

Estimation of the Randomness of Continuous and Discrete Signals Using the Disentropy of the Autocorrelation

R. V. Ramos

rubens.ramos@ufc.br

Lab. of Quantum Information Technology, Department of Teleinformatic Engineering – Federal University of Ceara - DETI/UFC, C.P. 6007 – Campus do Pici - 60455-970 Fortaleza-Ce, Brazil.

Abstract — The amount of randomness in a signal generated by physical or non-physical process can reveal important information about that process. For example, the presence of randomness in ECG signals may indicate a cardiac disease. On the hand, the lack of randomness in a speech signal may indicate the speaker is a machine. Hence, to quantify the amount of randomness in a signal is an important task in many different areas. In this direction, the present work proposes to use the disentropy of the autocorrelation function as a measure of randomness. Examples using noisy and chaotic signals, economic and biological data are shown.

Keywords — Disentropy, autocorrelation, randomness, Lambert-Tsallis function.

1. Introduction

A basic task in modern science is the data analysis. Medical, astronomical, economic, seismological, among many others types of data, generated by physical or non-physical processes, have to be analysed every day. In all cases these data are (intentionally or not) contaminated with noise. This noise is the source of randomness in the data that, depending on the goal desired can be useful or not. The randomness in meteorological data can be very harmful since it makes harder the weather forecast. On the other hand, the randomness in a cryptographic key is highly desired. Thus, to study the amount of randomness in a (continuous or discrete) signal is an important task since it can inform, for example, the presence of a disease in a ECG signal, an illegal operation in the stock market, to denounce a fake image or speech or, simply, to show the data considered is useless, like a cryptographic key with low randomness.

An important mathematical tool in the analysis of random signals is the autocorrelation function. For a continuous signal $s(t)$, the autocorrelation function is defined as

$$R(\tau) = \int_{-\infty}^{\infty} s(t) s^*(t - \tau) dt. \quad (1)$$

Similarly, the autocorrelation at lag k of the discrete signal s_t is defined as

$$r_k = \frac{E[(s_t - \bar{s})(s_{t+k} - \bar{s})]}{\sigma_s^2} = \frac{1}{N} \frac{\sum_{t=1}^{N-k} (s_t - \bar{s})(s_{t+k} - \bar{s})}{\frac{1}{N} \sum_{t=1}^N (s_t - \bar{s})^2}, \quad (2)$$

where \bar{s} and σ_s^2 are, respectively, the mean value and variance of s_t . The symbol E denotes the expected value. Since the autocorrelation function shows the similarity between a function and its delayed version, thereby showing the degree by which its value at one time is similar to its value at a certain later time, it is natural to use it to analyse the randomness contained in a signal: the larger the randomness, the smaller the similarity. However, the autocorrelation is a function and not a variable. Thus, in order to get a randomness measure based on the autocorrelation, one needs a function that maps the autocorrelation function to a real number. In this work, the disentropy is used for this task. As it will be show latter, the disentropy is a good candidate once it has a physical meaning, it can be easily calculated and unlike entropy its argument can be negative. The latter is very important because the autocorrelation function usually has negative values.

The present work is outlined as follows: In Section 2 the deformed Lambert functions and their associated disentropies are briefly reviewed; In Section 3 the disentropy of the autocorrelation function of some random signals is considered; Section 4 shows the disentropy of the autocorrelation of chaotic signals; In Section 5 the usage of randomness measure in a miscellaneous of situations is considered. At last, the conclusions are drawn in Section 6.

2. The generalized Lambert functions and their associated disentropies

The Lambert W function is an important elementary mathematical function that has been used in different areas of mathematics, computer Science and physics [1-6]. Basically, the Lambert W function is defined as the solution of the equation

$$W(z) e^{W(z)} = z. \quad (3)$$

Equation (3) has infinite solutions, however, only two of them return a real value when the argument z is real. In the interval $-1/e \leq z \leq 0$ there exist two real values of $W(z)$. The branch for which $W(z) \geq -1$ is the principal branch named $W_0(z)$ while the branch satisfying $W(z) \leq -1$ is named $W_{-1}(z)$. For $z \geq 0$ only $W_0(z)$ is real and for $z < -1/e$ there are not real solutions. The point $(z_b = -1/e, W(z_b) = -1)$ is the branch point where the solutions W_0 and W_{-1} have the same value and $dW/dz = \infty$. The plot of $W(z)$ versus z is shown in Fig. 1.

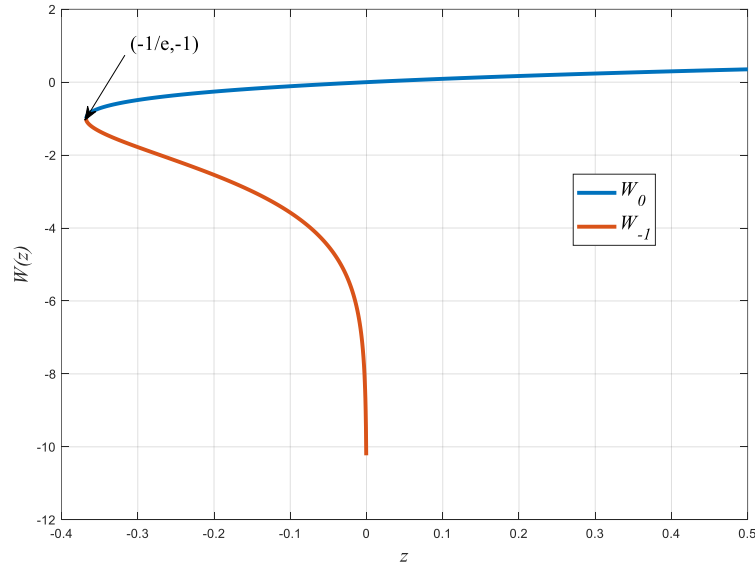


Fig. 1 – $W(z)$ versus z

Taking the logarithm in both sides of eq. (3) one gets

$$\log(z) = W(z) + \log[W(z)]. \quad (4)$$

Thereby, the entropy can be written as

$$S = \sum_i p_i \log(p_i) = \sum_i p_i W(p_i) + \sum_i p_i \log[W(p_i)]. \quad (5)$$

where $\{p_1, p_2, \dots, p_n\}$ is a probability distribution. The term

$$D = \sum_i p_i W(p_i) \quad (6)$$

is named disentropy [7]. When the disentropy is minimal entropy is maximal and vice-versa. Equation (6) is the disentropy related to Boltzmann-Gibbs entropy. If eq. (3) is modified to

$$R_2(z) 2^{R_2(z)} = z, \quad (7)$$

the solution of eq. (7) is

$$R_2(z) = \log_2(e) W\left(\frac{z}{\log_2(e)}\right) \quad (8)$$

and the disentropy related to Shannon entropy is

$$D = \sum_i p_i R_2(p_i). \quad (9)$$

One can modify eq. (3) once more by exchanging the exponential by Tsallis q -exponential [7]

$$W_q(z) e_q^{W_q(z)} = z, \quad (10)$$

where Tsallis q -exponential is given by

$$e_q^z = \begin{cases} e^z & q = 1 \\ [1 + (1-q)z]^{1/(1-q)} & q \neq 1 \text{ \& } 1 + (1-q)z \geq 0 \\ 0 & q \neq 1 \text{ \& } 1 + (1-q)z < 0 \end{cases} \quad (11)$$

The function $W_q(z)$ is named Lambert-Tsallis function and it also has two branches with real solutions.

Taking the q -logarithm in both sides of (10), one gets

$$\log_q(z) = W_q(z) + \log_q[W_q(z)] + (1-q)W_q(z)\log_q[W_q(z)]. \quad (12)$$

where

$$\log_q(z) = \begin{cases} \log(z) & x > 0 \text{ \& } q = 1 \\ \frac{x^{(1-q)} - 1}{1-q} & x > 0 \text{ \& } q \neq 1 \\ \text{undefined} & x \leq 0 \end{cases} \quad (13)$$

Therefore, Tsallis q -entropy [8] can be written as

$$S_T = -\sum_i p_i^q \log_q(p_i) = -\sum_i p_i^q W_q(p_i) - \sum_i p_i^q \log_q[W_q(p_i)] - (1-q)\sum_i p_i^q W_q(p_i) \log_q[W_q(p_i)]. \quad (14)$$

The term

$$D_q = \sum_i p_i^q W_q(p_i) \quad (15)$$

is the disentropy related to Tsallis q -entropy.

At last, one can also modify eq. (3) by using the q -product operation [9]

$$R_{qQ}(z) \times_Q e_q^{R_{qQ}(z)} = z, \quad (16)$$

where

$$a \times_q b = \max \left\{ \left[a^{(1-q)} + b^{(1-q)} - 1 \right]^{1/(1-q)}, 0 \right\}. \quad (17)$$

Doing $q = Q$ and taking the q -logarithm in both sides of (16) one gets

$$\log_q(z) = R_{qq}(z) + \log_q[R_{qq}(z)]. \quad (18)$$

Thus, Tsallis q -entropy can now be written as

$$S_T = -\sum_i p_i^q \log_q(p_i) = -\sum_i p_i^q R_{qq}(p_i) - \sum_i p_i^q \log_q[R_{qq}(p_i)]. \quad (19)$$

The term

$$D_{qq} = \sum_i p_i^q R_{qq}(p_i) \quad (20)$$

is a second disentropy related to Tsallis q -entropy.

Some examples (upper branch) for W_q and R_{qq} are shown below [7,9]:

$$W_2(z) = \frac{z}{z+1}, \quad z > -1, \quad (21)$$

$$W_{3/2}^+(z) = \frac{2(z+1) + 2\sqrt{2z+1}}{z}, \quad z > -1/2, \quad (22)$$

$$R_{2,2}(z) = -\frac{1}{2z} \pm \frac{1}{2} \sqrt{\frac{1}{z^2} + 4}, \quad z \geq 0, \quad (23)$$

$$R_{1/2,1/2}(z) = 2(z^{1/2} + 1) - 2\sqrt{2z^{1/2} + 1}, \quad z \geq 0. \quad (24)$$

It can be shown the branch point of the Lambert-Tsallis W_q function is $(z_b = \exp_q(1/(q-2))/(q-2), W_q(z_b) = 1/(q-2))$, for $q \neq 2$. There is no branch point with finite z_b for $q = 2$. For $q = 1$, the branch point of Lambert W function is recovered. The solution in the interval $z_b \leq z < 0$ is $W_q^-(z)$ while the solution in the interval $z_b \leq z < \infty$ is $W_q^+(z)$. For example, Fig. 2 shows the plot of $W_{q=3/2}$ versus z .

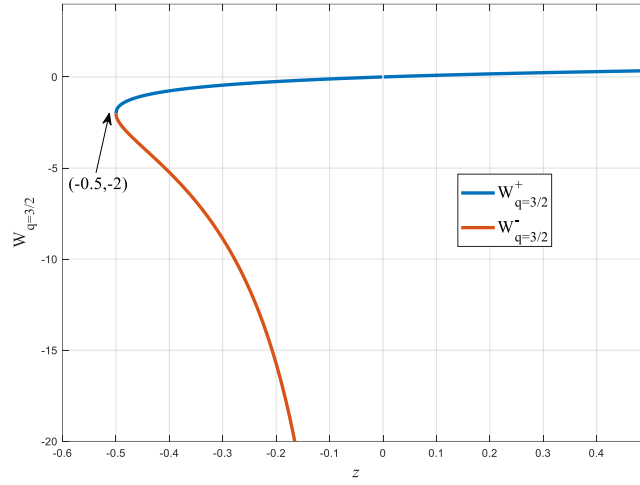


Fig.2. $W_{q=3/2}(z)$ versus z .

One may note it is also possible to modify eq. (3) by using the Kaniadakis κ -exponential [10]. In this case, the solutions will be the Lambert-Kaniadakis W_κ functions, and another disentropy based on W_κ can be introduced [11].

The disentropy can be used everywhere entropy is used [12], however, there are situations where only the disentropy can be used, like the quantumness measure proposed in [13] and the randomness measure proposed in this work. At last, although only the discrete disentropies were described, their continuous version is straightforward. For example, if $p(x)$ is a probability density distribution, then its continuous disentropy D_q is given by

$$D_q = \int_{-\infty}^{\infty} p^q(x) W_q(p(x)) dx. \quad (25)$$

3. The disentropy of the autocorrelation function of noisy signals

In this work only the disentropy $D_{q=2}$ (eqs. (15) and (21)) will be used, since it can be easily calculated and the autocorrelation function is inside of its domain. Thus, the randomness measure of a continuous signal $s(t)$ is

$$D_2 = \int_{-\infty}^{\infty} \frac{R^3(\tau)}{R(\tau)+1} d\tau. \quad (26)$$

while for a discrete series s_t one gets

$$D_2 = \sum_{n=1}^N \frac{r_n^3}{r_n + 1}, \quad (27)$$

where r_n is the n -th value of the discrete autocorrelation function of s_t . For a maximally random discrete signal, the autocorrelation function is a delta function with magnitude equal to one hence, according to eq. (27), the disentropy value of a maximally random discrete signal is $D_2 = 0.5$. If one desires a normalized randomness measure, one can use $R = 0.5/|D_2|$. In this case one gets $0 < R \leq 1$. In this work only the non-normalized version will be used.

Let us start using the signal given in eq. (28).

$$s(t) = \sin(2\pi t) + 3\sin(3\pi t) + A\varepsilon(t). \quad (28)$$

In eq. (28) A is the variable that controls the noise power (the noise power is proportional to A^2) and $\varepsilon(t)$ is a random variable with uniform distribution in the interval $[0,1]$. The curve disentropy versus A for $s(t)$ is shown in Fig. 3.

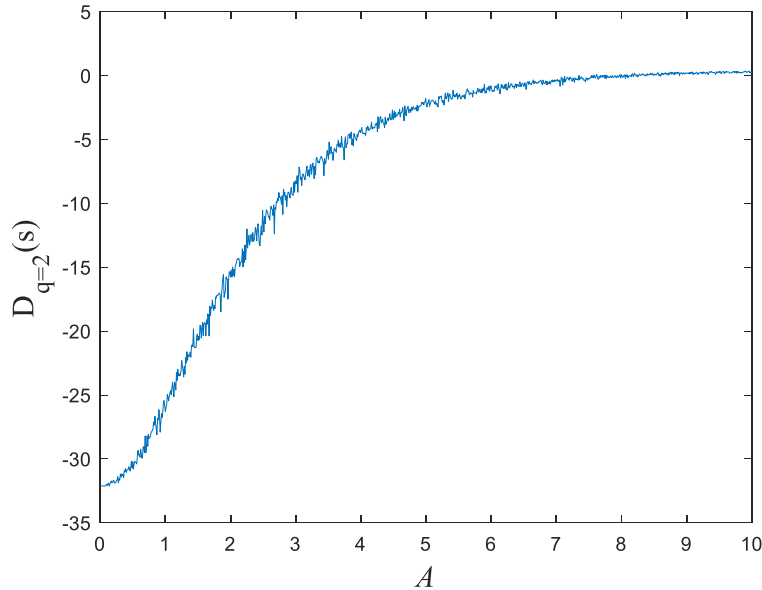


Fig. 3 – $D_2(s)$ versus A for $s(t)$ given in eq. (28).

For each value of A , eq. (28) was simulated only once, this is the reason the curve is not so smooth. However, its behaviour is very clear, the larger the noisy power the larger the randomness, until the saturation is reached.

The second signal to be considered is given in eq. (29), a signal with noise in the frequency.

$$s(t) = 5 \sin\left(2\pi\left[5 + A\varepsilon(t)\right]t\right). \quad (29)$$

The curve disentropy versus A for $s(t)$ is shown in Fig. 4.

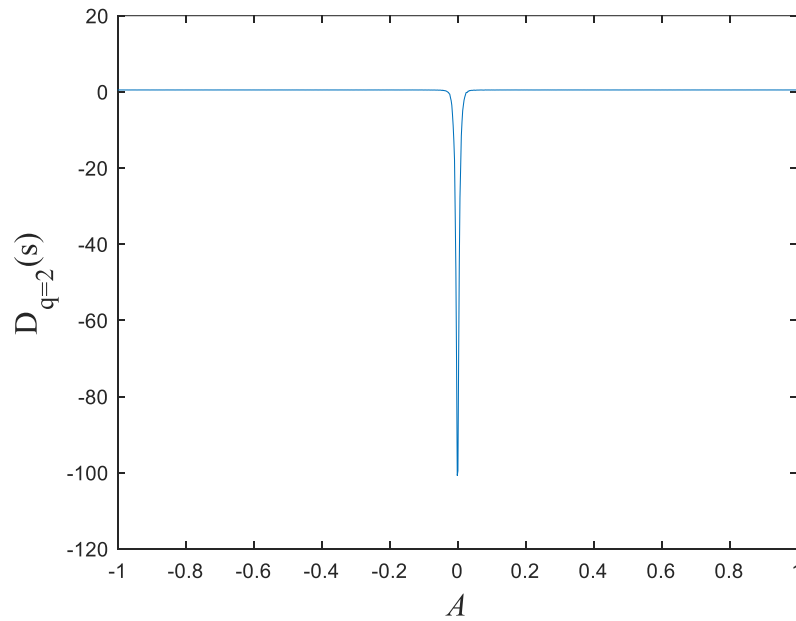


Fig. 4 – $D_2(s)$ versus A for $s(t)$ given in eq. (29).

As one can see in Fig. 4, once more the larger the noise power the larger the randomness.

At last, eq. (30) shows a signal with phase noise.

$$s(t) = 5 \sin\left(2\pi t + A\varepsilon(t)\pi\right). \quad (30)$$

The curve disentropy versus A for $s(t)$ is shown in Fig. 5.

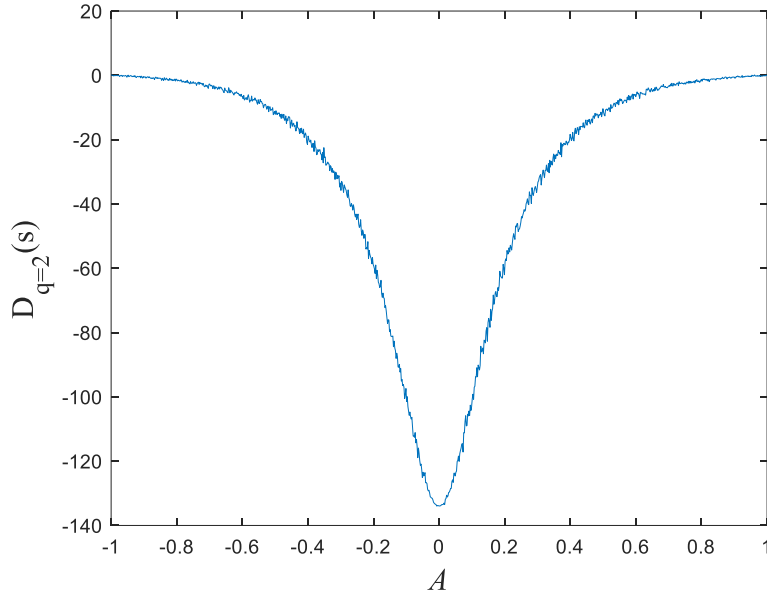


Fig. 5 – $D_2(s)$ versus A for $s(t)$ given in eq. (30).

As before, the larger the noise power the larger the randomness. Comparing Figs. 3, 4 and 5, one can see that, for a telecommunication system, the noise is more harmful when it acts in the frequency or phase. In other words, coherent telecommunication systems are more susceptible to noise than telecommunication systems that use amplitude (or power) modulation.

4. The disentropy of the autocorrelation function of chaotic signals

As it is well known, a non-linear system working in the chaotic regime can simulate a random variable, in the sense that the chaotic signal, even being deterministic, it can pass in some statistical tests of randomness. Hence, it is interesting to analyse the randomness of a chaotic signal using the disentropy of the autocorrelation. The non-linear system here considered is the logistic map

$$x_{n+1} = \lambda x_n (1 - x_n). \quad (31)$$

The curve disentropy versus λ for x_n is shown in Fig. 6. The chaotic behaviour of the logistic map starts for $\lambda \approx 3.56995$, the red vertical line in Fig. 6. However, the randomness is not maximal at this point. It is

maximal only for λ close to 3.7. Therefore, different chaotic regimes can show different randomness. In this case, if the logistic map is used in a chaotic cryptographic scheme, it will be more secure if $\lambda \geq 3.7$.

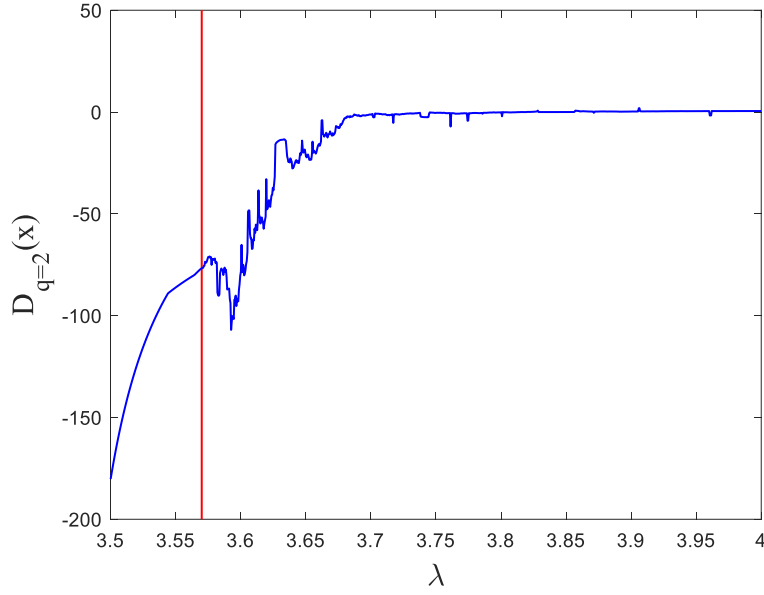


Fig. 6 – $D_2(x)$ versus λ for x given in eq. (31).

At last, consider the noisy logistic map shown in eq. (32).

$$x_{n+1} = k3.9x_n(1-x_n) + \sqrt{1-k^2}0.2\varepsilon_n. \quad (32)$$

The curve of the disentropy versus k for x_n is shown in Fig. 7.

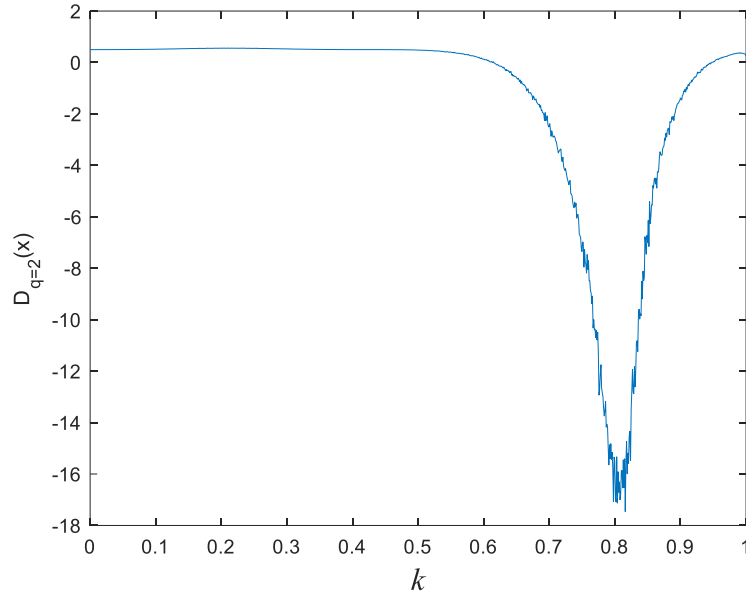


Fig. 7 – $D_2(x)$ versus k for x given in eq. (32).

Equation (32) is a mixture of a chaotic process with a random process hence, a large amount of randomness is expected. In fact, the minimal value of randomness, that occurs for $k \sim 0.8$, is much larger than the randomness measured when the logistic map is still working in the periodic regime ($\lambda \leq 3.56995$ in Fig. 6).

5. The disentropy of the autocorrelation function in a miscellaneous of problems

In this section, the randomness measure is used in very different situations. Initially Lena's image contaminated with Gaussian noise is considered. The n -th pixel of Lena's image has the value $x_n = x_{n0} + A\varepsilon$ where x_{n0} is the noiseless value and ε is a random variable with normal distribution ($0, \sigma^2 = 1$). Thus, the parameter A gives the standard deviation. In Fig. 8 one can see the disentropy D_2 versus A . As expected, the larger the variance of the noise the larger the randomness. In Fig. 9 one can see Lena's image without noise ($A = 0$ and $D_2 \sim 4.5$) and with a strong noise ($A = 1$ and $D_2 \sim 0.5$).

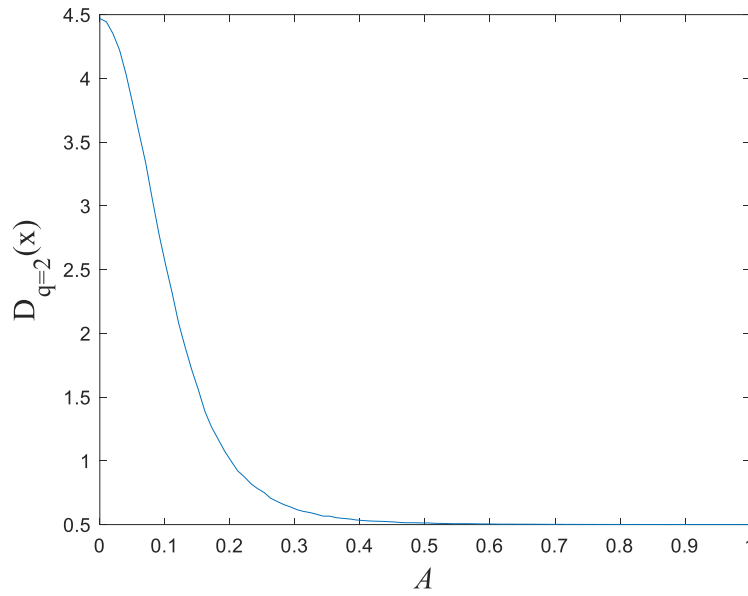


Fig. 8 – $D_2(x)$ versus standard deviation of the Gaussian noise with zero mean.



Fig. 9 – Lena's image without noise ($A = 0$) (left) and with a strong noise ($A = 1$) (right).

One should note the randomness of an image can be considered a signature, hence, it is a parameter that can be used to distinguish between real and false (or modified) images.

Consider now a periodic input signal with PSK (phase-shift key) modulation (two symbols), passing through a Rayleigh fading channel characterized by a Gaussian Doppler spectrum with normalized standard deviation $\sigma_N^2 = \sigma^2/f_d$, where f_d denotes the maximum Doppler shift. The input $x(t)$ (top) and output $y(t)$ (bottom) signals can be seen in Fig. 10. Figure 11 shows the randomness of the output signal, $y(t)$, versus σ_N^2 (two paths with average gain of 2 and 3 and delays of 0 and $1.5 \cdot 10^{-4}$ s).

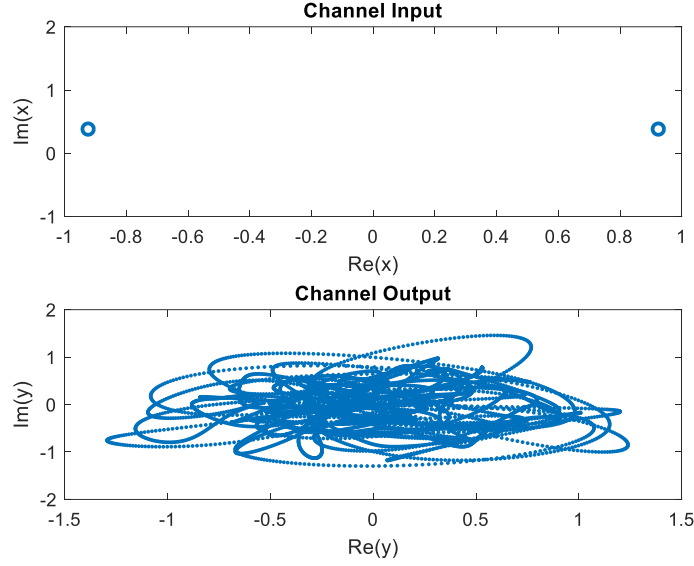


Fig. 10 – Input $x(t)$ (top) and output $y(t)$ (bottom) of a Rayleigh fading channel.

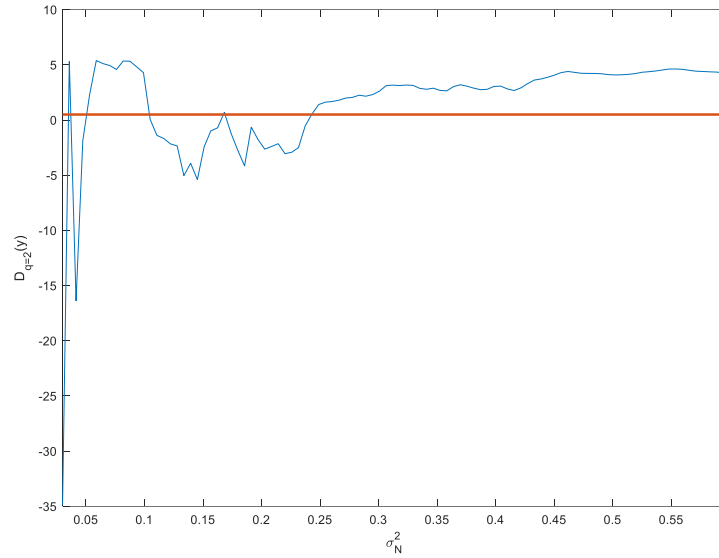


Fig. 11 – Randomness of the output signal, $y(t)$, versus the normalized standard deviation of the Gaussian Doppler spectrum, σ_N^2 .

The disentropy of the autocorrelation measures how much the input signal was deteriorated by the channel. In fact, if $x(t)$ and $y(t)$ are, respectively, the input and output signals of a noisy process (like a channel or an amplifier, for example) then one can use the randomness measure to create a new noise figure: $NF = D_2(y(t))/D_2(x(t))$.

In some situations, a random process is controlled aiming to keep the randomness inside of an allowed range. This happens for example in countries that control the dollar currency exchange rate. Figure 12 shows the price of dollar in Brazil (in Brazilian currency *reais*) during five thousands days (finishing in 2020 July, 2nd) (top) and the variation of the randomness (bottom) during the same period. The action of the Brazilian government along the last years kept the randomness low and almost constant.

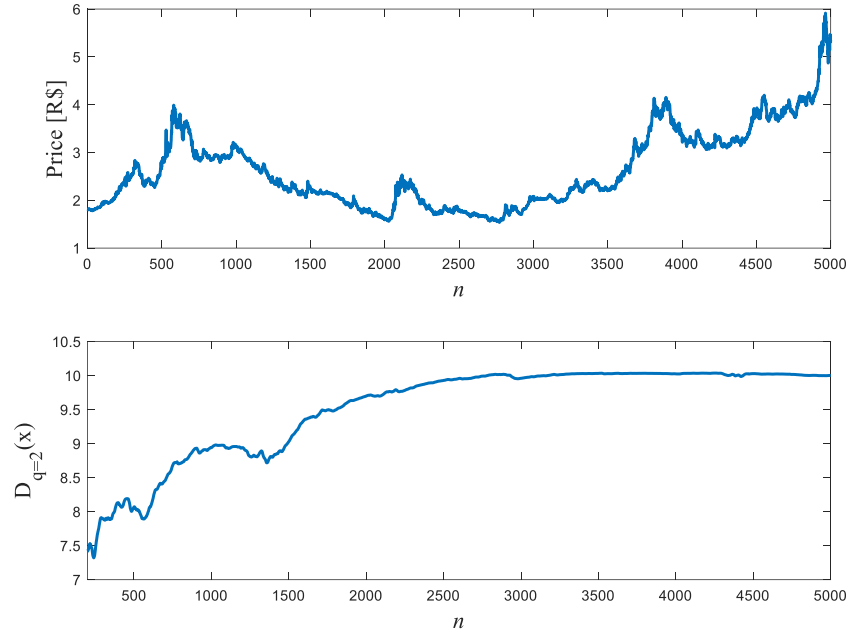


Fig. 12 – Variation of the dollar exchange rate in Brazil during 5000 days (top) and its randomness (bottom).

An example of a random process with controlled high randomness is given by eq. (33).

$$x_{n+1} = \begin{cases} 0.5x_n + \varepsilon_n, & n < 100 \\ 0.5x_n + |D_2(x_n)|\varepsilon_n & n \geq 100 \end{cases} \quad (33)$$

In (33) ε_n is once more a random variable with normal distribution. If the randomness increases (decreases), $|D_2(x)|$ decreases (increases), decreasing (increasing) also the variance of the Gaussian noise. If the variance decreases (increases) the randomness also decreases (increases). Hence, the feedback tends to keep the randomness constant. One can see this effect in Fig. 13. The variable x_n around $n = 100$ can be seen in Fig. 14.

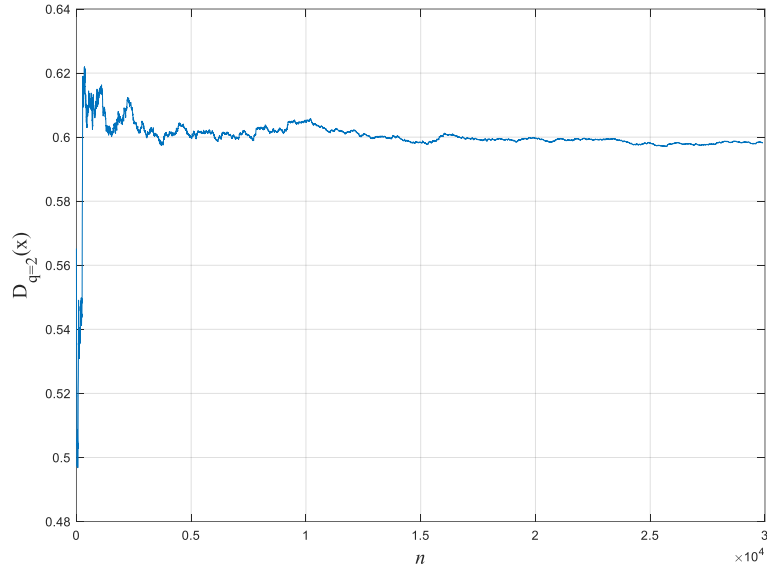


Fig. 13 – Randomness of random process with feedback given in eq. (33).

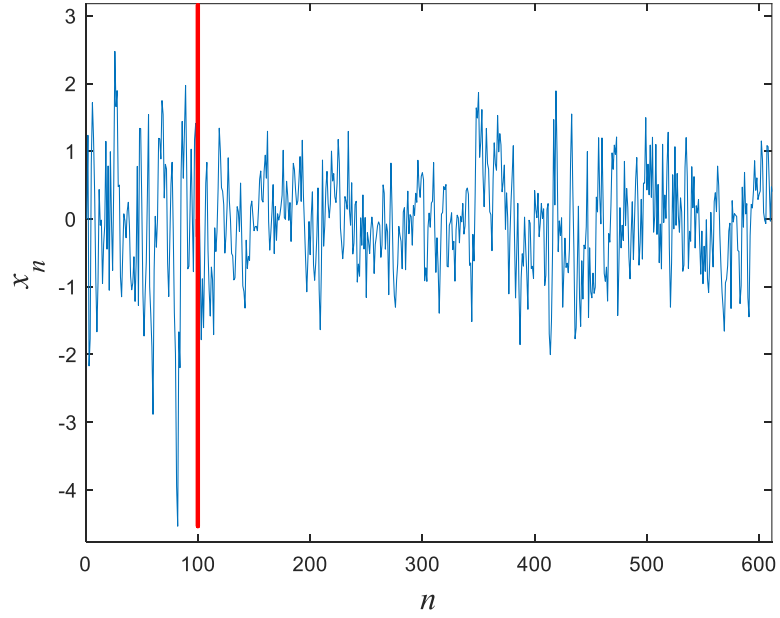


Fig. 14 – x_n versus n (eq. (33)).

At last, Fig. 15 shows the randomness of the sequence x_n formed by the first n thousands, $n \in [1, 210]$, nucleotides of the genome of the Chinese mitten crab [14]. As one can note, the randomness is very large along the sequence, showing, as expected, that randomness plays an important role in biology.

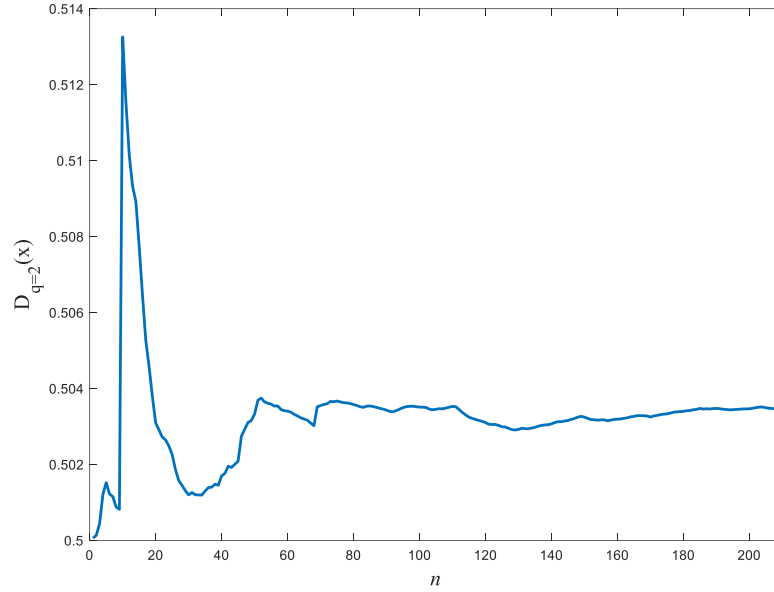


Fig. 15 – Randomness of the first n thousands nucleotides of the genome of the Chinese mitten crab.

6. Conclusions

In this work it was shown the disentropy of the autocorrelation function can be used as a measure of randomness of signals and time series. Initially three sinusoidal signals were considered since one can have noise in amplitude, frequency and phase. The result obtained clearly pointed that coherent communication systems are more susceptible to noise than amplitude modulated-based communication systems. This can be easily seen in optical systems for example, where up to now coherent systems still have problems to become commercial. On the other hand, since a chaotic signal can emulate a random signal, it is very natural to apply the randomness measure to analyse the real randomness of a chaotic signal. As it was pointed out, the randomness depends on the parameters values of the nonlinear system. This very important since a large number of works have used chaotic cryptography to protect secret data and the security of such systems depends on the randomness of the key used. Hence, the randomness measure here proposed can help cryptographers to optimize the parameters of the nonlinear system used.

While statistical tests can say if a binary sequence is random or not with some probability, the randomness measure using the disentropy returns a value (real number), not only ‘yes’ or ‘no’ with some probability of success. This permits the ease comparison between time series or signals. Furthermore,

since it can be applied to digital and analog signals, one can measure the randomness of a binary key, a speech signal, an image, an ECG signal, the first k digits of π or other mathematical special number, economic and biological data, among others.

At last, one can consider several important situations where the knowledge of the amount of randomness of a signal measured by the disentropy of the correlation is important. For example, the randomness of an image can be considered as a signature of it. In other words, it is a parameter that can be used to distinguish between real and fake image. One can also define a new noise figure: the ratio between the randomness at the output and input. This new noise figure would be very useful to characterize communication channels and electronic/optical amplifiers, for example. As third example, a medical doctor can use the amount of randomness of an ECG signal to decide what kind of medicine should be introduced to a patient.

Acknowledgements

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, and CNPq via Grant no. 307184/2018-8. Also, this work was performed as part of the Brazilian National Institute of Science and Technology for Quantum Information.

References

1. R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey and D. E. Knuth, *On the Lambert W function*, Advances in Computational Mathematics, vol. 5, 329 – 359, 1996.
2. S. R. Valluri, D. J. Jeffrey, R. M. Corless, *Some applications of the Lambert W function to Physics*, Canadian Journal of Physics, vol. 78 n° 9, 823-831, 2000.
3. D. C. Jenn, *Applications of the Lambert W function in Electromagnetics*, IEEE Antennas and Propagation Magazine, vol. 44, n° 3, 2002.
4. F. C.-Blondeau and A. Monir, *Numerical evaluation of the Lambert W function and application to generation of generalized Gaussian noise with exponent $1/2$* , IEEE Transactions on Signal Processing, vol. 50, no. 9, 2160-2165, 2002.
5. D. Veberic, *Having fun with Lambert $W(x)$ function*, GAP-2009-114 [Online]. Available: <http://arxiv.org/abs/1003.1628>, 2010.

6. K. Roberts, S. R. Valluri, Tutorial: *The quantum finite square well and the Lambert W function*, Canadian Journal of Physics, vol. 95, no. 2, 105-110, 2017.
7. G. B. da Silva and R.V. Ramos, *The Lambert-Tsallis W_q function*, Physica A, 525, 164-170, 2019.
8. C. Tsallis, *Possible generalization of Boltzmann-Gibbs statistics*, J. Stat. Phys. 52, 479, 1988.
9. R. V. Ramos, *The $R_{q,Q}$ function and the q -diode*, Physica A, 556, 124851 2020. DOI: 10.1016/j.physa.2020.124851
10. G. Kaniadakis, *Statistical mechanics in the context of special relativity*, Physical Review E, v. 66, n. 5, p. 056125 2002.
11. L. E. da Silva, G. B. da Silva. R. V. Ramos, *The Lambert-Kaniadakis W_κ function*, Phys. Lett. A, 2019. DOI: 10.1016/j.physleta.2019.126175
12. R. V. Ramos, Quantum and classical information theory with disentropy, ArXiv/quant-ph:1901.04331, 2020.
13. R. V. Ramos, *Disentropy of the Wigner function*, J. of Opt. Soc. of Am. B, 36, 8 2244, 2019.
14. Downloaded from <https://www.ncbi.nlm.nih.gov/nuccore/LQIF01000001.1?report=fasta>.