Constant Weight Codes for Correcting Symmetric Errors and Detecting Unidirectional Errors

Mao-Chao Lin, Member, IEEE

Abstract—We propose two classes of constant weight codes, which can be used for correcting $t$ symmetric errors and simultaneously detecting all unidirectional errors. Codes in the first class are in systematic form and codes in the second class are in quasi-systematic form. Since each codeword of codes in both classes can be divided into a data part and a parity check part, the proposed codes have the merit of easily mapping messages into codewords.

Index Terms—Constant weight codes, error-control coding, symmetric errors, $t$-EC/AUED codes, unidirectional errors.

I. INTRODUCTION

Error control coding [1]–[3] consists of important techniques for computer or communication systems to increase the reliability against errors. Different systems may be vulnerable to different types of errors. If the $1 \rightarrow 0$ errors and $0 \rightarrow 1$ errors occur with equal probability in each code bit, we classify this type of error as symmetric. If $1 \rightarrow 0$ errors and $0 \rightarrow 1$ errors are both likely to occur but not simultaneously in each codeword of a system, we classify this type of error as unidirectional. Conventional error detecting/correcting codes are usually designed for combating symmetric errors [1], [2]. It has been reported [4] that the most likely faults in VLSI chips cause unidirectional errors. Therefore, the number of symmetric errors is usually limited while the number of unidirectional errors can be very large. Hence, there are some research works [5]–[10] aiming at designing codes for correcting up to $t$ symmetric errors and detecting all unidirectional errors. These codes are called $t$-EC/AUED codes. Most research works about $t$-EC/AUED codes concentrate on systematic $t$-EC/AUED codes, which are constructed by applying Berger techniques [3], [12] to some systematic linear codes. Systematic codes have the merit of easily mapping messages into codewords. However, some nonsystematic codes also have the merit of easily mapping messages into codewords. For example, let us consider a code in quasi-systematic form, for which each codeword is divided into a data part and a parity check part, where the data part has a one-to-one correspondence with the message to be encoded. Clearly, mapping messages into codewords is easy if the size of message space is not large. In case the size of message space is very large, we can make a small compromise by slightly increasing the code length to achieve a much simplified implementation of mapping messages into codewords, which will be illustrated in Example 7 of this paper. Bose and Rao [5] have shown that constant weight codes with minimum distance $2t + 2$ are $t$-EC/AUED codes. In [5], Bose and Rao introduced a class of constant weight codes with minimum distance 4, which are in quasi-systematic form. In [13]–[15], some classes of quasi-systematic constant weight codes with distances of 6, 8, and 10 were proposed. In this paper, we present two classes of constant weight codes as $t$-EC/AUED codes. Codes in the first class are in quasi-systematic form, which are generalized from the Bose and Rao codes and are modified from the codes in [13]–[15]. Codes in the first class can be designed to have minimum distances of $2t$, where $t \geq 2$. Codes in the second class are in systematic form, which are constructed by combining the technique of designing the Bose and Rao codes and the technique of designing the systematic constant weight codes presented by Boinck and van Tilborg in [10]. Codes in the second class have the minimum distances of $4t + 2$, where $t \geq 1$. Some specific codes proposed in this paper are very efficient in the coding rate. For all the codes presented in this paper, we also design an efficient decoding algorithm based on the well-developed technique of decoding binary BCH codes.

II. QUASI-SYSTEMATIC CONSTANT WEIGHT CODES PROPOSED BY BOSE AND RAO

In this section, we describe the quasi-systematic constant weight codes of distance 4 proposed by Bose and Rao in [5]. Let $n$ be an integer. There are $\binom{n}{l}$ $n$-tuples of weight $l$, which can represent $\binom{n}{l}$ distinct messages. Let $N$ be the set of all the $n$-tuples of weight $l$ and $M$ be the set of $\binom{n}{l}$ messages. Let

$$f : M \rightarrow N$$

represent a one-to-one mapping. A quasi-systematic constant weight code $C$ of length $(n + p)$ which has $\binom{n}{l}$ codewords can be constructed by appending a $p$-tuple of weight $r$ to each of the $\binom{n}{l}$-tuples of weight $l$, which is firstly encoded from one of the $\binom{n}{l}$ messages. The $n$-tuple and $p$-tuple of a codeword are called the data part and the parity check part, respectively. The parity check part of a codeword is designed to maximize the minimum distance between $(n + l)$-tuples. Each codeword $\overline{c}$ of $C$ can be written in the form $[\overline{a}|\overline{b}]$, where $\overline{a} = (a_0, a_1, \ldots, a_{n-1}) \in N$ and $\overline{b} = (b_0, b_1, \ldots, b_{p-1}), a_i \in \mathbb{GF}(2)$.
Let \( \{a_0 = 0, a_1, a_2, \ldots, a_{n-1}\} \) be an abelian additive group. We mark the \( n \) positions of the data part by \( a_0, a_1, a_2, \ldots, a_{n-1} \). The code \( C \) can be designed to have minimum distance 4 by designing the parity check part as follows. We require \( \lceil \frac{n}{2} \rceil \geq n \). Then, we can map each of \( a_i, i = 0, 1, \ldots, n - 1 \), into a unique \( p \)-tuple of weight \( r \). We use
\[
g : \{a_0, a_1, a_2, \ldots, a_{n-1}\} \rightarrow P
\]
to represent the one-to-one mapping, where \( P \) is a subset of the set of all the \( p \)-tuples of weight \( r \) and the cardinality of \( P \) is \( n \). Suppose a message \( m \) is encoded into the data part \( a \in N \). We calculate
\[
\sum_{i=0}^{n-1} a_i \cdot a_i = \alpha_j
\]
for some \( j \in \{0, 1, \ldots, n - 1\} \), where
\[
a_i \cdot a_i = \begin{cases} a_i, & \text{for } a_i = 1, \\ 0, & \text{for } a_i = 0.\end{cases}
\]
The obtained \( \alpha_j \) is then mapped into the associated \( p \)-tuple \( b \) in \( P \), i.e., \( b = g(\alpha_j) \). Then, we have the codeword \( \tilde{b} = [a | b] \) encoded from \( m \). Let us check the distance property.

Let \( \tilde{c} = [b[a]b^*] \) be the codeword encoded from \( m' \neq m \). In case that \( b \neq b^* \), we have \( \tilde{a} \neq \tilde{a}' \). Let \( d(\tilde{b}, \tilde{b}') \) denote the Hamming distance between two vectors \( \tilde{b} \) and \( \tilde{b}' \). It is clear that \( d(\tilde{b}, \tilde{a}) \geq 2 \) and \( d(b, b^*) \geq 2 \). Hence, we have \( d(\tilde{c}, \tilde{c}') \geq 4 \).

In case that \( b = b^* \), we have
\[
\sum_{i=0}^{n-1} a_i \cdot a_i = \sum_{i=0}^{n-1} a_i^2 = \alpha_i
\]
Suppose that \( d(\tilde{a}, \tilde{a}') = 2 \), where at position \( j \) and \( k \) we have \( \alpha_j = 1, \alpha_k = 0 \) and \( a_j = 0, a_k = 1 \). This results in that \( \alpha_j = \alpha_k \) for \( j \neq k \), which is impossible. Thus, the distance between \( \tilde{a} \) and \( \tilde{a}' \) is at least 4. Hence, we have \( d(\tilde{c}, \tilde{c}') \geq 4 \) for all the distinct codewords of \( C \).

The decoding procedure for error correction can be easily implemented as described in the following. Let \( w(\tilde{c}) \) represent the weight of the vector \( \tilde{c} \). Let \( \tilde{c}^* = [a[a]a^*] \) be the possibly error-affected codeword. We decode \( \tilde{c}^* \) according to the following conditions:

i) \( w(\tilde{c}^*) = 0 \): Find \( g(\Sigma_{i=0}^{n-1} a_i^* \cdot a_i) = b^* \). If \( d(\tilde{b}^*, b^*) \leq 1 \), the corresponding message \( m = f^{-1}(\tilde{a}^*) \). Otherwise, an uncorrectable error pattern is detected.

ii) \( w(\tilde{c}^*) \in \{1, 2, 3, 4\} \) and \( w(\tilde{c}^*) = r \): We map \( b^* \) back to find \( g^{-1}(b^*) \). If \( g^{-1}(b^*) \) does not exit, we detect an uncorrectable error pattern in \( c^* \). Otherwise, we calculate \( S = (\Sigma_{i=0}^{n-1} a_i^* \cdot a_i) - g^{-1}(b^*) \) if \( w(\tilde{c}^*) = 1 \) and \( S = -(\Sigma_{i=0}^{n-1} a_i^* \cdot a_i) - g^{-1}(b^*) \) if \( w(\tilde{c}^*) = 1 \). Then, \( S \) indicates the location of the single error in \( a^* \).

iii) For any condition not in i and ii, we detect an uncorrectable error pattern in \( c^* \).

In [5], it has been shown that any constant weight code with minimum distance of 2\( r + 2 \) is a \( 1 \)-EC/AUED code. Using the above decoding algorithm, we are able to correct any single error in \( [\tilde{a}[^*]b^*] \) and detect other uncorrectable error patterns. Hence, the detectable but uncorrectable error patterns include all the unidirectional errors. Note that the detectable but uncorrectable error patterns also include all the error patterns of two symmetric errors. Therefore, this class of quasi-systematic constant weight codes with distance 4 proposed by Bose and Rao can be used as a 1-EC/AUED code. This class of quasi-systematic constant weight codes with distance 4 proposed by Bose and Rao can be used as 1-EC/AUED codes. This class of quasi-systematic 1-EC/AUED codes, which can be easily encoded and decoded, sometimes are more efficient than the best systematic 1-EC/AUED codes. For example, by taking \( n = 10, l = 5, p = 5 \), and \( r = 2 \), we have a 1-EC/AUED code of length 15 and \( [M] = 252 \). Let \( k = \lceil \log_2 252 \rceil = 7 \). The best known systematic 1-EC/AUED code with a 7-bit message has code length of 16 [10].

### III. Construction I: Quasi-Systematic Constant Weight Codes with Distance More than 4

The codes constructed in this section are generalized from the Bose and Rao codes in [5] and are modified from the codes in [13]-[15], which can be further divided into three subclasses, denoted by Construction I-B, Construction I-C and Construction I-D, respectively. Codes from Construction I-B have a minimum distance of 8 and codes from Construction I-C have a minimum distance of 10. Codes from Construction I-D have minimum distance of 2\( t \), where \( t \geq 2 \). For the convenience of illustration, we take the special case of Construction I-D with \( t = 3 \) as Construction I-A.

#### A. Construction I-A: Codes with Distance 6

Let \( C \) be a binary code of length \( n + 2p \) for which each codeword is of the form \( \tilde{c} = \tilde{a} | \tilde{b} \), where \( \tilde{a} \) is an \( n \)-tuple of weight \( l \) and \( \tilde{b} \) are both \( p \)-tuples of weight \( r \). We require that \( n \leq 2^p - 1 \) for some integer \( \mu \). Let \( \alpha \) be a primitive element of \( GF(2^p) \). We mark the \( \mu \) locations of \( \tilde{a} \) by \( a_0, a_1, a_2, \ldots, a_{n-1} \) and \( \tilde{b} = h(\Sigma_{i=0}^{n-1} a_i \cdot \alpha^j) \). We require that \( \lceil \mu \rceil \geq 2^p \). Let \( h \) be a one-to-one mapping from \( GF(2^p) \) to \( \tilde{P} \), i.e.,
\[
h : GF(2^p) \rightarrow \tilde{P},
\]
where \( \tilde{P} \) is a subset of the set of all the \( p \)-tuples of weight \( r \) and the cardinality of \( \tilde{P} \) is \( 2^p \). For each message \( m \) from \( M \), we have \( \tilde{a} = f(\tilde{m}) \), \( \tilde{b} = h(\Sigma_{i=0}^{n-1} a_i \cdot \alpha^j) \). Hence, \( C \) is a constant weight code in quasi-systematic form.

The distance property of \( C \) can be examined as follows. Let \( \tilde{c} = [\tilde{a} | \tilde{b}] \) be another codeword of \( C \). In case that \( \tilde{a} = \tilde{a}' \) and \( \tilde{b} = \tilde{b}' \), we have \( \Sigma_{i=0}^{n-1} (a_i - a'_i) \cdot \alpha^j = 0 \) and \( \Sigma_{i=0}^{n-1} (a_i - a'_i) \cdot (\alpha^j) = 0 \). Note that \( \Sigma_{i=0}^{n-1} (a_i - a'_i) \cdot 0 = 0 \). Hence, we have \( \Sigma_{i=0}^{n-1} (a_i - a'_i) \cdot (\alpha^j) = 0 \) for \( -2 \leq j \leq 2 \).

It follows from the BCH bound that the number of positions with nonzero \( a_i - a'_i \) is at least 6. In case that \( \tilde{a} = \tilde{a}' \) and \( \tilde{b} = \tilde{b}' \), we have \( d(\tilde{a}, \tilde{a}') \geq 4 \) from the BCH bound. We also have \( d(\tilde{a}, \tilde{b}) \geq 2 \). Therefore, \( d(\tilde{c}, \tilde{c}') \geq 6 \). Hence, \( d(\tilde{c}, \tilde{c}') \geq 2 \). Therefore, \( C \) is a quasi-systematic constant weight code with distance 6.
weight code of length \( n + 2p \), weight \( l + 2r \) and minimum distance 6.

Let \( \overline{a} = [\overline{a}^* | \overline{a}^**] \) be the vector to be decoded, which is possibly error-corrupted. We first compute \( S_0 = (\Sigma_{i=0}^{l-1} a_i) \mod 2 \), \( S_1 = (\Sigma_{i=0}^{l-1} a_i \cdot \alpha^i) - h^{-1}(\overline{b}^*) \) and \( S_2 = (\Sigma_{i=0}^{l-1} a_i \cdot (\alpha^{-1})^i) - h^{-1}(\overline{w}^*) \).

1. We check if \( S_2 = 0 \). If so, we perform the syndrome calculation to remove errors in \( \overline{a}^* \) and obtain the associated message \( \overline{a}^* \).

2. In step 1 an uncorrectable error condition in \( [\overline{a}^* | \overline{a}^**] \) is detected and in step 2 \( \overline{a}^** \) is decoded: The message \( \overline{a}^** = f^{-1}(\overline{a}^*) \).

3. After steps 1 and 2 are completed, we have four distinct conditions:
   a) In step 1, \( \overline{a}^** \) is decoded and in step 2 an uncorrectable error condition in \([\overline{a}^* | \overline{a}^**]\) is detected: The message \( \overline{a}^** \) is equal to \( \overline{a}^** \).
   b) In step 1 an uncorrectable error condition in \([\overline{a}^* | \overline{a}^**]\) is detected and in step 2 \( \overline{a}^** \) is decoded: The message \( \overline{a}^** \) is equal to \( \overline{a}^** \).
   c) In step 1 \( \overline{a}^** \) is decoded and in step 2 \( \overline{a}^** \) is decoded: If \( \overline{a}^** = \overline{a}^** \), then the message \( \overline{a}^** \) is equal to \( \overline{a}^** \). If \( \overline{a}^** \neq \overline{a}^** \) then an uncorrectable error condition in \( \overline{a}^* \) is detected.
   d) Uncorrectable error patterns are detected for both steps 1 and 2: Suppose that there are two symmetric errors both in \([\overline{a}^* | \overline{a}^**]\) and \([\overline{a}^* | \overline{a}^**]\). The two errors must be in \( \overline{a}^* \). We use \( S_2, S_1, S_0 \) and the decoding algorithm for 2-error-correcting BCH codes such as the famous Berlekamp-Massey algorithm to obtain the two errors in \( \overline{a}^* \) and obtain the associated message \( \overline{a}^** \). If the decoding algorithm for BCH codes cannot be processed, then an uncorrectable error condition in \( \overline{a}^* \) is detected.

4. Encode \( \overline{a}^** \) into \( \overline{c}^** = [\overline{a}^** | \overline{a}^***] \).
   If \( d(\overline{c}^*, \overline{c}^**) > 2 \), there are uncorrectable errors in \( \overline{c}^* \). Otherwise \( \overline{a}^** \) is the final decoded message.

5. For any condition not applicable to the previous steps, we detect an uncorrectable error pattern in \( \overline{a}^* \).

Using this decoding algorithm for \( C \), we can correct up to 2 symmetric errors and detect all unidirectional errors. Moreover, the detectable but uncorrectable error patterns include all the error patterns of 3 symmetric errors.

**Example 1:** Let \( n = 15, p = 6, l = 7 \), and \( r = 3 \). The code \( C \) is a quasi-systematic constant weight code of length 27, weight 13, and minimum distance 6. The size of the message space \( M \) is \( |M| = \binom{15}{7} = 6435 \). Let \( k = \lceil \log_2 |M| \rceil \). This code can represent all the \( k \)-bit messages, where \( k = 12 \). This code is a 2-EC/AUED code.

# Table I

<table>
<thead>
<tr>
<th>( k )</th>
<th>( n' )</th>
<th>( n' )</th>
<th>( n )</th>
<th>( n' )</th>
<th>( n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>16</td>
<td>14</td>
<td>20</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>17</td>
<td>16</td>
<td>21</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>6</td>
<td>20</td>
<td>19</td>
<td>22</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>7</td>
<td>22</td>
<td>21</td>
<td>23</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>8</td>
<td>23</td>
<td>23</td>
<td>26</td>
<td>16</td>
<td>16</td>
</tr>
<tr>
<td>9</td>
<td>24</td>
<td>25</td>
<td>27</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>10</td>
<td>25</td>
<td>28</td>
<td>29</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>11</td>
<td>26</td>
<td>30</td>
<td>31</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>12</td>
<td>27</td>
<td>32</td>
<td>32</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>13</td>
<td>30</td>
<td>34</td>
<td>34</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>14</td>
<td>31</td>
<td>36</td>
<td>35</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>20</td>
<td>37</td>
<td>49</td>
<td>41</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>22</td>
<td>39</td>
<td>53</td>
<td>46</td>
<td>33</td>
<td>33</td>
</tr>
<tr>
<td>23</td>
<td>40</td>
<td>55</td>
<td>48</td>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>26</td>
<td>43</td>
<td>61</td>
<td>51</td>
<td>38</td>
<td>38</td>
</tr>
<tr>
<td>27</td>
<td>44</td>
<td>63</td>
<td>52</td>
<td>39</td>
<td>39</td>
</tr>
<tr>
<td>28</td>
<td>45</td>
<td>65</td>
<td>54</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>50</td>
<td>70</td>
<td>110</td>
<td>76</td>
<td>62</td>
<td>62</td>
</tr>
</tbody>
</table>

**Table I**

2-EC/AUED Codes

- \( k \): number of message bits carried by each codeword.
- \( n' \): length of the \( r \)-ED/AUED code.
- \( n \): length of the base code.

**B. Construction I-B: Codes with Distance 8**

We now construct a quasi-systematic constant weight code \( C \) of distance 8. It is required that \( 2^{n} - 1 \) and \( n \) to be relatively prime and \( n \leq 2^{n} - 1 \). Each codeword \( \overline{a} \) of \( C \) is divided into five parts, \( \overline{a}, \overline{b}, \overline{w}, \overline{z}, \overline{e} \), where \( \overline{a} \) is an \( n \)-tuple of weight \( 1 \), \( \overline{b}, \overline{w}, \overline{z} \) are \( p \)-tuples of weight \( r \), and \( \overline{e} \) is a \( q \)-tuple of weight \( s \). Then, \( \overline{a} = [\overline{b} | \overline{w} | \overline{z} | \overline{e}] \). We require that \( \binom{n}{r} \geq p \).

Let \( \{ \beta_0 = 0, \beta_1, \beta_2, \ldots, \beta_{p-1} \} \) represent an abelian additive group. Write \( \overline{w} = (w_0, w_1, \ldots, w_{p-1}) \) and mark the \( p \) positions of \( \overline{e} \) by \( \beta_0, \beta_1, \beta_2, \ldots, \beta_{p-1} \), respectively. We map each of \( \beta_0, i = 0, 1, \ldots, p-1 \) into a unique \( q \)-tuple of weight \( s \). The mapping is

\[ \zeta : \{ \beta_0, \beta_1, \beta_2, \ldots, \beta_{p-1} \} \rightarrow Q \]

where \( Q \) is the subset of the set of all the \( q \)-tuples of weight \( s \), and the cardinality of \( Q \) is \( p \). For a message \( \overline{m} \in M \), we first encode it into \( \overline{w} = f(\overline{m}) \). Then, \( \overline{w} = h(\Sigma_{i=0}^{p-1} \alpha_i \cdot (\alpha^{-1})^i), \overline{w} = h(\Sigma_{i=0}^{p-1} \alpha_i \cdot \alpha^i) \), and \( \overline{e} = h(\Sigma_{i=0}^{p-1} \alpha_i \cdot (\alpha^3)^i) \). Furthermore, \( \overline{e} = \zeta(\Sigma_{i=0}^{p-1} v_i \cdot \beta_i) \), where the definition of \( v_i \cdot \beta_i \) is similar to that for \( \alpha_i \cdot \alpha \) described in Section II.
Let $\overline{c} = [\overline{w}^*|\overline{a}^*|\overline{v}^*|\overline{z}^*]$ be a codeword of $C$ different from $\overline{c}$. Let us check the following conditions to see the distance property of $C$:

1) $\overline{w} = \overline{w}', \overline{v} = \overline{v}', \overline{x} = \overline{x}'$: The BCH bound shows that $d(\overline{a}, \overline{a}') \geq 8$. 

2) $\overline{w} = \overline{w}', \overline{v} = \overline{v}', \overline{x} \neq \overline{x}'$: The BCH bound shows that $d(\overline{a}, \overline{a}') \geq 6$. Combining the fact that $d(\overline{x}, \overline{x}') \geq 2$, we have $d(\overline{a}, \overline{a}') \geq 8$.

3) $\overline{w} = \overline{w}', \overline{v} \neq \overline{v}'$: The BCH bound shows that $d(\overline{a}, \overline{a}') \geq 4$. If $\overline{x} = \overline{x}'$, then $d(\overline{v}, \overline{v}') \geq 4$ from the argument in Section II. If $\overline{x} \neq \overline{x}'$, we have $d(\overline{v}, \overline{v}') \geq 2$. For either case, we have $d(\overline{a}, \overline{a}') \geq 8$.

4) $\overline{w} \neq \overline{w}', \overline{v} = \overline{v}', \overline{x} \neq \overline{x}'$: The BCH bound shows that $d(\overline{a}, \overline{a}') \geq 6$. Since $d(\overline{w}, \overline{w}') \geq 2$, we have $d(\overline{a}, \overline{a}') \geq 8$.

5) $\overline{w} \neq \overline{w}', \overline{v} = \overline{v}', \overline{x} \neq \overline{x}'$: The BCH bound shows that $d(\overline{a}, \overline{a}') \geq 4$. Since $d(\overline{w}, \overline{w}') \geq 2$ and $d(\overline{x}, \overline{x}') \geq 2$, we have $d(\overline{a}, \overline{a}') \geq 8$.

6) $\overline{w} \neq \overline{w}', \overline{v} = \overline{v}', \overline{x} \neq \overline{x}':$ The BCH bound shows that $d(\overline{a}, \overline{a}') \geq 4$. Since $d(\overline{w}, \overline{w}') \geq 2$ and $d(\overline{v}, \overline{v}') \geq 2$, we have $d(\overline{a}, \overline{a}') \geq 8$.

7) $\overline{w} \neq \overline{w}', \overline{v} \neq \overline{v}', \overline{x} \neq \overline{x}':$ Clearly, we have $d(\overline{a}, \overline{a}') \geq 8$ from observing that $d(\overline{a}, \overline{a}') \geq 2, d(\overline{w}, \overline{w}') \geq 2, d(\overline{v}, \overline{v}') \geq 2, \text{ and } d(\overline{x}, \overline{x}') \geq 2$.

Combining 1-7, we see that $C$ has minimum distance of at least 8. Clearly, $C$ is a quasi-systematic constant weight code of length $n + 3p + q$ and weight $l + 3r + s$. This code can be used as a 3-EC/AUED code.

Let $\overline{a}^* = [\overline{a}_1^*|\overline{a}_2^*|\overline{a}_3^*|\overline{a}_4^*]$ be an error-corrupted codeword to be decoded. We compute $S_{-1} = (\sum_{i=0}^{n-1} \alpha_i^* \cdot (\alpha^{-1})^i) - h^{-1}(\overline{w}^*), S_0 = (\sum_{i=0}^{n-1} \alpha_i^* \cdot (\alpha^{-1})^i - h^{-1}(\overline{w}^*))$ and $S_3 = (\sum_{i=0}^{n-1} \alpha_i^* \cdot (\alpha^{-1})^i - h^{-1}(\overline{w}^*))$. If $h^{-1}(\overline{w}^*), h^{-1}(\overline{w}^*), h^{-1}(\overline{w}^*)$ exist. Then, the decoding of $C$ can be implemented as follows:

1) We decode $[\overline{w}^*|\overline{a}^*|\overline{v}^*|\overline{z}^*]$ according to the decoding technique for the quasi-systematic constant weight codes of Construction I-A with distance 6. If the number of symmetric errors in $[\overline{w}^*|\overline{a}^*|\overline{v}^*|\overline{z}^*]$ is no more than 2, the errors can be removed. Let $\overline{a}^*$ be the n-tuple corresponding to $\overline{a}^*$ obtained after decoding. The temporarily decoded message $\overline{m}_{1}^* = f^{-1}(\overline{a}^*)$. If there are three symmetric errors or unidirectional errors in $[\overline{w}^*|\overline{a}^*|\overline{v}^*|\overline{z}^*]$, the decoder can detect this condition.

2) We decode $[\overline{w}^*|\overline{v}^*|\overline{z}^*]$ with a procedure similar to that in step 1 and have a temporarily decoded message $\overline{m}_{2}^*$. If there are three symmetric errors or unidirectional errors in $[\overline{w}^*|\overline{v}^*|\overline{z}^*]$, the decoder can detect this condition.

3) Consider the following conditions:

4. If $\overline{a}^* = \overline{a}_1^*|\overline{a}_2^*|\overline{a}_3^*|\overline{a}_4^*$ is decoded and in step 2 an uncorrectable error condition in $[\overline{w}^*|\overline{a}^*|\overline{v}^*|\overline{z}^*]$ is detected: The message $\overline{m}_{3}^*$ is equal to $\overline{m}_{1}^*$.

b) In step 1, an uncorrectable error condition in $[\overline{w}^*|\overline{v}^*|\overline{z}^*]$ is detected and in 2) $\overline{m}_{2}^*$ is decoded: The message $\overline{m}_{3}^*$ is equal to $\overline{m}_{2}^*$.

c) In step 1 $\overline{m}_{1}^*$ is decoded and in step 2, $\overline{m}_{2}^* = \overline{m}_{3}^*$. Then the message $\overline{m}_{3}^*$ is equal to $\overline{m}_{1}^*$. If $\overline{m}_{1}^* \neq \overline{m}_{2}^*$, then an uncorrectable error condition in $\overline{w}^*$ is detected.

d) An uncorrectable error condition in $[\overline{w}^*|\overline{a}^*|\overline{v}^*|\overline{z}^*]$ and an uncorrectable condition in $[\overline{w}^*|\overline{v}^*|\overline{z}^*]$ are detected in step 1 and 2, respectively: Suppose that there are three symmetric errors in both $[\overline{w}^*|\overline{a}^*|\overline{v}^*|\overline{z}^*]$ and $[\overline{a}^*|\overline{v}^*|\overline{z}^*]$. Then the distribution of the three symmetric errors can be divided into four conditions: i) three errors in $\overline{a}^*$ and no error in $\overline{v}^*$; ii) two errors in $\overline{a}^*$ and a single error in $\overline{v}^*$; iii) one error in $\overline{a}^*$ and two errors in $\overline{v}^*$; iv) no error in $\overline{a}^*$ and three errors in $\overline{v}^*$. In all of i)–iv), there is no error in $\overline{w}^*$ and $\overline{z}^*$. If $S_0 = 0$ and $S_{-1} = 0$, we must be in condition iv. Then $\overline{a}^*$ is correct. Otherwise we are in conditions i, ii, and iii. We decode $[\overline{v}^*|\overline{z}^*]$ using the technique in Section II. In case we detect two errors in $[\overline{v}^*|\overline{z}^*]$, we are in condition iii. Then, the location of the single error in $\overline{a}^*$ is $(S_{-1})^{-1}$. In case of i and ii, there is at most a single error in $\overline{v}^*$, which can be removed in the decoding of $[\overline{w}^*|\overline{z}^*]$. Let $\overline{v}^*$ be the p-tuple corresponding to $\overline{v}^*$ obtained after the decoding of $[\overline{w}^*|\overline{z}^*]$. We calculate $S_1 = (\sum_{i=0}^{n-1} \alpha_i^* \cdot (\alpha^{-1})^i)$ and $S_3 = (\sum_{i=0}^{n-1} \alpha_i^* \cdot (\alpha^{-1})^i)$. If $S_0 = 0$ and $S_{-1} = 0$, we must be in condition iv. Then $\overline{a}^*$ is correct. Otherwise, $\overline{m}_{3}^*$ is the finally decoded message.
Example 2: Let \( n = 26, p = 7, q = 5, l = 13, r = 3, \) and \( s = 2. \) Then we have a constant weight code of length \( n' = n + 3p + q = 52, \) weight 24 and distance 8. The size of the message space \( M \) is \( |M| = \binom{26}{5}. \) Since \( |\log_2 |M| | = 23, \) we see that this code can represent all the 23-bit messages. This code is a quasi-systematic 3-EC/AUED code.

C. Construction I-C: Codes with Distance 10

In case that \( n \leq 2^{n'} - 1 \) with \( 2^n - 1 \) and 3 relatively prime, a constant weight code of distance 10 can be constructed by encoding each message \( m \) into the codeword \( \overline{c} = (\overline{w}_1|\overline{w}_2|\overline{w}_3|\overline{w}_4|\overline{w}_5), \) where \( \overline{a} = f(m), \overline{y} = h(\sum_{i=0}^{5} a_i \cdot (\alpha^3)^i), \overline{v} = h(\sum_{i=0}^{5} a_i \cdot (\alpha^2)^i), \overline{x} = h(\sum_{i=0}^{5} a_i \cdot (\alpha^{32})^i), \overline{z} = h(\sum_{i=0}^{5} a_i \cdot (\alpha^{33})^i). \) The decoding is to first decode \( [w_1|w_2|w_3|w_4|w_5] \) and \( v_j|v_j|w_j, \) respectively, and then check the conditions for which four errors are distributed in \( [w_1|w_2|w_3|w_4|w_5]. \)

Example 3: Let \( n = 26, p = 7, q = 5, l = 13, r = 3, \) and \( s = 2. \) Then we have a constant weight code of length \( n' = n + 4p + 2q = 64, \) weight 29, and distance 10. This code can represent all the 23-bit messages and is a quasi-systematic 4-EC/AUED code.

D. Construction I-D: Codes Generalized from Construction I-A

Now we represent the construction of quasi-systematic constant weight code with distance more than 4, which are in a generalized form of the Bose and Rao codes described in Section II and the codes of Construction I-A. The construction can be divided into the cases of distance \( 4t + 2 \) and \( 4t. \)

1) Codes with Distance \( 4t + 2 \): Let \( C \) be a quasi-systematic code for which each codeword has the form \( \overline{c} = (w_1|w_2|w_3|w_4|w_5), \) where \( \overline{a} \) is the data part and \( [w_1|w_2|w_3|w_4|w_5] \) is the parity check part. The data part \( \overline{a} \) is an \( n \)-tuple of weight \( l \) and hence the size of message space \( M \) for \( C \) is \( \binom{2^n}{l}. \) For \( i = 1, \ldots, t, \) \( \overline{w}_i \) is a \( p_i \)-tuple of weight \( r_i, \) and \( \overline{w}_i \) is a \( p_i \)-tuple of weight \( r_i. \) The vectors \( \overline{w}_i \) and \( \overline{w}_i \) in binary form are \( [w_{i,0}|w_{i,1}|\ldots|w_{i,n-1}] \) and \( [w_{i,0}|w_{i,1}|\ldots|w_{i,n-1}] \), respectively. The requirements for designing the parity check part are listed as follows:

1) Let \( \mu_1 \) be the integer such that \( 2^{\mu_1-1} - 1 < n \leq 2^{\mu_1}. \) Let \( \alpha \) be a primitive element in the field of \( GF(2^{\mu_1}). \) We mark the positions of \( \alpha, \alpha^2, \ldots, \alpha^{n-1}. \) Suppose that we need \( \mu_1, \mu_2, \ldots, \mu_t \) to represent \( \alpha, \alpha^2, \ldots, \alpha^{2t}. \) Let \( \mu = \mu_1 + \mu_2 + \ldots + \mu_t. \) We require that \( \mu_t \geq 2^t. \) Let \( P_t \) be the subset of the set of all \( p_t \)-tuples of weight \( r_t \) with cardinality \( 2^{\mu_1}. \) We need a one-to-one mapping

\[ g_t: \text{The set of all } \mu_t \text{-tuples } \rightarrow P_t. \]

2) For \( 1 \leq j < t, \) the design of \( \overline{w}_{t-j+1} \) and \( \overline{w}_{t-j+1} \) has the following requirement. Let \( \overline{v}_j \) be the integer such that \( 2^{\overline{v}_j-1} - 1 \leq p_{t-j+1} \leq 2^{\overline{v}_j-1} - 1. \) Let \( \beta_j \) be a primitive element in the field of \( GF(2^{\overline{v}_j}). \) We mark the \( p_{t-j+1} \) positions of \( \overline{w}_{t-j+1} \) and \( \overline{w}_{t-j+1} \) by \( \beta_j, \beta_j, \ldots, \beta_j, \beta_j. \) Suppose that we need \( \nu_{t-j+1} \) to represent \( \beta_j, \beta_j, \ldots, \beta_j, \beta_j. \) Let \( \nu_j = \nu_{t-j+1} + \nu_{t-j+2} + \ldots + \nu_{t-j}. \) We require that \( \binom{p_{t-j+1}}{p_{t-j}} \geq 2^{\nu_j}. \) Let \( P_{t-j} \) be the subset of the set of all \( p_{t-j} \)-tuples of weight \( r_{t-j} \) with cardinality \( 2^{\nu_j}. \)
for $-2t \leq j \leq 2t$. From the BCH bound, we see that $d(\overline{a}, \overline{a}) \geq 4t+2$. Hence, $d(\overline{G}, \overline{G}) \geq 4t+2$.

2) Suppose that $\overline{v}_i = \overline{v}_i', \overline{w}_{t-i} \neq \overline{w}_{t-i}'$ for $0 \leq i < j$ and $\overline{w}_{t-j} = \overline{w}_{t-j}'$. It can be easily checked that $d(\overline{a}, \overline{a}) \geq 2t + 2$ from the BCH bound. Clearly, $d(\overline{G}, \overline{G}) \geq 2$ for $0 \leq i < j$. Furthermore, $\overline{w}_{t-j} = \overline{w}_{t-j}'$ implies $\overline{y}_{t-j+1} = \overline{g}_{t-j}(\overline{w}_{t-j}) = \overline{g}_{t-j}(\overline{w}_{t-j}) = \overline{w}_{t-j+1}$. It follows that $d(\overline{a}, \overline{a}) \geq 2t + 2$. Hence, we see that $d(\overline{G}, \overline{G}) \geq 2t + 2$. If $\overline{w}_{t-j} = \overline{w}_{t-j}'$, the design of $G$ can be modified to be a mapping from the set of $\overline{w}_{t-j}$.

3) Suppose that $\overline{w}_i = \overline{w}_i', \overline{w}_{t-i} \neq \overline{w}_{t-i}'$ for $0 \leq i < j$ and $\overline{w}_{t-j} = \overline{w}_{t-j}'$. We then have the correct value of $\overline{w}_{t-j}$.

4) Suppose that $\overline{w}_{t-j} = \overline{G}_{t-j}'$ for $0 \leq i < j$, and $\overline{w}_{t-j} \neq \overline{w}_{t-j}'$. It is seen that $d(\overline{a}, \overline{a}) \geq 2t + 2$. Hence, $d(\overline{G}, \overline{G}) \geq 2t + 2$. Moreover, it follows from the BCH bound that $d(\overline{w}_{t-j}, \overline{w}_{t-j+1}) \geq 2t + 2$ and $d(\overline{w}_{t-j+1}, \overline{w}_{t-j+2}) \geq 2t + 2$. Hence, $d(\overline{G}, \overline{G}) \geq 2t + 2$. If $\overline{w}_{t-j+1} = \overline{w}_{t-j+1}'$, we have the correct value of $\overline{w}_{t-j+1}$.

5) Suppose that $\overline{w}_{t-j} \neq \overline{w}_{t-j}'$ for $0 \leq i < j$, and $\overline{w}_{t-j} \neq \overline{w}_{t-j}'$. It is seen that $d(\overline{a}, \overline{a}) \geq 2t + 2$. Hence, $d(\overline{G}, \overline{G}) \geq 2t + 2$. If $\overline{w}_{t-j+1} = \overline{w}_{t-j+1}'$, we have the correct value of $\overline{w}_{t-j+1}$.

6) Suppose that $\overline{w}_{t-j} \neq \overline{w}_{t-j}'$ for $0 \leq i < j$, and $\overline{w}_{t-j} \neq \overline{w}_{t-j}'$. It is seen that $d(\overline{a}, \overline{a}) \geq 2t + 2$. Hence, $d(\overline{G}, \overline{G}) \geq 2t + 2$. If $\overline{w}_{t-j+1} = \overline{w}_{t-j+1}'$, we have the correct value of $\overline{w}_{t-j+1}$.

7) We decode $\overline{a}$ into $\overline{a}^*$ using the decoding algorithm for a $t$-error-correcting BCH code. Encode $\overline{a}^*$ into $\overline{e}^* = [\overline{e}_1^* \cdots \overline{e}_i^* \cdots \overline{e}_n^*]$. If $\overline{d}(\overline{e}, \overline{e}^*) \leq 2t$, then $\overline{a}^*$ is correctly decoded; otherwise we proceed to step 8.

8) We decode $\overline{a}^*$ into $\overline{a}^*$ using the decoding algorithm for a $t$-error-correcting BCH code. Encode $\overline{a}^*$ into $\overline{e}^* = [\overline{e}_1^* \cdots \overline{e}_i^* \cdots \overline{e}_n^*]$. If $\overline{d}(\overline{e}, \overline{e}^*) \leq 2t$, then $\overline{a}^*$ is correctly decoded; otherwise we proceed to step 9.

9) We decode $\overline{a}^*$ into $\overline{a}^*$ using the decoding algorithm for a $(2t)$-error-correcting BCH code. Encode $\overline{a}^*$ into $\overline{e}^* = [\overline{e}_1^* \cdots \overline{e}_i^* \cdots \overline{e}_n^*]$. If $\overline{d}(\overline{e}, \overline{e}^*) \leq 2t$, then $\overline{a}^*$ is correctly decoded; otherwise we proceed to step 10.

10) Increase $i$ by 1 and proceed to step 2 and complete steps 6 to 9. If $\overline{a}^*$ is still not correctly decoded, we proceed to step 11.

11) Increase $j$ by 1 and proceed to step 3 and complete steps 6–10.

12) In case that we obtain $\overline{a}^*$ which results in the condition that $\overline{d}(\overline{e}, \overline{e}^*) \leq 2t$, we have the decoded message $\overline{m} = f^{-1}(\overline{e}^*)$. Otherwise, we detect an uncorrectable condition.

Assume now the validity of the above decoding procedure. Assume that the number of total symmetric errors in $\overline{a}^*$ is at most $2t$. Denote the number of symmetric errors in $\overline{a}^*$, $\overline{w}_{t-i}, \overline{w}_{t-j}$ by $e, e_{t-i}$, and $f_{t-j}$, respectively. Clearly, if $e = 0$, we can have the correct $\overline{a}^* = \overline{a}$ in step 4 of decoding. Hence, we assume $e \neq 0$ in the following conditions of ii–iv to be checked.

i) For $0 \leq i < t$ and $0 \leq j < t$, all the $e_{t-i}$ and $f_{t-j}$ are nonzero. Since the total number of errors is assumed to be $2t$, the data part $\overline{a}$ is correct.

ii) For $0 \leq i < t$ and $0 \leq j < J$, all the $e_{t-i}$ and $f_{t-j}$ are nonzero and $f_{t-j} = 0$ for some $0 \leq j < t$. Note that in this condition, we have $e \neq f_t + \cdots + f_{t-j+1}$ and $f_{t-j+1} < t$. Thus, each of $e, e_{t-i}, \cdots$, and $e_{t-j-1}$ cannot be greater than $t - J$. Note that $g_{r_{t-j}}(\overline{w}_{t-j}) = \overline{y}_{t-j+1}$ and $\overline{w}_{t-j+1}$ can be combined to correct $t - J$ errors in $\overline{w}_{t-j+1}$. We then have the correct value of $\overline{w}_{t-j+1}$. Then, we use $g_{r_{t-j+1}}(\overline{w}_{t-j+1}) = \overline{y}_{t-j+2}$ and $\overline{w}_{t-j+2}$ to correct errors in $\overline{w}_{t-j+2}$. Finally, we use $g_{r_{t-j+1}}(\overline{w}_{t-j+2}) = \overline{y}_{t-j+3}$ and $\overline{a}$ to correct errors in $\overline{a}$.
iii) For $0 \leq i < I$ and $0 \leq j < t$, all the $e_{t-i}$ and $f_{t-j}$ are nonzero and $e_{t-I} = 0$ for some $0 \leq I < t$. The situation is similar to that for the condition ii.

iv) For $0 \leq i < I$ and $0 \leq j < t$, all the $e_{t-i}$ and $f_{t-j}$ are nonzero and $e_{t-I} = 0$ for some $0 \leq I < t$ and some $0 \leq J < t$. Three cases need to be considered.

a) $e + e_{t+1} + \cdots + e_{t-I+1} \leq t - I$. From the correct $\tilde{\mu}_{I-II}$, we can have all $\tilde{\mu}_{I-II+1}^{*}, \tilde{\mu}_{I-II+2}^{*}$ and $\tilde{\mu}_{I-II+1}^{**}$ decoded correctly.

b) $e + f_{t+1} + \cdots + f_{t-I+1} \leq t - I$. From the correct $\tilde{\mu}_{I-II}$, we can have all $\tilde{\mu}_{I-II+1}^{*}, \tilde{\mu}_{I-II+2}^{*}$ and $\tilde{\mu}_{I-II+1}^{**}$ decoded correctly.

c) $e + e_{t+1} + \cdots + e_{t-I+1} > t$ and $e + f_{t+1} + \cdots + f_{t-I+1} > t$. In this case, we have all $\tilde{\mu}_{I-II+1}^{*}, \tilde{\mu}_{I-II+2}^{*}$ and $\tilde{\mu}_{I-II+1}^{**}$ decoded correctly.

Using $\tilde{\mu}_{I-II}^{*}[a]^{[\tilde{\mu}_{I-II}]}$ and the decoding algorithm for the 2t-error-correcting BCH code we can remove the possible 2t errors in $\tilde{\mu}$ and have the correct $\tilde{\mu}^{**}$.

Since $C$ is a constant weight code of distance $4t + 2$ and we can remove up to 2t symmetric errors, the uncorrectable but detectable error patterns all include the unidirectional errors.

**Example 4:** Let $t = 2, n = 23$, and $I = 11$. We require that $\mu_{1} = 5$ so that $2^{\mu_{1}} - 1 \geq n = 23$. Then, $\mu = \mu_{1} + \mu_{2} = 10$. Using $p_{2} = 13$ and $r_{2} = 6$, we have $\left(\frac{p_{2}}{r_{2}}\right) = 1716 \geq 2^{10}$. Using $p_{1} = 6$ and $r_{1} = 3$, we have $\left(\frac{p_{1}}{r_{1}}\right) = 20 \geq 13$. Thus, we have a code $C$ of length $n' = 23 + 2(13 + 6) = 61$ and weight $11 + 2(6 + 3) = 29$. The size of the message space $M$ is $2^{61}/2^{29} > 2^{32}$. Code $C$ has minimum distance $4t + 2 = 10$ and hence is a 4-EC/AUED code.

2) **Codes with Distance 4t:** Let $C$ be a quasi-systematic code for which each codeword has the form of $\tilde{\mu} = \tilde{\mu}_{I-II}^{*} \tilde{\mu}_{I-II+1}^{*} \tilde{\mu}_{I-II+2}^{*} \tilde{\mu}_{I-II+1}^{**}$, where $\tilde{\mu}$ is the data part and $\tilde{\mu}_{I-II}^{*} \tilde{\mu}_{I-II+1}^{*} \tilde{\mu}_{I-II+2}^{*} \tilde{\mu}_{I-II+1}^{**}$ is the parity check part. The data part $\tilde{\mu}$ is an n-tuple of weight $l$ and hence the size of message space $C$ is $\left(\frac{q}{q_{2}}\right)$. Unlike the symmetric structure in case 1, lengths of $\tilde{\mu}$ and $\tilde{\mu}_{I-II}$ may be distinct. For $i = 1, \ldots, t$, $\tilde{\mu}_{I-II}^{*}$ is a $p_{1}$-tuple of weight $r_{1}$, and for $1 \leq i \leq t - 1$, $\tilde{\mu}_{I-II}^{*}$ is a $q_{t-1}$-tuple of weight $s_{t}$. Vectors $\tilde{\mu}_{I-II}$ and $\tilde{\mu}_{I-II}$ in binary form are $[w_{I-II,0}, w_{I-II,1}, \ldots, w_{I-II,n_{I-II} - 1}]$ and $[w_{I-II,0}, w_{I-II,1}, \ldots, w_{I-II,n_{I-II} - 1}]$, respectively. The requirements for designing the parity check part are listed as follows:

1) The design of $\tilde{\mu}_{I-II}^{*}$ to $\tilde{\mu}_{I-II}$ is exactly the same as the design in case 1.

2) Let $\mu' = \mu_{1} + \mu_{2} + \cdots + \mu_{t-1}$. We require that $\left(\frac{n_{I-II} - 1}{2}\right) \geq 2^{\mu'}$. Let $\tilde{\mu}_{I-II}$ be the subset of the set of all $q_{t-1}$-tuples of weight $s_{t-1}$ with cardinality $2^{\mu'}$. We need a one-to-one mapping

$$h_{t-1}: \text{all of } \mu' \text{-tuples } \rightarrow \tilde{\mu}_{I-II}.$$

For $2 \leq j < t$, the design of $\tilde{\mu}_{I-II}$ has the following requirements. Let $\lambda_{1}$ be the integer such that $2^{\lambda_{1}} - 1 < q_{t-1} + 1 \leq 2^{\lambda_{1}} - 1$. Let $\gamma_{j}$ be a primitive element in the field of $GF(2^{\lambda_{k+1}})$. We mark the $q_{t-1} + 1$ positions of $\tilde{\mu}_{I-II}$ by $\gamma_{j}, \gamma_{j+1}, \ldots, \gamma_{j+q_{t-1} - 1}$, respectively. Suppose that we need $\lambda_{j,1}, \lambda_{j,2}, \ldots, \lambda_{j,q_{t-1} + 1} - 1$ to represent $\gamma_{j}, \gamma_{j+1}, \ldots, \gamma_{j+q_{t-1} - 1}$. Let $\lambda_{j} = \lambda_{j,1} + \lambda_{j,2} + \cdots + \lambda_{j,q_{t-1} + 1} - 1$. We require that $\left(\frac{n_{I-II}}{q_{t-1}}\right) \geq 2^{\lambda_{j}}$. Let $\tilde{\mu}_{I-II}$ be the subset of the set of all $q_{t-1}$-tuples of weight $s_{t-1}$ with cardinality $2^{\lambda_{j}}$. We need a one-to-one mapping

$$h_{t-1}: \text{all of } \mu_{j} \text{-tuples } \rightarrow \tilde{\mu}_{I-II}.$$
A. Construction II-A: Codes with Distance 6

We now construct systematic constant weight codes with distance 6 by adding some constant weight parity check part similar to those used in constructing quasi-systematic constant weight codes by Boinck and van Tilborg. Let $C$ be a code of length $2n + p$ for which each codeword is in the form of $[a][b][c]$, where $a$ is an $n$-tuple, $\bar{a}$ is the binary complement of $a$, and $b$ is a $p$-tuple of weight $r$. Let \( \{a_0, a_1, \ldots, a_{n-1}\} \) be an abelian additive group and $P$ be a subset of $p$-tuples of weight $r$, which has cardinality of $n$. Mark the $n$ positions of $\bar{a}$ by $a_0, a_1, \ldots, a_{n-1}$ and define the one-by-one mapping $g: \{a_0, a_1, \ldots, a_{n-1}\} \rightarrow P$ as we did in Section II. Here, $\bar{a}$ is a codeword of an $(n, n - 1, 2)$ binary systematic code $V$. The encoding of $C$ is done by first encoding each $(n - 1)$-bit message $\overline{m}$ into a codeword $\overline{a}$ of $V$. Then, $\bar{b}$ is equal to $g(\Sigma_{i=0}^{n-1} a_i \cdot \alpha_i)$ and $\bar{c}$ is obtained by taking the binary complement of $\bar{a}$. Let $\bar{c} = [\bar{a}][\bar{b}][\bar{c}]$ and $\bar{c}' = [\bar{a}'][\bar{b}'][\bar{c}']$ be two distinct codewords of $C$. If $\bar{b} = \bar{b}'$, then $d(\bar{a}, \bar{a}') \geq 4$ from the argument used in Section II. Hence, $d(\bar{c}, \bar{c}') = d(\bar{a}, \bar{a}') + d(\bar{b}, \bar{c}') \geq 8$. If $\bar{b} \neq \bar{b}'$, then $d(\bar{a}, \bar{a}') \geq 2, d(\bar{b}, \bar{b}') \geq 2, d(\bar{a}', \bar{c}') \geq 2$. Hence, $d(\bar{c}, \bar{c}') \geq 6$. Therefore, the code $C$ has minimum distance of 6. Note that $C$ is a systematic constant code of length $2n + p$ and weight $n + r$. Hence $C$ is a 2-EC/AUED code.

Example 5: Let $n = 6, p = 4$, and $r = 2$. Then, we have a constant weight code of length 16 and minimum distance 6. This code can represent all the five-bit messages. This is a systematic 2-EC/AUED code.

B. Construction II-B: Codes with Distance $4t + 2$

Let $C$ be a binary code of length $2n + p$ for which each codeword is of the form $[a][b][c]$, where $a$ is an $n$-tuple and $b$ is a $p$-tuple of weight $r$. Here, we require that $n \leq 2^r - 1$ and $|P| \geq 2^r$. The $n$ positions of $\bar{a}$ are marked by $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$, respectively. The encoding is done by encoding each $k$-bit message $\overline{m}$ into $\bar{a}$, which is a codeword of an $(n, k)$ systematic BCH code (or shortened BCH code) with designed distance of $2t, t \geq 1$. Then, $\bar{b}$ is obtained by $h(\Sigma_{i=0}^{n-1} a_i \cdot \alpha^i)$ and $\bar{c}$ is obtained by the binary complement of $\bar{a}$, where $h$ is the one-to-one mapping defined in Section III. Let $\bar{c}$ and $\bar{c}'$ be two distinct codewords of $C$. If $\bar{b} = \bar{b}'$, then we see that $h(\Sigma_{i=0}^{n-1} (a_i - a_i') \cdot \alpha^i) = 0$ for $0 \leq j \leq 2t$. It follows from the BCH bound, we have $d(\bar{a}, \bar{a}') \geq 2t + 2$. Since $\bar{a}$ and $\bar{a}'$ are binary complements of $\bar{a}$ and $\bar{a}'$, respectively, we have $d(\bar{a}, \bar{a}') \geq 2t + 2$. Therefore, $d(\bar{c}, \bar{c}') \geq 4t + 4$. Combining the previous two conditions, we note that the minimum distance of $C$ is at least $4t + 2$. The code $C$ is a systematic constant weight code of length $2n + p$ and weight $n + r$. Clearly, $C$ is a systematic $(2t)$-EC/AUED code.

Construction II-A is very similar to the special case of Construction II-B with $t = 1$ except that the former requires that $|P| \geq n$ and the latter requires that $|P| \geq 2^r > n$. Therefore, sometimes the former is more efficient than the latter. Both Constructions II-A and II-B can be decoded in the following way. Let $[\bar{a}][\bar{b}][\bar{c}]$ be the vector to be decoded.

1) Decode $\bar{a}^{**}$ by using the decoding algorithm of the $(t - 1)$-error-correcting BCH code.

2) Denote $[\bar{a}'][\bar{b}']$ by using the decoding algorithm of the $t$-error-correcting BCH code.

3) If the number of symmetric errors in $\bar{c}^*$ is at most $2t$, then at least one of the decodings from steps 1 to 2 will result in the correct decoding. Otherwise, an uncorrectable error pattern will be detected.

Example 6: Let $V$ be the $(n, k)$ shortened BCH code of distance 4 where $n = 15, k = 10$. Let $p = 6$ and $r = 3$. Then, we have a systematic constant weight code $C$ of length $n' = 2n + p = 36$ and weight 18. Since the minimum distance of $C$ is 10, $C$ is a systematic 4-EC/AUED code.

V. COMPARISONS

We list some codes constructed in this paper in Tables I, II, and III. For comparisons, we also list the codes constructed by Nikolo, Gaitanis, and Philokyprou (NGP codes) [7] and the systematic codes constructed by Boinck and van Tilborg (BCH codes) [10] in the tables. In the tables, the symbol $n'$ is used to denote the length of the $t$-EC/AUED code, $n$ is used to denote the length of the base code for constructing the NGP code and the BCH code, and $k = \lceil \log_2 |M| \rceil$ is the number of message bits, where $M$ is the size of the message space. Moreover, in the tables, symbols I-A, I-B, I-C, I-D, and II are used to indicate codes from Constructions I-A, I-B, I-C, I-D, and II of this paper, respectively. The symbols $B, N, Q, R$, and $PT$ represent primitive BCH codes, nonprimitive BCH codes, quadratic residue codes and Piret codes [11], respectively; where shortened codes are also included, for example, a shortened primitive BCH code is also indicated by $B$. It is seen that code rates of the proposed codes are superior to those of NGP codes and BCH codes in most cases when $t = 2, 3, 4$. Roughly speaking, codes of Construction II usually have efficient code rates at low $k$ situations and codes of Construction I-D usually have efficient code rates at high $k$ situations. However, NGP codes usually have efficient code rates at high $t$ situations.

In addition to the code rates, decoding complexities are also important parameters to be considered. For each code represented in this paper, the decoding is repeating the decoding procedures of several binary BCH codes of shorter lengths. For a $(t)$-EC/AUED code of Construction I, the decoding complexity is about the complexity of decoding a $t$-error-correcting binary BCH code plus decoding several $i$-error-correcting binary BCH codes, where $i = t - 1$. Note that for Construction I-D, the parameter $i \leq \lceil t/2 \rceil$. For a $(t)$-EC/AUED codes of Construction II, the decoding complexity is only the complexity of decoding a $(t/2)$-error-correcting binary BCH code plus decoding an $i$-error-correcting binary BCH code, where $i = \lceil t/2 \rceil - 1$. Clearly, the decoding complexity of a code in Construction II is very low. For a quasi-systematic code with large value of $k$, the one-to-one mapping between $M$ and $M$ may be complicated. We can make a small compromise by slightly increasing the code length to achieve a much simplified mapping algorithm. This technique can be explained by the following example.
Example 7: Consider the codes of Examples 2 and 3, where we map the distinct \( \binom{t}{2} \) message into all the 26-tuples of weight 13. If we consider 23-bit messages, only 2^26 of all the 26-tuples of weight 13 are mapped. We make a compromise by increasing \( n \), the length of \( \overline{n} \), by two to 28. We further divide the vector \( \overline{n} \) into two parts \( \overline{a}_1 \) and \( \overline{a}_2 \), where \( \overline{a}_1 \) and \( \overline{a}_2 \) are both 14-tuples. Note that \( \binom{28}{14} = \binom{28}{14} = 3003 \) and \( \log_2 3003 \approx 11 \). The mapping of 23-bit messages into the 28-tuples can be implemented as follows. Let a message \( \overline{m} \) be expressed as \( (m_0 | m_1 | m_2) \), where \( m_0 \) is a single bit, \( m_1 \) and \( m_2 \) are 11-tuples. If \( m_0 = 0 \), then \( \overline{m} \) is mapped into an \( \overline{a} \) of the form \( \overline{a}_1 \| \overline{a}_2 \), where \( \overline{a}_1 \) is a 14-tuple of weight 6 mapped from \( \overline{m}_1 \) and \( \overline{a}_2 \) is a 14-tuple of weight 8 mapped from \( \overline{m}_2 \). If \( m_0 = 1 \), then \( \overline{m} \) is mapped into an \( \overline{a} \) of the form \( \overline{a}_1 \| \overline{a}_2 \), where \( \overline{a}_1 \) is a 14-tuple of weight 8 mapped from \( \overline{m}_1 \) and \( \overline{a}_2 \) is a 14-tuple of weight 6 mapped from \( \overline{m}_2 \). For \( i = 1 \) and 2, the mapping between \( \overline{m}_i \) and \( \overline{a}_i \) can be easily implemented by a 11 x 14 ROM. With this modification, the code length is increased by 2 for either Example 2 or Example 3. For both codes modified from Example 2 and Example 3, they are still more efficient than the corresponding NGP and BV codes.

VI. CONCLUSION

We generalize the technique of constructing constant weight codes of distance 4 used by Bose and Rao to construct two classes of constant weight codes as t-EC/AUED codes. The proposed codes are in either quasi-systematic or systematic form. Therefore, the mapping between messages and code-words can be easily implemented. Moreover, decoding of each proposed code can be adapted from the well-developed technique of decoding BCH codes and hence the complexity is about the order of decoding a comparable binary BCH code. In particular, decoding complexity for the proposed code in systematic form is very low. In addition, code rates of the proposed codes are very efficient in some cases.

REFERENCES


Mao-Chau Lin (S'86-M'87) was born in Taipei, Taiwan, Republic of China, on December 24, 1954. He received the Bachelor and Master degree, both in electrical engineering, from National Taiwan University in 1977 and 1979, respectively. He received the Ph.D. degree in electrical engineering from the University of Hawaii in 1986. From 1979 to 1982, he was an assistant scientist of Chung-Shan Institute of Science and Technology at Lung-Tan, Taiwan. He is currently with National Taiwan University, serving as Associate Professor in the Department of Electrical Engineering. His research interests include coding theory and cryptography.