<sup>2</sup> which is particularly attractive in radio environments when a high-resolution (i.e., wide bandwidth) probing pulse is required. The main advantage of this approach is in the hardware implementation, since the measurement system requires very little control logic (see Ref. 2).

Information about the effect of actual implementations on the accuracy of this measurement technique is scarce. With this analysis simple formulas are derived that display the practical effects of the various system parameters, thus offering new tools to a system designer.

Section II of this paper reviews some basic concepts on measuring the impulse response of a system by using crosscorrelation methods, and illustrates Cox's methods from a heuristic point of view. This is analyzed in detail in Section III, where a closed form for the probing pulse is derived. In Section IV the analysis is particularized to different system configurations. Incidentally, we note that this analysis may also be of some interest in the area of synchronization of pseudonoise systems.<sup>5</sup>

## II. PRELIMINARIES

The pseudonoise waveform that is used to excite the system is usually obtained by amplitude modulating a pulse  $g(t)$  with a pseudorandom sequence  $\{a_n\}$ ,  $n = 0, 1, \dots, M - 1$ ;  $a_n \in \{-1, 1\}$  so that the autocorrelation sequence

$$c_a(k) \triangleq \frac{1}{M} \sum_{i=0}^{M-1} a_i a_{(i-k) \bmod M} \quad (1)$$

has the form

$$c_a(k) = \begin{cases} 1, & \text{for } (k) \bmod M = 0 \\ -\frac{1}{M} & \text{otherwise.} \end{cases} \quad (2)$$

For simplicity of implementation we consider the case where  $\{a_n\}$  is generated by using a shift register with feedback.<sup>4</sup> A block diagram of the entire system to measure the impulse response of an unknown system with impulse response  $\alpha(t)$  is illustrated in Fig. 1. If the output of the shift register with clock frequency  $f_c = 1/t_0$  is used to amplitude modulate a pulse  $g(t)$ , the output of the pseudorandom sequence generator  $r(t)$  is just the repetition with period  $T = Mt_0$  of the pseudorandom waveform  $s(t)$  given by

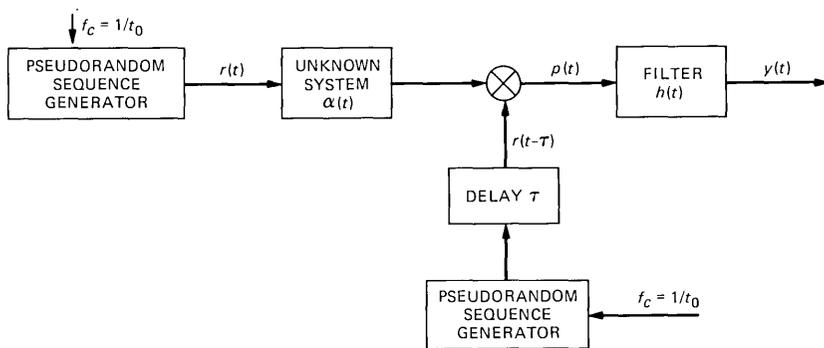


Fig. 1—Schematic diagram of the measuring system.

$$s(t) = \sum_{i=0}^{M-1} a_i g(t - it_0). \quad (3)$$

Thus

$$r(t) = \sum_{i=0}^{\infty} a_i g(t - it_0) \quad (4)$$

with the understanding that  $a_i = a_{i \bmod M}$ . The autocorrelation function of  $r(t)$ ,

$$\begin{aligned} c_r(t) &\triangleq \lim_{t_1 \rightarrow \infty} \frac{1}{2t_1} \int_{-t_1}^{t_1} r(\eta) r(\eta - t) d\eta \\ &= \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} r(\eta) r(\eta - t) d\eta, \end{aligned} \quad (5)$$

is periodic with period  $T$ . If  $g(t)$  has finite support, say  $g(t) = 0$  for  $t < 0$  and  $t > t_0$ , then  $c_r(t)$  has a simple expression given by

$$c_r(t) = \sum_{k=0}^M c_a(k) c_g(t - kt_0), \quad \text{for } 0 \leq t \leq T, \quad (6)$$

where

$$c_g(t) = \frac{1}{t_0} \int_0^{t_0} g(\eta) g(\eta - t) d\eta. \quad (7)$$

In the particular case where the output of the shift register feeds directly the "unknown system," we have that\*

\* The function  $\text{rect}(\cdot)$  is defined as

$$\text{rect} \frac{t-a}{b} = \begin{cases} 1, & \text{for } a - \frac{b}{2} \leq t \leq a + \frac{b}{2} \\ 0, & \text{elsewhere.} \end{cases}$$

$$g(t) = \text{rect} \frac{t - t_0/2}{t_0}. \quad (8)$$

Then

$$c_g(t) = \left(1 - \frac{|t|}{t_0}\right) \text{rect} \frac{t}{2t_0} \quad (9)$$

and  $c_r(t)$  has the well-known expression:

$$c_r(t) = -\frac{1}{M} + \left(1 + \frac{1}{M}\right) \left(1 - \frac{|t|}{t_0}\right) \text{rect} \frac{t}{2t_0} \quad \text{for } |t| \leq \frac{T}{2}, \quad (10)$$

as shown in Fig. 2. The “temporal resolution” of this pulse is seen to be on the order of  $t_0$ . Thus, to improve resolution, one might increase the shift register clock frequency.

For some applications it might be easier to change the pulse shape  $g(\cdot)$  rather than the shift register clock frequency; Fig. 3 shows  $c_r(t)$  when  $g(t) = \sqrt{2} \text{rect}[(t - t_0/4)/(t_0/2)]$ .

Referring again to Fig. 1, the output of the “unknown system” is multiplied by a time-shifted version of the input  $r(t - \tau)$  to give  $p(t)$ . This signal is then filtered with a low-pass filter having impulse response  $h(t)$ . If  $h(t)$  is an ideal integrator between 0 and  $T$ , i.e.,

$$h(t) = \frac{1}{T} \text{rect} \frac{t - \frac{T}{2}}{T}, \quad (11)$$

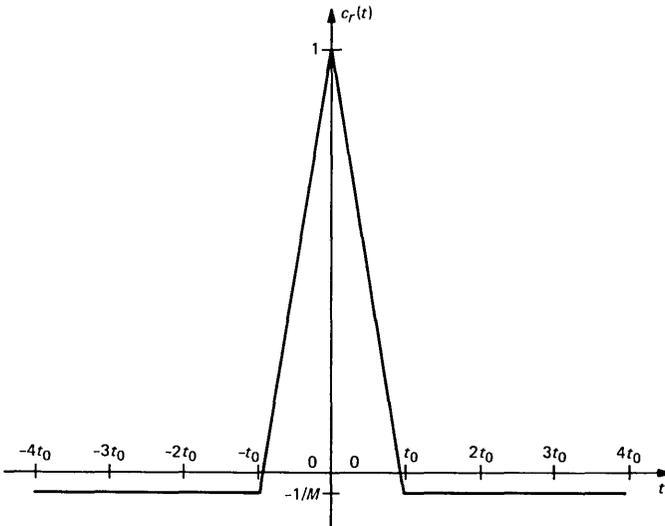


Fig. 2—Autocorrelation function for  $g(t) = \text{rect}[(t - t_0/2)/(t_0)]$ .

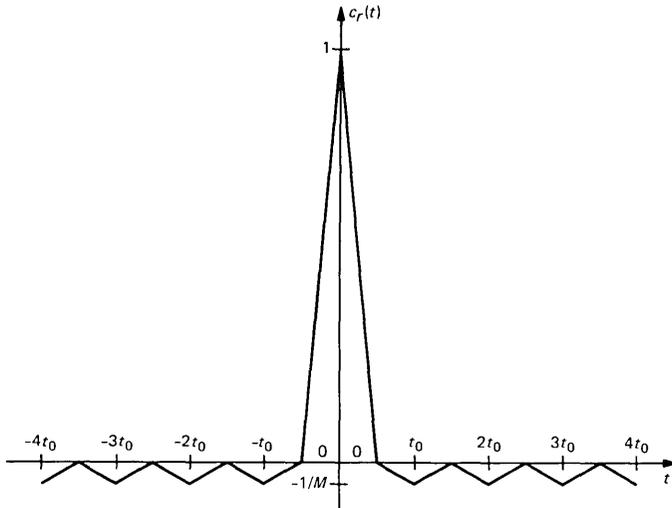


Fig. 3—Autocorrelation function for  $g(t) = \sqrt{2} \text{rect}[(t - t_0/4)/(t_0/2)]$ .

then

$$\begin{aligned}
 y_\tau \triangleq y(t) &= \frac{1}{T} \int_{t-T}^t p(\eta) d\eta \\
 &= \frac{1}{T} \int_{t-T}^t \int_0^{+\infty} [\alpha(\zeta) r(\eta - \zeta) d\zeta] [r(\eta - \tau)] d\eta \\
 &= \int_0^{+\infty} \alpha(\zeta) c_r(\tau - \zeta) d\zeta, \tag{12}
 \end{aligned}$$

i.e., the output has a constant value  $y_\tau$  equal to the convolution of  $\alpha(\cdot)$  with  $c_r(\cdot)$  evaluated at  $\tau$ .

Assuming  $f_c$  is larger than the highest frequency component of  $\alpha(\cdot)$ , and  $M$  is large enough, we have that  $y_\tau$  is approximately proportional to  $\alpha(\tau)$ .<sup>3</sup> From (12),  $c_r(\cdot)$  will be referred to as the probing pulse in the measurement system. A straightforward implementation of the system in Fig. 1 has been presented in Ref. 3.

An alternative configuration that has the benefit of a simple implementation has been proposed by Cox.<sup>2</sup> The baseband equivalent of Cox's carrier system scheme is shown in Fig. 4. The only difference with respect to the system of Fig. 1 is that the output of the "unknown system" is multiplied by a pseudorandom waveform that is identical both in character and in timing with the transmitter sequence, but it is produced by a different clock frequency  $f'_c$ . As in Ref. 2  $f'_c$  is related to  $f_c$  by

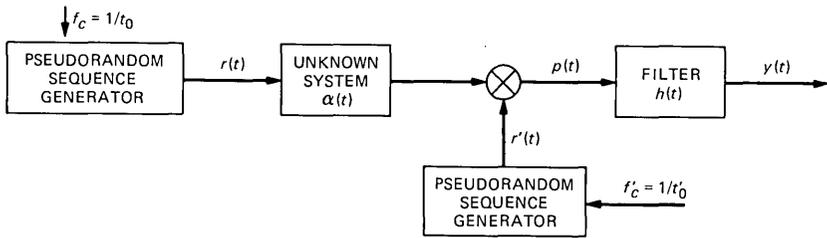


Fig. 4—Schematic diagram of the measuring system using Cox's method.

$$f'_c = f_c \left( 1 - \frac{1}{K} \right) \quad (13)$$

and  $K$  will be a parameter of our system.

In general, consider the function

$$c_{r'r}(\tau) \triangleq \frac{1}{T} \int_0^T r'(\eta)r(\eta - \tau)d\eta, \quad (14)$$

where  $\tau$  is the delay at time  $t = 0$  between the two sequences. As time elapses the delay decreases by the amount  $(t/t'_0)(t'_0 - t_0) = t/K$ , so that

$$\frac{1}{T} \int_{t-T}^t r'(\eta)r(\eta - \tau)d\eta \simeq c_{r'r} \left( \tau - \frac{t-T}{K} \right), \quad \text{for } t \geq T. \quad (15)$$

In (15) the approximation is due to the fact that the integral depends on the symbols  $\{a_n\}$  under integration. Now, if  $K$  is large enough we can assume that

$$c_{r'r}(t) = c_r(t). \quad (16)$$

In this case, for an integrate and dump filter, we have that

$$\begin{aligned} y(t) &= \int_0^{+\infty} \alpha(\zeta)c_{r'r} \left( \zeta - \frac{t-T}{K} \right) d\zeta \\ &\simeq \int_0^{+\infty} \alpha(\zeta)c_r \left( \frac{t-T}{K} - \zeta \right) d\zeta \end{aligned} \quad (17)$$

or

$$y(Kt' + T) \simeq \int_0^{\infty} \alpha(\zeta)c_r(t' - \zeta)d\zeta. \quad (18)$$

Note that the right-hand side of (18) coincides with (12). Thus if  $K$  is large enough such that assumptions (15) and (16) hold, this measurement system is equivalent to the system of Fig. 1. Unfortunately,  $K$  cannot be set arbitrarily large, as this would give rise to two conflicts.

First, the duration of the experiment [see eq. (18)] would become very long, with the consequence that the assumption of a time-invariant system may no longer be true. Second, a very large  $K$  would imply a costly frequency synthesizer from which  $f'_c$  is derived.<sup>2</sup>

In the next section we analyze the conditions under which assumptions (15) and (16) may be considered valid. We consider also the effect of using a low-pass filter as  $h(t)$  rather than an ideal integrator as in Ref. 2. In general terms we shall see that the probing signal can be written as a sum of two terms: a useful term plus noise. A closed-form expression for the useful term is derived for a general filter impulse response  $h(t)$ . Also, the dependence of the noise term upon the system parameters will be determined.

### III. PROBING PULSE DERIVATION

To determine the actual probing pulse for the system of Fig. 4, we can consider the output function of the block diagram in Fig. 5. We shall determine the conditions on  $K$  and filter impulse response  $h(t)$  such that  $q(\cdot)$  approximates  $c_r(\cdot)$ .

In Appendix A, a general expression for  $q(t)$  is derived for an arbitrary pulse shape  $g(t)$  (which for simplicity we assume to have finite support  $t_0$ ). It is assumed that  $h(t)$  has finite support, i.e.,

$$h(t) = 0, \quad \text{for } t < 0, \quad \text{and } t > t_h \quad (19)$$

with

$$t_h \leq K(M - 2)t_0. \quad (20)$$

As we shall see, this assumption does not impose any serious limitation on the filter type for practical purposes. Typically,  $t_h$  will be much less than  $K(M - 2)t_0$ , but the analysis applies as long as (20) is satisfied. Furthermore, we normalize  $h(t)$  such that it has unit area. Let

$$h'(t) \triangleq Kh(tK) \quad (21)$$

and

$$d(t) = \int_{-t_0}^{t_0} c_g(\eta)h'(\eta - t)d\eta. \quad (22)$$

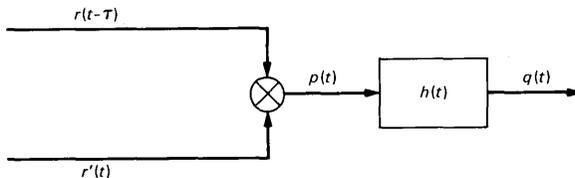


Fig. 5—Equivalent system to derive probing pulse.

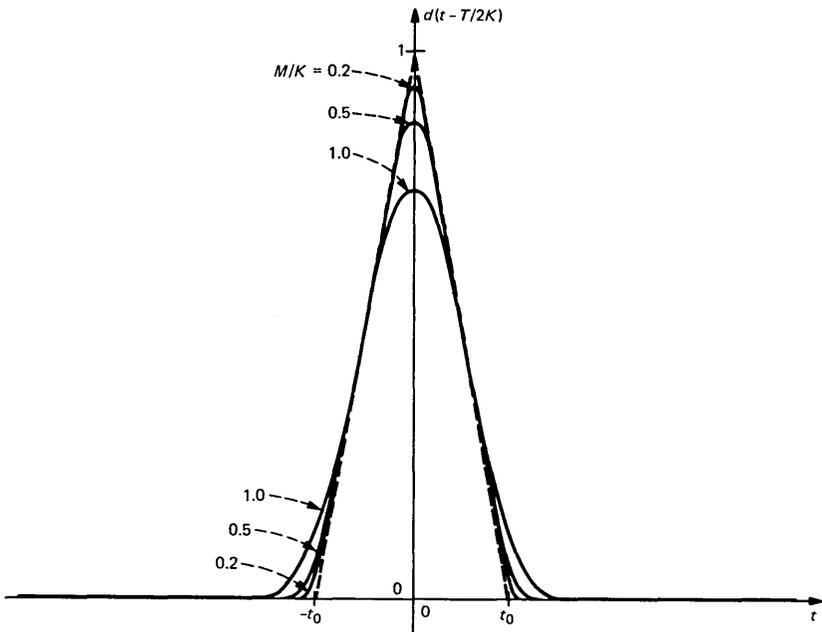


Fig. 6—The  $d(t)$  pulse for an integrate-and-dump filter.

In Appendix A it is shown that  $q(t)$  can be written as the sum of two terms:

$$q(t) = q_0(t) + q_n(t). \quad (23)$$

In particular, when  $g(t)$  is given by (8),  $q_0(t)$  has the form

$$q_0(t) = -\frac{1}{M} + \left(\frac{1}{M} + 1\right) d\left(\tau - \frac{t}{K}\right),$$

$$\text{for } -T + t_0 \leq \tau - \frac{t}{K} \leq t_0. \quad (24)$$

For a low-pass filter with  $t_h \ll KT$ , if  $K \gg 1$  we have from (19), (21), and (22) that  $d(t) \approx c_g(t)$ . Comparing to (10) we see that  $q_0(t)$  coincides with  $c_r[\tau - (t/K)]$ . Moreover, if the bandwidth of the filter is small enough,\* it is seen that the general term of the series (69) which defines  $q_n(t)$ ,  $\mathcal{A}(m) \cdot G(m/T) \cdot H[(l/T) - (m)/(KT)]$ , is very small, with the result that  $q_n(t)$  is negligible if  $K \gg M$  ( $G(f)$  and  $H(f)$  are the Fourier transforms of  $g(t)$  and  $h(t)$ , respectively). Thus, if the band-

\* In practice, this condition is not in conflict with (19) for  $K$ , a large number; i.e., we require for a general filter shape  $T \ll t_h \ll KT$ . For special filter shapes  $t_h$  may be as small as  $T$ ; e.g., the ideal integrator discussed in the next section.

width of  $H(f)$  is small enough, and  $K$  is large enough,  $q(t)$  coincides with  $c_r[\tau - (t/K)]$  as desired. A more quantitative analysis will be given in the next section for two particular impulse responses  $h(t)$ .

#### IV. EXAMPLES

For  $g(t)$  given by (8), we shall evaluate  $q_0(t)$  and  $q(t)$  for two extreme cases: first for an ideal integrate-and-dump filter, then for a single-pole Resistance-Capacitance (RC) filter.

##### 4.1 Integrate and dump filter

Here  $h(t)$  is given by (11), thus

$$h'(t) = \frac{K}{T} \operatorname{rect} \frac{t - T/(2K)}{T/K} \quad (25)$$

and for  $K > M$  we get

$$d\left(t - \frac{T}{2K}\right) = \begin{cases} \frac{K}{M} \left( \frac{M}{K} \left( 1 - \frac{M}{4K} \right) - \left( \frac{t}{t_0} \right)^2 \right), \\ \quad \text{for } 0 \leq |t| \leq t_0 \frac{M}{2K} \\ \left( 1 - \frac{t}{t_0} \right), \\ \quad \text{for } t_0 \frac{M}{2K} \leq |t| \leq t_0 \left( 1 - \frac{M}{2K} \right) \\ \frac{K}{2M} \left( 1 + \frac{M}{2K} - \frac{t}{t_0} \right)^2, \\ \quad \text{for } t_0 \left( 1 - \frac{M}{2K} \right) \leq |t| \leq t_0 \left( 1 + \frac{M}{2K} \right) \\ 0, \\ \quad \text{for } |t| \geq t_0 \left( 1 + \frac{M}{2K} \right). \end{cases} \quad (26)$$

Plotted in Fig. 6 is  $d(\cdot)$  for various values of  $M/K$ . We note that  $d(\cdot)$  depends only on the ratio  $M/K$  and already for  $M/K \leq 0.5$  the distortion introduced by  $h'(\cdot)$  is very small. For a comparison with the ideal case we also present  $c_g(t)$  as given by (9) (dashed curve in Fig. 6).

With regard to the noisy component  $q_n(t)$ , according to the series (69) the main contributions to  $q_n(t)$  are for  $l = \pm 1$ . In Fig. 7 we plot

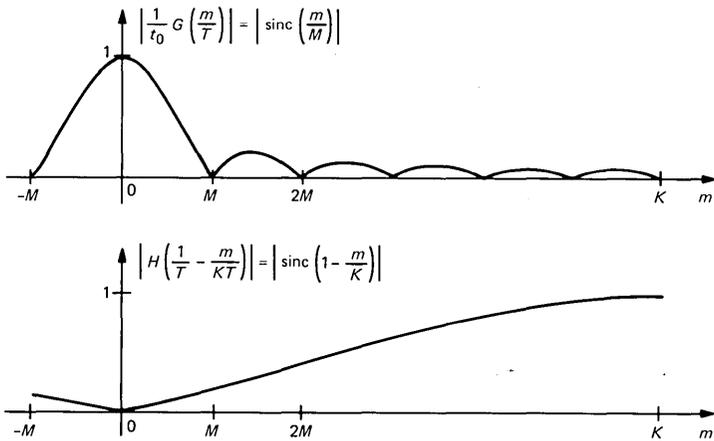


Fig. 7—Sequence profiles of noisy term  $q_n(t)$ .

the envelopes of the sequences  $(1/t_0)|G(m/T)|$  and  $|H[(1/T) - (m)/(KT)]|$  in (69) for  $K/M = 6$ . Since from (52)

$$|\mathcal{A}(m)| \leq \left( \frac{1}{M} + \frac{1}{M^2} \right)^{1/2} \quad (27)$$

we can say that although  $q_n(t)$  depends on  $M$ , it can be made negligible by increasing  $K/M$ .

Thus far, using general properties of a pseudonoise sequence, we have determined the exact shape of  $q_0(t)$  and got a rule of thumb of how to make  $q_n(t)$  negligible. However, if an exact determination of  $q_n(t)$ , and therefore  $q(t)$ , is desired, we need to know the sequence  $\{a_n\}$ ,  $n = 0, 1, \dots, M - 1$ , explicitly. In the following an exact expression for  $q(t)$  is derived as a function of the system parameters. The output of the filter  $h(t)$  will be computed only for times corresponding to filter integration intervals that begin at integral multiples of the receiver sequence  $r'(\cdot)$  subintervals  $t'_0$ . Since the width of  $q(t)$  is approximately  $2Kt_0$ , the computed values will be sufficiently closely spaced to completely characterize the output.

Here we consider only the case  $K > M$ ; similar expressions hold for  $K \leq M$ . Let

$$g'(t) = \text{rect} \frac{t - t'_0/2}{t'_0} \quad (28)$$

[modulating pulse for  $r'(t)$ ] and

$$g''(t) = \text{rect} \frac{t - t''_0/2}{t''_0}, \quad (29)$$

where

$$\begin{aligned} t_0'' &= T - (T' - t_0') \\ &= t_0 \frac{K - M}{K - 1}. \end{aligned} \quad (30)$$

Thus  $g''(t)$  is the pulse that must be added to  $M - 1$   $g'(t)$  pulses to provide a total duration of  $T$ . Moreover,  $c_{gg'}(t)$  and  $c_{gg''}(t)$  denote correlation functions, defined as

$$c_{ab}(t) = \frac{1}{T} \int_{-\infty}^{+\infty} a(\eta)b(\eta - t)d\eta. \quad (31)$$

We introduce the periodic function  $z_I(\eta)$  defined as

$$z_I(\eta) = \frac{1}{T} \int_0^T r'(\zeta)r(\zeta - \eta)d\zeta. \quad (32)$$

The subscript  $I$  indicates that the sequence  $\{a_n\}$  in  $r'(\zeta)$  starts at  $\zeta = 0$  with the symbol  $a_I$ . Substituting (4) for  $r(\cdot)$  and similarly for  $r'(\cdot)$  in (32),  $z_I(\eta)$  can be written, for  $0 < \eta \leq T$ , as

$$\begin{aligned} z_I(Lt_0 - \zeta) &= \sum_{i=0}^M \left\{ \sum_{j=0}^{M-2} a_{i-L+I} a_{j+I} c_{gg'}(jt_0 - it_0 + \zeta) \right. \\ &\quad \left. + a_{i-L+I} a_{M-1+I} c_{gg''}((M-1)t_0 - it_0 + \zeta) \right\}, \\ L &= 1, 2, \dots, M \quad \text{and} \quad 0 < \zeta \leq t_0. \end{aligned} \quad (33)$$

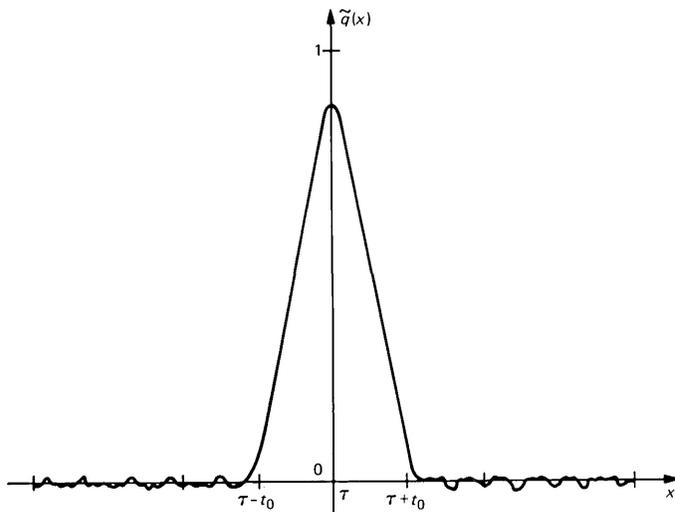


Fig. 8—Probing pulse for an integrate-and-dump filter:  $M = 511$ , and  $K = 1000$ .

Since

$$c_{gg'} \left( t - \frac{t'_0 - t_0}{2} \right) = \begin{cases} \frac{t_0}{T}, & \text{for } |t| \leq \frac{t'_0 - t_0}{2} \\ \frac{1}{T} \left( \frac{t'_0 + t_0}{2} - |t| \right), & \text{for } \frac{t'_0 - t_0}{2} \leq |t| \leq \frac{t'_0 + t_0}{2} \\ 0, & \text{for } |t| \geq \frac{t'_0 + t_0}{2} \end{cases} \quad (34)$$

and

$$c_{gg''} \left( t - \frac{t_0 - t''_0}{2} \right) = \begin{cases} \frac{t''_0}{T}, & \text{for } |t| \leq \frac{t_0 - t''_0}{2} \\ \frac{1}{T} \left( \frac{t_0 + t''_0}{2} - |t| \right), & \text{for } \frac{t_0 - t''_0}{2} \leq |t| \leq \frac{t_0 + t''_0}{2} \\ 0, & \text{for } |t| \geq \frac{t_0 + t''_0}{2} \end{cases} \quad (35)$$

have finite support, we can limit the summation in  $j$  to just a few terms, namely,

$$z_I(Lt_0 - \zeta) = \sum_{i=0}^M \left\{ \sum_{j=J_m}^{J_M} a_{i-L+I} a_{j+I} c_{gg'}(jt'_0 - it_0 + \zeta) + a_{i-L+I} a_{M-1+I} c_{gg''}((M-1)t'_0 - it_0 + \zeta) \right\}, \quad (36)$$

where

$$J_m = \min \left( \left[ \frac{it_0 - \zeta}{t'_0} - 1 \right]_C, M - 2 \right) \quad (37)$$

$$J_M = \min \left( \left[ \frac{(i+1)t_0 - \zeta}{t'_0} \right]_F, M - 2 \right). \quad (38)$$

$[\cdot]_C$  and  $[\cdot]_F$  are the "ceiling" and "floor" functions, respectively. Finally, for  $t = T + It'_0$ ,  $I$  an integer, we have

$$q(t) = z_I \left( \tau - \frac{t - T}{K} \right). \quad (39)$$

Plots of  $q(t)$  shifted in time by  $-[(T)/(2K)]$  [see eq. (26)] and scaled by the factor  $K$  are reported in Figs. 8 and 9 for  $M = 511$  and two values of  $K$ . Specifically for  $x$  a multiple of  $t'_0/K$  we plot

$$\begin{aligned}\tilde{q}(x) &= q \left[ K \left( x + \frac{T}{2K} \right) \right] \\ &= z_I \left( \tau + \frac{T}{2K} - x \right), \quad I = \frac{K}{t'_0} x.\end{aligned}\quad (40)$$

From these figures we can see that the noisy term is significant for  $M/K = 0.511$ . The maximum value of  $|q_n(t)|$  in Fig. 8 is about  $-33$  dB. However, for  $M/K = 0.128$  it is reduced to  $-43$  dB (see Fig. 9).

#### 4.2 Single-pole RC filter

Here, the filter impulse response is

$$h(t) = \begin{cases} \frac{1}{RC} e^{-\frac{t}{RC}}, & t \geq 0 \\ 0, & t < 0. \end{cases}\quad (41)$$

Note that in order to guarantee condition (19),  $(KT)/(RC)$  must be much greater than 1, as will be the case. Let

$$\beta = \frac{RC}{K \cdot t_0}.\quad (42)$$

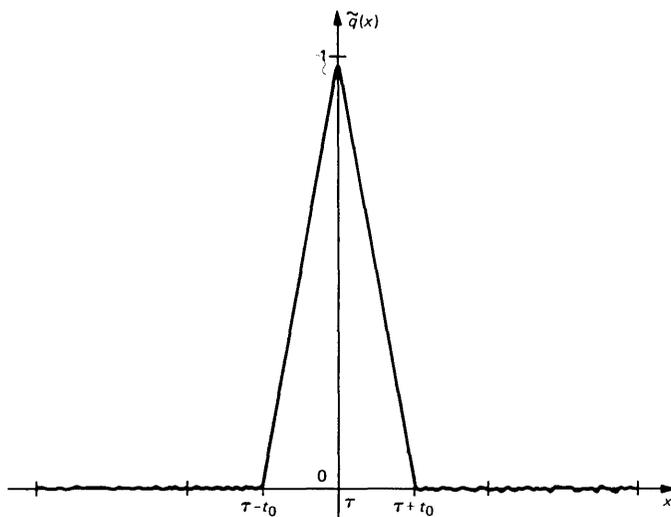


Fig. 9—Probing pulse for an integrate-and-dump filter:  $M = 511$ , and  $K = 4000$ .

Then, from (22),  $d(t)$  has the form

$$d(t) = \begin{cases} 0, & t \geq t_0 \\ 1 - \frac{1}{t_0} + \beta \left( e^{\frac{t/t_0-1}{\beta}} - 1 \right), & 0 \leq t \leq t_0 \\ 1 + \frac{t}{t_0} + \beta \left( 1 - 2e^{\frac{t/t_0}{\beta}} + e^{\frac{t/t_0-1}{\beta}} \right), & -t_0 \leq t \leq 0 \\ \beta e^{\frac{t/t_0}{\beta}} \left( e^{\frac{1}{\beta}} - 2 - e^{-\frac{1}{\beta}} \right), & t \leq -t_0. \end{cases} \quad (43)$$

The maximum value of  $d(t)$  occurs for

$$\begin{aligned} t = t_{\max} &= -t_0 \beta \ln 2 - e^{-\frac{1}{\beta}} \\ &\approx -0.69\beta t_0 \quad (\text{for small } \beta) \end{aligned} \quad (44)$$

and

$$\begin{aligned} d(t_{\max}) &= 1 - \beta \ln 2 \\ &= 1 - 0.69\beta. \end{aligned} \quad (45)$$

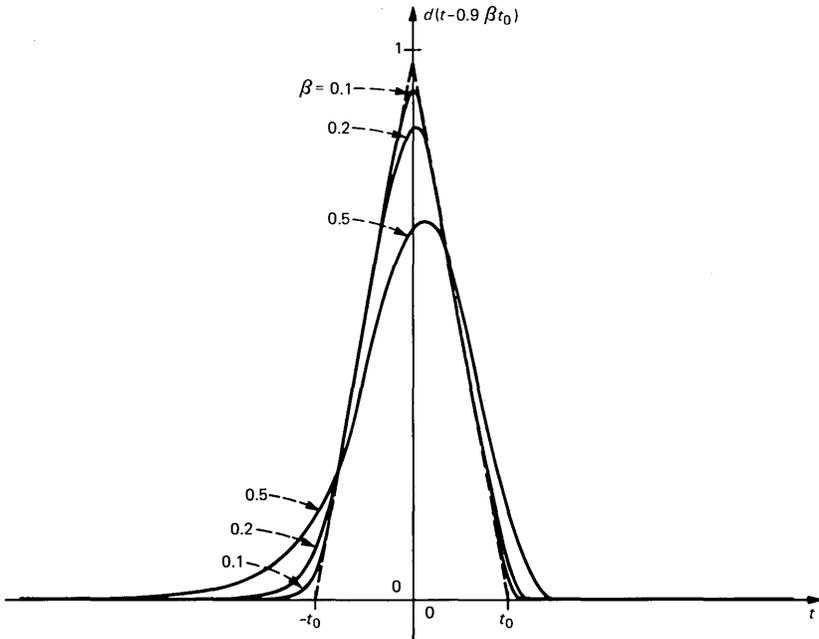


Fig. 10—The  $d(t)$  pulse for a single-pole RC filter.

However, for sufficiently small  $\beta$ ,  $d(\cdot)$  is approximately symmetric with respect to the abscissa  $t = t_d = -\beta t_0$ . In Fig. 10 we report  $d(t - 0.9\beta t_0)$  for three values of  $\beta$ . The dashed curve in Fig. 10 refers to  $c_g(t)$ . It is interesting to note that  $d(t)$  is independent of the sequence length  $M$  and is almost indistinguishable from  $c_g(\cdot)$  for  $\beta \leq 0.2$ .

The noisy term  $q_n(t)$  can be treated similarly to the prior case. Here, the condition for negligible  $q_n(t)$  is:

$$\frac{RC}{T} \gg 1. \quad (46)$$

For an exact evaluation of  $q(t)$  a similar procedure as above is used (details appear in Appendix B). Figures 11 and 12 show a shifted (by  $0.9\beta t_0$ ) and scaled (by  $K$ ) version of  $q(\cdot)$ ,  $\tilde{q}(\cdot)$ , for two values of  $RC$  and the same value of  $\beta$ . In Figs. 11 and 12 the maximum value of  $|q_n(t)|$  is  $-34$  dB and  $-44$  dB, respectively.

## V. CONCLUSION

Expressions for the probing pulse in a measurement system have been derived. The closed-form expression for the main component is particularly simple. This has allowed us to investigate the trade-off between the parameter  $K$  and the receiver filter complexity. In particular, for the same system configuration as in Ref. 2, we can replace the ideal integrate-and-dump filter with a simple RC filter, but  $K$  must

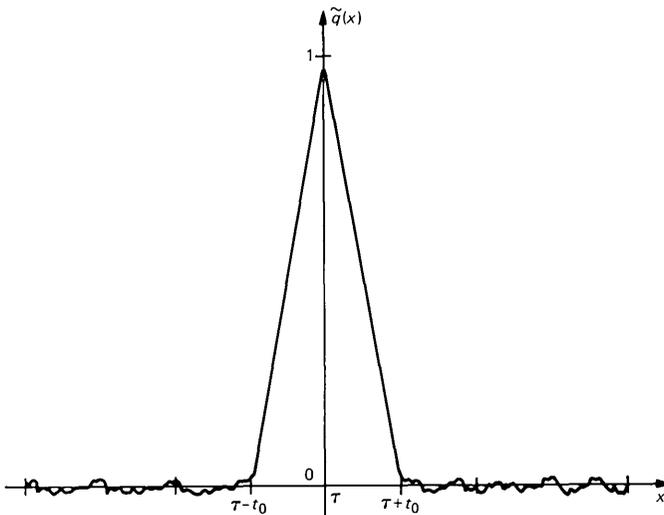


Fig. 11—Probing pulse for a single-pole RC filter:  $M = 511$ ,  $RC/T = 0.978$ , and  $\beta = 0.046$ .

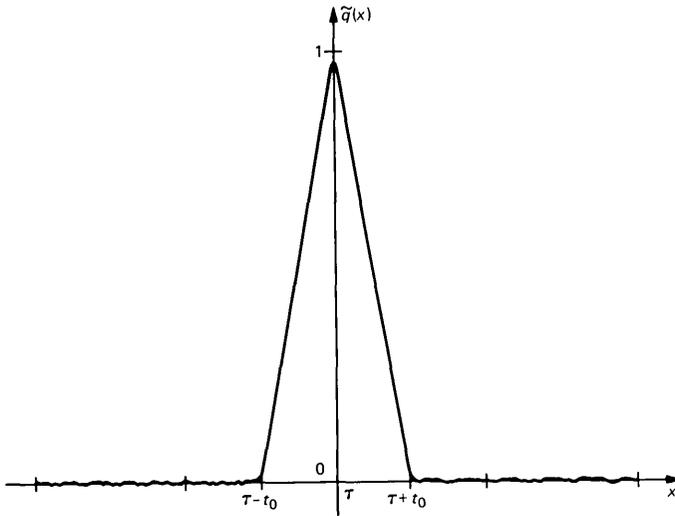


Fig. 12—Probing pulse for a single-pole RC filter.  $M = 511$ ,  $RC/T = 4.5$ , and  $\beta = 0.046$ .

be increased by more than an order of magnitude in order to maintain the same pulse shape and noise level. The larger  $K$  requires a more precise frequency synthesizer. The increased restrictions imposed by the RC filter relative to an ideal integrator may often be accommodated without modification of typical system performance capabilities.<sup>2</sup>

## VI. ACKNOWLEDGMENT

I am grateful to M. J. Gans for many helpful comments on an earlier version of this paper.

## REFERENCES

1. M. S. Gupta, "Applications of Electrical Noise," Proc. IEEE, 63 (July 1975), pp. 996-1010.
2. D. C. Cox, "Delay Doppler Characteristics of Multipath Propagation at 910 MHz in a Suburban Mobile Radio Environment," IEEE Trans. Antennas Propagat., AP-20 (September 1972), pp. 625-35.
3. J. D. Balcomb, H. B. Demuth, and E. P. Gyftopoulos, "A Cross-Correlation Method for Measuring the Impulse Response of Reactor Systems," Nuclear Sci. Eng., 11 (September-December 1961), pp. 159-66.
4. S. W. Golomb, ed., *Digital Communications With Space Applications*, Englewood Cliffs, N.J.: Prentice-Hall, 1964.
5. G. F. Sage, "Serial Synchronization of Pseudonoise Systems," IEEE Trans. Commun., COM-12 (December 1964), pp. 123-7.

## APPENDIX A

### Derivation of $q(t)$

Since  $r(t)$  is a periodic function of period  $T$ , we can consider its Fourier Series (FS) expansion:

$$r(t) = \sum_{n=-\infty}^{+\infty} \mathcal{R}(n) e^{j2\pi \frac{n}{T} t}. \quad (47)$$

If  $g(t)$  has finite support  $t_0$ , the Fourier Coefficients (FCs)  $\{\mathcal{R}(n)\}$  are given by

$$\mathcal{R}(n) = \mathcal{A}(n) \mathcal{G}(n), \quad (48)$$

where  $\mathcal{A}(n)$  is the discrete Fourier Transform (FT) of  $\{a_n\}$ ,  $n = 0, 1, \dots, M-1$ :

$$\mathcal{A}(n) = \frac{1}{M} \sum_{i=0}^{M-1} a_i e^{j2\pi \frac{i}{M} n} \quad (49)$$

and  $\mathcal{G}(n)$  is proportional to the FT of  $g(t)$ ,  $G(f)$ , evaluated at  $f = n/T$ ,

$$\mathcal{G}(n) = \frac{1}{t_0} G\left(\frac{n}{T}\right). \quad (50)$$

For later use, two facts follow. First,  $\mathcal{A}(n)$  is related to the discrete FT of  $\{c_a(k)\}$ ,  $\{\mathcal{L}_a(n)\}$ , by

$$\mathcal{L}_a(n) = |\mathcal{A}(n)|^2 \quad (51)$$

and from (2) it is shown that

$$\mathcal{L}_a(n) = \begin{cases} \frac{1}{M^2}, & \text{for } (n) \bmod M = 0 \\ \frac{1}{M} + \frac{1}{M^2}, & \text{otherwise.} \end{cases} \quad (52)$$

Second, we note that  $(1/M) |\mathcal{G}(n)|^2$  is just the  $n$ th FC in the FS expansion of  $c_g(t)$  [defined in (7)] in the interval  $[-(T/2), (T/2)]$ . Thus we can write

$$c_g(t) = \sum_{n=-\infty}^{+\infty} \frac{1}{M} |\mathcal{G}(n)|^2 e^{j2\pi \frac{n}{T} t}, \quad \text{for } |t| \leq \frac{T}{2}. \quad (53)$$

Since  $r'(\cdot)$  is related to  $r(\cdot)$  by the equation

$$r'(t) = r\left(t \frac{t_0}{t'_0}\right), \quad (54)$$

$r'(t)$  is also a periodic function but of period  $T' = Mt'_0 = [K/(K-1)]T$ , and its FS has the form:

$$r'(t) = \sum_{m=-\infty}^{+\infty} \mathcal{R}(n) e^{j2\pi \frac{m}{T'} t}. \quad (55)$$

Since  $K$  and  $K - 1$  are mutually prime numbers,

$$p(t) = r'(t)r(t - \tau) \quad (56)$$

is a periodic function of period  $KT$ . Its FS has the form

$$p(t) = \sum_{l=-\infty}^{+\infty} \mathcal{P}(l)e^{j2\pi\frac{l}{KT}t} \quad (57)$$

Upon substituting (47) and (55) into (56) we have

$$p(t) = \sum_{n=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} \mathcal{R}(n)\mathcal{R}(m)e^{-j2\pi\frac{n}{T}t} e^{j2\pi\frac{t}{KT}(Kn+(K-1)m)} \quad (58)$$

Comparison of (57) and (58) shows that  $\mathcal{P}(l)$  can be written as

$$\mathcal{P}(l) = \sum_{\substack{\text{all } (n,m) \\ Kn+(K-1)m=l}} \mathcal{R}(n)\mathcal{R}(m)e^{-j2\pi\frac{n}{T}\tau} \quad (59)$$

Making use of assumption (19) we can consider the FS of  $h(t)$  in the interval  $[0, KT]$ ; its FCs will be given by

$$\mathcal{A}(n) = \frac{1}{KT} H\left(\frac{n}{KT}\right), \quad (60)$$

where  $H(f)$  is the FT of  $h(t)$ .

Finally, since  $q(t)$  is the convolution of  $h(t)$  with  $p(t)$ ,  $q(t)$  is a periodic function of period  $KT$  and we can write

$$q(t) = \sum_{l=-\infty}^{+\infty} \mathcal{Q}(l)e^{j2\pi\frac{l}{KT}t}, \quad (61)$$

where coefficients  $\{\mathcal{Q}(l)\}$  are given by

$$\mathcal{Q}(l) = KT \mathcal{A}(l) \mathcal{P}(l). \quad (62)$$

Substituting (59) and (60) into (62), and this result into (61), we get

$$q(t) = KT \sum_{n=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} \mathcal{R}(n)\mathcal{R}(m)e^{-j2\pi\frac{n}{T}t} \cdot \mathcal{A}(Kn + (K - 1)m)e^{j2\pi\frac{Kn+(K-1)m}{KT}t} \quad (63)$$

Let  $l = n + m$ ; a more convenient expression for  $q(t)$  is:

$$q(t) = KT \sum_{l=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} \mathcal{R}(m) \mathcal{R}(l-m) \cdot \mathcal{L}(Kl-m) e^{-j2\pi \frac{l-m}{T} \tau} e^{j2\pi \left(\frac{l}{T} - \frac{m}{KT}\right) t} \quad (64)$$

In general,  $q(t)$  can be separated in a useful component  $q_0(t)$ , given by (64) for  $l = 0$ , and a noisy component  $q_n(t)$ , i.e.,

$$q(t) = q_0(t) + q_n(t), \quad (65)$$

where

$$q_0(t) = KT \sum_{m=-\infty}^{+\infty} \mathcal{R}(m) \mathcal{R}(-m) \mathcal{L}(-m) e^{j2\pi \frac{m}{T} \left(\tau - \frac{t}{K}\right)}. \quad (66)$$

A closed-form expression for  $q_0(t)$  is now derived. Since  $\mathcal{R}(-m) = \mathcal{R}^*(m)$ , using (48), (51) and (52) we get

$$q_0(t) = - \sum_{m=-\infty}^{+\infty} T \frac{1}{M} |\mathcal{L}(mM)|^2 K \mathcal{L}(-mM) e^{j2\pi \frac{mM}{T} \left(\tau - \frac{t}{K}\right)} + \left(\frac{1}{M} + 1\right) \sum_{m=-\infty}^{+\infty} T \frac{1}{M} |\mathcal{L}(m)|^2 K \mathcal{L}(-m) e^{j2\pi \frac{m}{T} \left(\tau - \frac{t}{K}\right)}. \quad (67)$$

Using general properties of FS, (67) becomes

$$q_0(t) = - \frac{1}{M} \sum_{i=-(M-1)}^{M-1} d\left(\tau - \frac{t}{K} - it_0\right) + \left(\frac{1}{M} + 1\right) d\left(\tau - \frac{t}{K}\right) \quad \text{for } -T + t_0 \leq \tau - \frac{t}{K} \leq t_0, \quad (68)$$

where  $d(\cdot)$  is defined in (22). From (65) and (66) the noisy term  $q_n(t)$  has the form:

$$q_n(t) = \sum'_{l=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} \mathcal{L}(m) \mathcal{L}(l-m) \frac{1}{t_0} G\left(\frac{m}{T}\right) \frac{1}{t_0} G\left(\frac{l-m}{T}\right) \cdot H\left(\frac{l}{T} - \frac{m}{KT}\right) e^{j2\pi \frac{l-m}{T} \tau} e^{j2\pi \left(\frac{l}{T} - \frac{m}{KT}\right) t}, \quad (69)$$

where the primed summation excludes the  $l = 0$  term. In the particular case when  $g(\cdot)$  is given by (8), (68) simplifies into

$$q_0(t) = - \frac{1}{M} + \left(\frac{1}{M} + 1\right) d\left(\tau - \frac{t}{K}\right) \quad \text{for } -T + t_0 \leq \tau - \frac{t}{K} \leq t_0. \quad (70)$$

## APPENDIX B

### Derivation of $\tilde{q}(x)$ for a Single-Pole RC Filter

From Fig. 5, and for  $h(t)$  given by (41), we have that

$$q(t) = \int_{-\infty}^t \frac{1}{RC} e^{-\frac{t-\zeta}{RC}} r'(\zeta) r(\zeta - \tau) d\zeta. \quad (71)$$

Let  $z_I(\eta)$  [ $I$  has the same meaning as in (32)] be a periodic function of period  $T$  defined as

$$z_I(\eta) = \int_0^{QT} \frac{1}{RC} e^{-\frac{QT-\zeta}{RC}} r'(\zeta) r(\zeta - \eta) d\zeta. \quad (72)$$

Let

$$QT \gg RC \quad (Q \text{ an integer}), \quad (73)$$

so that the infinite range integral in (71) may be truncated to  $\{t - QT, t\}$ . As in the case of the ideal integrator we evaluate  $q(t)$  at discrete times:

$$t = QT + It'_0 \quad (I \text{ an integer}). \quad (74)$$

Since

$$I(t'_0 - t_0) = I \frac{t'_0}{K} = \frac{t - QT}{K}, \quad (75)$$

we have for the times given in (74):

$$q(t) = z_I \left( \tau - \frac{t - QT}{K} \right). \quad (76)$$

The value of  $Q$  should be chosen as a compromise between accuracy in the results [see eq. (72)] and long computer execution time. For our purposes, a choice of  $Q = 8$  was more than adequate. In order to simplify matters, we are also going to assume that

$$K > MQ. \quad (77)$$

Under these conditions,  $r'(t)$  can be written as

$$r'(t) = \sum_{j=0}^{MQ-2} a_{j+I} g'(t - jt'_0) + a_{MQ-1+I} g''(t - (MQ - 1)t'_0), \quad (78)$$

for  $0 \leq t \leq QT$ ,

where

$$g''(t) = \text{rect} \frac{t - t''_0/2}{t''_0} \quad (79)$$

and

$$t''_0 = t_0 \frac{K - MQ}{K - 1}. \quad (80)$$

Also for  $0 < \eta \leq T$  we have

$$r(t - \eta) = \sum_{i=0}^{MQ} a_{i-L+I} g(t - it_0 + \zeta) \quad \text{for } 0 < t - \eta \leq QT, \quad (81)$$

where  $L$  and  $\zeta$  are related to  $\eta$  by

$$\eta = Lt_0 - \zeta, \quad L = 1, 2, \dots, M, \quad 0 < \zeta \leq t_0. \quad (82)$$

Upon substitution of (81) and (78) into (72) we get

$$z_I(Lt_0 - \eta) = \sum_{i=0}^{MQ} \left\{ \sum_{j=J_m}^{J_M} a_{i-L+I} a_{j+I} e_{gg'}(it_0 - \zeta, jt'_0) \right. \\ \left. + a_{i-L+I} a_{MQ-1+I} e_{gg'}[it_0 - \zeta, (MQ - 1)t'_0] \right\}, \quad (83)$$

where

$$J_m = \min \left[ \left( \frac{it_0 - \zeta}{t'_0} - 1 \right)_C, MQ - 2 \right] \quad (84)$$

$$J_M = \min \left\{ \left[ \frac{(i+1)t_0 - \zeta}{t'_0} \right]_F, MQ - 2 \right\} \quad (85)$$

$$e_{gg'} \left( a - \frac{t_0}{2}, b - \frac{t'_0}{2} \right) \\ = \int_{-\infty}^{+\infty} \frac{1}{RC} e^{-\frac{QT-t}{RC}} g \left( t - a + \frac{t_0}{2} \right) g' \left( t - b + \frac{t'_0}{2} \right) dt \\ = e^{-\frac{QT-a}{RC}} \left\{ \begin{array}{ll} 0, & b - a < -\frac{t'_0 + t_0}{2} \\ e^{\frac{b-a+t'_0/2}{RC}} - e^{\frac{t_0/2}{RC}}, & b - a > -\frac{t'_0 + t_0}{2} \\ e^{\frac{t_0/2}{RC}} - e^{\frac{t_0/2}{RC}}, & b - a < -\frac{t'_0 - t_0}{2} \\ e^{\frac{t_0/2}{RC}} - e^{\frac{t_0/2}{RC}}, & b - a > -\frac{t'_0 - t_0}{2} \\ e^{\frac{t_0/2}{RC}} - e^{\frac{b-a-t'_0/2}{RC}}, & b - a < \frac{t'_0 - t_0}{2} \\ e^{\frac{t_0/2}{RC}} - e^{\frac{b-a-t'_0/2}{RC}}, & b - a > \frac{t'_0 - t_0}{2} \\ 0, & b - a < \frac{t'_0 + t_0}{2} \\ 0, & b - a > \frac{t'_0 + t_0}{2} \end{array} \right. \quad (86)$$

$$\begin{aligned}
& e_{gg''} \left( a - \frac{t_0}{2}, b - \frac{t_0''}{2} \right) \\
&= \int_{-\infty}^{+\infty} \frac{1}{RC} e^{-\frac{QT-t}{RC}} g \left( t - a + \frac{t_0}{2} \right) g'' \left( t - b + \frac{t_0''}{2} \right) dt \\
&= e^{-\frac{QT-b}{RC}} \left\{ \begin{array}{ll} 0, & b - a < -\frac{t_0 + t_0''}{2} \\ e^{\frac{t_0''/2}{RC}} - e^{-\frac{b-a+t_0/2}{RC}}, & b - a > -\frac{t_0 + t_0''}{2} \\ e^{\frac{t_0''/2}{RC}} - e^{-\frac{t_0-t_0''}{2RC}}, & b - a < -\frac{t_0 - t_0''}{2} \\ e^{\frac{t_0''/2}{RC}} - e^{-\frac{t_0-t_0''}{2RC}}, & b - a > -\frac{t_0 - t_0''}{2} \\ e^{-\frac{b-a-t_0/2}{RC}} - e^{-\frac{t_0-t_0''}{2RC}}, & b - a < \frac{t_0 - t_0''}{2} \\ e^{-\frac{b-a-t_0/2}{RC}} - e^{-\frac{t_0-t_0''}{2RC}}, & b - a > \frac{t_0 - t_0''}{2} \\ 0, & b - a < \frac{t_0 + t_0''}{2} \\ 0, & b - a > \frac{t_0 + t_0''}{2} \end{array} \right. \quad (87)
\end{aligned}$$

Finally,  $\tilde{q}(x)$  as reported in Figs. 11 and 12 is defined only for  $x$  multiples of  $t_0'/K$  as

$$\begin{aligned}
\tilde{q}(x) &= q[K(x + 0.9\beta t_0)] \\
&= z_I \left( \tau + \frac{QT - 0.9K\beta t_0}{K} - x \right), \quad I = \frac{K}{t_0'} x. \quad (88)
\end{aligned}$$

#### AUTHOR

**Nevio Benvenuto**, Dr. Ing., University of Padova, Padova, Italy, 1976, Ph.D., University of Massachusetts, Amherst, 1983, both in Electrical Engineering. Mr. Benvenuto spent a year (1976-1977) at the University of Padova as a researcher in the communication area. Since 1983, he has been with AT&T Bell Laboratories, Holmdel, New Jersey, working on VLSI system design and signal analysis problems. Member, IEEE.

## Soft Decision Demodulation to Reduce the Effect of Transmission Errors in Logarithmic PCM Transmitted Over Rayleigh Fading Channels

By W. C. WONG,\* R. STEELE,<sup>†</sup> and C.-E. SUNDBERG<sup>‡</sup>

(Manuscript received May 16, 1984)

In this paper we investigate soft decision demodulation applied to  $\mu$ -law Pulse Code Modulated (PCM)-encoded signals transmitted over Rayleigh fading channels by means of coherent phase-shift keying modulation. Each bit in the  $\mu$ -law PCM word is assigned its own soft decision demodulation erasure threshold. These thresholds are theoretically determined as a function of the input power level, channel  $s/n$ , and the relative mean square error power  $\delta$  that occurs because of the replacement by interpolation or prediction of those samples discarded in the soft decision demodulation process. We find that there is no advantage in applying soft demodulation to more than the first 4 bits of the 8-bit  $\mu$ -law PCM words,  $\mu = 255$ . When the input signal level was  $-17$  dB (corresponding to peak overall speech  $s/n$  in the absence of transmission errors), the gains in overall speech  $s/n$  compared to basic  $\mu$ -law PCM and fixed weighted  $\mu$ -law PCM were 7 and 3.5 dB, respectively, when the channel  $s/n$  was 30 dB. More significantly, when the input signal level was reduced to  $-40$  dB, the corresponding gains in overall speech  $s/n$  were 18 and 12 dB. Simulations were performed using four concatenated speech sentences and a fading channel that was obtained from a mobile radio Rayleigh fading hardware simulator. The simulations were in reasonable agreement with our theoretical results.

### I. INTRODUCTION

Digital transmission of logarithmically companded Pulse Code Mod-

---

\* National University of Singapore. <sup>†</sup> University of Southampton, England. <sup>‡</sup> AT&T Bell Laboratories.

---

Copyright © 1984 AT&T. Photo reproduction for noncommercial use is permitted without payment of royalty provided that each reproduction is done without alteration and that the Journal reference and copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free by computer-based and other information-service systems without further permission. Permission to reproduce or republish any other portion of this paper must be obtained from the Editor.

ulation (log-PCM) over Rayleigh fading channels has been studied for Coherent Phase-Shift Keying (CPSK) and Noncoherent Frequency-Shift Keying (NCFSK).<sup>1</sup> The effect of transmission errors on the recovered analog signal can be most serious, particularly when the more significant magnitude bits are corrupted. To mitigate these effects numerous techniques can be employed, such as: channel protection coding,<sup>2</sup> bit scrambling prior to transmission in order to combat burst errors,<sup>1</sup> diversity schemes,<sup>3,4</sup> weighted PCM,<sup>5-7</sup> soft decision demodulation,<sup>8-10</sup> and so forth.

In this discourse we will focus on soft decision demodulation, a method that is implemented only at the receiver. As a preamble we need to distinguish between soft and hard decision demodulation. Let us commence with the observation that although the carrier is modulated by a binary log-PCM signal for its transmission over a Rayleigh fading channel, the demodulation process yields an analog rather than a digital signal. This loss of binary status is due to the channel imperfections and the noise in the front end of the receiver. The role of the regenerator may be viewed as an Analog-to-Digital Converter (ADC), identifying the binary signal that lies latent in the demodulated radio signal. When hard decision demodulation is employed, the output of the ADC is a logical one or a logical zero, the state being dependent on which side of the decision boundary the input analog level resides. By contrast, soft decision demodulation has more than two output states, and the occupancy of any state is associated with a reliability number, i.e., a description of our confidence as to whether the transmitted bit was a logical one or a logical zero. For the simple soft decision demodulator we have three states corresponding to three zones. The middle zone extends from  $-Z$  to  $+Z$ , and is known as an erasure zone. As the transmitted binary signal represents logical one and logical zero states by equal magnitude positive and negative voltages, respectively, a signal level that resides within the erasure zone at a sampling instant is deemed unreliable. This unreliable status is assigned because the level of the signal is not sufficiently positive or negative to warrant our confidence of assigning it as a logical one or logical zero. The other two zones stretch for input signal levels that are  $\geq Z$ , and  $\leq -Z$ , and signal levels at a sampling instant that are in these zones are regenerated with relative confidence as bits of logical one or logical zero, respectively.

Various strategies can be employed when the analog input falls into the erasure zone. Typically,<sup>8-10</sup> if one or more bits in the first  $M$  Most Significant Bit(s) (MSB) of an  $N$ -bit PCM word are deemed to be "unreliable", the entire word is rejected. The resulting missing decoded sample is replaced by means of interpolation or prediction from neighboring decoded samples. The last  $(N - M)$  bits of the PCM word

are not subjected to soft decision demodulation because an error in these bits is decoded into a sample that is often more acceptable than a sample that has been imperfectly introduced by interpolation or prediction.

It is our intention here to extend the work on soft decision demodulation for the Gaussian channel<sup>8-10</sup> to the case of the Rayleigh fading channel. We commence in Section II with a discussion concerning the digital noise power in log-PCM where the transmission is over a Gaussian channel and soft decision demodulation is employed. The problems associated with a Rayleigh fading channel are encountered in Section III. Specifically, the conventional fixed-width erasure zone is examined in Section 3.1, and in Section 3.2 we optimize a soft decision demodulation system where each bit is assigned its unique erasure zone. Section IV states our performance criterion, while Section V is concerned with presenting the theoretical performance of our 8-bit  $\mu$ -law PCM that utilizes an individual bit threshold soft decision demodulation system. Computer simulation results for  $\mu$ -law PCM speech transmitted over a mobile radio channel and utilizing soft decision demodulation are provided in Section VI. The final section summarizes our findings.

## II. BASIC CONCEPTS

When the log-PCM binary output signal is transmitted in the form of antipodal signals over an additive white Gaussian noise channel, the digital noise power produced when a soft decision demodulator is employed can be shown to be<sup>8</sup>

$$\epsilon_a^2 \cong P_D \sum_{l=1}^M A_l + P \sum_{l=M+1}^N A_l + P_R \delta S^2, \quad (1)$$

where we have assumed that the bit errors occur independently. The first term in this equation is the digital noise power due to the presence of only one erroneously regenerated bit in one of the first  $M$  bits of the  $N$ -bit log-PCM words. These bit errors occur because the received signal amplitude resides outside the erasure zone at the instant of bit regeneration. Consequently, the erroneously regenerated bits are not designated as unreliable. The probability of an undetected bit error is  $P_D$ , while  $A_l$  is called the  $l$ th single-error  $A$ -factor. For a full description of the single-error  $A$ -factors the reader is advised to consult Refs. 1, 2, 5, and 7. Suffice to state here that the definition of  $A_l$  is

$$A_l \triangleq E_i\{(x_i - x_{i,l})^2\}, \quad (2)$$

where  $E_i(\cdot)$  is the expectation of  $(\cdot)$ , and  $x_i$  or  $x_{i,l}$  is the recovered speech sample in the absence or presence of transmission errors,

respectively,  $i = 0, 1, \dots, 2^{N-1}$ . The expectation is made over all  $i = 2^N$  quantized levels. Note that for input speech samples  $x$ , the encoder noise power is given by

$$\epsilon_q^2 + \epsilon_c^2 = E\{(x - x_i)^2\}, \quad (3)$$

where  $\epsilon_q^2$  and  $\epsilon_c^2$  are the quantization and clipping noise power contributed by the encoder, respectively, and the expectation is formed over the source statistics. By contrast,  $A_l$  is the average noise power in the decoded output sequence due to the existence of a single error in the  $l$ th bit of all the transmitted  $N$ -bit log-PCM words. Observe that  $A_1, A_2, \dots, A_N$  are the first, second,  $\dots$ ,  $N$ th, single-error  $A$ -factors, respectively, and relate to a single transmission error in the first, second,  $\dots$ ,  $N$ th bits of the log-PCM words, respectively. We assume that the regenerated errors are statistically independent.

The contribution to the total digital noise power due to an error occurring in the last  $(N - M)$  bits of the log-PCM words is represented by the second term in eq. (1). These least significant  $(N - M)$  bits are not subjected to soft decision demodulation, i.e., they are detected by hard decision demodulation. The average bit error probability for the hard decision demodulation is  $P$ .

When one or more bits in the first  $M$  bits of a log-PCM word are deemed unreliable, the complete word is rejected. The probability of this event is  $P_R$ . The noise power due to replacing the original log-PCM words by those determined by interpolation or prediction is represented by the final term in eq. (1). Observe that this noise power is the product of the input signal power  $S^2$  and the relative mean square error  $\delta$  of the correction process, namely,

$$\delta S^2 = E\{(x - \hat{x})^2\}, \quad (4)$$

multiplied by its probability of occurrence,  $P_R$ . The value of  $\delta$  is determined by the quality of the correction process, having a low value when the corrections are accurate, and vice versa. For the type of correctors that are relatively easy to implement the value of  $\delta$  spans the range from 0.4 to 0.01. In eq. (4),  $x$  is the input sample, and  $\hat{x}$  is the sample at the output of the receiver after the interpolation or prediction processes have been applied. We also note that

$$P_R = 1 - (1 - P_z)^M \cong MP_z, \quad (5)$$

where  $P_z$  is the probability that a matched filter output voltage is within the erasure zone  $\pm Z$ , where

$$Z = T\sqrt{E} \quad (6)$$

and where  $T$  is the soft decision erasure threshold and  $E$  is the bit energy. When Coherent Phase-Shift Keying (CPSK) modulation is used, the probabilities  $P$ ,  $P_D$ , and  $P_z$  are<sup>8</sup>

$$P = Q\left(\sqrt{2\frac{E}{N_o}}\right) \quad (7)$$

$$P_D = Q\left(\sqrt{2\frac{E}{N_o}}(1+T)\right) \quad (8)$$

$$P_2 = Q\left(\sqrt{2\frac{E}{N_o}}(1-T)\right) - Q\left(\sqrt{2\frac{E}{N_o}}(1+T)\right), \quad (9)$$

where  $E/N_o$  is the channel s/n,  $N_o$  is the one-sided spectral density function for white Gaussian noise, and

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-u^2/2} du. \quad (10)$$

Threshold optimizations for the Gaussian channel are described in Ref. 8.

### III. RAYLEIGH FADING CHANNELS

When the bits constituting the log-PCM signal are scrambled prior to their transmission over a Rayleigh fading channel, the burst errors that would have frequently occurred in the regeneration process at the receiver appear as random errors. Because of the scrambling process we assume that the bit errors are statistically independent. The channel s/n,  $\gamma$ , namely  $E/N_o$ , is a random variable in a Rayleigh fading channel, and its Probability Distribution Function (PDF) is given by<sup>4</sup>

$$f(\gamma) = \frac{1}{\Gamma} e^{-\gamma/\Gamma}, \quad (11)$$

where

$$\Gamma = E\{\gamma\}. \quad (12)$$

The performance of the soft decision demodulation scheme will be compared with that of the basic log-PCM transmission system. Accordingly, we note that the digital noise power for the log-PCM system is<sup>1</sup>

$$\epsilon_a^2 = \sum_{w=1}^N P^w (1-P)^{N-w} S_w, \quad (13)$$

where  $S_w$  is the sum of the  $A$ -factors associated with the number of error patterns containing  $w$  bit errors in the  $N$ -bit words. If the binary log-PCM signal is transmitted using CPSK modulation, the bit error probability as a function of channel s/n  $\gamma$  is<sup>1</sup>

$$P(\gamma) = Q(\sqrt{2\gamma}). \quad (14)$$

The average bit error probability for the Rayleigh fading channel is given by

$$\mathbf{P} = \int_0^{\infty} f(\gamma)P(\gamma)d\gamma \quad (15)$$

and after substituting  $f(\gamma)$  and  $P(\gamma)$  into eq. (15),  $\mathbf{P}$  is substituted for  $P$  in eq. (13) to yield the digital noise power. We do this on the proviso that bit errors occur independently, i.e., ideal interleaving is used.<sup>1</sup>

### 3.1 Fixed soft demodulation threshold system

We now consider the application of soft decision demodulation to enhance the performance of log-PCM transmission over a Rayleigh fading channel. The soft decision demodulation system to be considered contains a fixed threshold  $T$  as described in Section II. We specify the value of this threshold as a function of the average s/n,  $\Gamma$ , such that the erasure zone width of  $\pm Z$ , see eq. (6), becomes

$$Z = T\sqrt{\mathbf{E}}, \quad (16)$$

where the average s/n,  $\Gamma$ , is related to  $\mathbf{E}$  by

$$\Gamma = E\{\gamma\} = \frac{\mathbf{E}}{N_o} \quad (17)$$

and  $N_o$  is the one-sided spectral density of the additive white Gaussian noise in the Rayleigh fading channel. The average bit error probability for CPSK with Rayleigh fading is from eqs. (7) and (11):

$$\begin{aligned} \mathbf{P} &= \int_0^{\infty} \frac{\gamma}{\Gamma} e^{-\gamma/\Gamma} Q(\sqrt{2\gamma})d\gamma \\ &= \frac{1}{2} \left[ 1 - \frac{\sqrt{\Gamma}}{\sqrt{1+\Gamma}} \right]. \end{aligned} \quad (18)$$

We now apply the average process to determine the probability  $\mathbf{P}_D$  of occurrence of an undetected bit error. From eqs. (8) and (16) we formulate

$$P_D(\gamma) = Q(\sqrt{2\gamma} + T\sqrt{2\Gamma}),$$

whence, with the aid of eq. (11),

$$\begin{aligned} \mathbf{P}_D &= \int_0^{\infty} \frac{1}{\Gamma} e^{-\gamma/\Gamma} Q(\sqrt{2\gamma} + T\sqrt{2\Gamma})d\gamma \\ &= Q(T\sqrt{2\Gamma} - \sqrt{\frac{\Gamma}{1+\Gamma}} e^{-T^2\left(\frac{\Gamma}{1+\Gamma}\right)}) Q\left(\frac{T\sqrt{2\Gamma}}{\sqrt{1+\Gamma}}\right). \end{aligned} \quad (19)$$

The average of  $P_z$  is

$$\mathbf{P}_z = \int_0^\infty \frac{1}{\Gamma} e^{-\gamma/\Gamma} P_z(\gamma) d\gamma \quad (20)$$

and from eqs. (8) and (9)

$$P_z(\gamma) = Q(\sqrt{2\gamma} - T\sqrt{2\Gamma}) - P_D(\gamma) \quad (21)$$

and, hence,

$$\mathbf{P}_z = \int_0^\infty \frac{1}{\Gamma} e^{-\gamma/\Gamma} Q(\sqrt{2\gamma} - T\sqrt{2\Gamma}) d\gamma - \mathbf{P}_D. \quad (22)$$

This expression can be simplified to

$$\mathbf{P}_z = 1 - 2Q(T\sqrt{2\Gamma}) - \sqrt{\frac{\Gamma}{1+\Gamma}} e^{-T^2 \frac{\Gamma}{1+\Gamma}} \left[ 1 - 2Q\left(\frac{T\sqrt{2\Gamma}}{\sqrt{1+\Gamma}}\right) \right]. \quad (23)$$

By replacing  $P_D$ ,  $P$ , and  $P_R \approx MP_z$  in eq. (1) by the average values  $\mathbf{P}_D$ ,  $\mathbf{P}$ , and  $\mathbf{P}_z$  given by eqs. (19), (18), and (23), respectively, we formulate the digital noise power as

$$\begin{aligned} \epsilon_a^2 = & \left[ Q(T\sqrt{2\Gamma}) - \sqrt{\frac{\Gamma}{1+\Gamma}} e^{-T^2 \frac{\Gamma}{1+\Gamma}} Q\left(T\sqrt{2\Gamma} \sqrt{\frac{\Gamma}{1+\Gamma}}\right) \right] \\ & \cdot \left[ \sum_{l=1}^M A_l - M\delta S^2 \right] \\ & + \frac{1}{2} \left( 1 - \sqrt{\frac{\Gamma}{1+\Gamma}} \right) \sum_{l=M+1}^N A_l \\ & + \left\{ [1 - Q(T\sqrt{2\Gamma})] - \sqrt{\frac{\Gamma}{1+\Gamma}} e^{-T^2 \frac{\Gamma}{1+\Gamma}} \right. \\ & \left. \cdot \left[ 1 - Q\left(T\sqrt{2\Gamma} \sqrt{\frac{\Gamma}{1+\Gamma}}\right) \right] \right\} M\delta S^2. \quad (24) \end{aligned}$$

Observe that if only hard decision demodulation is used, the digital noise power becomes

$$\epsilon_a^2 = \frac{1}{2} \left( 1 - \sqrt{\frac{\Gamma}{1+\Gamma}} \right) \sum_{l=1}^N A_l. \quad (25)$$

### 3.2 Individual bit soft demodulation thresholds

Rather than employ the same threshold  $T$  for each of the  $M$  first bits in the  $N$ -bit PCM word, we now consider the case when the first,

second,  $\dots$ ,  $M$ th bits, are assigned thresholds  $T_1, T_2, \dots, T_M$ , respectively. When this procedure is adopted, the digital noise power of eq. (24) becomes

$$\epsilon_a^2 = \sum_{l=1}^M \alpha_l + \frac{1}{2} \left( 1 - \sqrt{\frac{\Gamma}{1+\Gamma}} \right) \sum_{l=M+1}^N A_l + \delta S^2 \sum_{l=1}^M \beta_l, \quad (26)$$

where

$$\alpha_l = \left[ Q(T_l \sqrt{2\Gamma}) - \sqrt{\frac{\Gamma}{1+\Gamma}} e^{-T_l^2 \frac{\Gamma}{1+\Gamma}} Q \left( T_l \sqrt{2\Gamma} \sqrt{\frac{\Gamma}{1+\Gamma}} \right) \right] \cdot [A_l - \delta S^2] \quad (27)$$

and

$$\beta_l = 1 - Q(T_l \sqrt{2\Gamma}) - \sqrt{\frac{\Gamma}{1+\Gamma}} e^{-T_l^2 \frac{\Gamma}{1+\Gamma}} \left[ 1 - Q \left( T_l \sqrt{2\Gamma} \sqrt{\frac{\Gamma}{1+\Gamma}} \right) \right]. \quad (28)$$

The terms

$$\alpha_l + \delta S^2 \beta_l; \quad l = 1, 2, \dots, M$$

are minimized with respect to the soft decision erasure threshold  $T_l$  for a given  $\Gamma$ . The  $M$  values of  $T_l$  found this way are substituted into eqs. (27) and (28) to give  $\alpha_l$  and  $\beta_l$ , which in turn are used to estimate the minimum digital noise power with the aid of eq. (26). The procedure is repeated for different values of  $\Gamma$ . Thus, for a soft demodulation system where each bit is assigned a unique threshold  $T_l$ , the digital noise power, and hence overall s/n, is determined as a function of channel s/n,  $\Gamma$ . Observe that our optimization process causes the individual thresholds  $T_l$  to be independent of the value of  $M$  for a given average channel s/n,  $\Gamma$ . For example, if we compare two soft demodulation systems having  $M = 3$  and  $M = 6$ , we find that they have the same  $T_1, T_2$ , and  $T_3$ . However, the system having  $M = 3$  has  $T_4 = T_5 = \dots = T_8 = 0$ , whereas the system with  $M = 6$  has only  $T_7 = T_8 = 0$ .

#### IV. PERFORMANCE CRITERION

We use as our objective performance criterion the overall s/n, defined as the power  $S^2$  of the input signal to the noise power of the overall error signal, viz:

$$s/\hat{n} = \frac{S^2}{\epsilon_a^2 + \epsilon_q^2 + \epsilon_c^2}, \quad (29)$$

where  $\epsilon_q^2$  and  $\epsilon_c^2$  are defined in connection with eq. (3), and  $\epsilon_a^2$  is the digital noise power. Although  $\epsilon_q^2$  and  $\epsilon_c^2$  are constant for a given log-PCM encoder and for an input signal having a given PDF,  $\epsilon_a^2$  is dependent on the transmission and reception techniques and on the channel. Its value is given by eqs. (24) and (26), and as described in Section III.

## V. THEORETICAL PERFORMANCE

The logarithmic PCM codec considered in our theoretical calculations was 8-bit  $\mu$ -law PCM,  $\mu = 255$ . The binary code employed was binary-folded PCM, where the MSB was the polarity bit of the quantized sample, and the remaining bits represented the sample magnitude. The range of the quantizer was from  $-1$  to  $+1$ , and the input signal was assumed to have an exponential PDF,

$$p_x(x) = \frac{1}{S\sqrt{2}} \exp\left(-\frac{\sqrt{2}x}{S}\right), \quad (30)$$

a density function that is known to be representative of the long-term PDF of speech signals. We considered two input power levels of  $-17$  dB and  $-40$  dB, which corresponded to the input signal having a standard deviation of  $\sqrt{2}/10$  and  $0.01$ , respectively. The  $-17$  dB was the input level to give maximum overall s/n in the absence of transmission errors, whereas the lower input level of  $-40$  dB had a correspondence with low-level input speech signals. The single-error A-factors employed in our calculations were those derived in Ref. 1. The digital modulation was CPSK. We emphasize that our results are for the Rayleigh fading channel.

We now present in Section 5.1 a series of tabulated results calculated from the equations provided in Sections III and IV. In Section 5.2 we provide additional theoretical results, but this time in a graphical format.

### 5.1 Soft demodulation using individual bit thresholds

Table I shows the values of the individual thresholds  $T_1, T_2, \dots, T_8$ , as a function of channel s/n, for an input power level of  $-17$  dB and a relative mean square error  $\delta$  of  $0.1$ . This table relates to the individual bit soft demodulation threshold scheme described in Section 3.2. The nature of log-PCM is that the bit that generates the greatest digital noise power when it is erroneously received is the most significant magnitude bit. This bit is the second bit in the word, and when soft demodulation is applied it has a threshold  $T_2$ . Thus, because the second MSB is capable of generating the largest error it is assigned the largest threshold. In general, the higher the contribution to the

Table I—Individual bit thresholds,  $T_0$ , as a function of channel  $s/n$ ,  $\Gamma$ , for input level  $-17$  dB,  $\delta = 0.1$ 

$\Gamma$	$T_1$	$T_2$	$T_3$	$T_4$	$T_5$	$T_6$	$T_7$	$T_8$
0	1.398231	1.594590	1.431157	0.687725	0.000000	0.000000	0.000000	0.000000
5	0.637306	0.727069	0.653243	0.313625	0.000001	0.000001	0.000001	0.000001
10	0.327333	0.373505	0.335553	0.161287	0.000001	0.000001	0.000001	0.000001
15	0.178123	0.203837	0.182689	0.087959	0.000003	0.000003	0.000003	0.000003
20	0.099486	0.113294	0.101778	0.048928	0.000001	0.000006	0.000006	0.000001
25	0.055756	0.063378	0.056953	0.027489	0.000007	0.000007	0.000007	0.000007
30	0.031123	0.035811	0.032048	0.015410	0.000001	0.000020	0.000007	0.000006
35	0.017558	0.020109	0.017823	0.008295	0.000009	0.000019	0.000020	0.000019
40	0.009995	0.011319	0.011901	0.004986	0.000061	0.000073	0.000071	0.000008
45	0.005612	0.007441	0.005312	0.002778	0.000035	0.000010	0.000011	0.000015
50	0.003194	0.004631	0.002849	0.001673	0.000057	0.000039	0.000035	0.000034
55	0.001817	0.001943	0.001618	0.001069	0.000046	0.000494	0.000474	0.000459
60	0.001228	0.000779	0.001206	0.000662	0.000160	0.000109	0.000105	0.000062

Table II—Individual bit thresholds,  $T_0$ , as a function of channel  $s/n$ ,  $\Gamma$ , for input level  $-17$  dB,  $\delta = 0.01$ 

$\Gamma$	$T_1$	$T_2$	$T_3$	$T_4$	$T_5$	$T_6$	$T_7$	$T_8$
0	2.160955	2.327636	2.192461	1.582595	0.980817	0.308022	0.000000	0.000000
5	0.987638	0.901698	0.998271	0.721279	0.447031	0.140304	0.000001	0.000001
10	0.506489	0.547609	0.514840	0.371179	0.229991	0.072350	0.000001	0.000001
15	0.275695	0.298701	0.279855	0.202130	0.125606	0.039297	0.000003	0.000003
20	0.154197	0.166714	0.155108	0.112054	0.069540	0.021877	0.000006	0.000006
25	0.086858	0.093416	0.086822	0.062541	0.039325	0.012327	0.000007	0.000007
30	0.048570	0.052675	0.050988	0.035585	0.021794	0.006826	0.000018	0.000008
35	0.027583	0.029888	0.027333	0.019425	0.012291	0.003812	0.000021	0.000020
40	0.013697	0.015763	0.016203	0.010958	0.007052	0.002211	0.000066	0.000012
45	0.012576	0.008794	0.010522	0.007447	0.004531	0.001325	0.000023	0.000005
50	0.004466	0.004910	0.004599	0.004615	0.002283	0.000647	0.000044	0.000036
55	0.003976	0.003906	0.002797	0.001813	0.001078	0.000372	0.000423	0.000242
60	0.002335	0.002483	0.002307	0.000783	0.000809	0.000355	0.000097	0.000098

digital noise power by an erroneous bit, the wider its erasure zone. We see in Table I that the erasure zones for  $T_5$  to  $T_8$  were very narrow, justifying our decision to only perform soft demodulation on the  $M$  (here 4) most significant bits. For high values of channel s/n,  $\Gamma$ , the probability of making an error in the bit regeneration process was small. Consequently the erasure zones were narrow for all the  $N$ -bits. As the channel s/n was decreased the erasure zones for the four MSB's widened, and for a channel s/n of 0 dB we observed that the boundaries of the erasure zone exceeded  $\pm\sqrt{E}$ .

When a good corrector was employed such that  $\delta = 0.01$ , the threshold values increased as shown in Table II, all other parameters being those associated with Table I. As the corrections were more accurate it was safer to activate the corrector more frequently compared to when the corrector having  $\delta = 0.1$  was used, and this activity is reflected by the larger threshold values in Table II compared to those in Table I.

A poor corrector was associated with  $\delta = 0.4$ , and the application of such a corrector was to yield a high amount of "correction noise" because the replacement of rejected samples was not very accurate. The intervention of the poor corrector was less likely to occur if the thresholds were reduced, and this was what happened, as can be seen in Table III, where  $\delta = 0.4$ .

When the input level was reduced to  $-40$  dB,  $\delta = 0.1$ , the threshold values were increased (except for the polarity bit), as is apparent by comparing Tables I and IV. This occurred because, for the same magnitude bit in error, more noise tended to be produced when the input power was low. We do not present tables for  $\delta = 0.01$  and  $0.04$  when the lower input level of  $-40$  dB was used, but the same trends observed in Tables I to III for the higher input were evident at this lower input power.

## 5.2 Graphical results

We commence our presentation of theoretical graphical results by displaying in Fig. 1 the variation of overall s/n, namely  $s/n$ , as a function of channel s/n for an input signal level of  $-17$  dB, and soft decision demodulation applied to the four MSB's. The curve for the basic 8-bit  $\mu$ -law PCM, i.e., no soft demodulation, provided a reference from which the soft demodulation performance can be judged. We also display an additional reference curve that relates to a fixed-bit weighting strategy applied to  $\mu$ -law PCM. In this weighting scheme every  $l$ th bit in the PCM word has its magnitude multiplied by  $\sqrt{\phi_l}$ ,  $l = 1, 2, \dots, N$ . Reference 5 presents the method for determining  $\sqrt{\phi_l}$  for CPSK modulation and a Rayleigh fading channel in order to minimize

Table III—Individual bit thresholds,  $T_0$ , as a function of channel  $s/n$ ,  $\Gamma$ , for input level  $-17$  dB,  $\delta = 0.4$

$\Gamma$	$T_1$	$T_2$	$T_3$	$T_4$	$T_5$	$T_6$	$T_7$	$T_8$
0	0.860252	1.085411	0.900163	0.000000	0.000000	0.000000	0.000000	0.000000
5	0.392162	0.494988	9.410554	0.000001	0.000001	0.000001	0.000001	0.000001
10	0.201742	0.254390	0.211293	0.000001	0.000001	0.000001	0.000001	0.000001
15	0.109656	0.138744	0.115081	0.000003	0.000003	0.000003	0.000003	0.000003
20	0.061201	0.077085	0.063834	00.00001	0.000000	0.000000	0.000001	0.000000
25	0.034235	0.043139	0.035805	0.000007	0.000007	0.000007	0.000007	0.000007
30	0.019297	0.024132	0.020214	0.000002	0.000024	0.000014	0.000026	0.000023
35	0.011055	0.013769	0.011485	0.000052	0.000020	0.000020	0.000019	0.000021
40	0.005944	0.007815	0.006511	0.000061	0.000070	0.000068	0.000014	0.000004
45	0.003554	0.004531	0.003261	0.000093	0.000022	0.000017	0.000005	0.000014
50	0.002395	0.002428	0.001588	0.000099	0.000037	0.000035	0.000034	0.000041
55	0.001068	0.001738	0.001073	0.000106	0.000235	0.000469	0.000477	0.000473
60	0.000619	0.000813	0.000732	0.000118	0.000097	0.000145	0.000098	0.000097

Table IV—Individual bit thresholds,  $T_0$ , as a function of channel  $s/n$ ,  $\Gamma$ , for input level  $-40$  dB,  $\delta = 0.1$

$\Gamma$	$T_1$	$T_2$	$T_3$	$T_4$	$T_5$	$T_6$	$T_7$	$T_8$
0	1.397515	2.799741	1.564703	0.907276	0.074134	0.000000	0.000000	0.000000
5	0.637817	1.285825	0.714349	0.413849	0.033814	0.000001	0.000001	0.000001
10	0.327897	0.659582	0.366937	0.212889	0.017447	0.000001	0.000001	0.000001
15	0.178242	0.361774	0.199849	0.115725	0.009486	0.000003	0.000003	0.000003
20	0.099536	0.198387	0.110899	0.064638	0.005377	0.000001	0.000006	0.000006
25	0.055780	0.108670	0.062466	0.036221	0.002834	0.000007	0.000007	0.000007
30	0.013398	0.062405	0.035221	0.020151	0.001512	0.000021	0.000020	0.000021
35	0.017505	0.038389	0.019635	0.011423	0.000924	0.000020	0.000020	0.000019
40	0.010103	0.023034	0.011127	0.006421	0.000536	0.000012	0.000069	0.000013
45	0.004531	0.014624	0.005947	0.003738	0.000420	0.000000	0.000005	0.000006
50	0.003000	0.005471	0.004557	0.002030	0.000189	0.000048	0.000037	0.000036
55	0.001844	0.005023	0.001552	0.001066	0.000149	0.000472	0.000484	0.000242
60	0.001283	0.000775	0.001254	0.000766	0.000245	0.000111	0.000062	0.000097

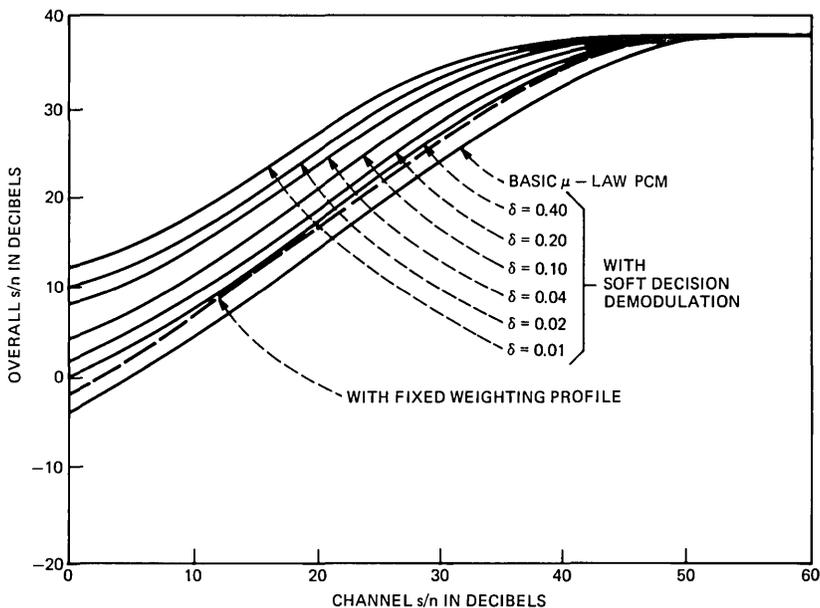


Fig. 1—Variation of overall  $s/n$  as a function of channel  $s/n$  for  $M = 4$ , input power level =  $-17$  dB, 8-bit  $\mu$ -law PCM.

the digital noise power. The curves of Fig. 1 therefore show the gain in  $\hat{s}/n$  due to using weighting as compared to the basic  $\mu$ -law PCM, and the further gains in  $\hat{s}/n$  that soft decision demodulation provide.

When a poor corrector corresponding to a  $\delta$  of 0.4 was used, we achieved a marginally better performance using soft demodulation than when only a fixed weighting profile was used. Improving the corrector, i.e., reducing  $\delta$ , yielded greater gains in  $\hat{s}/n$ . While  $\delta = 0.1$  is relatively easy to achieve, values of  $\delta < 0.01$  require complex corrector algorithms. We observe that for  $\delta = 0.1$  our soft demodulation strategy of individual soft thresholds yielded a gain of 7 dB and 4 dB over conventional and fixed weighted  $\mu$ -law PCM, respectively, over a wide range of channel  $s/n$ .

Reducing the input power level to  $-40$  dB resulted in a severe reduction in  $\hat{s}/n$  for conventional  $\mu$ -law PCM and for fixed-bit weighted  $\mu$ -law PCM. However, by using soft demodulation the  $\hat{s}/n$  was approximately the same as for the  $-17$  dB input level. This can be seen in Fig. 2, where  $M$  was again 4. For  $\delta = 0.1$  the gains in  $\hat{s}/n$  over conventional log-PCM and fixed weighted log-PCM were 19 dB and 13 dB, respectively. Figure 3 shows the effect of varying the number  $M$  of bits subjected to soft demodulation. The input level was  $-17$  dB, and the value of  $\delta$  was 0.1. Extending  $M$  from 2 to 3 is worthwhile, but there is no advantage in using  $M$  in excess of 4. Thus

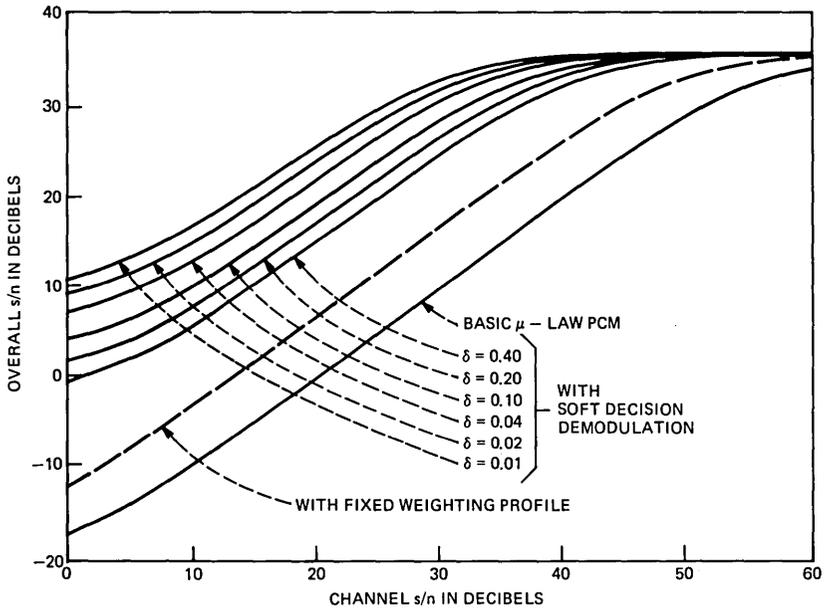


Fig. 2—Variation of overall  $s/n$  as a function of channel  $s/n$  for  $M = 4$ , input power level =  $-40$  dB, 8-bit  $\mu$ -law PCM.

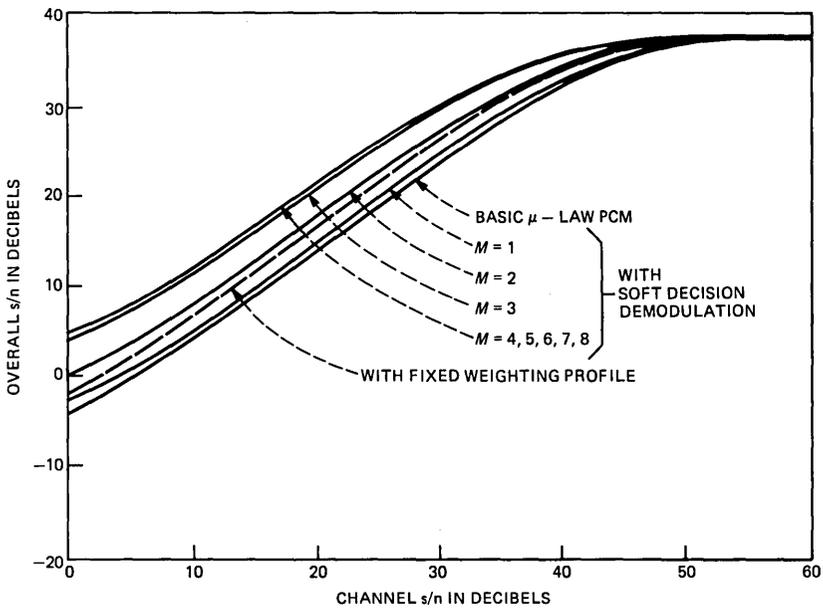


Fig. 3—Variation of overall  $s/n$  as a function of channel  $s/n$  for  $\delta = 0.1$ , input power level =  $-17$  dB, 8-bit  $\mu$ -law PCM.

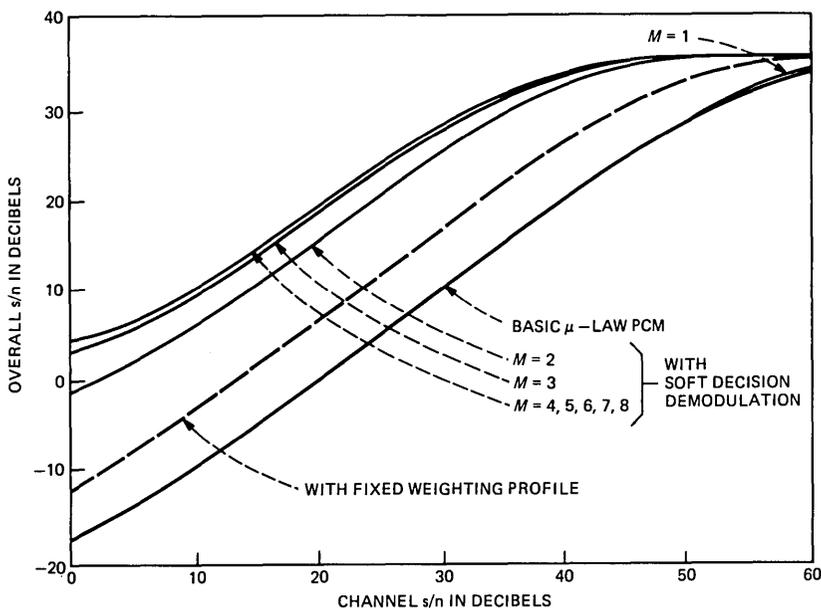


Fig. 4—Variation of overall  $s/n$  as a function of channel  $s/n$  for  $\delta = 0.1$ , input power level = -40 dB, 8-bit  $\mu$ -law PCM.

our gain in  $s/n$  of 7 dB over basic  $\mu$ -law PCM was derived from the application of individual soft demodulation on the first three MSBs. Observe that when soft decision demodulation was applied to only the polarity bit, i.e., the case of  $M = 1$ , the performance was worse than for fixed weighted  $\mu$ -law PCM. It is important to provide soft decision demodulation on both the polarity and most significant magnitude bits ( $M \geq 2$ ). For the lower input power of -40 dB, the curve for  $M = 1$  yielded no gain in  $s/n$ , compared to conventional  $\mu$ -law PCM. The value of  $M$  was required to be 4, as shown in Fig. 4, to achieve the maximum gain in  $s/n$ .

## VI. COMPUTER SIMULATION RESULTS

In our simulations we employed four concatenated speech sentences, two spoken by males, two by females, that were bandlimited between 200 and 3200 Hz, and sampled at 8 kHz. These samples constituted our input speech sequence. The speech sequence was 8-bit  $\mu$ -law PCM,  $\mu = 255$ , encoded, and weighted where appropriate; two-level CPSK modulation ensued. The modulated signal was subjected to Rayleigh fading and contaminated by additive white Gaussian noise. The Rayleigh fading envelope  $C(t)$  was obtained from a hardware simulator of

a frequency-selective Rayleigh-fading mobile radio channel.<sup>11</sup> Different envelope functions were available corresponding to different vehicular speeds when the propagation frequency was 900 MHz. The first  $M$  bits in every  $N$  bits were initially examined. If the  $l$ th bit in the  $k$ th word satisfied

$$C(k)B_l(k) + I(k) \geq |Z|, \quad (31)$$

the bit was regenerated according to

$$\hat{B}_l(k) = \begin{cases} C(k)B_l(k) + I(k) \geq Z; & \text{logical 1 generated} \\ C(k)B_l(k) + I(k) \leq -Z; & \text{logical 0 generated} \end{cases} \quad (32)$$

for  $l = 1, 2, \dots, M$ , where  $I(k)$  represented the additive interference level. If

$$C(k)B_l(k) + I(k) < |Z|; \quad \text{ERASURE} \quad (33)$$

and complete word was rejected. If no erasure occurred during the examination of the first  $M$  bits, the last  $(N - M)$  bits were regenerated according to

$$\hat{B}_i(k) = \begin{cases} C(k)B_i(k) + I(k) \geq 0; & \text{logical 1 generated} \\ C(k)B_i(k) + I(k) < 0; & \text{logical 0 generated} \end{cases} \quad (34)$$

for  $i = N - M, \dots, N - 1, N$ .

The bit regenerating process, therefore, examined the first  $M$  bits, and if the input signal level  $C(k)B_l(k) + I(k)$  resided outside the erasure level for  $l = 1, 2, \dots, M$ , the bits were regenerated as prescribed by inequalities (32) and (34). Those  $N$ -bit words so formulated were  $\mu$ -law PCM decoded into samples. However, when an erasure of one of the first  $M$  bits in any word occurred, the entire word was rejected, and consequently there was no corresponding decoded sample. Samples so discarded by the soft decision demodulation process were replaced by means of prediction and/or interpolation using those samples that had been  $\mu$ -law PCM decoded. Various complex prediction and interpolation algorithms were available to us,<sup>12-14</sup> but we opted for a simple scheme. Knowing that interpolation is superior to prediction if there are two correctly received samples adjacent to the rejected sample, we replaced the discarded sample  $x_r$  by means of nearest neighbor averaging

$$\hat{x}_r = 0.5(\hat{x}_{r-1} + \hat{x}_{r+1}), \quad (35)$$

where  $\hat{x}_{r-1}$  and  $\hat{x}_{r+1}$  were the decoded samples in the  $(r - 1)$ -th and  $(r + 1)$ -th sampling instants, respectively. When the words at both the  $(r - 1)$ -th and  $(r + 1)$ -th sampling instants were designated to be

soft,  $\hat{x}_r$  was generated by prediction from the previously accepted correct sample, viz:

$$\hat{x}_r = \rho \hat{x}_{r-1}, \quad (36)$$

where  $\rho$  is the normalized first-order autocorrelation coefficient given by

$$\rho = \frac{E\{\hat{x}_r \hat{x}_{r+1}\}}{S^2}. \quad (37)$$

Figure 5 shows the variation of overall s/n as a function of channel s/n when 8-bit  $\mu$ -law PCM encoded speech was conveyed by two-level CPSK modulation over a Rayleigh fading mobile radio channel. The input speech level was  $-17$  dB. The threshold values used in our experiments were the optimal ones, assuming a particular value of  $\delta$ , and a given channel s/n. Thus Fig. 5 displays a collection of curves that use threshold values associated with  $\delta = 0.01, 0.02, 0.04, 0.1, 0.2,$  and  $0.4$ . The curves shown in Figs. 1 and 5 are in close agreement over the usable range of channel s/n, typically when the channel s/n exceeds 20 dB. Our corrector tends to operate almost exclusively using nearest neighbor averaging when the number of rejected samples is small. As the channel s/n decreases, eq. (36) is utilized more frequently. Thus the  $\delta$  associated with our corrector increases as the channel s/n falls,

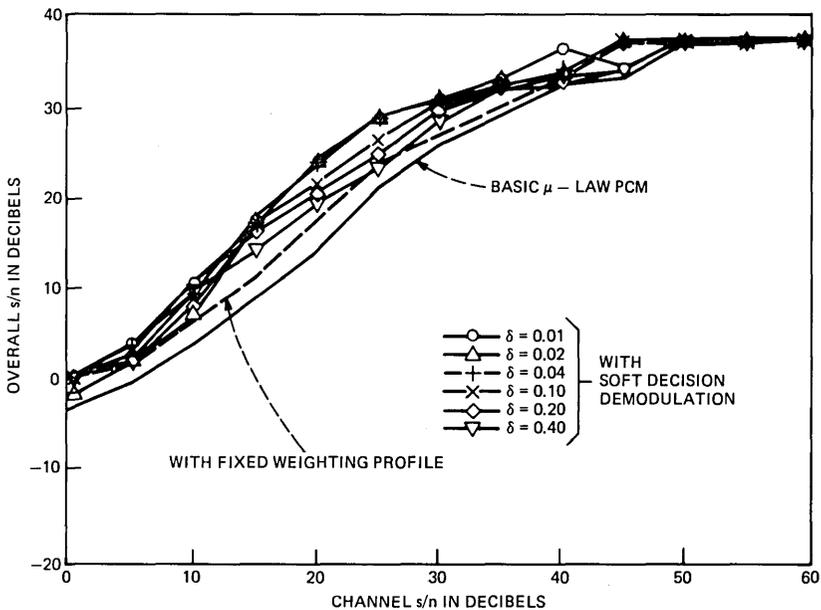


Fig. 5—Simulation of mobile radio channels with a speech input signal. Variation of overall s/n as a function of channel s/n for  $M = 4$ , input level of  $-17$  dB.

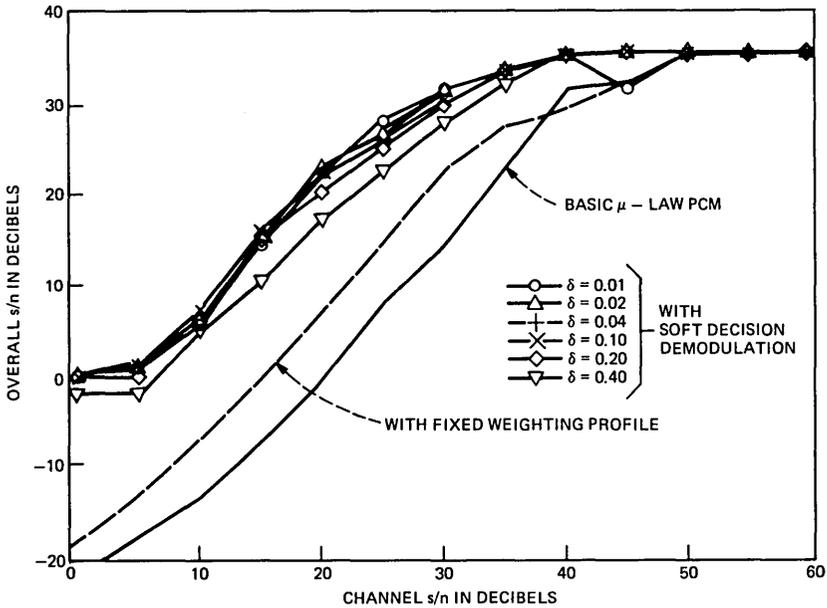


Fig. 6—Simulation of mobile radio channels with a speech input signal. Variation of overall  $s/n$  as a function of channel  $s/n$  for  $M = 4$ , input power level =  $-40$  dB.

i.e., the quality of the corrector process decreases with decreasing channel  $s/n$ . Figure 6 shows the situation when the input speech level is reduced to  $-40$  dB, and again we observe a correspondence between the curves in this figure and those in Fig. 2.

Because our corrector has a value of  $\delta$  that depends upon the channel  $s/n$ , employing either interpolation [eq. (35)] or prediction [eq. (36)], it is difficult to calculate the optimum thresholds  $T_i$ . However, over the range of channel  $s/n$  that yield an acceptable overall  $s/n$ ,  $\hat{s}/n$ , it appears appropriate to use the values of  $T_i$  that correspond to  $\delta$  of 0.1.

## VII. DISCUSSION

New theory has been presented for the use of soft decision demodulation applied to  $\mu$ -law PCM signals transmitted over Rayleigh fading channels using CPSK modulation. Each bit in the  $\mu$ -law PCM word has been assigned its individual erasure threshold, and these thresholds have been calculated for high and low input power levels, as a function of the channel  $s/n$ , and for correctors of differing quality. These results are displayed in tabular form in Tables I through IV, and in a graphical format in Figs. 1 through 4. In our theory we do not specify the corrector algorithms, i.e., those methods of replacing samples discarded by the soft decision demodulator, because the num-

ber of possible correctors is legion. Instead we calculate the individual erasure thresholds for specified values of mean square error power  $\delta$  that is associated with the correction process. We acknowledge that this technique does have disadvantages when the actual correction algorithm employed is dynamically dependent on which samples are being corrected. In this later case it is desirable to adjust the values of the thresholds as the corrector algorithm varies if near optimum performance is to be maintained.

From our theoretical deliberations we observed the following salient points:

1. There is no advantage to be gained by applying soft decision demodulation to more than 3 or 4 of the most significant bits of the 8-bit  $\mu$ -law PCM,  $\mu = 255$ , words (see Figs. 3 and 4).

2. We used two bench markers to measure the performance of our soft decision demodulation scheme. These were the basic  $\mu$ -law PCM, and the fixed-profile weighted  $\mu$ -law PCM. At the high input power level of  $-17$  dB (which corresponded to a peak  $\hat{s}/n$  in the absence of transmission errors) the gain in  $\hat{s}/n$  due to soft decision demodulation using  $\delta = 0.1$  over basic  $\mu$ -law PCM and weighted  $\mu$ -law PCM was 7 and 3 dB, respectively, when the channel  $s/n$  was 30 dB. Observe that with the soft decision demodulation the  $\hat{s}/n$  was 30 dB, a value commensurate with toll-quality speech. For the lower input level of  $-40$  dB, a level that has some correspondence with unvoiced speech, the corresponding gains in  $\hat{s}/n$  were 18 and 12 dB, respectively. Figure 4 shows that soft decision demodulation is particularly important at the low input level, achieving an  $\hat{s}/n$  of 28 dB for a channel  $s/n$  of 30 dB. Indeed, the  $\hat{s}/n$  deteriorated by only 8 dB, the same value observed for the higher input level, when the channel  $s/n$  decreased from the ideal channel condition ( $\approx 60$  dB) to the lower channel  $s/n$  of 30 dB.

3. Provided  $M = 4$ , soft decision demodulation always gave a gain in  $\hat{s}/n$  compared to our reference systems, even when  $\delta = 0.4$ . The ability that soft decision demodulation has to maintain its performance for both high and low input levels is strikingly demonstrated in Figs. 1 and 2. Observe that improving the corrector's performance by a factor of ten compared to  $\delta = 0.1$  caused the  $\hat{s}/n$  to increase by 5 dB when the channel  $s/n$  was 30 dB.

4. Four concatenated speech sentences were used in our simulations. The speech was  $\mu$ -law PCM encoded prior to CPSK modulation, and the Rayleigh fading channel was obtained from a real-time simulator. The results, displayed in Figs. 5 and 6, demonstrate the advantage of using soft decision demodulation. We observed gains in  $\hat{s}/n$  of 6 and 4 dB compared to the basic  $\mu$ -law PCM and the fixed weighted  $\mu$ -law PCM, respectively, when the input power was  $-17$  dB and the channel  $s/n$  was 30 dB. The corresponding gains when the input power was

-40 dB were 17 and 9 dB, respectively. Although the corrector used in our simulations could switch its algorithm [see eqs. (35) and (36)], for the gains in  $s/\hat{n}$  quoted above eq. (35) was used most of the time, and  $\delta$  was approximately 0.1.

From our theoretical results and simulations we conclude that the application of soft decision demodulation, where each bit in the  $\mu$ -law PCM word is assigned its unique threshold value, does offer significant advantages in  $s/\hat{n}$  when the transmission is by CPSK modulation over Rayleigh fading channels. The gains in  $s/\hat{n}$  are achieved by only a marginal increase in complexity at the receiver, while the transmitter is unchanged. Careful optimization of a combination of soft decision demodulation at the receiver with a weighting strategy at the transmitter leads us to expect further improvements in  $s/\hat{n}$ , although at the expense of added system complexity.

## REFERENCES

1. R. Steele, C.-E. Sundberg, and W. C. Wong, "Transmission Errors in Companded PCM Over Gaussian and Rayleigh Fading Channels," *AT&T Bell Lab. Tech. J.*, 63, No. 6, Part 1 (July-August 1984), pp. 955-90.
2. C.-E. Sundberg and N. Rydbeck, "Pulse Code Modulation With Error-Correcting Codes for TDMA Satellite Communication Systems," *Ericsson Technics*, 31, No. 1 (1976), pp. 3-56.
3. W. C. Wong, R. Steele, B. Glance, and D. Horn, "Time Diversity With Adaptive Error Detection to Combat Rayleigh Fading in Digital Mobile Radio," *IEEE Trans. Commun. COM-31*, No. 3 (March 1983), pp. 378-87.
4. W. C. Jakes, ed., *Microwave Mobile Communications*, New York: J. Wiley & Sons, 1974.
5. W. C. Wong, C.-E. Sundberg, and R. Steele, "Weighting Strategies for Companded PCM Transmitted Over Rayleigh Fading and Gaussian Channels," *AT&T Bell Lab. Tech. J.*, 63, No. 4 (April 1984), pp. 587-626.
6. E. Bedrosian, "Weighted PCM," *IRE Trans. Inform. Theory*, IT-4 (March 1958), pp. 45-9.
7. C.-E. Sundberg, "Optimum Weighted PCM for Speech Signals," *IEEE Trans. Commun., COM-26*, No. 6 (June 1978), pp. 872-81.
8. C.-E. Sundberg, "Soft Decision Demodulation for PCM Encoded Speech Signals," *IEEE Trans. Commun., COM-26*, No. 6 (June 1978), pp. 854-9.
9. C.-E. Sundberg, "Algorithms for Reducing the Effect of Transmission Errors in PCM Encoded Speech Signals by Means of Soft Demodulation Techniques," *Conf. Rec. Toronto, Canada, ICC 78*, June 1978, pp. 8.4.1-5.
10. N. S. Jayant, R. Steele, N. W. Chan, and C. E. Schmidt, "On Soft Decision Demodulation for PCM and DPCM-Encoded Speech," *IEEE Trans. Commun., COM-28*, No. 2 (February 1980), pp. 308-11.
11. W. F. Bodtmann and H. W. Arnold, "Fade Duration Statistics of a Rayleigh-Distributed Wave," *IEEE Trans. Commun., COM-30* (March 1982), pp. 549-53.
12. R. Steele and F. Benjamin, "Sample Reduction and Subsequent Adaptive Interpolation of Speech Signals," *B.S.T.J.*, 62, No. 6, Part 1 (July-August 1983), pp. 1365-98.
13. H. S. Hon and H. C. Andrews, "Cubic Splines for Image Interpolation and Digital Filtering," *IEEE Trans. Acoust. Speech Signal Proc., ASSP-26*, No. 6 (December 1978), pp. 508-17.
14. M. V. Mathews, "Extremal Coding for Speech Transmission," *IRE Trans. Info. Theory*, IT-5 (September 1959), pp. 129-36.

## AUTHORS

**Raymond Steele**, (SM '80), B.S. (Electrical Engineering) from Durham University, Durham, England, in 1959, the Ph.D. degree in 1975, and the

Doctor of Science Degree (DSc) in 1983. Prior to his enrollment at Durham University, he was an indentured apprenticed Radio Engineer. After research and development posts at E. K. Cole Ltd., Cossor Radar and Electronics, Ltd., and The Marconi Company, all in Essex, England, he joined the lecturing staff at the Royal Naval College, Greenwich, London, England. Here he lectured in telecommunications to NATO and the External London University degree courses. His next post was as Senior Lecturer in the Electronic and Electrical Engineering Department of Loughborough University, Loughborough, Leics., England, where he directed a research group in digital encoding of speech and television signals. In 1975 his book, *Delta Modulation Systems* (Pentech Press, London), was published. He was a consultant to the Acoustics Research department at Bell Laboratories in the summers of 1975, 1977, and 1978, and in 1979 he joined the company's Communications Methods Research department, Crawford Hill Laboratory, Holmdel, N.J. In 1983 he became Professor of Communications in the Electronics department at the University of Southampton, England.

**Carl-Erik W. Sundberg**, M.S.E.E., 1966, and Dr. Techn., 1975, Lund Institute of Technology, University of Lund, Sweden; University of Lund, 1963–1984; AT&T Bell Laboratories, 1984—. Mr. Sundberg was an Associate Professor in the Department of Telecommunication Theory, University of Lund, from 1977 to 1984 and a consultant in his field. During 1976 he was with the European Space Research and Technology Centre (ESTEC), Noordwijk, The Netherlands, as an ESA Research Fellow. He has been a Consulting Scientist at LM Ericsson and SAAB-SCANIA, Sweden, and at AT&T Bell Laboratories. His research interests include source coding methods, channel coding (especially decoding techniques), digital modulation methods, fault-tolerant systems, digital mobile radio systems, spread spectrum systems, and digital satellite communication systems. He has published a large number of papers in these areas during the last few years. Senior Member, IEEE; member, SER, Sveriges Elektroingenjörers Riksförening.

**Wai Choong Wong**, B.Sc. (Electronic and Electrical Engineering), Ph.D. (Electronic Engineering), Loughborough University of Technology, Loughborough, England, in 1976 and 1980, respectively; AT&T Bell Laboratories, 1980–1983; National University of Singapore, 1983—. At AT&T Bell Laboratories, Mr. Wong was a Member of Technical Staff in the Communications Methods Research department. He is now a lecturer at the National University in Singapore. His current interest is in speech signal processing, particularly applied to mobile radio environment, digital modulation techniques, and simultaneous transmission of data and speech signals.



## Linear Equalization Theory in Digital Data Transmission Over Dually Polarized Fading Radio Channels

By N. AMITAY and J. SALZ\*

(Manuscript received May 16, 1984)

A theory of linear least-mean-squares equalization in digital data communications operating over two coupled linear dispersive channels, with particular application to dually polarized terrestrial radio systems is presented. We jointly optimize transmitter and receiver matrix filters when the inputs are two independent quadrature-amplitude-modulated data signals of fixed total average power. Formulas for minimum total mean-square error, upper bounds on probability of error, matched filter bounds, and the Shannon limit are provided. Using this theory in conjunction with a proposed propagation model for the dually polarized radio channel, we provide estimates of probability distributions of the data rates that can be supported by the optimum receiver structures as well as the jointly optimized transmitter-receiver. The distribution of data rates predicted from the Shannon limit as well as from the matched filter bounds are also given. For the dually polarized radio channel, we find that linear equalization structures can practically eliminate the effects of cross-polarization interference contributed by the antennas and their feeds. At very low outage probabilities, the combined-channels data rate achieved with the jointly optimized structures is about 4 b/s/Hz less than predicted from their matched filter bounds. Also, the combined data rate associated with the Shannon limit is about 3 b/s/Hz higher than predicted by the matched filter bound at very low outage probabilities. Sensitivity analyses reveal that about 1 b/s/Hz can be gained for each 3-dB increase in carrier-to-noise ratio. For each order-of-magnitude decrease in the required error rate, about 0.3 b/s/Hz is given up.

---

\* Authors are employees of AT&T Bell Laboratories.

---

Copyright © 1984 AT&T. Photo reproduction for noncommercial use is permitted without payment of royalty provided that each reproduction is done without alteration and that the Journal reference and copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free by computer-based and other information-service systems without further permission. Permission to reproduce or republish any other portion of this paper must be obtained from the Editor.

## I. INTRODUCTION

Communications via dually polarized radio channels is an effective means of reusing existing bandwidth. It is well known that one can transmit and receive two orthogonally polarized waves, e.g., linear, circular, or elliptical. Information can be conveyed by independently modulating the signals carried by each of these waves. In conventional terrestrial microwave radio systems, linear orthogonal polarizations—i.e., horizontal and vertical—are usually employed.

Digital communications over a single radio channel (single polarization) is adversely affected by multipath fading, which causes Inter-symbol Interference (ISI). In dually polarized digital radio communications, one also has to contend with Cross-Polarization Interference (CPI). This interference, which is dispersive in general, is due to antenna misalignments, imperfect waveguide feeds, fading, and other propagation path anomalies. In order to utilize the dually polarized radio channels to their maximum capacity, it is necessary to mitigate the deleterious effects of ISI and CPI.

Various approaches to this problem have been suggested.<sup>1-16</sup> Some of them address the problem of CPI cancellation only, some minimize CPI and noise, and others treat nondispersive channels only. Often these methods are based on cross-channel frequency response estimations, thereby facilitating at least a partial cancellation of CPI. The chief purpose of this paper is to establish a sound theoretical base for realizing effective algorithms and devices for the simultaneous and optimal compensation of cross-polarization and intersymbol interferences in the presence of noise.

The dually polarized channel is modeled as a  $2 \times 2$  complex matrix with frequency-dependent elements followed by an additive noise vector. The diagonal entries represent the cochannel (in-line) frequency transfer characteristics, while the off-diagonal terms represent the cross-channel transfer characteristics. The four entries in this matrix channel can be viewed as a sample from a random collection of frequency functions. The presence of noise implies that each matrix channel in the collection is limited to a maximum data rate, for a given error rate and specific communication method. Because of the random nature of the matrix channel, a meaningful performance measure for a particular communication method is the probability distribution of data rates that can be supported at a certain error rate objective.

Figure 1 depicts the problem considered. Two data sources are first passed through a matrix transmitter filter. The diagonal entries control the cochannel signal shaping, while the off-diagonal terms control the cross-channel signal shaping as well as a possible power distribution (transfer) between the two channels. The two emerging output

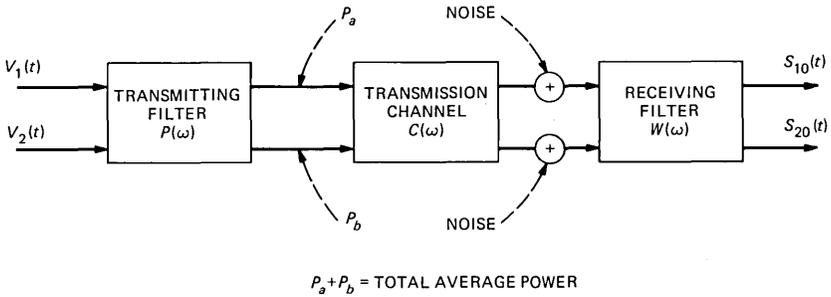


Fig. 1—Dually polarized radio transmission.

signals, represented by a vector, are then transmitted over the matrix channel. At the receiver, independent noise sources are added, and the resultant vector signal is processed by a linear matrix filter, which acts as a linear equalizer/canceler.

We pose these general questions. How does one choose the receiver matrix filter so as to minimize a reasonable cost function? Next, how does one choose an optimum transmitter matrix filter under a reasonable power constraint? Finally, how does the performance of these optimum structures compare with the ultimate performance predicted by information theory (Shannon limit)?

After a discussion of the channel model in the next section, we describe the modulation and the transmitter models in Section III. In Section IV we discuss the receiver optimization problem, while in Section V we derive the theory for joint transmitter-receiver optimization. The Matched Filter (MF) bounds on performance are derived in Section VI, and in Section VII the information theory limit is developed. Description and discussion of the propagation model advanced in this paper can be found in Section VIII. In Section IX we give numerical evaluations and discuss the results. Section X contains our summary and conclusions.

## II. THE CHANNEL MODEL

A general linear, baseband-equivalent, dually polarized transmission channel can be characterized by the matrix impulse response

$$c(t) = \begin{pmatrix} c_{11}(t) & c_{21}(t) \\ c_{12}(t) & c_{22}(t) \end{pmatrix}, \quad (1)$$

or its Fourier transform

$$C(\omega) = \int_{-\infty}^{\infty} c(t)e^{i\omega t} dt = \begin{pmatrix} C_{11}(\omega) & C_{21}(\omega) \\ C_{12}(\omega) & C_{22}(\omega) \end{pmatrix}, \quad (2)$$

which is the frequency transfer characteristic of the transmission medium. The diagonal entries in the matrix  $c(t)$  are the complex-valued impulse responses of the co-polarized channels, while the off-diagonal entries are the complex-valued cross-polarized channel impulse responses. To complete the channel characterization, we assume that the noises at the two output ports are additive white Gaussian processes represented by the complex vector,

$$\nu(t) = \begin{pmatrix} \nu_1(t) \\ \nu_2(t) \end{pmatrix}, \quad (3)$$

and characterized by the covariance matrix

$$E[\nu(t_1)\nu^\dagger(t_2)] = N_o\delta(t_1 - t_2)I, \quad (4)$$

where  $\nu_1(t)$  and  $\nu_2(t)$  are baseband-equivalent, zero-mean, complex Gaussian processes with common double-sided spectral density,  $N_o$ . The notation,  $E[\cdot]$ , stands for mathematical expectation,  $\delta(\cdot)$  is the Dirac delta function,  $\dagger$  stands for conjugate transpose, and  $I$  is the identity matrix. This channel model is illustrated schematically in Fig. 2.

### III. MODULATION AND TRANSMITTER MODELS

Our investigation focuses on the class of efficient linear modulation methods known as Quadrature Amplitude Modulation (QAM). A convenient representation of these modulated signals is

$$s(t) = \text{Re}\{(\exp(i2\pi f_o t) \sum_n a_n g(t - nT))\}, \quad (5)$$

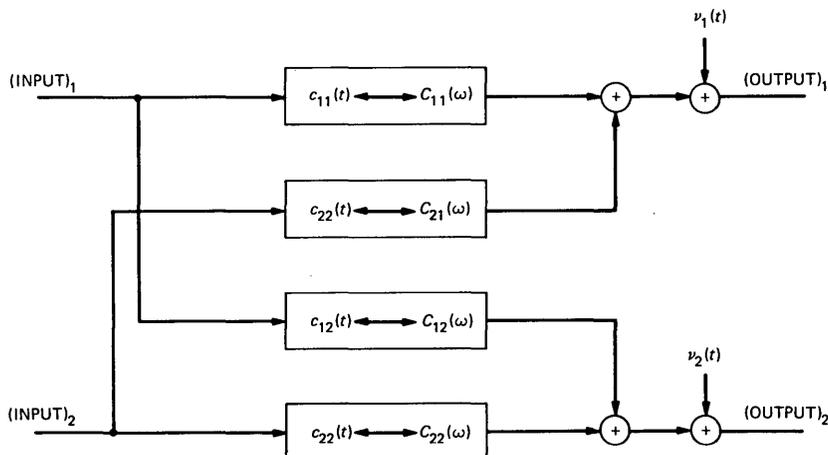


Fig. 2—Linear channel model for dually polarized radio transmission.

where  $\text{Re}(\cdot)$  stands for the “real part” and  $f_o$  is the carrier frequency. The data symbols,  $\{a_n\}$ , transmitted at  $T$ -second intervals take on values on a two-dimensional lattice and are therefore complex valued. In general, the pulse  $g(t)$  is also complex valued.

In a dually polarized communications system, two independent QAM data signals as individually represented in (5) are transmitted, one on the horizontal channel and the other on the vertical channel. In this situation, the transmitted data signal is modeled as a two-dimensional vector

$$V(t) = \begin{pmatrix} V_1(t) \\ V_2(t) \end{pmatrix} = \text{Re} \left\{ \exp(i2\pi f_o t) \sum_n p(t - nT) A_n \right\}, \quad (6)$$

where

$$A_n = \begin{pmatrix} a_n \\ b_n \end{pmatrix} = \begin{pmatrix} a_{nr} + ia_{nl} \\ b_{nr} + ib_{nl} \end{pmatrix} \quad (7)$$

is the  $n$ th pair of data symbols and  $p(t)$  is the transmitter matrix filter impulse response,

$$p(t) = \begin{pmatrix} p_{11}(t) & p_{21}(t) \\ p_{12}(t) & p_{22}(t) \end{pmatrix}. \quad (8)$$

The diagonal entries in (8) are cochannel transmitter impulse responses acting as signal shapers, while the off-diagonal impulse responses act as cross-channel signal shapers as well as power distributors.

For two independent input data streams, the covariance matrix of the data symbols is diagonal,

$$E[A_n A_n^\dagger] = \begin{pmatrix} E|a_n|^2 & 0 \\ 0 & E|b_n|^2 \end{pmatrix} = \begin{pmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_b^2 \end{pmatrix}, \quad (9a)$$

and  $E[A_n A_m^\dagger] = 0$ ,  $n \neq m$ . For a rectangular lattice consisting of odd positive and negative integers,

$$\sigma_a^2 = 2 \left( \frac{L^2 - 1}{3} \right), \quad (9b)$$

where  $L^2$  is the total number of points in the constellation. We now refer to the cochannels as channel “ $a$ ” and channel “ $b$ ”, respectively. By use of (8) and (9), the average transmitted power in each of the channels is, by a straightforward calculation,

$$P_a = \frac{\sigma_a^2}{T} \int_{-\infty}^{\infty} (|p_{11}(t)|^2 + |p_{21}(t)|^2) dt$$

and

$$P_b = \frac{\sigma_d^2}{T} \int_{-\infty}^{\infty} (|p_{22}(t)|^2 + |p_{12}(t)|^2) dt. \quad (10)$$

The total average transmitted power is, therefore,

$$\begin{aligned} P_{av} = P_a + P_b &= \frac{\sigma_d^2}{T} \int_{-\infty}^{\infty} \text{trace}[p^\dagger(t)p(t)] dt \\ &= \frac{\sigma_d^2}{T} \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}[P^\dagger(\omega)P(\omega)] d\omega, \end{aligned} \quad (11)$$

where  $P(\omega)$  is the Fourier transform of  $p(t)$ . The quantity  $P_{av}$  in (11), the total average power, is constrained to be fixed throughout this work.

#### IV. THE OPTIMUM RECEIVER PROBLEM

Let  $H(\omega)$  be the overall matrix transfer function of the transmitter filter in cascade with the channel  $C$ ,

$$H(\omega) = C(\omega)P(\omega). \quad (12a)$$

Consequently, the equivalent baseband signal plus noise vector at the receiver can be written

$$S(t) = \sum_n H(t - nT)A_n + \nu(t). \quad (12b)$$

If  $W(t)$  denotes the impulse response of the receiver matrix filter, as shown in Fig. 1, the output vector signal is then represented as

$$\begin{aligned} S_o(t) &= \sum_n \int_{-\infty}^{\infty} W(t - \tau)H(\tau - nT)A_n d\tau \\ &+ \int_{-\infty}^{\infty} W(t - \tau)\nu(\tau) d\tau = \begin{pmatrix} s_{10}(t) \\ s_{20}(t) \end{pmatrix} \end{aligned} \quad (13)$$

and the sampled vector output at  $t = 0$  is  $S_o(0)$ . This is a representative sample since it possesses the same statistics as the  $n$ th sample because of stationarity.

Next, we define the error vector  $\epsilon$  as the difference between  $S_o(0)$  and the vector data symbol  $A_o$ ,

$$\epsilon = S_o - A_o = \begin{pmatrix} s_{10}(0) - a_o \\ s_{20}(0) - b_o \end{pmatrix},$$

and we form the average-squared error matrix,

$$\left. \begin{aligned} \xi &= E[\epsilon\epsilon^\dagger] \\ \text{MSE} &= \text{MSE}_1 + \text{MSE}_2 = \text{trace } \xi \end{aligned} \right\}. \quad (14)$$

The diagonal elements of  $\xi$  are the individual mean-square errors of the two channels,  $\text{MSE}_1 = E |s_{10}(0) - a_o|^2$  and  $\text{MSE}_2 = E |s_{20}(0) - b_o|^2$ . Substituting (13) into (14) and making use of (4) and (9), we get

$$\begin{aligned} \xi = \sigma_d^2 \left\{ I - \int_{-\infty}^{\infty} W(-\tau)H(\tau)d\tau - \int_{-\infty}^{\infty} [W(-\tau)H(\tau)]^\dagger d\tau \right. \\ \left. + \sum_n \int_{-\infty}^{\infty} W(-\tau)H(\tau - nT)d\tau \int_{-\infty}^{\infty} [W(-\tau)H(\tau - nT)]^\dagger d\tau \right. \\ \left. + N_o \int_{-\infty}^{\infty} W(-\tau)W^\dagger(-\tau)d\tau \right\}. \end{aligned} \quad (15)$$

It should be noted that  $\text{MSE}_1$  is a function of  $W_{11}$  and  $W_{21}$  only while  $\text{MSE}_2$  is a function of  $W_{12}$  and  $W_{22}$  only. Our objective is to minimize MSE, expressed in (14), with respect to the elements of the filter matrix  $W$ , i.e., to determine the elements  $W_{jl}$ ;  $j, l = 1, 2$ , which minimize MSE in (14). From the previous discussion it is clear that  $\text{MSE}_1$  and  $\text{MSE}_2$  are independent of one another. Therefore, minimization of (14) is tantamount to the independent minimization of  $\text{MSE}_1$  and  $\text{MSE}_2$ .

The justification for using the quadratic cost function, MSE, is due to its mathematical simplicity and tractability. Moreover, as will be seen later, minimizing this cost implies minimizing a reasonable upper bound on error rate. Additionally, its use is practically motivated, since it lends itself to an easy estimation method often used to update adaptive equalizer tap coefficients in practical systems.

Since trace  $\xi$  is a positive quadratic form in  $W(t)$ , we set its gradient equal to zero and are thus guaranteed to obtain a global stationary point. Proceeding with the calculus of variation, we replace the matrix  $W$  by

$$W_o + \begin{pmatrix} \epsilon_{11} \delta W_{11} & \epsilon_{21} \delta W_{21} \\ \epsilon_{12} \delta W_{12} & \epsilon_{22} \delta W_{22} \end{pmatrix} \quad (16a)$$

and set

$$\begin{pmatrix} \frac{\partial}{\partial \epsilon_{11}} \text{trace } \xi & \frac{\partial}{\partial \epsilon_{21}} \text{trace } \xi \\ \frac{\partial}{\partial \epsilon_{12}} \text{trace } \xi & \frac{\partial}{\partial \epsilon_{22}} \text{trace } \xi \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad (16b)$$

at  $\epsilon_{11} = \epsilon_{21} = \epsilon_{12} = \epsilon_{22} = 0$ .

After considerable detailed calculations, summarized in Appendix A, we get an integral equation for the optimum matrix filter,  $W_o(t)$ :

$$W_o(-\tau) = \frac{1}{\sigma^2} H^\dagger(\tau) - \frac{1}{\sigma^2} \sum_n U_n H^\dagger(\tau - nT) \quad (17a)$$

with

$$U_n = \int_{-\infty}^{\infty} W_o(-t)H(t - nT)dt, \quad \sigma^2 = N_o/\sigma_d^2. \quad (17b)$$

Equation (17) can be interpreted as a representation of the optimum filter,  $W_o(t)$ , by a matrix matched filter, with impulse response  $H^\dagger(-t)$ , followed by a sampler at  $t_n = nT$  and a matrix transversal filter, or a nonrecursive matrix digital filter, with matrix tap coefficients  $U_n$ . The matrix representation in (17) is symbolically similar to the scalar case,<sup>17</sup> with matrices playing the role of scalars. When the channel matrix impulse response (1) is diagonal, i.e., cross-polarization interference is absent, (17) reduces to two scalar equations representing the two optimum cochannel equalizers. In the general case, (17) represents a cross-coupled transversal structure admitting equalization as well as cancellation.

When (17) is postmultiplied by  $W_o^\dagger(-t)$ , integrated, and then substituted into (15), we get

$$\xi_o = \sigma_d^2 \left[ I - \int_{-\infty}^{\infty} W_o(-t)H(t)dt \right] = \sigma_d^2(I - U_o) \quad (18a)$$

and

$$\text{MSE}_o = \text{Min}_W \text{trace}\{\xi\} = \text{trace}\{\xi[W_o(t)]\} = \text{trace}\{\xi_o\}. \quad (18b)$$

Furthermore, explicit equations for the optimum matrix tap coefficients,  $U_n$ , are obtained by postmultiplying (17) by  $H(\tau - kT)$ , for all  $k$ , and then integrating. The result is a set of linear equations in the matrices  $U_n$ ,

$$\sigma^2 U_k = R_k - \sum_n U_n R_{k-n}, \quad \text{all } k, \quad (19)$$

where

$$R_k = \int_{-\infty}^{\infty} H^\dagger(\tau)H(\tau - kT)d\tau = R_{-k}^\dagger.$$

The discrete matrix convolution equation (19) has a simple unique solution obtained by taking the Fourier series of both sides. Defining

$$\begin{pmatrix} U(\omega) \\ R(\omega) \end{pmatrix} = \sum_n \begin{pmatrix} U_n \\ R_n \end{pmatrix} e^{-i\omega nT} \quad (20)$$

with

$$\begin{pmatrix} U_n \\ R_n \end{pmatrix} = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} \begin{pmatrix} U(\omega) \\ R(\omega) \end{pmatrix} e^{i\omega nT} d\omega \quad (21)$$

and substituting (21) into (19), we obtain

$$U(\omega) = R(\omega)[\sigma^2 + R(\omega)]^{-1}. \quad (22)$$

Again, (19) and (22) are strikingly similar in structure to the case obtained in the scalar situation,<sup>17</sup> with the proviso that matrices play the role of scalars and with the obvious condition that the matrices do not necessarily commute.

We now express (18) more explicitly by obtaining  $U_o$  from (21) and (22)

$$\xi_o = \sigma_d^2 \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} \left[ I + \frac{R(\omega)}{\sigma^2} \right]^{-1} d\omega$$

and

$$\text{MSE}_{U_o} = \text{trace}\{\xi_o\}. \quad (23)$$

The connection between  $R(\omega)$ , given in (20) and (21), and the matrix channel transfer function  $H(\omega)$  is made by equating  $R_n$  from (21) to the definition in (19),

$$\begin{aligned} R_n &= \int_{-\infty}^{\infty} H^\dagger(\tau)H(\tau - nT)d\tau = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} R(\omega)e^{-i\omega nT} d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} H^\dagger(\omega)H(\omega)e^{-i\omega nT} d\omega \\ &= \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} \sum_k H^\dagger\left(\omega - \frac{2\pi k}{T}\right) H\left(\omega - \frac{2\pi k}{T}\right) e^{-i\omega nT} d\omega. \end{aligned} \quad (24)$$

Equation (24) implies that

$$R(\omega) = \frac{1}{T} \sum_k H^\dagger\left(\omega - \frac{2\pi k}{T}\right) H\left(\omega - \frac{2\pi k}{T}\right). \quad (25)$$

This is recognized as the aliased, or folded, matrix spectrum about the Nyquist frequency  $\pi/T$ . For the nonexcess bandwidth transmission case, i.e.,

$$H(\omega) = 0 \quad \text{for} \quad |\omega| \geq \pi/T,$$

eq. (25) reduces to

$$R(\omega) = \frac{1}{T} H^\dagger(\omega)H(\omega). \quad (26)$$

At this point we introduce simplifying normalizations and perform some sanity checks. Focusing attention on the nonexcess bandwidth case with constant and identical transmitter filters, i.e.,

$$\begin{aligned} P(\omega) = P(z) &= \begin{pmatrix} K & 0 \\ 0 & K \end{pmatrix}; \\ |\omega| \leq \pi/T \quad \text{or} \quad |z| \leq \frac{1}{2} \quad \text{with} \quad H &= KC(\omega), \end{aligned} \quad (27)$$

where the normalized frequency variable is

$$z = \frac{T}{2\pi} \omega = fT, \quad (28)$$

we find from (11) that the average power per channel is

$$P_a = P_b = K^2 \sigma_d^2 / T^2 = P_{av} / 2. \quad (29)$$

The quantity

$$\sigma^2 T = \frac{N_o T}{\sigma_d^2} = \frac{N_o / T}{P_a} K^2 = \frac{K^2}{\rho}, \quad (30)$$

where  $\rho$  is the clear air carrier-to-noise ratio (CNR) in each channel, being typically  $2 \times 10^6$ , or 63 dB. Applying the normalization defined in (26) through (30), we rewrite (23) as

$$\xi_o = \sigma_d^2 \int_{-1/2}^{1/2} [I + \rho C^+(z)C(z)]^{-1} dz$$

and

$$\text{MSE}_o = \sigma_d^2 \int_{-1/2}^{1/2} \text{trace}[I + \rho C^+(z)C(z)]^{-1} dz. \quad (31)$$

Again, it is evident that when  $H$  in (27) is diagonal (no cross-polarization interference), (31) reduces to the scalar case.<sup>18</sup> As we expected, when  $\sigma^2 \rightarrow 0$ , ( $\rho \rightarrow \infty$ ),  $\text{MSE}_o \rightarrow 0$ ; and when  $\sigma^2 \rightarrow \infty$ , ( $\rho \rightarrow 0$ ),  $\text{MSE}_o \rightarrow 2\sigma_d^2$ .

In the next section we use the closed-form expression for  $\text{MSE}_o$  in (31) to optimize the transmitter matrix filter.

## V. THE TRANSMITTER OPTIMIZATION PROBLEM

For the general transmitter matrix filter,  $P(\omega)$ , the optimization problem has proved thus far to be intractable. This is in contrast to the scalar case where a general solution exists.<sup>18</sup> However, for the nonexcess bandwidth case [i.e., where there is only one term in the sum (25)], we have been able to arrive at a complete solution.

We begin our analysis by substituting (26) and (28) into (23) to obtain

$$\xi_o = \sigma_d^2 \int_{-1/2}^{1/2} [I + P^+(z)C^+(z)C(z)P(z)/\sigma^2 T]^{-1} dz. \quad (32)$$

Our problem now is to minimize  $\text{MSE}_o = \text{trace}[\xi_o]$  over all matrices,  $P$ , in the class of ideally bandlimited filters. More explicitly, we wish to determine  $P_o$ , which minimizes  $\text{MSE}_o$  subject to an average power constraint,

$$\left. \begin{aligned} \text{MSE}_{op} &= \text{Min}_P \int_{-1/2}^{1/2} \text{trace}(\xi_o[W_o, P])dz \\ &= \int_{-1/2}^{1/2} \text{trace}(\xi_o[W_o, P_o])dz \\ \text{with } \int_{-1/2}^{1/2} \text{trace}[P^\dagger(z)P(z)]dz &= \text{constant} \end{aligned} \right\} \quad (33)$$

The search is for  $P_o$ , the class of filters that optimally distribute the fixed available average power between the two channels as well as across the frequency band of each channel.

We observe that (33) in conjunction with (29) and (30) leads to the matrix minimization of the functional

$$\text{Min}_P [\text{trace}\{(I + \rho_o P^\dagger C^\dagger C P)^{-1} + \bar{\lambda} P^\dagger P\}], \quad (34)$$

where  $\bar{\lambda}$  is a Lagrange multiplier to be determined from the power constraint and  $\rho_o = \rho/2K^2$ .

For each value of  $z$ , the matrix  $C^\dagger C$  is Hermitian and can be diagonalized by a unitary transformation  $\psi$ ,

$$C^\dagger C = \psi^\dagger \Lambda \psi, \quad (35)$$

whose nonnegative eigenvalues are

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Defining

$$\begin{aligned} G &= \psi P \\ P^\dagger \psi^\dagger \Lambda \psi P &= G^\dagger \Lambda G \end{aligned} \quad (36)$$

and observing that for 2 x 2 nonsingular matrices

$$\text{trace}[\cdot]^{-1} = \frac{\text{trace}[\cdot]}{\det[\cdot]}; \quad \text{trace}(AB) = \text{trace}(BA),$$

we express (34) in an equivalent form,

$$\begin{aligned} \text{Min}_G [\text{trace}\{(I + \rho_o G^\dagger \Lambda G)^{-1} + \lambda \rho_o G G^\dagger\}] \\ = \text{Min}_Q \left[ \frac{2 + \text{trace}(Q\Lambda)}{1 + \text{trace}(Q\Lambda) + \det(Q\Lambda)} + \lambda \text{trace}(Q) \right], \end{aligned} \quad (37)$$

where  $Q = \rho_o G G^\dagger$  and  $\lambda = \bar{\lambda}/\rho_o$ .

Careful inspection of (37) reveals that since  $Q$  is Hermitian and positive definite, and a nonzero off-diagonal entry in  $Q$  affects only the determinant and not the traces, *the minimizing matrix  $Q$  must be diagonal.*

Therefore, denoting

$$\mathbf{Q} = \begin{pmatrix} q_{11} & 0 \\ 0 & q_{22} \end{pmatrix}$$

with  $q_{11} \geq 0$  and  $q_{22} \geq 0$ , the original minimization problem now reduces to a much simpler one, namely,

$$\text{Min}_{\substack{q_{11}, q_{22} \\ q_{11}, q_{22} \geq 0}} \left\{ \frac{1}{\lambda_1 q_{11} + 1} + \frac{1}{\lambda_2 q_{22} + 1} + \lambda(q_{11} + q_{22}) \right\}, \quad (38)$$

which, with the power constraint  $2\rho = \text{trace}[\int_{-1/2}^{1/2} \mathbf{Q} dz]$ , immediately yields

$$\left. \begin{aligned} q_{11} &= \text{Max} \left[ 0, \quad \sqrt{\frac{\lambda_1}{\lambda}} - 1 \right] / \lambda_1 \\ q_{22} &= \text{Max} \left[ 0, \quad \sqrt{\frac{\lambda_2}{\lambda}} - 1 \right] / \lambda_2 \end{aligned} \right\}. \quad (39)$$

The procedure for obtaining  $\lambda$ ,  $q_{11}$ , and  $q_{22}$  can be found in Appendix B.

We now write the closed-form expression for  $\text{MSE}_{op}$  of (33) as

$$\text{MSE}_{op} = \sigma_d^2 \int_{-1/2}^{1/2} \left[ \frac{1}{1 + q_{11}\lambda_1} + \frac{1}{1 + q_{22}\lambda_2} \right] dz. \quad (40)$$

Recall that  $\mathbf{Q} = \rho_o \mathbf{G} \mathbf{G}^\dagger$  and  $\mathbf{G} = \psi \mathbf{P}$ . For the optimum  $\mathbf{Q}$ , we can therefore write

$$\left. \begin{aligned} \mathbf{Q} &= \rho_o \psi \mathbf{P}_o \mathbf{P}_o^\dagger \psi^\dagger \quad \text{and} \quad \mathbf{P}_o \mathbf{P}_o^\dagger = \psi^\dagger (\mathbf{Q}^{1/2} / \sqrt{\rho_o}) (\mathbf{Q}^{1/2} / \sqrt{\rho_o}) \psi \\ \mathbf{Q}^{1/2} &= \begin{pmatrix} \sqrt{q_{11}} & 0 \\ 0 & \sqrt{q_{22}} \end{pmatrix} \end{aligned} \right\}. \quad (41)$$

Equation (41) yields the functional form of the optimizing transmitter matrix filter,  $\mathbf{P}_o$ , given by

$$\mathbf{P}_o = \frac{1}{\sqrt{\rho_o}} \psi^\dagger \mathbf{Q}^{1/2} \mathbf{S}; \quad \mathbf{S} \text{ being any unitary matrix.} \quad (42)$$

Since  $\mathbf{S}$  is arbitrary, there is an infinite number of matrices,  $\mathbf{P}_o$ , that would give the same  $\text{MSE}_{op}$ . The particular choice of  $\mathbf{S}$  would determine  $\text{MSE}_1$  and  $\text{MSE}_2$ , the diagonal entries in  $\xi_o(\mathbf{W}_o, \mathbf{P}_o)$ , albeit their sum,  $\text{MSE}_{op}$ , is the same.

When  $\mathbf{S} = \mathbf{I}$ ,  $\mathbf{P}^\dagger \mathbf{C}^\dagger \mathbf{C} \mathbf{P} = \mathbf{Q} \mathbf{\Lambda}$  is diagonal and

$$\left. \begin{aligned} \xi_o(W_o, P_o) &= \begin{pmatrix} \text{MSE}_1 & 0 \\ 0 & \text{MSE}_2 \end{pmatrix}, \\ \text{MSE}_1 &= \sigma_d^2 \int_{-1/2}^{1/2} \frac{dz}{1 + q_{11}\lambda_1}; \quad \text{MSE}_2 = \sigma_d^2 \int_{-1/2}^{1/2} \frac{dz}{1 + q_{22}\lambda_2} \end{aligned} \right\} \quad (43)$$

From (18a) we obtain  $U_o$  as

$$U_o = I - \frac{1}{\sigma_d^2} \xi_o(W_o, P_o) = \begin{pmatrix} 1 - \text{MSE}_1/\sigma_d^2 & 0 \\ 0 & 1 - \text{MSE}_2/\sigma_d^2 \end{pmatrix}. \quad (44)$$

As we shall later see,  $U_o$  is an important quantity in the determination of the system error rate.

When  $S = \psi$ ,  $P_o$  is Hermitian. In general, however,

$$\left. \begin{aligned} \text{MSE}_1 &= \sigma_d^2 \int_{-1/2}^{1/2} \frac{1 + |S_{11}|^2 q_{11}\lambda_1 + |S_{12}|^2 q_{22}\lambda_2}{(1 + q_{11}\lambda_1)(1 + q_{22}\lambda_2)} dz \\ \text{MSE}_2 &= \sigma_d^2 \int_{-1/2}^{1/2} \frac{1 + |S_{21}|^2 q_{11}\lambda_1 + |S_{22}|^2 q_{22}\lambda_2}{(1 + q_{11}\lambda_1)(1 + q_{22}\lambda_2)} dz \end{aligned} \right\} \quad (45)$$

are real quantities, and so are the diagonal entries in the matrix  $U_o$ . The matrices  $\xi_o(W_o, P_o)$  and  $U_o$  will not be diagonal in general.

An interesting case arises when

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}; \quad |S_{jl}|^2 = \frac{1}{2}; \quad j, l = 1, 2, \quad (46)$$

since one can associate this matrix with the transmission of circular polarization. Consequently, from (45) and (46),  $\text{MSE}_1 = \text{MSE}_2$ . Thus, if we desire to have identical MSEs when the two channels are different, a power conditioning filter,  $\psi^\dagger Q^{1/2}/\sqrt{\rho_o}$  in (42), followed by a circular polarizer in the transmitter will achieve it.

It is instructive at this point to investigate certain limiting cases. For the no cross-polarization case,

$$\psi = I \quad \text{and} \quad C^\dagger C = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \quad (47)$$

From (31), (38), (39), and (40) we can write

$$\left. \begin{aligned} \text{MSE}_o &= \sigma_d^2 \int_{-1/2}^{1/2} \left[ \frac{1}{1 + \rho\lambda_1} + \frac{1}{1 + \rho\lambda_2} \right] dz \\ \text{MSE}_{op} &= \sigma_d^2 \int_{-1/2}^{1/2} \left[ \frac{1}{1 + q_{11}\lambda_1} + \frac{1}{1 + q_{22}\lambda_2} \right] dz \end{aligned} \right\} \quad (48)$$

It is therefore possible to view  $q_{11}$  and  $q_{22}$  as equivalent CNRs in the latter case. As long as the Lagrange multiplier  $\lambda$  obeys

$$\lambda < \text{Min}_{|z| \leq \frac{1}{2}} [\lambda_1(z), \lambda_2(z)],$$

we obtain (see Appendix B) from the power constraint

$$\sqrt{\lambda} = \frac{\int_{-1/2}^{1/2} [1/\sqrt{\lambda_1} + 1/\sqrt{\lambda_2}] dz}{2\rho + \int_{-1/2}^{1/2} [1/\lambda_1 + 1/\lambda_2] dz} \quad (49)$$

and

$$\text{MSE}_{op} = \sigma_d^2 \frac{\left\{ \int_{-1/2}^{1/2} [1/\sqrt{\lambda_1} + 1/\sqrt{\lambda_2}] dz \right\}^2}{2\rho + \int_{-1/2}^{1/2} (1/\lambda_1 + 1/\lambda_2) dz} \quad (50)$$

For nondispersive fades, (48) and (50) reduce to

$$\left. \begin{aligned} \text{MSE}_o &= \sigma_d^2 \left[ \frac{1}{1 + \rho\lambda_1} + \frac{1}{1 + \rho\lambda_2} \right] \\ \text{MSE}_{op} &= \sigma_d^2 \frac{[1/\sqrt{\lambda_1} + 1/\sqrt{\lambda_2}]^2}{2\rho + 1/\lambda_1 + 1/\lambda_2} \end{aligned} \right\} \quad (51)$$

As can be verified, when  $\rho \rightarrow \infty$ , both MSEs  $\rightarrow 0$ . For identical fades, i.e.,  $\lambda_1 = \lambda_2$ , we get the expected result

$$\text{MSE}_o = \text{MSE}_{op} = \frac{2\sigma_d^2}{1 + \rho\lambda_1}.$$

For independent fades with  $\sqrt{\lambda_1} \ll \sqrt{\lambda_2}$  and sufficiently large CNR, i.e.,  $\rho \rightarrow \infty$ , the improvement in mean-square errors is

$$\lim_{\rho \rightarrow \infty} \frac{\text{MSE}_o}{\text{MSE}_{op}} = \frac{2/\lambda_1}{1/\lambda_1} = 2.$$

This is due to power transfer from the good channel to the bad one.

For two identical fading channels, the transmitter optimization is tantamount to power redistribution within the individual channels without power transfer between the channels. The asymptotic improvement in MSE, utilizing an optimized transmitter filter, can be calculated as

$$\lim_{\rho \rightarrow \infty} \frac{\text{MSE}_o}{\text{MSE}_{op}} = \lim_{\rho \rightarrow \infty} \frac{2 \left[ 2\rho + 2 \int_{-1/2}^{1/2} \frac{dz}{\lambda_1} \right] \int_{-1/2}^{1/2} \frac{dz}{1 + \rho\lambda_1}}{4 \left[ \int_{-1/2}^{1/2} \frac{dz}{\sqrt{\lambda_1}} \right]^2} = \frac{\int_{-1/2}^{1/2} \frac{dz}{\lambda_1}}{\left[ \int_{-1/2}^{1/2} \frac{dz}{\sqrt{\lambda_1}} \right]^2} \geq 1,$$

by Schwarz's inequality.

## VI. THE MATCHED FILTER BOUND

A useful bound on the performance of equalizers can be obtained by assuming that one can detect each data vector symbol optimally by a matched matrix filter without incurring the penalty of intersymbol interference. Clearly, it is impossible for any equalizer to perform better than this fictional system.

We begin by noting from eq. (17) that in the absence of intersymbol interference the optimum equalizer filter  $W_o(-t)$  satisfies the equation

$$\begin{aligned} W_o(-t) &= \frac{1}{\sigma^2} H^\dagger(t) - \frac{1}{\sigma^2} U_o H^\dagger(t) \\ &= \frac{1}{\sigma^2} (I - U_o) H^\dagger(t). \end{aligned} \quad (52)$$

This is recognized as a matched filter followed by a multiplicative matrix coefficient.

The optimum mean-square error matrix is, again, from (18)

$$\xi_o = \sigma_d^2 (I - U_o), \quad (53)$$

where in this case,  $U_o$  satisfies (19) for  $k = 0$ ,

$$\sigma^2 U_o = (I - U_o) R_o$$

and

$$\left. \begin{aligned} \xi_o &= \sigma_d^2 \left[ I + \frac{R_o}{\sigma^2} \right]^{-1} \\ \text{MSE}_o &= \text{trace}[\xi_o] \end{aligned} \right\} \quad (54)$$

For bandlimited transmission

$$\begin{aligned} R_o &= \int_{-\infty}^{\infty} H^\dagger(\tau) H(\tau) d\tau = \frac{1}{2\pi} \int_{-\pi/T}^{\pi/T} H^\dagger(\omega) H(\omega) d\omega \\ &= \frac{1}{T} \int_{-1/2}^{1/2} P^\dagger(z) C^\dagger(z) C(z) P(z) dz. \end{aligned} \quad (55)$$

Consistent with our previous notation in (27) through (30) and the definition of  $R_o$  in (55), the average-square error matrix  $\xi_o$  is given by

$$\xi_o = \sigma_d^2 \left[ I + \rho \int_{-1/2}^{1/2} C^\dagger C dz \right]^{-1}. \quad (56)$$

Writing

$$\int_{-1/2}^{1/2} C^\dagger C dz = \begin{pmatrix} M_{11} & M_{21} \\ M_{12} & M_{22} \end{pmatrix},$$

where  $M_{12} = M_{21}^*$ , we obtain

$$\xi_o = \sigma_d^2 \left[ I + \rho \begin{pmatrix} M_{11} & M_{21} \\ M_{12} & M_{22} \end{pmatrix} \right]^{-1}. \quad (57)$$

The individual and the total mean-square errors are thus given by

$$\left. \begin{aligned} \text{MSE}_1 &= \sigma_d^2 \frac{1 + \rho M_{22}}{(1 + \rho M_{11})(1 + \rho M_{22}) - |\rho M_{21}|^2} \\ \text{MSE}_2 &= \sigma_d^2 \frac{1 + \rho M_{11}}{(1 + \rho M_{11})(1 + \rho M_{22}) - |\rho M_{21}|^2} \\ \text{MSE}_o &= \text{MSE}_1 + \text{MSE}_2 \end{aligned} \right\}. \quad (58)$$

As before, it is again possible to optimize the transmitter matrix filter  $P$ . The problem at hand, while it appears similar to the one already solved for the general case of cross-polarization and intersymbol interferences, is different and cannot be deduced from the general solution. Here the problem reduces to determining  $P$  such that

$$\left. \begin{aligned} &\sigma_d^2 \text{trace} \left\{ \left[ I + \frac{1}{\sigma_d^2 T} \int_{-1/2}^{1/2} P^\dagger(z) C^\dagger(z) C(z) P(z) dz \right]^{-1} \right\} \\ &\text{is a minimum, subject to} \\ &\text{trace} \left\{ \int_{-1/2}^{1/2} P^\dagger(z) P(z) dz \right\} = \text{constant} \end{aligned} \right\}. \quad (59)$$

Note that the matrix  $A = \int_{-1/2}^{1/2} P^\dagger C^\dagger C P dz$  is semidefinite Hermitian, and as before, it can be represented as [eq. (36)]

$$A = \int_{-1/2}^{1/2} G^\dagger \Lambda G dz,$$

where  $G = \psi P$ .

As we have already argued in Section V, minimizing (56) is equivalent to maximizing

$$\text{Max}_P \det \left[ \int_{-1/2}^{1/2} G^\dagger \Lambda G dz \right] \quad (60)$$

subject to trace  $\int G^\dagger G = \text{constant}$ . Since the integrand of  $A$  is Hermitian and positive, the maximization is equivalent to requiring that the off-diagonal terms of  $\int G^\dagger \Lambda G$  be zero. This immediately leads to the conclusion that both  $G$  and  $Q$  are diagonal,  $Q = G^\dagger G$ .

Thus, the equivalent minimization problem now reduces to

$$\begin{aligned} \text{Min trace}_{q_{11}, q_{22}} \left[ I + \frac{1}{\sigma^2 T} \int_{-1/2}^{1/2} Q \Lambda dz \right]^{-1} \\ = \text{Min trace}_{q_{11}, q_{22}} \begin{pmatrix} \frac{1}{1 + \frac{1}{\sigma^2 T} \int_{-1/2}^{1/2} q_{11} \lambda_1 dz} & 0 \\ 0 & \frac{1}{1 + \frac{1}{\sigma^2 T} \int_{-1/2}^{1/2} q_{22} \lambda_2 dz} \end{pmatrix} \end{aligned} \quad (61)$$

subject to  $\int_{-1/2}^{1/2} (q_{11} + q_{22}) dz = (P_a + P_b) / (\sigma_d^2 / T^2) = P_{av} / (\sigma_d^2 / T^2)$ .

The functions  $q_{11}(z)$  and  $q_{22}(z)$ , which minimize (61), would maximize the following integrals:  $\int_{-1/2}^{1/2} q_{11}(z) \lambda_1(z) dz$  and  $\int_{-1/2}^{1/2} q_{22}(z) \lambda_2(z) dz$ , subject to the same power constraint. This is a straightforward optimization problem resulting in

$$q_{11}(z) = x \delta(z_1) \quad \text{and} \quad q_{22}(z) = y \delta(z_2); \quad x, y \geq 0, \quad (62)$$

where  $z_1$  and  $z_2$  are the normalized frequencies where the eigenvalues  $\lambda_1$  and  $\lambda_2$  attain their respective maximal values,  $\lambda_{1m}$  and  $\lambda_{2m}$ . Physically, these results imply that power in each band should be concentrated at the frequency of peak transmission. Again,  $\delta(\cdot)$  is the Dirac delta function and the positive constants  $x$  and  $y$  have yet to be determined.

Substituting (62) into (61) yields

$$\text{Min MSE}_{op} = \text{Min}_{x, y \geq 0} \sigma_d^2 \left[ \frac{1}{1 + \frac{x}{\sigma^2 T} \lambda_{1m}} + \frac{1}{1 + \frac{y}{\sigma^2 T} \lambda_{2m}} \right] \quad (63)$$

subject to  $x + y = \text{constant}$ .

Carrying out the above minimization, we find that

$$\rho_1 = \frac{x}{\sigma^2 T} = \frac{1}{\sqrt{\lambda_{1m}} + \sqrt{\lambda_{2m}}} [2\rho \sqrt{\lambda_{2m}} + (\sqrt{\lambda_{1m}} - \sqrt{\lambda_{2m}}) / \sqrt{\lambda_{1m} \lambda_{2m}}]$$

and

$$\rho_2 = \frac{y}{\sigma^2 T} = \frac{1}{\sqrt{\lambda_{1m}} + \sqrt{\lambda_{2m}}} [2\rho \sqrt{\lambda_{1m}} - (\sqrt{\lambda_{1m}} - \sqrt{\lambda_{2m}}) / \sqrt{\lambda_{1m} \lambda_{2m}}]. \quad (64)$$

The quantities  $\rho_1$  and  $\rho_2$  can be interpreted as the equivalent CNRs in the respective channels. Their sum is constant and equal to  $2\rho$ . The conditions for  $x$  and  $y$  to be nonnegative can readily be verified to be

$$2\rho\sqrt{\lambda_{2m}} \geq \frac{1}{\sqrt{\lambda_{1m}}} - \frac{1}{\sqrt{\lambda_{2m}}}, \quad \lambda_{2m} \geq \lambda_{1m}$$

and

$$2\rho\sqrt{\lambda_{1m}} \geq \frac{1}{\sqrt{\lambda_{2m}}} - \frac{1}{\sqrt{\lambda_{1m}}}, \quad \lambda_{1m} \geq \lambda_{2m}. \quad (65)$$

Using (61) and the optimized values of  $x, y$  given in (64), we finally get for the minimum mean-squared errors

$$\left. \begin{aligned} \text{MSE}_1 &= \frac{\sigma_d^2}{1 + \rho_1 \lambda_{1m}} = \sigma_d^2 \frac{\sqrt{\lambda_{2m}}(\sqrt{\lambda_{1m}} + \sqrt{\lambda_{2m}})}{2\rho\lambda_{1m}\lambda_{2m} + \lambda_{1m} + \lambda_{2m}} \\ \text{MSE}_2 &= \frac{\sigma_d^2}{1 + \rho_2 \lambda_{2m}} = \sigma_d^2 \frac{\sqrt{\lambda_{1m}}(\sqrt{\lambda_{1m}} + \sqrt{\lambda_{2m}})}{2\rho\lambda_{1m}\lambda_{2m} + \lambda_{1m} + \lambda_{2m}} \\ \text{MSE}_{op} &= \text{MSE}_1 + \text{MSE}_2 = \sigma_d^2 \frac{(\sqrt{\lambda_{1m}} + \sqrt{\lambda_{2m}})^2}{2\rho\lambda_{1m}\lambda_{2m} + \lambda_{1m} + \lambda_{2m}} \end{aligned} \right\}. \quad (66)$$

As can be seen, for nondispersive fades where there is no intersymbol interference (66) is identical to (51).

## VII. INFORMATION THEORY LIMIT

Here we discuss the information theoretic capacity of dually polarized radio channels. This quantity is the maximum number of essentially error-free bits per cycle that can be attained for a given matrix channel transfer characteristic and a given clear air CNR. The problem is to calculate the Shannon capacity of a multi-input, multi-output linear dispersive channel with additive Gaussian noise. This general problem has been considered by Brandenburg and Wyner.<sup>19</sup> Recently, Foschini and Vannucci<sup>20</sup> called our attention to some technical results that make the evaluation of this quantity in our case relatively simple.

We first recall the capacity formula for a scalar dispersive channel with frequency transfer characteristic  $C(f)$

$$C = \int_w \log_2 \left\{ 1 + \frac{\sigma_s^2}{N_o} |P(f)|^2 |C(f)|^2 \right\} df, \text{ b/s}, \quad (67)$$

where  $\sigma_s^2$  is the average signal power at the input to the transmitter filter  $P(f)$ , and  $N_o$  is the double-sided spectral density of the additive Gaussian noise. The transmitter filter  $P(f)$  is selected to maximize  $C$  subject to an average transmitted power constraint.

In dually polarized communications, the channel frequency characteristic is a  $2 \times 2$  complex matrix. To facilitate the calculation of capacity, we use the polar factorization theorem<sup>21</sup> to represent  $C(f)$  as

$$C(f) = U(f)\sqrt{C^\dagger(f)C(f)}, \quad (68)$$

where  $U(f)$  is unitary. (This is the key idea suggested by Foschini and Vannucci.<sup>20</sup>) Moreover, (35) is used to obtain the representation

$$\sqrt{C^\dagger(f)C(f)} = \psi^\dagger \sqrt{\Lambda} \psi, \quad (69)$$

where

$$\sqrt{\Lambda} = \begin{pmatrix} \sqrt{\lambda_1} & 0 \\ 0 & \sqrt{\lambda_2} \end{pmatrix}.$$

Since the product of two unitary transformations is unitary and since unitary transformations do not alter either the total average transmitted power or the noise power, the original channel can be represented mathematically as two parallel channels, each with frequency characteristic  $\lambda_1^{1/2}(f)$  and  $\lambda_2^{1/2}(f)$ , respectively. This is evident from the representation expressed in Ref. 69.

Since the mathematically equivalent representation is comprised of two parallel scalar channels, the capacities add and consequently the total channel capacity is

$$\begin{aligned} C = & \frac{1}{\ln 2} \int_{w_1} \ln \left\{ 1 + \frac{\sigma_1^2}{N_o} |P_1(f)|^2 \lambda_1(f) \right\} df \\ & + \frac{1}{\ln 2} \int_{w_2} \ln \left\{ 1 + \frac{\sigma_2^2}{N_o} |P_2(f)|^2 \lambda_2(f) \right\} df, \text{ b/s,} \end{aligned} \quad (70)$$

where  $P_1$  and  $P_2$  are the filter characteristics that optimize  $C$  subject to a total average power constraint

$$P_{av} = 2\sigma_1^2 \int_{w_1} |P_1(f)|^2 df + 2\sigma_2^2 \int_{w_2} |P_2(f)|^2 df, \quad (71)$$

where  $\sigma_1^2$  and  $\sigma_2^2$  are the average input powers to the transmitter filters.

In our applications, the channels are assumed to be strictly band-limited to the band  $W = (1/T)$ , and therefore, it is convenient to normalize the capacity formula by dividing (70) by  $W$ . We denote the normalized capacity—i.e., the efficiency index, sometimes called the Shannon limit—by  $I_s$  in units of b/s/Hz.

The total CNR obtained via the constraint (71) is

$$2\rho = P_{av}/2N_o W,$$

and, therefore, (70) and (71) can be rewritten in the normalized notation as

$$\mathbf{I}_s = \frac{1}{\ln 2} \left\{ \int_{-1/2}^{1/2} \ln \left[ 1 + \frac{\sigma_1^2}{N_o} |P_1(z)|^2 \lambda_1(z) \right] dz + \int_{-1/2}^{1/2} \ln \left[ 1 + \frac{\sigma_2^2}{N_o} |P_2(z)|^2 \lambda_2(z) \right] dz \right\} \quad (72)$$

with  $2\rho = \frac{\sigma_1^2}{N_o} \int_{-1/2}^{1/2} |P_1(z)|^2 dz + \frac{\sigma_2^2}{N_o} \int_{-1/2}^{1/2} |P_2(z)|^2 dz$

The mathematical optimization can thus be cast as finding

$$\text{Max}_{|P_1|^2, |P_2|^2 \geq 0} \left[ \mathbf{I}_s[P_1, P_2] - \lambda \left[ \frac{\sigma_1^2}{N_o} |P_1|^2 + \frac{\sigma_2^2}{N_o} |P_2|^2 \right] \right], \quad (73)$$

where  $\lambda$  is a Lagrange multiplier. Using the calculus of variation on (73), we obtain the following two equations

$$\left. \begin{aligned} |P_1(z)|^2 &= \text{Max} \left[ 0, \frac{N_o}{\sigma_1^2} \left[ \frac{1}{\lambda} - \frac{1}{\lambda_1} \right] \right] \\ |P_2(z)|^2 &= \text{Max} \left[ 0, \frac{N_o}{\sigma_2^2} \left[ \frac{1}{\lambda} - \frac{1}{\lambda_2} \right] \right] \end{aligned} \right\} \quad (74)$$

The Lagrange multiplier is evaluated from the power constraint in (72),

$$2\rho = \int_{-1/2}^{1/2} \left[ \text{Max} \left\{ 0, \left[ \frac{1}{\lambda} - \frac{1}{\lambda_1} \right] \right\} + \text{Max} \left\{ 0, \left[ \frac{1}{\lambda} - \frac{1}{\lambda_2} \right] \right\} \right] dz. \quad (75)$$

The iterative procedure outlined in Appendix B can accommodate the evaluation of  $\lambda$  in this case, and the proper intervals  $Z_1$  and  $Z_2$ . These are the intervals over which  $|P_1(z)|^2 > 0$  and  $|P_2(z)|^2 > 0$ , respectively. We can now calculate the Shannon limit,  $\mathbf{I}_s$ , from (74) and (72) as

$$\mathbf{I}_s = \frac{1}{\ln 2} \left\{ \int_{Z_1} \ln \left[ \frac{\lambda_1}{\lambda} \right] dz + \int_{Z_2} \ln \left[ \frac{\lambda_2}{\lambda} \right] dz \right\}. \quad (76)$$

### VIII. PROPAGATION MODEL

To evaluate the efficacy of the optimized dually polarized data communications system, one would require a detailed knowledge of the functional and statistical behavior of the complex matrix channel characteristics. Since fading events are not very frequent, proper complex matrix channel characterization typically requires a few years of coherent measurements. Such experiments have not been performed to date and only partial data, mainly amplitude characteristics of the

in-line channels  $|C_{11}|$  and  $|C_{22}|$ , have been measured over a limited number of propagation paths.

The propagation model that we shall employ incorporates present knowledge with suggested extensions for the phase as well as the cross-polarization characteristics. Although this model is incomplete, we can still obtain meaningful qualitative results and use them for comparative evaluations of various equalization methods.

The behavior of  $|C_{11}|$  and  $|C_{22}|$ , the single channel (in-line) amplitude characteristics, are well established and documented.<sup>22-25</sup> To review briefly, the generic form of the single channel characteristic is

$$C(z) = a[1 - b \exp(-2\pi iz\tau/T)]; \quad |z| \leq \frac{1}{2}. \quad (77)$$

In the above representation  $\tau$  is a fixed parameter (6.3 ns),  $a$  is real, and  $b$  is complex valued, ( $|b| < 1$ ). The following definitions and statistics are usually associated with the parameters in (77),

$$\begin{aligned} A &= -20 \log a \\ B &= -20 \log(1 - |b|) \\ \phi &= \arg(b). \end{aligned} \quad (78)$$

The parameters  $B$  and  $\phi$  are taken to be independent random variables possessing probability densities

$$P(B) = \frac{1}{3.8} \exp(-B/3.8)$$

and

$$P(\phi) = \begin{cases} 5/6 \pi, & |\phi| \leq \pi/2 \\ 1/6 \pi, & \pi/2 \leq |\phi| \leq \pi \end{cases}. \quad (79)$$

The parameter  $A$  is a Gaussian random variable with average value dependent on  $B$

$$E[A] = 24.6(500 + B^4)/(800 + B^4) \text{ in dB} \quad (80)$$

and standard deviation

$$\sigma_A = 5 \text{ in dB}. \quad (81)$$

The single channels can be viewed as being in one of two states with associated probabilities,

$$\begin{aligned} \text{Pr}(\text{unfaded state}) &= 0.99689 \\ \text{Pr}(\text{faded state}) &= 0.00311. \end{aligned} \quad (82)$$

The unfaded state is characterized by  $C(z) = 1$ .

The structure of the off-diagonal terms of the matrix  $C(z)$ , repre-

senting the cross-polarization interference, is based upon some experimental evidence and heuristic arguments:<sup>26-30</sup>

$$\begin{aligned} |C_{12}(z)| &= |k_1 C_{11}(z) + k_2 C_{22}(z) + R_3| \\ |C_{21}(z)| &= |k_4 C_{11}(z) + k_5 C_{22}(z) + R_6|, \end{aligned} \quad (83)$$

where  $k_1$ ,  $k_2$ ,  $k_4$ , and  $k_5$  are constants that incorporate the nonideal properties of the antennas and waveguide feeds at both ends of the channel. We assume that dispersion due to multimode conversion, propagation, and reflections, in the waveguide feeds, has either been eliminated (in the 6-GHz and higher bands) or is nonexistent (in the 4-GHz band). Typical values of these constants, which vary from one hop to another, are in the  $-35$  to  $-20$  dB range. The terms  $R_3$  and  $R_6$  are hypothesized to be independent complex Gaussian random variables. One may view each of these variables as the result of an independent ray, which would contribute a nondispersive cross-polarization response except for a time delay. Typical values of the variance  $\sigma_R$ , associated with these Rayleigh random variables, is in the  $-45$  to  $-35$  dB range.

Current data<sup>27</sup> and previous experience<sup>5</sup> indicate that  $|C_{11}|$  and  $|C_{22}|$  are highly correlated, i.e., they practically fade in unison. It has been observed in some cases that  $|C_{11}|$  and  $|C_{22}|$  are displaced relative to one another across the frequency band.

Based on the above discussion, we use for our numerical evaluations the following extended model (to include amplitude as well as phase characteristics) for the dually polarized fading radio channel

$$\left. \begin{aligned} C_{11}(z) &= a[1 - b \exp(-2\pi iz\tau/T)], \quad |z| \leq 1/2 \\ C_{22}(z) &= a[1 - b \exp(-2\pi i(z - \Delta z)\tau/T)]; \\ &\quad \Delta z \text{ is uniformly distributed over } |\Delta z| \leq \Delta z_{\max} \\ C_{12}(z) &= k_1 C_{11}(z) + k_2 C_{22}(z) + R_3 \exp(-2\pi izD_1/T) \\ C_{21}(z) &= k_4 C_{11}(z) + k_5 C_{22}(z) + R_6 \exp(-2\pi izD_2/T) \end{aligned} \right\} \quad (84)$$

As can be seen, we have also associated different time delays,  $D_1$  and  $D_2$ , with  $R_3$  and  $R_6$ . Their values and influence will be discussed in a later section. We shall use (84) to evaluate the normalized data rates of dually polarized radio channels.

## IX. NUMERICAL SIMULATION STUDIES AND RESULTS

A meaningful performance measure of a particular communication technique is the probability distribution of normalized data rate that can be supported at a certain error rate objective. The normalized data rate is  $\mathbf{I} = (\log_2 L^2)/WT$  in b/s/Hz, where  $L^2$  is the total number of points in the QAM constellation and  $WT = 1$  in our case.  $\mathbf{I}$  can be obtained from the Chernoff bounds on the probability of error devel-

oped in Appendix C. For a fixed error rate we determine the maximum number of b/s/Hz supported by the particular channel at a given CNR. By computing the maximum number of b/s/Hz supported by each member of a large representative population of matrix channels, we obtain an estimate of the probability distribution of normalized data rates. This distribution is then used to estimate the probability of outage.

More specifically, if  $F(r)$  is the probability distribution of normalized data rates associated with a particular equalizer/canceler structure and an outage objective  $\alpha$  is set, then the value  $r_\alpha$  for which  $F(r_\alpha) = \alpha$  represents the maximum data rate at which it is possible to communicate data and meet the set outage objectives.

We evaluate, via Monte Carlo simulation, the normalized data rates that can be supported by our model of a dually polarized radio channel employing various equalization/cancellation structures at a prescribed error rate. Initial sanity checks for correlated and independent channels subjected to flat fading were used in the early stages of the program debugging. As we intended to simulate a large number (25,000) of fading events, we have utilized the CRAY computer.

Initial studies were devised to evaluate the sensitivity of the normalized mean-square errors,  $MSE_o/\sigma_d^2$  and  $MSE_{op}/\sigma_d^2$  from (31) and (40), to the propagation parameters  $\Delta z_{\max}$ ,  $D_1/T$ , and  $D_2/T$  suggested in (84). For a small subset of fading channels, we varied  $\Delta z_{\max}$  from 0 to 0.3 and  $D_1/T$  and  $D_2/T$  from 0 up to 3. The results indicated that the normalized mean-square errors were not very sensitive to these parameters when the other propagation parameters satisfy  $\{|k_j|\} \leq -20$  dB,  $j = 1, 2, 4, 5$  and  $\sigma_R < -35$  dB. (In the practical case of finite tap equalizers, there may be more sensitivity to  $D_1/T$  and  $D_2/T$ .) In the same calculations we have also observed insensitivity to the phases associated with  $\{k_j\}$  in (84).

Based upon these observations, we chose the following worst values for our subsequent evaluations

$$k_1 = k_2 = k_4 = k_5 = 0.1; \quad \sigma_R = -35 \text{ dB.} \quad (85)$$

Other parameters were chosen as follows:

$$D_1/T = D_2/T = 0.27 \quad \text{and} \quad \Delta z_{\max} = 0.15.$$

We have generated two sets of 25,000 simulated frequency characteristics, each being obtained via Monte Carlo simulation of the random variables  $A$ ,  $B$ ,  $\phi$ ,  $R_3$ , and  $R_6$  in (78) and (84). One set was generated for the case of the cross-polarized radio channels as described in (84). In this case the channels are correlated. The other set was generated for the case of a joint fade of two independent channels where the  $A$ ,  $B$ , and  $\phi$  parameters of  $C_{11}$  and  $C_{22}$  in (84) are chosen

independently with identical statistics. The interest in this extremal and “nonphysical” dual radio channel case is to demonstrate the inherent optimization advantages obtained when the two channels are independent.

For given probabilities of errors,  $P_{e1}$  and  $P_{e2}$ , which we had designated as outage thresholds (typically  $10^{-4}$ ), we calculated the various normalized channels data rates for each simulated frequency characteristic of the appropriate set. This was accomplished by evaluating the individual mean-square errors in (31), (43), (58), and (64) and inserting them into (116) and (117). As we pointed out in Appendix C, these probabilities are upper bounds since exact expressions are not possible to obtain. We thus calculated the individual and total normalized channels data rates for the optimized receiver and jointly optimized transmitter-receiver equalizer/canceler, as well as their Matched Filter (MF) bounds, using these estimates. In addition, we also calculated the corresponding Shannon (information theoretic) limit from (76). The results are presented in the form of curves of outage probability versus normalized data rate, in which we incorporated the probabilities of being in a fading state given in (82). In Figs. 3 and 4, we obtained these curves for various situations, namely, no CPI;  $\{k_j\} = 0$  and  $\sigma_R = -35$  dB; and  $\{k_j\} = 0.1$  and  $\sigma_R = -35$  dB. Also in Figs. 3 and 4  $\text{CNR} = 63$  dB;  $\tau/T = 0.189$  (30-MHz bandwidth);  $D_1 = D_2 = 0.27T$  and  $P_{e1} = P_{e2} = 10^{-4}$ .

Figures 3a and b exhibit normalized data-rate distributions for correlated channels and for both optimized receiver and jointly optimized transmitter-receiver, respectively. For the optimized receiver we note that, in the absence of CPI (solid curves), there is no significant difference between the data rates attained with the equalizer and the matched filter bound, about 0.7 b/s/Hz per channel. (When CPI is absent and the channels are correlated, the performance—outage probability versus normalized data rate—of each individual radio channel is identical to that of the single radio channel.<sup>24</sup> We can, therefore, assess the penalties associated with dually polarized radio transmission from performance degradation relative to the case when CPI is absent.) The corresponding combined channels rates are displaced by approximately 5 b/s/Hz relative to the Shannon limit. In addition, the combined channels rate is twice that of an individual channel. When CPI consists of the Rayleigh distributed variables only (dashed curves), we note an improvement in the matched filter bound and the information theory limit while an individual channel rate with the equalizer degrades by about 1.5 b/s/Hz. (This degradation is the penalty of dually polarized radio transmission using a linear receiver equalizer only. The matched filter bound as well as Shannon’s limit indicate that CPI can be used to our advantage rather than detriment.)

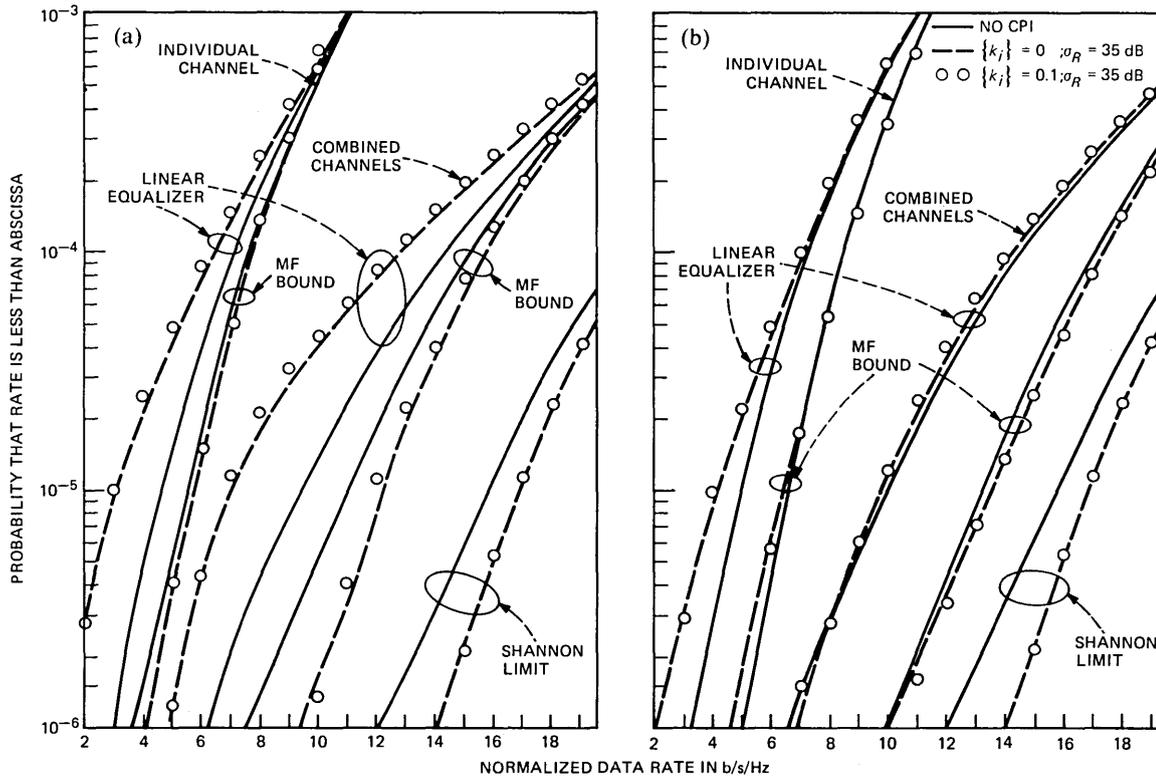


Fig. 3—(a) Receiver optimized equalizer/canceler. Outage probability versus normalized data rate for correlated channels. (b) Jointly optimized transmitter-receiver equalizer/canceler. Outage probability versus normalized data rate for correlated channels.

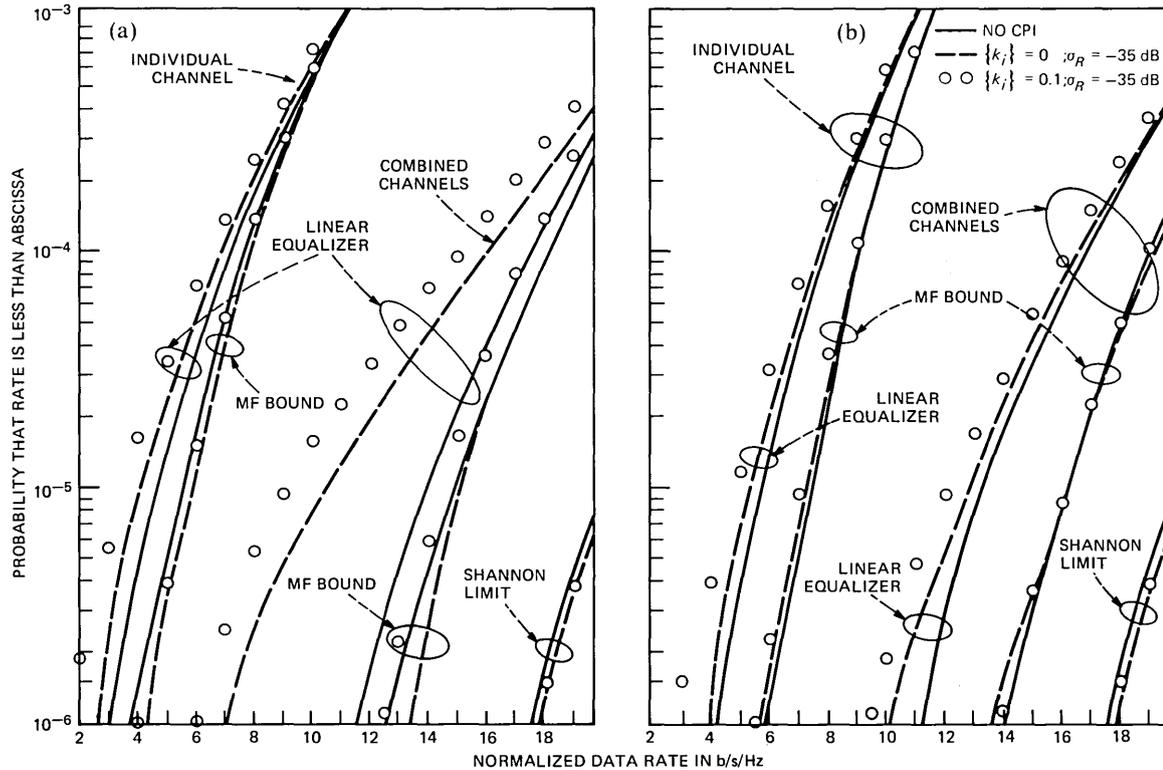


Fig. 4—(a) Receiver optimized equalizer/canceler. Outage probability versus normalized data rate for independent channels. (b) Jointly optimized transmitter-receiver equalizer/canceler. Outage probability versus normalized data rate for independent channels.

Note that the combined channels data rate is greater than twice the individual channel rate, as the combined rate is degraded by only 1.5 to 2.5 b/s/Hz. The physical interpretation of this phenomenon is that the independent CPI components  $R_3$  and  $R_6$  provide a certain degree of diversity that can be exploited to advantage as in the MF bound and the Shannon limit, albeit the equalizer in the optimized linear receiver is able to utilize it to a very limited extent.

When we add to the CPI the additional terms  $\{k_j\} = -20$  dB, all the curves are only slightly perturbed. This implies that the contribution of the antennas and their feeds can be practically neutralized by the equalizer.

Figure 3b demonstrates the performance of the jointly optimized transmitter-receiver for the same set of simulated channel frequency characteristics used in Fig. 3a. We note an improvement of 2 to 2.5 b/s/Hz over Fig. 3a in the matched filter bound in the absence of CPI (solid curves) while the corresponding improvement in the combined rate with the equalizer is under 1 b/s/Hz. The payoff for joint transmitter-receiver optimization results when CPI is present (dashed curves). As can be seen, about 1 b/s/Hz in individual channel rate and 2 to 3 b/s/Hz in combined rate are gained over Fig. 3a, in the critical region given by probability  $\leq 3 \times 10^{-5}$ . The differences between the rates with and without CPI are much smaller in Fig. 3b, indicating barely any penalty for dually polarized transmission with jointly optimized equalization. When the data rate is  $< 8$  b/s/Hz, the equalizer appears to utilize the CPI to yield a small advantage over the case without CPI. We conclude that the joint optimization utilizes the diversity, provided by the random components of the CPI, to a much greater extent. Note that there is still a 3-b/s/Hz difference relative to the matched filter bound in Fig. 3b. Again, the effects of CPI contributed by the antennas and their feeds have been eliminated for all practical purposes.

In Fig. 5 we compare the normalized data rates of the correlated channels, under severe CPI, for the optimized receiver and jointly optimized transmitter receiver. In this figure  $\text{CNR} = 63$  dB;  $\tau/T = 0.189$  (30-MHz bandwidth);  $\{k_j\} = -20$  dB;  $\sigma_R = 35$  dB;  $D_1 = D_2 = 0.27T$ ; and  $P_{e1} = P_{e2} = 10^{-4}$ . For an individual channel, an improvement of about 1 b/s/Hz can be achieved, in the region of outage probability less than  $6 \times 10^{-5}$ , with joint transmitter-receiver optimization. This advantage can amount to 50 percent or more of relative improvement when outage probability is  $\leq 3 \times 10^{-6}$ , and diminishes in significance as the outage probability increases. Note that the matched filter bounds of the individual channel rates are better by 2 to 2.5 b/s/Hz relative to their corresponding equalizer rates for outage probability  $< 5 \times 10^{-5}$ . Such a large relative difference indicates that perhaps

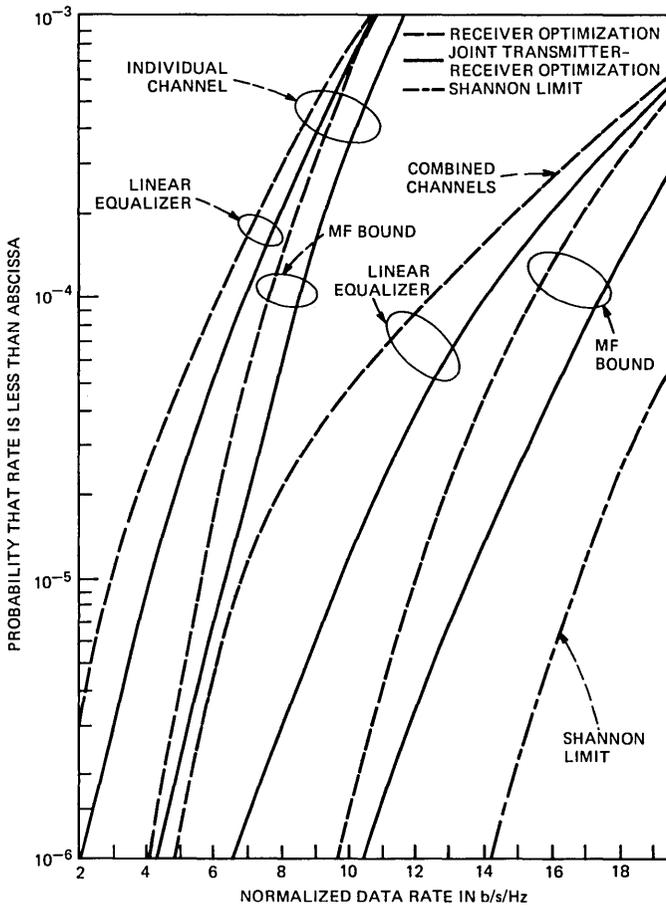


Fig. 5—Outage probability versus normalized data rate for correlated channels.

other equalization/cancellation methods and/or Viterbi decoding<sup>31</sup> could increase the individual rate of the channels.

The effects of partial diversity, due to the random component of the CPI, where the combined channels rate is more than three times the individual rate for the lower values of outage probability can again be observed. The total combined rate of the jointly optimized transmitter-receiver equalizer is lower by about 4 b/s/Hz and 7 b/s/Hz relative to the matched filter bound and the Shannon limit, respectively. The combined channels rate of the optimized receiver equalizer can be lower by as much as 5 b/s/Hz relative to its matched filter bound. These facts indicate that an additional improvement may be feasible by employing Viterbi decoding.<sup>31</sup> If higher values of probability of outage are acceptable, an optimized receiver equalizer can come rela-

tively close in performance to a jointly optimized transmitter-receiver equalizer.

Figures 4a and b exhibit the independent channels normalized data-rate distributions for an optimized receiver and jointly optimized transmitter-receiver, respectively. In this case there is diversity between the channels, and consequently, the combined channels data rate is much higher than twice that of an individual channel.

Again we note that for the optimized receiver there is no significant difference between the data rates supported by the equalizer and the matched filter bound in the absence of CPI interference. The total data rate is about three to four times that of an individual channel. There is, however, a 5-b/s/Hz difference between the Shannon limit relative to the correlated channels case of Fig. 3. When CPI is present, in the form of  $R_3$  and  $R_6$  only, (84), we notice a slight degradation in the equalized individual channel data rate with a correspondingly severe degradation in the total channels data rate. This seemingly strange phenomenon can be explained as follows: since the channels are independent, the probability of a joint bad fade is very low. Thus, in the absence of CPI, we expect one channel to be relatively good—capable of supporting a very high data rate—while the other channel is subject to a severe fade. When CPI is introduced, the data rate of the good channel degrades at a much faster pace, with the equalizer not being able to stem this pace as effectively as it can equalize the bad channel. Note that the contribution of the antennas and their feed is more noticeable here.

In Fig. 4b we show the performance of the jointly optimized transmitter-receiver for the same set of simulated channel responses as in Fig. 4a. We note a 1- to 2-b/s/Hz improvement in the individual channel data rate with the equalizer and the matched filter bound in the absence of CPI. In this case the total data rate attainable by the matched filter bound has improved by about 1 to 1.5 b/s/Hz relative to the optimized receiver in the absence of CPI indicated in Fig. 4a. On the other hand, the total data rate with the equalizer is slightly inferior to the optimized receiver case. This occurs because we have minimized the total mean-square error, which is dominated by the bad channel in each pair of simulated responses. We thus rob from the good channel some power and give it to the bad one. Consequently, the data rate of the good channel is decreased by more than the improvement obtained from the bad channel. This effect totally disappears when CPI is present and the superiority of the joint transmitter-receiver optimization can be observed to substantially counteract CPI.

In Fig. 6 we compare the normalized data rates of the independent channels with the optimized receiver and jointly optimized transmit-

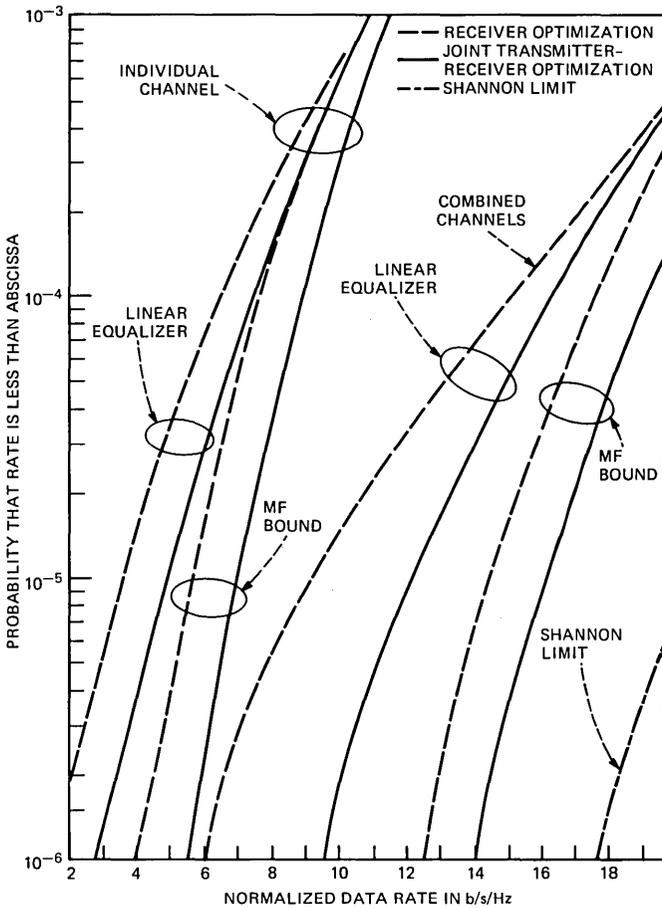


Fig. 6—Outage probability versus normalized data rate for independent channels.

ter-receiver equalizers, subject to the same severe CPI as in Fig. 5. In Fig. 6 the  $CNR = 63$  dB;  $\tau/T = 0.189$  (30-MHz bandwidth);  $\{k_j\} = -20$  dB;  $\sigma_R = -35$  dB;  $D_1 = D_2 = 0.27T$ ;  $P_{e1} = P_{e2} = 10^{-4}$ . An improvement of 1 to 1.5 b/s/Hz can be achieved in the region of outage probability  $< 6 \times 10^{-5}$  with joint optimization. This improvement is comparable to the correlated channels case, Fig. 5, although the rates themselves are slightly higher in Fig. 6. For the individual channel rates, we note again a 2- to 2.5-b/s/Hz difference between the two equalizers and their corresponding match filter bound.

The rate of the combined channels, being much higher than twice the rate of the individual channel, demonstrates the advantages that can be obtained by utilizing the diversity of the independent channels. Operation at a fixed data rate per channel would not utilize this

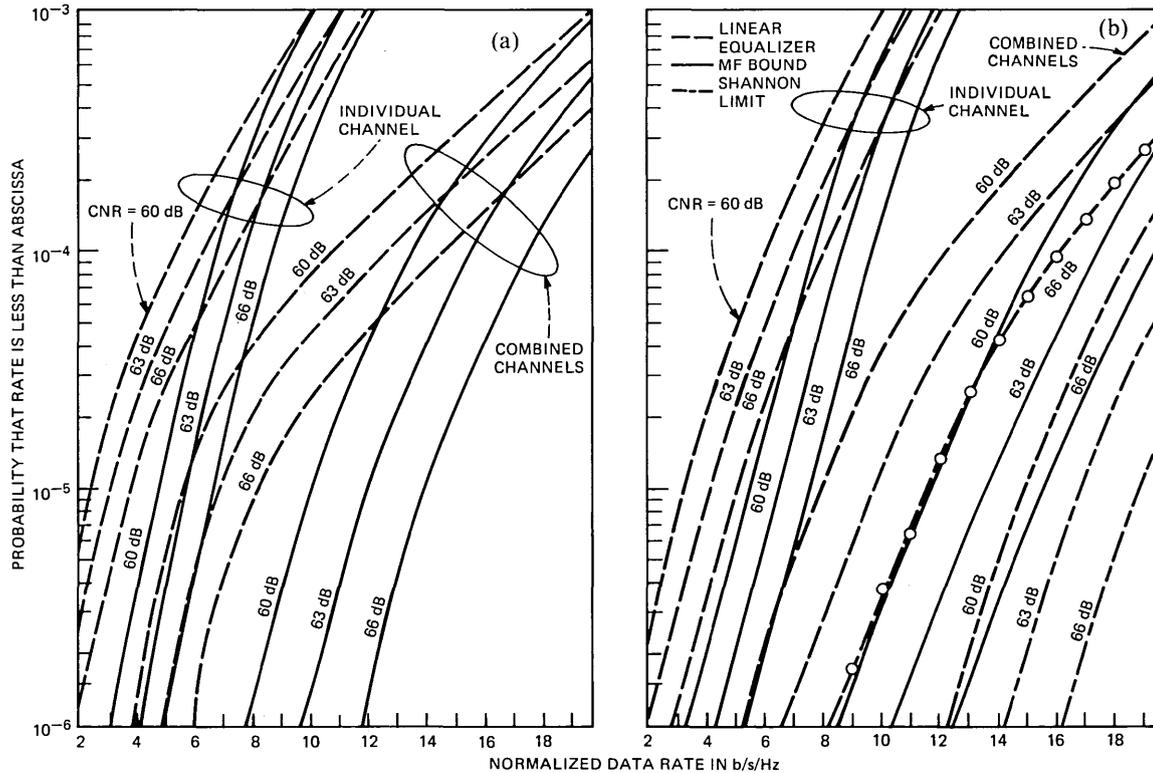


Fig. 7—(a) Receiver optimized equalizer/canceler. Sensitivity to clear air CNR. Outage probability versus normalized data rate for correlated channels. (b) Jointly optimized transmitter-receiver equalizer/canceler. Sensitivity to clear air CNR. Outage probability versus normalized data rate for correlated channels.

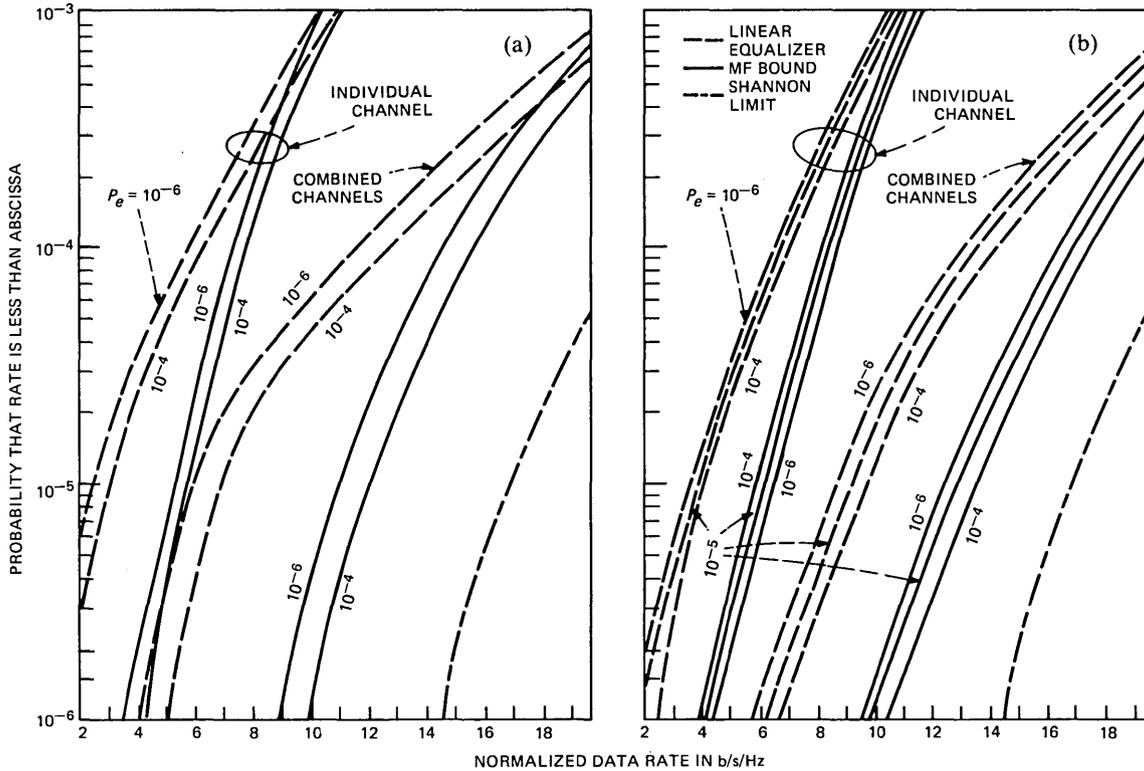


Fig. 8—(a) Receiver optimized equalizer/canceler. Sensitivity to threshold error rate— $P_{e1} = P_{e2} = P_e$ . Outage probability versus normalized data rate for correlated channels. (b) Jointly optimized transmitter-receiver equalizer/canceler. Sensitivity to threshold error rate— $P_{e1} = P_{e2} = P_e$ . Outage probability versus normalized data rate for correlated channels.

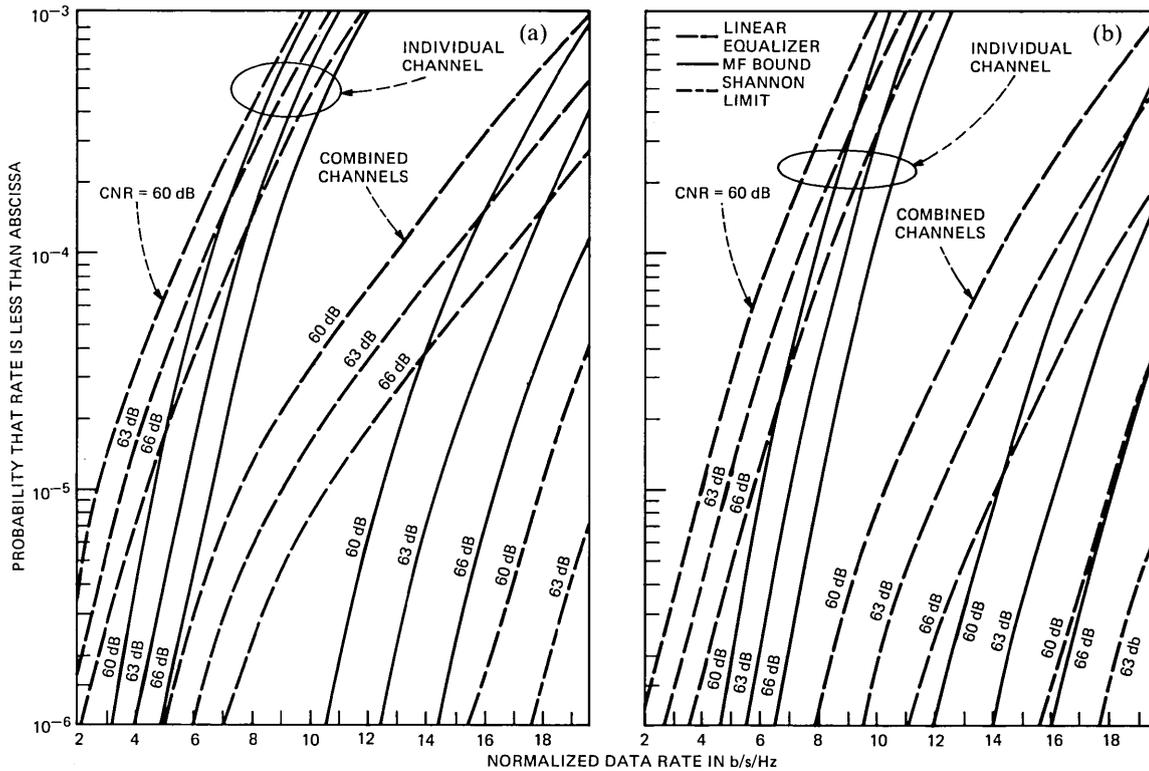


Fig. 9—(a) Receiver optimized equalizer/canceler. Sensitivity to clear air CNR. Outage probability versus normalized data rate for independent channels. (b) Jointly optimized transmitter-receiver equalizer/canceler. Sensitivity to clear air CNR. Outage probability versus normalized data rate for independent channels.

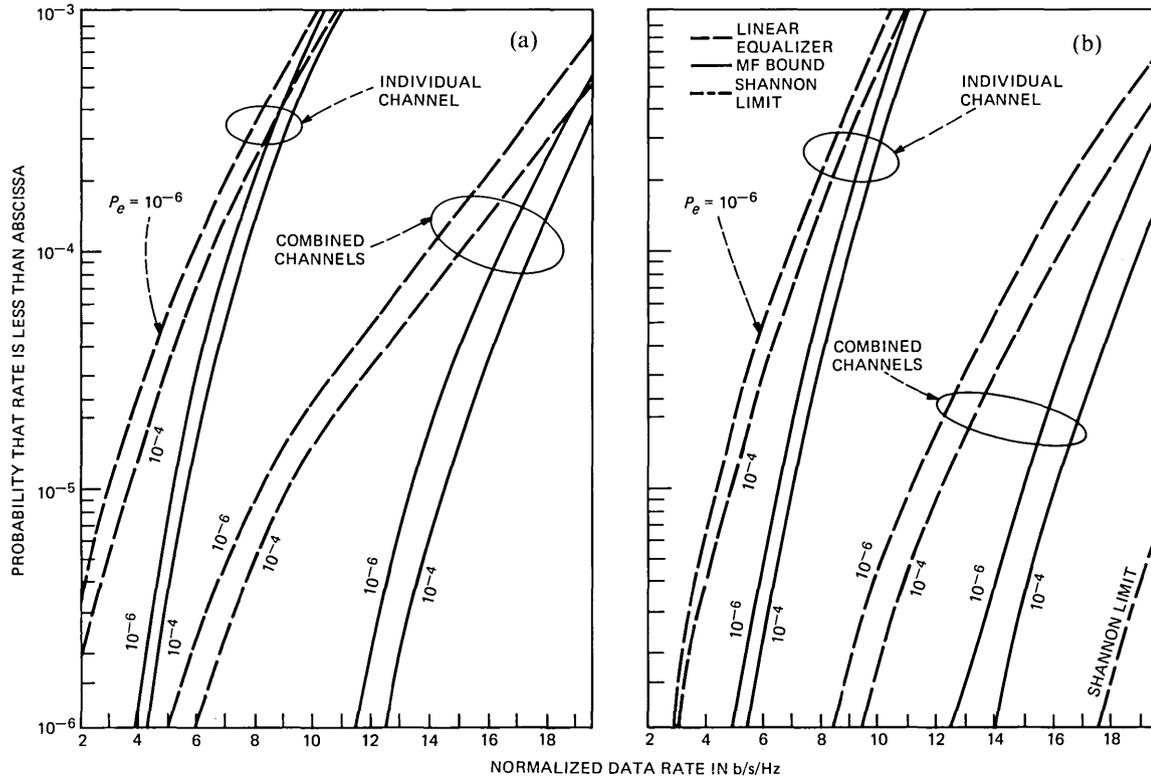


Fig. 10—(a) Receiver optimized equalizer/canceler. Sensitivity to threshold error rate— $P_{e1} = P_{e1} = P_e$ . Outage probability versus normalized data rate for independent channels. (b) Jointly optimized transmitter-receiver equalizer/canceler. Sensitivity to threshold error rate— $P_{e1} = P_{e2} = P_e$ . Outage probability versus normalized data rate for independent channels.

diversity, and design considerations would be based upon the individual channel performance. If we assume that Viterbi decoding can practically realize the matched filter bound, then the corresponding single-channel curves in Fig. 6 indicate an approximate 1- to 1.5-b/s/Hz advantage for joint optimization. In Fig. 5 the corresponding curves for the correlated channels indicate a marginal advantage for joint optimization. Operation at variable data rate, where the rate is adjusted according to the states of the channels, can result in a manyfold increase of total data rate. Judging again by the combined channels matched filter bounds in Figs. 5 and 6, we see no substantial advantage in employing joint transmitter-receiver optimization.

In Figs. 7 and 8 we exhibit the sensitivity of the normalized data rates to variations in CNR (clear air carrier-to-noise ratio) and  $P_e$  (threshold error rates) for the correlated radio channels case. In Figs. 7 and 9  $\tau/T = 0.189$ ;  $\{k_j\} = -20$  dB;  $\sigma_R = -35$  dB;  $D_1 = D_2 = 0.27T$ ; and  $P_{e1} = P_{e2} = 10^{-4}$ . In Figs. 8 and 10 CNR = 63 dB;  $\tau/T = 0.189$ ;  $\{k_j\} = -20$  dB;  $\sigma_R = -35$  dB; and  $D_1$  and  $D_2 = 0.27T$ . Figures 9 and 10 indicate the corresponding sensitivity for the independent channels case. For the individual channels, these figures indicate that, irrespective of mode of equalization and correlation between channels, a 3-dB increase (decrease) in CNR correspond to 1-b/s/Hz increase (decrease) in data rate. Also, a ten-fold decrease in  $P_e$  corresponds to an approximate 0.3-b/s/Hz decrease in data rate. These statements hold except for very low levels of outage probabilities,  $< 5 \times 10^{-6}$ , where less sensitivity is observed. These observations are similar to the case of single radio channels.<sup>24</sup>

## X. SUMMARY AND CONCLUSIONS

We have presented a theory of optimal (in a least-mean-square sense) equalization/cancellation in digital data transmission over two cross-coupled linear dispersive channels. This research was primarily motivated by problems associated with digital data communications over dually polarized fading radio channels, and most of our numerical work is directed toward this application.

Viewing the channel as a two-input port, two-output port linear network followed by additive noise, we determined the optimum linear receiver structure when the transmitted signals are two independent quadrature amplitude modulated data waves. The figure of merit used in the optimization was the total mean-square error between the desired data symbol and the post-filtered signal sample. We found that the structure of the optimum filter is comprised of a 2 by 2 matrix matched filter followed by a matrix tapped delay line. The tap coefficients are described by a sequence of constant matrices. The utility of this structure is that it can be approximated by a finite matrix

transversal filter whose matrix taps can be adjusted adaptively. For a given channel matrix, a transmitter matrix filter and the optimum receiver filter, we have derived a closed-form expression for the minimum attainable total mean-square error as a functional of the transmitter matrix. This also made it possible to optimize the transmitter subject to a total average power constraint

The extensive numerical work exhibited in the various graphs is based on our suggested propagation model. This model is rich enough to allow meaningful performance comparisons of the various equalization/cancellation methods and their ultimate limits.

Our major conclusions are:

1. The effects of Cross-Polarization Interference (CPI), contributed by the antennas and their feeds, can be practically eliminated by the linear equalizer/canceler structures discussed in this paper. This may have a significant bearing upon the cross-polarization requirements placed on the antennas.

2. In principle, the random components of CPI can provide partial diversity reception. They could be exploited to increase the total data rates supported by the channels relative to the case where CPI is absent.

3. A jointly optimized transmitter-receiver equalizer/canceler is effective when low outage probabilities are required. At relatively high outage probabilities it is of marginal benefit.

4. Over a range from  $10^{-6}$  to  $10^{-5}$  outage probability, the combined channels data rate achieved by the jointly optimized transmitter-receiver equalizer/canceler is 4 b/s/Hz less than its matched filter bound. Similarly, when only the receiver is optimized, the difference is 5 b/s/Hz. This observation applies to both correlated and independent channels. This significant difference in data rates warrants further investigation of other equalization/cancellation methods, such as decision feedback<sup>18</sup> and/or Viterbi decoding.<sup>31</sup>

5. In the same range of outage probability as in 4, the combined channels data rate associated with the matched filter bounds differ from the Shannon limit by about 3 b/s/Hz when jointly optimized transmitter-receiver are used and there is an additional loss of 1 b/s/Hz when only the receiver filter is optimized. This observation is again valid for both correlated and independent channels. The only way known to recoup some of this loss is by channel coding.

6. Irrespective of equalization methods and correlation between the channels, a 3-dB increase (decrease) in CNR corresponds to 1 b/s/Hz increase (decrease) in the achievable data rate of an individual channel. Also, a ten-fold decrease in error rate corresponds to an approximate 0.3-b/s/Hz decrease in individual channel data rate.

## XI. ACKNOWLEDGMENT

Discussions with G. J. Foschini, D. J. Goodman, L. J. Greenstein, D. O. Reudink, A. A. M. Saleh, A. D. Wyner, and Y. S. Yeh were valuable during the course of this work.

## REFERENCES

1. A. F. Culmone, "Polarization Diversity With Adaptive Channel Decoupling," Nat. Telecommun. Conf., New Orleans, December 1975, pp. 25-22-27.
2. B. D. Cullen et al., "Spectrum-Reuse by Adaptive Polarization Separation," Nat. Telecommun. Conf., New Orleans, December 1975, pp. 43-18-25.
3. C. A. Baird and G. Pelchat, "Cross Polarization Techniques Investigation," Harris Corporation Report No. RADC-TR-77-244, July 1977.
4. Howard E. Nichols, Arthur A. Giordano, and John G. Proakis, "MLD and MSE Algorithms for Adaptive Detection of Digital Signals in the Presence of Interchannel Interference," IEEE Trans. Inform. Theory, *IT-23*, No. 5 (September 1977).
5. N. Amitay, "Signal-to-Noise Ratio Statistics for Nondispersive Fading in Radio Channels With Cross Polarization Interference Cancellation," IEEE Trans. Commun., *COM-27*, No. 2 (February 1979), pp. 498-502.
6. B. E. Gillingham et al., "Cross Polarization Interference Reduction Techniques," Harris Corporation Report No. RADC-TR-79-154, June 1979.
7. J. Namiki and S. Takahara, "Adaptive Receiver for Cross-Polarized Digital Transmission," Int. Conf. Commun., Denver, June 14-18, 1981, Paper 46.3.1.
8. K. T. Wu and T. S. Giuffrida, "Feasibility Study for an Interference Canceller for Co-Channel Cross Polarization Operation of Digital Radio," 1982 ICC, June 1982, pp. 2B.7.1-5.
9. M. L. Steinberger, "Design of a Terrestrial Cross-Pol Canceller," Int. Conf. Commun., Philadelphia, June 1982, pp. 2B.6.1-5.
10. H. E. Lin, private communication.
11. P. D. Karabinis, private communication.
12. M. Kavehrad and C. A. Siller, private communication.
13. H. E. Lin, private communication.
14. M. Kavehrad, private communication.
15. M. Kavehrad, "Adaptive Cross-Polarization Interference Cancellation for Dual-Polarized M-QAM Signals," GLOBECOM Conf., San Diego, November 1983.
16. N. Amitay and L. J. Greenstein, private communication.
17. M. S. Mueller and J. Salz, "A Unified Theory of Data-Aided Equalization," B.S.T.J., *60*, No. 9 (November 1981), pp. 2023-38.
18. J. Salz, "Optimum Mean-Square Decision Feedback Equalization," B.S.T.J., *52*, No. 8 (October 1973), pp. 1341-73.
19. L. H. Brandenburg and A. D. Wyner, "Capacity of the Gaussian Channel With Memory: The Multivariate Case," B.S.T.J., *53*, No. 5 (May-June 1974), pp. 745-78.
20. G. J. Foschini and G. Vannucci, private communication.
21. H. Minc and M. Marcus, *A Survey of Matrix Theory and Matrix Inequalities*, Boston: The Prindle, Weber and Schmidt Complementary Series in Mathematics, 1964.
22. W. D. Rummler, "A New Selective Fading Model: Application to Propagation Data," B.S.T.J., *58*, No. 7 (May-June 1979), pp. 1037-71.
23. W. D. Rummler, "More on the Multipath Fading Channel Model," IEEE Trans. Commun., *COM-29*, No. 3 (March 1981), pp. 346-52.
24. G. J. Foschini and J. Salz, "Digital Communications Over Fading Radio Channels," B.S.T.J., *62*, No. 2 (February 1983), pp. 429-56.
25. N. Amitay and L. J. Greenstein, "Multipath Outage Performance of Digital Radio Receivers Using Finite-Tap Adaptive Equalizers," IEEE Trans. Commun., *COM-32*, No. 5 (May 1984), pp. 597-608.
26. S. H. Lin, "Impact of Microwave Depolarization During Multipath Fading on Digital Radio Performance," B.S.T.J., *56*, No. 5 (May 1977), pp. 645-74.
27. P. E. Butzien, private communication.
28. K. T. Wu, "Measured Statistics on Multipath Dispersion of Cross Polarization Interference," Int. Conf. Commun., Amsterdam, May 14-17, 1984, Paper 46.3.
29. M. L. Steinberger, private communication.

30. T. D. Mottl, "Dual-Polarized Channel Outages During Multipath Fading," B.S.T.J., 56, No. 5 (May 1977), pp. 675-701.
31. G. D. Forney, Jr., "The Viterbi Algorithm," Proc. IEEE, 61 (1973), pp. 268-78.
32. B. R. Saltzberg, "Intersymbol Interference Error Bounds With Application to Ideal Bandlimited Signaling," IEEE Trans. Inform. Theory, IT-14, No. 4 (July 1968), pp. 563-8.

## APPENDIX A

### Integral Equation for Optimum Matrix Filter

To evaluate (16), express trace  $\xi$  explicitly in terms of  $W_{ij}$  and  $h_{ij}$ ,  $i, j = 1, 2$ , the entries of matrices  $W$  and  $H$ , respectively:

$$\begin{aligned} \text{trace } \xi = 2 - 2\text{Re} \left\{ \int [W_{11}(-\tau)h_{11}(\tau) + W_{12}(-\tau)h_{21}(\tau) \right. \\ \left. + W_{21}(-\tau)h_{12}(\tau) + W_{22}(-\tau)h_{22}(\tau)]d\tau \right\} \\ + \sigma^2 \int [ |W_{11}(\tau)|^2 + |W_{12}(\tau)|^2 + |W_{21}(\tau)|^2 + |W_{22}(\tau)|^2 ]d\tau \\ + \sum_n \left| \int [W_{11}(-\tau)h_{11}(\tau - nT) + W_{21}(-\tau)h_{12}(\tau - nT)]d\tau \right|^2 \\ + \sum_n \left| \int [W_{11}(-\tau)h_{21}(\tau - nT) + W_{21}(-\tau)h_{22}(\tau - nT)]d\tau \right|^2 \\ + \sum_n \left| \int [W_{12}(-\tau)h_{11}(\tau - nT) + W_{22}(-\tau)h_{12}(\tau - nT)]d\tau \right|^2 \\ + \sum_n \left| \int [W_{12}(-\tau)h_{21}(\tau - nT) + W_{22}(-\tau)h_{22}(\tau - nT)]d\tau \right|^2. \quad (86) \end{aligned}$$

We proceed as follows. Set the variations

$$\partial_{W_{ijr}} \text{trace } \xi = 0, \quad i, j = 1, 2$$

and

$$\partial_{W_{ij\mathcal{I}}} \text{trace } \xi = 0, \quad i, j = 1, 2, \quad (87)$$

where  $W_{ijr} = \text{Re}(W_{ij})$  and  $W_{ij\mathcal{I}} = \text{Im}(W_{ij})$ . The variations expressed in (87) imply that  $\partial_w \text{trace } \xi = 0$ , which is tantamount to equating

$$\frac{\partial}{\partial \epsilon} \text{trace } [\xi(W + \epsilon \delta W)] = 0, \quad (88)$$

where  $W + \epsilon \delta W$  is given in (16a) and the partial derivatives are with respect to  $\epsilon_{ij}$ ,  $i = j = 1, 2$ . Detailed calculations of

$$\frac{\partial}{\partial \epsilon_{11}} \text{trace } \xi = 0$$

yield the integral equation,

$$\begin{aligned}
h_{11}(\tau) &= \sigma^2 W_{11}^*(-\tau) + \sum_n h_{11}(\tau - nT) \\
&\cdot \int [W_{11}(-z)h_{11}(z - nT) + W_{21}(-z)h_{12}(z - nT)]^* dz \\
&+ \sum_n h_{21}(\tau - nT) \\
&\cdot \int [W_{11}(-z)h_{21}(z - nT) + W_{21}(-z)h_{22}(z - nT)]^* dz; \quad (89) \\
&\frac{\partial}{\partial \epsilon_{12}} \text{trace } \xi = 0
\end{aligned}$$

yields

$$\begin{aligned}
h_{21}(\tau) &= \sigma^2 W_{12}^*(-\tau) + \sum_n h_{11}(\tau - nT) \\
&\cdot \int [W_{12}(-z)h_{11}(z - nT) + W_{22}(-z)h_{12}(z - nT)]^* dz \\
&+ \sum_n h_{21}(\tau - nT) \\
&\cdot \int [W_{12}(-z)h_{21}(z - nT) + W_{22}(-z)h_{22}(z - nT)]^* dz; \quad (90) \\
&\frac{\partial}{\partial \epsilon_{21}} \text{trace } \xi = 0
\end{aligned}$$

yields

$$\begin{aligned}
h_{12}(\tau) &= \sigma^2 W_{21}^*(-\tau) + \sum_n h_{12}(\tau - nT) \\
&\cdot \int [W_{11}(-z)h_{11}(z - nT) + W_{21}(-z)h_{12}(z - nT)]^* dz \\
&+ \sum_n h_{22}(\tau - nT) \int [W_{11}(-z)h_{21}(z - nT) \\
&+ W_{21}(-z)h_{22}(z - nT)]^* dz; \quad (91)
\end{aligned}$$

and

$$\frac{\partial}{\partial \epsilon_{22}} \text{trace } \xi$$

yields

$$\begin{aligned}
h_{22}(\tau) &= \sigma^2 W_{22}^*(-\tau) + \sum_n h_{12}(\tau - nT) \\
&\cdot \int [W_{12}(-z)h_{11}(z - nT) + W_{22}(-z)h_{12}(z - nT)]^* dz \\
&+ \sum_n h_{22}(\tau - nT) \\
&\cdot \int [W_{12}(-z)h_{21}(z - nT) + W_{22}(-z)h_{22}(z - nT)]^* dz, \quad (92)
\end{aligned}$$

where \* stands for complex conjugate. Equations (89) through (92) can be put into a matrix form:

$$\left[ \frac{\partial \text{trace } \xi}{\partial \epsilon_{ij}} \right]_{ij} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad i, j = 1, 2$$

yields

$$\begin{pmatrix} h_{11}^*(\tau) & h_{12}^*(\tau) \\ h_{21}^*(\tau) & h_{22}^*(\tau) \end{pmatrix} = \sigma^2 \begin{pmatrix} W_{11}(-\tau) & W_{21}(-\tau) \\ W_{12}(-\tau) & W_{22}(-\tau) \end{pmatrix} \\ + \sum_n \int_{-\infty}^{\infty} \begin{pmatrix} W_{11}(-z) & W_{21}(-z) \\ W_{12}(-z) & W_{22}(-z) \end{pmatrix} \begin{pmatrix} h_{11}(z - nT) & h_{21}(z - nT) \\ h_{12}(z - nT) & h_{22}(z - nT) \end{pmatrix} dz \\ \cdot \begin{pmatrix} h_{11}^*(\tau - nT) & h_{12}^*(\tau - nT) \\ h_{21}^*(\tau - nT) & h_{22}^*(\tau - nT) \end{pmatrix},$$

which is exactly eq. (17) in the text.

## APPENDIX B

### Determination of the Matrix Q

In (39) we give the elements of the diagonal matrix **Q** in terms of the eigenvalues  $\lambda_1$  and  $\lambda_2$ , and the Lagrange multiplier  $\lambda$ . We shall evaluate  $\lambda$  from the constraint of (38)

$$2\rho = \int_{-1/2}^{1/2} [q_{11}(z) + q_{22}(z)] dz. \quad (93)$$

From (39) it is clear that there may be a situation where  $\lambda_1$  or  $\lambda_2 \leq \lambda$  over a limited range of the interval  $|z| \leq 1/2$ . Here we shall develop an iterative procedure for determining  $\lambda$  and the appropriate subintervals of  $z$  over which  $q_{11}$  and  $q_{22} \geq 0$ .

In our first iteration, we equate  $\lambda = \lambda^{(0)}$  to the smallest eigenvalue in the interval, i.e.,

$$\lambda^{(0)} = \text{Min}_{|z| \leq \frac{1}{2}} [\lambda_1(z), \lambda_2(z)]. \quad (94)$$

This will ensure that  $q_{11}$  and  $q_{22}$  in (39) are nonnegative.

Substituting (39) in (93), we obtain the next iterative value of  $\lambda$ :

$$\lambda^{(1)} = \left[ \frac{\int_{-1/2}^{1/2} \frac{dz}{\sqrt{\lambda_1}} + \int_{-1/2}^{1/2} \frac{dz}{\sqrt{\lambda_2}}}{2\rho + \int_{-1/2}^{1/2} \frac{dz}{\lambda_1} + \int_{-1/2}^{1/2} \frac{dz}{\lambda_2}} \right]^2. \quad (95)$$

If  $\lambda^{(1)} \leq \lambda^{(0)}$ , then  $\lambda = \lambda^{(1)}$  and we can now determine  $q_{11}$  and  $q_{22}$  in (39). On the other hand, if  $\lambda^{(1)} > \lambda^{(0)}$ , we have to determine the subintervals  $sp_1$  and  $sp_2$  (supports) in  $|z| \leq 1/2$  over which  $q_{11}$  and  $q_{22}$  are respectively nonnegative. We determine these supports from (39). The next iterative value of  $\lambda$  is now

$$\lambda^{(2)} = \left[ \frac{\int_{sp_1} \frac{dz}{\sqrt{\lambda_1}} + \int_{sp_2} \frac{dz}{\sqrt{\lambda_2}}}{2\rho + \int_{sp_1} \frac{dz}{\lambda_1} + \int_{sp_2} \frac{dz}{\lambda_2}} \right]^2. \quad (96)$$

From  $\lambda^{(2)}$  we obtain the new supports  $sp_1$  and  $sp_2$  and continue iterating for  $\lambda^{(3)}$  and so on until

$$1 - \delta \leq \frac{\lambda^{(m)}}{\lambda^{(m+1)}} \leq 1 + \delta; \quad \delta \ll 1. \quad (97)$$

In our subsequent calculations, we used  $\delta = 0.000001$ . Typically, we obtained  $\lambda$  within two iterations. The highest number of iterations that we observed was four.

## APPENDIX C

### Upper Bound on Error Rate and Channel Data Rate

If  $U(t)$  is the overall matrix system impulse response with Fourier transform  $TU(\omega)$ , then the vector signal sample is

$$S(0) = \sum_n U(-nT)A_n. \quad (98)$$

The sample in "channel  $a$ " is

$$\begin{aligned} S(0)_a = & (a_{0r} + ia_{0\mathcal{I}})(U_{0r}^{11} + iU_{0\mathcal{I}}^{11}) \\ & + \sum'_n [(a_{nr} + ia_{n\mathcal{I}})(U_{nr}^{11} + iU_{n\mathcal{I}}^{11}) + (b_{nr} + ib_{n\mathcal{I}})(U_{nr}^{21} + iU_{n\mathcal{I}}^{21})] \\ & + (b_{0r} + ib_{0\mathcal{I}})(U_{0r}^{21} + iU_{0\mathcal{I}}^{21}) \end{aligned} \quad (99)$$

and in "channel  $b$ "

$$\begin{aligned} S(0)_b = & (b_{0r} + ib_{0\mathcal{I}})(U_{0r}^{22} + iU_{0\mathcal{I}}^{22}) \\ & + \sum'_n [(b_{nr} + ib_{n\mathcal{I}})(U_{nr}^{22} + iU_{n\mathcal{I}}^{22}) + (a_{nr} + ia_{n\mathcal{I}})(U_{nr}^{12} + iU_{n\mathcal{I}}^{12})] \\ & + (a_{0r} + ia_{0\mathcal{I}})(U_{0r}^{12} + iU_{0\mathcal{I}}^{12}), \end{aligned} \quad (100)$$

where  $A_n = \begin{pmatrix} a_n \\ b_n \end{pmatrix}$  and the complex data symbols are

$$a_n = a_{nr} + ia_{n\mathcal{J}}$$

and

$$b_n = b_{nr} + ib_{n\mathcal{J}},$$

while the complex matrices are denoted

$$U_n = [U_{nr}^{ij} + iU_{n\mathcal{J}}^{ij}], \quad l, j = 1, 2$$

$$\text{with } \sum'_n = \sum_{n=0}^n.$$

Channels  $a$  and  $b$  have real and imaginary parts corresponding to the in-phase and quadrature components in QAM systems. Thus, the four data signals are expressed as

$$\begin{aligned} \text{Re}[S(0)_a] = & a_{0r}U_{0r}^{11} - \alpha_{0\mathcal{J}}U_{0\mathcal{J}}^{11} \\ & + \sum'_n [(a_{nr}U_{nr}^{11} - a_{n\mathcal{J}}U_{n\mathcal{J}}^{11}) + (b_{nr}U_{nr}^{21} - b_{n\mathcal{J}}U_{n\mathcal{J}}^{21})] \\ & + b_{0r}U_{0r}^{21} - b_{0\mathcal{J}}U_{0\mathcal{J}}^{21}, \quad (101) \end{aligned}$$

$$\begin{aligned} \text{Im}[S(0)_a] = & a_{0\mathcal{J}}U_{0r}^{11} + a_{0r}U_{0\mathcal{J}}^{11} \\ & + \sum'_n [(a_{n\mathcal{J}}U_{nr}^{11} + a_{nr}U_{n\mathcal{J}}^{11}) + (b_{n\mathcal{J}}U_{nr}^{21} + b_{nr}U_{n\mathcal{J}}^{21})] \\ & + b_{0\mathcal{J}}U_{0r}^{21} + b_{0r}U_{0\mathcal{J}}^{21}, \quad (102) \end{aligned}$$

$$\begin{aligned} \text{Re}[S(0)_b] = & b_{0r}U_{0r}^{22} - b_{0\mathcal{J}}U_{0\mathcal{J}}^{22} \\ & + \sum'_n [(b_{nr}U_{nr}^{22} - b_{n\mathcal{J}}U_{n\mathcal{J}}^{22}) + (a_{nr}U_{nr}^{12} - a_{n\mathcal{J}}U_{n\mathcal{J}}^{12})] \\ & + a_{0r}U_{0r}^{12} - \alpha_{0\mathcal{J}}U_{0\mathcal{J}}^{12}, \quad (103) \end{aligned}$$

and

$$\begin{aligned} \text{Im}[S(0)_b] = & b_{0\mathcal{J}}U_{0r}^{22} + b_{0r}U_{0\mathcal{J}}^{22} \\ & + \sum'_n [(b_{n\mathcal{J}}U_{nr}^{22} + b_{nr}U_{n\mathcal{J}}^{22}) + (a_{n\mathcal{J}}U_{nr}^{12} + a_{nr}U_{n\mathcal{J}}^{12})] \\ & + \alpha_{0\mathcal{J}}U_{0r}^{12} + a_{0r}U_{0\mathcal{J}}^{12}. \quad (104) \end{aligned}$$

Equations (101) through (104) are the distorted data symbols. In the ideal case, when there is no intersymbol and/or cross-polarization interference

$$U_n = \delta_{no}I \quad (105)$$

and, as can be verified,

$$\begin{aligned} \text{Re}[S(0)_a] = a_{0r}, \quad \text{Im}[S(0)_a] = \alpha_{0\mathcal{J}}, \\ \text{Re}[S(0)_b] = b_{0r} \quad \text{and} \quad \text{Im}[S(0)_b] = b_{0\mathcal{J}}. \end{aligned}$$

We now assume that the data symbols  $\{a_{nr}, a_{n\mathcal{J}}, b_{nr}, b_{n\mathcal{J}}\}$  take on values from the positive and negative integers,  $\pm 1 \pm 3 \pm \dots \pm (L - 1)$ ,  $L$  even. Consequently, we can calculate

$$E(a_{nr}^2) = E(a_{n\mathcal{J}}^2) = E(b_{nr}^2) = E(b_{n\mathcal{J}}^2) = \frac{L^2 - 1}{3}. \quad (106)$$

From the noise model described in (3) and (4) in the text we deduce that the output noise in channels  $a$  and  $b$  can be represented, respectively, as

$$\nu_{0a}(0) = \int_{-\infty}^{\infty} W_{11}(-\tau)\nu_a(\tau)d\tau + \int_{-\infty}^{\infty} W_{21}(-\tau)\nu_b(\tau)d\tau \quad (107)$$

and

$$\nu_{0b}(0) = \int_{-\infty}^{\infty} W_{12}(-\tau)\nu_a(\tau)d\tau + \int_{-\infty}^{\infty} W_{22}(-\tau)\nu_b(\tau)d\tau.$$

Thus, the four noise samples added to (101) through (104) are  $\text{Re}[\nu_{0a}(0)]$ ,  $\text{Im}[\nu_{0a}(0)]$ ,  $\text{Re}[\nu_{0b}(0)]$ , and  $\text{Im}[\nu_{0b}(0)]$ . They represent four zero-mean Gaussian random variables and the respective variances can be readily calculated from (107):

$$\begin{aligned} E[\text{Re}\{\nu_{0a}(0)\}]^2 &= E[\text{Im}\{\nu_{0a}(0)\}]^2 = \sigma_{0ar}^2 \\ &= \sigma_{0a\mathcal{J}}^2 = \frac{N_o}{2} \left[ \int_{-\infty}^{\infty} (|W_{11}(t)|^2 + |W_{21}(t)|^2) dt \right], \end{aligned} \quad (108)$$

while

$$\begin{aligned} E[\text{Re}\{\nu_{0b}(0)\}]^2 &= E[\text{Im}\{\nu_{0b}(0)\}]^2 = \sigma_{0br}^2 \\ &= \sigma_{0b\mathcal{J}}^2 = \frac{N_o}{2} \left[ \int_{-\infty}^{\infty} (|W_{22}(t)|^2 + |W_{12}(t)|^2) dt \right]. \end{aligned} \quad (109)$$

In accordance with the definition in (106), which presumes that thresholds are set at  $0, \pm 2, \pm 4, \dots, \pm(L - 2)$ , an upper bounding on the probability of error in each of the four channels can be obtained using Chernoff bounding techniques.<sup>32,24</sup>

The resulting bounds are

$$P_{e1} \leq \exp \left( - \frac{\frac{1}{2} \frac{(U_0^{11})^2}{\frac{N_o}{2T} (|W_{11}|^2 + |W_{21}|^2) + \frac{\sigma_a^2}{2}}}{\left[ \sum_n (|U_n^{11}|^2 + |U_n^{21}|^2) + |U_0^{21}|^2 \right]} \right) \quad (110)$$

and

$$P_{e2} \leq \exp \left( - \frac{1}{2} \frac{N_o (U_0^{22})^2}{2T (\langle |W_{22}|^2 + |W_{12}|^2 \rangle) + \frac{\sigma_d^2}{2}} \cdot \left[ \sum'_n (|U_n^{22}|^2 + |U_n^{12}|^2) + |U_o^{12}|^2 \right] \right), \quad (111)$$

where we used the following notation:

$$\langle W_{ij} \rangle = \int_{-1/2}^{1/2} W_{ij}(z) dz, \quad i, j = 1, 2.$$

Note that since the matrices  $U_n$  are Hermitian,  $U_0^{11}$  and  $U_0^{22}$  are real valued. Also, from (101) through (104) it is clear that the error rate in the quadrature component is equal to the in-phase component in each of the respective channels. But, clearly,  $P_{e1}$  need not equal  $P_{e2}$ .

We now relate the denominator of the exponents in (110) and (111) to the individual mean-square errors. Utilizing (17a) and (17b) in (15), we obtain for the optimum squared error matrix

$$\xi_o = \sigma_d^2 \left[ I - U_o - U_o^\dagger + \sum_n U_n U_n^\dagger + \frac{N_o}{T} \int_{-\infty}^{\infty} W_o(z) W_o^\dagger(z) dz \right]. \quad (112)$$

Since the diagonal entries of  $\xi$ ,  $\xi_{11}$ , and  $\xi_{22}$  are equal to  $\text{MSE}_1$  and  $\text{MSE}_2$ , respectively, we reduce (110) and (111) with the aid of (18a) to

$$P_{e1} \leq \exp \left[ - \frac{(U_o^{11})^2}{\xi_{11} - \sigma_d^2 (1 - U_o^{11})^2} \right] = \exp \left\{ - \frac{\mu_1}{\sigma_d^2} \right\} \quad (113)$$

$$P_{e2} \leq \exp \left[ - \frac{(U_o^{22})^2}{\xi_{22} - \sigma_d^2 (1 - U_o^{22})^2} \right] = \exp \left\{ - \frac{\mu_2}{\sigma_d^2} \right\} \quad (114)$$

with

$$\mu_j = (1 - \xi_{jj}/\sigma_d^2)^2 / [\xi_{jj}/\sigma_d^2 - (\xi_{jj}/\sigma_d^2)^2], \quad j = 1, 2.$$

As can be seen, when the mean-square errors are very small we get the well-known relationship<sup>24</sup>

$$P_{ej} \leq \exp[-1/\text{MSE}_j]; \quad j = 1, 2. \quad (115)$$

One can easily ascertain that in all the formulas previously derived,  $\xi/\sigma_d^2$  and  $\xi_{op}/\sigma_d^2$  are independent of  $\sigma_d^2$ . Utilizing (106) in (113) and (114), we obtain after some manipulations the normalized data rates of the two channels,  $\mathbf{I}_1$  and  $\mathbf{I}_2$ ,

$$\mathbf{I}_1 = \log_2(1 + 1.5\mu_1 / |\ln P_{e1}|) \quad (116)$$

$$\mathbf{I} = \log(1 + 1.5\mu / |\ln P|). \quad (117)$$

These two quantities, sometimes referred to as efficiency indices, will be used to compare the performance of the various equalization methods.

#### **AUTHORS**

**Noach Amitay**, B.Sc. and Diplome Ingenieur (Electrical Engineering), Technion, Israel Institute of Technology, in 1953 and 1954, respectively; M.Sc. and Ph.D. (Electrical Engineering), Carnegie Institute of Technology, Pittsburgh, in 1957 and 1960, respectively; AT&T Bell Laboratories, 1962—. From 1954 to 1956 Mr. Amitay directed the Electronic Measuring and Aligning Instruments department of the Signal Corps, Defense Army of Israel. He was a Research Assistant in 1957, a Project Engineer in 1958, and an Instructor in 1959, all at Carnegie Institute of Technology. He served as a consultant on electronic instrumentation for Magnetics, Inc., Butler, Pennsylvania, from 1957 to 1958 and from 1960 to 1961. From 1960 to 1962 he was an Assistant Professor in the Department of Electrical Engineering, Carnegie Institute of Technology and a part-time employee of the New Products Laboratories of Westinghouse Electric Corporation, Pittsburgh. At AT&T Bell Laboratories, Mr. Amitay first conducted research in phased array radar and communication antennas, electromagnetic theory, and numerical methods and analysis. In recent years, he has been involved in studies of satellite communication antennas and digital radio communications. At present, Mr. Amitay is a Member of the Technical Staff in the Radio Systems Research department at AT&T Bell Laboratories. Fellow, IEEE; Member, International Scientific Radio Union.

**Jack Salz**, B.S.E.E., 1955, M.S.E., 1956, and Ph.D., 1961, University of Florida; AT&T Bell Laboratories, 1961—. Mr. Salz first worked on the electronic switching system. Since 1968 he has supervised a group engaged in theoretical studies in data communications and is currently a member of the Communications Methods Research department. During the academic year 1967–68, he was on leave as Professor of Electrical Engineering at the University of Florida. He was a visiting lecturer at Stanford University in Spring 1981 and a visiting MacKay Lecturer at the University of California, at Berkeley, in Spring 1983.



# PAPERS BY AT&T BELL LABORATORIES AUTHORS

## COMPUTING/MATHEMATICS

- Arthurs E., Stuck B. W., **A Modified Access Policy for Ethernet Version 1.0 Data Link Layer (Letter)**. *IEEE Commun* 32(8):977-979, Aug 1984.
- Baker B. S., Coffman E. G., **Insertion and Compaction Algorithms in Sequentially Allocated Storage**. *SIAM J Comp* 13(3):600-609, Aug 1984.
- Balkanski M., Haro E., Espinosa G. P., Phillips J. C., **Raman-Scattering Study of Glass Crystallization Kinetics**. *Sol St Comm* 51(8):639-642, Aug 1984.
- Barron E. N., Evans L. C., Jensen R., **Viscosity Solutions of Isaacs Equations and Differential Games With Lipschitz Controls**. *J Diff Equa* 53(2):213-233, Jul 30 1984.
- Belic M. R., Vukovic S., Lax M., **New Efficient Algorithm for Solution of the Driven Nonlinear Schrodinger Equation**. *Comput Phys* 32(3):239-243, Jun 1984.
- Bentley J., **Programming Pearls**. *Comm ACM* 27(7):630-636, Jul 1984.
- Borcherds R. E., Conway J. H., Queen L., Sloane N. J. A., **A Monster Lie Algebra**. *Adv Math* 53(1):75-79, Jul 1984.
- Calderbank A. R. et al., **Increasing Sequences With Nonzero Block Sums and Increasing Paths in Edge-Ordered Graphs**. *Discr Math* 50(1):15-28, Jun 1984.
- Cardelli L., **A Semantics of Multiple Inheritance**. *Lect N Comp* 173:51-67, 1984.
- Chen R. W., Hwang F. K., **Some Theorems, Counterexamples, and Conjectures in Multinomial Selection Theory**. *Comm St-The* 13(10):1289-1298, 1984.
- Coffman E. G., Yannakakis M., **Permuting Elements Within Columns of a Matrix in Order to Minimize Maximum Row Sum**. *Math Oper R* 9(3):384-390, Aug 1984.
- Conway J. H., Sloane N. J. A., **On the Voronoi Regions of Certain Lattices**. *SIAM J Alg* 5(3):294-305, Sep 1984.
- Desarbo W. S., Green P. E., **Concepts, Theory, and Techniques—Choice-Constrained Conjoint Analysis**. *Decision Sc* 15(3):297-323, Sum 1984.
- Fishburn P. C., **On Harsanyi Utilitarian Cardinal Welfare Theorem**. *Theor Decis* 17(1):21-28, Jul 1984.
- Gay D. M., **A Trust-Region Approach to Linearly Constrained Optimization**. *Lect N Math* 1066:72-105, 1984.
- Gehani N. H., **Broadcasting Sequential Processes (BSP)**. *IEEE Soft E* 10(4):343-351, Jul 1984.
- Gilbert J. R. et al., **A Separator Theorem for Chordal Graphs**. *SIAM J Alg* 5(3):306-313, Sep 1984.
- Harrison J. M., Shepp L. A., **A Tandem Storage System and Its Diffusion Limit**. *Stoch Pr Ap* 16(3):257-274, Mar 1984.
- Kaufman L., Warner D. D., **High-Order, Fast-Direct Methods for Separable Elliptic Equations**. *SIAM J Num* 21(4):672-694, Aug 1984.
- Lagarias J. C., **Performance Analysis of Shamir Attack on the Basic Merkle-Hellman Knapsack Cryptosystem**. *Lect N Comp* 172:312-323, 1984.
- Leung J. Y. T. et al., **On Some Variants of the Bandwidth Minimization Problem**. *SIAM J Comp* 13(3):650-667, Aug 1984.
- Logan B. F., **On the Eigenvalues of a Certain Integral Equation**. *SIAM J Math* 15(4):712-717, Jul 1984.
- Nair V. N., **Confidence Bands for Survival Functions With Censored Data—A Comparative Study**. *Technomet* 26(3):265-275, Aug 1984.
- Odlyzko A. M., **Cryptanalytic Attacks on the Multiplicative Knapsack Cryptosystem and on Shamir Fast Signature Scheme**. *IEEE Info T* 30(4):594-601, Jul 1984.
- Ozarow L. H. et al., **An Achievable Region and Outer Bound for the Gaussian Broadcast Channel With Feedback (Letter)**. *IEEE Info T* 30(4):667-671, Jul 1984.
- Ozarow L. H., **The Capacity of the White Gaussian Multiple Access Channel With Feedback**. *IEEE Info T* 30(4):623-629, Jul 1984.

- Paige R., Tarjan R. E., **A Linear Time Algorithm to Solve the Single Function Coarsest Partition Problem.** Lect N Comp 172:371-379, 1984.
- Polak E., Wardi Y. Y., **A Study of Minimizing Sequences.** SIAM J Con 22(4):599-609, Jul 1984.
- Reiman M. I., **Open Queuing Networks in Heavy Traffic.** Math Oper R 9(3):441-458, Aug 1984.
- Ritchie D. M., **Reflections on Software Research.** Comm ACM 27(8):758-760, Aug 1984.
- Snead B., Ho F., Engram B., **Operating System Features Real Time and Fault Tolerance.** Comput Des 23(9):177+, Aug 1984.
- Sternberg S., Knoll R. L., **Perception, Production, and Imitation of Time Ratios by Skilled Musicians.** Ann NY Acad 423(May):429-441, May 11 1984.
- Tarjan R. E., Yannakakis M., **Simple Linear-Time Algorithms to Test Chordality of Graphs, Test Acyclicity of Hypergraphs, and Selectively Reduce Acyclic Hypergraphs.** SIAM J Comp 13(3):566-579, Aug 1984.
- Thompson K., **Reflections on Trusting Trust.** Comm ACM 27(8):761-763, Aug 1984.
- Wallace J. J., Barnes W. W., **Designing for Ultrahigh Availability—The UNIX RTR Operating System.** Computer 17(8):31-39, Aug 1984.
- Willard D. E., **New Trie Data Structures Which Support Very Fast Search Operations.** J Comput Sy 28(3):379-394, Jun 1984.
- Wingo D. R., **Fitting Three-Parameter Lognormal Models by Numerical Global Optimization—An Improved Algorithm.** Comput Stat 2(1):13-25, Jun 1984.
- Wyner A. D., Landau H. J., **Optimum Waveform Signal Sets With Amplitude and Energy Constraints.** IEEE Info T 30(4):615-622, Jul 1984.

## ENGINEERING

- Abidi A. A., **The Effects of Small Noise on Implicitly Defined Nonlinear Dynamical Systems—Comment (Letter).** IEEE Circ S 31(8):750-751, Aug 1984.
- Agrawal G. P., Olsson N. A., Dutta N. K., **Reduced Chirping in Coupled-Cavity-Semiconductor Lasers.** Appl Phys L 45(2):119-121, Jul 15 1984.
- Anderson W. T., **Consistency of Measurement Methods for the Mode Field Radius In a Single-Mode Fiber.** J Lightw T 2(2):191-197, Apr 1984.
- Bosch M. A., Herbst D., Grogan J. K., Lemons R. A., Tennant D. M., Tewksbury S. K., **NMOS Silicon Gate Transistors in Large-Area Laser-Crystallized Silicon Layers.** IEE Proc-I 131(4):121-124, Aug 1984.
- Bowers J. E., Wilt D. P., **Optimum Design of 1.55- $\mu$ m Double Heterostructures and Ridge-Waveguide Lasers.** Optics Lett 9(8):330-332, Aug 1984.
- Capasso F., **Multilayer Avalanche Photodiodes and Solid-State Photomultipliers.** Laser Foc B 20(7):84+, Jul 1984.
- Capasso F., Cho A. Y., Foy P. W., **Low-Dark-Current Low-Voltage 1.3-1.6  $\mu$ m Avalanche Photodiode With High-Low Electric-Field Profile and Separate Absorption and Multiplication Regions by Molecular-Beam Epitaxy.** Electr Lett 20(15):635-637, Jul 19 1984.
- Corwin W. L., **A Communications Network for the Summer Olympics.** IEEE Spectr 21(7):38-44, Jul 1984.
- Decker G. A., Giangrossi R. V., **High-Strength Low-Loss Flame Fusion Splicing of Singlemode Fibers for Undersea Cable.** P Soc Photo 479:53-55, 1984.
- Dutta N. K., Olsson N. A., Liou K. Y., **Effect of External Optical Feedback on Spectral Properties of External Cavity Semiconductor Lasers.** Electr Lett 20(14):588-589, Jul 5 1984.
- Eisenstein G., Tucker R. S., Kaminow I. P., Lee T. P., Burrus C. A., **InGaAsP 1.3  $\mu$ m Optical Amplifier-Modulator Integrated With a Fiber-Resonator Mode-Locked Laser.** Electr Lett 20(15):624-625, Jul 19 1984.
- Forrest S. R., Deimel P. P., Glacet J. Y., Logan R. A., **Narrow Spectral Width Surface Emitting LED for Long Wavelength Multiplexing Applications.** IEEE J Q El 20(8):906-912, Aug 1984.
- Forrest S. R., Kaplan M. L., Schmidt P. H., **Organic-on-Inorganic Semiconductor Contact Barrier Diodes. 2. Dependence on Organic Film and Metal Contact**

- Properties.** *J Appl Phys* 56(2):543-551, Jul 15 1984.
- Hasegawa A., **Generation of a Train of Soliton Pulses by Induced Modulational Instability** *Optics Lett* 9(7):288-290, Jul 1984.
- Hegarty J., Poulsen S. D., Jackson K. A., Kaminow I. P., **Low-Loss Single-Mode Wavelength Division Multiplexing With Etched Fiber Arrays.** *Electr Lett* 20(17):685-686, Aug 16 1984.
- Henry C. H., Kazarinov R. F., **Stabilization of Single Frequency Operation of Coupled-Cavity Lasers.** *IEEE J Q El* 20(7):733-744, Jul 1984.
- Huang A., **Architectural Considerations Involved in the Design of an Optical Digital Computer.** *P IEEE* 72(7):780-786, Jul 1984.
- Johnson, R., **The Wheatstone Bridge Reduction in Network Reliability Computations.** *IEEE Reliab* 32(4):374-378, Oct 1983.
- Jordan A. S., Osullivan T. D., **Determination of the Reliability of Aerospace Ni-Cd Batteries From Survival Data on Cells Fabricated Between 1964 and 1977.** *J Elchem So* 131(7):1492-1498, Jul 1984.
- Kleinman D. A., Auston D. H., **Theory of Electrooptic Shock Radiation in Non-linear Optical Media.** *IEEE J Q El* 20(8):964-970, Aug 1984.
- Lee T. P., Burrus C. A., Eisenstein G., Sessa W. B., Besomi P., **Amplifier Modulator Integrated With a Cleaved-Coupled-Cavity Injection Laser.** *Electr Lett* 20(15):625-627, Jul 19 1984.
- Levine B. F., Bethea C. G., Campbell J. C., **Near Room-Temperature 1.3- $\mu$ m Single Photon Counting With a InGaAs Avalanche Photodiode.** *Electr Lett* 20(14):596-598, Jul 5 1984.
- Lin C. L., **Microwave-Frequency Intensity Modulation and Gain Switching in Semiconductor Injection Lasers.** *P Soc Photo* 477:2-11, 1984.
- Luryi S., **New Infrared Detector on a Silicon Chip.** *IEEE Device* 31(9):1135-1139, Sep 1984.
- Luss H. **A Capacity Expansion Model With Applications to Multiplexing in Communication Networks.** *IEEE Syst M* 14(3):419-423, May-Jun 1984.
- Nozari A., Arnold S. F., Pegden C. D., **Control Variates for Multipopulation Simulation Experiments.** *IIE Trans* 16(2):159-169, Jun 1984.
- Olsson N. A., Logan R. A., Johnson L. F., **Transmission Experiment at 3 GBit/s With Close Spaced Wavelength Division Multiplexed Single Frequency Lasers at 1.5  $\mu$ m.** *Electr Lett* 20(17):673-674, Aug 16 1984.
- Saksena V. R. et al., **Singular Perturbations and Time-Scale Methods in Control Theory—Survey, 1976-1983.** *Automatica* 20(3):273-293, May 1984.
- Sandberg I. W., **Existence and Evaluation of Almost Periodic Steady-State Responses of Mildly Nonlinear Systems.** *IEEE Circ S* 31(8):689-701, Aug 1984.
- Schlough S. D., Oconnell E. F., **Cleaning Processes for HICs With Solder Paste.** *IEEE Compon* 7(2):176-188, Jun 1984.
- Sears F. M., Cohen L. G., Stone J., **Interferometric Measurements of Dispersion-Spectra Variations in a Single-Mode Fiber.** *J Lightw T* 2(2):181-184, Apr 1984.
- Shively R. R., Robinson W. V., Orton D. E., **Cascading Transmission Gates to Enhance Multiplier Performance (Letter).** *IEEE Comput* 33(7): 677-679, Jul 1984.
- Stolen R. H., Ashkin A., Pleibel W., Dziedzic J. M., **In-Line Fiber-Polarization-Rocking Rotator and Filter.** *Optics Lett* 9(7):300-302, Jul 1984.
- Stolen R. H., Lee C., Jain R. K., **Development of the Stimulated Raman-Spectrum in Single-Mode Silica Fibers.** *J Opt Soc B* 1(4):652-657, Aug 1984.
- Tewksbury S. K., Biazzo M. R., Harrison T. R., Lindstrom T. L., Tennant D. M., Storz F. G., **Static and Nonequilibrium Transient Conductance at Strong Carrier Freeze-Out in a Buried Channel, Metal-Oxide-Semiconductor Transistor.** *J Appl Phys* 56(2):511-516, Jul 15 1984.
- Venkatesan T., Lemaire P. J., Wilkens B., Soto L., Gossard A. C., Wiegmann W., Jewell J. L., Gibbs H. M., Tarng S. S., **All-Optical Data Switching in an Optical-Fiber Link Using a GaAs Optical Bistable Device.** *Optics Lett* 9(7):297-299, Jul 1984.
- Wong W. C., Greenstein L. J., **Multipath-Fading Models and Adaptive Equalizers in Microwave Digital Radio.** *IEEE Commun* 32(8):928-934, Aug 1984.
- Zucker J. E., Finczuk A., Chemla D. S., Gossard A., Wiegmann W., **Delocalized Excitons in Semiconductor Heterostructures.** *Phys Rev B* 29(12):7065-7068, Jun 15 1984.

## PHYSICAL SCIENCES

- Agrawal G. P., **Atomic Coherence Effects in a Two-Mode Laser With Coupled Transitions.** Phys Rev A 30(2):884-889, Aug 1984.
- Agrawal G. P., Dutta N. K., **Optical Bistability in Coupled-Cavity Semiconductor Lasers.** J Appl Phys 56(3):664-669, Aug 1 1984.
- Andrei E. Y., Grimes C. C., Adams G., **Electrons on Helium—The Polaron Transition.** Surf Sci 142(1-3):104-106, Jul 1984.
- Aur S., Kofalt D., Waseda Y., Egami T., Chen H. S., Teo B. K., Wang R., **Use of a Semiconductor Detector in Anomalous (Resonance) X-Ray-Scattering Measurement of Local Structure of an Amorphous Alloy.** Nucl Inst A 222(1-2):259-261, May 15 1984.
- Auston D. H., Cheung K. P., Smith P. R., **Picosecond Photoconducting Hertzian Dipoles.** Appl Phys L 45(3):284-286, Aug 1 1984.
- Baraff G. A., Schluter M., **Calculation of the Total Energy of Charged Point Defects Using the Greens-Function Technique.** Phys Rev B 30(4):1853-1866, Aug 15 1984.
- Brawer S. A., **Theory of Relaxation in Viscous Liquids and Glasses.** J Chem Phys 81(2):954-975, Jul 15 1984.
- Bucksbaum P. H., Bokor J., **Rapid Melting and Regrowth Velocities in Silicon Heated by Ultraviolet Picosecond Laser Pulses.** Phys Rev L 53(2):182-185, Jul 9 1984.
- Budny R. et al., **Initial Results From the Scoop Limiter Experiment in PDX.** J Nucl Mat 121:294-303, 1984.
- Camlibel I., Chin A. K., Guggenheim H., Singh S., Van Uitert L. G., Zydzik G. J., **Fabrication of Low Dark-Current Planar Photodiodes Using an Open-Tube Method for Zn Diffusion Into InP and In<sub>0.53</sub>Ga<sub>0.47</sub>As.** J Elchem So 131(7):1687-1688, Jul 1984.
- Capaccioli M., Fasano G., Lake G., **Kinematical Tests for the Intrinsic Shapes of Galaxies.** M Not R Ast 209(2):317-334, Jul 15 1984.
- Capasso F., **New Multilayer and Graded Gap Optoelectronic and High-Speed Devices by Band-Gap Engineering.** Surf Sci 142(1-3):513-528, Jul 1984.
- Celli V., **Atom-Surface Interaction in Inelastic Scattering—He/Ag(111) (Letter).** Surf Sci 143(1):L376-L382, Jul 1984.
- Chabal Y. J., Raghavachari K., **Surface Infrared Study of Si(100)-(2×1)H.** Phys Rev L 53(3):282-285, Jul 16 1984.
- Chang A. M., Berglund P., Tsui D. C., Stormer H. L., Hwang J. C. M., **Higher-Order States in the Multiple-Series, Fractional, Quantum Hall Effect.** Phys Rev L 53(10):997-1000, Sep 3 1984.
- Chang A. M., Paalanen M. A., Stormer H. L., Hwang J. C. M., Tsui D. C., **Fractional Quantum Hall Effect at Low Temperatures.** Surf Sci 142(1-3):173-178, Jul 1984.
- Chen C. H., Phillips J. C., Bridenbaugh P. M., Aboav D. A., **Large-Scale Domain-Structure in Evaporated Thin Films of Chalcogenide Alloy Glasses.** J Non-Cryst 65(1):1-28, Jun 1984.
- Chin A. K., Zipfel G. L., Geva M., Camlibel I., Skeath P., Chin B. H., **Direct Evidence for the Role of Gold Migration in the Formation of Dark-Spot Defects in 1.3- $\mu$ m Inp/InGaAsP Light-Emitting Diodes.** Appl Phys L 45(1):37-39, Jul 1 1984.
- Cohen R. L., West K. W., **Flaws in Aluminum Spot Welds Observed by Electrical Measurements.** Welding J 63(8):21-23, Aug 1984.
- Cross M. C., Newell A. C., **Convection Patterns in Large Aspect Ratio Systems.** Physica D 10(3):299-328, Mar 1984.
- Davis G. T. et al., **Hysteresis in Copolymers of Vinylidene Fluoride and Trifluoroethylene.** Ferroelectr 57(1-4):73-84, 1984.
- Degani J., Wilt D. P., Besomi P., **Effect of Photocarrier Spreading on the Photoluminescence of Double Heterostructure Material.** J Appl Phys 56(2):468-476, Jul 15 1984.
- Delalande C., Ziemelis U. O., Bastard G., Voos M., Gossard A. C., Wiegmann W., **Photoluminescence and Excitation Spectroscopy in Coupled GaAs-Ga(Al)As Quantum Wells.** Surf Sci 142(1-3):498-503, Jul 1984.
- Dicenzo S. B., Wertheim G. K., Buchanan D. N., **Site Dependence of Core-Electron**

- Binding Energies of Adsorbates—I/Pt(111).** Phys Rev B 30(2):553-557, Jul 15 1984.
- Digiuseppe M. A., Chin A. K., Chin B. H., Lourenco J. A., Camlibel I., **The Effect of Melt-Carry-Over on the LPE Growth of Planar Buried InGaAsP/InP Double Heterostructures.** J Cryst Gr 67(1):1-7, Jun 1984.
- Dutt B. V., Roccasecca D. D., Temkin H., Bonner W. A., **A Novel Multislice LPE Boat. 1. Preliminary Results on InGaAs Alloys.** J Cryst Gr 66(3):525-530, May 1984.
- Dutta N. K., Nelson R. J., Wright P. O., Craft D. C., **Criterion for Improved Linearity of 1.3- $\mu$ m InGaAsP InP Buried-Heterostructure Lasers.** J Lightw T 2(2):160-164, Apr 1984.
- Dutta N. K., Olsson N. A., **Electroabsorption in InGaAsP-InP Double Heterostructures.** Electr Lett 20(15):634-635, Jul 19 1984.
- Farrell H. H., Stucki F., Anderson J., Frankel D. J., Lapeyre G. J., Levinson M., **Electronic Excitations on Si(100) (2  $\times$  1).** Phys Rev B 30(2):721-725, Jul 15 1984.
- Feibelman P. J., Hamann D. R., **Quantum-Size Effects in Work Functions of Freestanding and Adsorbed Thin Metal Films.** Phys Rev B 29(12):6463-6467, Jun 15 1984.
- Fisher D. S., **Random Walks in Random Environments.** Phys Rev A 30(2):960-964, Aug 1984.
- Fuoss P. H., Robinson I. K., **Apparatus for X-Ray-Diffraction in Ultrahigh Vacuum.** Nucl Inst A 222(1-2):171-176, May 15 1984.
- Gilmer G. H., **Models of Impurity Trapping During Rapid Solidification.** Mater Sci E 65(1):15-25, Jul 1984.
- Gossmann H. J., Feldman L. C., Gibson W. M., **Reordering of Reconstructed Si-Surfaces Upon Ge Deposition at Room Temperature.** Phys Rev L 53(3):294-297, Jul 16 1984.
- Gottscho R. A., Donnelly V. M., **Optical-Emission Actinometry and Spectral-Line Shapes in RF Glow Discharges.** J Appl Phys 56(2):245-250, Jul 15 1984.
- Greenblatt M., McCarroll W. H., Neifeld R., Croft M., Waszczak J. V., **Quasi Two-Dimensional Electronic-Properties of the Lithium Molybdenum Bronze. Li<sub>0.9</sub>Mo<sub>6</sub>O<sub>17</sub>.** Sol St Comm 51(9):671-674, Sep 1984.
- Greenside H. S., Coughran W. M., **Nonlinear Pattern Formation Near the Onset of Rayleigh-Bernard Convection.** Phys Rev A 30(1):398-428, Jul 1984.
- Haight R., Feldman L. C., Buck T. M., Gibson W. M., **Neutralization of Energetic He Ions Scattered From Clean and Cs-Covered Si(100).** Phys Rev B 30(2):734-740, Jul 15 1984.
- Hamm R. A., Vandenberg J. M., **A Study of the Initial Growth Kinetics of the Copper-Aluminum Thin-Film Interface Reaction by Insitu X-Ray-Diffraction and Rutherford Backscattering Analysis.** J Appl Phys 56(2):293-299, Jul 15 1984.
- Heffner W., **Glaucus Equation and Thermodynamics of a Bistable Liquid-Crystal Cell.** J Appl Phys 56(2):286-292, Jul 15 1984.
- Hensel J. C., Dynes R. C., Halperin B. I., Tsui D. C., **Scattering and Adsorption of Ballistic Phonons by the Electron Inversion Layer in Silicon—Theory and Experiment.** Surf Sci 142(1-3):249-255, Jul 1984.
- Hensel J. C., Tung R. T., Poate J. M., Unterwald F. C., **Transport Studies in Single-Crystal Films of CoSi<sub>2</sub> and NiSi<sub>2</sub>—A New Class of Quasi-Two-Dimensional Metals.** Surf Sci 142(1-3):37-42, Jul 1984.
- Huang A., **Parallel Algorithms for Optical Digital Computers.** P Soc Photo 422:13-17, 1984.
- Huse D. A., **Renormalization-Group Analysis of Layering Transitions in Solid Films.** Phys Rev B 30(3):1371-1376, Aug 1 1984.
- Jackson K. A., **Crystal-Growth Kinetics.** Mater Sci E 65(1):7-13, Jul 1984.
- Jackson S. A., Platzman P. M., **Polaronic Aspects of Spin Polarized Hydrogen on Films of Liquid He.** Sol St Comm 51(9):733-737, Sep 1984.
- Jackson S. A., Platzman P. M., **The Polaronic State of Two-Dimensional Electrons on the Surface of Liquid Helium.** Surf Sci 142(1-3):125-129, Jul 1984.
- Johnson P. D., Stoffel N. G., Klaffky R., Smith N. V., **A Normal-Incidence Monochromator Branch Line on the FEL Undulator for Experimental Studies in the 5-30 eV Region.** Nucl Inst A 222(1-2):66-69, May 15 1984.

- Kastalsky A., Hwang J. C. M., **Study of Persistent Photoconductivity Effect in N-Type Selectively Doped AlGaAs/GaAs Heterojunction.** *Sol St Comm* 51(5):317-322, Aug 1984.
- Katz H. E., Starnes W. H., **Quasi-Concerted Allylic Rearrangement in the Reaction of Allylic Chlorides With Methyltin Tris(Methanethiolate).** *J Org Chem* 49(15):2758-2761, Jul 27 1984.
- Kaye S. M. et al., **Attainment of High Confinement in Neutral Beam Heated Divertor Discharges in the PDX Tokamak.** *J Nucl Mat* 121:115-125, 1984.
- Kevan S. D., Stoffell N. G., **Metal Insulator Transition on the Ge(001) Surface.** *Phys Rev L* 53(7):702-705, Aug 13 1984.
- Kurtze D. A., Vansaarloos W., Weeks J. D., **Front Propagation in Self-Sustained and Laser-Driven Explosive Crystal Growth—Stability Analysis and Morphological Aspects.** *Phys Rev B* 30(3):1398-1415, Aug 1 1984.
- Lines M. E., Eibschutz M., **Theory of Hyperfine-Field and Exchange-Field Distributions in Amorphous Spermagnets With Application to Amorphous Yttrium Iron Garnet.** *Phys Rev B* 30(3):1416-1423, Aug 1 1984.
- Lourenco J. A., **A Defect Etchant for (100) InGaAsP.** *J Elchem So* 131(8):1914-1916, Aug 1984.
- Luryi S., Kastalsky A., **Anomalous Photomagneto-resistance Effect in Modulation-Doped AlGaAs GaAs Heterostructures.** *Appl Phys L* 45(2):164-167, Jul 15 1984.
- Marchut L., Buck T. M., Wheatley G. H., McMahon C. J., **Surface-Structure Analysis Using Low-Energy Ion Scattering. 1. Clean Fe(001).** *Surf Sci* 141(2-3):549-566, Jun 1984.
- Mattheiss L. F., Hamann D. R., **Bulk and Surface Electronic Structure of Hexagonal WC.** *Phys Rev B* 30(4):1731-1738, Aug 15 1984.
- Messina F. D., **Fluorescence Detection of Thin-Film Electrical Contact Lubricants.** *ASLE Trans* 27(3):237-242, Jul 1984.
- Miller D. A. B., Chemla D. S., Damen T. C., Gossard A. C., Wiegmann W., Wood T. H., Burrus C. A., **Novel Hybrid Optically Bistable Switch—The Quantum-Well Self-Electro-Optic Effect Device.** *Appl Phys L* 45(1):13-15, Jul 1 1984.
- Miller R. C., **Photoluminescence From GaAs-Al<sub>x</sub>Ga<sub>1-x</sub>As Quantum Wells With Nonuniform P-Type Doping Profiles.** *J Appl Phys* 56(4):1136-1140, Aug 15 1984.
- Miller R. C., Kleinman D. A., Gossard A. C., **Energy-Gap Discontinuities and Effective Masses for GaAs-Al<sub>x</sub>Ga<sub>1-x</sub>As Quantum Wells.** *Phys Rev B* 29(12):7085-7087, Jun 15 1984.
- Moerner W. E., Chraplyvy A. R., Sievers A. J., **Anharmonic Vibrational-Relaxation Dynamics for a Molecular Impurity Mode in Alkali-Halide Crystals.** *Phys Rev B* 29(12):6694-6708, Jun 15 1984.
- Murarka S. P., **Transition-Metal Silicides—Low Resistivity Alternatives for Polysilicon and Metals in Integrated Circuits.** *J Metal* 36(7):57-60, Jul 1984.
- Nakahara S., Gallagher P. K., Felder E. C., Lawry R. B., **Interaction Between Zinc Metallization and Indium-Phosphide.** *Sol St Elec* 27(6):557-564, Jun 1984.
- Orenstein J., Vardeny Z., Baker G. L., Eagle G., Etamad S., **Mechanism for Photo-generation of Charge Carriers in Polyacetylene.** *Phys Rev B* 30(2):786-794, Jul 15 1984.
- Paalanen M. A., Tsui D. C., Lin B. J., Gossard A. C., **Localization of Two-Dimensional Electrons in GaAs-Al<sub>x</sub>Ga<sub>1-x</sub>As Heterostructures.** *Surf Sci* 142(1-3):29-36, Jul 1984.
- Pearsall T. P., **Two-Dimensional Electronic Systems for High-Speed Device Applications.** *Surf Sci* 142(1-3):529-544, Jul 1984.
- Pearton S. J., **Neutralization of Shallow Acceptor Levels in Silicon by Atomic Hydrogen—Comment (Letter).** *Phys Rev L* 53(8):855, Aug 20 1984.
- Peeters F. M., **Two-Dimensional Wigner Crystal of Electrons on a Helium Film—Static and Dynamic Properties.** *Phys Rev B* 30(1):159-165, Jul 1 1984.
- Penna A. F. S., Shah J., Digiovanni A. E., Dentai A. G., **Luminescence in High-Purity In<sub>0.53</sub>Ga<sub>0.47</sub>As.** *Sol St Comm* 51(4):217-220, Jul 1984.
- Penna T. C., Tamargo M. C., Swartzwelder W. L., **A Study of the Growth of High-Purity InGaAs by Conventional LPE.** *J Cryst Gr* 67(1):27-30, Jun 1984.
- Phillips J. C., **Composite Ionberg-Iceberg Model of Aqueous Nonpolar Solvation and Water Exchange Reactions.** *J Chem Phys* 81(1):478-483, Jul 1 1984.

- Pinczuk A., Shah J., Stormer H. L., Miller R. C., Gossard A. C., Wiegmann W., **Investigation of Optical Processes in a Semiconductor Two-Dimensional Electron Plasma.** *Surf Sci* 142(1-3):492-497, Jul 1984.
- Portal J. C. et al., **Two-Dimensional Magnetophonon Resonance in GaInAs-InP and GaInAs-AlInAs Heterojunctions and Superlattices.** *Surf Sci* 142(1-3):368-374, Jul 1984.
- Raghavachari K., **Rotational Potential Surface for Alkanes—Basis Set and Electron Correlation Effects on the Conformations of Normal Butane.** *J Chem Phys* 81(3):1383-1388, Aug 1 1984.
- Raghavan R. S., **Raman Identified (Letter).** *Phys Today* 37(8):84, Aug 1984.
- Reents W. D., Muijsce A. M., **Ion Molecule Reactions of Silicon Tetrafluoride.** *Int J Mass* 59(1):65-75, Jun 25 1984.
- Reimann C. T., Johnson R. E., Brown W. L., **Sputtering and Luminescence in Electronically Excited Solid Argon.** *Phys Rev L* 53(6):600-603, Aug 6 1984.
- Rossetti R., Brus L. E., **Time-Resolved Raman-Scattering Study of Adsorbed, Semioxidized Eosin-Y Formed by Excited-State Electron Transfer into Colloidal TiO<sub>2</sub> Particles.** *J Am Chem S* 106(16):4336-4340, Aug 8 1984.
- Schumer R. A., Julesz B., **Binocular Disparity Modulation Sensitivity to Disparities Offset From the Plane of Fixation.** *Vision Res* 24(6):533-542, 1984.
- Schwartz B., Koszi L. A., Anthony P. J., Hartman R. L., **Thermal Annealing of Proton-Bombarded GaAs and (Al, Ga)As.** *J Elchem So* 131(7):1703-1707, Jul 1984.
- Schwartz G. P., Gualtieri G. J., Dubois L. H., Bonner W. A., Ballman A. A., **Free Carrier Reduction in Vacuum-Annealed S-Doped, Sn-Doped, and Ge-Doped (100) InP.** *J Elchem So* 131(7):1716-1720, Jul 1984.
- Shapira Y., Erillson L. J., Heller A., **Origin of Surface and Metal-Induced Interface States in InP.** *Phys Rev B* 29(12):6824-6832, Jun 15 1984.
- Shelnutt J. A. et al., **Core Expansion and Electronic Structure of the Porphyrin in the Neutral pH Form of Copper Cytochrome-C.** *Biochem* 23(17):3946-3954, Aug 14 1984.
- Sinvani M. et al., **Desorption of Helium Atoms From Thin Films.** *Phys Rev B* 30(3):1231-1248, Aug 1 1984.
- Skocpol W. J., Jackel L. D., Howard R. E., Craighead H. G., Fetter L. A., Mankiewich P. M., Grabbe P., Tennant D. M., **Magnetoconductance and Quantized Confinement in Narrow Silicon Inversion Layers.** *Surf Sci* 142(1-3):14-18, Jul 1984.
- Slusher R. E. et al., **New Fluctuation Phenomena in the H-Mode Regime of Poloidal-Diverter Tokamak Plasmas.** *Phys Rev L* 53(7):667-670, Aug 13 1984.
- Stark A. A., **Kinematics of Molecular Clouds. 1. Velocity Dispersion in the Solar Neighborhood.** *Astrophys J* 281(2):624-633, Jun 15 1984.
- Stavola M., Levinson M., Benton J. L., Kimerling L. C., **Extrinsic Self-Trapping and Negative U in Semiconductors—a Metastable Center in InP.** *Phys Rev B* 30(2):832-839, Jul 15 1984.
- Stillinger F. H., Weber T. A., **Packing Structures and Transitions in Liquids and Solids.** *Science* 225(4666):983-989, Sep 7 1984.
- Stormer H. L., **Novel Physics in Two Dimensions With Modulation-Doped Heterostructures.** *Surf Sci* 142(1-3):130-146, Jul 1984.
- Swaminathan V., Zilko J. L., Nygren S. F., **Doping of Al<sub>0.35</sub>Ga<sub>0.65</sub>As Grown by Metal Organic-Chemical Vapor Deposition With Zn or Se.** *Mater Lett* 2(48):308-312, May 1984
- Takase Y. et al., **Observation of Parametric Decay by Use of CO<sub>2</sub> Laser Scattering From a Plasma.** *Phys Rev L* 53(3):274-277, Jul 16 1984.
- Thiry P., Bennett P. A., Kevan S. D., Royer W. A., Chaban E. E., Rowe J. E., Smith N. V., **A 6m Toroidal-Grating-Monochromator Beam Line for High Momentum Resolution Photoelectron Spectroscopy.** *Nucl Inst A* 222(1-2):85-90, May 15 1984.
- Thiry P., Smith N. V., **Nonstigmatic Input Optics for Toroidal Grating Monochromators.** *Nucl Inst A* 222(1-2):91-94, May 15 1984.
- Thomsen C., Strait J., Vardeny Z., Maris H. J., Tauc J., Hauser J. J., **Coherent Phonon Generation and Detection by Picosecond Light Pulses.** *Phys Rev L* 53(10):989-992, Sep 3 1984.
- Thurston R. N., Cheng J., Meyer R. B., Boyd G. D., **Physical Mechanisms of dc Switching in a Liquid-Crystal Bistable Boundary-Layer Display.** *J Appl Phys* 56(2):263-272, Jul 15 1984.

- Tukey J. W., **Use of Spatial Analysis in Mineral-Resource Evaluation—Comments (Letter)**. *J Int A Mat* 16(6):591-594, Aug 1984.
- Tyson J. A., Valdes F., Jarvis J. F., Mills A. P., **Galaxy Mass Distribution From Gravitational Light Deflection**. *Astrophys J* 281(2):L 59-L 62, Jun 15 1984.
- Venkatesan T. et al., **Measurement of Thermodynamic Parameters of Graphite by Pulsed-Laser Melting and Ion Channeling**. *Phys Rev L* 53(4):360-363, Jul 23 1984.
- Vonseggern H., Gross B., Berkley D. A., **Constant Hole Schubweg in Teflon FEP (Fluorinated Ethylene Propylene Copolymer)**. *Appl Phys A* 34(3):163-166, Jul 1984.
- Walden R. W., Kolodner P., Passner A., Surko C. M., **Nonchaotic Rayleigh-Benard Convection With Four and Five Incommensurate Frequencies**. *Phys Rev L* 53(3):242-245, Jul 16 1984.
- Warren W. W., **Comment on the Enhanced Magnetic Susceptibility of Expanded Liquid Cesium (Letter)**. *Phys Rev B* 29(12):7012-7013, Jun 15 1984.
- Wesner D. A., Johnson P. D., Smith N. V., **Photoemission Spectra and Band Structures of D-Band Metals. 11. Inverse Photoemission From Pd(111)**. *Phys Rev B* 30(2):503-506, Jul 15 1984.
- Wilt D. P., Karlicek R. F., Strege K. E., Dautremont-Smith W. C., Dutta N. K., Flynn E. J., Johnston W. D., Nelson R. J., **Channelled-Substrate Buried Heterostructure InGaAsP InP Lasers With Vapor-Phase Epitaxial Base Structure and Liquid-Phase Epitaxial Regrowth**. *J Appl Phys* 56(3):710-712, Aug 1 1984.
- Woolery G. L., Powers L., Peisach J., Spiro T. G., **X-Ray Absorption Study of Rhus Laccase—Evidence for a Copper-Copper Interaction, Which Disappears on Type-2 Copper Removal**. *Biochem* 23(15):3428-3434, Jul 17 1984.
- Woolery G. L., Powers L., Winkler M., Solomon E. I., Lerch K., Spiro T. G., **Extended X-Ray Absorption Fine-Structure Study of the Coupled Binuclear Copper Active Site of Tyrosinase From Neurospora-Crassa**. *Bioc Biop A* 788(2):155-161, Jul 31 1984.
- Worklock J. M., Maciel A. C., Petrou A., Perry C. H., Aggarwal R. L., Smith M., Gossard A. C., Wiegmann W., **Magneto-Optical Studies of Two-Dimensional Electrons In MQW Heterostructures**. *Surf Sci* 142(1-3):486-491, Jul 1984.
- Yoshioka D., **Ground-State of the Two-Dimensional Charged Particles in a Strong Magnetic Field and the Fractional Quantum Hall Effect**. *Phys Rev B* 29(12):6833-6839, Jun 15 1984.
- Yurke B. et al., **Back-Action-Evading Measurement of Optical Fields With the Use of a Four-Wave Mixer**. *Phys Rev A* 30(2):895-900, Aug 1984.

## SOCIAL AND LIFE SCIENCES

- Foltz G. S., Poltrock S. E., Potts G. R., **Mental Comparison of Size and Magnitude—Size Congruity Effects**. *J Exp Psy L* 10(3):442-453, Jul 1984.
- Gelperin A., Culligan N., **Invitro Expression of Invivo Learning by an Isolated Molluscan CNS**. *Brain Res* 304(2):207-213, Jun 25 1984.
- Gerjuoy E., Hohenberg P., Feshbach H., Heicklen J., Jackson J. D., Lebowitz J. L., Pershan P. S., **We Cannot Ignore Soviet Rights Violations (Letter)**. *Phys Today* 37(7):11+, Jul 1984.
- Herrstrom D. S., **Technical Writing as Mapping Description Onto Diagram—The Graphic Paradigms of Explanation**. *J Tech Wr C* 14(3):223-240, 1984.
- Johnson M. K., Kahan T. L., Raye C. L., **Dreams and Reality Monitoring**. *J Exp Psy G* 113(3):329-344, Sep 1984.
- Kat D., Samuel A. G., **More Adaptation of Speech by Nonspeech**. *J Exp Psy P* 10(4):512-525, Aug 1984.
- Niklasson G. A., Craighead H. G., **Threshold Laser Powers of Textured Optical Storage Media**. *Opt Eng* 23(4):443-447, Jul-Aug 1984.
- Pearson D. S., Helfand E., **Rheological Behavior of Branched Polymer Molecules**. *Faraday Sym* (18):189-197, 1983.

## **SPEECH/ACOUSTICS**

Haavasoja T., Stormer H. L., Bishop D. J., Narayanamurti V., Gossard A. C., Wiegmann W., **Magnetization Measurements on a Two-Dimensional Electron System.** Surf Sci 142(1-3):294-297, Jul 1984.

Penna A. F. S. et al., **Spatial Variation of Bandgap Energy in  $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ .** Sol St Comm 51(6):425-428, Aug 1984.



# CONTENTS, JANUARY 1985

## Part 1

Optimal Resource Allocation for Two Processes

C. Courcoubetis

An Infant Mortality and Long-Term Failure Rate Model for Electronic Equipment

D. P. Holcomb and J. C. North

Waiting Time Convexity in the M/G/1 Queue

D. L. Jagerman

Accumulation of Jitter: A Stochastic Model

C. Chamzas

Nonlocal Input-Output Expansions

I. W. Sandberg

Effects of Channel Impairment on the Performance of an In-band Data-Driven Echo Canceler

J. J. Werner

Effects of Biases on Digitally Implemented Data-Driven Echo Cancelers

J. M. Cioffi and J. J. Werner

## Part 2

### SYSTEM 75

Introduction and Overview

A. Feiner, E. J. Rodriguez, and C. D. Weiss

Communications and Control Architecture

L. A. Baxter, P. R. Berkowitz, C. A. Buzzard, J. J. Horenkamp, and F. E. Wyatt

Physical Architecture and Design

A. S. Loverde, H. D. Frisch, C. R. Lindemulder, and D. Baker

Switch Services Software

W. Densmore, R. J. Jakubek, M. J. Miracle, and J. H. Sun

System Management

H. K. Woodland, G. A. Reisner, and A. S. Melamed

Maintenance Architecture

K. S. Lu, J. D. Price, and T. L. Smith

The Oryx/Pecos Operating System

G. R. Sager, J. A. Melber, and K. T. Fong

Project Development Environment

T. S. Kennedy, D. A. Pezzutti, and T. L. Wang

Software Development Tools

T. J. Pedersen, J. E. Ritacco, and J. A. Santillo

GAMUT: A Message Utility System for Automatic Testing

C. J. Lake, J. J. Shanley, and S. M. Silverstein

Introduction Activities and Results

M. A. McFarland and J. A. Miller

STATEMENT OF OWNERSHIP, MANAGEMENT AND CIRCULATION

(Required by 39 U.S.C. 3685)

1. Title of publication: AT&T Bell Laboratories Technical Journal.
- 1B. Publication No. 005-8580.
2. Date of filing: November, 1984.
3. Frequency of issue: Monthly, except combined May-June and July-August.
4. Publication office: AT&T Bell Laboratories, 101 J. F. Kennedy Pkwy, Short Hills, NJ 07078.
5. Headquarters: AT&T, 550 Madison Ave., New York, NY 10022.
6. Publisher: AT&T, 550 Madison Ave., New York, NY 10022. Editor: Bernard G. King, AT&T Bell Laboratories, 101 J. F. Kennedy Pkwy, Short Hills, NJ 07078. Managing Editor: Pierce Wheeler, AT&T Bell Laboratories, 101 J. F. Kennedy Pkwy, Short Hills, NJ 07078.
7. Owner: AT&T, 550 Madison Ave., New York, NY 10022.
8. Bondholders, mortgages and other security holders: none.
9. Extent and nature of circulation:

	Average number of copies of each issue during preceding 12 months	Actual number of copies of single issue published nearest to filing date
A. Total no. copies printed (Net press run) .....	<u>12,295</u>	<u>12,597</u>
B. Paid circulation		
1. Sales through dealers and carriers, street vendors and counter sales .....	<u>2,313</u>	<u>2,494</u>
2. Mail subscriptions .....	<u>4,486</u>	<u>4,836</u>
C. Total paid circulation .....	<u>6,799</u>	<u>7,330</u>
D. Free distribution by mail, carrier or other means, samples, complimentary and other free copies .....	<u>4,701</u>	<u>4,472</u>
E. Total distribution (sum of C and D) .....	<u>11,500</u>	<u>11,802</u>
F. Copies not distributed		
1. Office use, left over, unaccounted, spoiled after printing .....	<u>795</u>	<u>795</u>
2. Returns from news agents .....	None	None
G. Total (sum of E and F—should equal net press run shown in A) .....	<u>12,295</u>	<u>12,597</u>

I certify that the statements made by me above are correct and complete.

Pierce Wheeler, Managing Editor

**AT&T BELL LABORATORIES TECHNICAL JOURNAL** is abstracted or indexed by *Abstract Journal in Earthquake Engineering, Applied Mechanics Review, Applied Science & Technology Index, Chemical Abstracts, Computer Abstracts, Current Contents/Engineering, Technology & Applied Sciences, Current Index to Statistics, Current Papers in Electrical & Electronic Engineering, Current Papers on Computers & Control, Electronics & Communications Abstracts Journal, The Engineering Index, International Aerospace Abstracts, Journal of Current Laser Abstracts, Language and Language Behavior Abstracts, Mathematical Reviews, Science Abstracts (Series A, Physics Abstracts; Series B, Electrical and Electronic Abstracts; and Series C, Computer & Control Abstracts), Science Citation Index, Sociological Abstracts, Social Welfare, Social Planning and Social Development, and Solid State Abstracts Journal.* Reproductions of the Journal by years are available in microform from University Microfilms, 300 N. Zeeb Road, Ann Arbor, Michigan 48106.



**AT&T**  
Bell Laboratories