



An Evaluation of Intrinsic Mode Function Characteristic of Non-Gaussian Autorregresive Processes

Fernando Pose¹, Javier Zelechower², Marcelo Risk¹,
and Francisco Redelico^{1,3}(✉)

¹ Instituto de Medicina Traslacional e Ingeniería Biomédica,
Hospital Italiano de Buenos Aires, Instituto Universitario del Hospital Italiano
de Buenos Aires, CONICET, Perón 4190 - (C1199ABB),
Ciudad Autónoma de Buenos Aires, Argentina
francisco.redelico@hospitalitaliano.org.ar

² Department of Information Systems Engineering, National Technological
University, Buenos Aires, Argentina

³ Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes,
Roque Sáenz Peña 352, B1876BXD Bernal, Buenos Aires, Argentina

Abstract. Empirical mode decomposition (EMD) is a suitable transformation to analyse non-linear time series. This work presents an empirical study of intrinsic mode functions (IMFs) provided by the empirical mode decomposition. We simulate several non-gaussian autoregressive processes to characterize this decomposition. Firstly, we studied the probability density distribution, Fourier spectra and the cumulative relative energy to each IMF as part of the study of empirical mode decomposition. Then, we analyze the capacity of EMD to characterize, both the autocorrelation dynamics and the marginal distribution of each simulated stochastic process. Results show that EMD seems not to only discriminate autocorrelation but also the marginal distribution of simulated processes. Results also show that entropy based EMD is a promising estimator as it is capable to distinguish between correlation and probability distribution. However, the EMD entropy does not reach its maximum value in stochastic processes with uniform probability distribution.

Keywords: Empirical mode decomposition (EMD) · Entropy-complexity informational plane · Intrinsic mode function (IMF) · Stochastic Processes

1 Introduction

Empirical mode decomposition was introduced for the first time by, Huang and collaborators in [9]. It is a fully adaptive technique useful to analyze non-linear and non-stationary time series. The basic idea of EMD method is decompose the actual time series into a number of intrinsic mode functions (IMFs) as the sum of these IMFs will recover the original time series. Even this technique has been have a great impact in analysing time series, it has the drawback to be a totally

empirical technique and, from our knowledge, it does not have a completely mathematical formulation that allow us to explore its properties, so their should be evaluated empirically [16].

Despite of non-gaussian process are ubiquitous within science and technology in such diverse areas as random number generators [11], modeling irregularly spaced transaction financial data [5], foreign exchange rate volatility modeling [8], studying nervous systems mechanism (Spike sorting) [6] among others, the properties of IMFs generated from them seems to be limited evaluated [16].

For the reasons above, we propose analyse the statistical features of the IMFs generated from autoregressive process of order 1 endowed with Gaussian, Exponential or Uniform marginal probability distribution as prototype of non-gaussian autoregressive processes [21]. We divide our analysis in two parts, in the first part, we analyse the time-frequency domain computing the probability distribution, mean, peaks, periods and the Fourier spectrum of each intrinsic characteristic function of the respective stochastic process and within the second part we extract energetic and entropy like features of the IMFs, the density energy of each IMF, the accumulated energy, EMD entropy and the relative position in the entropy-complexity plane shows the dynamical nature of each IMF.

The paper reads as follows: Sect. 2 briefly explain the EMD technique and the autoregressive process generation algorithm, Sect. 3 shows the statistical features of the IMF from the series explained in the previous section and finally Sect. 4 is devoted to conclusions.

2 Empirical Mode Decomposition Features and Stochastic Processes

We describe the stochastic processes simulated and the features from the respective IMF along with the EMD algorithm in order to make the article selfcontained and accessible for a wider audience.

2.1 Stochastic Processes

We present all the necessary results to simulate the stochastic processes analysed in the next section, readers are referred to the original paper of each random process [3, 12] or to [22] for a deeper explanation. There should be better algorithm implementations, in terms of memory use and algorithmical complexity, however these algorithms are good enough for clarity and for reproducibility purposes.

A general linear stochastic process is modeled as generated by a linear aggregation of random shocks [2] as the autoregressive model of order p ,

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + a_t \quad (1)$$

where the current value z_t is expressed as a finite, linear aggregate of previous values of the process $\{z_{t-1}, z_{t-2}, \dots, z_{t-p}\}$ and a random shock a_t , distributed with mean 0 and finite variance σ^2 .

Three kind of autoregressive process are considered within this paper: Gaussian, Exponential and Uniform. The difference between them is random shock a_t form. Within this paper, for sake of simplicity we set the order of the autoregressive process $p = 1$ leading to the first-order autoregressive process AR(1).

In the Gaussian process, Eq. 2 takes the form:

$$z_t = \phi_1 z_{t-1} + a_t \quad (2)$$

where the current value z_t is expressed as a finite, linear aggregate of the previous values of the process z_{t-1} and an independent and identically distributed (i.i.d.) random shocks a_t that has marginal Gaussian distribution with mean 0 and variance σ^2 . We use the standard R command [15] `arima.sim()` to simulate this process.

For the Exponential marginally distributed AR(1) process, we use the NEARA(1) model [12] what is schematic present in Algorithm 3 and Eq. 2 takes the form:

$$z_t = a_t + \begin{cases} \beta \cdot z_{t-1} & \text{w.p } \alpha \\ 0 & \text{w.p } 1 - \alpha \end{cases} \quad (3)$$

with

$$a_t = \begin{cases} e_t & \text{w.p } \frac{1-\beta}{1-(1-\alpha)\beta} \\ (1-\alpha) \cdot \beta \cdot e_t & \text{w.p } \frac{\alpha\beta}{1-(1-\alpha)\beta} \end{cases} \quad (4)$$

where *w.p.* stands by *with probability*, $\alpha > 0$ and $\beta > 0$ are free correlation parameters such as $\rho = \alpha\beta$, providing that α and β are not both equal to one, a_t has a p Exponential distribution. In Algorithm 3 we sketch the main features of the simulation. We simulate a time series of length N (line 1) starting with a realization of a random variable exponentially distributed with $\lambda = 1$ (line 2), we set the desire correlation coefficient, in our simulations (line 3) and compute α y β accordingly (lines 4 y 5), within the loop (lines 7–14) we resolve Eqs. 3 and 4. In order to generate the random variable a_t , we generate an exponential random variable with $\lambda = 1$ (line 8) and the selected it according a Binomial distribution with parameter $\frac{1-\beta}{(1-\beta+\alpha*\beta)}$ (lines 10–12). Lines 13 and 14 shows Ec. 3 when finally the stochastic process is fully generated.

Algorithm 1. Exponential Autoregressive Process generation

```

1:  $N \leftarrow$  Set the desired time series spam)
2:  $x[1] \sim \text{Exp}(\lambda = 1)$ 
3:  $\rho \leftarrow$  Set the desired autocorrelation coefficient)
4:  $\alpha \leftarrow \frac{2*\rho}{(1+\rho)}$ 
5:  $\beta \leftarrow \frac{1}{(2-\alpha)}$ 
6:  $t \leftarrow 2$ 
7: for  $t := 2$  to  $N$  do
8:
    $En \sim \text{Exp}(\lambda = 1)$ 
9:  $pr \leftarrow \frac{1-\beta}{(1-\beta+\alpha*\beta)}$ 
10:  $bn \sim \text{Bin}(n = 1, p = pr)$ 
11:  $a_t[bn == 1] \leftarrow En[bn == 1]$ 
12:  $a_t[bn == 0] \leftarrow En[bn == 0]$ 
13:  $bn \sim \text{Bin}(n = 1, p = \alpha)$ 
14:  $x[t] < -a_t + bn * \beta * x[t - 1]$ 

```

Algorithm 2. Uniform Autorregresive Process generation

```

1:  $N \leftarrow$  Set the desired time series spam)
2:  $k \leftarrow$  Set  $k$ , which determine the desired autocorrelation coefficient)
3:  $\rho \leftarrow \frac{1}{k}$ 
4:  $x[1] \sim \text{sample}(\text{seq}(0, (k-1)/k, 1/k), 1, \text{replace}=\text{TRUE})$ 
5: for  $t := 2$  to  $N$  do
6:
    $x[t] < -rho * x[t - 1] + \text{sample}(\text{seq}(0, (k - 1)/k, 1/k), 1, \text{replace} = \text{TRUE})$ 

```

In the Uniform autoregressive model of order 1 UAR(1) [3], depicted in Algorithm 2. Eq. 2 follows,

$$z_t = \frac{1}{k} z_{t-1} + a_t \quad (5)$$

with $k \geq 2$. It has been shown in [3] that z_t would shield continuous $U(0,1)$ marginal distribution if the i.i.d. random shocks a_t is sampled as a $1/k$ uniform distribution over the set $\{0, 1/k, 2/k, \dots, \frac{k-1}{k}\}$. Algorithm 3 shows the pseudocode for the computer implementation of Eq. 5. The desired time spam and correlation coefficient are sets in line 1 and 2–3, respectively. In line 4 the Eq. 5 is realized for $t = 1$ and them in the loop for all time spam.

2.2 Empirical Mode Decomposition Features

In 1998, Huat et al. proposed empirical mode decomposition [10]. It is a fully adaptive technique that can be applied to nonlinear and non-stationary process. The basic idea of EMD method is decompose the complicated time series $x(t)$ into a number of IMFs as,

$$x(t) = \sum_{i=1}^M h_i(t) + r(t) \quad (6)$$

where M is the cardinality of the IMF set and $r(t)$ is the residual after transformation.

EMD method through which EMD decomposes the original time series into a series of IMF components with different timescales is shown in Algorithm 3. Each IMF must satisfy the following two conditions [7]: (1) the number of extreme and the number of zero crossings must be equal or differ at most by one (in the entire signal length); (2) the average value of the two envelope defined by the local maximum and the minimum must be zero at any moment. Finally, the last level is the residue of the time series which is related with the trend from time series.

Algorithm 3. EMD Algorithm

```

1:  $D(t) \leftarrow x(t)$ 
2:  $i \leftarrow 1$ 
3: while  $D(t)$  is not monotonic do
4:
    $E_{max}(t) \leftarrow interpolation(max\{x(t)\})$ 
5:  $E_{min}(t) \leftarrow interpolation(min\{x(t)\})$ 
6:  $m(t) \leftarrow (E_{max}(t) + E_{min}(t))/2$ 
7:  $D(t) \leftarrow x(t) - m(t)$ 
8: if  $D(t)$  satisfy the conditions of IMF then
9:
   else
   -
    $IMF_i(t) \leftarrow D(t)$ 
10:  $r(t) \leftarrow x(t) - IMF_i(t)$ 
11:  $x(t) = r(t)$ 
12:  $i \leftarrow i + 1$ 
13:  $x(t) \leftarrow D(t)$ 

```

2.3 Time-Frequency Related Features

In order to characterize the IMF set for several non-gaussian autoregressive processes we follow [20] and estimate:

1. Density probability distribution for each IMF using their respective histogram, using the Scott's rule [20].
We use Scott's rule for determining the bin number because this rule is optimal in the sense that asymptotically minimizes the integrated mean squared error.
2. The number of peaks of each IMF attended in the lines 4 and 5 inside the while loop in Algorithm 3 applied to each process studied and which allows determines the mean period of the function by counting the number of peaks of the function [23].
3. Mean period in terms of the mean number of samples for each IMF and each process (Gaussian, Uniform and Exponential).

4. The effect produced by EMD method for each level of decomposition from the time series was calculated with the spectrum of frequencies $X(\Omega)$ using the Discrete Fourier Transform defined as

$$X(\Omega) = \sum_{n=-\infty}^{\infty} x[t]e^{-j\Omega n} \quad (7)$$

2.4 Energy-Entropy Related Features

In [7] the EMD-entropy and EMD-statistical complexity are proposed as follows, The total energy of each IMF is proposed as $E_i = \sum_{t=1}^n [h_i(t)]^2$ where n is the length of the i -th IMF and $h_i(t)$ denotes the value n of the i IMF. Finally, the total energy of signal is the sum over all total energies E_i , $E = \sum_{i=1}^m E_i$. Hence the relative energy of each IMF is defined as

$$p_i = \frac{E_i}{E} \quad (8)$$

and following the usual discrete form of Shannon entropy, the EMD-entropy is defined using p_i in Eq. 8 $S_{EMD}(l) = -\sum_{i=1}^l p_i \times \log(p_i)$ where l should stand for the order of EMD-entropy. As the reconstruction could be done using the first $l, l < m$ IMF, and regarding all other IMF rest as the rest in Eq. 6 this new entropy has this free parameter. This entropy $H_{EMD}(l)$ is an unnormalized quantity, so to restrict its value to the $[0, 1]$ interval, it is redefined as,

$$H_{EMD}(l) = \frac{S_{EMD}(l)}{\log(l)} \quad (9)$$

as $H_{EMD}(l)$ gets its maximum value when $P = (p_1, \dots, p_l) = (1/l, \dots, 1/l)$. Another defined quantity in [13, 14] is the statistical complexity, using the Jensen-Shannon divergence, i.e.

$$C_{EMD}(l) = Q_{EMD}(l)H_{EMD}(l) \quad (10)$$

where $Q_{EMD}(l) = Q_0 \times JSD[P, P_e]$, P_e is a reference probability distribution $P = (p_1, \dots, p_l) = (1/l, \dots, 1/l)$ and Q_0 a normalization constant. Both quantities form the entropy-complexity plane

$H_{EMD}(l) \times C_{EMD}(l)$ plane, There is a vast literature regarding this plane [13, 14, 17–19]

3 Numerical Results and Discussion

The processes presented in the previous section were simulated, varying their autocorrelation coefficients value to evaluate their IMFs using the features described in the previous section as an evaluation methodology [4]. For the Gaussian processes five positively autocorrelated time series were simulated, with the

corresponding color code $\rho = ((\text{yellow}), 0, (\text{magenta}), 0.2, (\text{cyan}) 0.4, (\text{red}) 0.6, (\text{green}) 0.8, (\text{blue}) 0.9)$, all these processes were simulated with normally distributed shocks a_t with $\sigma = 1$ and $\mu = 0$. The Exponential processes were simulated with suitable values for α and β (see lines 4 and 5 in Algorithm 3) such for exponential process are $\rho = (\text{yellow}) 0, (\text{magenta}) 0.125, (\text{cyan}) 0.25, (\text{red}) 0.5, (\text{green}) 0.75$ and the autoregressive Uniform process was simulated for $\rho = (\text{yellow}) 0, (\text{magenta}) 0.1, (\text{cyan}) 0.2, (\text{red}) 0.25, (\text{green}) 0.33, (\text{blue}) 0.50$, plus the uncorrelated data with Uniform marginal distribution. All series have length $N = 10^{15}$ points. IEEE 754 double precision floating point numbers was used for all computations.

Figure 1 shows the probability density function for IMFs 1, 3 and 5 of each stochastic process. Surprisingly, Uniform distribution, no matter the autocorrelation coefficient (see Fig. 1a, 1d and 1g,) does not lead an equiprobable distribution of p_i (in Eq. 8) since the density probability distribution for the first IMF is bi-modal. This fact has a profound impact on the entropy estimation based on EMD and especially in the entropy-complexity plane as it can be seen later in this Section. Gaussian process behaves as reported previously [7] and exponential processes seem to be more likely as equiprobable distribution of p_i . By equiprobable distribution we mean that all the IMF has the same probability distribution

In Tables 1 and 2 show the number of peaks and the mean period respectively for each simulated stochastic processes. The results show that, for all processes, the mean period of any IMF component is almost exactly double that of the previous one. This result is consistent with the result obtained by Wu et al. (2004) and Flandrin et al. (2003) for Gaussian distributions. Also, the results indicate that when the autocorrelation increases the mean difference between any two IMF component increases, no matter the probability distribution. Standard deviation also increases as the autocorrelation coefficient increase, too. Table 1 shows that for times series where the value of autocorrelation is higher there are fewer peaks in each IMF component so that, the number of peaks can be related to how correlated the samples of time series are and no matter about the time series marginal distribution.

Figure 3 shows Fourier spectra (Eq. 7) from simulated stochastic processes for IMFs 1, 3 and 6. It can be observed that when the IMF component increases the content of information from the process is found in low frequency which corresponds with [23]. Furthermore, we can see in IMF 1 component that Exponential distribution has more frequency content in low frequency than Gaussian distribution and Uniform distribution. It is interesting to note that when the value of the autocorrelation coefficient increases, the frequency content increases in low frequency so that the information about autocorrelation from the processes is mainly in the lower part of the spectra. The study of the stochastic processes with Gaussian distribution indicates that when the IMF component increases the spectra is concentrated on low frequency and its amplitude increases as if the energy decreases slowly in each IMF component. For a stochastic process with uniform or exponential distribution, it does not happen.

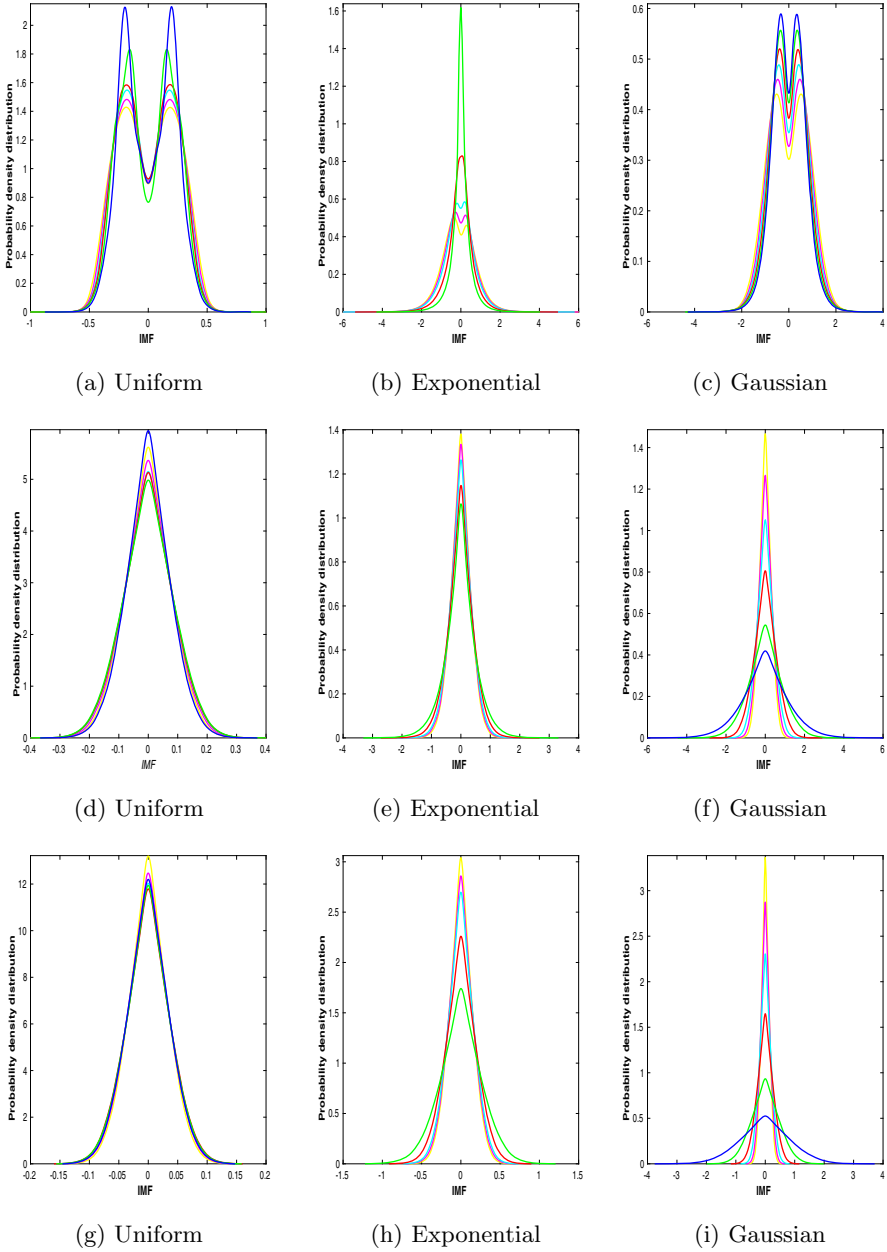


Fig. 1. Probability density distribution plots for the relative energy of IMFs 1-3-6 (Eq. 8) for the simulated stochastic processes. Color codes are for $\rho = 0, 0.10, 0.2, 0.25, 0.33, 0.5, 0.75$ are yellow, magenta, cyan, red, green, blue, respectively. (Color figure online)

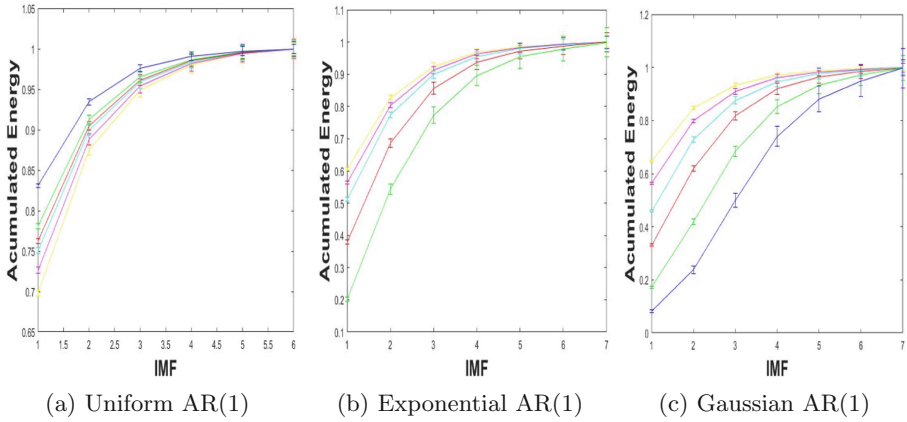


Fig. 2. Cumulative plots for the relative Energy (Eq. 8) for the simulated stochastic processes. Color codes are for $\rho = 0, 0.10, 0.2, 0.25, 0.33, 0.5, 0.75$ are yellow, magenta, cyan, red, green, blue, respectively. (Color figure online)

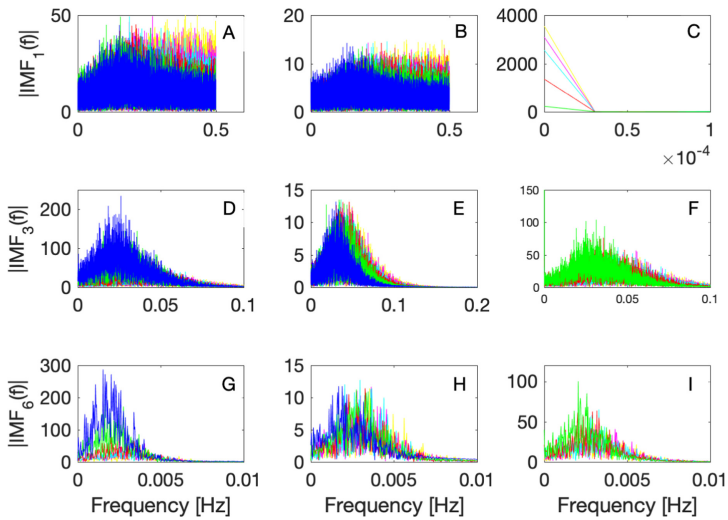


Fig. 3. Fourier spectra (Eq. 7) of IMFs 1,3 and 6 for simulated stochastic processes. A, D, G: Gaussian process. B, E, H: Uniform process. C, F, I: Exponential process. Color codes are for $\rho = 0, 0.10, 0.2, 0.25, 0.33, 0.5, 0.75$ are yellow, magenta, cyan, red, green, blue, respectively. (Color figure online)

Table 1. Number of peaks of IMFs 1–6 obtained from the simulated stochastic processes of 10^{15} data points and $\rho = (0, 0.4, 0.9)$ for Gaussian and Uniform distributions, $\rho = (0, 0.1, 0.2)$ for Exponential distribution and the standard deviation.

Gaussian distribution			
IMF	r = 0	r = 0.4	r = 0.9
1	22559 (90.5)	20897.5 (136.5)	19847.5 (252.5)
2	9126.5 (73.5)	7931 (65)	6542.5 (473.5)
3	3767 (44.5)	3206.5 (52)	2313.5 (152)
4	1569 (30.5)	1323.5 (31)	892 (58)
5	651 (19.5)	550 (14.5)	361 (25)
6	272 (12)	228.5 (12.5)	149 (12)
Uniform distribution			
IMF	r = 0	r = 0.4	r = 0.9
1	22435 (88)	20965 (115.5)	16831 (127)
2	9075.5 (53.5)	8121 (72)	6492.5 (63)
3	3757.5 (44.5)	3345 (46.5)	2632 (44.5)
4	1562.5 (30.5)	1392 (26.5)	1086.5 (23.5)
5	651 (19.5)	576.5 (18)	448 (20)
6	271 (11)	238.5 (12)	187 (10)
Exponential distribution			
IMF	r = 0	r = 0.1	r = 0.2
1	22289 (108)	20867.5 (93.5)	17300 (110)
2	8771.5 (88.5)	8135.5 (81)	6561 (94)
3	3606 (45)	3327.5 (50)	2633.5 (50)
4	1500 (31.5)	1377 (22)	1073.5 (25.5)
5	622 (20)	575 (16)	440 (16)
6	259 (11.5)	237 (10)	183 (9)

In Fig. 2, cumulative energy plots are displayed for each of the three simulated stochastic processes. As EMD allows reconstructing time series up to a certain m in Eq. 6, this plot is just a summation of the relative energy values. It is shown that with only the first five IMFs we recover most than 80% of energy for all simulated processes. It is interesting to note that the Uniform stochastic processes retrieve the information with less IMFs than the other processes and the process which needs more IMF's to retrieving the information is the Gaussian. In turn, as the correlation increases, the number of IMFs necessary to retrieve the information also increases. This characteristic could be an advantage of this method since it would allow knowing the size of the entropy by means of an explicit criterion. In other words, if the researcher were to agree to recover 90% of the time series information, would search for the value of m that makes

Table 2. Mean periods of IMFs 1–6 obtained from the simulated stochastic processes of 10^{15} data points and $\rho = (0, 0.4, 0.9)$ for Gaussian and Uniform distributions, $\rho = (0, 0.1, 0.2)$ for Exponential distribution in terms of the number of samples and the standard deviation.

Gaussian distribution			
IMF	r = 0	r = 0.4	r = 0.9
1	1.4525 (0.0058)	1.568 (0.0102)	1.651 (0.021)
2	3.5904 (0.0289)	4.1316 (0.0339)	5.0085 (0.3446)
3	8.6987 (0.1026)	10.2192 (0.1655)	14.1639 (0.9013)
4	20.8846 (0.4064)	24.7586 (0.5791)	36.7354 (2.354)
5	50.3349 (1.5092)	59.5782 (1.5695)	90.7701 (6.2761)
6	120.4706 (5.3175)	143.4055 (7.7988)	219.9195 (17.6223)
Uniform distribution			
IMF	r = 0	r = 0.4	r = 0.9
1	1.4606 (0.0057)	1.5630 (0.0086)	1.9449 (0.0146)
2	3.6106 (0.0213)	4.0350 (0.0357)	5.0471 (0.049)
3	8.7207 (0.1035)	9.7961 (0.1362)	12.4498 (0.2107)
4	20.9715 (0.4103)	23.5402 (0.4467)	30.1592 (0.6552)
5	50.3349 (1.5092)	56.8396 (1.7537)	73.1429 (3.2308)
6	120.9151 (4.8920)	137.3926 (6.8883)	175.2299 (9.3773)
Exponential distribution			
IMF	r = 0	r = 0.1	r = 0.2
1	1.4701 (0.0071)	1.5703 (0.0070)	1.8941 (0.0120)
2	3.7357 (0.0378)	4.0278 (0.0400)	4.9944 (0.0715)
3	9.0871 (0.1134)	9.8476 (0.1480)	12.4428 (0.2348)
4	21.8453 (0.4593)	23.7967 (0.3799)	30.5245 (0.7255)
5	52.6817 (1.6971)	56.9878 (1.5943)	74.4727 (2.6967)
6	126.5174 (5.6308)	138.2616 (5.7875)	179.0601 (8.8600)

that recovery and then compute Shannon entropy using the p_i set. We have arbitrarily selected the value of $m = 5$ and we let $m = 6$ as the residual energy in the plot.

Figure 4 shows three entropy-complexity planes. It is interesting that the Exponential processes are located in the right part and below the plane (see Fig. 3b) instead the uniform processes are located in the middle part of the plane (entropy = 0.6). This differs significantly from previous results referring to the same processes but located in the plane using the permutation entropy [22]. On the one hand, using the symbolization of Band and Pompe [1] to calculate the probabilities, the location of the three processes in the plane is practically indistinguishable, suggesting that the symbolization proposed by Band and

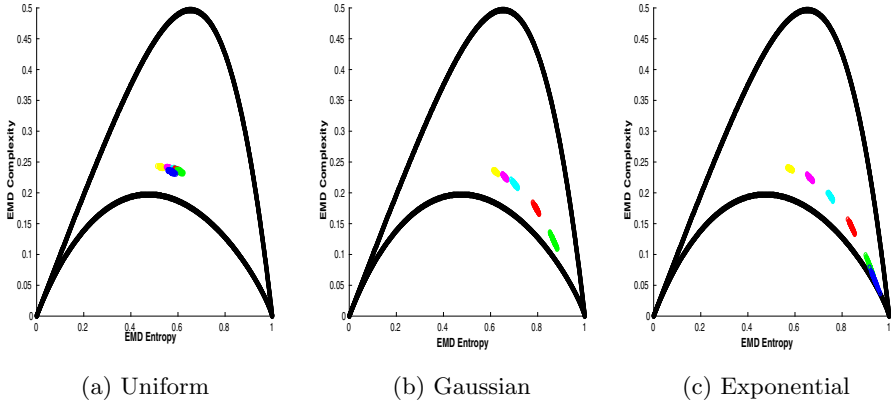


Fig. 4. Several EMD Entropy-Complexity planes for the simulated stochastic processes. Color codes are for Uniform are $\rho = (0, 0.1, 0.2, 0.25, 0.33, 0.50)$, for Exponential process are $\rho = (0, 0.125, 0.25, 0.5, 0.75)$ and for Gaussian process are $\rho = (0, 0.2, 0.4, 0.6, 0.8, 0.9)$ (Color figure online)

Pompe does not distinguish between probability distributions if not only between different correlations; on the other hand, in the methodology proposed in [7] entropy recognizes both characteristics, the marginal distribution of the stochastic process and autocorrelation. On the other hand, regarding the densities shown in Fig. 3, the order of the distributions is not as expected, one would have expected the maximum entropy to manifest in uniform stochastic processes, however, it is shown that these processes have the minimum entropy among the three analyzed and that the maximum entropy occurs in exponential stochastic processes.

4 Conclusions

Exploring features for IMF descriptions is an interesting task, useful in order to characterize actual time series. If a real time series is tough as a realization of an stochastic process endowed with a correlation structure and a marginal probability distribution, the IMF features could reveal some features of both.

Regarding the distribution if the energy for each IMF component and each stochastic process with different distribution is calculated, for a process with Gaussian distribution the energy loss in each IMF is lower compared with the previous IMF than the process with uniform or Exponential distribution. Thus, the EMD method in a Non-Gaussian process produces a large decrease in the energy when IMF component increment. We also observe Exponential distribution has more frequency content in low frequency than Gaussian and Uniform distribution.

According to the correlation value from distribution we observe for high values of autocorrelation coefficient that the frequency content increases in low frequency while for low values the frequency content decreases.

EMD also seems not to only discriminate autocorrelation but also the marginal distribution of simulated processes, as can be seen in Fig. 4. This would give it good performance as a classifier shown in [7] and should be explored. However, this strategy to calculate the probabilities shows a challenge and that is that the maximum entropy is not found in the scenario of greater uncertainty, that is, in the uniform distribution, as it can be seen in Fig. 4. This characteristic must be taken into account if the intention of the investigation is the characterization between chaos and noise. In particular, statistical complexity includes a measure of distance from a reference probability, and this reference is generally established as the equilibrium probability and this is not the case if we follow the proposal in [7].

This is a step ahead in using EMD to characterize several stochastic processes presents in science and technology. Even more research is needed in validating this approach with several probability distribution and correlation structures we think the results present in this contribution are general enough to a first glimpse into the subject.

References

1. Bandt, C., Pompe, B.: Permutation entropy: a natural complexity measure for time series. *Phys. Rev. Lett.* **88**(17), 174102 (2002)
2. Box, G.E., Jenkins, G.M., Reinsel, G.C., Ljung, G.M.: *Time Series Analysis: Forecasting and Control*. Wiley, Hoboken (2015)
3. Chernick, M.R.: A limit theorem for the maximum of autoregressive processes with uniform marginal distributions. *The Annals of Probability*, pp. 145–149 (1981)
4. Davidson, E.J.: *Evaluation Methodology Basics: The Nuts and Bolts of Sound Evaluation*. Sage, Thousand Oaks (2005)
5. Engle, R.F., Russell, J.R.: Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica*, pp. 1127–1162 (1998)
6. Farashi, S.: Spike sorting method using exponential autoregressive modeling of action potentials. *World Acad. Sci. Eng. Technol. Int. J. Med. Health Biomed. Bioeng. Pharmaceutical Eng.* **8**(12), 864–870 (2015)
7. Gao, J., Shang, P.: Analysis of complex time series based on EMD energy entropy plane. *Nonlinear Dyn.* **96**(1), 465–482 (2019)
8. Hafner, C.: *Nonlinear time series analysis with applications to foreign exchange rate volatility*. Springer Science & Business Media (2013)
9. Huang, N.E., Shen, Z., Long, S., Wu, M., Shih, H., Zheng, Q., Tung, C., Liu, H.: The empirical mode decomposition and hilbert spectrum for nonlinear and nonstationary time series analysis. *Proc. Roy. Soc. A* **545**(1971), 903–995 (1998)
10. Huang, N.E., et al.: The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. Roy. Soc. London. Series A: Math. Phys. Eng. Sci.* **454**(1971), 903–995 (1998)
11. Lawrance, A.: Uniformly distributed first-order autoregressive time series models and multiplicative congruential random number generators. *J. Appl. Probability* **29**, 896–903 (1992)
12. Lawrance, A., Lewis, P.: A new autoregressive time series model in exponential variables (near (1)). *Advances in Applied Probability*, pp. 826–845 (1981)

13. Lopez-Ruiz, R.: Complexity in some physical systems. *Int. J. Bifurcation Chaos* **11**(10), 2669–2673 (2001)
14. Lopez-Ruiz, R., Mancini, H., Calbet, X.: A statistical measure of complexity. arXiv preprint nlin/0205033 (2002)
15. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2020). <https://www.R-project.org/>
16. Rilling, G., Flandrin, P., Goncalves, P., et al.: On empirical mode decomposition and its algorithms. In: *IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing*, vol. 3, pp. 8–11. NSIP-03, Grado (I) (2003)
17. Rosso, O., Larrondo, H., Martin, M., Plastino, A., Fuentes, M.: Distinguishing noise from chaos. *Phys. Rev. Lett.* **99**(15), 154102 (2007)
18. Rosso, O.A., Carpi, L.C., Saco, P.M., Ravetti, M.G., Plastino, A., Larrondo, H.A.: Causality and the entropy-complexity plane: Robustness and missing ordinal patterns. *Physica A* **391**(1), 42–55 (2012)
19. Rosso, O.A., Olivares, F., Zunino, L., De Micco, L., Aquino, A.L., Plastino, A., Larrondo, H.A.: Characterization of chaotic maps using the permutation bandt-pompe probability distribution. *Eur. Phys. J. B* **86**(4), 1–13 (2013)
20. Schlotthauer, G., Torres, M.E., Rufiner, H.L., Flandrin, P.: Emd of gaussian white noise: effects of signal length and sifting number on the statistical properties of intrinsic mode functions. *Adv. Adapt. Data Anal.* **1**(04), 517–527 (2009)
21. Sengupta, D., Kay, S.: Efficient estimation of parameters for non-gaussian autoregressive processes. *IEEE Trans. Acoust. Speech Signal Process.* **37**(6), 785–794 (1989)
22. Traversaro, F., Redelico, F.O.: Characterization of autoregressive processes using entropic quantifiers. *Physica A* **490**, 13–23 (2018)
23. Wu, Z., Huang, N.E.: A study of the characteristics of white noise using the empirical mode decomposition method. *Proc. Roy. Soc. London. Series A: Math. Phys. Eng. Sci.* **460**(2046), 1597–1611 (2004)