Channel Coding Theory

Laura Cottatellucci

EURECOM

laura.cottatellucci@eurecom.fr

Course Information

Instructor	Laura Cottatellucci Office 333 Phone: 04 93 00 81 37 Any course related e-mal should contain the subject line "Channel Coding Theory" to get a quicker response.
Class Time	Wednesday 13:30–15:00 Wednesday 15:15–16:45
Location	Classroom 102
Important Dates	Exercise session: to be defined Final Exams: to be defined
Course Web Page	http://intranet.eurecom.fr/accademicAffaris/cursus_d/ courses_content/techcourses/Coding.

Course Information (2)

Goal The goal of this course is to introduce channel coding theory and its applications in communication systems. This is done by the design and performance analysis of encoders and decoders in both wireline and wireless communication systems. The course will address classical topics of algebraic coding theory (e.g. block codes and convolutional codes) and modern coding approaches based on soft decision.

Exams Home-works will be assigned during the course. Homework will be graded. There will be an exercise session (TD) and a final exam. Dates and rules of the exercise session and the exam will be announced in class and by e-mail.

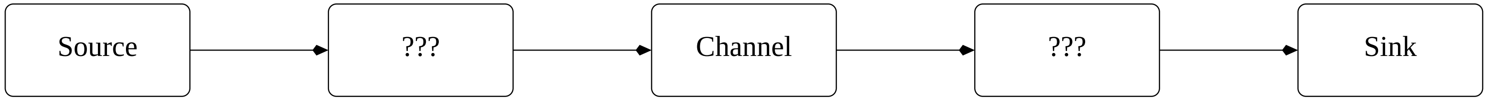
Grading Just as an indication: Reports of exercise session 40%, Home-works 20%, Final exam 40%.

Course Information (3)

Related Courses	Digital communications (Raymond Knopp) Information Theory (David Gesbert)
Coursenotes	You can find slides used during lectures in the web page.
Recommended Texts	At the end of each topic detailed references will be provided. Reference copies of the texts are available in the library or in my office. When possible, the referenced documents will be in the web page.

Outlines

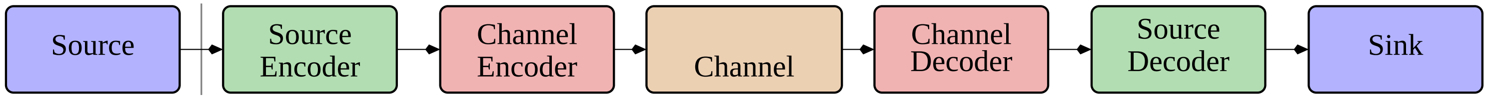
Basic Point-to-point Communication Problem



A source transmits via a noisy channel to a sink

Problem: Reliable transmission, i.e. we want to recreate the transmitted information with as little distortion as possible at the sink, at high speed.

Comm. Problem and Source-Channel Separation



Source Coding Theorem: **For a given source and distortion measure there exists a minimum rate $R = R(d)$ (bit per emitted source symbol) which is necessary and sufficient to describe this source with distortion not exceeding d.**

Channel Coding Theorem: **Given a channel there exists a maximum rate (bits per channel use) at which information can be transmitted reliably, i.e. with vanishing error probability, over the channel.**

The maximum rate is called the capacity of the channel and is denoted by C.

Comm. Problem and Source-Channel Separation (2)

Source-Channel Separation Coding Theorem: Given a channel of capacity C, a source can be reconstructed with a distortion at most of d at the receiver if $R(d) < C$, i.e. if the rate required to represent the given source with the allowed distortion is smaller than the channel capacity. Conversely, no scheme can do better.

Under the assumptions of the source-channel separation theorem

The approach of separating source and channel coding is optimum.

Comm. Problem and Source-Channel Separation (3)

Examples of situations where source channel coding separation is not optimum:

- **Delay constraints;**
- **Multi-user scenarios.**

Joint source coding could be substantially better in terms of complexity and delay.

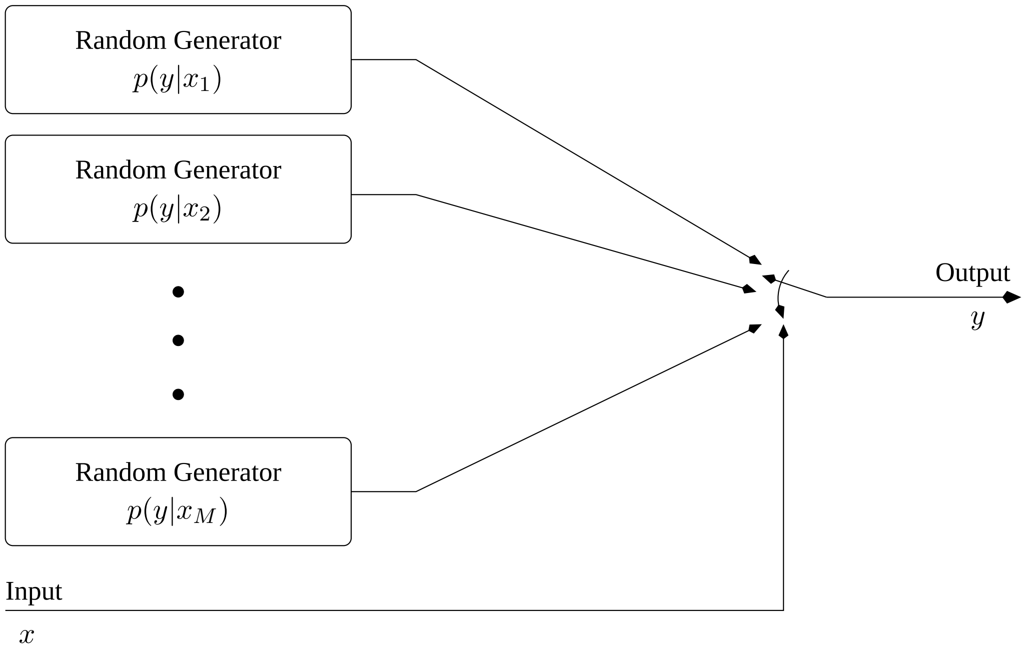
In the following we do not concern about source coding and we assume:

- **Source coding problem optimally solved.**
- **The source emits a sequence of i.i.d. bits.**

The Channel (1)

The channel is modelled as a random mapping from the input x to the output y

It is described by a cumulative distribution function $P(x,y)$ or a probability density(mass) function (p.d.(m.) f.) $p(x,y)$.



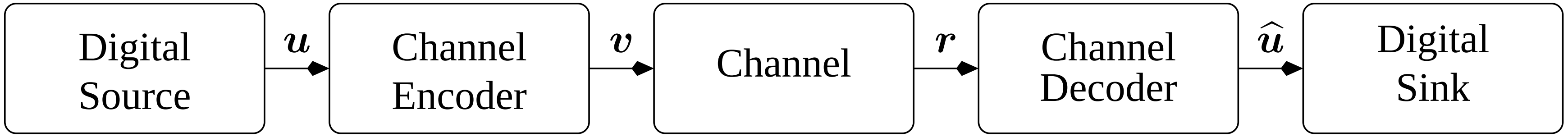
Channel: Taxonomy

with feedback	without feedback
e.g. Automatic Repeat Request (ARQ)	e.g. Forward Error Correction (FEC)
memoryless	with memory
e.g. flat fading, additive white noise, etc.	e.g. frequency selective fading, additive colored noise
without spatial diversity	with spatial diversity
e.g. single receive antenna	e.g. multiple transmit, multiple receive antennas
binary input	continuous waveform input
discrete output	continuous output

	Binary input	Continuous waveform input
Discrete output	BEC, BSC	
Continuous output	BIAWGN	AWGN, Fading channel

Discrete Input – Discrete Output Channels

Communication Model



$\boldsymbol{u} = (u_0, u_1, \ldots u_{k-1})$: binary k-tuple called message

$\boldsymbol{v} = (v_0, v_1, \ldots v_{n-1})$: n-tuple of discrete symbols called codeword

$\boldsymbol{r} = (r_0, r_1, \ldots r_{n-1})$: n-tuple of discrete symbols at the channel output

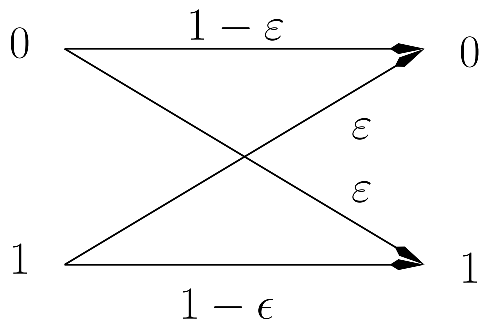
$\widehat{\boldsymbol{u}} = (\widehat{u}_0, \widehat{u}_1, \ldots \widehat{u}_{k-1})$: binary k-tuple called decoded message

$R = \frac{k}{n}$ **[bit/symbol] is the code rate**

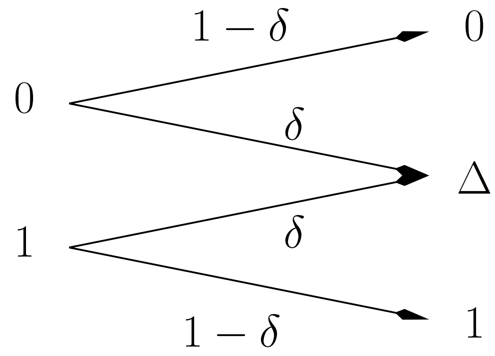
The code rate is the number of bits entering in the encoder per transmitted symbol.

Examples

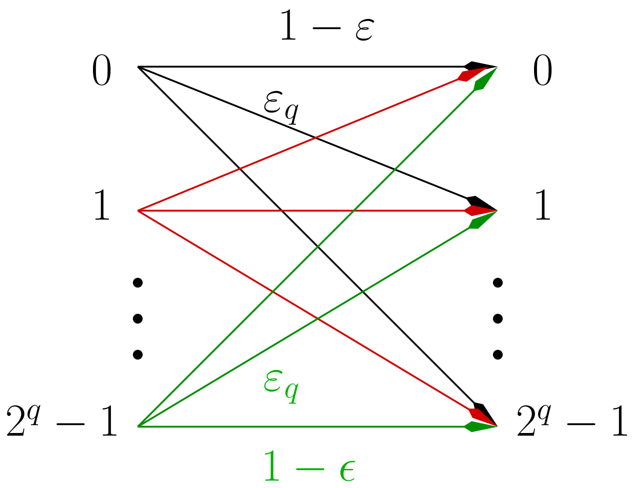
Binary Symmetric Channel (BSC)



Binary Erasure Channel (BEC)



Universal Binary Symmetric Channel (UBSC)



$\varepsilon_q = \frac{\varepsilon}{2^q-1}$

Some Capacities of Discrete Channels

From the general equation for the capacity of a channel $P_{X|Y}(x|y)$

$$C = \max_{P_X(x)} I(X;Y) = \max_{P_X(x)} \sum_y \sum_x P_{X,Y}(x,y) \log_2 \frac{P_{X,Y}(x,y)}{P_X(x)P_Y(y)}$$

$C_{\mathrm{BSC}} = 1 - \mathcal{H}(\varepsilon)$ with

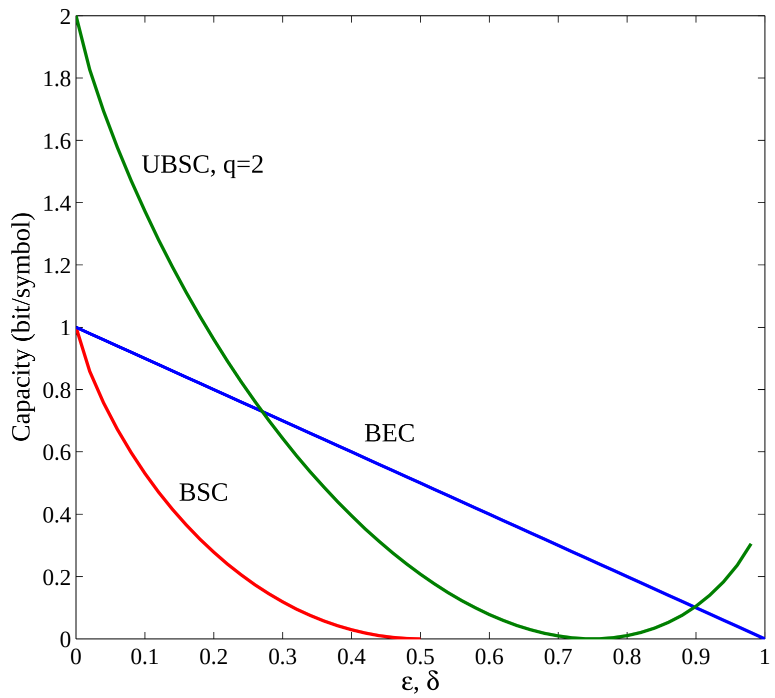
$\mathcal{H}(x) = -x \log_2(x) - (1-x)\log_2(1-x)$

being the entropy of a binary source

$C_{\mathrm{BEC}} = 1 - \delta$

$C_{\mathrm{UBSC}} = q + \varepsilon \log_2 \varepsilon_q + (1-\varepsilon)\log_2(1-\varepsilon)$

$\leq q$



The Coding Problem

Capacity is achievable with random codes for messages of infinite length...

...However, the complexity of the decoder is not affordable!

Example of a simple code: Repetition Code

	Decoded bit	Error probability	(rate, error prob.)
First trial:	$\widehat{u} = \widehat{u}_{\mathrm{MAP}}(r_1) = r_1$	$P_{e,1} = \varepsilon$	$(1, \varepsilon)$
Third trial:	$\widehat{u} = \widehat{u}_{\mathrm{MAP}}(r_1, r_2, r_3) = \mathrm{majority}(r_1, r_2, r_3)$	$P_{e,3} = \binom{3}{2} \varepsilon^2 (1-\varepsilon) + \varepsilon^3$	$\left(\frac{1}{3}, P_{e,3}\right)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
kth trial (k odd):	$\widehat{u} = \widehat{u}_{\mathrm{MAP}}(r_1, \ldots, r_k) = \mathrm{majority}(r_1, \ldots, r_k)$	$P_{e,k} = \sum_{i > \frac{k}{2}} \binom{k}{i} \varepsilon^i (1-\varepsilon)^{k-i}$	$\left(\frac{1}{k}, P_{e,k}\right)$

Low decoding complexity but in order to get $P_e \to 0$ also the rate vanishes!

Object: Design codes enabling low/affordable encoding and decoding complexity and vanishing error probability with positive rate.

Some Definitions

Definition of Code: A code $\mathcal{C}$ of length n and cardinality M over a field (set endowed with some specific operation) $\mathbb{F}$ is a subset of $\mathbb{F}^n$ with M elements.

$$\mathcal{C}(n, M) = \{\boldsymbol{v}^{[1]}, \boldsymbol{v}^{[2]}, \ldots, \boldsymbol{v}^{[M]}\}, \quad \boldsymbol{v}^{[m]} \in \mathbb{F}^n,\ 1 \leq m \leq M.$$

The elements of the code are called codewords. The parameter n is called the blocklength.

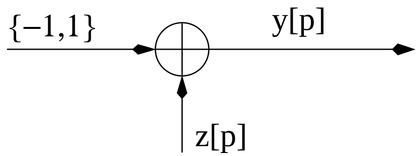
Example: For $\mathbb{F} = \{0, 1\}$, $n = 5$ and $M = 4$

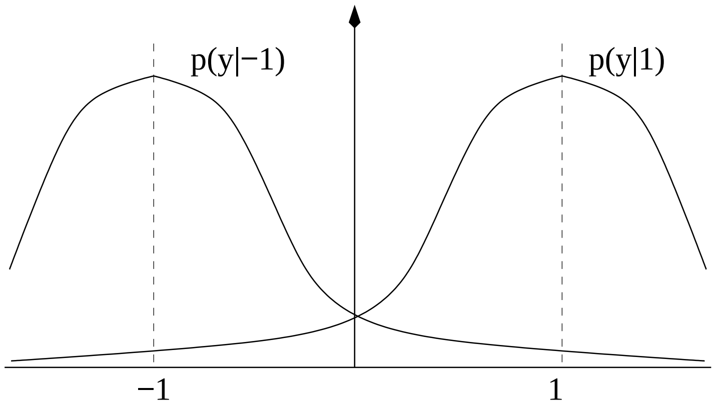
$$\mathcal{C} = \{01001, 11001, 11011, 10101\}$$

Definition of Encoder: An encoder $\mathcal{E}$ is the set of the M pairs $(\boldsymbol{u}, \boldsymbol{v})$ where $\boldsymbol{u}$ is the message or data word and $\boldsymbol{v}$ is the codeword belonging to a code $\mathcal{C}$.

Definition of Rate: The rate (measured as information symbol per transmitted symbol) of a code $\mathcal{C}(n, M) = \frac{1}{n}\log_{|\mathbb{F}|} M$.

Binary-Input Additive White Gaussian Channel





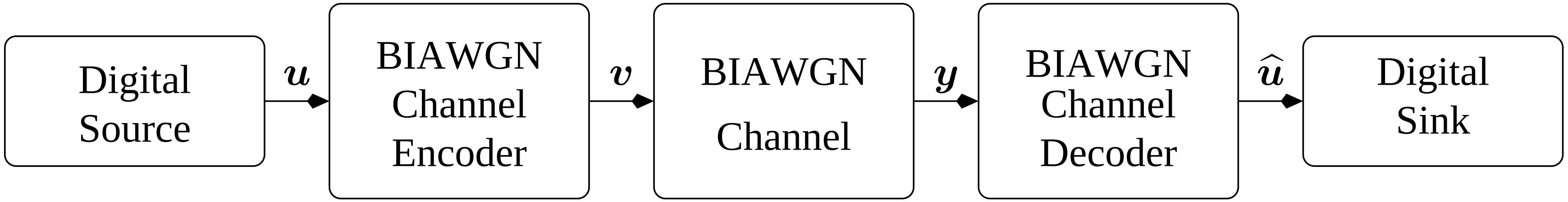
$z[p]$ is a random Gaussian sequence with uncorrelated elements (white noise), zero mean, and variance σ^2.

$$p_Z(z) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{z^2}{2\sigma^2}}$$

Then, the conditioned pdf describing the channel are also Gaussian.
Note that we implicitly assumed that the channel is symmetric.
If $\mathcal{E}_s$ is the energy of the transmitted symbol, i.e. we transmit $\pm\sqrt{\mathcal{E}_s}$, then

$$p_{Y|\pm\sqrt{\mathcal{E}_s}}(y|\pm\sqrt{\mathcal{E}_s}) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(y\mp\sqrt{\mathcal{E}_s})^2}{2\sigma^2}}$$

Optimum Communication System Model



$\boldsymbol{u}$, $\boldsymbol{v}$ and $\widehat{\boldsymbol{u}}$ are defined as for the discrete-input discrete-output channel.

$\boldsymbol{y} = (y_0, y_1, \ldots y_{n-1})$: n-tuple of received signals at the channel output with values in $\mathbb{R}$.

In contrast to a channel decoder for discrete output channels...

...a BIAWGN channel decoder processes analog sequences.

Such decoders are called soft decoders.

Capacity of a BIAWGN Channel (1)

Note that for a BIAWGN channel

$$p_{Y|X}(y|x = -\sqrt{\mathcal{E}}) = p_{Y|X}(-y|x = \sqrt{\mathcal{E}}).$$

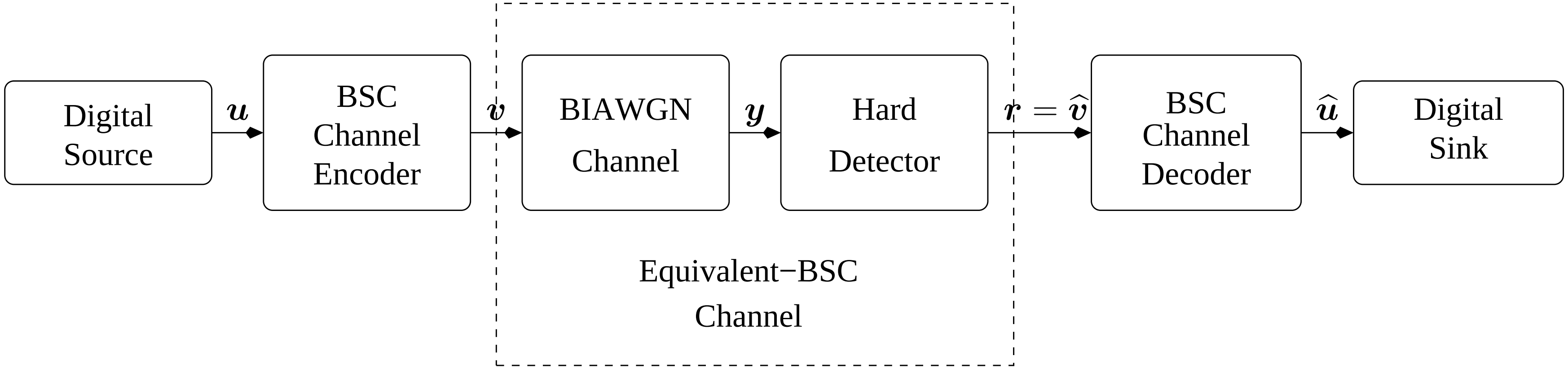
Thanks to the symmetry of the channel, the capacity is obtained for equiprobable inputs.

Capacity of a BIAWGN Channel (2)

$$
\begin{aligned}
C_{\text{BIAWGN}} &= \max_{P_X(x)} I(X;Y) \\
&= I(X;Y) \qquad \text{for} \quad P_X(x) = \frac{1}{2} \text{ and } x = \pm\sqrt{\mathcal{E}} \\
&= H(X) - H(X|Y) \\
&= 1 - \sum_{x=\pm\sqrt{\mathcal{E}}} \int_{\mathbb{R}} p_{XY}(x,y) \log \frac{1}{p_{X|Y}(x|y)} \mathrm{d}y \\
&= 1 - \sum_{x=\pm\sqrt{\mathcal{E}}} \int_{\mathbb{R}} p_{XY}(x,y) \log \frac{p_Y(y)}{p_{Y|X}(y|x) p_X(x)} \mathrm{d}y \\
&= 1 - \sum_{x=\pm\sqrt{\mathcal{E}}} \int_{\mathbb{R}} p_{XY}(x,y) \log \frac{\sum_{x=\pm\sqrt{\mathcal{E}}} p_{Y|X}(y|x) p_X(x)}{p_{Y|X}(y|x) p_X(x)} \mathrm{d}y \\
&= 1 - \int_{\mathbb{R}} p_{Y|X}(y|x=\pm\sqrt{\mathcal{E}}) \log \frac{2\sum_{x=\pm\sqrt{\mathcal{E}}} p_{Y|X}(y|x) p_X(x)}{p_{Y|X}(y|x=\pm\sqrt{\mathcal{E}})} \mathrm{d}y \qquad \text{for} \quad x = -\sqrt{\mathcal{E}} \text{ or } x = \sqrt{\mathcal{E}}
\end{aligned}
$$

* $H(X)$ is the entropy of X. In case of binary inputs $H(X) = \mathcal{H}(x)$.

A Suboptimum Approach



$\boldsymbol{u}$, $\boldsymbol{v}$, and $\widehat{\boldsymbol{u}}$ are defined as for the discrete-input discrete-output channel.

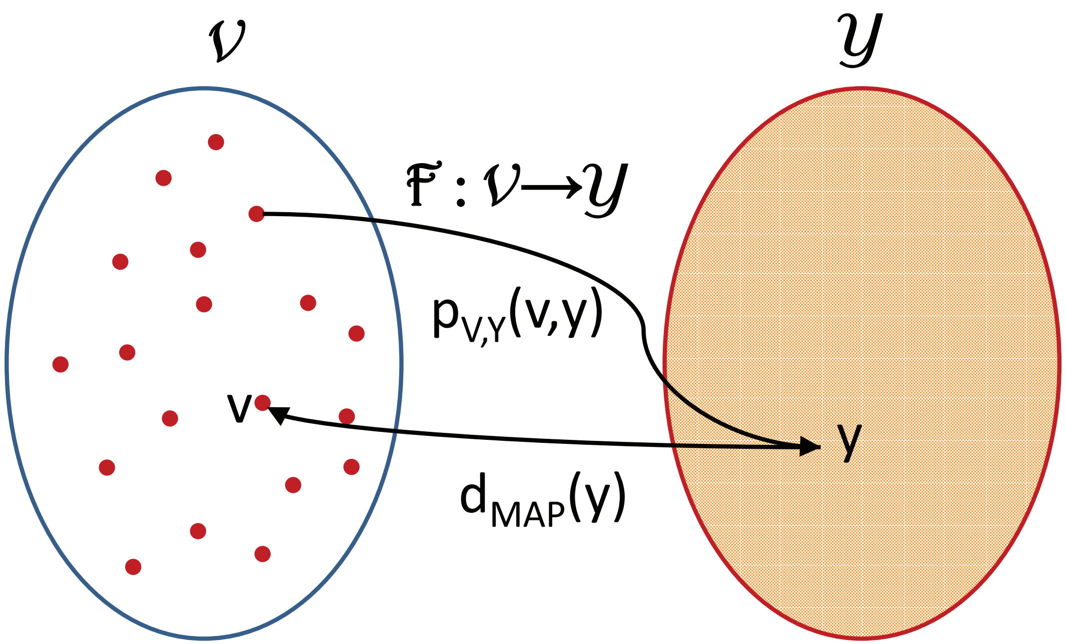
$\boldsymbol{r}$ is defined as in the optimum communication system for BIAWGN channel.

$\widehat{\boldsymbol{v}} = (\widehat{v}_0, \widehat{v}_1, \ldots . \widehat{v}_{n-1})$: n-tuple of binary symbols corresponding to the detected transmitted signals $\boldsymbol{v}$.

The hard detector maps each received real element of the vector y onto an element of a binary set.

Maximum A Posteriori (MAP) Probability Detector

Definition of Detector: **Given a random mapping from a discrete set $\mathcal{V}$ to a discrete or continuous set $\mathcal{Y}$ a detector $d(y)$ of $v \in \mathcal{V}$ given $y \in \mathcal{Y}$ is a deterministic mapping $d : \mathcal{Y} \rightarrow \mathcal{V}$.**



Maximum A Posteriori (MAP) Probability Detector (2)

If the random mapping $\mathcal{F} : \mathcal{V} \to \mathcal{Y}$ is described by the pdf (pms) $p_{V,Y}(v, y)$, then the MAP detector is given by

$$d_{\mathrm{MAP}}(y) = \underset{v}{\mathrm{argmax}} p_{V|Y}(v|y).$$

Using the Markov rule $p_{V,Y}(v, y) = p_{V|Y}(v|y)p_Y(y) = p_{Y|V}(y|v)p_V(v)$

$$\begin{aligned} d_{\mathrm{MAP}}(y) &= \underset{v}{\mathrm{argmax}} \frac{p_{Y|V}(y|v)p_V(v)}{p_Y(y)} \\ &= \underset{v}{\mathrm{argmax}}\, p_{Y|V}(y|v)p_V(v) \end{aligned}$$

Maximum A Posteriori Probability Detector (3)

If V is an equiprobable random variable, i.e. $\forall v \in \mathcal{V}\ \ P_V(v) = \frac{1}{|\mathcal{V}|}$

$$d_{\mathrm{MAP}}(y) = \underset{v}{\operatorname{argmax}}\, p_{Y|V}(y|v)$$

$d_{\mathrm{MAP}}(y)$ coincides with the definition of the maximum likelihood detector $d_{\mathrm{ML}}(y)$.

If $|\mathcal{V}| = 2$

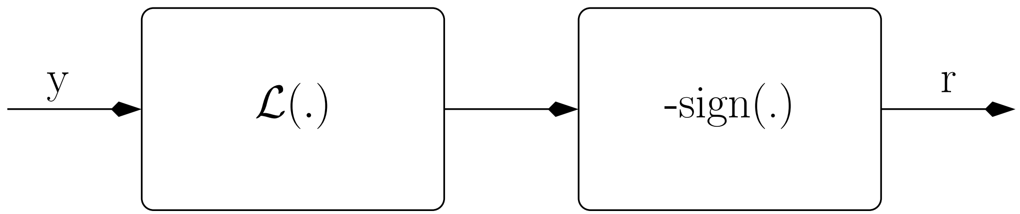
$$\begin{aligned} d_{\mathrm{MAP}}(y) &= \begin{cases} 0\ (-1) & \text{for } \frac{p_{Y|V}(y|0)}{p_{Y|V}(y|1)} > 1, \\ 1 & \text{otherwise} \end{cases} \\ &= \begin{cases} 0\ (-1) & \text{for } \log \frac{p_{Y|V}(y|0)}{p_{Y|V}(y|1)} > 0, \\ 1 & \text{otherwise} \end{cases} \end{aligned}$$

Maximum A Posteriori Probability Detector (4)

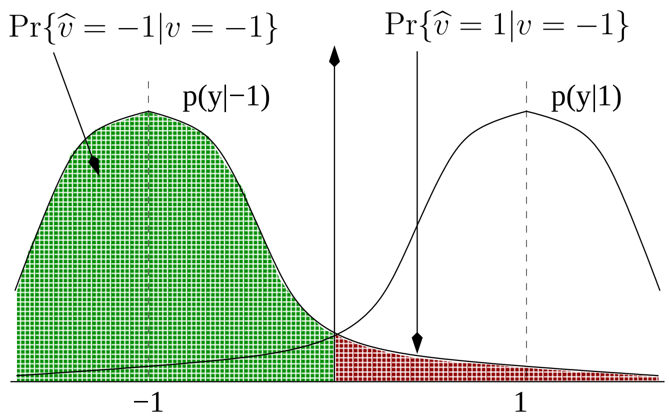
The ratio $\frac{p_{Y|V}(y|0)}{p_{Y|V}(y|1)}$ is called likelihood ratio.

$\mathcal{L}(y) = \log \frac{p_{Y|V}(y|0)}{p_{Y|V}(y|1)}$ is the log-likelihood ratio.

Hard Detector for a BIAWGN channel



Capacity of the Suboptimum Scheme



$\mathcal{E}$: average energy of the transmitted signal.
σ^2 : variance of the AWGN.
$\mathrm{SNR}=\frac{\mathcal{E}}{\sigma^2}$

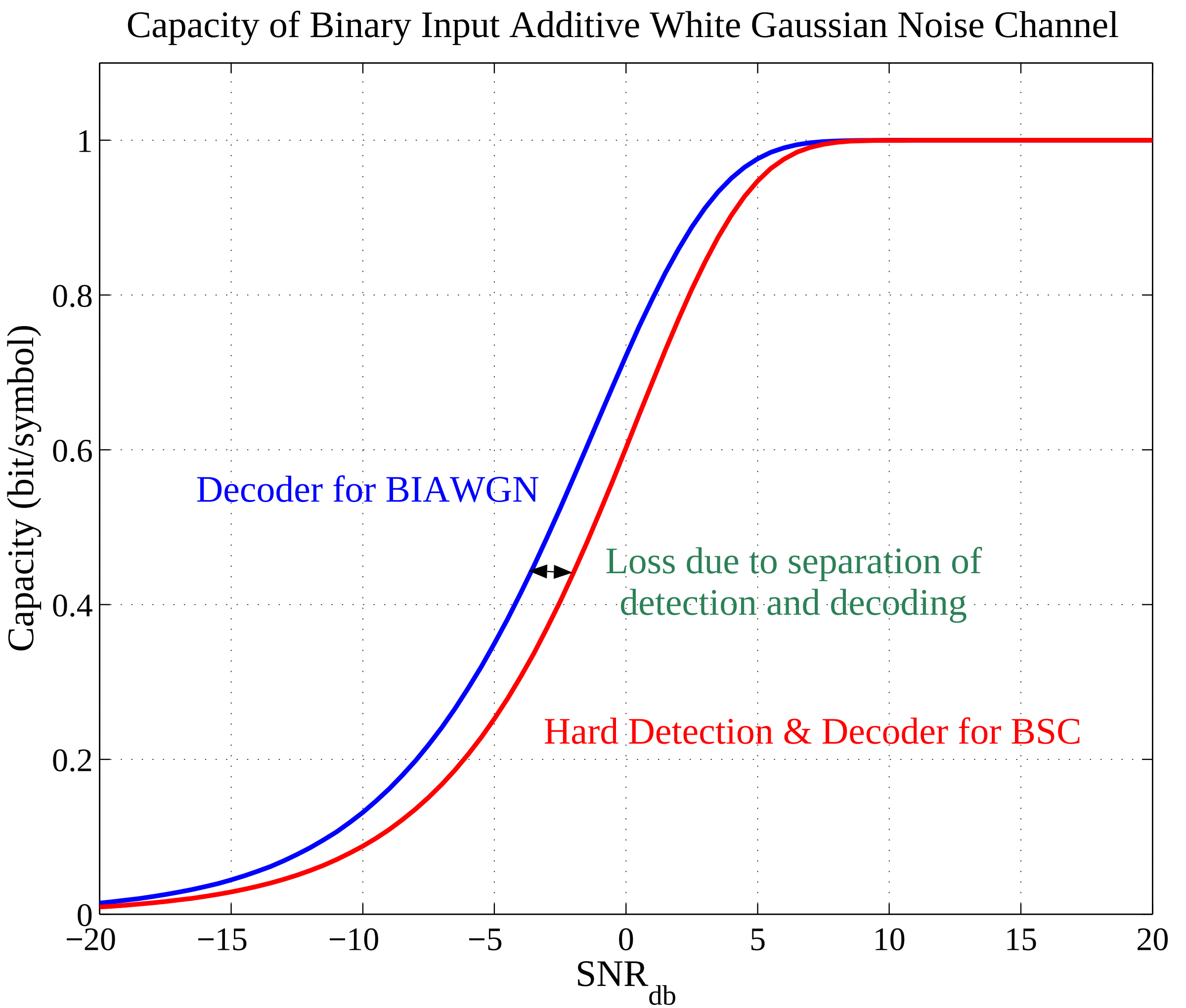
Probability of error of the hard detector

$$
\begin{aligned}
P_e &= \Pr\{\widehat{v}=-1|v=1\}P_V(v=1)+\Pr\{\widehat{v}=+1|v=-1\}P_V(v=-1)\\
&= \int_0^{+\infty}\frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-\frac{(y+\sqrt{\mathcal{E}})^2}{2\sigma^2}}\mathrm{d}y\\
&= \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{\mathrm{SNR}}{2}}\right)
\end{aligned}
$$

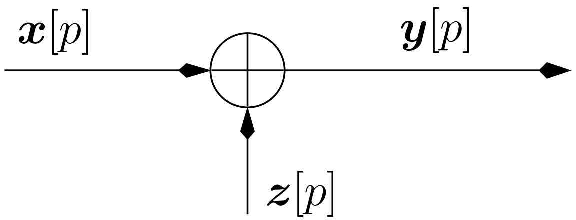
The capacity of the suboptimum scheme is given by the capacity of the equivalent BSC with $\varepsilon=P_e$.

$$
C_{\mathrm{eq-BSC}}=1-\mathcal{H}\left(\frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{\mathrm{SNR}}{2}}\right)\right)
$$

BIAWGN Channel versus the Equivalent-BSC



Additive White Gaussian Channel (Discrete-Time)



Channel Input: $\{\boldsymbol{x}[p],\ p \in \mathcal{Z}\}$, sequence of elements in $\mathbb{R}^N$, i.e. $\boldsymbol{x}[p] \in \mathbb{R}^N$.

For $N = 1$, $\{\boldsymbol{x}[p]\}$ is a sequence of reals. For $N = 2$, $\{\boldsymbol{x}[p]\}$ is a sequence of complex numbers. For $N > 2$, $\{\boldsymbol{x}[p]\}$ is a sequence of vectors (typical case: MIMO channels).

$$\boldsymbol{Y}|\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{x}, \boldsymbol{Q})$$

Additive Noise: $\{\boldsymbol{z}[p],\ p \in \mathcal{Z}\}$, is a random process (discrete time) of independent elements in $\mathbb{R}^N$ described by the pdf

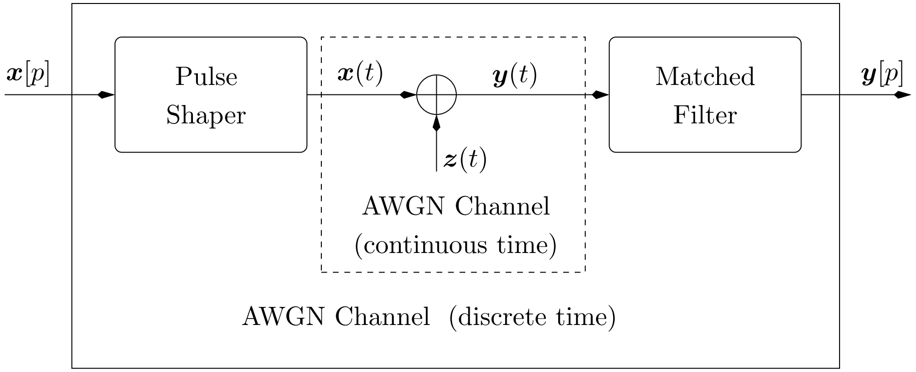
$$P_{\boldsymbol{Z}}(\boldsymbol{z}) = \frac{1}{\sqrt{(2\pi)^N \det \boldsymbol{Q}}} \mathrm{e}^{-\frac{\boldsymbol{z}^H \boldsymbol{Q}^{-1} \boldsymbol{z}}{2}} \qquad (\mathcal{N}(\mathbf{0}, \boldsymbol{Q}))$$

being $\boldsymbol{Q} \in \mathbb{R}^{N \times N}$ a diagonal matrix. We assume $\boldsymbol{Q} = \sigma^2 \boldsymbol{I}$ (i.i.d. random variables).

Channel Output: $\{\boldsymbol{y}[p],\ p \in \mathbb{Z}\}$ random sequence of elements in $\mathbb{R}^N$.

Discrete-time versus Continuous-time AWGN Channels

An AWGN channel in the discrete-time domain is typically obtained from a continuous time AWGN channel via a pulse shaper and a matched filter.



Pulse waveforms with unit energy

$\Downarrow$

- $\mathcal{E}_p$, the energy of continuous waveform transmitted in pT equals the energy of the discrete signal $\boldsymbol{x}[p]$, i.e. $\mathcal{E}_p = \|\boldsymbol{x}[p]\|^2$;
- The variance σ^2 of the discrete Gaussian noise per dimension equals $\frac{N_0}{2}$, the power spectral density per dimension of the continuous noise, i.e. $\sigma^2 = \frac{N_0}{2}$.

Capacity of an AWGN Channel (Discrete Time)

Capacity in an AWGN channel can be achieved with Gaussian distributed input.

Under the constraint of average energy $\mathcal{E}$, $p_{\boldsymbol{X}}(\boldsymbol{x}) = \left(\frac{2\pi\mathcal{E}}{N}\right)^{-\frac{N}{2}} \mathrm{e}^{-\frac{|\boldsymbol{x}|^2}{2\mathcal{E}}}$

Let us focus on $N = 1$ and recall that the entropy of a Gaussian random variable X is

$$H(x) = \frac{1}{2}\log 2\pi\mathrm{e}\mathrm{E}\{x^2\}$$

Since

$$Y \sim \mathcal{N}\left(0, \mathcal{E} + \frac{N_0}{2}\right) \qquad Z \sim \mathcal{N}\left(0, \frac{N_0}{2}\right)$$

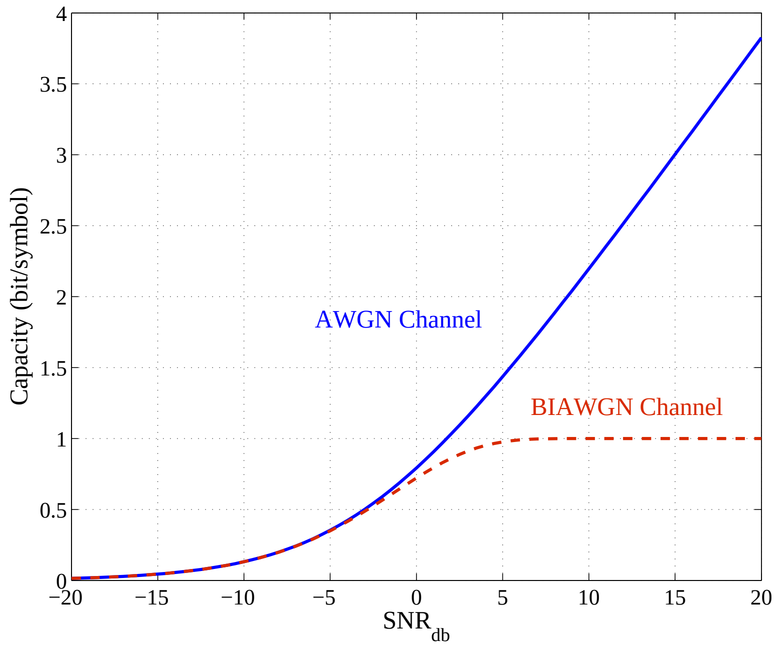
$$\Downarrow$$

$$C = H(Y) - H(Y|X) = \frac{1}{2}\log\left(2\pi\mathrm{e}\left(\mathcal{E} + \frac{N_0}{2}\right)\right) - \frac{1}{2}\log\left(2\pi\mathrm{e}\frac{N_0}{2}\right)$$
$$= \frac{1}{2}\log\left(1 + \frac{2\mathcal{E}}{N_0}\right)$$

being $\frac{2\mathcal{E}}{N_0} = \mathrm{SNR}$.

AWGN Channel versus BIAWGN Channel: Capacity

Transmission of antipodal signals in a AWGN channel $\Rrightarrow$ BIAWGN channel



Transmission of codewords with antipodal signals (BIAWGN channel) almost optimum only at low SNR.

How to use the channel at high SNR effectively?

Coding in a signal space!

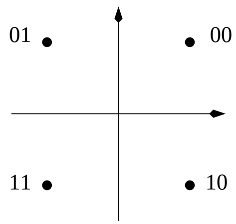
Coding in a Signal Space

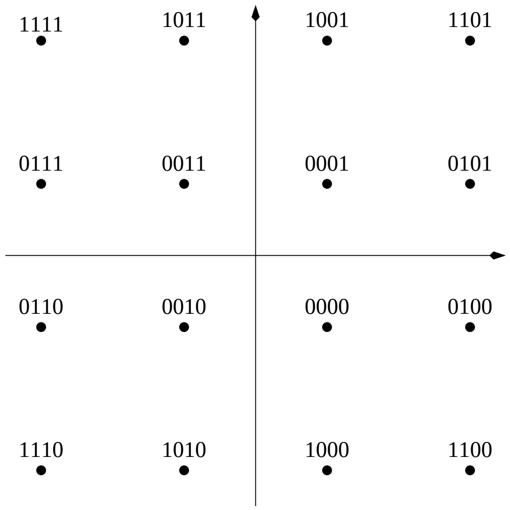
Definition: A finite signal constellation is a finite set $\mathcal{S} = \{x\}$ of vectors in $\mathbb{R}^N$ having cardinality $M = |\mathcal{S}|$. The vectors x are called points or signals.

For simplicity we assume $M = 2^m$.

Definition: Labelling is a one-to-one map, which associates with every element $x \in \mathcal{S}$ an n-tuple of binary digits.

Examples



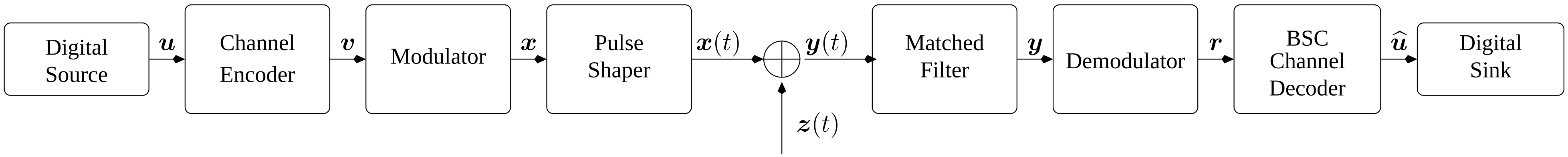


Coding in a Signal Space (2)

Definition: Modulation is the association of a n-tuple of binary digits with the elements of $\mathcal{S}$, based on labelling.

Definition: The demodulator transforms the received vector y back to a sequence of m binary digits. Note that in general the dimension of y may differ from the dimension of x.

System Model for an AWGN Channel



$\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{y}$, $\boldsymbol{r}$, and $\widehat{\boldsymbol{u}}$ are as previously defined.

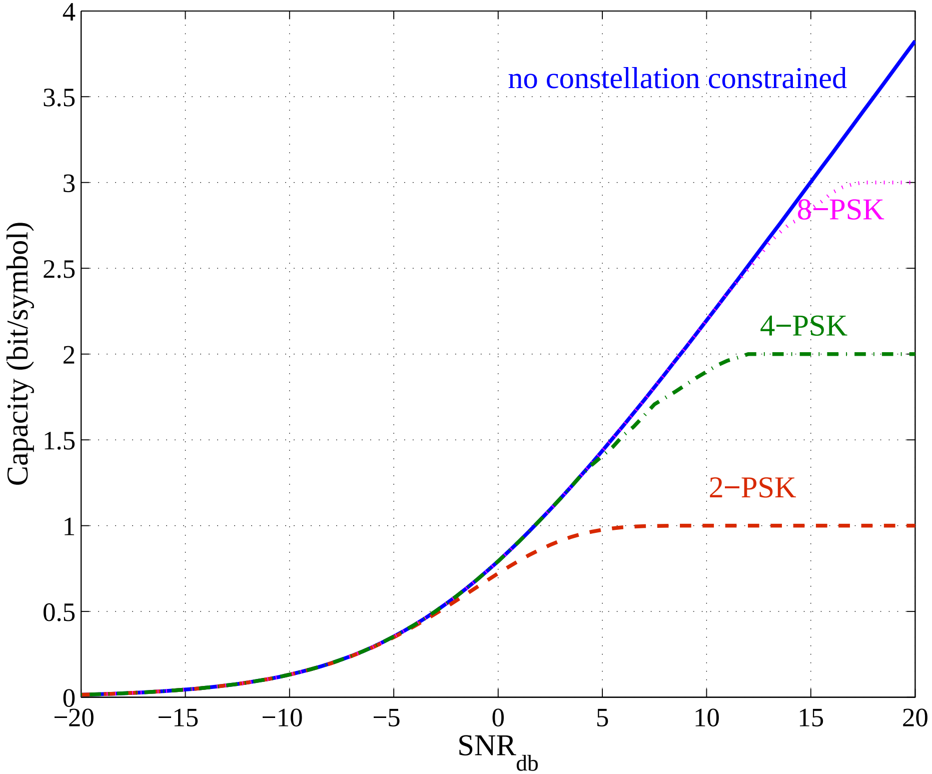
$\{\boldsymbol{x}\}$ is a sequence of elements in the signal space $\mathcal{S}$.

Capacity of a Constellation-Constrained AWGN Channel

To avoid the maximization of the mutual information over the a priori probability, we make the simplifying assumption that the signals in $|\mathcal{S}|$ are transmitted with equal probabilities.

As for the BIAWGN channel

$$
\begin{aligned}
I(\boldsymbol{X};\boldsymbol{Y}) &= H(\boldsymbol{X}) - H(\boldsymbol{X}|\boldsymbol{Y}) \\
&= \log|\mathcal{S}| - \mathrm{E}\left\{\log\frac{1}{p_{\boldsymbol{X}|\boldsymbol{Y}}(\boldsymbol{x}|\boldsymbol{y})}\right\} \\
&= \log|\mathcal{S}| - \mathrm{E}\left\{\log\frac{\sum_{\boldsymbol{x}'\in\mathcal{S}} p_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}')}{p_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x})}\right\}
\end{aligned}
$$



Performance Measures for a Channel Decoder

We are interested mainly in two performance measures:

- **Message (symbol) error probability P_e (or SER)**

$$P_e = \sum_{\boldsymbol{u}} \Pr(\boldsymbol{u} \neq \widehat{\boldsymbol{u}})$$

- **Bit error probability P_b (or BER)**

$$P_b = \sum_{\boldsymbol{u}} \sum_{k=1}^{K} \Pr(u_k \neq \widehat{u}_k)$$

Relation between P_e and P_b

$$\frac{P_e}{K} \leq P_b \leq P_e$$

Maximum a Posteriori Decoders

Object: Minimize $P_e = \sum_{\boldsymbol{u}\in\mathbb{F}^K} \sum_{\boldsymbol{r}\in\mathcal{S}} \Pr(\boldsymbol{u} \neq \widehat{\boldsymbol{u}}|\boldsymbol{r})\Pr(\boldsymbol{r})$

P_e **is minimized if** $\Pr(\boldsymbol{u} \neq \widehat{\boldsymbol{u}}|\boldsymbol{r})$ **is minimized, or, equivalently, if** $\Pr(\boldsymbol{u} = \widehat{\boldsymbol{u}}|\boldsymbol{r})$ **is maximized. Then, the decoder that minimizes the message error probability is**

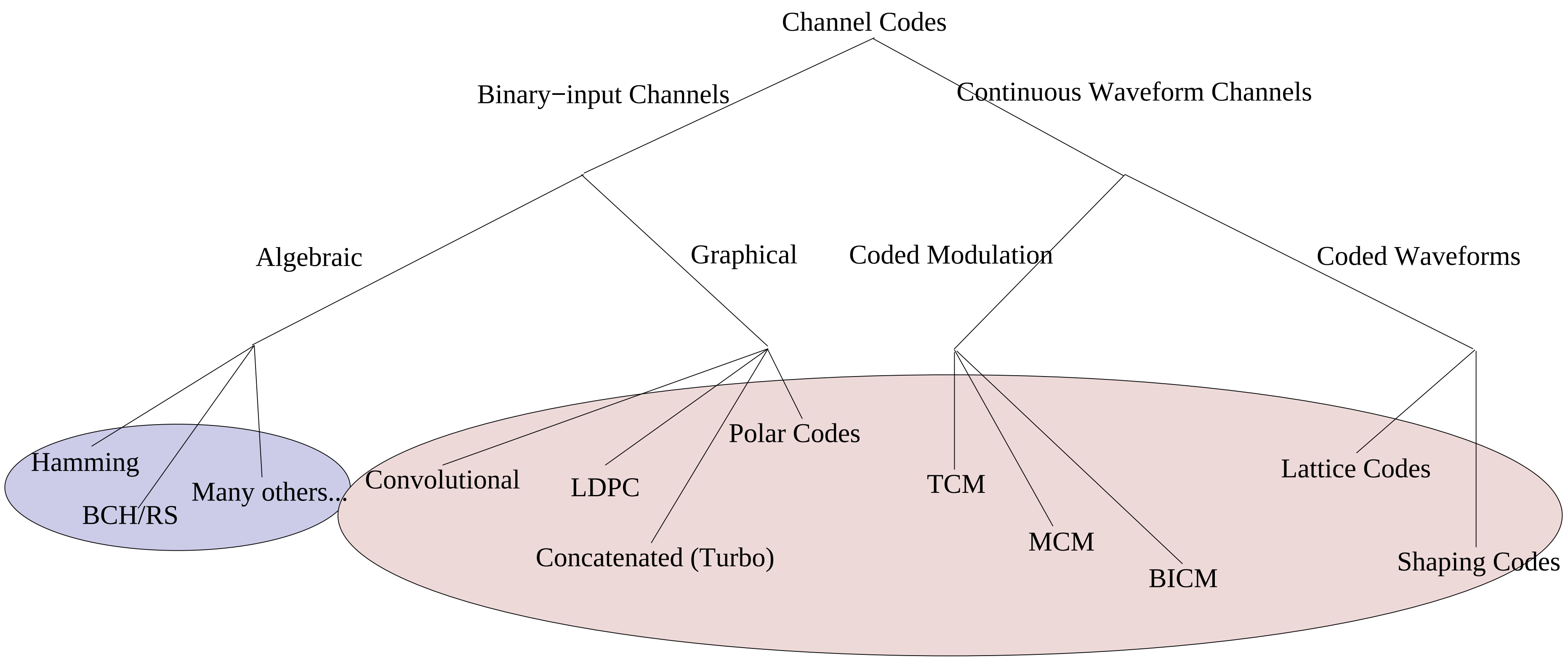
$$\begin{aligned}\widehat{\boldsymbol{u}}_{\mathrm{MAP}} &= \underset{\boldsymbol{u}}{\operatorname{argmax}} \mathrm{P}_{\boldsymbol{U}|\boldsymbol{R}}(\boldsymbol{u}|\boldsymbol{r}) \\ &= \underset{\boldsymbol{u}}{\operatorname{argmax}} \frac{\mathrm{P}_{\boldsymbol{R}|\boldsymbol{U}}(\boldsymbol{r}|\boldsymbol{u})\mathrm{P}_{\boldsymbol{U}}(\boldsymbol{u})}{\mathrm{P}_{\boldsymbol{R}}(\boldsymbol{r})} \\ &= \underset{\boldsymbol{u}}{\operatorname{argmax}} \mathrm{P}_{\boldsymbol{R}|\boldsymbol{U}}(\boldsymbol{r}|\boldsymbol{u})\mathrm{P}_{\boldsymbol{U}}(\boldsymbol{u})\end{aligned}$$

If the transmitted messages are equiprobable

$$\widehat{\boldsymbol{u}}_{\mathrm{MAP}} = \underset{\boldsymbol{u}}{\operatorname{argmax}} \mathrm{P}_{\boldsymbol{R}|\boldsymbol{U}}(\boldsymbol{r}|\boldsymbol{u})$$

For equiprobable messages the MAP decoder coincides with the maximum likelihood decoder.

Channel Code Taxonomy



Outlines

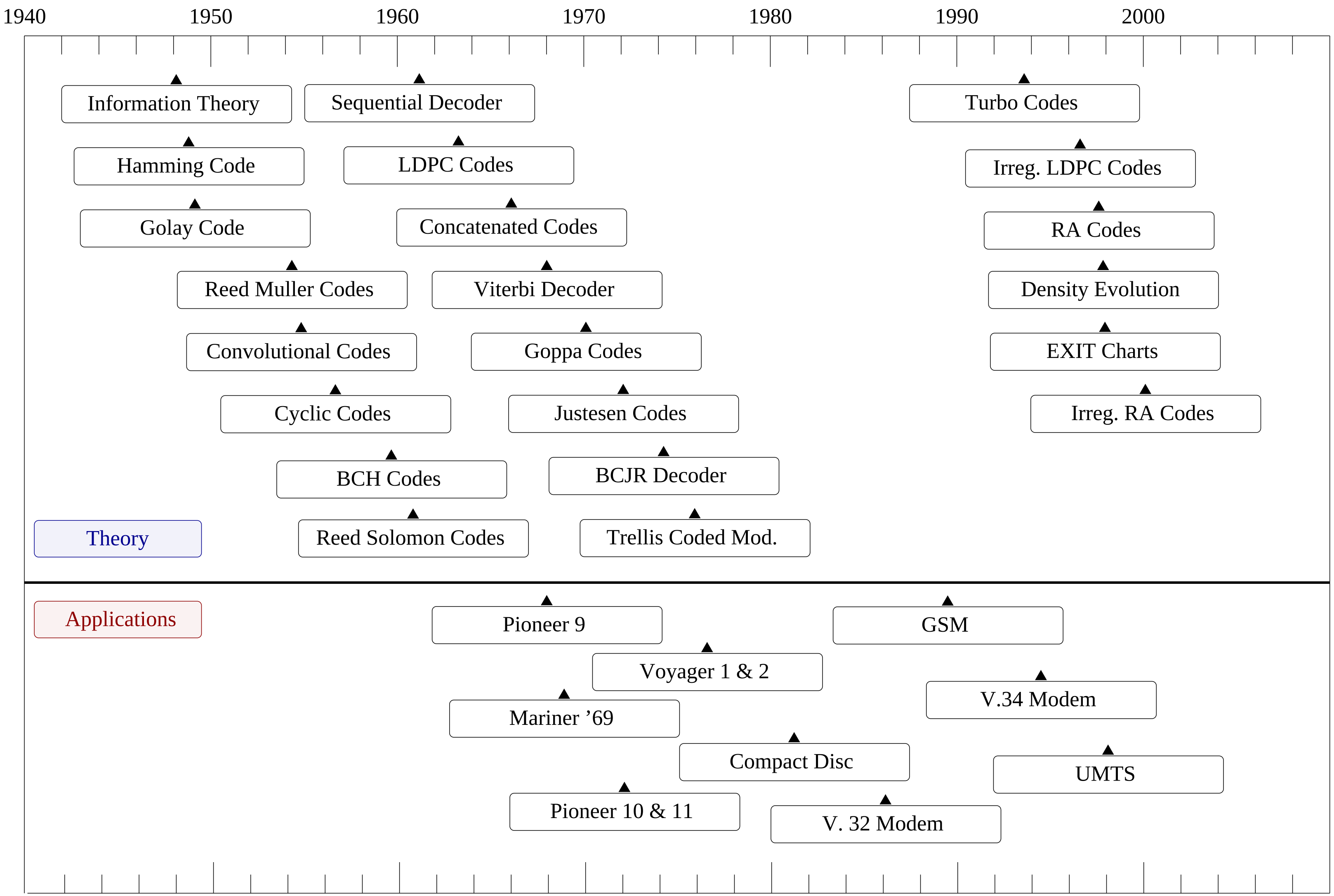
- **Block codes**
- **Convolutional codes**
- **Serially and parallel concatenated (turbo) codes**
- **Low density parity check (LDPC) and repeat-accumulate codes**

Two Milestones in the History of Channel Coding

- **C.E. Shannon, A Mathematical Theory of Communication,** *The Bell System Technical Journal,* **vol. 27, pp. 379-423, 623-656, July, October, 1948.**

- **R.W. Hamming, Error-Detecting and Error-Correcting Codes,** *The Bell System Technical Journal,* **vol. 29, pp. 147-160, 1950.**

Timeline


1940
1950
1960
1970
1980
1990
2000
Information Theory
Sequential Decoder
Turbo Codes
Hamming Code
LDPC Codes
Irreg. LDPC Codes
Golay Code
Concatenated Codes
RA Codes
Reed Muller Codes
Viterbi Decoder
Density Evolution
Convolutional Codes
Goppa Codes
EXIT Charts
Cyclic Codes
Justesen Codes
Irreg. RA Codes
BCH Codes
BCJR Decoder
Theory
Reed Solomon Codes
Trellis Coded Mod.
Applications
Pioneer 9
GSM
Voyager 1 & 2
V.34 Modem
Mariner '69
Compact Disc
UMTS
Pioneer 10 & 11
V. 32 Modem

Codes Used in Deep-Space Communications

Year	Mission	Code
1968	Pioneer 9	$R = 1/2$ convolutional, memory 20, sequential decoder
1969	Mariner '69	(32,6) Reed-Muller, correlation decoder
1972	Pioneer 10 & 11	$(2, 1, 32)$ convolutional, memory 31, sequential decoder
1977	Voyager 1 & 2	$(2, 1, 7)$ convolutional, Viterbi decoder
1987	CCSDS* Telemetry Standard	concatenation, $(255, 223)$ Reed Solomon and $(2, 1, 7)$ convolutional
1989	Galileo	concatenation, $(255, 223)$ Reed Solomon and $(4, 1, 15)$ convolutional
1999	CCSDS Telemetry Recommendation	adds Turbo code specification

* CCSDS: Consultative Committee for Space Data System, http://www.ccsds.org
From G. Kramer, Turbo codes and Iterative Decoding, Mini course at TUWien, Vienna.

References

- S. Lin and D. Costello, **Error Control Coding,** *Prentice Hall,* 1983. Chapter 1.
- S. Benedetto and E. Biglieri, **Principles of Digital Transmission with Wireless Applications,** , Chapter 10.
- E. Biglieri, **Coding in a Signal Space,** Springer-Verlag, 2005, Chapter 3.
- T. Richardson and R. Urbanke, **Modern Coding Theory,** *http://lthcwww.epfl.ch/mct/index.php,* Chapter 1.
- J.G. Proakis, Digital Communications, McGraw Hill.
- G. Caire, Lecture notes for the course on Information Theory and Coding, 2003.
- D.J. Costello, J. Hagenauer, H. Imai, S.B. Wicker, **Applications of Error-Control Coding,** *IEEE Transactions on Information Theory,* vol. 44, n. 6, pp. 2531-2560, October 1998.