# CUT-OFF RATE CALCULATIONS FOR THE OUTER CHANNEL IN A CONCATENATED CODING SYSTEM

Mark A. Herro[1], Daniel J. Costello, Jr.[2], Laizhao Hu[3]
Department of Electrical Engineering
University of Notre Dame
Notre Dame, IN 46556

## ABSTRACT

Concatenated codes have long been used as a practical means of achieving long block or constraint lengths for combating errors on very noisy channels. The inner and outer encoders are normally separated by an interleaver, so that decoded error bursts coming from the inner decoder are randomized before entering the outer decoder. In this paper we examine the effectiveness of this interleaver by calculating the cut-off rate of the "outer channel" seen by the outer decoder with and without interleaving. The results show that interleaving can never hurt the performance of a concatenated code, and that when the inner code rate is near the cut-off rate of the "inner channel", interleaving can significantly improve code performance.

## I. INTRODUCTION

Forney [1] first introduced concatenated codes as a practical means of implementing codes with long block or constraint lengths. Further use has shown these codes to be extremely powerful as means of combating errors on very noisy channels. In this paper, we examine the performance of concatenated coding systems as measured by their effective channel cut-off rate.

Concatenated coding systems are usually implemented by employing two levels of coding, as illustrated in Figure 1 below.

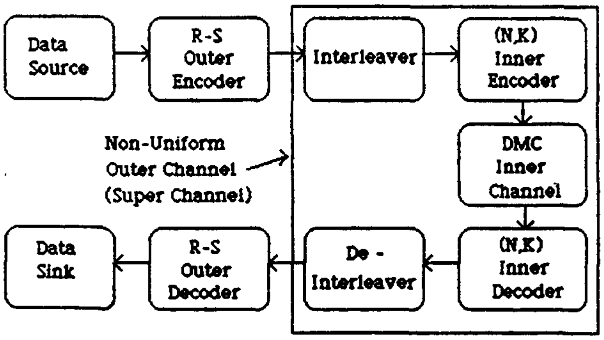


Figure 1. Concatenated Coding System

Binary data from the information source is serially partitioned into K-bit blocks that are subsequently used as input signals to a $2^K$-ary block encoder known as the outer encoder. Usually $Q=2^K$-ary Reed-Solomon (RS) codes are used for this purpose. The output of the RS encoder (Q-ary symbols) is converted back into bits and serially encoded by a second encoder (the inner encoder), which may be either block or convolutional, and the resultant sequence of channel symbols is sent over the physical channel. Decoding is accomplished in the reverse order.

For purposes of illustration, we consider the inner channel to be a Binary Symmetric Channel (BSC) derived from forcing hard decisions on an additive white gaussian noise (AWGN) channel. This channel is representative of the deep space channel where concatenated codes have met with a great deal of success. The outer channel (the channel presented to the RS code) is no longer memoryless, but has been transformed into a non-uniform (time-varying) channel by the inner decoder. To calculate the performance of the overall concatenated coding scheme requires investigating the channel produced by the inner encoder-BSC-inner decoder combination which Forney [1] has called the "superchannel".

## II. CHANNEL MODELS

McEliece and Stark [2] have suggested some models for channels with "block interference". In a block interference channel, the noise statistics are assumed to be constant for the time required to send K bits of data, but vary independently from one block of K bits to the next. In particular, McEliece and Stark have evaluated the channel capacity and cut-off rate for a two state block interference channel. When the channel is in the quiet state, no errors are made in transmission. The noisy state is represented by a BSC with a crossover probability of s. (A totally noisy channel has $s=1/2$). However, even when the channel is in the noisy state, it is possible to receive the K bits of data correctly with probability $(1-s)^K$. In this model, the probability of being in a particular state, p(s), depends on the physical channel (fading, burst noise, frequency hopping, etc) and is fixed and independent of the block length K.

To evaluate the performance of a concatenated code, we propose a modified form of the above model that provides a better match to this coding application. We assume that a rate $R=K/N$ block code is used on the inner channel and that a decoding error occurs with probability $P_{BE}(K)$. That is, the channel is in the noisy state with this probability. When the channel is in the noisy state, each information bit is decoded incorrectly with probability s. Various methods can be used to estimate the decoded information bit error rate

[1]This work was supported in part by NASA Grant NCA2-1R562-401
[2]On leave from the Illinois Institute of Technology, Chicago, IL 60616. This work was supported in part by NASA Grant NAG 5-234
[3]On leave from the Electronic Technology Institute of Southwest China, People's Republic of China

from the block error probability (see, for example, Clark and Cain [3]). For non-systematic codes, $s=1/2$ gives a good estimate for the decoded information bit error rate, given that a block error has occurred. For systematic codes, an improved estimate is $s=d/N$, where d is the minimum distance of the code. The channel is in the quiet state ($s=0$) with probability $1-P_{BE}(K)$, and no decoding errors are made. With probability $P_{BE}(K)$ the channel is in the noisy state and decoding errors are always made. The important difference between this model and that used by McEliece and Stark is that in this model the probability of being in the noisy state is dependent on the length of the "interference" blocks (information word length).

When convolutional codes are used on the inner channel, a more complex model of the superchannel is needed to compute performance measures. The reason is that error events in convolutional coding have different lengths. Forney [4] has derived some random coding results on the lengths of these error events. To illustrate this point, consider the error environment generated with an (N,1,K) (1 N output length K shift register convolutional code). Using hard decisions on the AWGN channel, Viterbi [5] shows that the decoded event error probability at any time during maximum likelihood decoding is bounded by

$$P(E) \leq T(X)\Big|_{X = 2\sqrt{p(1-p)}}, \tag{1}$$

where $T(X)$ is one form of the code generating function, and p is the crossover probability forced by using hard decisions. The decoded information bit error probability is bounded by

$$P_b(E) \leq \frac{\partial T(X,Y)}{\partial Y}\Bigg|_{X = 2\sqrt{p(1-p)},\ Y=1.} \tag{2}$$

For small values of p (high signal-to-noise ratios), the most likely error event (when the all-zero sequence is sent) is that the minimum weight path is decoded instead of the all-zero path. If the information sequence corresponding to the minimum weight path has weight b and length L, then the typical event error causes b bit errors and has length L. Therefore, we can use the above block code model with "block length"=L, and $s=b/L$. For larger values of p (lower signal-to-noise ratios), longer event errors become more likely and cannot be ignored. In this case a more general channel model for the outer channel is required to evaluate the cut-off rate.

## III. THE CUT-OFF RATE

The objective is to find the maximum achievable coding rate for the concatenated coding system. Define the cut-off rate of the inner channel as $R_{0i}$ and that of the outer channel as $R_{0o}$. Now if a rate R code is used on the inner channel, the maximum overall code rate is given by

$$R_{max} = R \cdot R_{0o}. \tag{3}$$

If the crossover probability of the inner BSC is p, then the cut-off rate, $R_{0i}$, is given by [6]:

$$R_{0i} = 1-\log_2[1+2\sqrt{p(1-p)}]. \tag{4}$$

For inner code rates $R<R_{0i}$, the random coding bound on the block error probability is given by

$$P_{BE}(K) \leq 2^{-N(R_{0i}-R)} = 2^{-K[(R_{0i}/R)-1]}, \tag{5}$$

where N, K, and R are the block length, information block length and code rate of the inner code, respectively. For non-systematic codes, the decoded information bit error rate for the block code is equal to 1/2, given that a block decoding error has occurrred. Therefore, the unconditional bit error probability of the decoded output is

$$P_b = (1/2)\cdot P_{BE}(K) \leq 2^{-1-K[(R_{0i}/R)-1]}. \tag{6}$$

We now proceed to evaluate the cut-off rate of the outer code. Two possibilities are considered. If full interleaving is used, then the outer channel can be treated as a memoryless channel and the cut-off rate can be calculated in the usual way. If interleaving is not used, then the channel is treated as a $2^K$-ary memoryless channel and the cut-off rate is evaluated by methods to be presented in the next section.

### A. Interleaving

If we interleave the bits entering the superchannel, the channel seen by the outer coding system becomes a memoryless BSC with a bit error probability given by eqn. (6) above. Note that we are still using coding, so that the block error probabilities given are unchanged. Under this assumption, we calculate the cut-off rate for the outer channel by

$$R_{0o} = 1-\log_2[1+2\sqrt{P_b(1-P_b)}]$$

$$\geq 1-\log_2\left[1+\sqrt{2^{-K[(R_{0i}/R)-1]}\cdot(2-2^{-K[(R_{0i}/R)-1]})}\right] \tag{7}$$

Some sample calculations of $R_{0o}$ are presented in Table 1 below.

| $R/R_{0i}$ \ K | 4 | 8 | 16 | 32 | 64 | 128 | 256 | 512 |
|---|---|---|---|---|---|---|---|---|
| 0.1 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 0.2 | .9921 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 0.3 | .9219 | .9968 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 0.4 | .7660 | .9685 | .9995 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 0.5 | .5692 | .8779 | .9921 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 0.6 | .3783 | .7115 | .9503 | .9987 | 1.000 | 1.000 | 1.000 | 1.000 |
| 0.7 | .2186 | .4932 | .8223 | .9825 | .9998 | 1.000 | 1.000 | 1.000 |
| 0.8 | .1000 | .2676 | .5692 | .8779 | .9921 | 1.000 | 1.000 | 1.000 |
| 0.9 | .0260 | .0832 | .2295 | .5109 | .8364 | .9853 | .9999 | 1.000 |

Table 1. $R_{0o}$ with Interleaving

The outer channel cut-off rate is calculated as a function of the ratio of the inner code rate to the inner channel cut-off rate, and the information word length K. The probability of being in the noisy state (block decoding error) is obtained from the random coding bound (5) for the particular values of K and $R/R_{0i}$ in the table. Table 2 gives the maximum attainable coding rate of the overall system, normalized by the inner channel cut-off rate, $R_{0i}$.

| $R/R_{0i}$ \ K | 4 | 8 | 16 | 32 | 64 | 128 | 256 | 512 |
|---|---|---|---|---|---|---|---|---|
| 0.1 | .1000 | .1000 | .1000 | .1000 | .1000 | .1000 | .1000 | .1000 |
| 0.2 | .1984 | .2000 | .2000 | .2000 | .2000 | .2000 | .2000 | .2000 |
| 0.3 | .2766 | .2991 | .3000 | .3000 | .3000 | .3000 | .3000 | .3000 |
| 0.4 | <u>.3064</u> | .3874 | .3998 | .4000 | .4000 | .4000 | .4000 | .4000 |
| 0.5 | .2846 | <u>.4390</u> | .4960 | .5000 | .5000 | .5000 | .5000 | .5000 |
| 0.6 | .2270 | .4269 | .5702 | .5992 | .6000 | .6000 | .6000 | .6000 |
| 0.7 | .1530 | .3453 | <u>.5756</u> | .6878 | .6999 | .7000 | .7000 | .7000 |
| 0.8 | .0800 | .2141 | .4554 | <u>.7023</u> | <u>.7936</u> | .8000 | .8000 | .8000 |
| 0.9 | .0234 | .0749 | .2066 | .4598 | .7528 | <u>.8868</u> | <u>.8999</u> | <u>.9000</u> |

Table 2. $R_{max}/R_{0i}$ with Interleaving

The maximum values for each value of K are underlined. For a given value of K, note that the maximum value of $R_{max}/R_{0i}$ does not occur at the lowest value of R (which presents the best channel to the outer coder). Instead, there is a local maximum for each K that drifts toward the higher rates as K increases. This is reasonable since the code performance should improve for increasing block lengths as long as the coding rate R is below $R_{0i}$. In addition, $R_{max}$ increases as R increases, given that K is sufficiently large. For the shorter block lengths (e.g., K=4 or K=8) a lower rate inner code is required to "clean up" the inner channel, and as a consequence the overall coding rate suffers.

The limiting values are of interest. From Table 1, we see that $R_{0o}$ approaches 1 as K increases, and from Table 2 we note that $R_{max}/R_{0i}$ tends toward $R/R_{0i}$ with increasing K. When K is small, lower inner code rates R yield a higher overall coding rate.

## B. Inner Channel as a $2^K$-ary Memoryless Channel

If no interleaving is used, it is difficult to justify using a random coding bound of the $R_0$ type since the channel is not memoryless. However, if we think of the Kth extension of the channel as producing a memoryless channel whose input and output symbols correspond to blocks of length K, then we can calculate the channel cutoff rate for this memoryless $2^K$-ary channel. Following Massey [6] we have

$$R_{0o} = \max_{Q(\underline{x})} \{-\log_M [\sum_{\underline{y}} [\sum_{\underline{x}} \sqrt{f(\underline{y}|\underline{x})}Q(\underline{x})]^2]\} \qquad (8)$$

where $\underline{x}$ and $\underline{y}$ are the channel input and output K-tuples, respectively, $f(\underline{y}|\underline{x})$ is the channel transition probability, $Q(\underline{x})$ is the channel input distribution, and $M=2^K$. The calculation in eq. (8) begins by first calculating

$$f(\underline{y}|\underline{x}) = \sum_s p(s) f(\underline{y}|\underline{x},s) \qquad (9)$$

That is, $f(\underline{y}|\underline{x})$ is calculated for an average channel. (This assumes that the receiver has no "side information" about the state of the channel, as defined by McEliece and Stark [2]).

As an example, consider a two state channel ($s=0$ <=> "quiet", and $s\neq 0$ <=> "noisy"). Let $Pr(s\neq 0) = P_{BE}(K)=\varepsilon$, and $Pr(s=0) = 1-P_{BE}(K) = 1-\varepsilon$. Given that no block error occurs ($s=0$),

$$\begin{aligned} f(\underline{y}|\underline{x},s=0) &= 1 \ , \ \underline{y} = \underline{x} \\ &= 0 \ , \ \underline{y} \neq \underline{x} \end{aligned} \qquad (10)$$

If there is a block decoding error, then we assume that the event $\{\underline{y}=\underline{x}\}$ never occurs. In this case, the entries for each element in the channel transition matrix for which $\underline{y}\neq\underline{x}$ must be normalized by the probability that $\underline{y}\neq\underline{x}$. Let $\alpha = 1-(1-s)^K$ be the probability that $\underline{y}\neq\underline{x}$ for a block of K bits. Then

$$\begin{aligned} f(\underline{y}|\underline{x},s\neq 0) &= \frac{s^i(1-s)^{K-i}}{\alpha} \ , \ \underline{y} \neq \underline{x} \\ &= 0 \ , \ \underline{y} = \underline{x} \end{aligned} \qquad (11)$$

where $i = d_H(\underline{x},\underline{y})$ = Hamming distance between $\underline{x}$ and $\underline{y}$.

The average channel transition probability is obtained as follows:

$$\begin{aligned} f(\underline{y}|\underline{x}) &= \varepsilon f(\underline{y}|\underline{x},s\neq 0) + (1-\varepsilon) f(\underline{y}|\underline{x},s=0) \\ &= (1-\varepsilon) \ , \ \underline{y} = \underline{x} \\ &= \frac{\varepsilon}{\alpha} s^i(1-s)^{K-i} \ , \ \underline{y} \neq \underline{x}. \end{aligned} \qquad (12)$$

If eq. (12) is put into matrix form, it is easy to see that the channel is symmetric, and therefore $Q(\underline{x})=1/M$ maximizes $R_{0o}$ in eq. (8). Substituting in eq. (8) we obtain

$$R_{0o} = -\frac{1}{K}\left\{\log_2 \sum_{\underline{y}} \frac{1}{M^2}\left[\sqrt{1-\varepsilon} + \sum_{i=1}^{K}\binom{K}{i}\sqrt{\frac{\varepsilon}{\alpha} s^i(1-s)^{K-i}}\right]^2\right\} \qquad (13)$$

Since the term inside of the square brackets is independent of $\underline{y}$, the summation over $\underline{y}$ reduces to a multiplication by M, and

$$R_{0o} = 1-\frac{2}{K}\log_2\left[\sqrt{1-\varepsilon}+\sum_{i=1}^{K}\binom{K}{i}\sqrt{\frac{\varepsilon}{\alpha} s^i(1-s)^{K-i}}\right] \qquad (14)$$

With $\varepsilon=P_{BE}(K) \leq 2^{-N(R_{0i}-R)}$, and choosing $s=1/2$ to represent the decoding of a non-systematic code, we have

$$R_{0o} \geq 1-\frac{2}{K}\log_2\left\{\sqrt{1-2^{-K[(R_{0i}/R)-1]}} + \frac{2^{-(K/2)[(R_{0i}/R)-2]}}{\sqrt{1-2^{-K}}}\right\} \qquad (15)$$

Some sample calculations of $R_{0o}$ are shown in Table 3. The corresponding values for $R_{max}/R_{0i}$ are found in Table 4. As in the case of interleaving, $R_{0o}$ approaches 1 for $R/R_{0i} < 1/2$, and in all cases increases with increasing block length for fixed rate. With K fixed, $R_{0o}$ decreases with increasing $R/R_{0i}$.

| $R/R_{0i}$ \ K | 4 | 8 | 16 | 32 | 64 | 128 | 256 | 512 |
|---|---|---|---|---|---|---|---|---|
| 0.1 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 0.2 | .9892 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 0.3 | .8981 | .9912 | .9999 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 0.4 | .7190 | .9197 | .9891 | .9996 | 1.000 | 1.000 | 1.000 | 1.000 |
| 0.5 | .5233 | .7507 | .8750 | .9375 | .9687 | .9844 | .9922 | .9961 |
| 0.6 | .3522 | .5479 | .6403 | .6645 | .6666 | .6667 | .6667 | .6667 |
| 0.7 | .2143 | .3648 | .4212 | .4284 | .4286 | .4286 | .4286 | .4286 |
| 0.8 | .1076 | .2136 | .2473 | .2500 | .2500 | .2500 | .2500 | .2500 |
| 0.9 | .0304 | .0916 | .1100 | .1111 | .1111 | .1111 | .1111 | .1111 |

Table 3. $R_{0o}$ without Interleaving

| $R/R_{0i}$ \ K | 4 | 8 | 16 | 32 | 64 | 128 | 256 | 512 |
|---|---|---|---|---|---|---|---|---|
| 0.1 | .1000 | .1000 | .1000 | .1000 | .1000 | .1000 | .1000 | .1000 |
| 0.2 | .1978 | .2000 | .2000 | .2000 | .2000 | .2000 | .2000 | .2000 |
| 0.3 | .2694 | .2974 | .3000 | .3000 | .3000 | .3000 | .3000 | .3000 |
| 0.4 | .2876 | .3679 | .3956 | .3999 | .4000 | .4000 | .4000 | .4000 |
| 0.5 | .2616 | .3754 | .4375 | .4687 | .4844 | .4922 | .4961 | .4980 |
| 0.6 | .2113 | .3287 | .3842 | .3987 | .4000 | .4000 | .4000 | .4000 |
| 0.7 | .1500 | .2554 | .2948 | .2999 | .3000 | .3000 | .3000 | .3000 |
| 0.8 | .0861 | .1709 | .1978 | .2000 | .2000 | .2000 | .2000 | .2000 |
| 0.9 | .0273 | .0824 | .0990 | .1000 | .1000 | .1000 | .1000 | .1000 |

Table 4. $R_{max}/R_{0i}$ without Interleaving

From Table 4 we see that the normalized value of $R_{max}$ also increases with K for a fixed value of $R/R_{0i}$. However, for a fixed K, the value of $R_{max}/R_{0i}$ does not monotonically decrease with $R/R_{0i}$, but increases to a maximum value at $R/R_{0i}=1/2$, and then decreases as the inner code rate increases. For a fixed K, there is some optimum value of $R/R_{0i}$ that maximizes the total overall coding rate achievable.

It is of interest to note the asymptotic behavior of $R_{0o}$ as K becomes large. If we take the limit as $K\to\infty$ of $R_{0o}$ in eq. (15), we see that

$$R_{0o} \to 1-\frac{2}{K}\log_2\left\{1+2^{-(K/2)[(R_{0i}/R)-2]}\right\} \qquad (16)$$

From Eq. (16) above, it is easy to see that

$$R_{0o} \to 1 \qquad \text{when } R < \frac{R_{0i}}{2}$$

and

$$R_{0o} \to \frac{R_{0i}}{R} - 1 \quad \text{when } \frac{R_{0i}}{2} < R < R_{0i}. \qquad (17)$$

Therefore,

$$R_{max} \to R \qquad \text{when } R < \frac{R_{0i}}{2}$$

and

$$R_{max} \to R_{0i} - R \quad \text{when } \frac{R_{0i}}{2} < R < R_{0i}. \qquad (18)$$

Figure 2 shows the normalized values of $R_{max}$ when interleaving is used. Figure 3 shows the normalized $R_{max}$ when interleaving is not used.

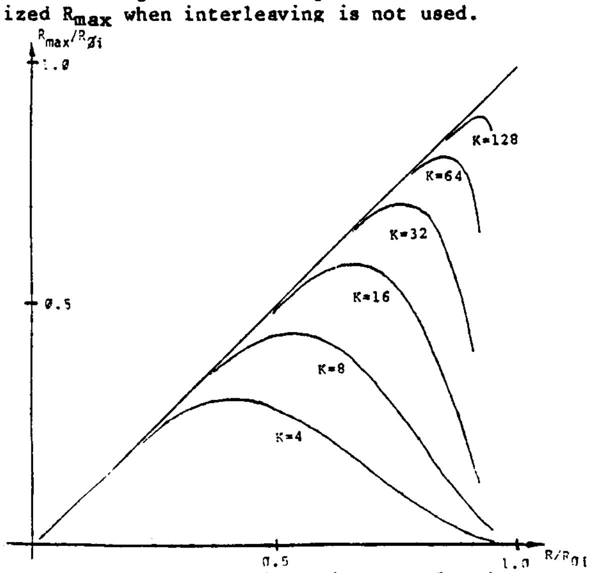


Figure 2. $R_{max}/R_{0i}$ with Interleaving

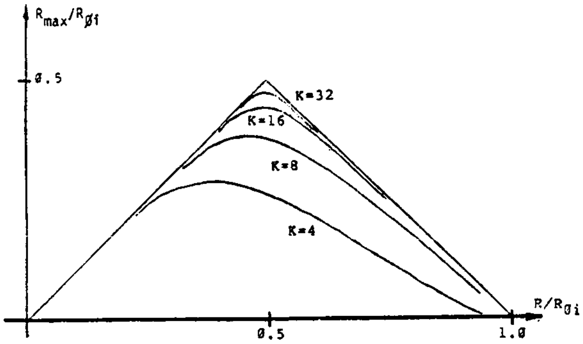


Figure 3. $R_{max}/R_{0i}$ without Interleaving

## IV. DISCUSSION OF RESULTS

From Figures 2 and 3 we see that there is some optimum inner code rate R which maximizes the overall coding rate achievable. When interleaving is used, this optimum inner code rate increases with increasing K, approaching the cut-off rate $R_{0i}$ of the inner channel as K becomes large. On the other hand, when interleaving is not used, the optimum inner code rate is always $\leq 1/2\ R_{0i}$. In both cases, the maximum overall coding rate achievable increases with increasing K, approaching a maximum of $R_{0i}$ with interleaving and $1/2\ R_{0i}$ without interleaving.

It is instructive to compare the interleaved and non-interleaved cases on the same graph. Figure 4 shows both the interleaved case (dotted curves) and the non-interleaved case (solid curves) for three different values of K: K=8, K=32, and $K\to\infty$. We see that the maximum achievable coding rate $R_{max}$ is larger with interleaving than without interleaving for any finite K and all inner code rates R. The difference is particularly significant when $R > 1/2\ R_{0i}$. For large $K(K\to\infty)$, the two curves are merged when $R\leq 1/2\ R_{0i}$, but when $R > 1/2\ R_{0i}$, $R_{max} \to R_{0i}$ for the interleaved curve whereas $R_{max} \to 0$ for the non-interleaved curve. These results clearly indicate that interleaving should be used between the inner and outer codes in order to maximize the overall achievable coding rate.

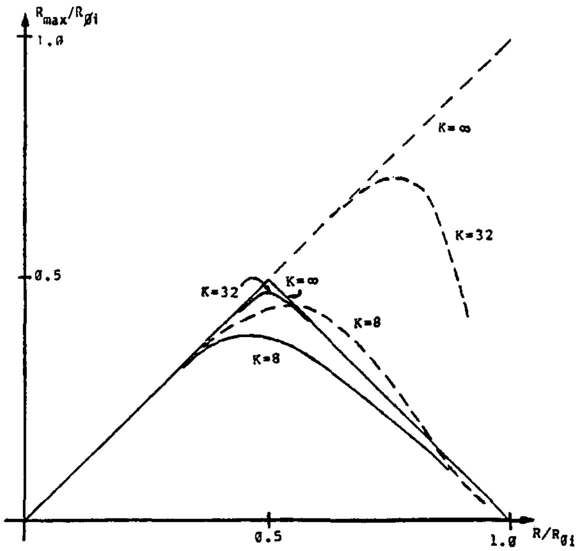


Figure 4. $R_{max}/R_{0i}$ for K=8, 32, and ∞
-------- with interleaving
________ without interleaving

The situation may not be as clear if we consider the channel capacity of the outer channel instead of the cut-off rate. McEliece and Stark [2] have shown that in the case of a block interference channel with a fixed probability (independent of K) of being in the noisy state, interleaving causes the channel capacity to decrease while the cut-off rate increases. An analysis of some specific codes led them to the conclusion that capacity was a better measure of code performance in that case, and hence that using a code designed to correct burst errors gives better performance than a random-error-correcting code with interleaving. We are currently investigating channel capacity for the concatenated code case, and hope to report the results of this research in the near future.

Some other areas under current investigation include extending the above results to the case of convolutional inner codes and the use of side information obtained from the inner decoder by the outer decoder. As pointed out in Section II, the model must be modified to include more than 2 states when the inner code is convolutional, since the error events are of different lengths. In the case of side information, the inner decoder can deliver a reliability indicator along with each decoded block to the outer decoder. We are currently investigating the effects these alterations in the model will have on the calculations of cut-off rate and capacity, and on the performance of a concatenated coding system both with and without interleaving.

## REFERENCES


[1] G.D. Forney, Jr., Concatenated Codes, MIT Press, Cambridge, MA, 1966.

[2] R.J. McEliece and W.E. Stark, "Channels with Block Interference", IEEE Trans. on Info. Theory, IT-30, pp. 44-53, January, 1984.

[3] G.C. Clark and J.B. Cain, Error-Correcting Coding for Digital Communications, Plenum Press, New York, NY, 1981.

[4] G.D. Forney, Jr., "Convolutional Codes II: Maximum Likelihood Decoding" and "Convolutional Codes III: Sequential Decoding", Information Systems Laboratory, Tech. Report No. 7004-1, Stanford University, June, 1972.

[5] A.J. Viterbi, "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm," IEEE Trans. Info. Theory, IT-13, pp. 260-269, April 1967.

[6] J.L. Massey, "Coding and Modulation in Digital Communications," in Proc. Int. Zurich Sem. Digital Commun., Zurich, Switzerland, March, 1974.