CHAPTER 3

CONSTRUCTION OF NOVEL ENHANCED HAMMING CODE FOR CRYPTO-CODING ARCHITECTURES

3.1 Introduction

The conventional way of transferring the secret message, from one end to another end, over the noisy channel is composed of channel coding [63] and cryptography [117] techniques. But, the properties and behavior of these two methods are entirely different. Secured message transmission over the wireless communication channel is done by the forward error correcting codes and cryptography techniques. A good encryption algorithm will penetrate more number of errors on retrieving the original message due to channel error. It leads to the need of most efficient error control algorithm to transfer the secret data over the wireless communication channel. In the past few decades, many theoretical links between coding theory and cryptography have been expressed [127]. The present digital scenario of “System on Chip” drives the need for combining the functions of coding and cryptography into a single algorithm without sacrificing required security and error control. This idea will reduce the considerable amount of delay and power consumptions in digital devices. Coding and Cryptography functions have distinct scopes to achieve their reliability higher. Hence, very few reliable methods have been identified to merge coding and cryptography functions into one function as Crypto-coding. The employability of cryptography and error correction functions is at the different purpose to each other. A cipher needs the property of avalanche effect on bit inversion of the original message with false key and channel coder possess the property of bit error control on the original message. The reliability of avalanche effect of chipper and error control of channel coder depends on the amount of creation of bit inversion due to the wrong
usage of keys and amount of correction of bit inversion due to channel noise respectively.

Although these two functions exhibit differences in their inherent properties, researchers have done much work to combine them to form a single function. This function is achieved by looking into few of their common properties such as bit length modulation, non-linearity propagation, uniqueness in the set of output. This paper describes one such method of Crypto-coding. In the past few years, diffusion properties of certain error correcting codes have been used to create ciphers [94]. For example, the Mix Column operation of the Advanced Encryption Standard (AES) cipher is generated using Maximum Distance Separable (MDS) codes [19]. Mathematically derived channel codes are bounded by systematic characteristics. So, they do not possess the adequate amount of diffusion required by ciphers. However, researchers are still working on to find a single efficient algorithm for error control and cryptography.

In this thesis, we establish a novel Crypto-coding schemes as a tool for constructing cryptographically secured block cipher with the good non-linearity property. We have proposed a novel Mixed Parity Code (MP Code) to develop our design. The behavior of this proposed MP code is investigated for designing a novel Enhanced Hamming Code along with high-security strength.

For the above-mentioned investigation, we analyze three major factors as (1) similarity between bit symbols of message input ‘k’ and bit symbols of codeword output ‘n’ of \((n, k, q)\) MP Code, (2) occurrence of non-linearity in data recovery between input block size ‘k’ and output block of size ‘n’ of MP code, and (3) Hamming Distance between the set of output blocks ‘\(n_j\)’ (where \(j \in \{1, 2, 3, \ldots, (1-2^n)\}\)) of MP code.
3.2 Preliminary Concepts for Proposed Mixed Parity Codes

In wireless communication systems, the coding theory is used for forward error correction technique of transmission data over the noisy channel. It is a technique used for controlling errors in data transmission over unreliable or noisy communication channels. The principal idea is the sender encodes their message with redundant bits by using an error-correcting code (ECC). As per the coding theory, error correction capability can be improved by adding redundancy bits. This redundancy addition leads to the higher computational complexity of forward error control algorithms.

3.2.1 Coding Theory

A block code is a set of words that has a well-defined mathematical property or structure, and where each word is a sequence of a fixed number of bits. The words belonging to a block code are called codewords. Examples of simple block code with 4-bit codewords are BCD codes, Gray Codes. A word with ‘n’ bits is referred to as an n-bit word. The idea in the formation of a codeword is to add redundancy to the message or information in order to be able to detect and correct the errors. We use an encoding algorithm to add this redundancy and a decoding algorithm to reconstruct the initial message. A message of length ‘k’ is transformed into a codeword ‘c’ of length ‘n’ with $n > k$ and $n = k + r$ where ‘r’ is the number of redundancy bits in a codeword. A code, whose codewords have ‘k’ information bits or message bits, ‘r’ parity bits and n-bit codewords where $n = k + r$, is referred as an $(n, k)$ block code where $n$ and $k$ are the block length and information length of the code respectively. The position of the parity bits ‘r’ within a codeword is quite arbitrary. They can be dispersed within the information bits or kept together and placed on either side of the information bits. A codeword, whose information bits are kept together, is said to be systematic code. A codeword, whose parity bits are dispersed within the information bits, is referred to as non-systematic code.
The Hamming Weight or weight of a word ‘v’ is defined as the number of nonzero components of ‘v’ and is denoted by \( w(v) \). The Hamming Distance or distance between two words ‘\( v_1 \)’ and ‘\( v_2 \)’, having the same number of bits, is defined as the number of places in which they differ and is denoted by \( d(v_1, v_2) \).

For example, the words \( v_1 = (011010) \) and \( v_2 = (101000) \) have weights of 3 and 2 respectively and are separated by a distance of 3.

The minimum distance (\( d_{min} \)) of a block code is the smallest distance between codewords. Hence codewords differ by \( d_{min} \) or more bits. The minimum distance is found by taking a pair of codewords, determining the distance between them and then repeating this for all pairs of different codewords. The smallest value obtained is the minimum distance of the code.

Hamming Distance is used to assess the error control ability of a code. The error correction limits ‘\( n \)’ and error detection limits ‘\( l \)’ are bounded by Hamming Distance or minimum distance (\( d_{min} \)) of a code. Codes with error correction limit ‘\( t \)’ and error detection limit ‘\( l \)’ are referred to as \( t \)-error correcting codes and \( l \)-error detecting codes respectively. Mathematically, \( l = \lfloor d_{min} - 1 \rfloor \) and \( t = \lceil \frac{1}{2} (d_{min} - 1) \rceil \).

A word with \( n \)-bits can be represented by a vector with \( n \)-components. For example, a 4-bit word can be \( (1010), (1110), (0010) \) and so on. A set of words is called code. For an \( (n, k) \) block code the input to the encoder is the information word \( i = (i_1, i_2, i_3, ..., i_k) \) where \( i_j = 0 \) or \( 1 \) and \( j \) is an integer \( 1 \leq j \leq k \). The encoder determines \( r = n-k \) parity check bits \( p_1, p_2, ..., p_r \) according to the encoding rule of the code and appends them to the information bits or disperse them among the information bits so giving the codeword as shown below.

\[
c = (i_1, i_2, i_3, ..., i_k, p_1, p_2, ..., p_r)
\] (3.1)
The encoding rule of a code is such that the combination of the parity bits and the information bits (i.e. the codeword) has the mathematical property required by the code.

It is usual to represent codeword as 
\[ c = (c_1, c_2, \ldots, c_n) \]  
(3.2)

where \( c_j = i_j \) for \( 1 \leq j \leq k \) and \( c_j = p_{j-k} \) for \( k < j \leq n \).

\[ e = (e_1, e_2, \ldots, e_n) \]  
(3.3)

where \( e_j = 1 \), if there is an error in the \( j^{th} \) position or \( e_j = 0 \), if the \( j^{th} \) position is error free. A codeword \( 'c' \) that incurs an error \( 'e' \) results in the word as \( v = c + e \) where the components of \( 'v' \) are given by the components of \( c \) and \( e \) added pair wise as
\[ v = (c_1, c_2, \ldots, c_n) + (e_1, e_2, \ldots, e_n) \]
\[ = (c_1+e_1, c_2+e_2, \ldots, c_n+e_n) \]
\[ = (v_1, v_2, \ldots, v_n) \]  
(3.4)

where \( v_j = c_j + e_j \) for \( 1 \leq j \leq n \) and where modulo-2 addition is used when adding \( c_j \) and \( e_j \) together.
The word ‘\(v\)’ represents the codeword after it has been subjected to the error ‘\(e\)’. If all the components of \(e\) are zero, then \(v_1 = c_1, v_2 = c_2, \ldots, v_n = c_n\) and therefore, \(v = c\) \hspace{1cm} (3.5)

The equation \(v = c + e\) is central to the decoding process. A decoder has no prior knowledge of ‘\(c\)’, the only information that it has is the word ‘\(v\)’ that it receives. It is referred to as the received word or decoder input. For an error detecting code, the task of the decoder is to establish whether ‘\(v\)’ is a codeword. This decoding can be achieved by checking ‘\(v\)’ against a table of codewords or by checking whether ‘\(v\)’ has the mathematical property required by the code.

For an error correcting code, the decoder has to estimate or guess the codeword from ‘\(v\)’. If the decoder’s estimate of the error pattern is \(\hat{e}\) then, from the equation \(v = c + e\), its estimate of the codeword is \(\hat{c} = v - \hat{e}\), and if, modulo-2 addition is used, then \(\hat{c} = v + \hat{e}\) is the decoder’s estimate of ‘\(c\)’. However, whether a decoder can determine the correct codeword from ‘\(v\)’ depends upon the code, the errors incurred and the decoding algorithm.

Finite fields are referred to as Galois Fields after the mathematician Evariste Galois (1811-1832). It is used to identify the set of components within the field to perform the addition and multiplication operation of a code. The fields are usually expressed as \(GF(p^m)\) where \(p\) is the number of elements in the base field, which is referred to as the field’s characteristics and \(m\) is the degree of the polynomial whose root is used to construct the fields. The order of the field is given by \(q = p^m\). In the digital communication, where binary digits are only employed, \(p\) is always 2 and \(m = 1, 2, 3, \ldots\) so on. In normal course, block codes are usually represented with \(GF\) fields as \((n, k, q)\) block code. For example, if \(q = 2^l\), then the block code is bounded by 0s and 1s.
only as field components. Otherwise, if $q=2^3$, the field components are $0, 1, \alpha, \alpha^2, \alpha^3, \alpha^4, \alpha^5$ and $\alpha^6$ where $\alpha$ is the root lies within a finite field $GF(2^3)$.

Figure 3.2 [105] shows the concept of Binary Symmetric Channel (BSC). It is a channel that transports 1’s, and 0’s from the transmitter ($T_x$) to the receiver ($R_x$). It makes an error occasionally, with probability $p$.

![Figure 3.2 Binary Symmetric Channel Model](image)

A BSC flips a 1 to 0 and vice-versa with equal probability. Let $X$ and $Y$ be binary random variables that represent the input and output of this BSC respectively. Let the input symbol be equally likely, and the output symbols depend upon the input according to the channel transition probabilities as given below.

\[
P(Y=0|X=0) = 1-p \tag{3.6}
\]
\[
P(Y=0|X=1) = p \tag{3.7}
\]
\[
P(Y=1|X=1) = 1-p \tag{3.8}
\]
\[
P(Y=1|X=0) = p \tag{3.9}
\]

Above equations (3.6) to (3.9) implies that the probability of a bit getting flipped (i.e. in error) when transmitted over this BSC is $p$. 
3.2.1.1 Significance of Focusing on Binary Symmetric Channel over other Noisy Channels

Crepeau (1997) has described that both commitment schemes and oblivious transfer protocols are realizable in some Binary Symmetric Channel (BSC) models [21], [22], in which random noise is added to these channels in both directions with some known probability. One of these simple and reliable Binary Symmetric Channels with known probability is shown Figure 3.2.

Mangesh Abhimanyu Ingale (2003) has mentioned [65] that the improved transmission capacity in fiber-optic systems has drawn the attention of error correcting coding researchers in the field of communication systems to develop efficient Forward Error Correction (FEC) for optical communication systems in the recent years. It has also noted that the International Telecommunication Union (ITU-T G.975) has recommended Reed-Solomon RS (255, 239) Code to be used in most complex optical communication systems. This thesis described error correction analysis of the RS (255,239) code for completely random distribution of errors over an Additive White Gaussian Noise (AWGN) channel with Binary Phase Shift Keying (BPSK) signaling and hard decision decoding. The Reed-Solomon Codes (RS Codes) are mainly used for burst error correction on the codewords when they are transmitted over more complex noisy channels unlike Binary Symmetric Channels [132]. However, the burst error correction capability of the RS code can be resembled by using our proposed Enhanced Hamming Code as described in the sections from 3.3 to 3.7.

Later, Damgard, I., Fehr, S., et al. (2004) showed that commitment schemes and oblivious transfer protocols can also be implemented, under certain conditions, in the weaker and Unfair Noisy Channel model [27],[28]. However, in theses channels, the error probability is not known precisely to the honest parties, and furthermore can be influenced by the third party other than the sender and receiver. Therefore, selection
of Unfair Noisy Channel model for developing cryptographic algorithm may reduce the feasibility of designing good Cryptosystems or it would make the cryptography algorithm to have the complex structures and complicated implementations.

Tal Moran and Gil Segev (2008) have mentioned in their paper [124] that the researchers already have done some contribution for finding commitment schemes and oblivious transfer protocols on the physical properties of communication channels. These contributions are mainly performed by using the random noise in a communication channel as the basis for security.

Jurg Wullschleger (2009) said that, in a cryptographic context, noisy channels can become a valuable resource. The basic idea of implementing all the cryptographic protocols over the particular noisy channel is very similar to all the noisy channels [49]. For example, first, the researchers construct some kind of erasure channel. Then, this erasure channel is frequently used to implement Oblivious Transfer (OT). The correctness and the security are guaranteed by proper selection and usage of error correcting codes and developing privacy amplification schemes to strengthen the cryptographic algorithm.

3.2.2 ASCII Character Set

The American Standard Code for Information Interchange (ASCII) assigns values between 0 and 255 for upper case letters, numeric digits, punctuation marks, and other symbols. ASCII characters can be split into the following sections:

- **0 – 31** for Control codes;
  
  Examples are 000 – NUL, 002 – SOH, 027 – ESC

- **32 – 127** for Standard, implementation – independent characters;
  
  Examples are 032 – Space, 049 – 1, 065 – A, 097 – a, 127 – delete
• 128 – 255 for Special symbols, international character sets, non-standard characters; Examples are 131 - f, 137 – %, 251 - ©

Among the above whole ASCII characters, 32 – 127 are the standard, independent alphanumeric characters being used by the common people for every day.

3.2.3 Error Probability in Channel Coding

Definition 1 [105]: Let $C$ be an $(n, k)$ code over $GF(q)$ and ‘$a$’ be any vector of length ‘$n$’.

Then the set

$$ a + C = \{a+x \mid x \in C\} $$

(3.10)

The above equation is called a coset of $C$. ‘$a$’ and ‘$b$’ are said to be in the same coset if $(a-b) \in C$.

Definition 2 [105]: The vector having the minimum weight in a coset is called the coset leader. If there are more than one vector with the minimum weight, one of them is chosen at random and is declared the coset leader.

Definition 3 [105]: A standard array for an $(n, k)$ code $C$ is a $q^{n-k} \times q^k$ array of all vectors in $GF(q)^n$ in which the first row consists of code $C$ with ‘0’ on the extreme left. In this same array, other rows are formed by the cosets $(a_i + C)$, each arranged in corresponding order, with the coset leader on the left.

Definition 4 [105]: The probability of error or word error rate $P_{err}$ for any decoding scheme is the probability that the decoder output is a wrong codeword. It is also called as the residual error rate.
Suppose there are $M$ codewords of length $n$ which are used with equal probability. Let the number of coset leaders with weight $i$ be denoted by $\alpha_i$. Here BSC channel is assumed with symbol error probability $p$. A decoding error occurs if the error vector $e$ is not a coset leader. Therefore, the probability of correct decoding will be

$$P_{\text{cor}} = \sum_{i=0}^{n} \alpha_i p^i (1-p)^{n-i}$$

(3.11)

Hence, the probability of error will be

$$P_{\text{err}} = 1 - \sum_{i=0}^{n} \alpha_i p^i (1-p)^{n-i}$$

(3.12)

### 3.3 Construction and Properties of Mixed Parity Codes

**Definition 5 [63]:** For any vector $u$ in $GF(q)^n$ and any integer $r \geq 0$, the sphere of radius $r$ and centre $u$, denoted by $S(u,r)$, is the set:

$$\{v \in GF(q)^n | d(u,v) \leq r\}$$

(3.13)

**Theorem 1 [63]:** A sphere of radius $r$ $(0 \leq r \leq n)$ contains exactly following number of vectors.

$$\begin{bmatrix} n \\ 0 \end{bmatrix} + \begin{bmatrix} n \\ 1 \end{bmatrix} (q-1) + \begin{bmatrix} n \\ 2 \end{bmatrix} (q-1)^2 + \ldots + \begin{bmatrix} n \\ r \end{bmatrix} (q-1)^r$$

(3.14)

**Theorem 2 [63]:** A $q$-ary $(n, k)$ code with $M$ codewords and minimum distance $(2t +1)$ satisfies,

$$M \left\{ \begin{bmatrix} n \\ 0 \end{bmatrix} + \begin{bmatrix} n \\ 1 \end{bmatrix} (q-1) + \begin{bmatrix} n \\ 2 \end{bmatrix} (q-1)^2 + \ldots + \begin{bmatrix} n \\ t \end{bmatrix} (q-1)^t \right\} \leq q^n$$

(3.15)
Definition 6 [63]: For binary codes, the Hamming bound will be

\[
M \left( \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{t} \right) \leq 2^n
\] (3.16)

Definition 7 [63]: A perfect code is one which achieves the Hamming bound as,

\[
M \left( \binom{n}{0} + \binom{n}{1} (q-1) + \binom{n}{2} (q-1)^2 + \ldots + \binom{n}{t} (q-1)^t \right) = q^n
\] (3.17)

3.3.1 Binary Hamming Codes

The Binary Hamming Codes have the property that \((n, k) = (2^m - 1, 2^m - 1 - m)\) where ‘m’ is any positive integer [37]. For example, for \(m=3\), \((7, 4)\) code is the Hamming Code. The parity check matrix of an \((n, k)\) code has ‘\(n-k\)’ rows and ‘\(n\)’ columns. For the binary \((n, k)\) Hamming Code, the \(n = 2^m - 1\) columns consist of all possible binary vectors with \(n-k = m\) elements, expect the all zero vector. The minimum Hamming Distance \(d_{\text{min}}\) of a \((7, 4)\) Hamming Code is equal to 3, which implies that it is a single – error correcting code. Hamming Codes are called as Perfect Codes.

3.3.2 Modified Binary Hamming Codes

By adding overall parity bit, an \((n, k)\) Hamming Code can be modified to yield an \((n+1, k)\) code with \(d_{\text{min}} = 4\). Also, an \((n, k)\) Hamming Code can be shortened to an \((n- l, k – l)\) code by removing \(l\) columns of its parity check matrix \(H\).

Definition 8: The formal definition of Hamming Code is given as,

Let \(n = (q^k - 1) / (q -1)\). The \((n, k)\) Hamming Code over \(GF (q)\) is a code for which the parity check matrix has columns that are pair wise linearity independent over \(GF (q)\).
3.3.3 Limitations on Codeword Block Length

According to Shannon’s theorem, if $C(p)$ represents the capacity of a BSC with probability of bit error equal to ‘p’, then for arbitrarily low probability of symbol error, code rate $R$ must be less than $C(p)$.

Even though the channel capacity provides an upper bound on the achievable code rate as $R = k/n$, evaluating a code exclusively against channel capacity may be misleading. The block length of the code, which translates directly into the delay, is also an important parameter. It has been observed that, if block length of the code is increased, the bounds on code rate are closer to channel capacity.

However, longer block lengths imply longer delays in decoding. The reason for this delay is because decoding of a codeword cannot begin until the receiving of the entire codeword. The maximum delay allowable is limited by practical constraints. For example, a codeword with very large block lengths cannot be used in mobile radio communications where the packets of data are restricted to fewer bits.

3.3.4 Construction of Mixed Parity Codes

Let us consider an $[n, k, q]$ block code on the binary Galois field $GF(q)$ of order 2, where $n = r + k$ and ‘n’ is the number of bits in output Crypto-codeword, ‘k’ is the number of input message bits, and ‘r’ is the number of parity bits.

Then the Mixed Parity Codes are defined as follows.

**Definition 9:** A $[B, k, q]$ code $C$ where block number $B = [(r/k) + 1]$ is said to be a Mixed Parity Code with the encoding operation $\mathcal{O}$, if it satisfies the following two conditions, for all $i$ and $j$ with $i \neq j$, $i, j \in \{1, 2, 3, \ldots, (1-2^k)\}$:

As the first condition, $B = n/k \geq 2$ \hspace{1cm} (3.18)
As the second condition, \( \min H_d(f(m_i), f(m_j)) = \min H_d(f(C_i), f(C_j)) \) \hspace{1cm} (3.19)

where ‘\( m \)’ is the input message and ‘\( C \)’ is the Crypto-codeword generated by the function \( f(O(m)) \)

Equation (3.19) shows that block number ‘\( B \)’ of \( O \) is lower bound by 2 for the existence of MP codes. Because minimum output length of a Crypto-codeword is to be double the length of input message bits for efficient Crypto-coding of MP codes. For the construction of mixed parity code, this output block length \( (n) \) is achieved by mixing the parity bits with the input block length \( (k) \) such that the minimum bound of \( B \) is 2.

Block number \( B \) must be a high value for the good security of the Mixed Parity Code. Also, it should be a real number to avoid concatenation of extra redundancy bits at Crypto-codeword output for message interpretation.

The \( (n, k) \) Mixed Parity Codes (MP Codes) are the resemblance of \( (n, k) \) Binary Hamming Codes where the parity bits are dispersed among the information bits unlike the traditional method of appending the parity bits to the left most or right most sides of information bits. The purpose of deliberately dispersing the parity bit among the information bits is to scramble the originality of the information bits for crypto-encoding function without diluting the error control capability of the MP codes. It is the responsibility of the decoder to identify the position of parity bits to control over the error correction on received word vector ‘\( v \)’ after performing crypto-decoding.

These MP Codes use the common character set, such as ASCII character set, for performing Crypto-coding. This MP code will be a tool to design the Crypto-coding system. In this paper, ASCII character set is considered for analysis. Because any
character in ASCII character set can be segregated as BCD format that is required for error correction technique. For example, the symbol ‘A’ is represented as 65 and 01000001 in decimal and binary formats respectively. Suppose, 01000001 are grouped in terms of 4-bit BCD code, there will be two 4-bit BCD codes as 0100 and 0001. Also, we may have the more different set of two 4-bit BCD combinations out of this ASCII code such as 0000 : 1000, 1000 : 0010 and so on. This process can be called as BCD Coded ASCII (BCA).

In common, if any character in ASCII set is represented by $A = a_7a_6a_5a_4a_3a_2a_1a_0$, then the possible set of 4-bit BCD combinations are $a_7a_6a_5a_4 : a_3a_2a_1a_0$, $a_4a_5a_6a_7 : a_3a_2a_1a_0$, $a_7a_6a_5a_4 : a_6a_1a_2a_3$ and so on.

If (7, 4) Hamming Code is used to add the parity bits with each 4-bit BCA, a 4-bit BCA can be stretched to 7-bit longer by appending or dispersing three parity bits for error correction. An even or add parity is added for this whole 7-bit at the LSB and then there will be a new 8-bit code. This way, a two 8-bit codewords can be generated for each ASCII character.

For example, if an ASCII code 01000001 is converted BCA as 0100: 0001, then applying (7, 4) Hamming Code technique, 0100 is converted as 0100111 and 0001 is converted as 0001011. Further, even parity is chosen to be placed at the MSB of resultant codes, then 0100111 becomes $A_1 = 00100111$ and 0001011 becomes $A_2 = 10001011$.

If $A_1$ and $A_2$ are interpreted as ASCII Coded BCD (ACB), 39 is the decimal value of ASCII character set for $A_1$ and its equivalent message is ‘’ (single quote). Similarly for $A_2$, after removing even parity bit at MSB, 11 is the decimal value of ASCII character set, and its equivalent message is ‘VT’ (Vertical Tab \v).
For the same ASCII character ‘A’, if some other combination of two four bit BCD is interpreted, we will have two different ACB codes. Similarly, if parity bits are dispersed among the information bits, again there will be new sets of ACB codes. Further, some bits are intentionally inverted with the limits bounded by type of code chosen for adding parity bits. For example, (7, 4) Hamming Code can allow one-bit inversion among a 7-bit codeword that can be identified and corrected during crypto-decryption. These operations can be selectively performed by crypto-encoder based on the selective keys to ensure the security of transmission of information from one end to another.

**3.3.5 Criteria for Mixed Parity Code**

The main purpose of our work is to construct a code that possesses combined error correction and cryptographic properties. This code should not compromise the required error control and security strength. Hence three measures are developed to satisfy these requirements. They are (1) Criterion for security, (2) Criterion for error control and (3) Criterion for Crypto-coding. A good Crypto-coding must satisfy all these three measures.

The proposed code will be mainly used for providing some effects, such as bit error penetration and message misinterpretation, on the huge section of the output block for an intruder without having the correct key. Block number ‘B’ is introduced to measure the message dissemination and bit error penetration rates.

The block number of a crypto-encoding function $\mathcal{O}$, with the input vector $x$ and the output vector $\mathcal{O}(x)$ is defined as below.

$$B = \frac{\text{length of } \mathcal{O}(x_i)}{\text{length of } x_i},$$

(3.20)

where $i \in \{1, 2, 3, \ldots, (1-2k)\}$ and ‘k’ is the total number of bits per input message bits.
The error control rate of the proposed code is determined by the pair wise Hamming Distance between the set of Crypto-codewords. If the Hamming Distance is high value, it guarantees a large amount of error control rates for both error detection and error correction.

The MP code must satisfy the following condition for reliable error correction capability.

If \( \Theta(x_i) \) is the Crypto-codeword generated by MP code of length \( n = b \cdot k \),

\[
\min H_d \{ f(x_i), f(x_j) \} = \min H_d \{ f(\Theta(x_i)), f(\Theta(x_j)) \},
\]

where \( i \neq j, i, j \in \{1,2,3,\ldots,(1-2k)\} \) and \( H_d \) is the Hamming Distance

In equation (3.21), \( f(x_i) \) is the function that selects the number of bits equal to one character bits of ASCII code among the input \( x_i \). Also, \( f(\Theta(x_i)) \) is the function that selects the number of bits equal to one character bits of ASCII code among the output \( \Theta(x_i) \).

In general, the number of bits, equal to one character bits of ASCII code, is seven.

It is described through increasing the maximum number of probabilities for retrieving message bits without using the correct key and high value of Hamming Distance for Crypto-codewords with good error control capability. These properties can be accomplished by the following three methods.

(1) wrapping the Crypto-codeword length by dispersing parity bits among message bits, (2) intentional bit inversion over Crypto-codewords based on Hamming Distance, and (3) increasing the security reliability of Crypto-coding by doing the bitwise
Boolean XOR operation, bit rotation and bit permutation functions over the Crypto-codewords.

3.3.6 Properties of Mixed Parity Codes

In this section, we illustrate that the MP Codes possess the maximum possible message dissemination and error correction capability as desired in the design requirement.

As per the definition of MP Codes, it has the block number \( B \) equal to \( n/k \). Also, the Crypto-coding operation \( \emptyset \) from ‘\( k \)’ bits into ‘\( n \)’ bits is done by the following three consecutive operations.

1. performing error control functions such as parity bits generation of MP codes, (2) outputting the Crypto-codeword of length ‘\( n \)’ by random parity bits mixing with input message bits, and (3) performing key-based cryptographic functions such as bits permutations, bits rotations and Modulo-2 arithmetic functions over Crypto-codewords. Randomness in all these functions is the vital part of a disseminating operation of the proposed code. The strength of this disseminating power is proportional to the block number \( B \).

For example, if the input message bits length \( k \) is 63 as there are nine ASCII characters in \( k \) and each character having seven binary bits, then the output block length \( n \) is 126, where the block number is 2. If these bits are separated by seven binary bits each, then there are eighteen ASCII characters in the output block length ‘\( n \)’.

The optimal error correction capability of Mixed Parity Codes depends on the selection of existing channel code for construction of Mixed Parity Codes and effective usage of their parity bits.
Theorem 3: An \([n,k,q]\) MP Code ‘C’ with encoding operation \(\varnothing\) is a Maximum Disseminating Code with a number of correctable error \('t'\) is directly proportional to \(b(\varnothing)\).

Proof: Generally, if \(H_d(C_i, C_j) = h\), where \(h\) is the Hamming Distance and \(i \neq j, i, j \in \{1,2,3,..., (1-2^n)\}\); Then, error correction limit \((t)\) of \((7, 4)\) Hamming Code is

\[ t = \lfloor (h-1)/2 \rfloor = 1 \text{ since } h=3. \] (3.22)

Equation (3.22) shows that 1-bit error correction is possible for every 7-bit codeword.

If the \((56,28)\) MP code ‘C’ with \(b(\varnothing)= 2\) is constructed using \((7, 4)\) Hamming Code as a base code, each codeword has 56-bit out of which eight ASCII characters can be realized for every 7-bit. Therefore, eight error corrections can be achieved among the 56-bit at the rate of 1 error correction for every 7-bit.

Then, error correction capacity of MP codes derived from \((7, 4)\) Hamming Code is

\[ [(n/7).t] \text{ where } n = (b(\varnothing).k) \] (3.23)

This will lead to the probability of altering all the bits but limited to one bit per 7 symbols which is bound by chosen Hamming Code for construction of MP Codes. However, if error correcting capability \((t)\) of chosen codes is further increased, then the probability of inverting the bits per set of symbols can be greatly increased.

The potential in good Crypto-coding based on MP Code depends on proper selection of existing channel code. Since the priority of our design is cryptography through error penetration and correction capabilities by mixed parity bits, the chosen code must possess the good message misinterpretation phenomenon on the Crypto-codewords of...
MP code. For example, if the input block is collection of ASCII values, Hamming Code will exhibit all the required properties of Crypto-coding.

Further, the strength of Crypto-coding through MP Code can be increased by key based randomization in parity bit addition and bit inversion on the output bits of MP code. Also, if any Crypto-coding operation on output block produces the meaningful message that will lead to the data misinterpretation to the intruder. This is an added advantage to the strength of Crypto-coding. The strength of good MP codes depends on the process of resemblances of original message bits from the codeword of MP codes.

The basic requirement for existence of MP Code is $n = [b.k]$ where $b$ is block number, ‘$n$’ is the length of the codeword and ‘$k$’ is the length of message over Galois field of order ‘$q$’ i.e., $[GF (q)]$. The following considerations are valid for $GF (q)$ for all MP code.

**Definition 10:** $q^x \geq q + 1$ is valid for any $q > 1$ and $x \geq 1$, Then, there exist an $(n, k, q^x)$ MP code with the condition $n > k > x$ for finite Galois field.

**Definition 11:** All the $2^k$ messages can always be assigned a codeword of length $(b.k)$ such that all the codewords are in the domain of all the possible input messages of $2^k$. This property is one of the major requirements of Crypto-coding for the purpose of message misinterpretation towards the cryptanalysis by the intruder.

If $(8, 4, 2^1)$ MP Code is used, then $2^8 = 2^4 = 16$ and $b = 8 / 4 = 2$. Therefore, all the 16 Crypto-codewords will assume 16 ASCII characters where each Crypto-codeword is one among the 128 combinations between 0-127 decimal values of ASCII character set.
Proof:

we show that \[ b. (1- \frac{r}{n}) \] = 1 where the factor \( 1- \frac{r}{n} \) is the probability of existence of original message symbols among the total length of the output symbol ‘n’.

The probability factor \( 1- \frac{r}{n} \) multiplied with block number \( b \) always yields the result one such that the function \( O \) produces a multiple numbers of message blocks in the output of MP Code. Again this will lead to the required data misinterpretation towards the cryptanalysis by the intruder.

By substituting \( k = n - r \) for \( n = [b \cdot k] \) gives \( n = [b \cdot (n-r)] \) where ‘r’ is number of redundancy bits among ‘n’.

If equation (3.24) is rearranged, we will get \[ b \cdot (1-\frac{r}{n}) \] = 1

(3.25)

**Theorem 4:** For a given \([n,k,q]\) MP Code over \( GF(2^l) \) where \( n \) is one output block size and \( k \) is one input block size, and \( X \) is the total number of output block sizes of MP Code generated by the Crypto-encoding function \( O \) where \( n / b = k \),

Then the function, \( (X_j / k) \) always produces ‘m’ number of input block size in the finite set of \( x \in \{ x_1, x_2, x_3, ..., x_p \} \) where \( p = (1-2^k) \) and \( j \in \{1,2,3, ..., (1-2^k)\} \). Also, ‘m’ is always divisible by the block number \( b \).

Therefore, \( f(O(x_i)) = X_j = \{x_{i1} \parallel x_{i2} \parallel x_{i3} \parallel ... x_{im}\} \)

(3.26)

where \( m = X_j / k \) and \( i \in \{1,2,3, ..., (1-2^k)\} \).

For example, if two output blocks of length \( X_j \) 112 is generated by the \((56, 28, 2^l)\) MP code,
then, \( m = 112 / 28 = 4 \) and \( m / b = 2 \).
3.3.7 Choice of Existing Codes for Construction of Mixed Parity Codes

The major consideration for construction of MP Code mainly depends on its usage for Crypto-coding. Since the MP Code will undergo traditional cryptographic processes such as bit permutation, bit rotation, bit wise Boolean XOR operation, it will show the security strength against well-known attacks such as plaintext attack, ciphertext attack, and brute force attack. Also, the error correction capability of the MP code must play the dual role. If redundancy bits are not flipped purposefully at the sender side, the whole error correction capacity of MP code will be retained for receiver side to correct the errors occurred during the transmission of Crypto-codewords over noisy channel.

Table 3.1 Comparison of Various Codes for their Error Correction Features

<table>
<thead>
<tr>
<th>Type of code</th>
<th>Block size ( (n, k) )</th>
<th>Galois field ( GF(q) )</th>
<th>Number of output bits in a codeword ( n )</th>
<th>Number of error correction bits</th>
<th>Error correction capacity ( {[(5) / (4)] \times 100} % )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hamming Code (Single error correction)</td>
<td>((7, 4))</td>
<td>(GF(2^1))</td>
<td>7</td>
<td>1</td>
<td>14.28571</td>
</tr>
<tr>
<td>BCH Code (Single error correction)</td>
<td>((15, 11))</td>
<td>(GF(2^4))</td>
<td>60</td>
<td>4</td>
<td>6.66667</td>
</tr>
<tr>
<td>BCH Code (Double error correction)</td>
<td>((15, 7))</td>
<td>(GF(2^4))</td>
<td>60</td>
<td>8</td>
<td>13.33333</td>
</tr>
<tr>
<td>Reed-Solomon Code (Single error correction)</td>
<td>((7, 5))</td>
<td>(GF(2^3))</td>
<td>21</td>
<td>3</td>
<td>14.28571</td>
</tr>
<tr>
<td>Reed-Solomon Code (Double error correction)</td>
<td>((15, 11))</td>
<td>(GF(2^3))</td>
<td>45</td>
<td>6</td>
<td>13.33333</td>
</tr>
</tbody>
</table>
But, if part of the redundancy bits is flipped purposefully so as to increase the message dissemination capacity at the sender side using key based functions, then the receiver side will have the reverse key functions to retrieve the original redundancy bits. Then the error correction technique is applied to restore the original message.

Since the common message input format to computers is based on ASCII values whose binary symbol length is seven for each character, the (7, 4) Hamming Code is chosen here to construct the MP Code. The MP Codes can be constructed by using any one of the existing channel codes. However, the choice of selection depends on number of symbols in the input block, number of symbols in the output block, possession of cyclic and non-linearity properties after adding redundancy bits, maximum number of bit penetration with minimum number of bit inversion. As for as MP code is concerned, the importance is given to plaintext dissemination and misinterpretation with parity bits mixing by key-based function rather than simply the length of the output with parity bits during the outcome of encryption process.

Also, Table 3.1 shows that the selection of (7, 4) is the better choice among some of the other existing codes as for as specified selection of input symbol block length is concerned. Also, it reveals that if the behavior of the communication channel is known, choice of Hamming Code is better than any other commonly used codes in terms of required number of reduced redundancy bits for Crypto-coding with specified length.

The block size is so chosen that the output length of the block must be able to produce the meaningful ASCII characters by subdividing total length in terms of seven symbols per characters. This is the function that is required for the property of plaintext misinterpretation at out Crypto-coding technique. So, the total number of input bits must be divisible by both four and seven as well as the total number of output bits must be divisible by seven. Therefore, a number of binary input symbols per encoding
function $\mathcal{O}$ can be 28, 56, and 84… so on. The increase in the input block size leads to the increase in the probability of combinations of characters using output block size.

If we choose, 4 ASCII characters for the encoding function $\mathcal{O}$, then the input block size will be 28. Then, (7, 4) Hamming Code is used for generation of parity bits, there are 21 parity bits for each 4-bit of input bits. Now, the resultant bits are multiples of 7. That is, in our case, resultant bits are 49. Again another 7 redundancy bits can be generated such that one odd or even parity bit for every 7 symbols. This last 7 set of parity bits employs the vital role in bit inversion and rotation process in Crypto-coding. Because, if we are able to retrieve the first generated 49-bit without any error on the decryption side, even though the whole 7-bit are inverted at encryption side, we will retrieve it using respective parity check algorithm. So, if this scrambled parity symbols block of 7-bit is effectively used for bit error penetration throughout the encryption process at Crypto-coding, which will strengthen the overall security of the proposed MP code. Here we have generated 8 ASCII characters with codeword size of 56-bit as the outcome of function $\mathcal{O}$ from the input of 4 ASCII characters with message size of 28-bit. So, the bit magnify number for the described example is 2. If the bit magnify number is further increased, the number of outcome of ASCII character also gets increased. This function will show the strength to the Crypto-coding when the generated parity bits and ASCII characters are placed randomly among themselves to have message dissemination.

### 3.3.8 Error Correction Mechanism of Mixed Parity codes

As illustrated in Table 3.1, Hamming Code possesses the maximum efficiency with specified number of bits. However, it has the limitations on bulk error correction in the single codeword. But, as described in the section 2.2.2, randomly placed parity bits along with the message bits may show bulk error correction probability up to some extent over the controllable noisy channel or known communication channel. In our
example described in the above section 3.3.6, with the aid of Table 3.1, we can conclude that there is the possibility of correcting the maximum of 7-bit but limited to one bit per seven binary symbols of first 49 parity mixed message bits. Remaining seven parity bits show the strength to the bulk error correction as that can be corrected with the knowledge of correctly recovered first 49-bit.

Since the good Crypto-coding requires the non-systematic successive outputs, concatenation of parity bits to the message bits in MP codes need not be a systematic operation. So, these parity bits can be randomly placed in anywhere among the total length of output bits. But, only the knowledge of placement of parity bits is to be preserved such a way that it can be identified at the decoding process for retrieving original message from codeword or ciphertext.

**3.3.9 Security of Crypto-coding using Mixed Parity Codes**

The codeword output of the MP codes is separable into distinct words in terms of specified input symbol size. It provides the greater plaintext permutation. Further, if error correction capability of the MP code is decisively used to perform the selective bit inversion process, it will provide the higher degree of non-linearity in the output of Crypto-coding.

The procedure of parity bits mixing with message bits makes the MP Code as one of the most efficient tool for Crypto-coding. For example, the single codeword with 7-bit of (7, 4) Hamming Code will have three parity bits and four message bits. If these three parity bits are placed in between the message bits with different combinations, the total number of combinations is 210. This number of combinations can be further increased by increasing input length size. Therefore, if the key round is used to make these parity mixing combinations to preserve the reverse process at the decryption, this property of MP codes provides greater data misinterpretation for the cryptanalysis.
As described in the section 3.3.7, in built error correction technique of MP codes can be trickily used to strength the Crypto-coding algorithm. Although, if the length of the output block size is large enough, some of the mixed parity bits can be deliberately inverted to penetrate further data manipulation and rest of the mixed parity bits can be left out for overcoming the noisy communication channel. However, number of bits utilized for these two purposes depends on choice of the base code for the construction of MP code, input message block size and output codeword block size of selected MP Code.

If the functions, namely parity mixing, bit inversion, bit rotation and output block segregation with respect to input block size, are efficiently performed with the help of key rounds as required by cryptography, the traditional method of substitution boxes for bit permutation can be removed. The alternative for the substitution boxes is one of the major criteria for the present cryptography algorithm, as this technique requires the larger look-up table for high throughput and occupies a large amount of memory size hardware implementation of traditional algorithms.

Since the process of mixed parity bits provides the large number of combinations with minimum amount of bits, it shows the security strength against Known Plaintext Attack. Also, inclusion of bit inversion at encryption process and error correction at the decryption process with specified block size increases the probability of getting erroneous messages during Known cipher Attack. These two security properties collectively showing the strength to another well-known attack called Brute Force Attack which is mainly based on searching with all possible keys. Because, a single key can provide more number of combinations of message text for different Crypto-coding functions of MP codes such as parity mixing, bit inversion, bit rotation and output block segregation for the specified set of block. This will lead to the computationally infeasible process with Brute Force Attack to retrieve the original
message block with the stipulated time period as required by the present cryptography algorithm.

### 3.3.10 Factors Influenced on Mixed Parity Codes for good Crypto-coding

The following four factors are to be mainly considered for the design of cryptographically secured block cipher with the high rate of error resilience using Crypto-coding. (1) Choice of existing code-based on maximum error correcting capability with fixed length of input and output block size, (2) Choice of input block size based on the chosen character set of plaintext with relevant to the output block size of chosen code, (3) Proper placement of the parity bits among the plaintext with key based preservation for achieving high error control reliability over noisy channel, and (4) Round key based boolean functions with bit rotations for penetrating non-linearity in ciphertext

### 3.4 Extended Hamming Code and its Error Control Coding Techniques

A (8, 4) Extended Hamming Code ‘C’ is constructed by transforming all the 4-bit of information (i) over GF(2^4) into sixteen 8-bit codewords and

\[ C \in \{c_j : 0 \leq j \leq 15 \} \text{, where } c_j \text{ is a codeword} \]  

#### 3.4.1 Construction of Extended Hamming Code

If 4-bit information \( i = (i_1, i_2, i_3, i_4) \),

then codeword ‘c’ = \( iG = (i_1, i_2, i_3, i_4, p_1, p_2, p_3, p) \) \hspace{1cm} (3.28)

where generator matrix ‘G’ = 

\[
\begin{pmatrix}
1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\
0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\
\end{pmatrix}
\]

and \( p_1 = i_1 + i_2 + i_3, \ p_2 = i_2 + i_3 + i_4, \ p_3 = i_1 + i_2 + i_4, \ p = i_1 + i_3 + i_4, \) ‘G’ is 4x8 Generator matrix and ‘+’ is modulo-2 addition.
From the equation (3.28), for C is the (8, 4) Extended Hamming Code, then all the vectors over $GF(2^4)$ will have the 16 codewords of code C such that $C \in \{00,17,2d,3a,4e,59,63,74,8b,9c,a6,b1,c5,d2,e8,ff\}$.

If $v = (v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8)$, then, its error syndrome is $s = (s_1, s_2, s_3, s)$.

Here, $s = (v_1 + v_2 + v_3 + v_4 + v_5 + v_6 + v_7 + v_8)$ is the overall parity-check sum and $s_1, s_2, s_3$ are the same as the parity-check sums given by the (7, 4) code.

### 3.4.2 Randomized Double Error Correction Technique

Since the parity bits are separated uniquely as $s$, $s_1$, $s_2$ and $s_3$, the error correcting capability of the (8, 4) Extended Hamming Code has two distinguish schemes.

For example, let codeword 'c' = $(i_1, i_2, i_3, i_4, p_1, p_2, p_3, p)$ be (8, 4) Enhanced Hamming Code. Then the set of parity bits $s_1$, $s_2$, $s_3$ can be used to correct the single error among the codeword set $(i_1, i_2, i_3, i_4, p_1, p_2, p_3)$. After finding the correct set, using the equation $p = i_1 + i_3 + i_4$, error at the location 'p' of codeword 'c' can be found and corrected.

As mentioned above, this stipulated double error correction capability of (8, 4) Extended Hamming Code is effectively used for more permutation enhancement in the proposed design by optimal randomized bit placement of parity bits among message bits and its conditional intentional bit inversion with the modulo-2 addition of code’s error vectors as required in the efficient Crypto-coding technique.

### 3.5 Construction of Proposed Enhanced Hamming Codes

We construct the Enhanced Hamming Code from custom Hamming Codes by randomized parity bits placement among the message bits.
3.5.1 Formation of (7, 4) Enhanced Hamming Code

For Crypto-coding application, the choice of code depends on its linear and cyclic behavior. An Enhanced Hamming Code is developed as shown in Table 3.2. Each 7-bit codeword of (7, 4) Hamming Code would exhibit linear and cyclic properties for its error correction. However, concatenation of these codewords (Enhanced word) would not exhibit any mathematical properties.

Table 3.2 Enhanced Hamming Code Formation Scheme for Block Size (28, 16)

<table>
<thead>
<tr>
<th>Enhanced Codewords formation scheme of (7,4) Hamming Code from 16 Message Bits (For any one-bit inversion and its reversal per 7-bit codeword)</th>
<th>General form of (28,16) Enhanced Hamming Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parity Bits</td>
<td>Message Bits</td>
</tr>
<tr>
<td>P</td>
<td>C</td>
</tr>
<tr>
<td>11</td>
<td>27</td>
</tr>
<tr>
<td>10</td>
<td>21</td>
</tr>
<tr>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

In Table 3.2, there are four codewords are used to form the (28, 16) Enhanced Hamming Code. These codes can be effectively used to perform the proposed Cryptosystem schemes.

Selectively chosen Enhanced Hamming Codewords are not bounded by any linear or cyclic behavior. That is modulo-2 sum of any two words will not produce another codeword. Also, the cyclic shift of any codeword will not produce any other codeword. Some form of selectively chosen Enhanced Hamming Codewords are shown in Tables 3.3 and 3.4. These dual non-mathematical properties of Enhanced words of Enhanced
Hamming Code (EnHC) are effectively used for doing Crypto-coding. Because non-
mathematical properties based algorithms of Crypto-coding are not vulnerable to the
most of the known cryptanalysis at present.

### 3.5.2 Randomly Concatenated (7, 4) Enhanced Hamming Codes

It has been proved that, for any linear cum cyclic block code, there is an algorithm to
correct the limited number of errors on its codewords [105]. Therefore, all the possible
256 combinations of (14, 8) Enhanced Hamming Code from codewords of (7, 4)
Hamming Code is constructed to analysis its mathematical properties. Some of the
combinations of different Enhanced codewords are shown in Tables 3.3 and 3.4
illustrative purposes.

The mathematical analysis of each word in the above combination reveals that these
set of words are not obeyed to the linear and cyclic properties. That is the modulo-2
addition of any two words in Table 3.3 and 3.4 will not produce any other word of
Enhanced Hamming Code. Also, a cyclic shift of any word will not produce any other
word in an Enhanced Hamming Code.

According to the coding theory, any set of words that is not bounded by the linear and
cyclic properties would not have an efficient algorithm for its error correction. These
non-mathematical properties of the Enhanced Hamming Code is effectively used to
make the cryptography application on the above set of words by making deliberate
errors on them to dilute the originality of the initial key information.

### 3.5.3 Formation of (8, 4) Enhanced Hamming Code

An (8, 4) Enhanced Hamming Code is constructed from (8, 4) Extended Hamming
Code. If the overall parity bit is placed randomly among four message bits of (8, 4)
Extended Hamming Code, then the resultant code is called as (8, 4) Enhanced Hamming Code.

Table 3.3 Some Words of (14, 8) Enhanced Hamming Codes Scheme

| Formation of (14, 8) Enhanced Hamming Code from Codewords CW0 and CW1 (16 codewords among 256 non-repeated words) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| M7 | M6 | M5 | M4 | M3 | M2 | M1 | M0 | P5 | P4 | P3 | P2 | P1 | P0 |
| 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 1  | 0  | 0  | 0  | 1  | 1  |
| 0  | 0  | 0  | 0  | 0  | 0  | 1  | 0  | 0  | 0  | 0  | 1  | 1  | 0  |
| 0  | 0  | 0  | 0  | 0  | 0  | 1  | 1  | 0  | 0  | 0  | 1  | 0  | 1  |
| 0  | 0  | 0  | 0  | 0  | 1  | 0  | 0  | 0  | 0  | 1  | 1  | 1  | 1  |
| 0  | 0  | 0  | 0  | 0  | 1  | 0  | 1  | 0  | 0  | 0  | 1  | 0  | 0  |
| 0  | 0  | 0  | 0  | 0  | 0  | 1  | 1  | 0  | 0  | 0  | 0  | 0  | 1  |
| 0  | 0  | 0  | 0  | 0  | 0  | 1  | 1  | 1  | 0  | 0  | 0  | 1  | 0  |
| 0  | 0  | 0  | 0  | 0  | 1  | 0  | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| 0  | 0  | 0  | 0  | 1  | 0  | 1  | 1  | 0  | 0  | 0  | 0  | 0  | 0  |
| 0  | 0  | 0  | 0  | 0  | 1  | 1  | 1  | 0  | 0  | 0  | 0  | 1  | 0  |
| 0  | 0  | 0  | 0  | 1  | 1  | 1  | 0  | 1  | 0  | 0  | 0  | 0  | 1  |
| 0  | 0  | 0  | 0  | 0  | 1  | 1  | 1  | 0  | 0  | 0  | 0  | 1  | 0  |
| 0  | 0  | 0  | 0  | 0  | 1  | 1  | 1  | 1  | 0  | 0  | 0  | 1  | 1  |

For example, let codeword ‘c’ = (i1, i2, i3, i4, p1, p2, p3, p) be an (8, 4) Extended Hamming Code. Then, some of the codewords of (8, 4) Enhanced Hamming Code formats are:

‘c’ = (i1, i2, i3, p i4, p1, p2, p3), ‘c’ = (i1, p1, i2, i3, i4, p2, p3, p), ‘c’ = (i1, p2, p3, i2, i3, i4, p1, p).
Table 3.4 Some Words of (21, 12) Enhanced Hamming Codes Scheme

<table>
<thead>
<tr>
<th>$M_{11}$</th>
<th>$M_{10}$</th>
<th>$M_9$</th>
<th>$M_8$</th>
<th>$M_7$</th>
<th>$M_6$</th>
<th>$M_5$</th>
<th>$M_4$</th>
<th>$M_3$</th>
<th>$M_2$</th>
<th>$M_1$</th>
<th>$M_0$</th>
<th>$P_8$</th>
<th>$P_7$</th>
<th>$P_6$</th>
<th>$P_5$</th>
<th>$P_4$</th>
<th>$P_3$</th>
<th>$P_2$</th>
<th>$P_1$</th>
<th>$P_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

$M_0$ to $M_{11}$ – Message bits; $P_0$ to $P_8$ – Parity bits; (3 parity bits per 4 message bits)

3.6 Choice of Concatenated Enhanced Hamming Code for Proposed Design

An optimal cryptography algorithm requires a large amount of diffusion by bit inversion within the minimum word size. This requirement is improved by Concatenated Enhanced Hamming Code (CEnHC) other than existing codes as shown in Table 3.5. In this Table 3.5, (7, 4) cyclic code is equivalent to (7, 4) Hamming Code.

A (28, 16) CEnH Code can correct any 4-bit on 28-bit size. In particular, as shown in Table 3.6.
Table 3.5 Comparative Analysis of the Proposed Enhanced Hamming Code with other Codes for Different Block Size

<table>
<thead>
<tr>
<th>No. of bits available for inversion and correction</th>
<th>Required Block size over GF(2^4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cyclic code[105] over GF(2^4)</td>
</tr>
<tr>
<td>1</td>
<td>(7, 4)</td>
</tr>
<tr>
<td>2</td>
<td>(15,7)</td>
</tr>
<tr>
<td>3</td>
<td>-----</td>
</tr>
<tr>
<td>4</td>
<td>-----</td>
</tr>
</tbody>
</table>

Selective bit placement among 28-bit word will provide 4-bit burst diffusion and correction. For the same 4-bit correction capacity, Reed-Solomon Code [96] and BCH code [38], [11] take 60-bit. Burst error correction capacity of the enhanced code depends on the correct reassembling of the original enhanced code at the decoder to apply error control algorithm.

Table 3.6 Selective Bit Placement on (28, 16) Enhanced Hamming Code

| Selective bit placement on (28,16) Enhanced Hamming Code for any 4-bit inversion |
|-----------------------------------------------|---------------|
|                                               |               |
| P 11                                          | P 8           |
| P 5                                           | P 10          |
| P 7                                           | P 4           |
| P 9                                           | P 1           |
| P 3                                           | M 15          |
| P 0                                           | M 11          |
| M 7                                           | M 3           |
| M 10                                          | M 14          |
| M 6                                           | M 2           |
| M 14                                          | M 9           |
| M 8                                           | M 5           |
| M 12                                          | M 12          |
| M 4                                           | M 10          |
| M 0                                           |               |

Bit Shuffling of 4 codewords CW3, CW2, CW1, CW0

Since, (8, 4) Extended Hamming Code is one of the coset code of (8, 4) Enhanced Hamming Code as per the examples that are shown in section 3.5.3, these two codes are interchangeably used in this thesis to represent the (8, 4) block code.
3.7 Significance of Focusing Hamming Code and Enhanced Hamming Code over other Existing Codes in our Proposed Design

As described in the sections and 2.1 and 3.2.1, we have chosen the Binary Symmetric Channel (BSC) as the primary scheme of our proposed Crypto-coding. The best available error correcting code to describe this Binary Symmetric Channel is Hamming Code. As stated by the Jurg Wullschleger (2009), the correctness and the security is guaranteed by proper selection and usage of error correcting codes relevant to the chosen noisy channel and developing privacy amplification schemes to strengthen the cryptographic algorithm, we have developed the Enhanced Hamming Code from the concept of Extended Hamming Code [105]. To strengthen our choice of Hamming Code and Enhanced Hamming Code for the proposed design, we have provided the Table 3.1 in section 3.3.7 and Table 3.5 in section 3.6.

As shown in Table 3.1, the (7, 4) Hamming Code is a binary code over Galois field GF(2^1) that takes four message bits (k) to produce 7-bit codeword (n). This codeword can be used to correct any single bit error that may happen among it when this codeword is transferred over the noisy channel with known error probability like Binary Symmetric Channel. It seems that the error correction capacity of the (7, 4) Hamming Code is that 1-bit can be corrected for every 7-bit. The code with the similar error correction capacity on Table 3.1 is (7, 5) Reed-Solomon Code. But, this code is a non-binary code over the Galois Field GF (2^3). Generally, this code is used for burst error correction in a particular domain of its codeword. For example, (7, 5) Reed-Solomon Code takes 15-bit message word (k) to produce 21-bit codeword (n). Among these 21-bit codeword, the correct decoder of (7, 5) Reed-Solomon Code would be able to correct any one of the three consecutive set of bits error over it. Suppose, the codeword output of the (7, 5) Reed-Solomon Code is labelled as \( C_{20}, C_{19}, C_{18}, C_{17}, C_{16}, C_{15}, C_{14}, C_{13}, C_{12}, C_{11}, C_{10}, C_{9}, C_{8}, C_{7}, C_{6}, C_{5}, C_{4}, C_{3}, C_{2}, C_{1}, C_{0} \), then the three consecutive set of bits of them are \( \{C_{20}, C_{19}, C_{18}\}, \{C_{17}, C_{16}, C_{15}\}, \{C_{14}, C_{13}, C_{12}\}, \{C_{11}, \)
These seven sets may be called as sub-codewords. Therefore, \((7, 5)\) Reed-Solomon Code would be able to correct any one 3-bit error among the seven set of sub-codewords. This 3-bit error correction on any one of the sub-codewords is called as single error correction over \(\text{GF}(2^3)\). As described in the section 3.3.4, by using Hamming Code, appropriate selection of parity bits placement among the codewords of Hamming Code as Mixed parity Codes and proper designing of the decoder for that Mixed parity Codes, we can reassemble the Reed-Solomon Code-like codes with same complexity. Similarly, we can reassemble any other codes that have been described in Table 3.1. However, as shown in the same Table 3.1, there error correction capacity is less than that of \((7, 4)\) Hamming Code and \((7, 5)\) Reed-Solomon Code. Hence, the other codes are not consider for our proposed design and we have chosen the \((7, 4)\) Hamming Code with proper enhancement over it to penetrate the necessary cryptographic reliability in our design against known cryptanalysis.

As shown in Table 3.5, our proposed Enhanced Hamming Code can be further modified into Concatenated Enhanced Hamming Code (C\text{E}_{\text{HC}}) as described in the section 3.5.2. The last column of Table 3.5 shows that the Concatenated Enhanced Hamming Code can be used to form the different block size over \(\text{GF}(2^4)\) to have the required number of bits correction on their codewords. Also, this code has the minimum number of block size when compared to other codes for the same bits correction rate. For example, the Concatenated Enhanced Hamming Code needs only \((28, 16)\) block size, which takes 16-bit message word and produces 28-bit codeword for 4-bit error correction over this codeword. But, the Bose, Chaudhuri, and Hocquenghem (BCH) Code needs \((60, 44)\) block size, which takes 44-bit message word and produces 60-bit codeword and the Reed-Solomon Code (RS Code) needs \((60, 52)\), which takes 52-bit message word and produces 60-bit codeword for the same 4-bit error correction over their codewords. Therefore, the main advantage of our
proposed code is that it requires a minimum number of block size for the higher error correction rate over the Galois Filed $GF(2^l)$ when compared with other existing codes as far as implementation cryptographic properties are considered by Crypto-coding.

3.8 Summary

A newly developed Mixed Parity Code with its properties and applications were described for Crypto-coding technique. The bit magnify number was introduced to deal with the adoptability of the proposed Mixed Parity Code to be used as a tool for efficient implementation of Crypto-coding algorithm. Plaintext dissemination measured by block number ‘$B$’ for data misinterpretation and error resilience measured by minimum distance between codewords were described. It has been shown that the choice of existing code to construct the Mixed Parity Code depends on its error correcting capacity in terms of minimal block size. Properties and consideration for the Mixed Parity Code were described by the three different criteria. Further, the mechanism of using Mixed Parity Code for constructing efficient block cipher was expressed. The possibility of dual role, namely intentional bit error penetration at encryption and unknown bit error correction at decryption, of Mixed Parity Code was explained to increase the security strength against some of the well-known attacks.

A newly developed Enhanced Hamming Code with its constructions and properties were described for designing of proposed VLSI architectures for secured data distribution. The failure rate of error control of Enhanced Hamming Code can be decreased by doing enhancement in the single bit error correction code for fixed transmission rate of the original message. Multiple and burst error correction capabilities of the Enhanced Hamming Code was described. Selective bits placement of the codewords of Enhanced Hamming Code technique was used for the increase in deliberated error control success rate.