IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 14, 2020, accepted March 31, 2020, date of publication April 7, 2020, date of current version April 23, 2020.

*Digital Object Identifier 10.1109/ACCESS.2020.2986273*

# Blind Estimation of Interleaver Parameter With a Limited Number of Data

**MINGYU JANG, GEUNBAE KIM, YOONJI KIM, AND DONGWEON YOON, (Senior Member, IEEE)**
Department of Electronics and Computer Engineering, Hanyang University, Seoul 04763, South Korea

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

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

**ABSTRACT** In non-cooperative contexts, one needs to estimate communication parameters by using collected data without any prior information. Particularly, when collecting only a limited number of data, estimation becomes a more challenging task. This paper presents a method for estimation of interleaver parameter when only a limited number of collected data are available. We first create additional data by combining a limited number of collected data. We then investigate the rank deficiency of the matrices composed of the collected and additionally-created data. Finally, we estimate the interleaver parameter by using the difference of average rank deficiencies. Through computer simulations, we validate the proposed method in terms of detection probability and the number of false alarms.

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

## I. INTRODUCTION

Blind estimation of communication parameters can make essential contributions to both cooperative and non-cooperative contexts, such as mobile wireless communications, surveillance systems, and cognitive radios [1]. For cooperative contexts, such as in 4G-LTE and 5G new radio mobile communication systems, some of the communication parameters should be blindly detected among the candidates. For example, user equipment performs blind detection for its corresponding physical downlink control channel in the search space to decode downlink control information which contains the physical layer transmission parameters, e.g., adaptive modulation and coding, power control, and resource information [2]–[5]. Meanwhile, blind estimation of communication parameters has played a more important role in non-cooperative contexts. For non-cooperative contexts, such as spectrum surveillance and cognitive radios, exploiting only the received data, we must blindly estimate the communication parameters used in a transmission system including source coding, channel coding, interleaving, scrambling, modulation and more, because the receiver does not know any parameters used in the transmitter [6]–[29].

Research on the blind estimation of communication parameters has been conducted, including but not limited to: modulation [6]–[11], spread spectrum sequence [12], [13], channel coding [14]–[17], interleaving [18]–[27], and source coding [28], [29]. One of the key ideas for estimation in the previous research is using the inherent linearity of the collected data relating to linear operations in the communication system. Among the various components of a communication system, the representative components relating to the linear property are channel coding and interleaving, which make data more resilient to the noise, fading, and interference induced in communication channels [30]–[34]. In particular, we focus on estimation of interleaver parameter for a non-cooperative context in this paper.

Interleaver parameter estimation has been studied for many years in many places in the literature [18]–[27]. [18] and [19] used rank of matrix and Gauss-Jordan elimination through pivoting for the estimation, respectively. References [20]–[24] estimated the interleaver parameter by evaluating the number of ones or zeros in each column (or row) of the matrix generated from collected data. Reference [20] proposed a blind reconstruction method of a helical scan interleaver with block channel coded data. For the case of convolutional channel coded data, [21] estimated block and helical scan interleavers, and [22] estimated convolutional and helical interleavers. For the cases of block and convolutional channel coded data, [23] addressed block interleaver and [24] estimated convolutional interleaver. Furthermore, [25]–[27] estimated the interleaver parameter based on

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

the rank deficiency distributions: [25] proposed an estimation algorithm choosing vectors having fewer errors; [26] used binomial distribution to compare rank deficiency distributions for the matrices generated from collected data and from random binary data; and [27] presented an enhanced estimation method that can be efficiently applied to more severe channel conditions. The previous methods generally estimate the interleaver parameters under the condition of a sufficient number of collected data. However, if the number of collected data falls below a certain limit, the previous methods become infeasible. In this case, estimation becomes more challenging.

In this paper, we present a method for the estimation of interleaver parameter when only a limited number of collected data are available. We first create additional data by combining a limited number of collected data. We then examine the rank deficiency of the matrices composed of the collected and additionally-created data. Finally, we estimate the interleaver parameter by using the difference of average rank deficiencies. Through computer simulations, we validate the proposed method in terms of detection probability and the number of false alarms.

This paper is organized as follows. Section II introduces the basic idea for the estimation of interleaver parameter. Section III proposes a method for the estimation of interleaver parameter when only a limited number of collected data are available. Section IV validates through computer simulations the proposed method in terms of detection probability and the number of false alarms. Section V offers conclusions.

## II. SYSTEM SETUP

Let us assume that transmitter uses an $(n_c, k_c)$ linear block code and an interleaver with period $P$ which is a multiple of the codeword length $n_c$, where a codeword consists of $k_c$ message bits and $n_c$-$k_c$ parity bits that are made by linear combinations of the message bits, and the interleaver changes the order of codeword bits in every period $P$. In a non-cooperative context, the $M$-bit collected data sequence $C$ can be expressed as

$$C = \{a_1, a_2, \ \dots \ , \ a_M\} \tag{1}$$

where $a_q$ is the $q$-th bit of the data sequence $C$ and $a_q \in \{0, 1\}$ for $1 \le q \le M$. If we sequentially divide the collected data sequence $C$ with a length of predicted interleaving period $\tilde{P}$ and make $w$ row vectors by using the divided data, then the vector multiset $S_O$ which consists of $w$ row vectors can be expressed as

$$S_O = \{\mathbf{s}_1, \ \mathbf{s}_2, \ \dots, \ \mathbf{s}_w\} \tag{2}$$

where $w = \left\lfloor \frac{M}{\tilde{P}} \right\rfloor$, $\lfloor \cdot \rfloor$ is the floor function, and the $i$-th vector $\mathbf{s}_i$ with the length of $\tilde{P}$ is

$$\mathbf{s}_i = [a_{(i-1)\tilde{P}+1} \ a_{(i-1)\tilde{P}+2} \ \dots \ a_{i\tilde{P}} \ ] \quad (1 \le i \le w). \tag{3}$$

Let $\mathbf{R}_O$ be a $\tilde{P} \times \tilde{P}$ square matrix generated by arbitrarily selecting $\tilde{P}$ different vectors from $w$ vectors in $S_O$ and arranging the selected $\tilde{P}$ different vectors row by row. In this case, the number of square matrices $\mathbf{R}_O$ we can generate is ${}_wC_{\tilde{P}}$, where ${}_xC_y$ is the binomial coefficient.

If the predicted interleaving period $\tilde{P}$ is different from the original interleaving period $P$, the linear dependence in a codeword is lost in $\mathbf{s}_i$, and message and parity bits are not aligned on columns of $\mathbf{R}_O$ [25]–[27]. In this case, rank deficiency of $\mathbf{R}_O$ is similar to that of a matrix generated from random binary data, and consequently, the rank deficiency distribution of $\mathbf{R}_O$ becomes similar to that of random binary matrices. It is well known that the probabilities that rank deficiencies of a random binary matrix become 0, 1, 2, 3, and 4 are about 0.288788, 0.577576, 0.128350, 0.005239, and 0.000047, respectively [35]. On the other hand, if $\tilde{P}$ is equal to $P$, the linear dependence in a codeword is maintained in $\mathbf{s}_i$, and message and parity bits are aligned on columns of $\mathbf{R}_O$ [25]–[27]. In this case, rank deficiency occurs in $\mathbf{R}_O$ as the number of parity bits in the interleaving period $P$, therefore, the rank deficiency distribution of $\mathbf{R}_O$ becomes different from that of random binary matrices.

The methods in [25]–[27] use the above properties to estimate interleaving period through comparison between rank deficiency distribution of $\mathbf{R}_O$ and that of random binary matrices, under the condition of a sufficient number of matrices for calculation of rank deficiency distribution. However, if the number of collected data is so limited that not even a single $P \times P$ matrix $\mathbf{R}_O$ can be generated, we cannot calculate the rank deficiency distribution of $\mathbf{R}_O$, unlike the methods in [25]–[27]. To solve that problem, we present in this paper a method for estimation of interleaving period by creating additional vectors through combining $\mathbf{s}_i$'s, and by using the difference of average rank deficiencies.

## III. PROPOSED METHOD

Now we propose a method to estimate the interleaving period under the condition of a limited number of collected data. First, we make the vector multiset $S_O$ of (2) from a limited number of collected data. Then, we create additional vectors by linear combinations of $n$ different vectors in $S_O$ to obtain a sufficient number of vectors for calculation of the rank deficiency distribution of $\tilde{P} \times \tilde{P}$ square matrices, where $n \ge 2$. Finally, we estimate the interleaving period by using the difference of average rank deficiencies under the condition of a limited number of collected data.

When only a limited number of collected data are available in a non-cooperative context, to obtain a sufficient number of vectors for calculation of the rank deficiency distribution, we propose creating additional vectors. If we select $n$ different vectors from the $w$ vectors in $S_O$ of (2) generated from a limited number of collected data, and take modulo-2 additions to the selected vectors in every combination, then we have ${}_wC_n$ additional vectors that compose the created vector multiset $S_C$

$$S_C = \left\{ \mathbf{c}_J = \bigoplus_{j \in J} \mathbf{s}_j \mid |J| = n \ , \ J \subset I \right\} \tag{4}$$

where $I$ is an index set of $I = \{1, 2, \dots, w\}$, $|J|$ is the cardinality of $J$, and $\oplus$ denotes the modulo-2 addition. By using

two vector multisets, $S_O$ of (2) and $S_C$ of (4), we can compose a new vector multiset, $S_N$

$$S_N = S_O + S_C. \tag{5}$$

Note that the total number of vectors we can acquire becomes $w + {}_wC_n$ in (5) and this may be a sufficient number of vectors to generate matrices and calculate rank deficiency distribution.

Let $\mathbf{R}_N$ be a $\tilde{P} \times \tilde{P}$ square matrix generated by arbitrarily selecting $\tilde{P}$ different vectors from $w + {}_wC_n$ vectors in $S_N$ of (5) and arranging the selected vectors row by row. For example, if $\tilde{P}$ is 49 and $w$ is 40, we can only obtain 40 vectors from (2). For this case, we cannot generate even a single $49 \times 49$ square matrix $\mathbf{R}_O$, and thus cannot calculate a rank deficiency distribution. On the other hand, if we select two (i.e., $n = 2$) different vectors from 40 vectors in $S_O$ of (2) in every combination, then we can create ${}_{40}C_2 = 780$ additional vectors from (4), acquire a total of 820 vectors from (5), and generate ${}_{820}C_{49}$ different $49 \times 49$ square matrices $\mathbf{R}_N$. Note that even when $n = 2$, we can obtain ${}_{820}C_{49}$ matrices, enough to calculate rank deficiency distribution.

After creating additional vectors, we have to examine the linear dependence among the vectors in the new vector multiset, $S_N$ of (5). To simplify explanation, we take an example when $n = 2$. In this case, $J = \{i, j\}$ and $\mathbf{c}_J$ becomes $\mathbf{c}_J = \mathbf{c}_{\{i,j\}} = \mathbf{s}_i \oplus \mathbf{s}_j$ in (4). If $\tilde{P}$ is equal to $P$, since $\mathbf{c}_{\{i,j\}}$ in $S_N$ of (5) is the modulo-2 addition of $\mathbf{s}_i$ and $\mathbf{s}_j$, the linear dependence in a codeword is also maintained in $\mathbf{c}_{\{i,j\}}$ as in $\mathbf{s}_i$ and $\mathbf{s}_j$. Therefore, if we generate a $\tilde{P} \times \tilde{P}$ square matrix $\mathbf{R}_N$ by arbitrarily selecting $\tilde{P}$ different vectors in $S_N$, the message and parity bits are aligned on columns of $\mathbf{R}_N$ as in the matrix $\mathbf{R}_O$, and the rank deficiency is the number of parity bits in the interleaving period $P$.

On the other hand, if $\tilde{P}$ is different from $P$, though the linear dependence in a codeword is lost, there is additional linear dependence among vectors in $S_N$, unlike $S_O$. This is due to the following two cases: one is the linear dependence among vectors $\mathbf{s}_i$, $\mathbf{s}_j$, and $\mathbf{c}_{\{i,j\}}$, because $\mathbf{c}_{\{i,j\}} = \mathbf{s}_i \oplus \mathbf{s}_j$ for $i \neq j$, and the other is the linear dependence among vectors $\mathbf{c}_{\{i,j\}}$, $\mathbf{c}_{\{i,k\}}$, and $\mathbf{c}_{\{j,k\}}$, because $\mathbf{c}_{\{i,j\}} \oplus \mathbf{c}_{\{i,k\}} = (\mathbf{s}_i \oplus \mathbf{s}_j) \oplus (\mathbf{s}_i \oplus \mathbf{s}_k) = \mathbf{s}_j \oplus \mathbf{s}_k = \mathbf{c}_{\{j,k\}}$ for $i \neq j \neq k$. Here we have investigated the linear dependence among the vectors in the new vector multiset, $S_N$ for the case $n = 2$, but it can be straightforwardly extended to the general cases for $n \geq 2$.

Consequently, if we compose $\mathbf{R}_N$ by arbitrarily selecting $\tilde{P}$ different vectors in $S_N$, the linearly dependent vectors can be included in $\mathbf{R}_N$, so that even if $\tilde{P}$ is different from $P$, unlike the rank deficiency of $\mathbf{R}_O$, the rank deficiency may additionally occur in $\mathbf{R}_N$. In other words, unlike the rank deficiency distribution of $\mathbf{R}_O$, even if $\tilde{P}$ is different from $P$, the rank deficiency distribution of $\mathbf{R}_N$ is different from that of the random binary matrices. Accordingly, when we estimate the interleaving period using $\mathbf{R}_N$, we cannot use conventional methods such as [26] and [27], which compare the rank deficiency distribution of $\tilde{P} \times \tilde{P}$ square matrices generated by collected data with that of the random binary matrices. Therefore, if we use $\mathbf{R}_N$ for estimation of the interleaving period under the condition of a limited number of collected data, we have to use a new method. We will now examine it in detail.

When $\tilde{P}$ is different from $P$, as $\tilde{P}$ decreases, the number of the vectors in $S_N$ increases, and the number of the linearly dependent vectors that are included in $\mathbf{R}_N$, composed of $\tilde{P}$ different vectors selected from $S_N$, will tend to decrease. On the other hand, as $\tilde{P}$ increases, the number of the vectors in $S_N$ decreases, and the number of the linearly dependent vectors included in $\mathbf{R}_N$ will tend to increase. Therefore, the rank deficiency of $\mathbf{R}_N$ will vary with $\tilde{P}$: there will be a tendency for the rank deficiency of $\mathbf{R}_N$ to increase, as $\tilde{P}$ increases.

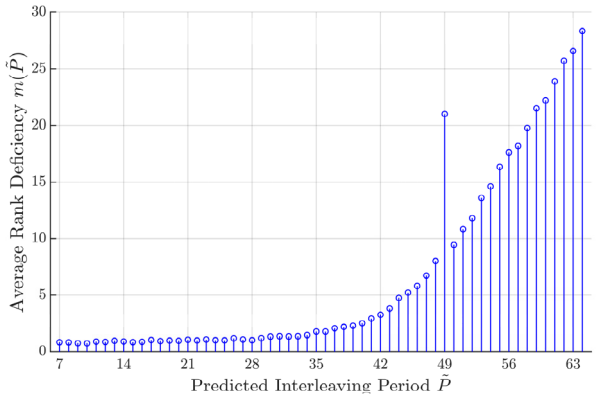



**FIGURE 1.** Average rank deficiency versus predicted interleaving period.

To investigate the tendency of the rank deficiency of $\mathbf{R}_N$ according to $\tilde{P}$, we show the rank deficiency of $\mathbf{R}_N$ by varying $\tilde{P}$ from 7 to 64 in Fig. 1, where the original interleaving period $P$ is 49, the number of collected data is 2401 ($49 \times 49$) bits, $n = 2$, and (7, 4) Hamming code is used. For notational convenience, we denote the average rank deficiency of $\mathbf{R}_N$ as $m(\tilde{P})$ and the difference of average rank deficiencies between $m(\tilde{P})$ and $m(\tilde{P}-1)$ as $\Delta m(\tilde{P})$, i.e., $\Delta m(\tilde{P}) = m(\tilde{P}) - m(\tilde{P}-1)$. In Fig. 1, $m(\tilde{P})$ is obtained by averaging the rank deficiencies of 100 matrices of $\mathbf{R}_N$ for a given $\tilde{P}$, and we can see that $m(\tilde{P})$ varies with $\tilde{P}$. In particular, as $\tilde{P}$ increases, since there is a possibility that the number of linearly dependent vectors included in $\mathbf{R}_N$ increases, $m(\tilde{P})$ tends to increase.

Meanwhile, it is also noteworthy that since the rank deficiency of $\mathbf{R}_N$ is determined by the number of parity bits in the interleaving period $P$ due to the linear dependence in a codeword when $\tilde{P}$ is equal to $P$, $\Delta m(\tilde{P})$ for $\tilde{P} = P$ is larger than the other $\Delta m(\tilde{P})$ for $\tilde{P} \neq P$. That is, $\Delta m(\tilde{P})$, the difference of average rank deficiencies between $m(\tilde{P})$ and $m(\tilde{P}-1)$, becomes the largest when $\tilde{P}$ is equal to $P$. This can be a clue for the estimation of the interleaving period. If $\tilde{P}$ is equal to $P$, the possibility that $\Delta m(\tilde{P})$ becomes the largest is maximized. Therefore, we can decide $\tilde{P}$ as the original interleaving period $P$ when $\Delta m(\tilde{P})$ has the largest value

$$P = \arg\max_{\tilde{P}} \Delta m(\tilde{P}). \tag{6}$$

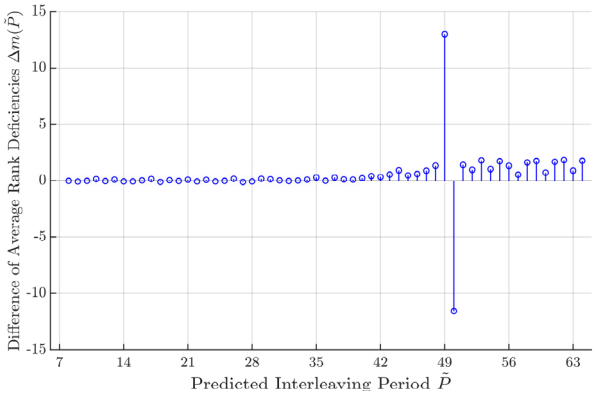



**FIGURE 2.** Difference of average rank deficiencies versus predicted interleaving period.

To verify this, we depict $\Delta m(\tilde{P})$ in Fig. 2 under the same conditions as in Fig. 1. From Fig. 2, we can see that, when $\tilde{P}$ is 49, $\Delta m(\tilde{P})$ becomes the largest and the estimated interleaving period obtained from (6) is equal to the original interleaving period $P$.

In a noisy channel, however, $\Delta m(\tilde{P})$ may not reliably be largest even when $\tilde{P}$ is equal to $P$ because the linear dependence in a codeword can be lost due to the erroneous bits caused by the noise. In this case, the estimated interleaving period obtained by (6) may not be equal to the original interleaving period $P$ and a false alarm may occur. In this paper, we control the false alarm occurrence by using the gap between the largest and the second largest values of $\Delta m(\tilde{P})$. For this, we set the threshold $\gamma$ to $\gamma = \alpha + k$, where $\alpha$ is the second largest value of $\Delta m(\tilde{P})$ and $k$ is a design parameter. If the largest value of $\Delta m(\tilde{P})$ is greater than $\gamma$, that is, if the gap is larger than $k$, then we declare that the estimated interleaving period obtained by (6) is the original interleaving period. Note that as $k$ increases, the number of false alarms decreases.

In Algorithm 1, we summarize the method for estimation of the interleaving period under the condition of a limited number of collected data, where $P_{\min}$ and $P_{\max}$ are minimum and maximum values of $\tilde{P}$, respectively, and $D$ is the number of matrices $\mathbf{R}_N$.

After estimating the interleaving period which is the most important interleaver parameter, other interleaver parameters can straightforwardly be estimated by using the previous results, such as the results of [20], [21], and [23]. For example, when the interleaving period is correctly estimated, the numbers of rows and columns of helical scan interleaver can be estimated by the methods in [20] and [21], and those of block interleaver can be estimated by the methods in [21] and [23].

## IV. SIMULATION RESULTS

We validate the proposed method in terms of detection probability and the number of false alarms through computer simulations. In the simulations, we assume a random interleaver

**Algorithm 1** Estimation of the Interleaving Period by Using the Difference of Average Rank Deficiencies

**Input**: Collected data sequence $C$
1: **for** $\tilde{P} = P_{\min} : P_{\max}$ **do**
2: Generate vector multiset $S_O$ of (2) by sequentially dividing $C$ to the vectors with length $\tilde{P}$
3: Combine $n$ different vectors in $S_O$ by modulo-2 addition and create vector multiset $S_C$ of (4)
4: Composing vector multiset $S_N$ of (5)
5: **for** $i = 1 : D$ **do**
6: Construct matrices $\mathbf{R}_N$ by choosing $\tilde{P}$ vectors from $S_N$ and arranging them row by row
7: Calculate the rank deficiency of $\mathbf{R}_N$
8: **end**
9: Average the rank deficiencies of matrices $\mathbf{R}_N$ to calculate $m(\tilde{P})$
10: **end**
11: Calculate$\Delta m(\tilde{P})$, where $\Delta m(\tilde{P}) = m(\tilde{P}) - m(\tilde{P}-1)$
12: Decide $P$ in (6) and record the largest value of $\Delta m(\tilde{P})$ and $\alpha$ which is the second largest value of $\Delta m(\tilde{P})$
13: Calculate threshold $\gamma$, where $\gamma = \alpha + k$
14: If the largest value of $\Delta m(\tilde{P}) > \gamma$, declare $P$ in (6) as the original interleaving period
**Output**: Estimated interleaving period $P$

with the interleaving period $P$, binary phase shift keying modulation, and an additive white Gaussian noise (AWGN) channel. We conduct the simulations when the number of collected data, $M$, is equal to or less than $P \times P$ bits, and include the results of [19] for comparisons. Note that other conventional methods, such as [23] and [25]–[27], cannot estimate the interleaving period even when $M$ is $P \times P$ bits because they generally need more than $P \times P$ bits of collected data for estimation: method in [23] is based on an $L \times P$ rectangular matrix for $L > P$; and methods in [25]–[27] use the rank deficiency distribution for estimation, which requires a large number of $P \times P$ square matrices.

**TABLE 1.** Average rank deficiencies and their differences for $n$.

| $n$ | $m(P\text{-}1)$ | $m(P)$ | $\Delta m(P)$ |
|---|---|---|---|
| 2 | 4.54 | 8.10 | 3.56 |
| 3 | 2.33 | 8.21 | 5.88 |
| 4 | 1.95 | 8.04 | 6.09 |
| 5 | 1.49 | 8.02 | 6.53 |

Before examining the detection performance of the proposed method, we first consider $\Delta m(P)$, the value of $\Delta m(\tilde{P})$ when $\tilde{P} = P$, with regard to $n$, which is the number of vectors selected from $S_O$ for creating additional vectors of $S_C$. For this, we tabulate $\Delta m(P)$, $m(P)$, and $m(P-1)$ for $n = 2$, 3, 4, and 5 in Table 1, where $P$ is 28, $M$ is 784 ($28 \times 28$) bits, and (7, 4) Hamming code is used when the signal-to-noise

ratio (SNR) is 5 dB. Recall that, for the correct estimation of interleaving period $P$, the largest value of $\Delta m(\tilde{P})$ must occur when $\tilde{P} = P$.

From Table 1, we see that as $n$ increases, $m\ (P-1)$ decreases and $m(P)$ does not change significantly, and consequently, $\Delta m(P)$ increases as $n$ increases. The explanation of this follows from Section III: if $\tilde{P} \neq P$, e.g., in case of $m\ (P-1)$ in Table 1, though the linear dependence in a codeword is lost in the vectors of $S_N$, the rank deficiency may additionally occur due to the linearly dependent vectors included in $\mathbf{R}_N$, which is composed of $\tilde{P}$ different vectors selected from $S_N$. Therefore, as $n$ increases, $m\ (P-1)$ decreases, since the number of vectors in $S_N$ increases and the possibility that the linearly dependent vectors can be included in $\mathbf{R}_N$ decreases. On the other hand, if $\tilde{P} = P$, i.e., in case of $m(P)$ in Table 1, the rank deficiency occurs mainly due to the parity bits aligned on columns of $\mathbf{R}_N$ because the linear dependence in a codeword is maintained in the vectors of $S_N$. Accordingly, even if $n$ increases, $m(P)$ does not change significantly.

Consequently, as $n$ increases, the possibility that $\Delta m(P)$ becomes the largest value increases and we can expect that the detection performance can be improved.

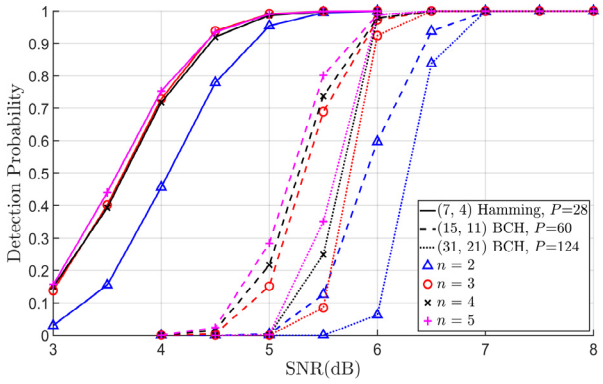



**FIGURE 3.** Detection probability for $n = 2, 3, 4$, and 5 in an AWGN channel.

Now, we examine the detection performance of the proposed method according to $n$. In Fig. 3, we show the detection probabilities by varying $n$ from 2 to 5, where we use the interleaving period 28 with (7, 4) Hamming code, interleaving period 60 with (15, 11) BCH code, and interleaving period 124 with (31, 21) BCH code when $M$ is $P \times P$ bits and $k$ is 1. From Fig. 3, we find that the detection performances improve as $n$ increases, as we expected. Performances improve significantly when $n$ varies from 2 to 3, and there are only relatively small improvements when $n$ increases from 3 to more.

In Table 1, we found that as $n$ increases, $m\ (P-1)$ decreases and $m(P)$ does not change significantly. For a given codeword length, $m(P)$ becomes smaller as the code rate increases because the number of parity bits in the codeword decreases as the code rate increases. Therefore, for a high code rate, to improve detection performance, i.e., to increase $\Delta m(P)$, we should take lager values of $n$ that can significantly decrease $m\ (P-1)$. We set $n$ to a design parameter, because the performance improvements can be noticeable when $n$ increases from 3 to more for higher code rates.

Under the same simulation conditions, to investigate the detection performance of the proposed method according to $k$, we depict the detection probabilities and the number of false alarms of the proposed method for $k = 0$, 0.5, and 1 when $n = 3$ in Figs. 4 and 5, respectively. From Figs. 4 and 5, we see that the detection probabilities decrease but false alarms also decrease as $k$ increases as we expected in Section III.

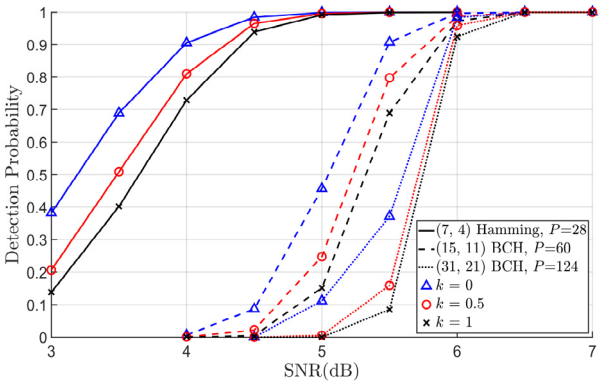



**FIGURE 4.** Detection probability for $k = 0$, 0.5, and 1 in an AWGN channel.

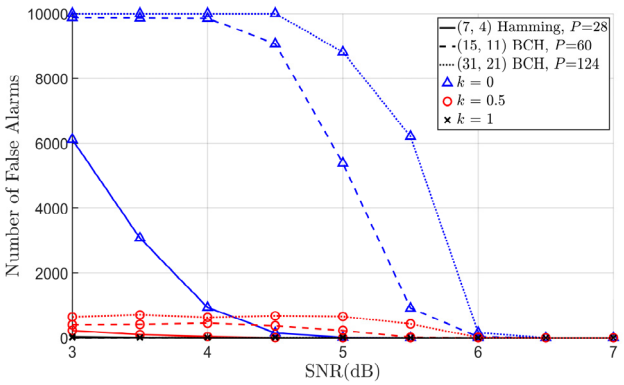



**FIGURE 5.** Number of false alarms in 10,000 iterations for $k = 0$, 0.5, and 1 in an AWGN channel.

To analyze the detection performance according to $M$, the number of collected data, which is the main concern of this paper, we show the detection probabilities for interleaving period of 28 with (7, 4) Hamming code, interleaving period of 60 with (15, 11) BCH code, and interleaving period of 124 with (31, 21) BCH code, in Figs. 6, 7, and 8, respectively. We also depict the number of false alarms in 10,000 iterations for the interleaving period of 124 with (31, 21) BCH code in Fig. 9. In the simulations, we vary $M$ from $P \times P$ to $P \times P \times \beta$ bits for $0 < \beta < 1$, and set $k$ to 1 and $n$ to 3. For comparisons, the performance for the conventional method of [19] is included.

When $M$ is $P \times P$ bits, at a detection probability of 0.9, the proposed method achieves SNR gains of about 2.0 dB,

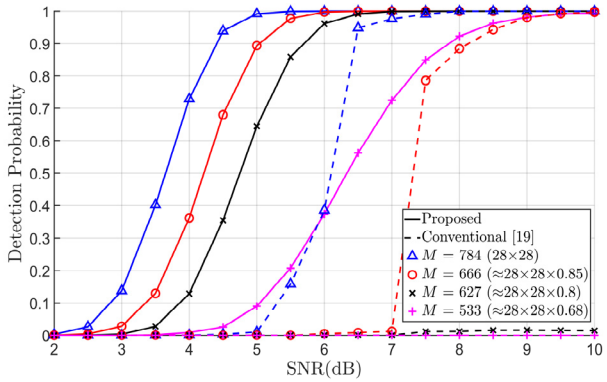



**FIGURE 6.** Detection probability for $M$ when interleaving period is 28 and (7, 4) Hamming code is used in an AWGN channel.

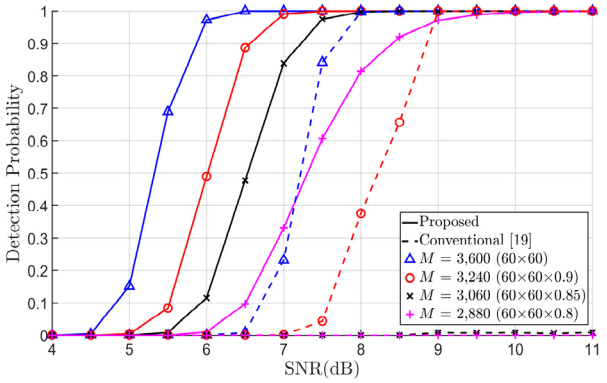



**FIGURE 7.** Detection probability for $M$ when interleaving period is 60 and (15, 11) BCH code is used in an AWGN channel.

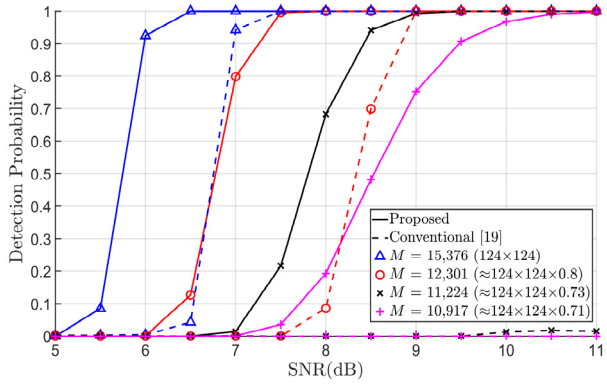



**FIGURE 8.** Detection probability for $M$ when interleaving period is 124 and (31, 21) BCH code is used in an AWGN channel.

1.8 dB, and 1.0 dB over [19] in Figs. 6, 7, and 8, respectively. We see that, with smaller values of $M$, the proposed method achieves greater SNR gains over [19]: at a detection probability of 0.9, about 3.1 dB, 2.3 dB, and 1.6 dB over [19] when $M$ are 666 ($\approx 28 \times 28 \times 0.85$) bits in Fig. 6, 3,240 ($60 \times 60 \times 0.9$) bits in Fig. 7, and 12,301 ($\approx 124 \times 124 \times 0.8$) bits in Fig. 8, respectively. In more severe cases of $M$, when $M$ are 627 ($\approx 28 \times 28 \times 0.8$) bits in Fig. 6, 3,060 ($60 \times 60 \times 0.85$) bits in Fig 7, and 11,224 ($\approx 124 \times 124 \times 0.73$) bits in Fig. 8, the detection probabilities of the proposed method can reach 0.9 at SNRs of 5.7 dB, 7.2 dB, and 8.4 dB respectively, whereas

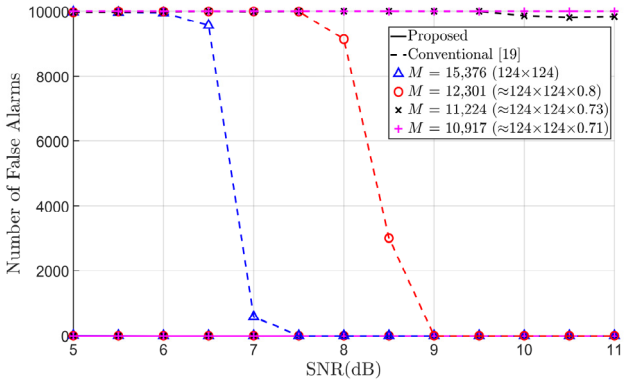



**FIGURE 9.** Number of false alarms in 10,000 iterations for $M$ when interleaving period is 124 and (31, 21) BCH code is used in an AWGN channel.

the method in [19] cannot give any meaningful detection results. From Fig. 9, we also see that the proposed method has few false alarms even when the number of collected data, $M$, is small.

From the results, we can expect that the proposed method can effectively estimate interleaver parameter even when only relatively few data are collected.

## V. CONCLUSIONS

In this paper, we presented a novel method for blind estimation of interleaver parameter when only a relatively small number of collected data are available. Under the condition of collecting a limited number of data that not even a single $P \times P$ square matrix can be generated, the previous estimation methods became infeasible or had degraded estimation performances. To solve this problem, we first created additional data by combining the collected data. Then, we investigated the rank deficiency of the matrices composed of the collected and additionally-created data. Finally, we estimated the interleaver parameter by using the difference of average rank deficiencies when only a limited number of data are available. We validated the proposed method through computer simulations, showing that the proposed method is applicable under the condition of collecting a limited number of data. Our method for the estimation of interleaver parameter with a limited number of collected data, however, is not restricted to interleaver parameter estimation. The proposed method can be applied to the blind estimation of other communication parameters, such as channel coding parameter, by using the linearity hidden inside the collected data.

## ACKNOWLEDGMENT

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

## REFERENCES

[1] Z. Zhu and A. K. Nandi, *Automatic Modulation Classification: Principles, Algorithms and Applications*. Hoboken, NJ, USA: Wiley, 2015.

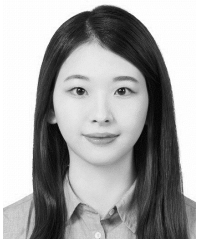
[2] S. Ahmadi, *5G NR: Architecture, Technology, Implementation, and Operation of 3GPP New Radio Standards*. Cambridge, MA, USA: Academic, 2019.

[3] S. Ye, S. Horng Wong, and C. Worrall, "Enhanced physical downlink control channel in LTE advanced release 11," *IEEE Commun. Mag.*, vol. 51, no. 2, pp. 82–89, Feb. 2013.

[4] C. Condo, S. A. Hashemi, A. Ardakani, F. Ercan, and W. J. Gross, "Design and implementation of a polar codes blind detection scheme," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 66, no. 6, pp. 943–947, Jun. 2019.
[5] A. Jalali and Z. Ding, "Joint detection and decoding of polar coded 5G control channels," *IEEE Trans. Wireless Commun.*, vol. 19, no. 3, pp. 2066–2078, Mar. 2020.
[6] O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "Survey of automatic modulation classification techniques: Classical approaches and new trends," *IET Commun.*, vol. 1, no. 2, pp. 137–156, Apr. 2007.
[7] W. S. Lin and K. J. R. Liu, "Modulation forensics for wireless digital communications," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Las Vegas, NV, USA, Mar. 2008, pp. 1789–1792.
[8] 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.
[9] 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.
[10] 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.
[11] 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.
[12] X.-B. Liu, S. N. Koh, X.-W. Wu, and C.-C. Chui, "Reconstructing a linear scrambler with improved detection capability and in the presence of noise," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 1, pp. 208–218, Feb. 2012.
[13] Y. Ma, L.-M. Zhang, and H.-T. Wang, "Reconstructing synchronous scrambler with robust detection capability in the presence of noise," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 2, pp. 397–408, Feb. 2015.
[14] 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.
[15] R. Moosavi and E. G. Larsson, "Fast blind recognition of channel codes," *IEEE Trans. Commun.*, vol. 62, no. 5, pp. 1393–1405, May 2014.
[16] A. Refaey, R. Niati, X. Wang, and J. Yves-Chouinard, "Blind detection approach for LDPC, convolutional, and turbo codes in non-noisy environment," in *Proc. IEEE Conf. Commun. Netw. Secur.*, San Francisco, CA, USA, Oct. 2014, pp. 502–503.
[17] 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.
[18] 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.
[19] G. Sicot, S. Houcke, and J. Barbier, "Blind detection of interleaver parameters," *Signal Process.*, vol. 89, no. 4, pp. 450–462, Apr. 2009.
[20] 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.
[21] 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.
[22] 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, 2017.
[23] 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.
[24] 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, 2019.
[25] C. Choi and D. Yoon, "Enhanced blind interleaver parameters estimation algorithm for noisy environment," *IEEE Access*, vol. 6, pp. 5910–5915, 2018.
[26] 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.
[27] 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, 2019.
[28] M. Tagliasacchi and S. Tubaro, "Blind estimation of the QP parameter in H.264/AVC decoded video," in *Proc. WIAMIS*, Desenzano del Garda, Italy, 2010, pp. 1–4.
[29] 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.
[30] B. Sklar, *Digital Communications: Fundamentals and Applications*. Upper Saddle River, NJ, USA: Prentice-Hall, 2001.
[31] G. Proakis, *Digital Communications*, 4th ed. New York, NY, USA: McGraw-Hill, 2001.
[32] L. I. Bluestein, "Interleaving of pseudorandom sequences for synchronization," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-4, no. 4, pp. 551–556, Jul. 1968.
[33] J. Ramsey, "Realization of optimum interleavers," *IEEE Trans. Inf. Theory*, vol. IT-16, no. 3, pp. 338–345, May 1970.
[34] 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.
[35] V. F. Kolchin, *Random Graphs*. New York, NY, USA: CUP, 1999.

**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 Electronics and Computer Engineering. His research interests are in 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 at the Signal Intelligence Research Center, Hanyang University. His research interests include channel coding, signal intelligence, and wireless communications.



**YOONJI KIM** received the B.S. degree in electronic engineering from Hanyang University, Seoul, South Korea, in 2019, where she is currently pursuing the M.S. degree with the Department of Electronics and Computer Engineering. Her research interests are in digital communication theory and wireless communications.

**DONGWEON YOON** (Senior Member, IEEE) 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 with the Faculty of Hanyang University, Seoul, where he is currently a Professor at 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 wireless communications.