# Flexible IR-HARQ Scheme for Polar-Coded Modulation

Peihong Yuan, Fabian Steiner, Tobias Prinz

Institute for Communications Engineering Technical University of Munich Munich, Germany {peihong.yuan,fabian.steiner,tobias.prinz}@tum.de

Abstract—A flexible incremental redundancy hybrid automated repeat request (IR-HARQ) scheme for polar codes is proposed based on dynamically frozen bits and the quasi-uniform puncturing (QUP) algorithm. The length of each transmission is not restricted to a power of two. It is applicable for the binary input additive white Gaussian noise (biAWGN) channel as well as higher-order modulation. Simulation results show that this scheme has similar performance as directly designed polar codes with QUP and outperforms LTE-turbo and 5G-LDPC codes with IR-HARQ.

Index Terms-polar coding, coded modulation, IR-HARQ

#### I. INTRODUCTION

Many communication channels are time-varying and unknown to the transmitter. Incremental redundancy hybrid automated repeat request (IR-HARQ) as shown in Fig. 1 is a scheme that transmits additional redundancy bits until the data bits can be reconstructed. For turbo codes (such as those used in LTE), a low rate mother code is punctured with different patterns for several transmissions. The coding scheme for enhanced mobile broadband (eMBB) in 5G uses protographbased, Raptor-like LDPC codes [1] that allow for both flexible block length and code rate adaptation. The standard defines two base matrices that offer optimized performance for different operating regimes.

With cyclic redundancy check (CRC) outer codes and successive cancellation list (SCL) decoding [2], polar codes [3], [4] outperform state-of-the-art turbo and LDPC codes in the short to medium length regime. Polar-coded modulation (PCM) is discussed in [5]–[7]. The performance comparison and efficient code design methods of three polar-coded modulation schemes are presented in [8, Fig. 11]. Multilevel polar coding (MLPC) with set partitioning (SP) labeling in [5] performs best and is around 1 dB more power efficient than an AR4JA [9] LDPC code decoded with 200 iterations.

A quasi-uniform puncturing (QUP) algorithm was proposed in [10] to efficiently design length-flexible polar codes, i.e., polar codes where the number of coded bits is not limited to be a power of two.

In [11], a scheme the authors called "polar codes with incremental freezing" and in [12] "parallel-concatenated polar codes" is proposed. The capacity-achievability of this scheme is proved in [11] by using the nesting property. In [13], a

Georg Böcherer Mathematical and Algorithmic Sciences Lab Huawei Technologies France Paris, France georg.boecherer@ieee.org







polar code extension method is presented which outperforms the scheme in [11], [12] for finite block length.

In this work, we prove that the scheme in [13] can achieve capacity asymptotically in the block length under some design constraints. In addition, a length-flexible IR-HARQ scheme based on dynamically frozen bits and QUP is proposed. This scheme is extended to polar-coded modulation with amplitude shift keying (ASK) and quadrature amplitude modulation (QAM) constellations. Simulation results show that the polar codes designed by our algorithm have similar error correction performance as directly designed polar codes.

This work is organized as follows. In Sec. II, we review polar codes, PCM and QUP. We discuss existing and proposed IR-HARQ schemes in Sec. III. Sec. IV provides design examples and numerical results. The performance of the proposed scheme is compared with directly designed polar codes and 4G/5G codes [14], [15] in additive white Gaussian noise (AWGN) channels. We conclude in Sec. V.

## **II. PRELIMINARIES**

## A. Polar Coding

In this paper, uppercase letters X, Y, B, U denote the random variables (RV) while the corresponding lowercase letters are their realizations. The notation  $c_1^N$  is short for  $c_1c_2...c_N$ . For an arbitrary subset S of  $\{1, ..., N\}$ ,  $S^{\complement}$  is the complement of S and  $c_S$  denotes the vector of  $c_1^N$  formed by the elements with indices in S.

A binary polar code of block length N and dimension k is defined by the polar transform with matrix  $\mathbb{F}^{\otimes \log_2 N}$  and

$$I^{-} = I(U_{2}; Y_{1}Y_{2}) + I_{1} = I(B_{1}; Y_{1})$$
  
$$I^{+} = I(U_{2}; Y_{1}Y_{2}|U_{1}) + I_{2} = I(B_{2}; Y_{2})$$

Figure 2: MIs of the basic polar transform

N-k frozen positions, where  $\mathbb{F}^{\otimes \log_2 N}$  denotes the  $\log_2 N$ -fold Kronecker power of the kernel

$$\mathbb{F} = \begin{bmatrix} 1 & 0\\ 1 & 1 \end{bmatrix}.$$
 (1)

Polar encoding can be represented by

$$c_1^N = u_1^N \mathbb{F}^{\otimes \log_2 N}.$$
<sup>(2)</sup>

The vector  $c_1^N$  denotes the codeword. The vector  $u_1^N$  includes k information bits  $u_A$  and N - k predefined frozen bits  $u_{A^{\complement}}$ . A and  $A^{\complement}$  are called the information set and frozen set defined in [4]. SC decoding uses the channel observation and previous estimates  $\hat{u}_1, \ldots, \hat{u}_{i-1}$  to decode  $u_i$ . Both encoding and SC decoding have complexity  $\mathcal{O}(N \log_2 N)$  [4].

The polar code construction finds the most reliable bits in  $u_1^N$  under SC decoding. The Monte Carlo (MC) construction was introduced in [3], [4], and needs extensive simulations. An information theoretical construction was introduced in [4]. The reliability of the *i*th bit can be quantified by the mutual information (MI) I  $(U_i; Y^N | U_1^{i-1})$ . We can calculate these MIs for all  $i \in \{1, 2, ..., N\}$  by recursively calculating the MIs of the basic polar transform displayed in Fig. 2.

For the binary input additive white Gaussian noise (bi-AWGN) channel, density evolution [16] with Gaussian approximation (GA) [17] has much lower complexity and performs very close to the MC construction. The update rule for the basic polar transform is given by

$$I^{-} = 1 - J\left(\sqrt{\left[J^{-1}(1 - I_{1})\right]^{2} + \left[J^{-1}(1 - I_{2})\right]^{2}}\right)$$
(3)

$$I^{+} = J\left(\sqrt{\left[J^{-1}(I_{1})\right]^{2} + \left[J^{-1}(I_{2})\right]^{2}}\right)$$
(4)

The numerical approximations in [18] can be used for  $J(\cdot)$  and  $J^{-1}(\cdot)$ . Additionally, the frame error rate (FER) under SC decoding can be estimated by

$$\Pr\left(\hat{U}_{i} \neq U_{i} | \hat{U}_{1}^{i-1} = U_{1}^{i-1}\right) = Q\left(\frac{1}{2}J^{-1}\left(I\left(U_{i}; Y^{n} | U_{1}^{i-1}\right)\right)\right)$$
(5)  

$$\operatorname{FER}_{\mathrm{SC,est}} = 1 - \prod_{i \in \mathcal{A}} \left(1 - \Pr\left(\hat{U}_{i} \neq U_{i} | \hat{U}_{1}^{i-1} = U_{1}^{i-1}\right)\right)$$
(6)

where

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_{x}^{\infty} e^{\frac{u^{2}}{2}} \mathrm{d}u.$$
 (7)

The nesting property is introduced in [11]. The authors show that the reliability order of the polarized bits is independent on the channel quality for infinite length. Note that the nesting property generally does not hold for finite length.

# B. Polar-Coded Modulation

We consider the following memoryless AWGN channel model with  $2^m$ -ASK constellations  $X \in \mathcal{X} = \{\pm 1, \pm 3, \dots, \pm (2^m - 1)\}.$ 

$$Y = X + \sigma Z \tag{8}$$

where Z is zero mean Gaussian noise with variance one. Note that ASK constellations can be orthogonally extended to QAM constellations.

Polar-coded modulation (PCM) schemes are discussed in [5]–[7]. In [8], the performance of three schemes is compared and their efficient design algorithms are presented. MLPC with SP labeling in [5] provides the best performance. MLPC with code length mN works as following:

# • Encoding

1. Put k information bits in vector  $u_1^{mN}$  and define

$$u_{1,j}^N = u_{(j-1)N+1}^{jN}, \ j = 1, \dots, m.$$
 (9)

2. Encode m polar codes

$$c_{1,j}^N = u_{1,j}^N \mathbb{F}^{\otimes \log_2 N}, \ j = 1, \dots, m.$$
 (10)

3. Map the code words to symbols for i = 1, ..., N

$$c_{i,1}, c_{i,2}, \dots, c_{i,m} = b_{i,1}, b_{i,2}, \dots, b_{i,m}$$
  
=  $(b_1, b_2, \dots, b_m)_i \stackrel{\text{SP}}{\mapsto} x_i.$  (11)

- Decoding
  - 1. Demap and decode level 1

$$\ell_{i,1} = \log \frac{p_{B_{i,1}|Y_i}(0|y_i)}{p_{B_{i,1}|Y_i}(1|y_i)}, \ i = 1, \dots, N$$

$$\hat{c}_{1,1}^N = \text{polarSCDecode}\left(l_{1,1}, \dots, l_{N,1}\right).$$
(12)

j. Demap and decode level j for  $j = 2, \ldots, m$ 

$$\ell_{i,j} = \log \frac{p_{B_{i,j}|Y_iB_{i,1}...B_{i,j-1}} \left(0|y_i\hat{c}_{i,1}...\hat{c}_{i,j-1}\right)}{p_{B_{i,j}|Y_iB_{i,1}...B_{i,j-1}} \left(1|y_i\hat{c}_{i,1}...\hat{c}_{i,j-1}\right)}, \\ i = 1, \dots, N$$
$$\hat{c}_{1,j}^N = \text{polarSCDecode} \left(l_{1,j}, \dots, l_{N,j}\right)$$
(13)

where l denotes the input log-likelihood ratio (LLR).

MI demapper GA (MI-DGA) construction calculates I  $(B_j; YB_1^{j-1})$  for j = 1, ..., m and uses them to find the most reliable bits in  $u_1^{mN}$  with GA. Therefore, the required code rates of the *m* polar codes  $k_1/N, ..., k_m/N$  are given by MI-DGA. Note that the overall code rate of the system is  $(k_1 + \cdots + k_m)/mN = k/mN$  and the transmission rate is k/N bits/channel use.

# C. Rate-Matched Polar Codes

Because of the recursive structure of  $\mathbb{F}^{\otimes \log_2 N}$ , polar codes usually have a block length that is a power of two. Punctured polar codes are introduced in [19], [20]. The punctured (n, k)polar codes can be decoded with standard polar decoders for length  $N = 2^{\lceil \log_2 n \rceil}$ . For punctured polar codes, N - n bits in  $c_1^N$  are not transmitted and the corresponding LLRs are set to zero. For GA construction, the initial MIs of the punctured bits are zero.

With the QUP algorithm, the first N - n bits in  $c_1^N$  are not transmitted, i.e.,

GA construction: set  $I(B_i; Y) = 0$ , i = 1, ..., N - nafter encoding: transmit  $c_{N-n+1}^N$ before decoding: set  $l_1^{N-n} = 0$ .

This algorithm is called QUP because the punctured position in bit-reversal representation [4] looks like uniformly distributed in  $\{1, \ldots, N\}$ . Let (n, k, N) denote a QUP polar code with dimension k, block length n and punctured from N bits mother codes with QUP algorithm, where N has to the power of two.

**Theorem 1.** For an (n, k, N) QUP-polar code, the first N - n bits in  $u_1^N$  are frozen.

*Proof.* For the MI update of the basic polar transform in Fig. 2, it is easy to show that

$$I^{-} \leq \min\{I_1, I_2\}, \ I^{+} \geq \max\{I_1, I_2\}.$$
 (14)

As MI is always non-negative, we have

$$I^{-} = 0$$
, if  $I_1 = 0$  or  $I_2 = 0$  (15)

$$I^{-} = I^{+} = 0$$
, if  $I_{1} = I_{2} = 0$ . (16)

Thus the number of channels with zero capacity is invariant. With QUP we set  $I_1^{N-n}$  to zero. Because of the recursive structure of  $\mathbb{F}^{\otimes \log_2 N}$  the zeros will propagate through the transform which causes I  $(U_i; Y^N | U_1^{i-1}) = 0$  for  $i = 1, \ldots, N-n$ .

**Theorem 2.** All  $(n, k, 2^j N)$  QUP-polar codes have the same encoding and decoding complexity as (n, k, N) QUP-polar codes, where j is a natural number and  $N = 2^{\lceil \log_2 n \rceil}$ .

*Proof.* The decoder for (n, k, 2N) QUP-polar codes is shown in Fig. 3. Obviously, more than N bits are punctured. Thus,  $u_1^N$  are all frozen because of Theorem 1. According to the SC decoding, we first decode  $u_1^N$  (upper decoder) and then decode  $u_{N+1}^{2N}$  (lower decoder) based on  $\hat{u}_1^N$ . The input of the lower decoder is  $(1-0) \cdot \ell_1^N + \ell_{N+1}^{2N} = \ell_{N+1}^{2N}$ . Thus for this QUPpolar code, we just need to run the lower decoder (dashed box in Fig. 3) which is the decoder for (n, k, N) QUP-polar codes. With the same idea we can extend the theorem to mother code length  $2^j N$ .

## 

## III. IR-HARQ WITH POLAR CODES

The basic idea of IR-HARQ is displayed in Fig. 1. The decoder receives  $n_t$  bits from the *t*th transmission and then decodes  $\left(\sum_{q=1}^t n_q, k\right)$  code. Using the CRC, the receiver may detect a decoding failure, in which case it requests the (t+1)th transmission from the sender.



Figure 3: Equivalence of (n, k, N) and (n, k, 2N) QUP-polar code.

## A. Existing Schemes: Polar Codes with Incremental Freezing

In [11], [12], polar codes with incremental freezing are proposed. The main idea is as follows. An  $(N_1, k)$  polar code is transmitted first. When retransmission occurs, we additionally freeze k' "most unreliable" information bits in the  $(N_1, k)$  code and send them with an  $(N_2, k')$  code in next transmission. After the 2nd transmission we decode the  $(N_2, k')$  code and the first code successively. Note that the first code becomes a  $(N_1, k - k')$  code with the estimation of the "most unreliable" information bits. For the 3rd transmission (if needed), we freeze k'' "most unreliable" information bits in both the  $(N_1, k - k')$  code and the  $(N_2, k')$  code. The retransmissions are continued with the same manner until the decoding is successful. For infinite code length, this scheme achieves capacity [11] with the nesting property. We note that this scheme is equivalent to dividing the  $\left(\sum_{q=1}^{t} N_q, k\right)$  code into t separated polar codes and decoding them successively, which causes a huge performance loss in the finite length regime.

#### B. Existing Schemes: Polar Extension

A polar code extension method is presented in [13]. Consider the 1st transmission with  $(N_1, k)$  polar code  $c(1)_1^{N_1} = u(1)_1^{N_1} \mathbb{F}^{\otimes \log_2 N_1}$ , where  $c(t)_1^N$  denotes the codeword after the *t*th transmission. When retransmission occurs, we design a  $(2N_1, k)$  polar code  $c(2)_{N_1+1}^{2N_1} = u(2)_1^{2N_1} \mathbb{F}^{\otimes \log_2(2N_1)}$ , let the information set in  $u(2)_{N_1+1}^{2N_1}$  be a subset of the information set in  $u(1)_1^{N_1}$  and set  $u(2)_{N_1+1}^{2N_1} = u(1)_1^{N_1}$ . Then copy the bits, which are reliable in  $u(1)_1^{N_1}$  but unreliable in  $u(2)_{N_1+1}^{2N_1}$  to the most reliable positions in  $u(2)_1^{N_1}$ . The code word can be presented by

$$c(2)_{N_1+1}^{2N_1} = c(1)_1^{N_1} \tag{17}$$

$$c(2)_1^{N_1} = c(1)_1^{N_1} \oplus u(2)_1^{N_1} \mathbb{F}^{\otimes \log_2 N_1}.$$
 (18)

Thus, only the first  $N_1$  bits  $c(2)_1^{N_1}$  need to be transmitted, because the receiver already knows the remaining  $N_1$  bits of the  $(2N_1, k)$  polar code from the previous transmission. The receiver first decodes the information bits in  $u(2)_1^{N_1}$ , copies them to  $u(2)_{N_1+1}^{2N_1}$  as frozen bits and then decodes the remaining bits. The only difference between a directly generated  $(2N_1, k)$  polar code and the code after the 2nd transmission is the information set in  $u(2)_{N_1+1}^{2N_1}$  which has to be a subset of the information set in  $u(1)_1^{N_1}$ . In this scheme, we decode an  $(N_2, k)$  code after the 2nd transmission  $(N_2 = 2N_1)$ . This extension is repeated until the decoding is successful. This extension method is not flexible because the transmission length has to be the same as the sum of all previous transmissions, i.e.,

$$N_t = \sum_{q=1}^{t-1} N_q, \ t = 2, 3, \dots$$
 (19)

This scheme is capacity-achieving under some constraints. According to the information theoretical construction in [4], designing a (2N, k) polar code for a channel with mutual information I is equivalent to designing a  $(N, k_1)$  polar code for  $I^-$  and a  $(N, k_2)$  polar code for  $I^+$ , where

$$k_1 + k_2 = k \tag{20}$$

$$k_1/k_2 = I^-/I^+. (21)$$

Let  $I_t$  be the design MI for the tth transmission and  $2N_t$  the code length after the *t*th transmission ( $t \ge 2$ ). Note that only  $N_t$  bits are sent in the *t*th transmission. Under the constraints

$$I_t^+ \le I_{t-1}, \ t = 2, 3, \dots$$
 (22)

the information set in  $u(t+1)_{N_t+1}^{2N_t}$  is always a subset of the information set in  $u(t)_1^{N_t}$  because of the nesting property, i.e., every polar code after tth transmission is an optimal  $\left(\sum_{q=1}^{t} N_q, k\right)$  polar code.

## C. Proposed Scheme

Polar codes with dynamically frozen bits are proposed in [21] to improve the distance properties of polar codes. The idea is to predetermine some frozen bits as linear combinations of previous information bits. The corresponding "dynamically frozen constraint" is needed to encode and decode. Our scheme is based on this technique and QUP. There are no constraints regarding the block length of any transmission.

We assume the system is designed for a maximum of  $t_{max}$ transmissions. Let N be the length of a mother polar code

$$N = 2^{\left\lceil \log_2\left(\sum_{q=1}^{t_{\max}} n_q\right)\right\rceil},\tag{23}$$

where  $\left[\cdot\right]$  denotes the ceiling function and  $n_q$  denotes the length of the qth transmission. In the proposed scheme, after the *t*th transmission, a  $\left(\sum_{q=1}^{t} n_q, k, N\right)$  QUP-polar code is decoded. The main structure and design algorithm are displayed in Fig. 4 and Algorithm 1 respectively. The set  $\mathcal{I}_t$ is

$$\left\{N - \sum_{q=1}^{t} n_q + 1, \dots, N - \sum_{q=1}^{t-1} n_q\right\},$$
 (24)

and  $\mathcal{A}_t$  and  $\mathcal{A}_t^{\complement}$  denote the information and frozen set of the  $\left(\sum_{q=1}^{t} n_q, k, N\right)$  QUP-polar code after the *t*th transmission. The output dynamically frozen constraint is used for encoding and decoding. Note that  $\mathcal{A}_t \cup \mathcal{A}_t^{\complement} = \bigcup_{a=1}^t \mathcal{I}_q$ . First



 $N - \sum_{q=1}^{t} n_q$  bits are frozen to zero (Theorem 1), but their indices are neither in  $\mathcal{A}_t$  nor  $\mathcal{A}_t^{\complement}$ .

Algorithm 1: Design tth transmission
<b>Input</b> : message length $k$ , mother code length $N$ ,
th transmission length $n_t$ , design MI $I_t$ ,
previous code length $n_t = \sum_{q=1} n_q$ ,
previous frozen set $\mathcal{A}_t^{\prime o} = \bigcup_{q=1}^{\sigma-1} \mathcal{A}_q^{o}$ ,
previous information set $\mathcal{A}'_t = \bigcup_{q=1}^{t-1} \mathcal{A}_q$
<b>Output:</b> information set $A_t$ ,
frozen set $\mathcal{A}_t^{L}$ ,
dynamically frozen constraint
1 Initialize design MI: $I_1^{N-n_t-n'_t} = 0, I_{N-n_t-n'_t+1}^N = I_t.$
2 Estimate $\operatorname{Pe}_{i} := \operatorname{Pr}\left(\hat{U}_{i} \neq U_{i}   \hat{U}_{1}^{i-1} = U_{1}^{i-1}\right).$
3 Set Pe $_{4'}$ = 1. Find k smallest in Pe $_{N-n_1-n'+1}^N$ and put their
indices in $\mathcal{A}_t$ .
4 Frozen set $\mathcal{A}_t^{\complement} = \bigcup_{q=1}^t \mathcal{I}_q \setminus \mathcal{A}_t.$
5 if $t \neq 1$ then
6 Dynamically frozen constraint is given by $u_{\mathcal{A}'_t \setminus \mathcal{A}_t} = u_{\mathcal{A}_t \setminus \mathcal{A}'_t}$ .
7 end

Obviously, for t = 1, the code is a normal QUP-polar code. Other codes are extended from previous codes with dynamically frozen bits. The bits which are already frozen have to be frozen for all extensions. This scheme is equivalent to the polar code extension method in [13] if  $n_1$  is a power of two and  $n_t = \sum_{q=1}^{t-1} n_q$ ,  $t = 2, 3, \ldots, t_{\text{max}}$ . For example, we consider k = 5 and n = (7, 5). The

information bits are  $m_{1,\dots,5}$ .

- For the first transmission,  $A_1 = \{12, 13, 14, 15, 16\}$  for the (7,5,16) QUP-polar code. Thus,  $u_{\mathcal{A}_1} = m_1^5$  and  $u_{\{1,\dots,16\}\setminus A_1} = 0$ . We encode the vector by  $c(1)_1^{16} =$  $u_1^{16}\mathbb{F}^4$  and transmit the last 7 bits  $c(1)_{10}^{16}$ .
- For the second transmission,  $\mathcal{A}_2 = \{8, 12, 14, 15, 16\}$  for • the (12, 5, 16) QUP-polar code. We precode  $u_1^{16}$  with  $u_{\mathcal{A}_2 \setminus \mathcal{A}_1} = u_{\mathcal{A}_1 \setminus \mathcal{A}_2}$ , which is  $u_8 = u_{13} = m_2$  in this case. We encode  $c(2)_1^{16} = u_1^{16} \mathbb{F}^4$  and send  $c(2)_5^9$ . Because  $\mathbb{F}^4$  is a lower triangular matrix,  $u_8 = u_{13}$  does not change the first transmitted bits, i.e.,  $c(2)_{10}^{16} = c(1)_{10}^{16}$ .

Table I: Explanation for the curves in Fig. 5, Fig. 6, Fig. 7.

	direct polar	proposed polar	5G-LDPC	LTE-turbo
1st	- + -	-	-	— <del>×</del> —
2nd	- + -			— <del>×</del> —
3rd	- + -	-		— <del>×</del> —
4th	- + -			— <del>×</del> —

At receiver, we decode the (12, 5, 16) QUP-polar code from the noisy version of  $c(2)_5^{16}$ . Note that  $u_{13}$  is now a dynamically frozen bit with constraint  $u_{13} = u_8$ .

In this example, the QUP-polar codes with optimal information set are decoded after every single transmission by using all received information.

## D. Extension to MLC

With  $2^m$ -ASK,  $n_t$  symbols ( $mn_t$  bits) are transmitted in the *t*th transmission. We replace line 1,2 in Algorithm 1 with an MLPC construction algorithm (MI-DGA in this work).

## IV. DESIGN EXAMPLES AND SIMULATION RESULTS

In this section, three design examples for  $t_{\text{max}} = 4$  are shown in Fig. 5, Fig. 6, Fig. 7. We use 16 bits CRC with generator polynomial " $x^{16} + x^{12} + x^5 + 1$ " for error detection. The polar codes are decoded by min-sum approximated SCL decoding with list size 32. Log-MAP decoding with 10 iterations and belief propagation (BP) with 50 iterations are used for LTE-turbo and 5G-LDPC codes, respectively. In the 8-ASK example, Bit-Interleaved Coded Modulation (BICM) [22] is used for turbo and LDPC codes. Note that the directly designed polar codes (dashed curves) are  $\left(\sum_{q=1}^{t} n_q, k, \cdot\right)$ QUP-polar codes only serve as a reference and can not work for an IR-HARQ scheme.

The simulation results show that the polar codes generated by the proposed algorithm perform very similar to directly designed polar codes. In the 8-ASK example, the proposed scheme performs approximately 1dB better than 5G-LDPC codes after two to four transmissions.

Because of the extension constraint, it is hard to extend a heavily punctured polar codes. Consider an (N + w, k, 2N) QUP-polar code, where w is a positive integer and  $w \ll N$ . Normally, the first w bits are unreliable, while these bits are almost perfect in the extended (2N, k) code. This effect degrades the performance for all further extensions. For example, for  $n_1^4 = (1000, 100, 100, 100)$ , the 3rd and 4th polar codes perform much worse than directly designed code because  $\sum_{q=1}^{2} n_q = 1100$  and  $1100 - 1024 = 76 \ll 1024$ . Therefore, we should avoid using heavily punctured polar code for the  $\{1, \ldots, t_{\text{max}} - 1\}$ th transmissions in biAWGN channel. However, this effect disappears for MLPC. We can design very good codes for  $n_1^4/m = (1000, 100, 100, 100, 100)$ . The reason should be the automatically controlled code rate for m polar codes.

## V. CONCLUSION

In this paper, an IR-HARQ scheme based on QUP and dynamically frozen bits for biAWGN channel and PCM is proposed. Simulation results show that the rate-matched polar codes generated by the proposed algorithm perform very similar to directly designed QUP-polar codes.

For future work, this scheme can be applied for a fading channel, i.e. the channel information estimated by training symbols of the previous transmissions could be used to design polar codes for next transmission.

#### REFERENCES

- T. Y. Chen, K. Vakilinia, D. Divsalar, and R. D. Wesel, "Protograph-Based Raptor-Like LDPC Codes," vol. 63, no. 5, pp. 1522–1532, May 2015.
- [2] I. Tal and A. Vardy, "List decoding of polar codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2213–2226, May 2015.
- [3] N. Stolte, "Rekursive Codes mit der Plotkin-Konstruktion und ihre Decodierung," Ph.D. dissertation, TU Darmstadt, 2002.
- [4] E. Arıkan, "Channel polarization: A method for constructing capacityachieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, Jul. 2009.
- [5] M. Seidl, A. Schenk, C. Stierstorfer, and J. B. Huber, "Polar-coded modulation," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4108–4119, Sep. 2013.
- [6] H. Mahdavifar, M. El-Khamy, J. Lee, and I. Kang, "Polar coding for bit-interleaved coded modulation," *IEEE Trans. Veh. Technol.*, vol. 65, no. 5, pp. 3115–3127, Jun. 2015.
- [7] S. R. Tavildar, "Bit-permuted coded modulation for polar codes," *IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2017.
- [8] G. Böcherer, T. Prinz, P. Yuan, and F. Steiner, "Efficient polar code construction for higher-order modulation," *IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2017.
- [9] D. Divsalar, S. Dolinar, C. Jones, and K. Andrews, "Capacityapproaching protograph codes," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 6, pp. 876–888, Aug. 2009.
- [10] K. Niu, K. Chen, and J.-R. Lin, "Beyond turbo codes: Rate-compatible punctured polar codes," *IEEE Int. Conf. Commun. (ICC)*, pp. 3423–3427, Jun. 2013.
- [11] B. Li, D. Tse, K. Chen, and H. Shen, "Capacity-achieving rateless polar codes," *IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 46–50, Jul. 2016.
- [12] S.-N. Hong, D. Hui, and I. Marić, "Capacity-achieving rate-compatible polar codes," *IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 41–45, Jul. 2016.
- [13] L. Ma, J. Xiong, Y. Wei, and M. Jiang, "An incremental redundancy HARQ scheme for polar code," arXiv preprint arXiv:1708.09679, 2017.
- [14] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and channel coding," 3rd Generation Partnership Project (3GPP), TS, Sep. 2008.
- [15] "LS on channel coding," 3GPP TSG RAN WG1 Meeting, R1-1715317, Prague, Czech Republic, Tech. Rep. 90, Aug. 2017.
- [16] R. Mori and T. Tanaka, "Performance of polar codes with the construction using density evolution," *IEEE Commun. Lett.*, vol. 13, no. 7, pp. 519–521, Jul. 2009.
- [17] S. ten Brink, G. Kramer, and A. Ashikhmin, "Design of low-density parity-check codes for modulation and detection," *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 670–678, May 2004.
- [18] F. Brännström, L. K. Rasmussen, and A. J. Grant, "Convergence analysis and optimal scheduling for multiple concatenated codes," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3354–3364, Aug. 2005.
- [19] A. Eslami and H. Pishro-Nik, "A practical approach to polar codes," *IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 16–20, Aug. 2011.
- [20] R. Wang and R. Liu, "A novel puncturing scheme for polar codes," *IEEE Commun. Lett.*, vol. 18, no. 12, pp. 2081–2084, Oct. 2014.
- [21] P. Trifonov and V. Miloslavskaya, "Polar subcodes," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 2, pp. 254–266, Feb. 2016.
- [22] A. G. i Fàbregas, A. Martinez, and G. Caire, "Bit-interleaved coded modulation," *Foundations and Trends*® in *Communications and Information Theory*, vol. 5, pp. 1–153, 2008.



Figure 5: biAWGN,  $k = 128, n_1^4 = (250, 250, 200, 140)$ , design SNR = (3, -1, -2.5, -3)dB



Figure 6: biAWGN, k = 848,  $n_1^4 = (1000, 1000, 1500, 800)$ , design SNR = (6.5, 1, -2, -3)dB



Figure 7: 8-ASK, k = 896,  $n_1^4 = (1200, 600, 1200, 900)$ , design SNR = (16.25, 11.25, 6.75, 5)dB