Information Theory in Physics Applications

TA5-WP3 Workshop on Dynamic Archiving

Vladimir Lenok

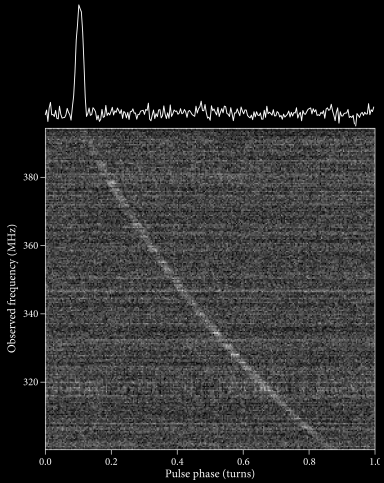
January 18, 2024

Bielefeld University, Faculty of Physics

UNIVERSITÄT
BIELEFELD

PUNCH
4 N F D I

Introductory Example — Pulsar Signal Observation (Single Pulse)



- Time series itself does not bring information (information is in the modulation of the radio wave)
- Pieces of information:
    - dispersion measure
    - spectral shape

Basic terminology

- **Information** — knowledge or data that are subject of some operations
- **Message** — a form of presenting information for its storage, processing, transformation, or direct use
- **Signal** — a form of information representation for transmission via a channel
- **Channel** — a set of means of information transmission that includes the physical medium

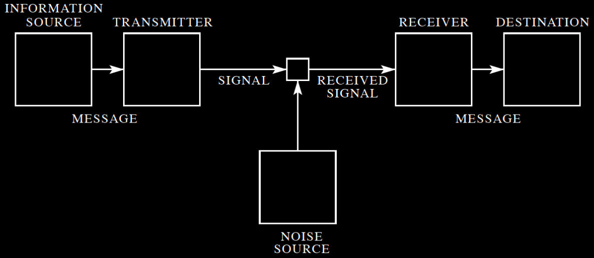
Distress calls (information → message → signal → channel):

“Disaster is happening” → SOS (· · · — — — · · ·) → modulated radio waves → medium → ...

“Disaster is happening” → MAYDAY MAYDAY MAYDAY→ modulated radio waves → medium → ...

Shannon Information Theory (I)

- comes from the work on cryptography and secrecy systems (Bell Labs, 1940s)
- operates with discrete “alphabet” (i.e., a finite set over which a discrete sequence is defined)
- uses statistical approach



The focus of the theory is on how to encode the messages for efficient transmission

Shannon Information Theory (II) — Abstract Formulation (Kolmogorov)

$$\overbrace{\xi \longrightarrow \eta}^{\text{encoding}} \mkern-18mu \underbrace{ \longrightarrow \eta'}_{\text{transmission}} \mkern-18mu \overbrace{ \longrightarrow \xi'}^{\text{decoding}}$$

- $\eta \longrightarrow \eta' : P_{\eta'|\eta}(B'|y) = P(\eta' \in B'|\eta = y)$ (transmission)
- $\xi \longrightarrow \eta : P_{\eta|\xi}(B|x) = P(\eta \in B|\xi = x)$ (encoding)
- $\eta' \longrightarrow \xi' : P_{\xi'|\eta'}(A'|y') = P(\eta' \in A'|\eta' = y')$ (decoding)
- $P_\xi(A) = P(\xi \in A)$ (incoming messages)
- ξ, η, η', ξ' are Markov sequences from fixed alphabets.

The main results are for asymptotic cases of stationary processes

Shannon Information Theory (III) — Definition of Term “Amount of Information”

List of Kolmogorov:

- Combinatorial approach
- Probabilistic approach

} Shannon information theory

- Algorithmic approach (out of scope for now, may be not “practical”)

Useful concepts:

- Entropy as a concept for “amount of information” estimation (additivity principle)
- Random and averaged entropy

Features of the Shannon Scheme and Theory

$$\overbrace{\xi \longrightarrow \rlap{$\underbrace{\eta \longrightarrow \eta'}_{\text{transmission}}$}\eta}^{\text{encoding}} \longrightarrow \overbrace{\eta' \longrightarrow \xi'}^{\text{decoding}}$$

- Encoding and decoding are ciphering
- The scheme works on alphabet level
- We know how the transmission works (how to transform a message to a signal and back) (!)
- Amount of information is not conserved in transformations

Observations in Radio Astronomy (I)

Observation scheme for pulsar-like signals (e.g., pulsars and FRB)

$$\overbrace{\xi \longrightarrow }^{\text{encoding}}\!\!\!\underbrace{\eta \longrightarrow \text{modulation} \longrightarrow \text{propagation} \longrightarrow \text{demodulation} \longrightarrow \eta'}_{\text{transmission}}\!\!\!\overbrace{ \longrightarrow \xi'}^{\text{decoding}}$$

Problems:

- There is no alphabet. Information is encoded differently: in radio wave modulation
- Distributions are continuous, but not discrete (minor problem)
- Signals are highly dispersed ($\therefore$ amount of information is zero in the basic time domain case)

What do we know?

- Class of the signals (chirp signals)
- Searching for “alphabet” = correct/optimal signal processing (optimal demodulation)

Observations in Radio Astronomy (II) — Pulsar Observation as Example

$$\overbrace{\xi \longrightarrow \eta}^{\text{encoding}}\kern-0.9em\underbrace{ \longrightarrow \text{modulation} \longrightarrow \text{propagation} \longrightarrow \text{demodulation} \longrightarrow }_{\text{transmission}}\kern-0.9em\overbrace{\eta' \longrightarrow \xi'}^{\text{decoding}}$$

The radio signal by itself does not bring any information

Pieces of information (message) from a single pulse (η')

- Dispersion measure (type of the phase modulation and its parameter)
- Shape of spectrum (amplitude modulation)

This message is followed by interpretation

Statistical Signal Processing

Bayesian signal detection is the foundation

minimal risk = maximum likelihood ratio = matched filter

Summary

- Information theory shows how to compute useful information related quantities
- There is no magic quantity that shows amount of information in arbitrary case
- Signal itself does not bring any information (requires optimal demodulation/processing)
- Problem of absence of “alphabet” (search for optimal transformations/demodulations)
- Non-optimal transformations lead to losing of information (amount of information $\neq$ energy)

Solution $\rightarrow$ Statistical signal processing + Information theory

Discussion point — General Observation Scheme

$$\rlap{$\overbrace{}^{\text{encoding}}$}\xi \longrightarrow \underbrace{\eta \longrightarrow \text{modulation} \longrightarrow \text{propagation} \longrightarrow \text{demodulation} \longrightarrow \eta'}_{\text{transmission}} \llap{$\overbrace{}^{\text{decoding}}$}\!\!\!\!\!\!\!\!\!\! \longrightarrow \xi'$$

The problem:

Formulate the observational scheme appearing in physics applications in a way that incorporates the known developments from statistical signal processing and information theory

Why should we do it this way?

- These theories are highly developed and already provided practically useful results. (We should not reinvent the wheel.)
- Individual components of these theories seem to describe well the components of data analysis developed independently in radio astronomy and other fields.