Plausibility Analysis of Shift-Sum Decoding for Cyclic Codes

Jiasheng Yuan †, Jiongyue Xing ‡, 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: yuanjsh@mail2.sysu.edu.cn, xingjyue@mail2.sysu.edu.cn, chenli55@mail.sysu.edu.cn

2021 IEEE International Symposium on Information Theory (ISIT) | 978-1-5386-8209-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ISIT45174.2021.9518200

***Abstract*—Using the minimum weight dual codewords (MWDCs) of a cyclic code, the shift-sum decoding can correct errors beyond half of the code's minimum Hamming distance. It utilizes the frequency of the syndrome polynomials' coefficients to identify the erroneous positions and correct the errors. This paper analyzes the plausibility of the shift-sum decoding for both binary and non-binary cyclic codes. It first determines the probability distributions of the frequency of the syndrome polynomials' coefficients as well as their expected values for the erroneous and non-erroneous positions. Based on these characterizations, this work further provides an analysis for the iterative shift-sum decoding, unveiling the statistical rationale on the shift-sum decoding's capability of correcting errors beyond the half distance bound.**

***Index Terms*—Cyclic codes, minimum weight dual codewords, plausibility analysis, shift-sum decoding**

I. Introduction

Cyclic codes are a wide class of channel codes including Reed-Solomon (RS) codes and BCH codes [1]. They have galvanized research interests due to their algebraic structures, simple encoding and efficient decoding algorithms. The traditional syndrome-based decoding, e.g., the Berlekamp-Massey (BM) algorithm [2], can correct errors up to half of the code's minimum Hamming distance. Interpolation based algebraic list decoding [3] [4] can correct errors beyond the above distance bound but with a higher complexity. With soft received information, Chase decoding [5] and ordered statistics decoding [6] can achieve competent performance with a moderate complexity. To improve the decoding performance, several adjusted belief-propagation (BP) algorithms have been proposed to decode cyclic codes. Jiang and Narayanan proposed an adaptive BP algorithm with the aid of the BM algorithm to achieve a near maximum-likelihood (ML) decoding performance for short RS codes [7]. The multiple-bases BP algorithm utilizes several parity-check matrices for the decoding, also achieving a near ML performance [8] [9].

Recently, shift-sum decoding that utilizes the minimum weight dual codewords (MWDCs) has been proposed by Bossert [10] [11]. For binary BCH codes, simulations show that it can correct errors beyond the half distance bound. This decoding was later extended to decode non-binary cyclic codes, which also demonstrates an advanced decoding performance [12]. However, the rationale on why the shift-sum decoding exhibits such an advanced error-correction capability has not been fully understood. This paper provides our recent study findings and analyzes the plausibility of the shift-sum decoding. In general, each MWDC used for the shift-sum decoding produces a syndrome polynomial. Their coefficients indicate the erroneous positions and the corresponding magnitudes, which are considered as syndrome random variables (SRVs) during the plausibility analysis. The probability distributions of SRVs at the erroneous and the non-erroneous positions will be determined, respectively. Their expected values for both binary and non-binary cyclic codes will also be derived, improving the earlier results of [11]. Finally, we provide a probabilistic analysis of the iterative shift-sum decoding, and reveal the reason for why the shift-sum decoding can correct errors beyond the half distance bound.

II. Preliminaries

Let $\mathbb{F}_q = \{\sigma_0, \sigma_1, \dots, \sigma_{q-1}\}$ denote a finite field of size q and $\mathbb{F}_q[x]$ denote the univariate polynomial ring over $\mathbb{F}_q$. For simplicity, this paper considers the binary extension fields, i.e., $q = 2^s$, where $s \in \mathbb{N}$, and the codes with length $n = 2^s - 1$. Let $C(n, k, d)$ denote a cyclic code defined over $\mathbb{F}_{2^p}$ with a dimension k and the minimum Hamming distance d, where $1 \le p \le s$. Note that when $p = 1$, C is a binary BCH code. When $p > 1$, C is a non-binary BCH code. Especially, when $p = s$, C becomes an RS code. Its dual code is denoted as $C^{\perp}(n, n-k, d^{\perp})$. Let $\boldsymbol{c} = (c_0, c_1, \dots, c_{n-1}) \in C(n, k, d)$ denote a codeword, which can also be represented as $c(x) = c_0 + c_1 x + \cdots + c_{n-1}x^{n-1}$. The support of $c(x)$ is defined as $\text{supp}(c(x)) = \{j | c_j \neq 0, \forall j\}$.

Definition I: Given two distinct codewords $c^{(1)}(x)$, $c^{(2)}(x) \in C$, they are cyclically different if

$$c^{(2)}(x) \neq \sigma x^j c^{(1)}(x) \mod (x^n - 1),$$

for all $j \in \{0, 1, \dots, n-1\}$ and $\sigma \in \mathbb{F}_q \backslash \{0\}$.

Encoding of cyclic codes is defined by its generator polynomial $g(x)$ with degree $n - k$. Given message polynomial $m(x) = m_0 + m_1 x + \cdots + m_{k-1}x^{k-1}$, and $m(x) \in \mathbb{F}_q[x]$, codeword $c(x)$ can be generated by

$$c(x) = m(x)g(x). \tag{1}$$

The check polynomial $h(x)$ is defined as

$$h(x) = \frac{x^n - 1}{g(x)}. \tag{2}$$

978-1-5386-8209-8/21/$31.00 ©2021 IEEE

Note that dual code $C^\perp$ is also a cyclic code and its generator polynomial would be $h(x)$. The MWDCs are the codewords in $C^\perp$ with a Hamming weight of $d^\perp$. Denoting the non-zero coefficients of a MWDC as $\beta_{b_0}, \beta_{b_1}, \cdots, \beta_{b_{d^\perp-1}}$, it can be written as

$$b(x) = \beta_{b_0} x^{b_0} + \beta_{b_1} x^{b_1} + \cdots + \beta_{b_{d^\perp-1}} x^{b_{d^\perp-1}}. \tag{3}$$

Note that $c(x)b(x) = 0 \bmod (x^n - 1)$.

III. Shift-Sum Decoding

This section briefly reviews the shift-sum decoding which utilizes a number of cyclically different MWDCs to determine the erroneous positions and their magnitudes [11].

Since $C^\perp$ is linear and cyclic, recalling (3), we can assume $b_0 = 0$ and $\beta_{b_0} = 1$. Suppose τ erroneous positions are $e_0, e_1, \ldots, e_{\tau-1}$ and their corresponding magnitudes are $\varepsilon_{e_0}, \varepsilon_{e_1}, \ldots, \varepsilon_{e_{\tau-1}}$, respectively. The error polynomial can be written as

$$\varepsilon(x) = \varepsilon_{e_0} x^{e_0} + \varepsilon_{e_1} x^{e_1} + \cdots + \varepsilon_{e_{\tau-1}} x^{e_{\tau-1}}, \tag{4}$$

and the received polynomial is $r(x) = c(x) + \varepsilon(x)$, which can be written as

$$r(x) = r_0 + r_1 x + \cdots + r_{n-1} x^{n-1}. \tag{5}$$

The syndrome polynomial $w(x)$ is defined by

$$\begin{aligned} w(x) &= r(x)b(x) \\ &= (c(x) + \varepsilon(x))b(x) \\ &= \varepsilon(x)b(x) \mod (x^n - 1)\,. \end{aligned} \tag{6}$$

Elaborating on $w(x)$ will lead to

$$\begin{aligned} w(x) =& \beta_{b_0} x^{b_0} \varepsilon(x) + \cdots + \beta_{b_{d^\perp-1}} x^{b_{d^\perp-1}} \varepsilon(x) \mod (x^n - 1) \\ =& \varepsilon_{e_0} x^{e_0} + \cdots + \varepsilon_{e_{\tau-1}} x^{e_{\tau-1}} + \\ & \beta_{b_1} \varepsilon_{e_0} x^{e_0 + b_1} + \cdots + \beta_{b_1} \varepsilon_{e_{\tau-1}} x^{e_{\tau-1} + b_1} + \\ & \vdots \\ & \beta_{b_{d^\perp-1}} \varepsilon_{e_0} x^{e_0 + b_{d^\perp-1}} + \cdots + \beta_{b_{d^\perp-1}} \varepsilon_{e_{\tau-1}} x^{e_{\tau-1} + b_{d^\perp-1}}, \end{aligned} \tag{7}$$

where the exponents are calculated mod n. Note that any non-zero coefficient of polynomial $w(x)$ is an error or a shifted scalar error. Therefore, the above $w(x)$ can indicate the erroneous positions and their magnitudes. Let $b_h(x) = \frac{x^{-h}}{\beta_h} b(x)$, where $h \in \mathrm{supp}(b(x))$, we can obtain $d^\perp$ syndrome polynomials $w_h(x) = \frac{x^{-h}}{\beta_h} w(x)$. This implies that non-zero coefficients of $w(x)$ can be shifted back to their original positions. Therefore, the errors would occur more frequently among the coefficients of polynomials $w_h(x)$.

Based on the above observations, we can use a number of cyclically different MWDCs to determine the erroneous positions and their magnitudes. Assume that there are L cyclically different MWDCs

$$b^{(\ell)}(x) = 1 + \beta^{(\ell)}_{b_1} x^{b_1} + \cdots + \beta_{b^{(\ell)}_{d^\perp-1}} x^{b_{d^\perp-1}}, \tag{8}$$

where $\ell = 1, 2, \ldots, L$. They can be utilized to generate L syndrome polynomials $w^{(\ell)}(x)$ as

$$w^{(\ell)}(x) = r(x) b^{(\ell)}(x) = \varepsilon(x) b^{(\ell)}(x) \mod (x^n - 1)\,. \tag{9}$$

Moreover, each of them has $d^\perp$ shifted counterparts, which are represented as

$$w^{(\ell)}_h(x) = \frac{x^{-h}}{\beta^{(\ell)}_h} r(x) b^{(\ell)}(x) \mod (x^n - 1)\,. \tag{10}$$

For each $w^{(\ell)}_h(x)$, its coefficients $w^{(\ell)}_{h,j}$ can be written as

$$w^{(\ell)}_{h,j} = \frac{1}{\beta^{(\ell)}_h} \sum_{u \in \mathrm{supp}\left(b^{(\ell)}(x)\right)} \beta^{(\ell)}_u r_{(j+h-u) \bmod n}, \tag{11}$$

where $j = 0, 1, \ldots, n-1$. The shift-sum decoding counts the frequency of each element σ_i ($\sigma_i \in \mathbb{F}_q$) at position j based on $w^{(\ell)}_{h,j}$, denoted as $\phi_{i,j}$. Let Φ denote the frequency matrix with entry $\phi_{i,j}$ as

$$\Phi = \begin{bmatrix} \phi_{0,0} & \phi_{0,1} & \cdots & \phi_{0,n-1} \\ \phi_{1,0} & \phi_{1,1} & \cdots & \phi_{1,n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{q-1,0} & \phi_{q-1,1} & \cdots & \phi_{q-1,n-1} \end{bmatrix}. \tag{12}$$

Note that the sum of each column is $d^\perp L$, i.e., $\sum_{i=0}^{q-1} \phi_{i,j} = d^\perp L$. The iterative shift-sum decoding has been proposed for binary and non-binary cyclic codes in [10] and [12], respectively. Their simulation results have shown that the decoding can correct errors beyond half of the code's minimum Hamming distance.

Recently, Bossert has given a plausibility analysis of this decoding mechanism for binary codes [11], where the coefficients of syndrome polynomial $w^{(l)}_h(x)$ are either zero or one. Given that there are τ errors, the expected weight v of each $w(x)$ is [11]

$$\mathbb{E}[v] = \frac{n \sum_{i=1}^{\tau} \binom{d^\perp}{i} \binom{n-d^\perp}{\tau-i}}{\binom{n}{\tau}}, \tag{13}$$

where i is odd, $i \le d^\perp$ and $\tau - i \le n - d^\perp$. Let $\Phi_{\mathrm{e}}(\tau)$ and $\Phi_{\mathrm{c}}(\tau)$ denote the frequency of one among all coefficients $w^{(\ell)}_{h,j}$ for the erroneous and the non-erroneous positions, respectively. Their expected values can be obtained by

$$\mathbb{E}_0 \left[\Phi_{\mathrm{e}}(\tau)\right] = \frac{\mathbb{E}[v]}{\tau} L \tag{14}$$

and

$$\mathbb{E}_0 \left[\Phi_{\mathrm{c}}(\tau)\right] = \frac{d^\perp \left(\mathbb{E}[v] - \frac{\mathbb{E}[v]}{d^\perp}\right)}{n - \tau} L, \tag{15}$$

respectively. However, these characterizations deviate from the simulation results for large τ. The following section gives a more accurate characterization.

IV. Plausibility Analysis

This section analyzes plausibility of the shift-sum decoding, which determines the statistical distributions of the frequency matrix Φ. In particular, fixing a particular error weight τ, we characterize the expected values of entries in Φ at the erroneous and non-erroneous positions for both binary and non-binary codes.

A. Binary Codes

Given τ errors, the error polynomial can be written as $\varepsilon(x) = x^{e_0} + x^{e_1} + \cdots + x^{e_{\tau-1}}$. Based on (7), we have

$$w_j = \varepsilon_j + \varepsilon_{j-b_1} + \varepsilon_{j-b_2} + \cdots + \varepsilon_{j-b_{d^\perp-1}}, \tag{16}$$

where the addition is performed over $\mathbb{F}_2$. It can be seen that if there are an odd number of ones among ε_{j-b_1}, ε_{j-b_2}, $\ldots, \varepsilon_{j-b_{d^\perp-1}}$, $w_j \neq \varepsilon_j$. Let $M_\mathrm{e}(\tau)$ denote the number of such cases for the erroneous positions ($\varepsilon_j = 1$). It can be determined by

$$M_\mathrm{e}(\tau) = \sum_{i \text{ is odd}} \tbinom{d^\perp-1}{i}\tbinom{n-d^\perp}{\tau-1-i}. \tag{17}$$

Note that $1 \le i \le d^\perp - 1$ and $\tau - 1 - i \le n - d^\perp$. Similarly, let $M_\mathrm{c}(\tau)$ denote the number for the non-erroneous cases ($\varepsilon_j = 0$), which can be determined by

$$M_\mathrm{c}(\tau) = \sum_{i \text{ is odd}} \tbinom{d^\perp-1}{i}\tbinom{n-d^\perp}{\tau-i}. \tag{18}$$

Note that $1 \le i \le d^\perp - 1$ and $\tau - i \le n - d^\perp$. It can be seen that $M_\mathrm{c}(\tau) = M_\mathrm{e}(\tau + 1)$.

Further let w_{j_e} and w_{j_c} denote the coefficients at the erroneous positions ($\varepsilon_j = 1$) and the non-erroneous positions ($\varepsilon_j = 0$), respectively. The probability of w_{j_e} being one, denoted as $p_\mathrm{e}(\tau)$, can be calculated by

$$p_\mathrm{e}(\tau) = \Pr(w_{j_\mathrm{e}} = 1) = 1 - \frac{M_\mathrm{e}(\tau)}{\binom{n-1}{\tau-1}}. \tag{19}$$

Similarly, the probability of w_{j_c} being one can be calculated by

$$p_\mathrm{c}(\tau) = \Pr(w_{j_\mathrm{c}} = 1) = \frac{M_\mathrm{c}(\tau)}{\binom{n-1}{\tau}}. \tag{20}$$

The above equations show that $p_\mathrm{e}(\tau)$ and $p_\mathrm{c}(\tau)$ are relevant to the minimum Hamming distance of the dual code and codeword length. Let us define

$$\Delta(\tau) = p_\mathrm{e}(\tau) - p_\mathrm{c}(\tau) \tag{21}$$

as the discrepancy between $p_\mathrm{e}(\tau)$ and $p_\mathrm{c}(\tau)$. The values of $\Delta(\tau)$ indicates the statistical difference between the erroneous and the non-erroneous positions. Fig. 1 shows for a cyclic code of length $n = 63$, how $\Delta(\tau)$ varies with τ under different $d^\perp$. It can be seen that $\Delta(\tau)$ increases as $d^\perp$ decreases.

For binary case, we define the SRV $W_\mathrm{e}(w_{j_\mathrm{e}})$ associated with the coefficients w_{j_e} as follows,

$$W_\mathrm{e}(w_{j_\mathrm{e}}) = \begin{cases} 1 - p_\mathrm{e}(\tau), & \text{if } w_{j_\mathrm{e}} = 0, \\ p_\mathrm{e}(\tau), & \text{if } w_{j_\mathrm{e}} = 1, \end{cases} \tag{22}$$

where the coefficients $w_{j_\mathrm{e}} \in \mathbb{F}_2$. Similarly, the SRV $W_\mathrm{c}(w_{j_\mathrm{c}})$ is defined by

$$W_\mathrm{c}(w_{j_\mathrm{c}}) = \begin{cases} 1 - p_\mathrm{c}(\tau), & \text{if } w_{j_\mathrm{e}} = 0, \\ p_\mathrm{c}(\tau), & \text{if } w_{j_\mathrm{c}} = 1. \end{cases} \tag{23}$$

where the coefficients $w_{j_\mathrm{c}} \in \mathbb{F}_2$. From eqs. (22) and (23), SRVs $W_\mathrm{e}(w_{j_\mathrm{e}})$ and $W_\mathrm{c}(w_{j_\mathrm{c}})$ produced by the shift-sum decoding satisfy the Bernoulli distribution [14]. Suppose L cyclically different MWDCs are used for the decoding. The $d^\perp L$ SRVs at erroneous positions are denoted as $W_\mathrm{e}^{(0)}, W_\mathrm{e}^{(1)}, \ldots, W_\mathrm{e}^{(d^\perp L-1)}$, while the $d^\perp L$ SRVs at non-erroneous positions are denoted as $W_\mathrm{c}^{(0)}, W_\mathrm{c}^{(1)}, \ldots, W_\mathrm{c}^{(d^\perp L-1)}$. Since $\Phi_\mathrm{e}(\tau)$ is the summation of the $d^\perp L$ SRVs $W_\mathrm{e}^{(0)}, W_\mathrm{e}^{(1)}, \ldots, W_\mathrm{e}^{(d^\perp L-1)}$, its expected value can be determined by

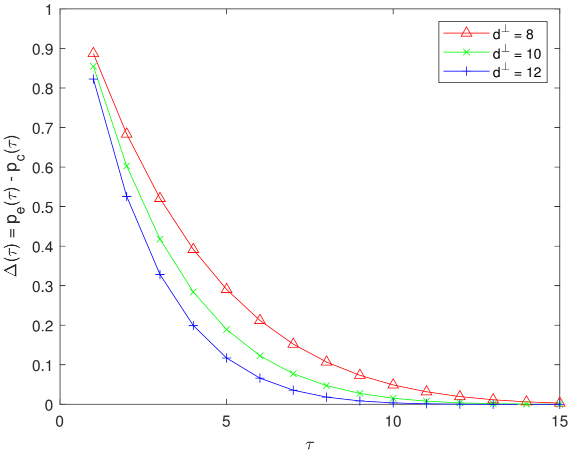

Fig. 1. The discrepancy $\Delta(\tau)$ with $n = 63$ and $d^\perp = 8, 10, 12$.

$$\begin{aligned} \mathbb{E}[\Phi_\mathrm{e}(\tau)] &= \mathbb{E}[\sum_{l=0}^{d^\perp L-1} W_\mathrm{e}^{(l)}] \\ &= \sum_{l=0}^{d^\perp L-1} \mathbb{E}[W_\mathrm{e}^{(l)}] \\ &= p_\mathrm{e}(\tau) d^\perp L. \end{aligned} \tag{24}$$

The expected value of $\Phi_\mathrm{c}(\tau)$ can be determined by

$$\begin{aligned} \mathbb{E}[\Phi_\mathrm{c}(\tau)] &= \mathbb{E}[\sum_{l=0}^{d^\perp L-1} W_\mathrm{c}^{(l)}] \\ &= \sum_{l=0}^{d^\perp L-1} \mathbb{E}[W_\mathrm{c}^{(l)}] \\ &= p_\mathrm{c}(\tau) d^\perp L. \end{aligned} \tag{25}$$

Example 1: Given a BCH code $C(63, 24, 15)$, its dual code is also a BCH code $C(63, 39, 8)$. There are $L = 35$ cyclically different dual codewords with $d^\perp = 8$. The average value (AV) of $\Phi_\mathrm{e}(\tau)$ and $\Phi_\mathrm{c}(\tau)$ were obtained through simulations by running 10 000 decoding events for each τ. Fig. 2 shows that our characterizations of $\mathbb{E}[\Phi_\mathrm{e}(\tau)]$ and $\mathbb{E}[\Phi_\mathrm{c}(\tau)]$ agree well with the empirical results of AV$[\Phi_\mathrm{e}(\tau)]$ and AV$[\Phi_\mathrm{c}(\tau)]$, respectively. These characteristics improve over the results of [11], i.e., $\mathbb{E}_0[\Phi_\mathrm{e}(\tau)]$ and $\mathbb{E}_0[\Phi_\mathrm{c}(\tau)]$.

B. Non-binary Codes

For a non-binary code defined over $\mathbb{F}_q$, the error polynomial is $\varepsilon(x) = \varepsilon_{e_0} x^{e_0} + \varepsilon_{e_1} x^{e_1} + \cdots + \varepsilon_{e_{\tau-1}} x^{e_{\tau-1}}$. Since $w(x) = b(x)\varepsilon(x) \bmod (x^n - 1)$, its coefficients w_j can be calculated by

$$w_j = \varepsilon_j + \beta_{b_1}\varepsilon_{j-b_1} + \beta_{b_2}\varepsilon_{j-b_2} + \cdots + \beta_{b_{d^\perp-1}}\varepsilon_{j-b_{d^\perp-1}}, \tag{26}$$

where the addition is performed over $\mathbb{F}_q$.

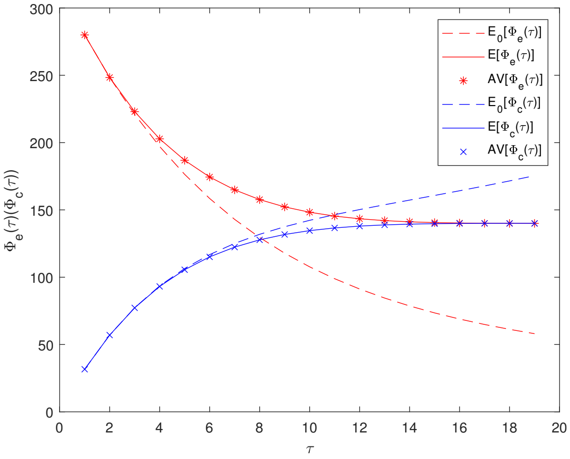

Fig. 2. Expected and empirical values of $\Phi_e(\tau)$ and $\Phi_c(\tau)$.

Lemma 1: Suppose the error magnitude is equally drawn from $\mathbb{F}_q\backslash\{0\}$. Let A_i denote the probability of $\sum_{s=1}^{i}\beta_{b_s}\varepsilon_{j-b_s}$ to be nonzero. For all i, we have

$$A_i = 1 - \frac{1}{q} + \frac{1}{q}\left(\frac{1}{q-1}\right)^{i-1}(-1)^{i-1}. \tag{27}$$

Proof: Let $\sigma_{j,s} = \beta_{b_s}\varepsilon_{j-b_s}$. Since $\beta_{b_s} \neq 0$, $\sigma_{j,s}$ is also equally drawn from $\mathbb{F}_q\backslash\{0\}$. If the sum of the first i symbols equals to zero, i.e., $\sum_{s=1}^{i}\sigma_{j,s} = 0$, then $\sum_{s=1}^{i+1}\sigma_{j,s} \neq 0$ since $\sigma_{j,i+1} \neq 0$. On the other hand, if $\sum_{s=1}^{i}\sigma_{j,s} \neq 0$, $\sum_{s=1}^{i+1}\sigma_{j,s} = 0$ only if $\sum_{s=1}^{i}\sigma_{j,s} = \sigma_{j,i+1}$. Since $\sigma_{j,i+1}$ is equally drawn from $\mathbb{F}_q\backslash\{0\}$, the probability of $\sum_{s=1}^{i+1}\sigma_{j,s} \neq 0$ is $\frac{q-2}{q-1}$. Therefore, the relation between A_i and A_{i+1} is

$$A_{i+1} = \frac{q-2}{q-1}A_i + (1 - A_i).$$

The above equation can be rewritten as

$$A_{i+1} - \frac{q-1}{q} = -\frac{1}{q-1}\left(A_i - \frac{q-1}{q}\right).$$

Hence, the sequence $A_i - \frac{q-1}{q}$ is a geometric sequence. With the initial condition $A_1 = 1$, we can obtain (27). □

Based on the above lemma, the probability of $w_{j_e} = \varepsilon_j$ can be calculated by

$$\Pr(w_{j_e} = \varepsilon_j) = 1 - \frac{\sum_{i=1}^{\tau-1} A_i\binom{d^{\perp}-1}{i}\binom{n-d^{\perp}}{\tau-1-i}}{\binom{n-1}{\tau-1}}, \tag{28}$$

where $i \leq d^{\perp}-1$ and $\tau-1-i \leq n-d^{\perp}$. For the other elements $\sigma \in \mathbb{F}_q$ and $\sigma \neq \varepsilon_j$, the probability is

$$\Pr(w_{j_e} = \sigma, \sigma \neq \varepsilon_j) = \frac{\sum_{i=1}^{\tau-1} A_i\binom{d^{\perp}-1}{i}\binom{n-d^{\perp}}{\tau-1-i}}{(q-1)\binom{n-1}{\tau-1}}. \tag{29}$$

Similarly, the probability of $w_{j_c} = 0$ is

$$\Pr(w_{j_c} = 0) = 1 - \frac{\sum_{i=1}^{\tau} A_i\binom{d^{\perp}-1}{i}\binom{n-d^{\perp}}{\tau-i}}{\binom{n-1}{\tau}}, \tag{30}$$

where $i \leq d^{\perp}-1$ and $\tau - i \leq n - d^{\perp}$. Finally, for the elements $\sigma \in \mathbb{F}_q$ and $\sigma \neq 0$, the probability is

$$\Pr(w_{j_c} = \sigma, \sigma \neq 0) = \frac{\sum_{i=1}^{\tau} A_i\binom{d^{\perp}-1}{i}\binom{n-d^{\perp}}{\tau-i}}{(q-1)\binom{n-1}{\tau}}. \tag{31}$$

Note that the above derivations hold for every dual codeword. Assume L cyclically different MWDCs are used for decoding, the expected values of every element in matrix Φ can also be obtained, similar to the binary case. Unfortunately, the number of cyclically different MWDCs for most of the cyclic codes is unknown, except the RS codes. Given an RS code $C(n,k,d)$, the number of cyclically different MWDCs is [13]

$$L_{\text{RS}} = \frac{1}{n}\sum_{j|\text{GCD}(n-k-1,n)}\varphi(j)\binom{n/j}{(n-k-1)/j}, \tag{32}$$

where $\varphi(\cdot)$ is the Euler's totient function and $\text{GCD}(n-k-1,n)$ is the greatest common divisor of $n-k-1$ and n.

Example 2: Given an RS code $C(15,5,11)$, it can correct five errors using the BM algorithm. The dual code is also an RS code $C(15,10,6)$, which has $L_{\text{RS}} = 335$ MWDCs. Fig. 3 shows the four probabilities derived from eqs. (28) – (31). When $\tau \leq 7$, $\Pr(w_{j_e} = \varepsilon_j)$ is larger than $\Pr(w_{j_e} = \sigma, \sigma \neq \varepsilon_j)$ and $\Pr(w_{j_c} = \sigma, \sigma \neq 0)$. This indicates that the decoding can identify the erroneous positions and their magnitudes with a higher probability, even when the error number is larger than the half distance bound.

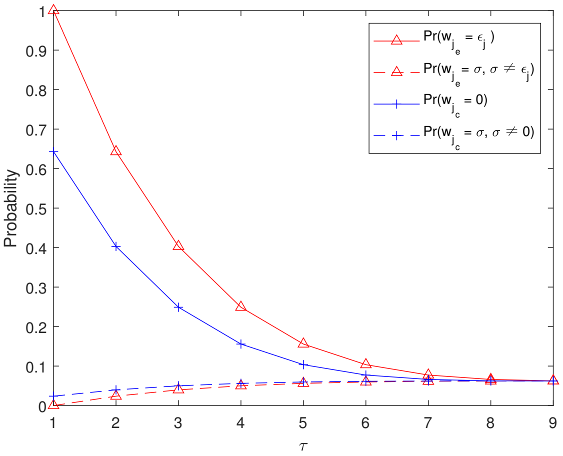

Fig. 3. Probability distributions of w_{j_e} and w_{j_c} of the RS code $C(15,10,6)$.

V. Iterative Decoding Analysis

Armed with the above knowledge, we can now analyze the iterative shift-sum decoding of binary codes from a perspective probability.

Based on the analysis of $\Phi_e(\tau)$ and $\Phi_c(\tau)$, a heuristic iterative decoding strategy can be derived accordingly. Recall the frequency of one at j^{th} positions is denoted by $\phi_{1,j}$. A larger value of $\phi_{1,j}$ indicates the position is more likely to be erroneous. In [10] – [12], the received bits r_j corresponding the λ largest $\phi_{1,j}$ would be flipped at each iteration, where λ is a preset parameter in the decoding. The frequency matrix Φ will be computed again through multiplying the updated $r(x)$ by the cyclically different MWDCs. Repeat this process until a valid codeword is obtained or the maximum number of iterations is reached. For simplicity, we consider $\lambda = 1$ in this section. In this decoding strategy, if the flipping bit is erroneous, the error number will be reduced by one. Otherwise, it will be increased by one. Let $P_{\mathrm{red}}(\tau)$ and $P_{\mathrm{inc}}(\tau)$ denote the error reduction and error increase probabilities, respectively. Note that $P_{\mathrm{red}}(\tau) + P_{\mathrm{inc}}(\tau) = 1, \forall \tau$. The iterative shift-sum decoding functions if $P_{\mathrm{red}}(\tau) > P_{\mathrm{inc}}(\tau)$ holds at each iteration. We will analyze $P_{\mathrm{red}}(\tau)$ in the following.

To simplify the analysis, we assume the SRVs $W_{\mathrm{e}}^{(0)}$, $W_{\mathrm{e}}^{(1)}$, $\ldots, W_{\mathrm{e}}^{(d^{\perp}L-1)}$ and $W_{\mathrm{c}}^{(0)}$, $W_{\mathrm{c}}^{(1)}$, $\ldots, W_{\mathrm{c}}^{(d^{\perp}L-1)}$ are independent and identically distributed. Therefore, $\Phi_{\mathrm{e}}(\tau)$ and $\Phi_{\mathrm{c}}(\tau)$ should satisfy the binomial distributions $B(d^{\perp}L, p_{\mathrm{e}}(\tau))$ and $B(d^{\perp}L, p_{\mathrm{c}}(\tau))$, respectively. Hence, their probability mass functions (PMFs) are

$$\begin{aligned} \Phi_{\mathrm{e}}{}^{(\tau)}(\xi) &= \Pr(\Phi_{\mathrm{e}}(\tau) = \xi) \\ &= \binom{d^{\perp}L}{\xi} p_{\mathrm{e}}^{\xi}(\tau)(1 - p_{\mathrm{e}}(\tau))^{d^{\perp}L-\xi}, \end{aligned} \tag{33}$$

and

$$\begin{aligned} f_{\Phi_{\mathrm{c}}}^{(\tau)}(\xi) &= \Pr(\Phi_{\mathrm{c}}(\tau) = \xi) \\ &= \binom{d^{\perp}L}{\xi} p_{\mathrm{c}}^{\xi}(\tau)(1 - p_{\mathrm{c}}(\tau))^{d^{\perp}L-\xi}, \end{aligned} \tag{34}$$

respectively, where $0 \le \xi \le d^{\perp}L$. Furthermore, the cumulative distribution functions (CDFs) can be determined by

$$\begin{aligned} F_{\Phi_{\mathrm{e}}}^{(\tau)}(\xi) &= \Pr(\Phi_{\mathrm{e}}(\tau) \le \xi) \\ &= \sum_{l=0}^{\xi} \binom{d^{\perp}L}{l} p_{\mathrm{e}}(\tau)^{l}(1 - p_{\mathrm{e}}(\tau))^{d^{\perp}L-l}, \end{aligned} \tag{35}$$

and

$$\begin{aligned} F_{\Phi_{\mathrm{c}}}^{(\tau)}(\xi) &= \Pr(\Phi_{\mathrm{c}}(\tau) \le \xi) \\ &= \sum_{l=0}^{\xi} \binom{d^{\perp}L}{l} p_{\mathrm{c}}(\tau)^{l}(1 - p_{\mathrm{c}}(\tau))^{d^{\perp}L-l}, \end{aligned} \tag{36}$$

respectively. Let $\Phi_{\mathrm{e,max}}(\tau)$ denote the largest value among the τ erroneous positions. Based on (35), the CDFs of $\Phi_{\mathrm{e,max}}(\tau)$ can be derived as

$$F_{\Phi_{\mathrm{e,max}}}^{(\tau)}(\xi) = \Pr(\Phi_{\mathrm{e,max}}(\tau) \le \xi) = (F_{\Phi_{\mathrm{e}}}^{(\tau)}(\xi))^{\tau}. \tag{37}$$

Similarly, the CDFs of the largest value $\Phi_{\mathrm{c,max}}(\tau)$ among the non-erroneous positions can be computed by

$$F_{\Phi_{\mathrm{c,max}}}^{(\tau)}(\xi) = \Pr(\Phi_{\mathrm{c,max}}(\tau) \le \xi) = (F_{\Phi_{\mathrm{c}}}^{(\tau)}(\xi))^{n-\tau}. \tag{38}$$

Based on the relationship between the CDFs and PMFs, the PMFs of $\Phi_{\mathrm{e,max}}(\tau)$ can be determined by

$$\begin{aligned} f_{\Phi_{\mathrm{e,max}}}^{(\tau)}(\xi) &= \Pr(\Phi_{\mathrm{e,max}}(\tau) \le \xi) - \Pr(\Phi_{\mathrm{e,max}}(\tau) < \xi) \\ &= (F_{\Phi_{\mathrm{e}}}^{(\tau)}(\xi))^{\tau} - (F_{\Phi_{\mathrm{e}}}^{(\tau)}(\xi) - f_{\Phi_{\mathrm{e}}}^{(\tau)}(\xi))^{\tau}. \end{aligned} \tag{39}$$

Meanwhile, the PMFs of $\Phi_{\mathrm{c,max}}(\tau)$ can be determined by

$$\begin{aligned} f_{\Phi_{\mathrm{c,max}}}^{(\tau)}(\xi) &= \Pr(\Phi_{\mathrm{c,max}}(\tau) \le \xi) - \Pr(\Phi_{\mathrm{c,max}}(\tau) < \xi) \\ &= (F_{\Phi_{\mathrm{c}}}^{(\tau)}(\xi))^{n-\tau} - (F_{\Phi_{\mathrm{c}}}^{(\tau)}(\xi) - f_{\Phi_{\mathrm{c}}}^{(\tau)}(\xi))^{n-\tau}. \end{aligned} \tag{40}$$

Based on the above analysis, we can obtain the error reduction probability $P_{\mathrm{red}}(\tau)$ as

$$P_{\mathrm{red}}(\tau) = \sum_{\xi=0}^{d^{\perp}L} f_{\Phi_{\mathrm{e,max}}}^{(\tau)}(\xi)(F_{\Phi_{\mathrm{c,max}}}^{(\tau)}(\xi) - f_{\Phi_{\mathrm{c,max}}}^{(\tau)}(\xi)). \tag{41}$$

Note that $P_{\mathrm{inc}}(\tau) = 1 - P_{\mathrm{red}}(\tau)$. The following example validates our analysis as well as the advanced error correction capability of the shift-sum decoding.

Example 3: Given a BCH code $\mathcal{C}(63, 24, 15)$, it can correct up to seven errors using the BM algorithm. Fig. 4 compares our characterizations of $P_{\mathrm{red}}(\tau)$ and $P_{\mathrm{inc}}(\tau)$ with the simulation results of $\mu_{\mathrm{red}}(\tau)$ and $\mu_{\mathrm{inc}}(\tau)$, which are obtained by counting the frequencies of error reduction and error increase for a given τ. It shows that our analysis is a tight estimation of the empirical results. Moreover, it should be pointed out that when $\tau \le 11$, $P_{\mathrm{red}}(\tau)$ is larger than $P_{\mathrm{inc}}(\tau)$. This indicates that in this region the shift-sum decoding has a higher probability of flipping an erroneous position than a non-erroneous position, resulting in the error-correction. Therefore, it shows the shift-sum decoding can correct errors beyond half of the code's minimum Hamming distance.

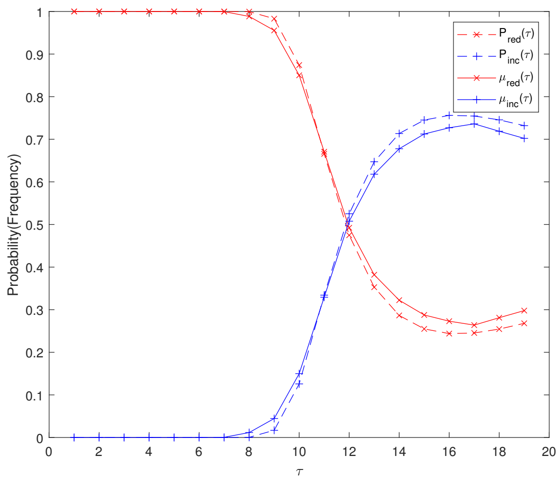

Fig. 4. Comparison between analytical and simulation results.

VI. Conclusion

This paper has analyzed the plausibility of the shift-sum decoding for both binary and non-binary cyclic codes. By regarding coefficients of the syndrome polynomials as random variables, their probability distributions and the corresponding expected values have been derived for both the erroneous and the non-erroneous positions. Finally, we have analyzed the iterative shift-sum decoding scheme of binary codes, revealing the rationale on the shift-sum decoding's capability of correcting errors beyond the half distance bound.

Acknowledgment

This work is sponsored by National Natural Science Foundation of China (NSFC) with project ID 62071493.

References

[1] F. MacWilliams and N. Sloane, *The theory of error correcting codes*, Elsevier, 1977.

[2] J. Massey, "Shift register synthesis and BCH decoding," *IEEE Trans. Inform. Theory*, vol. 15, no. 1, pp. 122-127, Jan. 1969.

[3] V. Guruswami and M. Sudan, "Improved decoding of Reed-Solomon and algebraic-geometric codes," *IEEE Trans. Inform. Theory*, vol. 45, no. 6, pp. 1757-1767, Sept. 1999.

[4] Y. Wu, "New list decoding algorithms for Reed-Solomon and BCH codes," *IEEE Trans. Inform. Theory*, vol. 54, no. 8, pp. 3611-3630, Aug. 2008.

[5] D. Chase, "Class of algorithms for decoding block codes with channel measurement information," *IEEE Trans. Inform. Theory*, vol. 18, no. 1, pp. 170-182, Jan. 1972.

[6] M. P. Fossorier and S. Lin, "Soft-decision decoding of linear block codes based on ordered statistics," *IEEE Trans. Inform. Theory*, vol. 41, no. 5, pp. 1379-1396, Sept. 1995.

[7] J. Jiang and K. R. Narayanan, "Iterative soft-input soft-output decoding of Reed-Solomon codes by adapting the parity-check matrix," *IEEE Trans. Inform. Theory*, vol. 52, no. 8, pp. 3746-3756, Aug. 2006.

[8] T. Hehn, O. Milenkovic, S. Laendner and J. B. Huber, "Permutation Decoding and the Stopping Redundancy Hierarchy of Cyclic and Extended Cyclic Codes," *IEEE Trans. Inform. Theory*, vol. 54, no. 12, pp. 5308-5331, Dec. 2008.

[9] T. Hehn, J. B. Huber, O. Milenkovic, and S. Laendner, "Multiple-bases belief-propagation decoding of high-density cyclic codes," *IEEE Trans. Commun.*, vol. 58, no. 1, pp. 1-8, Jan. 2010.

[10] M. Bossert, "An iterative hard and soft decision decoding algorithm for cyclic codes," *in Proc. 12th Int. ITG Conf. Syst. Commun. Coding (SCC)*, Rostock, Germany, pp. 263-268, Feb. 2019.

[11] M. Bossert, "On decoding using codewords of the dual code," *arXiv preprint:2001.02956*, Jan. 2020.

[12] J. Xing, M. Bossert, S. Bitzer and L. Chen, "Iterative decoding of non-binary cyclic codes using minimum-weight dual codewords," in *Proc. IEEE Int. Symp. Inform. Theory (ISIT)*, CA, U.S.A, pp. 333-337, June. 2020.

[13] G. Pólya and R. Read, *Combinatorial enumeration of groups, graphs, and chemical compounds*. New York, NY, USA: Springer, 1987.

[14] Grinstead, Charles Miller, and James Laurie Snell, *Introduction to probability*. American Mathematical Soc., 2012.