Technical Research Report

Adaptive Forward Error Correction for Channels with Feedback

by Bernhard G. Dorsch

T.R. 86-67

INSTITUTE FOR SYSTEMS RESEARCH

Sponsored by
the National Science Foundation
Engineering Research Center Program,
the University of Maryland,
Harvard University,
and Industry

TR-86-67

ADAPTIVE FORWARD ERROR CORRECTION FOR CHANNELS WITH FEEDBACK

by

Bernhard G. Dorsch

Inst. f. Netzwerk u. Signal-Theory
Techn. Hochschule Darmstadt
W. Germany

Visiting Scientist at the Systems Research Center of the
University of Maryland

August 1986

ABSTRACT

Forward error correction (FEC) without feedback always has to transmit maximum redundancy for the worst case of errors to be corrected. With feedback, the redundancy can be adapted to actual channel disturbances. For Maximum-Distance-Separable (MDS) algebraic block codes, including Reed-Solomon (RS) codes, methods are described and analyzed, where parity symbols are transmitted only until enough redundancy is received to correct and detect reliably all errors and erasures in the transmitted sequence. The expected coderate is shown to be much higher than without feedback, even close to capacity for a variety of channels with and without memory, and reasonably better than the cutoff rate estimate for FEC. The method proves to be very robust against the length distribution of error bursts.

1. Introduction

Pure Forward Error Correction (FEC) is used for channels where feedback is either not available or not advisable, as, for example, due to a long time delay in deep space data transmission. Good results with FEC are achieved when the disturbances are fairly constant, for example Additive White Gaussian Noise (AWGN). But near earth channels often are dominated by heavy time varying noise or error bursts, often with unpredictable time behavior, caused for example by multipath effects, interference, manmade noise, radar, hostile jammers, etc. In such cases, a feedback signal from the receiver to the transmitter telling whether a message was received reliably, eventually requiring some kind of retransmission may be essential for high reliability and data throughput.

The simplest way of using feedback for error control is Automatic Repeat Request (ARQ) as shown in fig. 1a. A string i of information data together with some few parity symbols p for error detection only are transmitted. When errors are detected, retransmission is required. The block length has to be adjusted carefully to the channel. Long blocks may almost always have some errors somewhere, requiring many retransmissions and a poor data throughput. For short blocks, the header (for control data like address, frame synchronization, etc.) decreases the data throughput.

As an alternative philosophy, some commercial terminals for mobile communications transmit more redundancy (e.g., a code of rate 1/2) in order to correct some few "typical" errors by FEC as shown in fig. 1b. Longer codes of reasonable rate also are able to detect reliably when more errors occurred than the code can correct, even if the full error correcting capability of the code is used for correction. But for bursty channels, the redundancy transmitted for FEC often is wasted because either no errors occur in clean parts of the channel or the redundancy is not enough to correct long error bursts.

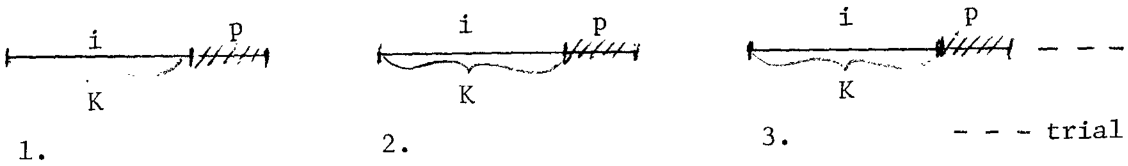

Fig. 1a: ARQ

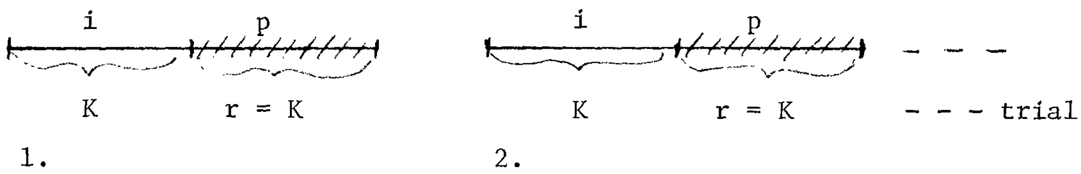

Fig. 1b: ARQ + FEC, code rate R = K/(K+r) = 1/2

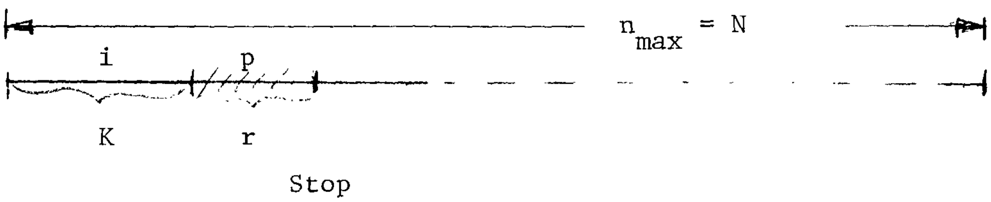

Fig. 1c: Adaptive FEC, continuous transmission

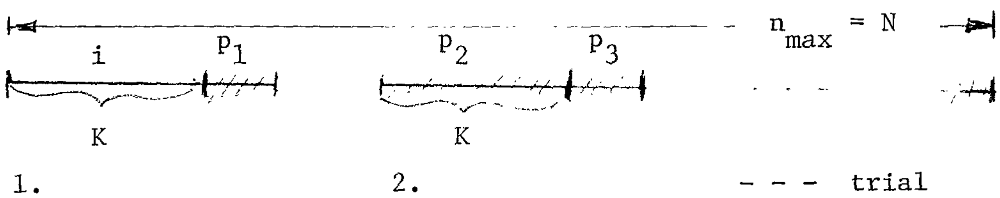

Fig. 1d: Adaptive FEC, packet transmission

Fig. 1: Error control philosophies for channels with feedback.

Methods which combine some advantages of ARQ and FEC were developed, e.g., by S. Lin [1],[23], D. Chase [3].

In this paper, a method for Adaptive FEC with algebraic blockcodes will be described and analyzed. The basic principle [4] is shown in fig. 1c. First, all K information symbols of a relatively long blockcode are transmitted followed by a variable number r of parity symbols. In the limiting case of zero-delay, error-free feedback with continuous data transmission, the receiver tries to decode the received sequence of length K + r after each received parity symbol. If all errors can be corrected reliably, the receiver causes the transmitter to stop transmission of more parity symbols and to start transmitting the next codeword. So the number of transmitted parity symbols is adapted to the noise in each codeword resulting in a small average number K + r of symbols per codeword. Especially appropriate for such a scheme are MDS codes, inclduing Reed-Solomon (SR) codes, which have optimal error correcting and detecting capability for any K + r received symbols of a codeword. We will use RS codes with elements from a binary extension field $GF(2^s)$ where each symbol is an s-bit-byte. An error is defined as a wrong symbol (with any number of wrong bits) at an unknown position. Erasures are unreliable symbols at known positions defined by the receiver, based, for example, on the received signal level, indicated by an Automatic Gain Control (AGC) signal, or any other "side information" from the front end receiver, bit synchronizer, or demodulator. Long error bursts usually can be detected reliably as erasures. Within K + r received symbols, up to $a \leq r$ erasures can be corrected together with t errors as long as $a + 2t \leq r$. In other words, two parity symbols are necessary for each error to be corrected but only one for each erasure. If $a + 2t > r$, the decoder may make an undetected decoding error with probability P_u. With probability $(1 - P_u)$ such disturbances are detected as uncorrectable resulting in transmission of more parity elements. As shown in [5],[6], for a symbol length $s \geq 6$ the undetected error probability is $Pu < 10^{-6}$ for $r - a > 10$. For $r - a \leq 10$, only one error less up to $t \leq (r - a)/2 - 1$ errors should be corrected to achieve a small $P_u < 10^{-6}$, i.e., two additional parity symbols are used for reliable error detection in this case. The maximum codelength of RS codes is $K + r \leq 2^s - 1 = :N$ (extendable to 2^s). We will investigate two schemes, one in which the number $n = K + r$ of transmitted symbols per codeword is limited to the maximum codelength $n \leq N$, with a certain block error probability Pe (when $n > N$ symbols would be necessary), and a second scheme in which a new codeword is started and combined with the previous ones until the message is decoded successfully, n limited to any number $L < \infty$.

In practical applications and packet transmission the scheme may be modified as indicated in fig. 1d. For example, with a symbol length of $s = 6$ bits (maximum codelength $N = 127$ symbols) first $K = 40$ information symbols together with 2 parity symbols for error detection are transmitted in a first transmission trial. If errors are detected, a second block of 40 + 2 check symbols is transmitted. In a cyclic block code, the next symbol at any position always has the same linear dependence from the previous K symbols. Therefore, the last 2 symbols in the second block can be used for error detection in the same way as in the first block. If no errors are detected in the second block, the information is reconstructed by regarding the first block as erased or by "inversion" [1]. If errors are detected t errors plus a erasures can be corrected in the first two blocks together, if $2t + a \leq r = 2 + 40 + 2$ resp. $2(t - 1) + a \leq 44$ for $44 - a \leq 10$. Otherwise, a third block with 40 + 2 parity symbols will be transmitted, again first with error detection within the third block (regarding the first two blocks as erased, if no errors are detected in the third block) resp. correction of $2t + a \leq 2 + 40 + 2 + 40 + 2$ errors resp. erasures in all three blocks.

After describing the channel assumptions by simple models used to represent channels with error bursts in section 2, we discuss an analytical method to investigate the performance of the proposed Adaptive FEC for the continuous case, fig. 1c, with n limited to $n \leqslant N$. In section 3.2, a transmission scheme with n limited to any arbitrary number L is described in more detail. Numerical results for the expected average number $\overline{n}$ of transmitted symbols, resp. the code rate $R := K/\overline{n}$, and the block error probability $P_e := Pr(n > L)$ are discussed in chapter 5 and compared to theoretical limits and estimates, which are explained in section 4.

2. Channels.

Feedback is essential especially for channels with memory (error bursts) for reliable data transmission. The simplest model for such a channel is a two-state first-order Markov chain as shown in Fig. 2, often referred to as Gilbert-Elliott (GE) model (which might or might not represent the real world, but is a model which is simple enough to be handled mathematically). The GE model has two channel states, z = G ($\hat{=}$ good) with a fairly low bit error probability p_G, and z = B ($\hat{=}$ bad, burst) with a considerably higher $p_B >> p_G$. For RS codes (with symbols of length s-bits), we regard the transition probabilities $P(z_j = B/z_{j-1} = G) = g$, resp $P(z_j = b/z_{j-1} = B) = b$ as symbol-transition probabilities. This seems to be a valid assumption if the average burst-length (number of successive bits in the bad state) is long compared to the symbol length s.

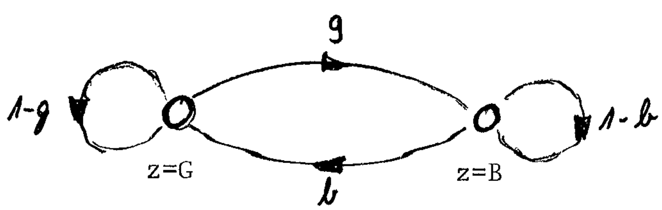

Fig. 2: A simple model for channels with memory.

The channel has the well known following properties:

The average burst length is

$$i_B = \sum_{i=1}^{\infty} i\cdot(1 - b)^{i-1}\cdot b = \frac{1}{b} \text{ symbols} \tag{1}$$

The conditional probability that state $z_{i+\ell}$ will be $z = G$ if $z_i = B$ (e.g., for ℓ-symbol interleaving) is

$$b_\ell = \frac{b}{b + g} \cdot \left[1 - (1-b-g)^{\ell}\right]$$

$$g_\ell = \frac{g}{b + g} \cdot \left[1 - (1-b-g)^{\ell}\right] \tag{2}$$

with $b_1 = b$, $g_1 = g$, and

$$g_\infty = \frac{g}{b + g} = :\delta \ , \quad b_\infty = \frac{b}{b + g} = 1 - \delta \tag{3}$$

as absolute probabilities of being in the bad (resp. good) state. If each of the two states is regarded as a Binary Symmetric Channel (BSC), the symbol error probabilities for $z = G$ resp. $z = B$ are

$$u_z = 1 - (1 - p_z)^s \tag{4}$$

What a bad or good state is may be defined by the receiver, based, for example, on the received signal strength, the number of detected errors in the sync pattern, the variance of the soft decision demodulater output, phase-jitter in the Phase-Locked-Loops (PLL) of the receiver, noise level in neighbor frequency bands, or any other "side information". For error-and-erasures-correction with RS codes, we regard signals in the bad state as erased. The good state still may have errors besides correct symbols. For error-correction-only, the channel state is assumed not to be known (only the channel parameters p_z resp. u_z for $z = G/B$, b and g).

One of the most serious causes for error bursts in some practical channels (as, e.g., in car mobile, ship to satellite, aircraft to satellite communication) are multipath effects where the received signal is the sum of many reflected paths of the transmitted carrier frequency, each with random amplitude and phase, resulting in a Rayleigh distribution of the received amplitude (if no path dominates). With Additive White Gaussian Noise (AWGN), antipodal modulation and coherent demodulation, e.g., coherent binary phase modulation (BPSK), the demodulated binary signals $r_{1/2}$ (for transmitted binary signals $s_1 = +1$, $s_2 = -1$) are (perfect syndromization assumed !)

$$r_k = a \cdot s_k + n \ , \quad k = 1,2 \tag{5}$$

where a has a Rayleigh distribution with probability density function

$$f(a) = 2a \cdot \exp(-a^2) \ , \quad a \geqslant 0 \tag{6}$$

normalized on an expected value $E[a^2] = 1$. The additive noise n is zero-mean Gaussian with variance $\sigma^2 = N_o/(2E_s)$, N_o = one-sided spectral noise power density, E_s = average received signal energy per bit. Usually, the received signal amplitude a can be well estimated by the receiver if the fading bandwidth is small compared to the data bandwidth, indicated, for example, by the Automatic Gain Control (AGC). If binary signals $s_{1/2}$ are transmitted with equal probability 1/2 the optimum decision rule for the demodulated bit is $\hat{s}_k = \text{sign } r_K$. The bit error probability conditioned on a given signal amplitude a is

$$p(a) = Q(a \cdot \sqrt{2\gamma_o}) \ , \quad \text{with} \quad \gamma_o : = E_s/N_o \tag{7}$$

and

$$Q(x) := \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2}\, dt \tag{8}$$

resulting in an average bit error probability

$$p = \int_0^\infty f(a)p(a)da = \frac{1}{2} \cdot (1 - \sqrt{\gamma_o/(1+\gamma_o)}) \tag{9}$$

We define here a "bad" state $z = B$ by a weak signal strength $a \leq a_T$, where a_T is a (optimized) threshold, $z = G$ for $a > a_T$. The probability δ of being in state $z = B$ then is

$$\delta = \int_0^{a_T} f(a)da = 1 - \exp(-a_T^2) \tag{10}$$

The conditioned bit error probabilities p_G, p_B in state $z = B/G$ are

$$p_G = \frac{1}{1-\delta} \int_{a_T}^\infty f(a) \cdot p(a)da = \frac{1}{2 \cdot \exp(-a_T^2)} \left[\exp(-a_T^2) \cdot Q(a_T\sqrt{2\gamma_o}) - \sqrt{\gamma_o/(1+\gamma_o)} \cdot Q(a_T \cdot \sqrt{2+2\gamma_o})\right] \tag{11}$$

$$p_B = \frac{1}{2 \cdot (1-\exp(-a_T^2))} \left[1 - \sqrt{\gamma_o/(1+\gamma_o)} - \exp(-a_T^2)Q(a_T\sqrt{2\gamma_o}) + \sqrt{\gamma_o/(1+\gamma_o)}Q(a_T\sqrt{2+2\gamma_o})\right] \tag{12}$$

These probabilities do not express yet how long fades are. If we want to model the distribution of burst-lengths by a GE model as in Fig. 2, we assume a certain average burst length i_B and calculate the transition probabilities b, resp. g by (1) resp (3) with (10).

2. Reliable error detection.

RS codes with K information and $r = n - K$ parity symbols can correct up to $a \leq r$ erasures, where the number a and the positions of the erasures are known to the receiver. For $a < r$ in addition up to

$$t = \lfloor (r-a)/2 \rfloor \tag{13}$$

errors can be corrected corresponding to error-correction-only in a shortened code of length $n' = K + r - a$. If only up to $t' \leq t$ errors are corrected, $t \geq t'$ errors are always detected as "undecodable" as long as $t \leq r - a - t'$. If the number of transmission errors t in the n' non-erased positions is $t > r - a - t'$, then the probability for undetected decoding errors (conditioned under $t > r - a - t'$) is bounded by

$$P_u \leq \sum_{i=0}^{t'} \binom{n'}{i}(2^s - 1)^i / 2^{s(r-a)} \tag{14}$$

as shown by Kressel [6], based on [7], [5]. For $s \geq 6$ and $r - a > 10$, for example, (14) results in $P_u < 10^{-6}$ for $t' = t$ (i.e., even if all errors up to the full error correcting capability (13) are corrected). For $(r-a) \leq 10$, $s \geq 6$, only up to $t' = t - 1$ errors should be corrected for a comparable low limit of P_u in (14). This is equivalent to using two of the remaining (r-a) parity symbols not for error correction but for additional error detection only. In the simulations $t' = t$ for $(r-a) > 10$ and $t' = t - 1$ for $(r-a) \leq 10$ was used. In the calculations of section 3, $t' = t$ and $t' = t - 1$ results in upper and lower limits R_u and R_ℓ for the adaptive code rate R.

3. Code rate, block error probability for $n \leq 2^s - 1$.

If the number of transmitted symbols n is limited to $n \leq 2^s - 1 = N$ (the maximum length of an RS codeword with symbols from $GF(2^s)$), the expected number $\bar{n}$ and the corresponding code rate $R = K/\bar{n}$ as well as the remaining block error probability P_e (the probability that $n > N$ transmissions would be needed) can be calculated by recursion even for channels with memory (represented, e.g., by a GE model, Fig. 2).

Let us begin with the assumption $t' = t$. After K information symbols $r = 0,1,\ldots,r_{max} = 2^s-1-K$ parity symbols are transmitted. Let t_r resp. a_r be the number of errors resp erasures in the first $n = K + r$ transmitted symbols and $i_r := a_r + 2 \cdot t_r - r$. Because all errors and erasures are corrected, if $a_r + 2t_r \leq r$, exactly $n = K + r$ symbols are to be transmitted, if $i_r = 0$ and $i_j > 0$ for all $j < r$. The probabilities $P(i_j, z_j)$ can be calculated from the $P(i_{j-1}, z_{j-1})$ by recursion. If the j-th symbol s_j is correct ($\hat{=}$c), erased ($\hat{=}$a) or in error ($\hat{=}$e)

$$i_j = i_{j-1} \begin{cases} -1 & \\ \pm 0 & \text{for } s_j \\ +1 & \end{cases} \begin{matrix} = c \\ = a \\ = e \end{matrix} \tag{15}$$

With known probabilities $P(s_j, z_j/z_{j-1}) = P(z_j/z_{j-1}) \cdot P(s_j/z_j)$, where $P(z_j/z_{j-1})$ are the transition probabilities of channel states and

$$\begin{matrix} P(s_j=c/z_j=G) = 1-u_G & P(s_j=c/z_j=B) = 0 \\ P(s_j=a/z_j=G) = 0 & P(s_j=a/z_j=B) = 1 \\ P(s_j=e/z_j=G) = u_G & P(s_j=e/z_j=B) = 0 \end{matrix} \Bigg\} \text{ if } Z_j \text{ known resp. } \begin{matrix} = 1 - u_B \\ = 0 \\ = u_B \end{matrix} \Bigg\} \text{ if } z_j \text{ unknown} \tag{16}$$

the corresponding joint probabilities $P(j_j, z_j)$ are

$$\begin{aligned} P(i_j, z_j) = \sum_{z_{j-1}=G,B} \cdot \{ & P(i_{j-1} = i_j - 1, z_{j-1}) \cdot P(s_j = c, z_j/z_{j-1}) \\ & + P(i_{j-1} = i_j, \ z_{j-1}) \cdot P(s_j = a, z_j/z_{j-1}) \\ & + P(i_{j-1} = i_j + 1, z_{j-1}) \cdot P(s_j = e, z_j/z_{j-1}) \} \end{aligned} \tag{17}$$

setting $P(i_{j-1} = 0, z_{j-1})$ to zero before, because transmission is assumed to be finished in this case after $r = j - 1$ parity symbols. The probability for exactly $r = j$ parity symbols to be transmitted is $P(i_j = 0) = P(i_j = 0, z_j = G) + P(i_j = 0, z_j = B)$.

The remaining block error probability $P_e = P(i_{r_{max}} > 0)$ where $r_{max} = 2^{s-1} - K$ parity symbols are not sufficient to correct all errors and erasures is about the same as for FEC with a fixed rate $R = K/(K + r_{max})$, but slightly smaller, because Adaptive FEC would be successful in (rare) cases where most of the uncorrectable errors for FEC are concentrated at the very end of the code word.

The expected number of parity symbols, therefore, is

$$\bar{r} = \sum_{j=0}^{r_{max}} j \cdot P(i_j = 0) + P_e \cdot r_{max} \qquad (18)$$

The corresponding upper limit of the code rate yields

$$R_u = K/(K + \bar{r}) \qquad (19)$$

A lower limit R_ℓ of the achievable code rate can be given by the assumption that in each case only up to $t' = t - 1$ errors will be corrected, corresponding to two additional parity symbols for reliable error detection in each case, yielding

$$R_\ell = K/(K + \bar{r} + 2) \qquad (20)$$

Numerical results are given together with the simulation results for an arbitrary limitation of $n \leqslant L$, any $L < \infty$ in section 5.

3.2 Limitation $n \leqslant L$, any $L < \infty$.

The scheme described in section 3.1 with n limited to $n \leqslant N = 1^s - 2$ (the maximum in length of an RS-code) may be appropriate for applications with a limited number of transmission trials. For example, in car mobile communications with packet transmission according to Fig. 1.d, if some few transmission trials are not successful, the reason may be some shadowing effect, and transmission should be interrupted rather than occupying the channel without success. In this chapter, we investigated the performance of a scheme where the codeword symbols are repeated cyclically until the message is transmitted successfully. The encoder just goes on as it does for $n \leqslant N$. By using a linear feedback shift register of length K, based on the check polynomial $h(x): = (x^N-1)/g(x)$ ($g(x)$ = generation polynomial), each cyclically following symbol of a codeword is generated from the previous K symbols in the same way. The decoder stores up to N received symbols, $\hat{s}_1,\ldots,\hat{s}_N$ beginning with all N symbols marked as "erased", a = N erasures. When the j-th transmitted symbol s_j is received, $\hat{s}_{j\,\mathrm{mod}N}$ is replaced by s_j, if it was marked as erased before. If $\hat{s}_{j\,\mathrm{mod}N}$ was not marked as erased and s_j is not erased, but $s_j \neq \hat{s}_{j\,\mathrm{mod}N}$ (i.e., if either s_j or $\hat{s}_{j\,\mathrm{mod}N}$ is in error, $\hat{s}_{j\,\mathrm{mod}N}$ will be marked as erased; otherwise (for $s_j = \hat{s}_{j\,\mathrm{mod}N}$) remains unchanged. As explained before, all errors and erasures can be corrected as soon as $a + 2t \leqslant N - K$ (a = number of erasures, t = number of errors in the codeword), resp $t \leqslant (N-K-a)/2 - 1$ if $N - K - a \leqslant 10$, $s \geqslant 6$ (for a small probability P_u for undetected decoding errors).

The probabilities $P(n = j)$ that $n = j$ symbols per codeword are sufficient were estimated by computer simulations. The probability $P(n > L)$ represents the probability of transmission failures if n is limited to $n \leqslant L$. The average number of transmitted symbols is $\overline{n}(L) = \sum_{j=K+2}^{L} j \cdot P(n = j) + L \cdot P(n > L)$, approaching asymptotically a limit $\overline{n}(L = \infty)$. How the coderate $R(L): = K/\overline{n}(L)$ and $P(n > L)$ depend on the channel and code parameters will be discussed in section 5.

4. Channel Capacity, Cutoffrate R_0.

Before the results of Adaptive FEC will be discussed, we derive some useful limits to compare with. Channel capacity is the absolute limit for coderates R (with small block error probability), whereas R_0 is regarded as a more practical estimate of achievable coderates R, [8] (but neither an upper nor a lower limit!). This is considered to be true for pure FEC without feedback. But with feedback and Adaptive FEC, it will be shown that R gets close to the appropriate capacity C but is much better than the corresponding R_0 of the forward channel.

If the state z = G/B is known to the receiver, and therefore part of the received information, channel capacity C is defined as

$$C = \sum_{x} P(x) \sum_{y,z} P(y,z/x)\log_2 \frac{P(y,z/x)}{P(y,z)} \tag{22}$$

where the alphabet size for the transmitted signal x is 2^s, but $2^s + 1$ for the received signal y (including y = erasure). For a constant discrete memoryless channel (CDMC) in each state z = G/B, (20) results in $C = (1-\delta)\cdot C_G + \delta\cdot C_B$, where δ is the probability of z = B and $C_{G/B}$ the conditional channel capacity for z = G/B. If the channel state z is known, notice that C does not depend on the memory in the transition of channel states. That means C is independent of the lengths and distribution of error bursts for δ = const. It also is the same for channels with and without feedback. But that is not the case for cutoff rate R_o, defined as

$$R_o = -\log_2\{\sum_{y,z} [\sum_{x} P(x)\sqrt{P(y,z/x)}\,]^2\} \tag{23}$$

(23) represents R_o for the forward channel without feedback and may be different for the channel with feedback (which we have not calculated yet).

What are the probabilities P(y,z/x) (defining the forward channel) we have to look at in order to get an appropriate measure for the performance of the proposed transmission scheme? There are at least 3 channels, i.e., sets of $P(y,z/x) = P(z)\cdot P(y/x,z)$ for which RS-codes with error and erasure correction perform exactly the same.

(a) Each z = G/B corresponding to a BSC with bit error probability p_z and

$$P(y/x,z) = p_z^i(1-p_z)^{s-i} \tag{24}$$

for each binary error pattern (y-x) of weight i. Let's label this channel as "Bit" channel, with $C = :C_{ps}$, $R_o = :R_{ps}$.

(b) a channel, where each error pattern $(y-x) \neq 0$ has the same probability

$$P(y/x,z) = \begin{cases} 1 - u_z & \text{for } y = x \\ u_z/(2^s-1) & \text{for } y \neq x \end{cases} \tag{25}$$

Let's label this channel as "symbol" channel, with $C = :C_{us}$, $R_o = :R_{us}$.

(c) a channel, where each y is regarded as erased for $z = B$: $P(y/x,z = G)$ the same as in (25), but $P(y = \text{erasure}/x,z = B) = 1$, labeled as "Erasure" channel, $C = :C_{\delta s}$, $R_o = :R_{\delta s}$.

It is not surprising that the channels become "worse" in this sequence; that means that $C_p \geq C_{us} \geq C_{\delta s}$, $R_{ps} \geq R_{us} \geq R_{\delta s}$. In order to get the closest estimate of limits for the performance of a coding scheme, we should refer to capacity, resp. R_o, of the "worst" channel for which a specific coding/decoding scheme performs exactly the same, in this case $C_{\delta s}$ and $R_{\delta s}$. The $P(y/x,z)$ defined above result in

$$C_{ps} = C_{pl} =: C_p = (1-\delta) \cdot (1-H_2(p_G)) + \delta \cdot (1-H_2(p_B)) \tag{26}$$

where $H_2(x) = -x\log_2(x) - (1-x)\log_2(1-x)$ is the binary entropy function. All C and R_o are given in "bit per transmitted binary signal."

$$C_{us} = (1-\delta) \cdot C_{usG} + \delta \cdot C_{usB} \tag{27}$$

with

$$C_{usz}: = \left[s-H_2(u_z) - u_z \cdot \log_2(2^s-1)\right]/s \tag{28}$$

and, finally,

$$C_{\delta s} = (1-\delta) \cdot C_{usG} \tag{29}$$

The corresponding cutoff rates R_o are

$$R_{ps} = \{s - \log_2[(1-\delta)\cdot(1 + 2\sqrt{p_G(1-p_G)})^s + \delta(1 - 2\sqrt{p_B(1-p_B)})^s]\}/s \qquad (30)$$

Notice that (in contrast to capacity)

$$R_{ps} \leqslant (1-\delta)R_{psG} + \delta\cdot R_{psB} \quad \text{with}$$

$$R_{psG} = 1 - \log_2(1 + 2\sqrt{p_G(1-p_G)}) = R_{ps}(\delta=0)$$

$$R_{psB} = 1 - \log_2(1 + 2\sqrt{p_B(1-p_B)}) = R_{ps}(\delta=1)$$

because log(x) is a convex function of x. Special cases:

$s = 1$, $\delta = 0$, $p = p_G$ (errors only): $R_{p1} = 1 - \log(1 + 2\sqrt{p(1-p)})$

$s = 1$, $p_G = 0$, $p_B = 0.5$ (erasures only): $R_{p1} = 1-\log(1 + \delta)$

For the "symbol" channel R_o yields

$$R_{us} = \{s - \log_2[(1-\delta)v_{sG} + \delta\cdot v_{sB}]\}/s \qquad (31)$$

with $v_{sz} = 1 + u_z\cdot(2^s-2) + 2\sqrt{u_z(1-u_z)(2^s-1)}$ for $z = G/B$.

Finally, for the "Erasure" channel,

$$R_{\delta s} = \{s - \log_2[(1-\delta)v_{sG} + \delta\cdot 2^s]\}/s \qquad (32)$$

For Rayleigh-fading, as described in section 2, the maximum cutoff rate $R_{o,max}$ is achieved for a channel where in eq. (5) the value r_k ("soft decision demodulator output") as well as the signal strength a are known exactly to the receiver. For this case, [9]

$$R_{o,max} = 1 - \log_2[1 - 1/(1+\gamma_o)] \qquad (33)$$

often is referred to as "R_o of a Rayleigh-fading channel."

5. Performance.

In this chapter, we discuss numerical results for the expected number $\overline{n} = K + \overline{r}$ of transmitted symbols, resp. the average coderate $R(L) = k/\overline{n}$ (n limited to a maximum of L transmitted symbols per codeword) and the block error probability $P_e(L) = P(n > L)$, i.e., the probability that $n \leq L$ transmitted symbols are not enough to correct and detect all errors and erasures reliably. For $L = 2^s - 1 = N$ (the maximum codeword length) $R(L)$, $P_e(L)$ were calculated analytically by using eqs. (17),(18). For all L (including $L < N$ and up to $L = 20000$), computer simulations were run, the results of both methods confirming each other. How $R(L)$, $P_e(L)$ depend on channel and code parameters and compare to theoretical limits will be discussed in sections 5.1 through 5.5.

5.1. Influence of burst-length distribution.

The proposed scheme mainly was designed to cope with heavy error bursts, often of unpredictable length and distribution. One of the most interesting questions, therefore, is: How do $R(L)$, $P_e(L)$ depend on the burstlength distribution in the GE-model, fig. 2? The overall probability of being in bad state $z = B$ is $\delta = g/(g+b)$, the average burstlength $i_B = 1/b$. Numerical results for a typical example with $\delta = 40\%$, symbol error rate $u_G = 0.03$ in good state $z = G$ (erasures in bad state $z = B$, i.e., $p_B = 0.5$) are given in fig. 3, for a code with a symbol length of $s = 6$ bits and $K = 20$ information symbols. Beginning with statistically independent state transitions (i.e., $g = \delta = 0.4$, $b = 1 - g = 0.6$, $i_B = 1/b = 1.7$), average burstlengths i_B up to 333 symbols (i.e., more than five times the maximum codelength $N = 2^s - 1 = 63$) are regarded. For a reasonable limitation L such that $P(n > L)$ is small (say, $<10^{-3}$), $R(L) \approx R(\infty)$. As one of the most surprising results, the average number of symbols per codeword $\overline{n}$ $(L = \infty)$, resp. $R = K/\overline{n}$, is seen to be almost independent of the burstlength i_B. This turned out to be true also for other

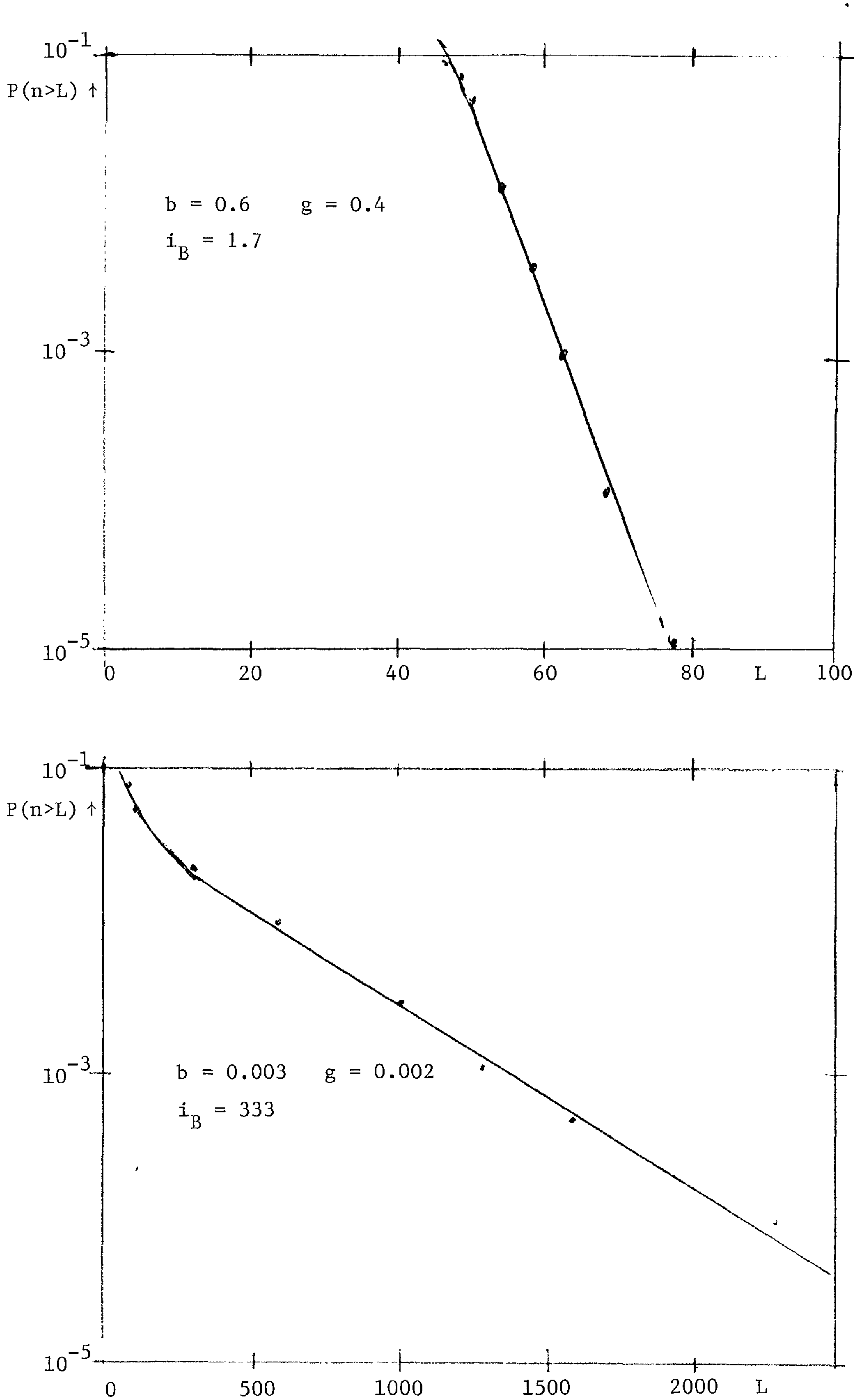

Fig. 4: P(n>L) vs L for i_B = 1.7 and 333.

investigated values of δ and u_G.

But, of course, i_B influences the probability distribution $P(n=j)$ resp. $P_e(L) = P(n > L)$. If i_B becomes larger, $P(n = j)$ increases for larger j. That means in long bursts much more symbols are to be transmitted. But we then also often have long strings of symbols in the good state (for δ = constant), where only a small number n of symbols is required. For example, $P(n = K+2 = 22 = \text{min. } n) \approx 10^{-5}$ for $i_B = 1.7$ but 45% for $i_B = 333$. On the other hand, $P(n > 200) \approx 0$ for $i_B = 1.7$ but $\approx 2.5\%$ for $i_B = 333$.

g	b	i_B	$n(\infty)$	$R(\infty)$	P(n=22)	P(n>80)	P(n>200)	P(n>2800)
.4	.6	1.7	38.9	.514	$\approx 10^{-5}$	≈ 0	–	–
.02	.03	33	39.4	.507	.25	$\approx 10^{-2}$	$\approx 10^{-5}$	≈ 0
.002	.003	333	39.6	.505	.45	$4 \cdot 10^{-2}$	$2.5 \cdot 10^{-2}$	$\approx 10^{-5}$

Fig. 3: Influence of average burstlength i_B on $n(\infty)$, $R(\infty)$, $P(n>L)$.

Error exponent E: For most coding schemes with FEC, the block error probability P_e usually decreases almost exponentially with codelength N, i.e., $P_e(n) \approx 2^{-N \cdot E(\ldots)}$, where $E(\ldots)$ depends on the channel and code parameters but not on N (for large N). $E(\ldots)$ is called the "error exponent". Figure 4 shows $\log P_e(L) = \log P(n > L)$ also depending about linearly on L, corresponding to a constant error exponent E (independent of L). Of course, $E(i_B)$ decreases with increasing i_B (i.e., becomes worse).

5.2. Influence of Code Parameters.

With straight FEC without feedback, the code parameters must carefully be adapted in order to achieve a low error probability P_e and a high coderate R (data throughput). But Adaptive FEC will be shown to adapt to the channel very independently of the code parameters s and K if they are chosen within a wide, reasonable range.

Figure 5 shows the coderate $R = R(L) \approx R(\infty)$ (for reasonable limitations L) for s = 6 (max. codelength N = 63) and s = 8 (N = 255) and various K (number of information symbols). If K is only a small multiple of 2, then the $r \geq 2$ parity symbols which are always necessary for reliable error detection, decrease the coderate. If K becomes close to N, the code does not have much error correcting capability anymore resulting in a poor coderate. But for $2 \ll K \ll N$ the resulting coderate is seen to depend only weakly on K and s.

s	K	$R(\infty)$
6	10	.47
6	20	.51
6	40	.45

s	K	$R(\infty)$
8	10	.47
8	20	.51
8	40	.54
8	100	.55
8	150	.52
8	200	.38

Fig. 5: Influence of code parameters s,K on $R(\infty)$.

5.3. Influence of Channel Parameters δ, u_G.

We have seen in sections 5.1 and 5.2 that R mainly is a function of the symbol error probability u_G and the overall probability $\delta = P(z = B)$ of being in the bad state. For $\delta \neq 0$, $R(u_G, \delta)$ can be approximated from knowing $R(u_G, \delta = 0)$. For $\delta \neq 0$, of course, at least as much symbols in state z = G are required as for $\delta = 0$ because erased symbols in state z = B do not contain any information. Therefore,

$$n_\delta \cdot (1-\delta) \geq n_{\delta=0} \qquad (34)$$

We expect the inequality in eq. (34) close to equality because an erasure and a transmitted parity symbol may be regarded as canceling each other by eq. (13). The simulation results also showed results close to equality in (34). For example, with $s = 6$, $K = 20$, $u_G = 0.03$, the simulations for $\delta = 0$ yield $\overline{n} = 23.4$ (or $R = K/\overline{n} = 0.855$). For $\delta = 40\%$, equality in (34) would result in $\overline{n} = 23.4/(1-0.4) = 39.0$ (or $R = K/\overline{n} = 0.855*(1-0.4) = 0.51$) close to the simulation results in fig. 3. So,

$$R_\delta \approx R_{\delta=0} \cdot (1-\delta) \tag{35}$$

can be regarded as a good estimate of the coderate for $\delta \neq 0$.

A correspondingly simple estimate for R as function of u_G only could be found and shown to be valid for small values of u_G with correspondingly small $\overline{n} \ll N$. For error correction only ($\delta = 0$), the expected number $\overline{r}$ of parity symbols is about twice the expected number of errors, $\overline{r} = 2u_G \cdot \overline{n}$. With 2 additional parity symbols for reliable error detection if $\overline{n}$ is small, this results in $\overline{r} = \overline{n} - K \approx 2 + 2u_G\overline{n}$ or

$$R_{\delta=0} = K/\overline{n} \approx (1 - 2u_G)/(1 + 2/K) \text{ for } u_G \ll 1,\ \overline{n} \ll N \tag{36}$$

For example, with $s = 6$, $K = 20$, $u_G = 0.03$, (36) yields $R_{\delta=0} = 0.855$ as in the simulations. For $\overline{n} > N$, approximation (36) is not valid anymore. For example, with $s = 6$, $K = 20$, $p_G = 0.1$ corresponding to $u_G = 0.47$, (36) would result in $R \approx 0.06$, but the simulations revealed $R \approx 0.13$.

5.4. Rayleigh-Fading.

For Rayleigh-fading the threshold a_T for the signal level a was optimized such that capacity Cp, eq. (26), is maximum. When other optimization criteria were used (as C_{us}, $C_{\delta s}$, P_e, R), the results revealed not to be very sensitive to a_T, indicated also by a very flat optimum in C, P_e, and R. For various values of $\gamma_o = E_s/N_o$ and given s,K, the corresponding values p_G, p_B, u_G, u_B, δ according to eqs. (5) to (12) were calculated. By computer simulation, the

s	K	E_s/N_o	R	E_b/N_o	
4	8	0 dB	.15	8.2 dB	
		2 dB	.25	8.1 dB	
		4 dB	.37	8.3 dB	
6	20	0 dB	.19	7.2 dB	
		2 dB	.35	6.5 dB	≈opt.
		4 dB	.47	7.2 dB	
8	20	0 dB	.14	8.5 dB	
		2 dB	.30	7.2 dB	
		4 dB	.45	7.5 dB	
8	40	0 dB	.13	8.6 dB	
		2 dB	.31	7.1 dB	
		3 dB	.39	7.0 dB	
		4 dB	.47	7.3 dB	

Fig. 6: E_b/N_o as function of E_s/N_o and code parameters s,K.

resulting values R(L) were found (for independent state transitions, i.e., $g = \delta$, $b = 1 - \delta$). For $R \approx R(\infty)$, the resulting $E_b/N_o = E_s/(N_o \cdot R)$ are shown in fig. 6.

As seen from fig. 6, the optimum $E_s/N_o \approx 2$dB results in $E_b/N_o \approx 6.5$ to 7 dB for small P_e. With fixed rate FEC (without feedback), only about $E_b/N_o \approx 10$ to 11 dB can be achieved. Compared to uncoded CBPSK with AWGN $E_b/N_o = 9.6$ dB and with Rayleigh-fading $E_b/N_o \approx 44$ dB are necessary for $P_e \leq 10^{-5}$. So with feedback and Adaptive FEC with RS-codes, a Rayleigh-fading channel with coding needs less signal energy than an AWGN channel without coding.

5.5. Comparison to Capacity and Cutoff Rate.

The coderates R achieved by Adoptive FEC were compared to capacity $C = Cp$, C_{us}, $C_{\delta s}$ and cutoff rate $R_o = R_{ps}$, R_{us}, $R_{\delta s}$, and R_{max}, e.g., (26) through (33). As explained in chapter 4, $C_{\delta s}$, eq. (29), is the closest capacity limit for any error correcting scheme which corrects s-bit symbols (treating each wrong symbol the same way) as errors and erasures. $R_{\delta s}$, eq. (32), is the corresponding cutoff rate. For a GE-channel with $u_G = 0.03$, $\delta = 40\%$ and code parameters $s = 6$, $r = 20$, a coderate $R \approx 0.51$ was achieved by Adaptive FED. The corresponding values $C_{\delta s}$, $R_{\delta s}$ are $C_{\delta s} = 0.56$ (relatively close to R) but $R_{\delta s} = 0.19$ (much too pessimistic). Similar results were seen for Rayleigh-fading. For $\gamma_o = 2dB \approx (E_s/N_o)opt$, $a_T = 0.65$ (such that $Cp = max$) and a symbol-length of $s = 6$ bits the values of capacity C (and cutoff rate R_o together with the corresponding $E_b/N_o = \gamma_o$) C resp R_o are shown in fig. 7.

C =	⇔ E_b/N_o	R_o =	⇔ E_b/N_o
$C_{pl} = C_{ps} = C_p$		R_{pl} = .369	6.33 dB
C_p = .568	4.45 dB	R_{ps} = .295	7.31 dB
C_{us} = .459	5.38 dB	R_{us} = .156	10.07 dB
$C_{\delta s}$ = .423	5.74 dB	$R_{\delta s}$ = .135	10.71 dB

Fig. 7: Capacities and R_o's for Rayleigh-fading, $E_s/N_o = 2$ dB, $a_T = 0.65$, symbols-length $s = 6$ bits.

The coderate R achieved by Adaptive FEC with RS-codes $s = 6$, $K = 20$ is $R - 0.353$ resulting in $E_b/N_o = \gamma_o/R = 6.52$ dB. This is only 0.8 dB worse than the corresponding capacity limit $C_{\delta s}$ but more than 4 dB better than the R_o estimate from $R_{\delta s}$. (The maximum $R_{0,max} = .528$ corresponds to $E_b/N_o = 4.77$ dB for $E_s/N_o = 2$ dB.)

6. Conclusions and Open Questions.

For a two-state GE-channel-model with or without memory, it was shown that the proposed Adaptive FEC scheme has a coderate R which is almost independent from the length distribution of bursts. The coderate depends mainly on the symbol error probability u_G in the "good" state and the overall probability δ of being in the "bad" state. For $\delta \neq 0$, about as many symbols in the "good" state are necessary for reliable error and erasure correction and detection as for

$\delta = 0$, resulting in $R_\delta \approx R_{\delta=0} \cdot (1-\delta)$. For small values of $\overline{n}$, a simple approximation of R could be given. Code parameters s and K were seen not to influence the performance much, due to the adaptive nature of the correcting scheme. For error bursts caused by miltipath effects and modeled as Rayleigh-fading with a binary channel state quantization, the required signal-to-noise-ratio E_b/N_o is only 6.5 dB compared to 10 to 11 dB without feedback and 44 dB without coding. It could be shown that this is only about 0.8 dB worse than the corresponding channel capacity $C_{\delta s}$, but more than 4 dB better than the appropriate cutoff rate $R_o = R_{\delta s}$. Whereas capacity is the same for channels with or without feedback, this is not the case for R_o. Here we only were able to calculate R_o for channels without feedback. To calculate cutoff rate for channels without feedback is an issue for further research.

Of course, the distribution of bursts and the average burstlength i_B influences the error probability $P_e = P(n>L)$ if n is limited to a maximum of L transmitted symbols per codeword. By computer simulations, P_e could be shown to decrease exponentially with L, described by a constant error exponent E (independent from L), which decreases with increasing burstlength i_B. Because E is of great practical importance for the required buffer length of the receiver and the decoding delay, more analytical work should be done in order to get realistic estimates of E for a variety of parameters without tedious computer simulations. Some first information theoretic approaches by P. Narayan [10], using random codebook arguments, show results close to the values seen by simulations, encouraging further work in this direction. All results in the paper were calculated under the often unrealistic assumption of a zero-delay error-free feedback signal. The influence of feedback delay and feedback errors also may be an important issue for future work.