ELEC 405
Error Control Coding and Sequences

Binary Linear Block Codes

Basic Concept

- The techniques used to protect information against noise and interference can be complex, but the basic principles are easily understood.

- The key idea is to encode the message by adding some redundant information or parity to the message.

- In such a case, even if the message is corrupted, the redundancy in the encoded message added at the transmitter can be used for error detection and/or correction at the receiver.

Modular Arithmetic

In mod (modulo) 2 arithmetic, 2 is the base (modulus) and there are no numbers other than 0 and 1. Any higher number mod 2 is obtained by dividing it by 2 and taking the remainder.

For example, $3 \equiv 1 \bmod 2$ and $4 \equiv 0 \bmod 2$.

Mod 2 addition

+	0	1
0	0	1
1	1	0

} same as logical XOR ($\oplus$)

Mod 2 multiplication

*	0	1
0	0	0
1	0	1

} same as logical AND

Binary Linear Block Codes

- Linear codes are the most important class of error correcting codes
  - simple description
  - nice properties
  - easy encoding
  - conceptually easy decoding
- Binary linear code: mod 2 sum of any two codewords is a codeword
- Block code: codewords have a finite length *n*
- The number of codewords in a code *C* is

$$|C| = M = 2^k$$

- Each codeword of *n* bits represents *k* data bits
- The code rate is

$$R = \frac{\log_2 M}{n} = \frac{k}{n}$$

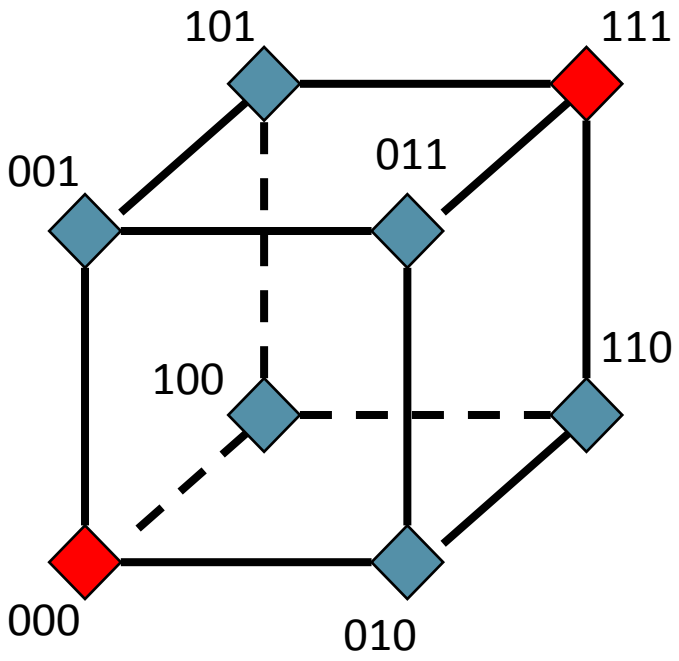
Triple Repetition Code

Data	Codeword		
0	0	0	0
1	1	1	1

Triple Repetition Codes – Decoding

Received Word			Codeword			Error Pattern		
0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	1
0	1	0	0	0	0	0	1	0
1	0	0	0	0	0	1	0	0
1	1	1	1	1	1	0	0	0
1	1	0	1	1	1	0	0	1
1	0	1	1	1	1	0	1	0
0	1	1	1	1	1	1	0	0

Triple Repetition Code



Vector Spaces

- Set of n-tuples over an alphabet *A*
  - n-dimensional vector space
- Example: binary n-tuples of length 5 - V_5
  - 5-dimensional vector space

$$\left.\begin{array}{l} 00000 \\ 00001 \\ 00010 \\ 00011 \\ 00100 \\ \vdots \\ 11111 \end{array}\right\} 32 \quad 5\text{-tuples}$$

Vector Space Operations

vector addition

$$\begin{array}{r} 11001 \\ +10011 \\ \hline 01010 \end{array}$$

scalar multiplication

$$a \bullet \bar{v}$$

$$0 \bullet (11001) = 00000$$

$$1 \bullet (11001) = 11001$$

$$a \in A$$

The space is closed under vector addition and scalar multiplication

Inner Product

$$\bar{u} \bullet \bar{v} = \sum_{i=0}^{n-1} u_i \bullet v_i$$

$$\begin{aligned} (11001) \bullet (10011) &= 1 \bullet 1 + 1 \bullet 0 + 0 \bullet 0 + 0 \bullet 1 + 1 \bullet 1 \\ &= 2 \\ &= 0 \bmod 2 \end{aligned}$$

(11001) and (10011) are orthogonal

Vector Subspaces

- A smaller vector space which is closed under vector addition and scalar multiplication
- Example: subspace of V_5

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \end{matrix}$$

$$\begin{array}{r} 00111 \\ +11011 \\ \hline 11100 \end{array}$$

Basis

- A minimal number of linearly independent vectors that span the space

$$\begin{bmatrix} 00111 \\ 11100 \end{bmatrix}$$

$$\begin{aligned} 0 \bullet (00111) + 0 \bullet (11100) &= 00000 \\ 0 \bullet (00111) + 1 \bullet (11100) &= 11100 \\ 1 \bullet (00111) + 0 \bullet (11100) &= 00111 \\ 1 \bullet (00111) + 1 \bullet (11100) &= 11011 \end{aligned}$$

- Any vector in the space is a linear combination of basis vectors

Dual Spaces

- Set of vectors orthogonal to a vector space

S	$S^{\perp}$
0000	0000
0101	1010
0001	1000
0100	0010

Vector Space Dimensions

- If a basis has *k* elements then the vector space is said to have dimension *k*

$$S \qquad\qquad S^{\perp}$$

$$\begin{bmatrix} 0001 \\ 0100 \end{bmatrix} \qquad\qquad \begin{bmatrix} 1000 \\ 0010 \end{bmatrix}$$

$$\dim(S) + \dim(S^{\perp}) = \dim(V)$$

Example

- For the subspace generated by the basis

$$\begin{bmatrix} 00111 \\ 11100 \end{bmatrix}$$

what is the dual space?

Self-Dual Spaces

$$S = S^{\perp}$$

- Example

S	$S^{\perp}$
0000	0000
1010	1010
0101	0101
1111	1111

Linear Codes as Vector Spaces

Codewords can be considered as vectors in the vector space V_n of binary vectors of length n.

Definition A subset $C \subseteq V_n$ is a binary linear block code if

(1) $u + v \in C$ for all $u, v \in C$

(2) $au \in C$ for all $u \in C$, $a \in \{0,1\}$

C is a k-dimensional subspace of V_n

C is called a linear (n,k) code

C has 2^k codewords

The single parity check codes and repetition codes are examples of binary linear block codes.

Basis Matrices

- Triple repetition code $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$

- Single parity check code

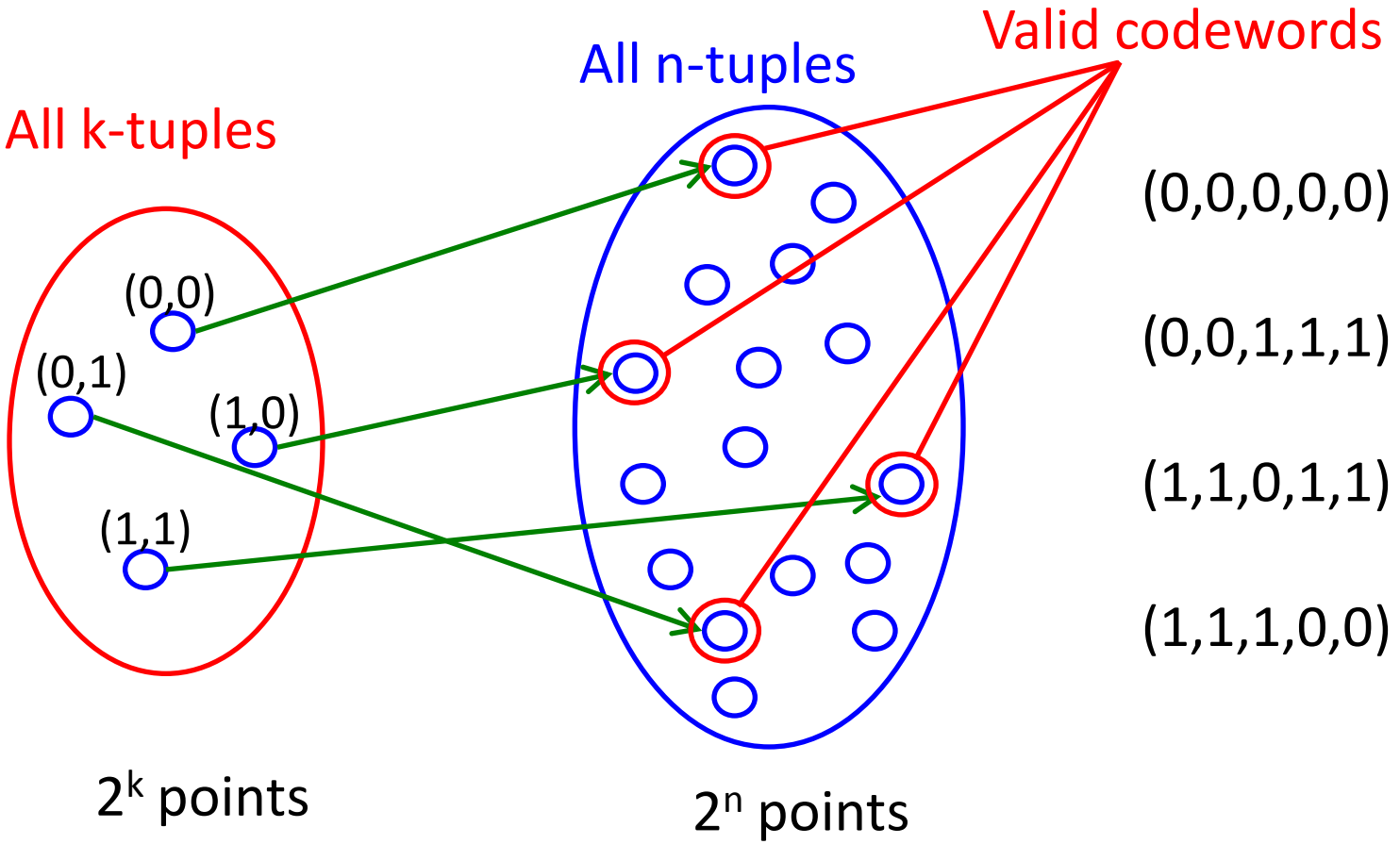
$$\begin{bmatrix} & & 1 \\ & & 1 \\ & & 1 \\ & I_7 & 1 \\ & & 1 \\ & & 1 \\ & & 1 \end{bmatrix}$$

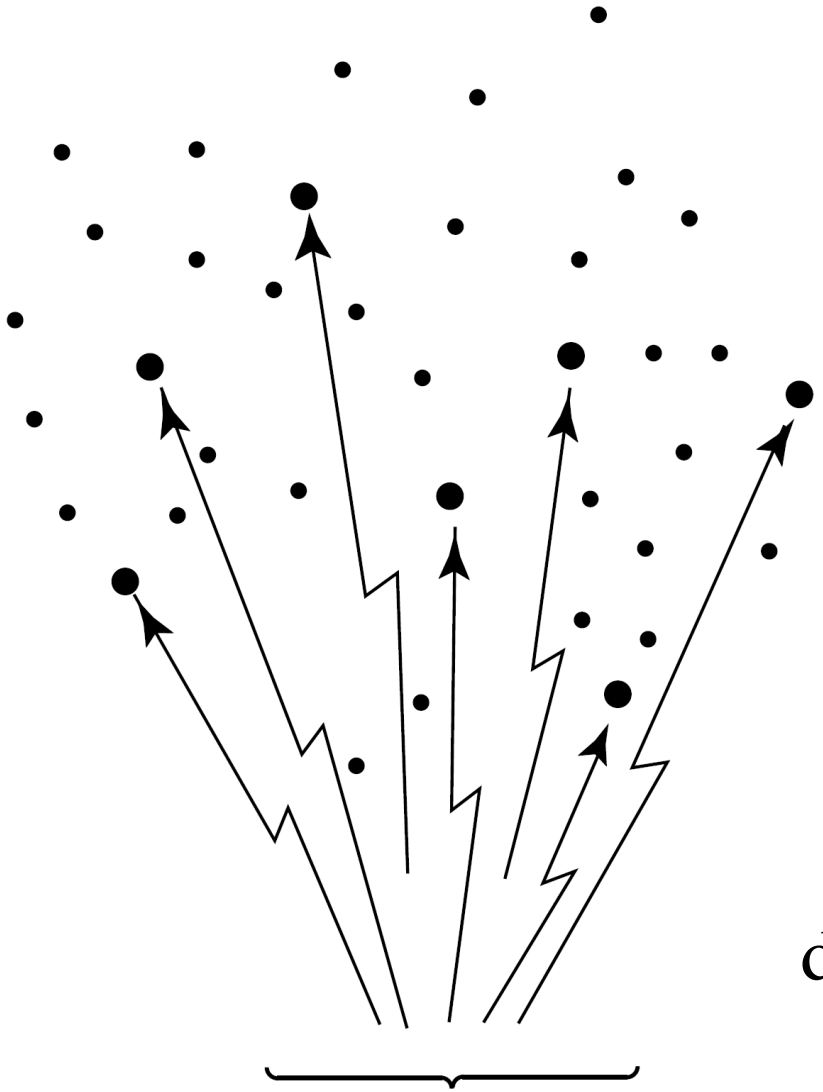
E = 1000101 c = 10001011

G = 1000111 c = 10001110

Linear Codes as Vector Spaces (Cont.)

$$(m_0, m_1, \ldots, m_{k-1}) \rightarrow (c_0, c_1, \ldots, c_{n-1})$$





2^n n-tuples constitute the entire space V_n

2^k n-tuples constitute the subspace of codewords

$$\dim(C) = k \qquad \dim(V) = n$$

Which of the following binary codes are not linear?

C_1 = {00, 01, 10, 11}

C_2 = {000, 011, 101, 110}

C_3 = {00000, 11100, 00111, 11011}

C_4 = {101, 111, 011}

C_5 = {000, 001, 010, 011}

C_6 = {0000, 1001, 0110, 1110}

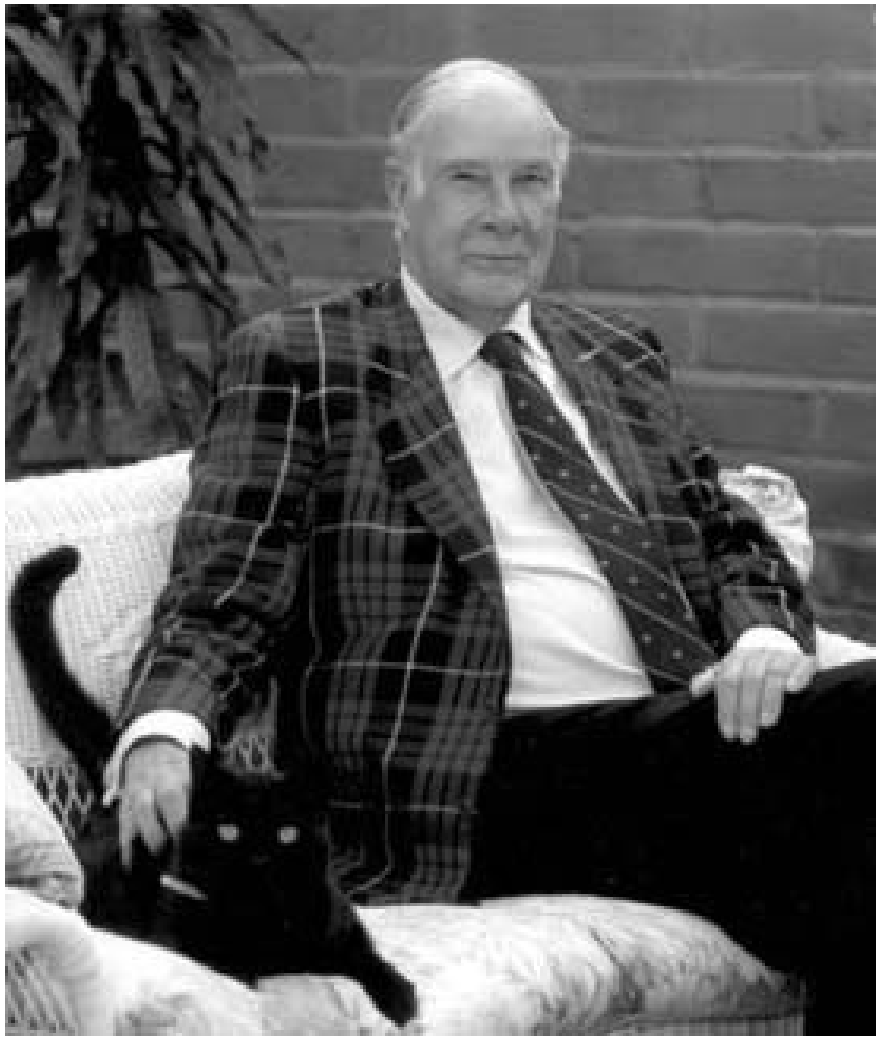
(5,2) Binary Linear Code

- *k* x *n* Generator Matrix

$$G = \begin{matrix} \\ \\ 2 \end{matrix} \begin{matrix} 5 \\ \left[\begin{array}{c} 00111 \\ 11100 \end{array}\right] \end{matrix}$$

	$\bar{m}$	$\bar{c}$
c_0	00	00000
c_1	01	11100
c_2	10	00111
c_3	11	11011

Richard W. Hamming (1915-1998)



Hamming at Bell Labs

- The development of error correcting codes began in 1947 at Bell Laboratories
- Hamming had access to a mechanical relay computer on some weekends
- The computer employed an error detecting code, but with no operator on duty during weekends, the computer simply stopped or went on to the next problem when an error occurred

“Two weekends in a row I came in and found that all my stuff had been dumped and nothing was done.” And so I said, “Damn it, if the machine can detect an error, why can't it locate the position of the error and correct it?”

Hamming Weight and Distance

- The concept of "closeness" of two codewords is formalized through the Hamming distance.
- Let x and y be any two codewords in C

  $x = 00111 \quad y = 11100$
- The Hamming weight of a codeword is defined as the number of nonzero elements in the codeword

  $w(x) = w(00111) = 3 \quad w(y) = w(11100) = 3$
- The Hamming distance between two codewords is defined as the number of places in which they differ

  $d(x,y) = d(00111,11100) = 4$

Hamming Distances for Linear Codes

- For a binary linear code, the mod 2 sum of any two codewords is another codeword

  $x + y = z$ 00111 + 11100 = 11011

- Thus

  $d(x,y) = w(x+y) = w(z) = w(11011) = 4$

- Since we are concerned with the error correcting capability of a code *C*:

  What is the most important criteria for a linear (*n,k*) code?

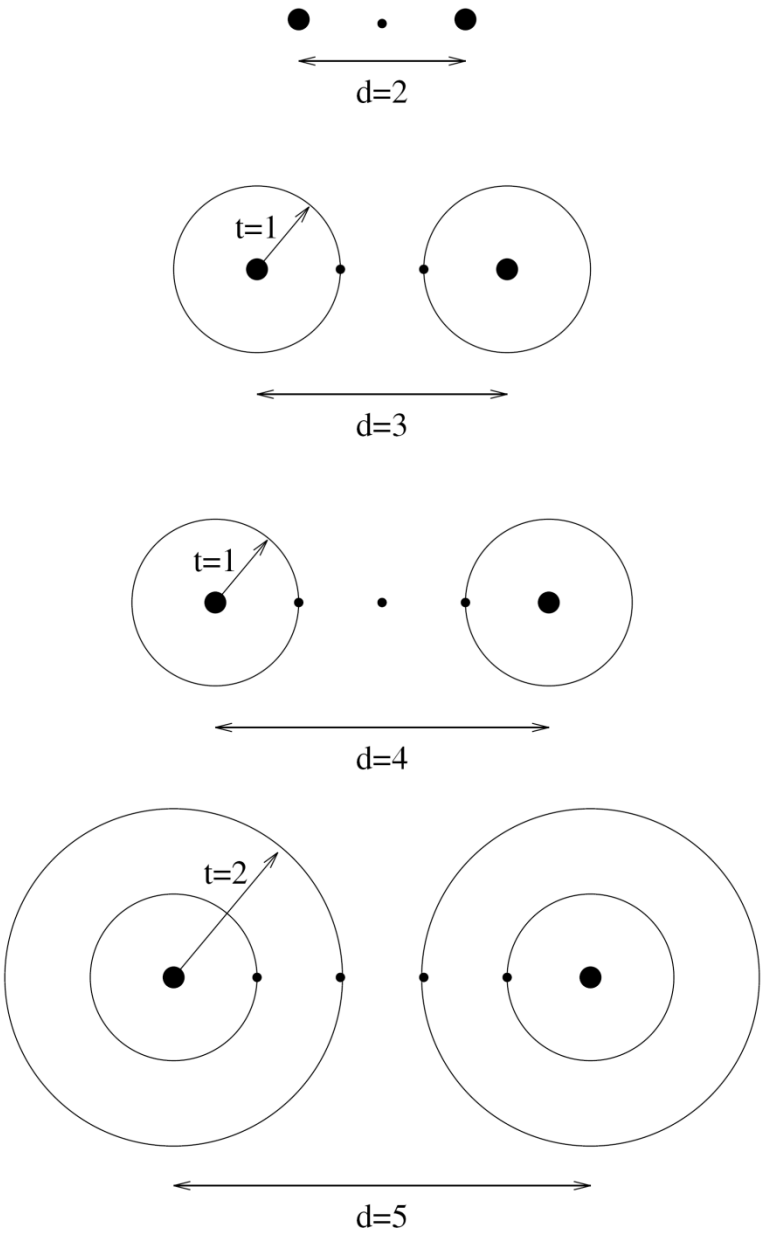
(5,2,3) Code

- *k* x *n* Generator matrix

$$G = \begin{matrix} \\ 2 \end{matrix} \begin{matrix} \overset{5}{\left[\begin{array}{c} 00111 \\ 11100 \end{array}\right]} \end{matrix}$$

	$\bar{m}$	$\bar{c}$	w
c_0	00	00000	0
c_1	01	11100	3
c_2	10	00111	3
c_3	11	11011	4

- d(*C*) = 3


d=2
t=1
d=3
t=1
d=4
t=2
d=5

Minimum Hamming Distance

- An important parameter of a code C is its minimum distance

  $$d(C) = \min \{d(x, y) \mid x,y \in C, x \neq y\}$$

  (also called d_{min})

- For a linear code

  $$d(C) = \min \{w(x) \mid x \in C, x \neq 0\}$$

- A code C can detect up to v errors if $d(C) \geq v + 1$
- A code C can correct up to t errors if $d(C) \geq 2t + 1$

Notation and Examples

An (n,k,d) code C is a linear code such that

- n - is the length of the codewords
- k - is the number of data symbols in a codeword
- d - is the minimum distance of C

$$d = d(C)$$

Examples:

C_1 = {00, 01, 10, 11} is a (2,2,1) code.

C_2 = {000, 011, 101, 110} is a (3,2,2) code.

C_3 = {00000, 11100, 00111, 11011} is a (5,2,3) code.

A good (n,k,d) code has small n, large k and large d.

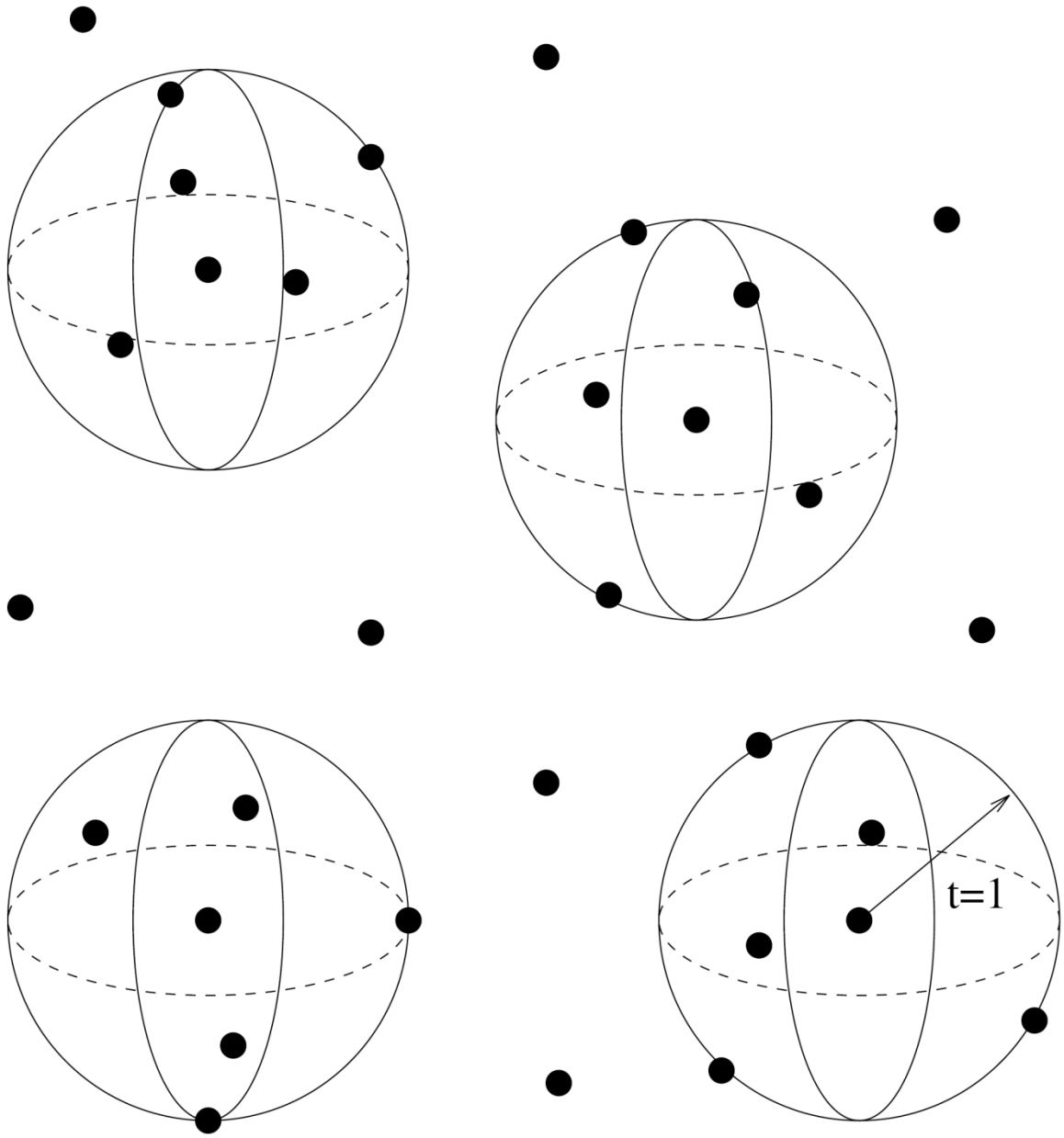
(5,2,3) Code

- *k* x *n* Generator matrix

$$G = \begin{matrix} \\ 2 \end{matrix} \overset{5}{\begin{bmatrix} 00111 \\ 11100 \end{bmatrix}}$$

	$\bar{m}$	$\bar{c}$	w
c_0	00	00000	0
c_1	01	11100	3
c_2	10	00111	3
c_3	11	11011	4

- $d = 3,\ t = \left\lfloor \frac{3-1}{2} \right\rfloor = 1$


t=1

Advantages of Linear Block Codes

1. The minimum distance d(*C*) is easy to compute if *C* is a linear code.
2. Linear codes have simple specifications.

- To specify a non-linear code usually all codewords have to be listed.
- To specify a linear (*n*,*k*) code it is enough to list *k* linearly independent codewords.

Definition A $k \times n$ matrix whose rows form a basis for a linear (*n*,*k*) code (subspace) *C* is said to be a generator matrix for *C*.

Example A generator matrix for the code

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

and for the code

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

3. There are simple encoding/decoding procedures for linear codes.

Important Linear Block Codes

There are many classes of practical linear block codes:

- Hamming codes
- Cyclic codes
- Reed-Solomon codes
- BCH codes
- LDPC codes
- Turbo codes
- …

How to Create a Linear Code

If S is a set of vectors of a vector space, then let $\langle S \rangle$ be the set of all linear combinations of vectors from S.

$\langle S \rangle$ is a linear space that consists of the following words:

- the all-zero word
- all words in S
- all sums of two or more words in S

Example:

S = {0100, 0011, 1100}

$\langle S \rangle$ = {0000, 0100, 0011, 1100, 0111, 1011, 1000, 1111}