

# A coding scheme for additive noise channels with feedback, Part II: Band-limited signals

**Citation for published version (APA):**

Schalkwijk, J. P. M. (1968). A coding scheme for additive noise channels with feedback, Part II: Band-limited signals. *IEEE Transactions on Information Theory*, IT-12(2), 183-189. <https://doi.org/10.1109/TIT.1966.1053880>

**DOI:**

[10.1109/TIT.1966.1053880](https://doi.org/10.1109/TIT.1966.1053880)

**Document status and date:**

Published: 01/01/1968

**Document Version:**

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

# A Coding Scheme for Additive Noise Channels with Feedback

## Part II: Band-Limited Signals

J. PIETER M. SCHALKWIJK, MEMBER, IEEE

**Abstract**—In Part I of this paper, we presented a scheme for effectively exploiting a noiseless feedback link associated with an additive white Gaussian noise channel with no signal bandwidth constraints. We now extend the scheme for this channel, which we shall call the wideband (WB) scheme, to a band-limited (BL) channel with signal bandwidth restricted to  $(-W, W)$ . Our feedback scheme achieves the well-known channel capacity,  $C = W \ln(1 + P_{av}/N_0W)$ , for this system and, in fact, is apparently the first deterministic procedure for doing this. We evaluate the fairly simple exact error probability for our scheme and find that it provides considerable improvements over the best-known results (which are lower bounds on the performance of sphere-packed codes) for the one-way channel. We also study the degradation in performance of our scheme when there is noise in the feedback link.

### I. INTRODUCTION

IN THIS PAPER a band-limited (BL) channel with feedback is considered. The signal bandwidth is restricted to  $(-W, W)$ .

A general introduction has been given in Part I, with particular attention to wideband (WB) channels.

The BL coding scheme developed here, is as far as we know, the first deterministic coding procedure to achieve the well-known capacity

$$C = W \ln [1 + (P_{av}/N_0W)].$$

To our knowledge, the only other results pertaining to the band-limited (BL) channel have been published by Elias [3]. He divided the channel into  $K$  subchannels of bandwidth  $w = W/K$ . If noiseless feedback is available and if  $K \rightarrow \infty$ , information can be sent at a rate equal to  $W \ln [1 + (P_{av}/N_0W)]$ . However, since the signal bandwidth is  $w$  instead of  $W$ , the coding and decoding complexity for the feedback scheme becomes an arbitrarily small fraction of that required without feedback.

### II. A FEEDBACK COMMUNICATION SYSTEM WITH A CONSTRAINT ON THE BANDWIDTH

Let  $T$  be the time in seconds necessary for the transmission of a particular message. For the WB coding scheme discussed in Part I, as for orthogonal codes in one-way transmission, the bandwidth  $W(T)$  of the transmission is an exponential function of the coding delay  $T$ .

This work was supported by NASA NsG 377.

The author is with the Applied Research Laboratory, Sylvania Electronic Systems, a division of Sylvania Electric Products Inc., Waltham, Mass.

In order to make the probability of error vanish for a fixed relative rate smaller than one, a large bandwidth is required.

Suppose now that one is given a fixed bandwidth  $W$ , which the transmission is not supposed to exceed. With this additional transmitter constraint imposed, the channel capacity  $C$  is no longer  $P_{av}/N_0$  as before, but it now given by  $W \ln[1 + (P_{av}/N_0W)]$ , nats/second. For small values of  $P_{av}/N_0W$  the latter capacity approaches  $P_{av}/N_0$  as it should, for when  $W \rightarrow \infty$ , both channels are identical.

Shannon [1] derives the capacity formula,  $W \ln [1 + (P_{av}/N_0W)]$ , by a random coding argument, and up till now no deterministic way was known for constructing a code achieving the critical rate for a band-limited white Gaussian noise channel with or without feedback. In this part the first such code will be developed for the case where noiseless feedback is available.

As in Part I, an optimization for finite block-length is carried through, the results are compared with bounds on one-way transmission plotted by Slepian [4], and the deterioration of the present scheme due to feedback noise is considered.

#### A. The BL Coding Scheme

In the WB coding scheme discussed in Part I, the variance of the estimate  $X_{N+1}$  for the message point  $\theta$  was inversely proportional to the number  $N$  of iterations. The critical rate was  $R_{crit} = (\ln N)/2T$  nats/second, and in order to achieve a constant rate one had to choose  $N = e^{2AT}$ , that is, the number of transmissions had to increase exponentially with time.

Now suppose one has to meet a bandwidth constraint  $W$  in cycles per second. In this case the number of independent transmissions can only increase linearly with time. The highest number of independent transmissions per second is approximately equal to  $2W$ . Substituting  $N = 2WT$  in the equation for the critical rate above gives  $R_{crit} = (\ln 2WT)/2T$  nats/second. Hence,  $R_{crit} \rightarrow 0$  with increasing  $T$ , and so the system discussed in Part I has to be modified in order to achieve a constant rate different from zero in the band-limited case.

Two useful observations can be made at this point. First, while the critical rate approaches zero when we take  $2W$  iterations per second the asymptotic relation  $R_{crit}(T) \approx P_{av}(T)/N_0$  is still valid. In other words, both the rate and the average power approach zero for

increasing  $T$ . The limit of their ratio, however, is equal to the constant  $N_0$ . The second observation is that  $X_{N+1}$  can be looked at as the maximum likelihood estimate of  $\theta$  having observed  $Y_1(X_1)$  through  $Y_n(X_n)$ , and assuming Gaussian noise, as explained in Part I-Section II.

With these two observations in mind, we shall present a coding scheme for the band-limited white Gaussian noise channel.

Suppose that transmissions take place at integer values of time, the time unit being  $1/(2W)$  second. Numbers are sent again by amplitude modulating some basic waveform of bandwidth  $W$  and unit energy. The disturbance is white Gaussian noise (with spectral density  $N_0/2$ ), and reception takes place using a matched filter.

The coding scheme starts out the same as in Part I-Section II. At the transmitter:

1) divide the unit interval  $[0, 1]$  into  $M$  disjoint message intervals of equal length; let  $\theta$  be the midpoint of the message interval corresponding to the particular message to be transmitted, and

2) at instant one, transmit  $\alpha(X_{11} - \theta)$ , where  $X_{11} = 0.5$  and  $\alpha$  is some constant to be determined later.

At the receiver:

1) receive  $Y_{11}(X_{11}) = \alpha(X_{11} - \theta) + Z_{11}$ , where  $Z_{11}$  is as before a Gaussian random variable with mean zero and variance  $\sigma^2 = N_0/2$ , and

2) compute  $X_{12} = X_{11} - \alpha^{-1}Y_{11}(X_{11})$ , then set  $X_{21} = X_{12}$  and send  $X_{21}$  back to the transmitter.

Up to this point everything is the same as for the coding scheme of Part I-Section II-B. In other words,  $X_{21} - \theta = -(1/\alpha)Z_{11}$ , where  $X_{21}$  is the maximum likelihood estimate of  $\theta$  having observed  $Y_{11}(X_{11})$ .

Now, in order to prevent the expected power per transmission from decreasing, as it did in the WB coding scheme, we shall take the next transmission as  $g\alpha(X_{21} - \theta)$  instead of  $\alpha(X_{21} - \theta)$ , where the constant  $g$  will be determined presently. The receiver obtains the noisy observation

$$Y_{21}(X_{21}) = g\alpha(X_{21} - \theta) + Z_{21}$$

and then computes

$$X_{22} = X_{21} - (g\alpha)^{-1}Y_{21}(X_{21}).$$

We now have two independent estimates of  $\theta$ :

$$X_{21} = \theta - \frac{1}{\alpha}Z_{11} \quad \text{and} \quad X_{22} = \theta - \frac{1}{g\alpha}Z_{21}.$$

For the value  $X_{31}$  to be sent back to the transmitter, we shall take the maximum likelihood estimate of  $\theta$  having observed  $Y_{11}(X_{11})$  and  $Y_{21}(X_{21})$ , that is, we shall set

$$X_{31} = \frac{\left(\frac{1}{g\alpha}\right)^2 X_{21} + \left(\frac{1}{\alpha}\right)^2 X_{22}}{\left(\frac{1}{\alpha}\right)^2 + \left(\frac{1}{g\alpha}\right)^2} = \frac{X_{21} + g^2 X_{22}}{1 + g^2}.$$

What is the variance of our successive maximum likelihood estimates  $X_{11}$ ,  $X_{21}$ , and  $X_{31}$ ? It is known that

$$X_{21} \sim N\left(\theta, \frac{\sigma^2}{\alpha^2}\right) \quad \text{and} \quad X_{31} \sim N\left(\theta, \frac{\sigma^2}{\alpha^2} \frac{1}{1 + g^2}\right).$$

If, however,  $g = (\alpha^2 - 1)^{1/2}$  is chosen, then

$$X_{31} \sim N\left(\theta, \frac{\sigma^2}{(\alpha^2)^2}\right).$$

In general,  $X_{i1}$ ,  $i = 2, 3, \dots$ , is sent back. The next transmission is

$$\alpha^{i-1}(\alpha^2 - 1)^{1/2}(X_{i1} - \theta) \quad (1)$$

but the receiver obtains

$$Y_{i1}(X_{i1}) = \alpha^{i-1}(\alpha^2 - 1)^{1/2}(X_{i1} - \theta) + Z_{i1}$$

and then computes

$$X_{i2} = X_{i1} - [\alpha^{i-1}(\alpha^2 - 1)^{1/2}]^{-1}Y_{i1}(X_{i1})$$

and

$$X_{i+1,1} = \frac{X_{i1} + (\alpha^2 - 1)X_{i2}}{\alpha^2}.$$

The maximum likelihood estimate  $X_{i+1,1}$  is normally distributed with mean  $\theta$  and variance  $\sigma^2/[(\alpha^2)^i]$ , that is,

$$X_{i+1,1} \sim N\left[\theta, \frac{\sigma^2}{(\alpha^2)^i}\right]. \quad (2)$$

From this point on, the analysis is very similar again to that of Part I. Suppose the transmitter sends one of  $M$  possible messages, that is, the interval  $[0, 1]$  is divided into  $M$  disjoint equal-length message intervals. The message point  $\theta$  is the midpoint of the message interval corresponding to the particular message being transmitted. The probability of the receiver deciding on the wrong message interval (i.e., the probability of  $X_{N+1}$  lying outside the correct interval) is

$$P_e = 2 \operatorname{erfc}\left(\frac{\frac{1}{2}M^{-1}}{\sigma/\alpha^N}\right).$$

Now pick  $M = \alpha^{N(1-\epsilon)}$ , that is,  $R = (\ln M)/N = (1 - \epsilon) \ln \alpha$ , nats/dimension (the time unit was  $1/(2W)$  seconds). This gives for the probability of error

$$P_e = 2 \operatorname{erfc}\left(\frac{\alpha^{N\epsilon}}{2\sigma}\right) \quad (3)$$

and thus,

$$\lim_{N \rightarrow \infty} P_e(N, \epsilon) = \begin{cases} 0 & \text{for } \epsilon > 0 \\ 1 & \text{for } \epsilon < 0. \end{cases}$$

In other words, the critical rate is equal to  $R_{\text{crit}} = \ln \alpha$ , nats/dimension. Putting  $\alpha = e^A$  gives  $R_{\text{crit}} = A$ .

Next let us derive an expression for the average power  $P_{av}$ .

$$P_{av} = \frac{1}{T} E \left\{ \alpha^2 (X_{11} - \theta)^2 + \sum_{i=2}^N [\alpha^{i-1} (\alpha^2 - 1)^{1/2}]^2 (X_{i,1} - \theta)^2 \right\}$$

$$= \frac{1}{T} \left\{ \alpha^2 E (X_{11} - \theta)^2 + \sum_{i=2}^N [\alpha^{i-1} (\alpha^2 - 1)^{1/2}]^2 \frac{\sigma^2}{(\alpha^2)^{i-1}} \right\}.$$

Substituting  $T = N/(2W)$  seconds,  $\sigma^2 = N_0/2$ ,  $\alpha = e^A$ , and  $E(X_{11} - \theta)^2 = 1/12$  (assuming a uniform prior distribution for  $\theta$ ), one gets

$$P_{av} = \frac{W e^{2A}}{6N} + \frac{N-1}{N} N_0 W (e^{2A} - 1). \quad (4)$$

Hence, asymptotically,

$$R_{crit} = \begin{cases} A = \frac{1}{2} \ln \left( 1 + \frac{P_{av}}{N_0 W} \right) & \text{nats/dimension, or} \\ 2WA = W \ln \left( 1 + \frac{P_{av}}{N_0 W} \right) & \text{nats/second} \end{cases} \quad (5)$$

which is the channel capacity as computed by Shannon [1]! This result proves that the BL coding system presented here achieves capacity for the band-limited white Gaussian noise channel. It is the first deterministic coding procedure to do so.

### B. Optimization for Finite Blocklength

As in Part I, let us now investigate how far one falls short of the ideal when only permitting a finite coding delay  $N$  (in time units of  $1/(2W)$  second).

In Section I-A, the slope  $\alpha_i$  at the  $i$ th transmission was taken as

$$\alpha_i = \alpha^{i-1+\delta_{i1}} (\alpha^2 - 1)^{1/2(1-\delta_{i1})}$$

where

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

In order to make an optimization possible, an additional factor  $a$  is introduced, hence,

$$\alpha_i = a \alpha^{i-1+\delta_{i1}} (\alpha^2 - 1)^{1/2(1-\delta_{i1})} \quad i = 1, 2, 3, \dots \quad (6)$$

The receiver now has

$$Y_{i1}(X_{i1}) = \alpha_i (X_{i1} - \theta) + Z_{i1}$$

and computes

$$X_{i2} = X_{i1} - \alpha_i^{-1} Y_{i1}(X_{i1}),$$

$$X_{21} = X_{12},$$

and

$$X_{i+1,1} = \frac{X_{i1} + (\alpha^2 - 1)X_{i2}}{\alpha^2} \quad \text{for } i = 2, 3, \dots$$

Effectively, the introduction of the factor  $a$  reduces  $N_0$  by a factor  $a^2$ .

By (3), minimizing the probability of error is equivalent to maximizing the expression

$$a^2 \frac{\alpha^{2N\epsilon}}{2N_0} \quad (7)$$

where  $\epsilon$  can be obtained from

$$R = (1 - \epsilon) \ln \alpha \text{ nats/dimension} \quad (8)$$

and  $\sigma^2 = N_0/2$  was substituted for the variance.

Substituting  $\alpha = e^A$  in (4) and allowing for the additional factor  $a$  leads to the following expression for the average power:

$$P_{av} = a^2 \left[ \frac{W \alpha^2}{6N} + \frac{N-1}{N} \frac{N_0}{a^2} W (\alpha^2 - 1) \right]$$

which can be modified as

$$\frac{P_{av}}{N_0 W} = a^2 \frac{\alpha^2}{6N N_0} + \frac{N-1}{N} (\alpha^2 - 1). \quad (9)$$

Now assuming  $C$ ,  $R$ ,  $W$ ,  $N_0$ , and  $N$  constant, let us maximize (7) with respect to  $a^2$ . Note that  $C$  and  $W$  constant implies  $P_{av}/N_0 W$  constant, for

$$C = W \ln \left( 1 + \frac{P_{av}}{N_0 W} \right).$$

Having gone through these preliminaries, one is now ready to perform the optimization. Set the derivative of  $a^2 (\alpha^{2N\epsilon}/2N_0)$  equal to zero,

$$\frac{d}{da^2} \left( a^2 \frac{\alpha^{2N\epsilon}}{2N_0} \right) = \frac{\alpha^{2N\epsilon}}{2N_0} + a^2 \frac{d}{da^2} \left( \frac{\alpha^{2N\epsilon}}{2N_0} \right) \frac{d\alpha^2}{da^2}$$

$$+ a^2 \frac{d}{d\epsilon} \left( \frac{\alpha^{2N\epsilon}}{2N_0} \right) \frac{d\epsilon}{da^2} \frac{d\alpha^2}{da^2} = 0. \quad (10)$$

From (8) it follows that

$$\epsilon = 1 - \frac{2R}{\ln \alpha^2} \quad \text{and} \quad \frac{d\epsilon}{d\alpha^2} = \frac{2R}{\alpha^2} \frac{1}{(\ln \alpha^2)^2}$$

and from (9) it follows that

$$\frac{d\alpha^2}{da^2} = -\frac{\alpha^2}{a^2 + 6N_0(N-1)}.$$

Making these substitutions in (10) and putting the result equal to zero finally gives, after some algebra, the following simple expression for the optimum value  $a_0^2$  of  $a^2$ :

$$a_0^2 = 6N_0. \quad (11)$$

For the probability of error, substituting  $\sigma^2 = N_0/2$  and  $a_0^2 = 6N_0$  in (3), one has

$$P_e = 2 \operatorname{erfc} [(3\alpha^{2N\epsilon})^{1/2}].$$

Solving for  $\alpha^2$  from (9) and (11) gives

$$\alpha^2 = \frac{N-1}{N} + \frac{P_{av}}{N_0 W}.$$

By (8) one has

$$\frac{R}{2W} = \ln \alpha^{1-\epsilon} \quad \text{or} \quad \alpha^{1-\epsilon} = \exp\left(\frac{R}{2W}\right)$$

where  $R$  is now in nats/second. Hence,

$$\alpha^\epsilon = \frac{\alpha}{\exp\left(\frac{R}{2W}\right)} = \frac{\left(\frac{N-1}{N} + \frac{P_{av}}{N_0W}\right)^{1/2}}{\exp\left(\frac{R}{2W}\right)}$$

and finally,

$$P_e = 2 \operatorname{erfc} \left\{ \sqrt{3} \left[ \frac{\frac{N-1}{N} + \frac{P_{av}}{N_0W}}{\exp\left(\frac{R}{W}\right)} \right]^{N/2} \right\}. \quad (12)$$

This final result will be compared in the next section with the bounds on one-way communication as obtained by Slepian [4].

### C. Comparison with Slepian's Results

In 1963 Slepian [4] plotted lower bounds on communication in the one-way case based on a geometrical approach to the coding problem for band-limited white Gaussian noise channels used by Shannon [2]. That is, there is no one-way communication system whose performance is any better than that plotted by Slepian. Figures 1 through 6 compare Slepian's curves (dashed lines) with the results described by (12) (solid lines). Note that the solid curves are exact, that is, they are not a bound as Slepian's curves are. The graphs presented are described in the following.

1) Figure 1 shows the signal-to-noise ratio  $S/N = 10 \log_{10} (P_{av}/N_0W)$  in decibels vs. the rate  $R/W$  in dits/cycle, as given by Shannon's capacity equation,  $\ln [1 + (P_{av}/N_0W)]$ .

2) Figures 2(a) to 2(c) indicate the additional signal-to-noise ratio, in decibels above the value indicated in Fig. 1, required for a finite coding delay  $N$ , as a function of the rate in dits per cycle. The probability of error for the three figures is, respectively,  $P_e = 10^{-2}$ ,  $10^{-4}$ , and  $10^{-6}$ . It is seen that a large improvement is obtained by going from  $N = 5$  to  $N = 15$ , especially in the feedback scheme. Increasing the coding delay further does not result in much improvement.

3) Figures 3(a), and 3(b) are plots of the additional signal-to-noise ratio in decibels above the ideal value indicated in Fig. 1 vs. the coding delay  $N$ , for different values of the probability of error  $P_e$  and for a rate of  $R/W = 0.2$  dit per cycle. Figure 3(b) represents a plot for the bounds computed by Slepian. Note that the curves for the feedback scheme in Fig. 3(a) indicate a much lower relative (to the ideal, given in Fig. 1) signal-to-noise ratio, except for extremely small values of  $N$ .

4) Figures 4(a) and (b) are plots of the probability of error vs. the coding delay  $N$ , with the signal-to-noise

ratio in decibels above the ideal as the parameter. The rate is  $R/W = 0.2$  dit per cycle. Note the difference in shape between the two sets of curves.

5) Figure 5 is a plot of the relative rate  $R/C$  vs. the rate  $R/W$  in dits per cycle for different values of the coding delay. The probability of error is  $P_e = 10^{-4}$ .

6) Figures 6(a) and (b) are plots of the relative rate  $R/C$  vs. the coding delay  $N$  for different values of the signal-to-noise ratio.

### D. Influence of Feedback Noise on the BL Coding Scheme

In this section, only the configuration in which  $Y'_n(X'_n)$  (the received "number") is sent back will be investigated. The results for the case where  $X_n$  (the receiver's estimate) is sent back are similar to those in Part I in that the rate drops off to zero quickly.

Using the same notation as in Part I, it follows easily that

$$X'_{N+1,1} = X'_{N+1,1} + \alpha^{-1} Z'_{11} + \frac{\alpha^2 - 1}{\alpha^2} \sum_{i=2}^N \alpha_i^{-1} Z_{i1} \quad (13)$$

where  $\sum_{i=2}^1 = 0$ , and  $\alpha_i$  is given by (6). Hence,

$$X'_{N+1,1} \sim N \left\{ \theta, \frac{1}{\alpha^2} \left[ \frac{\sigma'^2 + \sigma''^2}{\alpha^{2N}} + \frac{\sigma''^2}{\alpha^2} + \sigma''^2 (\alpha^{2(N-1)} - 1) \right] \right\}.$$

The variance  $\sigma_i^2$  of the estimate  $X'_{N+1}$  of  $\theta$ , as computed by the receiver, is

$$\sigma_i^2 = \frac{1}{\alpha^2} \left[ \frac{\sigma'^2 + \sigma''^2}{\alpha^{2N}} + \frac{\sigma''^2}{\alpha^2} + \sigma''^2 (\alpha^{2(N-1)} - 1) \right]. \quad (14)$$

For the probability of error one has, from (3),

$$P_e = 2 \operatorname{erfc} \left[ \frac{\alpha^{N(1-\epsilon)}}{2\sigma_i} \right] \quad (15)$$

where again  $R = (1 - \epsilon) \ln \alpha$ , nats/dimension.

The expression for the signal-to-noise ratio in the forward direction is, from (9),

$$\frac{P_{av}}{N_0W} = \alpha^2 \frac{\alpha^2}{6N_0} + \frac{N-1}{N} \frac{N'_0 + N''_0}{N'_0} (\alpha^2 - 1). \quad (16)$$

Figure 7 presents curves for the probability of error  $P_e$  vs. the coding delay  $N$  for  $R/W = 0.2$  dit per cycle, and different values of the feedback noise relative to the forward noise,  $N''_0/N'_0$ . For  $\alpha^2$  the value  $\alpha^2 = 6N'_0$  as given by (11) is used. Hence, the curves present the degradation due to feedback noise of a system that is optimum for the noiseless feedback case.

### III. CONCLUDING REMARKS

The WB (wideband) coding scheme, discussed in Part I, was suggested by the Robbins-Monro stochastic approximation procedure. In the Gaussian case it turns out that this coding procedure determines the maximum likelihood estimate of the message point  $\theta$  recursively. Since the maximum likelihood estimate approaches 0

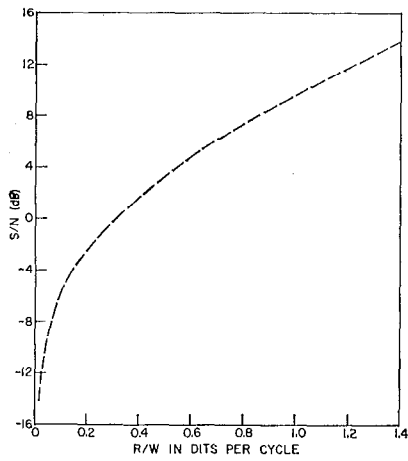
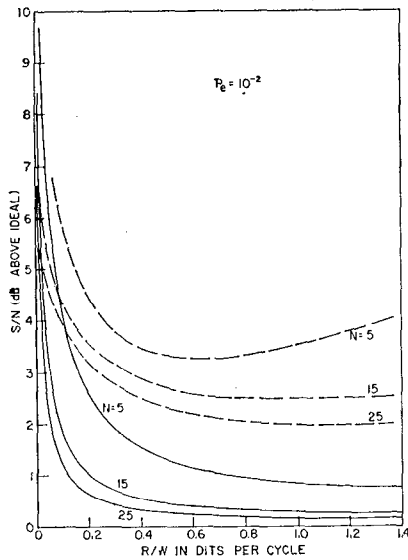
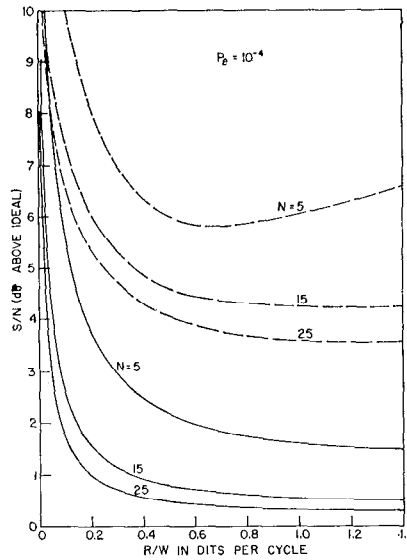


Fig. 1. The signal-to-noise ratio required by Shannon's capacity equation.

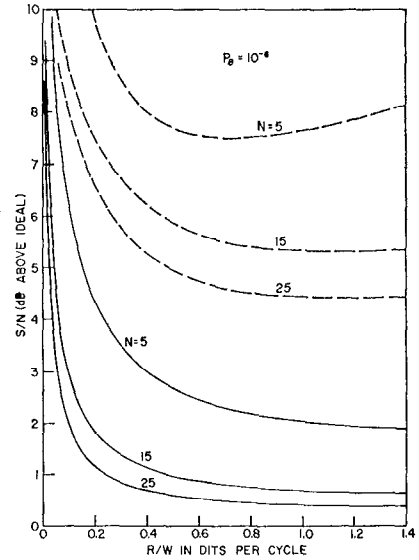
Fig. 2. The additional signal-to-noise ratio required when using a finite coding delay. (a)  $P_e = 10^{-2}$ . (b)  $P_e = 10^{-4}$ . (c)  $P_e = 10^{-6}$ .



(a)

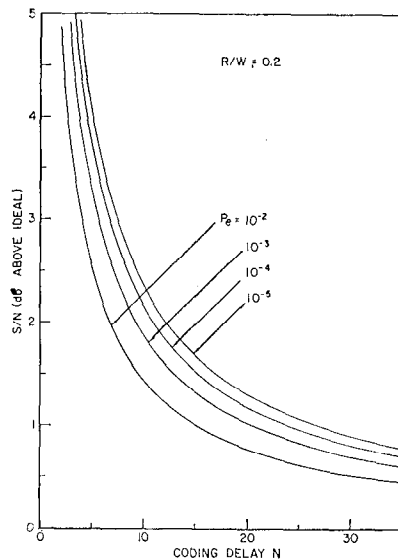


(b)

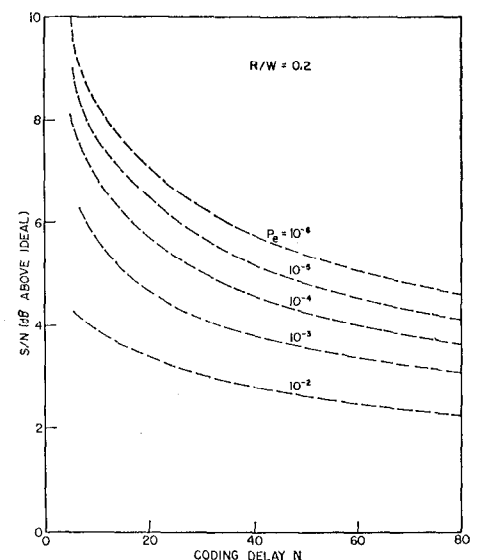


(c)

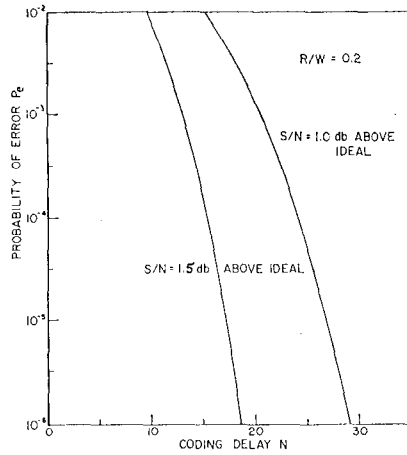
Fig. 3. The additional signal-to-noise ratio as a function of the coding delay for different values of the probability of error. (a) BL coding scheme. (b) Bounds on one-way communication.



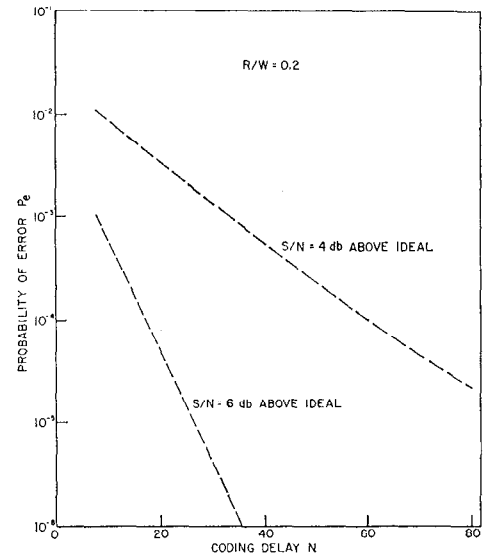
(a)



(b)



(a)



(b)

Fig. 4. The probability of error as a function of the coding delay for different values of the relative signal-to-noise ratio. (a) BL coding scheme. (b) Bounds on one-way communication.

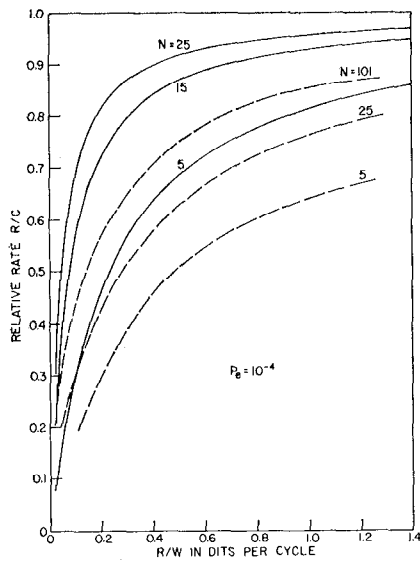
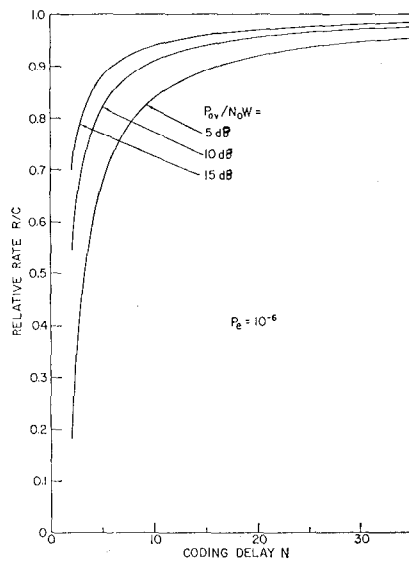
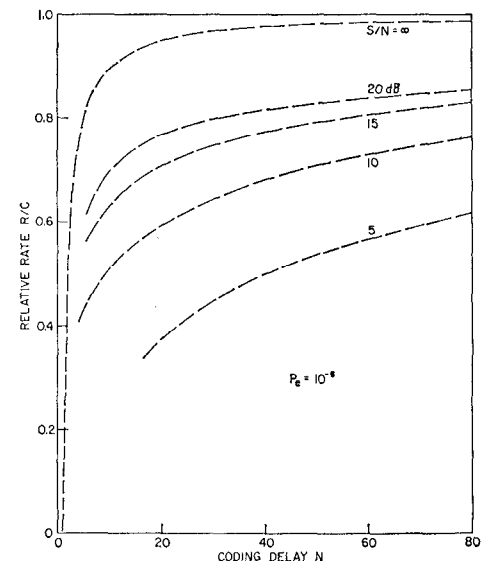


Fig. 5. The relative rate vs. the rate per unit bandwidth for different values of the coding delay.

Fig. 6. The relative rate vs. the coding delay for different values of the signal-to-noise ratio. (a) BL coding scheme. (b) Bounds on one-way communication.



(a)



(b)

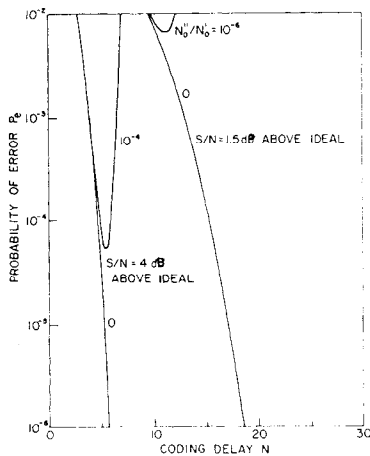


Fig. 7. The probability of error as a function of the coding delay for different values of the relative signal-to-noise ratio in the presence of feedback noise.

and the transmitted power is proportional to the square of the difference, the expected transmitted power per iteration decreases in this scheme. Retaining the maximum likelihood property but making up for the transmitted power in order to make the expected power per trans-

mission a constant, leads to the BL (band-limited) coding schemes. This simple scheme is the first deterministic procedure to achieve the channel capacity,  $W \ln [1 + (P_{av}/N_0W)]$ , of the band-limited white Gaussian noise channel.

It is believed that this approach of recursive maximum likelihood estimation to the coding problem with feedback has a much wider area of application, for example, channels with unknown parameters, fading channels, dependences between the noises in forward and feedback links, and so on. The method is ideally suited for noiseless feedback and it may well be possible to find an extension that is in some sense optimum for the noisy feedback case.

#### REFERENCES

- [1] C. E. Shannon, "A mathematical theory of communication," *Bell Sys. Tech. J.*, vol. 27, pp. 379-424 and 623-657, July-October 1948.
- [2] —, "Probability of error for optimal codes in a Gaussian channel," *Bell Sys. Tech. J.*, vol. 38, pp. 611-656, May 1959.
- [3] P. Elias, "Channel capacity without coding," in *Lectures on Communication System Theory*, Baghdady, Ed. New York: McGraw-Hill, 1961.
- [4] D. Slepian, "Bounds on communication," *Bell Sys. Tech. J.*, vol. 42, pp. 681-707, May 1963.
- [5] J. P. M. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback—Part I: no bandwidth constraint," this issue, page 172.

## Coding for a Class of Unknown Channels

IRVIN G. STIGLITZ, MEMBER, IEEE

**Abstract**—A channel which is selected for each use (without knowledge of past history) to be one of a given set of discrete memoryless channels is to be used by an ignorant communicator, i.e., the transmitter and receiver are assumed to have no knowledge of the particular channels selected. For this situation an upper bound on the insurable average error probability for block codes of length  $n$  is obtained which exponentially approaches zero for all rates less than capacity. Communication design techniques for achieving these results are discussed.

FOR OUR PURPOSES, a statistically describable channel model is one for which the statistics of the channel output are known for each possible channel input. For a nonstatistically describable channel model, this statement does not hold.

The problem of designing communication systems for statistically describable channel models has been widely investigated. C. E. Shannon showed, for a large class of such models, that information can be transmitted over such a channel with arbitrarily small error probability for any rate less than a maximum rate called capacity.

The author is with the Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Mass. (Operated with support from the U. S. Air Force.)

For statistically describable discrete memoryless channel models, the minimum error probability,  $P_e$ , achievable with a block code of length  $n$ , has been overbounded for rates  $R$  less than capacity by [1], [2]

$$P_e \leq e^{-nE_1(R)} \quad (1)$$

where  $E_1(R)$  is a function of the channel statistics and is a positive convex downward<sup>1</sup> function of  $R$  for  $R$  less than capacity.

The evaluation of capacity for a large class of nonstatistically describable channel models has been investigated by Blackwell, Breiman, and Thomasian [3], [4] and by Wolfowitz [5]. We investigate in the sequel the following class of discrete nonstatistically describable channel models. For each use the channel is selected in a fashion unknown to the transmitter or receiver to be one of a given fixed set of statistically describable channels. The channel selection mechanism is permitted to change from use to use; however, channels are assumed to be

<sup>1</sup> A function  $f(z)$  is said to be convex downward if every cord lies on or above the function. If  $f(z)$  is convex downward then  $-f(z)$  is said to be convex upward.