Parameter identification of Reed-Solomon codes over noisy environment

Swaminathan, Ramabadrar; Madhukumar, A. S.; Wang, Guohua; Ting, Shang Kee

2017

Swaminathan, R., Madhukumar, A. S., Wang, G., & Ting, S. K. (2017). Parameter identification of Reed-Solomon codes over noisy environment. Proceedings of the 2017 IEEE 86th Vehicular Technology Conference (VTC-Fall), 1-5.
doi:10.1109/VTCFall.2017.8287916

<https://hdl.handle.net/10356/144740>

<https://doi.org/10.1109/VTCFall.2017.8287916>

© 2017 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The published version is available at:
<https://doi.org/10.1109/VTCFall.2017.8287916>

Downloaded on 26 Apr 2023 19:52:43 SGT

Parameter Identification of Reed-Solomon Codes over Noisy Environment

Swaminathan R[†], A.S.MadhuKumar[†], Wang Guohua*, and Ting Shang Kee*

[†]School of Computer Science and Engineering, *Temasek Laboratories

Nanyang Technological University, Singapore

Email: {sramabadran, asmadhukumar, guowang, skting}@ntu.edu.sg

Abstract—Blind reconstruction of code parameters plays a vital role in non-cooperative applications. It also provides additional advantages in applications such as adaptive modulation and coding, reconfigurable receivers, etc. In this paper, blind estimation algorithm is proposed to identify Reed-Solomon (RS) code parameters at the receiver over noisy channels. Simulation results validating the proposed algorithm are given for various test cases. Further, the accuracy of estimation of RS code parameters using the proposed algorithm is also evaluated for various M -ary quadrature amplitude modulation schemes. From the simulation results, it is observed that the accuracy of estimation improves with decrease in code dimension and modulation order and the proposed algorithm outperforms other algorithms in the prior works.

Index Terms—Adaptive modulation and coding, blind estimation, non-cooperative applications, Reed-Solomon codes

I. INTRODUCTION

Forward error correcting (FEC) codes are mainly used to improve the error performance of communication and storage systems by counteracting the random errors introduced by the noisy channel with the help of redundant bits. Reed-Solomon (RS) codes, which is a special class of non-binary Bose-Chaudhuri-Hocquenghem (BCH) codes, are widely used in data storage, digital video broadcasting, and deep-space systems. In most of the applications, FEC code parameters are known at the receiver to decode the received message symbols. However, in a non-cooperative applications [1], it is mandatory to estimate the FEC code parameters. This is because, only limited knowledge about the FEC codes used for transmission might be available at the receiver. Further, the blind estimation of code parameters provide additional advantages in applications such as adaptive modulation and coding (AMC), reconfigurable receivers, etc. In AMC-based systems, the code parameters are transmitted via control channel to the receiver. Therefore, the blind identification of code parameters will lead to conservation of channel resources and improve the spectral efficiency of the AMC-based systems. It is a known fact that the reconfigurable receiver systems need to adapt itself to the variations in the channel coding scheme used for transmission. Thus, the blind recognition of code parameters finds potential applications in modern reconfigurable radio systems. Further, designing dedicated receivers for each applications is a costly and tedious process. Therefore, blind estimation of code parameters is mandatory

for any future communication and storage systems.

Code classification and code parameter estimation techniques for different FEC codes over non-erroneous scenario were studied in [1] with greater details. The blind recognition of convolutional code parameters over noisy transmission environment was reported in [2]. Further, the blind recognition of codeword length for different non-binary error correcting codes was extensively investigated over noisy environment in [3]. In [4], the algorithm for blind estimation of RS code parameters was reported by adding symbol errors over RS encoded data without using any modulation schemes. The algorithm was based on Barbier's dual code method developed for binary codes in [5] and the same was generalized in [4] to non-binary codes. In [6], [7], and [8], the novel code parameter recognition technique based on the average log-likelihood ratio (LLR) of syndrome a posteriori probability (SPP) was proposed for RS codes, low-density parity-check (LDPC) codes, and convolutional codes, respectively. Finally, using iterative expectation-maximization and least square methods, the parameters of turbo convolutional encoders were estimated in [9] and [10], respectively.

A. Motivations

The main motivations behind the proposed work are given as follows:

- It is always necessary to estimate the code parameters from the transmitted bit stream in non-cooperative applications such as military and spectrum surveillance. In most of such scenarios, some partial information about the code (e.g. type of the code) is known at the receiver.
- In some of the prior works, the parameter estimation algorithm proposed for RS codes can estimate only codeword length and the same cannot estimate generator polynomial and code dimension, which are mandatory for decoding RS coded symbols. In addition, the algorithm for blind estimation of RS code parameters was reported by adding symbol errors over RS coded data without using any modulation schemes, which is not accurate and realistic. Because, in a practical digital communication systems, the FEC coded data will be modulated and then only sent for transmission.
- In the proposed LLR-based methodology [6] to identify the true RS encoder, a predefined set of encoders was

assumed to be known at the transmitter and receiver. Hence, the proposed method was not strictly blind and the delay of the received modulated symbols to achieve synchronization was not recovered.

- Blind recognition algorithm for RS codes considering M -ary quadrature amplitude modulation (M -QAM) scheme over noisy transmission environment has not been reported in any of the prior works to the best of our knowledge.

B. Contributions

The main contributions of the proposed work are given as follows:

- In this paper, we propose innovative algorithms for the blind recognition of RS encoder over noiseless and noisy channels considering M -QAM scheme, where M denotes the modulation order. The estimated code parameters are codeword length n , code dimension k , number of bits per symbol m , primitive polynomial p , and generator polynomial $g(x)$.
- The present paper proposes an innovative approach for synchronization compensation through appropriate bit positioning and is integrated with the proposed code parameter estimation algorithms.
- Simulation results are given for different test cases validating the proposed algorithms.
- The performance of the algorithm in terms of accuracy of estimation is compared with the performance of the Barbier's method proposed for non-binary codes in [4] and the LLR-based technique proposed in [6].
- In order to show the robustness of the proposed algorithm, the accuracy of parameter estimation of RS codes by varying the signal-to-noise ratio (SNR) values is given for different values of M and k .

II. REED-SOLOMON CODES AND PARAMETER ESTIMATION PROCESS

The parameters of t error correcting RS code with symbols from Galois field $GF(q)$, where $q=2^m$ and $m \geq 3$, are given below [11].

- Codeword length $n = q - 1$
- Number of parity check symbols $n - k = 2t$
- Code dimension $k = q - 1 - 2t$
- Minimum hamming distance $d_{\min} = 2t + 1$

Firstly, we denote the RS encoded data symbols as $RS(n, k, m, p)$. Let α denotes the primitive element of $GF(q)$ such that $\alpha^{q-1} = 1$. It is a known fact that $\alpha, \alpha^2, \dots, \alpha^{2t}$ are the roots of the generator polynomial $g(x)$. Hence, $g(x)$ can be written as

$$\begin{aligned} g(x) &= (x - \alpha)(x - \alpha^2) \cdots (x - \alpha^{2t}) \\ &= g_0 + g_1 x + g_2 x^2 + \cdots + g_{2t-1} x^{2t-1} + x^{2t} \end{aligned} \quad (1)$$

The block diagram of the parameter estimation process of RS codes is shown in Fig. 1 (see next page). Firstly, the RS encoded data symbols are converted into bit streams

and modulated using M -QAM scheme for transmission. The modulated symbols are transmitted over additive white Gaussian noise (AWGN) channel. At the receiving end, the RS code parameters are estimated from the binary coded data symbols after demodulation. Once all the parameters are recognized, the same can be used to decode the incoming coded symbols.

III. PARAMETER ESTIMATION OVER NON-ERRONEOUS SCENARIO

After demodulation, the RS code parameters and number of bits to be shifted for synchronization are estimated using Algorithm 1 from the binary coded data symbols.

Algorithm 1: Estimation of RS code parameters (non-erroneous or noiseless scenario)

Notations: Let ϕ denotes the adjustment of bit position to achieve synchronization. $a, b, \rho(m, p, \phi)$, and $\rho'(m, p, \phi)$ denote the number of rows, columns, rank, and rank ratio of data matrix S , respectively.

Assumptions: $a \geq 2b$, $m \in [m_{\min}, m_{\max}]$, $p \in [p_{\min}, p_{\max}]$, and $\phi \in [0, ((q-1)m) - 1]$

```

1: for  $m = m_{\min} : m_{\max}$  do
2:    $p = \text{primpoly}(m, 'all')$ , where  $\text{primpoly}(\cdot, 'all')$  represents all primitive polynomials over  $GF(q)$ 
3:   for  $p = p_{\min} : p_{\max}$  do
4:     for  $\phi = 0 : ((q-1)m) - 1$  do
5:       Shift the binary data symbols by  $\phi$  bit positions and convert the same into non-binary symbols such that each symbol is an integer between 0 and  $q-1$ , where  $q = 2^m$ ;
6:       Create a GF array from the non-binary data symbols using primitive polynomial  $p$ ;
7:       Reshape the RS encoded GF array elements into a data matrix  $S$  of size  $a \times b$ , where  $b = q - 1$ ;
8:       Use finite-field Gauss elimination process [12] in  $GF(q)$  to convert  $S$  into column echelon form  $F$ ;
9:       Compute  $\rho(m, p, \phi)$  from the number of non-zero columns in  $F$ ;
10:      Compute  $\rho'(m, p, \phi) = \rho(m, p, \phi) / b$ ;
11:    end
  end
  end
  Obtain  $[m_{\text{est}}, p_{\text{est}}, \phi_{\text{est}}] = \underset{m, p, \phi}{\text{argmin}}(\rho'(m, p, \phi))$ ,
   $n_{\text{est}} = 2^{m_{\text{est}}} - 1$ , and  $k_{\text{est}} = \rho(m_{\text{est}}, p_{\text{est}}, \phi_{\text{est}})$ ;

```

Note that the number of primitive polynomials will vary according to the value of m . For correct values of m, p , and ϕ , rank deficiency will be observed in step 11 of Algorithm 1. However, full rank will be obtained for incorrect values of m, p , and ϕ . The reason for rank deficiency is explained using a proposition given below. It is to be observed that b also varies while varying m , since $b = 2^m - 1$.

Fig. 1. Block diagram of parameter estimation process of RS codes

Proposition:

Let S be a data matrix of RS coded symbols with b columns. Assuming m , p , and ϕ are estimated correctly, if $b = n_{\text{est}} = 2^{m_{\text{est}}} - 1$, then rank deficiency will be obtained. Now $\rho(m_{\text{est}}, p_{\text{est}}, \phi_{\text{est}})$ and $\rho'(m_{\text{est}}, p_{\text{est}}, \phi_{\text{est}})$ for $b = n_{\text{est}}$ are, respectively, given by

$$\begin{aligned}\rho(m_{\text{est}}, p_{\text{est}}, \phi_{\text{est}}) &= k_{\text{est}}, \\ \rho'(m_{\text{est}}, p_{\text{est}}, \phi_{\text{est}}) &= \frac{k_{\text{est}}}{b} = r,\end{aligned}\quad (2)$$

where r is the code rate of RS codes. Similarly, if b is a multiple of n_{est} (i.e. $b = \alpha n_{\text{est}}$), then the deficient rank value is given by αk_{est} . However, if $b \neq \alpha n_{\text{est}}$, then full rank will be obtained.

Explanation:

The output n data symbols depend only on k input symbols in the case of RS codes. Therefore, αn_{est} output symbols of S depend on αk_{est} input symbols. Further, when $b = \alpha n_{\text{est}}$, then α codewords in all the rows will be aligned properly in the same column. Hence, after converting S into F through finite field Gauss elimination process [12], all $\alpha(n_{\text{est}} - k_{\text{est}})$ dependent columns will be eliminated, which will give rise to rank deficiency. If $b \neq \alpha n_{\text{est}}$, then α codewords will not be properly aligned in the same column across all the rows, which will lead to full rank. Note that (2) is a special case for which $\alpha = 1$.

IV. PARAMETER ESTIMATION OVER ERRONEOUS SCENARIO

The dependent columns in S will be converted into all-zero-columns in F using finite-field Gauss elimination process and the number of non-zero columns gives the deficient rank of S in the case of non-erroneous scenario. The Algorithm 1 proposed for non-erroneous scenario fails for erroneous scenario, since full rank will be obtained irrespective of correct or incorrect values of m , p , and ϕ due to erroneous symbols. It has been observed that the data matrix S with rank deficiency under erroneous channel conditions will have less number of non-zero elements compared to the full-rank matrix, which will have zero and non-zero elements with equally-likely probability. Therefore, the rank-deficient data matrix is identified based on evaluating the non-zero-mean-ratio of F in the case of erroneous scenario. The detailed steps for estimating the RS code parameters and bit position adjustment to achieve synchronization at the receiver for erroneous scenario are given using Algorithm 2 (refer to next page).

As already mentioned for non-erroneous case, the rank deficiency will be observed for correct values of m , p , and ϕ . Since rank-deficient data matrix will have less number of non-zero elements, $\mu'(m_{\text{est}}, p_{\text{est}}, \phi_{\text{est}})$ will be smaller compared to other incorrect possible combinations of $[m, p, \phi]$ as shown in step 11 of Algorithm 2. It is to be noted that the code and

Algorithm 2: Estimation of RS code parameters (erroneous or noisy scenario)

Notations: The mean and normalized mean values of number of non-zero elements in c^{th} column of F are denoted as $\sigma(c, m, p, \phi)$ and $\sigma'(c, m, p, \phi)$, respectively.

Assumptions: $a \geq t b$, where t is a constant, $m \in [m_{\min}, m_{\max}]$, $p \in [p_{\min}, p_{\max}]$, $\phi \in [0, ((q-1)m) - 1]$, and the erroneous incoming bit stream is assumed to be RS encoded.

```

1: for  $m = m_{\min} : m_{\max}$  do
2:    $p = \text{primpoly}(m, 'all')$ ;
3:   for  $p = p_{\min} : p_{\max}$  do
4:     for  $\phi = 0 : ((q-1)m) - 1$  do
5:       Shift the erroneous incoming binary data
       symbols by  $\phi$  bit positions and convert the
       same into non-binary symbols, where each
       symbol is an integer between 0 and  $q-1$ ;
6:       GF array is created using primitive
       polynomial  $p$  from the non-binary data
       symbols;
7:       RS encoded GF array elements are
       reshaped into a data matrix  $S$  of size  $a \times b$ ,
       where  $b = q-1$ ;
8:       Using finite-field Gauss elimination
       process,  $S$  is converted into  $F$ ;
9:        $\sigma(c, m, p, \phi)$ , where  $c \in \{1, 2, \dots, b\}$ , is
       evaluated and the obtained values are
       normalized with respect to maximum value;
10:      Calculate the normalized
       non-zero-mean-ratio, which is given by
        $\mu'(m, p, \phi) = \frac{\sum_{c=1}^b \sigma'(c, m, p, \phi)}{b}$ ;
11:    end
  end
end
11: Obtain  $[m_{\text{est}}, p_{\text{est}}, \phi_{\text{est}}] = \underset{m, p, \phi}{\text{argmin}}(\mu'(m, p, \phi))$  and
 $n_{\text{est}} = 2^{m_{\text{est}}} - 1$ ;
```

generator polynomials of RS codes have equal number of roots and is given by $n - k = 2t$. Hence, by finding the number of roots of code polynomial, $n_{\text{est}} - k_{\text{est}}$ is identified. As n_{est} is already estimated from step 11, k_{est} can be recognized from $n_{\text{est}} - k_{\text{est}}$. After recognizing the number of roots of code polynomial, $g(x)$ is obtained from (1).

V. SIMULATION RESULTS AND DISCUSSIONS

In our simulation study, we assume $a = 2b$ for non-erroneous case and $t = 5000$ for erroneous case. It is to be noted that higher the number of rows or the value of t , better is the accuracy for erroneous case. Moreover, if the receiver starts receiving the modulated data symbols at Δ^{th} position of t^{th} RS code word, then frame synchronization is achieved after

Fig. 2. Variation of $\rho'(m, p, \phi)$ with respect to $[m, p, \phi]$ for RS(63, 45, 6, 67) assuming 64-QAM scheme, $\Delta = 6$, and non-erroneous scenario

Fig. 3. Variation of $\mu'(m, p, \phi)$ with respect to $[m, p, \phi]$ for RS(31, 15, 5, 37) assuming 4-QAM scheme, $\Delta = 7$, and SNR = 10 dB

demodulation by shifting $\phi = (tm(q-1)+1) - (\Delta M_1 - M_1 + 1)$ bit positions, where $M_1 = \log_2 M$.

In Fig. 2, the variation of $\rho'(m, p, \phi)$ is shown for all possible combinations of $[m, p, \phi]$, where $m \in [4, 6]$, assuming RS(63, 45, 6, 67), 64-QAM scheme (i.e. $M = 64$ and $M_1 = 6$), $\Delta = 6$, $t = 1$, and non-erroneous scenario. From the curves, it is observed that for $[m, p, \phi] = [6, 67, 348]$, $\rho'(m, p, \phi) = 0.714$, which is minimum among all other possible combinations of $[m, p, \phi]$. Hence, RS code parameters $m_{\text{est}} = 6$, $n_{\text{est}} = 63$, $p_{\text{est}} = 67$ and $k_{\text{est}} = \rho(m, p, \phi) = 45$ are estimated successfully using Algorithm 1. Further, the adjustment of bit position to achieve synchronization (i.e. $\phi_{\text{est}} = 348$) is also estimated correctly using Algorithm 1.

The variation of $\mu'(m, p, \phi)$ with respect to $[m, p, \phi]$, where $m \in [3, 5]$, is shown in Fig. 3 assuming RS(31, 15, 5, 37), 4-QAM scheme (i.e. $M = 4$ and $M_1 = 2$), $\Delta = 7$, $t = 1$, and SNR = 10 dB. From the plot, it is inferred that $\mu'(m, p, \phi)$ reaches minimum at $[m, p, \phi] = [5, 37, 143]$. Therefore, the RS code parameters $m_{\text{est}} = 5$, $n_{\text{est}} = 31$, and $p_{\text{est}} = 37$ are estimated successfully using Algorithm 2. Further, the number of roots of code polynomial is also estimated and is given by $n_{\text{est}} - k_{\text{est}} = 16$. The code dimension $k_{\text{est}} = 15$ is recognized correctly from $n_{\text{est}} - k_{\text{est}}$. Finally, the adjustment of bit position to achieve synchronization (i.e. $\phi_{\text{est}} = 143$) is also identified correctly using Algorithm 2.

In Fig. 4, the accuracy of estimation of RS codes considering RS(15, 9, 4, 19) is given for M -QAM scheme considering different values of M . We notice from the figure that as M increases, the estimation accuracy deteriorates, as expected. In Fig. 5, the accuracy of estimation of RS codes assuming $n = 15$, $m = 4$, and $p = 19$ with 16-QAM constellation is given for different values of code dimension. From the plots, it is observed that the accuracy of estimation deteriorates with increase in the value of k or r . This is because, it is always difficult to classify rank-deficient and full-rank data matrices based on $\mu'(m, p, \phi)$ as r increases due to less number of dependent columns (i.e. $n - k$).

In Fig. 6, the performance of the proposed algorithm is compared with the performance of the LLR-based methodology reported in [6]. The code parameter recognition technique proposed in [6] was based on average LLR of SPP. Here, we have compared the accuracy of estimation of two RS codes namely RS(15, 7, 4, 19) and RS(31, 19, 5, 37) assuming 16-QAM and 32-QAM schemes, respectively. From the performance curves,

Fig. 4. Accuracy of estimation of RS codes RS(15, 9, 4, 19) for different M -QAM schemes

Fig. 5. Accuracy of estimation of RS codes assuming $n = 15$, $m = 4$, $p = 19$, and 16-QAM scheme for different values of code dimension k

Fig. 6. Performance comparison of the proposed algorithm with the algorithm proposed in [6]

TABLE I
COMPARISON OF PROBABILITY OF CORRECT DETECTION OF RS CODE RS(15, 7, 4, 19)

SER	Probability of detection ([4])	Probability of detection (proposed algorithm)
0.001	1	1
0.01	1	1
0.05	1	1
0.075	0.52	0.92
0.1	0	0.40

it is inferred that the proposed algorithm outperforms the LLR-based methodology. For example, the algorithm proposed in [6] works with 100% and 40% accuracy for RS(15, 7, 4, 19) when SNR ≥ 20 dB and SNR = 15 dB, respectively. However, our algorithm can estimate RS code parameters with 100% accuracy for SNR ≥ 15 dB.

In Table I, the probability of correct detection of RS code RS(15, 7, 4, 19) using the proposed algorithm is compared with Barbier's dual-code method proposed in [4] for different symbol error rate (SER) values. The Barbier's method is based on identifying potential dual-codewords and finding the code parameters which maximize rank of a matrix whose rows are constituted of all dual-codewords. In our proposed algorithm, RS code parameters are identified based on rank deficiency and normalized non-zero-mean-ratio values. From the tabulated values, improvement in the probability of detection is observed using the proposed algorithm compared to the algorithm proposed in [4] for SER = 0.075 and SER = 0.1.

VI. CONCLUSIONS

In this paper, blind estimation algorithms have been proposed for estimating RS code parameters based on rank deficiency and normalized non-zero-mean-ratio values for noiseless and noisy environments, respectively. The bit positioning adjustment is also integrated with the the proposed code parameter estimation algorithms to achieve synchronization. The simulation studies show that the proposed algorithms can successfully estimate RS code parameters for various test cases. Further, the accuracy of estimation plots are shown for different M -QAM schemes and code dimension values. It has been inferred that the accuracy of parameter estimation improves with decrease in code dimension values. Further, the lower modulation order schemes perform better than the higher modulation order schemes. It has also been noted that the proposed algorithm for noisy environment consistently outperforms the algorithms proposed in the prior works.

REFERENCES

- [1] J. F. Ziegler, "Automatic Recognition and Classification of Forward Error Correcting Codes," M.S. thesis, The Pennsylvania State University, 2000.
- [2] M. Marazin, R. Gautier, and G. Burel, "Blind recovery of k/n rate convolutional encoders in a noisy environment," *EURASIP J. Wirel. Commun. and Netw.*, vol. 2011:168, pp. 1–9, 2011.

- [3] Y. Zrelli, M. Marazin, R. Gautier, E. Rannou, and E. Radoi, "Blind identification of code word length for non-binary error-correcting codes in noisy transmission," *EURASIP J. Wirel. Commun. Netw.*, vol. 2015:43, pp. 1–16, 2015.
- [4] A. Zahedi and G-R. Mohammad-Khani, "Reconstruction of a non-binary block code from an intercepted sequence with application to Reed-Solomon codes," *IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences*, VOL.E95-A, no. 11, pp. 1873–1880, Nov. 2012.
- [5] G. Sicot, S. Houcke, and J. Barbier, "Blind detection of interleaver parameters," *Signal Process.*, vol. 89, pp. 450–462, April 2009.
- [6] H. Zhang, H-C. Wu, and H. Jiang, "Novel blind encoder identification of Reed-Solomon codes with low computational complexity," in *proc. IEEE GLOBECOM*, 2013, pp. 3294–3299.
- [7] T. Xia and H.-C. Wu, "Novel blind identification of LDPC codes using average LLR of syndrome a posteriori probability," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 632–640, Feb. 2014.
- [8] R. Moosavi and E. G. Larsson, "Fast blind recognition of channel codes," *IEEE Trans. Commun.*, vol. 62, no. 5, pp. 1393–1405, May 2014.
- [9] Y. G. Debessu, H.-C. Wu, and H. Jiang, "Novel Blind Encoder Parameter Estimation for Turbo Codes," *IEEE Commun. Lett.*, vol. 16, no. 12, pp. 1917–1920, Dec. 2012.
- [10] P. Yu, J. Li, and H. Peng, "A least square method for parameter estimation of RSC sub-codes of turbo codes," *IEEE Commun. Lett.*, vol. 18, no. 4, pp. 644–647, Apr. 2014.
- [11] S. Lin and D. J. Costello, "Error Control Coding," 2nd ed. New Jersey, USA: Prentice hall publication, 2004.
- [12] L. Lu, K. H. Li, and Y. L. Guan, "Blind detection of interleaver parameters for non-binary coded data streams," in *Proc. IEEE ICC*, 2009, pp. 1–4.