2020 Short Course on Error-Correcting Codes

Introduction to Linear Block Codes

February 10, 2020

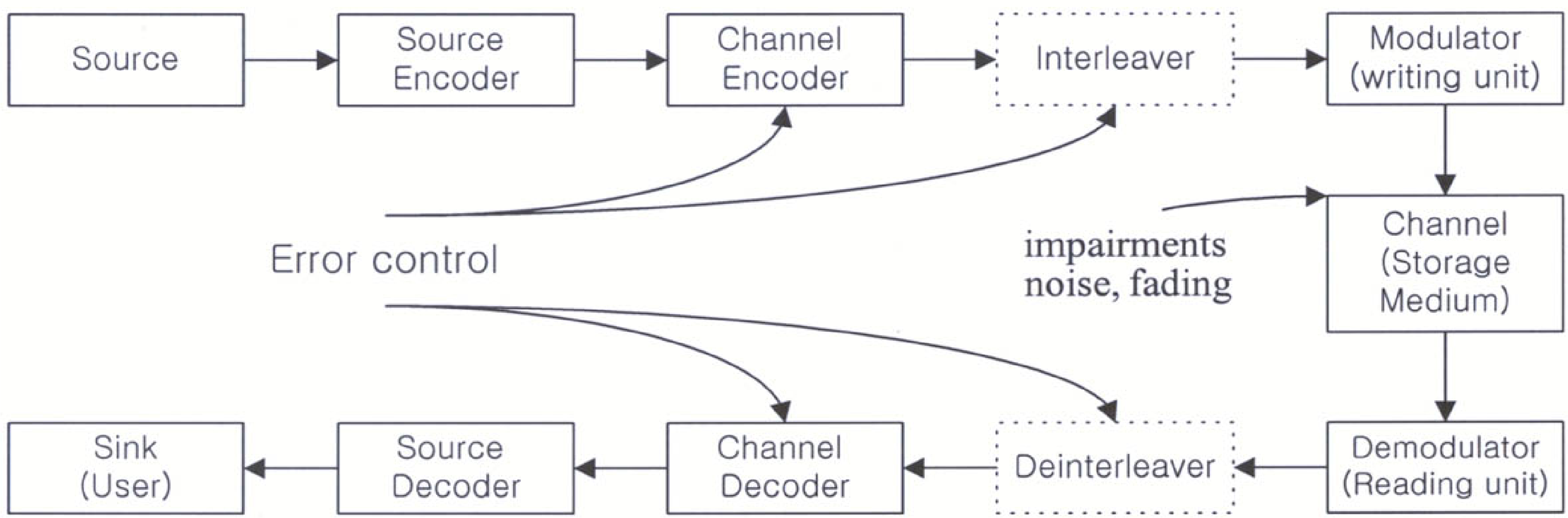
Kyeongcheol Yang

Department of Electrical Engineering

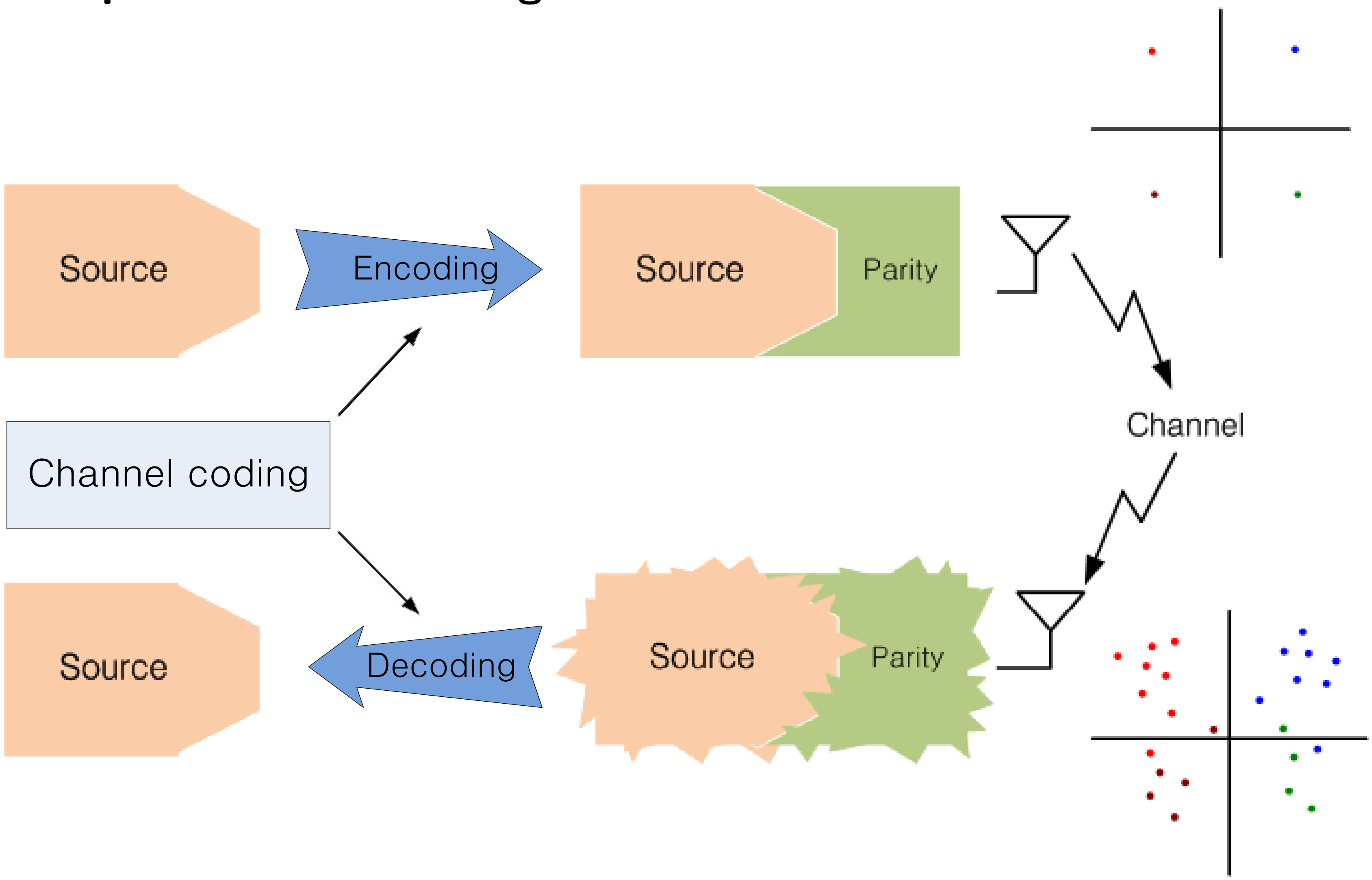
Pohang University of Science and Technology (POSTECH)

Introduction to Error-Correcting Codes

□ Digital Communication Systems



□ Concept of Channel Coding



□ Limitations in Communication Systems

- Bandwidth limitations
- Power limitations
- Channel impairments (attenuation, distortion, interference, noise and fading)

$\Rightarrow$ *Error control techniques* are employed in digital communication systems for reliable transmission under these limitations

□ Physical Channels

- Communication channels: *here to there*
- Storage channels: *now to then*

□ Advantages of Error Control Coding

- In principle:

  Every channel has a capacity C. If you transmit information at a rate $R < C$, then *error-free transmission* is possible.

- In practice:

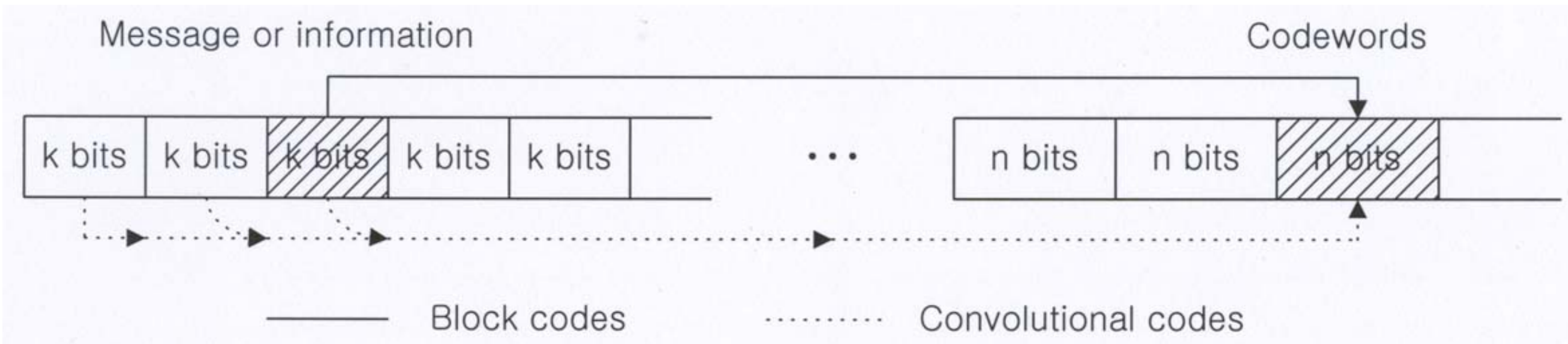
  – Increase the operational range of a communication system

  – Reduce the error rates

  – Reduce the transmitted power requirements

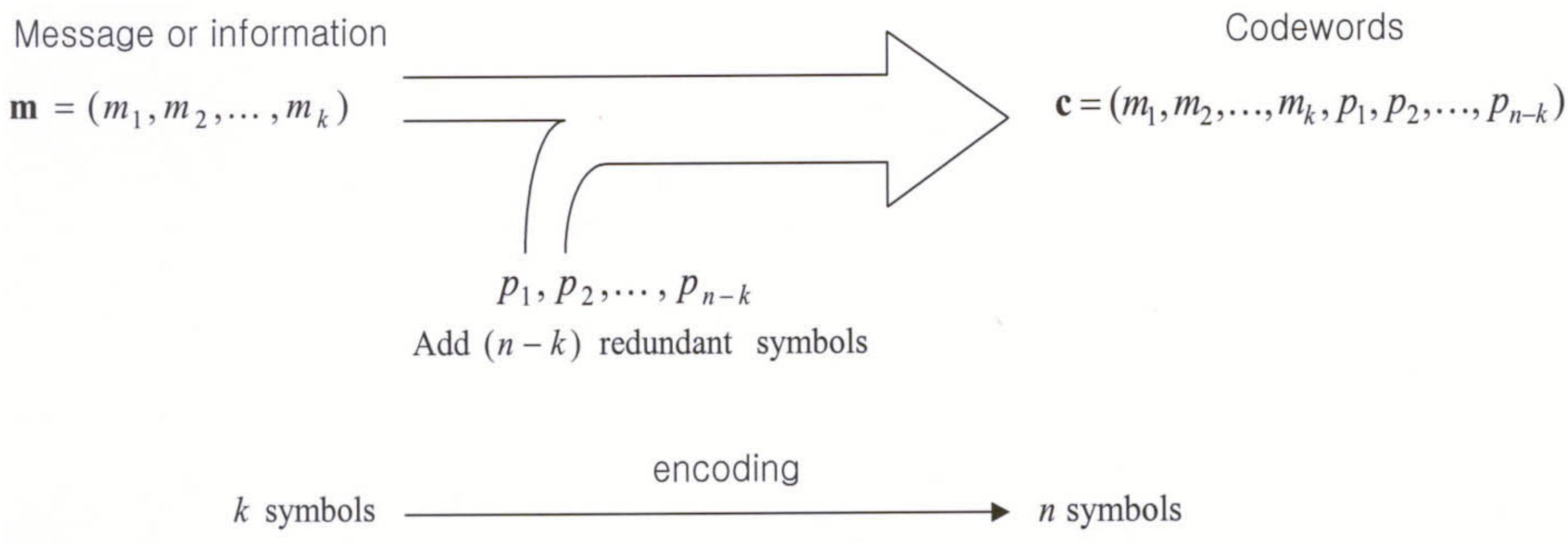
□ Error Control Techniques

- *Forward Error Correction (FEC)*



- *Error Detection*
  - Cyclic Redundancy Check (CRC)
  - Syndrome checking
  - Applications: Automatic Repeat reQuest (ARQ)

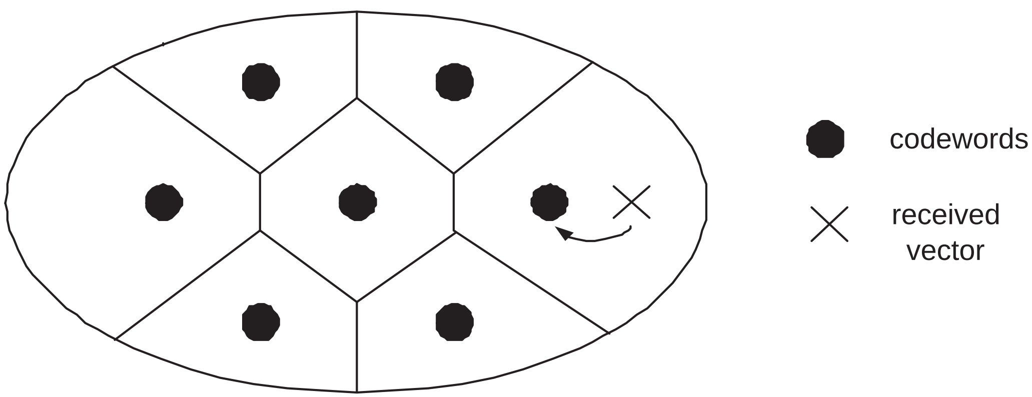
□ Encoding of an $[n, k]$ Block Code

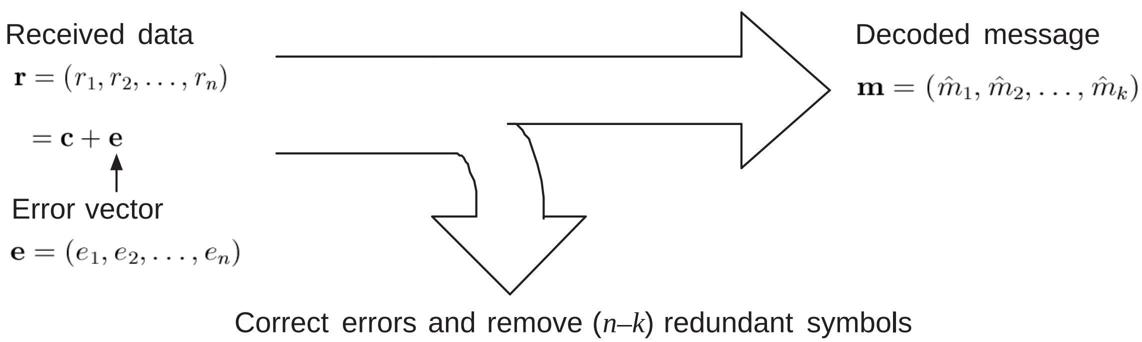


- *Redundancy* $\quad r = n - k$
- *Code rate* $\quad R = k/n$

□ Decoding of an $[n, k]$ Block Code

- Decide what the transmitted information was
- Optimum decoding rule: Minimum distance decoding in a memoryless channel





□ Example of Decoding: $[6, 3]$ code

message	codeword	distance
000	000000	4
100	100101	1
010	010011	5
110	110110	4
001	001111	2
101	101010	3
011	011100	3
111	111001	2

comparing with 101101

transmit the information: 100

$\Downarrow$

choose the codeword: 100101

$\Downarrow$

101101 is received

$\Downarrow$

the closest codeword: 100101

$\Downarrow$

extract the information: 100

□ Measure of Distance

- *Hamming distance* = the number of positions at which symbols are different in the two vectors

  Example: $\mathbf{u} = (101000), \ \ \mathbf{v} = (111010) \quad \Rightarrow \quad d(\mathbf{u}\ , \mathbf{v}) = 2$

- *Hamming weight* = the number of nonzero elements in a vector

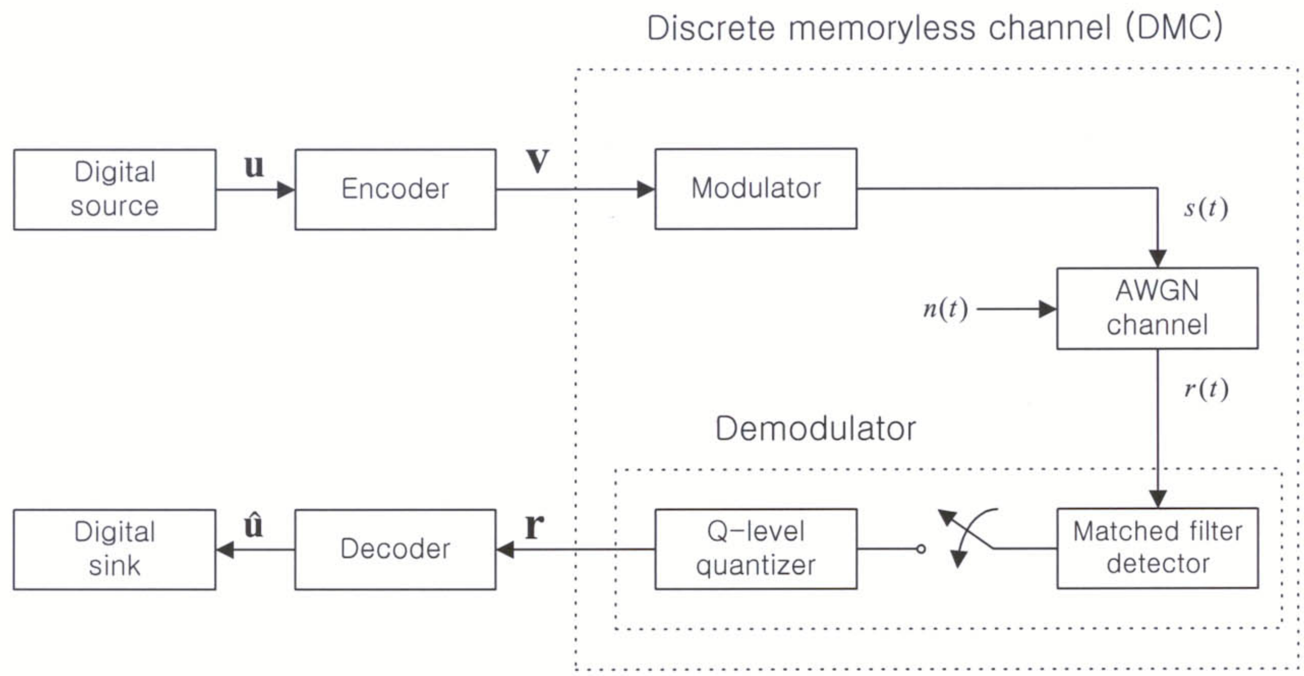
  Example: $w(\mathbf{u}) = 2, \ \ w(\mathbf{v}) = 4$

- *Binary case* : $d(\mathbf{u}\ , \mathbf{v}) = w(\mathbf{u} \oplus \mathbf{v})$ ($\oplus$ means (bitwise) exclusive OR)

  Example: $\mathbf{u} \oplus \mathbf{v} = (010010)$

  $$d(\mathbf{u}\ , \mathbf{v}) = w((010010)) = 2$$

☐ Maximum-Likelihood Decoding (MLD)



- $\hat{\mathbf{u}} = \mathbf{u} \;\Leftrightarrow\; \hat{\mathbf{v}} = \mathbf{v}$

  $\hat{\mathbf{v}}$ = an estimate of the codeword $\mathbf{v}$, given $\mathbf{r}$

- Assume the codeword $\mathbf{v}$ was transmitted.

  A decoding error occurs $\;\Leftrightarrow\; \hat{\mathbf{v}} \neq \mathbf{v}$.

□ Optimum Receiver: MAP decoder

- The conditional error prob. of the decoder given $\mathbf{r}$: $P(E|\mathbf{r}) = P(\hat{\mathbf{v}} \neq \mathbf{v}|\mathbf{r})$
- The error probability of the decoder

$$P(E) = \sum_{\mathbf{r}} P(E|\mathbf{r})\, P(\mathbf{r})$$

Note that $P(\mathbf{r})$ is independent of decoding rule.

- Criterion: minimize $P(E)$

$\Leftrightarrow$ minimize $P(E|\mathbf{r}) = P(\hat{\mathbf{v}} \neq \mathbf{v}|\mathbf{r})$ for each $\mathbf{r}$

$\Leftrightarrow$ maximize $P(\hat{\mathbf{v}} = \mathbf{v}|\mathbf{r})$ for each $\mathbf{r}$

- Optimum decoding rule:

$$\hat{\mathbf{v}} = \mathbf{v} \;\Leftrightarrow\; P(\mathbf{v}|\mathbf{r}) = \max_{\mathbf{s}} P(\mathbf{s}|\mathbf{r})$$

$\Rightarrow$ *MAP (maximum a posteriori probability) decoder*

□ Maximum-Likelihood Decoding (MLD)

- Assume that $P(\mathbf{v})$ is constant, i.e., $\mathbf{v}$ is equally likely.

- Bayes' Rule:

$$P(\mathbf{v}|\mathbf{r}) = \frac{P(\mathbf{r}|\mathbf{v})P(\mathbf{v})}{P(\mathbf{r})}$$

- The MAP decoder is equivalent to the following rule:

$$\hat{\mathbf{v}} = \mathbf{v} \ \Leftrightarrow P(\mathbf{r}|\mathbf{v}) = \max_{\mathbf{s}} P(\mathbf{r}|\mathbf{s})$$

$\Rightarrow$ *ML (maximum-likelihood) decoder*

□ DMC (Discrete Memoryless Channel)

- Given that $\mathbf{v}$ was transmitted, the conditional probability of $\mathbf{r}$ is

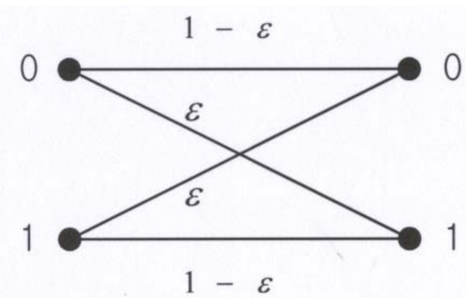
$$P(\mathbf{r}|\mathbf{v}) = \prod_i P(r_i|v_i)$$

$$\log P(\mathbf{r}|\mathbf{v}) = \sum_i \log P(r_i|v_i)$$

- *ML decoder for a DMC:*

$$\hat{\mathbf{v}} = \mathbf{v} \;\Leftrightarrow\; \log P(\mathbf{r}|\mathbf{v}) = \max_{\mathbf{s}} \; \log P(\mathbf{r}|\mathbf{s})$$

- Example: BSC (binary symmetric channel)



$r_i \neq v_i : P(r_i \mid v_i) = \varepsilon$

$r_i = v_i : P(r_i \mid v_i) = 1 - \varepsilon$

☐ ML Decoding for BSC

- The conditional probability is

$$\begin{aligned}\log P(\mathbf{r}|\mathbf{v}) &= d(\mathbf{r},\mathbf{v})\log\epsilon + (n - d(\mathbf{r},\mathbf{v}))\log(1-\epsilon) \\ &= d(\mathbf{r},\mathbf{v})\log\frac{\epsilon}{1-\epsilon} + n\log(1-\epsilon)\end{aligned}$$

where $\log\frac{\epsilon}{1-\epsilon} < 0$ for $\epsilon < \frac{1}{2}$ and $n\log(1-\epsilon)$ is constant for all $\mathbf{v}$.

- ML decoding: $\text{maximize } P(\mathbf{r}|\mathbf{v}) \Leftrightarrow \text{minimize } d(\mathbf{r},\mathbf{v})$

$$\hat{\mathbf{v}} = \mathbf{v} \quad \Leftrightarrow \quad d(\mathbf{r},\mathbf{v}) = \min_{\mathbf{s}} d(\mathbf{r},\mathbf{s})$$

- The optimum decoding rule over the BSC is the *Minimum Distance Decoding*.

□ Communication Channels

- *Physical Channels*: Memoryless channel, Symmetric channel, Additive white Gaussian noise (AWGN) channel, Bursty channel, Compound (or diffuse) channel

- *Random error channels*: Memoryless channels such as deep-space channels, satellite channels

  ⇒ Use random-error-correcting codes

- *Burst error channels*: Channels with Memory

  – Radio channels: signal fading due to multipath transmission

  – Wire and cable transmission: impulse switching noise, crosstalk

  – Magnetic recording: tape dropouts due to surface defects and dust particles

  ⇒ Use burst-error-correcting codes

□ Code Performance and Coding Gain

- Performance measure
  - *Bit error rate* (BER) in the information after decoding
  - *Signal-to-noise power ratio* (SNR): E_b/N_0 [dB]

    $E_b =$ signal energy per bit

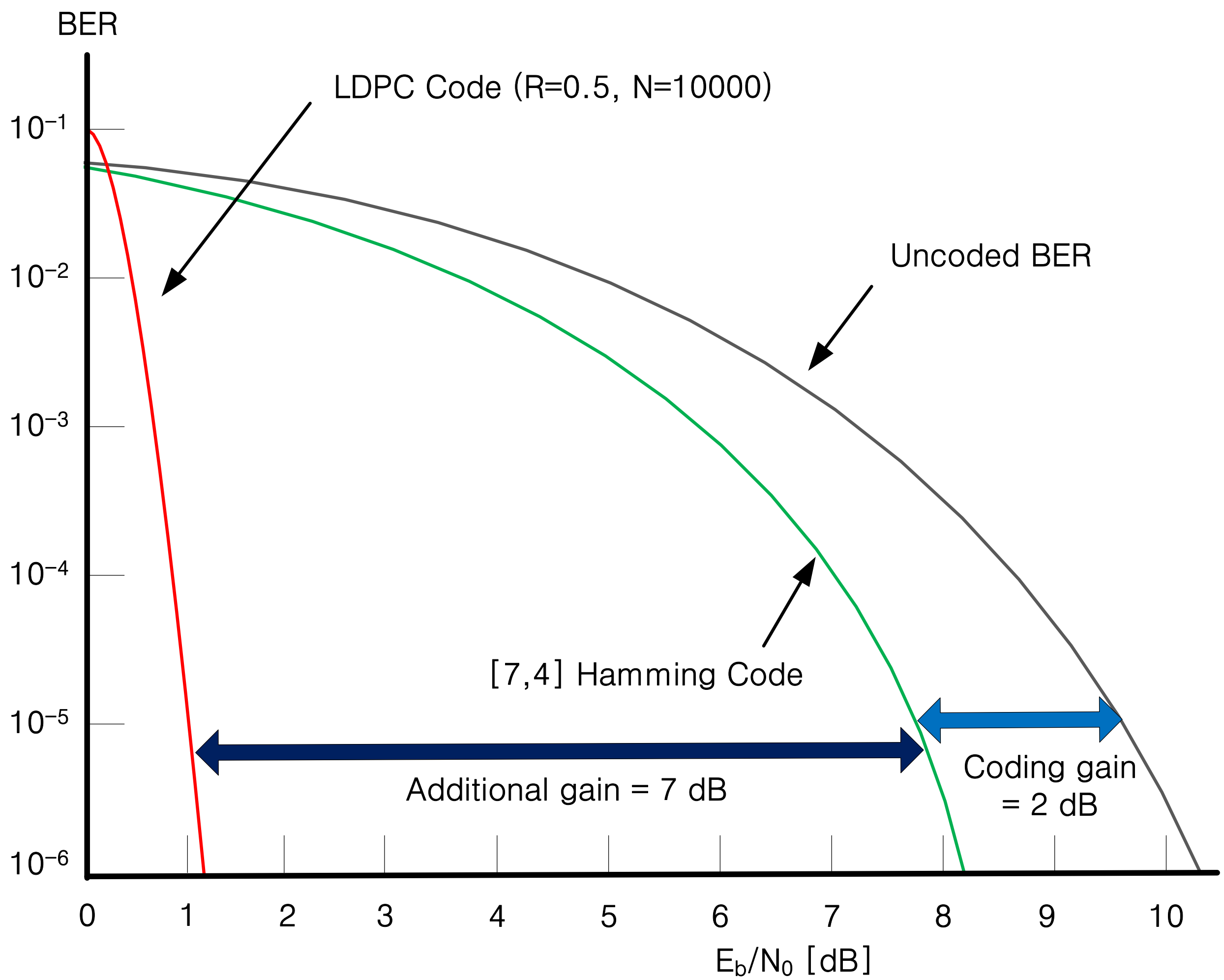
    $N_0 =$ one sided noise power spectral density in the channel

- For a given BER in the communication system, the *coding gain* G is defined by

$$G = \left.\frac{E_b}{N_0}\right|_{\text{w/o FEC}} - \left.\frac{E_b}{N_0}\right|_{\text{with FEC}} \text{ [dB].}$$

At a given BER, we can save the transmission power by G over the uncoded system.

□ BER Performance Curve



□ Basic Problems in Coding Theory

- *To find a good code* (e.g., capacity-achieving or capacity-approaching)
- *To find its decoding algorithm* with low complexity
- *To find a way of implementing the decoding algorithm*

Note:

If we use an $[n, k]$ code, the transmission rate increases by n/k.

$\Rightarrow$ The required channel bandwidth increases by n/k or
the message transmission rate decreases by k/n.

$\Rightarrow$ Cost for FEC

□ Classification of FEC

- Block codes: Hamming, BCH, RS, Golay, Algebraic geometric codes
  Low-density parity-check (LDPC) codes

  Tree codes: Convolutional codes, turbo codes, repeat-accumulate (RA) codes

- Linear codes

  Nonlinear codes: Nordstrom-Robinson code (1967), Preparata codes (1968),
  Kerdock(1972), etc.

- Systematic codes

  Nonsystematic codes

□ History of Coding Theory

- Shannon (1948) proved by the random coding arguments:

  If $R < C$, it is possible to transfer information at error rates that can be reduced to any desired level.

  Here, R is the transmission rate of data and C is the channel capacity.

- The *channel capacity* C of the AWGN channel is given by

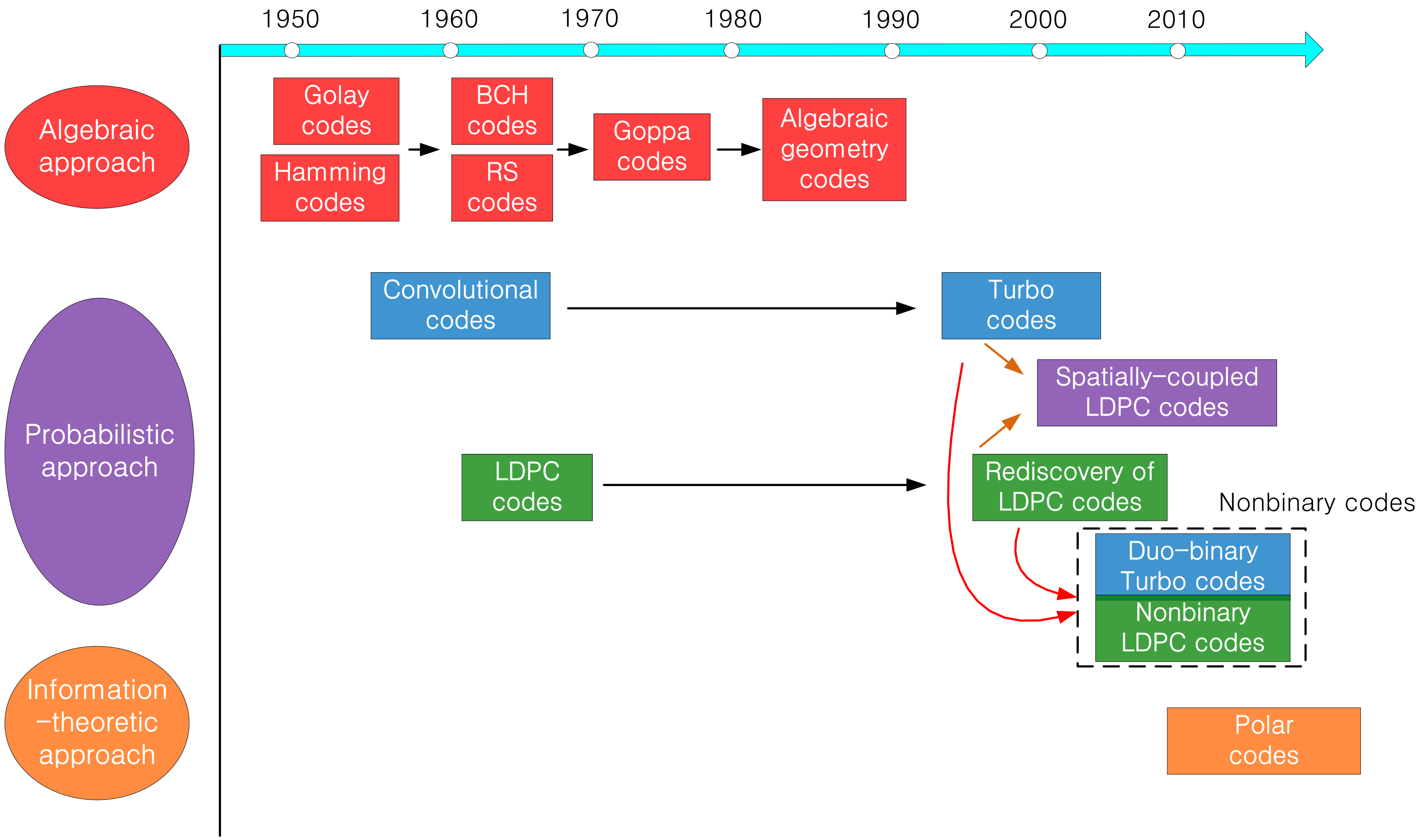
$$C = B \log_2(1 + S/N)$$

  where B is the bandwidth, S is the signal power, and N is the noise power.

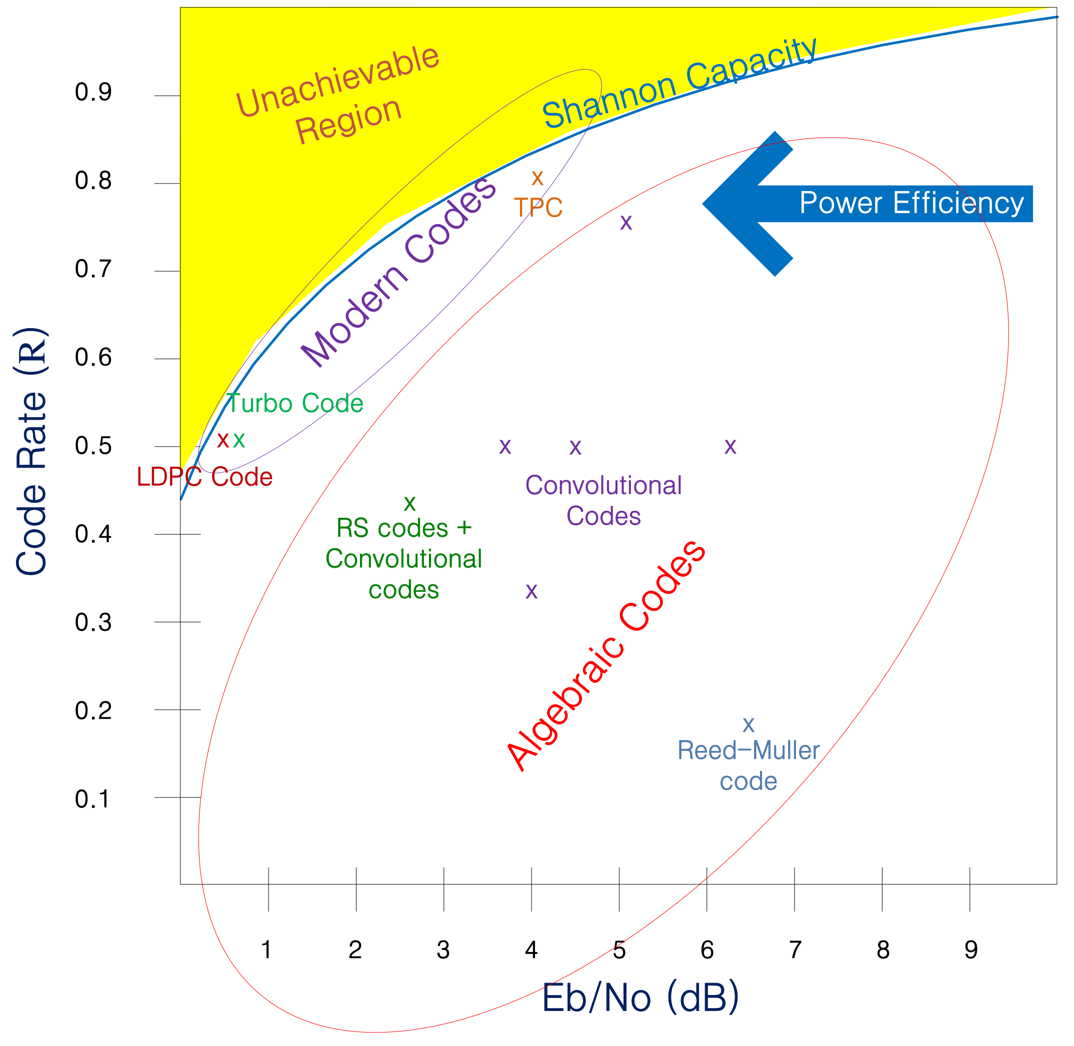
  It is required that $S/N = E_b/N_0 \geq -1.6$ dB.

- **Major Developments of Codes**
  - Hamming codes (1950)
  - Convolutional codes (Elias, 1955)
  - BCH (1960), RS codes (1960)
  - Low-density parity-check (LDPC) codes (Gallager, 1962)
  - Goppa codes (1970)
  - Algebraic-geometric codes (early 1980's)
  - Turbo codes (1993)
  - Turbo-like codes: LDPC codes (rediscovered in 1995),
    RA (repeat-accumulate) codes (1998)

□ Major Approaches to Coding Theory



☐ How Close to the Channel Capacity? (AWGN, BPSK)



Linear Block Codes

□ Main Topics on Linear Block Codes

- Linear codes and vector spaces
- Description of linear codes: generator and parity-check matrices
- Standard array and decoding
- Bounds on the parameters of codes

□ Block Codes

- An (n, M) *block code* $\mathcal{C}$ of size M and length n over the alphabet $\mathcal{A}$ is a set of M vectors of length n with components in $\mathcal{A}$.
- A vector in the code is called a *code vector* or a *codeword*.

- The *rate* (or *code rate*) of an (n, M) block code is defined by $\log_q M/n$, where q is the size of $\mathcal{A}$. It is the number of information symbols per channel symbol.

Example: $\mathcal{A} = \{0, 1\}$, $n = 3$, $M = 4$.

$$\mathcal{C} = \{(000), (011), (101), (110)\}.$$

$$\text{rate} = (\log_2 4)/3 = \tfrac{2}{3}.$$

Example: $\mathcal{A} = \{0, 1, 2\}$, $n = 4$, $M = 3$.

$$\mathcal{C}_1 = \{(0000), (1111), (2222)\}, \quad \mathcal{C}_2 = \{(0112), (2011), (0221)\}.$$

$$\text{rate} = (\log_3 3)/4 = \tfrac{1}{4}.$$

□ Linear Block Codes

- An $[n, k]$ *linear block code* over $\mathbf{F}_q = \mathrm{GF}(q)$ is a k-dimensional subspace of the n-dimensional vector space

$$V_n(\mathbf{F}_q) = \mathbf{F}_q^n \triangleq \{(x_1, x_2, \ldots, x_n) \mid x_i \in \mathbf{F}_q\}.$$

  – n is called the *length* (or *code length*) of the code;

  – k is called the *dimension* of the code;

  – The *rate* (or *code rate*) of an $[n, k]$ linear code is given by

$$\frac{\log_q q^k}{n} = \frac{k}{n}.$$

- **Necessary and sufficient conditions for a code $\mathcal{C}$ to be linear**
  - If $\mathbf{u} \in \mathcal{C}$ and $\mathbf{v} \in \mathcal{C}$, then $\mathbf{u} + \mathbf{v} \in \mathcal{C}$. (vector addition)
  - If $\mathbf{u} \in \mathcal{C}$ and $a \in \mathbf{F}_q$, then $a\mathbf{u} \in \mathcal{C}$. (scalar multiplication)

- **Major concepts for a vector space**
  - Linear Independence
  - Span
  - Basis
  - Dimension
  - Subspace
  - Linear transformation, etc.

□ Generator Matrix G for an $[n, k]$ Code $\mathcal{C}$

- Let $\{\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_k\}$ be a basis for $\mathcal{C}$, where

$$\mathbf{g}_i = (g_{i1}, g_{i2}, \ldots, g_{in}) \quad i = 1, 2, \ldots, k,$$

and $g_{ij} \in \mathbf{F}_q$ for all $i,\ j$.

- Then any codeword $\mathbf{c} = (c_1, c_2, \ldots, c_n)$ can be expressed as a linear combination of $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_k$, i.e.,

$$\mathbf{c} = m_1\mathbf{g}_1 + m_2\mathbf{g}_2 + \ldots + m_k\mathbf{g}_k.$$

- In matrix notation,

$$\mathbf{c} = \begin{bmatrix} m_1 & m_2 & \cdots & m_k \end{bmatrix} \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \\ \vdots \\ \mathbf{g}_k \end{bmatrix} = \begin{bmatrix} m_1 & m_2 & \cdots & m_k \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ g_{k1} & g_{k2} & \cdots & g_{kn} \end{bmatrix}$$

That is,

$$\mathbf{c} = \mathbf{m}\,\mathbf{G}$$

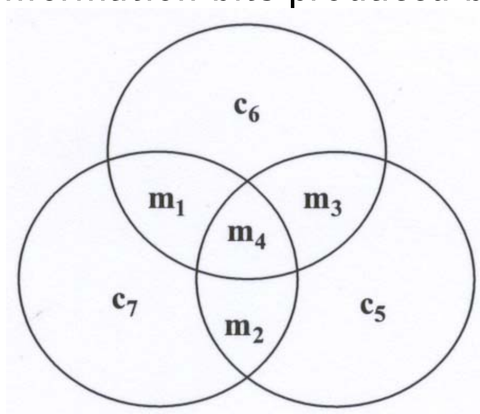
where $\mathbf{c}$: $1 \times n$, $\mathbf{m}$: $1 \times k$, and $\mathbf{G}$: $k \times n$.

- The $k \times n$ matrix $\mathbf{G}$ has k basis vectors for $\mathcal{C}$ as its rows and is called a *generator matrix* of the code.

- $\mathbf{m} \in \mathbf{F}_q^k$ is called a message to be encoded.

☐ Idea of a [7,4] Hamming code (1950)

- Let m_1, m_2, m_3, m_4 be the 4 information bits produced by source and parity bits be constructed as follows:



- The 7 bits to be transmitted are:

$$\underbrace{\begin{aligned} c_1 &= m_1 \\ c_2 &= m_2 \\ c_3 &= m_3 \\ c_4 &= m_4 \end{aligned}}_{\text{information symbols}} \qquad \underbrace{\begin{aligned} c_5 &= m_2 + m_3 + m_4 \pmod 2 \\ c_6 &= m_1 + m_3 + m_4 \pmod 2 \\ c_7 &= m_1 + m_2 + m_4 \pmod 2 \end{aligned}}_{\text{redundant symbols}}$$

“parity-check” symbols)

- The codeword to be transmitted is

$$\mathbf{c} = (c_1, c_2, c_3, \ldots, c_7) \;=\; \underbrace{\begin{bmatrix} m_1 & m_2 & m_3 & m_4 \end{bmatrix}}_{\mathbf{m}} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & \vdots & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & \vdots & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & \vdots & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & \vdots & 1 & 1 & 1 \end{bmatrix}}_{\mathbf{G}}$$

- **Systematic code**: $G = [I_k \ P]$
  - Elementary row operations on $\mathbf{G}$ do not change a code $\mathcal{C}$ because $\mathcal{C}$ is the row space of G.
  - Any generator matrix can be reduced to a "row-reduced echelon form",
  - *Every linear block code can always be considered to be equivalent to a systematic code by applying elementary column operations, if necessary.*

□ Parity-Check Matrix for a $[7,4]$ Hamming code (continued)

- Conditions for $\mathbf{c} = (c_1, c_2, \ldots, c_7)$ to be a codeword:

$$\begin{aligned} c_i &= m_i\,, & i = 1,2,3,4 \\ c_5 &= m_2 + m_3 + m_4 & \pmod 2 \\ c_6 &= m_1 + m_3 + m_4 & \pmod 2 \\ c_7 &= m_1 + m_2 + m_4 & \pmod 2 \end{aligned} \quad \Rightarrow \quad \begin{aligned} c_2 + c_3 + c_4 + c_5 &= 0 \\ c_1 + c_3 + c_4 + c_6 &= 0 \\ c_1 + c_2 + c_4 + c_7 &= 0 \end{aligned}$$

$$\Rightarrow \quad \underbrace{\begin{bmatrix} 0 & 1 & 1 & 1 & \vdots & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & \vdots & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & \vdots & 0 & 0 & 1 \end{bmatrix}}_{=\,\mathbf{H}} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_7 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

- The parity-check matrix $\mathbf{H}$ has size $(n-k) \times n$ and

$$\mathbf{H}\,\mathbf{c}^t \;=\; \mathbf{0}.$$

□ Parity-Check Matrix of an $[n, k]$ Linear Block code

- Let $\mathcal{C}$ be an $[n, k]$ linear code over $\mathbf{F}_q$. A matrix $\mathbf{H}$ with the property that $\mathbf{H}\mathbf{x}^{\mathrm{t}} = 0$ iff $\mathbf{x} \in \mathcal{C}$ is called a *parity-check matrix* for $\mathcal{C}$.

  In general, $\mathbf{H}$ has size $(n-k) \times n$.

  The code $\mathcal{C}$ is the null space of $\mathbf{H}$, denoted by $\mathcal{C} = \mathcal{N}(\mathbf{H})$.

- If $\mathcal{C}$ is an $[n, k]$ systematic code,

$$\mathbf{G} = \left[\mathbf{I}_k \ \vdots \ \mathbf{P} \ \right] \ \longleftrightarrow \ \mathbf{H} = \left[-\mathbf{P}^t \ \vdots \ \mathbf{I}_{n-k} \right]$$

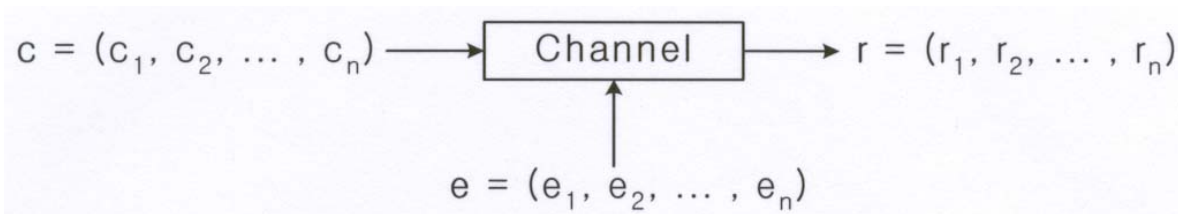
- For any linear code over $\mathbf{F}_q$,

$$\mathbf{G}\mathbf{H}^t = \mathbf{0}.$$

□ Syndrome and Error Detection

- Assume that a codeword $\mathbf{c}$ is transmitted and an error vector $\mathbf{e}$ is added to $\mathbf{c}$ (The channel is assumed to be an *additive* BSC.) Then the received vector $\mathbf{r}$ is given by

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



- The decoder gets the information on the unknown $\mathbf{e}$ from the observation $\mathbf{r}$: The decoder computes the so-called *syndrome*, defined by

$$\mathbf{s} = (s_1, s_2, \ldots, s_{n-k}) \triangleq \mathbf{r}\mathbf{H}^{\mathbf{t}}.$$

- The syndrome value depends only on the errors, but not on the transmitted codeword, since

$$\mathbf{s} = \mathbf{r}\mathbf{H}^{\mathrm{t}} = (\mathbf{c} + \mathbf{e})\mathbf{H}^{\mathrm{t}} = \underbrace{\mathbf{c}\mathbf{H}^{\mathrm{t}}}_{=\,\mathbf{0}} + \mathbf{e}\mathbf{H}^{\mathrm{t}} = \mathbf{e}\mathbf{H}^{\mathrm{t}}.$$

- **Error Detection by Syndrome**:
  - If $\mathbf{s} = \mathbf{0}$, $\mathbf{e} = \mathbf{0}$ (No error) or Undetectable.

    $\Longrightarrow$ Decide that no error occured.

  - If $\mathbf{s} \neq \mathbf{0}$, $\mathbf{e} \neq \mathbf{0}$: Errors are detected.

□ Syndrome for Systematic Codes

- In a systematic code, *the syndrome is the difference* between the received parity bits and the parity bits calculated from the received information bits:

Codeword to be transmitted: $\mathbf{c} = (\underbrace{c_1, c_2, \ldots, c_k,}_{\text{information}}\ \underbrace{c_{k+1}, \ldots, c_n}_{\text{parity}})$

Received vector: $\mathbf{r} = (\underbrace{r_1, r_2, \ldots, r_k}_{\text{received information}}, \underbrace{r_{k+1}, \ldots, r_n}_{\text{received parity bits}}) = (\mathbf{r}_1\ \mathbf{r}_2)$

- Generator and parity-check matrices for a systematic code:

$$\mathbf{G} = [\,\mathbf{I}_k\ \mathbf{P}\,], \qquad \mathbf{H} = [\,-\mathbf{P}^t\ \ \mathbf{I}_{n-k}\,]$$

- The syndrome is

$$\mathbf{s} = \mathbf{r}\,\mathbf{H}^t = \begin{bmatrix}\mathbf{r}_1 & \mathbf{r}_2\end{bmatrix}\begin{bmatrix}-\mathbf{P}\ \mathbf{I}_{n-k}\end{bmatrix} = \underbrace{\mathbf{r}_2}_{\text{received parity bits}} - \underbrace{\mathbf{r}_1\mathbf{P}}_{\substack{\text{parity bits recalculated}\\ \text{the received information}}}$$

□ Example for Syndrome Computation

- Consider the $[7, 4]$ Hamming code with parity-check matrix given by

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

- If $\mathbf{r} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$ is received, the corresponding syndrome is

$$\mathbf{s} = \mathbf{r}\mathbf{H}^t = (0\ 1\ 0).$$

- Decide that *the sixth position is in error*.

□ Hamming Distance and Hamming Weight

- The *Hamming weight* $w_H(\mathbf{x})$ (or $w(\mathbf{x})$) of a vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ is the number of nonzero symbols in $\mathbf{x}$.

- The *Hamming distance* $d_H(\mathbf{x},\ \mathbf{y})$ between two vectors $\mathbf{x}$ and $\mathbf{y}$ is the Hamming weight of their difference vectors $\mathbf{x} - \mathbf{y}$, i.e.,

$$d_H(\mathbf{x},\ \mathbf{y}) = w_H(\mathbf{x} - \mathbf{y}).$$

- The Hamming distance $d_H(\mathbf{x},\ \mathbf{y})$ is a metric. That is,

(a) $d_H(\mathbf{x},\ \mathbf{y}) \geq 0$ for all $\mathbf{x}, \mathbf{y}$ with equality iff $\mathbf{x} = \mathbf{y}$

(b) $d_H(\mathbf{x},\ \mathbf{y}) = d_H(\mathbf{y},\ \mathbf{x})$ (symmetric)

(c) $d_H(\mathbf{x},\ \mathbf{y}) \leq d_H(\mathbf{x},\ \mathbf{z}) + d_H(\mathbf{z},\ \mathbf{y})$ (triangle inequality)

□ Minimum Distance of a Linear Block Code

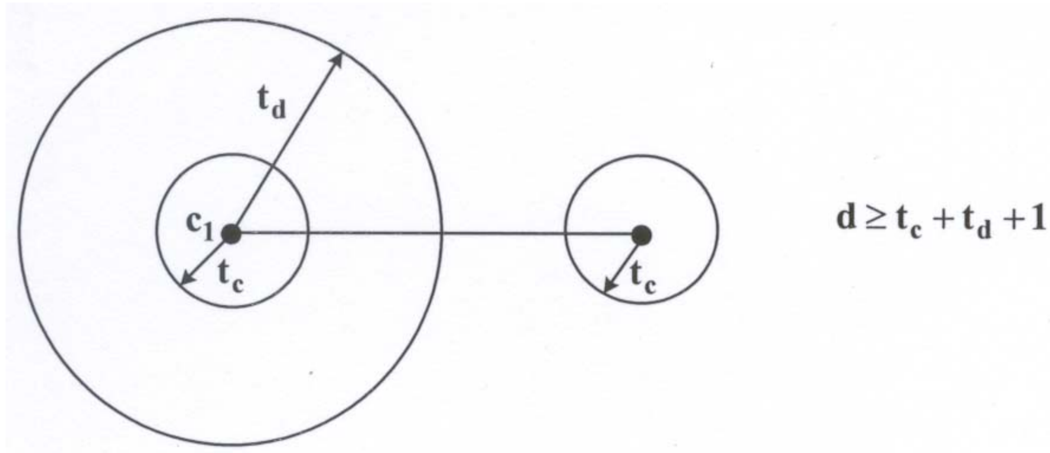
- The *minimum* (*Hamming*) *distance* d_{min} of a linear code $\mathcal{C} \subset \mathbf{F}_q^n$ is the minimum (Hamming) distance between any two distinct codewords, i.e.,

$$\begin{aligned} d_{\min}(\mathcal{C}) &= \min \ \{d_H(\mathbf{x}_1,\ \mathbf{x}_2) \mid \mathbf{x}_1 \in \mathcal{C},\ \mathbf{x}_2 \in \mathcal{C},\ \mathbf{x}_1 \neq \mathbf{x}_2\} \\ &= \min \ \{w_H(\mathbf{x}) \mid \mathbf{x} \in \mathcal{C},\ \mathbf{x} \neq \mathbf{0}\} \end{aligned}$$

- A t_c*-error-correcting* and t_d*-error-detecting* code with $t_d \geq t_c$ is a code that can correct all combinations of ν errors ($\nu \leq t_c$) and detect all combinations of μ errors ($\mu \leq t_d$).

- The code $\mathcal{C}$ has $(t_c,\ t_d)$-error-correction/detection capability iff

$$t_c + t_d + 1 \leq d_{\min}.$$



- A code with minimum distance $d_{\min}$ can correct any patterns of ν errors $(\nu \leq t_c)$ iff

$$2t_c + 1 \leq d_{min}.$$

The number $t_c = \lfloor \frac{d_{min}-1}{2} \rfloor$ is called the *error-correction capability*.

Theorem: Let $\mathcal{C}$ be an $[n, k]$ code with parity check matrix $\mathbf{H}$. There is a codeword of weight w if and only if there are w linearly dependent column of $\mathbf{H}$.

Example: Consider the $[7,4]$ Hamming code with p.c.m. $\mathbf{H}$ given by

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

Determine its minimum distance analytically.

(Solution) Apply Theorem.

- No zero columns $\Rightarrow d_{\min} \geq 2$.
- All columns are distinct.

  $\Rightarrow$ No linear combination of 2 columns is zero.

  $\Rightarrow d_{\min} \geq 3$.
- But, column 1 + column 6 + column 7 $= 0 \quad \Rightarrow \; d_{\min} \leq 3$.

  Therefore, $d_{\min} = 3$. $\square$

□ Minimum Distance of Simple Linear Codes

- $[n, n-1]$ single-parity check code

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

$$d_{\min} = 2 \quad \Rightarrow \quad (t_c, t_d) = (0, 1)$$

- $[7, 4]$ Hamming code.

$$d_{\min} = 3 = t_c + t_d + 1, \;\; t_c \leq t_d$$

$$\Rightarrow \quad (t_c, t_d) = (0, 2) \text{ or } (1, 1)$$

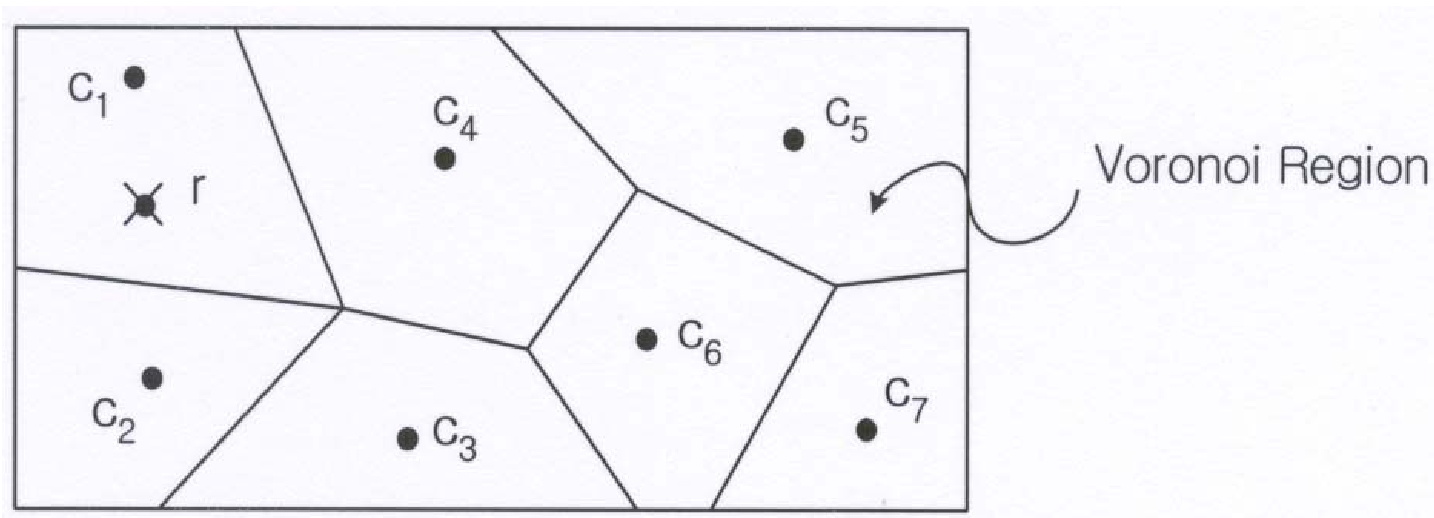
- $[7, 1]$ repetition code

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

$$d_{\min} = 7 \quad \Rightarrow \quad (t_c, t_d) = (0, 6), (1, 5), (2, 4), (3, 3)$$

□ Standard Array and Decoding.

- Known results
  - Optimum decoding rule: *minimum distance decoding*
  - Available information on the error pattern from the given $\mathbf{r}$: *syndrome*
- Decision region



- Let $\mathbf{z}+\mathcal{C}$ be the coset of $\mathcal{C}$ containing $\mathbf{z}$, defined by

$$\mathbf{z}+\mathcal{C} = \{\mathbf{x} \in \mathbf{F}_q^n \mid \mathbf{x} = \mathbf{z}+\mathbf{c}, \mathbf{c} \in \mathcal{C}\}.$$

  – Each vector in $\mathbf{z}+\mathcal{C}$ has the same syndrome as $\mathbf{z}$, since

    $\mathbf{x}\mathbf{H}^t = (\mathbf{z}+\mathbf{c})\mathbf{H}^t = \mathbf{z}\mathbf{H}^{\mathrm{t}}$ for any $\mathbf{x} \in \mathbf{z}+\mathcal{C}$.

  – The most likely error pattern in a coset $\mathbf{z}+\mathcal{C}$ (i.e., the minimum weight error vector in $\mathbf{z}+\mathcal{C}$) is called the *coset leader* of $\mathbf{z}+\mathcal{C}$.

  – If $\mathbf{z}_0$ is the coset leader of $\mathbf{z}+\mathcal{C}$, then $\mathbf{z}_0+\mathcal{C} = \mathbf{z}+\mathcal{C}$.

- **Syndrome Decoding Algorithm**

  1) Compute the syndrome $\mathbf{s} = \mathbf{r}\mathbf{H}^t$.

  2) Find a minimum-weight vector in the coset corresponding to $\mathbf{s}$. Call it $\mathbf{z}_0$.

  3) Output the codeword $\hat{\mathbf{c}} = \mathbf{r} - \mathbf{z}_0$.

□ Construction of Standard Array

- List all codewords in the top row with $\mathbf{0}$ being the first. Set $\mathbf{e}_1 = \mathbf{0}$. The vectors in the top row form the coset $\mathbf{e}_1 + \mathcal{C}$.
- Choose a minimum weight vector $\mathbf{e}_2$ which is not a codeword. List all vectors of $\mathbf{e}_2 + \mathcal{C}$ in the second row so that $\mathbf{e}_2 + \mathbf{c}$ lies below $\mathbf{c}$ for any $\mathbf{c} \in \mathcal{C}$.
- Choose a minimum weight vector $\mathbf{e}_3 \notin \bigcup_{i=1}^{2}(\mathbf{e}_i + \mathcal{C})$ and list $\mathbf{e}_3 + \mathcal{C}$ as before, and repeat the process until no vectors are left.

$$\vdots$$

□ Standard Array for an $[n, k]$ Code $\mathcal{C}$ over $\mathbf{F}_q$

$$\begin{aligned} \text{Number of } n\text{-tuples} &= q^n \\ \text{Number of codewords} &= q^k \triangleq M \\ \text{Number of cosets} &= q^n \Big/ q^k = q^{n-k} \triangleq L \end{aligned}$$

$$\boxed{\begin{array}{ccccc} \mathbf{e}_1 = \mathbf{0} = \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{c}_3 & \cdots & \mathbf{c}_M \\ \mathbf{e}_2 & \mathbf{e}_2 + \mathbf{c}_2 & \mathbf{e}_2 + \mathbf{c}_3 & \cdots & \mathbf{e}_2 + \mathbf{c}_M \\ \mathbf{e}_3 & \mathbf{e}_3 + \mathbf{c}_2 & \mathbf{e}_3 + \mathbf{c}_3 & \cdots & \mathbf{e}_3 + \mathbf{c}_M \\ & & \vdots & & \\ \mathbf{e}_j & \mathbf{e}_j + \mathbf{c}_2 & \mathbf{e}_j + \mathbf{c}_3 & \cdots & \mathbf{e}_j + \mathbf{c}_M \\ & & \vdots & & \\ \mathbf{e}_L & \mathbf{e}_L + \mathbf{c}_2 & \mathbf{e}_L + \mathbf{c}_3 & \cdots & \mathbf{e}_L + \mathbf{c}_M \end{array}} \begin{array}{cc} \Rightarrow & \mathcal{C} = \mathbf{0} + \mathcal{C} \\ \Rightarrow & \mathbf{e}_2 + \mathcal{C} \\ \Rightarrow & \mathbf{e}_3 + \mathcal{C} \\ & \vdots \\ \Rightarrow & \mathbf{e}_j + \mathcal{C} \\ & \vdots \\ \Rightarrow & \mathbf{e}_L + \mathcal{C} \\ & \text{"coset"} \end{array}$$

□ Properties of Standard Array

1) The coset leader $\mathbf{e}_i$ has minimum weight in the corresponding coset (row).

2) Any two vectors in a coset have the same syndrome.

3) No two n-tuples in the same row are identical. Each n-tuple appears only once in the array.

(Proof) Suppose $\mathbf{e}_i + \mathbf{c}_j = \mathbf{e}_i + \mathbf{e}_m$. Then $\mathbf{c}_j = \mathbf{c}_m$, so $j = m$. Now, suppose that $\mathbf{e}_i + \mathbf{c}_j = \mathbf{e}_l + \mathbf{c}_m$, where $i < l$. Then

$$\mathbf{e}_l = \mathbf{e}_i + \mathbf{c}_j - \mathbf{c}_m \in \mathbf{e}_i + \mathcal{C}.$$

This is a contradiction to our choice of $\mathbf{e}_l$. □

4) Every $[n, k]$ linear block code is capable of correcting 2^{n-k} error patterns

(Proof) Assume $\mathbf{r} = \mathbf{c}_i + \mathbf{e}_j$ is received. Then what codeword is the closest to $\mathbf{r}$? If $d(\mathbf{r}, \mathbf{c}_l) < d(\mathbf{r}, \mathbf{c}_i)$ for some $l \neq i$, then $w(\mathbf{r} - \mathbf{c}_l) < w(\mathbf{r} - \mathbf{c}_i)$.

$\Rightarrow w(\underbrace{\mathbf{c}_i - \mathbf{c}_l}_{\in \mathcal{C}} + \mathbf{e}_j) < w(\mathbf{e}_j)$

$\Rightarrow$ Contradiction to our choice of $w(\mathbf{e}_j)$

Hence, $\mathbf{c}_i$ is the closest to $\mathbf{r}$ in terms of Hamming distance and $\mathbf{r}$ is decoded into $\mathbf{c}_i$.

$\Rightarrow$ Number of correctable error patterns $= L = 2^{n-k}$. $\square$

5) Each column contains just one codeword that should be the decoder output for any sequence in the column.

□ Decoding by Standard Array

Step 1: Calculate a syndrome $\mathbf{s}$ by

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

Step 2: Find the coset leader of the corresponding coset $\mathbf{e}$ to $\mathbf{s}$.

Step 3: Compute $\hat{\mathbf{c}} = \mathbf{r} - \mathbf{e}$

Remark:

1) Standard array decoding: syndrome decoding or table look-up decoding.

   This achieves ML decoding or minimum distance decoding.

2) For large $n - k$, this method may be impossible!

□ Example of Standard Array

Consider the [5,2] binary code with generator matrix given by

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

A standard array for the code is constructed as follows:

0 0 0 0 0	1 0 1 0 1	0 1 0 1 1	1 1 1 1 0
1 0 0 0 0	0 0 1 0 1	1 1 0 1 1	0 1 1 1 0
0 1 0 0 0	1 1 1 0 1	0 0 0 1 1	1 0 1 1 0
0 0 1 0 0	1 0 0 0 1	0 1 1 1 1	1 1 0 1 0
0 0 0 1 0	1 0 1 1 1	0 1 0 0 1	1 1 1 0 0
0 0 0 0 1	1 0 1 0 0	0 1 0 1 0	1 1 1 1 1
1 1 0 0 0	0 1 1 0 1	1 0 0 1 1	0 0 1 1 0
1 0 0 1 0	0 0 1 1 1	1 1 0 0 1	0 1 1 0 0

□ Bounded Distance Decoder and Complete Decoder

- Given a (t_c, t_d)-code, a decoder that corrects all patterns of t_c or less errors and detects all patterns of t_d or less errors is called a *bounded distance decoder*.
  Example: BCH/RS decoder using Euclidean algorithm

- A decoder that performs minimum distance decoding is a *complete decoder*.
  Example: Decoder using the standard array.

□ The Dual Code of a Linear Block Code

- Let $\mathcal{C}$ be an $[n, k]$ linear code over $\mathbf{F}_q$. Then the set

$$\mathcal{C}^{\perp} = \{\mathbf{y} \in \mathbf{F}_q^n \,|\, \mathbf{x}^{\mathrm{t}}\mathbf{y} = 0, \ \forall \mathbf{x} \in \mathcal{C}\}$$

  is called the *dual code* of $\mathcal{C}$.

- Let $\mathbf{G}, \mathbf{H}$ be a generator matrix and a parity check matrix of $\mathcal{C}$, respectively. Then

$$\mathbf{H} = \text{a generator matrix for } \mathcal{C}^{\perp}$$
$$\mathbf{G} = \text{a parity-check matrix for } \mathcal{C}^{\perp}$$

  Therefor, $\mathcal{C}^{\perp}$ is an $[n, n-k]$ linear code.

☐ Simple Codes and Their Dual Codes

- The $[n, 1]$ *repetition code* has the following generator matrix $\mathbf{G}$:

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

  The dual code has $\mathbf{G}$ as its p.c.m., so it is the single parity check matrix of an *even parity check code*.

- Let $\mathbf{H}$ be an $m \times (2^m - 1)$ binary matrix with distinct nonzero columns. Then the $[2^m - 1, 2^m - 1 - m]$ linear code with p.c.m. $\mathbf{H}$ is called a *binary Hamming code* and has $d_{\min} = 3$.

  – The Hamming code has high rate:

$$\text{rate} = \frac{2^m - 1 - m}{2^m - 1} \longrightarrow 1 \quad \text{(high)}$$

  – The dual code of a binary Hamming code is a *simplex code*.

□ Weight Enumerator

- Let $\mathcal{C}$ be an $[n, k]$ linear code and let A_i be the number of codewords of weight i. The *weight enumerator* A of $\mathcal{C}$ is defined by

$$A(z) = A_0 + A_1 z + \cdots + A_n z^n.$$

- $A(1) = A_0 + A_1 + \cdots + A_n = q^k = \left|\mathcal{C}\right| =$ Number of codewords.
  $A_0 = 1$

- **MacWilliams Identity**: Let $A(z)$ be the weight enumerator of an $[n, k]$ linear code $\mathcal{C}$ and $B(z)$ be the weight enumerator of the dual code $\mathcal{C}^\perp$ of $\mathcal{C}$. Then

$$q^k B(z) = \left[1 + (q-1)z\right]^n A\left(\frac{1-z}{1+(q-1)z}\right)$$

Example: Consider the $[7,4]$ binary Hamming code $\mathcal{C}$ defined by

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

It is easily shown that $A_0 = 1, A_3 = A_4 = 7, A_7 = 1; \quad A_i = 0,$ elsewhere. Thus,

$$A(z) = 1 + 7z^3 + 7z^4 + z^8.$$

On the other hand, its dual code $\mathcal{C}^\perp$ has $B_0 = 1, B_4 = 7; \quad B_i = 0,$ elsewhere. This leads to $B(z) = 1 + 7z^4$.

Check:

$$2^4 \; B(z) \; = \; (1+z)^7 \; A\left(\frac{1-z}{1+z}\right)$$

$\square$

□ Probability of Error in Linear Codes

- Because of linearity, the performance of a linear code (with respect to the error probability) is the same regardless of the particular codeword transmitted, i.e.,

$$P(E \,|\, \mathbf{c}) = P(E \,|\, \mathbf{0})$$

for any $\mathbf{c} \in \mathcal{C}$.

- Therefore, we will assume that $\mathbf{0} = (00 \cdots 0)$ is the transmitted codeword.

- The average error probability is given by

$$P(E) \; = \; \textstyle\sum_{\mathbf{c} \in \mathcal{C}} P(E \,|\, \mathbf{c}) \cdot P(\mathbf{c}) \; = \; P(E \,|\, \mathbf{0})$$

$\hookrightarrow$ linear

□ Probability of Undetected Error (over BSC)

- Assume that the code is used only for error detection. The decoding algorithm at the receiver declares:

$$\mathbf{r} \in \mathcal{C} \;\Rightarrow\; \text{No errors have occurred.}$$

$$\mathbf{r} \notin \mathcal{C} \;\Rightarrow\; \text{Errors have occurred.}$$

- Let P_u be the *probability of undetected error*, that is, the probability of failing to detect an error when an error has taken place. In other words, P_u is the probability that the error pattern is one of the non-zero codewords.

- Let $A(z) = \sum_{i=0}^{n} A_i z^i$ be the weight enumerator of the code $\mathcal{C}$. Then

$$P_u = \sum_{i=1}^{n} A_i \epsilon^i (1-\epsilon)^{n-i} = (1-\epsilon)^n \cdot \sum_{i=1}^{n} A_i \left(\frac{\epsilon}{1-\epsilon}\right)^i = (1-\epsilon)^n \cdot \left(A(z) \Big|_{z=\rho} - 1 \right).$$

Example: Consider the $[6, 3, 3]$ code $\mathcal{C}$ generated by

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

The code $\mathcal{C}$ has weight enumerator given by

$$A(z) = 1 + 4z^3 + 3z^4$$

Therefore, the undetected error probability P_u is given by

$$\begin{aligned} P_u &= (1-\epsilon)^6 \left(4\rho^3 + 3\rho^4\right) \\ &= \epsilon^3 (1-\epsilon)^2 (4-\epsilon) \end{aligned}$$

where $\rho = \epsilon/(1-\epsilon)$. □

□ Probability of Incorrectly Decoding

- Assume that minimum distance decoding is employed and that the code is used only for error-correction.

- Let P_w be the *probability of incorrectly decoding*, in other words, the probability that the error pattern is not a coset leader. Then

$$P_w = 1 - \text{Prob}\{\text{the error pattern is a coset leader}\}.$$

P_w is also called the *probability of word error* (or *probability of block error*)

Example: For the [5,2] code over the BSC with $\rho = \epsilon/(1-\epsilon)$,

$$\begin{aligned} 1 - P_w &= (1-\epsilon)^5 + 5(1-\epsilon)^4\rho \; + \; 2(1-\epsilon)^3\rho^2 \\ &= (1-\epsilon)^5 \left(1 + 5\rho + 2\rho^2\right). \end{aligned}$$

- BER analysis is more complicated.

□ Modification of Linear Block Codes

• Extension

– Parameters: $[n, k, d] \quad \longrightarrow \quad [n+1, k, \geq d]$

– Extending procedure: $\mathbf{c} = (c_1, c_2, \ldots, c_n) \quad \longrightarrow \quad (c_1, c_2, \ldots, c_n, c_{n+1})$

where $c_{n+1} = c_1 + c_2 + \cdots + c_n$. (overall parity-check bit)

– If d is odd, we get an $[n+1, k, d+1]$ code.

Example: $[7, 4, 3]$ Hamming code $\quad \longrightarrow \quad$ $[8, 4, 4]$ extended Hamming code

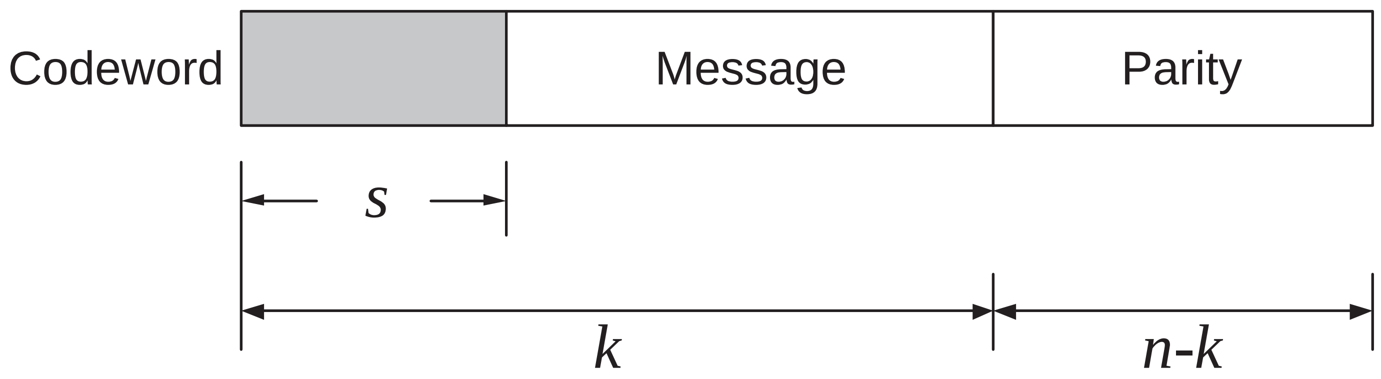
– Parity-check matrix for an extended code

$$H_E = \begin{bmatrix} & & & & 0 \\ & & H & & \vdots \\ & & & & 0 \\ 1 & 1 & \cdots & 1 & 1 \end{bmatrix}$$

- **Shortening**

  – Parameters: $[n, k, d] \quad \longrightarrow \quad [n-s, k-s, \geq d]$

  – Structure of codewords



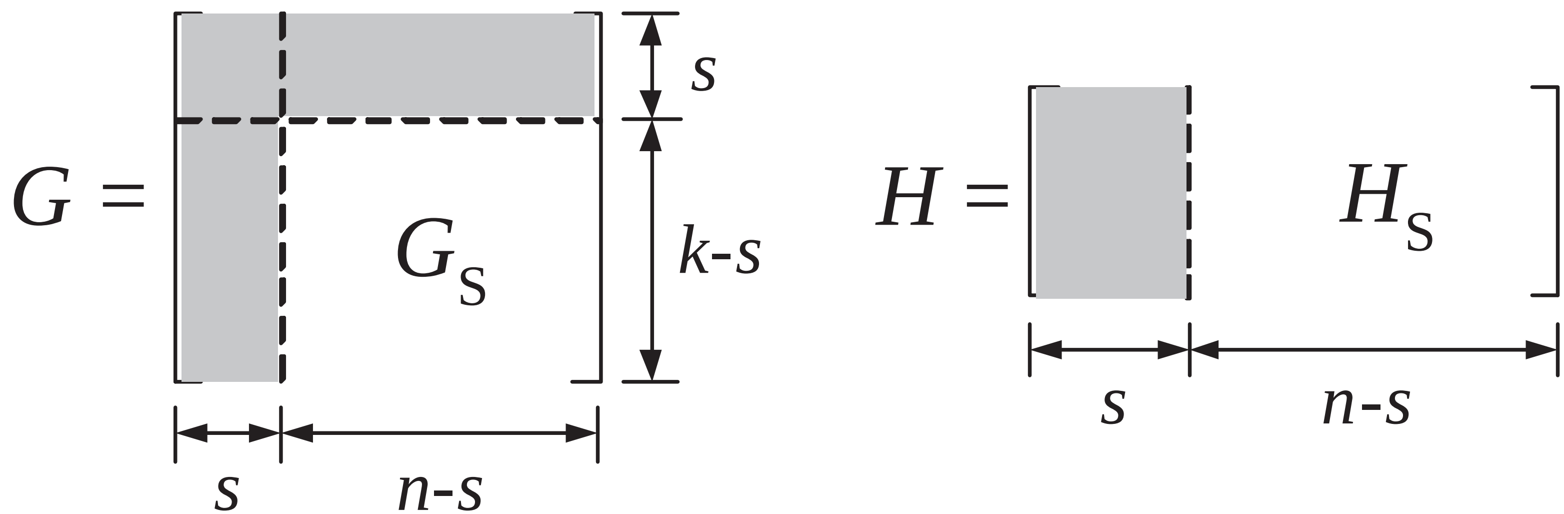
  – Shortening procedure

    1) Assume $G = \begin{bmatrix} I_k & P \end{bmatrix}$ and $H = \begin{bmatrix} -P^t & I_{n-k} \end{bmatrix}$.

    2) Set the first s bits of the information part to zero.

    3) Generate a codeword and then remove the first s bits from it.

– Generator matrix and parity-check matrix for a shortened code



- **Puncturing**
  - Parameters: $[n, k, d] \quad \longrightarrow \quad [n-1, k, \geq d-1]$
  - Puncturing procedure

    1) Delete any fixed coordinate from each codeword.

    2) If $d > p$, then any s coordinates can be deleted without changing the dimension of the code in general. In that case the code has the following parameters:

$$[n, k, d] \quad \longrightarrow \quad [n-p, k, \geq d-p]$$

□ Bounds on Block Codes

• Singleton bound

– Key idea: Let $\mathbf{G} = [\ \mathbf{I}_k\ \mathbf{P}\ ]$ be a generator matrix of an $[n, k, d_{\min}]$ linear block code over $\mathbf{F}_q$. Every row in $\mathbf{G}$ is a codeword and has at most $(n-k)+1$ nonzero components.

– For any $[n, k, d_{\min}]$ linear block code over $\mathbf{F}_q$, we have

$$d_{\min} \leq n - k + 1.$$

– If $\mathcal{C}$ is a linear code satisfying the Singleton bound with equality, then $\mathcal{C}$ is called an *MDS* (*maximum distance separable*) code.

Example: $[n, 1, n]$ repetition code, $[n, n-1, 2]$ single parity check code,
Reed-Solomon codes.

- **Hamming bound**

  – If an (n, M) code over $\mathbf{F}_q$ can correct any pattern of t or less errors, then the spheres of radius t centered at codewords must be disjoint.

  – In a sphere of radius t centered at each codeword, there are

  $$\binom{n}{0}(q-1)^0 + \binom{n}{1}(q-1)^1 + \binom{n}{2}(q-1)^2 + \cdots + \binom{n}{t}(q-1)^t$$

  vectors. Since the total number of vectors in the space $\mathbf{F}_q^n$ is q^n, we have

  $$\left[\sum_{i=0}^{t}\binom{n}{i}(q-1)^i\right]\cdot M \;\le\; q^n$$

  – In an $[n, k, d]$ linear code over $\mathbf{F}_q$,

  $$\sum_{i=0}^{t}\binom{n}{i}(q-1)^i \;\le\; q^{n-k} \quad (= \text{number of correctable error patterns})$$

- **Perfect codes**:

  An (n, M, d) code is said to be *perfect* if the parameters n, M, d satisfy the Hamming bound with equality.

  Examples of perfect codes

  – The $[2^r - 1, 2^r - 1 - r, 3]$ binary Hamming code is perfect.

  – The $[23, 12, 7]$ binary Golay code is perfect.

  – The $[11, 6, 5]$ ternary Golay Code is perfect.

  – Trivial binary perfect codes

  1) $[n, 1, n]$ repetition code with odd length;

  2) a code with only one codeword (can correct n errors);

     for example, $\mathcal{C} = \{\mathbf{0}\}$ is assumed to have minimum distance ∞.

  3) the code with all vectors, that is, the entire space $\mathcal{C} = \mathbf{F}_2^n$.

- **Plotkin Bound**

  – Key idea: Minimum distance $\leq$ average weight of all nonzero codewords. This implies that

  $$d_{\min} \leq \frac{1}{|\mathcal{C}|-1} \sum_{\mathbf{x}\in\mathcal{C}} w(\mathbf{x}).$$

  – Key fact: Let $\mathcal{C}$ be an $[n,k]$ code over $\mathbf{F}_q$. For $a \in \mathbf{F}_q$ and any $i = 1, 2, \cdots, n$, let $\mathcal{C}_i(a) = \{\mathbf{c} = (c_1, \cdots, c_n) \in \mathcal{C} \mid c_i = a\}$. Then either $|\mathcal{C}_i(0)| = |\mathcal{C}|$ or $|\mathcal{C}_i(a)| = \frac{1}{q}|\mathcal{C}|$.

  – For a binary linear $[n,k,d]$ code, $d \leq n(2^k - 2^{k-1})/(2^k - 1)$.

  – Plotkin bound for a binary (n, M, d) code: If $n < 2d$, then

  $$M \leq 2\left\lfloor \frac{d}{2d-n} \right\rfloor.$$

- **Griesmer bound**

  Let $N(k,d)$ be the smallest n for a linear code $\mathcal{C}$ of dimension k and minimum distance d. Then

$$N(k,d) \;\geq\; \sum_{i=0}^{k-1} \left\lceil \frac{d}{2^i} \right\rceil.$$

  (Sketch of proof)

  – WLOG, we can assume that a generator matrix for an $[N(k,d),k,d]$ code $\mathcal{C}$ is

$$\mathbf{G} \;=\; \underbrace{\begin{bmatrix} 1 & 1 & \cdots & 1 & \vdots & 0 & 0 & \cdots & 0 \\ & & \mathbf{G}_1 & & \vdots & & \mathbf{G}_2 & & \end{bmatrix}}_{\longleftarrow\; d \;\longrightarrow \longleftarrow N(k,d)-d \longrightarrow}$$

  where $\mathbf{G}_2$ generates an $[N(k,d)-d,k-1]$ code $\mathcal{C}_2$ of minimum distance d_2.

– Let $\mathbf{u} \in \mathcal{C}_2$ be such that $w(\mathbf{u}) = d_2$ and choose $\mathbf{v}$ such that $(\mathbf{v} \vdots \mathbf{u}) \in \mathcal{C}$. Then

$$w(\mathbf{v} \vdots \mathbf{u}) = w(\mathbf{v}) + w(\mathbf{u}) = w(\mathbf{v}) + d_2 \geq d. \quad (1)$$

Also, we have $(\mathbf{1} + \mathbf{v} \vdots \mathbf{u}) \in \mathcal{C}$ by linearity. Therefore,

$$d - w(\mathbf{v}) + d_2 \geq d. \quad (2)$$

From (1) and (2), we have $2d_2 \geq d$, i.e., $d_2 \geq \lceil d/2 \rceil$.

– From the existence of $\mathcal{C}_2$, we have

$$\begin{aligned} N(k,d) - d &\geq N(k-1, d_2) \\ &\geq N(k-1, \lceil d/2 \rceil) \end{aligned}$$

Therefore, we have a recursion: $N(k,d) \geq d + N(k-1, \lceil d/2 \rceil))$.

– By applying the process iteratively, we get the bound. □

- **Gilbert-Varshamov bound**
  - *Best packing* of radius $d-1$ in volume $\mathbf{F}_q^n$:

$$M \cdot \sum_{i=0}^{d-1} \binom{n}{i} (q-1)^i \geq q^n$$

  - Key idea for construction: Let $\mathbf{H} = \begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \cdots & \mathbf{h}_n \end{bmatrix}$ be a parity check matrix. Then there exists s linearly dependent columns iff there is a codeword of weight s.
  - Constructive bound: An $[n, k, d]$ code exists if

$$\sum_{i=0}^{d-2} \binom{n-1}{i} (q-1)^i < q^{n-k} \leq \sum_{i=0}^{d-1} \binom{n-1}{i} (q-1)^i.$$

□ Asymptotic Bounds

- Let $A(n,d)$ be the *maximum number of codewords* in any binary code (linear or nonlinear) of length n and minimum distance d between codewords.
  The *relative minimum distance* δ is defined as $\delta = \lim_{n\to\infty} \frac{d}{n}$.
  The rates $\bar{R}(\delta)$ and $\underline{R}(\delta)$ are given by

$$\begin{aligned} \bar{R}(\delta) &= \lim_{n\to\infty} \sup \frac{1}{n} \log_2 A(n,d), \\ \underline{R}(\delta) &= \lim_{n\to\infty} \inf \frac{1}{n} \log_2 A(n,d). \end{aligned}$$

- Asymptotic bounds
  - Hamming bound: $\bar{R}(\delta) \le 1 - H_2(\delta/2), \quad 0 \le \delta \le 1$
  - Plotkin bound: $\bar{R}(\delta) \le 1 - 2\delta, \quad 0 \le \delta \le 1/2$
  - G-V bound: $\underline{R}(\delta) \ge 1 - H_2(\delta), \quad 0 \le \delta \le 1/2$

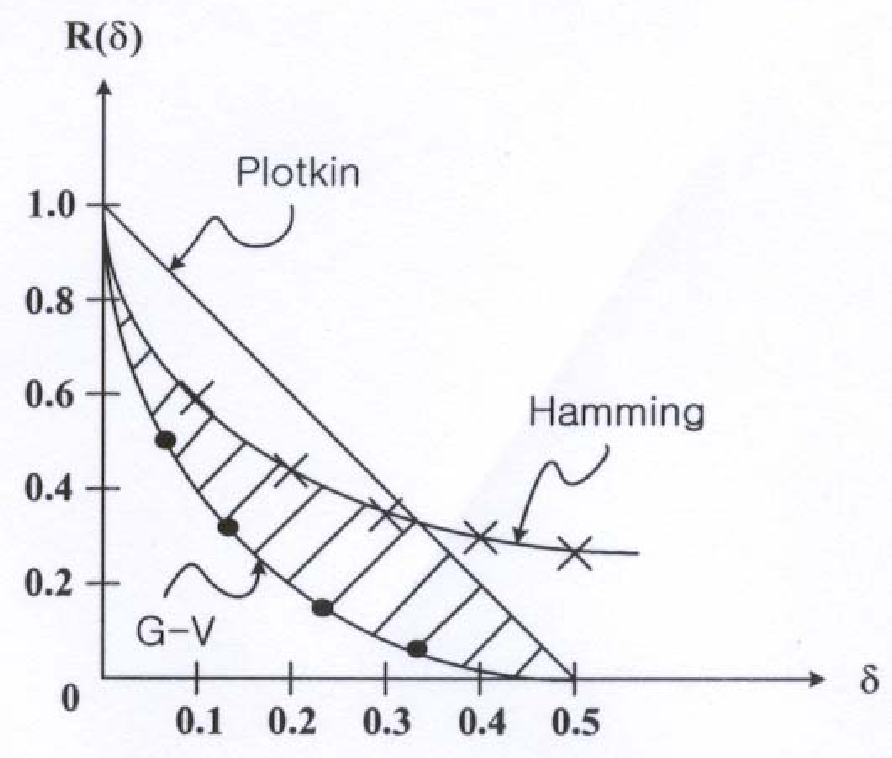
  where $H_2(\lambda) = -\lambda \log_2 \lambda - (1-\lambda)\log_2(1-\lambda)$.

Remark: $\underline{R}(\delta) \leq \bar{R}(\delta)$

$$\delta = 0: \quad \bar{R} \leq 1, \quad \underline{R} \geq 1 \;\Rightarrow\; \bar{R} = \underline{R} = 1$$

$$\delta = 1/2: \quad \bar{R} \leq 0, \quad \underline{R} \geq 0 \;\Rightarrow\; \bar{R} = \underline{R} = 0$$

δ	Hamming	Plotkin	G-V
0	1	1	1
.1	.714	.8	.531
.2	.531	.6	.278
.3	.390	.4	.119
.4	.278	.2	.029
.5	.189	0	0



□ Summary of Linear Block Codes

- A linear block code over a finite field F is a subspace of the vector space F^n
- A linear block code is the row space of a generator matrix.
- A linear block code is the null space of a parity-check matrix.
- A linear block code can be decoded using the standard array.

  (But, the complexity is too high.)
- There is a trade-off among the parameters of a code.