The Gaussian Channel with Noisy Feedback

Young-Han Kim
Department of ECE
University of California, San Diego
La Jolla, CA 92093, USA
yhk@ucsd.edu

Amos Lapidoth
Signal and Information Processing Lab.
ETH Zurich
8092 Zurich, Switzerland
lapidoth@isi.ee.ethz.ch

Tsachy Weissman
Information Systems Lab.
Stanford University
Stanford, CA 94305, USA
tsachy@stanford.edu

***Abstract*— Upper and lower bounds are derived on the reliability function of the additive white Gaussian noise channel with output fed back to the transmitter over an independent additive white Gaussian noise channel. Special attention is paid to the regime of very low feedback noise variance and it is shown that the reliability function is asymptotically inversely proportional to the feedback noise variance. This result shows that the noise in the feedback link, however small, renders the commnication with noisy feedback fundamentally different from the perfect feedback case. For example, it is demonstrated that with noisy feedback, linear coding schemes fail to achieve any positive rate. In contrast, an asymptotically optimal coding scheme is devised, based on a three-phase detection/retransmission protocol, which achieves an error exponent inversely proportional to the feedback noise variance for any rate less than capacity.**

I. INTRODUCTION

This paper studies the discrete-time additive white Gaussian noise channel with a noisy feedback link depicted in Fig. 1. We

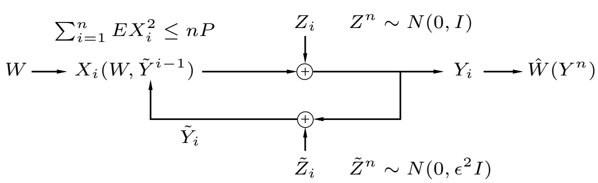

Fig. 1. Gaussian channel with noisy feedback.

wish to communicate a message index $W \in \{1, 2, \dots, e^{nR}\}$ over the *forward* additive white Gaussian noise channel

$$Y_i = X_i + Z_i, \tag{1}$$

where X_i, Z_i, Y_i, respectively, denote the channel input, the additive Gaussian noise, and the channel output. Let further $\tilde{Y}_i$ denote a noisy version of Y_i over the *feedback (backward)* additive white Gaussian noise channel

$$\tilde{Y}_i = Y_i + \tilde{Z}_i, \tag{2}$$

where $\tilde{Z}_i$ is the Gaussian noise in the backward link. We assume that the forward noise process $\{Z_i\}_{i=1}^\infty$ and the backward noise process $\{\tilde{Z}_i\}_{i=1}^\infty$ are independent of each other, and respectively independent and identically distributed according to $N(0, 1)$ and $N(0, \epsilon^2)$.

We define an (e^{nR}, n) code with the encoding functions

$$f_i^{(n)} : (W, \tilde{Y}^{i-1}) \mapsto X_i, \qquad i = 1, 2, \dots, n, \tag{3}$$

satisfying the expected average power constraint

$$\frac{1}{n} \sum_{i=1}^{n} EX_i^2(W; \tilde{Y}^{i-1}) \le P, \tag{4}$$

and the decoding function $\phi^{(n)} : Y^n \mapsto \hat{W}$. Thus, the encoder has the causal access to the noisy feedback $\tilde{Y}^n$.

The probability of error $P_e^{(n)}$ is defined by

$$\begin{aligned} P_e^{(n)} &= \Pr(W \ne \hat{W}(Y^n)) \\ &= \frac{1}{e^{nR}} \sum_{w=1}^{e^{nR}} \Pr(W \ne \hat{W}(Y^n) | W = w), \end{aligned}$$

where W is distributed uniformly over $\{1, \dots, e^{nR}\}$ and is independent of $(Z^n, \tilde{Z}^n)$.

A rate R (nats per transmission) is said to be achievable if there exists a sequence of (e^{nR}, n) codes such that the associated probability of error $P_e^{(n)}$ tends to 0 as $n \to \infty$. The feedback capacity C_{FB} is defined as the supremum of all achievable rates.

Because the forward channel (1) is memoryless, the availability of feedback—even if noiseless (i.e., $\epsilon^2 = 0$)—does not increase the capacity as shown by Shannon [15]. Thus, we focus on the *reliability function* $E_{\text{FB}}(R; P, \epsilon^2)$ of the Gaussian channel with noisy feedback as a function of transmission rate R, power constraint P, and feedback noise variance ϵ^2.

The reliability function $E_{\text{FB}}(R; P, \epsilon^2)$ is defined as the rate of decay for the error probability of the optimal sequence of (e^{nR}, n) codes, i.e.,

$$E_{\text{FB}}(R; P, \epsilon^2) = \limsup_{n \to \infty} -\frac{1}{n} \log P_{e,\text{opt}}^{(n)}(R),$$

where $P_{e,\text{opt}}^{(n)}(R)$ denotes the infimum of the probability of error over all (e^{nR}, n) codes. We will use the notation $E(R; P)$ to denote the reliability of the additive white Gaussian noise channel *without* feedback under signal-to-noise ratio P.

While perfect feedback does not increase the capacity of a memoryless channel, it can improve the reliability dramatically. In their celebrated work [14], Schalkwijk and Kailath showed that a simple linear feedback coding scheme achieves the capacity. More surprisingly, the associated probability of error decays doubly exponentially. In terms of the reliability function, we have the infinite reliability, i.e., $E_{\text{FB}}(R; P, \epsilon^2 = 0) = \infty$. This fascinating result has been later

1-4244-1429-6/07/$25.00

extended in many directions. Pinsker [13], Kramer [9], and Zigangirov [17] showed that the probability of error can be made to decay as fast as an arbitrary level of nested exponentials. For nonwhite Gaussian channels, Butman [3] studied a variant of the Schalkwijk–Kailath coding scheme that achieves a rate higher than the nonfeedback capacity, which was later shown to be optimal [8]. The Schalkwijk–Kailath scheme has been also extended optimally to two-user Gaussian multiple access channels with feedback by Ozarow [12] and to Costa's writing on dirty paper channel [4] with feedback [11].

Much less explored is how the feedback noise affects the reliability function. There are a few papers in the literature on noisy feedback, including Kashyap [7], Lavenberg [10], and Draper and Sahai [6], but these papers have a fundamentally different nature from the current work. While Kashyap [7] focused on the reliability, his setup allowed the coding over the feedback channel with *exponentially* large power. Lavenberg [10] considered the Gaussian channel with infinite bandwidth and his focus was on orthogonal keying both in the forward channel and on the feedback channel. A recent work by Draper and Sahai [6] deals with the *variable-length* coding over discrete channels with noisy feedback under a different error criterion.

In comparison, this paper studies the block coding over the discrete-time Gaussian channel with uncoded feedback. The reason is clear to exclude variable-length coding schemes, in which the duration of communication is a random variable that depends on the channel behavior. Such schemes, which were studied by Burnashev [2] and others in the noiseless feedback case, turn out to be extremely fragile with noisy feedback. Due to the noise in the feedback channel, the transmitter and receiver may not be in agreement as to whether the communication has been concluded, and this may not only cause a decoding error of the present message, but also create havoc in all subsequent uses of the coding scheme.

In order to understand the effect of the feedback noise on the reliability function $E_{\text{FB}}(R;P,\epsilon^2)$, we focus especially on the regime $0<\epsilon^2\ll 1$, i.e., the feedback noise has a small, but nonzero variance. We will ultimately establish the following result.

Theorem 1. *For all rates $0<R<C(P)$ and all $\epsilon^2>0$, there exist constants $0<K_1(R,P), K_2(R,P)<\infty$ such that the reliability function $E_{\text{FB}}(R;P,\epsilon^2)$ is bounded by*

$$\frac{K_1(R,P)}{\epsilon^2} < E_{\text{FB}}(R;P,\epsilon^2) < \frac{K_2(R,P)}{\epsilon^2}. \tag{5}$$

To prove Theorem 1, we present upper and lower bounds on the reliability function. In Section II, three upper bounds on the reliability function are given. The first method, which is derived on a genie-aided setup, yields an upper bound on the noisy feedback reliability function $E_{\text{FB}}(R;P,\epsilon^2)$ in terms of the nonfeedback reliability function $E(R;P)$ at rate R and signal-to-noise ratio P:

$$E_{\text{FB}}(R;P,\epsilon^2) \le E\left(R;\frac{P(1+\epsilon^2)}{\epsilon^2}\right). \tag{6}$$

This bound suffices to demonstrate that very noisy feedback is like no feedback

$$\lim_{\epsilon^2\to\infty} E_{\text{FB}}(R;P,\epsilon^2) = E(R;P)$$

and that the reliability function $E_{\text{FB}}(R;P,\epsilon^2)$ is finite for every $\epsilon^2>0$. (On the other hand, recall that $E_{\text{FB}}(R;P,0)=\infty$.) In fact, (6) implies that

$$\overline{\lim_{\epsilon^2\to 0}}\, \epsilon^2 E_{\text{FB}}(R;P,\epsilon^2) < \infty,$$

or, written informally,

$$E_{\text{FB}}(R;P,\epsilon^2) = O(1/\epsilon^2). \tag{7}$$

The bound (6) is, however, rather loose in that it is positive even for $R>C$.

To remedy this deficiency we introduce a second upper bound

$$E_{\text{FB}}(R;P,\epsilon^2) \le f(R,P,\epsilon^2), \tag{8}$$

for all $C(P/(1+\epsilon^2))<R<C(P)$, where $C(P)$ is defined as $\frac{1}{2}\log(1+P)$ and $f(R,P,\epsilon^2)$ tends to zero as $R\to C(P)$. Thus, as a function of the rate R, $E_{\text{FB}}(R;P,\epsilon^2)$ is continuous at $R=C(P)$ for each $\epsilon^2>0$. (On the other hand, we have $E_{\text{FB}}(R;P,0)=\infty$ for all $R<C$.) The derivation of (8) is based on a change-of-measure argument, an application of Cramér's theorem from large deviations theory, and the strong converse for channels with perfect feedback at rates exceeding capacity.

Our third upper bound is also based on a change-of-measure argument. By changing the measure so that the feedback is useless, we can show that

$$E_{\text{FB}}(R;P,\epsilon^2) \le E(R;P/\sigma^2) + \gamma \tag{9}$$

for any γ,σ^2 satisfying $\Lambda^*_{\epsilon^2,\sigma^2}(\gamma) = E(R;P/\sigma^2)$, where $\Lambda^*_{\epsilon^2,\sigma^2}(\gamma)$ is the Fenchel–Legendre transform of a certain random variable and can be characterized analytically. This third upper bound can be easily generalized to arbitrary non-Gaussian channels.

These upper bounds show that the feedback noise, however small it is, renders the feedback communication fundamentally different from the perfect feedback case. This has important implications on the design of optimal codes. As a corollary of (6), in particular, $E_{\text{FB}}(R;P,\epsilon^2)<\infty$ for all $\epsilon^2>0$, we will prove in Section III that linear feedback coding schemes such as the Schalkwijk–Kailath coding scheme, which for $f=0$ achieves not only the capacity but the double-exponentially decaying probability of error, fail to achieve any positive rate.

Motivated by (7), we present in Section IV a coding scheme that achieves

$$\underline{\lim_{\epsilon^2\to 0}}\, \epsilon^2\Big(\overline{\lim_{n\to\infty}} -\frac{1}{n}\log P_e^{(n)}\Big) > 0, \quad R<C(P), \tag{10}$$

or written informally,

$$E_{\text{FB}}(R;P,\epsilon^2) = \Omega(1/\epsilon^2).$$

Combined with (7), this establishes Theorem 1.

The coding scheme that achieves (10) is based on an idea of simple detection/retransmission protocol and consists of three phases—1) nonfeedback communication, 2) acknowledgment, 3) retransmission. Each phase contains, however, a unique idea that hinges on the noise in the feedback link.

In the following sections, we provide some technical details on the upper and lower bounds introduced above.

II. Upper Bounds on the Reliability Function

In this section, we give three upper bounds on the reliability function $E_{\text{FB}}(R; P, \epsilon^2)$ introduced in Section I. Since $\tilde{Y}_i = X_i + Z_i + \tilde{Z}_i$, we can recast the encoding functions (3) in the form:

$$\tilde{f}_i^{(n)} : (W, S^{i-1}) \to X_i, \tag{11}$$

where $S_i = Z_i + \tilde{Z}_i$. Recall that $(Z_i, S_i), i = 1, 2, \ldots$ are independent and identically distributed (i.i.d.) with

$$Z_i \sim N(0,1), \quad S_i \sim N(Z_i, \epsilon^2). \tag{12}$$

Therefore, given an (e^{nR}, n) code, we can write

$$\Pr(W \neq \hat{W}(Y^n) | W = w) = \Pr((Z^n, S^n) \in A_w), \tag{13}$$

for each $w \in \{1, \ldots, e^{nR}\}$ with appropriately chosen sets $A_1, \ldots, A_{e^{nR}}$ that partition $\mathbb{R}^n \times \mathbb{R}^n$.

Under this notation, we start our discussion with the third upper bound (9).

A. Upper Bounds via Change of Measure

Let (Z, S) be a generic pair of random variables distributed as the pair (Z_i, S_i) in (12), and let (Z', S') be a pair of *independent* Gaussians with $Z' \sim N(0, \sigma^2)$ and $S' \sim N(0, 1+\epsilon^2)$. Let further f and f' denote the respective densities of (Z, S) and (Z', S'). Finally, let $\Lambda^*_{\epsilon^2,\sigma^2}$ denote the Fenchel–Legendre transform (see, e.g., [5]) of the random variable

$$\log\bigl(f'(Z', S')/f(Z', S')\bigr).$$

This function $\Lambda^*_{\epsilon^2,\sigma^2}(\gamma)$ can be characterized explicitly as follows. Let $\tilde{\Lambda}^*_{\epsilon^2,\sigma^2}$ be the Fenchel–Legendre transform of $U^2 \frac{\sigma^2+\sigma^2\epsilon^2-\epsilon^2}{\epsilon^2} - \frac{2\sigma\sqrt{1+\epsilon^2}}{\epsilon^2} UV + \frac{1}{\epsilon^2} V^2$ with independent standard Gaussians U, V, which can be written in closed form as

$$\begin{aligned}
&\tilde{\Lambda}^*_{\epsilon^2,\sigma^2}(\alpha) \\
&= \frac{1}{4}\Bigl(-2 - \alpha(1 + \sigma^2 + \epsilon^2(\sigma^2 - 1)) \\
&\qquad + \sqrt{4 + \alpha^2(\epsilon^4(\sigma^2-1)^2 + (\sigma^2+1)^2 + 2\epsilon^2(1+\sigma^4))}\Bigr) \\
&\quad + \frac{1}{2}\log \frac{-2 + \sqrt{4 + \alpha^2(\epsilon^4(1-\sigma^2)^2 + (1+\sigma^2)^2 + 2\epsilon^2(1+\sigma^4))}}{\alpha^2\epsilon^2}.
\end{aligned}$$

Then we have $\Lambda^*_{\epsilon^2,\sigma^2}(\gamma) = \tilde{\Lambda}^*_{\epsilon^2,\sigma^2}(2\gamma + \log[\sigma^2\epsilon^2(1+\epsilon^2)])$.

Proposition 1. *For each $\sigma^2 > 0$,*

$$E_{\text{FB}}(R; P, \epsilon^2) \leq \gamma + E\left(R; \frac{P}{\sigma^2}\right), \tag{14}$$

where $\gamma \geq D(f'||f)$ is the unique solution to

$$\Lambda^*_{\epsilon^2,\sigma^2}(\gamma) = E(R; P/\sigma^2)$$

and $D(f'||f)$ is the relative entropy between the densities f' and f.

Proof sketch.. Fix an (e^{nR}, n) code $(\{f_i^{(n)}\}, \phi^{(n)})$ and let $A_1, \ldots, A_{e^{nR}}$ be the associated error events defined as in (13). Let P denote the measure associated with the noises (Z_i, S_i) i.i.d. $\sim (Z, S)$ and let P' denote the measure associated with (Z_i', S_i') i.i.d. $\sim (Z', S')$. Suppose we use the same code $(\{f_i^{(n)}\}, \phi^{(n)})$ for the following communication scenario:

$$Y_i = X_i(W, S_1', \ldots, S_{i-1}') + Z_i'. \tag{15}$$

That is, the channel noises come from the measure P' instead of P. Then, because $\{Z_i'\}$ and $\{S_i'\}$ are *independent*, the feedback $\{S_i'\}$ is completely useless and the error exponent achieved in (15) is no better than the nonfeedback reliability function $E(R; P/\sigma^2)$. (Because $S_i \overset{d}{=} S_i'$, the power used under the new setting is identical to that used in the original setting.)

We define the set

$$B_\gamma = \left\{(z^n, s^n) : \frac{1}{n}\log\frac{f'(z^n, s^n)}{f(z^n, s^n)} \leq \gamma\right\}.$$

Then, we have

$$\begin{aligned}
P'(A_w) &\leq P'(A_w \cap B_\gamma) + P'(B_\gamma^c) \\
&= \int_{A_w \cap B_\gamma} f'(z^n, s^n) dz^n ds^n + P'(B_\gamma^c) \\
&\leq e^{n\gamma}\int_{A_w \cap B_\gamma} f(z^n, s^n) dz^n ds^n + P'(B_\gamma^c) \\
&= e^{n\gamma} P(A_w) + P'(B_\gamma^c).
\end{aligned}$$

When averaged over $w \in \{1, \ldots, e^{nR}\}$, this implies

$$P_e' \leq e^{n\gamma} P_e + P'(B_\gamma^c), \tag{16}$$

where P_e and P_e' denote the error probabilities under respective measures. We now take an arbitrary sequence of (e^{nR}, n) codes and choose γ such that

$$\Lambda^*_{\epsilon^2,\sigma^2}(\gamma) = \lim_{n\to\infty} -\frac{1}{n}\log P'(B_\gamma^c) > E(R; P/\sigma^2).$$

Since $E(R; P/\sigma^2) \geq \limsup -\frac{1}{n}\log P_e'^{(n)}$ by definition of the reliability function, (16) imples that

$$\begin{aligned}
\limsup_{n\to\infty} -\frac{1}{n}\log P_e^{(n)} &\leq \gamma - \frac{1}{n}\log\bigl(P'^{(n)}_{e,\text{ opt}} - P'(B_\gamma)\bigr) \\
&\leq \gamma + E(R; P/\sigma^2).
\end{aligned}$$

Finally the continuity and strict monotonicity of $\Lambda^*_{\epsilon^2,\sigma^2}(\gamma)$ in γ prove the desired result. □

Now we move on to the second upper bound (8). The proof technique is again based on a change-of-measure argument. This time, we define the new measure P' as

$$(Z', \tilde{Z}') \sim f_\delta' = N(0, 1+\delta) \times N(0, \epsilon^2 - \delta) \tag{17}$$

for $\delta \in (0, \epsilon^2)$, instead of the original measure

$$(Z, \tilde{Z}) \sim f = N(0,1) \times N(0, \epsilon^2). \tag{18}$$

Proposition 2. *Let $\delta \in (0, \epsilon^2)$ be given. Then, for all $R \in (C(P/(1+\delta)), C(P))$, we have*

$$E_{\text{FB}}(R; P, \epsilon^2) \leq D(f' || f), \tag{19}$$

where $D(f'_\delta || f)$ is the relative entropy between the densities f' and f defined in (17) *and* (18).

Proof sketch. Proceeding along the similar lines as the proof of Proposition 1, we can reach (16). (Again observe that $S' \overset{d}{=} S$ so that the power consumption stays the same.) Now when $R > C(P/(1+\delta))$, then by the strong converse of the coding theorem, $P_e'^{(n)} \to 1$ as $n \to \infty$. Thus we have $E_{\text{FB}}(R; P, \epsilon^2) \leq \gamma$, as long as B_γ^c is a large deviations event, namely, $\gamma > D(f'_\delta || f)$. □

By solving R for $R = C(P/(1+\delta))$, we can rewrite the upper bound (19) as a function of (R, P, ϵ^2), leading to (8). If $R \to C$ with ϵ^2 held fixed, or equivalently, $\delta \to 0$, then we have $D(f'_\delta || f) \to 0$. This proves that $E_{\text{FB}}(R; P, \epsilon^2) \to 0$ as $R \to C$, so the reliability function is continuous at $R = C$ for every $\epsilon^2 > 0$.

B. Upper Bound via Genie

Consider a genie-aided scheme in which the encoding functions are allowed to depend on the S_i sequence noncausally, that is, $X_i = f_i(W, S^n)$ instead of $f_i(W, S^{i-1})$. Assume further that the decoder is also given access to S^n in addition to Y^n, i.e., $\hat{W} = \phi(Y^n, S^n)$. By conditioning on S^n we then see that this new channel is equivalent to the standard nonfeedback additive white Gaussian noise channel with noise variance equal to $\text{var}(Z_i | S_i) = \text{var}(Z_i | Z_i + \tilde{Z}_i) = \epsilon^2/(\epsilon^2 + 1)$. Obviously, the reliability function of this new problem dominates that of the original problem, since here the encoder and decoder are supplied with more information. Therefore, we have proved the following statement.

Proposition 3.

$$E_{\text{FB}}(R; P, \epsilon^2) \leq E\left(R; \frac{P(\epsilon^2+1)}{\epsilon^2}\right). \tag{20}$$

Simple as the argument leading to it may be, the bound (6) is tighter than (9) in many cases. Furthermore, from (6), we can conclude that the noisy feedback can be no more useful than having the power increase by the factor of $(1+\epsilon^2)/\epsilon^2$ in the absence of feedback. This observation, combined with the sphere packing bound on the nonfeedback reliability function [16], [1], implies that

$$E_{\text{FB}}(R; P, \epsilon^2) \leq \frac{P(1+\epsilon^2)}{2\epsilon^2}. \tag{21}$$

III. Sensitivity of Linear Feedback Coding Schemes

It is relatively well-known that the Schalkwijk–Kailath coding scheme is sensitive to the noise in the feedback link; see, for example, Schalkwijk [14, Part II, Section II-D] or Lavenberg [10, p. 479]. In this section, we give a simple proof of this observation and extend it to a class of coding schemes based on linear encoding of message and feedback signals.

The basic idea is that any "successful" linear feedback coding scheme that achieves a positive rate $R^* > 0$ under $\epsilon^2 > 0$ must have the infinite reliability for all rates $R < R^*$. This clearly contradicts the upper bound $E_{\text{FB}}(R; P, \epsilon^2) = O(1/\epsilon^2)$ obtained in the previous section, and therefore a linear feedback coding scheme cannot achieve any positive rate.

We now make this simple argument more precise. First a sequence of (M_n, n) codes is called linear if the encoding functions $f_i^{(n)}$ can be expressed in the following form:

$$X_i = f_i^{(n)}(W, \tilde{Y}^{i-1}) = L_i^{(n)}(\theta(W), \tilde{Y}^{i-1}), \quad i = 1, \ldots, n,$$

where $\theta(W)$ takes values in $\mathbb{R}^k$ with dimension k independent of n, and $L_i^{(n)}$ is a linear function of $(\theta, \tilde{Y}^{i-1})$.

Proposition 4. *For any sequence of (M_n, n) linear feedback codes, if $P_e^{(n)} \to 0$, then $M_n/n \to 0$.*

Proof sketch. Suppose there exists a sequence of linear feedback codes that achieves $R^* > 0$. For simplicity, we assume $k = 1$, i.e., $\theta \in \mathbb{R}$. Now from the positive achievable rate, there exists $\alpha > 0$ such that $\frac{1}{n} I(\theta; Y^n) \geq \alpha$ for all n. But from the linear structure of the encoding functions and the additive nature of the channel, we can represent the channel as $Y_i = \alpha_i \theta + \xi_i$, $i = 1, 2, \ldots, n$, for appropriately chosen Gaussian random variables $(\xi_1, \ldots, \xi_n)$ independent of θ. Since a Gaussian input maximizes the mutual information over the Gaussian channel, we thus have

$$\frac{1}{n} I(\theta_G; Y^n) \geq \frac{1}{n} I(\theta; Y^n) \geq \alpha$$

where θ_G is a Gaussian random variable with the same mean and variance as θ. But from joint Gaussianity of (θ_G, Y^n), it can be easily verified that there exists a linear function $L(Y^n)$ that can be written as

$$L(Y^n) = \theta_G + \xi \tag{22}$$

for some Gaussian random variable ξ independent of θ_G, so that $I(\theta_G; L(Y^n)) = I(\theta_G; Y^n) \geq n\alpha$. This further implies that $E\xi^2 \leq (E\theta^2)/(e^{2n\alpha} - 1)$. This implies that for the channel (22), we can use the uniform message constellation as in the original Schalkwijk–Kailath coding scheme, which achieves $P_e^{(n)} \dot{\leq} \exp(-e^{2n(\alpha - R)})$ for any rate $R < \alpha$. In particular, $E_{\text{FB}}(R; P, \epsilon^2) = \infty$ for $R < \alpha$, which clearly contradicts the fact that $E_{\text{FB}}(R; P, \epsilon^2) < \infty$ for all R and all $\epsilon^2 > 0$. □

From a similar argument, Proposition 4 can be further extended to concatenated codes with a linear feedback code as the inner code.

IV. Lower Bounds on the Reliability Function

In this section, we study lower bounds on the reliability function $E_{\text{FB}}(R; P, \epsilon^2)$. In particular, we obtain a coding scheme that achieves the error exponent $\Omega(1/\epsilon^2)$.

We first consider the case of of two-message communication, which becomes useful later and is interesting on its own.

A. Feedback Coding for Two Messages

Suppose we wish to communicate one of two messages, namely, $W \in \{1, 2\}$, over the noisy feedback channel (1), (2). From Proposition 3, the optimal error exponent is upper bounded by

$$E_{\text{FB}}(\text{binary}; P, \epsilon^2) := \lim_{n\to\infty} -\frac{1}{n} \log P_e^{(n)} \le \frac{P(1+\epsilon^2)}{2\epsilon^2}.$$

On the other hand, we have the following lower bound, which meets the upper bound up to the same constant.

Proposition 5.

$$E_{\text{FB}}(\text{binary}; P, \epsilon^2) \ge \frac{P}{2\epsilon^2}.$$

Proof sketch. Suppose $W = 1$ is sent. At time 1, we send $X_1 = a$, and subsequently send $X_i = a - Z_{i-1} - \tilde{Z}_{i-1}$ for some positive constant a, the value of which will be specified later. In words, each transmission cancels the effect of noises in both forward and backward channels. If $W = 2$ is to be communicated, then we use $-a$ instead of a.

After n transmissions, the decoder declares $\hat{W} = 1$ if $\sum_{i=1}^n Y_i > 0$ and declares $\hat{W} = 2$ otherwise. Conditioned on $W = 1$, it is easy to verify that $\sum_i Y_i \sim N(na, 1+(n-1)\epsilon^2)$. By symmetry, this implies that $P_e^{(n)} \doteq \exp(-na^2/(2\epsilon^2))$, while the expected power satisfies $E\left[\frac{1}{n}\sum_{i=1}^n (X_i)^2\right] \le a^2 + 1 + \epsilon^2 =: Q$. In other words, we can achieve

$$\lim_{n\to\infty} -\frac{1}{n} \log P_e^{(n)} = \frac{Q - 1 - \epsilon^2}{2\epsilon^2}.$$

Now by operating only fraction γ of the time, we can achieve the error exponent $\gamma(Q - 1 - \epsilon^2)/(2\epsilon^2)$ under the power constraint γQ. Finally, by taking $\gamma \to 0$ with $P = \gamma Q$ held fixed, we have the desired result. □

B. Three-Phase Protocol

Here we sketch a coding scheme that achieves the error exponent $\limsup_{n\to\infty} -\frac{1}{n}\log P_e^{(n)} = \Omega(1/\epsilon^2)$. This coding scheme has three phases. In the first phase, the channel is used as a standard Gaussian channel without feedback and at the end of the first phase, the encoder guesses from the noisy feedback $\tilde{Y}^n$ whether the receiver Y^n has decoded the message correctly. Then, depending on this decision, the encoder communicates in the second phase a binary message—SUCCESS or FAILURE, in a manner that will be described shortly. Finally in the third phase, if a FAIRLURE has been declared, the message is retransmitted, this time with very high power for a very short period of time; otherwise, nothing is transmitted. This coding scheme turns out to achieve the desired error exponent under the given power constraint.

Because the difference $\tilde{Z}^n = \tilde{Y}^n - Y^n$ between the true output and the noisy observation has a small variance ϵ^2, we can devise a decoding scheme for the first phase (nonfeedback coding) so that Y^n and $\tilde{Y}^n$ can agree upon the correct message with high probability. There are two type of events to be dealt with. The first type, called Type I (false positive), is the event that the noisier receiver $\tilde{Y}^n$ decodes the message correctly, but the true receiver Y^n does not. The second type, called Type II (false negative), is the event that $\tilde{Y}^n$ decodes the message incorrectly, so the FAILURE is declared. The basic ingredient for the first-phase code design is the feedback decoder (on the transmitter side) that achieves a reasonably small Type II error probability (with positive error exponent), but a very small Type I error probability (with the error exponent $\sim 1/\epsilon^2$).

As we saw before, the second phase also has the exponent $\sim 1/\epsilon^2$. Finally, for the third phase, we use a large amount of power (exponential in block size) and retransmit the message via the uniform message constellation coding as in the Schalkwijk–Kailath coding. The error exponent for the third phase is infinite. But the average power spent can be made to stay the same, since the probability of retransmission, which is less than the probability of Type II error, is also exponentially small.

REFERENCES

[1] A. E. Ashikhmin, A. Barg, and S. N. Litsyn, "A new upper bound on the reliability function of the Gaussian channel," *IEEE Trans. Inf. Theory*, vol. IT-46, no. 6, pp. 1945–1961, 2000.

[2] M. V. Burnashev, "Data transmission over a discrete channel with feedback: Random transmission time," *Problems of Information Transmission*, vol. 12, no. 4, pp. 10–30, 1976.

[3] S. Butman, "A general formulation of linear feedback communication systems with solutions," *IEEE Trans. Inf. Theory*, vol. IT-15, no. 3, pp. 392–400, May 1969.

[4] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 439–441, 1983.

[5] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, 2nd ed. New York: Springer-Verlag, 1998.

[6] S. C. Draper and A. Sahai, "Noisy feedback improves communication reliability," in *Proc. IEEE International Symposium on Information Theory*, Seattle, WA, July 2006, pp. 69–73.

[7] R. L. Kashyap, "Feedback coding schemes for an additive noise channel with a noisy feedback link," *IEEE Trans. Inf. Theory*, vol. IT-14, no. 3, pp. 1355–1387, 1968.

[8] Y.-H. Kim, "Feedback capacity of stationary Gaussian channels," submitted to *IEEE Trans. Inf. Theory,* February 2006.

[9] A. J. Kramer, "Improving communication reliability by use of an intermittent feedback channel," *IEEE Trans. Inf. Theory*, vol. IT-15, pp. 52–60, Jan. 1969.

[10] S. S. Lavenberg, "Feedback communication using orthogonal signals," *IEEE Trans. Inf. Theory*, vol. IT-15, pp. 478–483, 1969.

[11] N. Merhav and T. Weissman, "Coding for the feedback Gel'fand–Pinsker channel and the feedforward Wyner–Ziv source," *IEEE Trans. Inf. Theory*, vol. IT-52, no. 9, pp. 4207–4211, Sept. 2006.

[12] L. H. Ozarow, "The capacity of the white Gaussian multiple access channel with feedback," *IEEE Trans. Inf. Theory*, vol. IT-30, no. 4, pp. 623–629, 1984.

[13] M. S. Pinsker, "The probability of error in block transmission in a memoryless Gaussian channel with feedback," *Problemy Peredači Informacii*, vol. 4, no. 4, pp. 3–19, 1968.

[14] J. P. M. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback," *IEEE Trans. Inf. Theory*, vol. IT-12, pp. 172–182, 183–189, Apr. 1966.

[15] C. E. Shannon, "The zero error capacity of a noisy channel," *IRE Trans. Inf. Theory*, vol. IT-2, no. 3, pp. 8–19, Sept. 1956.

[16] ——, "A note on a partial ordering for communication channels," *Information and Control*, vol. 1, pp. 390–397, 1958.

[17] K. S. Zigangirov, "Upper bounds for the probability of error for channels with feedback," *Problemy Peredači Informacii*, vol. 6, no. 2, pp. 87–92, 1970.