

Random Coding for Additive Gaussian Channels with Feedback

LAWRENCE H. OZAROW, MEMBER, IEEE

Abstract—A random coding strategy for discrete-time additive Gaussian noise channels with feedback is analyzed. It has long been known that feedback may increase the capacity of such channels as long as the additive noise process is not white. We prove that a strictly positive gain is always achieved. In addition, we prove that as the signal power goes to zero the ratio of feedback capacity to capacity without feedback may be strictly greater than one if the noise spectrum has a null. This is not the case when the noise spectrum is bounded away from zero. We also demonstrate that random coding, where the codewords are chosen from an ensemble of stationary Gaussian sequences, does not achieve capacity.

I. INTRODUCTION

WHILE it has always been clear that feedback may increase capacity when the noise process is not white, the extent to which it can do so remains only partially explored. Butman [1] outlined a deterministic strategy, which he was able to evaluate only for first-order Markov noise, and demonstrated a strict increase in capacity. Schalkwijk and Tiernan [2] obtained upper bounds for the first-order Markov case, and Ozarow [3] has obtained bounds for the general case. Ozarow's result applied to the Markov case yields upper bounds that are quite close to Butman's lower bounds.

In the present work we analyze a random coding strategy for discrete-time additive Gaussian noise channels with feedback and obtain an expression for the power required as a function of rate. We have successfully evaluated this quantity for first-order Markov noise and for first-order moving average noise. Our result may be expressed as follows: let the power spectral density of the noise be represented as

$$S_n(\theta) = \hat{\sigma}_n^2 |P(e^{i\theta})|^2$$

where $\hat{\sigma}_n^2$ is the entropy power of the process $\{n\}$, and

$$P(e^{i\theta}) = \sum_0^\infty p_k e^{i\theta k}$$

where $p_0 = 1$ and $P(z)$ (the analytical extension of $P(e^{i\theta})$ into the unit disk) has no zeros inside the unit disk. Let $Q(e^{i\theta})$ be a polynomial:

$$Q(e^{i\theta}) = \sum_0^M q_k e^{i\theta k}$$

such that $Q(z)$ ($Q(e^{i\theta})$ extended to the interior of the unit disk) satisfies

$$Q(0) = e^{-R}$$

$$|Q(z)|^2 \leq 1$$

and

$$Q(z) \neq 0, \quad \text{for } |z| < 1.$$

Then let

$$\gamma = \frac{P}{\hat{\sigma}_n^2}$$

be the ratio of average signal power to the noise process entropy power $\hat{\sigma}_n^2$, and denote by $\{f(e^{i\theta})\}_+$ the function derived from $f(e^{i\theta})$ by discarding negative powers of its Fourier series expansion. We show that

$$\gamma = \frac{\sigma_n^2}{\hat{\sigma}_n^2} + \left| e^R - \frac{1}{2\pi} \int_{-\pi}^{\pi} P(e^{i\theta}) Q(e^{-i\theta}) d\theta \right|^2 - \frac{1}{2\pi} \int_{-\pi}^{\pi} |\{P(e^{i\theta}) Q(e^{-i\theta})\}_+|^2 d\theta \quad (1)$$

is achievable.

While (1) is not particularly enlightening, it does allow fairly straightforward solutions in at least two cases: first-order Markov noise (as in Butman [1]) and noise whose spectrum is given by

$$S_n(\theta) = \hat{\sigma}_n^2 |1 + \rho e^{i\theta}|^2,$$

i.e., a first-order moving average process.

Our results demonstrate a number of interesting facts. First, we show that, provided the noise process is not white, capacity is strictly increased. Second, we can show by use of (1) applied to the moving average channel that cases exist for which the following holds. Denote the power required to achieve rate R without feedback by $P_{NF}(R)$ and the power with feedback by $P_F(R)$; then

$$\lim_{R \rightarrow 0} \frac{P_F(R)}{P_{NF}(R)} = l < 1.$$

This behavior is exhibited when $S_n(\theta)$ is zero for some value of θ . Conversely, it can be shown that if $S_n(\theta) \geq \delta > 0$ for all θ , then

$$\lim_{R \rightarrow 0} \frac{P_F(R)}{P_{NF}(R)} = 1,$$

Manuscript received April 7, 1989.
The author is with AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, NJ 07974.
IEEE Log Number 8932716.

i.e., that feedback shows no gain asymptotically as R (or P) goes to zero.

In addition our results demonstrate that the intuition obtained from channels without feedback is contradicted when feedback is available. The random codes we derive in the next section pile all of their energy on the point at which the noise spectrum is at its minimum, quite unlike the case without feedback, where the signal is distributed as evenly as possible.

Yet another interesting corollary of what follows is that random coding cannot achieve capacity for the feedback channel. In [1] Butman demonstrated a deterministic scheme similar to that introduced in [5] for the white channel M th feedback. For the first-order Markov case, the only one he was able to compute, he achieves a rate strictly greater than ours.

Our scheme is optimum among codes with codewords selected from the traditional random ensemble, with the addition of linear feedback. Since Ihara [6] has shown that Gaussian signals with linear feedback maximize mutual information, we are led to conclude that the best codes are either of the deterministic type employed in [1] or with codewords drawn from some hitherto—unexamined ensemble.

II. RESULTS

The coding scheme we will use is as follows. Let $\{W^j\}_{j=1}^{2^{NR}}$ be a codeword set whose code-words are drawn from a stationary Gaussian distribution with power spectral density $S_w(\theta)$. Provided

$$R < \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{2} \log \frac{S_w(\theta) + S_n(\theta)}{S_n(\theta)} d\theta, \quad (2)$$

reliable communication is possible when the codewords are transmitted over a one-way additive Gaussian channel whose noise samples $\{n_j\}$ have power spectral density $S_n(\theta)$.

Now let $\{\alpha_j\}_{j=1}^{\infty}$ be an arbitrary sequence, and at time i transmit the variable

$$x_i = w_i + \sum_1^{\infty} \alpha_j w_{i-j} + \sum_1^{\infty} \alpha_j n_{i-j}$$

where w_k and n_k are taken to be zero for $k < 0$. Of course, since we have assumed that feedback is available, then n_j is known at the transmitter for all $j < i$. The received datum at time i is

$$y_i = w_i + \sum_1^{\infty} \alpha_j w_{i-j} + \sum_1^{\infty} \alpha_j n_{i-j} + n_i.$$

Define

$$v^j = \left. \begin{array}{c} 0 \\ 0 \\ 0 \\ 1 \\ \alpha_1 \\ \vdots \\ \alpha_{N-j} \end{array} \right\} j-1 \text{ terms,}$$

i.e., v^j is the N vector with $j-1$ leading zeros followed by a unity element followed by $(\alpha_1, \alpha_2, \dots, \alpha_{N-j})$. Then the N -vector of received variables is

$$y = A(x + n)$$

where A is a matrix in which the j th column (starting with 0) is v^j .

Since A is subdiagonal and has ones on the diagonal, it is nonsingular and may be inverted. Premultiplying y by A^{-1} yields

$$y' = x + n,$$

and if (1) is satisfied, reliable communication is possible.

One would expect that the best transmitted datum, from the point of view of average power at time i , would be

$$x_i = w_i - E[w_i | Y_{i-1}, Y_{i-2}, \dots, Y_0].$$

However, if we follow this scheme, then $w_i + n_i$ is a causal linear function of the $\{Y_i\}$, so that the reverse is also true: Y_i is a causal linear function of the $\{w_i + n_i\}$. Therefore, $E[w_i | Y_{i-1}, Y_{i-2}, \dots, Y_0] = E[w_i | w_{i-1} + n_{i-1}, \dots, w_0 + n_0]$, which is the representation we shall use. Thus the performance we achieve is the best possible using random codes.

The power required on the i th transmission is

$$P_i = E(X_i^2) = \sigma_w^2 + \alpha_i^T M_{w+n} \alpha_i + 2\alpha_i^T R_w \quad (3)$$

where σ_w^2 is the variance of w_i , α_i is the i -vector with elements $\alpha_1, \alpha_2, \dots, \alpha_i$, R_w is the autocorrelation vector of process $\{w\}$ omitting the $R(0)$ term, and M_{w+n} is the autocorrelation matrix of the stationary process $\{w + n\}$.

We now digress to prove the following.

Corollary 1: Feedback strictly increases capacity when the noise process is not white.

Equations (2) and (3) (with $\alpha_i = 0$) are exactly the parametric expressions for capacity in the absence of feedback. Provided the noise process is not white, the optimizing $S_w(\theta)$ in the absence of feedback is not flat. Hence some j exists for which $R_w(j)$ is not zero. Therefore, by picking

$$\alpha_j = \frac{-R_w(j)}{\sigma_w^2 + \sigma_n^2},$$

and all other $\alpha_i = 0$, we achieve the same rate as without feedback, but using

$$P' = \sigma_w^2 - \frac{R_w^2(j)}{\sigma_w^2 + \sigma_n^2},$$

strictly less than the power required without feedback. Since (nonfeedback) capacity is strictly monotonically increasing in power, capacity as a function of P is strictly increased.

Returning to the main theorem, if we define

$$r_w(e^{i\theta}) = \sum_1^{\infty} R_w(j) e^{i\theta(j-1)}$$

and

$$\alpha_k(e^{i\theta}) = \sum_1^k \alpha_j e^{i\theta(j-1)},$$

then

$$E(x_k^2) = \sigma_w^2 + \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[|a_k(e^{i\theta})|^2 S_{w+n}(\theta) + 2a_k^*(e^{i\theta}) r_w(e^{i\theta}) \right] d\theta. \quad (4)$$

Let us choose $S_w(\theta)$ so that

$$\begin{aligned} S_{w+n}(\theta) &\triangleq S_w(\theta) + S_n(\theta) \\ &= \frac{S_n(\theta)}{|Q(e^{i\theta})|^2} \end{aligned}$$

where $Q(\cdot)$ is as defined in the introduction.

The last integrated term in (4) may be written as

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} a_k^*(e^{i\theta}) r_w(e^{i\theta}) d\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} a_k^*(e^{i\theta}) S_w(\theta) e^{-i\theta} d\theta$$

because

$$S_w(\theta) = \sum_{j=-\infty}^{\infty} R_w(j) e^{i\theta j}$$

and negative powers of $e^{i\theta}$ will not contribute to the integral. Since

$$S_w(\theta) = S_{w+n}(\theta) - S_n(\theta),$$

we may write

$$P_k = \sigma_w^2 + \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[|a_k(e^{i\theta})|^2 S_{w+n}(\theta) + 2a_k^*(e^{i\theta}) (S_{w+n}(\theta) - S_n(\theta)) e^{-i\theta} \right] d\theta.$$

Factoring $S_{w+n}(\theta)$ as

$$S_{w+n}(\theta) = |S_{w+n}^+(e^{i\theta})|^2,$$

we can write

$$\begin{aligned} P_k &= \sigma_w^2 + \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[|a_k(e^{i\theta}) S_{w+n}^+(e^{i\theta})|^2 + 2a_k^*(e^{i\theta}) S_{w+n}^-(e^{i\theta}) \left(\frac{S_{w+n}(\theta) - S_n(\theta)}{S_{w+n}^-(e^{i\theta})} \right) e^{-i\theta} \right] d\theta \\ &= \sigma_w^2 + \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ |a_k(e^{i\theta}) S_{w+n}^+(e^{i\theta})|^2 + 2a_k^*(e^{i\theta}) S_{w+n}^-(e^{i\theta}) (S_{w+n}^+(e^{i\theta}) - S_n^+(e^{i\theta}) Q^*(e^{i\theta})) e^{-i\theta} \right\} d\theta \end{aligned}$$

where

$$S_{w+n}^-(e^{i\theta}) = (S_{w+n}^+(e^{i\theta}))^*.$$

Defining

$$g(e^{i\theta}) = (S_{w+n}^+(e^{i\theta}) - S_n^+(e^{i\theta}) Q^*(e^{i\theta})) e^{-i\theta},$$

we get

$$P_k = \sigma_w^2 + \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[|a_k(e^{i\theta}) S_{w+n}^+(e^{i\theta})|^2 + 2a_k^*(e^{i\theta}) S_{w+n}^-(e^{i\theta}) \{g(e^{i\theta})\}_+ \right] d\theta.$$

Completing the square in the integral yields

$$\begin{aligned} P_k &= \sigma_w^2 - \frac{1}{2\pi} \int_{-\pi}^{\pi} |\{g(e^{i\theta})\}_+|^2 d\theta \\ &\quad + \frac{1}{2\pi} \int_{-\pi}^{\pi} |a_k(e^{i\theta}) S_{w+n}^+(e^{i\theta}) + \{g(e^{i\theta})\}_+|^2 d\theta. \end{aligned}$$

Since the coefficients of $a_k(e^{i\theta})$ may be chosen freely, we pick them to cancel the first k terms in $\{g(e^{i\theta})\}_+$. If $\{g\}_+$ is in L_2 , then as $k \rightarrow \infty$ the last term may be driven to zero. Now

$$\begin{aligned} &\frac{1}{2\pi} \int_{-\pi}^{\pi} |\{g(e^{i\theta})\}_+|^2 d\theta \\ &\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |g(e^{i\theta})|^2 d\theta \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} [S_{w+n}(\theta) + S_n(\theta) |Q(e^{i\theta})|^2 - 2S_n(\theta)] d\theta. \end{aligned}$$

Since $|Q|^2 \leq 1$ and $S_n(\theta) \geq 0$, we see that

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} |\{g(e^{i\theta})\}_+|^2 &\leq \frac{1}{2\pi} \int_{-\pi}^{\pi} S_w(\theta) d\theta \\ &< \infty, \end{aligned}$$

so T_k may be driven to zero. The average power may therefore be written as

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \sum E(x_k^2) &\triangleq P = \sigma_w^2 - \frac{1}{2\pi} \int_{-\pi}^{\pi} |\{(S_{w+n}^+(e^{i\theta}) - S_n^+(e^{i\theta}) Q^+(e^{i\theta})) e^{-i\theta}\}_+|^2 d\theta. \end{aligned}$$

Noting that

$$\{F(e^{i\theta}) e^{-i\theta}\}_+ = \{F(e^{i\theta}) - F_0\}_+ e^{-i\theta}$$

where

$$F_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(e^{i\theta}) d\theta,$$

we have

$$\begin{aligned} &\frac{1}{2\pi} \int_{-\pi}^{\pi} |\{F(e^{i\theta}) e^{-i\theta}\}_+|^2 \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |\{F(e^{i\theta})\}_+|^2 d\theta - F_0^2, \end{aligned}$$

so

$$\begin{aligned} P &= \sigma_w^2 - \frac{1}{2\pi} \int_{-\pi}^{\pi} |\{S_{w+n}^+(e^{i\theta}) - S_n^+(e^{i\theta}) Q^*(e^{i\theta})\}_+|^2 d\theta \\ &\quad + \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} S_{w+n}^+(e^{i\theta}) d\theta - \frac{1}{2\pi} \int_{-\pi}^{\pi} S_n^+(e^{i\theta}) Q^*(e^{i\theta}) d\theta \right)^2. \end{aligned}$$

Since $|S_{w+n}^+(e^{i\theta})|^2$ is the power spectral density of $w+n$, we have

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} S_{w+n}^+(e^{i\theta}) d\theta = \hat{\sigma}_{w+n},$$

which from (2) is $\hat{\sigma}_n e^R$. Therefore,

$$P = \sigma_w^2 - \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(S_{w+n}(\theta) + |\{S_n^+ Q^*(e^{i\theta})\}_+|^2 - 2S_{w+n}^+(e^{i\theta})S_n^-(e^{i\theta})Q(e^{i\theta}) \right) d\theta + \left(\hat{\sigma}_n e^R - \frac{1}{2\pi} \int_{-\pi}^{\pi} S_n^+(e^{i\theta})Q^+(e^{i\theta}) d\theta \right)^2.$$

Noting that

$$S_{w+n}^+(e^{i\theta})S_n^-(e^{i\theta})Q(e^{i\theta}) = S_n(\theta),$$

we deduce that

$$P = \sigma_w^2 - \sigma_w^2 - \sigma_n^2 + 2\sigma_n^2 + \left(\hat{\sigma}_n e^R - \frac{1}{2\pi} \int_{-\pi}^{\pi} S_n^+(e^{i\theta})Q^*(e^{i\theta}) d\theta \right)^2 - \frac{1}{2\pi} \int_{-\pi}^{\pi} |\{S_n^+ Q^*(e^{i\theta})\}_+|^2 d\theta.$$

Dividing by $\hat{\sigma}_n^2$ and recalling that

$$S_n(\theta) = \hat{\sigma}_n^2 |P(e^{i\theta})|^2,$$

we have

$$\frac{P}{\hat{\sigma}_n^2} = \frac{\sigma_w^2}{\hat{\sigma}_n^2} + \left(e^R - \frac{1}{2\pi} \int_{-\pi}^{\pi} P(e^{i\theta})Q^*(e^{i\theta}) d\theta \right)^2 - \frac{1}{2\pi} \int_{-\pi}^{\pi} |\{P(e^{i\theta})Q^*(e^{i\theta})\}_+|^2 d\theta$$

which is (1). In Sections III and IV we examine the first-order Markov and moving average noise models in detail.

III. MARKOV NOISE

We evaluate (1) for the case

$$S_n(\theta) = \hat{\sigma}_n^2 |P(e^{i\theta})|^2 = \frac{\sigma_n^2(1-\rho^2)}{|1-\rho e^{i\theta}|^2}.$$

If $Q(e^{i\theta})$ has the Fourier series representation

$$Q(e^{i\theta}) = \sum_0^{\infty} q_k e^{i\theta k},$$

then

$$\begin{aligned} \{S_n^+(e^{i\theta})Q^*\}_+ &= \hat{\sigma}_n \left\{ \frac{Q^*(e^{i\theta})}{1-\rho e^{i\theta}} \right\}_+ \\ &= \hat{\sigma}_n \sum_0^{\infty} \frac{\rho^k q_k}{1-\rho e^{i\theta}} = S_n^+(e^{i\theta})Q(\rho) \end{aligned}$$

where $Q(\rho)$ is the value of $Q(z)$, the analytic continuation of $Q(e^{i\theta})$ into the unit disk, evaluated at $z = \rho$. Since

$$\int S_n^+(e^{i\theta})Q^*(e^{i\theta}) d\theta = \int \{S_n^+ Q^*\}_+ d\theta$$

and

$$\frac{1}{2\pi} \int |S_n^+(e^{i\theta})|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_n(\theta) d\theta = \sigma_n^2,$$

(1) becomes

$$\gamma = \frac{1}{1-\rho^2} + (e^R - Q(\rho))^2 - \frac{1}{1-\rho^2} Q^2(\rho)$$

and average power is minimized by maximizing $Q(\rho)$. Now the analytic function $Q(z)$ satisfies

$$\begin{aligned} Q(0) &= e^{-R} \\ |Q(z)|^2 &\leq 1, \quad |z|^2 \leq 1. \end{aligned} \quad (5)$$

Since we have required that $Q(z)$ have no zeros, then

$$q(z) \triangleq \log Q(z)$$

is analytic on the disk, and (3) implies

$$q(0) = -R$$

$$\operatorname{Re} q(z) \leq 0.$$

If we define

$$w(z) = \frac{q(z) - q(0)}{q(z) + q(0)},$$

then $w(z)$ is analytic on the disk and

$$w(0) = 0$$

$$|w(z)|^2 \leq 1 \text{ for all } z \text{ such that } |z|^2 \leq 1,$$

so by Schwarz's lemma [4]

$$|w(\rho)| \leq |\rho|,$$

which implies that

$$Q(\rho) \leq \exp\left(-R \frac{1-|\rho|}{1+|\rho|}\right).$$

By choosing

$$Q(z) = \exp\left[-R \frac{1 - S \operatorname{sgr}(\rho)z}{1 + S \operatorname{sgn}(\rho)z}\right],$$

we can achieve

$$Q(\rho) = \exp\left[-R \left(\frac{1 - S|\rho|}{1 + S|\rho|}\right)\right]$$

for S arbitrarily close to one. Therefore, power arbitrarily close to

$$\gamma = \frac{1}{1-\rho^2} + \left(e^R - \exp\left(-R \frac{1-|\rho|}{1+|\rho|}\right) \right)^2 - \frac{1}{1-\rho^2} \exp\left(-2R \frac{1-|\rho|}{1+|\rho|}\right)$$

is achievable. In Table I we compare some achievable rates

TABLE I
ACHIEVABLE RATES AND UPPER BOUNDS FOR MARKOV CHANNELS

P	C_{NF}	Achievable		Upper Bound (Ozarow)
		Butman	Ozarow	
		$\rho = 0.1$		
0.1	0.1033	0.1224	0.1223	0.1880
1.0	0.4860	0.5792	0.5757	0.6074
10.0	1.3428	1.4365	1.4254	1.4415
100.0	2.4514	2.4905	2.4797	2.4915
1000.0	3.5982	3.6115	3.6055	3.6118
		$\rho = 0.9$		
0.1	0.3629	0.4229	0.4213	0.5390
1.0	1.0521	1.1380	1.1325	1.1960
10.0	2.0280	2.0944	2.0892	2.1211
100.0	3.1379	3.1709	3.1678	3.1810
1000.0	4.2847	4.2965	4.2945	4.3002

with those of Butman [1] and upper bounds on capacity derived by Ozarow.

IV. MOVING-AVERAGE NOISE

We now evaluate (1) when

$$\begin{aligned} S_n(\theta) &= \hat{\sigma}_n^2 |P(e^{i\theta})|^2 \\ &= \frac{\sigma_n^2}{1 + \rho^2} |1 + \rho e^{i\theta}|^2. \end{aligned}$$

As before, if

$$Q(e^{i\theta}) = \sum_0^{\infty} q_k e^{i\theta k},$$

then

$$\begin{aligned} \{S_n^+(e^{i\theta})Q^*(e^{i\theta})\}_+ &= \hat{\sigma}_n^2 (q_0 + \rho(q_0 e^{i\theta} + q_1)) \\ &= \hat{\sigma}_n^2 (q_0 + \rho q_1 + \rho q_0 e^{i\theta}). \end{aligned}$$

Therefore,

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} S_n^+(e^{i\theta})Q^*(e^{i\theta}) d\theta = \hat{\sigma}_n^2 (q_0 + \rho q_1)$$

and

$$\frac{1}{2\pi} \int |\{S_n^+(e^{i\theta})Q^*(e^{i\theta})\}_+|^2 d\theta = \hat{\sigma}_n^2 ((q_0 + \rho q_1)^2 + \rho^2 q_0^2).$$

Since by hypothesis

$$q_0 = e^{-R},$$

(1) becomes

$$\begin{aligned} \gamma &= \sigma_n^2 + (e^R - (q_0 + \rho q_1))^2 - (q_0 + \rho q_1)^2 - \rho^2 q_0^2 \\ &= \sigma_n^2 + e^{2R} - 2e^R (q_0 + \rho q_1) - \rho^2 q_0^2 \\ &= \frac{\sigma_n^2}{\hat{\sigma}_n^2} + e^{2R} - 2 - \rho^2 e^{-2R} - 2\rho e^R q_1. \end{aligned}$$

Again, we can apply Schwarz's lemma to

$$w(z) = \frac{\log Q(z) - \log Q(0)}{\log Q(z) + \log Q(0)}$$

to get $|w'(z)| \leq 1$, which implies that

$$|Q'(0)| = |q_1| \leq 2R e^{-R}.$$

Again by choosing

$$Q(z) = \exp\left(-R \frac{1 - S_{\text{sgn}(\rho)}(\rho)z}{1 + S_{\text{sgn}(\rho)}(\rho)z}\right)$$

we can achieve

$$q_1 = 2R \text{sgn}(\rho) e^{-R},$$

so that

$$\gamma = \frac{\sigma_n^2}{\hat{\sigma}_n^2} - 2 + e^{2R} - \rho^2 e^{-2R} - 4|\rho|SR \quad (6)$$

can be achieved for any $S < 1$.

As a special case consider $\rho = 1$, $\gamma \rightarrow 0$. Equation (4) becomes

$$\gamma = e^{2R} - e^{-2R} - 4SR.$$

If γ is small, so is R , so asymptotically

$$\begin{aligned} \gamma &= \left(1 + 2R + 4R^2 + \frac{4}{3}R^3 \dots\right) \\ &\quad - \left(1 - 2R + 4R^2 - \frac{4}{3}R^3 \dots\right) - 2SR \\ &= 4(1 - S)R + \frac{8}{3}R^3. \end{aligned}$$

The solution to this equation is continuous in S and may be driven arbitrarily close to

$$R = \left[\frac{3\gamma}{8}\right]^{1/3}.$$

Since $\hat{\sigma}_n^2 = \sigma_n^2/2$, then rates close to

$$R' = \left[\frac{3P}{4\sigma_n^2}\right]^{1/3}$$

can be attained.

On the other hand, without feedback, if we assume for convenience that

$$S_n(\theta) = \sigma_n^2 (1 - \cos \theta),$$

then for small power all the signal power will be distributed near the origin, where

$$S_n(\theta) \sim \sigma_n^2 \frac{\theta^2}{2}.$$

We can solve the parametric expressions for capacity without feedback given by

$$\begin{aligned} P &= \frac{1}{2\pi} \int_{-\theta'}^{\theta'} (K - S_n(\theta)) d\theta \\ R &= \frac{1}{2\pi} \int_{-\theta'}^{\theta'} \frac{1}{2} \log \frac{K}{S_n(\theta)} d\theta \end{aligned}$$

where

$$S_n(\theta') = S_n(-\theta') = K,$$

to obtain

$$C_{NF} = \left[\frac{3}{\pi^2} \frac{P}{\sigma_n^2} \right]^{1/3}.$$

Therefore, as $P/\sigma_n^2 \rightarrow 0$ the use of feedback achieves a gain of $[\pi^2/4]^{1/3}$ relative to the case of no feedback. This asymptotic gain cannot occur if $S_n(\theta)$ is bounded away from zero.

V. CONCLUSION

We have examined the performance of random coding for additive Gaussian channels with feedback and colored noise. We obtained a new achievable rate and established that feedback strictly increases the capacity of any non-white stationary Gaussian channel, and that random coding cannot achieve capacity. Also, we showed that if the noise power spectral density has a null, feedback can achieve an asymptotic gain as the power goes to zero, i.e.,

$$\lim_{P \rightarrow 0} \frac{C_{NF}}{C_F} < 1.$$

This is not the case if the power spectral density is bounded away from zero.

ACKNOWLEDGMENT

The author wishes to acknowledge helpful discussions with H. J. Landau.

REFERENCES

- [1] S. Butman, "A general formulation of linear feedback communication systems with solutions," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 392-400, May 1969.
- [2] J. P. Schalkwijk and Tiernan, "An upper bound to the capacity of band-limited Gaussian autoregressive channel with noiseless feedback," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 311-316, May 1974.
- [3] L. H. Ozarow, "Upper bounds on the capacity of Gaussian channels with feedback," this issue, pp. 156-161.
- [4] W. Rudin, *Real and Complex Analysis*. New York: McGraw Hill, 1974.
- [5] J. Schalkwijk, "A coding scheme for additive noise channels with feedback—Part II: Band-limited channels," *IEEE Trans. Inform. Theory*, vol. IT-12, pp. 183-189, Apr. 1966.
- [6] S. Ihara, "On the capacity of the discrete time Gaussian channel with feedback," in *Trans. 8th Prague Conf. Information Theory*, Czech Acad. Sci., 1979.