IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received November 15, 2017, accepted December 23, 2017, date of publication January 4, 2018, date of current version March 9, 2018.

Digital Object Identifier 10.1109/ACCESS.2017.2788887

Joint Source-Channel Polarization With Side Information

LIQIANG JIN, (Student Member, IEEE), AND HONGWEN YANG, (Member, IEEE)
School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

Corresponding author: Hongwen Yang(yanghong@bupt.edu.cn)

This work was supported by the National Key Technology Research and Development Program of China under Grant 2017ZX03001012-003.

ABSTRACT As an extension of source polarization and channel polarization, this paper considers joint source-channel polarization, which results in a joint source-channel coding (JSCC) scheme using a quasi-uniform systematic polar code (SPC). In this JSCC scheme, the source with side information is encoded as a systematic polar codeword and only parity bits are transmitted through the channel. The indices of systematic bits are quasi-uniform, which enable the source and the channel to be jointly polarized to either a high entropy part or a low entropy part. The analysis reveals that the quasi-uniform SPC cannot be constructed via original polar coding. To solve this problem, additional bit-swap coding is introduced to modify original polar coding and construct this kind of SPCs. The proposed JSCC scheme can asymptotically approach the information-theoretical limit. For the noiseless channel, the proposed scheme is degraded into classic Slepain-Wolf coding or lossless source coding based on parity approach.

INDEX TERMS Joint source-channel polarization, joint source-channel coding, side information, systematic polar codes.

I. INTRODUCTION

The source channel separation theorem [1] indicates that for a point-to-point communication, the separate design of source coding and channel coding is asymptotically optimal as the code length goes to infinity. However, in a practical transmission system where the delay, the complexity, and the code length are constrained, joint source-channel coding (JSCC) may be more efficient than separate source-channel coding (SSCC) [2], [3]. For example, in wireless sensor networks (WSNs), the complexity and power consumption of wireless sensors are strictly constrained. In this scenario, the performance of JSCC is usually better than, at least equals to that of SSCC. A wireless sensor adopting SSCC requires two pair of encoders and decoders for source coding and channel coding, which exacerbates the burden of the sensor. Therefore, JSCC has attracted an increasing number of attention among researchers [4]–[6].

Polar codes, invented in [7] using a technique called channel polarization, are capable of achieving the symmetric capacity of any binary-input discrete memoryless channel (B-DMC) with $O(N \log N)$ encoding and decoding complexity. The construction of polar codes has been extended to general polar kernels, non-binary alphabets, and asymmetric setting (asymmetric channels or distortion measures) [8]–[10]. The concept of source polarization was introduced in [11] as the complement of channel polarization. One immediate application of source polarization is the design of polar codes for source coding. Polar codes used for lossless/lossy and multiple-users source coding have been studied in [12]–[14]. In [15], Wang [15] have proposed a joint source-channel decoding (JSCD) scheme based on nonsystematic polar codes, where a joint decoder takes the prior information of the source into consideration at each decoding level. Obviously, JSCC/D based on nonsystematic polar codes can not achieve the information-theoretical limit under successive cancellation (SC) decoding ($O(N \log N)$ complexity) because the source and the channel can not be jointly polarized to either a high entropy part or a low entropy part (two extremes) that is the necessary condition of achieving the information-theoretical limit under SC decoding. Two practical JSCC/D schemes based on systematic polar codes (SPCs) were proposed in[16] and [17]. However, the question of whether the SPC could achieve the information-theoretical limit of JSCC remains unsolved. Actually, we can think that there are two kinds of channels in the codeword if SPCs are applied for JSCC. The positions of systematic bits directly determine whether the entropy of component channels are polarized to two extremes. Therefore, to achieve the information-theoretical limit, the positions of systematic bits have to be designed carefully.

In practical applications, the performance of polar codes is rather poor under the SC decoder, due to the finite code length. To further improve the performance, Arikan [18] proposed a belief propagation (BP) decoder and compared the performance of polar codes with Reed-Muller codes. To reduce decoding complexity and keep soft outputs available, soft cancellation (SCAN) decoder was proposed in [19]. The SC list (SCL) decoding algorithm [20] and the SC stack (SCS) decoding algorithm [21] can approach the maximum likelihood performance at a high SNR. It is worth mentioning that with the aid of cyclic redundancy check (CRC), the performance of polar codes under the SCL/SCS decoder is improved significantly, even comparable to the state-of-the-art turbo or LDPC codes.

To consider the optimality of SPCs for the JSCC problem, it is inevitable to analyze joint source-channel polarization. Therefore, we consider joint source-channel polarization with side information in this paper, which results in a JSCC scheme using a quasi-uniform SPC. Our system model is depicted in Fig. 1, where a binary source V with side information S is transmitted over a symmetric B-DMC $W(y|x)$. Our goal is to transmit this source through the channel only using a SPC and achieve the information-theoretical limit under the SC decoder. The main contributions of this paper are summarized as follows:

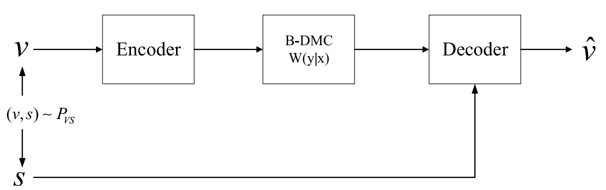


FIGURE 1. The system model of JSCC with side information.

- At first, a definition called quasi-uniform index set and its extension property are provided. Based on this definition, the positions of systematic bits of SPCs are designed carefully. Further, we introduce a class of SPCs called quasi-uniform SPCs.
- We analyze joint source-channel polarization based on a quasi-uniform SPC and prove that the source and the channel are jointly polarized to either a high entropy part or a low entropy part.
- To solve encoding problem of quasi-uniform SPCs, we propose bit-swap coding and an efficient encoding algorithm. And we prove that JSCC based on a quasi-uniform SPC can approach the information-theoretical limit of JSCC.
- The quasi-uniform SPC is extended to general binary polar kernels.

II. PRELIMINARIES

This section provides a brief review on source polarization and channel polarization. Also, the encoding of the SPC is described. All random variables (RVs) are denoted by capital letters and their samples are denoted by small letters. Some notations that will be used are shown in Table 1.

TABLE 1. Notations.

Notation	Description
$H(X)$	Entropy function
$Z(X\|Y)$	Source Bhattacharyya parameter
$Z_{\text{Bha}}(X\|Y)$	Channel Bhattacharyya parameter
$U_{\mathcal{A}}, \widetilde{U}_{\mathcal{A}}$	Information bits
$U_{\mathcal{A}^c}, \widetilde{U}_{\mathcal{A}^c}$	Frozen bits
$X_{\mathcal{B}}$	Systematic bits
$X_{\mathcal{B}^c}$	Parity bits
$Q(N, K)$	Quasi-uniform index set

A. POLAR TRANSFORMATION

Let $(X, Y) \sim P_{X,Y}$ be an arbitrary pair of RVs over $\mathcal{X} \times \mathcal{Y}$, where $\mathcal{X} = \{0, 1\}$ and $\mathcal{Y}$ is an arbitrary countable set. When source polarization is considered, Y is regarded as side information of the source X. When channel polarization is considered, Y is regarded as the output of a B-DMC $W(y|x)$ with the distribution $P_{X,Y}(x, y) = P_X(x)W(y|x)$. Let $(X_1, Y_1), (X_2, Y_2), \ldots, (X_N, Y_N)$ be $N = 2^n$ i.i.d. copies of (X, Y). In [7] and [11], X_1^N are transformed to U_1^N by

$$X_1^N = U_1^N G_N, \tag{1}$$

where $G_N = B_N F^{\otimes n} = F^{\otimes n} B_N$, B_N is a permutation matrix known as the bit-reversal operation and $F^{\otimes n}$ denotes n-order Kronecker power of $F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$. As $N \to \infty$, both $H(U_i|Y_1^N U_1^{i-1})$ and $Z(U_i|Y_1^N U_1^{i-1})$ are polarized to either 0 or 1. And for any $0 < \epsilon < \frac{1}{2}$,

$$\begin{aligned}
&\lim_{N\to\infty} \frac{|\{i : H(U_i|Y_1^N U_1^{i-1}) \in [0, \epsilon)\}|}{N} \\
&\quad = \lim_{N\to\infty} \frac{|\{i : Z(U_i|Y_1^N U_1^{i-1}) \in [0, \epsilon)\}|}{N} = 1 - H(X|Y), \\
&\lim_{N\to\infty} \frac{|\{i : H(U_i|Y_1^N U_1^{i-1}) \in (1-\epsilon, 1]\}|}{N} \\
&\quad = \lim_{N\to\infty} \frac{|\{i : Z(U_i|Y_1^N U_1^{i-1}) \in (1-\epsilon, 1]\}|}{N} = H(X|Y),
\end{aligned} \tag{2}$$

where $H(\cdot)$ is entropy function, $Z(X|Y) \in [0, 1]$ is a source Bhattacharyya parameter defined as

$$Z(X|Y) = 2 \sum_{y\in\mathcal{Y}} \sqrt{P_{XY}(0, y)P_{XY}(1, y)}. \tag{3}$$

If equiprobable X is the channel input of $W(y|x)$, $Z(X|Y)$ coincides with the channel Bhattacharyya parameter defined as

$$Z_{\text{Bha}}(X|Y) = \sum_{y\in\mathcal{Y}} \sqrt{W(y|0)W(y|1)}. \tag{4}$$

If X is nonuniformly distributed, define a virtual channel $\bar{W}_{YE|\bar{X}}(ye|\bar{x}) = W(y|\bar{x})P_{E|\bar{X}}(e|\bar{x})$ in which $E = X \oplus \bar{X}$, $\bar{X}$ is uniformly distributed and independent of (X, Y). Then $Z(X|Y)$ can be calculated by

$$\begin{aligned}
Z(X|Y) &= Z_{\text{Bha}}(\bar{X}|YE) \\
&= \sum_{y\in\mathcal{Y}, e\in\{0,1\}} \sqrt{\bar{W}_{YE|\bar{X}}(ye|0)\bar{W}_{YE|\bar{X}}(ye|1)}.
\end{aligned} \tag{5}$$

$Z(X|Y)$ and $H(X|Y)$ have the property [22]

$$Z^2(X|Y) \leq H(X|Y) \leq \log_2(1+Z(X|Y)), \quad (6)$$

and $H(X|Y) = 0 \Leftrightarrow Z(X|Y) = 0$, $H(X|Y) = 1 \Leftrightarrow Z(X|Y) = 1$. Besides, the following proposition shows that source Bhattacharyya parameters of component channels/sources can also be calculated by channel Bhattacharyya parameters.

Proposition 1 [10]: Let $E_1^N = X_1^N \oplus \bar{X}_1^N$, where $\oplus$ denotes bit-wise mod-2 addition, $\bar{X}_1^N$ are i.i.d. with equal probabilities and independent of $(X_1, Y_1), (X_2, Y_2), \ldots, (X_N, Y_N)$, then $Z(U_i|Y_1^N U_1^{i-1}) = Z_{\text{Bha}}(\bar{U}_i|Y_1^N E_1^N \bar{U}_1^{i-1})$.

When two RVs with different distributions are transformed by F, Proposition 2 gives the relationship of source Bhattacharyya parameters.

Proposition 2: Given two pairs of RVs $(X_1, Y_1) \sim P_{X_1Y_1}, (X_2, Y_2) \sim P_{X_2Y_2}$, $X_1, X_2 \in \{0, 1\}$, define $U_1 = X_1 \oplus X_2$ and $U_2 = X_2$, then

$$\begin{aligned} &Z(U_1|Y_1^2) \leq Z(X_1|Y_1) + Z(X_2|Y_2) - Z(X_1|Y_1)Z(X_2|Y_2), \\ &Z(U_2|Y_1^2U_1) = Z(X_1|Y_1)Z(X_2|Y_2). \end{aligned} \quad (7)$$

Proof: See Appendix.

B. SYSTEMATIC POLAR CODES

Let $\mathcal{A} \subseteq \{i|H(U_i|Y_1^N U_1^{i-1}) \in [0, \epsilon)\}$ denote information set. If $\mathcal{A} \subsetneq \{i|H(U_i|Y_1^N U_1^{i-1}) \in [0, \epsilon)\}$, some "good bits" in U_1^N are frozen. Let $\mathcal{B}$ denote an index set containing the indices of systematic bits, $|\mathcal{A}| = |\mathcal{B}| = K$. $\mathcal{A}^c$ and $\mathcal{B}^c$ denote the complementary sets of $\mathcal{A}$ and $\mathcal{B}$, respectively. For nonsystematic polar codes, source message is transmitted over the information bits $U_{\mathcal{A}}$. For SPCs [23], source message is transmitted over the systematic bits $X_{\mathcal{B}}$. The (N, K) SPC encoder first calculates the information bits $U_{\mathcal{A}}$ by

$$U_{\mathcal{A}} = \left(X_{\mathcal{B}} - U_{\mathcal{A}^c}G_N(\mathcal{A}^c\mathcal{B})\right)(G_N(\mathcal{A}\mathcal{B}))^{-1}, \quad (8)$$

where $G_N(\mathcal{A}\mathcal{B})$ denotes the submatrix of G_N with rows indexed by $\mathcal{A}$ and columns indexed by $\mathcal{B}$, and $U_{\mathcal{A}^c}$ are frozen bits (realized by a pseudo random sequence and known by the sender and the receiver). Then the codeword is obtained by

$$X_1^N = (U_{\mathcal{A}}G_N(\mathcal{A})) \oplus (U_{\mathcal{A}^c}G_N(\mathcal{A}^c)), \quad (9)$$

where $G_N(\mathcal{A})$ and $G_N(\mathcal{A}^c)$ denote the rows of G_N indexed by $\mathcal{A}$ and $\mathcal{A}^c$, respectively. Note that the SPC can be constructed if and only if $G_N(\mathcal{A}\mathcal{B})$ is invertible. In [23], Arikan suggested letting $\mathcal{B}$ be the image of $\mathcal{A}$ under the bit-reversal permutation to ensure the invertibility of $G_N(\mathcal{A}\mathcal{B})$.

III. JOINT SOURCE-CHANNEL POLARIZATION

This section introduces joint source-channel polarization. At first, a definition called quasi-uniform index set and its property are given. Then based on the assumption that the indices of systematic bits of the SPC are quasi-uniform, we study a class of joint source-channel polarization.

Definition 1: Suppose that there is a binary sequence L of length $N = 2^n$ that has K ones. The index set $\mathcal{B}$ contains all positions of ones in L. If $\mathcal{B}$ is the image of $\{i : N-K+1 \leq i \leq N\}$ under the bit-reversal permutation defined as

$$\mathcal{B} = \left\{ b = 1 + \sum_{i=0}^{n-1} b_i 2^i \,\middle|\, 1 + \sum_{i=0}^{n-1} b_i 2^{n-1-i} \in \{i : N-K+1 \leq i \leq N\}, \quad b_i \in \{0,1\} \right\}, \quad (10)$$

we say that $\mathcal{B}$ is a quasi-uniform index set and denote it as $\mathcal{B} = Q(N, K)$ for short.

Example 1: If $\mathcal{B}$ is $Q(16, 4)$,

$$\begin{aligned} &\{13, 14, 15, 16\} \leftrightarrow \{1100, 1101, 1110, 1111\} \\ &\xrightarrow{\text{bit-reversal}} \{0011, 1011, 0111, 1111\} \leftrightarrow \{4, 12, 8, 16\} = \mathcal{B}. \end{aligned}$$

And the sequence corresponding to $\mathcal{B}$ is $L = 0001000100010001$.

Proposition 3: After $M = 2^m$ copies of a sequence L with $\mathcal{B} = Q(N, K)$, the sequence $L_M = \underbrace{LL \ldots L}_{M}$ has the index set $\mathcal{B}_M = Q(MN, MK)$ that contains the positions of ones in L_M.

Proof: We first prove that $\mathcal{B}_2 = Q(2N, 2K)$ with respect to (w.r.t.) L_2. Assume that $b \in \mathcal{B} = Q(N, K)$ and $b = 1 + \sum_{i=0}^{n-1} b_i 2^i$ with $b_i \in \{0, 1\}$. For the left half of $L_2 = LL$, the index set containing the positions of ones is $\mathcal{B}_2^L = \mathcal{B}$. For the right half of L_2, the index set containing the positions of ones is $\mathcal{B}_2^R = \{b' = b + 2^n | b \in \mathcal{B}\}$. Obviously, we have

$$\mathcal{B}_2 = \mathcal{B}_2^L \cup \mathcal{B}_2^R = \{b' = b + b_n 2^n | b_n = 0 \text{ or } 1, b \in \mathcal{B}\}. \quad (11)$$

It is readily seen that

$$\begin{aligned} b' = b + b_n 2^n &\xrightarrow{\text{bit-reversal}} 1 + \sum_{i=0}^{n} b_i 2^{n-i} = 1 + 2\sum_{i=0}^{n-1} b_i 2^{n-1-i} \\ &+ b_n = \begin{cases} 1 + 2\sum_{i=0}^{n-1} b_i 2^{n-1-i}, & b_n = 0 \\ 1 + 2\sum_{i=0}^{n-1} b_i 2^{n-1-i} + 1, & b_n = 1 \end{cases} \end{aligned} \quad (12)$$

in which

$$1 + \sum_{i=0}^{n-1} b_i 2^{n-1-i} \in \{i : N-K+1 \leq i \leq N\} \quad (13)$$

because $b \in \mathcal{B} = Q(N, K)$. Thus the image of $\mathcal{B}_2$ under the bit-reversal permutation is $\left\{1 + \sum_{i=0}^{n} b_i 2^{n-i}\right\} = \{i : 2N - 2K + 1 \leq i \leq 2N\}$, i.e., $\mathcal{B}_2 = \left\{b' = 1 + \sum_{i=0}^{n} b_i 2^i\right\} = Q(2N, 2K)$. We can recursively prove that $\mathcal{B}_4 = Q(4N, 4K)$, $\mathcal{B}_8 = Q(8N, 8K)$,..., $\mathcal{B}_M = Q(MN, MK)$ w.r.t. $L_4, L_8, \ldots, L_M$.

Example 2: The index set $\mathcal{B}$ w.r.t. $L = 0001$ is $Q(4, 1) = \{4\}$ and the index set $\mathcal{B}_2$ w.r.t. $L_2 = LL = 00010001$ is $Q(8, 2) = \{4, 8\}$.

Putting the code construction aside, we assume that there is a (MN, MK) SPC whose indices of systematic bits are

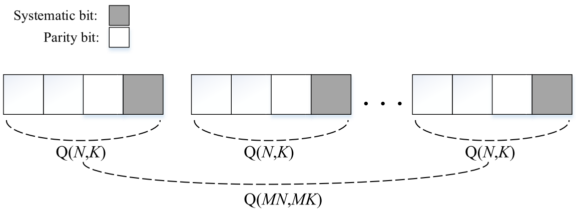


FIGURE 2. The codeword structure of $Q(MN, MK)$ SPC with $N = 4, K = 1$.

$Q(MN, MK)$, abbreviated as $Q(MN, MK)$ SPC or quasi-uniform SPC. According to Proposition 3, we know that the codeword structure of this SPC can be depicted as Fig. 2. Consider a system model in Fig. 1. The source message V_1^{MK} is encoded to a codeword X_1^{MN} by a $Q(MN, MK)$ SPC, in which $X_{\mathcal{B}} = V_1^{MK}$. Only the parity bits $X_{\mathcal{B}^c}$ are transmitted over a symmetric B-DMC $W(y|x)$ and the decoder has an access to side information S_1^{MK} w.r.t. V_1^{MK}. For simplicity of notations, let $Y_{\mathcal{B}^c}$ be channel outputs of $W(y|x)$ and $Y_{\mathcal{B}} = S_1^{MK}$. The SPC encoding rule in (8) shows that $U_{\mathcal{A}}$ are dependent on $X_{\mathcal{B}}$ and $U_{\mathcal{A}^c}$, and thus $H(U_{\mathcal{A}}|U_{\mathcal{A}^c}) = H(X_{\mathcal{B}})$. Assume that frozen bits $U_{\mathcal{A}^c}$ are i.i.d. with equal probabilities and independent of the source message $X_{\mathcal{B}}$. Then we have $|\mathcal{A}^c| = H(U_{\mathcal{A}^c}) = H(X_{\mathcal{B}^c}|X_{\mathcal{B}}) \le H(X_{\mathcal{B}^c})$ due to $H(X_{\mathcal{A}}X_{\mathcal{A}^c}) = H(X_{\mathcal{B}}X_{\mathcal{B}^c})$. Thus the parity bits $X_{\mathcal{B}^c}$ are also i.i.d. with equal probabilities, and $C(W) = 1 - H(X|Y)$ where $C(W)$ denotes the capacity of $W(y|x)$.

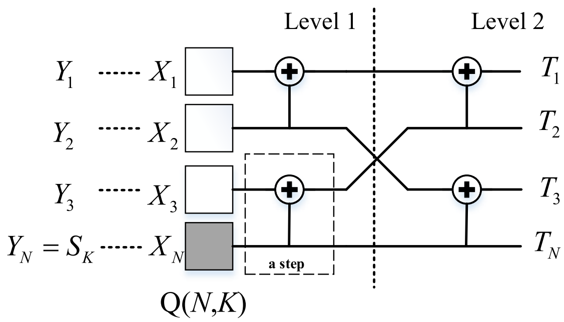


FIGURE 3. The dependency of $Q(N, K)$ SPC with $N = 4, K = 1$.

To clearly analyze the polarization of the $Q(MN, MK)$ SPC under above system model, let us describe it graphically. As shown in Fig. 3, X_1^N are transformed as T_1^N at level n and

$$
\begin{aligned}
\sum_{i=1}^{N} H(T_i|Y_1^N T_1^{i-1}) &= \sum_{i=1}^{N} H(X_i|Y_i) \\
&= (N-K)H(X|Y) + KH(V|S). \qquad (14)
\end{aligned}
$$

For the first n levels, there are some steps (each step corresponds to a polar kernel F) to transform two RVs with different distributions. However, for the level $n+1$ in Fig. 4 and the following levels ($> n+1$), each step always transforms two identical RVs, which is the original polarization method in [7]. The argument in (2) shows that, as $M \to \infty$,

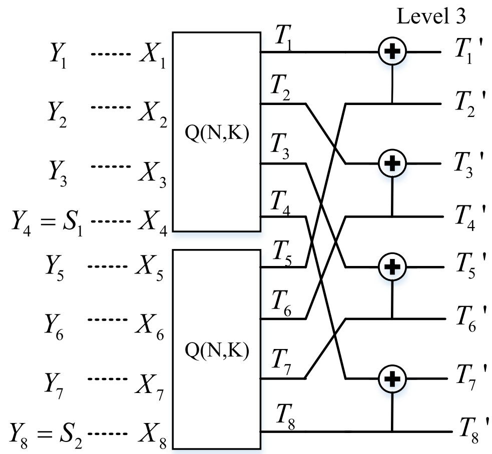


FIGURE 4. The dependency of $Q(MN, MK)$ SPC with $N = 4, K = 1, M = 2$.

for each $i \in \{i : 1 \le i \le N\}$,

$$
\sum_{j=(i-1)M+1}^{iM} H(U_j|Y_1^{MN} U_1^{j-1}) = MH(T_i|Y_1^N T_1^{i-1}), \qquad (15)
$$

and

$$
H(U_j|Y_1^{MN} U_1^{j-1}) \in [0, \epsilon) \cup (1-\epsilon, 1] \qquad (16)
$$

for any $0 < \epsilon < \frac{1}{2}$. The formal proof need to use entropy martingale and its convergence property [22], which is similar to the proof of channel polarization. The above result implies the following theorem.

Theorem 1: If a $Q(MN, MK)$ SPC exits and the system model is taken as above mentioned, for any $0 < \epsilon < \frac{1}{2}$,

$$
\begin{aligned}
&\lim_{M \to \infty} \frac{1}{MN} \left| \left\{ i : H(U_i|Y_1^{MN} U_1^{i-1}) \in [0, \epsilon) \right\} \right| \\
&\quad = 1 - \frac{KH(V|S) + (N-K)H(X|Y)}{N}, \\
&\lim_{M \to \infty} \frac{1}{MN} \left| \left\{ i : H(U_i|Y_1^{MN} U_1^{i-1}) \in (1-\epsilon, 1] \right\} \right| \\
&\quad = \frac{KH(V|S) + (N-K)H(X|Y)}{N}. \qquad (17)
\end{aligned}
$$

Proof: The previous analysis has shown that

$$
H(U_j|Y_1^{MN} U_1^{j-1}) \in [0, \epsilon) \cup (1-\epsilon, 1], \qquad (18)
$$

for any $0 < \epsilon < \frac{1}{2}$. The total entropy can be calculated by

$$
\begin{aligned}
&\sum_{i=1}^{N} \sum_{\substack{j= \\ (i-1)M+1}}^{iM} H(U_j|Y_1^{MN} U_1^{j-1}) \\
&\quad = \sum_{i=1}^{N} MH(T_i|Y_1^N T_1^{i-1}) \\
&\quad = M(N-K)H(X|Y) + MKH(V|S). \qquad (19)
\end{aligned}
$$

Therefore, the fraction of the high entropy part ($H \approx 1$) is $\frac{M(N-K)H(X|Y)+MKH(V|S)}{MN} = \frac{(N-K)H(X|Y)+KH(V|S)}{N}$, and the fraction of the low entropy part ($H \approx 0$) is $1 - \frac{(N-K)H(X|Y)+KH(V|S)}{N}$.

We can know from Proposition 2 that if two identical RVs are transformed, $Z(U_2|Y_1^2U_1) = Z^2(X_1|Y_1) = Z^2(X_2|Y_2)$. However, if two RVs with different distributions are transformed and any one of $Z(X_1|Y_1)$ and $Z(X_2|Y_2)$ approaches 0 or 1, $Z(U_1|Y_1^2) \approx \max(Z(X_1|Y_1), Z(X_2|Y_2))$, $Z(U_2|Y_1^2U_1) \approx \min(Z(X_1|Y_1), Z(X_2|Y_2))$. This indicates that a $Q(MN, MK)$ SPC goes through n slow polarization and m fast/original polarization. As $M = 2^m \to \infty$, the fast/original polarization ensures that the source and the channel are jointly polarized to either a high entropy part or a low entropy part. Thus the quasi-uniform SPC has potential to achieve the information-theoretical limit of JSCC under the SC decoder.

However, the existence and the construction problems of the quasi-uniform SPC have not been solved. Intuitively, $\mathcal{A}$ depends only on $P_{VS}(v, s)$ and $W(y|x)$, thus $G_{MN}(\mathcal{AB})$ is not necessarily invertible. It seems that the $Q(MN, MK)$ SPC can not be constructed at least via original polar coding. To address this intractable problem, an additional coding, called bit-swap, will be introduced in the next section. And we will show that based on multiple bit-swaps, the $Q(MN, MK)$ SPC can be constructed successfully.

IV. THE CONSTRUCTION OF THE QUASI-UNIFORM SPC AND JSCC SCHEME

In this section, bit-swap is first introduced to solve the problem of code construction. Then we propose an efficient encoding algorithm for the $Q(MN, MK)$ SPC, the complexity of which is less than $O(MN(1 + \log(MN)))$. Finally, we prove that the JSCC scheme based on $Q(MN, MK)$ SPCs can approach the information-theoretical limit.

A. BIT-SWAP AND CODE CONSTRUCTION

Assume that U_1^{MN} are transformed by $X_1^{MN} = U_1^{MN}G_{MN}$ to a high entropy part $\mathcal{A}^c$ and a low entropy part $\mathcal{A}$. Consider a pair of bits U_{t_1} and U_{t_2} for $t_1 \leq MN - MK < t_2 \in \mathcal{A}^c$ and $t_1 \in \mathcal{A}$. According to the original encoding rule, U_{t_2} should be frozen. However, as shown in Fig. 5(a), $\widetilde{U}_{t_1} = U_{t_1} \oplus U_{t_2}$ and $\widetilde{U}_{t_2} = U_{t_2}$, which is called bit-swap coding, and $\widetilde{U}_{t_1}$ is frozen instead. For this code, the decoder still attempts to decode U_1^{MN} and finally outputs $\widetilde{U}_1^{MN}$ according to the bijection from U_1^{MN} to $\widetilde{U}_1^{MN}$. Specifically, an original SC decoder is applied to decode U_1^{MN} and there is no difference for decoding $U_1^{t_2-1}U_{t_2+1}^{MN}$ compared to a polar code without bit-swap coding. When this SC decoder is decoding U_{t_2}, it selects a path that satisfies $U_{t_1} \oplus U_{t_2} = \widetilde{U}_{t_1}$ ($\widetilde{U}_{t_1}$ is frozen now). Note that for a polar code without bit-swap coding, the SC decoder directly selects a path by the frozen value U_{t_2}. Then the decoder outputs $\widetilde{U}_1^{MN}$ by $\widetilde{U}_i = U_i$ for $i \neq t_1$ and $\widetilde{U}_i = U_i + U_{t_2}$ for $i = t_1$. It is clear that the bit error rate (BER) $P_e(U_{t_1}) = P_e(\widetilde{U}_{t_2})$ under this decoder. It can be seen that the bit-swap coding is equivalent to adding t_1-th row to t_2-th row for G_{MN}. The generator matrix becomes $\widetilde{G}_{MN} = D_1G_{MN}$, where D_1 is a row elementary transformation matrix that performs as adding t_1-th row to t_2-th row.

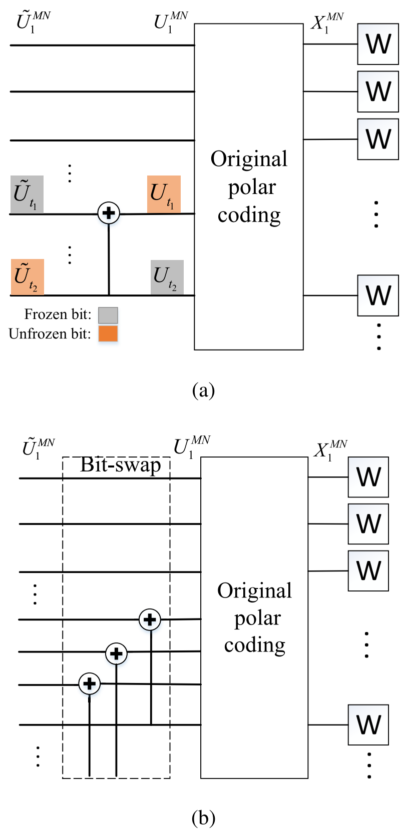


FIGURE 5. The illustration of bit-swap coding. (a) One bit-swap. (b) Multiple bit-swaps.

The previous encoding operation only has one bit-swap, let us consider a more complex case where multiple bit-swaps are implemented as shown in Fig. 5(b). Define two index sets as

$$\mathcal{C} = \{i : MN - MK + 1 \leq i \leq MN\} \cap \mathcal{A}^c,$$
$$\bar{\mathcal{C}} = \{i : 1 \leq i \leq MN - MK\} \cap \mathcal{A}. \quad (20)$$

From Theorem 1, as $M \to \infty$, we have

$$|\mathcal{C}| = \sum_{i=MN-MK+1}^{MN} H(U_i|Y_1^{MN}U_1^{i-1})$$
$$= \sum_{i=N-K+1}^{N} MH(T_i|Y_1^N T_1^i), \quad (21)$$
$$|\bar{\mathcal{C}}| = \sum_{i=1}^{MN-MK} (1 - H(U_i|Y_1^{MN}U_1^{i-1}))$$
$$= M(N-K) - \sum_{i=1}^{N-K} MH(T_i|Y_1^N T_1^i). \quad (22)$$

Here we only discuss the case of $|\bar{\mathcal{C}}| \geq |\mathcal{C}|$, since $|\bar{\mathcal{C}}| < |\mathcal{C}|$ already exceeds the theoretical limit $\frac{C(W)}{H(V|S)}$ in source symbols/channel use (we will prove it later). Next, we need to select $|\mathcal{C}|$ bits from $\bar{\mathcal{C}}$ and form $|\mathcal{C}|$ bit-swap pairs, each bit-swap pair corresponds to a row elementary transformation matrix D_i. Obviously, there are $(|\mathcal{C}|)! \cdot \binom{|\bar{\mathcal{C}}|}{|\mathcal{C}|}$ options, i.e., $(|\mathcal{C}|)! \cdot \binom{|\bar{\mathcal{C}}|}{|\mathcal{C}|}$ bit-swap schemes. The generator matrix becomes $\widetilde{G}_{MN} = DG_{MN}$, where $D = D_1 D_2 \ldots D_{|\mathcal{C}|}$ and the information set becomes $\widetilde{\mathcal{A}} = \{i : MN - MK + 1 \leq i \leq MN\}$.

Now, we are prepared to construct a $Q(MN, MK)$ SPC. Formally, source message $X_{\mathcal{B}} = V_1^{MK}$ are encoded by

$$\widetilde{U}_{\widetilde{\mathcal{A}}} = \left(X_{\mathcal{B}} - U_{\widetilde{\mathcal{A}}^c}\widetilde{G}_{MN}(\widetilde{\mathcal{A}}^c\mathcal{B})\right)\left(\widetilde{G}_{MN}(\widetilde{\mathcal{A}}\mathcal{B})\right)^{-1} = X_{\mathcal{B}}\left(\widetilde{G}_{MN}(\widetilde{\mathcal{A}}\mathcal{B})\right)^{-1}, \tag{23}$$

then the codeword is obtained by

$$X_1^{MN} = (\widetilde{U}_{\widetilde{\mathcal{A}}}\widetilde{G}_{MN}(\widetilde{\mathcal{A}})) \oplus (\widetilde{U}_{\widetilde{\mathcal{A}}^c}\widetilde{G}_{MN}(\widetilde{\mathcal{A}}^c)). \tag{24}$$

Proposition 4 shows that such $Q(MN, MK)$ SPC does exist.

Proposition 4: There exist multiple bit-swaps $D = D_1 D_2 \ldots D_{|\mathcal{C}|}$, such that $\widetilde{G}_{MN} = DG_{MN}$, and $\widetilde{G}_{MN}(\widetilde{\mathcal{A}}\mathcal{B})$ is invertible.

Proof: After multiple bit-swaps coding D, the generator matrix and the information set become $\widetilde{G}_{MN} = DF^{\otimes(m+n)}B_{MN}$ and $\widetilde{\mathcal{A}} = \{i : MN - MK + 1 \leq i \leq MN\}$, respectively. Since $\mathcal{B} = Q(MN, MK)$, the image of $\mathcal{B}$ under the bit-reversal permutation is $\{i : MN - MK + 1 \leq i \leq MN\} = \widetilde{\mathcal{A}}$, we only need to prove the reversibility of $DF^{\otimes(m+n)}(\widetilde{\mathcal{A}}\widetilde{\mathcal{A}})$. It is clear that $D = D_1 D_2 \ldots D_{|\mathcal{C}|}$ where D_i is a row elementary transformation matrix that performs as adding one row to another row below. Therefore, D, $F^{\otimes(m+n)}$, and $DF^{\otimes(m+n)}$ are all lower triangular matrixes and all diagonal entries are 1. Thus $DF^{\otimes(m+n)}(\widetilde{\mathcal{A}}\widetilde{\mathcal{A}})$ is invertible, i.e., $\widetilde{G}_{MN}(\widetilde{\mathcal{A}}\mathcal{B})$ is invertible.

For this $Q(MN, MK)$ SPC, the decoder tries to decode U_1^{MN} and finally outputs $\widetilde{U}_1^{MN}$ or X_1^{MN} by $X_1^{MN} = U_1^{MN} G_{MN}$. The specific decoding process is similar to the description for one bit-swap. The SC decoder selects a path that satisfies the parity relationship for decoding $U_i, i \in \mathcal{C}$ and there is no difference for decoding other bits. We can observe that the word error rate (WER) can be expressed by

$$P_{\mathrm{WER}} = 1 - \prod_{i \in \widetilde{\mathcal{A}}}\left(1 - P_e(\widetilde{U}_i)\right) = 1 - \prod_{i \in \mathcal{A}}\left(1 - P_e(U_i)\right) \tag{25}$$

under this decoder.

Similar to the technique in section III, it can be proven that $X_{\mathcal{B}^c}$ are i.i.d. with equal probabilities for the SPC based on $\widetilde{G}_{MN}$. However, it is unclear how the distribution of $X_{\mathcal{B}^c}$ remains unchanged after multiple bit-swaps coding. To answer this question, let us consider it reversedly. At first, we assume that parity bits $X_{\mathcal{B}^c}$ are i.i.d. with equal probabilities and the dummy channel $W(y|x)$ has the capacity $C(W) = 0$, and the perfect side information is $Y_{\mathcal{B}} = X_{\mathcal{B}}$. According to Proposition 2, we know that in Fig. 3,

$$\begin{aligned} H(T_i|Y_1^N T_1^{i-1}) &= 1 \text{ for } 1 \leq i \leq N - K, \\ H(T_i|Y_1^N T_1^{i-1}) &= 0 \text{ for } N - K + 1 \leq i \leq N. \end{aligned} \tag{26}$$

As $M \to \infty$,

$$\begin{aligned} H(U_i|Y_1^{MN} U_1^{i-1}) &= 1 \text{ for } i \in \widetilde{\mathcal{A}}^c, \\ H(U_i|Y_1^{MN} U_1^{i-1}) &= 0 \text{ for } i \in \widetilde{\mathcal{A}}. \end{aligned} \tag{27}$$

Thus $H(U_i|Y_1^{MN} U_1^{i-1}) = H(U_i|X_{\mathcal{B}} U_1^{i-1}) = 1$ for $i \in \bar{\mathcal{C}}$, then we have Proposition 5. To obtain the samples x_1^{MN} of RVs (X_1^{MN}, Y_1^{MN}) (i.e. $X_{\mathcal{B}^c}$ are i.i.d. with equal probabilities), we can first choose a value of U_i uniformly for $i \in \widetilde{\mathcal{A}}^c$, then determine U_i for $i \in \widetilde{\mathcal{A}}$ based on $Y_1^{MN} U_1^{i-1}$, finally transform U_1^{MN} to X_1^{MN} by $X_1^{MN} = U_1^{MN} G_{MN}$. Since $U_{\widetilde{\mathcal{A}}} = X_{\mathcal{B}}\left(G_{MN}(\widetilde{\mathcal{A}}\mathcal{B})\right)^{-1}$, we only need to let $U_{\widetilde{\mathcal{A}}^c}$ be i.i.d. RVs with equal probabilities.

Proposition 5: If $i \in \bar{\mathcal{C}}$, then $H(U_i|X_{\mathcal{B}} U_1^{i-1}) \in (1 - \epsilon, 1]$ for any $0 < \epsilon < \frac{1}{2}$.

Recall that $\widetilde{U}_{t_1} = U_{t_1} \oplus U_{t_2}$, and $\widetilde{U}_{t_2} = U_{t_2}$ for $t_1 \in \bar{\mathcal{C}}, t_2 \in \mathcal{C}$ by the definition of bit-swap. The value of U_{t_2} remains since $\widetilde{U}_{t_2} = U_{t_2}$. $\widetilde{U}_{t_1}$ is frozen as an independent RV with equal probabilities, and thus $U_{t_1} = \widetilde{U}_{t_1} \oplus \widetilde{U}_{t_2}$ is also independent and equiprobable. Essentially, bit-swap does not change the joint distribution of U_1^{MN}, since $P(U_{\widetilde{\mathcal{A}}^c}|X_{\mathcal{B}}) = \prod_{i=1}^{M(N-K)} P(U_i) = (\frac{1}{2})^{M(N-K)}$. That's why $X_{\mathcal{B}^c}$ are still i.i.d. with equal probabilities after bit-swaps.

B. THE EFFICIENT ENCODING ALGORITHM

So far we have shown that the $Q(MN, MK)$ SPC does exist and the way to construct such code is presented. However, the encoding in (23) involves the inverse of $\widetilde{G}_{MN}(\widetilde{\mathcal{A}}\mathcal{B})$ that is a high density matrix: the inverse of high density matrix $\widetilde{G}_{MN}(\widetilde{\mathcal{A}}\mathcal{B})$ is very difficult to find when the dimension is large. For this reason, an efficient encoding algorithm is proposed in the following.

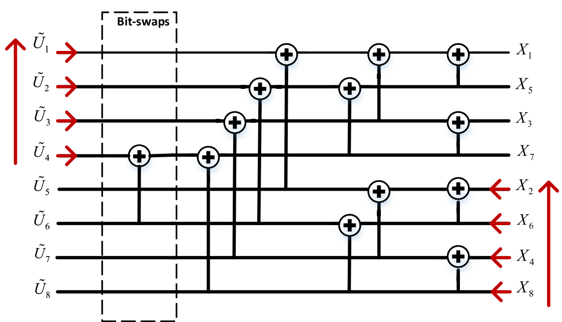


FIGURE 6. The efficient encoder of $Q(MN, MK)$ SPC, $N = 2, K = 1, M = 4, \mathcal{A} = \{4, 5, 7, 8\}, \widetilde{\mathcal{A}} = \{5, 6, 7, 8\}$.

The factor graph representation for the transformation $DF^{\otimes(m+n)}$ is illustrated in Fig. 6. Since $\mathcal{B} = Q(MN, MK)$, the values of systematic bits $X_{\mathcal{B}} = V_1^{MK}$ are assigned on the last MK nodes in the right-most column. The values of frozen bits $U_{\widetilde{\mathcal{A}}^c}$ are assigned on the first $M(N - K)$ nodes in the left-most column. The calculation starts row by row from the bottom to the top. As indicated by the red arrow, at each horizontal connection, the calculation starts from a side that has been assigned a value. It can be seen that the encoding complexity is $O(MN\log(MN)+|\mathcal{C}|) < O(MN(\log(MN)+1))$.

C. JSCC SCHEME

According to information theory, the theoretical limit of JSCC is $\frac{C(W)}{H(V|S)}$ in source symbols/channel use [24]. A codeword of the $Q(MN, MK)$ SPC occupies $M(N-K)$ channel uses (only parity bits are transmitted over the channel). To reliably transmit the source message, the code rate R should satisfy $\frac{C(W)}{H(V|S)} \geq \frac{MK}{MN-MK} = \frac{R}{1-R}$. Then we have $R \leq R_{\text{limit}} = \frac{C(W)}{C(W)+H(V|S)}$. From Lemma 1 below, asymptotically, we know that the code rate R can arbitrarily approach R_{limit}. If $R \leq R_{\text{limit}}$, it can be proven that $R \leq 1 - \frac{KH(V|S)+(N-K)H(X|Y)}{N}$, which means that $\widetilde{\mathcal{A}} \subseteq \{i|P_e(\widetilde{U}_i) \in [0,\epsilon)\}$ and $P_{\text{WER}} \to 0$ under the SC decoder. If $\widetilde{\mathcal{A}} \subsetneq \{i|P_e(\widetilde{U}_i) \in [0,\epsilon)\}$, some "good bits" are frozen and wasted. In section IV-A, we assume $|\bar{\mathcal{C}}| \geq |\mathcal{C}|$, which is equivalent to $\left|\{i|H(U_i|Y_1^{MN}U_1^{i-1}) \in [0,\epsilon)\}\right| \geq MK$. Thus $|\bar{\mathcal{C}}| \geq |\mathcal{C}|$ holds if and only if $R \leq R_{\text{limit}}$. Then JSCC based on a $Q(MN, MK)$ SPC asymptotically approaches the information-theoretical limit.

Lemma 1: For any $\epsilon > 0$, there exists a positive number $N(\epsilon) = 2^{n(\epsilon)}$, such that $|R - R_{\text{limit}}| \leq \epsilon$.

Proof: Let $n = \lceil \log_2 \frac{1}{\epsilon} \rceil$, $N = 2^n$ and $K = \lfloor NR_{\text{limit}} \rfloor$, it can be verified that $R = \frac{MK}{MN} = \frac{K}{N}$ and $|R - R_{\text{limit}}| < \frac{1}{N} \leq \epsilon$.

V. EXTENSION TO THE GENERAL BINARY KERNEL

The polarization can be generalized to the general binary kernel F [7]. If $l \times l$ matrix F is a binary kernel of polar codes, the recursive definition of G_N is

$$G_N = (I_{N/l} \otimes F)R_N(I_l \otimes G_{N/l}), \tag{28}$$

where $G_1 = I_1$, $N = l^n$, and R_N is a permutation matrix known as the l-ary reverse shuffle operation. R_N sorts the components w.r.t. mod-l residue classes of indices and ensures that l identical RVs are transformed by F. Since $(I_{N/l} \otimes F)R_N = R_N(F \otimes I_{N/l})$,

$$\begin{aligned} G_N &= R_N(F \otimes G_{N/l}) \\ &= R_N(F \otimes (R_{N/l}(F \otimes G_{N/l^2}))) \\ &= R_N(I_l \otimes R_{N/l})(F^{\otimes 2} \otimes G_{N/l^2}) \\ &\dots \\ &= B_N F^{\otimes n}, \end{aligned} \tag{29}$$

where $B_N = R_N(I_l \otimes R_{N/l})(I_{l^2} \otimes R_{N/l^2}) \dots (I_{N/l} \otimes R_l)$ is a permutation matrix known as the l-ary bit-reversal operation. It can been proven that G_N is a l-ary bit-reversal invariant matrix, $G_N = B_N F^{\otimes n} = F^{\otimes n} B_N$. The proof is similar to the case of $l = 2$, and thus the detailed derivation is omitted here.

Assuming that $l \times l$ matrix F is a binary kernel of polar codes, we can repeatedly permute the columns and add a row above it while keep the partial distance sequence unchanged [9]. Finally, F can be converted to a lower triangular matrix, known as the standard form of the polar kernel, and all diagonal entries are 1. Obviously, bit-swap can also be generalized to polar codes with a standard form of binary polar kernel. The definition of $Q(N, K)$ becomes that $\mathcal{B}$ is the image of $\{i : N-K+1 \leq i \leq N\}$ under the l-ary bit-reversal permutation. Based on the redefined quasi-uniform index set and multiple bit-swaps, the $Q(MN, MK)$ SPC with the general binary kernel can also be constructed. The error probability of this SPC under the SC decoder is $O\left(\beta(n)2^{-l^{mE(F)}}\right)$, where $E(F)$ is the error exponent of F, and $\beta(n)$ is related to n slow polarization. If $F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, $E(F) = \frac{1}{2}$. It has been shown in [8] that $E(F)$ increses when a larger l is considered. We deduce that the proposed JSCC has a better performance when the $Q(MN, MK)$ SPC is constructed with a larger polar kernel.

VI. SIMULATION RESULTS

A. THE UPPER BOUND

We first investigate the upper bound of $Q(MN, MK)$ SPCs under the SC decoder, which is obtained by

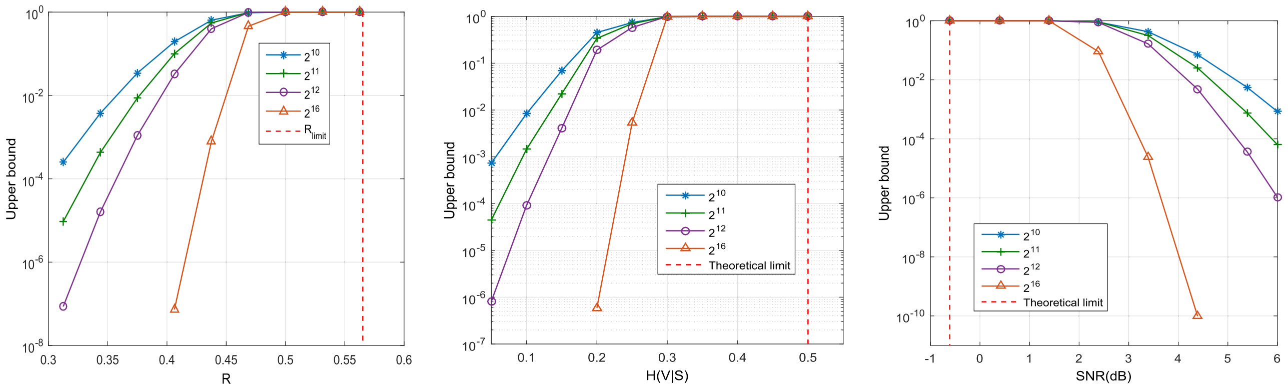
$$\begin{aligned} P_{\text{WER}} &= 1 - \prod_{i \in \widetilde{\mathcal{A}}}(1 - P_e(\widetilde{U}_i)) = 1 - \prod_{i \in \mathcal{A}}(1 - P_e(U_i)) \\ &\leq 1 - \prod_{i \in \mathcal{A}}(1 - \frac{1}{2}Z(U_i|Y_1^{MN}U_1^{i-1})). \end{aligned} \tag{30}$$

In the simulation, $P_{S|V}(s|v)$ is considered as a binary symmetric channel (BSC) with a crossover probability q, V is uniformly distributed, and $W(y|x)$ is a BI-AWGN channel. According to Proposition 1, $Z(U_i|Y_1^{MN}U_1^{i-1})$ can be calculated by $Z_{\text{Bha}}(\bar{U}_i|Y_1^{MN}E_1^{MN}\bar{U}_1^{i-1})$. Due to the symmetry of $P_{V|S}$ and $W(y|x)$, $Z_{\text{Bha}}(\bar{U}_i|Y_1^{MN}E_1^{MN}\bar{U}_1^{i-1})$ can be estimated with the method in [25]. Since the complexity of this method is controllable, we can estimate $Z_{\text{Bha}}(\bar{U}_i|Y_1^{MN}E_1^{MN}\bar{U}_1^{i-1})$ with a large M to observe the asymptotic performance.

Fig. 7 shows the upper bound versus R, $H(V|S)$, and the SNR of BI-AWGN channel, respectively. The code rate R of $Q(MN, MK)$ SPC varies with $\frac{K}{N} = \frac{10}{32}, \frac{11}{32}, \dots, \frac{18}{32}$. The labels of "$2^{10}$", "$2^{11}$","$2^{12}$","$2^{16}$" stand for the code lengths MN with $M = 32, 64, 128, 2048$, respectively. The red dashed lines labeled "R_{limit}" and "Theoretical limit" denote the theoretical limits of reliably reconstructing the source for corresponding conditions. The simulation results show that the upper bounds become better and get close to the theoretical limits with increasing M. If $Z_{\text{Bha}}(\bar{U}_i|Y_1^{MN}E_1^{MN}\bar{U}_1^{i-1})$ is perfectly estimated, the curve will tend to the theoretical limit (the red dashed line) as M goes to infinity.

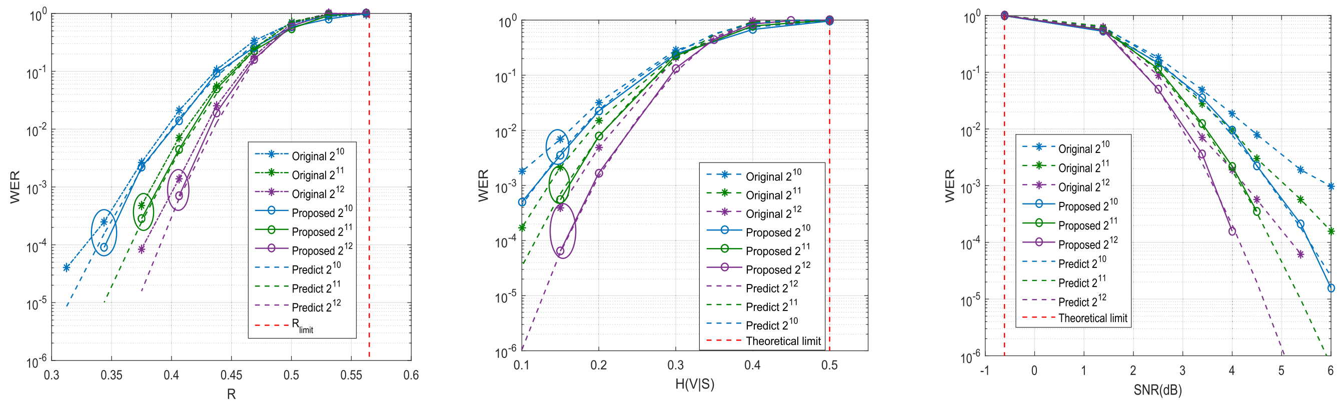
B. THE PERFORMANCE UNDER THE SC DECOER

Next, we investigate WERs of $Q(MN, MK)$ SPCs under the SC decoder. Although the density evolution (DE) [26] is impractical for the polar code with a long code length due to high complexity, $P_e(U_i)$ can be accurately estimated via DE for the short or moderate code length. Therefore, in the following simulation, the $Q(MN, MK)$



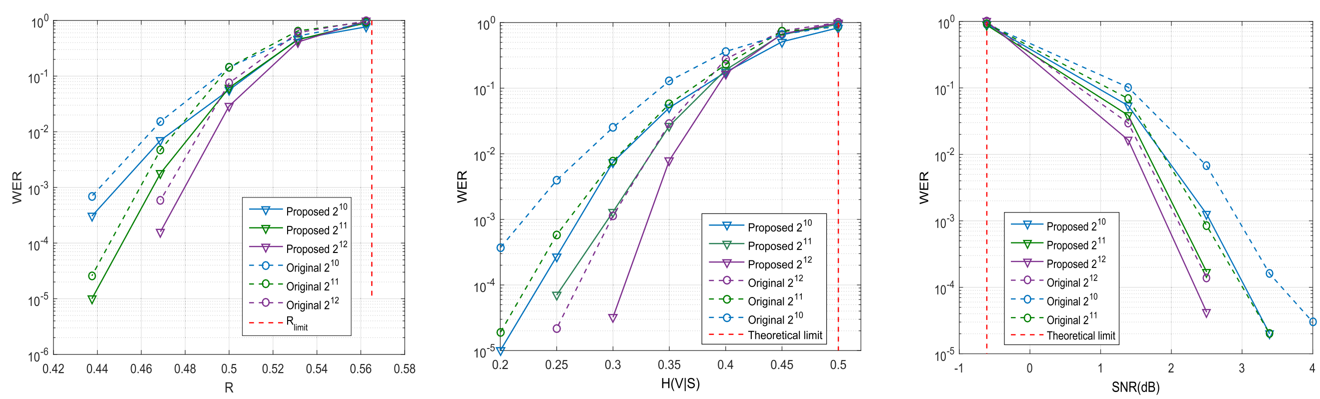
(a) (b) (c)

FIGURE 7. The upper bound of the WER versus R, $H(V|S)$, and SNR. (a) BSC(0.11), BI-AWGN(2.1dB). (b) $R = \frac{18}{32}$, BI-AWGN(2.1dB). (c) $R = \frac{15}{32}$, BSC(0.11).



(a) (b) (c)

FIGURE 8. The WER versus R, $H(V|S)$, and SNR under the SC decoder. (a) BSC(0.11), BI-AWGN(2.1dB). (b) $R = \frac{18}{32}$, BI-AWGN(2.1dB). (c) $R = \frac{15}{32}$, BSC(0.11).



(a) (b) (c)

FIGURE 9. The WER versus R, $H(V|S)$, and SNR under the CA-SCL decoder. (a) BSC(0.11), BI-AWGN(2.1dB). (b) $R = \frac{18}{32}$, BI-AWGN(2.1dB). (c) $R = \frac{15}{32}$, BSC(0.11).

SPC is constructed via DE. As a contrast, the original SPC [23] is constructed via Gaussian approximation [27] over a BI-AWGN channel with the capacity $C_{equ} = R \cdot (1 - H(V|S)) + (1 - R) \cdot C(W)$. Fig. 8 shows the WER performance of two different schemes and the performance prediction versus R, $H(V|S)$, and SNR under the

$$
\begin{aligned}
Z(U_1|Y_1^2) &= 2\sum_{y_1^2}\sqrt{\sum_{u_2}P_{XY}(u_2,y_1)P_{XY}(u_2,y_2)\sum_{v_2}P_{XY}(v_1\oplus 1,y_1)P_{XY}(v_2,y_2)}\\
&= 2\sum_{y_1^2}\sqrt{P_{XY}(0,y_1)P_{XY}(0,y_2)+P_{XY}(1,y_1)P_{XY}(1,y_2)}\sqrt{P_{XY}(1,y_1)P_{XY}(0,y_2)+P_{XY}(0,y_1)P_{XY}(1,y_2)}\\
&\leq 2\sum_{y_1^2}\left[\sqrt{P_{XY}(0,y_1)P_{XY}(0,y_2)}+\sqrt{P_{XY}(1,y_1)P_{XY}(1,y_2)}\right]\left[\sqrt{P_{XY}(1,y_1)P_{XY}(0,y_2)}+\sqrt{P_{XY}(0,y_1)P_{XY}(1,y_2)}\right]\\
&\quad -4\sum_{y_1^2}\sqrt{P_{XY}(0,y_1)P_{XY}(1,y_1)P_{XY}(0,y_2)P_{XY}(1,y_2)}\\
&= P_{X_2}(0)Z(X_1|Y_1)+P_{X_1}(0)Z(X_2|Y_2)+P_{X_1}(1)Z(X_2|Y_2)+P_{X_2}(1)Z(X_1|Y_1)-Z(X_1|Y_1)Z(X_2|Y_2)\\
&= Z(X_1|Y_1)+Z(X_2|Y_2)-Z(X_1|Y_1)Z(X_2|Y_2)\\
Z(U_2|Y_1^2U_1) &= 2\sum_{y_1^2,u_1}\sqrt{P_{XY}(u_1,y_1)P_{XY}(0,y_2)P_{XY}(u_1\oplus 1,y_1)P_{XY}(1,y_2)}\\
&= \sum_{y_1,u_1}\sqrt{P_{XY}(u_1,y_1)P_{XY}(u_1\oplus 1,y_1)}\cdot 2\sum_{y_2}\sqrt{P_{XY}(0,y_2)P_{XY}(1,y_2)}\\
&= Z(X_1|Y_1)Z(X_2|Y_2). \qquad (32)
\end{aligned}
$$

SC decoder, respectively. The curves labeled "Proposed" represent JSCC schemes using $Q(MN, MK)$ SPCs. The curves labeled "Predict" represent the predictions of WERs through DE and (30). And the curves labeled "Original" represent JSCC schemes using original SPCs. In general, the WER performance coincides with the prediction except for some minor differences. This results nearly verify (25) because P_{WER} is estimated by $P_e(U_i)$ and the simulation result is about $\widetilde{U}_i$. Besides, from these figures, we can see that the quasi-uniform SPC outperforms the original SPC.

C. THE PERFORMANCE UNDER THE CA-SCL DECODER

Although the upper bound shows the excellent asymptotic performance under the SC decoder. For the short or moderate code length, the performance is not satisfied. Thus the WER performance under the CRC-aid SCL (CA-SCL) decoder is presented in Fig. 9. For the $Q(MN, MK)$ SPC, the systematic bits contain 16 bits CRC, i.e., the length of source message is $MK - 16$. At the decoder, the log likelihood ratios (LLRs) of 16 bits CRC are initialized as 0 and the list size is set as 32. We can see that compared with the SC decoder, the performance is improved significantly and the quasi-uniform SPC also outperforms the original SPC.

VII. CONCLUSION

The quasi-uniform SPC has been introduced in this paper. As an extension of source polarization and channel polarization, we have considered a class of joint source-channel polarization based on a $Q(MN, MK)$ SPC. The indices of systematic bits are quasi-uniform, which enable the source and the channel to be jointly polarized to either a high entropy part with the fraction $RH(V|S) + (1 - R)H(X|Y))$ or a low entropy part with the fraction $1 - RH(V|S) - (1 - R)H(X|Y)$. Through the reasonable choice of N and K, JSCC based on a $Q(MN, MK)$ SPC asymptotically approaches the information-theoretical limit under the SC decoder.

To solve the encoding problem of the quasi-uniform SPC, bit-swap coding and an efficient encoding algorithm have been proposed. The complexity of efficient encoding algorithm is less than $O((MN)(\log(MN) + 1))$. When the transmission channel is noiseless, the proposed scheme is degraded into classic source coding (e.g. Slepain-Wolf coding, lossless source coding) based on parity approach. Also, the $Q(MN, MK)$ SPC has been extended to general binary polar kernels.

APPENDIX
THE PROOF OF PROPOSITION 2

By the definition,

$$
\begin{aligned}
P_{Y_1^2U_1}(y_1^2u_1) &= \sum_{u_2}P_{XY}(u_1\oplus u_2,y_1)P_{XY}(u_2,y_2)\\
P_{Y_1^2U_1U_2}(y_1^2u_1u_2) &= P_{XY}(u_1\oplus u_2,y_1)P_{XY}(u_2,y_2), \qquad (31)
\end{aligned}
$$

then we have (32), as shown at the top of this page.

REFERENCES

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 4, pp. 623–656, Oct. 1948.
[2] Y. Zhong, F. Alajaji, and L. L. Campbell, "Source side information can increase the joint source-channel coding error exponent," in *Proc. 10th Can. Workshop Inf. Theory (CWIT)*, Edmonton, AB, Canada, 2007, pp. 140–143.
[3] Y. Zhong, F. Alajaji, and L. L. Campbell, "On the joint source-channel coding error exponent for discrete memoryless systems," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1450–1468, Apr. 2006.
[4] J. Garcia-Frias, "Joint source-channel decoding of correlated sources over noisy channels," in *Proc. Data Compress. Conf. (DCC)*, Snowbird, UT, USA, 2001, pp. 283–292.
[5] F. Daneshgaran, M. Laddomada, and M. Mondin, "LDPC-based channel coding of correlated sources with iterative joint decoding," *IEEE Trans. Commun.*, vol. 54, no. 4, pp. 577–582, Apr. 2006.

[6] A. Aaron and B. Girod, "Compression with side information using turbo codes," in *Proc. Data Compress. Conf. (DCC)*, 2002, pp. 252–261.
[7] E. Arıkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3051–3073, Jul. 2009.
[8] H.-P. Lin, S. Lin, and K. A. S. Abdel-Ghaffar, "Linear and nonlinear binary kernels of polar codes of small dimensions with maximum exponents," *IEEE Trans. Inf. Theory*, vol. 61, no. 10, pp. 5253–5270, Oct. 2015.
[9] R. Mori and T. Tanaka, "Channel polarization on q-ary discrete memoryless channels by arbitrary kernels," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, USA, Jun. 2010, pp. 894–898.
[10] J. Honda and H. Yamamoto, "Polar coding without alphabet extension for asymmetric models," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 7829–7838, Dec. 2013.
[11] E. Arikan, "Source polarization," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, USA, Jun. 2010, pp. 899–903.
[12] H. S. Cronie and S. B. Korada, "Lossless source coding with polar codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, USA, Jun. 2010, pp. 904–908.
[13] S. B. Korada and R. L. Urbanke, "Polar codes are optimal for lossy source coding," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1751–1768, Apr. 2010.
[14] S. B. Korada and R. Urbanke, "Polar codes for slepian-wolf, Wyner-Ziv, and Gelfand-Pinsker," in *Proc. IEEE Inf. Theory Workshop Inf. Theory (ITW, Cairo)*, Cairo, Egypt, 2010, pp. 1–5.
[15] Y. Wang, M. Qin, K. R. Narayanan, A. Jiang, and Z. Bandic, "Joint source-channel decoding of polar codes for language-based sources," in *Proc. IEEE Global Commun. Conf.*, Washington, DC, USA, Dec. 2016, pp. 1–6.
[16] C. Yaacoub and M. Sarkis, "Systematic polar codes for joint source-channel coding in wireless sensor networks and the Internet of Things," in *Proc. Comput. Sci.*, vol. 110, pp. 266–273, Jan. 2017.
[17] L. Jin, P. Yang, and H. Yang, "Distributed joint source-channel decoding using systematic polar codes," *IEEE Commun. Lett.*, vol. 22, no. 1, pp. 49–52, Jan. 2018.
[18] E. Arıkan, "A performance comparison of polar codes and Reed-Muller codes," *IEEE Commun. Lett.*, vol. 12, no. 6, pp. 447–449, Jun. 2008.
[19] U. U. Fayyaz and J. R. Barry, "Polar codes for partial response channels," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Budapest, Hungary, Jun. 2013, pp. 4337–4341.
[20] I. Tal and A. Vardy, "List decoding of polar codes," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2213–2226, May 2015.
[21] K. Niu and K. Chen, "Stack decoding of polar codes," *Electron. Lett.*, vol. 48, no. 12, pp. 695–697, Jun. 2012.
[22] E. Sasoglu, "Polar coding theorems for discrete systems," Ph.D dissertation, EPFL, Suisse, Lausanne, Switzerland, 2011.
[23] E. Arıkan, "Systematic polar coding," *IEEE Commun. Lett.*, vol. 15, no. 8, pp. 860–862, Aug. 2011.
[24] S. Shamai and S. Verdu, "Capacity of channels with uncoded-message side-information," in *Proc. IEEE Int. Symp. Inf. Theory*, Whistler, BC, USA, Sep. 1995, p. 7.
[25] I. Tal and A. Vardy, "How to construct polar codes," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6562–6582, Oct. 2013.
[26] 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.
[27] P. Trifonov, "Efficient design and decoding of polar codes," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3221–3227, Nov. 2012.

LIQIANG JIN is currently pursuing the Ph.D. degree with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China. His current research interests include information theory, channel codes, and joint source-channel codes.

HONGWEN YANG was born in 1964. He is currently the Director with the Wireless Communication Center, School of Information and Communication Engineering, Beijing University of Posts and Telecommunications. His research mainly concentrates on wireless physical aspects, including modulation and channel codes, MIMO, CDMA, OFDM, etc.