Linear Block Codes

The parity bits of linear block codes are linear combination of the message. Therefore, we can represent the encoder by a linear system described by matrices.

Basic Definitions

- Linearity:

$$\text{If} \quad \mathbf{m}_1 \rightarrow \mathbf{c}_1 \ \text{and} \ \mathbf{m}_2 \rightarrow \mathbf{c}_2$$
$$\text{then} \ \mathbf{m}_1 \oplus \mathbf{m}_2 \rightarrow \mathbf{c}_1 \oplus \mathbf{c}_2$$

where **m** is a *k*-bit information sequence
c is an *n*-bit codeword.
$\oplus$ is a bit-by-bit mod-2 addition without carry

- Linear code: The sum of any two codewords is a codeword.
- Observation: The all-zero sequence is a codeword in every linear block code.

Basic Definitions (cont'd)

- Def: The weight of a codeword $\mathbf{c}_i$, denoted by $w(\mathbf{c}_i)$, is the

  number of of nonzero elements in the codeword.

- Def: The minimum weight of a code, w_{min}, is the smallest

  weight of the nonzero codewords in the code.

- Theorem: In any linear code, $d_{\text{min}} = w_{\text{min}}$

- Systematic codes

n-k check bits	*k* information bits

Any linear block code can be put in systematic form

linear Encoder.

By linear transformation

$\boldsymbol{c} = \boldsymbol{m} \cdot \boldsymbol{G} = m_0 g_0 + m_1 g_0 + \ldots\ldots + m_{k-1} g_{k-1}$

The code C is called a k-dimensional subspace.

G is called a generator matrix of the code.

Here G is a $k \times n$ matrix of rank k of elements from GF(2), g_i is the i-th row vector of G.

The rows of G are linearly independent since G is assumed to have rank k.

Example:

($7, 4$) Hamming code over GF(2)

The encoding equation for this code is given by

$$c_0 = m_0$$
$$c_1 = m_1$$
$$c_2 = m_2$$
$$c_3 = m_3$$
$$c_4 = m_0 + m_1 + m_2$$
$$c_5 = m_1 + m_2 + m_3$$
$$c_6 = m_0 + m_1 + m_3$$

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

Linear Systematic Block Code:

An (*n, k*) *linear systematic code is completely specified by a k* × *n* generator matrix of the following form.

$$G = \begin{bmatrix} \overline{g}_0 \\ \overline{g}_1 \\ \vdots \\ \overline{g}_{k-1} \end{bmatrix} = \left[I_k P \right]$$

where I_k *is the k × k identity matrix.*

Linear Block Codes

- the number of codeworde is 2^k since there are 2^k distinct messages.
- The set of vectors $\{g_i\}$ are linearly independent since we must have a set of unique codewords.
- linearly independent vectors mean that no vector g_i can be expressed as a linear combination of the other vectors.
- These vectors are called baises vectors of the vector space C.
- The dimension of this vector space is the number of the basis vector which are *k*.
- G_i ϵ C→ the rows of G are all legal codewords.

Hamming Weight

the minimum hamming distance of a linear block code is equal to the minimum hamming weight of the nonzero code vectors.

Since each $g_i \in C$,we must have $W_h(g_i) \geq d_{min}$ this a necessary condition but not sufficient.

Therefore, if the hamming weight of one of the rows of G is less than d_{min}, $\rightarrow d_{min}$ is not correct or G not correct.

Generator Matrix

- All 2^k codewords can be generated from a set of k linearly independent codewords.
- The simplest choice of this set is the k codewords corresponding to the information sequences that have a single nonzero element.
- Illustration: The generating set for the (7,4) code:

  1000 ===> 1101000

  0100 ===> 0110100

  0010 ===> 1110010

  0001 ===> 1010001

Generator Matrix (cont'd)

- Every codeword is a linear combination of these 4 codewords.

  That is: **c** = **m G,** where

$$\mathbf{G} = \begin{bmatrix} \underbrace{\begin{matrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{matrix}}_{k\times(n-k)} & \underbrace{\begin{matrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{matrix}}_{k\times k} \end{bmatrix} = \left[\mathbf{P} \mid \mathbf{I}_{\mathbf{k}}\right]$$

- Storage requirement reduced from $2^k(n+k)$ to $k(n-k)$.

Parity-Check Matrix

For $\mathbf{G} = [\ \mathbf{P} \mid \mathbf{I}_k\ ]$, define the matrix $\mathbf{H} = [\mathbf{I}_{n-k} \mid \mathbf{P}^{\mathrm{T}}]$

(The size of **H** is $(n-k)$xn).

It follows that $\mathbf{G}\mathbf{H}^{\mathrm{T}} = \mathbf{0}$.

Since $\mathbf{c} = \mathbf{m}\mathbf{G}$, then $\mathbf{c}\mathbf{H}^{\mathrm{T}} = \mathbf{m}\mathbf{G}\mathbf{H}^{\mathrm{T}} = \mathbf{0}.$

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

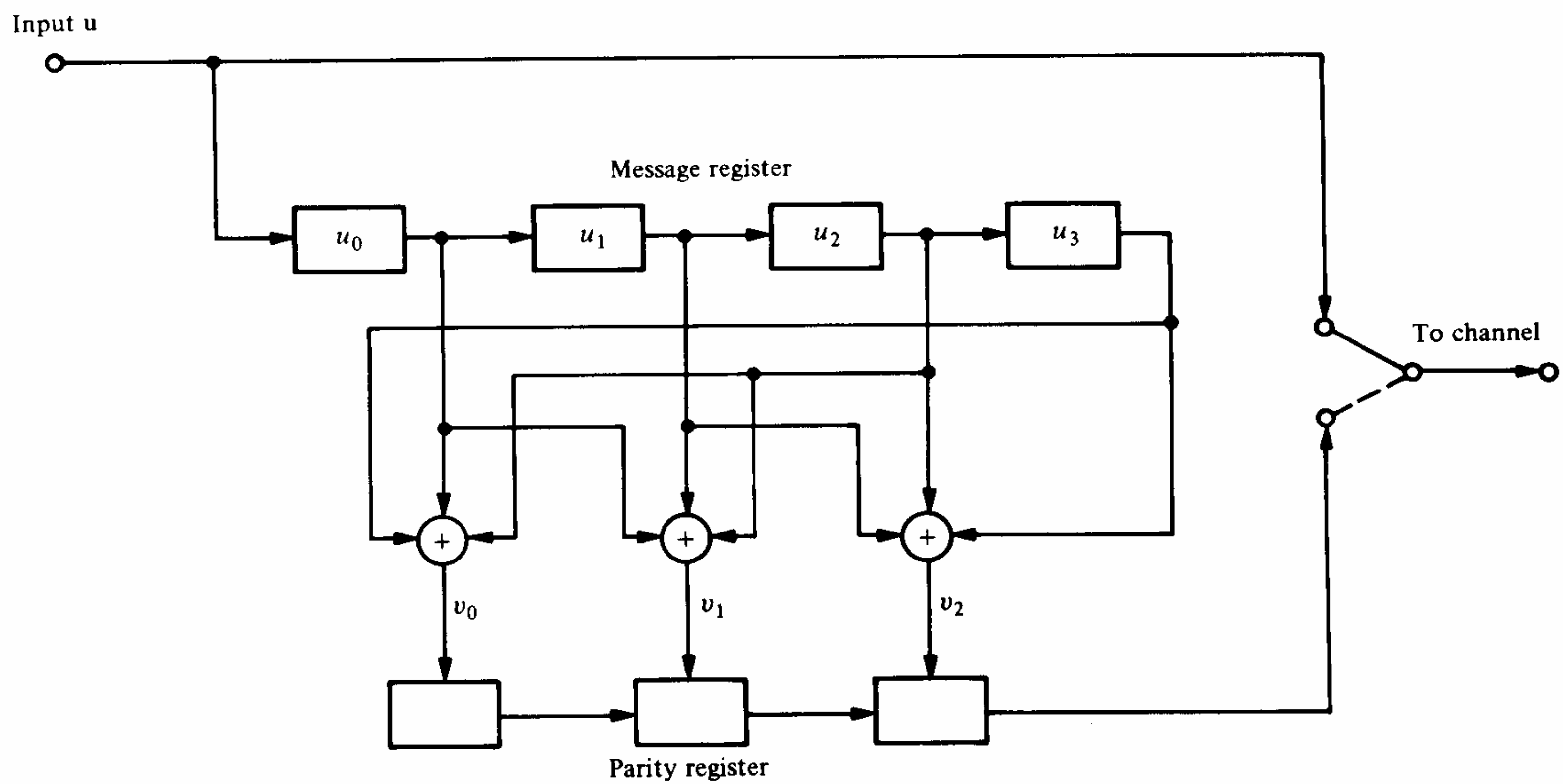
Encoding Using H Matrix

$$\begin{bmatrix} c_1 & c_2 & c_3 & c_4 & c_5 & c_6 & c_7 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \mathbf{0}$$

$$c_1 \quad c_2 \quad c_3 \quad \underbrace{c_4 \quad c_5 \quad c_6 \quad c_7}_{\text{information}}$$

$$\begin{aligned} c_1 + c_4 + c_6 + c_7 &= 0 \\ c_2 + c_4 + c_5 + c_6 &= 0 \\ c_3 + c_5 + c_6 + c_7 &= 0 \end{aligned} \quad \Rightarrow \quad \begin{aligned} c_1 &= c_4 + c_6 + c_7 \\ c_2 &= c_4 + c_5 + c_6 \\ c_3 &= c_5 + c_6 + c_7 \end{aligned}$$

Encoding Circuit



The Encoding Problem (Revisited)

- Linearity makes the encoding problem a lot easier, yet:

  How to construct the G (or H) matrix of a code of minimum distance $d_{\min}$?
- The general answer to this question will be attempted later. For the time being we will state the answer to a class of codes: the Hamming codes.

Hamming Codes

- Hamming codes constitute a class of single-error correcting codes defined as:

  $n = 2^r\text{-}1, k = n\text{-}r, r > 2$
- The minimum distance of the code $d_{\min} = 3$
- Hamming codes are perfect codes.
- <u>Construction rule</u>:

  The H matrix of a Hamming code of order r has as its columns all non-zero r-bit patterns.

  Size of H: r x$(2^r\text{-}1)=(n\text{-}k)xn$