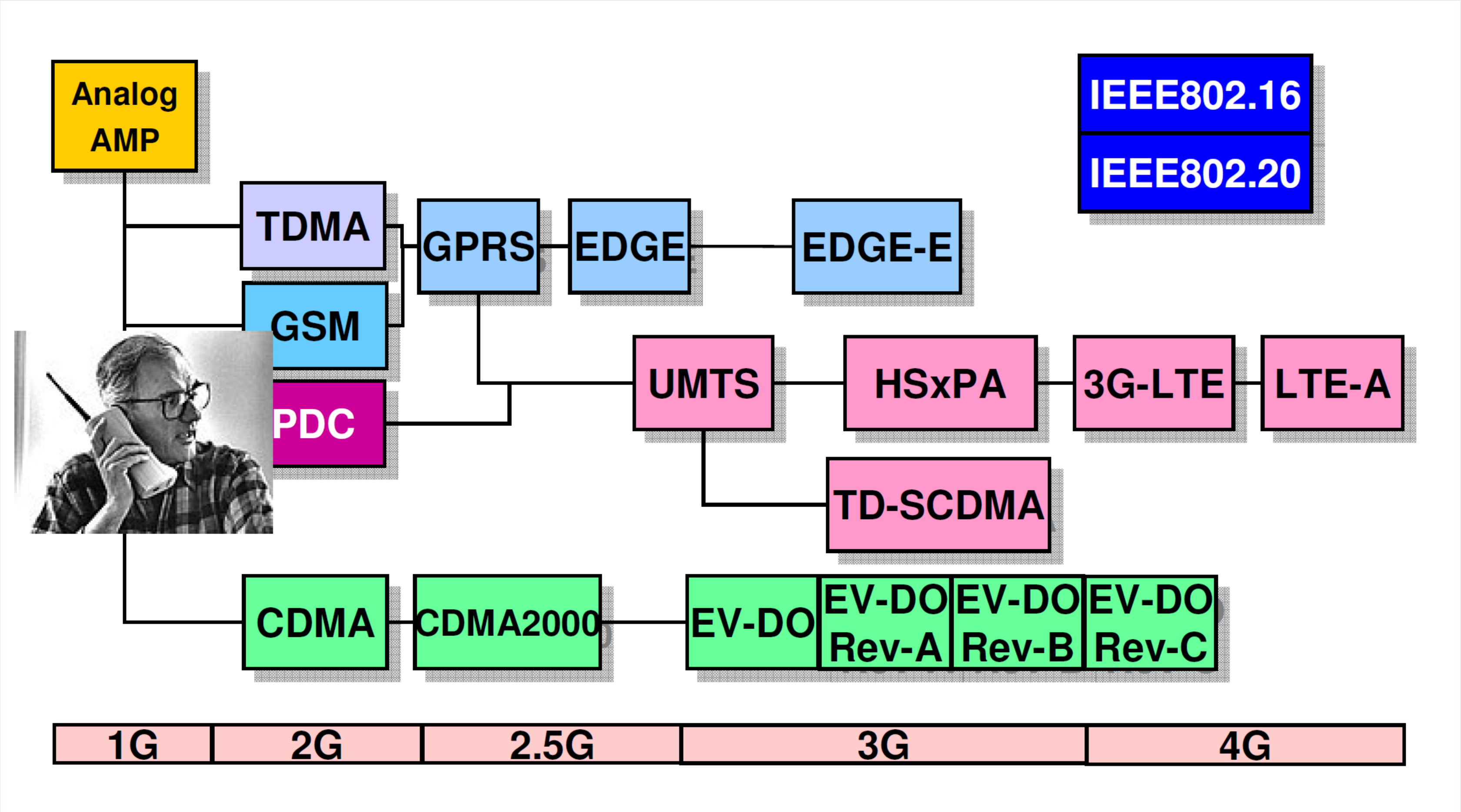
KING'S *College* LONDON

6CCS3COS Communication Systems: Chapter 1

Osvaldo Simeone

Where Are We Now?

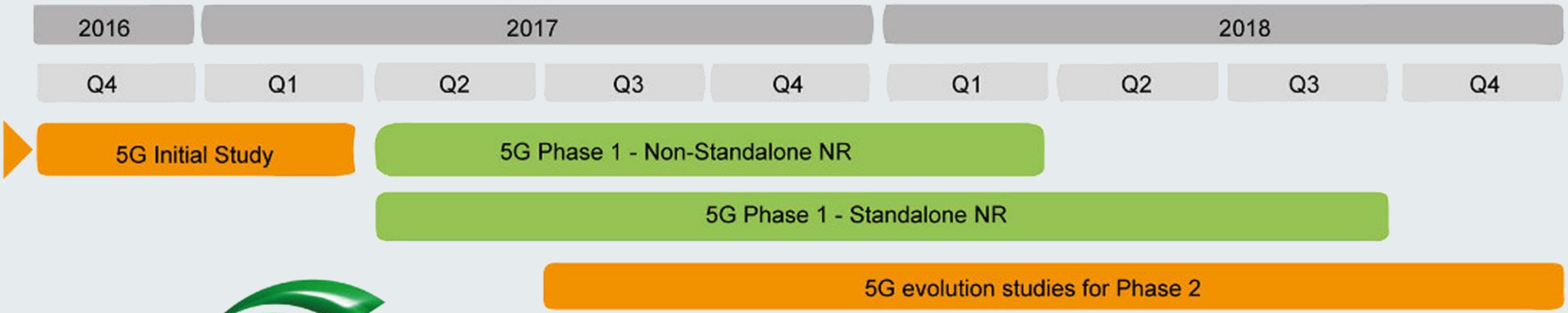
Evolution to 4G



[Neocific]

Where Are We Now?

Work in progress on a 5G standard...



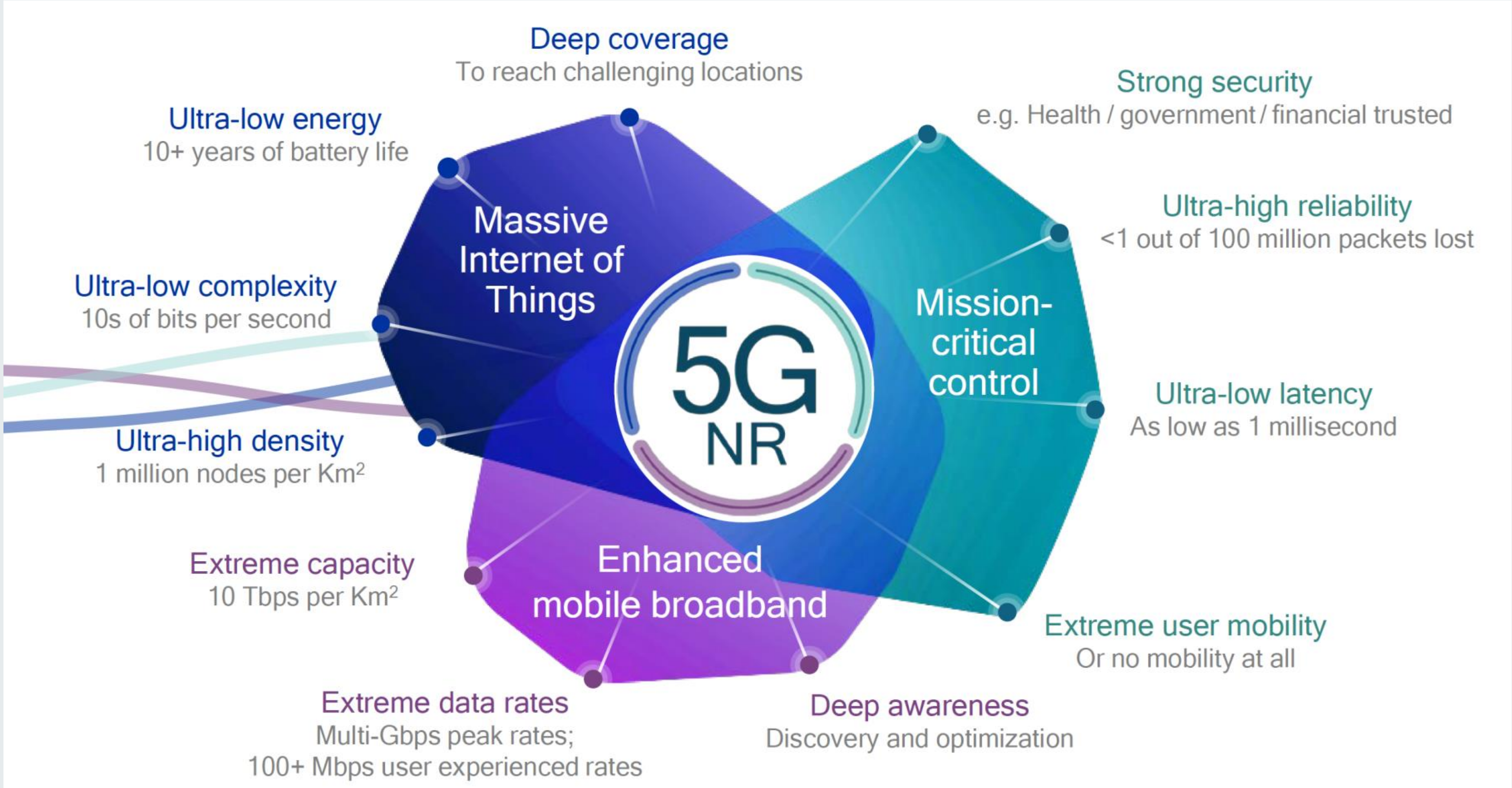
Non-Standalone
Uses LTE core and LTE radio anchor with a 5G small cell
Mobile BroadBand capacity boost

Standalone
Uses 5G core and 5G radio anchor
5G overlay
Expansion of the wireless ecosystem

"5G product supports either 'NG NR'
or 'access via a 5G CN'
or both."
Erik Guttman, 3GPP SA Chairman

Where Are We Now?

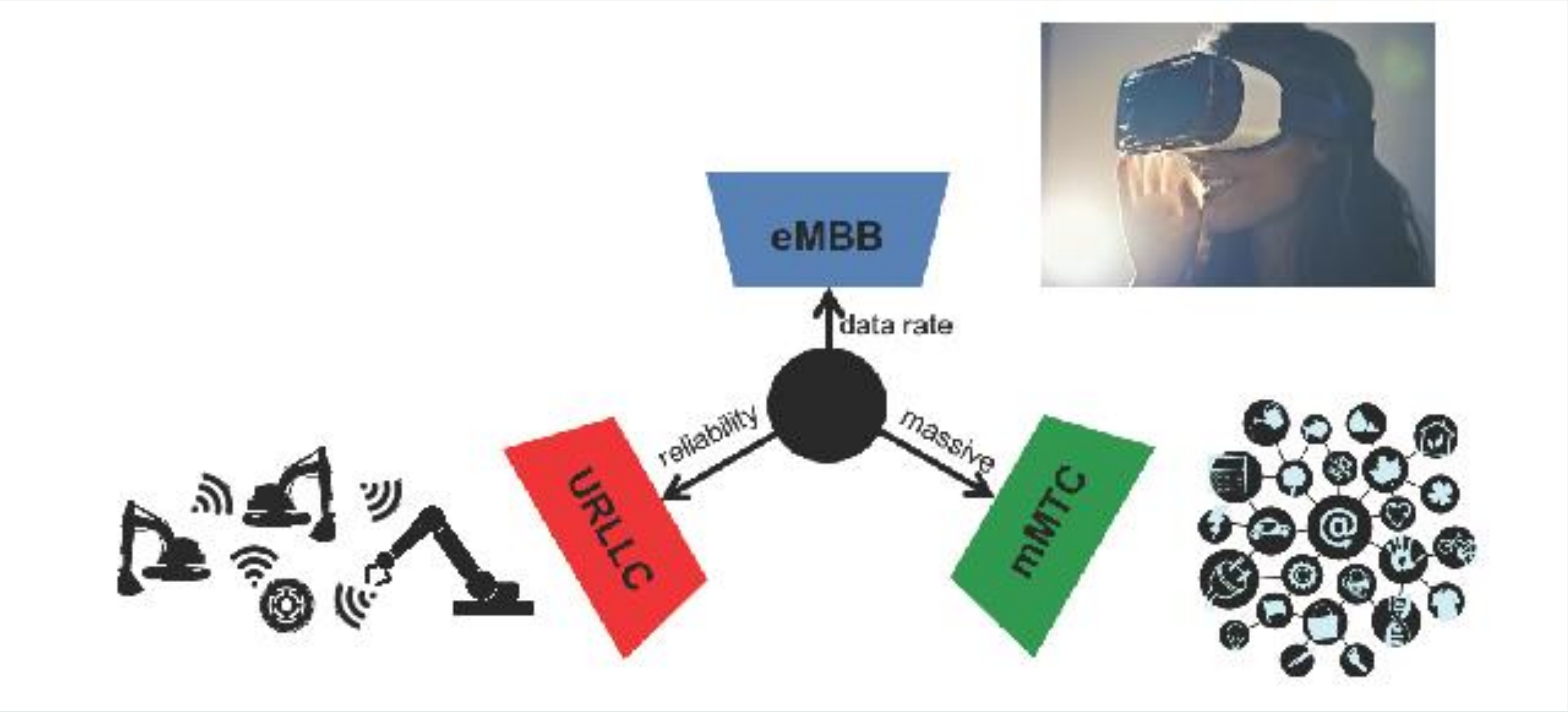
Work in progress on a 5G standard…



[Qualcomm '17]

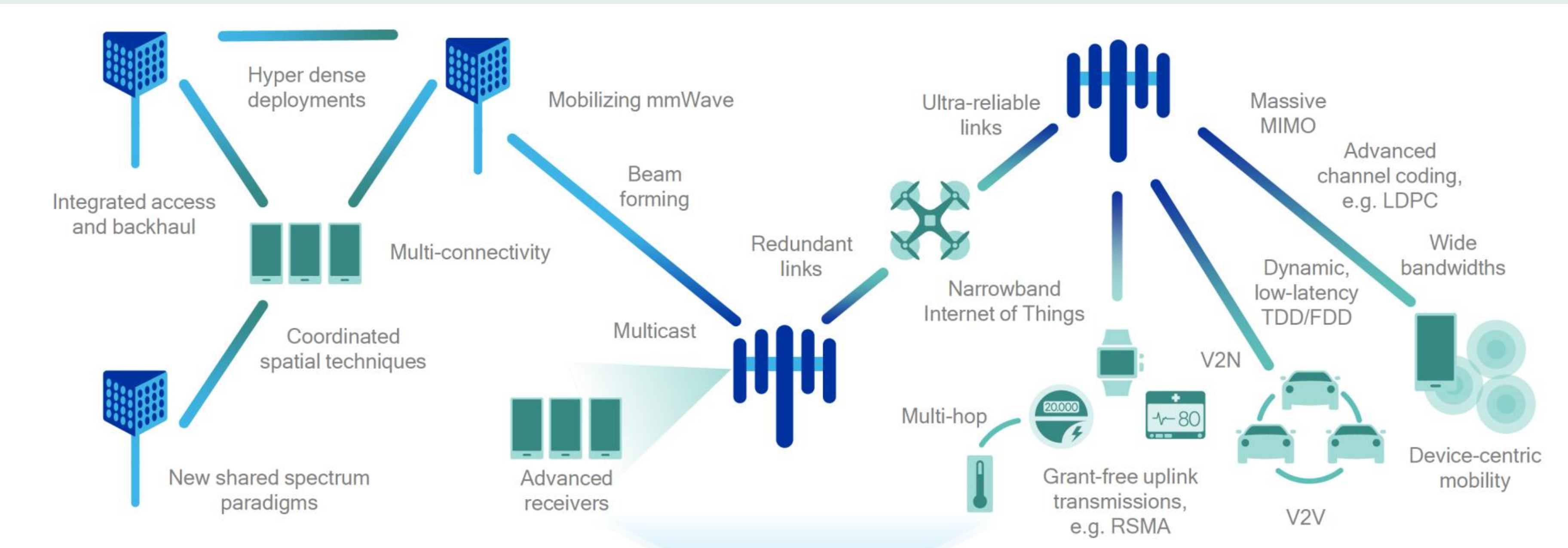
Where Are We Now?

Work in progress on a 5G standard...



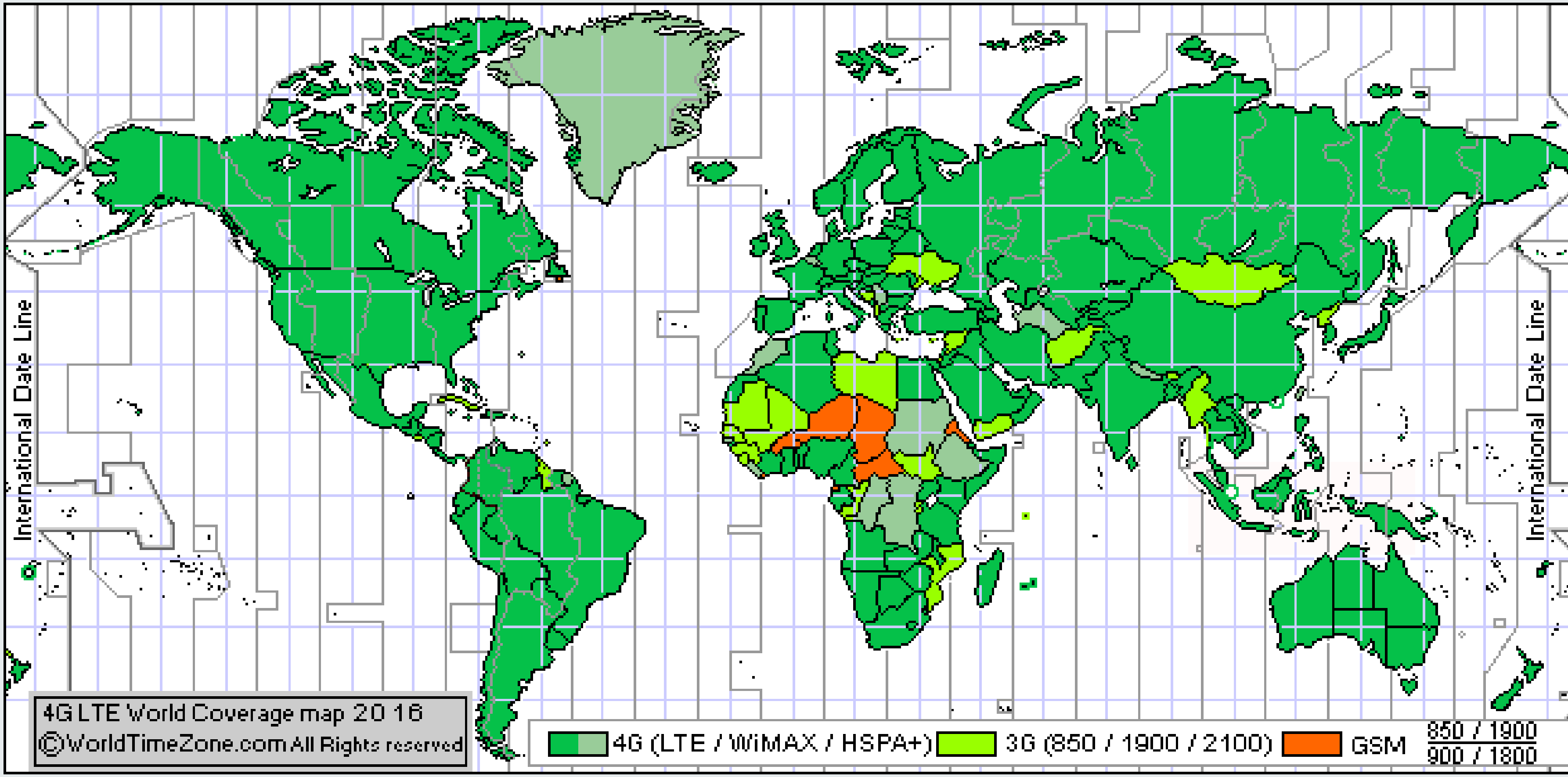
Where Are We Now?

Work in progress on a 5G standard…



[Qualcomm '17]

Where Are We Now?



Where Are We Now?

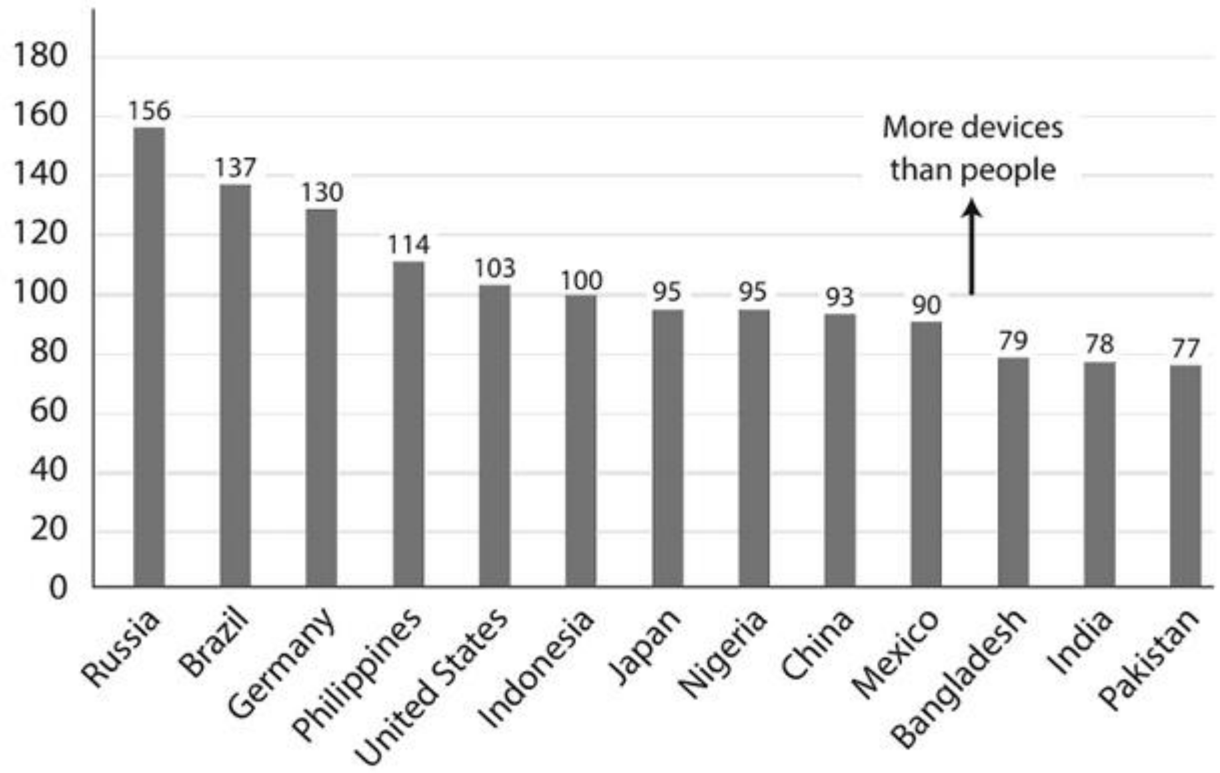


Illustration 1.1 Mobile penetration, or number of cell phones per person, in selected countries as of June 2015. Six countries have a penetration of at least 100%, meaning there are more phones than people.

[Brinton and Chiang '17]

How Did We Get Here?

Quiz: When was the word “bit” invented?

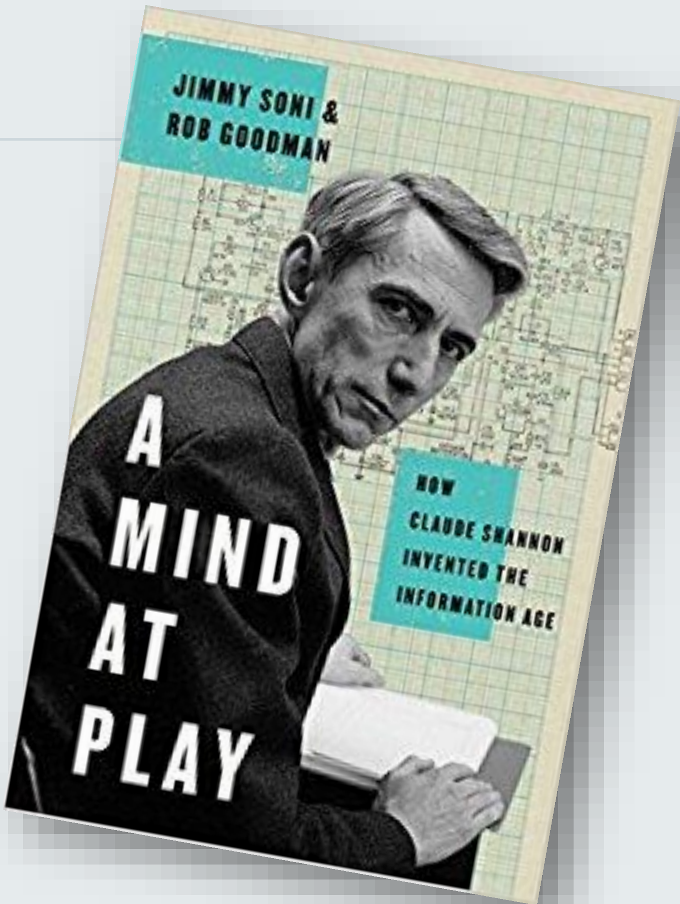
How Did We Get Here?

1948

Reprinted with corrections from *The Bell System Technical Journal*, Vol. 27, pp. 379–423, 623–656, July, October

A Mathematical Theory of Communication

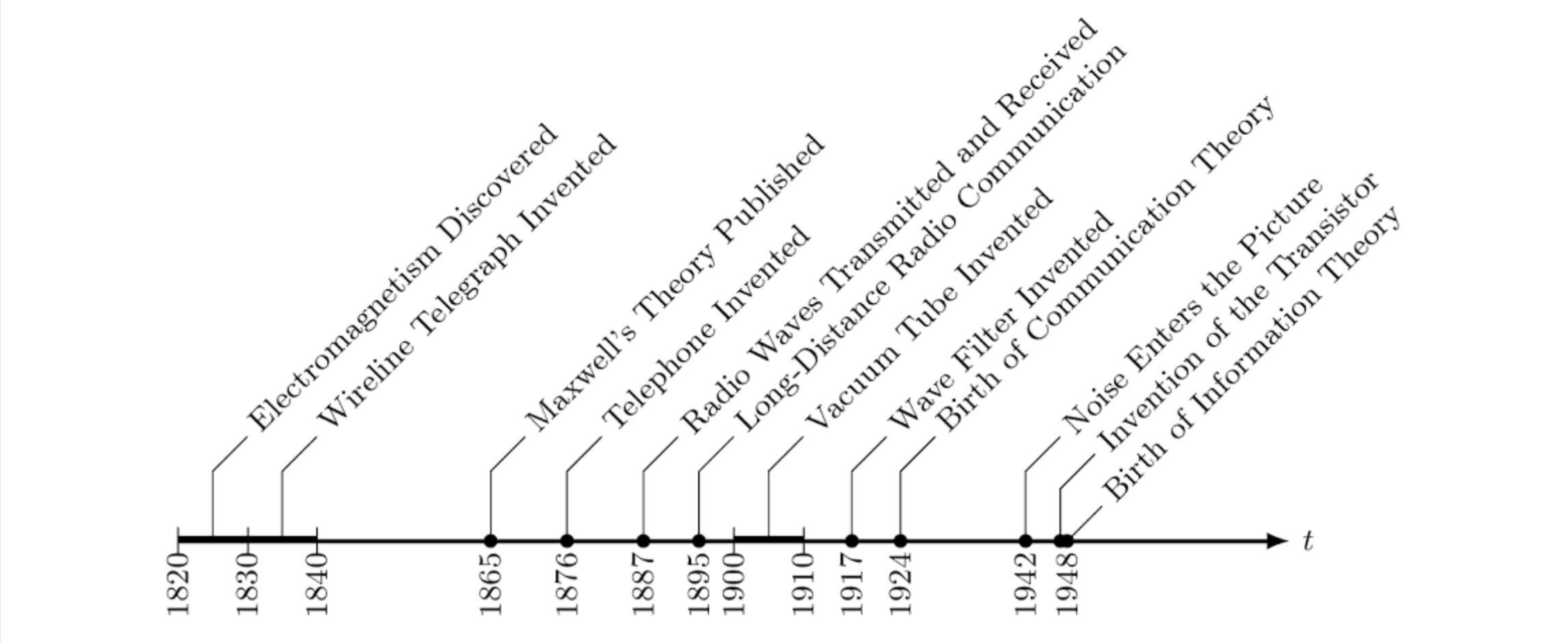
By C. E. SHANNON



The choice of a logarithmic base corresponds to the choice of a unit for measuring information. If the base 2 is used the resulting units may be called binary digits, or more briefly *bits*, a word suggested by J. W. Tukey. A device with two stable positions, such as a relay or a flip-flop circuit, can store one bit of information. N such devices can store N bits, since the total number of possible states is 2^N and $\log_2 2^N = N$.

How Did We Get Here?

Before 1948



[B. Rimoldi, Principles of Digital Communications, Cambridge University Press]

How Did We Get Here?

- **Pre-cellular mobile systems:**
  - 1948: Mobile Telephone Service from Bell Telephone
  - Analog voice
  - FDMA
  - $330 + per-call cost

https://smartphones.gadgethacks.com/news/from-backpack-transceiver-smartphone-visual-history-mobile-phone-0127134/

How Did We Get Here?

- **Pre-cellular mobile systems:**
  - 1956: Mobile System A (MTA) from Ericcson
  - Analog voice
  - FDMA
  - Weighs as much as 300 iPhones!

https://smartphones.gadgethacks.com/news/from-backpack-transceiver-smartphone-visual-history-mobile-phone-0127134/

How Did We Get Here?

- **Pre-cellular mobile systems:**
  - 1964: Improved Mobile Telephone Service (IMTS)
  - Analog voice
  - FDMA

https://smartphones.gadgethacks.com/news/from-backpack-transceiver-smartphone-visual-history-mobile-phone-0127134/

How Did We Get Here?

- **Enter cellular networks**

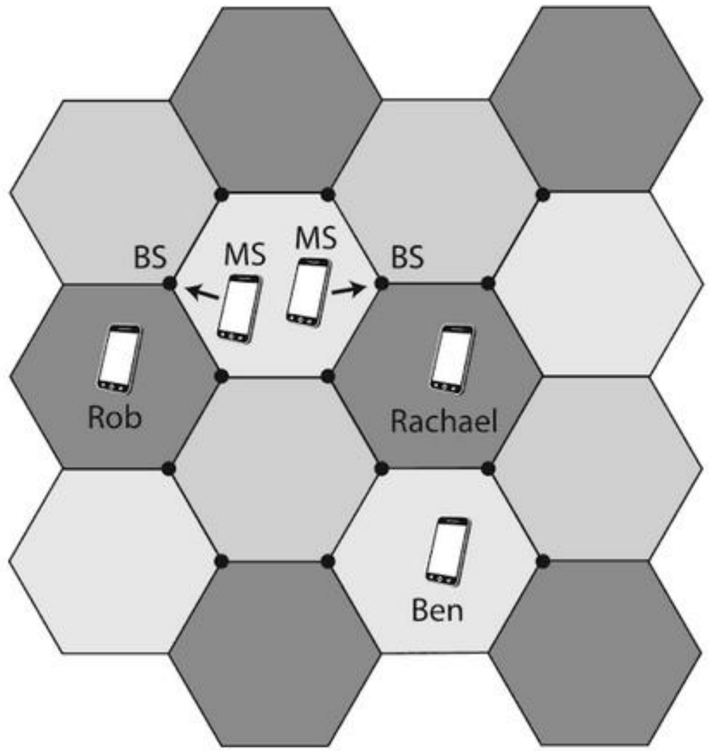
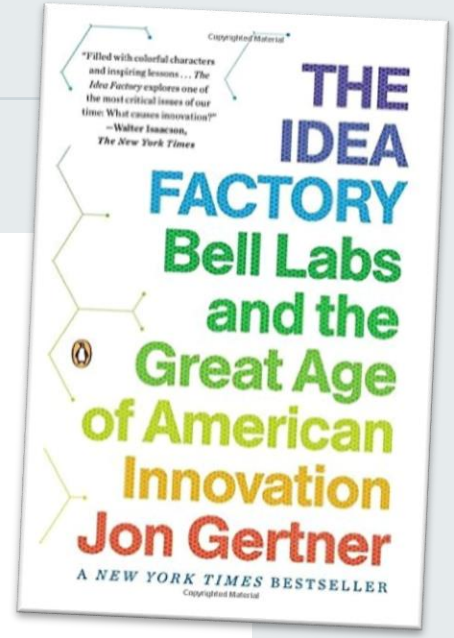


Illustration 1.6 This is a diagram of a cellular network. Each cell is a hexagon, with multiple mobile stations (MSs) and base stations (BSs). The shading of a cell indicates the frequency band that the cell is using. No two neighboring cells have the same shading, and so use different frequency bands to prevent interference.

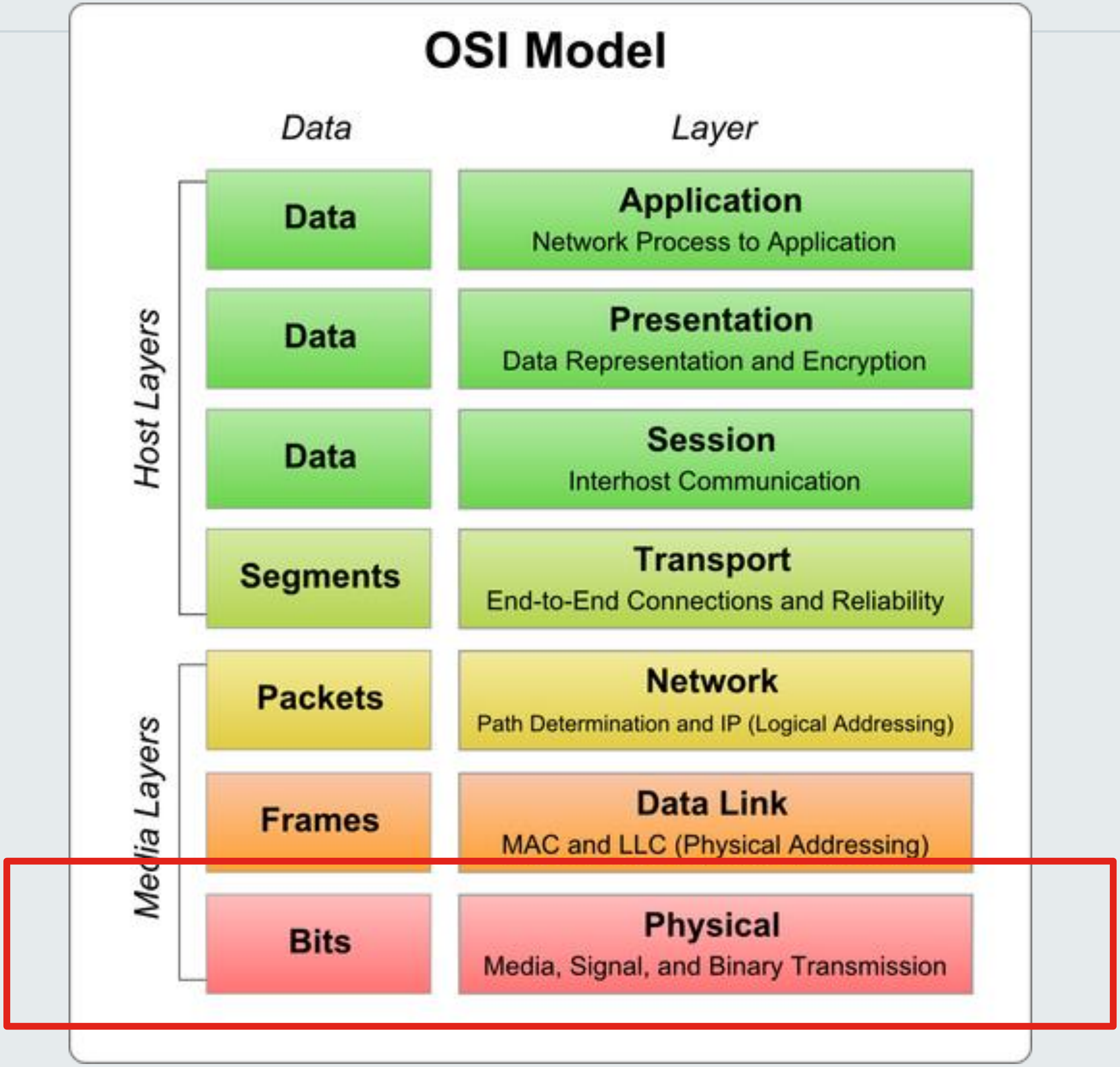


[Brinton and Chiang '17]

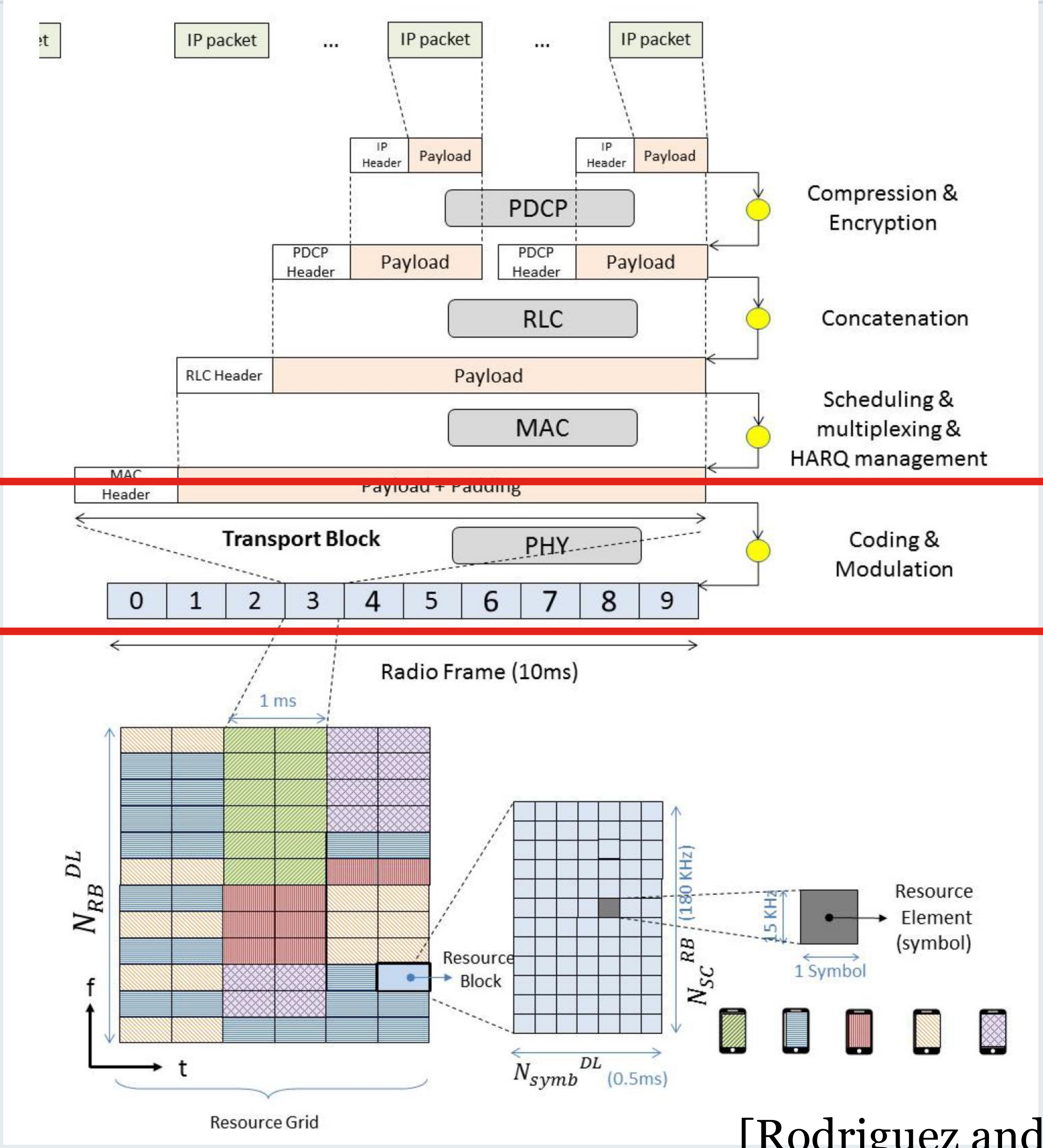
How Did We Get Here?

[Qualcomm ‘17]

What Is This Course About?

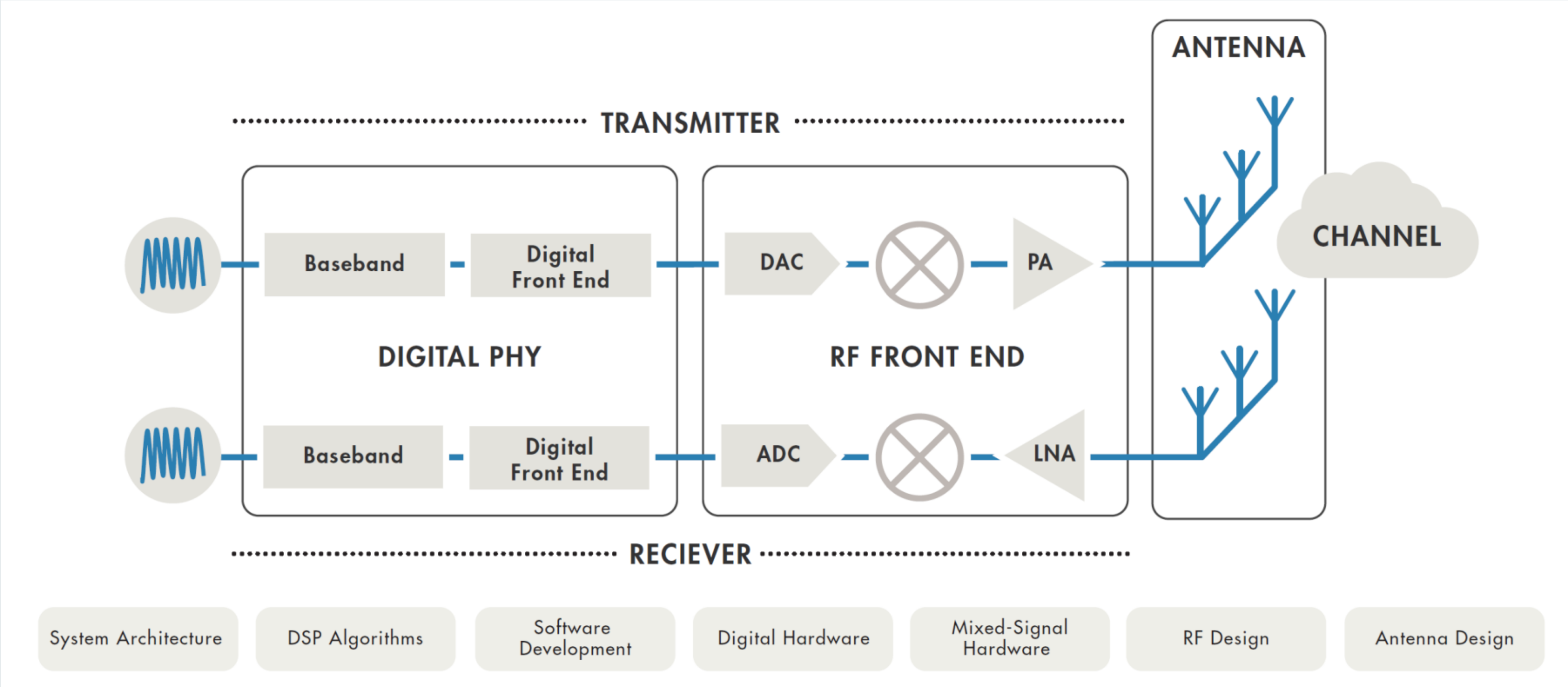


What Is This Course About?



[Rodriguez and Guillemin '17]

What Is This Course About?



[MathWorks]

What Is This Course About?

- **Overview**

  - 1. One-shot digital communications: Fundamentals

  - 2. One-shot digital communications: Passband Systems

  - 3. Stream digital communications

Main reference

- J. Cioffi, Lecture notes, Stanford Univ., Chapters 1, 2, 3

Additional reference

- B. Rimoldi, Principles of Digital Communications, Cambridge University Press

Expectations of inclusive behaviour

The Department of Informatics is committed to providing an inclusive learning and working environment.

Staff and students are expected to behave respectfully to one another – during lectures, outside of lectures and when communicating online or through email.

We won’t tolerate inappropriate or demeaning comments related to gender, gender identity and expression, sexual orientation, disability, physical appearance, race, religion, age, or any other personal characteristic.

If you witness or experience any behaviour you are concerned about, please speak to someone about it. This could be one of your lecturers, your personal tutor, a programme administrator, the Informatics equality & diversity lead (Elizabeth Black), or any other member of staff you feel comfortable talking to.

The College also has a range of different support and reporting procedures that you might find helpful: kcl.ac.uk/harassment

KING'S
College
LONDON

6CCS3COS Communication Systems: Chapter 2

Osvaldo Simeone

What Is This Course About?

- **Overview**

  - 1. One-shot digital communications: Fundamentals

  - 2. One-shot digital communications: Passband Systems

  - 3. Stream digital communications

Main reference

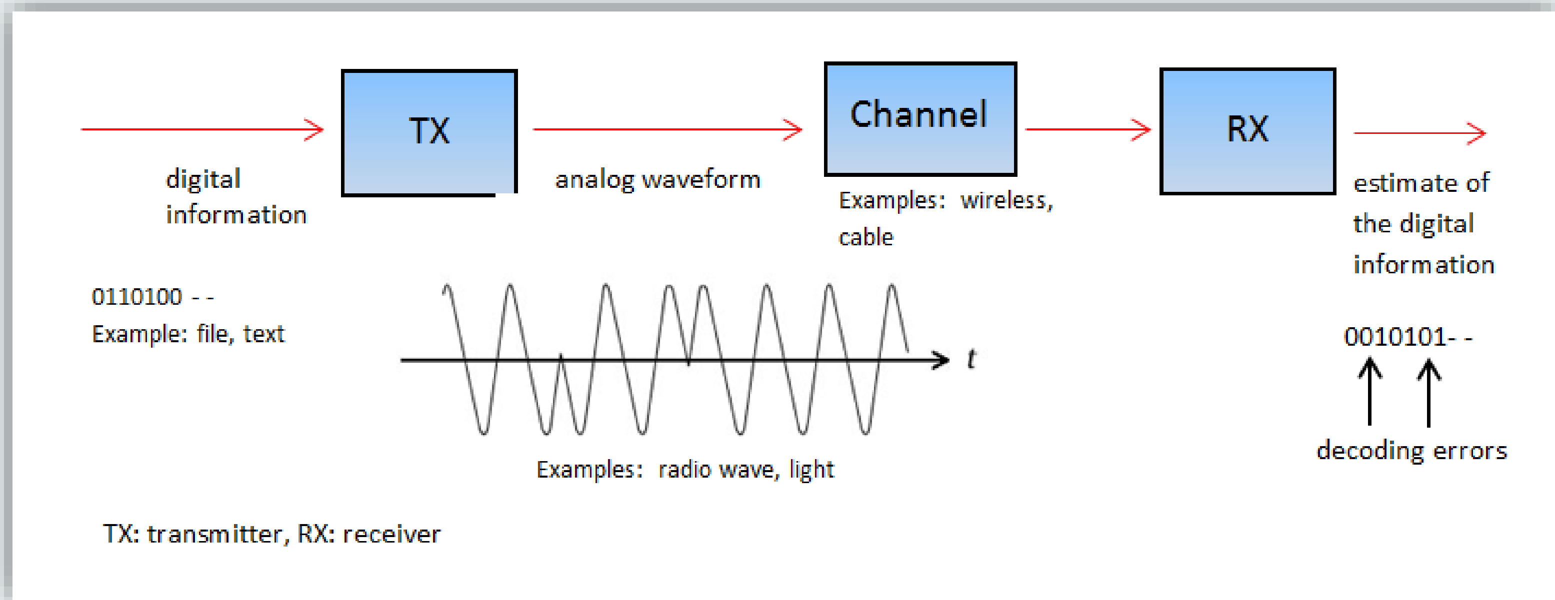
- J. Cioffi, Lecture notes, Stanford Univ., Chapters 1, 2, 3

Additional reference

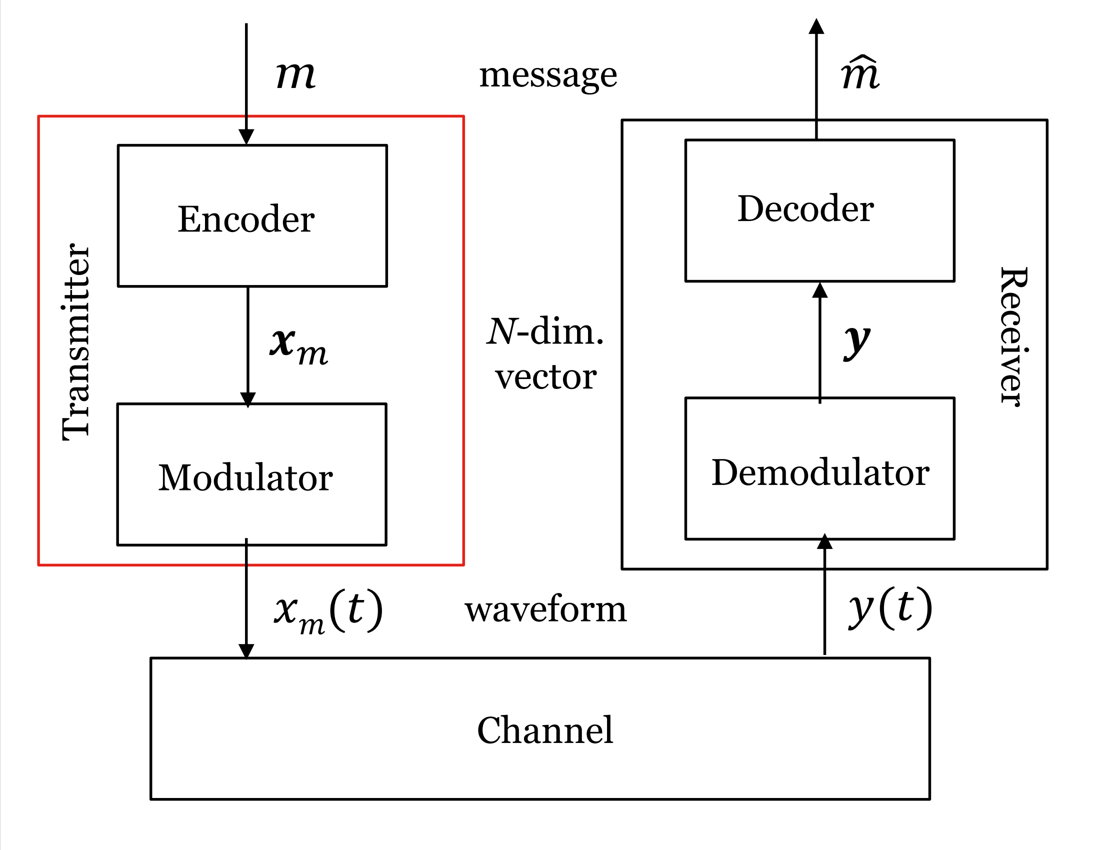
- B. Rimoldi, Principles of Digital Communications, Cambridge University Press

What Is One-Shot Digital Digital Communication?

- The transmitter (TX) selects one message from a finite set (e.g., a bit string) and sends a corresponding signal (or "waveform") through the communication channel via **coding and modulation**.
- The receiver (RX) decides the message sent by observing the channel output via **demodulation and decoding**.
- Optimum detection minimizes the probability of an erroneous receiver decision.



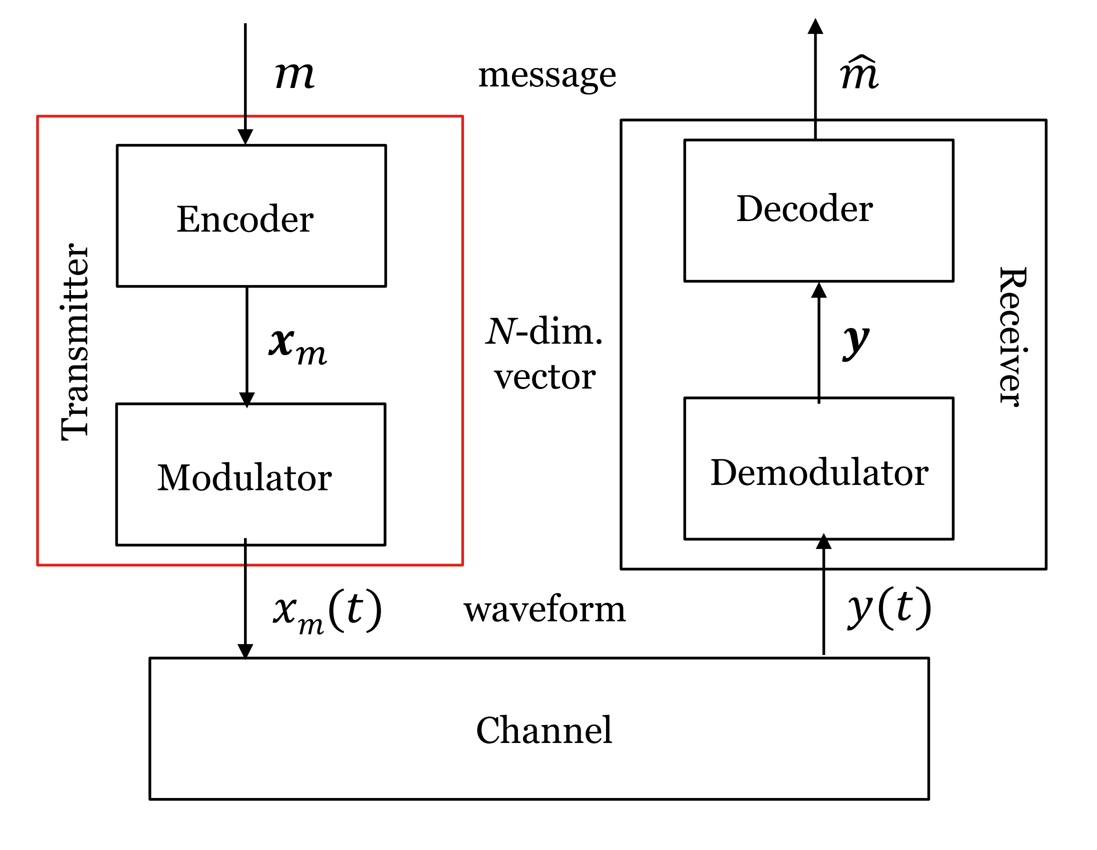
What Does the Transmitter Do?



- **Encoder**: Message $m \in \{0, \dots, M = 2^b - 1\}$ (b bits) → symbol (real vector) $\boldsymbol{x}_m$ in the signal space

- **Modulator**: Symbol $\boldsymbol{x}_m$ → transmitted waveform (analog and continuous-time) $x_m(t)$ of duration T seconds and bandwidth B Hz

What Does the Transmitter Do?



- **Symbol rate:** $1/T$ messages per second
- **Number of bits per symbol:** $b = \log_2 M$
- **Data rate:** $R = b/T$ bits per second

What Does the Transmitter Do?

- A modulator uses a bandwidth B to encode R bit/s
- Spectral efficiency

$$\eta = \frac{R}{B} \text{ (bit/s/Hz)}$$

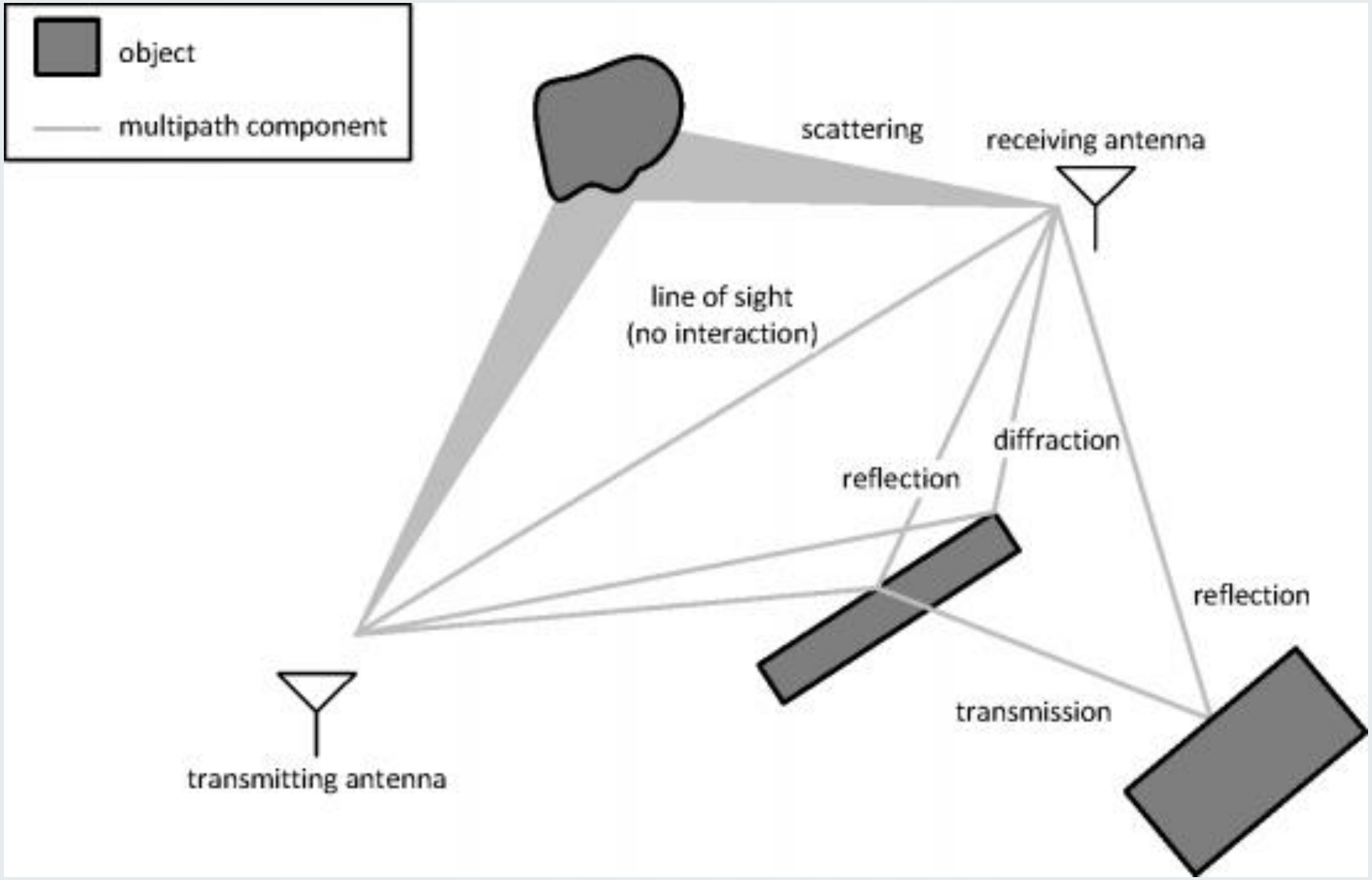
- Ex.: If R=10 Mb/s and B=10 MHz, the spectral efficiency is $\eta = 1$.
- The spectral efficiency ranges from values smaller than 1 (wireless channels with low SNR) to values larger than 10 (e.g., wireless broadcasting).

What Is the Channel?

- Adds noise due to interference and receiver circuitry.
- Attenuates the signal.

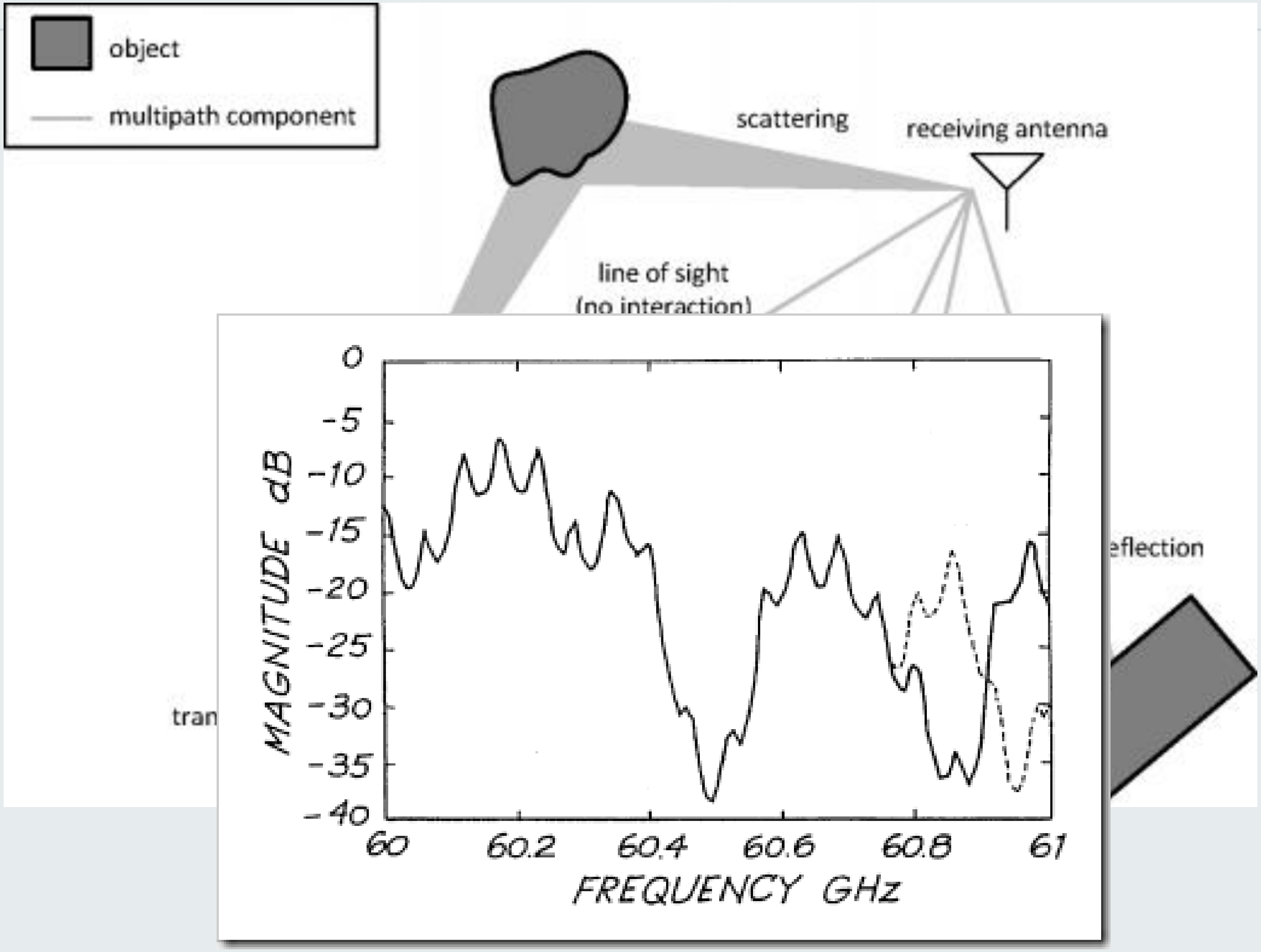
What Is the Channel?

- Adds noise due to interference and receiver circuitry.
- Attenuates the signal.
- Distorts the signal due to multiple propagation paths.



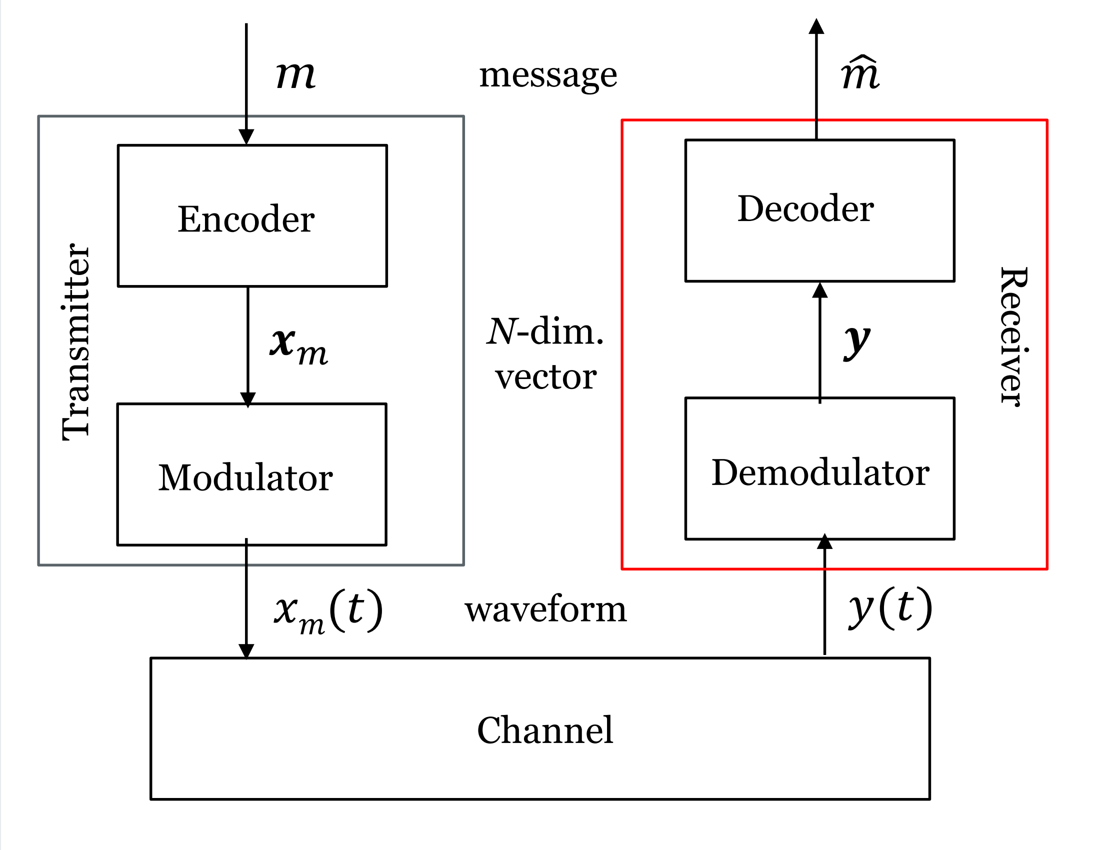
http://www.wica.intec.ugent.be/research/propagation/physical-radio-channel-models

What Is the Channel?



- **Frequency selectivity**: Non-uniform frequency response
- Time selectivity: Time-variability

What Does the Receiver Do?



- **Demodulator**: Channel output waveform (analog and continuous-time) $y(t)$ → channel output vector $\mathbf{y}$ in the signal space

- **Decoder:** Channel output vector $\mathbf{y}$ → estimate $\hat{m}$ of the message m (b bits)

How Much Power to Transmit?

- A basic question in the design of transmitter and receiver is: how much power should the transmitter use in order to ensure a given **reliability (probability of correct detection)**?

- This is known as **link budget**.

- The starting point is that the receiver needs to be guaranteed a certain signal-to-noise ratio (SNR)

$$\text{SNR} = \frac{\text{received power (mW)}}{\text{noise power (mW)}} = \frac{P_r \text{ (mW)}}{\sigma^2 \text{ (mW)}}$$

despite the attenuation caused by the channel.

How Much Power to Transmit?

- If the channel gain is $L \leq 1$, then the received power is

$$P_r = L \times P_x$$

where P_x is the transmitted power.

- Therefore, the required transmitted power is given as

$$\text{SNR} = \frac{P_r}{\sigma^2} = \frac{L \times P_x}{\sigma^2}$$

$$\rightarrow \boxed{P_x = \sigma^2 \frac{\text{SNR}}{L}}$$

How Much Power to Transmit?

- Ex.: If SNR = 100, L=10^{-8}, and $\sigma^2 = 10^{-11}$ mW, then the required power is

$$P_x = 10^{-11} \times 100 \times 10^{8} = 0.1 \text{ mW}$$

- These are realistic values and demonstrate the need to deal with large numbers (e.g., 10^{8}) and very small numbers (e.g., 10^{-11}).

- To this end, communication engineers work with decibel (dB) measures.

How Much Power to Transmit?

- The decibel (symbol: dB) is a unit of measurement used to express a ratio of powers on a logarithmic scale:

$$\text{power ratio (dB)} = 10\,\log_{10}(\text{power ratio})$$

dB	Power ratio
100	10 000 000 000
90	1 000 000 000
80	100 000 000
70	10 000 000
60	1 000 000
50	100 000
40	10 000
30	1 000
20	100
10	10
6	3 .981 ≈ 4
3	1 .995 ≈ 2
1	1 .259
0	1
0	1
−1	0 .794
−3	0 .501 ≈ ½
−6	0 .251 ≈ ¼
−10	0 .1
−20	0 .01
−30	0 .001
−40	0 .000 1
−50	0 .000 01
−60	0 .000 001
−70	0 .000 000 1
−80	0 .000 000 01
−90	0 .000 000 001
−100	0 .000 000 000 1

How Much Power to Transmit?

- Link budget calculations can be carried out fully in the logarithmic scale.

- Key properties: $\log(a \times b) = \log(a)+\log(b)$
  $\log(a/b) = \log(a)-\log(b)$

(note that $\log(b^c)=c\log(b)$ and hence the second property follows from the first)

How Much Power to Transmit?

- Link budget calculations can be carried out fully in the logarithmic scale.

- Key properties: $\log(a \times b) = \log(a)+\log(b)$
  $\log(a/b) = \log(a)\text{-}\log(b)$

  (note that $\log(b^c)$=c$\log(b)$ and hence the second property follows from the first)

- To this end, powers are measured in dBm:

$$\text{Power (dBm)} = 10\log_{10}\left(\frac{\text{power (mW)}}{1\text{ mW}}\right)$$

and we have

$$\boxed{P_x\ (\text{dBm}) = \sigma^2\ (\text{dBm}) + \text{SNR}\ (dB)\ - L\ (dB)}$$

How Much Power to Transmit?

- Ex.: For the same example above, working in the logarithmic scale, we have:

- If SNR = 100 = 20 dB,

  $L = 10^{-8} = -80$ dB,

  and $\sigma^2 = 10^{-11}$ mW = -110 dBm, then the required power is

$$P_x\ (\text{dBm}) = \sigma^2\ (\text{dBm}) + \text{SNR}\ (dB) - L\ (dB)$$

$$= -110 + 20 + 80 = -10\ \text{dBm}$$

How Much Power to Transmit?

- In wireless channels, the attenuation is given as a function of the distance d (m) between transmitter and receiver as

$$L\ (\text{dB}) = L_1\ (\text{dB}) - \gamma\ 10\log_{10}(d)$$

where L_1 is the attenuation at 1 m and γ is the path loss exponent (between 2 and 5 depending on the environment).

How Much Power to Transmit?

- In wireless channels, the attenuation is given as a function of the distance d (m) between transmitter and receiver as

$$L\ (\text{dB}) = L_1\ (\text{dB}) - \gamma\ 10\log_{10}(d)$$

where L_1 is the attenuation at 1 m and γ is the path loss exponent (between 2 and 5 depending on the environment).

- In wired channels, e.g., fiber optics cables, the attenuation can be typically written as

$$L\ (\text{dB}) = L_0\ \left(\frac{dB}{km}\right) \times\ d$$

where L_o is the attenuation in dB per km (e.g., -0.1 dB/km for fiber).

How Much Power to Transmit?

- Ex.: Indoor WLAN (e.g., Wi-Fi)

  Assuming γ=2 and L_1=-50 dB, and $\sigma^2 = -110\ \mathrm{dBm}$, what is the maximum distance at which can we ensure an SNR of 10 dB if the transmitted power is 0 dBm?

How Much Power to Transmit?

- Ex.: Indoor WLAN (e.g., Wi-Fi)

Assuming γ=2 and L_1=-50 dB, and $\sigma^2 = -110$ dBm, what is the maximum distance at which can we ensure an SNR of 10 dB if the transmitted power is 0 dBm?

In order to ensure SNR = 10 dB, the channel gain must be at least

$$P_x\ (\text{dBm}) = \sigma^2\ (\text{dBm}) + \text{SNR}\ (dB) - L\ (dB)$$

$$0 = -110 + 10 - L\ (dB)$$

$$\rightarrow \quad L\ (dB) = -100\ \text{dB}$$

and hence the maximum distance satisfies $-50 - 20\log_{10}(d) = -100$

$$\log_{10}(d) = 50/20 = 2.5 \quad \rightarrow d = 10^{2.5} = 316\ \text{m}$$

How Much Power to Transmit?

- Ex.: Fiber optics cable

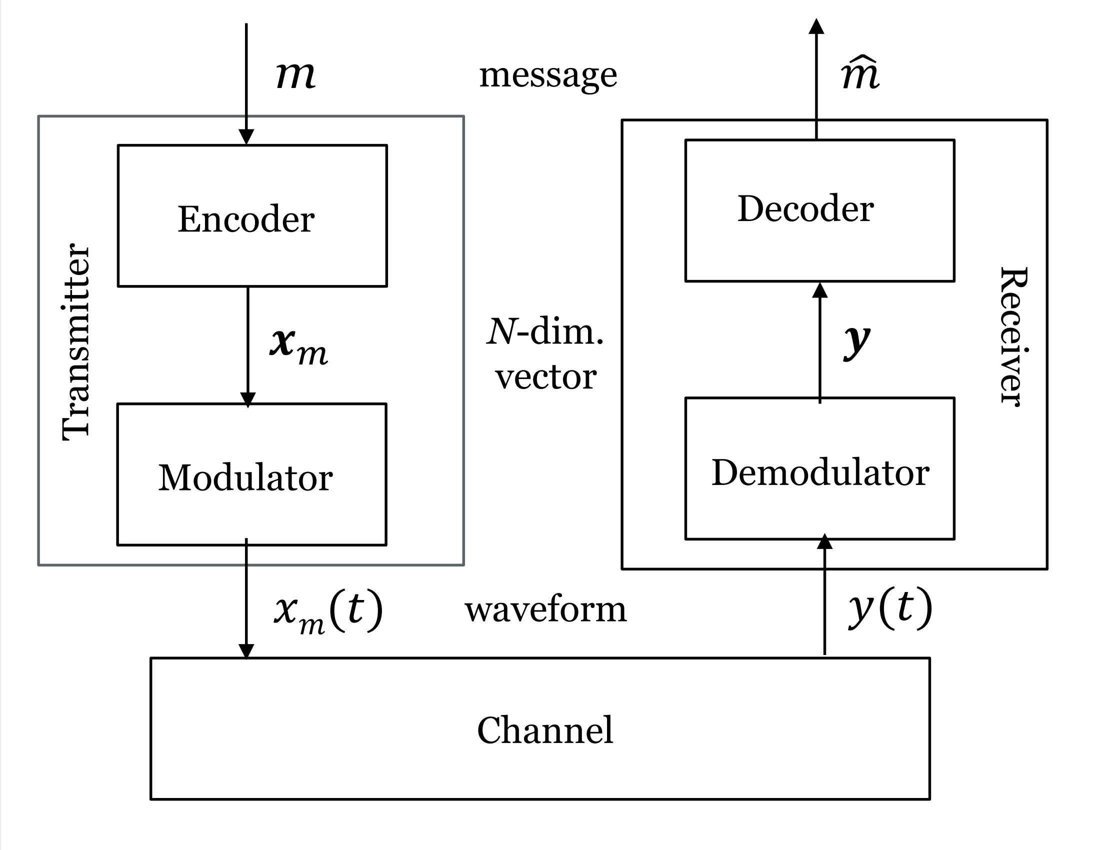
Assuming L_0=-1.5 dB/km, and $\sigma^2 = -110$ dBm, what is the maximum distance at which can we ensure an SNR of 10 dB if the transmitted power is 0 dBm?

In order to ensure SNR = 10 dB, the channel gain must be at least $L\ (dB)$ = -100 dB, and hence the minimum distance is

$$-1.5d = -100$$

$$\rightarrow d = 100/1.5 = 67 \text{ km}$$

How to Encode and Decode Information?



- Why is the performance dependent on the SNR?
- How much information can be transferred for a given SNR?
- How is the information encoded and decoded?

How to Encode?

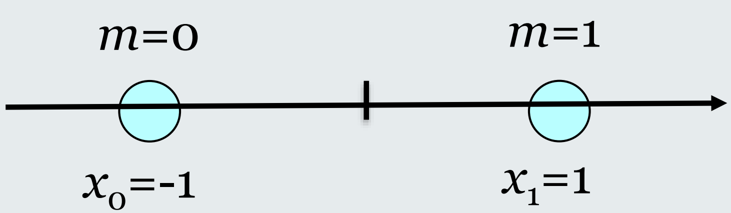
- **Encoder**:

$$\text{message } m \in \{0, \ldots, M-1\} \rightarrow \text{symbol (real vector) } \boldsymbol{x}_m = \begin{bmatrix} x_{m,1} \\ x_{m,2} \\ \vdots \\ x_{m,N} \end{bmatrix}$$

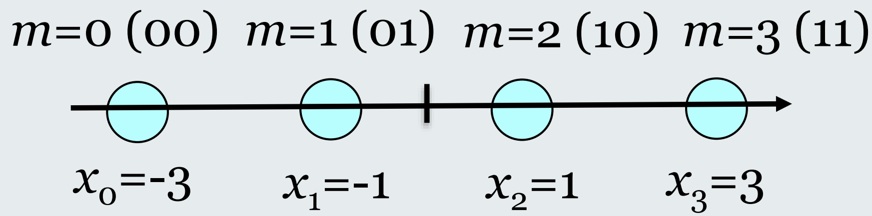
- The signal space is R^N
- The set of all symbols $\boldsymbol{x}_m$, m=1,...,M, is the signal **constellation**
- We expect that a constellation with symbols that are further apart will lead to a smaller probability of error.

How to Encode?

Examples: N=1



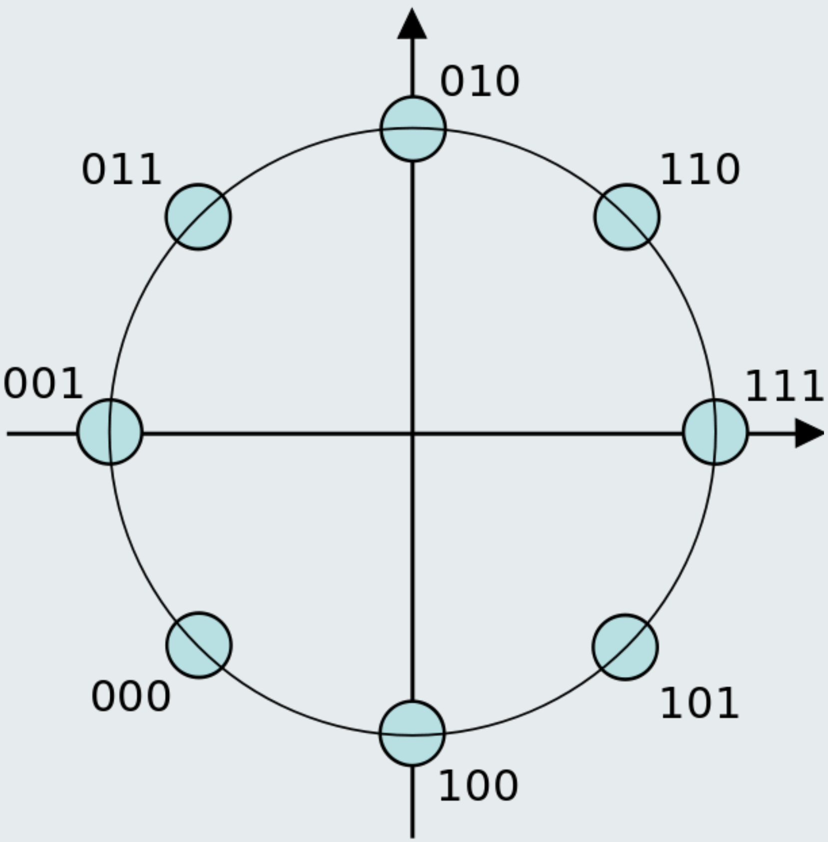
M=2 (b=1)
Binary Phase Shift Keying (BPSK)



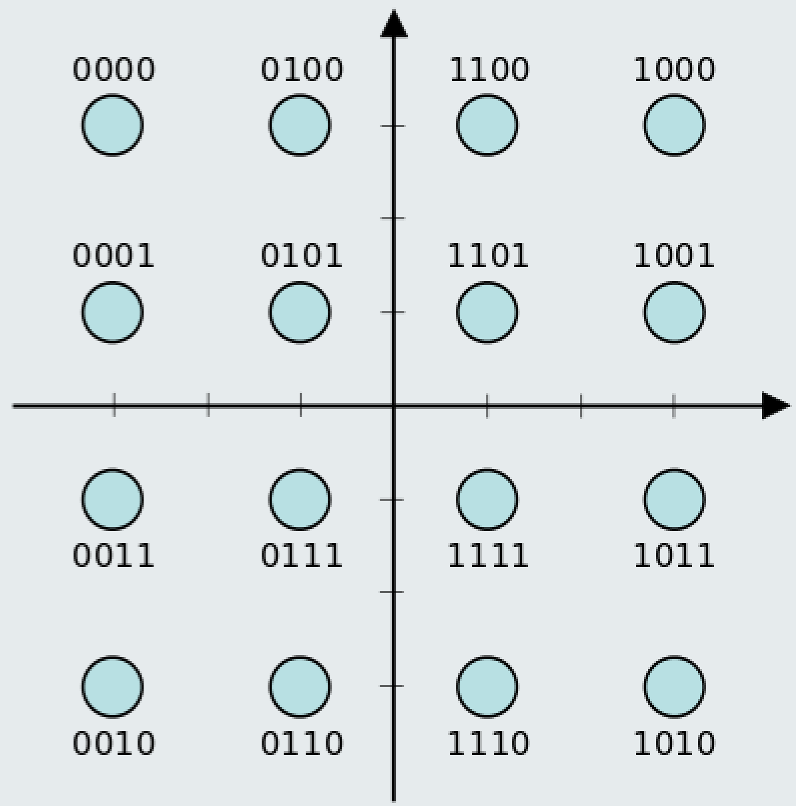
M=4 (b=2)
4-Pulse Amplitude Modulation (PAM)

How to Encode?

Examples: N=2



M=8 (b=3)
8-Phase Shift Keying (PSK)

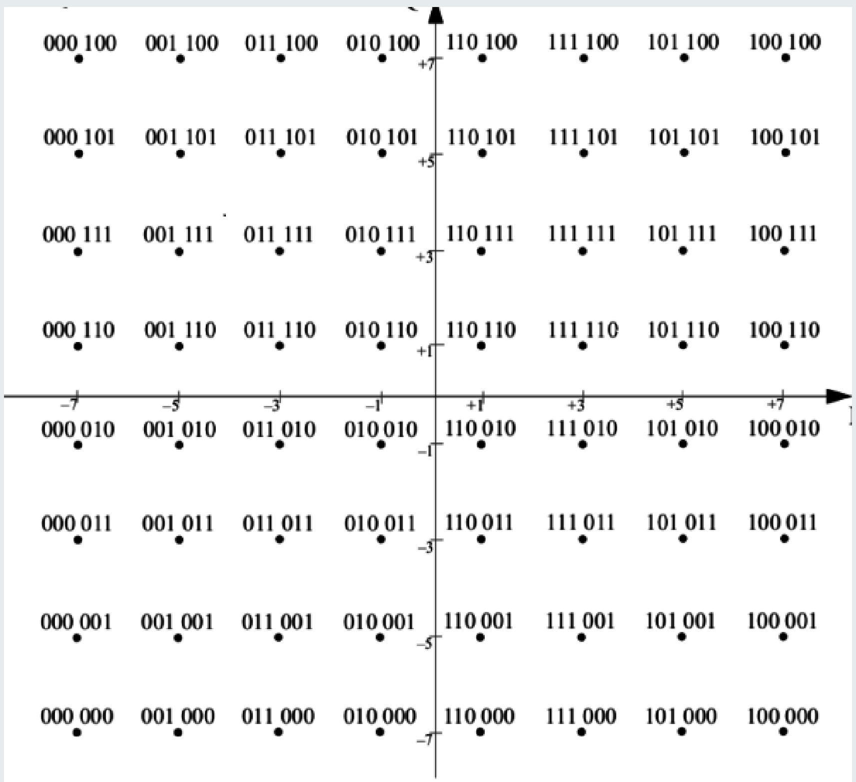


M=16 (b=4)
16-Quadrature Amplitude Modulation (QAM)

https://en.wikipedia.org/wiki/Constellation_diagram

How to Encode?

Examples: N=2



M=64 (b=6)
64-QAM

M=256 (b=8)
256-QAM

How to Encode?

- **Average energy** of a constellation: If each message m is selected with probability p_m, the average energy of a constellation $\{\boldsymbol{x}_m\}$ is defined as

$$E_x = \mathrm{E}[||\mathbf{x}||^2] = \sum_{m=0}^{M-1} p_m ||\mathbf{x}_m||^2$$

where

$$||\mathbf{x}_m||^2 = \sum_{n=1}^{N} |x_{m,n}|^2$$

is the energy of the mth symbol

- **Average power**:

$$P_x = \frac{E_x}{T}$$

How to Encode?

- When M=2 (or b=1), the energy per bit is

$$E_b = E_x$$

- More generally, we have

$$E_x = bE_b$$

- E_b is an important metric, since it relates cost (energy) to performance (bit rate).

- Note: E_x is also often referred to as the symbol energy (and denoted as E_s).

What Happens When the Energy Increases/ Decreases?

- Quiz: BPSK energy

  If N=1, M=2, the symbols are equally likely and the average energy is Eb<1, what are valid constellations?

  A) -1,+1
  B) $-\sqrt{E_b}, +\sqrt{E_b}$
  C) $0, +\sqrt{2E_b}$
  D) $-2\sqrt{E_b}, +2\sqrt{E_b},$

What Happens When the Energy Increases/ Decreases?

- Quiz: BPSK energy

  Considering that the distance between constellation points determines the probability of error, which constellation would you choose?

  A) $-\sqrt{E_b}, +\sqrt{E_b},$
  B) $0, +\sqrt{2E_b}$

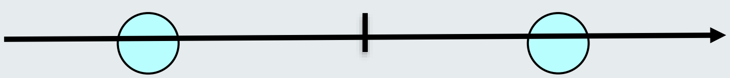
What Happens When the Energy Increases/ Decreases?

- If all symbols have the same probability, we have the condition

$$\frac{1}{M}\sum_{m=0}^{M-1}||\boldsymbol{x}_m||^2 = bE_b$$

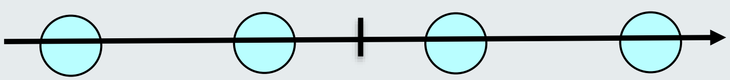
- **Examples:**

BPSK



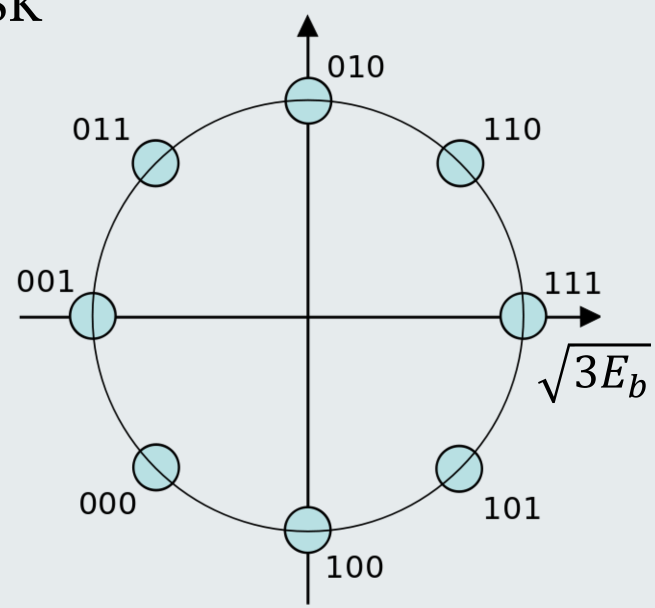
x_0=-$\sqrt{E_b}$ x_1=$\sqrt{E_b}$

4-PAM

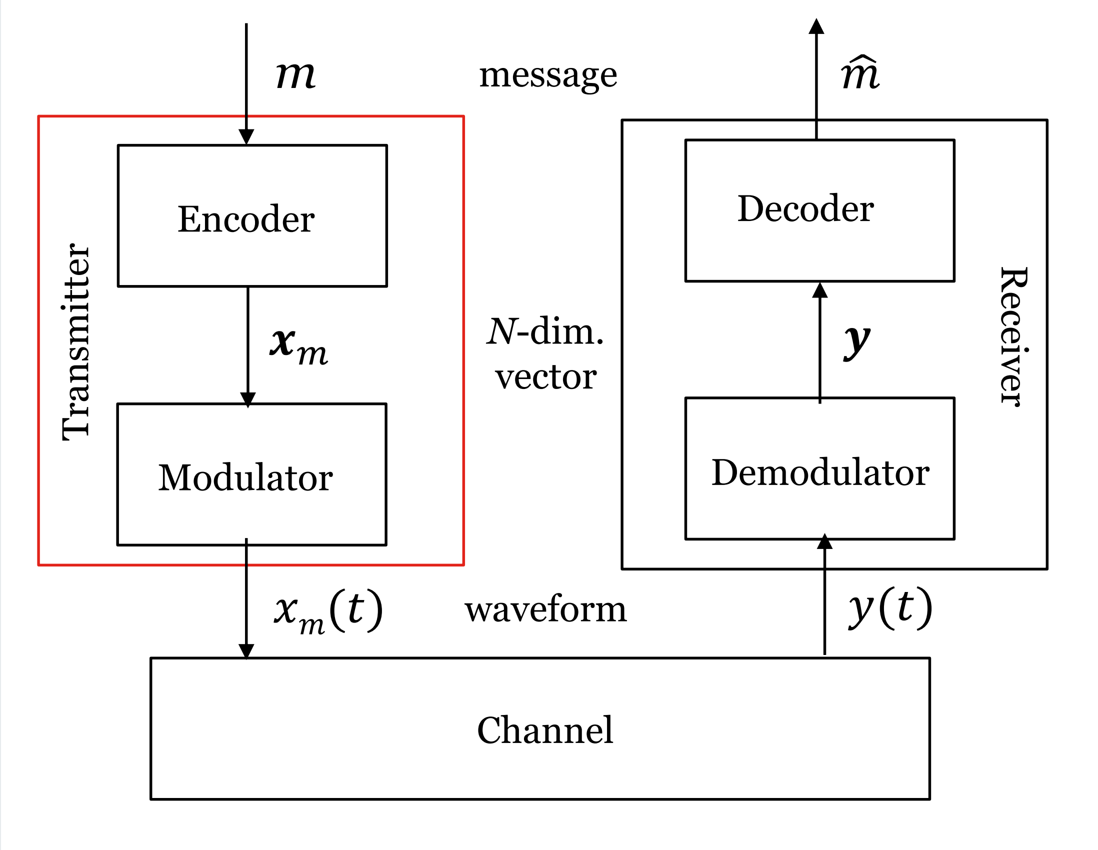


x_0=-3$\sqrt{\frac{2}{5}E_b}$ x_1=-$\sqrt{\frac{2}{5}E_b}$ x_2=$\sqrt{\frac{2}{5}E_b}$ x_3=3$\sqrt{\frac{2}{5}E_b}$

8-PSK



What Does the Transmitter Do?



- **Encoder**: Message $m \in \{0, \dots, M = 2^b - 1\}$ (b bits) → symbol (real vector) $\boldsymbol{x}_m$ in the signal space

- **Modulator**: Symbol $\boldsymbol{x}_m$ → transmitted waveform (analog and continuous-time) $x_m(t)$ of duration T seconds and bandwidth B Hz

Why Modulation?

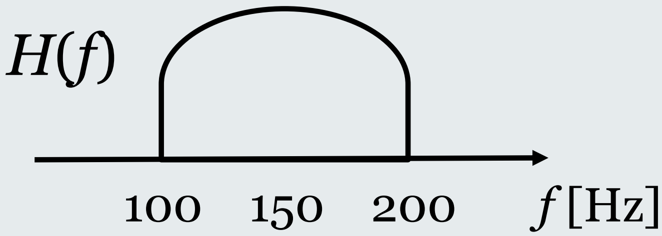
- To adapt transmitted signals to the channel

- Note that the probability of error depends on the distance between symbols in the constellation *after* the distortion caused by the channel (to be discussed).

How Does the Modulation Adapt to the Channel?

Example:

- Channel frequency response: passband in the interval 100 Hz and 200 Hz with 150 Hz having the largest gain



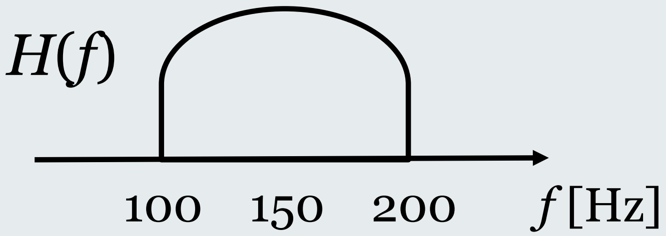
- b=1, or equivalently M=2 (binary transmission)
- Encoder: $0 \rightarrow x_0 = -1$ and $1 \rightarrow x_1 = 1$
- Modulator: $x_0 \rightarrow x_0(t) = -1$ and $x_1 \rightarrow x_1(t) = 1$ for all t

- Quiz: Is this a good choice for the modulator?

How Does the Modulation Adapt to the Channel?

Example:

- Channel frequency response: passband in the interval 100 Hz and 200 Hz with 150 Hz having the largest gain



- Quiz: Which waveforms are suitable to be chosen by the modulator for this channel?

A) $x_0(t)$=-cos(2*pi*t) and $x_1(t)$=cos(2*pi*t)

B) $x_0(t)$=-cos(300*pi*t) and x_1 (t)=cos(300*pi*t)

C) $x_0(t)$=-cos(100*pi*t) and x_1 (t)=cos(100*pi*t)

D) $x_0(t)$=-cos(240*pi*t) and x_1 (t)=cos(240*pi*t)

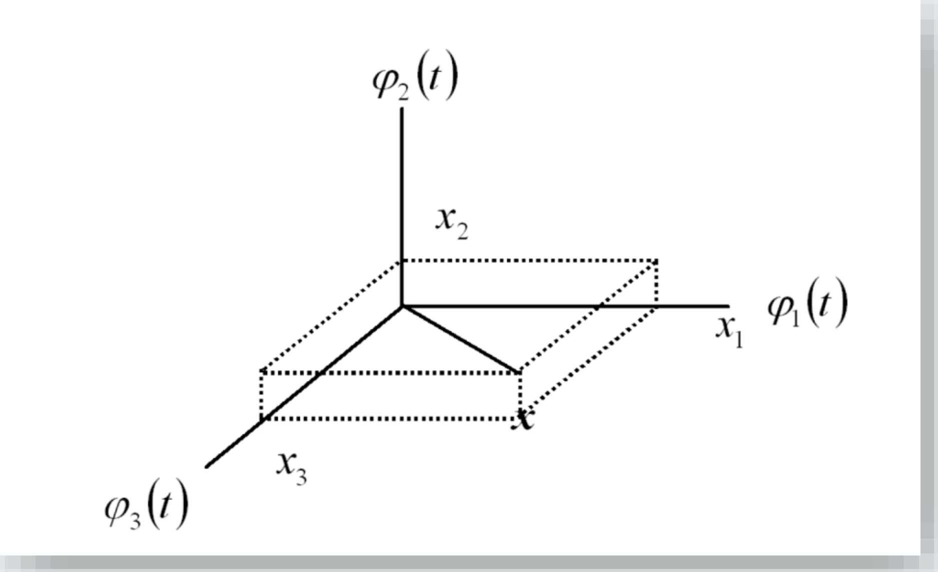
How to Modulate?

- **Modulator**:

  symbol $\boldsymbol{x}_m \rightarrow$ (analog and continuous-time) waveform $x_m(t)$ of duration T seconds (symbol period)

- The corresponding set of modulated waveforms $\{x_m(t)\}$, $m = 0, \dots, M - 1$, is a **signal set**.

- To map symbol to waveform, we need to associate each axis of the signal space R^N with a basis function in a set of **orthonormal basis functions**.



Some Math

- **Inner product or correlation:**

  - between real vectors

$$\langle \mathbf{v}, \mathbf{u} \rangle = \sum_{n=1}^{N} v_n u_n$$

  - between real functions

$$\langle v(t), u(t) \rangle = \int_{-\infty}^{\infty} v(t) u(t) dt$$

Some Math

- **Squared Euclidean norm or energy:**

  - for a real vector

$$||\mathbf{v}||^2 = \langle \mathbf{v}, \mathbf{v} \rangle = \sum_{n=1}^{N} v_n^2$$

  - for a real function

$$||v(t)||^2 = \langle v(t), v(t) \rangle = \int_{-\infty}^{\infty} (v(t))^2 dt$$

Some Math

- **Orthogonality:**

  - for real vectors: vectors **v** and **u** are orthogonal if

$$\langle \mathbf{v}, \mathbf{u} \rangle = 0$$

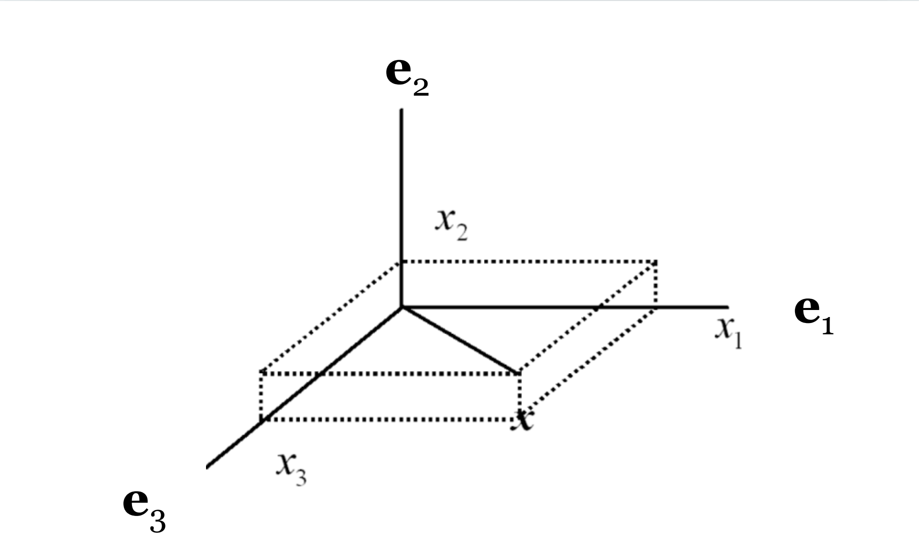
  - for real functions: functions $v(t)$ and $u(t)$ are orthogonal if

$$\langle v(t), u(t) \rangle = 0$$

How to Modulate?

Preliminaries:

- In R^N the vectors defining the coordinate axes are $\mathbf{e}_m$, m=1,…N.
- $\mathbf{e}_m$ contains all zeros except for a 1 in the mth position, i.e., $e_{m,m}$=1 and $e_{m,n}$=1 for m different from n.

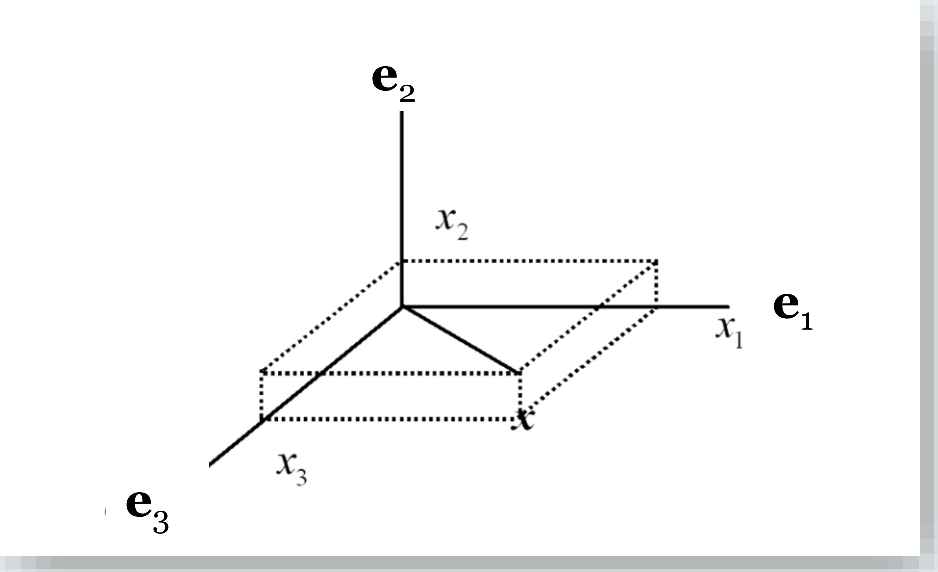


How to Modulate?

Preliminaries:

- These vectors constitute an orthonormal basis:

$$\langle \mathbf{e}_m, \mathbf{e}_n \rangle = 0 \text{ for } m \neq n \text{ (orthogonal)}$$
$$\langle \mathbf{e}_m, \mathbf{e}_m \rangle = ||\mathbf{e}_m||^2 = 1 \text{ (normalized)}$$



How to Modulate?

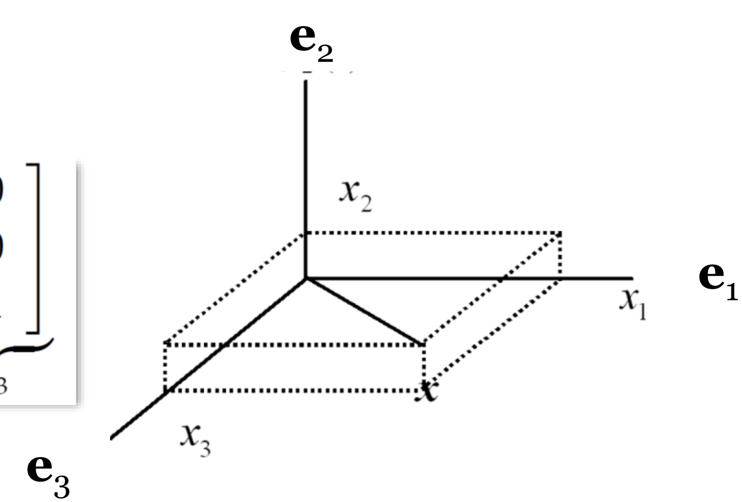
Preliminaries:

- In R^N we can write

$$\mathbf{x}_m = \sum_{n=1}^{N} x_{m,n}\mathbf{e}_n$$

- Ex.:

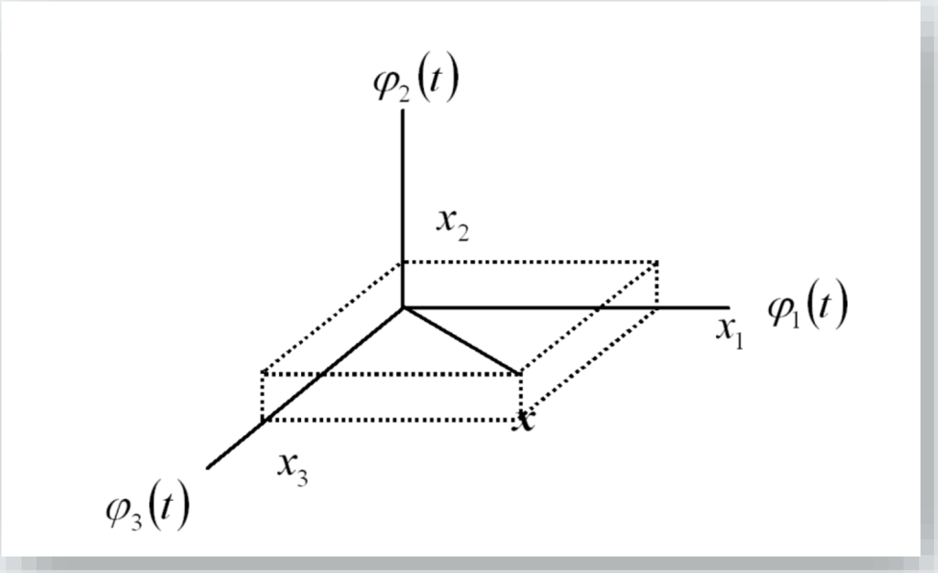
$$\mathbf{x}_m = \begin{bmatrix} x_{m,1} \\ x_{m,2} \\ x_{m,3} \end{bmatrix} = x_{m,1}\underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{e}_1} + x_{m,2}\underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{\mathbf{e}_2} + x_{m,3}\underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{e}_3}$$



How to Modulate?

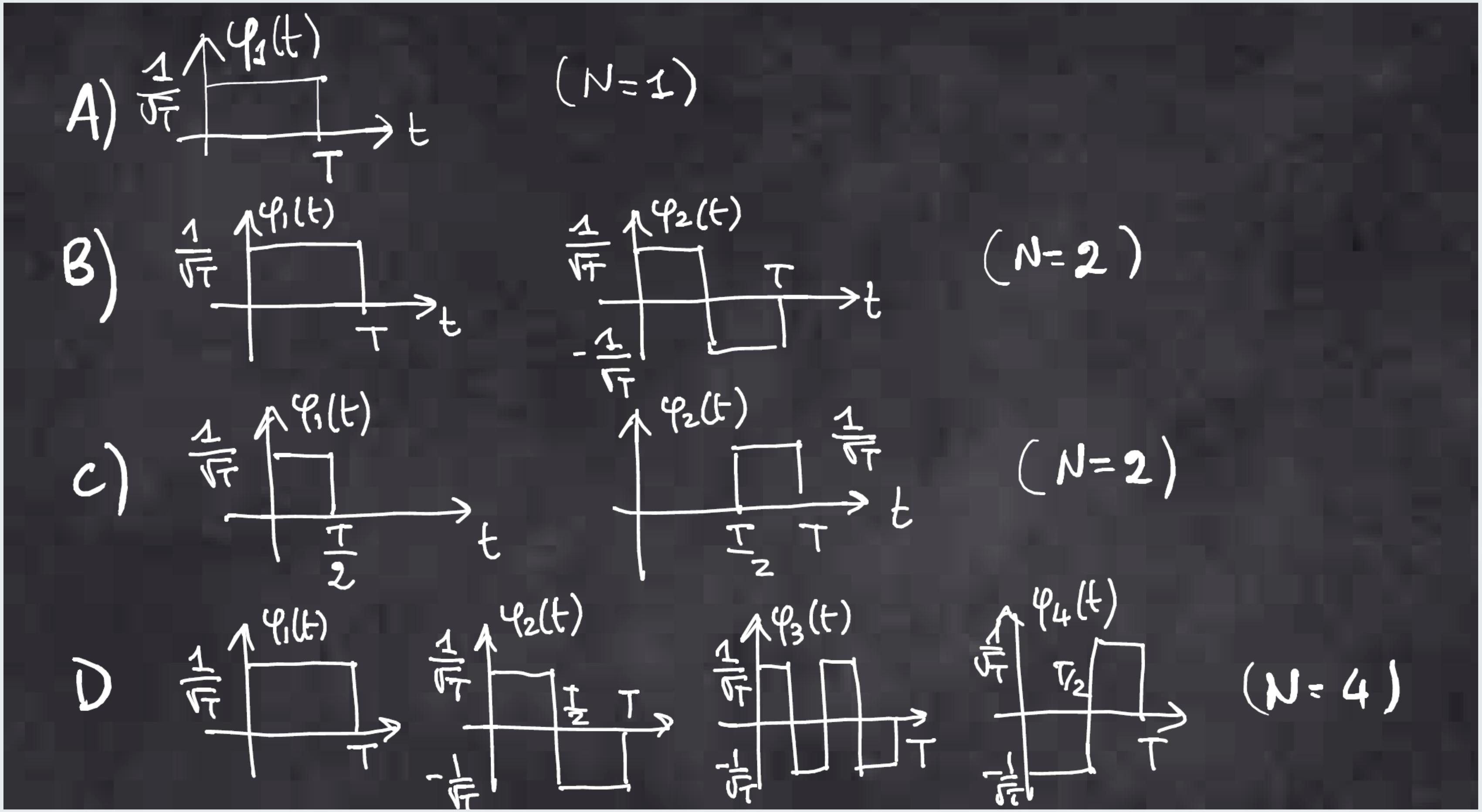
- A set of N function $\{\varphi_m(t)\}$, $m = 1, \dots, N$ is an N-dimensional **orthonormal basis** (and the functions are **orthonormal basis functions**) if it satisfies the conditions

$$\langle \varphi_m(t), \varphi_n(t) \rangle = 0 \text{ for } m \neq n \text{ (orthogonal)}$$
$$\langle \varphi_m(t), \varphi_m(t) \rangle = ||\varphi_m(t)||^2 = 1 \text{ (normalized)}$$



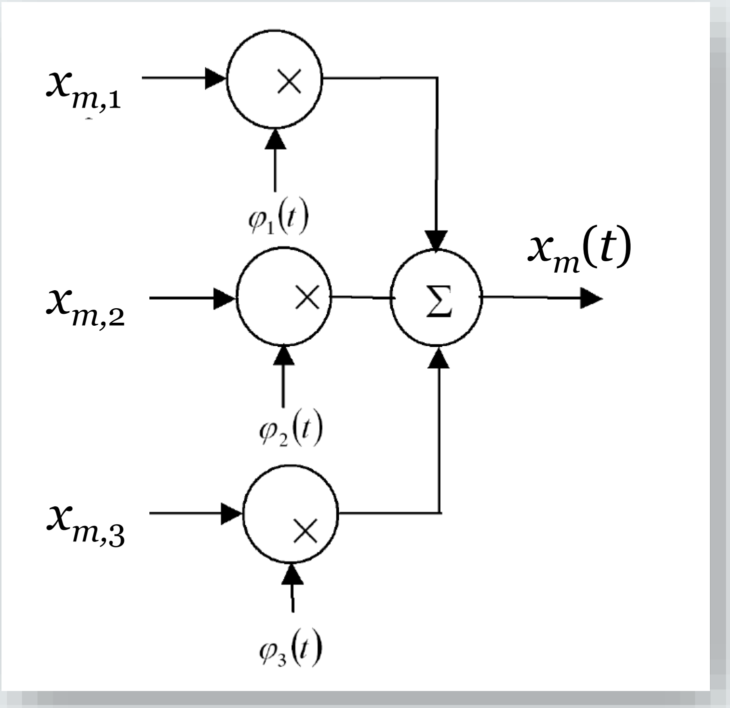
How to Modulate?

- Quiz: Which ones of the sets below is an orthonormal basis?



How to Modulate?

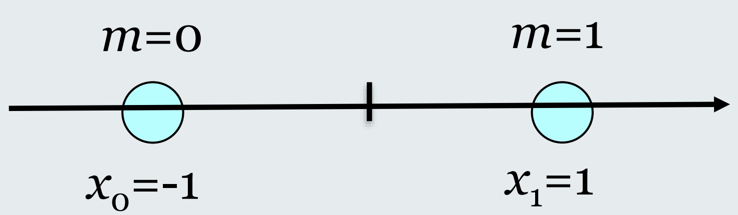
- General block diagram of a modulator:



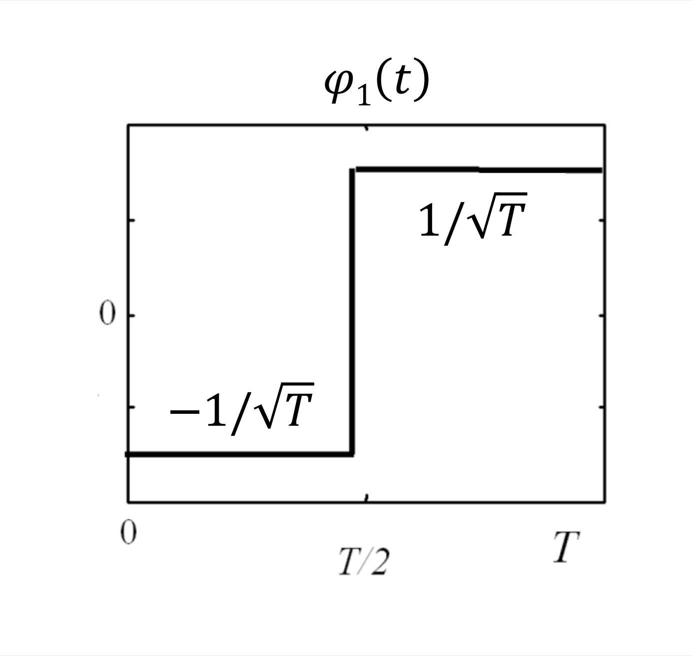
$$x_m(t) = \sum_{n=1}^{N} x_{m,n} \varphi_n(t)$$

How to Modulate?

- **Example:** N=1



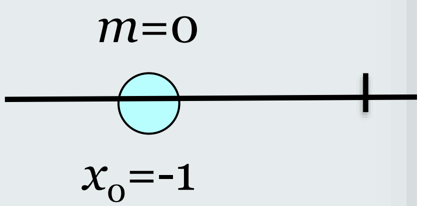
Constellation (E_b=1)

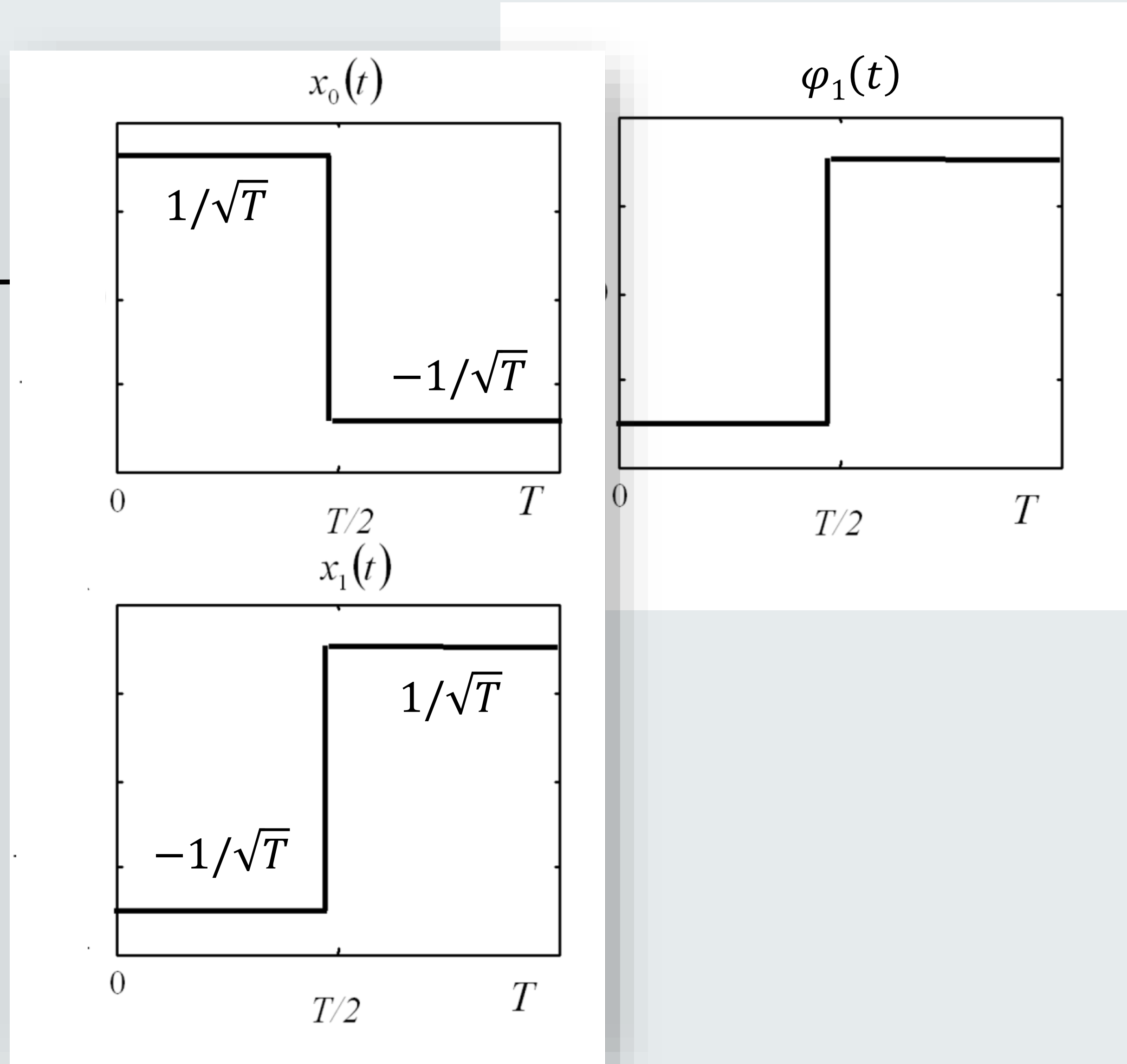


normalized basis function

How to Modulate?

- **Example:** N=1

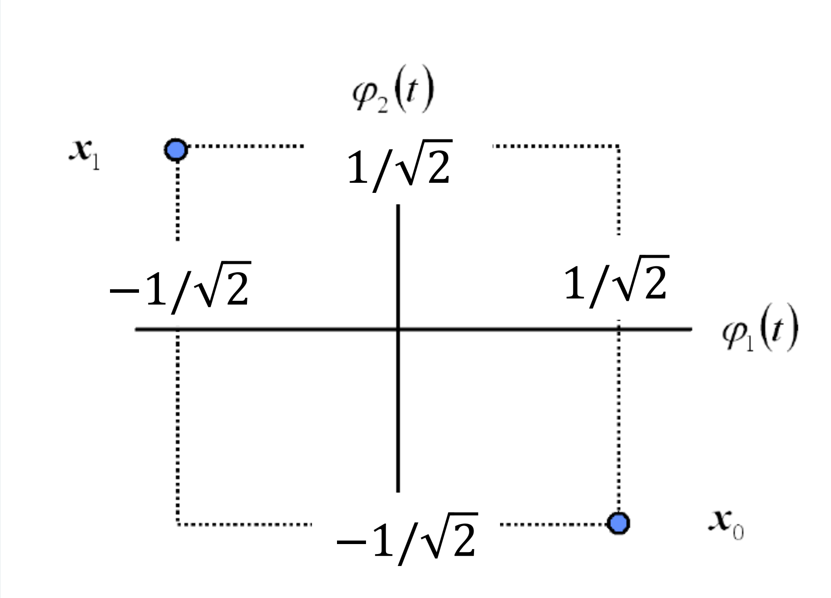




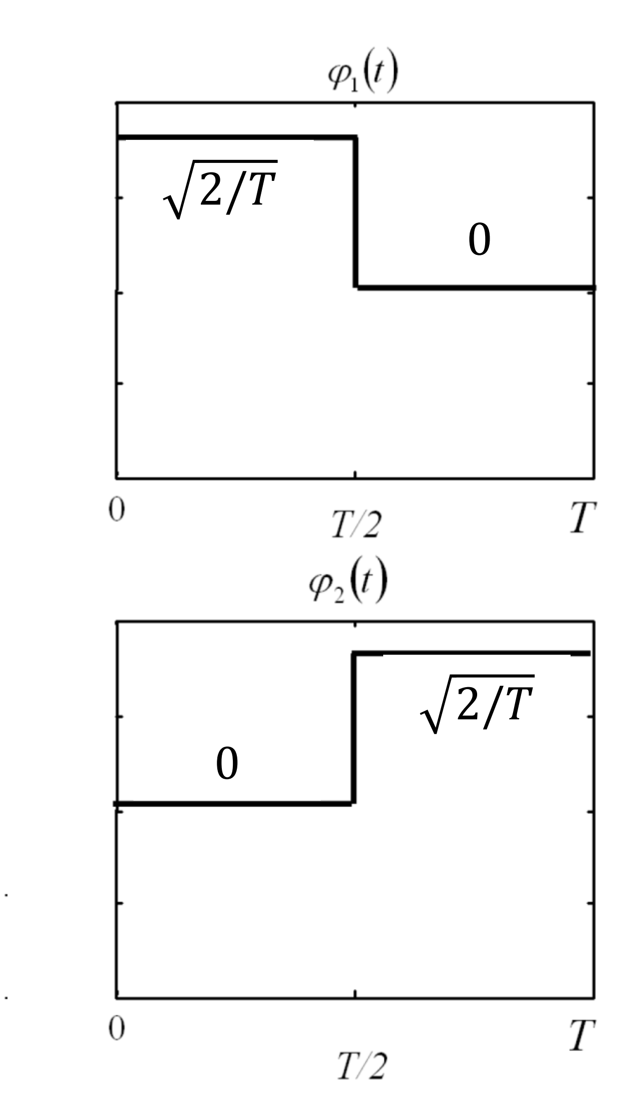
signal set

How to Modulate?

- **Example:** N=2



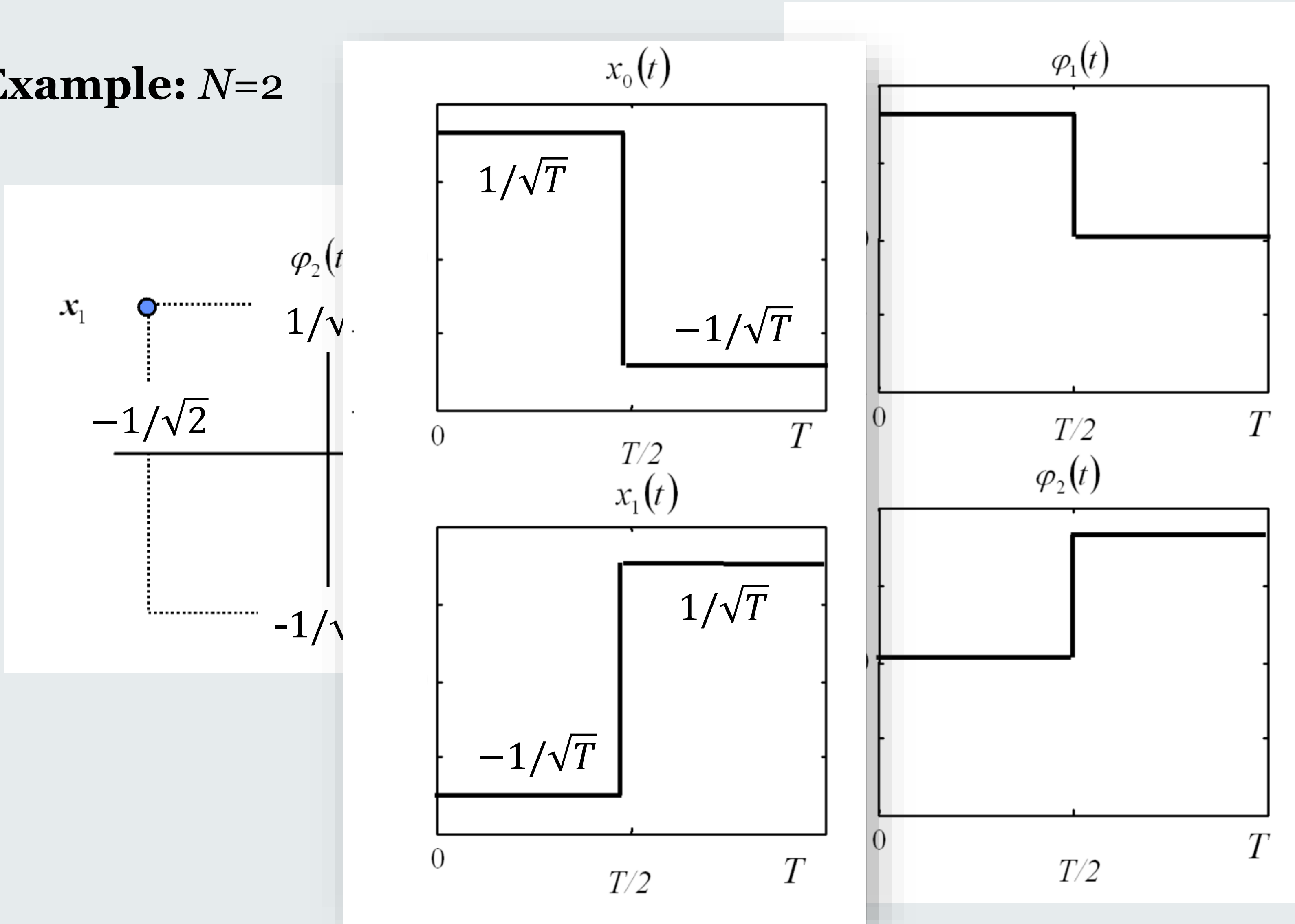
constellation (E_b=1)



orthonormal basis functions

How to Modulate?

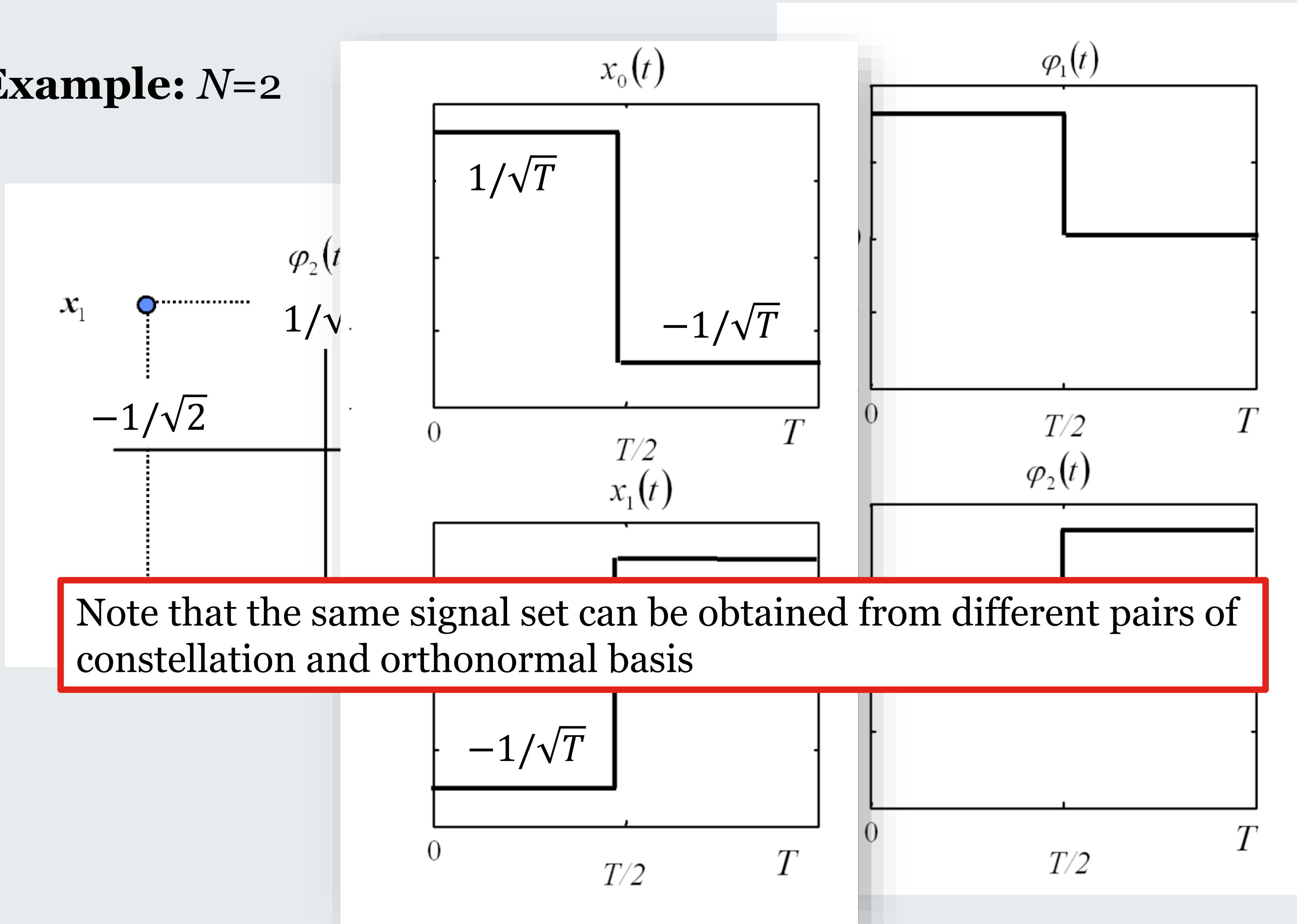
- **Example:** N=2



signal set

How to Modulate?

- **Example:** N=2

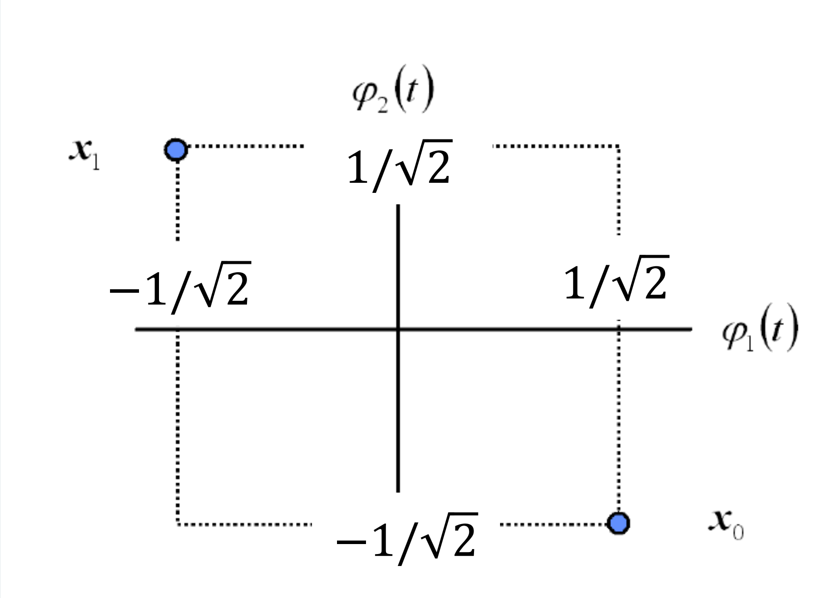


Note that the same signal set can be obtained from different pairs of constellation and orthonormal basis

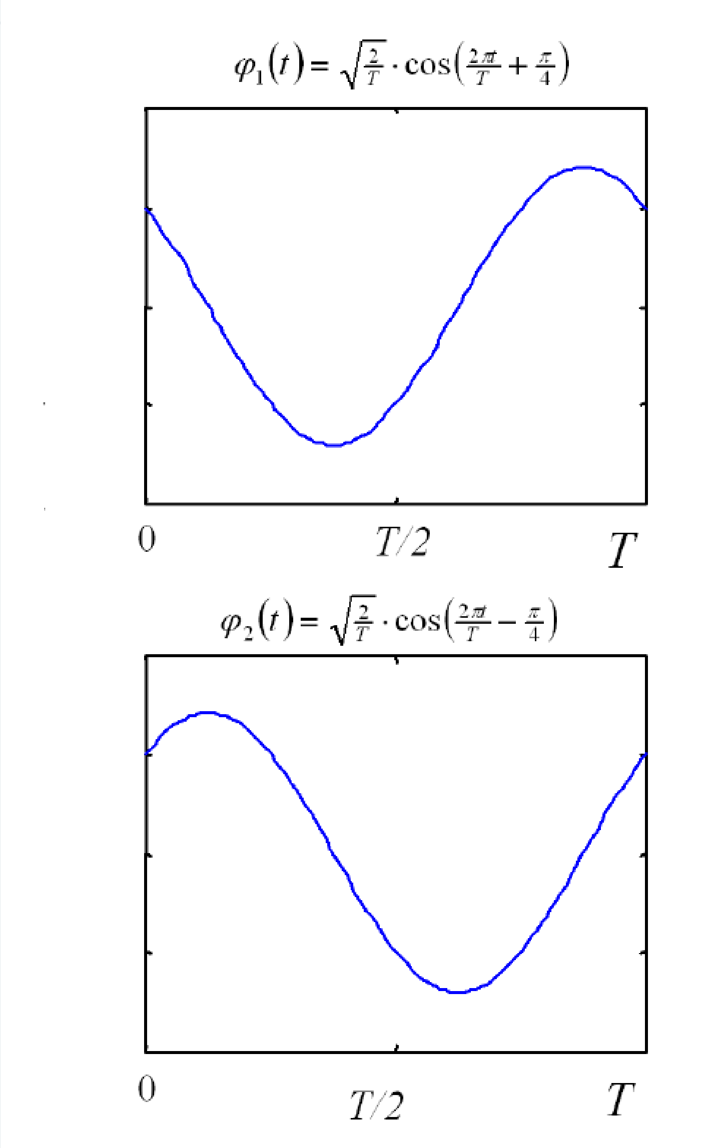
signal set

How to Modulate?

- **Example:** N=2



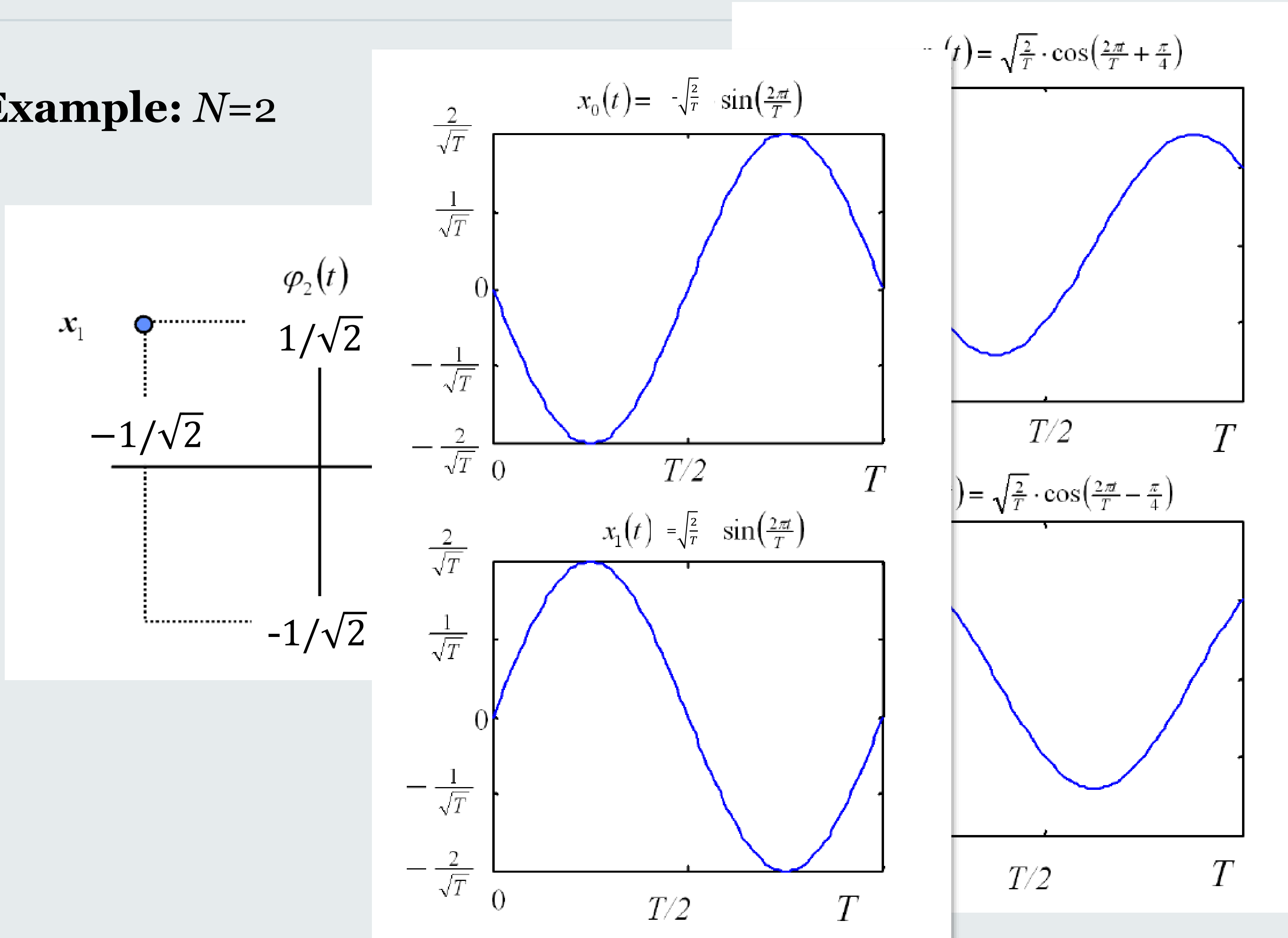
constellation (E_b=1)



orthonormal basis functions

How to Modulate?

- **Example:** N=2



signal set

Some Math

- **Invariance of the inner product:**

Theorem 1.1.1 (Invariance of the Inner Product) *If there exists a set of basis functions* $\varphi_n(t)$, $n = 1, \ldots, N$ *for some* N *such that* $u(t) = \sum_{n=1}^{N} u_n \varphi_n(t)$ *and* $v(t) = \sum_{n=1}^{N} v_n \varphi_n(t)$ *then*

$$\langle u(t), v(t) \rangle = \langle \boldsymbol{u}, \boldsymbol{v} \rangle \ . \tag{1.8}$$

where

$$\boldsymbol{u} \triangleq \begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix} \quad \textit{and} \quad \boldsymbol{v} \triangleq \begin{bmatrix} v_1 \\ \vdots \\ v_N \end{bmatrix} \quad . \tag{1.9}$$

Correlation in signal space = correlation between modulated signals

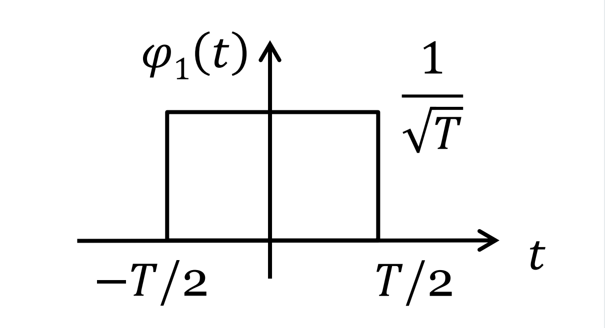
Is the Energy Modified by Modulation?

- Energy of the mth constellation point: $E_x = ||\boldsymbol{x}_m||^2$
- Modulation does not change the energy:

$$
\begin{aligned}
||x_m(t)||^2 &= \langle x_m(t), x_m(t) \rangle \\
&= \langle \mathbf{x}_m, \mathbf{x}_m \rangle \quad \text{(by invariance of inner product)} \\
&= E_m
\end{aligned}
$$

How to Modulate?

- The signals in the previous examples do not have good spectral properties for typical band-limited channels.
- Bandlimited channel: Channel that passes only a limited range of frequencies
- In fact, consider



$$\varphi_1(t) = \frac{1}{\sqrt{T}}\mathrm{rect}\left(\frac{t}{T}\right)$$

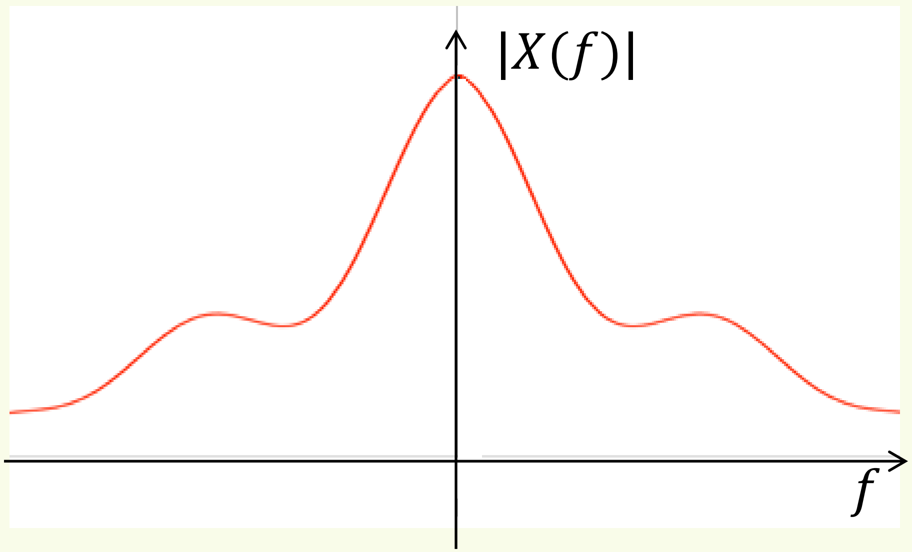
The Fourier transform is $\sqrt{T}\mathrm{sinc}(fT)$, which has an infinite bandwidth.

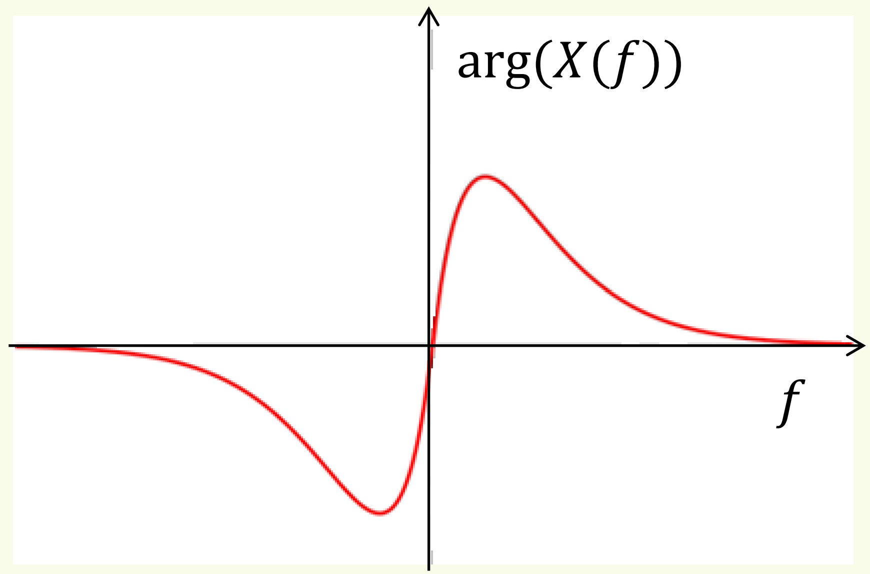
Background: Fourier Transform

- The Fourier transform represents any (energy-limited) signal as the sum of an infinite sum of complex sinusoidal signals

$$e^{j\,2\pi\,ft} = \cos(2\pi f t) + j\sin(2\pi f t)$$

with different amplitudes and phases





$$x(t) = \int_{-\infty}^{+\infty} X(f)\, e^{j\,2\pi f t} df$$

Background: Fourier Transform

- We will be mostly interested in the energy spectrum $G_{\mathrm{x}}(f) = |X(f)|^2$, which describes how the energy of a signal is distributed in the frequency domain.

- Rayleigh theorem

$$E_x = \int_{-\infty}^{+\infty} |x(t)|^2 \, dt = \int_{-\infty}^{+\infty} G_{\mathrm{x}}(f) df$$

- Computation of the Fourier transform

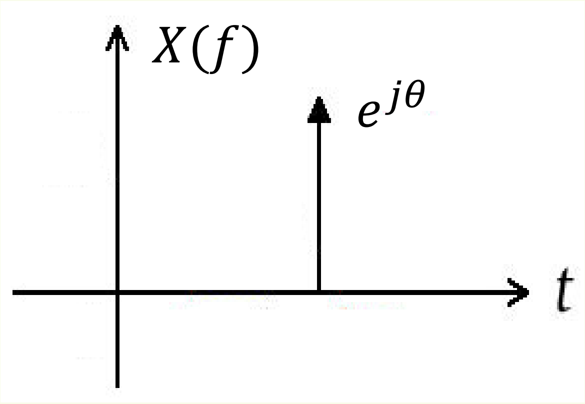
$$X(f) = \int_{-\infty}^{+\infty} x(t) \, e^{-j\,2\pi f t} dt = \mathcal{F}\{x(t)\}$$

- The Fourier transform of non-energy limited signals can also be defined by using the impulse function.

Background: Fourier Transform

- **Example:**

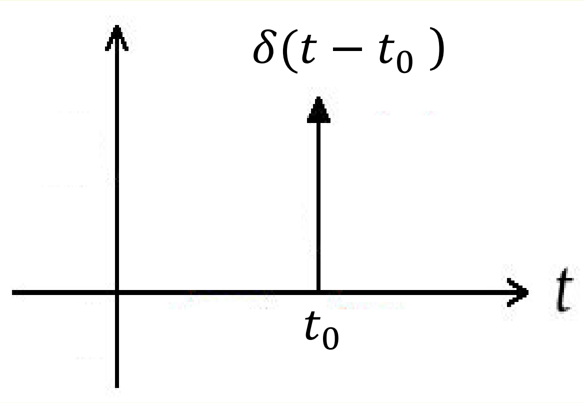
$$x(t) = e^{j(2\pi f_c t+\theta)}$$



$$X(f) = e^{j\theta}\delta(f - f_c)$$

(check by substituting in $x(t) = \int_{-\infty}^{+\infty} X(f)\, e^{j\,2\pi f t} df$ and using the sifting property)

Recall: Impulse function



$\delta(t - t_0\,) = 0$ for all t $\neq t_0$

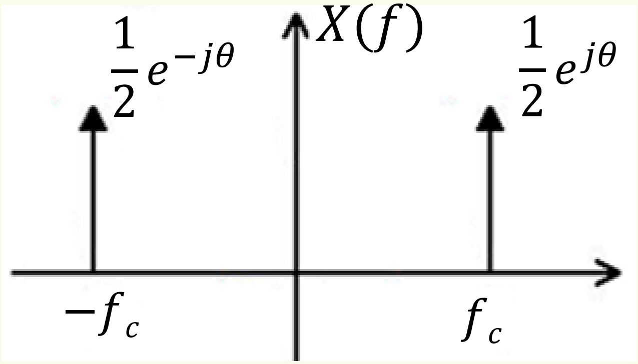
$$\int_{-\infty}^{+\infty} f(t)\,\delta(t - t_0\,)\,dt = f(t_0\,)$$

(sifting property)

Background: Fourier Transform

- **Example:**

$$x(t) = \cos(2\pi f_c t + \theta) = \frac{1}{2} e^{j(2\pi f_c t+\theta)} + \frac{1}{2} e^{-j(2\pi f_c t+\theta)}$$



$$X(f) = \frac{1}{2} e^{j\theta}\delta(f - f_c) + \frac{1}{2} e^{-j\theta}\delta(f + f_c)$$

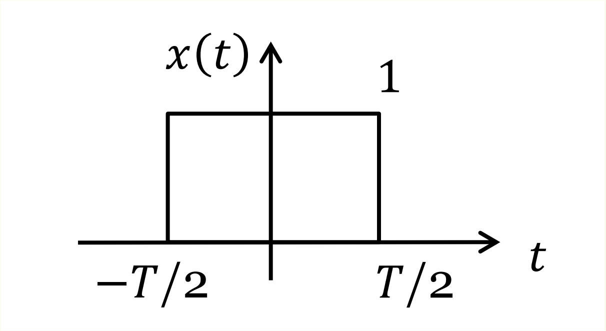
Note that

$$\sin(2\pi f_c t) = \cos\left(2\pi f_c t - \frac{\pi}{2}\right)$$

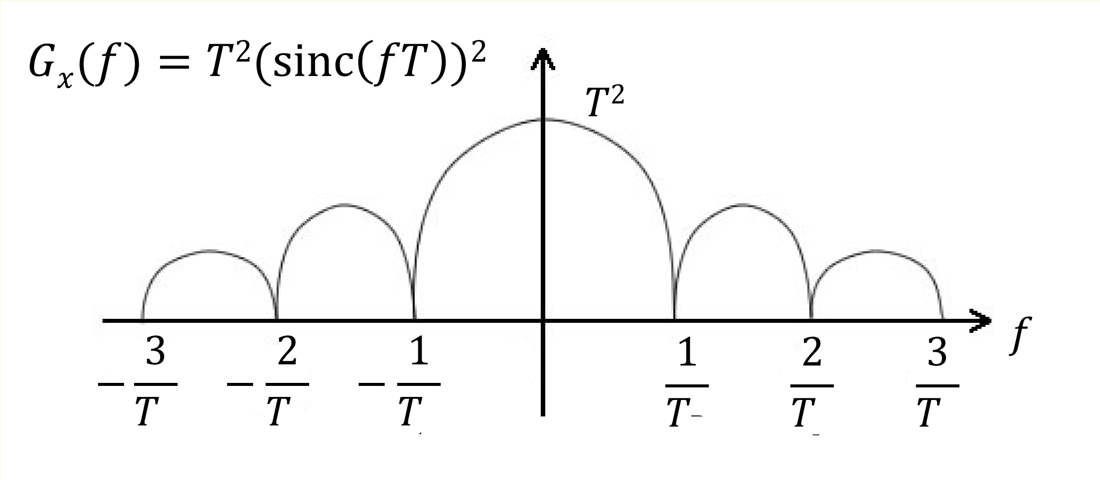
Background: Fourier Transform

- **Example:**

$$x(t) = \mathrm{rect}\left(\frac{t}{T}\right)$$



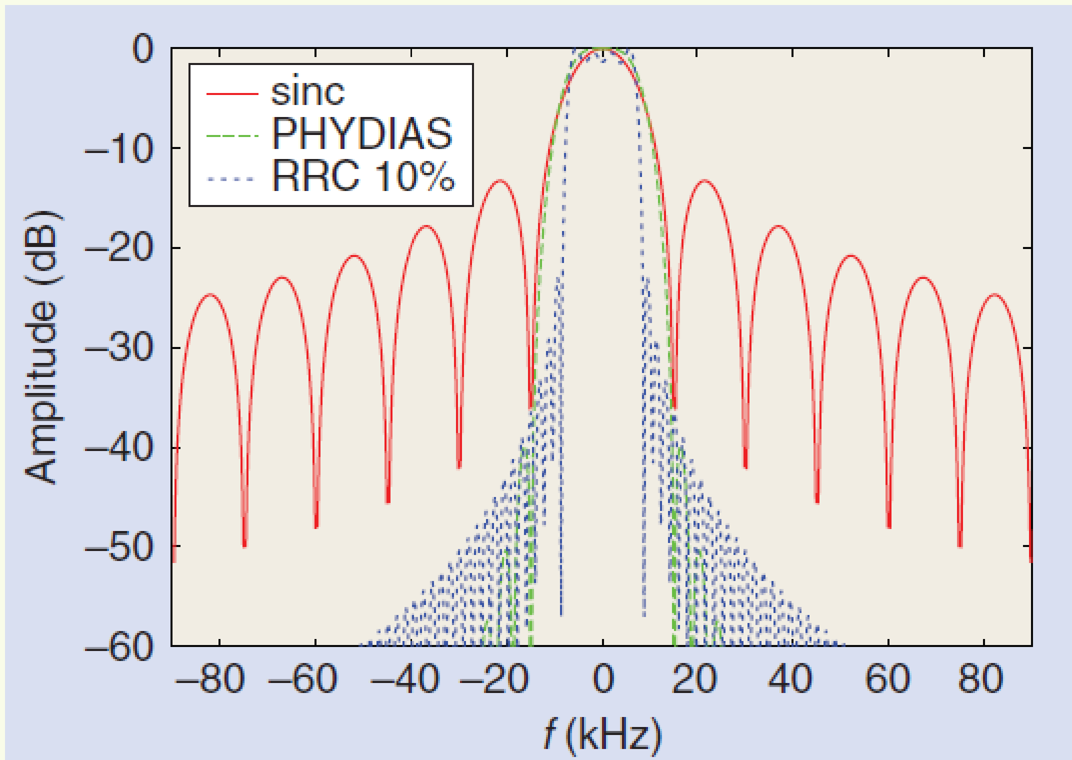
$$X(f) = \int_{-T/2}^{T/2} e^{-j\,2\pi f t}\,dt = T\,\mathrm{sinc}(f\,T)$$



Background: Fourier Transform

- The energy spectrum is typically measured in dB/Hz:

$$G_x(f)|_{\mathrm{dB}} = 10\log_{10} G_x(f)$$

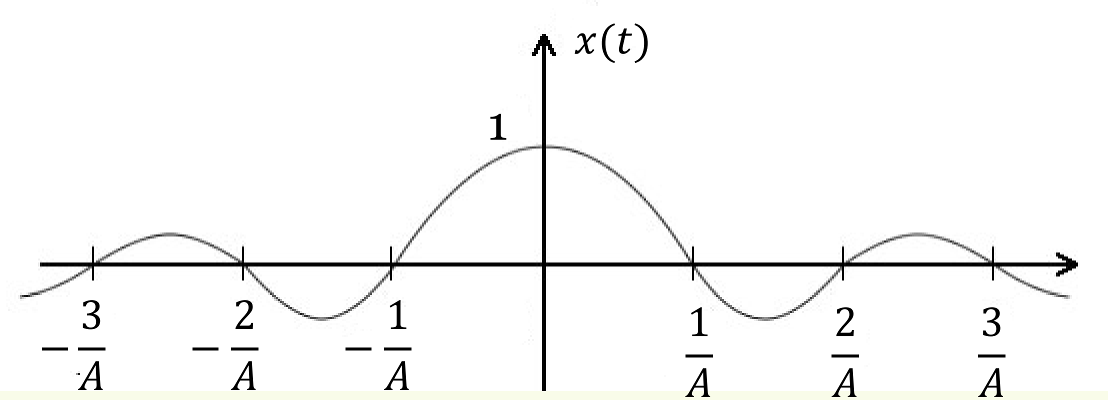


[Banelli et al '14]

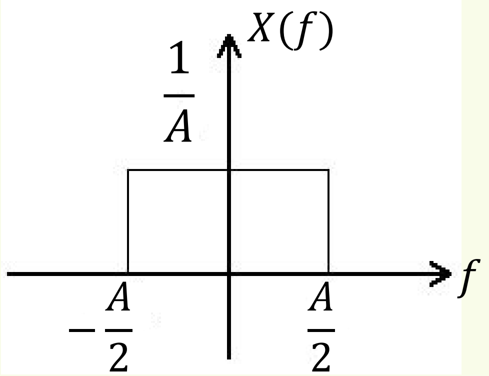
Background: Fourier Transform

- **Example:**

$$x(t) = \text{sinc}(At)$$



$$X(f) = \frac{1}{A}\text{rect}\left(\frac{f}{A}\right)$$



- Mnemonic trick: Energy of a sinc = squared value at peak × time of first zero

Background: Fourier Transform

Properties:

1) Hermitian symmetry: If $x(t)$ is real

$$X(f) = X^*(-f) \quad \text{or equivalently:}$$

$$\begin{cases} |X(f)| = |X(-f)| \\ \arg(X(f)) = -\arg(X(-f)) \end{cases}$$

Hermitian symmetry

Ex.: Rectangular function, sinc

Background: Fourier Transform

2) Frequency translation:

$$\mathcal{F}\left\{x(t)e^{j\,2\pi f_c t}\right\} = X(f - f_c)$$

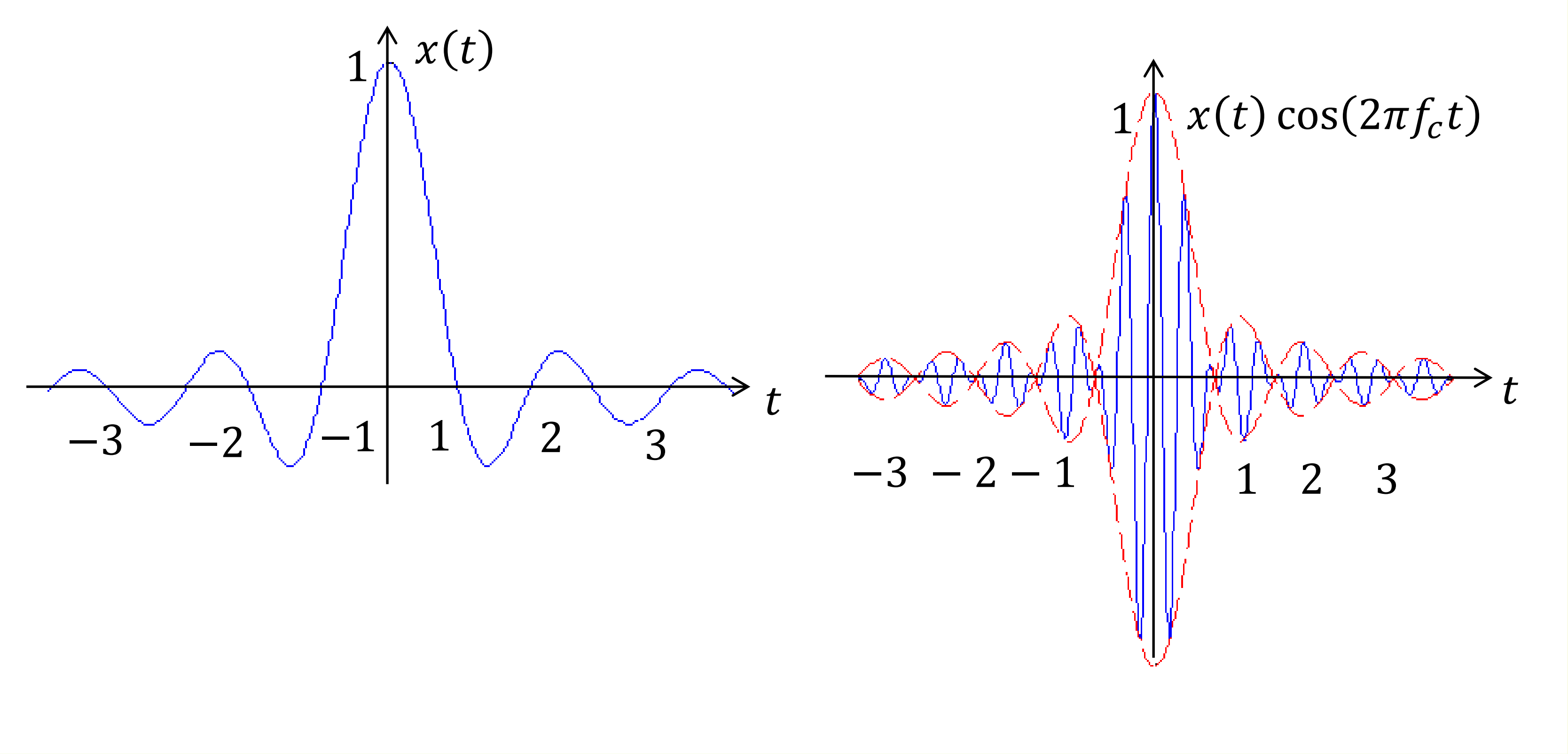
2') Corollary: $\mathcal{F}\{x(t)\cos(2\pi f_c t)\} = \frac{1}{2}(X(f - f_c) + X(f + f_c))$

Proof (corollary): $\cos(2\pi f_c t) = \frac{1}{2}(e^{j\,2\pi f_c t} + e^{-j\,2\pi f_c t})$

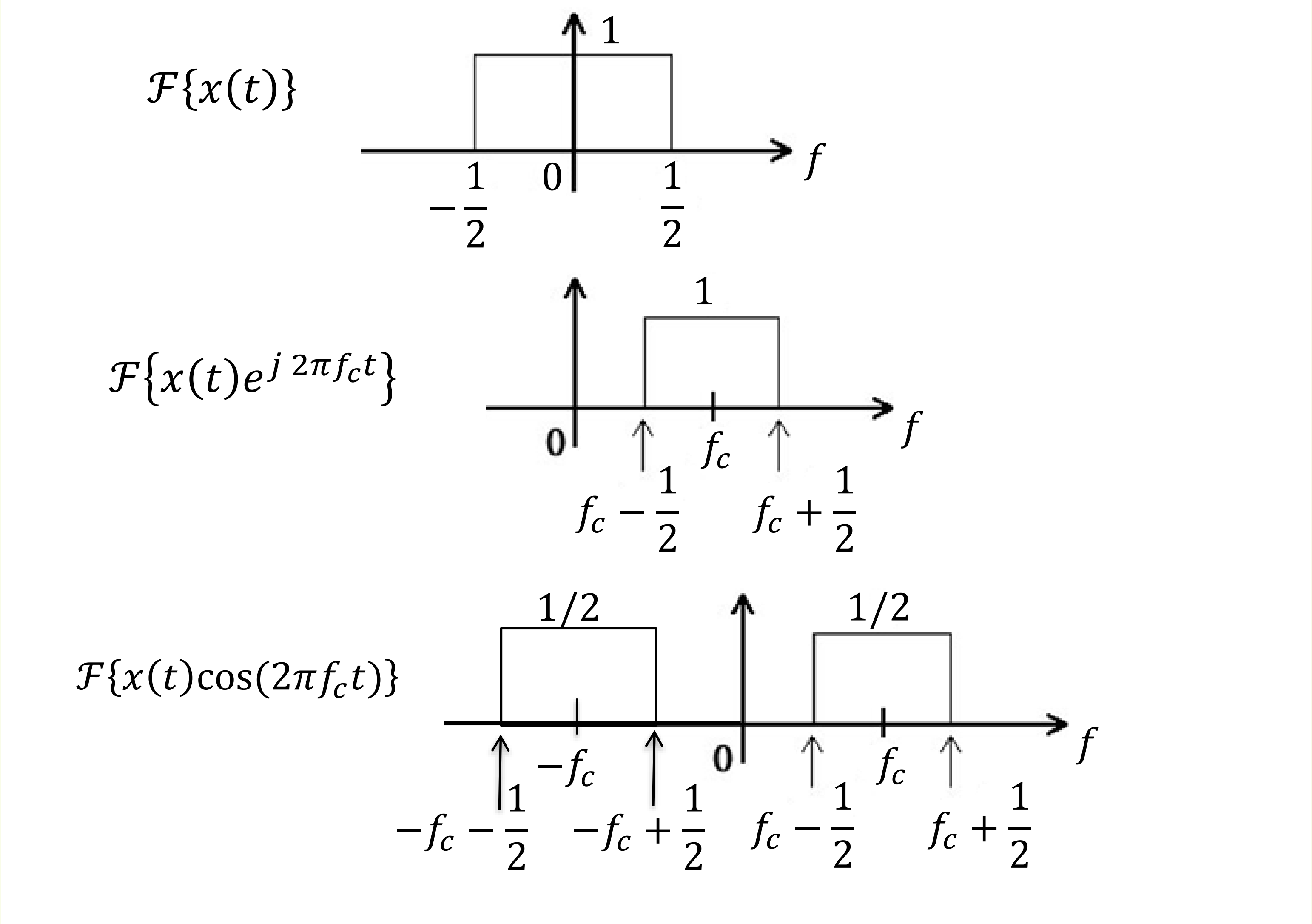
Background: Fourier Transform

Example:

$$x(t) = \mathrm{sinc}(t)$$



Background: Fourier Transform



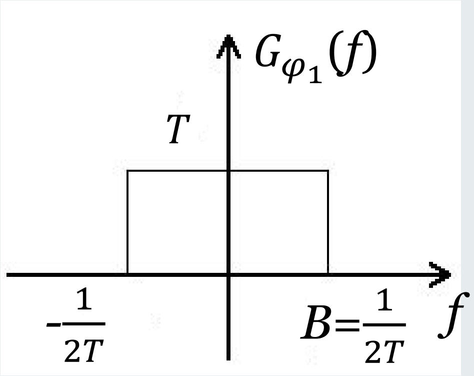
How to Modulate?

- We now provide examples of practically used waveforms over bandlimited channels.

- Baseband signals with N=1:

$$\varphi_1(t) = \frac{1}{\sqrt{T}} \operatorname{sinc}\left(\frac{t}{T}\right)$$

Bandwidth limited to $B=1/(2T)$.



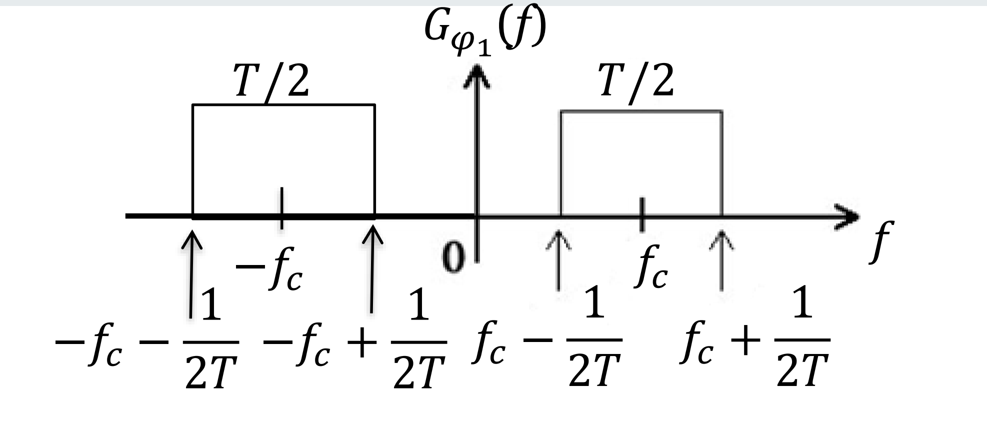
How to Modulate?

- We now provide examples of practically used waveforms over bandlimited channels.

- Passband signals with $N = 1$ and carrier frequency $f_c \gg \frac{1}{T}$:

$$\varphi_1(t) = \sqrt{\frac{2}{T}} \operatorname{sinc}\left(\frac{t}{T}\right) \cos(2\pi f_c t)$$

Bandwidth limited to B=1/T around carrier frequency f_c.



How to Modulate?

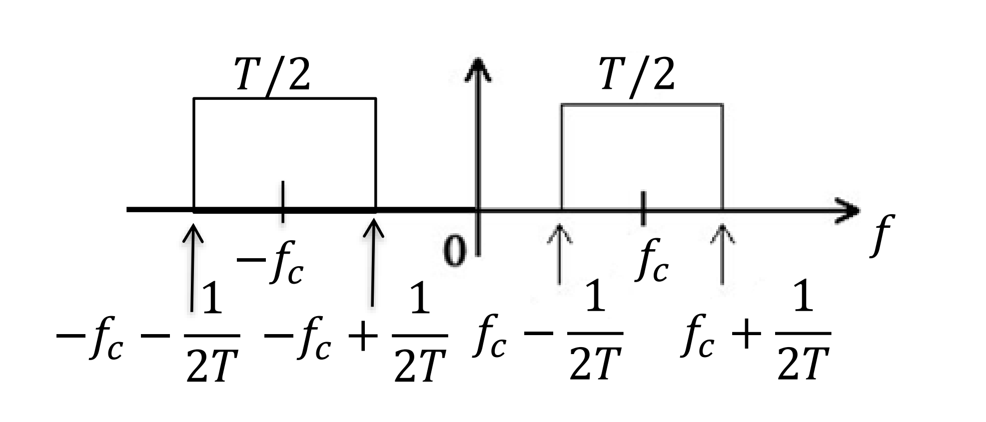
- Passband signals with N=2 and carrier frequency $f_c \gg \frac{1}{T}$

$$\varphi_1(t) = \sqrt{\frac{2}{T}} \operatorname{sinc}\left(\frac{t}{T}\right) \cos(2\pi f_c t)$$

$$\varphi_2(t) = \sqrt{\frac{2}{T}} \operatorname{sinc}\left(\frac{t}{T}\right) \sin(2\pi f_c t)$$

Bandwidth limited to B=1/T around carrier frequency f_c.

$G_{\varphi_1}(f)$=$G_{\varphi_2}(f)$



How to Modulate?

- Passband signals with N=2 and carrier frequency $f_c \gg \frac{1}{T}$

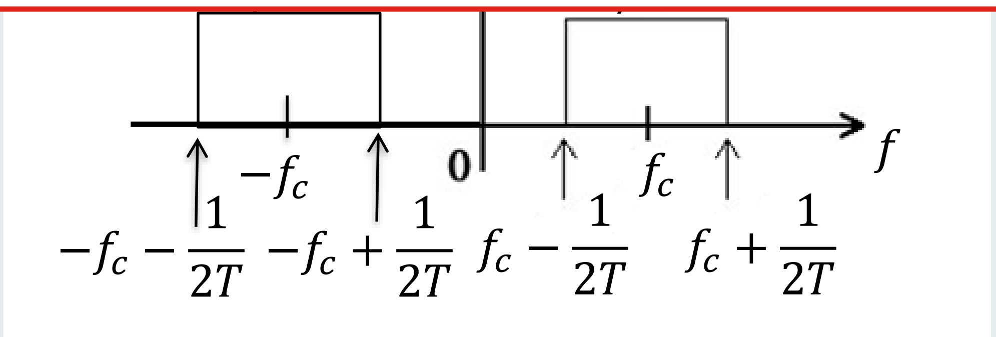
$$\varphi_1(t) = \sqrt{\frac{2}{T}}\,\mathrm{sinc}\left(\frac{t}{T}\right)\cos(2\pi f_c t)$$

$$\varphi_2(t) = \sqrt{\frac{2}{T}}\,\mathrm{sinc}\left(\frac{t}{T}\right)\sin(2\pi f_c t)$$

Bar

Note: The cos and sin carriers are orthogonal and said to be in quadrature. More generally, we can choose any two carriers whose phases differ by $\pi/2$.

$G_{\varphi_1}(f)$=$G_{\varphi_2}(f)$



How to Modulate?

- Passband waveforms with any *N*:

Orthogonal Frequency Division Multiplexing (OFDM), used in 4G, Wi-Fi,…

$$\varphi_m(t) = \sqrt{\frac{2}{T}}\,\mathrm{rect}\left(\frac{t}{T}\right)\cos\left(2\pi(f_c + \frac{m-1}{T} - \frac{N}{4T})t\right) \text{ for } m=1,\ldots,N/2 \text{ and}$$

$$\varphi_m(t) = \sqrt{\frac{2}{T}}\,\mathrm{rect}\left(\frac{t}{T}\right)\sin\left(2\pi(f_c + \frac{m-(\frac{N}{2}+1)}{T} - \frac{N}{4T})t\right) \text{ for } m=N/2+1,\ldots,N$$

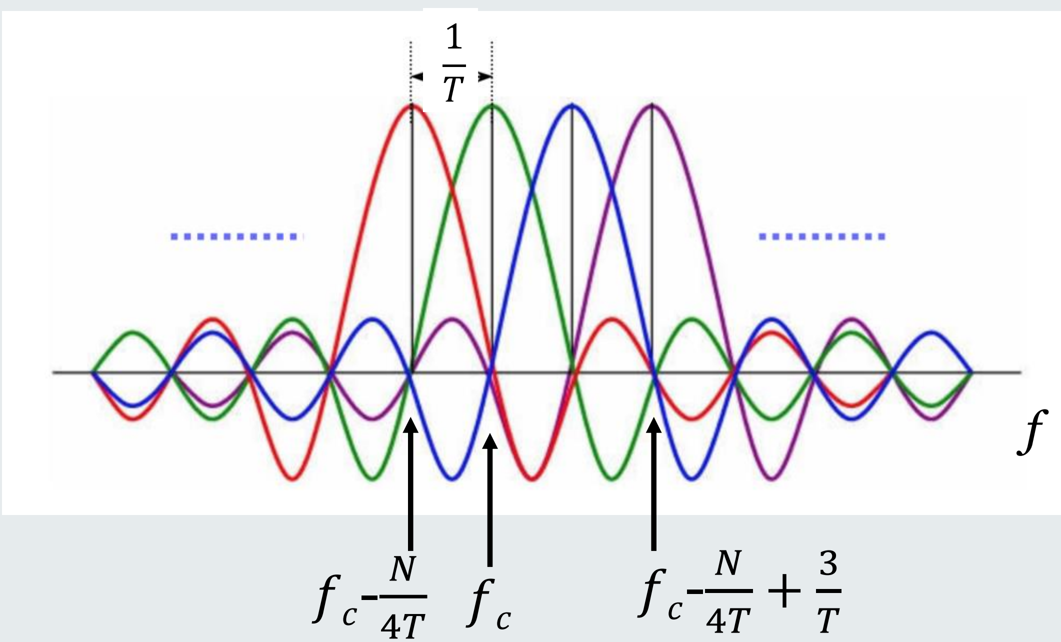
How to Modulate?

- Passband waveforms with any N:

Orthogonal Frequency Division Multiplexing (OFDM), used in 4G, Wi-Fi,...

$$\varphi_m(t) = \sqrt{\frac{2}{T}}\,\mathrm{rect}\left(\frac{t}{T}\right)\cos\left(2\pi\left(f_c + \frac{m-1}{T} - \frac{N}{4T}\right)t\right) \text{ for } m=1,\ldots,N/2 \text{ and}$$

$$\varphi_m(t) = \sqrt{\frac{2}{T}}\,\mathrm{rect}\left(\frac{t}{T}\right)\sin\left(2\pi\left(f_c + \frac{m-(\frac{N}{2}+1)}{T} - \frac{N}{4T}\right)t\right) \text{ for } m=N/2+1,\ldots,N$$



$$B \approx \frac{N}{2T}$$

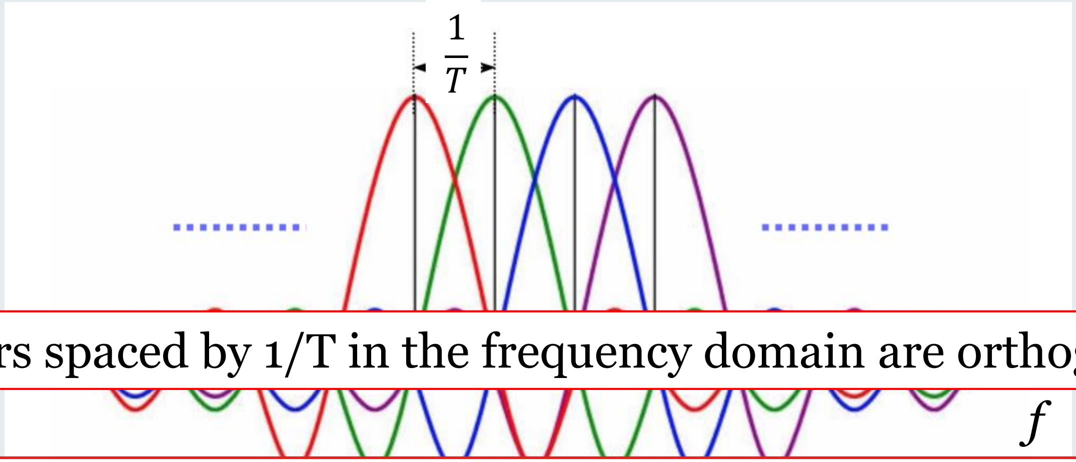
How to Modulate?

- Passband waveforms with any *N:*

Orthogonal Frequency Division Multiplexing (OFDM), used in 4G, Wi-Fi,...

$$\varphi_m(t) = \sqrt{\frac{2}{T}}\,\mathrm{rect}\left(\frac{t}{T}\right)\cos\left(2\pi\left(f_c + \frac{m-1}{T} - \frac{N}{4T}\right)t\right) \text{ for } m=1,\ldots,N/2 \text{ and}$$

$$\varphi_m(t) = \sqrt{\frac{2}{T}}\,\mathrm{rect}\left(\frac{t}{T}\right)\sin\left(2\pi\left(f_c + \frac{m-(\frac{N}{2}+1)}{T} - \frac{N}{4T}\right)t\right) \text{ for } m=N/2+1,\ldots,N$$



Note: Carriers spaced by 1/T in the frequency domain are orthogonal.

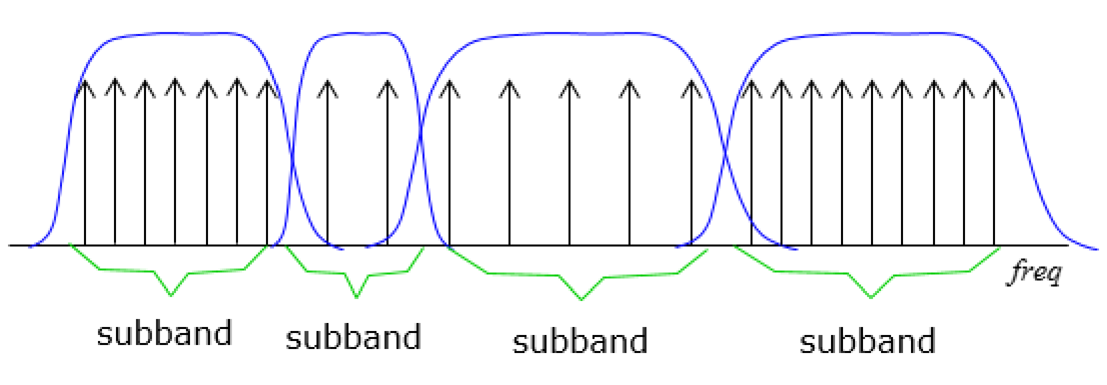
Remark: OFDM also includes a cyclic prefix that allows it to simplify the operation over frequency selective channels

How to Modulate?

- Passband waveforms with any *N:*

filtered-OFDM (f-OFDM), candidate waveform for 5G

$$\varphi_m(t) = \varphi_m^{\mathrm{OFDM}}(t) * h(\mathrm{t})$$



http://www.sharetechnote.com/html/5G/5G_Phy_Candidate_fOFDM.html

How to Modulate?

- It can be proven that the maximum number of orthogonal dimensions that can be accommodated in a bandwidth B over a time T is

$$N=2BT$$

- It follows that, except for $\varphi_1(t) = \sqrt{\frac{2}{T}}\,\mathrm{sinc}\left(\frac{t}{T}\right)\cos(2\pi f_c t)$ (N=1), the presented orthonormal basis functions use the available bandwidth and time in the most efficient way.

- In practice, it is generally preferable to choose orthonormal basis functions with $N<2BT$ that are easier to realize (the sharp spectral transitions of the sinc can only be approximated.)

How to Modulate?

- Quiz:

Which of the following functions are suitable for transmission on a passband channel that has a bandwidth of 1 MHz around a center frequency of 1 GHz?

A) $\varphi_1(t) = \sqrt{2\times 10^7}\,\mathrm{sinc}(t\times 10^7)\cos(2\pi 10^9 t)$

B) $\varphi_1(t) = \sqrt{2\times 10^6}\,\mathrm{sinc}(t\times 10^6)\cos(2\pi(1.001)\times 10^9 t)$

C) $\varphi_1(t) = \sqrt{2\times 10^6}\,\mathrm{sinc}(t\times 10^6)\cos(2\pi 10^9 t)$

How Effectively Is Bandwidth Used by a Transmitter?

- Recall that the spectral efficiency is defined as

$$\eta = \frac{R}{B} \text{ (bit/s/Hz)}$$

- Ex.: The modulator uses the waveform $\varphi_1(t)=1/\sqrt{T}\ \mathrm{sinc}(t/T)$ with a bandwidth of 1 MHz and the encoder uses either BPSK or 4-PAM. What is the spectral efficiency?

How Effectively Is Bandwidth Used by a Transmitter?

- Recall that the spectral efficiency is defined as

$$\eta = \frac{R}{B} \text{ (bit/s/Hz)}$$

- Ex.: The modulator uses the waveform $\varphi_1(t)=1/\sqrt{T}\ \mathrm{sinc}(t/T)$ with a bandwidth of 1 MHz and the encoder uses either BPSK or 4-PAM. What is the spectral efficiency?

- We have

$$\eta = \frac{R}{B} = \frac{b/T}{1/2T} = 2b$$

and hence $\eta = 2$ for BPSK and $\eta = 4$ for 4-PAM with this choice of waveform.

How Long Does It Take to Transmit a File?

- Ex.: Assume that we would like to transmit a file of size 2 Gbits. The modulator uses the baseband waveform $\varphi_1(t) = \frac{1}{\sqrt{T}} \operatorname{sinc}\left(\frac{t}{T}\right)$ with a bandwidth of 1 MHz and the encoder uses either BPSK or 4-PAM. How long does it take to complete transmission of the file (assuming that there are no errors)?

How Long Does It Take to Transmit a File?

- Ex.: Assume that we would like to transmit a file of size 2 Gbits. The modulator uses the baseband waveform $\varphi_1(t) = \frac{1}{\sqrt{T}} \text{sinc}\left(\frac{t}{T}\right)$ with a bandwidth of 1 MHz and the encoder uses either BPSK or 4-PAM. How long does it take to complete transmission of the file (assuming that there are no errors)?

The symbol period is T = 1/(2B) =0.5 × 10^{-6} s. The time need to download the file is given as

$$\frac{\text{file size}}{R} = \frac{2\times10^9}{b/T} = \frac{10^3}{b}$$

Hence, it takes 1000 seconds to download with BSK and 500 seconds with 4- PAM.

How Long Does It Take to Transmit a File?

- Ex.: Assume that we would like to transmit a file of size 2 Gbits. The modulator uses the passband waveform $\varphi_1(t) = \sqrt{\frac{2}{T}}\operatorname{sinc}\left(\frac{t}{T}\right)\cos(2\pi f_c t)$ with a bandwidth of 1 MHz and the encoder uses either BPSK or 4-PAM. How long does it take to complete transmission of the file (assuming that there are no errors)?

How Long Does It Take to Transmit a File?

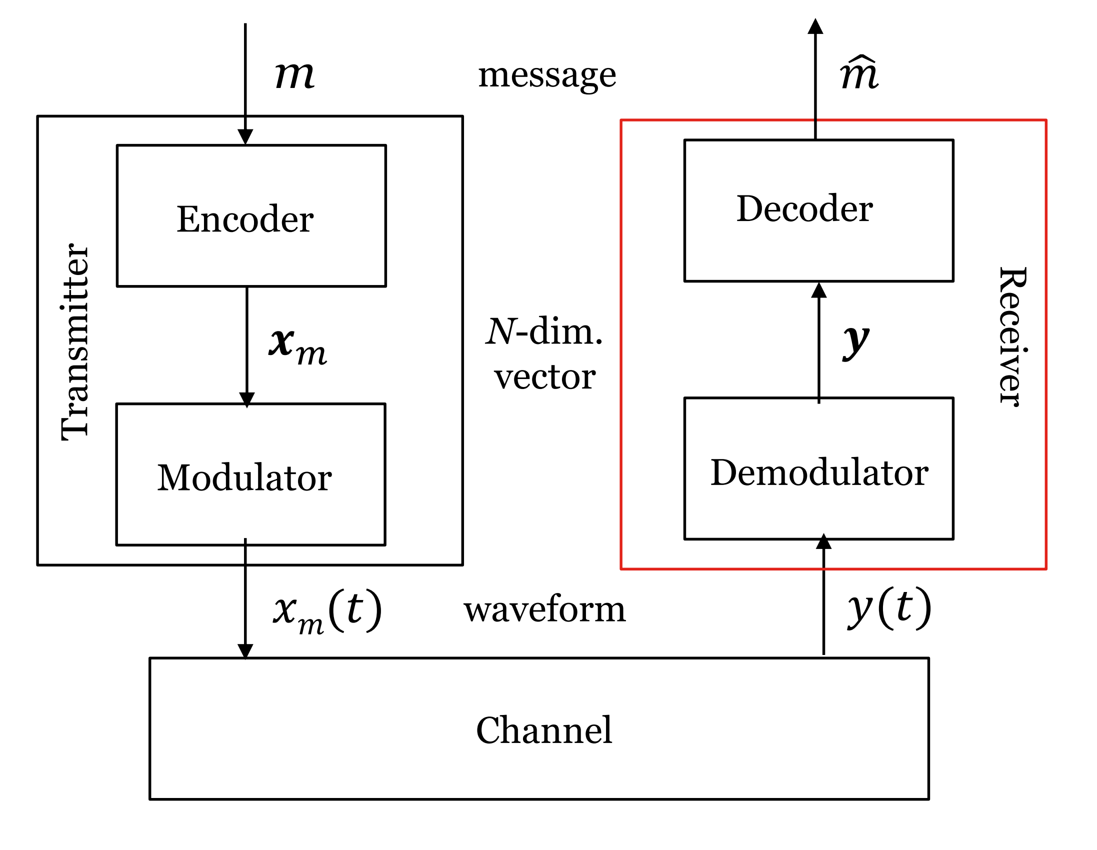
- Ex.: Assume that we would like to transmit a file of size 2 Gbits. The modulator uses the passband waveform $\varphi_1(t) = \sqrt{\frac{2}{T}}\,\mathrm{sinc}\left(\frac{t}{T}\right)\cos(2\pi f_c t)$ with a bandwidth of 1 MHz and the encoder uses either BPSK or 4-PAM. How long does it take to complete transmission of the file (assuming that there are no errors)?

The symbol period is T = 1/B = 10^{-6} s. The time need to download the file is given as

$$\frac{\text{file size}}{R} = \frac{2\times 10^9}{b/T} = \frac{2\times 10^3}{b}$$

Hence, it takes 2000 seconds to download with BSK and 1000 seconds with 4- PAM.

What Does the Receiver Do?



- **Demodulator**: (continuous-time analog) channel output signal $y(t) \rightarrow$ channel output vector $\mathbf{y}$ in the signal space
- **Detector:** channel output vector $\mathbf{y} \rightarrow$ estimate $\hat{m}$ of the message m

What Does the Receiver Do?

- Useful observations:

  1) **vector projection:** for a vector

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \sum_{n=1}^{N} x_n e_n$$

the nth component can be obtained via correlation with the basis vector e_n (projection of x into e_n)

$$< x, e_n >= x_n$$

What Does the Receiver Do?

- Useful observations:

1) **vector projection:** for a vector

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \sum_{n=1}^{N} x_n e_n$$

the *n*th component can be obtained via correlation with the basis vector e_n (projection of x into e_n)

$$< x, e_n >= x_n$$

2) modulating e_n yields $\varphi_n(t)$

3) **invariance of inner product:** correlation is equal in the signal space and on the modulated signals

How Can the Message Be Recovered Without Noise?

- Assume that the received signal $y(t)$ is noiseless and hence

$$y(t) = x(t) = \sum_{n=1}^{N} x_n \varphi_n(t)$$

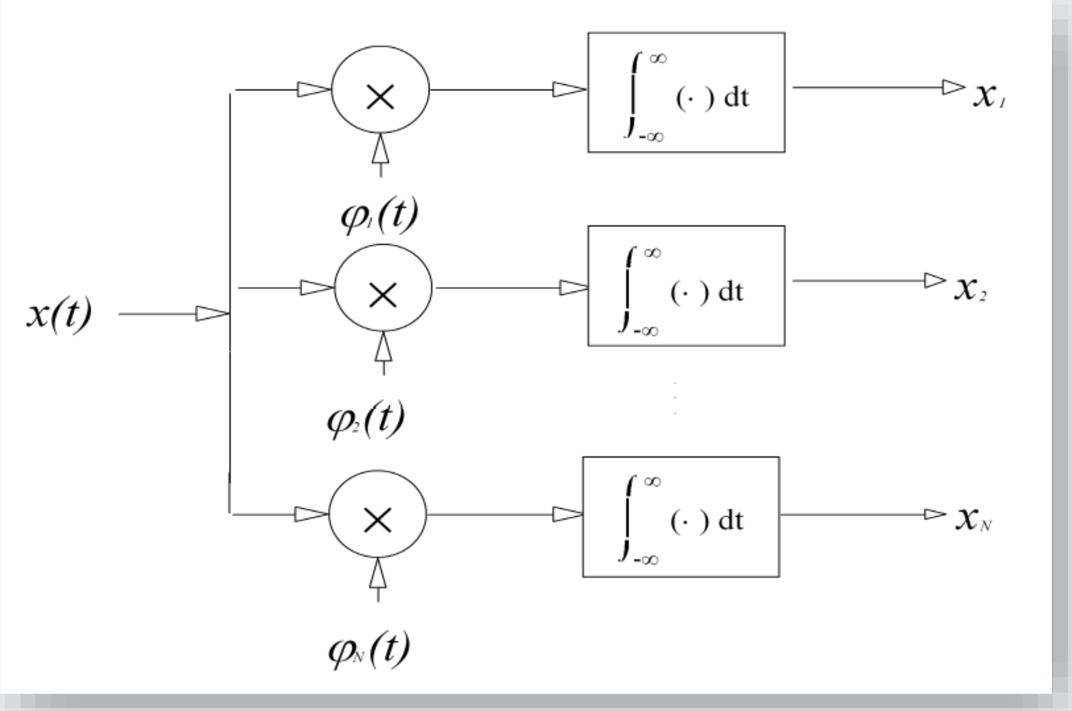
- Therefore, recovering each constellation coordinate x_n is equivalent to projecting y(t) into $\varphi_n(t)$

How Can the Message Be Recovered Without Noise?

- Assume that the received signal $y(t)$ is noiseless and hence

$$y(t) = x(t) = \sum_{n=1}^{N} x_n \varphi_n(t)$$

- The demodulator can recover the symbol **x** as follows:



correlative demodulator

$$\langle y(t), \varphi_n(t) \rangle = x_n$$

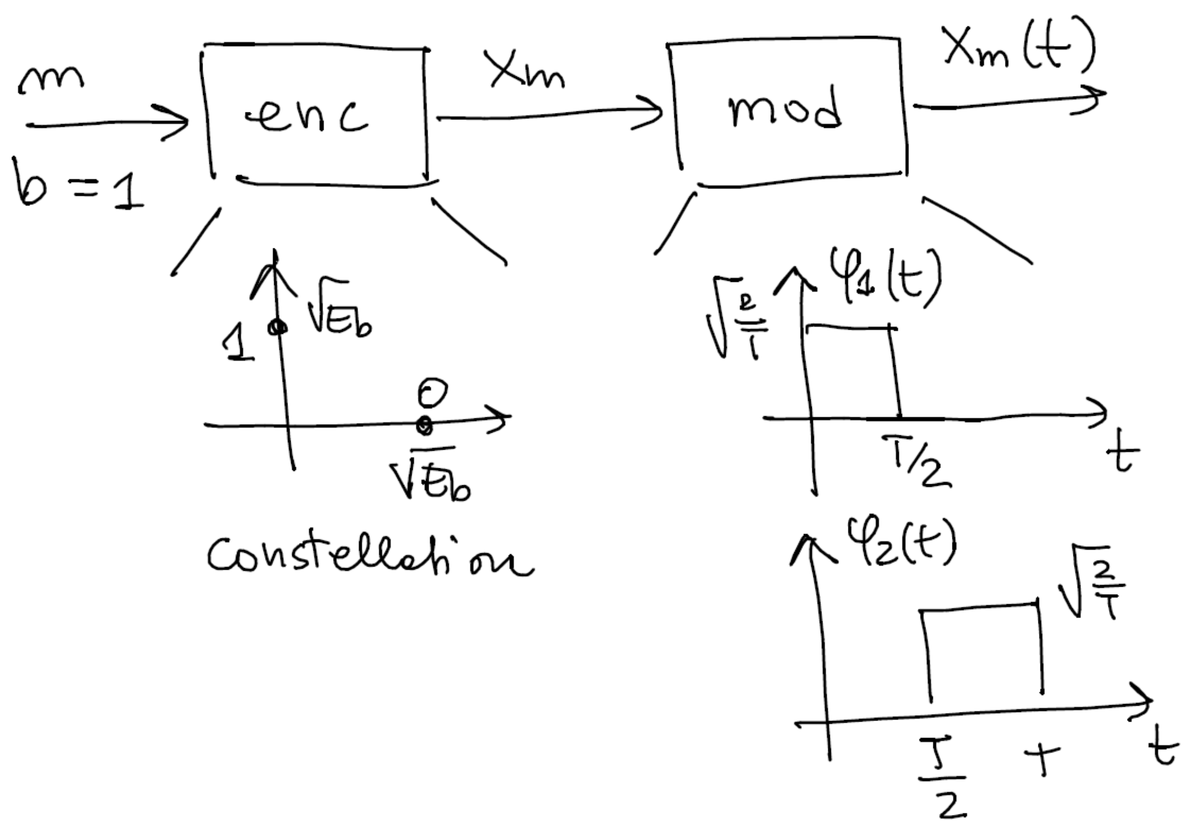
How Can the Message Be Recovered Without Noise?

- Proof:

$$< x(t), \varphi_n(t) > = < x, e_n > = x_n$$

How Can the Message Be Recovered Without Noise?

Example: Pulse Position Modulation (PPM)

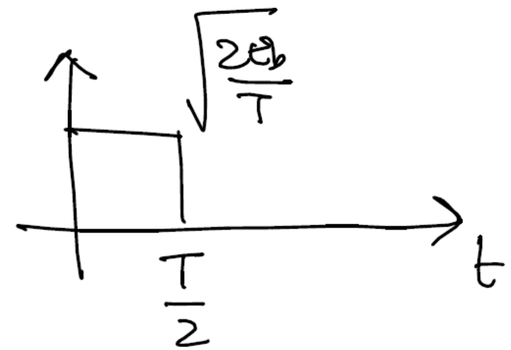


- What is the signal set? (Why is it called PPM?)

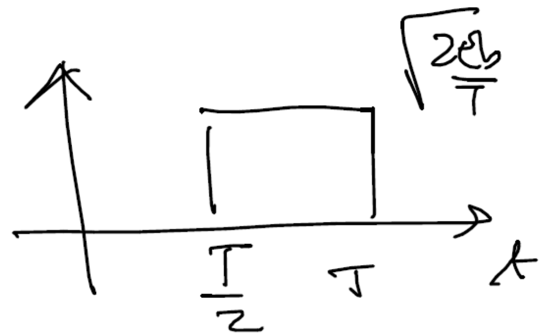
How Can the Message Be Recovered Without Noise?

Signal set :

$$x_0(t) = \sqrt{E_b}\,\varphi_1(t) =$$

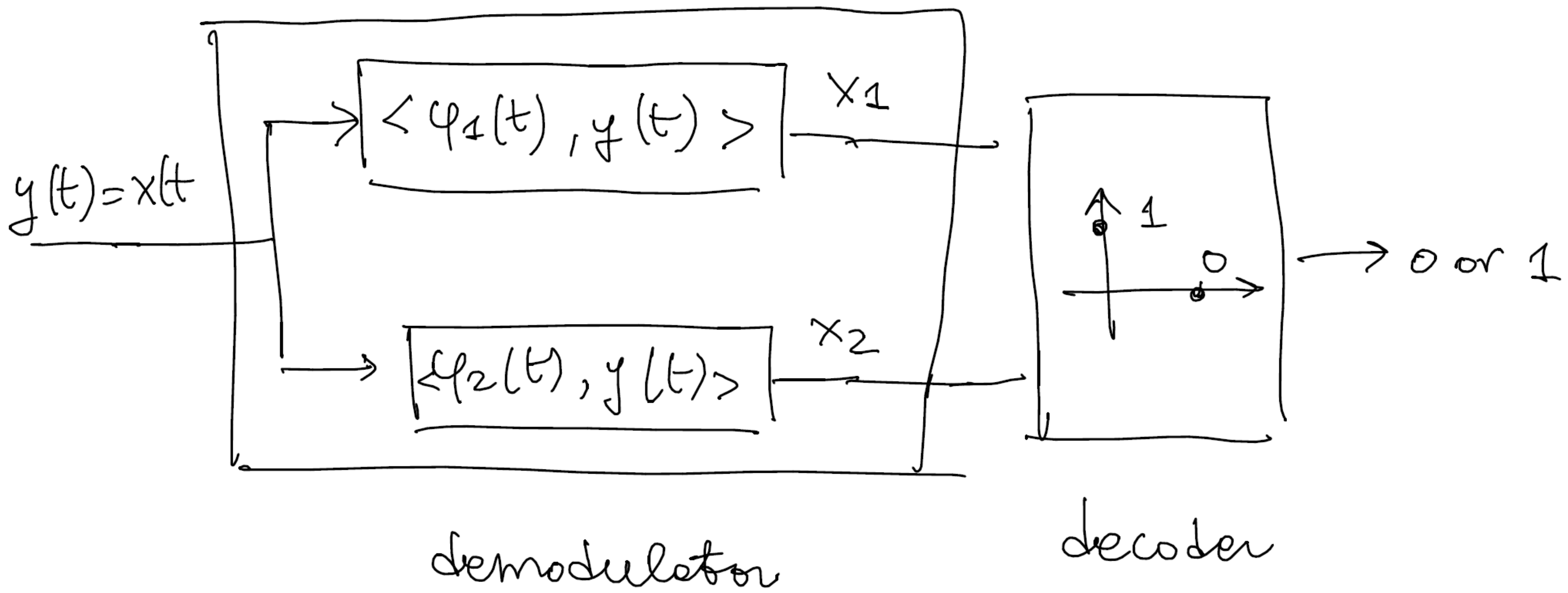


$$x_1(t) = \sqrt{E_b}\,\varphi_2(t) =$$



- What is the optimal receiver in the absence of noise?

How Can the Message Be Recovered Without Noise?



- Check that the receiver works when $m=0$.

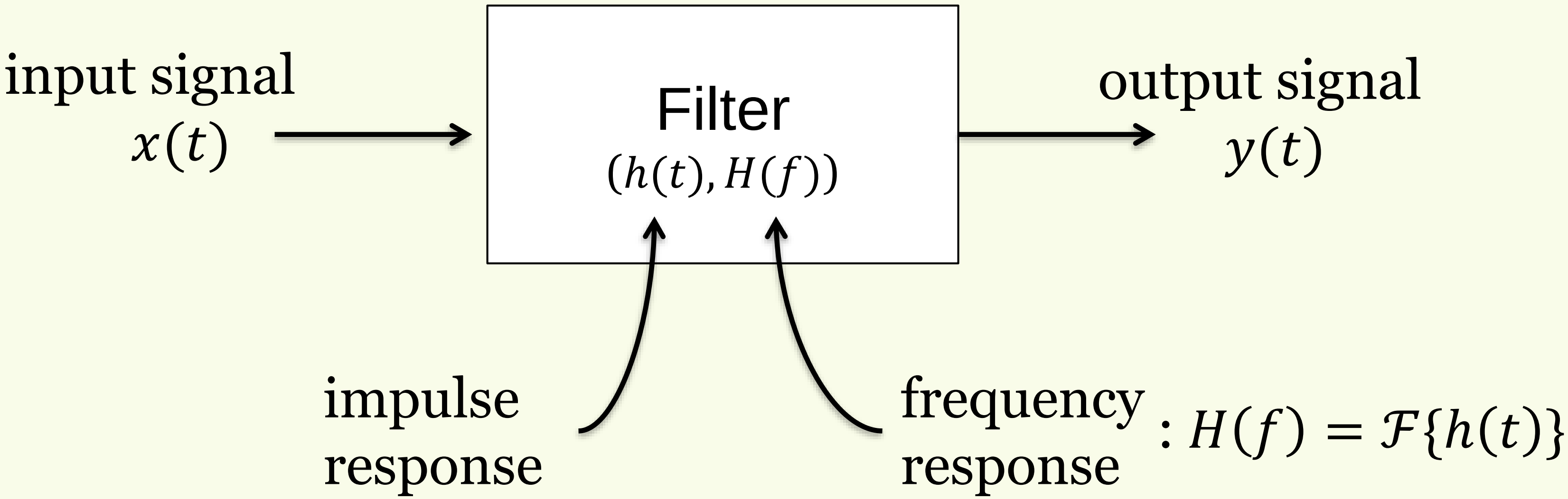
How Can the Message Be Recovered Without Noise?

If $m=0$, $y(t) = \sqrt{E_b}\,\varphi_1(t)$

$$\Rightarrow \quad \langle y(t), \varphi_1(t) \rangle = \sqrt{E_b}\,\|\varphi_1(t)\|^2 = \sqrt{E_b}$$

$$\langle y(t), \varphi_2(t) \rangle = \sqrt{E_b}\,\langle \varphi_1(t), \varphi_2(t) \rangle = 0$$

Filters (Linear Time Invariant Systems)



Filters (Linear Time Invariant Systems)

- In time domain:

$$y(t) = x(t) * h(t) = \int_{-\infty}^{+\infty} x(\lambda)h(t-\lambda)d\lambda$$
$$= \int_{-\infty}^{+\infty} h(\lambda)x(t-\lambda)d\lambda$$
$$= h(t) * x(t)$$

convolution

- In frequency domain

$$Y(f) = H(f)\,X(f)$$

and hence

$$G_y(f) = |H(f)|^2 G_x\,(f)$$

output energy spectral density

input energy spectral density

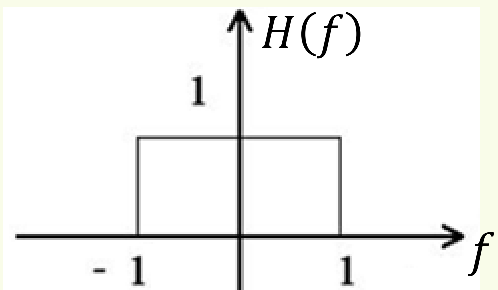
Filters (Linear Time Invariant Systems)

Examples:

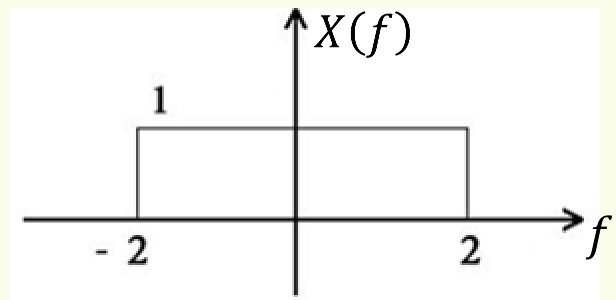
a) $h(t) = \delta(t)$ $\quad H(f) = 1$ $\quad Y(f) = X(f)$

$y(t) = x(t)$

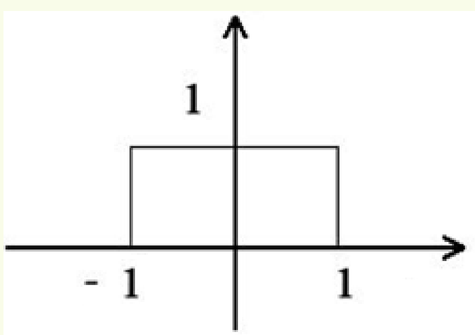
b)



(low-pass filter)



$x(t) = 4\ \mathrm{sinc}(4t)$

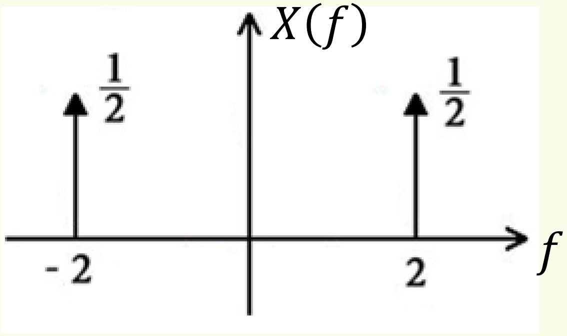


$Y(f) = H(f)X(f)$

and $y(t) = 2\ \mathrm{sinc}(2t)$

Filters (Linear Time Invariant Systems)

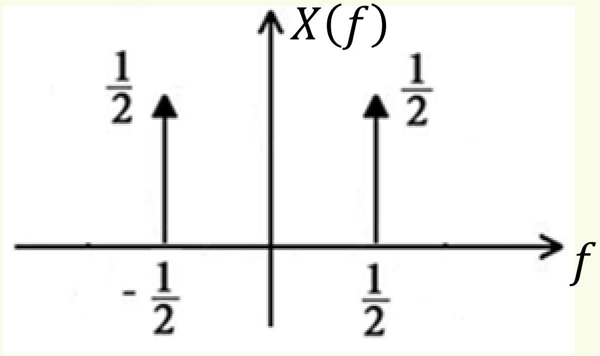
c) Same low-pass filter as above, but with input $x(t) = \cos(4\pi t)$



$\rightarrow$

$$Y(f) = H(f)\,X(f) = 0$$
$$y(t) = 0$$

d) Repeat for $x(t) = \cos(\pi t)$



$\rightarrow$

$$Y(f) = X(f)$$
$$y(t) = x(t)$$

How Can the Message Be Recovered Without Noise?

- Instead of using correlations, the demodulator can use the matched filter-based architecture.

- A filter with impulse response

$$h(t) = \varphi(-t)$$

is said to be matched to the waveform $\varphi(t)$.

How Can the Message Be Recovered Without Noise?

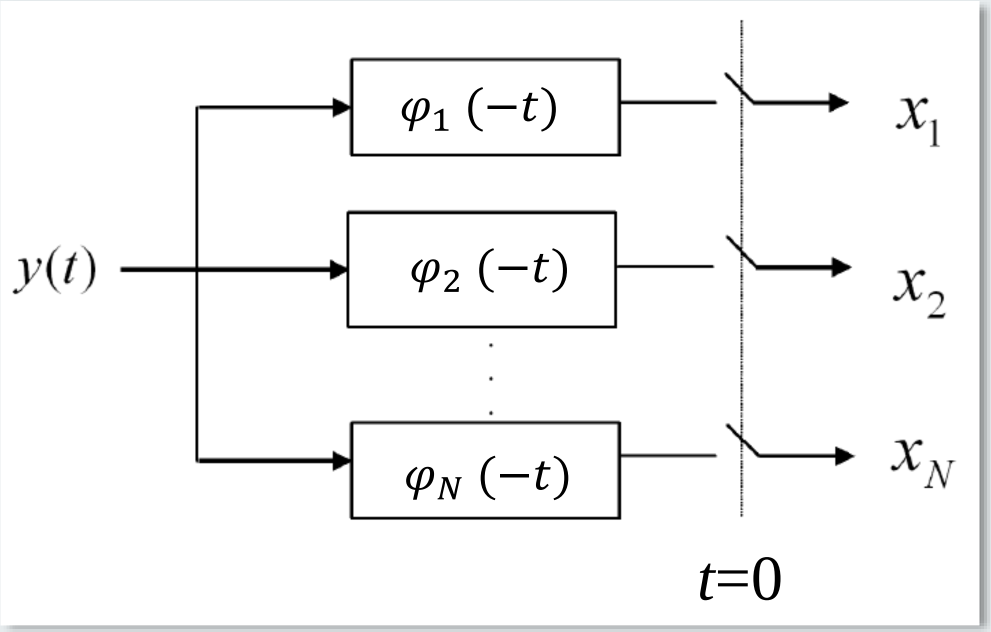
- Why "matched"?

$$\varphi(t) * h(t) = \varphi(t) * \varphi(-t) = \int_{-\infty}^{+\infty} \varphi(\lambda)\varphi(\lambda + t)d\lambda$$

- Hence, at t=0 (symbol peak), the matched filter recovers the energy of the pulse

$$\varphi(t) * h(t)|_{t=0} =< \varphi(t), \varphi(t) >= ||\varphi(t)||^2$$

How Can the Message Be Recovered Without Noise?



matched filter demodulator

How Can the Message Be Recovered Without Noise?

- Proof of equivalence between the two demodulators

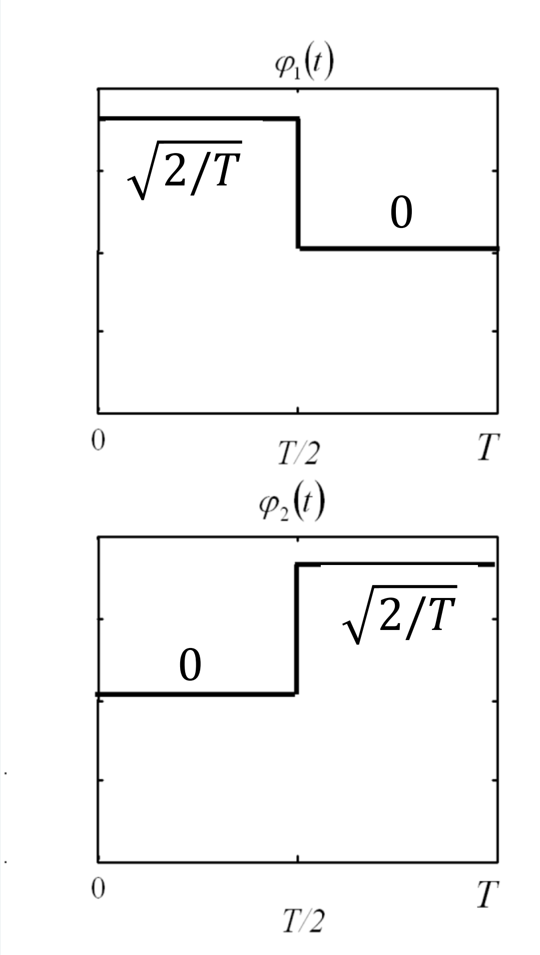
$$< x(t), \varphi_n(t) >= \int_{-\infty}^{+\infty} x(\lambda)\varphi_n(\lambda)d\lambda$$

$$= \int_{-\infty}^{+\infty} x(\lambda)\varphi_n(\lambda + t)d\lambda|_{t=0}$$

$$= x(t) * \varphi_n(-t)|_{t=0}$$

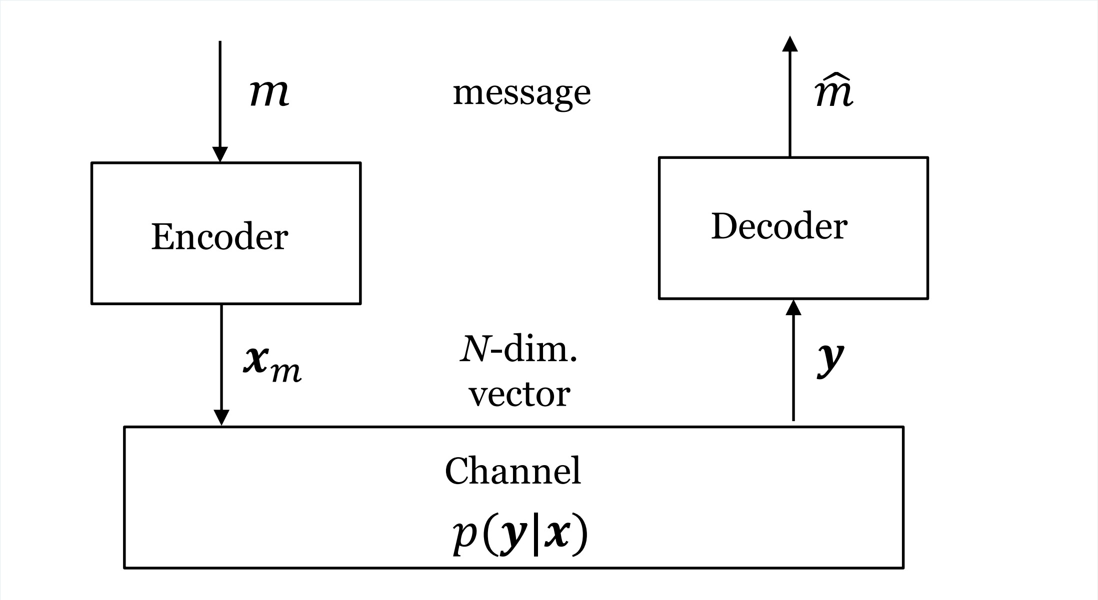
How Can the Message Be Recovered Without Noise?

- Quiz: Evaluate the matched filters for the waveforms below.



How to Detect on a Noisy Channel?

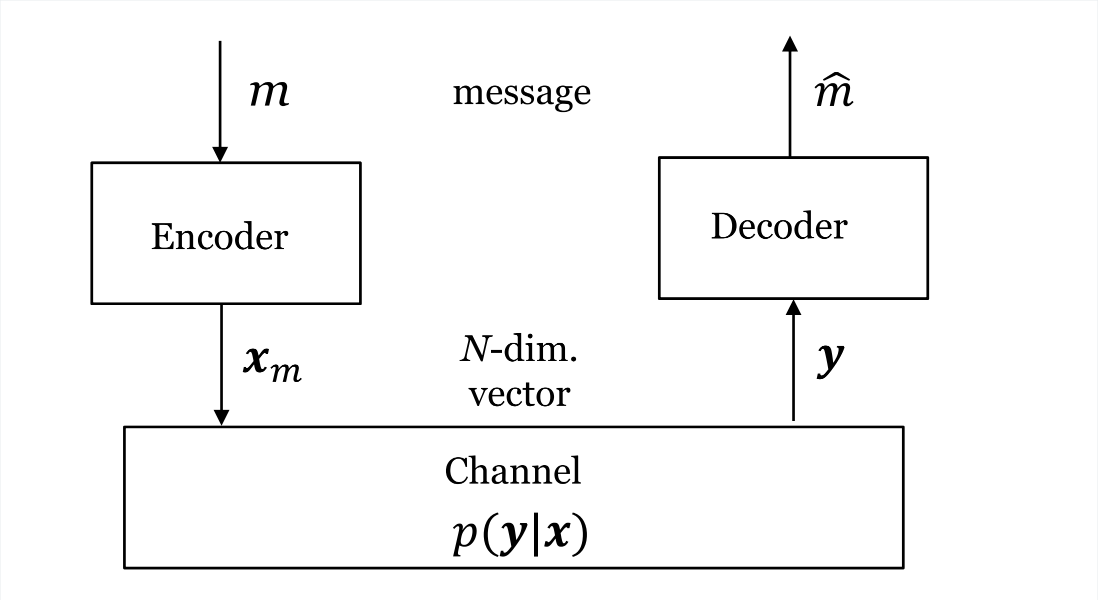
- Consider first the vector (or discrete) channel model



- The channel is described by the conditional probability distribution $p(\boldsymbol{y}|\boldsymbol{x})$

How to Detect on a Noisy Channel?

- Consider first the vector (or discrete) channel model



- The channel is described by the conditional probability distribution $p(\boldsymbol{y}|\boldsymbol{x})$

- Ex.: Gaussian channel, N=1

$$y = x + z \quad \text{with} \quad z \sim \mathcal{N}\left(0, \frac{N_0}{2}\right)$$

Gaussian Distribution

- Gaussian distribution

variance

$$X \sim N(m, \sigma^2)$$

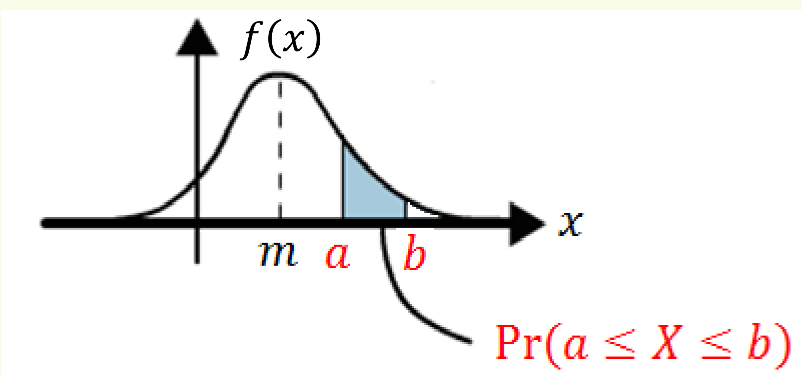
mean

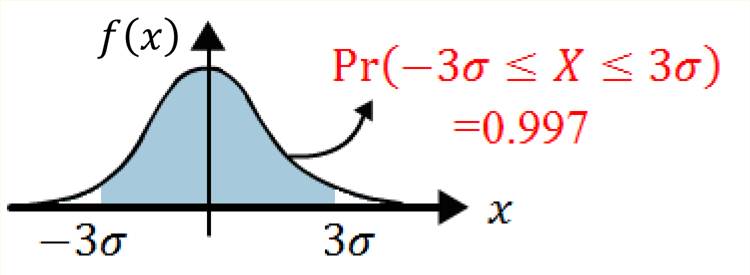
distributed as

Gaussian (Normal)

- Probability density function: $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$

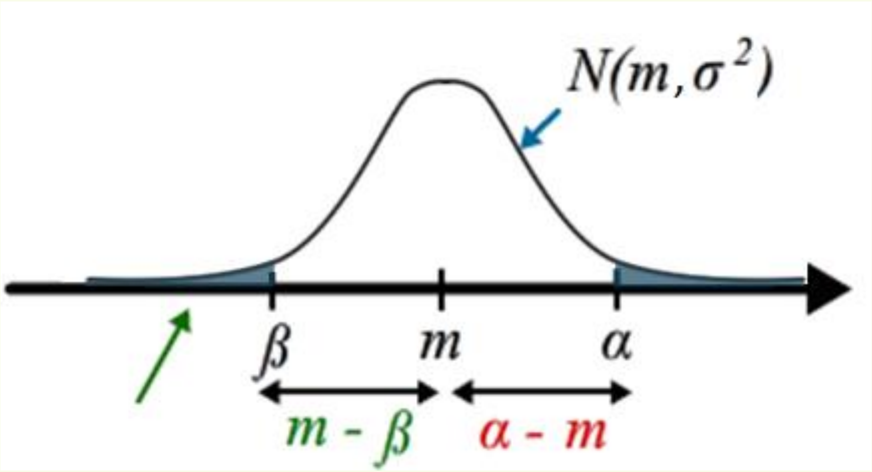
- Computing probabilities:





Gaussian Distribution

- Tail probabilities:

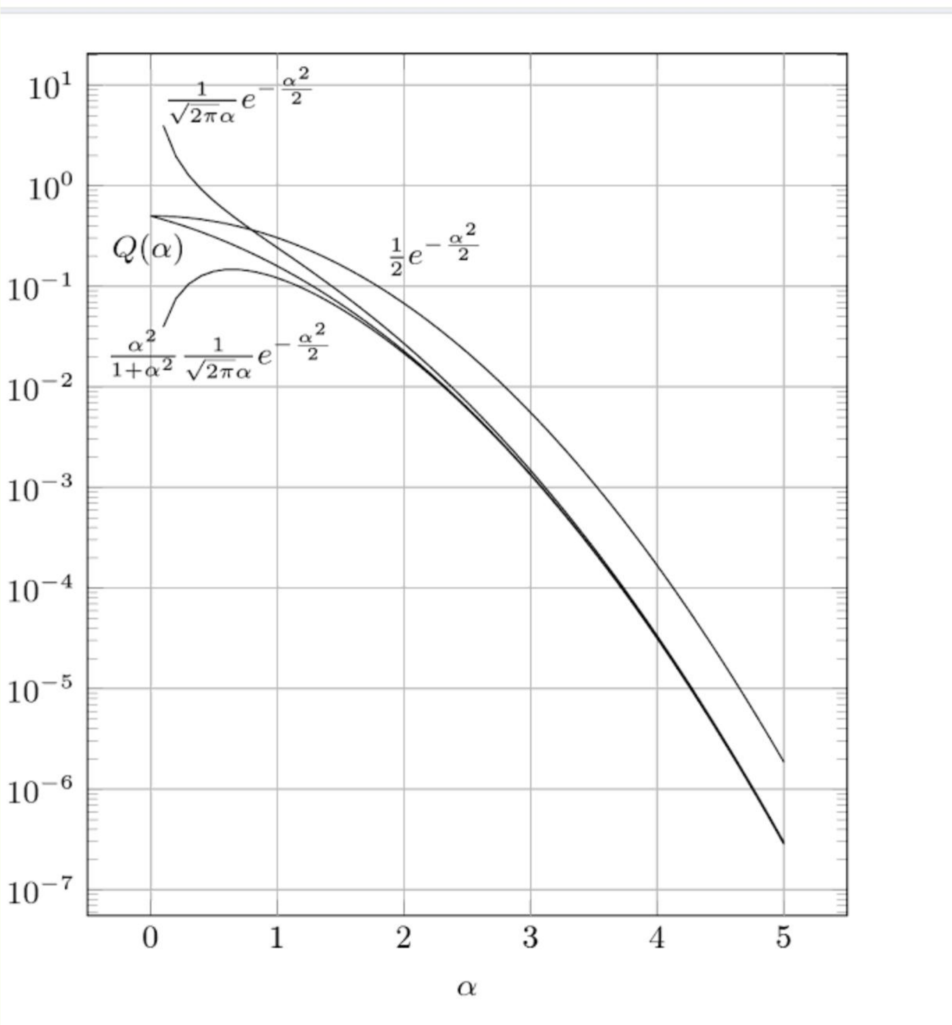


$$\int_{a}^{+\infty} f(x)dx = \mathrm{Q}\left(\frac{a-m}{\sigma}\right)$$

$$\int_{-\infty}^{\beta} f(x)dx = Q\left(\frac{m-\beta}{\sigma}\right)$$

Gaussian Distribution

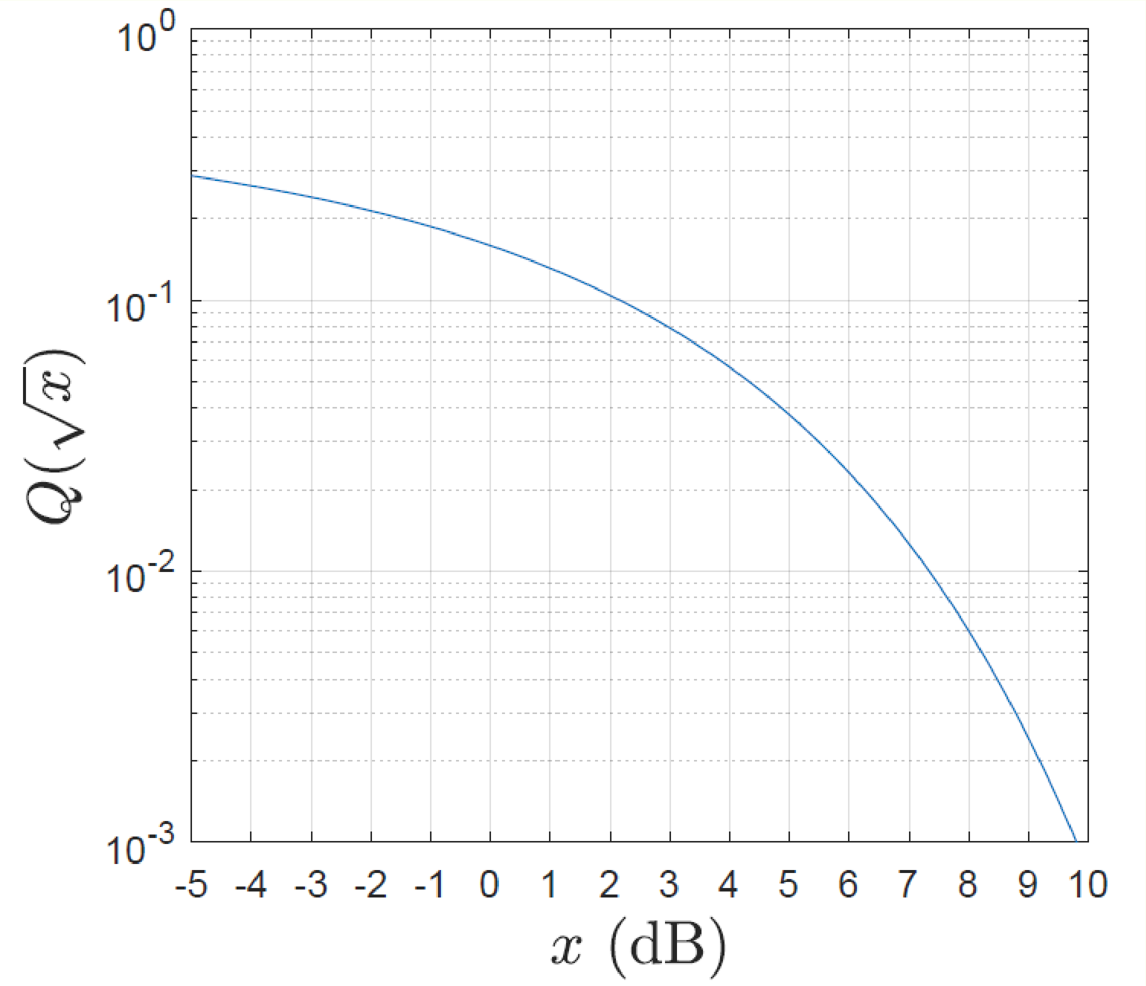
- Q function



[B. Rimoldi, Principles of Digital Communications, Cambridge University Press]

Gaussian Distribution

- Q function



[B. Rimoldi, Principles of Digital Communications, Cambridge University Press]

Digression: How to Design a Fire Alarm?

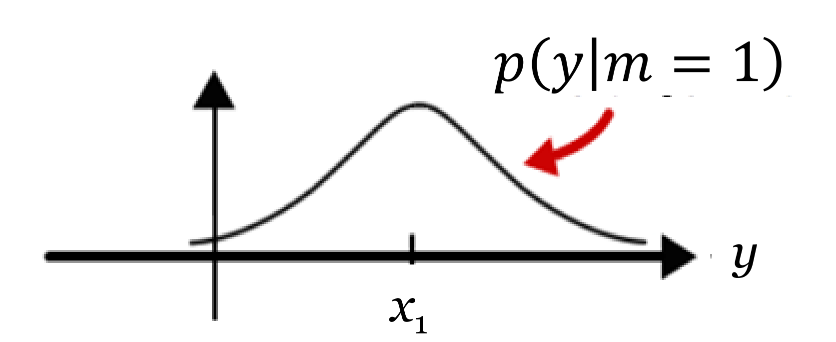
- Sensor measuring temperature

- If temperature > threshold γ -> alarm;
  if temperature < threshold γ -> no alarm

Digression: How to Design a Fire Alarm?

- Define m=0/1 (no fire/fire) and y=temperature

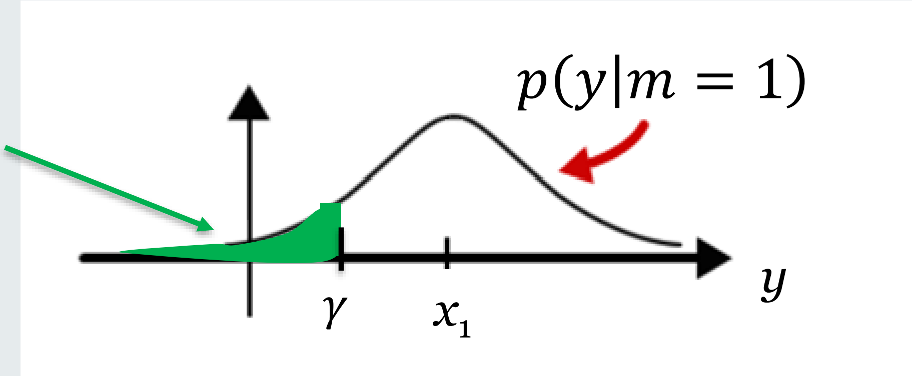
- Conditional probability density function (pdf)



x_1 = average temperature when there is fire

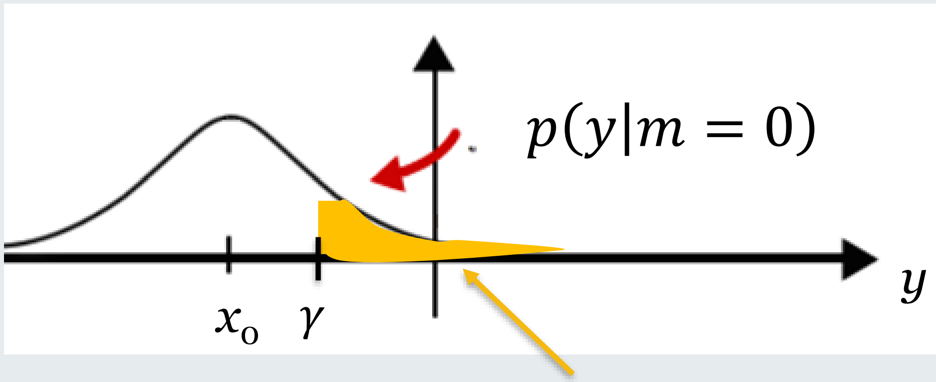
- Conditional probability of error:

$\Pr[\hat{m} \neq m | m = 1]$



Digression: How to Design a Fire Alarm?

- Conditional probability of error:



x_o = average temperature when there is no fire

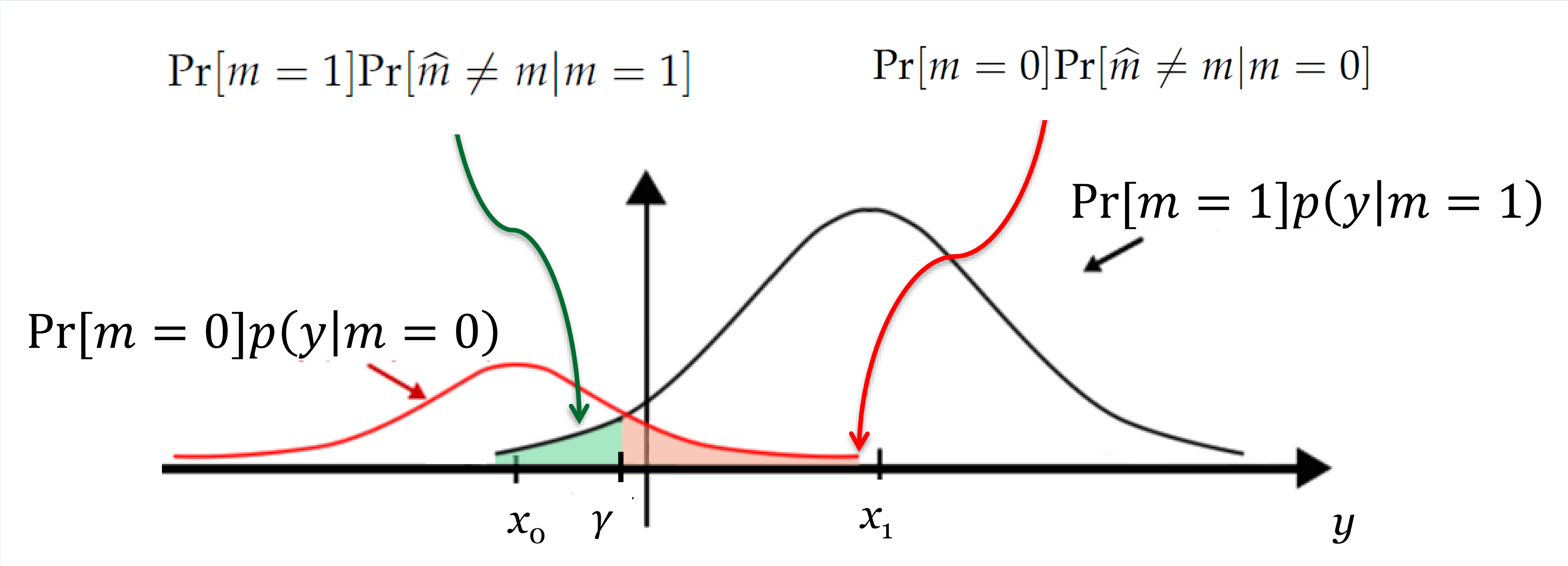
$$\Pr[\hat{m} \neq m | m = 0]$$

- We are interested in the overall probability of error.

- By the law of total probability

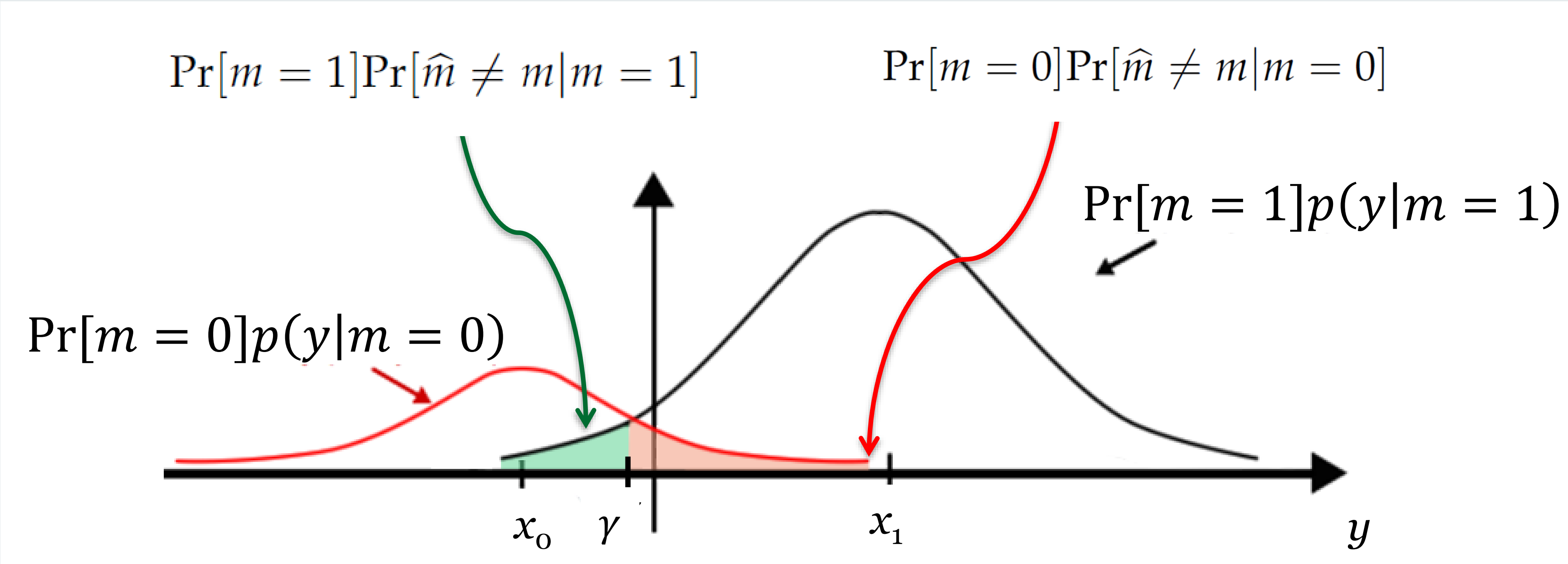
$$\begin{aligned} P_e &= \Pr[\hat{m} \neq m] \\ &= \Pr[m = 1]\Pr[\hat{m} \neq m | m = 1] + \Pr[m = 0]\Pr[\hat{m} \neq m | m = 0] \end{aligned}$$

Digression: How to Design a Fire Alarm?



- Minimizing P_e is equivalent to minimizing the colored area
- Intuitive arguments show that the optimal threshold should be such that the two curves take the same value.

Digression: How to Design a Fire Alarm?



- This yields the optimal rule

$$p(y|m=0)\Pr[m=0] \gtrless p(y|m=1)\Pr[m=1]$$

Digression: How to Design a Fire Alarm?

- Optimal rule

$$p(y|m=0)\Pr[m=0] \gtrless p(y|m=1)\Pr[m=1]$$

a priori probabilities

Digression: How to Design a Fire Alarm?

- Optimal rule

$$p(y|m=0)\Pr[m=0] \gtrless p(y|m=1)\Pr[m=1]$$

likelihood of each message for the received data y

a priori probabilities

a priori (from prior knowledge) × likelihood (from data)

Digression: How to Design a Fire Alarm?

- Optimal rule

$$\log\left(p(y|m=0)\right) + \log(\Pr[m=0]) \gtrless \log(p(y|m=1)) + \log(\Pr[m=1])$$

log-likelihood of each message for the received data y

log-a priori probabilities

log-prior (from prior knowledge) + log-likelihood (from data)

Digression: How to Design a Fire Alarm?

- The derived optimal rule is known as Maximum a Posterior (MAP).

- Posterior probability of the message given the received signal (Bayes theorem)

$$p(m|y) = \frac{p(y|m)p(m)}{p(y)}$$

- The optimal rule can be then equivalently expressed as

$$\begin{aligned}\hat{m} &= \underset{m\in\{0,1\}}{\operatorname{argmax}}\ p(m|y)\\ &= \underset{m\in\{0,1\}}{\operatorname{argmax}}\ p(y|m)p(m)\end{aligned}$$

which justifies the name MAP.

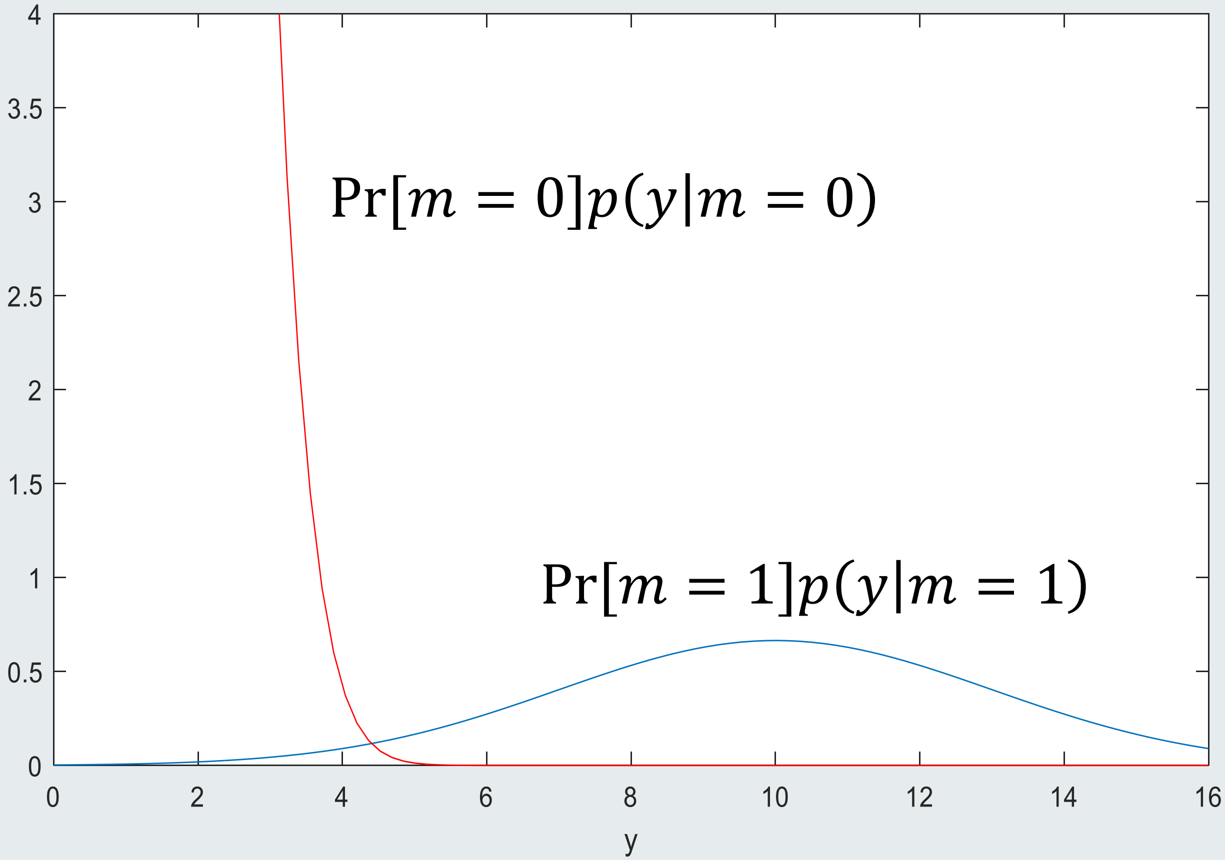
Digression: How to Design a Fire Alarm?

Example:

$$\Pr[m=1] = 0.05$$

$$p(y|m=1) = \mathcal{N}(y|10,3)$$

$$p(y|m=0) = \mathcal{N}(y|1,1)$$



Digression: How to Design a Fire Alarm?

- If $y = 4$, should the alarm go off?

First, note that

$$\log \mathcal{N}(x \mid \mu, \sigma^2) = \log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) - \frac{(x-\mu)^2}{2\sigma^2}$$

Therefore, the MAP rule can be written as

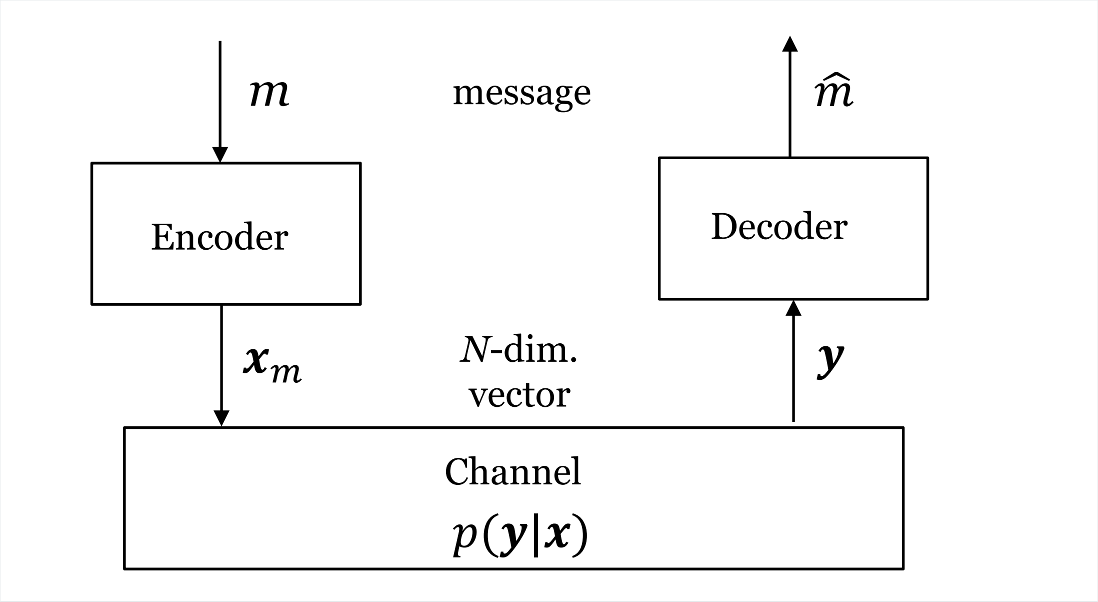
$$\log(0.05) + \log\left(\frac{1}{\sqrt{2\pi 3}}\right) - \frac{(4-10)^2}{6}$$

$$\gtrless \log(0.95) + \log\left(\frac{1}{\sqrt{2\pi 1}}\right) - \frac{(4-1)^2}{2}$$

$$\Rightarrow -7.01 < -5.47 \Rightarrow \text{no fire } (\hat{m} = 0)$$

- Repeat for $y = 5$

How to Detect on a Noisy Channel?



- Problem: Choose a decision rule y→ $\hat{m}$ that minimizes the probability of error

How to Detect on a Noisy Channel?

- Using again the law of total probability, we have

$$\begin{aligned}\min_{\hat{m}(\mathbf{y})} P_e &= \min_{\hat{m}(\mathbf{y})} \Pr[\widehat{m} \neq m] \\ &= \min_{\hat{m}(\mathbf{y})} \int p(\mathbf{y}) \Pr[m \neq \widehat{m}(\mathbf{y})|\mathbf{y}] d\mathbf{y} \text{ (by the law of total prob.)} \\ &= \int p(\mathbf{y}) \left(\min_{\hat{m}(\mathbf{y})} \Pr[m \neq \widehat{m}(\mathbf{y})|\mathbf{y}] \right) d\mathbf{y}\end{aligned}$$

- And hence the optimal decoding rule is

$$\min_{\hat{m}(\mathbf{y})} \Pr[m \neq \widehat{m}(\mathbf{y})|\mathbf{y}] \Longleftrightarrow \max_{\hat{m}(\mathbf{y})} \Pr[m = \widehat{m}(\mathbf{y})|\mathbf{y}]$$

How to Detect on a Noisy Channel?

- It follows that the decision that minimizes the probability of error P_e is the MAP rule

$$\hat{m}(\mathbf{y}) = \underset{m \in \{0,\ldots,M-1\}}{\operatorname{argmax}} \; p(\mathbf{x}_m | \mathbf{y})$$

where the posterior probability of m is

$$p(\mathbf{x}_m|\mathbf{y}) = \frac{p(m)p(\mathbf{y}|\mathbf{x}_m)}{p(\mathbf{y})} \propto p(m)p(\mathbf{y}|\mathbf{x}_m)$$

a priori (from prior knowledge) × likelihood (from data)

- As seen above, the rule can also be written in terms of log –probabilities.

How to Detect on a Noisy Channel?

- When $p(m)$=1/M, the MAP rule reduces to the Maximum Likelihood (ML) rule

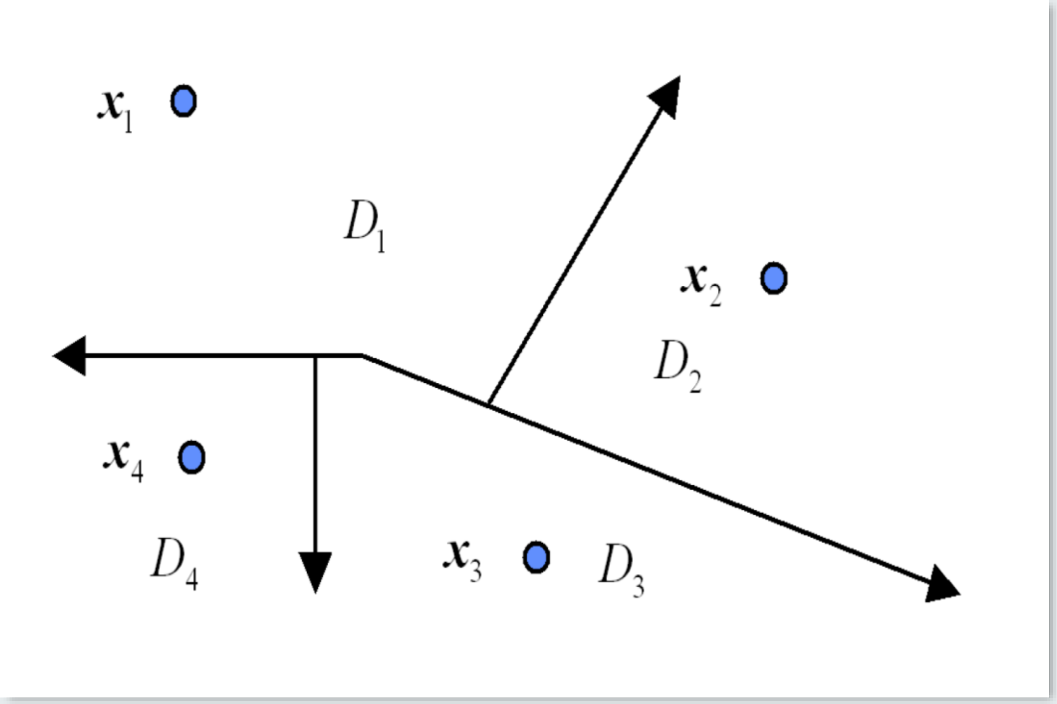
$$\hat{m}(\mathbf{y}) = \underset{m \in \{0,\ldots,M-1\}}{\operatorname{argmax}} \; p(\mathbf{y}|\mathbf{x}_m)$$

likelihood (from data)

How to Detect on a Noisy Channel?

- Both ML and MAP rules partition the space of the received signal $\mathbf{y}$ into decision regions, one for each message m

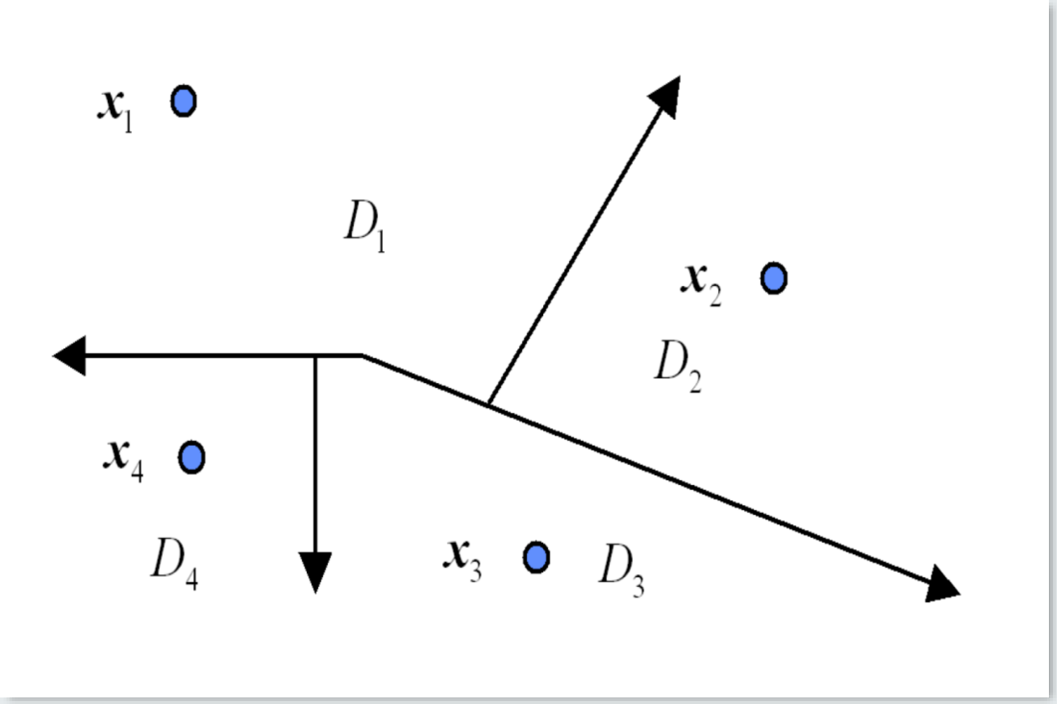
MAP: $$\mathcal{D}_m = \left\{ \mathbf{y} \in \mathbb{R}^N : \; m = \underset{m' \in \{0,\ldots,M-1\}}{\operatorname{argmax}} \; p(\mathbf{x}_{m'} \mid \mathbf{y}) \right\}$$



How to Detect on a Noisy Channel?

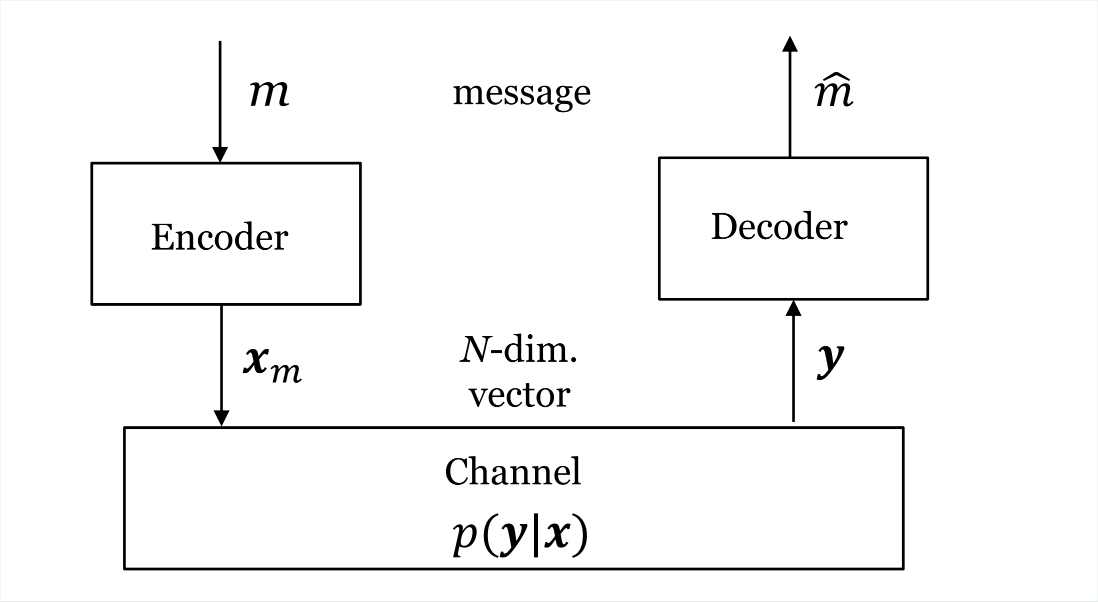
- Both ML and MAP rules partition the space of the received signal $\mathbf{y}$ into decision regions, one for each message m

ML: $$\mathcal{D}_m = \left\{ \mathbf{y} \in \mathbb{R}^N : \; m = \underset{m' \in \{0,\ldots,M-1\}}{\operatorname{argmax}} \; p(\mathbf{y} \mid \mathbf{x}_{m'}) \right\}$$



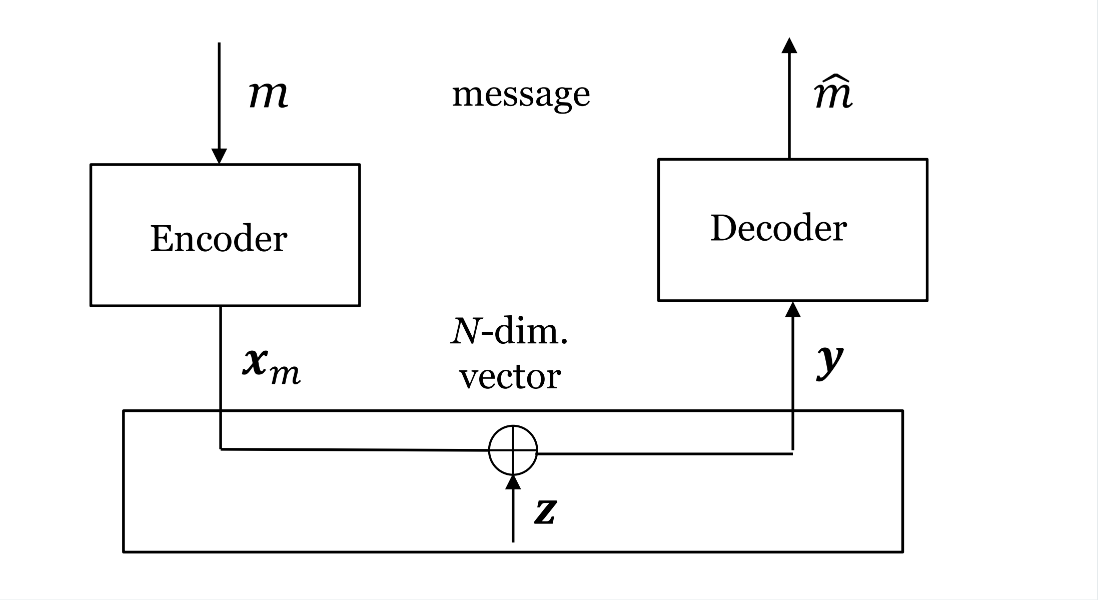
How to Detect on a Noisy Channel?

- Consider now the important case of the additive Gaussian noise channel



How to Detect on a Noisy Channel?

- Consider now the important case of the additive Gaussian noise channel



$\mathbf{y} = \mathbf{x} + \mathbf{z}$ with $z_n \sim \mathcal{N}\left(0, \frac{N_0}{2}\right)$ i.i.d.

How to Detect on a Gaussian Channel?

- The likelihood of the message m for data $\mathbf{y}$ is

$$
\begin{aligned}
p(\mathbf{y}|\mathbf{x}_m) = p_{\mathbf{z}}(\mathbf{y} - \mathbf{x}_m) &= \frac{1}{(\pi N_0)^{\frac{N}{2}}} \exp\left(-\frac{1}{N_0} \sum_{n=1}^{N} (y_n - x_{m,n})^2 \right) \\
&= \frac{1}{(\pi N_0)^{\frac{N}{2}}} \exp\left(-\frac{1}{N_0} \|\mathbf{y} - \mathbf{x}_m\|^2 \right)
\end{aligned}
$$

and hence the log-likelihood is

$$
\log p(\mathbf{y}|\mathbf{x}_m) = -\frac{1}{N_0} \|\mathbf{y} - \mathbf{x}_m\|^2 + \text{const}
$$

How to Detect on a Gaussian Channel?

- For the Gaussian channel, the ML rule is simply a minimum distance rule

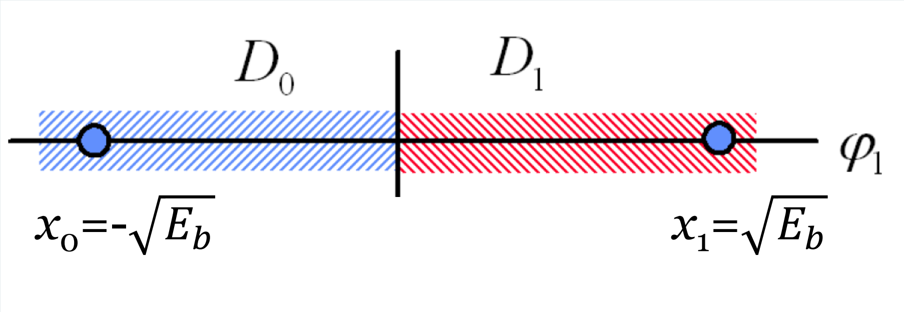
$$\hat{m}(\mathbf{y}) = \underset{m \in \{0,\ldots,M-1\}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{x}_m\|^2$$

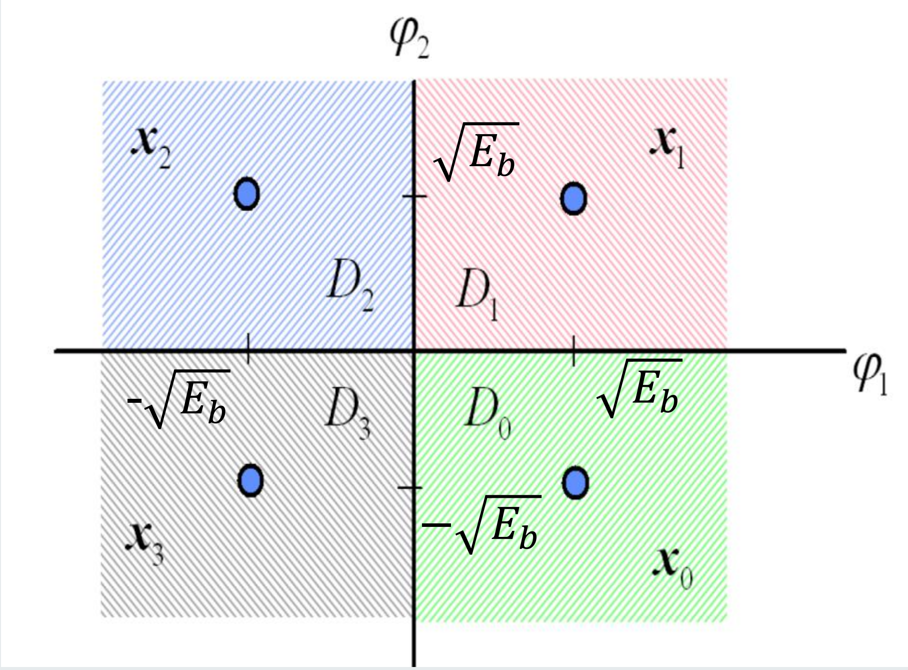
and hence the decision regions are

$$\mathcal{D}_m = \left\{ \mathbf{y} \in \mathbb{R}^N : \, m = \underset{m' \in \{0,\ldots,M-1\}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{x}_{m'}\|^2 \right\}$$

How to Detect on a Gaussian Channel?

Examples: N=1 N=2



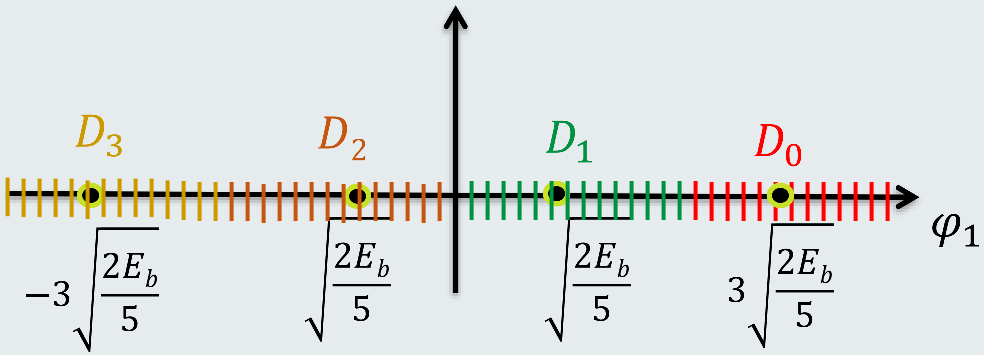


M=2
BPSK

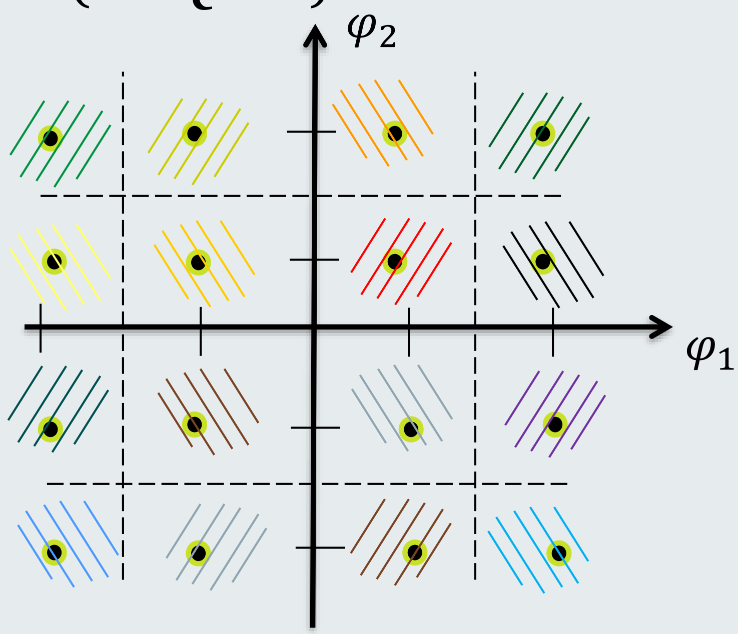
M=4
4-PSK

How to Detect on a Gaussian Channel?

Example: N=1, M=4 (4-PAM)



Example: N=2, M=16 (16-QAM)

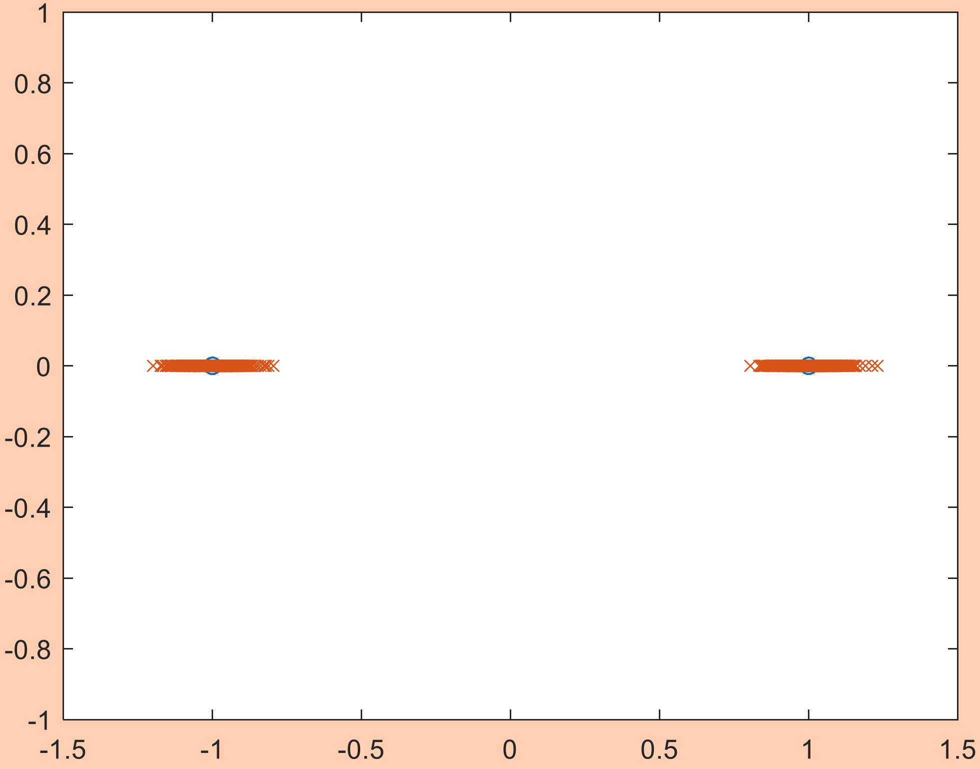


Matlab: Implementing an ML Threshold Detector

BPSK transmission

```
%parameters
Eb=1;
N0=0.01;%noise variance – try changing this parameter!
L=1000; %number of bits

%simulation
m=randi(2,L,1)-1; %generate independent bits
x=sqrt(Eb)*2*(m-1/2); %generate signal vector
plot(x,zeros(size(x)),'o'); %plot transmitted constellation points
z=randn(L,1)*sqrt(N0/2); %generate noise
y=x+z; %received signal
hold on; plot(y,zeros(size(y)), 'x');
mhat=(sign(y)+1)*1/2; %decoded bits
error_rate=sum(m~=mhat)/L
```

Matlab: Implementing an ML Threshold Detector

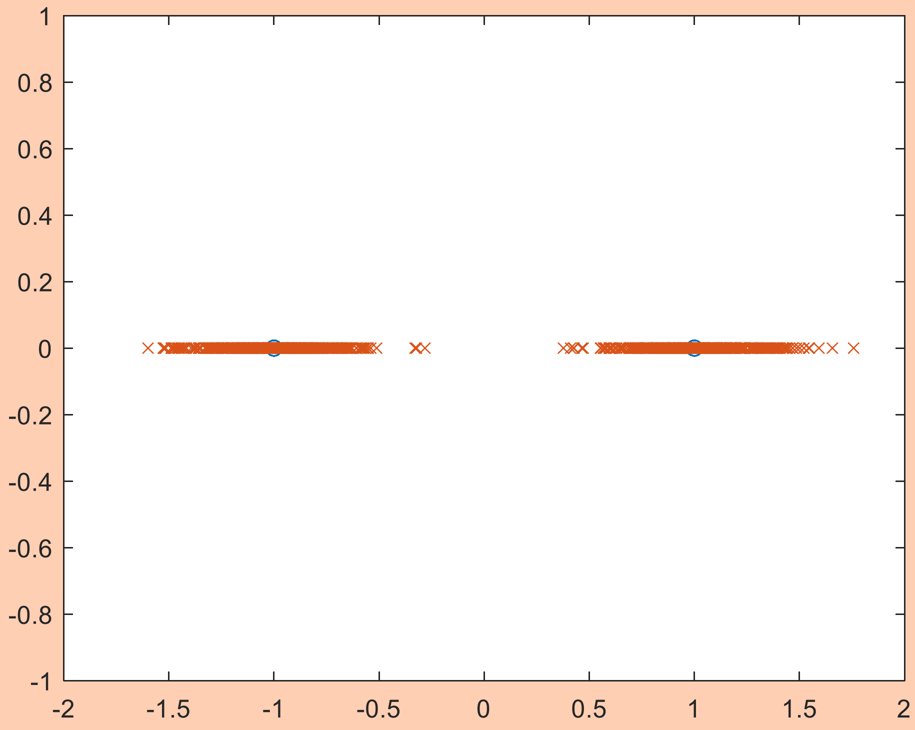


error_rate =

0

Matlab: Implementing an ML Threshold Detector

N0=0.1

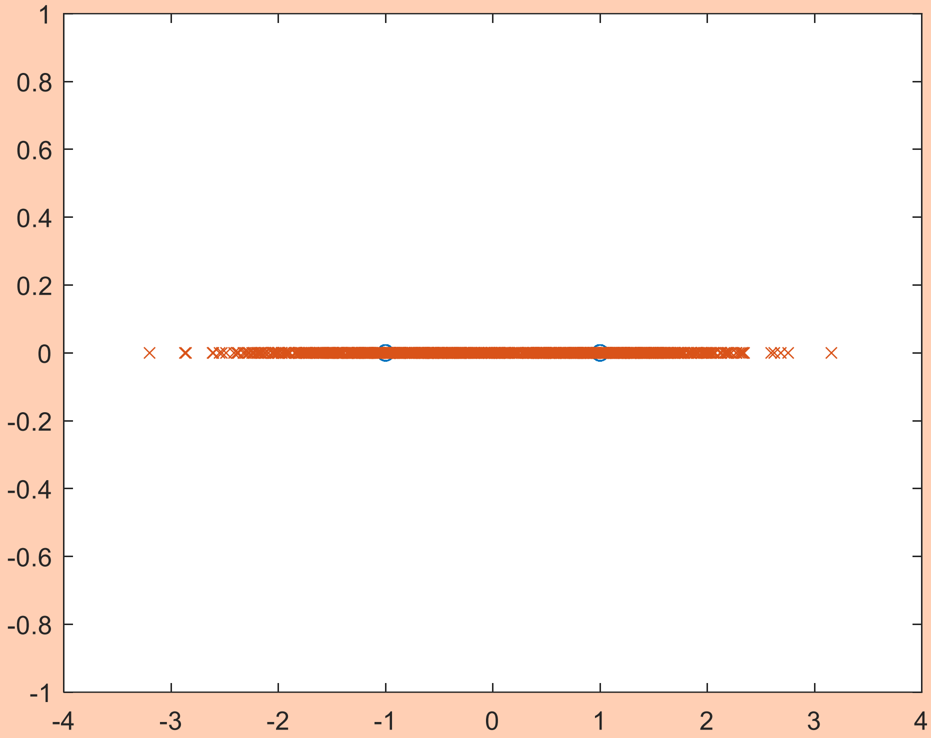


error_rate =

0

Matlab: Implementing an ML Threshold Detector

N0=1



error_rate =

0.0630

How to Detect on a Gaussian Channel?

- For the Gaussian channel, the MAP rule is

$$\hat{m}(\mathbf{y}) = \underset{m \in \{0,\ldots,M-1\}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{x}_m\|^2 - N_0 \log(p(m))$$

and the decision regions are accordingly defined.

- Example: For BPSK, the new threshold becomes (try to prove it!)

$$\frac{N_0}{4\sqrt{E_b}} \log\left(\frac{p(0)}{p(1)}\right)$$

How to Detect on a Gaussian Channel?

- Using the equality

$$\|\mathbf{y}-\mathbf{x}\|^2 = \|\mathbf{y}\|^2 + \|\mathbf{x}\|^2 - 2\langle \mathbf{x}, \mathbf{y}\rangle$$

we can rewrite the MAP rule also in terms of correlations only as

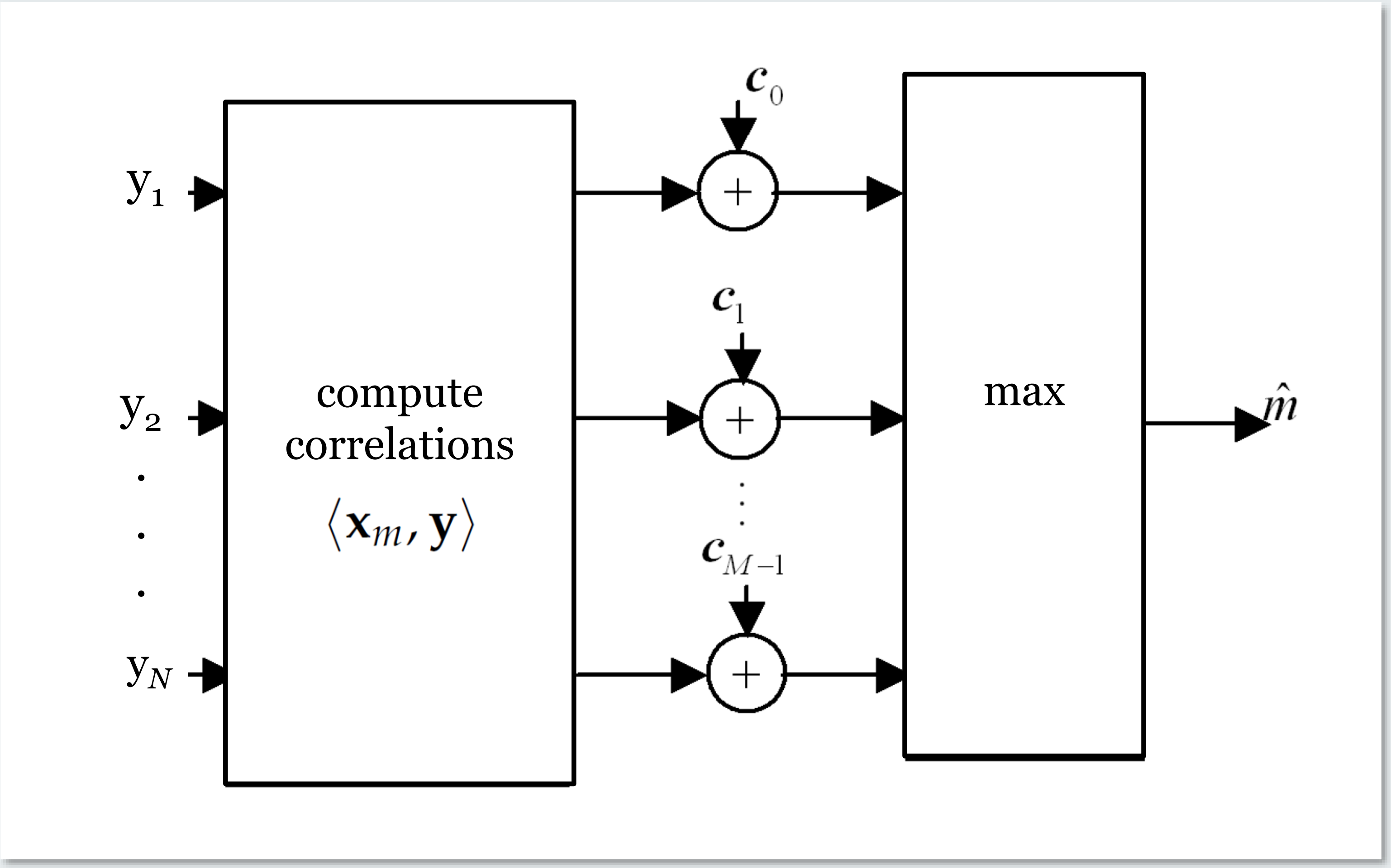
$$\hat{m}(\mathbf{y}) = \underset{m\in\{0,\ldots,M-1\}}{\operatorname{argmax}} \langle \mathbf{x}_m, \mathbf{y}\rangle + c_m$$

where

$$c_m = \frac{N_0}{2}\log(p(m)) - \frac{\|\mathbf{x}_m\|^2}{2}$$

How to Detect on a Gaussian Channel?

- Block diagram of a MAP decoder



correlative decoder

Matlab: Implementing an ML Correlative Decoder

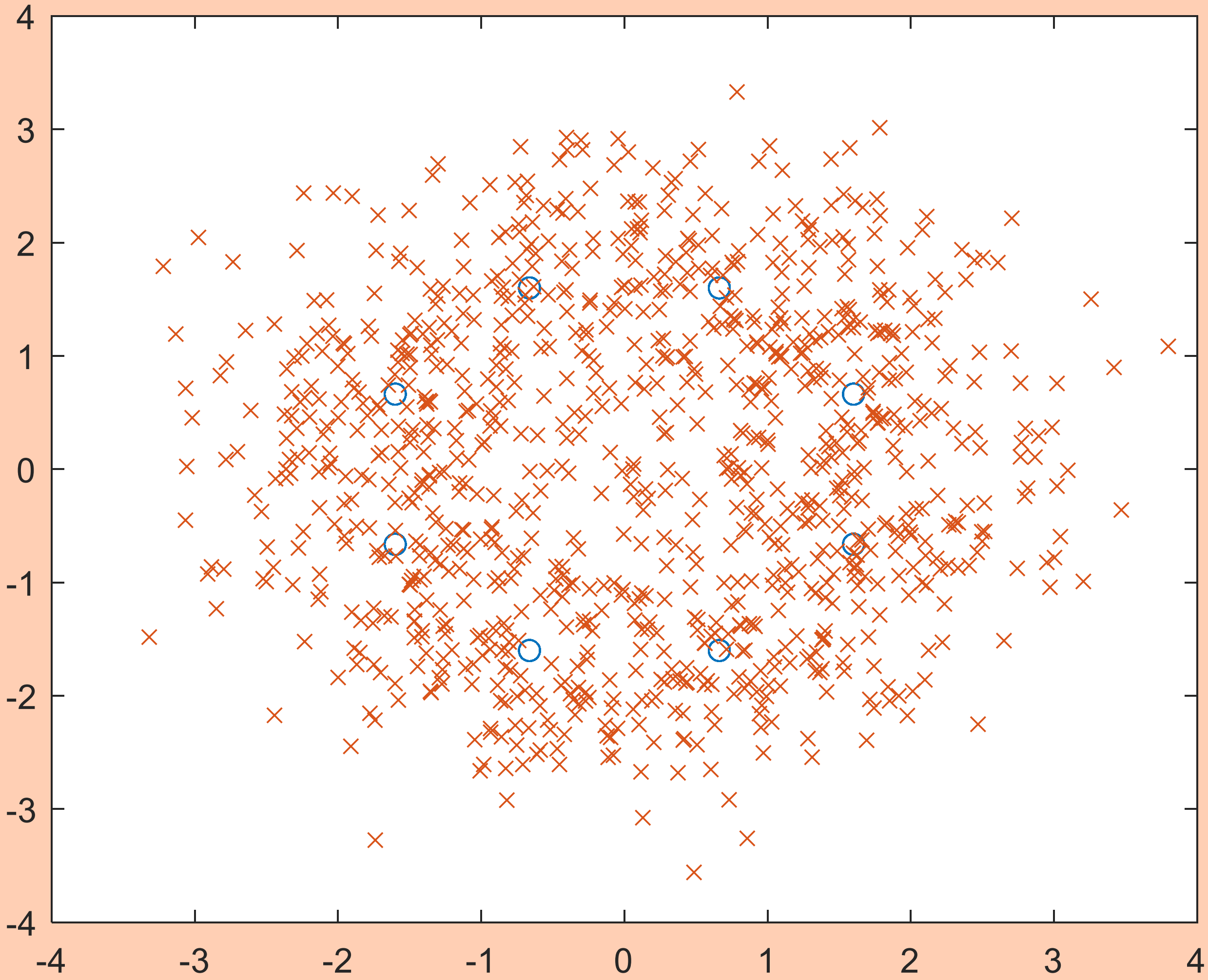
8-PSK transmission

```
%parameters
Eb=1;
N0=1;%noise variance – try changing this parameter!
L=1000; %number of symbols

%simulation
m=randi(8,L,1)-1; %generate independent symbols
x(:,1)= sqrt(3*Eb)*cos(pi*(2*m+1)/8);
x(:,2)= sqrt(3*Eb)*sin(pi*(2*m+1)/8);
%generate signal vector
plot(x(:,1),x(:,2),'o'); %plot transmitted constellation points
z=randn(L,2)*sqrt(N0/2); %generate noise
y=x+z; %received signal
hold on; plot(y(:,1),y(:,2), 'x');
```

Matlab: Implementing an ML Correlative Decoder

```
Xmat(:,1)= sqrt(3*Eb)*cos(pi*(2*[0:7]+1)/8);
Xmat(:,2)= sqrt(3*Eb)*sin(pi*(2*[0:7]+1)/8);

for l=1:L %for each transmitted symbol
score=Xmat*y(l,:)'; %compute the score
[smax,imax]=max(score); %find the maximum score
mhat(l)=imax-1;
end

error_rate=sum(m~=mhat')/L
```

Matlab: Implementing an ML Correlative Decoder



error_rate =

0.3440

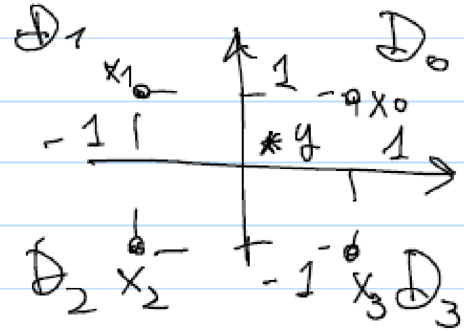
How to Detect on a Gaussian Channel?

Example: QPSK, $E_b = 1$ J, $N_0 = 0.3$ J

$$p(m) = \frac{1}{4}, \quad m = 0, 1, 2, 3$$

- What is the output of the optimal decoder of $y = \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix}$?

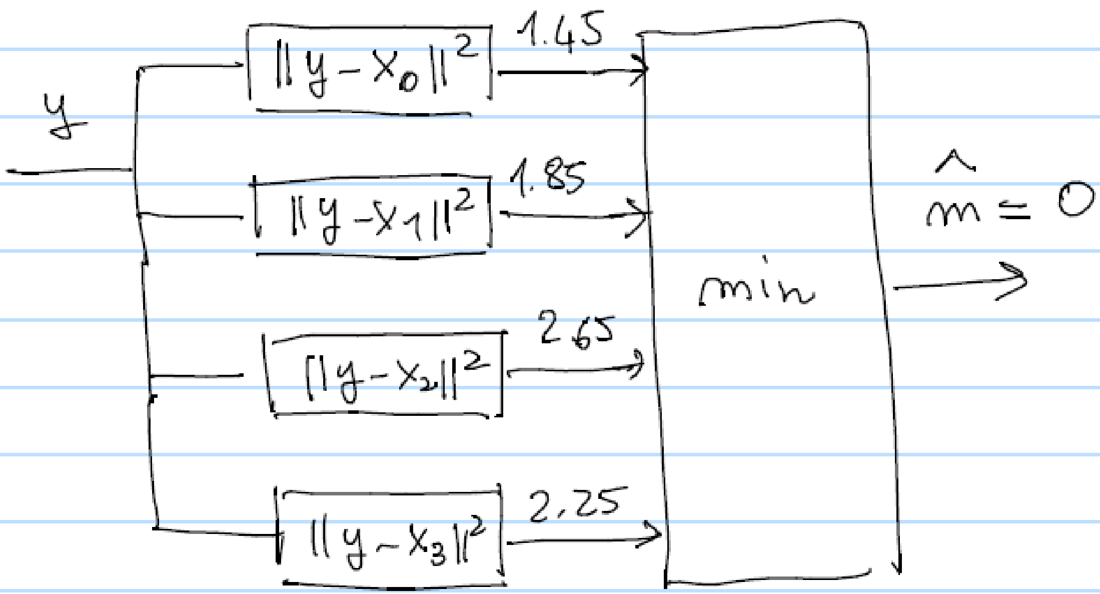
Solution 1: Since we have equiprobable messages, the decision regions of the optimal ML decoder can be easily obtained:



$$\Rightarrow \hat{m} = 0$$

How to Detect on a Gaussian Channel?

Solution 2 : Apply ML decoder using minimum distance decoder:

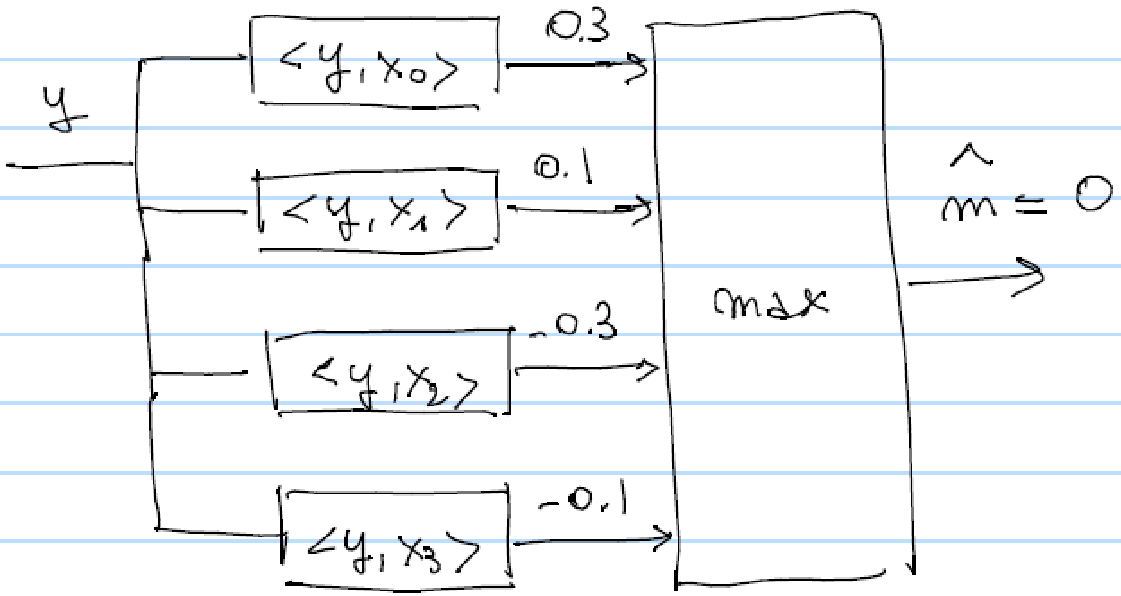


For example,

$$\|y-x_0\|^2 = (0.1-1)^2 + (0.2-1)^2 = 1.45$$

How to Detect on a Gaussian Channel?

Solution 3 : Apply ML decoder using correlative decoder:



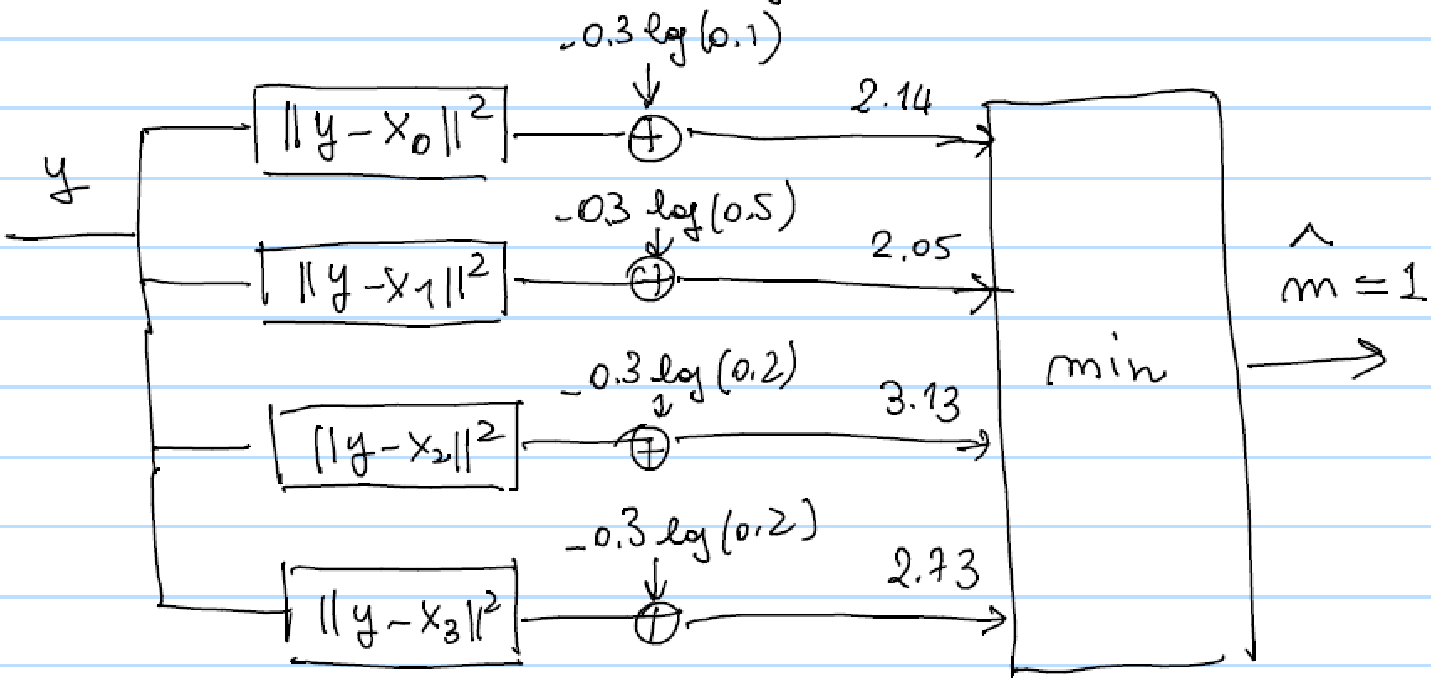
for example,

$$\langle y, x_0 \rangle = 0.1 \times 1 + 0.2 \times 1 = 0.3$$

How to Detect on a Gaussian Channel?

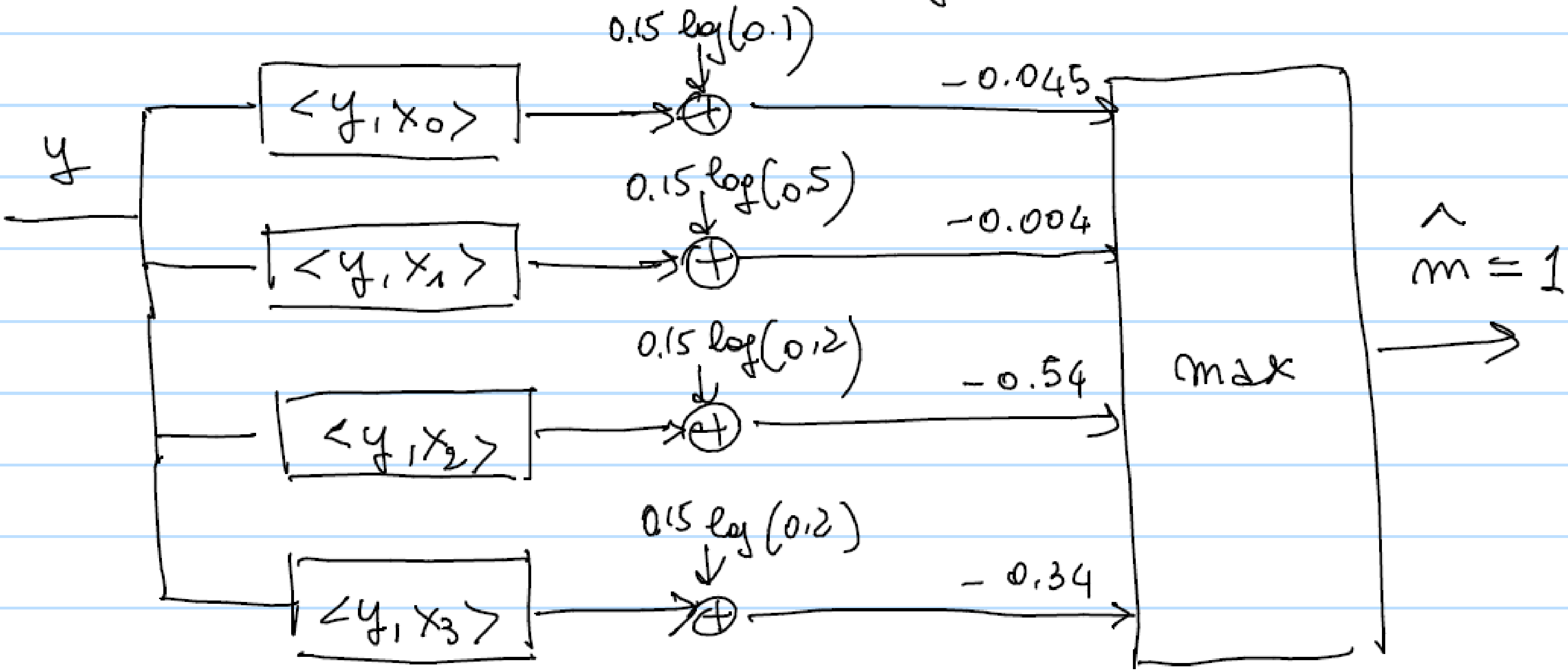
- Repeat with $p(0)=0.1$, $p(1)=0.5$, $p(2)=0.2$, $p(3)=0.2$

Solution 1 : Apply MAP decoder using minimum distance decoder:

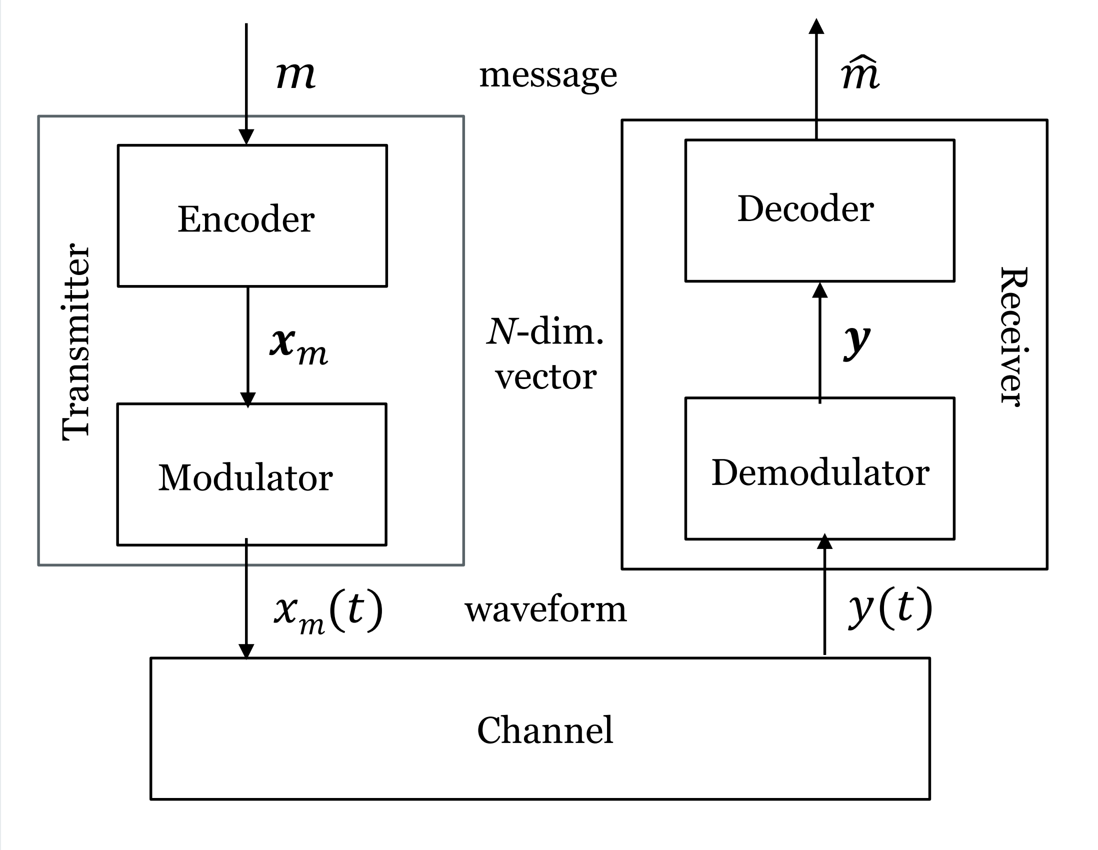


How to Detect on a Gaussian Channel?

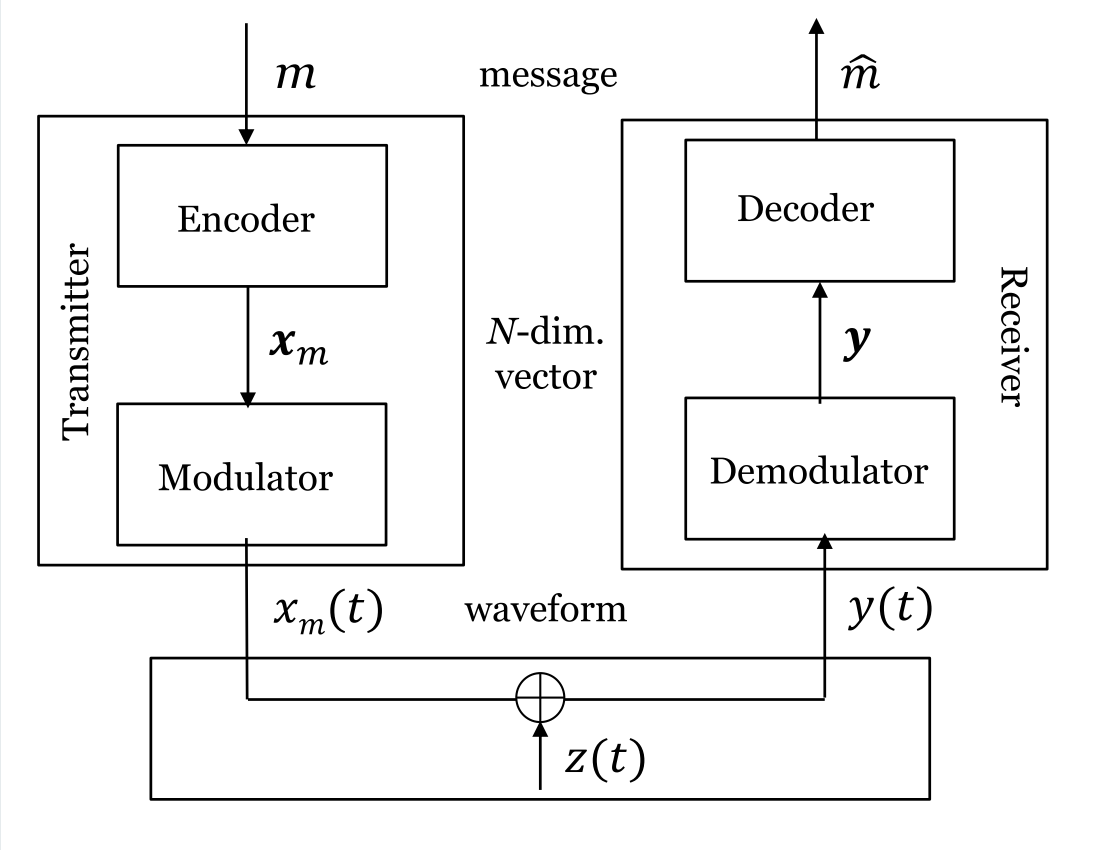
Solution 2 : Apply MAP decoder using correlative decoder:



Additive White Gaussian Noise Channel (AWGN)



Additive White Gaussian Noise Channel (AWGN)



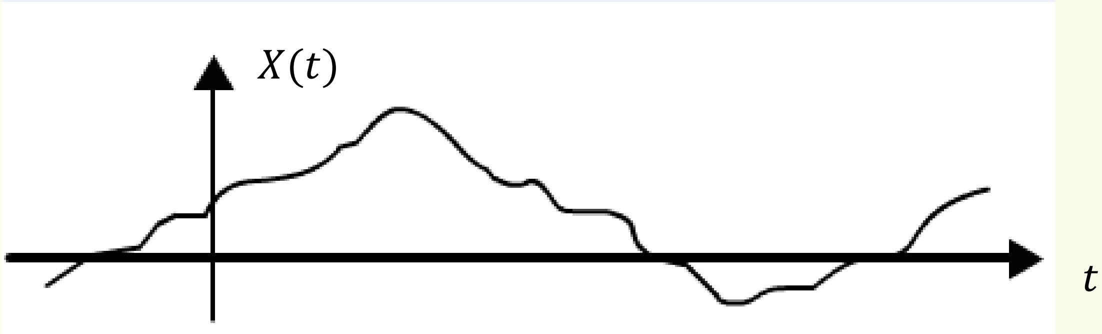
$$y(t) = x(t) + z(t)$$

with WGN $z(t)$:

$$\mathrm{E}[z(t)z(t-\tau)] = \frac{N_0}{2}\delta(\tau)$$

Random Processes

- A random process $X(t)$ is a collection of random variables indexed by time t.
- Another way to thinking about is in terms of random functions.
- We will be interested in stationary random processes: the probabilistic description of the process does not change with time.

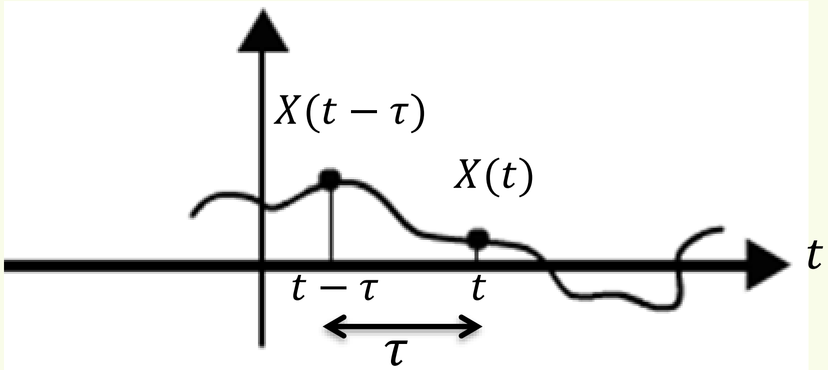


a realization of a random process

Random Processes

- Two key quantities that characterize a stationary random process $X(t)$ are:

1) Correlation Function
   - $R_x(\tau) = E[X(t)X(t-\tau)]$



- $R_x(\tau)$ tells us how predictable $X(t)$ is based on $X(t-\tau)$:

$$\uparrow |R_x(\tau)| \rightarrow \downarrow \text{ "randomness"}$$

Random Processes

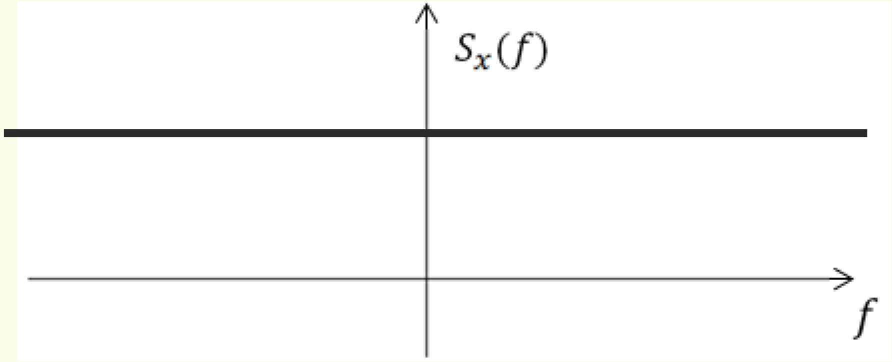
2) Power spectral density

- $S_x(f) = \mathcal{F}\{R_x(\tau)\}$
- Real and positive
- Symmetric $(S_x(f) = S_x(-f))$

Random Processes

- White Noise:

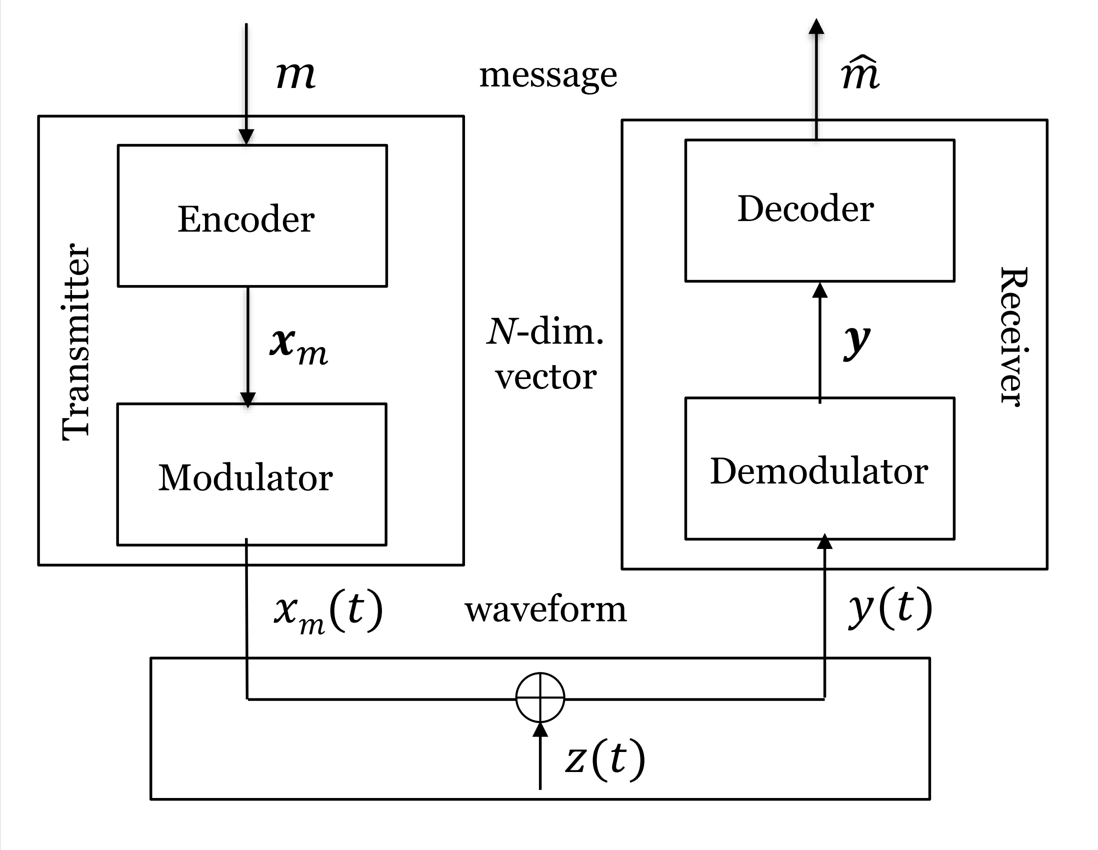
  Uniform power spectral density



  and impulsive correlation function

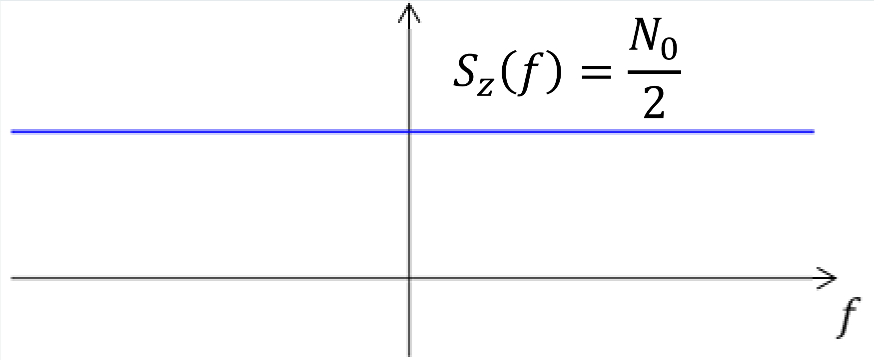
$$R_x(\tau) = \text{const} \times \delta(\tau)$$

Additive White Gaussian Noise Channel (AWGN)

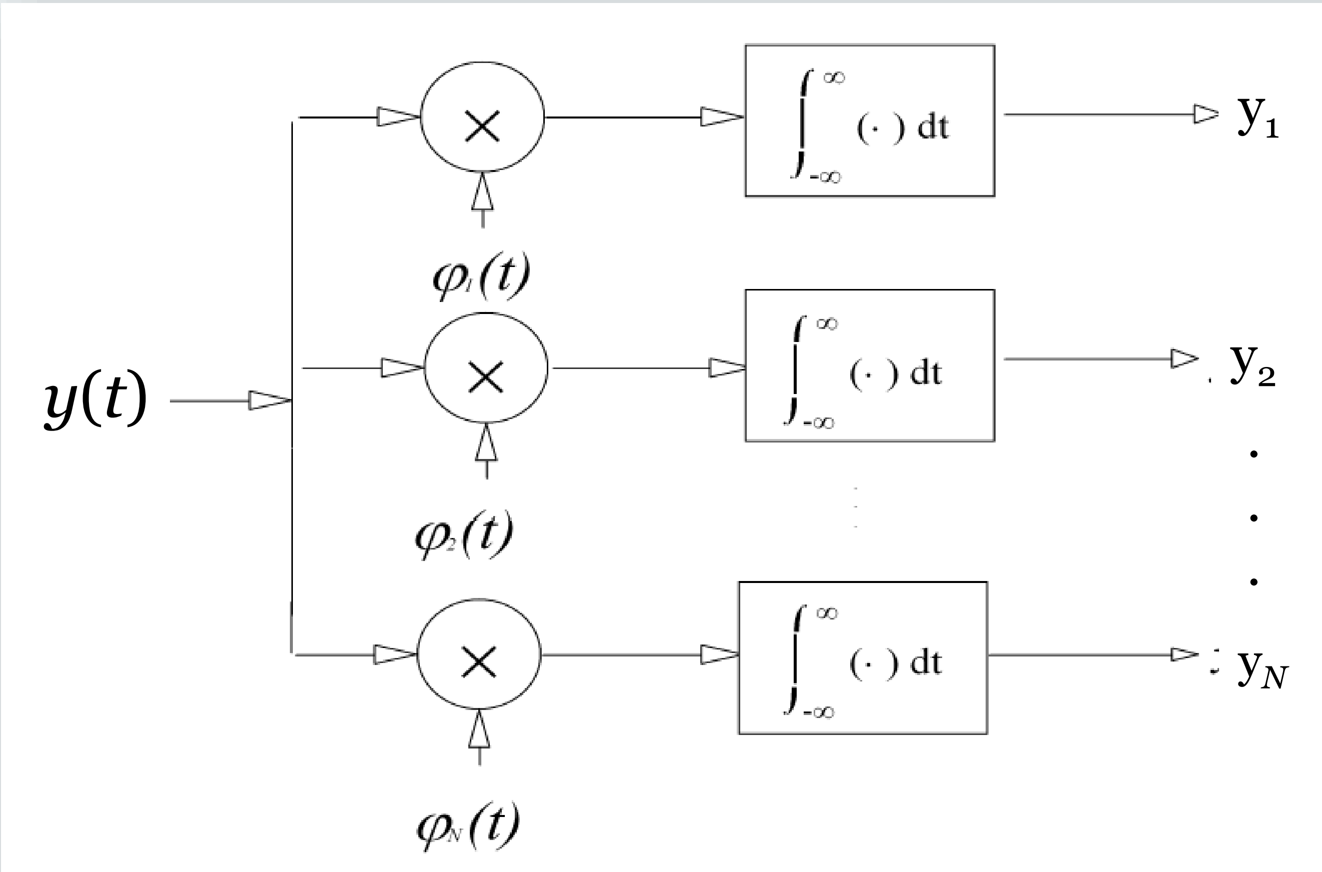


$$y(t) = x(t) + z(t)$$

with WGN $z(t)$: $$\mathrm{E}[z(t)z(t-\tau)] = \frac{N_0}{2}\delta(\tau)$$



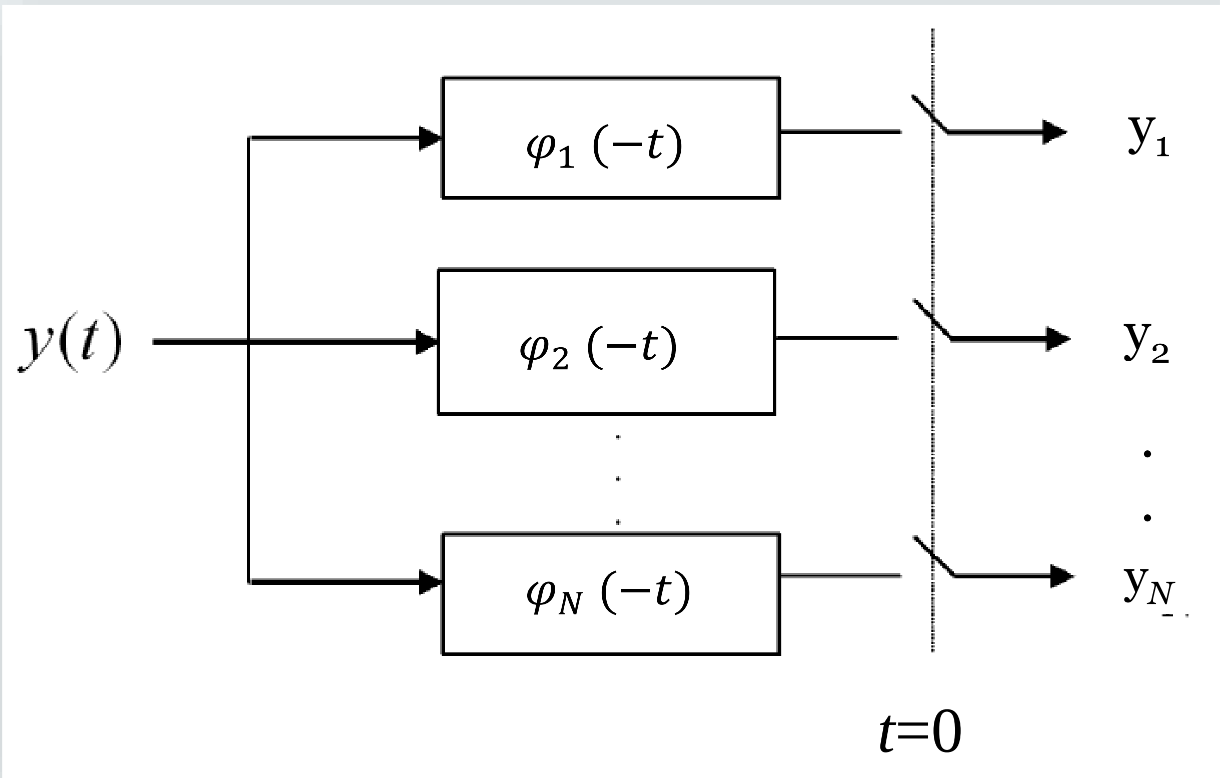
How to Demodulate on an AWGN Channel?



correlative demodulator

$$\langle y(t), \varphi_n(t) \rangle = y_n$$

How to Demodulate on an AWGN Channel?



$$\langle y(t), \varphi_n(t) \rangle = y_n$$

matched filter demodulator

How to Demodulate on an AWGN Channel?

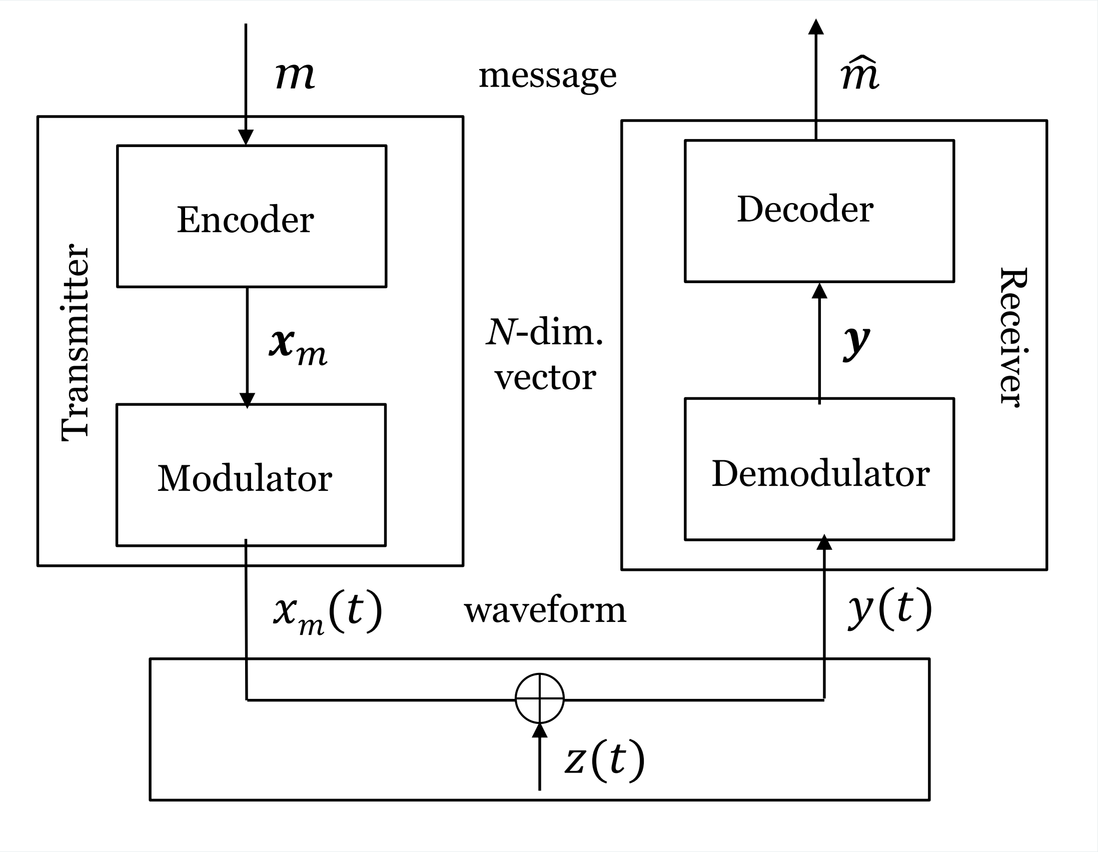
- Demodulated signal vector:

$$
\begin{aligned}
y_n = \langle y(t) = x(t) + z(t), \varphi_n(t) \rangle &= \langle x(t), \varphi_n(t) \rangle + \langle z(t), \varphi_n(t) \rangle \\
&= x_n + z_n
\end{aligned}
$$

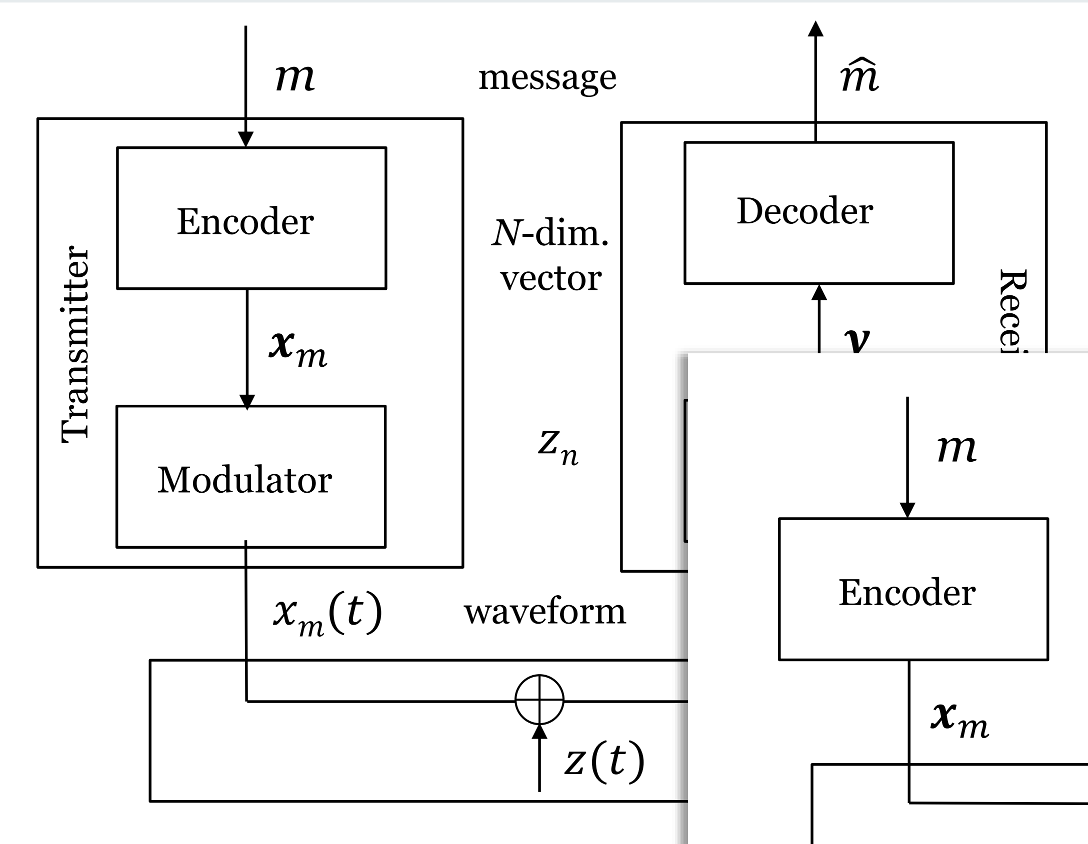
- Noise components are zero-mean Gaussian (linear combinations of Gaussian variables are Gaussian) with correlation:

$$
\begin{aligned}
\mathrm{E}[z_n z_{n'}] &= \mathrm{E}\left[\int \int z(t) z(t') \varphi_n(t) \varphi_{n'}(t') dt dt' \right] \\
&= \int \int \mathrm{E}[z(t) z(t')] \varphi_n(t) \varphi_{n'}(t') dt dt' \\
&= \frac{\mathrm{N}_0}{2} \int \varphi_n(t) \varphi_{n'}(t) dt \\
&= \frac{\mathrm{N}_0}{2} \delta_{nn'}
\end{aligned}
$$

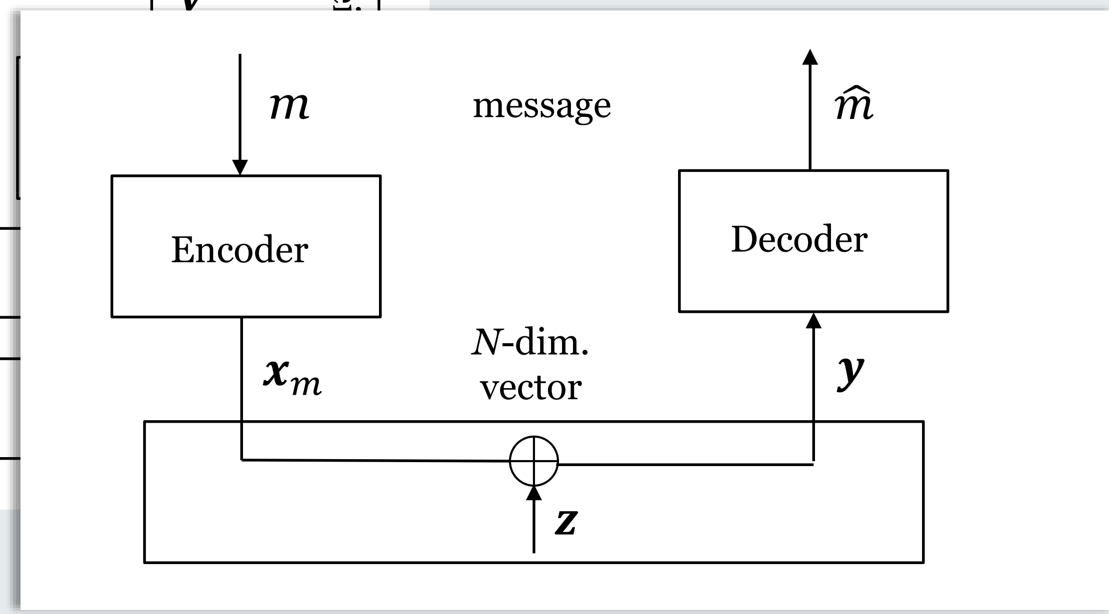
How to Demodulate on an AWGN Channel?



How to Demodulate on an AWGN Channel?



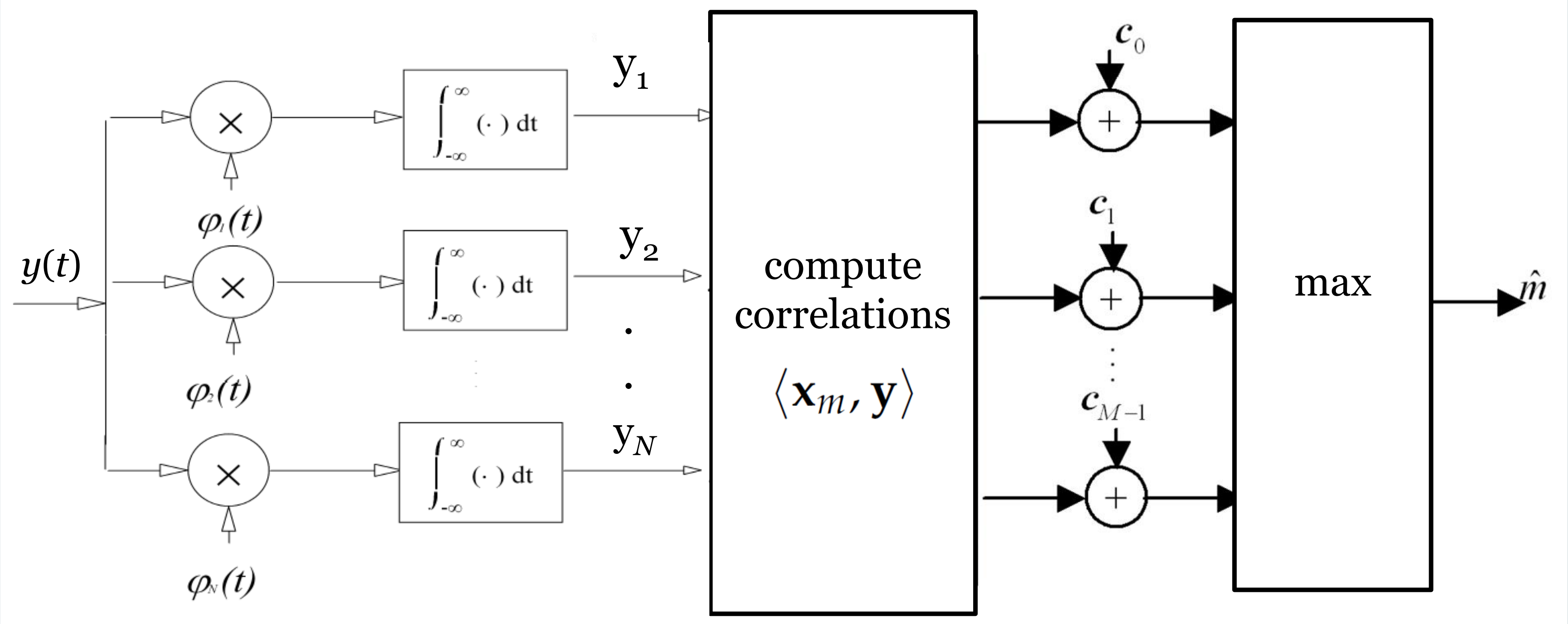
correlative/ matched filter demodulator



$\mathbf{y} = \mathbf{x} + \mathbf{z}$ with $z_n \sim \mathcal{N}\left(0, \frac{N_0}{2}\right)$ i.i.d.

How Does an Optimal Receiver Work?

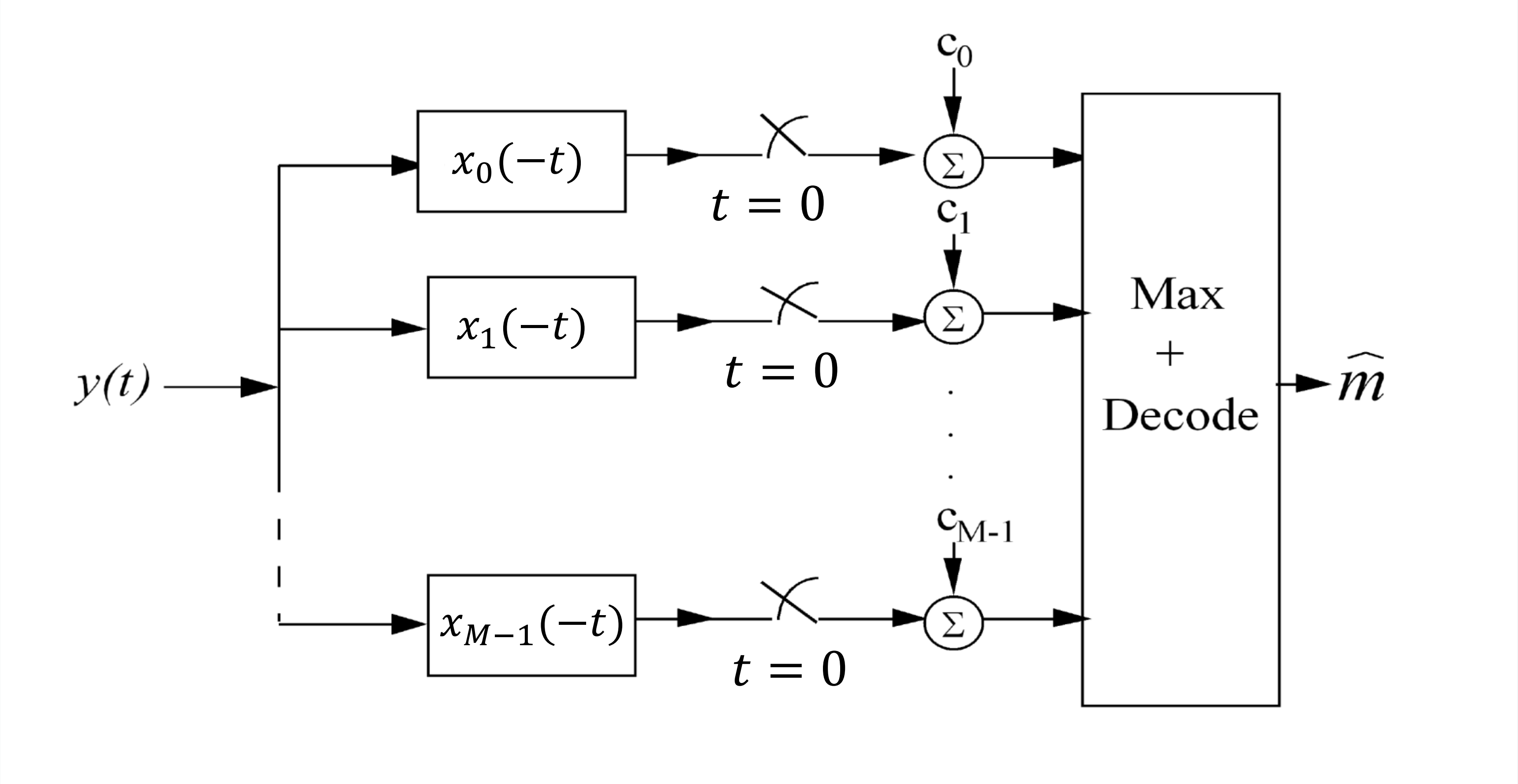
- As a result, the cascade of optimal demodulator and decoder at the receiver can be summarized as



or equivalently with the matched filter demodulator in lieu of the correlative demodulator.

How Does an Optimal Receiver Work?

- Using invariance property of the inner product, the optimal cascade of demodulator and decoder can also be implemented directly as



- Note that this architecture requires M, which is typically much larger than N, analog correlators or matched filters.

Are we Forgetting Anything?

- The demodulator captures the signal in full in the sense that

$$x(t) = \sum_{n=1}^{N} x_n \varphi_n(t)$$

but this is not the case for the noise component since

$$z(t) \neq \sum_{n=1}^{N} z_n \varphi_n(t)$$

- Can any other noise component obtained from

$$\tilde{z}(t) = z(t) - \sum_{n=1}^{N} z_n \varphi_n(t)$$

be useful for decoding the signal?

Are we Forgetting Anything?

- The short answer is: No, because the other noise components are independent of the signal and of the noise components $\mathbf{z}$ that affect the signal.

- The longer answer follows.

Which Parts of the Received Signal Can Be Neglected?

- Partition the received signal vector components into two parts as

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix}$$

- If

$$p(\mathbf{y}_2 | \mathbf{y}_1, \mathbf{x}) = p(\mathbf{y}_2 | \mathbf{y}_1)$$

then $\mathbf{y}_2$ is said to be irrelevant for the estimate of $\mathbf{x}$ when $\mathbf{y}_1$ is given, while $\mathbf{y}_1$ is said to be a sufficient statistic for $\mathbf{x}$.

- Irrelevant components can be neglected with no loss of optimality, or, equivalently, the detector can focus solely on sufficient statistics.

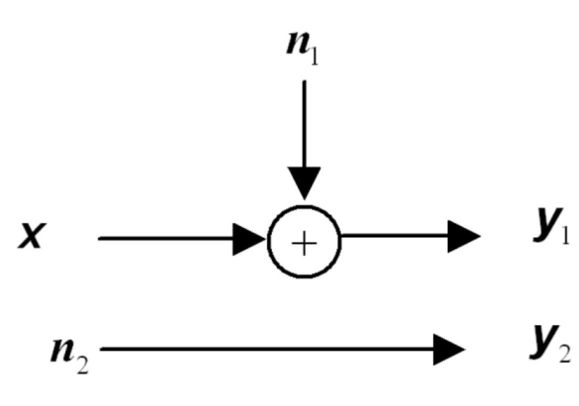
- Proof:

$$\begin{aligned} p(\mathbf{y} | \mathbf{x}_m) &= p(\mathbf{y}_1 | \mathbf{x}_m) p(\mathbf{y}_2 | \mathbf{y}_1, \mathbf{x}_m) \\ &= p(\mathbf{y}_1 | \mathbf{x}_m) p(\mathbf{y}_2 | \mathbf{y}_1) \end{aligned}$$

chain rule of prob.

Which Parts of the Received Signal Can Be Neglected?

- As an application of the previous result, consider the discrete channel below:



$\mathbf{n}_1$ and $\mathbf{n}_2$ are mutually independent and independent of $\mathbf{x}$

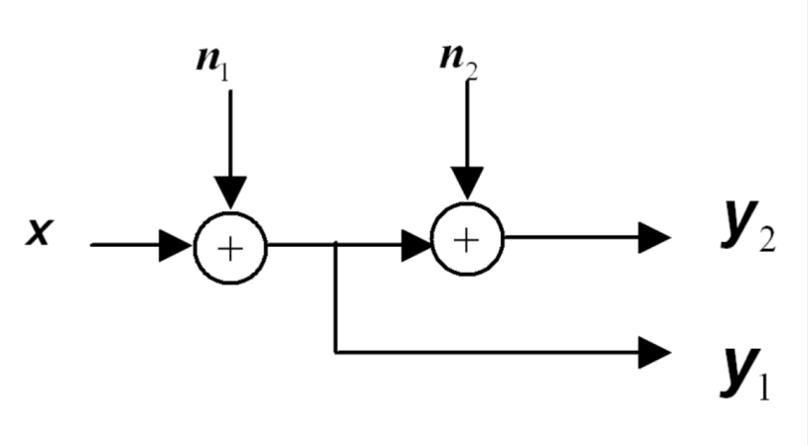
- We have

$$\begin{aligned} p(\mathbf{y}_2|\mathbf{y}_1, \mathbf{x}) &= p(\mathbf{n}_2|\mathbf{x}+\mathbf{n}_1, \mathbf{x}) \\ &= p(\mathbf{n}_2) && \text{independence of } \mathbf{n}_2 \text{ and } (\mathbf{x},\mathbf{n}_1) \\ &= p(\mathbf{y}_2) \\ &= p(\mathbf{y}_2|\mathbf{y}_1) && \text{independence of } \mathbf{y}_2 \text{ and } (\mathbf{y}_1) \end{aligned}$$

and hence $\mathbf{y}_2$ is irrelevant and can be neglected.

Which Parts of the Received Signal Can Be Neglected?

- As another application, show that $\mathbf{y}_2$ is irrelevant also for the discrete channel below.



$\mathbf{n}_1$ and $\mathbf{n}_2$ are mutually independent and independent of $\mathbf{x}$

Are we Forgetting Anything?

- No, because $\tilde{z}(t)$ is independent of $\sum_{n=1}^{N} z_n \varphi_n(t)$ and of $x(t)$, and hence any component extracted from it is irrelevant.

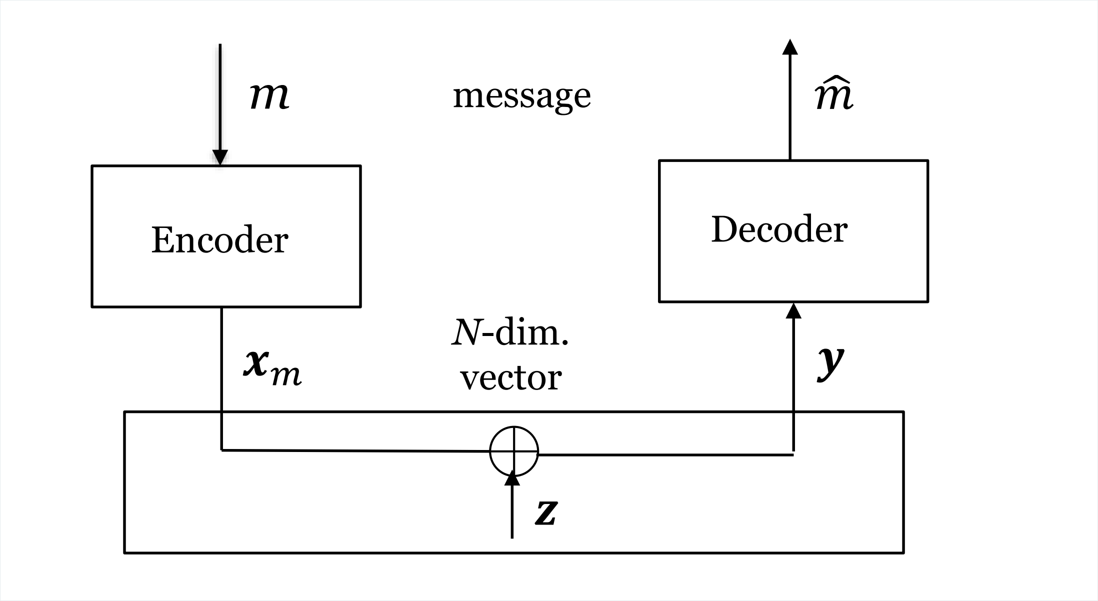
- Proof: We need to show that $\tilde{z}(t)$ is independent of $\sum_{n=1}^{N} z_n \varphi_n(t)$:

$$
\begin{aligned}
\mathrm{E}\left[\tilde{z}(t')z_n\right] &= \mathrm{E}\left[\left(z(t') - \sum_{n'=1}^{N} z_{n'}\varphi_{n'}(t)\right) z_n\right] \\
&= \mathrm{E}\left[z(t')z_n\right] - \mathrm{E}\left[\left(\sum_{n'=1}^{N} z_{n'}\varphi_{n'}(t)\right) z_n\right] \\
&= \int \mathrm{E}[z(t)z(t')]\varphi_n(t)dt - \sum_{n'=1}^{N} \mathrm{E}\left[z_{n'}z_n\right]\varphi_{n'}(t) \\
&= \frac{N_0}{2}\int \delta(t-t')\varphi_n(t)dt - \frac{N_0}{2}\sum_{n'=1}^{N} \delta_{n,n'}\varphi_{n'}(t) \\
&= \frac{N_0}{2}\left(\varphi_n(t') - \varphi_n(t')\right) = 0
\end{aligned}
$$

How Do We Design a Coding Scheme?

- In order to design the constellation, we need to evaluate the performance of a given constellation in the presence of an optimal receiver.

- As seen, we can concentrate on the discrete-time additive Gaussian model



$\mathbf{y} = \mathbf{x} + \mathbf{z}$ with

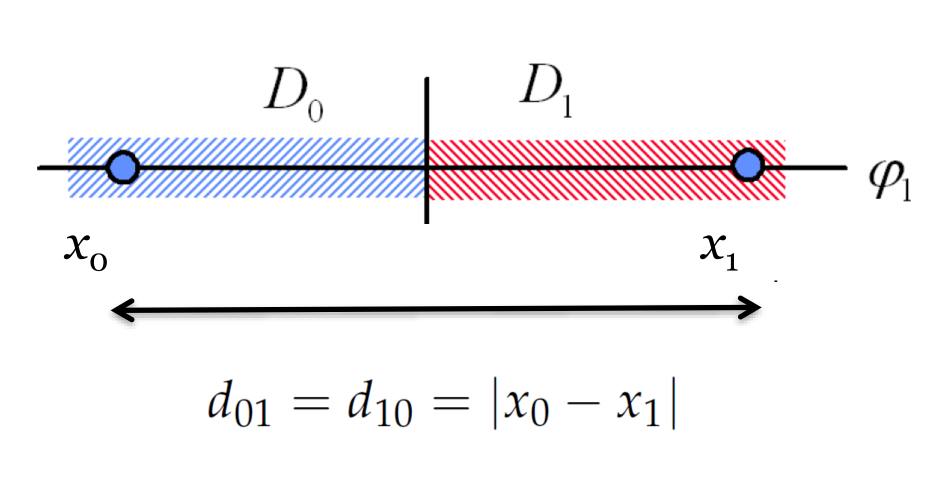
$z_n \sim \mathcal{N}\left(0, \frac{N_0}{2}\right)$ i.i.d.

How Do We Design a Coding Scheme?

- We will focus on the case of equiprobable messages for which MAP (optimal decoder) equals ML.

How Do We Evaluate the Performance of a Coding Scheme?

- Consider first N=1 and M=2 (binary communications)

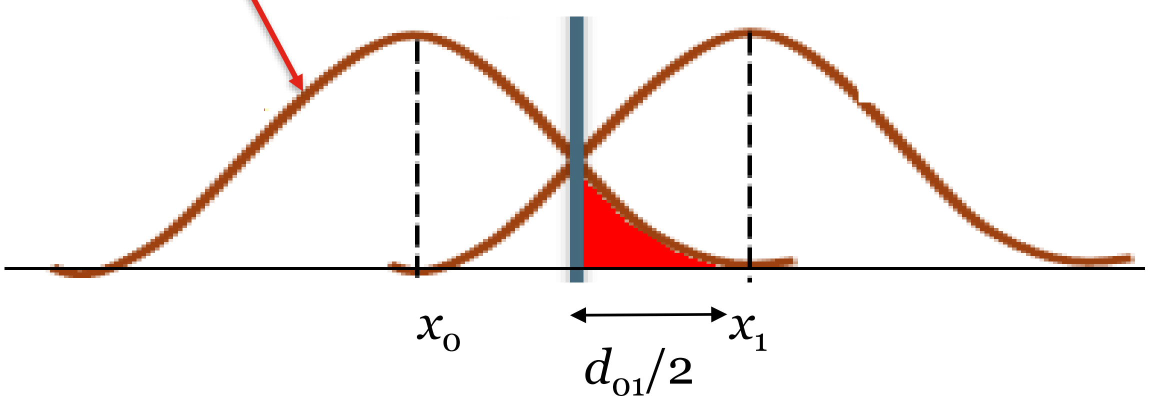


- d_{ij} = distance between constellation points $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$

How Do We Evaluate the Performance of a Coding Scheme?

- Consider first N=1 and M=2 (binary communications)

$$p(y|m=0) = N\left(x_0, \frac{N_0}{2}\right)$$

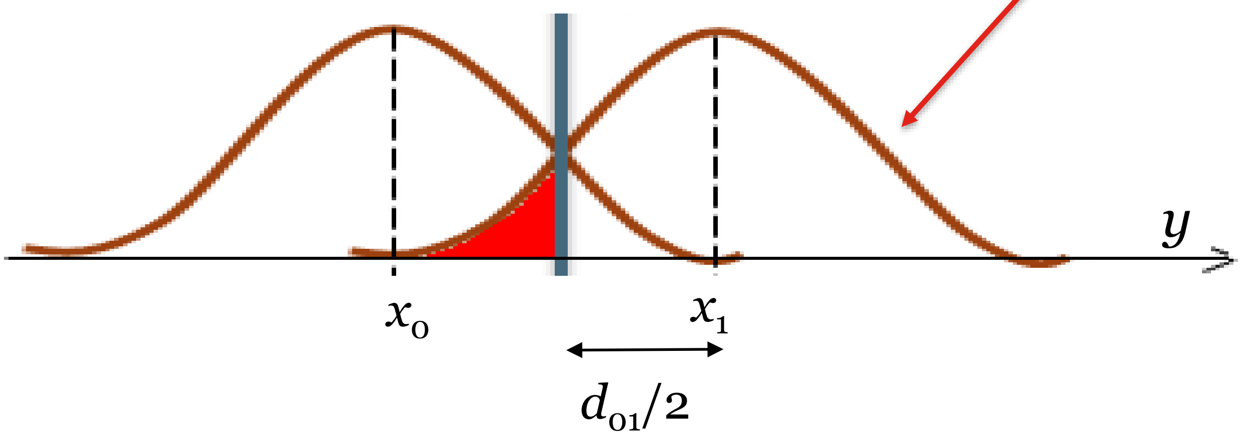


$$P^B_{e|0\to 1} = \Pr[\widehat{m} = 1 | m = 0]$$
$$= Q\left(\frac{d_{01}/2}{\sqrt{N_0/2}}\right) = Q\left(\frac{d_{01}}{\sqrt{2N_0}}\right)$$

How Do We Evaluate the Performance of a Coding Scheme?

- Consider first N=1 and M=2 (binary communications)

$$p(y|m=1) = N\left(x_1, \frac{N_0}{2}\right)$$



$$P^B_{e|1\to 0} = \Pr[\hat{m}=0|m=1]$$
$$= Q\left(\frac{d_{01}}{\sqrt{2N_0}}\right)$$

How Do We Evaluate the Performance of a Coding Scheme?

- The probability of error with N=1 and M=2 is hence given by

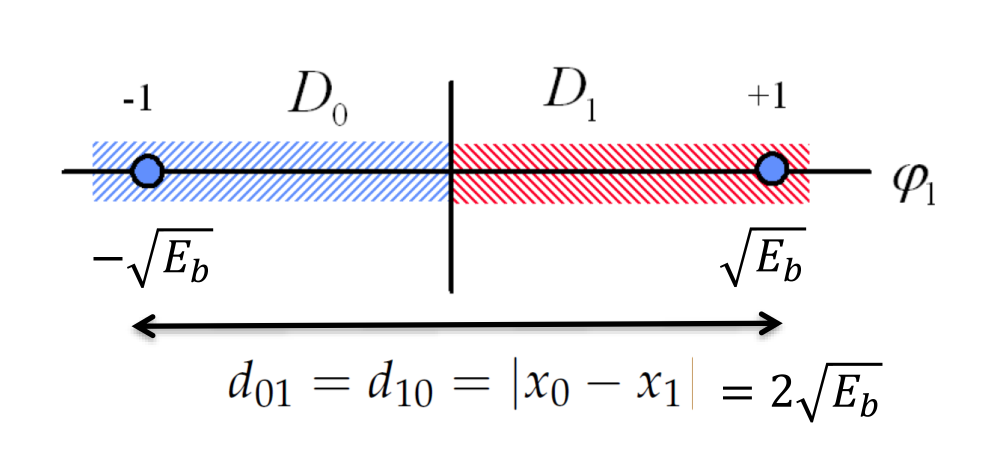
$$P_e = \frac{1}{2} P^B_{e|0\to 1} + \frac{1}{2} P^B_{e|1\to 0} = Q\left(\frac{d_{01}}{\sqrt{2N_0}}\right)$$

How Do We Evaluate the Performance of a Coding Scheme?

- As we have seen, the distance depends on the transmission energy

- Recall: E_b = energy per bit

- Ex.: BPSK



- The probability of error is hence

$$P_e = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$$

How Do We Evaluate the Performance of a Coding Scheme?

- Ex.: On-Off Keying (OOK)

$$P_e = Q\left(\sqrt{\frac{E_b}{N_0}}\right)$$

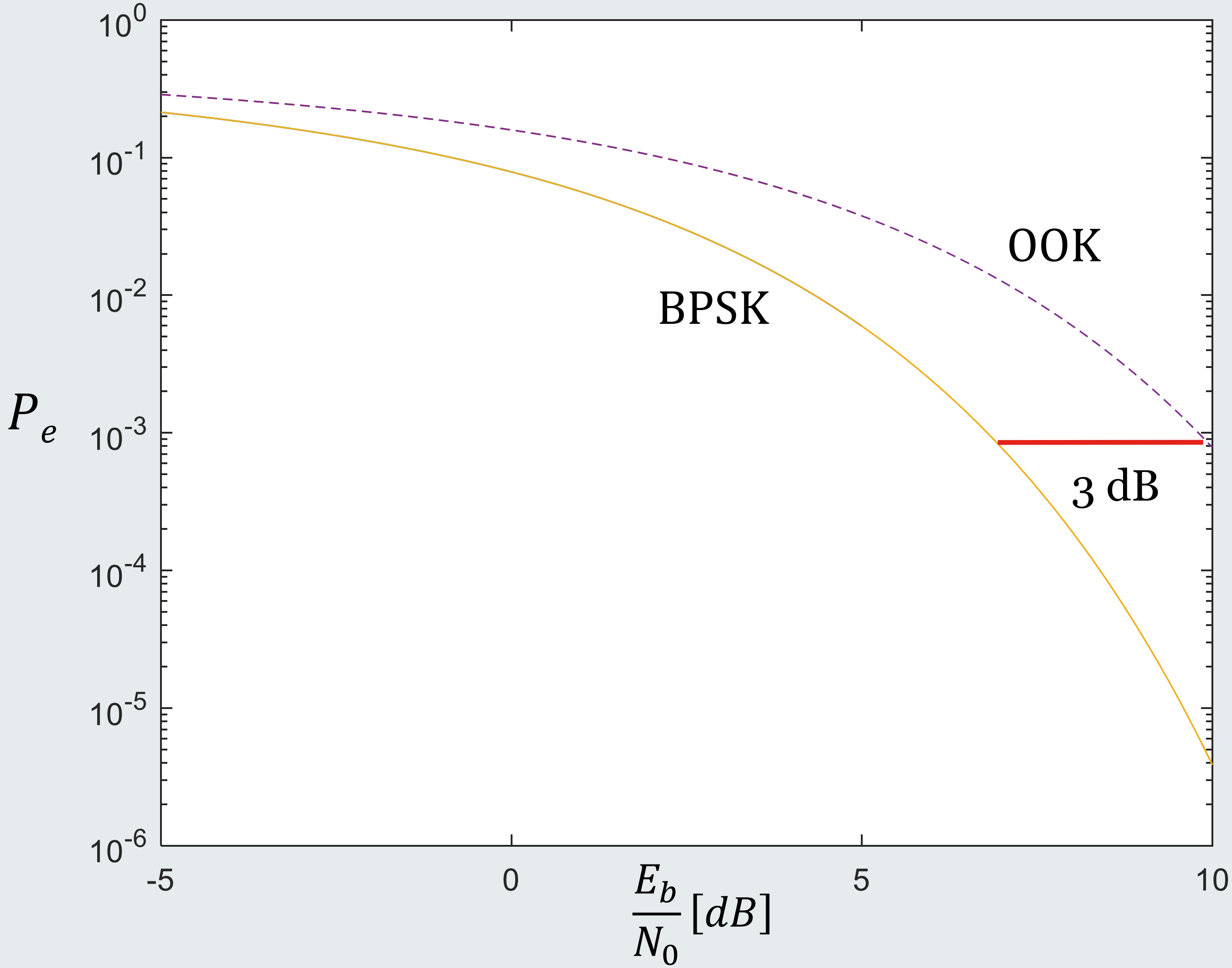
- Since we need double E_b to obtain the same probability of error, ON-OFF keying has a loss of 3dB ($= 10\log_{10}2$) with respect to the optimal modulation

Matlab: Plotting the Probability of Bit Error

BPSK and OOK

```
EbN0dB=linspace(-5,10,100); %x axis in dB
EbN0=10.^(EbN0dB./10); %x axis in linear scale

for s=1:length(EbN0)

E=EbN0(s);
Pebpsk(s)=qfunc(sqrt(2*E));
Peook(s)=qfunc(sqrt(E));

end

semilogy(EbN0dB,Pebpsk);
hold on
semilogy(EbN0dB,Peook,'--');
```

How Do We Evaluate the Performance of a Coding Scheme?



How Do We Evaluate the Performance of a Coding Scheme?

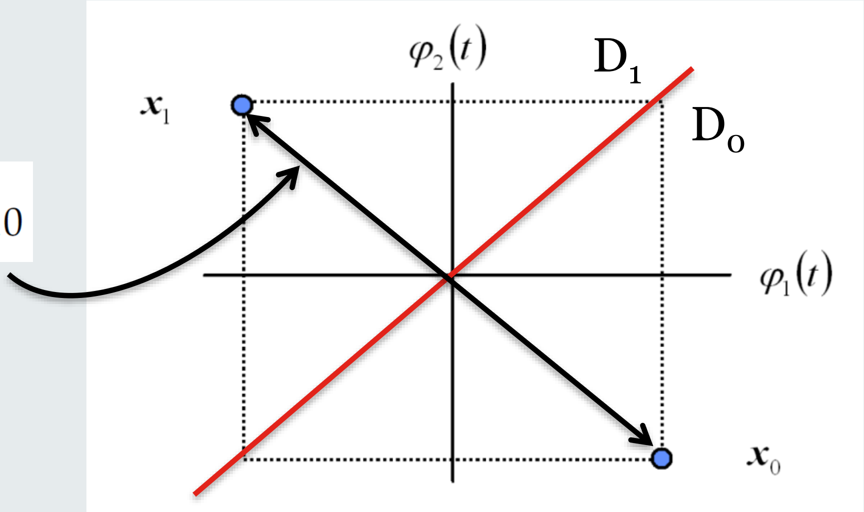
- Generalizing to any N, if M=2 (b=1), the probability of error is given by

$$P_e = Q\left(\frac{d_{01}}{\sqrt{2N_0}}\right) \quad \text{where} \quad d_{01} = d_{10} = ||\mathbf{x}_0 - \mathbf{x}_1||$$

for any N.

- Illustration:

$d_{01} = d_{10}$

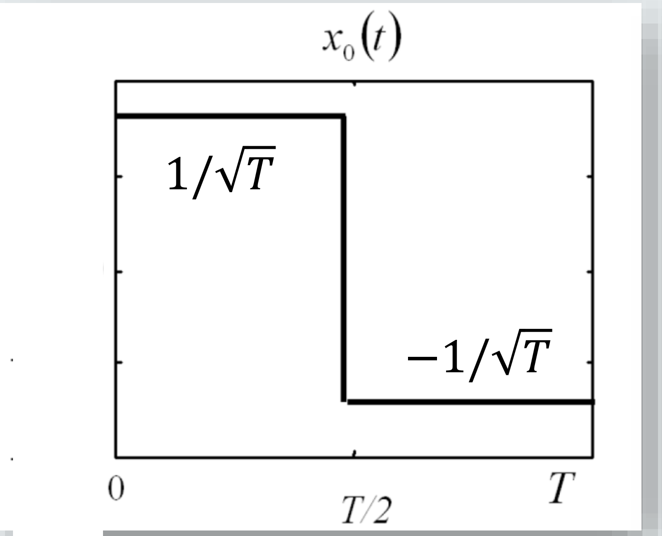


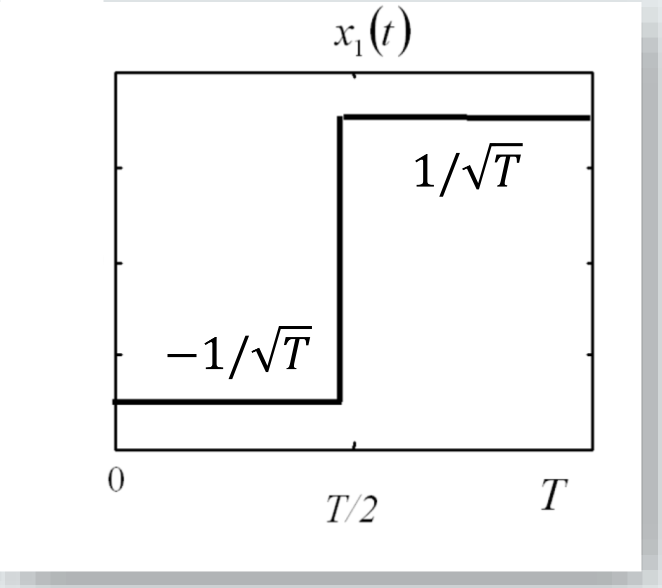
How Do We Evaluate the Performance of a Coding Scheme?

- Note that, by the invariance of the inner product, we can compute the distance directly between the analog waveforms:

$$d_{01} = d_{10} = ||\mathbf{x}_0 - \mathbf{x}_1|| = \sqrt{\int (x_0(t) - x_1(t))^2 dt}$$

- Ex.: BPSK





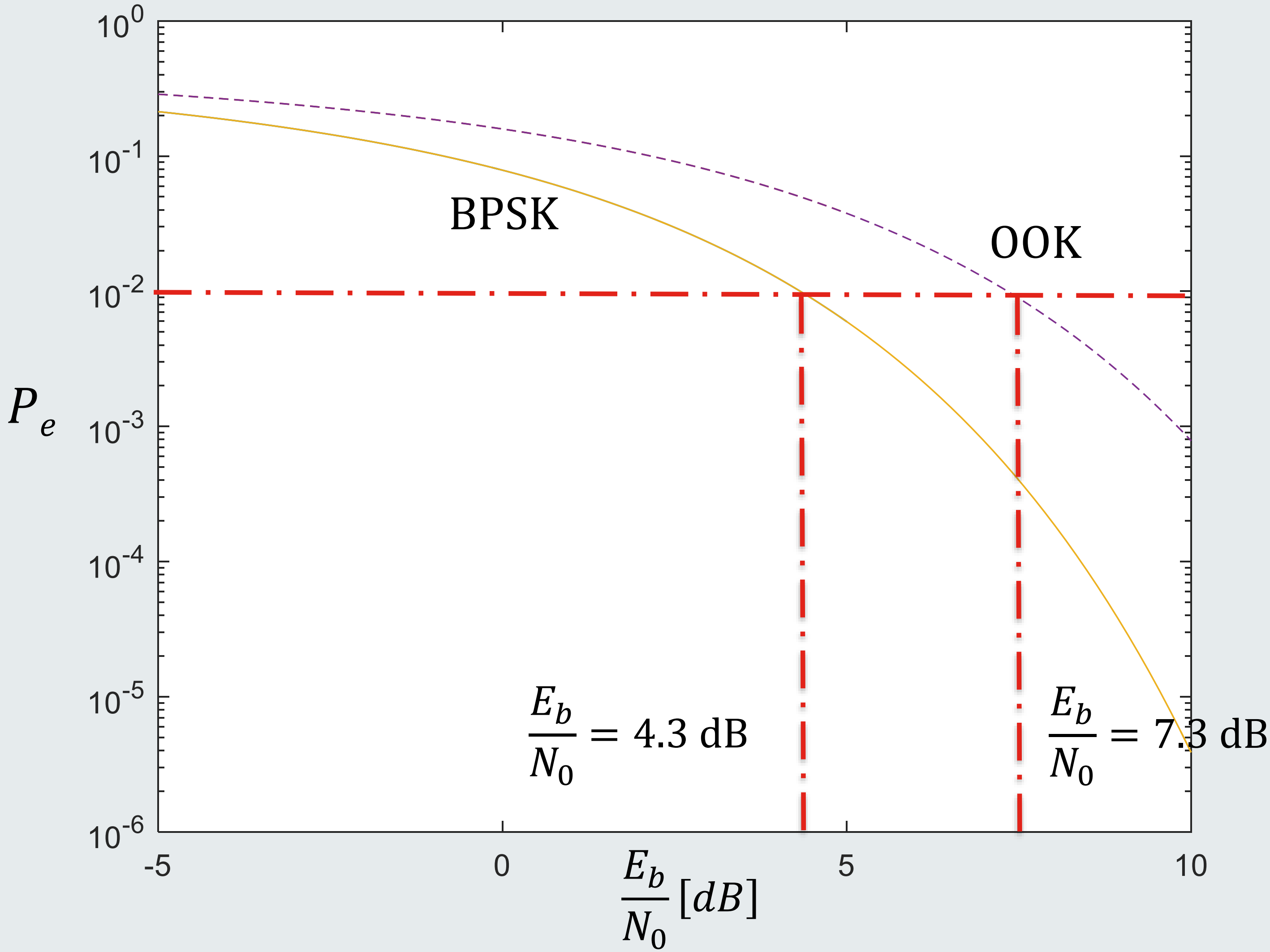
$d_{01}=d_{10}=2$

How Far Can You Go From a Wi-Fi Access Point?

- Ex.: Consider a Wi-Fi access point with a transmission power $P_x = 0$ dBm operating over a channel with attenuation at 1 m equal to $L_1 = -50$ dB and path loss $\gamma = 2$. Assume that the access point uses BPSK or OOK with modulator $\varphi_1(t)=\sqrt{(2/T)}\,\mathrm{sinc}(t/T)\cos(2\pi f c\ t)$ (N=1) with bandwidth 1 MHz and that the power spectral density of the noise is $N_0 = -170$ dBm/Hz. How far can you be if you wish to receive at a probability of error no larger than 10^{-2}?

How Far Can You Go From a Wi-Fi Access Point?

- From the plot, using tables, or qfuncinv in MATLAB, we compute the required Eb/No:



How Far Can You Go From a Wi-Fi Access Point?

- It follows that the required received energy per bit is

$$E_\mathrm{b} = 4.3 - 170 = -165.7 \text{ dBm} \;\text{ for BPSK}$$

$$E_\mathrm{b} = 7.3 - 170 = -162.7 \text{ dBm for OOK}$$

- The required received power is

$$P_\mathrm{r} = \frac{E_\mathrm{b}}{T/b} = \frac{E_\mathrm{b}}{T} = -165.7 - (-60) = -105.7 \text{ dBm} \;\text{ for BPSK}$$

$$P_\mathrm{r} = \frac{E_\mathrm{b}}{T/b} = \frac{E_\mathrm{b}}{T} = -162.7 - (-60) = -102.7 \text{ dBm} \;\text{ for OOK}$$

- The maximum distance is obtained as

$$\begin{aligned} P_\mathrm{r} &= P_x + L_1 \;\text{(dB)} - \gamma\, 10\log_{10}(d) = P_x - 50 - 20\log_{10}(d) \\ &= 0 - 50 - 20\log_{10}(d) = -50 - 20\log_{10}(d) \end{aligned}$$

and so d=609.5 m for BPSK and so d=431.5 m are the maximum distances.

How Do We Evaluate the Performance of a Coding Scheme?

- Let's consider now any M

- Probability of error:

$$P_e = \frac{1}{M} \sum_{m=0}^{M-1} P_{e|m}$$
$$= \frac{1}{M} \sum_{m=0}^{M-1} \sum_{m' \neq m} P_{e|m \to m'}$$

where

$$P_{e|m} = \Pr[\hat{m} \neq m | m]$$

$$P_{e|m \to m'} = \Pr[\hat{m} = m' | m]$$

How Do We Evaluate the Performance of a Coding Scheme?

- Computing

$$P_{e|m \to m'} = \Pr[\hat{m} = m' | m]$$

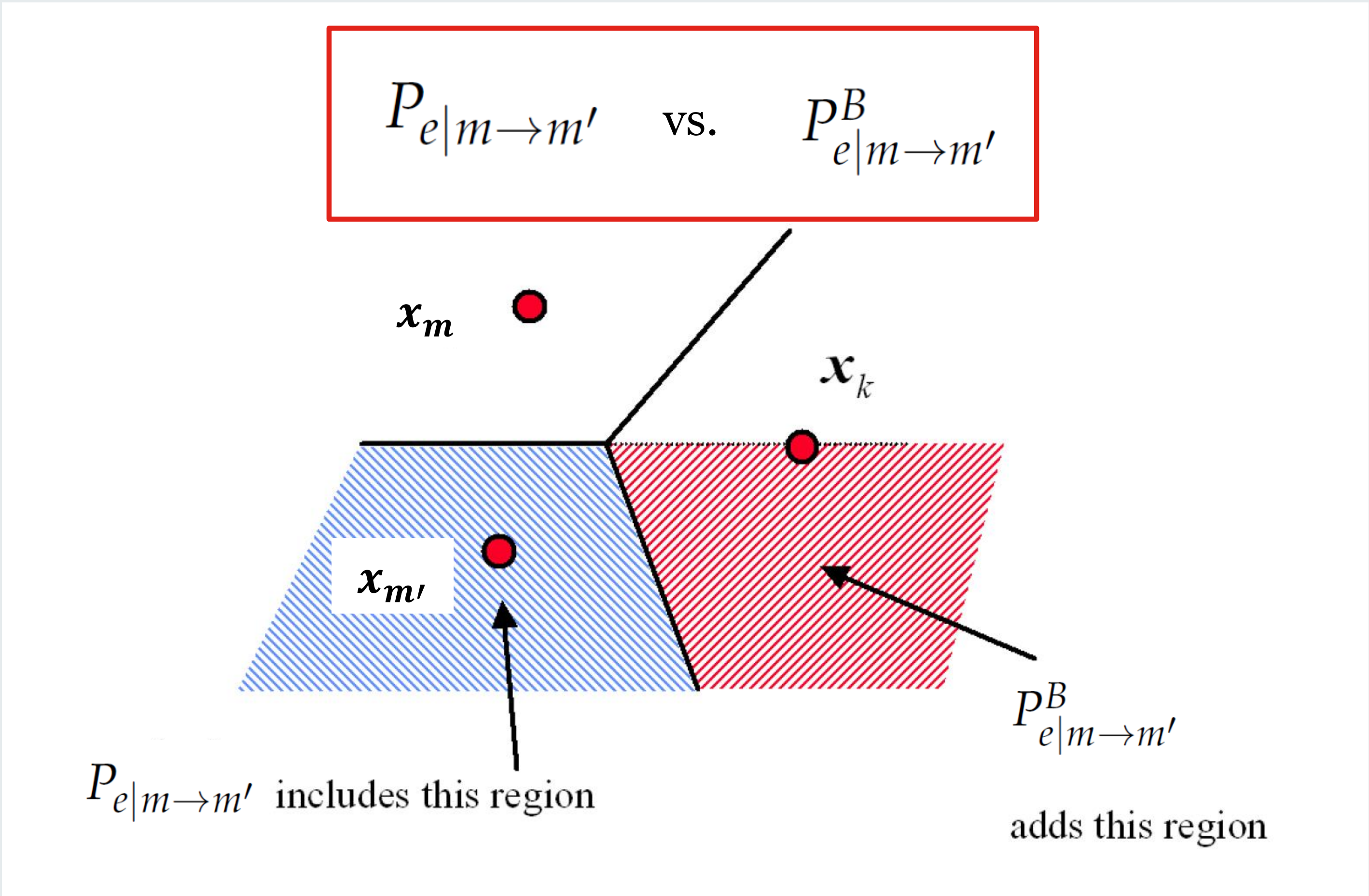
is generally difficult.

- We hence consider instead an upper bound that is easy to compute

$$P_{e|m \to m'} \leq P^{B}_{e|m \to m'}$$

- The upper bound is obtained by assuming that only messages *m* and *m'* exist, and hence the system is binary.
- Having an upper bound is useful, since the real probability of error is smaller than the bound.

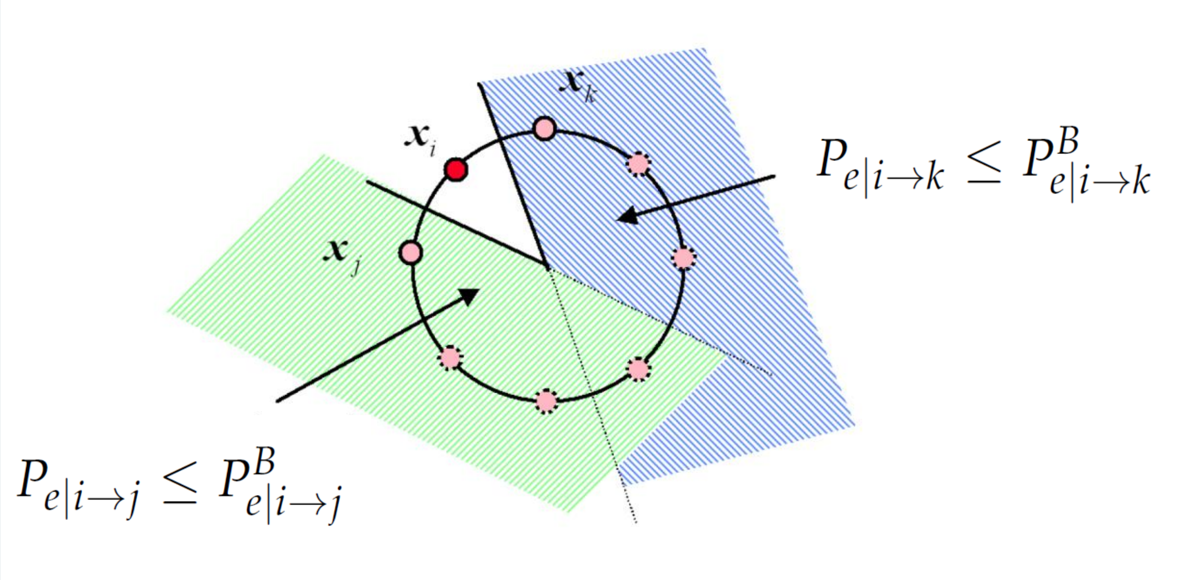
How Do We Evaluate the Performance of a Coding Scheme?



$$P_{e|m\to m'} \le P^B_{e|m\to m'} = Q\left(\frac{d_{mm'}}{\sqrt{2N_0}}\right)$$

How Do We Evaluate the Performance of a Coding Scheme?

- Ex.: 8-PSK



How Do We Evaluate the Performance of a Coding Scheme?

- **Union bound:**

$$P_e \leq \frac{1}{M} \sum_{m=0}^{M-1} \sum_{m' \neq m} Q\left(\frac{d_{mm'}}{\sqrt{2N_0}}\right)$$

- **Approximate union "bound":**

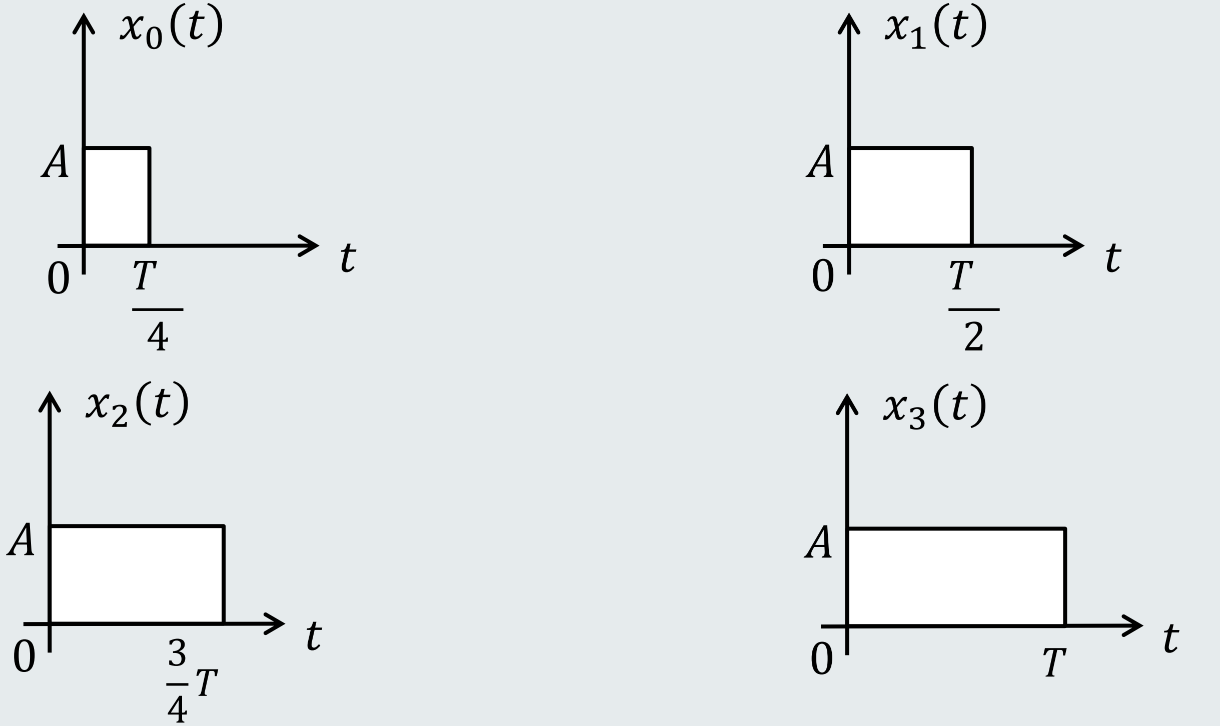
$$P_e \cong \frac{1}{M}(\text{num. pairs at min. distance})Q\left(\frac{d_{\min}}{\sqrt{2N_0}}\right)$$

$$d_{\min} = \min_{i \neq j} d_{ij}$$ minimum distance

- **Important:** By the invariance of the correlation, distances can be computed either in the signal space or on the waveforms of the signal set.

How Do We Evaluate the Performance of a Coding Scheme?

- Ex.: 4-Pulse Width Modulation (PWM)



$$A = \sqrt{\frac{16}{5}\frac{E_b}{T}}$$

How Do We Compute the Union Bound?

Ex.: 4-PWM ($b = 2$)

1) Conditional squared distance spectrum:

for $m = 0$

$$d_{01}^2 = \int_{T/4}^{T/2} \frac{16}{5}\frac{E_b}{T_P}\,dt = \frac{4E_b}{5}$$

$$d_{02}^2 = \frac{8E_b}{5}$$

$$d_{03}^2 = \frac{12E_b}{5}$$

$$\Rightarrow \left[\left\{\frac{4E_b}{5}, 1\right\},\ \left\{\frac{8E_b}{5}, 1\right\}, \left\{\frac{12E_b}{5}, 1\right\}\right]$$

How Do We Compute the Union Bound?

for m= 1 and m = 2

$$\left[\left\{\frac{4E_b}{5}, 2\right\},\ \left\{\frac{8E_b}{5}, 1\right\}\right]$$

for m = 3

$$\left[\left\{\frac{4E_b}{5}, 1\right\},\ \left\{\frac{8E_b}{5}, 1\right\}, \left\{\frac{12E_b}{5}, 1\right\}\right]$$

2) Squared distance spectrum

$$\left[\left\{\frac{4E_b}{5}, 6\right\},\ \left\{\frac{8E_b}{5}, 4\right\}, \left\{\frac{12E_b}{5}, 2\right\}\right]$$

How Do We Compute the Union Bound?

3) Union bound:

$$P_e \le \frac{1}{4}\left(6Q\left(\sqrt{\frac{4E_b}{10N_0}}\right) + 4\mathrm{Q}\left(\sqrt{\frac{8E_b}{10N_0}}\right) + 2\mathrm{Q}\left(\sqrt{\frac{12E_b}{10N_0}}\right)\right)$$

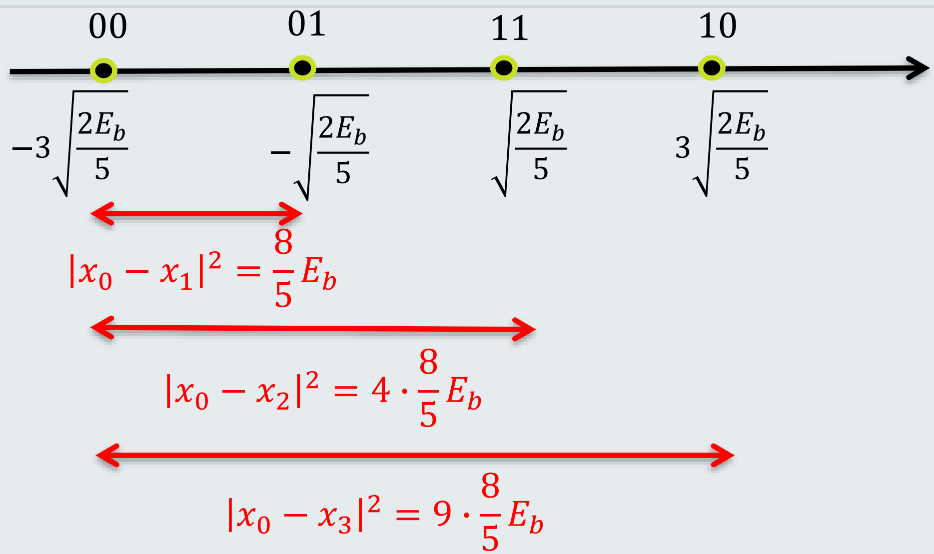
$$= \frac{3}{2}\,\mathrm{Q}\left(\sqrt{\frac{2E_b}{5N_0}}\right) + \mathrm{Q}\left(\sqrt{\frac{4E_b}{5N_0}}\right) + \frac{1}{2}\mathrm{Q}\left(\sqrt{\frac{6E_b}{5N_0}}\right)$$

4) Approximate union bound:

$$P_e \simeq \frac{3}{2}\,\mathrm{Q}\left(\sqrt{\frac{2E_b}{5N_0}}\right)$$

How Do We Compute the Union Bound?

- Ex.: 4-PAM



The squared distance spectrum is then given as

$$\left[\left\{\frac{8E_b}{5}, 6\right\}, \left\{\frac{32E_b}{5}, 4\right\}, \left\{\frac{72E_b}{5}, 2\right\}\right]$$

and the union bound is

$$P_e \leq \frac{1}{4}\left(6\,Q\left(\sqrt{\frac{4E_b}{5N_0}}\right) + 4Q\left(\sqrt{\frac{16E_b}{5N_0}}\right) + 2Q\left(\sqrt{\frac{36E_b}{5N_0}}\right)\right)$$

$$\simeq \frac{3}{2}\,Q\left(\sqrt{\frac{4E_b}{5N_0}}\right)$$

Union bound approximation

Matlab: Plotting the Union Bound

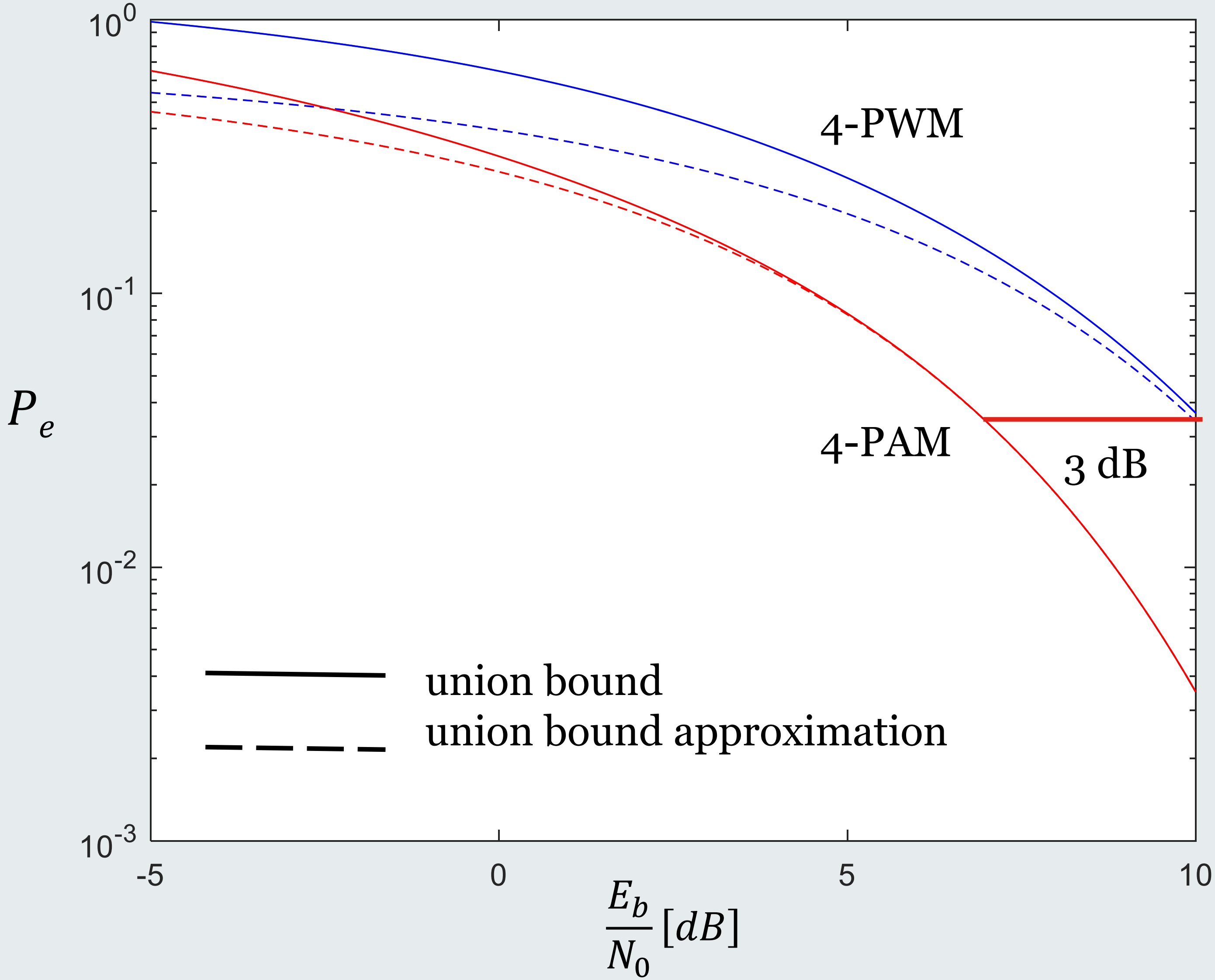
Union bound for 4-PWM and 4-PAM

```
EbN0dB=linspace(-5,10,100); %x axis in dB
EbN0=10.^(EbN0dB./10); %x axis in linear scale

for s=1:length(EbN0)

E=EbN0(s);
Pubpwm(s)=3/2*qfunc(sqrt(2/5*E))+qfunc(sqrt(4/5*E))+1/2*qfunc(sqrt
(6/5*E));
Papppwm(s)=3/2*qfunc(sqrt(2/5*E));

Pubpam(s)=3/2*qfunc(sqrt(4/5*E))+qfunc(sqrt(16/5*E))+1/2*qfunc(sqr
t(36/5*E));
Papppam(s)= 3/2*qfunc(sqrt(4/5*E));
end
```

Matlab: Plotting the Union Bound

```
semilogy(EbNodB,Pubpwm,'b');
hold on
semilogy(EbNodB,Papppwm,'b--');
semilogy(EbNodB,Pubpam, 'r');
semilogy(EbNodB,Papppam, 'r--');
```

How Do We Compute the Union Bound?



How Do We Compute the Union Bound?

- The union bound approximation is increasingly accurate for large SNR values to the exponential decay of the Q function.

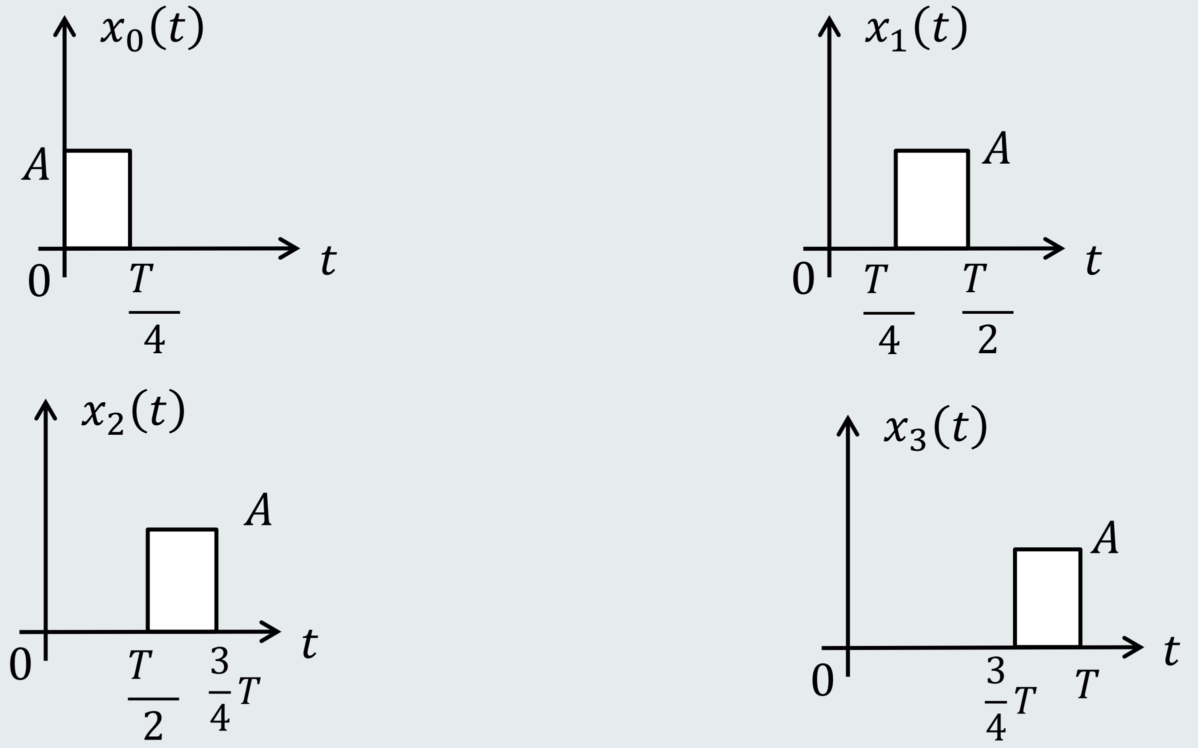
- The gain/ loss of a constellation as compared to another can be well approximated by considering only the arguments of the Q function in the union bound approximation.

- Ex.: Loss of 4-PWM as compared to 4-PAM

$$10\log_{10}\left(\frac{4/5}{2/5}\right) = 3\text{ dB}$$

How Do We Evaluate the Performance of a Coding Scheme?

- Ex.: Try with 4-Pulse Position Modulation (PPM)



$$A = \sqrt{\frac{8\,E_b}{T}}$$

What About the Bit Error Rate?

- The bit error rate (or probability is given as):

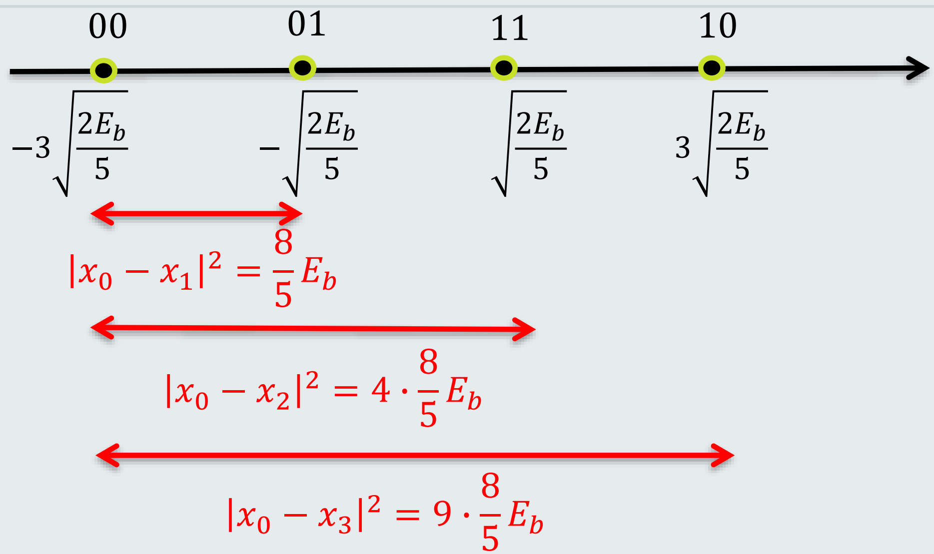
$$
\begin{aligned}
P_b &= \frac{1}{M}\sum_{m=0}^{M-1} P_{b|m} \\
&= \frac{1}{M}\sum_{m=0}^{M-1}\sum_{m'\neq m} \frac{(\text{number of different bits between m and m}')}{\log_2 M} P_{e|m\to m'}
\end{aligned}
$$

which can be bounded by using the union bound.

- Unlike the probability of error, the probability of bit error depends on the mapping between bits and symbols.

What About the Bit Error Rate?

- Ex.: 4-PAM



Union bound on the probability of bit error:

$$P_b \leq \frac{1}{4}\left(6\frac{1}{2}\,\mathrm{Q}\left(\sqrt{\frac{4E_b}{5N_0}}\right) + 4\frac{2}{2}Q\left(\sqrt{\frac{16E_b}{5N_0}}\right) + 2\frac{1}{2}\mathrm{Q}\left(\sqrt{\frac{36E_b}{5N_0}}\right)\right)$$

$$\simeq \frac{3}{4}\,\mathrm{Q}\left(\sqrt{\frac{4E_b}{5N_0}}\right)$$

Union bound approximation

KING'S
College
LONDON

KING'S College LONDON

6CCS3COS Communication Systems: Chapter 3

Osvaldo Simeone

What is This Course About?

- **Overview**

  - 1. One-shot digital communications: Fundamentals

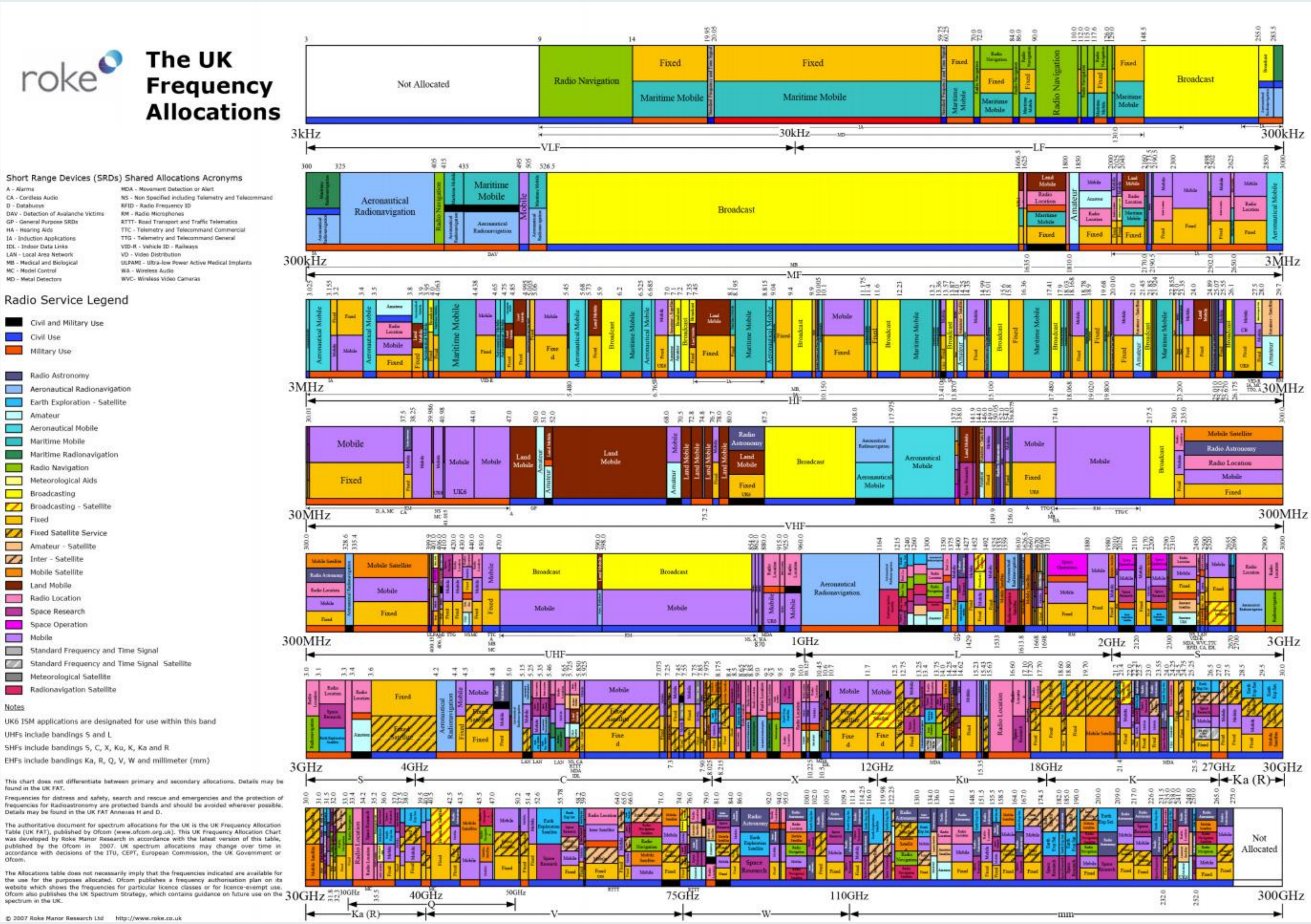
  - 2. One-shot digital communications: Passband Systems

  - 3. Stream digital communications

Main references

- J. Cioffi, Lecture notes, Stanford Univ., Chapters 1, 2, 3

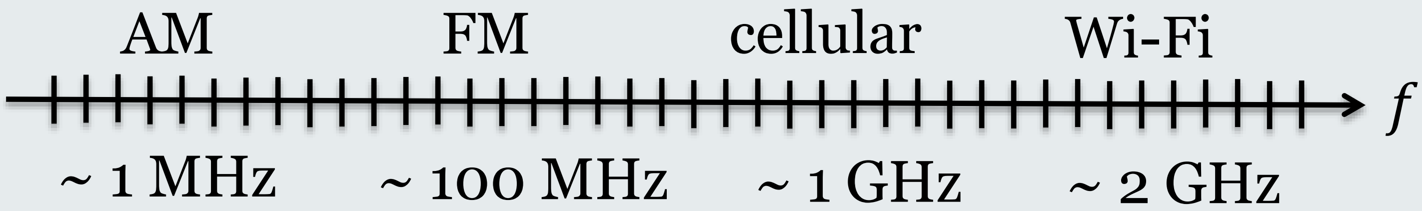
Why Passband Communications?



[http://www.extremetech.com]

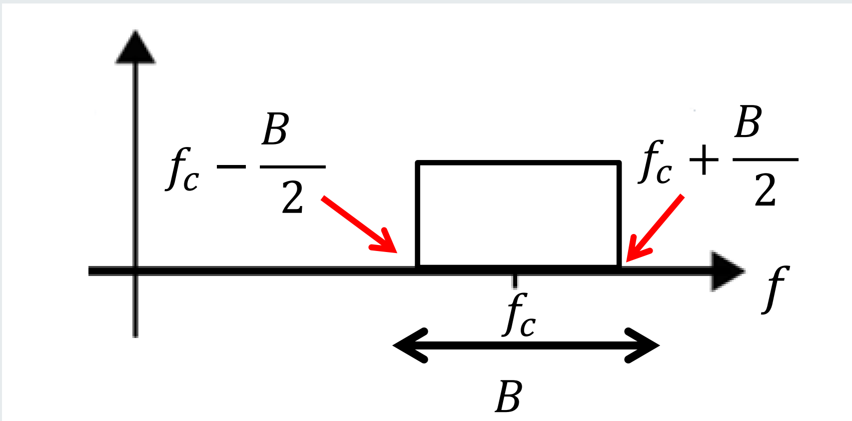
Why Passband Communications?

- 1) Frequency Division Multiplexing (FDM)



...different information streams (e.g., radio stations) modulated on different carriers

Each carrier corresponds to a different passband channel



Why Passband Communications?

System	Carrier Frequency f_c
AM radio	530-1600 kHz
FM radio	88-108 MHz
Cellular	~900 MHz, ~1-2 GHz
Wi-Fi	2.4 GHz
Satellite	~3-6 GHz
Fiber optics	200 THz

legend: $k \rightarrow 10^3, M \rightarrow 10^6, G \rightarrow 10^9, T \rightarrow 10^{12}$, $Hz = \text{cycles}/s$

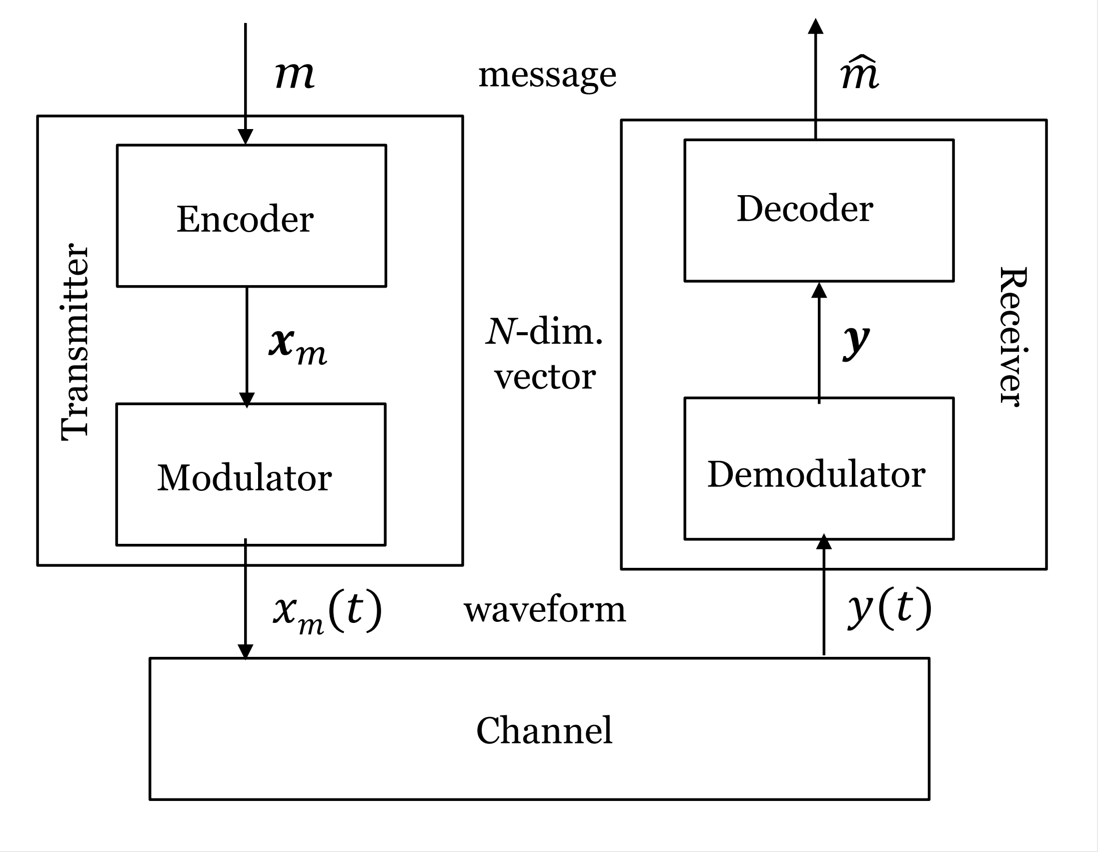
Why Passband Communications?

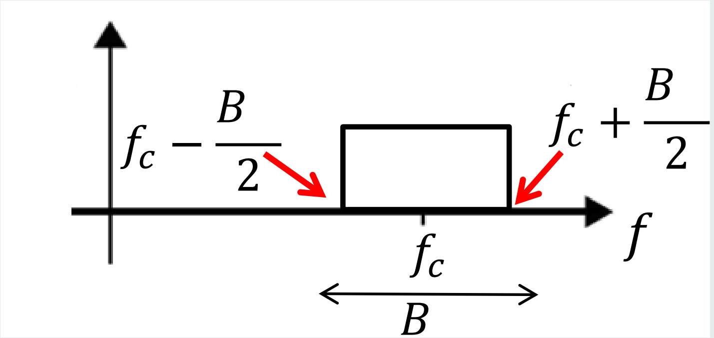
2) The antenna size depends on the wavelength

$$\lambda = \frac{c}{f_c} \quad (c = 3 \times 10^8 \text{ m/s})$$

System	**Wavelength λ**
AM	~ 300 m
FM	~ 3 m
Cellular	~ 0.3 m
Wi-Fi	~ 0.1 m

Why Passband Communications?





Why Passband Communications?

- Note: The passband filter in practice is implemented at the receiver, but it is convenient to think of it as being part of the channel.

How To Carry Out Link Budgets for Passband Communications?

- The path loss at 1 m is typically given by Friis formula

$$L_1\ (\mathrm{dB}) = \mathrm{G}\ (\mathrm{dB}) + \gamma\ 10\ \log_{10}(\frac{\lambda}{4\pi})$$

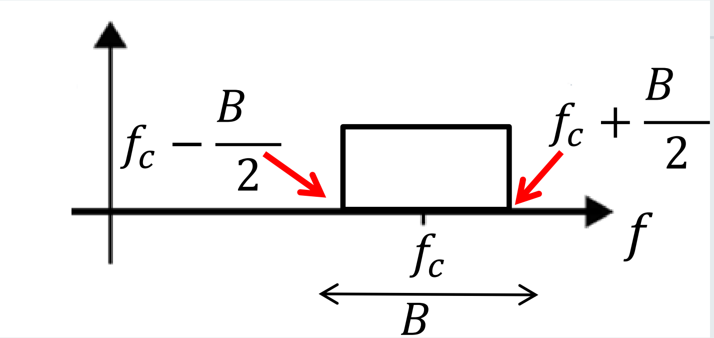
where G is the antenna gain and $\lambda = c/f_c$ ($c = 3 \times 10^8$ m/s).

- The overall path loss is given as

$$L\ (\mathrm{dB}) = L_1\ (\mathrm{dB}) - \gamma\ 10\ \log_{10}(d)$$

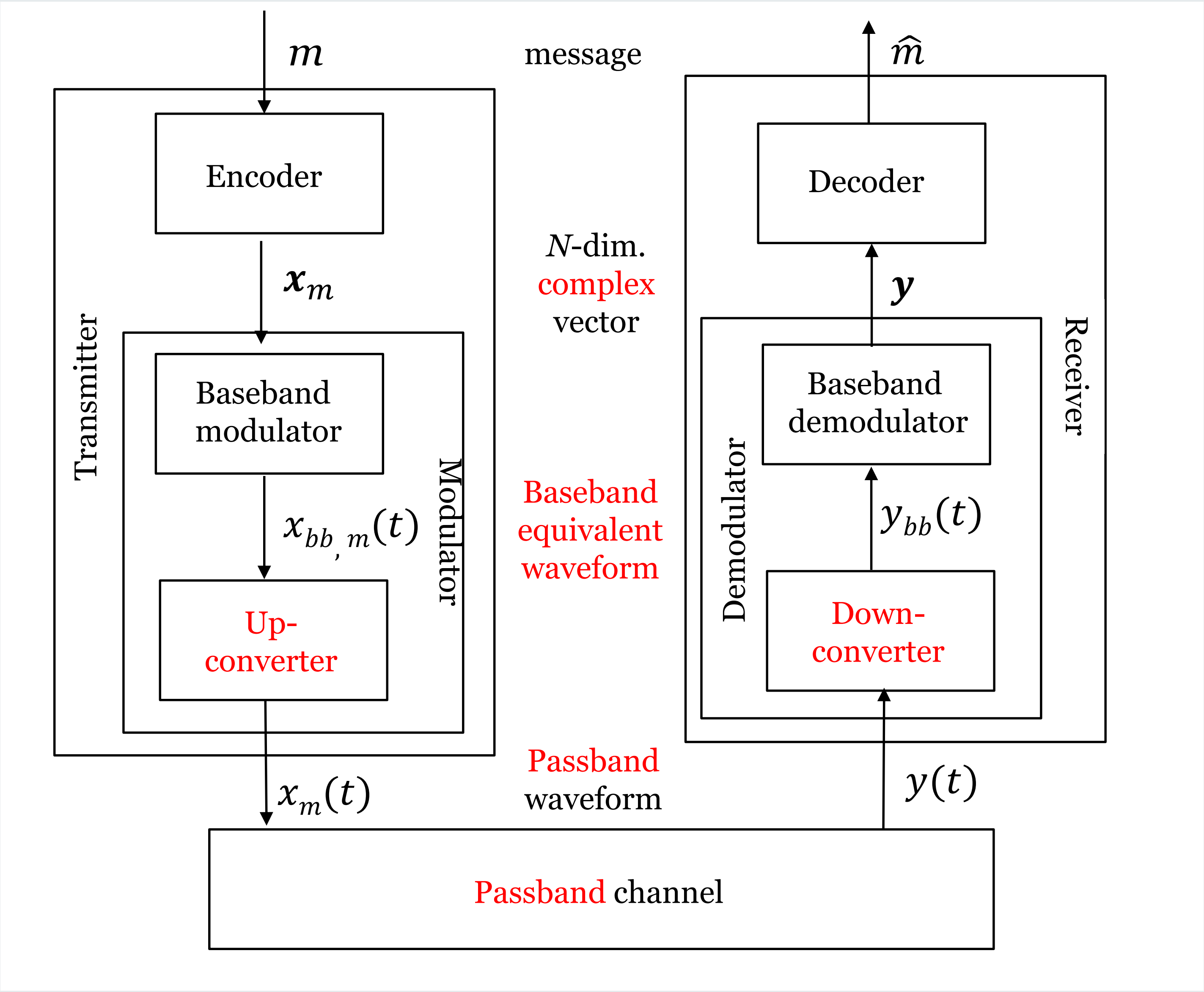
with d measured in meters

What Does This Mean for Coding and Modulation?

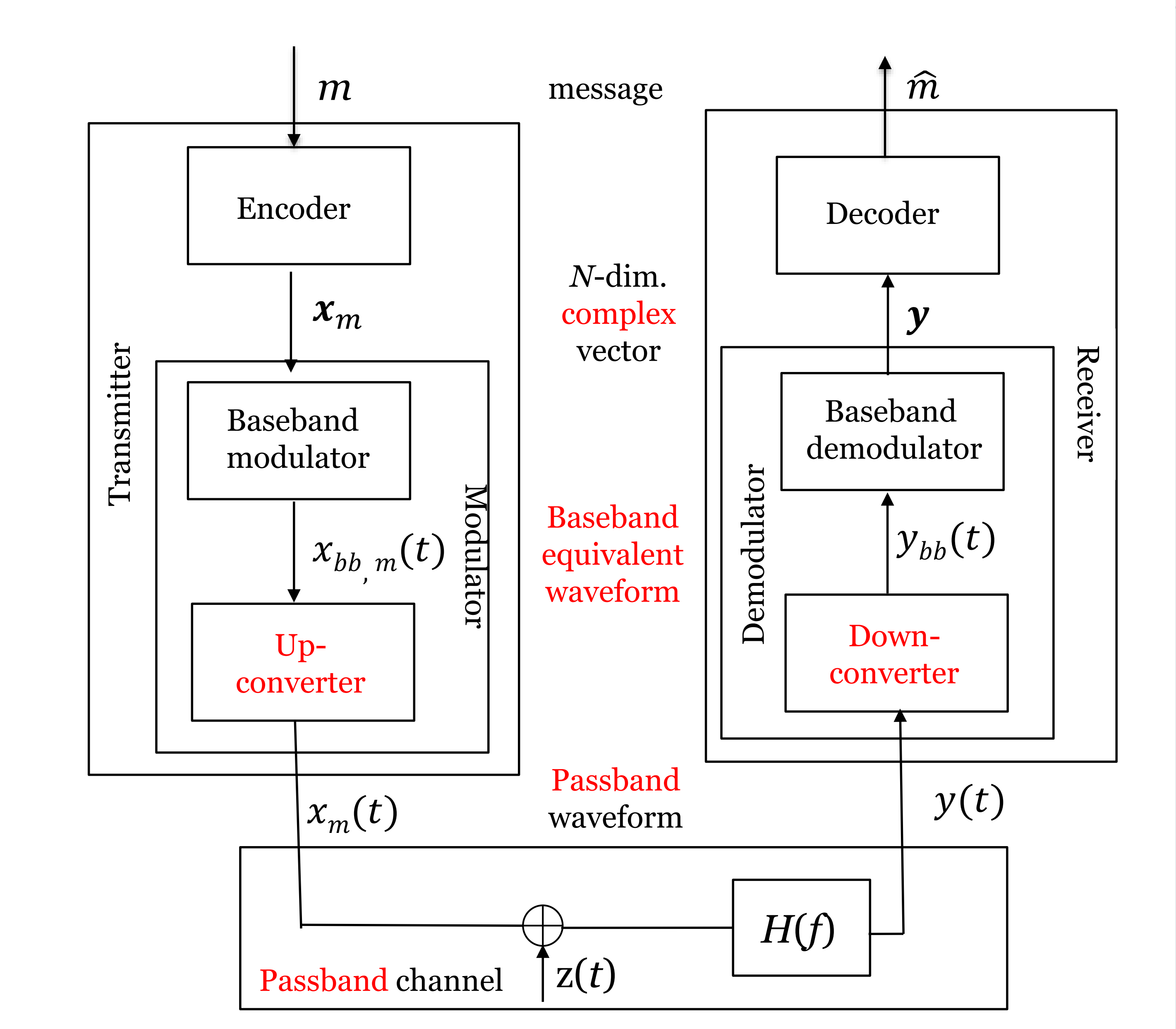


- Passband signals are difficult to process directly due to the large carrier frequency, especially in the digital domain.
- Operating the modulator and demodulator directly on passband signals requires tuning to the specific carrier frequency.
- Solution: Perform modulation and demodulation on baseband signals and carry out upconversion and downconversion separately.
- Upconversion/ downconversion are done via multiplication with sinusoidal signals at the carrier frequency.

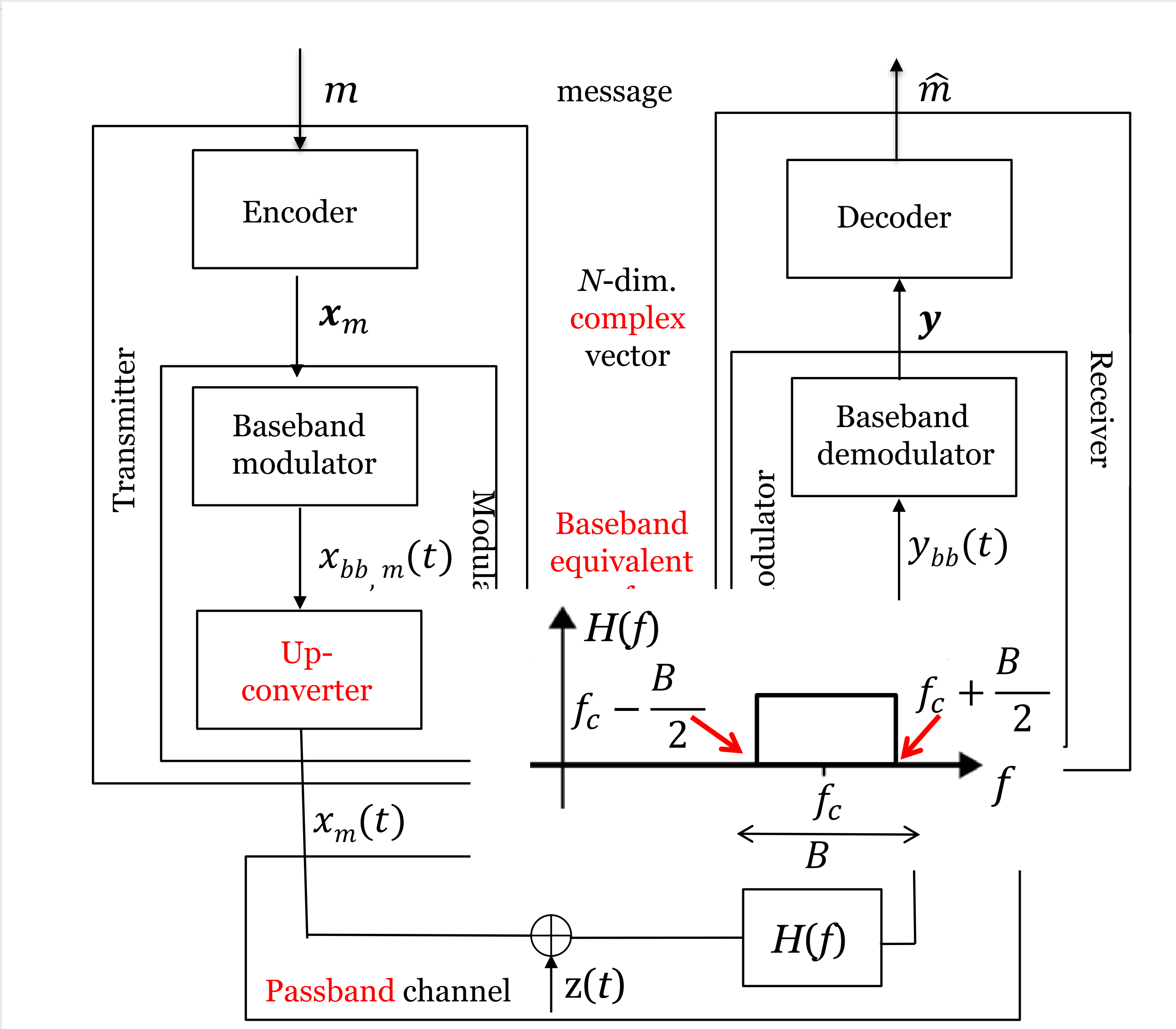
What Does This Mean for Coding and Modulation?



What Does This Mean for Coding and Modulation?

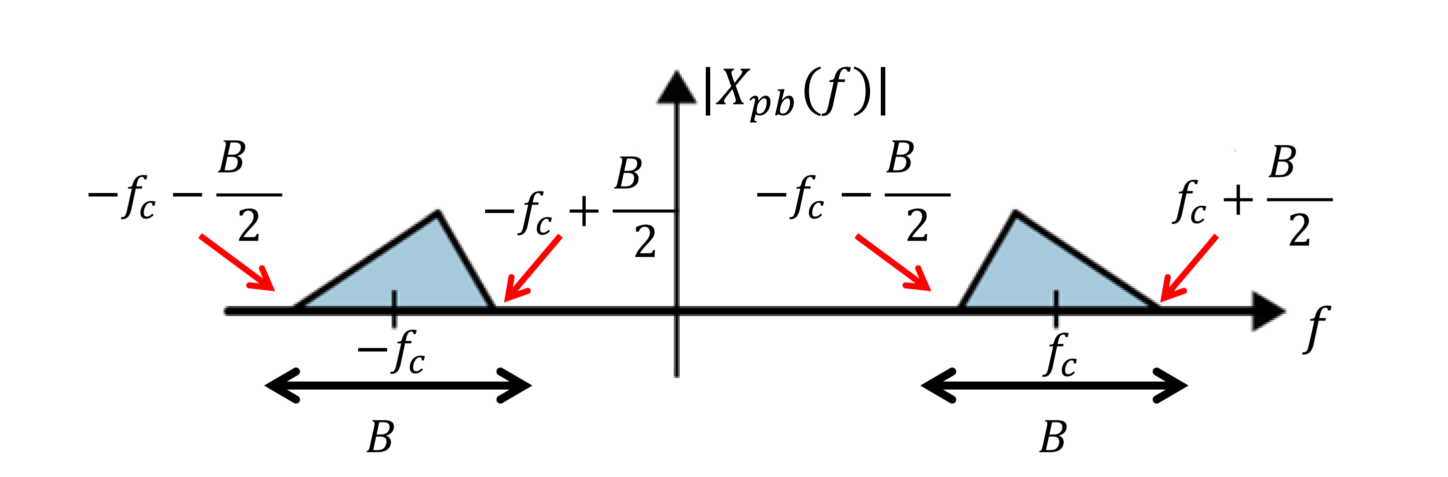


What Does This Mean for Coding and Modulation?



Why is the Baseband Equivalent Complex?

- Passband signal: Fourier transform is non-zero only in a bandwidth B around the carrier frequency $\pm f_c$ $\left(\frac{B}{2} < f_c\right)$

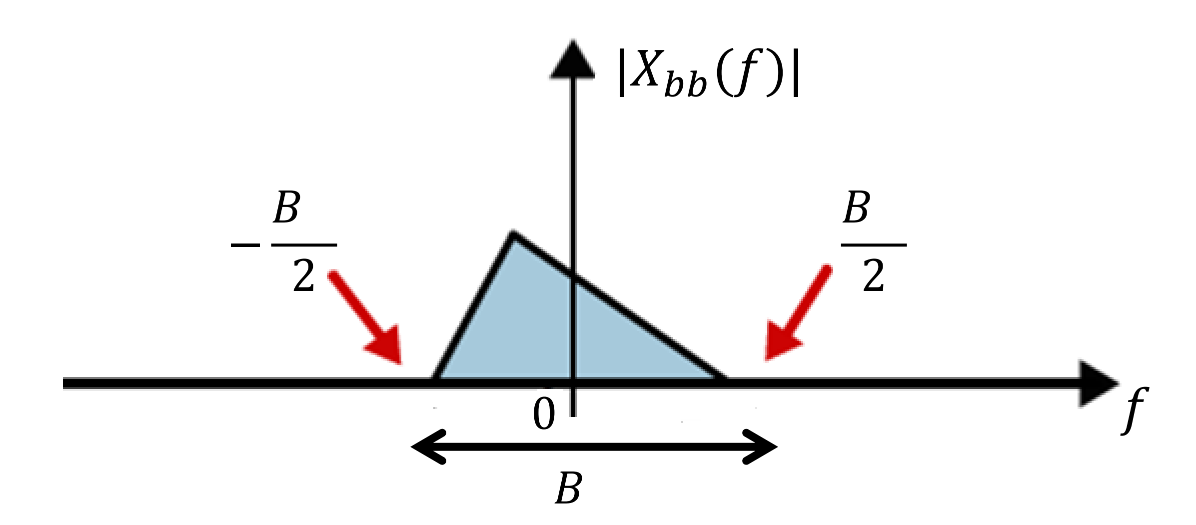


… suitable for transmission on a passband channel

- Passband signals are sent on a physical channel, and hence they are real. Therefore, their Fourier transform has Hermitian symmetry.

Why is the Baseband Equivalent Complex?

- Baseband signal: Fourier transform is non-zero only in a bandwidth B around the zero frequency



… not suitable for transmission on passband channel

- Given the generally asymmetric Fourier transform (as in the figure), baseband equivalent signals are complex in the time domain.

What is a Passband Signal?

- Mathematically, a bandpass signal is defined as:

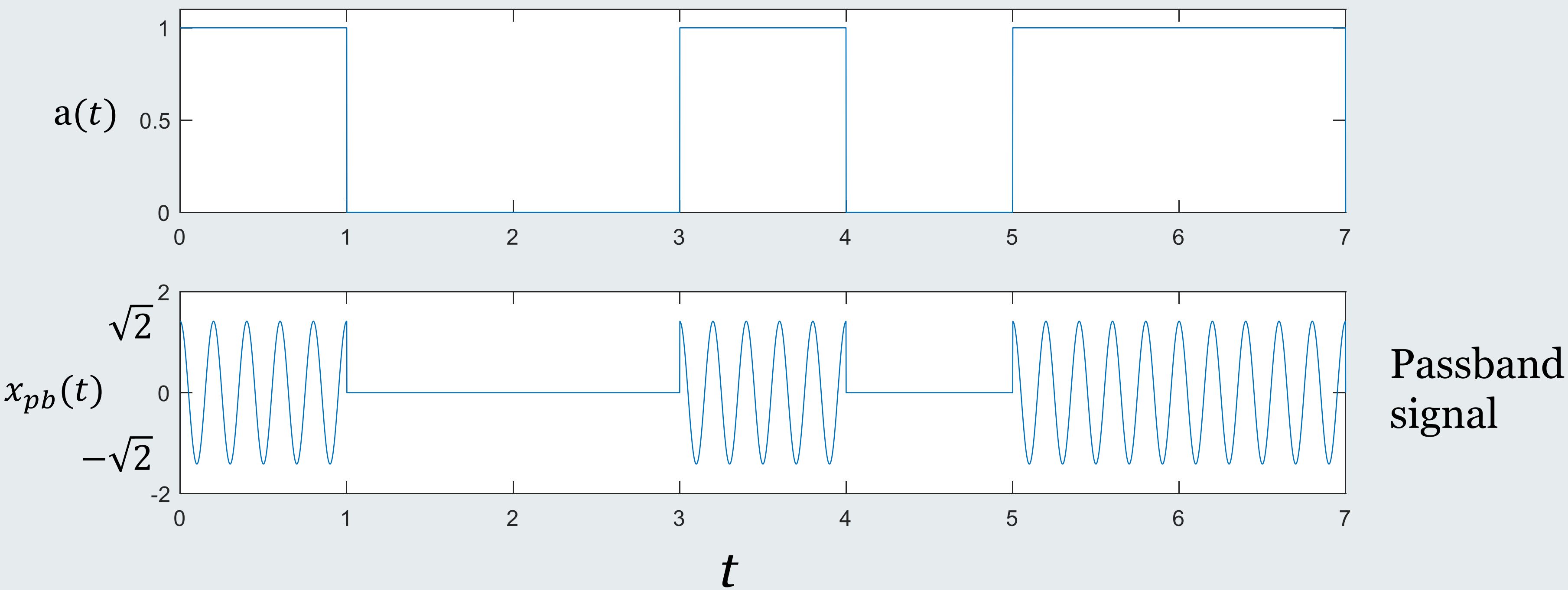
$$x_{pb}(t) = \sqrt{2}\, a(t) \cos(2\pi f_c t + \theta(t))$$

- Information is encoded by two baseband signals: amplitude $a(t)$ and phase $\theta(t)$.

- Note: The additional term $\sqrt{2}$ as compared to the expression in J. Cioffi's notes is introduced to simplify some of the later derivations.

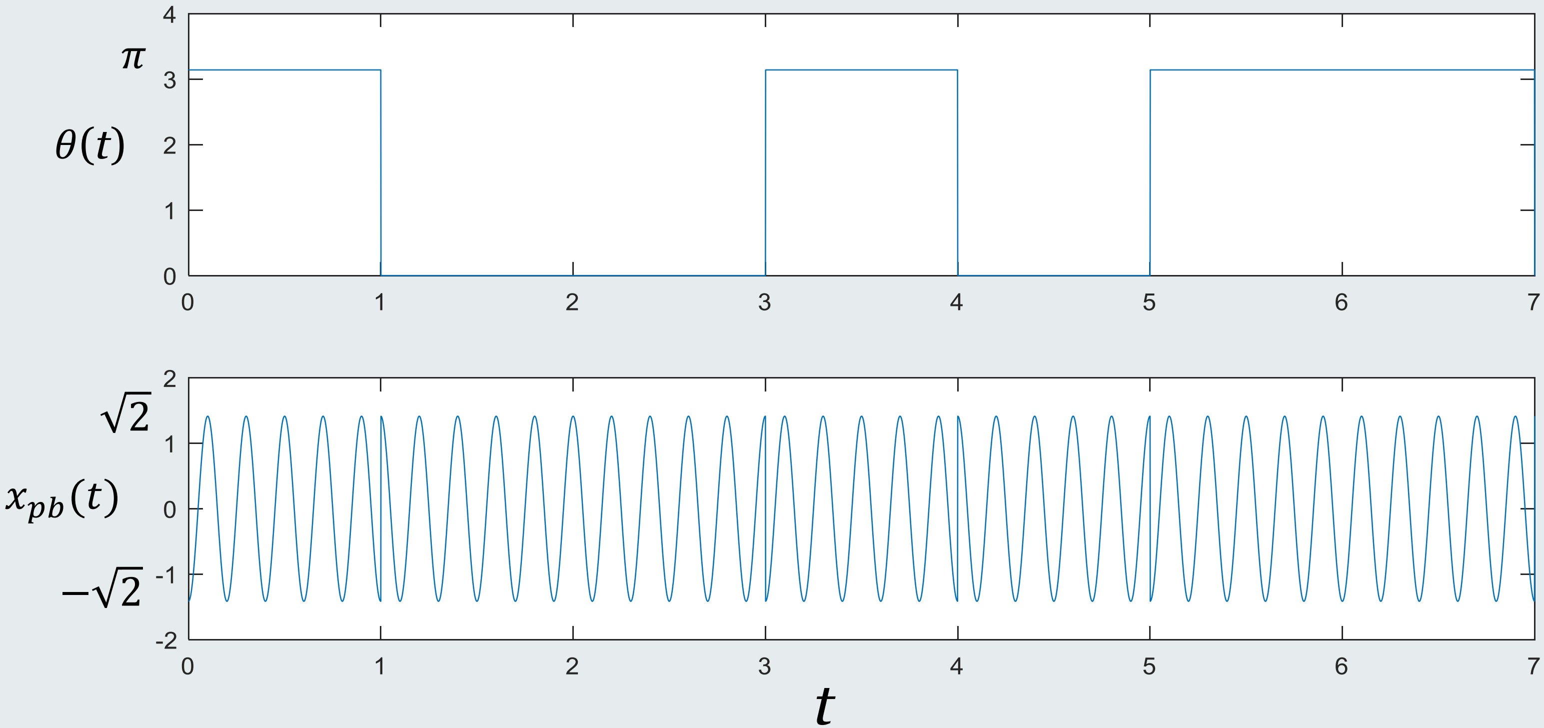
What is a Passband Signal?

- Example: Amplitude modulation ($\theta(t) = 0$)



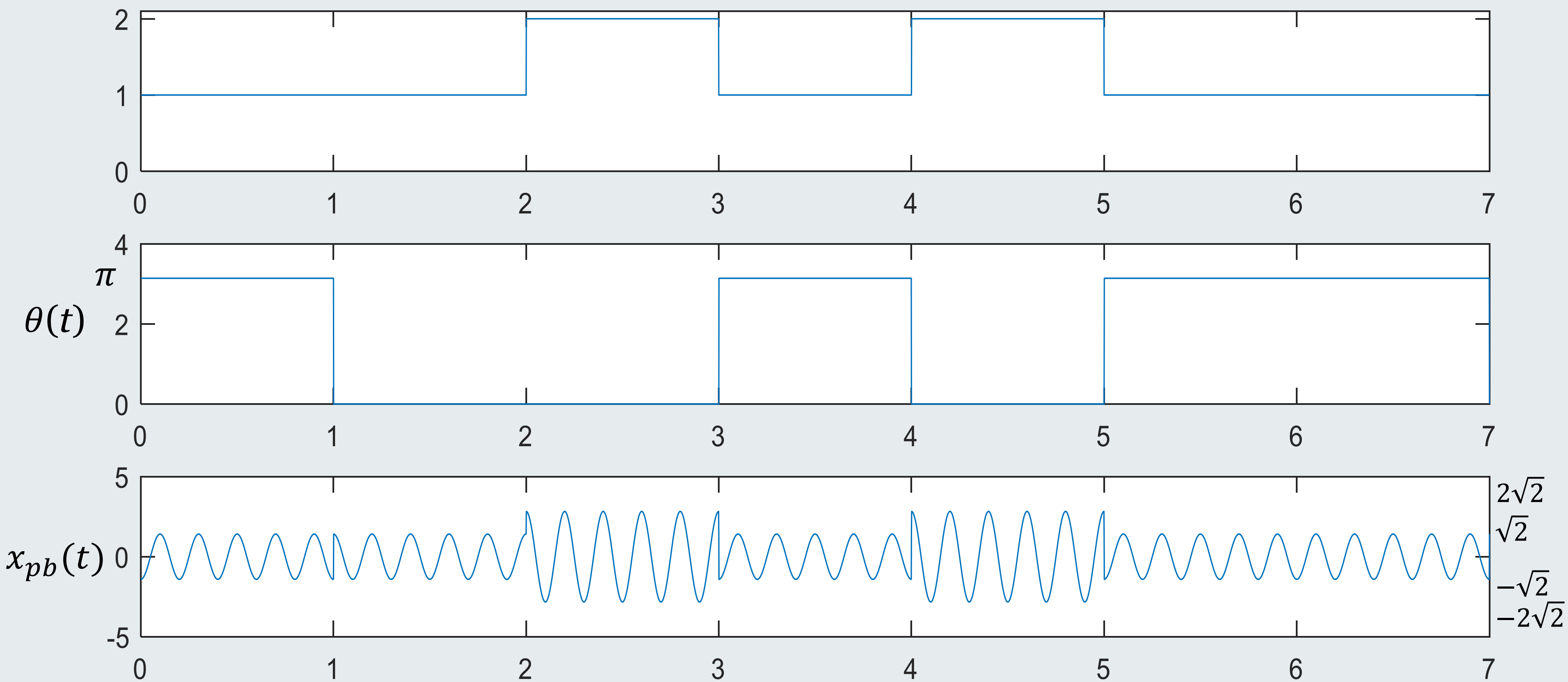
What is a Passband Signal?

- Example: Phase modulation ($a(t) = 1$)



What is a Passband Signal?

- Example: Amplitude and phase modulation



What is a Passband Signal?

- Alternative form of the bandpass signal

$$x_{pb}(t) = \sqrt{2}\, x_I(t) \cos(2\pi f_c t) - \sqrt{2}\, x_Q(t) \sin(2\pi f_c t)$$

in-phase, or I, component

quadrature, or Q, component

- Information is encoded by the two baseband signals $x_I(t)$ and $x_Q(t)$

– $x_I(t)$ modulates the in-phase carrier $\cos(2\pi f_c t)$

– $x_Q(t)$ modulates the quadrature carrier $-\sin(2\pi f_c t)$

What is a Passband Signal?

- From amplitude/phase representation to in-phase quadrature representation:

Recall that $\cos(a+b) = \cos a \cos b - \sin a \sin b$

Using this formula, we obtain

$$
\begin{aligned}
x_c(t) &= \sqrt{2}\, a(t) \cos(2\pi f_c t + \theta(t)) \\
&= \sqrt{2}\, \underbrace{a(t) \cos\theta(t)}_{x_I(t)} \cos(2\pi f_c t) - \sqrt{2}\, \underbrace{a(t) \sin\theta(t)}_{x_Q(t)} \sin(2\pi f_c t)
\end{aligned}
$$

and hence

$$
\begin{aligned}
x_I(t) &= a(t) \cos\theta(t) \\
x_Q(t) &= a(t) \sin\theta(t)
\end{aligned}
$$

What is a Passband Signal?

- From in-phase / quadrature representation to amplitude / phase representation:

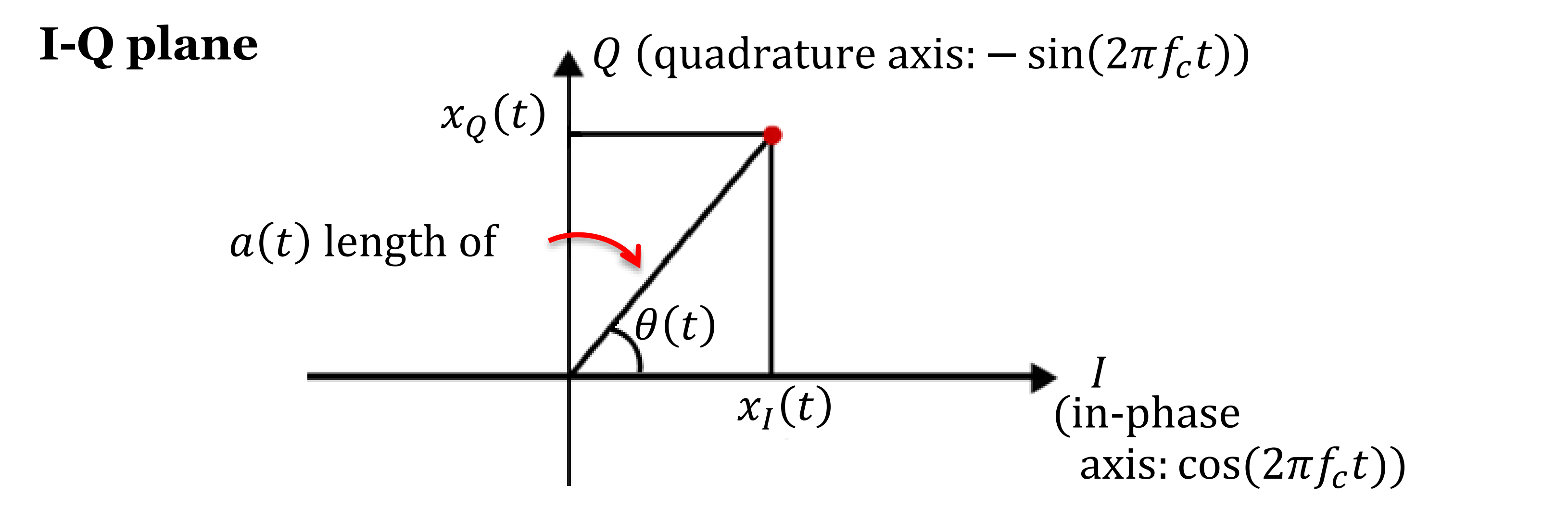
  - From the equations on the previous slide, we get

$$a(t) = \sqrt{x_I(t)^2 + x_Q(t)^2}$$

$$\theta(t) = \arg(x_I(t) + \mathrm{j}x_Q(\mathrm{t}))$$

Why Do We Use Complex Numbers to Represent Passband Signals?

- Illustration:



Information signal described by

$(x_I(t), x_Q(t))$ … cartesian coordinates

$(a(t), \theta(t))$ … polar coordinates

Remark: As t increases, the point ● moves on the I-Q plane

Remark: $\cos\left(2\pi f_c t + \frac{\pi}{2}\right) = -\sin(2\pi f_c t)$

What is the Baseband Equivalent Signal?

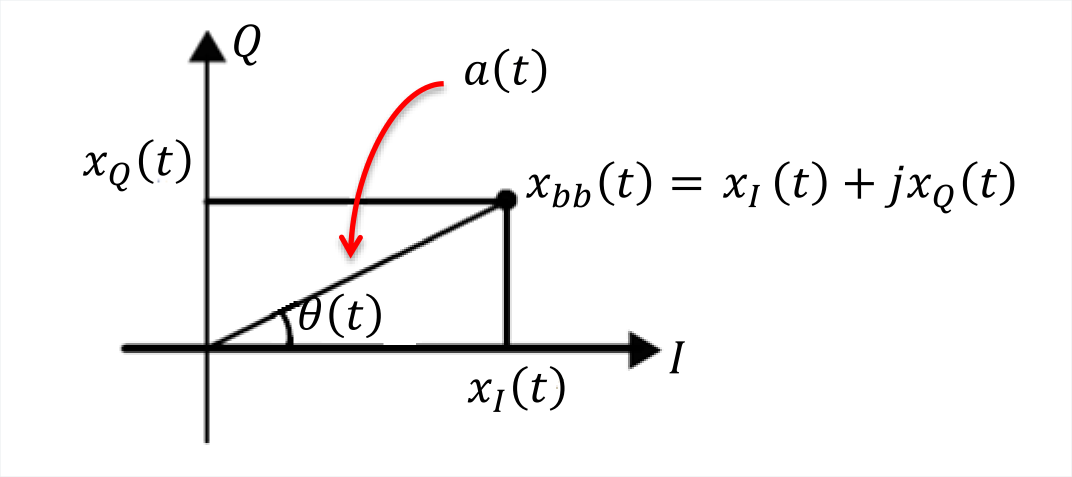
- Based on the discussion from the previous slide, the information signal is completely described by the pair of baseband signals $(a(t), \theta(t))$ or $\left(x_I(t), x_Q(t)\right)$.

- Baseband equivalent signal

$$x_{bb}(t) = x_I(t) + j\,x_Q(t) = a(t)e^{j\theta(t)}$$

What is the Baseband Equivalent Signal?

- Complex baseband representation

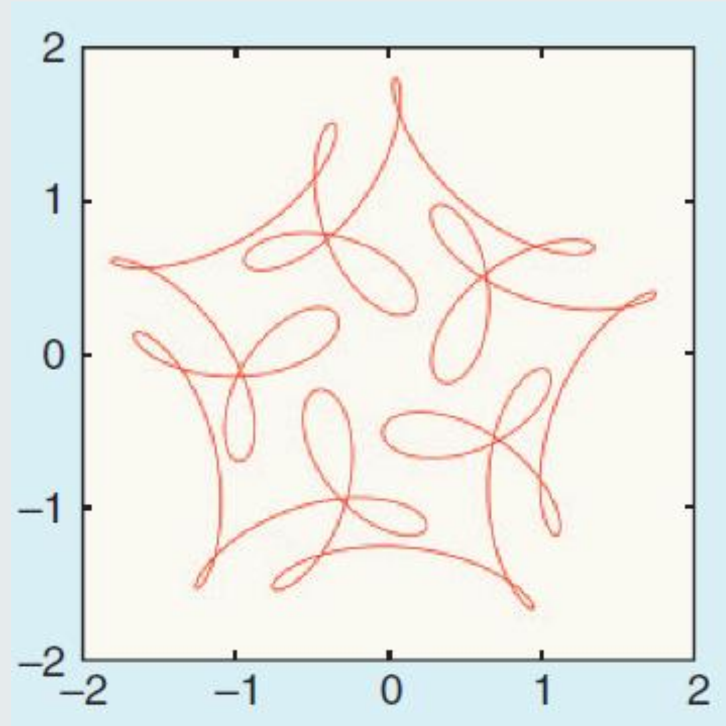


- $x_{bb}(t)$ is complex and baseband

What is the Baseband Equivalent Signal?

- Exercise: Draw the baseband equivalent signals in the complex plane for the examples in slides 17, 18 and 19.

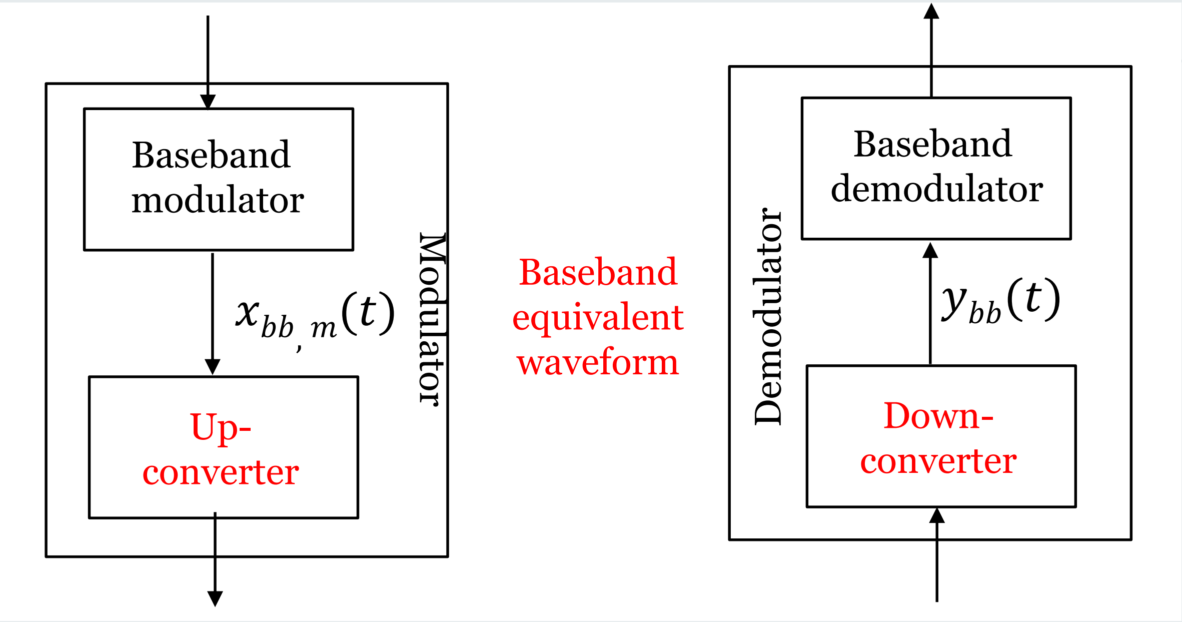
What is the Baseband Equivalent Signal?



$$
\begin{aligned}
c(t) &= x(t) + iy(t) \\
&= \left(\cos(t) + \frac{\cos(6t)}{2} + \frac{\sin(14t)}{3}\right) \\
&\quad + i\left(\sin(t) + \frac{\sin(6t)}{2} + \frac{\cos(14t)}{3}\right) \\
&= e^{it} + \frac{e^{i6t}}{2} + \frac{ie^{-14t}}{3}. \qquad (2)
\end{aligned}
$$

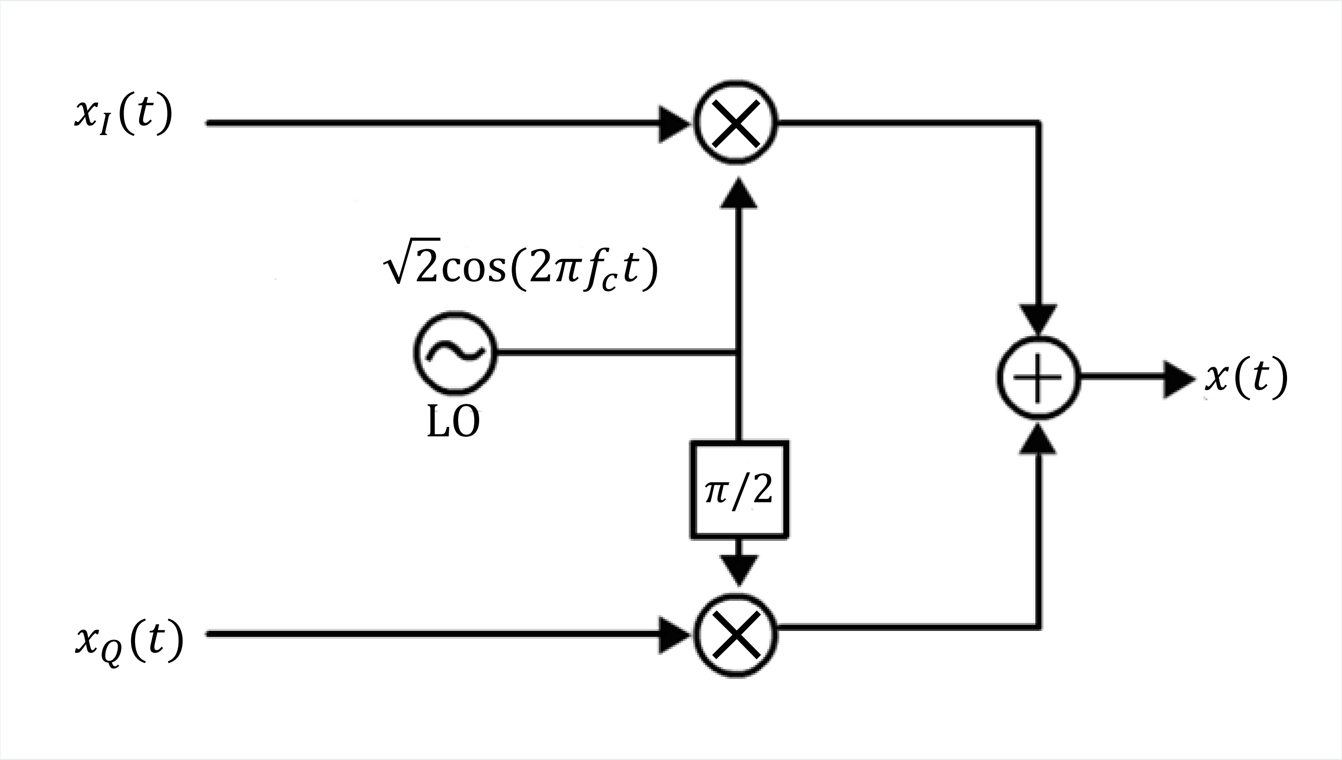
[Pei and Chang '17]

Why Should We Use the Baseband Equivalent Again?



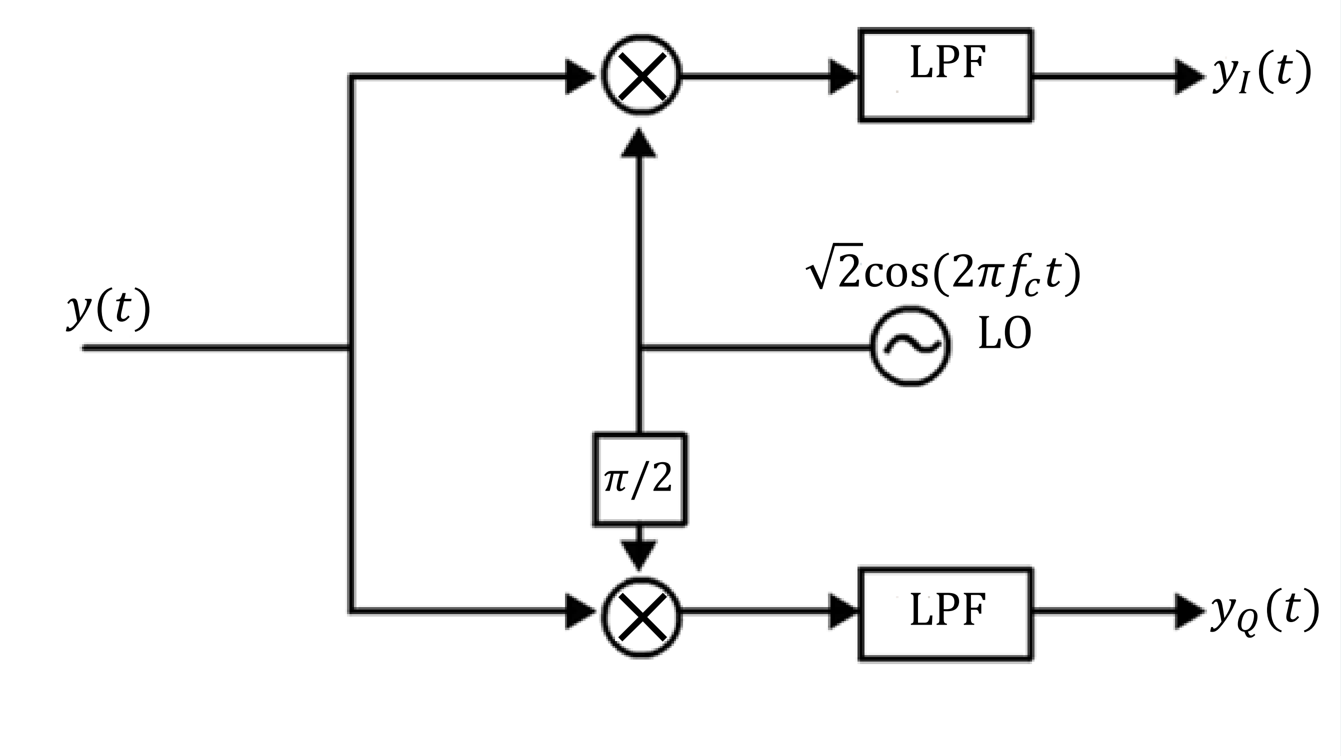
- Baseband signals are much easier (cheaper) to process than passband signals given the smaller frequencies involved.
- Baseband signals can be processed in the digital domain with Analog-to-Digital and Digital-to-Analog converters operating at feasible frequencies.
- Baseband signals are independent of the carrier frequencies and hence changing the carrier frequencies only requires to modify the up/down-converters.

What Does the Up-Converter Do?



LO = Local Oscillator

What Does the Down-Converter Do?



- Note: Up- and down-converters are fixed and need not be designed.

What Does the Down-Converter Do?

- **Why do we need LPF?** Consider the noiseless case $y(t)=x(t)$

$$y(t)\sqrt{2}\cos(2\pi f_c t) = 2x_I(t)\,(\cos(2\pi f_c t))^2$$

from LO

$$- 2\,x_Q(t)\sin(2\pi f_c t)\cos(2\pi f_c t)$$

$$= x_I\,(t) + \boxed{x_I(t)\cos(4\pi f_c t) - x_Q(t)\sin(4\pi f_c t)}$$

$\cos a \cos b = 1/2\,(\cos(a-b) + \cos(a+b))$

$\cos a \sin b = 1/2\,(\sin(a+b) - \sin(a-b))$

Removed by LPF

And similarly for $y(t)\left(-\sqrt{2}\sin(2\pi f_c t)\right) =$

$$= x_Q(t) - \boxed{x_Q(t)\cos(4\pi f_c t) - x_I(t)\cos(4\pi f_c t)}$$

Can We Write Directly $x_{pb}(t)$ as a Function of $x_{bb}(t)$?

- Define the analytic equivalent signal

$$x_A(t) = x_{bb}(t)e^{j2\pi f_c t}$$

- We then have:

$$\begin{aligned} x_{pb}(t) &= \sqrt{2}\, a(t)\cos(2\pi f_c t + \theta(t)) \\ &= \sqrt{2}\ x_I(t)\cos(2\pi f_c t) - \sqrt{2}\ x_Q(t)\sin(2\pi f_c t) \\ &= \sqrt{2}\ Re\{x_A(t)\} \end{aligned}$$

- To summarize, a passband signal can be represented in the following ways:

1. magnitude, phase	$a(t), \theta(t)$
2. inphase, quadrature	$x_I(t), x_Q(t)$
3. complex baseband	$x_{bb}(t)$
4. analytic	$x_A(t)$

How To Represent Passband Signals?

Example: Compute in-phase and quadrature components, as well as amplitude of

$$x(t) = \mathrm{sinc}(10^6 t) \cdot \cos(2\pi 10^7 t) + 3\mathrm{sinc}(10^6 t) \cdot \sin(2\pi 10^7 t)$$

How To Represent Passband Signals?

Example: Compute in-phase and quadrature components, as well as amplitude of

$$x(t) = \mathrm{sinc}(10^6 t) \cdot \cos(2\pi 10^7 t) + 3\mathrm{sinc}(10^6 t) \cdot \sin(2\pi 10^7 t)$$

$$x_I(t) = \frac{\mathrm{sinc}(10^6 \mathrm{t})}{\sqrt{2}}$$

$$x_{\mathrm{Q}}(t) = -\frac{3\mathrm{sinc}(10^6 \mathrm{t})}{\sqrt{2}}$$

$$x_{bb}(t) = \frac{\mathrm{sinc}(10^6 \mathrm{t})}{\sqrt{2}} - \mathrm{j}\frac{3\mathrm{sinc}(10^6 \mathrm{t})}{\sqrt{2}}$$

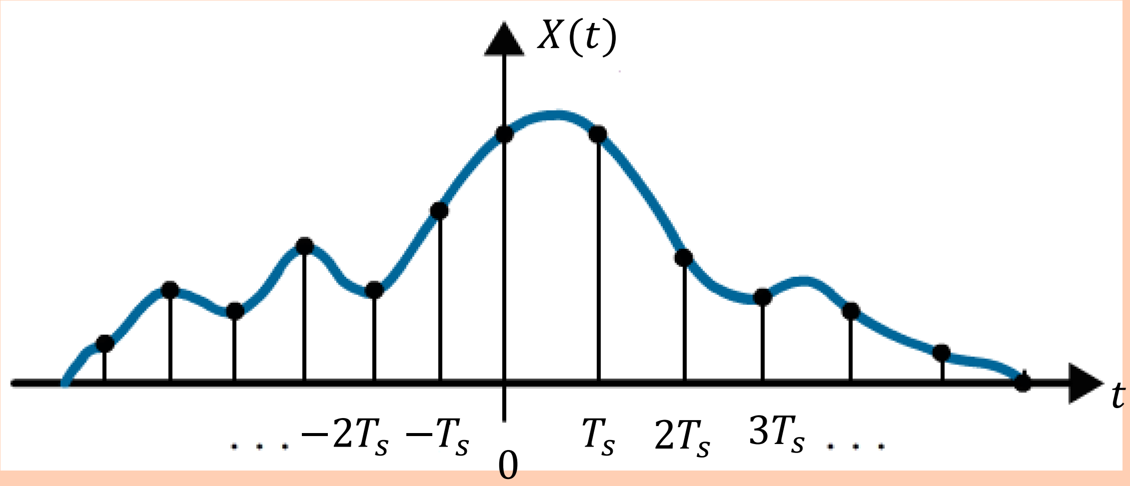
$$a(t) = \sqrt{5}\,|\mathrm{sinc}(10^6 \mathrm{t})|$$

Matlab: Plotting Baseband and Passband Signals

- A signal $x(t)$ is represented as the vector

$$x = [\ldots x(-2T_s), x(-T_s), x(0), x(T_s), x(2T_s), \ldots]$$

by sampling



- In order to guarantee that no information loss is incurred by sampling, the sampling rate needs to satisfy the condition of the Nyquist-Shannon theorem:

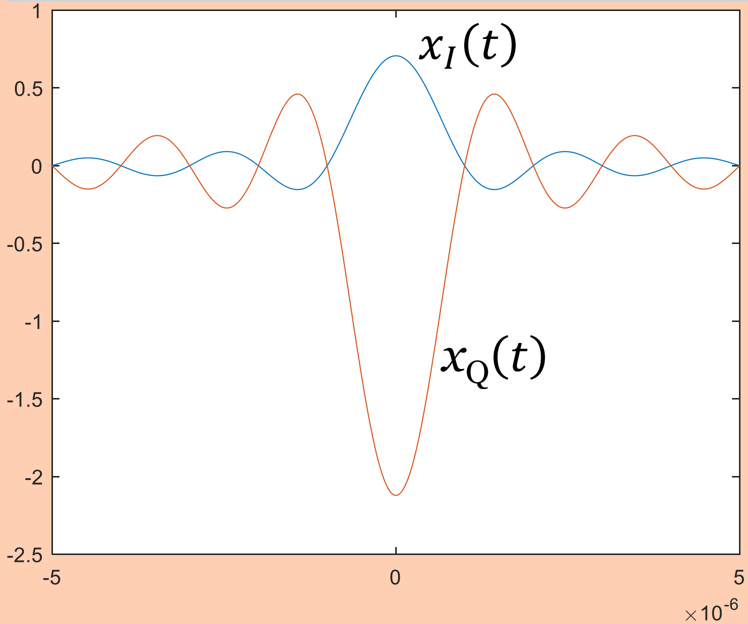
$$\frac{1}{T_s} \geq 2 \times \text{ highest frequency of } X(f)$$

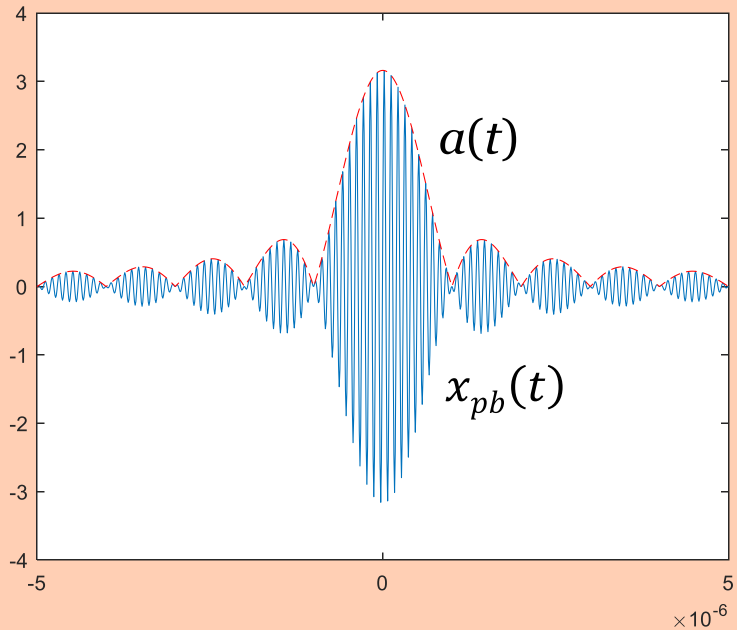
Matlab: Plotting Baseband and Passband Signals

Plotting signal from the previous example

```
Ts=10^-8; %sampling interval (it satisfies Shannon-Nyquist)
t=[-5*10^-6:Ts:5*10^-6];

x=sinc(10^6*t).*cos(2*pi*10^7*t)+3*sinc(10^6*t).*sin(2*pi*10^7*t);
plot(t,x)
hold on
a=sqrt(5)*abs(sinc(10^6*t));
plot(t,sqrt(2)*a, 'r--');

figure
xi=sinc(10^6*t)/sqrt(2);
xq=-3*sinc(10^6*t)/sqrt(2);
plot(t,xi);
hold on; plot(t,xq, 'r');
```

Matlab: Plotting Baseband and Passband Signals





How Do We Obtain the Fourier Transform of the Baseband Equivalent?

- Using the frequency translation property of the Fourier transform, we can calculate

$$
\begin{aligned}
X_{pb}(f) &= \mathcal{F}\{x_{pb}(t)\} \\
&= \sqrt{2}\,\mathcal{F}\{x_I(t)\cos(2\pi f_c t)\} - \sqrt{2}\,\mathcal{F}\{x_Q(t)\sin(2\pi f_c t)\} \\
&= \frac{1}{\sqrt{2}}\big(X_I(f-f_c) + X_I(f+f_c)\big) - \frac{1}{\sqrt{2}\,j}\Big(X_Q(f-f_c) - X_Q(f+f_c)\Big) \\
&= \frac{X_I(f-f_c) + jX_Q(f-f_c)}{\sqrt{2}} + \frac{X_I(f+f_c) - jX_Q(f+f_c)}{\sqrt{2}} \\
&= \frac{1}{\sqrt{2}}\,X_{bb}(f-f_c) + \frac{1}{\sqrt{2}}X_{bb}^*(-f-f_c)
\end{aligned}
$$

where the last equality follows from the Hermitian symmetry of $X_I(f)$ and $X_Q(f)$, since

$$X_{bb}^*(-f-f_c) = X_I^*(-f-f_c) - j\,X_Q^*(-f-f_c) = X_I(f+f_c) - j\,X_Q(f+f_c)$$

How Do We Obtain the Fourier Transform of the Baseband Equivalent?

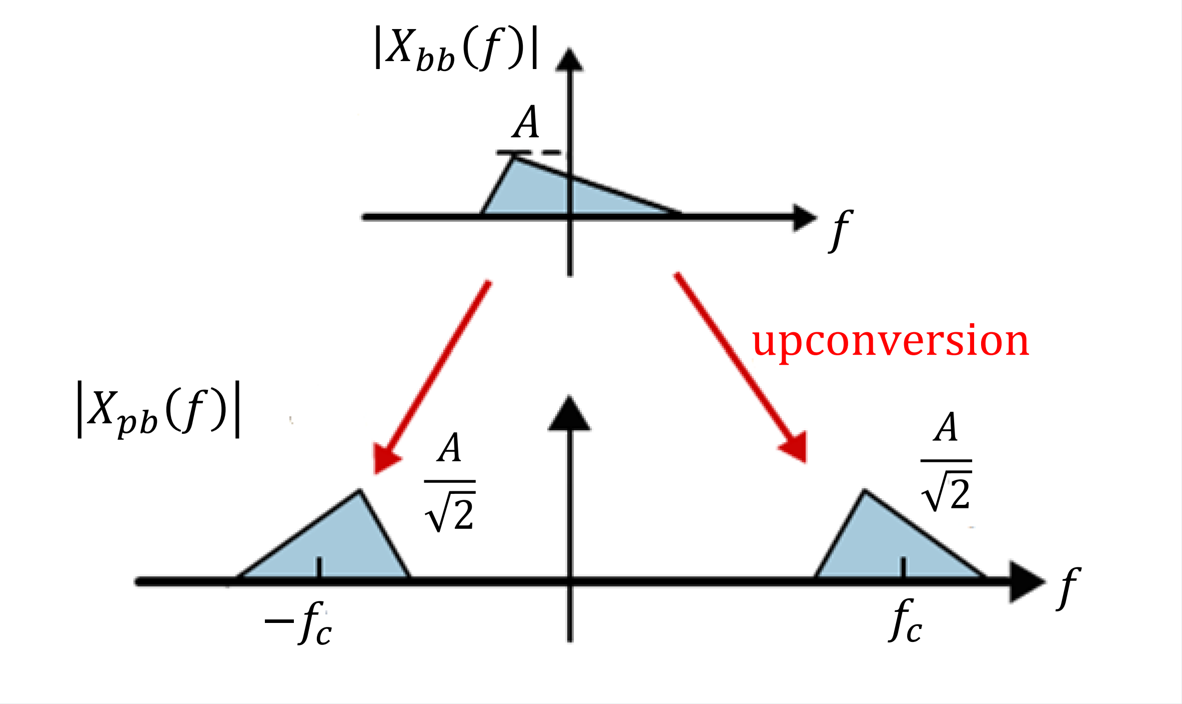
- We have shown that

$$X_{pb}(f) = \frac{1}{\sqrt{2}} X_{bb}(f - f_c) + \frac{1}{\sqrt{2}} X_{bb}^*(-f - f_c)$$

- Remark: $X_c(f)$ satisfies Hermitian symmetry

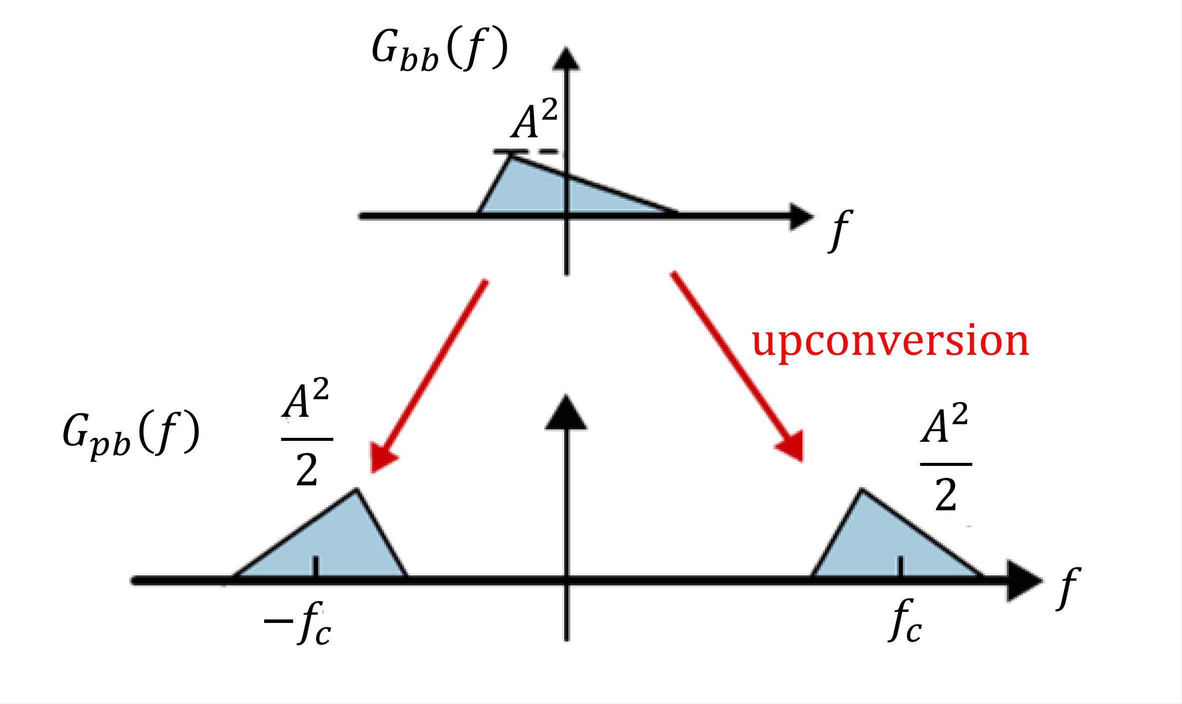
How Do We Obtain the Fourier Transform of the Baseband Equivalent?

- The relationship between $X_{bb}(f)$ and $X_{pb}(f)$ consists of an upconversion operation that guarantees Hermitian symmetry.

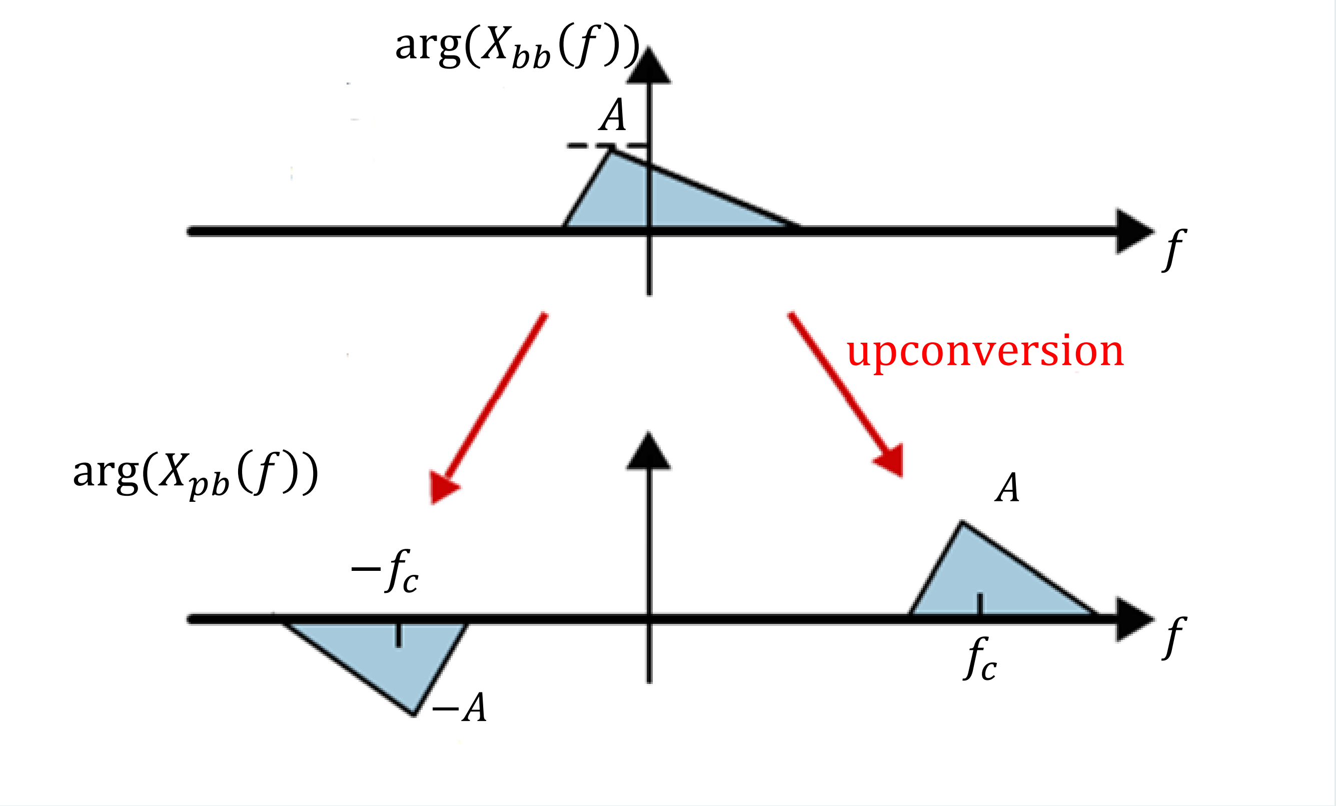


How Do We Obtain the Fourier Transform of the Baseband Equivalent?

- The relationship between $X_{bb}(f)$ and $X_{pb}(f)$ consists of an upconversion operation that guarantees Hermitian symmetry.

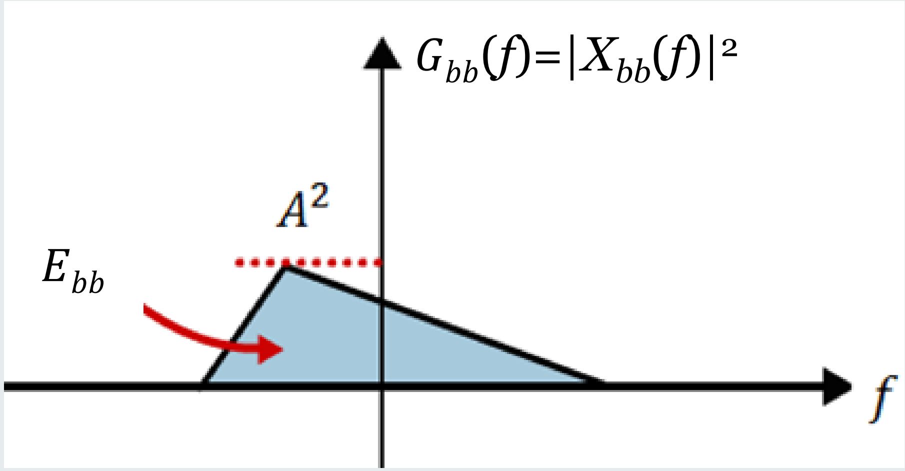


How Do We Obtain the Fourier Transform of the Baseband Equivalent?



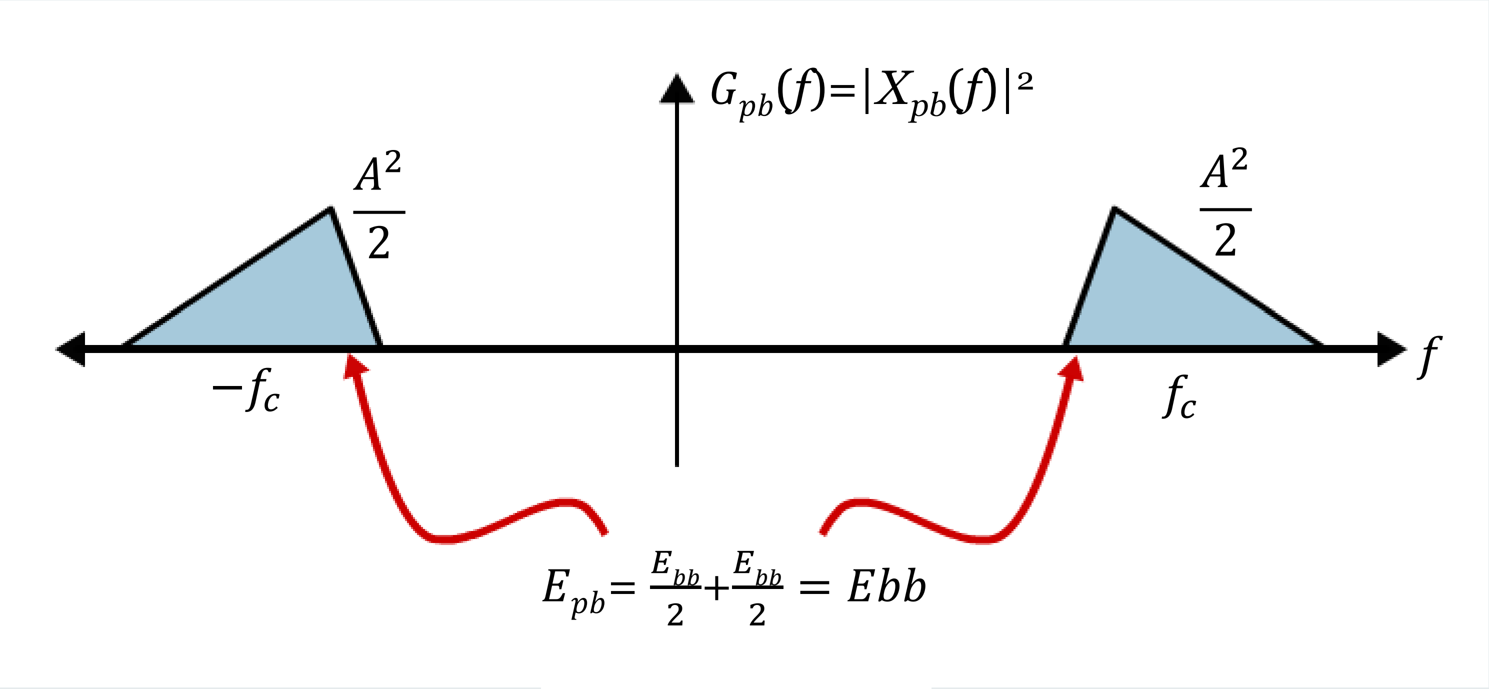
Why Did We Add That sqrt(2) Term Again?

- In this way, the energy of $x_{pb}(t)$ is the same as the energy of $x_{bb}(t)$.
- This can be seen using Rayleigh theorem, as illustrated below.



energy spectrum and energy of the baseband signal

Why Did We Add That sqrt(2) Term Again?

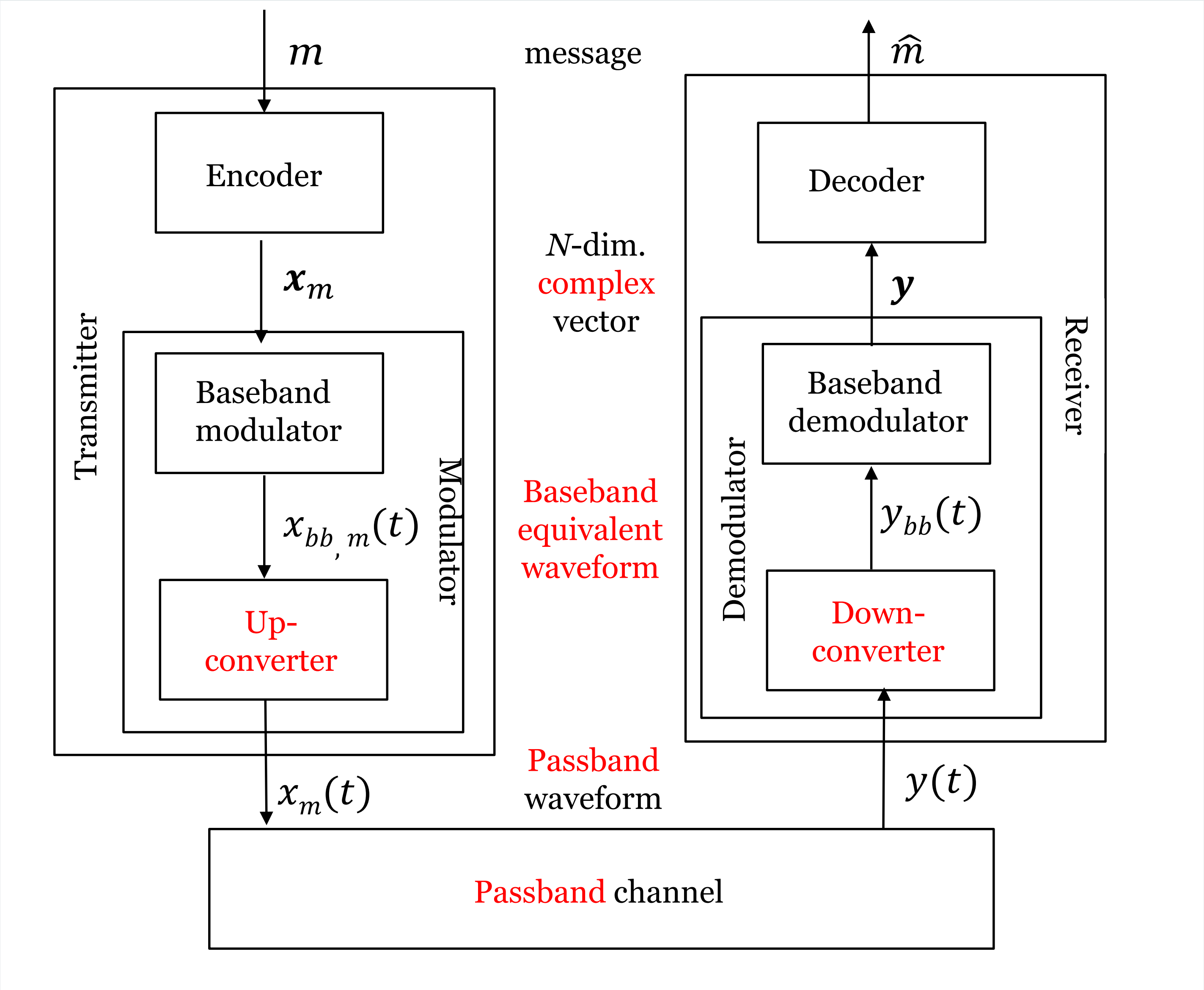


What Does This Mean for Coding and Modulation?

- The signal set produced by the baseband modulator is given by complex functions $x_{bb,m}(t)$ for $\mathrm{m} = 0,1,\ldots,2^{M-1}$.

- Therefore, both constellation and orthonormal basis functions are generally complex.

What Does This Mean for Coding and Modulation?



How to Encode?

- **Encoder**:

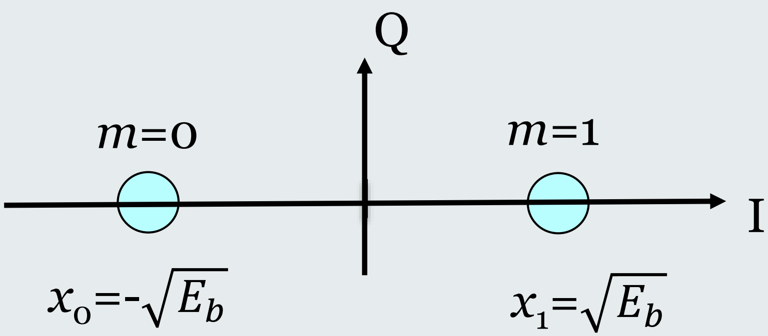
$$\text{message } m \in \{0, \ldots, M-1\} \rightarrow \text{symbol } (\textbf{complex} \text{ vector}) \; \boldsymbol{x}_m = \begin{bmatrix} x_{m,1} \\ x_{m,2} \\ \vdots \\ x_{m,N} \end{bmatrix}$$

- The signal space is $\mathbb{C}^N$

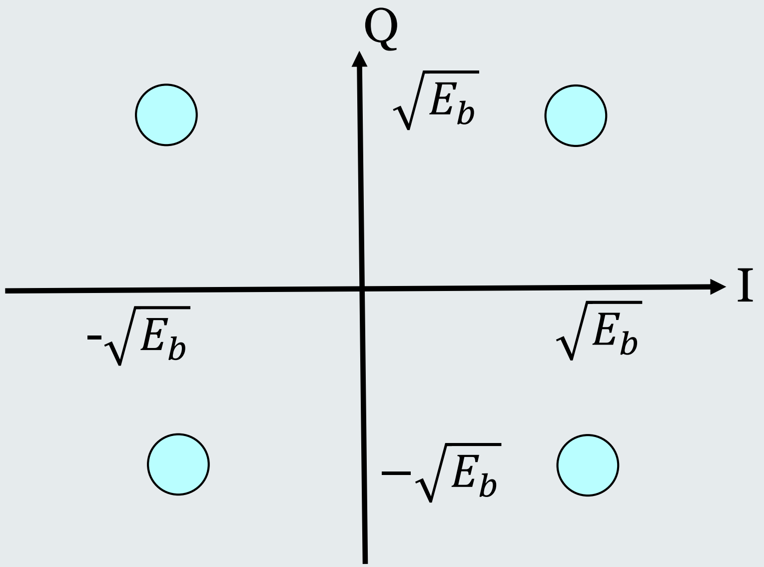
- The set of all symbols $\boldsymbol{x}_m$, m=1,…,M, is the signal constellation

How to Encode?

Examples: N=1



M=2
BPSK



M=4
4-PSK or QPSK

- More generally, any constellation that we have encountered earlier can be reinterpreted by considering every pair of real dimensions as a complex dimension.
- The reason for the (I,Q) labeling will be detailed shortly.

How to Modulate?

- **Modulator**:

  symbol $\boldsymbol{x}_m \rightarrow$ (analog and continuous-time) **complex** waveform $x_m(t)$ of duration T seconds (symbol period)

- Signal set:

$$x_m(t) = \sum_{n=1}^{N} x_{m,n} \varphi_n(t)$$

- The set $\{\varphi_n(t)\}$, $n = 1, \dots, N$ is an N-dimensional **orthonormal basis**:

$$\langle \varphi_m(t), \varphi_n(t) \rangle = 0 \text{ for } m \neq n \text{ (orthogonal)}$$
$$\langle \varphi_m(t), \varphi_m(t) \rangle = ||\varphi_m(t)||^2 = 1 \text{ (normalized)}$$

Some Math

- **Inner product or correlation:**

  - between complex vectors

$$\langle \mathbf{v}, \mathbf{u} \rangle = \sum_{n=1}^{N} v_n u_n^*$$

  - between complex functions

$$\langle v(t), u(t) \rangle = \int_{-\infty}^{\infty} v(t) u^*(t) dt$$

Some Math

- **Squared Euclidean norm or energy:**

  - for a complex vector

$$||\mathbf{v}||^2 = \langle \mathbf{v}, \mathbf{v} \rangle = \sum_{n=1}^{N} |v_n|^2$$

  - for a complex function

$$||v(t)||^2 = \langle v(t), v(t) \rangle = \int_{-\infty}^{\infty} |v(t)|^2 dt$$

Some Math

- **Orthogonality:**

  - for complex vectors: vectors **v** and **u** are orthogonal if

$$\langle \mathbf{v}, \mathbf{u} \rangle = 0$$

  - for complex functions: functions $v(t)$ and $u(t)$ are orthogonal if

$$\langle v(t), u(t) \rangle = 0$$

- Note that orthogonal pairs of real functions are also orthogonal when interpreted as complex functions (with zero imaginary part).

How To Modulate?

Examples:

1) N=1

$$\varphi_1(t) = \frac{1}{\sqrt{T}} \operatorname{sinc}\left(\frac{t}{T}\right)$$

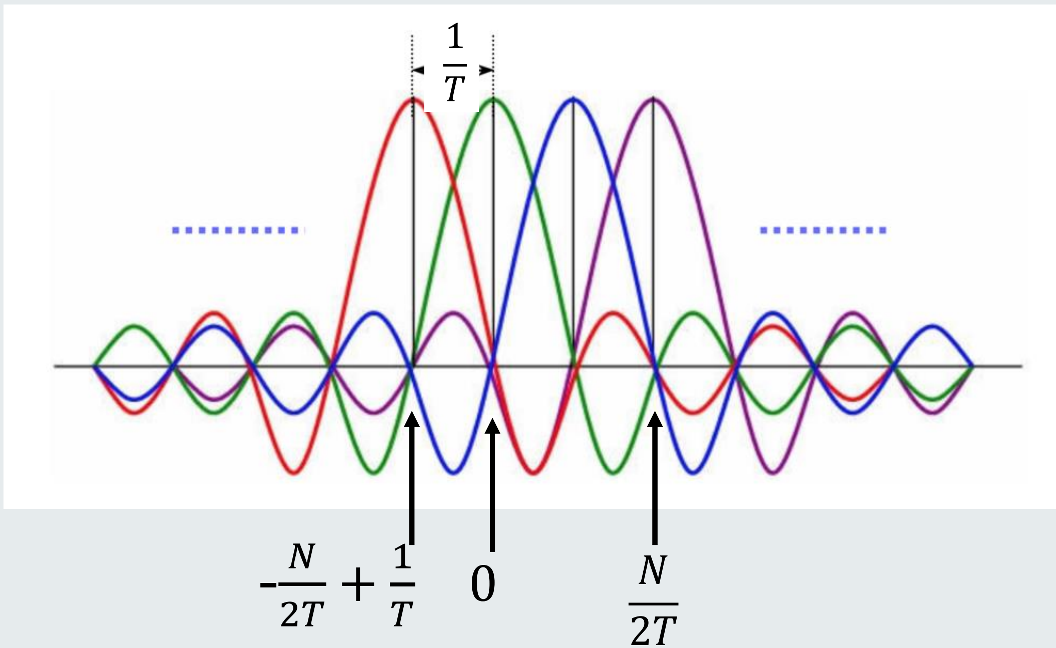
2) Any N: Stream modulation

$$\varphi_n(t) = \frac{1}{\sqrt{T}} \operatorname{sinc}\left(\frac{t-nT}{T}\right) \text{ for } n=1,\ldots,N$$

3) Any N: Orthogonal Frequency Division Multiplexing (OFDM)

$$\varphi_n(t) = \frac{1}{\sqrt{T}} \operatorname{rect}\left(\frac{t}{T}\right) \exp\left(j2\pi\left(\frac{n}{T} - \frac{N}{2T}\right)t\right) \text{ for } n=1,\ldots,N$$

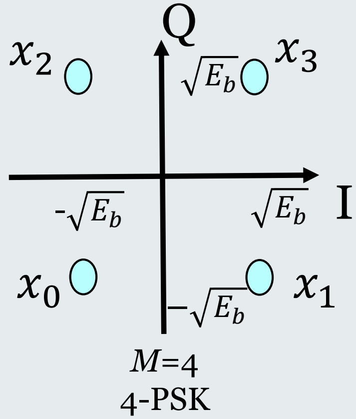
How To Modulate?



$$B \approx \frac{N}{T}$$

How To Modulate?

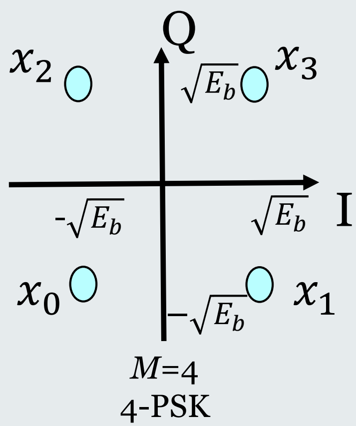
Example: Given the 4-PSK constellation and the basis function indicated below, find the signal set. Explain why the axes of the constellation are labeled as I and Q.



$$\varphi_1(t) = \frac{1}{\sqrt{T}} sinc\left(\frac{t}{T}\right)$$

How To Modulate?

Example: Given the 4-PSK constellation and the basis function indicated below, find the signal set. Explain why the axes of the constellation are labeled as I and Q.



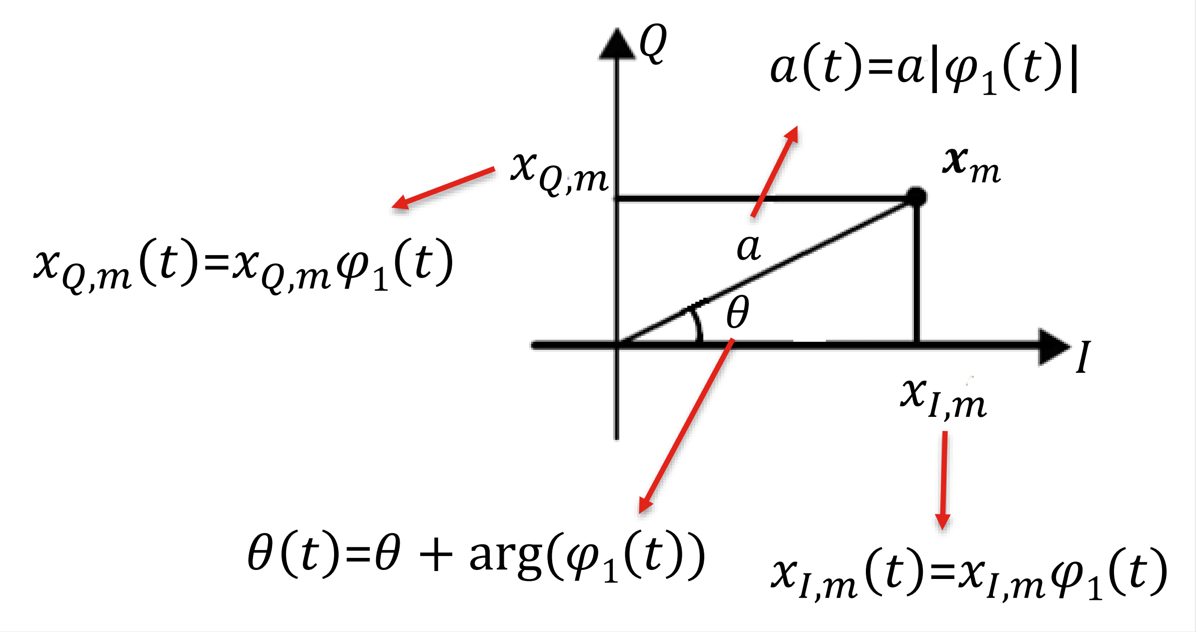
$$\varphi_1(t) = \frac{1}{\sqrt{T}} sinc\left(\frac{t}{T}\right)$$

$$x_{bb,0}(t) =- \sqrt{\frac{E_b}{T}}\,\text{sinc}\left(\frac{t}{T}\right) - j\sqrt{\frac{E_b}{T}}\,\text{sinc}\left(\frac{t}{T}\right),\quad x_{bb,1}(t) = \sqrt{\frac{E_b}{T}}\,\text{sinc}\left(\frac{t}{T}\right) - j\sqrt{\frac{E_b}{T}}\,\text{sinc}\left(\frac{t}{T}\right)$$

$$x_{bb,2}(t) =- \sqrt{\frac{E_b}{T}}\,\text{sinc}\left(\frac{t}{T}\right) + j\sqrt{\frac{E_b}{T}}\,\text{sinc}\left(\frac{t}{T}\right),\quad x_{bb,3}(t) = \sqrt{\frac{E_b}{T}}\,\text{sinc}\left(\frac{t}{T}\right) + j\sqrt{\frac{E_b}{T}}\,\text{sinc}\left(\frac{t}{T}\right)$$

How To Modulate?

- As shown in the previous example, with a real basis function, the real part of the constellation point is proportional to the I-component of the baseband modulated signal and the imaginary part is proportional to the Q-component.



How To Modulate?

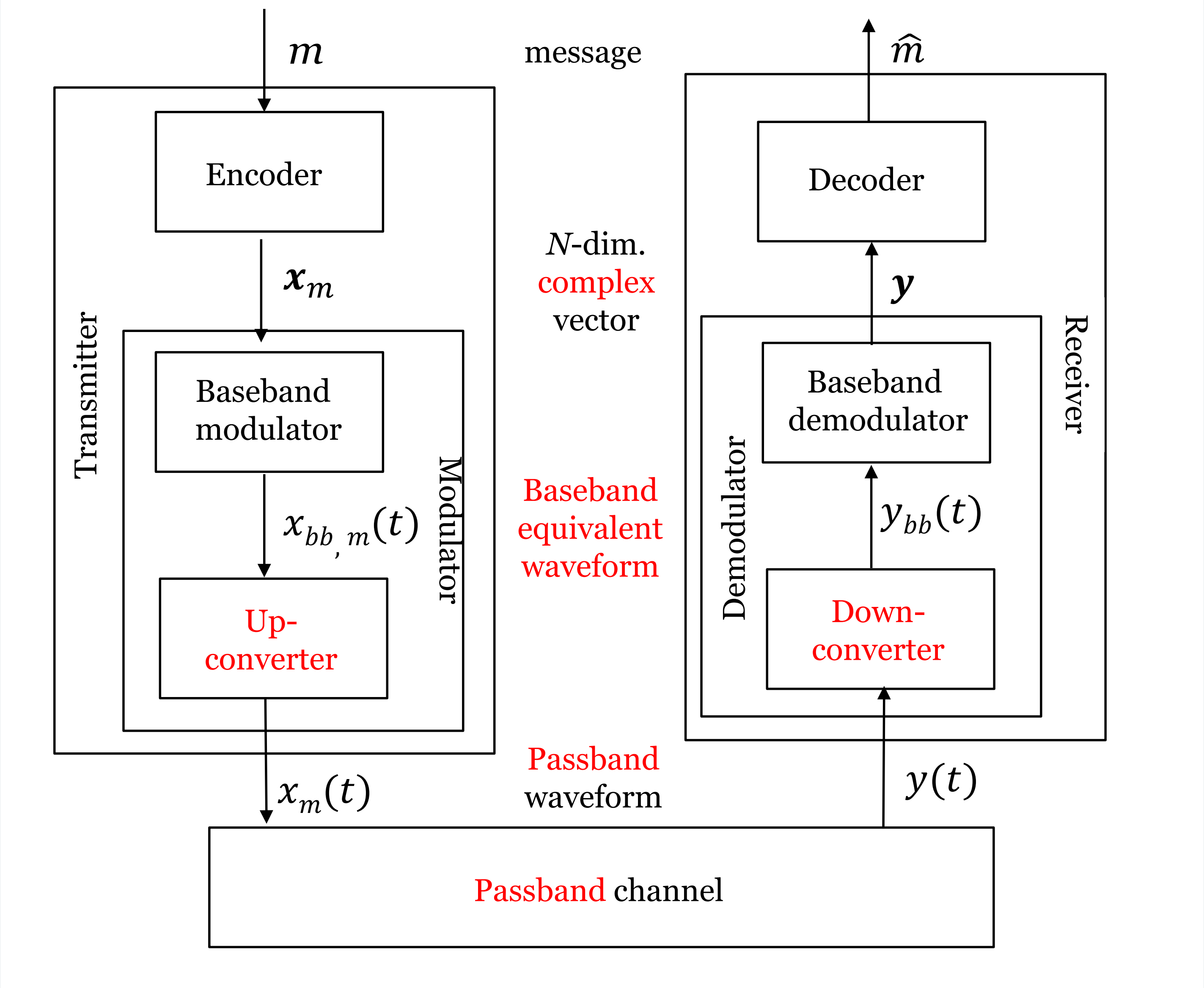
- The same applies for stream modulation for each symbol, as well as to OFDM for each subcarrier.

How To Modulate?

- The maximum number of complex dimensions is given as $N=BT$, where B is the bandwidth of the passband signal, or equivalently the bandwidth of the baseband equivalent including also the negative frequencies.

- This is consistent with the general results in Chapter 2 since one complex dimension amounts to two real dimensions. In other words, the number of real dimensions is still $N=2BT$.

What Does This Mean for Decoding and Demodulation?



What Does This Mean for Decoding and Demodulation?

- Following the same reasoning as in the previous chapter, the optimal receiver can be obtained as the cascade of
    - Down-converter
    - Correlative or matched filter baseband demodulator (with complex correlations or impulse responses)
    - MAP decoder (or ML in case of equally likely messages)

How to Demodulate?

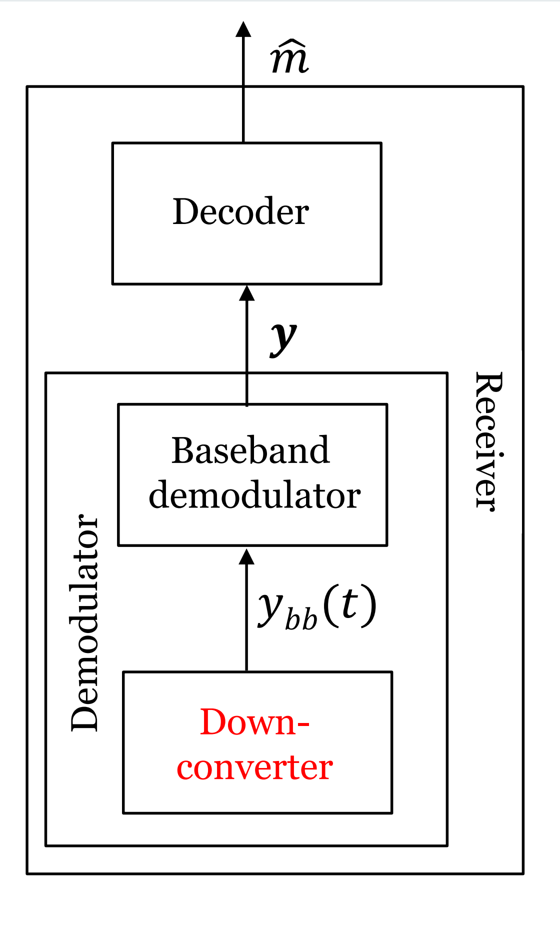
- After down-conversion, we have

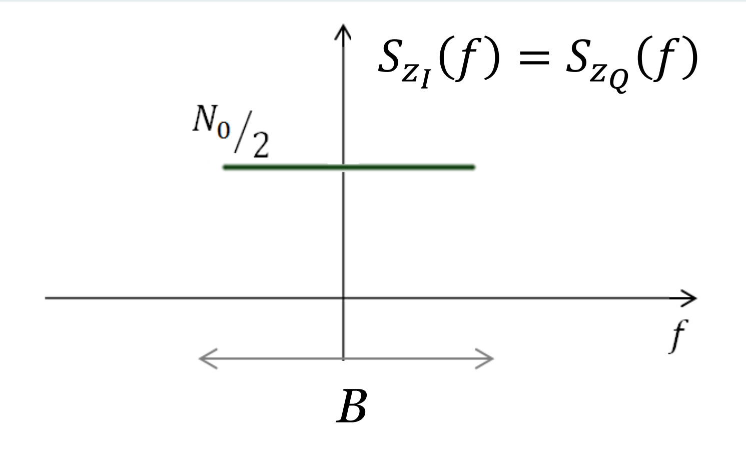
$$y_{bb}(t) = x_{bb}(t) + z_{bb}(t)$$

with complex baseband noise

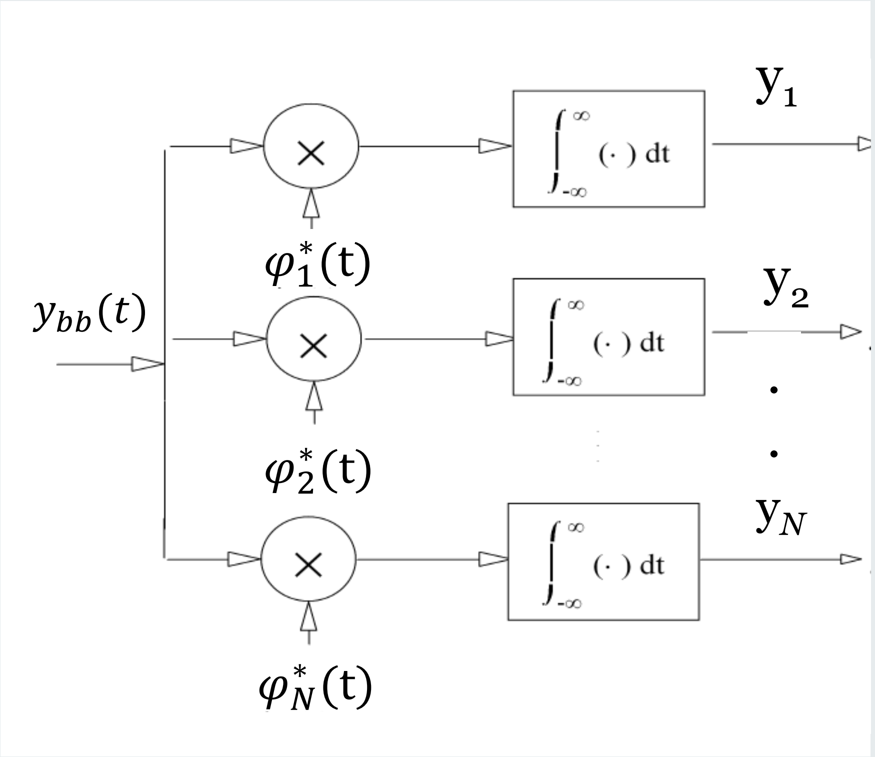
$$z_{bb}(t) = z_I(t) + jz_Q(t)$$

- The in-phase and quadrature components of the noise are independent WGN processes with power spectral density N0/2 (to be discussed).



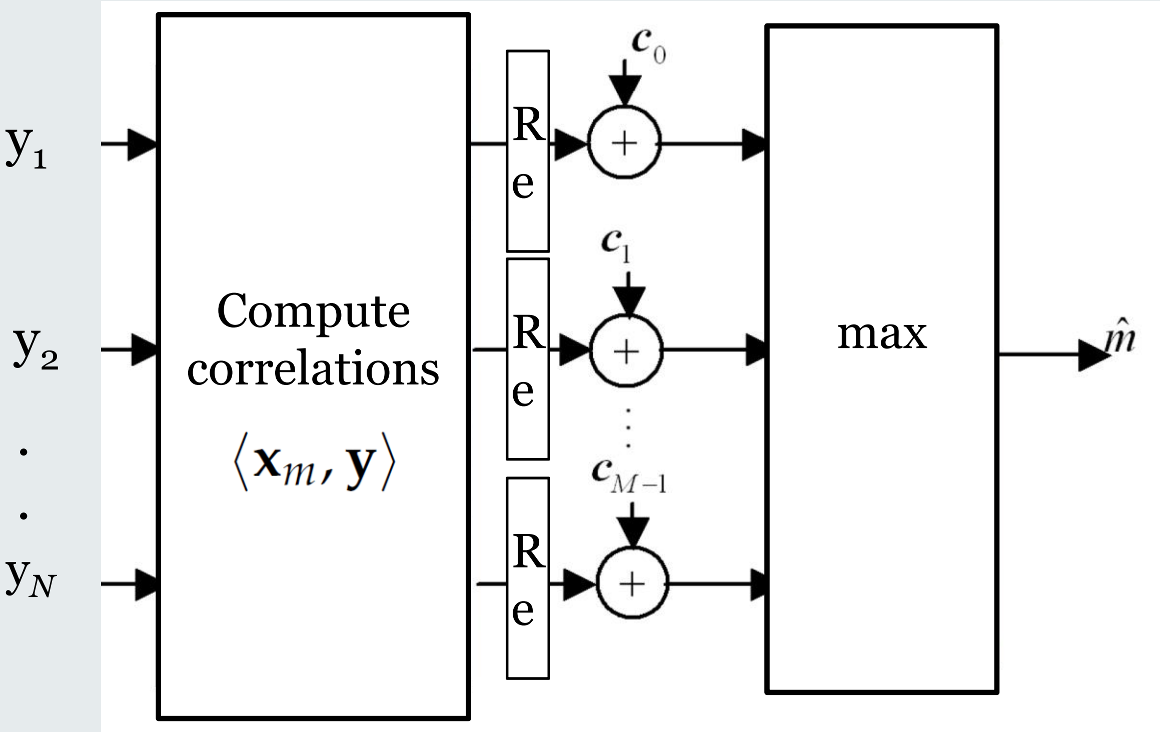


How to Demodulate?



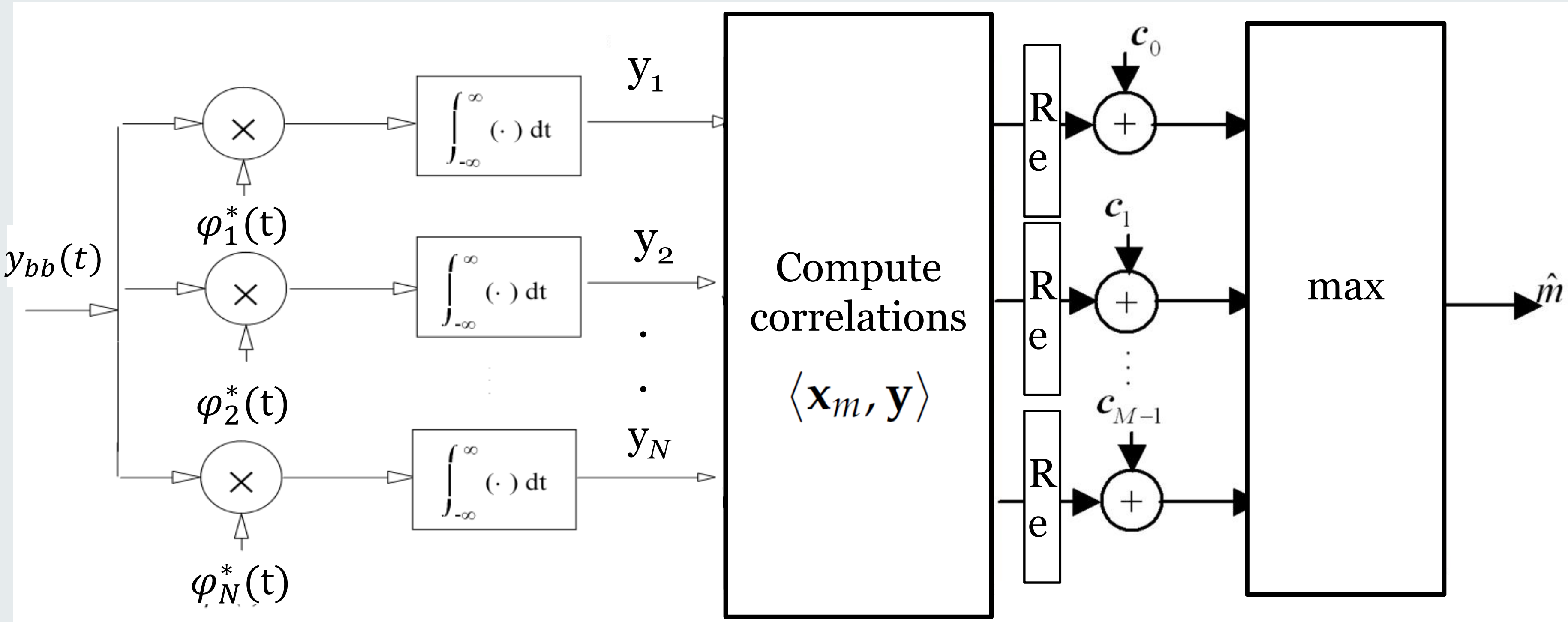
correlative demodulator

How to Decode?



$$c_m = \frac{N_0}{2}\log(p(m)) - \frac{\|\mathbf{x}_m\|^2}{2}$$

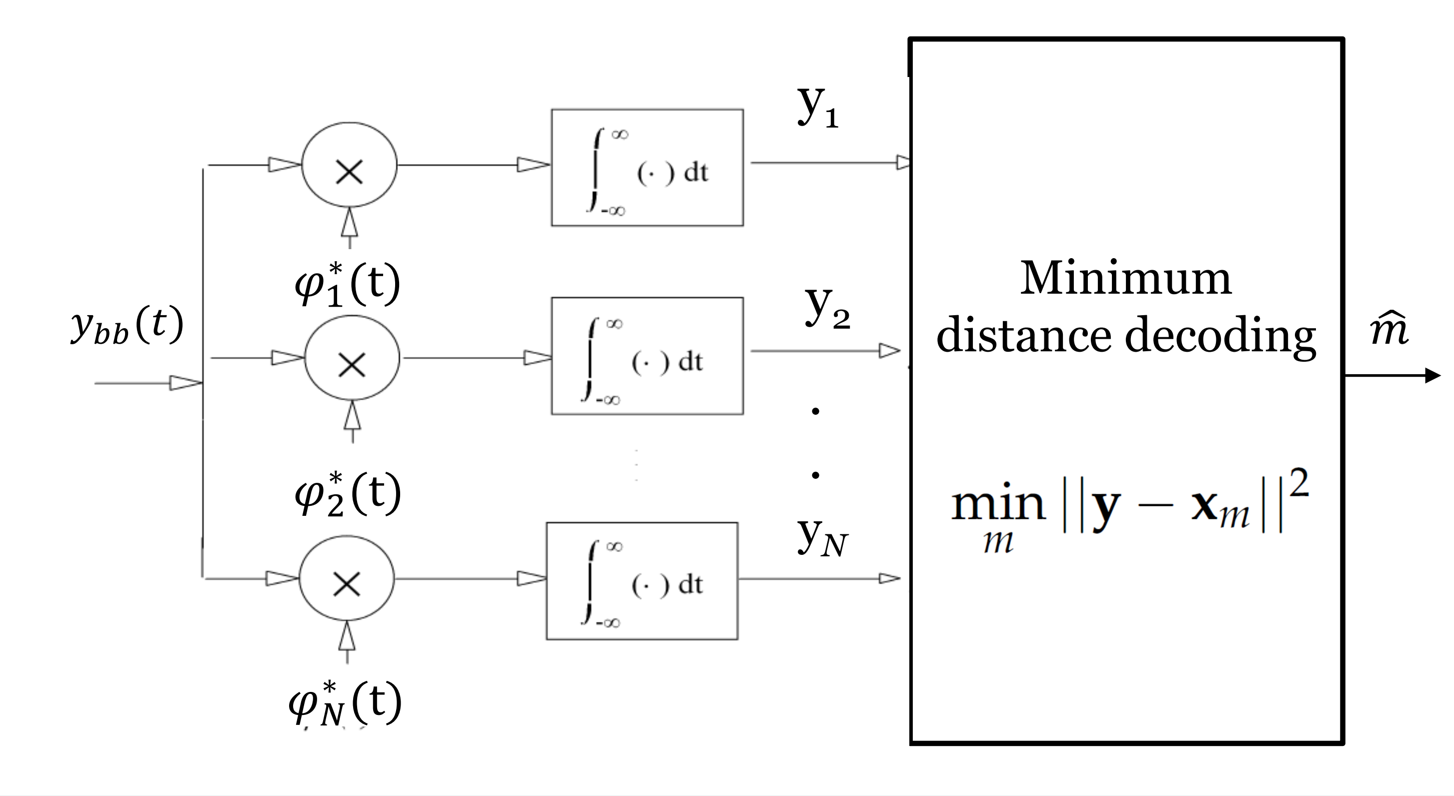
How to Demodulate and Decode?



$$c_m = \frac{N_0}{2}\log(p(m)) - \frac{\|\mathbf{x}_m\|^2}{2}$$

How to Demodulate and Decode?

- When the messages are equally likely, we can also use minimum distance decoding.



How to Demodulate and Decode?

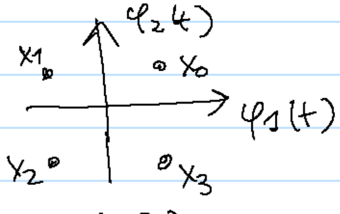
- **Problem:** Prove that the scores of the four messages in 4-PSK obtained by the receiver with a correlative decoder are the same as those obtained by the optimal receiver studied in the previous chapter in the absence of noise. Consider as an example the transmission of message m=0.

How to Demodulate and Decode?

Solution:

1) Following Chapter 2:

encoder:



$(N=2)$

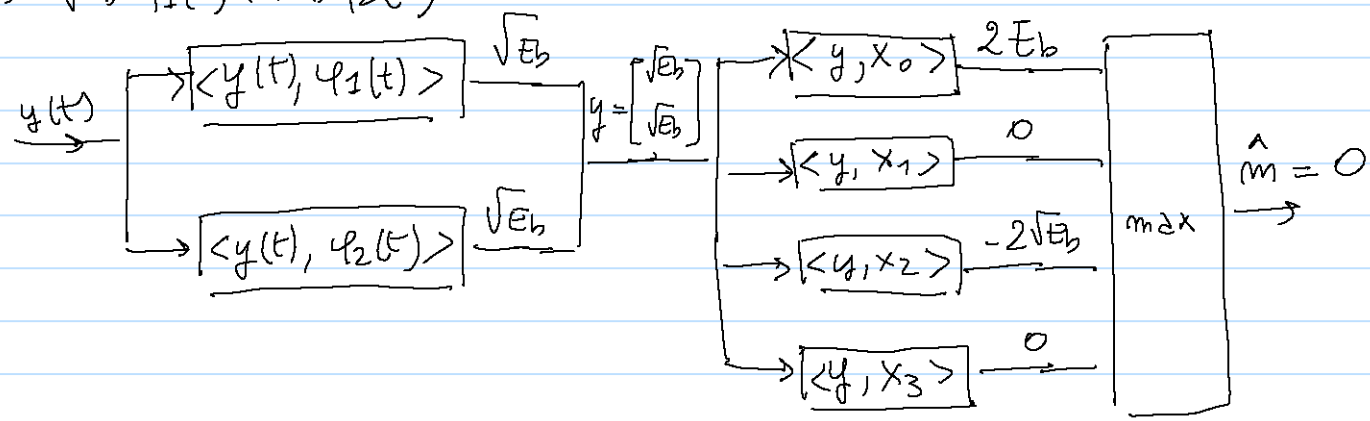
modulator:

$$\varphi_1(t) = \sqrt{\frac{2}{T}}\,\mathrm{sinc}\left(\frac{t}{T}\right)\cos(2\pi f_c t)$$

$$\varphi_2(t) = -\sqrt{\frac{2}{T}}\,\mathrm{sinc}\left(\frac{t}{T}\right)\sin(2\pi f_c t)$$

receiver:

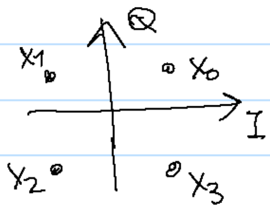
$$y(t) = \sqrt{E_b}\,\varphi_1(t) + \sqrt{E_b}\,\varphi_2(t)$$



How to Demodulate and Decode?

2) Using baseband equivalent transmitter and receiver

encoder:



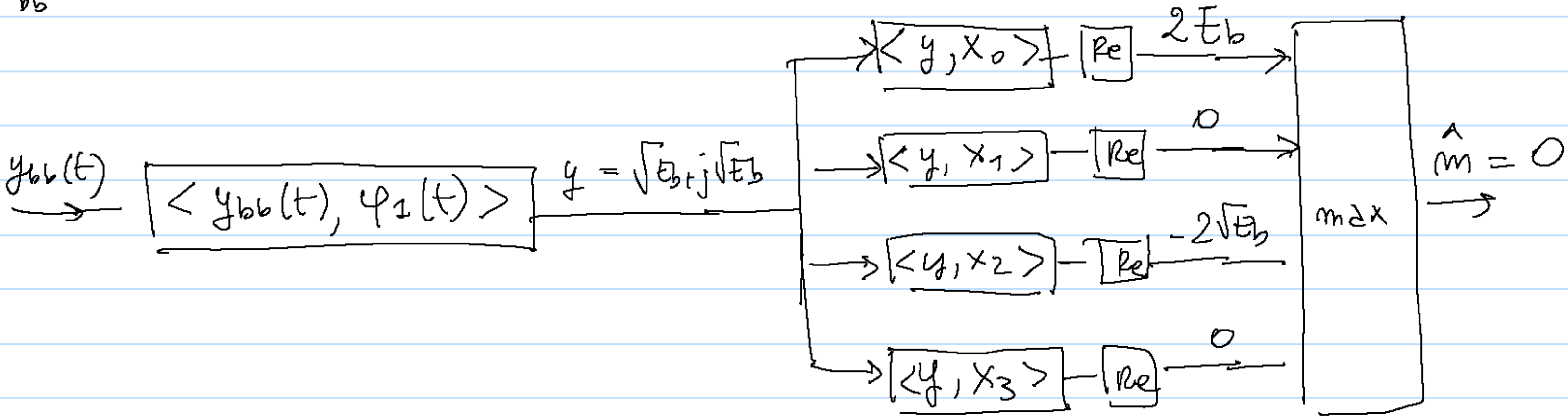
(N=1)

modulator:

$$\varphi_1(t) = \frac{1}{\sqrt{T}}\,\mathrm{sinc}\left(\frac{t}{T}\right)$$

receiver:

$$y_{bb}(t) = (\sqrt{E_b} + j\sqrt{E_b})\,\varphi_1(t)$$



How to Demodulate and Decode?

for instance: $\langle y, x_0 \rangle = (\sqrt{E_b} + j\sqrt{E_b})(\sqrt{E_b} - j\sqrt{E_b})$

$= (E_b + E_b) + j(E_b - E_b) = 2E_b$

$\langle y, x_1 \rangle = (\sqrt{E_b} + j\sqrt{E_b})(-\sqrt{E_b} - j\sqrt{E_b})$

$= (-E_b + E_b) + j(-E_b - E_b) = -2jE_b$

How to Demodulate and Decode?

- **Problem:** Consider 8-PSK and that all symbols have the same probability. If Eb=1 and N0=0.1, what is the optimal decision if the received signal after demodulation is y=3+j0.01?

Matlab: Implementing a MAP Correlative Detector

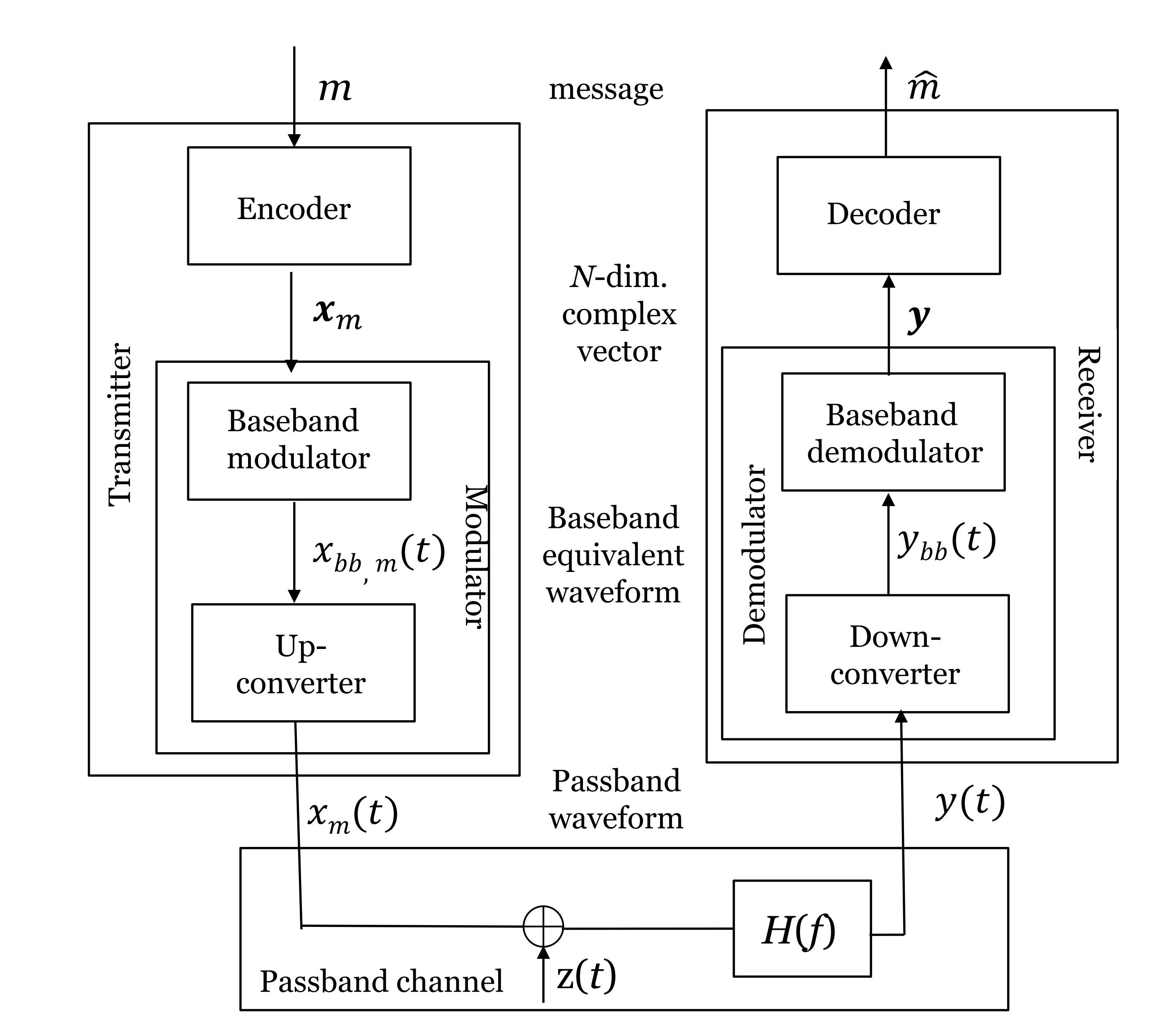
8-PSK transmission (equal probability)

```
%parameters
Eb=1;
N0=0.1;%noise variance – try changing this parameter!
L=1000; %number of symbols

%simulation
m=randi(8,L,1)-1; %generate independent symbols
x= sqrt(3*Eb)*cos(pi*(2*m+1)/8)+j* sqrt(3*Eb)*sin(pi*(2*m+1)/8);
%generate signal vector
plot(real(x),imag(x),'o'); %plot transmitted constellation points
z=randn(L,1)*sqrt(N0/2)+j*randn(L,1)*sqrt(N0/2); %generate noise
y=x+z; %received signal
hold on; plot(real(y),imag(y), 'x');
```

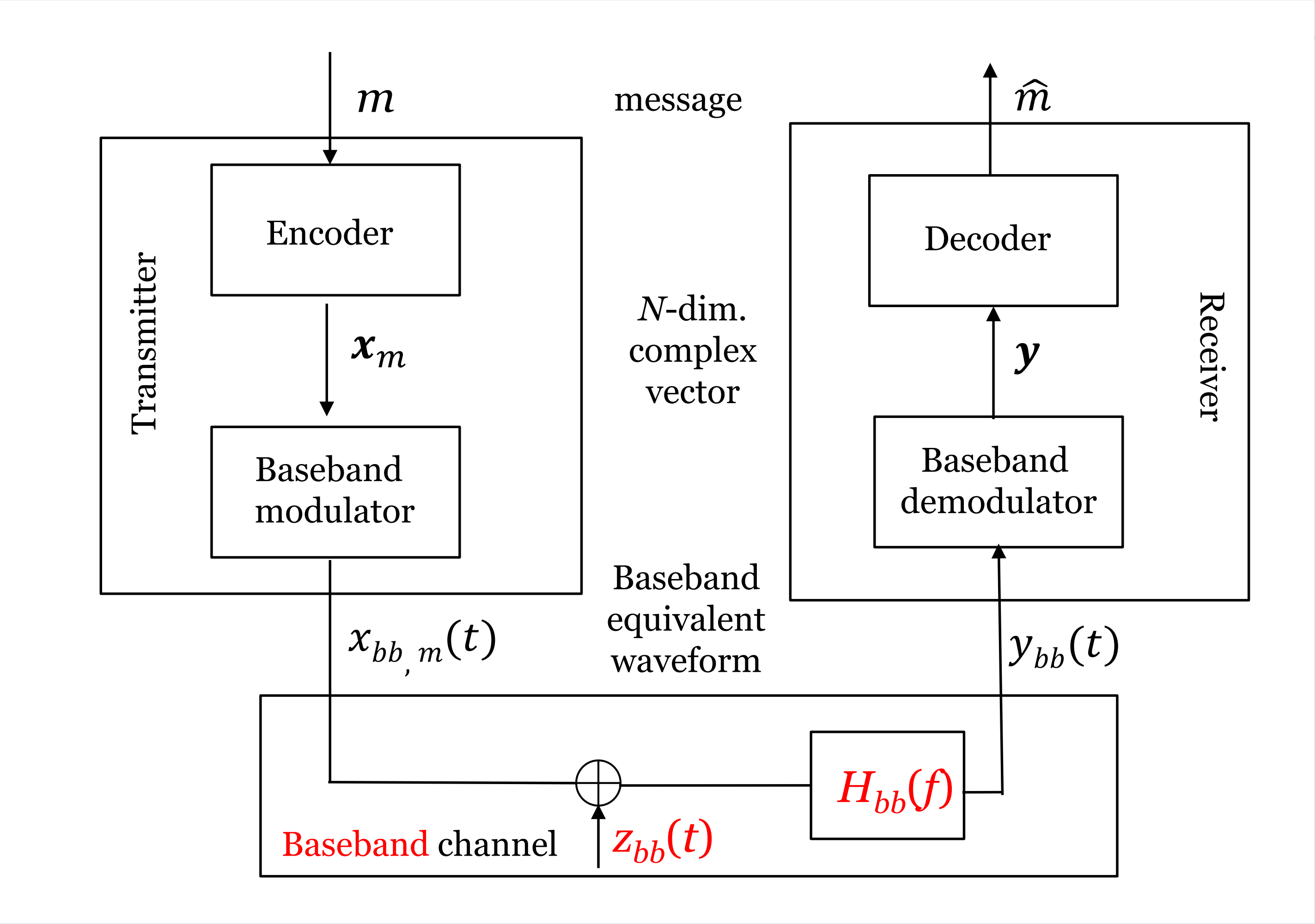
Matlab: Implementing a MAP Correlative Detector

```
Xmat=sqrt(3*Eb)*cos(pi*(2*[0:7]+1)/8)'+j*sqrt(3*Eb)*sin(pi*(2*[0:7]+
1)/8)';

for l=1:L %for each transmitted symbol
score=real(Xmat*y(l,:)'); %compute the score
[smax,imax]=max(score); %find the maximum score
mhat(l)=imax-1;
end

error_rate=sum(m~=mhat')/L
```

Can We Represent the Entire System with a Baseband Equivalent?



Can We Represent the Entire System with a Baseband Equivalent?



- To design the baseband modulator and demodulator, it is useful to operate directly with a baseband equivalent channel model.

- What is the equivalent baseband noise $z_{bb}(t)$? What is the equivalent baseband filter $H_{bb}(f)$?

Why Would We Want to Use An Equivalent Baseband System Again?

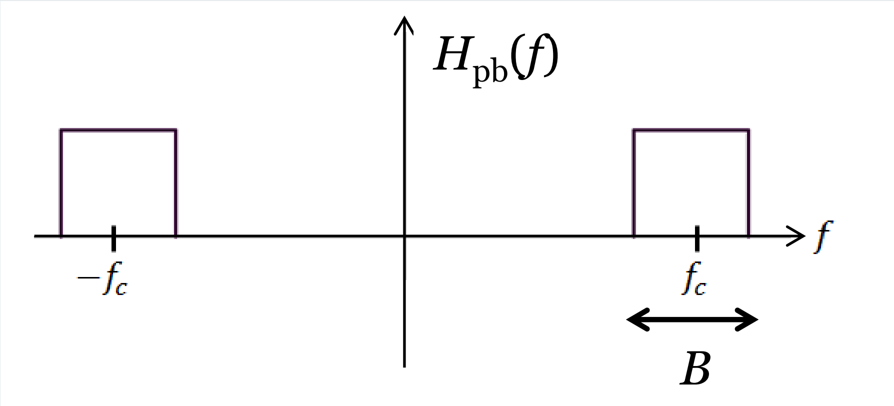
- As discussed, the main processing steps take place at the baseband modulator and demodulator. Therefore, for a communication engineer focusing on signal processing, up- and down-converters need not be included in the design.

- The baseband equivalent is carrier frequency-independent.

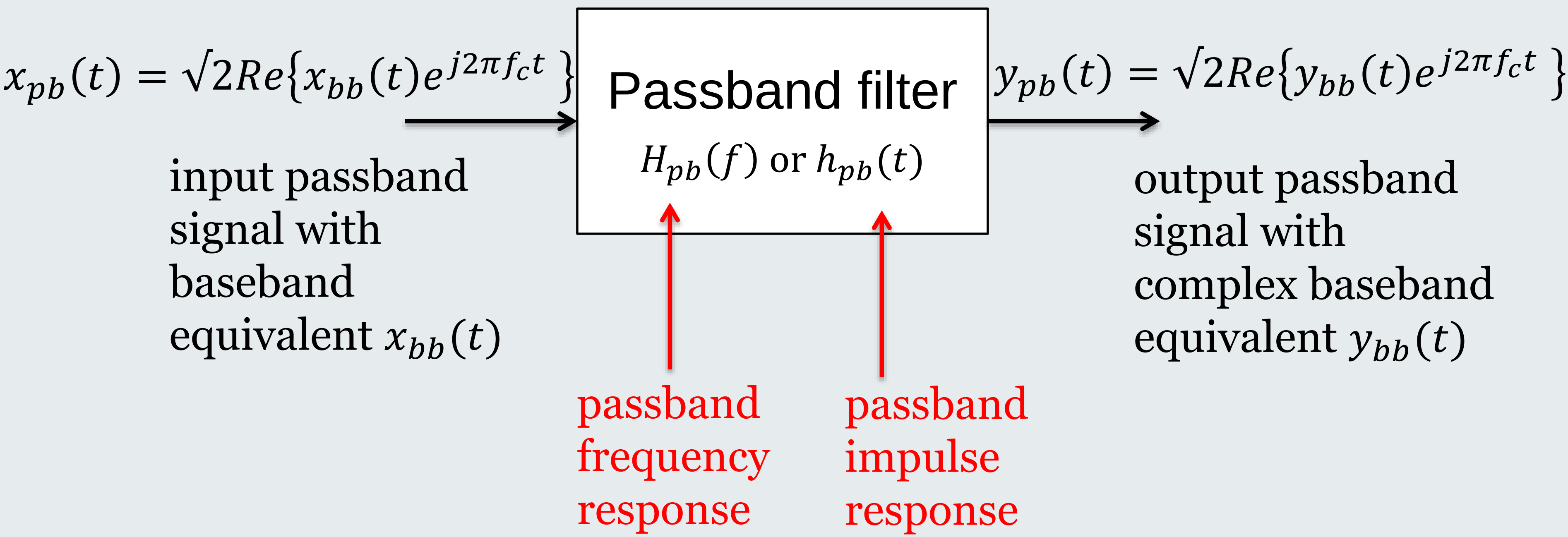
- (In practice, the channel model may depend on the carrier frequency.)

How Do We Obtain a Baseband Equivalent Filter?

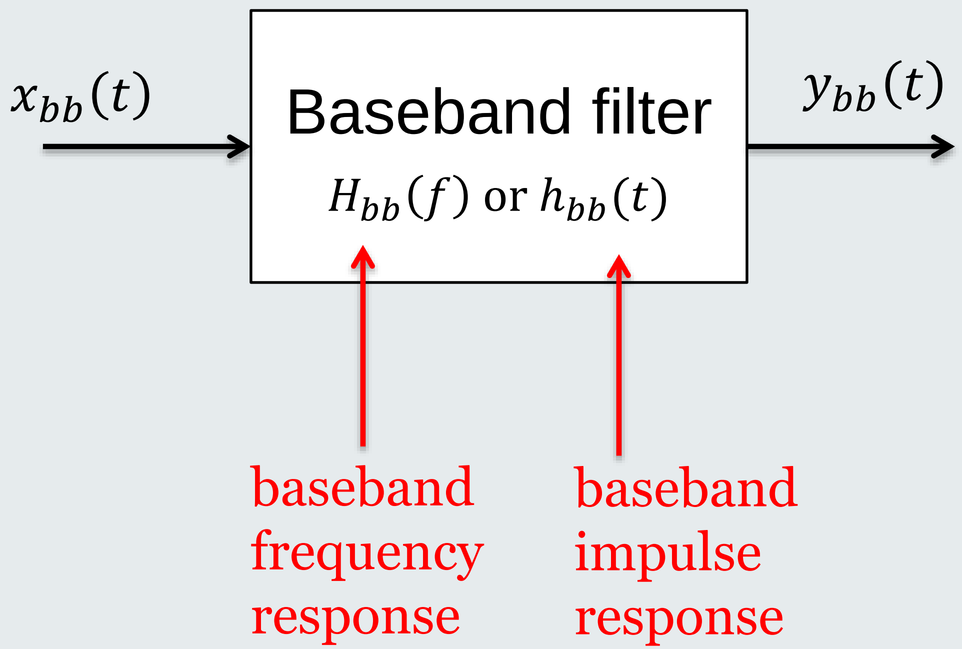
- A passband filter has a non-zero frequency response only around the carrier frequency f_c.
- Example:



How Do We Obtain a Baseband Equivalent Filter?



How Do We Obtain a Baseband Equivalent Filter?



How Do We Obtain a Baseband Equivalent Filter?

- How to choose the baseband filter so that we have the equivalence at the previous slide?

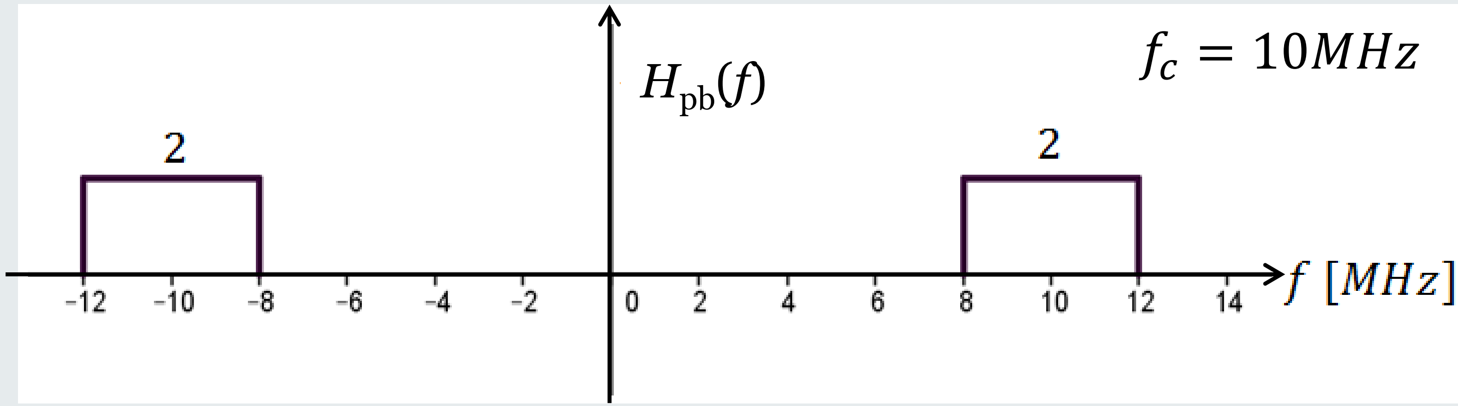
- It can be easily seen that we need

$$H_{pb}(f) = H_{bb}(f - f_c) + H_{bb}^*(-f - f_c)$$

- Remark: Unlike for signals, there is no $\frac{1}{\sqrt{2}}$ term.

How Do We Obtain a Baseband Equivalent Filter?

Consider the bandpass filter in the figure below.



a) If the input signal is

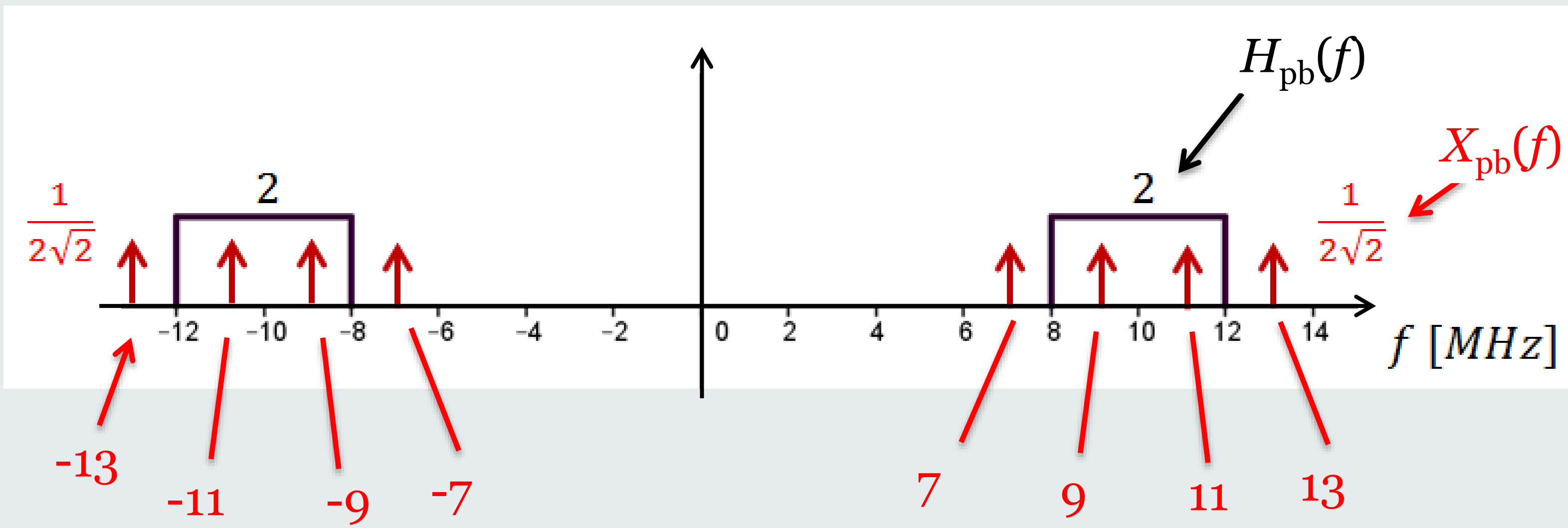
$$x_{pb}(t) = \sqrt{2}\ x_I(t)\cos(2\pi 10^7 t)$$

with $x_I(t) = \cos(2\pi 10^6 t) + \cos(6\pi 10^6 t)$

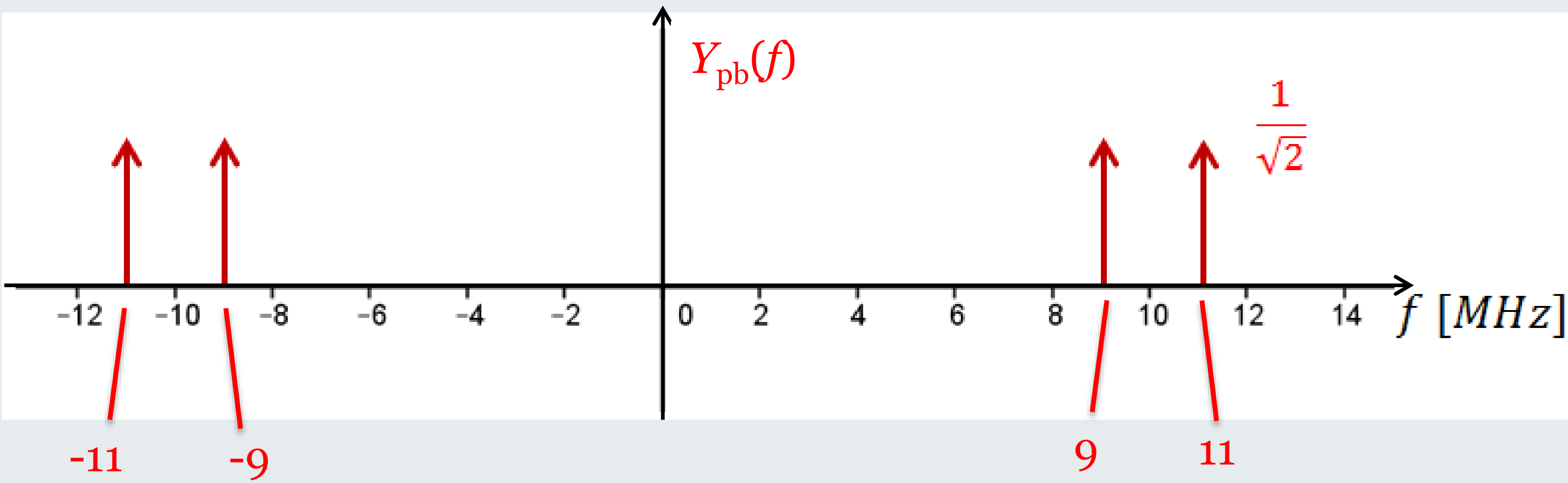
find the output $y_{pb}(t)$.

How Do We Obtain a Baseband Equivalent Filter?

Input and filter:



How Do We Obtain a Baseband Equivalent Filter?



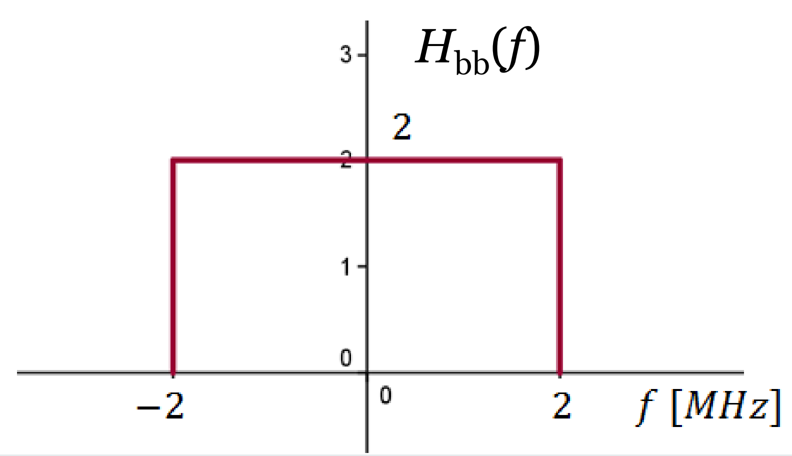
$$\Rightarrow y_{pb}(t) = \sqrt{2}\ \underbrace{(2\cos(2\pi 10^6 t))}_{y_I(t) = y_{bb}(t)} \cos(2\pi 10^7 t)$$

How Do We Obtain a Baseband Equivalent Filter?

b) Calculate the complex baseband equivalent of the output $y_{pb}(t)$:

$$y_{bb}(t) = 2\cos(2\pi 10^6 t)$$

c) Calculate the equivalent baseband filter $H_{bb}(f)$ and $h_{bb}(t)$:

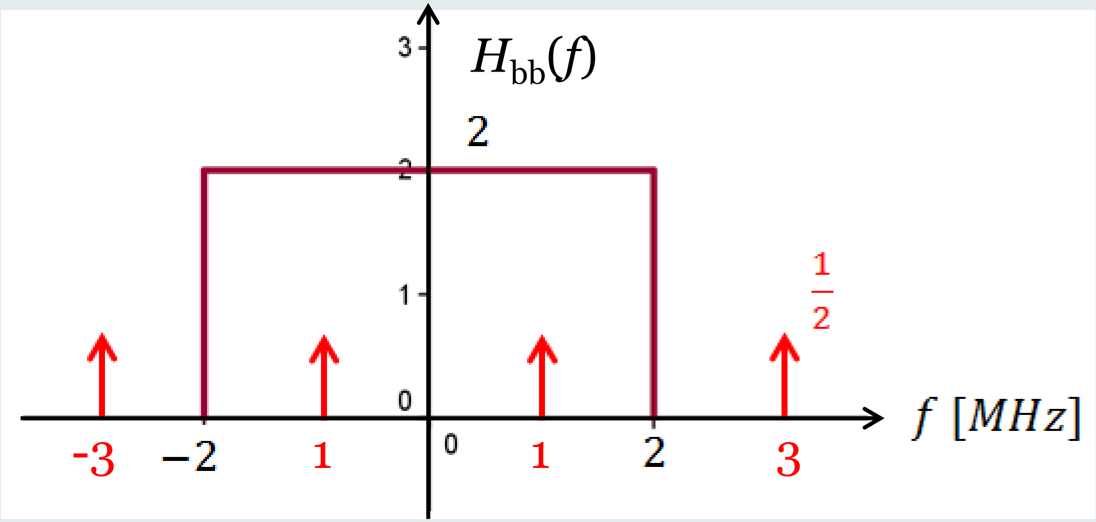


$$\Rightarrow h_{bb}(t) = 8 \times 10^6 \times \text{sinc}(4 \times 10^6 t)$$

How Do We Obtain a Baseband Equivalent Filter?

d) Using the baseband filter, calculate $y_{bb}(t)$. Compare with the results at point b).

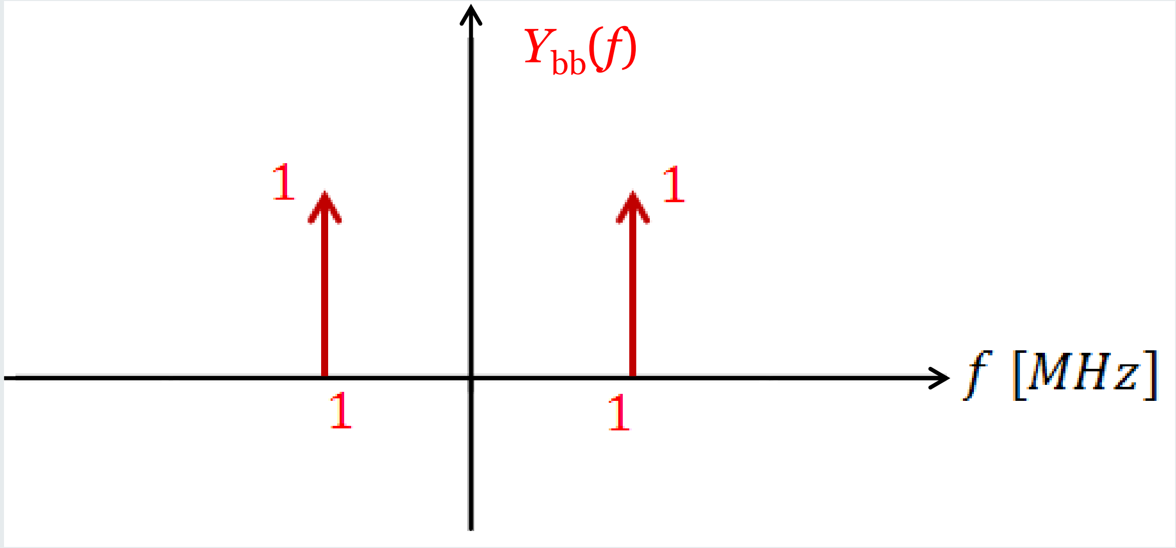
Input and filter:



How Do We Obtain a Baseband Equivalent Filter?

d) Using the baseband filter, calculate $y_{bb}(t)$. Compare with the results at point b).

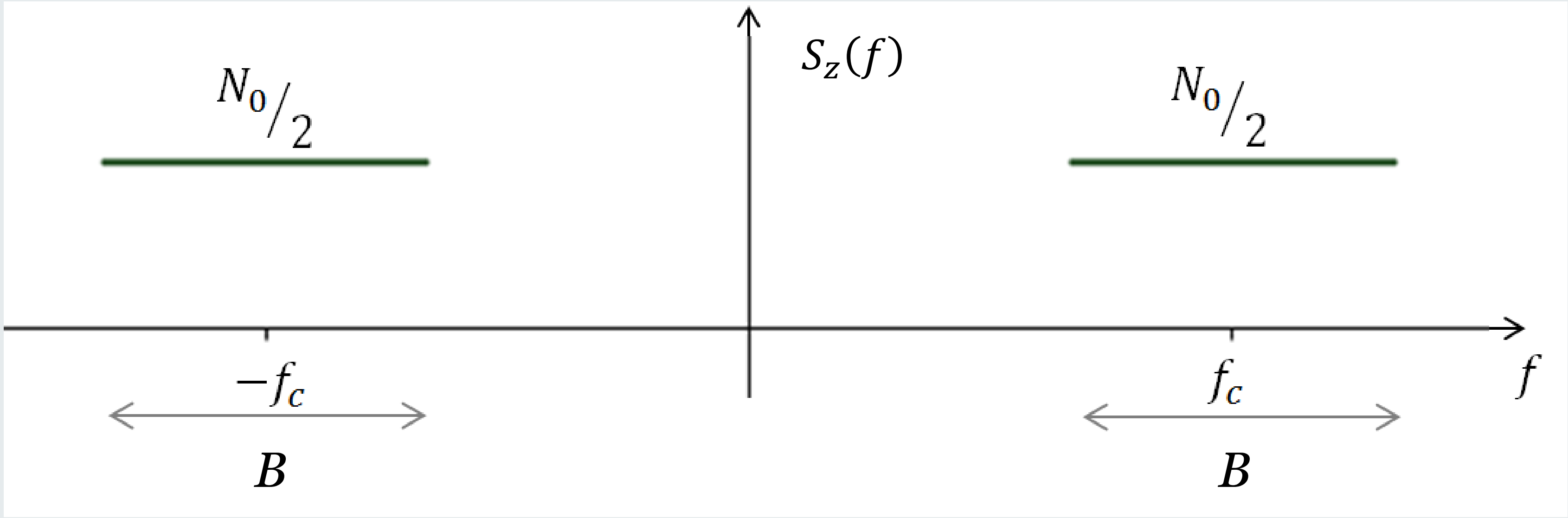
Input and filter:



$\Rightarrow y_{bb}(t) = 2\cos(2\pi 10^6 t)$, as at point b).

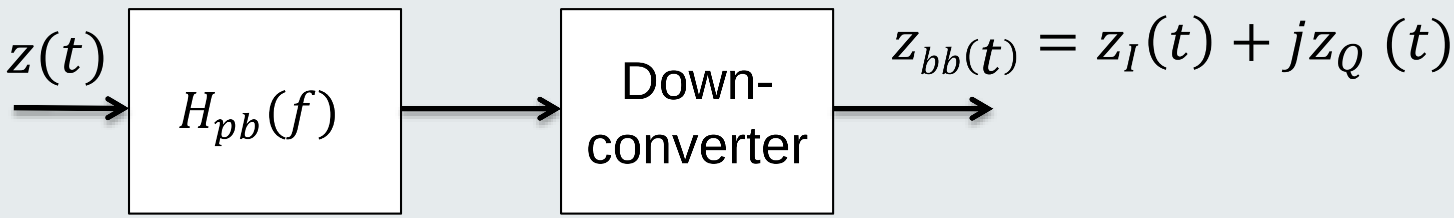
How Do We Obtain the Baseband Equivalent Noise?

- The bandpass noise $z(t)$ is a WGN with power spectral density $S_z(f) = \frac{N_0}{2}$ within the bandwidth of the channel



How Do We Obtain the Baseband Equivalent Noise?

- The baseband noise is the output of the following blocks:

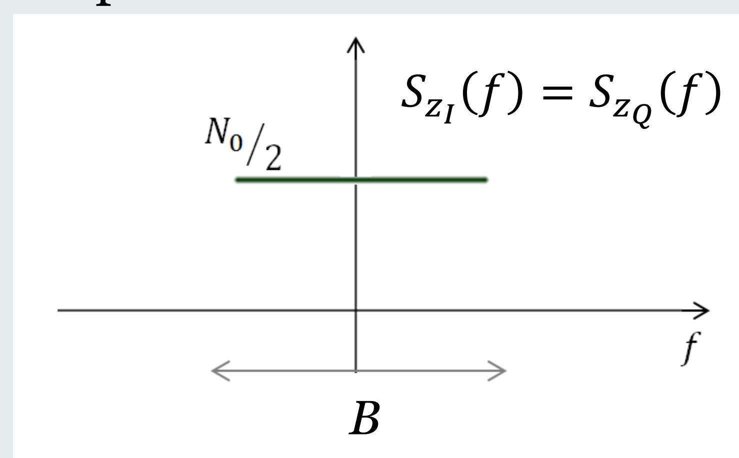


- Assuming that $H_{pb}(f)$ is an ideal passband filter, it can be proved that

a) $z_I(t)$ and $z_Q(t)$ are WGN with power spectral densities

$$S_{z_I}(f) = S_{z_Q}(f) = \frac{N_0}{2}$$

within the bandwidth $\left(-\frac{B}{2}, \frac{B}{2}\right)$



b) $z_I(t)$ and $z_Q(t)$ are independent

How Do We Obtain the Baseband Equivalent Noise?

Remark:

$$E[z(t)^2] = \frac{N_0}{2} \cdot 2B \quad = N_0 B$$

$$\left.\begin{aligned} E[z_I(t)^2] &= \frac{N_0}{2} B \\ E[z_Q(t)^2] &= \frac{N_0}{2} B \end{aligned}\right\} \qquad \begin{aligned} &E[|z_{bb}(t)|^2] \\ &= E[z_I(t)^2] + E[z_Q(t)^2] \\ &= N_0 B \end{aligned}$$

The power of the bandpass noise $E[z(t)^2]$ is hence equal to the power of the baseband signal $E[|z_{bb}(t)|^2]$.

This relationship is akin to that between the energy for bandpass and baseband signals.

Matlab: Generating Noise

- Generating a baseband WGN $W_z(t)$ with power $P_{W_z}(= N_0 B_R)$ in MATLAB

$W_I = sqrt(P_{W_z}/2)^*$randn (N,1);

% generates N samples of $W_I(t)$

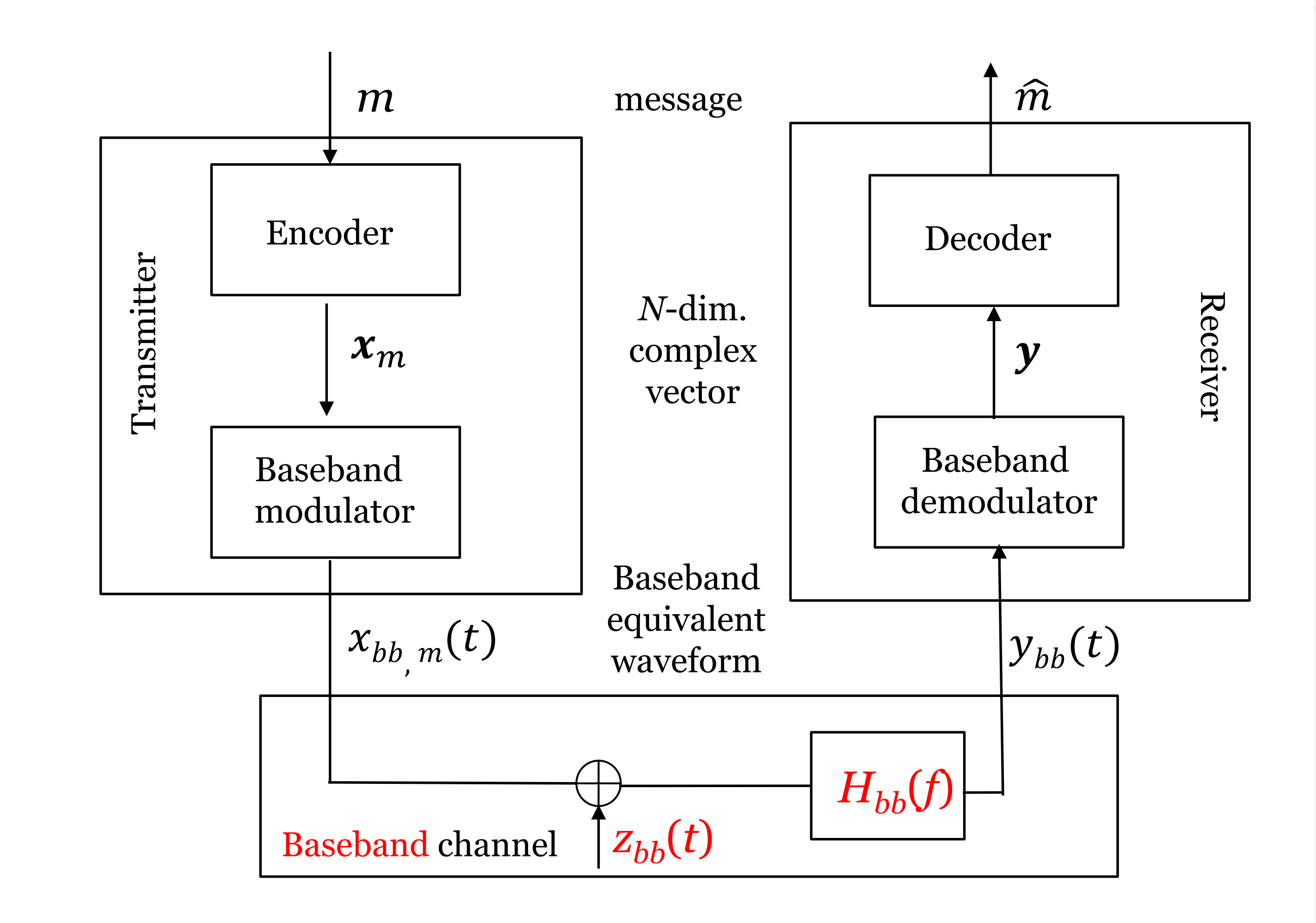
$W_Q = sqrt(P_{W_z}/2)^*$randn (N,1);

% generates N samples of $W_Q(t)$

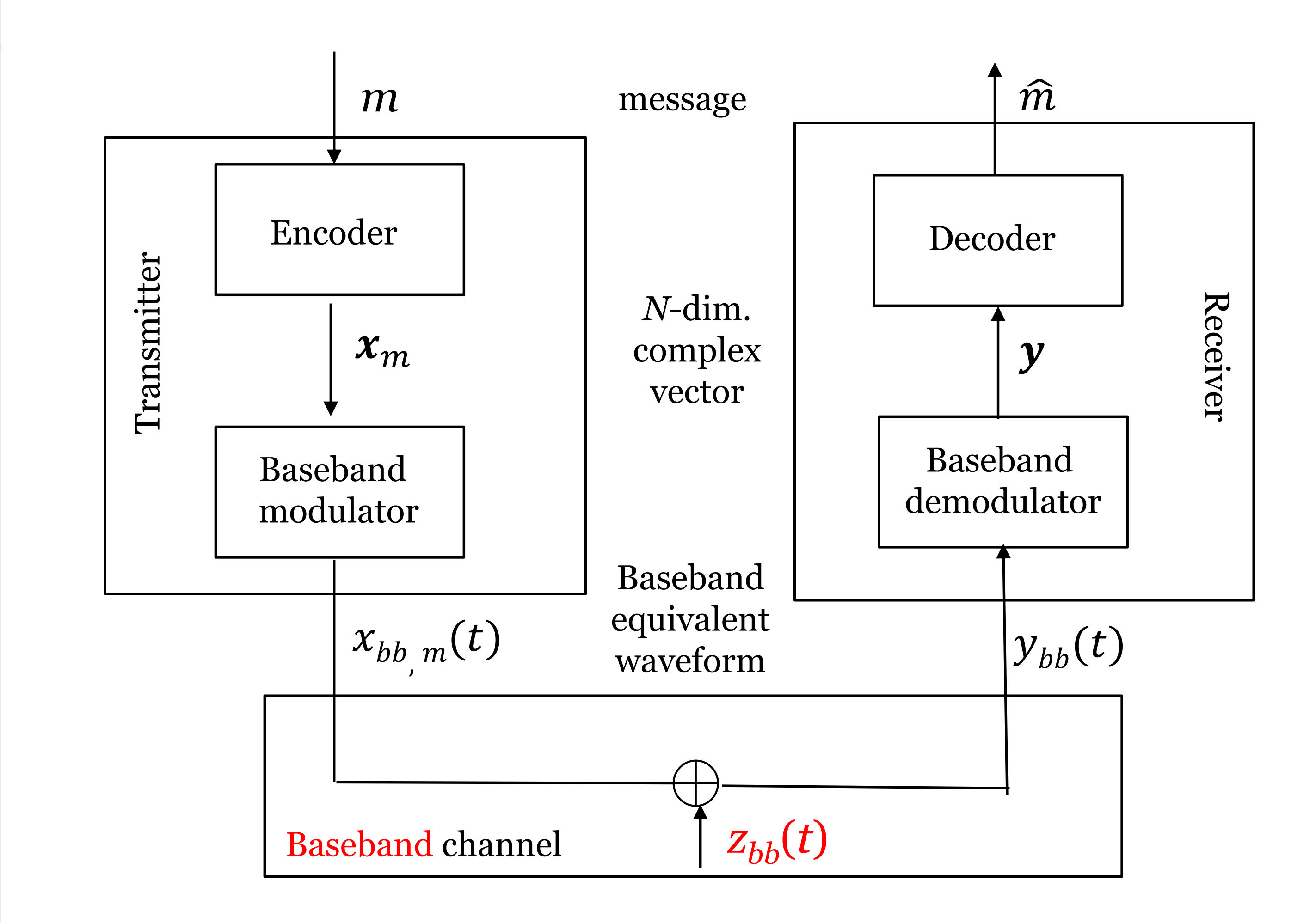
$W_z = W_I + j * W_Q$;

% generates N samples of $W_z(t)$

Can We Represent the Entire System with a Baseband Equivalent?



Can We Represent the Entire System with a Baseband Equivalent?



- If the signals are bandlimited within the bandwidth B and the filter is ideal in this bandwidth, we can just write

$$y_{bb}(t) = x_{bb,m}(t) + z_{\text{bb}}(t)$$

where $z_{\text{bb}}(t)$ is AWGN

Can We Represent the Entire System with a Baseband Equivalent?

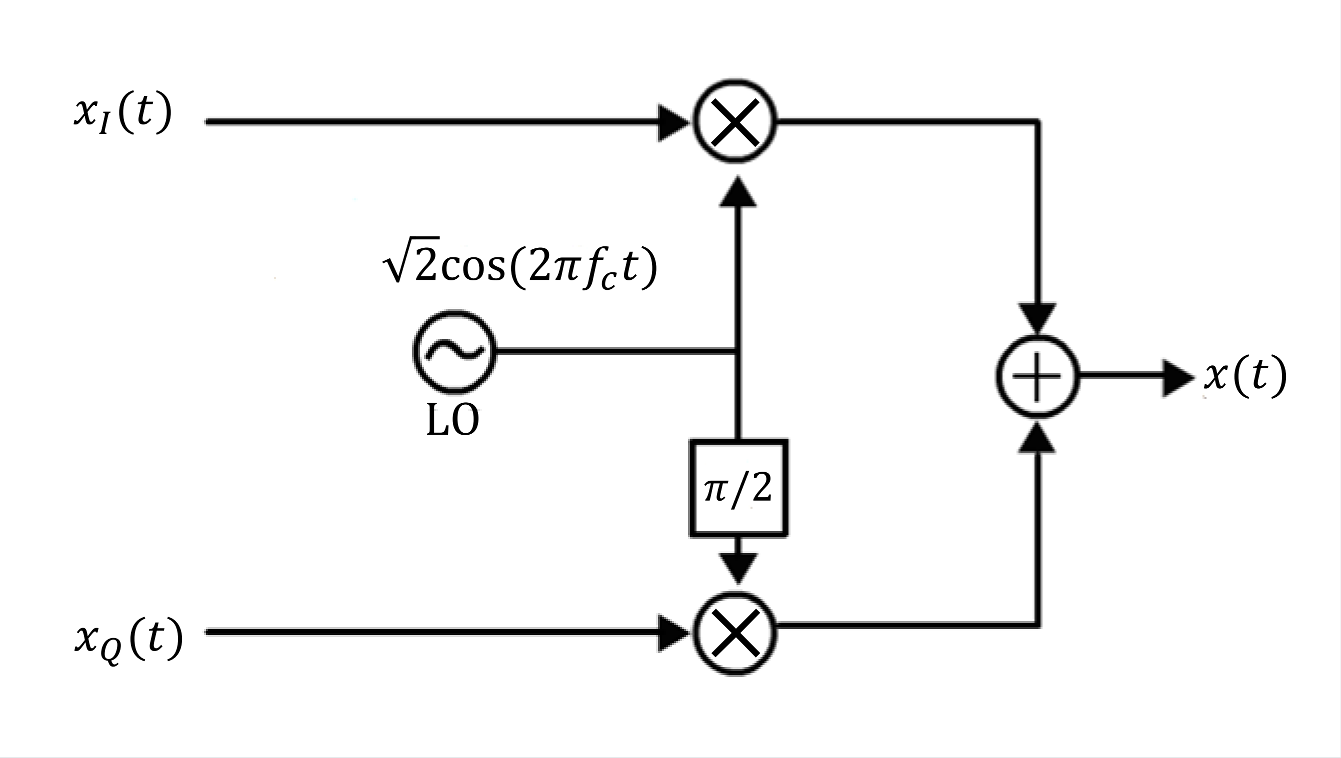
- From the viewpoint of modulator and demodulator, we have then obtained an AWGN channel as studied in the previous chapter with the only caveat that symbols and functions can be complex.

- From the viewpoint of encoder and decoder, the channel is additive Gaussian as studied in the previous chapter with one complex number representing two real dimensions.

- The optimal receiver is hence the one discussed in the previous slides.

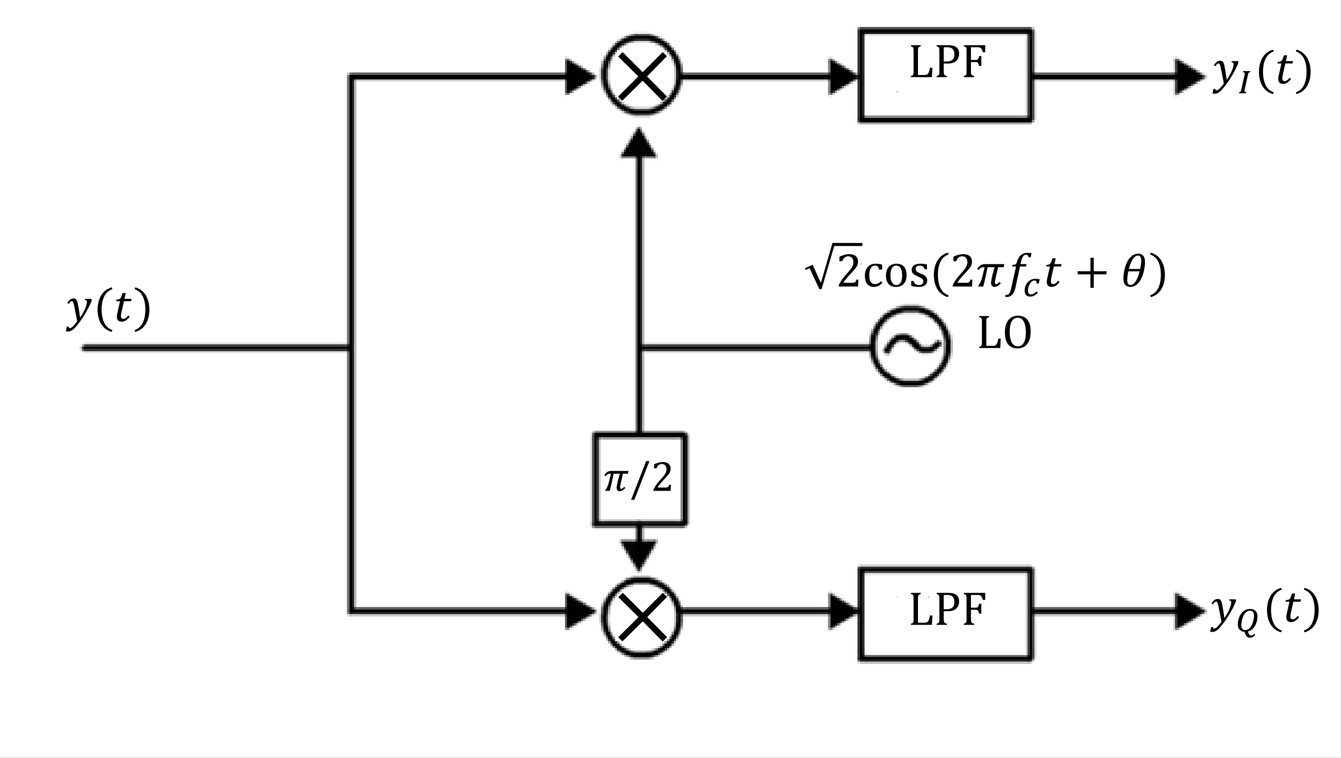
- The probability of error can be computed as seen in the previous chapter.

What is the Impact of Lack of Synchronization?



LO = Local Oscillator

What is the Impact of Lack of Synchronization?



Phase asynchronism with phase offset θ

What is the Impact of Lack of Synchronization?

- Consider the noiseless case $y(t)$=$x(t)$

$$y(t)\sqrt{2}\cos(2\pi f_c t + \theta) = 2x_I(t)\cos(2\pi f_c t)\cos(2\pi f_c t + \theta)$$

$$- 2\, x_Q(t)\sin(2\pi f_c t)\cos(2\pi f_c t + \theta)$$

$$= x_I\,(t)\cos(\theta) + x_I(t)\cos(4\pi f_c t + \theta) - x_Q(t)\sin(4\pi f_c t + \theta)$$
$$+x_Q(t)\sin(\theta)$$
$$= x_I\,(t)\cos(\theta) + x_Q(t)\sin(\theta) \text{ after LPF}$$
$$= y_I\,(t)$$

And similarly for $y(t)\left(-\sqrt{2}\sin(2\pi f_c t)\right)$:

$$y_Q(t) = x_Q(t)\cos(\theta) - x_I(t)\sin(\theta)$$

$$\cos a \cos b = 1/2\,(\cos(a-b) + \cos(a+b))$$
$$\cos a \sin b = 1/2\,(\sin(a+b) - \sin(a-b))$$

What is the Impact of Lack of Synchronization?

- Lack of phase synchronization hence causes interference between I and Q components:

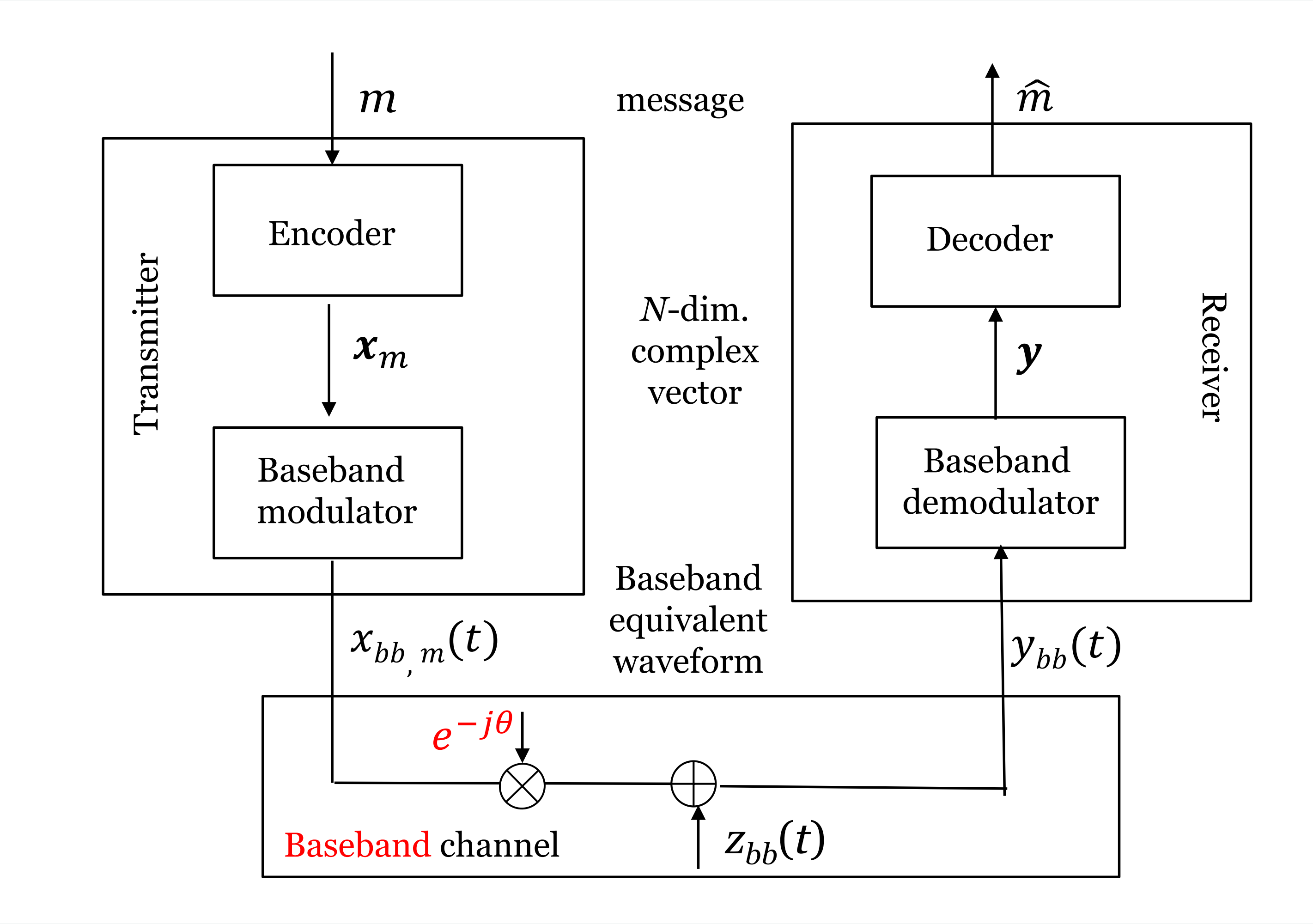
$$y_I(t) = x_I(t)\cos(\theta) + x_Q(t)\sin(\theta)$$
$$y_Q(t) = -x_I(t)\sin(\theta) + x_Q(t)\cos(\theta)$$

- The previous relationship can also be written as

$$y_{bb}(t) = x_{bb}(t)e^{-j\theta}$$

What is the Impact of Lack of Synchronization?

- Equivalent baseband system



What is the Impact of Lack of Synchronization?

- How to detect a lack of phase synchronization?

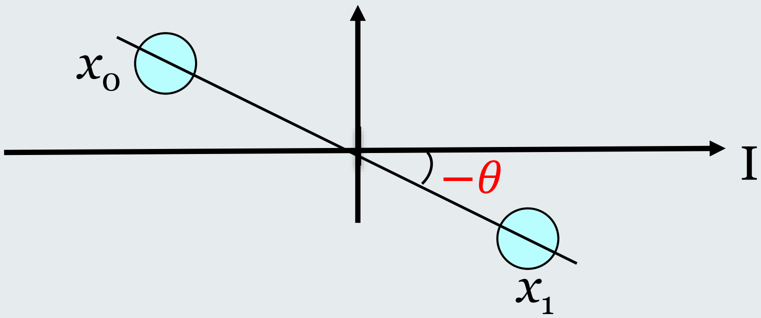
- Consider BPSK. We have

$$y_{bb}(t) = x_m \varphi_1(t) e^{-j\theta}$$

- Hence, after demodulation we obtain

$$y = x_m e^{-j\theta}$$

- This implies that the received constellation is rotated.



What is the Impact of Lack of Synchronization?

- As a result, a decoder applying the decision regions of BPSK would fail completely if the phase offset is 90 degrees.
- A coherent decoder tries to estimate the phase offset and then compensates it before performing decoding.
- Estimation of the phase offset based on the reception of the data only has an irreducible ambiguity of 180 degrees.
- To resolve this ambiguity, we need to transmit pilot symbols prior to the data. Pilot symbols are known to the receiver in advance and do not carry any information.

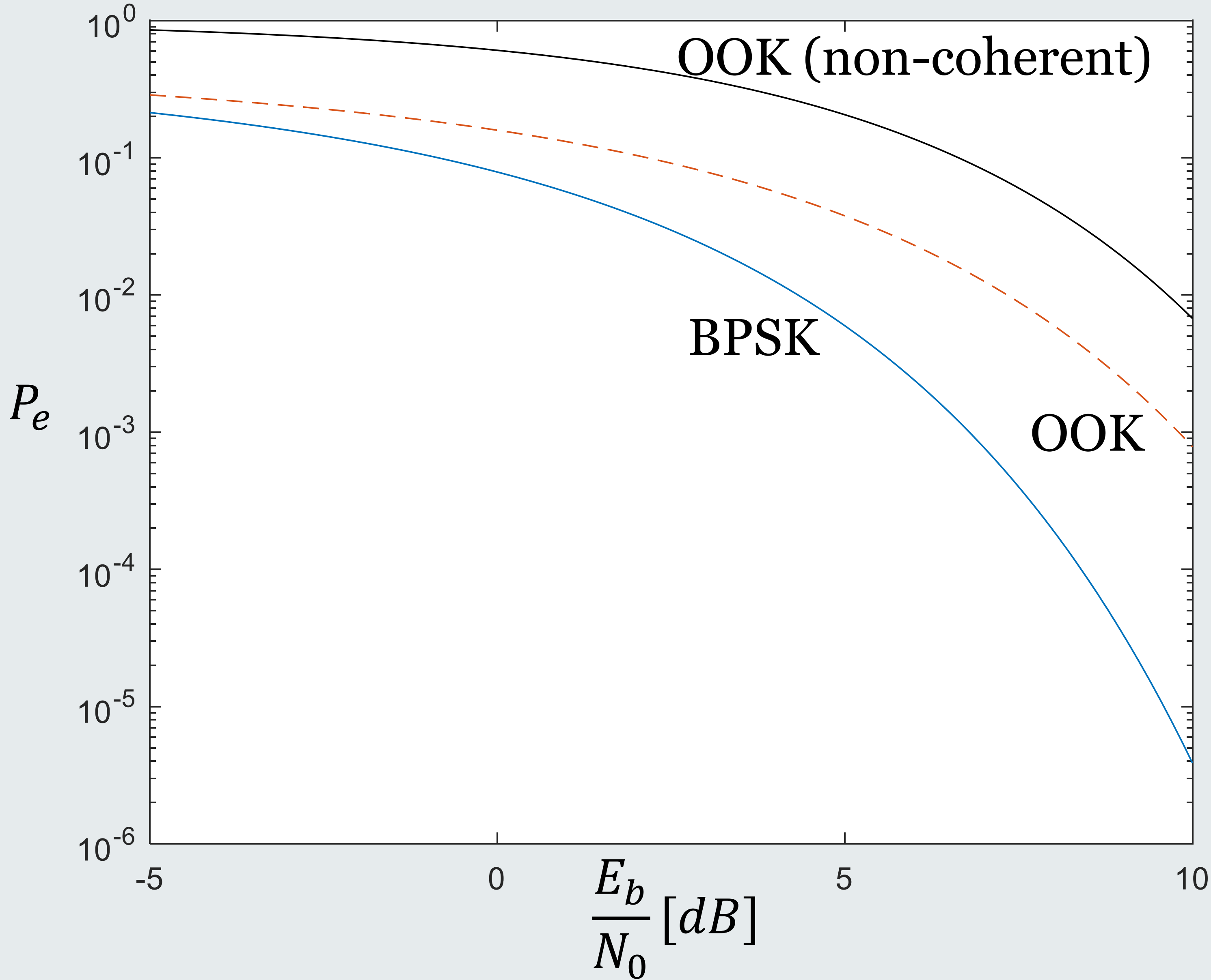
What is the Impact of Lack of Synchronization?

- The same applies to other constellations, but the ambiguity is more pronounced.

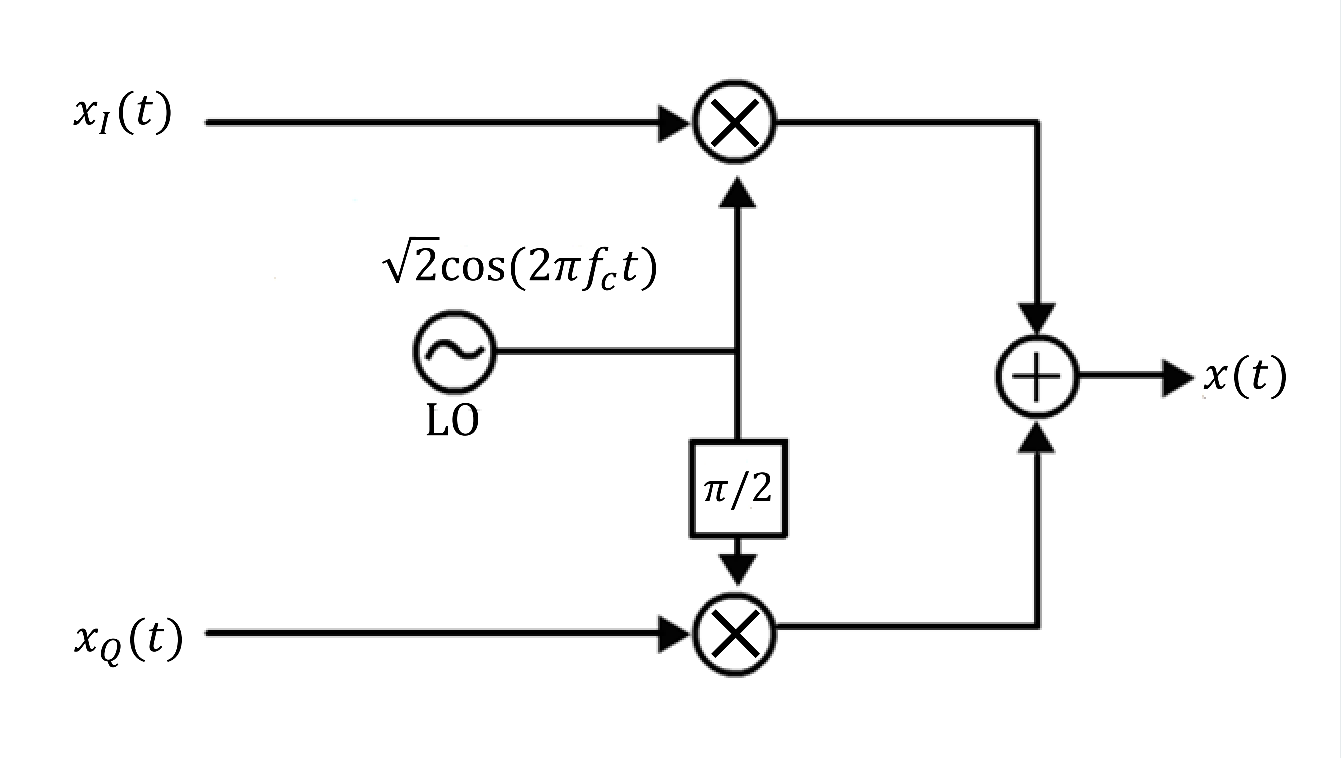
- **Problem:** What is the maximum phase offset detectable without training symbols for 4-PSK? How about 8-PSK?

- **Problem (Non-coherent transmission):** What is the impact of a phase offset on on-off modulation? Can you design a coherent decoder with no pilots? Can you design a non-coherent decoder that works irrespective of the phase offset for equally likely messages?

What is the Impact of Lack of Synchronization?

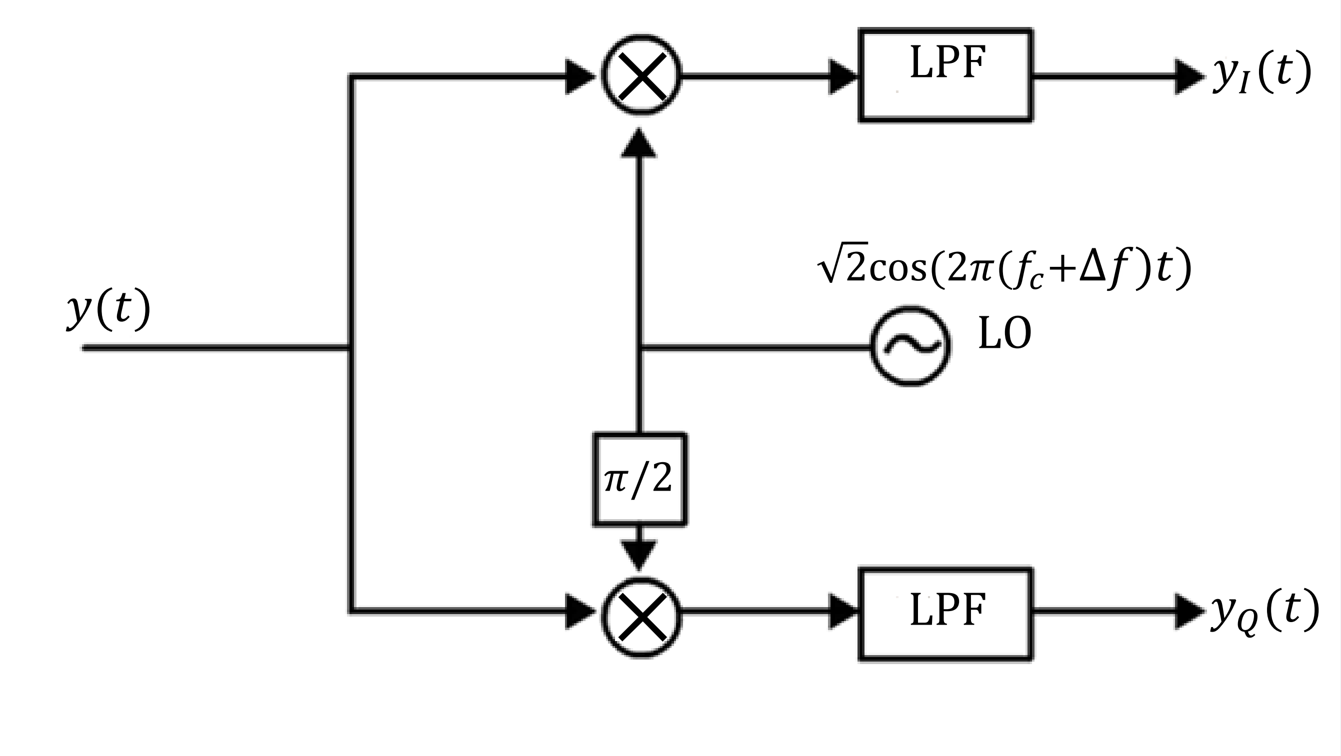


What is the Impact of Lack of Synchronization?



LO = Local Oscillator

What is the Impact of Lack of Synchronization?



Frequency asynchronism with frequency offset Δf

What is the Impact of Lack of Synchronization?

- Consider the noiseless case $y(t)=x(t)$

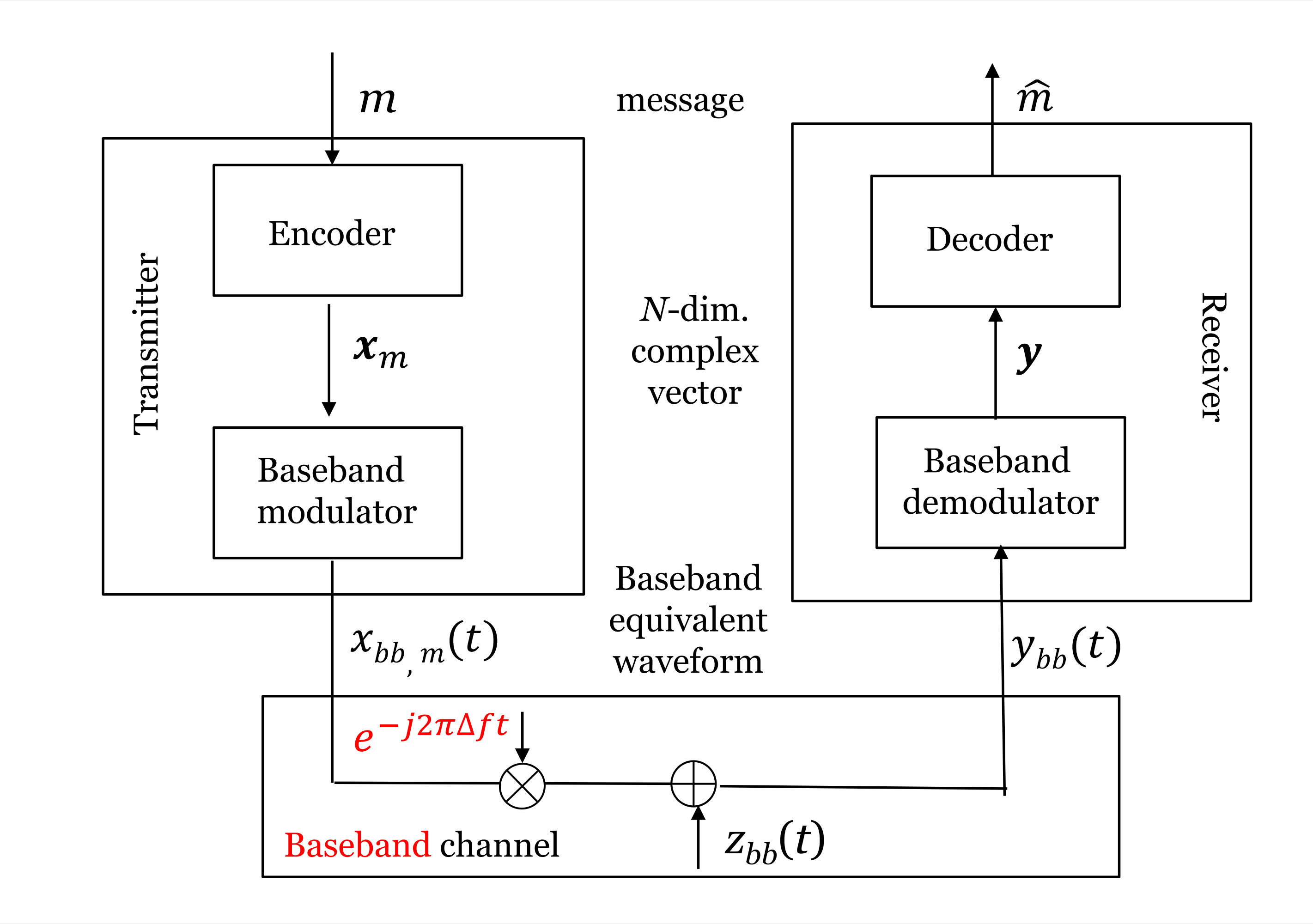
$$y(t)\sqrt{2}\cos(2\pi(f_c+\Delta f)t)$$
$$= x_I(t)\cos(2\pi\Delta f t) + x_Q(t)\sin(2\pi\Delta f t) \text{ after LPF}$$

And similarly for $y(t)\left(-\sqrt{2}\sin(2\pi(f_c+\Delta f)t)\right) =$

$$= x_Q(t)\cos(2\pi\Delta f t) - x_I(t)\sin(2\pi\Delta f t)$$

What is the Impact of Lack of Synchronization?

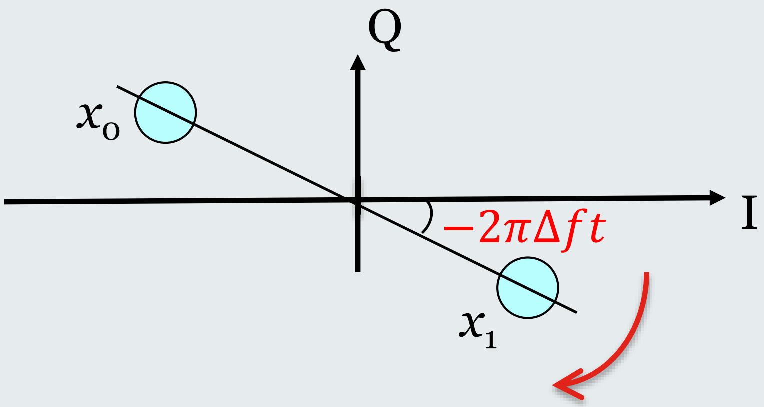
- Equivalent baseband system



What is the Impact of Lack of Synchronization?

- How to detect a lack of frequency synchronization?

- Consider BPSK – After downconversion and demodulation, the constellation looks as follows



- The constellation rotates at a frequency equal to the frequency offset. The frequency offset can hence be estimated in the frequency domain by observing received training symbols.

KING’S
College
LONDON

6CCS3COS Communication Systems: Chapter 4

Osvaldo Simeone

What is This Course About?

- **Overview**

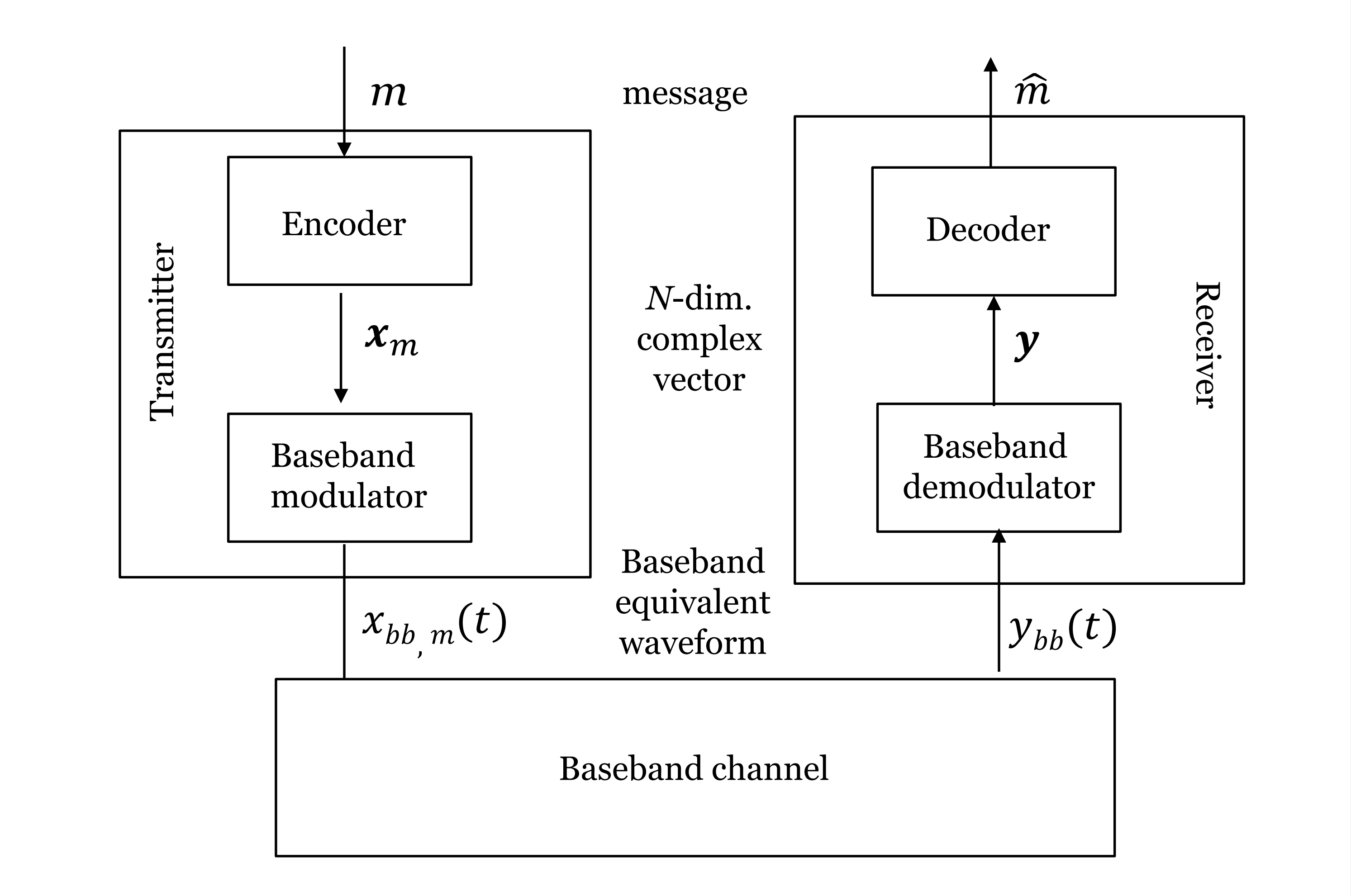
  - 1. One-shot digital communications: Fundamentals
  - 2. One-shot digital communications: Passband Systems
  - 3. Stream digital communications

Main references

- J. Cioffi, Lecture notes, Stanford Univ., Chapters 1, 2, 3

What Have We Learned So Far?

- We have mostly focused so far on "one-shot" transmission with N=1 complex dimension (or N=2 real dimensions).



How to Stream Data?

- Time-domain transmission:

$$\varphi_n(t) = \varphi(t - nT) \text{ for } n\text{=1,…,}N$$

for some unitary-energy waveform $\varphi(t)$

How to Stream Data?

- Time-domain transmission:

$$\varphi_n(t) = \varphi(t - nT) \text{ for } n\text{=1,…,}N$$

for some unitary-energy waveform $\varphi(t)$

- Frequency-domain transmission: Orthogonal Frequency Division Multiplexing (OFDM)

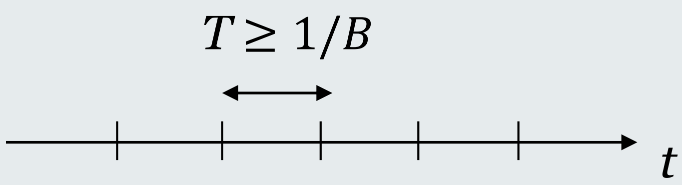
$$\varphi_n(t) = \varphi(t)\exp\left(j2\pi\left(\frac{n}{T} - \frac{N}{2T}\right)t\right) \text{ for } n\text{=1,…,N}$$

for some unitary-energy waveform $\varphi(t)$

- Each symbol carries one complex dimension.

How to Stream Data?

- Time-domain transmission:



- T must be larger than 1/B since each symbol carries one complex dimension.

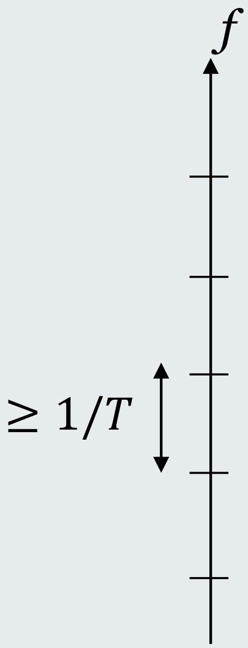
- The choice

$$\varphi(t) = \frac{1}{\sqrt{T}} \operatorname{sinc}(t/T)$$

maximizes the symbol rate, i.e., uses the minimum T for a given B.

How to Stream Data?

- Frequency-domain transmission (OFDM):

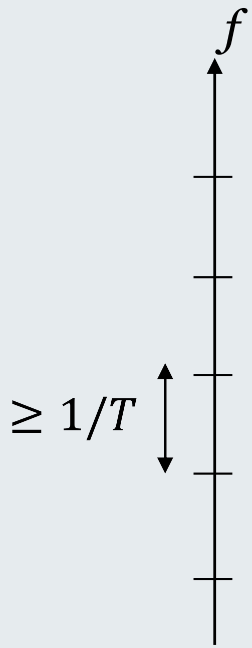


- The spacing between subcarriers must be at least 1/T, since each symbol carries one complex dimension.

- The choice $\varphi(t) = \frac{1}{\sqrt{T}} \mathrm{rect}\left(\frac{t}{T}\right)$ uses the spectrum in the most efficient way, i.e., it minimizes the subcarrier spacing for a given T.

How to Stream Data?

- Frequency-domain transmission (OFDM):

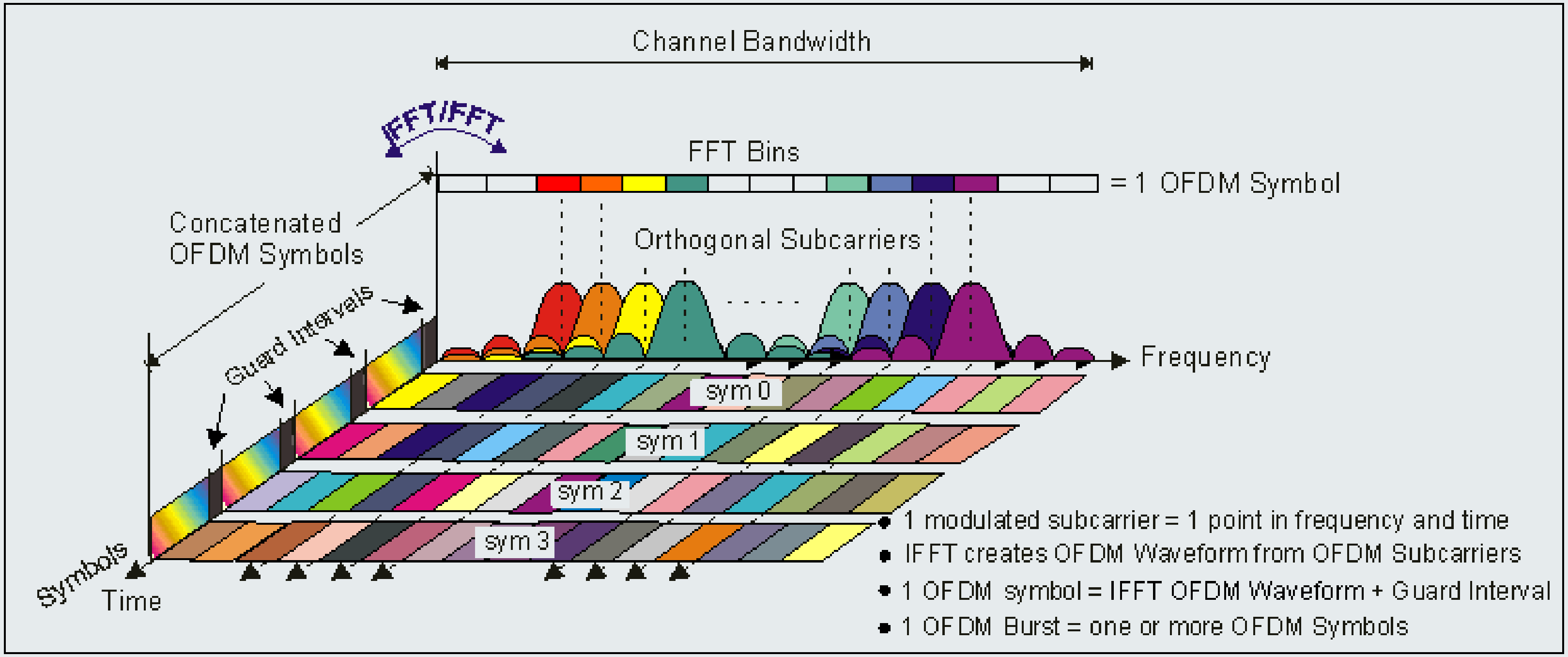


In practice, OFDM symbols, of duration T, are sent back to back with guard periods or cyclic prefixes

- The spacing between subcarriers must be at least 1/T, since each symbol carries one complex dimension.

- The choice $\varphi(t) = \frac{1}{\sqrt{T}} \text{rect}\left(\frac{t}{T}\right)$ uses the spectrum in the most efficient way, i.e., it minimizes the subcarrier spacing for a given T.

How to Stream Data?



Frequency-Time Representative of an OFDM signal

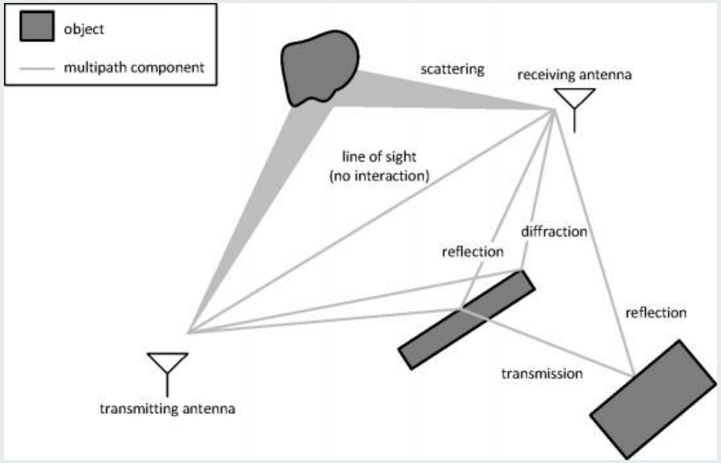
http://rfmw.em.keysight.com/wireless/helpfiles/89600b/webhelp/subsystems/wlan-ofdm/content/ofdm_basicprinciplesoverview.htm

How to Stream Data?

- In this chapter, we will talk about two key aspects of data streaming.

- 1. Effect of channel distortions:

  How does the channel affect the reception of successive symbols in time or frequency domain?



- 2. Coding over many dimensions:

  When streaming, one has available a large number N of dimensions. Can this be used to reduce the probability of error?

What is the Effect of the Channel on Data Streaming?

- Let us start with time domain transmission.
- The design and analysis considered up to now applies only if the successive symbols do not interfere with one another.
- We know that this is the case if the delayed symbols are orthogonal.
- We will see that orthogonality in practice depends also on the channel and not only on the modulator.
- When orthogonality does not hold, the interference between successive transmissions is called **intersymbol interference (ISI)**. ISI can severely complicate the implementation of an optimum detector.

How to Stream Data in Time Domain?

- The message transmissions are separated by T units in time, where T is called the **symbol period**.
- $1/T$ is called the **symbol rate**.
- The **data rate** is

$$R = \frac{\log_2 M}{T} = \frac{b}{T} \text{ (bit/s)}$$

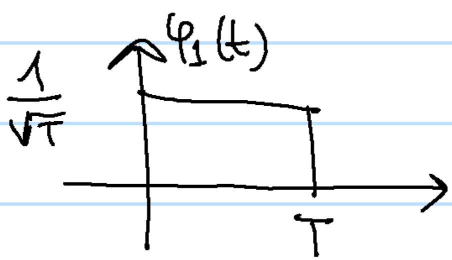
What is the Effect of the Channel on Data Streaming?

Introductory example:

encoder: BPSK $m[1] \rightarrow x[1]$ and $m[2] \rightarrow x[2]$

modulator: $\varphi_1(t) = \frac{1}{\sqrt{T}} \, \mathrm{rec}\left(\frac{t}{T}\right) = \varphi(t)$

$$\varphi_2(t) = \frac{1}{\sqrt{2}} \, \mathrm{rect}\left(\frac{t-T}{T}\right) = \varphi(t-T)$$



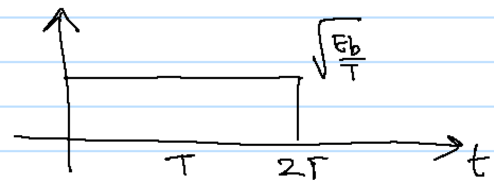
What is the Effect of the Channel on Data Streaming?

Signal set:

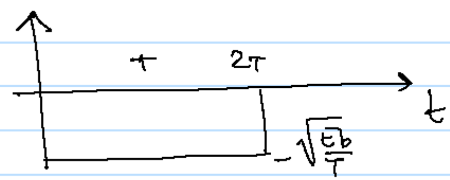
$$x_m(t) = x[1]\,\varphi(t) + x[2]\,\varphi(t-T)$$

$m = 1, 0$

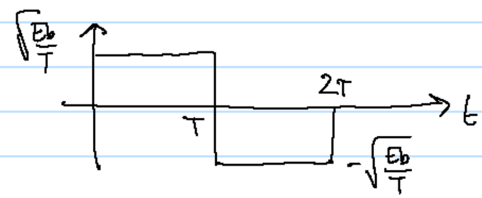
$m[1]$ $m[2]$



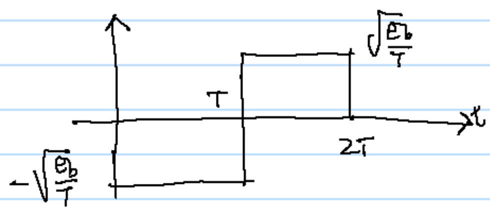
$m = 0, 0$



$m = 1, 0$



$m = 0, 1$

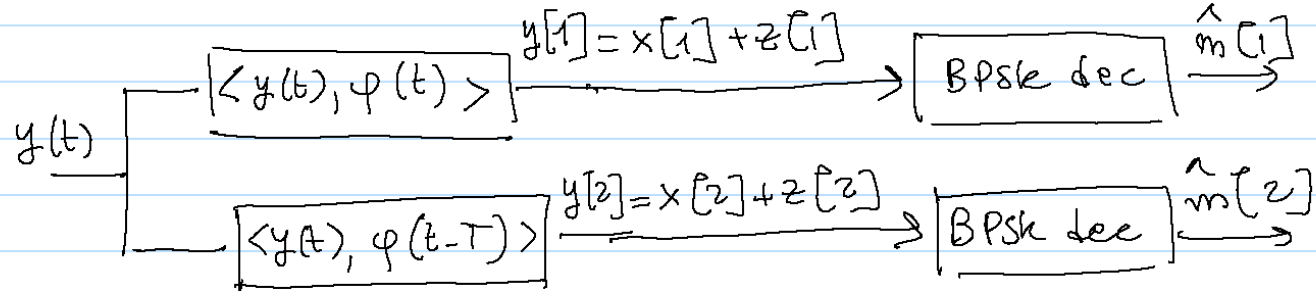


$\Rightarrow$ the first symbol encodes $m[1]$, while the second symbol encodes $m[2]$

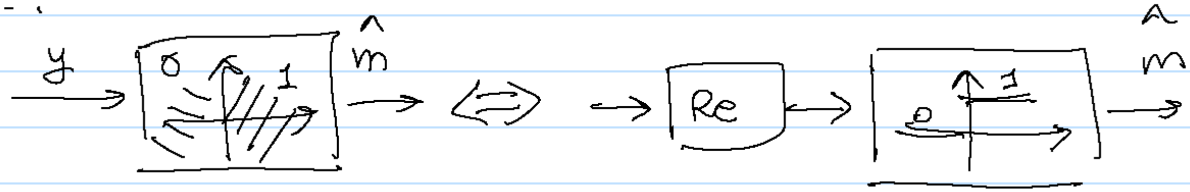
$\Rightarrow$ we have two successive BPSK transmissions

What is the Effect of the Channel on Data Streaming?

The optimal decoder, assuming equally likely symbols, is



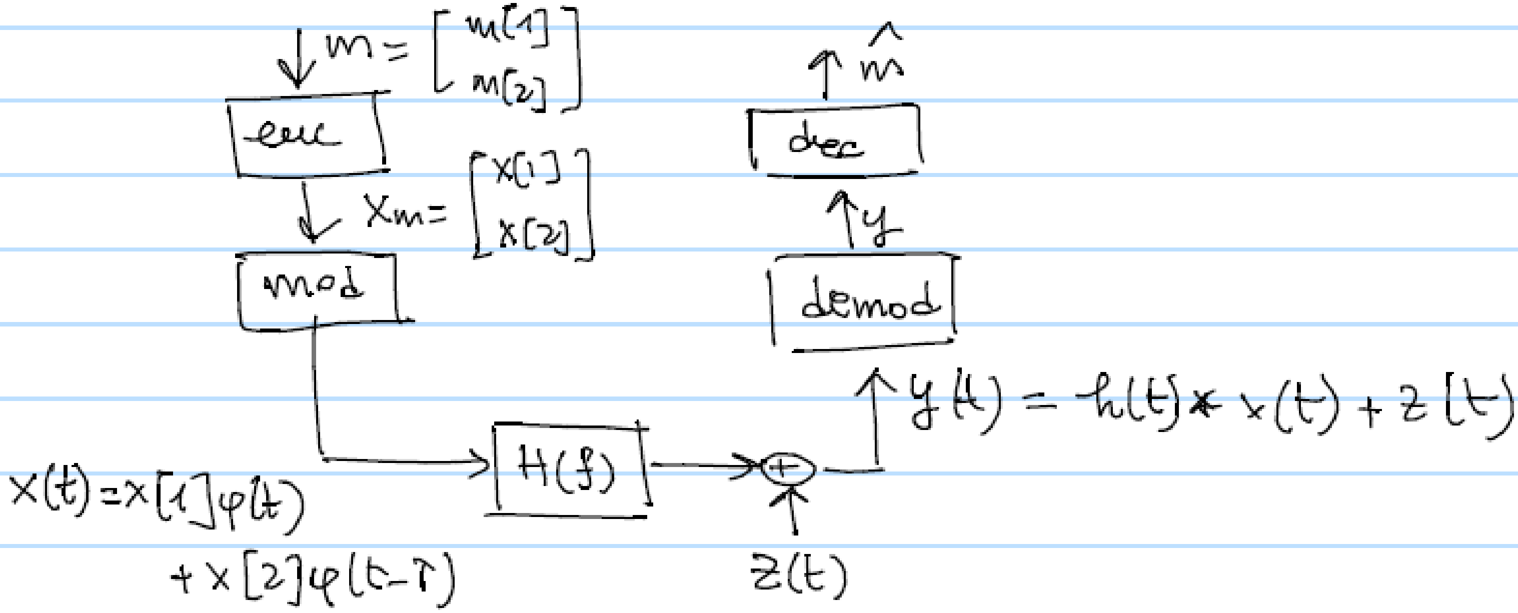
BPSK dec. :



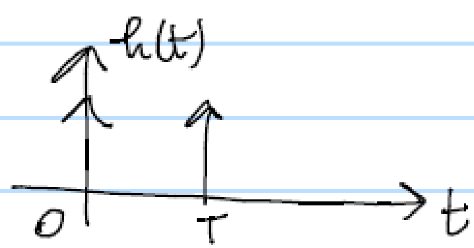
$\Rightarrow$ decoding can be carried out symbol by symbol without loss of optimality and the probability of bit error is $P_b = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$

What is the Effect of the Channel on Data Streaming?

Assume now that there is a multipath channel



with impulse response $h(t) = \delta(t) + \delta(t-T)$



What is the Effect of the Channel on Data Streaming?

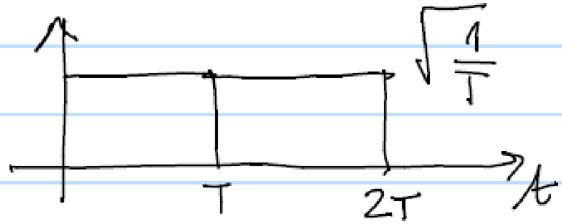
The received signal can be written as

$$y(t) = x(t) * h(t) + z(t)$$

$$= \underbrace{x[1]\big(\varphi(t) * h(t)\big)}_{= p(t)} + \underbrace{x[2]\big(\varphi(t-T) * h(t)\big)}_{= p(t-T)} + z(t)$$

$$= x[1]\, p(t) + x[2]\, p(t-T) + z(t)$$

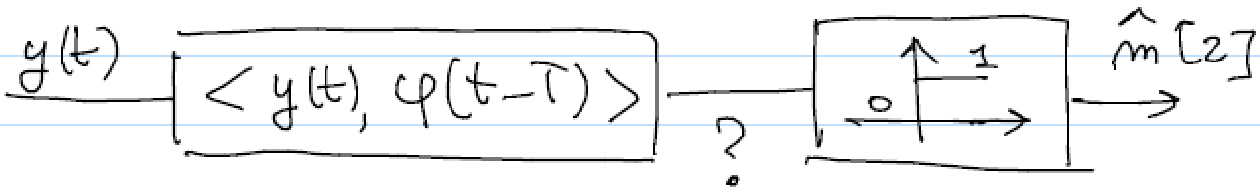
with $p(t) = \varphi(t) * h(t)$

$=$ 

$= \sqrt{\frac{1}{T}}\ \mathrm{rect}\left(\frac{t}{2T}\right)$

What is the Effect of the Channel on Data Streaming?

Let us try to apply the symbol-by-symbol decoder for the second symbol:



$$\langle x(t), \varphi(t-T)\rangle = \langle x[1]\,p(t) + x[2]\,p(t-T), \varphi(t-T)\rangle$$

$$= x[1]\underbrace{\langle p(t), \varphi(t-T)\rangle}_{=1} + x[2]\underbrace{\langle p(t-T), \varphi(t-T)\rangle}_{=1}$$

$$= x[1] + x[2]$$

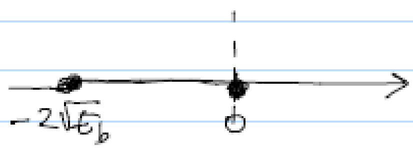
$\Rightarrow$ we have inter-symbol interference (ISI)!

What is the Effect of the Channel on Data Streaming?

Probability of bit error :

$$P_b = Pr[m[1]=0] \times Pr[\hat{m}[2] \neq m[2] \mid m[1]=0]$$
$$+ Pr[m[1]=1] \times Pr[\hat{m}[2] \neq m[2] \mid m[1]=1]$$

- if $m[1]=0 \Rightarrow x[1] = -\sqrt{E_b}$ and the equivalent constellation is

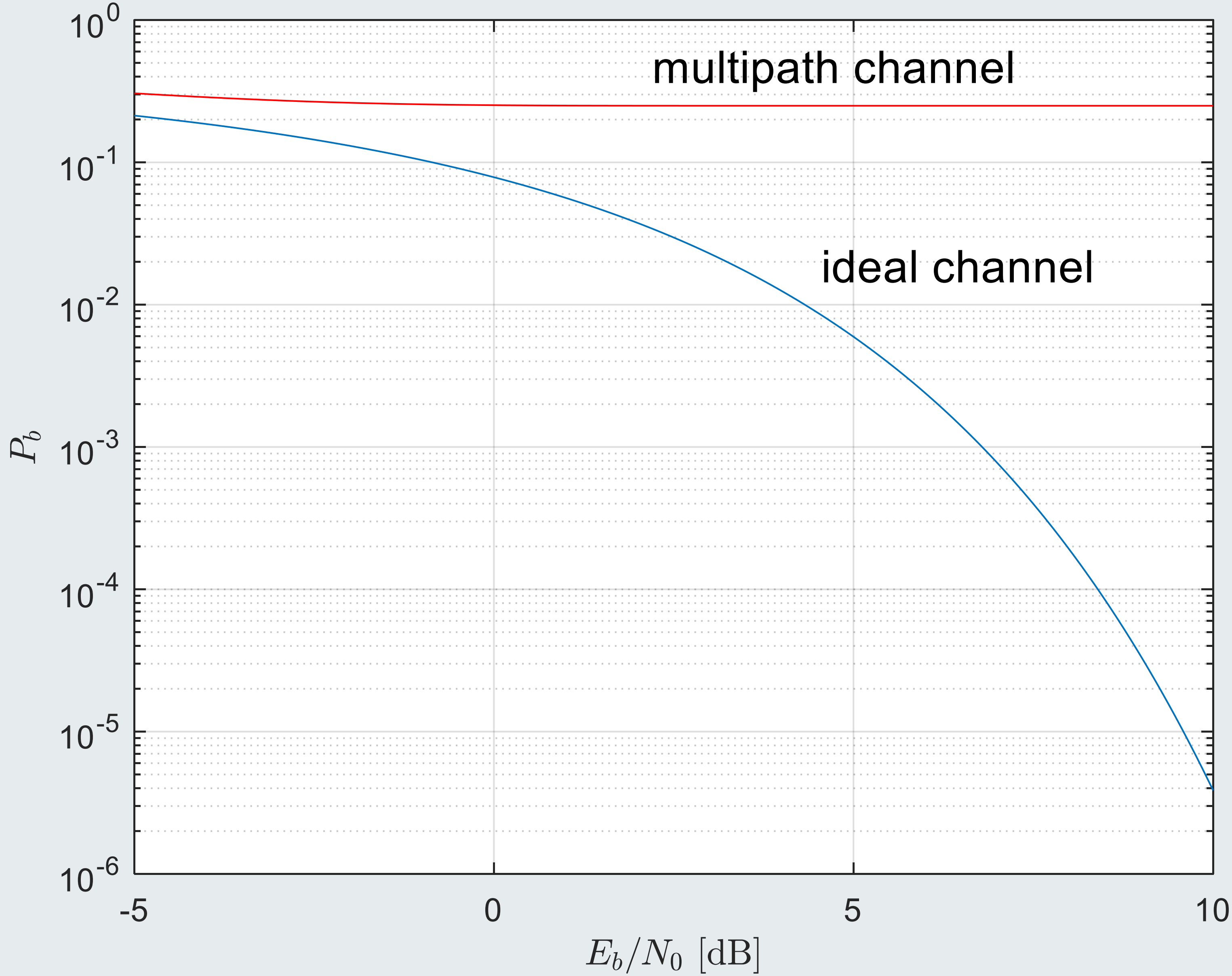


$$Pr[\hat{m}[2] \neq m[2] \mid m[1]=0] = \frac{1}{2} \times Q\left(\frac{2\sqrt{E_b}}{\sqrt{N_0/2}}\right) + \frac{1}{2} \times \frac{1}{2}$$
$$= \frac{1}{2} Q\left(\sqrt{\frac{8E_b}{N_0}}\right) + \frac{1}{4}$$

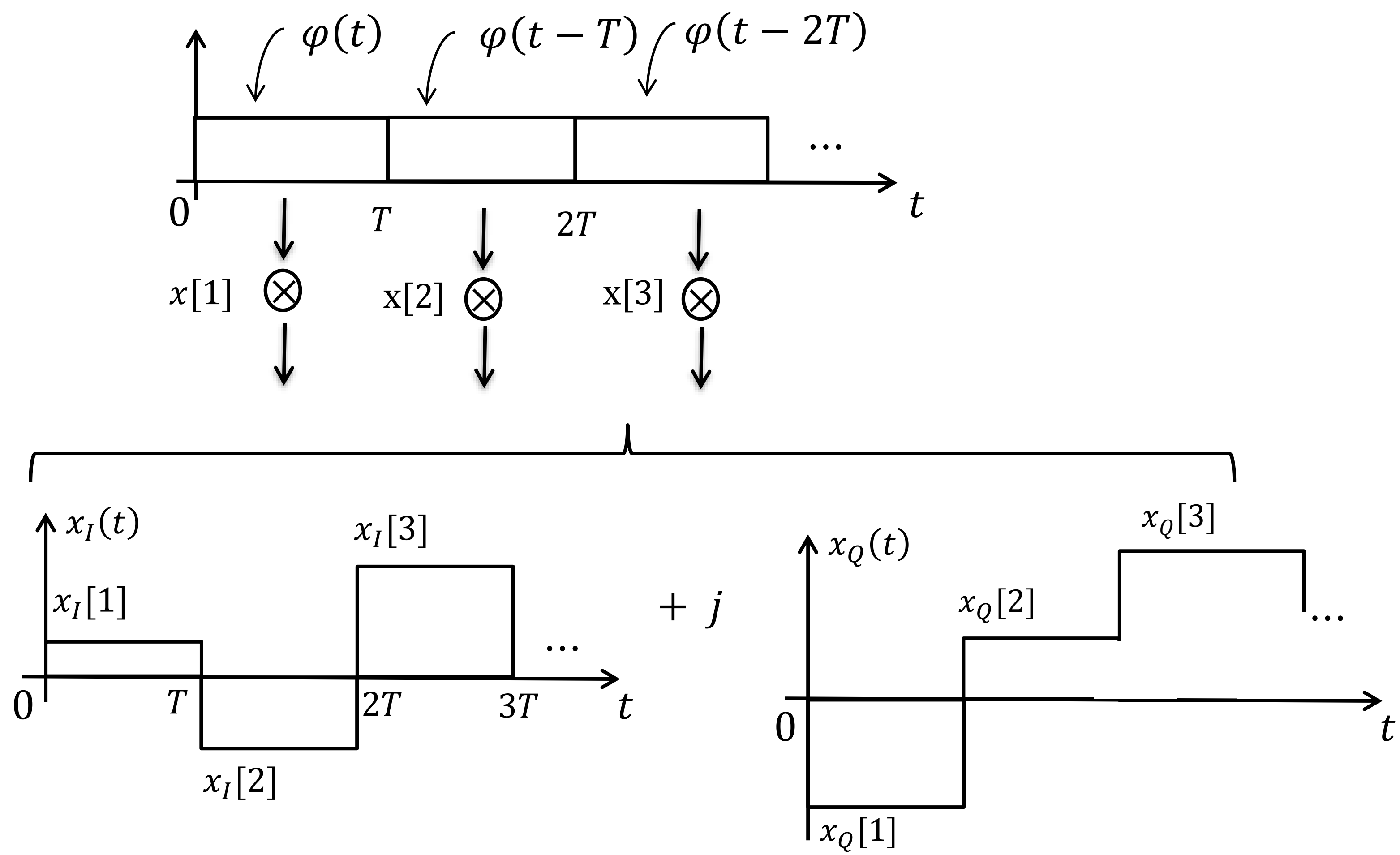
- Same probability if $m[1]=1$

$$\Rightarrow P_b = \frac{1}{2} Q\left(\sqrt{\frac{8E_b}{N_0}}\right) + \frac{1}{4}$$

What is the Effect of the Channel on Data Streaming?



How to Stream Data in Time Domain?



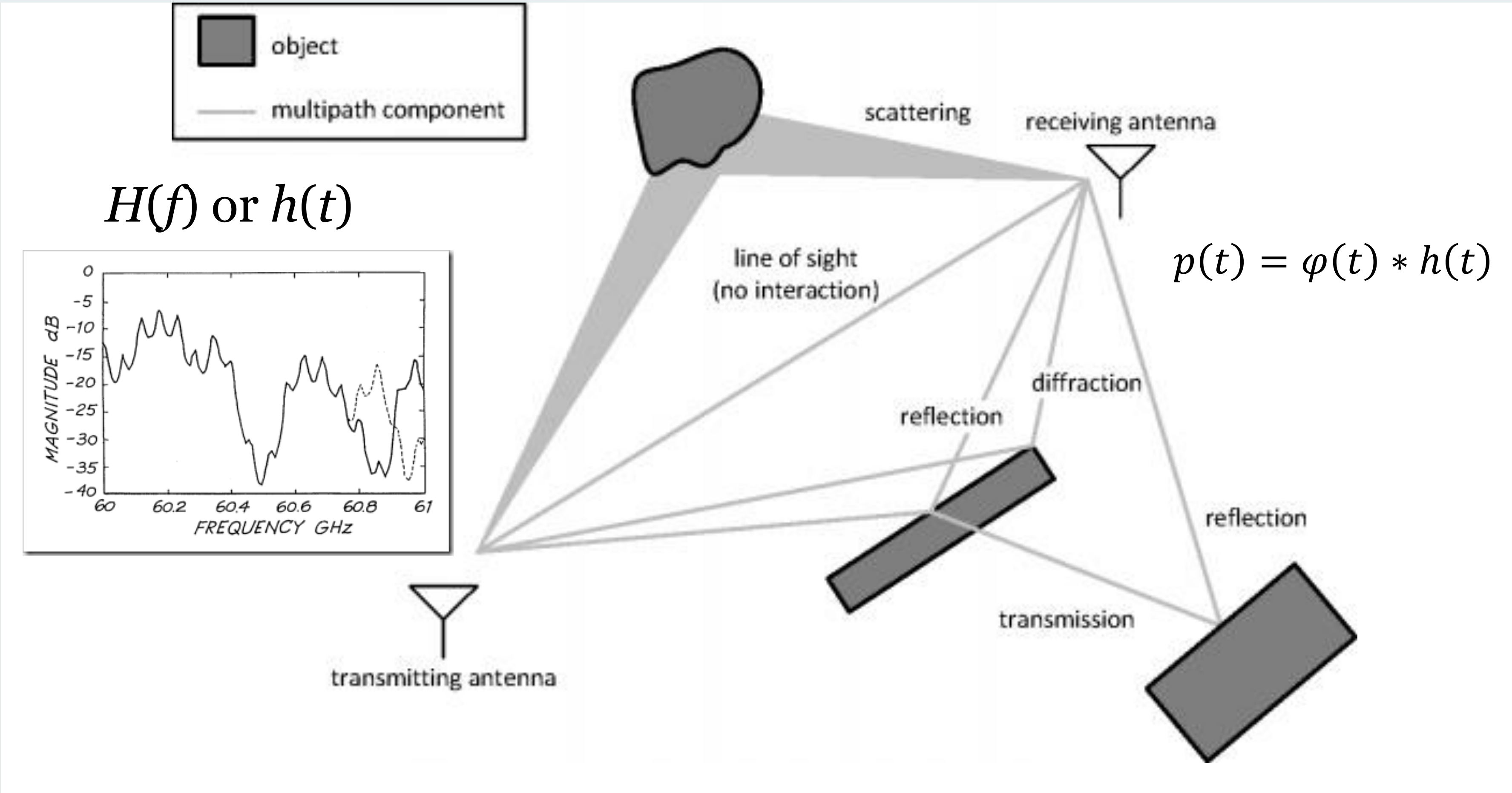
- Note: With rectangular waveforms: no ISI (with an ideal channel), but generally unacceptable spectrum.

How to Stream Data in Time Domain?

- The baseband equivalent transmitted baseband signal can be written as (dropping the subscript “bb”):

$$x(t) = \sum_{l} x[l]\, \varphi(t - lT)$$

What is the Effect of the Channel?



- **Frequency selectivity:** If the transmitter uses waveform $\varphi(t)$, the **effective waveform** is given by the convolution $p(t) = \varphi(t) * h(t)$.

What is the Effect of the Channel?

- As a result of the channel, the received signal is

$$y(t) = \sum_{l} x[l]\, p(t - lT) + z(t)$$

where $p(t) = \varphi(t) * h(t)$.

What is the Effect of the Channel?

- **Example:** As seen, if

$\varphi(t)=1/\sqrt{T}\mathrm{rect}(t/T)$
$h(\mathrm{t})=\delta(t) + \delta\ (t - T\ )$

we have

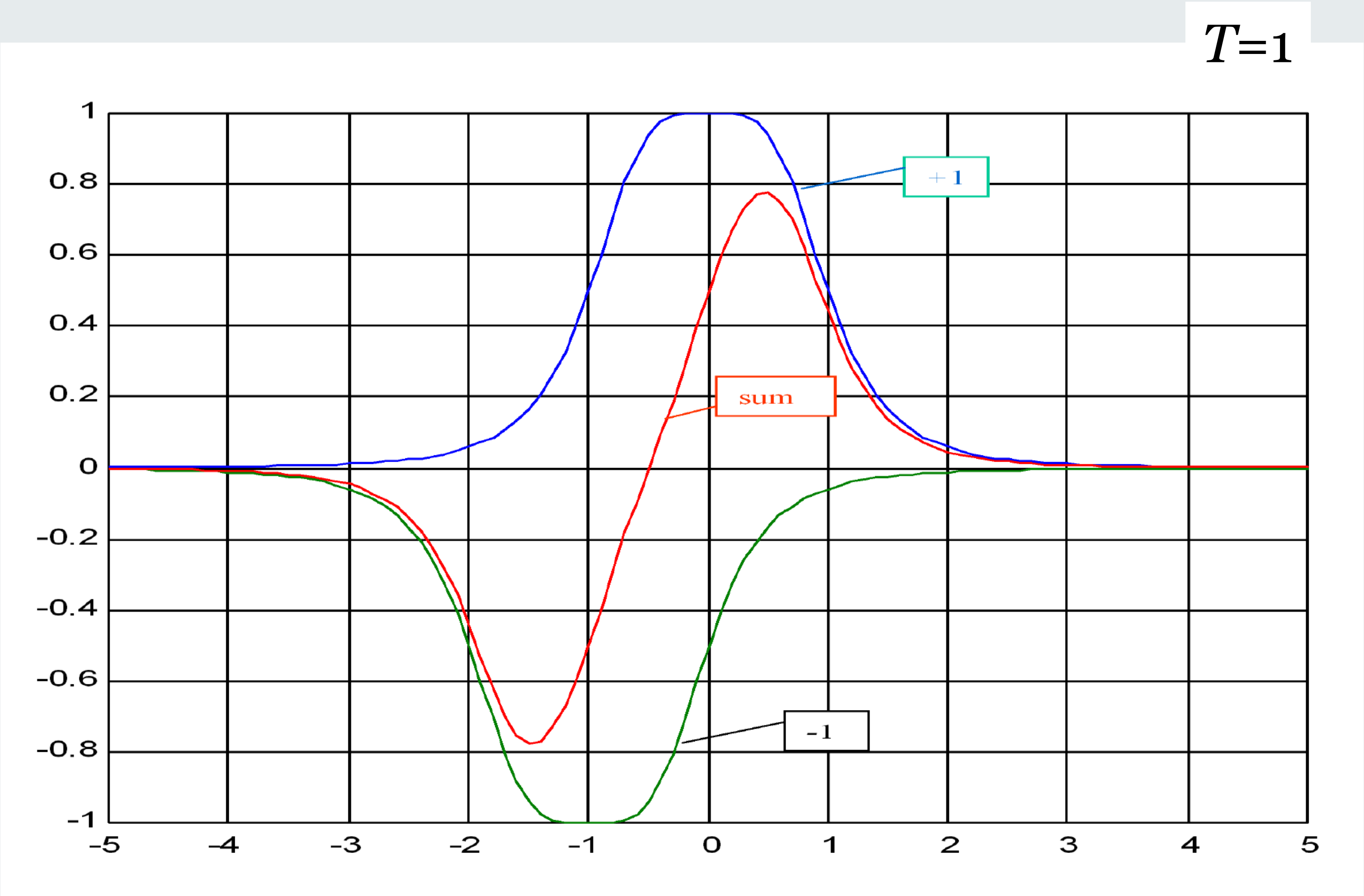
$p(t) = 1/\sqrt{2T}\ \mathrm{rect}(t/(2T))$

... the effective waveform suffers from ISI while the original waveform does not. In other words, the ISI is created by the multipath channel.

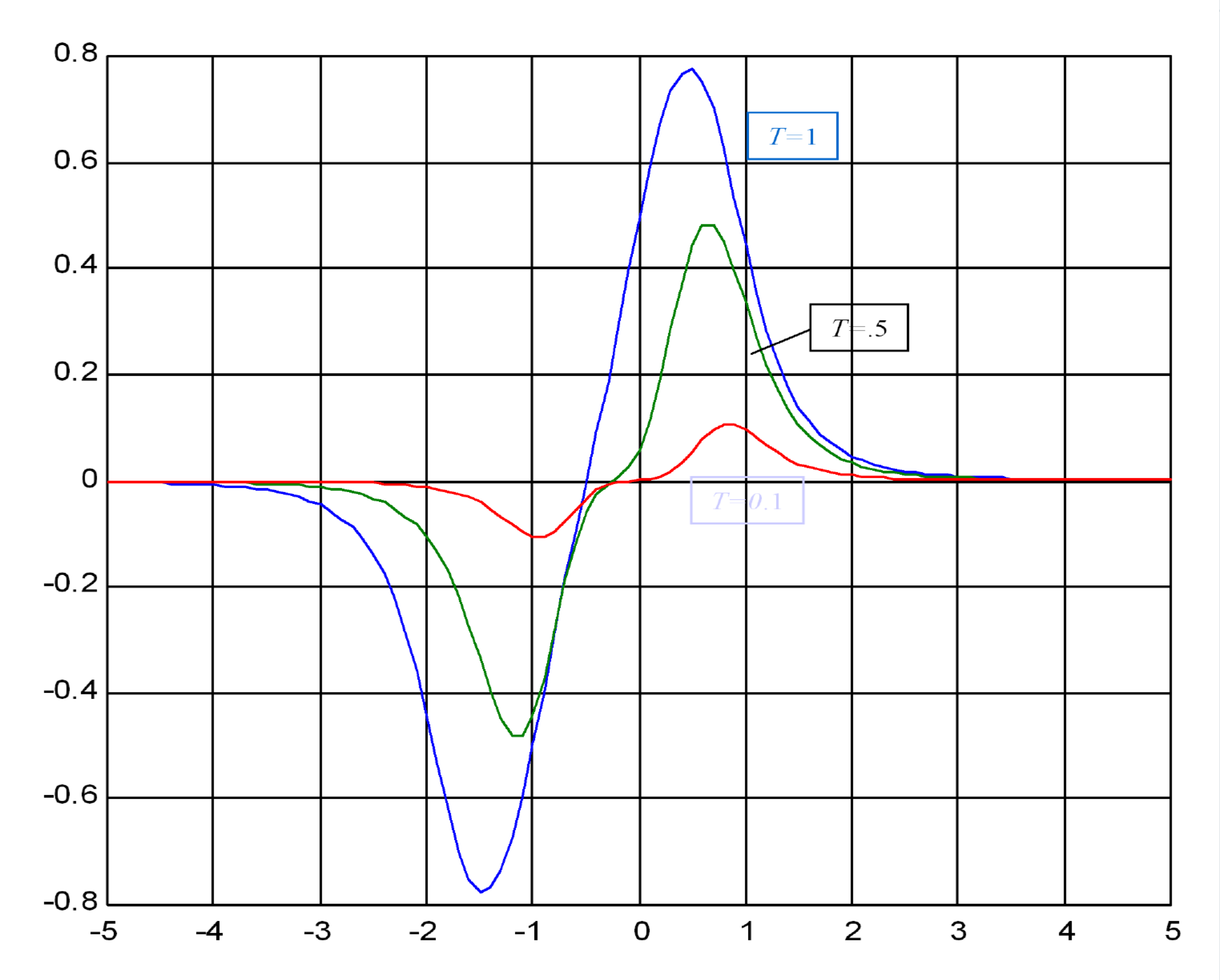
What is the Effect of the Channel?

- **Example:** ISI can limit the transmission rate

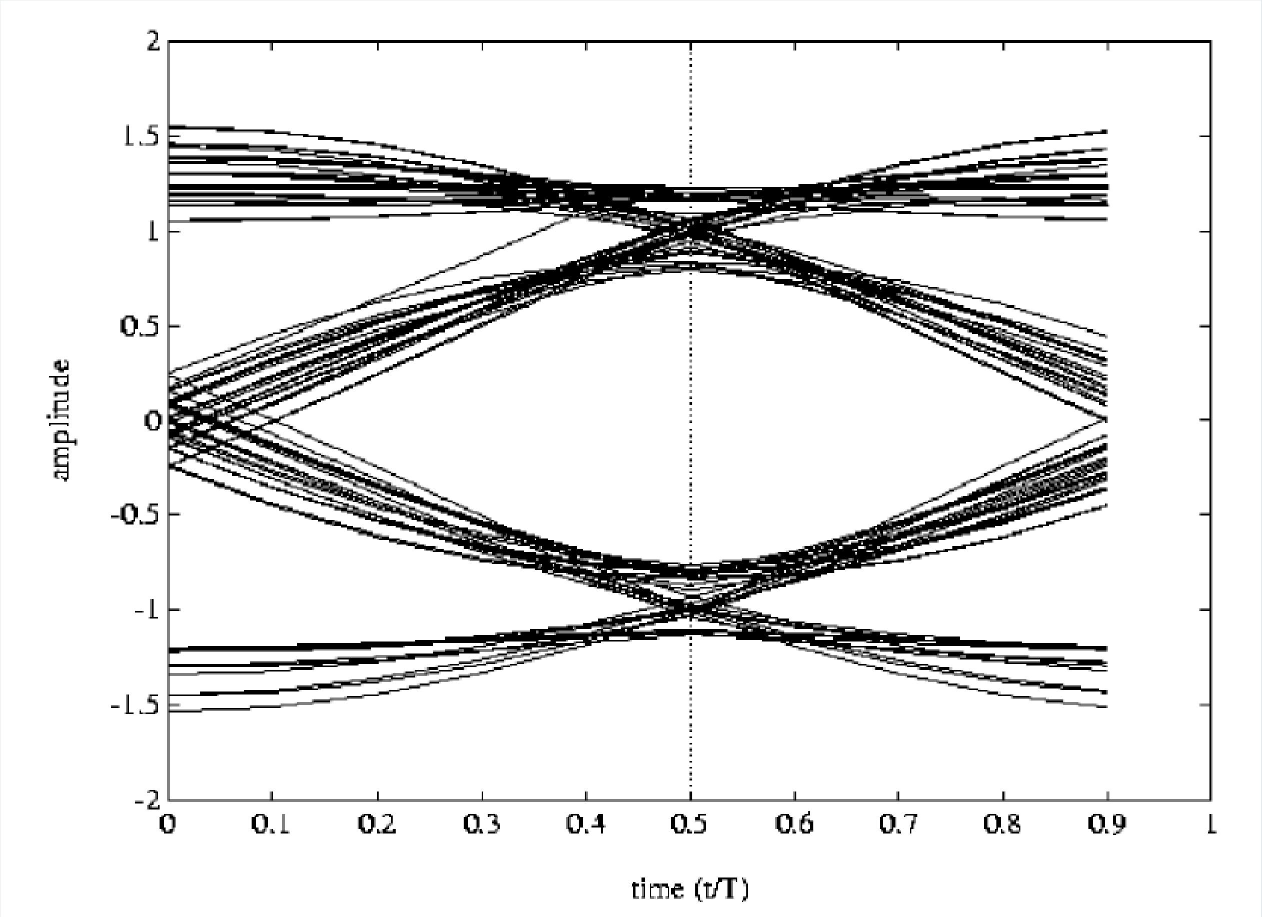
$p(t) = 1/(1+t^4)$; BPSK; $x(-T)=-1$ and $x(0)=1$



What is the Effect of the Channel?

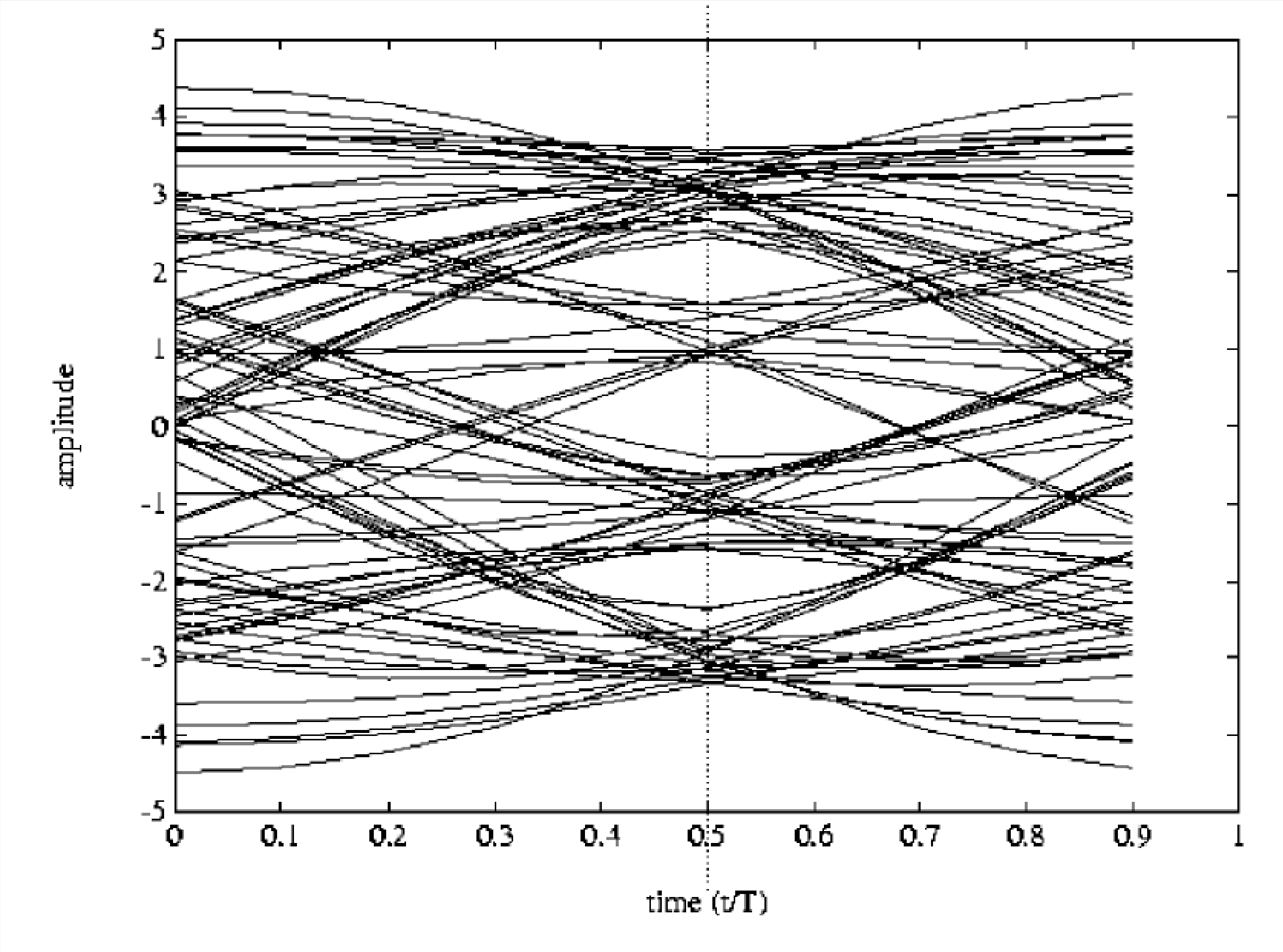


What is the Effect of the Channel?



eye diagram

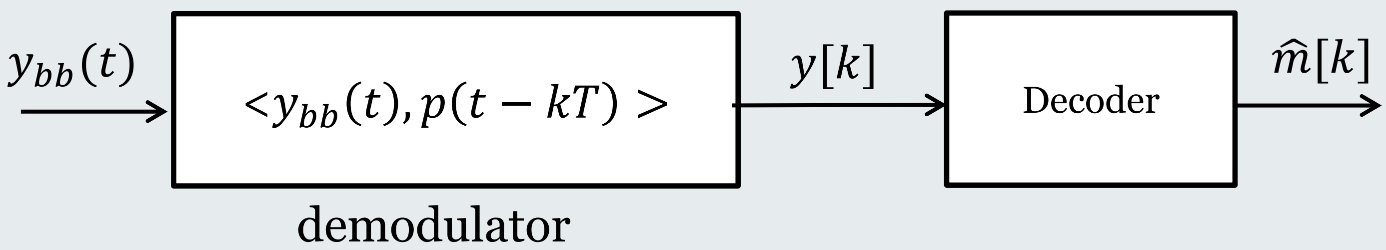
What is the Effect of the Channel?



eye diagram (for 4-PAM)

How Can We Avoid ISI?

- In the absence of ISI, the optimal correlative demodulator would operate as follows:

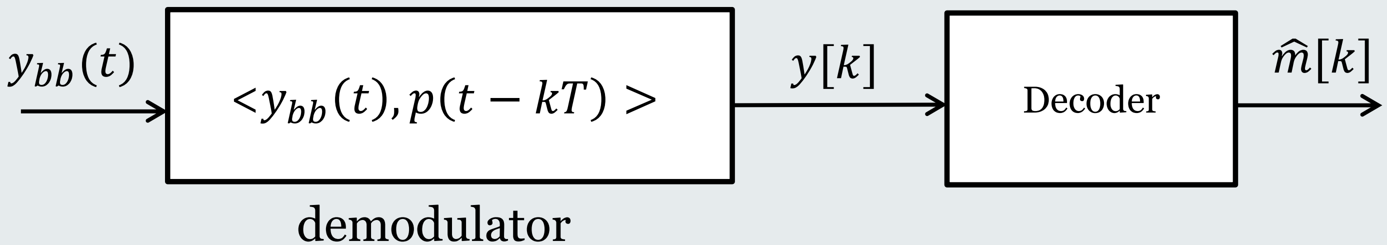


- Note that each symbol is encoded separately and hence it can also be decoded separately if there is no ISI.

- Note also that, unlike the simpler demodulator considered in the introductory example, this demodulator requires knowledge of the channel and is hence a coherent decoder.

How Can We Avoid ISI?

- In the absence of ISI, the optimal correlative demodulator would operate as follows:



- Let us calculate $y[k]$ assuming no noise for simplicity:

$$y[k] = \sum_{l} x[l] < p(t-lT), p\,(t-kT) >$$

$$= \mathrm{x}[k] \underbrace{< p(t-kT), p\,(t-kT) >}_{E_p}$$

$$\underbrace{+ \sum_{l \neq k} x[l] < p(t-lT), p\,(t-kT) >}_{\text{inter-symbol interference (ISI)}}$$

How Can We Avoid ISI?

- In the absence of ISI, the optimal correlative demodulator would operate as follows:

$$y_{bb}(t) \longrightarrow \underbrace{\boxed{<y_{bb}(t), p(t-kT)>}}_{\text{demodulator}} \xrightarrow{y[k]} \boxed{\text{Decoder}} \xrightarrow{\hat{m}[k]}$$

- Let us calculate $y[k]$ assuming no noise for simplicity:

$$y[k] = \sum_{l} x[l] < p(t-lT), p\,(t-kT) >$$

$$= \mathrm{x}[k] \underbrace{< p(t-kT), p\,(t-kT) >}_{E_p}$$

$$\underbrace{+ \sum_{l \neq k} x[l]\, R_p\big((l-k)T\big)}_{\text{inter-symbol interference (ISI)}}$$

autocorrelation function

$$R_p(\tau) =< p(t), p(t-\tau) >$$

How Can We Avoid ISI?

- As a first observation, the useful signal x[k] is multiplied by the energy of the waveform Ep. Therefore, this decoder is able to collect all the energy created by the channel through multiple propagation paths.

- In order to have zero ISI, we need to ensure that the effective waveform $p(t)$ is such that its correlation function satisfies

$$\boxed{R_p(kT) = 0 \qquad \text{for all } k \neq 0}$$

- Nyquist criterion for zero ISI: $p(t)$ should be orthogonal to all its time shifts at multiples of the symbol time T.

How Can We Avoid ISI?

- As a first observation, the useful signal x[k] is multiplied by the energy of the waveform Ep. Therefore, this decoder is able to collect all the energy created by the channel through multiple propagation paths.

- In order to have zero ISI, we need to ensure that the effective waveform $p(t)$ is such that its correlation function satisfies

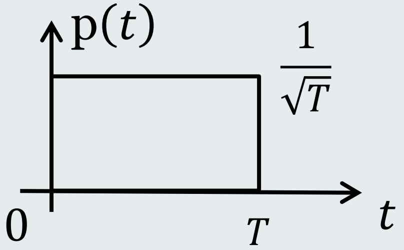
$$\boxed{R_p(kT) = 0 \qquad \text{for all } k \neq 0}$$

- Nyquist criterion for zero ISI: $p(t)$ should be orthogonal to all its time shifts at multiples of the symbol time T.

- More generally, ISI from one symbol l to another symbol k depends on how far two symbols are through the autocorrelation $Rp\ ((l-k)T)$.

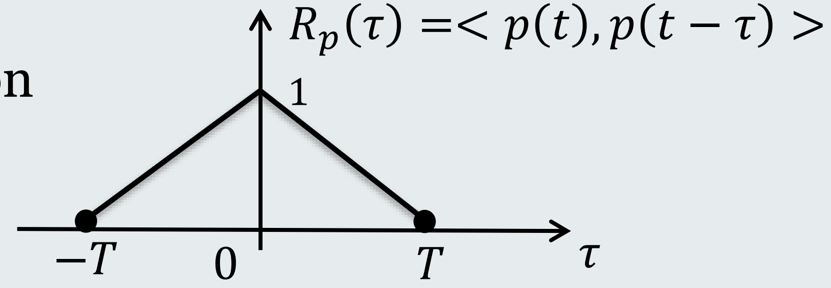
How Can We Avoid ISI?

Ex.: a)



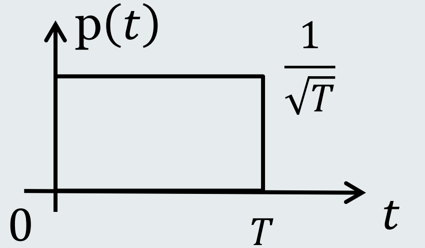
satisfies the Nyquist criterion

since:

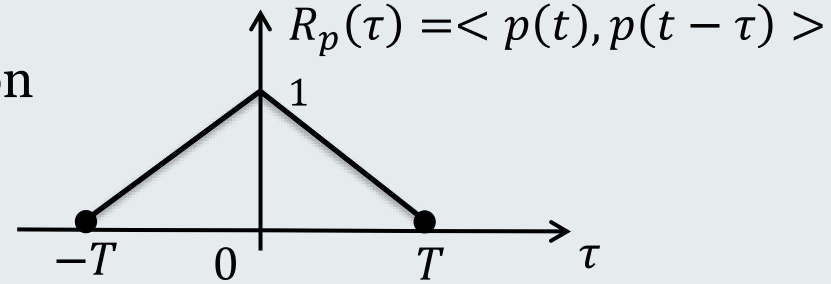


How Can We Avoid ISI?

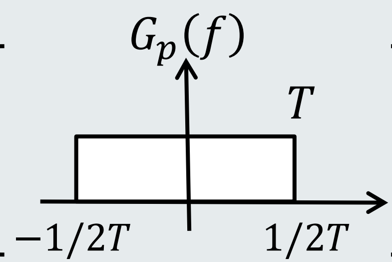
Ex.: a)



satisfies the Nyquist criterion since:

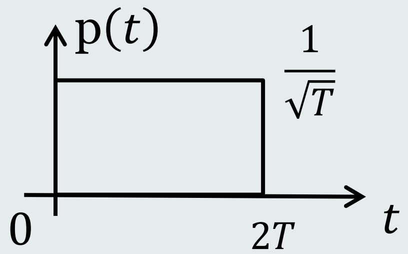


b) $p(t) = \frac{1}{\sqrt{T}} \text{sinc}\left(\frac{t}{T}\right)$ satisfies the Nyquist criterion since

$R_p(\tau) = \mathcal{F}^{-1}$ {  } $= \text{sinc}\left(\frac{\tau}{T}\right)$

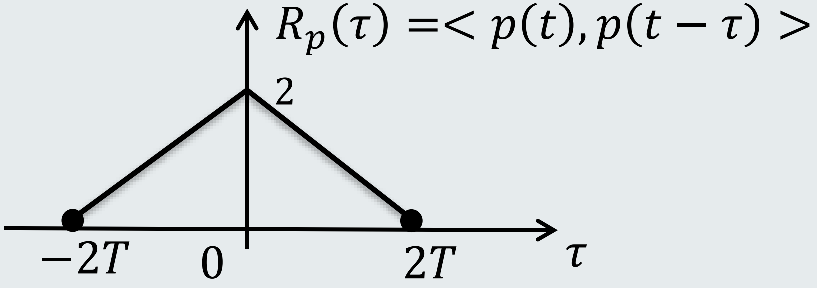
How Can We Avoid ISI?

Ex.: c)



does not satisfy the Nyquist criterion

since:



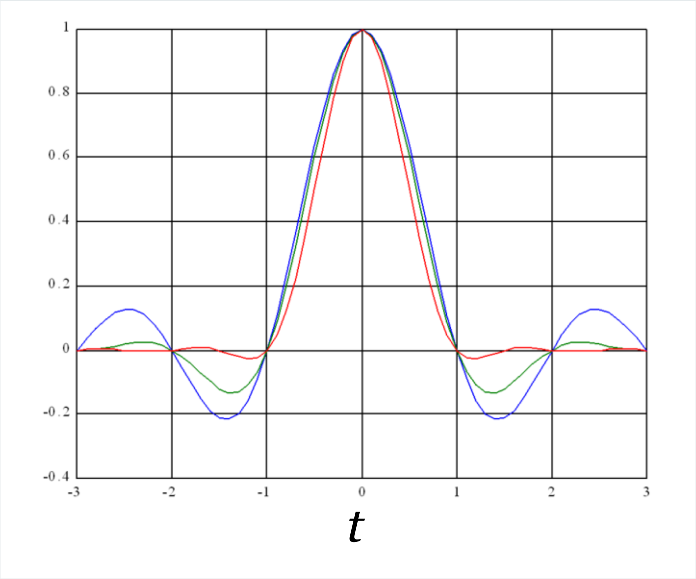
How Can We Avoid ISI?

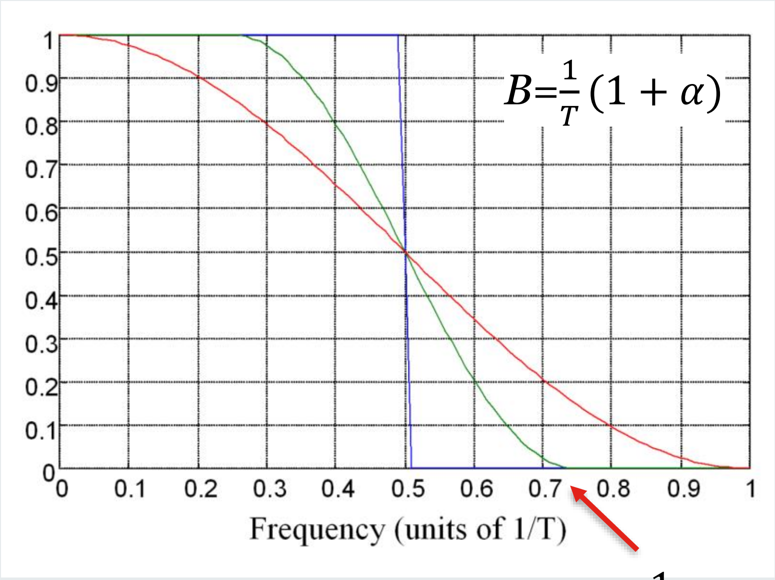
- d) We can reduce the impact of ISI by choosing waveforms with lower sidelobes than the sinc.

- A typical example is given by raised-cosine waveforms

$$p(t) = sinc\left(\frac{t}{T}\right) \cdot \left[\frac{\cos\left(\frac{\alpha\pi t}{T}\right)}{1-\left(\frac{2\alpha t}{T}\right)^2}\right]$$

α = roll-off factor

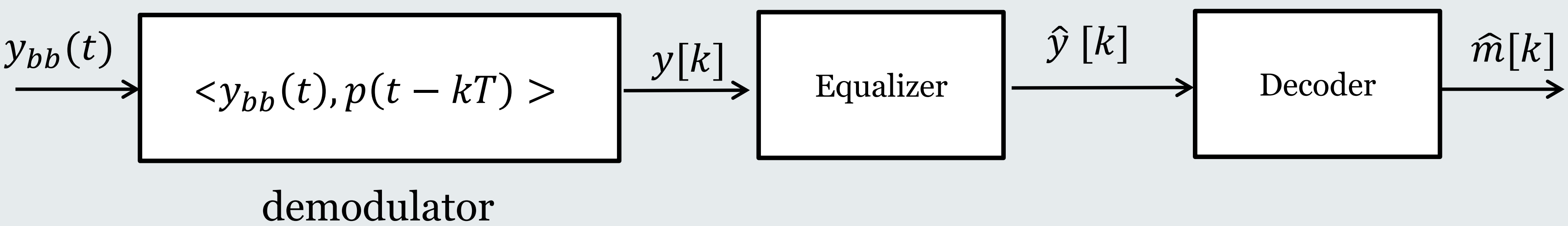




$$\frac{1}{2T}(1+\alpha)$$

What to Do if ISI is Present?

- **Equalization** methods are used by communication engineers to mitigate the effects of the intersymbol interference.



- The equalizer attempts to compensate for the channel

- Equalizers can be linear or non-linear.

What to Do if ISI is Present?

- Following the introductory example assume that

$$y[k] = x[k] + x[k-1] + z[k]$$

- Linear equalizer:

$$\hat{y}[k] = -\hat{y}[k-1] + y[k]$$

What to Do if ISI is Present?

What does this do?

$$\hat{y}[0] = 0$$

$$\hat{y}[1] = y[1]$$

$$\hat{y}[2] = -\hat{y}[1] + y[2]$$

$$= -y[1] + y[2]$$

$$= \cancel{-x[1]} - \underbrace{x[0]}_{0} - z[1] + x[2] + \cancel{x[1]} + z[2]$$

$$= x[2] + z[2] - z[1]$$

What to Do if ISI is Present?

What does this do?

$$\hat{y}[0] = 0$$

$$\hat{y}[1] = y[1]$$

$$\hat{y}[2] = -\hat{y}[1] + y[2]$$

$$= -y[1] + y[2]$$

$$= -\cancel{x[1]} - \underbrace{\cancel{x[0]}}_{0} - z[1] + x[2] + \cancel{x[1]} + z[2]$$

$$= x[2] + z[2] - z[1]$$

no ISI! ✓

noise has doubled in power! ✗

What to Do if ISI is Present?

⇒ Signal-to-noise-plus-interference ratio (SINR):

- without equalization:

$$\mathrm{SINR} = \frac{E_x}{E_x + N_0} = \frac{1}{1 + N_0/E_x}$$

- with equalization

$$\mathrm{SINR} = \frac{E_x}{2N_0}$$

- Note: The SINR of the linear equalizer scheme decreases for successive symbols: equalization causes noise to accumulate (try!).

What to Do if ISI is Present?

- Non-linear equalizer:
  - from $y[k-1]$ decode $x[k-1] \rightarrow \hat{x}[k-1]$
  - cancel $\hat{x}[k-1]$:

$$\hat{y}[k] = -\hat{x}[k-1] + y[k]$$

What to Do if ISI is Present?

- What does this do?

$$\hat{y}[2] = -\hat{x}[1] + y[2]$$

- if the decision $\hat{x}[1]$ is correct:

$$\hat{y}[2] = \cancel{-x[1]} + x[2] + \cancel{x[1]} + z[2]$$

$$= x[2] + z[2]$$

no ISI! ✓

no increase in noise! ✓

What to Do if ISI is Present?

- What does this do?

$$\hat{y}[2] = -\hat{x}[1] + y[2]$$

– if the decision $\hat{x}[1]$ is not correct

$$\hat{y}[2] = -\hat{x}[1] + x[2] + x[1] + z[2]$$

$$= x[2] + (x[1] - \hat{x}[1]) + z[2]$$

increased interference!

X

What to Do if ISI is Present?

⇒ SINR:

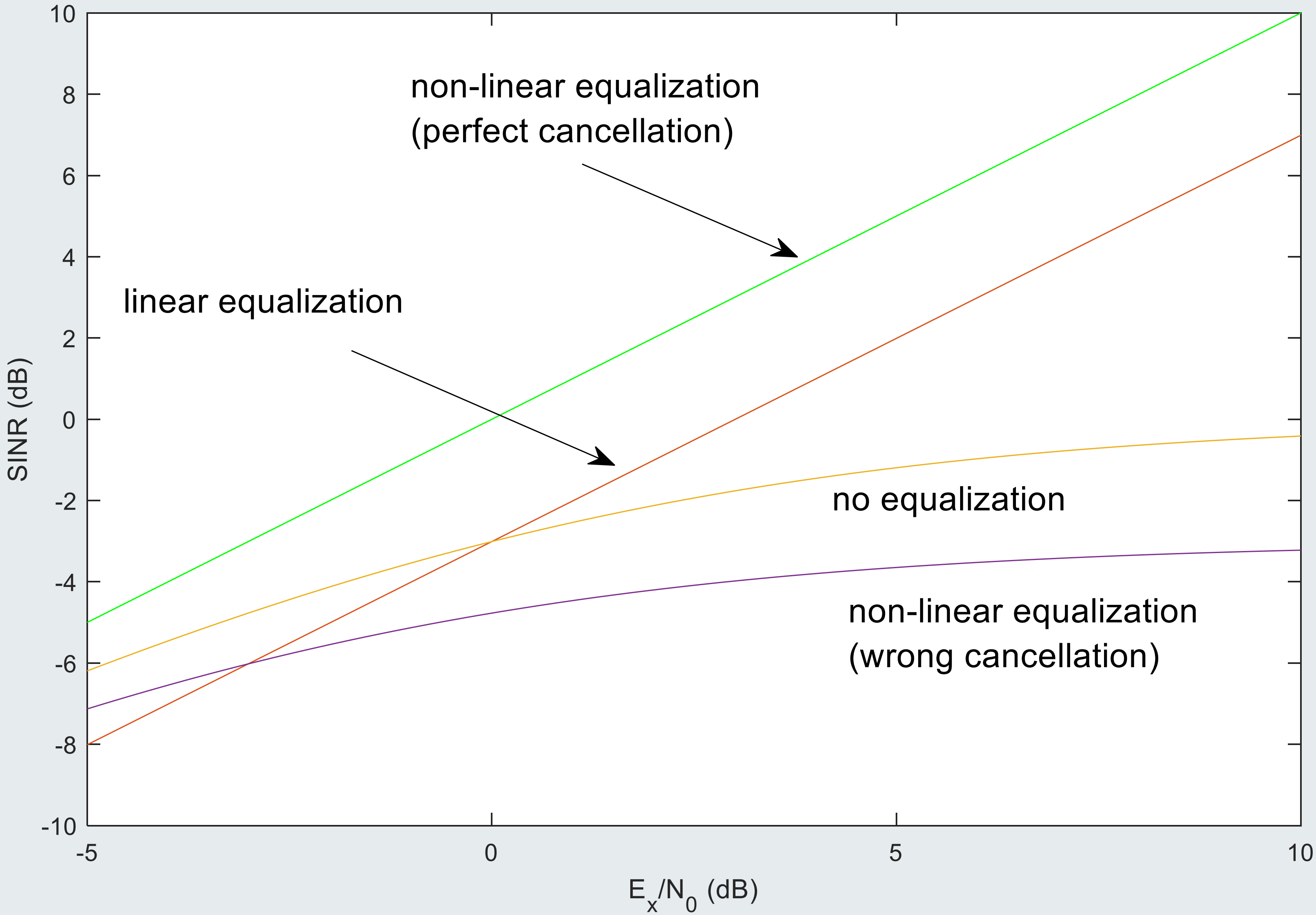
- with perfect cancellation:

$$\mathrm{SINR} = \frac{E_x}{N_0}$$

- with incorrect cancellation:

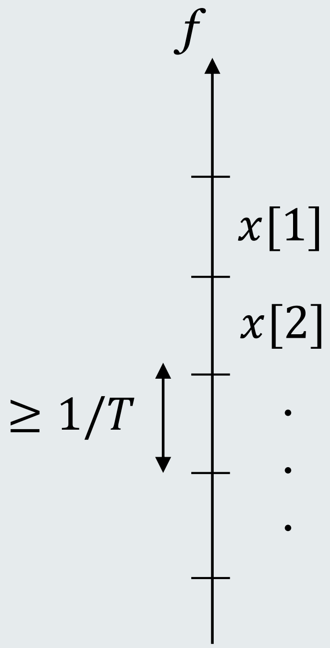
$$\mathrm{SINR} = \frac{E_x}{2E_x + N_0} = \frac{1}{2 + N_0/E_x}$$

What to Do if ISI is Present?



How to Stream Data in the Frequency Domain?

- Frequency-domain transmission:



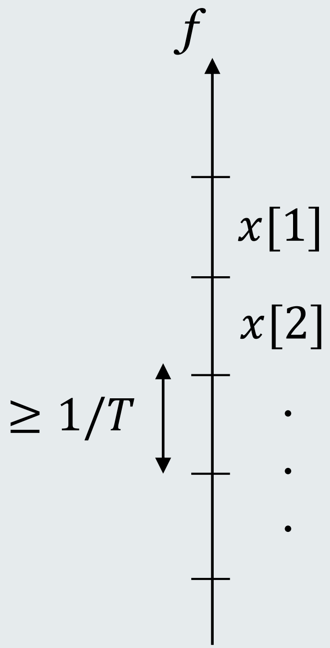
$$x(t) = \sum_{n=1}^{N} x[n]\, \varphi_n(t)$$

$$\varphi_n(t) = \mathrm{u}(t) \exp\left(j2\pi \left(\frac{n}{T} - \frac{N}{2T}\right) t\right)$$

where u(t) has energy equal to one and its Fourier transform guarantees no ISI in the frequency domain

How to Stream Data in the Frequency Domain?

- Frequency-domain transmission:



$$x(t) = \sum_{n=1}^{N} x[n]\,\varphi_n(t)$$

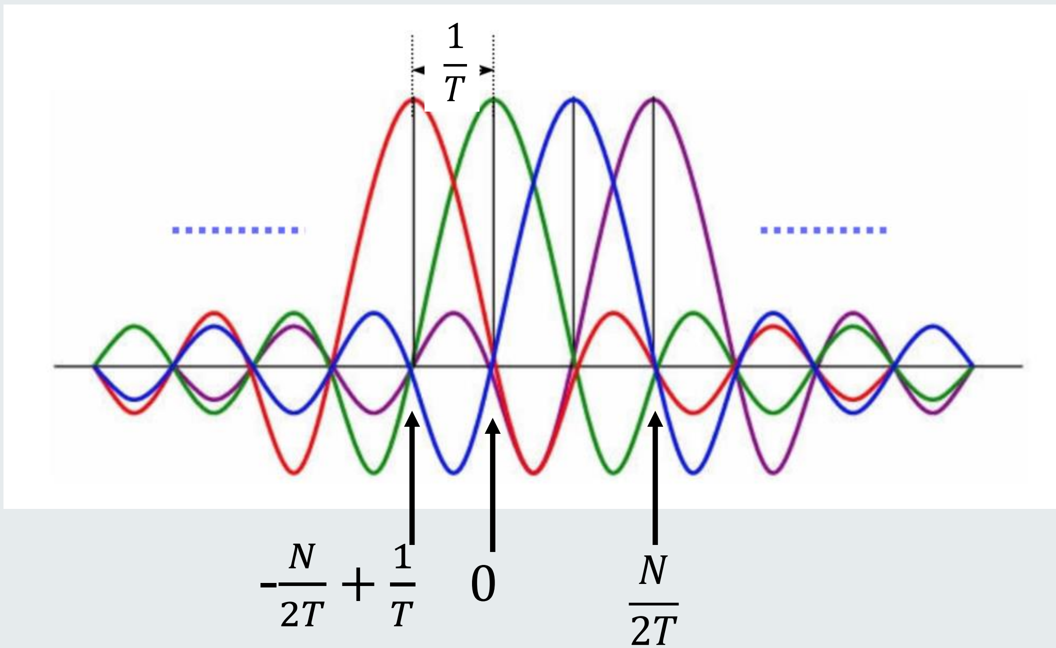
$$\varphi_n(t) = \mathrm{u}(t)\exp\left(j2\pi\left(\frac{n}{T} - \frac{N}{2T}\right)t\right)$$

where u(t) has energy equal to one and its Fourier transform guarantees no ISI in the frequency domain

- Waveform u(t) can be chosen as a rectangle with duration T or as the Fourier transform of a raised-cosine waveform.

How to Stream Data in the Frequency Domain?

- Frequency-domain transmission:



$$B \approx \frac{N}{T}$$

What is the Effect of the Channel?

- Convolutive channels (i.e., filters) $H(f)$ do not affect the orthogonality of OFDM subcarriers.

- This is one of the key advantages of OFDM, which has motivated its adoption in most modern systems.

- Each OFDM subcarrier f_n is merely multiplied by the value $H(f_n)$.

What is the Effect of the Channel?

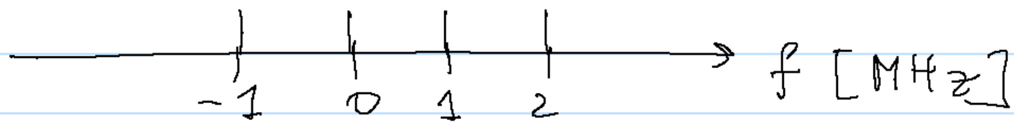
Example:

Assume an OFDM system with symbol of duration 1 μs. We have $N=4$ subcarriers. What are the channel gains on the four subcarriers if the channel is

$$h(t) = \delta(t) + \delta(t-T)\ ?$$

What is the Effect of the Channel?

Subcarriers:



channel gains:

$$H(f) = 1 + e^{-j2\pi f T}$$

$$= 1 + e^{-j2\pi f 10^{-6}}$$

$$\Rightarrow |H(f_1)| = |H(-10^6)| = |1 + e^{+j2\pi}| = 2$$

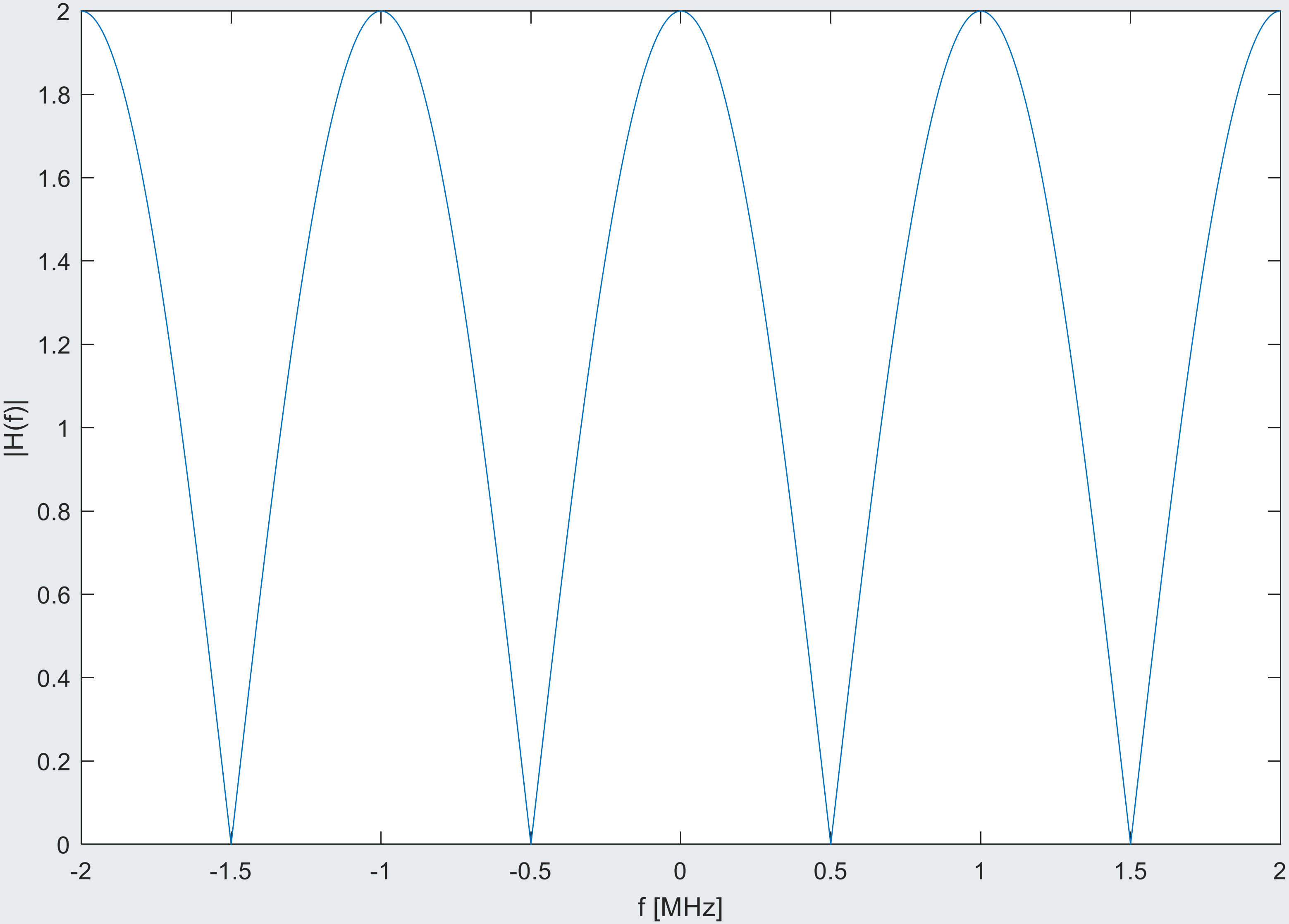
$$|H(f_2)| = |H(0)| = |1 + e^{-j0}| = 2$$

$$|H(f_3)| = |H(10^6)| = 2$$

$$|H(f_4)| = |H(2 \times 10^6)| = |1 + e^{-j4\pi}| = 2$$

No ISI and increased channel gain!

What is the Effect of the Channel?



What is the Effect of the Channel?

How about with the channel

$$h(t) = \delta(t) + \delta\left(t - \frac{T}{2}\right) ?$$

What is the Effect of the Channel?

Channel gains :

$$H(f_1) = \left|1 + e^{+j 2\pi 10^6 \frac{10^{-6}}{2}}\right| = |1 - 1| = 0$$

zero channel gain!

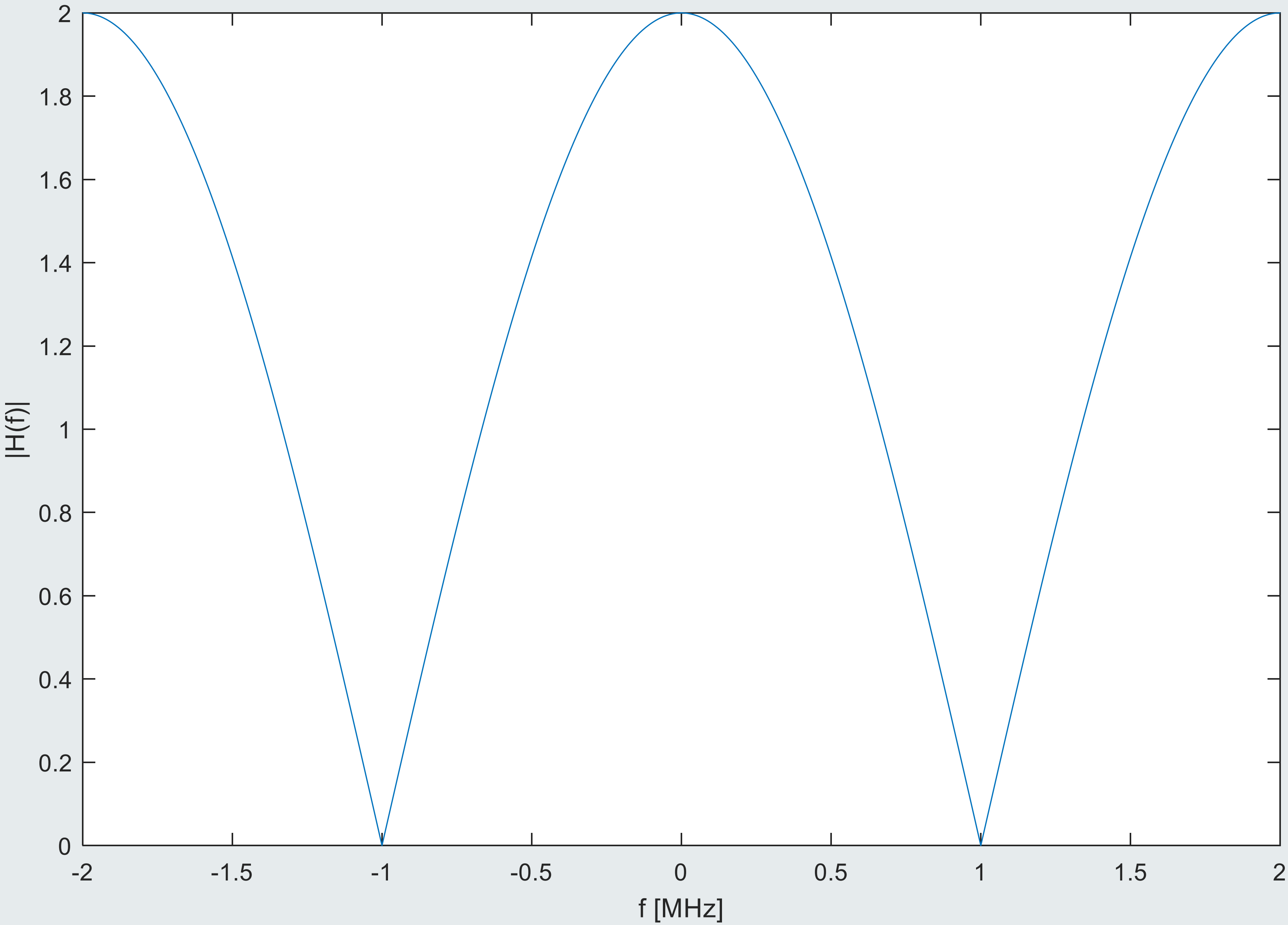
$$H(f_0) = |1 + e^{j0}| = 2$$

$$H(f_2) = \left|1 + e^{-j 2\pi 10^6 \frac{10^{-6}}{2}}\right| = 0$$

$$H(f_3) = \left|1 + e^{-j 4\pi 10^6 \frac{10^{-6}}{2}}\right| = |1 + e^{-j 2\pi}| = 2$$

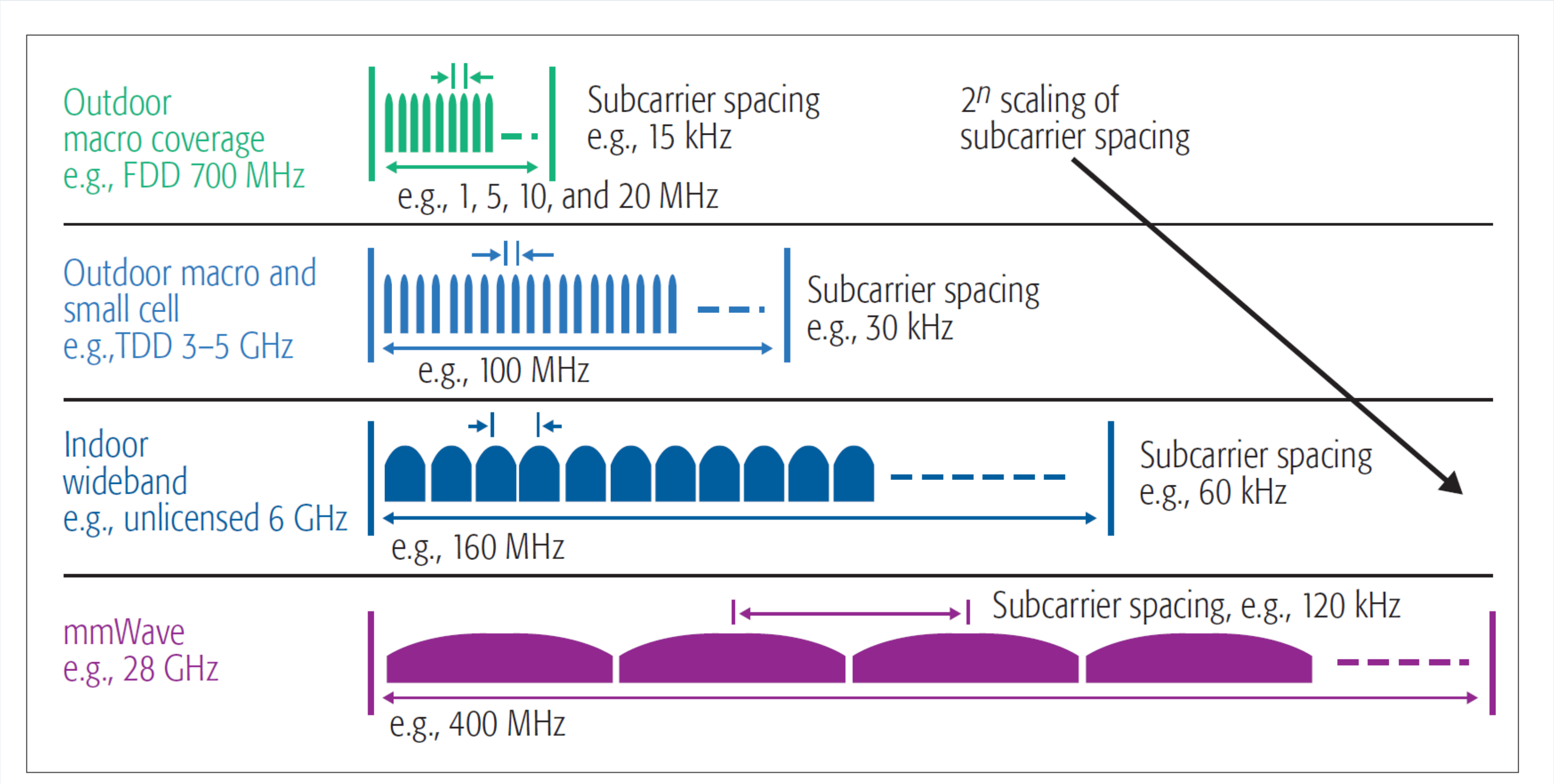
the frequency response can be zero in some subcarriers.

What is the Effect of the Channel?



What About 5G?

- The available bandwidth depends on the carrier frequency.



- Subcarrier spacing increases (and hence T decreases) with the overall bandwidth in order to avoid N being too large for complexity reasons.

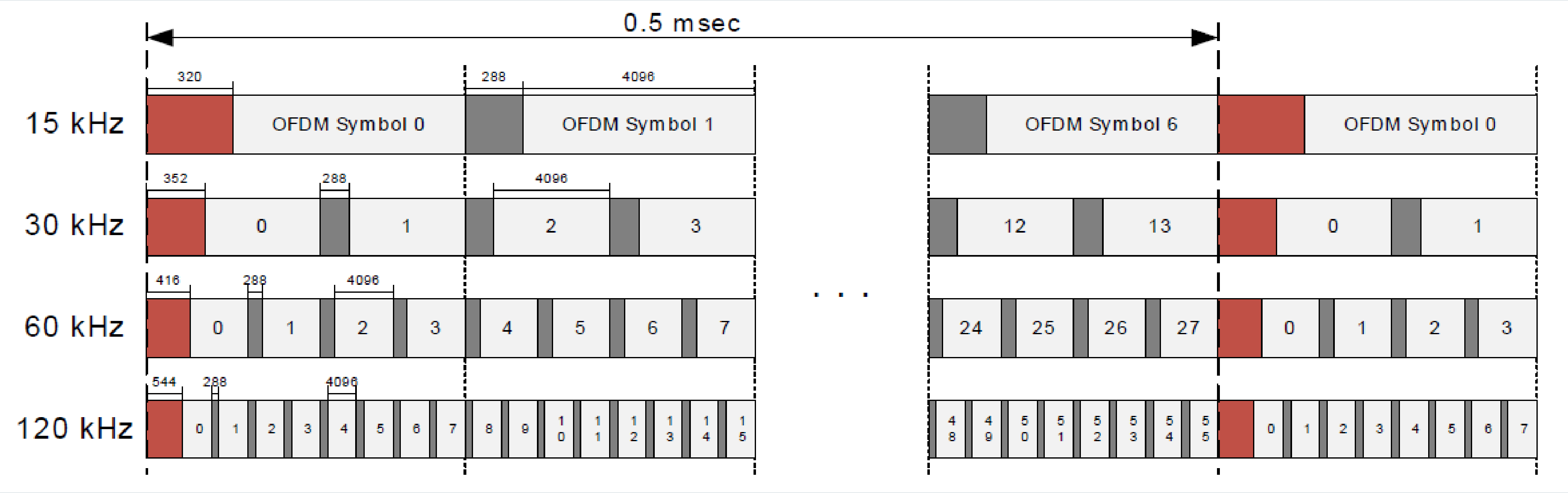
[Bhushan et al ’17]

What About 5G?

μ	$\Delta f = 2^{\mu} \cdot 15$ kHz
0	15 kHz
1	30 kHz
2	60 kHz
3	120 kHz
4	240 kHz
5	480 kHz

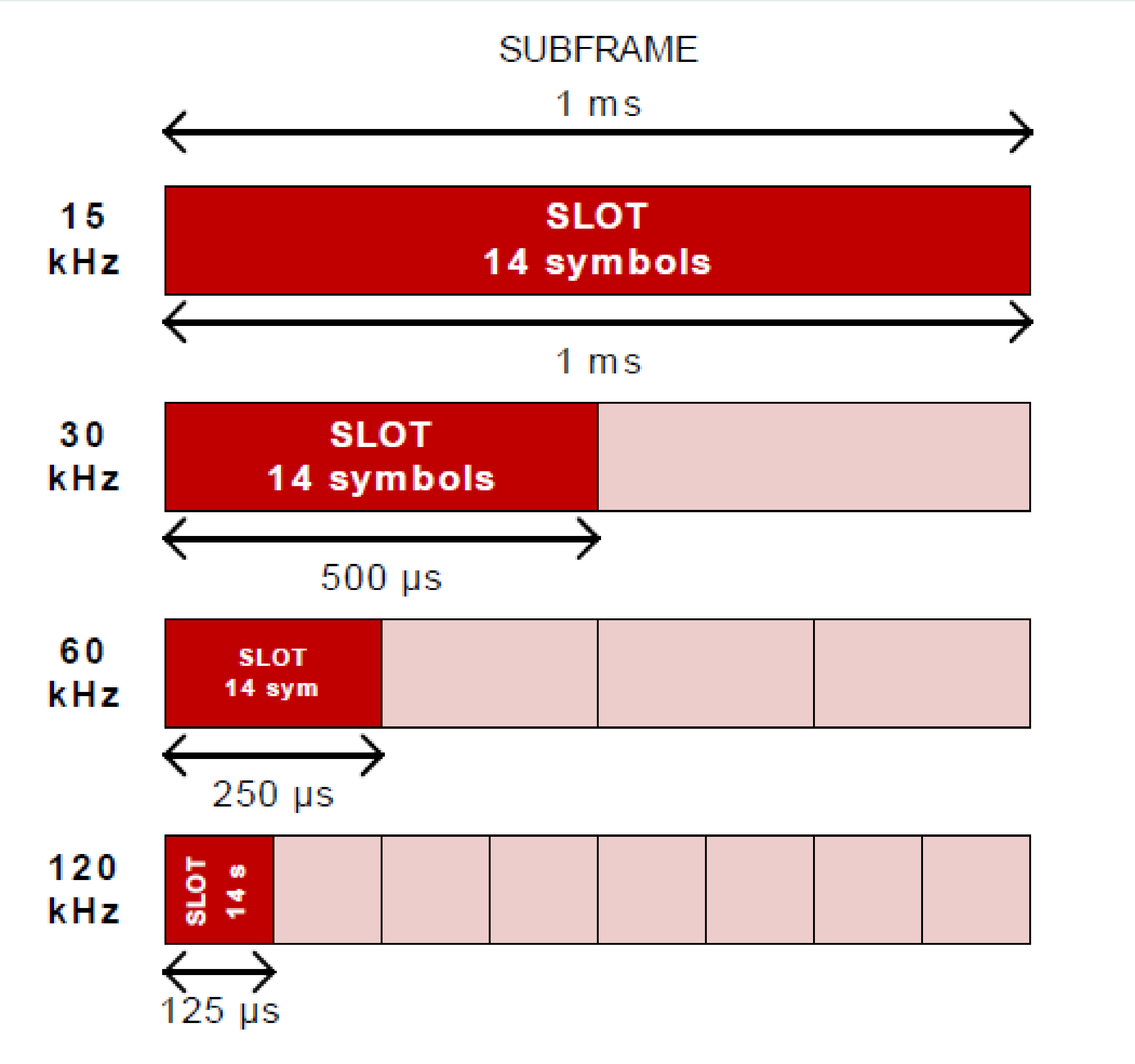
[Keysight Technologies '18]

What About 5G?



[Keysight Technologies ‘18]

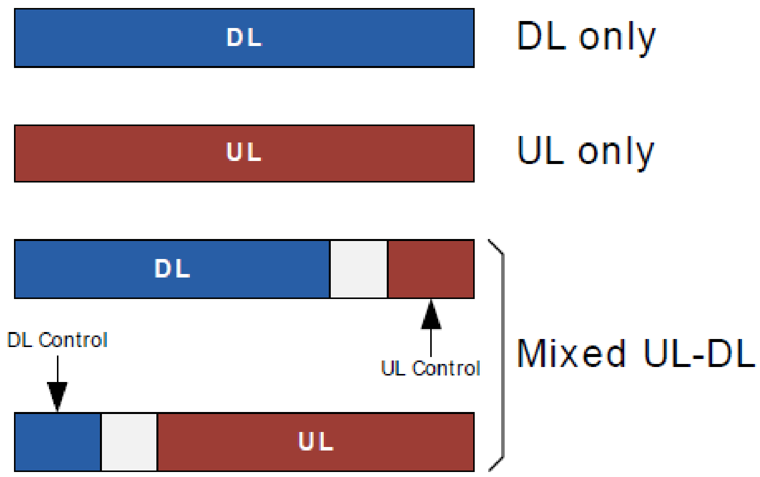
What About 5G?



[Keysight Technologies '18]

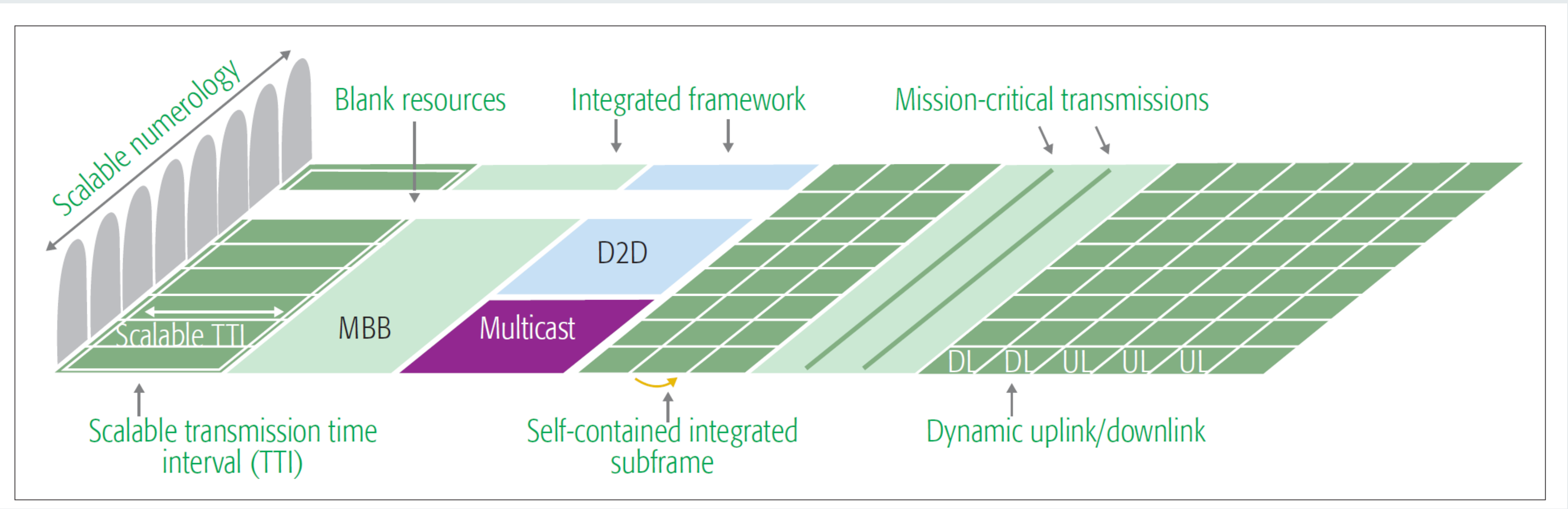
What About 5G?

- A slot can be:
  - All downlink
  - All uplink
  - Mixed downlink and uplink
    - Static, semi-static or dynamic
- Slot aggregation is supported
  - Data transmission can be scheduled to span one or multiple slots



What About 5G?

- Different services will share the same time-frequency resources.



- MBB = Mobile BroadBand
- D2D = Device-to-Device

[Bhushan et al ’17]

Can We Use a Large N to Improve the Performance?

- The encoders and decoders studied up to now operate on one symbol, i.e., on two real dimensions or one complex dimension, at a time.
- The probability of error is hence determined by the minimum distance between constellation points in the two dimensional signal space.
- This type of systems are known as being **uncoded**.
- Can we improve the probability of error by **coding** and decoding over multiple symbols (i.e., over $N>1$ complex dimensions)?
- In other words, can we improve the system performance by coding?

How Do We Define Performance?

- **1) Bandwidth or spectral efficiency**

$$\eta = \frac{R}{B} \text{ (bit/s/Hz)}$$

- Since N=BT (complex dimensions), T=1/B is the minimum symbol durations for a complex dimension or two real dimensions. Therefore, we also write

$$\eta = RT = \frac{R}{B} \text{ (bit/2D)}$$

How Do We Define Performance?

- **1) Bandwidth or spectral efficiency**

$$\eta = \frac{R}{B} \text{ (bit/s/Hz)}$$

- Since N=BT (complex dimensions), T=1/B is the minimum symbol durations for a complex dimension or two real dimensions. Therefore, we also write

$$\eta = RT = \frac{R}{B} \text{ (bit/2D)}$$

- **2) Power or energy efficiency**

$$\frac{E_b}{N_0} \quad \text{for a given } P_e$$

- **3) Complexity**

Can We Use a Large N to Improve the Performance?

Example: Repetition coding

$$\xrightarrow[b \text{ bits}]{m} \boxed{\text{Code}} \longrightarrow \boxed{\text{Const.}} \longrightarrow x[1] = x_m, \dots, x[N] = x_m$$

$$m[1] = m, \dots, m[N] = m$$

- Spectral efficiency:

$$R = \frac{b}{NT}, \quad B = \frac{1}{T} \Rightarrow \eta = \frac{b}{N} \text{ (bit/s/Hz or bit/2D)}$$

η decreases with N

Can We Use a Large N to Improve the Performance?

- Energy efficiency:

d_c = minimum distance between points in the constellation

d_{min} = minimum distance between encoded symbols ?

$$d^2_{m,m'} = \sum_{n=1}^{N} |x_m[n] - x_{m'}[n]|^2$$

$$\Rightarrow d^2_{min} = N d_c^2$$

$$\Rightarrow P_e \approx Q\left(\sqrt{\frac{N d_c^2}{2N_0}}\right)$$

Can We Use a Large N to Improve the Performance?

For example, for BPSK

$$P_e \approx Q\left(\sqrt{\frac{2E_b}{N_0} \times N}\right)$$

$\Rightarrow$ the energy efficiency has improved by N

- We have concluded that repetition coding offers:
  - a coding gain (that is, an improvement of the energy efficiency) of $10\log_{10}N$ dB
  - a reduction of spectral efficiency by $1/N$

What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?

1948

A Mathematical Theory of Communication

By C. E. SHANNON

Who is Claude Shannon Again?

- He's the most important genius you've never heard of, a man whose intellect was on par with Albert Einstein and Isaac Newton.

- At the age of 21, he published what's been called the most important master's thesis of all time, explaining how binary switches could do logic. It laid the foundation for all future digital computers.

- At the age of 32, he published "A Mathematical Theory of Communication," which has been called "the Magna Carta of the information age." Shannon's masterwork invented the *bit*, or the objective measurement of information, and explained how digital codes could allow us to compress and send any message with perfect accuracy.

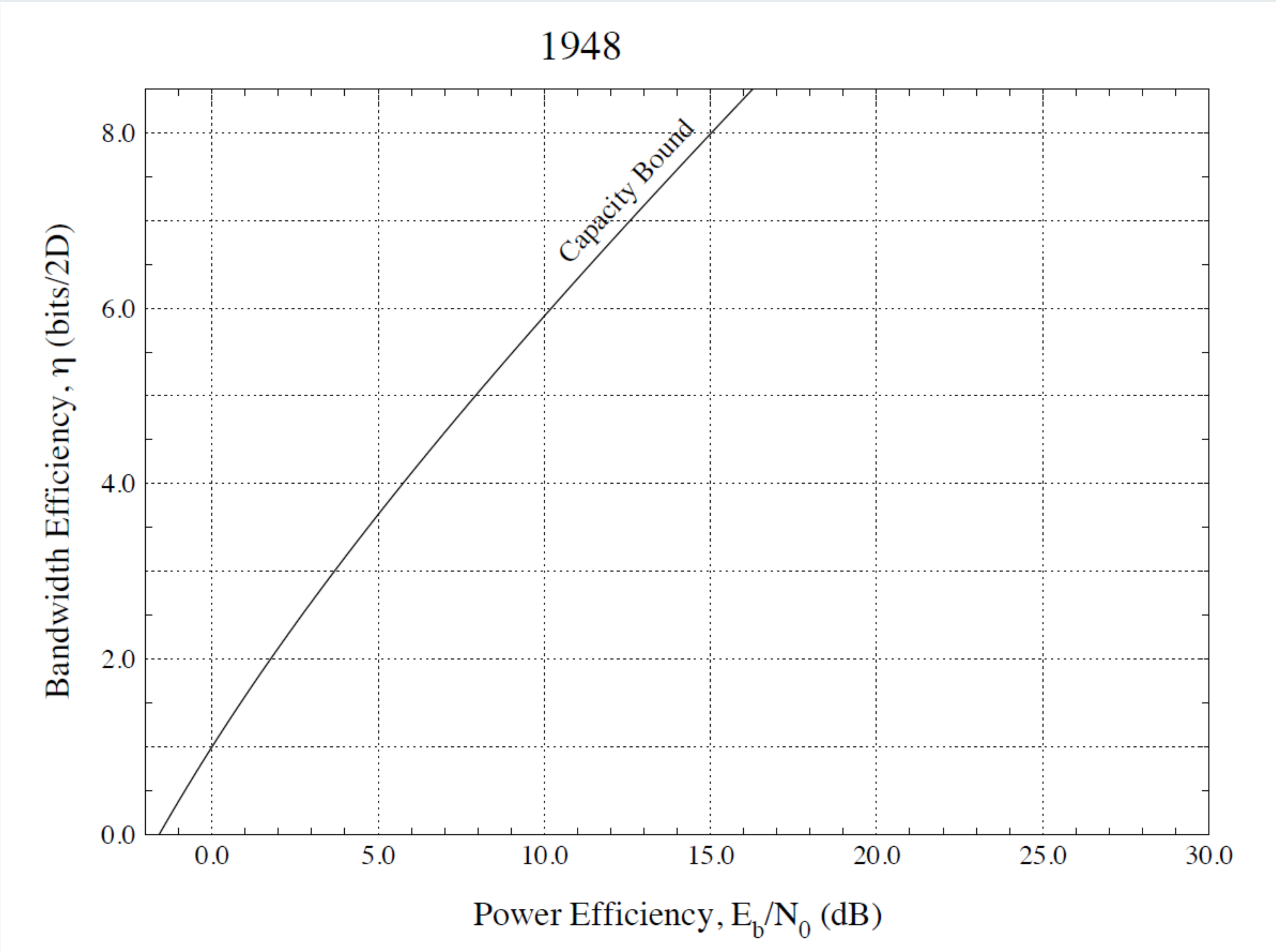
Excerpts from this post

Who is Claude Shannon Again?

- He worked on the top-secret transatlantic phone line connecting FDR and Winston Churchill during World War II and co-built what was arguably the world's first wearable computer. He learned to fly airplanes and played the jazz clarinet. He rigged up a false wall in his house that could rotate with the press of a button, and he once built a gadget whose only purpose when it was turned on was to open up, release a mechanical hand, and turn itself off. Oh, and he once had a photo spread in *Vogue* magazine.

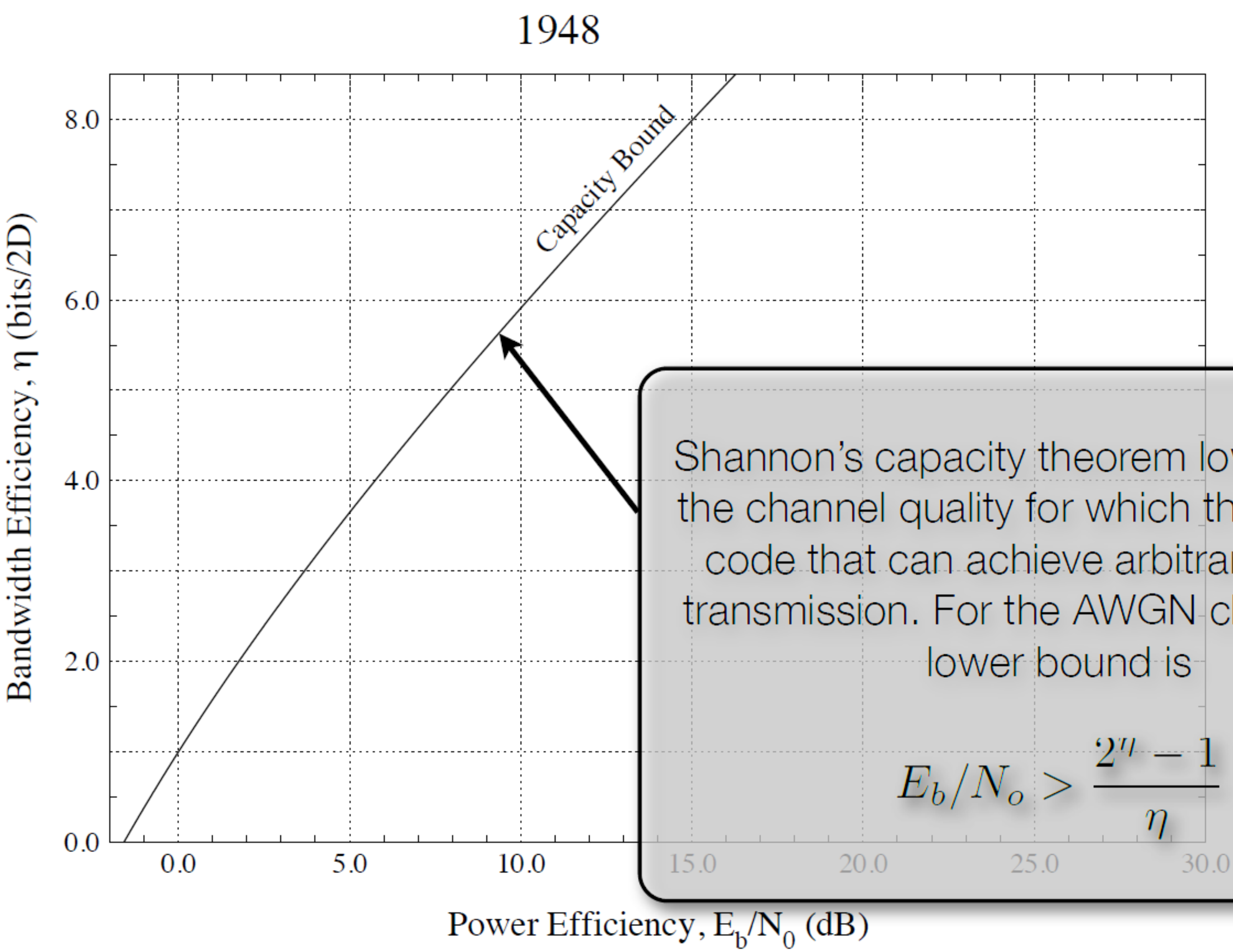
Excerpts from this post

What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?



@ D. Costello

What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?

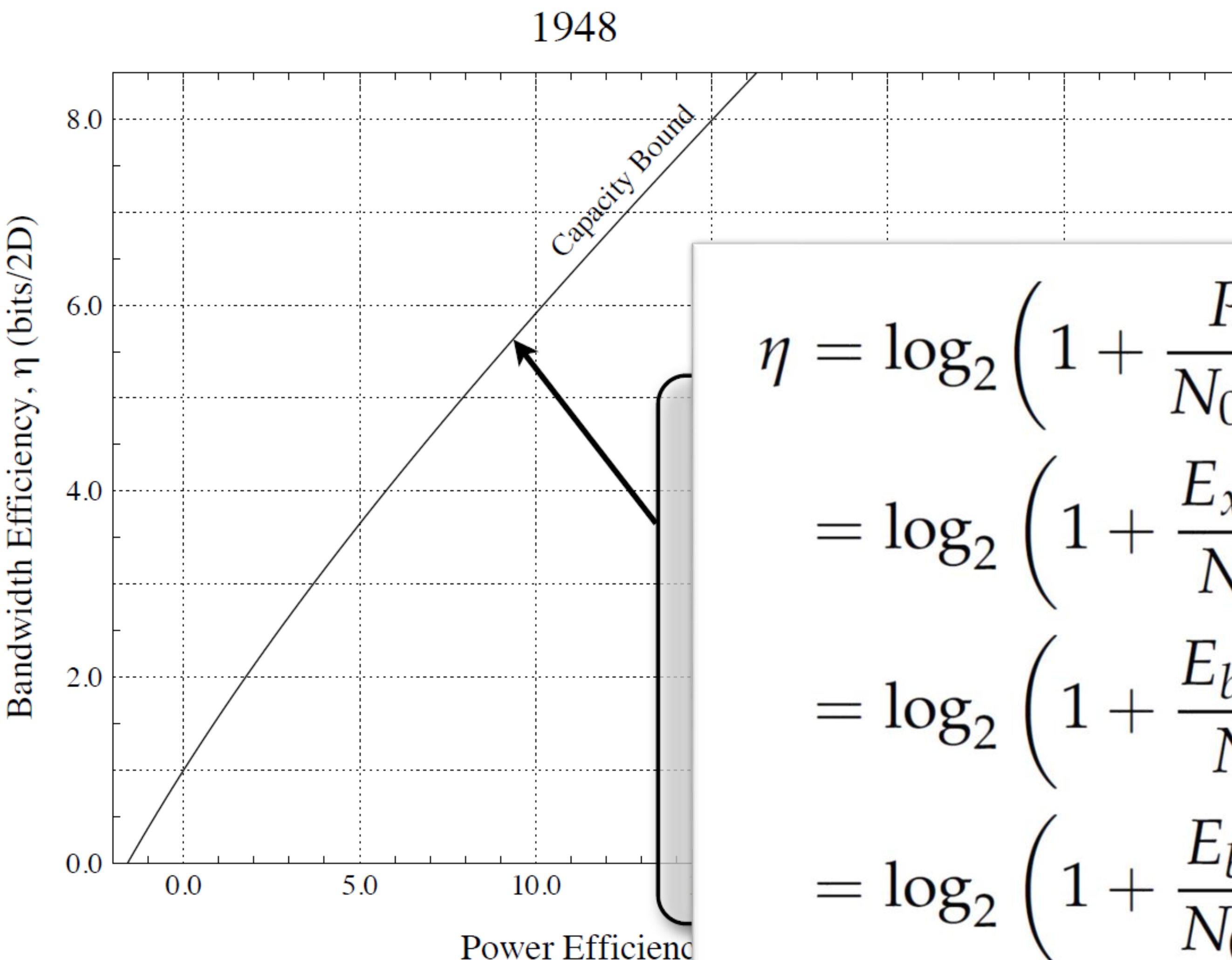


Shannon's capacity theorem lower bounds the channel quality for which there exists a code that can achieve arbitrarily reliable transmission. For the AWGN channel, this lower bound is

$$E_b/N_o > \frac{2^\eta - 1}{\eta}$$

@ D. Costello

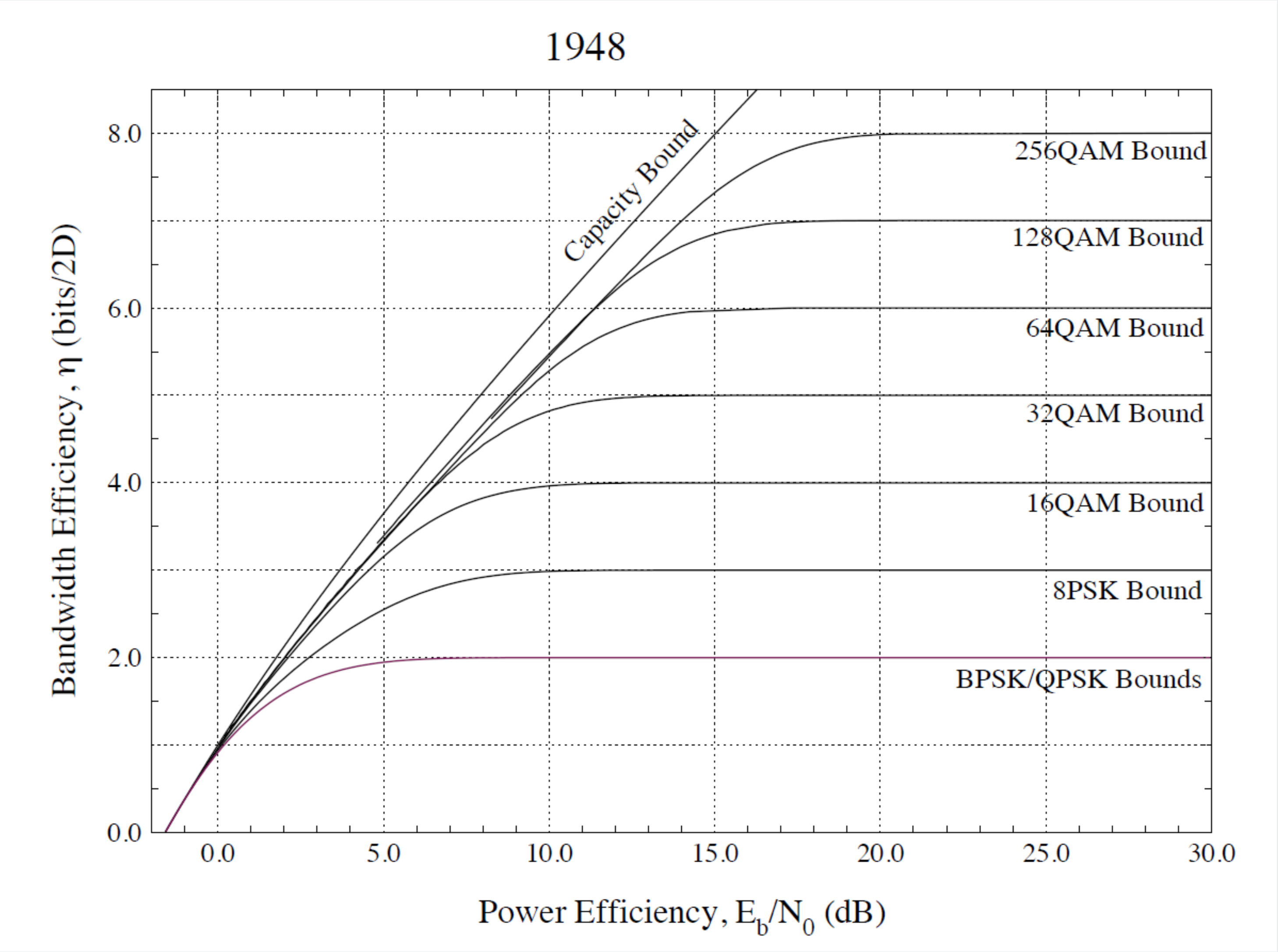
What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?



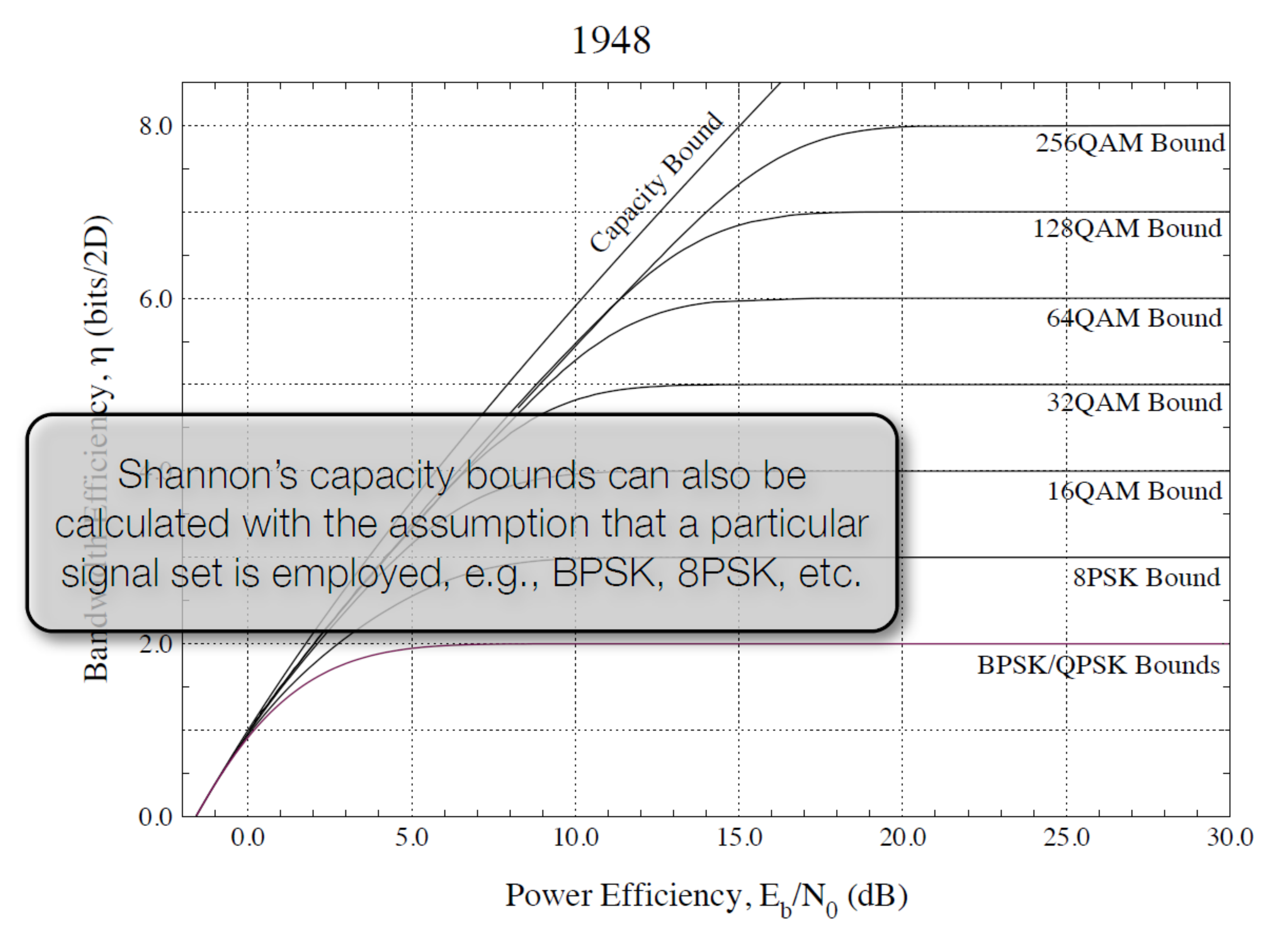
$$
\begin{aligned}
\eta &= \log_2\left(1 + \frac{P}{N_0 B}\right) \\
&= \log_2\left(1 + \frac{E_x/T}{N_0 B}\right) \\
&= \log_2\left(1 + \frac{E_b b/T}{N_0 B}\right) \\
&= \log_2\left(1 + \frac{E_b}{N_0}\eta\right)
\end{aligned}
$$

@ D. Costello

What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?

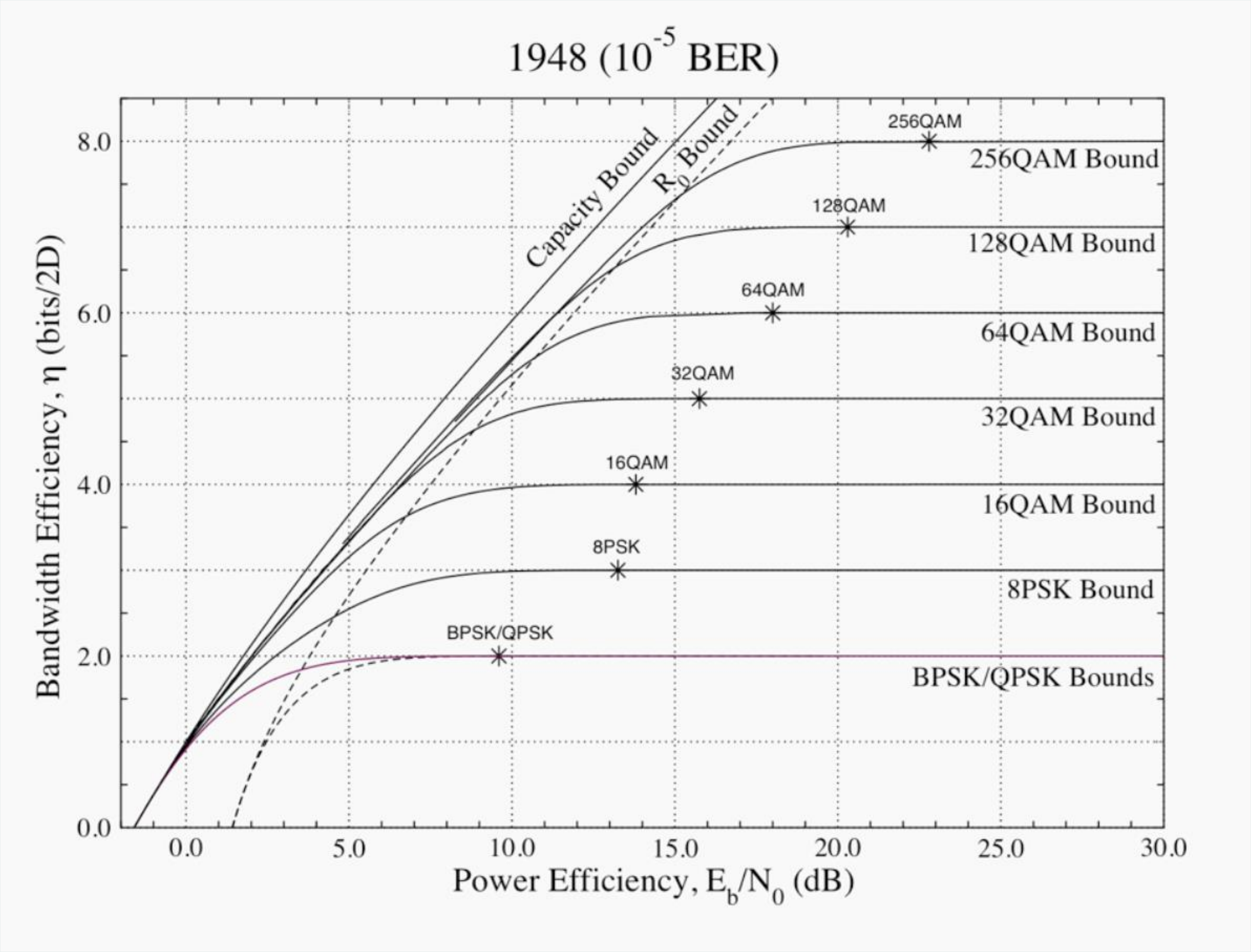


What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?

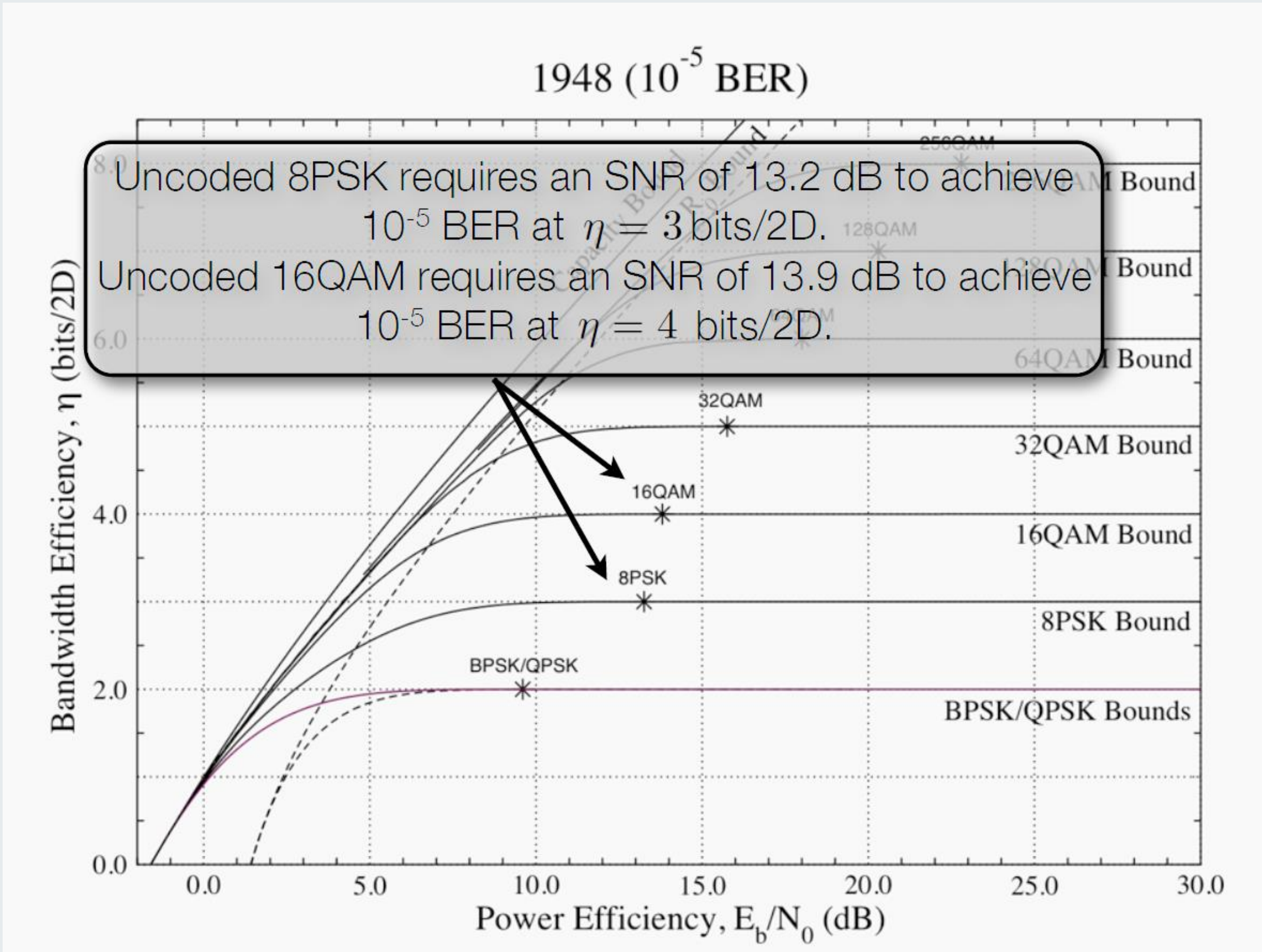


Shannon's capacity bounds can also be calculated with the assumption that a particular signal set is employed, e.g., BPSK, 8PSK, etc.

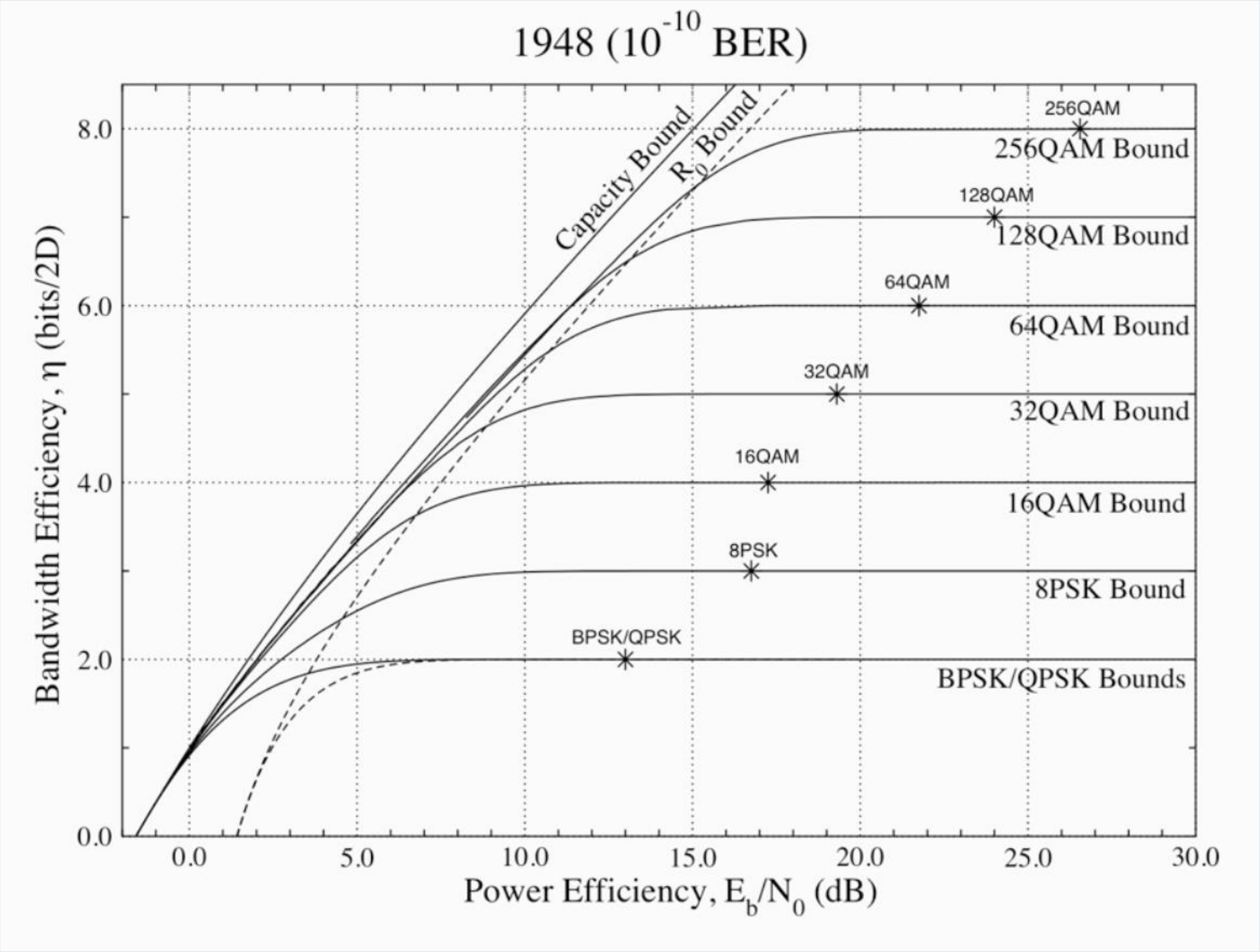
What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?



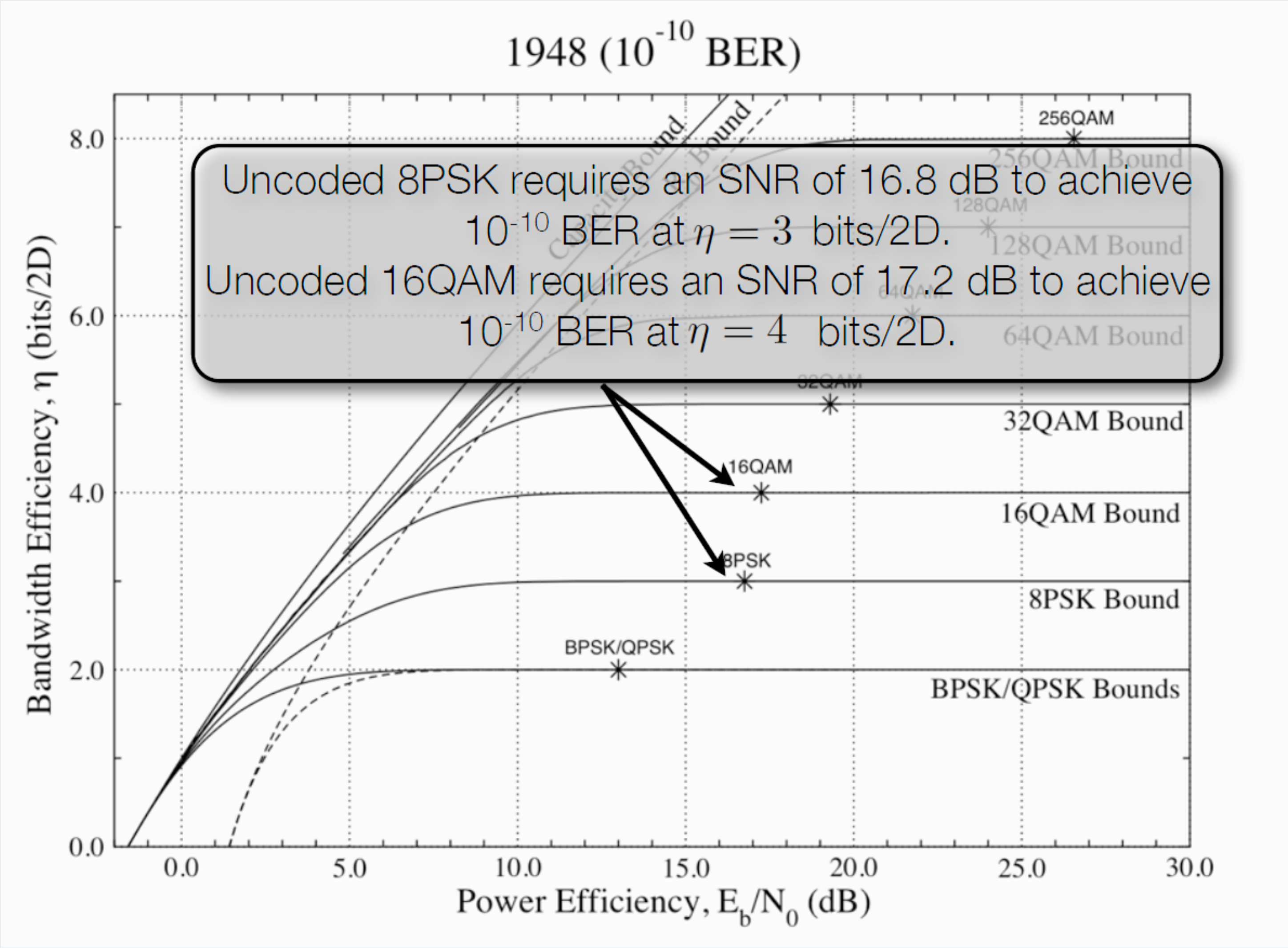
What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?



What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?



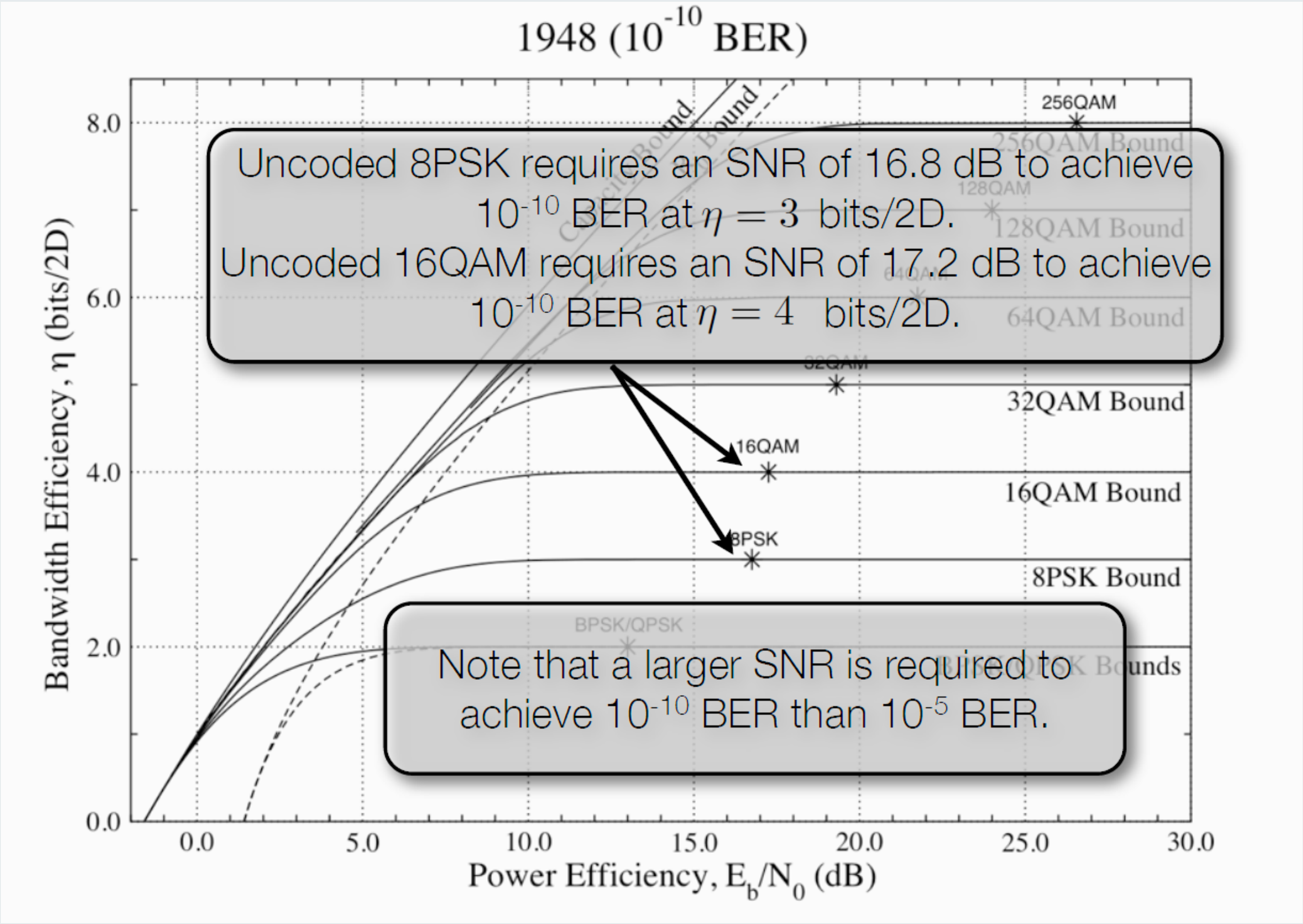
What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?



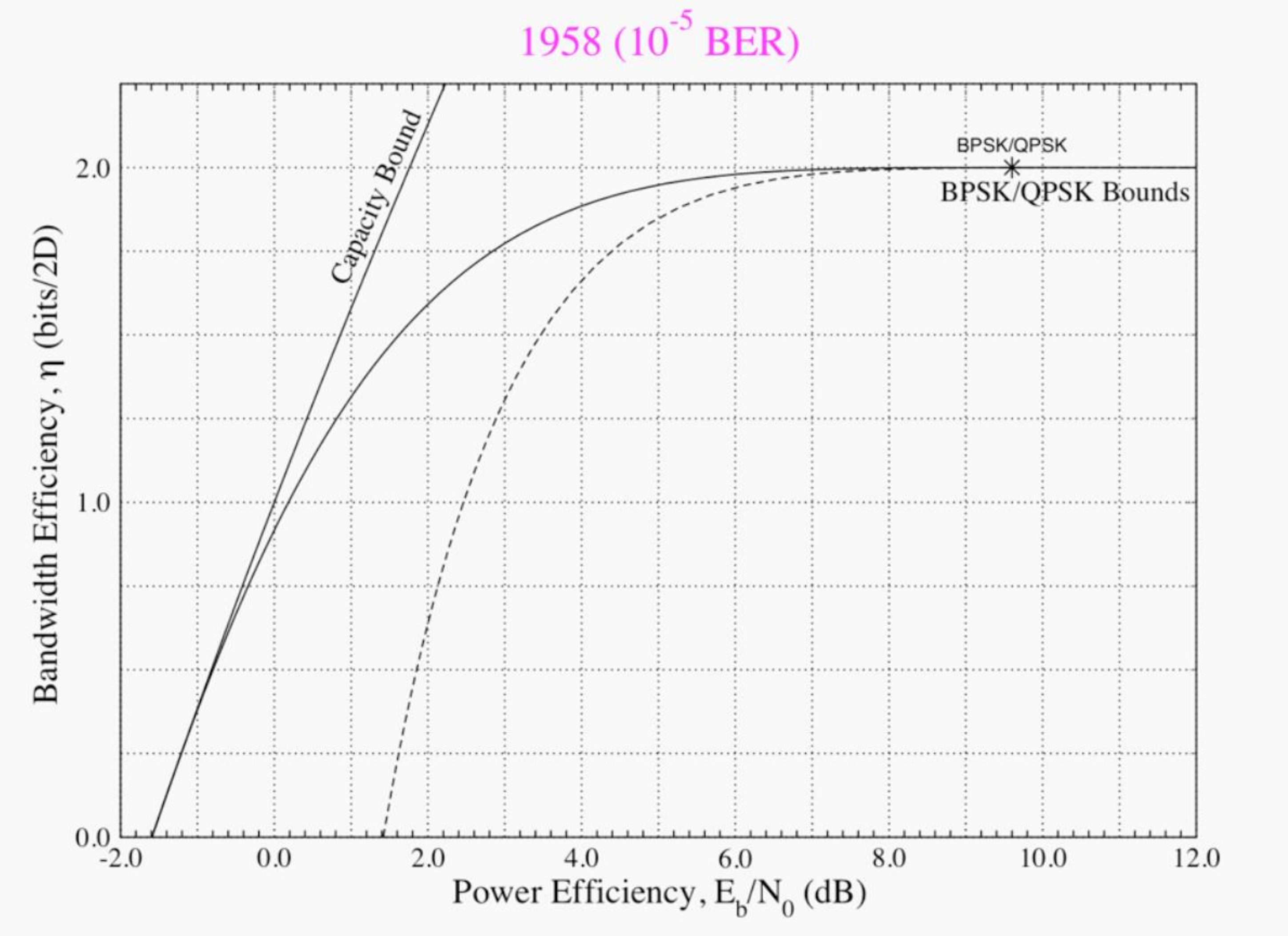
Uncoded 8PSK requires an SNR of 16.8 dB to achieve 10^{-10} BER at $\eta = 3$ bits/2D.

Uncoded 16QAM requires an SNR of 17.2 dB to achieve 10^{-10} BER at $\eta = 4$ bits/2D.

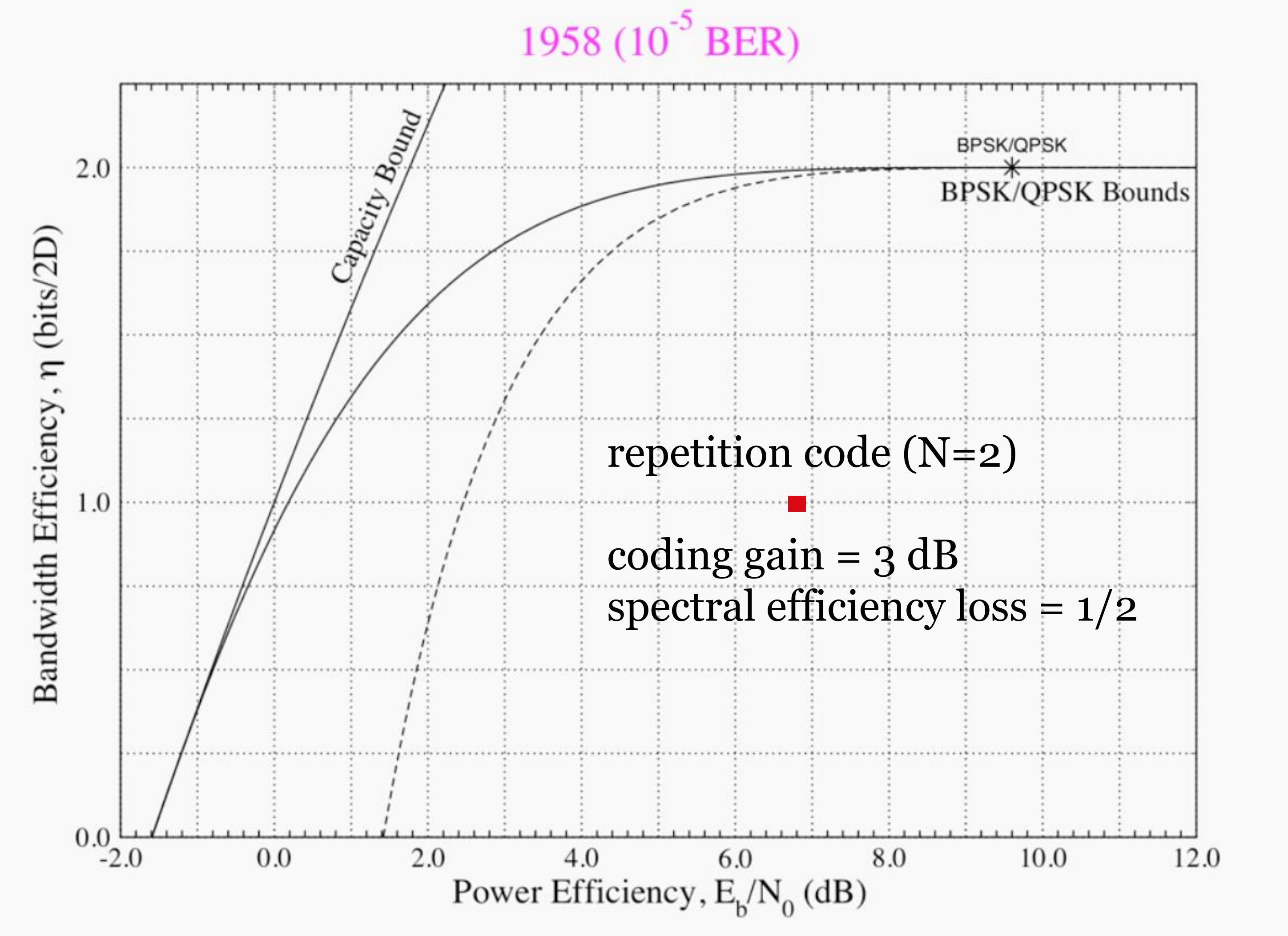
What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?



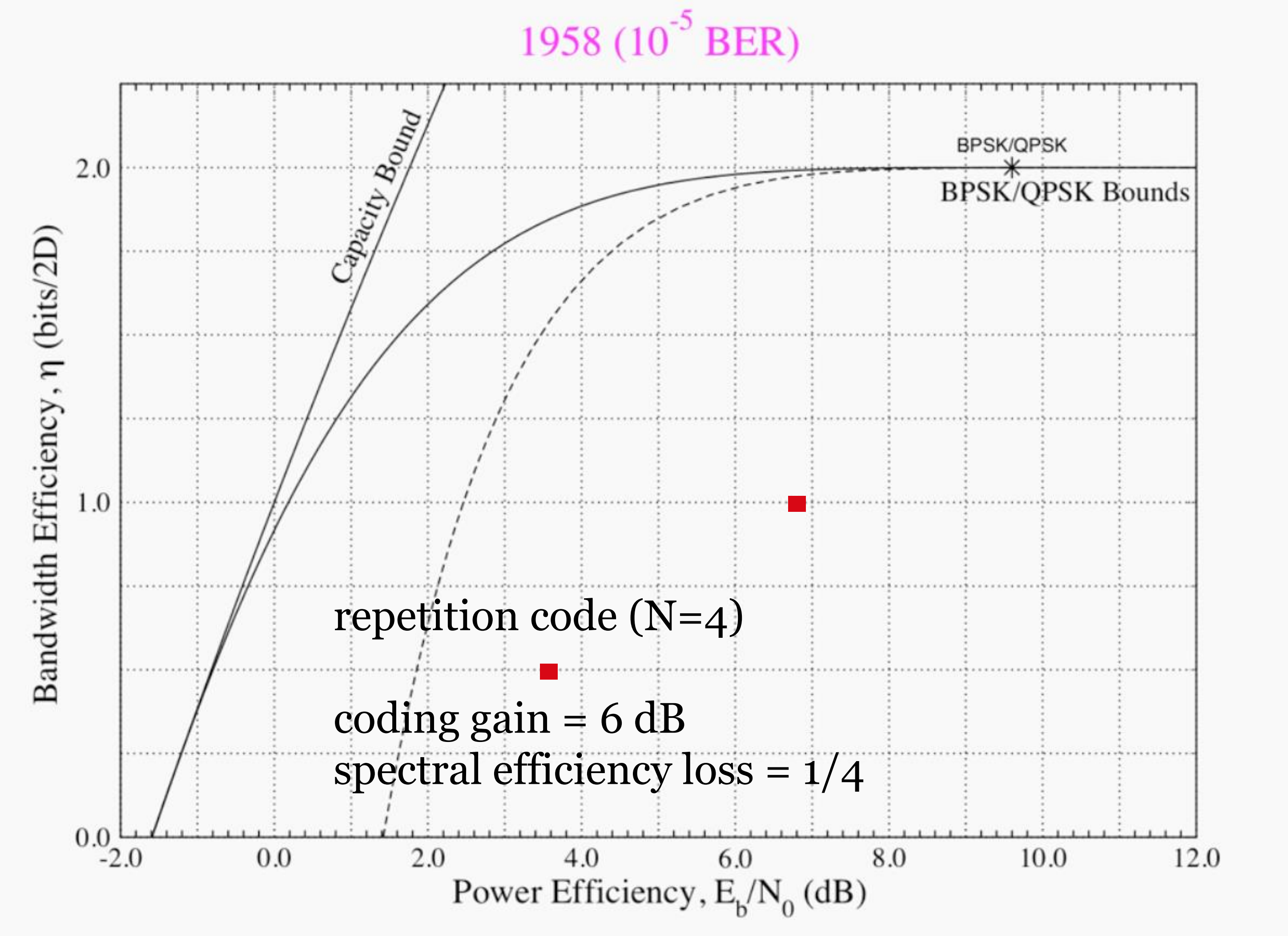
What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?



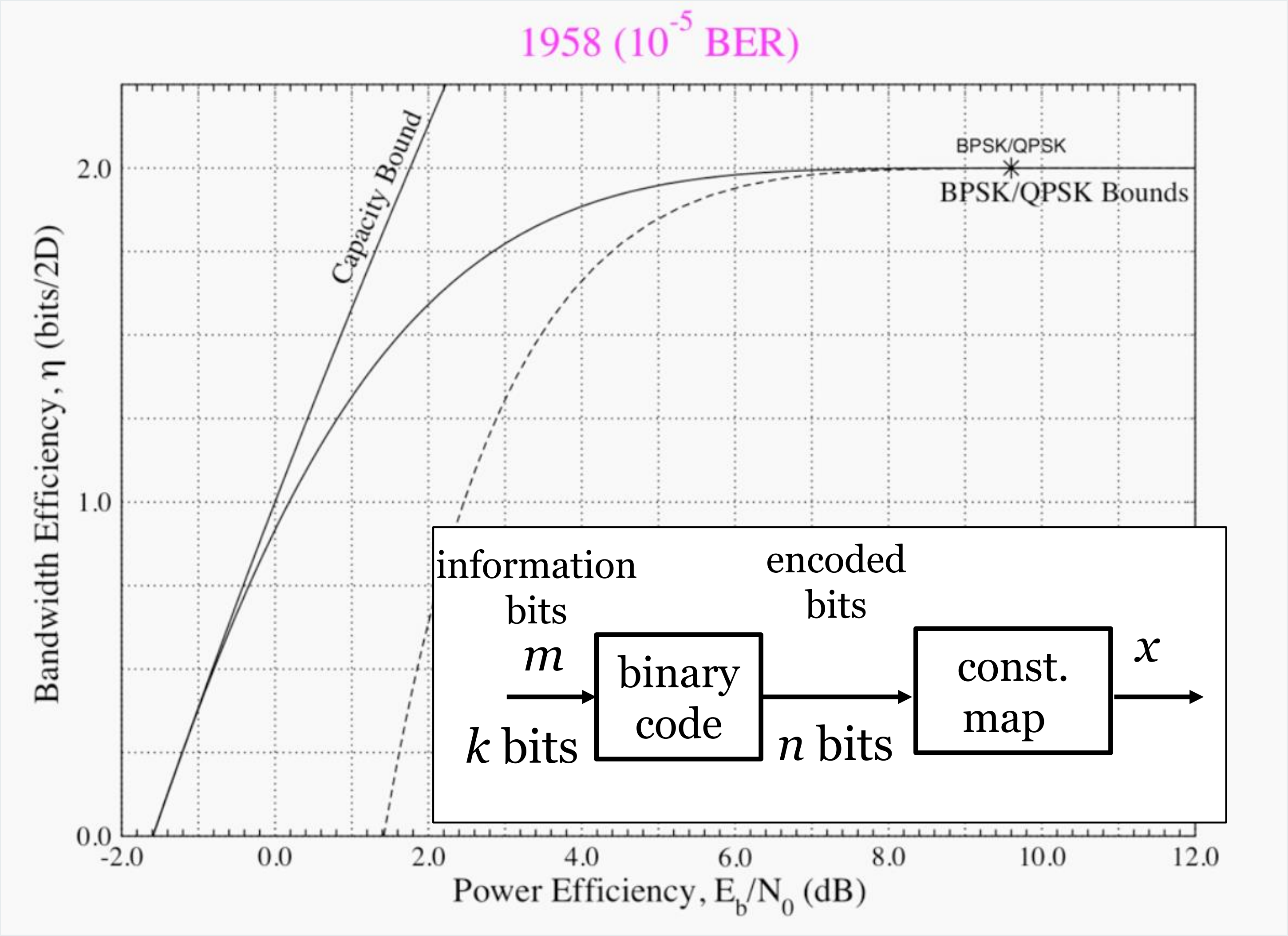
What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?



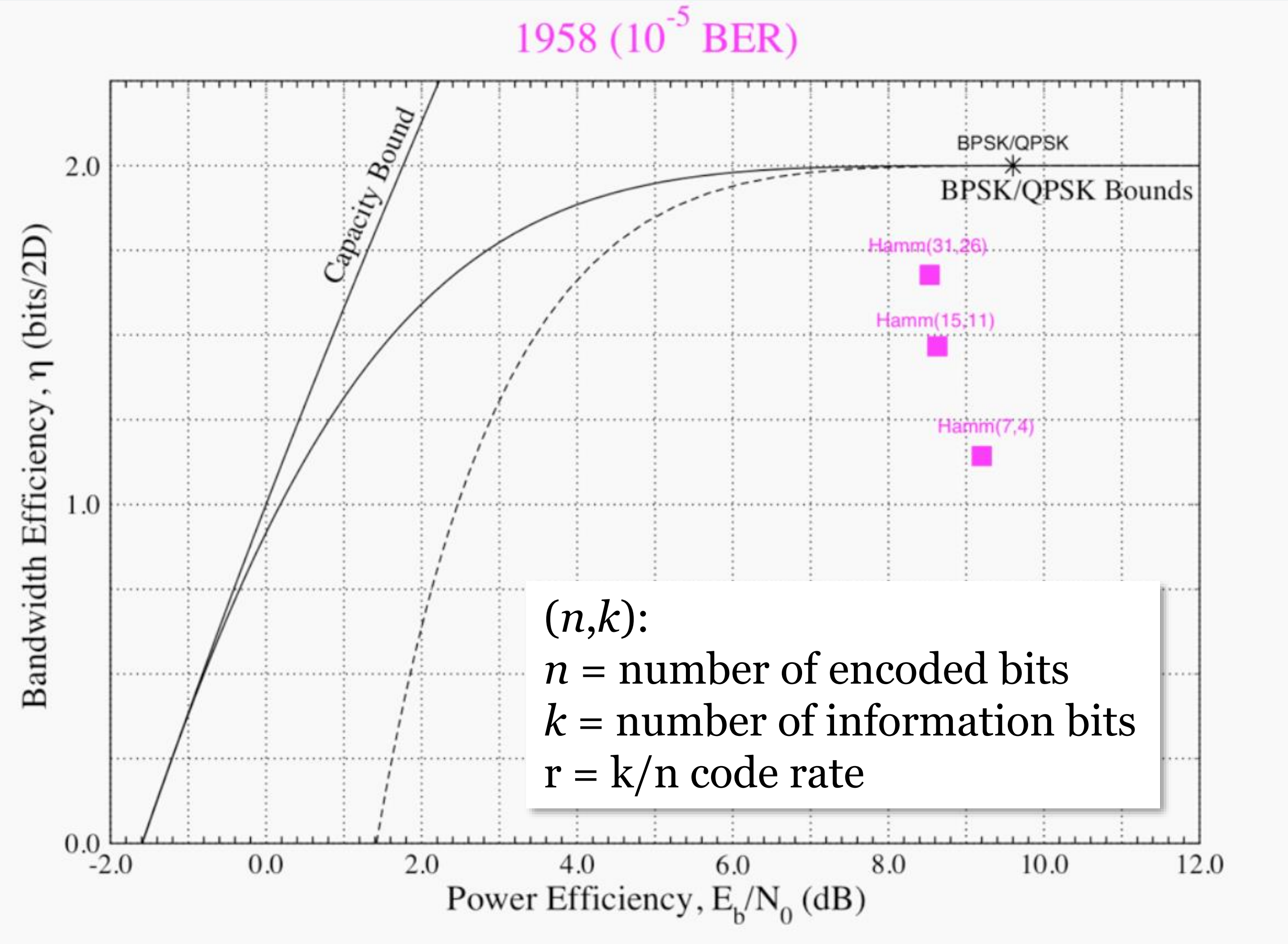
What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?



What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?

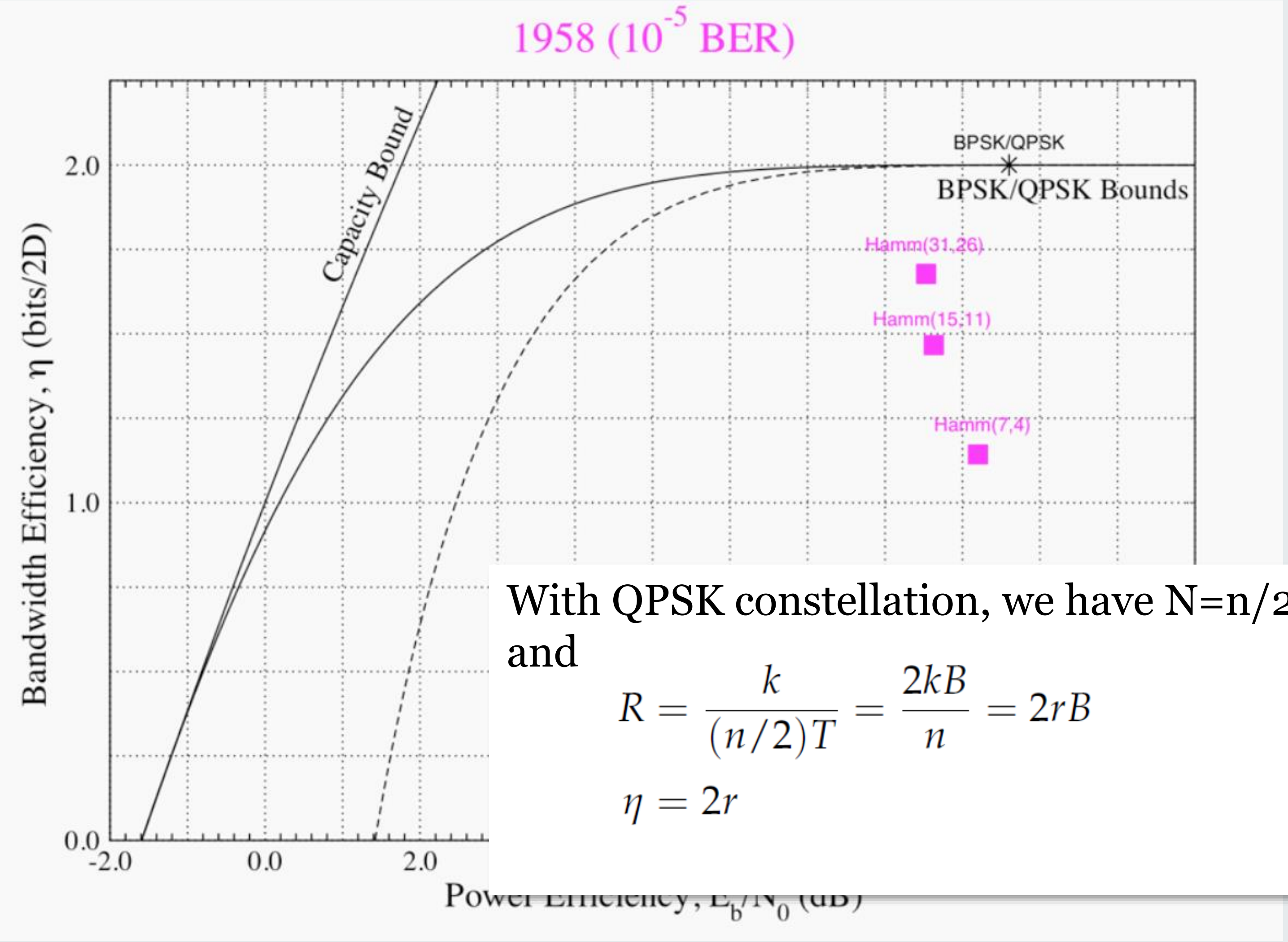


What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?



(n,k):
n = number of encoded bits
k = number of information bits
r = k/n code rate

What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?

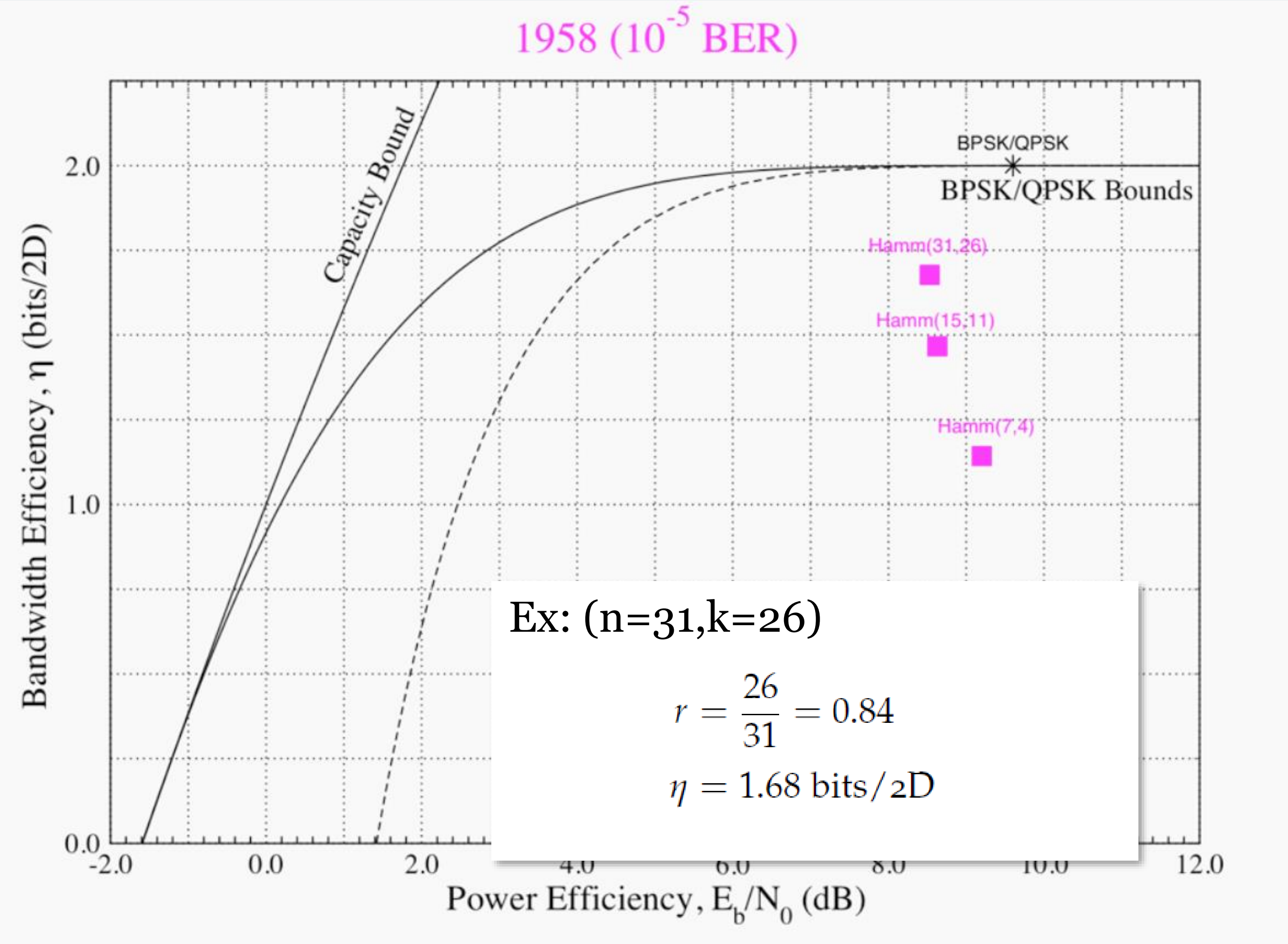


With QPSK constellation, we have N=n/2 and

$$R = \frac{k}{(n/2)T} = \frac{2kB}{n} = 2rB$$

$$\eta = 2r$$

What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?

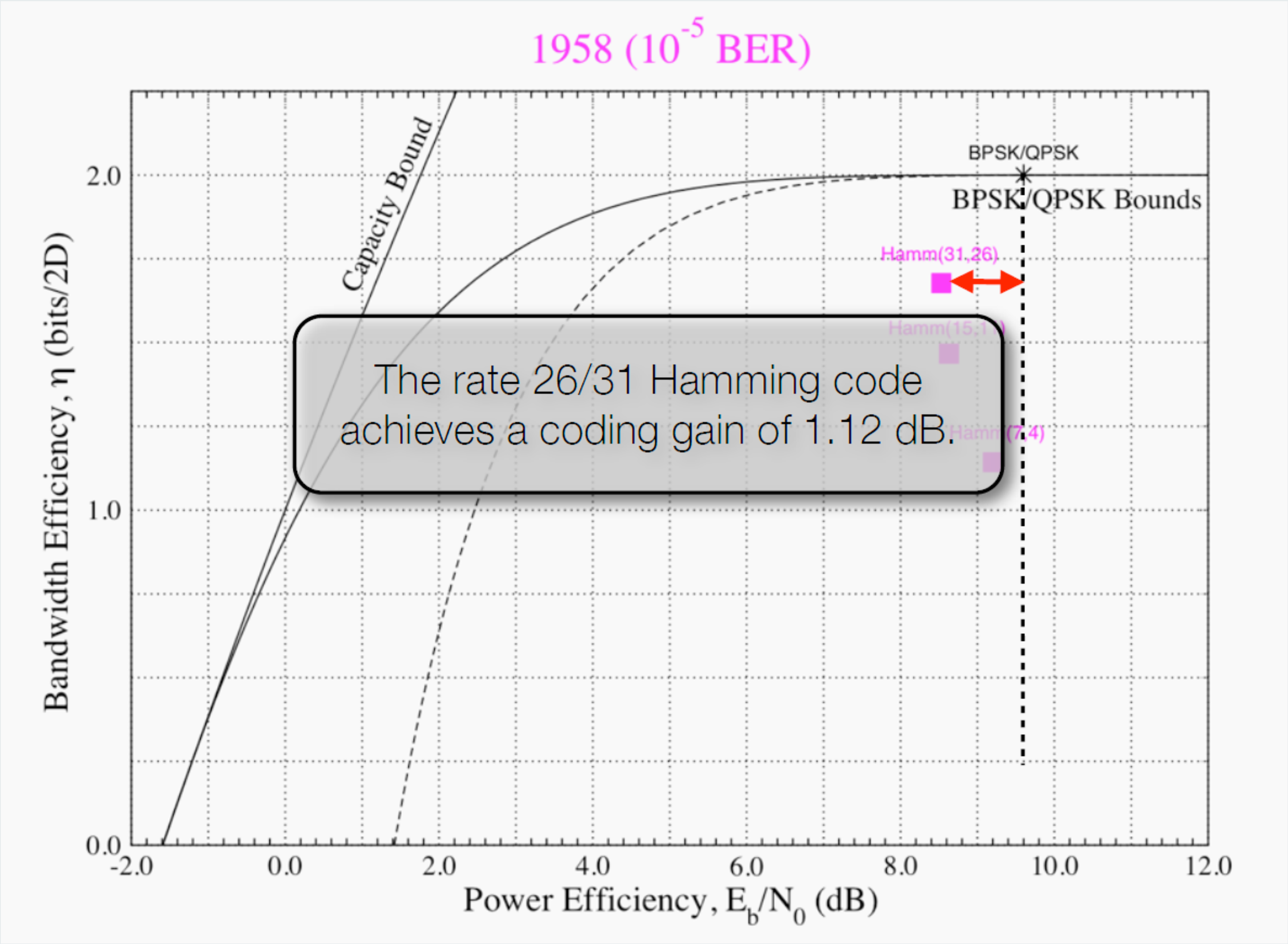


Ex: (n=31,k=26)

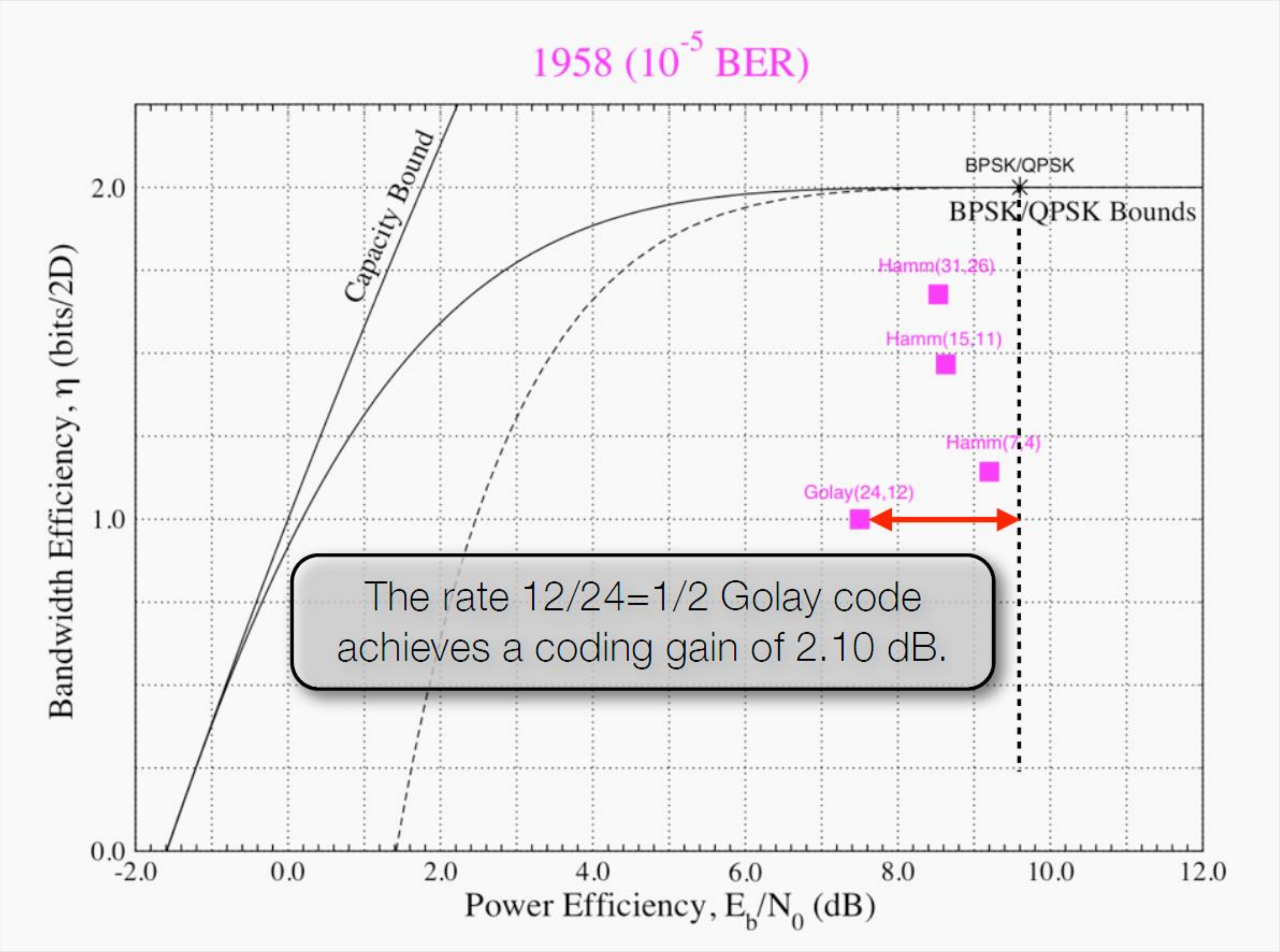
$$r = \frac{26}{31} = 0.84$$

$$\eta = 1.68 \text{ bits}/2\text{D}$$

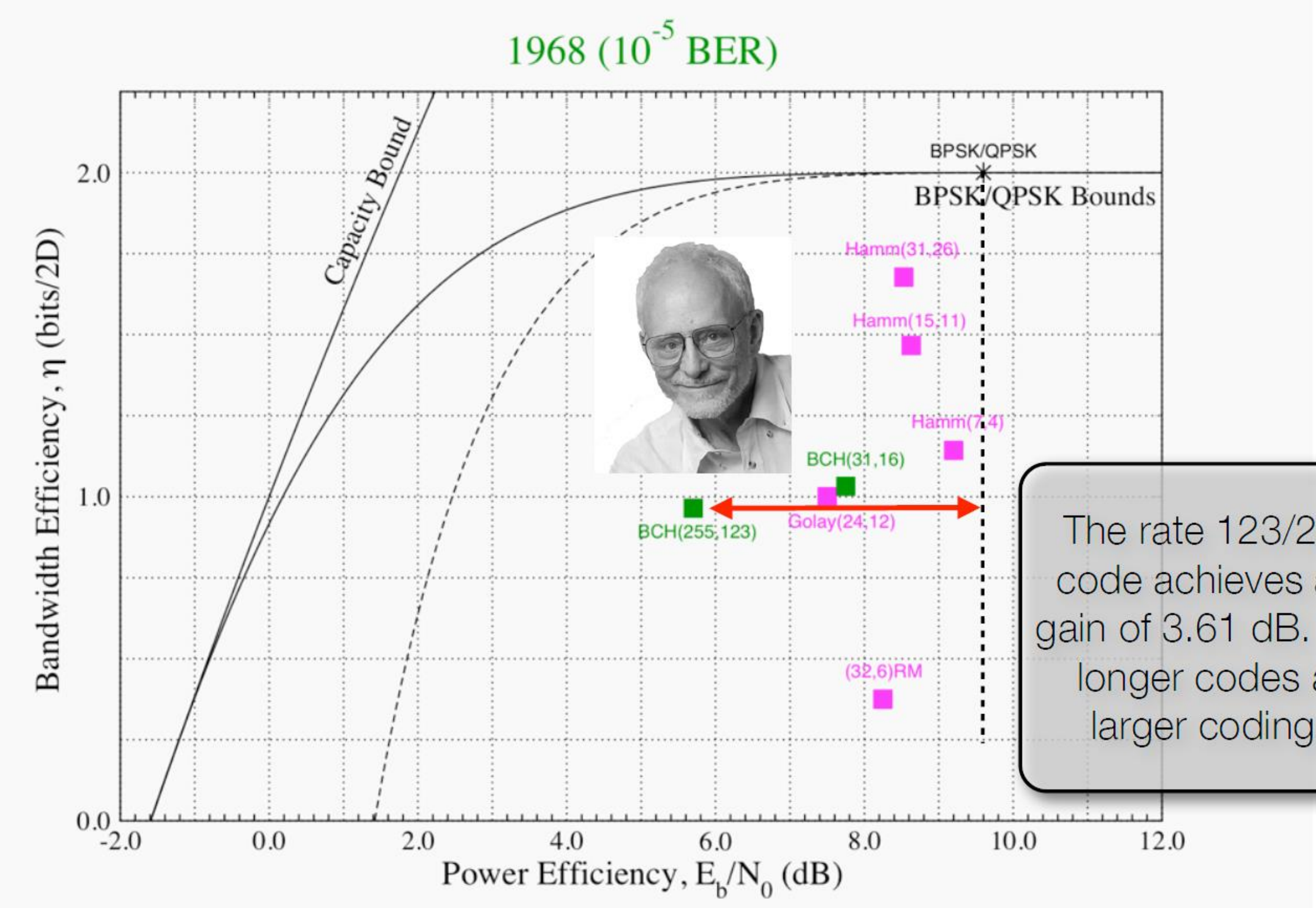
What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?



What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?

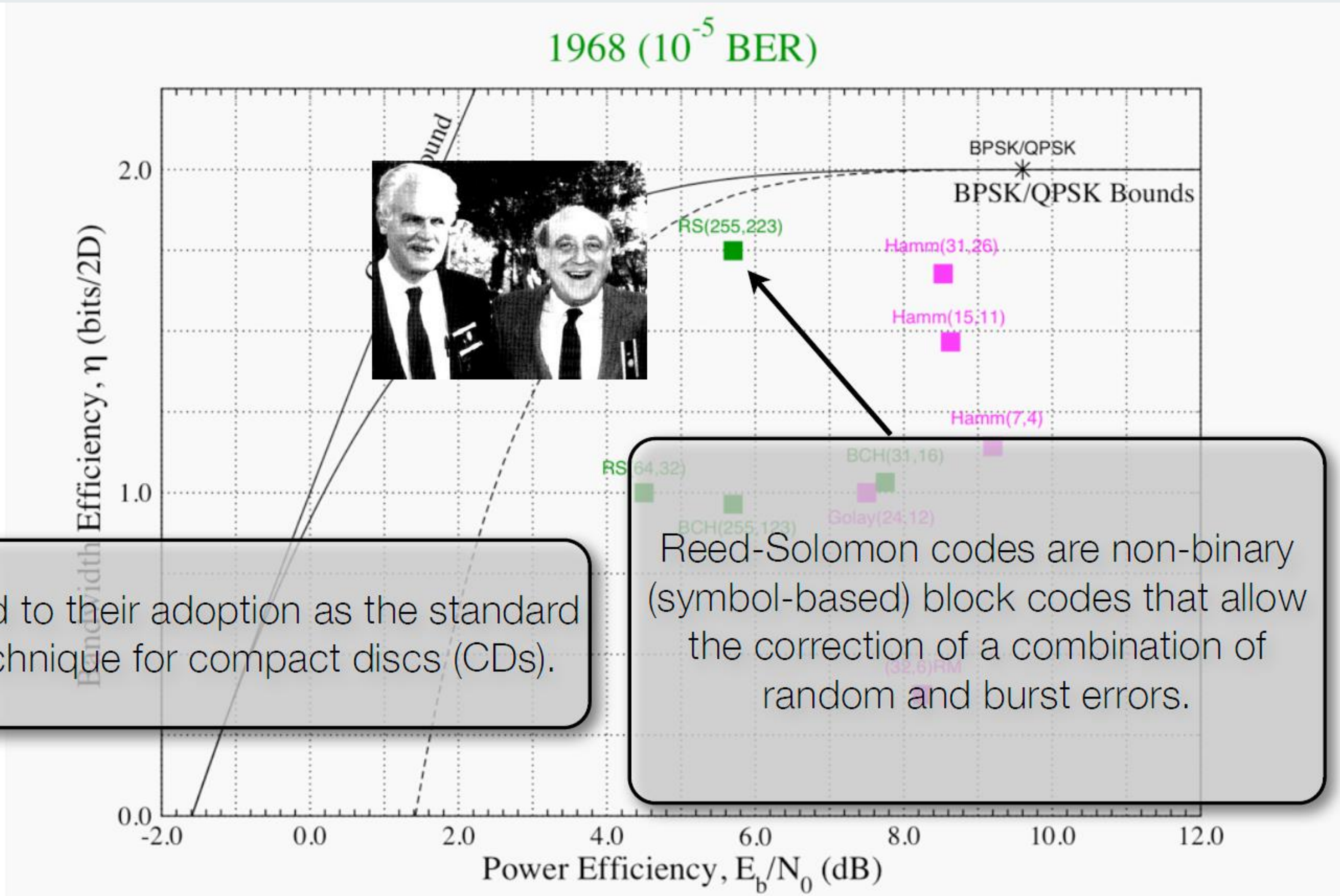


What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?



The rate 123/255 BCH code achieves a coding gain of 3.61 dB. Note that longer codes achieve larger coding gains.

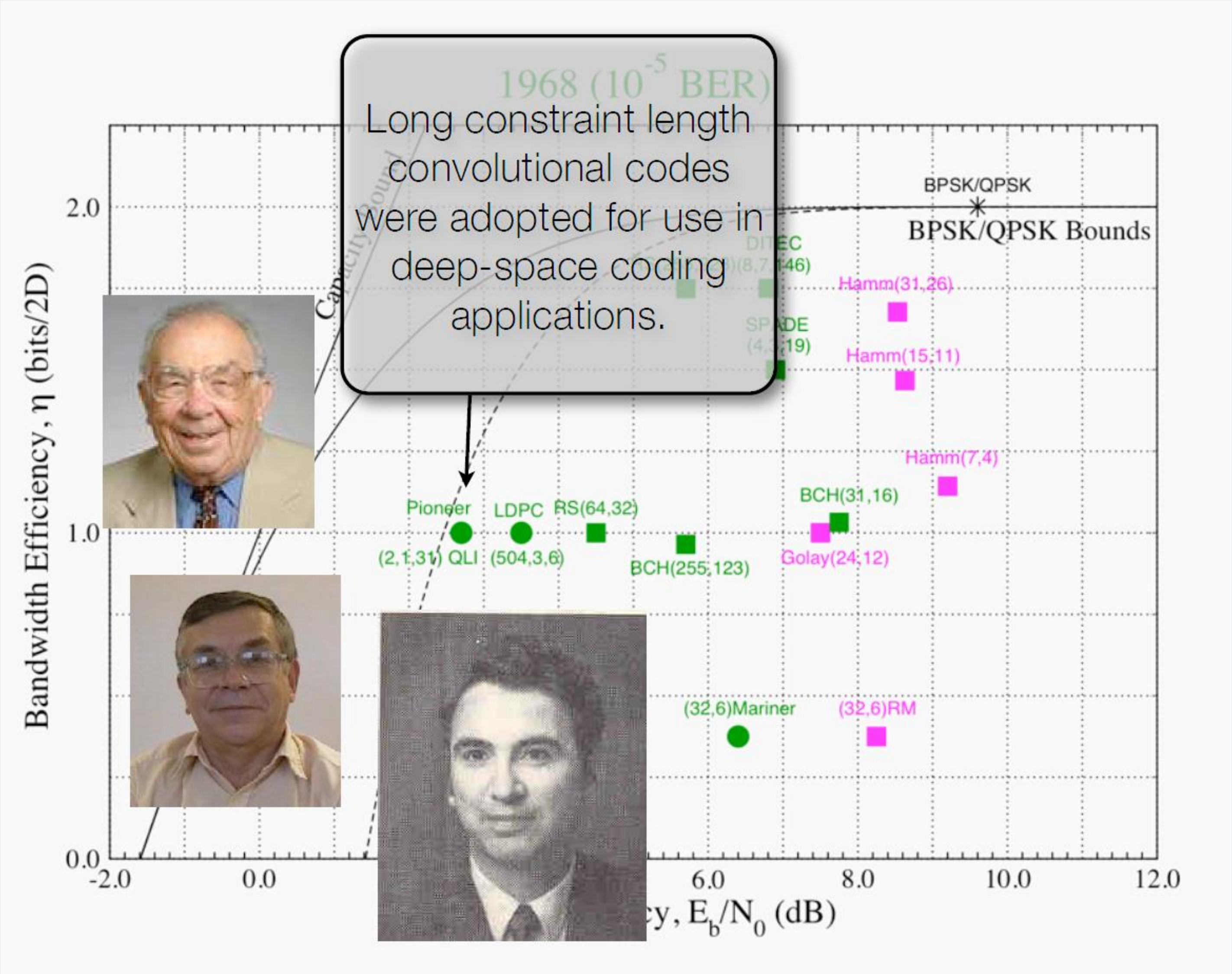
What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?



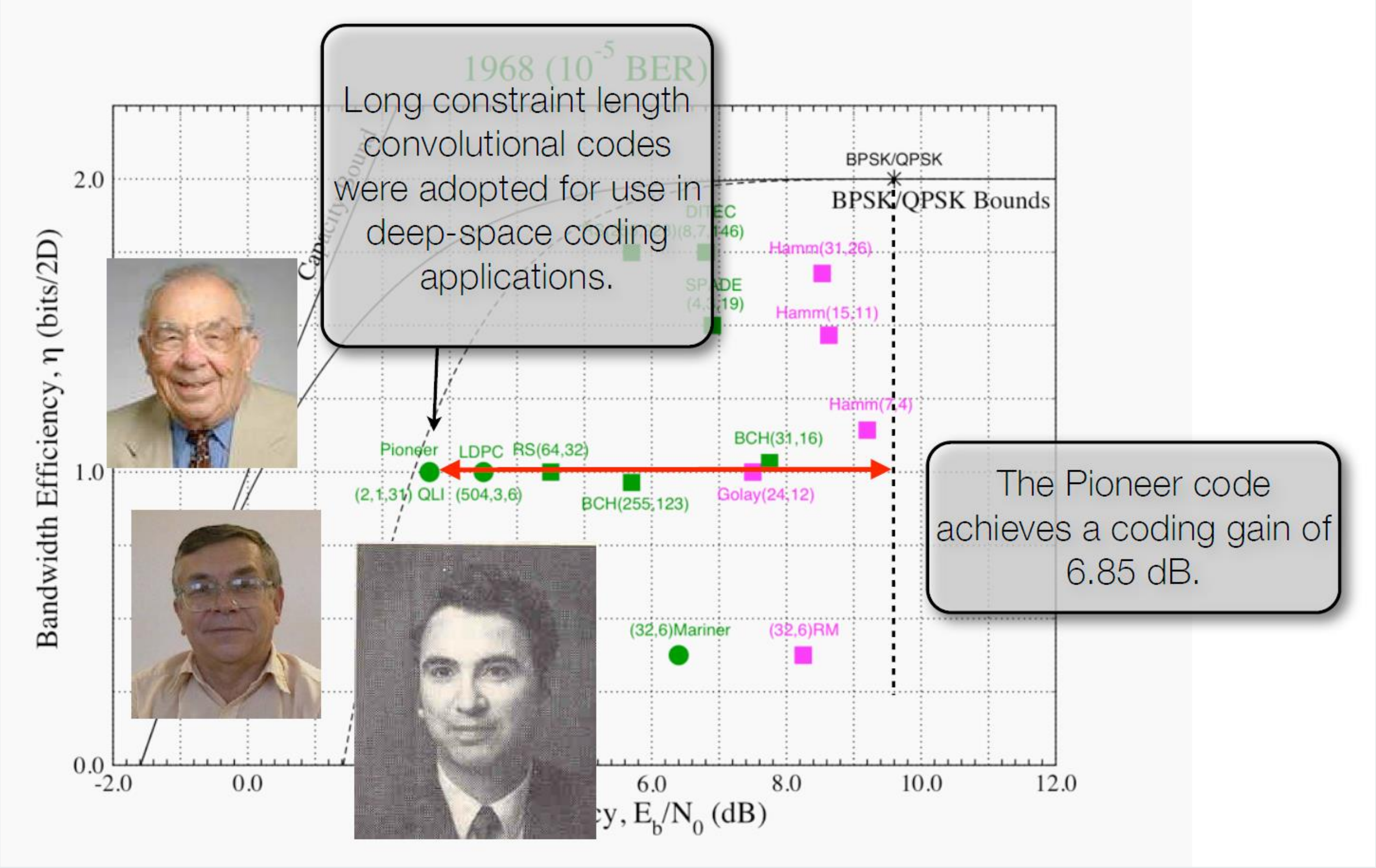
This fact led to their adoption as the standard coding technique for compact discs (CDs).

Reed-Solomon codes are non-binary (symbol-based) block codes that allow the correction of a combination of random and burst errors.

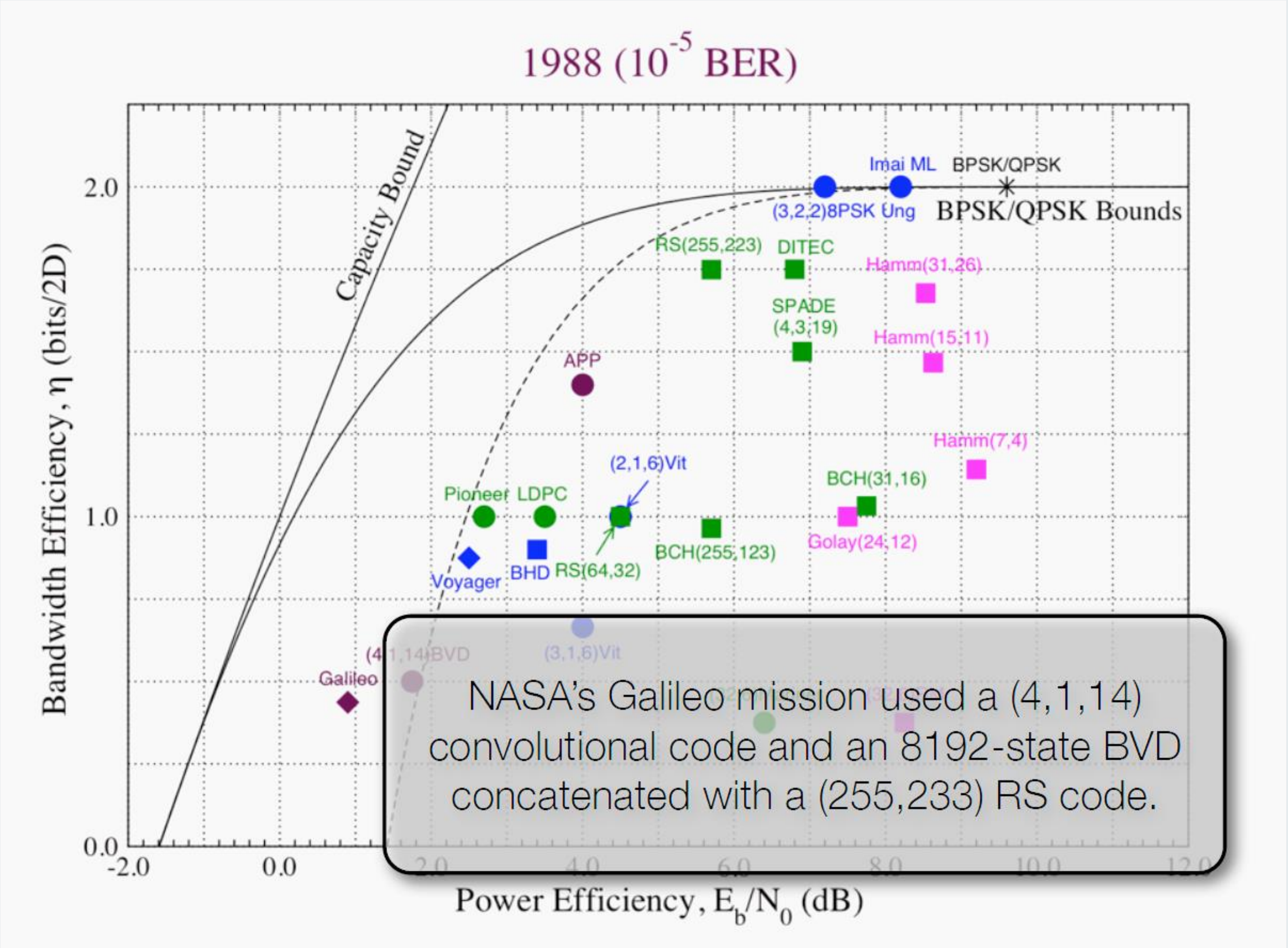
What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?



What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?

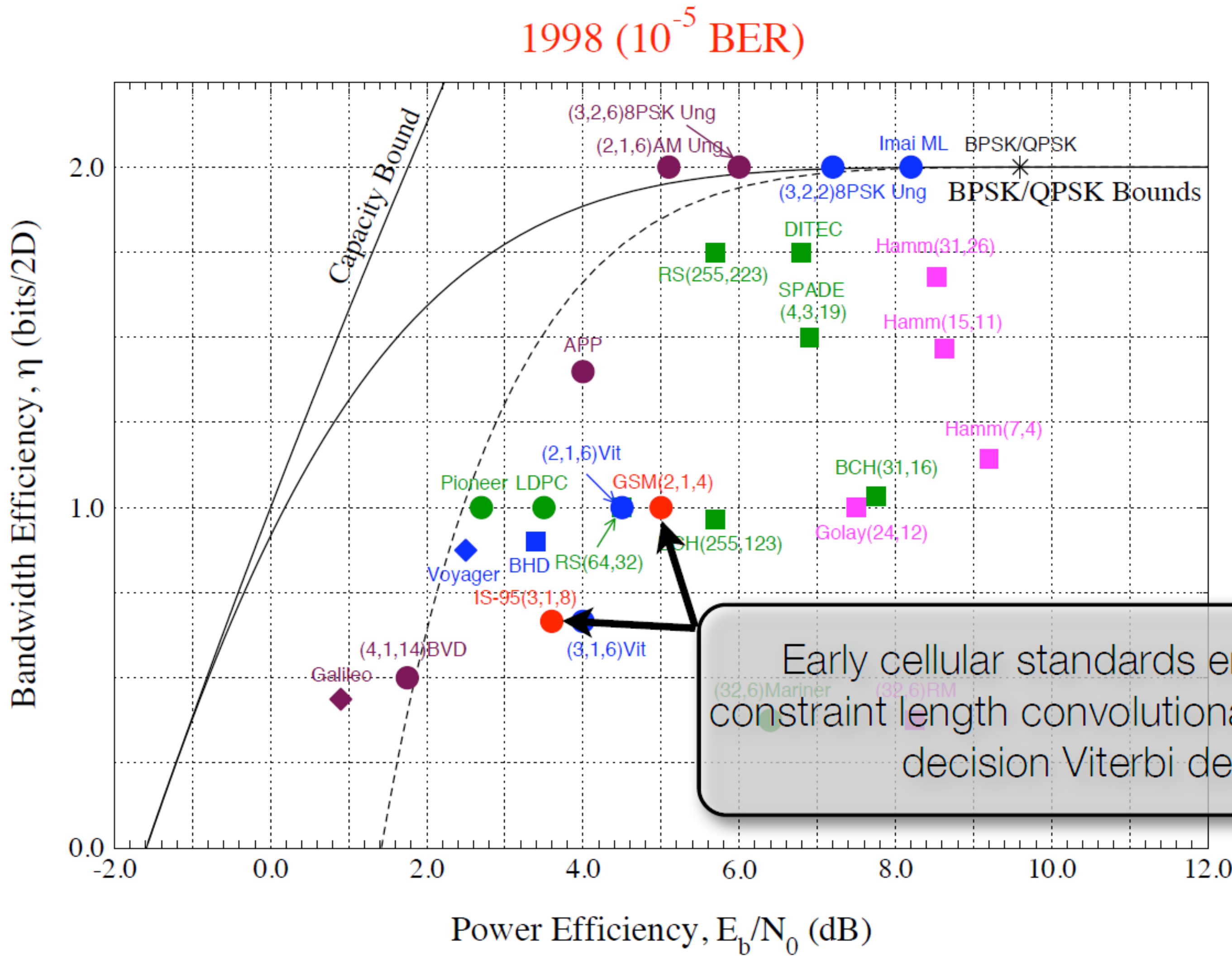


What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?



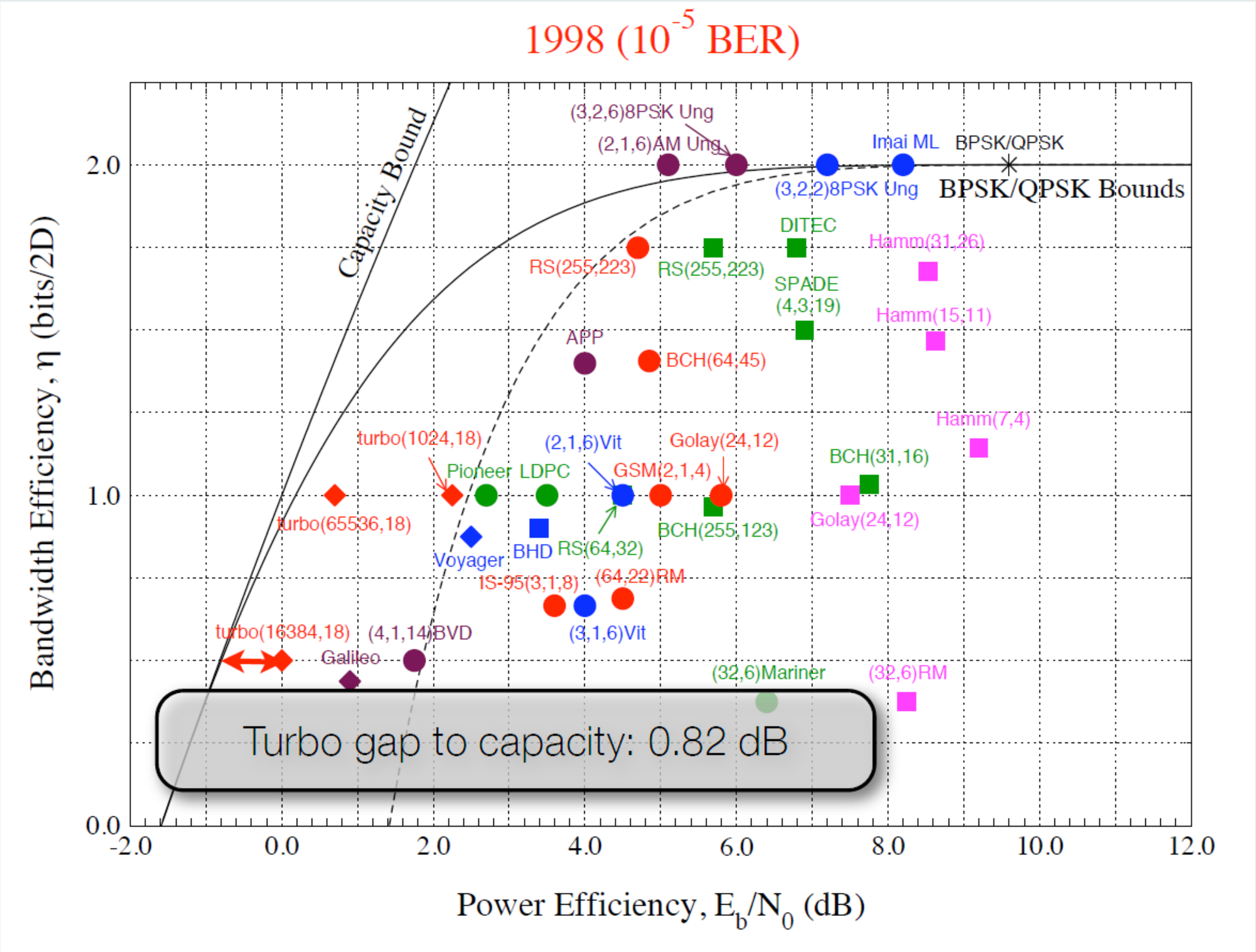
NASA's Galileo mission used a (4,1,14) convolutional code and an 8192-state BVD concatenated with a (255,233) RS code.

What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?

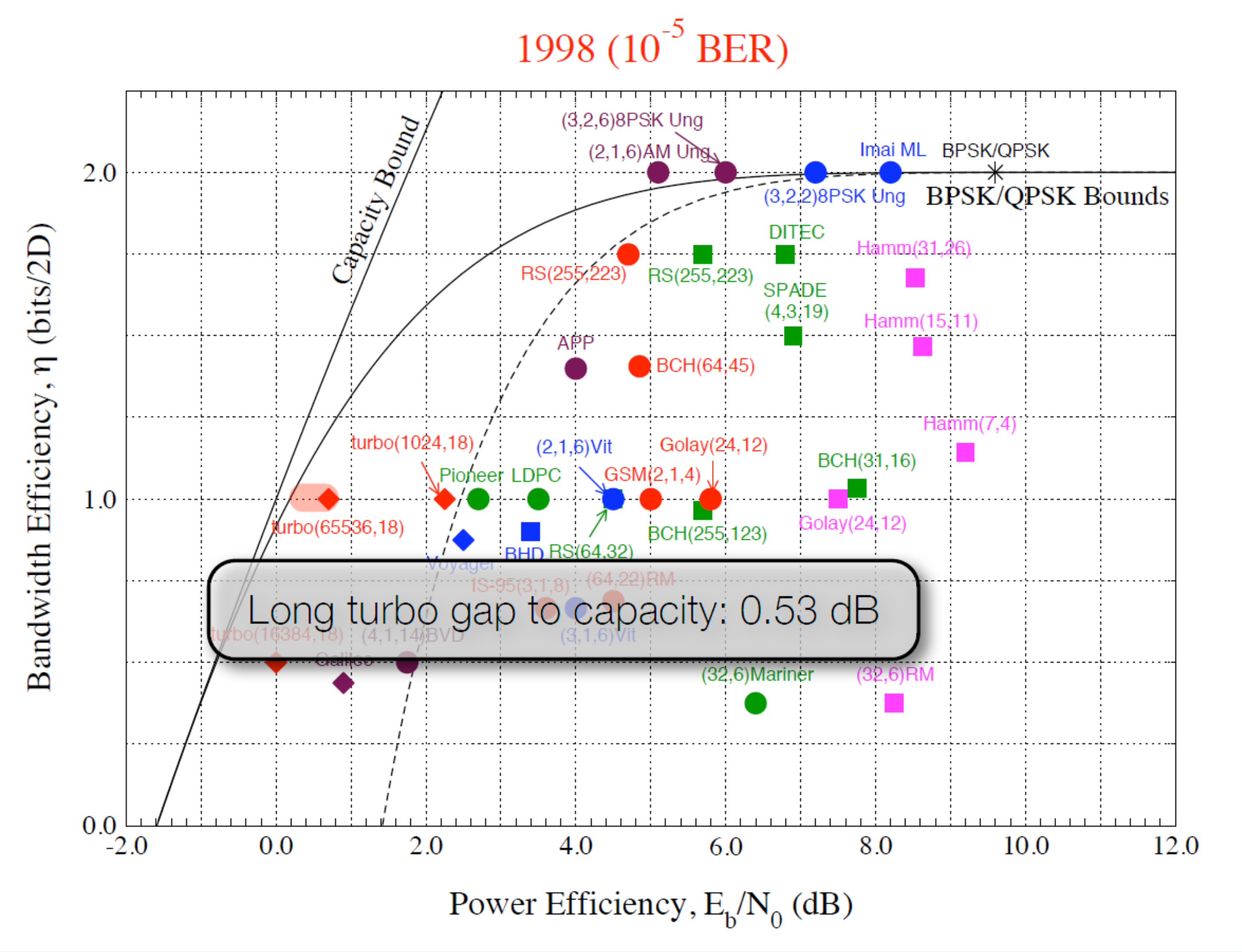


Early cellular standards employed short constraint length convolutional codes with soft-decision Viterbi decoding.

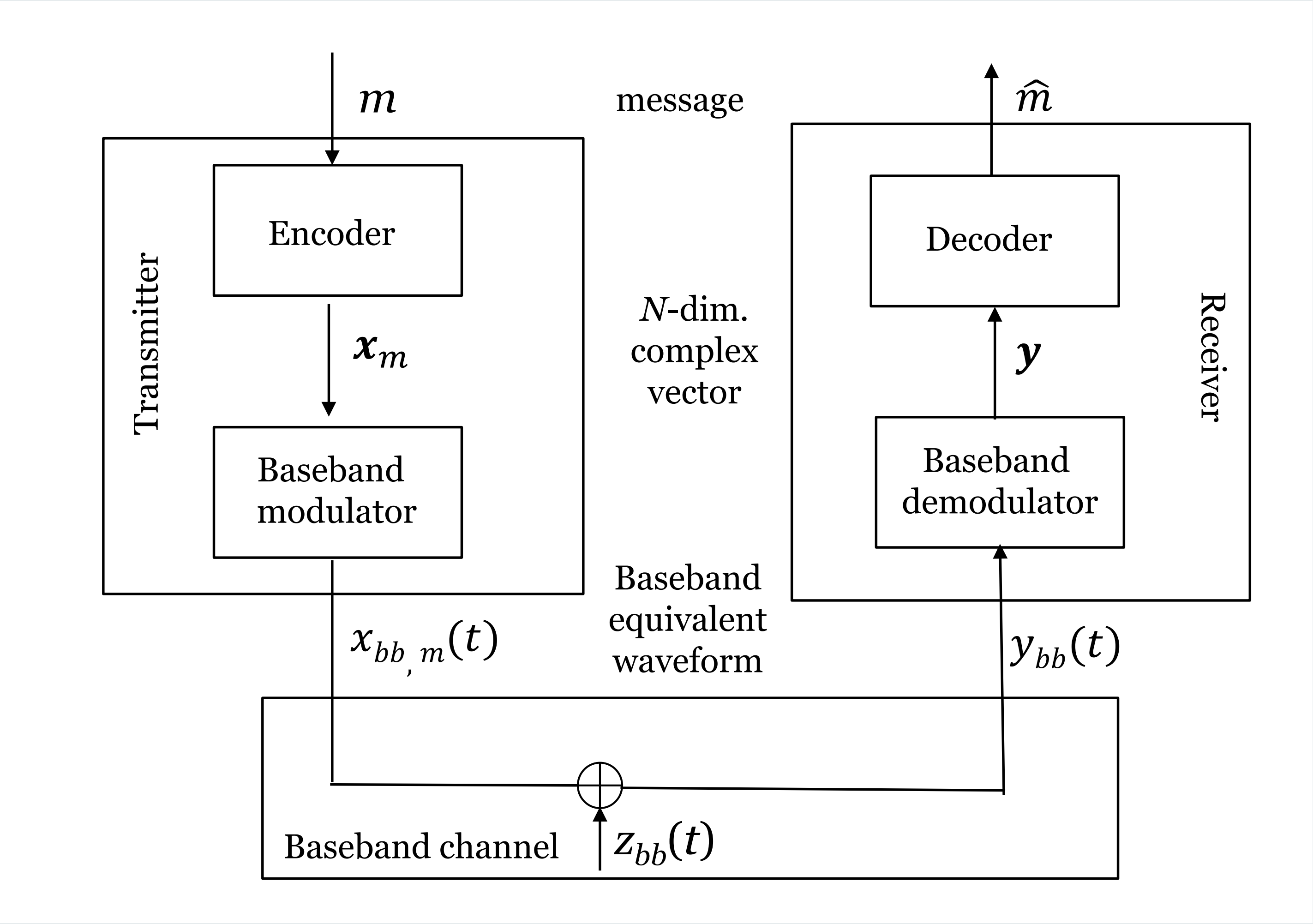
What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?



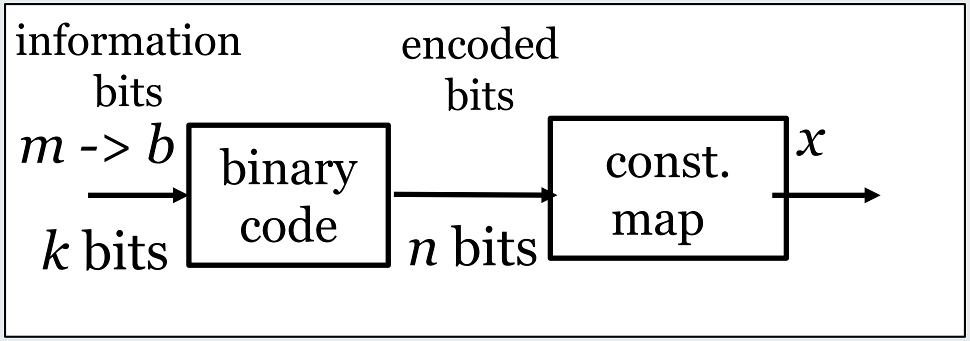
What is the Optimal Trade-Off between Energy and Bandwidth Efficiencies?



How to Encode over Multiple Dimensions?



How Do Convolutional Codes Work?



- Assume BPSK, and encode the information bits m in a stream $b[1]$, $b[2]$, … where $b[i]=1$ if $m[i]=1$ and $b[i]=-1$ if $m[i]=0$.

- Example: $m=(m[1]=0, m[2]=1, m[3]=1, m[4]=0,\dots, m[k]=1)$ message
  $b= (b[1]=-1, b[2]=1, b[3]=1, b[4]=-1,\dots, b[k]=1)$ uncoded (signed) bits

- A convolutional encoder takes as input k information bits and outputs n encoded bits.

How Do Convolutional Codes Work?

- The rate of the code is defined as $r=k/n$.

- Note that, with BPSK modulation, the spectral efficiency is computed as follows:

$$R = \frac{k}{nT} = \frac{kB}{n} = rB$$

$$\eta = \frac{R}{B} = r$$

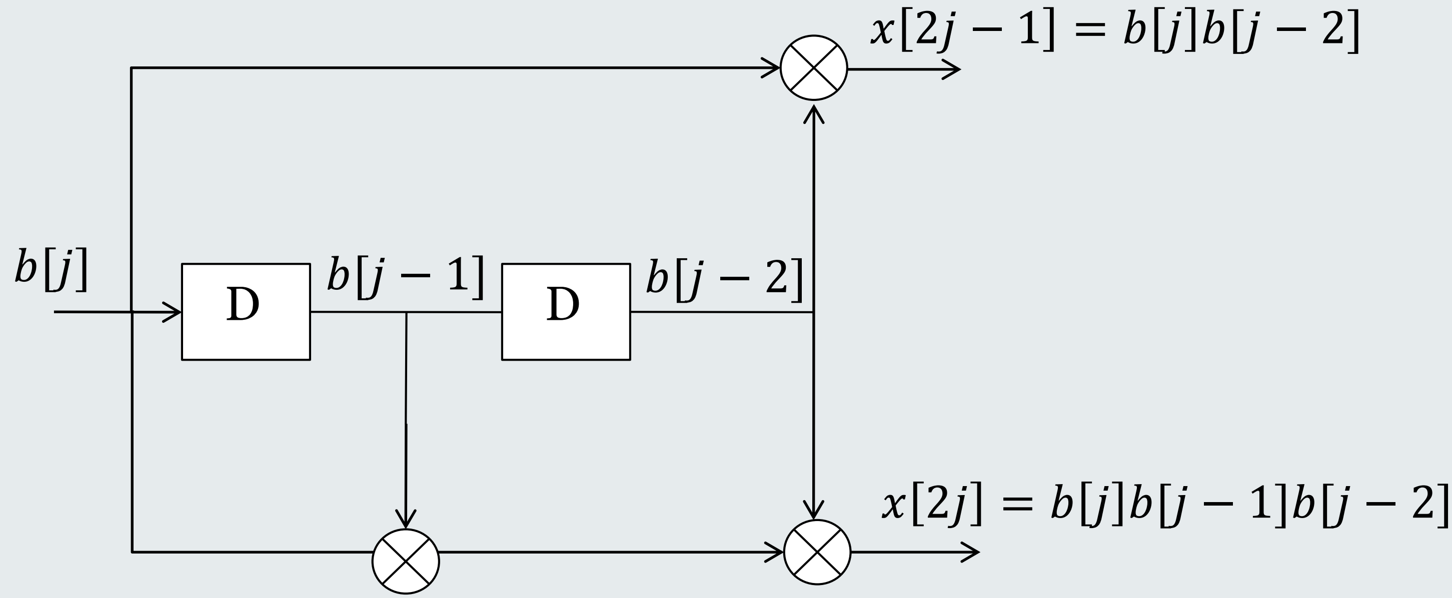
- As seen, this can be improved to

$$\eta = 2r$$

with QPSK.

How Do Convolutional Codes Work?

- **Example**: r=1/2 and constraint length (memory) = 3



input: uncoded (signed) bits

$b[j]$

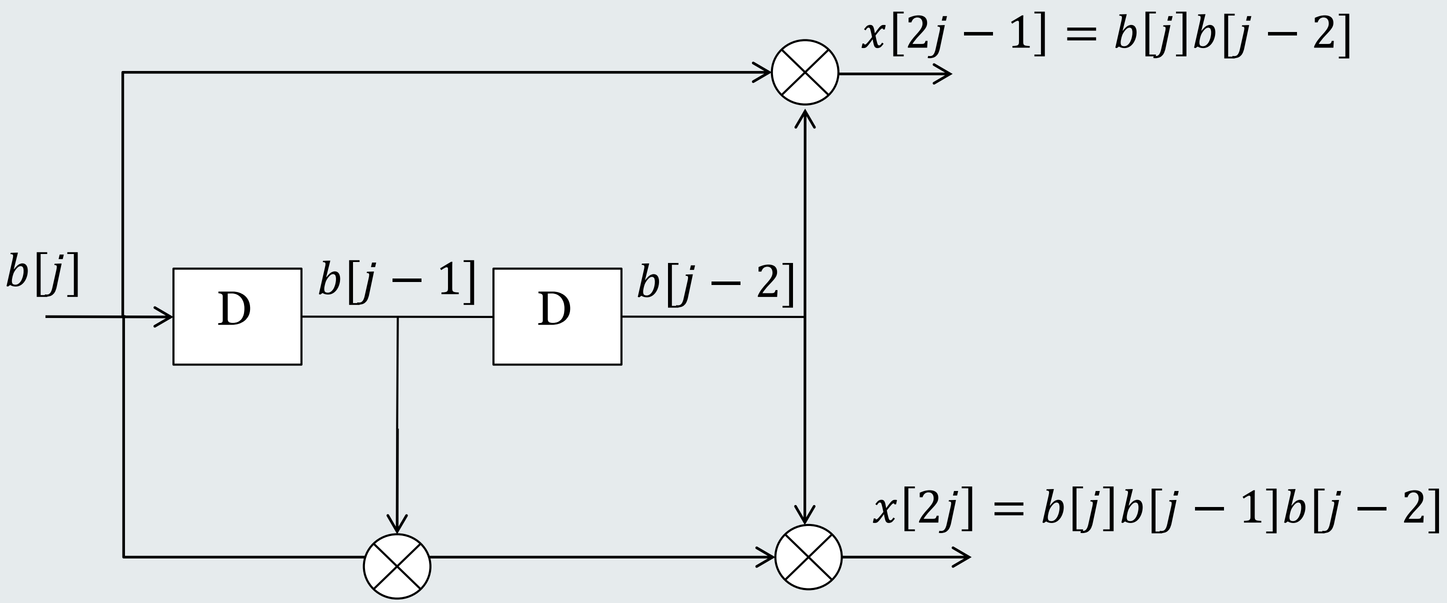
output: coded (signed) bits

$\boldsymbol{x}[j]=(x[2j\text{-}1], x[2j])$

convolutional encoder

How Do Convolutional Codes Work?

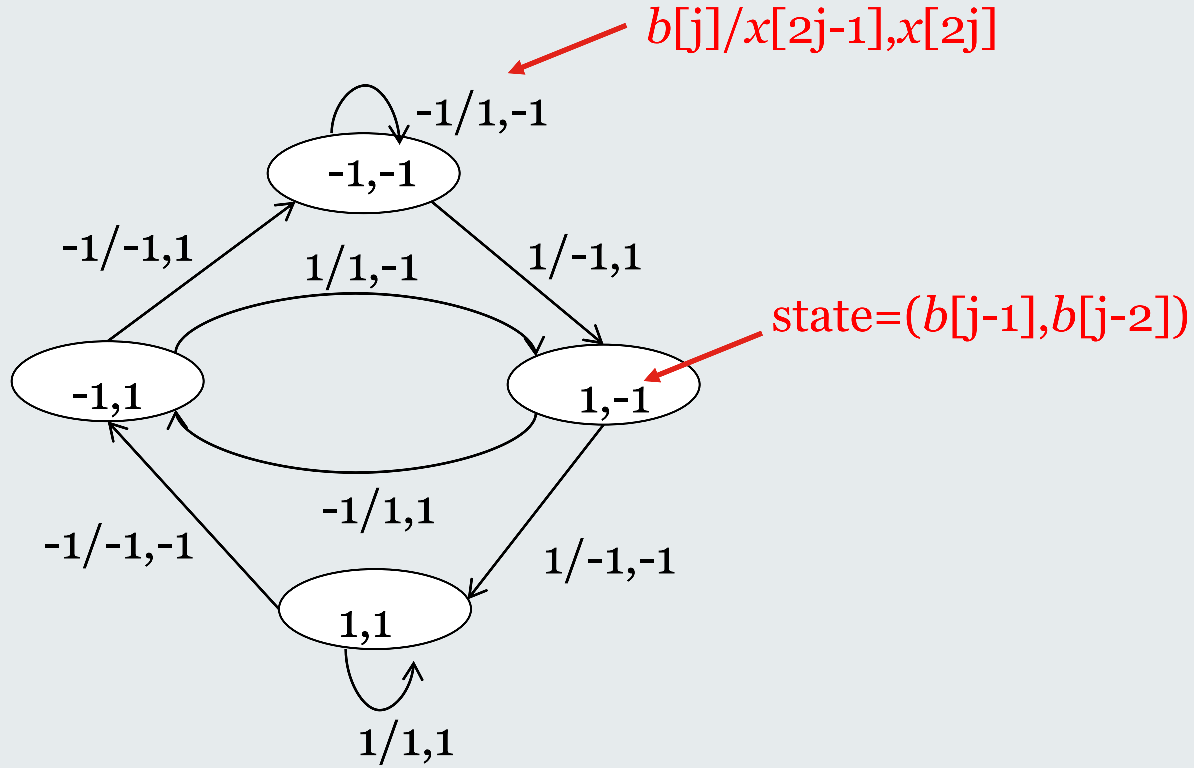
- **Example**: r=1/2 and constraint length (memory) = 3



- All registers are initialized to 1.

How Do Convolutional Codes Work?

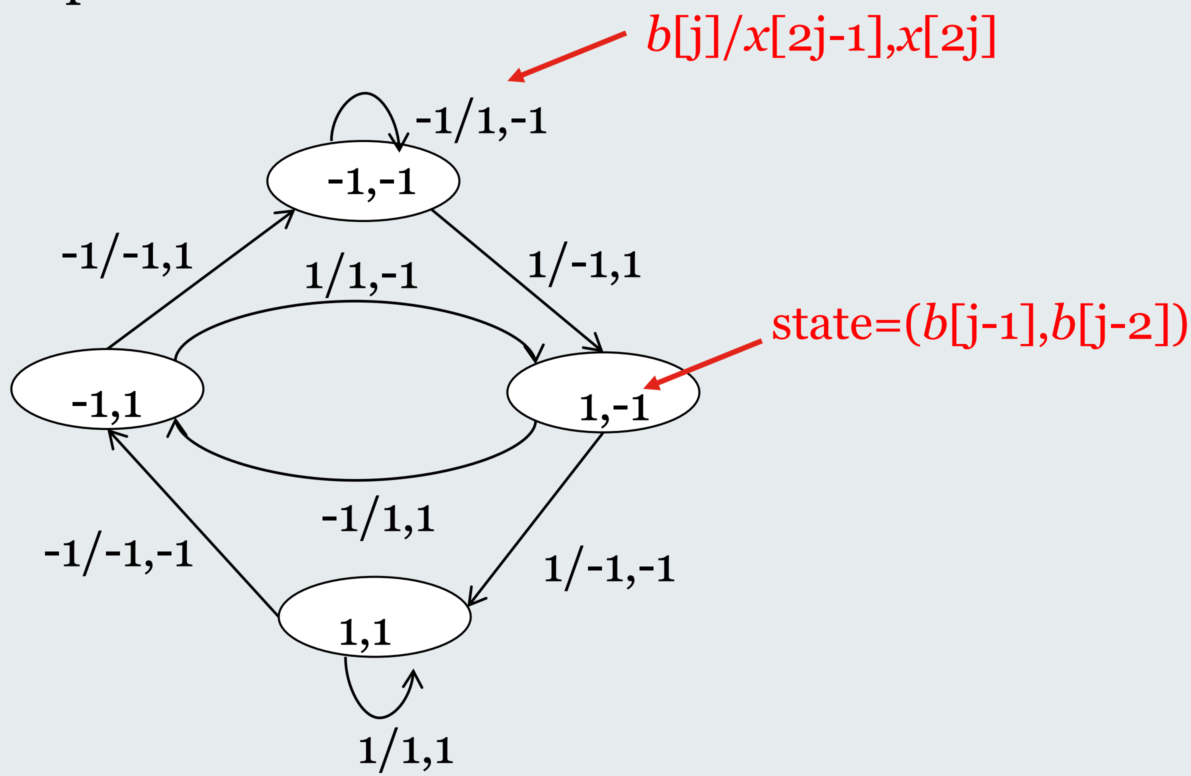
- State diagram description



state diagram

How Do Convolutional Codes Work?

- State diagram description



state diagram

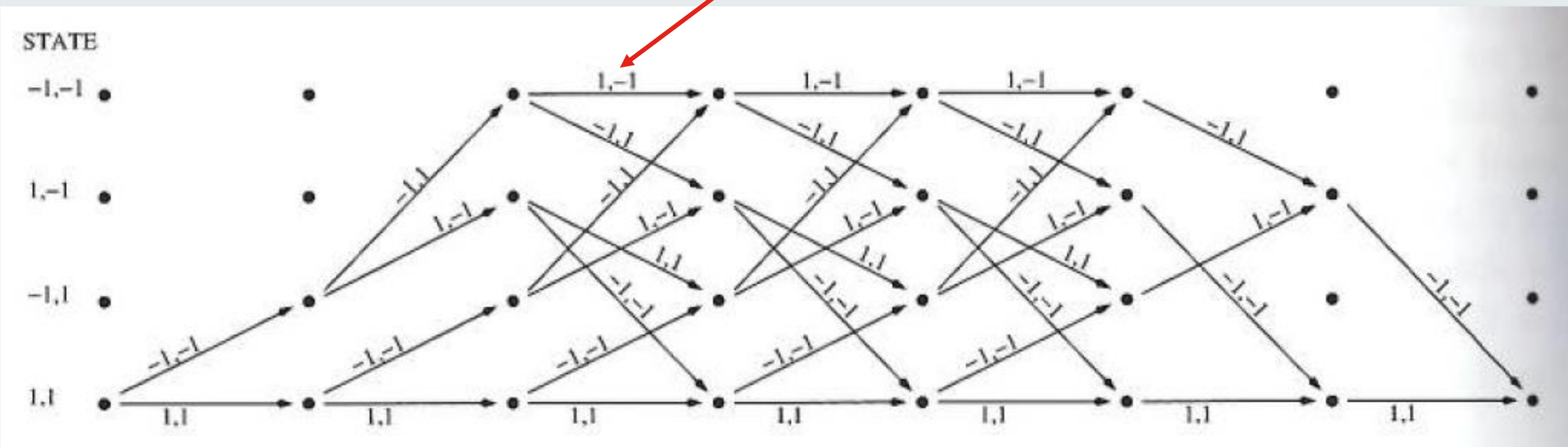
- Example:

b=[1, - 1, 1, 1, -1, 1, 1]

x=[1, 1, -1, -1, 1, -1, -1, -1. -1, -1, 1, -1, -1, -1]

How Do Convolutional Codes Work?

x[j]=(x[2j-1], x[2j])



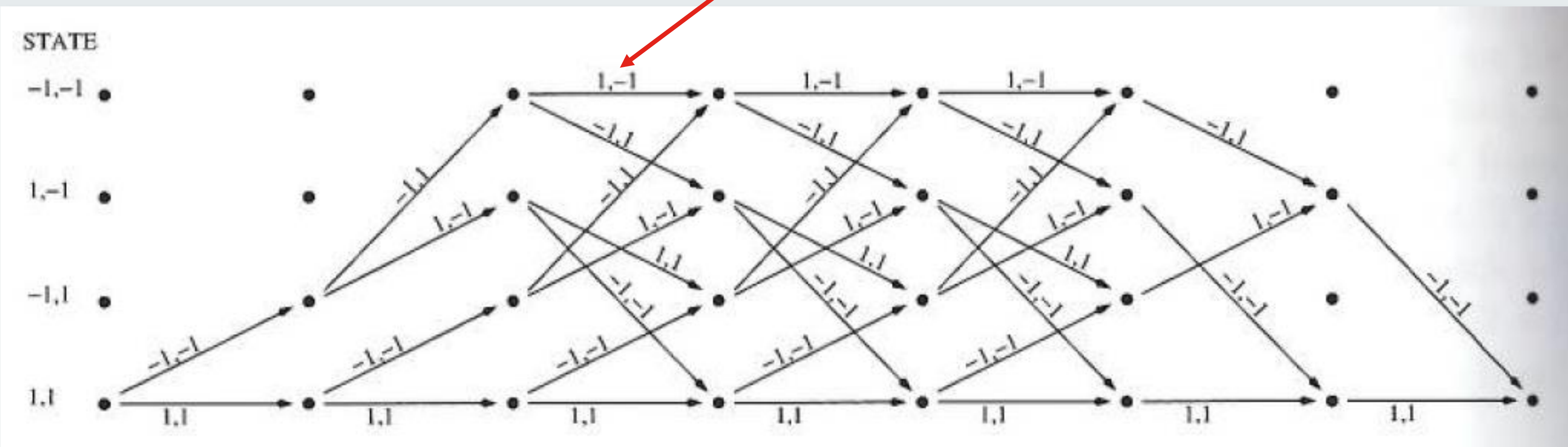
j

lower transitions correspond to b[j]=1 and higher transitions to b[j]=-1

trellis diagram

How Do Convolutional Codes Work?

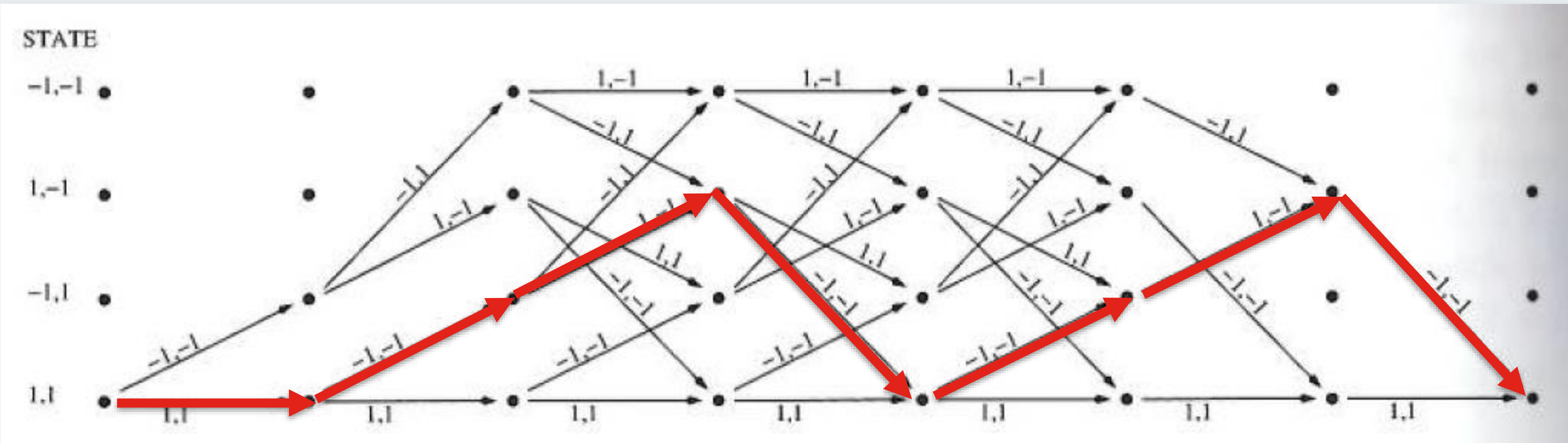
$\mathbf{x}[j]=(x[2j-1], x[2j])$



$\longrightarrow$ j

- The last two transmitted bits are fixed so as to end at state (1,1).

How Do Convolutional Codes Work?



b=[1, - 1, 1, 1, -1, 1, 1]

x=[1, 1, -1, -1, 1, -1, -1, -1. -1, -1, 1, -1, -1, -1]

trellis diagram

How to Decode a Convolutional Code?

- Minimum distance decoding would require the computation of 2^k distances

$$\min_m \sum_j ||\mathbf{y}[j] - \mathbf{x}_m[j]||^2$$

or correlations

$$\max_m \sum_j \langle \mathbf{y}[j], \mathbf{x}_m[j] \rangle$$

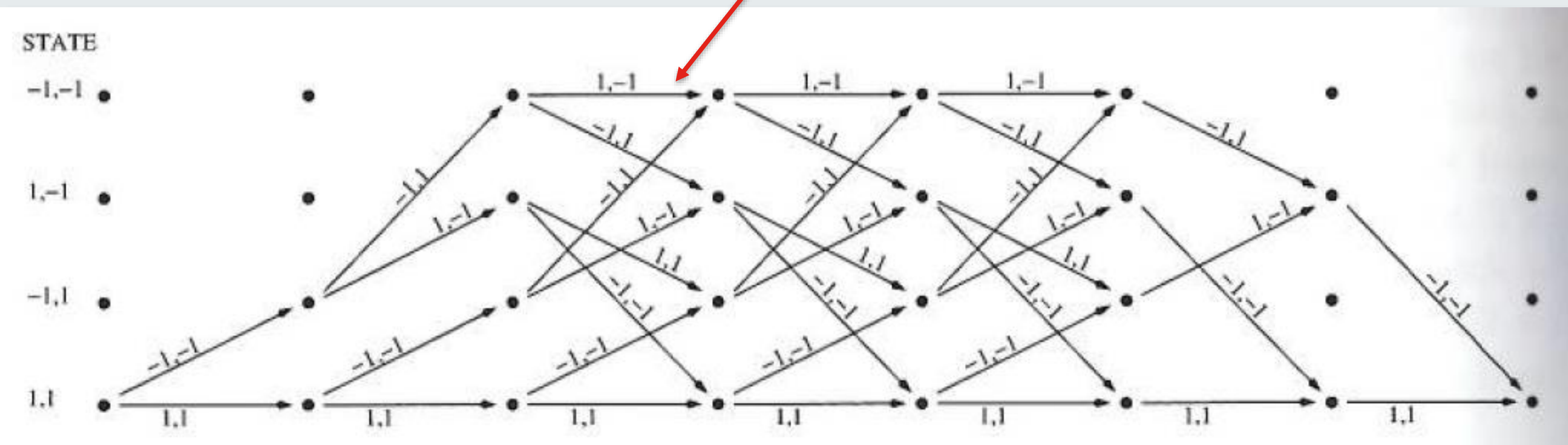
where $\boldsymbol{x}_m[j]$ is the jth encoded symbol for message m=0,1,....,$2^k - 1$.

- The Viterbi algorithm allows us to solve the problem with a complexity that is linear in n.

How to Decode a Convolutional Code?

- **1) Compute branch metrics**

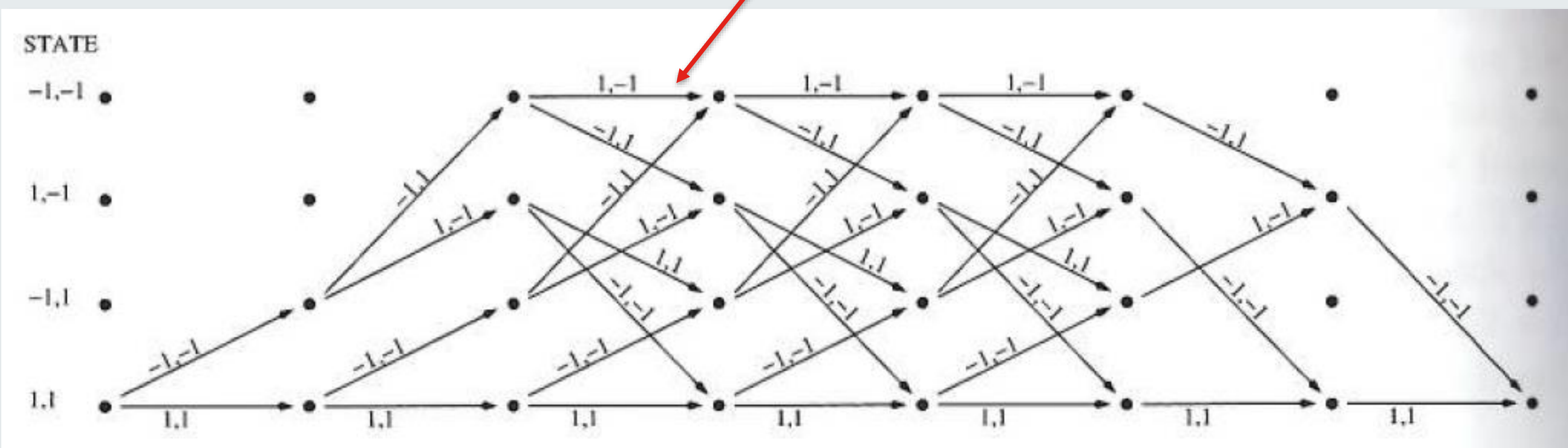
$\langle \mathbf{y}[j], \mathbf{x}_m[j] \rangle$ branch metric



How to Decode a Convolutional Code?

- Example:

$\langle \mathbf{y}[j], \mathbf{x}_m[j] \rangle$ branch metric

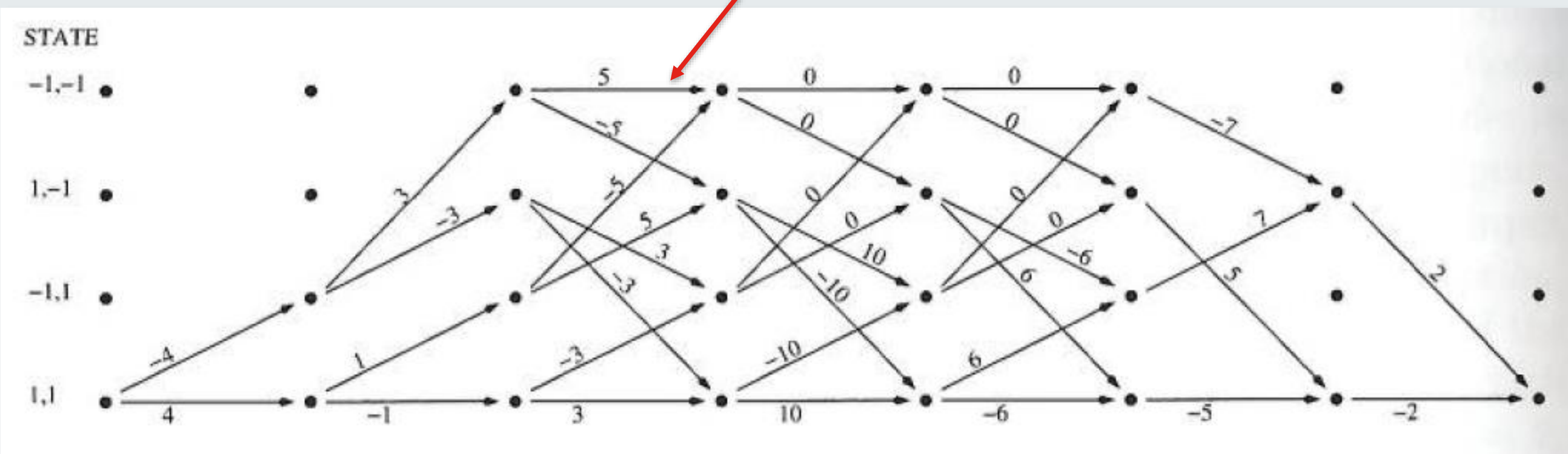


y=[(1,3), (-2,1), (4,-1), (5,5), (-3,-3), (1,-6), (2,-4)]

How to Decode a Convolutional Code?

- Example:

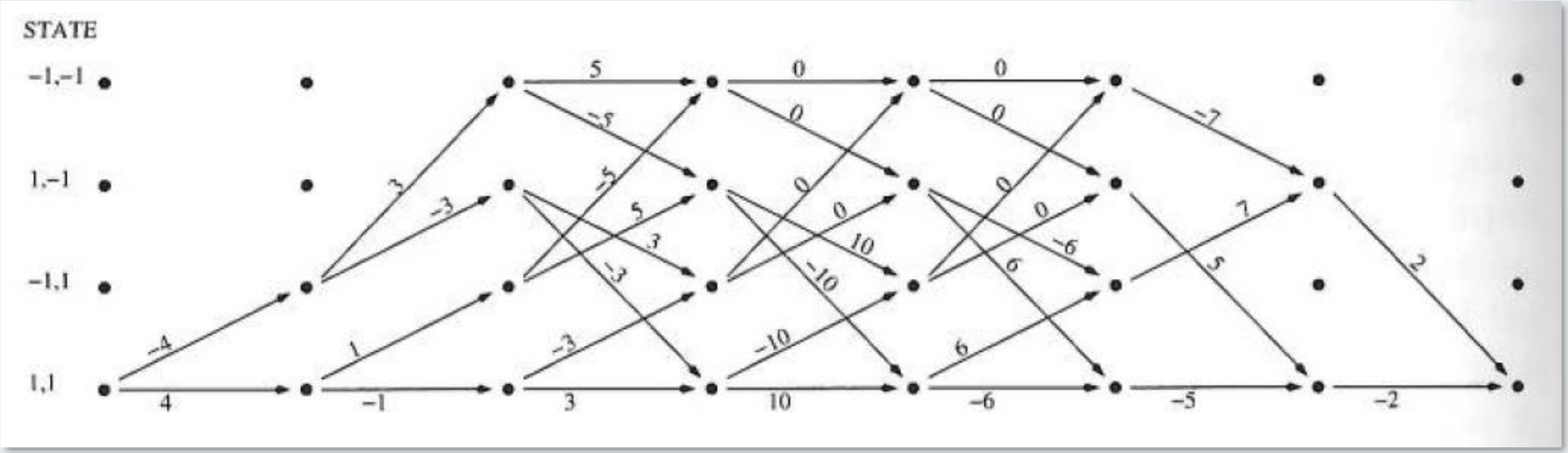
$\langle \mathbf{y}[j], \mathbf{x}_m[j] \rangle$ branch metric



$\mathbf{y}$=[(1,3), (-2,1), (4,-1), (5,5), (-3,-3), (1,-6), (2,-4)]

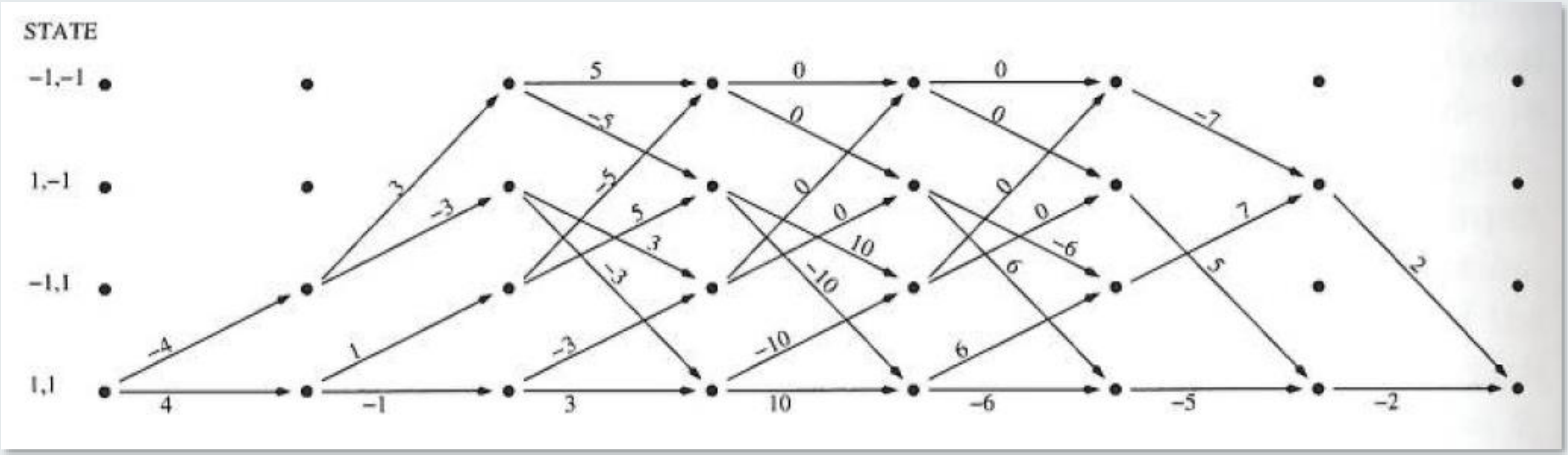
How to Decode a Convolutional Code?

- **2) Feedforward pass: for each state compute path with maximum metric (survivor)**



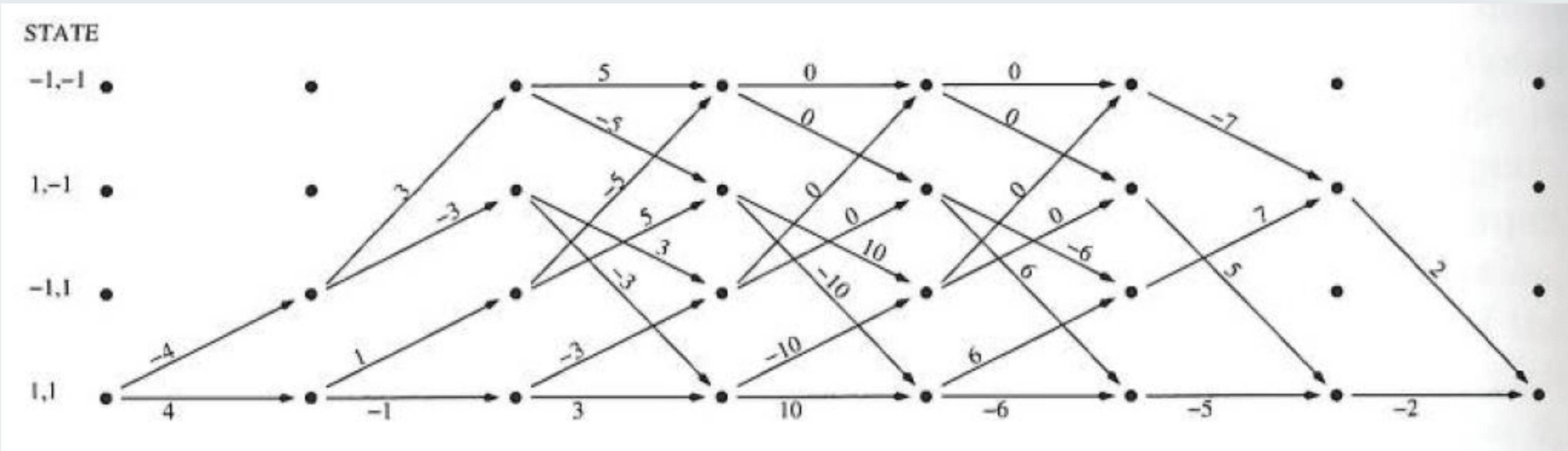
How to Decode a Convolutional Code?

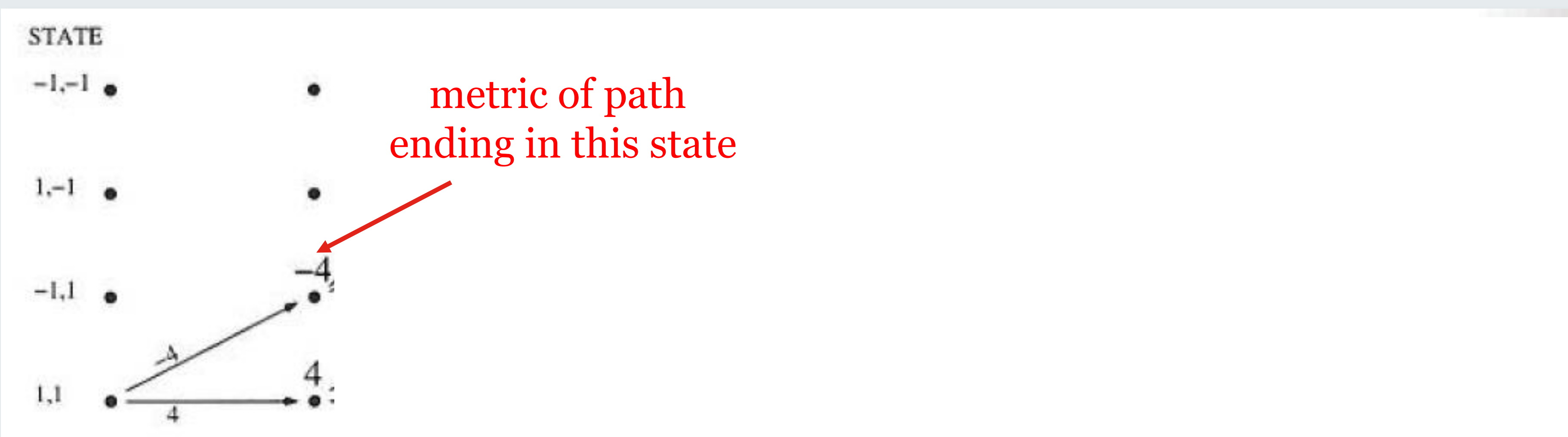
- Example:



How to Decode a Convolutional Code?

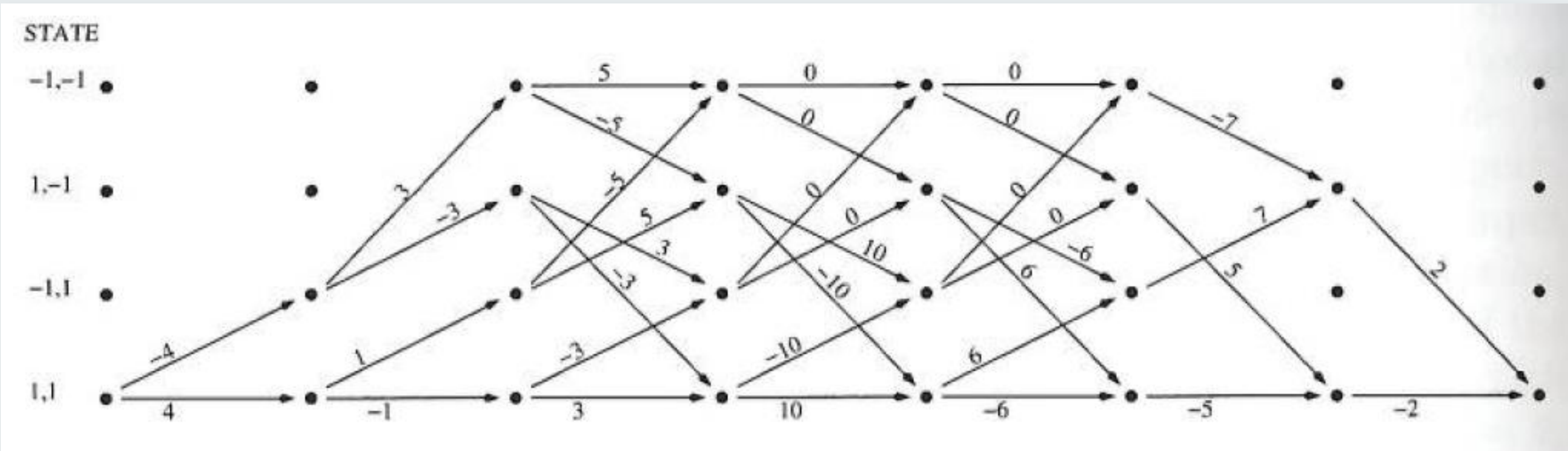
- Example:

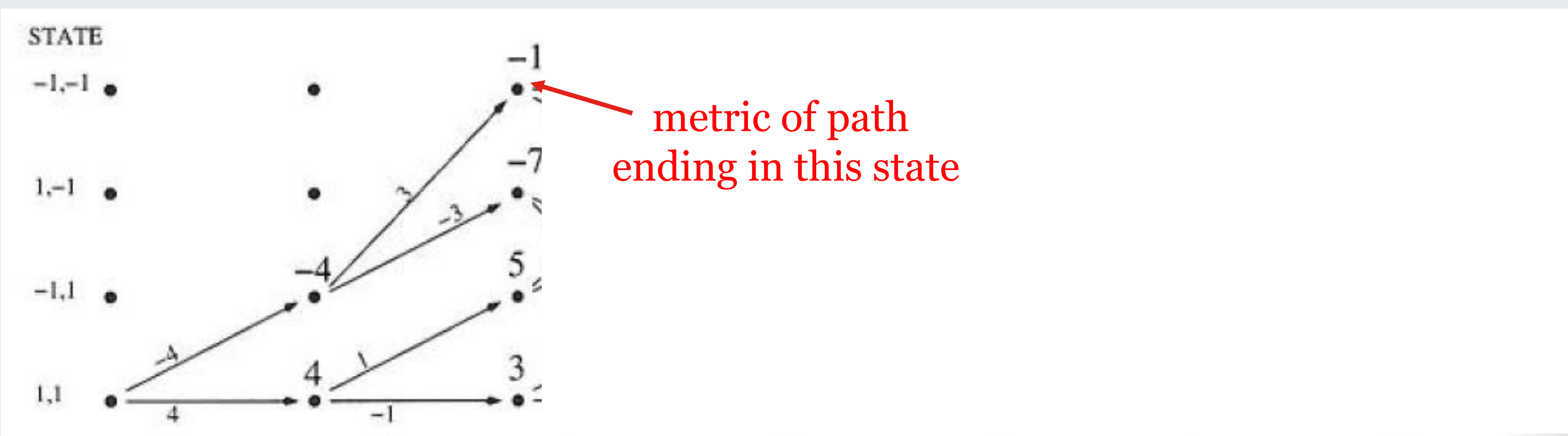




How to Decode a Convolutional Code?

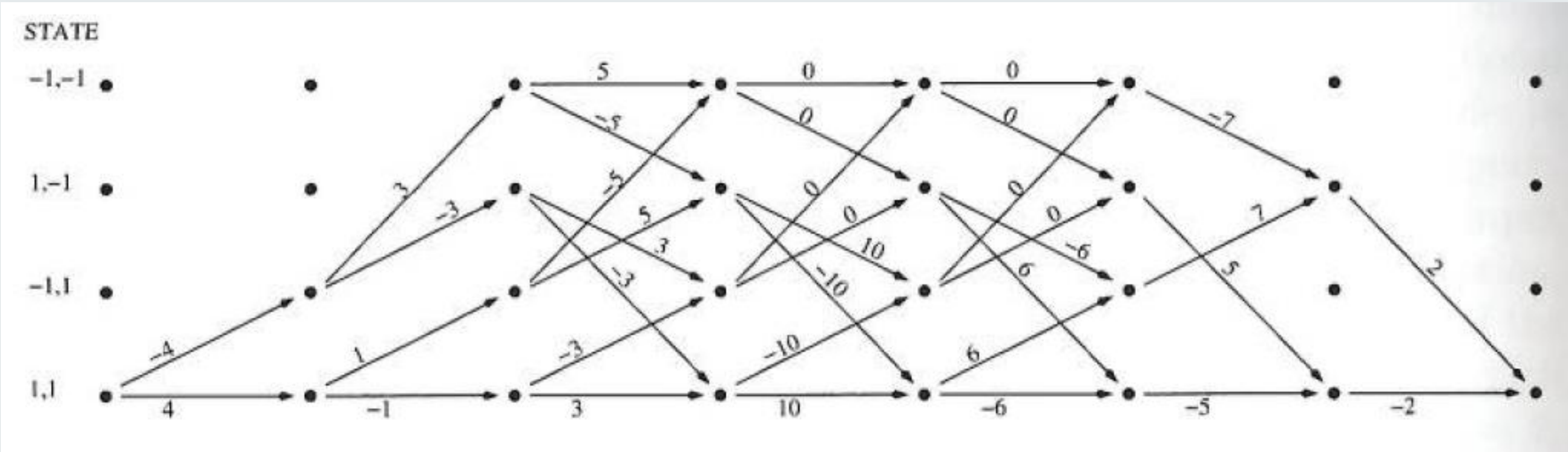
- Example:

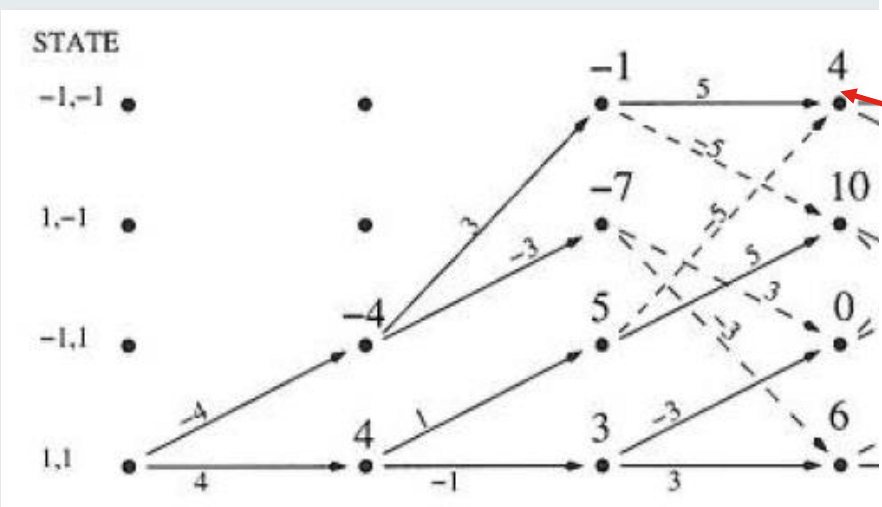




How to Decode a Convolutional Code?

- Example:

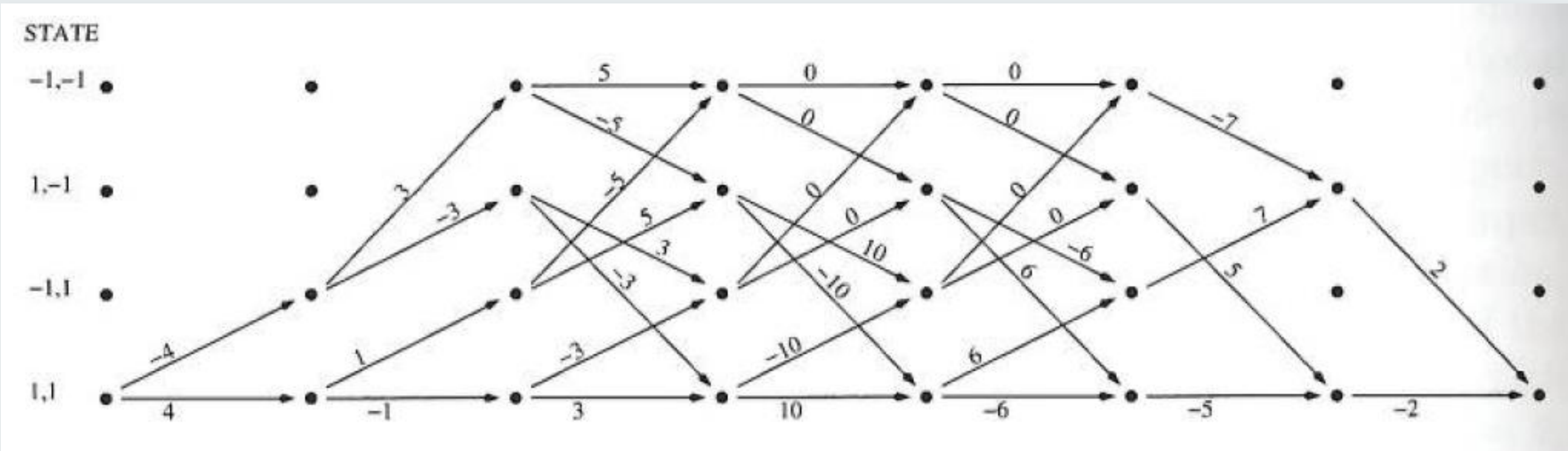




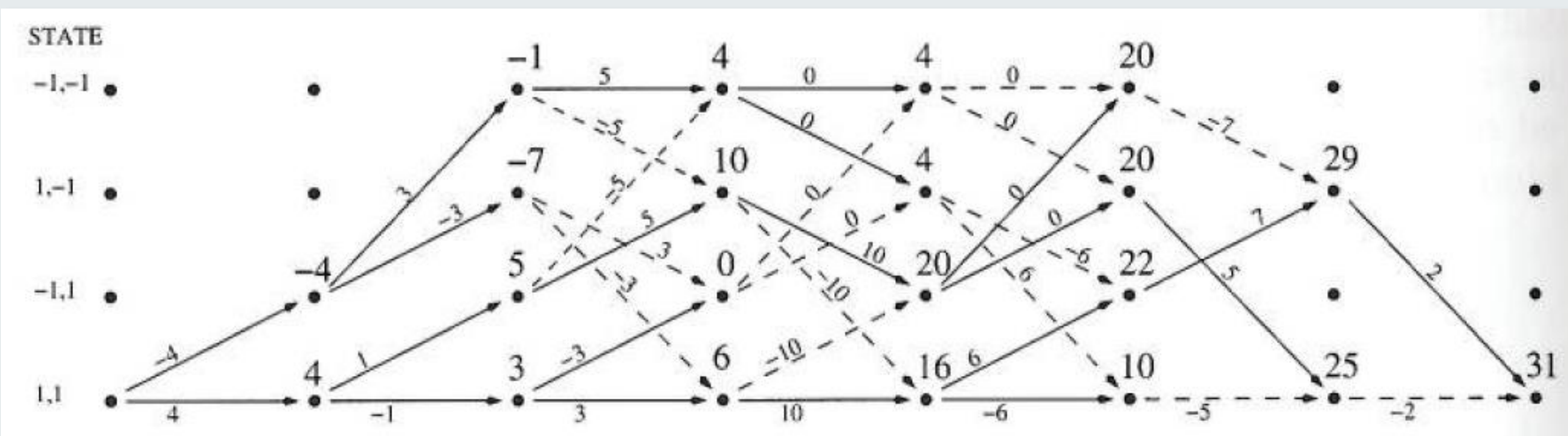
- choose maximum between -1+5 and 5-5
- prune non-surviving path (dashed)

How to Decode a Convolutional Code?

- Example:

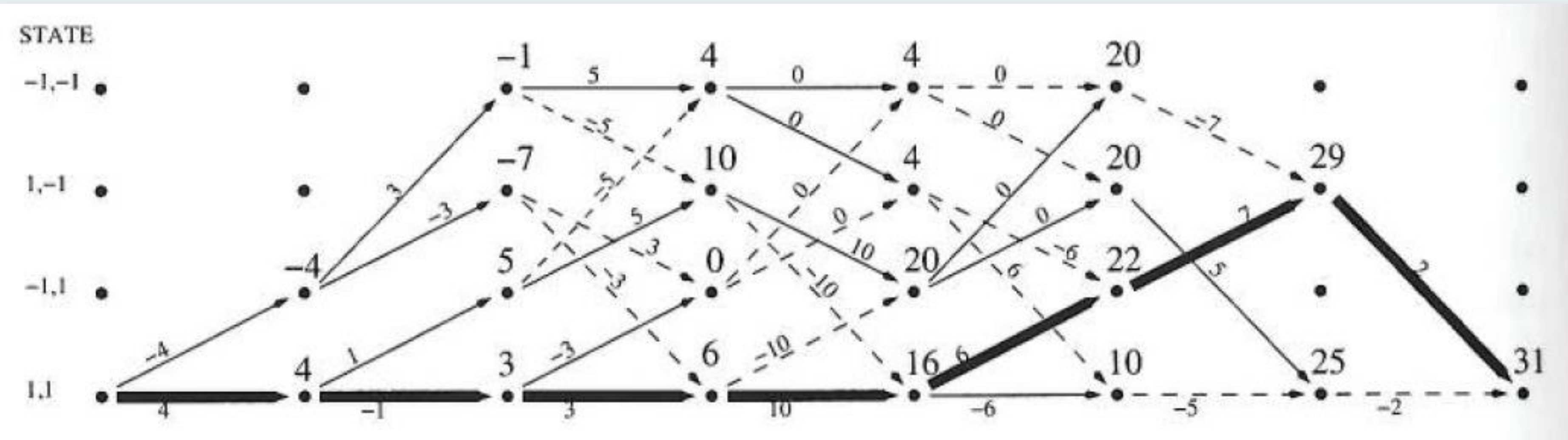


and so on…



How to Decode a Convolutional Code?

- **3) Backward pass: Backtracking to find the decoded sequence**



$\hat{m}$=[1, 1, 1, 1, -1, 1, 1]

What About Eb?

- In the previous discussion, we have chosen the coded BPSK symbols to take the values +1 and -1.

- In practice, the values of the encoded symbols x depend on Eb. How?

- Set x to be +A and –A.

What About Eb?

- In the previous discussion, we have chosen the coded BPSK symbols to take the values +1 and -1.
- In practice, the values of the encoded symbols x depend on Eb. How?
- Set *x* to be +A and –A.
- Note that we have

$$\boxed{kE_b = \text{total energy}}$$

- We conclude that

$$\boxed{nA^2 = kE_b \Rightarrow A = \sqrt{rE_b}}$$

How to Compute the Probability of Error?

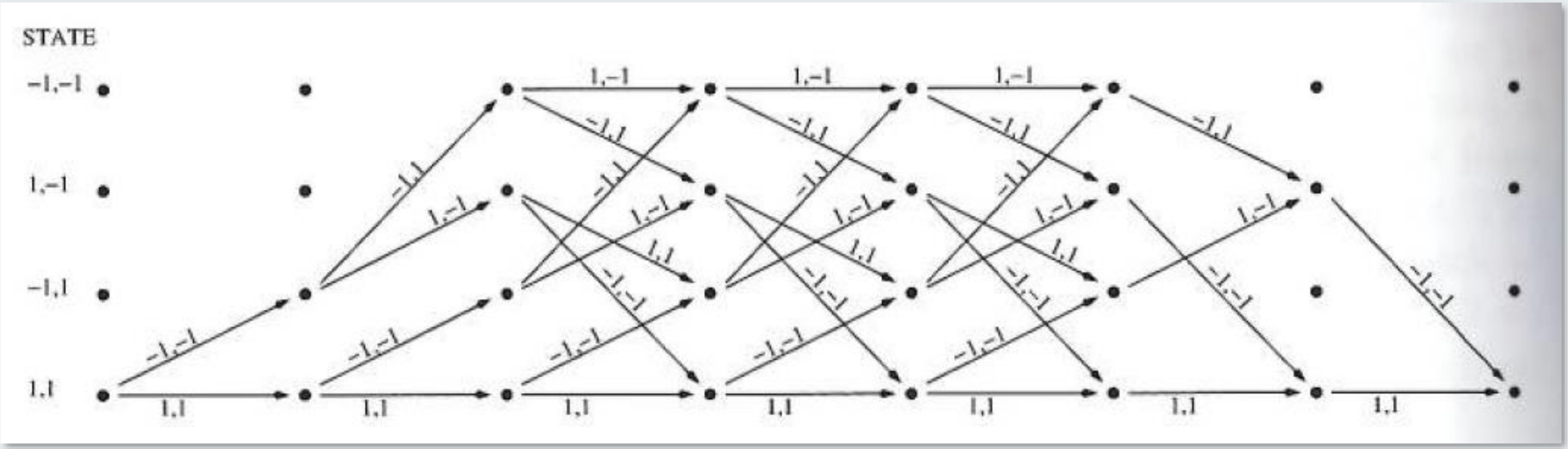
- To compute the probability of error, and hence the energy efficiency, we can use the union bound.

$$P_e \cong (\text{number of pairs at the min distance})\, Q\left(\sqrt{\frac{d_{min}^2}{2N_0}}\right)$$

- How to compute the minimum distance?

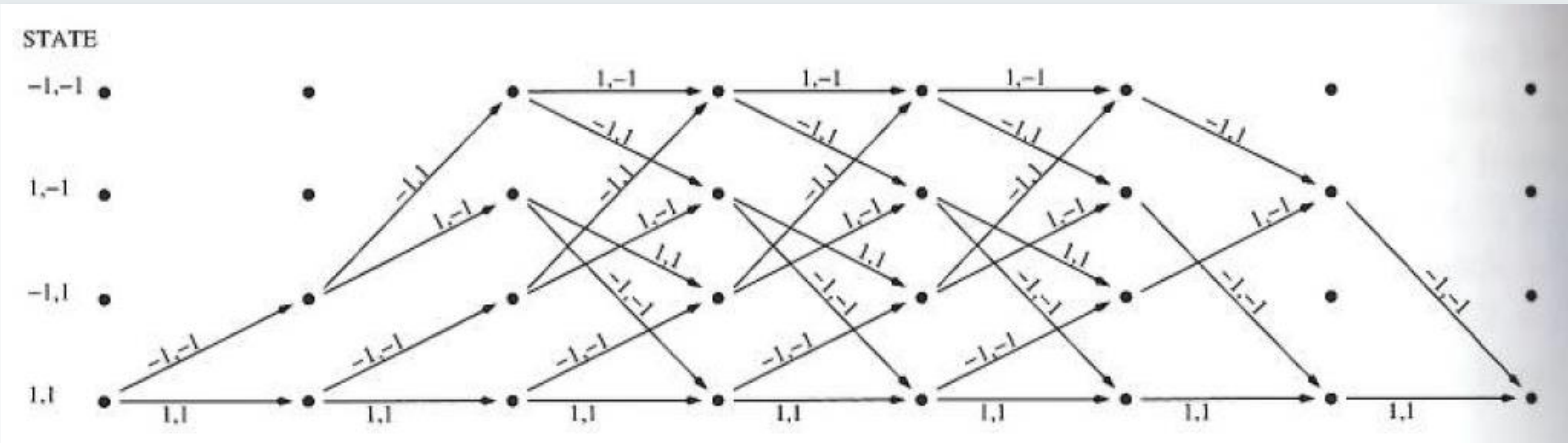
How to Compute the Probability of Error?

- Each codeword $\boldsymbol{x}_m$, encoding a message m, corresponds to a path on the trellis.



How to Compute the Probability of Error?

- Each codeword $\boldsymbol{x}_m$, encoding a message m, corresponds to a path on the trellis.



- The distance between two codewords is equal to

$$d^2(\mathrm{x}_m, \mathrm{x}_{m'}) = \sum_j ||\mathrm{x}_m[j] - \mathrm{x}_{m'}[j]||^2$$
$$= 4rE_b d_H(\mathrm{x}_m, \mathrm{x}_{m'})$$

Hamming distance between two codewords
= number of differences

How to Compute the Probability of Error?

- To compute the probability of error, and hence the energy efficiency, we can use the union bound.

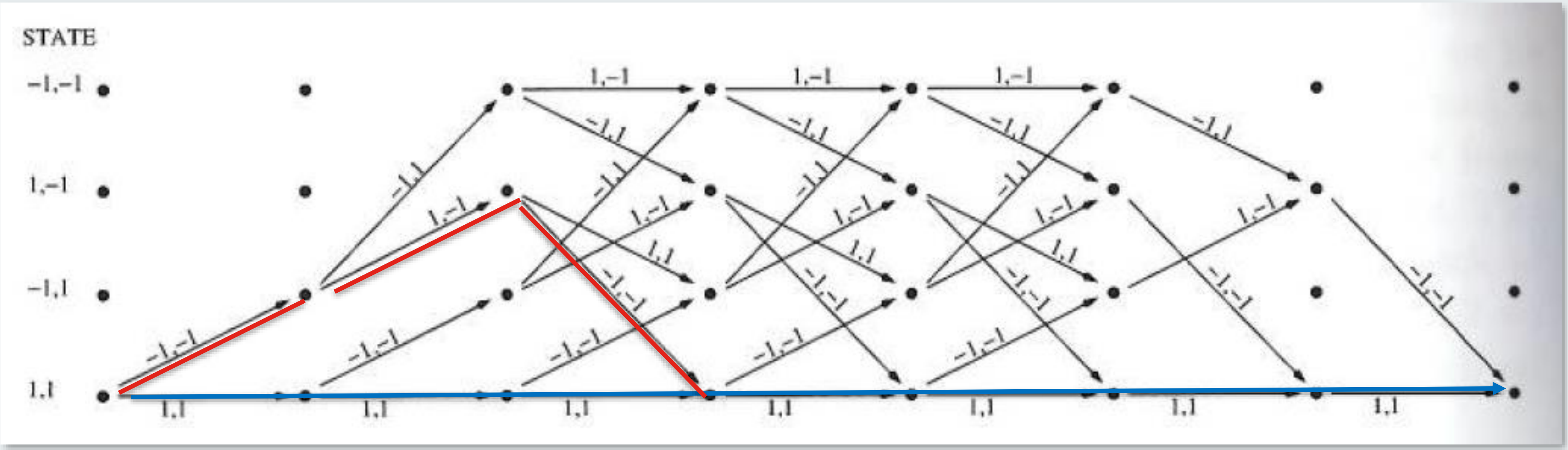
$$P_e \cong \frac{1}{2^k}(\text{number of pairs at the min distance})\,Q\left(\sqrt{\frac{d_{min}^2}{2N_0}}\right)$$

- How to compute the minimum distance?

- Consider one codeword as a reference, namely the all-1 codeword.

- We are interested in finding the **detour** on the trellis that has the minimum Hamming distance.

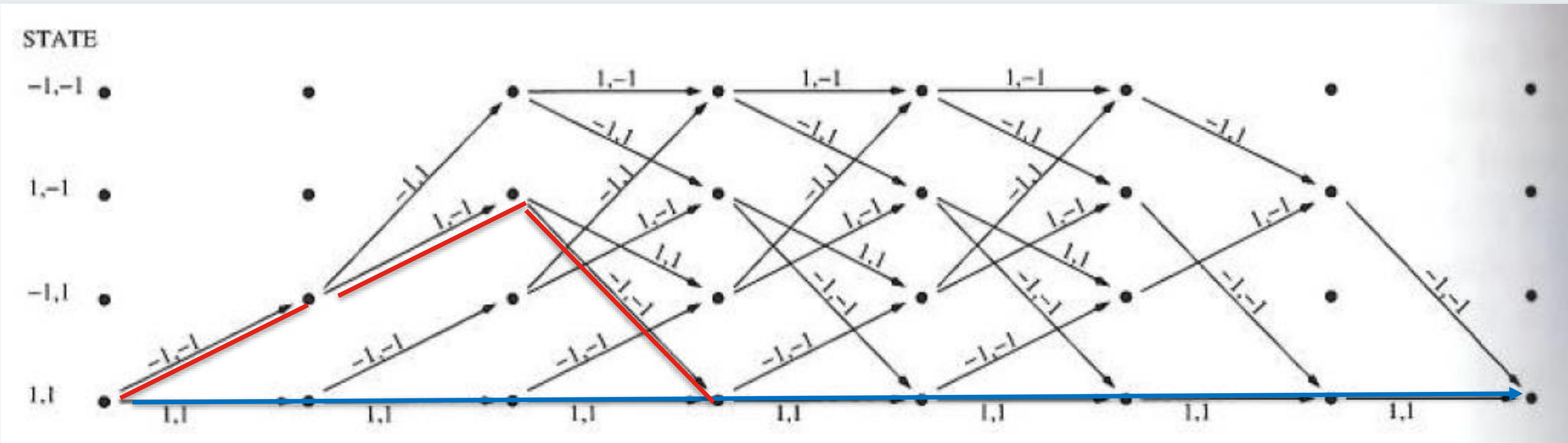
How to Compute the Probability of Error?



a detour

reference path

How to Compute the Probability of Error?

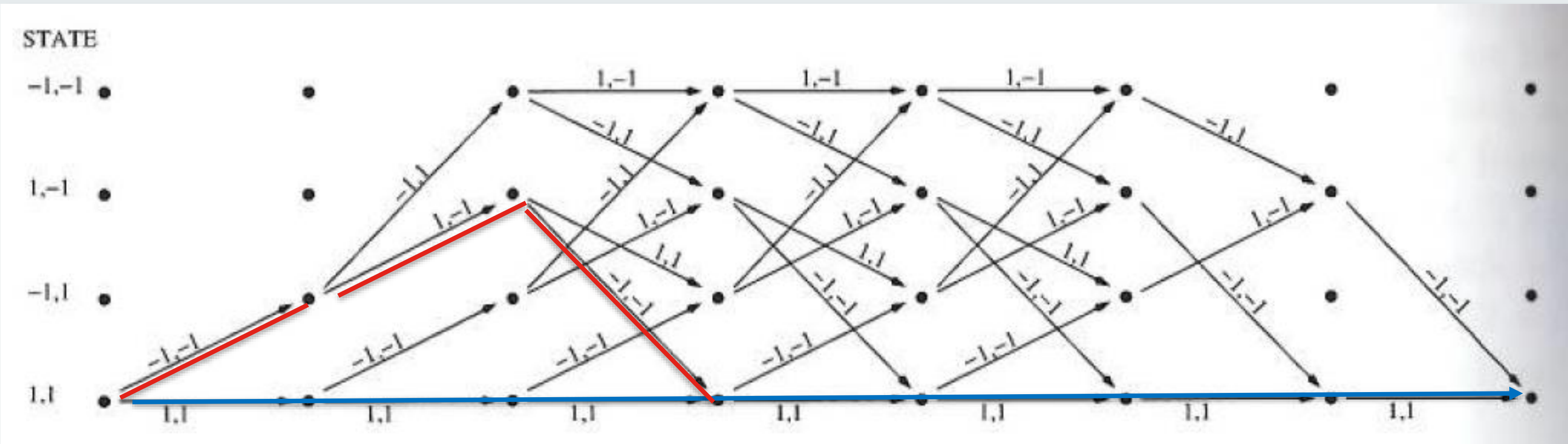


a detour

reference path

$$d_H(\mathbf{x}_{ref}, \mathbf{x}_{det}) = 5$$

How to Compute the Probability of Error?



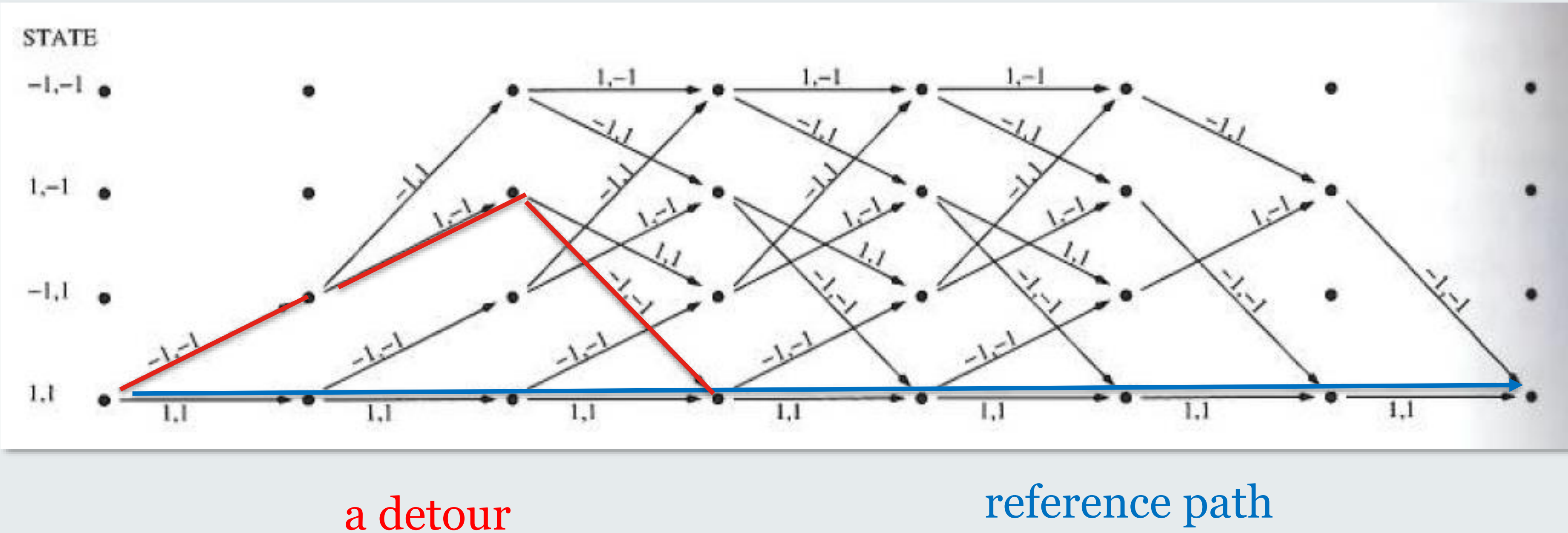
a detour

reference path

$$d_H(\mathbf{x}_{ref}, \mathbf{x}_{det}) = 5$$

- It is not difficult to see that there is no detour at a smaller Hamming distance (why?).

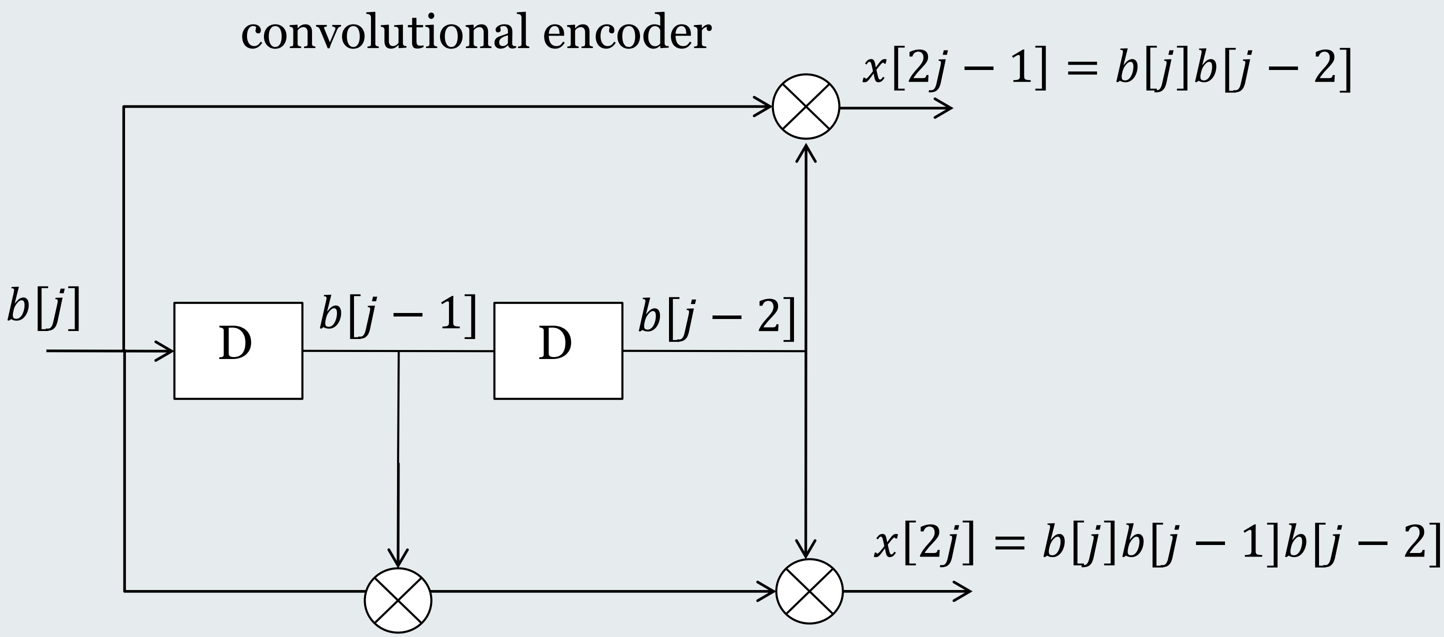
How to Compute the Probability of Error?



$$P_e \cong \frac{1}{2^k}(\text{number of pairs at the min distance})Q\left(\sqrt{\frac{4rE_b \times 5}{2N_0}}\right)$$

$$= \frac{1}{2^k}(\text{number of pairs at the min distance})Q\left(\sqrt{\frac{5E_b}{N_0}}\right)$$

In Summary…



- This convolutional code has spectral efficiency of ½ with BPSK and 1 with QPSK and a **coding gain** with respect to BPSK of

$$10\log_{10}\left(\frac{5}{2}\right) \cong 4 \text{ dB}$$

- This improves over repetition code with the same spectral efficiency, which has a coding gain of 3 dB

More Examples of Convolutional Codes

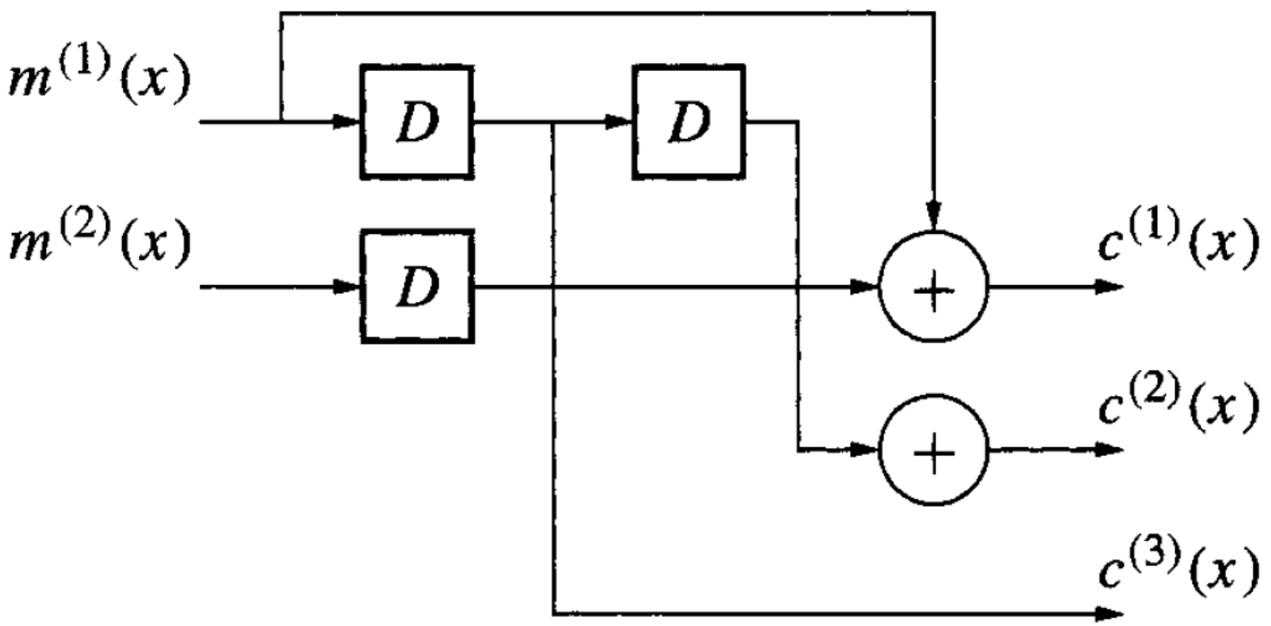


Figure 12.7: A feedforward $r = 2/3$ encoder.

[Moon '05]

More Examples of Convolutional Codes

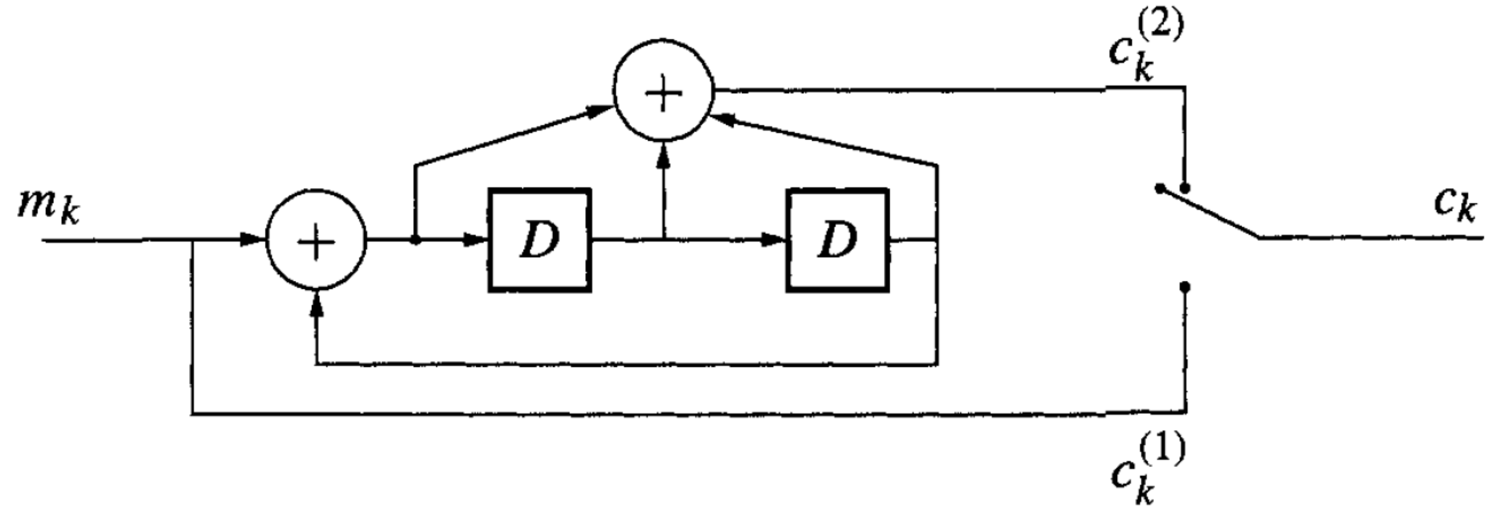


Figure 12.2: A systematic $r = 1/2$ encoder.

[Moon '05]

More Examples of Convolutional Codes

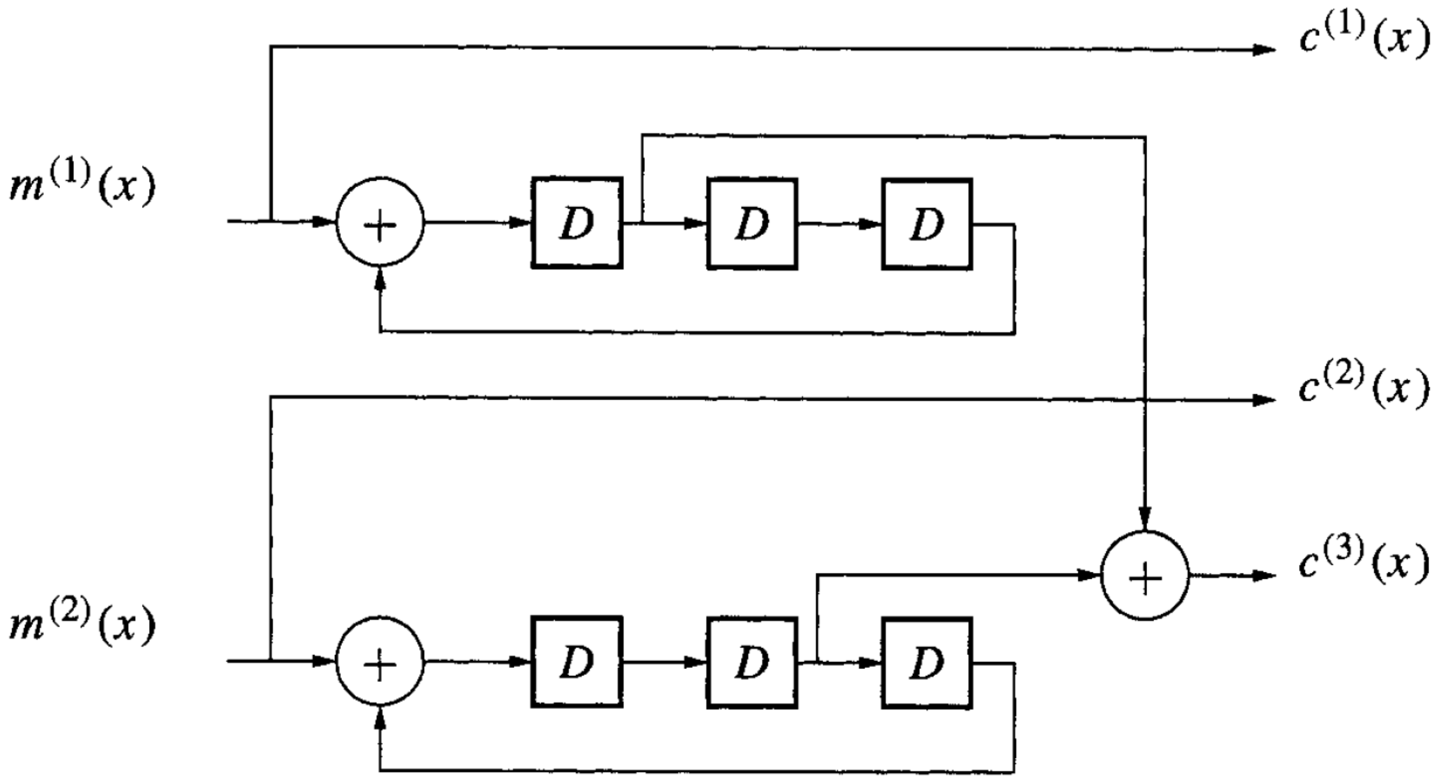
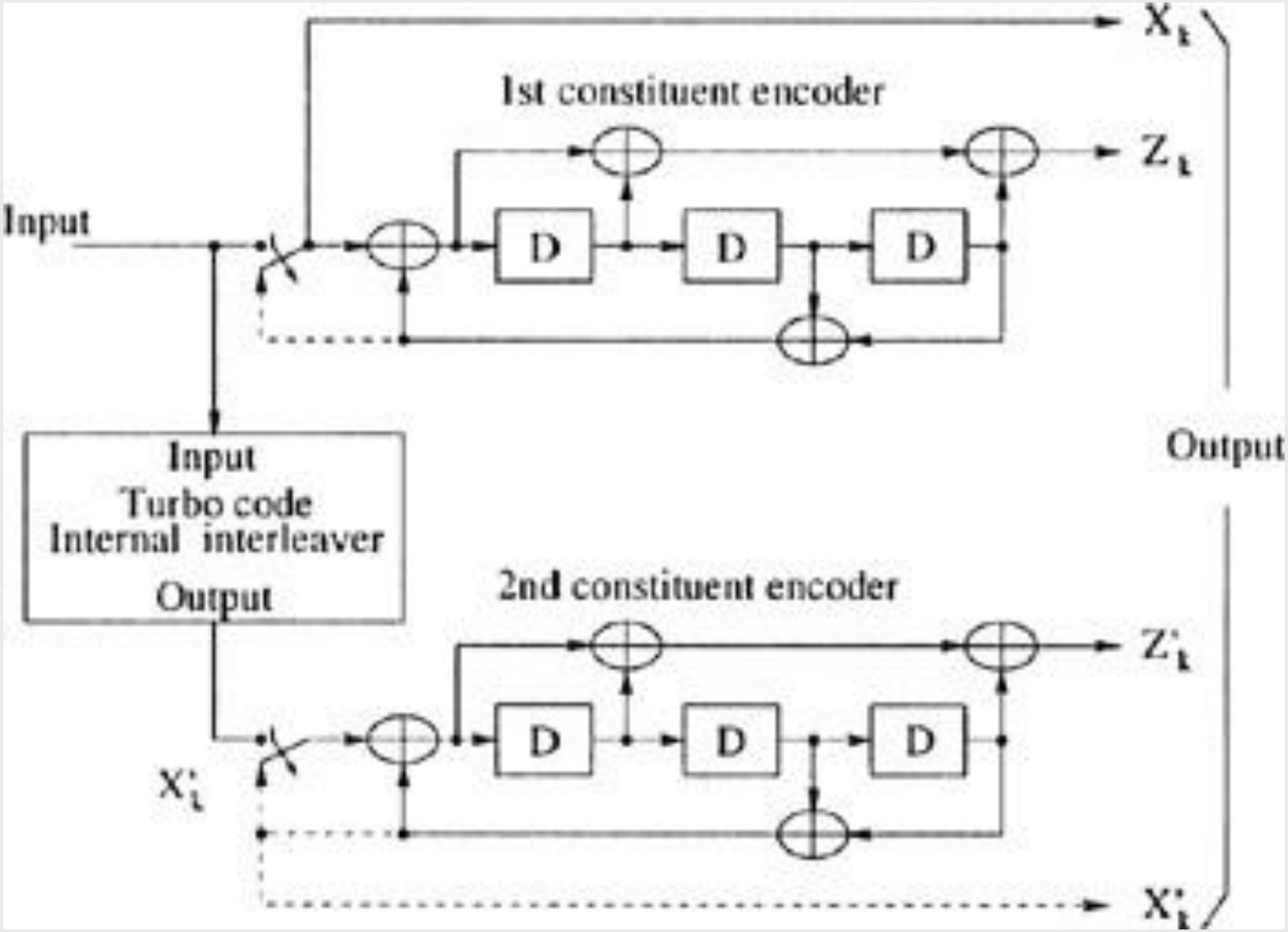


Figure 12.3: A systematic $r = 2/3$ encoder.

[Moon '05]

Turbo Codes



X_k
1st constituent encoder
Z_k
Input
D
D
D
Input
Turbo code
Internal interleaver
Output
Output
2nd constituent encoder
Z'_k
D
D
D
X'_k
X'_k

How to Increase the Spectral Efficiency?

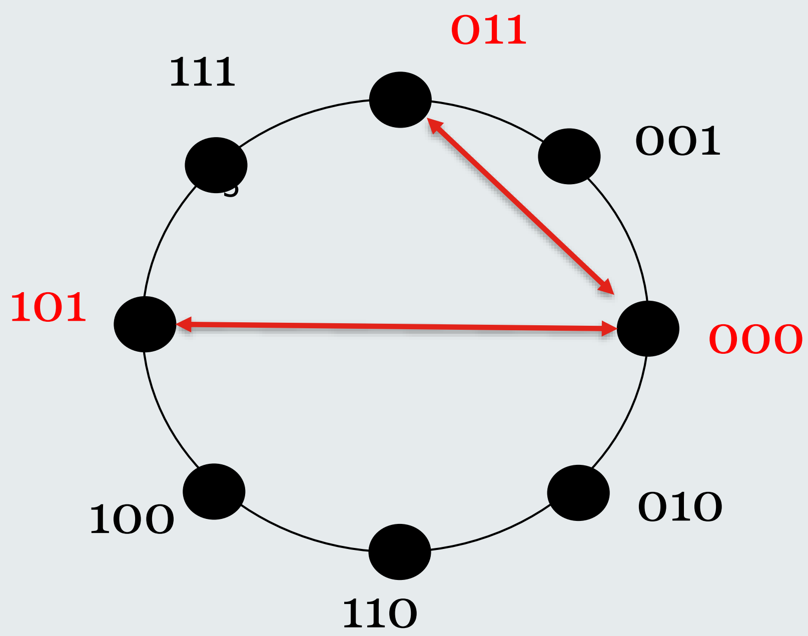
- In order to increase the spectral efficiency, we need to use large constellations.

- In fact, if we used a constellation carrying b bits, the spectral efficiency would be br.

- However, with a constellation that is different from QPSK, the distances between codewords would **not** be proportional to the corresponding Hamming distances.

How to Increase the Spectral Efficiency?

- Example:

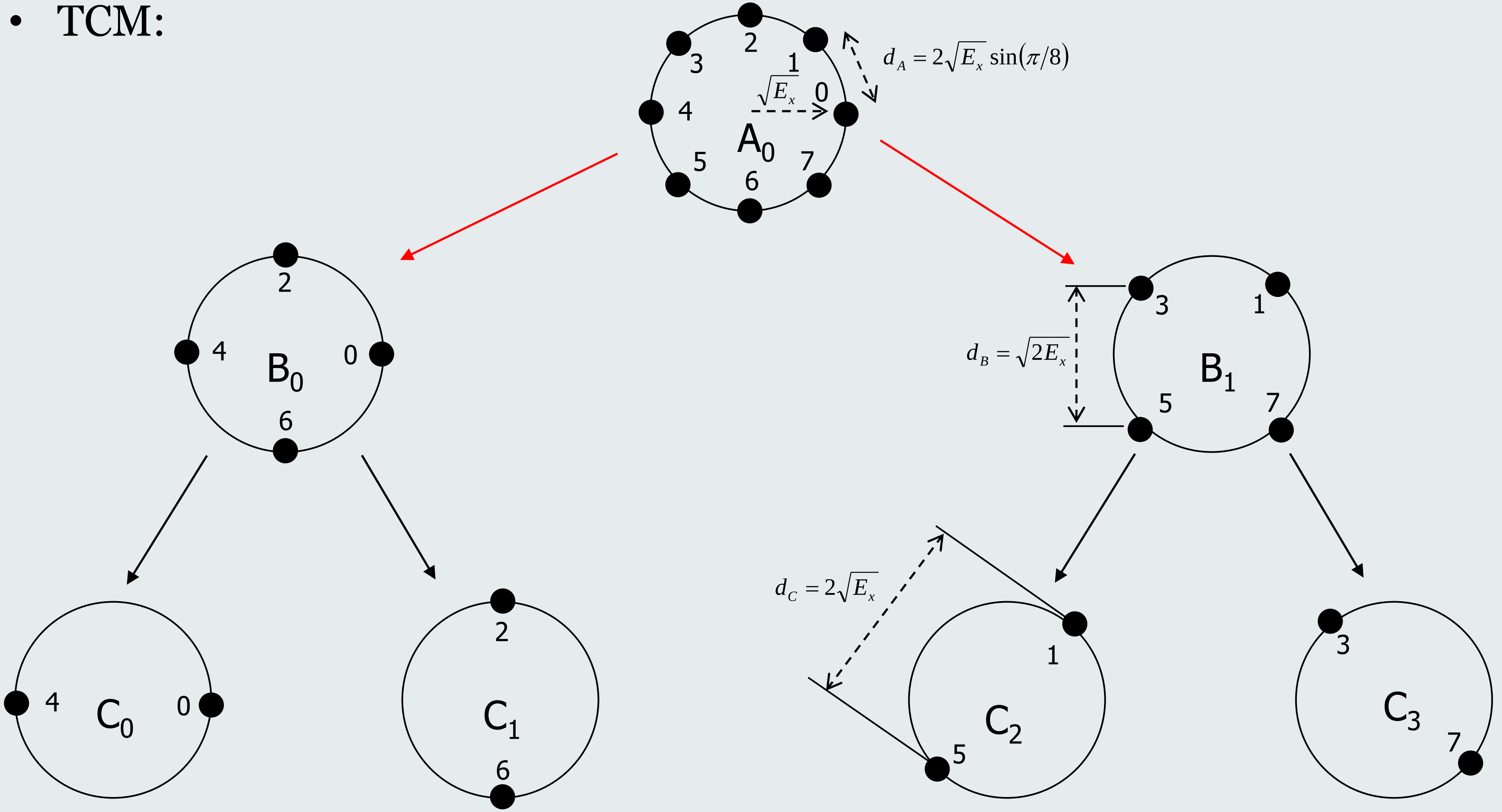


How to Increase the Spectral Efficiency?

- Three main solutions:
  - Trellis Coded Modulation (TCM)
  - Multilevel modulation
  - Bit Interleaved Coded Modulation (BICM)

How to Increase the Spectral Efficiency?

- TCM:



code only chooses subset at the first level

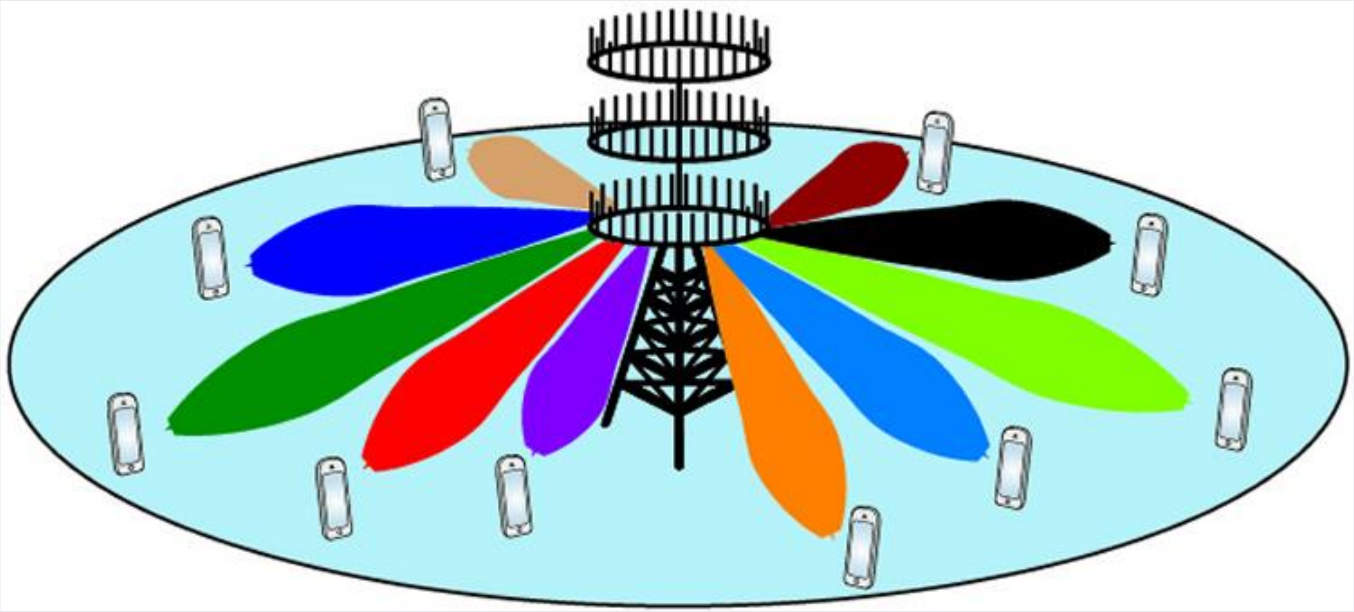
Additional Material

What if We Have Multiple Antennas?

- Modern base stations and access points have multiple antennas.
- Mobile devices at higher frequencies can also pack multiple antennas.

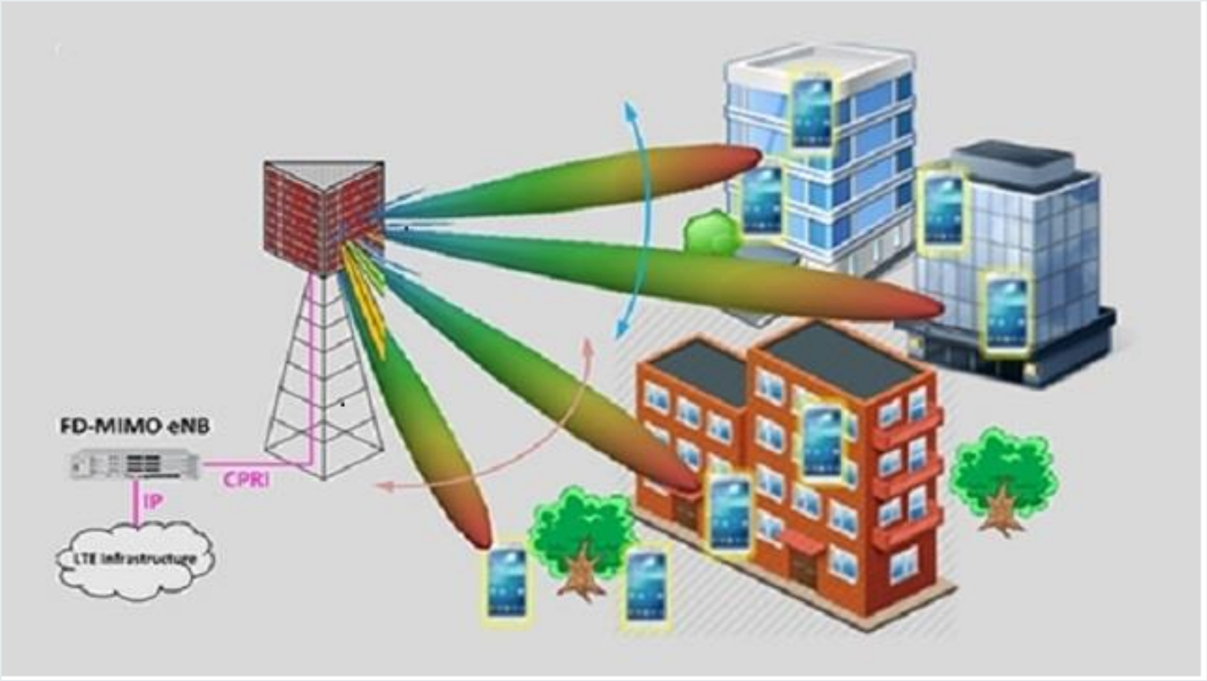
What if We Have Multiple Antennas?

- Massive MIMO, i.e., the deployment of base stations with hundreds of antennas, is a key technology for 5G.



What if We Have Multiple Antennas?

- Multiple antennas at the transmitting end of the communication link enable:
  - **Diversity**: transmission from multiple antennas can be leveraged so as to reduce the change of deep fades in the channel
  - **Space-Division Multiple Access**: serving multiple users in the same time-frequency resources by using beamforming (see figure)

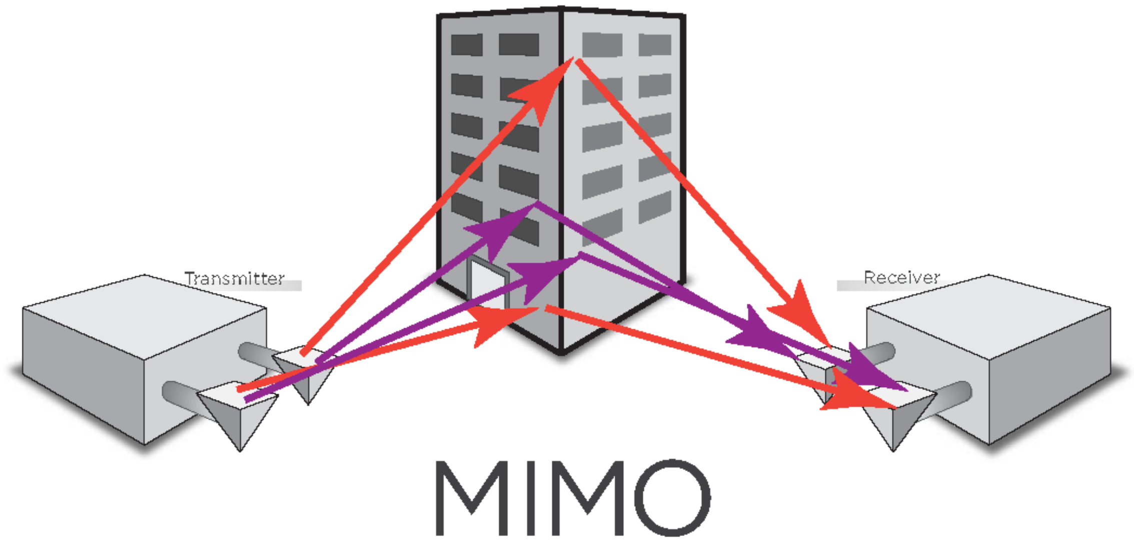


What if We Have Multiple Antennas?

- Multiple antennas at the receving end of the communication link enable:
  - **Diversity**: reception from multiple antennas can be combined so as to reduce the change of deep fades in the channel

What if We Have Multiple Antennas?

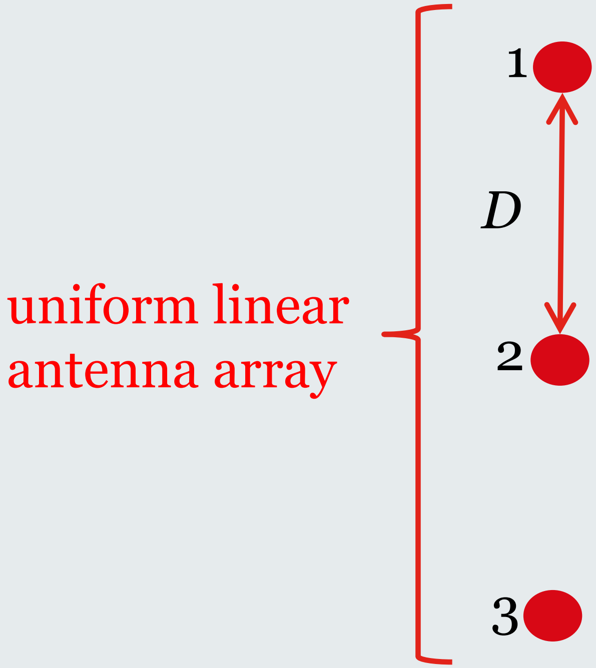
- Multiple antennas at both ends of a communication link enable transmission of multiple streams between a transmitter and a receiver.



(Multiple Input Multiple Output)

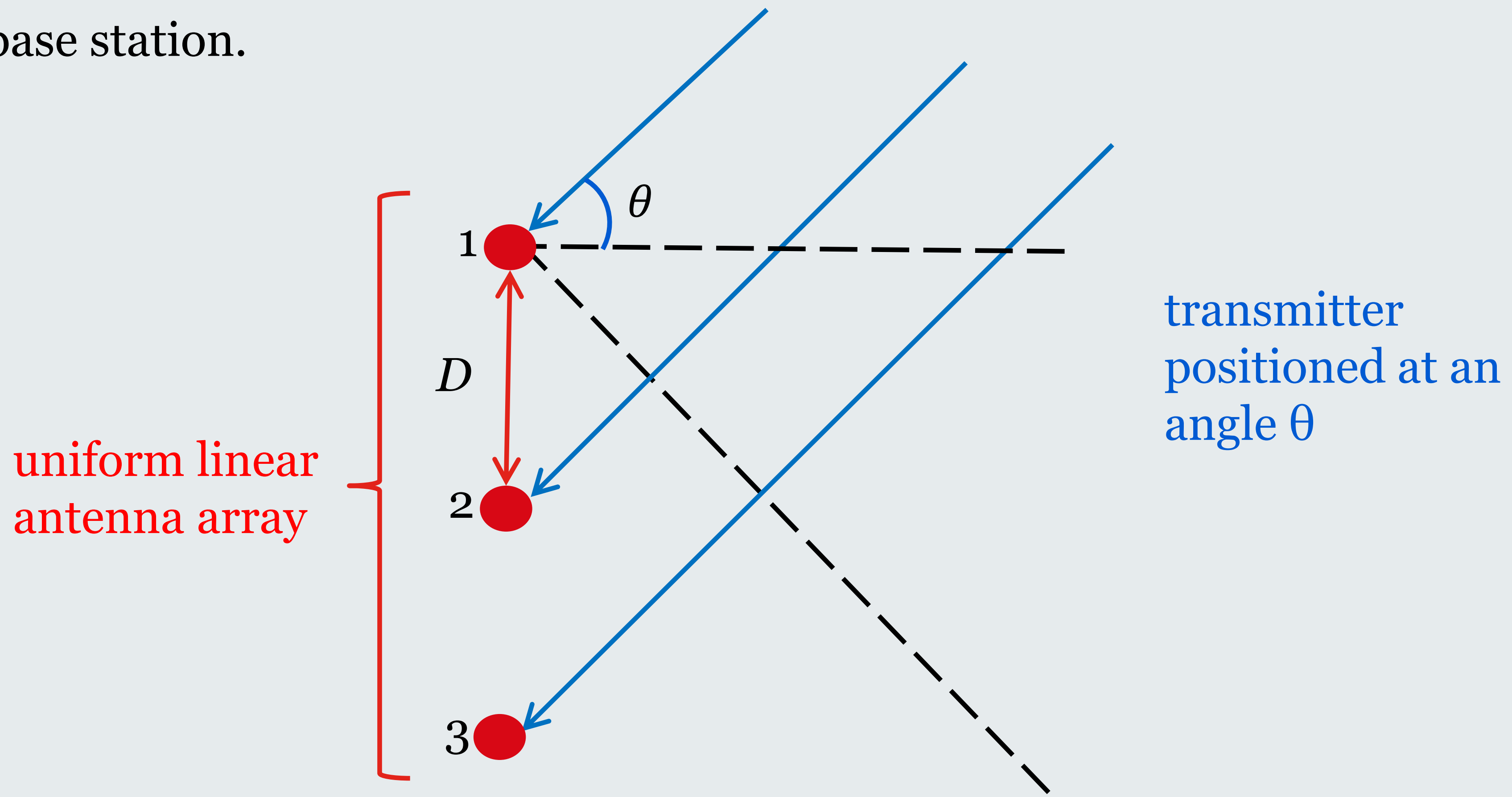
What is the Equivalent Baseband Model?

- Consider a single-antenna device communicating to a multi-antenna base station.



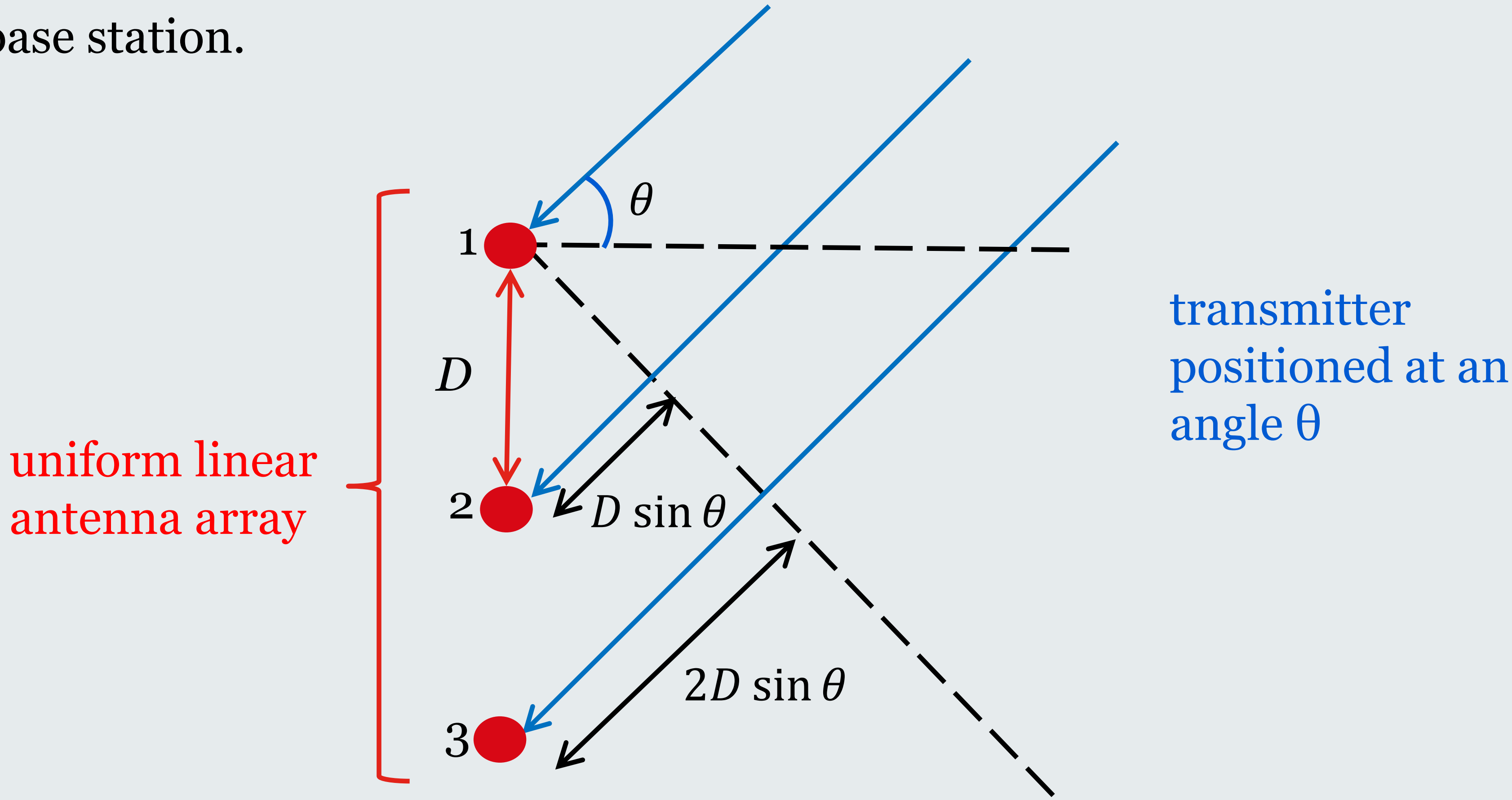
What is the Equivalent Baseband Model?

- Consider a single-antenna device communicating to a multi-antenna base station.



What is the Equivalent Baseband Model?

- Consider a single-antenna device communicating to a multi-antenna base station.



- Relative propagation delay for antenna n

$$\tau_n = \frac{D \sin(\theta)}{c}(n-1)$$

What is the Equivalent Baseband Model?

- Passband signal received at first antenna

$$y_1(t) = \sqrt{2}\mathrm{Re}\{x_{bb}(t)\exp(j2\pi f_c t)\}$$

- Passband signal received at the *n*th antenna

$$\begin{aligned} y_n(t) &= \sqrt{2}\mathrm{Re}\{x_{bb}(t-\tau_n)\exp(j2\pi f_c(t-\tau_n))\} \\ &\cong \sqrt{2}\mathrm{Re}\{x_{bb}(t)\exp(j2\pi f_c(t-\tau_n))\} \\ &= \sqrt{2}\mathrm{Re}\{x_{bb}(t)\exp(-j2\pi f_c\tau_n)\exp(j2\pi f_c t)\} \end{aligned}$$

What is the Equivalent Baseband Model?

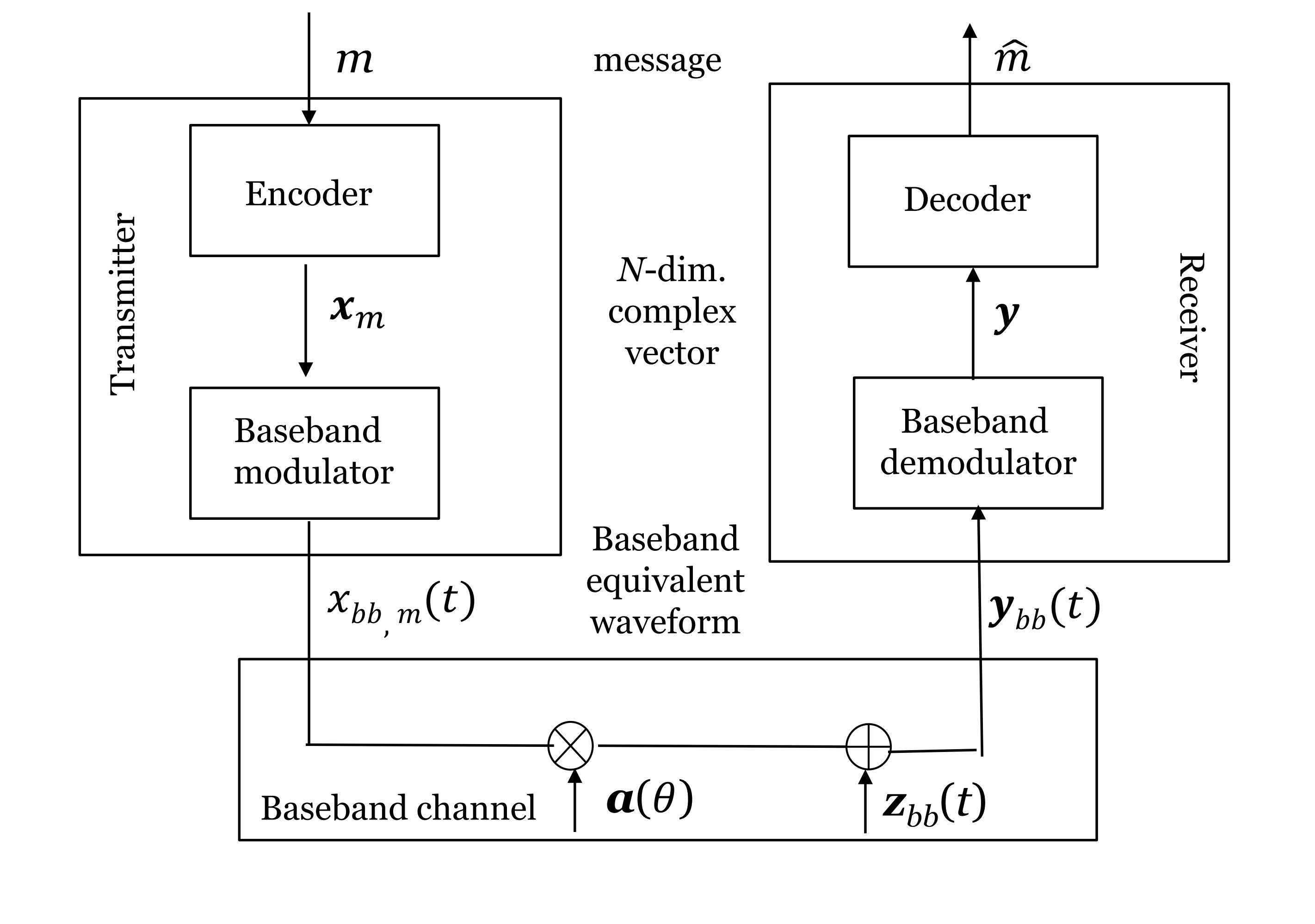
- Baseband signal received at the nth antenna

$$
\begin{aligned}
y_{bb,n}(t) &= x_{bb}(t)\exp(-j2\pi f_c \tau_n) \\
&= x_{bb}(t)\exp\left(-j2\pi f_c \frac{D\sin(\theta)}{c}(n-1)\right) \\
&= x_{bb}(t)\exp\left(-j2\pi \frac{D\sin(\theta)}{\lambda}(n-1)\right)
\end{aligned}
$$

- Define the steering vector (N_a = number of antennas)

$$
\mathbf{a}(\theta) = \begin{bmatrix} 1 \\ \exp\left(-j2\pi\frac{D\sin(\theta)}{\lambda}\right) \\ \exp\left(-j4\pi\frac{D\sin(\theta)}{\lambda}\right) \\ \vdots \\ \exp\left(-j2(N_a-1)\pi\frac{D\sin(\theta)}{\lambda}\right) \end{bmatrix}
$$

What is the Equivalent Baseband Model?



How to Demodulate and Decode?

- Consider transmission in the absence of ISI and communication using a time-domain waveform

$$x_{bb,m}(t) = x_m \varphi(t)$$

- Received signals on the antennas

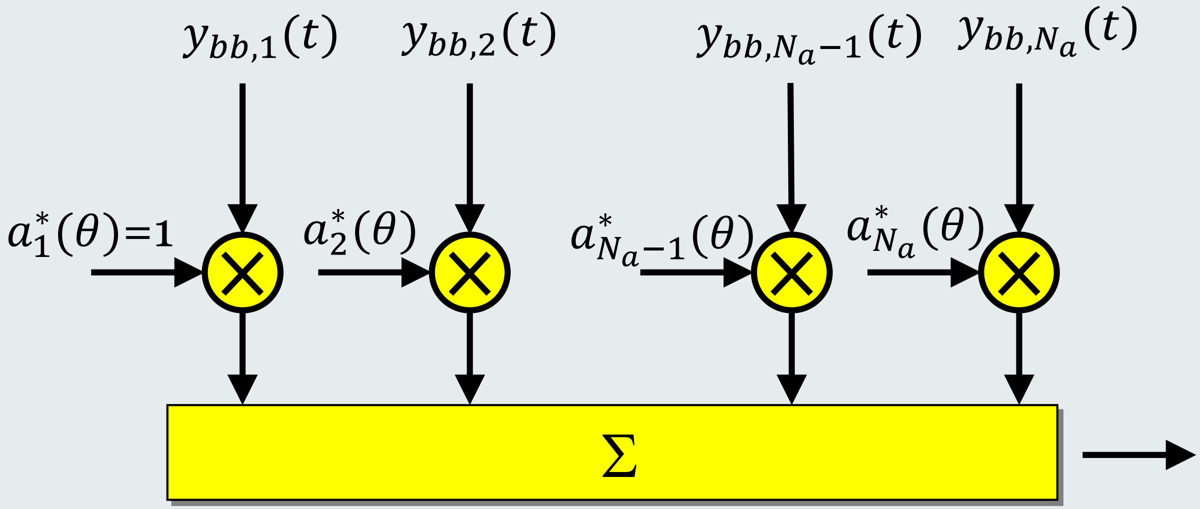
$$\mathbf{y}_{bb}(t) = x_m \varphi(t) \mathbf{a}(\theta) + \mathbf{z}_{bb}(t)$$

where

$$\mathbf{y}_{bb}(t) = \begin{bmatrix} y_{bb,1}(t) \\ y_{bb,2}(t) \\ \vdots \\ y_{bb,N_a}(t) \end{bmatrix}$$

How to Demodulate and Decode?

- Maximum ratio combining maximizes the signal-to-noise ratio



How to Demodulate and Decode?

- Consider transmission in the absence of ISI and communication using a time-domain waveform

$$x_{bb,m}(t) = x_m \varphi(t)$$

$$\mathbf{y}_{bb}(t) = \mathbf{a}(\theta) x_m \varphi(t) + \mathbf{z}_{bb}(t)$$

- Maximum ratio combining maximizes the signal-to-noise ratio

$$\begin{aligned}
\mathbf{a}^{\dagger}(\theta)\mathbf{y}_{bb}(t) &= \mathbf{a}^{\dagger}(\theta)\mathbf{a}(\theta) x_m \varphi(t) + \mathbf{a}^{\dagger}(\theta)\mathbf{z}_{bb}(t) \\
&= ||\mathbf{a}(\theta)||^2 x_m \varphi(t) + \mathbf{a}^{\dagger}(\theta)\mathbf{z}_{bb}(t) \\
&= N_a x_m \varphi(t) + \mathbf{a}^{\dagger}(\theta)\mathbf{z}_{bb}(t)
\end{aligned}$$

How to Demodulate and Decode?

- It can be proved that

$$
\begin{aligned}
Re\{\mathbf{a}^{\dagger}(\theta)\mathbf{z}_{bb}(t)\} &\sim \mathcal{N}(0, N_a \times N_0/2) \\
Im\{\mathbf{a}^{\dagger}(\theta)\mathbf{z}_{bb}(t)\} &\sim \mathcal{N}(0, N_a \times N_0/2)
\end{aligned}
$$

since the power of independence random variables is equal to the sum of the individual powers.

What is the Performance Gain?

- It can be proved that

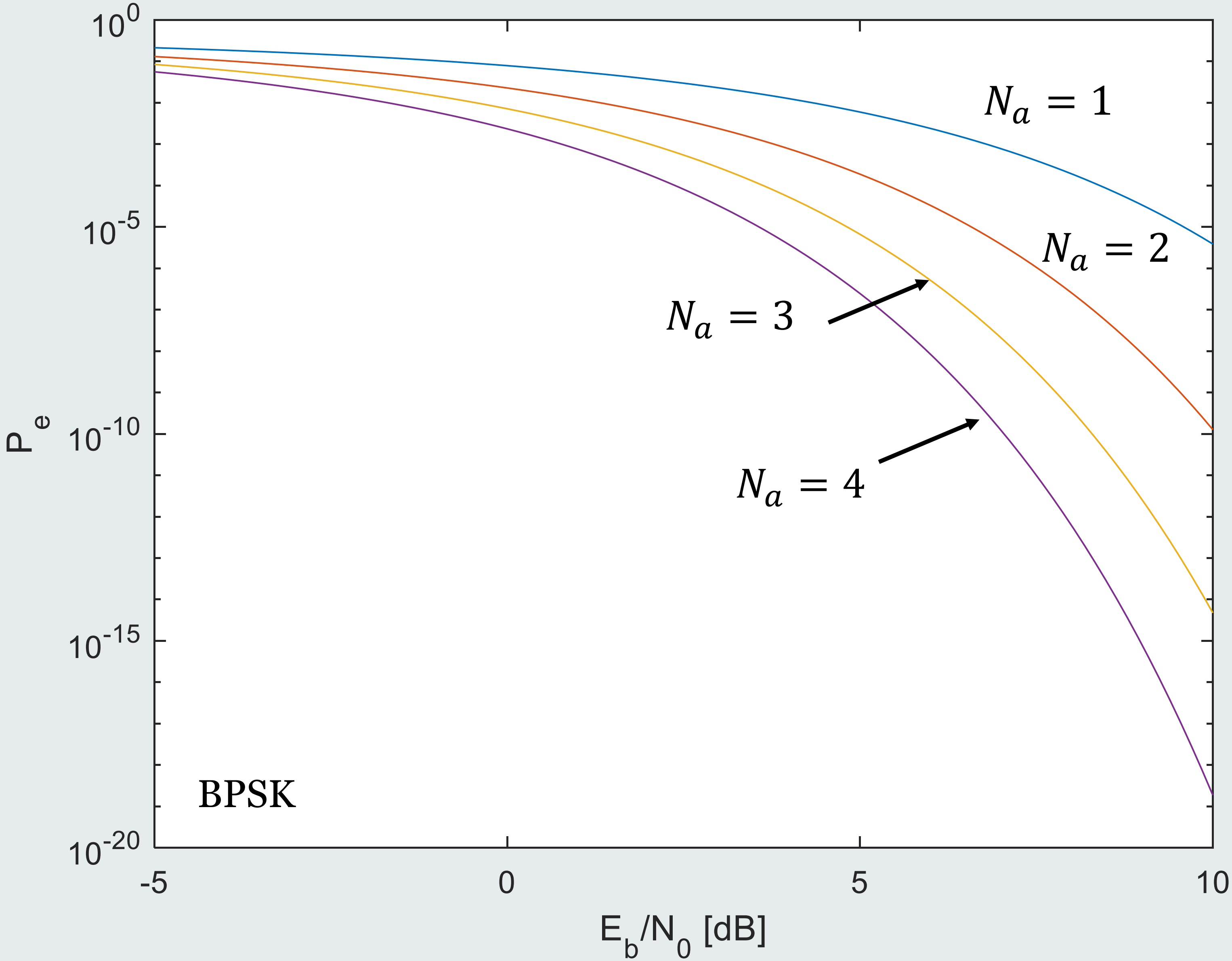
$$Re\{\mathbf{a}^{\dagger}(\theta)\mathbf{z}_{bb}(t)\} \sim \mathcal{N}(0, N_a \times N_0/2)$$
$$Im\{\mathbf{a}^{\dagger}(\theta)\mathbf{z}_{bb}(t)\} \sim \mathcal{N}(0, N_a \times N_0/2)$$

since the power of independence random variables is equal to the sum of t the individual powers.

- Hence, the SNR is improved by a factor equal to the number of antennas:

$$\frac{N_a^2 E_b}{N_a N_0} = N_a \times \frac{E_b}{N_0}$$

What is the Performance Gain?



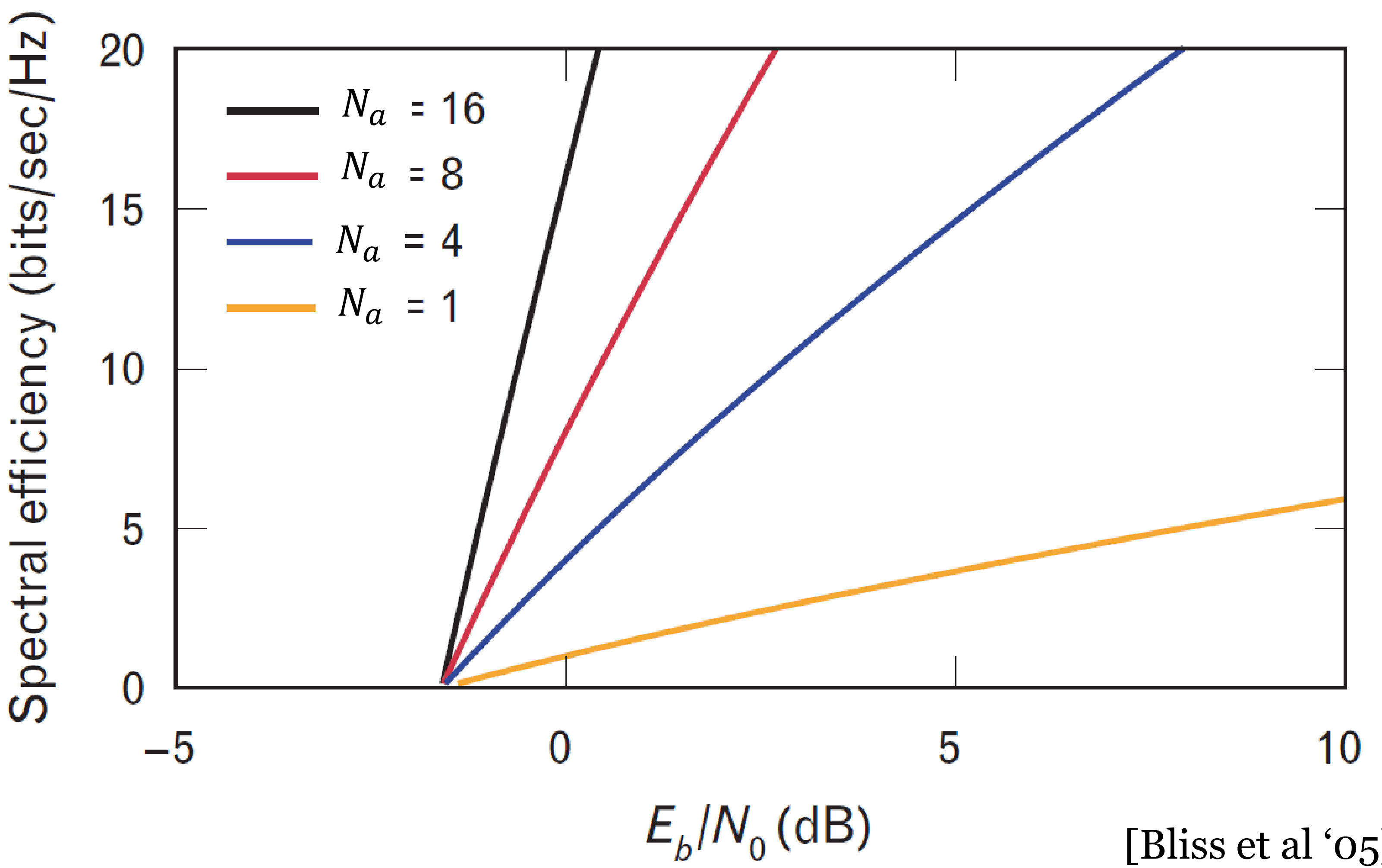
What Can We Gain with MIMO?

- With MIMO with N_a antennas at the transmitter and receiver, we can communicate N_a data streams simultaneously.

- In this way, the spectral efficiency is multiplied by N_a.

What Can We Gain with MIMO?

- MIMO provides a multiplexing gain and not merely a coding gain.



[Bliss et al '05]