Available online at www.sciencedirect.com
SCIENCE DIRECT®
PHYSICS LETTERS A
Physics Letters A 312 (2003) 27–33
www.elsevier.com/locate/pla

Analysis of signals in the Fisher–Shannon information plane

C. Vignat [a,b,*], J.-F. Bercher [b]

[a] *EECS, University of Michigan, Ann Arbor, MI, USA*
[b] *Équipe Signal et Information, ESIEE and UMLV, France*

Received 21 February 2003; accepted 26 March 2003

Communicated by C.R. Doering

Abstract

We show that the analysis of complex, possibly non-stationary signals, can be carried out in an information plane, defined by both Shannon entropy and Fisher information. Our study is exemplified by two large families of distributions with physical relevance: the Student-t and the power exponentials.

PACS: 02.50.Cw; 02.50.Ey; 65.40.Gr

Keywords: Complexity; Fisher information; Shannon–Boltzmann entropy; Power exponential; Student-t distribution

1. Introduction

Fisher Information Measure (FIM) was introduced by Fisher in 1925 [9] in the context of statistical estimation. In the last ten years, a growing interest for this information measure has arisen in theoretical physics. In a seminal paper [1], Frieden has characterized FIM as a versatile tool to describe the evolution laws of physical systems; one of his major results is that the classical evolution equations (e.g., the Schrödinger wave equation, the Klein–Gordon equation, the Helmoltz wave equation, the diffusion equation, the Boltzmann and Maxwell–Boltzmann law) can be derived from the minimization of FIM under proper constraint. As another important result, a version of the H-theorem has been extended—under the name I-theorem—to the notion of Fisher information [3]. Until recently, Shannon (Boltzmann) entropy was considered as the major tool to describe the informational behavior and complexity of physical systems. The theoretical contributions cited above suggest that this judgment should be revised and that FIM appears as an appealing alternative to Shannon entropy.

Since FIM allows an accurate description of the behavior of dynamic systems, its application to the characterization of complex signals issued from these systems appears quite natural. This approach was adopted for example by Martin et al. [2] for the characterization of EEG signals. One of the interesting results of their study was that FIM allowed the detection of some non-stationary behavior in situations where the Shannon entropy shows limited dynamics.

* Corresponding author. Université de Marne-la-Vallée, Cité Descartes, 77454 Champs-sur-Marne cedex 2, France.

E-mail addresses: vignat@univ-mlv.fr (C. Vignat), bercherj@esiee.fr (J.-F. Bercher).

0375-9601/03/$ – see front matter
doi:10.1016/S0375-9601(03)00570-X

Motivated by this work, we define a Fisher–Shannon information plane and show that the simultaneous examination of both Shannon entropy and FIM may be required to characterize the non-stationary behavior of a complex signal. More precisely, we exhibit two families of physically relevant signals, the Tsallis signals and the power exponential signals which have the unexpected property that their temporal trajectory in the Fisher–Shannon (FS) information plane can be arbitrarily designed. As a consequence, any non-stationarity measurement device based on only one of these two measures would yield a suboptimal result.

This Letter is organized as follows: in Section 2, we review some basic properties of entropy and Fisher information measures and introduce the notion of the Fisher–Shannon plane. In Section 3, we study two families of parameterized random variables and their locations in the Fisher–Shannon plane. These results are used in Section 4 to build explicitly two families of signals whose trajectories in the FS plane can be designed arbitrarily.

2. Fisher's information measure and Shannon entropy power

In the following, we consider a random variable X whose probability density function is denoted as $f_X(x)$. Its Shannon entropy writes

$$H_X = -\int f_X(x) \log f_X(x)\, dx \tag{1}$$

and its Fisher information measure writes

$$I_X = \int \left(\frac{\partial}{\partial x} f_X(x) \right)^2 \frac{dx}{f_X(x)}. \tag{2}$$

For convenience, we will use, rather than entropy, the alternative notion of entropy power (see [4]), defined by

$$N_X = \frac{1}{2\pi e} e^{2H_X}. \tag{3}$$

Among other properties, both measures N_X and I_X verify a set of resembling inequalities.

- *Evolution law*: if X represents the state of a system, the second law of the thermodynamics writes

$$\delta N_X \leqslant 0 \tag{4}$$

whereas the I-theorem writes accordingly

$$\delta I_X \leqslant 0. \tag{5}$$

- *Superadditivity property*: the Shannon entropy of the sum of two independent random variables verifies the entropy power inequality [4]

$$N_{X+Y} \geqslant N_X + N_Y \tag{6}$$

and the corresponding Fisher information inequality [4] writes

$$I_{X+Y}^{-1} \geqslant I_X^{-1} + I_Y^{-1}. \tag{7}$$

Two additional properties will help us to track the information trajectory of a random signal.

- *The scaling property*: when scaling a random variable X by a scalar factor $a \in \mathbb{C}^*$, the entropy power and FIM transform as follows

$$N_{aX} = |a|^2 N_X,$$
$$I_{aX}^{-1} = |a|^2 I_X^{-1}.$$

- *The uncertainty property*:

$$N_X I_X \geqslant 1 \tag{8}$$

with equality if and only if X is a Gaussian random variable. A proof of this property can be found, for example, in [4].

Both scaling and uncertainty properties enlight the fact that FIM and Shannon entropy are intrinsically linked, so that the characterization of signals should be improved when considering their location in the Fisher–Shannon (FS) plane.

Definition 1. The FS area, denoted as $\mathcal{D}$, is the set of all reachable values of FIM and Shannon entropies, namely

$$\mathcal{D} = \left\{ (N, I) \mid N \geqslant 0,\ I \geqslant 0 \text{ and } NI \geqslant 1 \right\}. \tag{9}$$

Next, as a consequence of the scaling property, we note that a scaled version aX of a random variable X

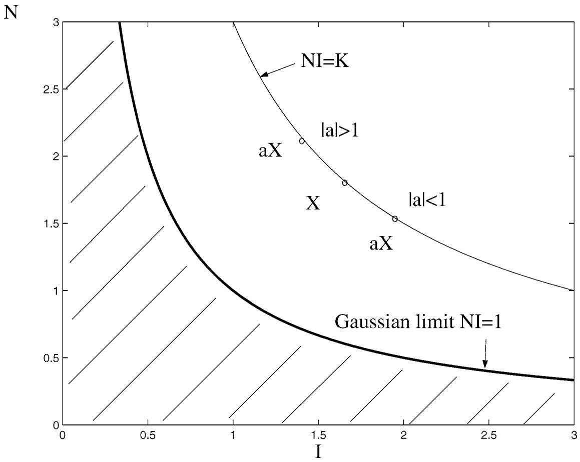

Fig. 1. The FS area.

belongs to the same $NI = K$ curve as X, as illustrated on Fig. 1.

3. Two families of random variables

In this part, we study two families of probability densities, namely the Student-t and the power exponential distributions. Both families include heavy tailed densities and the Student-t family is a member of the larger class of power laws. Furthermore, the Gaussian distribution is a particular case of both families and the uniform and exponential distributions are special cases of the power exponential family.

For both types of distributions, we compute the Shannon entropy power and FIM and the area spanned by these distributions in the FS plane.

3.1. Student-t random variable

The importance of Student-t distributions in statistical physics has been highlighted in the seminal work of Tsallis (see [7] and references therein). A part of the versatility of these distributions is due to the fact that they maximize under variance constraint a relevant statistical quantity, namely the Tsallis entropy

$$H_X^{(q)} = \frac{1}{q-1}\left(1 - \int f_X^q(x)\,dx\right), \tag{10}$$

where $q > 0$ is the extensivity parameter. Their exact expression is the following:

$$f_S(m,\sigma,x) = \frac{\Gamma\left(\frac{m+1}{2}\right)}{\sigma\sqrt{m-2}\,\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{1}{2}\right)} \times \left(1 + \frac{x^2}{(m-2)\sigma^2}\right)^{-\frac{m+1}{2}}$$
$$\text{for } (m > 2, 1 > q > 0, x \in \mathbb{R}) \tag{11}$$

where $m = \frac{1+q}{1-q}$ and σ is a scale parameter. These distributions and their properties have been extensively characterized, including in the multivariate case, in [6]. Note that the Gaussian law can be recovered from the Student-t family by letting $m \to +\infty$ or $q \to 1^-$. For convenience, we cite here two properties of these distributions that are helpful for the numerical simulations considered in Section 4.

- If X is distributed according to a Student-t law $f_S(m,\sigma,x)$, then a stochastic representation of X writes
$$X = \frac{\mathcal{N}}{\chi_m}, \tag{12}$$
where χ_m is a gamma distributed random variable with parameter m, independent of $\mathcal{N}$ which is Gaussian with zero mean and variance $\sigma^2(m-2)$.
- Moreover, if m is integer, χ_m is distributed as a chi variable with m degrees of freedom and a stochastic representation of X is thus
$$X = \frac{\mathcal{N}}{\sqrt{\sum_{k=1}^m \mathcal{N}_k^2}}, \tag{13}$$
where the $\{\mathcal{N}_k\}_{1 \leqslant k \leqslant m}$ are independent Gaussian random variables with unit variance.

The information measures of a Student-t law are characterized in the following theorem.

Theorem 1. *The entropy power and the Fisher information measure of the Student-t distribution with parameters m and σ are*

$$N_S(m,\sigma) = \frac{1}{2\pi e}\left(\frac{\sigma\sqrt{m-2}\,\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{m+1}{2}\right)}\right)^2 \times e^{(m+1)\left(\psi\left(\frac{m+1}{2}\right)-\psi\left(\frac{m}{2}\right)\right)}, \tag{14}$$

$$I_S(m,\sigma)=\frac{1}{\sigma^2}\frac{m(m+1)}{(m-2)(m+3)}, \tag{15}$$

where ψ denotes the digamma function.

Proof. By direct computation. Result (14) has been already obtained, for example, in [8]. □

We are now in position to determine the area spanned by all Student-t laws in the Fisher–Shannon plane.

Theorem 2. *Denote $\mathcal{D}_S$ the area of the FS plane defined by*

$$\begin{cases} 1\leqslant I_S N_S\leqslant \frac{3e^5}{80\pi}, \\ I_S\geqslant 0, \quad N_S\geqslant 0. \end{cases} \tag{16}$$

Then the application

$$\begin{aligned} &]2,+\infty[\times[0,+\infty[\,\to\mathcal{D}_S,\\ &(m,\sigma)\mapsto(I_S,N_S) \end{aligned} \tag{17}$$

is a bijection.

Proof. The product $I_S(m,\sigma)N_S(m,\sigma)$ depends only on m since σ is scale parameter. Denote $h_S(m)$ this function so that

$$h_S(m)=\frac{m(m+1)}{2e(m+3)}\frac{\Gamma^2\left(\frac{m}{2}\right)}{\Gamma^2\left(\frac{m+1}{2}\right)} \times e^{(m+1)\left(\psi\left(\frac{m+1}{2}\right)-\psi\left(\frac{m}{2}\right)\right)} \quad (m>2). \tag{18}$$

It is easy to check that $h_S(m)$ decreases from $+\infty$ to 1 for $2<m<+\infty$ and that $\lim_{m\to+\infty}h_S(m)=1$ and $h_S(2)=\frac{3e^5}{80\pi}$. Thus the equation

$$I_S N_S=h_S(m) \tag{19}$$

determines uniquely $2\leqslant m\leqslant+\infty$. Then, solving (15) in variable σ^2 yields a unique positive value of σ^2. □

3.2. Power exponential random variables

Let us now consider a second kind of random variable denoted as $X_{\mathrm{PE}}(\lambda,\gamma)$ whose probability density is of the power exponential type:

$$f_{\mathrm{PE}}(\lambda,\gamma,x)=\frac{\gamma\lambda^{1/\gamma}}{2\Gamma\left(\frac{1}{\gamma}\right)}\exp\left(-\lambda|x|^\gamma\right) \quad x\in\mathbb{R}, \tag{20}$$

where $\lambda>0$ is a scale parameter and $\gamma>1$ is a shape parameter. Note that $\gamma=2$ gives the Gaussian case while $\gamma\to+\infty$ leads to a uniform distribution. This versatility explains that these distributions are used to model accurately some physically realistic quantities, for instance the amplitudes of wavelets coefficients. These distributions can also be shown to maximize the Shannon entropy under moment constraint [10].

The information measures associated with this probability density function are characterized in the following theorem.

Theorem 3. *The entropy power and the FIM of the power exponential law with parameters λ and γ are*

$$N_{\mathrm{PE}}(\lambda,\gamma)=\frac{2}{2\pi e}\left(\frac{1}{\gamma}\Gamma\left(\frac{1}{\gamma}\right)\left(\frac{e}{\lambda}\right)^{1/\gamma}\right)^2, \tag{21}$$

$$I_{\mathrm{PE}}(\lambda,\gamma)=\frac{\Gamma\left(1-\frac{1}{\gamma}\right)}{\Gamma\left(\frac{1}{\gamma}\right)}\gamma(\gamma-1)\lambda^{2/\gamma}. \tag{22}$$

Proof. By direct computation. Result (21) has already been derived in [5]. □

Now given any point (I,N) in the Fisher–Shannon area, the following theorem proves that an exponential power random variable $X_{\mathrm{PE}}(\lambda,\gamma)$ having I and N as respective FIM and entropy power can be found.

Theorem 4. *Both applications*

$$\begin{aligned} &]0,+\infty[\times]2,+\infty[\,\to\mathcal{D},\\ &(\lambda,\gamma)\mapsto(I,N) \end{aligned} \tag{23}$$

and

$$\begin{aligned} &]0,+\infty[\times]1,2[\,\to\mathcal{D},\\ &(\lambda,\gamma)\mapsto(I,N) \end{aligned} \tag{24}$$

are bijections.

Proof. Suppose that we are given a couple $(I,N)\in\mathcal{D}$ and we want to determine (λ,γ). As λ is a scale parameter, the product $N_{\mathrm{PE}}(\lambda,\gamma)I_{\mathrm{PE}}(\lambda,\gamma)\geqslant 1$ is a function of the shape parameter γ only: denote $h_{\mathrm{PE}}(\gamma)$ this function so that

$$h_{\mathrm{PE}}(\gamma)=\frac{2e^{\frac{2}{\gamma}-1}}{\pi}\Gamma\left(\frac{1}{\gamma}\right)\Gamma\left(2-\frac{1}{\gamma}\right). \tag{25}$$

A straightforward computation shows that function h_{PE} decreases from $+\infty$ to 1 for $1 < \gamma \leqslant 2$ and increases from 1 to $+\infty$ for $2 \leqslant \gamma \leqslant +\infty$. Thus equation $h_{\mathrm{PE}}(\gamma) = NI$ has two solutions γ_1 and γ_2 such that $1 < \gamma_1 \leqslant 2$ and $2 \leqslant \gamma_2$. Given one of these solutions, λ can be determined uniquely by solving equation (21) or (22). □

4. Application to signal analysis

We are now in position to build non-stationary signals having arbitrary time trajectories in the FS plane. An an example, we have designed explicitly two signals:

- A first signal, based on independent Student-t samples, whose entropy power and FIM describe through time a step-like trajectory in the domain $\mathcal{D}_S$ of the FS plane.
- A second signal, based on independent power exponential samples, whose entropy and FIM evolve through time along a circle in the domain $\mathcal{D}$ of the FS plane.

4.1. The steps trajectory

The signal $s_{\mathrm{steps}}(n)$ is built from 112 000 independent samples of a Student-t with non-stationary parameters $m(n)$ and $\sigma(n)$ chosen as follows:

- Over the first 28 000 samples (A to B), parameter $m(n)$ increases linearly from $m = 3$ to $m = 30$, of one unit each 1000 samples; parameter $\sigma(n)$ evolves according to the following rule

$$\sigma^2(n) = I_S\big(m(n), 1\big) \tag{26}$$

  so that the FIM of $s_{\mathrm{steps}}(n)$ remains equal to 1 over all samples.
- Over the next 28 000 samples (B to C), parameter $m\,(n)$ decreases linearly from $m = 30$ to $m = 3$, of one unit each 1000 samples; parameter $\sigma(n)$ evolves according to the following rule

$$\sigma(n) = \frac{1}{\sqrt{N_S(m(n), 1)}} \tag{27}$$

  so that the entropy power remains equal to 1 over all samples.
- These two first steps are repeated accordingly over the 56 000 remaining samples so that the final trajectory in the FS plane looks as shown on Fig. 2: the first 28 000 of them (C to D) have a linearly decreasing entropy power with FIM fixed to 1.3 and conversely for the last part of the signal (D to E) where the entropy power is fixed to $N = 0.75$. Fig. 3 shows the corresponding signal $s_{\mathrm{steps}}(n)$.

This example shows that any measurement based on one only of the two information measures would yield a suboptimal detection.

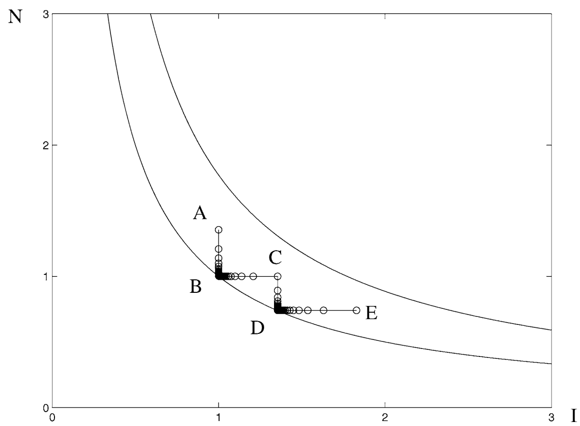

Fig. 2. The step-like trajectory of $s_{\mathrm{steps}}(n)$.

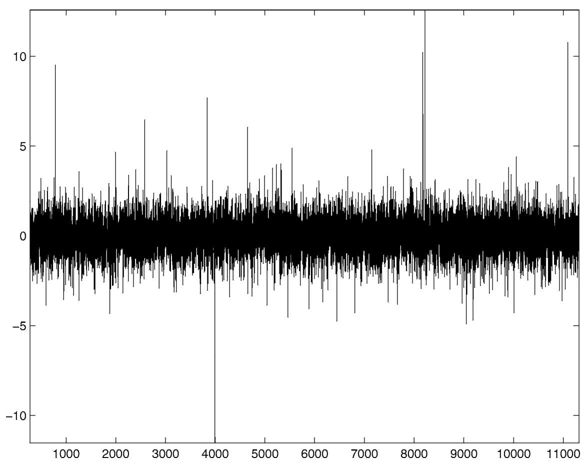

Fig. 3. Signal $s_{\mathrm{steps}}(n)$.

4.2. The circle trajectory

A signal was generated by drawing independent samples of power exponential distribution such as their entropy powers and FIM describe the circle trajectory in the FS plane as depicted on Fig. 4: to each point of the circle correspond 1000 samples, the circle being described clock-wise starting from angle $\theta = 0$, and step $\Delta\theta = \pi/10$. Details of the whole signal, corresponding to angles $\theta = \pi/5$ and $\theta = \pi + \pi/5$, are shown on Fig. 5.

This example shows once again that any trajectory can be designed in the FS plane and that a non-stationarity tracking device should inspect both FIM and entropy power to provide a good characterization of the signal.

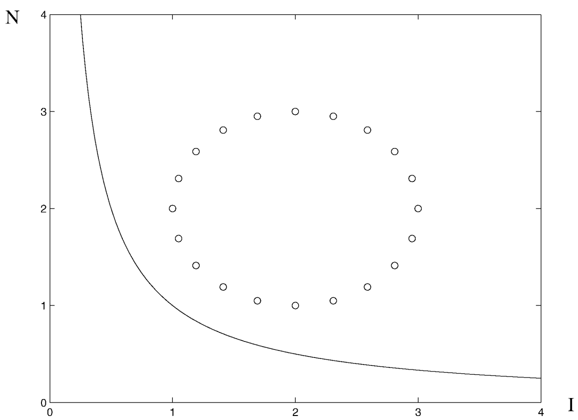

Fig. 4. Trajectory of $s_{\text{circle}}(n)$.

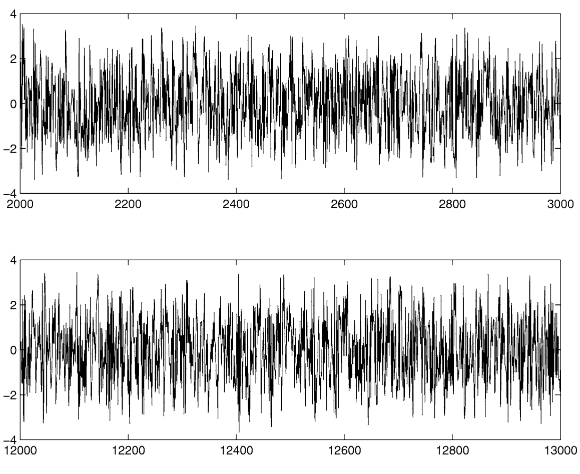

Fig. 5. Signal $s_{\text{circle}}(n)$.

4.3. Variance analysis

Given any point (I, N) in the area $\mathcal{D}_S$ of the FS plane, it follows from Theorems 2 and 4 that there exists a unique Student-t and a unique power exponential distribution having (I, N) as coordinates. But the variances of these distributions may also differ or not.

Fig. 6 shows constant variance trajectories of Student-t and power exponentials distributions: curve (1) (respectively (2)) corresponds to the family of Student-t (respectively power exponentials) with variance equal to 0.4, while curve (3) (respectively (4)) corresponds to a variance of 0.3. We deduce from this graph that there exist Student-t and power exponential laws that may be distinguished (separated) in the FS plane, although their variances coincide. Hence, in a non-stationary context, analysis in the FS plane may be useful.

On Fig. 7, a signal with a constant variance $\sigma^2 = 2$ is presented: first and last 2000 samples were generated according to a power exponential with parameters $(\lambda, \gamma) = (0.25, 2.05)$ and coordinates $(0.975, 1.74)$ in the FS plane; the 2000 center samples were generated according to a Student-t with parameters $(m, \sigma^2) = (3.5, 2)$ and coordinates $(0.807, 1.567)$.

Conversely, curves (5) and (6) in Fig. 6 correspond respectively to the trajectories of all Student-t distributions with variance equal to 36 and to all power expo-

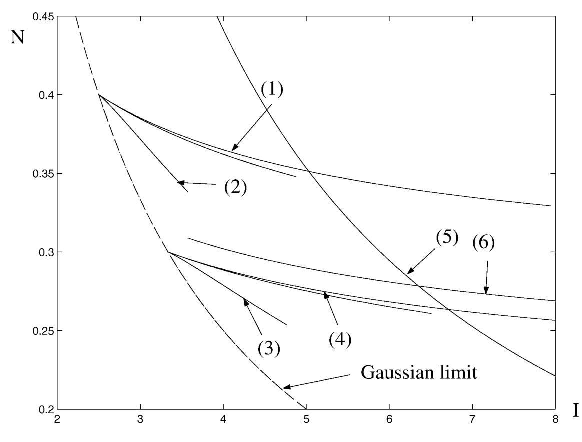

Fig. 6. Variance study in the FS plane.

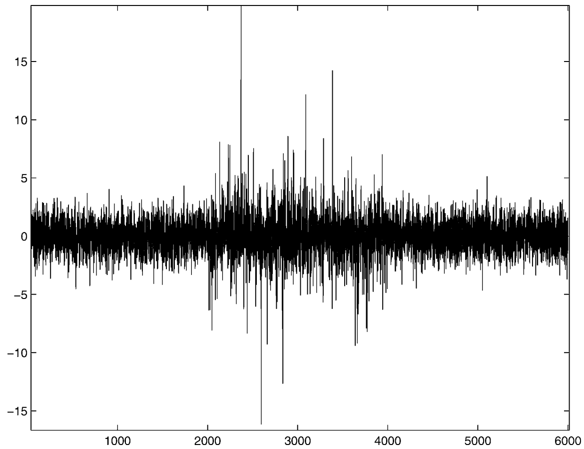

Fig. 7. Signal $s(n)$ with constant variance but different FS locations.

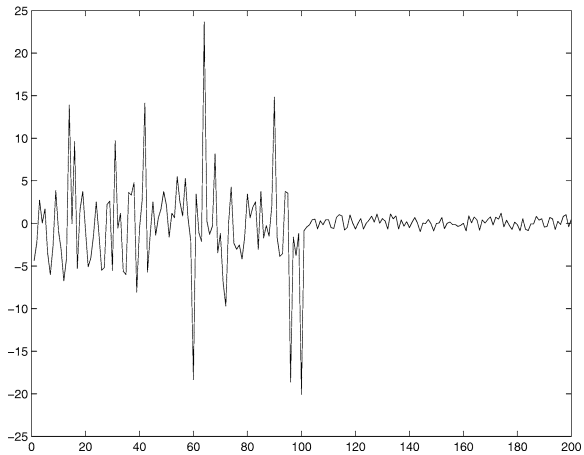

Fig. 8. Signal $s(n)$ generated according to distributions with same FS location but different variances.

nentials with variance 0.317. The intersection of these two curves correspond to a Student with parameters $m = 2.005$ and $\sigma = 6$ and to a power exponential with parameters $\lambda = 1$ and $\gamma = 6.25$. These two distributions are thus indistinguishable in the FS plane, while their variances are very different. A signal $s(n)$ generated from 1000 independent samples of the first distribution followed by 1000 samples of the second distribution, is shown on Fig. 8.

From this study, we conclude that in order to capture the dynamics of the complex signal, the additional information provided by variance may be useful, resulting in the analysis in the extended space (σ^2, I, N).

5. Conclusion

In this Letter, we have shown that two information measures, the Fisher information measure and the Shannon entropy power, can be used jointly in a context of non-stationarity detection. We have exhibited families of relevant signals that can behave arbitrarily in the Fisher–Shannon plane, proving that a device based on only one of these measures may fail to track the non-stationarity of the signal. We insist on the fact that the Student-t and power exponential distributions used here are not trivial examples, since both families span a large set of classical distributions and as such are often used to model complex data.

We can conclude that the EEG signals exhibited in [2] have a nearly horizontal trajectory in the FS plane and thus are better described by the inspection of their Fisher information rather by their entropy (power). We are currently working on the interpretation of such a property (and of its converse) for real data and trying to check its versatility for a larger class of physically relevant phenomena. We are also studying the geometrical characterization of signals in the FS plane, in terms of information distances and projections.

Acknowledgement

The authors wish to thank Dr. Bryan Griffiths who helped correcting the manuscript.

References

[1] B.R. Frieden, Phys. Rev. A 41 (1990) 4265.
[2] M.T. Martin, F. Pennini, A. Plastino, Phys. Lett. A 256 (1999) 173.
[3] A.R. Plastino, M. Casas, A. Plastino, Phys. Lett. A 246 (1998) 498.
[4] A. Dembo, T.M. Cover, J.A. Thomas, IEEE Trans. Inform. Theory 37 (1991) 1501.
[5] A. Kagan, Z. Landsman, Statist. Probab. Lett. 42 (1999) 7.
[6] J.A. Costa, A.O. Hero, C. Vignat, in: IEEE International Symposium on Information Theory, ISIT, Lausanne, 2002, p. 263.
[7] A.M.C. de Souza, C. Tsallis, Physica A 236 (1997) 52.
[8] K. Zografos, J. Multivariate Anal. 71 (1999) 67.
[9] R.A. Fisher, Proc. Cambridge Philos. Soc. 22 (1925) 700.
[10] C. Vignat, J.-F. Bercher, ESIEE Internal report number SIGTEL-SI 2003-2, 2003.