Abstract— This paper presents an error tolerant hardware efficient VLSI architecture for bit parallel systolic multiplication over dual base, which can be pipelined. This error tolerant architecture is well suited to VLSI implementation because of its regularity, modular structure, and unidirectional data flow. The length of the largest delay path and area of this architecture are less compared to the bit parallel systolic multiplication architectures reported earlier. The architecture is implemented using Austria Micro System’s 0.35um CMOS technology. This architecture can also operate over both the dual-base and polynomial base.

Keywords— Finite Field, RS codes, bit parallel, systolic, error correction, VLSI Testing.

I. INTRODUCTION

Finite field also known as Galois Field arithmetic operations over $GF(2^m)$ find increasing applications in public-key cryptography, error detecting and correcting code[9], VLSI testing[10], digital signal processing[11]. There are different equivalent representations of the elements of the finite field over $GF(2^m)$ e.g. Polynomial base (PB), normal base, and dual base. Dual-basis operators frequently have the lowest hardware requirements of all available operators [18-19]. Two basic operations over $GF(2^m)$ are addition and multiplication. Addition over $GF(2^m)$ is relatively straightforward to implement, requiring at most m XOR gates. Multiplication operation is much more expensive in terms of gate count and clock cycle. Other operations of the $GF(2^m)$ fields like exponentiation, division, and inversion can be performed by repeated multiplications. Based on different base representation, a variety of architectures for multiplication have been proposed. For high speed VLSI implementation, the preferred architecture for polynomial basis (PB) multiplier is systolic array architecture. In this type of architecture, a basic cell is repeated in an array and signals flow unilaterally between neighbours. PB systolic array multipliers in $GF(2^m)$ can be classified into four categories, namely bit serial, bit-parallel, hybrid and digit-serial. The bit serial architecture has minimum area and minimum throughput among all the categories. The problem with serial architecture is its latency. The bit-serial architecture, which processes one bit of input data per clock cycle, is area-efficient and suitable for low-speed applications. The most widely used bit serial multiplier is dual basis Berlekamp bit serial multiplier [12]. This multiplier requires less hardware. PB bit-serial and bit-parallel systolic multipliers were presented in [8, 13]. A bit-serial dual basis systolic multiplier over $GF(2^2)$ was presented in [3], which requires higher hardware compared to that needed for multiplier proposed in [6] and does not support pipelining. To support pipelining, a modified version which requires less hardware is presented in [14]. The bit parallel multiplier needs largest area and provides maximum throughput. Bit-parallel architecture, capable of processing one whole word of input data per clock cycle, is ideal for high-speed applications when pipelined at the bit-level. These architectures are typical examples of the area-speed tradeoff paradigm. Mastrovito has proposed an algorithm along with its hardware architecture for PB multiplication [7] known as the Mastrovito algorithm/multiplier. A formulation for Polynomial basis multiplication and generalized bit-parallel hardware architecture for special reduction polynomials has been presented in [2]. A testable polynomial basis bit parallel multiplier circuits over $GF(2^m)$ was presented in [21]. Although bit-serial dual basis multipliers have been widely employed in applications such as RS encoders [3], it has been proven in [19] that it is advantageous of employing bit-parallel dual basis multipliers, particularly in more complex circuits such as RS decoders and syndrome calculators. Bit-parallel dual basis multipliers therefore allow for reduced complexity constant multipliers. In this paper, we present a hardware efficient fast bit parallel systolic architecture with error detecting capability using parity prediction technique over dual base which can be pipelined.

II PRELIMINARIES

a) Polynomial Multiplication

Let $GF(N)$ denote a set of N elements, where N is a power of a prime number, with two special elements 0 and 1 representing the additive and multiplicative identities respectively and two operator addition ‘+’ and multiplication ‘·’. The $GF(N)$ defines a finite field, if it forms a commutative ring with identity over these two operators in which every element has a multiplicative inverse. Finite fields can be generated with primitive polynomials of the form $P(x) = x^m + \sum_{i=0}^{m-1} p_i x^i$, where $p_i \in GF(2)^{[9]}$. It is conventional to represent the elements of $GF(2^m)$ as a power of the primitive element $\alpha$ where $\alpha$ is the root of $P(x)$, i.e. $P(\alpha) = 0$. The set $\{1, \alpha, \ldots, \alpha^{m-1}\}$ is referred to as polynomial basis or standard basis. Each element $\alpha^j \in GF(2^m)$ can be expressed with respect to the PB as a polynomial of degree m over $GF(2)$, i.e. $\alpha^j = \sum_{i=0}^{m-1} a_i \alpha^i$, where $a_i \in GF(2)$. Given $A \in GF(2^m)$, $B \in GF(2^m)$, PB multiplication over $GF(2^m)$ can be defined as $C(x) = A(x).B(x) \mod P(x)$. In practice C(x) is obtained in two steps: polynomial multiplication and modulo reduction.

b) Dual Basis Multiplication

Let $F_{\alpha}$ denote the set of all linear function $f:GF(p^m)\rightarrow GF(p)$. A well known linear function is the trace function which is frequently used to produce the finite field multipliers. Rather than trace function there are a number of other linear function. We follow the definition of the duality of two bases [16-17] as given below.

Definition: Let $\{\alpha_i\}$ and $\{\mu_j\}$ be bases for $GF(2^m)$, let $f:GF(2^m)\rightarrow GF(2)$ be a linear function and let $\beta \in GF(2^m)$. Then the bases are said to be dual with respect to $f$ and $\beta$ if

$$f(\beta \lambda, \mu_i) = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases}$$

In this case $\{\alpha_i\}$ is the standard basis and $\{\mu_i\}$ is the dual basis. We now restate the multiplication algorithm utilized here. This result was first presented in the context of division [16] but has subsequently
been used to describe finite-field multiplication [15]. Furthermore, as observed in [1], the following represents a generalized and alternative representation of Berlekamp bit-serial multiplier.

Theorem 1 [14]: Let a, b, c ∈ GF(p^m) such that c = ab. Further, let α be a root of the defining irreducible polynomial for the field, let β ∈ GF(2^m), f ∈ F_2^m and represent c over the polynomial basis by

\[ \sum_{i=0}^{m-1} a_i \alpha^i \]

Then the following relation holds.

\[
\begin{bmatrix}
 f(b_0) & f(b_1) & \cdots & f(b_{m-1}) \\
 f(b_0 \alpha) & f(b_1 \alpha) & \cdots & f(b_{m-1} \alpha) \\
 \vdots & \vdots & \ddots & \vdots \\
 f(b_0 \alpha^m) & f(b_1 \alpha^m) & \cdots & f(b_{m-1} \alpha^m)
\end{bmatrix}
\begin{bmatrix}
 a_0 \\
 a_1 \\
 \vdots \\
 a_{m-1}
\end{bmatrix} =
\begin{bmatrix}
 f(c_0) \\
 f(c_1) \\
 \vdots \\
 f(c_{m-1})
\end{bmatrix}
\]

(1)

We have modified eqn. (1) as follows.

\[
\begin{bmatrix}
 b_0 & b_1 & \cdots & b_{m-1} & a_0 \\
 b_0 & b_1 & \cdots & b_{m-1} & a_1 \\
 \vdots & \vdots & \ddots & \vdots & \vdots \\
 b_0 & b_1 & \cdots & b_{m-1} & a_{m-1}
\end{bmatrix} =
\begin{bmatrix}
 c_0 \\
 c_1 \\
 \vdots \\
 c_{m-1}
\end{bmatrix}
\]

(2)

Where \( b_k = f(b \alpha^k) \) (k = 0, 1, 2m - 2) and \( c_j = f(c \beta^j) \) (k = 0, 1, m - 1). If \( \beta \) and \( \alpha \) are taken as in the preceding definition, \( c_0 \) and \( b_k \) (k = 0, 1, 2m - 1) in eqn. 1 are the dual-basis coefficients of c and b, respectively. Thus to make use of eqn. 1 in a systolic multiplier one must first generate the values of \( b_k \) (k = m, m + 1, ..., 2m - 2).

If \( p(s) = \sum_{j=0}^{m} p_j s^j \) is the defining irreducible polynomial for the field then \( b_w = f(b \alpha^w) = f(b \sum_{j=0}^{m-1} p_j \alpha^j) = \sum_{j=0}^{m-1} p_j f(b \alpha^j) = \sum_{j=0}^{m-1} p_j b_j \)

and then \( b_{w+k} = f(b \alpha^w \alpha^k) = f(b \beta^k (\alpha^w \beta^k)) = \sum_{j=0}^{m-1} p_j f(b \alpha^j \beta^k) = \sum_{j=0}^{m-1} p_j b_{j+k} \)

Then in general \( b_{w+k} = \sum_{j=0}^{m-1} p_j b_{j+k} \)

(3)

where \( b_k \) (k = 0, 1, ..., m - 1) are the dual basis coefficients of b and \( \alpha \) is root of \( p(s) \). Having generated these values of \( b_k \) from eqn. 2 one need to carry out the matrix multiplication given in eqn. 1. Now consider the implementation of this multiplication algorithm in the design of a bit-parallel systolic multiplier.

III BIT PARALLEL DUAL BASIS MULTIPLIER

a) Proposed Architecture

Let a, b, c ∈ GF(2^m) such that c = ab and \{μ_j\} be the dual basis to the polynomial basis for \( \beta \in GF(2^m) \) and \( f \in F_2^m \). Representing ‘b’ over the dual basis by \( b = \sum_{i=0}^{m-1} b_i \mu_i \) and ‘a’ over the polynomial basis by \( a = \sum_{i=0}^{m-1} a_i \alpha^i \). We can derive followings from eqn. 2.

\( c_0 = \sum_{i=0}^{m-1} a_i b_i \mu_i \)

\( c_j = \sum_{i=0}^{m-1} a_i b_i \mu_i \) where \( \mu_i = \sum_{j=0}^{m-1} a_j \alpha^j \)

Then in general

\( c_j = \sum_{j=0}^{m-1} a_j \mu_j \)

(4)

A bit-parallel dual basis multiplier over GF(2^m) can, therefore, be constructed using two cells. We introduce cell-1 as shown in Fig. 2 to generate eqn. (3) and also introduce a cell-2 for generating eqn. (2) as shown in Fig. 1. An example of such a multiplier over GF(2^4) is given below.

Example 1: Let \( p(x) = x^4 + x + 1 \) be the defining irreducible polynomial and let ‘α’ be a root of \( p(x) \). From eqns. (4), we can write as follows:

\( h(b, a) = b_0 a_0 + b_1 a_1 + b_2 a_2 + b_3 a_3 \) ... (5)

This equation can be implemented by the circuit as shown in Fig. 2. From \( p(x) = x^4 + x + 1 \) and eqns. (3) and (4), we can derive the values of \( b_k, b_0, b_1 \) as follows:

\( b_2 = b_0 + b_1 + b_2 \) \( b_3 = b_1 + b_2 \)

The eqn. (2) for this example is given below.

Fig. 1: Generation of Partial Products of eqn. 1

Fig. 2: Generation of the sum of partial products of eqn. 2

Fig. 3: Arrangement of systolic cells for bit-parallel multiplier for GF(2^4).

The m' cells of Fig 1 and m cells of Fig 2 are then combined to form the full bit-parallel dual basis multiplier for GF(2^m) as shown in Fig 3.

If \( b = \sum_{i=0}^{m-1} b_i \mu_i \) is the dual basis representation of \( b \) and a

\( a = \sum_{i=0}^{m-1} a_i \alpha^i \) is the polynomial basis representation of a. The product bits \( c_j (i = 0, 1, 2, 3) \) become available on the output lines. In the architecture the \( b_k, b_0, b_1, b_2 \) is generated by the block diagram of Fig. 2. In general Fig. 2, represents the equation (2) \( b_{w+k} = \sum_{j=0}^{m-1} p_j b_{j+k} \) where \( k = 0, 1, ..., m - 2 \). The partial sum in the matrix multiplication in eqn. (1) is generated by the block diagram of Fig. 1.

In BP Systolic dual basis multiplier design of [14], there exist two datapath, one is horizontal and the other is vertical. The vertical datapath generates partial sum in matrix multiplication of eqn. (1). The horizontal data path generates partial sum of eqn. (2).
bottleneck to support pipelining in this design. The horizontal data path consists of AND-XOR binary tree, the depth of tree is O(m). We try to modify the horizontal path by replacing the binary tree of depth O(m) with a binary tree of depth of O(log2m). For this purpose, we introduce a new cell [Fig. 2] to generate the eqn. (2). The complete circuit for dual basis systolic multiplier over GF(2^m) is shown in Fig. 3. Latches are introduced in Fig. 3, to make this architecture suitable for pipelining. There is m- clock cycle delay between ‘b’, ‘c’ entering in the multiplier and becoming available in the output lines. After the initial delay, results can be produced continuously one per clock cycle.

**b) Hardware and Delay Analysis**

We compare our proposed architecture with the bit parallel architecture described in [16]. Total hardware required for the architecture presented consists of m^2 cells. Each cell consists of two 2 input AND gates and two 2 input EXOR gates. Total circuit consists of 2m^2 AND gates and 2m^2 EXOR gates. Our proposed design requires 2 cells. First cell consists of one AND gate and one EXOR gate. Second cell consists of m AND gates and (m-1) EXOR gates. For m bit multipliers, the proposed architecture consists of m^2 first cells and m second cells. Total 2m^2 AND gates and (2m^2-m) EXOR gates are required. Overall saving in hardware is m EXOR gate.

**Table 1: Comparison between two-bit parallel systolic multipliers**

<table>
<thead>
<tr>
<th>Properties</th>
<th>Reference [19]</th>
<th>Presented here</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of cells</td>
<td>m^2</td>
<td>m \cdot m^2 + m</td>
</tr>
<tr>
<td>Circuit complexity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No of 2 input AND gate</td>
<td>2m^2</td>
<td>2m^2</td>
</tr>
<tr>
<td>No of 2 input XOR gate</td>
<td>2m^2</td>
<td>2m^2 + m</td>
</tr>
<tr>
<td>Largest delay path</td>
<td>(2m^2)[D_a + D_x]</td>
<td>m \cdot D_a + (log_2^m)m-1</td>
</tr>
</tbody>
</table>

Let D_a be the delay through a two-input AND gate and D_x be the delay through a two-input-XOR gate. The longest delay path is given in the eqn. (6). Longest Delay = \[(D_a + log_2^m - 1) + D_x \cdot (log_2^m + 1)\].

**Table 2: Hardware requirements and delays of dual basis Bit parallel multiplier (DPM) presented in [16] and the proposed multiplier (DPM)**

<table>
<thead>
<tr>
<th>m</th>
<th>AND XOR Delay</th>
<th>DPM [16]</th>
<th>AND XOR delay</th>
<th>DPM [PROPOSED]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>8</td>
<td>89.5 D_a</td>
<td>8</td>
<td>6 \cdot 2^D_a</td>
</tr>
<tr>
<td>3</td>
<td>18</td>
<td>18.5 D_a</td>
<td>18</td>
<td>15 \cdot 3D_a</td>
</tr>
<tr>
<td>4</td>
<td>32</td>
<td>32.5 D_a</td>
<td>32</td>
<td>15 \cdot 4D_a</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>50.5 D_a</td>
<td>50</td>
<td>45 \cdot 5D_a</td>
</tr>
<tr>
<td>6</td>
<td>72</td>
<td>72.5 D_a</td>
<td>72</td>
<td>66 \cdot 6D_a</td>
</tr>
<tr>
<td>7</td>
<td>98</td>
<td>98.5 D_a</td>
<td>98</td>
<td>91 \cdot 7D_a</td>
</tr>
<tr>
<td>8</td>
<td>128</td>
<td>128.5 D_a</td>
<td>128</td>
<td>120 \cdot 8D_a</td>
</tr>
<tr>
<td>9</td>
<td>162</td>
<td>162.5 D_a</td>
<td>162</td>
<td>153 \cdot 9D_a</td>
</tr>
<tr>
<td>10</td>
<td>200</td>
<td>200.5 D_a</td>
<td>200</td>
<td>190 \cdot 10D_a</td>
</tr>
</tbody>
</table>

From the table we can conclude that in this architecture, the number of AND gates are same compared to previous architecture [19], but for m-bit dual basis systolic multiplier m no. of XOR gates are less required in this architecture as well as the longest path delay of this architecture is also reduced by m-bit for AND gates and for XOR gates delay is reduced by log_2^m instead of m.

In Table 2, the hardware complexity and delays of the DPM [19] and the proposed DPM architecture are given for GF(2^m) for (m = 2, 3, . . . , 10). From Table 1, it can be seen that for every case, the hardware complexity and delays of our proposed DPM architecture are less compared to those of the DPM architecture [19].

**IV. Error Detection Using Parity Checking**

We use error detection scheme with a very high probability of detecting faults in the bit-parallel systolic multiplication over GF (2^m) using dual base with some additional outputs, called the check-bits as shown in Fig. 4. We assume that no interconnections or buses have any fault and each test phase with the test-circuits is separately controllable. At first, we attach parity-bits to the input elements: b_p and a_p and multiplying (AND) the inputs we have, b_p = b_0 \oplus b_1 \oplus b_2 \oplus b_3 , a_p = a_0 \oplus a_1 \oplus a_2 \oplus a_3

b_p = \{b_0 \oplus b_1 \oplus b_2 \oplus b_3\} \cdot (a_0 \oplus a_1 \oplus a_2 \oplus a_3) = \{b_a a_0 \oplus b_a a_1 \oplus b_a a_2 \oplus b_a a_3\} \oplus \{b_d a_0 \oplus b_d a_1 \oplus b_d a_2 \oplus b_d a_3\} \oplus \{b_x a_0 \oplus b_x a_1 \oplus b_x a_2 \oplus b_x a_3\}.

Now, from eqn. (2) of the previous architecture, we get

c_0 = b_0 a_0 \oplus b_0 a_1 \oplus b_0 a_2 \oplus b_0 a_3 ; c_1 = b_1 a_0 \oplus b_1 a_1 \oplus b_1 a_2 \oplus b_1 a_3 \cdot c_2 = b_2 a_0 \oplus b_2 a_1 \oplus b_2 a_2 \oplus b_2 a_3 ; c_3 = c_0 \oplus b_3 a_1 \oplus b_3 a_2 \oplus b_3 a_3 .

Now, we denote the modulo2 addition of these outputs of the multiplier by, r = c_0 \oplus c_1 \oplus c_2 \oplus c_3.

Here, we add some extra lines and gates for the testing purposes which constitute the feedback lines y_1. Lines b_1, b_2, b_3 and some XOR and AND gates are used to produce the circuit suitable for the testing. Some lines are used as feedback and are denoted by (y_1, y_2, y_3, y_4, y_5, y_6). So, some of the terms are eliminated when the b_p, a_p are added by modulo 2 addition to form the parity check in the output line with the feedback lines.

The y_lines are given as: y_1 = b_a a_1 \oplus b_d a_2 \oplus b_d a_3 ; y_2 = b_d a_1 \oplus b_a a_2 \oplus b_1 a_3 ; y_3 = b_d a_1 \oplus b_1 a_2 \oplus b_a a_3 .

The q line is derived from modular addition of b_p and the y_lines.

q = b_p \oplus y_1 \oplus y_2 \oplus y_3 \oplus y_4 \oplus y_5 \oplus y_6 = b_a a_0 \oplus b_d a_1 \oplus b_d a_2 \oplus b_d a_3 \oplus b_a a_0 \oplus b_a a_1 \oplus b_a a_2 \oplus b_a a_3 \oplus b_1 a_0 \oplus b_1 a_1 \oplus b_1 a_2 \oplus b_1 a_3 \oplus b_2 a_0 \oplus b_2 a_1 \oplus b_2 a_2 \oplus b_2 a_3 \oplus b_3 a_0 \oplus b_3 a_1 \oplus b_3 a_2 \oplus b_3 a_3 .

Now, rearranging, we see that q and r are same:

q = b_0 a_0 \oplus b_0 a_1 \oplus b_1 a_0 \oplus b_1 a_3 \oplus b_2 a_0 \oplus b_2 a_3 \oplus b_3 a_0 \oplus b_3 a_1 \oplus b_1 a_1 \oplus b_2 a_2 \oplus b_3 a_2 \oplus b_1 a_2 \oplus b_2 a_1 \oplus b_3 a_3 .

A parity checking circuit is presented in the figure which is correctly functioning for the Bit-parallel systolic multiplication over GF (2^m) using dual base. If the circuit operation is correct then q and r will agree and p = r \oplus q = 0. If any cell in the circuit is faulty, that will change the output lines and that fault reflects in the r line, as q remains unaltered, so p = 1 and the fault is detected. And if there is any failure in the y_i line that can also be detected by p = 1. Actually few of the y_i terms cancel the output parity checking operation as because they appear an even number of times in the coefficient of the output and cancelled out in the parity-checking operation. It can be improved further as the y_i terms are the sum of the results of some of the individual cells. So, if it is possible to temporarily disconnect those cells and connect with some lines to produce the desired feedback lines then the extra gates will not be required for the check
line q. Then the circuit complexity will be reduced and less time will be required.

DELAY: As the architecture is pipelined, so the path delays of each stage is same, except the last stage. The last has the maximum path delay. This can be calculated as for m-bit architecture:
\[ T_d = 2mT_{\text{XOR}} + T_{\text{AND}} \]
In our example in fig. 1, we calculate the path delay as \( T_d = 8T_{\text{XOR}} + T_{\text{AND}} \)

a) Simulation Result
We have modeled our proposed architecture in VHDL. The design was simulated in “Model Sim XE III 6.3c” and checked the functionality of the multiplier for different values of m. The post CTS-post detailed route layout of design for \( GF(2^2) \) is shown in Fig. 5.

V. CONCLUSIONS
The paper presented a fast dual-basis error tolerant bit-parallel systolic multiplier architecture over \( GF(2^m) \), which can be pipelined and which requires less hardware compared to that required in the multiplier architecture proposed earlier. Our proposed multiplier can also operate over both the dual-base and polynomial base. The proposed multiplier provides shorter longest delay path compared to that provided by the architecture presented earlier. A simple and efficient error detection procedure using parity checking has been incorporated with some additional AND- XOR gates.

![Fig.5: Layout of Bit-parallel Dual Basis systolic Multiplier for \( GF(2^2) \) with Error Checking Circuit](image)

REFERENCE