FEEDBACK COMMUNICATION USING ORTHOGONAL SIGNALS

Thesis by

Stephen Stuart Lavenberg

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1968

(Submitted April 26, 1968)

ACKNOWLEDGMENT

I greatly benefited from the guidance and advice of my research advisor, Dr. T. L. Grettenberg, and from several useful discussions with Mr. S. Farber, a fellow student. My studies were conducted under a National Science Foundation Graduate Fellowship and a National Science Foundation Graduate Traineeship, and I am grateful for the financial assistance they have provided. I wish to thank my wife, whose constant encouragement and good cheer provided great impetus for my work. Finally, the typing of Mrs. Doris Schlicht is appreciated.

ABSTRACT

This research is concerned with block coding for a feedback communication system in which the forward and feedback channels are independently disturbed by additive white Gaussian noise and average power constrained. Two coding schemes are proposed in which the messages to be coded for transmission over the forward channel are realized as a set of orthogonal waveforms. A finite number of forward and feedback transmissions (iterations) per message is made. Information received over the feedback channel is used to modify the waveform transmitted on successive forward iterations in such a way that the expected value of forward signal energy is zero on all iterations after the first. Similarly, information is sent over the feedback channel in such a way that the expected value of feedback signal energy is also zero on all iterations after the first. These schemes are shown to achieve a lower probability of error than the best one-way coding scheme at all rates up to the forward channel capacity, provided only that the feedback channel capacity be greater than the forward channel capacity. These schemes make more efficient use of the available feedback power than existing feedback coding schemes, and therefore require less feedback power to achieve a given error performance.

TABLE OF CONTENTS

I. INTRODUCTION

This research is concerned with block coding for a communication system consisting of a forward and a feedback channel. A block coding scheme for such a system is shown in Figure 1.

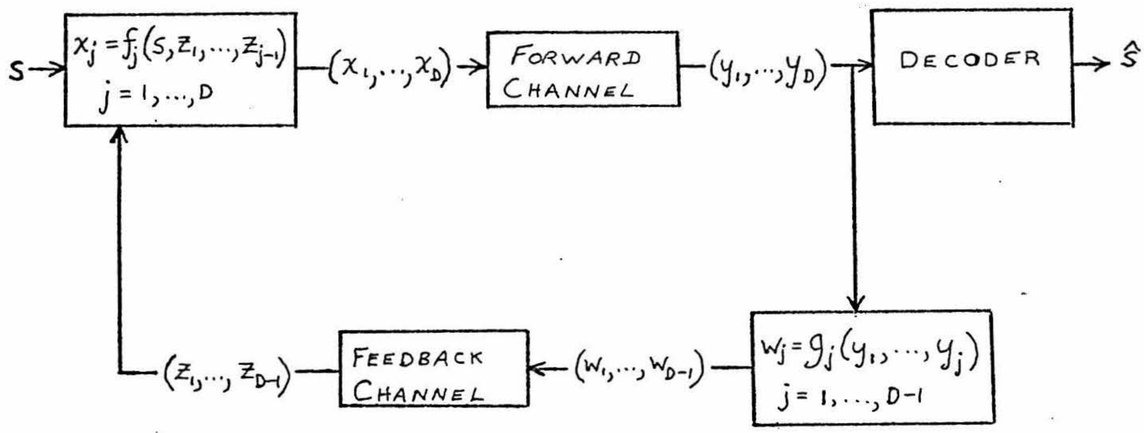

Figure 1. Feedback Coding Scheme.

The message s is one of a set of M equiprobable messages to be coded for transmission over the forward channel. We code each message into a sequence of forward channel inputs as shown. The forward channel inputs are functions of the message being coded and of the previous feedback channel outputs. The feedback channel inputs are functions of the previous forward channel outputs. These functions are fixed for a given feedback coding scheme. The decoder makes a decision, $\hat{s}$, as to the message being coded based on the received sequence of forward channel outputs. An error is made if $\hat{s}$ is not equal to s. We wish to choose the functions $\{f_j : j=1, \cdots D\}$, $\{g_j : j=1, \cdots ,$

D-1}, and the decoder so as to achieve reliable transmission of information over the forward channel.

Feedback coding schemes have been analyzed under a variety of assumptions regarding the forward and feedback channels [1-6]. These schemes achieve a lower probability of error than that attainable without feedback. The presence of the feedback channel does not, however, increase the capacity of the forward channel. This result holds for a wide class of feedback communication systems (see Appendix A).

In the remainder of this paper, we consider a feedback communication system in which the forward and feedback channels are independently disturbed by additive white Gaussian noise and average power constrained. While feedback coding does not increase the maximum rate at which reliable information may be sent over the forward channel of this system, it can improve error performance. If the feedback channel is noiseless, Schalkwijk [2,3], Kailath [2], Omura [4], and Butman [5] have devised schemes in which the probability of error in transmitting information over the forward channel is lower than the minimum probability of error attainable without feedback. This result holds at all rates up to the forward channel capacity. These schemes use scalar signals, that is each message is realized as a point on the real line. The forward channel inputs are linear combinations of the scalar message point being coded and the previous feedback channel outputs. The feedback channel inputs are linear combinations of the previous forward channel outputs. Decoding is accomplished by taking an appropriate linear combination of the forward channel outputs. Then, denoting this combination by α, we choose the message point closest to α as representing the message being coded. In particular, the maximum likelihood decoder may be implemented in this way. (Reference 5 contains the complete formulation of this linear coding scheme.)

Note that α is of the form

$$\alpha = \theta + n$$

where n is a Gaussian random variable with mean zero and variance σ_n^2, and θ is the message point being coded. If the message points are equispaced about the origin, the probability of error, P_e, for this linear coding scheme is

$$P_e = \frac{2}{\sqrt{2\pi}} \int_{\sqrt{\frac{3\gamma}{e^{2RT}-1}}}^{\infty} e^{-x^2/2}\, dx$$

where T is the time to transmit a message, $R = \frac{\ln M}{T}$ is the information rate of the forward channel, and γ is the output signal to noise ratio. ($\gamma = \frac{\sigma_\theta^2}{\sigma_n^2}$, where σ_θ^2 is the variance of the set of equiprobable message points.) Clearly, as T becomes large, it is necessary that γ increase exponentially with time in order to attain a vanishingly small probability of error for any non-zero rate R.

The above expression for the probability of error for a linear coding scheme is also valid in the presence of feedback channel noise. In this case γ may be upper bounded by the sum of the ratio of forward signal energy to forward channel noise and the ratio of feedback signal energy to feedback channel noise. (This bound follows from a result obtained by Elias [8]. An alternate derivation of this bound, which makes use of Butman's matrix formulation of the linear coding scheme [5], is due to Farber [9].) Thus, with the average power constrained in both the forward and feedback channels, γ can increase no faster than linearly with time, and reliable transmission cannot be maintained over the forward channel at non-zero rates. Equivalently, for a finite amount of power in the forward direction, an infinite amount of feedback power is required to maintain any non-zero rate, R, and achieve a zero probability of error. This is a severe limitation of linear coding schemes.

Kramer [6] has recently analyzed a feedback coding scheme in which each message is realized as one of a set of orthogonal waveforms. Information received over the feedback channel is used to modify the waveform transmitted on successive forward iterations in such a way that the expected value of forward signal energy is zero on all iterations after the first. His scheme also achieves a lower probability of error than the best one-way coding scheme at all rates up to the forward channel capacity. However, even in the presence of feedback noise, only a finite amount of feedback power is required to achieve this improved performance. Thus, this scheme is of particular interest. In Chapters II and III of this paper, feedback coding schemes are introduced which further reduce the amount of feedback power required. This is accomplished by sending information over the feedback channel in such a way that the expected value of feedback signal energy is also zero on all iterations after the first. Chapter IV contains a discussion of the performance of these schemes.

II. FEEDBACK CODING SCHEME 1.

2.1. Preliminaries.

The forward and feedback channels are the vector channel equivalents of the time continuous additive white Gaussian noise channel. (Chapter 4 of Reference 11 contains a discussion of the equivalence of the vector and time continuous channel models.) Both channels are assumed to have no bandwidth constraints, and the forward and feedback noises are assumed to be statistically independent. Every T seconds we wish to code and transmit over the forward channel one of M equiprobable messages from the message set

$$\mathcal{S} = \{s_i : i=1, \cdots, M\}$$

Let

$$\Theta = \{\theta_i : i=1, \cdots , M\}$$

be a set of orthogonal M dimensional vectors representing M orthogonal waveforms over a time interval τ with $\|\theta_i\|^2 = E$. Let

$$\Phi = \{\phi_i : i=1, \cdots , M\}$$

be a similar set with $\|\phi_i\|^2 = E'$. We associate each message s_i in $\mathcal{S}$ with a vector θ_i in Θ. Let x_k denote the M dimensional vector input to the forward channel at the k^{th} transmission, n_k the

M dimensional forward channel noise and $y_k = x_k + n_k$ the forward channel received vector. The corresponding quantities for the feedback channel are w_k, m_k and $z_k = w_k + m_k$. (See Figure 2.) In addition, denote by N_o the one-sided power spectral density of the forward noise and by P_{av} the average signal power in the forward direction. Let N_o' and P_{FB} be the corresponding feedback quantities. The components of the forward and feedback noise vectors are independent zero mean Gaussian random variables with variance $\frac{N_o}{2}$ and $\frac{N_o'}{2}$ respectively. The forward channel capacity is $C = \frac{P_{av}}{N_o}$.

$$\begin{tikzcd} & n_k \arrow[d] & \\ x_k \arrow[r] & \oplus \arrow[r] & y_k \\ z_k & \oplus \arrow[l] & w_k \arrow[l] \\ & m_k \arrow[u] & \end{tikzcd}$$

Figure 2. The Forward and Feedback Vector Channels.

We now proceed to the description of feedback coding Scheme 1.

2.2. Description of Scheme 1.

Assume $s \in \mathscr{S}$ is the message to be coded for forward transmission and $\theta \in \Theta$ is the M dimensional vector associated with s. We make N forward transmissions and N-1 feedback transmissions, each of time duration τ, as follows:

$$\begin{tikzcd} & n_1 \arrow[d] & \\ x_1 = \theta \arrow[r] & \oplus \arrow[r] & y_1 = \theta + n_1 \\ z_1 = \hat{\phi}_1 + m_1 & \oplus \arrow[l] & w_1 = \hat{\phi}_1 \arrow[l] \\ & m_1 \arrow[u] & \end{tikzcd}$$

$$\begin{tikzcd} & n_2 \arrow[d] & \\ x_2 = \theta - \theta_1^* \arrow[r] & \oplus \arrow[r] & y_2 = \theta - \theta_1^* + n_2 \end{tikzcd}$$

$$\begin{tikzcd} z_2 = \hat{\phi}_2 - \hat{\phi}_1 + m_2 & \oplus \arrow[l] & w_2 = \hat{\phi}_2 - \hat{\phi}_1 \arrow[l] \\ & m_2 \arrow[u] & \end{tikzcd}$$

$$\vdots$$

$$\begin{tikzcd} & n_k \arrow[d] & \\ x_k = \theta - \theta_{k-1}^* \arrow[r] & \oplus \arrow[r] & y_k = \theta - \theta_{k-1}^* + n_k \end{tikzcd}$$

$$\begin{tikzcd} z_k = \hat{\phi}_k - \hat{\phi}_{k-1} + m_k & \oplus \arrow[l] & w_k = \hat{\phi}_k - \hat{\phi}_{k-1} \arrow[l] \\ & m_k \arrow[u] & \end{tikzcd}$$

$$\vdots$$

$$\begin{tikzcd} & n_N \arrow[d] & \\ x_N = \theta - \theta_{N-1}^* \arrow[r] & \oplus \arrow[r] & y_N = \theta - \theta_{N-1}^* + n_N \end{tikzcd}$$

where $\hat{\phi}_k$ and θ_k^* are determined as follows:

Let

$$\lambda_k = y_1 + \sum_{j=2}^{k} (y_j + \hat{\theta}_{j-1})$$

Then $\hat{\theta}_k$ is that member of Θ which maximizes

$$\langle \lambda_k, \theta_j \rangle \qquad \text{over} \qquad \theta_j \in \Theta$$

and if $\hat{\theta}_k = \theta_\ell$ then $\hat{\phi}_k = \phi_\ell$.

Let

$$\beta_k = z_k + \phi^*_{k-1} \qquad (\phi^*_o = 0)$$

Then ϕ^*_k is that member of Φ which maximizes

$$\langle \beta_k, \phi_j \rangle \qquad \text{over} \qquad \phi_j \in \Phi$$

and if $\phi^*_k = \phi_\ell$ then $\theta^*_k = \theta_\ell$.

Finally, if $\hat{\theta}_N = \theta_\ell$ the receiver decides s_ℓ was the message coded. An error is made if $\hat{\theta}_N \neq \theta$. Note that the total time T to transmit s is

$$T = N\tau \quad .$$

2.3. Analysis of Scheme 1.

We wish to determine the probability of error, $P_N(e)$, the average forward power and the average feedback power for this scheme. In particular we wish to determine the behavior of these quantities as we let $\tau \to \infty$ while the rate of transmission, $R = \frac{\ln M}{N\tau}$, and N are held constant. Bounds on these quantities will be obtained in terms of P_{efb} and P_{ek} defined below:

Let

$$f(x) = 1 - \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\alpha-\sqrt{2x\tau})^2} \left[\int_{-\infty}^{\alpha} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \right]^{M-1} d\alpha$$

We define

$$P_{efb} = f\left(\frac{S'}{N_o'}\right) \quad \text{where} \quad S' = \frac{E'}{\tau}$$

and

$$P_{ek} = f\left(\frac{kS}{N_o}\right) \quad \text{where} \quad S = \frac{E}{\tau}$$

The properties of $f(x)$ are well known and the reader unfamiliar with them should consult Chapter 5 of Reference 11. In particular, $f(x)$ is monotone decreasing in x.

In the previous section we defined vector quantities λ_k and β_k which were used in determining $\hat{\theta}_k$ and ϕ_k^*. If s is the message being coded, then

$$\lambda_k = k\theta + \sum_{j=1}^{k} n_j + \sum_{j=1}^{k-1} (\hat{\theta}_j - \theta_j^*)$$

and if $\hat{\theta}_j = \theta_j^*$ for $j = 1, \cdots, k-1$ it follows that

$$\lambda_k = k\theta + \sum_{j=1}^{k} n_j$$

In this case the probability that $\hat{\theta}_k \neq \theta$ is known to be P_{ek}.

Similarly,

$$\beta_k = \hat{\phi}_k + m_k + \phi^*_{k-1} - \hat{\phi}_{k-1}$$

and if $\phi^*_{k-1} = \hat{\phi}_{k-1}$ then

$$\beta_k = \hat{\phi}_k + m_k \quad .$$

In this case the probability that $\phi^*_k \neq \hat{\phi}_k$ is simply P_{efb}. Hence, P_{ek} is the probability that $\hat{\theta}_k \neq \theta$ given that $\theta^*_j = \hat{\theta}_j$ for $j=1, \cdots, k-1$ and P_{efb} is the probability that $\theta^*_k \neq \hat{\theta}_k$ given that $\theta^*_{k-1} = \hat{\theta}_{k-1}$. In what follows let $\varepsilon(\ )$ be an operator denoting statistical expectation and $P(\ )$ denote the probability of the event in parentheses.

Let

$$P_k(e) = P(\hat{\theta}_k \neq \theta)$$

Then

$$P_k(e) = \frac{1}{M} \sum_{i=1}^{M} P_k(e/s_i)$$

where

$$P_k(e/s_i) = P(\hat{\theta}_k \neq \theta_i/s_i \text{ is being coded})$$

An upper bound to $P_k(e)$ may be derived as follows:

$$P_k(e/s_i) = P_k(e/s_i, \theta_j^* = \hat{\theta}_j \text{ for } j=1,\cdots,k-1)\, P(\theta_j^* = \hat{\theta}_j \text{ for } j=1,\cdots,k-1/s_i) + P_k(e/s_i, \theta_j^* \neq \hat{\theta}_j \text{ for some } j)\, P(\theta_j^* \neq \hat{\theta}_j \text{ for some } j/s_i) = P_{ek}(1-P_{efb})^{k-1} + P_k(e/s_i, \theta_j^* \neq \hat{\theta}_j \text{ for some } j)(1-(1-P_{efb})^{k-1}) \tag{2.1}$$

since

$$P(\theta_j^* = \hat{\theta}_j \; j=1,\cdots,k-1/s_i) = \prod_{j=1}^{k-1} P(\theta_j^* = \hat{\theta}_j/s_i,\; \theta_\ell^* = \hat{\theta}_\ell \; \ell=1,\cdots,j-1)$$

$$= \prod_{j=1}^{k-1} P(\theta_j^* = \hat{\theta}_j/s_i,\; \theta_{j-1}^* = \hat{\theta}_{j-1})$$

$$= (1-P_{efb})^{k-1}$$

Noting that

$$1 - (1-P_{efb})^{k-1} \leq (k-1)\, P_{efb}$$

it follows from (2.1) that

$$P_k(e) \leq P_{ek} + (k-1)\, P_{efb} \tag{2.2}$$

It also follows directly from (2.1) that

$$P_k(e) \geq P_{ek}(1-P_{efb})^{k-1} \tag{2.3}$$

We now obtain bounds on P_{av} where

$$\begin{aligned} P_{av} &= \frac{1}{N\tau} \sum_{k=1}^{N} \varepsilon(\|x_k\|^2) \\ &= \frac{1}{N\tau}\left(E + \sum_{k=2}^{N} \varepsilon(\|x_k\|^2)\right) \end{aligned} \tag{2.4}$$

For $k \geq 2$,

$$\varepsilon(\|x_k\|^2/s_i) = 2EP(\theta^*_{k-1} \neq \theta_i/s_i)$$

If $\theta^*_{k-1} \neq \theta_i$ then either $\theta^*_{k-1} \neq \hat{\theta}_{k-1}$ or $\hat{\theta}_{k-1} \neq \theta_i$. Hence,

$$\begin{aligned} P(\theta^*_{k-1} \neq \theta_i/s_i) &\leq P(\theta^*_{k-1} \neq \hat{\theta}_{k-1} \text{ or } \hat{\theta}_{k-1} \neq \theta_i/s_i) \\ &\leq P(\theta^*_{k-1} \neq \hat{\theta}_{k-1}/s_i) + P(\hat{\theta}_{k-1} \neq \theta_i/s_i) \end{aligned}$$

Therefore,

$$\varepsilon(\|x_k\|^2) \leq 2E(P_{k-1}(e) + P(\theta^*_{k-1} \neq \hat{\theta}_{k-1})) \tag{2.5}$$

where

$$P(\theta^*_{k-1} \neq \hat{\theta}_{k-1}) = P(\theta^*_{k-1} \neq \hat{\theta}_{k-1}/\theta^*_{k-2} \neq \hat{\theta}_{k-2})\ P(\theta^*_{k-2} \neq \hat{\theta}_{k-2})$$

$$+ P(\theta^*_{k-1} \neq \hat{\theta}_{k-1}/\theta^*_{k-2} = \hat{\theta}_{k-2})\ P(\theta^*_{k-2} = \hat{\theta}_{k-2})$$

$$\leq P(\theta^*_{k-2} \neq \hat{\theta}_{k-2}) + P_{efb}$$

and by induction, noting that $P(\theta^*_1 \neq \hat{\theta}_1) = P_{efb}$, we have

$$P(\theta^*_{k-1} \neq \hat{\theta}_{k-1}) \leq (k-1)\ P_{efb} \tag{2.6}$$

It follows from (2.2), (2.4), (2.5) and (2.6) that

$$\frac{E}{N\tau} \leq P_{av} \leq \frac{E}{N\tau}\left(1+2\sum_{k=2}^{N}\left(P_{ek-1} + (k-2)\ P_{efb} + (k-1)\ P_{efb}\right)\right)$$

Hence,

$$\frac{S}{N} \leq P_{av} \leq \frac{S}{N}\left(1 + 2(N-1)^2 P_{efb} + 2\sum_{k=1}^{N-1} P_{ek}\right) \tag{2.7}$$

We now bound P_{FB} where

$$P_{FB} = \frac{1}{N\tau}\sum_{k=1}^{N-1} \varepsilon(\|w_k\|^2)$$

$$= \frac{1}{N\tau}\left(E' + \sum_{k=2}^{N-1} \varepsilon\ (\|w_k\|^2)\right) \tag{2.8}$$

For $k \geq 2$,

$$\varepsilon(\|w_k\|^2/s_i) = 2E'P(\hat{\theta}_k \neq \hat{\theta}_{k-1}/s_i)$$

Now if $\hat{\theta}_k \neq \hat{\theta}_{k-1}$ then either $\hat{\theta}_k \neq \theta_i$ or $\hat{\theta}_{k-1} \neq \theta_i$. It follows from this that

$$P(\hat{\theta}_k \neq \hat{\theta}_{k-1}/s_i) \leq P(\hat{\theta}_k \neq \theta_i/s_i) + P(\hat{\theta}_{k-1} \neq \theta_i/s_i)$$

Therefore,

$$\varepsilon(\|w_k\|^2) \leq 2E'(P_k(e) + P_{k-1}(e)) \tag{2.9}$$

It follows from (2.2), (2.8) and (2.9) that

$$\frac{E'}{N\tau} \leq P_{FB} \leq \frac{E'}{N\tau}\left(1 + 2\sum_{k=2}^{N-1}\left(P_{ek} + (k-1)\,P_{efb} + P_{ek-1} + (k-2)\,P_{efb}\right)\right)$$

Finally, noting that $P_{ek} \leq P_{ek-1}$ we obtain

$$\frac{S'}{N} \leq P_{FB} \leq \frac{S'}{N}\left(1 + 2(N-2)^2\,P_{efb} + 4\sum_{k=1}^{N-2} P_{ek}\right) \tag{2.10}$$

We now wish to determine the asymptotic behavior of this scheme for large τ. It follows from the properties of the function $f(x)$ that

$$P_{ek} < P_{e1} \qquad \text{for} \qquad k > 1$$

$$P_{efb} \leq P_{e1} \qquad \text{for} \qquad \frac{S'}{N_o'} \geq \frac{S}{N_o}$$

and

$$P_{e1} \to 0 \quad \text{as} \quad \tau \to \infty \quad \text{for} \quad 0 \le \frac{\ln M}{\tau} < \frac{S}{N_o}$$

Therefore, for

$$\frac{S'}{N_o'} \ge \frac{S}{N_o} \qquad \text{and} \qquad 0 \le R = \frac{\ln M}{N\tau} < \frac{S}{NN_o}$$

we have

$$P_{av} \to \frac{S}{N} \qquad \text{as} \quad \tau \to \infty \tag{2.11}$$

$$P_{FB} \to \frac{S'}{N} \qquad \text{as} \quad \tau \to \infty \tag{2.12}$$

To observe the asymptotic behavior of the probability of error we examine the behavior of the channel reliability function, $E(R)$, where

$$E(R) = \lim_{\tau\to\infty} \left(- \frac{1}{N\tau} \ln P_N(e) \right)$$

From Eq.(2.2) we have that

$$P_N(e) \le P_{eN} + (N-1)\, P_{efb}$$

Let us now consider the following two cases:

A) $\frac{S'}{N_o'} = r \frac{S}{N_o}$ for $1 \le r < N$ (r need not be an integer)

In this case

$$P_{efb} = P_{er} \qquad \text{and} \qquad P_{eN} < P_{er}$$

Therefore, $\qquad P_N(e) \le N\, P_{er}$

and

$$E(R) \ge \lim_{\tau\to\infty}\left(-\frac{1}{N\tau}\,\ln P_{er}\right) + \lim_{\tau\to\infty}\left(-\frac{1}{N\tau}\,\ln N\right)$$

$$= \frac{r}{N}\lim_{\tau\to\infty}\left(-\frac{1}{r\tau}\,\ln P_{er}\right)$$

Making use of the asymptotic expression for P_{er} we have

$$E(R) \ge \frac{r}{N}\begin{cases} \dfrac{S}{2N_o} - \dfrac{\ln M}{r\tau} & 0 \le \dfrac{\ln M}{r\tau} \le \dfrac{S}{4N_o} \\[2ex] \left(\sqrt{\dfrac{S}{N_o}} - \sqrt{\dfrac{\ln M}{r\tau}}\right)^2 & \dfrac{S}{4N_o} \le \dfrac{\ln M}{r\tau} < \dfrac{S}{N_o} \end{cases}$$

We now require that $R < \frac{S}{NN_o}$. Then using (2.11) and (2.12) the above result may be rewritten as:

If

$$\frac{P_{FB}}{N_o'} = r\,\frac{P_{av}}{N_o} \qquad 1 \le r < N \tag{2.13}$$

then

$$E(R) \geq \begin{cases} \dfrac{rP_{av}}{2N_o} - R & 0 \leq R \leq \min\left(\dfrac{rP_{av}}{4N_o}, \dfrac{P_{av}}{N_o}\right) \\ \left(\sqrt{\dfrac{rP_{av}}{N_o}} - \sqrt{R}\right)^2 & \min\left(\dfrac{rP_{av}}{4N_o}, \dfrac{P_{av}}{N_o}\right) \leq R < \dfrac{P_{av}}{N_o} \end{cases} \tag{2.14}$$

B) $\dfrac{S'}{N_o'} = r\dfrac{S}{N_o}$ $\quad r \geq N$ $\qquad$ (r need not be an integer)

In this case

$$P_{efb} \leq P_{eN}$$

so that

$$P_N(e) \leq NP_{eN}$$

and

$$E(R) \geq \lim_{\tau\to\infty}\left(-\frac{1}{N\tau}\ln P_{eN}\right) + \lim_{\tau\to\infty}\left(-\frac{1}{N\tau}\ln N\right)$$

$$= \lim_{\tau\to\infty}\left(-\frac{1}{N\tau}\ln P_{eN}\right)$$

$$= \begin{cases} \dfrac{S}{2N_o} - R & 0 \leq R \leq \dfrac{S}{4N_o} \\ \left(\sqrt{\dfrac{S}{N_o}} - \sqrt{R}\right)^2 & \dfrac{S}{4N_o} \leq R < \dfrac{S}{N_o} \end{cases}$$

In addition, from (2.3) we have that

$$P_N(e) \geq P_{eN}(1-P_{efb})^{N-1}$$

Therefore,

$$E(R) \leq \lim_{\tau\to\infty}\left(-\frac{1}{N\tau}\ln P_{eN}\right)+\lim_{\tau\to\infty}\left(-\frac{N-1}{N\tau}\ln(1-P_{efb})\right)$$

$$= \lim_{\tau\to\infty}\left(-\frac{1}{N\tau}\ln P_{eN}\right)$$

provided that $P_{efb} \to 0$ as $\tau \to \infty$. This will be true as long as $\frac{\ln M}{\tau} < \frac{S'}{N'_o}$.

Finally, if we require that $R < \frac{S}{NN_o}$, then using (2.11) and (2.12) the above results may be rewritten as:

If

$$\frac{P_{FB}}{N'_o} = r\,\frac{P_{av}}{N_o} \qquad\qquad r \geq N \tag{2.15}$$

then

$$E(R) = \begin{cases} \dfrac{NP_{av}}{2N_o} - R & 0 \leq R \leq \min\left(\dfrac{NP_{av}}{4N_o}, \dfrac{P_{av}}{N_o}\right) \\ \left(\sqrt{\dfrac{NP_{av}}{N_o}} - \sqrt{R}\right)^2 & \min\left(\dfrac{NP_{av}}{4N_o}, \dfrac{P_{av}}{N_o}\right) \leq R < \dfrac{P_{av}}{N_o} \end{cases} \tag{2.16}$$

The upper bound to $E(R)$ obtained here from (2.3) also applies to the case $1 \le r < N$. However, in that case the upper and lower bounds no longer coincide. A more exact analysis is required to obtain the true value of $E(R)$ in that case. In the following chapter a feedback coding scheme is introduced for which an exact expression for the channel reliability function is obtained.

A discussion of the performance of these schemes is postponed until Chapter IV.

III. FEEDBACK CODING SCHEME 2

3.1. Preliminaries.

Section 2.1 of Chapter II applies verbatim to Scheme 2. In addition, quantities defined in Chapter II are not redefined in this Chapter unless their meanings have changed. The feedback coding scheme introduced here permits an exact analysis of the channel reliability function. In addition, a simple modification of this scheme is considered and its effect on peak power is discussed.

We now proceed to the description of feedback coding Scheme 2.

3.2. Description of Scheme 2.

Assume $s \in \mathscr{S}$ is the message to be coded for forward transmission and $\theta \in \Theta$ is the M dimensional vector associated with s. We make N forward transmissions and only one feedback transmission, each of time duration τ, as follows:

$$\begin{tikzcd}
 & n_1 \arrow[d] & \\
x_1 = \theta \arrow[r] & \oplus \arrow[r] & y_1 = \theta + n_1 \\
z_1 = \hat{\phi}_1 + m_1 & \oplus \arrow[l] & w_1 = \hat{\phi}_1 \arrow[l] \\
 & m_1 \arrow[u] &
\end{tikzcd}$$

$$\begin{tikzcd}
 & n_2 \arrow[d] & \\
x_2 = \theta - \theta_1^* \arrow[r] & \oplus \arrow[r] \arrow[d, phantom, "\vdots"] & y_2 = \theta - \theta_1^* + n_2 \\
 & {} &
\end{tikzcd}$$

$$x_k = \theta - \theta_1^* \longrightarrow \oplus \longrightarrow y_k = \theta - \theta_1^* + n_k \quad (n_k \downarrow \oplus)$$

$$\vdots$$

$$x_N = \theta - \theta_1^* \longrightarrow \oplus \longrightarrow y_N = \theta - \theta_1^* + n_N \quad (n_N \downarrow \oplus)$$

In this scheme we let

$$\lambda_N = y_1 + \sum_{k=2}^{N} (y_k + \hat{\theta}_1) .$$

$\hat{\theta}_N$ is then determined as in Scheme 1. $\hat{\theta}_1$, $\hat{\phi}_1$, ϕ_1^* and θ_1^* are the same as in Scheme 1.

3.3. Analysis of Scheme 2.

We first obtain exact expressions for P_{av} and P_{FB}.

$$P_{av} = \frac{1}{N\tau} \sum_{k=1}^{N} \varepsilon(\|x_k\|^2)$$

$$= \frac{1}{N\tau}\left(E + \sum_{k=2}^{N} \varepsilon(\|x_k\|^2)\right)$$

For $k \geq 2$,

$$\varepsilon(\|x_k\|^2/s_i) = 2EP(\theta_1^* \neq \theta_i/s_i)$$

where

$$\begin{aligned}
P(\theta_1^* \neq \theta_i/s_i) &= P(\theta_1^* \neq \theta_i/s_i, \theta_1^* = \hat{\theta}_1)P(\theta_1^* = \hat{\theta}_1/s_i) \\
&+ P(\theta_1^* \neq \theta_i/s_i, \theta_1^* \neq \hat{\theta}_1)P(\theta_1^* \neq \hat{\theta}_1/s_i) \\
&= P(\hat{\theta}_1 \neq \theta_i/s_i, \theta_1^* = \hat{\theta}_1)P(\theta_1^* = \hat{\theta}_1/s_i) \\
&+ P(\theta_1^* \neq \theta_i/s_i, \theta_1^* \neq \hat{\theta}_1)P(\theta_1^* \neq \hat{\theta}_1/s_i) \\
&= P_{e1}(1-P_{efb}) + P(\theta_1^* \neq \theta_i/s_i, \theta_1^* \neq \hat{\theta}_1)\, P_{efb}
\end{aligned}$$

We wish to determine $P(\theta_1^* \neq \theta_i/s_i, \theta_1^* \neq \hat{\theta}_1)$ in terms of P_{e1} and P_{efb}. To do this we write

$$\begin{aligned}
P(\theta_1^* = \theta_i/s_i, \theta_1^* \neq \hat{\theta}_1) &= P(\theta_1^* = \theta_i/s_i, \theta_1^* \neq \hat{\theta}_1, \hat{\theta}_1 = \theta_i)P(\hat{\theta}_1 = \theta_i/s_i, \theta_1^* \neq \hat{\theta}_1) \\
&+ P(\theta_1^* = \theta_i/s_i, \theta_1^* \neq \hat{\theta}_1, \hat{\theta}_1 \neq \theta_i)P(\hat{\theta}_1 \neq \theta_i/s_i, \theta_1^* \neq \hat{\theta}_1) \\
&= 0 + \left(\frac{1}{M-1}\right) P_{e1}
\end{aligned}$$

Therefore,

$$\varepsilon(\|x_k\|^2) = 2E\left(P_{e1}(1-P_{efb}) + \left(1 - \frac{P_{e1}}{M-1}\right) P_{efb}\right)$$

and

$$P_{av} = \frac{S}{N}\left(1 + 2(N-1)\left(P_{e1} + P_{efb} - \left(\frac{M}{M-1}\right) P_{e1}P_{efb}\right)\right) \quad (3.1)$$

Also

$$P_{FB} = \frac{1}{N\tau}\, \varepsilon(\|w_1\|^2)$$

Therefore,

$$P_{FB} = \frac{S'}{N} \tag{3.2}$$

We now consider $P_N(e)$, the probability of error, for this scheme. Note that when s is the message being coded,

$$\lambda_N = N\theta + (N-1)(\hat{\theta}_1 - \theta_1^*) + \sum_{k=1}^{N} n_k$$

The probability of error therefore depends on the values of $\hat{\theta}_1$ and θ_1^*. The following is a table of the possible events associated with these values. It is assumed that s_i is the message being coded.

Event	$\hat{\theta}_1$	θ_1^*	λ_N
A	θ_i	θ_i	$\sum_{k=1}^{N} n_k + N\theta_i$
B	θ_i	θ_j $j\neq i$	$\sum_{k=1}^{N} n_k + (2N-1)\theta_i - (N-1)\theta_j$
C	θ_j $j\neq i$	θ_j	$\sum_{k=1}^{N} n_k + N\theta_i$
D	θ_j $j\neq i$	θ_ℓ $\ell\neq i,j$	$\sum_{k=1}^{N} n_k + N\theta_i + (N-1)\theta_j - (N-1)\theta_\ell$

E	θ_j $j \neq i$	θ_i	$\sum_{k=1}^{N} n_k + \theta_i + (N-1)\theta_j$

These events are disjoint and we may write

$$P_N(e/s_i) = P(\hat{\theta}_N \neq \theta_i/s_i, A)P(A/s_i) + P(\hat{\theta}_N \neq \theta_i/s_i, B)P(B/s_i)$$

$$+ P(\hat{\theta}_N \neq \theta_i/s_i, C)P(C/s_i) + P(\hat{\theta}_N \neq \theta_i/s_i, D)P(D/s_i)$$

$$+ P(\hat{\theta}_N \neq \theta_i/s_i, E)P(E/s_i) \tag{3.3}$$

We wish to consider the terms which make up this sum.

$$P(\theta_N \neq \theta_i/s_i, D) = P(\langle \lambda_N, \theta_r \rangle \geq \langle \lambda_N, \theta_i \rangle \text{ for some } r \neq i/s_i, D)$$

$$\geq P(\langle \lambda_N, \theta_j \rangle \geq \langle \lambda_N, \theta_i \rangle / s_i, D)$$

We now substitute the value of λ_N corresponding to event D in the above expression. When we do this we can condition the probability only on the relationship

$$\langle n_1, \theta_j \rangle \geq \langle n_1, \theta_i \rangle + E$$

implied by event D. Therefore,

$$P(\langle \lambda_N, \theta_j \rangle \geq \langle \lambda_N, \theta_i \rangle / s_i, D) = P\left(\sum_{k=1}^{N} \langle n_k, \theta_j \rangle + (N-1)\, E \geq \sum_{k=1}^{N} \langle n_k, \theta_i \rangle \right.$$

$$\left. + NE / \langle n_1, \theta_j \rangle \geq \langle n_1, \theta_i \rangle + E \right)$$

$$\geq P\left(\sum_{k=2}^{N} \langle n_k, \theta_j \rangle \geq \sum_{k=2}^{N} \langle n_k, \theta_i \rangle \right)$$

where $\sum_{k=2}^{N} \langle n_k, \theta_j \rangle$ and $\sum_{k=2}^{N} \langle n_k, \theta_i \rangle$ are independent identically distributed Gaussian random variables.

It therefore follows that

$$1 \geq P(\hat{\theta}_N \neq \theta_i/s_i, D) \geq \frac{1}{2} \tag{3.4}$$

A similar argument shows that

$$1 \geq P(\hat{\theta}_N \neq \theta_i/s_i, E) \geq \frac{1}{2} \tag{3.5}$$

The conditional probabilities of events D and E are

$$P(D/s_i) = P_{e1} P_{efb} \left(\frac{M-2}{M-1} \right) \tag{3.6}$$

and

$$P(E/s_i) = \frac{P_{e1} P_{efb}}{M-1} \tag{3.7}$$

Now

$$P(\theta_N \neq \theta_i/s_i, B) = 1 - P(\langle \lambda_N, \theta_i \rangle > \langle \lambda_N, \theta_r \rangle \quad \text{for all} \quad r \neq i/s_i, B)$$

$$= 1 - P\left(\sum_{k=1}^{N} \langle n_k, \theta_i \rangle + (2N-1)\, E > \sum_{k=1}^{N} \langle n_k, \theta_j \rangle - (N-1)\, E, \sum_{k=1}^{N} \langle n_k, \theta_r \rangle \right.$$

$$\text{for all} \quad r \neq i,j/\langle n_1, \theta_i\rangle + E > \langle n_1, \theta_r\rangle \quad \text{for all} \quad r \neq i \Big)$$

$$\leq 1 - P\left(\sum_{k=1}^{N} \langle n_k, \theta_i \rangle + NE > \sum_{k=1}^{N} \langle n_k, \theta_r\rangle \quad \text{for all} \quad r \neq i/\langle n_1, \theta_i\rangle \right.$$

$$\left. + E > \langle n_1, \theta_r\rangle \text{ for all } r \neq i \right)$$

$$= 1 - P(\langle \lambda_N, \theta_i\rangle > \langle \lambda_N, \theta_r\rangle \quad \text{for all} \quad r \neq i/s_i, A)$$

$$= P(\hat{\theta}_N \neq \theta_i/s_i, A)$$

The conditional probabilities of events A and B are

$$P(A/s_i) = (1-P_{e1})(1-P_{efb})$$

and

$$P(B/s_i) = (1-P_{e1})\ P_{efb}$$

If

$$0 \leq \frac{\ln M}{\tau} < \frac{S'}{N'_o}$$

then

$$P_{efb} \to 0 \qquad \text{as} \qquad \tau \to \infty$$

and for large τ

$$P_{efb} < 1 - P_{efb}$$

In this case

$$0 \le P(\hat{\theta}_N \ne \theta_i/s_i,B)\ P(B/s_i) \le P(\hat{\theta}_N \ne \theta_i/s_i,A)\ P(A/s_i) \tag{3.8}$$

Finally,

$$P(\hat{\theta}_N \ne \theta_i/s_i,A)\ P(A/s_i) + P(\hat{\theta}_N \ne \theta_i/s_i,C)\ P(C/s_i)$$

$$= P(\hat{\theta}_N \ne \theta_i/s_i, \theta_1^* = \hat{\theta}_1)\ P(\theta_1^* = \hat{\theta}_1/s_i)$$

$$= P_{eN}(1-P_{efb}) \tag{3.9}$$

It follows from (3.3) - (3.9) that

$$\frac{P_{e1}P_{efb}}{2} + P_{eN}(1-P_{efb}) \le P_N(e) \le P_{e1}P_{efb} + 2P_{eN}(1-P_{efb}) \tag{3.10}$$

where we have assumed τ is large and $0 \le \frac{\ell nM}{\tau} < \frac{S'}{N_o'}$ to obtain the upper bound.

Suppose now that

$$\frac{S'}{N_o'} = r\,\frac{S}{N_o} \quad \text{for} \quad r \ge 1 \quad (r \text{ need not be an integer})$$

This implies

$$P_{efb} < P_{el}$$

If, in addition, we require that

$$0 \le R < \frac{S}{NN_o}$$

then

$$P_{el} \to 0 \qquad \text{as} \qquad \tau \to \infty$$

$$P_{efb} \to 0 \qquad \text{as} \qquad \tau \to \infty$$

and from (3.1)

$$P_{av} \to \frac{S}{N} \qquad \text{as} \qquad \tau \to \infty \tag{3.11}$$

Equation (3.10) is valid with $P_{efb} = P_{er}$ and we have for the channel reliability function

$$E(R) = \lim_{\tau \to \infty} \left(- \frac{1}{N\tau} \ln P_N(e) \right)$$

$$= \min(E_1(R), E_2(R))$$

where

$$E_1(R) = \lim_{T\to\infty} \left(- \frac{1}{NT} \ln P_{eN} \right)$$

and

$$E_2(R) = \lim_{T\to\infty} \left(- \frac{1}{NT} \ln P_{el} \right) + \lim_{T\to\infty} \left(- \frac{1}{NT} \ln P_{er} \right)$$

Now, using (3.2), (3.11) and the asymptotic expressions for P_{el}, P_{er} and P_{eN}, the above results may be written as:

If

$$\frac{P_{FB}}{N_o'} = r \frac{P_{av}}{N_o} \qquad r \geq 1 \tag{3.12}$$

then

$$E(R) = \min(E_1(R), E_2(R)) \tag{3.13a}$$

where

$$E_1(R) = \begin{cases} \dfrac{NP_{av}}{2N_o} - R & 0 \leq R \leq \min\left(\dfrac{NP_{av}}{4N_o}, \dfrac{P_{av}}{N_o}\right) \\ \left(\sqrt{\dfrac{NP_{av}}{N_o}} - \sqrt{R}\right)^2 & \min\left(\dfrac{NP_{av}}{4N_o}, \dfrac{P_{av}}{N_o}\right) \leq R < \dfrac{P_{av}}{N_o} \end{cases} \tag{3.13b}$$

and

$$E_2(R) = \begin{cases} \dfrac{(r+1)P_{av}}{2N_o} - 2R & 0 \le R \le \dfrac{P_{av}}{4N_o} \\[2ex] \dfrac{(r+2)P_{av}}{2N_o} - 2\sqrt{\dfrac{R\ P_{av}}{N_o}} & \dfrac{P_{av}}{4N_o} \le R \le \min\left(\dfrac{rP_{av}}{4N_o}, \dfrac{P_{av}}{N_o}\right) \\[2ex] \dfrac{(r+1)P_{av}}{N_o} - 2\sqrt{\dfrac{R\ P_{av}}{N_o}}\,(1+\sqrt{r}) + 2R & \min\left(\dfrac{rP_{av}}{4N_o}, \dfrac{P_{av}}{N_o}\right) \le R < \dfrac{P_{av}}{N_o} \end{cases} \tag{3.13c}$$

These equations completely describe the behavior of $E(R)$ for this scheme. If $N \le r$ or $N \ge r + 1$ they can be simplified as follows:

Clearly

$$E_2(R) \ge \begin{cases} \dfrac{rP_{av}}{2N_o} - R & 0 \le R \le \min\left(\dfrac{rP_{av}}{4N_o}, \dfrac{P_{av}}{N_o}\right) \\[2ex] \left(\sqrt{\dfrac{rP_{av}}{N_o}} - \sqrt{R}\right)^2 & \min\left(\dfrac{rP_{av}}{4N_o}, \dfrac{P_{av}}{N_o}\right) \le R < \dfrac{P_{av}}{N_o} \end{cases}$$

so that

$$E(R) = E_1(R) \qquad \text{for} \qquad N \le r$$

as in Scheme 1.

It is also possible to show that

$$E_2(R) \le \begin{cases} \dfrac{(r+1)P_{av}}{2N_o} - R & 0 \le R \le \min\left(\dfrac{(r+1)P_{av}}{4N_o}, \dfrac{P_{av}}{N_o}\right) \\ \left(\sqrt{\dfrac{(r+1)P_{av}}{N_o}} - \sqrt{R}\right)^2 & \min\left(\dfrac{(r+1)P_{av}}{4N_o}, \dfrac{P_{av}}{N_o}\right) \le R < \dfrac{P_{av}}{N_o} \end{cases}$$

Therefore,

$$E(R) = E_2(R) \qquad \text{for} \qquad N \ge r + 1$$

Note that $E_2(R)$ is independent of N. Hence, $E(R)$ cannot be increased by further increasing N, for $N \ge r + 1$.

Before concluding this section it is of interest to consider the performance of this coding scheme when $\frac{P_{FB}}{N'_o} < \frac{P_{av}}{N_o}$. We show that in this case reliable transmission of information over the forward channel is not possible at all rates up to the forward channel capacity.

To see this, note that if

$$\frac{P_{FB}}{N'_o} \le R < \frac{P_{av}}{N_o}$$

then

$$\frac{\ln M}{\tau} \ge \frac{S'}{N'_o} \quad \text{so that} \quad P_{efb} \to 1 \quad \text{as} \quad \tau \to \infty$$

It then follows from (3.1) that

$$P_{av} \rightarrow \frac{S}{N}(2N-1) \quad \text{as} \quad \tau \rightarrow \infty$$

and from (3.10) that

$$P_N(e) \geq \frac{P_{e1}P_{efb}}{2} \rightarrow \frac{1}{2} \quad \text{as} \quad \tau \rightarrow \infty$$

for all rates R such that

$$\max\left(\frac{P_{FB}}{N_o'}, \frac{P_{av}}{(2N-1)N_o}\right) \leq R < \frac{P_{av}}{N_o}$$

Reliable transmission cannot be maintained for this range of rates.

3.4. Peak Power.

We define peak power for the feedback coding scheme as the maximum average power over any transmission interval τ. Let P_{PK} denote the peak power in the forward direction and P_{PK}' denote the peak power in the feedback direction. The peak power in the forward direction is the maximum value of

$$\frac{\|x_k\|^2}{\tau} \qquad k = 1, \cdots, N$$

If $\theta_1^* \neq \theta$ then

$$\frac{\|x_k\|^2}{\tau} = 2S \qquad k = 2, \cdots, N$$

Therefore,

$$P_{PK} = 2S$$

The peak power in the feedback direction is

$$P'_{PK} = \frac{\|w_1\|^2}{\tau} = S'$$

Hence, from (3.2)

$$\frac{P'_{PK}}{P_{FB}} = N \tag{3.14}$$

For the remainder of this section we assume that

$$\frac{S'}{N'_o} = r\,\frac{S}{N_o} \qquad\qquad r \geq 1$$

and

$$0 \leq R < \frac{S}{NN_o}$$

Then from (3.11)

$$\frac{P_{PK}}{P_{av}} \to 2N \quad \text{as} \quad \tau \to \infty$$

Thus the ratio of peak power to average power increases with N, the number of transmissions per message. Any physical communication system operates with a peak power limitation. Hence, the number of transmissions is limited.

Suppose now we modify feedback coding Scheme 2 by letting

$$x_k = g(\theta - \theta_1^*) \qquad k = 2, \cdots, N$$

where g is a fixed positive gain constant. For this scheme we let

$$\lambda_N = y_1 + \sum_{k=2}^{N} (y_k + g\hat{\theta}_1)$$

to determine $\hat{\theta}_N$. Note that

$$P_{av} \to \frac{S}{N} \qquad \text{as} \qquad \tau \to \infty$$

independent of the choice of g. Now, however,

$$\frac{P_{PK}}{P_{av}} = \max(2g^2 N, N) \tag{3.15}$$

Since,

$$\frac{P_{PK}}{P_{av}} \geq \frac{P'_{PK}}{P_{FB}}$$

we assume that only the constraint on the forward peak to average power ratio is critical. We wish to determine the channel reliability function for this scheme. Proceeding as in the previous section it can be shown that, If

$$\frac{P_{FB}}{N_o'} = \frac{rP_{av}}{N_o} \qquad r \geq 1$$

then

$$E(R) = \min(E_1(R), E_2(R))$$

where

$$E_1(R) = \begin{cases} \dfrac{P_{av}}{2N_o} \dfrac{((N-1)g+1)^2}{N} - R & 0 \leq R \leq \min\left(\dfrac{((N-1)g+1)^2}{N} \dfrac{P_{av}}{4N_o}, \dfrac{P_{av}}{N_o}\right) \\ \left(\sqrt{\dfrac{P_{av}}{N_o} \dfrac{((N-1)g+1)^2}{N}} - \sqrt{R}\right)^2 & \min\left(\dfrac{((N-1)g+1)^2}{N} \dfrac{P_{av}}{4N_o}, \dfrac{P_{av}}{N_o}\right) \leq R < \dfrac{P_{av}}{N_o} \end{cases}$$

and $E_2(R)$ is again given by (3.13c) and is independent of N and g.

$E_1(R)$ can be increased by increasing either N or g. However, if we fix the forward peak to average power ratio (see (3.15)), it can be shown that $E_1(R)$ increases with decreasing g (for $g \geq \sqrt{2}/2$). Therefore, it is reasonable to choose $g = 1$ as in feedback coding Scheme 2, and increase N in order to obtain improved error performance. $E_1(R)$ is in fact maximized, for fixed $\frac{P_{PK}}{P_{av}}$, by choosing $g = \sqrt{2}/2$ and $N = \frac{P_{PK}}{P_{av}}$.

IV. PERFORMANCE OF THE CODING SCHEMES

4.1. The Channel Reliability Function.

In Chapters II and III of this paper we analyzed two block coding schemes for a feedback communication system in which the forward and feedback channels are disturbed by independent additive white Gaussian noise and average power constrained. In particular, we focused our attention on the behavior of the channel reliability function, $E(R)$, for these schemes. (See Equations (2.13)-(2.16), (3.12) and (3.13).) This function is of particular interest since for large coding delay (time to transmit a message) T, the probability of error is given by

$$P_N(e) \simeq \exp(-E(R)T)$$

and $E(R)$ can be used to compare the performance of different coding schemes.

4.2. Comparison of Coding Schemes.

The channel reliability functions for the feedback coding schemes can be compared with the optimum reliability function attainable if the feedback channel were not available. Denoting this optimum one-way channel reliability function by $E'(R)$, we have (see Chapter 5 of Reference 11)

$$E'(R) = \begin{cases} \frac{P_{av}}{2N_o} - R & 0 \le R \le \frac{P_{av}}{4N_o} \\ \left(\sqrt{\frac{P_{av}}{N_o}} - \sqrt{R}\right)^2 & \frac{P_{av}}{4N_o} \le R < \frac{P_{av}}{N_o} \end{cases} \tag{4.1}$$

It is well known that signals orthogonal over the time interval T attain this performance. Note that for both feedback coding schemes, $E(R) > E'(R)$ at all rates R up to the forward channel capacity, $C = \frac{P_{av}}{N_o}$, provided only that the feedback channel capacity be greater than the forward channel capacity. Hence, for the same values of P_{av}, N_o, T, and $R < \frac{P_{av}}{N_o}$, the probability of error, $P_N(e)$, for the feedback schemes is less than the probability of error, $P'(e) \simeq \exp(-E'(R)T)$, for the best one-way scheme.

As T becomes arbitrarily large for these schemes, so does the number of dimensions per second used in coding, or equivalently, so does the bandwidth used [11]. In many practical systems we may be restricted to a given large but finite time-bandwidth product, or equivalently, to a given large but finite number of dimensions. It is interesting to compare the one-way and feedback schemes for the same values of P_{av}, N_o, $R < \frac{P_{av}}{N_o}$, and the same large but finite number of dimensions D. Letting M' denote the number of messages and T' the coding delay for the one-way orthogonal scheme, we have

$$R = \frac{\ln M'}{T'}$$

and

$$D = M' \quad .$$

Using (4.1), it then follows that

$$P'(e) \simeq \exp(-E'(R)T') = \begin{cases} \exp\left(-\left(\frac{P_{av}}{2N_oR} - 1\right)\ln D\right) & 0 \le R \le \frac{P_{av}}{4N_o} \\ \\ \exp\left(-\left(\sqrt{\frac{P_{av}}{N_oR}} - 1\right)^2 \ln D\right) & \frac{P_{av}}{4N_o} \le R < \frac{P_{av}}{N_o} \end{cases} \tag{4.2}$$

Letting M denote the number of messages, T the coding delay, and N the number of forward transmissions for the feedback coding schemes, we have

$$R = \frac{\ln M}{T}$$

and

$$D = MN \; .$$

Assuming $N \le r$, where $\frac{P_{FB}}{N_o'} = r\,\frac{P_{av}}{N_o}$, it follows using (2.16) or (3.13) that

$$P_N(e) \simeq \exp(-E(R)T) = \begin{cases} \exp\left(-\left(\dfrac{NP_{av}}{2N_oR} - 1\right)\ln\dfrac{D}{N}\right) & 0 \le R \le \min\left(\dfrac{NP_{av}}{4N_o}, \dfrac{P_{av}}{N_o}\right) \\ \exp\left(-\left(\sqrt{\dfrac{NP_{av}}{N_oR}} - 1\right)^2 \ln\dfrac{D}{N}\right) & \min\left(\dfrac{NP_{av}}{4N_o}, \dfrac{P_{av}}{N_o}\right) \le R < \dfrac{P_{av}}{N_o} \end{cases} \quad (4.3)$$

Using (4.2) and (4.3) it can be shown that $P_N(e) < P'(e)$ for the same P_{av}, N_o, $R < \frac{P_{av}}{N_o}$, and number of dimensions D which is assumed to be large.

We now discuss the relation of Schemes 1 and 2 to existing feedback coding schemes. In Chapter I we mentioned several existing coding schemes for the particular feedback communication system we have considered. It is worth repeating that in the presence of feedback noise the schemes of Schalkwijk [2,3], Kailath [2], Omura [4], and Butman [5] require an infinite amount of feedback power to maintain any non-zero rate and achieve a zero probability of error. Hence, the presence of feedback noise poses a severe limitation on the performance of these schemes. The scheme considered by Kramer [6] does not have this limitation, however. Even in the presence of feedback noise, his scheme requires only a finite amount of feedback power to achieve improved asymptotic performance over the best one-way scheme at all rates up to the forward channel capacity. His is the first feedback coding scheme with this property. Kramer's scheme uses N forward transmissions and $N-1$ feedback transmissions and is similar to Schemes 1 and 2. For his scheme Kramer shows that

If

$$\frac{P_{FB}}{N_o'} = r \frac{P_{av}}{N_o} \qquad r \geq N(N-1) \tag{4.4}$$

then

$$E(R) = \begin{cases} \dfrac{NP_{av}}{2N_o} - R & 0 \leq R \leq \min\left(\dfrac{NP_{av}}{4N_o}, \dfrac{P_{av}}{N_o}\right) \\[2ex] \left(\sqrt{\dfrac{NP_{av}}{N_o}} - \sqrt{R}\right)^2 & \min\left(\dfrac{NP_{av}}{4N_o}, \dfrac{P_{av}}{N_o}\right) \leq R < \dfrac{P_{av}}{N_o} \end{cases} \tag{4.5}$$

To obtain the same performance as in (4.5) for the same number, N, of forward transmissions, Schemes 1 and 2 require only that

$$\frac{P_{FB}}{N_o'} = r \frac{P_{av}}{N_o} \qquad r \geq N$$

(See Equations (2.15), (2.16), (3.12), and (3.13).) Of course the condition (4.4) on the amount of feedback power required is only a sufficient condition. As Kramer points out, it may in fact be possible to obtain the same $E(R)$ for a smaller value of the feedback power.

With this in mind, a lower bound to the probability of error for Kramer's scheme is obtained in Appendix B. (See Equations (B.8) and (B.9).) Suppose for simplicity that

$$\frac{P_{FB}}{N_o'} = r\,\frac{P_{av}}{N_o}$$

where r is an integer greater than 1 and that $N = r + 1$. Let $E_K(R)$ denote the channel reliability function for Kramer's scheme. It follows from (B.9) that

$$E_K(R) \le \lim_{T\to\infty} \left(-\frac{1}{N\tau}\ln(P_{el}^{N})\right)$$

$$= \begin{cases} \dfrac{(r+1)P_{av}}{2N_o} - (r+1)\,R & 0 \le R \le \dfrac{P_{av}}{4N_o} \\[2ex] \left(\sqrt{\dfrac{(r+1)P_{av}}{N_o}} - \sqrt{(r+1)R}\right)^2 & \dfrac{P_{av}}{4N_o} \le R < \dfrac{P_{av}}{N_o} \end{cases} \tag{4.6}$$

The channel reliability function, $E(R)$, for Scheme 2 with $N = r + 1$ is given by (3.13c), which is repeated here for convenience.

$$E(R) = \begin{cases} \dfrac{(r+1)P_{av}}{2N_o} - 2R & 0 \le R \le \dfrac{P_{av}}{4N_o} \\[2ex] \dfrac{(r+2)P_{av}}{2N_o} - 2\sqrt{\dfrac{RP_{av}}{N_o}} & \dfrac{P_{av}}{4N_o} \le R \le \min\left(\dfrac{rP_{av}}{4N_o}, \dfrac{P_{av}}{N_o}\right) \\[2ex] \dfrac{(r+1)P_{av}}{N_o} - 2\sqrt{\dfrac{RP_{av}}{N_o}}\,(1+\sqrt{r}) + 2R\,\min\left(\dfrac{rP_{av}}{4N_o}, \dfrac{P_{av}}{N_o}\right) & \le R < \dfrac{P_{av}}{N_o} \end{cases} \tag{3.13c}$$

It follows from the above that $E(R) > E_K(R)$ for $R > 0$. In particular, it can be shown that

$$E(R) \geq E_K(R) + (r-1)R \qquad 0 \leq R \leq \frac{P_{av}}{4N_o}$$

and for $r \geq 4$

$$E(R) \geq E_K(R) + \frac{P_{av}}{4N_o}(r-1) \qquad \frac{P_{av}}{4N_o} \leq R < \frac{P_{av}}{N_o}$$

Kramer uses equal amounts of energy on each feedback transmission rather than using all the available energy on the first transmission as in the forward channel. By sending information over the feedback channel in such a way that the expected value of feedback signal energy is zero on all transmissions after the first, Schemes 1 and 2 achieve a reduction in the amount of feedback power required.

Finally, it should be mentioned that for $N = 2$, Schemes 1 and 2 are identical to Kramer's scheme.

V. CONCLUSIONS

A feedback communication system in which the forward and feedback channels are independently disturbed by additive white Gaussian noise and average power constrained was considered. Feedback coding schemes were presented which make efficient use of the feedback power available to obtain improved error performance over existing coding schemes. The behavior of the probability of error is particularly dramatic at rates arbitrarily close to the forward channel capacity, since channel reliability functions were obtained which remain positive at capacity.

The messages to be coded were realized with a set of signals (in this case orthogonal signals) which allow reliable one-way transmission of information over both the forward and feedback channels. The expected value of signal energy in both the forward and feedback channels could then be made negligible on all iterations after the first. In this way all the available signal energy per message could be used on the first iteration, and the probability of error was decreased. This approach can be applied under other assumptions regarding the forward and feedback channels provided that signal sets exist which allow reliable one-way transmission of information over these channels. If the average power were the critical factor in determining the error probability for the forward channel, improved error performance should be obtainable in this way.

It should be pointed out that the coding schemes presented here, while effective, are not optimum. Several modifications are possible. Signal gain constants could be used. However, it was shown that peak power limitations make it reasonable to increase the number of iterations per message as a means of improving error performance rather than using gain constants. The decoder considered is optimum only if the feedback channel is noiseless, and it could be modified. It is, however, desirable that the decoder still be easy to implement and analyze. The decision rule used on the feedback channel could also be modified.

APPENDIX A

WEAK CONVERSE FOR A FEEDBACK COMMUNICATION SYSTEM

Consider the feedback coding scheme of Figure 1. If the forward channel is discrete and memoryless and the feedback channel is noiseless, Shannon [7] has shown that such a scheme cannot increase the capacity of the forward channel. This result is now extended to a system in which the forward and feedback channels are independent, time discrete, amplitude continuous, and memoryless. In what follows we assume that all random variables have bounded density functions and finite variances so that all integrals exist.

Let $p(y/x)$ be the conditional probability density describing the forward channel, where the channel inputs x and outputs y are points on the real line. Let $\mathcal{S}$ denote the set of M messages to be coded for transmission over the forward channel and V denote the space of forward channel output sequences (V is Euclidean D-space). We assume the forward channel inputs are constrained so that

$$\frac{1}{D} \sum_{j=1}^{D} \varepsilon(h(x_j)) \le K$$

where $\varepsilon(\ )$ is an operator denoting statistical expectation, h is a real-valued function, and K is a constant. Let

$$I(\mathcal{S};V) = \sum_{\mathcal{S}} \int_{V} P(s)\ p(v/s)\ \ln \frac{p(v/s)}{p(v)}\ dv$$

be the mutual information between the set of messages and the space of forward channel output sequences for a given feedback coding scheme.

Let

$$I(X;Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x)p(y/x) \ln \frac{p(y/x)}{p(y)} \, dy \, dx$$

be the forward channel mutual information, and

$$C = \underset{p(x)\,:\,\int_{-\infty}^{\infty} p(x)h(x)dx \leq K}{\mathrm{Max}} I(X;Y)$$

be the forward channel capacity.

We now prove the following

Lemma. $\qquad I(\mathcal{S};V) \leq DC$

Proof: Since the channels are independent and memoryless, the channel output y_j depends only on the channel input x_j. Therefore,

$$p(y_j/s, y_1, \cdots, y_{j-1}, x_j) = p(y_j/x_j)$$

Making use of this, Gallager has shown (see Appendix A of Reference 4)

$$I(\mathcal{S};V) \leq \sum_{j=1}^{D} I_j(X;Y)$$

The mutual information, $I_j(X;Y)$, is computed using the density, $p_j(x)$, on the j^{th} channel input for the given feedback coding scheme. We define a probability density function, $p'(x)$, as follows.

$$p'(x) = \frac{1}{D} \sum_{j=1}^{D} p_j(x)$$

Let $I'(X;Y)$ be the mutual information computed using this density function. Letting

$$p_j(y) = \int_{-\infty}^{\infty} p_j(x)p(y/x)dx \quad \text{and} \quad p'(y) = \int_{-\infty}^{\infty} p'(x)p(y/x)dx$$

we have

$$I'(X;Y) = \int_{-\infty}^{\infty} p'(y)\ln \frac{1}{p'(y)}\, dy - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p'(x)p(y/x)\, \ln \frac{1}{p(y/x)}\, dy\, dx$$

$$= \frac{1}{D} \sum_{j=1}^{D} \int_{-\infty}^{\infty} p_j(y)\ln \frac{1}{p'(y)}\, dy - \frac{1}{D} \sum_{j=1}^{D} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_j(x)p(y/x)\, \ln \frac{1}{p(y/x)}\, dy\, dx$$

$$\geq \frac{1}{D} \sum_{j=1}^{D} \left(\int_{-\infty}^{\infty} p_j(y)\ln \frac{1}{p_j(y)}\, dy - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p_j(x)p(y/x)\ln \frac{1}{p(y/x)}\, dy\, dx \right)$$

$$= \frac{1}{D} \sum_{j=1}^{D} I_j(X;Y)$$

Hence

$$I(\mathscr{S};V) \leq DI'(X;Y)$$

The forward channel inputs are constrained so that

$$\frac{1}{D} \sum_{j=1}^{D} \int_{-\infty}^{\infty} p_j(x)\ h(x)\ dx \le K$$

Therefore, $p'(x)$ satisfies

$$\int_{-\infty}^{\infty} p'(x)\ h(x)\ dx \le K$$

and

$$I'(X;Y) \le C$$

The proof is complete.

Assuming the messages are equiprobable, we now have the following.

Theorem (Weak Converse). If $R = \frac{\ln M}{D} > C$, the probability of error, $P(e)$, for the feedback coding scheme is bounded away from zero.

Proof: The proof is standard (see Chapter 8 of Reference 10). Let

$$H(\mathscr{S}/V) = \sum_{\mathscr{S}} \int_{V} P(s/v)p(v)\ \ln \frac{1}{P(s/v)}\ dv$$

Then

$$H(\mathscr{S}/V) \le P(e)\ln \frac{1}{P(e)} + (1-P(e))\ln \frac{1}{1-P(e)} + P(e)\ln(M-1)$$

This is Fano's inequality. We also have, for a set of equiprobable messages,

$$I(\mathcal{S};V) = \ln M - H(\mathcal{S}/V)$$

Application of the lemma now yields the weak converse.

Hence, feedback coding cannot increase the capacity, C, of the forward channel. In particular, the above results apply to the vector channel model for the feedback communication system considered in this paper.

APPENDIX B

A LOWER BOUND TO THE PROBABILITY OF ERROR FOR KRAMER'S SCHEME

The symbols to be used here have been previously defined in Chapter II. The description of feedback coding Scheme 1 (see Section 2.2) applies to Kramer's coding scheme [6] with the following changes. We transmit

$$w_k = \hat{\phi}_k \qquad k = 1, \cdots, N-1$$

over the feedback channel and let

$$\beta_k = z_k \qquad k = 1, \cdots, N-1$$

to determine ϕ_k^*. Note that the probability that $\phi_k^* \neq \hat{\phi}_k$ is simply P_{efb}, independent of the values of ϕ_ℓ^* and $\hat{\phi}_\ell$, $\ell=1, \cdots, k-1$.

It can be shown for this scheme that

$$P_{av} = \frac{S}{N}\left(1 + 2(N-1)\, P_{efb} + 2\left(1 - \left(\frac{M}{M-1}\right) P_{efb}\right)\sum_{k=1}^{N-1} P_k(e)\right) \tag{B.1}$$

$$P_{FB} = \frac{(N-1)S'}{N} \tag{B.2}$$

$$P_k(e) \leq P_{ek} + (k-1)\, P_{efb} \tag{B.3}$$

(See Chapter III of Reference 6.)

We now obtain a lower bound to the probability of error for this scheme. To do this we consider the following table of events (similar to that in Section 3.3) associated with the possible values of θ^*_{N-1} and $\hat{\theta}_{N-1}$. It is assumed that s_i is the message being coded.

Event	$\hat{\theta}_{N-1}$	θ^*_{N-1}	$\lambda_N = \lambda_{N-1} + y_N + \hat{\theta}_{N-1}$
A_{N-1}	θ_i	θ_i	$\lambda_{N-1} + n_N + \theta_i$
B_{N-1}	θ_i	θ_j $j\neq i$	$\lambda_{N-1} + n_N + 2\theta_i - \theta_j$
C_{N-1}	θ_j $j\neq i$	θ_j	$\lambda_{N-1} + n_N + \theta_i$
D_{N-1}	θ_j $j\neq i$	θ_ℓ $\ell\neq i,j$	$\lambda_{N-1} + n_N + \theta_i - \theta_\ell + \theta_j$
E_{N-1}	θ_j $j\neq i$	θ_i	$\lambda_{N-1} + n_N + \theta_j$

The method of analysis is similar to that of Section 3.3. The events are disjoint so that

$$\begin{aligned}P_N(e/s_i) = {} & P(\hat{\theta}_N \neq \theta_i/s_i,A_{N-1})\, P(A_{N-1}/s_i) + P(\hat{\theta}_N \neq \theta_i/s_i,B_{N-1})P(B_{N-1}/s_i) \\ & + P(\hat{\theta}_N \neq \theta_i/s_i,C_{N-1})\, P(C_{N-1}/s_i) + P(\hat{\theta}_N \neq \theta_i/s_i,D_{N-1})P(D_{N-1}/s_i) \\ & + P(\hat{\theta}_N \neq \theta_i/s_i,E_{N-1})\, P(E_{N-1}/s_i) \qquad \text{(B.4)}\end{aligned}$$

We now lower bound the terms in this sum. Proceeding as in Section 3.3,

$$P(\hat{\theta}_N \neq \theta_i/s_i, D_{N-1}) = P(\langle \lambda_N, \theta_r \rangle \geq \langle \lambda_N, \theta_i \rangle \quad \text{for some} \quad r \neq i/s_i, D_{N-1})$$

$$\geq P(\langle \lambda_N, \theta_j \rangle \geq \langle \lambda_N, \theta_i \rangle / s_i, D_{N-1})$$

$$= P(\langle \lambda_{N-1}, \theta_j \rangle + \langle n_N, \theta_j \rangle + E \geq \langle \lambda_{N-1}, \theta_i \rangle + \langle n_N, \theta_i \rangle + E / \langle \lambda_{N-1}, \theta_j \rangle \geq$$
$$\langle \lambda_{N-1}, \theta_i \rangle)$$
$$\geq P(\langle n_N, \theta_j \rangle \geq \langle n_N, \theta_i \rangle)$$
$$= \frac{1}{2}$$

Similarly, it can be shown that

$$P(\hat{\theta}_N \neq \theta_i/s_i, E_{N-1}) \geq \frac{1}{2}$$

The conditional probabilities of events D_{N-1} and E_{N-1} are

$$P(D_{N-1}/s_i) = P_{N-1}(e/s_i) \ P_{efb}(\frac{M-2}{M-1})$$

and

$$P(E_{N-1}/s_i) = \frac{P_{N-1}(e/s_i)P_{efb}}{M - 1}$$

Lower bounding the second term in (B.4) by zero, it follows from the above that

$$P_N(e/s_i) \geq P(\hat{\theta}_N \neq \theta_i/s_i, A_{N-1})P(A_{N-1}/s_i) + P(\hat{\theta}_N \neq \theta_i/s_i, C_{N-1})P(C_{N-1}/s_i)$$
$$+ \frac{1}{2} P_{N-1}(e/s_i) \ P_{efb} \tag{B.5}$$

It is difficult to obtain an exact expression for the first two terms in (B.5). However, a simple lower bound to these terms follows from noting that

$$P(\hat{\theta}_N \neq \theta_i/s_i, A_{N-1})P(A_{N-1}/s_i) + P(\hat{\theta}_N \neq \theta_i/s_i, C_{N-1})P(C_{N-1}/s_i)$$

$$= P(\hat{\theta}_N \neq \theta_i/s_i, \theta^*_{N-1} = \hat{\theta}_{N-1})P(\theta^*_{N-1} = \hat{\theta}_{N-1}/s_i)$$

$$\geq P(\hat{\theta}_N \neq \theta_i/s_i, \theta^*_k = \hat{\theta}_k \quad \text{for} \quad k = 1, \cdots, N-1)P(\theta^*_k = \hat{\theta}_k$$

$$\text{for} \quad k = 1, \cdots, N-1/s_i)$$

$$= P_{eN}(1-P_{efb})^{N-1}$$

Therefore,

$$P_N(e) \geq P_{eN}(1-P_{efb})^{N-1} + \frac{P_{N-1}(e)\ P_{efb}}{2} \tag{B.6}$$

A similar lower bound to $P_{N-1}(e)$ may be obtained. Substituting this lower bound in (B.6), lower bounding $P_{N-2}(e)$, and continuing in this way, the following lower bound to $P_N(e)$ is obtained.

$$P_N(e) \geq \sum_{k=1}^{N} P_{ek}(1-P_{efb})^{k-1}\left(\frac{P_{efb}}{2}\right)^{N-k} \tag{B.7}$$

Suppose now that

$$\frac{S'}{N'_o} = \left(\frac{r}{N-1}\right)\frac{S}{N_o} \qquad r \geq 1 \quad (r \text{ need not be an integer})$$

and

$$0 \leq R < \frac{S}{NN_o}$$

Then

$$P_{ek} < P_{e1} \to 0 \quad \text{as} \quad \tau \to \infty \quad \text{for} \quad k > 1$$

$$P_{efb} = P_{e\frac{r}{N-1}}$$

and if $N \leq r + 1$ then

$$P_{efb} \to 0 \quad \text{as} \quad \tau \to \infty$$

It then follows from (B.1) and (B.3) that

$$P_{av} \to \frac{S}{N} \quad \text{as} \quad \tau \to \infty$$

Using (B.2) and (B.7), it follows from the above that

If

$$\frac{P_{FB}}{N_o'} = r\,\frac{P_{av}}{N_o} \qquad r \geq 1 \quad \text{and} \quad N \leq r + 1 \tag{B.8}$$

then

$$P_N(e) \geq \sum_{k=1}^{N} P_{ek}\left(1 - P_{e\,\frac{r}{N-1}}\right)^{k-1}\left(\frac{P_{e\,\frac{r}{N-1}}}{2}\right)^{N-k} \tag{B.9}$$

This is the desired lower bound.

REFERENCES

[1] E. R. Berlekamp, "Block Coding with Noiseless Feedback", Ph.D. Dissertation, M.I.T., Department of Electrical Engineering, September 1964.

[2] J. P. M. Schalkwijk and T. Kailath, "A Coding Scheme for Additive Noise Channels with Feedback - Part I: No Bandwidth Constraint", I.E.E.E. Transactions on Information Theory, IT-12, 172-182, April 1966.

[3] J. P. M. Schalkwijk, "A Coding Scheme for Additive Noise Channels with Feedback - Part II: Band-Limited Signals", I.E.E.E. Transactions on Information Theory, IT-12, 183-189, April 1966.

[4] J. K. Omura, "Signal Optimization for Channels with Feedback", Report SEL-66-068, Stanford Electronics Labs., Stanford University, August 1966.

[5] S. Butman, "Optimum Linear Coding for Additive Noise Systems Using Information Feedback", Technical Report No. 1, Communications Theory Lab., California Institute of Technology, May 1967.

[6] A. J. Kramer, "Analysis of Communication Schemes Using an Intermittent Feedback Link", Report SEL-67-014, Stanford Electronics Labs., Stanford University, March 1967.

[7] C. E. Shannon, "The Zero-Error Capacity of a Noisy Channel", I.R.E. Transactions on Information Theory, IT-2, 8-19, September 1956.

[8] P. Elias, "Networks of Gaussian Channels with Applications to Feedback Systems", I.E.E.E. Transactions on Information Theory, IT-13, 493-501, July 1967.

[9] S. M. Farber, Private Communication, 1968 (to be published).

[10] R. B. Ash, Information Theory, John Wiley and Sons, Inc., New York, 1965.

[11] J. M. Wozencraft and I. M. Jacobs, Principles of Communication Engineering, John Wiley and Sons, Inc., New York, 1965.