ECE 405/511
Error Control Coding

Simple Encoding and Decoding

Equivalent Binary Linear Codes

Two binary linear codes are called equivalent if one can be obtained from the other by permuting the positions of the code.

Two $k \times n$ binary matrices generate equivalent linear codes if one matrix can be obtained from the other by a sequence of the following operations:

a) permutation of the rows

b) addition of one row to another

c) permutation of the columns

Equivalent Codes

- Distances between codewords are unchanged by the equivalence operations.

- Consequently, equivalent linear codes have the same parameters (n,k,d)
  - correct and detect the same number of errors

- Therefore the form of the code can be chosen to best suit the application.

Example (5,2,3) Codes

$$\mathbf{G} = \begin{bmatrix} 00111 \\ 11100 \end{bmatrix}$$

a) $$\mathbf{G}' = \begin{bmatrix} 11100 \\ 00111 \end{bmatrix}$$

b) $$\mathbf{G}'' = \begin{bmatrix} 00111 \\ 11011 \end{bmatrix}$$

c) $$\mathbf{G}''' = \begin{bmatrix} 10110 \\ 01101 \end{bmatrix}$$

Systematic Codes

Let **G** be a generator matrix of an (n,k) code. Then by the previous operations, **G** can be transformed into the form

$$[\ \mathbf{I}_k \mid \mathbf{P}\ ]$$

where $\mathbf{I}_k$ is a $k \times k$ identity matrix and **P** is a $k \times (n - k)$ matrix.

Example

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

Systematic Codes

Transforming **G** to systematic form

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} \rightarrow$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} \rightarrow ?$$

Systematic Codes

$$\rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} = [\mathbf{I}_4 \;\; \mathbf{P}]$$

Systematic Codes

- For a systematic code the data appears unaltered in the codeword
- If the data is on the left, the generator matrix has the structure

$$\mathbf{G} = \overbrace{\begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}}^{k} \overbrace{\begin{bmatrix} p_{0,0} & p_{0,1} & \cdots & p_{0,n-k-1} \\ p_{1,0} & p_{1,1} & \cdots & p_{1,n-k-1} \\ \vdots & \vdots & & \vdots \\ p_{k-1,0} & p_{k-1,1} & \cdots & p_{k-1,n-k-1} \end{bmatrix}}^{n-k} = \left[\mathbf{I}_k \mid \mathbf{P} \right]$$

- **P** is often referred to as the parity matrix

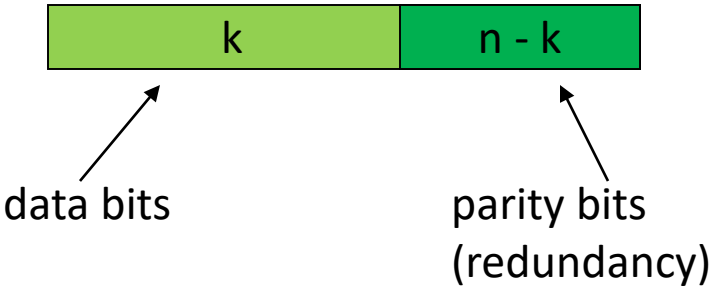
Encoding with a Systematic Code

$$\mathbf{c} = \mathbf{mG} = \mathbf{m}\left[\mathbf{I}_k \mid \mathbf{P}\right] = (\mathbf{m} \mid \mathbf{b})$$

$$\mathbf{G} = \begin{bmatrix} 1000 & 101 \\ 0100 & 111 \\ 0010 & 011 \\ 0001 & 110 \end{bmatrix} = [\mathbf{I}_4 \;\; \mathbf{P}]$$

$$\mathbf{m} = 0111 \;\; \mathbf{c} = \mathbf{mG} = 0111010 = \mathbf{m}010$$

$$\mathbf{m} = 1011 \;\; \mathbf{c} = \mathbf{mG} = 1011101 = \mathbf{m}101$$



Parity Check Matrix

- Define a parity check matrix $\mathbf{H}$ (n-$k \times n$) such that

$$\mathbf{G}\mathbf{H}^T = \mathbf{0}$$

- $\mathbf{H}$ is a basis for the dual space
- If $\mathbf{G}$ is in systematic form

$$\mathbf{G} = [\mathbf{I}_k \ \mathbf{P}]$$

$$\mathbf{H} = [-\mathbf{P}^T \ \mathbf{I}_{n-k}] \ (= [\mathbf{P}^T \ \mathbf{I}_{n-k}] \text{ in binary})$$

$$\mathbf{G}\mathbf{H}^T = -\mathbf{P}+\mathbf{P} = \mathbf{0}$$

- Every codeword $\mathbf{c}$ satisfies the parity check equations given by the rows of $\mathbf{H}$

$$\mathbf{c}\mathbf{H}^T = \mathbf{m}\mathbf{G}\mathbf{H}^T = \mathbf{0}$$

Examples

- (3,1,3) Repetition code

$$\mathbf{G} = \left[\begin{array}{c|cc} 1 & 1 & 1 \end{array}\right] \qquad \mathbf{H} = \left[\begin{array}{c|cc} 1 & 1 & 0 \\ \hline 1 & 0 & 1 \end{array}\right]$$

- (3,2,2) SPC code

$$\mathbf{G} = \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \end{array}\right] \qquad \mathbf{H} = \left[\begin{array}{cc|c} 1 & 1 & 1 \end{array}\right]$$

Examples

- (8,7,2) SPC code

$$\mathbf{G} = \left[\begin{array}{c:c} & 1 \\ & 1 \\ & 1 \\ \mathbf{I}_7 & 1 \\ & 1 \\ & 1 \\ & 1 \end{array}\right] \qquad \mathbf{H} = \left[\begin{array}{c:c} 1111111 & 1 \end{array}\right]$$

Systematic Parity Check Matrices

If $\mathbf{G} = [\mathbf{I}_k \mid \mathbf{P}]$ is the generator matrix of a binary (n,k) code C, then a parity check matrix for C is

$\mathbf{H} = [\mathbf{P}^T \mid \mathbf{I}_{n-k}]$

Example

$$\mathbf{G} = \left[\begin{array}{c|ccc} & 1 & 0 & 1 \\ & 1 & 1 & 1 \\ \mathbf{I}_4 & 0 & 1 & 1 \\ & 1 & 1 & 0 \end{array}\right] \quad \Rightarrow \quad \mathbf{H} = \left[\begin{array}{cccc|c} 1 & 1 & 0 & 1 & \\ 0 & 1 & 1 & 1 & \mathbf{I}_3 \\ 1 & 1 & 1 & 0 & \end{array}\right]$$

Dual Codes

If C is a linear (n,k) code generated by **G**, then the dual code $C^{\perp}$ is a linear $(n,n\text{-}k)$ code generated by **H**.

The dual code is just the dual space.

Example

$$C = \begin{matrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{matrix} \qquad C^{\perp} = \begin{matrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{matrix}$$

(3,1,3) code (3,2,2) code

Dual Codes

For the (n,1) repetition code C with generator matrix

$$\mathbf{G} = [1 \;\; 1 \;\; \ldots \;\; 1]$$

the dual code $C^{\perp}$ is an (n,n-1) SPC code with generator matrix

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

The rows of **G** are orthogonal to the rows of **H**

Duals of (5,2,3) Codes

- Consider two equivalent (5,2,3) codes
  - nonsystematic and systematic

$$\mathbf{G} = \begin{bmatrix} 00111 \\ 11100 \end{bmatrix} \rightarrow \mathbf{G}' = \left[\begin{array}{c|cc} 10 & 110 \\ 01 & 101 \end{array}\right] \quad \text{(permute columns 1 and 5)}$$

$$\mathbf{H} = \begin{bmatrix} 01101 \\ 00011 \\ 11000 \end{bmatrix} \qquad \mathbf{H}' = \left[\begin{array}{c|cc} 11 & 100 \\ 10 & 010 \\ 01 & 001 \end{array}\right]$$

- **H** and **H'** both generate (5,3,2) codes

Decoding Linear Codes

- After transmission through a noisy channel the received word is

  $$\mathbf{r} = \mathbf{c} + \mathbf{e}$$

  where

  c is the transmitted codeword

  e is the error pattern
- For error detection, the goal is to determine if **e** = 0
- For error correction, the goal is to find the codeword **c**' that is closest to **r**
  - there are 2^n possible received words
  - there are 2^k codewords

Decoding Linear Codes

- One possibility is a lookup table
- The received word **r** is used as an address

- For error detection, just store a binary value
  - 0: valid codeword (**e** = 0)
  - 1: not a codeword (**e** ≠ 0)
  - Example: (8,7,2) single parity check code

Address (**r**)	Entry
00000000	0
00000001	1
00000010	1
00000011	0
⋮	⋮

If the entry is 0, the data is the first $k = 7$ bits (assuming a systematic code, **G** = [**I P**])

Decoding Linear Codes

- For error correction, the table must provide the codeword (**c**), message (**m**), or error pattern (**e**)
- If the probability of error is small, the most likely error patterns are the ones with small Hamming weight
- How to determine the most likely received words associated with these error patterns?
- The easiest way is to make a table of possible combinations

$$\mathbf{r} = \mathbf{c} + \mathbf{e}$$

Standard Array

- The number of correctable error patterns $\mathbf{e}_i$ is 2^{n-k}
- The standard array is formed by choosing $\mathbf{e}_i$ to be the
  - all-zero pattern
  - 1 bit error patterns
  - 2 bit error patterns
    ⋮
- Ensure that each new error pattern is not already in the array

Standard Array

$\mathbf{c}_0$ (all zero)	$\mathbf{c}_1$	.....	$\mathbf{c}_{M-2}$	$\mathbf{c}_{M-1}$
$\mathbf{e}_1$	$\mathbf{c}_1+\mathbf{e}_1$	.....	$\mathbf{c}_{M-2}+\mathbf{e}_1$	$\mathbf{c}_{M-1}+\mathbf{e}_1$
$\mathbf{e}_2$	$\mathbf{c}_1+\mathbf{e}_2$	.....	$\mathbf{c}_{M-2}+\mathbf{e}_2$	$\mathbf{c}_{M-1}+\mathbf{e}_2$
$\mathbf{e}_3$	$\mathbf{c}_1+\mathbf{e}_3$	.....	$\mathbf{c}_{M-2}+\mathbf{e}_3$	$\mathbf{c}_{M-1}+\mathbf{e}_3$
⋮	⋮	.....	⋮	⋮
$\mathbf{e}_{2^{n-k}-1}$	$\mathbf{c}_1+\mathbf{e}_{2^{n-k}-1}$	.....	$\mathbf{c}_{M-2}+\mathbf{e}_{2^{n-k}-1}$	$\mathbf{c}_{M-1}+\mathbf{e}_{2^{n-k}-1}$

The array has

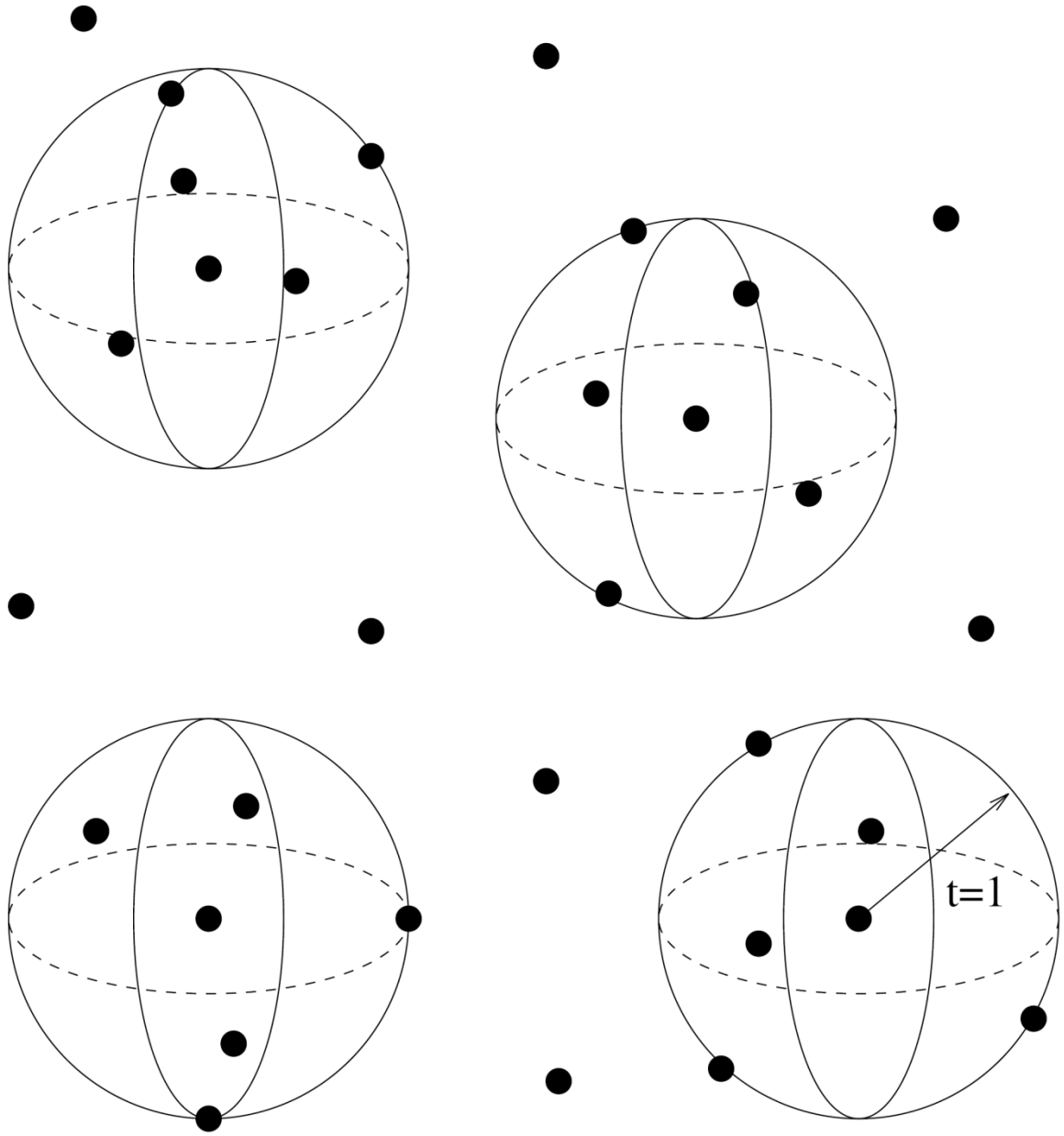
- 2^k columns (the number of codewords M)
- 2^{n-k} rows (the number of correctable error patterns)

Standard Array Decoding

- Assume that the received word is $\mathbf{r} = \mathbf{c}_1 + \mathbf{e}_3$ (shown in red in the standard array)
- The most likely codeword is the one at the top of the column containing **r**
- The corresponding error pattern is the coset leader at the start of the row containing **r**
- Can be implemented using a lookup table which maps all words in the array to the
  - error pattern **e** on the left of the row containing the received word, or the
  - message **m** corresponding to the codeword **c** at the top of the column containing the received word

Standard Array for the (5,2,3) Code

00000	10110	01101	11011
10000	00110	11101	01011
01000	11110	00101	10011
00100	10010	01001	11111
00010	10100	01111	11001
00001	10111	01100	11010
00011	10101	01110	11000
10001	00111	11100	01010


t=1

Standard Array

- The rows of the standard array are called cosets
- The first (left) column elements are called the coset leaders
  - the coset leaders are the correctable error patterns

Syndromes

- For a received word **r**, there is a simpler method to determine the closest codeword (or lowest weight error pattern)
- To do this we use the syndrome **s** of the received word **r**

$$\mathbf{s} = \mathbf{r}\mathbf{H}^T$$

- If **c** was corrupted by the error vector **e** then

$$\mathbf{r} = \mathbf{c} + \mathbf{e}$$

and

$$\mathbf{s} = \mathbf{r}\mathbf{H}^T = (\mathbf{c} + \mathbf{e})\mathbf{H}^T = \mathbf{c}\mathbf{H}^T + \mathbf{e}\mathbf{H}^T$$

$$\mathbf{s} = \mathbf{0} + \mathbf{e}\mathbf{H}^T$$

- The syndrome depends only on the error vector **e**

Syndromes

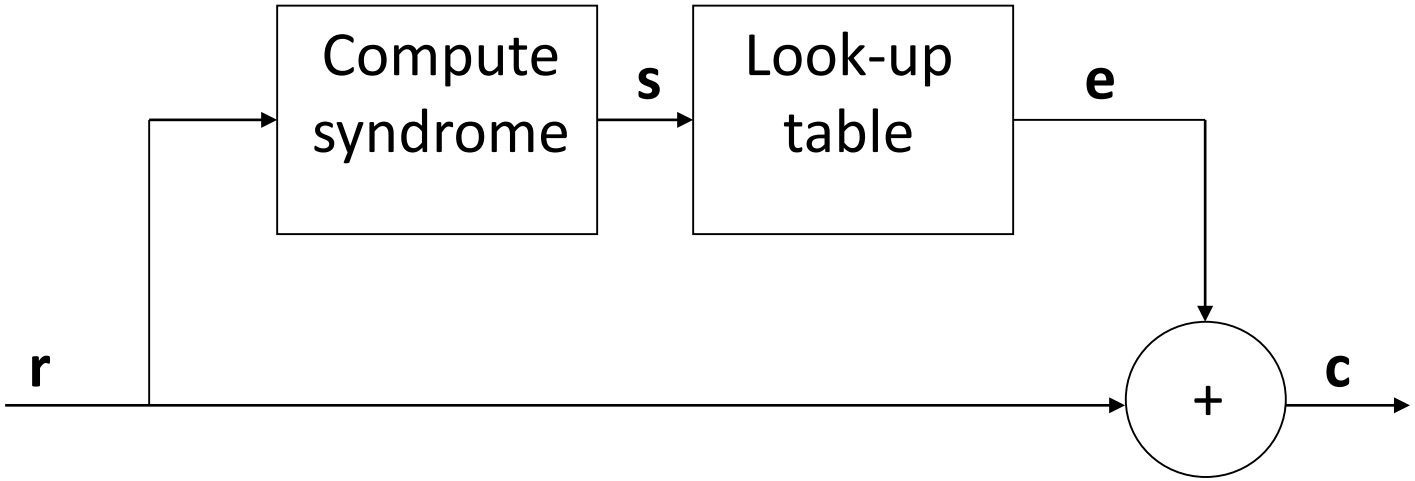
- The same error pattern added to different codewords gives the same syndrome.

- There are 2^n possible received words but only
- $2^{(n-k)}$ syndromes
  - (5,2,3) code has 32 received words and 8 syndromes
  - (7,4,3) code has 128 received words and 8 syndromes
  - (255,247,3) code has 2^{255} received words and $2^8 = 256$ syndromes

- Only need to determine which error pattern corresponds to the syndrome.

Syndrome Decoding

Block diagram of the syndrome decoder



Syndrome Decoding - (5,2,3) Code

$$\mathbf{G} = \left[\begin{array}{cc|ccc} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{array}\right]$$

$$\mathbf{H} = \left[\begin{array}{cc|ccc} 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{array}\right]$$

$\mathbf{s}_0 = (00000)\mathbf{H}^T = 000$
$\mathbf{s}_1 = (10000)\mathbf{H}^T = 110$
$\mathbf{s}_2 = (01000)\mathbf{H}^T = 101$
$\mathbf{s}_3 = (00100)\mathbf{H}^T = 100$
$\mathbf{s}_4 = (00010)\mathbf{H}^T = 010$
$\mathbf{s}_5 = (00001)\mathbf{H}^T = 001$
$\mathbf{s}_6 = (00011)\mathbf{H}^T = 011$
$\mathbf{s}_7 = (10001)\mathbf{H}^T = 111$

r = (10010) $\mathbf{rH}^T = 100$ **e** = (00100) **c** = **r**+**e** = (10110)
r = (11000) $\mathbf{rH}^T = 011$ **e** = (00011) **c** = **r**+**e** = (11011)

Syndrome Decoding – (7,4,3) Code

$$\mathbf{G} = \left[\mathbf{I}|\mathbf{P}\right] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$\mathbf{H} = \left[\mathbf{P}^{\mathrm{T}}|\mathbf{I}\right] = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Syndromes for the (7,4,3) Code

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

$\mathbf{s}_0 = (0000000)\mathbf{H}^T = 000$
$\mathbf{s}_1 = (1000000)\mathbf{H}^T = 011$
$\mathbf{s}_2 = (0100000)\mathbf{H}^T = 101$
$\mathbf{s}_3 = (0010000)\mathbf{H}^T = 110$
$\mathbf{s}_4 = (0001000)\mathbf{H}^T = 111$
$\mathbf{s}_5 = (0000100)\mathbf{H}^T = 100$
$\mathbf{s}_6 = (0000010)\mathbf{H}^T = 010$
$\mathbf{s}_7 = (0000001)\mathbf{H}^T = 001$

Syndrome Decoding

- Received word **r** = (1101001)

$$\mathbf{s} = \mathbf{r}\mathbf{H}^{T} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = 0 \quad 0 \quad 0$$

- **e** = (0000000) and **r** is a valid codeword

Syndrome Decoding

- Received word **r** = (1101000)

$$\mathbf{s} = \mathbf{r}\mathbf{H}^{\mathrm{T}} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{matrix} 0 & 0 & 1 \end{matrix}$$

- Syndrome 001 indicates an error in bit 7 of the received word so **c** = (1101001)