Shannon's Formula & Hartley's Rule:

Olivier Rioul* José Carlos Magossi†

*Télécom ParisTech †Unicamp

claude shannon

claude shannon

a sound channel

a sound channel

Shannon's formula:

$$\boxed{C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)} \quad \text{bits/symbol}$$

a sound channel

Shannon's formula:

or... $$C = W \log_2\left(1 + \frac{P}{N}\right)$$ bits/second

Claude Shannon

Shannon's formula:

$$\boxed{C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)}$$

"A Mathematical Theory of Communication," *The Bell System Technical Journal*, Vol. 27, pp. 623–656, October, 1948.

Claude Shannon

Shannon's formula:

$$\boxed{C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)}$$

"A Mathematical Theory of Communication," *The Bell System Technical Journal*, Vol. 27, pp. 623–656, October, 1948.

In the end, "The Mathematical Theory of Communication," [1] and the book based on it [25] came as a bomb, and something of a delayed-action bomb.

Claude Shannon

Shannon's formula:

$$C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)$$

"A Mathematical Theory of Communication," *The Bell System Technical Journal*, Vol. 27, pp. 623–656, October, 1948.

Note on the Theoretical Efficiency of Information Reception with PPM*

For small P/N ratios, the now classical expression for the information reception capacity of a channel

$$C = W \lg_2 (1 + P/N)$$

can be written, substituting kTW for N,

$$CT_0 = WT\, P/N \lg_2 e = \frac{PT_0}{kT} \lg_2 e = \frac{E}{kT} \lg_2 e$$

* Received by the Institute, February 23, 1949.

Ralph Hartley

20 years before... in the same journal...

Ralph Hartley

20 years before... in the same journal...

Hartley's rule:

$$\boxed{C' = \log_2\left(1 + \frac{A}{\Delta}\right)} \quad \text{bits/symbol}$$

"Transmission of Information," *The Bell System Technical Journal*, Vol. 7, pp. 535–563, July 1928.

Ralph Hartley

20 years before... in the same journal...

Hartley's rule:

or...
$$\boxed{C' = 2W \log_2\left(1 + \frac{A}{\Delta}\right)}$$
bits/second

"Transmission of Information," *The Bell System Technical Journal*, Vol. 7, pp. 535–563, July 1928.

Ralph Hartley

Hartley's rule:

$$C' = \log_2\left(1 + \frac{A}{\Delta}\right)$$

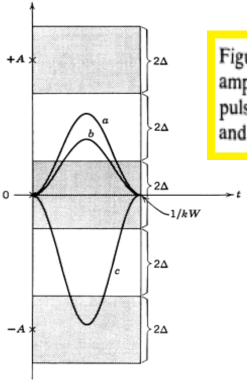


Figure 1.1 Distinguishable receiver amplitudes. Hartley considered received pulse amplitudes to be distinguishable only if they lie in different zones of width 2Δ. Thus pulses a and c are distinguishable but a and b are not. For the case shown, $A/\Delta = 4$ and there are five distinguishable zones.

(Wozencraft-Jacobs textbook, 1965)

Ralph Hartley

Hartley's rule:

$$\boxed{C' = \log_2\left(1 + \frac{A}{\Delta}\right)}$$

- amplitude "SNR" A/Δ (factor 1/2 is missing)
- no coding involved (except quantization)
- zero error

Hartley's formulation exhibits a simple but somewhat inexact interrelation among the time interval T, the channel bandwidth W, the maximum signal magnitude A, the receiver accuracy Δ, and the allowable number M of message alternatives. Communication theory is intimately concerned with the determination of more precise interrelations of this sort.

(Wozencraft-Jacobs textbook, 1965)

Outline

Hartley's $C' = \log_2\left(1 + \frac{A}{\Delta}\right)$ came 20 years before Shannon

Outline

Hartley's $C' = \log_2\left(1 + \frac{A}{\Delta}\right)$ came 20 years before Shannon

Shannon's $C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)$ came unexpected in 1948

Outline

Hartley's $C' = \log_2\left(1 + \frac{A}{\Delta}\right)$ came 20 years before Shannon

Shannon's $C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)$ came unexpected in 1948

Hartley's rule is inexact: $C' \neq C$

Outline

Hartley's $C' = \log_2\left(1 + \frac{A}{\Delta}\right)$ came 20 years before Shannon

Shannon's $C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)$ came unexpected in 1948

Hartley's rule is inexact: $C' \neq C$

Besides, C' is not the capacity of a noisy channel

Outline

Hartley's $C' = \log_2\left(1 + \frac{A}{\Delta}\right)$ came 20 years before Shannon

Shannon's $C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)$ came unexpected in 1948

Hartley's rule is inexact: $C' \neq C$

Besides, C' is not the capacity of a noisy channel

(no question)

Outline

Wrong!

Outline

This Hartley's rule $C' = \log_2\left(1 + \frac{A}{\Delta}\right)$ is not Hartley's

Outline

This Hartley's rule $C' = \log_2\left(1 + \frac{A}{\Delta}\right)$ is not Hartley's

Many authors independently derived $C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)$ in 1948.

Outline

This Hartley's rule $C' = \log_2\left(1 + \frac{A}{\Delta}\right)$ is not Hartley's

Many authors independently derived $C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)$ in 1948.

In fact, $C' = C$ (**a coincidence?**)

Outline

This Hartley's rule $C' = \log_2\left(1 + \frac{A}{\Delta}\right)$ is not Hartley's

Many authors independently derived $C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)$ in 1948.

In fact, $C' = C$ (**a coincidence?**)

Besides, C' is the capacity of the "uniform" channel

Outline

This Hartley's rule $C' = \log_2\left(1 + \frac{A}{\Delta}\right)$ is not Hartley's

Many authors independently derived $C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)$ in 1948.

In fact, $C' = C$ (**a coincidence?**)

Besides, C' is the capacity of the "uniform" channel

(and we can explain)

Outline

This Hartley's rule $C' = \log_2\left(1 + \frac{A}{\Delta}\right)$ is not Hartley's

Many authors independently derived $C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)$ in 1948.

In fact, $C' = C$ (**a coincidence?**)

Besides, C' is the capacity of the "uniform" channel

(and we can explain)

Hartley or not Hartley

Quote from Shannon, 1984:

> aspects of information theory. I started with information theory, inspired by Hartley's paper, which was a good paper, but it did not take account of things like noise and best encoding and probabilistic aspects.[3]

- In Hartley's paper, no mention of signal vs. noise or A vs. Δ
- Why was $C' = \log_2\left(1 + \frac{A}{\Delta}\right)$ mistakenly attributed to Hartley?

The first tutorial of information theory!

A HISTORY OF THE THEORY OF INFORMATION

By E. COLIN CHERRY, M.Sc., Associate Member.

(The paper was first received 7th February, and in revised form 28th May, 1951.)

⋮

increased. Although not explicitly stated in this form in his paper, Hartley[12] has implied that the quantity of information

⋮

distinguishable" amplitude change; in practice this smallest step may be taken to equal the *noise level*, *n*. Then the quantity of information transmitted may be shown to be proportional to

$$Bt \log\left(1 + \frac{a}{n}\right)$$

where *a* is the maximum signal amplitude, an expression given by Tuller,[23] being based upon Hartley's definition of information.

Outline

This Hartley's rule $C' = \log_2\left(1 + \frac{A}{\Delta}\right)$ is not Hartley's

Many authors independently derived $C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)$ in 1948.

In fact, $C' = C$ (**a coincidence?**)

Besides, C' is the capacity of the "uniform" channel

(and we can explain)

And then there were eight

Quote from Shannon, 1948:

> Formulas similar to $C = W \log \frac{P+N}{N}$ for the white noise case have been developed independently by several other writers, although with somewhat different interpretations. We may mention the work of N. Wiener,[7] W. G. Tuller,[8] and H. Sullivan in this connection.

1. Norbert Wiener, *Cybernetics*, early 1948

And then there were eight

Quote from Shannon, 1948:

> Formulas similar to $C = W \log \frac{P+N}{N}$ for the white noise case have been developed independently by several other writers, although with somewhat different interpretations. We may mention the work of N. Wiener,[7] W. G. Tuller,[8] and H. Sullivan in this connection.

1. Norbert Wiener, *Cybernetics*, early 1948
2. William G. Tuller, PhD Thesis, June 1948

And then there were eight

Quote from Shannon, 1948:

> Formulas similar to $C = W \log \frac{P+N}{N}$ for the white noise case have been developed independently by several other writers, although with somewhat different interpretations. We may mention the work of N. Wiener,[7] W. G. Tuller,[8] and H. Sullivan in this connection.

1. Norbert Wiener, *Cybernetics*, early 1948
2. William G. Tuller, PhD Thesis, June 1948
3. H. Sullivan, ?

And then there were eight

Quote from Shannon, 1948:

Formulas similar to $C = W \log \frac{P+N}{N}$ for the white noise case have been developed independently by several other writers, although with somewhat different interpretations. We may mention the work of N. Wiener,[7] W. G. Tuller,[8] and H. Sullivan in this connection.

1. Norbert Wiener, *Cybernetics*, early 1948
2. William G. Tuller, PhD Thesis, June 1948
3. H. Sullivan, ?
4. Jacques Laplume, April 1948

And then there were eight

Quote from Shannon, 1948:

> Formulas similar to $C = W \log \frac{P+N}{N}$ for the white noise case have been developed independently by several other writers, although with somewhat different interpretations. We may mention the work of N. Wiener,[7] W. G. Tuller,[8] and H. Sullivan in this connection.

1. Norbert Wiener, *Cybernetics*, early 1948
2. William G. Tuller, PhD Thesis, June 1948
3. H. Sullivan, ?
4. Jacques Laplume, April 1948
5. Charles W. Earp, June 1948

And then there were eight

Quote from Shannon, 1948:

> Formulas similar to $C = W \log \frac{P+N}{N}$ for the white noise case have been developed independently by several other writers, although with somewhat different interpretations. We may mention the work of N. Wiener,[7] W. G. Tuller,[8] and H. Sullivan in this connection.

1. Norbert Wiener, *Cybernetics*, early 1948
2. William G. Tuller, PhD Thesis, June 1948
3. H. Sullivan, ?
4. Jacques Laplume, April 1948
5. Charles W. Earp, June 1948
6. André G. Clavier, December 1948

And then there were eight

Quote from Shannon, 1948:

> Formulas similar to $C = W \log \frac{P+N}{N}$ for the white noise case have been developed independently by several other writers, although with somewhat different interpretations. We may mention the work of N. Wiener,[7] W. G. Tuller,[8] and H. Sullivan in this connection.

1. Norbert Wiener, *Cybernetics*, early 1948
2. William G. Tuller, PhD Thesis, June 1948
3. H. Sullivan, ?
4. Jacques Laplume, April 1948
5. Charles W. Earp, June 1948
6. André G. Clavier, December 1948
7. Stanford Goldman, May 1948

And then there were eight

Quote from Shannon, 1948:

> Formulas similar to $C = W \log \frac{P+N}{N}$ for the white noise case have been developed independently by several other writers, although with somewhat different interpretations. We may mention the work of N. Wiener,[7] W. G. Tuller,[8] and H. Sullivan in this connection.

1. Norbert Wiener, *Cybernetics*, early 1948
2. William G. Tuller, PhD Thesis, June 1948
3. H. Sullivan, ?
4. Jacques Laplume, April 1948
5. Charles W. Earp, June 1948
6. André G. Clavier, December 1948
7. Stanford Goldman, May 1948
8. Claude E. Shannon, ... July 1940 ????

Norbert Wiener

III

Time Series, Information, and Communication

There is a large class of phenomena in which what is observed is a numerical quantity, or a sequence of numerical quantities, dis-

An interesting problem is that of determining the information gained by fixing one or more variables in a problem. For example, let us suppose that a variable u lies between x and $x + dx$ with the probability $\exp(-x^2/2a)\, dx/\sqrt{2\pi a}$, while a variable v lies between the same two limits with a probability $\exp(-x^2/2b)\, dx/\sqrt{2\pi b}$. How much information do we gain concerning u if we know that $u + v = w$? In this case, it is clear that $u = w - v$, where w is

⋮

Norbert Wiener

The excess of information concerning x when we know w to be that which we have in advance is

$$\frac{1}{\sqrt{2\pi[ab/(a+b)]}}\int_{-\infty}^{\infty}\left\{\exp\left[-(x-c_2)^2\left(\frac{a+b}{2ab}\right)\right]\right\}$$
$$\times\left[-\frac{1}{2}\log_2 2\pi\left(\frac{ab}{a+b}\right)\right]-(x-c_2)^2\left[\left(\frac{a+b}{2ab}\right)\right]\log_2 e\Big]\,dx$$
$$-\frac{1}{\sqrt{2\pi a}}\int_{-\infty}^{\infty}\left[\exp\left(-\frac{x^2}{2a}\right)\right]\left(-\frac{1}{2}\log_2 2\pi a-\frac{x^2}{2a}\log_2 e\right)dx$$
$$=\frac{1}{2}\log_2\left(\frac{a+b}{b}\right) \qquad (3.091)$$

Norbert Wiener

The excess of information concerning x when we know w to be that which we have in advance is

$$\frac{1}{\sqrt{2\pi[ab/(a+b)]}} \int_{-\infty}^{\infty} \left\{\exp\left[-(x-c_2)^2\left(\frac{a+b}{2ab}\right)\right]\right\} \times \left[-\frac{1}{2}\log_2 2\pi\left(\frac{ab}{a+b}\right)\right] - (x-c_2)^2\left[\left(\frac{a+b}{2ab}\right)\right]\log_2 e\Big] dx - \frac{1}{\sqrt{2\pi a}}\int_{-\infty}^{\infty}\left[\exp\left(-\frac{x^2}{2a}\right)\right]\left(-\frac{1}{2}\log_2 2\pi a - \frac{x^2}{2a}\log_2 e\right)dx = \frac{1}{2}\log_2\left(\frac{a+b}{b}\right) \qquad (3.091)$$

Later. . . in 1956:

What is Information Theory?

NORBERT WIENER

INFORMATION THEORY has been identified in the public mind to denote the theory of information by bits, as developed by Claude E. Shannon and myself. This notion is certainly impor-

Jacques Laplume

Meanwhile (1948), far away...

PHYSIQUE MATHÉMATIQUE. — *Sur le nombre de signaux discernables en présence du bruit erratique dans un système de transmission à bande passante limitée.* Note de M. **Jacques Laplume**.

Soit a l'amplitude maximum de $u(t)$, et soit Δa la précision sur l'évaluation de u. Deux signaux $u(t)$ seront discernables s'ils diffèrent de Δa au moins pendant l'un des intervalles Δt. Dans chacun de ces intervalles, le signal peut avoir l'une quelconque des amplitudes discernables $0, \Delta a, 2\Delta a, \ldots, q\Delta a$, avec $q = (a/\Delta a)$.

Ces amplitudes discernables sont au nombre de $q+1$. Le nombre total des signaux discernables est donc

$$M = (q+1)^r,$$

d'où

$$\log M = r \log(q+1). \tag{4}$$

Charles W. Earp

Relationship Between Rate of Transmission of Information, Frequency Bandwidth, and Signal-to-Noise Ratio*

By C. W. EARP

Standard Telephones and Cables, Limited, London, England

nels, channel maximum signal to root-mean-square noise ratio $= S_{\mathrm{SSB}}/\sqrt{n}$ and maximum signal-to-*peak*-noise ratio $= S_{\mathrm{SSB}}/(p\sqrt{n})$.

In each channel, the available power may be used to provide N instantaneous values, this being achieved without ambiguity provided that

$$N < \left(\frac{S_{\mathrm{SSB}}}{p\sqrt{n}} + 1\right).$$

Charles W. Earp

Relationship Between Rate of Transmission of Information, Frequency Bandwidth, and Signal-to-Noise Ratio*

By C. W. EARP

Standard Telephones and Cables, Limited, London, England

nels, channel maximum signal to root-mean-square noise ratio $= S_{SSB}/\sqrt{n}$ and maximum signal-to-*peak*-noise ratio $= S_{SSB}/(p\sqrt{n})$.

In each channel, the available power may be used to provide N instantaneous values, this being achieved without ambiguity provided that

$$N < \left(\frac{S_{SSB}}{p\sqrt{n}} + 1\right).$$

* The present paper was written in original form in October, 1946, when the author had no knowledge of any practical development of pulse-code modulation, as the

André G. Clavier

Evaluation of Transmission Efficiency According to Hartley's Expression of Information Content*

By A. G. CLAVIER

Federal Telecommunication Laboratories, Incorporated, Nutley, New Jersey

small percentage of error due to noise. The total number of distinguishable levels on the ideal line is thus given by

$$\frac{S+\bar{N}\sqrt{2}}{\bar{N}\sqrt{2}}=1+\frac{S}{\bar{N}\sqrt{2}},$$

with a reasonable approximation. It follows that the amount of information transmittible on the ideal line is measured by

$$H_{lm}=k_0\cdot 2f_l\cdot t\cdot \log\left(1+\frac{S_l}{\bar{N}_l\sqrt{2}}\right).$$

André G. Clavier

Evaluation of Transmission Efficiency According to Hartley's Expression of Information Content*

By A. G. CLAVIER

Federal Telecommunication Laboratories, Incorporated, Nutley, New Jersey

* A symposium on "Recent Advances in the Theory of Communication" was presented at the November 12, 1947, meeting of the New York Section of the Institute of Radio Engineers. Four papers were presented by A. G. Clavier, Federal Telecommunication Laboratories; B. D. Loughlin, Hazeltine Electronics Corporation; and J. R. Pierce and C. E. Shannon, both of Bell Telephone Laboratories. The

Stanford Goldman

Some Fundamental Considerations Concerning Noise Reduction and Range in Radar and Communication*

STANFORD GOLDMAN†, SENIOR MEMBER, I.R.E.

The number of significant amplitude levels is usually determined by the noise in the system. If the system is of a linear nature, and the maximum signal amplitude is S, while the noise amplitude is N, then the number of significant amplitude levels is essentially

$$L = (S/N) + 1 \tag{2}$$

where the "1" is due to the fact that the zero signal level can be used.

Stanford Goldman

Some Fundamental Considerations Concerning Noise Reduction and Range in Radar and Communication*

STANFORD GOLDMAN†, SENIOR MEMBER, I.R.E.

The number of significant amplitude levels is usually determined by the noise in the system. If the system is of a linear nature, and the maximum signal amplitude is S, while the noise amplitude is N, then the number of significant amplitude levels is essentially

$$L = (S/N) + 1 \quad (2)$$

where the "1" is due to the fact that the zero signal level can be used.

[4] Equation (5) has been derived independently by many people, among them W. G. Tuller, from whom the writer first learned about it.

William G. Tuller

Theoretical Limitations on the Rate of Transmission of Information*

WILLIAM G. TULLER†, SENIOR MEMBER, IRE

recognizable.[14] Then, if N is the rms amplitude of the noise mixed with the signal, there are $1+S/N$ significant values of signal that may be determined. This sets s in

⋮

have from (1) the quantity of information available at the output of the system:

$$H = kn \log s = k2f_cT \log (1 + S/N). \qquad (2)$$

This is an important expression, to be sure, but gives

William G. Tuller

Theoretical Limitations on the Rate of Transmission of Information*

WILLIAM G. TULLER†, SENIOR MEMBER, IRE

recognizable.[14] Then, if N is the rms amplitude of the noise mixed with the signal, there are $1+S/N$ significant values of signal that may be determined. This sets s in

⋮

have from (1) the quantity of information available at the output of the system:

$$H = kn \log s = k2f_cT \log (1 + S/N). \qquad (2)$$

This is an important expression, to be sure, but gives

[10] C. E. Shannon, "Communication in the presence of noise," Proc. I.R.E., vol. 37, pp. 10–22; January, 1949.

[11] The existence of this work was learned by the author in the spring of 1946, when the basic work underlying this paper had just been completed. Details were not known by the author until the summer of 1948, at which time the work reported here had been complete for about eight months.

Communication in the Presence of Noise*

CLAUDE E. SHANNON†, MEMBER, IRE

THEOREM 2: *Let P be the average transmitter power, and suppose the noise is white thermal noise of power N in the band W. By sufficiently complicated encoding systems it is possible to transmit binary digits at a rate*

$$C = W \log_2 \frac{P + N}{N} \qquad (19)$$

with as small a frequency of errors as desired. It is not pos-

Claude E. Shannon

Communication in the Presence of Noise*

CLAUDE E. SHANNON†, MEMBER, IRE

* Decimal classification: 621.38. Original manuscript received by the Institute, July 23, 1940. Presented, 1948 IRE National Convention, New York, N. Y., March 24, 1948; and IRE New York Section, New York, N. Y., November 12, 1947.

Communication in the Presence of Noise*

CLAUDE E. SHANNON†, MEMBER, IRE

[10] A. Hodges, *Alan Turing: The Enigma*, New York: Simon and Schuster, 1983. [The following information was obtained from C. E. Shannon on March 3, 1984: "On p. 552, Hodges cites a Shannon manuscript date of 1940, which is, in fact, a typographical error. While results for coding statistical sources into noiseless channels using the $p\log(p)$ measure were obtained in 1940–1941 (at the Institute for Advanced Study in Princeton), first submission of this work for formal publication occurred soon after World War II."]

Who's formula?

The "Shannon-Hartley" formula

$$\boxed{C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)}$$

Who's formula?

The "Shannon-Hartley" formula

$$\boxed{C = \tfrac{1}{2}\log_2\left(1 + \frac{P}{N}\right)}$$

would actually be the

Shannon-Tuller-Wiener-Sullivan-Laplume-Earp-Clavier-Goldman formula

Who's formula?

The "Shannon-Hartley" formula

$$\boxed{C = \tfrac{1}{2}\log_2\Big(1 + \frac{P}{N}\Big)}$$

would actually be the

Shannon-Tuller-Wiener-Sullivan-Laplume-Earp-Clavier-Goldman formula

or simply the

Shannon-Tuller formula

Outline

This Hartley's rule $C' = \log_2\left(1 + \frac{A}{\Delta}\right)$ is not Hartley's

Many authors independently derived $C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)$ in 1948.

In fact, $C' = C$ (**a coincidence?**)

Besides, C' is the capacity of the "uniform" channel

(and we can explain)

“Hartley”’s argument

The channel input X is taking $M = 1 + A/\Delta$ equiprobable values in the set $\{-A, -A + 2\Delta, \ldots, A - 2\Delta, A\}$:

$$P = \mathbb{E}(X^2) = \frac{1}{M} \sum_{k=0}^{n} (M - 1 - 2k)^2 = \Delta^2 \frac{M^2 - 1}{3}.$$

The input is mixed with additive noise Z with accuracy $\pm\Delta$, i.e. having uniform distribution in $[-\Delta, \Delta]$:

$$N = \mathbb{E}(Z^2) = \frac{1}{2\Delta} \int_{-\Delta}^{\Delta} z^2 \mathrm{d}z = \frac{\Delta^2}{3}.$$

“Hartley”’s argument

The channel input X is taking $M = 1 + A/\Delta$ equiprobable values in the set $\{-A, -A + 2\Delta, \dots, A - 2\Delta, A\}$:

$$P = \mathbb{E}(X^2) = \frac{1}{M}\sum_{k=0}^{n}(M-1-2k)^2 = \Delta^2\frac{M^2-1}{3}.$$

The input is mixed with additive noise Z with accuracy $\pm\Delta$, i.e. having uniform distribution in $[-\Delta, \Delta]$:

$$N = \mathbb{E}(Z^2) = \frac{1}{2\Delta}\int_{-\Delta}^{\Delta} z^2 \mathrm{d}z = \frac{\Delta^2}{3}.$$

Hence

$$\log_2\Bigl(1+\frac{A}{\Delta}\Bigr) = \tfrac{1}{2}\log_2(1+M^2-1) = \tfrac{1}{2}\log_2\Bigl(1+\frac{3P}{\Delta^2}\Bigr) = \tfrac{1}{2}\log_2\Bigl(1+\frac{P}{N}\Bigr)$$

i.e., $\boxed{C' = C}$. A mathematical coïncidence?

Outline

This Hartley's rule $C' = \log_2\left(1 + \frac{A}{\Delta}\right)$ is not Hartley's

Many authors independently derived $C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)$ in 1948.

In fact, $C' = C$ (**a coincidence?**)

Besides, C' is the capacity of the "uniform" channel

(and we can explain)

The uniform channel

The capacity of $Y = X + Z$ with additive *uniform* noise Z is

$$\begin{aligned}\max_{X \text{ s.t. } |X|\leq A} I(X;Y) &= \max_X h(Y) - h(Y|X)\\ &= \max_X h(Y) - h(Z)\\ &= \max_{X \text{ s.t. } |Y|\leq A+\Delta} h(Y) - \log_2(2\Delta)\end{aligned}$$

Choose X^* to be discrete uniform in $\{-A, -A+2\Delta, \ldots, A\}$, then $Y = X^* + Z$ has uniform density over $[-A-\Delta, A+\Delta]$, which maximizes differential entropy:

$$\begin{aligned}&= \log_2(2(A+\Delta)) - \log_2(2\Delta)\\ &= \boxed{\log_2\left(1+\frac{A}{\Delta}\right)}\end{aligned}$$

What is the worst noise?

Thus $C' = \log_2\left(1+\frac{A}{\Delta}\right)$ is *correct* as the capacity of a communication channel! except that

- the noise is *not* Gaussian, but uniform;
- signal limitation is *not* on the power, but on the amplitude.

What is the worst noise?

Thus $C' = \log_2\left(1 + \frac{A}{\Delta}\right)$ is *correct* as the capacity of a communication channel! except that

- the noise is *not* Gaussian, but uniform;
- signal limitation is *not* on the power, but on the amplitude.

Further analogy:

What is the worst noise?

Thus $C' = \log_2\left(1 + \frac{A}{\Delta}\right)$ is *correct* as the capacity of a communication channel! except that

- the noise is *not* Gaussian, but uniform;
- signal limitation is *not* on the power, but on the amplitude.

Further analogy:

- Shannon used the entropy power inequality to show that under limited *power*, Gaussian noise is the worst possible noise one can inflict in the channel:

$$\frac{1}{2}\log_2\left(1+\alpha\frac{P}{N}\right) \leq C \leq \frac{1}{2}\log_2\left(1+\frac{P}{N}\right) + \frac{1}{2}\log_2\alpha,$$

  where $\alpha = N/\tilde{N} \geq 1$

What is the worst noise?

Thus $C' = \log_2\left(1+\frac{A}{\Delta}\right)$ is *correct* as the capacity of a communication channel! except that

- the noise is *not* Gaussian, but uniform;
- signal limitation is *not* on the power, but on the amplitude.

Further analogy:

- Shannon used the entropy power inequality to show that under limited *power*, Gaussian noise is the worst possible noise one can inflict in the channel:

$$\tfrac{1}{2}\log_2\left(1+\alpha\frac{P}{N}\right) \leq C \leq \tfrac{1}{2}\log_2\left(1+\frac{P}{N}\right)+\tfrac{1}{2}\log_2\alpha,$$

  where $\alpha = N/\tilde{N} \geq 1$
- We can show: under limited *amplitude*, *uniform* noise is the worst possible noise one can inflict in the channel:

$$\log_2\left(1+\frac{A}{\Delta}\right) \leq C' \leq \log_2\left(1+\frac{A}{\Delta}\right)+\log_2\alpha,$$

  where $\alpha = \Delta/\tilde{\Delta} \geq 1$.

Conclusion

Why is Shannon's formula ubiquitous?

Conclusion

Why is Shannon's formula ubiquitous?

- we can explain the coincidence by deriving necessary and sufficient conditions s.t. $C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)$.

Conclusion

Why is Shannon's formula ubiquitous?

- we can explain the coincidence by deriving necessary and sufficient conditions s.t. $C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)$.
- the uniform (Tuller) and Gaussian (Shannon) channels are not the only examples.

Conclusion

Why is Shannon's formula ubiquitous?

- we can explain the coincidence by deriving necessary and sufficient conditions s.t. $C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)$.
- the uniform (Tuller) and Gaussian (Shannon) channels are not the only examples.
- using B-splines, we can construct a sequence of such additive noise channels s.t.

$$\text{uniform channel} \longrightarrow \text{Gaussian channel}$$

Conclusion

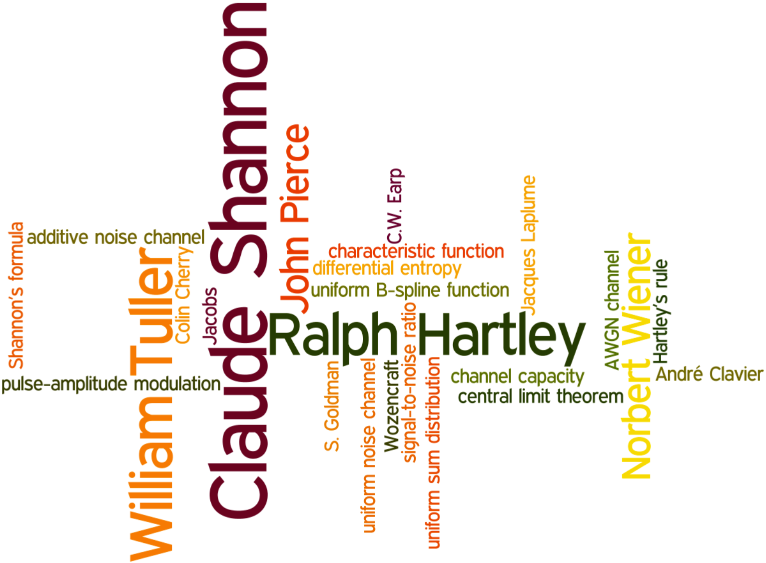
Why is Shannon's formula ubiquitous?

- we can explain the coincidence by deriving necessary and sufficient conditions s.t. $C = \frac{1}{2}\log_2\left(1+\frac{P}{N}\right)$.
- the uniform (Tuller) and Gaussian (Shannon) channels are not the only examples.
- using B-splines, we can construct a sequence of such additive noise channels s.t.

$$\text{uniform channel} \longrightarrow \text{Gaussian channel}$$

"On Shannon's formula and Hartley's rule: Beyond the mathematical coincidence,"
in Journal *Entropy*, Vol. 16, No. 9, pp. 4892-4910, Sept. 2014.
http://www.mdpi.com/1099-4300/16/9/4892/

Thank you!


Claude Shannon
Ralph Hartley
William Tuller
John Pierce
Norbert Wiener
Shannon's formula
additive noise channel
Colin Cherry
Jacobs
pulse-amplitude modulation
C.W. Earp
characteristic function
differential entropy
uniform B-spline function
Jacques Laplume
S. Goldman
uniform noise channel
Wozencraft
signal-to-noise ratio
uniform sum distribution
channel capacity
central limit theorem
AWGN channel
Hartley's rule
André Clavier

A characterization of $C = \frac{1}{2}\log_2\left(1 + \frac{P}{N}\right)$

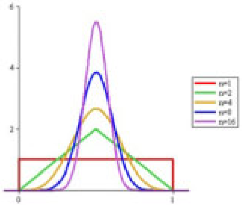
There exists $\alpha > 1$ such that the ratio of characteristic functions

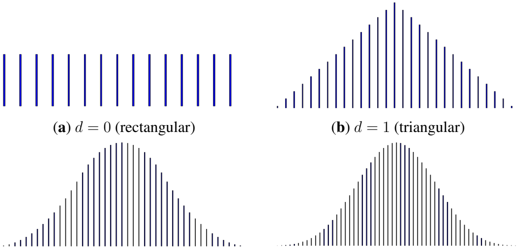
$$\frac{\Phi_Z(\alpha\omega)}{\Phi_Z(\omega)}$$

is itself a characterization function of a r.v. X^* — which attains capacity under an average cost per channel use $\mathbb{E}\{b(X)\} \leq C$, where

$$b(x) = \mathbb{E}\left\{\log_2\left(\frac{\alpha p_Z(Z)}{p_Z((x+Z)/\alpha)}\right)\right\}$$

B-splines channels





(a) $d = 0$ (rectangular) **(b)** $d = 1$ (triangular)

(c) $d = 2$ **(d)** $d = 3$