Error Exponent for Gaussian Channels with Partial Sequential Feedback

Manish Agarwal, Dongning Guo, and Michael L. Honig
Department of Electrical Engineering and Computer Science
Northwestern University
2145 Sheridan Road, Evanston, IL 60208 USA
{m-agarwal, dGuo, mh}@northwestern.edu

***Abstract*— We consider an additive white Gaussian noise (AWGN) channel with *partial* sequential feedback. Namely, for every fixed-length block of forward transmissions a fraction of the received symbols are fed back sequentially to the transmitter through a noiseless feedback link. It is well known that complete noiseless feedback can provide a dramatic improvement in reliability (i.e., double-exponential error rate with block length). We show that partial feedback can also provide a substantial improvement in error rate. Specifically, we propose a capacity-achieving coding scheme with partial feedback, in which the feedback is used to induce a prior distribution for the decoding of random forward error control (FEC) codewords. The error-exponent for this scheme is larger than the error-exponent with FEC coding only at all rates. For rates greater than those achieved by transmissions with feedback alone, we give an upper bound on the error exponent. Exponents close to this bound can be achieved with both the proposed scheme and a simple rate-splitting scheme. With finite block lengths, the proposed coding scheme achieves lower error rates than rate-splitting.**

I. Introduction

Feedback from the output of a noisy channel to the transmitter can both reduce the complexity of channel coding/decoding and improve reliability over the communication link. The availability of feedback, however, can depend on several factors, including the capacity and availability of a feedback link, and the processing delay at receiver. Here we consider an AWGN channel with *partial* noiseless feedback, i.e., the number of feedback transmissions, each representing a single output symbol, is less than the total number of forward transmissions. This may occur, for example, due to scheduling, bandwidth and delay constraints on the feedback link.

Channel coding with complete (sequential) noiseless feedback has been traditionally studied in a control theoretic framework, in which the objective is to drive the receiver towards a scalar value that corresponds to the message [1], [3], [4]. In contrast, without feedback efficient FEC codes have been designed and studied from an information theoretic perspective. For the partial feedback channel considered here, neither approach alone can generally achieve optimal performance, measured in terms of error rate. Namely, with a control-based coding scheme alone the transmitter has only partial observations of the receiver state (i.e., sequence of decoder inputs).[1] In general, the capacity cannot be achieved with such schemes (e.g., [1], [3]). Of course, the capacity can be achieved with an optimal (e.g., random) FEC code, however, the error rate then does not benefit from feedback.

Here we propose a fixed-length *hybrid* coding scheme with partial feedback, which is capacity-achieving and combines the control-based approach with random coding. We show that the error rate is substantially lower than that achievable with either approach alone. In this scheme a feedback (control-based) code is applied to symbols with feedback, and is used to generate a prior distribution for the decoding of an FEC (forward) code, which is applied to the symbols without feedback. The forward code corresponds to the entire set of messages, so that the forward code rate may exceed capacity. However, the receiver decodes the forward code with the maximum *a posteriori* (MAP) criterion using the prior distribution from the feedback code. In effect, each message is assigned two codewords, which jointly assist the decoder.

The transmitted symbols, which are fed back, can occur anywhere within the message block. Here we focus on the AWGN channel with Schalkwijk-Kailath (S-K) coding [1] for transmissions with feedback, and a random forward code. We compute a lower bound on the error exponent of the hybrid coding scheme, which is substantially better than the error exponent without feedback.[2] If the number of feedback transmissions is $N_f = fN$, where $0 < f < 1$, N is the block length and C is the channel capacity, then for transmission rate $R < fC$ the error exponent has an exponential form, similar to that derived in [1], but with an additional constant term provided by the forward code. At rates $R \geq fC$ the error exponent corresponds to a $(1-f)N$ length random forward code operating at rate $(R - fC)/(1-f)$. We also show that the corresponding sphere packing error exponent upper bounds the error exponent in this regime, hence the exponent for the proposed coding scheme is quite close to this upper bound.

The error rate of the proposed coding scheme is also compared with the *rate-splitting* scheme in which we have

This work was supported by the U.S. Army Research Office under grant DAAD19-99-1-0288, NSF under grant CCR-0310809 and the DARPA IT-MANET program Grant W911NF-07-1-0028.

[1] In principle, the receiver could map *any* set of received symbols to a single feedback symbol. With that assumption, the coding scheme proposed in [6] gives a doubly-exponential decay in error rate. In contrast, we assume that the feedback depends on a *subset* of received symbols.

[2] The error exponent is redefined to account for double-exponential decay, and can depend on block length.

independent feedback and feedforward codes operating at rates, which add up to give an average rate of R. Those results show that the proposed scheme can provide significantly lower error rates with small to moderate block lengths. We also show that feeding back a linear combination of received symbols (rather than a particular symbol) can reduce the peak-to-average power.

A similar partial sequential feedback model to that considered here is analyzed in [7] for the binary symmetric channel, although the coding scheme in [7] is substantially different from that presented here (and does not achieve capacity due to the use of convolutional codes). Other recent work on limited and noisy feedback models is presented in [8]–[11]. Our model differs in that the feedback can be sporadic, but is noiseless, and is not quantized. Our model also differs from the feedback models considered in [10]–[12]. Namely, those schemes allow variable decoding delay and/or the receiver must *compute* an appropriate feedback signal. Here we consider fixed-length block codes, where some the received symbols are fed back directly without processing.

II. The Hybrid Coding Scheme

We consider data transmission through an AWGN channel. If the scalar x is the input to the channel, then the channel output is given by

$$y = x + n \tag{1}$$

where n is zero-mean Gaussian noise with variance σ^2, and is independent for each channel use. Furthermore, the average power of the input x is no more than P. Given the data rate R nats per channel use and a block length of N channel uses, the transmitter transmits one of the $M = e^{RN}$ messages. The capacity is given by

$$C = \frac{1}{2}\log\left(1 + \frac{P}{\sigma^2}\right). \tag{2}$$

In the absence of feedback, this capacity can be achieved with a random codebook, where the input distribution is Gaussian with zero mean and variance P. In that case, the error exponent, averaged over random codebooks, can be lower-bounded by Gallager's exponent [13]

$$E_g[R] = \max_{0<\rho\leq 1}\left\{E_o(\rho) - \rho R\right\}, \qquad R < C \tag{3}$$

where, letting $g(x)$ and $g(y|x)$ denote the pdf of the Gaussian input and conditional pdf of output given input, respectively,

$$\begin{aligned} E_o(\rho) &= -\log\left[\int_y\left(\int_x (g(y|x))^{1/(1+\rho)} g(x)dx\right)^{(1+\rho)} dy\right] \\ &= \frac{\rho}{2}\log\left[1 + \frac{P}{\sigma^2(1+\rho)}\right]. \end{aligned} \tag{4}$$

Now suppose that $N_f < N$ received symbols are sequentially available at the transmitter. Those symbols can occupy arbitrary locations within the packet. The locations are known *a priori* to both the transmitter and receiver. It is known that feedback cannot increase the capacity of this channel [13]. Here we present a capacity-achieving coding scheme, which exploits the available feedback to improve the reliability of the decoded message.

Following [1], each message is mapped to one of M equally spaced points $\{\theta_i = \frac{i}{M} : i = 1, \dots M\}$ on the interval $[0,1]$. Assume that the k^{th} message, θ_k, is to be transmitted. The S-K coding scheme is used for the N_f transmitted symbols for which feedback is available, which directs the corresponding estimate at the receiver to the point θ_k. Let $\mathbf{X}_f = \left[x_{f1}, x_{f2}, \dots x_{fN_f}\right]$ denote the vector of transmitted symbols with feedback, and $\mathbf{Y}_f = \left[y_{f1}, y_{f2}, \dots y_{fN_f}\right]$ denote the corresponding vector of received symbols. The following power constraint is satisfied,

$$\frac{1}{N_f} E\left[\sum_{n=1}^{N_f} x_{fn}^2\right] \leq P \tag{5}$$

where the expectation is over the channel noise and message θ_k.

Proposition 1: Assuming that the message θ_k is transmitted and that all messages are equally likely, the likelihood ratio induced by the S-K coding scheme $\forall i \neq k$ after N_f transmissions is given by

$$\frac{p(\theta_i|\mathbf{Y}_f)}{p(\theta_k|\mathbf{Y}_f)} = \exp\left[-6\alpha^{2N_f}\left(\theta_i - \theta_k\right)\left(\theta_i + \theta_k - 2\,\hat{\theta}_k\right)\right] \tag{6}$$

where $p(\theta_i|\mathbf{Y}_f)$ is the conditional probability of message θ_i given $\mathbf{Y}_f$,

$$\alpha^2 = \frac{N_f - 1}{N_f} + \frac{P}{\sigma^2} \tag{7}$$

and $\hat{\theta}_k$ is function of $\mathbf{Y}_f$ only.

Proofs of our main results are omitted due to the space limitation. Using the fact that $\mathbf{Y}_f$ depends on θ_k and channel noise, it can be shown that $\hat{\theta}_k$ is a Gaussian random variable with mean θ_k and variance $1/(12\,\alpha^{2N_f})$.

The remaining $N_c = N - N_f$ transmissions are used to transmit an FEC codeword corresponding to the message k. The FEC code rate is therefore $R/(1-f)$ nats per channel use, which can exceed the channel capacity. However, we will see that the prior distribution induced by the feedback code effectively reduces this rate.

To simplify the analysis we assume a random, Gaussian FEC code-book for the M messages. Let $\mathbf{X}_c(i)$ denote the codeword corresponding to message i. Letting $\mathbf{Y}_c$ denote the vector of received symbols corresponding to the FEC code, and $p(\mathbf{X}_c(i)|\mathbf{Y}_c)$ denote the posterior probability of the codeword $\mathbf{X}_c(i)$ given $\mathbf{Y}_c$, the optimal (MAP) receiver decodes the message j if[3]

$$p(\mathbf{X}_c(j)|\mathbf{Y}_c) \geq p(\mathbf{X}_c(i)|\mathbf{Y}_c), \qquad \forall i = 1, \dots M. \tag{8}$$

If the receiver has the *a priori* message distribution $\{p(i) : i = 1, \dots, M\}$, then, according to the decoding rule (8), the receiver decodes message j if

$$g(\mathbf{Y}_c|\mathbf{X}_c(j))\, p(j) \geq g(\mathbf{Y}_c|\mathbf{X}_c(i))\, p(i), \; \forall i, \tag{9}$$

[3] Ties are broken arbitrarily.

where $g(\mathbf{Y}_c|\mathbf{X}_c(i))$ denotes the likelihood of receiving $\mathbf{Y}_c$ given $\mathbf{X}_c(i)$ is transmitted. Since the transmitted message is k, the receiver makes an error if the decoded message $j \neq k$.

Lemma 1: Let $p_e(k)$ denote the probability of decoding error given that message k is transmitted, averaged over randomly generated FEC codebooks. Given the prior message distribution $\{p(i) : i = 1, \ldots, M\}$, and the MAP decoding rule (9), for any $\rho \in (0, 1]$,

$$p_e(k) \leq \left[\sum_{i, i \neq k} \left(\frac{p(i)}{p(k)} \right)^{1/(1+\rho)} \right]^{\rho} \cdot \exp\left[-N_c E_o(\rho)\right]. \quad (10)$$

Now if the receiver uses the distribution induced by S-K coding as the *a priori* distribution, assuming the same message k is transmitted, that is $p(i) = p(\theta_i|\mathbf{Y}_f)\ \forall i$, then averaging (10) over the received vector $\mathbf{Y}_f$ gives

$$p_e(k) \leq L(\theta_k, \rho) \cdot \exp\left[-N_c E_o(\rho)\right] \quad (11)$$

for any $\rho \in (0, 1]$, where

$$L(\theta_k, \rho) = E_{\hat{\theta}_k} \left[\sum_{i, i \neq k} \left(\frac{p(\theta_i|\mathbf{Y}_f)}{p(\theta_k|\mathbf{Y}_f)} \right)^{1/(1+\rho)} \right]^{\rho} \quad (12)$$

where the expectation is over $\hat{\theta}_k$, defined after Proposition 1.

The quantity $L(\theta_k, \rho)^{1/\rho}$ represents the *effective* number of transmitted messages. Namely, in the absence of any feedback the *a priori* distribution is uniform over θ_i, that is, $p(\theta_i|\mathbf{Y}_f) = 1/M\ \forall i$, so that $L(\theta_k, \rho) = (M-1)^{\rho}$, and (11) reduces to the conventional random coding error exponent [13]. With feedback the prior distribution induced by the S-K scheme effectively reduces the number of candidate messages to be considered by the decoder.

The bound (11) serves as the basis for the subsequent lower bound on error exponent for the hybrid coding scheme.

III. Error Exponent with Partial Feedback

Let

$$P_e = \frac{1}{M} \sum_{k=1}^{M} p_e(k). \quad (13)$$

If the error probability decreases exponentially with N, then the exponent can be defined as $E = \lim_{N\to\infty} -(\log P_e)/N$. Because the error rate can decrease faster than exponentially with feedback, we use the following more general definition,

$$\lim_{N\to\infty} -\frac{\log P_e}{N} - E(N) = 0 \quad (14)$$

This definition is consistent with the preceding conventional definition when the error probability decays exponentially with N, but also allows the exponent to be a function of the code length N.

The following theorem gives a lower bound on the exponent for the hybrid coding scheme.

Theorem 1: For the hybrid coding scheme defined in Section II, given a fixed ratio of feedback transmissions $f = \frac{N_f}{N}$ and normalized data rate $r = \frac{\log M}{N\,C} = \frac{R}{C}$, the error exponent defined in (14) satisfies

$$E(N) \geq \begin{cases} \frac{3}{2\,N} e^{[2(f-r)C\,N]} + (1-f)\,E_o(1) & \text{if } 0 < r < f \\ (1-f)\,E_g\left[\left(\frac{r-f}{1-f}\right) C\right] & \text{if } f \leq r < 1 \end{cases} \quad (15)$$

For the extreme cases of no feedback ($f \to 0$) and full feedback ($f \to 1$), the bound (15) gives Gallager's random coding exponent and an exponential form similar to the one derived in [1], respectively. For rates $r < f$ the error exponent has an exponential form (which is same as the one in [1] except for a multiplicative factor of half) with an additional constant term.[4] That is, the error rate decays double-exponentially with block length. This is because the S-K scheme alone can achieve the rate rC without making use of the remaining $(1-f)N$ transmissions. The remaining channel symbols are used by the FEC code, which essentially operates at zero rate and gives the additional constant term in (15).

For $r > f$, the error exponent is a (rate-dependent) constant, i.e., the error rate decays exponentially with N. In this regime, the S-K scheme cannot achieve the target rate with vanishing error probability for large N. Hence the error rate is dominated by the performance of the FEC code, but with an effective rate of $(r-f)C/(1-f)$ nats per channel use. This is due to the prior distribution provided by the S-K coding scheme, which effectively reduces the number of messages to be considered as candidates for the FEC decoder. From the convexity of $E_g[\cdot]$ the exponent in this region is strictly larger than the exponent with FEC coding only, i.e., $E_g[rC]$.

Fig. 1 shows plots of the lower bound in Theorem 1 for different feedback fractions f. For all of the numerical plots we take $P = 2$ and $\sigma^2 = 1$. For $r < f$, the curves are shown as dotted vertical lines to indicate that in this region the exponent depends on N and is effectively infinite for large N. The figure shows the transition from double-exponential error rate for $r < f$ to exponential error rate for $r > f$.

The dotted straight line (across the corners) in the figure is $(1-r)E_o(1)$, which intersects the error exponent at $r = f$. It can be seen that the exponent with partial feedback $(0 < f < 1)$ is strictly larger than the exponent with no feedback $(f = 0)$ at all rates.

Fig. 2 shows plots of probability of error versus the feedback fraction f for the finite block length $N = 100$. Here we use the following bound on probability of error, which is derived as part of the proof of Theorem 1,

$$P_e \leq 8\left(2 + \frac{4\sqrt{2}}{d}\right)^{\rho} \exp\left[-\frac{\rho^2 d^2}{2(1+\rho)^2} - (1-f)E_o(\rho)N\right] \quad (16)$$

where $d = \sqrt{12}\alpha^{(f-r)N}$, α is given in (7), and $\rho \in (0, 1]$ can be chosen to minimize the right-hand side. Two plots are shown for $r = 0.5$ and 0.6. There is a steep fall, which occurs when the feedback ratio f crosses r. This is consistent with the transition from exponential to double-exponential decay.

[4] The first term differs from the one in [1] by a factor of half, which may be due to the looseness of the bound (10).

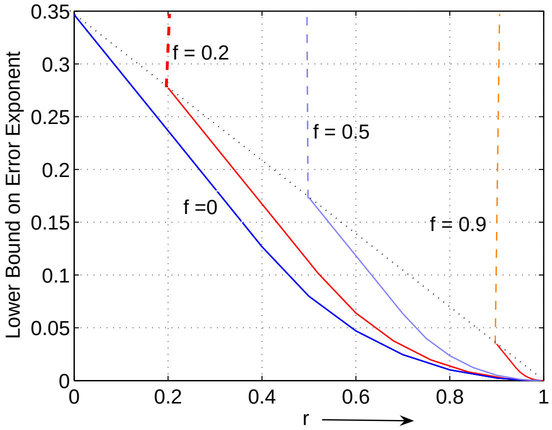

Fig. 1. The error exponent lower bound versus normalized rate, parameterized by feedback transmission ratio f.

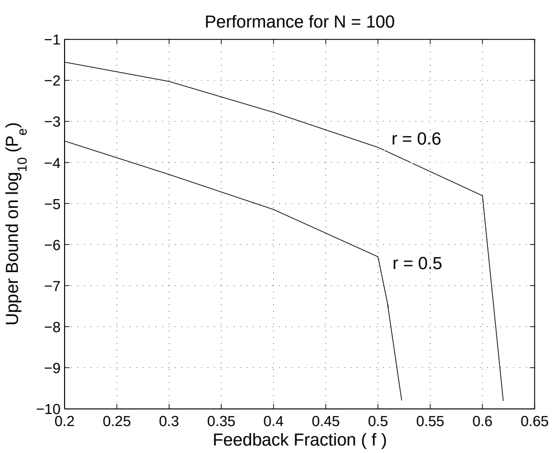

Fig. 2. The error probability upper bound versus feedback transmission ratio f for a code length of $N = 100$.

A. Comparison With Rate Splitting

A simple approach to coding with partial feedback is to split the rate as $r = r_{\mathsf{fb}} + r_{\mathsf{ff}}$ where r_{fb} and r_{ff} are the rates associated with the feedback and feedforward codes, respectively. That is, given the block length N, the fN-length feedback code has rate r_{fb}/f and the $(1-f)N$-length forward code has rate $r_{\mathsf{ff}}/(1-f)$. Letting $P_{e;\mathsf{fb}}(r_{\mathsf{fb}})$ and $P_{e;\mathsf{ff}}(r_{\mathsf{ff}})$ denote the error probabilities associated with these feedback and feedforward codes, respectively. We can select r_{fb} to minimize the probability of error $P_{e;\mathsf{rs}} = P_{e;\mathsf{fb}}(r_{\mathsf{fb}}) + P_{e;\mathsf{ff}}(r_{\mathsf{ff}}) - P_{e;\mathsf{fb}}(r_{\mathsf{fb}})\,P_{e;\mathsf{ff}}(r_{\mathsf{ff}})$. From [1] $P_{e;\mathsf{fb}}(r_{\mathsf{fb}})$ decays double-exponentially and $P_{e;\mathsf{ff}}(r_{\mathsf{ff}})$ is upper-bounded by the random coding bound [13].

For $r > f$, this rate-splitting scheme achieves the same error exponent as shown in (15). Namely, if r_{fb} is slightly smaller than f, then for large N, $P_{e;\mathsf{fb}}(r_{\mathsf{fb}})$ decays double exponentially, and $P_{e;\mathsf{ff}}(r_{\mathsf{ff}})$ is the dominant term. As N grows, r_{fb}, can be chosen arbitrarily close to f, which gives the exponent in (15). For $r < f$, using the feedback code alone gives $P_{e;\mathsf{rs}} = P_{e;\mathsf{fb}}(r_{\mathsf{fb}})$, which decays as $\exp(-3\exp[(f - r_{\mathsf{fb}})CN])$. For large N, the corresponding exponent is larger than that shown in (15) due to the extra factor of half in (15). For small to moderate N, and rates close to fC, the exponent in (15) contains an additional constant, and is therefore larger than that for feedback coding alone.

This is illustrated in Fig. 3, which compares the random coding upper bound on probability of error for the two schemes versus N with $f = 0.5$. Plots are shown for $r = 0.492$ and $r = 0.6$. For the rate-splitting scheme, r_{fb} is selected to minimize $P_{e;\mathsf{rs}}$. The plot shows that for $r = 0.6$ the error probability for the hybrid scheme is only slightly better than that for rate-splitting. In contrast, for $r = 0.492$, the error probability for the hybrid scheme is substantially smaller than that for rate-splitting or feedback coding alone. Also shown for comparison is the error probability with forward coding alone. These plots were generated by computing the random coding upper bound on the probability of error using (11), as opposed to using (16), which is slightly looser.

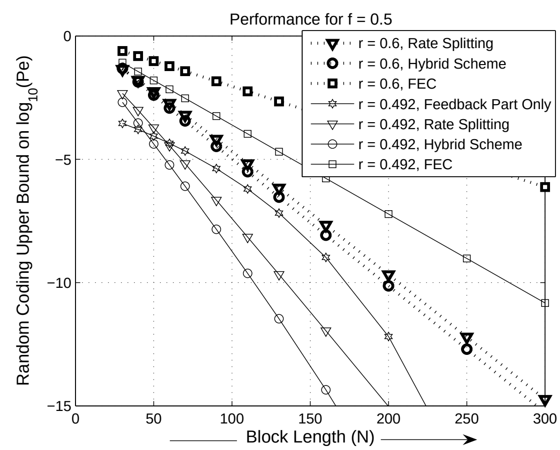

Fig. 3. The upper bound on the probability of error with random forward coding for the hybrid scheme and rate-splitting versus block length N.

IV. Converse: Upper Bound on Error Exponent

Theorem 1 and the preceding discussion indicate that as r increases there is a breakpoint at $r = f$, for which the error probability for the coding schemes considered transitions from double-exponential to exponential decay. The following theorem implies that this is characteristic of all good codes. Namely, for $r > f$ no coding scheme can achieve an error probability, which decays faster than exponential with block length. Although the following theorem is stated for discrete memoryless channels (DMCs) and AWGN channel, it applies to a wide class of continuous input/output alphabet channels.

Theorem 2: Consider a DMC or AWGN channel with capacity C_d and sphere packing bound on error exponent $E_{sp}[u]$, where $0 \le u \le C_d$. With partial feedback fraction $f < r$, the average probability of error satisfies

$$\lim_{N\to\infty} -\frac{\log P_e}{N} \leq (1-f)E_{sp}\left[\left(\frac{r-f}{1-f}\right)C_d\right]. \quad (17)$$

An outline of the proof follows. The probability of error is first lower bounded by applying [2, Thm 1], where all $N_f = fN$ feedback symbols are placed at the beginning of the block. The lower bound is the product of probability of error for the feedback and feedforward component codes. The feedback code produces a list of messages and a decoding error occurs only if the transmitted message is not in the list. This lower bound holds for any list size. Choosing the list size such that the feedback code operates at a rate slightly more than the capacity yields a tight lower bound on the probability of error and hence an upper bound on the exponent. Although [2, Thm 1] was proved only for DMCs without feedback, it can be shown to hold when feedback is present and/or input-output alphabets are continuous.

This implies that for $r > f$, the exponents for the coding schemes considered are close to optimal. The gap can perhaps be bridged by replacing the random forward codes with codes that achieve the sphere packing bound.

V. Linearly Processed Feedback

Here we extend the previous coding scheme to the scenario in which the receiver is capable of linearly processing the received symbols to construct an appropriate feedback signal.

Assume that the receiver feeds back a linear combination of received symbols every D forward transmissions. Without loss of generality, each feedback symbol can be seen as the first element of the processed output vector defined as $\bar{\mathbf{Y}} = \mathbf{U}\mathbf{Y}$, where $\mathbf{Y}$ is $D \times 1$ vector of received symbols and $\mathbf{U}$ is a $D \times D$ unitary matrix. Since we have an AWGN channel, the received vector $\mathbf{Y} = \mathbf{X} + \mathbf{n}$, where $\mathbf{X}$ is the $D \times 1$ vector of transmitted symbols and $\mathbf{n}$ is $D \times 1$ vector of noise samples. Further, without loss of generality, we can look at the input as a precoded vector $\mathbf{X} = \mathbf{U}^\dagger \bar{\mathbf{X}}$ so that the processed output can be written as,

$$\bar{\mathbf{Y}} = \bar{\mathbf{X}} + \bar{\mathbf{n}} \quad (18)$$

where, $\bar{\mathbf{n}} = \mathbf{U}\mathbf{n}$ is the processed noise vector which contains independent Gaussian entries with zero mean and variance σ^2. Note that (18) can be seen as an equivalent additive Gaussian noise channel on which every Dth received symbol is fed back to the transmitter. To summarize, an AWGN channel with linearly processed feedback is equivalent to another AWGN channel with partial feedback at feedback ratio $f = 1/D$. Of course, taking U as the Identity matrix corresponds to the case of no linear processing, which is our original feedback model. Clearly, by Theorem 2, even with linear processing for rates $r > 1/D$ only an exponential decay in error rate is possible.

Although linear processing does not improve the error exponent it may be beneficial in certain cases. For example, implementation of the S-K coding scheme in the time domain exhibits large peak-to-average power [5]. With linear processing, where $\mathbf{U}$ is the DFT matrix, we can implement the S-K coding in the frequency domain and hence reduce the peak-to-average power ratio roughly by a factor of D. Note that the size of the DFT (determined by D) trades off the peak power reduction with reliability.

VI. Conclusions

We have presented a capacity-achieving coding scheme for the AWGN channel with partial sequential feedback, which combines feedback-control and FEC. The scheme can achieve a substantially lower error probability than feedback or feedforward coding alone, and also typically performs better than rate-splitting with finite block lengths. An upper bound on the error exponent for high rates has been presented, which is approached by both the hybrid and rate-splitting schemes.

Extensions of this work may include optimizing the power constraint across the feedback and feedforward codes, and studying the effect of colored noise. In the latter case, linear processing of feedback may be useful. Finally, the approach to coding with partial feedback presented here can be used with any constituent feedback and FEC coding schemes, and can be applied to a wide variety of channels.

References

[1] J.P.M. Schalkwijk, “A coding scheme for additive noise channels with feedback,” *IEEE Trans. Inform. Theory*, vol. 12, pp. 183-189, April 1966.

[2] C.E.Shannon, R.G.Gallager, E.R.Berlekamp, ”Lower bounds to error probability for coding on discrete memoryless channels-I”, Information and Control 10, 65-103, 1967.

[3] M. Horstein, “Sequential transmission using noiseless feedback,” *IEEE Trans. Inform. Theory*, pp. 136-143, July 1963.

[4] N. Elia, “When Bode meets Shannon: Control-oriented feedback communication schemes,” *IEEE Trans. Automatic Control*, vol 49, No. 9, pp. 1477-1488, September 2004.

[5] A. D. Wyner, ”On the Schalkwijk-Kailath coding scheme with a peak energy constraint,” *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 129-134, January 1968.

[6] R. L. Kashyap, “Feedback coding schemes for an additive noise channels with a noisy feedback link,” *IEEE Trans. Inform. Theory*, vol. IT-14, No. 3, pp. 471-480, May 1968.

[7] M. S. Bargh, J.P.M. Schalkwijk, “Coding for channels with low rate noiseless feedback,” *IEEE International Symposium on Information Theory*, Ulm, Germany, June 1997.

[8] Y. H. Kim, A. Lapidoth, T. Weissman, ”Bounds on the error exponent of the AWGN channel with AWGN-corrupted feedback,” *Proceedings of 24th IEEE Convention of Electrical and Electronics Engineers in Israel*, Eilat, Israel, November 2006.

[9] N. C. Martins, T. Weissman, “Coding for additive white noise channels in the Presence of feedback corrupted by uniform quantization or bounded noise,” *Allerton Conference in Communication,Control and Computing*, Monticello IL, September 2006.

[10] S.Draper, K. Ramchandran, B. Rimoldi, A. Sahai, D. Tse, “Attaining maximal reliability with minimal feedback via joint channel-code and hash-function design,” *Allerton Conference in Communication,Control and Computing*, Monticello IL, September 2005.

[11] S. Draper, A.Sahai, “Beating Burnashev in delay with noisy feedback,” *Allerton Conference in Communication,Control and Computing*, Monticello IL, September 2006.

[12] J. M. Ooi, G. W. Wornell, ”Fast iterative coding techniques for feedback channels”, *IEEE Transactions on Information Theory* 44(7): 2960-2976, November 1998.

[13] R. G. Gallager, ”Information Theory and Reliable Communications”, Wiley and Sons, New York, 1968.