Privacy and Security in the Internet of Things: Theory and Practice

Bob Baxley; bob@bastille.io
HitB; 28 May 2015

Bastille

Internet of Things (IoT)

Bastille


Bastille
THE PROBLEM
By 2020
50 BILLION DEVICES
NO SECURITY


Lte
WEIGHTLESS
GSM
Digi
ZigBee
Wi Fi
enocean
ISA100 Wireless
WirelessHART
Powered by sedona FRAMEWORK
NFC
Z-WAVE
DECT
DIGITAL DECT
Hundreds of Protocols
THREAD
Billions of Devices
SMART TV
Bluetooth
Bastille

OSI Stack

OSI (Open Source Interconnection) 7 Layer Model

Layer	Application/Example		Central Device/ Protocols			DOD4 Model
Application (7) Serves as the window for users and application processes to access the network services.	**End User layer** Program that opens what was sent or creates what is to be sent Resource sharing • Remote file access • Remote printer access • Directory services • Network management		**User Applications** SMTP	GATEWAY Can be used on all layers		Process
Presentation (6) Formats the data to be presented to the Application layer. It can be viewed as the "Translator" for the network.	**Syntax layer** encrypt & decrypt (if needed) Character code translation • Data conversion • Data compression • Data encryption • **Character Set Translation**		JPEG/ASCII EBDIC/TIFF/GIF PICT			
Session (5) Allows session establishment between processes running on different stations.	**Synch & send to ports** (logical ports) Session establishment, maintenance and termination • Session support - perform security, name recognition, logging, etc.		**Logical Ports** RPC/SQL/NFS NetBIOS names			
Transport (4) Ensures that messages are delivered error-free, in sequence, and with no losses or duplications.	**TCP** Host to Host, Flow Control Message segmentation • Message acknowledgement • Message traffic control • Session multiplexing	PACKET FILTERING	TCP/SPX/UDP			Host to Host
Network (3) Controls the operations of the subnet, deciding which physical path the data takes.	**Packets** ("letter", contains IP address) Routing • Subnet traffic control • Frame fragmentation • Logical-physical address mapping • Subnet usage accounting		**Routers** IP/IPX/ICMP			Internet
Data Link (2) Provides error-free transfer of data frames from one node to another over the Physical layer.	**Frames** ("envelopes", contains MAC address) [NIC card — Switch — NIC card] (end to end) Establishes & terminates the logical link between nodes • Frame traffic control • Frame sequencing • Frame acknowledgment • Frame delimiting • Frame error checking • Media access control		**Switch Bridge WAP** PPP/SLIP	Land Based Layers		Network
Physical (1) Concerned with the transmission and reception of the unstructured raw bit stream over the physical medium.	**Physical structure** Cables, hubs, etc. Data Encoding • Physical medium attachment • Transmission technique - Baseband or Broadband • Physical medium transmission Bits & Volts		**Hub**			

Image source

Classical security mechanisms

Application-level access control

SSL; String checking

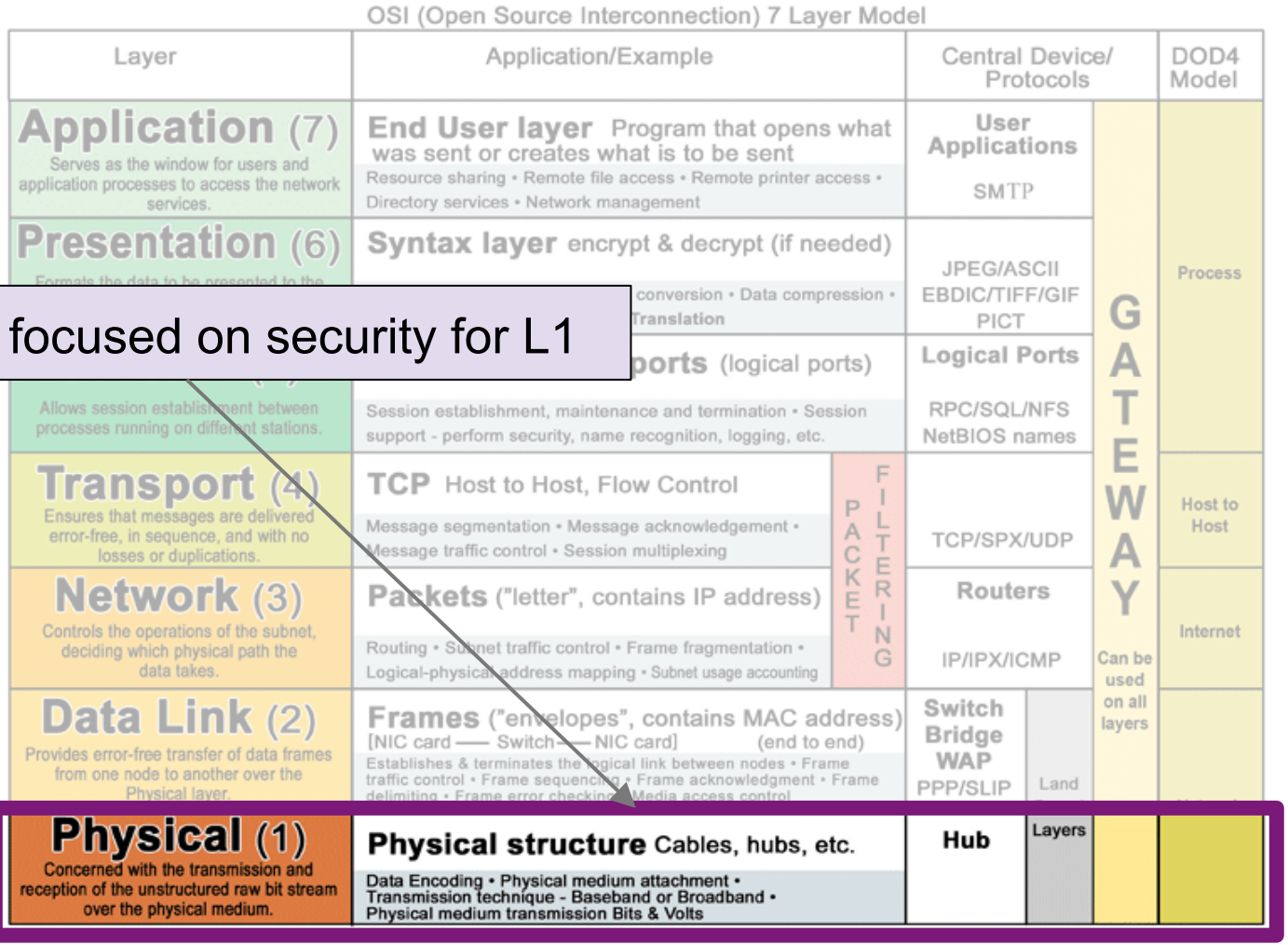
Encrypted password exchange

Firewalls

Route policy controls

MAC address filtering

Proximity physical security



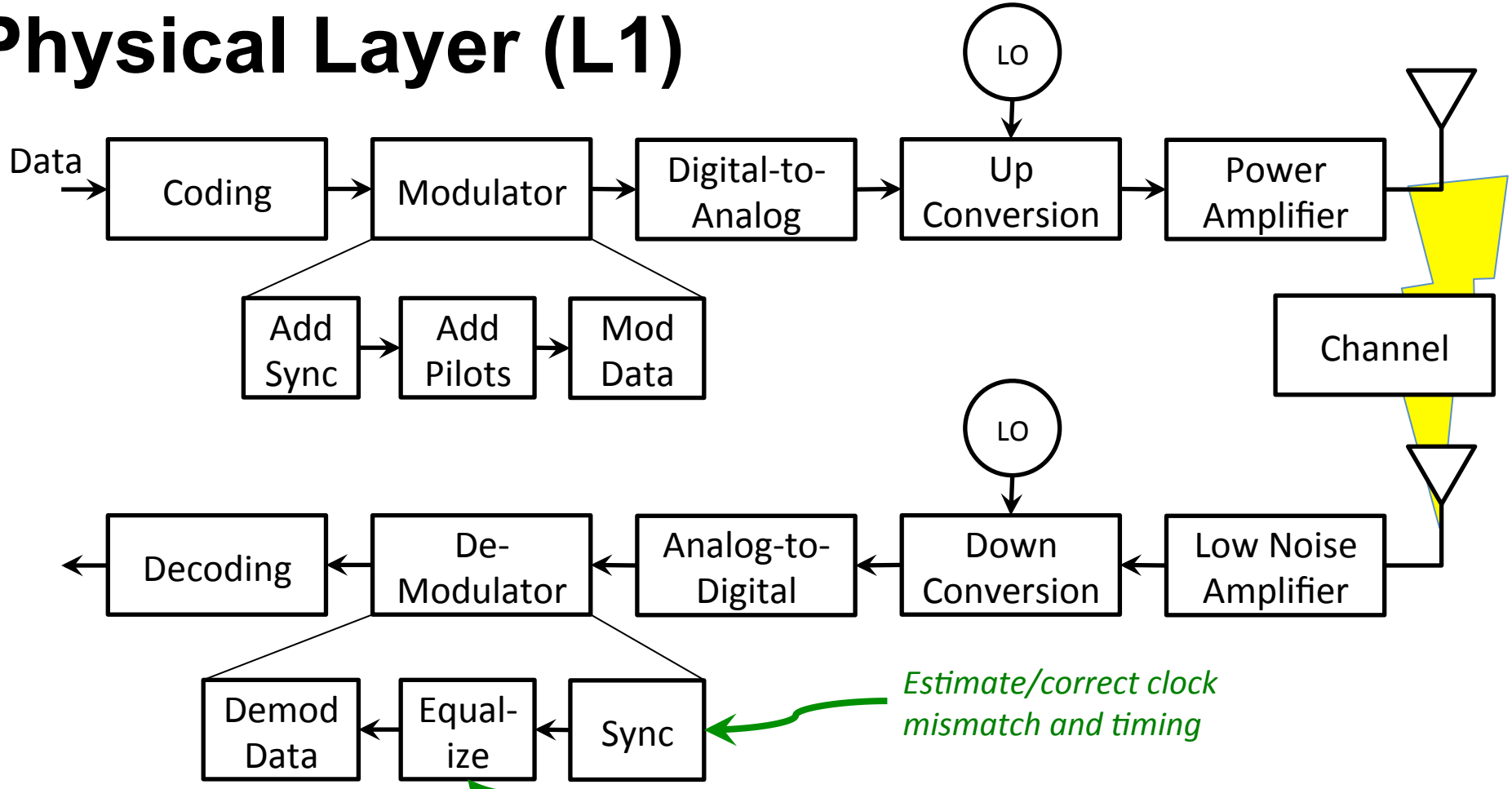
Bastille

Image source

Wireless

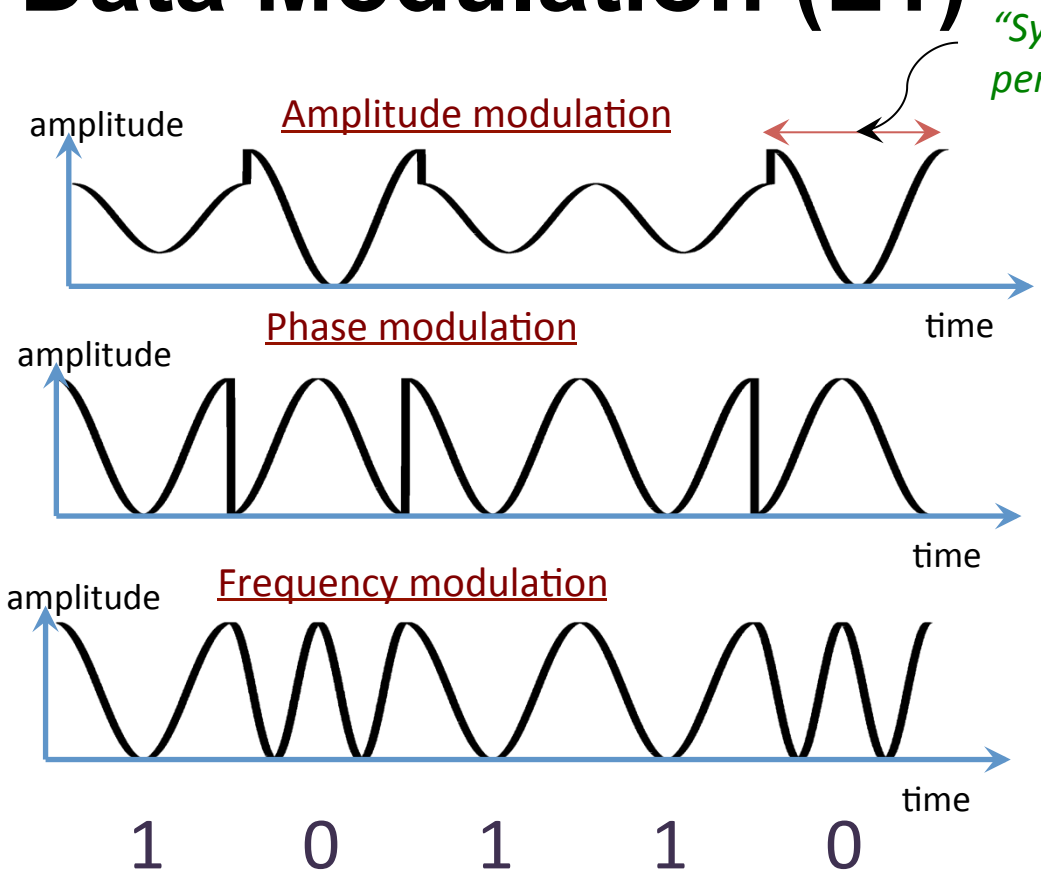
Bastille

Physical Layer (L1)

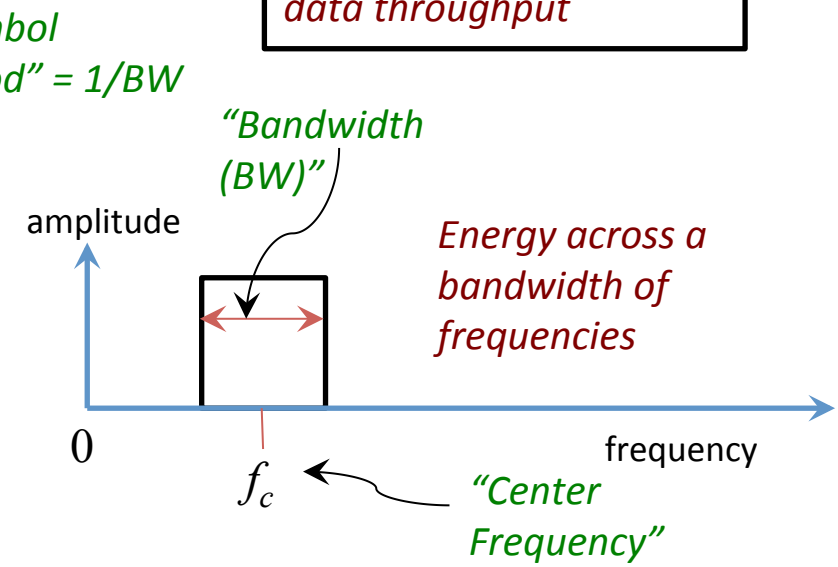


Bastille

Data Modulation (L1)



More bandwidth = more data throughput



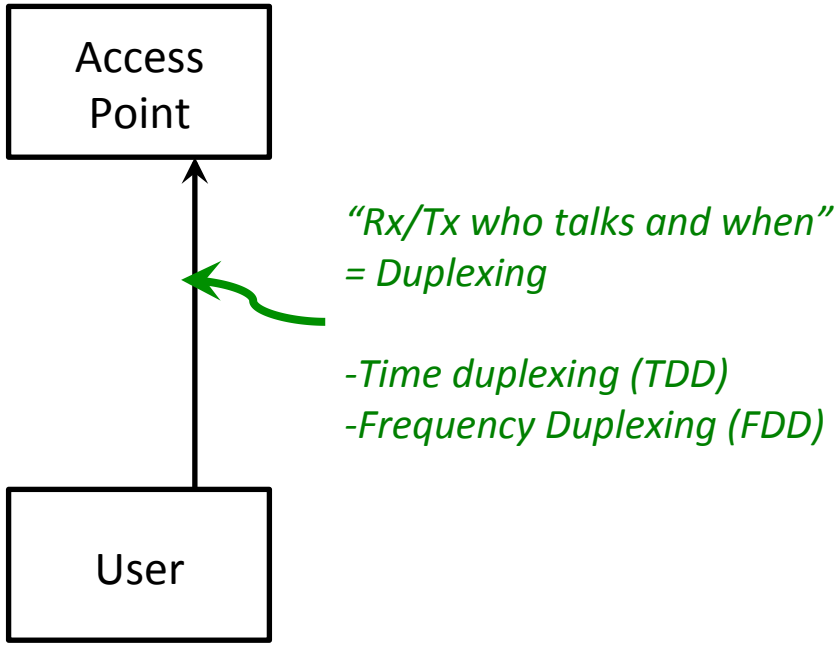
LTE (US): {5,10,20}MHz
Uplink is single carrier (phase+amp mod)
Downlink is OFDM (frequency)

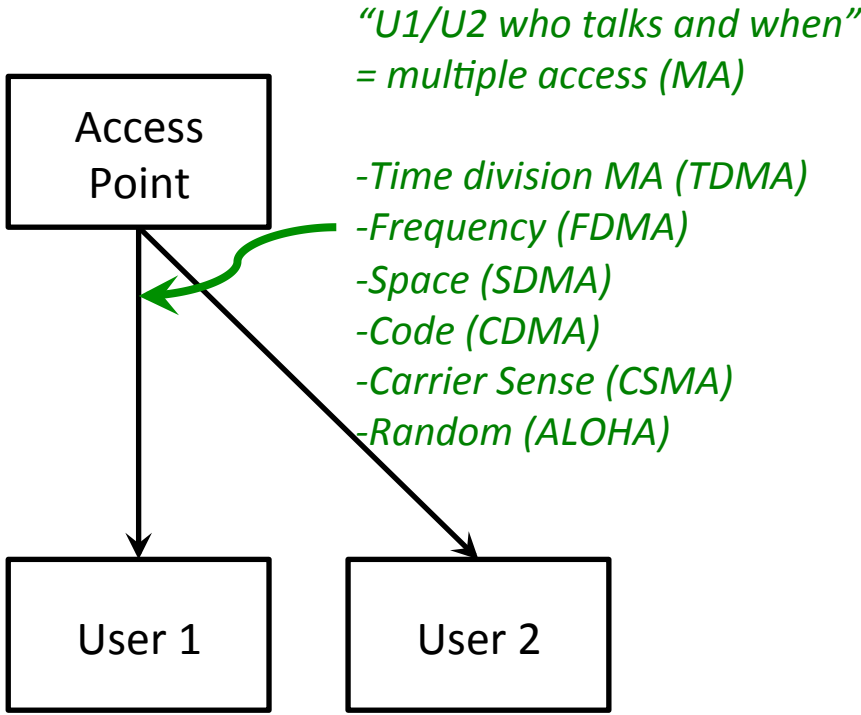
Wifi: OFDM – 20MHz channels

Bluetooth: Frequency shift keying (GFSK) – 1MHz channels

Bastille

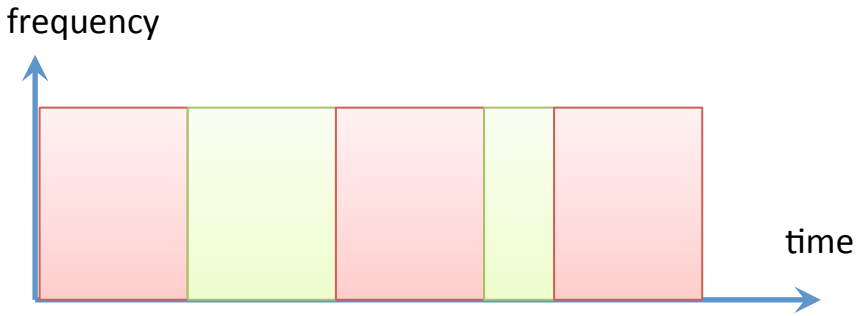
MAC in Wireless Comms (L2)





Medium Options (L2)

Time Division



Frequency Division

LTE (US): FDD
Up link is FDMA
Downlink is T/FDMA

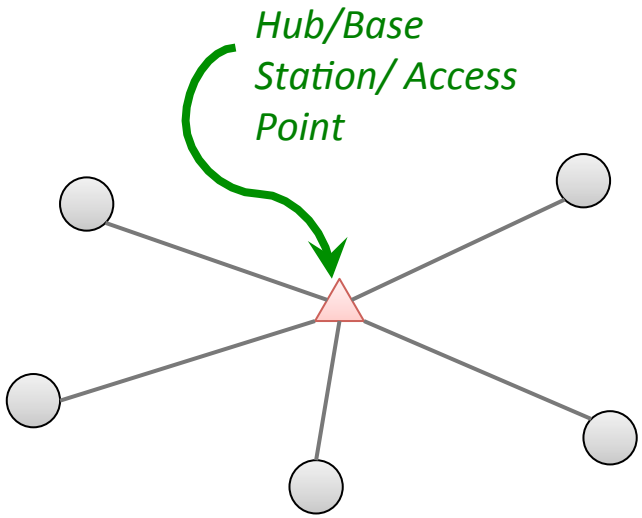
Wifi: TDD/CSMA

Bluetooth: TDD/FHSS-MA

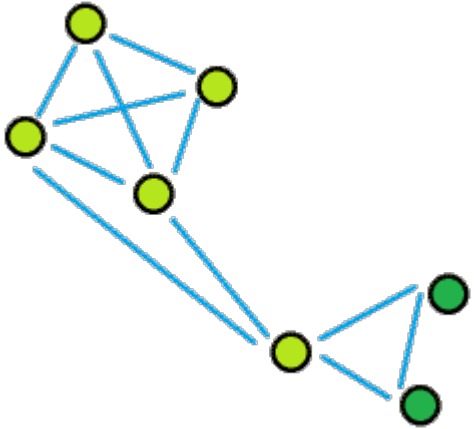
Bastille

Wireless Networks (L3)

Hub & Spoke: Brittle but simple



Mesh: robust but complex



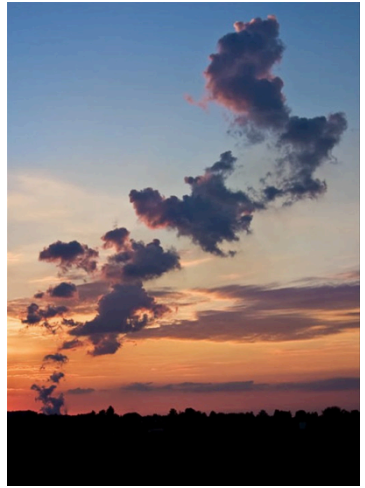
Attacks by Layer for Wireless

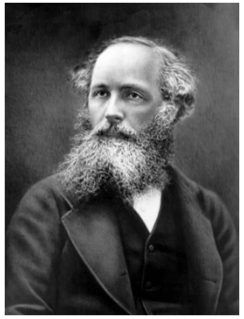
- **Layer 1**
  - Interference (jamming)
  - Impersonation
- **Layer 2**
  - Protocol cheating (not carrier sensing, not waiting between channel access)
  - Impersonation
  - False advertisement (changing identity, capabilities, etc.)
  - Wireless De-Auths
  - Promiscuous beaconing
- **Layer 3**
  - Protocol cheating (Sybil attack, wormholes/sinkholes, badmouthing, selective forwarding)
  - Impersonation

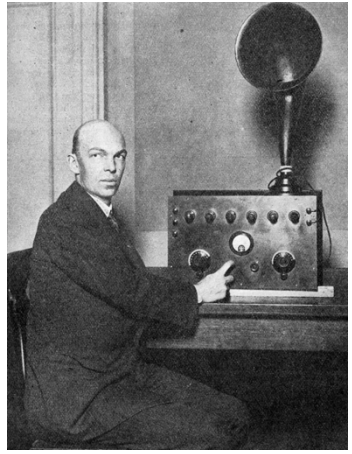
Information Theory

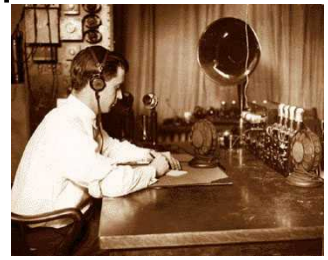
Brief History of Wireless Communications

http://wireless.ece.ufl.edu/jshea/HistoryOfWirelessCommunication.html









Long time ago: Optical (smoke signals)

1867: Maxwell predicts EM

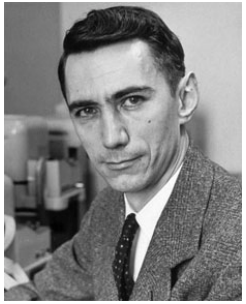
1887: Hertz proves EM waves exist

1896: Marconi demonstrates Wireless Telegraph

1925: Detroit starts using Radio for dispatch

1935: Armstrong demonstrates FM

Brief History of Wireless Communications Theory



1928: Nyquist showed we don't lose information by sampling. But the values of the signaling was still continuous.

1948: Shannon creates theory of information quantization--Bits!

1959: BCH and Reed-Solomon channel codes (error correction)

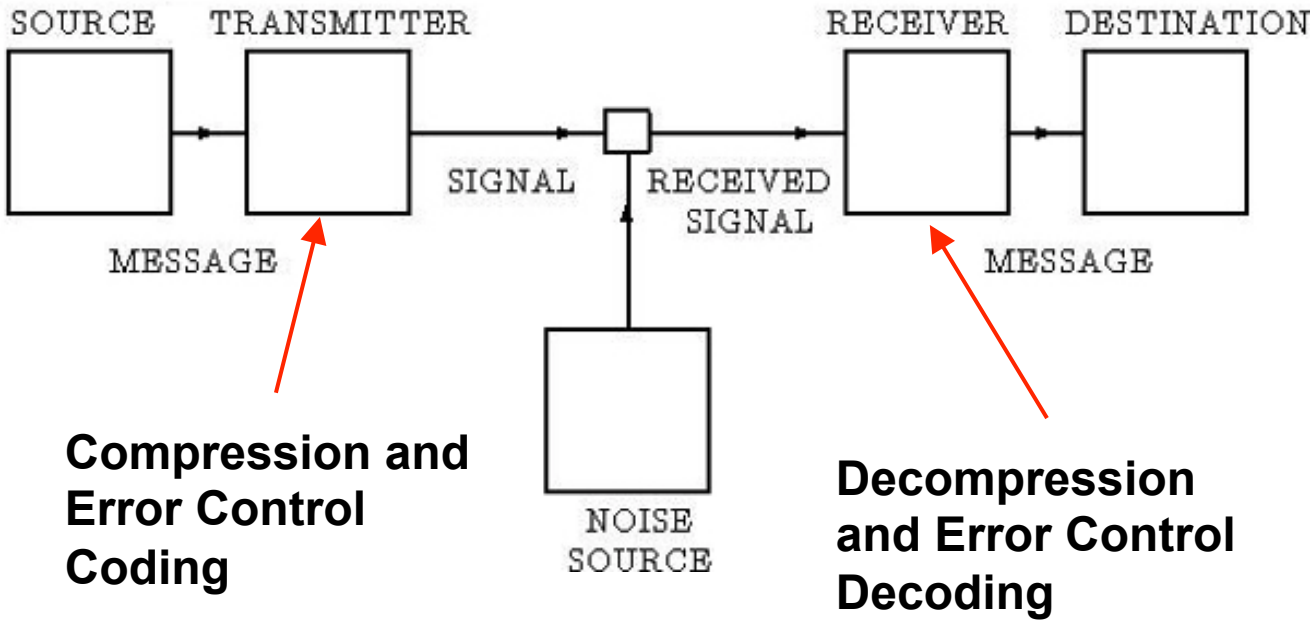
1977: Lempel-Ziv source coding (compression)

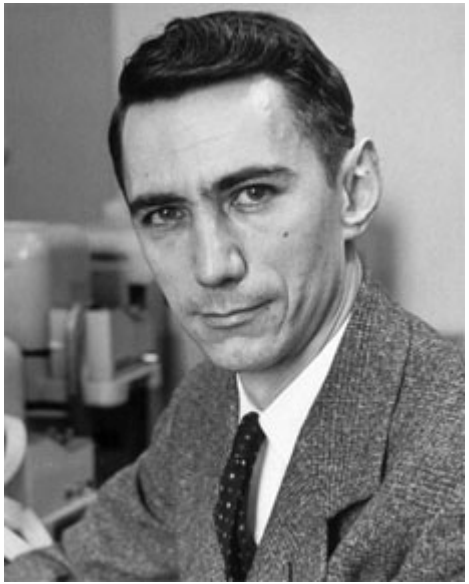
1993: Turbo codes; followed shortly by LDPC code rediscovery

Pre-Shannon

- Things were analog
- Communications systems were special purpose
- No comprehensive theory to dictate performance limits

Shannon's System



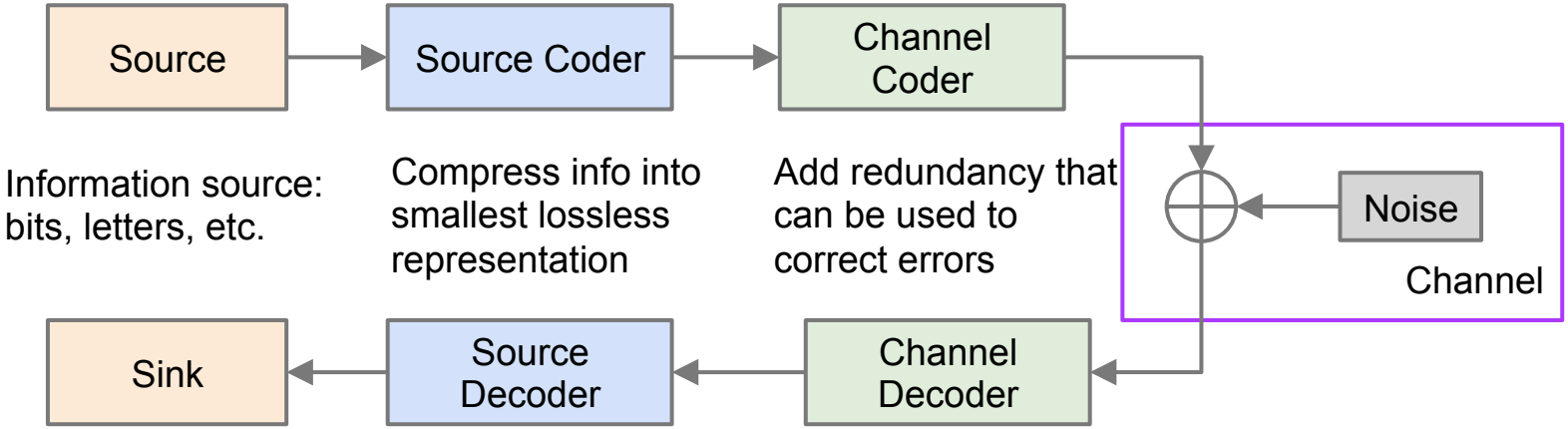


Compression and Error Control Coding

Decompression and Error Control Decoding

Bastille

Source Channel Separation

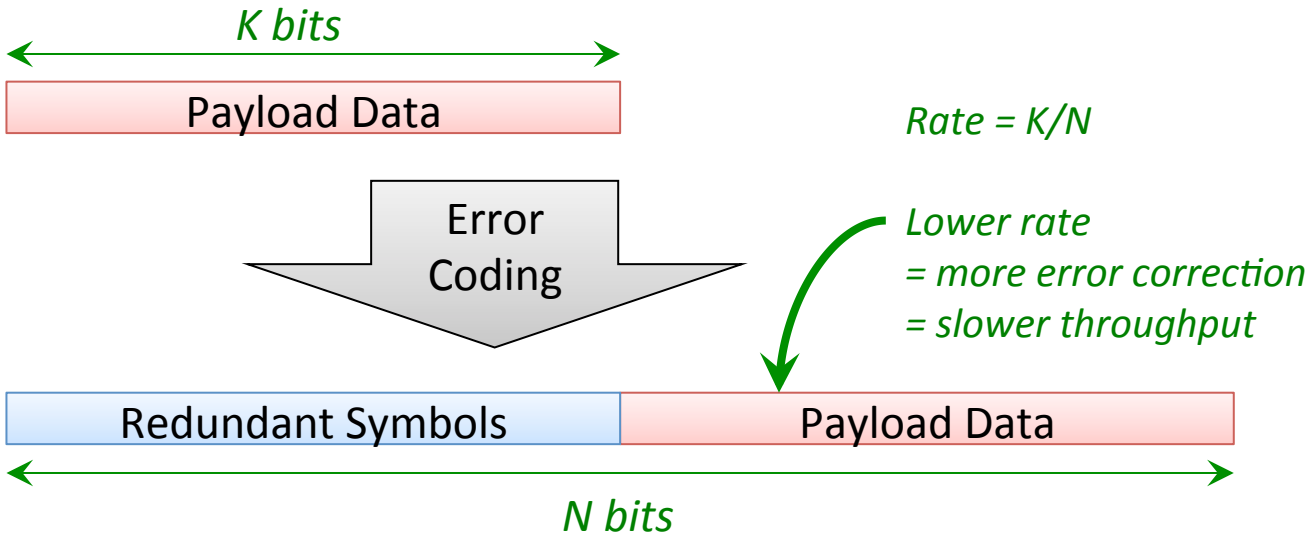


Source-Channel Separation Theorem: No performance loss by independently designing the source code and channel code.

This is a very nice result:

- Compression theorists can design compression algorithms without worrying about the channel.
- Error coding theorist can design codes assuming maximal entropy sources (i.e. compressed source)

Redundancy and Error Correction



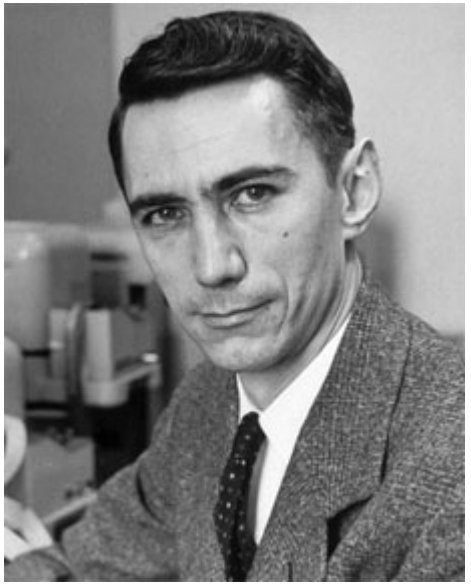
Is it possible to communicate with zero errors using a non-zero rate?

Big Result: Channel Capacity

Pre-Shannon: In order to have zero errors, we need infinite redundancy (R -> 0)

Shannon (1948): We can achieve zero errors with finite redundancy (R != 0) . The required rate for zero errors is called the "*channel capacity*"

Turbo codes and LDPC codes ~2003 were first to achieve capacity



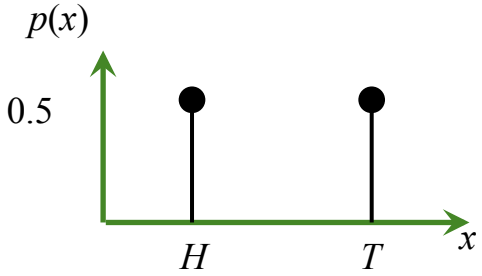
Max error free rate
= Channel Capacity
= BW*log(1+ SNR)

Information Theory Building Block: Entropy

Entropy is a measure of the average “randomness” or “uncertainty” of a random variable

If the random variable has a distribution p(x), then entropy is defined by

$$H(x) = -\sum_{x} p(x)\log_2(p(x))$$



Coin flip example: $x \in \{H, T\}$

Bastille

Entropy Properties

Entropy is measured in bits (or nats)

Source Coding Theorem: It is impossible to compress data to less than the entropy of the source

Entropy of a source with M states is at most $\log_2(M)$; this value is reached when the M states are equally likely

Example: English; 26 states -> max H(X) = 4.7 bits
Estimates of English entropy are ~1.3 bits per character

3.4 bits of redundancy per character in English!

From McKay, 2003

i	a_i	p_i
1	a	0.0575
2	b	0.0128
3	c	0.0263
4	d	0.0285
5	e	0.0913
6	f	0.0173
7	g	0.0133
8	h	0.0313
9	i	0.0599
10	j	0.0006
11	k	0.0084
12	l	0.0335
13	m	0.0235
14	n	0.0596
15	o	0.0689
16	p	0.0192
17	q	0.0008
18	r	0.0508
19	s	0.0567
20	t	0.0706
21	u	0.0334
22	v	0.0069
23	w	0.0119
24	x	0.0073
25	y	0.0164
26	z	0.0007
27	–	0.1928

Bastille

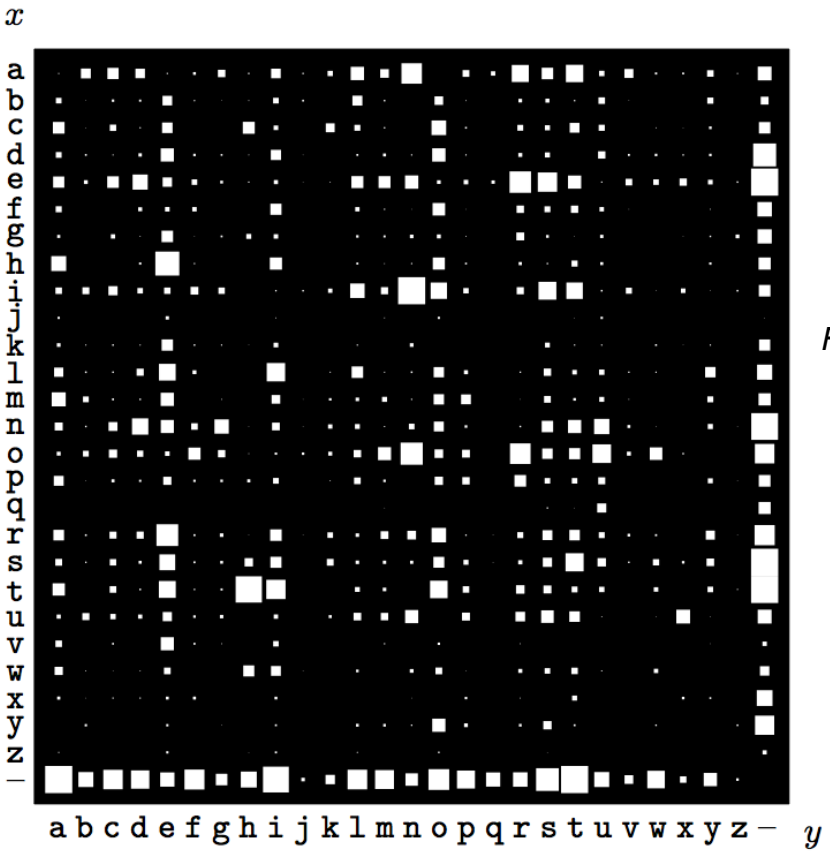
Entropy and Language

Language	Original Size (bytes)	Ratio (Original)	Compressed Size (bytes)	Ratio (Compressed)
English	1936473	1	363288	1
Spanish	1804756	0.932	360535	0.992
French	1896459	0.979	359903	0.991
Chinese	884860	0.457	341850	0.941
Korean	1259920	0.651	352440	0.970
Arabic	1875204	0.968	395242	1.09
Japanese	1519224	0.785	438135	1.206
Russian	1506920	0.778	362207	0.997

Table 3: Results for the Bible using the PPMZ compression algorithm

From: http://www.liafa.univ-paris-diderot.fr/~dxiao/docs/entropy.pdf

Entropy and Language II

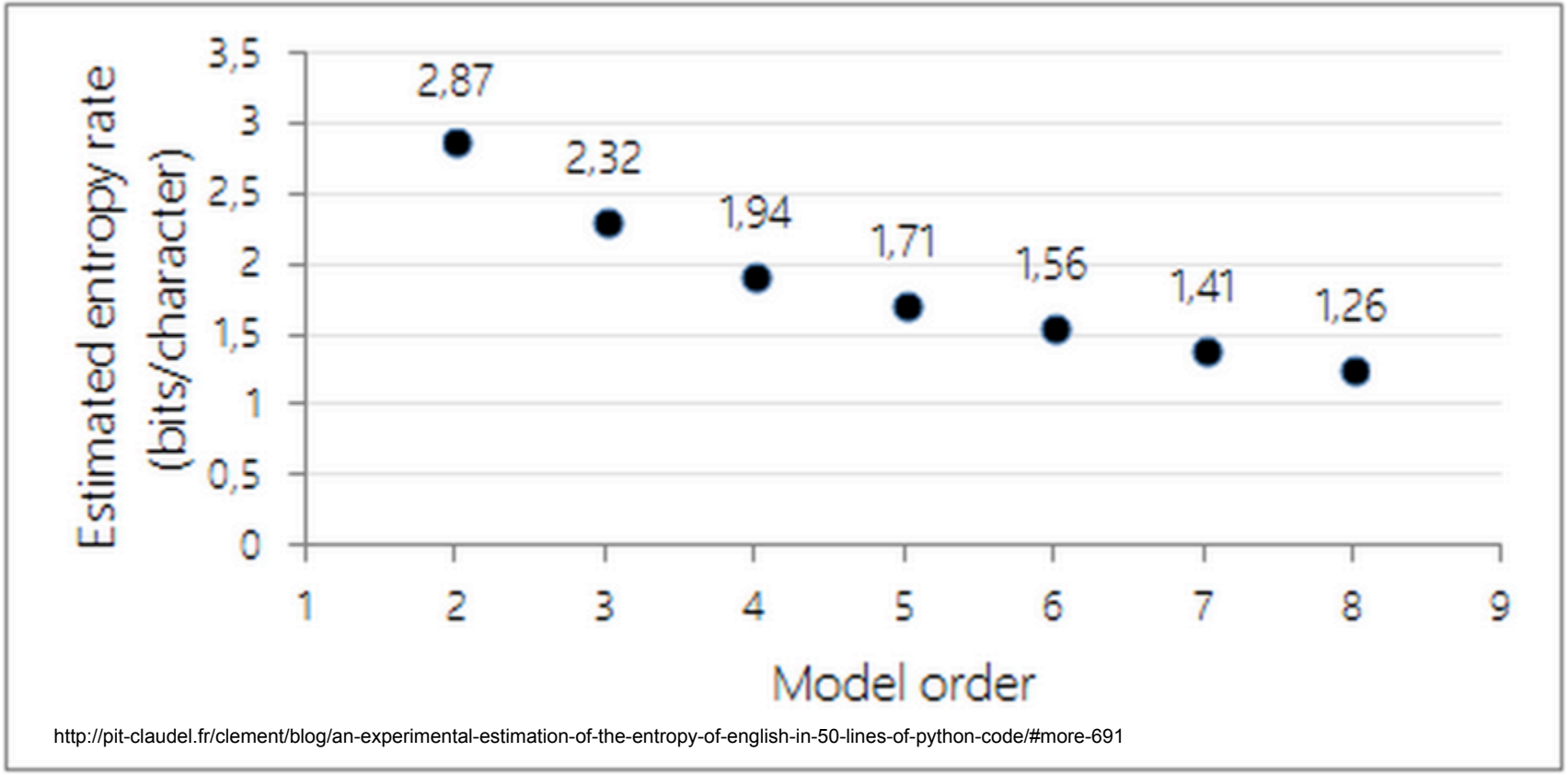


From McKay, 2003

i	a_i	p_i
1	a	0.0575
2	b	0.0128
3	c	0.0263
4	d	0.0285
5	e	0.0913
6	f	0.0173
7	g	0.0133
8	h	0.0313
9	i	0.0599
10	j	0.0006
11	k	0.0084
12	l	0.0335
13	m	0.0235
14	n	0.0596
15	o	0.0689
16	p	0.0192
17	q	0.0008
18	r	0.0508
19	s	0.0567
20	t	0.0706
21	u	0.0334
22	v	0.0069
23	w	0.0119
24	x	0.0073
25	y	0.0164
26	z	0.0007
27	–	0.1928

Bastille

English by Model Order



http://pit-claudel.fr/clement/blog/an-experimental-estimation-of-the-entropy-of-english-in-50-lines-of-python-code/#more-691

Information Theory Building Block: Mutual Information

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

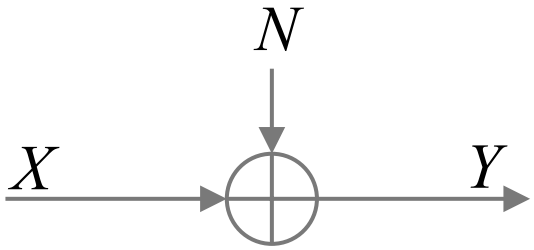
Mutual Information is the reduction in the uncertainty about X due to knowledge of Y

If knowing Y tells you a lot about X, then I(X;Y) is high

If Y is independent of X, then I(X;Y) = 0

The value of I(X;Y) is completely determined by the joint distribution of X and Y

Channel Capacity



Shannon showed that the maximum error-free rate of communications (the channel capacity) can be calculated by:

$$C = \max_{p(x)} I(X;Y)$$

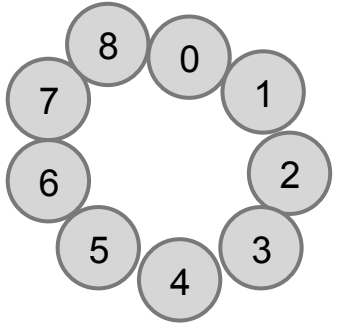
I.e. Pick codeword distribution that maximizes MI

Example: When N is Gaussian Noise, maximizing the mutual information requires that X have maximum entropy. The entropy maximizing distribution is Gaussian; so the optimal p(x) is Gaussian.

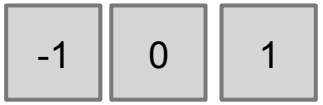
$$C_{AWGN} = B \log(1+SNR)$$

Example

Source Alphabet

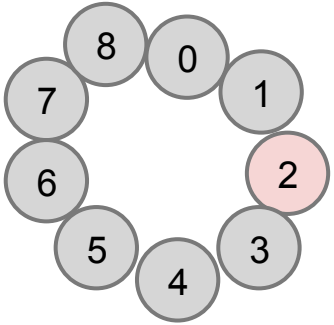


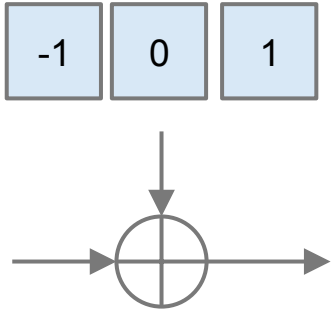
Noise Alphabet (mod 9)



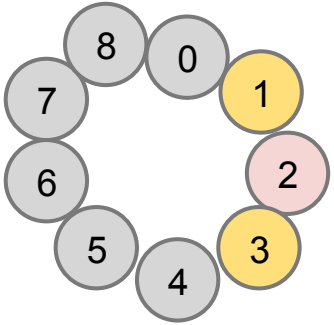
Trial Transmission:

Transmit “2”



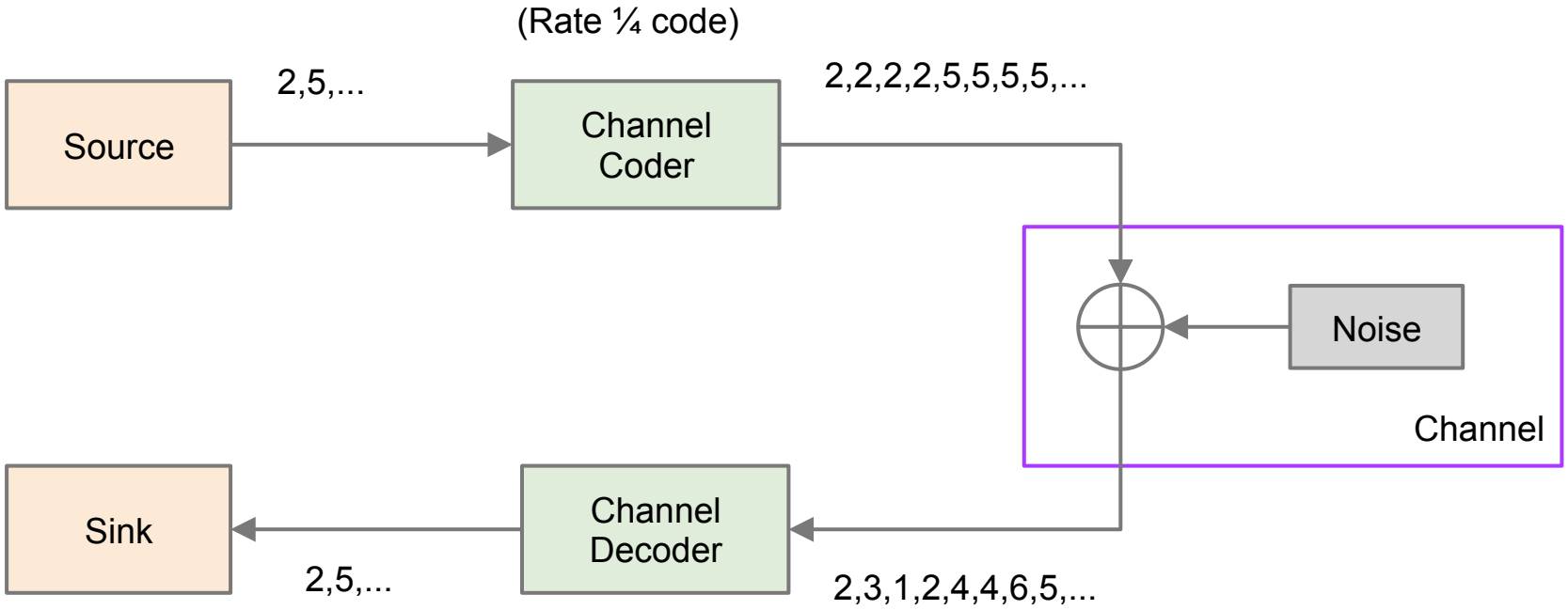


Receive one of {1,2,3}



Bastille

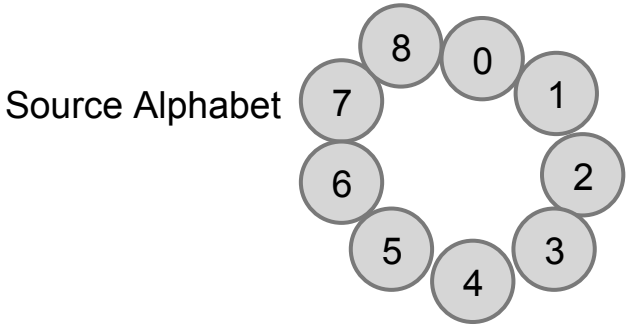
Repetition Code

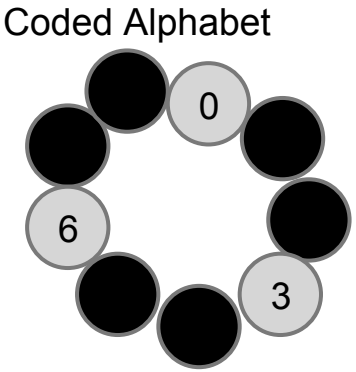


The problem: no matter the rate, there will be some probability that we will have a symbol error

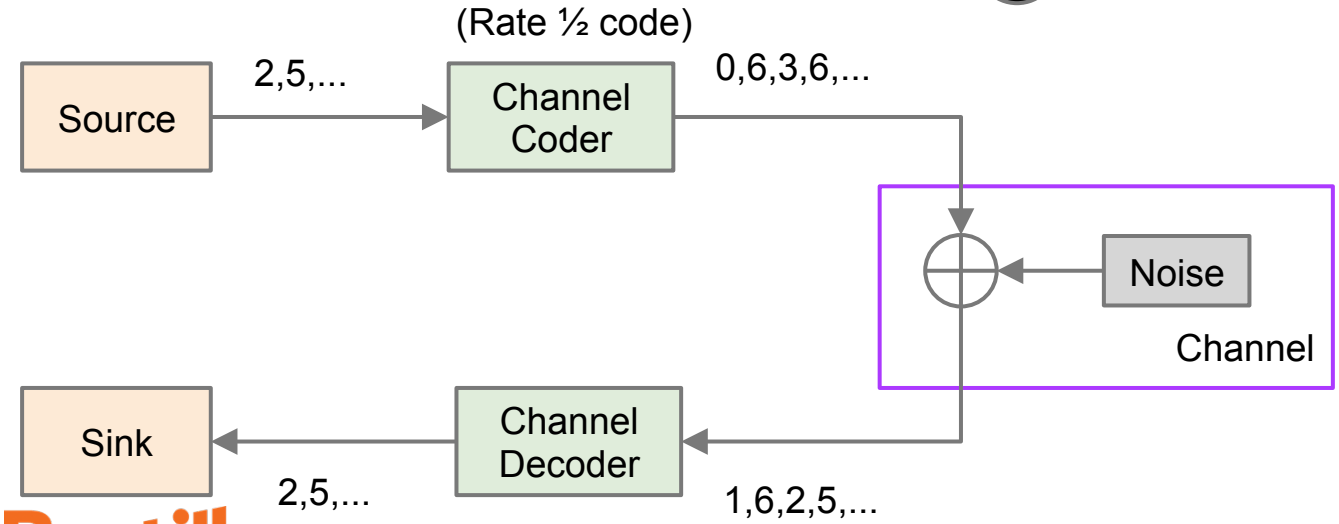
Bastille

Better Code





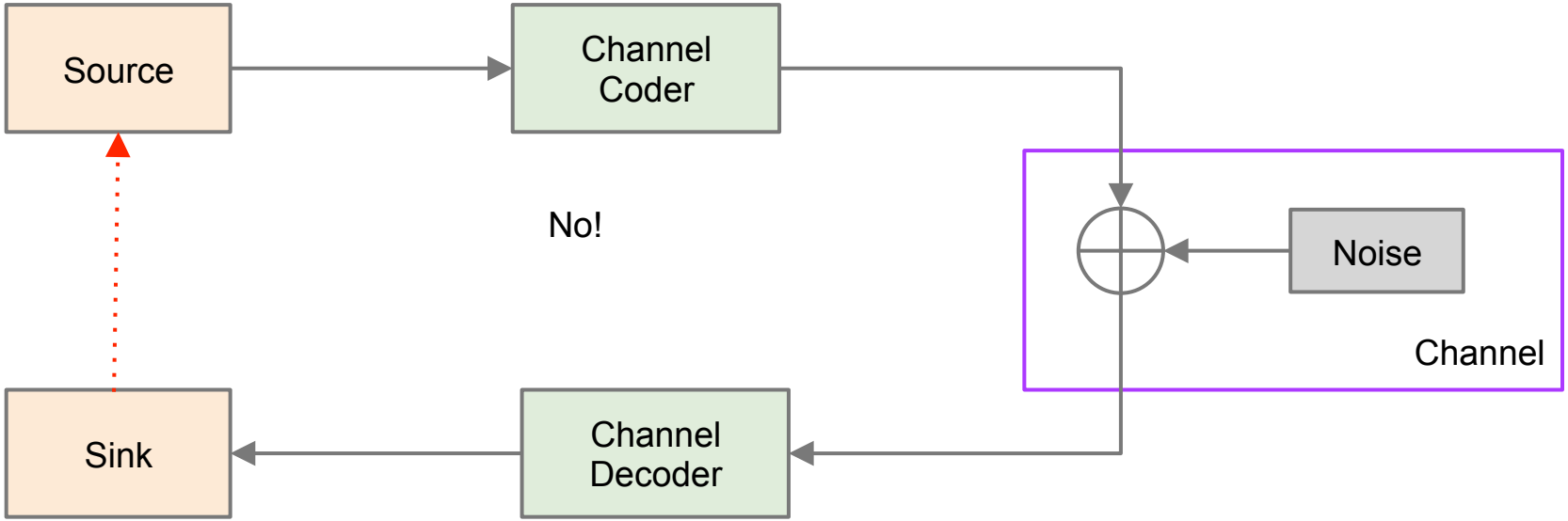
Source	Coded
0	00
1	03
2	06
3	30
4	33
5	36
6	60
7	63
8	66



Bastille

Neat Result

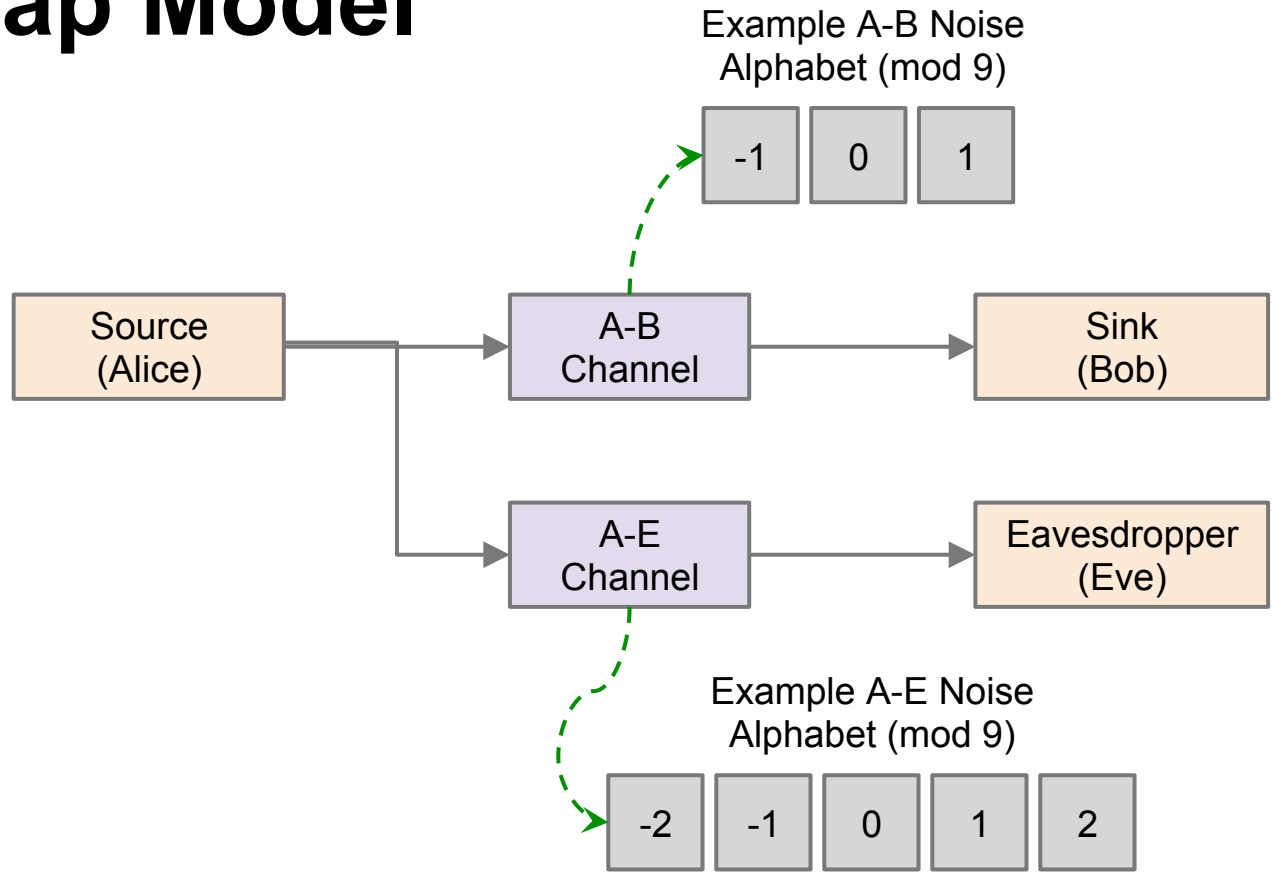
Does Feedback increase the capacity?



Physical-Layer Security

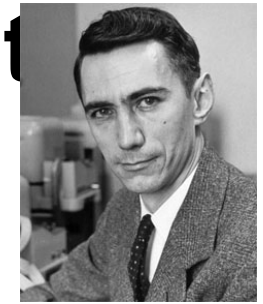
Bastille

Wiretap Model



Bastille

Brief History of Information Theoret Security



1948: Shannon presents concept of *perfect secrecy* using a one-time pad.

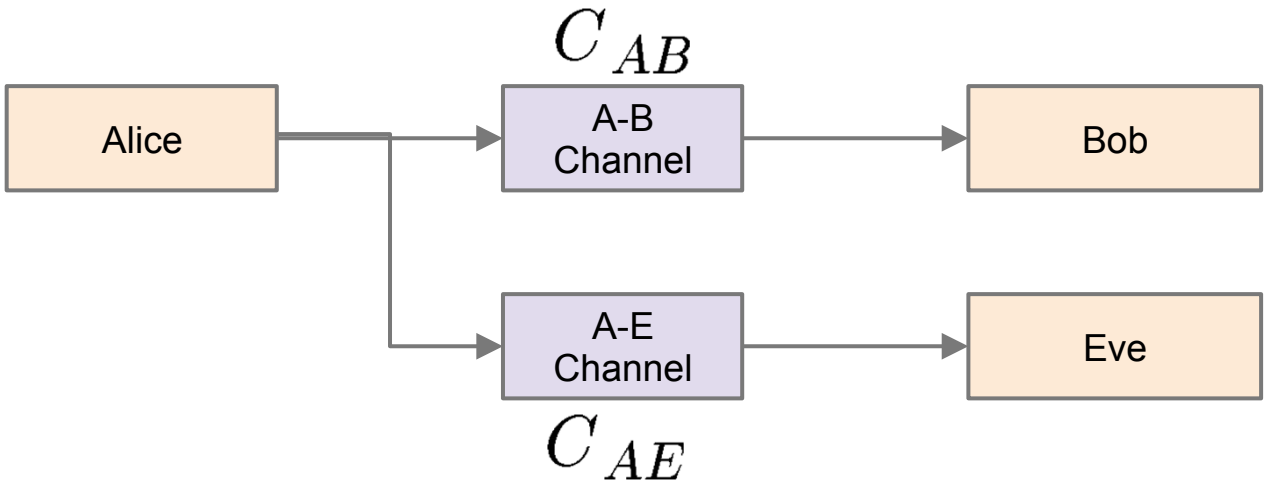
1975/1978: Wyner/Csiszár and Körner - channel codes exist that provide error robustness and confidentiality

1976: Diffie-Hellman public-key encryption

1993: Maurer: *Secret key agreement by public discussion from common information*

Secrecy Capacity a la Wyner/Csiszár and Körner

Recall, the capacity of a noisy channel is $C_{AWGN} = B \log(1+SNR)$

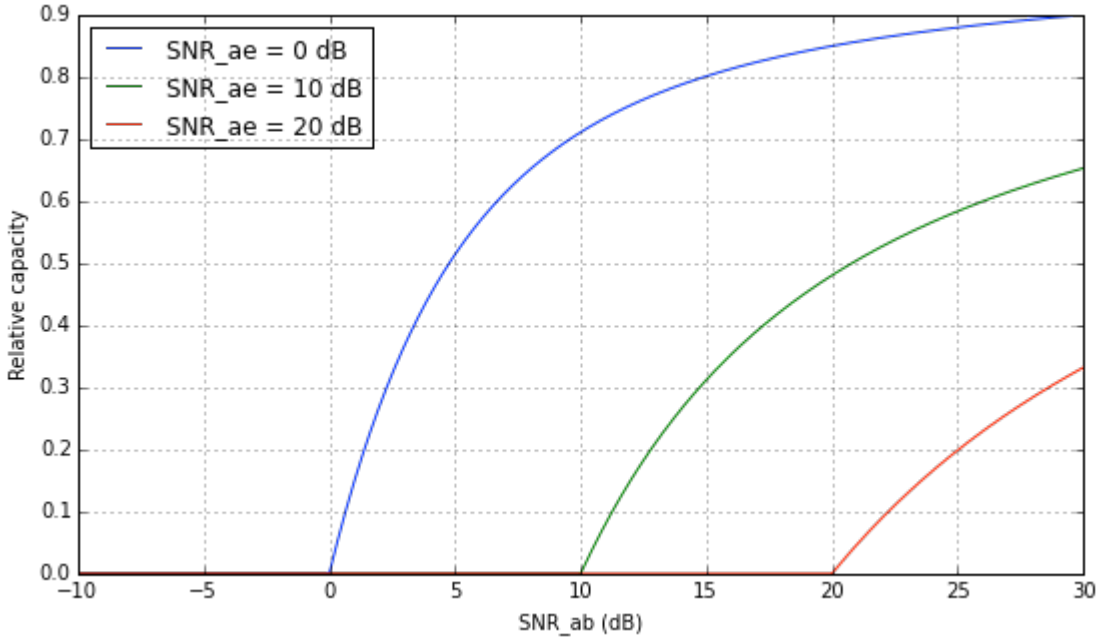


The *secrecy capacity* is simply the capacity advantage Bob has over Eve:

$$C_S = \max(0, C_{AB} - C_{AE})$$

Bastille

Secrecy Capacity



Limited applicability in practice:

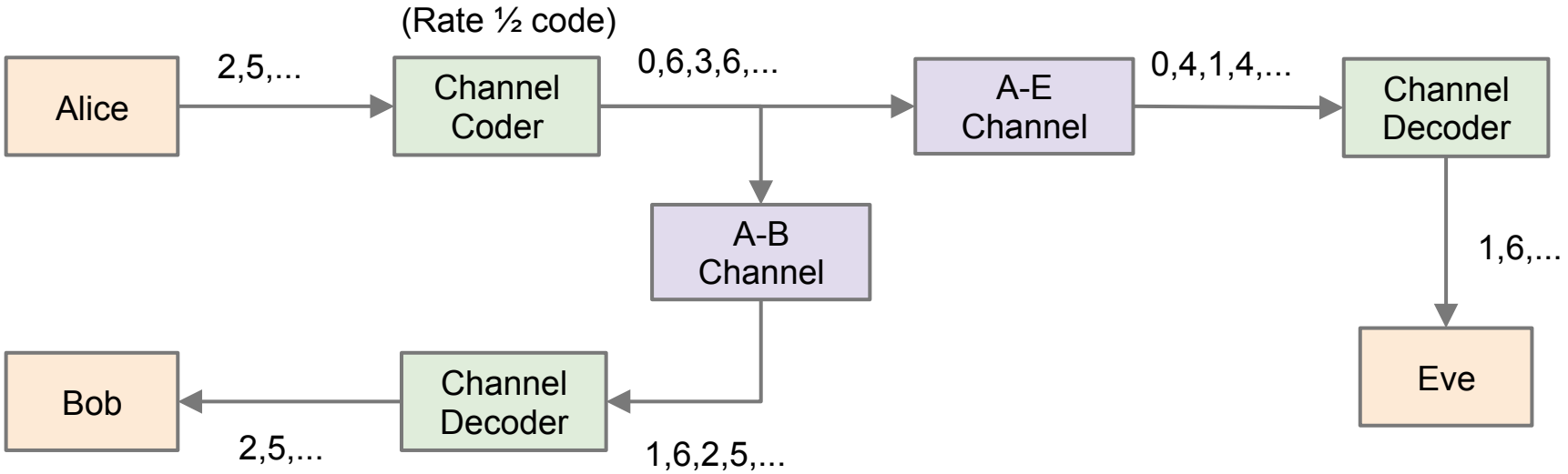
- Alice and Bob have to estimate Eve's SNR
- Then encode the data based on that estimate
- If they are wrong, then the secrecy no longer exists

Bastille

Plot reproduced from: Bloch, Matthieu, et al. "Wireless information-theoretic security." *Information Theory, IEEE Transactions on* 54.6 (2008): 2515-2534.

Take Away

We can make codes that “encrypt” and correct channel errors



Bastille

Physical-Layer Security with Channel Variations

Bastille

Propagation

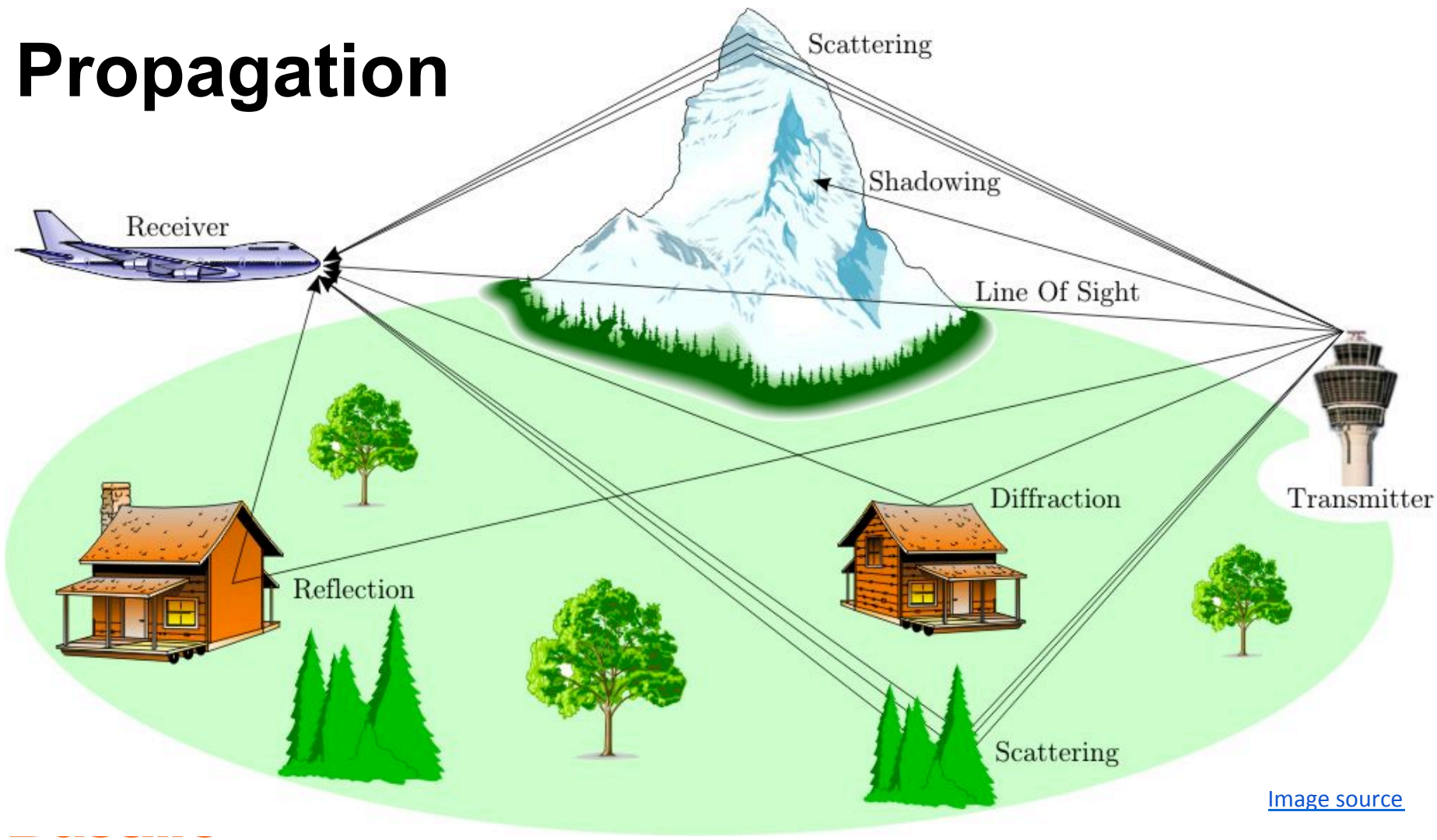
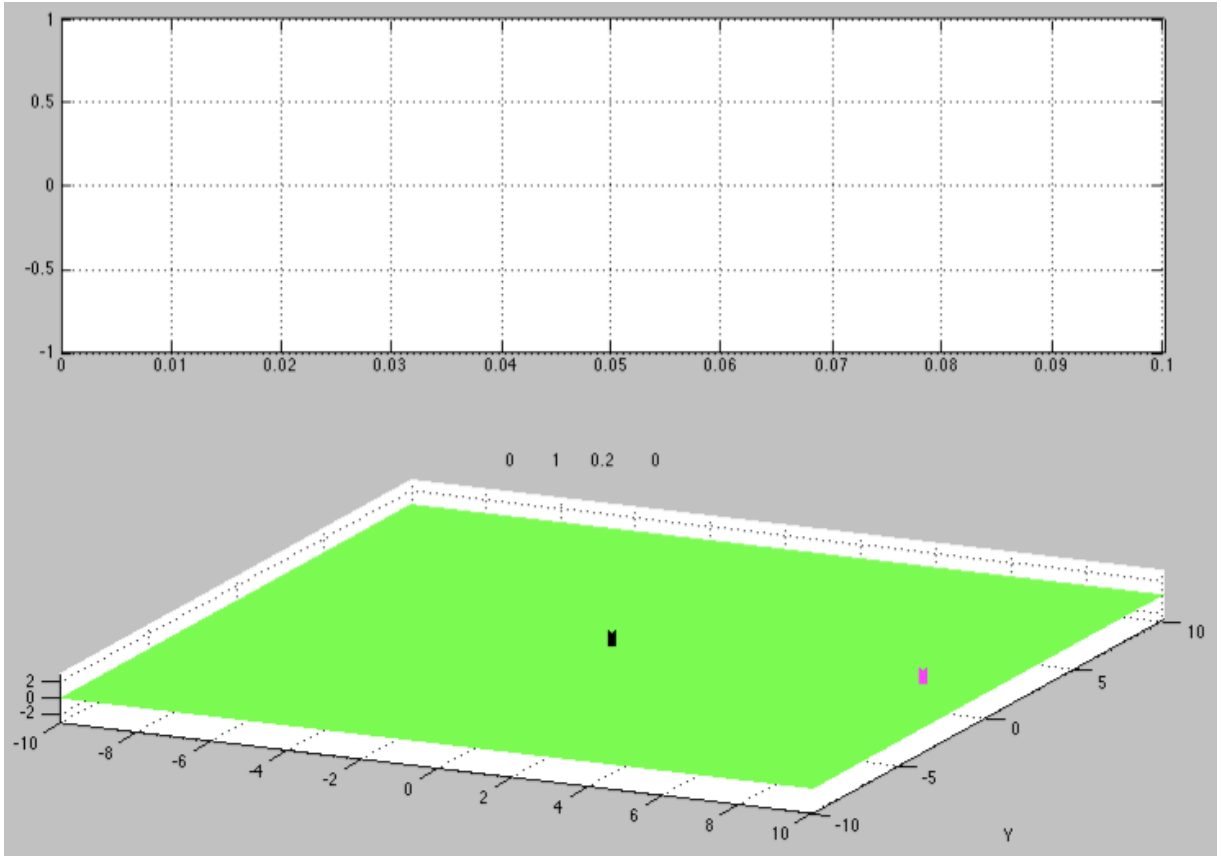


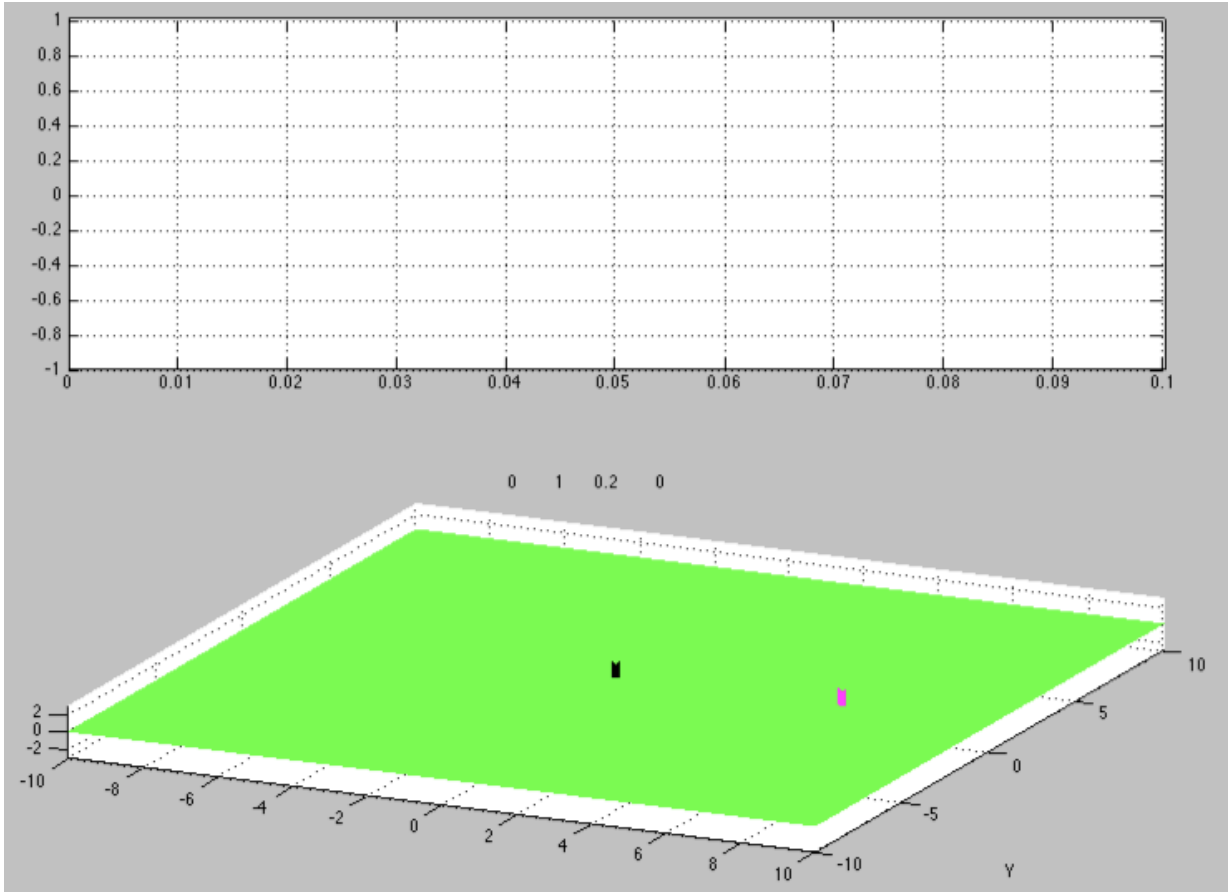
Image source

Constructive Channel Effect



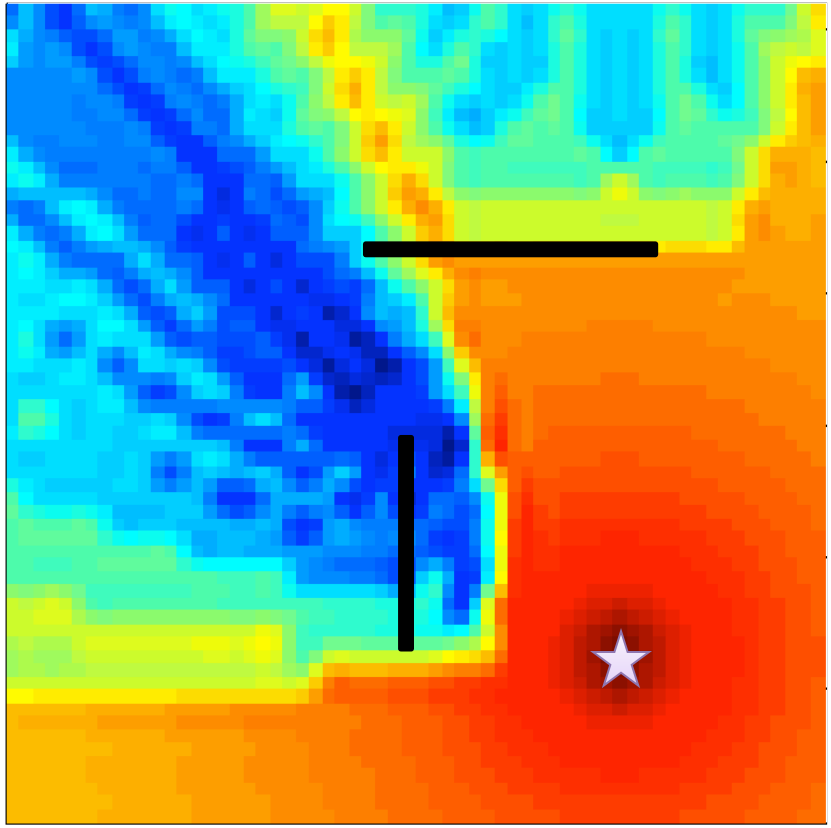
Bastille

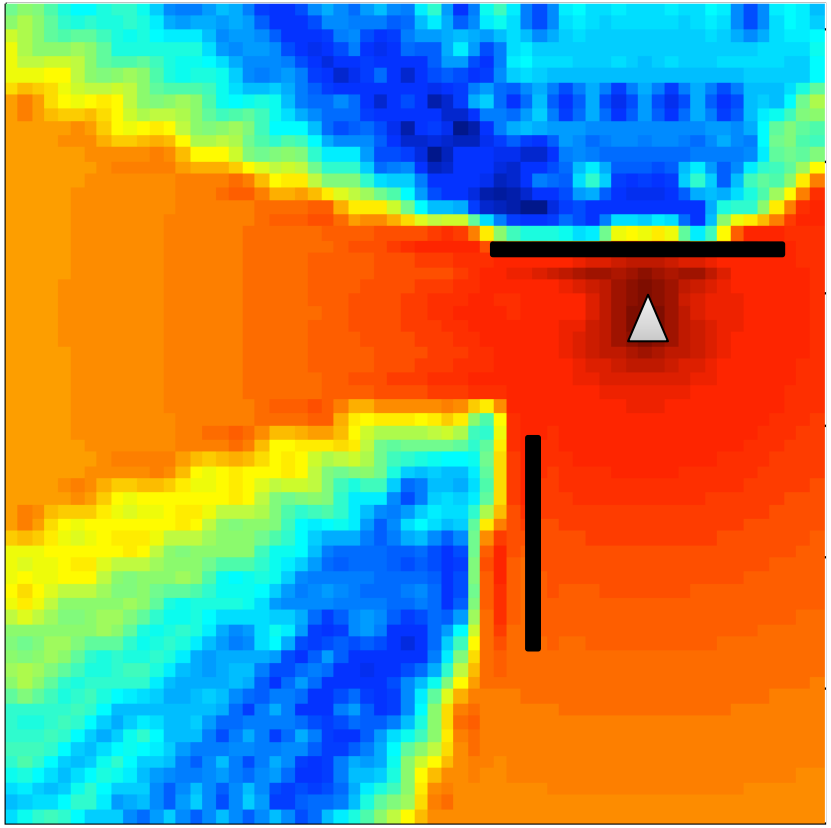
Destructive Channel Effect



Bastille

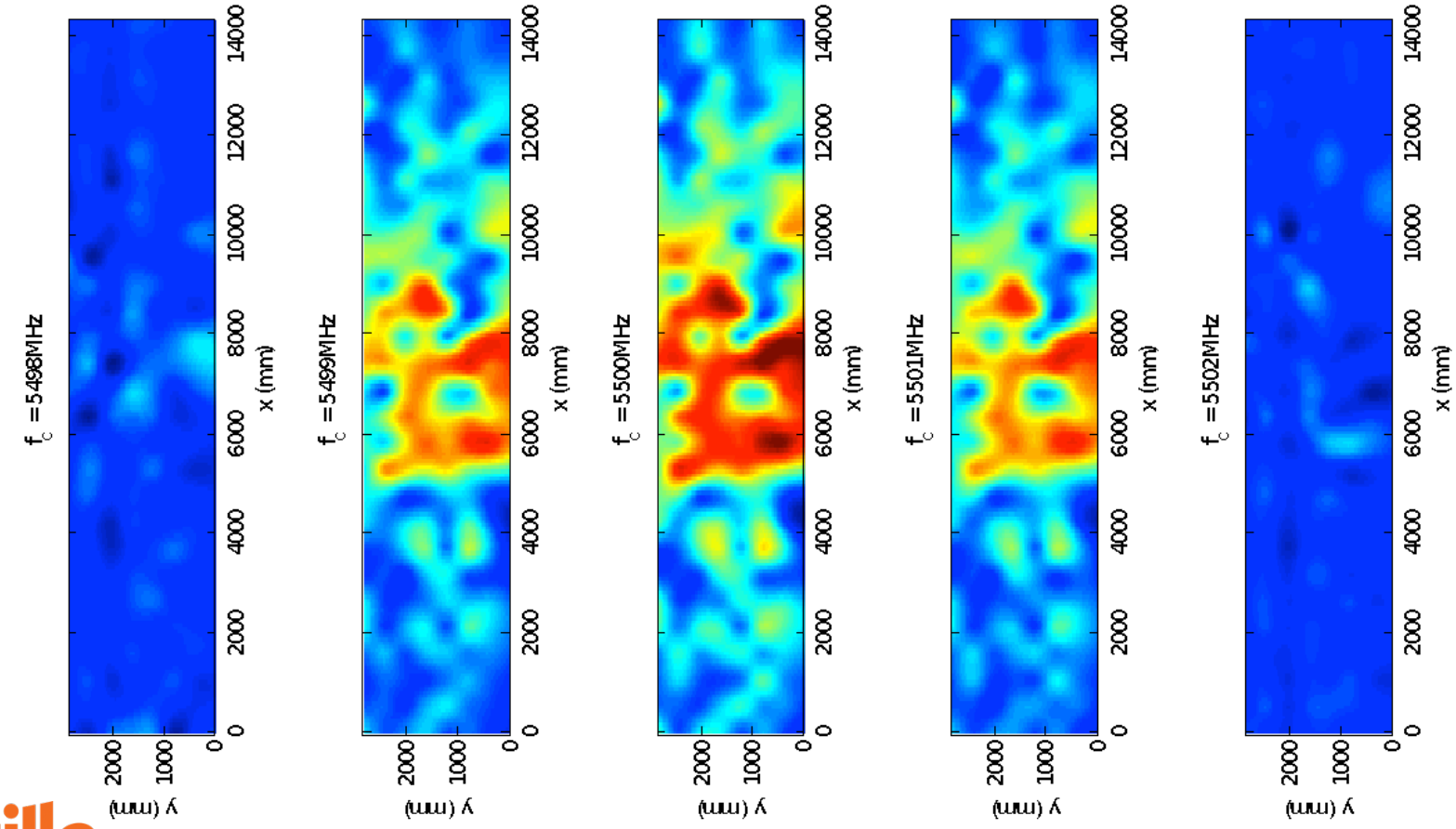
Shadowing Loss





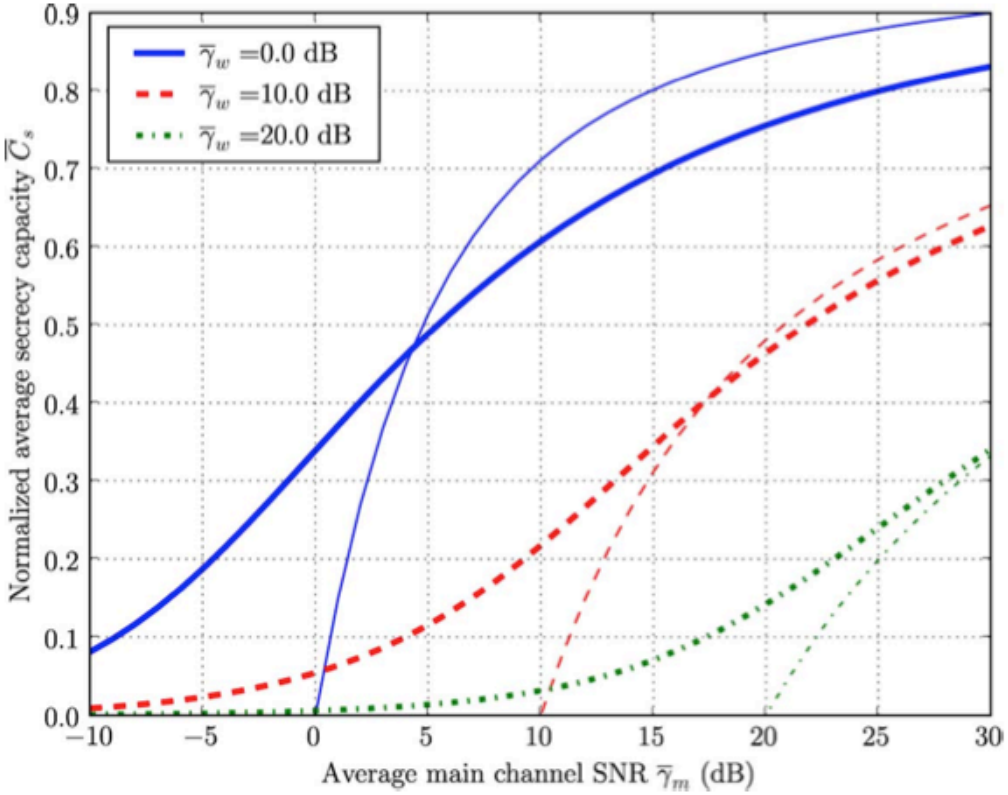
Bastille

Gain Map Example



Bastille

Secrecy Capacity in Fading Channels



In real channels positive secrecy rate is always possible!
(i.e. the thicker lines never go below zero)

Alice only transmits when she has an SNR advantage.

Bastille

Plot excerpted from: Bloch, Matthieu, et al. "Wireless information-theoretic security." *Information Theory, IEEE Transactions on* 54.6 (2008): 2515-2534.

Encryption Derived from Common Randomness

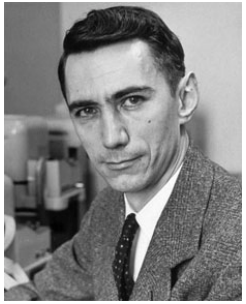
Bastille

One-Time Pad Encryption History

1882: Frank Miller described one-time pad for telegraphy

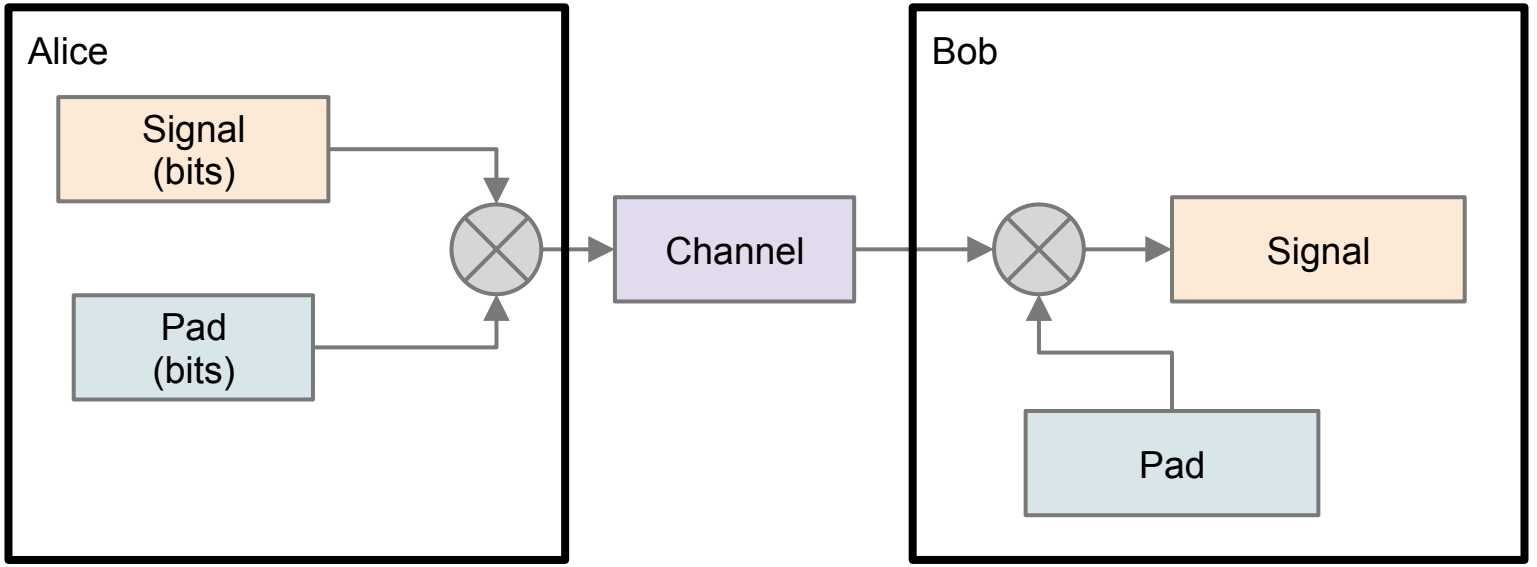
1917: Gilbert Vernam implemented and patented working system

1920s: Various implementations using dictionaries of one-time pads



1948: Claude Shannon proved that one-time pad achieves perfect secrecy

One-Time Pad



One-Time Pad Keys

- One large challenge in encryption is key exchange
- If an attack compromises the key-exchange side channel, the encryption is defeated

A source of common randomness that is unique to two communicators would be ideal!

Wireless Channels Gains are random and reciprocal--perfect!

1993: Maurer: *Secret key agreement by public discussion from common information*

Propagation

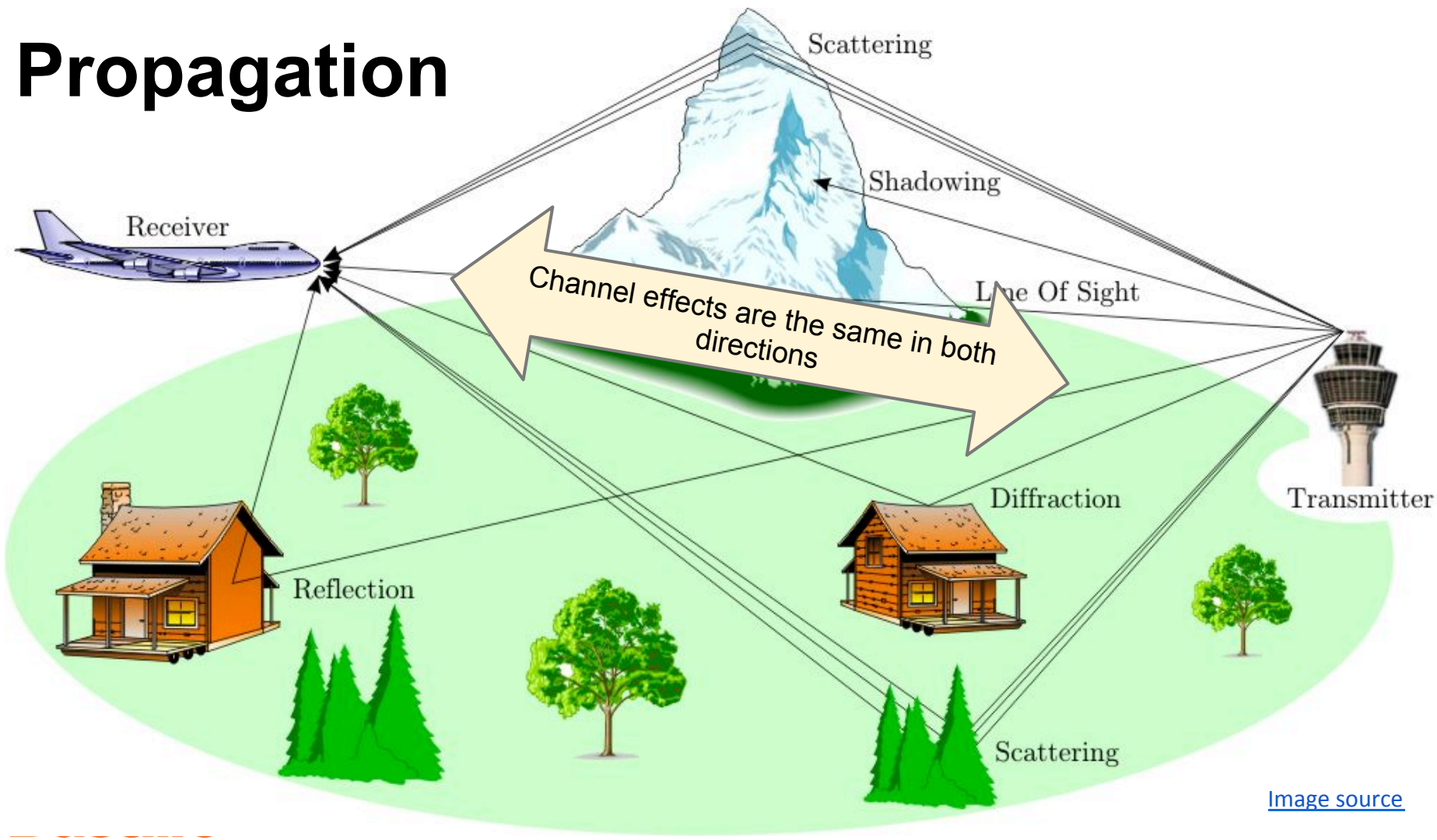
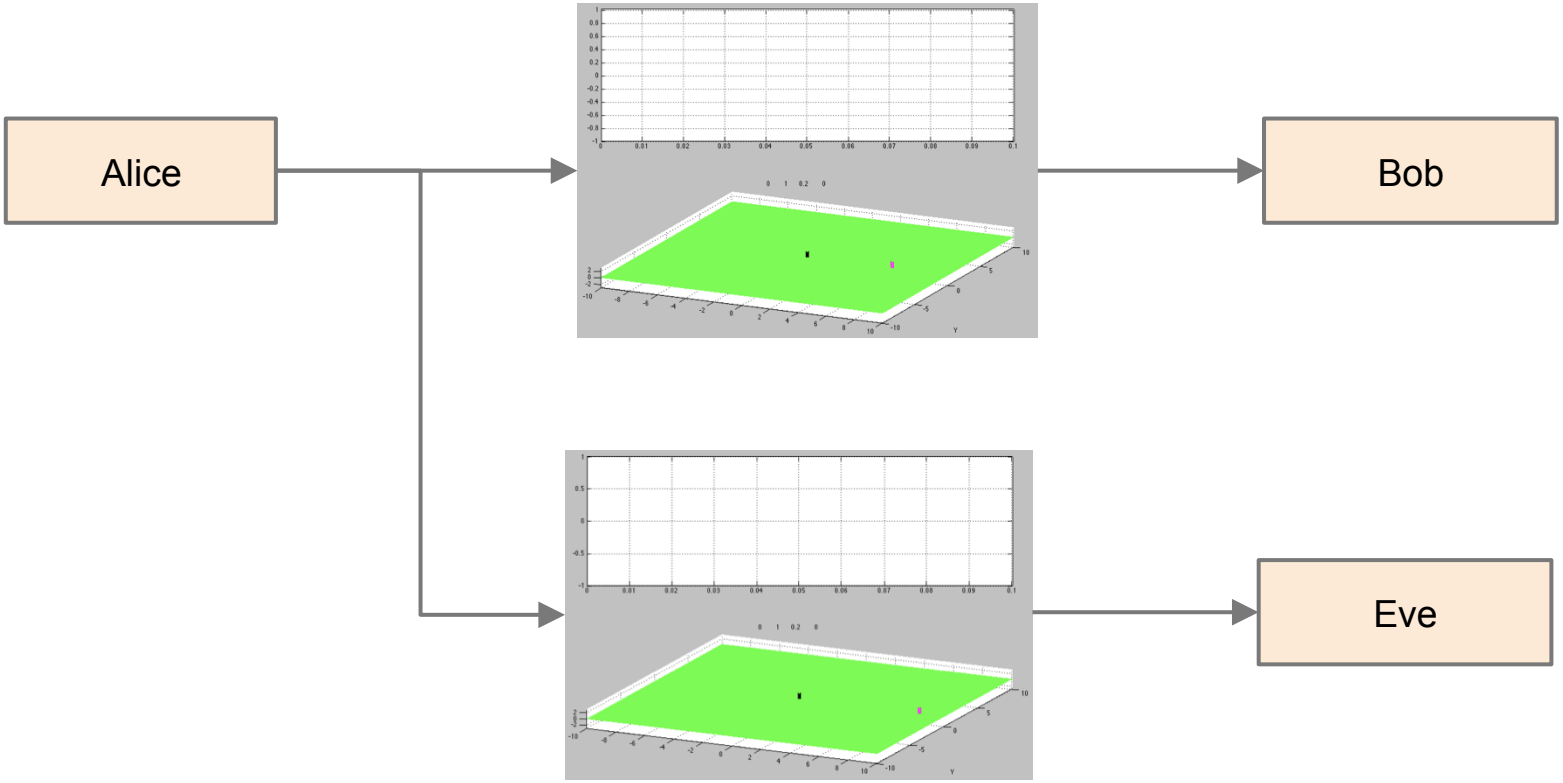


Image source

Channels are Random and Change by Location

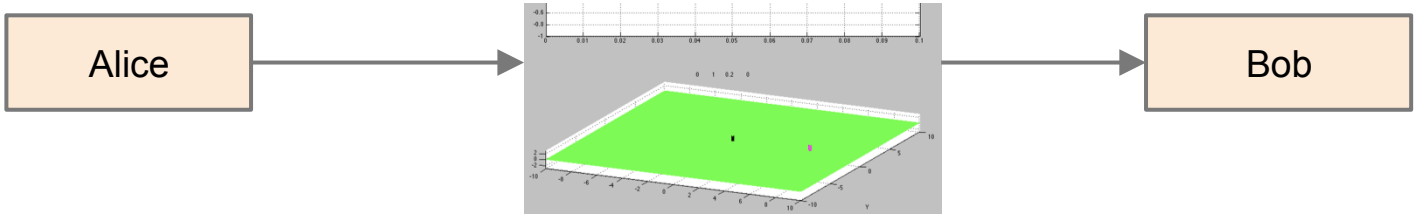


Bastille

Common Randomness Procedure

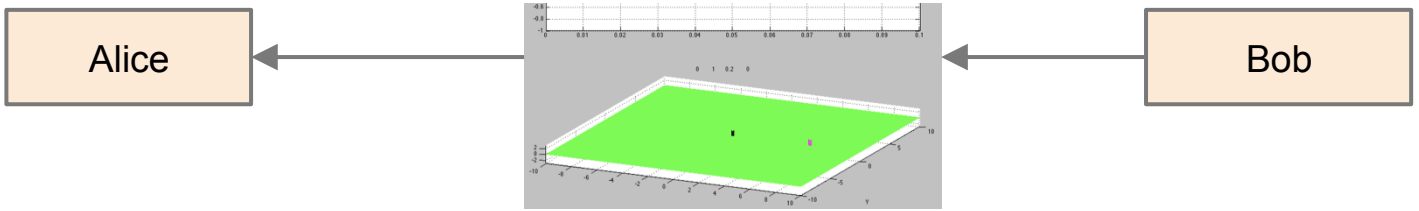
Alice *sounds* the channel to Bob

Channel gain = 5.2



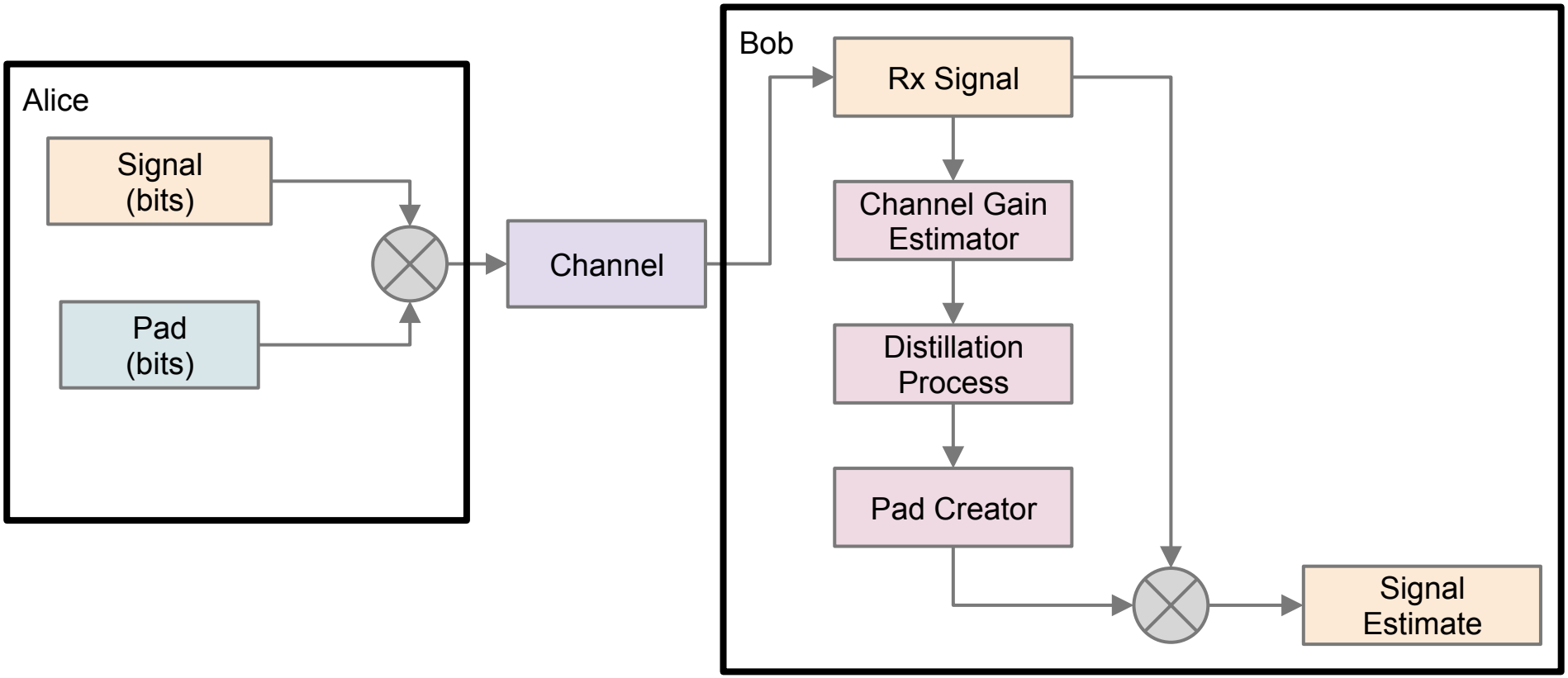
Bob *sounds* the channel to Alice

Channel gain = 5.3



- If Eve is sufficiently far away, she will calculate a completely different channel gain in this process
- Alice and Bob *reconcile* their estimates and *condense* them into a common encryption key

Channel Gains to Pad Bits



Idea presented in: U. Maurer, “Secret key agreement by public discussion from common information,” *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.

Summary

- Wireless security for L1-L3
- Information theory primer
- Physical layer security
- Encryption derived from physical layer

Thanks!

Presented by Bob Baxley
bob@bastille.io

Bastille