BCH Based U-UV Codes and Its Decoding

Jinjun Cheng †, Li Chen ‡
† School of Electronics and Communication Engineering, Sun Yat-sen University, Guangzhou, China
‡ School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou, China
Email: chengjj5@mail2.sysu.edu.cn, chenli55@mail.sysu.edu.cn

***Abstract*—This paper proposes the U-UV structured codes with BCH codes as its components. This coding construction will lead to polarized subchannels and rate of each component code can be designed accordingly. The component code will be decoded by the ordered statistic decoding (OSD), yielding multiple decoding outcomes. Integrated in a successive cancellation (SC) decoding mechanism, SC-list (SCL) decoding of the U-UV codes is further proposed. Our simulation results will show that over the short-to-medium length regime, SCL decoding of a BCH based U-UV code can outperform that of a similar rate polar code.**

***Index Terms*—BCH codes, successive cancellation list decoding, short-to-medium length codes, U-UV codes**

I. Introduction

To realize ultra-reliable low-latency communications, it is important to design good performing short-to-medium length codes. This paper investigates this aspect by exploring the (U|U+V) construction, where the U code and the V code are component codes of equal length [1]. For simplicity, we name it the U-UV construction, and the constructed codes the U-UV codes. They have been traditionally used for unequal error protection (UEP) [2–4], since the information carried by the U code is better protected than that carried by the V code.

Reed-Muller (RM) [5, 6] and polar [7] codes can also be interpreted by the U-UV structure. In particular, the research of polar codes has shown that the U-UV construction leads to polarized subchannels with their capacities approaching 0 or 1. High decoding performance can be achieved by designing the transmission according to the polarized subchannel capacities. This inspires a more general U-UV structured code that is yielded by a number of component codes, each of which is transmitted through one of the subchannels. Rates of the component codes can be designed based on the subchannel capacities. The U-UV codes can also be interpreted as the generalized concatenated codes [8, 9], where the polar code is used as the inner code. In addition, it has been shown that using algebraic-geometric codes as component codes and soft decoding, the U-UV codes can attain the discrete symmetric channel capacity [10]. However, a more practical U-UV structured code with simpler component codes is yet to be developed. Moreover, it has been proven that with the successive cancellation (SC) decoding, polar codes can achieve the capacity of the binary input symmetric discrete memoryless channel [7]. But this is conditioned on codeword length $n \rightarrow \infty$. When n remains small or medium, the incompleteness of channel polarization starts to weight, since there will be a larger portion of subchannels without a polarized capacity. This will downgrade the effectiveness of assigning transmission symbol wise as in polar coding. Instead, viewing polar coding as a generalized structure filled by component codes and designing the component code rates based on the subchannel capacities can overcome this challenge. Therefore, the U-UV coding offers an alternative solution for designing competent short-to-medium length codes.

This paper proposes the U-UV structured codes with BCH codes as its components. The polarized subchannel capacities will be estimated by Gaussian approximation (GA) [11], and rate of each component code can be designed accordingly. The component codes are decoded by the ordered statistic decoding (OSD) [12], providing close to maximum likelihood (ML) decoding performance. Meanwhile, multiple estimations for each component code will be provided. Integrated in an SC decoding mechanism, SC-list (SCL) decoding of the U-UV code is further proposed. Our simulation results will show that in the short-to-medium length regime, the SCL decoding of the U-UV code can outperform that of a similar rate polar code.

II. BCH Based U-UV Codes

This section introduces the code construction and its design.

A. Code Construction

Definition I. Let U code be an $(n, k_{\mathrm{U}}, d_{\mathrm{U}})$ linear block code, and V code an $(n, k_{\mathrm{V}}, d_{\mathrm{V}})$ one, where k_{U} and k_{V} are their dimensions, and d_{U} and d_{V} are their minimum Hamming distances, respectively. The U-UV code will be a $(2n, k, d)$ linear block code that is constructed by

$$(\mathrm{U}|\mathrm{U}+\mathrm{V}) = \{(\boldsymbol{u}|\boldsymbol{u}+\boldsymbol{v}); \boldsymbol{u} \in \mathrm{U} \text{ and } \boldsymbol{v} \in \mathrm{V}\}, \tag{1}$$

where $k = k_{\mathrm{U}} + k_{\mathrm{V}}$, and $d = \min\{2d_{\mathrm{U}}, d_{\mathrm{V}}\}$ is the minimum Hamming distance of the U-UV code. Without loss of generality, in this paper, U ($\boldsymbol{u}$) and V ($\boldsymbol{v}$) also denote the codebook (codeword) of the U code and the V code, respectively.

In the proposed coding scheme, the above U code and V code are two BCH codes. This construction can be extended recursively by involving more component BCH codes, forming a larger U-UV code. This requires a multilevel U-UV construction. In general, if a U-UV code is constructed by H levels, it consists of 2^H component BCH codes. For simplicity, let $\mathrm{U}_i^{(h)}$ ($\boldsymbol{u}_i^{(h)}$) denote the ith component code (codeword) at level-h, where $h = 0, 1, ..., H$ and $i = 1, 2, ..., 2^{H-h}$. In particular, when $h = 0$, the component code $\mathrm{U}_i^{(0)}$ becomes an (n, k_i) BCH code with rate $R_i = \frac{k_i}{n}$. Therefore, the ith component

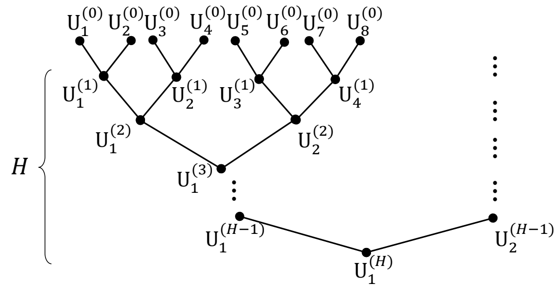

Fig. 1. Recursive construction of an H levels U-UV structured code.

code of level-h ($h \geq 1$) is constructed by

$$\mathrm{U}_i^{(h)} = (\mathrm{U}_{2i-1}^{(h-1)} | \mathrm{U}_{2i-1}^{(h-1)} + \mathrm{U}_{2i}^{(h-1)}). \tag{2}$$

Note that a component code of level-h has length-$2^h n$. Fig. 1 illustrates the recursive construction of an H levels U-UV structured code.

B. Component Code Rate Design

The above U-UV construction yields 2^{H-h} subchannels at level-h, each of which has a different capacity. Let $\mathrm{W}_i^{(h)}$ denote an equivalent subchannel at level-h and it is used to transmit component code $\mathrm{U}_i^{(h)}$. Let $I(\mathrm{W}_i^{(h)})$ denote the capacity of subchannel $\mathrm{W}_i^{(h)}$. As the number of levels H enlarges, the subchannel capacities at level-0, i.e., $I(\mathrm{W}_1^{(0)}), I(\mathrm{W}_2^{(0)}), ..., I(\mathrm{W}_{2^H}^{(0)})$, will polarize [7]. Since $\mathrm{U}_i^{(0)}$ is transmitted through $\mathrm{W}_i^{(0)}$, its code rate R_i can be chosen according to the subchannel capacity $I(\mathrm{W}_i^{(0)})$ as

$$R_i \leq I(\mathrm{W}_i^{(0)}). \tag{3}$$

GA can be utilized to estimate the subchannel capacities. Let $u_{i,j}^{(h)}$ denote the jth symbol of a component codeword $\boldsymbol{u}_i^{(h)}$, where $j = 0, 1, ..., 2^h n - 1$ and $L_{i,j}^{(h)}$ denote its log-likelihood ratio (LLR). Assume that codeword $\boldsymbol{u}_1^{(H)}$ is transmitted through the additive white Gaussian noise (AWGN) channel with a noise variance of σ^2. LLRs of its symbols, denoted as $L_{1,0}^{(H)}, L_{1,1}^{(H)}, ..., L_{1,2^H n-1}^{(H)}$, can also be considered as Gaussian variables with a mean of $\frac{2}{\sigma^2}$ and a variance of $\frac{4}{\sigma^2}$. It is denoted as $L_{1,j}^{(H)} \sim \mathcal{N}(\frac{2}{\sigma^2}, \frac{4}{\sigma^2})$. In general, statistics of the LLRs of a level-$(h-1)$ component code can be computed with that of level-h as [11]

$$\mathbb{E}[L_{2i-1,j'}^{(h-1)}] = 2\mathbb{E}[L_{i,j}^{(h)}], \tag{4}$$

$$\mathbb{E}[L_{2i,j'}^{(h-1)}] = \phi^{-1}(1 - (1 - \phi(\mathbb{E}[L_{i,j}^{(h)}]))^2), \tag{5}$$

where $j = 0, 1, \ldots, 2^h n - 1$, $j' = 0, 1, \ldots, 2^{h-1} n - 1$, and

$$\phi(x) = \begin{cases} 1 - \frac{1}{\sqrt{4\pi x}} \int \tanh(\frac{\nu}{2}) \exp(-\frac{(\nu - x)^2}{4x}) d\nu, & x > 0, \\ 1, & x = 0. \end{cases} \tag{6}$$

Performing the computations recursively through the levels, $\mathbb{E}[L_{i,j}^{(0)}]$ can be determined. They are utilized to further determined the subchannel capacities $I(\mathrm{W}_i^{(0)})$.

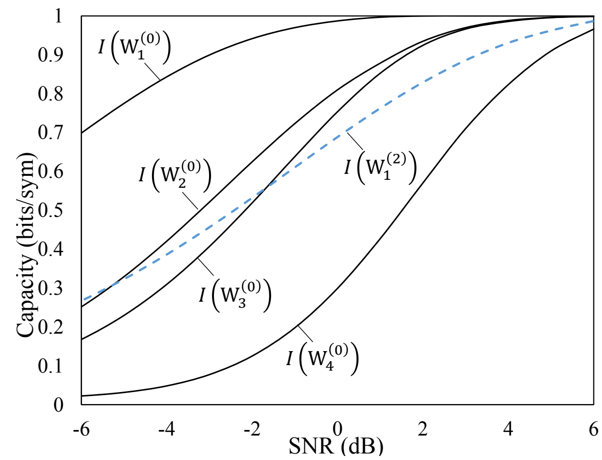

Fig. 2. Subchannel capacities $\mathrm{W}_i^{(0)}$ of 2 levels U-UV construction.

The subchannel capacity $I(\mathrm{W}_i^{(0)})$ can be determined using the Monte-Carlo simulation. With knowledge of $\mathbb{E}[L_{i,j}^{(0)}]$, the equivalent noise variance of subchannel $\mathrm{W}_i^{(0)}$ can be determined by $\sigma_i^2 = \frac{2}{\mathbb{E}[L_{i,j}^{(0)}]}$.

Considering $\mathrm{W}_i^{(0)}$ as an AWGN channel with a variance of σ_i^2, its capacity can be determined by

$$I(\mathrm{W}_i^{(0)}) = \frac{1}{2} \sum_{u_{i,j'}^{(0)} \in \{0,1\}} \mathbb{E}\left[\log_2 \frac{p(y_{j'} | u_{i,j'}^{(0)})}{\frac{1}{2} \sum_{u_{i,j}^{(0)} \in \{0,1\}} p(y_j | u_{i,j}^{(0)})}\right], \tag{7}$$

where y_j (y_j') is the received symbol that carries $u_{i,j}^{(0)}$ ($u_{i,j'}^{(0)}$), and $p(y_j | u_{i,j}^{(0)})$ ($p(y_{j'} | u_{i,j'}^{(0)})$) is the channel observation.

Fig. 2 shows the subchannel capacities $I(\mathrm{W}_i^{(0)})$ with $H = 2$. Rate of each component BCH code can be designed accordingly. For example, to achieve a transmission rate of 0.78bits/sym., the signal-to-noise ratio (SNR) should be at least 1dB, at which $I(\mathrm{W}_1^{(0)}) = 0.99$, $I(\mathrm{W}_2^{(0)}) = 0.87$, $I(\mathrm{W}_3^{(0)}) = 0.85$ and $I(\mathrm{W}_4^{(0)}) = 0.42$. Hence, based on (3), rates of $\mathrm{U}_1^{(0)}$, $\mathrm{U}_2^{(0)}$, $\mathrm{U}_3^{(0)}$ and $\mathrm{U}_4^{(0)}$ can be designed accordingly.

III. SCL Decoding

SCL decoding of the U-UV codes is built upon the OSD of its component codes and the SC decoding mechanism for the structured codes. They will be first introduced.

A. OSD of BCH Codes

Assumed that an (n, k) BCH codeword $\boldsymbol{c} = (c_0, c_1, \ldots, c_{n-1})$ is transmitted through a discrete memoryless channel using BPSK modulation. Let $\boldsymbol{y} = (y_0, y_1, \ldots, y_{n-1}) \in \mathbb{R}^n$ denote the received symbol vector and $\boldsymbol{L} = (L_0, L_1, \ldots, L_{n-1}) \in \mathbb{R}^n$ denote the corresponding LLR vector with entries defined as

$$L_j = \ln \frac{p(y_j | c_j = 0)}{p(y_j | c_j = 1)}, \tag{8}$$

where $j = 0, 1, ..., n-1$. Hard-decision r_j on each coded symbol c_j can be made. If $L_j \geq 0$, $r_j = 0$; otherwise, $r_j = 1$. Note that a larger $|L_j|$ implies the decision on c_j is more reliable. Therefore, based on $|L_j|$, k most reliable independent decisions are chosen to form an information vector $\boldsymbol{m}^{(0)}$. Meanwhile, a $k \times n$ generator matrix $\mathbf{G}$ that is in the reduced echelon form will be obtained, in which the k weight-1 columns correspond to the symbols of $\boldsymbol{m}^{(0)}$. The OSD of order $\tau (\tau \leq k)$ consists of τ decoding phases. Let β denote the OSD phase index, and $\beta = 1, \ldots, \tau$. At phase-β, β bits of $\boldsymbol{m}^{(0)}$ will be flipped, yielding information vectors $\boldsymbol{m}^{(1)}, \boldsymbol{m}^{(2)}, \ldots, \boldsymbol{m}^{\binom{k}{\beta}}$. Using $\mathbf{G}$, they will be encoded into codeword candidates $\hat{\boldsymbol{c}}^{(1)}, \hat{\boldsymbol{c}}^{(2)}, \ldots, \hat{\boldsymbol{c}}^{\binom{k}{\beta}}$, respectively. Therefore, the order-τ OSD will produce a list of

$$l_\tau = \binom{k}{0} + \binom{k}{1} + \cdots + \binom{k}{\tau} \tag{9}$$

decoding estimations. Likelihood of each estimation $\hat{\boldsymbol{c}}^{(\mu)} = (\hat{c}_0^{(\mu)}, \hat{c}_1^{(\mu)}, \ldots, \hat{c}_{n-1}^{(\mu)})$ can be measured by the correlation distance defined as [12]

$$\lambda(\boldsymbol{L}, \hat{\boldsymbol{c}}^{(\mu)}) = \sum_{j=0, r_j \neq \hat{c}_j^{(\mu)}}^{n-1} |L_j|, \tag{10}$$

where $\mu = 1, \ldots, l_\tau$. A smaller $\lambda(\boldsymbol{L}, \hat{\boldsymbol{c}}^{(\mu)})$ indicates the estimation $\hat{\boldsymbol{c}}^{(\mu)}$ is more likely to be the transmitted codeword.

B. The SC Decoding Mechanism

Given an H levels U-UV code, assume its codeword $\boldsymbol{u}_1^{(H)}$ is transmitted through a memoryless channel using BPSK modulation. Its received symbol vector is $\boldsymbol{y} = (y_0, y_1, ..., y_{2^H n-1}) \in \mathbb{R}^{2^H n}$. Let us define

$$\Omega(h) = \{1, 2, ..., 2^{H-h}\} \tag{11}$$

as the set of component code indices at level-h, and

$$\Omega_i(h) = \{i+1, i+2, ..., 2^{H-h}\} \tag{12}$$

as the subset of $\Omega(h)$. At level-h, the SC decoding exhibits a decoding order of $\mathrm{U}_{2^{H-h}}^{(h)} \to \mathrm{U}_{2^{H-h}-1}^{(h)} \to \cdots \to \mathrm{U}_1^{(h)}$. Hence, $\Omega_i(h)$ denotes the component codes that are decoded prior to $\mathrm{U}_i^{(h)}$. Note that at level-H, the overall U-UV code is constituted and $\Omega_i(H) = \emptyset$. Given $\boldsymbol{u}_i^{(h)} = (u_{i,0}^{(h)}, u_{i,1}^{(h)}, ..., u_{i,2^h n-1}^{(h)})$, under the SC decoding mechanism, LLR of its symbol is defined as

$$L_{i,j}^{(h)} = \ln \frac{p(\{y_{j+(\eta-1)2^h n}\}_{\eta \in \Omega(h)}, \{\hat{u}_{\rho,j}^{(h)}\}_{\rho \in \Omega_i(h)} | u_{i,j}^{(h)} = 0)}{p(\{y_{j+(\eta-1)2^h n}\}_{\eta \in \Omega(h)}, \{\hat{u}_{\rho,j}^{(h)}\}_{\rho \in \Omega_i(h)} | u_{i,j}^{(h)} = 1)}, \tag{13}$$

where $j = 0, 1, ..., 2^h n - 1$, $\hat{u}_{\rho,j}^{(h)}$ is an estimation of $u_{\rho,j}^{(h)}$ and $p(\{y_{j+(\eta-1)2^h n}\}_{\eta \in \Omega(h)}, \{\hat{u}_{\rho,j}^{(h)}\}_{\rho \in \Omega_i(h)} | u_{i,j}^{(h)})$ are the transition probabilities of subchannel $\mathrm{W}_i^{(h)}$. In particular, when $h = H$,

$$L_{1,j}^{(H)} = \ln \frac{p(y_j | u_{1,j}^{(H)} = 0)}{p(y_j | u_{1,j}^{(H)} = 1)}, \tag{14}$$

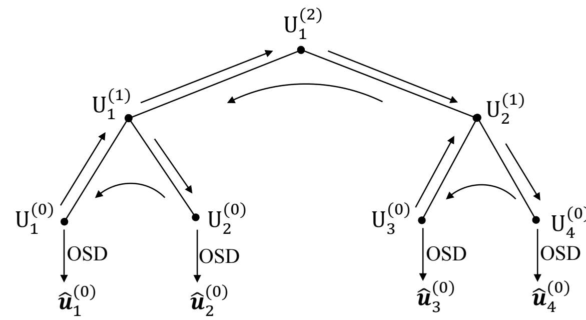

Fig. 3. SC Decoding of a 2 levels U-UV code.

which can be obtained based on $\boldsymbol{y}$. The LLRs will be updated level-by-level until $L_{i,j}^{(0)}$ are produced. In particular, let us define

$$f(\mathcal{X}, \mathcal{Y}) \triangleq \ln \frac{e^{\mathcal{X}} e^{\mathcal{Y}} + 1}{e^{\mathcal{X}} + e^{\mathcal{Y}}}, \tag{15}$$

where $\mathcal{X}, \mathcal{Y} \in \mathbb{R}$. With the U-UV construction, the LLRs of level-$(h-1)$ can be determined by those of level-h as

$$L_{2i,j'}^{(h-1)} = f(L_{i,j'}^{(h)}, L_{i,j'+2^{h-1}n}^{(h)}), \tag{16}$$

$$L_{2i-1,j'}^{(h-1)} = L_{i,j'}^{(h)} + (-1)^{\hat{u}_{2i,j'}^{(h-1)}} L_{i,j'+2^{h-1}n}^{(h)}, \tag{17}$$

where $j' = 0, 1, ..., 2^{h-1}n - 1$. The above LLR updates indicate that at level-$(h-1)$, component code $\mathrm{U}_{2i}^{(h-1)}$ is decoded prior to $\mathrm{U}_{2i-1}^{(h-1)}$, since the LLR updates of $\mathrm{U}_{2i-1}^{(h-1)}$ require the estimations of $\mathrm{U}_{2i,j}^{(h-1)}$. Once the LLRs of all component BCH codes are produced, the OSD will be used to decode the component codes.

Without loss of generality, the decoding of a component code $\mathrm{U}_i^{(h)} (h > 0)$ can be performed by the following key operations.

Code Decomposition: If $h > 0$, $\mathrm{U}_i^{(h)}$ is still a U-UV structured code. It should be decomposed into $\mathrm{U}_{2i}^{(h-1)}$ and $\mathrm{U}_{2i-1}^{(h-1)}$, with the LLR updates of (16) and (17), respectively.

Level-0 Component Code Decoding: The above decomposition continues until $h = 0$, where $\mathrm{U}_1^{(0)}, \mathrm{U}_2^{(0)}, \ldots, \mathrm{U}_{2^H}^{(0)}$ are component BCH codes. The OSD will be performed for each component code. Based on the correlation distance of (10), the most likely estimation will be chosen as the decoding output for the component code.

Code Reconstruction: Once component codes $\mathrm{U}_{2i-1}^{(0)}$ and $\mathrm{U}_{2i}^{(0)}$ have been decoded, $\mathrm{U}_i^{(1)}$ can be reconstructed. In general, once the estimations of level-$(h-1)$ component codes have been obtained, estimations of the level-h component code can be determined by

$$\hat{\boldsymbol{u}}_i^{(h)} = (\hat{\boldsymbol{u}}_{2i-1}^{(h-1)} | \hat{\boldsymbol{u}}_{2i-1}^{(h-1)} + \hat{\boldsymbol{u}}_{2i}^{(h-1)}). \tag{18}$$

In general, SC decoding of an H levels U-UV code is a recursive process starting by decomposing $\mathrm{U}_1^{(H)}$. Fig. 3 illustrates the SC decoding process of a 2 levels U-UV code. $\mathrm{U}_1^{(2)}$ is first decomposed into $\mathrm{U}_1^{(1)}$ and $\mathrm{U}_2^{(1)}$, where the latter will be further decomposed into $\mathrm{U}_3^{(0)}$ and $\mathrm{U}_4^{(0)}$. $\mathrm{U}_4^{(0)}$ will

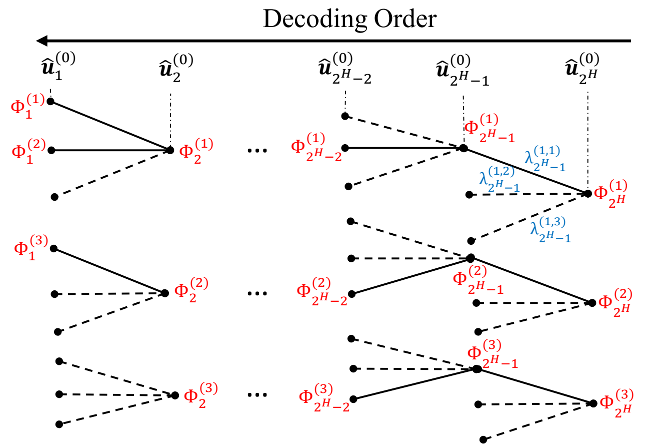

Fig. 4. SCL decoding tree with $l = 3$. The solid branches indicate the accumulated distances are being kept at the layers.

then be decoded by the OSD, following by the OSD of $\mathrm{U}_3^{(0)}$. Afterward, $\mathrm{U}_2^{(1)}$ can be reconstructed. The decoding of $\mathrm{U}_2^{(0)}$ and $\mathrm{U}_1^{(0)}$ follow similarly. Therefore, the component codes $\mathrm{U}_4^{(0)}, \mathrm{U}_3^{(0)}, \mathrm{U}_2^{(0)}, \mathrm{U}_1^{(0)}$ are successively decoded.

C. The SCL Decoding

Armed with the above knowledge, the SCL decoding of U-UV codes can be further introduced. The SCL decoding evolves from the above SC operations by considering the l most likely OSD candidates for each component code, resulting in a list of decoding estimations for the U-UV code.

The SCL decoding process can be illustrated by a multi-branch tree shown in Fig. 4. In the decoding tree, each layer corresponds to a component code and nodes of the layer denote the estimations of the component code. Consequently, the SCL decoding can be visualized as a layer-by-layer codeword estimation process over the tree, where the layer indices and the component code indices coincide. To prevent an exponentially growing decoding complexity, at each layer, only the l most likely estimations of the component code will be kept. At layer i, the l most likely estimations of component code $\mathrm{U}_i^{(0)}$ are denoted as $\hat{\boldsymbol{u}}_i^{(0)}(1), \hat{\boldsymbol{u}}_i^{(0)}(2), ..., \hat{\boldsymbol{u}}_i^{(0)}(l)$, respectively. Let $\hat{\boldsymbol{u}}_{i-1}^{(0)}(s')$ denote an estimation of $\mathrm{U}_{i-1}^{(0)}$ which is obtained based on one of the estimations of $\mathrm{U}_i^{(0)}$, i.e., $\hat{\boldsymbol{u}}_i^{(0)}(s)$. Note that $s, s' = 1, 2, ..., l$. Let $\boldsymbol{L}_{i-1}^{(0)}(s) = (L_{i-1,0}^{(0)}(s), L_{i-1,1}^{(0)}(s), ..., L_{i-1,n-1}^{(0)}(s))$ denote the corresponding LLR vector computed based on $\hat{\boldsymbol{u}}_i^{(0)}(s)$. The correlation distance between $\boldsymbol{L}_{i-1}^{(0)}(s)$ and $\hat{\boldsymbol{u}}_{i-1}^{(0)}(s')$ is

$$\lambda_{i-1}^{(s,s')} = \lambda(\boldsymbol{L}_{i-1}^{(0)}(s), \hat{\boldsymbol{u}}_{i-1}^{(0)}(s')), \tag{19}$$

where function $\lambda(\cdot)$ has been defined in (10). Over the decoding tree, $\lambda_i^{(s,s')}$ represents the metric of the s'th branch that emancipates from the sth node at layer i. Since $\mathrm{U}_{2^H}^{(0)}$ is the first component code to be decoded, its estimations are the roots of the decoding tree. Therefore, the correlation distances (branches) that lead to them are written as $\lambda_{2^H}^{(s,-)}$. These correlation distances are accumulated respectively along the decoding path. Let $\Phi_i^{(1)}, \Phi_i^{(2)}, ..., \Phi_i^{(l)}$ denote the l smallest accumulated distances at layer i of the tree. At layer 2^H, $\Phi_i^{(s)}$ are initialized as $\lambda_{2^H}^{(s,-)}$. At layer i where $i < 2^H$, let us define

$$\Lambda_{i-1}^{(s,s')} = \Phi_i^{(s)} + \lambda_{i-1}^{(s,s')} \tag{20}$$

as the accumulated correlation distance that is computed through a path from node s at layer i to node s' at layer $i-1$. Since $s, s' = 1, 2, ..., l$, l^2 accumulated correlation distances $\Lambda_{i-1}^{(s,s')}$ will be computed at layer-$(i-1)$. Only the l smallest ones will be kept and denoted as $\Phi_{i-1}^{(1)}, \Phi_{i-1}^{(2)}, ..., \Phi_{i-1}^{(l)}$, respectively. Decoding of the rest component codes follows until layer-1 is reached, indicating all component codes have been decoded. Note that to ensure the SCL decoding with a list size of l, the OSD order of code $\mathrm{U}_{2^H}^{(0)}$ should be chosen such that its $l_\tau \geq l$.

The above SCL decoding of an H levels U-UV code is further summarized as in **Algorithm 1**. The decoding will start by calling SCLD($\mathrm{U}_1^{(H)}$). From its output list, $\hat{\boldsymbol{u}}_1^{(H)}$ can be estimated. Note that after $\mathrm{U}_1^{(0)}$ is decoded, l smallest accumulated distances $\Phi_1^{(1)}, \Phi_1^{(2)}, ..., \Phi_1^{(l)}$ have been obtained. Let

$$\Phi_1^{(\star)} = \min\{\Phi_1^{(s)}, \forall s\} \tag{21}$$

denote the smallest value among them. The estimation $\hat{\boldsymbol{u}}_1^{(H)}$ can be retrieved from the path that yields $\Phi_1^{(\star)}$.

Algorithm 1 SCL Decoding of $\mathrm{U}_i^{(h)}$, SCLD($\mathrm{U}_i^{(h)}$)

Input: $\{L_{i,j}^{(h)} | j = 0, 1, ..., 2^h n - 1\}$, l;
Output: $\{\hat{\boldsymbol{u}}_i^{(h)}(s)\}$;

```
If h > 0
    Compute the LLRs of U_{2i}^{(h-1)} as in (16);
    Perform SCLD(U_{2i}^{(h-1)}), yielding {û_{2i}^{(h-1)}(s)};
    For each estimation û_{2i}^{(h-1)} do
        Compute the LLRs of U_{2i-1}^{(h-1)} as in (17);
        Perform SCLD(U_{2i-1}^{(h-1)}), yielding {û_{2i-1}^{(h-1)}(s)};
    For each pair of û_{2i-1}^{(h-1)} and û_{2i}^{(h-1)} do
        Reconstruct û_i^{(h)} as in (18);
    Form the estimation list {û_i^{(h)}(s)}.
Else
    Perform OSD, yielding {û_i^{(0)}};
    Determine λ_i^{(s,s')} and Λ_i^{(s,s')} as in (19) and (20);
    Keep the l smallest values Φ_i^{(1)}, Φ_i^{(2)}, ..., Φ_i^{(l)}, and yield
    the estimations {û_i^{(h)}(s)}.
```

IV. Performance and Complexity

This section shows the SCL decoding performance of the U-UV codes over the AWGN channel using BPSK modulation. They are compared with the SCL decoding of polar codes [13, 14]. The polar codes are designed using GA at the SNR of 0dB. The finite length performance limits represented by the normal approximation (NA) [15] are also shown for references.

Fig. 5 shows the frame error rate (FER) of the $(252, 183)$ U-UV code with the $(63, 57), (63, 51), (63, 51)$ and $(63, 24)$ BCH component codes. This 2 levels U-UV code is designed based on the subchannel capacity estimations at the SNR of 1dB, which has been shown in Fig. 2. For OSD of a BCH code with its minimum Hamming distance d, a decoding order of $\tau = \lfloor \frac{d}{4} \rfloor$ is sufficient for the OSD to yield an ML performance [12]. In our simulations, the OSD orders are chosen to meet this requirement so that the ML decoding performance can be realized for each component code. In particular, they are $1, 1, 1$ and 2 for $U_1^{(0)}, U_2^{(0)}, U_3^{(0)}$ and $U_4^{(0)}$, respectively. The performances of the SC and the SCL decoding with list size l (denoted as SCL (l)) are provided. It can be seen that as the list size enlarges, the SCL decoding achieves a larger coding gain over the SC decoding. With the same list size l, SCL decoding of the U-UV code outperforms that of the $(256, 185)$ polar code, where the latter is assisted by an 8-bit CRC (denoted as CA-SCL (l)). It should be pointed that the U-UV codes do not need to be assisted by CRC. This is due to its component code decoding have already provided the ML decoding feature for the structured code. The ML estimations are more likely to be validated by CRC.

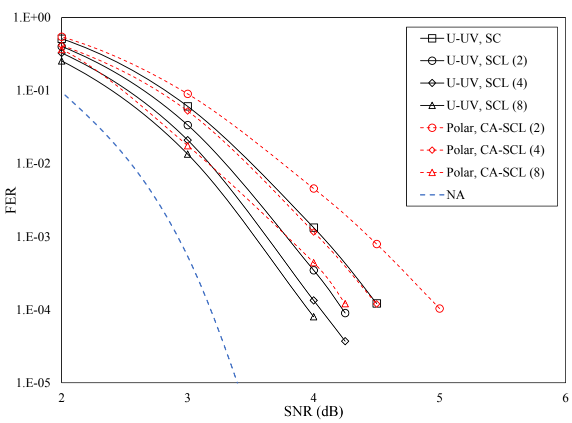

Fig. 5. SCL decoding performance of the $(252, 183)$ U-UV code.

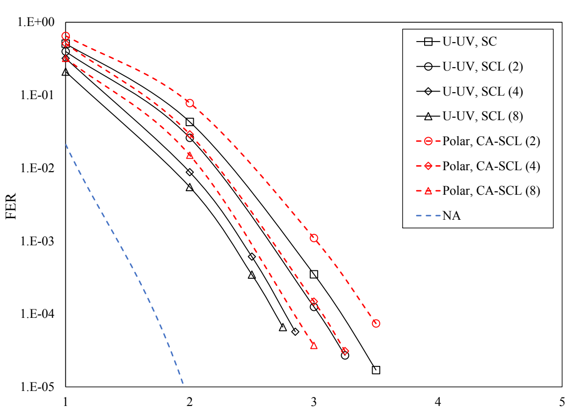

Fig. 6. SCL decoding performance of the $(504, 250)$ U-UV code.

Fig. 6 shows the FER performance of the $(504, 250)$ U-UV code. It is a 3 levels U-UV code with the $(63, 57), (63, 51), (63, 45), (63, 24), (63, 45), (63, 18), (63, 10), (63, 0)$ BCH component codes. Note that the $(63, 0)$ component code does not carry any information. It is acting as the frozen symbols of polar codes. The OSD orders of the other seven component codes are $1, 1, 1, 2, 1, 3$ and 3, respectively. Similarly, Fig. 6 shows that SCL decoding of the U-UV code can also outperform that of the $(512, 254)$ polar code assisted by a CRC of 8 bits.

It should be noted that, for the U-UV codes, the OSD dominates the overall decoding complexity. The complexity of OSD with order τ can be characterized as $O(n^\tau)$. Hence, for an H levels U-UV codes, the SCL(l) complexity is $O(n^\tau 2^H l)$. The decoding complexity will grow rapidly as the order increases. For the chosen BCH codes with $n = 63$, the most of the OSD orders would not be greater than two. In this case, the SCL(l) complexity will be $O(n^2 2^H l)$. The SCL decoding complexity of a polar code with a similar length of $n2^H$ is $O(n2^H \log_2(n2^H)l)$. Hence, that of the U-UV code will be more complex by a factor of $\frac{n}{\log_2(n2^H)}$. But this factor could be further decreased since the OSD complexity could be reduced by several existing methods from published work.

V. Conclusion

This paper has proposed the U-UV structured codes with BCH codes as their components, as well as the SCL decoding of the structured codes. Rate of each component code can be designed based on the estimated subchannel capacities. The SCL decoding has been proposed by integrating the OSD of the component codes and the SC decoding mechanism. Our simulation results have shown that over the short-to-medium length regime, SCL decoding of the U-UV code can outperform that of a similar rate polar code.

Acknowledgement

This work is sponsored by the National Natural Science Foundation of China (NSFC) with project ID 62071498.

References

[1] M. Plotkin, "Binary codes with specified minimum distance," *IRE Trans. Inf. Theory*, vol. 6, no. 4, pp. 445–450, Sept. 1960.

[2] W. van Gils, "Linear unequal error protection codes from shorter codes," *IEEE Trans. Inf. Theory*, vol. 30, no. 3, pp. 544–546, May 1984.

[3] V. Kumar and O. Milenkovic, "On unequal error protection LDPC codes based on Plotkin-type constructions," *IEEE Trans. Commun.*, vol. 54, no. 6, pp. 994–1005, Jun. 2006.

[4] J. S. Park, K. Park, and H. Song, "Rate allocation for component codes of Plotkin-type UEP codes," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Cambridge, MA, USA, Jul. 2012, pp. 1578–1582.

[5] I. Reed, "A class of multiple-error-correcting codes and the decoding scheme," *IRE Trans. Inf. Theory*, vol. 4, no. 4, pp. 38–49, 1954.

[6] D. E. Muller, "Application of Boolean algebra to switching circuit design and to error detection," *IRE Trans. Electron. Comput.*, vol. EC-3, no. 3, pp. 6–12, Sept. 1954.

[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] P. Trifonov and P. Semenov, "Generalized concatenated codes based on polar codes," in *Proc. IEEE Int. Symp. Wirel. Commun. Syst. (ISWCS)*, Aachen, Germany, Nov. 2011, pp. 442–446.

[9] K. R. N. Y. Wang, "Concatenations of polar codes with outer bch codes and convolutional codes," in *2014 52nd Annual Allerton Conference on Communication, Control, and Computing (ALLERTON)*, 2014, pp. 813–819.

[10] I. Mrquez-Corbella and J. Tillich, "Attaining capacity with iterated $(U|U+V)$ codes based on AG codes and Koetter-Vardy soft decoding," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, Jun. 2017, pp. 6–10.
[11] P. Trifonov, "Efficient design and decoding of polar codes," *IEEE Trans. Commum.*, vol. 60, no. 11, pp. 3221–3227, Nov. 2012.
[12] M. Fossorier and S. Lin, "Soft-decision decoding of linear block codes based on ordered statistics," *IEEE Trans. Inf. Theory*, vol. 41, no. 5, pp. 1379–1396, Sept. 1995.
[13] I. Tal and A. Vardy, "List decoding of polar codes," *IEEE Trans. Inf. Theory.*, vol. 61, no. 5, pp. 2213–2226, May 2015.
[14] K. Niu and K. Chen, "CRC-aided decoding of polar codes," *IEEE Commun. Lett.*, vol. 16, no. 10, pp. 1668–1671, Oct. 2012.
[15] Y. Polyanskiy, V. Poor, and S. Verdù, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.