

Partial Reed Solomon Codes for Erasure Channels

Shirish S Karande and Hayder Radha
 Department of Electrical and Computer Engineering
 Michigan State University
 East Lansing, MI - 48824, USA
 {karandes, radha}@egr.msu.edu

Abstract - In this paper, we introduce a new family of linear block codes, which we refer to as Partial Reed Solomon (PRS) Codes. These codes are specifically designed and optimized for real-time multimedia communication over packet-based erasure channels. Based on the constraints and flexibilities of real time applications, we define a performance measure, *message throughput* (τ_m), which is suitable for these applications. This measure differentiates the notion of optimum codes for the target multimedia applications as compared to performance measures that are used for non-realtime data. Based on the proposed measure, we combine the advantages of lowering the density of a code for near capacity performance with the high decoding efficiency of Reed Solomon (RS) codes, in order to design optimum PRS codes. Then, we demonstrate, through an example of a Binary Erasure Channel (BEC), that at near-capacity coding rates, appropriate design of a PRS code can outperform an RS-code. We extend this analysis and optimization for a general BEC over a wide range of channel conditions. Moreover, as compared with RS codes, the proposed PRS codes provide a significantly improved graceful degradation when the number of losses exceeds the number of parity symbols within the code block. This is a highly desirable feature for realtime multimedia applications.

I. INTRODUCTION

A fundamental requirement of any realtime application is the transmission of message data at a minimum desired rate R . In general, this minimum rate should be maintained to achieve a certain quality. The minimum rate requirement translates to the transmission of a minimum number of K message symbols within an N -symbol code block: $R=K/N$. Consequently, one of the constraints in the design of linear block codes for realtime applications is the usage of a maximum number ($N-K$) of parity symbols within the N -symbol block. Moreover, the maximum value of code block size N is also constrained. Therefore, unlike non-realtime applications that may have the flexibility in selecting N and $R=K/N$, realtime applications, in general, have to employ (adhere to) a block code with a pair-constraint (N, K).

Performance criteria for LBC codes, which are used for non-realtime data, are not always suitable for realtime applications. For example, a non-realtime LBC code can be evaluated based on the number of symbols needed to *perfectly* recover all of the original message symbols. In general, for realtime applications, *perfect* recovery, and consequently perfect reconstruction, of the original message symbols is not a hard requirement, because a wide range of

practical application-layer error resilience and concealment methods can be used to compensate for lost data [3]. Meanwhile, it is crucial to deliver the realtime application layer with the maximum number of the message symbols that are transmitted by the system. Therefore, the probability of a *message* symbol loss (after channel decoding) is a key performance parameter. We denote to this probability by p_m . Hence, the parameter $\tau_m=(1-p_m)$, which represents the probability of receiving a message symbol by the realtime application (after channel decoding), is a measure of the end-to-end message symbol throughput. One of the key objectives of the family of the Partial Reed Solomon (PRS) codes that are proposed in this paper is to maximize this throughput measure τ_m . (For the remainder of this paper, we will refer to τ_m as the *message throughput*.)

Practical multimedia error concealment and resilience methods usually become useless when the number of losses is beyond an application-dependent threshold. Consequently, it is very crucial for LBC codes to perform well when the number of lost message symbols is large by recovering the majority of these lost message symbols. Meanwhile, and although it is desirable, it is less crucial for these codes to provide perfect recovery when the number of losses (before or after channel decoding) is very small (e.g., one or few symbols) due to the maturity of powerful multimedia processing techniques. This desirable feature highlights one of the key problems with current LBC codes that are used widely for realtime video. It is well known, for example, that when a RS code block experience a number of losses that is larger than the number of parity symbols, then the code is incapable of recovering any of the lost message data.

Experiencing a number of losses that is larger than the number of parity symbols is quite feasible over channels with time-varying characteristics (e.g., the Internet and wireless networks [4]-[5]), even if, "on average", the message rate R is lower than the channel capacity. This is particularly true when the message rate R is close to (but may still be lower than) the channel capacity. Moreover, and due to (a) the large amount of data that is inherently needed for representing multimedia (in particular video) signals, and/or (b) the compressed representation of these signals is normally encoded ahead of time at a certain minimum rate that cannot be further reduced in realtime by the sender, it is quite often when multimedia applications operate very closely to channel capacity. Consequently, one of the main objectives of this work is the design of LBC codes that are capable of achieving high message throughput τ_m when the rate is close to (but still lower than) channel capacity and when the number of losses L exceeds the number of parity symbols ($N-K$).

II. PARTIAL REED SOLOMON CODES

It can be shown that the decoding of a codeword transmitted over an erasure channel is equivalent to solving a system of equations, represented by the parity check equations [1]. The erased symbols represent the unknowns in the system of equation. Thus for a given (N, K) and a given graph density representing an LBC code, as the probability of channel erasure p increases, the average number of unknowns in each parity check equation also increase. Also, as the number of unknowns in a parity check equation increase, the probability of that equation being successfully solved decreases. Due to this, when the coding rate is near (or above) channel capacity, it becomes necessary to reduce the number of message symbols that are protected by each parity symbol. This is equivalent to reducing the density of the code.

The iterative algorithms used for decoding current LDPC codes, limit the decoding process to decoding of graphs without short-cycles¹. This constraint has influenced the design of most of the current LDPC codes. If a code is based on $GF(2)$, then the above constraint of designing a graph without short cycles is not a severe one. But, for codes based on $GF(q)$, limiting the code design to graphs without short cycles, can lead to an over-compromise on the decoding process in terms of time-complexity. Moreover, it has been shown that codes designed on fields higher than $GF(2)$, can exhibit a much-improved performance [2] in terms of error recovery. Since a key objective of our effort is to maximize the message throughput (i.e., lost-symbol recovery), we did not want to constrain our code-design to graphs without cycles.

Meanwhile, decoding algorithms for a general code (with cycles) based on $GF(q)$ can have a very high time complexity. Thus we found it necessary to limit our code design to a family of codes, where the entire codeword could be broken down into sub-codes that resemble RS codes. This allows us to use algorithms developed for efficient decoding of RS codes, for decoding of these RS based sub-codes. Decoding of individual subcodes can facilitate the decoding of the entire codeword. After this brief discussion of the motivation for the proposed PRS codes, we introduce the general structure of these codes.

For a given realtime-pair constraint (N, K) , we denote a general PRS code of order s by $(N, K, \Lambda_s)_q$. Here q represents the underlying field². The order of the field is constrained by the equation $q > N$, where N represents the total number of symbols in a single codeword and K represents the number of message symbols in a codeword.

Λ_s represents a $2 \times (s+1)$ matrix given by $\begin{bmatrix} N_1 \cdots N_{s+1} \\ K_1 \cdots K_{s+1} \end{bmatrix}$.

¹ Each step in the iterative decoding algorithms is simplified into solving a single equation with a single unknown. Presence of cycles in a graph reduces the efficiency of this simplification.

² In all further discussion we shall drop q from the notation and assume that the order of the field on which the code is based has been pre-specified.

The entries of matrix Λ_s are constrained by the following equations:

$$N_i > K_i \forall i \in [1, s], \quad K_i > 0 \forall i \in [1, s], \quad N_{s+1} = K_{s+1}$$

$$\text{and } N = \sum_i N_i, \quad K = \sum_i K_i.$$

Thus Λ_s gives an s -partition on the set of parity symbols and a $(s+1)$ -partition on the set of message symbols. The code is designed such that, $\forall i \in [1, s]$, the pair (N_i, K_i) forms an RS-subcode over $GF(q)$ and the K_{s+1} number of message symbols are transmitted without any protection. Thus the code-graph can be divided into $(s+1)$ disjoint sub-graphs. Obviously such a code graph does not have full density and the density of the overall code has been lowered. It should be noted that an order 1 ($s=1$) PRS code with $N_2 = K_2 = 0$ is equivalent to the traditional full density RS code. In general, a PRS code with $N_{s+1} = K_{s+1} = 0$ does not include any subset of message symbols that are not protected.

III. OPTIMAL PRS CODES

In this section we identify the class of optimal PRS codes for a Binary Erasure Channel (BEC) based on the message throughput criterion. We show, and with the support of some experimental evidence that, for a BEC, the optimal PRS code is given by an order 1 PRS code (i.e., PRS-1). As mentioned above, the parameter used to measure performance of a code here, is message throughput. Thus a code that maximizes this parameter will be the optimal code. We shall also extend our optimality analysis to two simple feedback channels. Before we present our experimental evidence, we shall prove two lemmas, these lemmas help us to limit the ensemble of codes we have to consider to find the optima. The following notations and propositions are used by the lemmas. Let,

• Ψ_{N,K,K_3} be a set containing all PRS codes of order 1

with $\Lambda_1 = \begin{bmatrix} N_1 & K_2 \\ K_1 & K_2 \end{bmatrix}$ and all PRS codes of order 2 with

$\Lambda_2 = \begin{bmatrix} N_1 & N_2 & K_3 \\ K_1 & K_2 & K_3 \end{bmatrix}$. An example of Ψ_{N,K,K_3} is the set

$\Psi_{N,K,0}$. In addition to *all* possible PRS codes of order 1,

$\Psi_{N,K,0}$ includes only a subset of all PRS codes of order 2 (i.e., PRS-2). This subset represents PRS-2 codes where each message symbol is protected by at least one parity symbol. In other words, no message symbols in this particular PRS-2 subset, which is included in $\Psi_{N,K,0}$, is left unprotected.

• **Proposition 1 (P1):** $\forall (N, K)$ the optimal PRS code in the set $\Psi_{N,K,0}$ is an order 1 PRS code.

• **Proposition 2 (P2):** $\forall (N, K), \forall K_3 < K$ the optimal PRS code in the set Ψ_{N,K,K_3} is an order 1 PRS code.

• **Proposition 3 (P3):** $\forall (N, K), \exists$ an order s PRS code, that performs better than all order $(s+1)$ PRS codes.

LEMMA1: For a BEC $\mathbf{P1} \Rightarrow \mathbf{P2}$. In other words, if the optimal code within the set $\Psi_{N,K,0}$ is a PRS-1 code, then the optimal code in the more general set Ψ_{N,K,K_3} , $\forall K_3 < K$ is also a PRS-1 code.

Proof: Consider the optimal code on the set $\Psi_{(N-K_3),(K-K_3),0}$. $\mathbf{P1}$ implies that the optimal PRS code on this set is a PRS code of order 1. Since adding unprotected symbols to a block will not change the relative performance of two codes on a BEC, the optimal PRS code in the set Ψ_{N,K,K_3} is also an order 1 PRS code. Thus for a BEC $\mathbf{P1} \Rightarrow \mathbf{P2}$.

LEMMA 2: For a BEC $\mathbf{P1} \Rightarrow \mathbf{P3}$.

Proof: Let the optimal PRS code of order $(s+1)$ be given by

$$(N, K, \Lambda_{s+1}), \text{ such that } \Lambda_{s+1} = \begin{bmatrix} N_1 \cdots N_{s+2} \\ K_1 \cdots K_{s+2} \end{bmatrix}. \text{ Using } \mathbf{P1} \text{ we}$$

can conclude that optimal PRS code in $\Psi_{(N_1+N_2),(K_1+K_2),0}$ is an order 1 PRS code. For a BEC the relative performance of two codewords is not going to change due to addition of identical code sections. This implies that there exists $K^* < (K_1 + K_2)$ such that, the performance of (N, K, Λ_s) PRS code with

$$\Lambda_s = \begin{bmatrix} (N_1+N_2) & N_3 & \cdots & N_{s+1} & (K_{s+1}+K_1+K_2-K^*) \\ K^* & K_3 & \cdots & K_{s+1} & (K_{s+1}+K_1+K_2-K^*) \end{bmatrix}$$

will be better than any PRS code of order $(s+1)$. Thus we can conclude that for a BEC $\mathbf{P1} \Rightarrow \mathbf{P3}$.

Lemma's 1 and 2 reduce the ensemble of codes over which we need to search for an optimal code to the set $\Psi_{N,K,0}$. It should be noted that any PRS code of order 2 belonging to the set $\Psi_{N,K,0}$ can be represented by

$$\Lambda_2 = \begin{bmatrix} N_1 & N-N_1 & 0 \\ K_1 & K-K_1 & 0 \end{bmatrix}. \text{ Thus, the search space to find the}$$

optimal PRS code can be further reduced by noting that the performance of the above PRS code will be unchanged even if $\Lambda_2 = \begin{bmatrix} N-N_1 & N_1 & 0 \\ K-K_1 & K_1 & 0 \end{bmatrix}$. Thus we constraint the values of

N_1 and K_1 by the following equations: $(N_1 - K_1) > (N - K)/2$ and $K_1 > K/2$. Based on our simulation results we formulate a conjecture

CONJECTURE 1: For a BEC channel $\mathbf{P1}$ is true.

We verified the validity of conjecture 1 for different values of N , K and p , but due to brevity of space only some results for $N=100$ and $K=88$ are presented. Any PRS code of order 2 belonging to the set $\Psi_{100,88,0}$ can be

represented by $\Lambda_2 = \begin{bmatrix} N_1 & N-N_1 & 0 \\ K_1 & K-K_1 & 0 \end{bmatrix}$. Thus in all the

figures in this section the x-axis shows the value of $(N_1 - K_1)$, the y-axis shows the value of K_1 and the z-axis shows the message throughput of the corresponding code. Furthermore, in each figure it should be observed that $\mathbf{P1}$ is

validated if the code that has maximum message throughput satisfies $N_1 - K_1 = N - K$. This represents a PRS-1 code since all of the parity codes are being allocated to protect only one subset (with K_1 elements) of the message symbols.

The other subset of message symbols (with $K - K_1$ elements) is either empty (i.e., $K - K_1 = 0$) or not protected at all. In the case when $K - K_1 = 0$, we have a traditional RS code where all of the message symbols are protected by all of the parity symbols. Figure 1(a) shows the experimental results for $p = 0.05$ where the channel capacity is 0.90. It should be noted that in this case the coding rate (0.88) is below channel capacity. It can be seen in that the optimal PRS code for a BEC in $\Psi_{100,88,0}$ is an order 1 PRS code.

Thus using lemma's 1 and 2 it can be concluded that for a BEC the optimal code is given by PRS code of order 1. Although it is possible for an optimal PRS code to turn out to be equivalent to an RS code depending on the channel conditions, it should be noted that in figure 1(a) though the coding rate is lower than the channel capacity, the optimal code is given by a PRS code of order 1 that is not equivalent to a RS code.

It is also important to investigate the performance of PRS codes for coding rates greater than channel capacity. Since, even if "on-average" the coding rate is lower than channel capacity, the time varying nature of a channel can make the scenarios when the number of losses are greater than $N - K$, or when the coding rate is higher than channel capacity feasible. A possible way to mitigate this problem is to use some feedback information to adapt to the channel [6]. Most of the current adaptive FEC schemes change the rate of the underlying code. For real-time applications when the data transmission rate is constraint by K message symbols per N code symbols, such a scheme may not be suitable. Optimizing a PRS code to facilitate recovery when the number of losses are greater than $N - K$ or the data transmission rate is temporarily higher than the coding rate can facilitate the design of a more suitable adaptive block code. In the above scenarios it is impossible to recover all the lost information without changing the rate of the code, but adaptively changing the design of a PRS code without changing its rate can facilitate partial recovery of information. For multimedia applications partial recovery of information in adverse channel conditions is a desirable feature when compared to permanent losses of erasures.

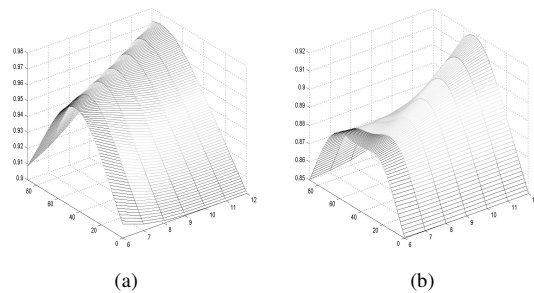


Figure 1 $N=100$, $K=88$, (a) $p=0.1$ (b) $p=0.15$. In the above figures the axis on the left is y-axis while the one on the right is x-axis.

Figure 1(b) shows the performance of PRS codes belonging to $\Psi_{100,88,0}$ for $p = 0.15$. It can be observed that even when the coding rate is greater than channel capacity the optimal PRS code is a PRS code of order 1.

IV. OPTIMAL PRS - 1 CODES

In this section we further evaluate and analyze the performance of PRS codes of order 1 (PRS -1). As the design of a PRS code is completely determined by our choice of K_1 , we use a shortened notation for order 1 PRS code. Thus a PRS code denoted by (N, K, K_1) is equivalent to a PRS code denoted by (N, K, Λ_1) where $\Lambda_1 = \begin{bmatrix} N - K + K_1 & K - K_1 \\ K_1 & K - K_1 \end{bmatrix}$. Thus the optimal PRS code will be obtained by choosing an optimal value of K_1 , denoted by K^* .

The probability of a message symbol loss (after channel decoding) for a (N, K, K_1) PRS-1 code over a BEC with probability of erasure p is given by

$$p_m = \left(\frac{1}{K} \right) \cdot \left((K - K_1) \cdot p + \left(\frac{K_1}{(N - K) + K_1} \right) \cdot \left(\sum_{i=(N-K)+1}^{N_1} i \cdot \binom{N_1}{i} \cdot p^i \cdot (1-p)^{(N-K)+K_1-i} \right) \right)$$

Equation 1

The optimal value of K_1 can be obtained by minimizing Equation-1. Since $\tau_m = (1 - p_m)$, this is equivalent to maximizing the message throughput. This equation can be used to show that there exist coding rates below channel capacity for which an optimal PRS code can outperform an RS code of similar rate and block-length. Figure 2 shows one such example. It can be clearly seen that for coding rate of 0.88 and block-length of 100, if the channel capacity is lesser than approximately 0.905 then the optimal PRS code outperforms the classical RS code. It can be seen that the optimum PRS codes maintain much better overall message throughput for coding rates beyond channel capacity as shown in Figure 3.

It was also observed that the dependence of optimal PRS codes on channel capacity and coding rate is symmetric. It can be concluded that for a given probability of erasure and block-length there exists a critical coding rate lesser than channel capacity, such that, for all coding rates above this critical value, there exists an optimal PRS code that can outperform the traditional RS code. Moreover, it can be shown for the PRS-1 codes, that as $N \rightarrow \infty$, the critical rate becomes equal to the channel capacity of the BEC.

V. GRACEFUL DEGRADATION

The performance of a PRS-1 code optimized for a particular channel condition degrades much more gracefully than a RS code with identical coding rate. This property can be suitably exploited in adaptive FEC schemes. The channel feedback information is an estimate of the channel conditions and hence it cannot be assumed to be always

accurate. In such circumstances a PRS -1 code is less susceptible to an inaccurate estimate of channel conditions as compared to an RS based code. (Due to space limitations, the results of this section are not provided here; however, they will be presented in a future paper.)

VI. CONCLUSIONS AND SUMMARY

In this paper, we introduced a family of codes, which we refer to as PRS codes. These codes can be considered to be a reduced density version of RS codes. It was observed that for all the channels considered in this paper, the optimal PRS code was an order 1 PRS code. It was shown that for a BEC channel, and for coding rates close to channel capacity, suitably designed PRS codes can perform better than the traditional RS codes. It was shown, that it is possible to design a "fixed coding rate" adaptive FEC scheme based on these PRS codes. The degradation in performance of PRS codes is more graceful than that of the RS codes. This feature can be suitably exploited to facilitate improved performance, for time varying channels and also satisfy the needs of modern robust multimedia communication schemes.

REFERENCES

- [1] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, D. A. Spielman, "Efficient Erasure Correcting Codes," IEEE Transactions on Information Theory, vol. 47, pp 569 - 584, 2001.
- [2] M. C Davey, D. J. C. MacKay, "Low Density parity check codes over GF (q)," IEEE Communication Letters, 1998.
- [3] Y. Wang, Q. Zhu, "Error Control and Concealment for Video Communications: A Review," Proceedings of the IEEE, 1998.
- [4] M. Yajnik, J. Kurose, D. Towsley, "Packet Loss Correlation in the Mbone Multicast Network," Proceedings of IEEE Global Internet Conference, 1996.
- [5] D. Loguinov, H. Radha, "End-to-End Internet Video Traffic Dynamics: Statistical Study and Analysis," IEEE INFOCOM, 2002.
- [6] J. C. Bolot, S. Fosse-Parisis, D. Towsley, "Adaptive FEC-based error control for Internet telephony," Proceedings of IEEE INFOCOM '99, vol. 3, pp. 1453-1460, 1999.

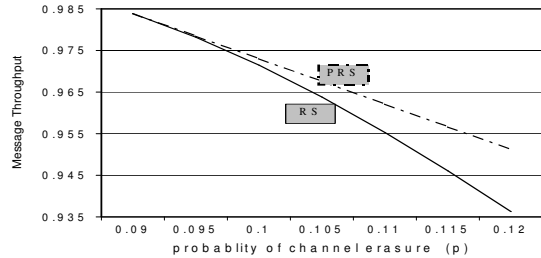


Figure 2 Message throughput performance of optimum $(N, K, K^*) = (100, 88, K^*)$ PRS codes as compared with the RS $(N, K) = (100, 88)$ code over different Binary Erasure Channel (BEC) conditions. The coding rate K/N is lower than the channel capacities

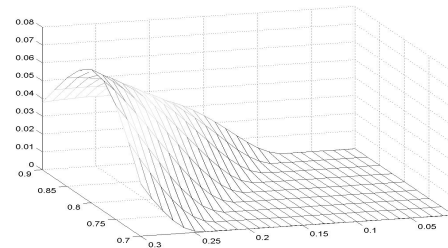


Figure 3 $N = 100$: Difference between PRS -1 and RS