Iterative Soft Decoding of Reed-Solomon Convolutional Concatenated Codes

Kwame Ackah Bohulu

November 15, 2018

1 Abstract

Reed-Solomon convolutional concatenated (RSCC) code is a popular coding scheme whose application can be found in wireless and space communications. However, iterative soft decoding of the concatenated code is yet to be developed. This paper proposes a novel iterative soft decoding algorithm for the concatenated code, aiming to better exploit its errorcorrection potential. The maximum a posteriori (MAP) algorithm is used to decode the inner convolutional code. Its soft output will be deinterleaved and then given to the soft-in-soft-out (SISO) decoder of the outer Reed-Solomon (RS) code. The RS SISO decoder integrates the adaptive belief propagation (ABP) algorithm and the Koetter-Vardy (KV) list decoding algorithm, attempting to find out the transmitted message. It feeds back both the deterministic and the extrinsic probabilities of RS coded bits, enabling the soft information to be exchanged between the inner and outer decoders. An extrinsic information transfer (EXIT) analysis of the proposed algorithm is presented, analyzing its iterative decoding behavior for RSCC codes. The EXIT analysis also leads to the design insight of inner code in the concatenation. Computational complexity of the proposed algorithm is also analyzed. Finally, the iterative decoding performance is shown and its advantage over the existing decoding algorithms is demonstrated.

2 Introduction

Concatenated codes were first introduced by Forney in [1]. It has been shown that concatenating a nonbinary outer code and a binary inner code could constitute a capacity approaching error-correction code with a polynomial-time decoding complexity. A popular case is the Reed-Solomon Convolutional Concatenated(RSCC) code where the Reed Solomon (R-S) code is the outer code and the convolutional code is the inner code. Because the inner code is good at correcting spread bit errors, while the outer code is good at correcting burst errors, this gives the RSCC codes strong error-correction capability and they are used in many wireless and space communication applications [2] - [4]. classic decoding of RSCC codes employs the Viterbi algorithm [5] and the Berlekamp-Massey (BM) algorithm [6] to decode the inner and outer codes respectively and a block interleaver (deinterleaver) is used between the inner and outer encoders (decoders) for the purpose of spreading the burst errors that are resulted from the Viterbi decoding. Over the years, many attempts have been made to improve the decoding of RSCC codes

An improved decoding algorithm that performs repeated decoding trials for RSCC codes was proposed in [7]. It is a primitive attempt to decode the concatenated codes iteratively. However, since the BM algorithm is used to decode the outer code, it is not possible for soft information to be feedback. Another attempt to improve the error-correction performance is to utilize a stronger RS decoding algorithm, e.g., the Guruswami-Sudan (GS) algorithm [8] [9] and Koetter- Vardy (KV) algorithm [10]. Utilizing the KV algorithm in iterative decoding process of [7] was considered in [11]. Meanwhile, utilizing the RS decoding output statistics to form the soft feedback information for iterative decoding the concatenated codes was proposed in [12]. Finally, collaborative decoding of RS codes has also been considered for the concatenated codes [13]. This enables different RS codewords to be decoded jointly, allowing the BM algorithm to correct symbol errors beyond the half distance bound for each RS code.

The study of Turbo codes has revealed that by iteratively exchanging soft information between 2 interleavers a code may be able to achive capacity approaching error correcting performance. Due to the challenge in designing a soft-input-soft-output (SISO) decoder for RS codes a truly iterative decoding algorithm for RSCC codes has yet to be found. Earlier attempts at SISO RS decoding include the maximum likelihood (ML) decoding that utilizes the codes binary image [15] [16]. But its complexity grows exponentially with the length of the code. Recently, SISO decoding of RS codes utilizing the adaptive belief propagation (ABP) algorithm was proposed in [17] [18]. The ABP algorithm enhances the reliability of the soft received information and passes it to the following algebraic decoding, i.e., the BM or KV algorithm. Such an approach was later extended to decode the general algebraic-geometric codes in [19].

The big advantage of using the ABP algorithm is that the extrinsic probabilities of RS coded bits can be calculated based on the adapted Tanner graph of the code with a moderate complexity. As a result, the soft information of RS coded bits can be iterated in a turbo decoding mechanism for RSCC codes.This paper proposes a novel iterative soft decoding algorithm for RSCC codes. The maximum a posteriori (MAP) [20] algorithm is used to decode the inner code, delivering the extrinsic probabilities for the interleaved RS coded bits. They are deinterleaved and mapped to the a priori probabilities of the RS coded bits. The RS SISO decoding has two successive stages. The first stage is the bit reliability oriented ABP algorithm that improves the reliability of the received information. Its output is the extrinsic probabilities and the a posteriori probabilities for the RS coded bits. The a posteriori probabilities will be utilized by the second RS decoding stage, i.e., KV algorithm. If the KV decoding is successful, deterministic probabilities(probability value of 0 or 1) of each RS coded bit will be given as the feedback. Otherwise, extrinsic probabilities that are yielded by the ABP algorithm will be fed back. They are then interleaved and mapped to the a priori probabilities of the interleaved RS coded bits for the next round MAP decoding. An extrinsic information transfer (EXIT) characteristics of the proposed algorithm is analyzed, leading to the insights of its iterative decoding behavior and design criteria of the inner and outer codes. The decoding complexity of the algorithm is also analyzed. Our simulation results show that it can outperform the classical Viterbi-BM algorithm with up to 2dB gain.

3 The RSCC Codes

Let $\mathcal{F}_q = \{\rho_1, \rho_2, ..., \rho_q\}$ denote the finite field of size q. In this paper, it is assumed that $\mathcal{F}_q$ is an extension field of $\mathcal{F}_2$ as $q = 2^\omega$, where ω is a positive integer. Let $\mathcal{F}_q[x]$ and $\mathcal{F}_q[x, y]$ denote the rings of univariate and bivariate polynomials defined over $\mathcal{F}_q$, respectively. The encoder block diagram of RSCC codes is shown by Fig. 1. There is a block interleaver between the inner and outer encoders, which has a vertical read-in and horizontal read-out interleaving pattern. Let D denote the depth of the block interleaver indicating there are D RS codewords being interleaved, and γ denote the index of the RS codeword where $1 \leq \gamma \leq D$. D is set as 10 in the paper unless otherwise specified.

4 MSRS decoding of IRS codes

To explain how the MSRS decoding can be applied to interleaved ReedSolomon codes, we first consider the classical case of a single ReedSolomon code $\mathcal{A} = \mathcal{RS}(q; N, K, D)$. Let $a(x) \in \mathcal{A}$ be some codeword of the RS codeword. We then assume that $a(x)$ is transmitted over a noisy channel, which adds an error polynomial $e(x)$ of degree smaller than N with coefficients from $\mathbb{F}_q$. We can then rewrite $r(x) = a(x) + e(x)$. In the case where $e(x) \notin \mathcal{A}$ some of the coefficients $R_K, ..., R_{N-1}$ will be non-zero, where $R(X) = \mathscr{F}(r(x))$. We may intepret them as syndrome coefficients and denote them by $S_j = R_{K+j}, j = 0, ..., N-K-1$

To decode t errors, the standard approach for algebraic ReedSolomon decoding is to define a polynomial $\lambda(X)$, such that the coefficient λ_j is zero whenever the corresponding coefficient e_j of $e(x)$ is not equal to zero, and non-zero, wheneve$e_j = 0$. Consequently, we have $\lambda_j \cdot e_j = 0 \; \forall j = 0, ..., N-1$.

Due to the properties of the Fourier Transform, this relation is transformed into

$$\Lambda(x) \cdot E(x) \equiv \quad \mod X^N - 1 \tag{1}$$

$\Lambda(x) = \Lambda_0 + \Lambda_1 X + \cdots + \Lambda_t X^t$ is called error locator polynomial. Since the roots of$\Lambda(x)$ are not modified by a multiplicative constant factor we set $\Lambda_0 = 1$ Equation (**??**) forms a linear system of equations, which contains t equations only dependent on the known syndrome coefficients S_t and the unknown coefficients $\Lambda_1, ..., \Lambda_t$. With this t equations we can write the matrix equation

$$\begin{bmatrix} S_0 & S_1 & \dots & S_{t-1} \\ S_1 & S_2 & \dots & S_t \\ & & & \vdots \\ S_{M-t-1} & S_{M-t} & \dots & S_{M-2} \end{bmatrix} \begin{bmatrix} \Lambda_t \\ \Lambda_{t-1} \\ \vdots \\ \Lambda_1 \end{bmatrix} = \begin{bmatrix} -S_t \\ -S_{t+1} \\ \vdots \\ -S_{M-1} \end{bmatrix} \tag{2}$$

$$\mathbf{S\Lambda} = \mathbf{T}$$

Where $M = N-K$ By solving Equation(**??**), we get the error locator polynomial and also the locations of the errors. Since the matrix $\mathbf{S}$ consists of $N-K-t$

rows and t columns, a unique solution never exists if $t > \frac{N-K}{2}$ and we aren't able to correct more than $\frac{\lfloor N-K \rfloor}{2}$.

We can rewrite Equation(**??**) as

$$S_k = -\sum_{j=1}^{t} \Lambda_j S_{k-j}, k = t, ..., N-K-1 \tag{3}$$

Now, the problem of calculating $\Lambda(x)$ is transformed to the problem of finding the connection weights $(\Lambda_1, ..., \Lambda_t)$ for the smallest possible t (or equivalently the shortest shift-register), which recursively generates the syndrome sequence $(S_0, ..., S_{N-K-1})$. This linear recursion synthesis problem is exactly what the BerlekampMassey algorithm is able to solve in a very efficient way.

In the case of an interleaved ReedSolomon code consisting of l codewords. Then we have l received vectors $r^(l) = a^(l)(x) + e^(l)(x)$ at the output of the channel with l different error polynomials $e^(l)(x)$ non-zero coefficients at the same positions. From these received vectors we calculate the syndrome sequences $(S_0^{(l)}, ..., S_{N-K-1}^{(l)}), l = 1, ..., l$, and use them to state the linear system of equations

$$\begin{bmatrix} S^{(1)} \\ S^{(2)} \\ \\ \vdots \\ S^{(l)} \end{bmatrix} \begin{bmatrix} \Lambda_t \\ \Lambda_{t-1} \\ \vdots \\ \Lambda_1 \end{bmatrix} = \begin{bmatrix} -T^{(1)} \\ -T^{(2)} \\ \\ \vdots \\ -T^{(l)} \end{bmatrix} \tag{4}$$
$$\mathbf{S}_l \boldsymbol{\Lambda} = \mathbf{T}_l$$

The linear system of Equations (**??**) consists of t unknowns and $l \cdot (N - K - t)$ equations. Hence, it may have a unique solution, provided that $t \leq \left\lfloor \frac{l}{l+1}(N-K) \right\rfloor$.

we can state Equation (**??**) in the form of the linear recursions

$$S_k^{(l)} = -\sum_{j=1}^{t} \Lambda_j S_{k-j}^{(l)}, k = t, ..., N-K-1, (l) = 1, ...l \tag{5}$$

However, we now have to find the connection weights $(\Lambda_1, ..., \Lambda_t)$ for the minimum t, which simultaneously generate the l different syndrome sequences $(S_0^{(l)}, ..., S_{N-K-1}^{(l)}), l = 1, ..., l$.

Algorithm 1 gives a description of the MSRS algorithm in pseudo code, suitable for locating errors in an interleaved RS code. For $l = 1$ Algorithm 1 reduces to the classical Berlekamp-Massey algorithm. For $l > 1$ the inner for-loop is repeated l times. Hence, the complexity of Algorithm 1 is approximately l times the complexity of the Berlekamp-Massey algorithm.

After calculating $\Lambda(x)$, we know where the errors are located. However, to complete the decoding, we still have to evaluate the error values. This can be done in the time domain by calculating the coefficients of the polynomials e^l with the well known Forney algorithm or whatever means that seems to be more suitable for a specific application.

Note that if the degree of $\Lambda(X) > \left\lfloor \frac{l}{l+1}(N-K) \right\rfloor$ after applying Algorithm 1, the linear system of equations (**??**) cannot have a unique solution and results in a decoding failure. However, if $\left\lfloor \frac{l}{l+1}(N-K) \right\rfloor < deg(\Lambda(X)) \leq \left\lfloor \frac{l}{l+1}(N-K) \right\rfloor$ Equation (7) will only have a unique solution, if rank $(\mathbf{S}_l) = t$ and if rank $(\mathbf{S}_l) < t$,the MSRS algorithm cannot detect that there does not exist a unique solution and will just create an error locator polynomial out of several possible solutions. This can be seen as a disadvantage of the MSRS algorithm .

5 Bounds on Concatenated Decoding Performance

Even though it is possible to decode beyond half the minimum distance of the RS code using IRS codes, the probabilistic nature of the decoder makes it difficult to guarantee that all error patterns of weight t can be corrected even if t is in the range $\lfloor \frac{N=K}{2} \rfloor < t \leq \left\lfloor \frac{l}{l+1}(N-K) \right\rfloor$ This is because some error patters create syndromes with rank less than t and therefore such error patterns can be decoded with some probability $P_f(t) < 1$

To analyze the probability $P_f(t)$ we assume that for any position of the interleaved ReedSolomon code any non-zero error pattern $\mathbf{e}_j = (e_j^1, ..., e_j^l)$ occurs with the same probability. Under this assumption, upper bounds for P_f have been derived in [2] and [10]. However since these bounds are not very tight or only tight for the maximum error correction radius $t = \left\lfloor \frac{l}{l+1}(N-K) \right\rfloor$ they are generally not suited very well for estimating the decoding performance of our concatenated codes. In [1], the probability $P_f(t)$ has been upper bounded dependent on t for the case of folded ReedSolomon codes. Using similar but simpler techniques of proving, we are able to derive the following theorem for interleaved RS codes.

Theorem 1 Let a codeword $(C) \in C$ be corrupted by some errors $e^{(l)}(x) = e_0^{(l)} + e_1^{(l)}x + \cdots + e_{N-1}^{(l)}x^{N-1}$ and let $\mathbf{E} = (\mathbf{e}_0, ..., \mathbf{e}_{N-1}), e_j = (e_j^{(1)}, ..., e_j^{(l)})^T \neq 0$ be an $(l \times n)$ matrix with t non-zero columns, whereas $\lfloor \frac{N-K}{2} \rfloor < t \leq \left\lfloor \frac{l}{l+1}(N-K) \right\rfloor$. Further assume that all non-zero error patterns $\mathbf{e}_j \neq 0$ occur equiprobable. Then, the probability for a decoding failure is upper bounded by

$$P_f(t) \leq (\frac{q^l - \frac{l}{q}}{q^l - 1}) \cdot \frac{q^{-\delta(t)}}{q-1}$$

where $\delta(t) = l \cdot (N-K-t) - t$ is the difference between the number of equations and the number of unknowns in Equation(7)

Since the proof for this theorem is too complex it was left out of the research paper.

Next, we assume that the concatenated code C is transmitted ove the AWGN channel with BPSK modulation and transmitted element-wise. At the output we observe the matrix $\mathbf{Y} = (y_0^T, ..., y_{N-1}^T) = (y_{i,j} = x_{i,j} + w_{i,j})$ The variables$w_{i,j})$ are statistically independent Gaussian random variables with zero mean and variance $\sigma^2 = N_0/2$. Next ML decoding is done on the columns of $\mathbf{Y}$. The word error probability p_w at the output of this inner decoder will be the input symbol error probability for the interleaved ReedSolomon code, i.e., the probability that a column of $\mathbf{R}$ will be erroneous. The probability p_w can be overbounded by a Union Bound, if the weight distribution of the code $\mathcal{B}$ is known. Let m_w be the number of codewords in $\mathcal{B}$, which have Hamming weight w. Then p_w can be overbounded by

$$p_w \leq \frac{1}{2} \sum_{w=d}^{n} m_w \cdot erfc\Big(\sqrt{\frac{w}{2\sigma^2}}\Big) = \hat{p}_w \tag{6}$$

If we decode each row of $\mathbf{R}$ independently by a standard ReedSolomon decoder we know, that we can correct up to $\lfloor \frac{N-K}{2} \rfloor$ errors in the l words. Therefore, the word error probability p_w after outer decoding is

$$P_W^{RS} = \sum_{\lfloor \frac{N-K}{2}+1 \rfloor}^{N} \binom{N}{i} \cdot p_w^i \cdot (1-p_w)^{N-i} \tag{7}$$

To overbound this expression, we define the function$f(p_w) = (p_w^i \cdot (1 - p_w)^{N-1})$ and calculate the derivative. From this we see that $f(p_w)$ is a monotonically nondecreasing function, if $p_w \leq \frac{i}{N}$. Therefore, we can overbound P_W^{RS} by

$$P_W^{RS} = \sum_{\lfloor \frac{N-K}{2}+! \rfloor}^{N} \binom{N}{i} \cdot \hat{p}_w^i \cdot (1-\hat{p}_w)^{(N-i)\cdot\epsilon(i)} \tag{8}$$

where $\epsilon(i) = 1$ if $i \geq \hat{p}_w \cdot N$ and $\epsilon(i) = 0$ otherwise.

If we use an interleaved ReedSolomon decoder to decode $\mathbf{R}$ we are able to correct $\lfloor \frac{N-K}{2} \rfloor < t \leq \left\lfloor \frac{l}{l+1}(N-K) \right\rfloor$ errors with the probability $P_f(t)$. Theorem 1 gives us an upper bound on $P_f(t)$ under the assumption of equiprobable distributed errors. Unfortunately, this assumption is anything but fulfilled after decoding the inner code. Due to the characteristics of ML decoding, low-weight error patterns will occur more frequently than high-weight patterns. To be able to apply Theorem 1 to our concatenated design, we have to modify it slightly to randomize the error patterns after the inner decoding. For this purpose, let $\mathcal{N}_1$ be the set of all $l \times l$matrices with elements from the field $\mathbb{F}_q$ and let $\mathcal{M}_l \subset \mathcal{N}_l$ be the subset of all nonsingular matrices. Now we modify the encoding rule given by Equation (1) into

$$C = \Big(g(\mathbf{M}_0\mathbf{a}_0), ..., g(\mathbf{M}_{N-1}\mathbf{a}_{N-1})\Big) = \Big(\mathbf{b}_0^T, ..., \mathbf{b}_{N-1}^T\Big)$$

whereas the matrices $\mathbf{M}_i$ are statistically independent random matrices, uniformly distributed in M_l The reverse mapping after inner decoding described by Equation (3) is modified to

$$\mathbf{R} = \Big(\mathbf{M}_0^{-1} \cdot g^{-1}(\hat{\mathcal{B}}_0^T), ..., \mathbf{M}_{N-1} \cdot g^{-1}(\hat{\mathcal{B}}_{N-1}^T)\Big)$$

By this randomization procedure, we do not influence the correct columns of $\mathbf{R}$, but we ensure that the error patterns after inner decoding are transformed to uniformly distributed error patterns. Since the number of erroneous columns is not changed by randomization, we do not expect a negative impact on the concatenated decoding performance. After randomizing, we can utilize Theorem 1 to derive an upper bound on the word error probability P_w^{IRS} w after outer interleaved RS decoding. To do this, we use the same technique as for Equation (10), but weight the i-th summand by the factor

$$\hat{P}_f(i) = \begin{cases} (\frac{q^l - \frac{l}{q}}{q^l - 1}) \cdot \frac{q^{-\delta(t)}}{q-1}, & i \leq \left\lfloor \frac{l}{l+1}(N-K) \right\rfloor \\ 1, & \text{otherwise} \end{cases}$$

obtained from Theorem 1. In this way, we overbound P_W^{IRS} by

$$P_W^{RS} = \sum_{\lfloor \frac{N-K}{2} + 1 \rfloor}^{N} \binom{N}{i} \cdot \hat{P}_f(i) \cdot \hat{p}_w^i \cdot (1 - \hat{p}_w)^{(N-i) \cdot \epsilon(i)} \tag{9}$$