Integer Codes Correcting High-Density Byte Asymmetric Errors

Aleksandar Radonjic and Vladimir Vujicic

Abstract—In optical networks without optical amplifiers the number of received photons never exceeds the number of sent ones. Hence, upon transmission, only asymmetric (1 → 0) errors can occur. Motivated by this fact, in this letter we present a class of integer codes capable of correcting high-density asymmetric errors within a b-bit byte. Unlike classical codes, these codes use integer and lookup table operations. As a result, they can be implemented "for free", i.e. without modifying the network hardware.

Index Terms—Integer codes, error correction, asymmetric errors, look-up table.

I. INTRODUCTION

Error control codes are usually designed for use on binary symmetric channels, where the error probabilities 0 → 1 and 1 → 0 are equal. However, in certain systems, the errors have a highly asymmetric nature. For the purpose of this paper, the most interesting example are optical networks without optical amplifiers (ONWOAs) (e.g. local and access networks) [1]. In these networks, the number of received photons never exceeds the number of transmitted ones. Hence, upon transmission only asymmetric (1 → 0) errors can occur [2], [3]. Besides this, it is known that these errors affect small number of bits. More precisely, the experiments showed that 99% to 99.9% of all errors are t-bit errors (1 ≤ t ≤ 4) confined to one or two adjacent bytes [4]-[6].

Another common feature of ONWOAs is high computing power of network nodes. This can be seen from the fact that exterior nodes (e.g. PCs and servers) contain general purpose processors, whereas interior nodes (e.g. switches and routers) are always equipped with network processors (NPs) [7], [8]. The only difference between these chips is that NPs integrate coprocessors for common kernels of computation (e.g. lookup and cryptographic operations). On the other hand, both these chips have integer execution units as well as the memory system including the caches [7], [8]. Hence, it can be said that they are designed for integer and lookup table operations.

Motivated by these facts, in this letter we present a new class of integer codes. The proposed codes, like those in [9]-[11], have several desirable properties including systematic structure, simple encoding/decoding procedures and fast error correction algorithm based on table lookups. However, unlike [9]-[11], the codes presented in this paper can correct two types of errors within a b-bit byte: single t/d asymmetric errors and double adjacent t′/d asymmetric errors, where 1 ≤ t < t < d and p = b/d ≥ 2. Thanks to this feature, they are more suitable for use in ONWOAs than the codes proposed in [9]-[11].

The organization of this paper is as follows: Section 2 deals with the construction of integer codes capable of correcting single t/d and double adjacent t′/d asymmetric errors in a b-bit byte (integer (S_{b/d}AEC-DA_{b/d}AEC)_{b} codes) Section 3 explains the implementation strategy for these codes, while Section 4 concludes the letter. Table 1 shows the notations used in this work.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>B_{i}</td>
<td>Integer value of the i-th b-bit data byte at the sender side</td>
</tr>
<tr>
<td>C_{B}</td>
<td>Integer value of the b-bit check-byte at the sender side</td>
</tr>
<tr>
<td>B_{i}</td>
<td>Integer value of the received i-th b-bit data byte</td>
</tr>
<tr>
<td>\hat{C}_{B}</td>
<td>Integer value of the received b-bit check-byte</td>
</tr>
<tr>
<td>C_{B}</td>
<td>Integer value of the b-bit check-byte at the receiver side</td>
</tr>
</tbody>
</table>

II. CODES CONSTRUCTION

A. Encoding and Decoding Procedures

Let \( Z_{2}^{b} = \{0, 1, \ldots, 2^{b} - 2\} \) be the ring of integers modulo \( 2^{b} - 1 \), and let \( C_{i} \) and \( C_{k+1} \) be integers such that \( C_{i} \in Z_{2}^{b} \setminus \{0, 1\} \) and \( C_{k+1} = -1 \). Now, suppose that the data are divided into \( k \) b-bit bytes. In that case, the encoder will compute the check-byte in the same way as in [9]-[11], i.e. by using the following operations:

\[
C_{B} = [C_{i} \cdot B_{1} + \cdots + C_{k} \cdot B_{k}] \pmod{2^{b} - 1} = \sum_{i=1}^{k} C_{i} \cdot B_{i} \pmod{2^{b} - 1} \tag{1}
\]

At the receiver, the decoder will perform the same calculation

\[
\hat{C}_{B} = [C_{i} \cdot \hat{B}_{1} + \cdots + C_{k} \cdot \hat{B}_{k}] \pmod{2^{b} - 1} = \sum_{i=1}^{k} C_{i} \cdot \hat{B}_{i} \pmod{2^{b} - 1} \tag{2}
\]

after which the syndrome \( S \) will be formed

\[
S = [C_{B} - \hat{C}_{B}] \pmod{2^{b} - 1} \tag{3}
\]

Obviously, when \( S \neq 0 \), the codeword is corrupted by one or more errors. Whether these errors can be corrected or not depends on the values of the coefficients \( C_{i} \). On the other hand, the coefficient \( C_{k+1} \) always has the same value, since it corresponds to the errors within the check-byte.
B. Necessary and Sufficient Conditions

**Definition 1.** An error is called t/d asymmetric error if t or fewer bits in a d-bit byte are in a 1 → 0 error, where 1 ≤ t < d.

**Definition 2.** An error is called low-density byte asymmetric (LDBA) error if there exists one t/d asymmetric error within a b-bit byte, where p = b/d ≥ 2.

**Definition 3.** An error is called high-density byte asymmetric (HDBA) error if there exists two adjacent t/d asymmetric errors within a b-bit byte, where 1 ≤ t' < t < d and p = b/d ≥ 2.

To make these definitions more clear, we give examples of LDBA and HDBA errors (Fig. 1).

![Fig. 1. Examples of (a) LDBA errors and (b) HDBA errors, where](image)

**Definition 4.** Let 0 ≤ x₁ ≤ ⋯ ≤ xₜ < d, 1 ≤ l ≤ t, 0 ≤ r ≤ p - 1, and let \( e_{i,r} = (2^{xₙ+\cdots+2^r}) \cdot 2^{d-r} \) be the difference between the integer values of the correct b-bit byte and its received erroneous counterpart affected by LDBA error. Then, the set of syndromes corresponding to LDBA errors is defined as

\[
s_i = \left\{ \left( \bigcup_{i=1}^{t} \bigcup_{r=0}^{p-1} (2^{c_i} \cdot e_{r,i}) \right) \mod 2^b \right\}
\]

(4)

**Definition 5.** Let 0 ≤ y₁ ≤ ⋯ ≤ yₘ, 0 ≤ z₁ ≤ ⋯ ≤ zₙ < d, 1 ≤ m, n ≤ t, 0 ≤ s ≤ p - 2, and let \( e_{m,n,s} = (2^{yₙ+\cdots+2^s}) \cdot 2^{d-s} \) be the difference between the integer values of the correct b-bit byte and its received erroneous counterpart affected by HDBA error. Then, the set of syndromes corresponding to HDBA errors is defined as

\[
s_s = \left\{ \left( \bigcup_{i=1}^{m} \bigcup_{j=1}^{n} \bigcup_{s=0}^{p-2} (2^{c_i} \cdot e_{m,n,s}) \right) \mod 2^b \right\}
\]

(5)

Now, we can prove the following theorem.

**Theorem 1.** The codes defined by (1)-(5) can correct all LDBA and HDBA errors if there exist k mutually different coefficients \( c_i \in Z_{2^d \setminus \{0\}} \) such that

1. \( |s_i| = (k+1) \cdot p \cdot \sum_{i=1}^{t} \frac{d}{i} \)
2. \( |s_s| = (k+1) \cdot (p-1) \cdot \sum_{n=1}^{\infty} \frac{d}{n} \)
3. \( s_i \cap s_s = \emptyset \)

where \( |A| \) denotes the cardinality of \( A \), and \( A \cap B \) the intersection of \( A \) and \( B \).

**Proof.** Condition 1 of this theorem says that LDBA errors generate \((k+1) \cdot p \sum_{i=1}^{t} \frac{d}{i}\) syndromes that are nonzero. To prove this, observe that the set \( s_i \) can be expressed as

\[
s_i = \bigcup_{i=1}^{t} M_i
\]

where

\[
M_1 = \left\{ \left( \bigcup_{i=1}^{t} \bigcup_{r=0}^{p-1} (2^{c_i} \cdot e_{r,i}) \right) \mod 2^b \right\}
\]

\[
M_2 = \left\{ \left( \bigcup_{i=1}^{t} \bigcup_{r=0}^{p-1} (2^{c_i} \cdot e_{r,i}) \right) \mod 2^b \right\}
\]

\[
M_{k+1} = \left\{ \left( \bigcup_{i=1}^{t} \bigcup_{r=0}^{p-1} (2^{c_i} \cdot e_{r,i}) \right) \mod 2^b \right\}
\]

Now, suppose that the coefficients \( c_i \) are chosen in such a way that each one multiplied (modulo \( 2^b-1 \)) by each \( e_{r,i} \) yields a different result. In that case, it will hold that

\[
|s_i| = \cdots = |M_k| = |M_{k+1}|
\]

As a consequence, the set \( s_i \) will have

\[
|s_s| = (k+1) \cdot |M_{k+1}| = (k+1) \cdot p \cdot \sum_{i=1}^{t} \frac{d}{i}
\]

nonzero elements. In a similar way Condition 2 says that HDBA errors generate \((k+1) \cdot (p-1) \cdot \sum_{n=1}^{\infty} \frac{d}{n} \) syndromes that are nonzero. To prove this, note that the set \( s_s \) can be expressed as

\[
s_s = \bigcup_{i=1}^{t} N_i
\]

where

\[
N_1 = \left\{ \left( \bigcup_{i=1}^{t} \bigcup_{j=1}^{n} \bigcup_{s=0}^{p-2} (2^{c_i} \cdot e_{m,n,s}) \right) \mod 2^b \right\}
\]

\[
N_2 = \left\{ \left( \bigcup_{i=1}^{t} \bigcup_{j=1}^{n} \bigcup_{s=0}^{p-2} (2^{c_i} \cdot e_{m,n,s}) \right) \mod 2^b \right\}
\]

\[
N_{k+1} = \left\{ \left( \bigcup_{i=1}^{t} \bigcup_{j=1}^{n} \bigcup_{s=0}^{p-2} (2^{c_i} \cdot e_{m,n,s}) \right) \mod 2^b \right\}
\]

Given this, suppose that the \( c_i \)s are chosen such that

\[
|N_1| = \cdots = |N_k| = |N_{k+1}|
\]

In that case, it is clear that the set \( s_s \) will have

\[
|s_s| = (k+1) \cdot |N_{k+1}| = (k+1) \cdot (p-1) \cdot \sum_{n=1}^{\infty} \frac{d}{n} \]

nonzero elements. Finally, Condition 3 is a necessary condition for distinguishing LDBA errors from HDBA errors. Therefore, the codes that satisfy the above conditions are \((kb + b, kb)\) integer \((S_{kb}\text{AEC-DA}_{kb}\text{AEC})_b\) codes. □
Theorem 2. Let \( \xi \) be the error set for \((kb + b, b)\) integer \((S_{kb}AEC-DA_{kb}AEC)\) codes. Then,
\[
|\xi| = |s_i| + |s_j| = (k + 1) \cdot \left[ p \cdot \sum_{l=1}^{d} \left( \frac{d}{l} \right) \right] + \left( p - 1 \right) \cdot \left[ \sum_{n=1}^{c} \left( \frac{d}{n} \right) \right] ^2.
\]

Proof. This theorem follows from Theorem 1. \( \square \)

From Theorems 1 and 2 it is easy to see that the elements of \( \xi \) cannot be generated without using a computer. Thus, it is clear that for some particular values of \( t, l, d, b \) and \( c \) we cannot know a priori the number and the values of coefficients \( C \). In this letter, we have restricted ourselves to practical codes, i.e. to codes with parameters \( t = 3, l = 4, d = 8, b = 32 \) and \( k \leq 64 \). The results of the corresponding search are given in Table 2.

<table>
<thead>
<tr>
<th>ELEMENTS OF SYNDROME TABLE</th>
<th>TABLE</th>
<th>NEEDED TO FIND</th>
<th>ERROR VECTOR</th>
<th>ERROR VECTOR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>( G )</td>
<td>( H )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>( M )</td>
<td>( N )</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>( P )</td>
<td>( Q )</td>
</tr>
</tbody>
</table>

C. Error Correction Procedure

From Theorem 2 we know that LDBA and HDBA errors generate \(|\xi|\) nonzero syndromes. In addition, from the same theorem, we implicitly know that the relationship between the nonzero syndrome (element of the set \( \xi \)), error location \( i \) and error vector \( e \) can be described using (4)-(5). Both these facts imply that the syndrome table requires \(|\xi| \cdot \left( 2 \cdot b + \left\lfloor \log_b (k + 1) \right\rfloor \right) \) bits (Fig. 2) to store the error correction data.

![Fig. 2. Bit-width of one syndrome table entry](image)

Given this, suppose that the data are received in error (\( S \neq 0 \)). In that case, the decoder will first search the syndrome table to find the appropriate entry. After that, it will execute one of the following operations:
- for LDBA errors within the \( t \)-th data byte
  \[
  B_i = [\hat{B}_i + e] \mod (2^b - 1), \quad 1 \leq i \leq k;
  \]
  \[
  e = [e_i] \mod (2^b - 1), \quad 1 \leq i \leq t, \quad 0 \leq r \leq p - 1;
  \]
- for LDBA errors within the check-byte
  \[
  C_B = [\hat{C}_B + e] \mod (2^b - 1);
  \]
  \[
  e = [e_i] \mod (2^b - 1), \quad 1 \leq i \leq t, \quad 0 \leq r \leq p - 1;
  \]
- for HDBA errors within the \( t \)-th data byte
  \[
  B_i = [\hat{B}_i + e] \mod (2^b - 1), \quad 1 \leq i \leq k;
  \]
  \[
  e = [e_{m,n}] \mod (2^b - 1), \quad 1 \leq m, n \leq t', \quad 0 \leq s \leq p - 2;
  \]
- for HDBA errors within the \( t \)-th check-byte
  \[
  C_B = [\hat{C}_B + e] \mod (2^b - 1);
  \]
  \[
  e = [e_{m,n}] \mod (2^b - 1), \quad 1 \leq m, n \leq t', \quad 0 \leq s \leq p - 2;
  \]

From the above it is clear that the efficiency of the error correction procedure (in terms of processing time) depends on the number of table lookups. For this reason it is desirable that the elements of \( \xi \) are sorted in increasing order. In that case it is possible to use binary search algorithm, which requires \( n_{TL} \) table lookups (1 \( \leq n_{TL} \leq \left\lfloor \log_b (|\xi| + 2) \right\rfloor \) [12].

Example 1. Let \( b = 8, \ d = 4, \ t = 3, \ t' = 1, \ k = 1 \) and \( C_1 = 2 \). According to Theorem 2, the syndrome table will have \(|\xi| = 88 \) entries. Given this, let us assume that we want to transmit 8 bits of data, \( D = 10110011 \). In that case, after calculating the value of check-byte \( C_B \)
\[
C_B = [C_1 \cdot B_1] \mod (2^b - 1) = [2 \cdot 179] \mod (255) = 103
\]
the codeword \( C_W = 10110011 01101111 \) will have 16 bits.

Now, let us analyze the following scenarios.

Scenario 1: Suppose that during data transmission an error on the 1\( \text{st} \), 3\( \text{rd} \) and 4\( \text{th} \) bit has occurred (\( C_W = 00000001 01101111 \)). In that case, the decoder will calculate
\[
C_B = [C_1 \cdot B_1] \mod (2^b - 1) = [2 \cdot 3] \mod (255) = 6
\]
\[
S = [C_W - C_B] \mod (2^b - 1) = [6 - 103] \mod (255) = 158
\]
in order to check whether the value \( S = 158 \) belongs to the set \( \xi \) (Table 3). After completing this task, it will perform error correction by using
\[
B_i = [\hat{B}_i + e] \mod (2^b - 1) = [3 + 176] \mod (255) = 179.
\]

Scenario 2: Let us assume that during data transmission an error on the 10\( \text{th} \) and 16\( \text{th} \) bit has occurred (\( C_W = 01101011 00100111 \)). Similar to the previous case, after calculating
\[
C_B = [C_1 \cdot \hat{B}_1] \mod (2^b - 1) = [2 \cdot 179] \mod (255) = 103
\]
\[
S = [C_W - C_B] \mod (2^b - 1) = [103 - 38] \mod (255) = 65
\]
the decoder will conclude that the value \( S = 65 \) indicates an error within the check-byte (Table 3). As a consequence, the following procedure will take place:
\[
C_B = [C_W + e] \mod (2^b - 1) = [38 + 65] \mod (255) = 103.
\]

III. IMPLEMENTATION STRATEGY

From (1)-(3) and (6)-(9) it is clear that the encoder/decoder uses integer and lookup table (LUT) operations. Since these operations are supported by all processors, it is interesting to discuss how the proposed codes can be implemented on modern architectures. Without loss of generality, we will restrict ourselves to eight-core processors (Fig. 3) and integer \((S_{kb}AEC-DA_{kb}AEC)\) codes.

![Fig. 3. Block diagram of eight-core processor](image)
1) L1 caches. The L1 cache is the smallest and fastest type of cache. Its size is limited to 64 KB (per core), whereas the data can be accessed in 1-5 clock cycles [8]. Both these facts suggest that the L1 can be used for storing the coefficient table (LUT₁) whose size is \( k \times b \) bits (Table 4).

2) L2 caches. The L2 cache is somewhat slower and larger than the L1. Precisely, its size is limited to 512 KB (per core), whereas the individual entries can be accessed in 8-15 clock cycles [8]. Based on these parameters, it is clear that the L2 cannot be used for storing the syndrome table (LUT₂).

3) L3 cache. Unlike L1 and L2 caches, the L3 is shared among all cores. Due to this reason, this cache has the highest access latency (25-50 clock cycles). On the other hand, its size is limited to 512 KB (per core), whereas the data can be accessed in 1-5 clock cycles [8]. Both these facts suggest that the L3 can be used for storing the coefficient table (LUT₁) whose size is \( k \times b \) bits (Table 4).

### IV. Conclusion

In this letter, we proposed a class of integer codes capable of correcting high-density asymmetric errors within a b-bit byte. We have shown that these codes can be implemented "for free", i.e., without modifying the network hardware. Thanks to this feature, the proposed codes have high potential to be used in practice. This primarily refers to optical networks without optical amplifiers in which all nodes possess powerful processors.

### REFERENCES


### TABLE III

THE SYNDROME TABLE (LUT₂) FOR (16, 8) INTEGER (S₄/₄AEC-DA₃/₄AEC)₈ DECODER.

<table>
<thead>
<tr>
<th>Element of the set ( \xi )</th>
<th>( i )</th>
<th>( e )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>17</td>
<td>17</td>
<td>2</td>
</tr>
<tr>
<td>18</td>
<td>18</td>
<td>2</td>
</tr>
<tr>
<td>19</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>20</td>
<td>20</td>
<td>2</td>
</tr>
<tr>
<td>21</td>
<td>21</td>
<td>2</td>
</tr>
<tr>
<td>22</td>
<td>22</td>
<td>2</td>
</tr>
</tbody>
</table>

### TABLE IV

LOOK-UP TABLE SIZES FOR SOME INTEGER (S₄/₄AEC-DA₃/₄AEC)₈ CODES.

<table>
<thead>
<tr>
<th>Code</th>
<th>Encoder</th>
<th>Decoder</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LUT₁</td>
<td>LUT₁</td>
</tr>
<tr>
<td></td>
<td>Size</td>
<td>Size</td>
</tr>
<tr>
<td>(512, 480)</td>
<td>60 B</td>
<td>60 B</td>
</tr>
<tr>
<td>(544, 512)</td>
<td>64 B</td>
<td>64 B</td>
</tr>
<tr>
<td>(1024, 992)</td>
<td>124 B</td>
<td>124 B</td>
</tr>
<tr>
<td>(1056, 1024)</td>
<td>128 B</td>
<td>128 B</td>
</tr>
<tr>
<td>(2048, 2016)</td>
<td>252 B</td>
<td>252 B</td>
</tr>
<tr>
<td>(2080, 2048)</td>
<td>256 B</td>
<td>256 B</td>
</tr>
</tbody>
</table>