Order Skipping Ordered Statistics Decoding and Its Performance Analysis

Xihao Li[†], Wenhao Chen[†], Li Chen[†‡], Yuan Li[§] and Huazi Zhang[§]
[†]School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou, China
[‡]Guangdong Province Key Laboratory of Information Security Technology, Guangzhou, China
[§]Hangzhou Research Center, Huawei Technologies Co. Ltd., Hangzhou, China
Email:{lixh98, chenwh85}@mail2.sysu.edu.cn, chenli55@mail.sysu.edu.cn, {liyuan299, zhanghuazi}@huawei.com

***Abstract*—This paper proposes a reduced complexity ordered statistics decoding (OSD) algorithm for linear block codes, the namely order skipping (OS)-OSD algorithm. An approximated correlation distance lower bound (CDLB) is derived by utilizing likelihood of the received symbols over the least reliable positions (LRPs). It enables the assessment of whether the higher-order decoding can yield a more likely codeword estimation. If not, they can be skipped. Error-correction performance of the OS-OSD is analyzed. In particular, the decoding error probability of OS-OSD with order one is theoretically characterized. Our simulation results verify that the OS-OSD can achieve a significant complexity reduction over the state-of-the-art OSD without compromising the decoding performance.**

***Index Terms*—Linear block codes, maximum-likelihood decoding, ordered statistics decoding, reduced complexity**

I. Introduction

The realization of ultra-reliable low-latency communication (URLLC) requires the support of competent short-to-medium length channel codes. The transmission limit of a finite length coded system has been characterized in [1]. Recent researches in short-to-medium length codes have shown that the ordered statistics decoding (OSD) [2] [3] of BCH codes can yield a performance that is close to the finite length transmission limit [4]. In the OSD, the most reliable independent positions (MRIPs) of a received word are first identified. A number of test error patterns (TEPs) are then superimposed onto the hard-decision of the MRIPs, forming a list of test messages. They are re-encoded by a systematic generator matrix (SGM) in which columns of the MRIPs form an identity submatrix, yielding a list of codeword candidates. Note that the identification of the MRIPs and the generation of the SGM are realized by Gaussian elimination (GE). The correlation distance between the codeword candidate and the received symbol sequence is utilized to assess the likelihood of the codeword candidate. The candidate that yields the smallest correlation distance will be selected as the decoding output. However, it should be noted that the number of TEPs increases exponentially with the decoding order, leading to an exponential complexity of the OSD. Therefore, despite its competency in decoding BCH codes, the OSD complexity remains challenging for practice. In order to reduce the complexity, several operational skipping rules and stopping rules have been proposed [5]–[8]. The skipping rules facilitate the decoding by eliminating some TEPs, while the stopping rules identify the optimal codeword candidate in the decoding output list, hence the decoding can be terminated earlier. Besides, a segmentation-discarding rule has been proposed in [9], dividing the MRIPs into several segments for reducing the complexity. The OSD variants utilizing the constraint of the parity-check matrix to reduce the number of TEPs have been proposed in [10]–[11]. There also exist several approaches to reduce the decoding order by utilizing information outside the MRIPs [12]–[14]. Meanwhile, OSD complexity reduction can also be achieved through reducing the GE complexity [15]–[16]. Recent research of [17] has shown that the BCH codeword candidates can be obtained by the SGM of a Reed-Solomon code and GE is no longer needed.

In order to eliminate the redundant decoding effort, an order skipping rule has been proposed in the so-called fast and scalable OSD (FOSD) in [18]. It estimates the correlation distance lower bound (CDLB) of the codeword candidates that are generated in the higher-order decoding. Consequently, the higher-order decoding can be skipped if they are unlikely to yield a more likely codeword. However, this CDLB is empirical. Its inevitable looseness results in limited complexity reduction. In this work, a more general and accurate CDLB is characterized. It utilizes likelihood of the received symbols over the least reliable positions (LRPs), adapting the CDLB to the channel. As a result, it can better identify the redundant high order decoding attempts, leading to a more significant complexity reduction. The decoding error probability upper bound of this order skipping (OS)-OSD is analyzed. In particular, the decoding error probability of the OS-OSD with a decoding order of one is theoretically characterized. Our simulation results verify the above analysis and demonstrate the complexity merit of the proposed OS-OSD. They also show that the affiliated performance loss over the prototype OSD is negligible. In return, this demonstrates the accuracy of our proposed CDLB.

II. Ordered Statistics Decoding

Let $\mathbb{F}_2 = \{0, 1\}$ denote the binary field. Let $\mathcal{C}(n, k, d)$ denote a binary linear block code of length n, dimension k and minimum Hamming distance d. Let $\mathbf{G} \in \mathbb{F}_2^{k\times n}$ denote the generator matrix of $\mathcal{C}(n, k, d)$. We use $[\beta]_a^b$ to denote a β-sequence $(\beta_a, \beta_{a+1}, \cdots, \beta_b)$. Assume a codeword $\boldsymbol{c} = [c]_1^n \in$

2024 IEEE Information Theory Workshop (ITW) | 979-8-3503-4893-4/24/$31.00 ©2024 IEEE | DOI: 10.1109/ITW61385.2024.10807042

$\mathbb{F}_2^n$ is modulated with binary phase shift keying (BPSK) with the mapping of $0 \mapsto 1$, $1 \mapsto -1$. The modulated symbol sequence is transmitted through an additive white Gaussian noise (AWGN) channel with a noise variance of $N_0/2$. Let $\boldsymbol{r} = [r]_1^n \in \mathbb{R}^n$ denote the received symbol sequence. That says $r_j = (-1)^{c_j} + w_j$ for $j = 1, 2, \cdots, n$, where w_j is the AWGN with zero mean and variance $N_0/2$. Let $P(r_j|c_j)$ denote the channel observations. The log-likelihood ratio (LLR) of r_j is defined as $L_j = \ln \frac{P(r_j|c_j=0)}{P(r_j|c_j=1)}$, which can be further derived as $L_j = \frac{4r_j}{N_0}$. Accordingly, let $\boldsymbol{y} = [y]_1^n \in \mathbb{F}_2^n$ denote the hard-decision received word, where $y_j = 0$ if $L_j > 0$, or $y_j = 1$ otherwise. Let $\boldsymbol{\alpha} = [\alpha]_1^n$ denote the reliability sequence of $\boldsymbol{r}$, where $\alpha_j = |r_j|$. A greater α_j indicates the decision made based on r_j is more reliable.

In the OSD, the received symbols are first sorted in a descending order of their reliabilities, yielding the sorted symbol sequence $\tilde{\boldsymbol{r}} = [\tilde{r}]_1^n = \Pi(\boldsymbol{r})$, where Π denotes the permutation function. Accordingly, columns of $\mathbf{G}$ are permuted, yielding $\tilde{\mathbf{G}} = \Pi(\mathbf{G})$. The GE is performed on $\tilde{\mathbf{G}}$, yielding the SGM $\tilde{\mathbf{G}}_{\rm s} = [\mathbf{I}_k \ \tilde{\mathbf{P}}]$, where $\mathbf{I}_k$ is a k-dimensional identity submatrix and $\tilde{\mathbf{P}}$ is the parity submatrix of size $k \times (n-k)$. Note that if the first k columns of $\tilde{\mathbf{G}}$ are not linearly independent, an additional column permutation is required to obtain $\tilde{\mathbf{G}}_{\rm s}$. For simplicity, we assume that the first k columns of $\tilde{\mathbf{G}}$ are ensured with the linear independence. Therefore, the first k positions carried by $\tilde{\boldsymbol{r}}$ are the MRIPs. We refer to the remaining $n-k$ positions as the LRPs. Let $\tilde{\boldsymbol{c}} = [\tilde{c}]_1^n = \Pi(\boldsymbol{c})$, $\tilde{\boldsymbol{y}} = [\tilde{y}]_1^n = \Pi(\boldsymbol{y})$ and $\tilde{\boldsymbol{\alpha}} = [\tilde{\alpha}]_1^n = \Pi(\boldsymbol{\alpha})$ denote the permuted version of $\boldsymbol{c}$, $\boldsymbol{y}$ and $\boldsymbol{\alpha}$, respectively. Given a β-sequence $[\beta]_1^n$, let $\boldsymbol{\beta}_{\rm B} = [\beta]_1^k$ and $\boldsymbol{\beta}_{\rm P} = [\beta]_{k+1}^n$ denote the MRIP segment and the LRP segment, respectively.

The OSD with a decoding order τ is denoted as OSD(τ). It generates codeword candidates by re-encoding test messages. Let $\boldsymbol{e} \in \mathbb{F}_2^k$ denote a TEP. The OSD proceeds by generating the TEPs with increasing weights. They are superimposed onto $\tilde{\boldsymbol{y}}_{\rm B}$, yielding the test messages for re-encoding. Specifically, in phase-i, all the weight-i TEPs are generated, where $i = 0, 1, ..., \tau$. With a TEP $\boldsymbol{e} \in \mathbb{F}_2^k$, the test message is generated as $\boldsymbol{e} \oplus \tilde{\boldsymbol{y}}_{\rm B}$. The corresponding (permuted) codeword candidate $\tilde{\boldsymbol{c}}_e = [\tilde{c}_e]_1^n$ is generated by

$$\tilde{\boldsymbol{c}}_e = (\boldsymbol{e} \oplus \tilde{\boldsymbol{y}}_{\rm B})\tilde{\mathbf{G}}_{\rm s}. \tag{1}$$

The correlation distance between $\tilde{\boldsymbol{c}}_e$ and $\tilde{\boldsymbol{r}}$ is defined as

$$\mathcal{D}(\tilde{\boldsymbol{c}}_e, \tilde{\boldsymbol{r}}) = \sum_{i=1}^{n} (\tilde{c}_{e,i} \oplus \tilde{y}_i)\tilde{\alpha}_i. \tag{2}$$

A smaller correlation distance indicates $\tilde{\boldsymbol{c}}_e$ is more likely to be the (permuted) transmitted codeword $\tilde{\boldsymbol{c}}$, and vice versa. In OSD(τ), the total number of codeword candidates is

$$\Omega_\tau = \sum_{i=0}^{\tau} \binom{k}{i}. \tag{3}$$

Among these candidates, the one with the minimum correlation distance to $\tilde{\boldsymbol{r}}$ is denoted as $\tilde{\boldsymbol{c}}_{\rm opt}$. The decoding produces the most likely codeword estimation as $\Pi^{-1}(\tilde{\boldsymbol{c}}_{\rm opt})$, where Π^{-1} is the inverse function of Π. For a code with rate $k/n \geq 1/2$, an order of $\lceil d/4 - 1 \rceil$ is sufficient for the OSD to yield a near-ML decoding performance [3].

III. Order Skipping OSD

This section introduces the OS-OSD. The conditional expectation of the partial correlation distance that is associated with the LRPs of $\tilde{\boldsymbol{r}}$ is first derived. Based on this, the order skipping condition is introduced, formulating the OS-OSD.

A. Expectation of the Partial Correlation Distance

Without loss of generality, we assume that the all-zero codeword $\mathbf{0}$ is transmitted. Thereby, $r_j = 1 + w_j$. The probability density function (PDF) of r_j is

$$f_r(x) = \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{(x-1)^2}{N_0}\right). \tag{4}$$

Let $\tilde{\boldsymbol{e}} = [\tilde{e}]_1^n \in \mathbb{F}_2^n$ denote the error vector introduced by the channel, i.e., $\tilde{\boldsymbol{y}} = \tilde{\boldsymbol{c}} \oplus \tilde{\boldsymbol{e}}$. Based on (4), the conditional error probability of $\tilde{y}_j$ can be estimated as

$$\Pr(\tilde{e}_j = 1|\tilde{\alpha}_j) = \frac{f_r(-\tilde{\alpha}_j)}{f_r(\tilde{\alpha}_j) + f_r(-\tilde{\alpha}_j)} = \frac{1}{1 + \exp(\frac{4\tilde{\alpha}_j}{N_0})}. \tag{5}$$

The conditional expectation of the partial correlation distance associated with the LRPs can be derived as

$$\begin{aligned} \mathbb{E}[\mathcal{D}(\tilde{\boldsymbol{c}}_{\rm P}, \tilde{\boldsymbol{r}}_{\rm P})|\tilde{\boldsymbol{\alpha}}] &= \mathbb{E}\left[\sum_{j=k+1}^{n} (\tilde{c}_j \oplus \tilde{y}_j)\cdot\tilde{\alpha}_j \,\middle|\, \tilde{\boldsymbol{\alpha}}\right] \\ &= \sum_{j=k+1}^{n} \mathbb{E}[(\tilde{c}_j \oplus \tilde{y}_j)|\tilde{\alpha}_j]\cdot\tilde{\alpha}_j, \end{aligned} \tag{6}$$

where the second equality is realized based on that $\tilde{\alpha}_j$ and $\tilde{\boldsymbol{\alpha}}$ provide the same information in determining the distribution of $\tilde{y}_j$. Since $\tilde{e}_j = \tilde{c}_j \oplus \tilde{y}_j$, it can be further derived that

$$\begin{aligned} \mathbb{E}[(\tilde{c}_j \oplus \tilde{y}_j)|\tilde{\alpha}_j] &= \mathbb{E}[\tilde{e}_j|\tilde{\alpha}_j] \\ &= \Pr(\tilde{e}_j = 1|\tilde{\alpha}_j)\cdot 1 + \Pr(\tilde{e}_j = 0|\tilde{\alpha}_j)\cdot 0 \\ &= \Pr(\tilde{e}_j = 1|\tilde{\alpha}_j). \end{aligned} \tag{7}$$

Based on (5) and (7), we have

$$\mathbb{E}[\mathcal{D}(\tilde{\boldsymbol{c}}_{\rm P}, \tilde{\boldsymbol{r}}_{\rm P})|\tilde{\boldsymbol{\alpha}}] = \sum_{j=k+1}^{n} \frac{\tilde{\alpha}_j}{1 + \exp(\frac{4\tilde{\alpha}_j}{N_0})}. \tag{8}$$

B. Order Skipping Condition

In OSD, the phase-i re-encoding generates $\binom{k}{i}$ codeword candidates. This implies the higher re-encoding phases dominate the overall complexity. To alleviate this complexity challenge, an order skipping condition is proposed for the OSD. It is developed based on the approximated CDLB of the codeword candidates that are generated in the higher re-encoding phases.

Algorithm 1: OS-OSD (τ)

Input: $\mathbf{G}, \tau, \boldsymbol{r}$
Output: $\hat{\boldsymbol{c}}$

Sort $\boldsymbol{r}$ and generate $\tilde{\boldsymbol{r}} = \Pi(\boldsymbol{r})$
Generate $\tilde{\mathbf{G}}_{\mathrm{s}}$ by performing GE on $\Pi(\mathbf{G})$
Generate $\tilde{\boldsymbol{y}}$ by making hard decisions on $\tilde{\boldsymbol{r}}$
For $i = 0, 1, \cdots, \tau$ **do**
  For *every TEP* $\boldsymbol{e} \in \mathbb{F}_2^k$ *of weight* i **do**
    Generate $\tilde{\boldsymbol{c}}_e$ as in (1)
    Compute $\mathcal{D}(\tilde{\boldsymbol{c}}_e, \tilde{\boldsymbol{r}})$ as in (2)
    If $i = 0$ ***or*** $\mathcal{D}(\tilde{\boldsymbol{c}}_e, \tilde{\boldsymbol{r}}) < \mathcal{D}(\tilde{\boldsymbol{c}}_{\mathrm{b}}, \tilde{\boldsymbol{r}})$ **then**
      Let $\tilde{\boldsymbol{c}}_{\mathrm{b}} = \tilde{\boldsymbol{c}}_e$
  Compute $\mathcal{D}_{\mathrm{OSD}}^{(i+1)}$ as in (11)
  If $\mathcal{D}(\tilde{\boldsymbol{c}}_{\mathrm{b}}, \tilde{\boldsymbol{r}}) < \mathcal{D}_{\mathrm{OSD}}^{(i+1)}$ **then**
    Break
Produce $\hat{\boldsymbol{c}} = \Pi^{-1}(\tilde{\boldsymbol{c}}_{\mathrm{b}})$

For a codeword candidate $\tilde{\boldsymbol{c}}_e$, its correlation distance to $\tilde{\boldsymbol{r}}$ can be decomposed as

$$\mathcal{D}(\tilde{\boldsymbol{c}}_e, \tilde{\boldsymbol{r}}) = \mathcal{D}(\tilde{\boldsymbol{c}}_{e,\mathrm{B}}, \tilde{\boldsymbol{r}}_{\mathrm{B}}) + \mathcal{D}(\tilde{\boldsymbol{c}}_{e,\mathrm{P}}, \tilde{\boldsymbol{r}}_{\mathrm{P}}). \tag{9}$$

In the OSD, the TEPs are generated with an increasing weight. For $\tilde{\boldsymbol{c}}_e$ generated in phase-i, $\mathcal{D}(\tilde{\boldsymbol{c}}_{e,\mathrm{B}}, \tilde{\boldsymbol{r}}_{\mathrm{B}})$ is lower bounded by only counting the i smallest reliabilities of the MRIPs, i.e.,

$$\mathcal{D}(\tilde{\boldsymbol{c}}_{e,\mathrm{B}}, \tilde{\boldsymbol{r}}_{\mathrm{B}}) \geq \sum_{j=k-i+1}^{k} \tilde{\alpha}_j. \tag{10}$$

Further approximating $\mathcal{D}(\tilde{\boldsymbol{c}}_{e,\mathrm{P}}, \tilde{\boldsymbol{r}}_{\mathrm{P}})$ as in (8), the approximated CDLB of the codeword candidates generated in phase-i is

$$\mathcal{D}_{\mathrm{OSD}}^{(i)} = \sum_{j=k-i+1}^{k} \tilde{\alpha}_j + \mathbb{E}[\mathcal{D}(\tilde{\boldsymbol{c}}_{\mathrm{P}}, \tilde{\boldsymbol{r}}_{\mathrm{P}}) | \tilde{\boldsymbol{\alpha}}]. \tag{11}$$

Hence, for $\tilde{\boldsymbol{c}}_e$ generated in phase-i, if $\tilde{\boldsymbol{c}}_e \neq \tilde{\boldsymbol{c}}$, it has a high probability to exhibit $\mathcal{D}(\tilde{\boldsymbol{c}}_e, \tilde{\boldsymbol{r}}) > \mathcal{D}_{\mathrm{OSD}}^{(i)}$. In other words, if $\mathcal{D}(\tilde{\boldsymbol{c}}_e, \tilde{\boldsymbol{r}}) < \mathcal{D}_{\mathrm{OSD}}^{(i)}$, it is highly probable that $\tilde{\boldsymbol{c}}_e = \tilde{\boldsymbol{c}}$. Note that for $i' > i$, $\mathcal{D}_{\mathrm{OSD}}^{(i')} > \mathcal{D}_{\mathrm{OSD}}^{(i)}$. Armed with this, the following order skipping condition can be introduced.

At the end of phase-i, let $\tilde{\boldsymbol{c}}_{\mathrm{b}}$ denote the codeword candidate that has the minimum correlation distance to $\tilde{\boldsymbol{r}}$. If

$$\mathcal{D}(\tilde{\boldsymbol{c}}_{\mathrm{b}}, \tilde{\boldsymbol{r}}) < \mathcal{D}_{\mathrm{OSD}}^{(i+1)}, \tag{12}$$

it implies the following decoding phases may not be able to generate a more likely codeword candidate. They can be skipped. Hence, the OSD terminates and produces the estimated codeword as $\hat{\boldsymbol{c}} = \Pi^{-1}(\tilde{\boldsymbol{c}}_{\mathrm{b}})$. Otherwise, the OSD proceeds into phase-$(i+1)$. The order-τ OSD armed with the above order skipping rule is denoted as OS-OSD (τ) and summarized in *Algorithm 1*.

If the approximated CDLB is underestimated, the probability of skipping the higher-order decoding will be small. It results in a limited complexity reduction. Conversely, if it is overestimated, a decoding performance loss will be led to. Compared with the CDLB of the FOSD [18] whose LRP segment is estimated with an empirical factor β, the proposed CDLB of (11) is a more general and accurate approximation. It is obtained based on likelihood of received symbols, which can be applied to the most of the OSD variants. Our following theoretical analysis and numerical results will verify its accuracy.

IV. Performance Analysis of OS-OSD

A. Ordered Statistics

Based on (4) and $\alpha_j = |r_j|$, the PDF of α_j is

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

The cumulative distribution function (CDF) of α_j can be further derived as

$$F_\alpha(x) = \begin{cases} 0, & x < 0; \\ 1 - Q\left(\frac{2x+2}{\sqrt{2N_0}}\right) - Q\left(\frac{2x-2}{\sqrt{2N_0}}\right), & x \geq 0, \end{cases} \tag{13}$$

where $Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x$ is the standard normal tail function. Furthermore, the PDF of $\tilde{\alpha}_j$ is given by [19]

$$f_{\tilde{\alpha}_j}(x) = \frac{n!}{(j-1)!(n-j)!} (1 - F_\alpha(x))^{j-1} F_\alpha(x)^{n-j} f_\alpha(x). \tag{14}$$

Let $q = 1, 2, \cdots, n$ and $q < j$, the joint PDF of $\tilde{\alpha}_q$ and $\tilde{\alpha}_j$ can be derived as [19]

$$\begin{aligned} f_{\tilde{\alpha}_q, \tilde{\alpha}_j}(x, y) = & \frac{n!}{(q-1)!(j-q-1)!(n-j)!} \\ & \cdot (1 - F_\alpha(x))^{q-1} \left(F_\alpha(x) - F_\alpha(y)\right)^{j-q-1} \\ & \cdot F_\alpha(y)^{n-j} f_\alpha(x) f_\alpha(y) \mathbf{1}_{[0,x]}(y). \end{aligned} \tag{15}$$

where $\mathbf{1}_{[a,b]}(y) = 1$ if $a \leq y \leq b$, and $\mathbf{1}_{[a,b]}(y) = 0$ otherwise. For a received symbol whose reliability is greater than x, the probability of its hard decision being erroneous is derived as

$$P_\alpha(x) = \frac{Q\left(\frac{2x+2}{\sqrt{2N_0}}\right)}{Q\left(\frac{2x+2}{\sqrt{2N_0}}\right) + Q\left(\frac{2x-2}{\sqrt{2N_0}}\right)}. \tag{16}$$

Given $\tilde{\boldsymbol{e}} = [\tilde{\boldsymbol{e}}_{\mathrm{B}} \quad \tilde{\boldsymbol{e}}_{\mathrm{P}}]$, let w denote the weight of $\tilde{\boldsymbol{e}}_{\mathrm{B}}$. Consequently, the probability that there are s errors in the MRIPs is given by [6]

$$\Pr(w = s) = \int_0^\infty \binom{k}{s} P_\alpha(x)^s (1 - P_\alpha(x))^{k-s} f_{\tilde{\alpha}_{k+1}}(x) \mathrm{d}x. \tag{17}$$

B. Decoding Error Probability of OS-OSD

The decoding error probability of OS-OSD (τ), denoted as P_{e}, is upper bounded by

$$P_{\mathrm{e}} \leq P_{\mathrm{ML}} + P_{\mathrm{list,OS\text{-}OSD}}(\tau), \tag{18}$$

where P_{ML} is the ML decoding error probability and $P_{\mathrm{list,OS\text{-}OSD}}(\tau)$ is the probability that the intended codeword $\tilde{\boldsymbol{c}}$ is not included in the decoding output list of the OS-OSD (τ). It is also called the list error probability. Note that

the upper bound for P_{ML} can be obtained by the existing bounding techniques, e.g., the tangential bound [20]. They usually require knowledge of the code's weight spectrum.

In OS-OSD (τ), there are two cases that will lead to a list error event. The first case is when there are more than τ errors in the MRIPs. None of the Ω_τ re-encoding attempts can yield the intended codeword $\tilde{\boldsymbol{c}}$. The second case is when there are s errors in the MRIPs, but the order skipping condition is satisfied before reaching phase-s. Let S_i denote the event that the decoding terminates at the end of phase-i, where $i < \tau$. Therefore, the list error probability can be characterized as

$$P_{\text{list,OS-OSD}}(\tau) = \Pr(w>\tau) + \sum_{i=0}^{\tau-1} \sum_{s=i+1}^{\tau} \Pr(S_i, w=s). \quad (19)$$

Based on (17), the first term of (19) can be computed as

$$\Pr(w > \tau) = 1 - \sum_{i=0}^{\tau} \Pr(w = i). \quad (20)$$

For the second term of (19), it characterizes the performance loss over the OSD, which is incurred by skipping the higher-order decoding. It can be aware that this depends on the accuracy of the proposed CDLB. Please note that it remains challenging to characterize $\Pr(S_i, w)$ for $i > 0$. In the following, we focus on the case of $\tau = 1$, for which the theoretical characterization retreats to $\Pr(S_0, w = 1)$. It provides a valid theoretical assessment for OS-OSD (1).

C. Decoding Error Probability of OS-OSD (1)

For $\tau = 1$, the list error probability of (19) is simplified as

$$P_{\text{list,OS-OSD}}(1) = \Pr(w > 1) + \Pr(S_0, w = 1). \quad (21)$$

We further analyze $\Pr(S_0, w = 1)$. In phase-0, the codeword candidate is generated as $\tilde{\boldsymbol{c}}_0 = \tilde{\boldsymbol{y}}_{\mathrm{B}} \tilde{\mathbf{G}}_{\mathrm{s}} = [\tilde{\boldsymbol{y}}_{\mathrm{B}} \ \tilde{\boldsymbol{y}}_{\mathrm{B}} \tilde{\mathbf{P}}]$. Therefore, $\tilde{\boldsymbol{c}}_0 \oplus \tilde{\boldsymbol{y}} = [\mathbf{0} \ \tilde{\boldsymbol{y}}_{\mathrm{B}} \tilde{\mathbf{P}} \oplus \tilde{\boldsymbol{y}}_{\mathrm{P}}]$. Since $\tilde{\boldsymbol{y}} = \tilde{\boldsymbol{c}} \oplus \tilde{\boldsymbol{e}}$ and $\tilde{\boldsymbol{c}}_{\mathrm{B}} \tilde{\mathbf{P}} = \tilde{\boldsymbol{c}}_{\mathrm{P}}$, we further have $\tilde{\boldsymbol{c}}_0 \oplus \tilde{\boldsymbol{y}} = [\mathbf{0} \ \tilde{\boldsymbol{e}}_{\mathrm{B}} \tilde{\mathbf{P}} \oplus \tilde{\boldsymbol{e}}_{\mathrm{P}}]$. Let a set $\mathcal{Z}$ collect the non-zero positions of $\tilde{\boldsymbol{c}}_0 \oplus \tilde{\boldsymbol{y}}$, i.e.,

$$\mathcal{Z} = \operatorname{supp}([\mathbf{0} \quad \tilde{\boldsymbol{e}}_{\mathrm{B}} \tilde{\mathbf{P}} \oplus \tilde{\boldsymbol{e}}_{\mathrm{P}}]). \quad (22)$$

Hence, the correlation distance between $\tilde{\boldsymbol{c}}_0$ and $\tilde{\boldsymbol{r}}$ can be written as $\mathcal{D}(\tilde{\boldsymbol{c}}_0, \tilde{\boldsymbol{r}}) = \sum_{j \in \mathcal{Z}} \tilde{\alpha}_j$. Given $\mathcal{Z}$, the event S_0 happens if $\mathcal{D}(\tilde{\boldsymbol{c}}_0, \tilde{\boldsymbol{r}}) < \mathcal{D}^{(1)}_{\mathrm{OSD}}$. Based on (11), we have

$$\Pr(S_0|\mathcal{Z}) = \Pr\left(\sum_{j \in \mathcal{Z}} \tilde{\alpha}_j < \tilde{\alpha}_k + \sum_{j=k+1}^{n} \frac{\tilde{\alpha}_j}{1+\exp(\frac{4\tilde{\alpha}_j}{N_0})} \right). \quad (23)$$

Note that $\Pr(S_0) = \sum_{\mathcal{Z}} \Pr(S_0|\mathcal{Z}) \Pr(\mathcal{Z})$. Therefore, $\Pr(S_0, w = 1)$ of (21) can be written as

$$\begin{aligned} \Pr(S_0, w=1) &= \sum_{\mathcal{Z}} \Pr(S_0, w = 1, \mathcal{Z}) \\ &= \sum_{\mathcal{Z}} \Pr(S_0|\mathcal{Z}, w = 1) \Pr(\mathcal{Z}, w = 1) \\ &\approx \sum_{\mathcal{Z}} \Pr(S_0|\mathcal{Z}) \Pr(\mathcal{Z}|w=1) \Pr(w=1), \end{aligned} \quad (24)$$

in which we approximate $\Pr(S_0|\mathcal{Z}, w = 1)$ as $\Pr(S_0|\mathcal{Z})$ by ignoring the effect of $w = 1$.

For BCH codes, their weight distributions can be tightly approximated by the binomial distributions. Therefore, given $w > 0$, $\tilde{\boldsymbol{e}}_{\mathrm{B}} \tilde{\mathbf{P}}$ has equal probability to be any binary vector of length $n - k$. Subsequently, $\tilde{\boldsymbol{e}}_{\mathrm{B}} \tilde{\mathbf{P}} \oplus \tilde{\boldsymbol{e}}_{\mathrm{P}}$ is also equally likely to be any binary vector of length $n - k$. We have

$$\Pr(\mathcal{Z}|w > 0) \approx \Pr(\mathcal{Z}|w = 1) \approx \frac{1}{2^{n-k}}. \quad (25)$$

Therefore, (24) can be approximated as

$$\Pr(S_0, w=1) \approx \frac{1}{2^{n-k}} \Pr(w=1) \sum_{\mathcal{Z}} \Pr(S_0|\mathcal{Z}), \quad (26)$$

where $\Pr(w = 1)$ can be determined by (17). For the codes whose weight spectrum cannot be approximated by the binomial distribution, the probability of $\Pr(\mathcal{Z}|w = 1)$ can be computed by using their weight spectrum [6].

We further analyze the computation of $\Pr(S_0|\mathcal{Z})$. Let $\mathcal{Z}^{\mathrm{c}} = \{k+1, k+2, \cdots, n\} \backslash \mathcal{Z}$ and $g(x) = \frac{x}{1+\exp(\frac{4x}{N_0})}$. By defining

$$\varphi_{\mathcal{Z}}([\tilde{\alpha}]_k^n) = \tilde{\alpha}_k + \sum_{j \in \mathcal{Z}} (g(\tilde{\alpha}_j) - \tilde{\alpha}_j) + \sum_{j \in \mathcal{Z}^{\mathrm{c}}} g(\tilde{\alpha}_j), \quad (27)$$

$\Pr(S_0|\mathcal{Z})$ of (23) can be re-characterized as

$$\Pr(S_0|\mathcal{Z}) = \Pr(\varphi_{\mathcal{Z}}([\tilde{\alpha}]_k^n) > 0). \quad (28)$$

Since $\tilde{\alpha}_j$ can be approximated as a Gaussian distributed random variable [6], $g(\tilde{\alpha}_j)$ can similarly be considered Gaussian distributed, given the approximated linearity of $g(x)$. Let

$$\theta_j = \begin{cases} \tilde{\alpha}_j, & j = k; \\ g(\tilde{\alpha}_j), & j \in \mathcal{Z}^{\mathrm{c}}; \\ g(\tilde{\alpha}_j) - \tilde{\alpha}_j, & j \in \mathcal{Z} \end{cases} \quad (29)$$

denote each term in the summation of (27). Using (14) and (15), the expectation and variance of $\varphi_{\mathcal{Z}}([\tilde{\alpha}]_k^n)$ be computed as

$$\mathbb{E}[\varphi_{\mathcal{Z}}([\tilde{\alpha}]_k^n)] = \sum_{j \in \{k, k+1, \cdots, n\}} \mathbb{E}[\theta_j]; \quad (30)$$

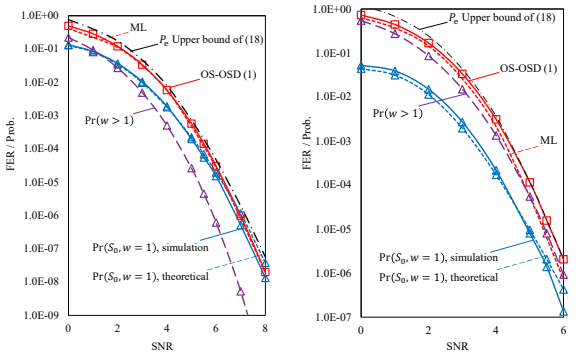
$$\mathbb{V}[\varphi_{\mathcal{Z}}([\tilde{\alpha}]_k^n)] = \sum_{p,j \in \{k, k+1, \cdots, n\}} \mathbb{V}[\theta_p, \theta_j], \quad (31)$$

where $\mathbb{V}[\theta_p, \theta_j]$ denotes the covariance of two random variables θ_p and θ_j. By assuming $\varphi_{\mathcal{Z}}([\tilde{\alpha}]_k^n)$ to be Gaussian distributed, $\Pr(S_0|\mathcal{Z})$ of (28) can be estimated as

$$\Pr(\varphi_{\mathcal{Z}}([\tilde{\alpha}]_k^n) > 0) \approx Q\left(\frac{-\mathbb{E}[\varphi_{\mathcal{Z}}([\tilde{\alpha}]_k^n)]}{\sqrt{\mathbb{V}[\varphi_{\mathcal{Z}}([\tilde{\alpha}]_k^n)]}} \right). \quad (32)$$

Therefore, $\sum_{\mathcal{Z}} \Pr(S_0|\mathcal{Z})$ can be estimated by traversing over all 2^{n-k} possible supports $\mathcal{Z}$. The decoding error probability upper bound of OS-OSD (1) given by (18) can be estimated. Its tightness will be verified via simulations in Section V, which will also show the performance loss incurred by this order skipping rule is negligible.

For the OS-OSD with $\tau > 1$, the performance analysis



(a) $(31, 21)$ BCH code (b) $(63, 45)$ BCH code

Fig. 1. Performance of OS-OSD (1).

becomes challenging. First of all, for $0 < i < \tau$, the event S_i can occur only if no order skipping occurs prior to phase-i. Moreover, the most likely codeword candidate $\tilde{\boldsymbol{c}}_{\rm b}$ is selected from a list. These two prerequisites of S_i make characterizing $\Pr(S_i, w)$ of (19) even more difficult. However, our simulation results of Section V will verify that the OS-OSD with $\tau > 1$ can yield a significant complexity reduction without compromising the decoding performance.

V. Simulation Results

Fig. 1(a) shows the frame error rate (FER) performance of the $(31, 21)$ BCH code with OS-OSD (1) over the AWGN channel. The ML decoding performance of the code [21] and the OS-OSD (1) decoding error probability ($P_{\rm e}$) upper bound of (18) are also shown. Fig. 1(a) shows that OS-OSD (1) can yield a near-ML decoding performance for this code. The FER performance of OS-OSD (1) is tightly upper bounded by the characterization of (18). The theoretical characterization of $\Pr(S_0, w = 1)$ is in line with the simulation at most of the SNR region. As the SNR continues to increase, it starts to deviate. This is due to the assumption that $\varphi_{\mathcal{Z}}([\tilde{\alpha}]_k^n)$ being Gaussian distributed becomes invalid as the SNR increases. For this code, $\Pr(S_0, w = 1) > \Pr(w > 1)$ appears in most of the shown SNR region, indicating $\Pr(S_0, w = 1)$ dominates in the list error probability of (21). However, with $\Pr(S_0, w = 1) < P_{\rm ML}$, the performance loss incurred by this order skipping is negligible. Fig. 1(b) further shows the performance of the $(63, 45)$ BCH code with OS-OSD (1). Again, the FER performance of OS-OSD (1) is tightly upper bounded by (18), verifying our performance analysis of OS-OSD (1). The theoretical characterization of $\Pr(S_0, w = 1)$ is again in line with the simulation results at most of the SNR region. For this code, it appears that $\Pr(S_0, w = 1) < \Pr(w > 1)$. Hence, $\Pr(w > 1)$ dominates in the list error probability of (21). The performance loss indicated by $\Pr(S_0, w = 1)$ is negligible.

Fig. 2 shows the FER performance of different algorithms for decoding the $(127, 64)$ BCH code over the AWGN channel. The OSD (τ) armed with probabilistic necessary condition (PNC) [5] and probabilistic sufficient condition (PSC) [8], denoted as PNC+PSC (τ), and the order-τ FOSD [18] with

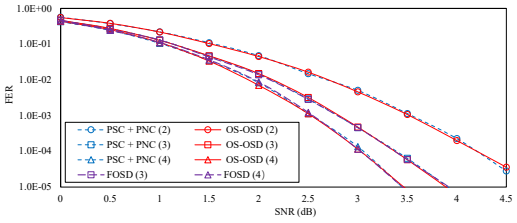


Fig. 2. Performance of the $(127, 64)$ BCH code.

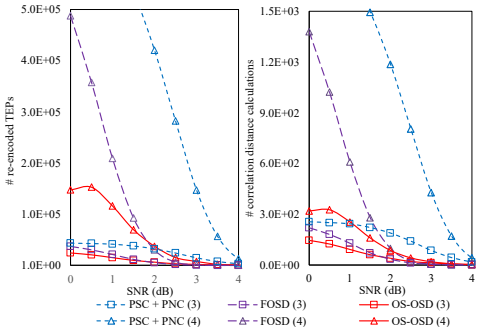


Fig. 3. Complexity of the $(127, 64)$ BCH code.

TABLE I
$\Pr(S_i)$ in OS-OSD (4) for the $(127, 64)$ BCH Code.

SNR(dB)	$\Pr(S_0)$	$\Pr(S_1)$	$\Pr(S_2)$	$\Pr(S_3)$
2	36.91%	32.53%	17.70%	8.01%
3	60.05%	27.17%	9.19%	2.75%
4	78.74%	17.25%	3.38%	0.54%
5	90.54%	8.61%	0.79%	0.05%

its optimal empirical factor $\beta = 4.3$, denoted as FOSD (τ), are compared. For PSC, we set $\tau_{\rm E} = 10$ for early termination and $\tau_{\rm PSC} = 25$ for skipping the correlation distance calculation as in [18]. Note that this PSC is also utilized in the OS-OSD for this code. Fig. 2 shows that the proposed OS-OSD can yield a similar FER performance as other algorithms for the code. Fig. 3 further shows the complexity of these decoding algorithms, in which the average numbers of re-encoded TEPs of (1) and correlation distance calculations of (2) are measured. Compared with the PSC+PNC (4) and FOSD (4), the OS-OSD (4) exhibits a significant complexity reduction, especially in the low SNR regime. This implies that the proposed CDLB is more accurate than that of the FOSD. Thus, the OS-OSD can more effectively prevent the higher-order re-encoding. Table I presents the statistical results of $P(S_i)$ during decoding the $(127, 64)$ BCH code with OS-OSD (4) (without PSC). It shows that as the SNR increases, $\Pr(S_0)$ becomes dominant. This implies most OS-OSD decoding events terminate after the phase-0 re-encoding, significantly reducing complexity.

Acknowledgment

This work is supported in part by the National Natural Science Foundation of China (NSFC) with project ID 62071498; and in part by the Natural Science Foundation of Guangdong Province with project ID 2024A1515010213.

REFERENCES

[1] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.

[2] B. Dorsch, "A decoding algorithm for binary block codes and J-ary output channels," *IEEE Transactions on Information Theory*, vol. 20, no. 3, pp. 391–394, May 1974.

[3] 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, 1995.

[4] M. C. Coşkun *et al.*, "Efficient error-correcting codes in the short blocklength regime," *Phys. Commun.*, vol. 34, pp. 66–79, 2019.

[5] Y. Wu and C. N. Hadjicostis, "Soft-decision decoding of linear block codes using preprocessing and diversification," *IEEE Transactions on Information Theory*, vol. 53, no. 1, pp. 378–393, 2007.

[6] C. Yue *et al.*, "A revisit to ordered statistics decoding: Distance distribution and decoding rules," *IEEE Trans. Inf. Theory*, vol. 67, no. 7, pp. 4288–4337, 2021.

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

[8] W. Jin and M. Fossorier, "Probabilistic sufficient conditions on optimality for reliability based decoding of linear block codes," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Seattle, WA, USA, July. 2006, pp. 2235–2239.

[9] C. Yue *et al.*, "Segmentation-discarding ordered-statistic decoding for linear block codes," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Waikoloa, HI, USA, Dec. 2019, pp. 1–6.

[10] ——, "Linear-equation ordered-statistics decoding," *IEEE Trans. Commun.*, vol. 70, no. 11, pp. 7105–7123, 2022.

[11] J. Liang *et al.*, "A low-complexity ordered statistic decoding of short block codes," *IEEE Commun. Lett.*, vol. 27, no. 2, pp. 400–403, 2023.

[12] A. Valembois and M. Fossorier, "Box and match techniques applied to soft-decision decoding," *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 796–810, 2004.

[13] M. Fossorier, "Reliability-based soft-decision decoding with iterative information set reduction," *IEEE Trans. Inf. Theory*, vol. 48, no. 12, pp. 3101–3106, 2002.

[14] W. Jin and M. P. C. Fossorier, "Reliability-based soft-decision decoding with multiple biases," *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 105–120, 2007.

[15] C. Choi and J. Jeong, "Fast soft decision decoding algorithm for linear block codes using permuted generator matrices," *IEEE Commun. Lett.*, vol. 25, no. 12, pp. 3775–3779, 2021.

[16] C. Yue *et al.*, "Ordered-statistics decoding with adaptive Gaussian elimination reduction for short codes," in *Proc. IEEE Global Commun. Conf. Workshops (GLOBECOM Wkshps)*, Rio de Janeiro, Brazil, Dec. 2022, pp. 492–497.

[17] L. Yang and L. Chen, "Low-latency ordered statistics decoding of BCH codes," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Mumbai, India, Nov. 2022, pp. 404–409.

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

[19] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*. New York, NY, USA: McGraw-Hill, 2002.

[20] G. Poltyrev, "Bounds on the decoding error probability of binary linear codes via their spectra," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1284–1292, 1994.

[21] M. Helmling *et al.*, "Database of channel codes and ML simulation results," www.uni-kl.de/channel-codes, 2019.