Received February 26, 2020, accepted March 8, 2020, date of publication March 12, 2020, date of current version March 23, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2980321

Irregular Marker Codes for Insertion/ Deletion-AWGN Channels

YUAN LIU[1,2], YASHUO HE[2], XIAONAN ZHAO[1,2], MENG XIE[2], YI HONG[2], AND CUIPING ZHANG[1,2]

[1]Tianjin Key Laboratory of Wireless Mobile Communications and Power Transmission, Tianjin Normal University, Tianjin 300387, China
[2]College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387, China

Corresponding author: Yuan Liu (liuyuan@tjnu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61801327, in part by the Doctor Fund of Tianjin Normal University under Grant 043135202-XB1711, in part by the Natural Science Foundation of Tianjin City under Grant 18JCYBJC86400, and in part by the Tianjin Higher Education Creative Team Funds Program.

ABSTRACT A concatenated coding scheme employing an irregular marker code as the inner code is designed to improve the ability of correcting insertions/deletions. In this scheme, bits associated with each marker symbol are allocated to the symbol of the LDPC code non-uniformly. Since the non-binary marker symbol at the irregular position provides reliable forward/backward quantities, significant amount of insertions and deletions can be detected and corrected by the presented method. Simulation results show that the proposed scheme has an improved performance with only a very small penalty in coding rate compared with the traditional regular marker code.

INDEX TERMS Concatenated coding scheme, marker codes, insertions/deletions, forward-backward algorithm.

I. INTRODUCTION

Most error-correcting codes are constructed to correct random substitution errors, which cannot change the length of the received sequences. In this paper, random insertions/ deletions are considered in the channel model. For this type of errors, extra uncertain symbols will be inserted into the sequences and the information symbol may be deleted randomly due to the non-perfect synchronization [1]. Besides, this insertion/deletion channel model can also describe situations such as the segmented edit channels [2], [3], DNA sequencing/data storage [4], [5], bit-patterned media recording [6], [7], the synchronization with applications to speech watermarking [8], and errors in differential pulse-position modulation (DPPM) [9]. Insertions/deletions inevitably arise to the unknown boundaries and bring a great challenge for the design of the coding [10]–[16].

For the binary insertions/deletions, marker codes can be retraced to the early research designed by Sellers [17]. The best known concatenated code was presented by Davey and MacKay [18], which uses an inner watermark code to eliminate insertions/deletions in the received blocks and employs a low-density parity-check (LDPC) code to correct the residual errors [19]–[22]. The concatenated code could determine the block boundaries exactly which the receiver is not told. In order to further obtain an improved performance, a symbol-level watermark decoder considering the information codebook was designed [23]. In this method, the symbol-level forward-backward algorithm allows soft a priori input, and it is possible to use an iterative decoding scheme to achieve the significant improvements.

Yazdani devised an efficient construction for combating non-binary insertions/deletions and additive white Gaussian noise (AWGN) [24]. In this approach, the marker sequence is uniformly inserted into the LDPC code by the inner encoder. Using the binary inner encoding scheme, the concatenated code is then mapped into two different sets of the signal constellation. The receiver is not told the position of the block boundaries. The beginning of blocks is inferred according to the forward/backward quantities at the end of the last block. As insertion/deletion probability gets larger, the accumulated errors in the identification of the symbols increase, which causes that the start of consecutive blocks will be shifted at error positions. Therefore, synchronization and decoding become more difficult.

In order to improve the ability of the marker to correct non-binary insertions/deletions and the reliability of the soft information output by the marker decoder, an irregular marker

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

code is designed and the modified forward-backward algorithm is proposed in this paper. The known marker code is inserted into the sequences non-uniformly. Then, the mapping subsets in the signal constellation are dynamic according to the symbol position and the degree of marker. Since the minimum distance between points in each subset increases, the ability to detect insertions/deletions is developed and the reliability of the soft information received by LDPC decoder is improved. Furthermore, by varying the degree of the irregular marker, easier synchronization and stronger correction can be obtained with reducing the rate of the concatenated code.

The rest of this paper is organized as follows. Section II introduces the system and the channel models. Section III describes the proposed marker encoding method in detail, and shows the modified efficient forward-backward decoding scheme. Section IV illustrates the simulation results with the empirical performance of irregular marker, and compares the performance with previous work of the regular marker code. Finally, some conclusions are drawn in Section V.

II. BACKGROUND

In this section, the system and channel models are described before beginning the discussion.

A. SYSTEM MODEL

As shown in Figure 1, the information sequence $\underline{b}$ with the length of N_b bits is encoded by the binary LDPC encoder producing the codeword sequence $\underline{d}$ with the length of N_L bits, where LDPC code (N_b, N_L) with the rate of $R_L = N_b/N_L$ is used in this paper. The marker code $\underline{m}$ with the length of N bits is inserted into $\underline{d}$ non-uniformly according to the degree $\underline{w} = (w_1, w_2, \cdots, w_j, \cdots, w_{N_s})$, where $w_j \geq 1$. Therefore, this code is called as an irregular marker code. In particular, let $w_j = 2$ for $w_j \geq 1$ in this paper. Each symbol has λ bits including w_j bits of the marker and $(\lambda - w_j)$ bits of the LDPC code, where $2^\lambda = M$ and M is the size of the signal set. The length of the concatenated code $N_s = (N_L - \varepsilon'\beta)/\varepsilon + \beta$ symbols, and $N_c = \lambda \times [(N_L - \varepsilon'\beta)/\varepsilon + \beta]$ bits. β is the number of positions with $w_j > 1$. ε' is the number of LDPC bits allocated to each marker symbol with $w_j > 1$, and ε is the number of LDPC bits allocated to each marker symbol with $w_j = 1$. The symbol position associated with $w_j > 1$ can be regarded as the irregular position. We rewrite $N = (N_L - \varepsilon'\beta)/\varepsilon + 2\beta$.

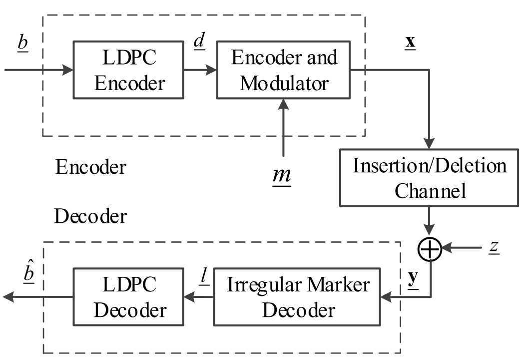


FIGURE 1. The system model using the irregular marker code.

Each symbol is then mapped into the signal point according to the signal constellations. In this paper, we take the system using 8-PSK as an example.

Transmitted and received words $\underline{\mathbf{x}} = (\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_{N_s-1})$ and $\underline{\mathbf{y}} = (\mathbf{y}_0, \mathbf{y}_1, \cdots)$, respectively. $\underline{\mathbf{x}}$ is corrupted by the noises and generating $\underline{\mathbf{y}}$, where the noises include insertions, deletions, and AWGN.

The irregular marker decoder outputs the soft information $\underline{l}$, i.e., log-likelihood ratios (LLRs) for the LDPC decoder by employing the forward-backward algorithm. In this step, the irregular marker decoder recovers the sequences and shifts the block boundaries to an estimated position. Then, the LDPC decoder produces the information estimation $\underline{\hat{b}}$.

B. CHANNEL MODEL

Let us consider three situations to illustrate the loss of synchronization, which is depicted in Figure 2. Extra symbols are inserted into the data sequence, which is named as *insertion*. For example, $\dot{\mathbf{y}}_1^+$ and $\dot{\mathbf{y}}_6^+$ are insertions in Figure 2. And symbols are removed from the data sequence, which is named as *deletion*. If there are no insertions and deletions at one symbol interval, $\mathbf{x}_i$ is transmitted correctly. By suffering from two types of errors above, the synchronization between the input $\underline{\mathbf{x}}$ and output $\underline{\dot{\mathbf{y}}}$ of the insertion/deletion channel is lost, and the symbol boundaries of $\underline{\dot{\mathbf{y}}}$ are unknown at the receiver. The length of $\underline{\dot{\mathbf{y}}}$ is denoted by N_s^*. $N_s^* = N_s + N_{\mathrm{I}} - N_{\mathrm{D}}$, where N_{I} is the number of insertions, and the N_{D} is the number of deletions occurred in one block, respectively. The maximum number of insertions is limited to I_{max} at each symbol interval.

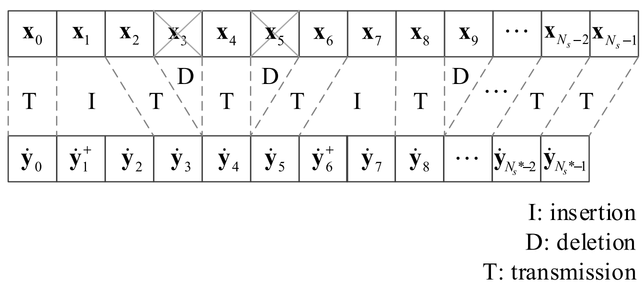


FIGURE 2. The length of the transmitted sequence may be changed by insertions and deletions.

The insertion/deletion-AWGN channel model is represented in Figure 3. Under this channel model, the i-th transmitted symbol $\mathbf{x}_i$ may be inserted with the insertion probability P_{i}, or deleted with the deletion probability P_{d}, or transmitted correctly with the transmission probability P_{t}. $P_{\mathrm{t}} = 1 - P_{\mathrm{i}} - P_{\mathrm{d}}$. The resulting sequence $\underline{\dot{\mathbf{y}}}$ is finally affected by $z \sim N(0, \sigma^2)$, and producing $\underline{\mathbf{y}} = \underline{\dot{\mathbf{y}}} + \underline{z}$ at the receiver.

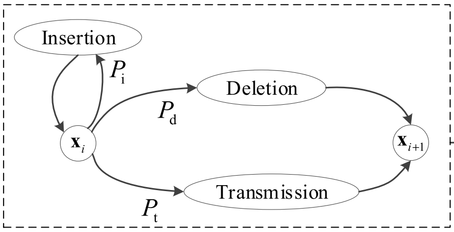


FIGURE 3. The channel model with insertions/deletions and AWGN.

III. PROPOSED ENCODING AND DECODING SCHEMES

A. IRREGULAR MARKER ENCODER

As is depicted in Figure 4, the inner marker code is embedded into the outer LDPC code non-uniformly according to the degree of the irregular marker code $\underline{w} = (w_1, w_2, \cdots, w_j, \cdots, w_{N_s})$. The value of w_j is varying with the symbol position j. The irregular position is recorded in the vector $\boldsymbol{\rho} = (\rho_0, \rho_1, \cdots, \rho_\theta, \cdots \rho_{\beta-1})$. If $j \in \boldsymbol{\rho}$, then $w_j > 1$. If $j \notin \boldsymbol{\rho}$, then $w_j = 1$.

According to $\underline{w}$, appropriate mappings η and η^* are chosen to design the modulator, which is shown in Figure 5. $\chi^{m_i} = \{\mathbf{x}|\mathbf{x} \in \chi; m_i\}$ denotes the subset in the constellation using η. Since the symbol has only one marker bit, χ includes two subsets χ^0 and χ^1 having $M/2$ signal points. Moreover, $\chi^{m_{i-1},m_i} = \{\mathbf{x}|\mathbf{x} \in \chi; m_{i-1}, m_i\}$ denotes the subset in the constellation using η^*, and $\chi = \chi^{00} \cup \chi^{10} \cup \chi^{01} \cup \chi^{11}$ includes four disjoint subsets. Each subset has $M/4$ signal points.

The minimum distance between the signal points in each subset is denoted as d_2, and the minimum distance between the signal points in different subsets is denoted as d_1. Notice that d_2 determines the ability of marker code to detect and correct insertions/deletions. By inserting more marker bits at the symbol interval, the mapping η^* increases d_2 compared with η (see Figure 5). Therefore, it is likely that significant improvements could be achieved with such an irregular marker technique.

Using this irregular marker encoding method, the rate of the concatenated code

$$R_c = \frac{N_b}{N_c} = \frac{N_b}{\lambda \times [(N_L - \varepsilon'\beta)/\varepsilon + \beta]}. \tag{1}$$

Rate R_c is reduced by increasing β, which making the marker code easier to identify insertions/deletions. For lower rate R_c, much higher probabilities of insertions, deletions can be tolerated.

B. FORWARD-BACKWARD DECODING SCHEME

Before discussing the algorithm of the modified forward-backward decoding scheme for the irregular marker code, we need to explain the process of the index mapping which plays a key role in the proposed system.

1) THE INDEX MAPPING

The irregular marker decoder receives $\mathbf{y}$, and compares the strings with the signals in the subsets to identify insertions/deletions and compute the drift probabilities. For the j-th symbol $\mathbf{x}_j$, it has w_j bits corresponding to the marker code and $(\lambda - w_j)$ bits of the information sequence. $0 \le j < N_s$. The signal subset associated with j-th symbol $\mathbf{x}_j$ is obtained according to $\boldsymbol{\rho} = (\rho_0, \rho_1, \cdots, \rho_\theta, \cdots \rho_{\beta-1})$, where $\boldsymbol{\rho}$ is sorted in the ascending order. Take the case in Figure 4 as an example, the signal subset

$$\begin{cases} \chi^{m_{i-1}m_i}, & w_j > 1 \; or \; j \in \boldsymbol{\rho} \\ \chi^{m_i}, & w_j = 1 or j \notin \boldsymbol{\rho}. \end{cases} \tag{2}$$

For the first situation in Eq. (2), if $w_j > 1$ or $j \in \boldsymbol{\rho}$,

$$i = j + \theta + 1, \quad \text{where } \rho_\theta = j. \tag{3}$$

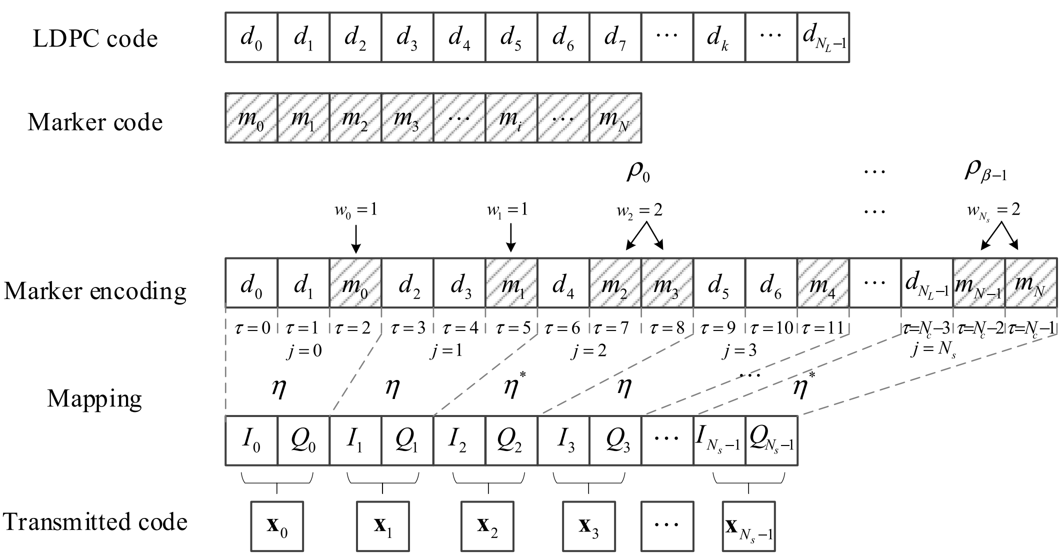
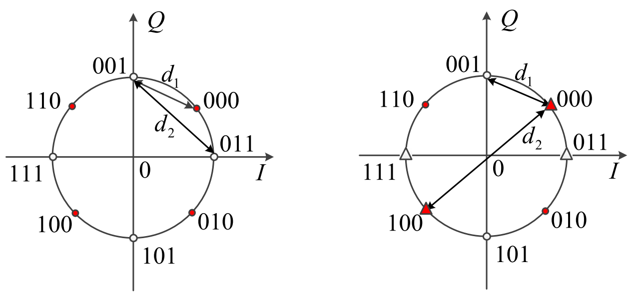


FIGURE 4. The encoding method of the irregular marker. $\varepsilon = 2$, $\varepsilon' = 1$, $\lambda = 3$.



(a) mapping η (b) mapping η^*

FIGURE 5. **8-PSK is taken for example in this paper. Signal constellations corresponds to symbols. (a) mapping η, where $\varepsilon = 2$, $w_j = 1$, the rightmost bit is the marker bit. (b) mapping η^*, $\varepsilon' = 1$, $w_j = 2$, the leftmost bit is the information bit. $M = 8$. $0 \le j < N_S$.**

Then, for the second situation, if $w_j = 1$ or $j \notin \boldsymbol{\rho}$,

$$\begin{cases} i = j, & j < \rho_0 \\ i = j + \theta, & j < \rho_\theta \\ i = j + \theta + 1, & otherwise, \end{cases} \tag{4}$$

where, $1 \le \theta < \beta, 0 \le i < N$.

The index mapping from the transmitted code to the LDPC code is described as follows. If $w_j > 1$ or $j \in \boldsymbol{\rho}$, the index of d_k

$$k = j \times \varepsilon - \theta, \quad \text{where } \rho_\theta = j. \tag{5}$$

If $w_j = 1 or j \notin \boldsymbol{\rho}$, the index of (d_{k-1}, d_k)

$$\begin{cases} k = j + 1, & j < \rho_0 \\ k = j - \theta + 1, & j < \rho_\theta \\ k = j - \theta, & otherwise, \end{cases} \tag{6}$$

where, $1 \le \theta < \beta, 0 \le k < N_L$.

2) CALCULATION FOR LLRs

The LLR at the k-th bit of the LDPC code is calculated as follows, which initializes the LDPC decoder to correct the remaining errors. $0 \le k < N_L$. Since $P(d_k = 0) = P(d_k = 1)$,

$$\begin{aligned} l_k &= \ln \frac{P(d_k = 0|\underline{\mathbf{y}}, \underline{w})}{P(d_k = 1|\underline{\mathbf{y}}, \underline{w})} \\ &= \ln \frac{P(\underline{\mathbf{y}}|d_k = 0, \underline{w})}{P(\underline{\mathbf{y}}|d_k = 1, \underline{w})}, \end{aligned} \tag{7}$$

where,

$$P(\underline{\mathbf{y}}|d_k = 0, \underline{w}) = \begin{cases} \sum\limits_{\mathbf{x}_j \in \chi^{m_{i-1}m_i}(d_k=0)} P(\underline{\mathbf{y}}|\mathbf{x}_j, \underline{w}), & w_j > 1 \\ \sum\limits_{\mathbf{x}_j \in \chi^{m_i}(d_k=0)} P(\underline{\mathbf{y}}|\mathbf{x}_j, \underline{w}), & w_j = 1 \end{cases} \tag{8}$$

The conditional probability $P(\underline{\mathbf{y}}|\mathbf{x}_j, \underline{w})$ is obtained by multiplying the forward and backward quantities by the middle quantity.

$$P(\underline{\mathbf{y}}|\mathbf{x}_j, \underline{w}) = \sum_{a,b} F_j(t_j = a) M^j_{ab}(\tilde{\mathbf{y}}|\mathbf{x}_j) B_{j+1}(t_{j+1} = b), \tag{9}$$

where the forward quantity $F_j(t_j = a)$ is the probability that the drift at the j-th time is a and that the first symbols generated by the channel is $\mathbf{y}_0, \cdots, \mathbf{y}_{j-1+a}$. $\tilde{\mathbf{y}} = (\mathbf{y}_{j+a}, \cdots, \mathbf{y}_{j+b})$. In this paper, the drift and maximum number of drifts at each time unit are denoted as t_j and $t_{\max}$, respectively. t_j also represents the state at j-th symbol in the decoding trellis. $F_j(t_j = a)$ is computed recursively as

$$F_j(t_j = a) = \sum_{c \in \{a - I_{\max}, \cdots, a+1\}} F_{j-1}(t_{j-1} = c) \times P_{ca} Q^{j-1}_{ca}(\mathbf{y}_{j-1+c}, \cdots, \mathbf{y}_{j-1+a}), \tag{10}$$

The calculation of the backward quantities are done similarly.

$$B_j(t_j = b) = \sum_{c \in \{b-1, \cdots, b+I_{\max}\}} B_{j+1}(t_{j+1} = c) \times P_{bc} Q^j_{bc}(\mathbf{y}_{j+b}, \cdots, \mathbf{y}_{j+c}). \tag{11}$$

In Eq. (10),

$$\begin{aligned} &P_{ca} Q^{j-1}_{ca} \\ &= P(\mathbf{y}_{j-1+c}, \cdots, \mathbf{y}_{j-1+a}, t_j = a | t_{j-1} = c, w_{j-1}) \\ &= \alpha_I P_{\mathrm{i}}^{a-c+1} P_{\mathrm{d}} \prod_{\xi = j-1+c}^{j-1+a} \gamma_\xi + \alpha_I P_{\mathrm{i}}^{a-c} P_{\mathrm{t}} \vartheta_{j-1+a} \prod_{\xi = j-1+c}^{j-1+a-1} \gamma_\xi. \end{aligned} \tag{12}$$

where, $\alpha_I = 1/(1 - P_{\mathrm{i}}^{I_{\max}})$,

$$\gamma_\xi = \frac{1}{M} \sum_{\mathbf{x} \in \chi} \frac{1}{2\pi\sigma^2} \exp\left(-\frac{|\mathbf{y}_\xi - \mathbf{x}|^2}{2\sigma^2}\right), \tag{13}$$

and

$$\vartheta_{j-1+a} = \begin{cases} \dfrac{4}{M} \sum\limits_{\mathbf{x} \in \chi^{m_{i-1}m_i}} \dfrac{1}{2\pi\sigma^2} \exp\left(-\dfrac{|\mathbf{y}_{j-1+a} - \mathbf{x}|^2}{2\sigma^2}\right), \\ \qquad w_{j-1} > 1 \\ \dfrac{2}{M} \sum\limits_{\mathbf{x} \in \chi^{m_i}} \dfrac{1}{2\pi\sigma^2} \exp\left(-\dfrac{|\mathbf{y}_{j-1+a} - \mathbf{x}|^2}{2\sigma^2}\right), \\ \qquad w_{j-1} = 1 \end{cases} \tag{14}$$

As shown in Eq. (14), for $w_{j-1} > 1$, more bits in a symbol are known to the receiver compared with that for $w_{j-1} = 1$. Thus, the better performance can be expected to be obtained owing to the larger Euclidean distance.

The middle quantity

$$\begin{aligned} M^j_{ab}(\tilde{\mathbf{y}}|\mathbf{x}_j) &= P(\tilde{\mathbf{y}}, t_{j+1} = b | t_j = a, \mathbf{x}_j) \\ &= \alpha_I P_{\mathrm{i}}^{b-a+1} P_{\mathrm{d}} \prod_{\xi = j+a}^{j+b} \gamma_\xi + \alpha_I P_{\mathrm{i}}^{b-a} P_{\mathrm{t}} \dot{\vartheta}_{j+b} \prod_{\xi = j+a}^{j+b-1} \gamma_\xi. \end{aligned} \tag{15}$$

where

$$\dot{\vartheta}_{j+b} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{|\mathbf{y}_{j+b} - \mathbf{x}_j|^2}{2\sigma^2}\right). \quad (16)$$

Finally, the LDPC decoder is initialized by $\underline{l}$ and produces the estimation of the information sequence.

IV. SIMULATION RESULTS

In this section, some examples were used to evaluate the performance of the proposed scheme using the irregular marker code. Table 1 shows the parameters of all outer codes used in the concatenated codes, and Table 2 gives the parameters of concatenated codes whose performance are reported in this paper. Let's take 8-PSK for example. A binary pseudorandom marker vector was created. $I_{max} = 5$. $P_{\rm i} = P_{\rm d}$. The signal-to-noise ratio (SNR) was set to 20 dB. A belief-propagation algorithm in log-domain was used in the LDPC decoder and the maximum number of iterations was set to 20. Since the drift was set to zero at the start, the forward quantities for 0-th symbol in the first block are calculated as follows.

$$F_0(t) = \begin{cases} 1, & t = 0, \\ 0, & \text{otherwise.} \end{cases} \quad (17)$$

TABLE 1. Parameters of outer codes used in the concatenated codes.

Concatenated Code	Construction of the outer code	N_L (bits)	R_L
A	Irregular quasi-cyclic LDPC code [25]	576	1/2
B	Irregular quasi-cyclic LDPC code [25]	2304	1/2
C	Regular LDPC code constructed by the progressive edge growth algorithm [26]	4000	1/2

TABLE 2. Parameters of concatenated codes reported in this paper.

Concatenated Code	N (bits)	N_c (bits)	R_c
A	$288+3\beta/2$	$864+3\beta/2$	$288/(864+3\beta/2)$
B	$1152+3\beta/2$	$3456+3\beta/2$	$1152/(3456+3\beta/2)$
C	$2000+3\beta/2$	$6000+3\beta/2$	$2000/(6000+3\beta/2)$

The backward quantities in each block are initialized with equal probabilities since the accurate block boundary is unknown at the receiver.

Firstly, we demonstrate the efficiency of the proposed irregular marker code. Results are given for several settings of β in Figures 6, 7 and 8. In each figure, the performance of codes with various length of markers are compared. As depicted in figures, the proposed scheme performs substantially better than the regular version, i.e., $\beta = 0$. It is also shown that the performance of the code using the irregular marker code is improved with the increasing of β. The reasons are given as follows. By increasing the number of the irregular positions, the length of the marker

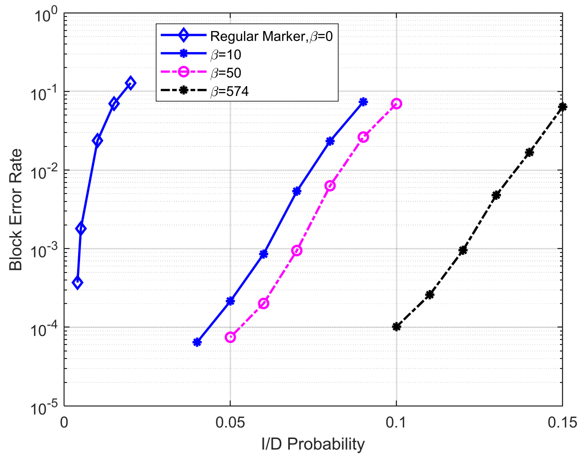


FIGURE 6. Performance of Code A.

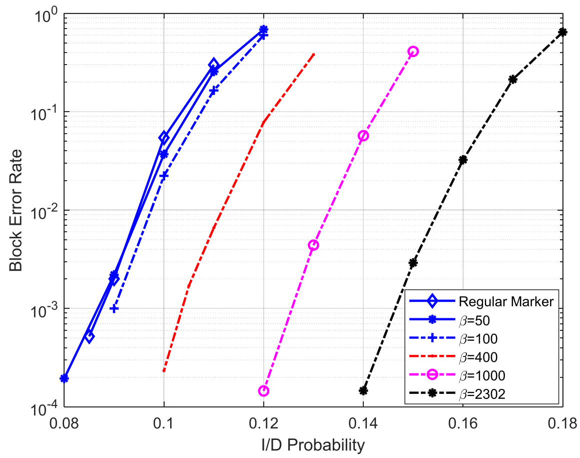


FIGURE 7. Performance of Code B.

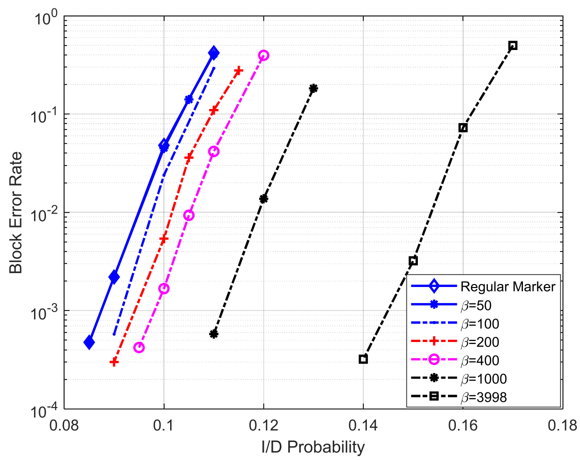


FIGURE 8. Performance of Code C.

code increases, making it easier to identify and correct insertions/deletions. Furthermore, residual errors could be eliminated by the LDPC decoder since that LLRs output by the marker decoder are more accurate. Notice that it is also a regular marker code if $\beta = N_L$. In this situation, the inner code can be regarded as a non-binary marker code. Nevertheless, it will bring a huge rate loss. A better balance between the performance and the coding rate can be achieved by selecting β in the practice.

Then, the rate of the proposed concatenated codes with various β is discussed as follows. According to Eq. (1), as β increases, the rate R_c decreases, which are shown in Table 3. In particular, $\beta = 0$ which corresponds to the regular non-binary marker coder. As observed from Table 3, performance gains can be obtained with very slight rate loss by selecting an appropriate β.

TABLE 3. **Rate of concatenated codes reported in this paper.**

Concatenated Code	β	R_c
	0	0.333
A	10	0.328
	574	0.167
	0	0.333
B	100	0.319
	2302	0.167
	0	0.333
C	100	0.325
	3998	0.167

When decoding marker codes, it is necessary to recover synchronization. The ability of marker code to estimate the drift positions of block boundaries has a great effect on the performance of the concatenated decoder. Figure 9 compares the accuracy of code C detecting boundaries for $\beta = 400$ with that of the regular marker code. As shown in the figure, improvements could be achieved by using the irregular constructions for the marker codes. Results show that the irregular marker code is extremely effective at identifying the block boundaries and recovering the synchronization over channels with insertions and deletions.

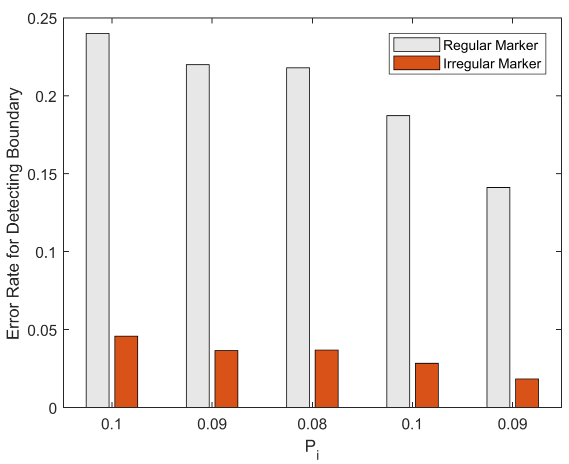


FIGURE 9. **The comparison of the ability to detecting boundaries. Code C, $\beta = 400$.**

V. CONCLUSION

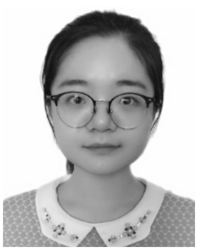
In this paper, an irregular marker code for correcting non-binary insertions/deletions and AWGN was presented, and the modified forward-backward decoding algorithm was designed. Each symbol has unequal number of marker bits while keeping the number of symbols unchanged in this scheme. By embedding marker code non-uniformly, the minimum distances among the symbols in the subset increase at the irregular positions, which provides improvements in recovering the synchronization for large insertion/deletion probabilities. The performance of the scheme using the irregular marker code was evaluated. Results show that significant amount of errors could be corrected by this proposed scheme with a very slight coding rate loss.

REFERENCES

[1] H. Mercier, V. Bhargava, and V. Tarokh, "A survey of error-correcting codes for channels with symbol synchronization errors," *IEEE Commun. Surveys Tuts.*, vol. 12, no. 1, pp. 87–96, 1st Quart., 2010.
[2] Q. Wang, S. Jaggi, M. Medard, V. R. Cadambe, and M. Schwartz, "File updates under Random/Arbitrary insertions and deletions," *IEEE Trans. Inf. Theory*, vol. 63, no. 10, pp. 6487–6513, Oct. 2017.
[3] Q. Wang, M. Medard, and M. Skoglund, "Efficient compression algorithm for file updates under random insertions and deletions," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Cambridge, U.K., Sep. 2016, pp. 434–438.
[4] A. Lenz, P. H. Siegel, A. Wachter-Zeh, and E. Yaakobi, "Coding over sets for DNA storage," in *Proc. IEEE Int. Symp. Inf. Theory*, Vail, CO, USA, Dec. 2018, pp. 1–5.
[5] C. Schoeny, A. Wachter-Zeh, R. Gabry, and E. Yaakobi, "Codes correcting a burst of deletion or insertions," *IEEE Trans. Inf. Theory*, vol. 63, no. 4, pp. 1–19, 4th Quart., 2017.
[6] X. Jiao and M. A. Armand, "Soft-input inner decoder for the davey-MacKay construction," *IEEE Commun. Lett.*, vol. 16, no. 5, pp. 722–725, May 2012.
[7] T. Wu and M. A. Armand, "Marker codes on BPMR write channel with data-dependent written-in errors," *IEEE Trans. Magn.*, vol. 51, no. 8, pp. 1–7, Aug. 2015.
[8] D. J. Coumou and G. Sharma, "Insertion, deletion codes with feature-based embedding: A new paradigm for watermark synchronization with applications to speech watermarking," *IEEE Trans. Inf. Forensics Secur.*, vol. 3, no. 2, pp. 153–165, Jun. 2008.
[9] W. Chen and Y. Liu, "Efficient transmission schemes for correcting insertions/deletions in DPPM," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kuala Lumpur, Malaysia, May 2016, pp. 1–6.
[10] C. Schoeny, A. Wachter-Zeh, R. Gabrys, and E. Yaakobi, "Codes correcting a burst of deletions or insertions," *IEEE Trans. Inf. Theory*, vol. 63, no. 4, pp. 1971–1985, Apr. 2017.
[11] J. Ullman, "Near-optimal, single-synchronization-error-correcting code," *IEEE Trans. Inf. Theory*, vol. IT-12, no. 4, pp. 418–424, Oct. 1966.
[12] A. S. J. Helberg and H. C. Ferreira, "On multiple insertion/deletion correcting codes," *IEEE Trans. Inf. Theory*, vol. 48, no. 1, pp. 305–308, 2002.
[13] T. G. Swart and H. C. Ferreira, "A note on double insertion/deletion correcting codes," *IEEE Trans. Inf. Theory*, vol. 49, no. 1, pp. 269–273, Jan. 2003.
[14] M. Mansour and A. Tewfik, "Convolutional decoding in the presence of synchronization errors," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 2, pp. 218–227, Feb. 2010.
[15] V. Buttigieg and N. Farrugia, "Improved bit error rate performance of convolutional codes with synchronization errors," in *Proc. IEEE Int. Conf. Commun. (ICC)*, London, U.K., Jun. 2015, pp. 4077–4082.
[16] M. F. Mansour and A. H. Tewfik, "A turbo coding scheme for channels with synchronization errors," *IEEE Trans. Commun.*, vol. 60, no. 8, pp. 2091–2100, Aug. 2012.
[17] F. Sellers, "Bit loss and gain correction code," *IEEE Trans. Inf. Theory*, vol. 8, no. 1, pp. 35–38, Jan. 1962.
[18] M. C. Davey and D. J. C. Mackay, "Reliable communication over channels with insertions, deletions, and substitutions," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 687–698, 2001.
[19] E. A. Ratzer, "Marker codes for channels with insertions and deletions," *Ann. Telecommun.*, vol. 60, no. 1, pp. 29–44, Feb. 2005.
[20] T. Wu and M. A. Armand, "The davey-MacKay coding scheme for channels with dependent insertion, deletion, and substitution errors," *IEEE Trans. Magn.*, vol. 49, no. 1, pp. 489–495, Jan. 2013.

[21] F. Wang, D. Fertonani, and T. M. Duman, "Symbol-level synchronization and LDPC code design for insertion/deletion channels," *IEEE Trans. Commun.*, vol. 59, no. 5, pp. 1287–1297, May 2011.
[22] X. Jiao and M. A. Armand, "Interleaved LDPC codes, reduced-complexity inner decoder and an iterative decoder for the Davey-MacKay construction," in *Proc. IEEE Int. Symp. Inf. Theory*, Saint Petersburg, Russia, Jul. 2011, pp. 742–746.
[23] J. A. Briffa, H. G. Schaathun, and S. Wesemeyer, "An improved decoding algorithm for the davey-MacKay construction," in *Proc. IEEE Int. Conf. Commun.*, Cape Town, South Africa, May 2010, pp. 1–5.
[24] R. Yazdani and M. Ardakani, "Reliable communication over non-binary Insertion/Deletion channels," *IEEE Trans. Commun.*, vol. 60, no. 12, pp. 3597–3608, Dec. 2012.
[25] *IEEE Standard for Local and Metropolitan Area Networks—Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems—Amendment for Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands*, IEEE Standard 802.16e-2005, IEEE Computer Society and the IEEE Microwave Theory and Techniques Society, New York, NY, USA, Feb. 2006.
[26] X.-Y. Hu, E. Eleftheriou, and D. M. Arnold, "Regular and irregular progressive edge-growth tanner graphs," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 386–398, Jan. 2005.

YUAN LIU received the B.E. degree from Tianjin Polytechnic University, Tianjin, China, in 2010, and the Ph.D. degree from Tianjin University, Tianjin, in 2017. She is currently a Lecturer with the School of Electronic and Communication Engineering, Tianjin Normal University. Her research interests include communication theory and synchronization error correcting scheme.



YASHUO HE is currently pursuing the bachelor's degree with the College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin, China. She is participating in the Excellent Students' Training Project and the Future Engineer's Training Project of Tianjin Normal University. Her current research interests include synchronization error correcting schemes and wireless electricity transmission.

XIAONAN ZHAO received the Ph.D. degree from Tianjin University, Tianjin, in 2015. He is currently working with the College of Electronic and Communication Engineering, Tianjin Normal University. His research interest includes wireless communication channel measurement and modeling.

MENG XIE is currently pursuing the bachelor's degree with the College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin, China. She is participating in the Future Engineer's Training Project of Tianjin Normal University. Her current research interests include wireless communications and synchronization error correcting schemes.

YI HONG is currently pursuing the bachelor's degree with the College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin, China. She is participating in the Future Engineer's Training Project of Tianjin Normal University. Her current research interests include the information theory and synchronization error correcting schemes.

CUIPING ZHANG received the M.S. degree from Tianjin Normal University, Tianjin, China, in 2018. She is currently working with the College of Electronic and Communication Engineering, Tianjin Normal University. Her main research interests include wireless sensor networks and image processing.