IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received November 2, 2020, accepted November 27, 2020, date of publication December 1, 2020,
date of current version December 15, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3041795

Blind Interleaver Parameter Estimation From Scant Data

MINGYU JANG, GEUNBAE KIM, DONGYEONG KIM, AND DONGWEON YOON
Department of Electronic Engineering, Hanyang University, Seoul 04763, South Korea

Corresponding author: Dongweon Yoon (dwyoon@hanyang.ac.kr)

This work was supported by the Research Fund of the Signal Intelligence Research Center supervised by the Defense Acquisition Program Administration and Agency for Defense Development of Korea.

ABSTRACT A method for blind estimation of interleaver parameter was recently reported which made additional data from a limited amount of received data. However, the process of making additional data creates undesirable linearity which degrades estimation performance. Promise for improved estimation therefore lies in enhancing blind estimation of interleaver parameter without making additional data. In this paper, we propose an improved method to blindly estimate interleaver parameter under the condition of scant data. We first generate a matrix by using the received data. From this matrix we then make square submatrices and obtain their rank deficiency distribution. Finally, we estimate the interleaver parameter by comparing the rank deficiency distribution of the square submatrices and that of random binary matrices. Through computer simulations, we validate the proposed method in terms of detection probability and the number of false alarms. Simulation results show that the proposed method works better than the conventional method given scarce received data.

INDEX TERMS Blind detection, non-cooperative context, remote sensing, spectrum surveillance.

I. INTRODUCTION

In non-cooperative contexts, a receiver cannot identify any information from the received data because the receiver lacks all information about transmission parameters. To recover information in a non-cooperative context, the receiver has to blindly estimate communication parameters by using only the received data. It is an extensive work requiring much effort to estimate even a single communication parameter. Therefore, blind estimation of communication parameters has been researched separately: source coding [1]–[5], channel coding [6]–[11], interleaving [12]–[23], spreading sequence [24]–[26], scrambling [27]–[29], and modulation [30]–[38]. In this paper, we focus on the blind estimation of interleaver parameter.

Interleaving along with channel coding is essential to establish reliable communication performance because it enables transmitted signals to better withstand the effects of various channel impairments [39]–[43]. Blind estimation of interleaver parameter has been researched extensively [12]–[23]. Reference [12] estimated block interleaver parameter using the rank of matrix composed of received data for noiseless channels and [13] extended the estimation method to noisy channels by using Gauss-Jordan elimination through pivoting. Further, [14]–[19] estimated various types of interleavers by using the number of ones or zeros in each row (or column). Reference [14] estimated helical scan interleaver for block channel coded data. Regarding convolutional channel coded data, [15] estimated block and helical scan interleavers, and [16] estimated convolutional and helical interleavers. Further, [17] jointly recognized the type of channel codes and interleaver parameters, and [18] jointly estimated Reed-Solomon code and block interleaver parameters. Recently, [19] proposed a blind estimation method of a convolutional interleaver with denoising algorithm.

Instead of using the number of ones or zeros in each row (or column), [20]–[22] proposed improved blind estimation methods based on the rank deficiencies of the square matrices generated from the received data: [20] estimated interleaver parameter by choosing vectors having fewer errors, [21] used binomial distribution to compare rank deficiency distributions, and [22] proposed an improved estimation method which can be effectively applied to more severe channel conditions. The estimation methods in [12]–[22] assumed

The associate editor coordinating the review of this manuscript and approving it for publication was Marco Martalo.

a volume of received data sufficient for the blind estimation of interleaver parameter. Given scant received data, the conventional methods become infeasible or have degraded estimation performance [23].

Very recently, [23] proposed a novel method for blind estimation of interleaver parameter when only a limited amount of received data is available. The method in [23] first made additional data by combining the received data. Then it estimated the interleaver parameter by using the average rank deficiency of square matrices composed of the received data and the created additional data. However, in the process of making additional data, a linearity arises which impairs estimation performance. Therefore, improved blind estimation will presumably result from a method that does not produce additional data from scant received data.

In this paper, we propose an improved method to blindly estimate interleaver parameter without making additional data under the condition of a limited amount of received data. We first generate a matrix by using the received data. We then make square submatrices from the matrix composed of the received data and obtain rank deficiency distribution of the square submatrices. Finally, we estimate the interleaver parameter by comparing rank deficiency distribution of the square submatrices and that of random binary matrices. To validate the proposed method, we show the simulation results in terms of detection probability and the number of false alarms for various limited amounts of received data.

The rest of the paper is organized as follows. Section II introduces the system model and Section III proposes an improved method for the blind estimation of interleaver parameter from scant data. Section IV shows the simulation results to validate the proposed method, followed by the conclusions in Section V.

II. SYSTEM MODEL

Let us assume a transmitter using an (n_c, k_c) linear block code and an interleaver with period P, where P is a multiple of the codeword length n_c; a codeword consists of k_c message bits and $n_c - k_c$ parity bits. The parity bits are generated by the linear combination of message bits and the interleaver changes the order of codeword bits in every period P. We further assume that a total number of the received data bits is M and the predicted interleaving period is L in a non-cooperative context.

If we sequentially partition the received data as w row vectors of length L, then the i-th row vector $\mathbf{s}_i$ can be written as

$$\mathbf{s}_i = [b_1^i \ b_2^i \ \cdots \ b_L^i] \tag{1}$$

where $w = \left\lfloor \frac{M}{L} \right\rfloor$, $\lfloor \cdot \rfloor$ is the floor function, b_j^i is the j-th bit of $\mathbf{s}_i$, $b_j^i \in \{0, 1\}$ for $1 \leq i \leq w$ and $1 \leq j \leq L$, and $w > L$. If we arbitrarily select L different vectors from the w row vectors and pile up the selected L vectors row by row, we can generate an $L \times L$ square matrix $\mathbf{R}$. By repeating the process, we can generate a total number of $\binom{w}{L}$ square matrices $\mathbf{R}$'s, where $\binom{x}{y}$ is the binomial coefficient.

If the predicted interleaving period L is different from the original interleaving period P, the message bits and the parity bits are not aligned in columns of $\mathbf{R}$ and the linearity in a codeword is lost [12]; in this case, the rank deficiency of $\mathbf{R}$ becomes similar to that of a random binary matrix composed of elements in the Galois field GF(2). Therefore, if L is not equal to P, the rank deficiency distribution of matrices $\mathbf{R}$'s will also become similar to that of random binary matrices. The probabilities that the rank deficiency of random binary matrix becomes ε are known to be 0.288788, 0.577576, 0.128350, 0.005239, and 0.000047 when ε equals 0, 1, 2, 3, and 4, respectively [44]. On the other hand, if the predicted interleaving period L is equal to the original interleaving period P, the message bits and the parity bits are aligned in columns of $\mathbf{R}$ and the linearity in a codeword is maintained [12]. Therefore, if L is equal to P, the rank deficiency distribution of matrices $\mathbf{R}$'s will differ from that of random binary matrices [21], [22].

Using the above properties, the methods in [21] and [22] compared the rank deficiency distribution of matrices $\mathbf{R}$'s with that of random binary matrices to estimate interleaving period. These methods generally do estimate the interleaver parameter given sufficient collected data. When the amount of received data is limited such that w is smaller than P, not even a single $P \times P$ square matrix $\mathbf{R}$ can be generated. Therefore, the methods in [21] and [22] become infeasible given a limited amount of data [23].

To solve that problem, [23] proposed a novel method for blind estimation of interleaver parameter when only a limited amount of received data is available. The method in [23] makes additional row vectors $\mathbf{c}_J$'s by linear combination of n different $\mathbf{s}_i$'s, where $\mathbf{c}_J = \bigoplus_{j \in J} \mathbf{s}_j$, J is the index set of size n, and $\oplus$ denotes modulo-2 addition. For example, when $n = 2$ and $J = \{x, y\}$, $\mathbf{c}_J$ becomes $\mathbf{s}_x \oplus \mathbf{s}_y$, where $1 \leq x \leq w$, $1 \leq y \leq w$, and $x \neq y$. Note that, in this case, there is a total number of $w + \binom{w}{n}$ vectors, which may be enough vectors to generate square matrices [23]. Meanwhile, in the process of making $\mathbf{c}_J$'s, additional linearity among vectors occurs. For example, when $n = 2$ and $J = \{x, y\}$, there is a linear relation among vectors $\mathbf{s}_x$, $\mathbf{s}_y$, and $\mathbf{c}_J$ since $\mathbf{c}_J = \mathbf{s}_x \oplus \mathbf{s}_y$. Therefore, if we denote the $L \times L$ square matrix composed of L different vectors from $\mathbf{c}_J$'s and $\mathbf{s}_i$'s as $\mathbf{R}_\mathrm{N}$ for notational convenience, unlike the rank deficiency of $\mathbf{R}$, there are two types of linearity relating to the rank deficiency of $\mathbf{R}_\mathrm{N}$: one is the linearity in a codeword and the other is the linearity among vectors, which additionally occurs in the process of making $\mathbf{c}_J$'s [23]. The latter linearity degrades the estimation performance. Therefore, it can be expected that the estimation performance can be improved if there is a blind estimation method without making additional data under the condition of a limited amount of received data. We propose just such an enhanced method in Section III.

III. PROPOSED METHOD

When the amount of received data is limited such that w is smaller than P, we cannot obtain rank deficiency distribution because not even a single $P \times P$ square matrix can be generated for this case. The method in [23] made additional data to generate a sufficient number of square matrices under the condition of scant data as explained in Section II. The cost of this is degraded estimation, because of the linearity arising from the process of making additional data. To solve this problem, we propose a method by using the rank deficiency distribution of square submatrices obtained from the matrix composed of received data, without making additional data. By comparing the rank deficiency distribution of the square submatrices and that of the random binary matrices, we blindly estimate interleaver parameter without making additional data, even with scant received data. To do this, we first generate a matrix by using the received data. We then make square submatrices from the matrix composed of received data and obtain rank deficiency distribution of the square submatrices. Finally, we estimate the interleaver parameter by comparing rank deficiency distribution of the square submatrices and that of random binary matrices. We develop the proposed method with the following steps.

If we arrange the w row vectors of (1) row by row, we can generate a $w \times L$ matrix $\mathbf{R}_\mathrm{a}$ as

$$\mathbf{R}_\mathrm{a}=\begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \vdots \\ \mathbf{s}_w \end{bmatrix}=\begin{bmatrix} b_1^1 & b_2^1 & \cdots & b_L^1 \\ b_1^2 & b_2^2 & \cdots & b_L^2 \\ \vdots & \vdots & \ddots & \vdots \\ b_1^w & b_2^w & \cdots & b_L^w \end{bmatrix}. \tag{2}$$

By arbitrarily deleting some rows and columns from the matrix $\mathbf{R}_\mathrm{a}$, we can compose different $k \times k$ square submatrices $\mathbf{R}_\mathrm{s}$'s [45], where k is an integer smaller than min(w, L), and min(x, y) is the min operation. Note that, for a given k, we can generate a total number of $\binom{w}{k} \times \binom{L}{k}$ different square submatrices $\mathbf{R}_\mathrm{s}$'s and this may be a sufficient number of square submatrices to obtain rank deficiency distribution. For example, when we receive 26×28 data bits and the predicted interleaving period L is 28, we only have $w = \frac{26\times 28}{28} = 26$ row vectors. In this case, not even a single 28×28 square matrix can be generated. However, if we compose 24×24 square submatrices $\mathbf{R}_\mathrm{s}$'s, i.e. $k = 24$, then we can generate $\binom{26}{24} \times \binom{28}{24} = 6\,654\,375$ different square submatrices $\mathbf{R}_\mathrm{s}$'s, and this is enough to obtain rank deficiency distribution.

We examine the rank deficiency distribution of $\mathbf{R}_\mathrm{s}$, which is generated by deleting some rows and columns from $\mathbf{R}_\mathrm{a}$, for two cases: one is for $L \neq P$ and the other is for $L = P$. When $L \neq P$, the message bits and parity bits are not aligned in columns of $\mathbf{R}_\mathrm{a}$ in (2). Consequently, if we generate a $k \times k$ square submatrix $\mathbf{R}_\mathrm{s}$ when $L \neq P$, neither the message bits nor parity bits are aligned in columns of $\mathbf{R}_\mathrm{s}$ and the linearity in a codeword is lost. In this case, the rank deficiency of $\mathbf{R}_\mathrm{s}$ becomes similar to that of a random binary matrix. Therefore, if $L \neq P$, the rank deficiency distribution of submatrices $\mathbf{R}_\mathrm{s}$'s will become similar to that of random binary matrices.

On the other hand, when $L = P$, the message bits and parity bits are aligned in columns of $\mathbf{R}_\mathrm{a}$ in (2). If we compose a $k \times k$ square submatrix $\mathbf{R}_\mathrm{s}$ when $L = P$, some of the message bits and parity bits in $\mathbf{R}_\mathrm{a}$ are also aligned in columns of $\mathbf{R}_\mathrm{s}$ and we expect the linearity in a codeword to be maintained. To show that the linearity in a codeword is maintained in $\mathbf{R}_\mathrm{s}$, we investigate the rank deficiency of $\mathbf{R}_\mathrm{s}$ mathematically. If the expected value of rank deficiency of $\mathbf{R}_\mathrm{s}$ is larger than that of a random binary matrix, we can consider that the rank deficiency distribution of $\mathbf{R}_\mathrm{s}$ is different from that of random binary matrices, and the linearity in a codeword is maintained in $\mathbf{R}_\mathrm{s}$. Therefore, we derive the expected value of minimum rank deficiency of $\mathbf{R}_\mathrm{s}$ and show that it is larger than the expected value of rank deficiency of a random binary matrix when $L = P$.

We start from a simple case for $P = n_c$ and then generalize the result to the case for $P = \alpha n_c$, where α is the number of codewords in an interleaving period. When $P = n_c$, there is only one codeword in an interleaving period. In this case, if we arrange the w row vectors of (1) row by row, we can generate a $w \times n_c$ matrix $\mathbf{R}_\mathrm{a}$ of (2) and the maximum rank of $\mathbf{R}_\mathrm{a}$ becomes k_c which is the dimension of (n_c, k_c) linear block code. If we make a $k \times k$ square submatrix $\mathbf{R}_\mathrm{s}$ by deleting some rows and columns from the matrix $\mathbf{R}_\mathrm{a}$, since the rank of $\mathbf{R}_\mathrm{s}$ is less than or equal to the rank of $\mathbf{R}_\mathrm{a}$, the maximum rank of $\mathbf{R}_\mathrm{s}$ also becomes k_c, in other words, the minimum rank deficiency of $\mathbf{R}_\mathrm{s}$ becomes $k - k_c$. Therefore, when $L = P = n_c$, the minimum rank deficiency of $\mathbf{R}_\mathrm{s}$, $f(k)$, can be obtained by

$$f(k) = \begin{cases} k - k_c, & k_c \leq k \leq n_c \\ 0, & 0 \leq k < k_c \end{cases} \tag{3}$$

where k is the number of rows (or columns) of $\mathbf{R}_\mathrm{s}$.

Next, we consider the general case for $P = \alpha n_c$, where there are α codewords in an interleaving period. In this case, if we arrange the w row vectors of (1) row by row, we can generate a $w \times \alpha n_c$ matrix $\mathbf{R}_\mathrm{a}$ of (2), and the maximum rank of $\mathbf{R}_\mathrm{a}$ becomes αk_c, since there are α codewords in each row of $\mathbf{R}_\mathrm{a}$. By deleting some rows and columns from the matrix $\mathbf{R}_\mathrm{a}$, we can obtain a $k \times k$ square submatrix $\mathbf{R}_\mathrm{s}$. If we denote k_i as the number of columns in $\mathbf{R}_\mathrm{s}$ relating to the i-th codeword in each row of $\mathbf{R}_\mathrm{a}$ where $1 \leq i \leq \alpha$ and $0 \leq k_i \leq n_c$, then $\sum_{i=1}^{\alpha} k_i = k$ and the minimum rank deficiency of $\mathbf{R}_\mathrm{s}$ can be calculated as $\sum_{i=1}^{\alpha} f(k_i)$ from (3). Since the number of possible k_i columns in $\mathbf{R}_\mathrm{s}$ relating to the i-th codeword in each row of $\mathbf{R}_\mathrm{a}$ for each k_i is $\binom{n_c}{k_i}$ and the number of codewords in each row of $\mathbf{R}_\mathrm{a}$ is α, there are a total number of $\prod_{i=1}^{\alpha} \binom{n_c}{k_i}$

square submatrices $\mathbf{R}_s$'s for given k_i's. Finally, by considering all possible k_i's for a given k, we can formulate the expected value of minimum rank deficiency of $\mathbf{R}_s$, $r_{\min}(k)$, when $L = P = \alpha n_c$ as

$$r_{\min}(k) = \binom{L}{k}^{-1} \sum_{k_1+k_2+\cdots+k_\alpha=k} \left[\prod_{j=1}^{\alpha} \binom{n_c}{k_j} \sum_{i=1}^{\alpha} f(k_i) \right]. \quad (4)$$

Note that when $\alpha = 1$, (4) reduces to (3).

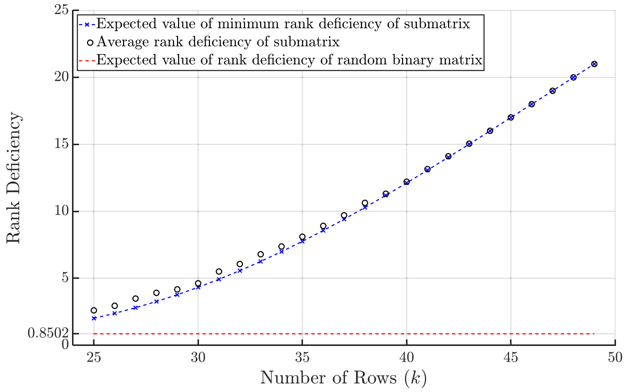


FIGURE 1. Rank deficiencies of submatrix and random binary matrix versus the number of rows (*k*) when interleaving period is 49 and (7, 4) Hamming code is used.

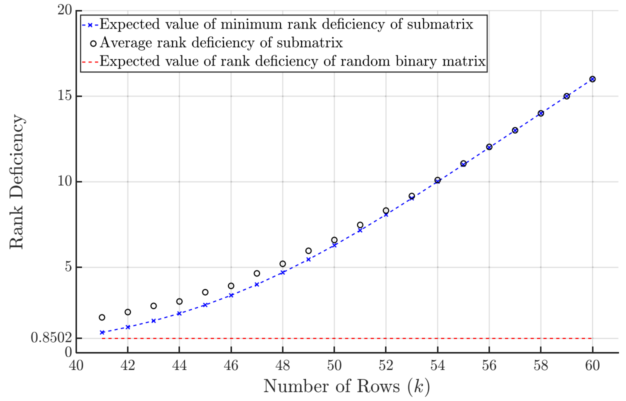


FIGURE 2. Rank deficiencies of submatrix and random binary matrix versus the number of rows (*k*) when interleaving period is 60 and (15, 11) BCH code is used.

To validate $r_{\min}(k)$, we compare $r_{\min}(k)$ obtained from (4) and the average rank deficiency of $\mathbf{R}_s$ obtained from computer simulation by varying k, in Figs. 1 and 2. We assume the interleaving period of 49 and (7, 4) Hamming code in Fig. 1, and the interleaving period of 60 and (15, 11) BCH code in Fig. 2. The average rank deficiency of $\mathbf{R}_s$ is obtained from computer simulation by averaging the rank deficiencies of 100 $\mathbf{R}_s$'s for a given k. From Figs. 1 and 2, we can see that $r_{\min}(k)$ is less than or equal to the average rank deficiency of $\mathbf{R}_s$. To be specific, the gap between $r_{\min}(k)$ and the average rank deficiency of $\mathbf{R}_s$ decreases as k increases and the gap becomes 0 when k is equal to the original interleaving period. We also show the expected value of rank deficiency of a random binary matrix in Figs. 1 and 2, where the expected value of rank deficiency of the random binary matrix is about 0.8502 [44]. From Figs. 1 and 2, we can also see that, when $L = P$, $r_{\min}(k)$ is different from the expected value of rank deficiency of a random binary matrix. Therefore, we can confirm that when $L = P$, the linearity in a codeword is maintained in $\mathbf{R}_s$ and the rank deficiency distribution of submatrices $\mathbf{R}_s$'s becomes different from that of random binary matrices. Consequently, by using the rank deficiency distribution of $\mathbf{R}_s$'s and that of random binary matrices, we can estimate the interleaver parameter under the condition of a limited amount of received data.

For the comparison of the rank deficiency distribution of $\mathbf{R}_s$'s and that of random binary matrices, we adopt mean square error as a measure, which is calculated as

$$D_{MSE} = \frac{1}{k+1} \sum_{i=0}^{k} \{P(X = i) - P(Y = i)\}^2 \quad (5)$$

where $P(X = i)$ and $P(Y = i)$ are the rank deficiency distribution of $\mathbf{R}_s$'s and that of random binary matrices respectively, i is rank deficiency, and k is the number of rows (or columns) of $\mathbf{R}_s$. If $P(X = i)$ and $P(Y = i)$ are similar, i.e., the rank deficiency distribution of $\mathbf{R}_s$'s and that of random binary matrices are similar, D_{MSE} will become relatively small, on the other hand, if two distributions are different from each other, D_{MSE} will become relatively large. Therefore, by obtaining D_{MSE} for each L, we can estimate L as the original interleaving period P when D_{MSE} has the largest value.

After estimating L as the original interleaving period P, we compare once more the rank deficiency distribution of $\mathbf{R}_s$'s and that of random binary matrices. The mean square error between the rank deficiency distribution of $\mathbf{R}_s$'s and that of random binary matrices, D_{MSE}, may have the largest value even when $L \neq P$ because of the erroneous bits caused by noise. In this case, a false alarm occurs. To control the false alarm, we adopt the Kullback-Leibler divergence (KLD), which is typically used to check the similarity of the two probability distributions [22]. KLD is the relative entropy between the two probability distributions; therefore, it has a non-negative near-zero value for similar distributions and relatively large value as the difference between the distributions increases. When KLD becomes larger than γ, we finally declare that L is the original interleaving period P, where the threshold γ is a design parameter to control false alarms. The larger γ becomes, the less often false alarms occur.

Finally, we can summarize the proposed method step by step in Algorithm 1.

IV. SIMULATION RESULTS

In this section, we validate the proposed method through computer simulations in terms of detection probability and the number of false alarms. In the simulations, a random interleaver with interleaving period P, binary phase shift keying modulation, and an additive white Gaussian noise (AWGN)

Algorithm 1 Estimation of the Interleaving Period by Using Square Submatrices

```
Notation of Variable: $Cnt$ denotes the number of matrices generated to calculate rank deficiency distribution. $L^*$ is the predicted interleaving period.
Input: $M$-bit received data
for $L^* = L_{\min} : L_{\max}$ do
    Sequentially partition the received data as $w$ row vectors $\mathbf{s}_i$'s with length $L^*$
    Arrange the $w$ row vectors row by row to generate matrix $\mathbf{R}_{\mathrm{a}}$ of (2)
    for $i = 1 : Cnt$ do
        Delete some rows and some columns from the matrix $\mathbf{R}_{\mathrm{a}}$ to obtain $k \times k$ square submatrix $\mathbf{R}_{\mathrm{s}}$
        Calculate the rank deficiency of $\mathbf{R}_{\mathrm{s}}$
    end
    Calculate and record rank deficiency distribution of $\mathbf{R}_{\mathrm{s}}$'s
    Calculate $D_{MSE}$ in (5)
    Update $L$ as $L^*$ when $D_{MSE}$ is larger than the previous maximum value of $D_{MSE}$
end
Calculate KLD by using the rank deficiency distribution corresponding to $L$
Declare $L$ as the original interleaving period $P$ when KLD is larger than $\gamma$
Output: Estimated interleaving period $L$
```

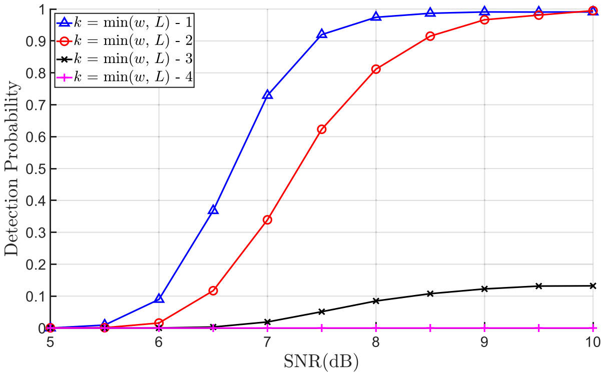


FIGURE 3. Detection probability according to k when interleaving period is 60 and (15, 11) BCH code is used in an AWGN channel.

channel are assumed. We limit the amount of received data, M, less than $P \times P$ bits in the simulations to validate that the proposed method can effectively estimate the interleaving period when there is only a limited amount of received data available. We examine the detection performance of the proposed method according to k, which is the number of rows (or columns) in a square submatrix $\mathbf{R}_{\mathrm{s}}$, γ, which is a threshold to control the false alarm, and M, which is the amount of received data. We include the results of the methods in [13] and [23] for comparison.

Firstly, we investigate the detection performance of the proposed method according to k where k is smaller than min(w, L). Generally, as k increases under the condition of k being less than min(w, L), the possibility of maintaining linearity in a codeword in $\mathbf{R}_{\mathrm{s}}$ also increases and we can expect better detection performance. We show the detection probability of the proposed method in Fig. 3 by varying k from min(w, L) – 1 to min(w, L) – 4, when the interleaving period is 60, (15, 11) BCH code is used, γ is 0.75, and M is 2520 ($60 \times 60 \times 0.7$) bits. From Fig. 3, we find that the detection performance improves as k increases, as we expected. Therefore, in general, we can choose k as min(w, L) – 1 to acquire better estimation performance. For some cases, such as a small values of w and L, if we set k to min(w, L) – 1, we cannot compose a sufficient number of $\mathbf{R}_{\mathrm{s}}$'s to obtain a rank deficiency distribution. In such a case, we can choose k smaller than min(w, L) – 1. Consequently, we set k to a design parameter, and choose k as large as possible to compose a sufficient number of $\mathbf{R}_{\mathrm{s}}$'s for calculating rank deficiency distribution.

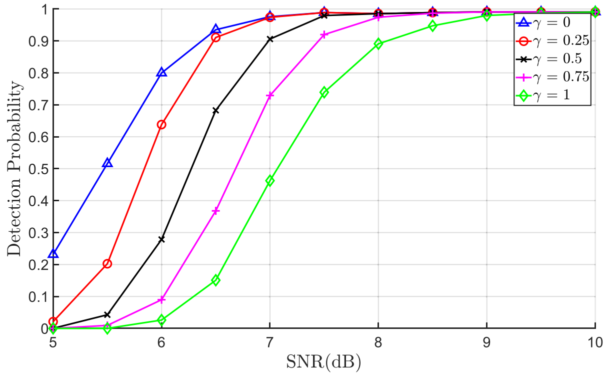


FIGURE 4. Detection probability according to γ when interleaving period is 60 and (15, 11) BCH code is used in an AWGN channel.

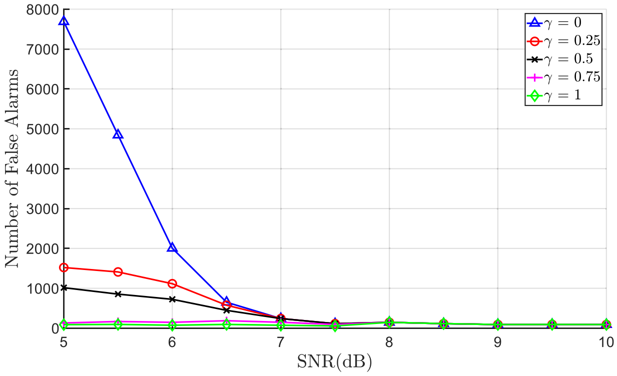


FIGURE 5. Number of false alarms in 10 000 iterations according to γ when interleaving period is 60 and (15, 11) BCH code is used in an AWGN channel.

We depict the detection probability and the number of false alarms of the proposed method for various values of γ in Figs. 4 and 5, respectively, when the interleaving period is 60, (15, 11) BCH code is used, k is min(w, L) – 1, and M is 2520 ($60 \times 60 \times 0.7$) bits. We can see from Fig. 4 that the detection probability increases as γ decreases. And from Fig. 5, we find that the number of false alarms decreases as γ increases. To be specific, when γ is 0, the detection probability becomes better but there are a number of false alarm occurrences. On the other hand, when γ is 1, the detection probability becomes worse but there are only a few false alarms. Therefore, we can confirm from Figs. 4 and 5 that there is trade-off between the detection probability and the number of false alarms according to the design parameter γ.

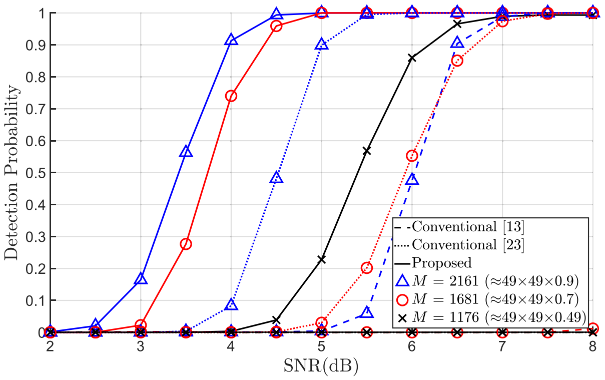


FIGURE 6. **Detection probability according to *M* when interleaving period is 49 and (7, 4) Hamming code is used in an AWGN channel.**

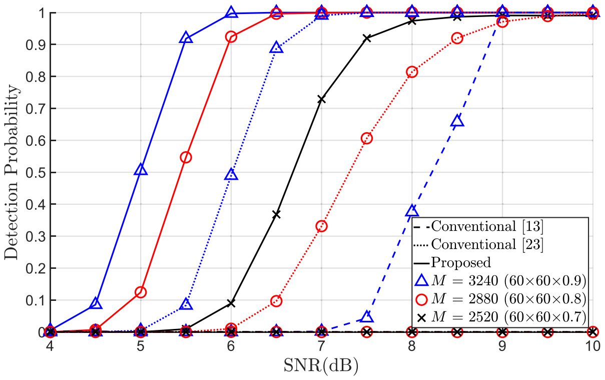
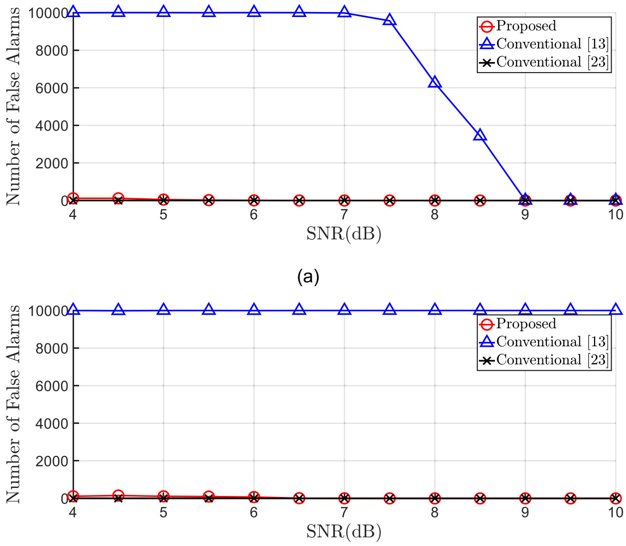


FIGURE 7. **Detection probability according to *M* when interleaving period is 60 and (15, 11) BCH code is used in an AWGN channel.**

Finally, we validate the performance of the proposed method in an AWGN channel according to the amount of received data M in Figs. 6, 7, and 8. We show the detection probabilities when the interleaving period is 49 and (7, 4) Hamming code is used in Fig. 6, and when the interleaving period is 60 and (15, 11) BCH code is used in Fig. 7. We also depict the number of false alarms when the interleaving period is 60 and (15, 11) BCH code is used in Fig. 8. In the simulations, we set M to $P \times P \times \beta$ bits for $0 < \beta < 1$, γ to 0.75, and k to min(w, L) – 1. For comparison, we include the performance of the conventional methods in [13] and [23].

From Figs. 6 and 7, we can see that the proposed method is superior to the conventional methods of [13] and [23] in detection probability. When M is $P \times P \times 0.9$ bits, at a detection probability of 0.9, the proposed method achieves signal-to-noise ratio (SNR) gains of about 2.5 dB and 3.4 dB over [13], 1.0 dB and 1.1 dB over [23], in Figs. 6 and 7, respectively. For more limited cases of M being 1176 ($\approx 49 \times 49 \times 0.49$) bits in Fig. 6 and 2520 ($60 \times 60 \times 0.7$) bits in Fig. 7, the detection probabilities of the proposed method reach 0.9 at SNRs of 6.2 dB and 7.4 dB, respectively, while the conventional methods in [13] and [23] cannot give any meaningful detection results. From Fig. 8, we can see that there are only a few false alarms with the proposed method, where the results are obtained from 10 000 iterations for each M.



(a)

(b)

FIGURE 8. **Number of false alarms in 10 000 iterations when interleaving period is 60 and (15, 11) BCH code is used in an AWGN channel. (a) *M* = 3240 (60 × 60 × 0.9) bits. (b) *M* = 2880 (60 × 60 × 0.8) bits.**

From the simulation results, we can see that the proposed method can effectively estimate the interleaving period when there is only a limited amount of received data available.

V. CONCLUSION

In this paper, we proposed an improved method for blind estimation of interleaver parameter from scant data. In cases where only a limited amount of received data was available, the previous method estimated interleaver parameter by making additional data with the received data. The cost of this was degraded estimation, because of the linearity arose in the process of making additional data. To solve this problem, we first generated a matrix by using the received data. We then made square submatrices from the matrix composed of the received data and obtained a rank deficiency distribution of the square submatrices. Finally, we estimated the interleaver parameter by comparing rank deficiency distribution of the square submatrices and that of random binary matrices. The proposed method was validated through computer simulations and gave better estimation performance than the conventional methods, given scant received data. Further, for more limited cases of the amount of received data, the proposed method could estimate interleaver parameter whereas the conventional methods in [13] and [23] could not give any meaningful detection results. Therefore, it is expected that the proposed method can effectively estimate interleaver parameter under the condition of scant received data.

ACKNOWLEDGMENT

(Mingyu Jang and Geunbae Kim contributed equally to this work.)

REFERENCES

[1] P. Bestagini, S. Milani, M. Tagliasacchi, and S. Tubaro, "Codec and GOP identification in double compressed videos," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 2298–2310, May 2016.

[2] X. Jiang, P. He, T. Sun, F. Xie, and S. Wang, "Detection of double compression with the same coding parameters based on quality degradation mechanism analysis," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 1, pp. 170–185, Jan. 2018.

[3] X. Liang, Z. Li, Y. Yang, Z. Zhang, and Y. Zhang, "Detection of double compression for HEVC videos with fake bitrate," *IEEE Access*, vol. 6, pp. 53243–53253, Sep. 2018.

[4] D. Vazquez-Padin, M. Fontani, D. Shullani, F. Perez-Gonzalez, A. Piva, and M. Barni, "Video integrity verification and GOP size estimation via generalized variation of prediction footprint," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 1815–1830, 2020.

[5] B. Kwon, H. Song, and S. Lee, "Accurate blind Lempel-Ziv-77 parameter estimation via 1-D to 2-D data conversion over convolutional neural network," *IEEE Access*, vol. 8, pp. 43965–43979, Mar. 2020.

[6] P. Yu, H. Peng, and J. Li, "On blind recognition of channel codes within a candidate set," *IEEE Commun. Lett.*, vol. 20, no. 4, pp. 736–739, Apr. 2016.

[7] A. D. Yardi, S. Vijayakumaran, and A. Kumar, "Blind reconstruction of binary cyclic codes from unsynchronized bitstream," *IEEE Trans. Commun.*, vol. 64, no. 7, pp. 2693–2706, Jul. 2016.

[8] A. Bonvard, S. Houcke, R. Gautier, and M. Marazin, "Classification based on Euclidean distance distribution for blind identification of error correcting codes in noncooperative contexts," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2572–2583, May 2018.

[9] D. Jo, S. Kwon, and D.-J. Shin, "Blind reconstruction of BCH codes based on consecutive roots of generator polynomials," *IEEE Commun. Lett.*, vol. 22, no. 5, pp. 894–897, May 2018.

[10] Z. Wu, Z. Zhong, and L. Zhang, "Blind recognition of cyclic codes based on average cosine conformity," *IEEE Trans. Signal Process.*, vol. 68, pp. 2328–2339, 2020.

[11] R. Swaminathan, A. S. Madhukumar, and G. Wang, "Blind estimation of code parameters for product codes over noisy channel conditions," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 56, no. 2, pp. 1460–1473, Apr. 2020.

[12] G. Burel and R. Gautier, "Blind estimation of encoder and interleaver characteristics in a non-cooperative context," in *Proc. IASTED*, Scottsdale, AZ, USA, 2003, pp. 275–280.

[13] G. Sicot, S. Houcke, and J. Barbier, "Blind detection of interleaver parameters," *Signal Process.*, vol. 89, no. 4, pp. 450–462, Apr. 2009.

[14] J. Jeong, D. Yoon, J. Lee, and S. Choi, "Blind reconstruction of a helical scan interleaver," in *Proc. 8th Int. Conf. Inf., Commun. Signal Process.*, Singapore, Dec. 2011, pp. 1–4.

[15] R. Swaminathan, A. S. Madhukumar, W. T. Ng, and C. M. S. See, "Parameter estimation of block and helical scan interleavers in the presence of bit errors," *Digit. Signal Process.*, vol. 60, pp. 20–32, Jan. 2017.

[16] S. Ramabadran, A. S. Madhukumar, N. Wee Teck, and C. M. S. See, "Parameter estimation of convolutional and helical interleavers in a noisy environment," *IEEE Access*, vol. 5, pp. 6151–6167, Mar. 2017.

[17] R. Swaminathan and A. S. Madhukumar, "Classification of error correcting codes and estimation of interleaver parameters in a noisy transmission environment," *IEEE Trans. Broadcast.*, vol. 63, no. 3, pp. 463–478, Sep. 2017.

[18] R. Swaminathan, A. S. Madhukumar, G. Wang, and T. Shang Kee, "Blind reconstruction of Reed–Solomon encoder and interleavers over noisy environment," *IEEE Trans. Broadcast.*, vol. 64, no. 4, pp. 830–845, Dec. 2018.

[19] Y. Xu, Y. Zhong, and Z. Huang, "An improved blind recognition method of the convolutional interleaver parameters in a noisy channel," *IEEE Access*, vol. 7, pp. 101775–101784, Jul. 2019.

[20] C. Choi and D. Yoon, "Enhanced blind interleaver parameters estimation algorithm for noisy environment," *IEEE Access*, vol. 6, pp. 5910–5915, Sep. 2018.

[21] C. Choi and D. Yoon, "Novel blind interleaver parameter estimation in a noncooperative context," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 4, pp. 2079–2085, Aug. 2019.

[22] G. Kim, M. Jang, and D. Yoon, "Improved method for interleaving parameter estimation in a non-cooperative context," *IEEE Access*, vol. 7, pp. 92171–92175, Jul. 2019.

[23] M. Jang, G. Kim, Y. Kim, and D. Yoon, "Blind estimation of interleaver parameter with a limited number of data," *IEEE Access*, vol. 8, pp. 69160–69166, Apr. 2020.

[24] J.-H. Liang, X. Wang, F.-H. Wang, and Z.-T. Huang, "Blind spreading sequence estimation algorithm for long-code DS-CDMA signals in asynchronous multi-user systems," *IET Signal Process.*, vol. 11, no. 6, pp. 704–710, Aug. 2017.

[25] B. Shen and J.-X. Wang, "Chip rate and pseudo-noise sequence estimation for direct sequence spread spectrum signals," *IET Signal Process.*, vol. 11, no. 6, pp. 727–733, Aug. 2017.

[26] H. Mirzadeh Sarcheshmeh, H. Khaleghi Bizaki, and S. Alizadeh, "PN sequence blind estimation in multiuser DS-CDMA systems with multipath channels based on successive subspace scheme," *Int. J. Commun. Syst.*, vol. 31, no. 12, p. e3591, Aug. 2018.

[27] H. WenJia, "Reconstructing the feedback polynomial of a linear scrambler with the method of hypothesis testing," *IET Commun.*, vol. 9, no. 8, pp. 1044–1047, May 2015.

[28] X. Gu, Z. Zhao, and L. Shen, "Blind estimation of pseudo-random codes in periodic long code direct sequence spread spectrum signals," *IET Commun.*, vol. 10, no. 11, pp. 1273–1281, Jul. 2016.

[29] D. Kim, J. Song, and D. Yoon, "On the estimation of synchronous scramblers in direct sequence spread spectrum systems," *IEEE Access*, vol. 8, pp. 166450–166459, Sep. 2020.

[30] Z. Wu, S. Zhou, Z. Yin, B. Ma, and Z. Yang, "Robust automatic modulation classification under varying noise conditions," *IEEE Access*, vol. 5, pp. 19733–19741, Aug. 2017.

[31] T. R. Kishore and K. D. Rao, "Automatic intrapulse modulation classification of advanced LPI radar waveforms," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 2, pp. 901–914, Apr. 2017.

[32] L. Han, F. Gao, Z. Li, and O. A. Dobre, "Low complexity automatic modulation classification based on order-statistics," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 400–411, Jan. 2017.

[33] J. Lee, J. Kim, B. Kim, D. Yoon, and J. Choi, "Robust automatic modulation classification technique for fading channels via deep neural network," *Entropy*, vol. 19, no. 9, p. 454, Aug. 2017.

[34] D. Zhu, V. J. Mathews, and D. H. Detienne, "A likelihood-based algorithm for blind identification of QAM and PSK signals," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 3417–3430, May 2018.

[35] B. Tang, Y. Tu, Z. Zhang, and Y. Lin, "Digital signal modulation classification with data augmentation using generative adversarial nets in cognitive radio networks," *IEEE Access*, vol. 6, pp. 15713–15722, Mar. 2018.

[36] F. Meng, P. Chen, L. Wu, and X. Wang, "Automatic modulation classification: A deep learning enabled approach," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 10760–10772, Nov. 2018.

[37] Y. Wang, M. Liu, J. Yang, and G. Gui, "Data-driven deep learning for automatic modulation recognition in cognitive radios," *IEEE Trans. Veh. Technol.*, vol. 68, no. 4, pp. 4074–4077, Apr. 2019.

[38] S. Zheng, P. Qi, S. Chen, and X. Yang, "Fusion methods for CNN-based automatic modulation classification," *IEEE Access*, vol. 7, pp. 66496–66504, May 2019.

[39] B. Sklar, *Digital Communications: Fundamentals and Applications*. Upper Saddle River, NJ, USA: Prentice-Hall, 2001.

[40] G. Proakis, *Digital Communications*, 4th ed. New York, NY, USA: McGraw-Hill, 2001.

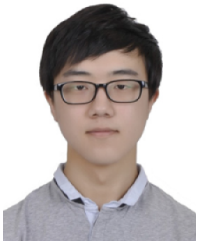
[41] L. I. Bluestein, "Interleaving of pseudorandom sequences for synchronization," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-4, no. 4, pp. 551–556, Jul. 1968.

[42] J. Ramsey, "Realization of optimum interleavers," *IEEE Trans. Inf. Theory*, vol. IT-16, no. 3, pp. 338–345, May 1970.

[43] R. Garello, G. Montorsi, S. Benedetto, and G. Cancellieri, "Interleaver properties and their applications to the trellis complexity analysis of turbo codes," *IEEE Trans. Commun.*, vol. 49, no. 5, pp. 793–807, May 2001.

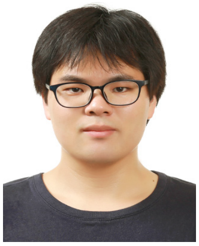
[44] V. F. Kolchin, *Random Graphs*. New York, NY, USA: CUP, 1999.

[45] C. Clapham and J. Nicholson, *The Concise Oxford Dictionary of Mathematics*, 4th ed. Oxford, U.K.: OUP, 2009.



MINGYU JANG received the B.S. degree in electronic engineering from Hanyang University, Seoul, South Korea, in 2019, where he is currently pursuing the M.S. degree with the Department of Electronic Engineering. His research interests include digital communication theory and wireless communications.

GEUNBAE KIM received the B.S., M.S., and Ph.D. degrees in electronic communications engineering from Hanyang University, Seoul, South Korea, in 1991, 1993, and 2012, respectively. He is currently a Research Professor with the Signal Intelligence Research Center, Hanyang University. His research interests include channel coding, signal intelligence, and wireless communications.



DONGYEONG KIM received the B.S. and Ph.D. degrees in mathematics from Hanyang University, Seoul, South Korea, in 2013 and 2020, respectively, under the supervision of Prof. J. Song. His research interests include cryptanalysis of symmetric-key cryptography, channel coding theory, and post quantum cryptography.

DONGWEON YOON received the B.S. *(summa cum laude)*, M.S., and Ph.D. degrees in electronic communications engineering from Hanyang University, Seoul, South Korea, in 1989, 1992, and 1995, respectively. From March 1995 to August 1997, he was an Assistant Professor with the Department of Electronic and Information Engineering, Dongseo University, Busan, South Korea. From September 1997 to February 2004, he was an Associate Professor with the Department of Information and Communications Engineering, Daejeon University, Daejeon, South Korea. Since March 2004, he has been on the faculty of Hanyang University, where he is currently a Professor with the Department of Electronic Engineering and the Director of the Signal Intelligence Research Center. His research interests include digital communications theory and systems, detection and estimation, satellite and space communications, and communication forensics.