ITW2004, San Antonio, Texas, October 24 – 29, 2004

On the variable-delay reliability function of discrete memoryless channels with access to noisy feedback

Anant Sahai and Tunç Şimşek
Electrical Engineering & Computer Science
University of California, Berkeley
{sahai,simsek}@eecs.berkeley.edu

Abstract — **We give a scheme for variable-delay reliable communication over a noisy discrete memoryless channel with a noisy feedback DMC being available between the receiver and the transmitter. This scheme works when the capacity of the feedback channel is greater than the capacity of the forward channel and the target rate of communication on the forward channel is less than its capacity. Rather than looking at a single block in isolation, we consider the scenario where we expect this system to carry an infinite stream of packets. We show that in the limit of nearly noiseless feedback channels, we approach the Burnashev bound[1] on the reliability of variable-delay channel codes.**

I. INTRODUCTION

Feedback can be an important part of the design of a communication system and it is widely believed that feedback can simplify the design of a communication system and achieve higher performance. It is therefore surprising that there a pair of negative results in classical information theory regarding feedback:

1. For discrete memoryless channels, even noiseless feedback does not increase the Shannon capacity of the channel.[2][1]

2. For sufficiently symmetric discrete memoryless channels, noiseless feedback does not improve the block-coding reliability function in the high-rate regime.[3]

While the first is well known, the second is less widely understood. The basic result for the reliability function is that the sphere-packing bound[2] continues to hold for block-codes over discrete memoryless channels in which causal noiseless feedback is available at the transmitter.[3] This result for DMCs stands in dramatic contrast to the situation for AWGN channels with an average power constraint where it is well known that access to causal noiseless feedback enables a doubly exponential drop in the probability of error with increasing block-length.[8]

Unlike Shannon capacity which is not tied to any particular class of communication schemes, the traditional reliability function is intimately connected to the fixed-delay block-coding setup for communication.[4] As such, it is a natural to wonder whether the negative results on reliability with feedback are tied to the fixed-delay block-coding framework. There are two main approaches which show that by changing the sense of reliability, one can see improvements with feedback. The first of these is the paradigm of variable-delay channel coding and the second is that of anytime coding.

In this paper, after first reviewing some existing results on variable-delay coding and anytime coding in the next two sections, we give a scheme for reliable variable-delay communication over a noisy discrete memoryless channel with a noisy feedback channel between the receiver and the transmitter. It uses a combination of expurgated block codes and an anytime code over the feedback link. We show that in the limit of nearly noiseless feedback channels, we approach the Burnashev bound on the reliability of variable-delay channel codes.

II. VARIABLE-DELAY CODING

Feedback allows the encoder to know when the decoder needs more help in decoding the data. This idea was studied rigorously by Forney[5] in [4] where he used the fact that block-decoding errors occur due to atypical channel behavior. This atypicality is often detectable at the decoder. Forney studied how this can be exploited by using low-rate feedback requesting retransmission of blocks. With the codeword length thus depending on the channel realizations, there is no longer a constant block-length that can be used in the definition of the reliability function. In its place, it is natural to use the expected block-length and Forney shows that it is possible to achieve reliabilities E_f above the sphere-packing bound by using this technique.

With high-rate instantaneous noiseless feedback available, the encoder does not need an explicit repeat-request. By seeing the channel outputs, it already knows that such a repeat is necessary. More significantly, the encoder can detect when the decoder is about to make an error even when the decoder does

[1] Or see exercise 4.6 in [5].

[2] Also called the volume bound

[3] See exercise 5.33 in [5] for an illustration of this in the context of the binary symmetric channel (BSC). It is even easier to see in the context of the binary erasure channel. The cause of block decoding errors is fundamentally the channel being atypically bad for the duration of that block. There is nothing that feedback can do during that block to get around the atypicality.

[4] Traditionally, this can be seen by noticing that the reliability bounds attained by convolutional codes as a function of constraint length are incompatible with the reliability function in terms of block-lengths.[5]

[5] The interested reader should also see exercise 5.19 in [5] for a short intuitive explanation.

not. The encoder can use this to actively steer the decoder away from the error. Burnashev [1] demonstrated that for any 'non-trivial' discrete-time memoryless channel the reliability function is upper bounded by

$$E_{opt}(\bar{R}) \leq C_1\left(1 - \frac{\bar{R}}{C}\right),\ 0 \leq \bar{R} \leq C, \tag{1}$$

where $\bar{R}$ is the expected rate, C is the Shannon capacity of the forward link in bits, and

$$C_1 = \max_{i,k} \sum_l p_{il} \log_2 \frac{p_{il}}{p_{kl}}. \tag{2}$$

where the p_{ij} represent the channel transition probabilities and we use $\log_2$ and hence powers of 2 in the definition of the reliability function. This bound is far better than what was achievable with the explicit repeat requests in [4].

[1] also gives an explicit scheme[6] that can be thought of as a generalization of the one given by Horstein in [6]. This scheme uses noiseless feedback and can approach the bound (1) arbitrarily closely. The scheme works in two phases. The first phase has the encoder explicitly track the decoder's posterior probability distribution over the possible messages. The encoder then transmits channel inputs that aim to maximize the reduction in entropy of that posterior[7] while having the posterior probability of the true message increase to 1. The second phase involves the encoder validating the decoder's state through a confirmation message. If the decoder is off track and about to make a mistake, the encoder attempts to signal to the decoder that it should expect a retransmission.

In [15], a simpler scheme is given which also approaches the Burnashev bound (1). The phases are illustrated in figure 1. The first consists of a block-code that works at a rate very close to capacity. By means of noiseless feedback[8], the encoder knows when the decoder has made an error in decoding the message. The second phase is a confirmation or deny signal followed by possible retransmission of the entire block in case of error.

1. During the time L, the encoder sends either a confirm or deny message by repeating one of the letters i or k that maximize (2). The decoder will "confirm" and decide to accept the block if the empirical type of the received signal is very close to the probability. By the AEP, there exists some suitably tiny $\delta > 0$ so that the probability of decoding a "confirm" as a "deny" is bounded above by δ. More importantly, a "deny" will be misread as a "confirm" with a probability[9] close to $2^{-C_1 L} = 2^{-(N+L)C_1 \frac{L}{N+L}}$ by Sanov's theorem.[2]
2. Because the block-code in phase 1 is at a rate below the capacity, the probability ϵ that the decoder makes an error for a single block can be made arbitrarily small. The probability of repeating a block is therefore bounded above by $\epsilon + \delta$. The stopping time then looks like a geometric random variable and the expected block length is bounded by $\frac{N+L}{1-\epsilon-\delta}$.
3. *By the noiseless nature of the feedback, the encoder always knows which block the receiver is expecting and so knows what to transmit.* The probability of error on the block is therefore only the probability that a "deny" is missed and received as a "confirm". This is $2^{-(N+L)C_1 \frac{L}{N+L}} \approx 2^{-(N+L)C_1(1-\frac{R}{C})}$ while the expected block length is $\frac{N+L}{1-\epsilon-\delta} \approx (N + L)$. Thus the Burnashev bound on reliability can be approached arbitrarily closely.

[6]The design of this scheme parallels the techniques used to prove the upper bound and so demonstrates how it is optimal beyond merely meeting the bound.

[7]It is an interesting fact that this is done by partitioning the messages so that the probability of the partitions approximates the capacity achieving input distribution to the channel. As such, it gives a nice operational interpretation to the capacity-achieving input distribution.

[8]It is important to realize that in this case, the noiseless feedback can be considered as being instantaneous decision-feedback rather than feedback of the exact sequence of channel outputs received. We will build on this property in our scheme.

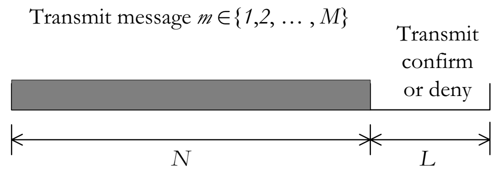

Figure 1: One transmission cycle for the asymptotically optimal encoder with access to noiseless feedback.

In addition to Yamamoto's scheme, there are variable-delay schemes given by Ooi in [7] where he uses ideas from source-coding to view variable-delay channel coding as a sequence of successive compressions of the data to be sent where the previously received symbols are viewed as side-information available at both the transmitter and receiver. With noiseless feedback, he is able to attain the Burnashev Bound (1). His goal is to keep complexity low and he expands this approach to more general channels and also has some initial work on the case of noisy feedback. His approach to noisy feedback is to go to a fixed-length code.

III. Anytime reliability

The fundamental idea of anytime reliability is to understand how the probability of error changes as we vary the delay that we are willing to tolerate.[11]

Definition 1 *The* anytime reliability $E_a(R)$ *achieved by a rate* R *code is the supremum of all* α *for which there exists a* $K > 0$

[9]Notice that the error probabilities are deliberately not symmetric in this. We are much more concerned with the probability of "missed detection" (deny as confirm) than we are with the probability of "false alarm" (confirm as deny). Missed detections let a bad block get out from the decoder. False alarms just cost us a little bit in the expected block length by causing more repeats.

such that for every $t, d > 0$,

$$P(\hat{M}_{t-d}(t) \neq M_{t-d}) \leq K 2^{-\alpha d}$$

where M_i represents the message that arrived at the encoder at time i and $\hat{M}_j(t)$ represents the best estimate at time t for the message M_j. The probability is taken over the channel noise and the messages M_i.

An easy consequence of the exponential decay of the probability of error is that it is equivalent to requiring that there exists some K' so that

$$P(\hat{M}_0^{t-d}(t) \neq M_0^{t-d}) \leq K' 2^{-\alpha d} \quad (3)$$

for all t and d. This means that by time t we can most likely correctly decode the entire sequence of messages up to d time steps ago. This implicit requirement to eventually get all the bits right is what makes anytime reliability special.

Anytime reliability was first introduced in the context of tracking an exponentially unstable scalar Markov process and it was shown in [9] and [11] that in order to communicate such unstable processes across a noisy channel, the channel requires not only enough Shannon capacity but also a high enough anytime reliability.[10] [10] studied the binary erasure channel and conjectured that anytime reliability is required when evaluating a noisy channel for use in a feedback link within a scalar control system trying to stabilize an unstable plant. This has subsequently been proven for scalar control over general channels in [12] and for control systems with vector valued states and linear observations in [13]. The key result that we will use here is that *there exist codes that have anytime reliability approaching the traditional random-coding error exponent even when the encoder has no access to feedback.*[11] [11]

Anytime reliability shows a clear benefit from having feedback since the anytime reliability achievable with feedback is significantly better than the best known bounds without feedback. Figure 2 shows how various bounds are related to each other for the BEC. Notice that the Burnashev exponent is infinite since zero-error transmissions are possible with variable delay. It has also recently been shown in [14] that a variation on the sphere-packing bound gives an upper bound on the anytime reliability even with feedback and this bound is significantly lower than the Burnashev bound. That bound, which is achievable for the BEC with noiseless feedback, generically has a concave $\cap$ shape like that shown in figure 2.

[10]The required rate is determined by the degree of exponential instability and the required anytime reliability depends on both the degree of exponential instability and which moment of the estimation error that we wish to keep bounded.

[11]An anytime code without feedback has a labeled infinite tree structure where branches of the tree correspond to the values of the bits to be communicated and the labels along the way correspond to the channel inputs to be sent. The decoder tries to keep track of the most likely path through the tree based on what it has received so far. By assigning the labels randomly, disjoint branches are independent of each other and using the Gallager-style union bound [5] tells us that the probability of any not-closely related path beating the true path is exponentially small. As such, the probability of error on a bit decays exponentially with delat. Such constructions also occur classically when considering infinite constraint length convolutional codes.

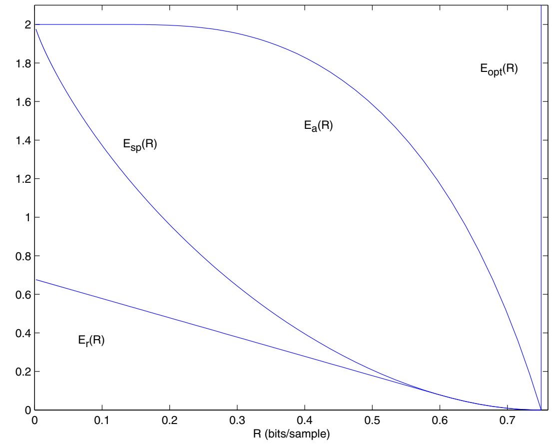

Figure 2: Random coding E_r, Sphere-packing E_{sp}, Anytime (with noiseless feedback) E_a and optimal variable-delay E_{opt} exponents for the BEC with erasure probability $\frac{1}{4}$.

IV. Variable-delay coding with noisy feedback

To understand the fundamental challenge in the case with noisy feedback, it is important to keep in mind the big picture of communication. *We assume that we have an infinite stream of bits that we want to communicate from the encoder to the decoder by repeatedly using the channels.* We choose to group the bits into packets/blocks for communication only so that we can keep the expected decoding delay finite.

Often, we information theorists like to analyze only the communication of a single packet. This is without loss of generality in the case of fixed-delay coding since it is always unambiguous to both the encoder and the decoder which block corresponds to which group of data bits. In the variable-delay case, the correspondence between received blocks and data bits becomes random due to the interactions with the channel. With instantaneous noiseless feedback, this is not a problem since the encoder always knows for certain which block the decoder is expecting and hence the single-shot analysis carries over to the repeated use scenario.

Unfortunately, it can not be taken for granted that any single-shot analysis will carry over to the repeated use case when a noisy feedback channel is involved. This is because the noise on the feedback link can make the encoder believe that the decoder has decoded K blocks correctly so far while the decoder has in fact only decoded $K' \neq K$ blocks correctly. As such, the probability of error on the next block is very high since the encoder is transmitting the wrong set of bits from the decoder's point of view![12] So the analysis of the case with noisy feedback is fundamentally interesting.

[12]In practical settings, this problem can be mitigated by the use of explicit sequence numbers on each block but theoretically we must notice that in the infinite-horizon setting, the sequence numbers will grow unboundedly in size and so the asymptotic fraction of channel uses devoted to communicating sequence numbers will go to 1 and the data rate will go to zero.

Overall, we use four key ideas to modify the scheme from [15] to work with the noisy feedback channel:

1. The phase 1 transmission is broken into sub-blocks of size K that are then "pipelined" with the decision feedback coming across the noisy feedback channel. This reduces the additional delay imposed by the noisy nature of the feedback. We only have to pay an extra K channel uses on the entire block. If we had not broken the transmission into sub-blocks, the decoder would have had to wait for the entire block before making a decision — and the encoder would have had to wait for a comparable time again before being able to see what that decision was.

2. An expurgated block code is used to feedback the decisions on the sub-blocks. This has the effect of allowing us to shrink the sub-block size and hence extra delay when the feedback channel is relatively noise free.

3. An anytime code is used to feedback the decisions regarding the confirm/deny messages. This has the effect of making sure that the encoder and decoder do not go out of synchrony with each other.

4. Blocks are interleaved with other blocks as though the feedback regarding the confirm/deny messages takes a long time to arrive. This plays a technical role in controlling the probability of error due to temporary loss of synchronization.

The formal problem, encoding and decoding procedure, and a brief sketch of the analysis are given next.

A. Problem setup Consider a streaming message transmission system as illustrated in Figure 3. The system consists of a message source that emits a randomly selected M-ary message at regularly spaced intervals of I units, a discrete-time memoryless forward channel $P = [p_{ij}]$ and a discrete-time memoryless feedback channel $Q = [q_{ij}]$. We assume that the Shannon capacity of the feedback channel is greater than or equal to the capacity of the forward channel. The problem is specified by selecting the triple (M, P, Q) and by prescribing a probability P_e. Based on that, the problem is to design an encoder $\{\mathcal{E}_n\}$ and a decoder $\{\mathcal{D}_n\}$ to minimize the average time required to transmit messages while maintaining a probability of decoding error less than or equal to P_e on each block. The transmission duration[13] for a particular message m is defined as the length of the time interval which starts at the first time information about m is emitted by the encoder and ends when the decoder produces an estimate m' of m.

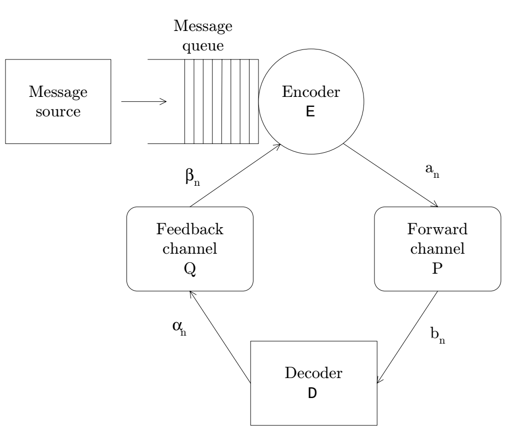

Figure 3: Conceptual block diagram of streaming message transmission system.

B. Notation

- P: Forward DMC transition probabilities p_{ij}
- Q: Feedback DMC transition probabilities q_{ij}
- M: The number of possible messages to be sent
- m: The true message is m, the decoder thinks it is m', and the encoder thinks the decoder has m''.
- N, K: The message is broken into $\frac{N}{K}$ submessages which are each transmitted using K channel uses on the forward link, and whose individual decisions are fed back using K channel uses on the feedback link.
- L: The final L channel uses are used to transmit a confirm/deny signal on the forward link, and to feedback the decoder's state on the feedback link.
- $W + 1$: The number of blocks in a single cycle before any one of them could possibly be retransmitted. Used to allow the encoder to get a good estimate of decoder state before selecting the next block of data bits to send.
- D: Number of time steps between when the decoder makes a finalization decision and when the encoder needs that decision.
- $x(n)$: The decoder's state $x(n)$ is the history of which blocks have been finalized (considered confirmed) and which blocks are awaiting retransmission. At the end of every block, the decoder's state can change by at most 1 bit — whether that block was finalized or not.
- C: The Shannon capacity of the forward channel P.
- R: The nominal rate R of phase 1 of both the forward and feedback channel use. $M = 2^{NR}$.
- E_r: The random coding reliability of the forward channel P.

[13] If we wanted to, we could analyze the total delay incurred by the block as the sum of the transmission time and the time it spent waiting in a queue awaiting transmission. However, it should be clear that the time spent waiting in the queue is negligible for much the same reason that the expected transmission duration is approximately the same as the block length.

E_{ex}: The expurgated bound for the feedback channel Q. The expurgated bound is used since it goes to infinity for almost noiseless channels while the random-coding bound can get stuck at finite reliabilities.

E_a: The anytime reliability of the Q channel when it is used without feedback.

C_1: The Burnashev exponent defined in (2).

δ: The probability of false alarm in the confirm/deny phase — that we sent a confirm but it was falsely decoded as a deny.

C. Our scheme:

Let L, N, K, W and $R < C$ be parameters to be determined later. Suppose at time n the encoder has just completed the transmission of a particular message. In the time interval $[n, n+1, \cdots, n+D]$ the decoder implicitly transmits information about its state[14] x_n over the feedback channel to the encoder. $D = W(N+K+L)$ represents the time spent in doing one cycle of the W other interleaved blocks.

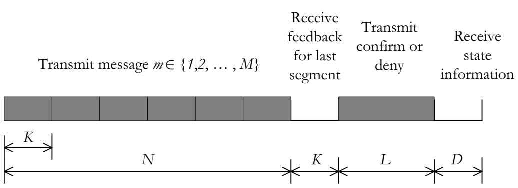

Figure 4: Encoder's timeline of a message transmission cycle with the interleaving of other cycles grouped into D.

At time $n+D$ the encoder makes an estimate of the decoder state x_n which is used to determine the message index i and begin the transmission of the message m_i. The message m_i is broken up into N/K sub-blocks encoding disjoint data bits with each segment of length K. From time $n+D$ to time $n+D+N$ the encoder emits the N/K segments to the decoder over the forward channel P by using an appropriate rate R random block code. By the time $n+D+N$, it has decoded the message as m' and the random coding bound tells us that the probability of decoding error $P_e^{(1)} = \mathbb{P}(m' \neq m)$,

$$P_e^{(1)} \leq \frac{N}{K} 2^{-E_r(R)K} \tag{4}$$

This plays the role of ϵ in the analysis of Yamamoto's scheme given earlier.

As each sub-block is received, the decoder makes a decision and feeds it back to the encoder over the feedback channel. For the feedback transmission the decoder uses a rate R block code that achieves the expurgated bound over the feedback channel.[15]

[14] The relevant decoder state x_n is simply the number of blocks that it has finalized up through time n.

[15] It is for this that we require Q to have a capacity greater than C of the forward link.

Having pipelined the N/K segments in this fashion, at time $n+D+N$ the encoder waits for an extra K time units while it receives the feedback for the last segment and at time $n+D+N+K$ it produces an estimate m'' of m'. The probability of making a mistake $P_e^{(2)} = \mathbb{P}(m = m'' \neq m')$ is bounded by $\mathbb{P}(m'' \neq m')$

$$P_e^{(2)} \leq 4\frac{N}{K} 2^{-E_{ex}(R)K}. \tag{5}$$

Finally, the encoder confirms or denies the decoder's estimate by transmitting L ones if $m'' = m$ or L zeros if $m'' \neq m$. Here we use "one" and "zero" to mean the two channel inputs[16] that achieve the maximization in (2). While this is being received on the forward link, the decoder uses a low rate $\frac{1}{L}$ anytime code to communicate information about the decoder state.

The decoder interprets the confirmation/deny message in exactly the same manner as Yamamoto's scheme and so the probability of mistakenly interpreting a deny as an accept is $\approx 2^{-C_1 L}$ while the probability of falsely viewing a confirm as a deny is δ. We then update the state of the receiver and the process repeats, but for the next block of data. This goes through all $W+1$ interleaved cycles before we return to have a possible new block or repeated transmission of the cycle we just went through.

D. Analysis: The decoder will make an error when it finalizes the block if one of the following occurs:

- The encoder sent a deny but it was interpreted as a confirm. This has probability $\approx 2^{-C_1 L}$
- The encoder thought that the decoder got it correct, but in fact the decoder got it wrong. This has probability bounded by $P_e^{(2)}$ as given in (5).
- The encoder was transmitting the wrong block.

We show that it is possible to choose parameters so that all the above are essentially equal in the exponent. The first two are easy. Set $K = (2 + \log_2 N + C_1 L)/E_{ex}(R)$. Notice that K is at most linear in L and so the number of sub-blocks is not exploding with increased block size.

To get the last term, recall that the encoder uses the anytime feedback to estimate the state of the decoder W blocks ago. The probability that the encoder selects the wrong message to transmit is therefore bounded by,

$$P_e^{(3)} \leq K' 2^{-E_a(\frac{1}{L})WL}. \tag{6}$$

So we can set $W = (\frac{\log_2 K'}{L} + C_1)/E_a(\frac{1}{L})$ and we are fine as long as $E_a > 0$ without feedback which it is known to be since it is bounded by the random coding exponent for cases without noiseless feedback. The total probability of decoding error is now $\approx 2^{-C_1 L + o(L)}$.

What remains is to bound the expected transmission duration. Notice that we accept the transmission with probability at least

[16] For a binary feedback channel, these will correspond to the usual one and zero.

$1-P_e^{(1)}-\delta$ just as before. The difference is that now the time is not distributed as a geometric random variable since the possible durations are now: $N+K+L, (W+1)(N+K+L), 2(W+1)(N+K+L), \ldots$. This has an expectation bounded by

$$(N+K+L)\left(1+\frac{P_e^{(1)}+\delta}{1-P_e^{(1)}-\delta}(W+1)\right)$$

which can be made as close to $N+K+L$ as we would like by making $P_e^{(1)}, \delta$ tiny. Putting it all together and dropping sub-exponential terms, the only thing that matters is the extra delay of $K \approx \frac{C_1 L}{E_{ex}(R)}$ and so after some simplification, our reliability lower bound becomes:

$$E_{noisy}(\bar{R}) \approx (\frac{1}{C_1}+\frac{1}{E_{ex}(R)})^{-1}\left(1-\frac{\bar{R}}{C}\right),\ 0 \le \bar{R} < C, \quad (7)$$

where E_{ex} is the expurgated reliability for the feedback channel. Notice that the new constant $(\frac{1}{C_1}+\frac{1}{E_{ex}(R)})^{-1}$ is one-half the harmonic mean between C_1 and $E_{ex}(R)$.

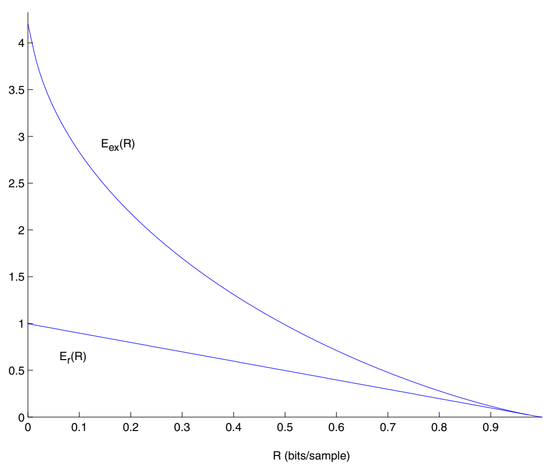

Figure 5: Illustration of the random coding E_r and expurgated E_{ex} exponents for a BSC with crossover probability 10^{-6}

As the feedback channel becomes less and less noisy, the expurgated exponent[5] will go to ∞ for rates less than the capacity. e.g. Figure 5 shows how the expurgated bound looks for the BSC with very small crossover probability. As such, the bound in (7) will approach (1).

V. Comments and Extensions

This paper was focused on showing that for the reliability of variable-delay decoding, as the feedback quality gets better, the reliability will approach the Burnashev bound for noiseless feedback. As such, we did not focus on getting the best possible reliability for the noisy feedback case.

With some additional work, (5) can be improved to $P_e^{(2)} \approx \frac{N}{K}2^{-(E_f(R)+1)K}$. This results in replacing E_{ex} in (7) with (E_f+1) where E_f is the Forney bound. To get this, we interpret the feedback message occasionally as an erasure as this only causes us to increase slightly the probability of retransmission. It is further improved by noticing that (5) is overly conservative since we interpret all errors as causing us to falsely believe that the submessage was correctly received. There are actually 2^K possible places we could have landed where only one would cause an undetected error.

It is also curious to see what happens in the other limit — as the feedback becomes more and more noisy. There, we can apply a similarly interleaved Forney scheme on the forward link and just use anytime feedback to maintain synchrony. This will result in the Forney bound E_f being achievable even with very noisy feedback. This gives rise to a surprising prospect — the limit of reliability with very noisy feedback might be different than the feedback-free case!

Acknowledgments

The authors would like to thank Bixio Rimoldi for introducing them to Burnashev's paper, as well as David Tse and Kannan Ramchandran for initial discussions on the nature of the Burnashev bound and the nature of variable delay decoding. An anonymous reviewer also gave very helpful comments on the presentation.

References

[1] M. V. Burnashev. Data transmission over a discrete channel with feedback, random transmission time. *Problemy Perdachi Informatsii*, 12(4):10-30, 1976.

[2] T. Cover and J. Thomas, *Elements of Information Theory*. John Wiley, New York, 1991

[3] R. L. Dobrushin, "An asymptotic bound for the probability error of information transmission through a channel without memory using the feedback." Problemy Kibernetiki, Vol 8, pp 161-168, 1962.

[4] G.D. Forney, "Exponential error Bounds for Erasure, List, and Decision Feedback Schemes," IEEE Trans. on Information Theory, Vol 14, pp 206-220, 1968.

[5] Robert Gallager. *Information theory and reliable communication*. John Wiley & Sons, New York, NY, USA, 1968.

[6] M. Horstein, "Sequential transmission using noiseless feedback." IEEE Trans. on Information Theory, Vol 9, pp 136-143, 1963.

[7] J.M. Ooi, *Coding for channels with feedback*. Kluwer Academic Publishers, Boston, MA, USA, 1998

[8] J. P. M. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback – I: No bandwidth constraint." IEEE Trans. on Information Theory, Vol 12, pp 172-182, April 1966

[9] A. Sahai, ""Any-time" Capacity and A Separation Theorem For Tracking Unstable Processes", ISIT 2000.

[10] A. Sahai, "Evaluating Channels for Control: Capacity Reconsidered." Proceedings of the 2000 American Control Conference, Vol 4 pp2358–2362, 2000

[11] A. Sahai, ""Any-time" Information Theory, PhD Dissertation, MIT 2001.

[12] A. Sahai, "The necessity and sufficiency of anytime capacity for scalar control over a noisy communication link." Accepted at the 2004 Conference on Decision and Control.

[13] A. Sahai, "The necessity and sufficiency of anytime reliability for control over a noisy communication link: Parts I and II." In preparation.

[14] T. Simsek, Any-time Channel Coding with Feedback, PhD Dissertation, UC Berkeley, 2004.

[15] Hirosuke Yamamoto and Kohji Itoh. Asymptotic Performance of a Modified Schalkwijk-Barron Scheme for Channels with Noiseless Feedback. *IEEE Transactions on Information Theory*, IT-25(6):729–733, 1979.