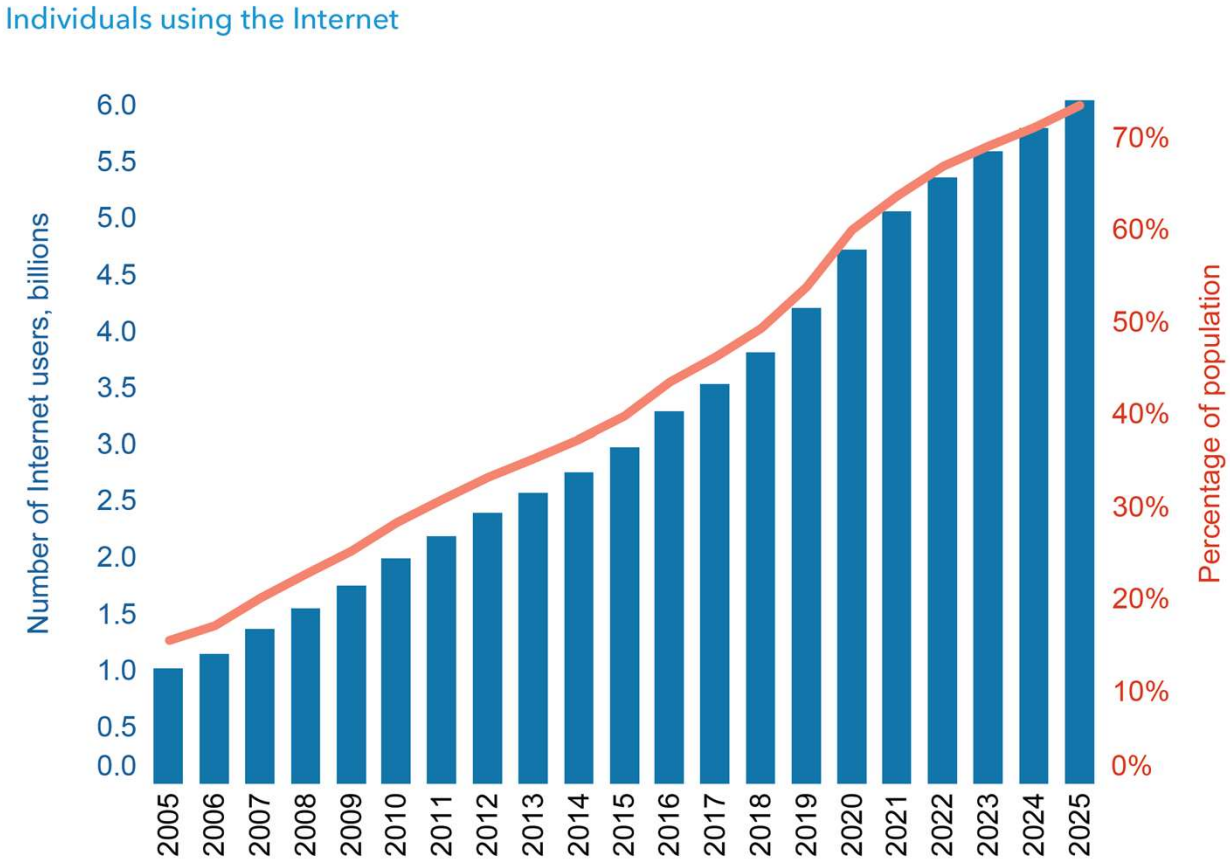
1014I – Communication systems and cybersecurity (2025/26)

A primer on digital communication systems

Giacomo Bacci
giacomo.bacci@unipi.it

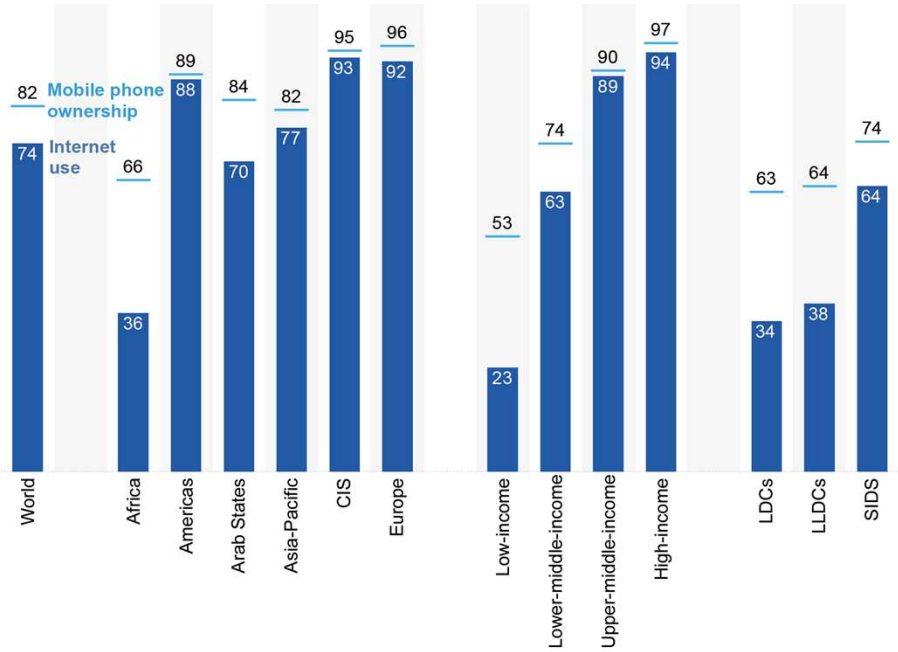
Introduction

Some (whopping) facts on communications (1/5)

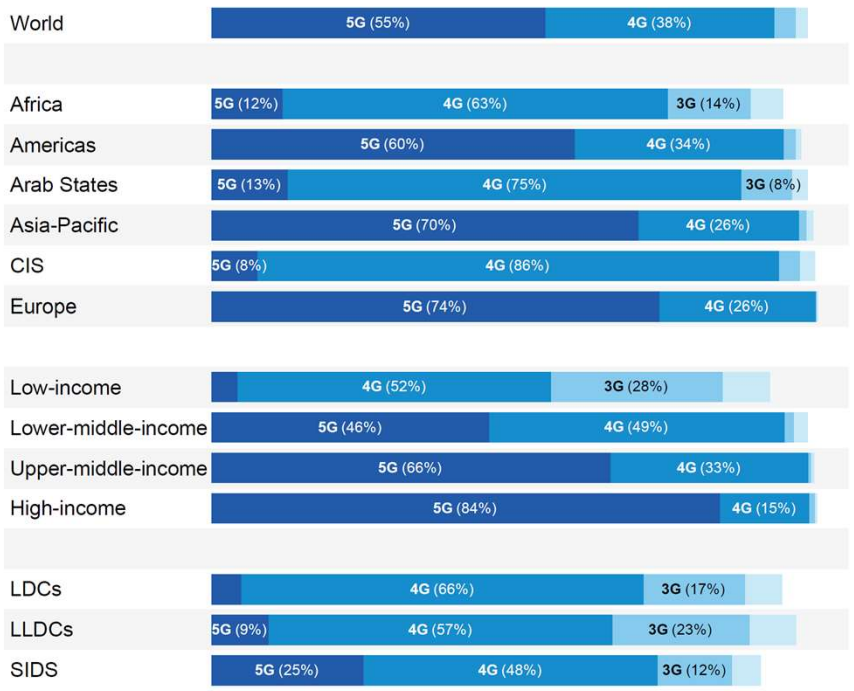


Some (whopping) facts on communications (2/5)

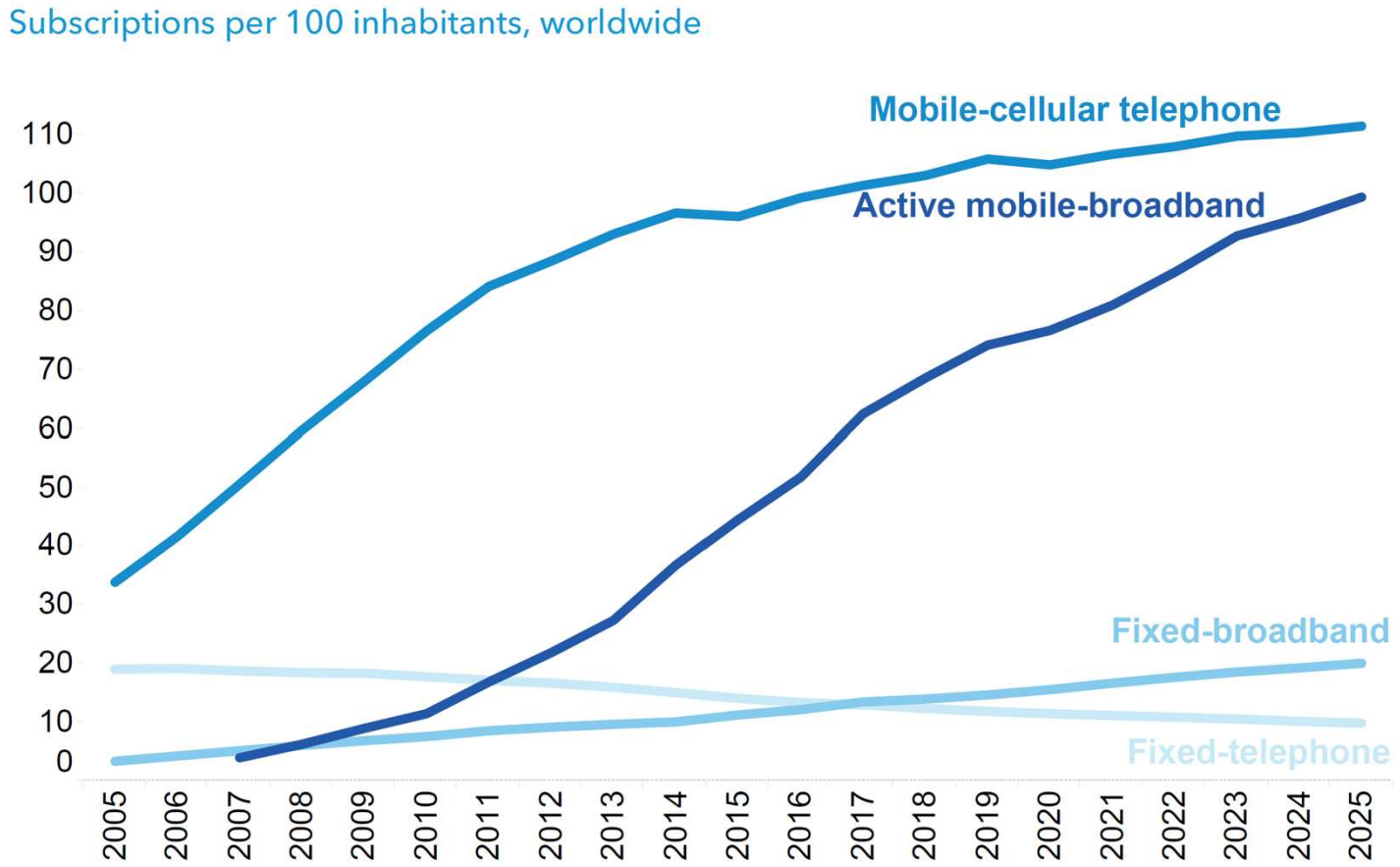
Percentage of individuals owning a mobile phone and using the Internet, 2025



Population coverage by type of mobile network, 2025

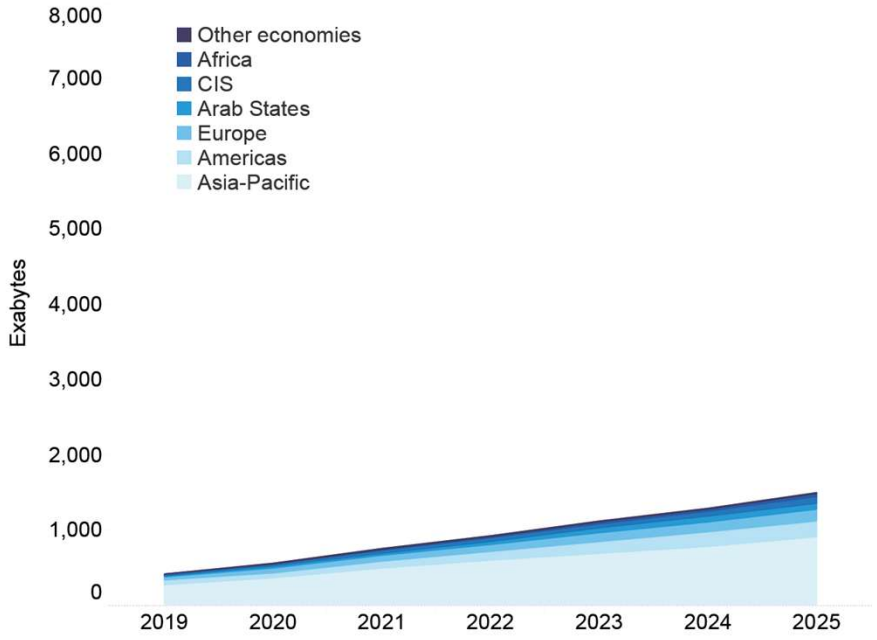


Some (whopping) facts on communications (3/5)

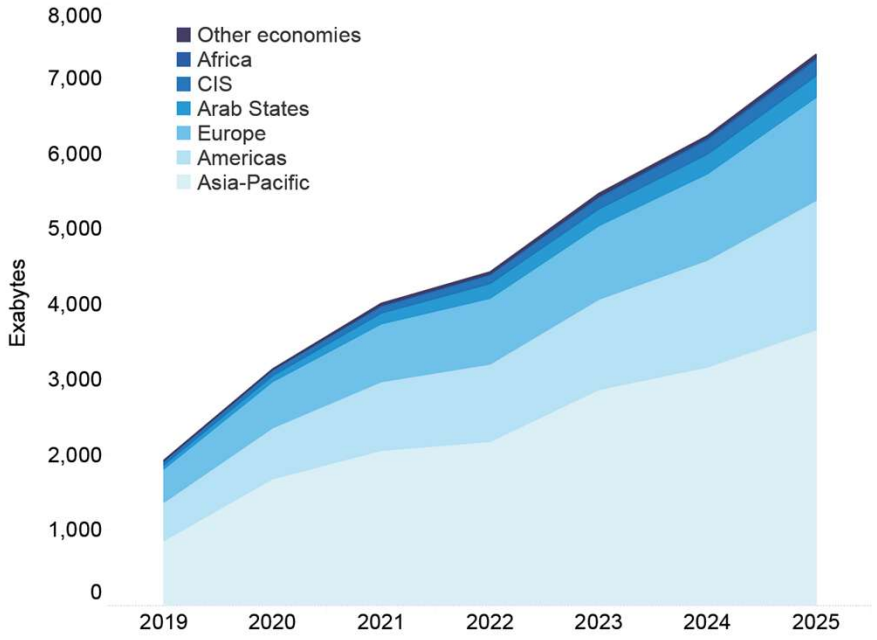


Some (whopping) facts on communications (4/5)

Mobile broadband traffic, EB



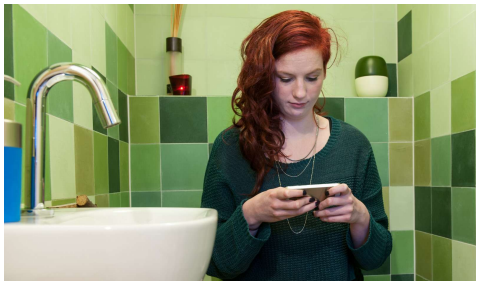
Fixed broadband traffic, EB



1 EB = 10^{18} bytes = 10^{12} MB

Some (whopping) facts on communications (5/5)

- The number of people with access to mobile communications is higher than those with access to working toilets (around 4.6 billions, source: World Health Organization)



- The number of people that owns a mobile phone is larger than the one that owns/uses a toothbrush (around 4 billions)

- Every second, more than eight hours of videos is uploaded to Youtube (as of 2022)

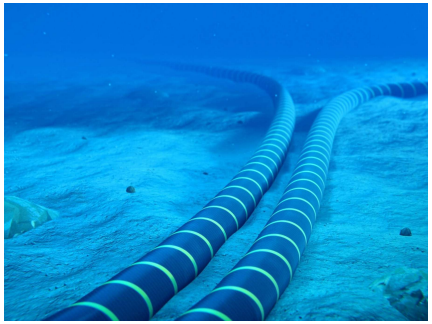
Outline of the lecture

- Overview of the architecture
- Building blocks of transmitter and receiver
- Analog-to-digital conversion
- Source coding
- Channel coding
- Modulation techniques
- Shannon capacity
- Multiplexing and multiple access

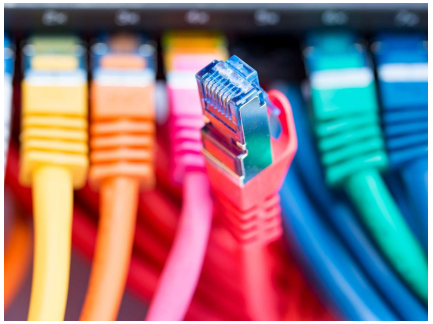
Categories of communication systems (1/3)

- Access vs. transport

vs.



- Wired vs. wireless



vs.

Categories of communication systems (2/3)

- Broadcast (point-to-multipoint) vs. unicast (point-to-point)

vs.

- Broadband vs. narrowband

vs.

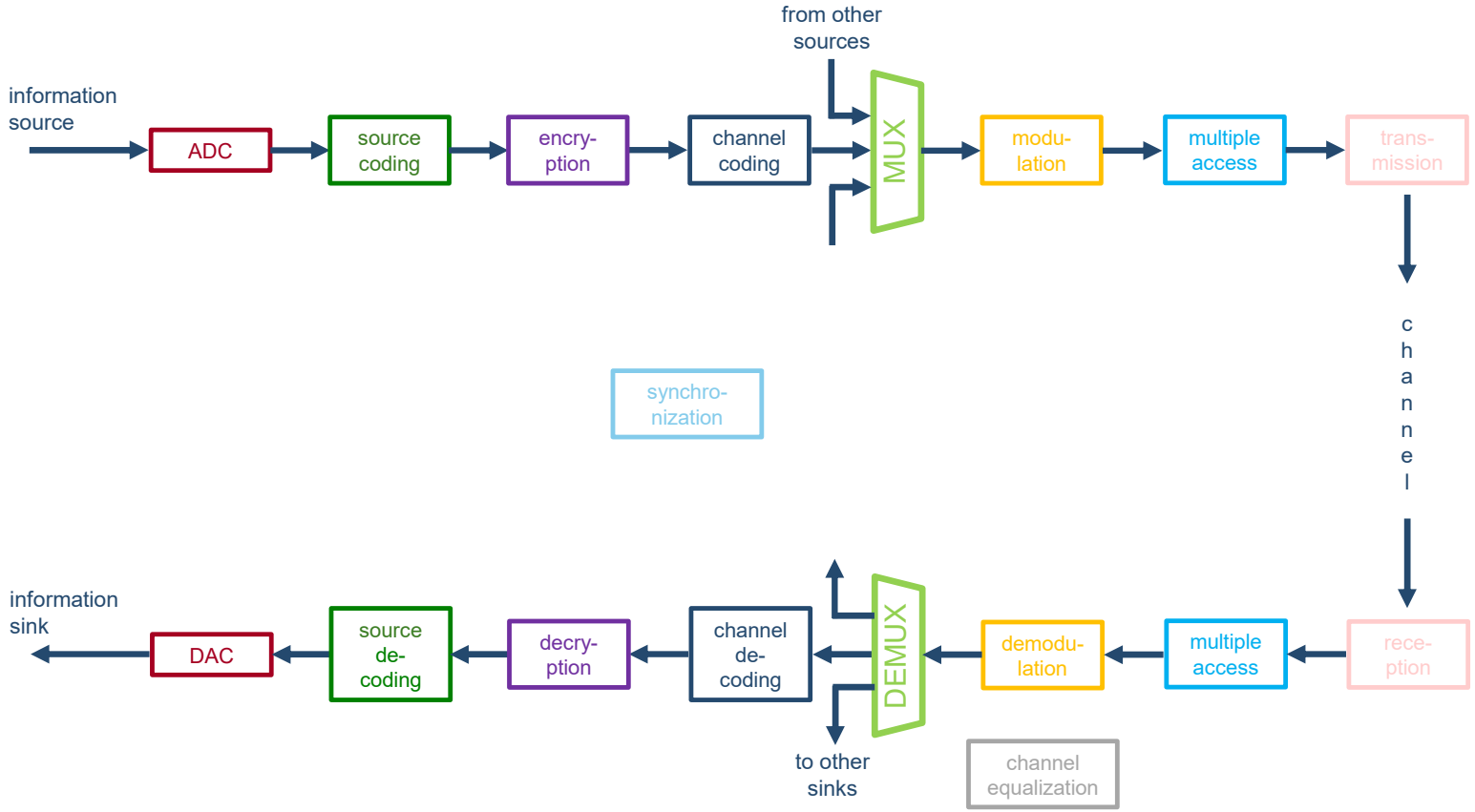
Categories of communication systems (3/3)

- Terrestrial vs. satellite

vs.

… and many more

Elements of a digital communication system

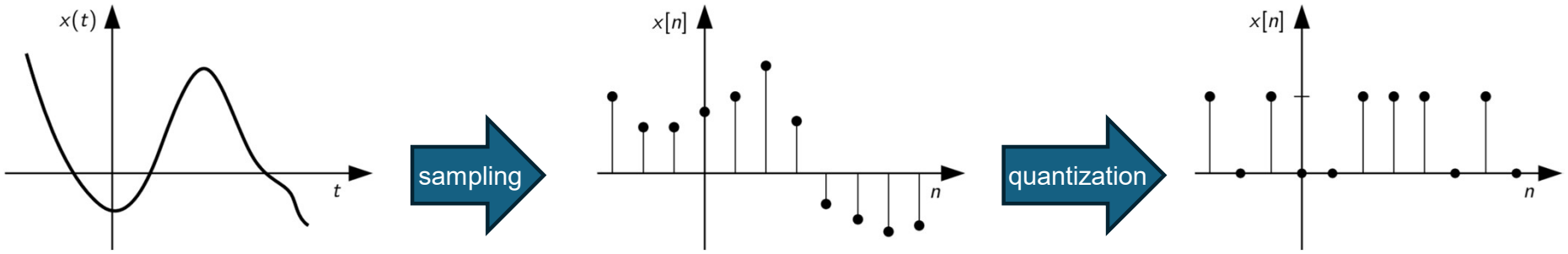


Analog-to-digital conversion (ADC)

Analog-to-digital conversion (ADC)

The ADC operation consists into two steps:

- Sampling, which makes the signal a sequence of values: time becomes discrete, values of the sequence are still in the real domain
- Quantization, which represents the values of the sequence with a finite number of bits: values now belong to a finite set

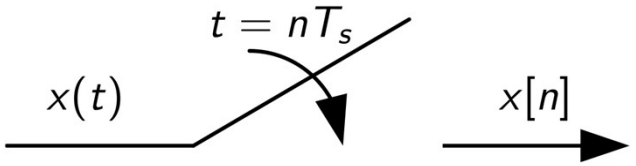


analog signal sequence digital signal

Sampling and interpolation

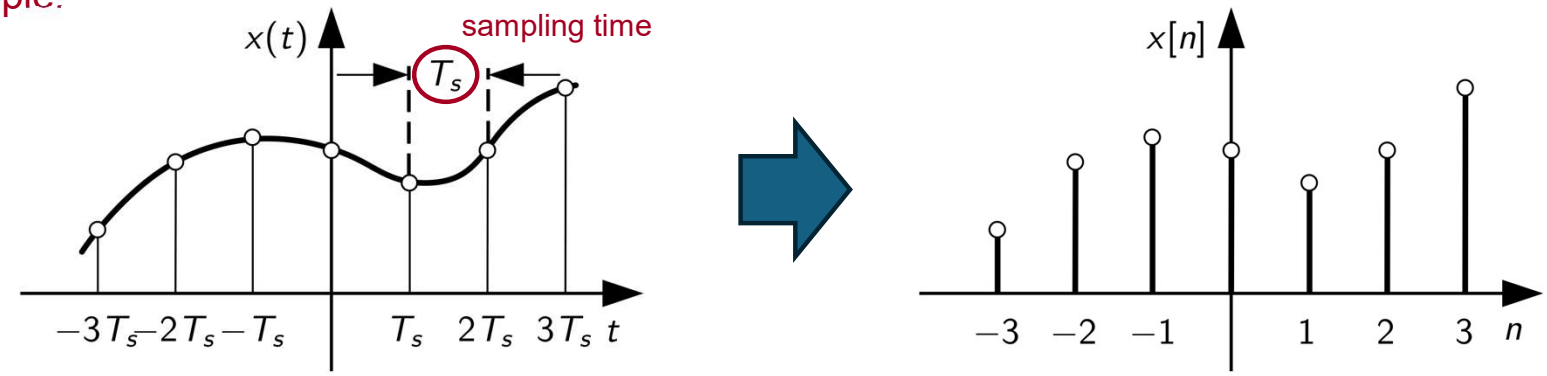
Sampling an analog signal

Sampling an analog signal $x(t)$ means collecting, one after the other, the sequence $x[n]$ of values taken by that signal at time instants nT_s:

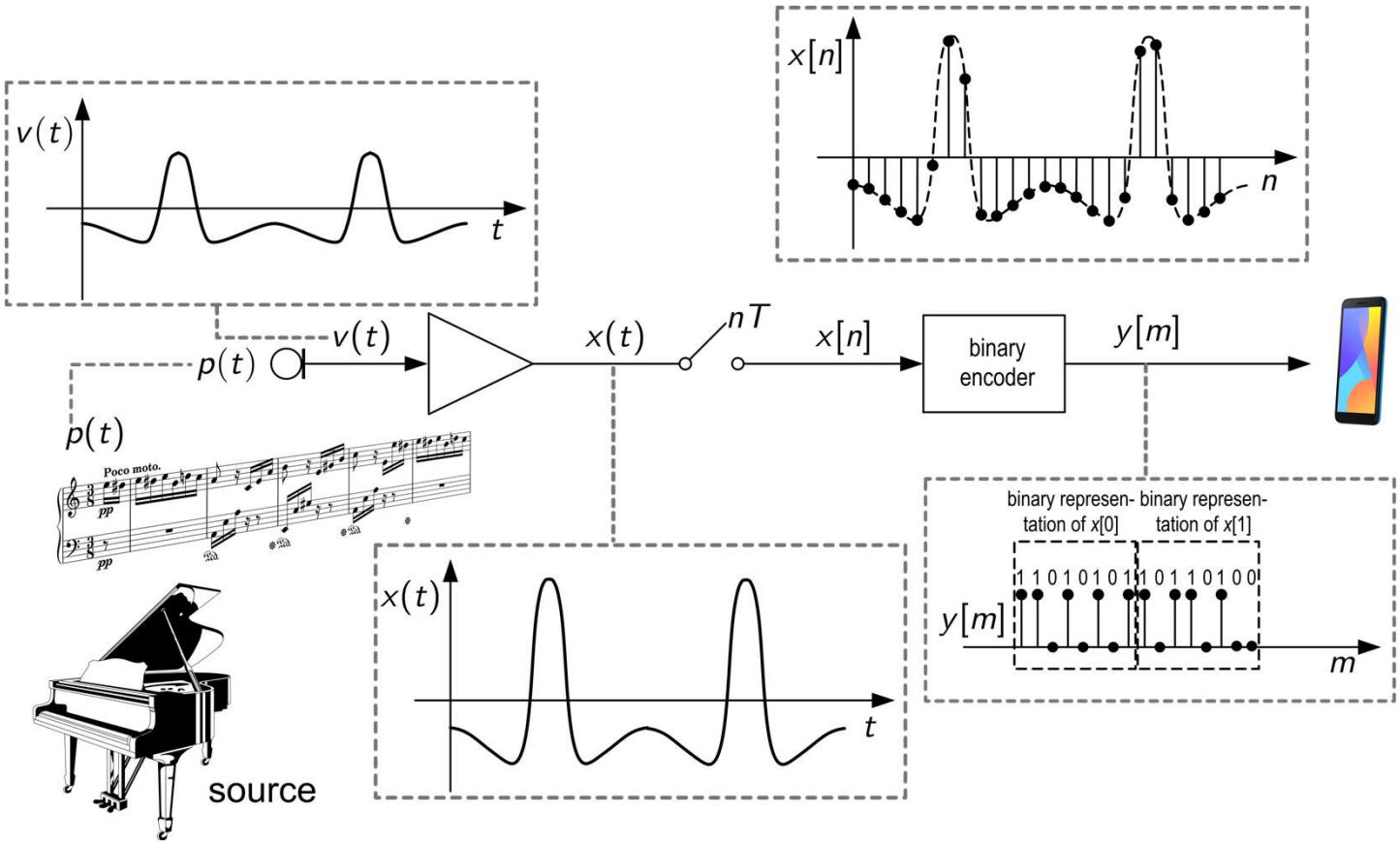


sampling frequency $f_s = 1/T_s$

Example:

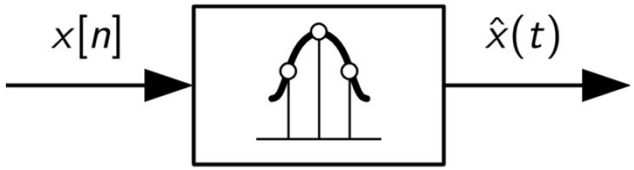


Example: digital recording of an audio signal

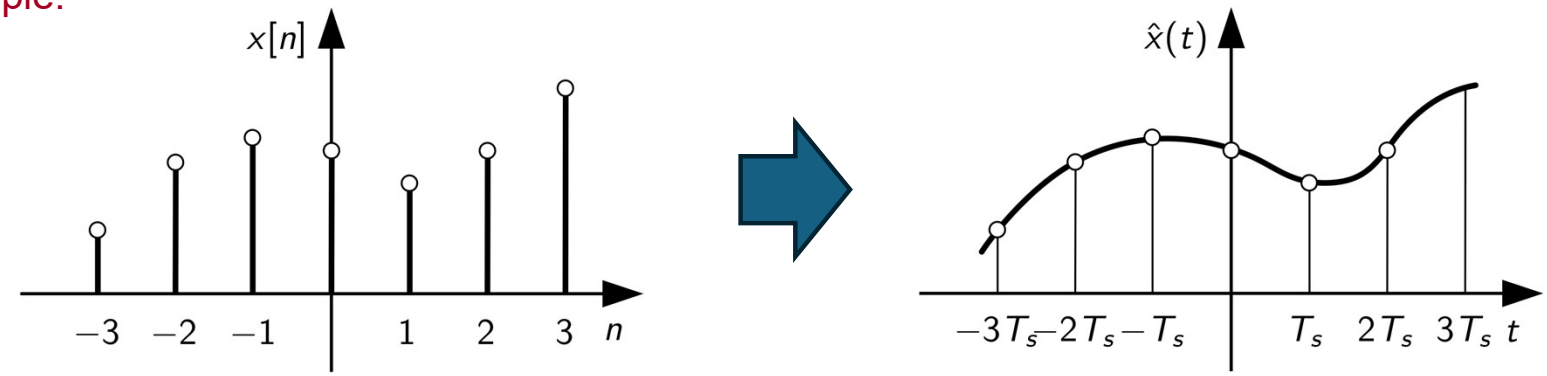


Back to the analog domain: interpolation

To reconstruct, with sufficiently high fidelity, the original signal starting from the sequence of samples, we can perform the operation of interpolation:

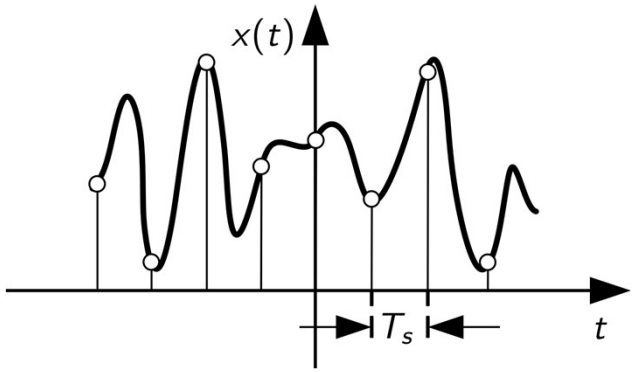


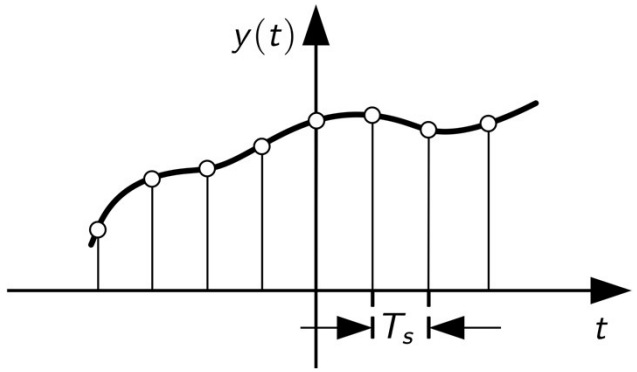
Example:



Sampling frequency (1/2)

The natural question is: how can we properly set the sampling frequency f_s?





The sampling period is adequate for $y(t)$, but it is clearly too large for $x(t)$!

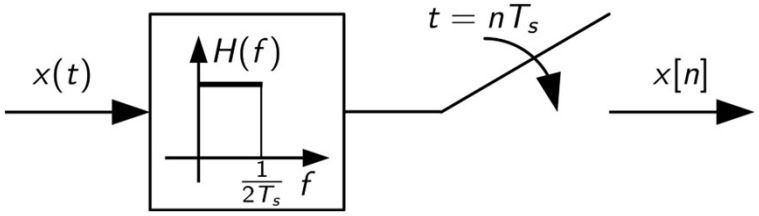
Note: increasing f_s (or equivalently, reducing T_s) means increasing the number of samples per unit of time, with drawbacks in terms of hardware and storage requirements

Sampling frequency (2/2)

To avoid aliasing errors when sampling a strictly bandlimited signal with bandwidth *B*, we need to satisfy the Nyquist condition:

$$f_s = \frac{1}{T_s} \geq 2B$$

In case of band-unlimited signals, we can place an anti-aliasing (low-pass) filter to limit the bandwidth of the analog signal, so as to avoid the aliasing error

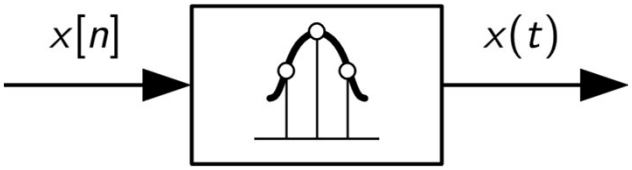


Interpolation

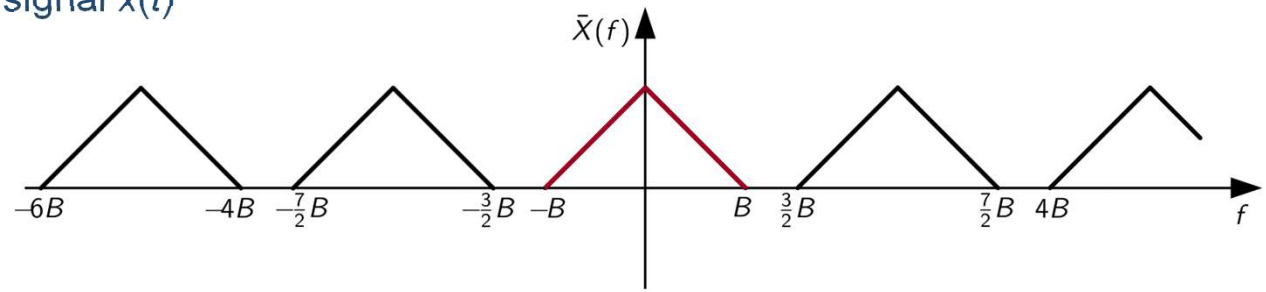
Interpolation (1/3)

How can we reconstruct the analog signal $x(t)$ starting from the sequence of samples $x[n]$?

This operation is called interpolation and is performed by the digital-to-analog converter (DAC)



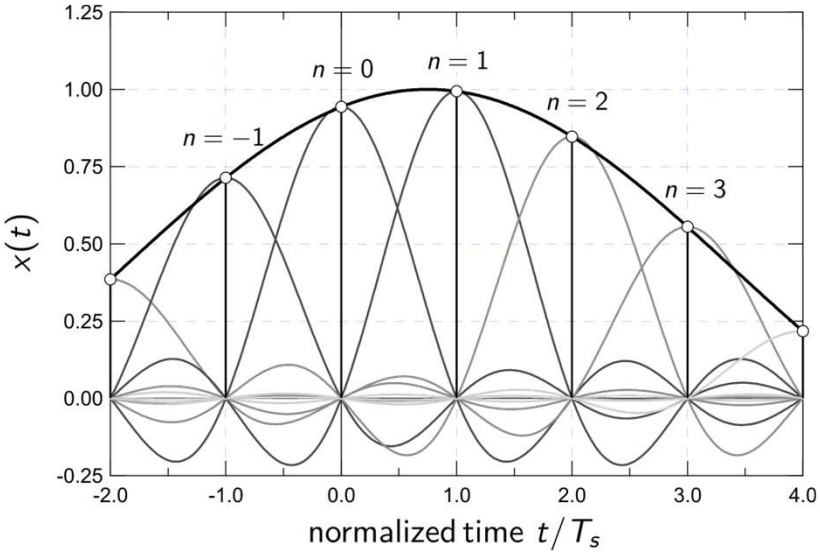
Ideally, we can apply a low-pass filtering on $x[n]$ to obtain back the original spectrum – and hence the original signal $x(t)$



Interpolation (2/3)

This goes under the name of cardinal interpolation, and is the result of the Shannon's sampling theorem:

$$x(t) = \sum_{n=-\infty}^{+\infty} x[n] \, \mathrm{sinc}\left(\frac{t - nT_s}{T_s}\right)$$

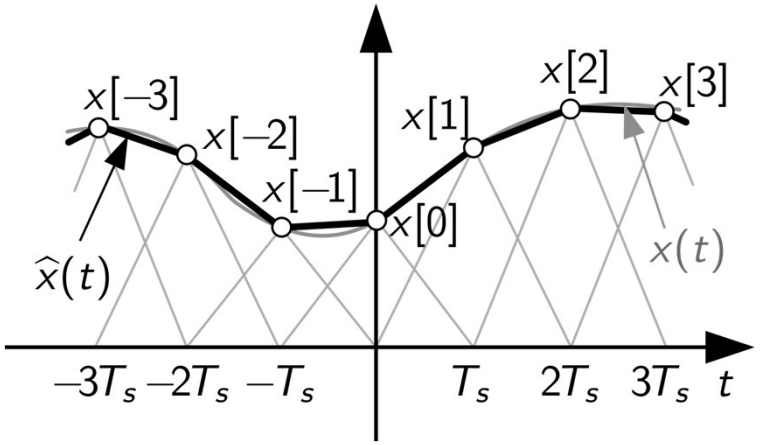


Interpolation (3/3)

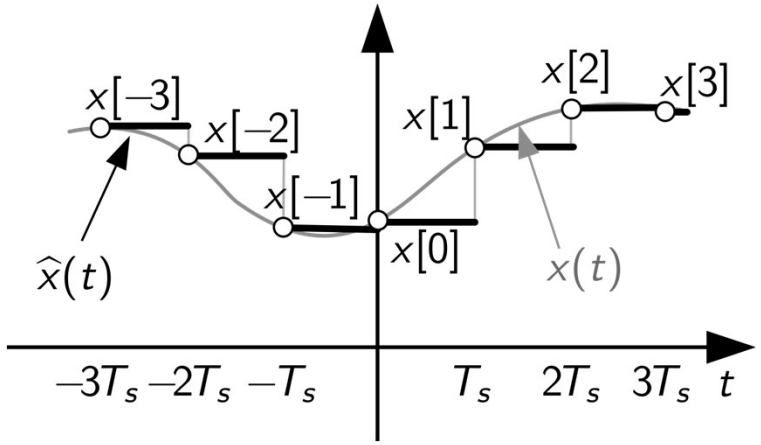
The cardinal interpolator calls for an infinite number of samples, including future ones.

Simpler methods, yet achieving good results that provide an approximated version $\hat{x}(t)$ of the original signal, include:

- linear interpolation



- sample & hold (S&H) interpolation



Quantization

Quantization (1/3)

To transform an analog signal into a digital one, the job is not done just after sampling: we need to express the sample over a finite number of bits

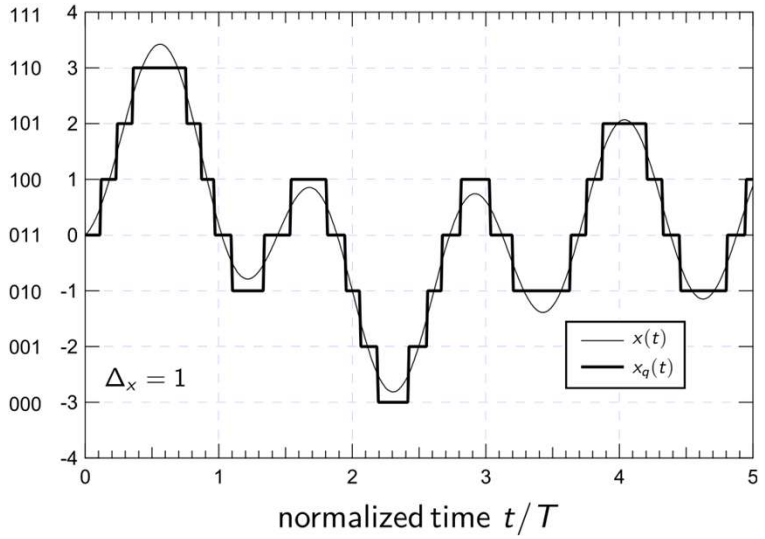
This task is called quantization, and works as follows:

$$x_q[n] = \Delta_x \cdot \left\lfloor \frac{x[n]}{\Delta_x} \right\rfloor$$

rounding

quantization step: $\Delta_x = \frac{D_x}{2^b - 1}$

b: number of quantization bits



Quantization (2/3)

The impact of a different number of bits (i.e., number of intensity levels) can be easily visualized in a 2D image:

8 bits / pixel

4 bits / pixel

2 bits / pixel

Quantization (3/3)

We can also experience it in a musical song (*Help!* by The Beatles):

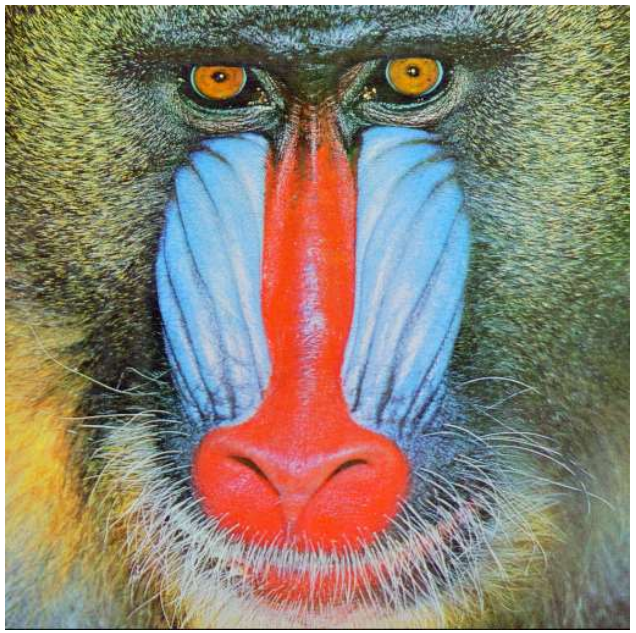
- 16 bits/sample:
- 4 bits/sample:
- 2 bits/sample:

Source coding

Source coding

Source coding aims at data compression, involving encoding information to use fewer bits than the original representation:

- Lossless compression identifies and eliminates statistical redundancy
- Lossy compression identifies and removes unnecessary information



Examples of lossless compression

- fixed-to-fixed code: ASCII code
- fixed-to-variable code: Morse code
- variable-to-fixed code: Lempel-Ziv-Welch algorithm

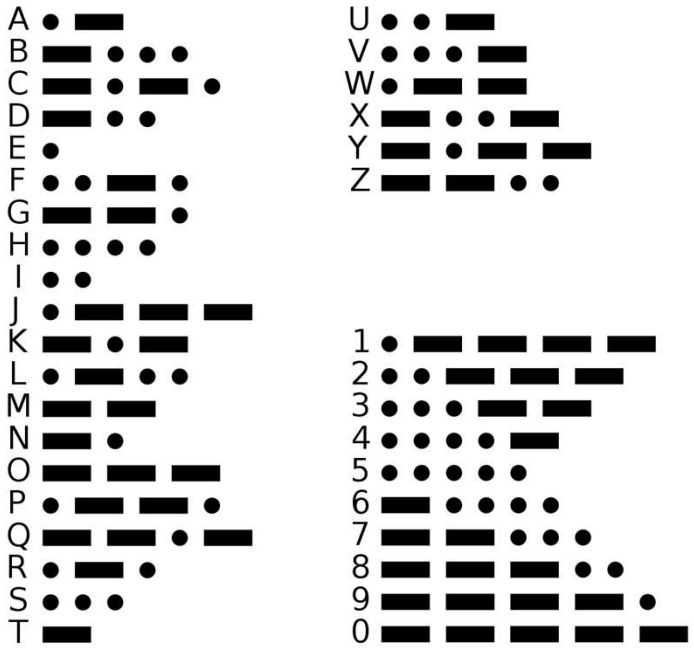
An example of lossless compression: The ASCII code

dec	hex	oct	char	dec	hex	oct	char	dec	hex	oct	char	dec	hex	oct	char
0	0	000	NULL	32	20	040	space	64	40	100	@	96	60	140	`
1	1	001	SOH	33	21	041	!	65	41	101	A	97	61	141	a
2	2	002	STX	34	22	042	"	66	42	102	B	98	62	142	b
3	3	003	ETX	35	23	043	#	67	43	103	C	99	63	143	c
4	4	004	EOT	36	24	044	$	68	44	104	D	100	64	144	d
5	5	005	ENQ	37	25	045	%	69	45	105	E	101	65	145	e
6	6	006	ACK	38	26	046	&	70	46	106	F	102	66	146	f
7	7	007	BEL	39	27	047	'	71	47	107	G	103	67	147	g
8	8	010	BS	40	28	050	(	72	48	110	H	104	68	150	h
9	9	011	TAB	41	29	051	)	73	49	111	I	105	69	151	i
10	a	012	LF	42	2a	052	*	74	4a	112	J	106	6a	152	j
11	b	013	VT	43	2b	053	+	75	4b	113	K	107	6b	153	k
12	c	014	FF	44	2c	054	,	76	4c	114	L	108	6c	154	l
13	d	015	CR	45	2d	055	-	77	4d	115	M	109	6d	155	m
14	e	016	SO	46	2e	056	.	78	4e	116	N	110	6e	156	n
15	f	017	SI	47	2f	057	/	79	4f	117	O	111	6f	157	o
16	10	020	DLE	48	30	060	0	80	50	120	P	112	70	160	p
17	11	021	DC1	49	31	061	1	81	51	121	Q	113	71	161	q
18	12	022	DC2	50	32	062	2	82	52	122	R	114	72	162	r
19	13	023	DC3	51	33	063	3	83	53	123	S	115	73	163	s
20	14	024	DC4	52	34	064	4	84	54	124	T	116	74	164	t
21	15	025	NAK	53	35	065	5	85	55	125	U	117	75	165	u
22	16	026	SYN	54	36	066	6	86	56	126	V	118	76	166	v
23	17	027	ETB	55	37	067	7	87	57	127	W	119	77	167	w
24	18	030	CAN	56	38	070	8	88	58	130	X	120	78	170	x
25	19	031	EM	57	39	071	9	89	59	131	Y	121	79	171	y
26	1a	032	SUB	58	3a	072	:	90	5a	132	Z	122	7a	172	z
27	1b	033	ESC	59	3b	073	;	91	5b	133	[	123	7b	173	{
28	1c	034	FS	60	3c	074	<	92	5c	134	\	124	7c	174	|
29	1d	035	GS	61	3d	075	=	93	5d	135	]	125	7d	175	}
30	1e	036	RS	62	3e	076	>	94	5e	136	^	126	7e	176	~
31	1f	037	US	63	3f	077	?	95	5f	137	_	127	7f	177	DEL

An example of lossless compression: The Morse code

International Morse Code

1. The length of a dot is one unit.
2. A dash is three units.
3. The space between parts of the same letter is one unit.
4. The space between letters is three units.
5. The space between words is seven units.

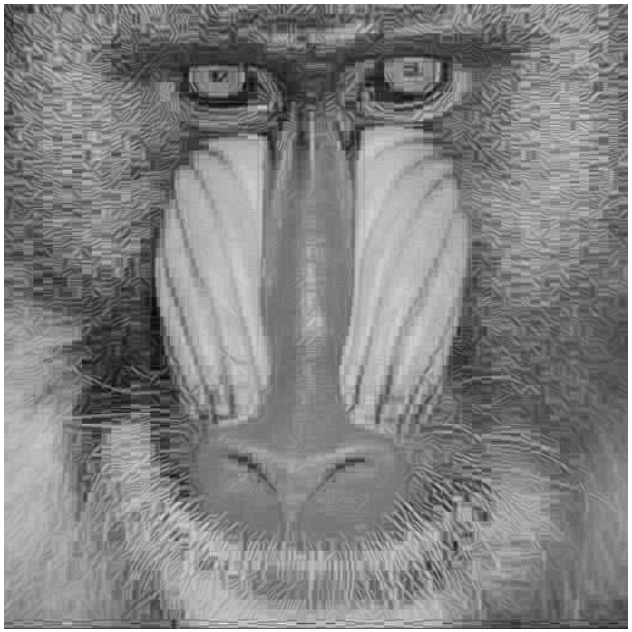


The Lempel-Ziv-Welch (LZW) algorithm

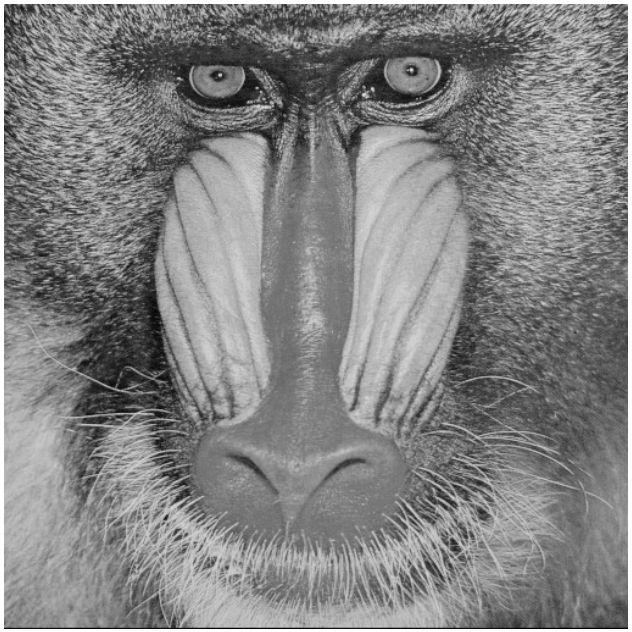
- Lempel and Ziv patented the algorithm in 1978 as LZ78
- Welch further improved the algorithm in 1984
- easy and universal implementation, widely used for zip (and its variants) and GIF formats
- the original LZW algorithm encodes sequences as fixed-length 12-bit sequences

An example of lossy compression: image coding

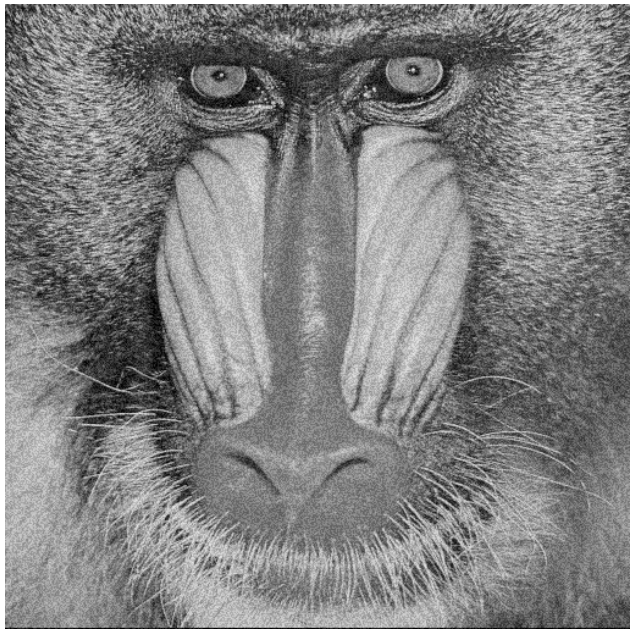
Image coding adopts a lossy compression, which is based on the specific properties of the human vision



altered #1



original

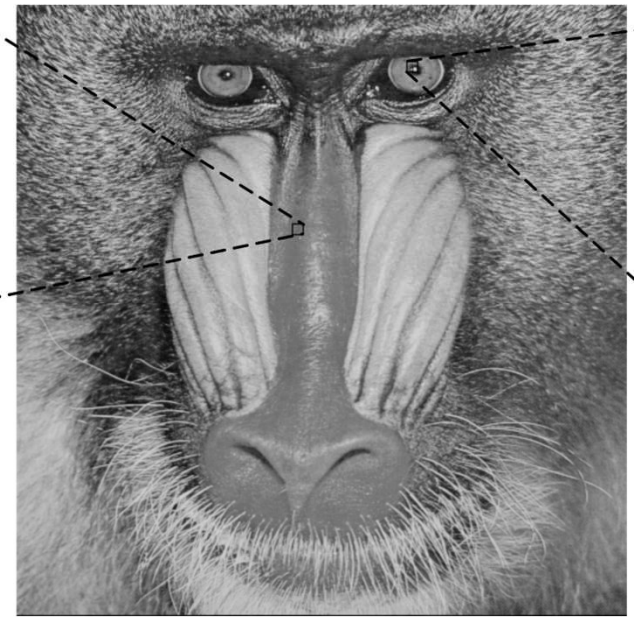


altered #2

Image representation

nose (intensity)

114	119	108	115	105	110	116	117
119	113	119	115	108	109	115	109
112	119	116	114	113	111	113	117
122	118	126	113	108	111	121	120
122	119	119	114	110	113	113	116
117	114	113	111	111	110	119	125
114	113	109	107	109	115	114	120
116	118	113	117	118	117	122	125



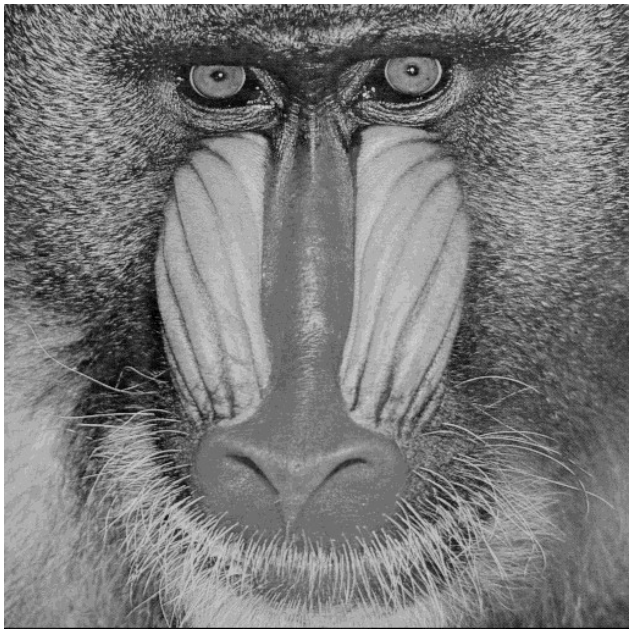
eye (intensity)

139	151	130	114	117	115	115	114
142	117	124	105	129	146	128	131
122	103	83	63	67	70	47	75
108	96	88	57	38	41	37	37
117	98	53	40	34	117	144	78
99	104	67	123	96	212	220	188
114	101	133	203	96	209	218	177
110	78	41	55	36	104	150	104

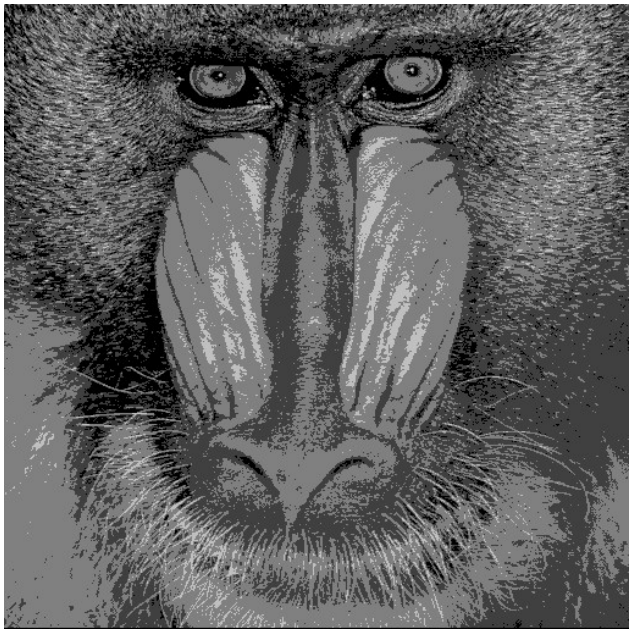
8 bits / pixel

Quantization in the space domain

Raw quantization:



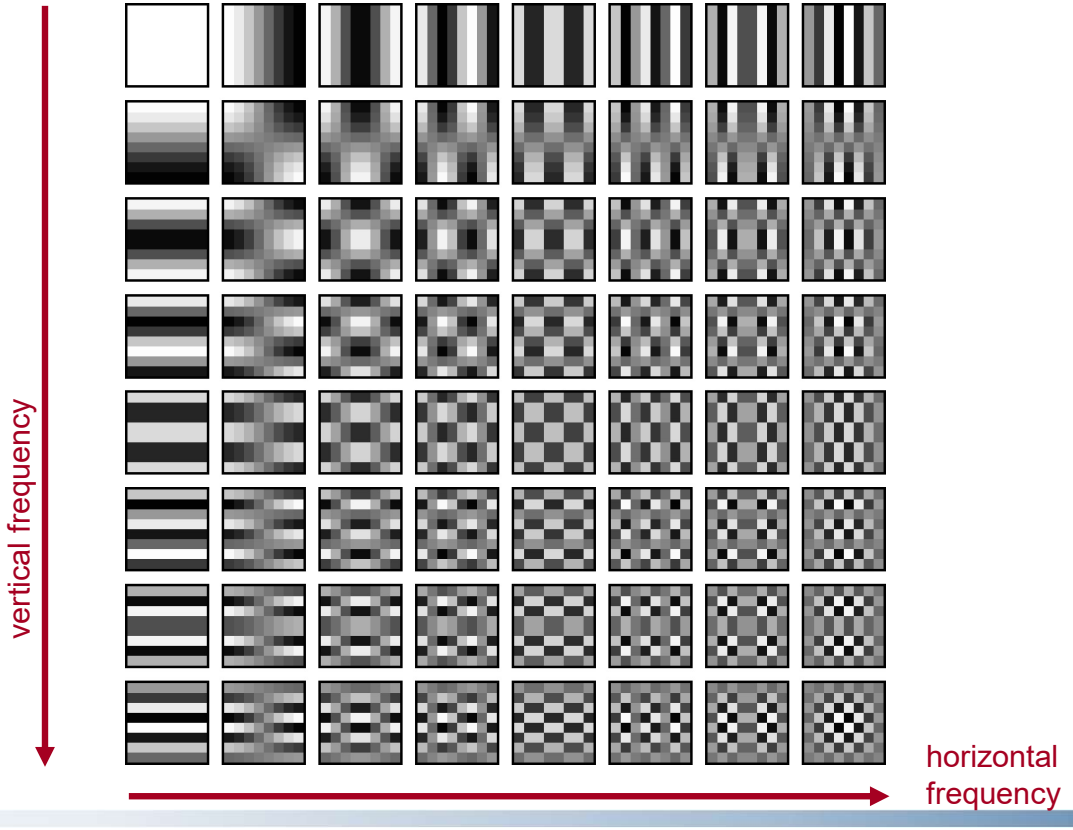
4 bits / pixel



2 bits / pixel

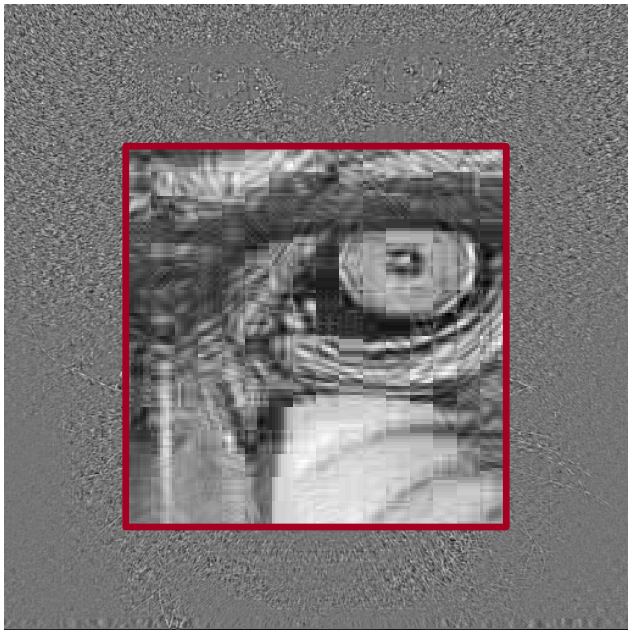
Discrete cosine transform (DCT) (1/3)

Contribution of DCT coefficients to the appearance of the 8×8 pixel block:



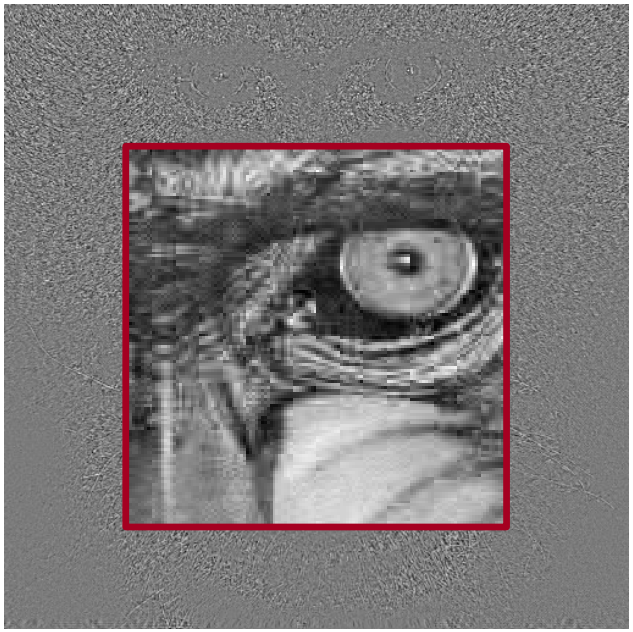
Discrete cosine transform (DCT) (2/3)

DFT



vs.

DCT



Discrete cosine transform (DCT) (3/3)

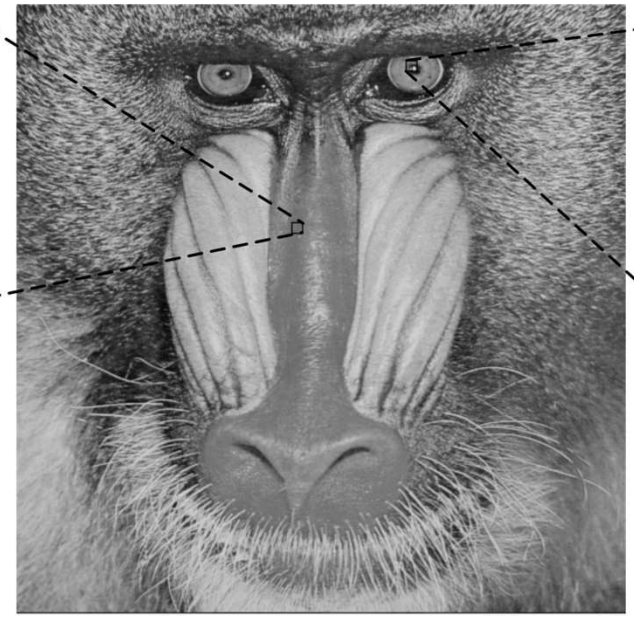
Image analysis using DCT:

nose (intensity)

114	119	108	115	105	110	116	117
119	113	119	115	108	109	115	109
112	119	116	114	113	111	113	117
122	118	126	113	108	111	121	120
122	119	119	114	110	113	113	116
117	114	113	111	111	110	119	125
114	113	109	107	109	115	114	120
116	118	113	117	118	117	122	125

nose (DCT spectrum)

920	1	19	−9	−2	3	−1	−1
−6	11	−3	−3	−3	5	−2	−1
−4	−8	−2	4	2	−1	−6	−6
−7	−1	6	1	2	1	0	−3
9	2	2	−1	−1	2	−1	−2
−3	−5	5	−1	−2	−3	0	−2
3	−6	0	−3	3	−5	−6	−2
−5	1	−3	3	0	−6	−1	−5



eye (intensity)

139	151	130	114	117	115	115	114
142	117	124	105	129	146	128	131
122	103	83	63	67	70	47	75
108	96	88	57	38	41	37	37
117	98	53	40	34	117	144	78
99	104	67	123	96	212	220	188
114	101	133	203	96	209	218	177
110	78	41	55	36	104	150	104

eye (DCT spectrum)

858	−47	90	51	−52	53	4	−47
−32	129	−27	−18	32	−72	17	48
116	−35	−16	−11	11	9	−4	−1
155	−91	−9	12	−29	16	−13	−27
−145	79	37	7	−14	−9	−20	27
28	35	−47	−50	15	11	7	−17
−54	−5	29	12	6	−28	−13	−2
−6	14	7	3	−8	12	−1	−4

Quantization: space vs. frequency (1/3)

Quantization tables:

1	1	1	2	4	8	16	32
1	1	2	2	4	8	16	32
1	2	4	4	8	16	32	64
2	2	4	8	16	32	32	64
4	4	8	16	16	32	64	128
8	8	16	32	32	128	128	256
16	16	32	32	64	128	256	256
32	32	64	64	128	256	256	256

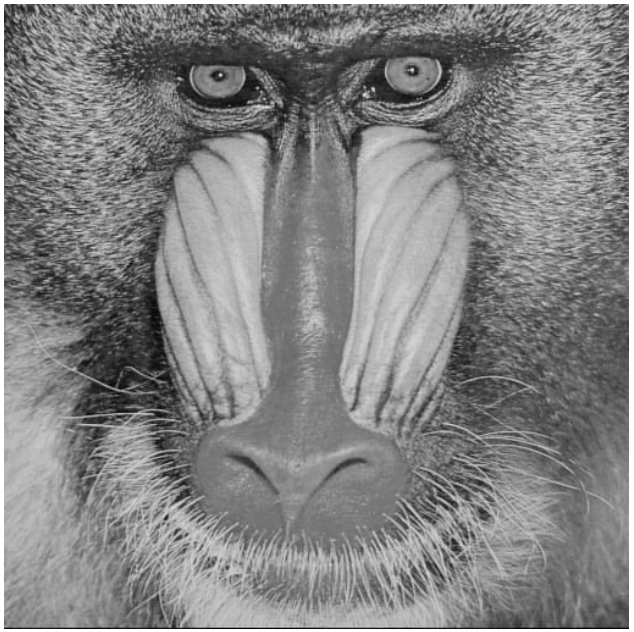
low compression

1	2	4	8	16	16	32	64
2	2	4	8	16	128	256	256
4	4	8	16	64	256	256	256
8	8	16	64	256	256	256	256
16	16	64	256	256	256	256	256
16	128	256	256	256	256	256	256
32	256	256	256	256	256	256	256
64	256	256	256	256	256	256	256

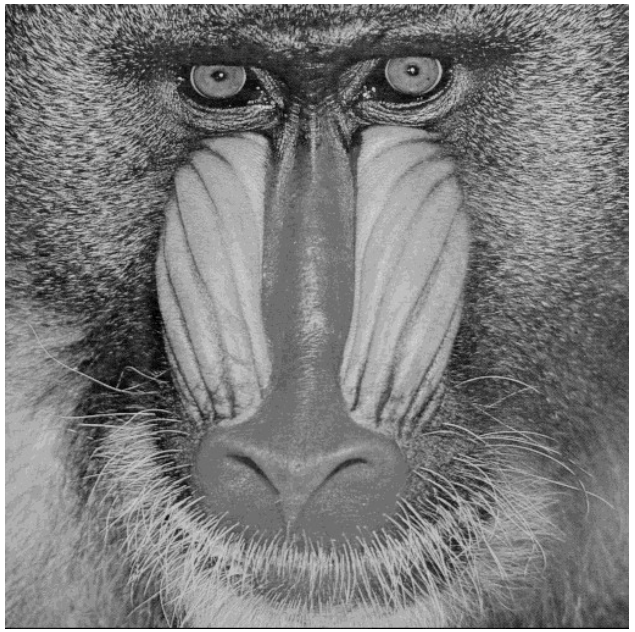
high compression

Quantization: space vs. frequency (2/3)

50% original file size



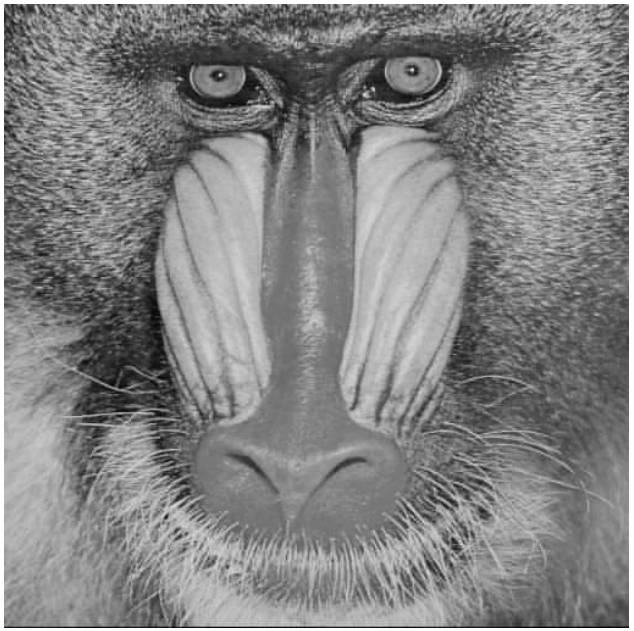
DCT-based quantization



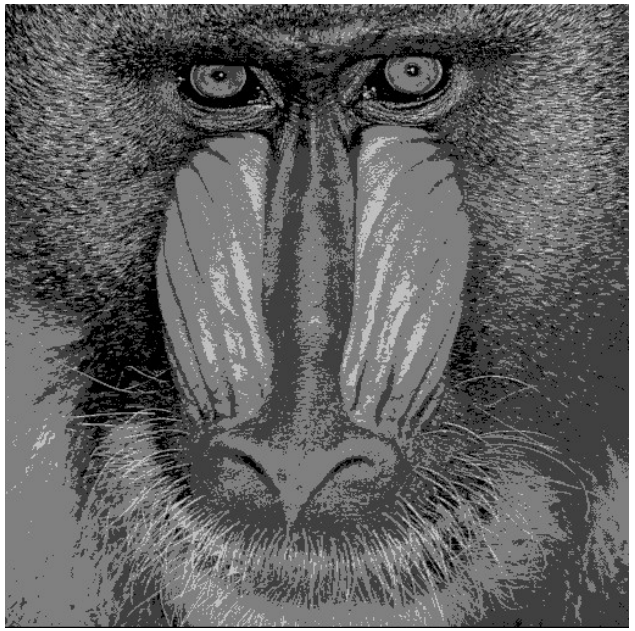
raw quantization

Quantization: space vs. frequency (3/3)

25% original file size



DCT-based quantization



raw quantization

The joint picture expert group (JPEG) algorithm

JPEG encoding scheme:

JPEG decoding scheme:

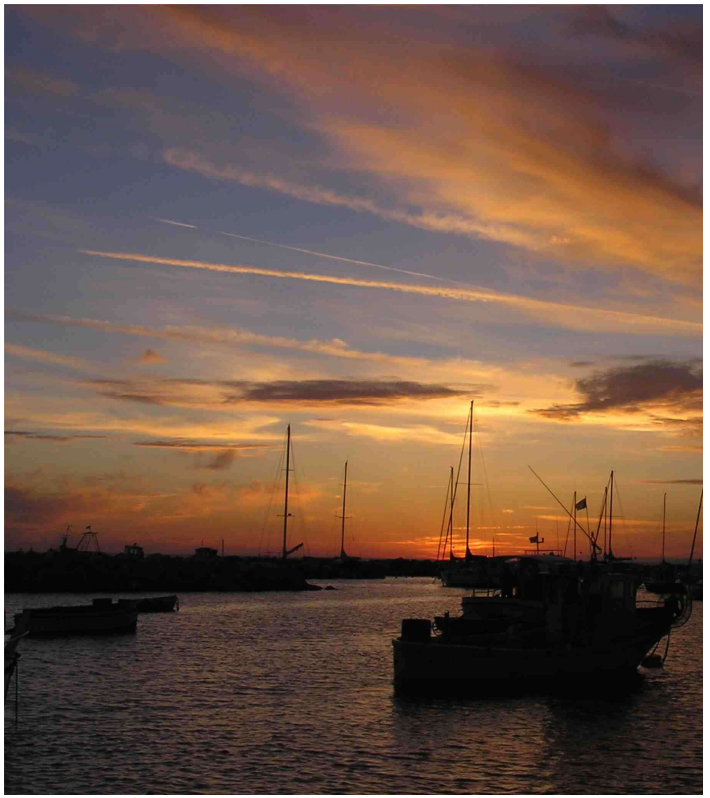
Quantization in the frequency domain

Quantization matrix (with zig-zag scanning):

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

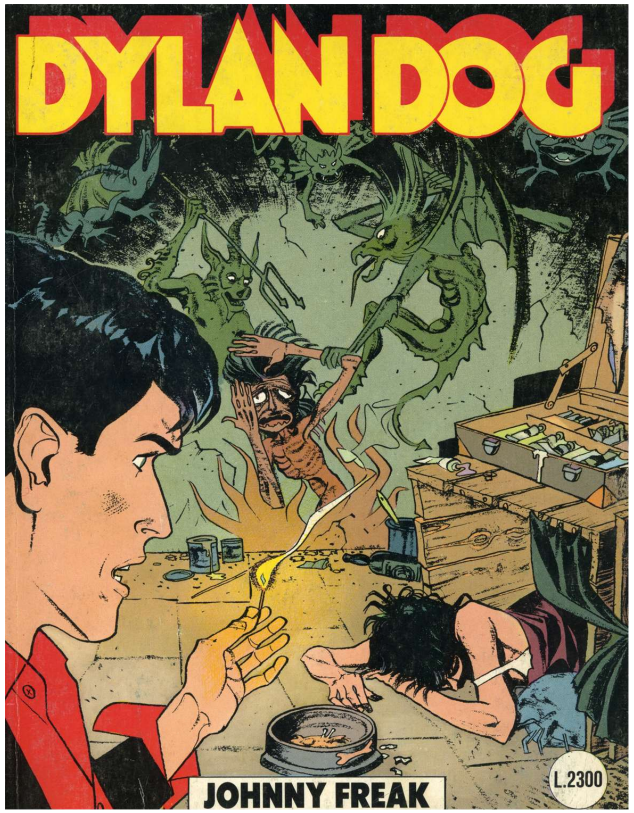
Examples of compressed images (1/4)

original image

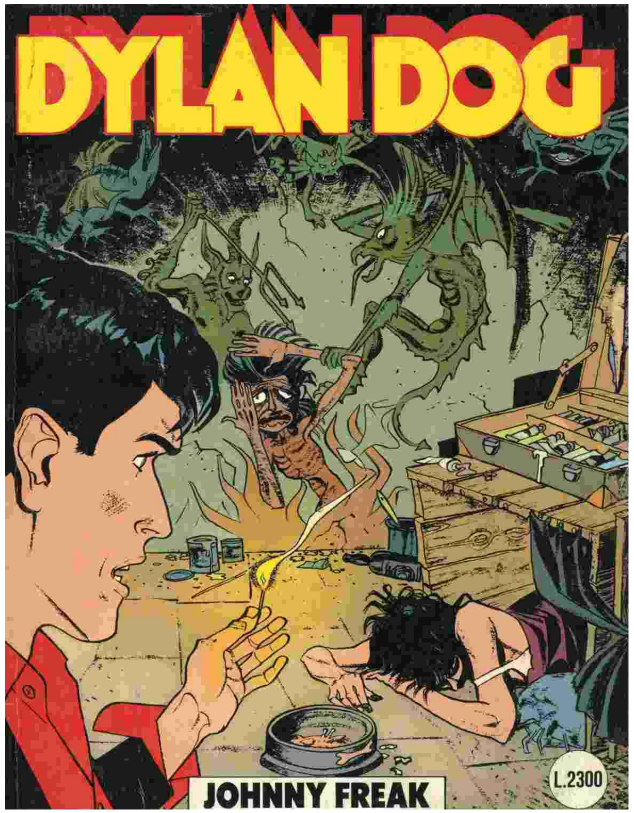


compressed image (ratio 1:10)

Examples of compressed images (2/4)

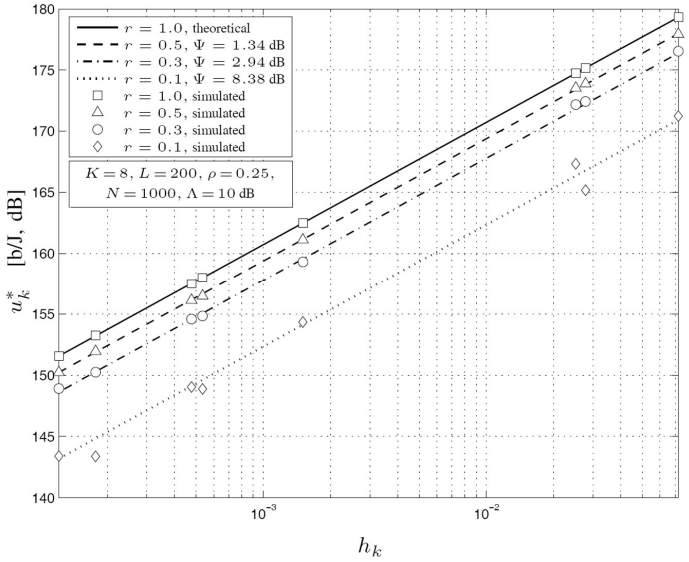


original image

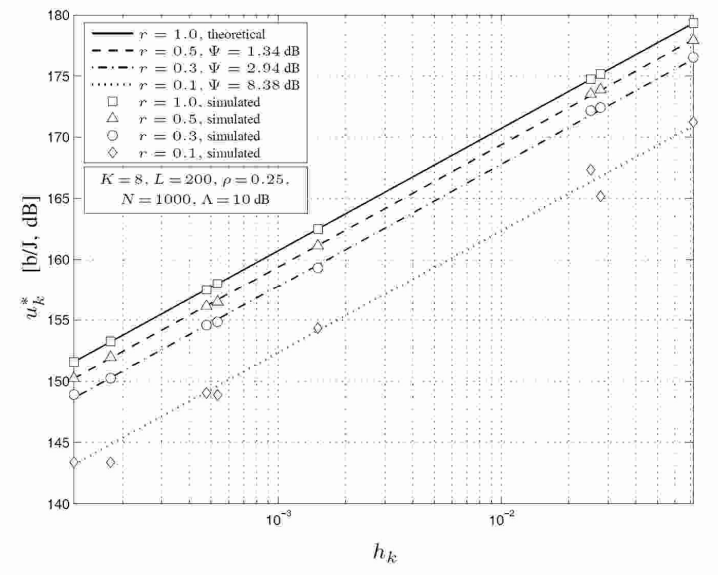


compressed image (ratio 1:10)

Examples of compressed images (3/4)



original image



compressed image (ratio 1:10)

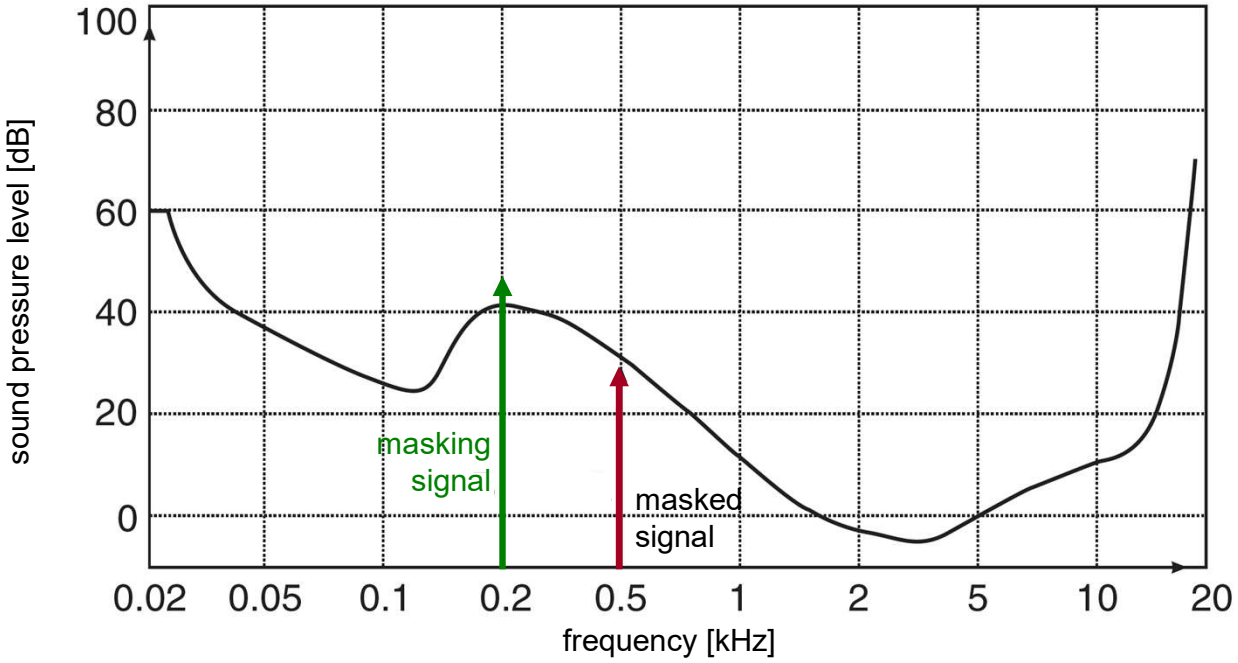
Examples of compressed images (4/4)

2.66 MB

7.02 MB

Basics of audio coding (1/2)

Similarly, audio coding adopts a lossy compression, again based on the psycho-acoustic model



Basics of audio coding (2/2)

Audio coding, like image coding, is based on the semantics of the source:

- **Uncompressed, CD-quality audio (1.41 Mb/s):**
- **Compressed, MPEG-1/2 audio layer III (MP3) audio (32 kb/s):**
- **Compressed using JPEG algorithm (32 kb/s):**

Why do we need compressed videos?

high-definition video (HD):

- component resolution (R, G, B): 8 bits / component
- color image composition: 3 components / pixel
- resolution: 1920 × 1080 pixels / frame
- image refresh rate: 60 frames / s

$$R_b = (8 \cdot 3 \cdot 1920 \cdot 1080 \cdot 60)\,\mathrm{b/s} \approxeq 2.99\,\mathrm{Gb/s}$$

ultra-high-definition video (UHD, aka 4k):

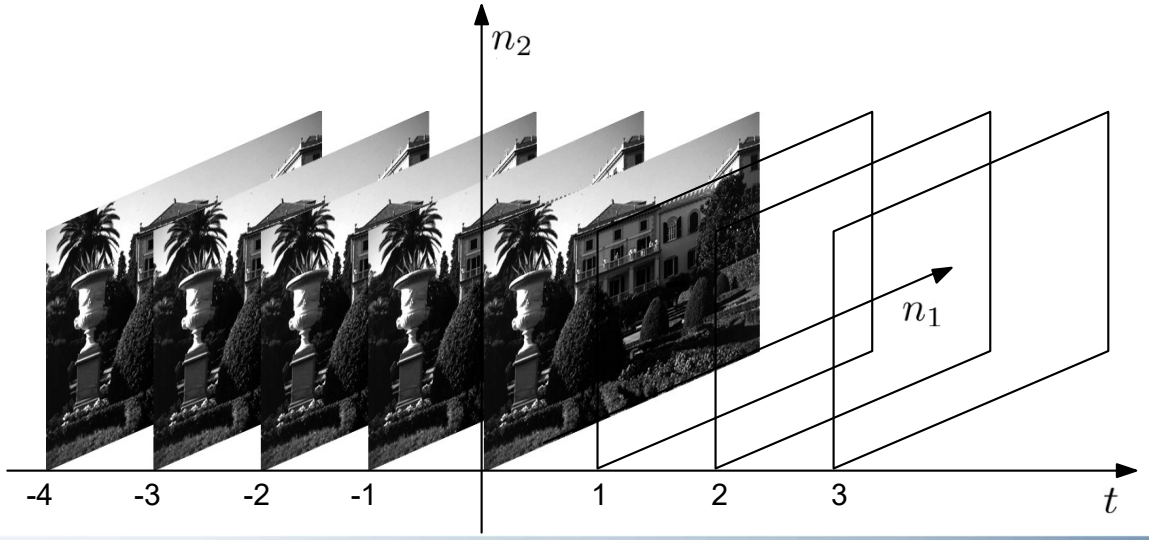
- 3840 × 2160 pixels / frame, using 120 f/s refresh rate and 48 bits per pixel

$$R_b = (16 \cdot 3 \cdot 3840 \cdot 2160 \cdot 120)\,\mathrm{b/s} \approxeq 47.78\,\mathrm{Gb/s}$$

Principles of video coding

We can compress a video stream, based on:

- **spatial correlation**, using the same principles exploited in image compressing
- **temporal correlation**, using time memory across successive frames



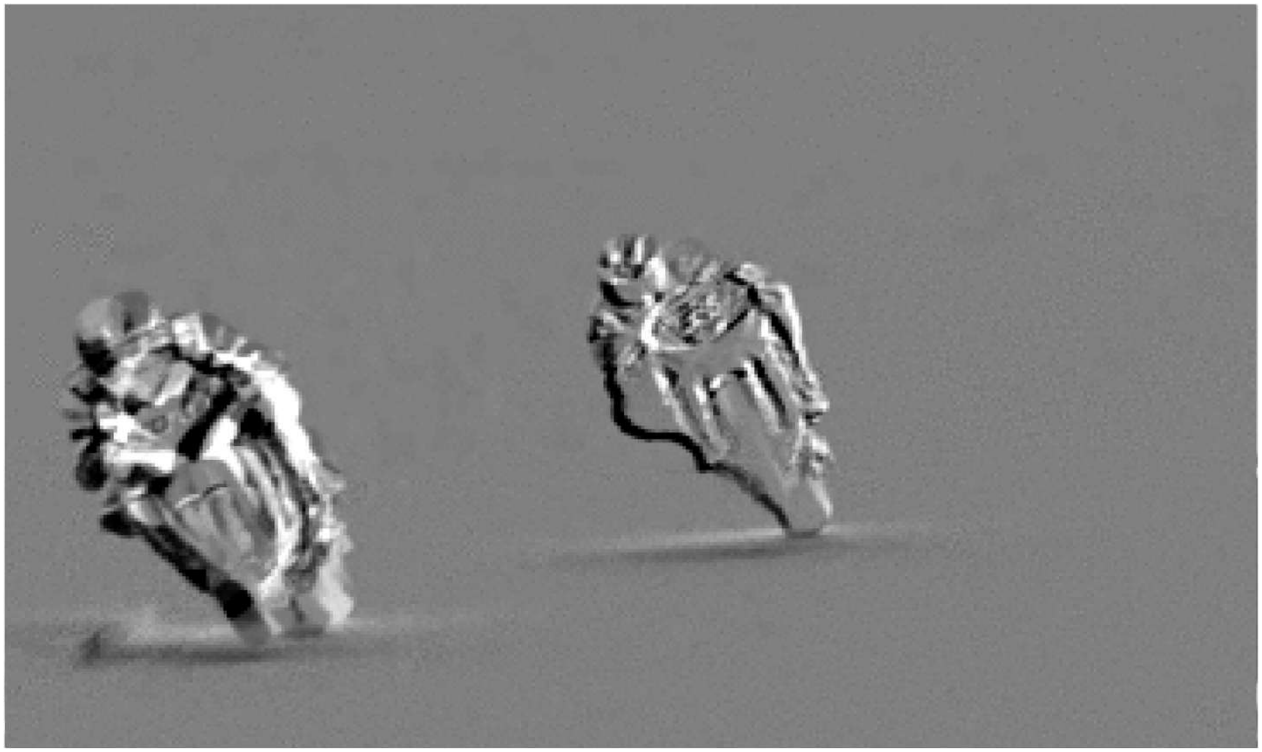
Examples of videos

poorly time-correlated video

highly time-correlated video

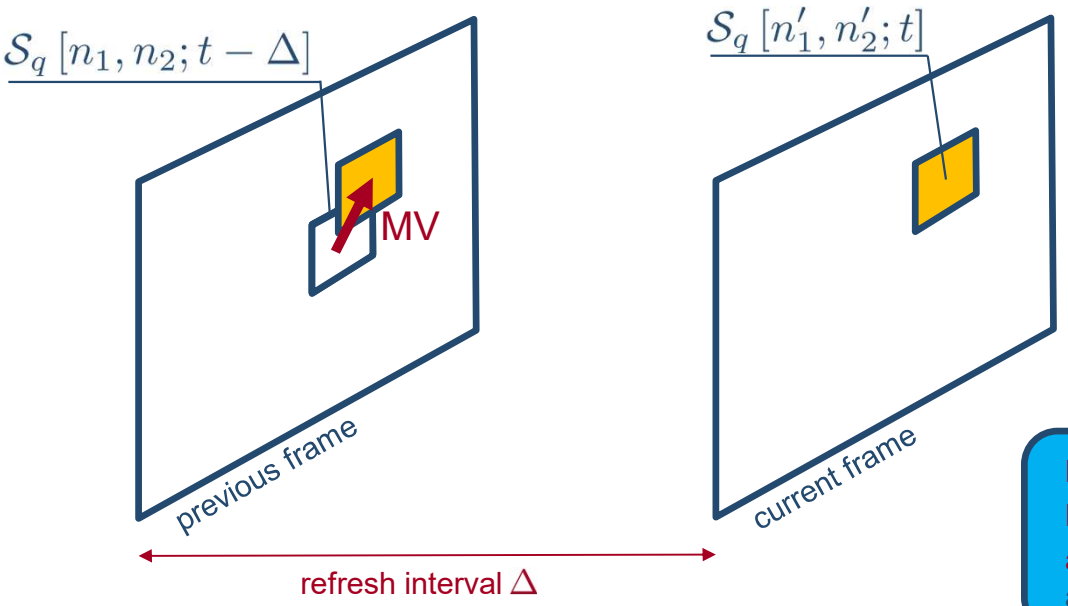
Time correlation

$$\mathcal{S}[n_1, n_2; t] - \mathcal{S}[n_1, n_2; t - \Delta]$$



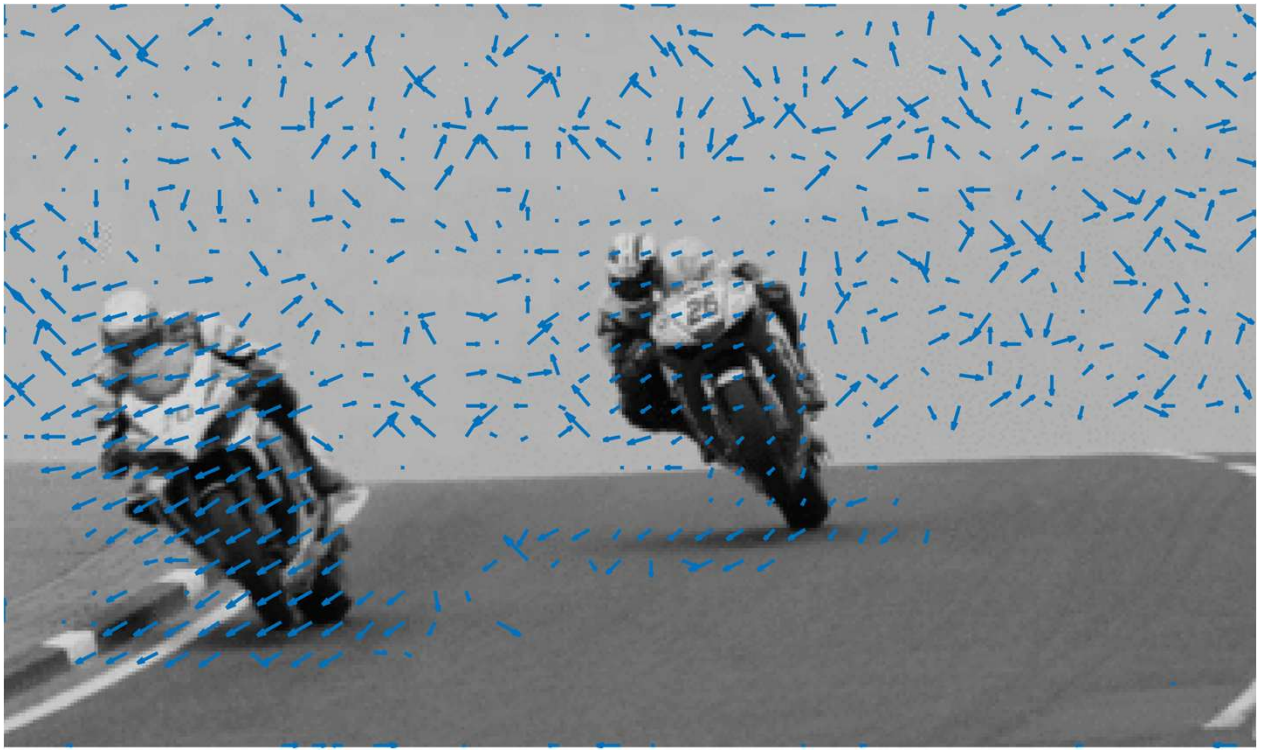
Motion prediction (1/3)

Effective video compression techniques rely on motion prediction based on motion vectors (MVs):



MPEG-1: 16×16 blocks
MPEG-2: 16×8 blocks
advanced standards: adaptive blocks

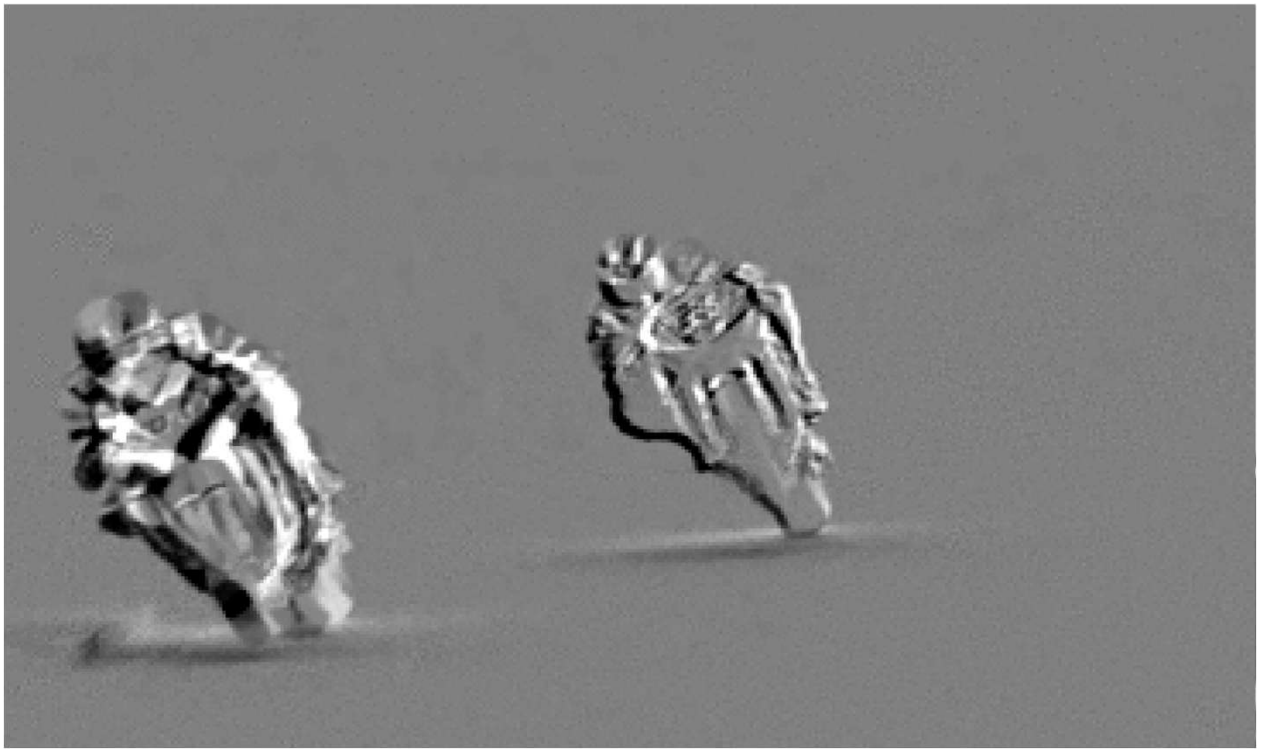
Motion prediction (2/3)



Motion prediction (3/3)

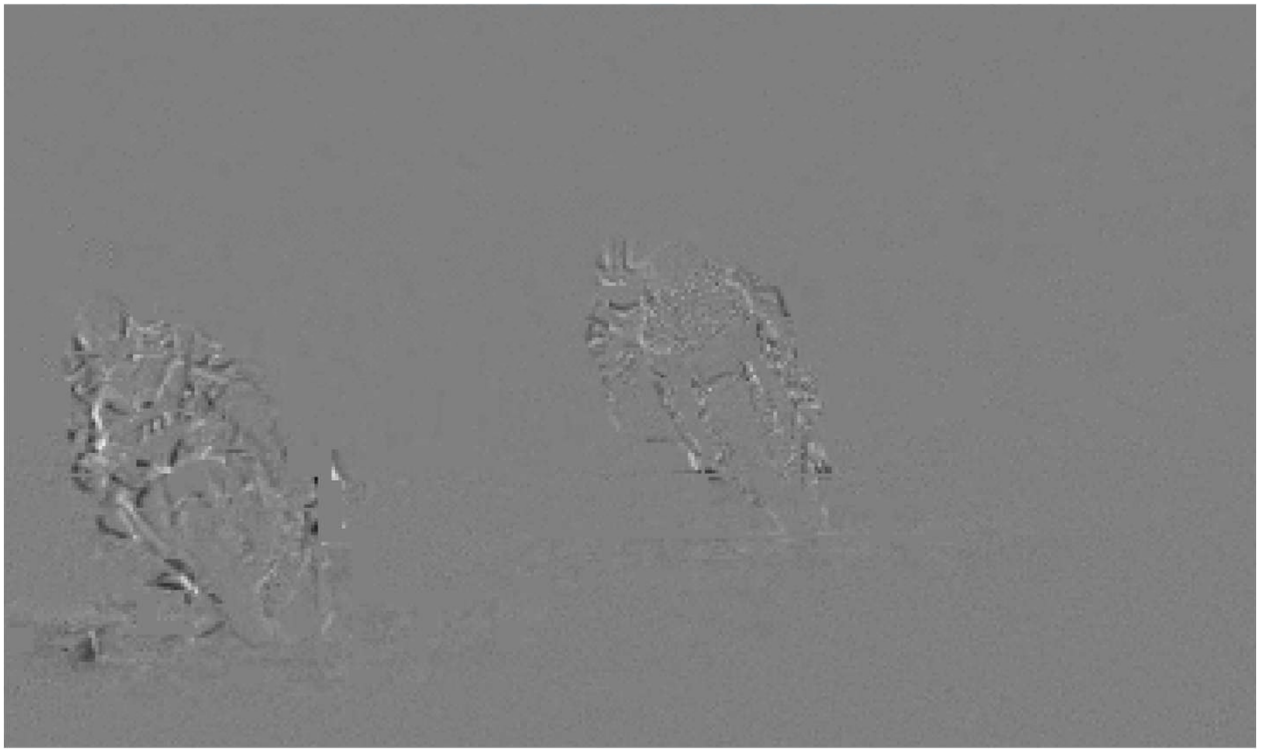
Motion compensation (1/2)

$$\mathcal{S}[n_1, n_2; t] - \mathcal{S}[n_1, n_2; t - \Delta]$$ **(without MC)**



Motion compensation (2/2)

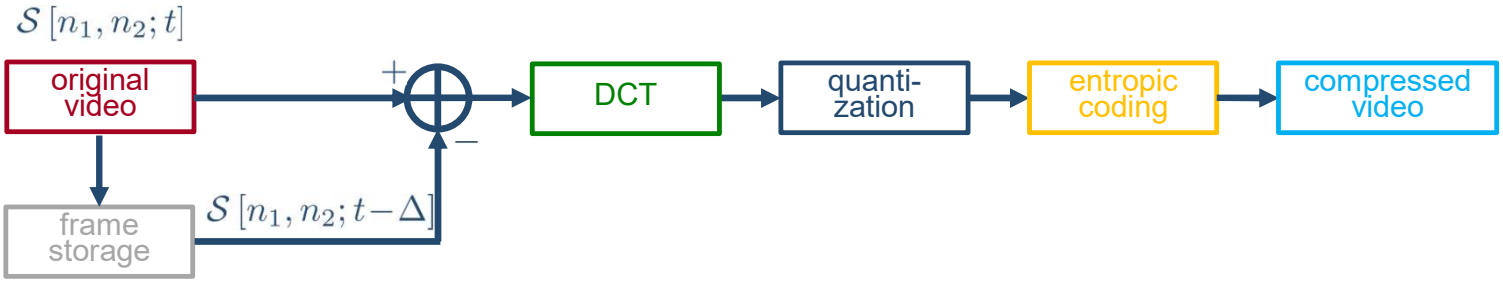
$$\mathcal{S}[n_1, n_2; t] - \hat{\mathcal{S}}[n_1, n_2; t]$$ **(with MC)**



Motion picture expert group (MPEG) scheme (1/3)

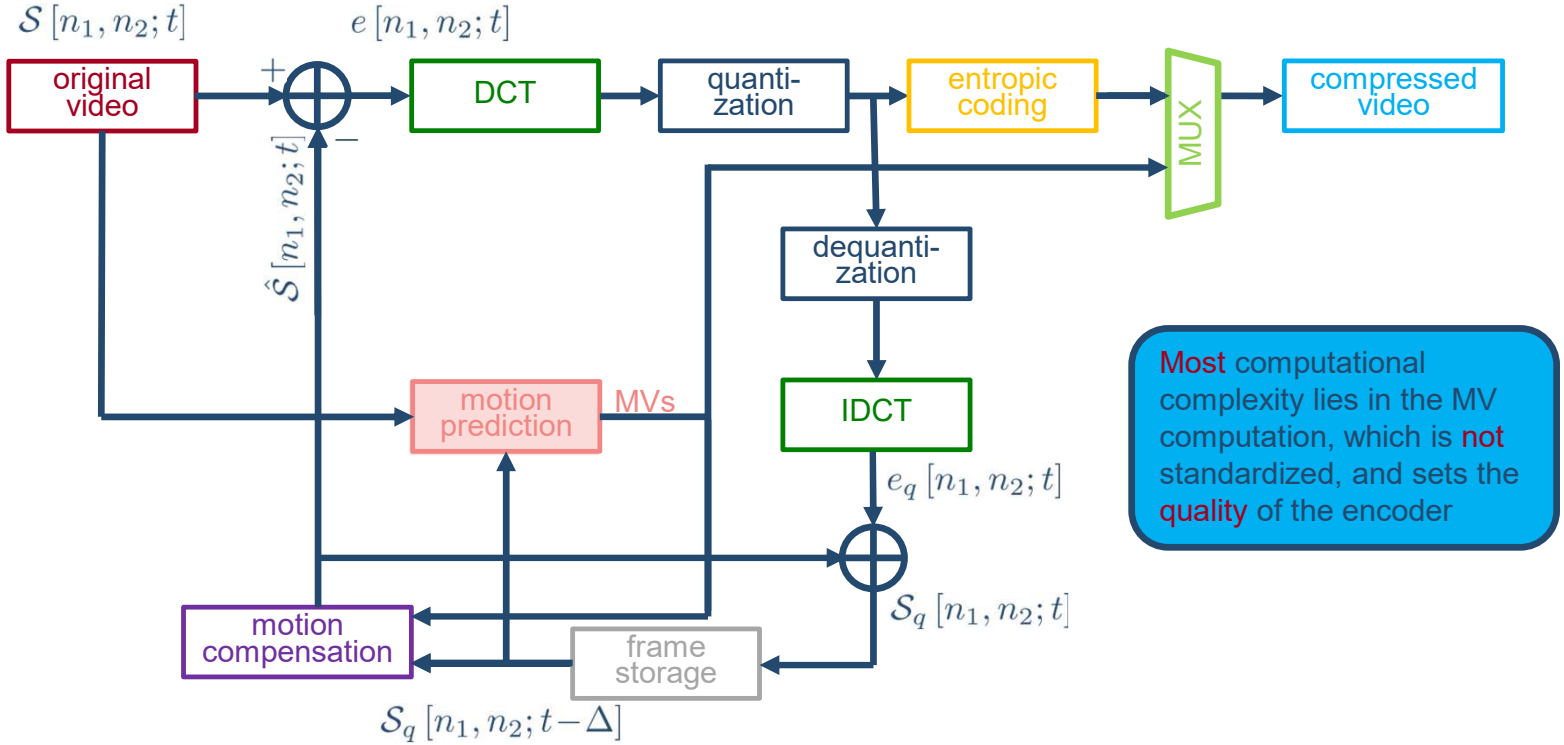
MPEG encoding scheme without time correlation:

MPEG encoding scheme without motion compensation (MC):



Motion picture expert group (MPEG) scheme (2/3)

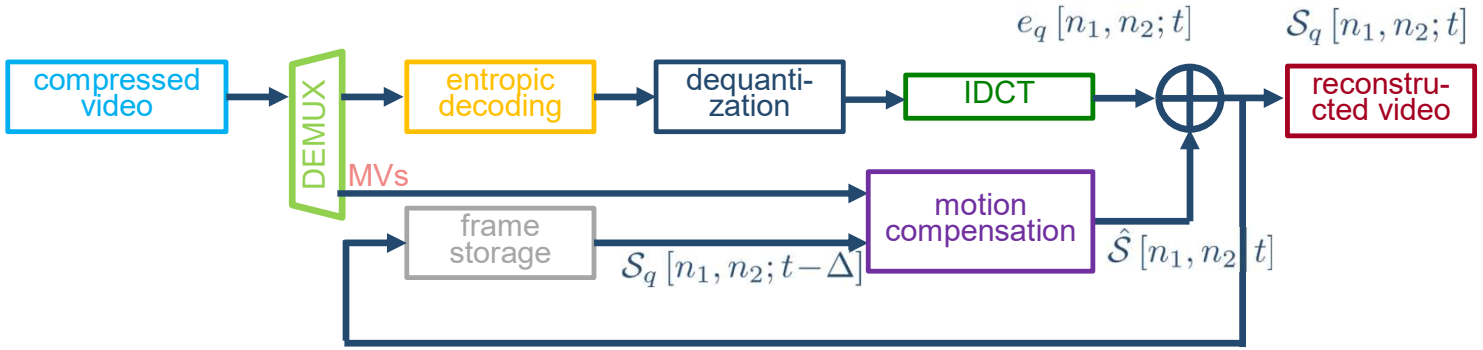
MPEG encoding scheme with MC:



Most computational complexity lies in the MV computation, which is not standardized, and sets the quality of the encoder

Motion picture expert group (MPEG) scheme (3/3)

MPEG decoding scheme:



- the MPEG encoder implements **locally** a decoder, in order to control the deviation between the two streams
- since there is **no** motion prediction at the decoder, it can be implemented with a **low-complexity architecture**

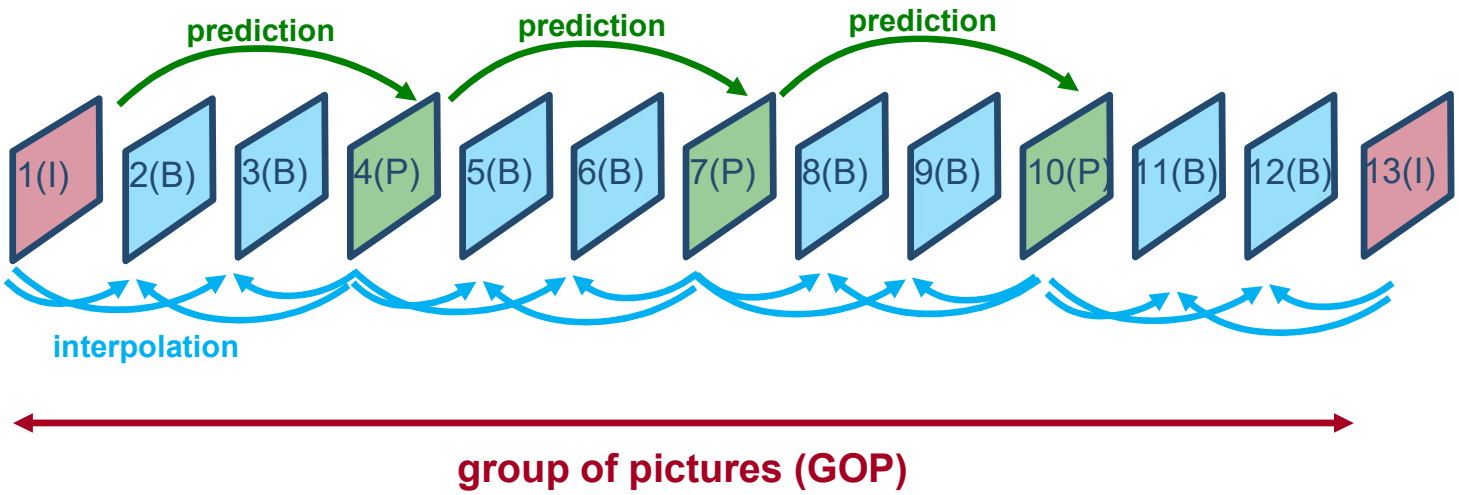
Group of pictures (GOP) (1/3)

Additional compression can be achieved by arranging the frames into:

- **I (intra) pictures, coded without reference to other pictures, similarly to JPEG, also used for random access – *low compression***
- **P (predicted) pictures, coded from preceding I or P frames, using MC – *medium compression***
- **B (bi-directionally predicted) pictures, coded by bi-directional interpolation between I and/or P frames which precedes and follows them – *high compression***

Group of pictures (GOP) (2/3)

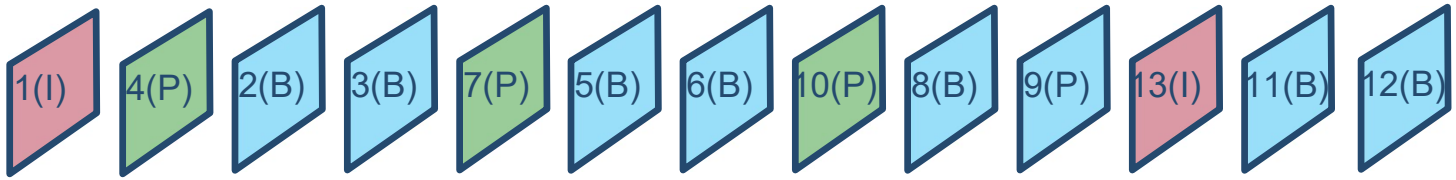
Concatenation of frames in the video stream:



In this example, the GOP is composed by 12 frames, with 3 frames between two successive P pictures

Group of pictures (GOP) (3/3)

To enable a correct decoding, the sequence of encoded frames is:



Picture re-ordering is needed at both the encoding and the decoding side (the latter needs to gather the information on the frame type based on the metadata included in the encoded stream)

Examples of MPEG-2 compressed videos (1/3)

6 MB/s

4 MB/s

1.5 MB/s

Examples of MPEG-2 compressed videos (2/3)

6 MB/s

4 MB/s

1.5 MB/s

Examples of MPEG-2 compressed videos (3/3)

6 MB/s

4 MB/s

1.5 MB/s

Evolution of video compression standards (1/4)

The **MPEG-2 standard** (a.k.a. H.222/H.262) was designed to target standard definition (SD) used for DVD videos (first release: 1996)

With the increasing demand for additional quality, other standards have been deployed:

- **advanced video coding (AVC)**, also referred to as H.264 or MPEG-4 Part 10 (first release: 2004), targets high definition (HD), thanks to adaptive macro-blocks, more effective entropic coding and lossy compression techniques
- **high efficiency video coding (HEVC)**, also known as H.265 and MPEG-H Part 2 (first release: 2013), targets ultra-high-definition (UHD), thanks to improved MV prediction and motion compensation

Evolution of video compression standards (2/4)

MPEG-2: 3.54 MB

Evolution of video compression standards (3/4)

AVC: 1.19 MB

Evolution of video compression standards (4/4)

HEVC: 1.04 MB

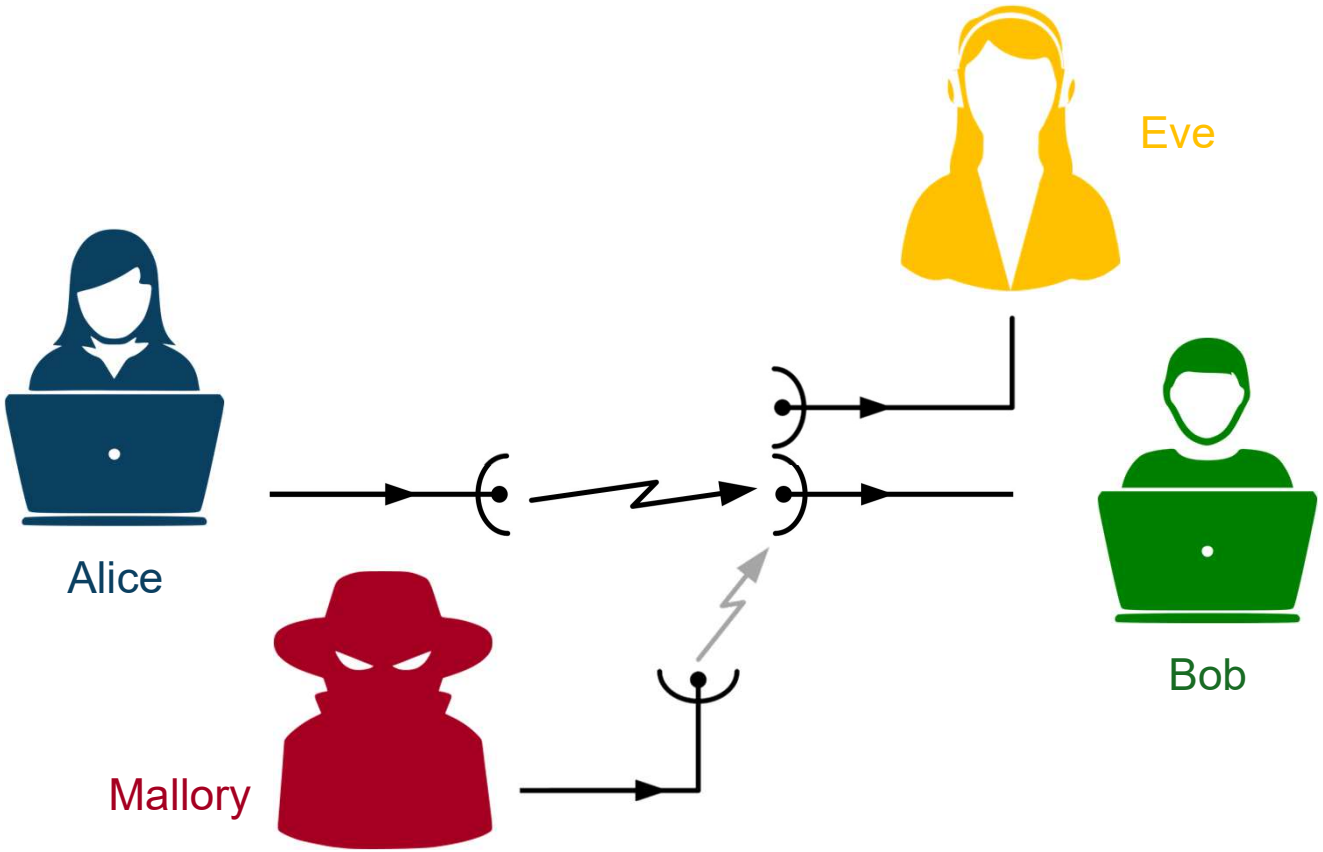
Encryption

Features of encrypted communications

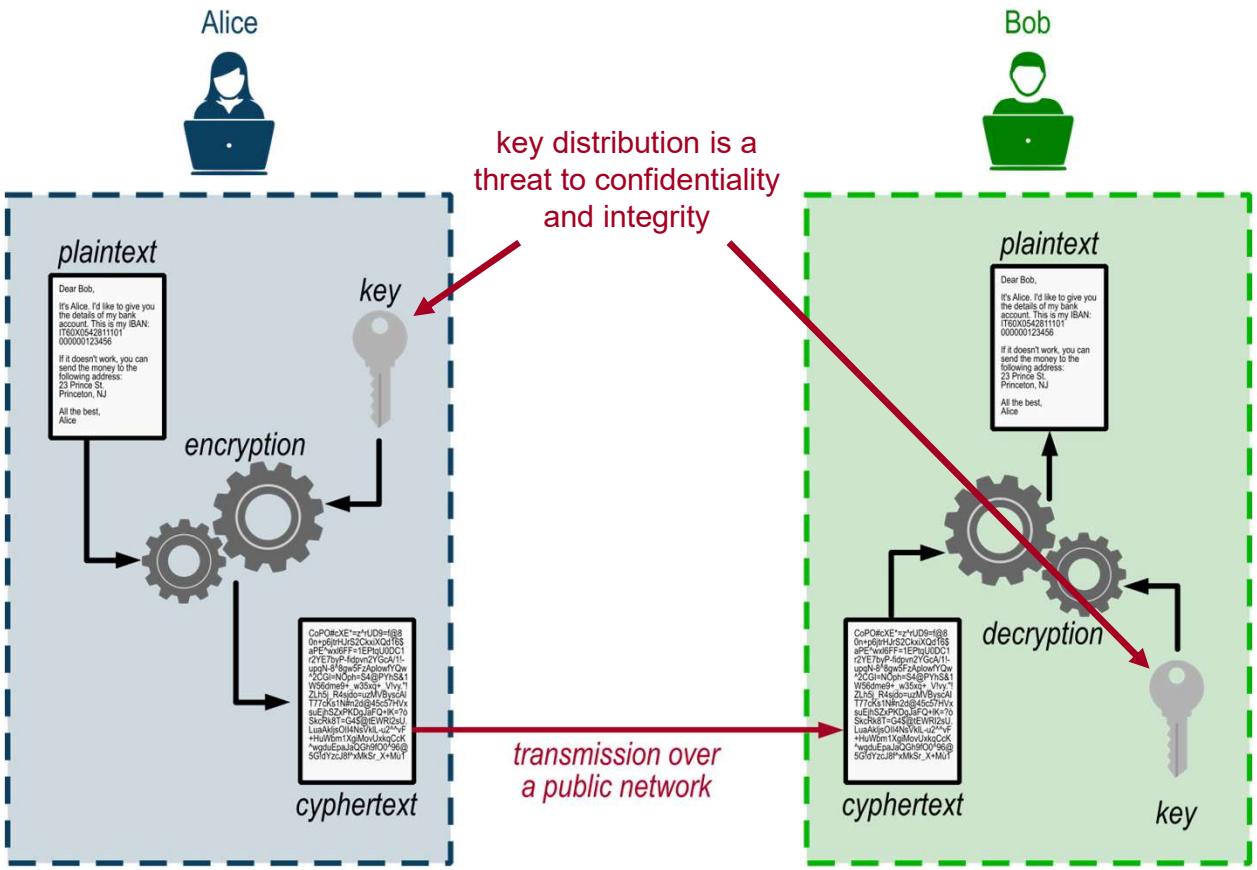
- Confidentiality: third-party cannot read exchanged data
  attacks: eavesdropping, sniffing
- Integrity: third-party cannot change exchanged data
  attacks: man-in-the-middle
- Authentication: each party is sure who is really communicating with
  attacks: masquerading, spoofing, traffic generation

- Availability: time the system is in a functioning condition
  attacks: denial of service (DoS), distributed DoS (DDoS)

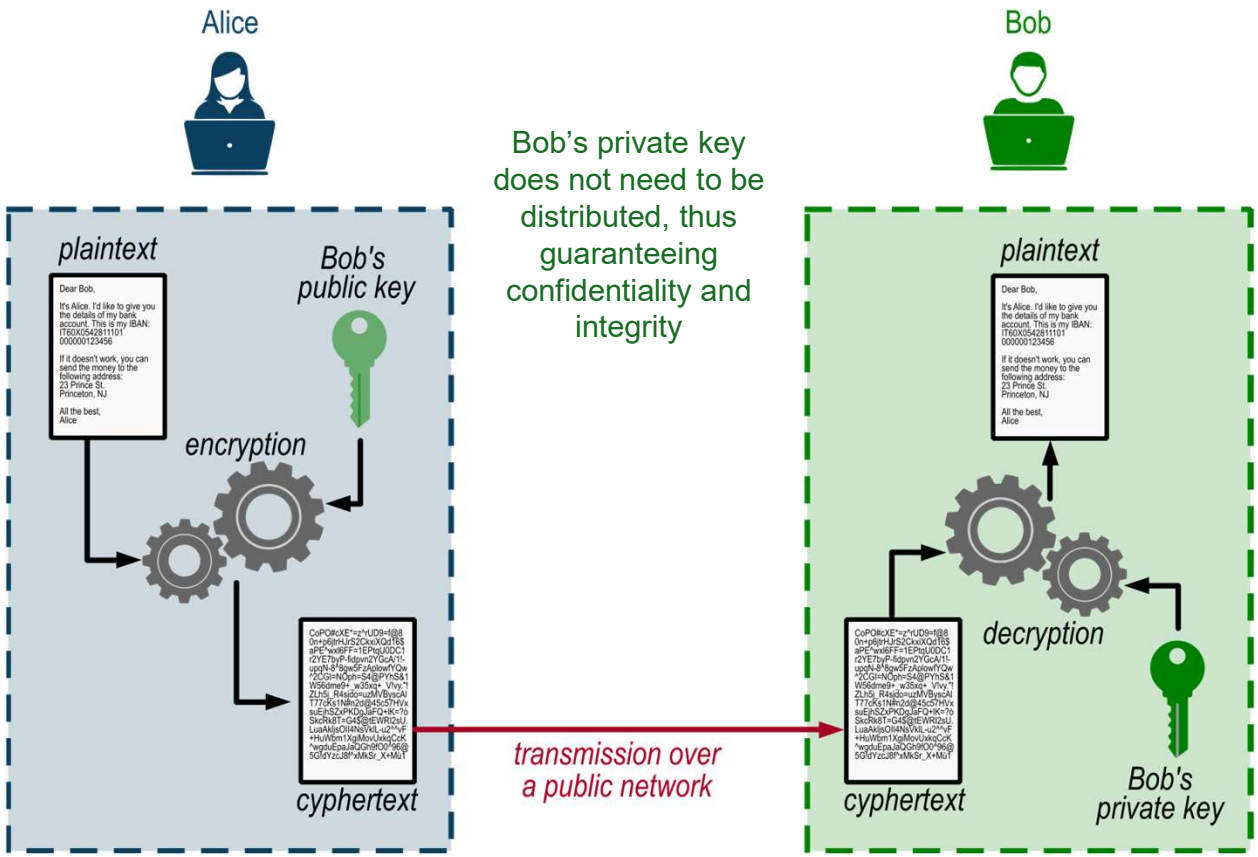
Players in a cybersecurity scenario



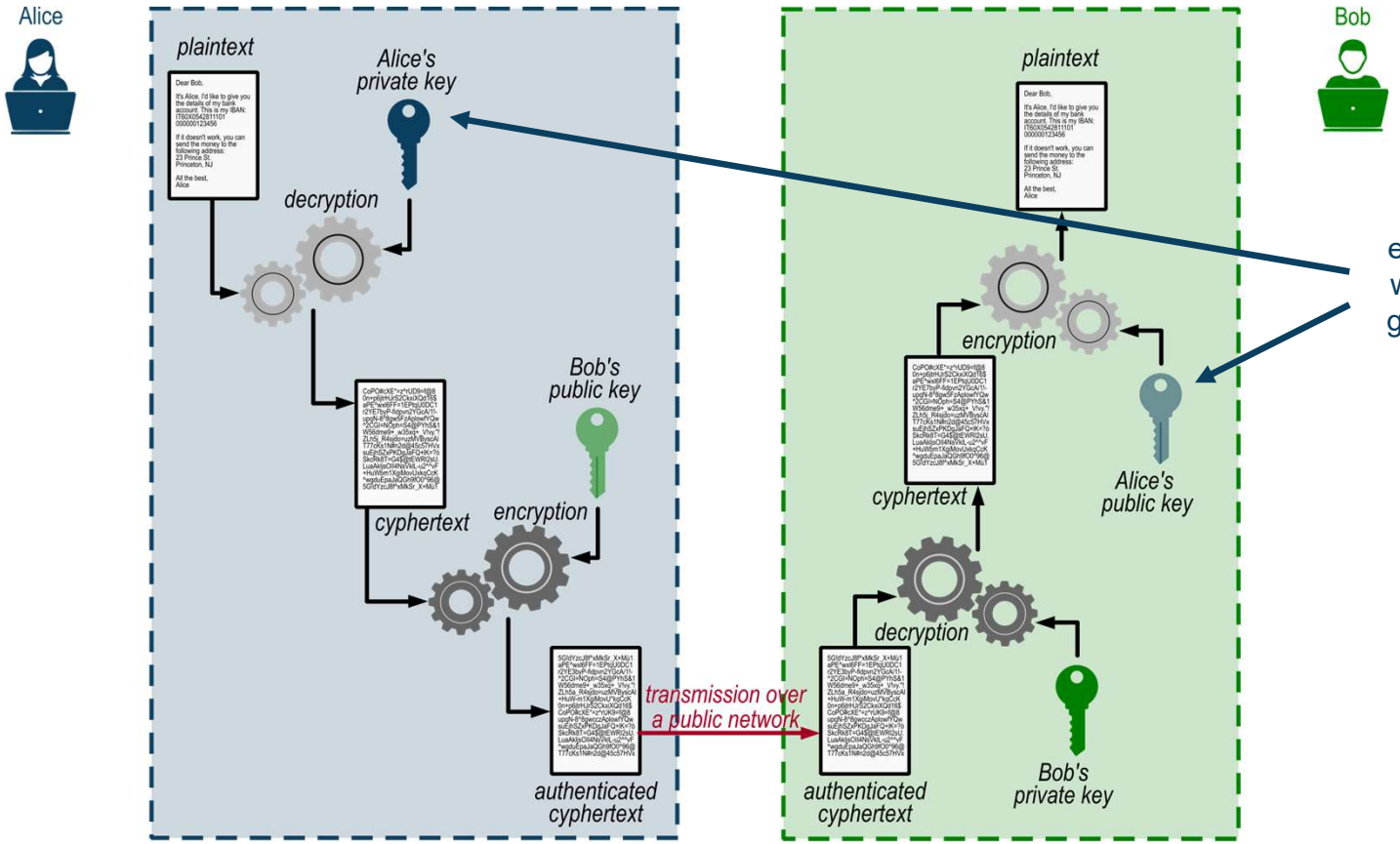
Encryption using a secret (symmetric) key



Encryption using a public (asymmetric) key



Adding authentication using a private key



encrypting the plaintext with Alice's private key guarantees authenticity to the message

Channel coding

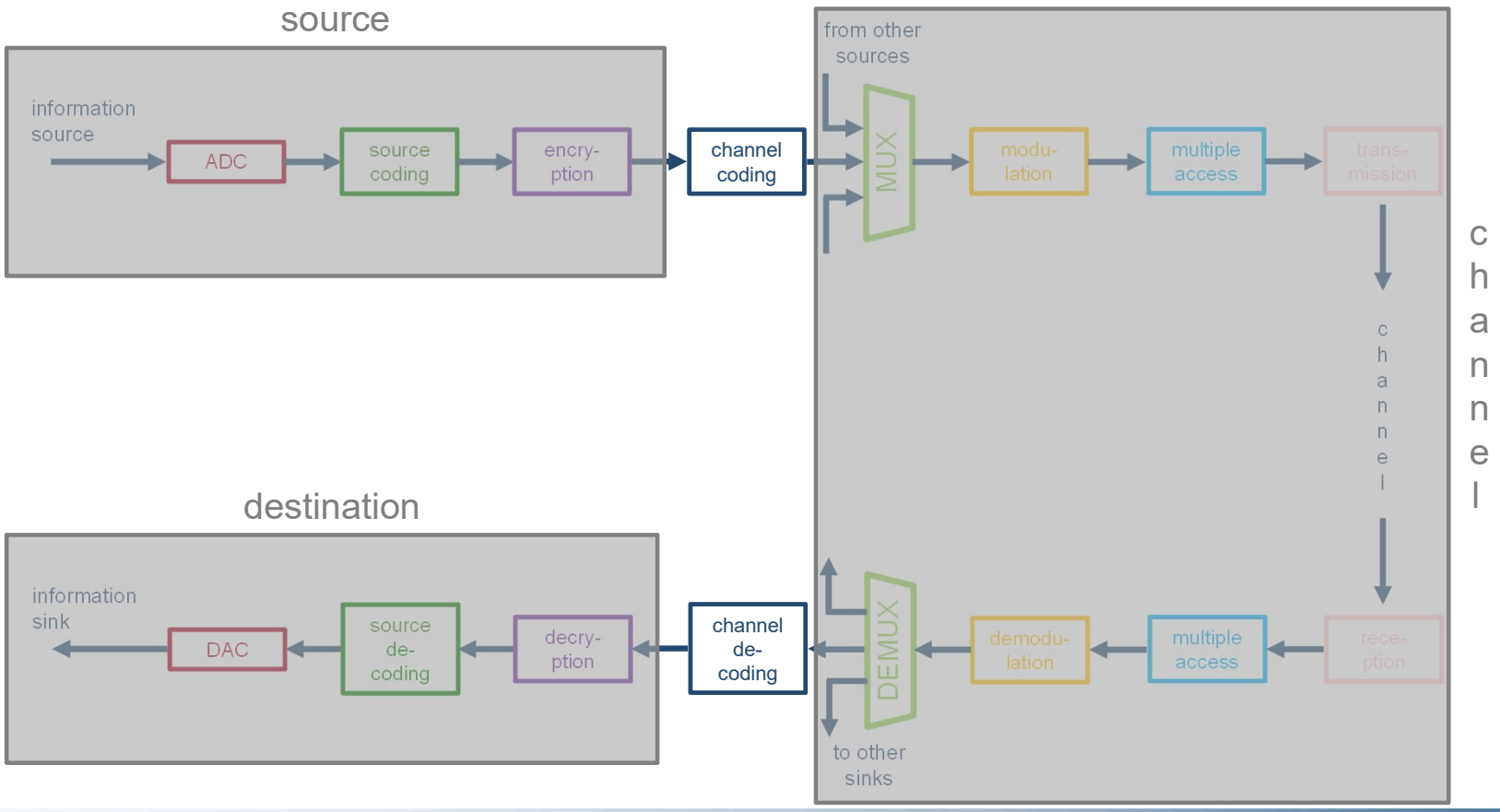
General principles (1/5)

Channel coding attempts to do the same: it gives the channel-encoded message a particular structure, such that bit errors introduced by signal propagation through the error-prone channel can be detected (and potentially corrected) at the receive side

Two ways of recovering the packets:

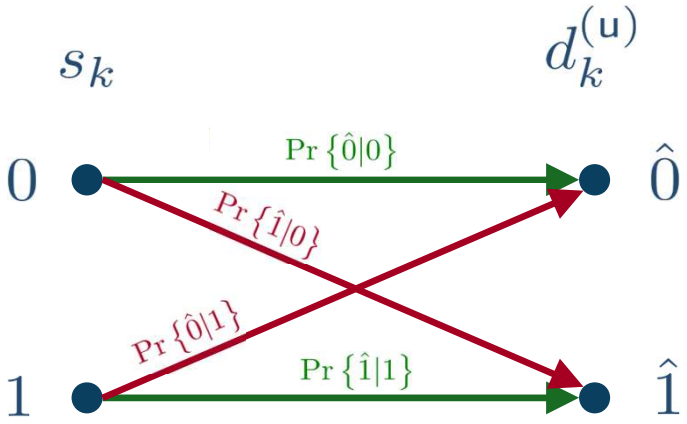
- forward error correction (FEC)
- automatic repeat request (ARQ)

General principles (2/5)



General principles (3/5)

Let us suppose we do not use channel coding:

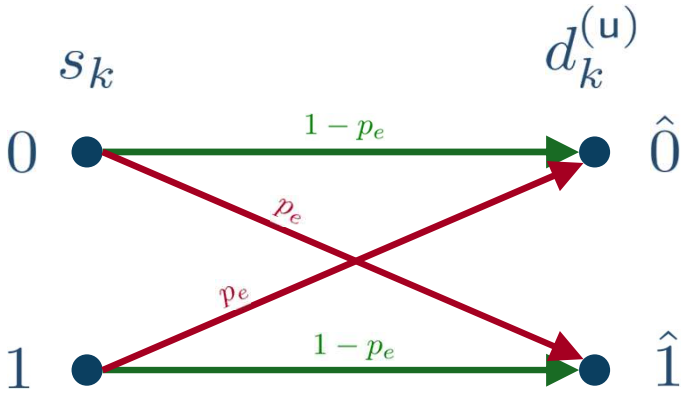


The quantities $\Pr\{\hat{1}|0\}$ and $\Pr\{\hat{0}|1\}$ are the probabilities that a source bit is flipped by the channel, i.e., the probabilities of errors

In practical systems, the probability of error can be measured in terms of bit error rate (BER)

General principles (4/5)

A customary assumption is to consider $\Pr\{\hat{1}|0\} = \Pr\{\hat{0}|1\} = p_e$, in the so-called binary symmetric channel (BSC):



Other channel models exist (e.g., the Z-channel that mimics an optical-fiber connection)

General principles (5/5)

Our goal is to introduce a channel coding stage such that the end-to-end error probability P_e is (much) lower than the bit-error probability p_e introduced by the BSC:

$$P_e = \Pr\left\{d_k^{(c)} \neq s_k\right\} < \Pr\left\{d_k^{(u)} \neq s_k\right\} = p_e$$

Channel coding for dummies: Repetition coding (1/4)

The simplest way to make the communication more “robust”, against bit flipping introduced by the channel, is to use the repetition code scheme: for each source bit, we send *N* copies

Example: *N*=5

$$s_k = 0 \Rightarrow \mathbf{c} = [0,0,0,0,0]$$
$$s_k = 1 \Rightarrow \mathbf{c} = [1,1,1,1,1]$$

At the receiver, we can apply a decision strategy based on a majority rule:

- If at least (*N*+1)/2 received coded bits are 0, we set $d_k^{(c)} = \hat{0}$, and $d_k^{(c)} = \hat{1}$ otherwise

Channel coding for dummies: Repetition coding (2/4)

What is the detection and correction performance of such approach?

$$P_e = \Pr\left\{\frac{N+1}{2} \text{ errors}\right\} + \Pr\left\{\frac{N+3}{2} \text{ errors}\right\} + \cdots + \Pr\{N \text{ errors}\}$$

Assuming that bit-flipping errors introduced by the channel are independent events,

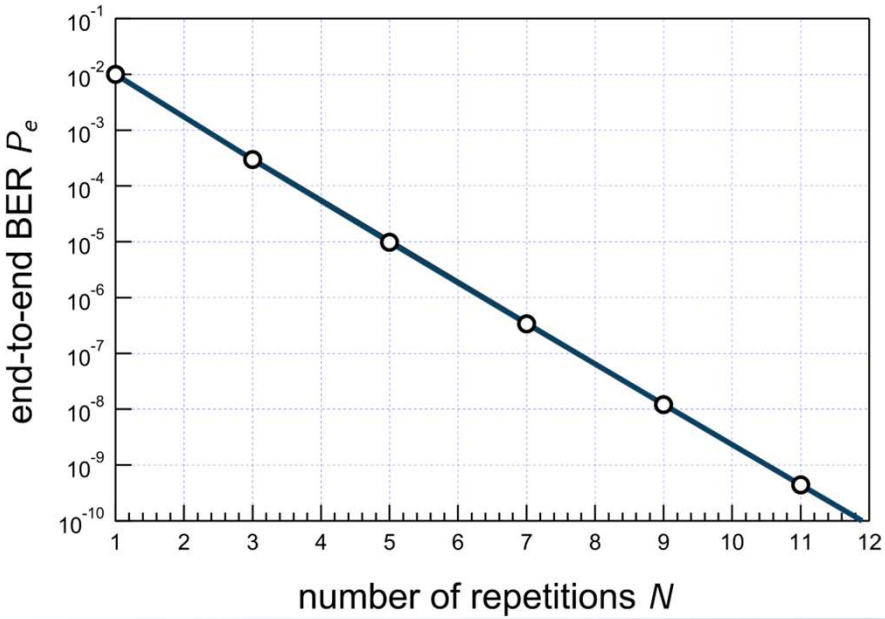
$$\Pr\{n \text{ errors}\} = \binom{N}{n} p_e^n \cdot (1-p_e)^{N-n} = \frac{N!}{n!(N-n)!} p_e^n \cdot (1-p_e)^{N-n}$$

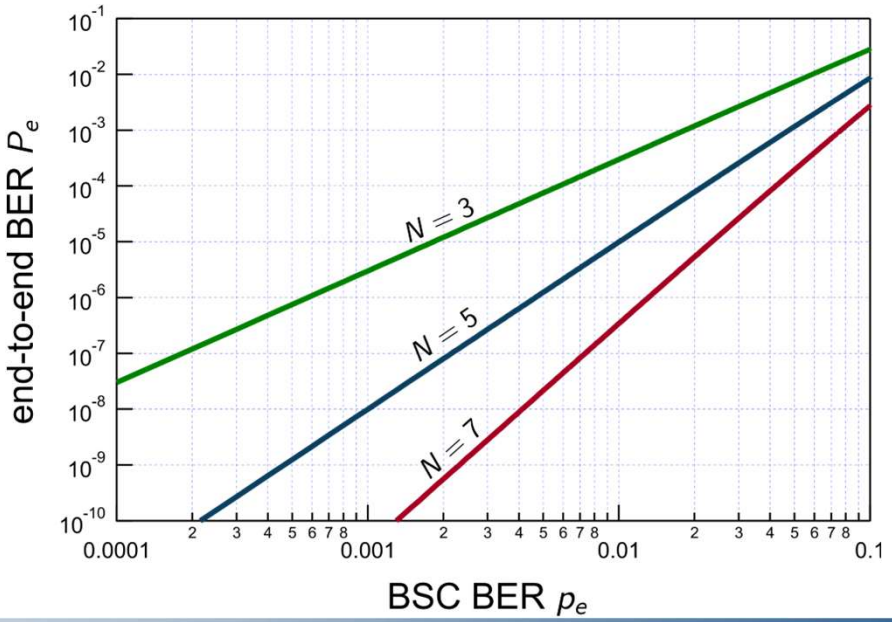
$$P_e = \sum_{n=\frac{N+1}{2}}^{N} \binom{N}{n} p_e^n \cdot (1-p_e)^{N-n} \approxeq \binom{N}{(N+1)/2} p_e^{(N+1)/2} \cdot (1-p_e)^{(N-1)/2}$$

Channel coding for dummies: Repetition coding (3/4)

What is the detection and correction performance of such approach?

$$P_e \approxeq \binom{N}{(N+1)/2} p_e^{(N+1)/2} \cdot (1-p_e)^{(N-1)/2}$$





Channel coding for dummies: Repetition coding (4/4)

What is the drawback? For each source bit, we are sending *N* coded bits!

We are increasing the amount of resources required for the transmission by a factor *N*

In channel coding, an important figure of merit is the coding rate *r*:

number of source bits at the input of the channel encoder

$$r = \frac{K}{N} \leq 1$$

number of coded bits at the output of the channel encoder

- the higher *r* (the closer to 1), the lower the amount of resources requested (i.e., the redundancy introduced) ☑
- the higher *r*, the lower the protection and the error detection/correction capabilities of the encoder ☒

For the repetition code, $r = 1/N$ (that decreases as *N* increases)

Parity check codes (1/3)

A way to improve the efficiency of the channel encoder (i.e., increasing r for a given performance P_e, or, equivalently, decreasing P_e for a given r) is through the use of 2D parity check codes, based on the XOR operation (row by row, and column by column):

A	B	$A \boxplus B$
0	0	0
0	1	1
1	0	1
1	1	0

coding rate

$$r = \frac{K}{N} = \frac{K}{(\sqrt{K}+1)^2} = \frac{16}{25}$$

Example: K=16

$$\mathbf{s} = [0\,1\,1\,1\,0\,0\,1\,0\,1\,1\,1\,0\,1\,0\,0\,1]$$

step 1: sort the bits into a $\sqrt{K} \times \sqrt{K}$ matrix

step 2: add parity bits

step 3: sort the matrix into a $(\sqrt{K}+1)^2$-element vector

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

$$\mathbf{c} = [0\,1\,1\,1\,0\,0\,1\,0\,1\,1\,1\,0\,1\,0\,0\,1\,0\,0\,1\,0\,1\,0\,1\,1\,1]$$

Parity check codes (2/3)

How does error detection and correction work?

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

case 1: ☑

case 2: ☑

case 3: ☒

What is the detection and correction performance of the 2D parity check code?

$$P_e \approxeq \frac{K(\sqrt{K}-1)^2}{4} p_e^4$$

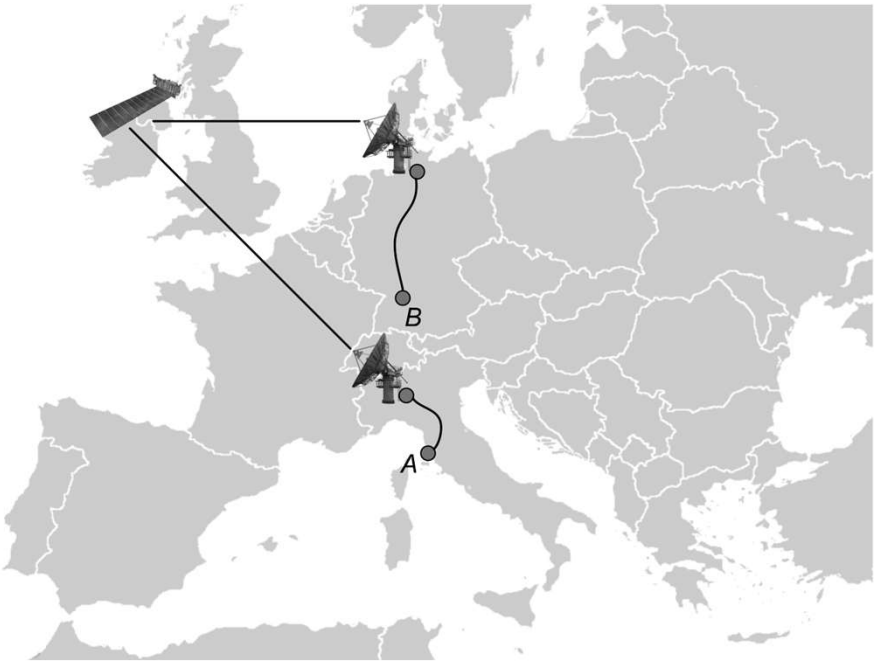
With K=16 and p_e=10^{-2}, P_e is comparable with a repetition code with N=7, but with a coding rate r = 16/25 ≈ 4.5·1/7

Parity check codes (3/3)

Parity check codes are used in both broadcast (e.g., satellite communications) and unicast (e.g., email) communications.

FEC is best suited for broadcast services, where a return channel is typically not available

ARQ mechanisms are better suited for unicast services



Linear block codes (1/2)

Linear block codes are a generalization of the parity check codes

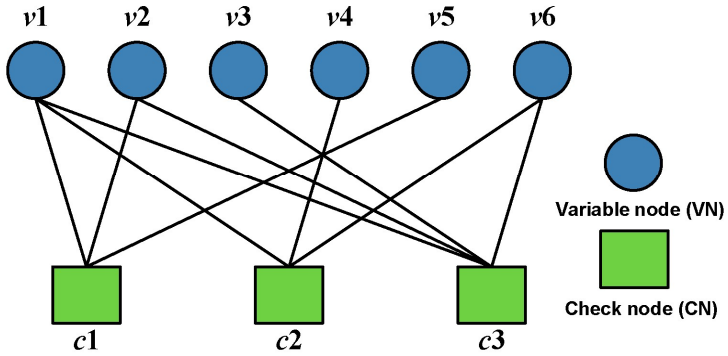
Error detection and correction can be performed using a syndrome vector, which can be derived from the generation matrix, and requires linear operations

Notable examples are:

- Reed-Solomon (RS) codes, e.g. used in satellite digital video broadcasting (DVB-S) services
- low-density parity check (LDPC) codes

Linear block codes (2/2)

- LDPC codes are linear block codes characterized by sparse parity check matrices



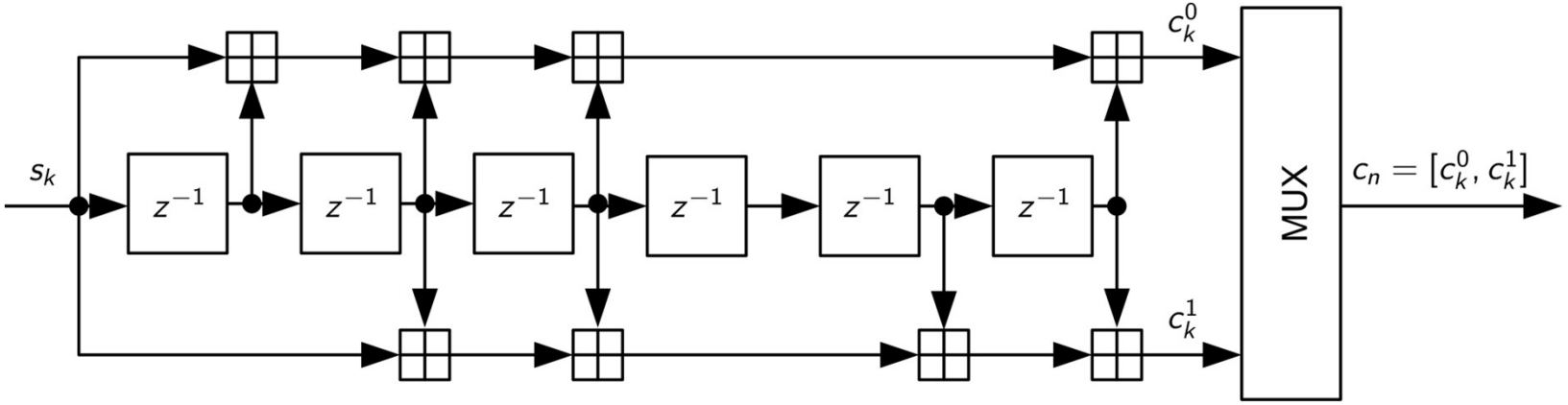
- Invented by Gallagher in 1960, but reproposed by MacKay and Neal in 1995, when technology of iterative decoding was mature enough
- Particularly performing with large block lengths (e.g., 64,800 bits)
- LDPC codes yield quasi-error-free (QEF) reception even in poor link conditions

Convolutional codes (1/4)

Convolutional codes introduce correlation between source bits, by using a feedforward architecture

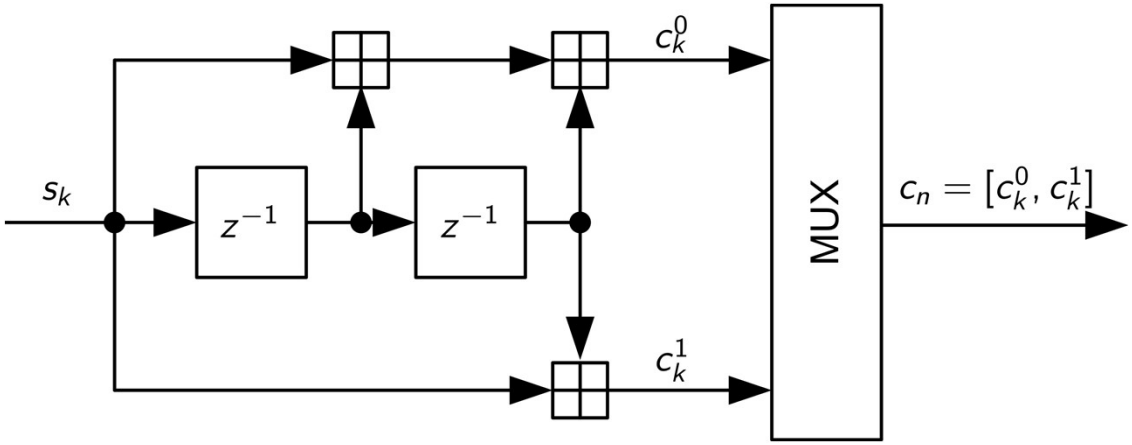
Notable examples:

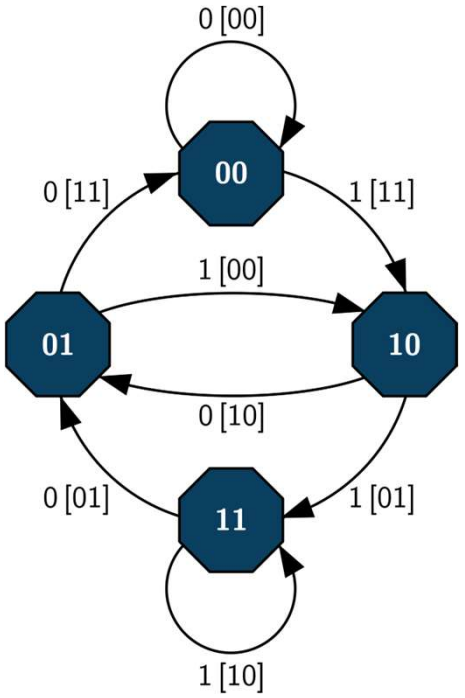
- convolutional code with rate r=1/2, used in WiFi, GPS/Galileo, and DVB-T
- Turbo codes



Convolutional codes (2/4)

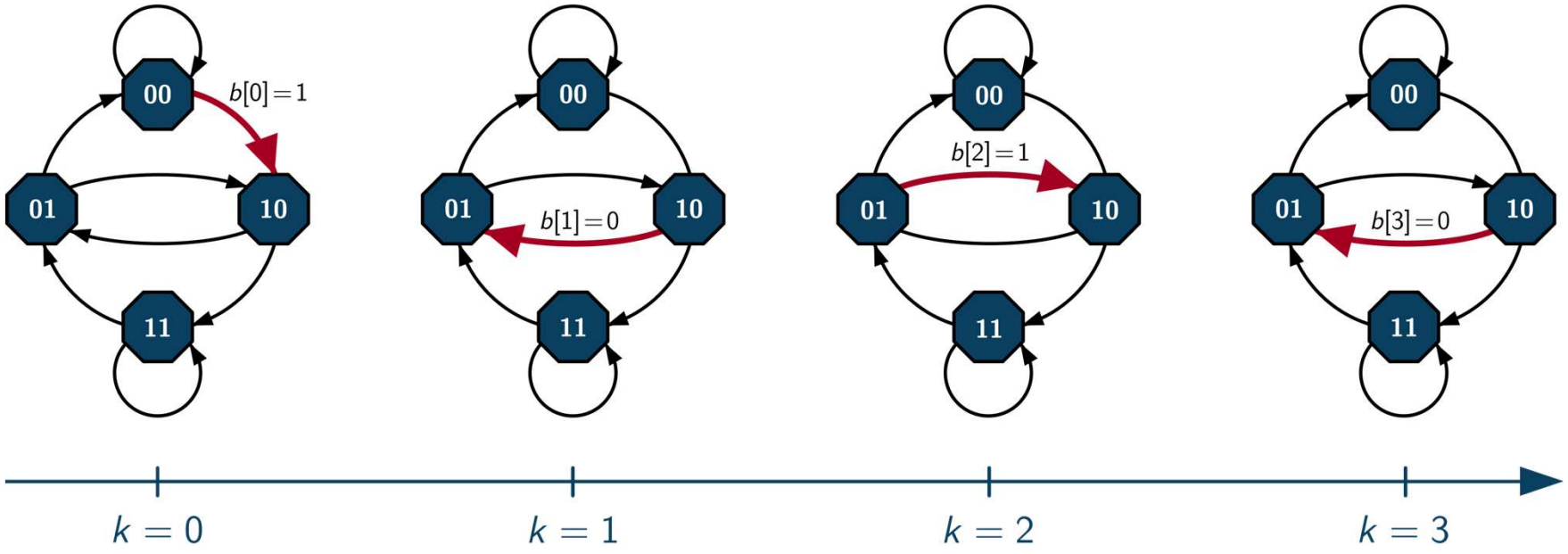
From the encoder structure, we can derive the corresponding finite-state machine:





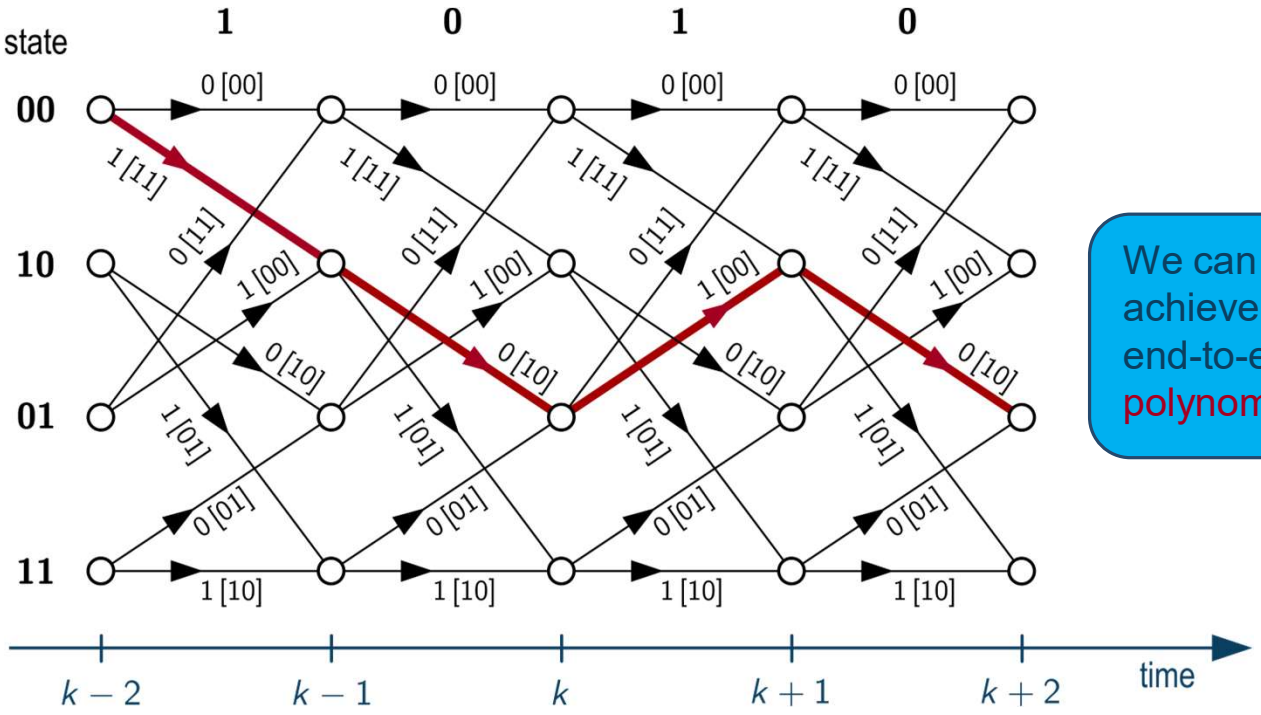
Convolutional codes (3/4)

Suppose the input bit sequence is **s** = [1 0 1 0]:



Convolutional codes (4/4)

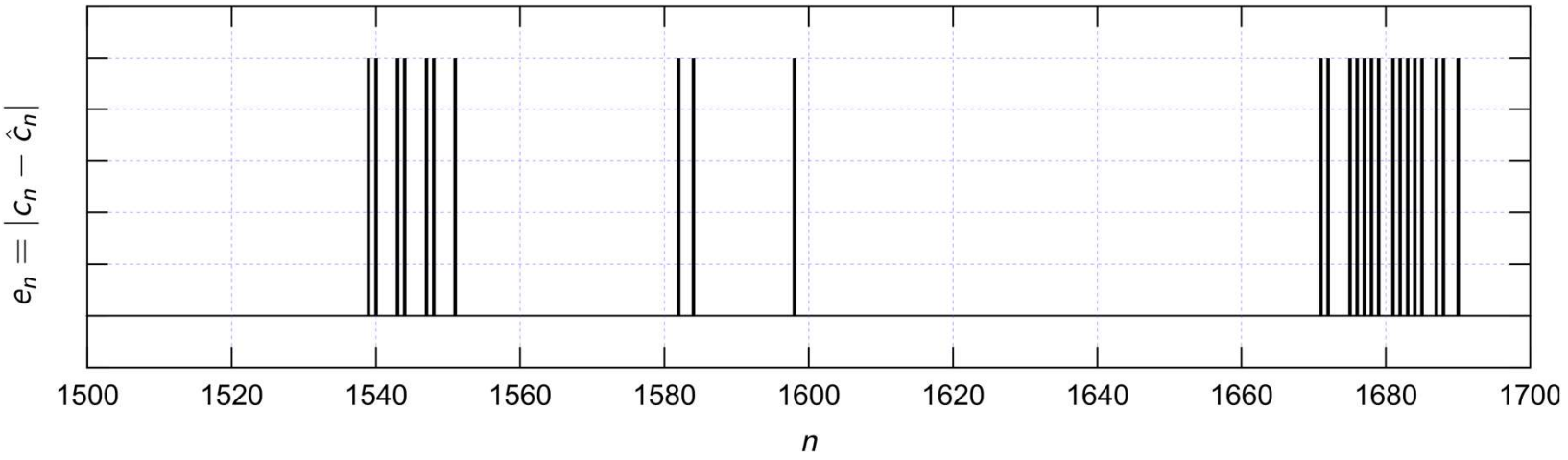
A more convenient representation is the trellis diagram, which unfolds the state transitions through time:



We can use the Viterbi algorithm to achieve exponentially decreasing end-to-end error probability P_e with polynomial-time decoding complexity

Interleaving (1/3)

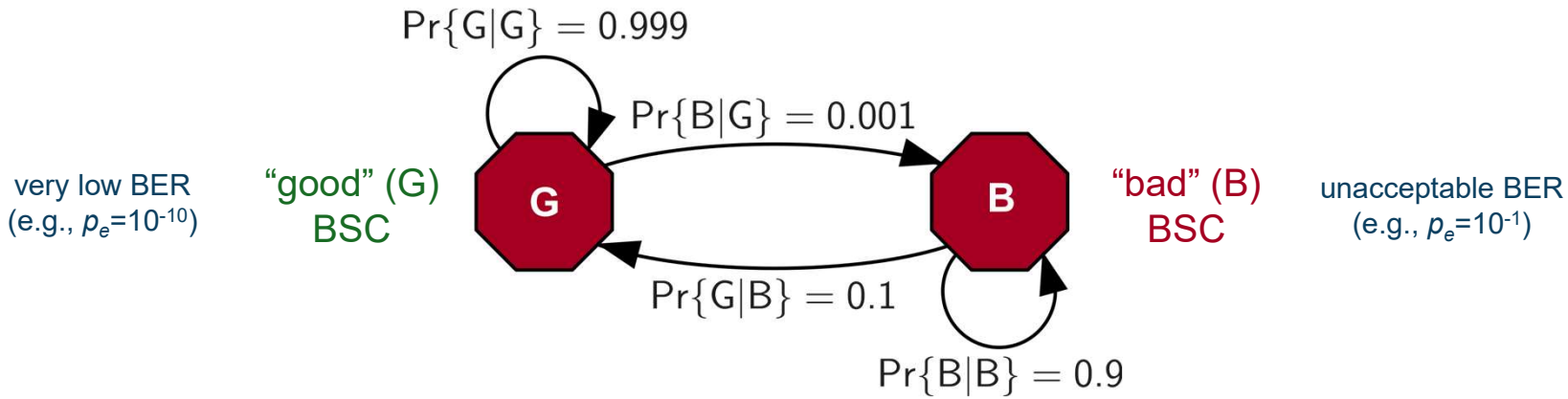
In many scenarios, the channel cannot be modeled as a BSC, due to the presence of error bursts (e.g., due to the sudden changes in the propagation environment)



Interleaving (2/3)

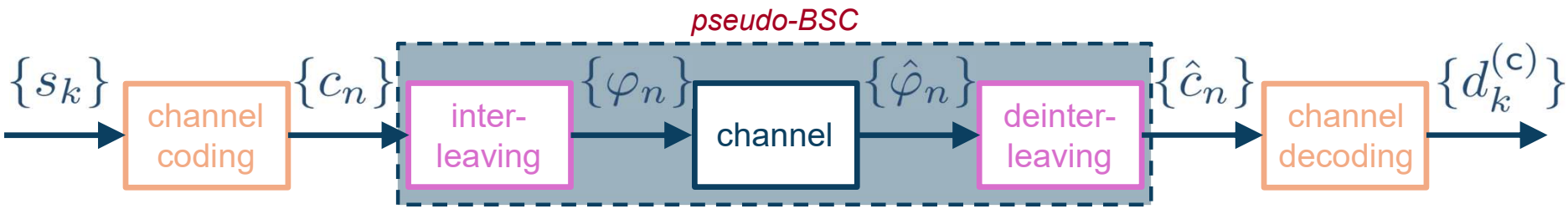
In many scenarios, the channel cannot be modeled as a BSC, due to the presence of error bursts (e.g., due to the sudden changes in the propagation environment)

This situation can be modeled as a two-state finite machine:



Interleaving (3/3)

To cope with error bursts, we can equip the channel encoder with an interleaver, which scrambles the outputs of the encoder so as to decorrelate the impact of the "bad" BSC:



Example: interleaving a block with length N=16

$$\mathbf{c} = [1011011100100110]$$

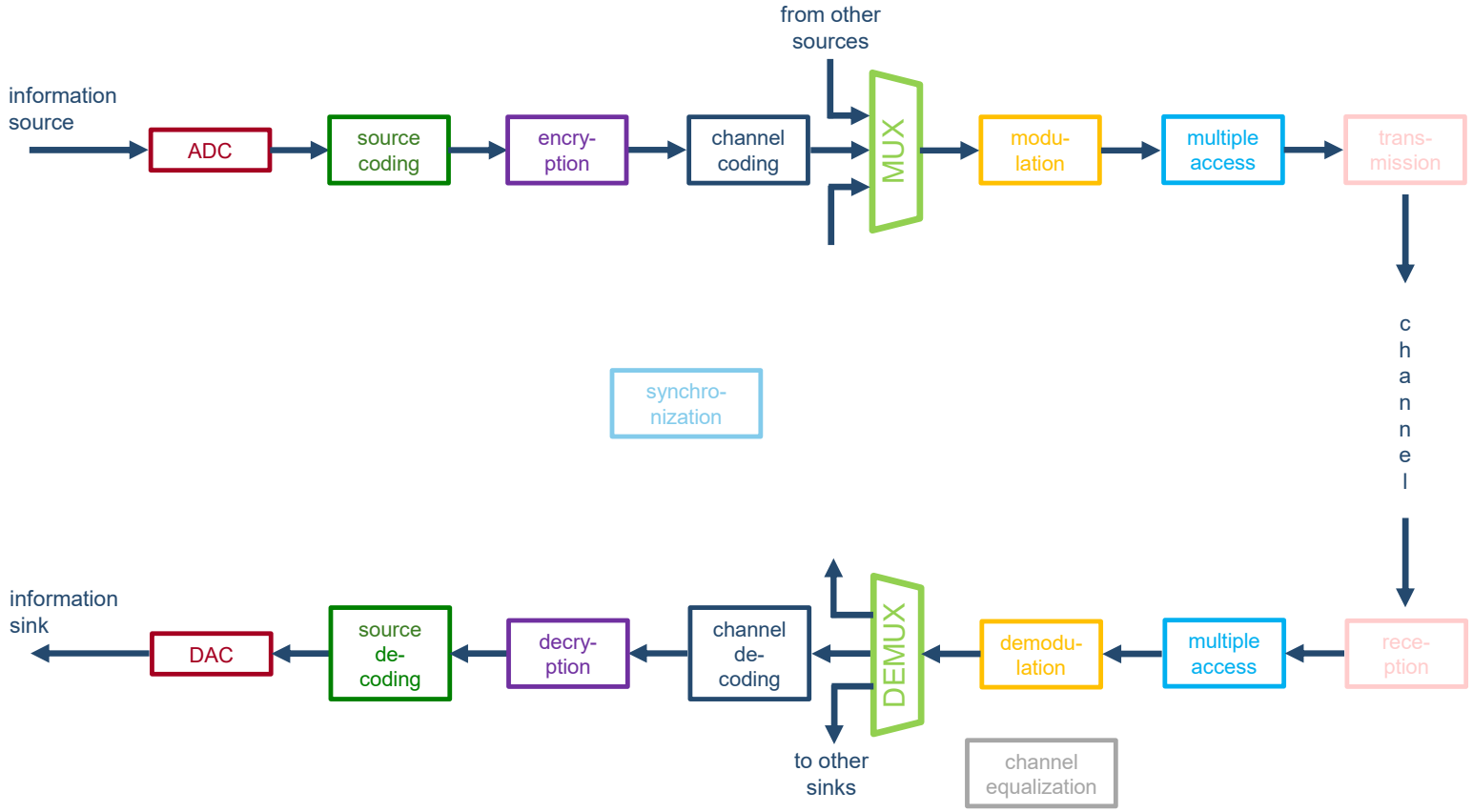
$$\boldsymbol{\varphi} = [1000010111111100]$$

step 1: sort **c** in a 4×4 matrix by rows

step 2: populate φ by reading the 4×4 matrix by columns

c_0=1	c_1=0	c_2=1	c_3=1
c_4=0	c_5=1	c_6=1	c_7=1
c_8=0	c_9=0	c_{10}=1	c_{11}=0
c_{12}=0	c_{13}=1	c_{14}=1	c_{15}=0

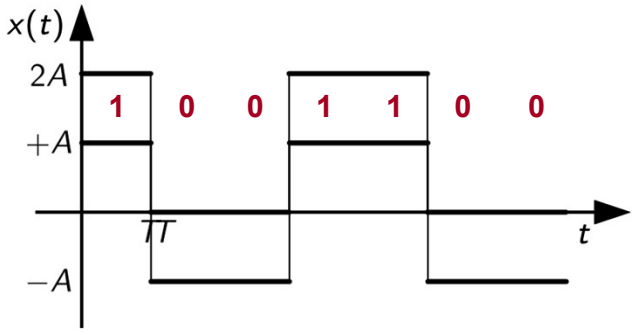
Elements of a digital communication system

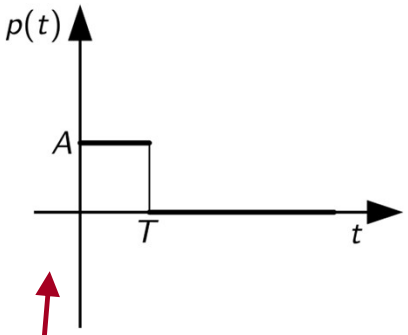


Modulation

Pulse shaping (1/3)

We need to adapt our sequence of bits into a signal which is suitable to be transmitted over a physical medium (air in wireless systems, a cable in wired systems)





symmetric NRZ signal: $x(t) = \sum_{k=-\infty}^{+\infty} (2b_k - 1)\, p(t - kT)$

shaping pulse

For convenience, let us use a symmetric NRZ modulation

Pulse shaping (2/3)

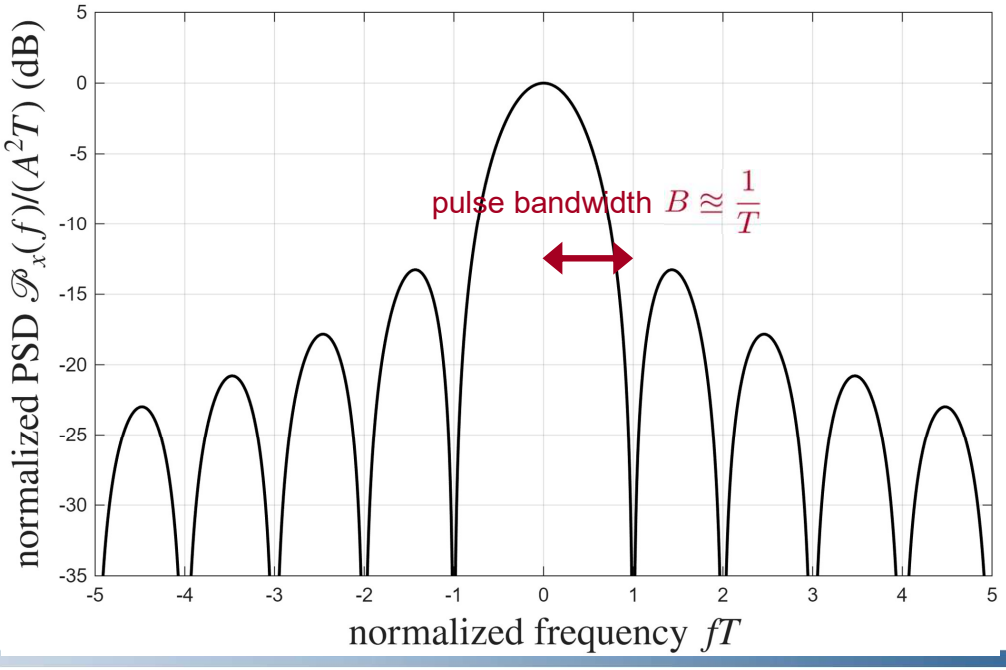
What is the (physical) bandwidth of a symmetric NRZ signal?

Let us consider the power spectral density (PSD) of $x(t)$:

$$\mathcal{P}_x(f) = \frac{A^2}{T}|P(f)|^2 = A^2\,T\,\mathrm{sinc}^2(fT)$$

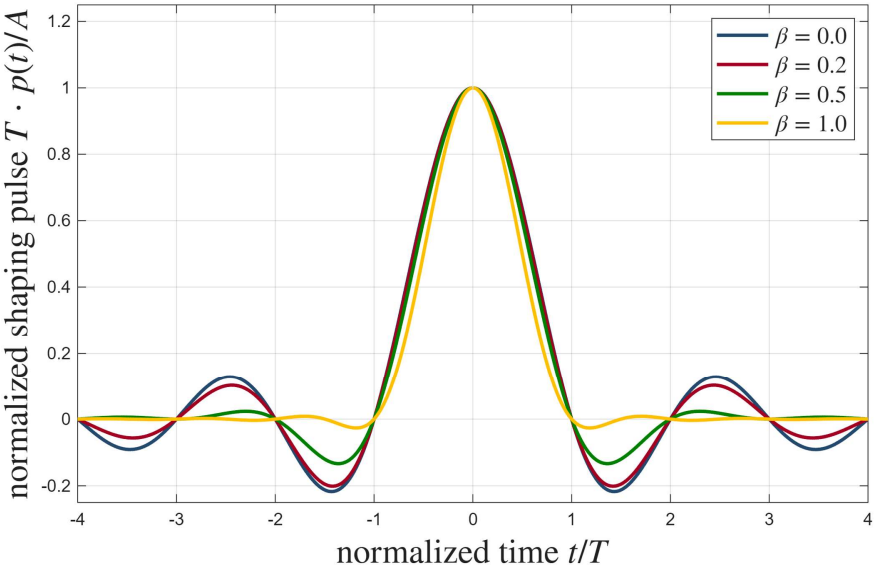
$$B \approx \frac{1}{T} = R = R_b$$

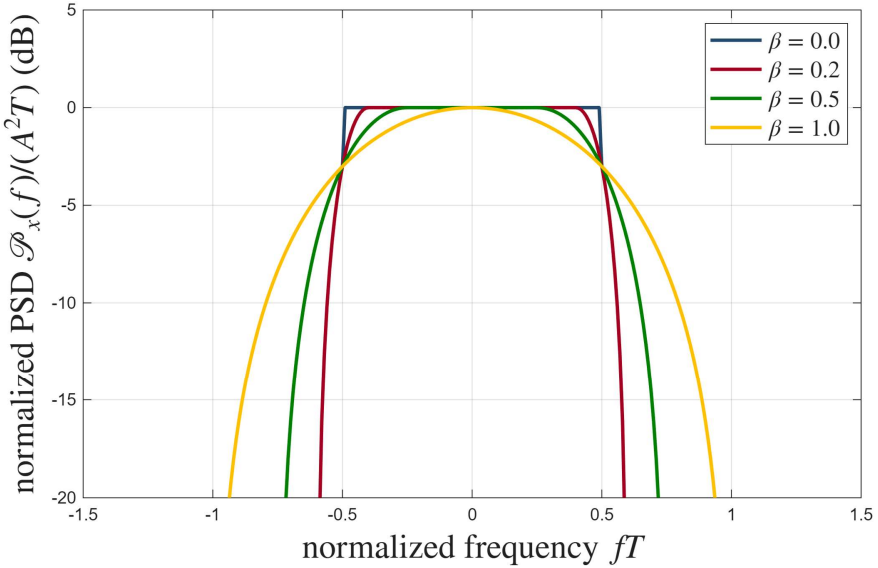
pulse bandwidth — symbol rate — bitrate



Pulse shaping (3/3)

Similar considerations can be drawn using other pulse formats, e.g, the square-root raised cosine (SRRC) with rolloff factor β:



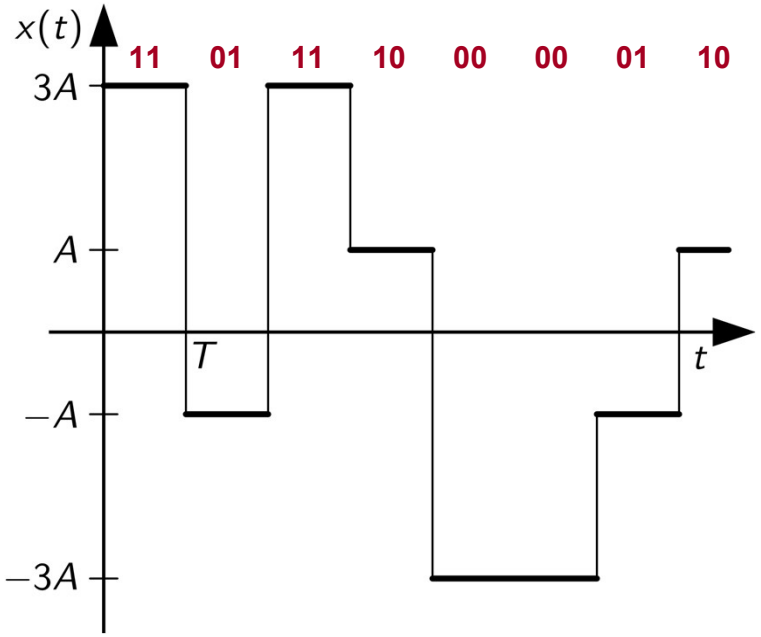


$$B \propto \frac{1}{T} = R = \cancel{R_b}$$

this is NOT true in generale!

Multilevel (baseband) modulations (1/2)

To improve the spectral efficiency R_b/B, we can map more bits into a single symbol:



the bitrate is now twice the symbol rate:

$$R_b = 2R = 2 \cdot \frac{1}{T} \approx 2B$$

$$T_b = \frac{1}{R_b} = \frac{T}{2}$$

Multilevel (baseband) modulations (2/2)

In general, when using an *M*-level mapping (i.e., when using $N_b = \log_2 M$ bits per symbol), we have

$$R_b = \log_2 M \cdot R = N_b \cdot R \approxeq N_b \cdot B$$

The spectral efficiency becomes

$$\eta_s \triangleq \frac{R_b}{B} \approxeq N_b \qquad [\text{b/s/Hz}]$$

This is particularly relevant in wireless communications, where the frequency spectrum is scarce!

Carrier modulation (1/4)

Why do we need to modulate?

- baseband is not sustainable, in terms of physical feasibility of the components: e.g., the antenna size needs to have the same order of magnitude of the signal wavelength λ (typical size: $\lambda/2$)

- bandpass yields additional benefits, e.g., in terms of multiple access

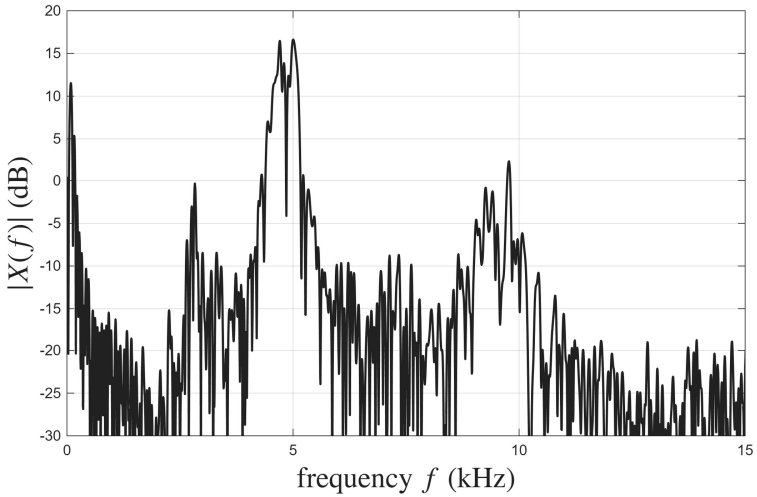
Carrier modulation (2/4)

Signal wavelength:

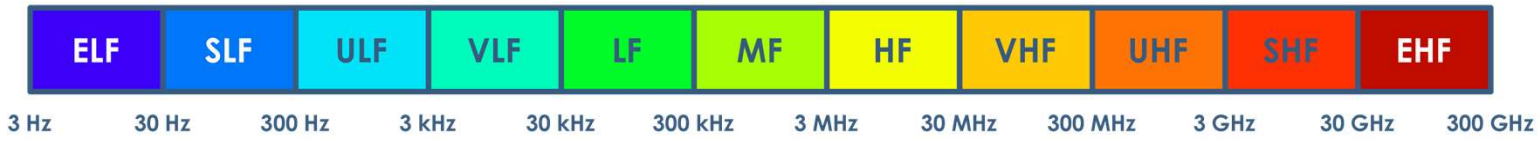
$$\lambda = \frac{c}{f}$$

speed of light, $3 \cdot 10^8$ m/s

frequency

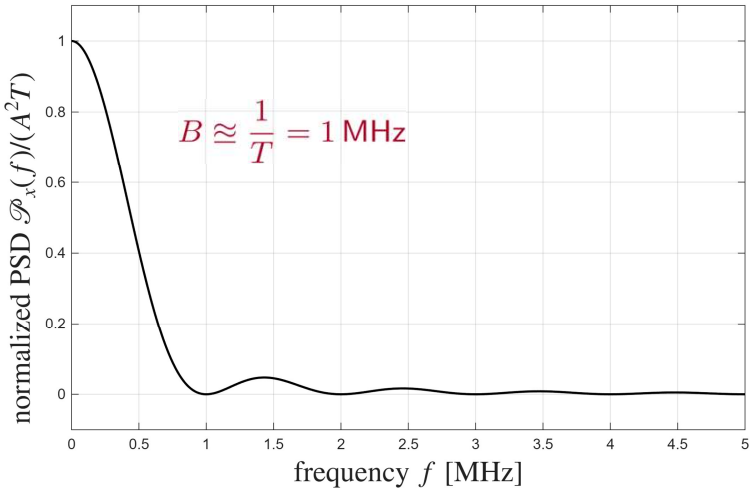


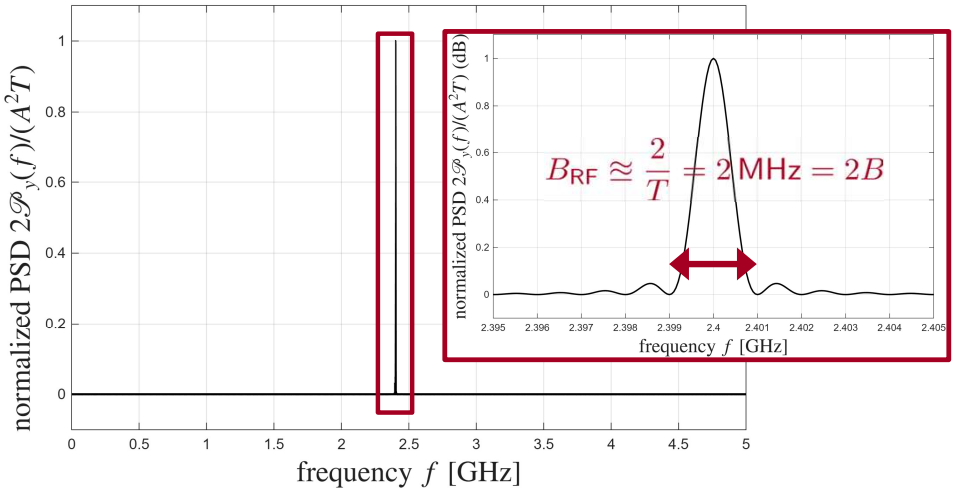
Wireless radio spectrum:



Carrier modulation (3/4)

$$x(t) = \sum_{k=-\infty}^{+\infty} (2b_k - 1)\, p(t - kT), \quad T = 1\,\mu\text{s} \qquad y(t) = x(t)\cos(2\pi f_0 t), \quad f_0 = 2.4\,\text{GHz}$$

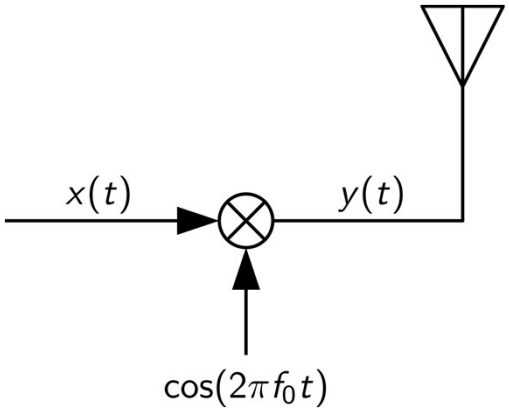


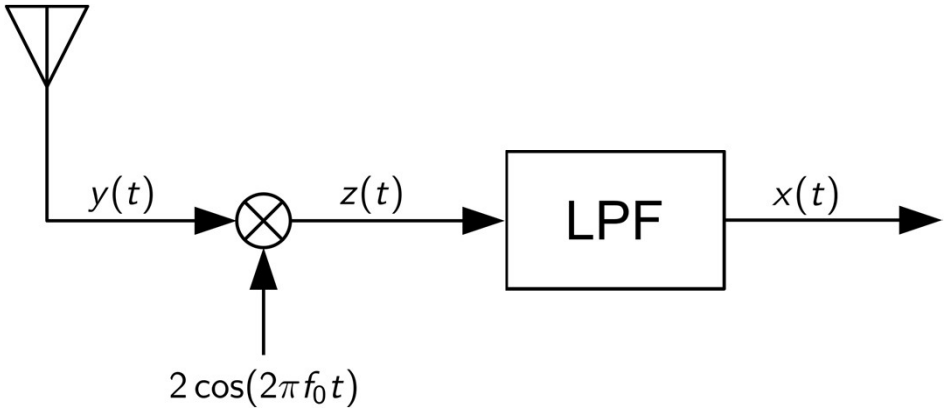


Modulation theorem: $x(t)\cos(2\pi f_0 t) \Leftrightarrow \dfrac{X(f - f_0) + X(f + f_0)}{2}$

Carrier modulation (4/4)

How can we demodulate the signal $x(t)$?





$$y(t) = x(t)\cos(2\pi f_0 t)$$

$$2\cos^2(\alpha) = 1 + \cos(2\alpha)$$

$$z(t) = y(t)2\cos(2\pi f_0 t) = 2x(t)\cos^2(2\pi f_0 t)$$

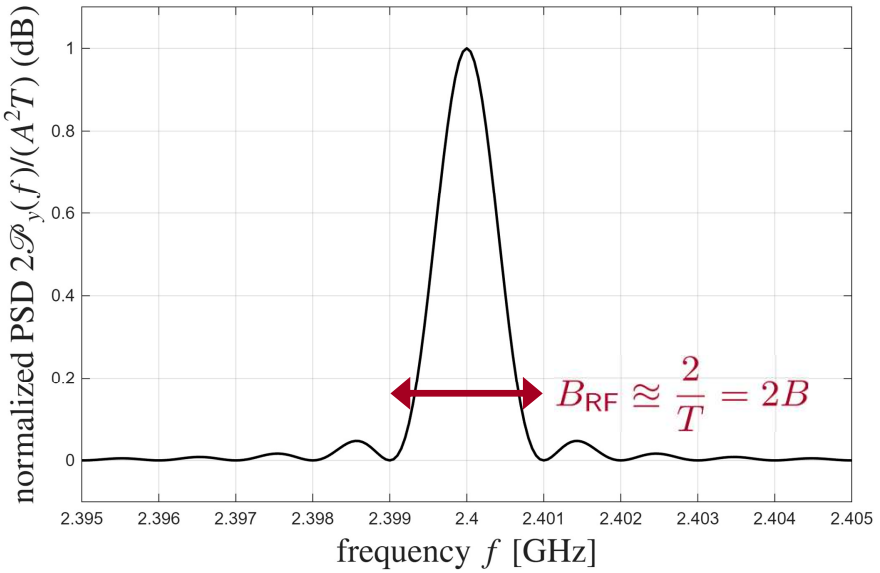
$$= x(t) + x(t)\cos\left[2\pi(2f_0)t\right]$$

I/Q modulation (1/4)

Modulation is the key to transmit a signal via radio frequency (RF) but… when moving to the RF, we decrease our spectral efficiency by a factor 2:

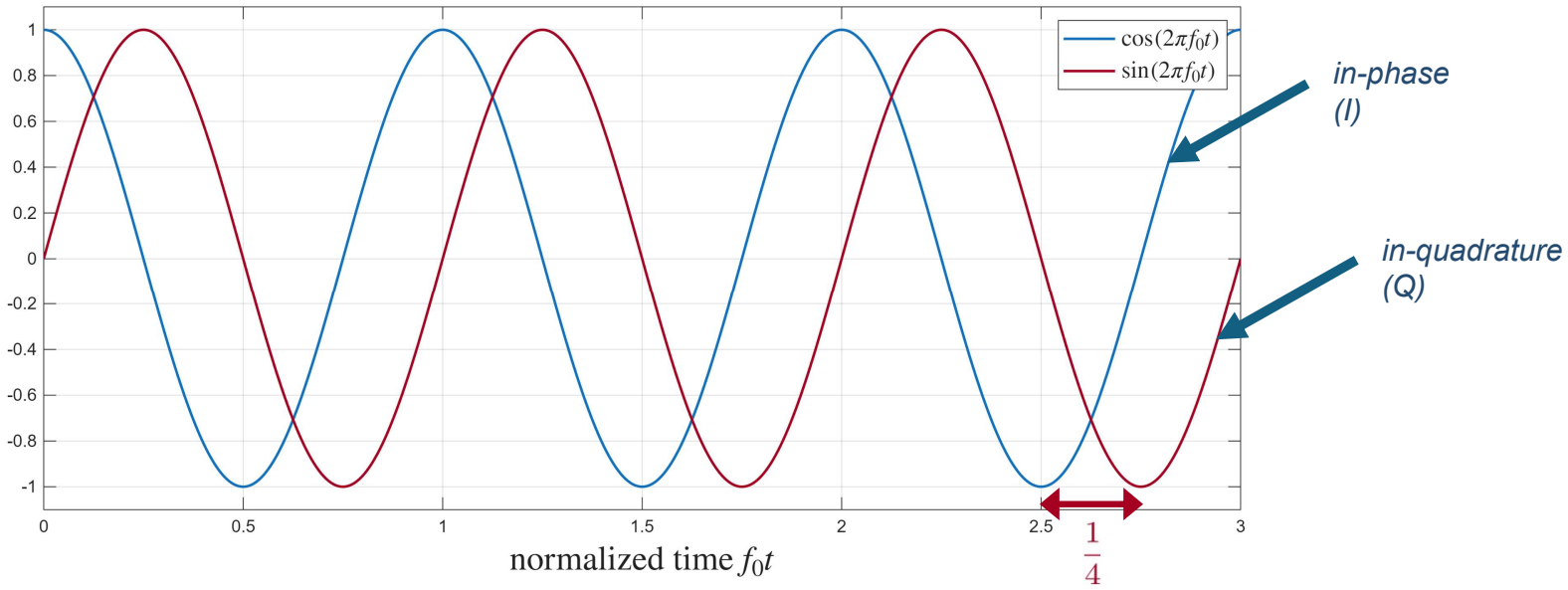
$$\eta_{\mathsf{s}} \triangleq \frac{R_b}{B_{\mathsf{RF}}} \approxeq \frac{N_b}{2} \qquad [\mathrm{b/s/Hz}]$$

How can we improve η_{s}, while exploiting the intuition of multiple bits per symbol?



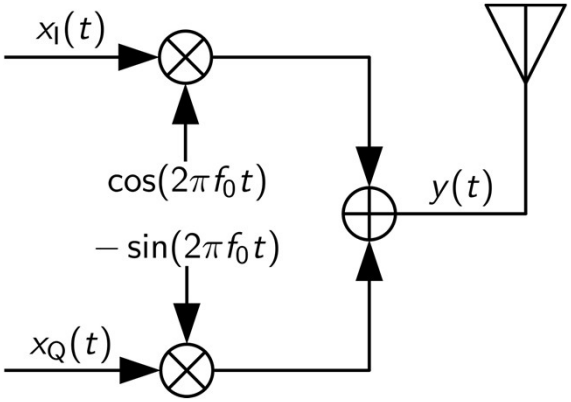
I/Q modulation (2/4)

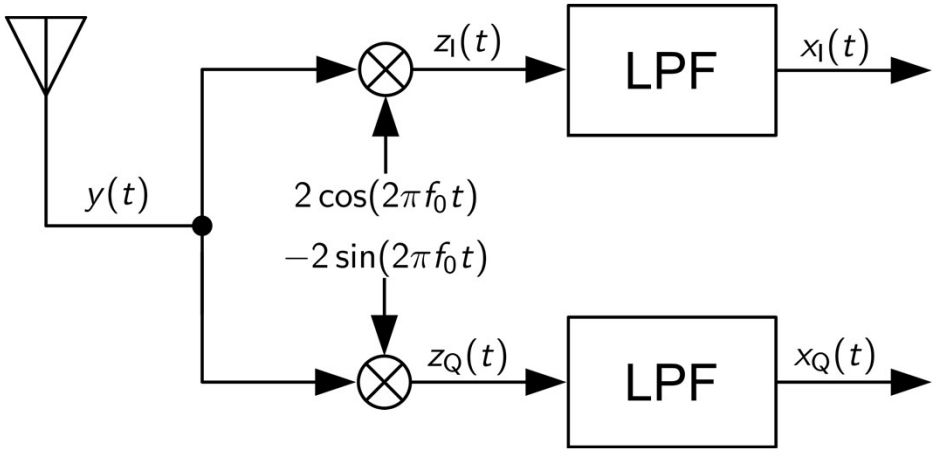
Let us multiplex two streams on the same carrier, by leveraging two orthogonal waveforms:



I/Q modulation (3/4)

Let us feed the modulator with two independent streams, each with bitrate R_b:





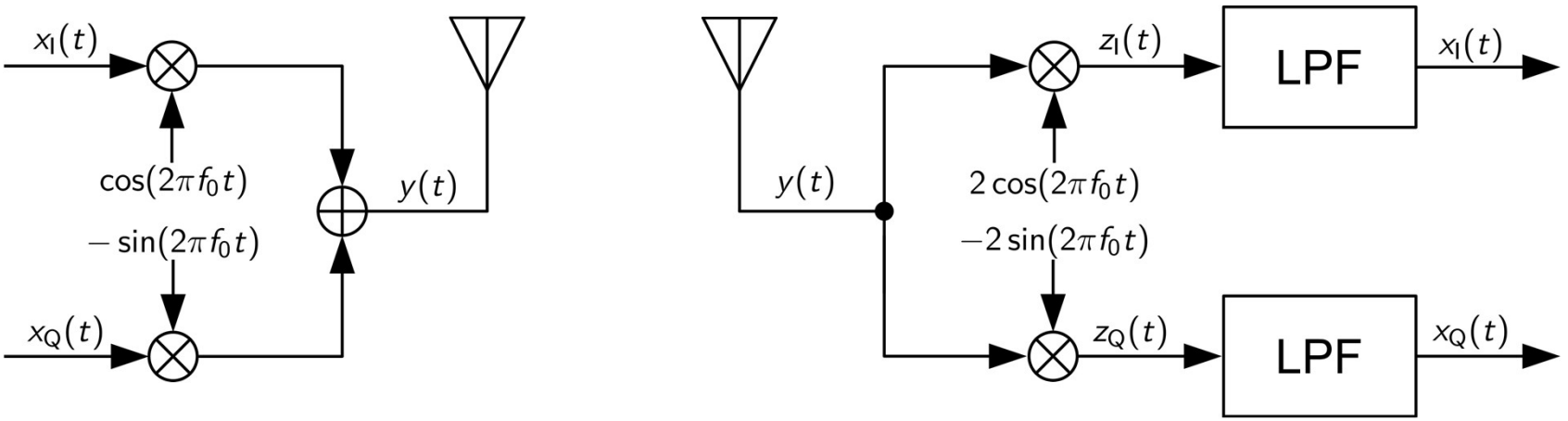
$$y(t) = x_I(t)\cos(2\pi f_0 t) - x_Q(t)\sin(2\pi f_0 t)$$

$$2\cos^2(\alpha) = 1 + \cos(2\alpha)$$
$$2\sin^2(\alpha) = 1 - \cos(2\alpha)$$
$$2\sin(\alpha)\cos(\alpha) = \sin(2\alpha)$$

$$z_I(t) = y(t)2\cos(2\pi f_0 t) = 2x_I(t)\cos^2(2\pi f_0 t) - 2x_Q(t)\sin(2\pi f_0 t)\cos(2\pi f_0 t)$$
$$= x_I(t) + x_I(t)\cos[2\pi(2f_0)t] - x_Q(t)\sin[2\pi(2f_0)t]$$
$$z_Q(t) = -y(t)2\sin(2\pi f_0 t) = 2x_Q(t)\sin^2(2\pi f_0 t) - 2x_I(t)\sin(2\pi f_0 t)\cos(2\pi f_0 t)$$
$$= x_Q(t) - x_Q(t)\cos[2\pi(2f_0)t] - x_I(t)\sin[2\pi(2f_0)t]$$

I/Q modulation (4/4)

Let us feed the modulator with two independent streams, each with bitrate R_b:



$$\eta_s \triangleq \frac{R_{b,I} + R_{b,Q}}{B_{RF}} = \frac{2R_b}{2B} \approx N_b \qquad [\text{b/s/Hz}]$$

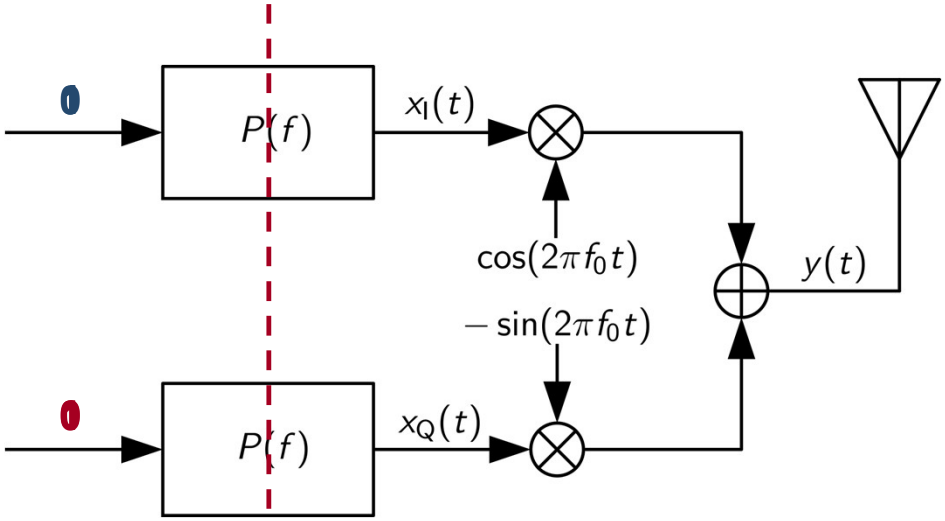
PSK and QAM constellations (1/5)

To cast one single binary stream into the I/Q scheme, for instance we can assign odd bits to the I branch and even bits to the Q branch:

01101110101100…

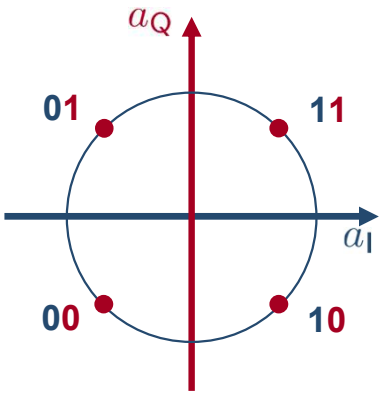
$$x_{\mathrm{I}}(t) = \sum_{\ell=-\infty}^{+\infty} (2b_{2\ell} - 1)\, p(t - \ell T)$$

$$x_{\mathrm{Q}}(t) = \sum_{\ell=-\infty}^{+\infty} (2b_{2\ell+1} - 1)\, p(t - \ell T)$$

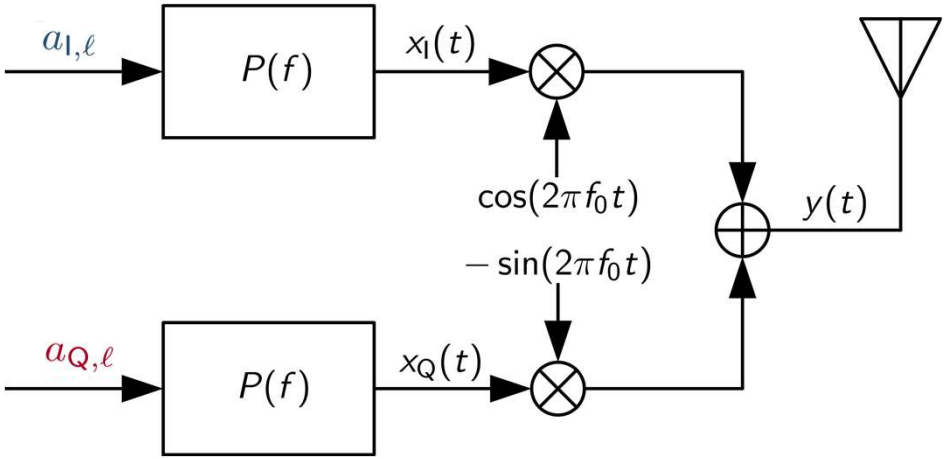


PSK and QAM constellations (2/5)

This is equivalent to map the input bits in complex symbols over the I/Q plane:



$b_{2\ell}$ $b_{2\ell+1}$	$a_{I,\ell}$	$a_{Q,\ell}$
00	-1	-1
01	-1	+1
10	+1	-1
11	+1	+1

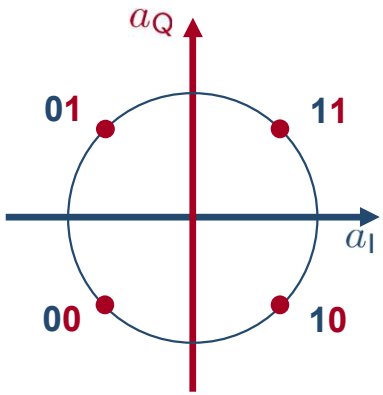


$$x_I(t) = \sum_{\ell=-\infty}^{+\infty} (2b_{2\ell} - 1)\, p(t - \ell T) = \sum_{\ell=-\infty}^{+\infty} a_{I,\ell}\, p(t - \ell T)$$
$$x_Q(t) = \sum_{\ell=-\infty}^{+\infty} (2b_{2\ell+1} - 1)\, p(t - \ell T) = \sum_{\ell=-\infty}^{+\infty} a_{Q,\ell}\, p(t - \ell T)$$

With this scheme, we can send N_b=$\log_2 4$=2 bits per symbol, thus achieving η_s=2 b/s/Hz

PSK and QAM constellations (3/5)

This scheme goes under the name of quaternary phase shift keying (QPSK):



$b_{2\ell}\, b_{2\ell+1}$	$a_{I,\ell}$	$a_{Q,\ell}$
00	-1	-1
01	-1	+1
10	+1	-1
11	+1	+1

Question: Can we generalize to an arbitrary number of bits (including odd N_b's)?

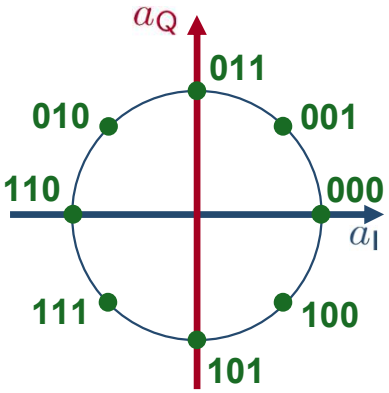
Answer: Yes, using phase shift keying (PSK) and quadrature amplitude modulation (QAM) constellations!

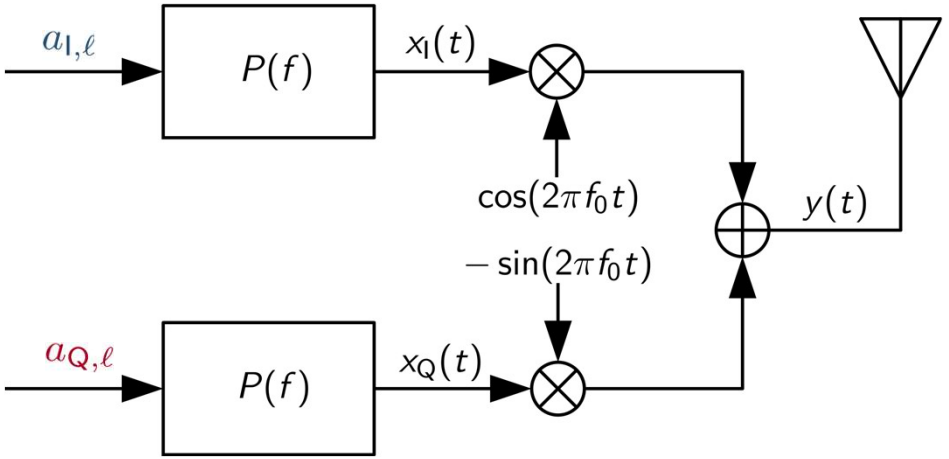
PSK and QAM constellations (4/5)

Important: The I/Q modulator block diagram remains the same!

M=8 (N_b=3)

→ 8-PSK:





$b_{2\ell}\, b_{2\ell+1}$	$a_{I,\ell}$	$a_{Q,\ell}$	$b_{2\ell}\, b_{2\ell+1}$	$a_{I,\ell}$	$a_{Q,\ell}$
000	$+1$	0	100	$+1/\sqrt{2}$	$-1/\sqrt{2}$
001	$+1/\sqrt{2}$	$+1/\sqrt{2}$	101	0	-1
010	$-1/\sqrt{2}$	$+1/\sqrt{2}$	110	-1	0
011	0	$+1$	111	$-1/\sqrt{2}$	$-1/\sqrt{2}$

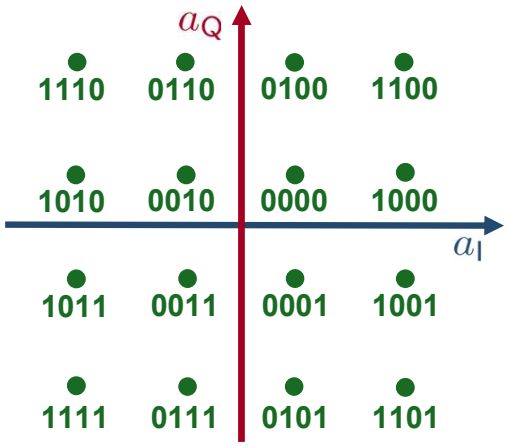
$$x_I(t) = \sum_{\ell=-\infty}^{+\infty} a_{I,\ell}\, p(t-\ell T)$$
$$x_Q(t) = \sum_{\ell=-\infty}^{+\infty} a_{Q,\ell}\, p(t-\ell T)$$

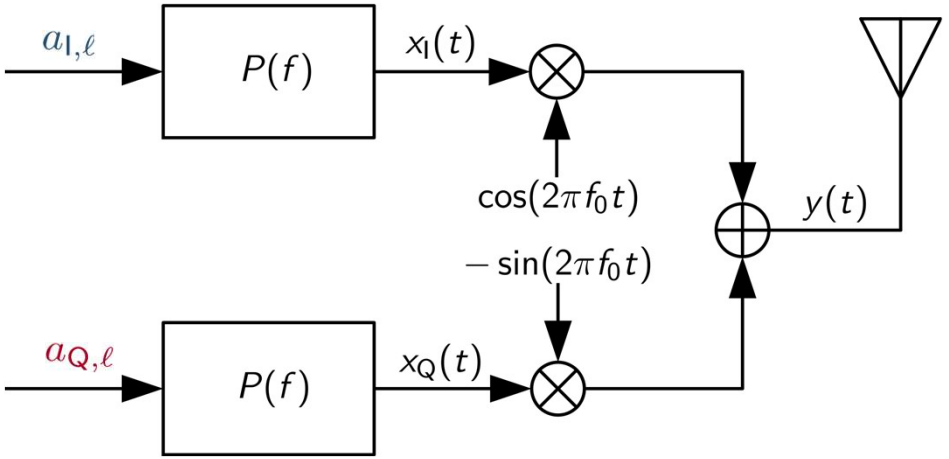
PSK and QAM constellations (5/5)

Important: The I/Q modulator block diagram remains the same!

M=16 (N_b=4)

→ 16-QAM:

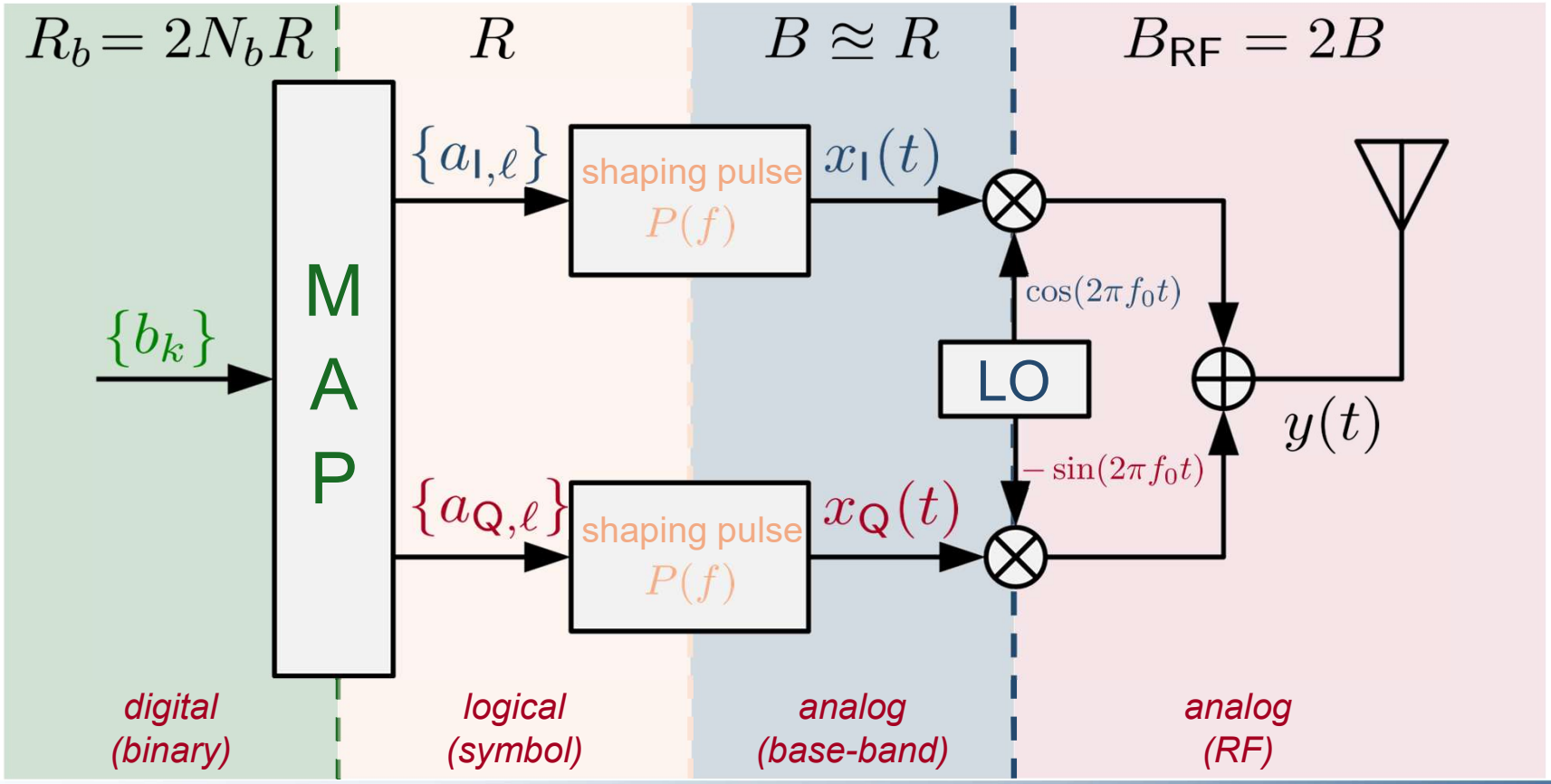




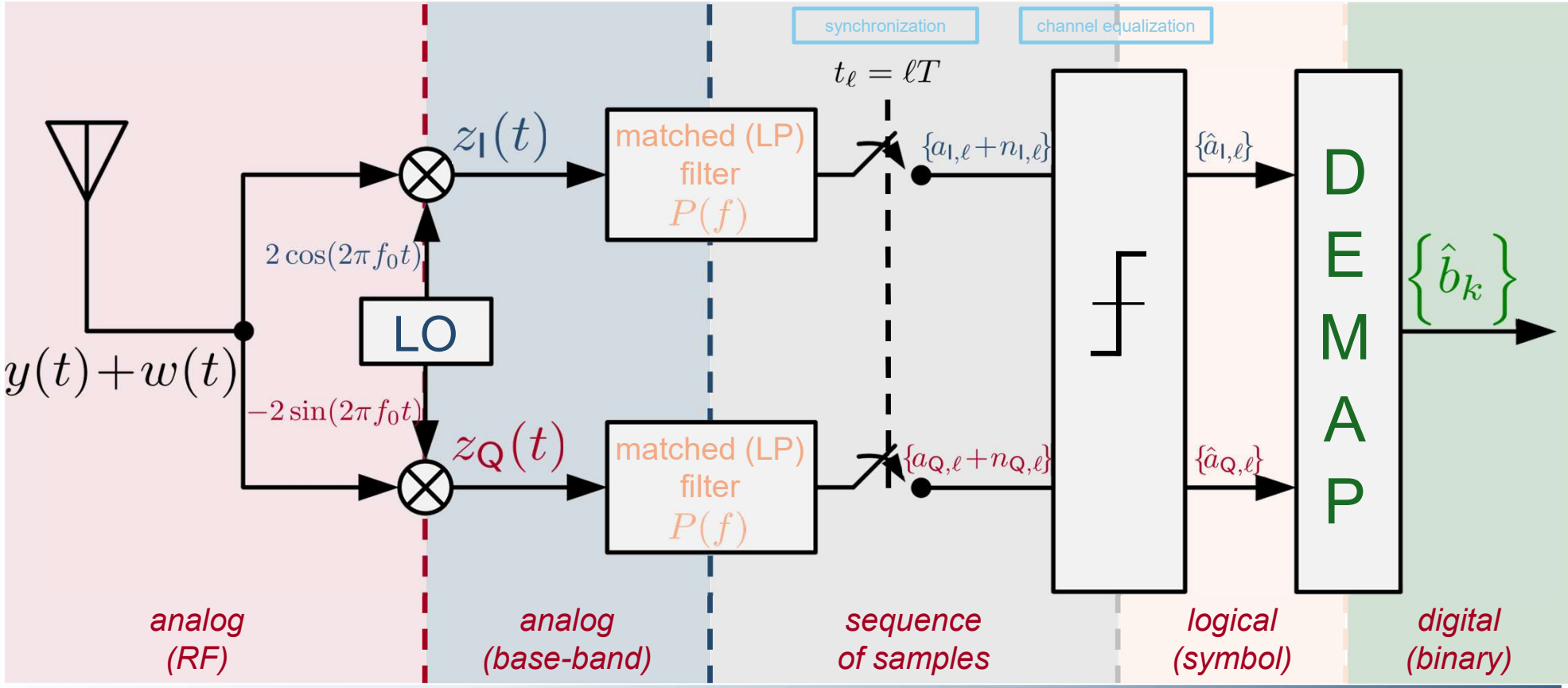
$$x_I(t) = \sum_{\ell=-\infty}^{+\infty} a_{I,\ell}\, p(t-\ell T)$$
$$x_Q(t) = \sum_{\ell=-\infty}^{+\infty} a_{Q,\ell}\, p(t-\ell T)$$

Exercise: Derive the mapping for the binary PSK (BPSK)

The I-Q modulator

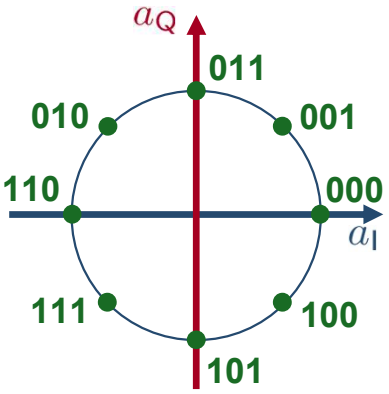


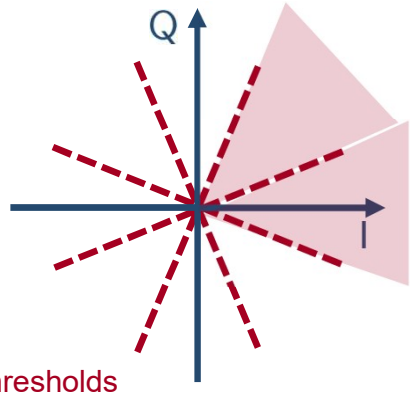
The I-Q demodulator (1/4)



The I-Q demodulator (2/4)

Example of threshold-based detection: let us suppose to use 8-PSK





decision thresholds

any sample with coordinates $(a_{I,\ell}+n_{I,\ell}, a_{Q,\ell}+n_{Q,\ell})$ falling in this region gives the symbol $\hat{a}_{I,\ell} = 1/\sqrt{2}, \hat{a}_{Q,\ell} = 1/\sqrt{2}$

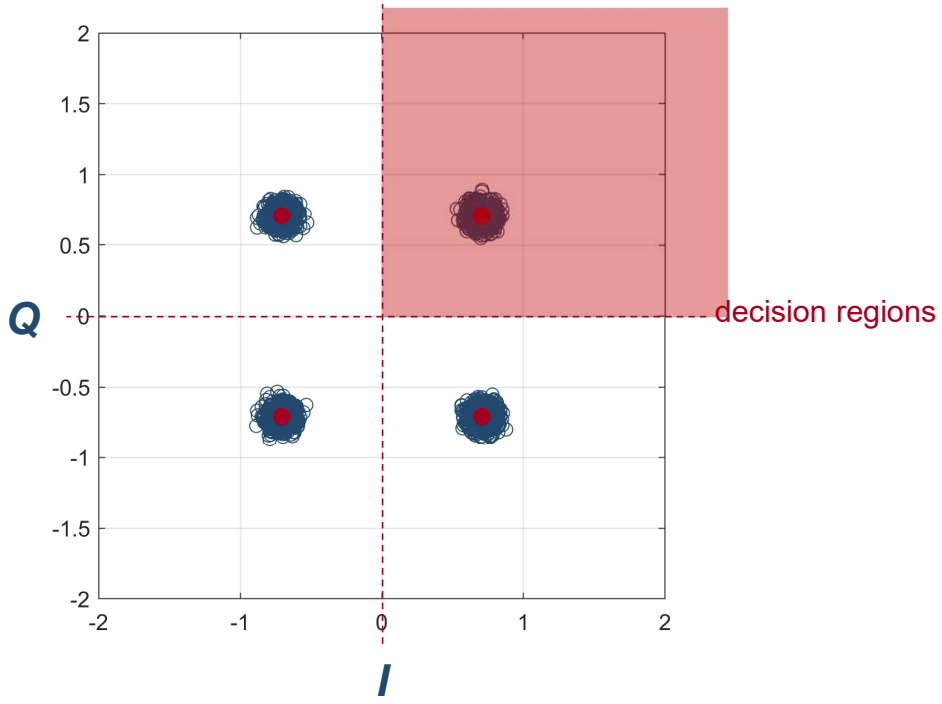
after de-mapping, we have **001**

any sample with coordinates $(a_{I,\ell}+n_{I,\ell}, a_{Q,\ell}+n_{Q,\ell})$ falling in this region gives the symbol $\hat{a}_{I,\ell} = 1, \hat{a}_{Q,\ell} = 0$

after de-mapping, we have **000**

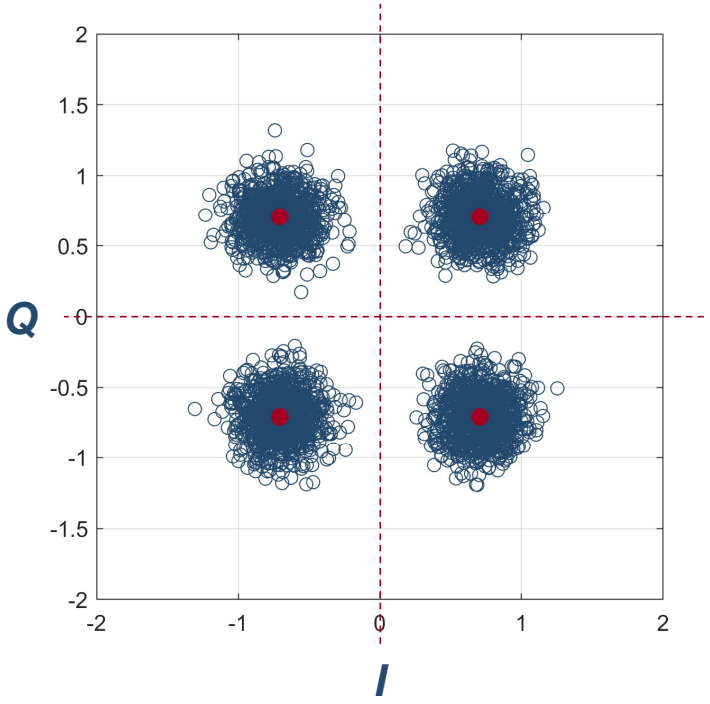
The I-Q demodulator (3/4)

What happens when using a QPSK modulation in a high signal-to-noise ratio (SNR) scenario?

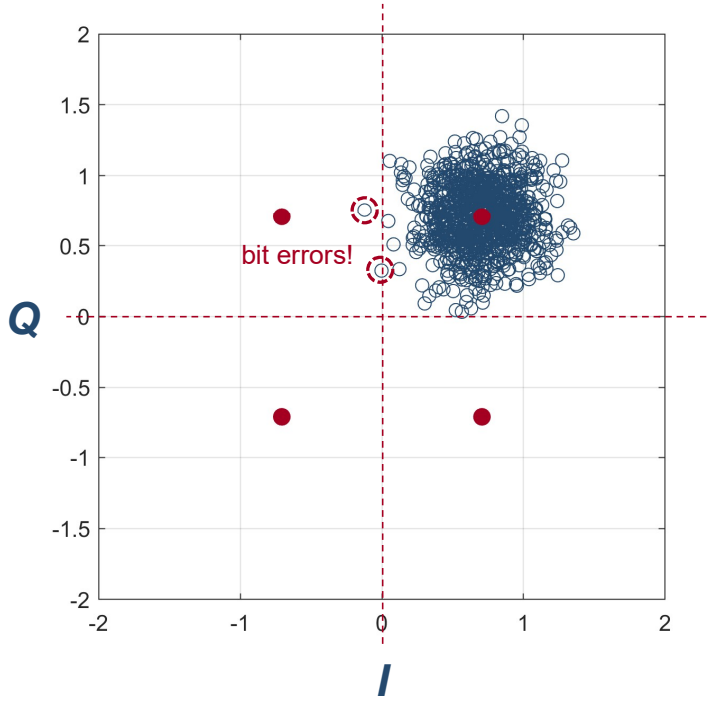


The I-Q demodulator (4/4)

medium-SNR scenario



low-SNR scenario



Impact of additive white Gaussian noise (AWGN) (1/5)

How can we quantify the impact of AWGN in the demodulation performance?

The main figure of merit becomes the signal-to-noise ratio (SNR), in the form of

$$\frac{E_b}{N_0} = \frac{S}{N} \cdot \frac{B}{R_b}$$

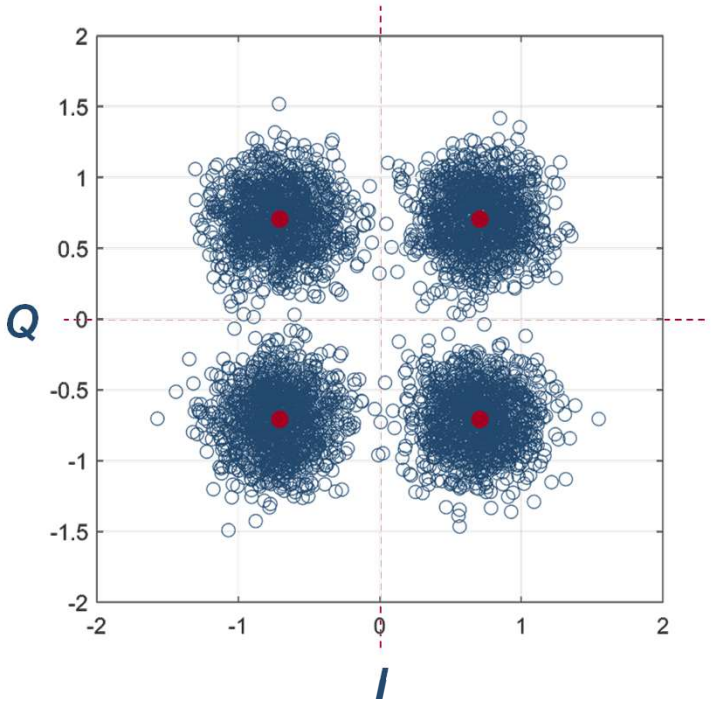
- E_b: energy per bit
- N_0: noise power spectral density
- S: useful signal power
- N: noise power
- B: signal bandwidth
- R_b: bit rate

The performance of a digital communication system can be assessed measuring the bit error rate (BER) – which occurs when $\hat{b}_k \neq b_k$ – and the symbol error rate (SER) – which occurs when $(\hat{a}_{\mathrm{I},\ell} + j\hat{a}_{\mathrm{Q},\ell}) \neq (a_{\mathrm{I},\ell} + ja_{\mathrm{Q},\ell})$

Impact of additive white Gaussian noise (AWGN) (2/5)

QPSK @

$E_b/N_0 = 10$ dB

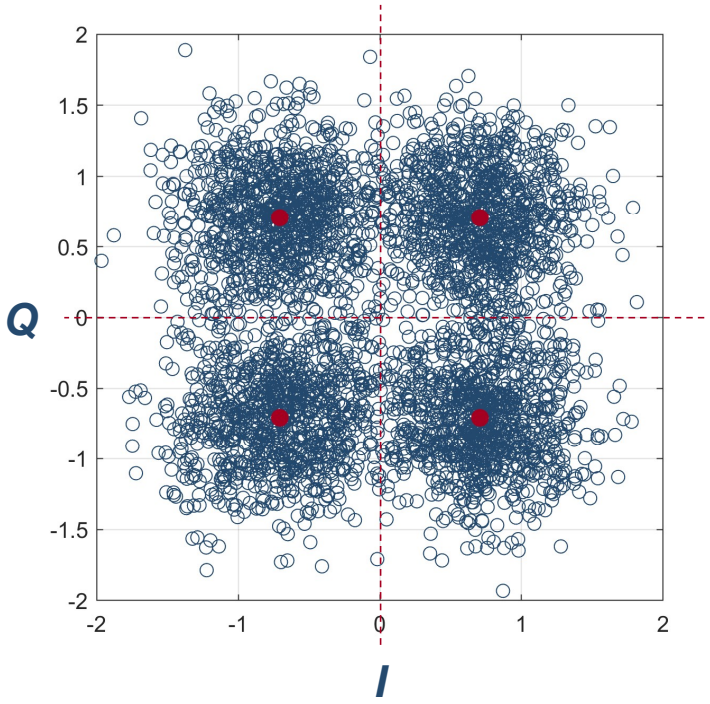


Impact of additive white Gaussian noise (AWGN) (3/5)

The situation becomes dramatic when either increasing the AWGN power…

QPSK @

$E_b/N_0 = 3\,\mathrm{dB}$

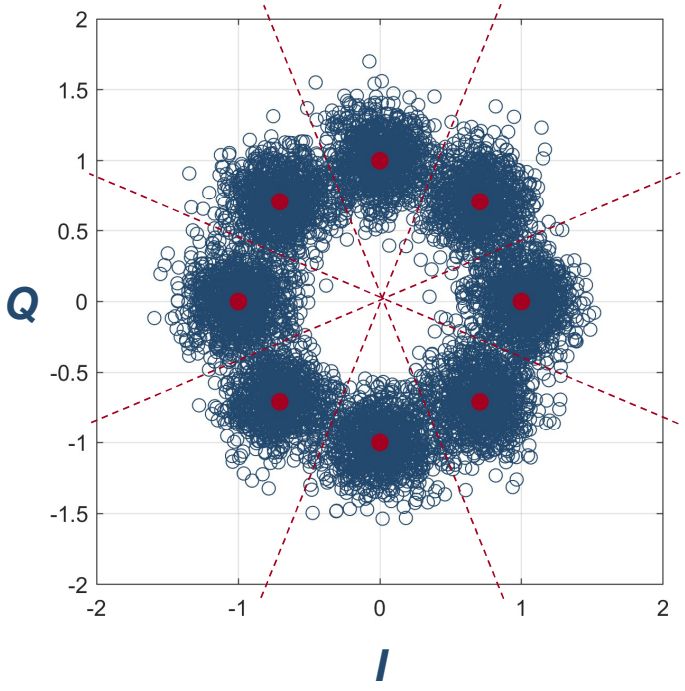


Impact of additive white Gaussian noise (AWGN) (4/5)

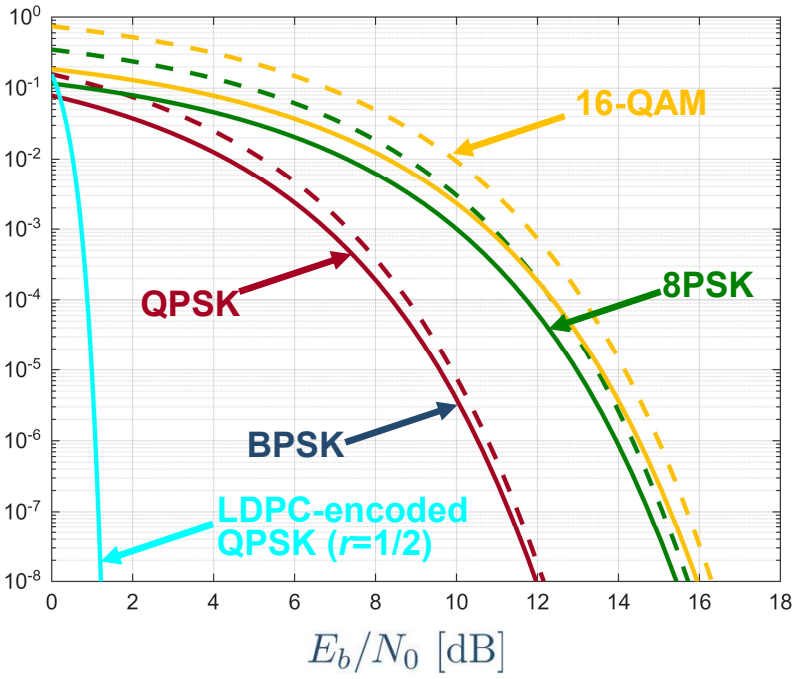
…or when increasing the modulation order

8PSK @

$E_b/N_0 = 7\,\mathrm{dB}$



Performance with AWGN



—— *BER* p_e

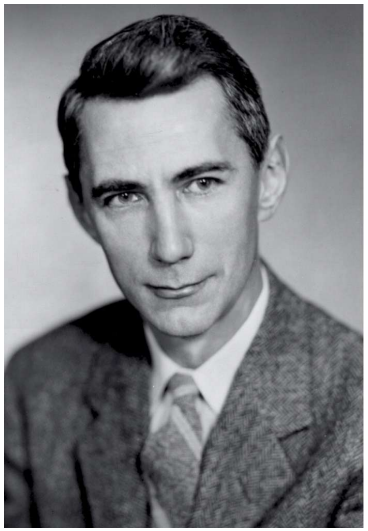
-------- *SER* $\approx N_b p_e$

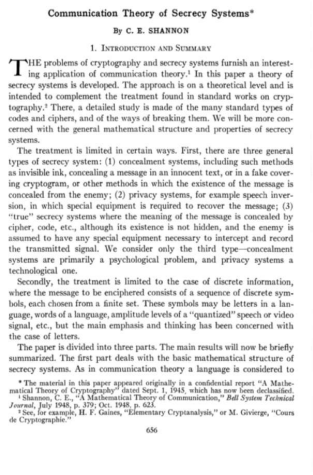
Shannon capacity (1/3)

How can we evaluate the performance of a system?

$$R_b \leq \mathcal{C} = B \cdot \log_2\left(1 + \frac{S}{N}\right) = B \cdot \log_2\left(1 + \frac{E_b}{N_0} \cdot \frac{R_b}{B}\right)$$

Shannon capacity





Communication Theory of Secrecy Systems*

By C. E. SHANNON

1. Introduction and Summary

THE problems of cryptography and secrecy systems furnish an interesting application of communication theory.[1] In this paper a theory of secrecy systems is developed. The approach is on a theoretical level and is intended to complement the treatment found in standard works on cryptography.[2] There, a detailed study is made of the many standard types of codes and ciphers, and of the ways of breaking them. We will be more concerned with the general mathematical structure and properties of secrecy systems.

The treatment is limited in certain ways. First, there are three general types of secrecy system: (1) concealment systems, including such methods as invisible ink, concealing a message in an innocent text, or in a fake covering cryptogram, or other methods in which the existence of the message is concealed from the enemy; (2) privacy systems, for example speech inversion, in which special equipment is required to recover the message; (3) "true" secrecy systems where the meaning of the message is concealed by cipher, code, etc., although its existence is not hidden, and the enemy is assumed to have any special equipment necessary to intercept and record the transmitted signal. We consider only the third type—concealment systems are primarily a psychological problem, and privacy systems a technological one.

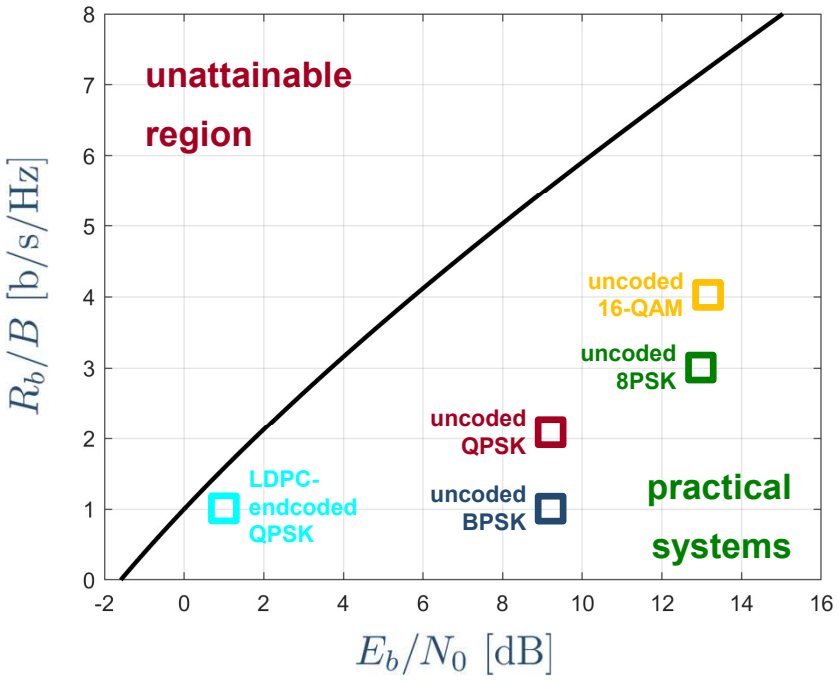
Secondly, the treatment is limited to the case of discrete information, where the message to be enciphered consists of a sequence of discrete symbols, each chosen from a finite set. These symbols may be letters in a language, words of a language, amplitude levels of a "quantized" speech or video signal, etc., but the main emphasis and thinking has been concerned with the case of letters.

The paper is divided into three parts. The main results will now be briefly summarized. The first part deals with the basic mathematical structure of secrecy systems. As in communication theory a language is considered to

* The material in this paper appeared originally in a confidential report "A Mathematical Theory of Cryptography" dated Sept. 1, 1945, which has now been declassified.
[1] Shannon, C. E., "A Mathematical Theory of Communication," *Bell System Technical Journal*, July 1948, p. 379; Oct. 1948, p. 623.
[2] See, for example, H. F. Gaines, "Elementary Cryptanalysis," or M. Givierge, "Cours de Cryptographie."

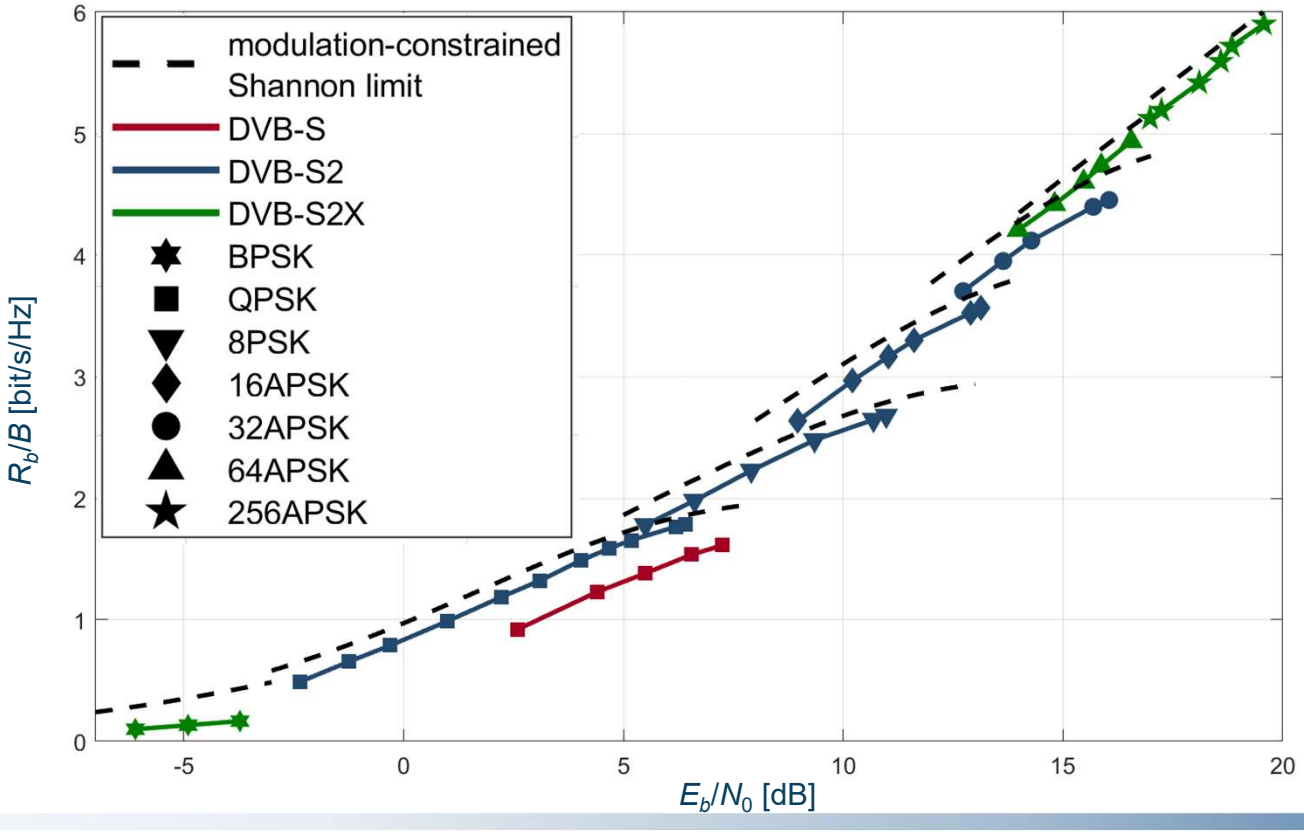
656

Shannon capacity (2/3)



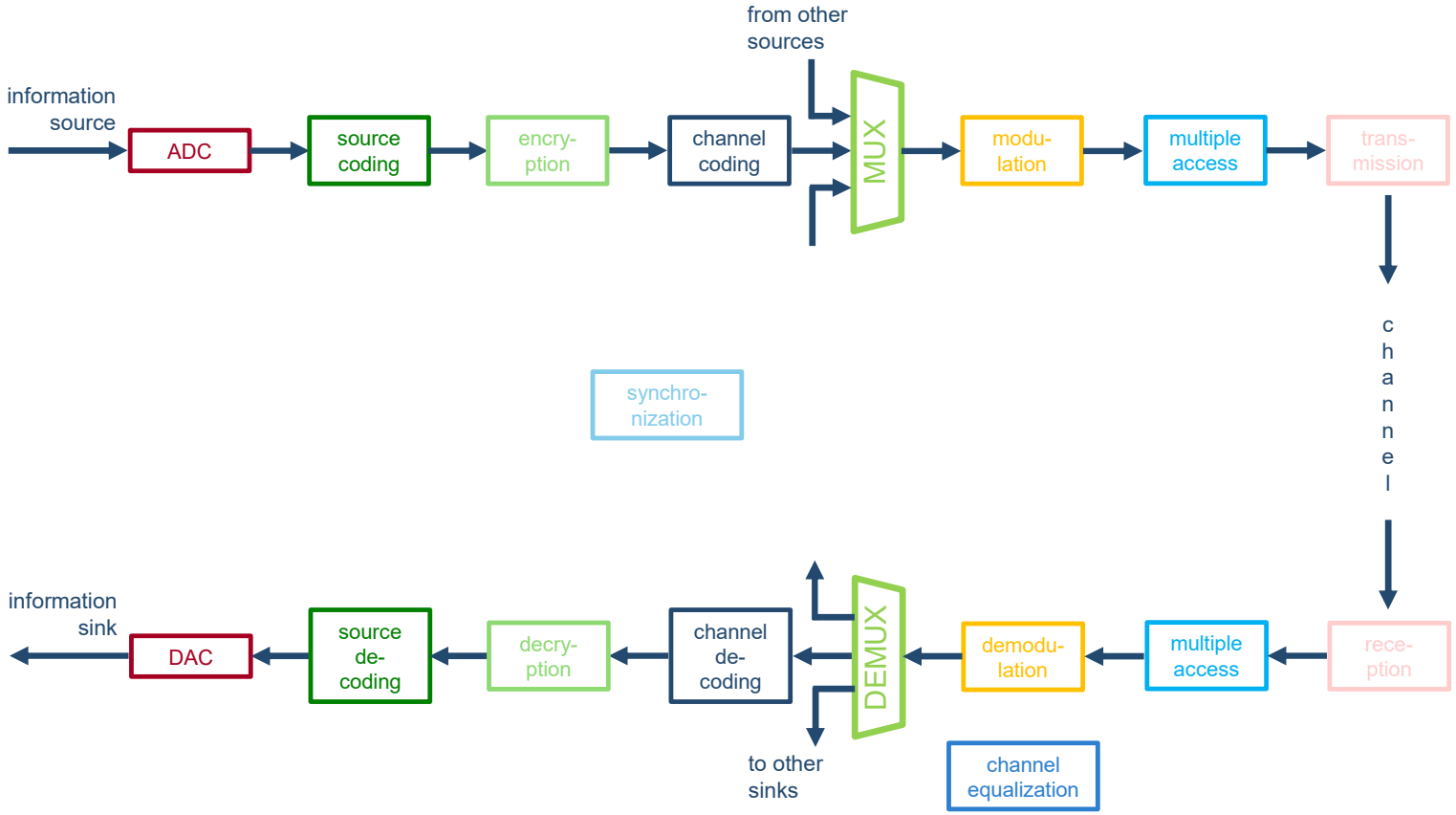
Shannon capacity (3/3)

Example: adaptive coding and modulation used by satellite standards



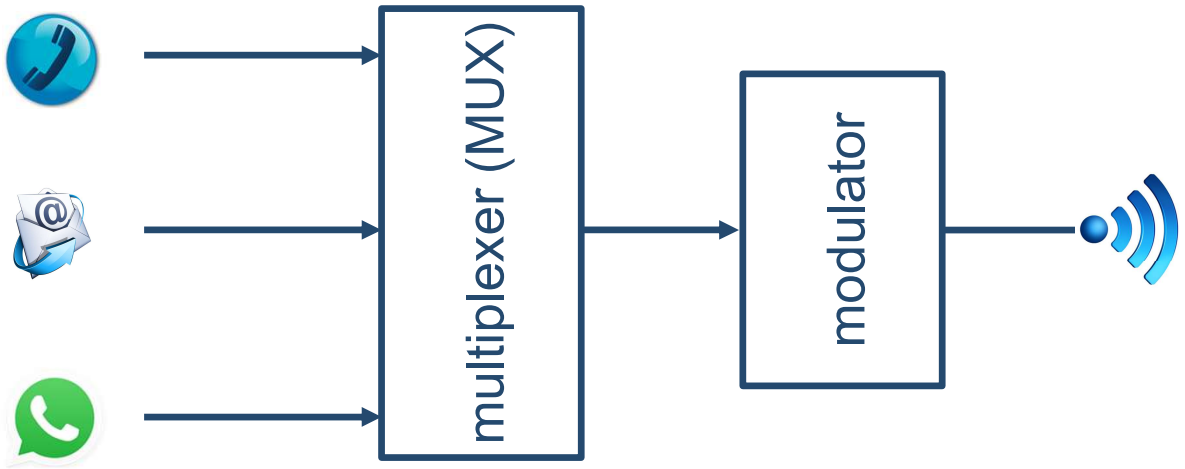
Multiplexing

Elements of a digital communication system



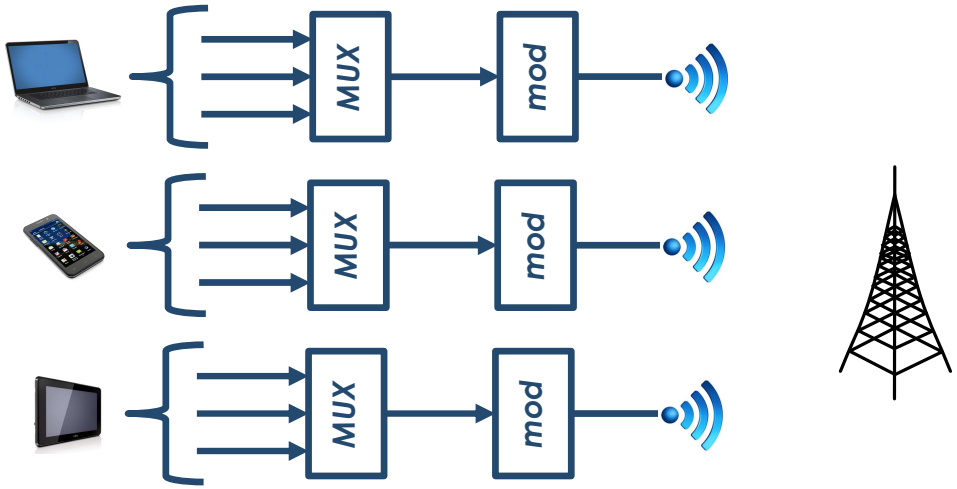
Multiplexing vs. multiple access (1/3)

Multiplexing: separating different flows at the same transmit side



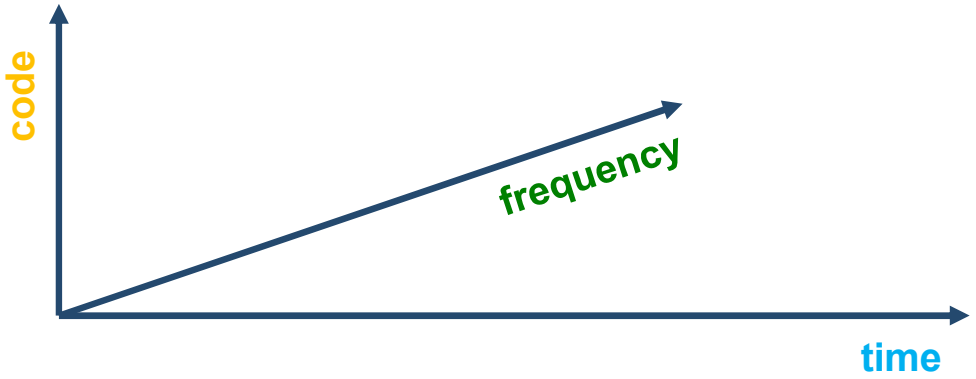
Multiplexing vs. multiple access (2/3)

Multiple access: separating different users at the receiver side



Multiplexing vs. multiple access (3/3)

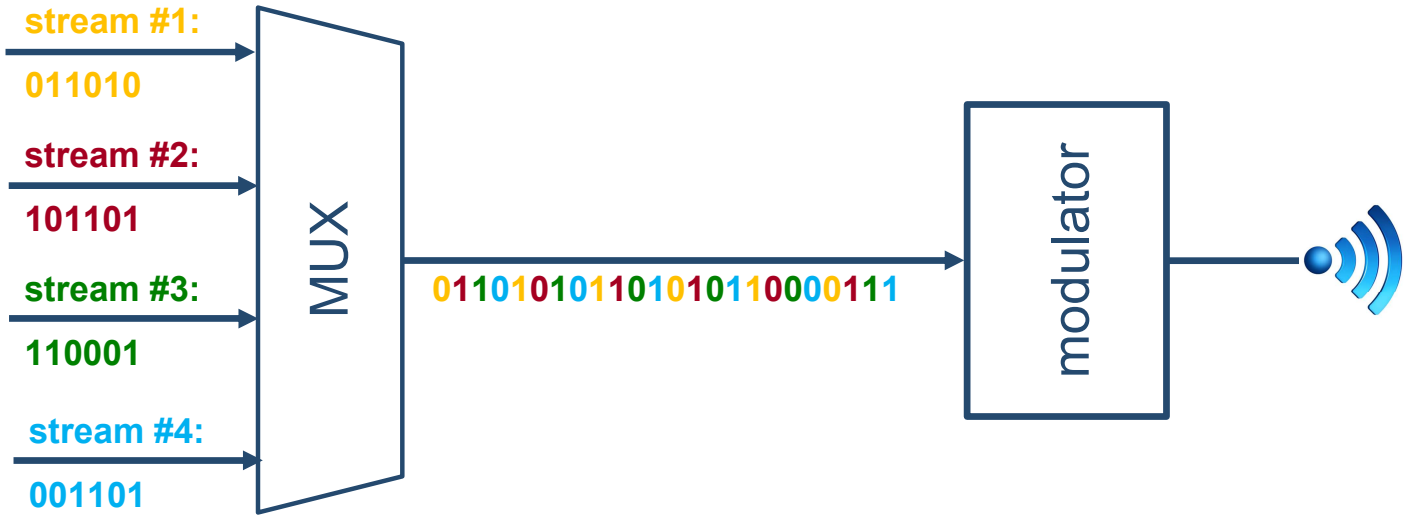
We can exploit several degrees of freedom: frequency, time, space, codes, etc.



Time division multiplexing (TDM)

Time division multiplexing (TDM) (1/2)

In practice, TDM works as follows:



$R = 1/T$ $\qquad$ $R^{(\mathrm{MUX})} = 4/T$

Time division multiplexing (TDM) (2/2)

Each stream makes use of the whole bandwidth using a round robin scheduling, with time slot duration $T^{(\mathrm{MUX})} = T/N$:

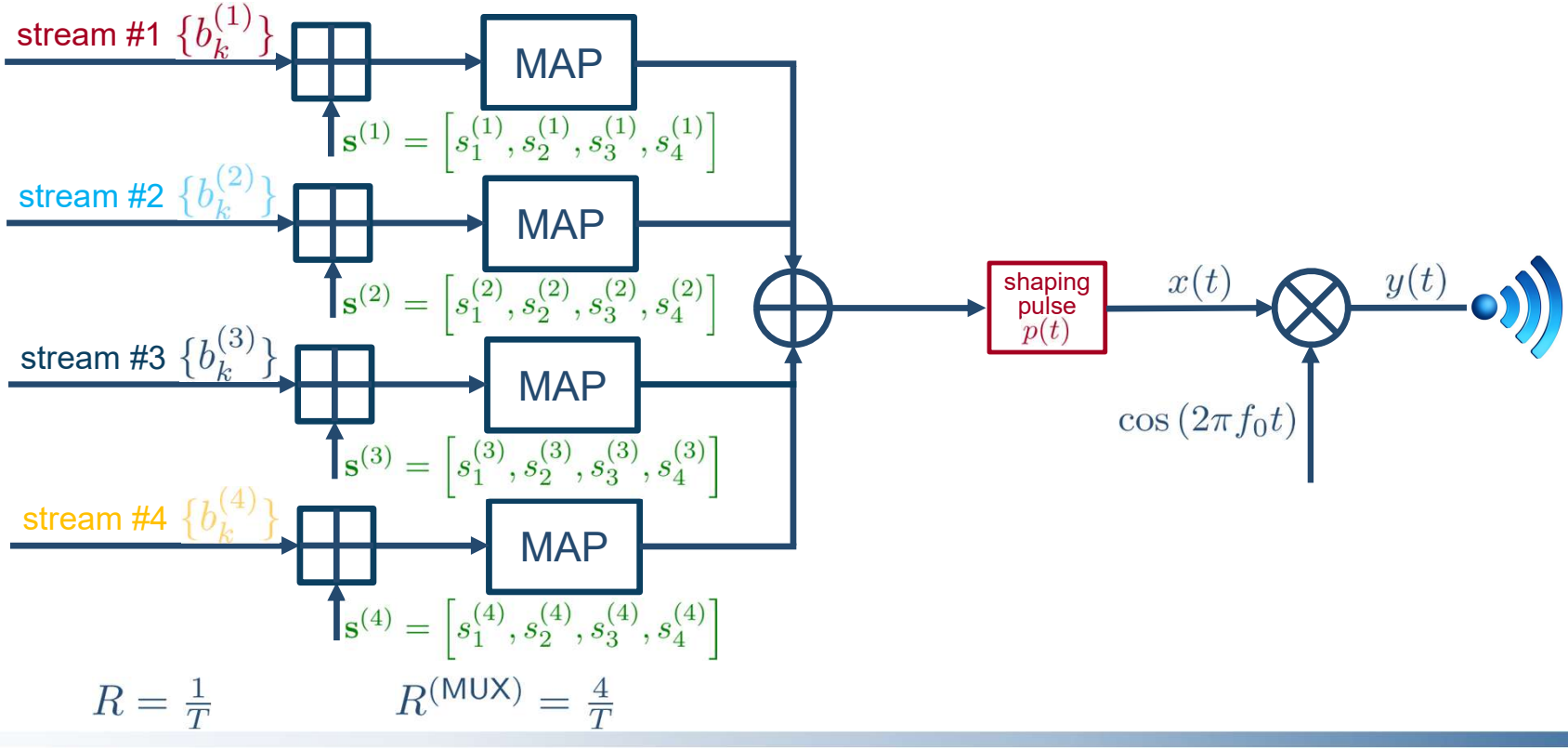
Practical applications of TDM

- E1 (Europe) & T1 (USA and Japan): byte-based TDM
- NRZ binary signals, with $R = 64\,\mathrm{kb/s}$ and $T = 15.625\,\mu\mathrm{s}$ are grouped byte by byte: $T_B = 8T = 125\,\mu\mathrm{s}$
- The E1 multiplex includes *N*=32 streams (30: data plane, 2: control plane)
- The TDM signal shows $R^{(\mathrm{MUX})} = N \cdot R = 2.048\,\mathrm{Mb/s}$ and $T^{(\mathrm{MUX})} = T/N \simeq 0.49\,\mu\mathrm{s}$ (for a re-clocked byte time $\simeq 3.9\,\mu\mathrm{s}$)

Code division multiplexing (CDM)

Code division multiplexing (CDM) (1/4)

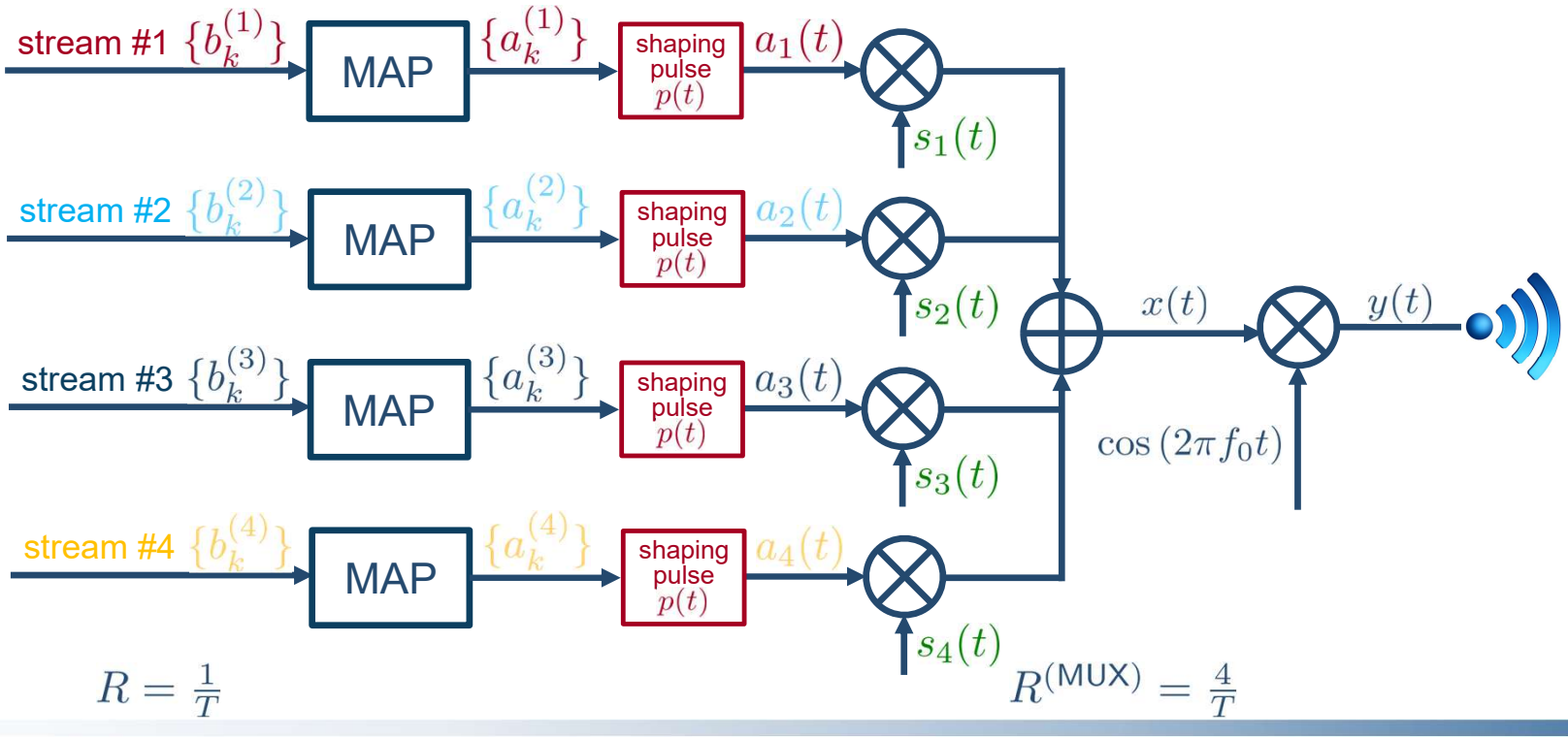
In practice, CDM works as follows:



$R = \frac{1}{T}$ $\quad R^{(\mathrm{MUX})} = \frac{4}{T}$

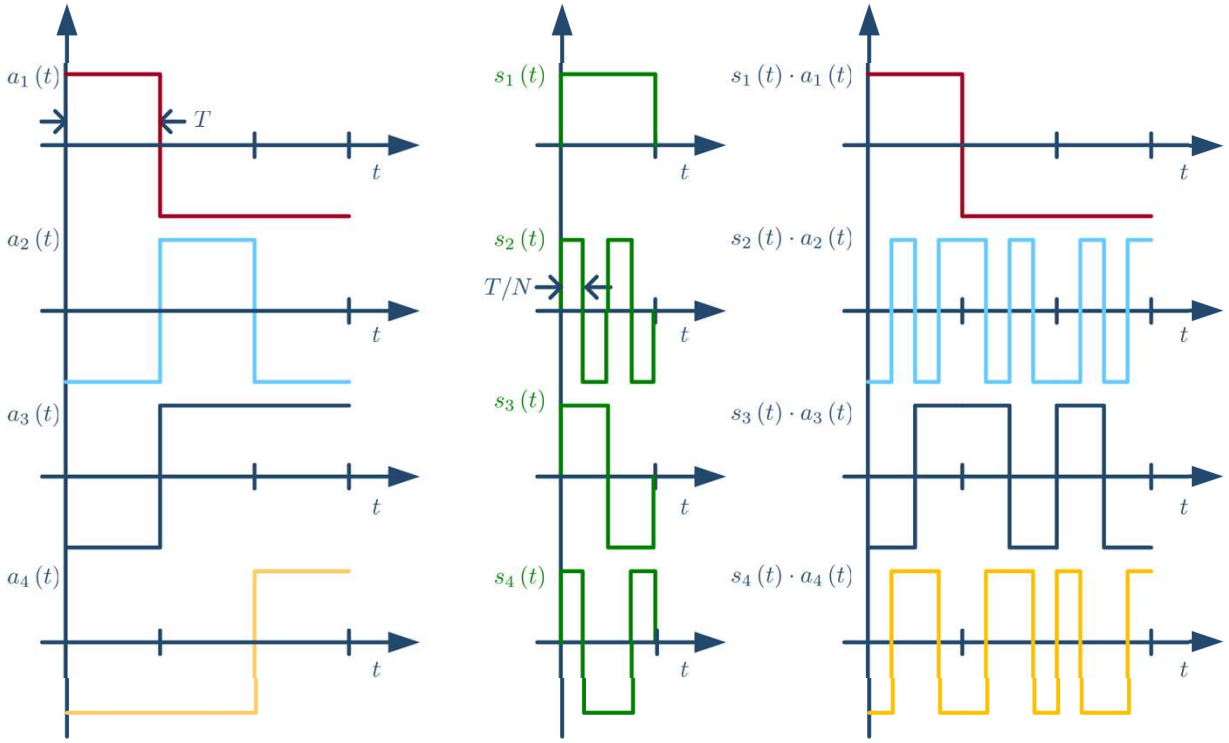
Code division multiplexing (CDM) (2/4)

To better visualize it, let us use the following equivalent scheme:



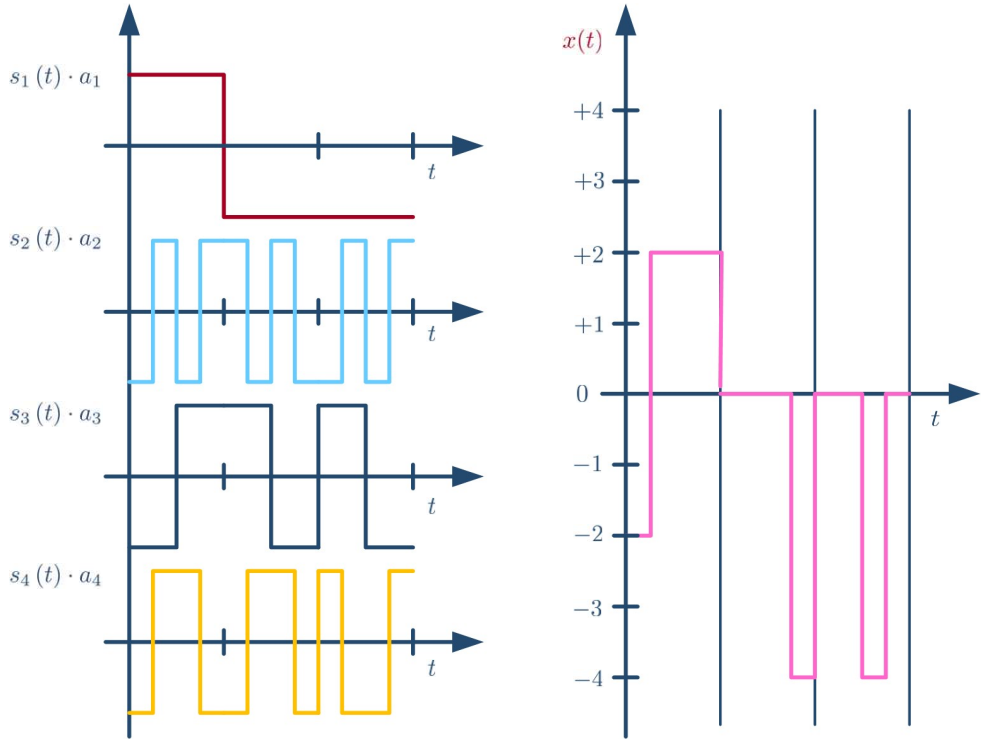
Code division multiplexing (CDM) (3/4)

In CDM, signatures are given by a code set $\{s_n(t)\}_{n=1}^{N}$:



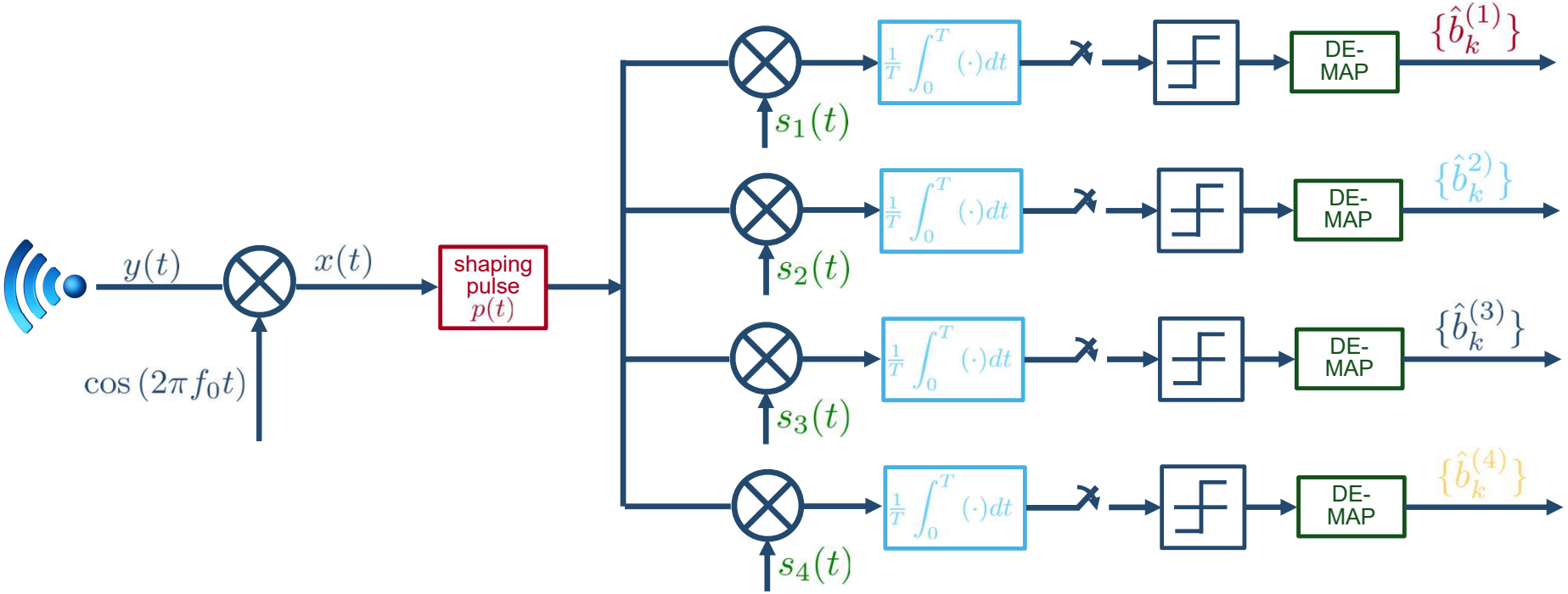
Code division multiplexing (CDM) (4/4)

The CDM-multiplexed signal is the summation of all coded streams:



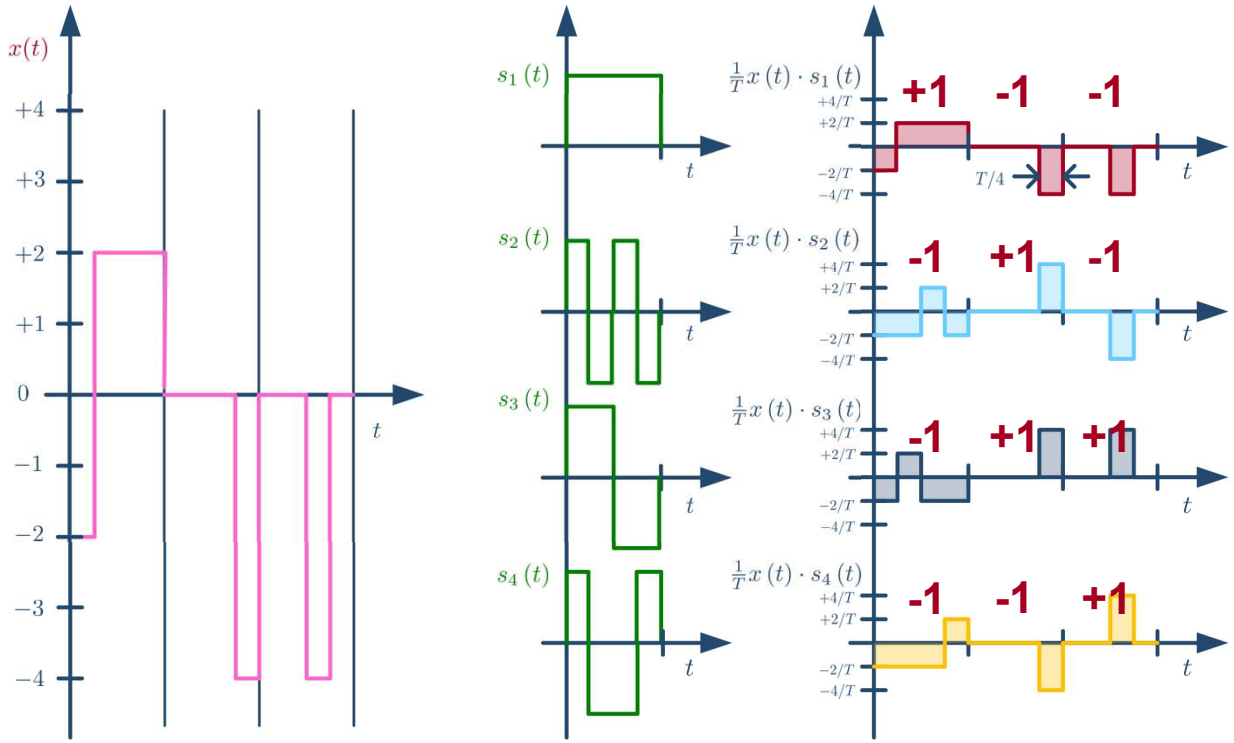
CDM demultiplexing (1/3)

At the receiver, the following operations needs to be done:



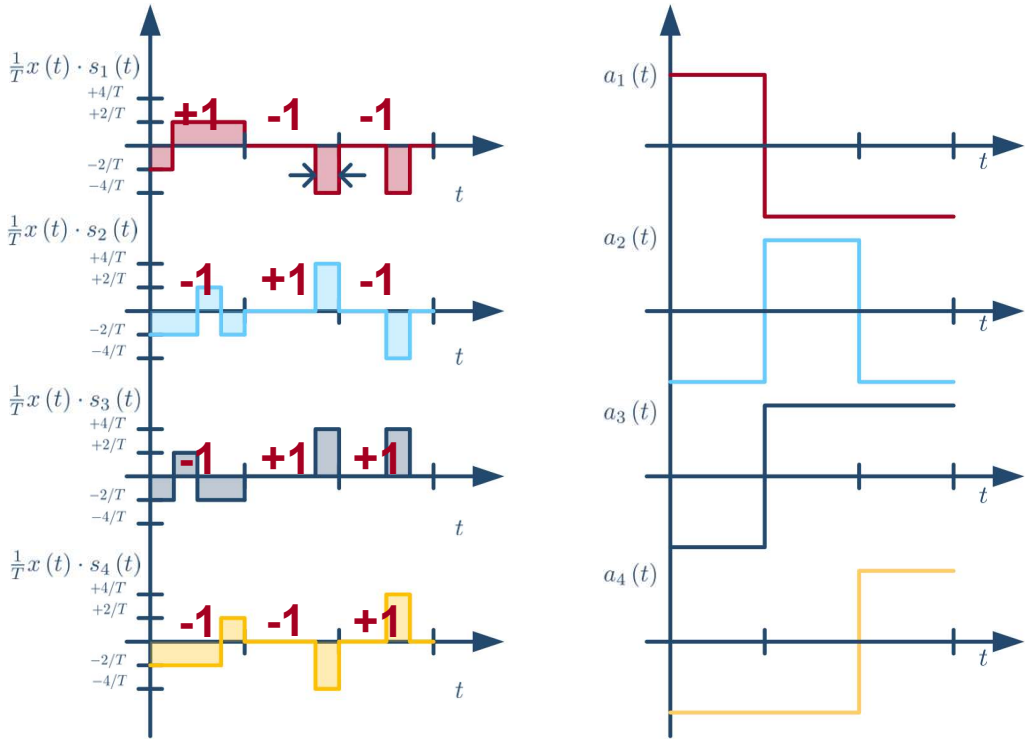
CDM demultiplexing (2/3)

To decode the *n*th stream, the receiver needs to know the code set:



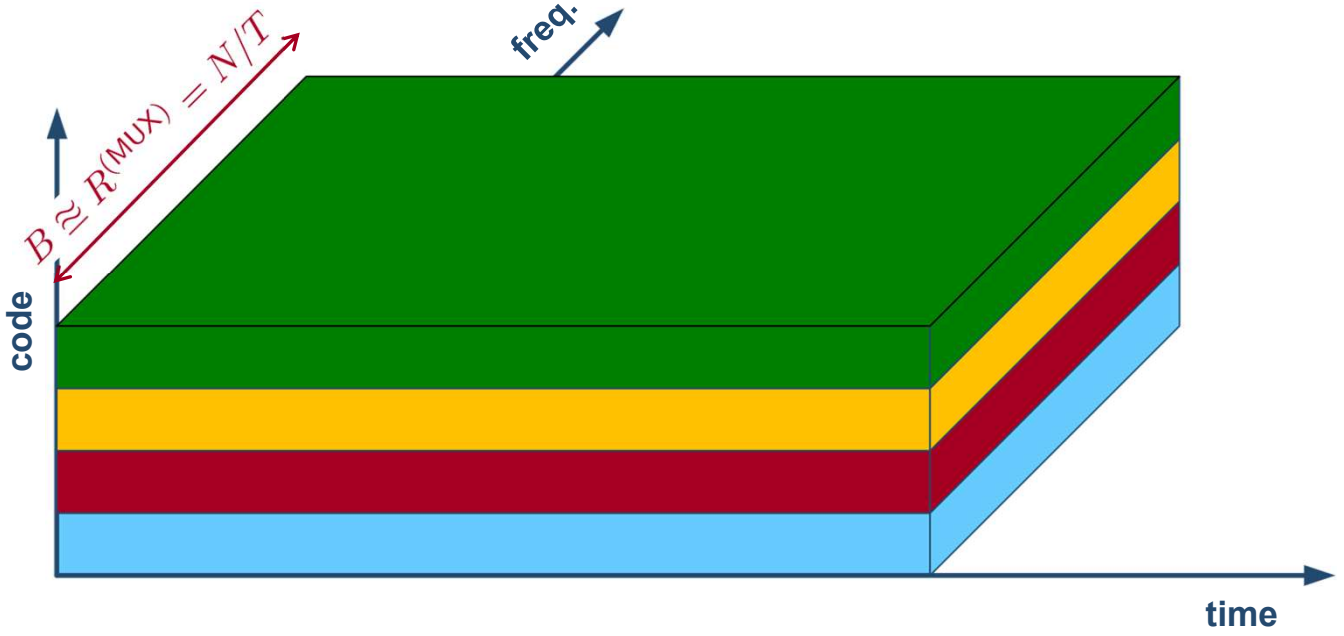
CDM demultiplexing (3/3)

To decode the *n*th stream, the receiver needs to know the code set:



CDM in 3D resource plan

This holds true for each CDM signal obtained with orthogonal codes: it makes continuous use of the whole bandwidth:



Walsh-Hadamard codes (1/4)

A useful set of signature codes for CDM is the Walsh-Hadamard (WH) code set

The nth code of the WH set is represented by the *n*th row of the *N*-order Hadamard matrix $\mathbf{H}_N$:

$$\mathbf{H}_4 = \begin{bmatrix} +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & -1 & +1 \end{bmatrix} \begin{matrix} \mathbf{s}^{(1)} \\ \mathbf{s}^{(2)} \\ \mathbf{s}^{(3)} \\ \mathbf{s}^{(4)} \end{matrix}$$

Walsh-Hadamard codes (2/4)

The *N*-order WH code set, with cardinality $N = 2^W$, with $W \in \mathbb{N}$, can be obtained in a recursive fashion:

$$\mathbf{H}_2 \triangleq \begin{bmatrix} +1 & +1 \\ +1 & -1 \end{bmatrix}, \quad \mathbf{H}_N = \mathbf{H}_2 \otimes \mathbf{H}_{N/2} = \begin{bmatrix} \mathbf{H}_{N/2} & \mathbf{H}_{N/2} \\ \mathbf{H}_{N/2} & \overline{\mathbf{H}}_{N/2} \end{bmatrix}$$

where $\otimes$ is the Kronecker product, and $\overline{\mathbf{H}}$ is the modulo-2 complement of the matrix $\mathbf{H}$

Walsh-Hadamard codes (3/4)

Examples:

$$\mathbf{H}_2 \triangleq \begin{bmatrix} +1 & +1 \\ +1 & -1 \end{bmatrix}$$

$$\mathbf{H}_4 = \begin{matrix} \mathbf{H}_2 \\ \mathbf{H}_2 \end{matrix} \begin{bmatrix} +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & -1 & +1 \end{bmatrix} \begin{matrix} \mathbf{H}_2 \\ \overline{\mathbf{H}}_2 \end{matrix}$$

$$\mathbf{H}_8 = \begin{matrix} \mathbf{H}_4 \\ \mathbf{H}_4 \end{matrix} \begin{bmatrix} +1 & +1 & +1 & +1 & +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 & +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 & +1 & +1 & -1 & -1 \\ +1 & -1 & -1 & +1 & +1 & -1 & -1 & +1 \\ +1 & +1 & +1 & +1 & -1 & -1 & -1 & -1 \\ +1 & -1 & +1 & -1 & -1 & +1 & -1 & +1 \\ +1 & +1 & -1 & -1 & -1 & -1 & +1 & +1 \\ +1 & -1 & -1 & +1 & -1 & +1 & +1 & -1 \end{bmatrix} \begin{matrix} \mathbf{H}_4 \\ \overline{\mathbf{H}}_4 \end{matrix}$$

Walsh-Hadamard codes (4/4)

Exercise: Show that this signature set $\{c_n(t)\}_{n=1}^{N}$ is orthogonal

$$\frac{1}{T}\int_0^T c_n(t)\cdot c_m^*(t)\,dt = \delta[n-m] \qquad \forall n,m$$

Thanks to this property, codes like the WH ones are called orthogonal codes

CDM vs. TDM (1/2)

Similarly to TDM, in CDM all streams need to be synchronized (hence, CDM is not suitable for analog systems)

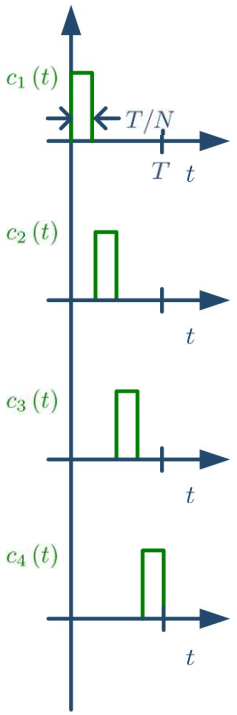
However, signatures in TDM are non-null only for a time slot, whereas codes in CDM are pseudo-random noise sequences:

TDM: $$\frac{1}{T}\int_0^T c_n(t)\,dt = \frac{1}{N}$$

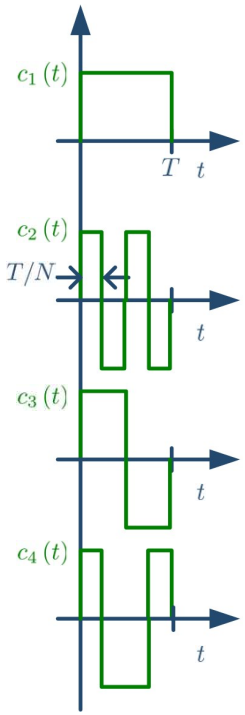
CDM: $$\frac{1}{T}\int_0^T c_n(t)\,dt = 0$$

CDM vs. TDM (2/2)

TDM: $\frac{1}{T}\int_0^T c_n(t)\,dt = \frac{1}{N}$



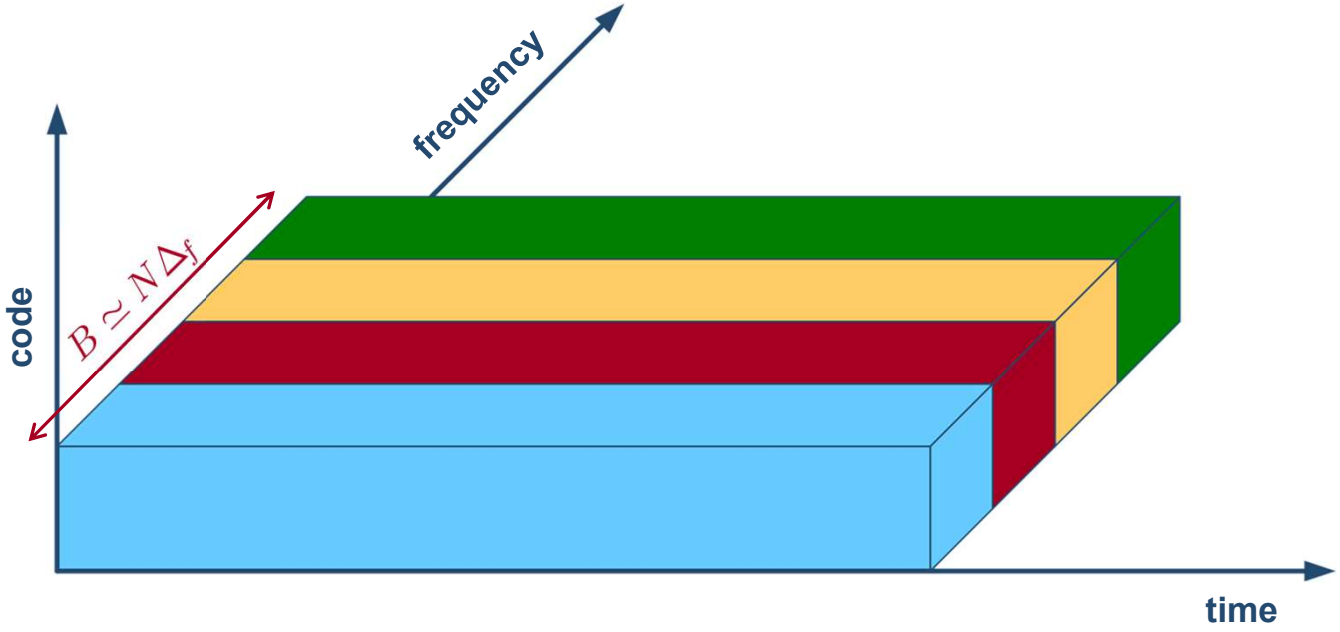
CDM: $\frac{1}{T}\int_0^T c_n(t)\,dt = 0$



Orthogonal frequency division multiplexing (OFDM)

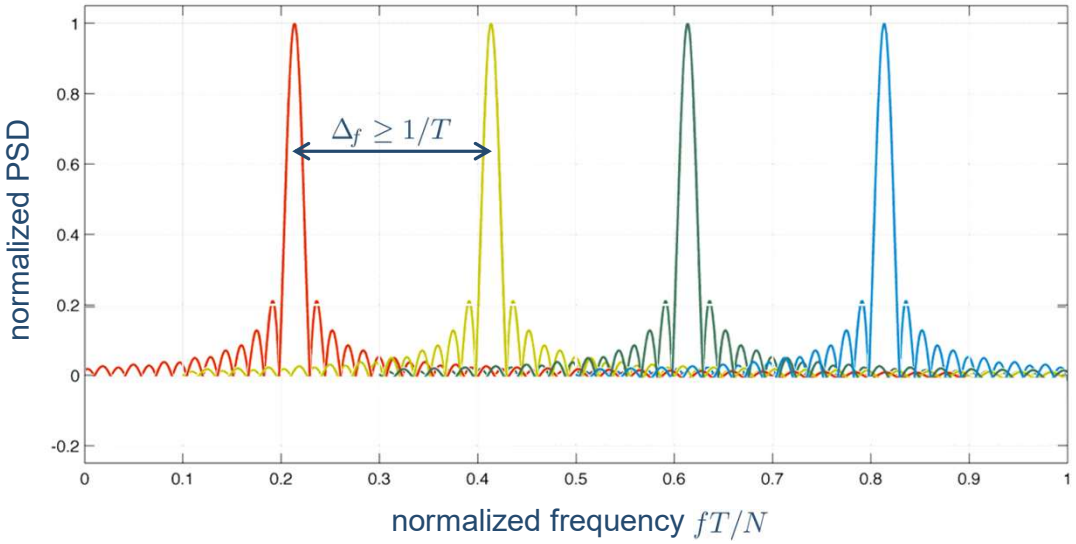
Frequency division multiplexing (FDM) (1/2)

We can also exploit the degree of freedom offered by the frequency, by separating N streams that do not overlap in the frequency domain, so as to form an FDM signal:



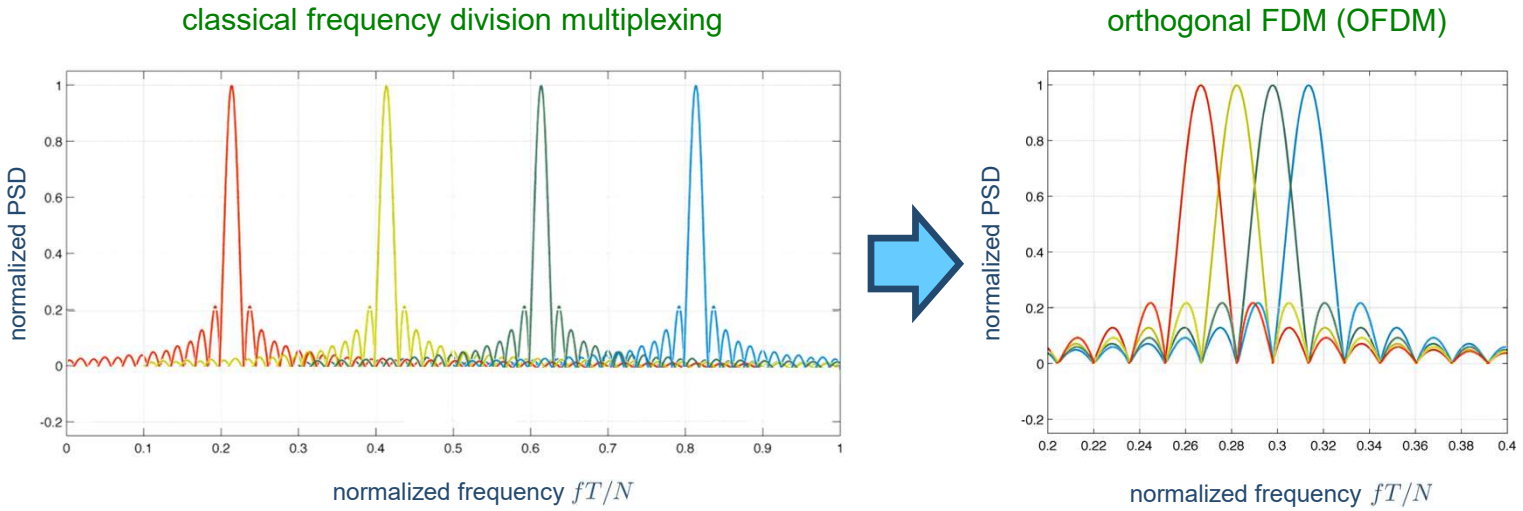
Frequency division multiplexing (FDM) (2/2)

The PSD of a frequency-division multiplexed signal is



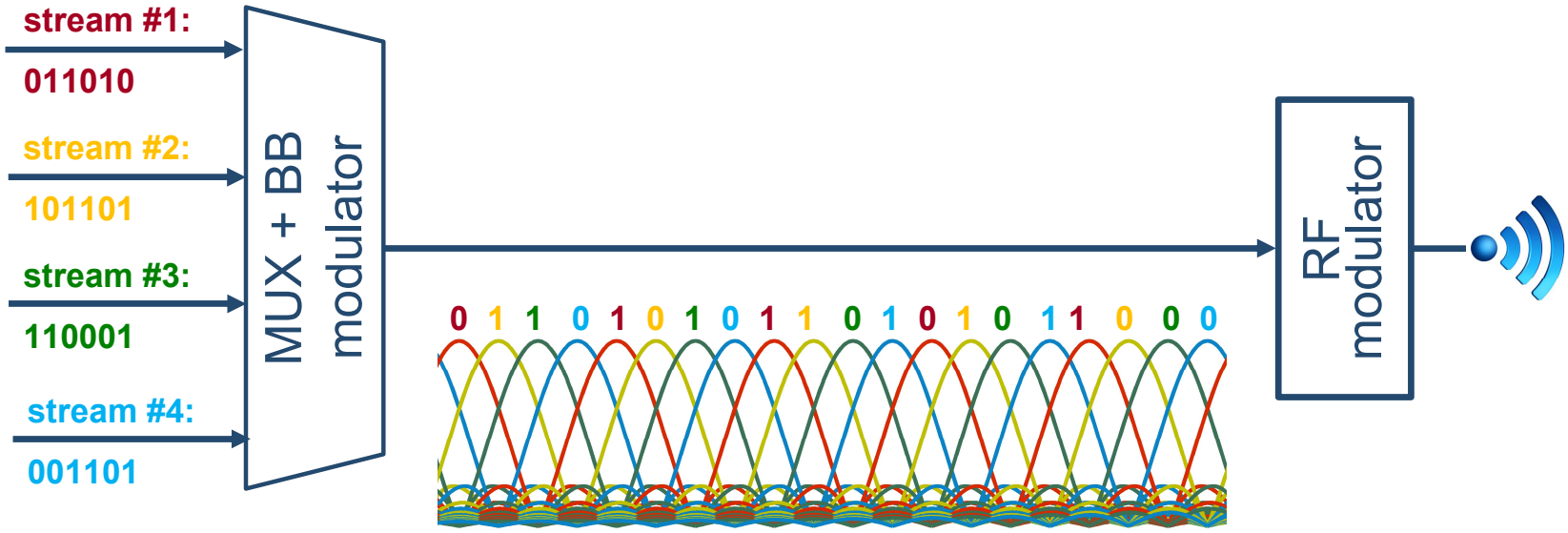
Orthogonal frequency division multiplexing (OFDM) (1/3)

The minimum bandwidth occupancy is given by $\Delta_f = 1/T$, corresponding to the first null of each stream's PSD:



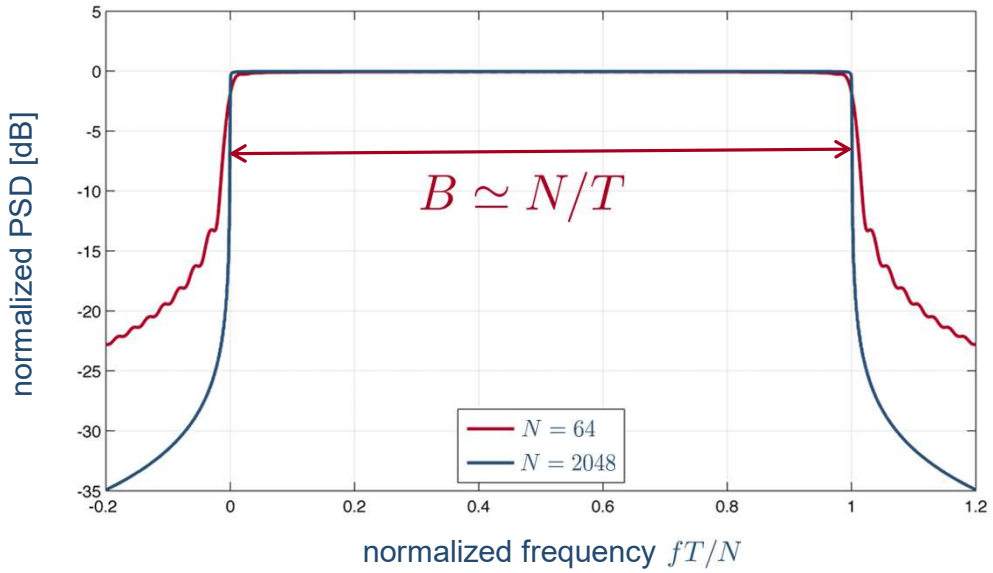
Orthogonal frequency division multiplexing (OFDM) (2/3)

OFDM works as follows:



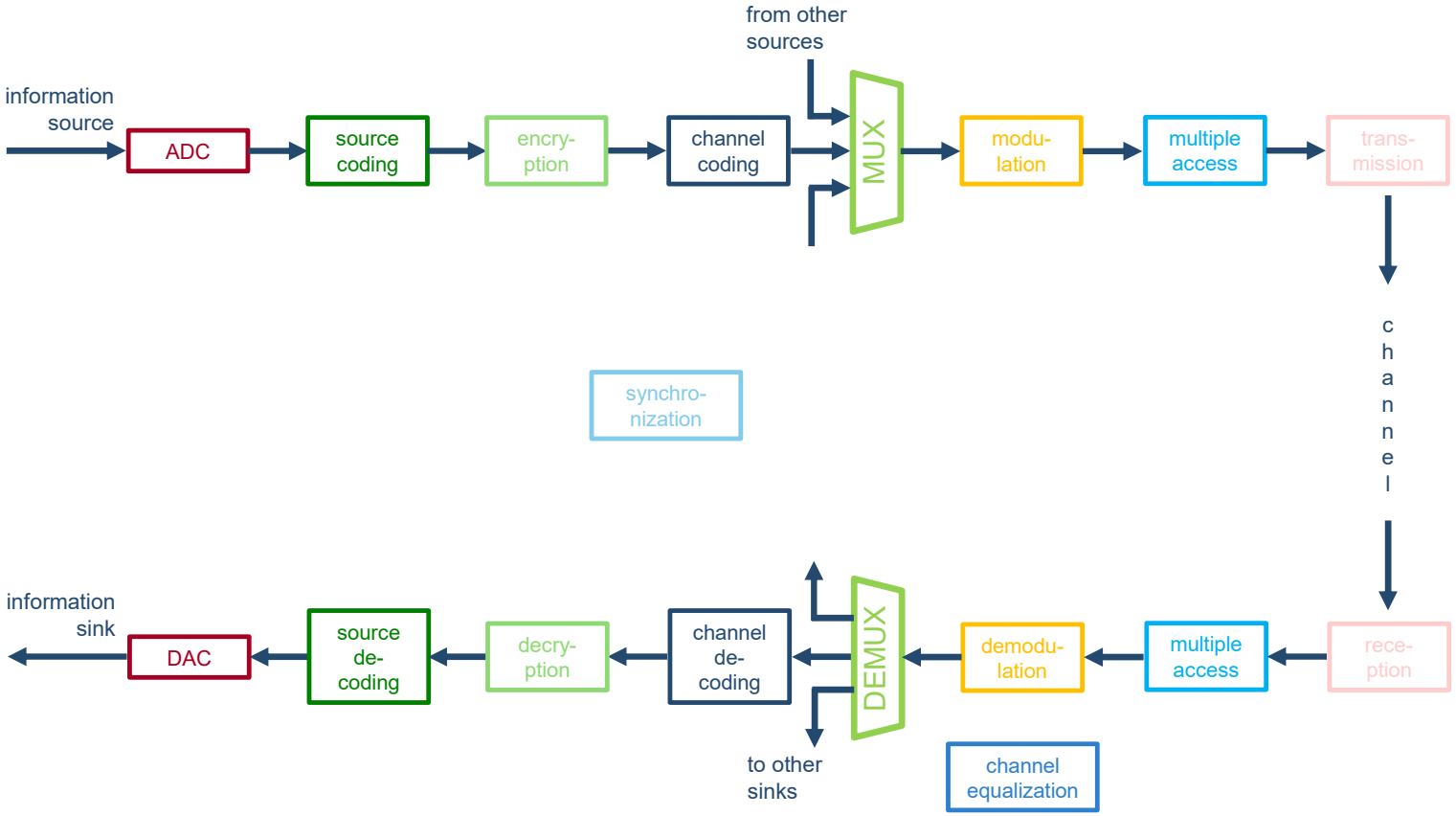
Orthogonal frequency division multiplexing (OFDM) (3/3)

The PSD of the OFDM signal is given by $\mathcal{P}_{\text{MUX}}(f) = \sum_{n=1}^{N} \mathcal{P}_n(f - n/T)$, where $\mathcal{P}_n(f) = T \cdot \text{sinc}^2(fT)$



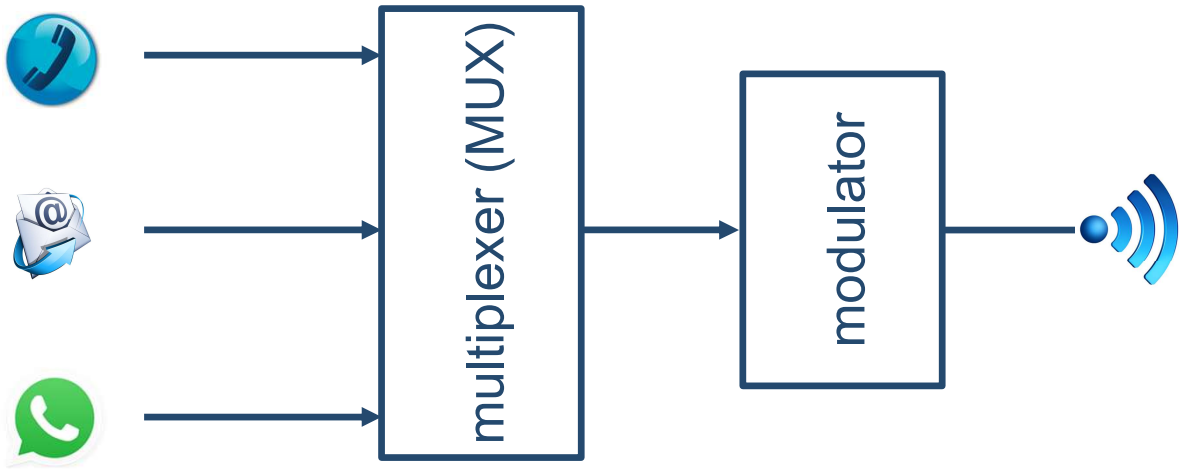
Multiple access

Elements of a digital communication system



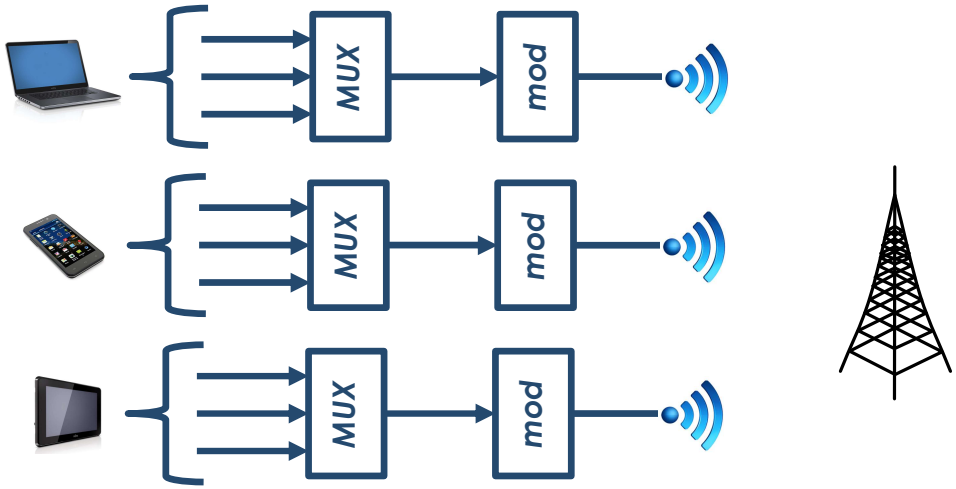
Multiplexing vs. multiple access (1/3)

Multiplexing: separating different flows at the same transmit side



Multiplexing vs. multiple access (2/3)

Multiple access: separating different users at the receiver side

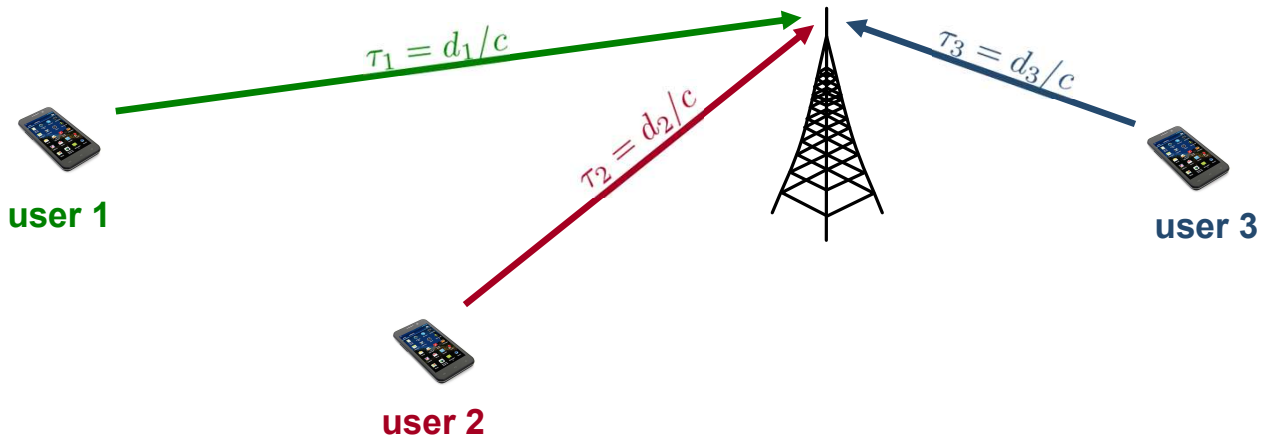


Time division multiple access (TDMA)

Time division multiple access (TDMA) (1/3)

To extend the TDM approach to multiple users, we need to account for a number of additional issues

Example: a typical cellular network



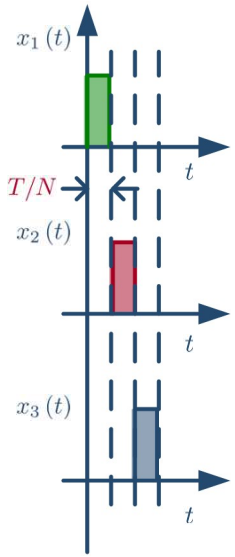
Time division multiple access (TDMA) (2/3)

- Users need to receive their own time slot allocation, assigned by the network through the control plane

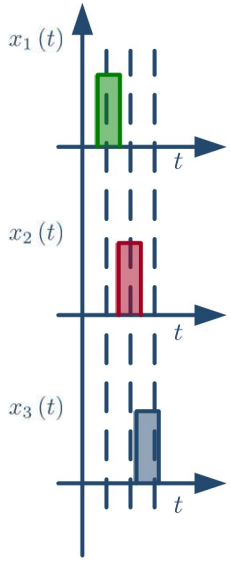
- There is the need for network synchronization to align the users, which can be geographically sparse (and hence with different propagation delays)

Time division multiple access (TDMA) (3/3)

Desired situation:



Actual situation:

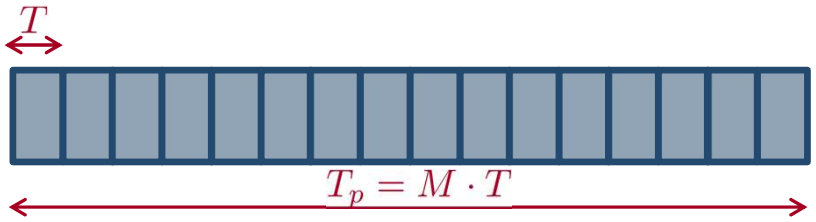


signals from different users overlap due to different propagation delays

Burst-based TDMA (1/2)

To simplify the synchronization tasks, time slots are grouped per burst instead of bit by bit

Each data burst is composed by a packet with duration $T_p = M \cdot T$, where M is the number of bits per packet

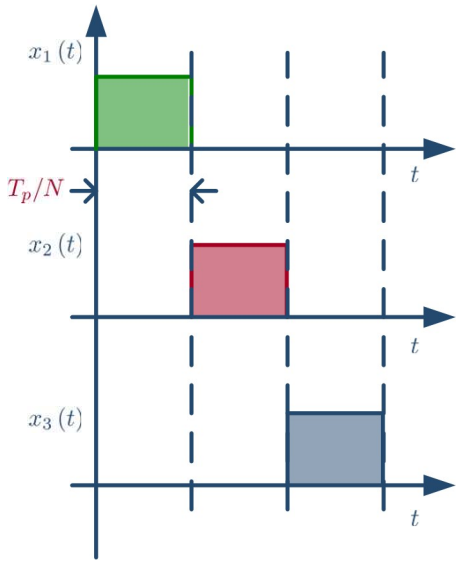


This approach allows the propagation times to be compared with T_p/N rather than with T/N (and hence there is a factor M)

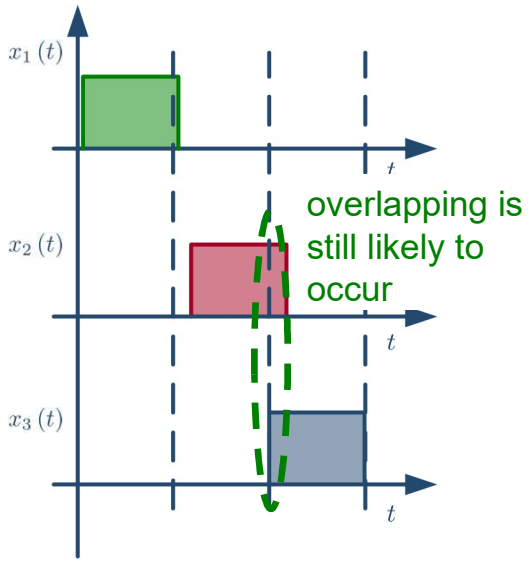
Burst-based TDMA (2/2)

The (centralized) network synchronization feeds back timing information to let the users anticipate or postpone transmissions

Desired situation:

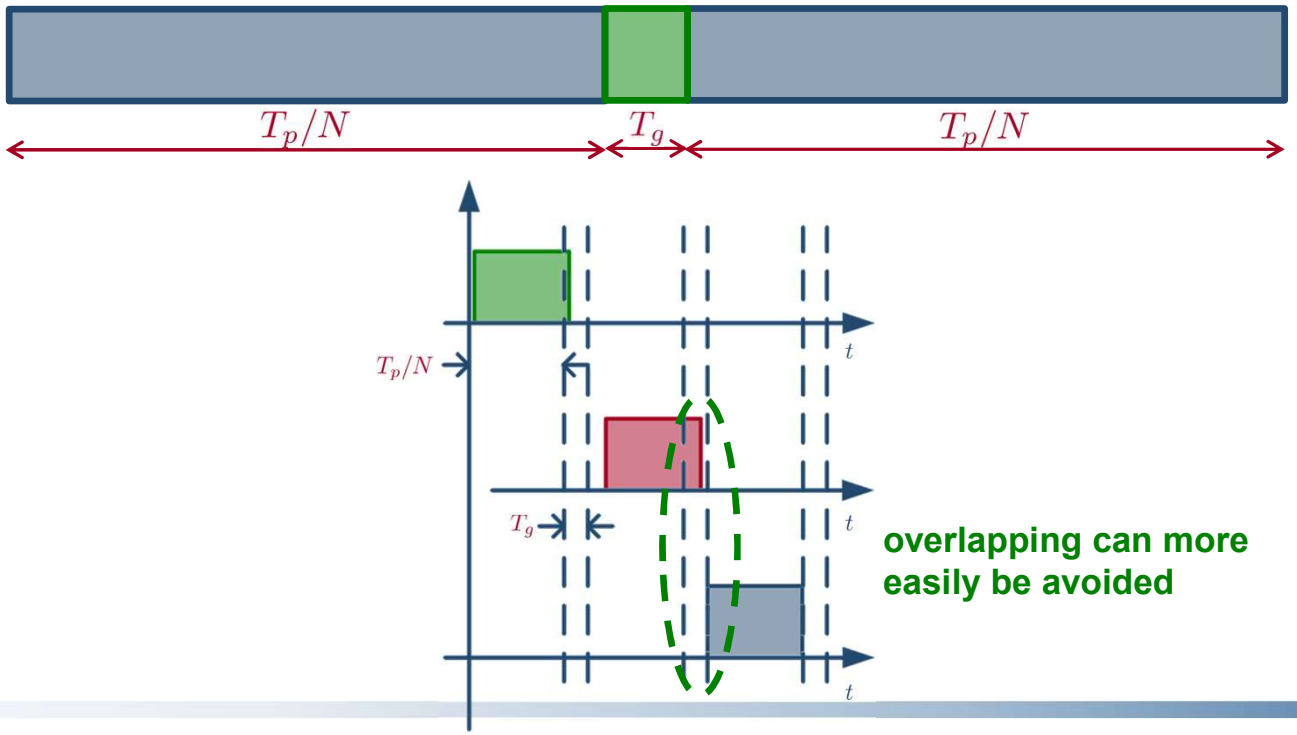


Actual situation:



Introducing a guard interval

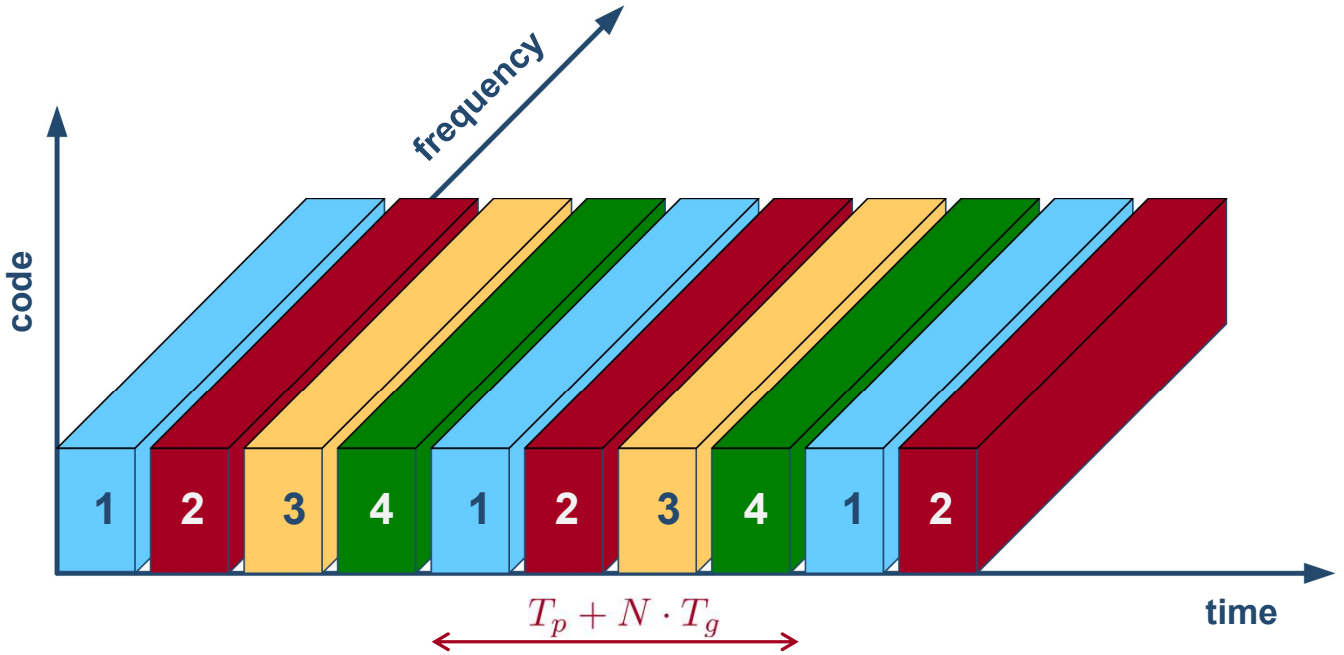
To relax the network synchronization requirements, we can introduce a guard interval T_g between each burst



overlapping can more easily be avoided

TDMA in 3D resource plan

The guard interval is a tradeoff between synchronization performance and optimal usage of network resources

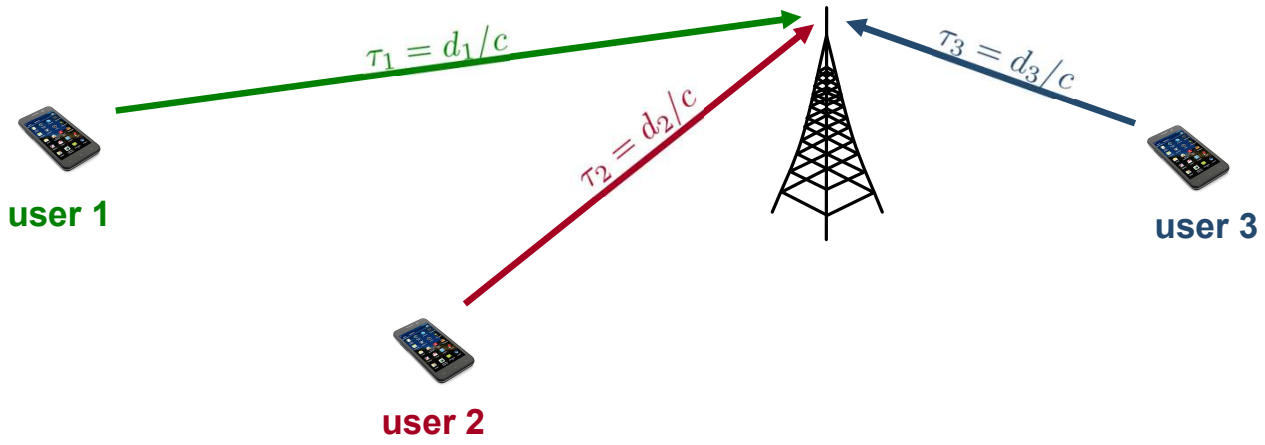


Code division multiple access (CDMA)

Code division multiple access (CDMA) (1/6)

The main difference between CDM and CDMA is due to uncoordinated times of arrival across the network users

Similarly to the TDMA scenario, different users experience different propagation delays:



Code division multiple access (CDMA) (2/6)

Let us evaluate the impact of asynchronous users on the aggregate signal measured at the receiver side:

$$x_{\text{MA}}(t) = \sum_{n=1}^{N} x_n(t-\tau_n) \cdot s_n(t-\tau_n)$$

For simplicity, let us focus on just two users and neglect the impact of AWGN and channel propagation on the output of the matched-filter receiver tailored for user #1 at sample $\ell = 0$:

$$\begin{aligned} a_1^{(0)} &= a_1(t)|_{t=\tau_1} \\ &= \frac{1}{T}\int_{\tau_1}^{T+\tau_1} x_{\text{MA}}(t) \cdot s_1^*(t-\tau_1)\,dt \end{aligned}$$

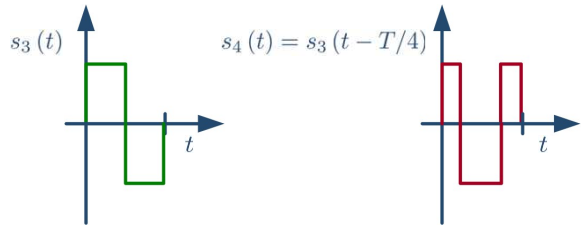
Code division multiple access (CDMA) (3/6)

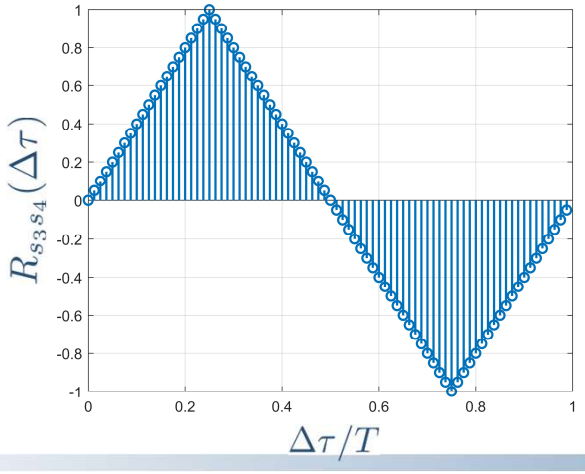
$$
\begin{aligned}
a_1^{(0)} &= b_1^{(0)} + \frac{1}{T}\int_{\tau_1}^{T+\tau_1} b_2^{(0)} \cdot s_2\left(t-\tau_2\right) \cdot s_1^*\left(t-\tau_1\right) dt \\
&= b_1^{(0)} + b_2^{(0)} \cdot \frac{1}{T}\int_0^{T} s_2\left(\alpha-\Delta\tau\right) \cdot s_1^*\left(\alpha\right) d\alpha \\
&= b_1^{(0)} + R_{s_1 s_2}\left(\Delta\tau\right) \cdot b_2^{(0)}
\end{aligned}
$$

where $\Delta\tau \triangleq \tau_2 - \tau_1$ is the interarrival time between user #2 and user #1, and $R_{s_1 s_2}\left(\Delta\tau\right)$ is the cross-correlation function between codes $s_1(t)$ and $s_2(t)$, computed at time $\Delta\tau$

Code division multiple access (CDMA) (4/6)

When considering orthogonal codes (such as the WH code set), cross-correlation is negligible if and only if $\Delta\tau \ll T/N = T_c$





Code division multiple access (CDMA) (5/6)

This means that, in practice, the requirement is $\Delta\tau \leq T_c/10$

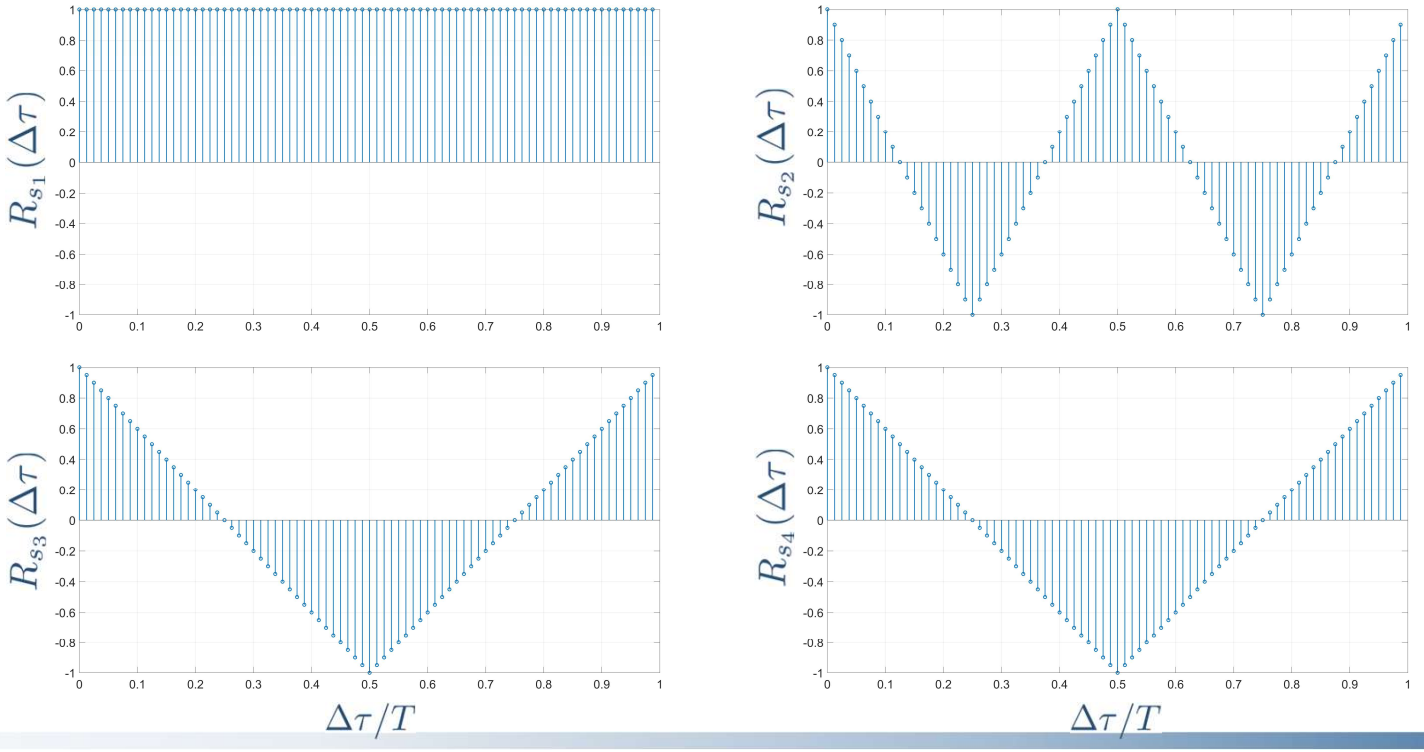
Example: in UMTS (3G), $T_c \approxeq 0.26\,\mu\text{s}$, which implies $\Delta\tau \leq 26\,\text{ns}$

This accuracy is not viable, especially if compared with that required by TDMA, in which $\Delta\tau \leq T_g \approx T$ (typically, accuracy is two-three orders of magnitude higher in CDMA)

Code division multiple access (CDMA) (6/6)

Moreover, the estimation of τ_n relies on the properties of the autocorrelation

$$R_{s_n}(\Delta\tau) \triangleq \frac{1}{T}\int_0^T s_n(\alpha-\Delta\tau)\cdot s_n^*(\alpha)\,d\alpha$$



Non-orthogonal codes (1/4)

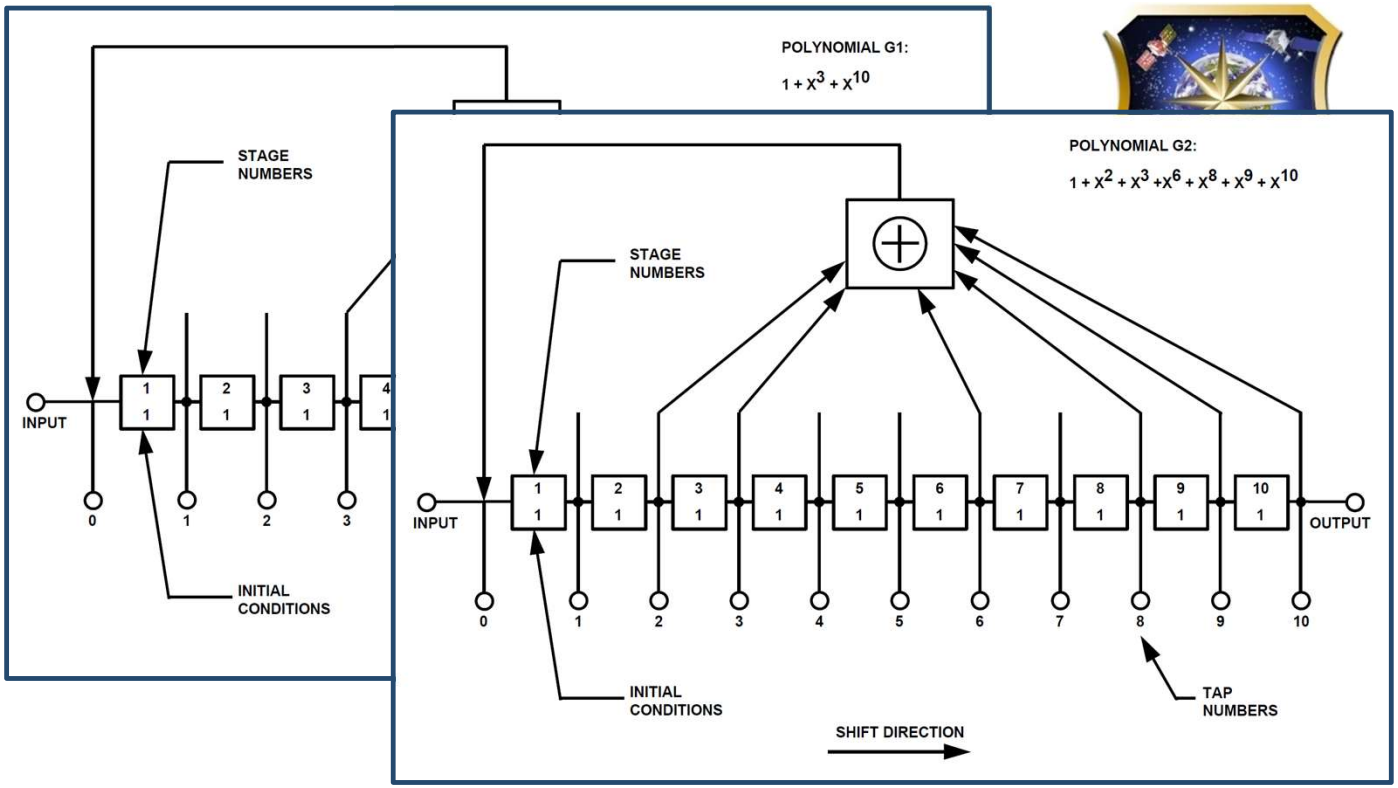
In this context, we need to identify signature codes that work pretty well in the asynchronous scenario

The class of pseudo-random noise (PRN) sequences, such as the Gold codes used in many systems, provide good performance in fully uncoordinated transmission patterns

PRN codes show similarities with the AWGN (e.g., the PSD is almost flat in the frequency domain), due to the pseudo-random nature of the chips; however, they are generated according to a well-defined algorithm, known to both transmit and receive sides

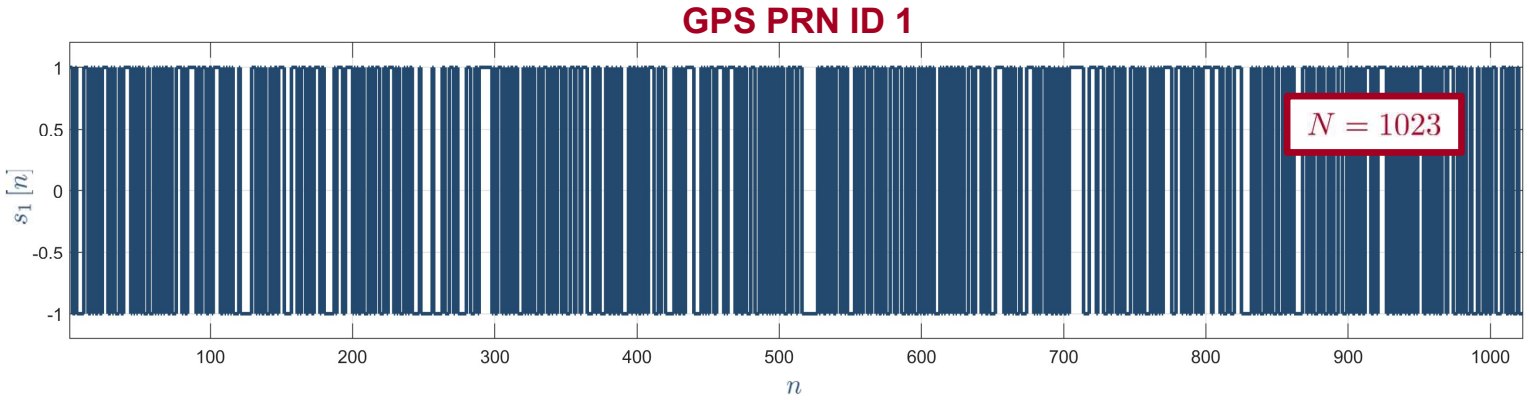
Non-orthogonal codes (2/4)

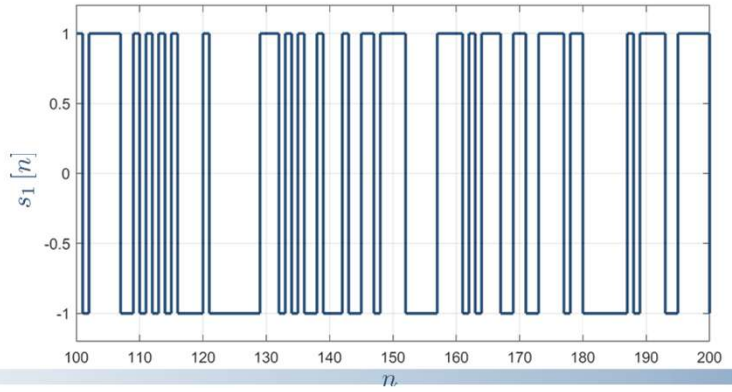
Example: PRN code generation in global positioning system (GPS)



Non-orthogonal codes (3/4)

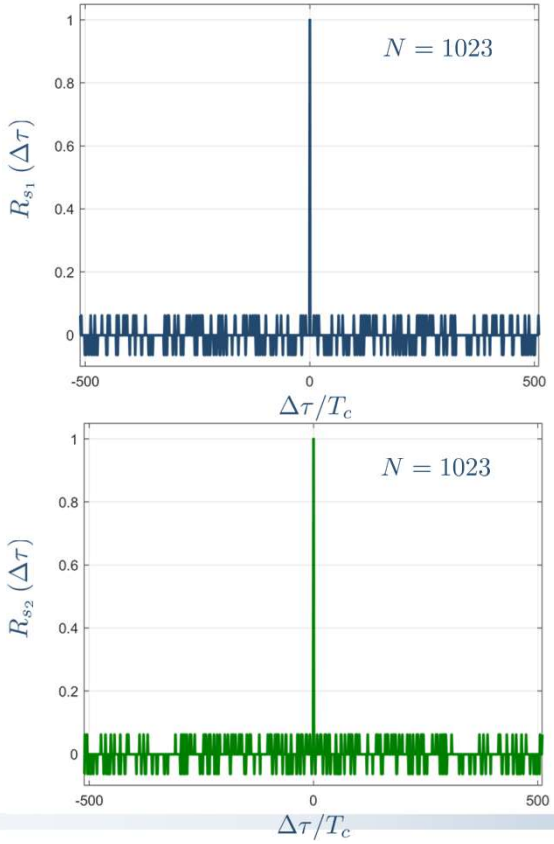
Example: PRN code generation in global positioning system (GPS)

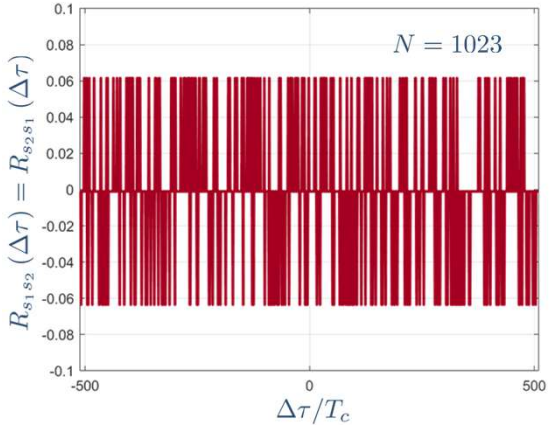




Non-orthogonal codes (4/4)

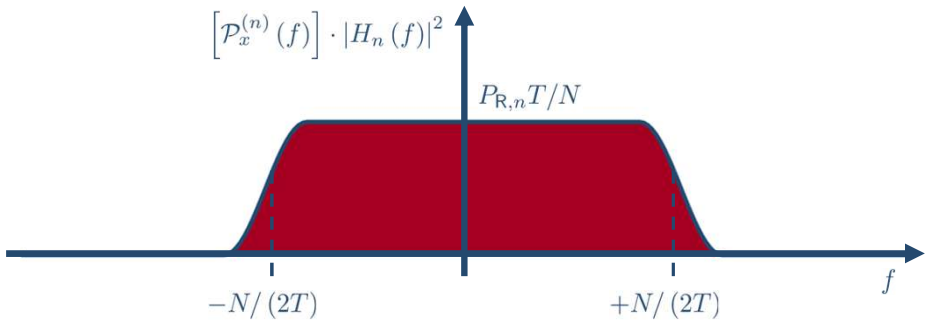
Example: PRN code generation in global positioning system (GPS)





Performance of CDMA (1/5)

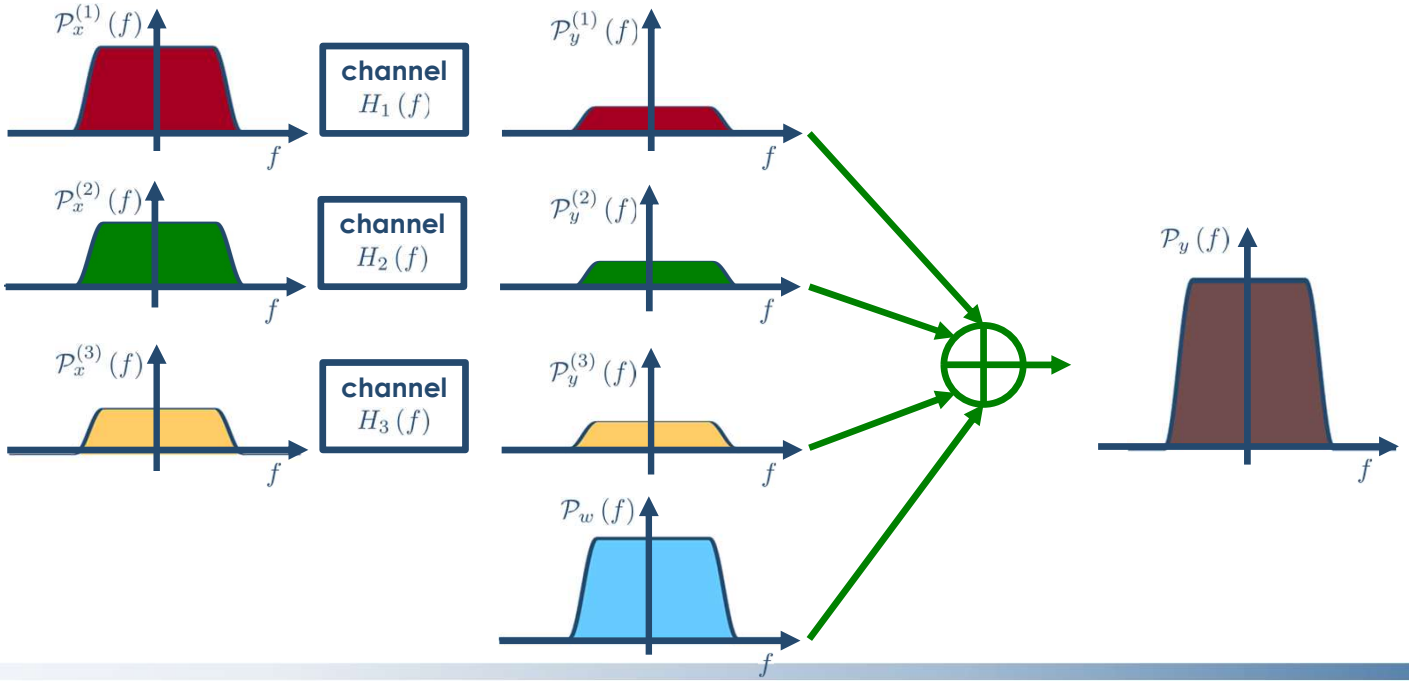
Thanks to the ad-hoc code chip generation, PRN codes show a flat PSD:



where $H_n(f)$ is the DFT of the channel experienced by user n, and $P_{\mathrm{R},n}$ is user n's received power

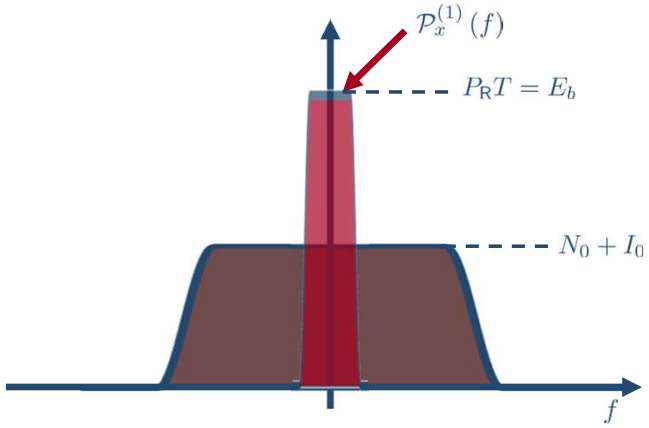
Performance of CDMA (2/5)

Hypothesis: our CDMA system adopts a closed-loop power contol that equalizes all received powers $\{P_{\mathrm{R},n}\}_{n=1}^{N_A} = P_{\mathrm{R}}$, where N_A is the number of users simultaneously active (note: $N_A \lessgtr N$)



Performance of CDMA (3/5)

Decoding the *n*th user by applying the matched-filter approach is equivalent to extract only the desired spectrum out of the compound signal:



where the interference to the other users (the so-called multiple access interference, MAI) can be considered as a white Gaussian process as soon as $N_A \gg 1$ (say, $N_A \geq 10$)

Performance of CDMA (4/5)

$$a_1^{(\ell)} = b_1^{(\ell)} + \sum_{n=2}^{N_A} b_n^{(\ell)} \cdot R_{s_1 s_n}(\Delta\tau_n) + z_1^{(\ell)}$$

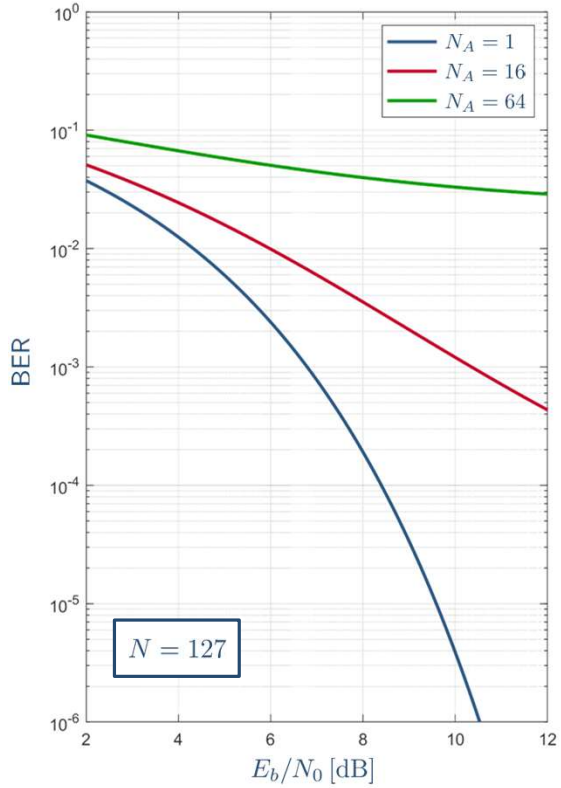
multiple access interference (MAI)

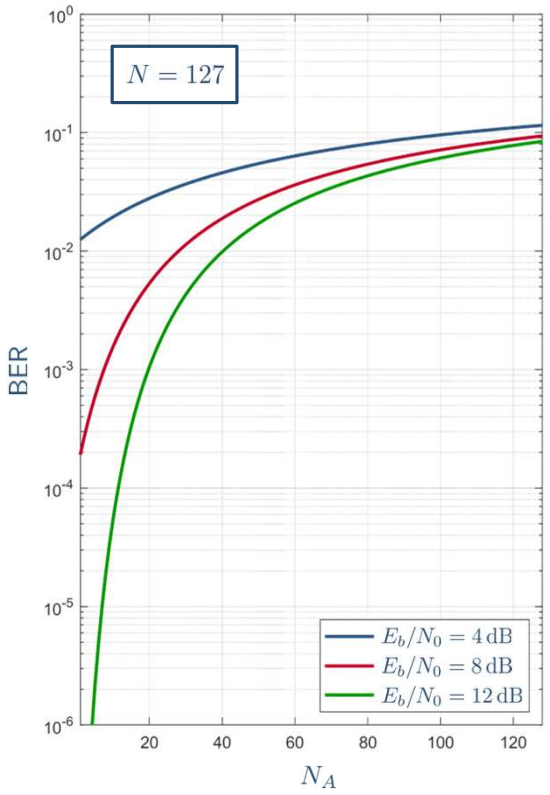
When considering a BPSK or a QPSK constellation,

$$\text{BER} = Q\left(\sqrt{\frac{2E_b}{N_0 + I_0}}\right) = Q\left(\sqrt{\frac{2E_b}{N_0 + (N_A - 1)\cdot \frac{P_R}{N/T}}}\right)$$

$$= Q\left(\sqrt{\frac{2E_b}{N_0}} \cdot \sqrt{\frac{1}{1 + \frac{N_A - 1}{N} \cdot \frac{E_b}{N_0}}}\right)$$

Performance of CDMA (5/5)





Differences wrt TDMA and FDMA

Non-orthogonality of the signature codes significantly degrades the performance of the network

However, CDMA allows users to access the network in a truly random manner, without the need for synchronization across the users

This places CDMA closer to statistical-based multiple-access techniques, such as packet-oriented ones

Packet-oriented multiple access

What is packet-oriented multiple access? (1/2)

Connection-oriented multiple access:

- coordinated access, such as in TDMA, FDMA, and, to some extent, CDMA
- deterministic approach, well suited for constant-bitrate communications

Packet-oriented multiple access:

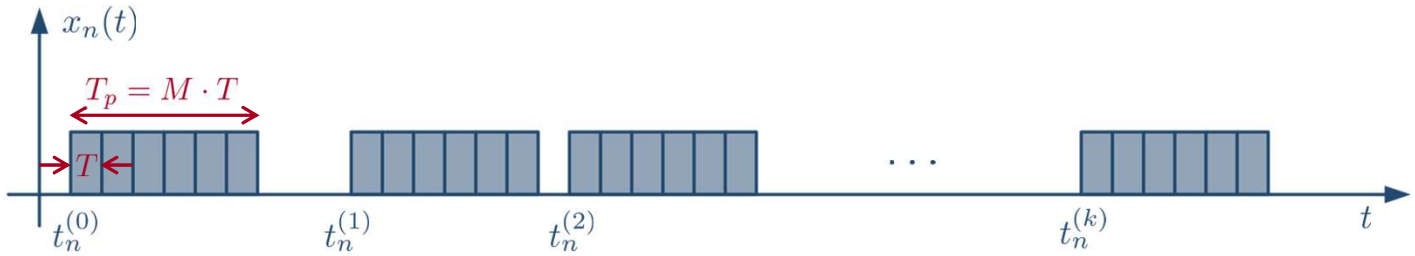
- uncoordinated access
- statistical approach, well suited for bursty-traffic communications

What is packet-oriented multiple access? (2/2)

Typical example of bursty/impulsive traffic: the IoT scenario

Packet-based MA: System model (1/4)

Let us suppose to have N active users, sending packets with constant duration T_p at random instants $\left\{t_n^{(k)}\right\}_{k=0}^{+\infty}$, using a fixed packet rate μ (measured in packets/second)



average bitrate $\overline{R}_b = M \cdot \mu\,[\mathrm{b/s}]$

Packet-based MA: System model (2/4)

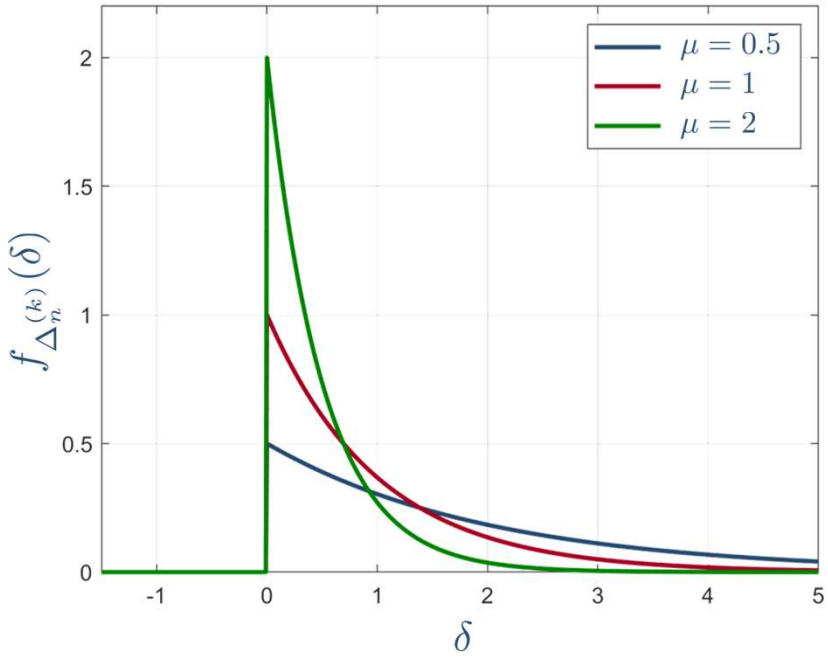
To properly investigate this packet-based scheme, we need to characterize the interarrival times

$$\Delta_n^{(k)} \triangleq t_n^{(k)} - t_n^{(k-1)}$$

A customary assumption in practical bursty systems is that interarrival times are independent and exponentially distributed

$$f_{\Delta_n^{(k)}}(\delta) = \mu e^{-\mu\delta} \cdot \mathrm{u}(\delta)$$

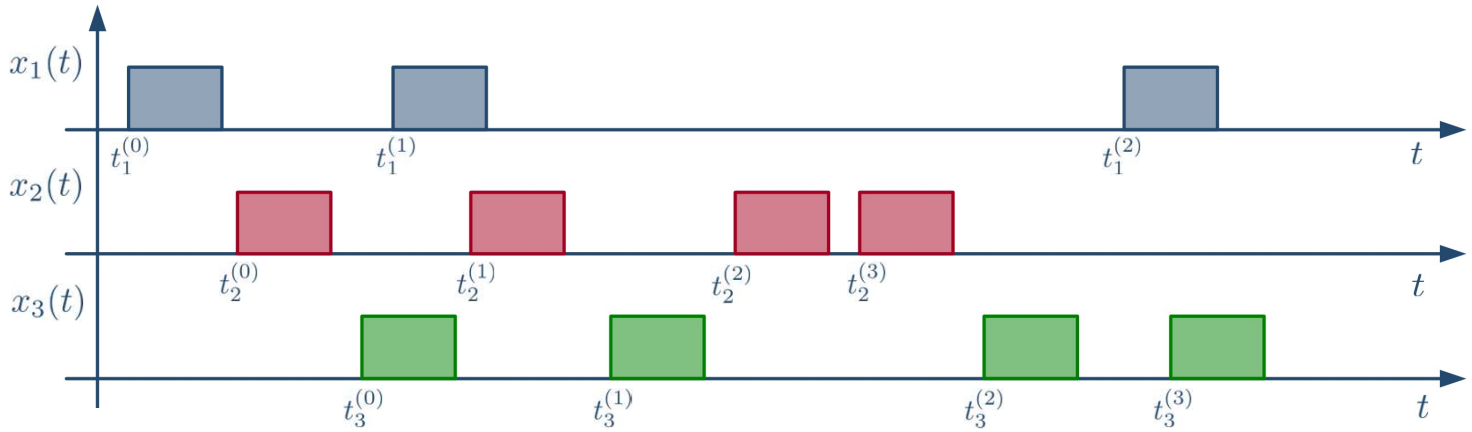
Packet-based MA: System model (3/4)



$\left\{\Delta_n^{(k)}\right\}_{k=0}^{+\infty}$ is a Poisson process with intensity μ

Packet-based MA: System model (4/4)

How can we exploit this to model the *N* users?



The aggregate system can be seen as a unique packet source, with packet rate $N\mu$ and average interarrival time $1/(N\mu)$

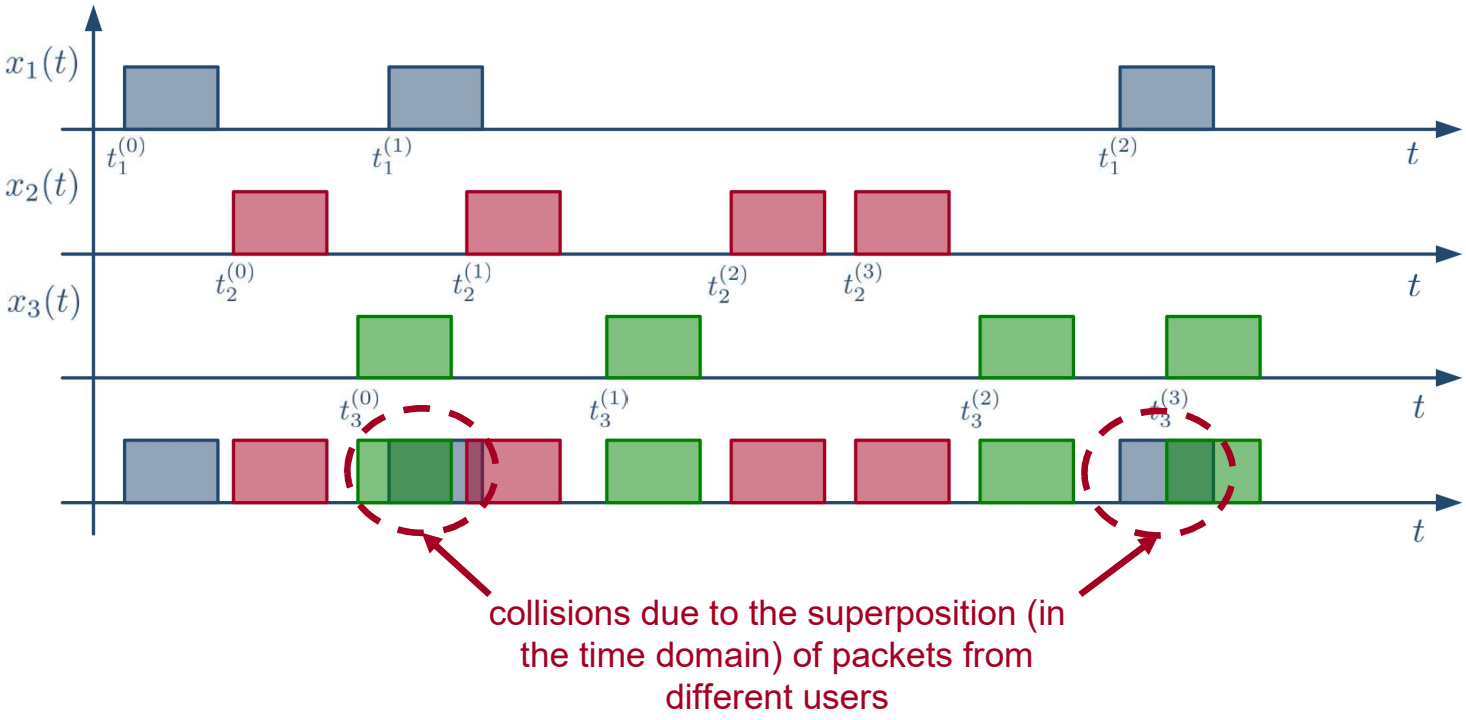
The ALOHA protocol and its variants

The ALOHA protocol (1/3)

Historically, the first protocol to address this issue was the ALOHAnet, or simply ALOHA, devised in the early 1970s

The main goal was to connect the research centers at the University of Hawaii, spread across the Hawaiian islands, with the main Oahu campus

The ALOHA protocol (2/3)



The ALOHA protocol (3/3)

Unlike connection-oriented multiple-access scheme, the ALOHA protocol needs a return channel, to receive feedback from the receiver (using acknowledgment of correct reception)

In case of a missing acknowledgement (corresponding to an event of collision), the re-transmission follows a specific backoff time (based on a statistical approach – out of scope in this lecture)

Performance of ALOHA (1/5)

Intuitively, if the packet rate μ is too high, there will be a large number of collisions (and hence of re-transmissions), which limit the efficiency of the system

To measure this tradeoff, let us compute the efficiency as

$$\rho = \frac{\text{bitrate successfully delivered}}{\text{used bandwidth}}$$

Performance of ALOHA (2/5)

Used bandwidth: $B = \frac{1}{T}$

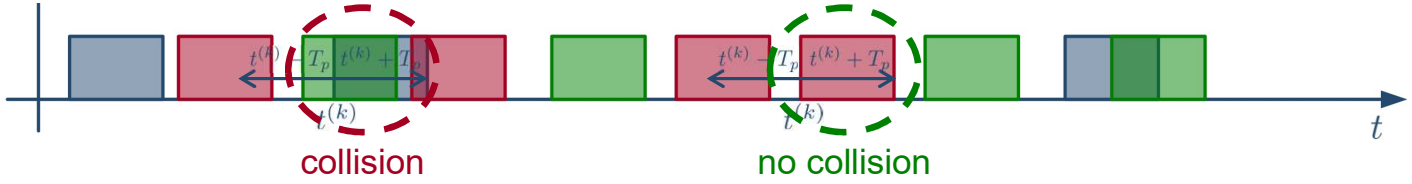
Bitrate successfully delivered: nominal bitrate $1/T$ times the average number of packets successfully received without a collision

The latter depends on the average number of transmitted packets, and the probability of no collision across packets

Average number of transmitted packets: $N\mu \cdot T_p \triangleq G$

Performance of ALOHA (3/5)

Collisions can be avoided if $t^{(k-1)} + T_p \leq t^{(k)} \leq t^{(k+1)} - T_p$



$$
\begin{aligned}
P_s &= \Pr\{\text{no collisions across packets}\} \\
&= \Pr\left\{t^{(k-1)} - T_p \leq t^{(k)} \leq t^{(k+1)} + T_p\right\} \\
&= \Pr\{\Delta \geq 2T_p\}
\end{aligned}
$$

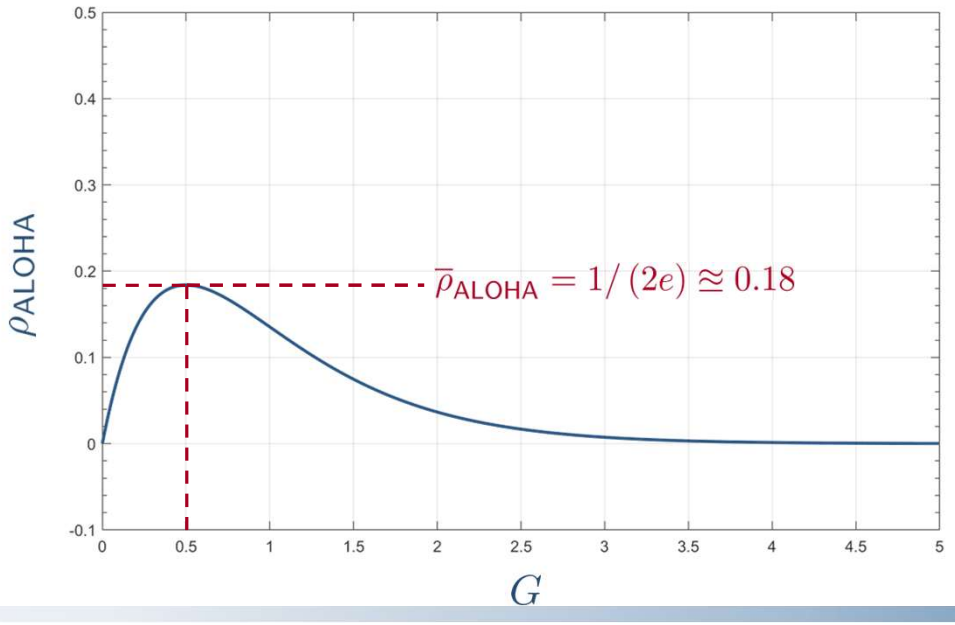
Performance of ALOHA (4/5)

Exercise: Show that $P_s = e^{-2G}$

Performance of ALOHA (5/5)

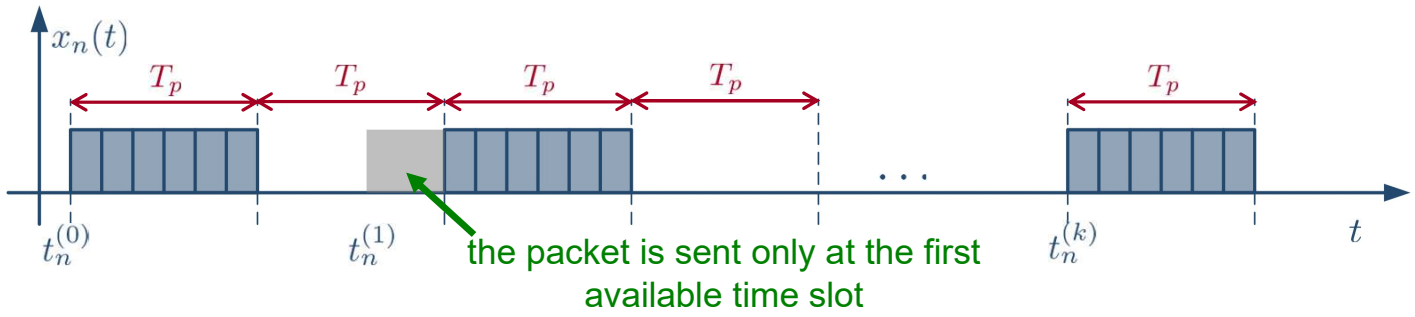
By collecting all pieces together, we get:

$$\rho_{\text{ALOHA}} = \frac{1/T \cdot G \cdot P_{\text{s}}}{1/T} = G \cdot e^{-2G}$$



The slotted-ALOHA (S-ALOHA) protocol (1/3)

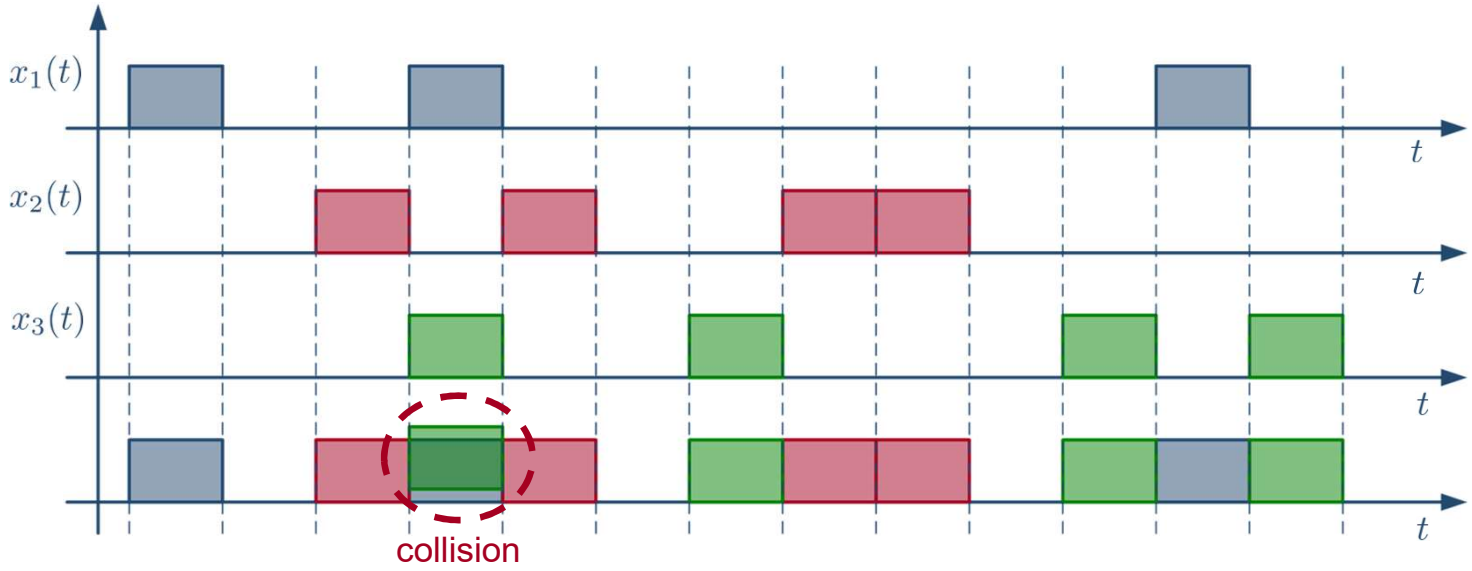
To reduce the occurrence of collision events, we can apply a centralized network synchronization similar to TDMA:



This modified version of ALOHA is called slotted ALOHA (S-ALOHA)

The slotted-ALOHA (S-ALOHA) protocol (2/3)

Collisions occur only when more than one source selects the same time slot to transmit the packet (i.e., partial collisions do not exist anymore)



The slotted-ALOHA (S-ALOHA) protocol (3/3)

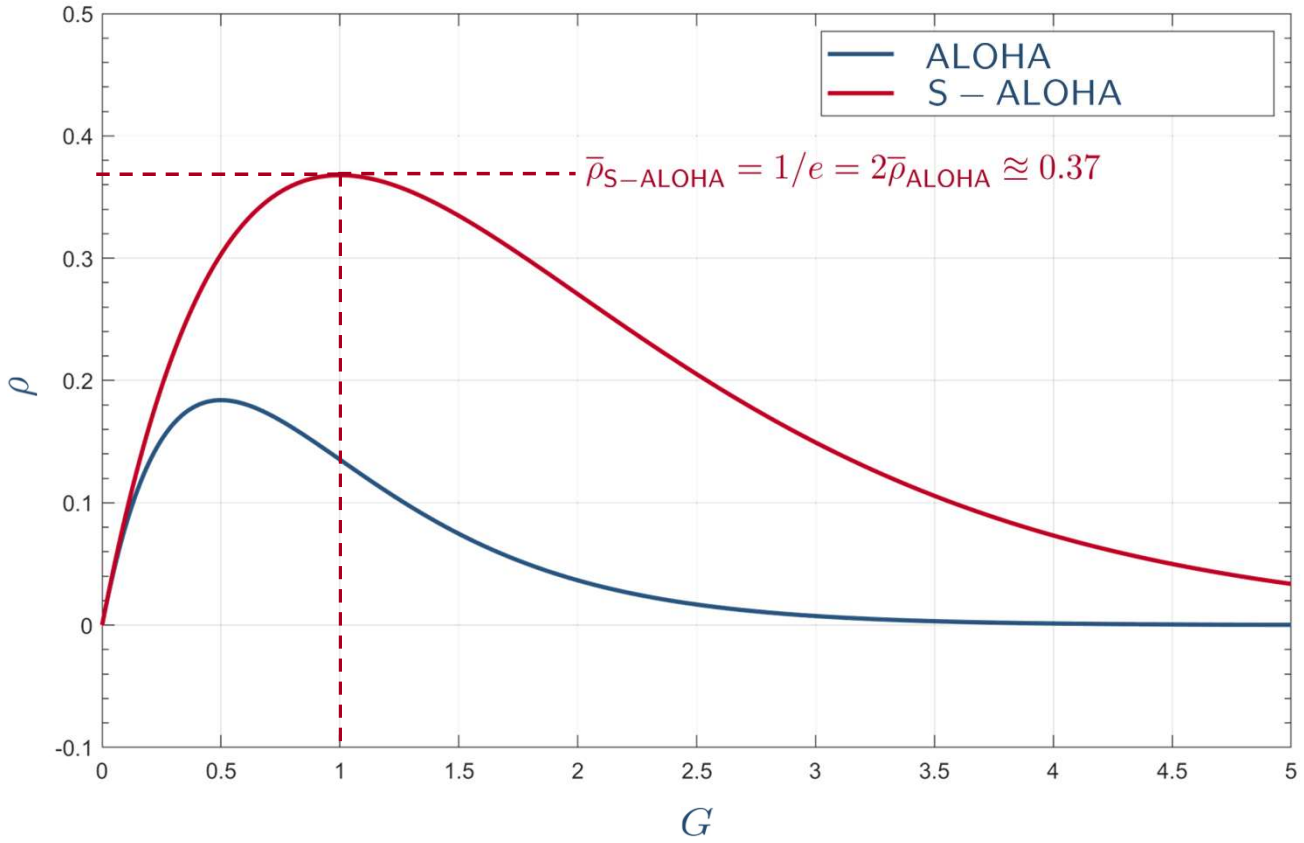
When considering time slotting, the probability of no collision across packets becomes

$$\begin{aligned} P_{\mathsf{s}} &= \Pr\{\Delta \geq T_p\} \\ &= \int_{T_p}^{+\infty} N\mu \cdot e^{-N\mu\delta} d\delta \\ &= e^{-G} \end{aligned}$$

Hence, we get:

$$\rho_{\mathsf{S-ALOHA}} = G \cdot e^{-G}$$

ALOHA vs. S-ALOHA



Efficiency comparison across multiplexing and multiple access techniques

Efficiency of deterministic-based approaches

It is now interesting to compare the packet-oriented and the connection-oriented approaches in terms of efficiency, using the definition introduced above:

$$\rho = \frac{\text{bitrate successfully delivered}}{\text{used bandwidth}}$$

Efficiency of multiplexing techniques

When considering frequency / time / code division multiplexing:

- used bandwidth: $B = N/T$
- bitrate successfully delivered: $R_b = N \cdot 1/T = N/T$

As a consequence,

$$\rho_{\text{FDM}} = \rho_{\text{TDM}} = \rho_{\text{CDM}} = 1$$

To sum up: multiplexing has unitary efficiency, thanks to its orthogonality

Efficiency of FDMA

When dealing with FDMA, we need to account for the guard bandwidth B_g:

- used bandwidth: $B = N \cdot [(1+\beta)/T + B_g]$
- bitrate successfully delivered: $R_b = N \cdot 1/T = N/T$

As a consequence,

$$\rho_{\text{FDMA}} = \frac{N/T}{N \cdot [(1+\beta)/T + B_g]} = \frac{1}{1+\beta+B_g T}$$

Example: with $\beta = 0.2$ and $B_g = 0.1/T$, we get $\rho_{\text{FDMA}} \approxeq 0.77$

Efficiency of TDMA

Similarly, when dealing with TDMA, efficiency is reduced by the guard interval T_g:

- used bandwidth: $B = N/T$
- bitrate successfully delivered: $R_b = \frac{N}{T} \cdot \frac{T_p/N}{T_p/N + T_g}$

This implies that

$$\rho_{\text{TDMA}} = \frac{N/T \cdot (MT/N)}{N/T \cdot (MT/N + T_g)} = \frac{1}{1 + NT_g/(MT)}$$

Example: with $T_g = 8.25T/N$ and $M = 148$ (GSM), we get $\rho_{\text{TDMA}} \approxeq 0.95$

Efficiency of CDMA (1/3)

With CDMA, the efficiency can be measured as follows:

- used bandwidth: $B = N/T$
- bitrate successfully delivered: $R_b = N_A \cdot 1/T$

which implies

$$\rho_{\text{CDMA}} = \frac{N_A}{N}$$

However, deriving the number of simultaneously active users N_A is not straight-forward, as it depends on the desired quality of service (QoS), which impacts the selected SNR E_b/N_0

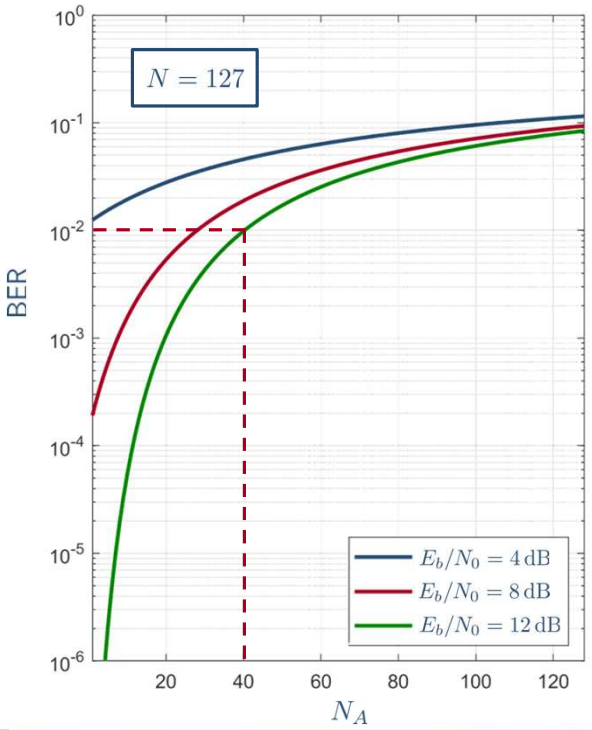
Efficiency of CDMA (2/3)

Example: $E_b/N_0 = 12\,\mathrm{dB}$, that yields $\mathrm{BER} < 10^{-7}$ in the single-user case

Target BER in the multi-user case (voice applications):

$\mathrm{BER} = 10^{-2}$

$$\rho_{\mathrm{CDMA}} = \frac{N_A}{N} = \frac{40}{127} \approxeq 0.31$$



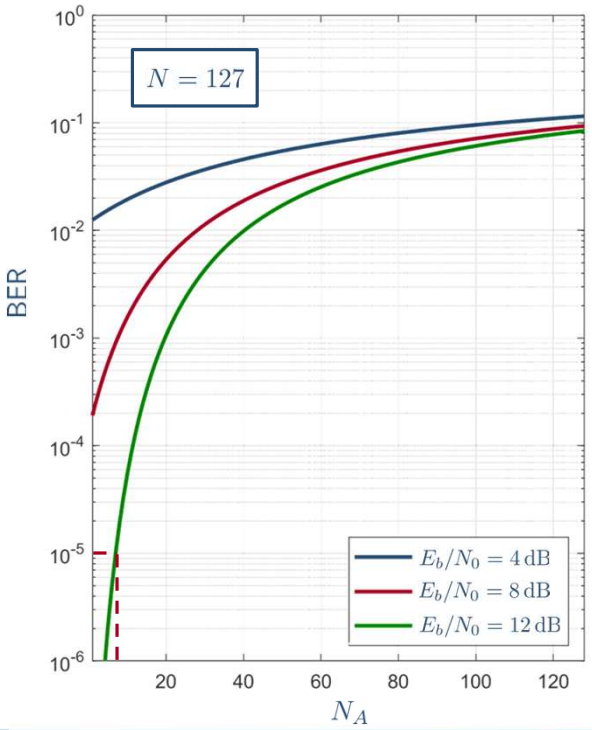
Efficiency of CDMA (3/3)

Example: $E_b/N_0 = 12\,\mathrm{dB}$, that yields $\mathrm{BER} < 10^{-7}$ in the single-user case

Target BER in the multi-user case (video applications):

$\mathrm{BER} = 10^{-5}$

$$\rho_{\mathrm{CDMA}} = \frac{N_A}{N} = \frac{7}{127} \approxeq 0.06$$



Efficiency of coded CDMA (1/3)

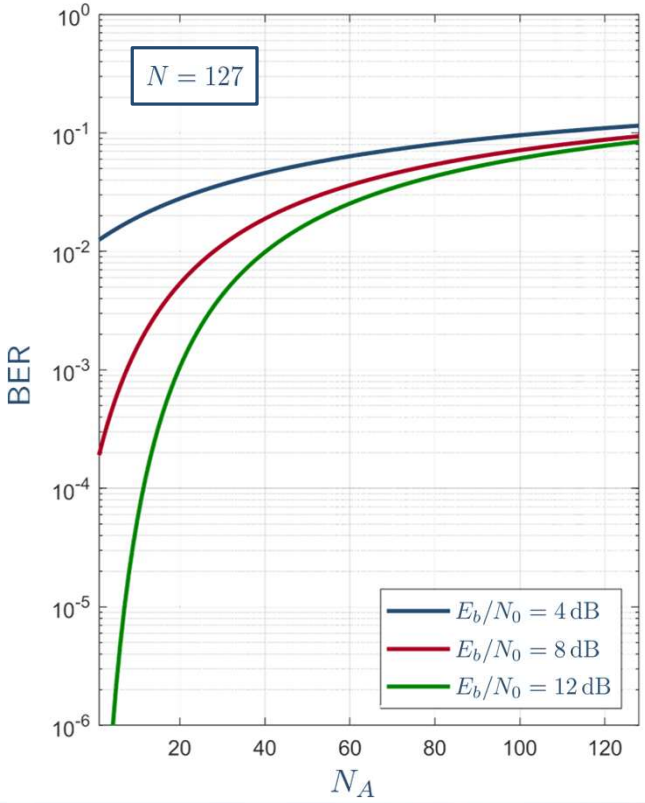
The impact of MAI is a very limiting factor for the efficiency of a CDMA-based system

We can improve the efficiency performance by adopting channel coding techniques

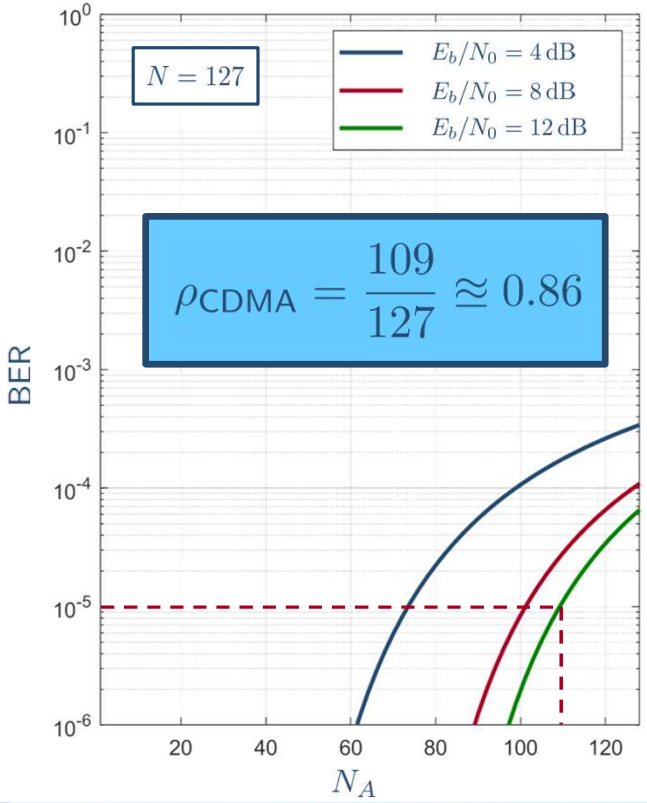
Example: low-density parity check (LDPC) code with rate 1/2, keeping the chip time equal to the uncoded case (and hence the used bandwidth)

Efficiency of coded CDMA (2/3)

uncoded case



coded case



Efficiency of coded CDMA (3/3)

With channel coding, with CDMA we can attain efficiencies in the same order as TDMA and FDMA (i.e., systems with orthogonal multiple access) (by relaxing the QoS requirements, even $\rho_{\text{CDMA}} > 1$ are possible!)

Please note that using channel coding with TDMA and FDMA does not further improve the efficiency, yet increasing the energy efficiency of the system given a certain QoS

Channel coding can also be applied to packet-based multiple access schemes, thus originating additional variants to the ALOHA protocol (e.g., spread-spectrum ALOHA)

History of wireless communications

A brief history of wireless communications (1/3)

- 1864: Maxwell proves the existence of electromagnetic waves
- 1887: Hertz sends and receives wireless waves, using a spark transmitter and a resonator receiver
- 1895: Guglielmo Marconi sends a radio signal over more than a mile, from the Isle of Wight to a tugboat 18 miles away
- 1904: Fleming patents the diode
- 1906: DeForest patents the triode amplifier; first speech wireless transmission, by Fessenden
- WW I: Rapid development of communications intelligence, intercept technology, cryptography
- 1920: Marconi discovers short-wave radio, with wavelengths between 10 and 100 meters
- 1935: Armstrong invents the frequency modulation (FM)

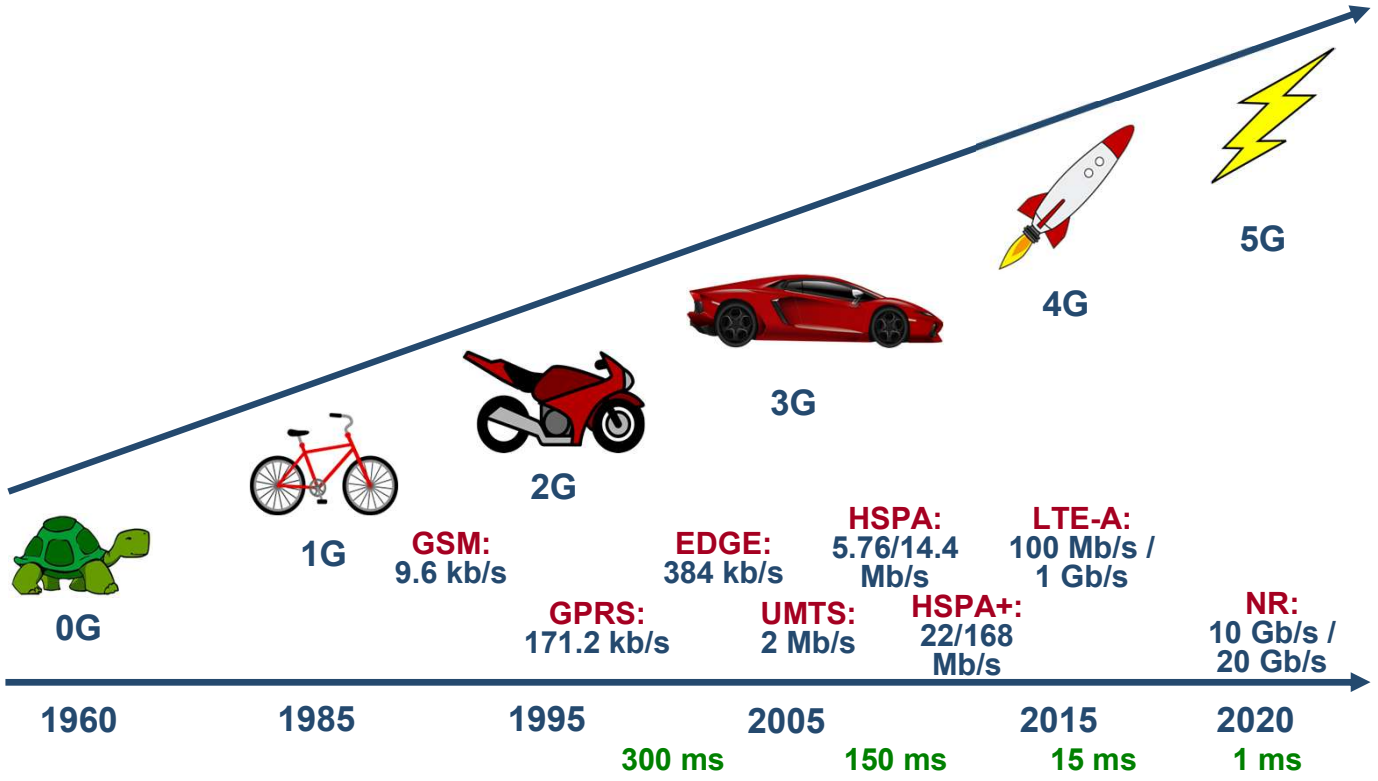
A brief history of wireless communications (2/3)

Mobile wireless systems ensure the communications between mobile nodes

In the last ninety years, wireless technology has evolved over many aspects:

- increased coverage distance
- increased quality (throughput, error rate performance, spectral efficiency)
- improved availability of services (broadband communications)
- decreased energy consumption (energy efficiency)
- reduced costs (for both service providers and subscribers/users)
- reduced device sizes and costs

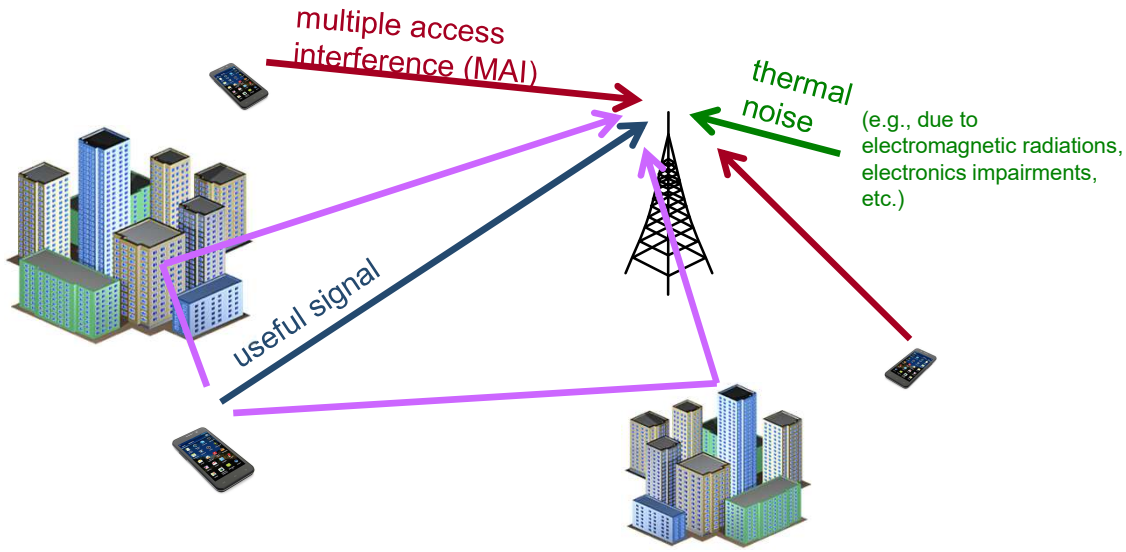
A brief history of wireless communications (3/3)



Basics of wireless propagation

The wireless propagation channel

Received contributions at the receiver (uplink):



The wireless channel between the transmitter and the receiver fluctuates randomly for a number of causes

Free-space propagation

If we consider the air as a perfectly uniform medium, the received power can be expressed as

$$P_{\mathsf{R}}(d) = \frac{P_{\mathsf{T}}}{L(d)}$$

where

$$L(d) = \frac{1}{G_{\mathsf{T}} G_{\mathsf{R}}} \left(\frac{4\pi d}{\lambda} \right)^2$$

tx-rx distance

gain of the transmit antenna

gain of the receive antenna

carrier wavelength

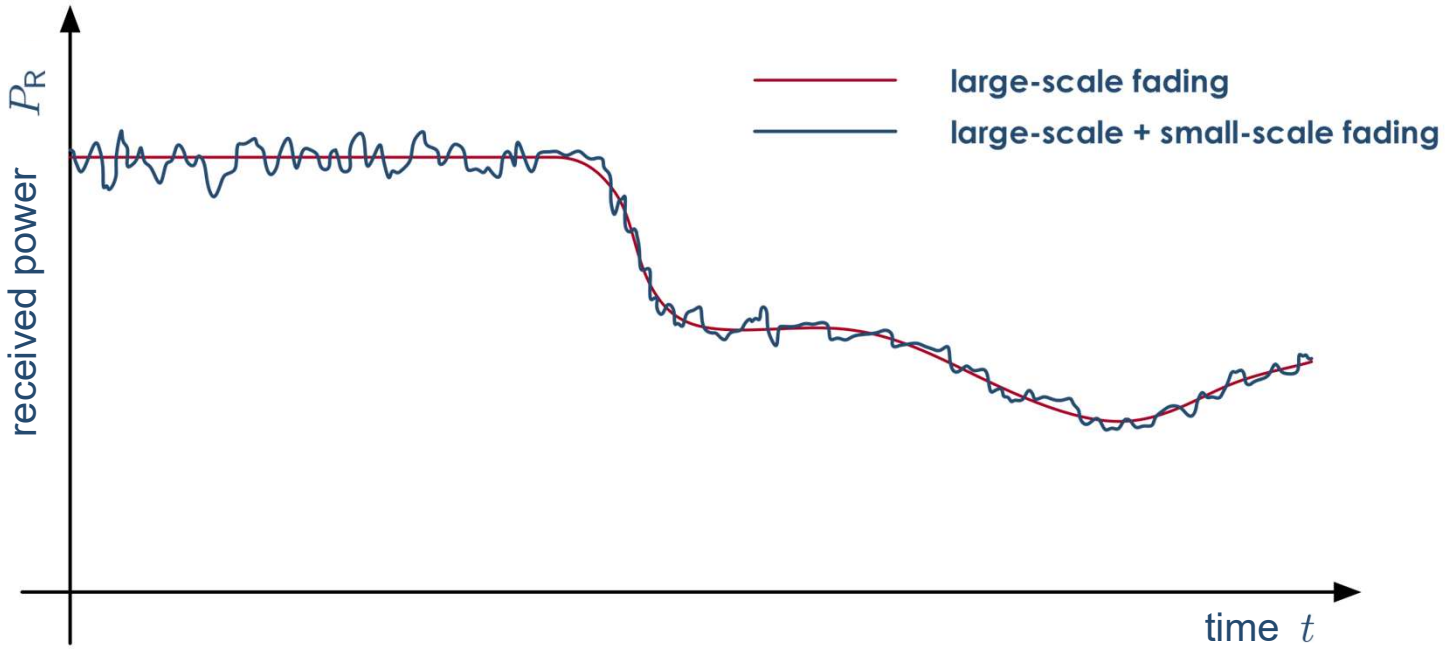
However, this model is not accurate to describe the wireless channel experienced by cellular-communication signals

Large- and small-scale models (1/2)

A better suited propagation model is composed by:

- large-scale models, that predict the average energy received in a wireless system as a function of the distance between the transmitter and the receiver
- small-scale models, that account for the instantaneous variations in the propagation conditions

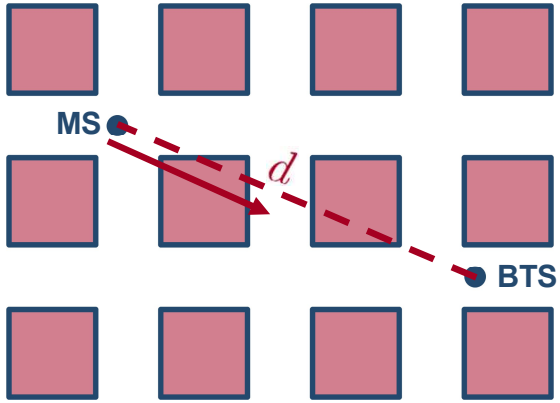
Large- and small-scale models (2/2)



Large scale fading

Large-scale fading models (1/3)

Average received power as a function of the MS-BTS distance d



Using the Hata and Okumura models,

$$\overline{P}_{\mathsf{R}}(d) = \begin{cases} P_{\mathsf{T}}/L(d) = \chi' \cdot P_{\mathsf{T}}/d^2, & \text{for } d \leq d_0 \\ \chi \cdot P_{\mathsf{T}}/d^n = \chi' \cdot P_{\mathsf{T}} \cdot d_0^{n-2}/d^n, & \text{for } d \geq d_0 \end{cases}$$

reference distance

Large-scale fading models (2/3)

Average received power as a function of the MS-BTS distance d

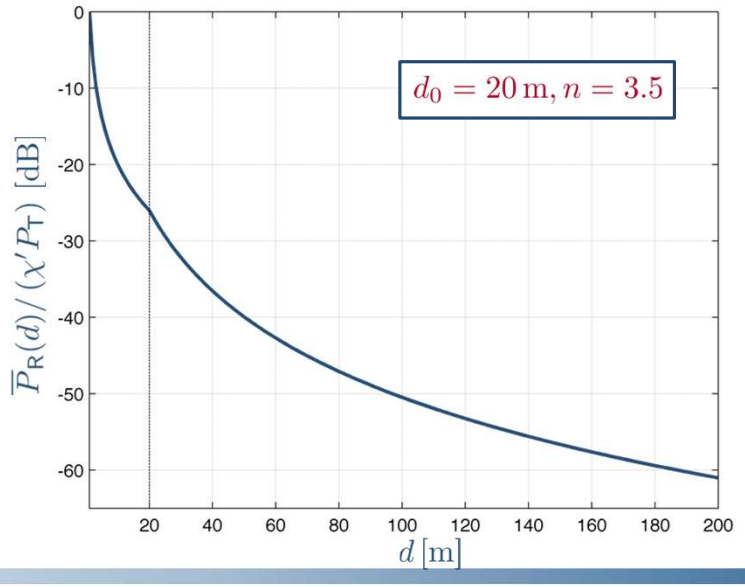
$$\overline{P}_{\mathsf{R}}(d) = \begin{cases} P_{\mathsf{T}}/L(d) = \chi' \cdot P_{\mathsf{T}}/d^2, & \text{for } d \leq d_0 \\ \chi \cdot P_{\mathsf{T}}/d^n = \chi' \cdot P_{\mathsf{T}} \cdot d_0^{n-2}/d^n, & \text{for } d \geq d_0 \end{cases}$$

$$\chi' = G_{\mathsf{T}} G_{\mathsf{R}} \left(\frac{\lambda}{4\pi}\right)^2$$

$$\chi = \chi' \cdot d_0^{n-2} = G_{\mathsf{T}} G_{\mathsf{R}} \left(\frac{\lambda}{4\pi}\right)^2 d_0^{n-2}$$

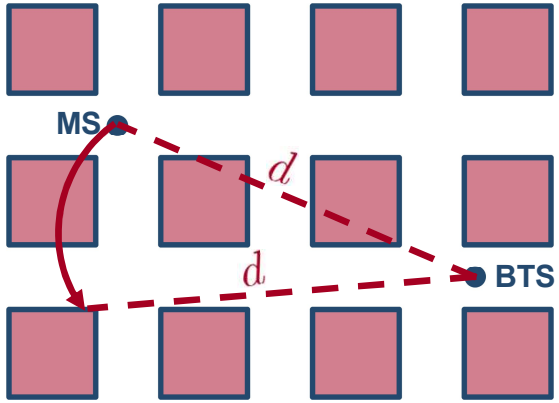
$$d_0 = \begin{cases} 20 \div 100\,\text{m}, & \textbf{urban scenarios} \\ 0.5 \div 1\,\text{km}, & \textbf{rural areas} \end{cases}$$

$$n = \begin{cases} 2, & \textbf{free space} \\ 2.7 \div 3.5 & \textbf{urban scenarios} \\ 4 \div 5 & \textbf{indoor} \end{cases}$$



Large-scale fading models (3/3)

Average received power as a function of a constant distance d



$$P_{\mathrm{R}}(d) = 10^{\nu(d)/10} \qquad \text{where} \qquad \nu(d) \sim \mathcal{N}\left(\overline{\nu}(d), \sigma_\nu^2\right)$$

dependent on the propagation scenario

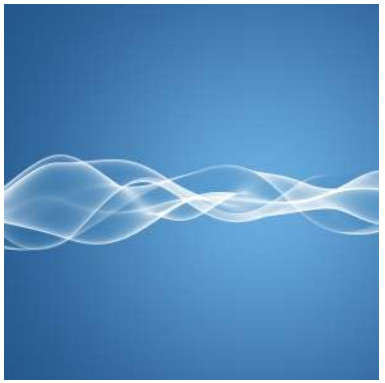
given by the Hata-Okumura model:
$\overline{\nu}(d) = 10 \log_{10} \overline{P}_{\mathrm{R}}(d)$

Small scale fading

Small-scale fading

The propagation laws can be computed using the Maxwell equations. However, this approach is not useful, due to:

- a large computational complexity, but also
- the need for a valid model for different propagation conditions, also considering the time variability



We need to identify a practical method to account for the macroscopic phenomena of wireless propagation

Wavelengths and frequencies in wireless systems (1/2)

Wireless radio spectrum:

Carrier wavelengths:

speed of light, $3 \cdot 10^8 \, \mathrm{m/s}$

$$\lambda = \frac{c}{f_0} = 1\,\mathrm{cm} \div 1\,\mathrm{m}$$

carrier frequency

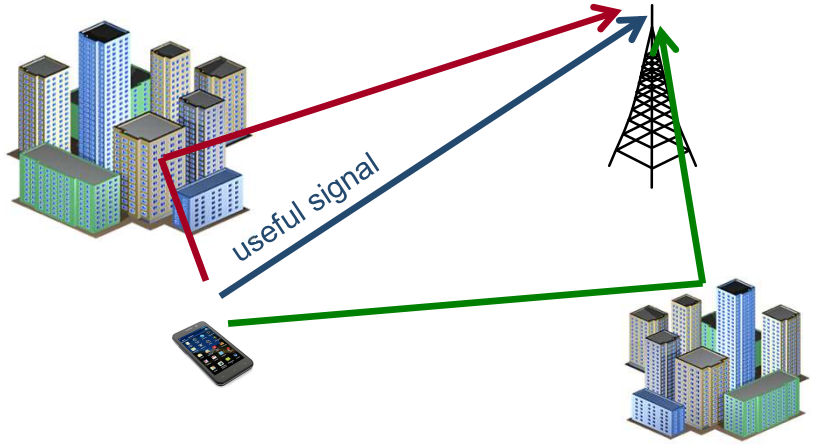
Wavelengths and frequencies in wireless systems (2/2)

Why are such frequencies particularly attractive?

- larger f_0's yield larger path losses: $L(d) \propto \frac{1}{\lambda^n}$
- smaller f_0's call for larger antennas (with size comparable with λ)
- smaller f_0's show favorable conditions for over-the-horizon (OTH) propagation, thus reducing the potential for frequency reuse
- such f_0's can accommodate large enough channel spacing and provide room for large user multiplexing and multiple access
- such f_0's have good indoor propagation

Multipath propagation

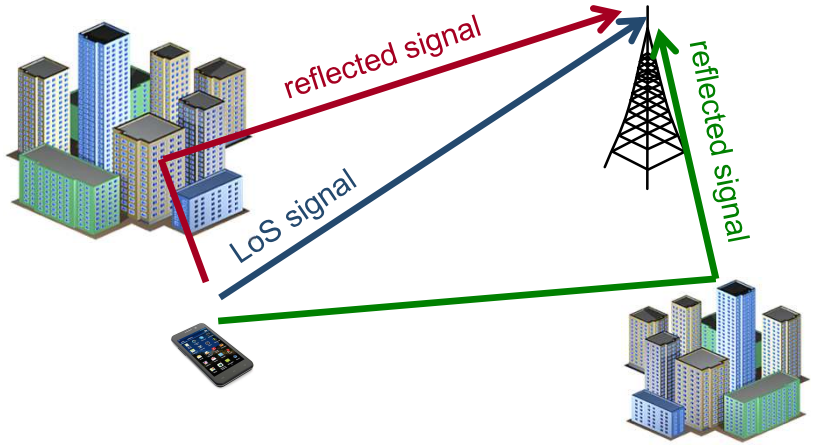
In this frequency spectrum, the wireless signal experiences a multipath propagation: the received signal is a linear combination of multiple paths



In addition to the direct path, a.k.a. line of sight (LoS) path, the wireless signal can propagate due to:

- reflection: $S \gg \lambda$
- shadowing: $S \simeq \lambda$
- scattering: $S \ll \lambda$

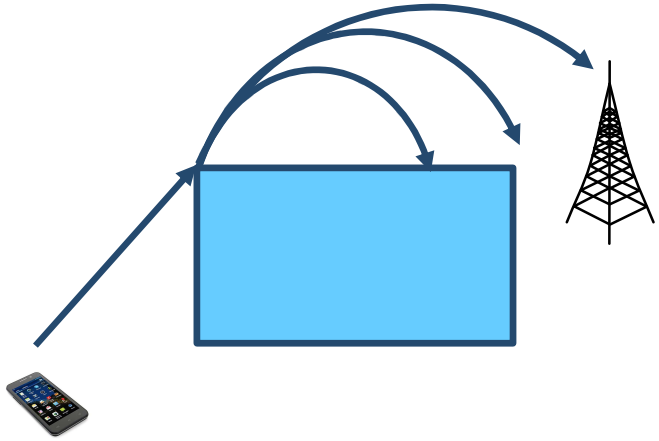
Reflection



Due to the different propagation lengths, for each path the reflection introduces:

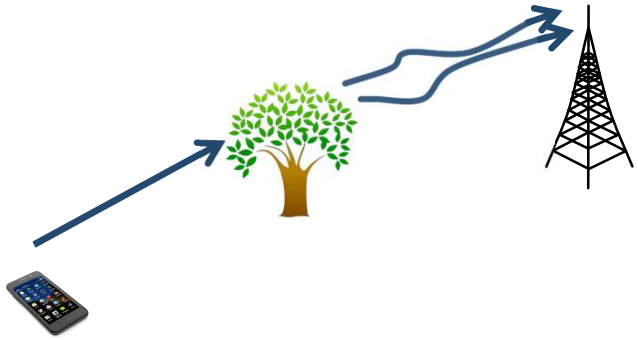
- amplitude attenuation
- group delay
- phase delay

Shadowing



Shadowing introduces additional paths when the transmitter and the receiver are not in visibility, thus affecting the statistics of the channel

Scattering



Similarly, scattering introduced a disordered reflection of the electro-magnetic waves, thus impacting on attenuations and phase delays

Multipath propagation: frequency selectivity

Multipath propagation model

To sum up, the received signal is a linear combination of a number of different propagation paths, each having its own attenuation, phase rotation, and time delay:

$$y(t) = \sum_{\ell=1}^{L(t)} \rho_\ell(t)\, e^{j\varphi_\ell(t)} x(t - \tau_\ell(t)) e^{-j2\pi f_0 \tau_\ell(t)}$$

$$= \sum_{\ell=1}^{L(t)} \rho_\ell(t)\, e^{j\theta_\ell(t)} x(t - \tau_\ell(t))$$

- $L(t)$: number of propagation paths
- $\rho_\ell(t)$: attenuation of the ℓ-th path
- $\theta_\ell(t)$: phase delay of the ℓ-th path
- $\tau_\ell(t)$: time delay of the ℓ-th path

(Preliminary) classification of the wireless channel

Time domain:

- static (time-invariant): its statistics change very slowly wrt signaling time
- time-varying: its statistics are a function of time

Frequency domain:

- frequency-flat: its behavior is similar across the frequency components of the signal
- frequency-selective: each frequency component of the signal is distorted in a different way by the wireless channel

Static frequency-flat channels (1/4)

A static channel means: the random processes $L(t)$, $\{\rho_\ell(t)\}$, $\{\theta_\ell(t)\}$, and $\{\tau_\ell(t)\}$ are not functions of time (i.e., they are just random variables):

$$y(t) = \sum_{\ell=1}^{L} \rho_\ell e^{j\theta_\ell} x(t - \tau_\ell)$$

Let us suppose that the standard deviation $\sigma_\tau = \sqrt{\mathbb{E}\{|\tau_\ell - \overline{\tau}|^2\}}$ is much smaller than the signaling interval T, where $\overline{\tau} = \mathbb{E}\{\tau_\ell\}$:

$$\sigma_\tau \ll T$$

With good accuracy, we can approximate

$$\tau_\ell \cong \overline{\tau} \quad \forall \ell$$

Static frequency-flat channels (2/4)

Hence,

$$y(t) = \sum_{\ell=1}^{L} \rho_\ell e^{j\theta_\ell} x(t-\tau_\ell) \approxeq x(t-\overline{\tau}) \cdot \sum_{\ell=1}^{L} \rho_\ell e^{j\theta_\ell}$$

$$= \overline{\rho} e^{j\overline{\theta}} x(t-\overline{\tau})$$

where

$$\overline{\rho} = \left| \sum_{\ell=1}^{L} \rho_\ell e^{j\theta_\ell} \right|, \qquad \overline{\theta} = \angle \left(\sum_{\ell=1}^{L} \rho_\ell e^{j\theta_\ell} \right)$$

In practice, the received signal is just a scaled copy of the transmitted signal $x(t)$, delayed by $\overline{\tau}$, attenuated by $\overline{\rho}$, and rotated by $\overline{\theta}$

Static frequency-flat channels (3/4)

This happens for each frequency component of the input signal $x(t)$:

$$Y(f) \triangleq \mathcal{F}\{y(t)\} = \mathcal{F}\left\{\overline{\rho} e^{j\overline{\theta}} x(t-\overline{\tau})\right\}$$
$$= \overline{\rho} e^{j\overline{\theta}} e^{-j2\pi f\overline{\tau}} X(f)$$

where

$$X(f) \triangleq \mathcal{F}\{x(t)\} = \int_{-\infty}^{+\infty} x(t) e^{-j2\pi ft} dt$$

is the Fourier transform of the signal $x(t)$

If $\sigma_\tau \ll T$, the channel is frequency-flat: $H(f) = \dfrac{Y(f)}{X(f)} = \overline{\rho} e^{j(\overline{\theta} - 2\pi f\overline{\tau})}$

Static frequency-flat channels (4/4)

What is the statistical distribution of $\overline{\tau}$, $\overline{\rho}$, and $\overline{\theta}$?

- usually, $\overline{\tau}$ depends on the propagation scenario, and it is determined by extensive measurement campaigns

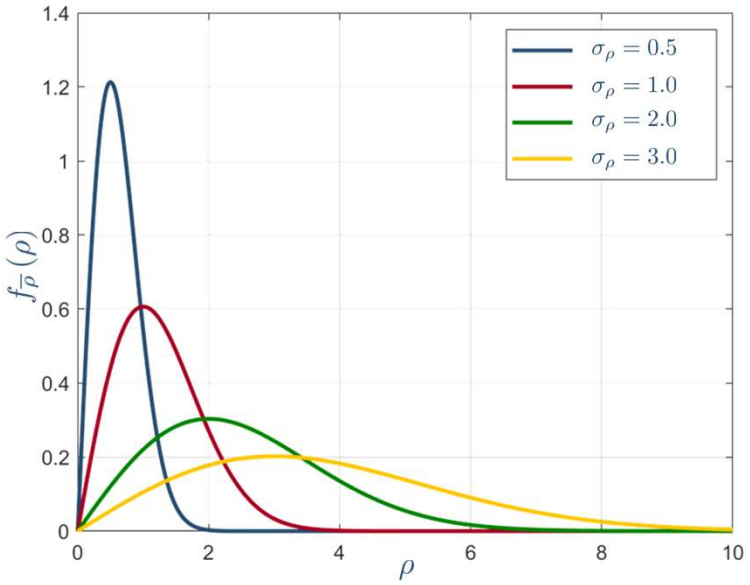
- the probability density function (pdf) of $\overline{\rho}$ and $\overline{\theta}$ depends on the two different situations:
  - non-LoS (NLoS) propagation (typically, urban scenarios)
  - LoS propagation (typically, rural areas)

NLoS scenarios

- the attenuation $\overline{\rho}$ follows a Rayleigh distribution:

$$f_{\overline{\rho}}(\rho) = \frac{\rho}{\sigma_\rho^2} e^{-\rho^2/(2\sigma_\rho^2)}\, \mathrm{u}(\rho)$$

where $\mathbb{E}\{\overline{\rho}^2\} = 2\sigma_\rho^2 = \overline{P}_{\mathrm{R}}(d)$ is the average value given by large-scale fading models



- the phase delay $\overline{\theta}$ is uniformly distributed in $[-\pi, +\pi]$

LoS scenarios

- the attenuation $\overline{\rho}$ follows a Rice distribution:

$$f_{\overline{\rho}}(\rho) = \frac{\rho}{\sigma_\rho^2} e^{-\rho^2/(2\sigma_\rho^2)} e^{-\kappa} \times \mathrm{I}_0\left(\sqrt{2\kappa}\frac{\rho}{\sigma_\rho}\right) \mathrm{u}(\rho)$$

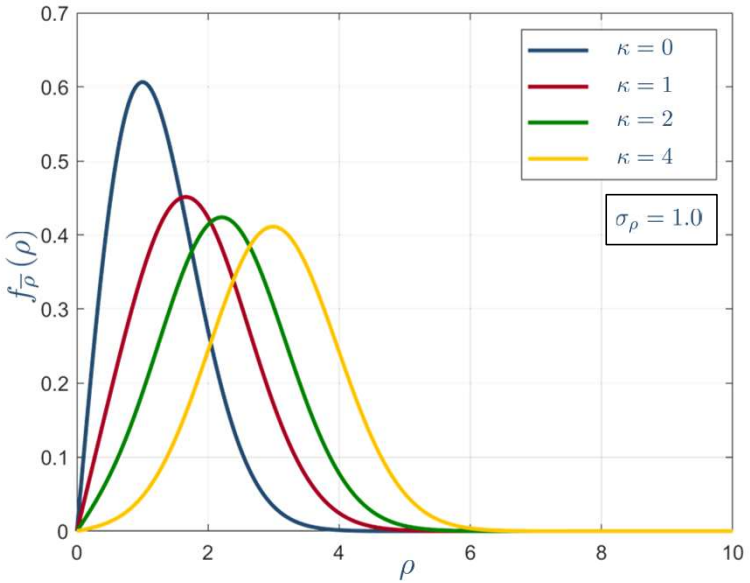
where

$$\kappa \triangleq \frac{\rho_1^2/2}{\sigma_\rho^2}$$

$\rho_1^2/2$: average power of the LoS path

σ_ρ^2: average power due to multipath propagation

is the Rice factor



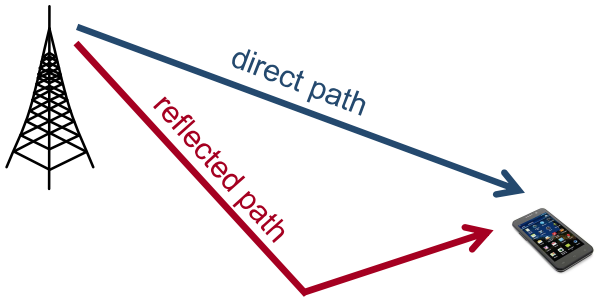
- the pdf of the phase delay $\overline{\theta}$ cannot be written in a closed form

Static frequency-selective channels (1/4)

Suppose now that the hypothesis $\sigma_\tau \ll T$ does not hold: this means that we now have

$$y(t) = \sum_{\ell=1}^{L} \rho_\ell e^{j\theta_\ell} x\left(t - \tau_\ell\right)$$

For the sake of simplicity, let's consider the two-ray channel, i.e., $L = 2$:



$$y(t) = \underbrace{\rho_1 e^{j\theta_1} x\left(t - \tau_1\right)}_{\text{direct (LoS) path}} + \underbrace{\rho_2 e^{j\theta_2} x\left(t - \tau_2\right)}_{\text{reflected path}}$$

Static frequency-selective channels (2/4)

To simplify the notation, let us take:

$$\rho_1 = 1, \quad \theta_1 = 0, \quad \tau_1 = 0$$
$$\rho_2 = \rho, \quad \theta_2 = \theta, \quad \tau_2 = \tau$$

received signal: $y(t) = x(t) + \rho e^{j\theta} x(t - \tau)$

Fourier transform: $Y(f) = X(f) \cdot \left[1 + \rho e^{j\theta} e^{-j2\pi f \tau}\right]$

The frequency response of the channel is

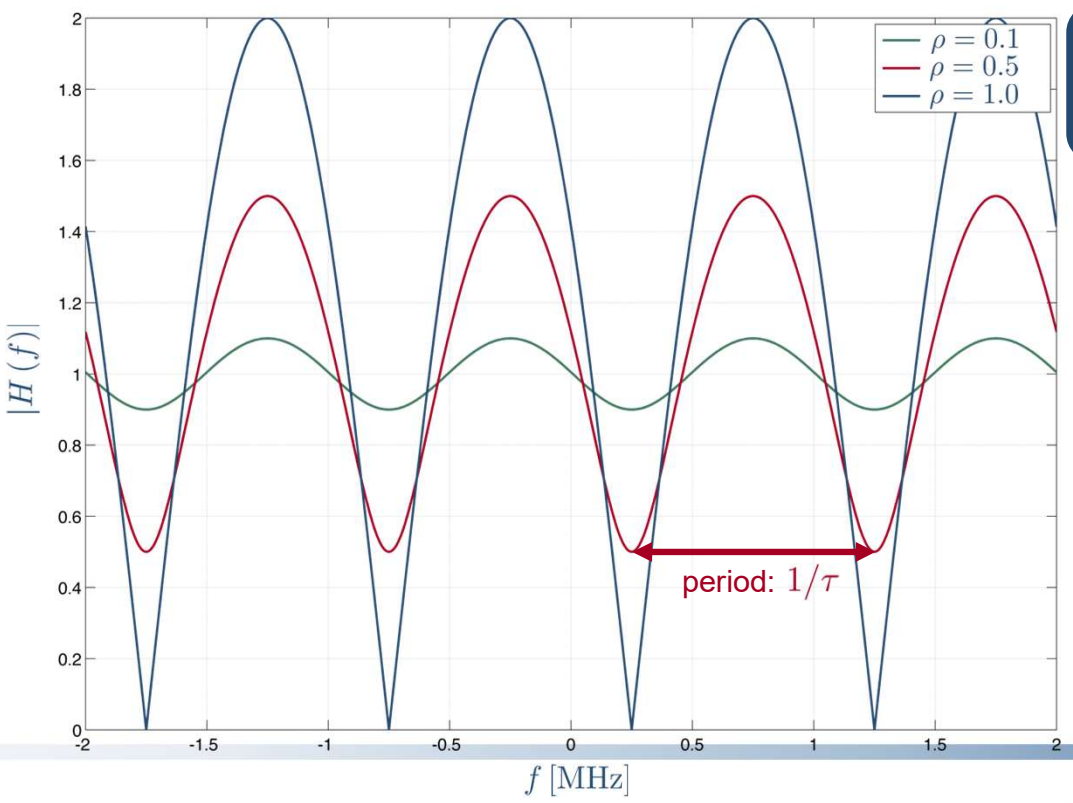
$$H(f) = \frac{Y(f)}{X(f)} = 1 - \rho e^{-j2\pi(f - f_N)\tau}$$

where $f_N = \frac{1}{2\tau} + \frac{\theta}{2\pi\tau}$ is the notch frequency of the channel

Static frequency-selective channels (3/4)

The amplitude response of the two-ray channel is

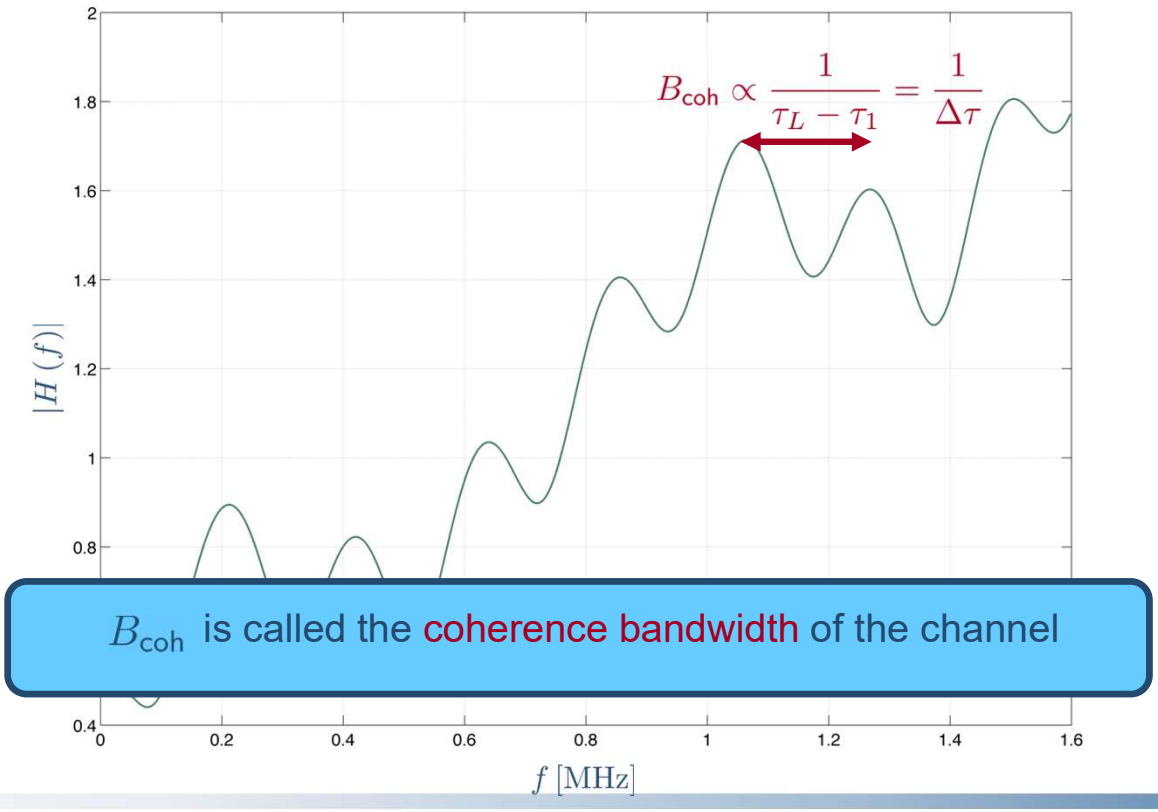
$$|H(f)| = \sqrt{1 + \rho^2 - 2\rho\cos\left[2\pi\left(f - f_{\mathrm{N}}\right)\tau\right]}$$



$\tau = 1\,\mu\mathrm{s}$
$f_{\mathrm{N}} = 0.25\,\mathrm{MHz}$
$\theta = -\pi/2$

Static frequency-selective channels (4/4)

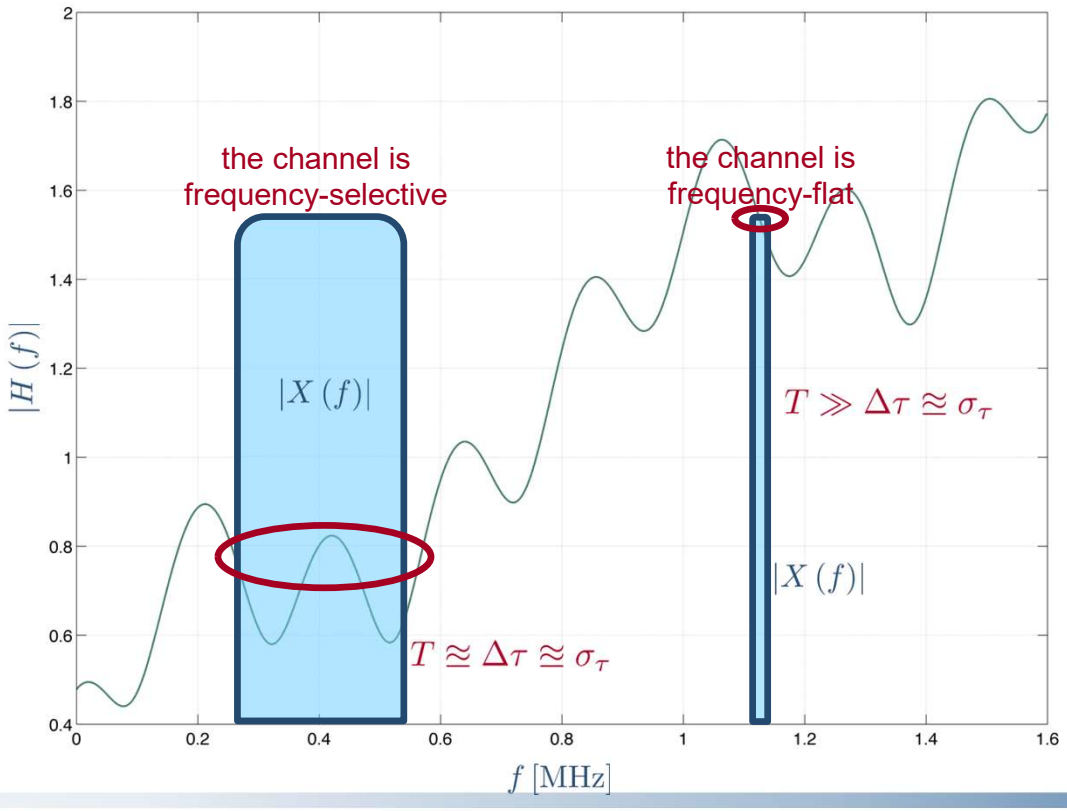
When extending the calculations to the L-ray channel, we get



B_{coh} is called the coherence bandwidth of the channel

The concept of frequency selectivity (1/2)

We know that the bandwidth of a signal $x(t)$ is $B \propto 1/T$



The concept of frequency selectivity (2/2)

The frequency selectivity depends on the statistics of the channel and of the input signal

There is a practical way to assess the frequency selectivity of a channel:

- $B \ll B_{\text{coh}} \Leftrightarrow T \gg \sigma_\tau$: frequency-flat channel
- $B \cong B_{\text{coh}} \Leftrightarrow T \cong \sigma_\tau$: frequency-selective channel

Example:

- urban scenarios: $\sigma_\tau \cong 1\,\mu\text{s} \Rightarrow B_{\text{coh}} \cong 1\,\text{MHz}$
- 3G and 4G signals: $B \geq 3.5\,\text{MHz}$

Some form of equalization is needed to combat the frequency selectivity

Multipath propagation: time selectivity

Multipath propagation model

In a multipath scenario, the received signal is a linear combination of a number of different propagation paths, each having its own attenuation, phase rotation, and time delay:

$$y(t) = \sum_{\ell=1}^{L(t)} \rho_\ell(t)\, e^{j\varphi_\ell(t)} x(t - \tau_\ell(t)) e^{-j2\pi f_0 \tau_\ell(t)}$$

$$= \sum_{\ell=1}^{L(t)} \rho_\ell(t)\, e^{j\theta_\ell(t)} x(t - \tau_\ell(t))$$

$L(t)$: number of propagation paths

$\theta_\ell(t)$: phase delay of the ℓ-th path

$\rho_\ell(t)$: attenuation of the ℓ-th path

$\tau_\ell(t)$: time delay of the ℓ-th path

Frequency selectivity

To assess the frequency selectivity of a channel, we need to compare the parameters of the input signal (bandwidth B, time interval T) with the characteristics of the channel (coherence bandwidth B_{coh}, delay spread σ_τ):

- $B \ll B_{\text{coh}} \Leftrightarrow T \gg \sigma_\tau$: frequency-flat channel
- $B \approxeq B_{\text{coh}} \Leftrightarrow T \approxeq \sigma_\tau$: frequency-selective channel

Time-varying frequency-flat channels (1/4)

Due to the relative motion between the transmitter and the receiver, the communication medium (the wireless channel) evolves through time:

$$y(t) = \sum_{\ell=1}^{L(t)} \rho_\ell(t)\, e^{j\theta_\ell(t)} x(t - \tau_\ell(t))$$

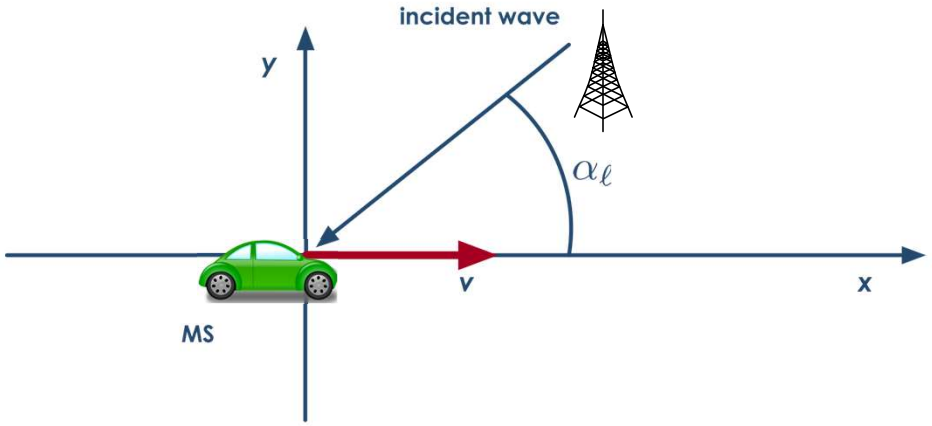
For simplicity, let's assume a frequency-flat channel: $\sigma_\tau \ll T \Rightarrow \tau_\ell(t) \cong \overline{\tau} \quad \forall \ell$

Similarly to the static case,

$$\begin{aligned} y(t) &\cong x(t - \overline{\tau}) \cdot \textstyle\sum_{\ell=1}^{L(t)} \rho_\ell(t)\, e^{j\theta_\ell(t)} \\ &= \overline{\rho}(t) \cdot e^{j\overline{\theta}(t)} \cdot x(t - \overline{\tau}) \\ &= \underbrace{A(t)}_{\text{fading process}} \cdot x(t - \overline{\tau}) \end{aligned}$$

Time-varying frequency-flat channels (2/4)

To study time and frequency characteristics of $A(t)$, let us use the kinematic model for the MS:



Due to Doppler effect, each band pass frequency $f \in \left[f_0 - \frac{B}{2}, f_0 + \frac{B}{2}\right]$, where f_0 is the carrier frequency, is shifted at the receive side by its Doppler shift Δf:

$$\Delta f = \frac{v}{c} \cdot f \cdot \cos\left(\alpha_\ell\right)$$

Time-varying frequency-flat channels (3/4)

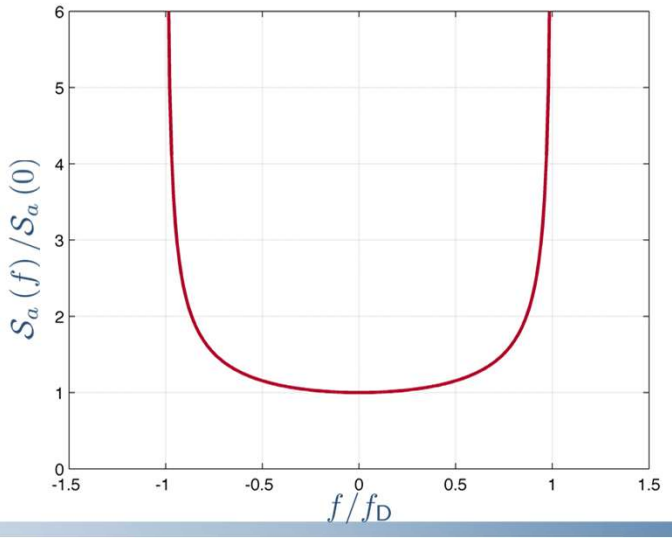
The behavior of $A(t)$ is given by the impact of the Doppler effect over all signal frequency components $f \in \left[f_0 - \frac{B}{2}, f_0 + \frac{B}{2}\right]$

A key parameter is the maximum Doppler shift at the carrier frequency f_0, called the Doppler spread f_D:

$$f_\mathrm{D} \triangleq \max_{\alpha_\ell} |\Delta f| = \frac{v}{c} \cdot f_0$$

Using the Clarke's model, we can compute the power spectral density (PSD) of the random process $A(t)$:

$$\mathcal{S}_a(f) = \frac{\sigma_\rho^2}{2\pi f_\mathrm{D}} \cdot \frac{1}{\sqrt{1-(f/f_\mathrm{D})^2}}$$

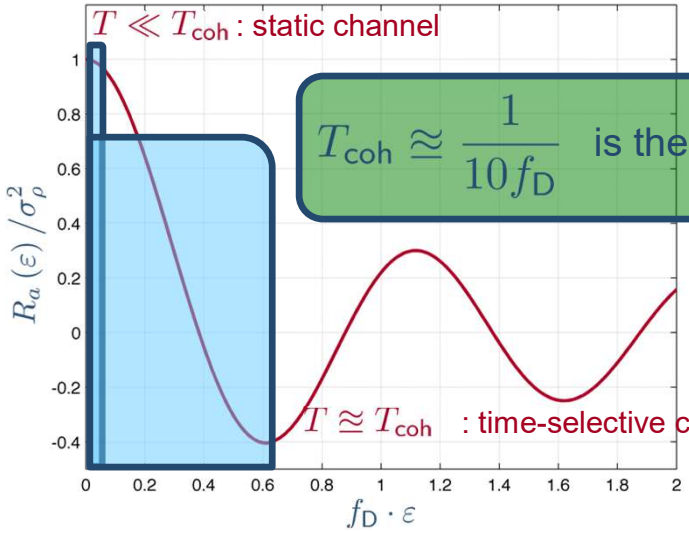


Time-varying frequency-flat channels (4/4)

Another useful statistical parameter to investigate the properties of $A(t)$ is its autocorrelation function:

$$R_a(\varepsilon) = \mathcal{F}^{-1}\{\mathcal{S}_a(f)\} = \mathbb{E}\{A(t) \cdot A(t+\varepsilon)\}$$

Using again the Clarke's model,



The concept of time selectivity

The time selectivity depends on the statistics of the channel and of the input signal

There is a practical way to assess the time selectivity of a channel:

- $B \gg f_{\mathrm{D}} \Leftrightarrow T \ll T_{\mathrm{coh}}$: static channel
- $B \cong f_{\mathrm{D}} \Leftrightarrow T \cong T_{\mathrm{coh}}$: time-selective channel

Example (3G systems):

- $v = 120\,\mathrm{km/h}, f_0 = 2\,\mathrm{GHz}: T_{\mathrm{coh}} = 0.45\,\mathrm{ms}$
- slot duration: $T_{\mathrm{slot}} = 0.66\,\mathrm{ms}$

Some design constraints must be added to combat the time selectivity

Multipath propagation: A summary

Frequency and time selectivity: A summary (1/5)

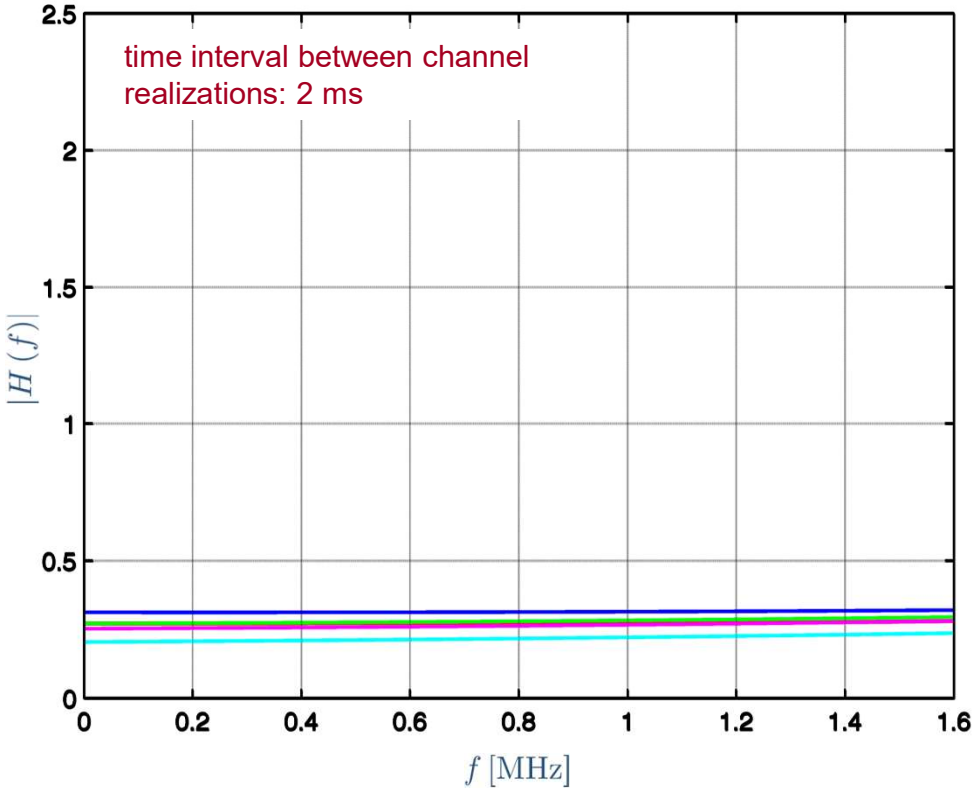
Both time- and frequency-selectivity are functions of both the channel and the input signal properties:

	frequency-flat	frequency-selective
static	$f_{\mathrm{D}} \ll B \ll B_{\mathrm{coh}}$	$f_{\mathrm{D}} \ll B \approxeq B_{\mathrm{coh}}$
time-selective	$f_{\mathrm{D}} \approxeq B \ll B_{\mathrm{coh}}$	$f_{\mathrm{D}} \approxeq B \approxeq B_{\mathrm{coh}}$

- frequency-selective = time-dispersive
- time-selective = frequency-dispersive
- frequency-selective ≠ time-selective!

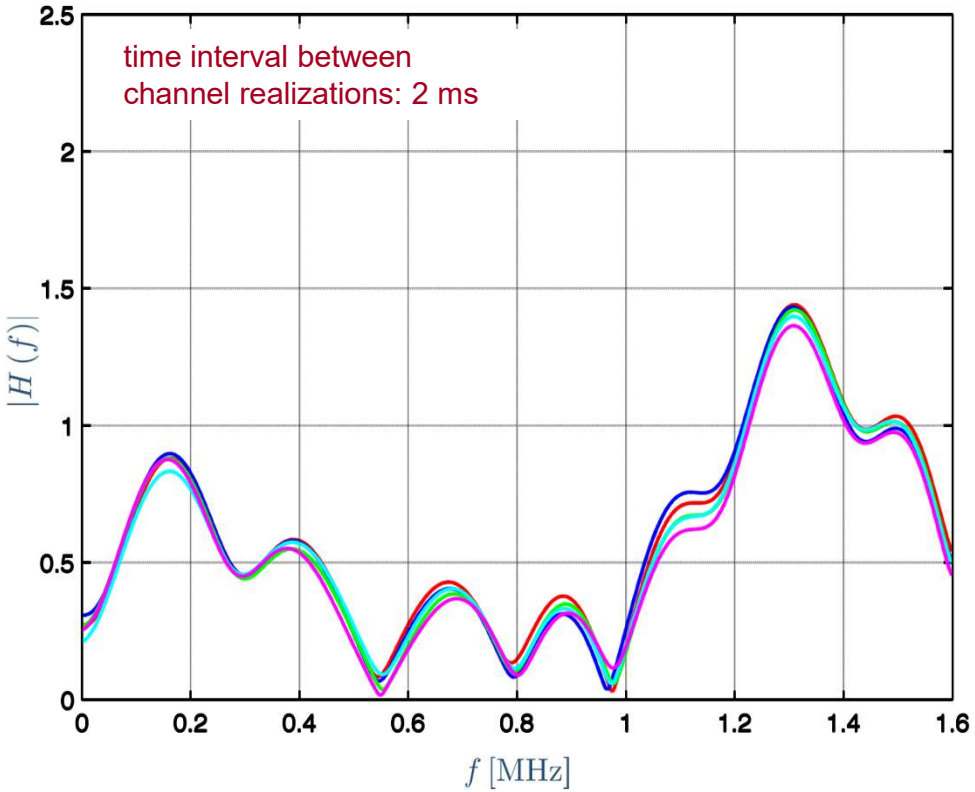
Frequency and time selectivity: A summary (2/5)

A static frequency-flat channel ($T_{\text{coh}} = 100\,\text{ms}, B_{\text{coh}} = 10\,\text{MHz}$):



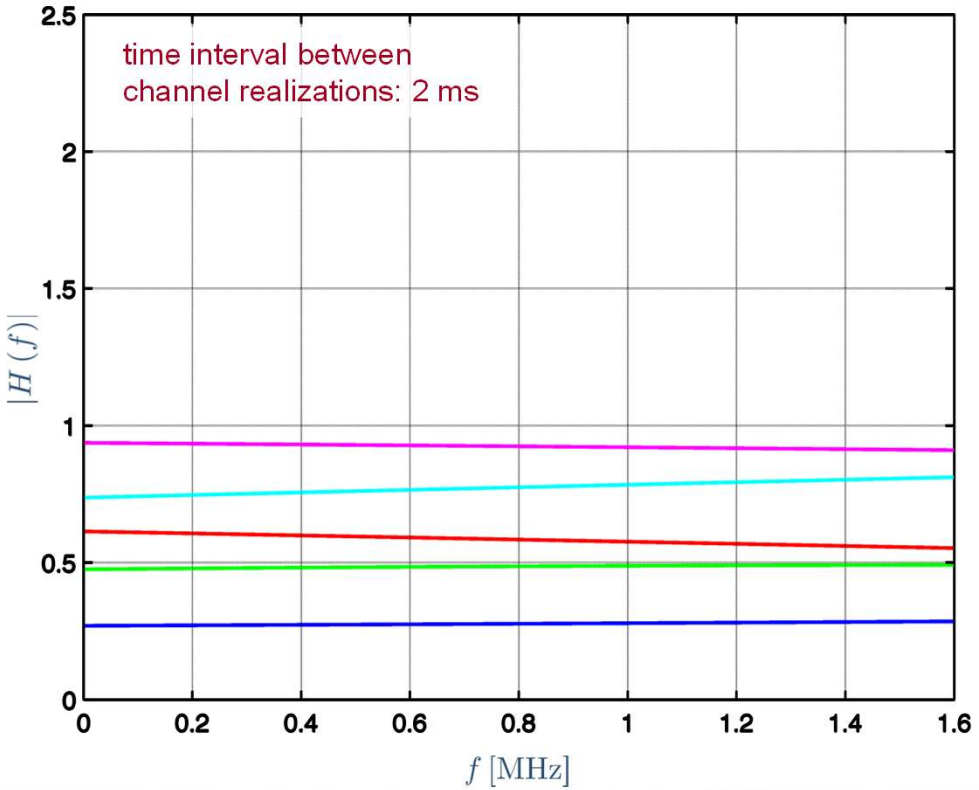
Frequency and time selectivity: A summary (3/5)

A static frequency-selective channel ($T_{\text{coh}} = 100\,\text{ms}, B_{\text{coh}} = 100\,\text{kHz}$):



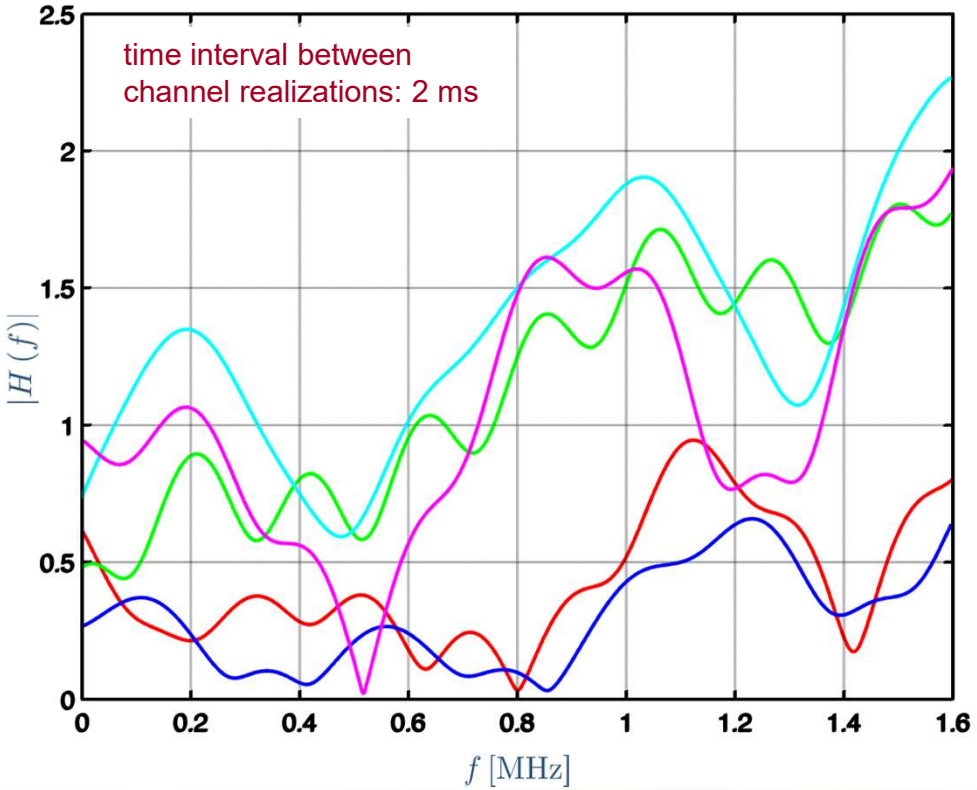
Frequency and time selectivity: A summary (4/5)

A time-selective frequency-flat channel ($T_{\text{coh}} = 1\,\text{ms}, B_{\text{coh}} = 10\,\text{MHz}$):



Frequency and time selectivity: A summary (5/5)

A doubly-selective channel ($T_{\text{coh}} = 1\,\text{ms}, B_{\text{coh}} = 100\,\text{kHz}$):



Bibliography

[01] B. Sklar and F. Harris, *Digital Communications* , 3rd ed. Upper Saddle River, NJ, US: Prentice Hall, 2021.

[02] T.S. Rappaport, *Wireless Communications: Principles and Practice,* 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 2002.

[03] A.J. Goldsmith, *Wireless Communications*. Cambridge, UK: Cambridge Univ. Press, 2005.

[04] J.G. Proakis and M. Salehi, *Digital Communications,* 5th ed. New York, NY: McGraw-Hill, 2007.

[05] A.F. Molisch, *Wireless Communications*. West Sussex, UK: J. Wiley & Sons, 2005.

[06] International Telecommunications Union (ITU), *Measuring digital development, Facts and Figures*, 2025. [Online] https://www.itu.int/itu-d/reports/statistics/facts-figures-2025/

[07] GSM Association, *The mobile economy Europe 2025*. [Online] https://www.gsma.com/solutions-and-impact/connectivity-for-good/mobile-economy/wp-content/uploads/2025/01/0125-Mobile-Economy-Europe-2025-web.pdf

[08] M. Luise, *Lecture notes on communication technologies.* 2022. [Online] https://docenti.ing.unipi.it/m.luise/ComTech/main_ComTech.pdf

[09] H. Benoit, *Digital Television: Satellite, Cable, Terrestrial, IPTV, Mobile TV in the DVB Framework*, 3rd ed. Burlington, MA: Elsevier, 2008.

[10] J.D. Parsons, *The Mobile Radio Propagation Channel.* Chichester, UK: J. Wiley & Sons, 2000.

[11] M. Pätzold, *Mobile Fading Channels.* Chichester, UK: J. Wiley & Sons, 2002.

[12] B. Sklar, “Rayleigh fading channels in mobile digital communication sys- tems, Part I: Characterization,” *IEEE Commun. Mag.*, vol. 35, no. 7, pp. 90–100, July 1997.

[13] B. Sklar, “Rayleigh fading channels in mobile digital communication systems, Part II: Mitigation,” *IEEE Commun. Mag.*, vol. 35, no. 7, pp. 102–109, July 1997.

[14] M. Hata, “Empirical formula for propagation loss in land mobile radio services,” *IEEE Trans. Veh. Technol.*, vol. 29, no. 3, pp. 317–325, Aug. 1980.