Shannon-inspired research tales

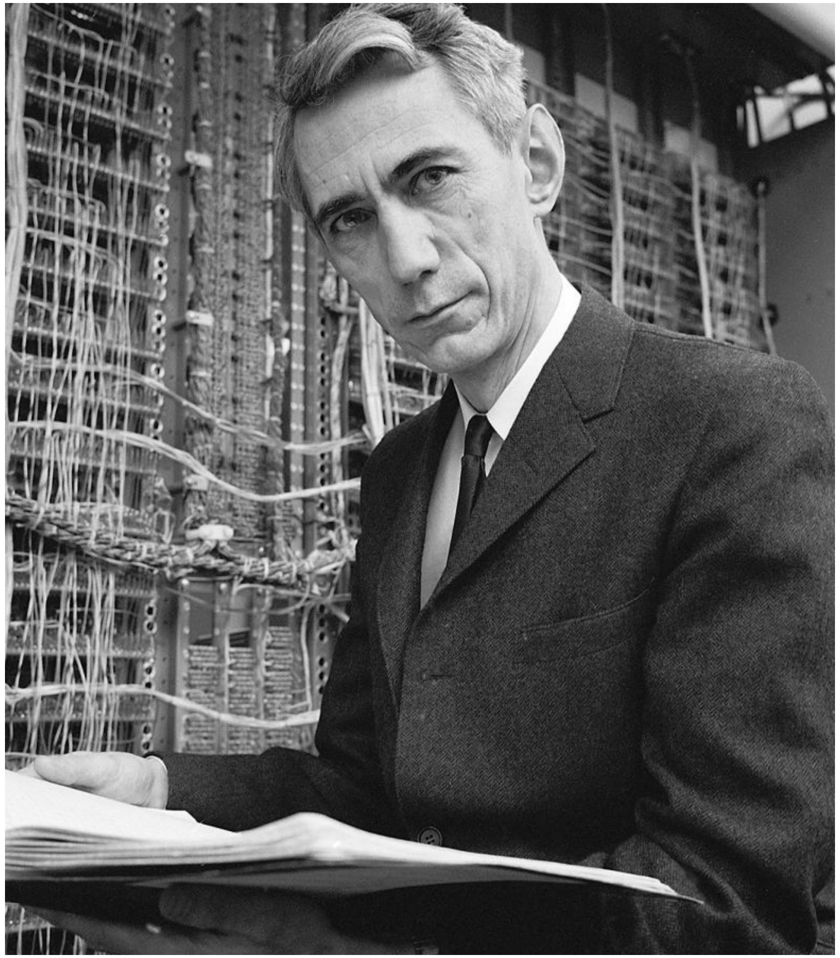
on Duality, Encryption, Sampling and Learning

Kannan Ramchandran
University of California, Berkeley

Berkeley Laboratory for Information and System Sciences

Shannon's incredible legacy

- A mathematical theory of communication
- Channel capacity
- Source coding
- Channel coding
- Cryptography
- Sampling theory
- …



(1916-2001)

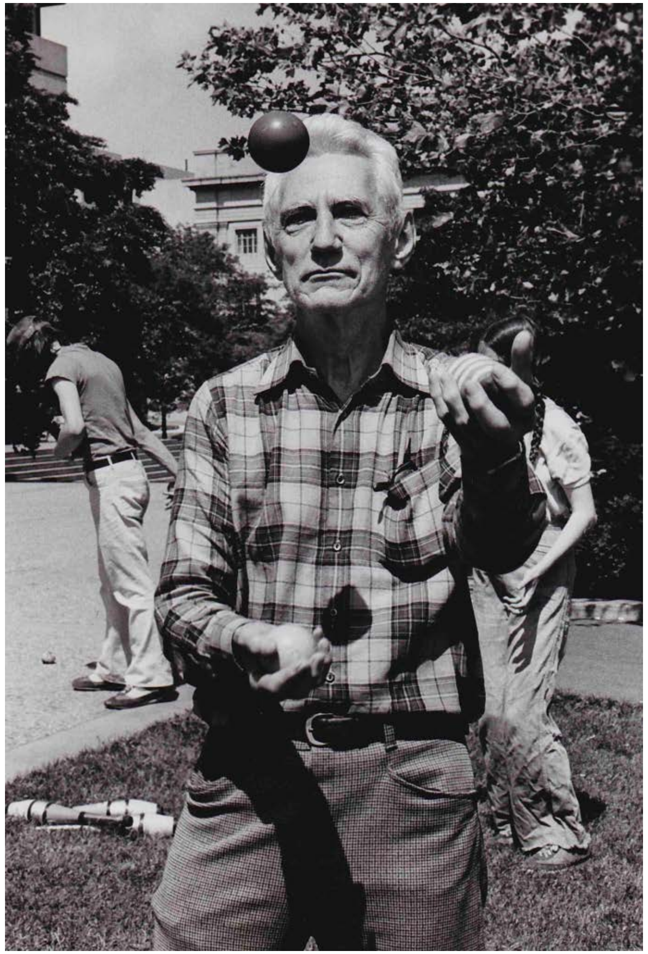
And many more…

- Boolean logic for switching circuits (MS thesis 1937)

- Juggling theorem:
  H(F+D)=N(V+D)

  ```
  F: the time a ball spends in the air,
  D: the time a ball spends in a hand,
  V: the time a hand is vacant,
  N: the number of balls juggled,
  H: the number of hands.
  ```

- …



(1916-2001)

Story: Shannon meets Einstein

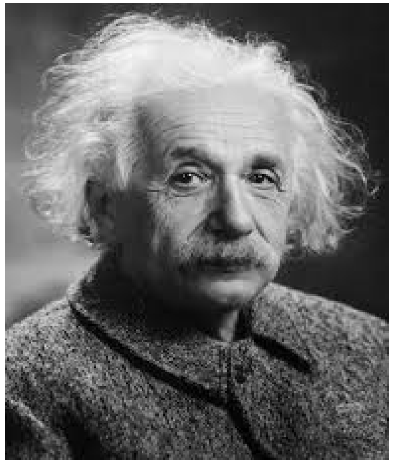
As narrated by Arthur Lewbel (2001)

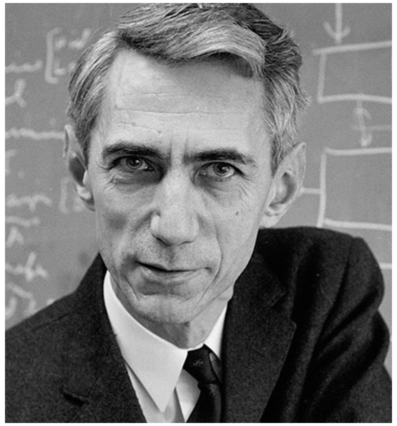
"

The story is that Claude was in the middle of giving a lecture to mathematicians in Princeton, when the door in the back of the room opens, and in walks **Albert Einstein**.

Einstein stands listening for a few minutes, whispers something in the ear of someone in the back of the room, and leaves. At the end of the lecture, Claude hurries to the back of the room to find the person that Einstein had whispered too, to find out what the great man had to say about his work.

The answer: Einstein had asked directions to the men's room.
"





Outline

Three "personal" Shannon-inspired research stories:

Chapter 1

Duality between source coding and channel coding – with side-information (2003)

Chapter 2

Encryption and **Compression** – swapping the order (2003)

Chapter 3

Sampling and **Learning** – Sampling below Nyquist rate and efficient learning (2014)

Sandeep Pradhan

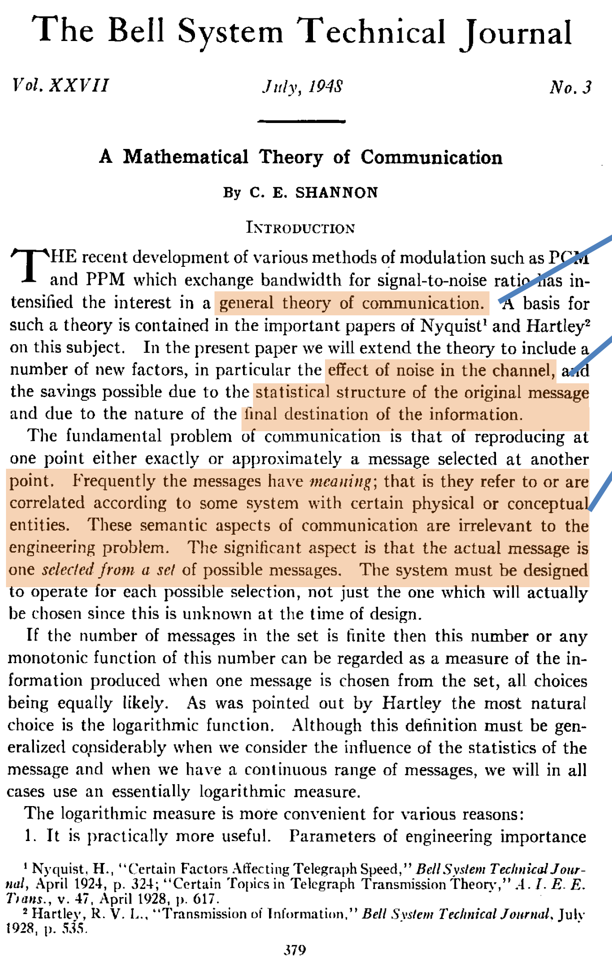
Jim Chou

Chapter 1

Duality

- source & channel coding
- with side-information

Shannon's celebrated 1948 paper



The Bell System Technical Journal

Vol. XXVII — July, 1948 — No. 3

A Mathematical Theory of Communication

By C. E. SHANNON

INTRODUCTION

THE recent development of various methods of modulation such as PCM and PPM which exchange bandwidth for signal-to-noise ratio has intensified the interest in a general theory of communication. A basis for such a theory is contained in the important papers of Nyquist[1] and Hartley[2] on this subject. In the present paper we will extend the theory to include a number of new factors, in particular the effect of noise in the channel, and the savings possible due to the statistical structure of the original message and due to the nature of the final destination of the information.

The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have *meaning*; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one *selected from a set* of possible messages. The system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design.

If the number of messages in the set is finite then this number or any monotonic function of this number can be regarded as a measure of the information produced when one message is chosen from the set, all choices being equally likely. As was pointed out by Hartley the most natural choice is the logarithmic function. Although this definition must be generalized considerably when we consider the influence of the statistics of the message and when we have a continuous range of messages, we will in all cases use an essentially logarithmic measure.

The logarithmic measure is more convenient for various reasons:

1. It is practically more useful. Parameters of engineering importance

[1] Nyquist, H., "Certain Factors Affecting Telegraph Speed," *Bell System Technical Journal*, April 1924, p. 324; "Certain Topics in Telegraph Transmission Theory," *A. I. E. E. Trans.*, v. 47, April 1928, p. 617.
[2] Hartley, R. V. L., "Transmission of Information," *Bell System Technical Journal*, July 1928, p. 535.

379

general theory of communication

communication system as source/channel/destination

abstraction of the concept of message

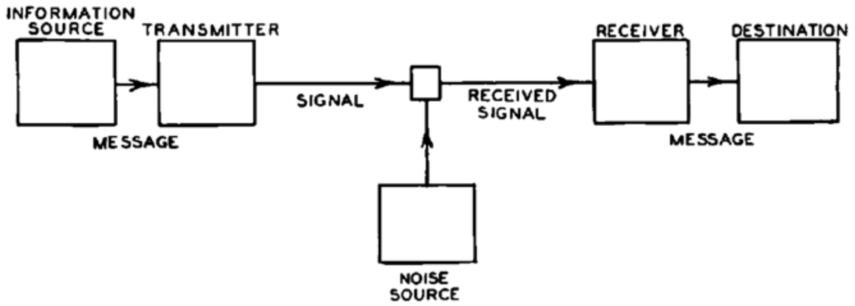
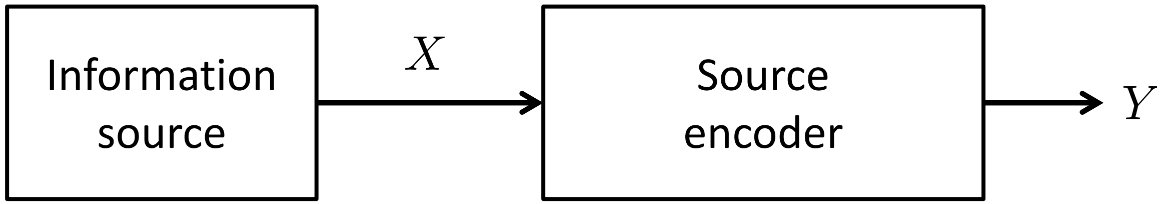


Fig. 1—Schematic diagram of a general communication system.

Source coding



$$H(X) = \mathbb{E}_X \left[\log \left(\frac{1}{p(X)} \right) \right]$$

Entropy of a random variable
= minimum number of bits required to represent the source

Rate-distortion theory - 1948

- Trade-off between ***compression rate*** and the ***distortion***

PART V: THE RATE FOR A CONTINUOUS SOURCE

27. FIDELITY EVALUATION FUNCTIONS

In the case of a discrete source of information we were able to determine a definite rate of generating information, namely the entropy of the underlying stochastic process. With a continuous source the situation is considerably more involved. In the first place a continuously variable quantity can assume an infinite number of values and requires, therefore, an infinite number of binary digits for exact specification. This means that to transmit the output of a continuous source with *exact recovery* at the receiving point requires, in general, a channel of infinite capacity (in bits per second). Since, ordinarily, channels have a certain amount of noise, and therefore a finite capacity, exact transmission is impossible.

This, however, evades the real issue. Practically, we are not interested in exact transmission when we have a continuous source, but only in transmission to within a certain tolerance. The question is, can we assign a definite rate to a continuous source when we require only a certain fidelity of recovery, measured in a suitable way. Of course, as the fidelity require-

Mutual information:

$\mathcal{H}(X)\text{-}\mathcal{H}(X|Y)$

$$R(D) = \min_{P_{Y|X}(y|x)} I(X;Y)$$

$$\text{subject to} \quad \mathbb{E}\left[d(X,Y)\right] \leq D$$

distortion measure

Channel coding

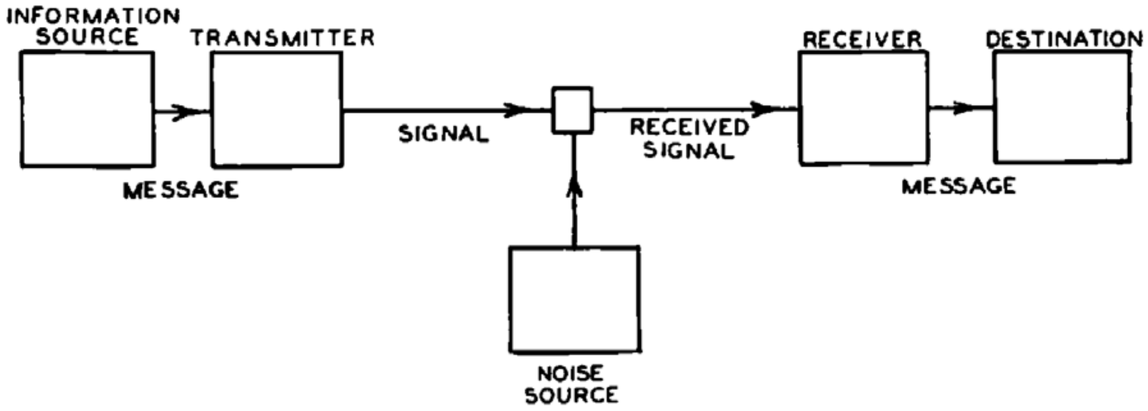


Fig. 1—Schematic diagram of a general communication system.

- For rates $R < C$, can achieve arbitrary small error probabilities
- Used to be thought one needs $R \to 0$

capacity

$$C(W) = \max_{P_X(x)} \quad I(X;Y)$$

$$\text{subject to} \quad \mathbb{E}\left[w(X)\right] \leq W$$

cost measure

Shannon’s breakthrough

- Communication before Shannon:
  - *Linear filtering* (Wiener) at receiver to remove noise
- Communication after Shannon:
  - Designing codebooks
  - *Non-linear estimation* (MLE) at receiver

Reliable transmission at rates approaching channel capacity

Shannon (1959)

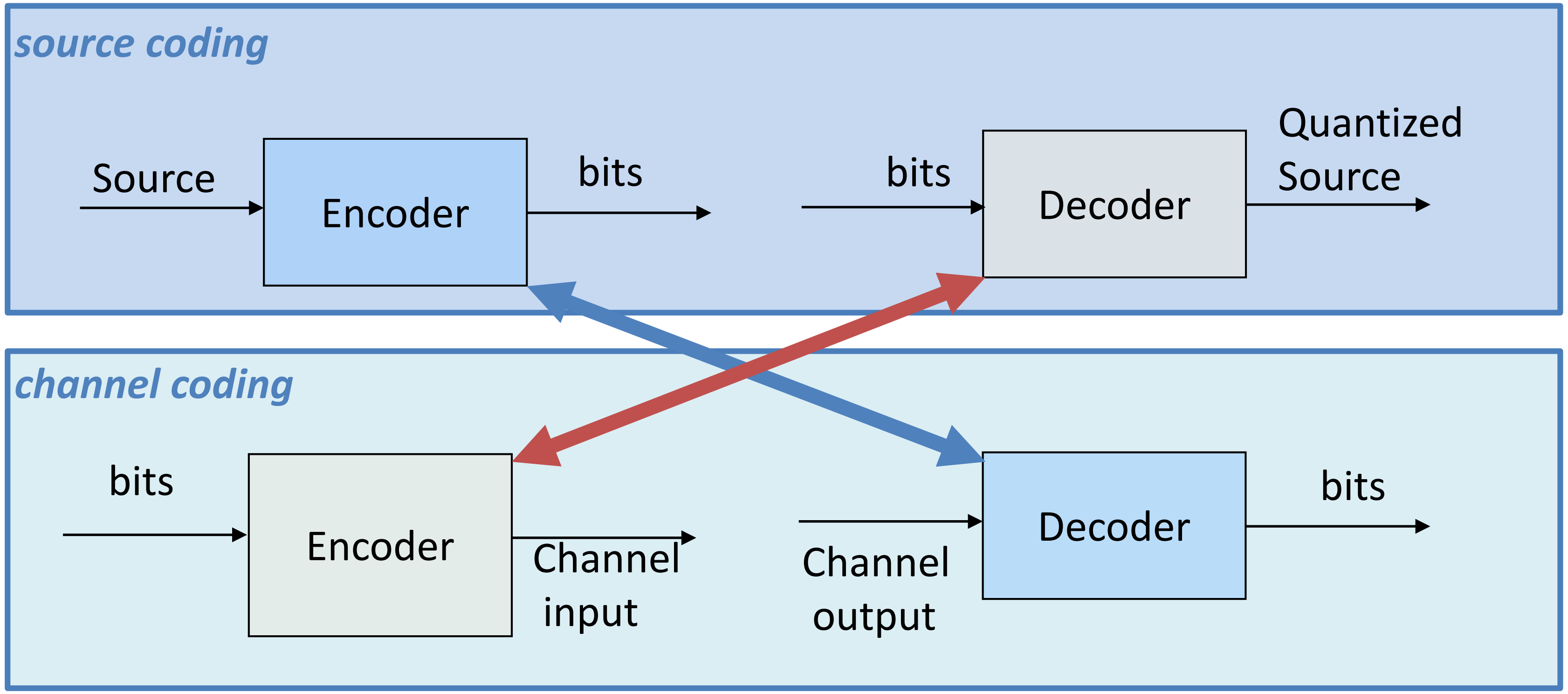
"***There is a curious and provocative duality between the properties of a source with a distortion measure and those of a channel.*** *This duality is enhanced if we consider channels in which there is a* ***cost*** *associated with the different input letters, and it is desired to find the capacity subject to the constraint that the expected cost not exceed a certain quantity…..*

Shannon (1959)

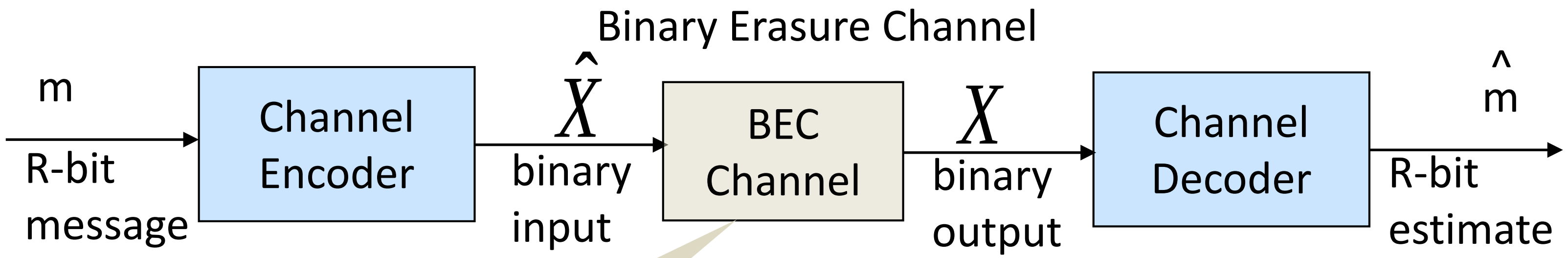
…This duality can be pursued further and is related to a duality between past and future and the notions of control and knowledge. ***Thus, we may have knowledge of the past but cannot control it; we may control the future but not have knowledge of it.***"

Functional duality

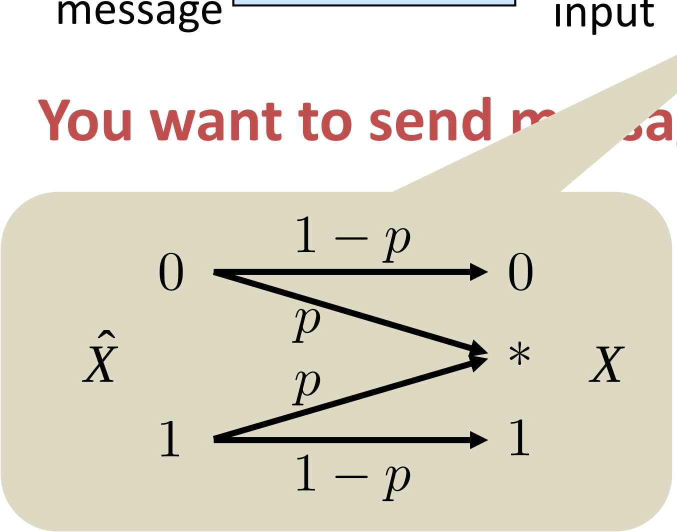
When is the ***optimal encoder*** for one problem functionally identical to the ***optimal decoder*** for the dual problem?



Duality example: Channel coding



You want to send message m: how big can you make R?



Shannon's result:

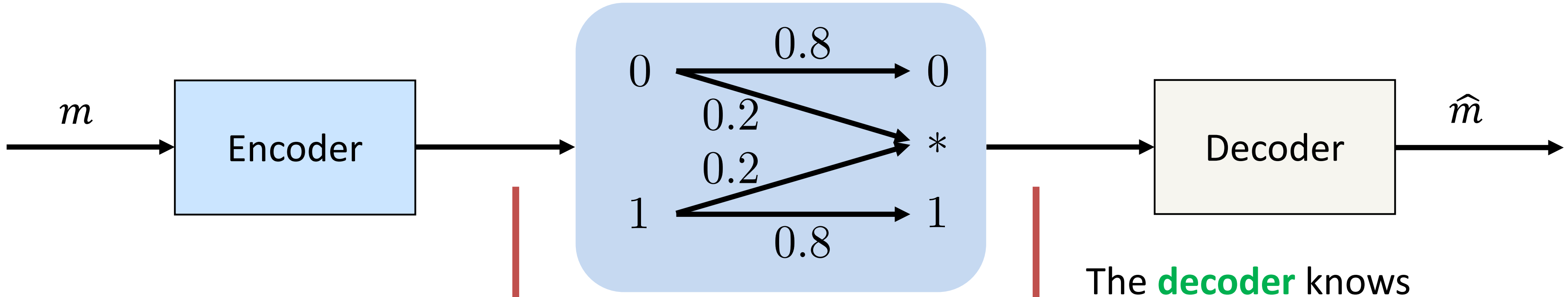
C_{BEC}**=(1-p)** *bits per channel use*

$p = 0.2$

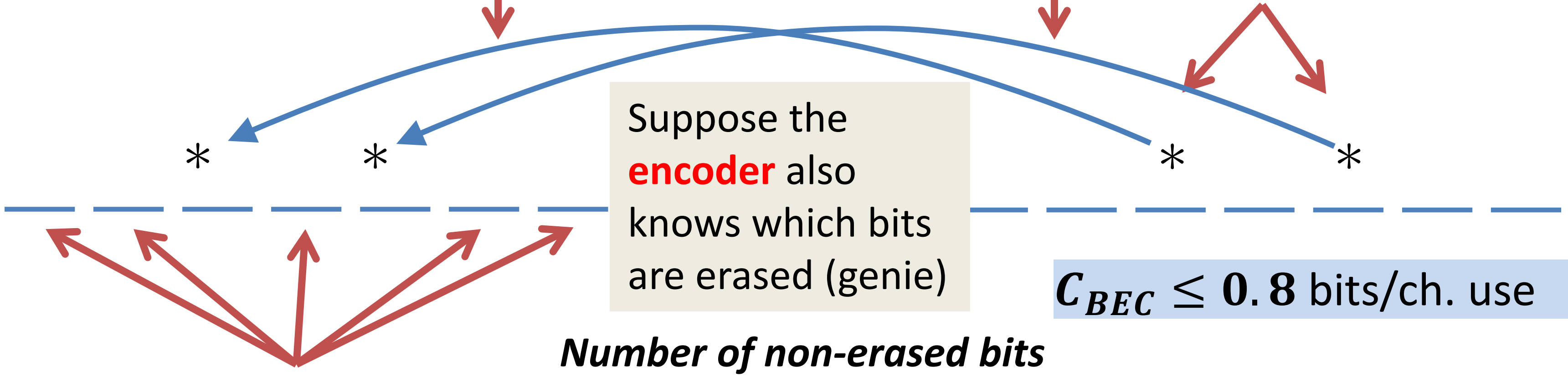
$\boldsymbol{Cost}\ (\mathbf{0}) = 1\,;\ \boldsymbol{Cost}\ (\mathbf{1}) = 1$

$\boldsymbol{Total\ budget} \leq 10{,}000$

What is the Shannon capacity?



The **decoder** knows which bits are erased (channel output)



Suppose the **encoder** also knows which bits are erased (genie)

Send information in non-erased locations

Number of non-erased bits

$\approx \mathbf{10,000 \times (1-p)}$

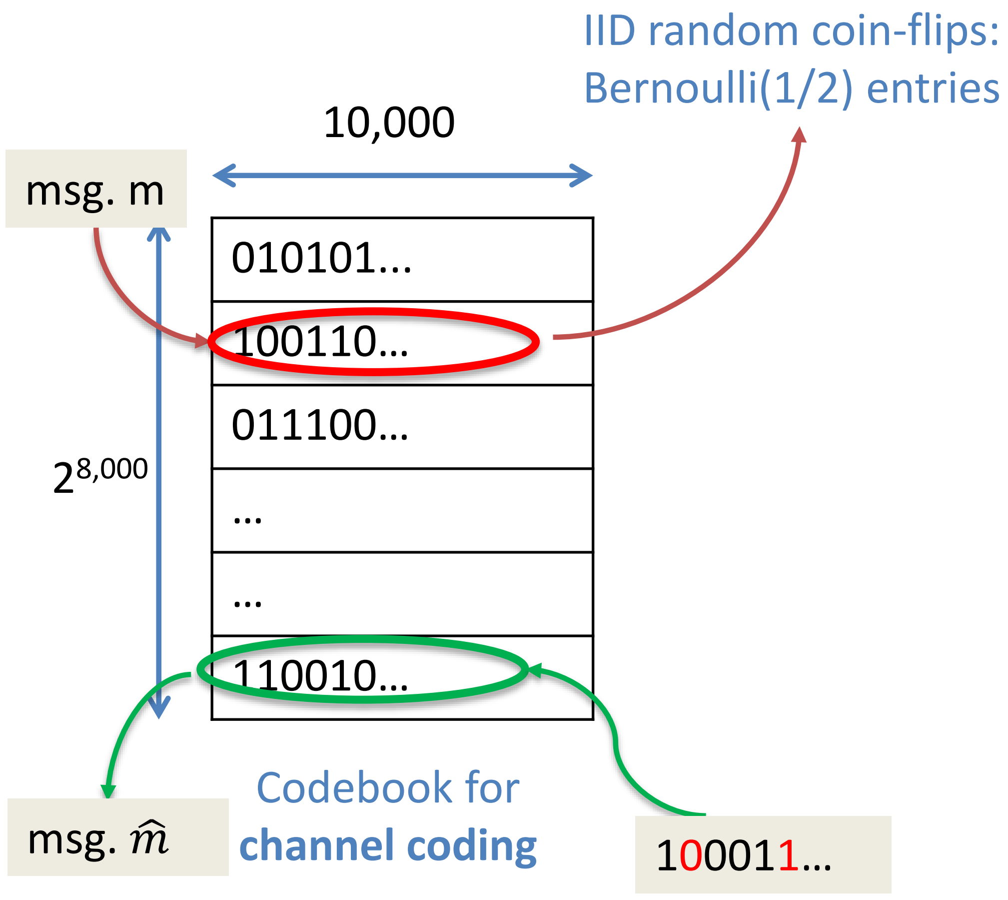
$= \mathbf{10,000 \times 0.8}$

$= \mathbf{8,000}$

$C_{BEC} \leq \mathbf{0.8}$ bits/ch. use

Surprise: ***the encoder does not need to know which bits are erased!***

Shannon's prescription: random coding



1) Encoder & Decoder agree on a random codebook

Shannon's random coding argument

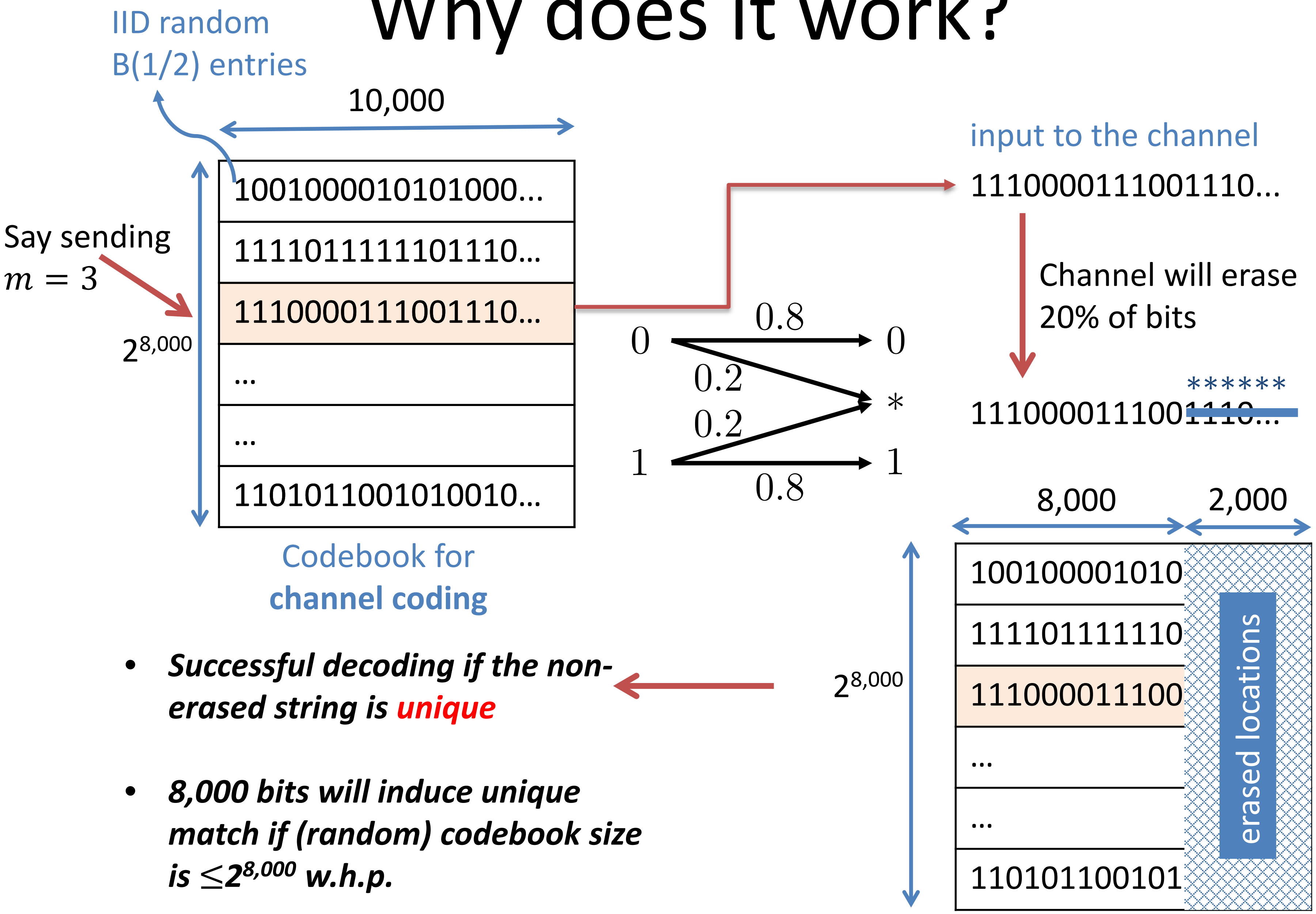
2) Encoder encodes message

Output the codeword corresponding to the index

3) Decoder decodes message

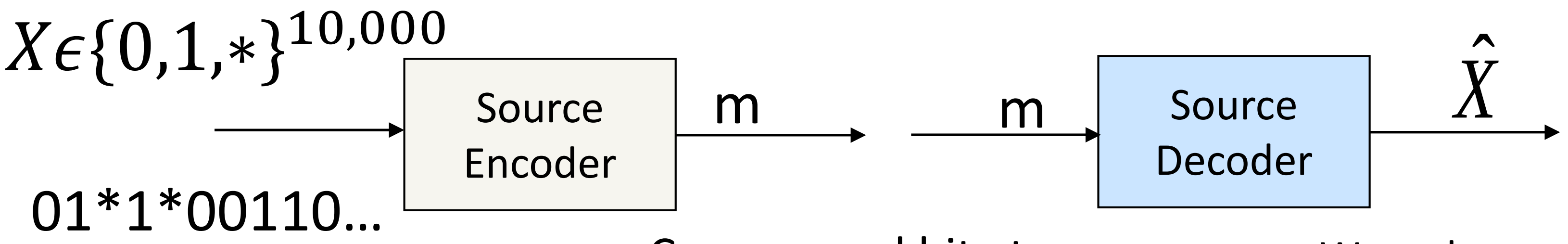
Output the index corresponding to the ***closest*** *codeword*

Why does it work?



- ***Successful decoding if the non-erased string is unique***
- ***8,000 bits will induce unique match if (random) codebook size is*** $\leq 2^{8,000}$ ***w.h.p.***

Source Coding Dual to the BEC: BEQ

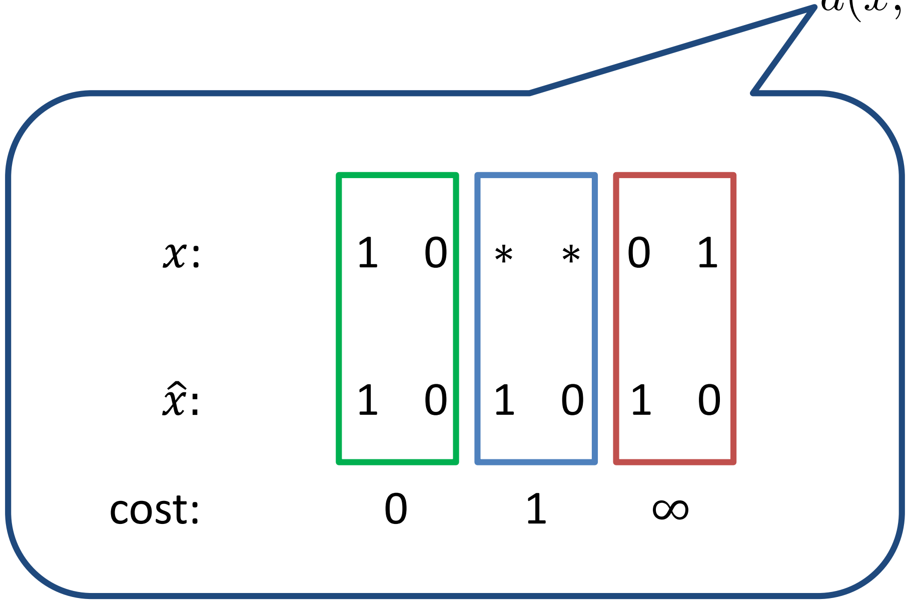


Compressed bit-stream 8,000 bits

Want the average distortion to be ≤ 0.2

$p(0) = p(1) = 0.4$; $p(*) = 0.2$

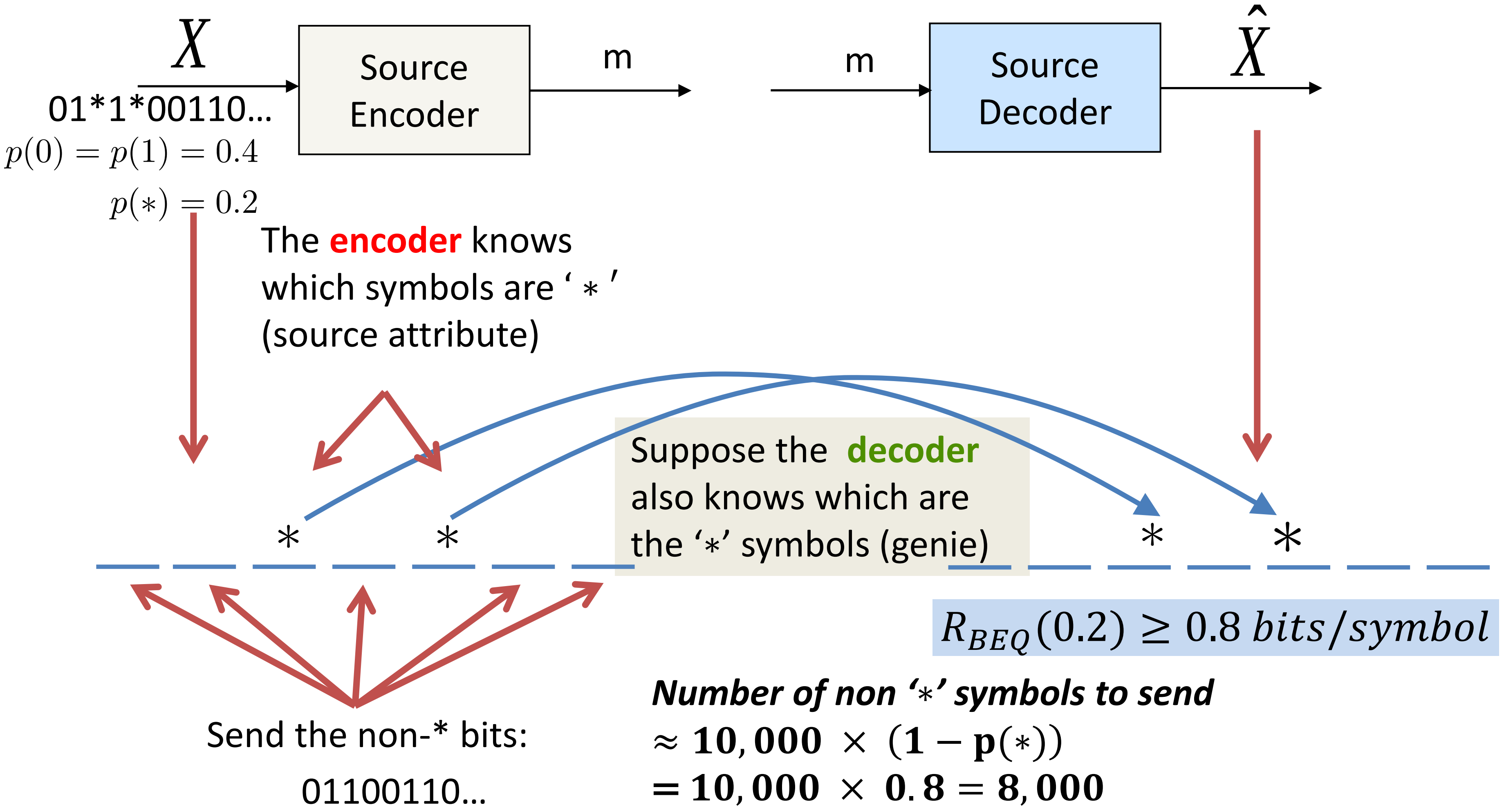
$$d(x, \hat{x}) = \begin{cases} 0 & \text{if } \hat{x} = x \text{ for } x \in \{0,1\} \\ \infty & \text{if } \hat{x} \neq x \text{ for } x \in \{0,1\} \\ 1 & \text{if } x = * \end{cases}$$



$*$ is like a "don't care" symbol (e.g., perceptually masked symbols). How can we exploit this for compression?

Martinian and Yedidia, 2004

Source Coding Dual to the BEC: BEQ



$R_{BEQ}(0.2) \geq 0.8\ bits/symbol$

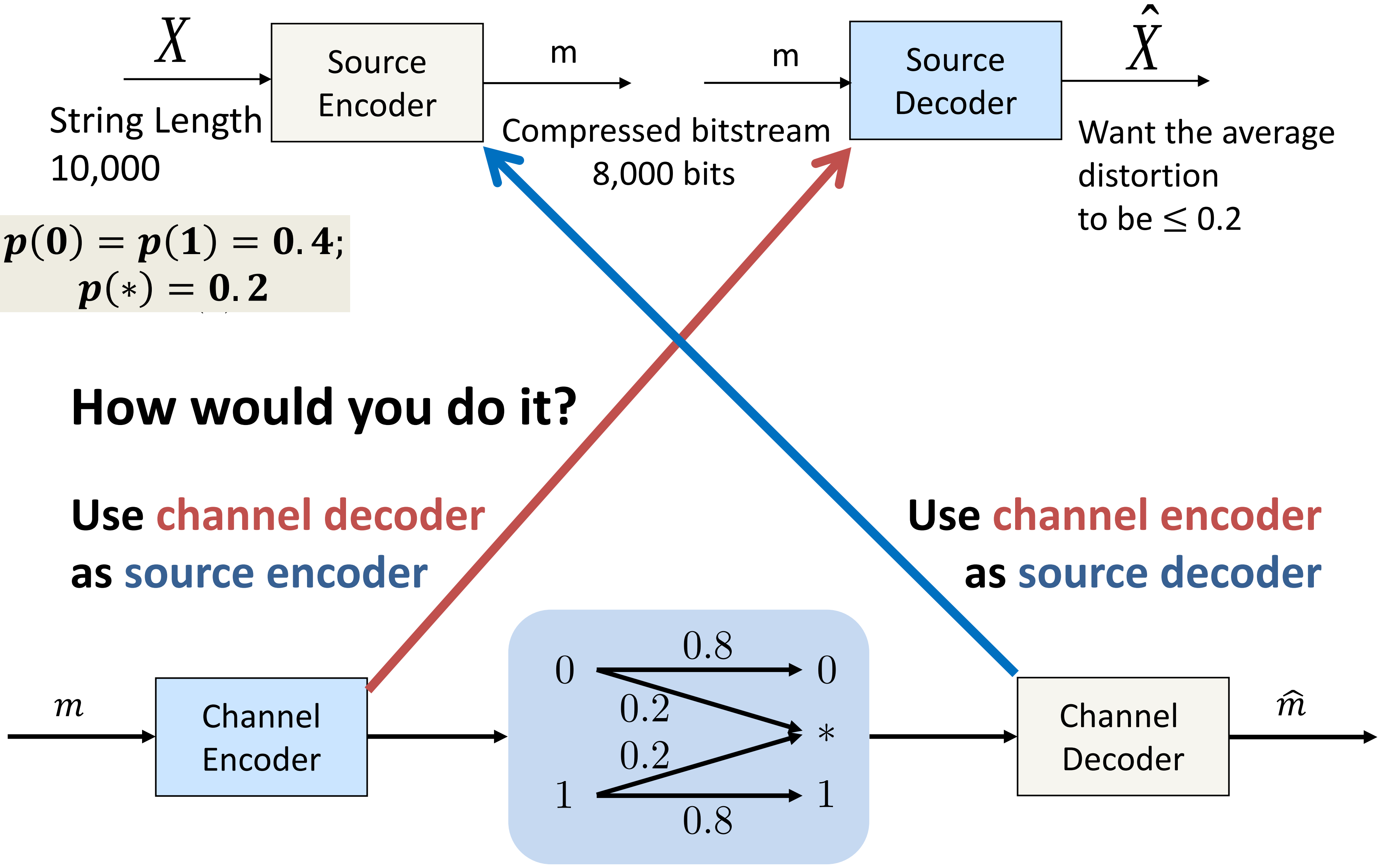
Number of non '$*$' symbols to send

$\approx \mathbf{10,000 \times (1 - p(*))}$

$= \mathbf{10,000 \times 0.8 = 8,000}$

Surprise: ***the decoder does not need to know which symbols are '$*$'!***

Source Coding Dual to the BEC: BEQ



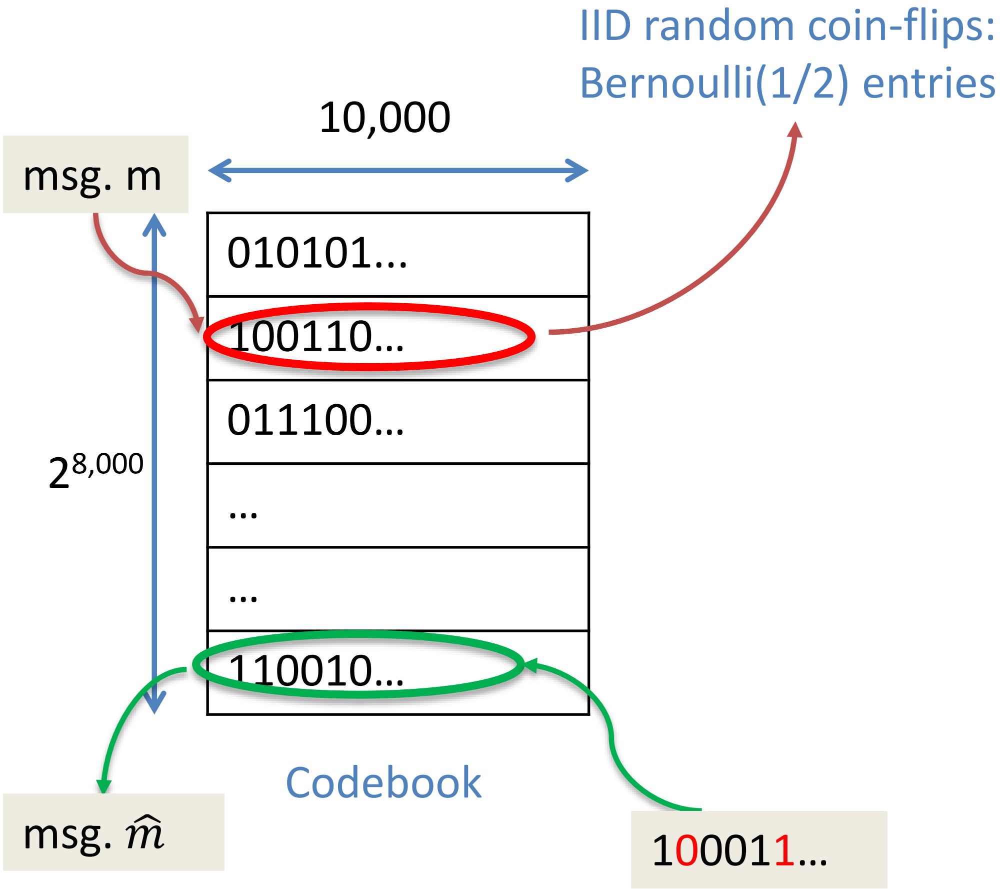
$$p(0) = p(1) = 0.4; \quad p(*) = 0.2$$

How would you do it?

Use channel decoder as source encoder

Use channel encoder as source decoder

Shannon's prescription: random coding



1) Encoder & Decoder agree on a random codebook

Shannon's random coding argument

2) Encoder encodes message

~~*Output the codeword corresponding to the index*~~

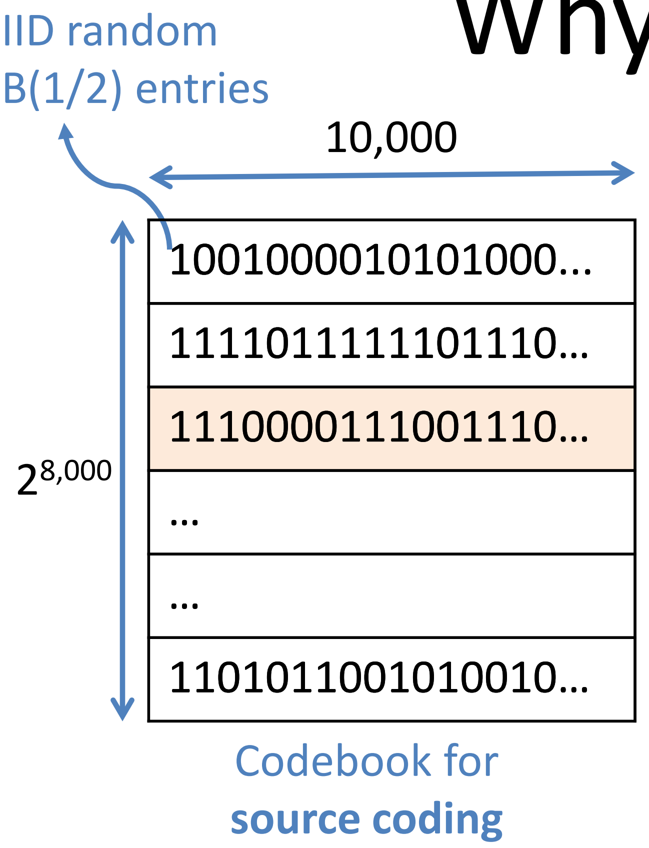
*Output the index corresponding to the **closest** codeword*

3) Decoder decodes message

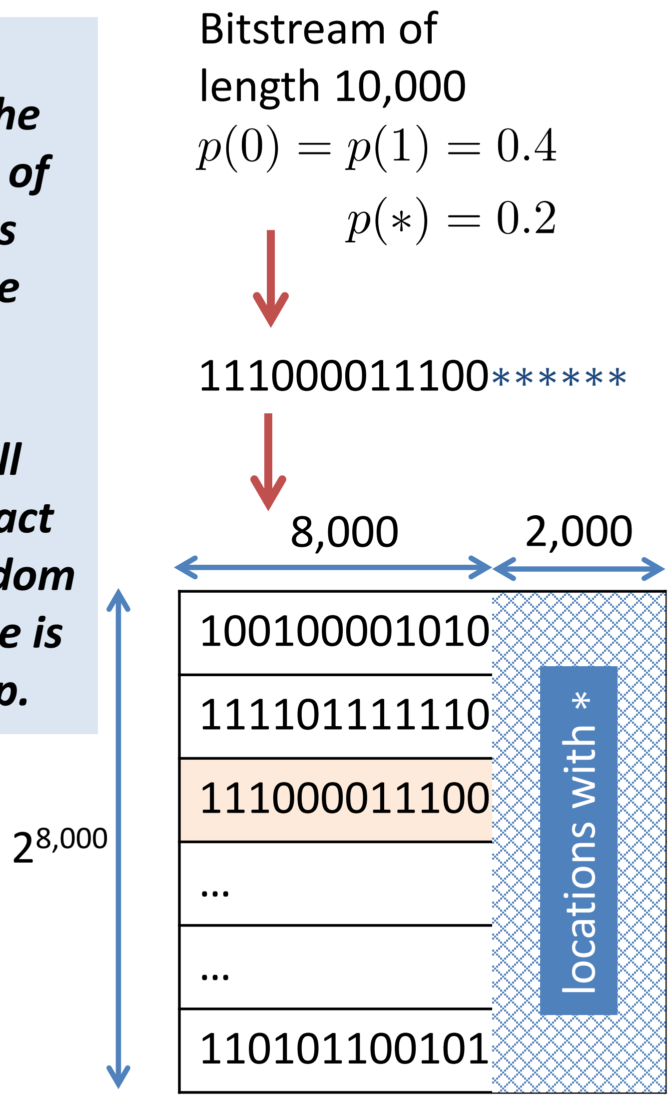
~~*Output the index corresponding to the **closest** codeword*~~

Output the codeword corresponding to the index

Why does it work?



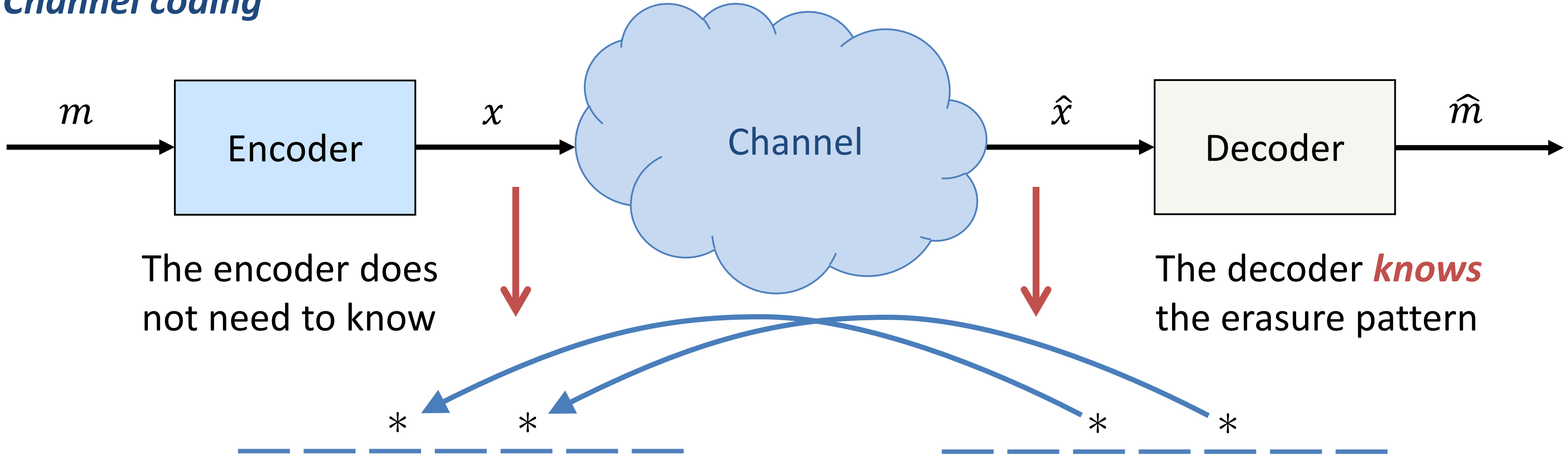
- ***Successful encoding if the "non-*" part of input string is present in the codebook***
- ***8,000 bits will induce an exact match if random codebook size is $\geq 2^{8,000}$ w.h.p.***



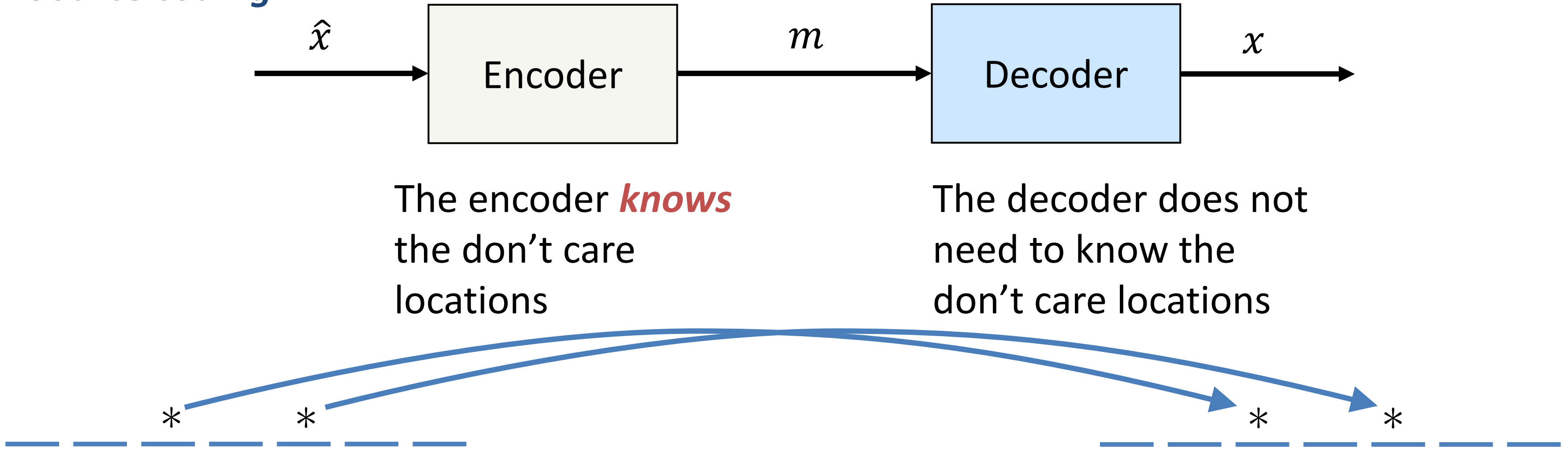
Index of the codeword that exactly matches the non-* part of input string

Knowledge of the erasure pattern

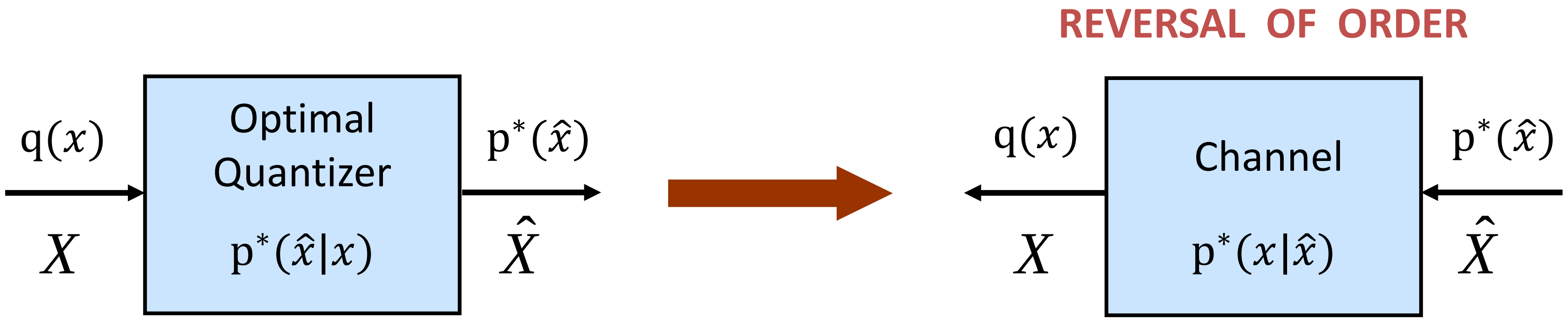
Channel coding



Source coding



Duality between source and channel coding



Given a ***source coding problem*** with source distribution $\mathrm{q}(x)$, optimal quantizer $\mathrm{p}^*(\hat{x}|x)$, distortion measure $d(x,\hat{x})$ and distortion constraint **D**

There is a ***dual channel coding problem*** with channel $\mathrm{p}^*(x|\hat{x})$ cost measure $w(\hat{x})$ and cost constraint **W** such that

$$\boldsymbol{R(D) = C(W)}$$

$$w(\hat{x}) = c_1 D(p^*(x|\hat{x})\,||q(x)) + \theta$$

$$W = E_{p^*(\hat{x})} w(\hat{X}).$$

Pradhan, Chou and R, 2003

Interpretation of functional duality

For ***any*** given source coding problem, there is a ***dual*** channel coding problem such that:

- both problems induce the ***same optimal joint distribution***
- the ***optimal encoder*** for one is ***functionally identical*** to the ***optimal decoder*** for the other
- an appropriate ***channel-cost measure*** is associated

Key takeaway

Source coding

distortion measure is as important as the ***source distribution***

Channel coding

channel cost measure is as important as the ***channel conditional distribution***

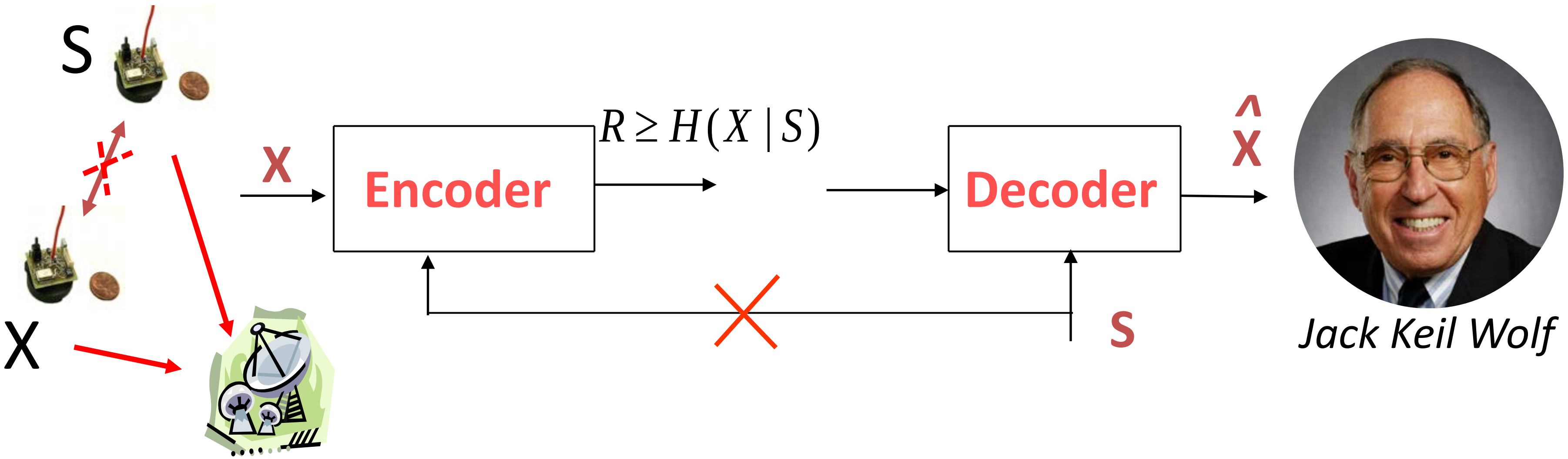
Duality between

source coding with side information

and

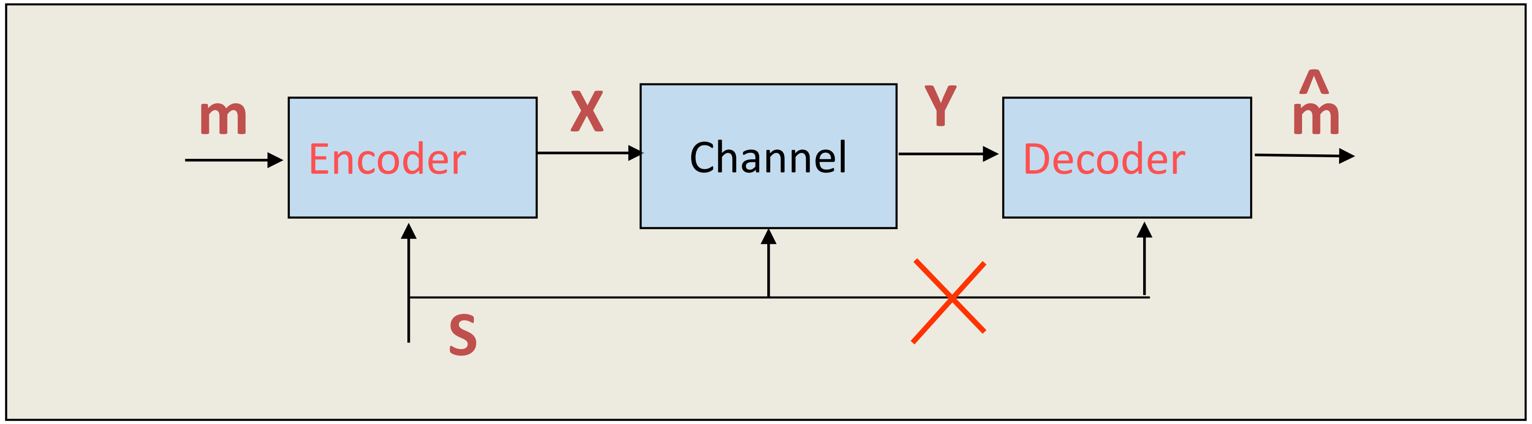
channel coding with side information

Source coding with side information (SCSI):



- (Only) decoder has access to side-information S
- Studied by Slepian-Wolf ‘73, Wyner-Ziv ’76, Berger ’77
- Applications: sensor networks (IoT), digital upgrade, secure compression.
- **No performance loss in some important cases**

Channel coding with side information (CCSI):



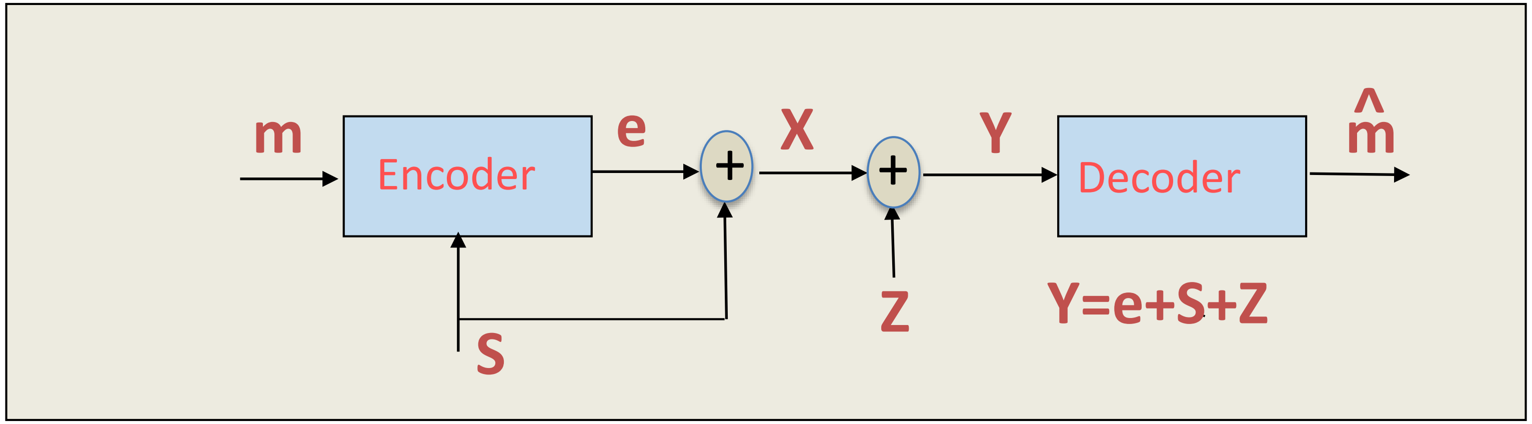
- (Only) encoder has access to ``interfering” side-information S

- Studied by Gelfand-Pinsker ‘81, Costa ‘83, Heegard-El Gamal ’85

- Applications: data hiding, watermarking, precoding for known interference, writing on dirty paper, MIMO broadcast.

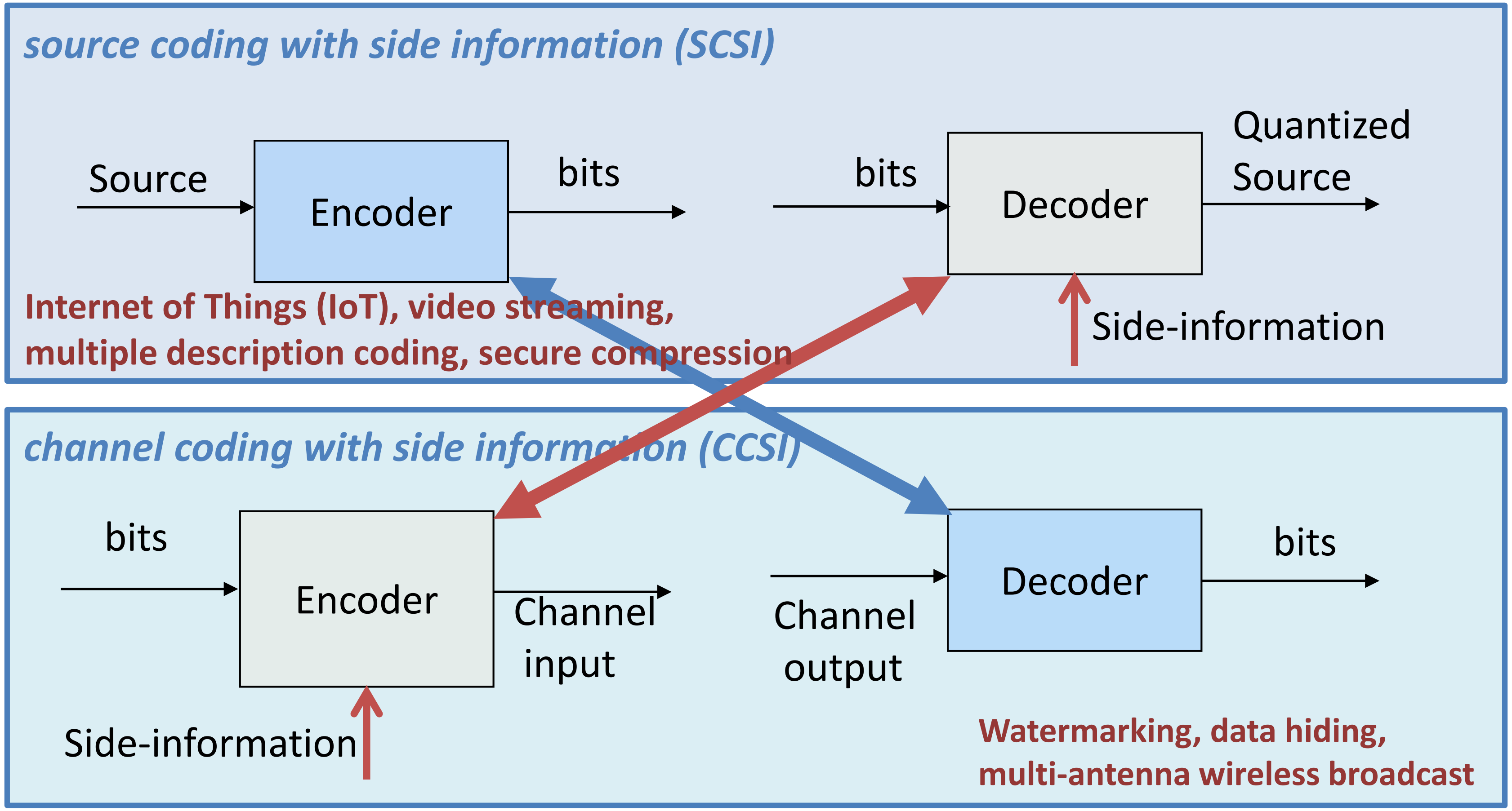
- **No performance loss in some important cases**

Channel coding with side information (CCSI):



- Encoder (only) has access to ``interfering” side-information S
- Studied by Gelfand-Pinsker ‘81, Costa ‘83, Heegard-El Gamal ’85
- Applications: data hiding, watermarking, precoding for known interference, writing on dirty paper, MIMO broadcast.
- **No performance loss in some important cases**

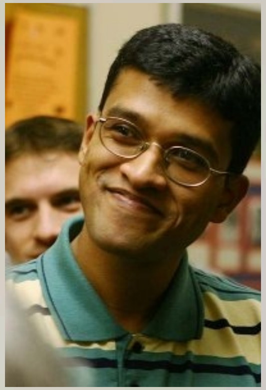
Duality between *source coding* & *channel coding with side information*



Pradhan, Chou and R, 2003

Mark Johnson

Prakash Ishwar



Vinod Prabhakaran

Chapter 2

Cryptography

- Compressing encrypted data

Cryptography – 1949

- Foundations of ***modern cryptography***
- All theoretically unbreakable ciphers must have the properties of one-time pad

Communication Theory of Secrecy Systems*

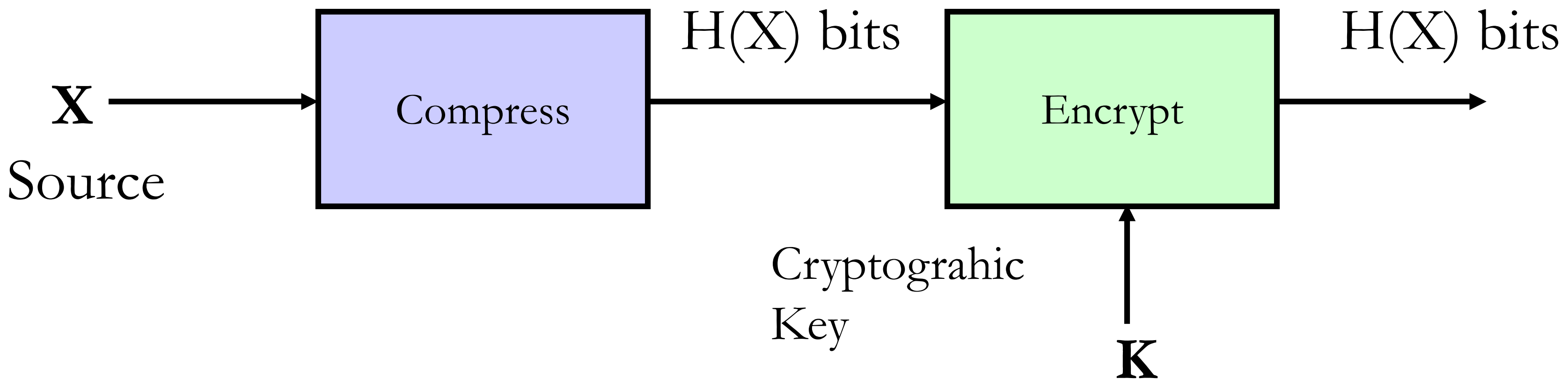
By C. E. SHANNON

1. Introduction and Summary

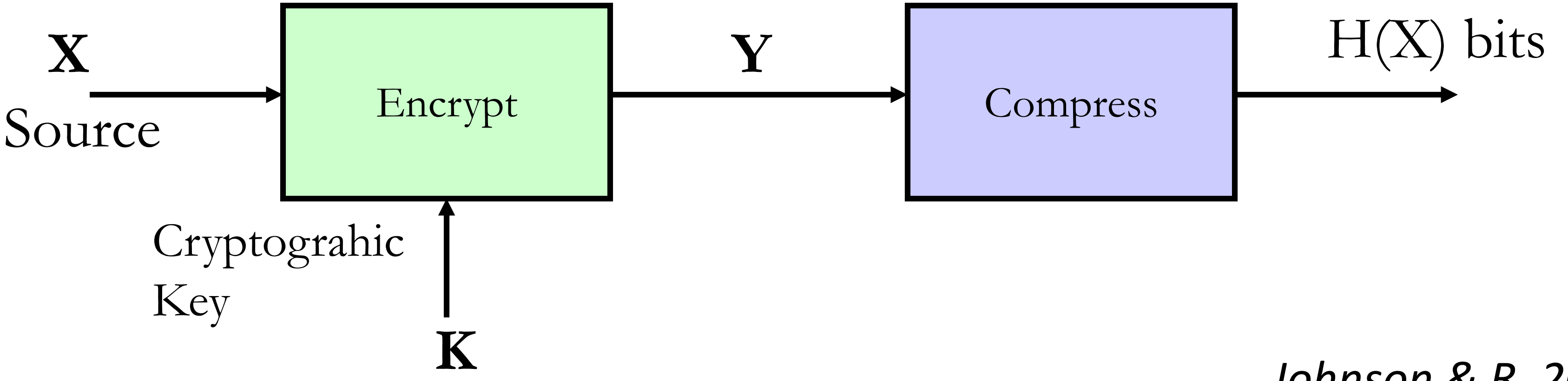
THE problems of cryptography and secrecy systems furnish an interesting application of communication theory.[1] In this paper a theory of secrecy systems is developed. The approach is on a theoretical level and is intended to complement the treatment found in standard works on cryptography.[2] There, 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.

Compressing Encrypted Data

"Correct" order



Wrong order?



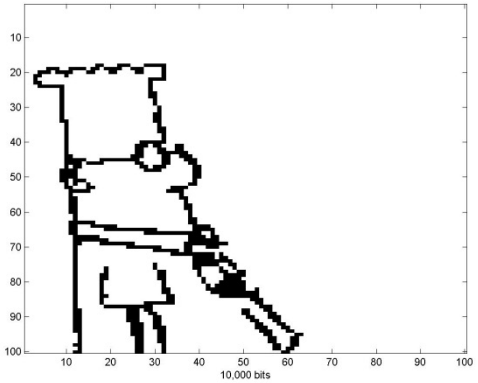
Johnson & R, 2003

Example

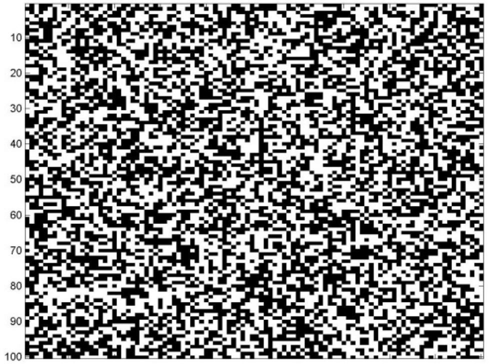
10,000 bits

5,000 bits

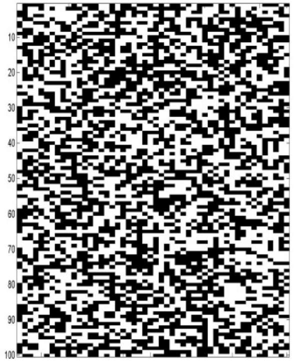
Original Image



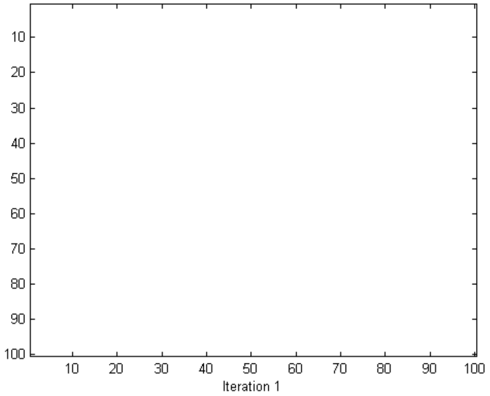
Encrypted Image



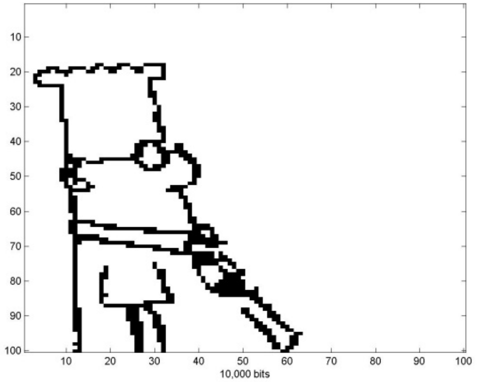
Compressed Encrypted Image

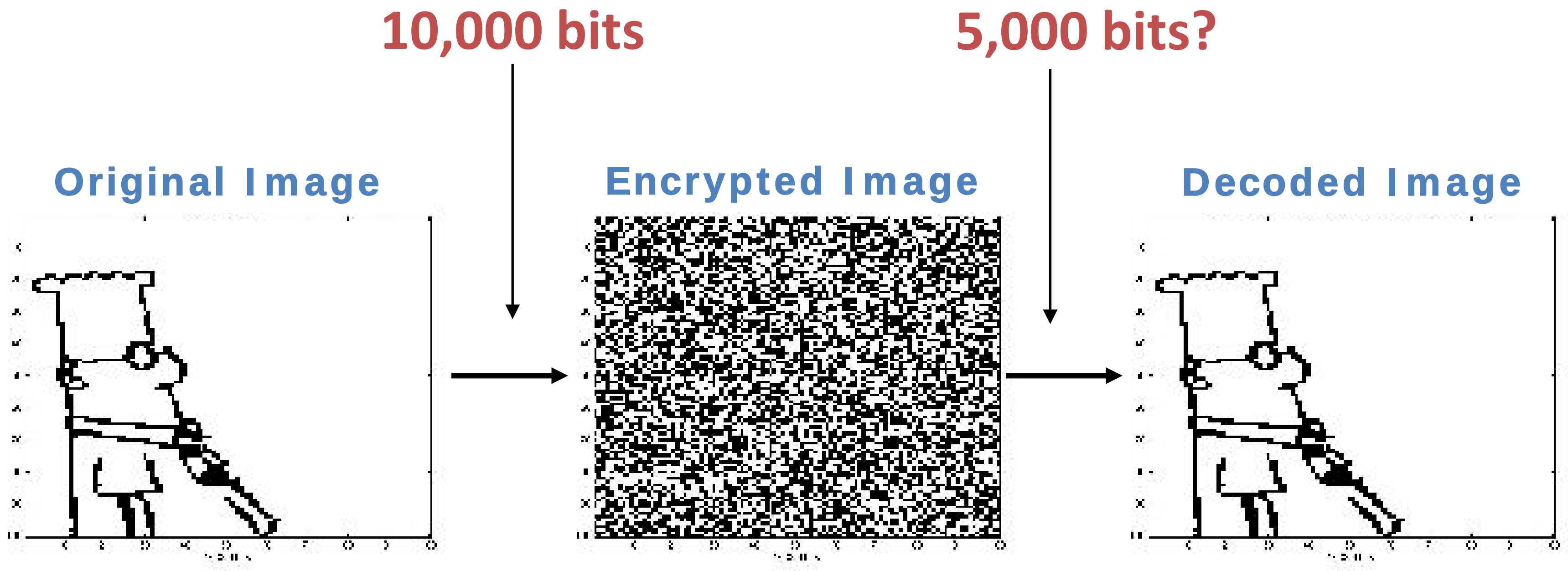


Decoding Compressed Image

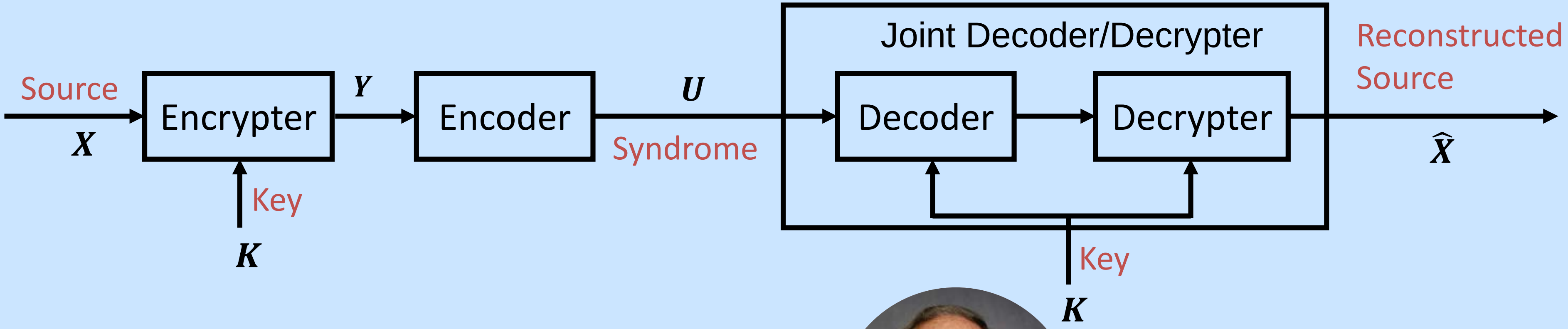


Final Reconstructed Image





Key Insight!

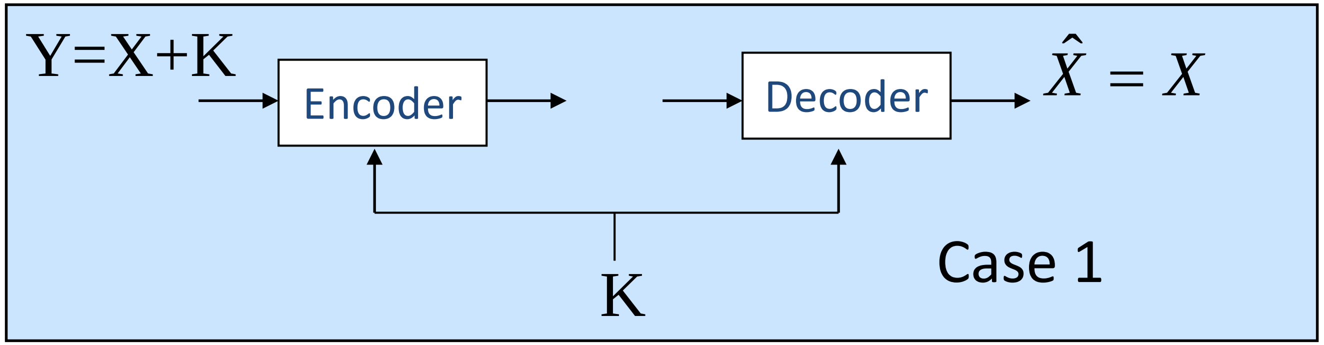


- Y = X + K where X is independent of K
- **Slepian-Wolf theorem**:
  can send X at rate H(Y|K) = H(X)

SCSI: binary example of noiseless compression

(Slepian-Wolf '73)

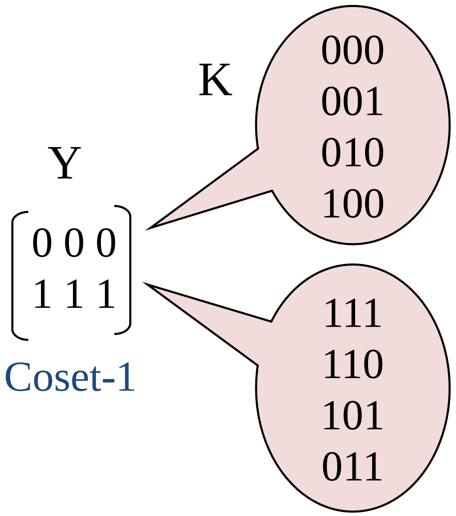
- **X** is uniformly chosen from {[000], [001], [010], [100]}
- **K** is a length-3 random key (equally likely in $\{0,1\}^3$)
- Correlation: Hamming distance between **Y** and **K** at most 1
- Example: when **K**=[0 1 0], **Y** => [0 1 0], [0 1 1], [0 0 0], [1 1 0]



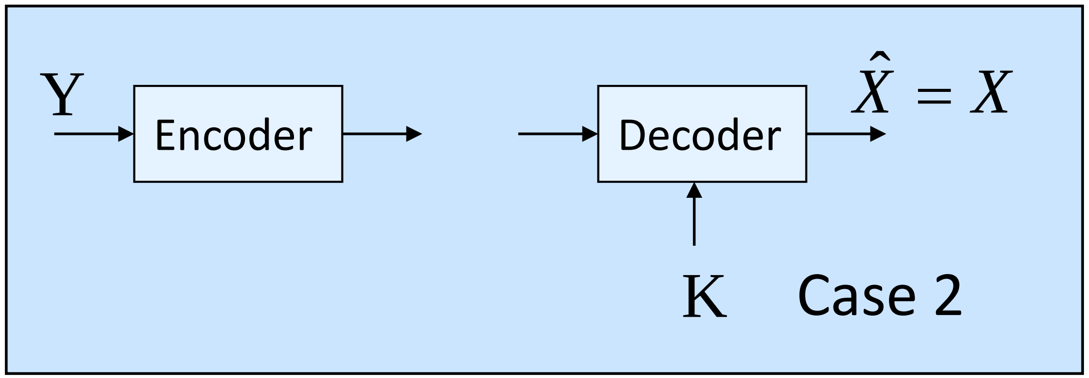
- **Encoder computes X=Y+K (mod 2)**
- **Encoder represents X using 2 bits**
- **Decoder outputs X (mod 2)**

00 → 000
01 → 001
10 → 010
11 → 100

=Y+K



(Slepian-Wolf '73)

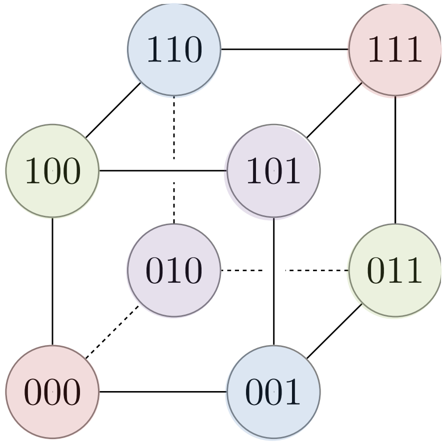


Coset-1 (00) $\begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$

Coset-2 (01) $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

Coset-3 (10) $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

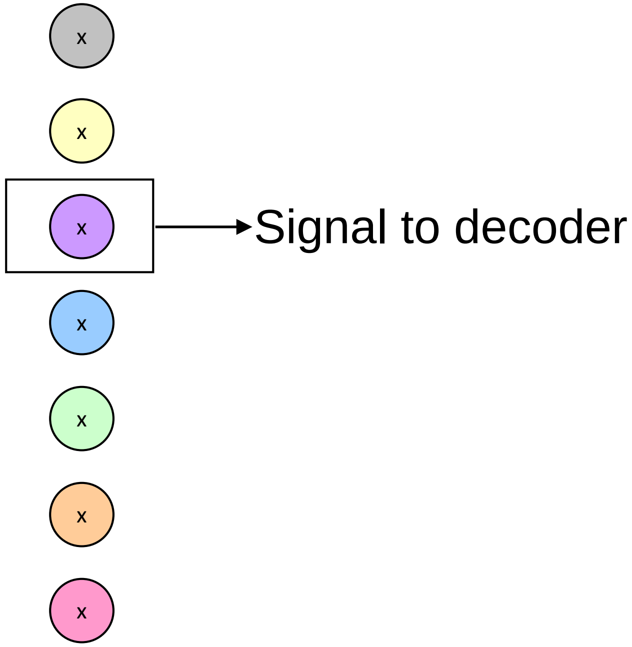
Coset-4 (11) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$



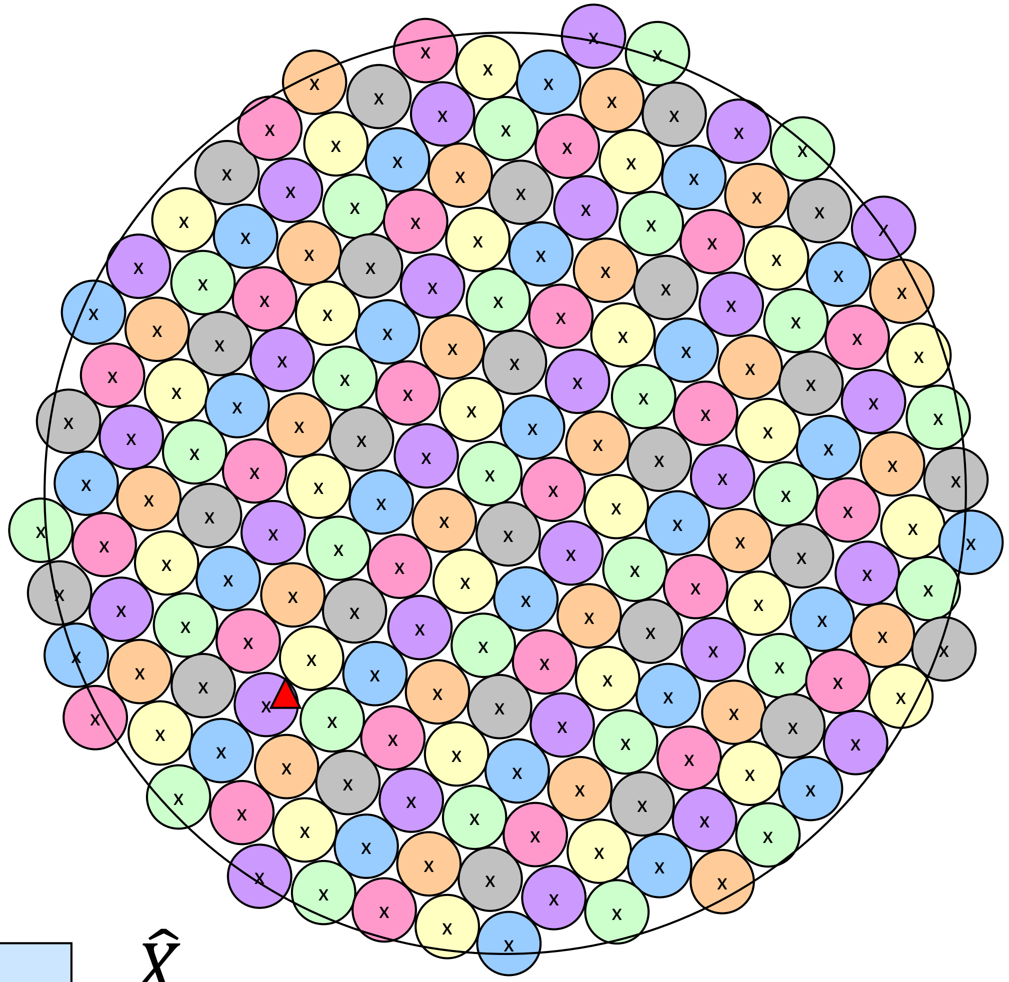
- Transmission at 2 bits/sample
- Encoder => send index of the coset containing X.
- Decoder => find a codeword in given coset closest to K

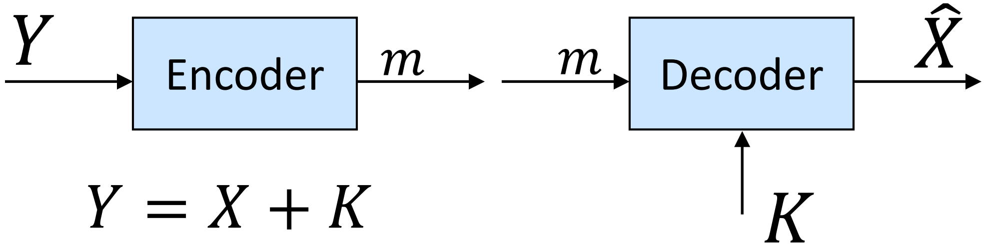
Example: Y=010 (K=110) => Encoder sends message 10

Geometric illustration



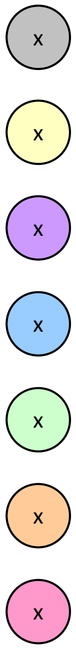
▲ Y (encrypted)





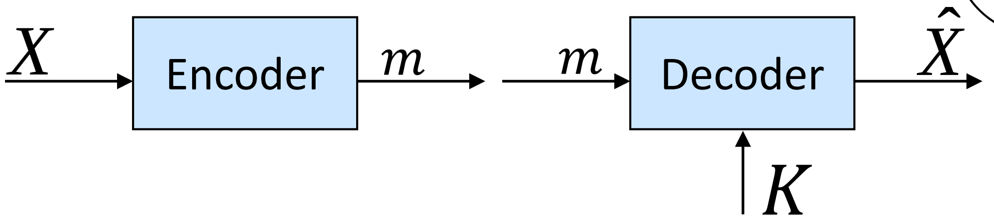
X (unencrypted & compressible)

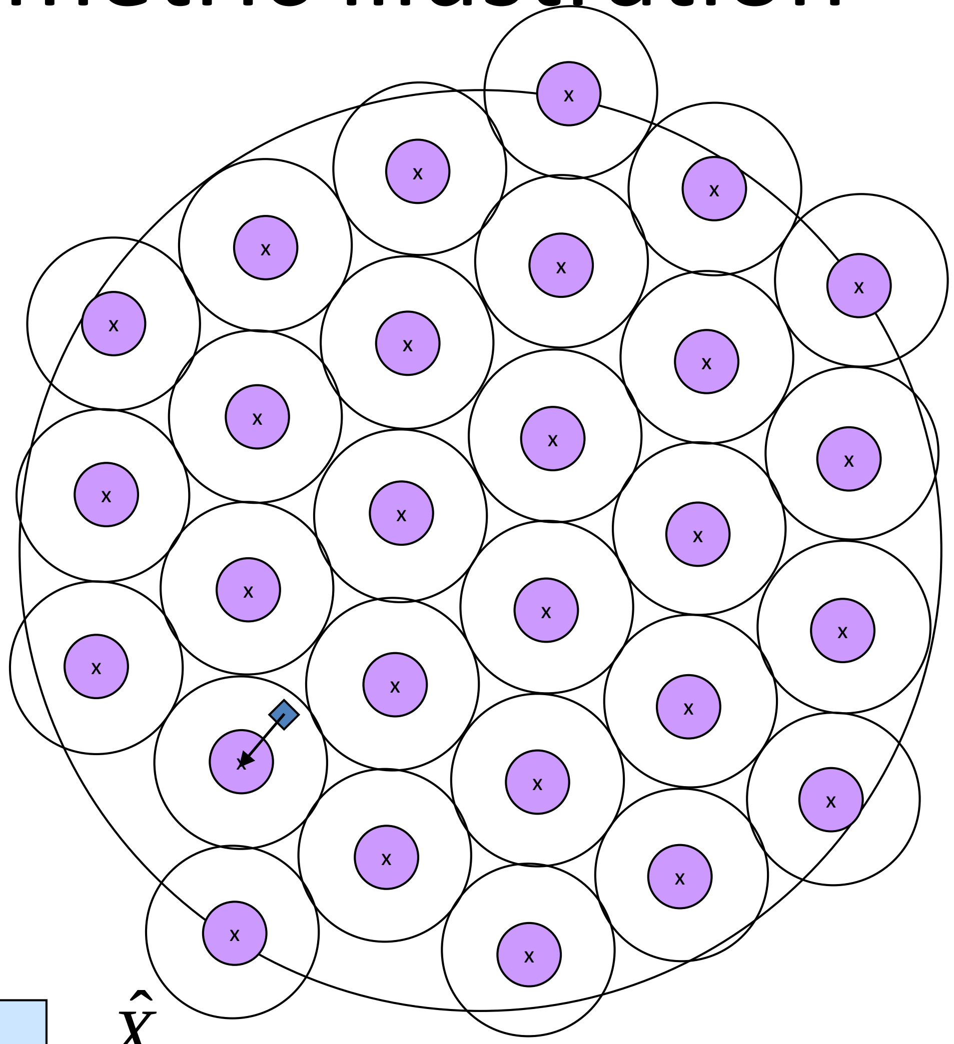
Example: geometric illustration



X

Side information K

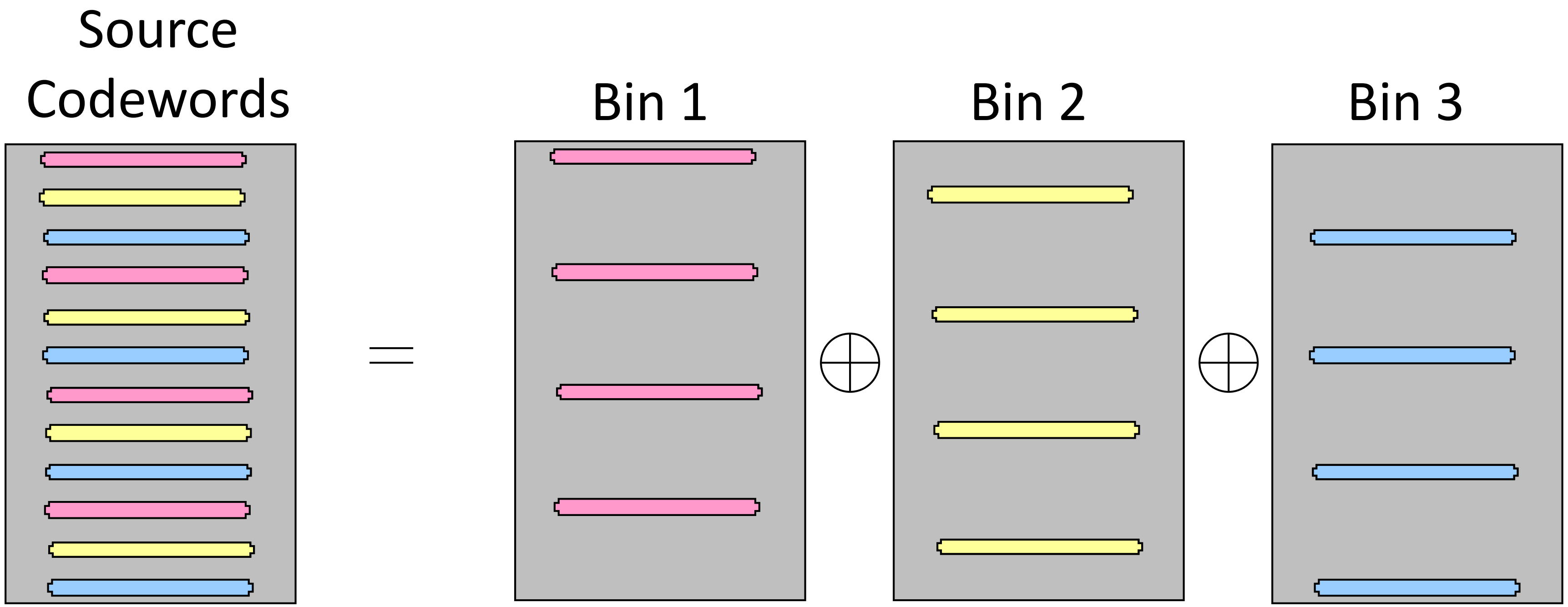




Practical Code Constructions

- Use a linear transformation (hash/bin)
- Design cosets to have maximal spacing
  - State of the art linear codes (LDPC codes)
- Distributed Source Coding Using Syndromes (DISCUS)*

**Pradhan & R, '03*



Orhan Ocal

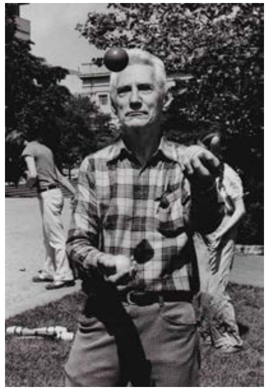
Xiao Li

Chapter 3

Sampling theory

- Sample and compute efficient sampling (and connections to learning)

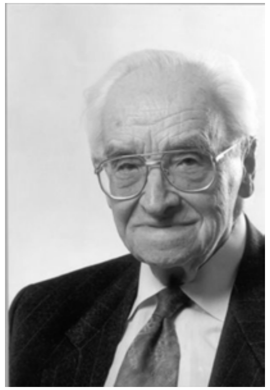
Sampling theorem



Shannon
1949

Nyquist
1928

Whittaker
1915



Kotelnikov
1933

Communication in the Presence of Noise

CLAUDE E. SHANNON, MEMBER, IRE

Theorem 1: If a function $f(t)$ contains no frequencies higher than W cps, it is completely determined by giving its ordinates at a series of points spaced 1/2 W seconds apart.

pointwise sampling!

...

Mathematically, this process can be described as follows. Let x_n be the nth sample. Then the function $f(t)$ is represented by

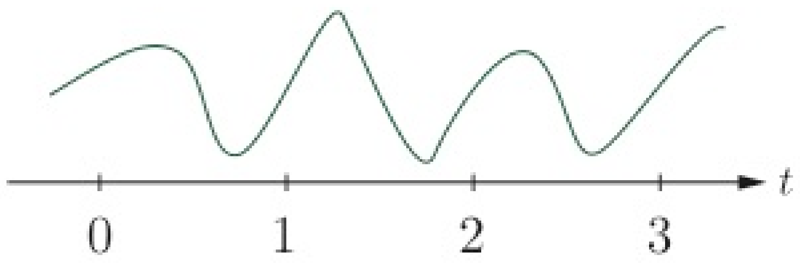
$$f(t) = \sum_{n=-\infty}^{\infty} x_n \frac{\sin \pi(2Wt - n)}{\pi(2Wt - n)}. \qquad (7)$$

linear interpolation!

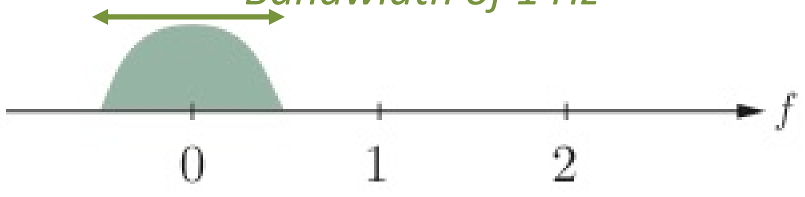
Aliasing phenomenon

Time domain | **Frequency domain**

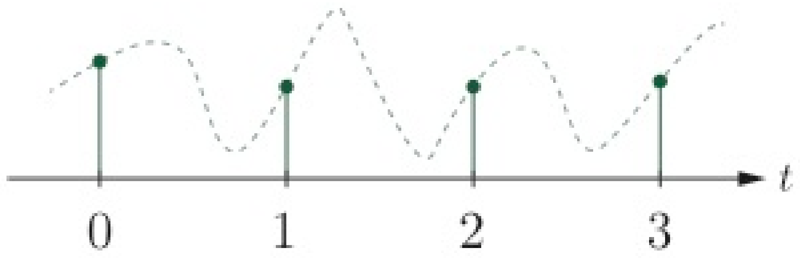
Input signal



Bandwidth of 1 Hz

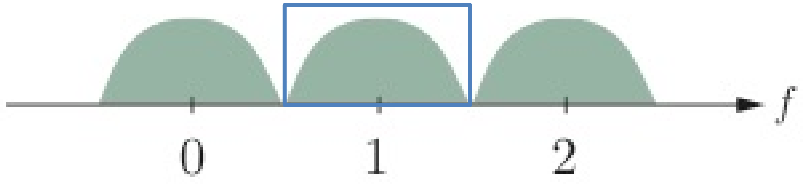


Sampling at rate 1

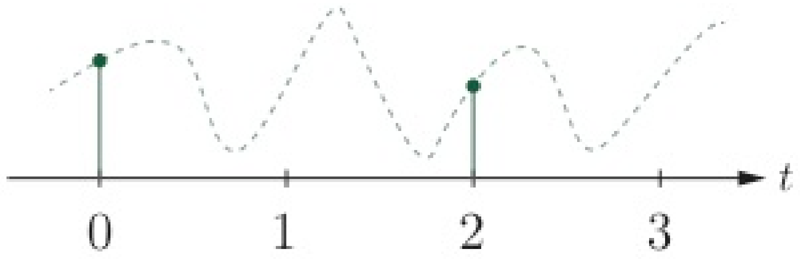


No aliasing

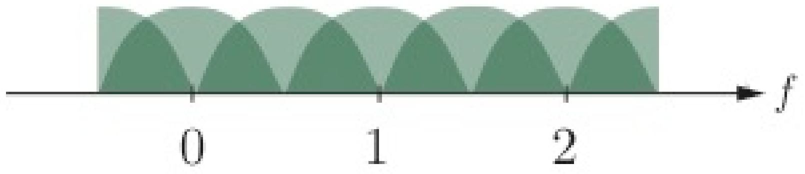
*– can recovery by **linear** filtering*



Sampling at rate 1/2

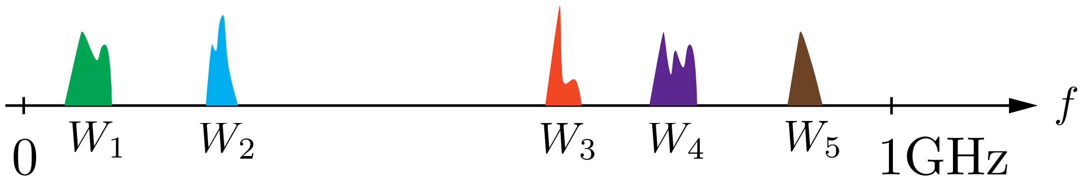


Spectrum is aliased!



But what if the spectrum is sparsely occupied?

Frequency domain



$$f_{occ} = \sum_{i=1}^{5} W_i = 100\text{MHz}$$

Henry Landau, 1967

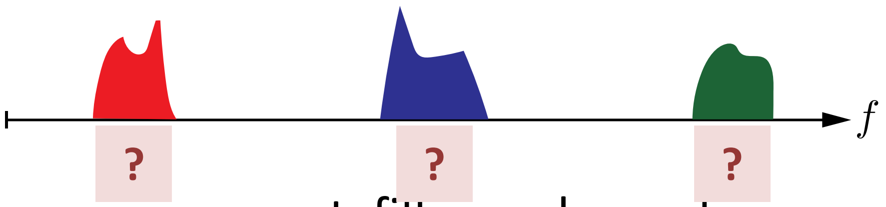
- Know the frequency support
- Sample at rate *"occupied bandwidth"* f_{occ} ***(Landau rate)***

When you do not know the support?

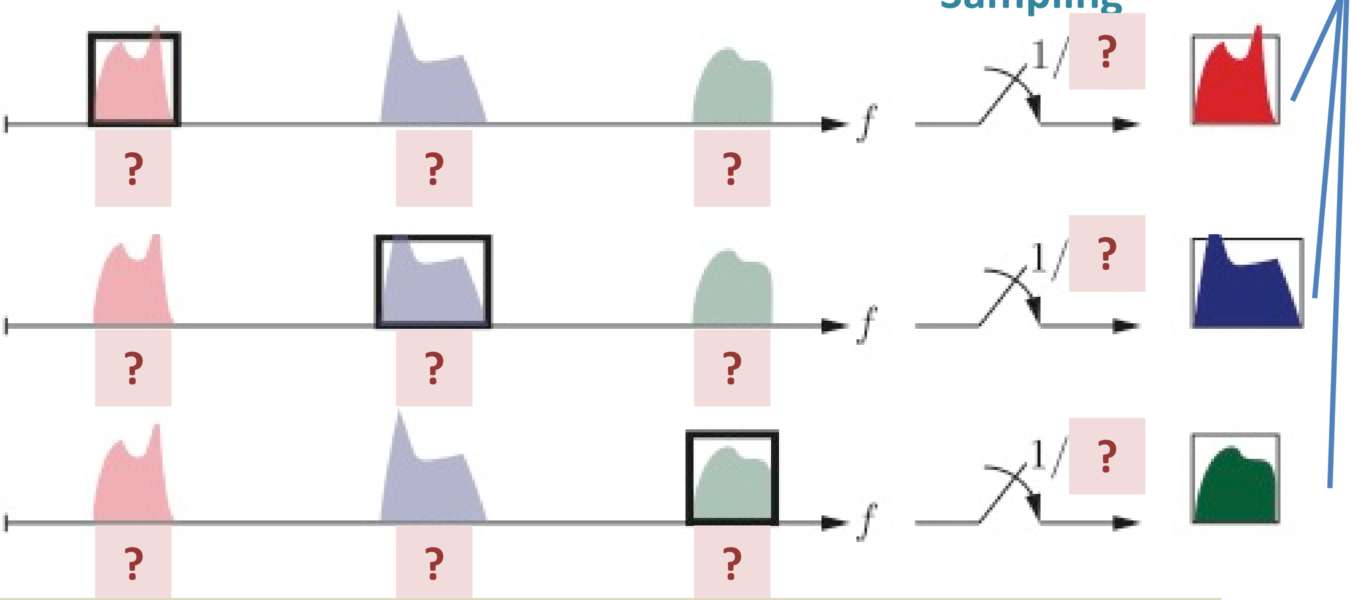
- *Feng and Bresler, 1996*
- *Lu and Do, 2008*
- *Mishali, Eldar, Dounaevsky and Shoshan, 2011*
- *Lim and Franceschetti, 2017*

Filter bank approach

Input in frequency domain



Know the frequency support, filter and sample



Sampling ***spectrum-blind?***

Requires $2f_{occ}$. ***Can we design a constructive scheme?***

Lu and Do, 2008

Puzzle: Gold thief

100 grams each

- One unknown thief
- Steals unknown but fixed amount from each coin
- What is min. no. of weighings needed ?
  - 2 are enough!

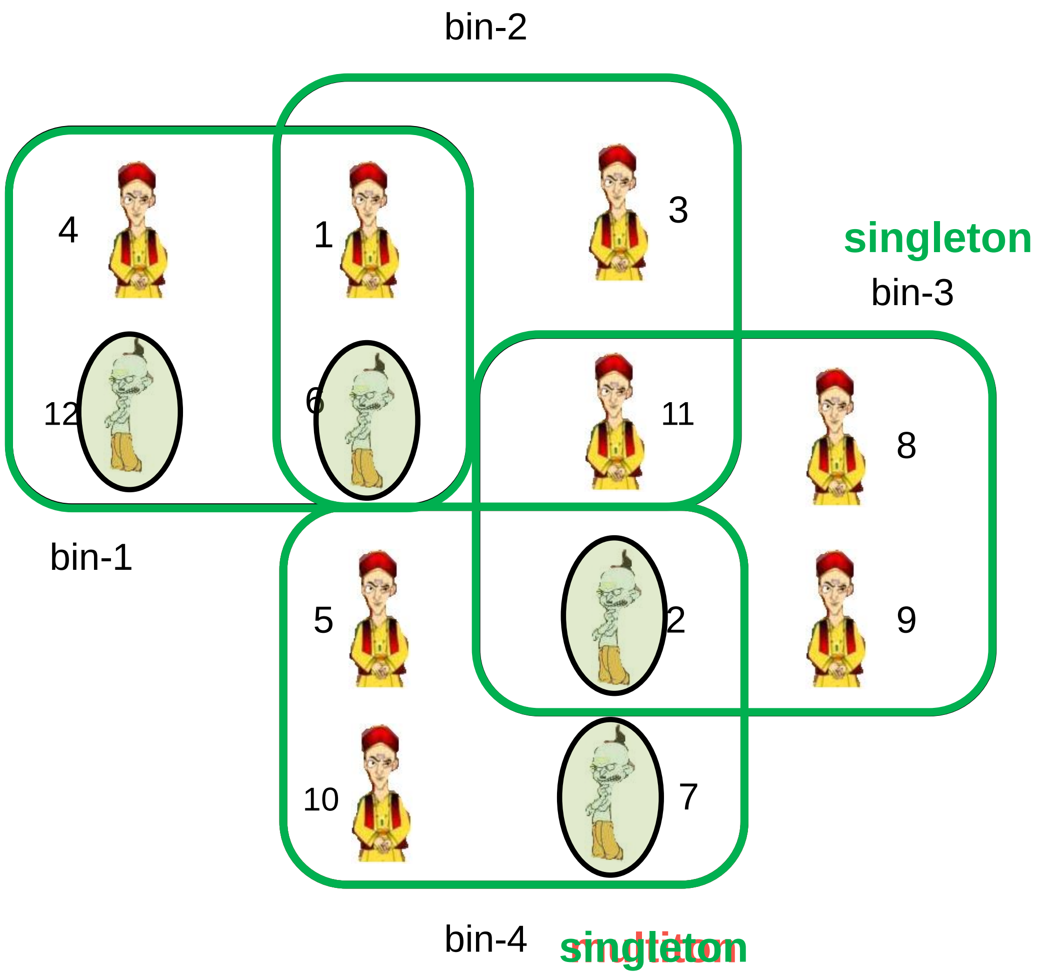
$\frac{y2}{y1}$ = 4

Differential weight

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix} = \begin{bmatrix} -5 \\ -20 \end{bmatrix} = \begin{bmatrix} y1 \\ y2 \end{bmatrix}$$

Ratio-test identifies the location

4-thieves among 12-treasurers



Key Ideas:

1. **Randomly group the treasurers.**
2. **If there is a single thief problem**
   - ✓ **Ratio test**
   - ✓ **Iterate.**

Questions:

1. **How many groups needed?**
2. **How to form groups?**
3. **How to identify if a group has a single thief?**

Main result

Any bandlimited signal $x(t) \in \mathbb{C}$ whose spectrum has occupancy f_{occ} can be sampled asymptotically at rate $f_s = 2f_{occ}$ by a randomized "*sparse-graph-coded filter bank*" with probability 1 using $O(f_{occ})$ operations per unit time.

Remarks

- Computational cost $O(f_{occ})$ ***independent of bandwidth***
- Requires mild assumptions (genericity)
- Can be made robust to sampling noise

Ocal, Li & R, 2016

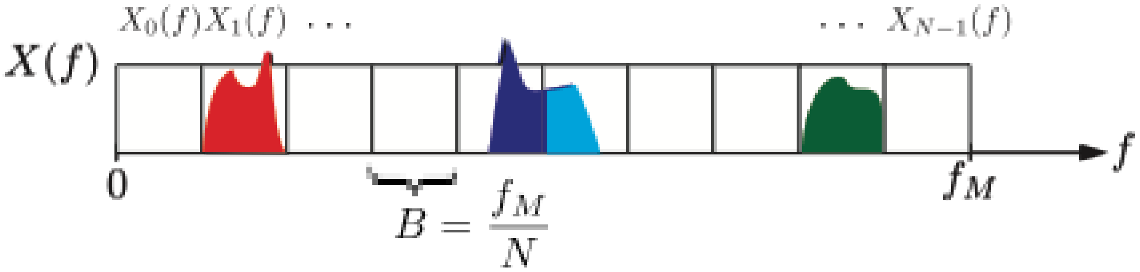
Key insight for spectrum-blind sampling

subsampling → aliasing

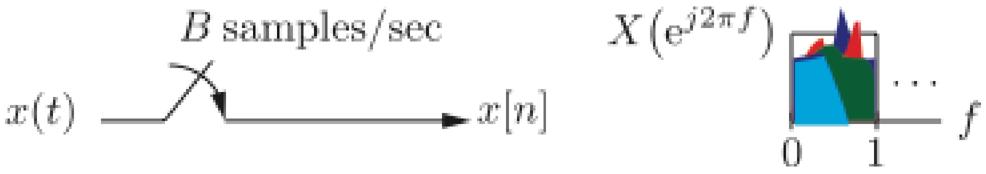
"judicious" filtering/subsampling → "good" aliasing

- To reduce sampling rate, ***subsample judiciously***
- ***Filter bank*** derived from ***capacity-achieving codes for the Binary Erasure Channel*** (BEC) (LDPC codes)
- Introduces aliasing (***structured noise***)
- ***Non-linear recovery*** instead of linear interpolation

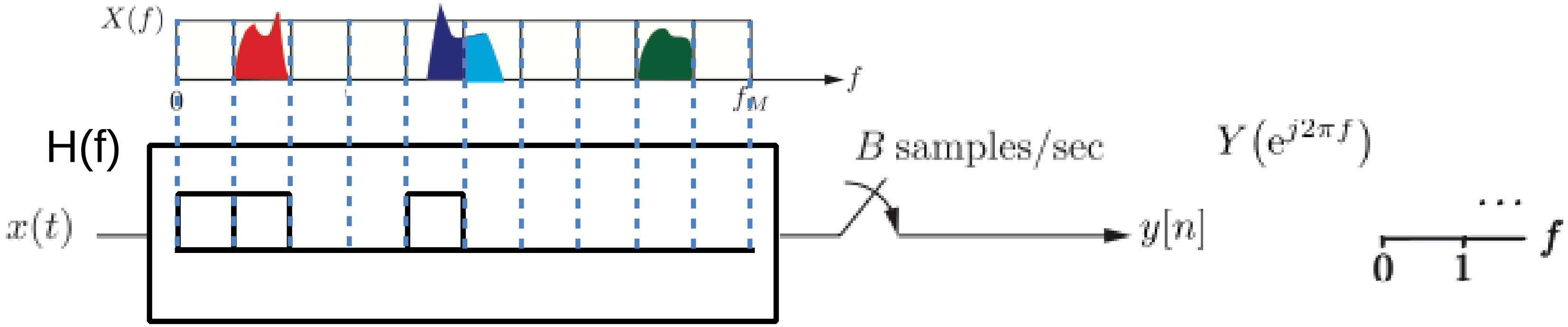
Filter bank for sampling



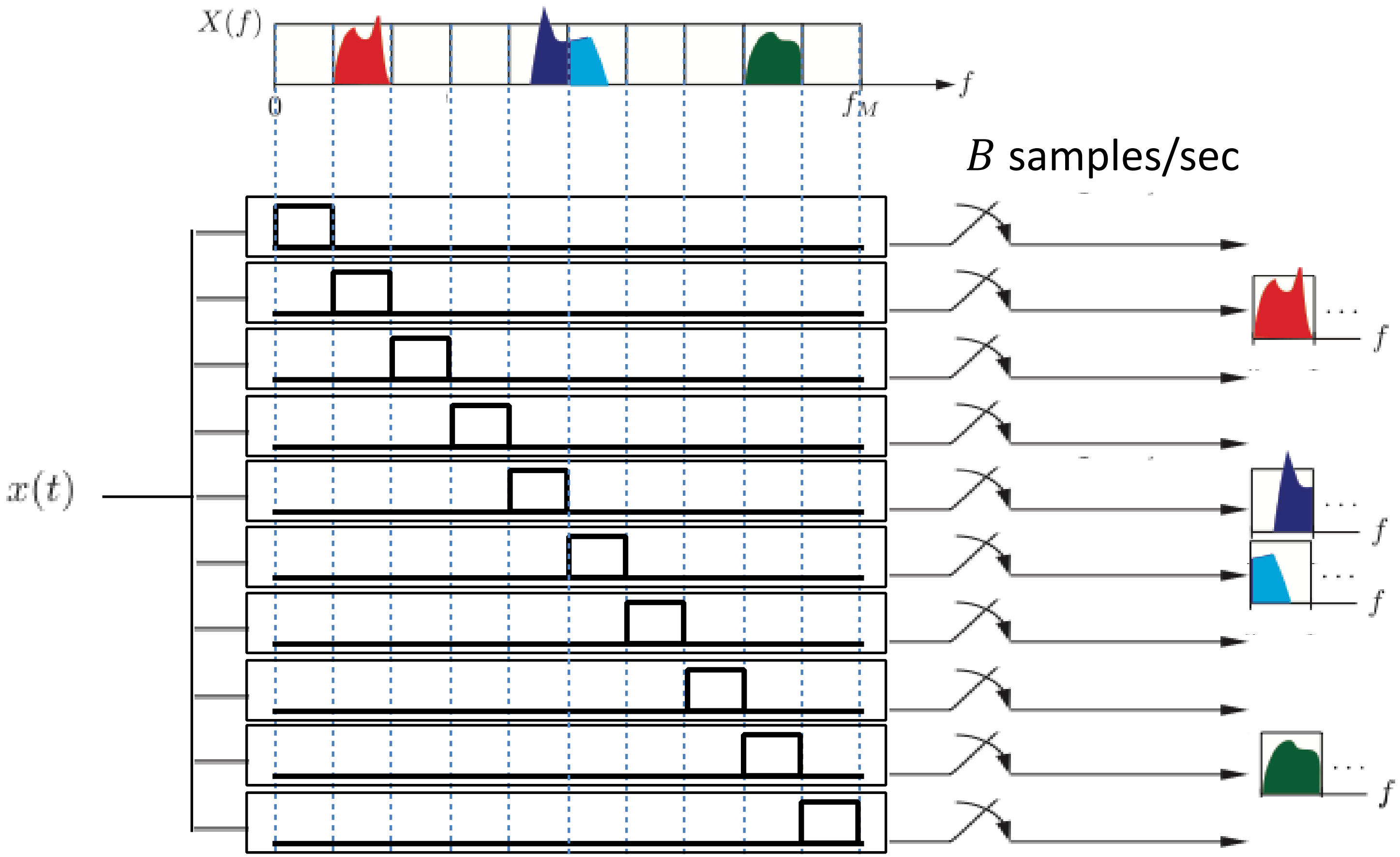
- Sample the signal at rate B



- **Filter and then sample at rate B**



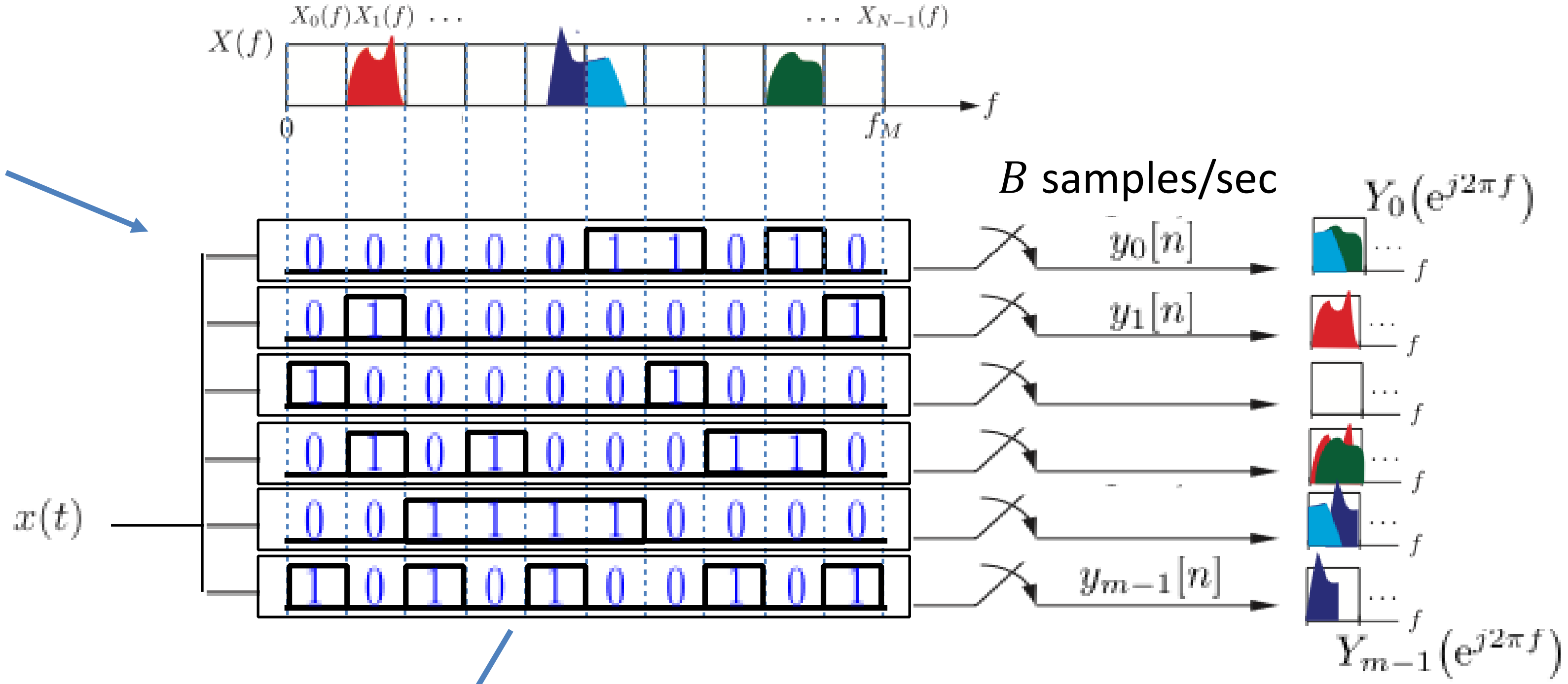
Filter bank for sampling



Aggregate sampling rate: $N\frac{f_M}{N} = f_M$ = Nyquist rate for $x(t)$

'*Sparse-graph-coded*' filter bank

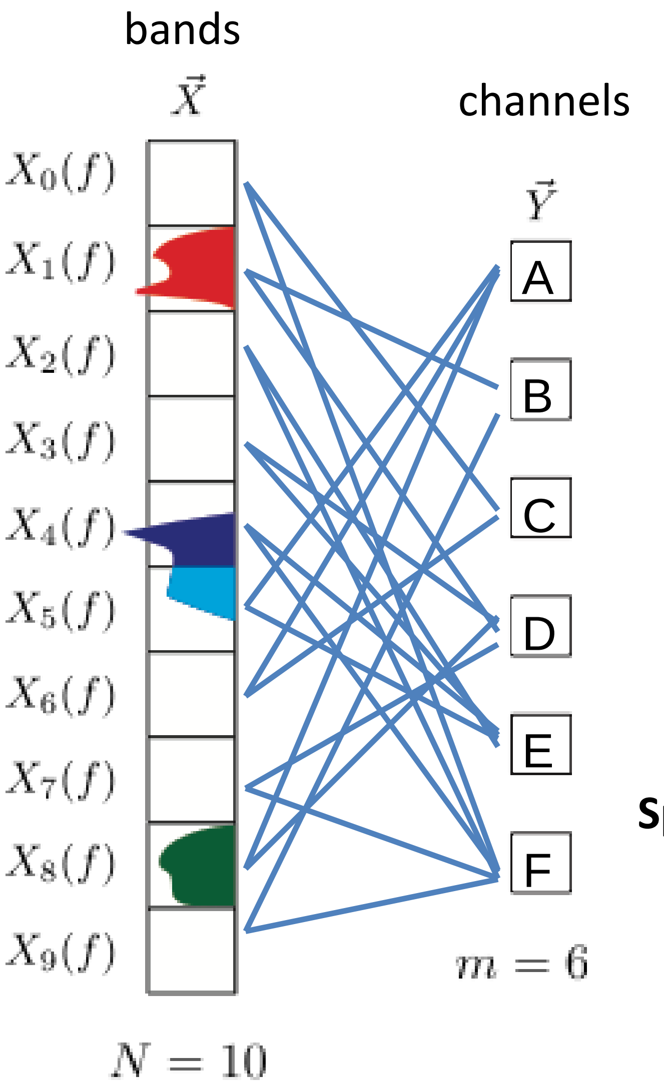
m filters
N bands



$$\vec{Y}(e^{j2\pi f}) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \end{pmatrix} \vec{X}(Bf) \quad \text{where} \quad \vec{X}(f) = \begin{pmatrix} X_0(f) \\ \vdots \\ X_{n-1}(f) \end{pmatrix}$$

$m \times N$ matrix

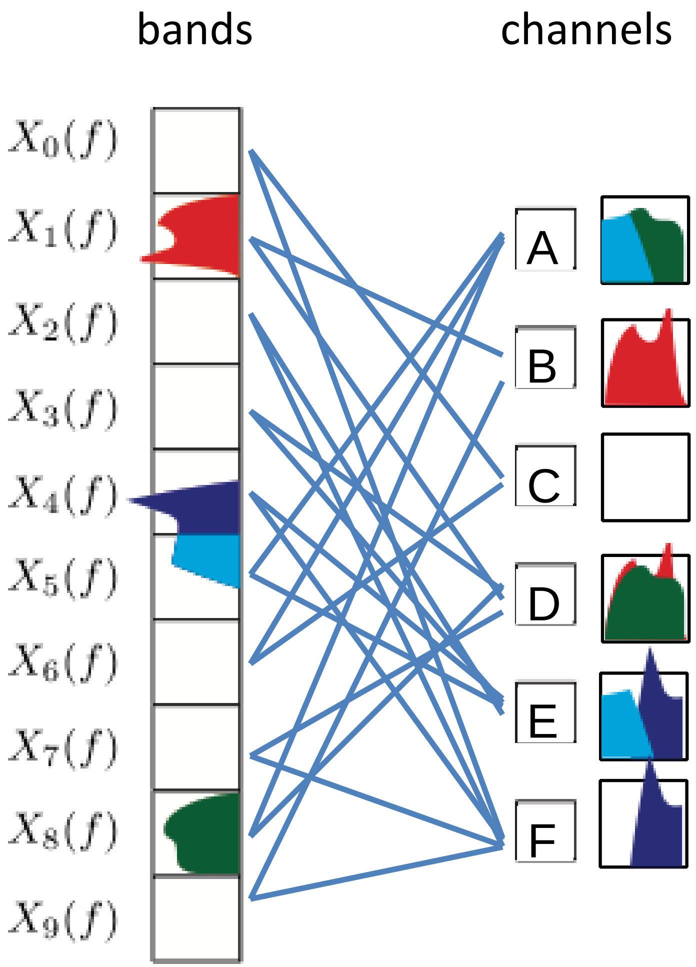
Example — sparse graph underlying the measurements



$$\vec{Y}(e^{j2\pi f}) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \end{pmatrix} \vec{X}(Bf)$$

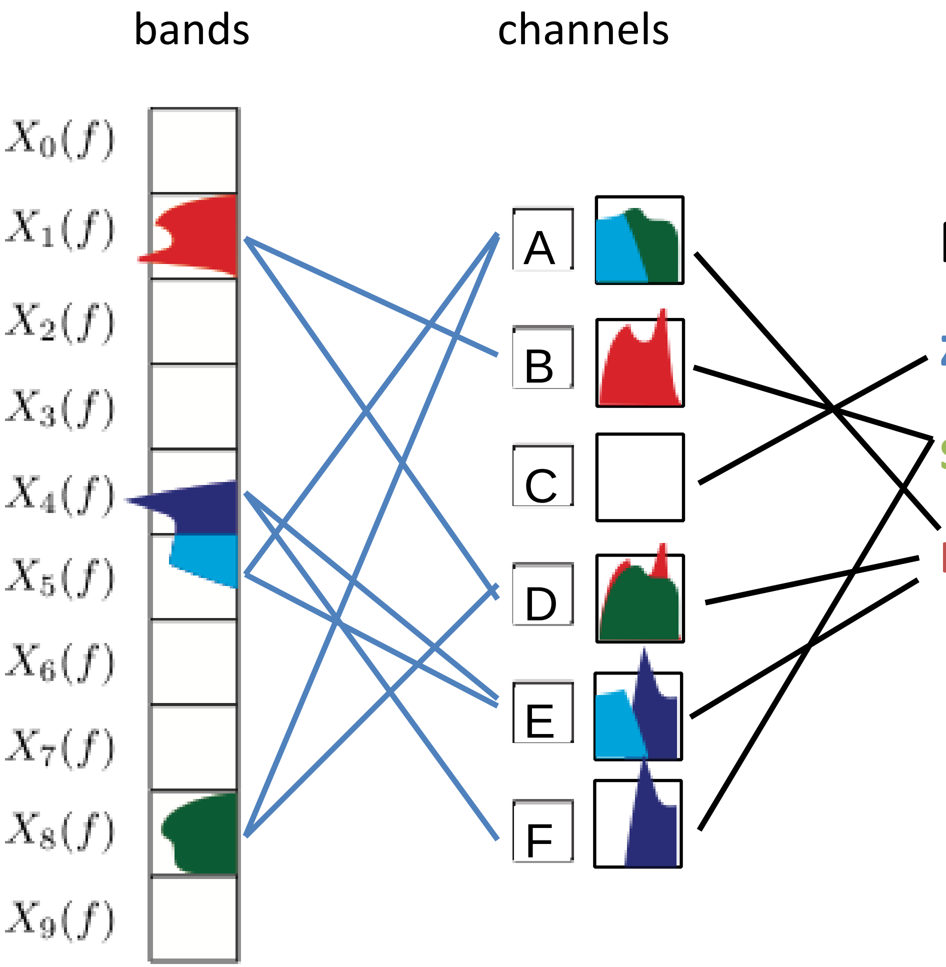
Sparse bipartite graph

Example — sparse graph underlying the measurements



visual cleaning for presentation: remove edges that connect to non-active bands

Example — peeling



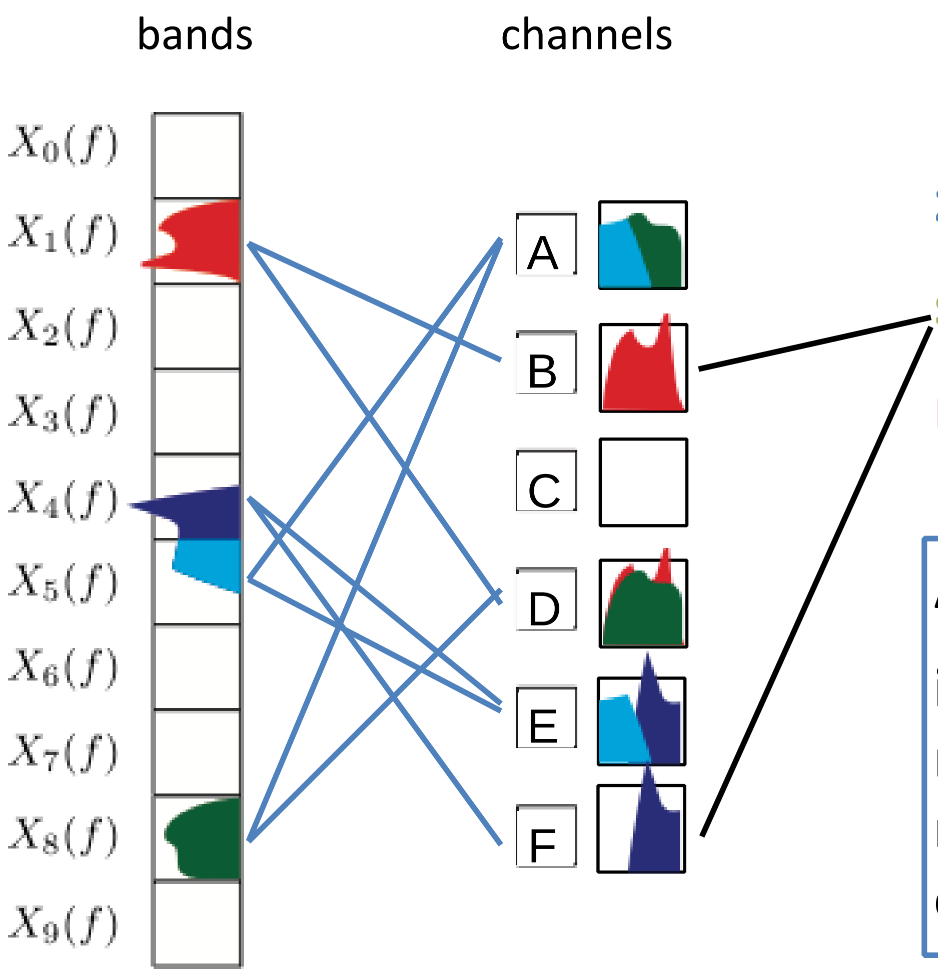
Measurement classification

zero-ton: no signal

single-ton: no aliasing

multi-ton: aliasing

Example — peeling



Measurement classification

zero-ton: **no signal**

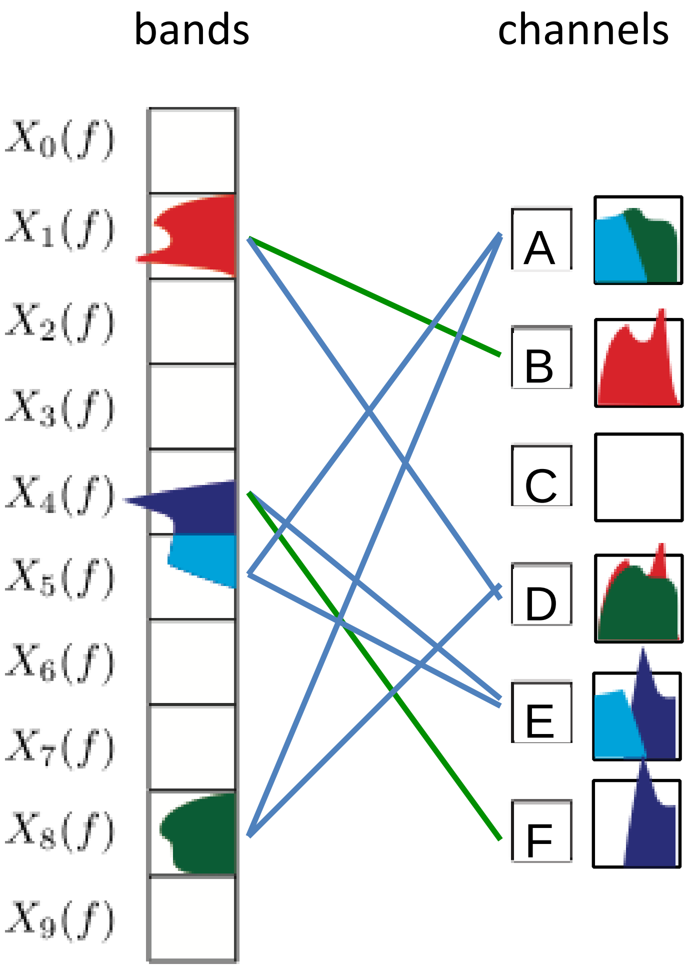
single-ton: **no aliasing**

multi-ton: **aliasing**

Assume a *mechanism*:

identifies which channels have no aliasing (here B and F) and maps them to which bands they came from (here 1 and 4 resp.)

Example — peeling



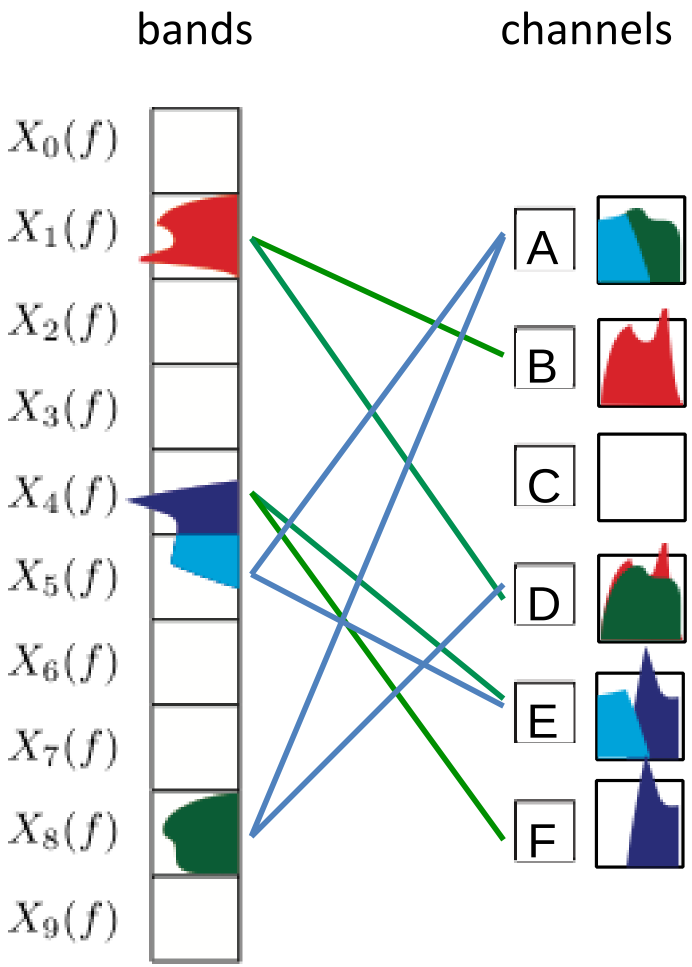
***mechanism*:**

identifies which channels have no aliasing and maps them to which bands they came from

output:

channel B: (red, index = 1)
channel F: (blue, index = 4)

Example — peeling



***mechanism*:**

identifies which channels have no aliasing and maps them to which bands they came from

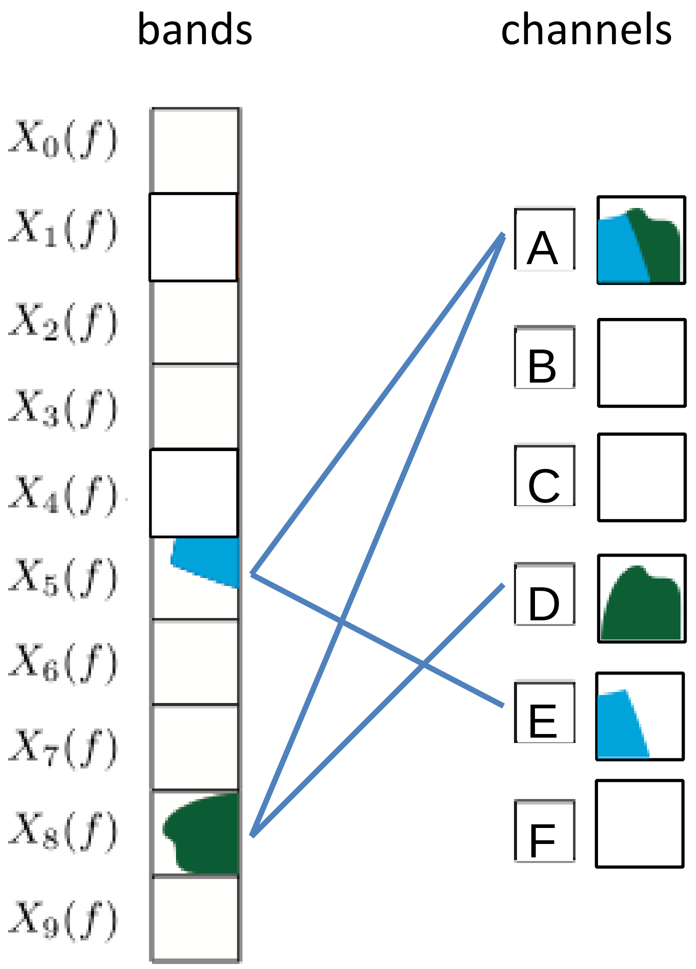
output:

channel B: (red, index = 1)

channel F: (blue, index = 4)

peel from channels they alias into!

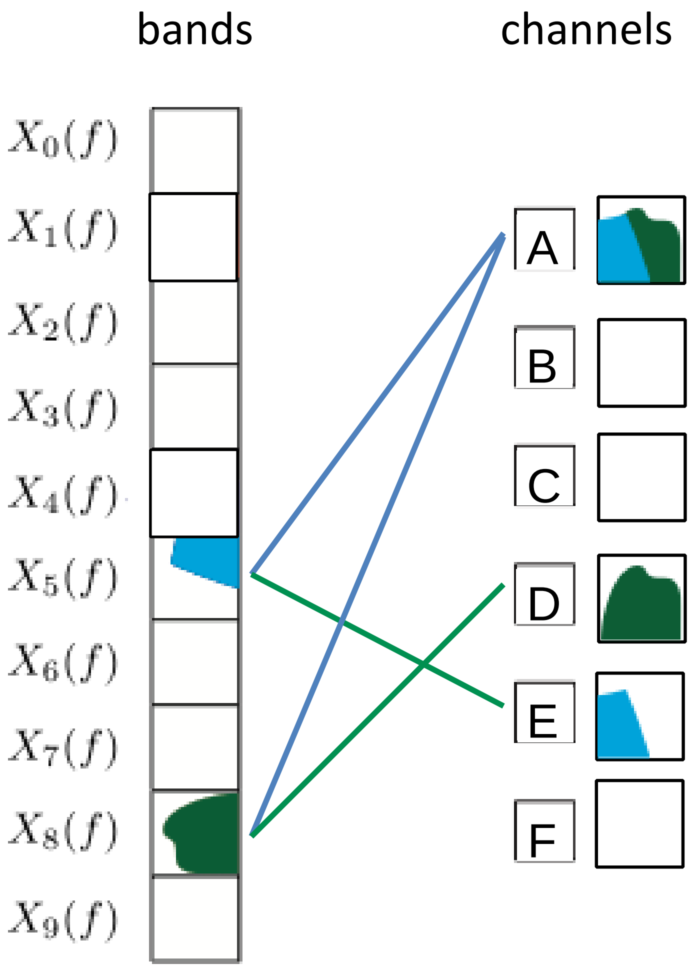
Example — peeling



***mechanism*:**

identifies which channels have no aliasing and maps them to which bands they came from

Example — peeling



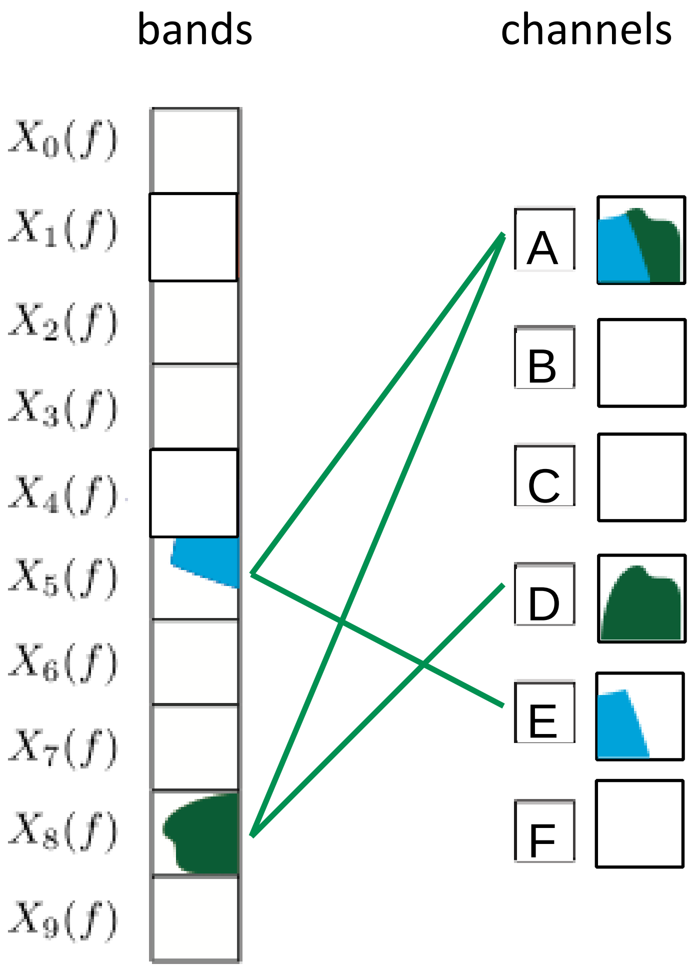
***mechanism*:**

identifies which channels have no aliasing and maps them to which bands they came from

output:

channel D: (green, index = 8)
channel E: (cyan, index = 5)

Example — peeling



***mechanism*:**

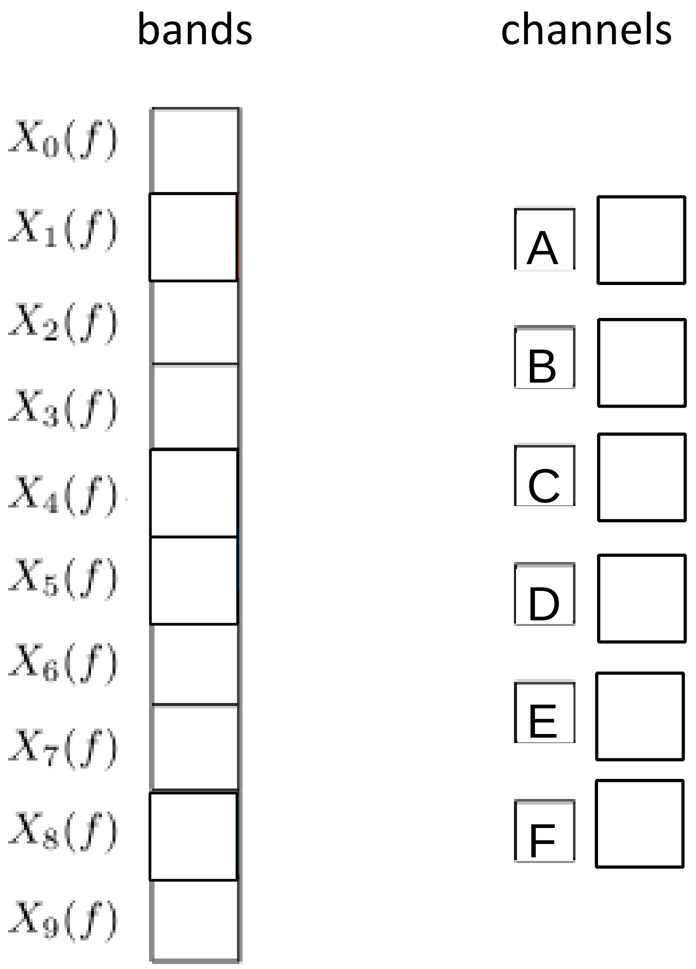
identifies which channels have no aliasing and maps them to which bands they came from

output:

channel D: (green, index = 8)
channel E: (cyan, index = 5)

peel from channels they alias into!

Example — peeling

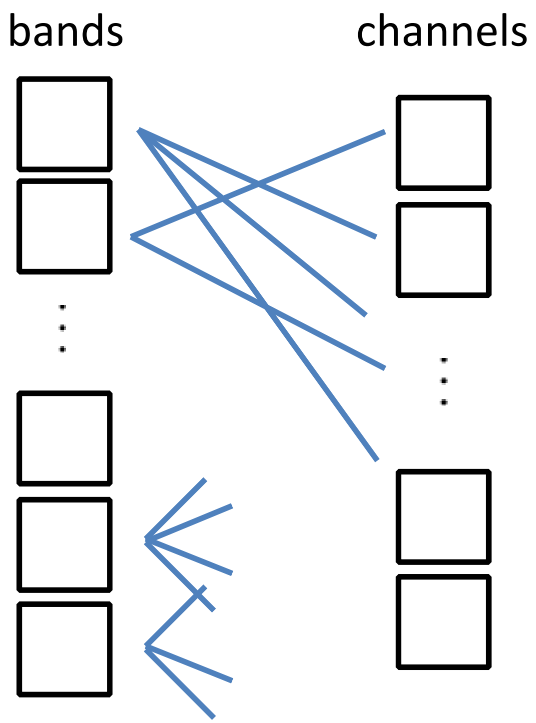


***mechanism*:**

identifies which channels have no aliasing and maps them to which bands they came from

signal is completely recovered!

Construction of the sparse-graph code

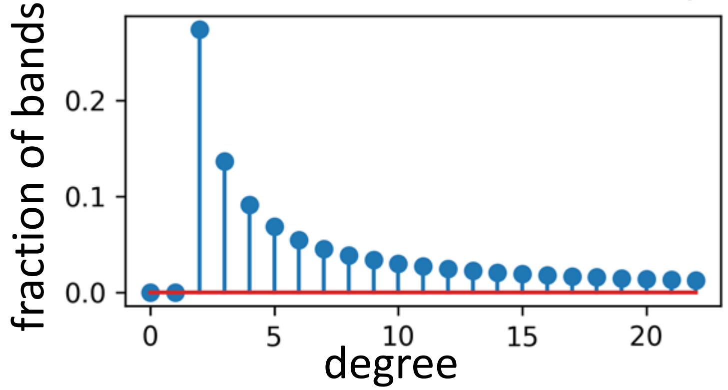


$$P(degree = j) \propto \frac{1}{j-1} \text{ for j=2,3,…D}$$

$$D > 1/\epsilon$$

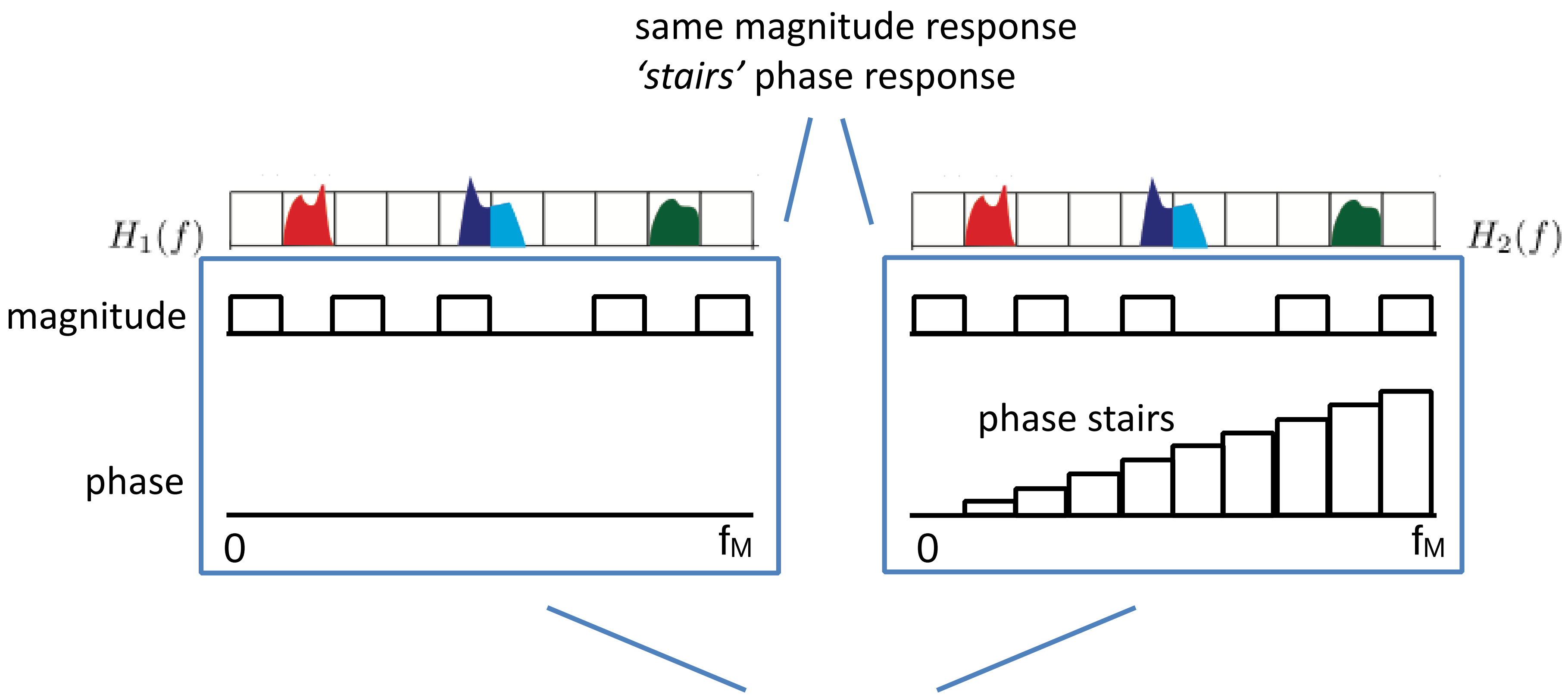
- Designed through ***capacity-approaching sparse-graph codes***
- Connect each ***band*** to ***channels*** at random according to a carefully chosen degree distribution.
- Asymptotically, ***number of channels*** is $(1 + \epsilon)$ times the ***number of active bands***

Degree distribution for $\epsilon = 1/20$



Realizing the *mechanism*

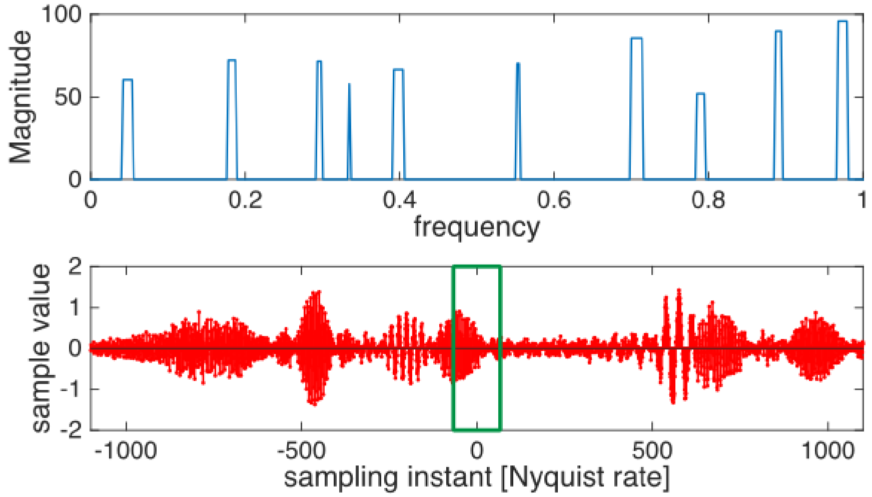
Identify which channels have no aliasing and map them to bands



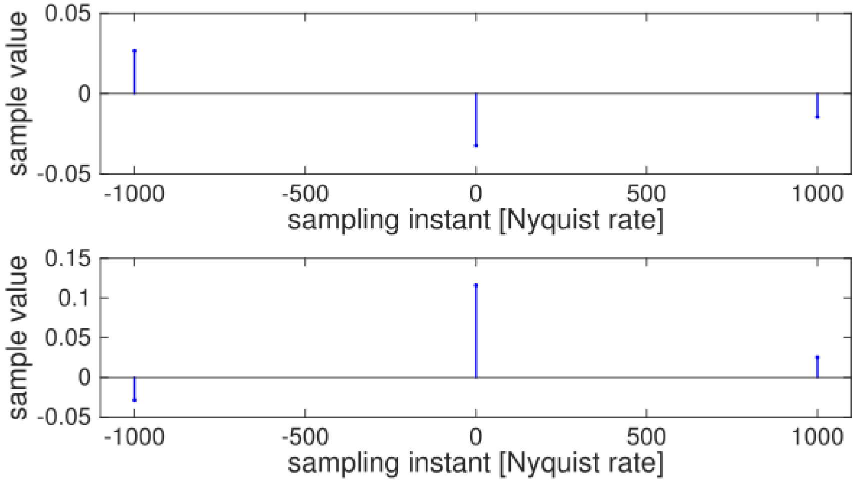
identifies dark blue band as a singleton

Numerical experiment

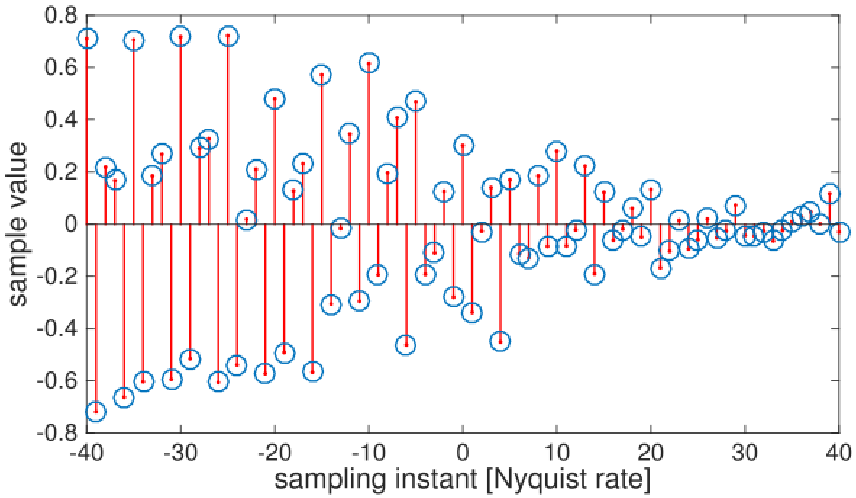
Input spectrum and time domain signal



Output from two sample channels

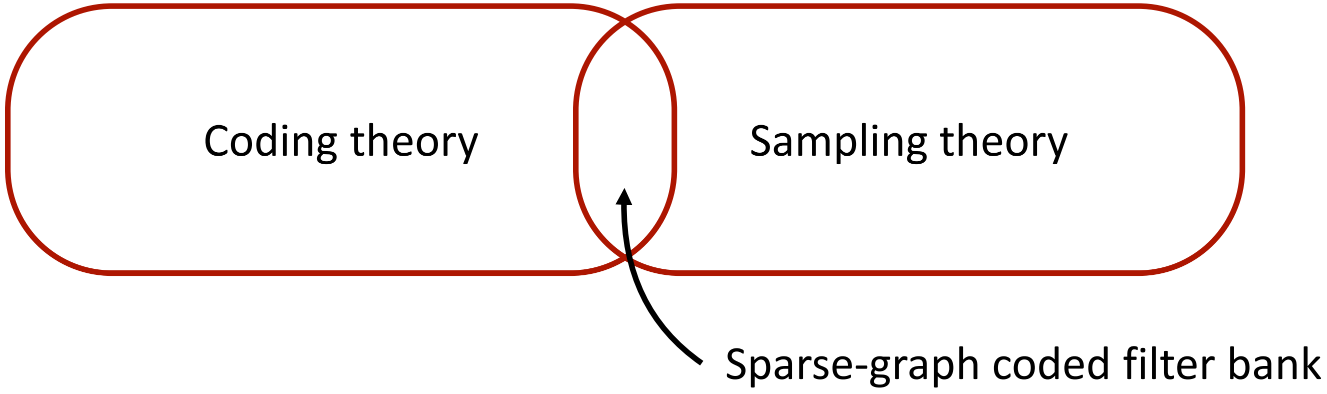


true signal ○ estimates

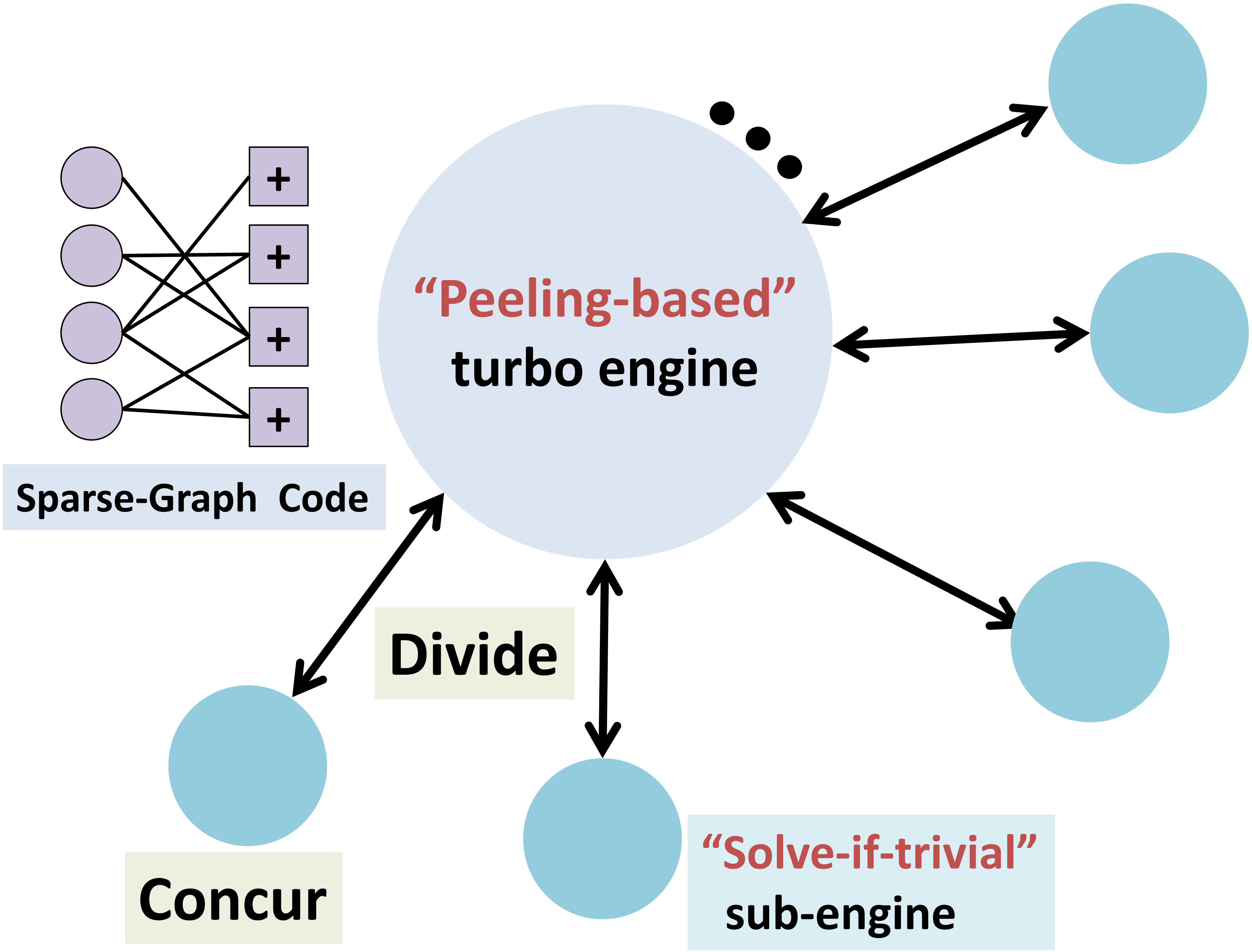


- Lebesgue measure $f_L = 0.1$
- Number of slices $N = 1000$
- Number of channels $M = 284$
- Sampling rate $f_S = 0.284$

Interesting connection

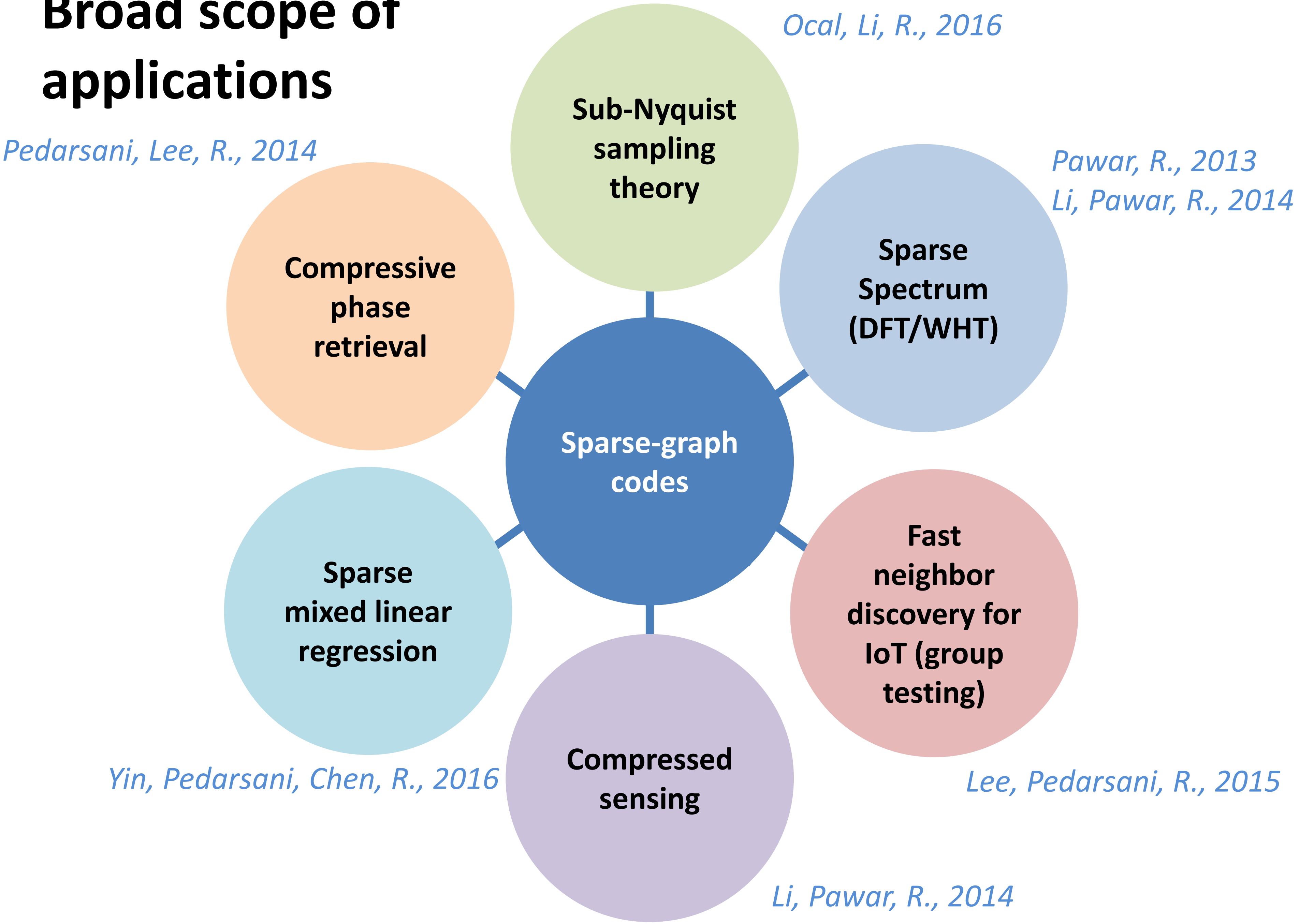


- *Minimum-rate spectrum-blind* sampling
- *Coding theory* and *sampling theory*
  - Capacity-approaching codes for erasure channels
  - Filter banks that approach Landau rate for sampling



+
+
+
+
Sparse-Graph Code
"Peeling-based"
turbo engine
Divide
Concur
"Solve-if-trivial"
sub-engine

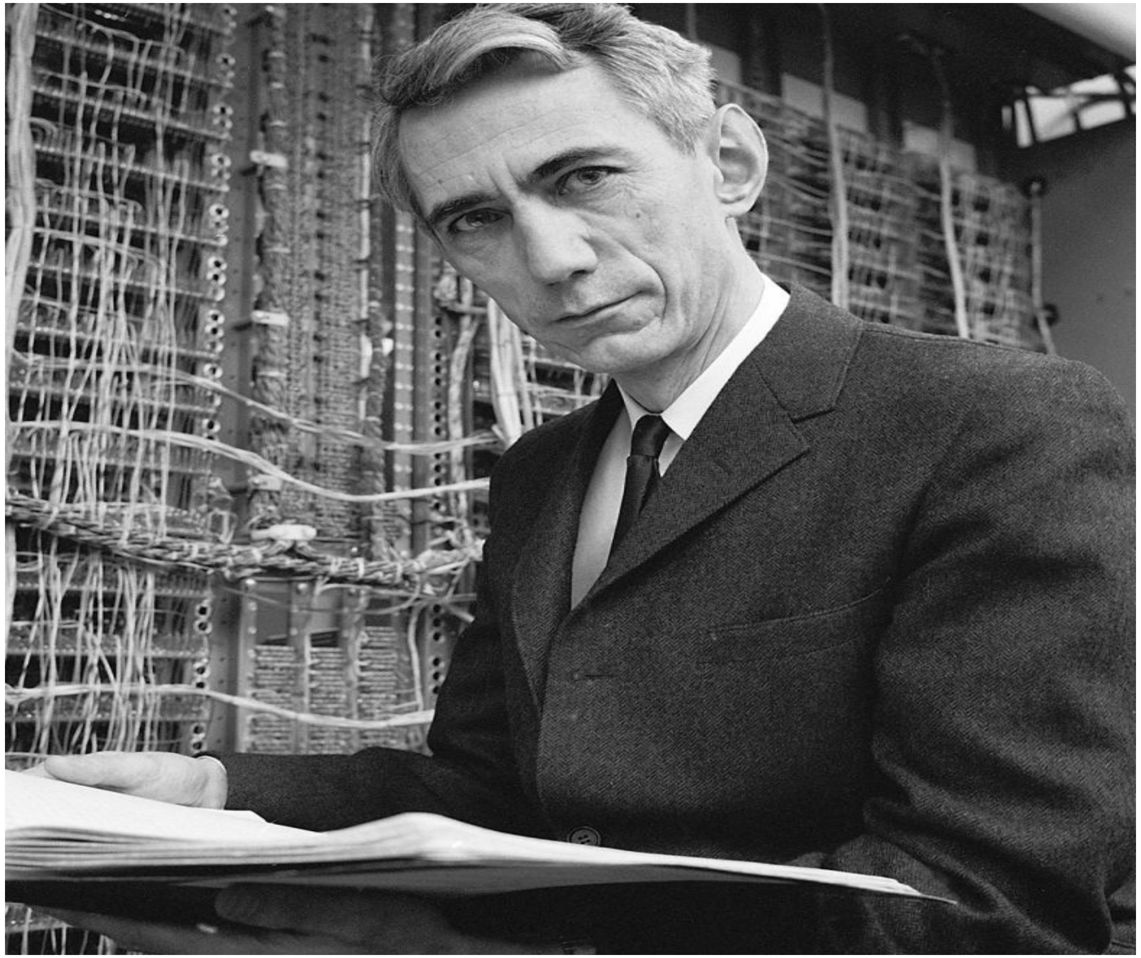
Broad scope of applications



Conclusion: Shannon's incredible legacy

- A mathematical theory of communication
- Channel capacity
- Source coding
- Channel coding
- Cryptography
- Sampling theory
- …

His legacy will last many more centuries!



(1916-2001)

Thank you!