Analytical approach for maximizing the average code rate of incremental redundancy schemes

Elisabeth Uhlemann and Lars K. Rasmussen

N.B.: When citing this work, cite the original article.

©2005 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.


DOI: http://dx.doi.org/10.1109/APCC.2005.1554106
Copyright: IEEE

Post-Print available at: Halmstad University DiVA
http://urn.kb.se/resolve?urn=urn:nbn:se:hh:diva-415
Analytical Approach for Maximizing the Average Code Rate of Incremental Redundancy Schemes

Elisabeth Uhlemann†,‡ and Lars K. Rasmussen§

†Centre for Research on Embedded Systems
Halmstad University
Box 823, SE-301 18 Halmsted, Sweden
Email: bettan@ide.hh.se

‡Volvo Technology Corporation
Transport and Telematics Services
Dept. 6600, M1:6, Götaävenksgatan 10
SE-405 08 Göteborg, Sweden

§Inst. for Telecommunication Research
University of South Australia
Mawson Lakes SA 5095, Australia
Email: Lars.Rasmussen@unisa.edu.au

Abstract – An analytical design procedure for maximizing the average code rate of incremental redundancy retransmission schemes is presented. Union bounds are used as the tool to select good puncturing patterns and to obtain bounds on frame error rate for the punctured codes. This, in turn, is used for determining the optimal packet size of each retransmission, so that the average code rate is maximized.

I. INTRODUCTION

A general design procedure for quality of service (QoS) based incremental redundancy (IR) hybrid automatic repeat request (HARQ) protocols was suggested in [1] and further extended in [2]. The QoS parameters identified for the design were deadline (delay) and required reliability. The specific deadline determines the maximum number of retransmissions allowed and a target frame error rate (FER) specifies the required reliability. Following the design procedure, a suitable error control code of rate $r_c = k/n$ is selected as the mother code for the IR scheme based on the target FER, $P_e$, and potential complexity constraints. The mother code is subsequently punctured according to a desired code rate profile, revealing a corresponding rate-compatible code family. The code rate profile is determined by complexity constraints, and forms the fundamental structure of the IR protocol. A codeword for any code family member is a particular partition of the corresponding mother codeword. Similarly, concatenating all the codewords in the rate-compatible code family results in the full mother codeword. Thus, for each retransmission a set of code bits (a packet) is transmitted and provides a codeword of lower rate at the receiver when concatenated with previously transmitted sets of code bits (previously transmitted packets).

In [1] a method for optimizing the packet lengths for each permitted transmission in IR-HARQ schemes was proposed. This method assumed a given puncturing pattern, i.e. a given order in which to transmit the code bits. To demonstrate the optimization process, a semi-analytical approach based on simulated FER curves for the members of the rate-compatible code family was presented in [2], where puncturing patterns were selected based on intuitive arguments.

In this paper, we replace the semi-analytical method in [2] with a rigorous analytical approach based on union bounds. Union bounds on punctured multiple parallel concatenated codes are used both to find good puncturing patterns and to obtain the FER as a function of the codeword length for these puncturing patterns. A three dimensional (3D) multiple parallel concatenated single parity check (PCSPC) code, i.e., three component codes, is used as the mother code in order to obtain a numerical example based on the analytical design procedure.

The paper is organized as follows. The adopted system model is presented in section II. In section III and IV methods for obtaining union bounds for unpunctured respective punctured multiple parallel concatenated codes are detailed. These bounds are then used to find good puncturing patterns in section V. Finally the bounds are used to obtain the FER for all the members of the selected rate-compatible code family, using the preferred puncturing pattern. This may in turn be used to find optimal packet lengths that maximize the average code rate as done in section VI. Section VII contains conclusions.

II. SYSTEM MODEL

Based on the retransmission criterion, it is determined whether or not a retransmission is required, which in this work is done by means of perfect error detection (PED). The retransmission request is then sent through the error free feedback channel. Since the use of an IR-HARQ scheme implies that no repetition of previously transmitted bits is made, the following notation is chosen. If the transmitter receives a retransmission request for a particular frame, a new packet will be transmitted. An information frame or a data frame is a set of $k$ information bits. Each data frame is passed through a rate $k/n$ encoder, generating a codeword of length $n$ code bits with $n-k$ redundant bits or parity bits. This codeword is then partitioned, by means of rate compatible puncturing, into $M$ IR packets. The packets are transmitted according to an IR strategy defined as follows. Let $t_i$ denote transmission $i$ and let $k_i$ and $c_i$ denote the number of information bits and parity bits sent in $t_i$ respectively. Note that $t_i$ must include at least $k$ bits according to $t_i : k_i + c_i$, whereas for $i > 1$, $t_i$ may include less than $k$ bits:

$$t_i : k_i + c_i , \quad t_2 : k_2 + c_2 , \quad t_3 : k_3 + c_3 , \ldots , \quad t_M : k_M + c_M$$  (1)
such that
\[
\sum_{i=1}^{m} k_i \leq k, \quad \sum_{i=1}^{m} c_i \leq n - k, \quad \text{and} \quad k + c_i \geq k, \tag{2}
\]
where \(k_i \in \{0, 1, 2, \ldots, k\}\) and \(c_i \in \{0, 1, 2, \ldots, n - k\}\). Following transmission \(i\), the decoder attempts to decode the corresponding codeword of length
\[
n_t = \sum_{i=1}^{k} k_i + c_i, \tag{3}
\]
If decoding is not successful, transmission \(t_{i+1}\) is made.

III. UNION BOUNDS FOR MULTIPLE PARALLEL CONCATENATED CODES

Let \(C\) be a linear block code with parameters \((n, k)\). The input-output weight distribution, \(B_{w,d}\), denotes the number of codewords in \(C\) with Hamming weight \(d\) generated by information frames of Hamming weight \(w\). The input-output weight enumerating function (IOWEF) of \(C\) is then [3]
\[
B^C(W, D) \triangleq \sum_{w=0}^{n} \sum_{d=0}^{\min(w,e)} B_{w,d} W^d. \tag{4}
\]
The IOWEF can be used to obtain an upper bound on the bit error probability for ML decoding of \(C\) transmitted using BPSK over an AWGN channel as
\[
P_b \leq \sum_{w=1}^{n} A_j Q \left( \frac{2dW E_{\text{b}}}{N_0} \right), \quad \text{where} \quad A_j \triangleq \sum_{k=0}^{j} B_{w,k}. \tag{5}
\]
where \(r = k/n\) is the code rate of the code \(C\). The frame error probability can then be obtained by
\[
P_f \leq \sum_{j=1}^{U} A_j Q \left( \frac{2dW E_{\text{b}}}{N_0} \right), \quad \text{where} \quad A_j \triangleq \sum_{k=0}^{j} B_{w,k}. \tag{6}
\]
For a systematic code having \(d = w + h\), where \(h\) is the weight of the parity bits, the input-redundancy weight distribution, \(B_{w,h}\), is defined as the number of codewords generated by information frames of Hamming weight \(w\), whose parity bits have Hamming weight \(h\), so that the overall Hamming weight is \(d = w + h\). Thus, the input-redundancy weight enumeration function (IRWEF) of a systematic \((n,k)\) block code \(C\) is [4]
\[
B^C(W, H) \triangleq \sum_{w=0}^{n} \sum_{h=0}^{\min(w,h)} B_{w,h} W^h. \tag{7}
\]
Similarly, the IRWEF can be used to obtain an upper bound on the bit error probability for ML decoding of a systematic linear block code \(C\) as
\[
P_b \leq \sum_{w+h=d} \sum_{k} W B_{w,h} Q \left( \frac{2dW E_{\text{b}}}{N_0} \right), \tag{8}
\]
Note that the IOWEF exists for a linear block code regardless whether or not it is in systematic form, whereas the IRWEF only exists for a systematic code.

Using the uniform interleaver introduced in [4], the IRWEF or the IOWEF for a concatenated code can be obtained based on knowledge of the IRWEF or the IOWEF of its constituent codes. Consider the conditional IRWEF, \(B^C(W, H)\) conditioned on \(w\), enumerating the parity bits generated by \(C\) corresponding to the respective information frames of weight \(w\). It is related to the IRWEF as [4]
\[
B^C(W, H) = \sum_{w} W^w B^C(W, H). \tag{9}
\]
From the above definitions it is apparent that the conditional IRWEF of the second code, \(B^{C_2}(w, H)\), becomes independent from that of the first code, \(B^{C_1}(w, H)\) due to the uniform randomization produced by the interleaver. Hence the conditional IRWEF for a 2D parallel concatenated block code, \(C_{2D}\), is [4]
\[
B^{C_{2D}}(w, H) = \frac{B^{C_2}(w, H) \cdot B^{C_1}(w, H)}{k}. \tag{10}
\]
This can be generalized to \(U\) component codes concatenated in parallel and separated by \(U - 1\) interleavers as
\[
B^{C_{U}}(w, H) = \prod_{j=1}^{U} \left( \frac{B^{C_j}(w, H)}{k} \right)^{j-1}, \quad \text{where} \quad k = l k_{C_j}. \tag{11}
\]
and \((n_{C_j}, k_{C_j})\) are the parameters of \(C_j\). This has also been concluded independently in [5, 6].

IV. UNION BOUNDS FOR PUNCTURED MULTIPLE PARALLEL CONCATENATED SPC CODES

Extending the above results, upper bounds on the FER can be derived for punctured multiple parallel concatenated codes. As an example, we consider a 3D PCSPC with an SPC(8,7) code as component code. The IRWEF for an SPC(8,7) code is:
\[
B^{C_{SPC}(8,7)}(W, H) = 1 + 7WH + 21W^2 + 35W^3H + 35W^4 + 21W^5H + 7W^6 + W^7H. \tag{12}
\]
Three of these SPC(8,7) codes used in a 3D PCSPC code, \(C_{3D}\), with \(k = 7\) = 343 implies \(k = l k_{C_{SPC(8,7)}}\), with \(l = 49\). Hence, the IRWEF for these longer component codes, \(C_{il}\) for \(i = 1, 2, 3\) can be obtained by convolving the IRWEF in (12) \(l = 49\) times. Thereafter the joint IRWEF can be obtained as
\[
B^{C_{il}}_{49}(w, H) = \frac{B^{C_{il}}(w, H) \cdot B^{C_{il}}(w, H) \cdot B^{C_{il}}(w, H)}{343^2}. \tag{13}
\]
Assume now that we want the joint IRWEF for the \(C_{2D}\)

482
with 25 parity bits punctured in the third dimension. The IRWEF for a punctured SPC(8,7) is then required, or rather the weights of an information frame of size seven, since puncturing a single parity bit from an SPC code implies that the information frame is now uncoded. The IRWEF for an SPC(8,7) with the parity bit punctured is given by

\[ B_{r_{spc}(8,7)}^{\text{spc}(8,7)}(W, H) = 1 + 7W + 21W^2 + 35W^3 + 35W^4 + 21W^5 + 7W^6 + W^7, \]  

(14)

where \( pp = 1 \) denotes that one parity bit is punctured. The IRWEF for the component code, \( C_{id} \), for \( l = 49 \) when 25 bits are punctured is obtained by convolving the IRWEF in (12) 24 times and then convolving with (14) 25 times. Consequently, for each parity bit we wish to puncture we will use the \( B_{r_{spc}(8,7)}^{\text{spc}(8,7)}(W, H) \) and for each parity bit we wish to keep the \( B_{r_{spc}(8,7)}^{\text{spc}(8,7)}(W, H) \) is used. This procedure can then be employed for each dimension we wish to puncture. Note that even though we may puncture several parity bits from one dimension, no more than one parity bit from each component codeword can be punctured, since we are dealing with SPC codes. If more advanced codes are used, each component codeword may contain more than one parity bit and that must be treated specially since the different puncturing patterns may result in different IRWEFs.

So far, when looking at bounds for parallel concatenated codes, the IRWEF has been used since each component code simply adds redundancy to one common information frame. However, when information bits are to be punctured the IOWEF is needed. This is partly due to the fact that \( d \) can no longer be obtained from \( w + h \) and partly since puncturing more information bits from a component codeword than the number of parity bits added, would lead to the weight of the redundancy, \( H \), being lower than zero. Obviously, puncturing more bits than what has been added in terms of redundancies makes little sense in a stand alone component code since the local invertibility of the code has to be guaranteed. When using the component code in a concatenated code this may, however, still be relevant.

A PCSPC code can be constructed either using a separate systematic part or, alternatively, as having the systematic part embedded in one of the component codes. For the first encoder, we would use three IRWEFs as in (13) when obtaining the joint IRWEF of the 3D PCSPC code, whereas for the second encoder, one IOWEF and two IRWEFs can be used to obtain the joint IOWEF of the equivalent 3D PCSPC code as

\[ B_{r_{spc}(3d)}^{\text{spc}(3d)}(w, D) = \frac{B_{r_{spc}(8,7)}^{C_{id}}(w, D)B_{r_{spc}(8,7)}^{C_{id}}(w, H)B_{r_{spc}(8,7)}^{C_{id}}(w, H)}{\binom{343}{w}}. \]  

(15)

Note that using the IOWEF of code \( C_1 \) results in the IOWEF, \( B_{r_{spc}(3d)}^{C_{id}}(w, D) \), rather than the IRWEF of \( C_{id} \) since the weights of the systematic part have already been added. Puncturing of information bits is now made in the component code contributing with its IOWEF, in this case \( C_1 \), whereas the puncturing of parity bits is made in the IRWEF or IOWEF of the corresponding code. Calculating the joint IOWEF as in (15) is only recommended if puncturing of information bits is needed. Otherwise (13) should be used since that makes the computation of the bound less complex.

Since an SPC code contains more than one information bit, two or more information bits belonging to the same component codeword may be punctured. Depending on particular interleaver realizations, two or more punctured information bits may also belong to the same component codeword in one or more additional dimensions of the overall code. The uniform interleavers used when calculating the joint IOWEF take into account that information bits may be paired up in the same component codeword again in another dimension. Consequently, we only need to consider the possibility of occurrence in the very first component code, before the first interleaver. Thus, the weight spectrum of all different puncturing scenarios that may occur in the first component code are required. Hence component code weight spectra with one or more information bits punctured – both when the corresponding single parity bit is punctured and unpunctured are required. Note that when more than one information bit are punctured from an SPC component codeword, the local invertibility for that component codeword is lost. However, the overall invertibility may still be maintained due to the addition of new parity bits in other dimensions.

V. SELECTION OF GOOD PUNCTURING PATTERNS

Given a family of rate-compatible codes obtained from puncturing a mother code, we want to select a good puncturing pattern in terms of maximizing the average code rate of the corresponding IR-HARQ system. By this is meant that a frame should be accepted after transmission of a minimum number of code bits. Several methods to select the better candidate given two alternative puncturing patterns and their respective weight spectrum have been suggested. Previous work has for example proposed to fit a regression line to the first 30 components of the joint IRWEF and then selected the puncturing pattern which has a regression line with the minimum slope, [7]. Another alternative is to pick the puncturing pattern that results in the maximal minimum distance, \( d_{\text{min}} \), and if there is a tie, pick the one with the lowest multiplicity at this distance, [7, 8]. Since the latter suggestion first maximizes \( d_{\text{min}} \) and secondly minimizes the multiplicities, it is likely to choose a pattern that performs well at high signal-to-noise ratios (SNRs) where \( d_{\text{min}} \) is the dominating factor. At low SNRs, higher weights than \( d_{\text{min}} \) yield a non-negligible contribution and hence the effects of multiplicities at the closest distances are more dominating in this area. The regression line suggestion provides a measure of the rate of growth of the weight multiplicities and it is thus likely to find patterns that perform better at low SNRs. Since an IR-HARQ system generates retransmissions when the SNR is low, we are interested in patterns that perform well in this area.
Once the joint IOWEF has been calculated, obtaining the bound is fast and simple. Further, the bound on FER for multiple concatenated SPC codes is found to be fairly tight in the relevant SNR region. Consequently, in this work the candidate puncturing pattern that results in the lowest bound on FER in the SNR region of interest for the chosen IR-HARQ scheme is selected. In order to obtain different candidate puncturing patterns to compare, we use the following procedure, which is similar to the procedure used in [7]:

1. Define an all ones vector, \( \mathbf{p} \), of length \( n \) and set \( i = 0 \) and \( b_i = 0 \).

2. Select the next code rate in the rate-compatible code family by setting \( i = i + 1 \) so that code rate \( k/(n-b_i) \) is considered, where \( \mathbf{b} = \{b_1, b_2, ..., b_i\} \) is a vector containing the number of bits to be punctured for obtaining the \( f \) different rates of the corresponding code rate profile. Note that \( b_i < h_{i+1} \).

3. Form a set of all \( \binom{n-h_i}{h_i} \) possible candidate puncturing patterns of this rate by temporarily resetting the appropriate number of positions, \( b_i - h_{i+1} \), among the remaining elements in \( \mathbf{p} \) that are still set to one, symbolizing that the corresponding bits are punctured.

4. Calculate the union bound on FER for each of the puncturing patterns obtained in Step 3.

5. Select the candidate puncturing pattern yielding a bound with the lowest FER in the relevant SNR region. If there is a tie, increase the SNR region in the appropriate direction.

6. Assign the best pattern to the corresponding rate, \( k/(n-b_i) \), and permanently reset the corresponding positions in \( \mathbf{p} \).

7. If \( i < f \), go back to Step 2, otherwise the search is finished.

Consider again the 3D PCSPC code with all possible code rates available, i.e., \( \{k/n, k/(n-1), k/(n-2), ..., 1\} \). If only parity bits are allowed to be punctured, the above search procedure concludes that the best puncturing pattern is dimension-wise puncturing, i.e., if we have started puncturing bits pertaining to one encoder, we should continue doing so until all these bits have been punctured. Hence parity bits from a “new” encoder should only be punctured once all the bits from the “old” encoder have been removed. This result is intuitive, since each time we start puncturing a new dimension, the minimum number of parity bits protecting an information bit is reduced. However, if we continue to puncture in the same dimension, we no longer affect this minimum number, but are only increasing the multiplicities of information bits encoded with fewer parity bits – something that has less influence at low SNRs. Due to the simple structure of SPC codes it is irrelevant which parity bit from a particular dimension that is punctured as long as random interleavers are used. Further, for PCSPC codes it is irrelevant which dimension is chosen first. This drastically reduces the number of possible candidate puncturing patterns in Step 3.

If we use the same mother code, but instead want a different family of rate-compatible codes, more than one bit may be punctured concurrently in Step 3. A search for the best puncturing pattern when say \( b_i - h_{i+1} = 30 \) parity bits are punctured simultaneously may not result in the same pattern. Therefore the resulting dimension-wise puncturing pattern is only guaranteed to be the best for the rate-compatible code family including all possible rates.

Puncturing information bits yields more candidate puncturing patterns in Step 3, since each component codeword has more than one information bit. When puncturing an information bit, \( d_{\text{min}} \) of all component codes is reduced, since the systematic information bit is part of all component codes. Consequently, at high SNR, where \( d_{\text{min}} \) is the dominating factor it is best to puncture parity bits. At lower SNRs the multiplicities are more important and hence it seems preferable to puncture information bits. However, puncturing information bits may result in the local invertibility being lost. Depending on the particular interleavers chosen, the other dimensions may or may not restore the invertibility. This means that the overall invertibility cannot be guaranteed, due to the use of uniform interleavers.

If we instead start the search for good puncturing patterns at the highest possible rate, one in the case with all rates, the best puncturing pattern is to transmit all \( k \) systematic bits first. This inherently results in the dimension-wise puncturing pattern that only allows puncturing of parity bits. It may be argued that for IR-HARQ schemes, we should start searching for a good puncturing pattern at the highest possible code rate, since higher rates are, in fact, transmitted first.

Note also that for the 3D PCSPC mother code, dimension-wise puncturing is the pattern yielding the lowest complexity, since not all decoders need to be activated until parity bits pertaining to them have been received.

VI. FINDING GOOD PACKET LENGTHS

The bounds on FER for different codeword lengths, using the dimension-wise puncturing pattern obtained in the previous section, are used to optimize the packet lengths in an IR-HARQ scheme to maximize the average code rate.

The code rate of a non-ARQ system is \( r_c = k/n \), but for an IR-HARQ system we can only get an average code rate since the retransmissions yield a variable code rate based on the current channel conditions. The average code rate is a function of the probability of a retransmission, \( P_{\text{ARQ}}(n) \), which in turn is a function of the received codeword or, if given a particular puncturing pattern, a function of the codeword length \( n \). Here, we assume that the criterion for declaring a retransmission is such that a frame accepted based on a codeword of length \( n \) will also be accepted based on a codeword of length \( n+1 \).

Assume that we have an IR-HARQ system that is limited to two packets, \( M = 2 \), e.g., first \( n_1 = k + c_1 \) is transmitted and thereafter \( c_1 \) parity bits are transmitted only when necessary, resulting in a codeword of length \( n_2 = k + c_1 + c_2 \). Consequently, when using this scheme, some of the frames have been accepted based on a code rate of \( k/(k + c_1 + c_2) \), whereas others needed a code rate of \( k/(k + c_1 + c_2) \). The average code rate of the IR-HARQ system can be expressed as a function of \( P_{\text{ARQ}}(n) \) for \( M = 2 \) as \( r_c^{M-1} = k/(k + c_1 + c_2 P_{\text{ARQ}}(n)) \). This
can be generalized to $M$ transmissions and the packet lengths, $c_i$, for $i=1,2,\ldots,M$ can be optimized so that the average code rate is maximized, [1]. For $M$ transmissions this is an $M$-dimensional discrete maximization problem [1]. Since $P_i$ is used to determine the total amount of redundancy needed to ensure the required QoS level, $n_p$, the last transmission should always include the remaining parity bits that have not yet been transmitted so that

$$c_M = (n_p - k) - \sum_{i=1}^{M-1} c_i, \quad \text{where} \quad c_i \in \{0,1,\ldots,n_p - k\}. \quad (16)$$

The discrete maximization problem then becomes $M-1$ dimensional according to

$$\left[ c_1, c_2, \ldots, c_{M-1} \right] = \arg \max_{c_1,c_2,\ldots,c_{M-1}} \frac{k + c_1 + \sum_{i=2}^{M-1} c_i P_{ARQ}(n_{i-1}) + c_{M-1} P_{ARQ}(n_{M-1})}{k + c_1 + \sum_{i=2}^{M-1} c_i P_{ARQ}(n_{i-1}) + c_{M-1} P_{ARQ}(n_{M-1})}. \quad (17)$$

Since we use PED, $P_{ARQ}(n_i)$ is equal to the FER and can be found by the union bounds described in the previous section. In Figure 1 the FER for the 3D PCSPC example code is plotted as a function of the received codeword length. Note that the gain is significant whenever a full dimension is completed. For $M=2$ and $n_p = k + 147$, using (17) and the FER curves in Figure 1 yields the result reported in Figure 2 for different values of $E_b/N_0$. It can be seen from Figure 2 that there is indeed an optimal value of $c_1$ that maximizes the average code rate. Note also that the value of $c_1$ changes for different $E_b/N_0$.

It can also be concluded that the more retransmissions that are allowed, the higher average code rate can be obtained. Hence, transmitting one code bit at a time yields an achievable upper bound on the average code rate. This way each information frame is accepted following the transmission of a minimum number of parity bits. The analytical results are also compared to Monte-Carlo simulations using the optimal packet lengths and the selected puncturing patterns in [9]. The average code rate obtained using simulations corresponds well with the average code rate obtained analytically.

VII. CONCLUSIONS

An analytical design method based on union bounds for maximizing the average code rate in IR-HARQ schemes is presented. First, a good puncturing pattern is selected and then the individual packet lengths are optimized. The more retransmissions that are allowed, the higher average code rate can be obtained.

REFERENCES