Stanford University

ENGR 76

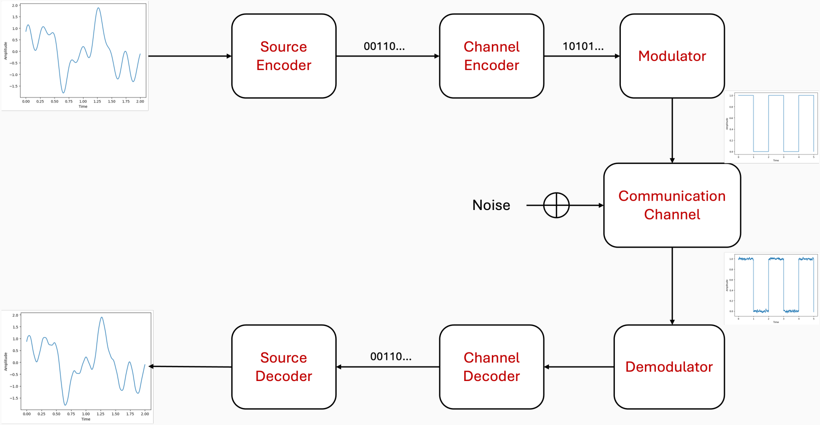
Information Science and Engineering

Lecture 13: Error Correction Codes

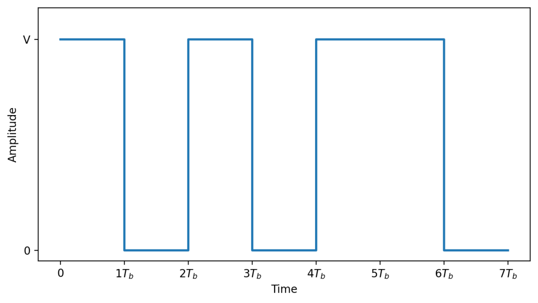
Siddharth Chandak

Recap

Digital Communication



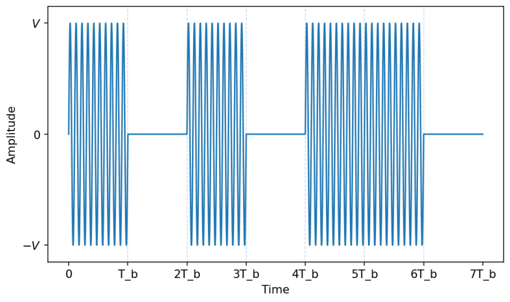
Modulation: On-Off Keying



- Bit Duration: T_b
- Bit Rate $R_b = \frac{1}{T_b}$

Modulation: Passband Signal

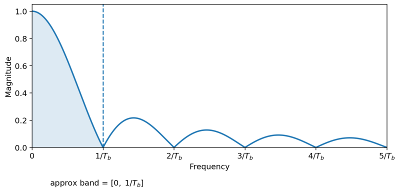
- $x(t)$ generated using on-off keying with bit duration T_b
- Let $s(t) = x(t) \times \sin(2\pi f_c t)$



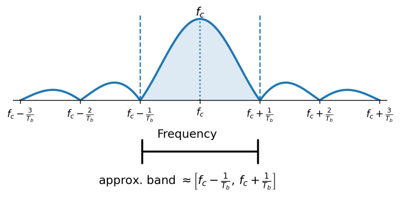
Carrier Wave and Carrier Frequency

- $x(t)$ is the message signal
- $\sin(2\pi f_c t)$ 'carries' it at high frequencies: **Carrier Wave**
- f_c: **Carrier Frequency**
    - Typically $f_c \gg 1/T_b$

Spectrum



(a) Spectrum of $x(t)$



(b) Spectrum of $s(t)$

- For the same rate $1/T_b$, bandwidth of baseband signal is $1/T_b$ and of passband signal is $2/T_b$

Demodulation: Power Detection and Thresholding

1. Divide time into bit intervals of length T_b, i.e.,

$$[(m-1)T_b,\, mT_b], \qquad m = 1, 2, \ldots$$

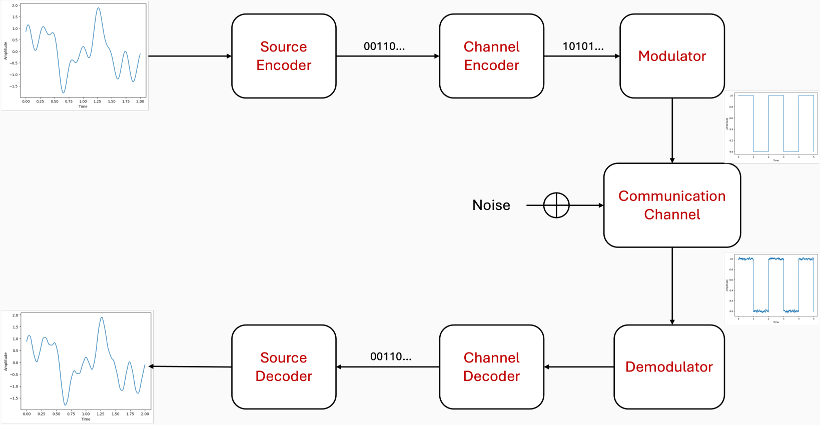
2. For each bit interval, compute the average received power

$$p_m \;=\; \frac{1}{T_b} \int_{(m-1)T_b}^{mT_b} y^2(t)\, dt.$$

3. Compare the power to a threshold and decide the bit:

$$\hat{b}_m = \begin{cases} 1, & p_m > p_{\textsf{thresh}}, \\ 0, & p_m \le p_{\textsf{thresh}}. \end{cases}$$

Digital Communication



Error Correction Codes

Error Correction Codes

- Adding redundancy (extra bits) so that errors caused by the channel can be handled at the receiver
- Two goals: Detect vs Correct?

Error Detection

Detection:

- Receiver checks consistency
  - Could this have been transmitted?
- If not valid: error declared
- Typical action: discard and request retransmission

Error Correction

Correction:

- Receiver finds the most likely transmitted codeword
    - What was actually transmitted?
- Can correct errors up to a certain limit
- No retransmission required
- Requires more redundancy as compared to detection

Example: Error Detection in Internet Packet

Ethernet II

Destination MAC 6 Bytes	Source MAC 6 Bytes	Type 2 Bytes	Data 46 – 1500 Bytes	Frame Check Sequence 4 Bytes

Example: Error Detection in Bar Codes

Example: Error Correction in QR Codes

Example: Error Correction in CDs

Example: Error Correction

- Cellular Communication: corrects bit errors due to noise
- Satellite Communication:
  - Retransmission impractical due to long delays
  - Strong error correction

Compression vs Error-Correction

- Both map messages to codewords
- **Compression**
  - Remove *statistical* redundancy to represent using fewer number of bits
- **Error-Correction**
  - Add *structured* redundancy to protect against channel noise

Repetition Codes

Simplest Scheme

- You want to transmit just one bit
    - Only two messages: 0 or 1
- Simplest idea for adding redundancy?

Repetition Codes

- Map 0 to 000
- Map 1 to 111
- Example:
  - Information Sequence: 1011
  - Encoded Sequence: 111000111111

How will we decode?

- Suppose received sequence is: 101100111100
- Decoding:
  - Divide into blocks of 3
  - Map each block to “closest” codeword
  - Map decoded codeword to message

How will we decode?

- Information Sequence: 1011
- Encoded Sequence: 111000111111
- Received Sequence: 101100111100
- Decoded Sequence: 111000111000
- Decoded Information: 1010

Why did we repeat by 3?

- Suppose we map 0 to 00 and 1 to 11
- Can we correct one bit flip?
- Can we detect one bit flip?

Why did we repeat by 3?

- Suppose we map 0 to 00 and 1 to 11
- Can we correct one bit flip?
  - No!
  - Suppose we receive 01 - then we have no idea to know what the transmitted bits were
- Can we detect one bit flip?
  - Yes!
  - If we receive 01 or 10, we can detect an error

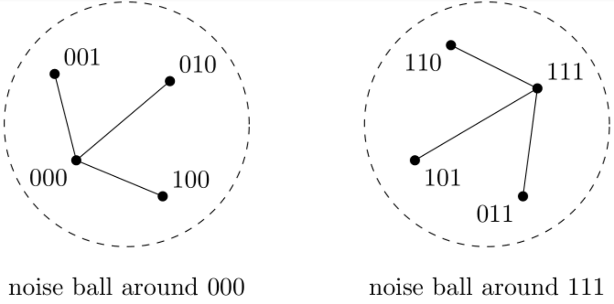
Why did we repeat by 3?

- Suppose we map 0 to 000 and 1 to 111
- Can we correct one bit flip?
- Can we correct two bit flips?

Why did we repeat by 3?

- Suppose we map 0 to 000 and 1 to 111
- Can we correct one bit flip?
  - Yes!
  - One bit flip can lead 000 to 100, 010, or 001
  - All of these are "closer" to 000 than 111
  - Decoded sequence will be 000
- Can we correct two bit flips?
  - No!
  - Two bit flips can lead 000 to 011 (for example)
  - Closer to 111 than 000
  - Decoded sequence will be 111 and decoded bit will be 1 - error not corrected
- Formalization?

Noise Balls



- Noise balls of 'radius' 1

Hamming Distance and Error-Correcting Capabilities

Hamming Distance

Definition (Hamming Distance)

For two binary vectors of equal length, the Hamming distance is the number of positions where they differ.

Examples

- $d_H(000, 111)$
- $d_H(01110, 10101)$

Hamming Distance

Definition (Hamming Distance)

For two binary vectors of equal length, the Hamming distance is the number of positions where they differ.

Examples

- $d_H(000, 111) = 3$
- $d_H(01110, 10101) = 4$

Minimum Distance of the Code

- The minimum Hamming distance between any two codewords
- Denoted by d_{min}
- Example:
  - Suppose $\mathcal{C} = \{00000, 00111, 11100, 11011\}$

Minimum Distance of the Code

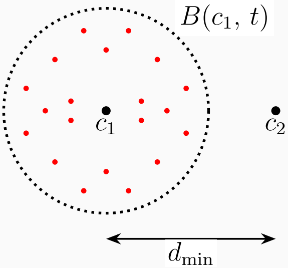
- The minimum Hamming distance between any two codewords
- Denoted by d_{min}
- Example:
  - Suppose $\mathcal{C} = \{00000, 00111, 11100, 11011\}$
    - $d_H(00000, 00111) = 3$
    - $d_H(00000, 11100) = 3$
    - $d_H(00000, 11011) = 4$
    - $d_H(00111, 11100) = 4$
    - $d_H(00111, 11011) = 3$
    - $d_H(11100, 11011) = 3$
  - $d_{min} = 3$

Alternate perspectives

- So far: adding redundancy to messages
- Equivalent view: a code is a subset of binary strings
- Good codes have codewords that are far apart
- Larger $d_{min} \implies$ better error-correction capabilities

Error Detection

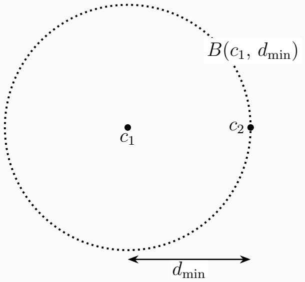
- How many bit flips can a code with minimum distance d_{min} detect?



- c_1 and c_2 are two codewords
- $B(c_1, t)$ is Hamming ball of radius t
- Red points inside the ball: Bit strings at most t bits away from c_1
  - Not a part of the code
- Any received word in this ball cannot equal another codeword $\Longrightarrow$ error detected

Error Detection

- How many bit flips can a code with minimum distance d_{min} detect?



- If d_{min} bit flips happen, then we can receive c_2 even though c_1 was transmitted
- Since $c_2 \in \mathcal{C}$, no error is detected
- Hence, we can detect at most $d_{min} - 1$ bit flips

Error Detection

Theorem (Error Detection)

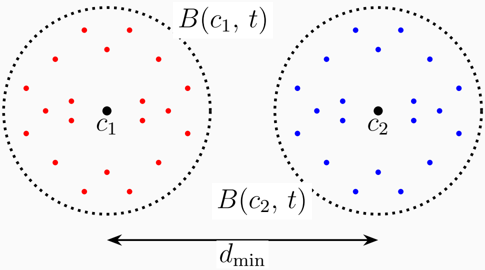
A code can detect up to t bit flips if and only if

$$d_{\min} \geq t + 1.$$

Equivalently, a code with minimum distance $d_{\min}$ can detect up to $d_{\min} - 1$ bit flips.

Error Correction

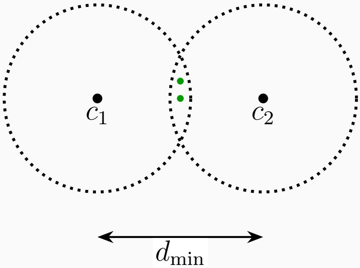
- How many bit flips can a code with minimum distance d_{min} correct?



- Red points: bit strings within Hamming distance t of c_1
- Blue points: bit strings within Hamming distance t of c_2
- We can correct up to t bit flips as long as these balls do not intersect

Error Correction

- How many bit flips can a code with minimum distance d_{min} correct?



- Cannot correct errors in case of intersection
- Non-intersection requires $d_{min} > 2t$

Error Correction

Theorem (Error Correction)

A code can correct up to t bit flips if and only if

$$d_{\min} \geq 2t + 1.$$

Equivalently, a code with minimum distance $d_{\min}$ can correct up to

$$\left\lfloor \frac{d_{\min} - 1}{2} \right\rfloor$$

bit flips.

- Here $\lfloor x \rfloor$ denotes the floor operator, largest integer smaller than or equal to x
- $\lfloor 1 \rfloor = 1$ and $\lfloor 1.5 \rfloor = 1$

Figures of Merit

Figures of Merit

- n: Number of bits in each codeword
- k: Number of information bits
  - $M = 2^k$ is the number of messages or codewords
- d_{min}: Minimum distance of the code
- Rate $= \frac{k}{n}$
  - Ratio of information bits to the total number of bits in a codeword

Examples

- Repetition code with 2 messages:
  - $0 \to 000$ and $1 \to 111$
- Repetition code with 4 messages:
  - $00 \to 000000, 01 \to 010101, 10 \to 101010, 11 \to 111111$
- Above example -
  - $00 \to 00000, 01 \to 00111, 10 \to 11100, 11 \to 11011$

Examples

- Repetition code with 2 messages:
  - $0 \to 000$ and $1 \to 111$
- $n = 3$ (number of bits in each codeword)
- $k = 1$ (one information bit)
- $M = 2$ (2 messages or 2 codewords)
- Rate $= 1/3$
- $d_{min} = 3$ (distance between codewords)

Examples

- Repetition code with 4 messages:
  - $00 \to 000000, 01 \to 010101, 10 \to 101010, 11 \to 111111$
- $n = 6$ (number of bits in each codeword)
- $k = 2$ (two information bits)
- $M = 4$ (4 codewords)
- Rate $= 1/3$
- $d_{min} = 3$ (distance between codewords)

Examples

- $00 \to 00000, 01 \to 00111, 10 \to 11100, 11 \to 11011$
- $n = 5$ (number of bits in each codeword)
- $k = 2$ (two information bits)
- $M = 4$ (4 codewords)
- Rate $= 2/5$
- $d_{min} = 3$ (distance between codewords)

Design Goals

We would like:

- Larger k and M: more information bits
- Small n: Want smaller transmissions
- Larger rate
- Larger d_{min}: Better error correction capabilities

Exam

Exam

- Will be held in class on February 26
  - Regular class time and location
  - Will start precisely at 9am
- **Syllabus**: Lectures 1-12 (inclusive)
  - Questions similar to Mini-PSets (1-6) and practice problems from PSet Discussions
- Allowed to bring one two-sided hand-written sheet for reference
  - No calculators
- Will release sample exam today

OAE Accommodations

- If you plan to use OAE accommodations for the exam, please email us by February 22
  - **Important:** You must email us this week even if you have already shared your OAE letter earlier in the quarter
  - Please include the specific accommodations you will require for this exam
  - Requests received after February 22 may not be possible to accommodate

Thank You!