ON THE PERFORMANCE OF A HYBRID CODING SCHEME

Johannes Huber, Murad Oenol
Institute of Communication Systems
Federal Armed Forces University
D-8014 Neubiberg, Federal Republic of Germany

ABSTRACT

The performance of a hybrid coding scheme for MDS codes (e.g. Reed-Solomon codes) with parity retransmission, which was proposed by B. Dorsch for time-variant channels such as mobile radio links, is analysed. For that the binary digital channel model by McCullough is used. Approximations of the throughput efficiency and the error rate are given. The results show that at high bit error rates this scheme is obviously better than codeword repetition. Additionally, the choice of good code parameters turns out to be less sensitive to the behaviour of the channel.

1. INTRODUCTION

In many digital transmission systems errors occur in bursts, for example in mobile radio links. For those channels FEC schemes are only efficient if very long codewords or interleaving are used. For the transmission of short messages without continuous data stream, such as links to or from data terminals, only ARQ or hybrid systems are practicable. Especially hybrid coding schemes in which additional check symbols are retransmitted after error detection instead of repetition of the codeword are well suited for time-variant channels [1],[2],[3],[4],[5].

This report deals with such an error control scheme, which was proposed by Dorsch for cyclic maximum distance separable (MDS) codes, for which an error and erasure decoding algorithm exists, e.g. Reed-Solomon codes. A systematic codeword with length n and k information symbols is divided into s parts A,B,C.... each with $n/s>k$ symbols. At first part A containing the k information symbols and m_A check symbols is transmitted. If the missing symbols are regarded as erasures, part A can be decoded with error correction and (or) detection according to its minimal weight m_A+1, because parts of the codeword also form MDS codes. In case of error detection block B with $m_B = n/s$ check symbols is retransmitted. As the code is cyclic, this part can be decoded in the same way as the first part. When decoded correctly, the reconstructed symbols contain the information. In addition it is possible to combine both parts (A B) and to correct and detect errors according to the minimal weight m_A+m_B+1. If the codeword is divided into several parts, there are many different ways of combining parts and a great deal of possible sequences for decoding trials. From that proposed in [1] Dorsch's coding scheme differs in the possibility of decoding each retransmitted part independent of the others. Compared to the schemes according to [2],[3] all check symbols of combined codeword parts can be used for error correction.

In Sections 2 and 3 approximations for the throughput efficiency and the error rate of this hybrid coding scheme applied with Reed-Solomon codes over GF ($q=2^h$) are derived for a model of binary channels with memory. In Section 4 code parameters and results are discussed.

2. CHANNEL MODEL AND ERROR DETECTION

Scheme [4] was proposed for time-variant channels, especially mobile radio links. Therefore a two state digital model for burst channels by McCullough [6] is used. (Some other digital channel models can be transformed into McCullough's model [7].) In this model the state can only be changed after an error. This restriction allows a simple determination of the model parameters from the error gap distribution and autocorrelation function [7]. For simplicity it is assumed that the choice of the state after an error is independent of the previous state (renewal model [8]). So the model parameters are:

p_1: bit error rate state 1

p_2: bit error rate state 2 (burst state)

w_1: Pr(state 1 after an error)

A more evident description is given by p_2 and:

p_m: average bit error rate

L_B: average error burst length

The original parameters result to [9]:

$$p_1 = p_m/(1+(p_2-p_m)\cdot(L_B-1))$$
$$w_1 = 1/(1+p_2(L_B-1)) \qquad (1)$$

Although the weight distribution A(i) of MDS codes is known [10], the probability p_u of undetected errors cannot be evaluated exactly, because the channel for code symbols is asymmetrical. To get an approximation a model referring to the code symbols is derived, which is also a McCullough model [9]. Its symbol error probabilities are:

$$p_{ci} = 1-(1-p_i)^h \qquad (2)$$

The transition probabilities z_{ij} from state i to state j after an error are:

$$z_{ij} = \sum_{k=1}^{h} \frac{(1-p_i)^{k-1} p_i}{p_{ci}} \sum_{m=1}^{2} w_m u_{mj}^{(h-k)} \qquad (3)$$

$u_{ij}^{(l)}$: Element of the l^{th} power of the transition matrix of the binary channel model

In [6] a recurrence formula for the probabilities

P(i,n) = Pr (a block with n code symbols contains i errors)

is given. The probability p_u for undetected errors is approximately:

$$p_u \approx \sum_{i=1}^{n} \frac{A(i)}{\binom{n}{i}(q-1)^i} P(i,n) \qquad (4)$$

In Figure 1 the quotient f of this approximation (eq. 4) and the exact value for a short Reed-Solomon code is plotted. Figure 1 and further calculations show that the approximation (eq.4) gets more precise for growing parameters p_m, L_B, n, k, and that this exactness is sufficient for estimations of the error rate of Dorsch's hybrid coding scheme.

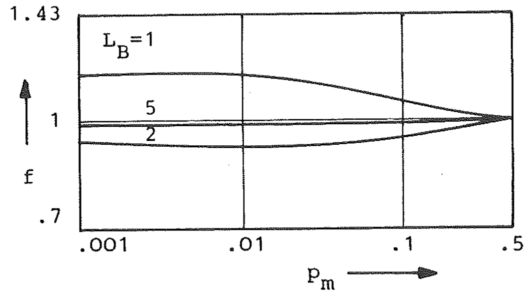

Fig.1. Quotient of the approximation eq.4 and the exact value for the probability of undetected errors (code parameters $n = 7,\ h = 3,\ k = 4$).

With correction up to t errors the probabilities p_u and p_d of undetected and detected errors are analogous to eq.4:

$$p_u \approx \sum_{i=0}^{n} u(i,t,n)P(i,n)$$

$$p_d \approx \sum_{i=0}^{n} d(i,t,n)P(i,n)$$

$$u(i,t,n) = \begin{cases} 0 & 0 \le i \le t \\ \dfrac{\sum_{e=0}^{t} B(e,i)}{\binom{n}{i}(q-1)^i} & t < i \le n \end{cases} \qquad (5)$$

$$d(i,t,n) = \begin{cases} 0 & 0 \le i \le t \\ 1-u(i,t,n) & t < i \le n \end{cases}$$

The weight distributions $B(e,i)$ of the code cosets, whose leaders have the weight e, can be evaluated from the weight distribution $A(i)$ [11]. (In eq.5 a test whether the corrected word is a codeword is assumed when error correction is applied up to $t=INT((d_{min}-1)/2)$.)

3. PERFORMANCE ANALYSIS

In this Section the throughput efficiency η and the word error rate p_e are determined for two types of decoding sequences. As the results should give a comparison of different code parameters, a selective repeat system with infinite buffer memory is assumed. (If there is no continuous data stream, a more exact analysis is not possible.) Moreover, a sufficient delay time between the transmissions of codeword parts is assumed so that the error patterns occur independently of one another. Equations are given for dividing the codewords in two parts; formulas for more parts are analogous.

3.1 Sequence I

The following decoding sequence is denoted by Ia.

Ia: A_1, (A_1B_2), B_2, (A_3B_2), A_3 ...

For this sequence the probabilities of incorrect decoding and error detection depend on the error patterns of previous transmissions. Applying eq.5, this dependence consists of the number of errors in the previous block. With the definition:

X_{il} = Pr (i errors at the l^{th} transmission and error detection for all decoding trials)

the joint probability $Pr(d_1 \cap d_2 \cap \ldots \cap d_l)$ of error detection after l transmitted blocks is:

$$Pr(d_1 \cap d_2 \ldots \cap d_l) = \sum_{i=0}^{n/2} X_{il} \qquad (6)$$

With

Y_{il} = Pr (i errors at the l^{th} transmission, incorrect decoding for this block and error detection for all previous decoding trials)

the joint probability for incorrect decoding is:

$$Pr(d_1 \cap d_2 \ldots \cap u_l) = \sum_{i=0}^{n/2} Y_{il} \qquad (7)$$

X_{il} and Y_{il} can be obtained by the following recurrence formula:

$$\begin{aligned} X_{i1} &= P(i,n/2)\cdot d(i,t_1,n/2) \\ Y_{i1} &= P(i,n/2)\cdot u(i,t_1,n/2) \\ Z_{il} &= P(i,n/2)\cdot \sum_{j=0}^{n/2} X_{jl-1} d(i+j,t_2,n) \\ X_{il} &= Z_{il}\cdot d(i,t_1,n/2) \\ Y_{il} &= P(i,n/2)\cdot \sum_{j=0}^{n/2} X_{jl-1} u(i+j,t_2,n) \\ &\quad + Z_{il} u(i,t_1,n/2) \end{aligned} \qquad (8)$$

t_i: error correction capability of the combination of i transmitted blocks

The recurrence can be stopped, if all X_{il_e} are small enough.

For a decoding sequence I_b (see [2])

Ib: A_1, B_2, (A_1B_2), A_3, (A_3B_2), ...

an analogous formula exists.

The average number $E[l]$ of transmitted blocks is

$$E[l] = Pr(\bar{d}_1) + 2\,Pr(d_1 \cap \bar{d}_2) + \ldots \qquad (9)$$

With

$$Pr(d_1 \cap \ldots \cap \bar{d}_l) = Pr(d_1 \cap \ldots \cap d_{l-1}) - Pr(d_1 \cap \ldots \cap d_l) \qquad (10)$$

eq. 9 leads to

$$E[l] \approx 1 + \sum_{l=1}^{l_e} \sum_{i=0}^{n/2} X_{il}. \qquad (11)$$

The throughput efficiency is therefore:

$$\eta \approx k/((n/2)E[l]) \qquad (12)$$

The error rate is given by:

$$p_e \approx \sum_{l=1}^{l_e} \sum_{i=0}^{n/2} Y_{il} \qquad (13)$$

3.2 Sequence II

Sequence II is characterized by a restarting decoding process at the retransmission of block A which contains the information symbols:

IIa: A_1, (A_1B_2), B_2, A_3, ...

IIb: A_1, B_2, (A_1B_2), A_3, B_4, ...

For these sequences only two steps of recurrence eq.8 are necessary, because for the joint probabilities of error detection is valid:

$$Pr(d_1 \cap d_2 .. \cap d_l) = (Pr(d_1 \cap d_2))^{INT(l/2)} \cdot (Pr(d_1))^{1-2INT(l/2)} \qquad (14)$$

With that eq.12 and eq.13 yield:

$$\eta \approx \frac{k}{n/2} \cdot \left(1+\sum_{i=0}^{n/2} X_{i1}\right)^{-1} \cdot \left(1-\sum_{i=0}^{n/2} X_{i2}\right) \qquad (15)$$

$$p_e \approx \left(\sum_{l=1}^{2}\sum_{i=0}^{n/2} Y_{il}\right)/\left(1-\sum_{i=0}^{n/2} X_{i2}\right) \qquad (16)$$

3.3 Comparison of Decoding Sequences

In Figure 2 the dependence of the throughput efficiency and the error rate on the bit error rate (BSC without memory: L_B=1) of a Reed-Solomon code of length 126 with 56 information symbols is shown for the decoding sequences Ia, Ib, IIa, IIb, and simple repetition of part A (III).

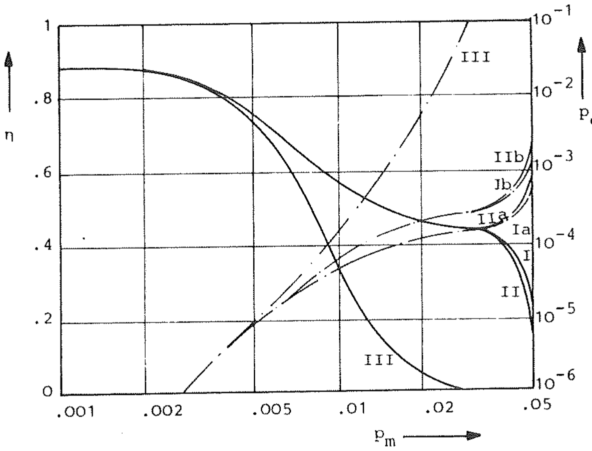

Fig.2. Throughput efficiency (——) and error rate (-·-·-) for the decoding sequences I, II, III (repetition) (L_B=1; code parameters s=2, n=126, k=56, t_1=3, t_2=35).

Especially at high bit error rates parity retransmission achieves a much greater efficiency, because further retransmissions are avoided by error corrections for block (A_1B_2). Only here sequence I in which one decoding trial more is carried out with block B is slightly superior to sequence II. For codewords divided into three parts this difference is neglectable. When retransmissions of part B are necessary, the sequences Ib, IIb show approximately twice the error rate of Ia, IIa, because almost all undetected errors occur at the decoding of single blocks. As long Reed-Solomon codes of rate ≈ 1/2 are packed extremely thinly, the probability of undetected errors for combinations of parts is neglectable, even when the full error correction capability is used. Therefore sequences in which long combinations are decoded first reach maximal reliability.

4. Code Parameters

In this Section advantageous code parameters, i.e. codeword length n, number k of information symbols and error correction capability t_1 for part A are given for several channels. Reed-Solomon codes of length n=30 (h=5), n=63, n=62 (h=6) and n=126 (h=7), divided into two (s=2) and three (s=3) parts, are dealt with here. For all cases results are given for the decoding sequences Ia and IIa. For dividing codewords into three parts these decoding sequences are A_1, (A_1B_2), B_2, $(A_1B_2C_3)$...; the possible combination (A_1C_3) is not included. For combinations of blocks always the maximal error correction capabilities t_2 and t_3 are assumed (see Section 3.3). As examples channels with low (p_m=0.001), medium (p_m=0.01) and high bit error rate (p_m= 0.05) without memory (L_B=1) and with long, dense error bursts (L_B=100, p_2=0.5) are chosen. For these channels Figures 3a-f show the throughput efficiency and the word error rate for several codes, whose parameters are given in Table I. The scheme with parity retransmission is marked with the symbol +. If the results for sequence Ia differ from that for sequence IIa additionally the sign □ is used for sequence Ia. For comparison simple repetition of part A of the same code is marked with the sign x in these Figures. Those codes are shown in Figure 3, with which for one of the six channels an optimum throughput efficiency can be achieved. Additionally, limit curves are plotted for Reed-Solomon codes of length n=30 up to n=126. For repetition of part A these limits are plotted with dashed lines.

For the memoryless channel with low error rate (Figure 3a) the greatest efficiency is obtained by long codes with a small error correction capability t_1 (code #10). The reliability is much enlarged by an additional check symbol (#11). Here parity retransmission leads to no further improvement; the limit curves are almost identical. For the channel with error bursts (Figure 3b) error detection t_1=0 for part A shows the highest efficiency (#8,#9), because in this case the probability of error-free blocks is much greater than for memoryless channels with the same error rate. The word error rate for error correction (#11) increases with growing average burst length. For the medium error rate (Figure 3c) a long code (#12) with high error correction capability is the optimum. For the codes #8 - #11 frequent parity retransmissions are necessary and therefore the efficiency is about 0.5, whereas for repetition of part A these codes are not useful. For the channel with error bursts (Figure 3d) again error detection t_1=0 is the optimum (#6,#7) but with shorter codes than for the low error rate (#8,#9).

Figure 3e shows that for the memoryless channel with high error rate short codes (#1, #2,#4) lead to the greatest efficiency. Only here there is a remarkable difference between the limit curves. Additionally, it turns out that for this channel a division of the codeword into three parts is superior to a division

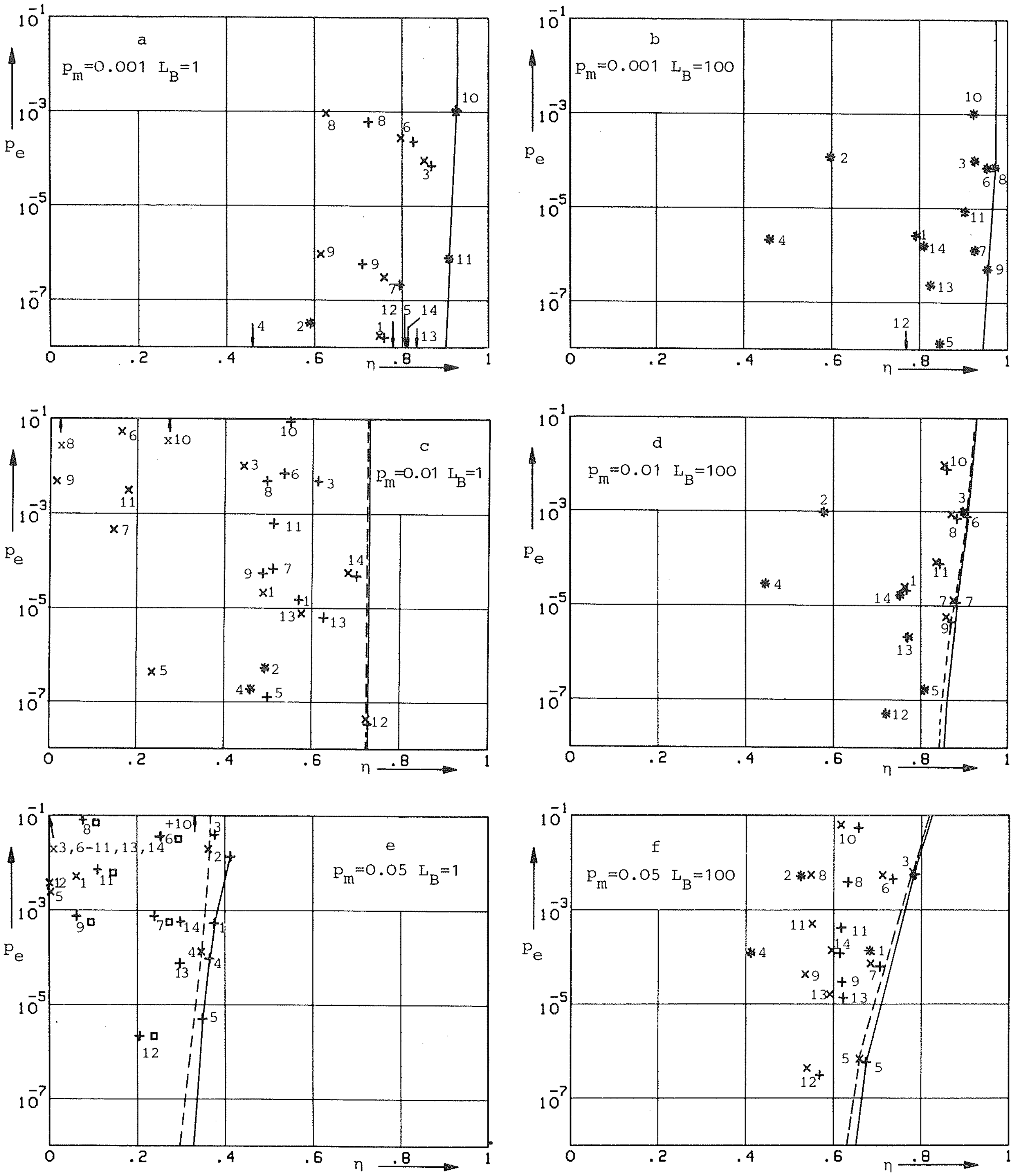

Fig.3. Word error rate p_e and throughput efficiency η of the codes listed in Table I for channels with different bit error rates p_m and average burst length L_B. (□: parity retransmission sequence Ia; +: parity retransmission sequence Ia and IIa; x: repetition of part A; ——: limit parity retransmission; ---: limit repetition)

Table I

Code Parameters

Code #	n	s	k	t_1	t_2	t_3
1	30	3	8	0	6	11
2	30	3	6	2	7	12
3	30	2	14	0	8	--
4	30	2	7	4	11	--
5	63	3	18	0	12	22
6	62	2	30	0	16	--
7	62	2	29	0	16	--
8	126	2	62	0	32	--
9	126	2	61	0	32	--
10	126	2	59	2	33	--
11	126	2	58	2	34	--
12	126	2	49	7	38	--
13	126	3	35	3	24	45
14	126	3	34	4	25	46

into two parts. With retransmission of part C an error correction is possible in most cases and therefore an efficiency of about 0.3 is achieved (#1, #2, #13, #14) whereas repetitions are necessary for two parts. Here the decoding sequence Ia offers an improvement for codes divided into two parts (sign □).

Figures 3a-f show that only for high error rates the scheme with parity retransmission has noticeably better limit curves than repetition. But the important advantage of the parity retransmission scheme is that the choice of good code parameters is much less sensitive to the behaviour of the channel. So the codes #1 and #5 are for almost all channels near to the limit curves. Only for p_m=0.01, L_B=1 (Figure 3c) error corrections applied for part A would improve the efficiency. Codes #13 and #14, which are not the best ones for any of the six channels, avoid this disadvantage and are quite near to all limit curves. In Figure 4 the throughput efficiency for code #14 is plotted for different burst length.

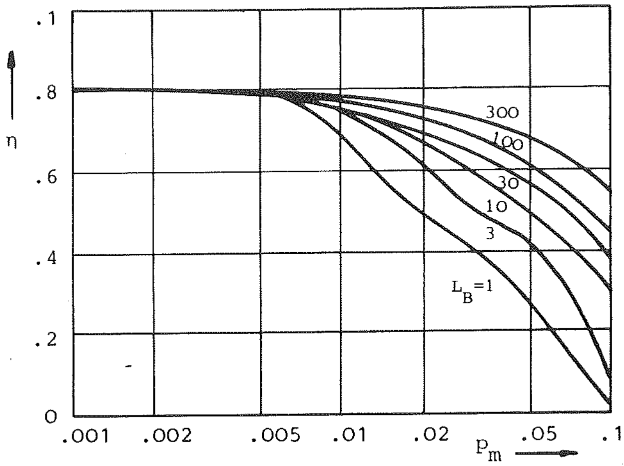

Fig.4. Throughput efficiency for code #14 (Table I) for channels with dense error bursts of different average length.

On the contrary, for repetition of part A good codes for high bit error rates (for example #4) are suboptimal at low errors and vice versa (#6 - #11).

5. CONCLUSION

The performance analysis for the hybrid coding scheme with parity retransmission [3], [4] which was derived in Sections 2 and 3 allows the determination of optimum code parameters for different channels without and with memory. It turns out that the choice of good code parameters is not sensitive to the channel behaviour and that there are several codes such as #5, #13, #14 of Table I for which a throughput efficiency near to the optimum value can be achieved for channels with different error rates and without or with error bursts.

REFERENCES

[1] D.M. Mandelbaum, "An Adaptive-Feedback Coding Scheme Using Incremental Redundancy," *IEEE Trans. Info. Theory*, vol. IT-20, pp. 388-389, May 1974.

[2] S. Lin and P.S. Yu, "A Hybrid ARQ Scheme with Parity Retransmission for Error Control of Satellite Channels," *IEEE Trans. Commun.*, vol. COM-30, pp. 1701-1719, July 1982.

[3] Y.M. Wang and S. Lin, "A Modified Selective-Repeat Type II Hybrid ARQ System and Its Performance Analysis," *IEEE Trans. Commun.* vol. COM-31, pp. 593-607, May 1983.

[4] B. Dorsch, "Kombinierte FEC-ARQ Systeme für den Mobilfunk," Lecture at the Discussion "*Mobiler Datenfunk*" at DFVLR, Oberpfaffenhofen, Germany, Sept. 16-17, 1982, and private communication.

[5] B. Dorsch, "Successive Check Digits Rather than Information Repetition", in *Proc. Int. Comm. Conf.* (ICC), Boston, pp. 323-327, June 1983.

[6] R.H. McCullough, "The Binary Regenerative Channel," *Bell Syst. Tech. J.*, vol. 47, pp. 1713-1735, Oct. 1968.

[7] J. Huber, "Zur Anwendung von Kanalmodellen für Codierverfahren," in *NTG-Fachberichte 84*, pp. 51-59, March 1983.

[8] L.N. Kanal and A.R.K. Sastry, "Models for Channels with Memory and Their Applications to Error Control," *Proc. IEEE*, vol. 66, pp. 724-744, July 1978.

[9] J. Huber, "Codierung für gedächtnisbehaftete Kanäle," Dissertation, Hochschule der Bundeswehr München, July 1982.

[10] F.J. MacWilliams and N.J.A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam: North-Holland Publ., 1978.

[11] Z.M. Hunctoon and A.M. Michelson, "On the Computation of the Probability of Post-Decoding Error Events for Block Codes," *IEEE Trans. Inf. Theory*, vol. IT-23, pp. 399-403, May 1977.