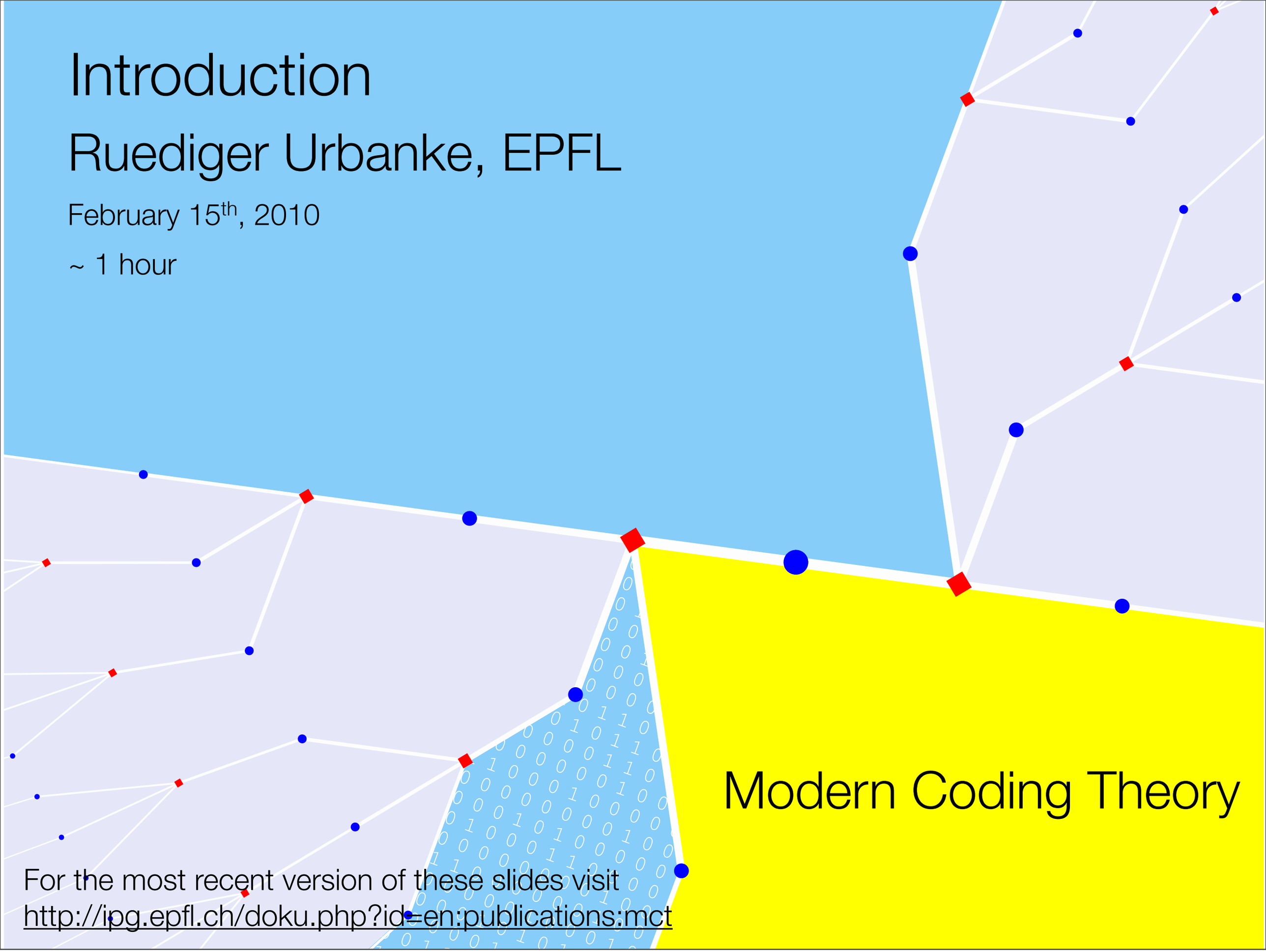
Introduction

Ruediger Urbanke, EPFL

February 15th, 2010

~ 1 hour



For the most recent version of these slides visit
http://ipg.epfl.ch/doku.php?id=en:publications:mct

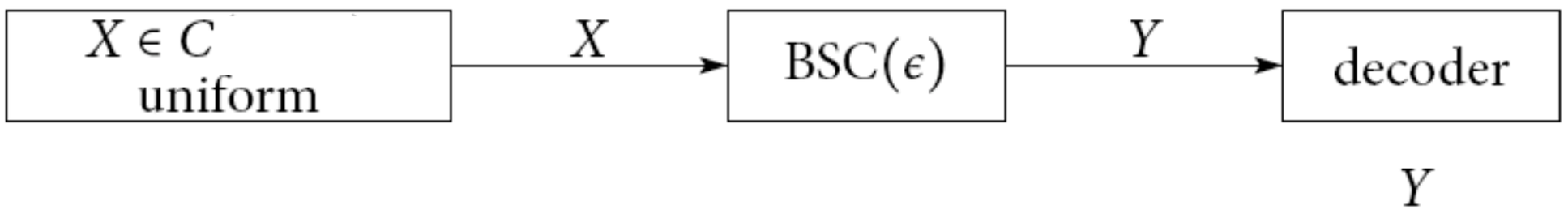
Coding

Coding

“information source” “channel”

Coding

“information source” “channel”



C={000, 010, 101, 111} block code

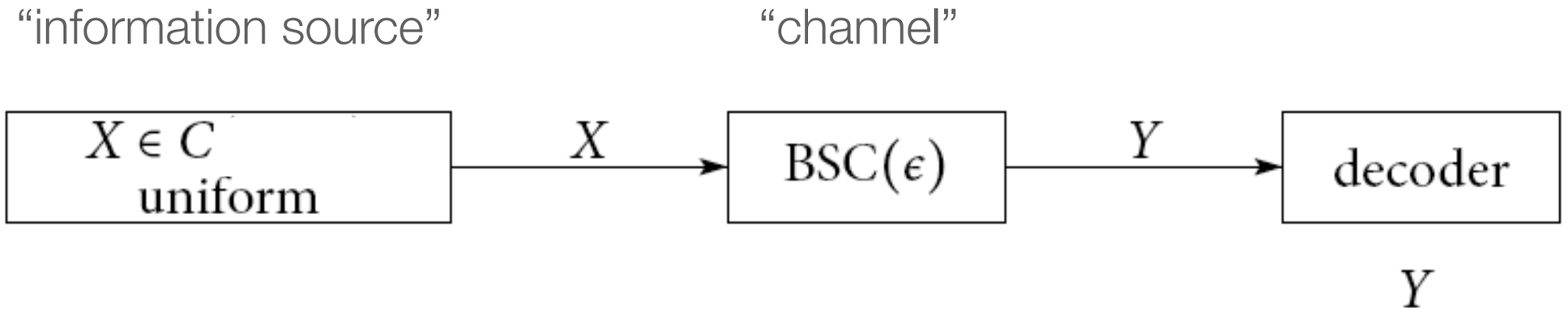
Coding

“information source” “channel”

C={000, 010, 101, 111} block code

codeword

Coding



C={000, 010, 101, 111} block code

codeword

n ... blocklength

How many codewords do we need?

How many codewords do we need?

Assume we want to send k bits of information.

How many codewords do we need?

Assume we want to send k bits of information.

Then we need to be able to distinguish between 2^k messages.

How many codewords do we need?

Assume we want to send k bits of information.

Then we need to be able to distinguish between 2^k messages.

Even for very small values of k, e.g., k=1000, 2^k is gigantic.

How many codewords do we need?

Assume we want to send k bits of information.

Then we need to be able to distinguish between 2^k messages.

Even for very small values of k, e.g., k=1000, 2^k is gigantic.

We need a systematic way of generating and storing codewords that can handle such large codes.

Linear Codes

generator matrix

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

Linear Codes

generator matrix

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

n=7

Linear Codes

generator matrix

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

k=4

n=7

Linear Codes

generator matrix

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

k=4

2^k=16 codewords

n=7

Linear Codes

generator matrix

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

k=4

2^k=16 codewords

n=7

x=u G

code word

information

Linear Codes

generator matrix

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

k=4

2^k=16 codewords

n=7

x=u G

u=(0110)

code word

information

Linear Codes

generator matrix

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

k=4

2^k=16 codewords

n=7

x=u G

code word

information

u=(0110)

x=(0110110)

Linear Codes

generator matrix

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

k=4

2^k=16 codewords

n=7

x=u G

code word

information

u=(0110)

x=(0110110)

Note: There are many matrices G that all define the same code. For an (n, k) code, how many such “equivalent” matrices are there?

Linear Codes

Linear Codes

$$C(G) = \left\{ x \in \mathbb{F}^n : x = uG, u \in \mathbb{F}^k \right\}$$

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

generator matrix

Linear Codes

$$C(G) = \left\{ x \in \mathbb{F}^n : x = uG, u \in \mathbb{F}^k \right\}$$

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

generator matrix

code = row space spanned by G

Linear Codes

$$C(G) = \left\{ x \in \mathbb{F}^n : x = uG, u \in \mathbb{F}^k \right\} \qquad C = \left\{ x \in \mathbb{F}^n : x = uG, u \in \mathbb{F}^k \right\} = \left\{ x \in \mathbb{F}^n : Hx^T = 0^T \right\}$$

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \qquad H = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

generator matrix

parity-check matrix

code = row space spanned by G

Linear Codes

$$C(G) = \left\{ x \in \mathbb{F}^n : x = uG, u \in \mathbb{F}^k \right\} \qquad C = \left\{ x \in \mathbb{F}^n : x = uG, u \in \mathbb{F}^k \right\} = \left\{ x \in \mathbb{F}^n : Hx^T = 0^T \right\}$$

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \qquad H = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

generator matrix

parity-check matrix

code = row space spanned by G

code = kernel of H

Linear Codes

$$C(G) = \{x \in \mathbb{F}^n : x = uG, u \in \mathbb{F}^k\} \qquad C = \{x \in \mathbb{F}^n : x = uG, u \in \mathbb{F}^k\} = \{x \in \mathbb{F}^n : Hx^T = 0^T\}$$

G=(I_{kxk}, P) H=($-P^T$, $I_{(n-k)x(n-k)}$)

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \qquad H = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}$$

generator matrix parity-check matrix

code = row space spanned by G code = kernel of H

Linear Codes

Linear Codes

Consider a (n, k) code C spanned by a rank-k matrix G, where G has the form $G=(I_{kxk}, P)$.

Linear Codes

Consider a (n, k) code C spanned by a rank-k matrix G,
where G has the form $G=(I_{kxk}, P)$.

Define the associated parity-check matrix H by $H=(-P^T, I_{(n-k)x(n-k)})$.

Linear Codes

Consider a (n, k) code C spanned by a rank-k matrix G,
where G has the form $G=(I_{kxk}, P)$.

Define the associated parity-check matrix H by $H=(-P^T, I_{(n-k)x(n-k)})$.

Let x be a codeword, i.e., x=u G, where u is an information word of length k.

Linear Codes

Consider a (n, k) code C spanned by a rank-k matrix G,
where G has the form $G=(I_{kxk}, P)$.

Define the associated parity-check matrix H by $H=(-P^T, I_{(n-k)x(n-k)})$.

Let x be a codeword, i.e., x=u G, where u is an information word of length k.

Then $H x^T=0$, since $H x^T= H G^T u^T$

$$=((-P^T, I_{(n-k)x(n-k)})(I_{kxk}, P^T))\ u^T$$
$$= (-P^T+ P)\ u^T$$
$$= 0$$

Linear Codes

Consider a (n, k) code C spanned by a rank-k matrix G,
where G has the form $G=(I_{kxk}, P)$.

Define the associated parity-check matrix H by $H=(-P^T, I_{(n-k)x(n-k)})$.

Let x be a codeword, i.e., x=u G, where u is an information word of length k.

Then $H x^T=0$, since $H x^T= H G^T u^T$

$$
\begin{aligned}
&=((-P^T, I_{(n-k)x(n-k)})(I_{kxk}, P^T))\ u^T \\
&= (-P^T+ P)\ u^T \\
&= 0
\end{aligned}
$$

So all 2^k codewords spanned by G are in the kernel of H. But the kernel of H only contains 2^k words since H has rank (n-k).

Linear Codes

Consider a (n, k) code C spanned by a rank-k matrix G,
where G has the form $G=(I_{kxk}, P)$.

Define the associated parity-check matrix H by $H=(-P^T, I_{(n-k)x(n-k)})$.

Let x be a codeword, i.e., x=u G, where u is an information word of length k.

Then $H x^T=0$, since $H x^T= H G^T u^T$

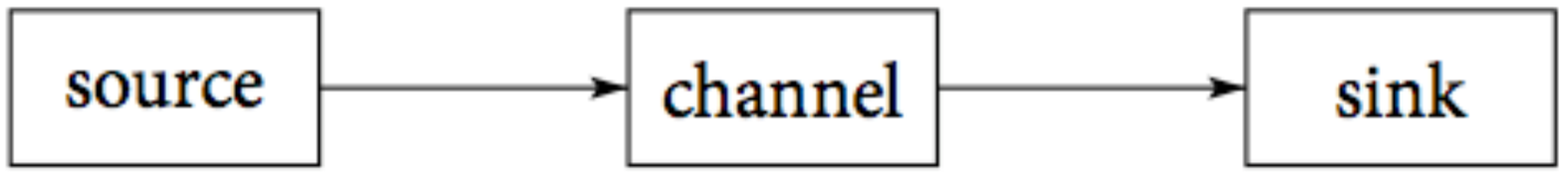
$$=((-P^T, I_{(n-k)x(n-k)})(I_{kxk}, P^T))\ u^T$$
$$= (-P^T+ P)\ u^T$$
$$= 0$$

So all 2^k codewords spanned by G are in the kernel of H. But the kernel of H only contains 2^k words since H has rank (n-k).

Note: There are many parity-check matrices H that all define the same code. For an (n, k) code, how many such “equivalent” matrices are there?

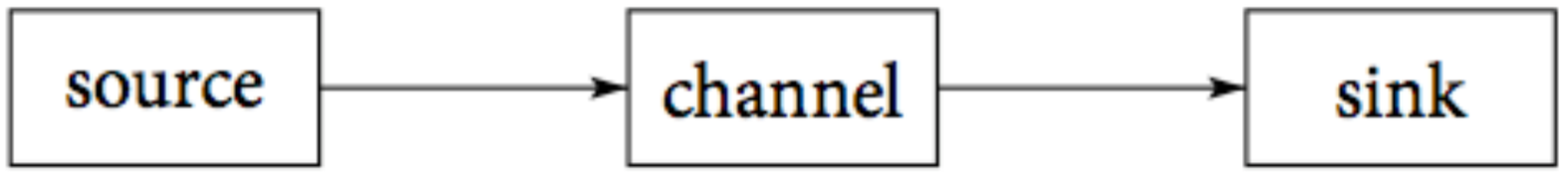
Standard Channel Model

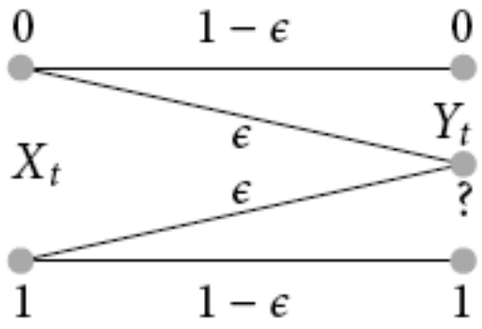
Shannon '48



Standard Channel Model

Shannon ’48

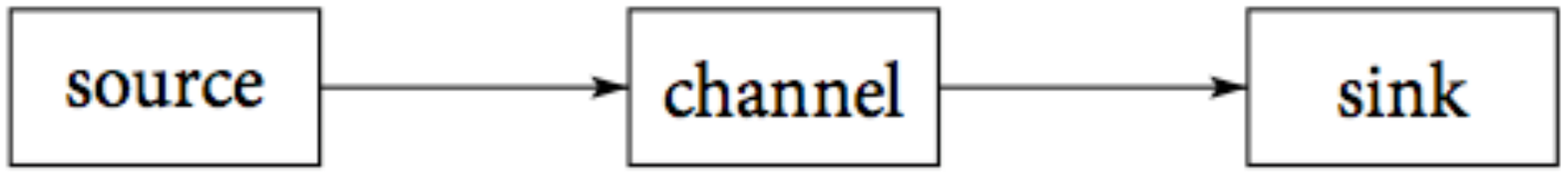


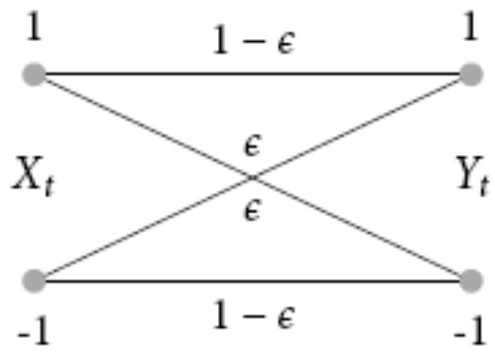


binary erasures channel
capacity: R≤1-**ε**

Standard Channel Model

Shannon ’48

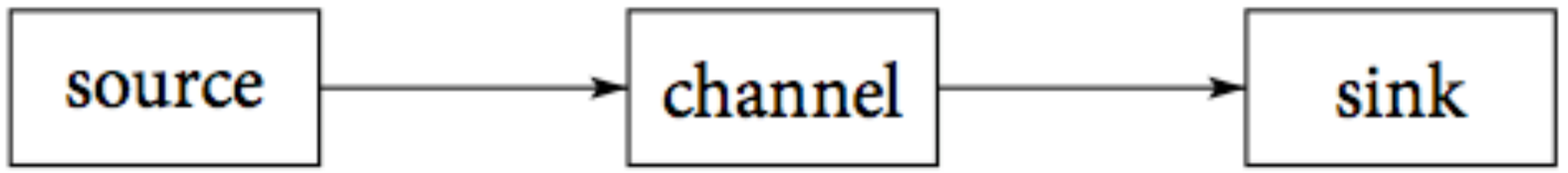


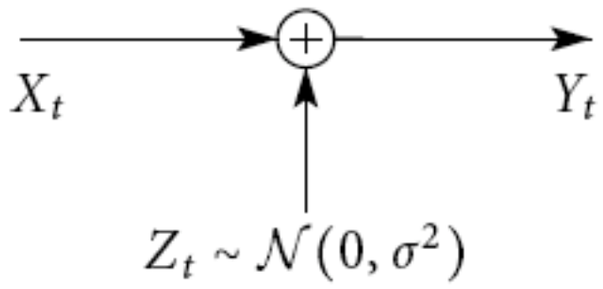


binary symmetric channel
capacity: R≤1-h(**ε**)

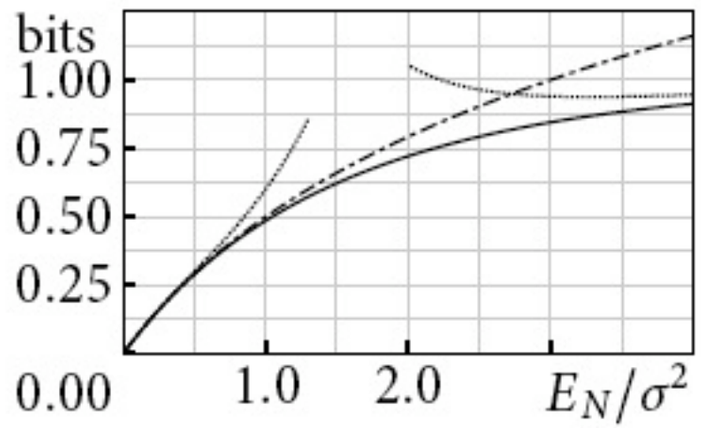
Standard Channel Model

Shannon ’48





BAWGN channel



$$1 + \frac{1}{\ln(2)}\left(\left(\frac{2}{\sigma^2} - 1\right)Q\left(\frac{1}{\sigma}\right) - \sqrt{\frac{2}{\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}} + \sum_{i=1}^{\infty}\frac{(-1)^i}{i(i+1)}e^{\frac{2i(i+1)}{\sigma^2}}Q\left(\frac{1+2i}{\sigma}\right)\right)$$

Memoryless Symmetric Channels

Memoryless Symmetric Channels

Definition 4.3 (Memoryless Channels). A channel, characterized by its *transition probability* $p_{Y|X}(y|x)$, is said to be *memoryless* if

$$p_{Y|X}(y|x) = \prod_t p_{Y_t|X_t}(y_t|x_t).$$

Memoryless Symmetric Channels

Definition 4.3 (Memoryless Channels). A channel, characterized by its *transition probability* $p_{Y|X}(y|x)$, is said to be *memoryless* if

$$p_{Y|X}(y|x) = \prod_t p_{Y_t|X_t}(y_t|x_t).$$

Definition 4.8 (Channel Symmetry). Assume $\mathcal{Y} \subset \bar{\mathbb{R}}$. We say that a binary memoryless channel is *symmetric* (more precisely, output-symmetric) if

$$p_{Y|X}(y|1) = p_{Y|X}(-y|-1).$$

Memoryless Symmetric Channels

DEFINITION 4.3 (MEMORYLESS CHANNELS). A channel, characterized by its *transition probability* $p_{Y|X}(y|x)$, is said to be *memoryless* if

$$p_{Y|X}(y|x) = \prod_t p_{Y_t|X_t}(y_t|x_t).$$

DEFINITION 4.8 (CHANNEL SYMMETRY). Assume $\mathcal{Y} \subset \bar{\mathbb{R}}$. We say that a binary memoryless channel is *symmetric* (more precisely, output-symmetric) if

$$p_{Y|X}(y|1) = p_{Y|X}(-y|-1).$$

EXAMPLE 4.10 (SYMMETRY OF STANDARD CHANNELS). Our three standard channels, the BEC, the BSC, as well as the BAWGNC, are all symmetric. ◇

MAP Decoding

MAP Decoding

MAP Decoding

$$\begin{aligned}\hat{x}^{\mathrm{MAP}}(y) &= \operatorname{argmax}_{x \in C} p_{X \mid Y}(x \mid y) \\ &= \operatorname{argmax}_{x \in C} p_{Y \mid X}(y \mid x) \frac{p_X(x)}{p_Y(y)} \\ &= \operatorname{argmax}_{x \in C} p_{Y \mid X}(y \mid x) p_X(x)\end{aligned}$$

MAP Decoding

$$\begin{aligned}\hat{x}^{\mathrm{MAP}}(y) &= \operatorname{argmax}_{x \in C} p_{X|Y}(x \mid y) \\ &= \operatorname{argmax}_{x \in C} p_{Y|X}(y \mid x) \frac{p_X(x)}{p_Y(y)} \\ &= \operatorname{argmax}_{x \in C} p_{Y|X}(y \mid x) p_X(x)\end{aligned}$$

THEOREM 1.27. The *ML Decision Problem* for the BSC is NP-complete.

MAP Decoding

$$\hat{x}^{\mathrm{MAP}}(y) = \operatorname{argmax}_{x \in C} p_{X|Y}(x|y)$$
$$= \operatorname{argmax}_{x \in C} p_{Y|X}(y|x) \frac{p_X(x)}{p_Y(y)}$$
$$= \operatorname{argmax}_{x \in C} p_{Y|X}(y|x) p_X(x)$$

Note: The MAP decoder minimizes the probability of error. This is why we would like to implement it. (In order to achieve capacity it is in fact not necessary to do MAP decoding.)

THEOREM 1.27. The *ML Decision Problem* for the BSC is NP-complete.

The most important parameters

The most important parameters

$(r, \mathrm{P}, \chi_E, \chi_D, n)$

rate, error probability,
encoding complexity,
decoding complexity,
blocklength

The most important parameters

$(r, \mathrm{P}, \chi_E, \chi_D, n)$

rate, error probability,
encoding complexity,
decoding complexity,
blocklength

$\delta = 1 - r/\mathrm{C}$

gap to capacity

$r = (1 - \delta)\mathrm{C}$

The most important parameters

$(r, \mathrm{P}, \chi_E, \chi_D, n)$

rate, error probability,
encoding complexity,
decoding complexity,
blocklength

$\delta = 1 - r/\mathrm{C}$

gap to capacity

$r = (1 - \delta)\mathrm{C}$

$(\delta, \mathrm{P}, \chi_E, \chi_D, n)$

Coding history according to Jim Massey

(Coding Theory: The Phoenix of Communications, 2001)

Coding history according to Jim Massey

(Coding Theory: The Phoenix of Communications, 2001)

Golay code (1948), Hamming codes (1949)
BCH codes (Hocquenghem, 1959, Bose and Ray-Chaudhuri, 1960)
RS codes (Reed and Solomon,1960)
Convolutional codes (Elias, 1960), sequential decoding (Wozencraft, 1961)

Coding history according to Jim Massey

(Coding Theory: The Phoenix of Communications, 2001)

Golay code (1948), Hamming codes (1949)
BCH codes (Hocquenghem, 1959, Bose and Ray-Chaudhuri, 1960)
RS codes (Reed and Solomon,1960)
Convolutional codes (Elias, 1960), sequential decoding (Wozencraft, 1961)

LDPC codes (Gallager, 1960)
Threshold decoding (Massey, 1963) } Codex
Concatenated Codes (Forney, 1966)
Pioneer 9 probe (convolutional codes and sequ. decoding, 1969)

Coding history according to Jim Massey

(Coding Theory: The Phoenix of Communications, 2001)

Golay code (1948), Hamming codes (1949)
BCH codes (Hocquenghem, 1959, Bose and Ray-Chaudhuri, 1960)
RS codes (Reed and Solomon,1960)
Convolutional codes (Elias, 1960), sequential decoding (Wozencraft, 1961)

LDPC codes (Gallager, 1960)
Threshold decoding (Massey, 1963) } Codex
Concatenated Codes (Forney, 1966)
Pioneer 9 probe (convolutional codes and sequ. decoding, 1969)

Viterbi Algorithm (Viterbi, 1969)
Trellis (Forney, 1969)

Coding history according to Jim Massey

(Coding Theory: The Phoenix of Communications, 2001)

Golay code (1948), Hamming codes (1949)
BCH codes (Hocquenghem, 1959, Bose and Ray-Chaudhuri, 1960)
RS codes (Reed and Solomon,1960)
Convolutional codes (Elias, 1960), sequential decoding (Wozencraft, 1961)

LDPC codes (Gallager, 1960)
Threshold decoding (Massey, 1963) } Codex
Concatenated Codes (Forney, 1966)
Pioneer 9 probe (convolutional codes and sequ. decoding, 1969)

Viterbi Algorithm (Viterbi, 1969)
Trellis (Forney, 1969)

Trellis Coded Modulation (Ungerboeck, 1976, 1982)

Coding history according to Jim Massey

(Coding Theory: The Phoenix of Communications, 2001)

Golay code (1948), Hamming codes (1949)
BCH codes (Hocquenghem, 1959, Bose and Ray-Chaudhuri, 1960)
RS codes (Reed and Solomon,1960)
Convolutional codes (Elias, 1960), sequential decoding (Wozencraft, 1961)

LDPC codes (Gallager, 1960)
Threshold decoding (Massey, 1963) } Codex
Concatenated Codes (Forney, 1966)
Pioneer 9 probe (convolutional codes and sequ. decoding, 1969)

Viterbi Algorithm (Viterbi, 1969)
Trellis (Forney, 1969)

Trellis Coded Modulation (Ungerboeck, 1976, 1982)

Turbo Codes (Berrou, Glavieux, Thitimashima, 1993)

... and the story continues

... and the story continues

rate-less codes
(Byers, Luby, Mitzenmacher, and Rege, 1998, Shokrollahi, 2004)

... and the story continues

rate-less codes
(Byers, Luby, Mitzenmacher, and Rege, 1998, Shokrollahi, 2004)

network coding
(Ahlswede, Cai, Li, and Yeung, 2000)

... and the story continues

rate-less codes
(Byers, Luby, Mitzenmacher, and Rege, 1998, Shokrollahi, 2004)

network coding
(Ahlswede, Cai, Li, and Yeung, 2000)

polar codes
(Arikan, 2008)