Rate-Compatible Shortened Polar Codes Based on RM Code-Aided

Chunjie Li[1,2], Haiqiang Chen[1,2](✉), Zelin Wang[1,2], Youming Sun[1,2], and Xiangcheng Li[1,2]

[1] School of Computer, Electronics and Information, Guangxi University, Nanning 530004, People's Republic of China
haiqiang@gxu.edu.cn

[2] Guangxi Key Laboratory of Multimedia Communications and Network Technology, Guangxi University, Nanning 530004, People's Republic of China

Abstract. The minimum Hamming distance is not considered for the traditional rate-compatible shortened polar (RCSP) codes, which may cause performance degradations. In this paper we propose a hybrid algorithm to construct RCSP codes based on Reed-Muller (RM) code-aided. The shortened bits and pre-frozen bits are jointly designed by the row weight property of the common generator matrix G_N for the RM/Polar code. First, the selected shortened bits are guaranteed to be uniquely depended upon the pre-frozen bits, which makes them completely be known by the decoder. Second, the proposed construction method is designed in such way, so that the minimum row weight of G_N can be maximized. More specifically, when multiple candidate positions satisfy the conditions (weight-1 column constraint), those rows having less weights are deleted to form the shortened/pre-frozen bits, which can reduce the number of rows with small weight and naturally, make the resulting RCSP codes have larger minimum Hamming distance in average. Simulation results show that the proposed RCSP codes perform better than the traditional shortened codes at low code rates. While at high code rates, the proposed RCSP codes can achieve better performance than that of the quasi uniform punctured (QUP) polar codes, especially at large signal-to-noise ratio (SNR) region. The proposed RCSP codes can find applications in future communications, such as the beyond 5th generation (B5G) and 6th generation (6G) systems.

Keywords: Polar codes · Rate-compatible · Reed-Muller codes · Hamming distance · Shortening

1 Introduction

Polar codes are the first family of codes which have been proven to achieve the capacity of any symmetric binary-input discrete memoryless channel (B-DMC) [1]. Polar codes have good structural characteristics and low encoding

This work was supported by NSF of China (No. 61761006, 61961004) and NSF of Guangxi (No. 2017GXNSFAA198263, 2017GXNSFAA198276, 2018GXNSFAA138079).

W. Xiang et al. (Eds.): BROADNETS 2021, LNICST 413, pp. 307–319, 2022.
https://doi.org/10.1007/978-3-030-93479-8_22

and decoding complexity. At the end of 2016, polar codes were selected as the candidate coding scheme for the 5G mobile communications, and were finally adopted as the coding standard for the uplink/downlink channel control [2].

However, the length of polar codes are limited to powers of 2 due to the original Kronecker power construction, which restricts their flexible applications in practice. Polar codes with arbitrary lengths and rates can be mainly obtained by puncturing, shortening and repetition, resulting the rate-compatible polar codes. These rate-compatible schemes are also recommended in 5G NR [3].

Punctured polar codes are first proposed in [4], where random puncturing and stopping-tree puncturing were both analysed and compared. Niu *et al.* proposed an efficient puncturing scheme [5], in which the puncturing positions are designed to be quasi-uniform distribution after bit-reversal permutation, thus called the quasi-uniform puncturing (QUP). The QUP has better row weight property than random puncturing and can achieve excellent decoding performances, especially at low code rates. The traditional shortening scheme was discussed in [6], where a simple shortening method was given. The last N_p coded bits, whose values are completely determined by the pre-frozen bits, are shortened to form the rate-compatible shortened polar (RCSP) codes.

It is shown that the minimum Hamming distance has a significant impact on error performance of polar codes. Thus, constructing polar codes with large minimum Hamming distance to improve decoding performance becomes possible. Similar work can be seen in [7], where the minimum Hamming distance is increased by joint optimization of the shortening pattern and the set of frozen symbols. Li *et al.* proposed the RM-Polar codes, which have much better distance property than polar codes and thus show better performance [8].

However, neither of the QUP scheme and the traditional shortening scheme consider the distance property in their construction and this may cause potential performance degradations. Actually, polar codes can be seen as a generalization of RM codes [9,10] and they share a common generator matrix [11]. This enable us to jointly optimize the distance property of the RCSP codes under row weight constraint of RM codes. Motivated by this, we propose a hybrid algorithm to construct the rate-compatible shortened polar codes by combining with the RM-rule constraint. The proposed algorithm maintains the superiority of the traditional shortened polar codes, *i.e.*, the shortened bits are designed to be completely known by the decoder thus the corresponding log-likelihood ratios (LLRs) can be set to infinity (or minus infinity) to ensure the decoding performance. To further improve the performance, a distance-greedy construction method is proposed to maximize the minimum row weight of G_N.

More specifically, those rows having less weights are first deleted during code construction to form the shortened/frozen bits, which can reduce the number of rows with small weight. Consequently, the constructed RCSP codes will have larger minimum Hamming distance in average. Simulation results show that the proposed RCSP codes have better frame error rate (FER) performance than the traditional shortened polar codes at low code rate. While at high code rate, the proposed RCSP codes outperform the QUP codes, especially at large signal-to-noise ratio (SNR) region.

The rest of this paper is organized as follows. In Sect. 2, we provide a short background on polar codes and RM codes, and introduce the RCSP codes construction system model. Section 3 gives a brief introduction to the traditional shortening method, then propose the RM code-aided hybrid RCSP algorithm. Section 4 gives the row property analysis of generator matrix and the simulation results. Section 5 concludes the paper.

2 Background

2.1 RM Codes

This subsection uses the Kronecker construction method to describe RM codes. Since RM code is a linear block code, which can be constructed by a generator matrix. Let $RM(n, n)$ denote the nth order RM code, and let $\mathbf{G}_N$ be the N-dimension generator matrix with $N = 2^n$, which can be defined as

$$\mathbf{G}_N = \mathbf{F}^{\otimes n} \tag{1}$$

where $\mathbf{F} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, $\mathbf{F}^{\otimes n}$ is the nth Kronecker power of $\mathbf{F}$. The rth order RM code $RM(n, r)$ can then be defined as the linear code with a sub-matrix of $\mathbf{G}_N$, which is obtained by selecting rows of $\mathbf{G}_N$ with Hamming weights $\geq 2^{n-r}$.

The row weight of the generator matrix $\mathbf{G}_N$ has the following constraint with the row index. Let i denote an integer, $i \in \{0, 1, \cdots, N-1\}$, and $\pi(i) = (b_{n-1}b_{n-2}\cdots b_1b_0)$ is the binary representation of i over n bits. Let $w_t(i)$ represent the Hamming weight of $\pi(i)$. The Hamming weight of ith row can be calculated by $w_r(i) = 2^{w_t(i)}$.

Since the RM code is a linear code, each row of the generator matrix can be regarded as a legal codeword. Therefore, the minimum row weight of the generator matrix corresponds to the minimum Hamming distance of the RM code. Actually, an RM code is equivalent to a special polar code which has the maximum row weight constraint. For example, an rth order RM code $RM(n, r)$ is equivalently a polar code with the frozen set $\mathcal{A}^c$ that satisfies the distance constraint $\mathcal{A}^c = \{i | w_r(i) < 2^{n-r}\}$. With this constraint, the minimum Hamming distance of the polar code is $d_{min} = \min\{w_r(i) | i \in \mathcal{A}\}$, where $\mathcal{A}$ is the complementary set of $\mathcal{A}^c$, called the information set.

2.2 Polar Codes

Given a B-DMC $W : \mathcal{X} \rightarrow \mathcal{Y}$, where $\mathcal{X} \in \{0, 1\}$ and $\mathcal{Y}$ denote the input and output alphabet, respectively. The channel transition probabilities can be defined as $W(y|x)$, $y \in \mathcal{Y}$, $x \in \mathcal{X}$. Let a_0^{N-1} denote a row vector $(a_0 \cdots a_{N-1})$, and $a_i^j = (a_i \cdots a_j)$ denote a subvector, $0 \leq i \leq j \leq N-1$. After channel combining and splitting operation on N independent uses of W, we get N successive uses of synthesized binary input channels $W_N^{(i)}$, $i \in \{0, 1, \cdots, N-1\}$, which can be defined by the transition probabilities as follows:

$$W_N^{(i)}(y_0^{N-1}, u_0^{i-1}|u_i) = \sum_{u_{i+1}^{N-1} \in \mathcal{X}^{N-i-1}} \frac{1}{2^{N-1}} W_N(y_0^{N-1}|u_0^{N-1}). \quad (2)$$

The N independent subchannels can be divided into two parts. One part of channels with capacity tends to be 1, called "noiseless channel", and the other part of channels with capacity tends to be 0, called "full noise channel". The reliability of each subchannel can be computed by using the Bhattacharyya parameter [1], density evolution (DE) [12], Gaussian approximation (GA) [13] or polarization weight [14]. The K most reliable subchannels with indices in $\mathcal{A}$ carry information bits and the rest subchannels in $\mathcal{A}^c$ are set to be fixed values, such as all zeros. For an (N, K) polar code with K message bits and N coded bits, the encoding process can be defined as

$$c_0^{N-1} = u_0^{N-1}\mathbf{G}_N, \quad (3)$$

where $u_0^{N-1} = (u_0, u_1, \cdots, u_{N-1})$ is the source information vector and $c_0^{N-1} = (c_0, c_1, \cdots, c_{N-1})$ is the polar codeword.

As mentioned above, polar codes can be seen as a generalization of RM codes and both of them are defined by the same generator matrix $\mathbf{G}_N$. However, they select the information bits according to different constraints. In particular, the Hamming distance is considered in the RM codes construction, which can be exploited to optimize the proposed RCSP codes in this paper.

2.3 System Model

The system model of the proposed RCSP codes construction is depicted in Fig. 1. In the transmitter, a K-bit information block is input into the polar encoder. After polar encoding, we get the N-bit polar codeword. To match arbitrary code length, the output polar code needs to be adjusted by shortening some bits from the N-bit encoded block, resulting in the M-bit RCSP codes. Then the RCSP codes with length-M is fed into the channel. In the receiver, we perform the opposite operation to get the corresponding estimated bits. Note that, the proposed RCSP codes are jointly designed with the encoding unit. The row weight constraint of RM code is employed to maximize the Hamming distance in the code construction. This is quite different from the traditional shortening scheme.

3 RM Code-Aided RCSP Codes

3.1 The Shortening Construction

Let $\mathbf{g}_j$ denote jth column vector of the generator matrix $\mathbf{G}_N$, where $j = 0, 1, \cdots, N-1$. Let $Q(\mathbf{g}_j)$ denote the index set of the "1" positions in $\mathbf{g}_j$. The vector $\mathbf{p} = (p_0, p_1, \cdots, p_{N-1})$ is the shortening pattern with $p_i \in \{0, 1\}$ and the

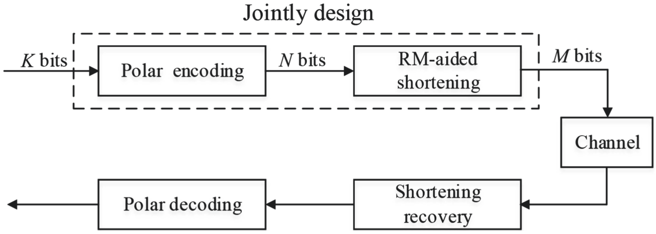


Fig. 1. The system model.

index $i = 0, 1, \cdots, N-1$, where the 1s imply the shortened positions. The index set of shortened positions can be represented as $Q(\mathbf{p})$. Let c_j be a code bit of c_0^{N-1}, which can be defined as follows:

$$c_j = \sum_{i \in Q(\mathbf{g}_j)} \oplus u_i. \tag{4}$$

Assume that c_j is selected as a shortened bit, then all the elements $i \in Q(\mathbf{g}_j)$ are designated to be the frozen bits. This is the key step to ensure that c_j is completely determined by the frozen bits, thus is known by the decoder. Define the pre-frozen set $\mathcal{A}_S^c$ as follows:

$$\mathcal{A}_S^c = \bigcup_{j \in Q(\mathbf{p})} Q(\mathbf{g}_j). \tag{5}$$

Note that the frozen positions in $\mathcal{A}_S^c$ are only determined by the shortening pattern, but not the sub-channel reliabilities. In order to minimize the number of pre-frozen bits, *i.e.*, the cardinality of $\mathcal{A}_S^c$, the *weight-1 first* criterion is introduced in [6]. Equivalently, the following equation should be satisfied,

$$|Q(\mathbf{g}_j)| = 1, \tag{6}$$

where $j \in Q(\mathbf{p})$. This implies that the shortening positions are always selected from the index of columns with weight-1. With this constraint, the number of shortened bits is exactly the pre-frozen bits, *i.e.*, $N_p = |\mathcal{A}_S^c|$. The pre-frozen set $|\mathcal{A}_S^c|$ with minimum cardinality can be determined by the following construction algorithm by N_p step. Let $\mathcal{A}_S^{c(k)}$ be the temporary set at the k-step, with $\mathcal{A}_S^{c(0)} = \emptyset$. Let $Q(\mathbf{p}^{(k)})$ be the corresponding shortening set at the k-step, with $Q(\mathbf{p}^{(0)}) = \emptyset$. Let $f^{(k)}$ be the selected pre-frozen bit at the k-step, then these two sets can be computed by

$$\mathcal{A}_S^{c(k)} = \mathcal{A}_S^{c(k-1)} \bigcup f^{(k)}, \tag{7}$$

and

$$Q(\mathbf{p}^{(k)}) = Q(\mathbf{p}^{(k-1)}) \bigcup f^{(k)}, \tag{8}$$

where $1 \leq k \leq N_p$. The pre-frozen bit $f^{(k)}$ can be selected from the temporary weight-1 set $\mathcal{W}^{(k)}$, which is determined by the pre-frozen construction function as follows:

$$\mathcal{W}^{(k)} = \underset{j' \in \mathcal{N}^{(k)}}{\arg} |Q(\mathbf{g}_{j'})| = 1, \tag{9}$$

where $\mathcal{N}^{(k)}$ is the index set after shortening, with $\mathcal{N}^{(k)} = \mathcal{N}^{(k-1)} - \mathcal{A}_S^{c(k-1)}$ and $\mathcal{N}^{(0)} = \{0, 1, \cdots, N-1\}$. The shortening construction algorithm can be described as follows.

Algorithm 1. The shortening construction

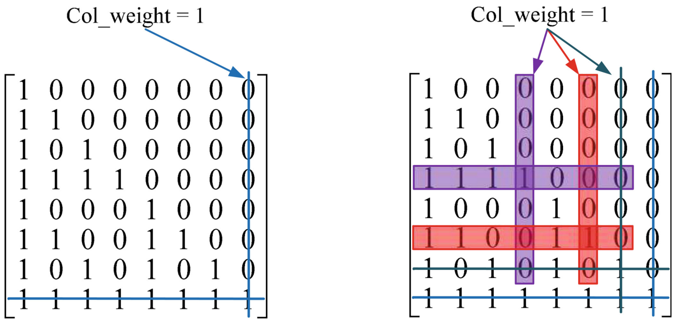
1: Given the required shortened code length M, the mother code length $N = 2^{\lceil \log_2 M \rceil}$, the generator matrix $\mathbf{G}_N$ and the number of shortening bits $N_p = N - M$
2: **Initialization:** $\mathcal{N}^{(0)} = \{0, 1, \cdots, N-1\}$, $\mathcal{A}_S^{c(0)} = \emptyset$
3: **for** $k = 1 : N_p$ **do**
4: Update $\mathcal{N}^{(k)} = \mathcal{N}^{(k-1)} - \mathcal{A}_S^{c(k-1)}$
5: Compute $\mathcal{W}^{(k)}$ according to (9)
6: Select the k-th pre-frozen bit $f^{(k)}$ from $\mathcal{W}^{(k)}$
7: Compute $Q(\mathbf{p}^{(k)})$ according to (8)
8: Compute $\mathcal{A}_S^{c(k)}$ according to (7)
9: **end for**

Remarks 1: In each step, only one bit is allowed to be selected from $\mathcal{W}^{(k)}$. However, the cardinality of $\mathcal{W}^{(k)}$ is greater than 1 in most case, which means there exist more than one possible scheme to pick out $f^{(k)}$. Specifically, if we modify the pre-frozen construction function as

$$\mathcal{W}^{*(k)} = \max \underset{j' \in \mathcal{N}^{(k)}}{\arg} |Q(\mathbf{g}_{j'})| = 1, \tag{10}$$

then the Algorithm 1 is equivalent to the shortening scheme presented in [6], where the last successive N_p indices are designated as the shortened bits and thus the pre-frozen bit positions.

Example 1: Consider a shortened polar code with $M = 6$, then we have $N = 8$ and $N_p = N - M = 2$, and the generator matrix is $\mathbf{G}_8 = \mathbf{F}^{\otimes 3}$. Figure 2 shows the shortening construction process with the pre-frozen construction function defined in (10). There are 2 steps to perform the construction. At the first step, only the last column $\mathbf{g}_7$ satisfies the weight-1 constraint. Thus we have $\mathcal{W}^{(1)} = \{7\}$ and $f^{(1)} = 7$. Obviously, $\mathbf{p}^{(1)} = (00000001)$ and $Q(\mathbf{p}^{(1)}) = \{7\}$, $\mathcal{A}_S^{c(1)} = \{7\}$, as shown in Fig. 2(a). It can be seen that the column 7 and row 7 are deleted from $\mathbf{G}_8$. At the second step, there exist 3 columns, $\mathbf{g}_3, \mathbf{g}_5, \mathbf{g}_6$ satisfy the weight-1 constraint. According to (10), only the maximum index 6 is selected, *i.e.*, $\mathcal{W}^{(2)} = \{6\}$ and thus $f^{(2)} = 6$. Similarly, $\mathbf{p}^{(2)} = (00000011)$ and $Q(\mathbf{p}^{(2)}) = \{6, 7\}$, $\mathcal{A}_S^{c(2)} = \{6, 7\}$, as shown in Fig. 2(b). At this step, column 6 and row 6 are deleted from $\mathbf{G}_8$.



(a) The first step.

(b) The second step.

Fig. 2. Shortening construction of Example 1.

3.2 RM Code-Aided Shortening Algorithm

As discussed in the previous subsection, the pre-frozen construction function $\mathcal{W}^{(k)}$ is of importance, since the pre-frozen bit is determined by this function. To guarantee the minimum number of the pre-frozen bits, it is required that $|\mathcal{A}_S^{c(N_p)}| = |Q(\mathbf{p}^{(N_p)})| = N_p$. Therefore, the weight-1 constraint is introduced for the pre-frozen construction function, as shown in (9). However, there exist more than one column with weight-1 at the k step when $k > 1$, implying that the cardinality $|\mathcal{W}^{(k)}| > 1$ thus $f^{(k)}$ may have different construction schemes. Although the scheme according to (10) has a simple shortening pattern, the construction does not take the Hamming distance into account, which may cause performance degradation.

In this subsection, we propose a distance-greedy shortening scheme with the help of RM codes, called RM code-aided shortening construction scheme. Since polar code is the linear block code obtained by the generator matrix G_N, the minimum Hamming distance is then depended on the minimum row weight of G_N. Moreover, polar codes are essentially a generalization of RM codes, and they share a common generator matrix G_N. We can jointly optimize the distance property of the RCSP codes under the row weight constraint of RM codes.

Let t be an element of $\mathcal{W}^{(k)}$ at the k step, which is also a candidate for $f^{(k)}$. According to the row weight property of RM code, different index t shows different weight for the t-row of G_N. In order to maximize the Hamming distance, the row having the minimum weight at each construction step is deleted first. Thus, the pre-frozen bit $f^{(k)}$ at the k step can be computed by

$$f^{(k)} = \min \underset{t \in \mathcal{W}^{(k)}}{\arg} \min w_r(t). \tag{11}$$

Equation (11) indicates that, the selected index t from the candidates corresponds to the t-row of G_N with minimum row weight. In other words, the shortened bit is selected to maximize the row weight of generator matrix and thus the Hamming distance of the resulting RCSP codes can be improved.

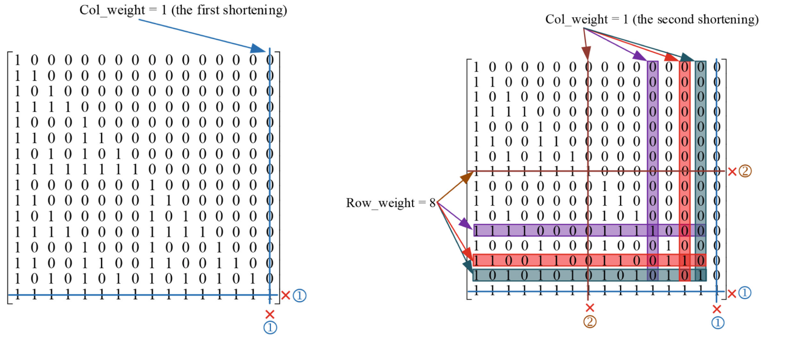
The proposed RM code-aided shortening construction can be described as follows.

Algorithm 2. The RM code-aided shortening construction

Given the required shortened code length M, the mother code length $N = 2^{\lceil \log_2 M \rceil}$, the generator matrix $\mathbf{G}_N$ and the number of shortening bits $N_p = N - M$
Initialization: $\mathcal{N}^{(0)} = \{0, 1, \cdots, N-1\}$, $\mathcal{A}_S^{c(0)} = \emptyset$
for $k = 1 : N_p$ **do**
Update $\mathcal{N}^{(k)} = \mathcal{N}^{(k-1)} - \mathcal{A}_S^{c(k-1)}$
Compute $\mathcal{W}^{(k)}$ according to (9)
Compute the k-th pre-frozen bit $f^{(k)}$ according to (11)
Compute $Q(\mathbf{p}^{(k)})$ according to (8)
Compute $\mathcal{A}_S^{c(k)}$ according to (7)
end for

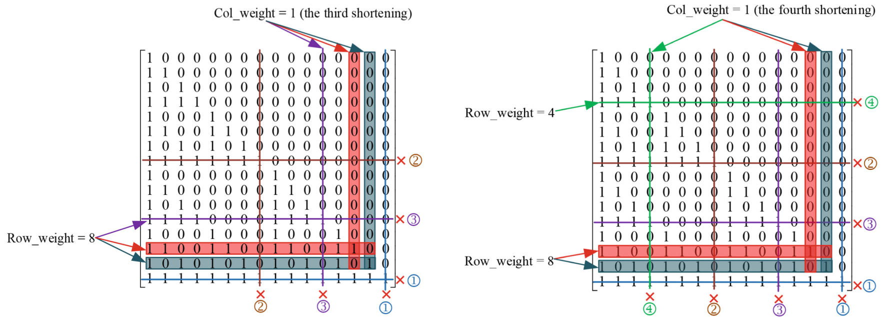
Remarks 2: Different from the Algorithm 1, the proposed construction algorithm is designed to be distance-greedy. When the candidates in $\mathcal{W}^{(k)}$ are greater than 1, we choose the one which has the minimum weight. Note that, if the index t produces the same minimum row weight $w_r(t)$, then the minimum index is selected, implying the uppermost row will be deleted.

Example 2: We construct shortened codes with code length $M = 12$, then we have $N = 16$ and $N_p = 4$, and the generator matrix is $\mathbf{G}_{16} = \mathbf{F}^{\otimes 4}$. Figure 3 shows the process of shortening construction with the pre-frozen construction function defined in (9) and $f^{(k)}$ defined in (11). We have four steps to construct the shortened code. At the first step, only the last column $\mathbf{g}_{15}$ satisfies the weight-1 constraint. Thus we have $\mathcal{W}^{(1)} = \{15\}$ and $f^{(1)} = 15$. Obviously, $\mathbf{p}^{(1)} = (0000000000000001)$ and $Q(\mathbf{p}^{(1)}) = \{15\}$, $\mathcal{A}_S^{c(1)} = \{15\}$, as shown in Fig. 3(a). It can be seen that the column 15 and row 15 are deleted from $\mathbf{G}_{16}$. At the second step, there exist 4 columns, $\mathbf{g}_7, \mathbf{g}_{11}, \mathbf{g}_{13}, \mathbf{g}_{14}$ satisfy the weight-1 constraint. According to (9) and (11), we have $\mathcal{W}^{(2)} = \{7, 11, 13, 14\}$ and thus $f^{(2)} = 7$. Similarly, $\mathbf{p}^{(2)} = (0000000100000001)$ and $Q(\mathbf{p}^{(2)}) = \{7, 15\}$, $\mathcal{A}_S^{c(2)} = \{7, 15\}$, as shown in Fig. 3(b). At this step, column 7 and row 7 are deleted from $\mathbf{G}_{16}$. The third step is shown in Fig. 3(c). There are 3 columns, $\mathbf{g}_{11}, \mathbf{g}_{13}, \mathbf{g}_{14}$ satisfy the weight-1 constraint. Accordingly, we have $\mathcal{W}^{(3)} = \{11, 13, 14\}$ and thus $f^{(3)} = 11$. Then, $\mathbf{p}^{(3)} = (0000000100010001)$ and $Q(\mathbf{p}^{(3)}) = \{7, 11, 15\}$, $\mathcal{A}_S^{c(3)} = \{7, 11, 15\}$. At this step, column 11 and row 11 are deleted from $\mathbf{G}_{16}$. At the fourth step, there exist 3 columns, $\mathbf{g}_3, \mathbf{g}_{13}, \mathbf{g}_{14}$ satisfy the weight-1 constraint. Thus, we have $\mathcal{W}^{(4)} = \{3, 13, 14\}$ and $f^{(4)} = 3$. Similarly, $\mathbf{p}^{(4)} = (0001000100010001)$ and $Q(\mathbf{p}^{(4)}) = \{3, 7, 11, 15\}$, $\mathcal{A}_S^{c(4)} = \{3, 7, 11, 15\}$, as shown in Fig. 3(d). At this step, column 3 and row 3 are deleted from $\mathbf{G}_{16}$. Finally, the RCSP code jointly designed by the pre-frozen set $\mathcal{A}_S^c = \{3, 7, 11, 15\}$ and the shortened pattern $\mathbf{p} = (0001000100010001)$ is constructed.



(a) Delete the last column and row

(b) Delete the uppermost row and the corresponding weight-1 column



(c) Delete the uppermost row and the corresponding weight-1 column

(d) Delete the row with the minimum weight and the corresponding weight-1 column

Fig. 3. RM code-aided shortening construction of Example 2.

4 Simulation Results

4.1 Row Weight Property

As described above, the minimum Hamming distance of a polar code can be obtained by the minimum row weight of G_N. Therefore, the Hamming distance of the resulted RCSP code could be known by the row weight property of G_N after being shortened and frozen. In this subsection, we analyze and compare the row weight distribution with the shortening scheme in [6] (marked by Wang14) and the proposed RM code-aided shortening construction scheme (marked by Proposed). The row weight distribution with different code lengths and code rates are considered.

Table 1 shows the row weight distribution of the polar code with $N = 512$, $M = 312$, $N_p = N - M = 200$, $R = 0.25$ and 0.8, respectively. Table 2 corresponds to the code with $N = 1024$, $M = 600$,$N_p = N - M = 424$, $R = 0.25$ and 0.8, respectively.

Table 1. Row weight distribution with RCSP code length $M = 312$ under different shortening algorithms

Row weight		4	8	16	32	64	128	256
Wang14	$R = 0.25$	0	0	9	32	28	8	1
	$R = 0.8$	8	47	87	68	31	8	1
Proposed	$R = 0.25$	0	0	3	37	29	8	1
	$R = 0.8$	3	51	88	68	31	8	1

Table 2. Row weight distribution with RCSP code length $M = 600$ under different shortening algorithms

Row weight		4	8	16	32	64	128	256	512
Wang14	$R = 0.25$	0	0	8	38	58	36	9	1
	$R = 0.8$	7	52	135	147	92	37	9	1
Proposed	$R = 0.25$	0	0	1	37	66	36	9	1
	$R = 0.8$	1	52	140	148	92	37	9	1

From Table 1, it can be seen that when the code rate is 0.25, the minimum row weight of the two shortening algorithms are both 16. However, the Wang14 scheme has 9 such rows with minimum weight 16, while the Proposed scheme has 3 such rows with minimum weight 16. Therefore, the number of rows with minimum weight of the proposed algorithm is less than that of the Wang14, which implies that the Hamming distance among the RCSP codewords can be improved in average. When the code rate changes to 0.8, we have the similar observations. That is, the number of rows with minimum weight (say, 4 in this case) of the proposed algorithm is still less than that of the Wang14 scheme, which shows a better distance property in average. As shown in Table 2, the code length now changes to $N = 1024$. Similarly, for the two code rates, the number of rows with minimum weight of the Proposed scheme are both less than that of the Wang14 scheme.

In summary, the Proposed algorithm can reduce the number of rows with minimum weight and thus can improve the Hamming distance property of the constructed RCSP codes in average. The improvement of the minimum Hamming distance may have a positive impact on the decoding performance, as shown in the performance analysis in the next subsection.

4.2 Decoding Performances

In this subsection, we compare the frame error rate (FER) performance of the proposed RM-aided shortening construction scheme (Proposed), the Wang14 [6] and the QUP [5] schemes under different code rates and different code lengths. In the simulations, the GA method is used for the channel reliability estimation and the SC decoding is performed. The additive white Gaussian noise (AWGN) channel with binary phase-shift keying (BPSK) modulation is considered. The total number of simulation frames is 10^7, and the maximum number of error frames is 200. When the simulation reaches the total number of frames or reaches the maximum number of error frames, the simulation is stopped.

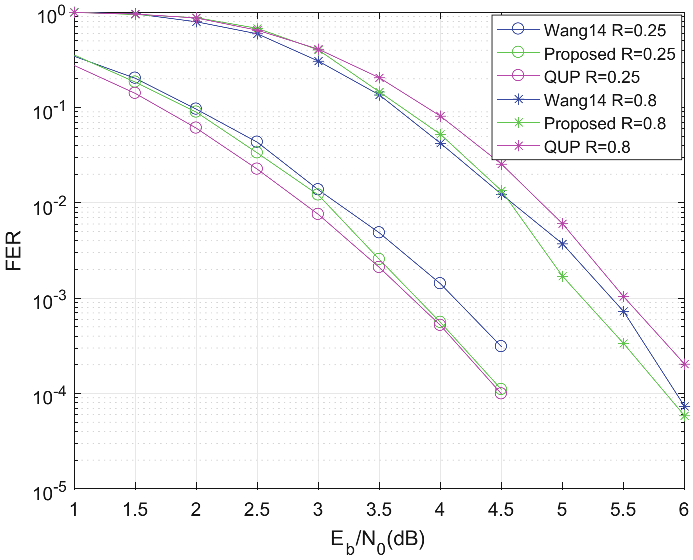


Fig. 4. Performance comparison under different rate matching algorithms with $M = 312$

Figure 4 shows the performances of the RCSP codes with the length $M = 312$ and mother code length $N = 512$. The code rates are set to be $R = 0.25$ and 0.8, respectively. We have the following observations.

– When the code rate is 0.25, the Proposed scheme has a significant performance gain compared with the Wang14 scheme. For example, at the FER $= 10^{-3}$, there is a gain of about 0.3 dB.
– Compared with the QUP scheme, the Proposed scheme shows a slightly performance degradation at low SNR region. However, the performance gap becomes smaller with the increasing of the SNR and can be ignored at high SNR region.
– For the code rate $R = 0.8$, the Proposed scheme performs as well as the Wang14 scheme but has a better FER performance than that of the QUP scheme. For example, when the FER is 10^{-3}, it has a performance gain about 0.35 dB compared with the QUP scheme.

In Fig. 5, we show the performances of the RCSP codes with the length $M = 600$ and mother code length $N = 1024$. The code rates are set to be $R = 0.25$ and 0.8, respectively. We have a similar observation as expected.

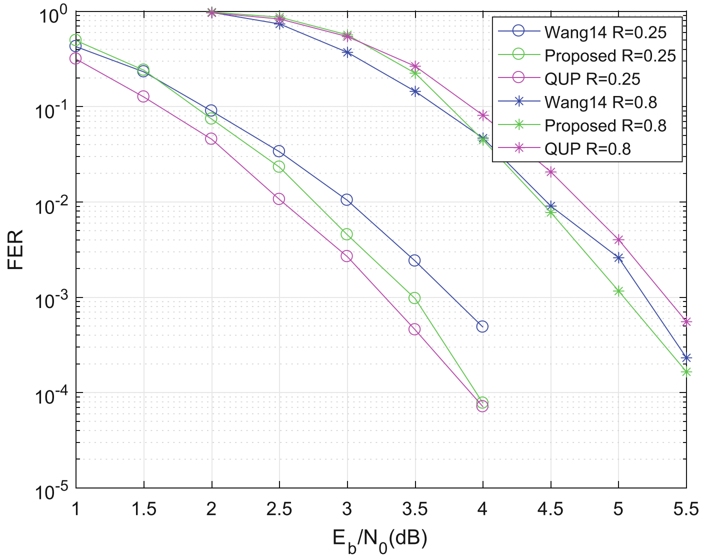


Fig. 5. Performance comparison under different rate matching algorithms with $M = 600$

- When the code rate is 0.25, the Proposed scheme has a better performance than the Wang14 scheme with a gain of about 0.3 dB at the FER $= 10^{-3}$.
- The QUP scheme performs better that the Proposed scheme at low SNR region, but the performance gap can be ignored at high SNR region.
- When the code rate is 0.8, the Proposed scheme has a comparable performance of the Wang14 scheme, but achieves a performance gain about 0.35 dB of the QUP scheme at the FER $= 10^{-3}$.

5 Conclusion

In this paper, we have proposed a hybrid algorithm to construct the rate-compatible shortened polar code by combining with the RM-rule constraint. First, the shortened bits are designed to be completely known by the decoder and thus the corresponding log-likelihood ratios (LLRs) can be set to infinity (or minus infinity) to guarantee the decoding performance. Second, a distance-greedy construction method is proposed to further improve the performance. When multiple candidate positions satisfy the weight-1 column constraint, we tend to choose the rows having less weights to form the shortened/pre-frozen bits. In this way, the generator matrix of the shortened code can be constructed with row weight as large as possible. The row weight distributions of generator matrix show that the Proposed scheme has less rows with minimum weight, which makes the Hamming distance of the constructed RCSP codes be larger in average. Simulation results show that the Proposed scheme can achieve performance gains at different levels when compared with the Wang14 scheme and the QUP scheme.

References

1. Arikan, E.: Channel polarization: a method for constructing capacity-achieving codes for symmetric binary-input memoryless channels. IEEE Trans. Inf. Theory **55**(7), 3051–3073 (2009)
2. 3GPP. Multiplexing and Channel Coding: 3GPP 38.212 vol. 15.1.0 [S] (2018)
3. Multiplexing and Channel Coding [S], document 3GPP TS38. 212 V15. 0.0 (2017)
4. Eslami, A., Pishro-Nik, H.: A practical approach to polar codes. In: 2011 IEEE International Symposium on Information Theory Proceedings, St. Petersburg, Russia, pp. 16–20. IEEE (2011)
5. Niu, K., Chen, K., Lin, J.: Beyond turbo codes: rate-compatible punctured polar codes. In: 2013 IEEE International Conference on Communications, Budapest, Hungary, pp. 3423–3427. IEEE (2013)
6. Wang, R., Liu, R.: A novel puncturing scheme for polar codes. IEEE Commun. Lett. **18**(12), 2081–2084 (2014)
7. Miloslavskaya, V.: Shortened polar codes. IEEE Trans. Inf. Theory **61**(9), 4852–4865 (2015)
8. Li, B., Shen, H., Tse, D.: A RM-Polar codes. arXiv preprint arXiv:1407.5483 (2014)
9. Muller, E.D.: Application of Boolean algebra to switching circuit design and to error detection. Trans. I.R.E. Prof. Group Electron. Comput. **EC-3**(3), 6–12 (1954)
10. Reed, I.: A class of multiple-error-correcting codes and the decoding scheme. Trans. IRE Prof. Group Inf. Theory **4**(4), 438–449 (1954)
11. Arikan, E.: A survey of reed-muller codes from polar coding perspective. In: 2010 IEEE Information Theory Workshop on Information Theory, Cairo, Egypt, pp. 1–5. IEEE (2010)
12. Mori, R., Tanaka, T.: Performance of polar codes with the construction using density evolution. IEEE Commun. Lett. **13**(7), 519–521 (2009)
13. Trifonov, P.: Efficient design and decoding of polar codes. IEEE Trans. Commun. **60**(11), 3221–3227 (2012)
14. He, G.: β-expansion: a theoretical framework for fast and recursive construction of polar codes. In: 2017 IEEE Global Communications Conference, Singapore, pp. 1–6. IEEE (2017)