

There is no Theoretical Limit for Noise Reduction in Digital Communications

János Ladvánszky¹

¹ Senior Member, IEEE, Budapest University of Technology and Economics, Department of Broadband Communications and Electromagnetic Theory, Hungary

Correspondence: János Ladvánszky, Senior Member, IEEE, Budapest University of Technology and Economics, Department of Broadband Communications and Electromagnetic Theory, Hungary. E-mail: Ladvanszky55@t-online.hu

Received: June 30, 2021 Accepted: July 11, 2021 Online Published: July 22, 2021

doi:10.5539/nct.v6n1p34

URL: <https://doi.org/10.5539/nct.v6n1p34>

Abstract

Shannon entropy is a basic characteristic of communications from energetic point of view. Entropy has been expressed here as a function of signal to noise ratio, and lower bound for entropy has been investigated. We prove that finite nonzero bound does not exist, therefore in case of M-QAM modulation, there is no theoretical limit for reduction of the effect of noise. In our investigations, averaging is considered, exploiting the zero expected value of the Gaussian noise.

Index Terms—Shannon entropy, probability of successful communications, bit error rate, signal to noise ratio, bound for noise reduction.

1. Introduction

In this year, we celebrate the 73rd anniversary of the Shannon theory (Shannon, 1948). An essential tool of the theory is the Shannon entropy (Shannon, 1949):

$$H = -\sum_i p_i \log_2 p_i \quad (1)$$

where p_i is the probability of the successful communication of the i^{th} message, \log_2 is the logarithm of base 2. For binary messages,

$$H(p) = -p \log_2 p - (1-p) \log_2 (1-p) \quad (2)$$

From thermodynamics, entropy is a measure of disorder. At complete order, entropy is zero, at complete disorder, entropy has a maximum. Suppose that this idea can be extended for the theory of communications as well. That means, when we try to decrease effect of noise, then we try to decrease entropy. At minimum effect of noise, entropy has a lower bound. So, our goal is to find the lower bound of entropy.

In Figure 1 we can see (2).

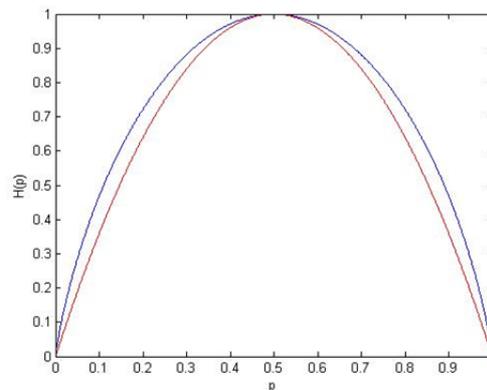


Figure 1. Shannon entropy as a function of probability of successful communication. Blue: $H(p)$ in Equation (2). Red: approximating quadratic curve

The curve of the entropy is almost a quadratic function. We can see that complete order is at $p=0$ and $p=1$, complete disorder is at $p=0.5$. Really, success of communications is surely detected at $p=0$ and $p=1$, and the uncertainty is the highest at $p=0.5$. Figure 1 has been provided using the following Matlab program:

```
p=0:0.001:1;
plot(p, -p.*log2(p) - (1-p) .*log2(1-p))
hold on
plot(p, 1-4.*(p-.5).^2, 'r')
```

In Section II, H has been expressed as a function of signal to noise ratio. In Section III, it has been proved that H has no extremum, as a function of the detection threshold. In Section IV, it is proved based on the previous Section, that there is no theoretical limit for reduction of the effect of noise. Throughout of our investigations, averaging is considered.

Entropy

i) Thermodynamical entropy (Lieb & Yngvason, 1999)

The concept of entropy was introduced by Clausius (Clausius, 1879). Goal was to distinguish natural processes that are not possible or irreversible, despite the law of energy conservation is not hurt. Second main law of thermodynamics says that in an isolated system, total entropy cannot decrease. Thus, entropy is a measure of disorder.

Clausius approach is macroscopic, complemented by Boltzmann's microscopic approach (Laranjeiras, Lucena, & Chiappin, 2020). The formula for Boltzmann entropy says that entropy is linearly proportional to the logarithm of number of microscopic states in a system, i. e. $S = k \ln W$ where k is the Boltzmann constant and W is the number of states. This approach is indispensable in determining energy distribution of different elementary particles. One of them, the Fermi-Dirac statistics (Dirac, 1926), is a part of the basis of semiconductor physics.

ii) Gibbs entropy (Jaynes, 1965)

The Gibbs entropy formula differs from that of Boltzmann, $H = k \sum p_i \ln p_i$ where k is the Boltzmann constant and p_i are probabilities of system fluctuations. It is shown that there are neglects in the Boltzmann formula that cannot be hold, and the correct formula is Gibbs's one.

iii) von Neumann entropy (Mackey, 2013)

The von Neumann entropy is an extension of Gibbs entropy for the case of quantum mechanics: $S = -tr(\rho \ln \rho)$, where ρ is the density matrix (*Wikipedia: Density matrix*), \ln is the natural logarithm (of base e), and $tr()$ is trace (sum of diagonal elements).

iv) Shannon entropy (Shannon & Weaver, 1949)

Shannon entropy is essentially the extension of Gibbs entropy from particles to messages in information transfer: $H = -\sum p_i \text{ld} p_i$ where ld is the logarithm of base 2 and p_i is the probability of the successful transfer of the i^{th} message.

It is an interesting question if entropy is in connection with the amount of transferred information. If self-information (*Wikipedia: Information content*) is defined as $I = -\sum \text{ld} p_i$ then entropy is the expected self-information $H = \sum p_i I_i$.

Now, if the probabilities p_i are identical, then we return to the Boltzmann entropy formula, namely, if $p_i = \frac{1}{W}$, then $H = \sum \frac{1}{W} \left(-\text{ld} \frac{1}{W}\right) = \text{ld} W$. So, the Boltzmann entropy of a series of messages equals to the expected self-information.

An expression for Shannon entropy

From (2) we have the following idea. i) The probability of a successful message p is in close connection with the bit error rate (BER). ii) BER can be expressed as a function of the signal to noise ratio (SNR) for any specified types of digital modulation (Proakis, 2001). iii) Then entropy can be expressed as a function of the detection threshold L , see please a few lines below. iv) Then extremum of entropy can be investigated.

This idea has been motivated by the problem if successful communication of bits is really limited by the condition when power density of the signal is identical to that of the noise. We suspect that the answer is negative, therefore we introduce the detection threshold $0 \leq L \leq 1$ and say, that the successful communication is limited by:

$$PDS = L * PDN \quad (3)$$

where PDS and PDN are the signal and noise power densities, respectively. That means if $PDS > L.PDN$ then the message is detected successfully, otherwise not. Our plan is to express $H(L)$ and try to determine the value of L corresponding to the minimum of H at a given SNR .

i) Denote g and b the number of good and bad detection of bits, respectively. Then

$$p = \frac{g}{g+b} \quad (4)$$

$$BER = \frac{b}{g+b} \quad (5)$$

By combining (4) and (5),

$$p = 1 - BER \quad (6)$$

From (2) and (6),

$$H(BER) = -BER \ln BER - (1 - BER) \ln(1 - BER) \quad (7)$$

ii) Expression for BER as a function of SNR is found in (Proakis, 2001):

$$BER = \frac{1}{2} \operatorname{erfc} \sqrt{\frac{E_b}{N_0}} \quad (8)$$

where E_b is bit energy, N_0 is noise power density and 4QAM modulation was assumed. Bit energy is the same as signal power density (Appendix III), so (8) can be written as follows:

$$BER = \frac{1}{2} \operatorname{erfc} \sqrt{\frac{PDS}{PDN}} \quad (9)$$

where PDS and PDN are the signal and noise power densities, respectively. In this paper, we define the signal to noise ratio SNR as

$$SNR = \frac{PDS}{PDN} \quad (10)$$

This definition, instead of $SNR = \frac{S}{N}$, where S and N are the signal and noise power, respectively, is based on our modification of the Shannon formula (Ladvánszky, 2020).

Assume that averaging is applied for our signal that decreases PDN and leaves PDS intact:

$$SNR' = \frac{PDS}{L.PDN} = \frac{1}{L} SNR \quad (11)$$

Thus from (9), $BER = \frac{1}{2} \operatorname{erfc} \sqrt{SNR'}$ is rewritten as

$$BER = \frac{1}{2} \operatorname{erfc} \sqrt{\frac{1}{L} \cdot SNR} \quad (12)$$

iii) Applying (7) and (12), $H(L)$ has been obtained, for a given SNR .

iv) Now it is to investigate if $H(L)$ has a finite nonzero minimum.

H has no extremum :

In Figure 2, we plotted the logarithmic derivative of $H(L)$.

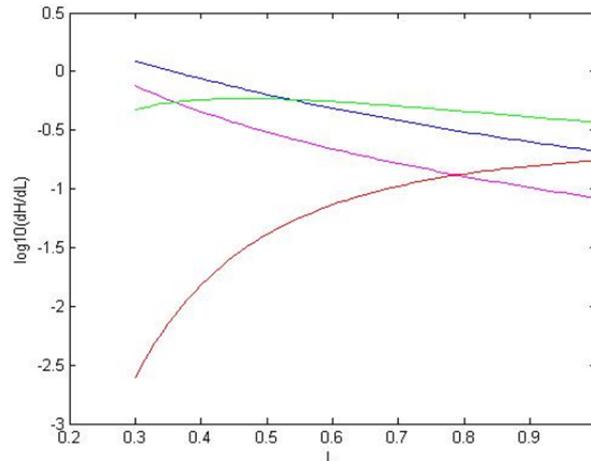


Figure 2. Horizontal axis: L . Vertical axis: $\log_{10} \frac{dH(L)}{dL}$, at some SNR values: .1 (magenta), .3 (blue), 1 (green), 3 (red). $\frac{dH}{dL} > 0$ thus $H(L)$ has no minimum, it is strictly monotonically increasing

Figure 2 has been provided using the following Matlab program:

```
SNR=.1;
L=0.3:0.01:1;
dL=0.001;
BER=(.5*erfc(sqrt(1./L.*SNR)));
BER1=(.5*erfc(sqrt(1./(L+dL).*SNR)));
H=-BER.*log2(BER)-(1-BER).*log2(1-BER);
H1=-BER1.*log2(BER1)-(1-BER1).*log2(1-BER1);
dH=H1-H;
plot(L,(log10(dH/dL)),'m');
```

This program has been repeated four times with different values of SNR .

Meaning of Figure 2 is the following.

L can be arbitrarily small. That is, effect of noise can be decreased arbitrarily by averaging. That is done by repeating the message in noise and taking the arithmetic mean of the detected signals. This can be done without any bounds.

Analytical proofs have been found in Appendix I and II.

5. Conclusions

We expected that by finding the extremum of $H(L)$, a specific L value can be found that is the bound of the averaging procedure. Surprisingly, we found that L can be arbitrarily small.

The result is significant. When trying to decrease the effect of noise by averaging, repetition can be done arbitrary times, without limit. This is reasonable, because increasing the number of averaging, effect of noise is further reduced. Thus, Equation (3) says that there is no theoretical limit for reduction of the effect of noise.

This result is in harmony with the similar result from the theory of coding. BER can be arbitrarily reduced by applying sufficient redundancy.

Note that we speak about lack of theoretical limit. Of course, practical limit is more rigorous. Otherwise signal to noise ratio could be made arbitrarily high, and there were no upper bound for the speed of information transfer over a noisy channel of finite bandwidth.

Although our calculations are for 4QAM, the same can be done for any kind of digital modulation. That is the explanation of the title.

Finally, we should emphasize that the term noise reduction does not mean here physical reduction of noise, but only that of its effect on communications.

One of the reviewers required more comparison with existing results. The author is not aware of any such study (Shannon, 1948; 1949; Proakis, 2001; Littlejohn & Foss, 2009). The nearest idea is, as we have already mentioned, proof of existence of an error correcting code if the BER is lower than the theoretical upper limit (0.5) (Shannon, 1948; 1949). Here we pointed out that theoretically we can achieve the same without coding. Thus, averaging and coding may result the same, as both exploit redundancy.

Acknowledgements

The author is grateful to the late Dr. A. Baranyi, and Dr. E. Simonyi who supported him in research for several decades.

Special thanks are due to Mr. Z. Huszka, who called the attention for aging of the useful signal in the averaging process.

References

- Clausius, R. (1879). *The mechanical theory of heat*. Retrieved from <https://www3.nd.edu/~powers/ame.20231/clausius1879.pdf>
- Dirac, P. A. M. (1926). On the theory of quantum mechanics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 112(762), 661-677. <https://doi.org/10.4135/9781412959384>
- Jaynes, E. T. (1965). Gibbs vs Boltzmann entropies. *American Journal of Physics*, 33(5), 391-398. <https://doi.org/10.1119/1.1971557>
- Ladvánszky, J. (2020). A Modification to the Shannon Formula. *Netw. Commun. Technol.*, 5(2), 1-6. <https://doi.org/10.5539/nct.v5n2p1>
- Laranjeiras, C. C., Lucena, J., & Chiappin, J. R. (2020). Boltzmann and the Heuristics of Representation in Statistical Mechanics. *Transversal: International Journal for the Historiography of Science*, 8. <https://doi.org/10.24117/2526-2270.2020.i8.07>
- Lieb, E. H., & Yngvason, J. (1999). The physics and mathematics of the second law of thermodynamics. *Physics Reports*, 310(1), 1-96. [https://doi.org/10.1016/S0370-1573\(98\)00082-9](https://doi.org/10.1016/S0370-1573(98)00082-9)
- Littlejohn, S. W., & Foss, K. A. (2009). *Encyclopedia of communication theory* (Vol. 1). Sage.
- Mackey, G. W. (2013). *Mathematical foundations of quantum mechanics*. Courier Corporation. <https://doi.org/10.1515/9781400889921>
- MathWorks. (n.d.) *Erfc function*. Retrieved from <https://mathworld.wolfram.com/Erfc.html>
- Meghdadi, V. (2008). BER calculation. *Wireless Communications*. Retrieved from https://www.unilim.fr/pages_perso/vahid/notes/ber_awgn.pdf
- Proakis, J. G. (2001). *Digital Communications*. McGraw-Hill, New York.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, 27(3), 379-423. <https://doi.org/10.1002/j.1538-7305.1948.tb00917.x>
- Shannon, C. E. (1949). Communication in the presence of noise. *Proceedings of the IRE*, 37(1), 10-21. <https://doi.org/10.1109/JRPROC.1949.232969>
- Shannon, C. E., & Weaver, W. (1949). *The Mathematical Theory of Communication*. Univ of Illinois Press.
- Wikipedia: *Density matrix*. Retrieved from https://en.wikipedia.org/wiki/Density_matrix#:~:text=A%20density%20matrix%20is%20a,a%20system%20in%20quantum%20mechanics.&text=Describing%20a%20quantum%20state%20by,a%20statistical%20ensemble%20of%20kets.
- Wikipedia: *Fourier transform*.
- Wikipedia: *Information content*. Retrieved from https://en.wikipedia.org/wiki/Information_content
- Wireless Pi. (2019). *Computing error rates*. Retrieved from <https://wirelesspi.com/computing-error-rates/>

Appendixes

Appendix I. Analytical proof of the strict monotonicity of $H(L)$ for 4QAM

Strict positivity of $\frac{dH(L)}{dL}$ should be proved. We achieve it in the following way:

$$\frac{dH(L)}{dL} = \frac{dH(p)}{dp} \frac{dp}{dBER} \frac{dBER}{dSNR'} \frac{dSNR'}{dL} \quad (\text{A.1.1})$$

Let us start with (2) above:

$$H(p) = -p \text{ld } p - (1-p) \text{ld}(1-p) \quad (\text{A.1.2})$$

(A.1.2) yields:

$$\begin{aligned} \frac{dH(p)}{dp} &= \frac{-1}{\ln 2} \frac{dp}{dp} (p \ln p + (1-p) \ln(1-p)) = \frac{-1}{\ln 2} \left[\left(\ln p + p \frac{1}{p} \right) + \left(-\ln(1-p) + (1-p) \frac{-1}{1-p} \right) \right] = \frac{1}{\ln 2} \ln \frac{1-p}{p} = \\ &\text{ld } \frac{1-p}{p} \end{aligned} \quad (\text{A.1.3})$$

Continue with (6) above:

$$p = 1 - BER \quad (\text{A.1.4})$$

$$\frac{dp}{dBER} = -1 \quad (\text{A.1.5})$$

Continue with (9) above:

$$BER \approx \frac{1}{2} \text{erfc} \sqrt{SNR'} \quad (\text{A.1.6})$$

$$\frac{dBER}{dSNR'} = \frac{1}{2} \left(\frac{-2}{\sqrt{\pi}} e^{-SNR'} \right) \frac{1}{2\sqrt{SNR'}} = \frac{-1}{2\sqrt{\pi SNR'}} e^{-SNR'} \quad (\text{A.1.7})$$

Continue with (11) above:

$$SNR' = \frac{PDS}{L \cdot PDN} = \frac{1}{L} SNR \quad (\text{A.1.8})$$

$$\frac{dSNR'}{dL} = \frac{-1}{L^2} SNR \quad (\text{A.1.9})$$

Combining the equations above:

$$\frac{dH(L)}{dL} = \text{ld } \frac{1-p}{p} (-1) \frac{-1}{2\sqrt{\pi SNR}} e^{-SNR} \frac{-1}{L^2} SNR \quad (\text{A.1.10})$$

$$\text{ld } \frac{1-p}{p} (-1) = \text{ld} \left(\frac{1}{BER} - 1 \right) \quad (\text{A.1.11})$$

As in (A.1.6),

$$BER < \frac{1}{2} \quad (\text{A.1.12})$$

thus in (A.1.11),

$$\text{ld} \left(\frac{1}{BER} - 1 \right) > 0 \quad (\text{A.1.13})$$

and in (A.1.10),

$$\frac{dH(L)}{dL} > 0 \quad (\text{A.1.14})$$

for any SNR values. Therefore $H(L)$ has no extremum, it is strictly monotonically increasing. That is, we wanted to prove.

Appendix II. Extension for QAM of arbitrary order

(A.1.6) is an approximation. That should be replaced with the exact formula that is valid for QAM of arbitrary order (Meghdadi, 2008; Wireless Pi, 2019).

Let us start with the exact formula for M-ary QAM symbol error rate (Meghdadi, 2005; Equation 15):

$$SER = 1 - \left(1 - \frac{2(\sqrt{M}-1)}{\sqrt{M}} Q \left(\sqrt{\frac{3\gamma_s}{M-1}} \right) \right)^2 \quad (\text{A.2.1})$$

where M is the order of QAM,

$$\gamma_s = \frac{E_s}{N_0} \quad (\text{A.2.2})$$

where E_s and N_0 are the average symbol energy and noise power density, respectively,

$$Q(x) = \frac{1}{2} \operatorname{erfc} \frac{x}{\sqrt{2}} \quad (\text{A.2.3})$$

$$\gamma_b = \frac{E_b}{N_0} \quad (\text{A.2.4})$$

where E_b is the average bit energy, and

$$E_s = E_b \log_2 M \quad (\text{A.2.5})$$

$$SER = BER \log_2 M \quad (\text{A.2.6})$$

Combination of (A.2.1-6) yields:

$$BER = \frac{1}{\log_2 M} \left[1 - \left(1 - \frac{(\sqrt{M}-1)}{\sqrt{M}} \operatorname{erfc} \left(\sqrt{\frac{3SNR' \log_2 M}{2(M-1)}} \right) \right)^2 \right] \quad (\text{A.2.7})$$

Based on (A1.2), (A1.4), (A2.7), (A1.8) the $H(L)$ function is shown in Fig. 3 for 4, 16 and 64 QAM.

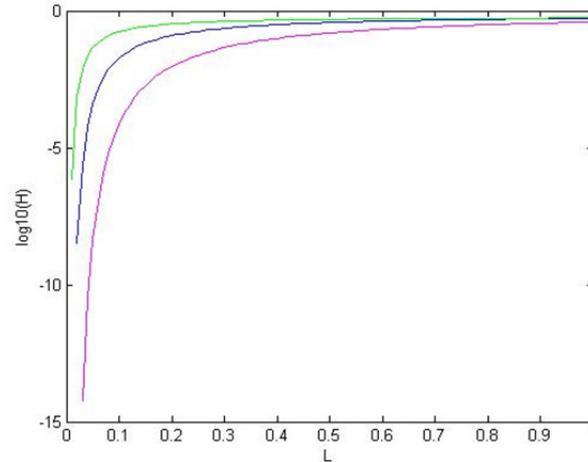


Figure 3. Shannon entropy as a function of the detection threshold for 4QAM (magenta), 16QAM (blue) and 64QAM (green)

Figure 3 has been plotted using the Matlab program below, repeated with different values of M :

```
M=4;
SNRprime_dB=0;
L=0.01:0.01:1
SNR=10.^(SNRprime_dB./20)./L;
BER=1./log2(M).*(1-(1-(sqrt(M)-1)/sqrt(M)*erfc(sqrt((3.*SNR.*log2(M))./(M-1))./2))).^2);
H=-BER.*log2(BER)-(1-BER).*log2(1-BER);
plot(L,log10(H),'m')
```

We use (MathWorks. n.d.):

$$\frac{d}{dz} \operatorname{erfc}(z) = -\frac{2e^{-z^2}}{\sqrt{\pi}} \quad (\text{A.2.8})$$

Then we replace (A.1.7) with the following:

$$\frac{dBER}{dSNR'} = \frac{-1}{ld M} \left[2 \left(1 - \frac{(\sqrt{M}-1)}{\sqrt{M}} \operatorname{erfc} \sqrt{\frac{3SNR' ld M}{2(M-1)}} \right) \left(-\frac{(\sqrt{M}-1)}{\sqrt{M}} \right) \left(-\frac{2}{\sqrt{\pi}} \right) e^{-\frac{3SNR' ld M}{2(M-1)}} \frac{1}{2 \sqrt{\frac{3SNR' ld M}{2(M-1)}}} \frac{3 ld M}{2(M-1)} \right] = \left[- \left(1 - \frac{(\sqrt{M}-1)}{\sqrt{M}} \operatorname{erfc} \left(\sqrt{\frac{3SNR' ld M}{2(M-1)}} \right) \right) \left(\frac{(\sqrt{M}-1)}{\sqrt{M}} \right) \left(\frac{1}{\sqrt{\pi}} \right) e^{-\frac{3SNR' ld M}{2(M-1)}} \frac{1}{\sqrt{\frac{3SNR' ld M}{2(M-1)}}} \frac{3}{M-1} \right] \quad (\text{A.2.9})$$

As $M \geq 2$, $\frac{dBER}{dSNR'} < 0$. Substituting this into (A.1.1):

$$\frac{dH(L)}{dL} = \frac{dH(p)}{dp} \frac{dp}{dBER} \frac{dBER}{dSNR'} \frac{dSNR'}{dL} > 0 \quad (\text{A.2.10})$$

because all terms are negative. Thus $H(L)$ is strictly monotonically increasing for QAM of order M .

Appendix III. Bit energy and signal power density

One of the reviewers asked a reference for the statement that bit energy is the same as signal power density. As we are not aware of a reference for this statement, in this Appendix we explain it.

Sketch of the calculation is that we determine bit energy E_b and average signal power density PSD of a periodic, rectangular signal of amplitude V_0 and bitrate R that means repetition frequency of $R/2$. Then we find

$$E_b = \alpha PSD \quad (A.3.1)$$

where α is a positive, real number. In the calculation, we assume an ideal rectangular lowpass filter following the mixer, with corner frequency R .

The exact value of α does not influence the conclusion of this paper, thus, without loss of generality, $\alpha = 1$ is assumed.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).