Low Complexity Successive Cancellation List Decoding of U-UV Codes

Chen Wenhao[1,2], Chen Li[1,2,*], Lin Jingyu[1,2], Zhang Huazi[3]

[1] School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou 510006, China
[2] Guangdong Province Key Laboratory of Information Security Technology, Guangzhou 510006, China
[3] Hangzhou Research Center, Huawei Technologies Co., Ltd., Hangzhou 310052, China
[*] The corresponding author, email: chenli55@mail.sysu.edu.cn

Cite as: W. Chen, L. Chen, *et al*., "Low complexity successive cancellation list decoding of u-uv codes," *China Communications*, 2025, vol. 22, no. 1, pp. 41-60. **DOI**: 10.23919/JCC.ja.2024-0052

Abstract: Constituted by BCH component codes and its ordered statistics decoding (OSD), the successive cancellation list (SCL) decoding of U-UV structural codes can provide competent error-correction performance in the short-to-medium length regime. However, this list decoding complexity becomes formidable as the decoding output list size increases. This is primarily incurred by the OSD. Addressing this challenge, this paper proposes the low complexity SCL decoding through reducing the complexity of component code decoding, and pruning the redundant SCL decoding paths. For the former, an efficient skipping rule is introduced for the OSD so that the higher order decoding can be skipped when they are not possible to provide a more likely codeword candidate. It is further extended to the OSD variant, the box-and-match algorithm (BMA), in facilitating the component code decoding. Moreover, through estimating the correlation distance lower bounds (CDLBs) of the component code decoding outputs, a path pruning (PP)-SCL decoding is proposed to further facilitate the decoding of U-UV codes. In particular, its integration with the improved OSD and BMA is discussed. Simulation results show that significant complexity reduction can be achieved. Consequently, the U-UV codes can outperform the cyclic redundancy check (CRC)-polar codes with a similar decoding complexity.

Keywords: ordered statistics decoding; successive cancellation list decoding; U-UV codes

I. INTRODUCTION

Future communication systems will realize ultra reliable and low-latency information transmission, in which the competent short-to-medium length channel codes will play a vital role. The Bose-Chaudhuri-Hocquenghem (BCH) codes [1, 2], the tail-biting convolutional codes [3], the cyclic redundancy check (CRC)-polar codes [4–7], and the more recent polarization adjusted convolutional (PAC) codes [8] are known to be good candidates in the short-to-medium length regime [9]. For BCH codes, the ordered statistics decoding (OSD) can achieve a near maximum likelihood (ML) decoding performance, but inherits a complexity that grows exponentially with its decoding order [10]. The CRC-polar codes can also achieve a near ML performance with the successive cancellation list (SCL) decoding [5, 6]. Meanwhile, it is well known that polar codes can achieve capacity of the binary input symmetric discrete memoryless channel when the codeword length is sufficiently large [4]. However, channel polarization remains incomplete for short-to-medium length polar codes. There exists a significant portion of subchannels without a polarized

Received: Jan. 19, 2024
Revised: Aug. 21, 2024
Editor: Xiao Lixia

capacity, which cannot be adequately utilized, leading to a capacity loss. To overcome this incomplete polarization, outer concatenation is introduced for the polar codes, including the parity-check-concatenated (PCC) polar codes [11–15] and the above mentioned PAC codes [8]. For ther latter, the outer convolutional transform helps leverage the transmission rates of the unpolarized subchannels.

Recently, the U-UV structural codes were proposed as another solution for overcoming the incomplete polarization [16–19]. It is a re-exploration of the classic Plotkin codes [20]. In this structure, the U codes and V codes are component codes of equal length and concatenated in a $(\mathrm{U} \,|\, \mathrm{U} + \mathrm{V})$ recursive manner. For succinctness, we use U-UV to refer this $(\mathrm{U} \,|\, \mathrm{U} + \mathrm{V})$ coding structure. Constituted by BCH component codes, the U-UV code is also known as the generalized BCH-polar concatenated codes [21]. This structural coding results in polarized subchannel capacities. Each of the subchannels conveys a component codeword. Based on this, the component code rates can be allocated accordingly [17, 18, 21]. Compared with CRC-polar codes, it does not rely on complete subchannel capacity polarization. Moreover, its decoding parallelism can be realized through the component code decoding, resulting in a greater potential of achieving a low decoding latency than the CRC-polar codes. It has been shown that in the short-to-medium length regime, SCL decoding of the U-UV codes can outperform that of the polar codes, but it is realized at the cost of decoding complexity [17]. Hence, the primary motivation of this work is to reduce the SCL decoding complexity.

In the SCL decoding of U-UV codes, the component codes are decoded by the OSD. The OSD can yield a near ML decoding performance for BCH component codes. However, its complexity grows exponentially with the decoding order, which will dominate the overall SCL decoding complexity. There exist several approaches to reduce the OSD complexity, e.g. by utilizing the skipping and stopping rules [22–29] and through the box-and-match algorithm (BMA) [30]. The OSD variants utilizing the constraint of the parity-check matrix have been proposed in [31, 32]. Meanwhile, OSD complexity reduction can also be realized through reducing the complexity of Gaussian elimination (GE) that yields the systematic generator matrix of the code [33–35]. Besides, a hybrid SCL decoding that combines the algebraic decoding and the OSD for the component codes has been proposed for U-UV codes [36]. On the other hand, the list decoding feature of SCL decoding also incurs a high complexity. For this, avoiding the expansion of unpromising SCL decoding paths is another approach to realize the complexity reduction [37]. These path pruning techniques have been investigated for the SCL decoding of polar codes [38–42].

This paper proposes the low complexity SCL decoding for U-UV codes. The major contributions of this work are summarized as follows:

- The order skipping (OS)-OSD is proposed for reducing the component code decoding complexity. Utilizing likelihood of received symbols over the least reliable positions (LRPs), the *a posteriori* lower bound of the correlation distance between a codeword estimation and the received symbols is characterized. With this, the higher order decoding can be skipped when they are not possible to provide a more likely codeword candidate. Through eliminating the redundant match operations, the skipping rule is further extended to the BMA, resulting in the low complexity OS-BMA.

- To eliminate the unpromising SCL decoding paths, the path pruning (PP)-SCL decoding is proposed. Integrating this PP-SCL decoding with the OS-OSD and OS-BMA, the low complexity SCL decoding is further proposed for U-UV codes. Exploiting the correlation distance distribution of component code decoding outputs, the *a priori* correlation distance lower bound (CDLB) that is obtained before conducting the component code decoding can be estimated. Subsequently, the SCL decoding path metric lower bound can be derived, which helps identify the unpromising decoding paths to be further pruned.

- Complexity of the proposed low complexity SCL decoding is analyzed, including that of the OS-OSD and OS-BMA. Utilizing BCH codes as the component codes, extensive simulations on the proposed low complexity SCL decoding have been conducted. Simulation results show that the low complexity SCL decoding can achieve a significant complexity reduction for U-UV codes with negligible performance loss. Moreover, with this low complexity SCL decoding, the U-UV codes can outperform the incumbent CRC-polar codes with a similar decoding complexity. It demonstrates the candidacy of the proposed U-UV codes as one of the promising short-to-medium length channel codes.

The rest of this paper is organized as follows. Section II introduces the U-UV codes and its SCL decoding. The OSD and BMA are also introduced. Section III proposes the OS-OSD and OS-BMA. Section IV proposes the PP-SCL decoding integrated with the OS-OSD and OS-BMA. Section V analyzes the decoding complexity. Section VI presents our simulation results and discussions. Finally, Section VII concludes the paper.

II. PRELIMINARIES

This section introduces the U-UV code construction and its successive cancellation list (SCL) decoding. Moreover, the ordered statistic decoding (OSD) and its complexity reducing variant, i.e., box-and-match algorithm (BMA), will be introduced for decoding the component codes.

2.1 U-UV Codes

Consider the U code and V code are two binary block codes of length n with dimension k_{U} and k_{V}, respectively. A single level U-UV code of length $2n$ and dimension $k_{\mathrm{U}} + k_{\mathrm{V}}$ is constructed as [20]

$$\mathcal{C}_{\text{U-UV}} = \{(\boldsymbol{c}_{\mathrm{U}} \,|\, \boldsymbol{c}_{\mathrm{U}} + \boldsymbol{c}_{\mathrm{V}}) : \boldsymbol{c}_{\mathrm{U}} \in \mathcal{C}_{\mathrm{U}}, \boldsymbol{c}_{\mathrm{V}} \in \mathcal{C}_{\mathrm{V}}\}, \quad (1)$$

where $\mathcal{C}_{\mathrm{U}}$ and $\mathcal{C}_{\mathrm{V}}$ denote the codebooks of the U code and V code, respectively, and $\boldsymbol{c}_{\mathrm{U}}$ and $\boldsymbol{c}_{\mathrm{V}}$ are their codewords. When more component codes are involved, this construction can be extended recursively, resulting in a larger U-UV code with multilevel construction. Figure 1 illustrates the construction of an H-level U-UV code. At level-0, there are $\gamma = 2^H$ component codes of length n. They are coupled in the U-UV manner as in (1) to construct the U-UV codes of level-1. The U-UV codewords of level-1 can again be coupled to construct the U-UV codewords of level-2. This U-UV construction can be extended recursively until level-H, yielding an H-level U-UV code of length $N = \gamma n$. By specifying the component codes at level-0, the U-UV code can also be seen as a generalized concatenated code (GCC) with inner polar codes [21, 43–46]. E.g., in this work, primitive BCH codes are utilized as the component codes, and the U-UV codes can be interpreted as generalized BCH-polar concatenated codes.

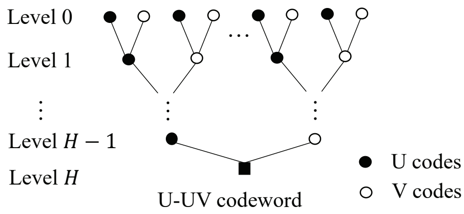

Figure 1. *Construction of an H-level U-UV code.*

Inheriting the polarization effect [4], this U-UV code construction results in γ subchannels with polarized capacities at level-0. They convey the component codes. The component code rates can be designed based on several rate allocation strategies [17, 18, 21]. In this work, the U-UV codes are designed based on the combined method in [18]. Let $\mathcal{C}^{(i)}$ denote the component code (the U code or the V code) of length n that is transmitted through the ith subchannel, where $i = 1, 2, ..., \gamma$. It has a dimension of k_i and a rate of $r_i = k_i/n$. As a result, the U-UV code has a dimension of $K = \sum_{i=1}^{\gamma} k_i$ and a rate of $R = K/N$. Let $\boldsymbol{c}^{(i)} = (c_1^{(i)}, c_2^{(i)}, ..., c_n^{(i)}) \in \mathbb{F}_2^n$ denote the codeword of $\mathcal{C}^{(i)}$. The U-UV codeword that is obtained by the above construction is denoted as $\boldsymbol{v} = (v_1, v_2, ..., v_N) \in \mathbb{F}_2^N$.

2.2 SCL Decoding of U-UV Codes

With the OSD for the component codes, which is a list decoding approach in nature, the successive cancellation list (SCL) decoding of the U-UV codes can be realized. This SCL decoding is parameterized by its decoding output list size l and denoted as SCL(l).

Figure 2 shows the SCL decoding of an H-level U-UV code. Assume that a U-UV codeword $\boldsymbol{v}$ of length N is transmitted over the additive white Gaussian noise (AWGN) channel using binary phase shift keying (BPSK). Let $\mathbf{r} = (\mathrm{r}_1, \mathrm{r}_2, \ldots, \mathrm{r}_N) \in \mathbb{R}^N$ denote the received symbol vector and $\boldsymbol{\mathcal{L}} = (\mathcal{L}_1, \mathcal{L}_2, ..., \mathcal{L}_N) \in \mathbb{R}^N$ denote the corresponding log-likelihood ratio (LLR) vector with entries defined as

$$\mathcal{L}_b = \ln \frac{P(\mathrm{r}_b \,|\, v_b = 0)}{P(\mathrm{r}_b \,|\, v_b = 1)}, \quad (2)$$

where $P(\mathrm{r}_b \,|\, v_b = 0)$ and $P(\mathrm{r}_b \,|\, v_b = 1)$ are the channel transition probabilities of v_b and $b = 1, 2, ..., N$. In the SCL decoding, the LLR values that correspond to the component codeword symbols at level-0 are computed via the SC decoding mechanism [4] based on the received LLRs at level-H that are defined as in (2). Once the LLRs of a component codeword at level-0 are determined, the component code will be decoded by the OSD. The OSD provides plural codeword estimations for each component code, substantiating the SCL decoding. Figure 2 indicates the component codes are decoded successively. Let $\boldsymbol{\mathcal{L}}^{(i)} = (\mathcal{L}_1^{(i)}, \mathcal{L}_2^{(i)}, ..., \mathcal{L}_n^{(i)}) \in \mathbb{R}^n$ and $\hat{\boldsymbol{c}}^{(i)} = (\hat{c}_1^{(i)}, \hat{c}_2^{(i)}, ..., \hat{c}_n^{(i)}) \in \mathbb{F}_2^n$ denote the LLR vector and the codeword estimation of component code $\mathcal{C}^{(i)}$, respectively. To assess the reliability of the codeword estimations, the correlation distance between $\boldsymbol{\mathcal{L}}^{(i)}$ and $\hat{\boldsymbol{c}}^{(i)}$ is needed, which is defined as

$$\Lambda(\boldsymbol{\mathcal{L}}^{(i)}, \hat{\boldsymbol{c}}^{(i)}) = \sum_{j:(1-2\hat{c}_j^{(i)})\cdot\mathcal{L}_j^{(i)}<0} |\mathcal{L}_j^{(i)}|, \tag{3}$$

where $j \in \{1, 2, ..., n\}$. A smaller correlation distance indicates the decoding estimation is more reliable.

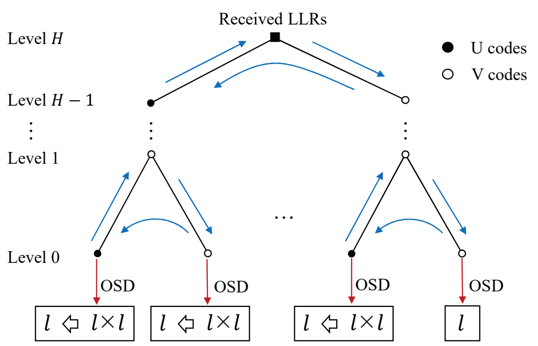

Figure 2. *The SCL decoding of an H-level U-UV code.*

In the SCL decoding of a U-UV code, the l most reliable estimations will be preserved at each decoding layer. This decoding path expansion can be illustrated by the SCL decoding tree as shown in Figure 3. In the decoding tree, each layer corresponds to a component code and nodes of the layer represent its estimations. Path expansion will be performed based on each estimation of a component code. An existing path can emancipate into multiple different paths, leading to an exponentially growing SCL decoding complexity. In order to rationalize the decoding complexity, at each layer, only the l most reliable expanded paths will be preserved. For this, the decoding path metric is defined based on the accumulated correlation distance (ACD) as [18]

$$\Phi^{(i)} = \sum_{i'=i}^{\gamma} \Lambda(\boldsymbol{\mathcal{L}}^{(i')}, \hat{\boldsymbol{c}}^{(i')}). \tag{4}$$

It indicates the reliability of a decoding path that reaches layer-i. The ACDs of the preserved paths at layer-i are denoted as $\Phi_\rho^{(i)}$, where $\rho = 1, 2, ..., l$. After the last component code is decoded, the U-UV codeword estimation that corresponds to the smallest path metric, i.e.,

$$\Phi_{\min}^{(1)} = \min\{\Phi_\rho^{(1)}, \forall\rho\}, \tag{5}$$

will be chosen as the decoding output $\hat{\boldsymbol{v}}$. Note that when $l = 1$, the SCL decoding degenerates into the SC decoding.

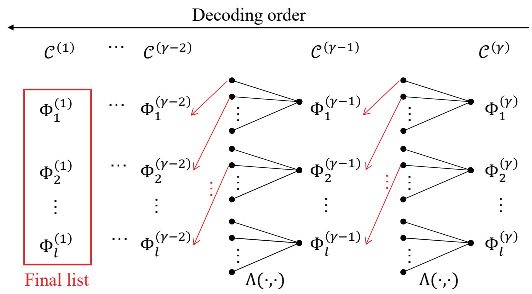

Figure 3. *Path expansion of the SCL decoding.*

2.3 OSD and Its BMA Variant

Let $\mathcal{C}(n, k, d)$ denote a binary linear block code of length n and dimension k with the minimum Hamming distance d. Its generator matrix $\mathbf{G}$ is a $k \times n$ binary matrix. Let $\boldsymbol{u} = (u_1, u_2, \ldots, u_k) \in \mathbb{F}_2^k$ denote the message. The codeword can be generated by $\boldsymbol{c} = \boldsymbol{u} \cdot \mathbf{G}$, and $\boldsymbol{c} = (c_1, c_2, \ldots, c_n) \in \mathbb{F}_2^n$. Let $\boldsymbol{z} = (z_1, z_2, \ldots, z_n)$ denote the BPSK modulated symbol sequence, where $z_j \in \{-1, 1\}$ and $j = 1, 2, ..., n$. The received symbol vector is denoted as $\boldsymbol{r} = (r_1, r_2, ..., r_n) \in \mathbb{R}^n$, where $r_j = z_j + w_j$ and w_j is the AWGN with a variance of $N_0/2$. The LLR vector of the received symbols is denoted as $\boldsymbol{L} = (L_1, L_2, ..., L_n) \in \mathbb{R}^n$ with entries $L_j = \ln \frac{P(r_j|c_j=0)}{P(r_j|c_j=1)}$. Under the BPSK paradigm, it can be further simplified as $L_j = \frac{4r_j}{N_0}$. Accordingly, the hard-decision received word $\boldsymbol{y} = (y_1, y_2, \ldots, y_n) \in \mathbb{F}_2^n$ can be obtained as $y_j = 0$, if $L_j > 0$, and $y_j = 1$ otherwise. The reliability of y_j can be evaluated by its LLR as $|L_j|$. A greater $|L_j|$ indicates the received symbol is more reliable.

2.3.1 The OSD

In the OSD, the received symbols are sorted in descending order of $|L_j|$, yielding the sorted LLR vector and received symbol vector, i.e., $\tilde{\boldsymbol{L}} = \Pi(\boldsymbol{L})$ and $\tilde{\boldsymbol{r}} = \Pi(\boldsymbol{r})$, where Π is the permutation function. Performing the same permutation on $\boldsymbol{y}$ and columns of $\mathbf{G}$, we obtain $\tilde{\boldsymbol{y}} = \Pi(\boldsymbol{y})$ and $\tilde{\mathbf{G}} = \Pi(\mathbf{G})$. Next, the systematic generator matrix $\tilde{\mathbf{G}}_{\mathrm{s}} = [\mathbf{I}_k \ \ \tilde{\mathbf{P}}]$ is obtained by performing GE on $\tilde{\mathbf{G}}$, where $\mathbf{I}_k$ is an identity submatrix of dimension k and $\tilde{\mathbf{P}}$ is the parity submatrix. If the first k columns of $\tilde{\mathbf{G}}$ are not linearly independent, an additional permutation may be performed. For simplicity, we assume that the first k columns of $\tilde{\mathbf{G}}$ are ensured with the linear independence. Therefore, the first k positions in $\tilde{\boldsymbol{y}}$ are called the most reliable independent positions (MRIPs), which also form the most reliable bases (MRB). The rest $n-k$ positions are called the least reliable positions (LRPs), as shown in Figure 4.

The OSD with order τ is referred to as OSD (τ). It means that there are $\tau + 1$ re-encoding phases during the decoding. Let $\tilde{\boldsymbol{y}}_{\mathrm{B}} = (\tilde{y}_1, \tilde{y}_2, ..., \tilde{y}_k) \in \mathbb{F}_2^k$ denote the first k positions of $\tilde{\boldsymbol{y}}$ that are associated with the MRB. In phase-0, the first codeword candidate $\tilde{\boldsymbol{c}}_0$ is obtained by $\tilde{\boldsymbol{c}}_0 = \tilde{\boldsymbol{y}}_{\mathrm{B}} \cdot \tilde{\mathbf{G}}_{\mathrm{s}}$. Let $\boldsymbol{e} = (e_1, e_2, ..., e_k) \in \mathbb{F}_2^k$ denote a test error pattern (TEP). In phase-t, all the TEPs with $w(\boldsymbol{e}) = t$ will be generated for re-encoding, i.e.,

$$\tilde{\boldsymbol{c}} = (\tilde{\boldsymbol{y}}_{\mathrm{B}} + \boldsymbol{e}) \cdot \tilde{\mathbf{G}}_{\mathrm{s}} = \boldsymbol{e} \cdot \tilde{\mathbf{G}}_{\mathrm{s}} + \tilde{\boldsymbol{c}}_0, \tag{6}$$

where $t = 1, 2, ..., \tau$ and $w(\cdot)$ denotes the Hamming weight of a vector. Let $\tilde{\boldsymbol{e}} = (\tilde{e}_1, \tilde{e}_2, ..., \tilde{e}_n) \in \mathbb{F}_2^k$ denote the hard-decision error pattern, i.e., $\tilde{\boldsymbol{e}} = \tilde{\boldsymbol{c}} + \tilde{\boldsymbol{y}}$. The correlation distance defined in (3) can also be written as

$$\Lambda(\tilde{\boldsymbol{e}}) = \sum_{j:\tilde{e}_j = 1} |\tilde{L}_j|. \tag{7}$$

Let $\tilde{\boldsymbol{c}}_{\mathrm{opt}}$ denote the codeword candidate with the smallest correlation distance. The decoding output can be obtained by $\hat{\boldsymbol{c}}_{\mathrm{opt}} = \Pi^{-1}(\tilde{\boldsymbol{c}}_{\mathrm{opt}})$, where Π^{-1} is the inverse of the permutation function Π. For an (n, k, d) binary linear block code with rate $k/n \geq 1/2$, a decoding order of

$$\tau = \min\left\{\left\lfloor \frac{d}{4} \right\rfloor, k\right\}, \tag{8}$$

would be sufficient to yield a near ML decoding performance for the code [10].

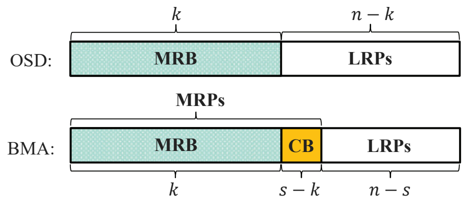

Figure 4. *Ordered received symbols in the OSD and BMA.*

2.3.2 The BMA

The BMA is a variant of the OSD. It further utilizes the information outside of the MRB with the matching techniques and reduces the OSD complexity but at the cost of storage [30]. In addition to considering all the codeword candidates associated with error patterns of weight at most τ in the MRB as the OSD, the BMA further considers the codeword candidates associated with error patterns of weight at most 2τ in the s most reliable positions (MRPs), where $k \leq s < n$. The $s - k$ positions outside the MRB form the control band (CB). Subsequently, length of the LRPs is $n - s$ in the BMA, which is also indicated as in Figure 4. Therefore, the BMA is explicitly defined as BMA (τ, s).

The BMA also consists of three steps, i.e., LLR sorting, GE operation and re-encoding. Note that the first two steps are the same as in the OSD. There are also $\tau + 1$ re-encoding phases in the BMA (τ, s). But the re-encoding operations in each phase (except phase-0) are divided into the standard process and the match process. Let $\tilde{\boldsymbol{e}}_{\mathrm{B}} = (\tilde{e}_1, \tilde{e}_2, ..., \tilde{e}_k)$, $\tilde{\boldsymbol{e}}_{\mathrm{S}} = (\tilde{e}_{k+1}, \tilde{e}_{k+2}, ..., \tilde{e}_s)$ and $\tilde{\boldsymbol{e}}_{\mathrm{L}} = (\tilde{e}_{s+1}, \tilde{e}_{s+2}, ..., \tilde{e}_n)$ denote the divided error patterns associated with the MRB, CB and LRPs, respectively. An error pattern $\tilde{\boldsymbol{e}}$ can be decomposed as $\tilde{\boldsymbol{e}} = (\tilde{\boldsymbol{e}}_{\mathrm{B}} \ \tilde{\boldsymbol{e}}_{\mathrm{S}} \ \tilde{\boldsymbol{e}}_{\mathrm{L}})$. In the standard process of phase-t, all the TEPs with $w(\boldsymbol{e}) = t$ will be generated for re-encoding as in (6), resulting in the codeword candidates associated with error patterns of $w(\tilde{\boldsymbol{e}}_{\mathrm{B}}) = t$. Subsequently, these TEPs will be stored with their corresponding $\tilde{\boldsymbol{e}}_{\mathrm{S}}$. Let $\boldsymbol{e}'$ denote a stored TEP. In the match process of phase-t, the re-encoding will be performed for the TEPs with weight greater than t. First, the TEPs with $w(\boldsymbol{e}) \le t$ will be generated to match the stored TEPs with $w(\boldsymbol{e}') = t$, constructing new TEPs that are denoted by $\boldsymbol{e}'' = \boldsymbol{e} + \boldsymbol{e}'$. Then, based on (6), the codeword candidates are generated by re-encoding $\boldsymbol{e}''$. It is also guaranteed that the error patterns generated in the match process should satisfy $w(\tilde{\boldsymbol{e}}_{\mathrm{B}}) + w(\tilde{\boldsymbol{e}}_{\mathrm{S}}) = 2t$ or $2t-1$. This match process can skip the unpromising TEPs of high weights and reduce the number of codeword candidates. Hence, compared to the OSD, the BMA can achieve a near ML decoding performance with a smaller decoding order and lower complexity. Table. 1 summarizes the Hamming weights of the error patterns that will be considered in re-encoding phase-t of the BMA. Note that the error patterns generated in the standard process satisfy $w(\tilde{\boldsymbol{e}}_{\mathrm{S}}) \ge t-1$, since the error patterns with $w(\tilde{\boldsymbol{e}}_{\mathrm{B}}) = t$ and $w(\tilde{\boldsymbol{e}}_{\mathrm{S}}) < t-1$ will be considered in the match process before phase-t.

Table 1. *The Hamming weights of the error patterns considered in phase-t of the BMA.*

	$w(\boldsymbol{e}_{\mathrm{B}})$	$w(\boldsymbol{e}_{\mathrm{S}})$	$w(\boldsymbol{e}_{\mathrm{B}}) + w(\boldsymbol{e}_{\mathrm{S}})$
Standard process	t	$\ge t-1$	$\ge 2t-1$
Match process	$1+t$ $2+t$ $\cdots$ $t+t$	$t-1, t-2$ $t-2, t-3$ $\cdots$ 0	$2t, 2t-1$

III. LOW COMPLEXITY COMPONENT CODE DECODING

The SCL decoding of U-UV codes is dominated by its component code decoding, i.e., the OSD. Its complexity increases exponentially with the decoding order. To reduce the component code decoding complexity, this section introduces the order skipping (OS) rule for both the OSD and its complexity reducing variant, the BMA. It estimates an approximated *a posteriori* correlation distance lower bound (CDLB). Consequently, some higher order phase re-encoding can be skipped if they cannot yield a more likely codeword. As for the BMA, it can further skip the redundant match process.

3.1 Ordered Statistics

Without loss of generality, we assume that an all-zero codeword is transmitted. Thus, over the received symbol vector $\boldsymbol{r}$, where $r_j = 1 + w_j$. Since w_j is the AWGN, the probability distribution function (pdf) of the received symbol r_j is given by

$$f_r(x) = \frac{1}{\sqrt{\pi N_0}} e^{-\frac{(x-1)^2}{N_0}}. \tag{9}$$

Recall that with BPSK modulation, the received LLRs can be simplified as $L_j = 4r_j/N_0$. Hence, the scaled magnitude of LLR can be utilized to assess the reliability of the received symbol r_j, i.e., $\alpha_j = |r_j|$. Accordingly, the ordered reliability sequence corresponding to $\tilde{\boldsymbol{r}}$ can be denoted by $\tilde{\boldsymbol{\alpha}} = (\tilde{\alpha}_1, \tilde{\alpha}_2, ..., \tilde{\alpha}_n)$. The pdf of α_j is

$$f_\alpha(x) = \begin{cases} 0, & x < 0; \\ \frac{1}{\sqrt{\pi N_0}} \left(e^{-\frac{(x+1)^2}{N_0}} + e^{-\frac{(x-1)^2}{N_0}}\right), & x \ge 0. \end{cases} \tag{10}$$

The cumulative distribution function (cdf) of α_j is further derived as

$$F_\alpha(x) = \begin{cases} 0, & x < 0; \\ 1 - Q\left(\frac{2x+2}{\sqrt{2N_0}}\right) - Q\left(\frac{2x-2}{\sqrt{2N_0}}\right), & x \ge 0, \end{cases} \tag{11}$$

where $Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)dx$ is the standard normal tail function. Furthermore, the pdf of $\tilde{\alpha}_j$ is given by [47]

$$f_{\tilde{\alpha}_j}(x) = \frac{n!}{(j-1)!(n-j)!} \cdot (1 - F_\alpha(x))^{j-1} F_\alpha(x)^{n-j} f_\alpha(x). \tag{12}$$

This pdf can be approximated as a Gaussian distribution and its expectation can be estimated by [25]

$$\mu_{\tilde{\alpha}_j} = F_\alpha^{-1}\left(1 - \frac{j}{n}\right). \tag{13}$$

Based on (9), the probability of the an erroneous decision on the j-th ordered bit, i.e., $\tilde{y}_j$, conditioning on $\tilde{\alpha}_j$, can be determined by

$$\begin{aligned} P(\tilde{e}_j = 1 \,|\, \tilde{\alpha}_j) &= \frac{f_r(-\tilde{\alpha}_j)}{f_r(\tilde{\alpha}_j) + f_r(-\tilde{\alpha}_j)} \\ &= \frac{1}{1 + \exp(4\tilde{\alpha}_j/N_0)}. \end{aligned} \tag{14}$$

3.2 Order Skipping OSD

In phase-t of the order τ OSD, i.e., OSD(τ), all TEPs with Hamming weight t will be generated for re-encoding as in (6). It generates $\binom{k}{t}$ codeword candidates. Therefore, the total number of codeword candidates generated by the OSD(τ) is $\sum_{t=0}^{\tau}\binom{k}{t}$. It can be seen that the re-encoding in the higher phases dominates the decoding complexity. The decoding complexity can be reduced if we can determine the higher phase OSD cannot provide a more likely candidate and skip running them. To realize this, the CDLB of the codeword candidates generated in high phases needs to estimated.

In the OSD, the error pattern $\tilde{\boldsymbol{e}}$ can be decomposed as $\tilde{\boldsymbol{e}} = (\tilde{\boldsymbol{e}}_{\mathrm{B}}\ \tilde{\boldsymbol{e}}_{\mathrm{P}})$, where $\tilde{\boldsymbol{e}}_{\mathrm{B}} = (\tilde{e}_1, \tilde{e}_2, ..., \tilde{e}_k)$ and $\tilde{\boldsymbol{e}}_{\mathrm{P}} = (\tilde{e}_{k+1}, \tilde{e}_{k+2}, ..., \tilde{e}_n)$. Consequently, its associated correlation distance $\Lambda(\tilde{\boldsymbol{e}})$ can be determined by

$$\Lambda(\tilde{\boldsymbol{e}}) = \Lambda(\tilde{\boldsymbol{e}}_{\mathrm{B}}) + \Lambda(\tilde{\boldsymbol{e}}_{\mathrm{P}}), \tag{15}$$

where $\Lambda(\tilde{\boldsymbol{e}}_{\mathrm{B}}) = \sum_{\substack{1\le j\le k\\ \tilde{e}_j=1}} |\tilde{L}_j|$ and $\Lambda(\tilde{\boldsymbol{e}}_{\mathrm{P}}) = \sum_{\substack{k+1\le j\le n\\ \tilde{e}_j=1}} |\tilde{L}_j|$ are associated with the MRB and the LRPs, respectively. In the re-encoding phase-$(t+1)$, the error patterns have the same Hamming weight on the MRB, i.e., $w(\tilde{\boldsymbol{e}}_{\mathrm{B}}) = t+1$. A lower bound of $\Lambda(\tilde{\boldsymbol{e}}_{\mathrm{B}})$ is given by

$$\Lambda(\tilde{\boldsymbol{e}}_{\mathrm{B}}) \ge \sum_{j=k-t}^{k} |\tilde{L}_j|. \tag{16}$$

Note that $\tilde{\boldsymbol{e}}_{\mathrm{P}}$ is unknown before the re-encoding. The conditional expectation can be utilized to estimate $\Lambda(\tilde{\boldsymbol{e}}_{\mathrm{P}})$. Based on (14), the expected probability of the j-th ordered received symbol being erroneous conditioning on its reliability can be determined by

$$\begin{aligned}\mathbb{E}\{\tilde{e}_j \,|\, \tilde{\alpha}_j\} &= P(\tilde{e}_j = 1|\tilde{\alpha}_j)\cdot 1 + P(\tilde{e}_j = 0|\tilde{\alpha}_j)\cdot 0\\ &= \frac{1}{1+\exp(|\tilde{L}_j|)}.\end{aligned} \tag{17}$$

Furthermore, the conditional expectation of $\Lambda(\tilde{\boldsymbol{e}}_{\mathrm{P}})$ can be estimated by

$$\begin{aligned}\mathbb{E}\{\Lambda(\tilde{\boldsymbol{e}}_{\mathrm{P}}) \,|\, \tilde{\boldsymbol{\alpha}}\} &= \sum_{j=k+1}^{n} \mathbb{E}\{\tilde{e}_j \,|\, \tilde{\alpha}_j\}\cdot|\tilde{L}_j|,\\ &= \sum_{j=k+1}^{n} \frac{|\tilde{L}_j|}{1+\exp(|\tilde{L}_j|)}.\end{aligned} \tag{18}$$

Let $\mathcal{D}_{\mathrm{OSD}}^{(t+1)}$ denote the CDLB associated with the codeword candidates generated in phase-$(t+1)$. Based on (16) and (18), it can be approximated as

$$\begin{aligned}\mathcal{D}_{\mathrm{OSD}}^{(t+1)} &\approx \sum_{j=k-t}^{k} |\tilde{L}_j| + \mathbb{E}\{\Lambda(\tilde{\boldsymbol{e}}_{\mathrm{P}}) \,|\, \tilde{\boldsymbol{\alpha}}\}\\ &= \sum_{j=k-t}^{k} |\tilde{L}_j| + \sum_{j=k+1}^{n} \frac{|\tilde{L}_j|}{1+\exp(|\tilde{L}_j|)}.\end{aligned} \tag{19}$$

Note that if $t' > t$, $\mathcal{D}_{\mathrm{OSD}}^{(t')} > \mathcal{D}_{\mathrm{OSD}}^{(t)}$. Let Λ^* denote the smallest correlation distance that is obtained in the decoding. If

$$\Lambda^* < \mathcal{D}_{\mathrm{OSD}}^{(t+1)}, \tag{20}$$

the phase-$t+1$ re-encoding is not able to yield a more likely codeword. Neither will the following higher phase re-encoding. The OSD can be terminated. This exemplifies the OS-OSD.

Compared with the order skipping rule of [26], in which its CDLB depends on an empirical parameter, the proposed CDLB of (19) is obtained based on likelihood of the received symbols over the LRPs. It is a more general and accurate approximation that can be applied over a wide ranges of signal-to-noise ratio (SNR) regime. Further considering the U-UV coding paradigm, the proposed CDLB estimation is also more effective for designing the OS-OSD for different component codes.

3.3 Order Skipping BMA

Two skipping rules are introduced for the OS-BMA. By determining the approximated *a posteriori* CDLB, the first rule skips the higher phase re-encoding, while the second rule skips the match process.

1) Skipping the Re-encoding: Similar to (15), the correlation distance in the BMA can be decomposed as

$$\Lambda(\tilde{\boldsymbol{e}}) = \Lambda(\tilde{\boldsymbol{e}}_{\mathrm{B}}) + \Lambda(\tilde{\boldsymbol{e}}_{\mathrm{S}}) + \Lambda(\tilde{\boldsymbol{e}}_{\mathrm{L}}), \tag{21}$$

where $\Lambda(\tilde{\boldsymbol{e}}_{\mathrm{S}}) = \sum_{\substack{k+1\le j\le s\\ \tilde{e}_j=1}} |\tilde{L}_j|$ and $\Lambda(\tilde{\boldsymbol{e}}_{\mathrm{L}}) = \sum_{\substack{s+1\le j\le n\\ \tilde{e}_j=1}} |\tilde{L}_j|$ are associated with the CB and LRPs, respectively. Let $\mathcal{D}_{\text{BMA}}^{(t+1)}$ denote the CDLB associated with the codeword candidates generated in phase-$(t+1)$ of the BMA. Recall (19), in order to estimate $\mathcal{D}_{\text{BMA}}^{(t+1)}$, the correlation distance lower bound over the s MRPs need to be known. Let $\tilde{\boldsymbol{e}}^{(a_1,a_2)} = (\tilde{\boldsymbol{e}}_{\text{B}}\ \tilde{\boldsymbol{e}}_{\text{S}})$ denote the error pattern associated with the s MRPs, where $w(\tilde{\boldsymbol{e}}_{\text{B}}) = a_1$ and $w(\tilde{\boldsymbol{e}}_{\text{S}}) = a_2$. The CDLB of $\tilde{\boldsymbol{e}}^{(a_1,a_2)}$ is given by

$$\Lambda(\tilde{\boldsymbol{e}}^{(a_1,a_2)}) \geq \sum_{j=k-a_1+1}^{k} |\tilde{L}_j| + \sum_{j=s-a_2+1}^{s} |\tilde{L}_j|. \quad (22)$$

Hence, this lower bound is evaluated by considering the reliability of the least reliable positions of the MRB and CB, respectively. To further estimate $\mathcal{D}_{\text{BMA}}^{(t+1)}$, the following proposition is introduced.

Proposition 1. *In the re-encoding phase-t of the BMA, where $0 < t \leq \tau$, the CDLB over the s MRPs is estimated by*

$$\Lambda(\tilde{\boldsymbol{e}}^{(t,t-1)}) \geq \sum_{j=k-t+1}^{k} |\tilde{L}_j| + \sum_{j=s-t+2}^{s} |\tilde{L}_j|. \quad (23)$$

Proof. The CDLB over the s MRPs is determined by the error patterns with the smallest Hamming weight of $\tilde{\boldsymbol{e}}_{\text{B}}$ and $\tilde{\boldsymbol{e}}_{\text{S}}$. Recall Table. 1, $\tilde{\boldsymbol{e}}^{(t,t-1)}$ is the error pattern with the smallest Hamming weight on the MRPs in the standard process, where $w(\tilde{\boldsymbol{e}}^{(t,t-1)}) = 2t-1$. In the match process, there are $t-1$ types of error patterns with a Hamming weight of $2t-1$ over the s MRPs, i.e., $\tilde{\boldsymbol{e}}^{(t+1,t-2)}, \tilde{\boldsymbol{e}}^{(t+2,t-3)}, ..., \tilde{\boldsymbol{e}}^{(2t-1,0)}$. Note that a received symbol of the MRB would be more reliable than anyone of the CB. Hence, among these error patterns, $\tilde{\boldsymbol{e}}^{(t,t-1)}$ is associated with the smallest CDLB over the s MRPs. Based on (22), the lower bound of (23) can be obtained.

Similar to (19), by further utilizing the conditional expectation of $\Lambda(\tilde{\boldsymbol{e}}_{\text{L}})$, $\mathcal{D}_{\text{BMA}}^{(t+1)}$ can be approximated as

$$\begin{aligned}\mathcal{D}_{\text{BMA}}^{(t+1)} &\approx \sum_{j=k-t}^{k} |\tilde{L}_j| + \sum_{j=s-t+1}^{s} |\tilde{L}_j| + \mathbb{E}\{\Lambda(\tilde{\boldsymbol{e}}_{\text{L}}) \,|\, \tilde{\boldsymbol{\alpha}}\} \\ &= \sum_{j=k-t}^{k} |\tilde{L}_j| + \sum_{j=s-t+1}^{s} |\tilde{L}_j| + \sum_{j=s+1}^{n} \frac{|\tilde{L}_j|}{1+\exp(|\tilde{L}_j|)}.\end{aligned}$$

Again let Λ^* denote the smallest correlation distance obtained so far in the decoding. If

$$\Lambda^* < \mathcal{D}_{\text{BMA}}^{(t+1)}, \quad (24)$$

the subsequent phase-$(t+1)$ BMA decoding can be skipped, and the decoding can be terminated.

2) Skipping the Match Process: Since the match process in the BMA can be regarded as the re-encoding in higher phases, the order skipping rule can also be applied to skip the redundant match process, resulting in a lower decoding complexity.

Before each match operation, Hamming weights of the error patterns over the MRB and CB can be known. Let $\tilde{\boldsymbol{e}}^{(t_1,t_2)}$ denote the error pattern, where (t_1,t_2) is the match pattern, $t_1 = w(\tilde{\boldsymbol{e}}_{\text{B}})$ and $t_2 = w(\tilde{\boldsymbol{e}}_{\text{S}})$. The match operation for $\tilde{\boldsymbol{e}}^{(t_1,t_2)}$ can be skipped, if

$$\Lambda^* < \mathcal{D}_{\text{BMA}}^{(t_1,t_2)}, \quad (25)$$

where

$$\begin{aligned}\mathcal{D}_{\text{BMA}}^{(t_1,t_2)} \approx &\sum_{j=k-t_1+1}^{k} |\tilde{L}_j| + \sum_{j=s-t_2+1}^{s} |\tilde{L}_j| \\ &+ \sum_{j=s+1}^{n} \frac{|\tilde{L}_j|}{1+\exp(|\tilde{L}_j|)},\end{aligned} \quad (26)$$

is the approximated CDLB associated with $\tilde{\boldsymbol{e}}^{(t_1,t_2)}$.

The above OS-BMA is summarized as in Algorithm 1, and the re-encoding in phase-t for $t > 0$ is summarized as in Algorithm 2. Note that in Algorithm 2, backtracking is utilized to generate the TEPs with Hamming weight not greater than t as in line 20. Details of the *Match* and *Store* operations in the BMA can be found in [30].

Algorithm 1. *OS-BMA.*

Input: $\mathbf{G}, \boldsymbol{L}, \boldsymbol{y}, s, \tau$
Output: $\hat{\boldsymbol{c}}_{\text{opt}}$
1: Sort received LLRs
2: Perform GE to obtain $\tilde{\mathbf{G}}_{\text{s}}$
3: Compute $\tilde{\boldsymbol{c}}^{(0)}$ and Λ^*
4: $\tilde{\boldsymbol{c}}_{\text{opt}} \leftarrow \tilde{\boldsymbol{c}}^{(0)}$
5: **for** $t = 1, 2, ..., \tau$ **do**
6:   **if** $\Lambda^* \leq \mathcal{D}_{\text{BMA}}^{(t)}$ **then**
7:     **break**
8:   **else**

9: $\quad\quad \boldsymbol{e} \leftarrow \mathbf{0}$
10: $\quad\quad q \leftarrow 0$
11: $\quad\quad$ Perform *Algorithm 2*
12: $\quad$ **end if**
13: **end for**
14: $\hat{\boldsymbol{c}}_{\text{opt}} \leftarrow \Pi^{-1}(\tilde{\boldsymbol{c}}_{\text{opt}})$

Algorithm 2. *OS-BMA re-encoding.*

Input: t, $\boldsymbol{e}$, q
1: $q' \leftarrow q$
2: $g \leftarrow w(\boldsymbol{e}) + 1$
3: **if** $g \leq t$ **then**
4: $\quad$ **for** $j = k, k-1, ..., q'+1$ **do**
5: $\quad\quad e_j \leftarrow 1$
6: $\quad\quad q \leftarrow j$
7: $\quad\quad$ **if** $w(\boldsymbol{e}) = t$ **then**
8: $\quad\quad\quad$ Re-encode $\boldsymbol{e}$ as in (6)
9: $\quad\quad\quad$ Update $\tilde{\boldsymbol{c}}_{\text{opt}}$ and Λ^*
10: $\quad\quad\quad$ *Store* $\boldsymbol{e}$ as $\boldsymbol{e}'$
11: $\quad\quad$ **end if**
12: $\quad\quad$ **if** $t > g$ and $\Lambda^* > \mathcal{D}_{\text{BMA}}^{(t+g,t-g-1)}$ **then**
13: $\quad\quad\quad$ *Match* $\tilde{\boldsymbol{e}}^{(t+g,t-g-1)}$, generate $\boldsymbol{e}'' = \boldsymbol{e} + \boldsymbol{e}'$
14: $\quad\quad\quad$ Re-encode $\boldsymbol{e}''$ as in **Step 8-9**
15: $\quad\quad$ **end if**
16: $\quad\quad$ **if** $\Lambda^* > \mathcal{D}_{\text{BMA}}^{(t+g,t-g)}$ **then**
17: $\quad\quad\quad$ *Match* $\tilde{\boldsymbol{e}}^{(t+g,t-g)}$, generate $\boldsymbol{e}'' = \boldsymbol{e} + \boldsymbol{e}'$
18: $\quad\quad\quad$ Re-encoding $\boldsymbol{e}''$ as in **Step 8-9**
19: $\quad\quad$ **end if**
20: $\quad\quad$ Call *Algorithm 2* again
21: $\quad\quad e_j \leftarrow 0$
22: $\quad$ **end for**
23: **end if**

IV. INTEGRATION WITH PATH PRUNING SCL DECODING

This section further proposes the path pruning (PP)-SCL decoding for the U-UV codes, It eliminates the unpromising SCL decoding paths through estimating the *a priori* CDLB. Unlike the *a posteriori* CDLB that is estimated based on the received information, this *a priori* CDLB is estimated by characterizing the correlation distance distribution of the OSD or BMA outputs. Therefore, this CDLB is obtained before the component code decoding. Finally, the proposed low complexity SCL decoding is substantiated by integrating the PP-SCL decoding and the above mentioned OS-OSD (or OS-BMA).

4.1 Path Pruning SCL Decoding

Recall the SCL decoding path expansion that is shown in Figure 3. In the PP-SCL decoding, the l smallest ACDs kept at layer-i are ordered as

$$\Phi_1^{(i)} \leq \Phi_2^{(i)} \leq \cdots \leq \Phi_l^{(i)}. \tag{27}$$

Subsequently, in decoding next component code $\mathcal{C}^{(i-1)}$, we prioritize to elaborate from the node with a smaller ACD. It means that $\mathcal{C}^{(i-1)}$ will first be decoded based on the estimation of $\mathcal{C}^{(i)}$ that corresponds to $\Phi_1^{(i)}$. Let $\Phi_{\rho,1}^{(i-1)}$, $\Phi_{\rho,2}^{(i-1)}$, $\ldots$, $\Phi_{\rho,l}^{(i-1)}$ denote the l smallest ACDs at layer-$(i-1)$. They are elaborated from the node with $\Phi_\rho^{(i)}$. By first elaborating from the node with $\Phi_1^{(i)}$, it yields $\Phi_{1,1}^{(i-1)}$, $\Phi_{1,2}^{(i-1)}$, $\ldots$, $\Phi_{1,l}^{(i-1)}$ at layer-$(i-1)$. Since so far they are the only l decoding paths at layer-$(i-1)$, they are sorted as in (27) and updated as $\Phi_1^{(i-1)}$, $\Phi_2^{(i-1)}$, $\ldots$, $\Phi_l^{(i-1)}$, respectively. The decoder then continues to decode $\mathcal{C}^{(i-1)}$ based on the node with $\Phi_2^{(i)}$. However, if the path starting from the node is unlikely to yield an ACD that is smaller than $\Phi_l^{(i-1)}$, it will be pruned. Its inherited component code decoding can be skipped.

Let $\Lambda_c^{(i)}$ denote the *a priori* CDLB of decoding $\mathcal{C}^{(i)}$. Once the l smallest ACDs $\Phi_1^{(i-1)}$, $\Phi_2^{(i-1)}$, $\ldots$, $\Phi_l^{(i-1)}$ have been updated based on the decoding that starts from layer-i, the following decoding of $\mathcal{C}^{(i-1)}$ will be assessed before their execution. For $\rho > 1$ and if

$$\Phi_\rho^{(i)} + \Lambda_c^{(i-1)} > \Phi_l^{(i-1)}, \tag{28}$$

it indicates that nodes of $\Phi_\rho^{(i)}$, $\Phi_{\rho+1}^{(i)}$, $\ldots$, $\Phi_l^{(i)}$ cannot lead to a more likely decoding path than the existing ones. Decoding path elaboration based on these nodes can be skipped. Otherwise, the decoder continues to decode $\mathcal{C}^{(i-1)}$ based on the node with $\Phi_\rho^{(i)}$, yielding the ACDs $\Phi_{\rho,1}^{(i-1)}$, $\Phi_{\rho,2}^{(i-1)}$, $\ldots$, $\Phi_{\rho,l}^{(i-1)}$ at layer-$(i-1)$. Afterward, the l smallest ACDs kept at layer-$(i-1)$ will be updated and sorted again as in (27). This process continues until the condition of (28) occurs or all the decoding path elaborations from layer-i have been completed. The decoder then moves onto decode the next component code $\mathcal{C}^{(i-2)}$. The above PP-SCL decoding is illustrated by Figure 5 as an example with $l = 3$.

Note that the *a priori* CDLB $\Lambda_c^{(i)}$ can be obtained

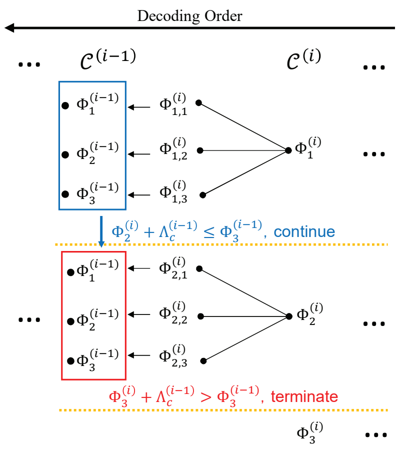

Figure 5. *Path expansion of PP-SCL decoding with $l = 3$.*

aforehand, which will be introduced in the following subsection through estimating the pdf of the correlation distance. Compared with the path pruning SCL decoding of [37], in which the unpromising decoding paths are eliminated under a tolerable decoding error performance loss, the proposed PP-SCL decoding does not require additional calculations on likelihood probabilities and results in a more effective path pruning for the list decoding.

4.2 PP-SCL with OS-OSD

To integrate the PP-SCL with the OS-OSD, the *a priori* CDLB $\Lambda_c^{(i)}$ is first estimated. Moreover, the OS-OSD is further compiled into the SCL decoding.

1) Estimation of $\Lambda_c^{(i)}$ in the OSD: Recall that $\alpha_j = |r_j|$ is utilized as the reliability of a received symbol. By this, the correlation distance can also be defined based on $\tilde{\boldsymbol{\alpha}}$ as

$$\lambda(\tilde{\boldsymbol{e}}) = \sum_{j:e_j=1} \tilde{\alpha}_j. \tag{29}$$

Based on $\tilde{\boldsymbol{\alpha}}$, let $\lambda_c^{(i)}$ denote the *a priori* CDLB of all codeword estimations in decoding $\mathcal{C}^{(i)}$. Hence, the estimation of $\Lambda_c^{(i)}$ can be converted into estimating $\lambda_c^{(i)}$ as

$$\Lambda_c^{(i)} = \frac{4\lambda_c^{(i)}}{N_0}, \tag{30}$$

where N_0 is the equivalent noise power of the polarized subchannels. It can be obtained through Gaussian approximation (GA) [21]. It has been shown that the pdf of λ in the OSD can be approximated as Gaussian distributed random variable [25, 48]. That says the pdf of λ can be approximated by

$$f_\lambda(x) = \frac{1}{\sqrt{2\pi}\sigma_\lambda} \exp\left(-\frac{(x-\mu_\lambda)^2}{2\sigma_\lambda^2}\right), \tag{31}$$

where μ_λ and σ_λ^2 are the expectation and variance of λ, respectively. Based on this, $\lambda_c^{(i)}$ can be estimated through

$$1 - \int_{\lambda_c^{(i)}}^{\infty} f_\lambda(x)dx = \theta, \tag{32}$$

where $\theta \in (0, 1)$. In practice, θ can be empirically chosen to achieve the best performance-complexity tradeoff.

Let $\mathbb{V}[\cdot]$ denote the variance of a random variable. To further determine μ_λ and σ_λ^2, the following lemmas are introduced.

Lemma 1. *([47]) Let $\{X_j\}$ denote a sequence of independent and identically distributed (i.i.d.) random variables, where X_j is a nonnegative integer. Let M denote a random variable with nonnegative integer value. The compound random variable W_M, called random sum, is defined as $W_M = \sum_{j=1}^{M} X_j$. The expectation and variance of W_M are given by*

$$\begin{aligned} &\mathbb{E}[W_M] = \mathbb{E}[X]\mathbb{E}[M]; \\ &\mathbb{V}[W_M] = \mathbb{V}[X]\mathbb{E}[M] + \mathbb{E}^2[X]\mathbb{V}[M]. \end{aligned} \tag{33}$$

Lemma 2. *([47]) Given two independent Gaussian distributed random variables denoted as $A \backsim \mathcal{N}(\mu_A, \sigma_A^2)$ and $B \backsim \mathcal{N}(\mu_B, \sigma_B^2)$. Let Z denote a random variable defined as $Z = A + B$. Z is Gaussian distributed and $Z \backsim \mathcal{N}(\mu_A + \mu_B, \sigma_A^2 + \sigma_B^2)$.*

Recall that in the OSD, $\lambda(\tilde{\boldsymbol{e}}) = \lambda(\tilde{\boldsymbol{e}}_{\mathrm{B}}) + \lambda(\tilde{\boldsymbol{e}}_{\mathrm{P}})$. Both $\lambda(\tilde{\boldsymbol{e}}_{\mathrm{B}})$ and $\lambda(\tilde{\boldsymbol{e}}_{\mathrm{P}})$ can be assumed as Gaussian distributed random variables. Based on the monotonicity of $\tilde{\boldsymbol{\alpha}}$ as in (12), the distribution of reliability α can be divided into two parts based on $\mu_{\tilde{\alpha}_k}$, which are associated with the MRB and LRPs, respectively. Let α_{B} and α_{P} denote the reliability on the MRB and LRPs, respectively. Their pdfs can be estimated by [48]

$$f_{\alpha_{\mathrm{B}}}(x) = \frac{f_\alpha(x)}{\int_{\mu_{\tilde{\alpha}_k}}^{\infty} f_\alpha(y)dy}, \ x \geq \mu_{\tilde{\alpha}_k}, \tag{34}$$

$$f_{\alpha_{\mathrm{P}}}(x)=\frac{f_\alpha(x)}{\int_0^{\mu_{\tilde{\alpha}_k}} f_\alpha(y)dy},\ 0\le x<\mu_{\tilde{\alpha}_k}. \quad (35)$$

Further let Q_{B} and Q_{P} denote the Hamming weight of the error pattern $\tilde{\boldsymbol{e}}$ over the MRB and LRPs, respectively, i.e., $Q_{\mathrm{B}}=w(\tilde{\boldsymbol{e}}_{\mathrm{B}})$ and $Q_{\mathrm{P}}=w(\tilde{\boldsymbol{e}}_{\mathrm{P}})$. In OSD ($\tau$), the distribution of λ is given in the following theorem.

Theorem 1. *([48]) The pdf of the correaltion distance λ generated by OSD (τ) is approximated as Gaussian distributed random variable. Its expectation and variance are given by:*

$$\mu_\lambda=\mathbb{E}[\alpha_{\mathrm{B}}]\mathbb{E}[Q_{\mathrm{B}}]+\mathbb{E}[\alpha_{\mathrm{P}}]\mathbb{E}[Q_{\mathrm{P}}], \quad (36)$$

$$\begin{aligned}\sigma_\lambda^2=&\mathbb{V}[\alpha_{\mathrm{B}}]\mathbb{E}[Q_{\mathrm{B}}]+\mathbb{E}^2[\alpha_{\mathrm{B}}]\mathbb{V}[Q_{\mathrm{B}}]+\\&\mathbb{V}[\alpha_{\mathrm{P}}]\mathbb{E}[Q_{\mathrm{P}}]+\mathbb{E}^2[\alpha_{\mathrm{P}}]\mathbb{V}[Q_{\mathrm{P}}].\end{aligned} \quad (37)$$

Proof. By assuming that $\lambda(\tilde{\boldsymbol{e}}_{\mathrm{B}})$ and $\lambda(\tilde{\boldsymbol{e}}_{\mathrm{P}})$ are independent and Gaussian distributed, they can be denoted as $\lambda(\tilde{\boldsymbol{e}}_{\mathrm{B}})\backsim\mathcal{N}(\mu_{\mathrm{B}},\sigma_{\mathrm{B}}^2)$ and $\lambda(\tilde{\boldsymbol{e}}_{\mathrm{P}})\backsim\mathcal{N}(\mu_{\mathrm{P}},\sigma_{\mathrm{P}}^2)$. Based on Lemma 1, we have $\mu_{\mathrm{B}}=\mathbb{E}[\alpha_{\mathrm{B}}]\mathbb{E}[Q_{\mathrm{B}}]$, $\sigma_{\mathrm{B}}^2=\mathbb{V}[\alpha_{\mathrm{B}}]\mathbb{E}[Q_{\mathrm{B}}]+\mathbb{E}^2[\alpha_{\mathrm{B}}]\mathbb{V}[Q_{\mathrm{B}}]$, $\mu_{\mathrm{P}}=\mathbb{E}[\alpha_{\mathrm{P}}]\mathbb{E}[Q_{\mathrm{P}}]$ and $\sigma_{\mathrm{P}}^2=\mathbb{V}[\alpha_{\mathrm{P}}]\mathbb{E}[Q_{\mathrm{P}}]+\mathbb{E}^2[\alpha_{\mathrm{P}}]\mathbb{V}[Q_{\mathrm{P}}]$. Based on Lemma 2 and $\lambda(\tilde{\boldsymbol{e}})=\lambda(\tilde{\boldsymbol{e}}_{\mathrm{B}})+\lambda(\tilde{\boldsymbol{e}}_{\mathrm{P}})$, the above characterization of μ_λ and σ_λ^2 can be obtained.

2) OS-OSD in PP-SCL: The OS-OSD is further utilized to decode the component codes. It should be ensured that the decoding of each component code can provide sufficient codeword estimations, i.e., the number of the decoding paths at each decoding layer should not be smaller than l. Hence, integrated with the PP-SCL, the order skipping rule as defined in (20) is only active for the phase $t\ge 1$. Since the smallest dimension of the BCH codes used in this work is $k=7$, re-encoding after phase-1 of the OSD can ensure that there are at least $\binom{k}{0}+\binom{k}{1}=8$ estimations. For a moderate decoding output list size $l\le 8$, such an operation ensures that there are sufficient decoding paths over the SCL decoding tree.

4.3 PP-SCL with OS-BMA

For the OS-BMA employed in the PP-SCL decoding, the order skipping rule defined in (24) is only active for phase $t\ge 1$. This can ensure that the decoding of each component code can provide sufficient codeword estimations to substantiate the list decoding.

Similarly, the *a priori* CDLB $\Lambda_c^{(i)}$ in the BMA can also be estimated through determining $\lambda_c^{(i)}$ as in (30) and (32). The distribution of λ in the BMA can be estimated as follows.

Recall that $\lambda(\tilde{\boldsymbol{e}})=\lambda(\tilde{\boldsymbol{e}}_{\mathrm{B}})+\lambda(\tilde{\boldsymbol{e}}_{\mathrm{S}})+\lambda(\tilde{\boldsymbol{e}}_{\mathrm{L}})$ in the BMA. We can follow the assumption that $\lambda(\tilde{\boldsymbol{e}}_{\mathrm{B}})$, $\lambda(\tilde{\boldsymbol{e}}_{\mathrm{S}})$ and $\lambda(\tilde{\boldsymbol{e}}_{\mathrm{L}})$ are Gaussian distributed. Subsequently, the distribution of reliability α can be divided into three parts associated with the MRB, CB and LRPs, respectively. With α_{B}, let α_{S} and α_{L} denote the reliability on the CB and LRPs, respectively. The pdf of α_{B} is given by (34). For α_{S} and α_{L} , their pdfs are given by

$$f_{\alpha_{\mathrm{S}}}(x)=\frac{f_\alpha(x)}{\int_{\mu_{\tilde{\alpha}_s}}^{\mu_{\tilde{\alpha}_k}} f_\alpha(y)dy},\ \mu_{\tilde{\alpha}_s}\le x\le\mu_{\tilde{\alpha}_k}, \quad (38)$$

and

$$f_{\alpha_{\mathrm{L}}}(x)=\frac{f_\alpha(x)}{\int_0^{\mu_{\tilde{\alpha}_s}} f_\alpha(y)dy},\ 0\le x\le\mu_{\tilde{\alpha}_s}, \quad (39)$$

respectively. With $Q_{\mathrm{B}}=w(\tilde{\boldsymbol{e}}_{\mathrm{B}})$, further let $Q_{\mathrm{S}}=w(\tilde{\boldsymbol{e}}_{\mathrm{S}})$ and $Q_{\mathrm{L}}=w(\tilde{\boldsymbol{e}}_{\mathrm{L}})$. The distribution of λ in the BMA is given in the following Theorem.

Theorem 2. *The pdf about the correlation distance λ generated by BMA (τ,s) is approximated as a Gaussian distribution. Its expectation and variance are given by:*

$$\mu_\lambda=\mathbb{E}[\alpha_{\mathrm{B}}]\mathbb{E}[Q_{\mathrm{B}}]+\mathbb{E}[\alpha_{\mathrm{S}}]\mathbb{E}[Q_{\mathrm{S}}]+\mathbb{E}[\alpha_{\mathrm{L}}]\mathbb{E}[Q_{\mathrm{L}}], \quad (40)$$

$$\begin{aligned}\sigma_\lambda^2=&\mathbb{V}[\alpha_{\mathrm{B}}]\mathbb{E}[Q_{\mathrm{B}}]+\mathbb{E}^2[\alpha_{\mathrm{B}}]\mathbb{V}[Q_{\mathrm{B}}]+\\&\mathbb{V}[\alpha_{\mathrm{S}}]\mathbb{E}[Q_{\mathrm{S}}]+\mathbb{E}^2[\alpha_{\mathrm{S}}]\mathbb{V}[Q_{\mathrm{S}}]+\\&\mathbb{V}[\alpha_{\mathrm{L}}]\mathbb{E}[Q_{\mathrm{L}}]+\mathbb{E}^2[\alpha_{\mathrm{L}}]\mathbb{V}[Q_{\mathrm{L}}].\end{aligned} \quad (41)$$

Proof. The proof is provided in Appendix A.

Figure 6 shows the distribution of the correlation distances in decoding the BCH codes of length 63 with different rates by the BMA. It can be seen that the estimation of Theorem 2 can precisely describe the distribution of λ for different code rates.

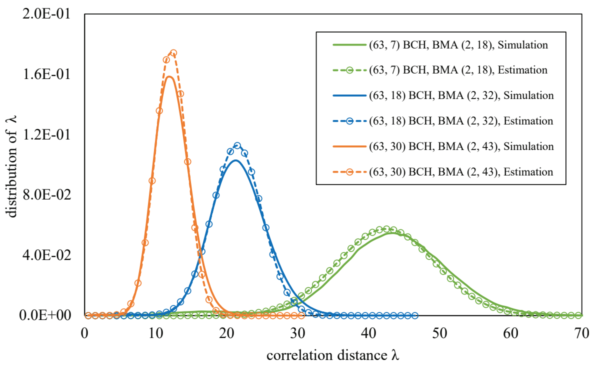

Figure 6. *Distribution of λ in decoding BCH codes with the BMA at SNR $= 2$ dB.*

V. COMPLEXITY ANALYSIS

In this section, complexity of the proposed low complexity SCL decoding is analyzed, including that of the OS-OSD and OS-BMA. We consider the complexity as the amount of required floating point operations (FLOPs) and binary operations (BOPs).

5.1 Complexity of OS-OSD and OS-BMA

In OS-OSD (τ) for an $(n,\, k,\, d)$ component code, the complexity attributes to the LLR sorting, the GE and the re-encoding. The same as its prototype [10], the LLR sorting requires $n\log_2 n$ FLOPs and the GE requires $n\cdot\min(k, n-k)^2$ BOPs. For the re-encoding complexity, it includes the FLOPs for correlation distance calculations and the BOPs for re-encoding. Let Ω_τ denote the total number of codeword candidates generated in the OSD, i.e.,

$$\Omega_\tau = \sum_{t=0}^{\tau}\binom{k}{t}. \tag{42}$$

Further let $P(S_t)$ denote the probability of the event that the re-encoding is terminated at the end of phase-t. The average number of FLOPs required by the correlation distance calculations can be estimated by

$$(n-k)\cdot\sum_{t=0}^{\tau}P(S_t)\Omega_t. \tag{43}$$

Let $\Theta_{\text{OSD}}(\tau)$ denote the number of BOPs required by the re-encoding in OSD (τ). It can be estimated by

$$\Theta_{\text{OSD}}(\tau) = (n-k)\left(k+\sum_{t=1}^{\tau}t\binom{k}{t}\right). \tag{44}$$

Therefore, the average number of BOPs required by the re-encoding of OS-OSD (τ) can be estimated by

$$\sum_{t=0}^{\tau}P(S_t)\Theta_{\text{OSD}}(t). \tag{45}$$

For the OS-BMA, complexity of the LLR sorting and the GE are the same as the above characterization. Its re-encoding complexity also consists of the correlation distance calculations and the re-encoding. To further analyze the re-encoding complexity of the OS-BMA, we consider the case that the skipping only occurs at the end of each phase. Let Ω'_τ denote the average number of codeword candidates that are generated in BMA (τ, s), which can be approximated by

$$\Omega'_\tau = \sum_{t=0}^{\tau}\binom{k}{t} + \sum_{t=\tau+1}^{2\tau}\binom{k}{t}\sum_{t'=0}^{2\tau-t}\frac{\binom{s-k}{t'}}{2^{s-k}}. \tag{46}$$

The average number of FLOPs required by the correlation distance calculations can be approximated by

$$(n-k)\cdot\sum_{t=0}^{\tau}P(S_t)\Omega'_t. \tag{47}$$

Moreover, let $\Theta_{\text{BMA}}(\tau)$ denote the average number of BOPs required by the re-encoding in BMA (τ, s). It can be approximated by

$$\Theta_{\text{BMA}}(\tau) = (n-k)\left(k+\sum_{t=1}^{2\tau}t\binom{k}{t}m(t)\right), \tag{48}$$

where

$$m(t) = \begin{cases} 1, & 0\le t\le\tau; \\ \displaystyle\sum_{t'=0}^{2\tau-t}\frac{\binom{s-k}{t'}}{2^{s-k}}, & \tau < t\le 2\tau. \end{cases} \tag{49}$$

The average number of BOPs required by the re-encoding can be approximated by

$$\sum_{t=0}^{\tau} P(S_t)\Theta_{\text{BMA}}(t). \tag{50}$$

Table 2 summarizes above complexity analysis, in which complexity of the OSD and BMA are also provided. Based on above analysis, it can be seen that $P(S_t)$ affects the average complexity of the OS-OSD and the OS-BMA. This implies the channel dependent feature of the component code decoding complexity. Our simulation results in Section VI will show that as the SNR increases, $P(S_0)$ increases and becomes dominant. Consequently, complexity of component code decoding is reduced.

Table 2. *Complexity of component code decoding.*

Operations	FLOPs	BOPs
LLR Sorting	$n\log_2 n$	–
GE	–	$n\cdot\min(k,n-k)^2$
OSD $(\tau)^\dagger$	$(n-k)\Omega_\tau$	$\Theta_{\text{OSD}}(\tau)$ in (44)
OS-OSD $(\tau)^\dagger$	$(n-k)\sum_{t=0}^{\tau}P(S_t)\Omega_t$	$\sum_{t=0}^{\tau}P(S_t)\Theta_{\text{OSD}}(t)$
BMA $(\tau,s)^\dagger$	$(n-k)\Omega'_\tau$	$\Theta_{\text{BMA}}(\tau)$ in (48)
OS-BMA $(\tau,s)^\dagger$	$(n-k)\sum_{t=0}^{\tau}P(S_t)\Omega'_t$	$\sum_{t=0}^{\tau}P(S_t)\Theta_{\text{BMA}}(t)$

† Complexity of the re-encoding.

5.2 Complexity of PP-SCL Decoding

For simplicity, we consider the FLOPs in analyzing the PP-SCL decoding complexity. Complexity of PP-SCL decoding attributes to the LLR updates between levels, the decoding of component codes and the decoding path sorting. With a decoding output list size of l, some of the l stored nodes of the layers may not be explored. Let $\bar{l}_i$ denote the average number of explored nodes at layer-i, where $1 \le \bar{l}_i \le l$.

The complexity of LLR updates of an H-level U-UV code can be characterized as that of decoding n polar codes of length $\gamma = 2^H$. Note that in decoding U-UV codes, the number of BOPs in LLR updates is negligible. The average required FLOPs can be approximated as

$$\frac{\sum_{i=1}^{\gamma}\bar{l}_i}{2^H} n 2^H \log_2 2^H = nH\sum_{i=1}^{\gamma}\bar{l}_i. \tag{51}$$

Let Γ_i denote the complexity of decoding component code $\mathcal{C}_i$. For the OSD, it can be characterized as $\Gamma_i = (n-k_i)\Omega_{\tau_i}$. Note that based on the analysis of Section 5.1, if the component codes are decoded by the OS-OSD, $\Gamma_i = (n-k_i)\sum_{t=0}^{\tau_i}P(S_t)\Omega_t$. Furthermore, if they are decoded by the OS-BMA, $\Gamma_i = (n-k_i)\sum_{t=0}^{\tau_i}P(S_t)\Omega'_t$. Hence, the complexity in decoding all component codes can be approximated as $\sum_{i=1}^{\gamma}\bar{l}_i\Gamma_i$.

The decoding path sorting exists in identifying the l most likely estimations from the component code decoding outputs and the subsequent $2l$-to-l decoding path sorting. Let Ψ_i denote the average number of codeword candidates generated in decoding $\mathcal{C}_i$. When the component codes are decoded by the OSD, $\Psi_i = \Omega_{\tau_i}$. If they are decoded by the OS-OSD or OS-BMA, it can be characterized as $\Psi_i = \sum_{t=0}^{\tau_i}P(S_t)\Omega_t$ or $\Psi_i = \sum_{t=0}^{\tau_i}P(S_t)\Omega'_t$, respectively. In order to identify the l most likely estimations from the decoding of $\mathcal{C}_i$, it requires $\Psi_i\log_2 l$ FLOPs. Hence, the sorting complexity can be approximated as $\sum_{i=1}^{\gamma}\bar{l}_i\Psi_i\log_2 l$. Furthermore, the $2l$-to-l decoding path sorting complexity is $\sum_{i=1}^{\gamma}\bar{l}_i \cdot l\log_2 l$. Hence, the total number of FLOPs required by the path sorting in PP-SCL decoding can be approximated as

$$\sum_{i=1}^{\gamma}\bar{l}_i(\Psi_i + l)\log_2 l. \tag{52}$$

The above analysis is summarized in Table 3, which also shows the prototype SCL decoding complexity [18]. It shows that the PP-SCL decoding realizes its low complexity feature by reducing the number of explored paths at each layer. Meanwhile, the OS-OSD and OS-BMA realize the complexity reduction by reducing the number of codeword candidates in component code decoding. Our simulations in Section VI will show that the complexity reduction brought by the proposed low complexity SCL decoding becomes more significant under good channel conditions.

Table 3. *SCL decoding complexity for U-UV codes.*

Operations	SCL	PP-SCL
LLR Update	$nH\gamma l$	$nH\sum_{i=1}^{\gamma}\bar{l}_i$
Component Code Decoding	$l\sum_{i=1}^{\gamma}\Gamma_i$	$\sum_{i=1}^{\gamma}\bar{l}_i\Gamma_i$
Path Sorting	$l\sum_{i=1}^{\gamma}(\Omega_{\tau_i}+l)\log_2 l$	eq. (52)

VI. SIMULATION RESULTS

The low complexity merit of the OS-OSD and OS-BMA will be first verified via simulations. Subsequently, they will be integrated with the PP-SCL in realizing low complexity SCL decoding for U–UV codes. The simulation results will show that significant complexity reduction can be achieved without losing decoding performance. Finally, the comparisons between the U-UV codes and the CRC-polar codes are provided.

1) Performance of the OS-OSD and the OS-BMA: Figure 7 shows the decoding frame error rate (FER) performance of the (63, 30) BCH code with different decoding schemes, in which their average number of required decoding FLOPs are also compared. Note that results of the fast OSD (FOSD) [26] is shown as a comparison benchmark, which is parameterized by the empirical factor $\beta = 0.05$. Figure 7 shows that compared with the original OSD, the OS-OSD yields a similar decoding performance and realizes a significant complexity reduction. In particular, the number of FLOPs required by the OS-OSD (3) converges to that of the OSD (0) at the high SNR regime. It indicates that as the SNR increases, $P(S_0)$ becomes dominant. That says in most cases, the first re-encoded codeword $\tilde{c}_0$ is the optimal one. The OS-OSD prevents further higher order decoding. Compared with the FOSD, the OS-OSD also exhibits a complexity advantage, especially at the low SNR regime. It shows that the proposed *a posteriori* CDLB is a more accurate correlation distance approximation than the FOSD. Moreover, both the BMA $(2, 43)$ [30] and the OS-BMA $(2, 43)$ yield a near ML decoding performance for the code, while the latter exhibits a lower complexity.

2) Performance of U-UV codes: In this work, the U-UV codes are designed based on the method of [18] with BCH component codes. In particular, the $(63, 57)$, $(63, 51)$, $(63, 45)$, $(63, 39)$, $(63, 36)$, $(63, 24)$, $(63, 18)$, $(63, 10)$ and $(63, 7)$ BCH codes are utilized as the component codes. Their OSD orders are $1, 1, 2, 2, 2, 3, 3, 3$ and 3, respectively. Their BMA orders and lengths of s, denoted as (τ, s), are (1,57), (1,51),(1,51), (1,45),(1,42), (2,30), (2,24), (2,16) and (2,18), respectively. Specifically, the 2-level $(252, 139)$ U-UV code is constructed by the $(63, 57)$, $(63, 39)$, $(63, 36)$ and $(63, 7)$ BCH codes.

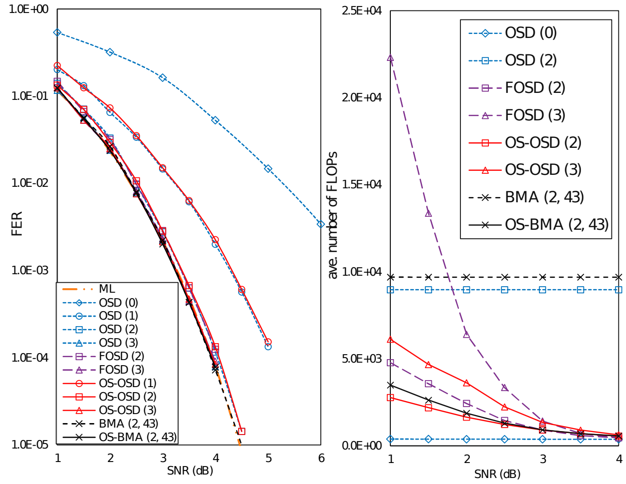

Figure 7. *Performance and complexity of the (63, 30) BCH code with different decoding schemes.*

The 3-level $(504, 250)$ U-UV code is constructed by the $(63, 57)$, $(63, 51)$, $(63, 45)$, $(63, 24)$, $(63, 45)$, $(63, 18)$, $(63, 10)$ and $(63, 0)$ BCH codes. For the PP-SCL decoding, $\theta = 10^{-4}$.

Figure 8 shows the decoding performance and complexity of the 3-level $(504, 250)$ U-UV code with different decoding schemes. The SCL decoding with the component code OSD employing the maximum likelihood criterion (MLC) [22] for early termination is denoted as MLC-SCL. Meanwhile, the SCL decoding with OS-OSD and OS-BMA are denoted as OS-OSD-SCL and OS-BMA-SCL, respectively. With a similar performance, the OS-OSD-SCL decoding requires a smaller number of FLOPs than the MLC-SCL decoding. This indicates that the proposed order skipping rule is more effective in identifying the optimal codeword candidate and terminating the decoding, especially at the low SNR regime. Moreover, the OS-BMA-SCL decoding can further reduce the complexity with a marginal performance loss.

Figure 9 further shows the performance and complexity of the $(504, 250)$ U-UV code with the low complexity SCL decoding schemes, i.e., the OS-OSD aided PP-SCL (OS-OSD-PP-SCL) decoding and the OS-BMA aided PP-SCL (OS-BMA-PP-SCL) decoding. It shows that, compared with the OS-BMA-SCL decoding, both the OS-OSD-PP-SCL decoding and OS-BMA-PP-SCL decoding suffer a slight performance loss. This is due to the fact that the accuracy of the estimated *a priori* CDLB is influenced by the pdf of correlation distance λ and the parameter θ as in

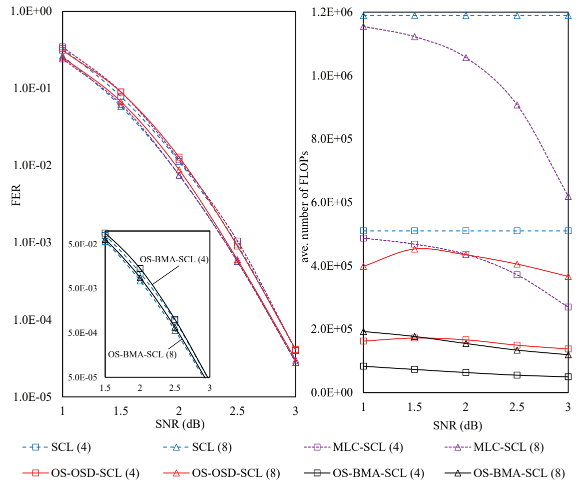

Figure 8. *Performance and complexity of the* $(504, 250)$ *U-UV code with different decoding schemes.*

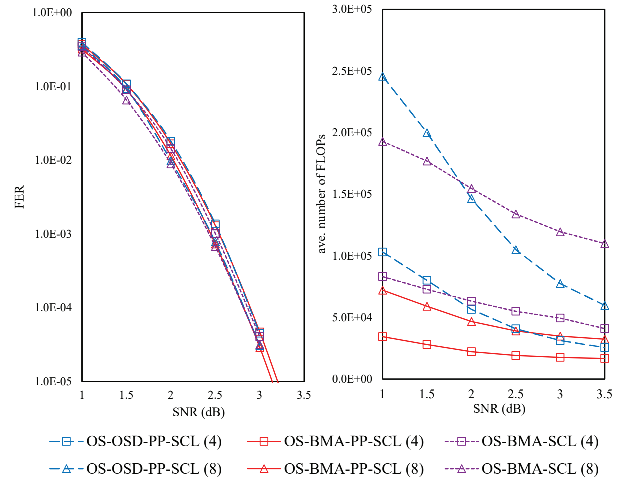

Figure 9. *Performance and complexity of the* $(504, 250)$ *U-UV code with the low complexity SCL decoding.*

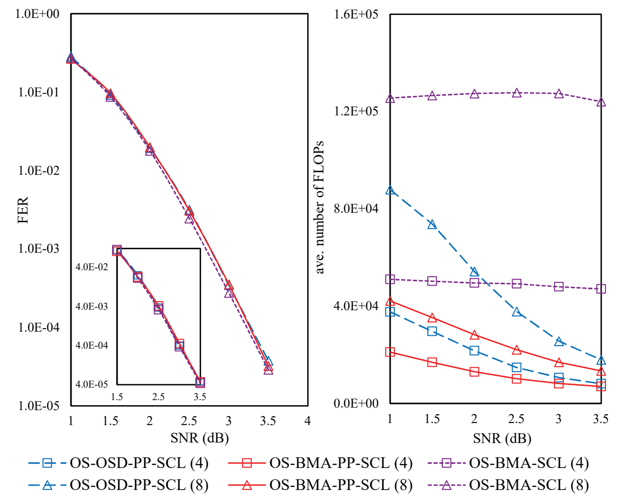

Figure 10. *Performance and complexity of the* $(252, 139)$ *U-UV code with the low complexity SCL decoding.*

(32). However, they both achieve a remarkable complexity reduction. It can also be seen that at the low SNR regime, the OS-OSD-PP-SCL decoding requires a larger amount of FLOPs than the OS-BMA-SCL decoding. This is because at the low SNR regime, the *a priori* CDLB of the component code decoding becomes larger. It results in the PP-SCL decoding being less effective in eliminating the unpromising decoding paths. Furthermore, Figure 10 shows the performance and complexity of the 2-level $(252, 139)$ U-UV code with different SCL decoding schemes. For this code, both the OS-OSD-PP-SCL decoding and OS-BMA-PP-SCL decoding require a smaller number of FLOPs than the OS-BMA-SCL decoding. Figures 9 and 10 indicate that integrating the PP-SCL decoding and the proposed OS-OSD or OS-BMA, low complexity SCL decoding becomes possible.

3) Comparison with polar codes: Figure 11 compares the proposed low complexity SCL decoding of the 3-level (504, 250) U-UV code with the CRC aided (CA)-SCL decoding of the $(512, 254)$ CRC-polar code. The CRC-polar code is designed by the 5th generation new radio (5G NR) standard and with a length-8 CRC outer concatenation. It shows that with the same decoding output list size, the U-UV code can achieve a coding gain over the CRC-polar code. Figure 12 further compares their the decoding complexity. The average number of decoding FLOPs and BOPs are shown. It shows that with the same decoding output list size, the OS-OSD-PP-SCL decoding and OS-BMA-PP-SCL decoding of the U-UV code require a similar number of FLOPs as the CA-SCL decoding of the CRC-polar code at the high SNR regime. It verifies the proposed low complexity SCL decoding, such that under a good channel condition, high order component code decoding are skipped and the unpromising SCL decoding path elaborations are curbed. Note that the number of BOPs required in decoding the U-UV code is greater than that of the CRC-polar code. This is due to the GE that is required by the component code decoding.

Furthermore, Figures 13 and 14 compare the performance and complexity of the 2-level (252,139) U-UV code with the (256,140) CRC-polar code that is assisted by a length-8 CRC code. It shows that with

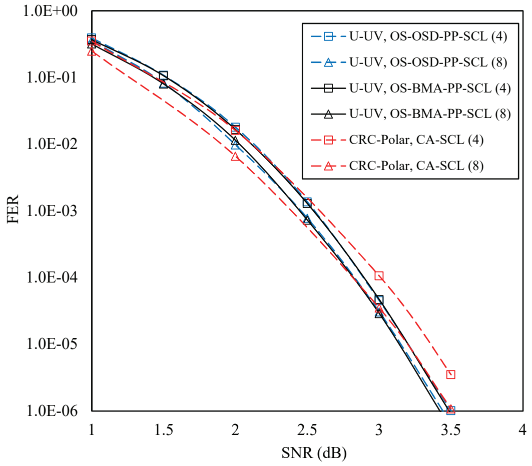

Figure 11. *Performance comparison between the (504, 250) U-UV code and the (512, 254) CRC-polar code.*

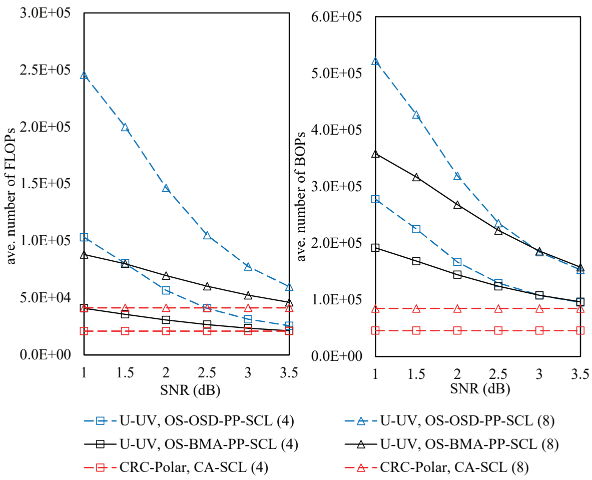

Figure 12. *Complexity comparison between the (504, 250) U-UV code and the (512, 254) CRC-polar code.*

the same decoding output list size, the U-UV code can also obtain a performance gain over the CRC-polar code. For this code, decoding the U-UV code still requires a greater number of BOPs. However, the number of FLOPs required by the U-UV code decoding becomes smaller than that required by the polar code decoding as the SNR increases. It demonstrates the effectiveness of the proposed low complexity SCL decoding. Therefore, it can be concluded that empowered by the proposed low complexity SCL decoding, the U-UV code is another competent short-to-medium length channel code that may be considered by the future communication systems.

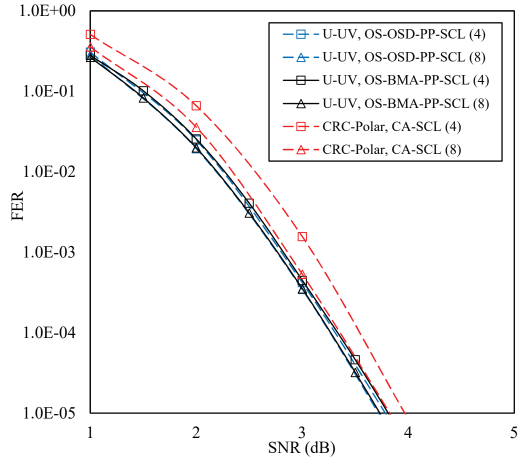

Figure 13. *Performance comparison between the (252, 139) U-UV code and the (256, 140) CRC-polar code.*

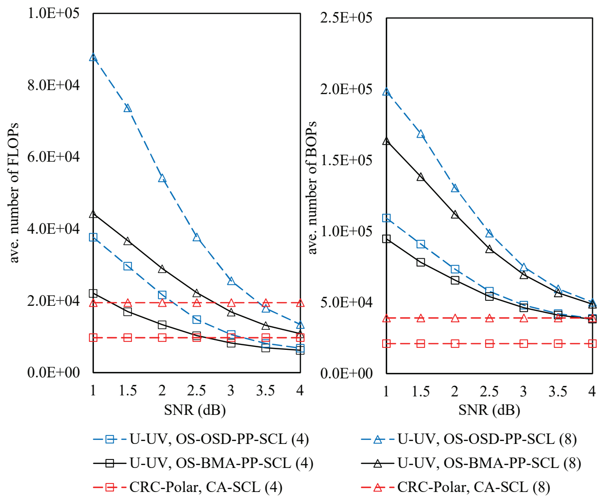

Figure 14. *Complexity comparison between the (252, 139) U-UV code and the (256, 140) CRC-polar code.*

VII. CONCLUSION

This paper has proposed the low complexity SCL decoding for the U-UV codes. The efficient order skipping rule has been proposed for the OSD through defining the *a posteriori* CDLB. This skipping rule has been further extended to the BMA, in which the redundant match operations can be skipped. Furthermore, substantiated the OS-OSD and OS-BMA component code decoding, the PP-SCL decoding is proposed through estimating the *a priori* CDLB of the component code decoding outputs. They lead to the proposed low complexity SCL decoding for U-UV codes. Complexity of the proposed SCL decoding is analyzed. Using BCH codes as the component codes, extensive simulation has been conducted. Our simulation results have shown that the low complexity SCL decoding can achieve a significant complexity reduction for U-UV codes with negligible performance loss. Armed with this low complexity SCL decoding, the U-UV codes can outperform the CRC-polar codes with a similar decoding complexity.

ACKNOWLEDGEMENT

This work was supported by the National Natural Science Foundation of China (NSFC) with project ID 62071498 and the Guangdong National Science Foundation (GDNSF) with project ID 2024A1515010213.

APPENDIX

A Proof of Theorem 2

Proof. Again assume that $\lambda(\tilde{e}_{\mathrm{B}})$, $\lambda(\tilde{e}_{\mathrm{S}})$ and $\lambda(\tilde{e}_{\mathrm{L}})$ are independent and Gaussian distributed. With the distribution of α_{B}, α_{S} and α_{L}, their expectations are determined by $\mathbb{E}[\alpha_{\mathrm{B}}] = \int_{\mu_{\tilde{\alpha}_k}}^{\infty} x f_{\alpha_{\mathrm{B}}}(x)dx$, $\mathbb{E}[\alpha_{\mathrm{S}}] = \int_{\mu_{\tilde{\alpha}_s}}^{\mu_{\tilde{\alpha}_k}} x f_{\alpha_{\mathrm{S}}}(x)dx$ and $\mathbb{E}[\alpha_{\mathrm{L}}] = \int_{0}^{\mu_{\tilde{\alpha}_s}} x f_{\alpha_{\mathrm{L}}}(x)dx$, respectively. Furthermore, their variances are given by $\mathbb{V}[\alpha_{\mathrm{B}}] = \int_{\mu_{\tilde{\alpha}_k}}^{\infty} (x - \mathbb{E}[\alpha_{\mathrm{B}}])^2 f_{\alpha_{\mathrm{B}}}(x)dx$, $\mathbb{V}[\alpha_{\mathrm{S}}] = \int_{\mu_{\tilde{\alpha}_s}}^{\mu_{\tilde{\alpha}_k}} (x - \mathbb{E}[\alpha_{\mathrm{S}}])^2 f_{\alpha_{\mathrm{S}}}(x)dx$ and $\mathbb{V}[\alpha_{\mathrm{L}}] = \int_{0}^{\mu_{\tilde{\alpha}_s}} (x - \mathbb{E}[\alpha_{\mathrm{L}}])^2 f_{\alpha_{\mathrm{L}}}(x)dx$, respectively.

Let Ω'_τ denote the average number of codeword candidates that are generated in BMA (τ, s), which can be approximated by

$$\Omega'_\tau = \sum_{t=0}^{\tau} \binom{k}{t} + \sum_{t=\tau+1}^{2\tau} \binom{k}{t} p(Q_{\mathrm{S}} \leq 2\tau - t), \tag{A.1}$$

where

$$p(Q_{\mathrm{S}} \leq 2\tau - t) = \sum_{t'=0}^{2\tau-t} \frac{\binom{s-k}{t'}}{2^{s-k}} \tag{A.2}$$

is the probability that Q_{S} is not greater than $2\tau - t$ under the binomial distribution assumption. Let W_{B} and W_{S} denote the total Hamming weight on the MRB and CB of all error patterns generated in the BMA, respectively. They can be approximated as

$$W_{\mathrm{B}} = \sum_{t=0}^{\tau} t\binom{k}{t} + \sum_{t=\tau+1}^{2\tau} t\binom{k}{t} p(Q_{\mathrm{S}} \leq 2\tau - t), \tag{A.3}$$

$$W_{\mathrm{S}} = \sum_{t=0}^{\tau} \binom{k}{t} \sum_{t'=0}^{s-k} \frac{t'\binom{s-k}{t'}}{2^{s-k}} + \sum_{t=\tau+1}^{2\tau} \binom{k}{t} \sum_{t'=0}^{2\tau-t} \frac{t'\binom{s-k}{t'}}{2^{s-k}}. \tag{A.4}$$

Subsequently, the expectations of Q_{B} and Q_{S} can be obtained by $\mathbb{E}[Q_{\mathrm{B}}] = W_{\mathrm{B}}/\Omega'_\tau$ and $\mathbb{E}[Q_{\mathrm{S}}] = W_{\mathrm{S}}/\Omega'_\tau$, respectively. It is also ensured that Q_{L} follow binomial distribution [25] with its probability mass function (pmf) given by

$$P(Q_{\mathrm{L}} = t) = \frac{\binom{n-s}{t}}{2^{n-s}}. \tag{A.5}$$

Hence, the expectation of Q_{L} can be approximated by

$$\mathbb{E}[Q_{\mathrm{L}}] = \sum_{t=0}^{n-s} \frac{t\binom{n-s}{t}}{2^{n-s}}. \tag{A.6}$$

Moveover, the variance of Q_{B} can be determined by

$$\mathbb{V}[Q_{\mathrm{B}}] = \sum_{t=0}^{2\tau} (t - \mathbb{E}[Q_{\mathrm{B}}])^2 P(Q_{\mathrm{B}} = t), \tag{A.7}$$

where $P(Q_{\mathrm{B}} = t)$ is the pmf of Q_{B}. Consider all the error patterns generated in the BMA, $P(Q_{\mathrm{B}} = t)$ can be calculated as

$$P(Q_{\mathrm{B}} = t) = \begin{cases} \dfrac{\binom{k}{t}}{\Omega'_\tau}, & 0 \leq t \leq \tau; \\ \dfrac{\binom{k}{t} \sum_{t'=0}^{2\tau-t} \frac{\binom{s-k}{t'}}{2^{s-k}}}{\Omega'_\tau}, & \tau < t \leq 2\tau. \end{cases} \tag{A.8}$$

For the variance of Q_{S}, it can be determined by

$$\mathbb{V}[Q_{\mathrm{S}}] = \sum_{t=0}^{s-k} (t - \mathbb{E}[Q_{\mathrm{S}}])^2 P(Q_{\mathrm{S}} = t), \tag{A.9}$$

where $P(Q_{\mathrm{S}} = t)$ is the pmf of Q_{S}. $P(Q_{\mathrm{S}} = t)$ can be calculated as

$$P(Q_{\mathrm{S}} = t) = \frac{\sum_{t'=0}^{2\tau} \binom{k}{t'} P(Q_{\mathrm{S}} = t \,|\, t')}{\Omega'_\tau}, \tag{A.10}$$

where $P(Q_{\mathrm{S}} = t \,|\, t')$ is probability of $w(\tilde{e}_{\mathrm{S}}) = t$ conditioning on $w(\tilde{e}_{\mathrm{B}}) = t'$ and it can be calculated as

$$P(Q_{\mathrm{S}} = t \,|\, t') = \begin{cases} \frac{\binom{s-k}{t}}{2^{s-k}}, & t' \leq \tau \text{ or } t + t' \leq 2\tau; \\ 0, & t + t' > 2\tau. \end{cases} \tag{A.11}$$

Based on (A.5), we have

$$\mathbb{V}[Q_{\mathrm{L}}] = \sum_{t=0}^{n-s} (t - \mathbb{E}[Q_{\mathrm{L}}])^2 \frac{\binom{n-s}{t}}{2^{n-s}}. \tag{A.12}$$

Summarize above analysis, we can obtain the expectations and variances of $\lambda(\tilde{e}_{\mathrm{B}})$, $\lambda(\tilde{e}_{\mathrm{S}})$ and $\lambda(\tilde{e}_{\mathrm{L}})$ based on Lemma 1. Further considering Lemma 2, we can obtain the Theorem 2.

References

[1] A. Hocquenghem, "Codes corecteurs d'erreurs," *Chiffres*, 1959, vol. 2, no. 2, pp. 147–156.

[2] C. Bose and K. RayChaudhuri, "On a class of error correcting binary group codes," *Information and Control*, 1960, vol. 3, no. 1, pp. 68–79.

[3] H. Ma and J. Wolf, "On tail biting convolutional codes," *IEEE Transactions on Communication*, 1986, vol. 34, no. 2, pp. 104–111.

[4] E. Arikan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Transactions on Information Theory*, 2009, vol. 55, no. 7, pp. 3051–3073.

[5] K. Niu and K. Chen, "Crc-aided decoding of polar codes," *IEEE Communications Letters*, 2012, vol. 16, no. 10, pp. 1668–1671.

[6] I. Tal and A. Vardy, "List decoding of polar codes," *IEEE Transactions on Information Theory*, 2015, vol. 61, no. 5, pp. 2213–2226.

[7] K. Niu, P. Zhang, *et al.*, "A golden decade of polar codes: From basic principle to 5g applications," *China Communications*, 2023, vol. 20, no. 2, pp. 94–121.

[8] E. Arıkan, "From sequential decoding to channel polarization and back again," *Preprint available as arXiv:1908.09594*, 2019.

[9] M. Coskun, G. Durisi, *et al.*, "Efficient error-correcting codes in the short blocklength regime," *Physical Communication*, 2019, vol. 34, pp. 66–79.

[10] M. Fossorier and S. Lin, "Soft-decision decoding of linear block codes based on ordered statistics," *IEEE Transactions on Information Theory*, 1995, vol. 41, no. 5, pp. 1379–1396.

[11] T. Wang, D. Qu, *et al.*, "Parity-check-concatenated polar codes," *IEEE Communications Letters*, 2016, vol. 20, no. 12, pp. 2342–2345.

[12] T. Jiang, Y. Liu, *et al.*, "Pcc polar codes for future wireless communications: Potential applications and design guidelines," *IEEE Wireless Communications*, Early Access, 2023, pp. 1–7.

[13] T. Wang, D. Qu, *et al.*, "An incremental redundancy hybrid arq scheme with non-catastrophic puncturing of polar codes," in *SCC 2019; 12th International ITG Conference on Systems, Communications and Coding*, 2019.

[14] Y. Liu, L. Xiao, *et al.*, "Channel coding for satellite-terrestrial integrated communication: Classic applications, key technologies, and challenges," *IEEE Wireless Communications*, 2024, vol. 31, no. 3, pp. 348–354.

[15] Z. Feng, C. Niu, *et al.*, "List-serial pipelined hardware architecture for SCL decoding of polar codes," *China Communications*, 2023, vol. 20, no. 3, pp. 175–184.

[16] M. Irene and J. Tillich, "Attaining capacity with iterated $(U|U+V)$ codes based on ag codes and koetter-vardy soft decoding," in *IEEE International Symposium on Information Theory (ISIT)*, 2017, pp. 6–10.

[17] J. Cheng and L. Chen, "Bch based u-uv codes and its decoding," in *2021 IEEE International Symposium on Information Theory (ISIT)*, 2021, pp. 1433–1438.

[18] W. Chen, J. Cheng, *et al.*, "Bch based u-uv codes and its scl decoding," *IEEE Transactions on Signal Processing*, 2024, vol. 72, pp. 1286–1300.

[19] C. Wu, W. Chen, *et al.*, "U-uv coding for bit-interleaved coded modulation," *IEEE Transactions on Communications*, 2023, vol. 71, no. 9, pp. 5065–5077.

[20] M. Plotkin, "Binary codes with specified minimum distance," *IRE Transactions on Information Theory*, 1960, vol. 6, no. 4, pp. 445–450.

[21] P. Trifonov, "Efficient design and decoding of polar codes," *IEEE Transactions on Communications*, 2012, vol. 60, no. 11, pp. 3221–3227.

[22] T. Kaneko, T. Nishijima, *et al.*, "An efficient maximum-likelihood-decoding algorithm for linear block codes with algebraic decoder," *IEEE Transactions on Information Theory*, 1994, vol. 40, no. 2, pp. 320–327.

[23] W. Jin and M. Fossorier, "Probabilistic sufficient conditions on optimality for reliability based decoding of linear block codes," in *IEEE International Symposium on Information Theory (ISIT)*, 2006, pp. 2235–2239.

[24] Y. Wu and C. Hadjicostis, "Soft-decision decoding using ordered recodings on the most reliable basis," *IEEE Transactions on Information Theory*, 2007, vol. 53, no. 2, pp. 829–836.

[25] C. Yue, M. Shirvanimoghaddam, *et al.*, "A revisit to ordered statistics decoding: Distance distribution and decoding rules," *IEEE Transactions on Information Theory*, 2021, vol. 67, no. 7, pp. 4288–4337.

[26] C. Choi and J. Jeong, "Fast and scalable soft decision decoding of linear block codes," *IEEE Communications Letters*, 2019, vol. 23, no. 10, pp. 1753–1756.

[27] C. Yue, M. Shirvanimoghaddam, *et al.*, "Segmentation-discarding ordered-statistic decoding for linear block codes," in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6.

[28] C. Kim, D. Kam, *et al.*, "Simplified ordered statistic decoding for short-length linear block codes," *IEEE Communications Letters*, 2022, vol. 26, no. 8, pp. 1720–1724.

[29] Y. Han, K. Niu, *et al.*, "Guessing error pattern ordered statistic decoding for short polar codes," *IEEE Communications

Letters, 2023, vol. 27, no. 11, pp. 2904–2908.
[30] A. Valembois and M. Fossorier, "Box and match techniques applied to soft-decision decoding," *IEEE Transactions on Information Theory*, 2004, vol. 50, no. 5, pp. 796–810.
[31] C. Yue, M. Shirvanimoghaddam, *et al.*, "Linear-equation ordered-statistics decoding," *IEEE Transactions on Communications*, 2022, vol. 70, no. 11, pp. 7105–7123.
[32] J. Liang, Y. Wang, *et al.*, "A low-complexity ordered statistic decoding of short block codes," *IEEE Communications Letters*, 2023, vol. 27, no. 2, pp. 400–403.
[33] C. Choi and J. Jeong, "Fast soft decision decoding algorithm for linear block codes using permuted generator matrices," *IEEE Communications Letters*, 2021, vol. 25, no. 12, pp. 3775–3779.
[34] C. Yue, M. Shirvanimoghaddam, *et al.*, "Ordered-statistics decoding with adaptive gaussian elimination reduction for short codes," in *2022 IEEE Globecom Workshops (GC Wkshps)*, 2022, pp. 492–497.
[35] L. Yang and L. Chen, "Low-latency ordered statistics decoding of bch codes," in *2022 IEEE Information Theory Workshop (ITW)*, 2022, pp. 404–409.
[36] D. Bailon, M. Bossert, *et al.*, "Concatenated codes based on the Plotkin construction and their soft-input decoding," *IEEE Transactions on Communication*, 2022, vol. 70, no. 5, pp. 2939–2950.
[37] W. Chen, L. Chen, *et al.*, "Reduced complexity scl decoding of u-uv codes through list pruning," in *IEEE/CIC International Conference on Communications in China (ICCC Workshops)*, 2023, pp. 1–6.
[38] K. Chen, K. Niu, *et al.*, "A reduced-complexity successive cancellation list decoding of polar codes," in *IEEE Vehicular Technology Conference (VTC Spring)*, 2013, pp. 1–5.
[39] Z. Zhang, L. Zhang, *et al.*, "A split-reduced successive cancellation list decoder for polar codes," *IEEE Journal on Selected Areas in Communications*, 2016, vol. 34, no. 2, pp. 292–302.
[40] K. Chen, B. Li, *et al.*, "Reduce the complexity of list decoding of polar codes by tree-pruning," *IEEE Communications Letters*, 2016, vol. 20, no. 2, pp. 204–207.
[41] J. Chen, Y. Fan, *et al.*, "Low-complexity list successive-cancellation decoding of polar codes using list pruning," in *IEEE Global Communications Conference (GLOBECOM)*, 2016, pp. 1–6.
[42] M. Rowshan and E. Viterbo, "Stepped list decoding for polar codes," in *IEEE International Symposium on Turbo Codes Iterative Information Processing (ISTC)*, 2018, pp. 1–5.
[43] H. Mahdavifa, M. ElKhamy, *et al.*, "Performance limits and practical decoding of interleaved reed-solomon polar concatenated codes," *IEEE Transactions on Communication*, 2014, vol. 62, no. 5, pp. 1406–1417.
[44] H. Saber and I. Marsland, "Design of generalized concatenated codes based on polar codes with very short outer codes," *IEEE Transactions on Vehicular Technology*, 2017, vol. 66, no. 4, pp. 3103–3115.
[45] D. Goldin and D. Burshtein, "Performance bounds of concatenated polar coding schemes," *IEEE Transactions on Information Theory*, 2019, vol. 65, no. 11, pp. 7131–7148.
[46] Q. Jan, Y. Chao, *et al.*, "Improved belief propagation decoder for ldpc-crc-polar codes with bit-freezing," *China Communications*, 2024, vol. 21, no. 7, pp. 135–148.
[47] A. Papoulis and S. Pillai, *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, 2002.
[48] F. Wang, J. Jiao, *et al.*, "Adjustable ordered statistic decoder for short block length code towards urllc," in *International Conference on Wireless Communications and Signal Processing (WCSP)*, 2021, pp. 1–5.

Biographies

Chen Wenhao received the B.Sc. degree in information engineering and the M.Sc. degree in information and communication engineering from Sun Yat-sen University, Guangzhou, China, in 2021 and 2024, respectively. His research interests include channel coding and data communications.

Chen Li received the B.Sc. degree in applied physics from Jinan University, China, in 2003, and the M.Sc. degree in communications and signal processing and the Ph.D degree in communications engineering from Newcastle University, U.K., in 2004 and 2008, respectively. From 2007 to 2010, he was a Research Associate with Newcastle University. In 2010, he returned China as a Lecturer of the School of Information Science and Technology, Sun Yat-sen University, Guangzhou. From 2011 to 2012, he was a Visiting Researcher with the Institute of Network Coding, the Chinese University of Hong Kong, Hong Kong, China. From 2011 and 2016, he was an Associate Professor and a Professor of the university. Since 2013, he has been the Associate Head of the Department of Electronic and Communication Engineering (ECE). From July 2015 to October 2015, he was a Visitor of the Institute of Communications Engineering, Ulm University, Germany. From October 2015 to June 2016, he was a Visiting Associate Professor with the Department of Electrical Engineering, University of Notre Dame, USA. From 2017 to 2020, he was the Deputy Dean of the School of Electronics and Communication Engineering. His research interests include information theory, error-correction codes, and data communications. He is a Senior Member of the Chinese Institute of Electronics (CIE). He is a member of the IEEE Information Theory Society Board of Governors and its External Nomination Committee and the Chair of its Conference Committee. He is also the Chair of the IEEE Information Theory Society Guangzhou Chapter. He is an Associate Editor of *IEEE Transaction on Communications*. He has been organizing several international conferences and workshops, including the 2018 IEEE Information Theory Workshop (ITW) at Guangzhou and the 2022 IEEE East Asian School of Information Theory (EASIT) at Shenzhen, for which he is the General Co-Chair. He is also the TPC Co-Chair of the 2022 IEEE/CIC International Conference on Communications in China (ICCC) at Foshan. He likes reading and photography.

Lin Jingyu received the B.Sc. degree in electronic information science and technology from Sun Yat-sen University, Guangzhou, China, in 2023. He is currently pursuing the M.Sc. degree in information and communication engineering in Sun Yat-sen University, Guangzhou, China. His research interests include channel coding and data communications.

Zhang Huazi received his B.Sc. and Ph.D degrees from Institute of Information and Communication Engineering in 2008 and 2013, respectively, from Zhejiang University. From 2011 to 2013, he was a visiting researcher with the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC, USA. From 2013 to 2014, he was a Research Fellow with the School of Computer Engineering, Nanyang Technological University, Singapore. From 2014 to 2015, he was a Research Scholar with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL, USA. He has over a hundred journal, conference publications and patent applications in the areas of coding techniques, wireless communications and signal processing. He has served on the Technical Program Committee of communications theory, wireless communications and signal processing symposia for IEEE Globecom and ICC since 2013. He was an Exemplary Reviewer for *IEEE Transactions on Communications* in the year of 2014. He joined Huawei Technologies Co., Ltd in 2015. Since then, he has engaged in research projects on advanced wireless communications involving channel coding and signal processing techniques, and engaged in multiple standardization activities. His contribution led to the adoption of many state-of-the-art research results into 5G standards, as well as the subsequent commercial rollout. His current research interests are channel coding, information theory and signal processing for wireless communications, with focus on theoretical analysis, algorithm design and hardware implementations for 5G and beyond.