Binary symmetric channel

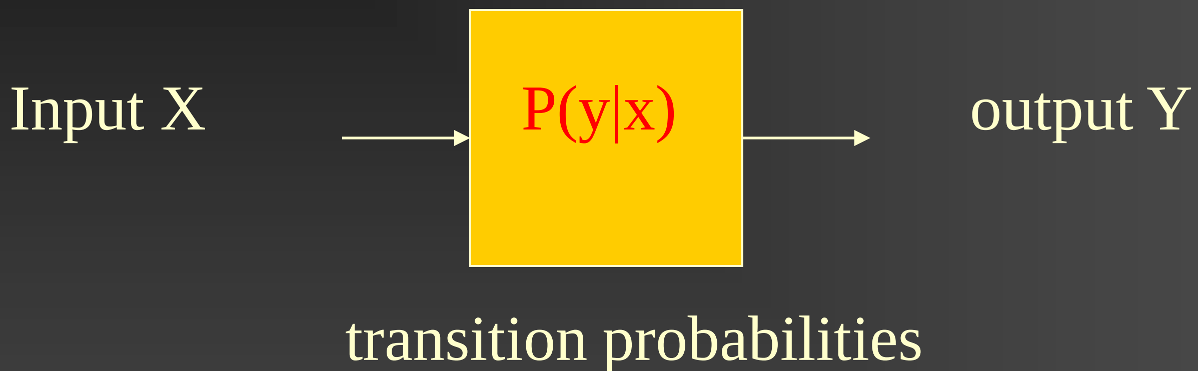
content

- Introduction
  - Entropy and some related properties
- Source coding
- Channel coding
- Multi-user models
- Constraint sequence
- Applications to cryptography

This lecture

- Some models
- Channel capacity
  - converse

some channel models



memoryless:

- output at time i depends only on input at time i
- input and output alphabet finite

channel capacity:

I(X;Y) = H(X) - H(X|Y) = H(Y) –H(Y|X) (Shannon 1948)

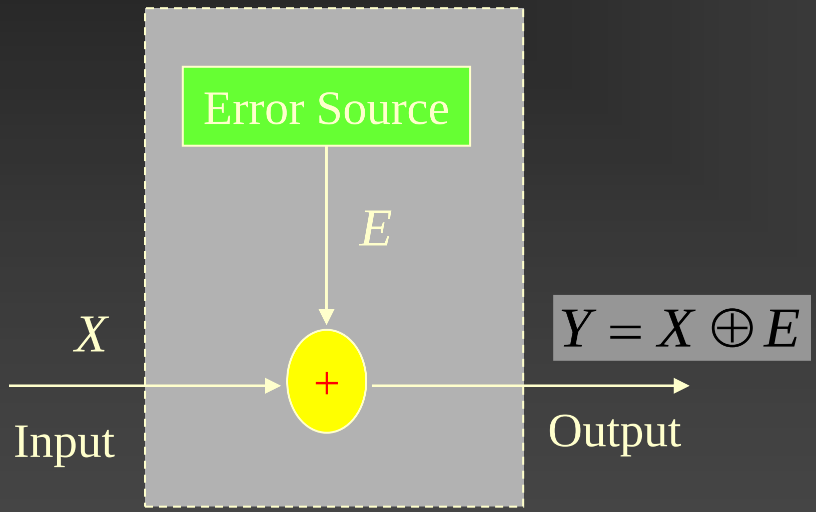
$$\max_{P(x)} I(X;Y) = \text{capacity}$$

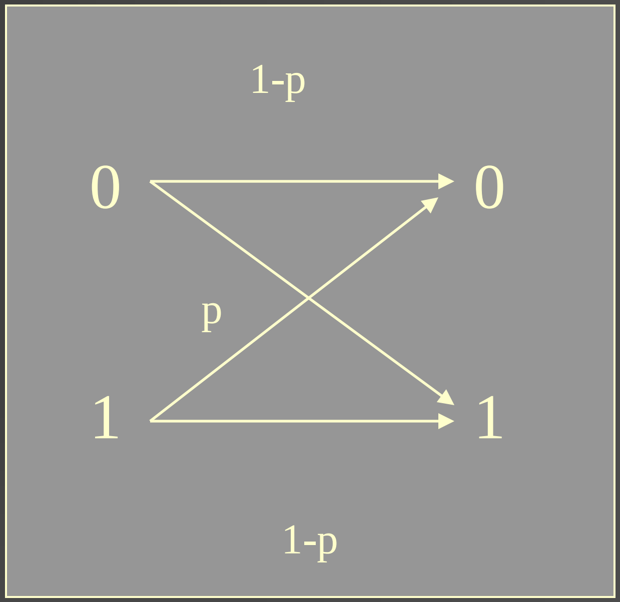
notes:

capacity depends on input probabilities

because the transition probabilites are fixed

channel model: binary symmetric channel





E is the binary error sequence s.t. P(1) = 1-P(0) = p

X is the binary information sequence

Y is the binary output sequence

burst error model

Random error channel; outputs independent

Error Source → $P(0) = 1 - P(1);$

Burst error channel; outputs dependent

Error Source →

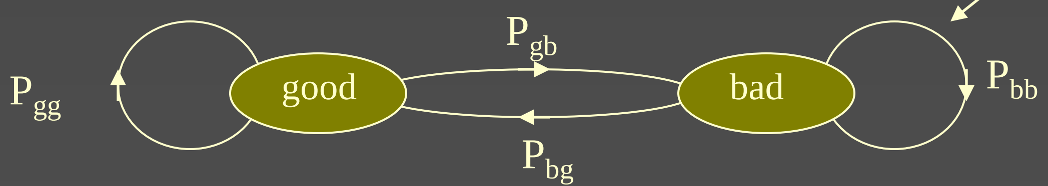
$P(0 \mid state = bad) = P(1 \mid state = bad) = 1/2;$

$P(0 \mid state = good) = 1 - P(1 \mid state = good) = 0.999$

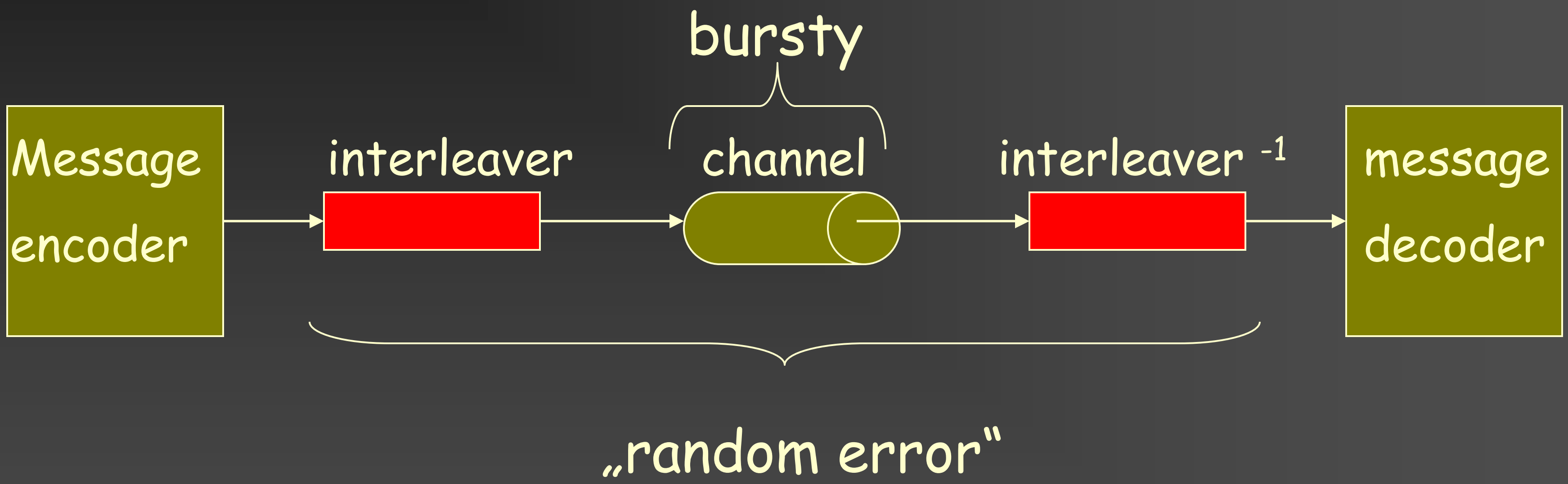
↑

State info: good or bad

transition probability



Interleaving:



Note: interleaving brings encoding and decoding delay

Homework: compare the block and convolutional interleaving w.r.t. delay

Interleaving: block

Channel models are difficult to derive:

- burst definition ?
- random and burst errors ?

for practical reasons: convert burst into random error

read in row wise

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

transmit

column wise

De-Interleaving: block

read in column wise

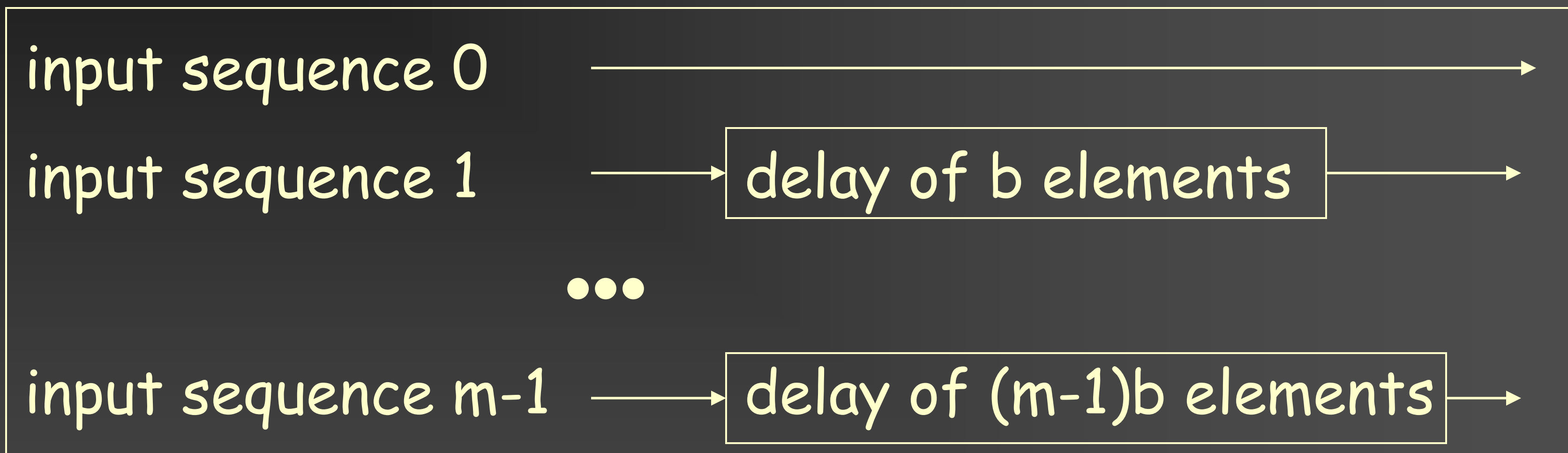
this row contains 1 error

1	0	1		e		1
0	1	e		e		0
0	0	e		1		0
1	0	e		1		1
1	1	e		0		1

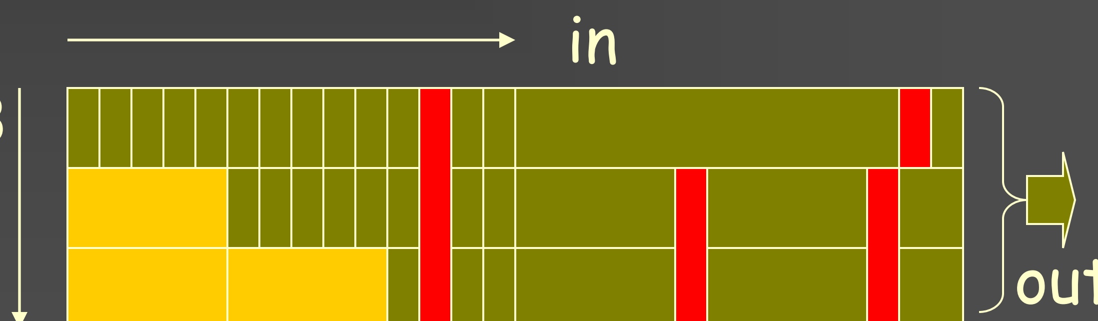
read out

row wise

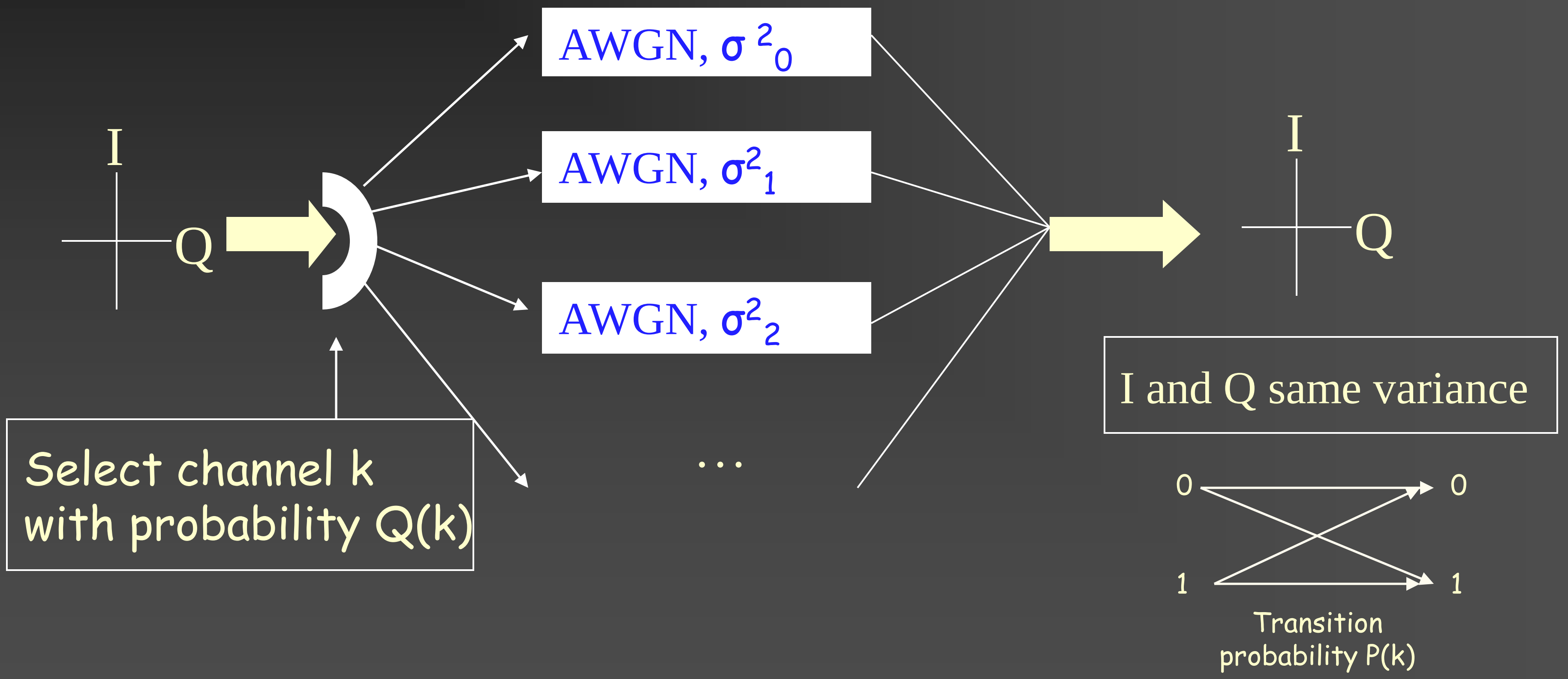
Interleaving: convolutional



Example: b = 5, m = 3



Class A Middleton channel model



Example: Middleton's class A

$\Pr\{\ \sigma = \sigma(k)\ \} = Q(k),\ \ k = 0, 1, \cdots$

$$\sigma(k) := \left(\frac{k\sigma_I^2 / A + \sigma_G^2}{\sigma_I^2 + \sigma_G^2}\right)^{1/2}$$

$$Q(k) := \frac{e^{-A} A^k}{k!}$$

A is the impulsive index

σ_I^2 and σ_G^2 are the impulsive and Gaussian noise power

Example of parameters

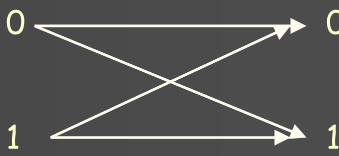
- Middleton's class A= 1; E = **σ = 1; $\sigma_I / \sigma_G = 10^{-1.5}$**

k	Q(k)	p(k) (= transition probability)
0	0.36	0.00
1	0.37	0.16
2	0.19	0.24
3	0.06	0.28
4	0.02	0.31

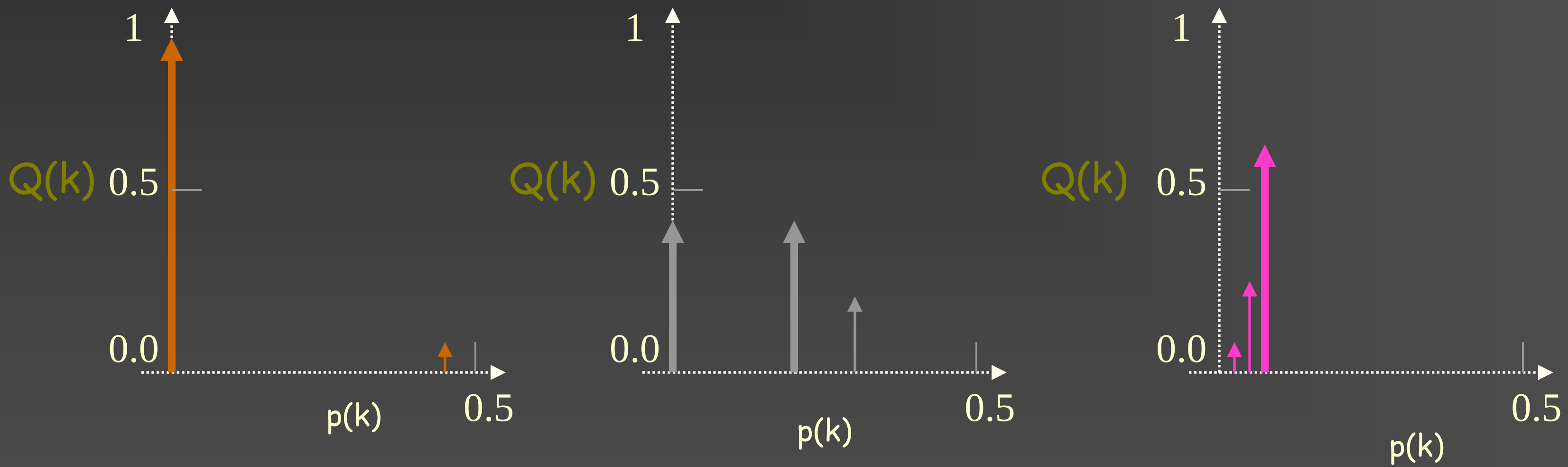
Average p = 0.124; Capacity (BSC) = 0.457

Example of parameters

Middleton's class A: $E = 1;\ \sigma = 1;\ \sigma_I / \sigma_G = 10^{-3}$



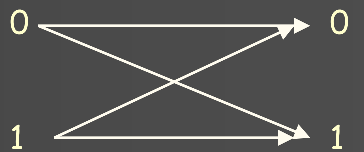
Transition probability P(k)



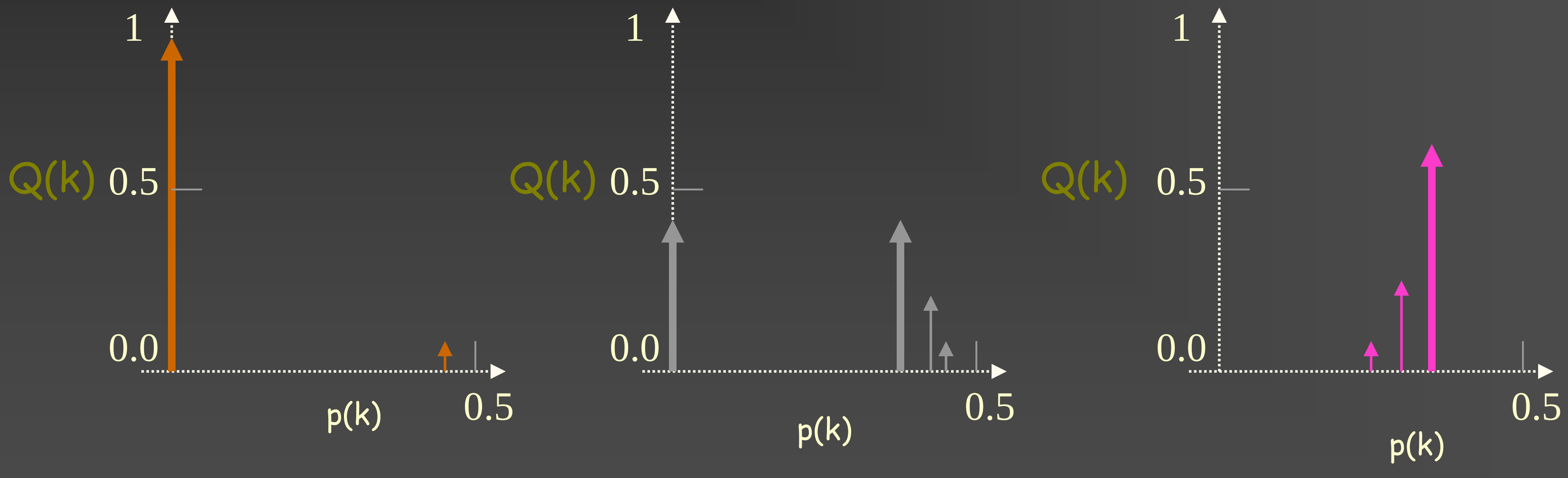
A = 0.1 A = 1 A = 10

Example of parameters

Middleton's class A: $E = 0.01$; $\sigma = 1$; $\sigma_I / \sigma_G = 10^{-3}$

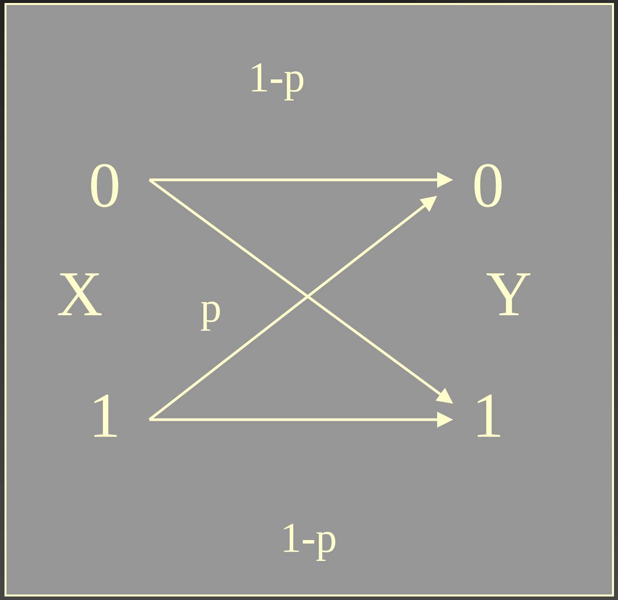


Transition probability P(k)



A = 0.1 A = 1 A = 10

channel capacity: the BSC



$I(X;Y) = H(Y) - H(Y|X)$

the maximum of $H(Y) = 1$

since Y is binary

$H(Y|X) = h(p)$

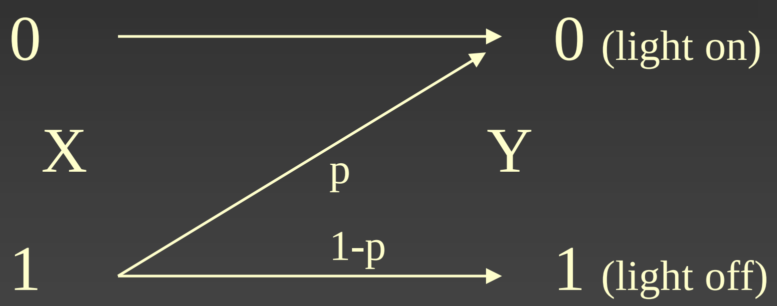
$= P(X=0)h(p) + P(X=1)h(p)$

Conclusion: the capacity for the BSC $C_{BSC} = 1 - h(p)$

Homework: draw C_{BSC}, what happens for $p > \frac{1}{2}$

channel capacity: the Z-channel

Application in optical communications



$P(X=0) = P_0$

$H(Y) = h(P_0 + p(1 - P_0))$

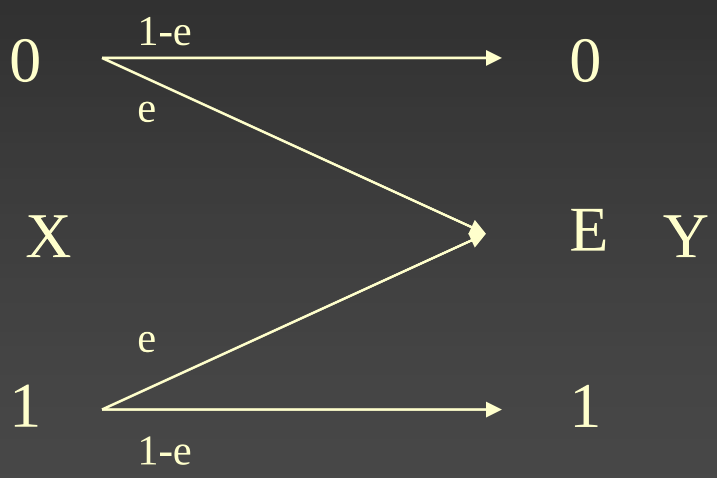
$H(Y|X) = (1 - P_0)\, h(p)$

For capacity,

maximize $I(X;Y)$ over P_0

channel capacity: the erasure channel

Application: cdma detection



$P(X=0) = P_0$

$I(X;Y) = H(X) - H(X|Y)$

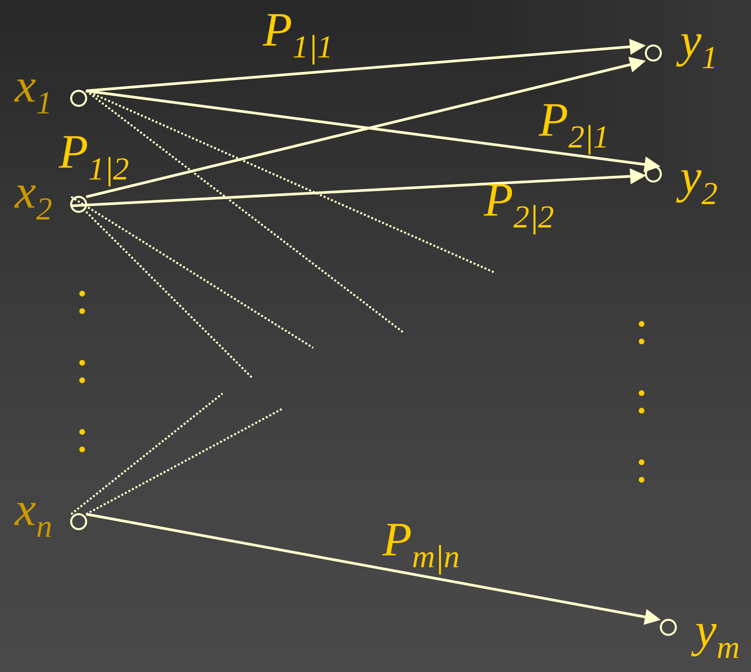
$H(X) = h(P_0)$

$H(X|Y) = e\, h(P_0)$

Thus $C_{erasure} = 1 - e$

(check!, draw and compare with BSC and Z)

channel models: general diagram



Input alphabet $X = \{x_1, x_2, \ldots, x_n\}$

Output alphabet $Y = \{y_1, y_2, \ldots, y_m\}$

$P_{j|i} = P_{Y|X}(y_j|x_i)$

In general:

calculating capacity needs more theory

clue:

I(X;Y)

is convex $\cap$ in the input probabilities

i.e. finding a maximum is simple

Channel capacity

Definition:

The rate R of a code is the ratio $\frac{k}{n}$, where

k is the number of information bits transmitted

in n channel uses

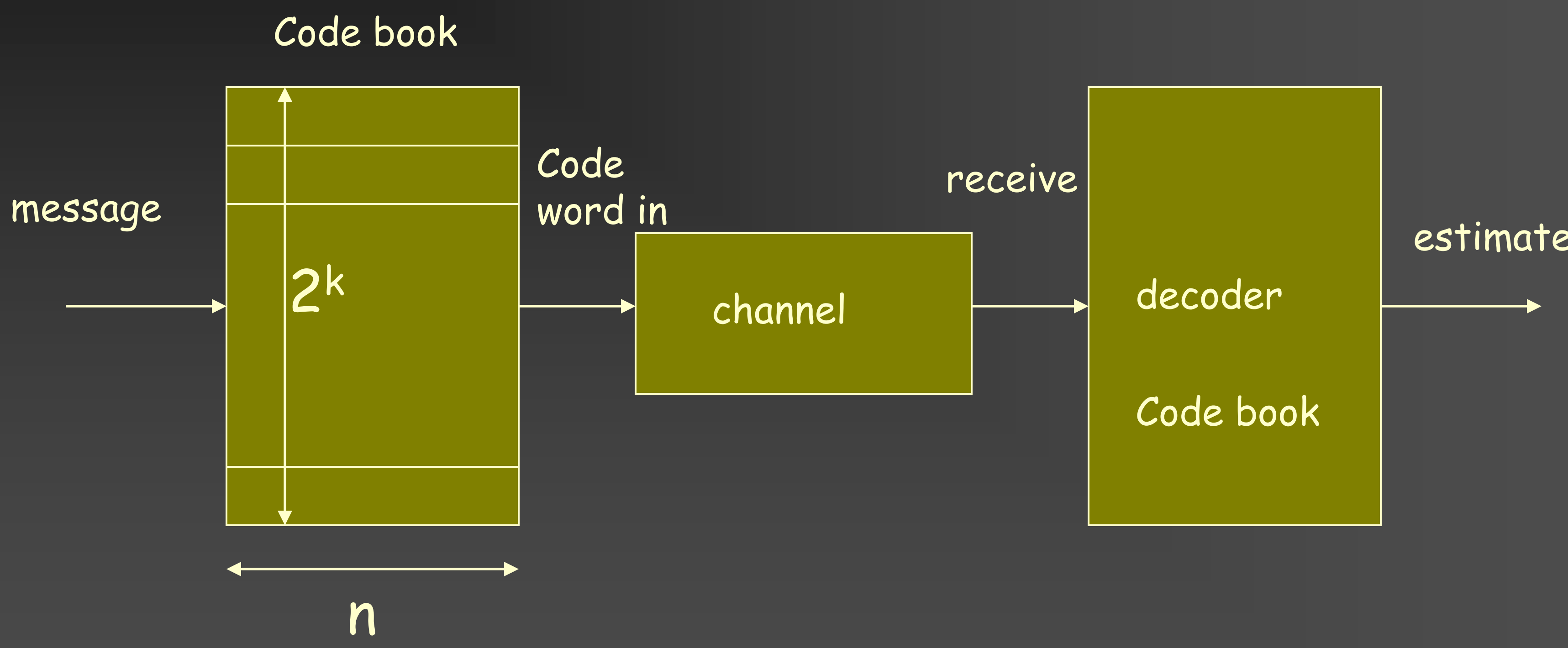
Shannon showed that: :

for $R \leq C$

encoding methods exist

with decoding error probability $\Rightarrow$ 0

System design



There are 2^k code words of length n

Channel capacity: sketch of proof for the BSC

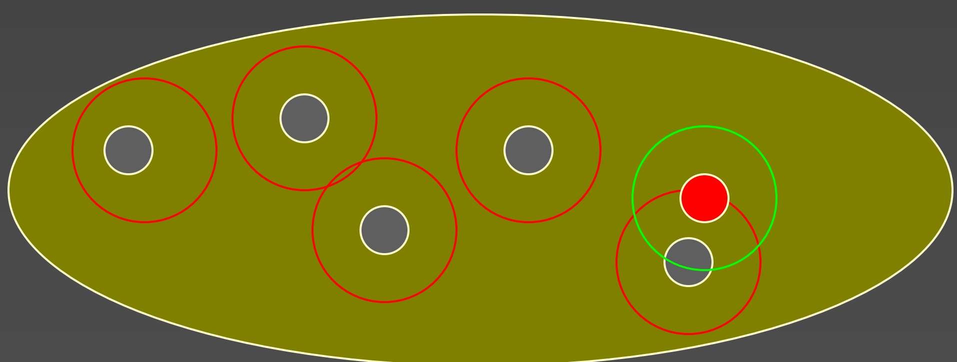
Code: 2^k binary codewords where $p(0) = P(1) = \frac{1}{2}$

Channel errors: $P(0 \rightarrow 1) = P(1 \rightarrow 0) = p$

i.e. # error sequences $\approx 2^{nh(p)}$

Decoder: search around received sequence for codeword

with $\approx$ np differences



space of 2^n binary sequences

Channel capacity: decoding error probability

1. For t errors: $|t/n-p| > \epsilon$

   $\rightarrow$ 0 for $n \rightarrow \infty$

   (law of large numbers)

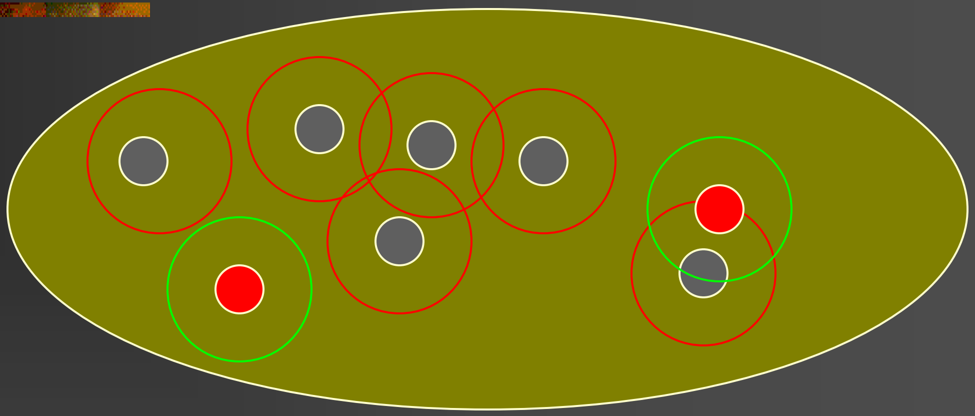
2. >1 code word in region

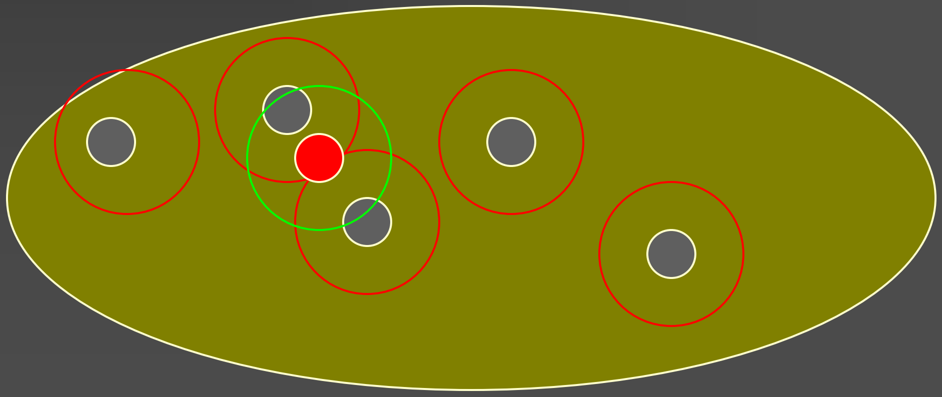
   (codewords random)

$$P(>1) \approx (2^k - 1)\frac{2^{nh(p)}}{2^n} \rightarrow 0$$

for $R = \frac{k}{n} < 1 - h(p)$

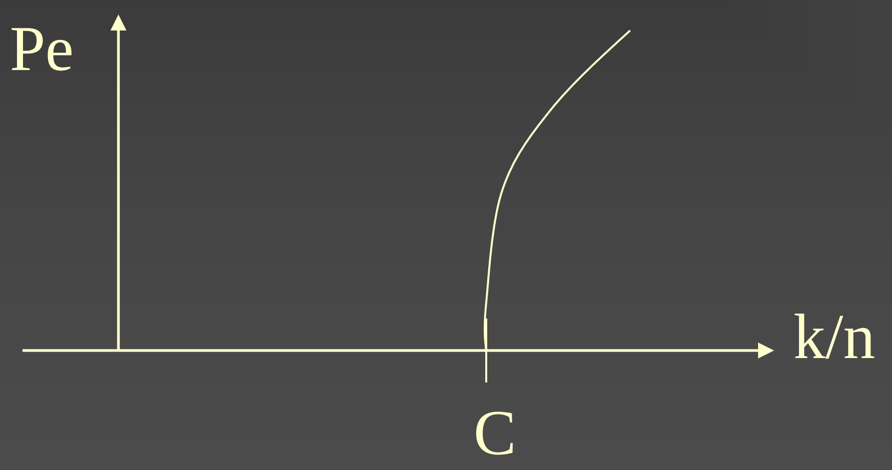
and $n \rightarrow \infty$



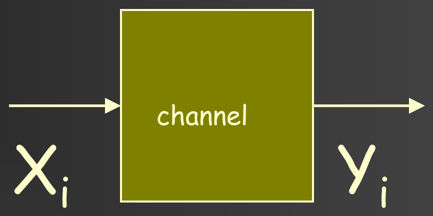


Channel capacity: converse

For R > C the decoding error probability > 0

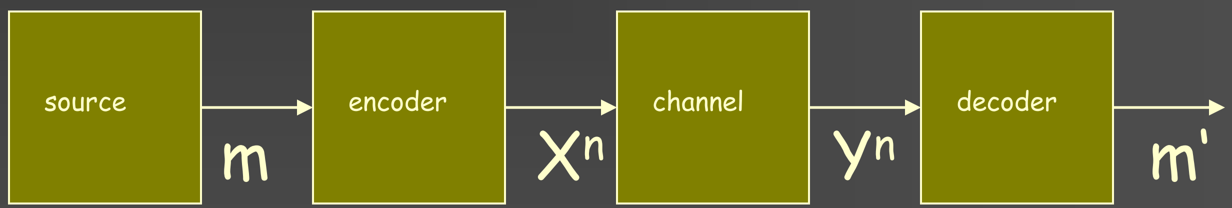


Converse: For a discrete memory less channel



$$I(X^n;Y^n) = H(Y^n) - \sum_{i=1}^{n} H(Y_i \mid X_i) \le \sum_{i=1}^{n} H(Y_i) - \sum_{i=1}^{n} H(Y_i \mid X_i) = \sum_{i=1}^{n} I(X_i;Y_i) \le nC$$

Source generates one out of 2^k equiprobable messages



Let Pe = probability that m' $\neq$ m

converse $R := k/n$

$$k = H(M) = I(M;Y^n) + H(M|Y^n)$$

$$\le \underbrace{I(X^n;Y^n)}_{X^n \text{ is a function of } M} + \underbrace{1 + k\,Pe}_{\text{Fano}}$$

$$\le nC + 1 + k\,Pe$$

$$1 - C\,n/k - 1/k \le Pe$$

$$Pe \ge 1 - C/R - 1/k$$

Hence: for large k, and $R > C$,

the probability of error $Pe > 0$

Appendix:

Assume:

binary sequence $P(0) = 1 - P(1) = 1-p$

t is the # of 1's in the sequence

Then $n \to \infty$, $\varepsilon > 0$

Weak law of large numbers

Probability ($|t/n - p| > \varepsilon$) $\to 0$

i.e. we expect with high probability pn 1's

Appendix:

Consequence:

1. $n(p-\varepsilon) < t < n(p+\varepsilon)$ with high probability

2. $$\log_2 \sum_{n(p-\varepsilon)}^{n(p+\varepsilon)} \binom{n}{t} \approx \log_2 (2n\varepsilon \binom{n}{pn}) \approx \log_2 2n\varepsilon + \log_2 2^{nh(p)}$$

3. $$\frac{1}{n}\log_2 2n\varepsilon + \frac{1}{n}\log_2 2^{nh(p)} \to h(p)$$

4. A sequence in this set has probability $\approx 2^{-nh(p)}$