Available online at www.sciencedirect.com
ScienceDirect
Physica A 380 (2007) 1–18
PHYSICA A
www.elsevier.com/locate/physa

Autocorrelation versus entropy-based autoinformation for measuring dependence in random signal

François Chapeau-Blondeau*

Laboratoire d'Ingénierie des Systèmes Automatisés (LISA), Université d'Angers, 62 Avenue Notre Dame du Lac, 49000 Angers, France

Received 28 November 2006; received in revised form 10 January 2007
Available online 6 March 2007

Abstract

This paper studies in conjunction correlation and entropy-based information measures for the characterization of statistical dependence in a random signal. Several simple reference models of random signal are presented, for which both the autocorrelation and autoinformation functions are calculated explicitly in analytical form. Conditions are investigated where a general relation is shown to exist between these two functions in asymptotic regime, and which especially apply to stationary signals. Another, recent, model of nonstationary random signal with long-range dependence, is also presented and analyzed. For this model, the autocorrelation and autoinformation functions are calculated and compared for the first time, and exhibit more complex asymptotic behavior. This paper is intended to provide essentially theoretical models and results useful for better appreciation of the potentialities of the autoinformation function, in complement to the more common autocorrelation function, for the study of structures, informational contents and properties of complex random signals.

Keywords: Correlation; Mutual information; Random signal analysis; Long-range dependence

1. Introduction

Measuring statistical dependence along a random signal is an important task, which is specially relevant in the study of many complex processes of current interest. It can serve various general purposes such as characterizing the structures and informational contents of a signal, separating informative versus noisy contributions in its fluctuations, identifying the dynamics of their underlying generating processes. Statistical dependence in a random signal is commonly characterized with autocorrelation measures. Such measures have interesting theoretical properties and are readily implementable in practice. They are employed on many occasions to characterize complex signals and processes [1–5]. Correlation measures have also some inherent limitations as they quantify statistical dependence only partially. Information–theoretic measures like the mutual information or autoinformation constitute a useful complementary approach [6–12]. Especially, these measures offer a complete quantification of the statistical dependence. The counterpart is that these

*Fax: +33 2 41 73 53 52.
E-mail address: chapeau@univ-angers.fr.

0378-4371/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2007.02.077

information–theoretic measures, on a general basis, are usually more difficult to implement practically and to handle theoretically. For example, while explicit theoretical expressions of the autocorrelation function are given in many places for many reference random signals, there are scant few places where comparable theoretical expressions are explicitly provided for the autoinformation function for instance.

Based on their complementary theoretical and practical properties, both types of dependence measures are simultaneously useful to contribute in the investigation of complex random signals. In the present paper, our contribution will be to collect and complement essentially theoretical properties of the autoinformation function, and to connect them with more standard results for the autocorrelation function. We will expose several simple models of random signal, for which both the autocorrelation function and the autoinformation function can be calculated explicitly in analytical form. These models can especially serve as reference models to test various estimation methods for these functions on experimental signals. In addition, on these simple reference models, we will exhibit a relation that exists between the autocorrelation and autoinformation functions in the asymptotic regime of large separation times when both functions are small. A theoretical argument will be given which demonstrates that this asymptotic relation also holds in a general class of stationary signals. We will also present a recent model of nonstationary random signal with long-range dependence, for which the autocorrelation and autoinformation functions are calculated analytically and compared for the first time, and which shows more complex asymptotic behaviors for these two functions. The global objective of this paper is to provide theoretical models and results useful to more precisely appreciate the potentialities of the autoinformation function, in complement to the autocorrelation function, to contribute in the study of structures, informational contents and properties of complex random signals.

The works in Refs. [13,14] also study the relation between correlation function and mutual information for random signals. These Refs. [13,14] concentrate essentially on symbolic sequences, and essentially binary and ternary sequences in Ref. [13], while here we also address analog signals. We also offer here a broader view with reference models not given in Refs. [13,14]. Also [13,14] assume stationarity of the signals, while here nonstationary signals are also considered, especially with new results concerning a recent model for random signal with long-range dependence.

2. Measuring statistical dependence

2.1. Autocorrelation function

The statistical dependence along a random signal $x(t)$ is commonly characterized by means of the autocorrelation function $R_{xx}(t,\tau)$ defined as the expectation

$$R_{xx}(t,\tau) = \mathrm{E}[x(t)x(t+\tau)], \tag{1}$$

or also by means of the autocovariance function $C_{xx}(t,\tau)$ defined as the autocorrelation of the centered signal $x(t) - \mathrm{E}[x(t)]$, i.e.,

$$C_{xx}(t,\tau) = \mathrm{E}(\{x(t) - \mathrm{E}[x(t)]\}\{x(t+\tau) - \mathrm{E}[x(t+\tau)]\}) = R_{xx}(t,\tau) - \mathrm{E}[x(t)]\mathrm{E}[x(t+\tau)]. \tag{2}$$

At any t and τ, from the Schwarz inequality, the autocovariance verifies $|C_{xx}(t,\tau)| \leqslant \mathrm{std}[x(t)]\ \mathrm{std}[x(t+\tau)]$, with $\mathrm{std}(\cdot)$ the standard deviation. A normalized version of the autocovariance is the autocorrelation coefficient

$$r_{xx}(t,\tau) = \frac{C_{xx}(t,\tau)}{\mathrm{std}[x(t)]\ \mathrm{std}[x(t+\tau)]}, \tag{3}$$

which takes its values in $[-1,1]$ for all t and τ. Especially, at any time t, the autocorrelation coefficient $r_{xx}(t,\tau)$ is at its maximum of 1 in $\tau = 0$. Also, for many random signals, there is usually when $\tau \to \pm\infty$, a loss of causal connection between $x(t)$ and $x(t+\tau)$ entailing statistical independence, and in this circumstance both $C_{xx}(t,\tau)$ and $r_{xx}(t,\tau)$ go to zero at this limit $\tau \to \pm\infty$. The evolution of $r_{xx}(t,\tau)$, between its maximum $r_{xx}(t,\tau=0) = 1$ and its asymptotics $r_{xx}(t,\tau=\pm\infty) = 0$, serves as a convenient measure of statistical similarity or dependence between $x(t)$ and $x(t+\tau)$. Also, the way $C_{xx}(t,\tau)$ or $r_{xx}(t,\tau)$ return to zero with increasing τ offers an interesting characterization of the statistical fluctuations in $x(t)$. An exponential decay of $C_{xx}(t,\tau)$ as $\exp(-\beta\tau)$, often observed in practice, identifies short-range dependence, over a characteristic time scale $1/\beta$. But slowler, power-law, decays as $\tau^{-\beta}$, with $\beta > 0$, are also frequently observed on various complex processes, and identify long-range dependence characterized by the exponent β.

A useful property of the autocorrelation function is that its behavior is completely predictable in the framework of linear systems theory [15, p. 308]: if a random signal $x(t)$ with autocorrelation $R_{xx}(t,\tau)$ is applied on the input of a time-invariant linear system with impulse response $h(t)$, then the random signal $y(t)$ at the output has an autocorrelation function $R_{yy}(t,\tau)$ given by

$$R_{yy}(t,\tau) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} R_{xx}(t-t', \tau + t' - t'')h(t')h(t'')\,\mathrm{d}t'\,\mathrm{d}t'', \tag{4}$$

which in the case of $x(t)$ stationary reduces to the convolution $R_{yy}(\tau) = R_{xx}(\tau) * R_{hh}(\tau)$, with $R_{hh}(\tau) = \int_{-\infty}^{\infty} h(t)h(t+\tau)\,\mathrm{d}t$. Similar relations exist for the autocovariance, for instance $C_{yy}(\tau) = C_{xx}(\tau) * R_{hh}(\tau)$. These general properties offer large possibilities to shape the autocorrelation of a random signal by means of the selection of a linear filter. For instance, a stationary white noise $x(t)$ at the input with delta correlation $R_{xx}(\tau) = \delta(\tau)$ will be transformed by the filter in a stationary random signal $y(t)$ with autocorrelation $R_{yy}(\tau) = R_{hh}(\tau)$ definable by the filter via $h(t)$ in a very flexible way. However, a linear system, whatever its complexity or high order, starting with an input signal $x(t)$ with no pre-existing correlation, i.e., a delta-correlated $R_{xx}(\tau) = \delta(\tau)$, is unable to produce an output with long-range correlation in $\tau^{-\beta}$, and will unavoidably produce short-range correlation in $\exp(-\beta\tau)$ at the output. The reason is as follows.

A linear system, in the time domain, can be characterized by an input–output linear differential equation, or equivalently by an impulse response $h(t)$. It can be announced that in general, at large t, the impulse response returns to zero as $h(t) \sim \exp(-\beta t)$. Laplace transform of the differential equation yields a transmittance $H(s)$ which comes as a rational function (ratio of two polynomials) in the Laplace variable s. Inverse Laplace transform of $H(s)$ yields $h(t)$. This occurs usually through an intermediate step which is a partial fraction expansion of $H(s)$ as the sum of terms in s of the form $1/(s-p_i)$, whose inverse Laplace transform induces temporal terms of the form $\exp(p_i t)$ in $h(t)$. At large t thus $h(t)$ is ruled by the dominant pole p_i that has the real part (negative for stability) with smaller absolute value, which is the coefficient $-\beta$ explaining the asymptotic form $h(t) \sim \exp(-\beta t)$, all the other temporal contributions in $h(t)$ having returned to zero faster. This asymptotic behavior of $h(t)$ carries over to the autocorrelation $R_{hh}(\tau)$ which is dominated asymptotically by an exponential decay[1] with similar form $\sim \exp(-\beta\tau)$. When no correlation pre-exists in the input, it is this exponential decay which remains, asymptotically, in the autocorrelation of the random output signal of any linear system. In particular, this leaves open the question of the origin of the long-range dependence frequently observed on signals from various complex processes, knowing that it cannot come from linear filtering of any order operating on delta-correlated or short-range dependent fluctuations. This also establishes the study of the autocorrelation of observed signals, and specially its short- or long-range behaviors, as a means providing indication on the possible underlying generating mechanisms, either linear or non.

2.2. *Autoinformation function*

Autocorrelation measures have thus interesting properties, but they also suffer from some limitations. Especially, $C_{xx}(t,\tau)$ as a measure of statistical dependence suffers from the limitation that, at given τ, meanwhile independence between $x(t)$ and $x(t+\tau)$ implies $C_{xx}(t,\tau) = 0$, on the converse $C_{xx}(t,\tau) = 0$ (decorrelation) does not imply independence between $x(t)$ and $x(t+\tau)$. To circumvent this limitation, another measure of statistical dependence between $x(t)$ and $x(t+\tau)$ is provided by a form of the mutual information $I_{xx}(t,\tau)$ between $x(t)$ and $x(t+\tau)$, or autoinformation function, which is definable from, when they exist, the one-time entropies as

$$I_{xx}(t,\tau) = H[x(t+\tau)] - H[x(t+\tau)|x(t)]. \tag{5}$$

[1] In case of a multiple pole, the associated temporal term in the impulse response and subsequently in the autocorrelation, rather has the form $\sim\tau^{n-1}\exp(-\beta\tau)$ with n related to the multiplicity of the pole, but this decay also is considered as short-range dependence.

In Eq. (5), the quantity $H[x(t+\tau)]$ is the marginal entropy of $x(t+\tau)$, while $H[x(t+\tau)|x(t)]$ is the conditional entropy of $x(t+\tau)$ given $x(t)$. The entropies in Eq. (5) follow the standard definitions of entropies for random variables, either discrete or continuous. For instance, when $x(t)$ is a continuous variable, characterized by the marginal probability density $p_x(u,t)$, one has

$$H[x(t)] = -\int p_x(u,t)\,\log p_x(u,t)\,\mathrm{d}u. \tag{6}$$

With the joint probability density $p_{xx}(u_1,u_2;t,\tau)$ for $x(t)$ and $x(t+\tau)$, one has

$$H[x(t+\tau)|x(t)] = -\int\!\!\int p_{xx}(u_1,u_2;t,\tau)\,\log\frac{p_{xx}(u_1,u_2;t,\tau)}{p_x(u_1,t)}\,\mathrm{d}u_1\,\mathrm{d}u_2. \tag{7}$$

Since the one-time entropy above differs from the entropy of a signal defined from multidimensional multi-time probabilities, we choose to name $I_{xx}(t,\tau)$ of Eq. (5) autoinformation rather than mutual information; another reason for this name is the strong parallelism between $I_{xx}(t,\tau)$ and the autocorrelation function. An interesting property of the autoinformation function is that it always verifies $I_{xx}(t,\tau)\geqslant 0$, with equality if and only if $x(t)$ and $x(t+\tau)$ are independent. In this respect, while the autocorrelation only measures partially the statistical dependence, the autoinformation offers a complete quantification of the two-time dependence in a random signal $x(t)$. The counterpart is that there is no general theoretical property predicting the behavior of the autoinformation that would parallel Eq. (4) achieved by linear systems theory for the autocorrelation. As a consequence, very scarce theoretical results are available that gather expressions for the autoinformation function for some reference random signals. It also results that the autocorrelation and autoinformation functions remain complementary approaches to characterize statistical dependence in a random signal. In the sequel, we will present several models of random signal for which both the autocorrelation and autoinformation functions can be calculated explicitly in analytical form, and we will study the relation between these two characteristics.

3. Some random signal models

3.1. Stationary Gaussian noise

When the random signal $x(t)$ is a stationary Gaussian noise, endowed with mean $\mathrm{E}[x(t)] = m$, standard deviation $\mathrm{std}[x(t)] = \sigma$ and autocorrelation coefficient $r_{xx}(\tau)$, then the joint probability density for $x(t)$ and $x(t+\tau)$ is [15, p. 127]

$$p_{xx}(u_1,u_2;\tau) = \frac{1}{2\pi\sigma^2\sqrt{1-r_{xx}^2(\tau)}}\exp\left[-\frac{(u_1-m)^2 - 2(u_1-m)(u_2-m)r_{xx}(\tau) + (u_2-m)^2}{2\sigma^2[1-r_{xx}^2(\tau)]}\right], \tag{8}$$

independent of t.

From Eq. (8), integrals (6)–(7) can be carried out to yield the entropies [15, p. 562]:

$$H[x(t)] = \ln(\sqrt{2\pi e}\,\sigma) \tag{9}$$

and

$$H[x(t+\tau)|x(t)] = \ln\left[\sqrt{2\pi e}\,\sigma\sqrt{1-r_{xx}^2(\tau)}\right]. \tag{10}$$

The autoinformation of Eq. (5) follows as:

$$I_{xx}(\tau) = -\tfrac{1}{2}\ln[1-r_{xx}^2(\tau)]. \tag{11}$$

It results from Eq. (11) that in this case of the stationary Gaussian noise, the autocorrelation $r_{xx}(\tau)$ uniquely determines the autoinformation $I_{xx}(\tau)$; and $I_{xx}(\tau) = 0$ if and only if $r_{xx}(\tau) = 0$, i.e., independence equivalues decorrelation. In addition, when $r_{xx}(\tau)\to 0$ at the limit $\tau\to\pm\infty$, then from Eq. (11) one gets

$$I_{xx}(\tau) \underset{\tau\to\pm\infty}{\longrightarrow} \tfrac{1}{2}r_{xx}^2(\tau). \tag{12}$$

3.2. Memoryless transform of Gaussian noise

It is possible to envisage the transformation of the Gaussian noise $x(t)$ of Section 3.1 by a memoryless one-point function $g(\cdot)$ to produce the random signal $y(t) = g[x(t)]$. In this case, a form of Price's theorem [15, p. 161] provides a connection between the autocorrelation functions through the differential equation

$$\frac{\mathrm{d}^n R_{yy}(\tau)}{\mathrm{d}C_{xx}(\tau)^n} = \mathrm{E}\left\{\frac{\mathrm{d}^n g[x(t)]}{\mathrm{d}x^n}\frac{\mathrm{d}^n g[x(t+\tau)]}{\mathrm{d}x^n}\right\} \tag{13}$$

for any positive integer n.

An interesting situation is when $g(\cdot)$ is an invertible function, because it leaves unchanged the autoinformation function, i.e., $I_{xx}(\tau) = I_{yy}(\tau)$. Thus, invertible transformations $g(\cdot)$ in association with Eq. (13), allows one to generate random signals with control on both their autocorrelation and autoinformation functions, to some extent.

For illustration, we consider the case of a zero-mean Gaussian noise $x(t)$ through the cubic transformation $g[x(t)] = x^3(t) = y(t)$. Exploitation of Eq. (13) at $n = 2$ gives

$$\frac{\mathrm{d}^2 C_{yy}(\tau)}{\mathrm{d}C_{xx}(\tau)^2} = \mathrm{E}[6x(t)6x(t+\tau)] = 36C_{xx}(\tau), \tag{14}$$

since $y(t)$ is also zero-mean. Integration of Eq. (14) yields

$$C_{yy}(\tau) = 6C_{xx}^3(\tau) + 9\sigma^4 C_{xx}(\tau), \tag{15}$$

where the two integration constants have been determined with the conditions $C_{yy}(\tau \to \pm\infty) = 0$ and $C_{yy}(\tau = 0) = \mathrm{E}[y^2(t)] = \mathrm{E}[x^6(t)] = 15\sigma^6$. For $y(t)$, from Eq. (15) the autocorrelation coefficient $r_{yy}(\tau) = C_{yy}(\tau)/\mathrm{E}[y^2(t)]$ follows as:

$$r_{yy}(\tau) = \tfrac{6}{15}r^3_{xx}(\tau) + \tfrac{9}{15}r_{xx}(\tau), \tag{16}$$

while its autoinformation $I_{yy}(\tau)$ remains, from Eq. (11) and invertibility of $g(x) = x^3$,

$$I_{yy}(\tau) = I_{xx}(\tau) = -\tfrac{1}{2}\ln[1 - r_{xx}^2(\tau)]. \tag{17}$$

In the asymptotic regime $\tau \to \pm\infty$ where $r_{xx}(\tau)$ goes to zero, Eq. (16) gives $r_{yy}(\tau) \approx 9r_{xx}(\tau)/15$, and Eq. (17) gives $I_{yy}(\tau) \approx r_{xx}^2(\tau)/2$, leading to

$$I_{yy}(\tau) \underset{\tau\to\pm\infty}{\longrightarrow} \left(\frac{15}{9}\right)^2 \times \frac{1}{2}r_{yy}^2(\tau) \tag{18}$$

which bears some similarity with Eq. (12).

Also, it is possible to choose the transformation $g(\cdot)$ in order to select the marginal probability density $p_y(u)$ of the transformed signal $y(t)$. The original Gaussian noise $x(t)$ has a Gaussian marginal density $p_x(u)$ derivable from Eq. (8) and associated to the marginal cumulative distribution function

$$F_x(u) = \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{u-m}{\sqrt{2}\sigma}\right). \tag{19}$$

This $F_x(u)$ is an invertible function, and the transformed signal $F_x[x(t)]$ has a probability density uniform over $[0, 1]$. The targeted density $p_y(u)$ is associated to a cumulative distribution $F_y(u)$ which is invertible when $p_y(u)$ contains no point masses of discrete nonzero probabilities. The resulting inverse function $F_y^{-1}(u)$ is an invertible function, that when used to transform the uniform signal $F_x[x(t)]$ produces a random signal $F_y^{-1}\{F_x[x(t)]\}$ endowed with the probability density $p_y(u)$. The composite transformation $g(\cdot) = F_y^{-1}[F_x(\cdot)]$ results as an invertible function, which sends the Gaussian noise $x(t)$ into a random signal $y(t) = F_y^{-1}\{F_x[x(t)]\}$ with prescribed marginal probability density $p_y(u)$ and prescribed autoinformation function $I_{yy}(\tau) = I_{xx}(\tau)$ defined by $r_{xx}(\tau)$ of the original $x(t)$ via Eq. (11). A rather general and flexible technique results to construct an analog random signal with prescribed marginal probability density and prescribed autoinformation function: Starting with a white noise, a linear filter is used to produce a Gaussian signal $x(t)$ with an autocorrelation $r_{xx}(\tau)$ prescribed via Eq. (4) and an autoinformation $I_{xx}(\tau)$ resulting via Eq. (11), next on $x(t)$ the nonlinear memoryless invertible transformation $F_y^{-1}\{F_x[x(t)]\} = y(t)$ produces a random signal $y(t)$ with both prescribed marginal $p_y(u)$ and autoinformation $I_{yy}(\tau) = I_{xx}(\tau)$.

3.3. Binarized stationary Gaussian noise

We now consider the stationary random signal $y(t)$ obtained from a binary quantization of the stationary Gaussian noise $x(t)$ of Section 3.1, with the quantization threshold at the mean m (which is also the median), according to the noninvertible transformation

$$y(t) = \mathrm{sign}[x(t) - m] = \pm 1. \tag{20}$$

Therefore $y(t)$ of Eq. (20) at all t assumes discrete values restricted to ± 1. From the joint density of Eq. (8), one can deduce the four joint probabilities $P_{i,j} = \Pr\{y(t) = i, y(t+\tau) = j\}$ for $(i,j) \in \{-1, 1\}^2$. For instance, one has

$$P_{1,1} = \int_{u_1=m}^{+\infty} \int_{u_2=m}^{+\infty} p_{xx}(u_1, u_2; \tau)\, \mathrm{d}u_1\, \mathrm{d}u_2, \tag{21}$$

which amounts to

$$P_{1,1} = \frac{1}{4} + \frac{1}{2\pi} \arcsin[r_{xx}(\tau)] = P_{-1,-1}. \tag{22}$$

In the same way, one obtains

$$P_{-1,1} = P_{1,-1} = \frac{1}{4} - \frac{1}{2\pi} \arcsin[r_{xx}(\tau)]. \tag{23}$$

The marginal probabilities for $y(t)$ are $\Pr\{y(t) = 1\} = \Pr\{y(t) = -1\} = \frac{1}{2}$, whence the mean $\mathrm{E}[y(t)] = 0$ and standard deviation $\mathrm{std}[y(t)] = 1$.

The autocorrelation coefficient for $y(t)$ follows as:

$$r_{yy}(\tau) = P_{1,1} + P_{-1,-1} - P_{-1,1} - P_{1,-1}, \tag{24}$$

giving the relation

$$r_{yy}(\tau) = \frac{2}{\pi} \arcsin[r_{xx}(\tau)] \tag{25}$$

known as the arcsine law [15, p. 307].

The marginal entropy for $y(t)$ is

$$H[y(t)] = \ln(2)\,\mathrm{nat} = 1\,\mathrm{bit}. \tag{26}$$

Together from the marginals and the joint probabilities of Eqs. (22)–(23), one easily deduces the conditional probabilities, and then the conditional entropy follows as:

$$H[y(t+\tau)|y(t)] = \ln(2) - \tfrac{1}{2}[1 + r_{yy}(\tau)] \ln[1 + r_{yy}(\tau)] - \tfrac{1}{2}[1 - r_{yy}(\tau)] \ln[1 - r_{yy}(\tau)]. \tag{27}$$

The autoinformation of Eq. (5) results as

$$I_{yy}(\tau) = \tfrac{1}{2}[1 + r_{yy}(\tau)] \ln[1 + r_{yy}(\tau)] + \tfrac{1}{2}[1 - r_{yy}(\tau)] \ln[1 - r_{yy}(\tau)]. \tag{28}$$

It follows from Eq. (28) that in this case also of the binarized stationary Gaussian noise, the autocorrelation $r_{yy}(\tau)$ uniquely determines the autoinformation $I_{yy}(\tau)$. From Eq. (28), $r_{yy}(\tau) = 0$ implies $I_{yy}(\tau) = 0$; conversely $I_{yy}(\tau) = 0$ implies independence which implies $r_{yy}(\tau) = 0$; therefore for this $y(t)$ also independence equivalues decorrelation. In addition, when $r_{yy}(\tau) \to 0$ at the limit $\tau \to \pm\infty$, then from Eq. (28) with $\ln(u) \approx u - u^2/2$ one obtains again

$$I_{yy}(\tau) \underset{\tau \to \pm\infty}{\longrightarrow} \tfrac{1}{2} r_{yy}^2(\tau). \tag{29}$$

3.4. Random telegraph signal

We now consider the so-called random telegraph signal $x(t) = \pm 1$, which alternates between the two levels $+1$ and -1 at random times t_i distributed according to a Poisson process of parameter λ. The number of points t_i in any interval $[t, t+\tau]$ thus equals n with probability $p(n) = \mathrm{e}^{-\lambda\tau}(\lambda\tau)^n/n!$. We assume the initial condition $x(t=0)=1$. At any $t>0$, $x(t)=1$ if there is an even number of points t_i in interval $[0, t]$, so one has the probability

$$\Pr\{x(t)=1\} = \sum_{n=0}^{+\infty} p(2n) = \mathrm{e}^{-\lambda t}\sum_{n=0}^{+\infty}\frac{(\lambda t)^{2n}}{(2n)!} = \mathrm{e}^{-\lambda t}\cosh(\lambda t) = \frac{1+\mathrm{e}^{-2\lambda t}}{2}. \tag{30}$$

Also $x(t)=-1$ for an odd number of points t_i in $[0, t]$, yielding the probability

$$\Pr\{x(t)=-1\} = \mathrm{e}^{-\lambda t}\sinh(\lambda t) = \frac{1-\mathrm{e}^{-2\lambda t}}{2}. \tag{31}$$

One gets the expectation $\mathrm{E}[x(t)] = \mathrm{e}^{-2\lambda t}$, and $\mathrm{E}[x^2(t)] = 1$.

In a similar way, for any $\tau \geqslant 0$, one has the conditional probabilities

$$\Pr\{x(t+\tau)=1 | x(t)=1\} = \Pr\{x(t+\tau)=-1 | x(t)=-1\} = \mathrm{e}^{-\lambda\tau}\cosh(\lambda\tau), \tag{32}$$

and

$$\Pr\{x(t+\tau)=1 | x(t)=-1\} = \Pr\{x(t+\tau)=-1 | x(t)=1\} = \mathrm{e}^{-\lambda\tau}\sinh(\lambda\tau). \tag{33}$$

Gathering the above results, one obtains the autocorrelation function

$$R_{xx}(t,\tau) = \mathrm{E}[x(t)x(t+\tau)] = \mathrm{e}^{-2\lambda\tau} \tag{34}$$

as also derived in Ref. [15, p. 292], and the autocorrelation coefficient $r_{xx}(t,\tau)$ could also be easily expressed.

From Eqs. (30) to (31), the marginal entropy of $x(t)$ is

$$H[x(t)] = -\mathrm{e}^{-\lambda t}\cosh(\lambda t)\ln[\mathrm{e}^{-\lambda t}\cosh(\lambda t)] - \mathrm{e}^{-\lambda t}\sinh(\lambda t)\ln[\mathrm{e}^{-\lambda t}\sinh(\lambda t)], \tag{35}$$

and via Eqs. (32)–(33), the conditional entropy $H[x(t+\tau)|x(t)]$ could also be easily expressed, leading to a rather bulky expression that we omit.

It is clear, for instance from $\mathrm{E}[x(t)] = \mathrm{e}^{-2\lambda t}$, that $x(t)$ is a nonstationary signal, due to the influence of the initial condition at $t=0$. However, sufficiently far from the origin $t=0$, i.e., at any $t \gg 1/\lambda$, one has access to useful approximations. Especially, one gets for the autocorrelation coefficient

$$r_{xx}(t,\tau) \approx \mathrm{e}^{-2\lambda\tau} - \mathrm{e}^{-4\lambda t}\mathrm{e}^{-2\lambda\tau} = \mathrm{e}^{-2\lambda\tau}(1-\mathrm{e}^{-4\lambda t}) \tag{36}$$

plus higher-order terms in $\mathrm{e}^{-\lambda t}$ that are negligible. Also, one finds for the marginal entropy

$$H[x(t)] \approx \ln(2) - \tfrac{1}{2}\mathrm{e}^{-4\lambda t}, \tag{37}$$

plus higher-order negligible terms in $\mathrm{e}^{-\lambda t}$. Now separately, at any t but for large delays $\tau \gg 1/\lambda$, one has for the conditional entropy

$$H[x(t+\tau)|x(t)] \approx \ln(2) - \tfrac{1}{2}\mathrm{e}^{-4\lambda\tau}, \tag{38}$$

plus higher-order negligible terms in $\mathrm{e}^{-\lambda\tau}$. Now in conjunction at large $t \gg 1/\lambda$ and $\tau \gg 1/\lambda$, from Eqs. (37) to (38) one finds for the autoinformation

$$I_{xx}(t,\tau) \approx \tfrac{1}{2}\mathrm{e}^{-4\lambda\tau} - \tfrac{1}{2}\mathrm{e}^{-4\lambda t}\mathrm{e}^{-4\lambda\tau} = \tfrac{1}{2}\mathrm{e}^{-4\lambda\tau}(1-\mathrm{e}^{-4\lambda t}). \tag{39}$$

Eqs. (36) and (39) illustrate that for the present nonstationary signal $x(t)$, there is in general no one-to-one relationship between, alone, the autocorrelation $r_{xx}(t,\tau)$ and the autoinformation $I_{xx}(t,\tau)$. One can however write $I_{xx}(t,\tau) \approx r_{xx}(t,\tau)\mathrm{e}^{-2\lambda\tau}/2$ showing that in the asymptotic regime $I_{xx}(t,\tau)$ decays faster with τ than $r_{xx}(t,\tau)$. Also, far from the origin $t=0$ and at large delays τ, the dominant asymptotic behaviors are, from Eq. (36), $r_{xx}(t,\tau) \approx \mathrm{e}^{-2\lambda\tau}$, and from Eq. (39), $I_{xx}(t,\tau) = \mathrm{e}^{-4\lambda\tau}/2$. At this limit, one can write

$$I_{xx}(t,\tau) \underset{t,\tau\to\infty}{\longrightarrow} \tfrac{1}{2} r_{xx}^2(t,\tau), \tag{40}$$

which matches the asymptotic relation of Eqs. (12) and (29) found for stationary signals.

4. Asymptotic behavior

After the reference models of Section 3, where explicit expressions were available for both the autocorrelation and autoinformation functions, we shall now examine, in an asymptotic regime, conditions where a general relation can exist between these two functions.

4.1. Arbitrary signal

We assume a random signal $x(t)$ which takes its values in a discrete set of N distinct values x_i for $i = 1$ to N. The random signals of Sections 3.3 and 3.4 are members of this category. However, the argument that will follow could easily be transposed to the case of analog signals $x(t)$ as in Sections 3.1 and 3.2, with discrete probabilities replaced by densities. We introduce the marginal probabilities $P_i(t) = \Pr\{x(t) = x_i\}$ for $i \in \{1,\ldots,N\}$, and the joint probabilities $P_{ij}(t,\tau) = \Pr\{x(t) = x_i, x(t+\tau) = x_j\}$ for $(i,j) \in \{1,\ldots,N\}^2$. When $x(t)$ and $x(t+\tau)$ are independent, one has the factorization $P_{ij}(t,\tau) = P_i(t)P_j(t+\tau)$ for any $(i,j) \in \{1,\ldots,N\}^2$.

We define, for any $(i,j) \in \{1,\ldots,N\}^2$, the quantity

$$\delta_{ij}(t,\tau) = \frac{P_{ij}(t,\tau) - P_i(t)P_j(t+\tau)}{P_i(t)P_j(t+\tau)}, \tag{41}$$

inspired from Ref. [14], and interpreted as an index of departure from statistical independence, since $\delta_{ij}(t,\tau) = 0$ at independence, for all $(i,j) \in \{1,\ldots,N\}^2$. We denote a statistical average over the two sets of marginal probabilities by the notation $\langle \cdot\cdot \rangle = \sum_{i,j} \cdot\cdot P_i(t)P_j(t+\tau)$, to be distinguished from the standard statistical expectation $\mathrm{E}(\cdot\cdot) = \sum_{i,j} \cdot\cdot P_{ij}(t,\tau)$. Because of the normalization of the two sets of probabilities $P_i(t)$ and $P_{ij}(t,\tau)$, i.e., $\sum_i P_i = 1$ and $\sum_{ij} P_{ij} = 1$, one has

$$\langle \delta_{ij}(t,\tau) \rangle = 0 \tag{42}$$

for any t and τ.

The autocovariance function of $x(t)$ results as

$$C_{xx}(t,\tau) = \sum_{i,j} x_i x_j [P_{ij}(t,\tau) - P_i(t)P_j(t+\tau)] = \langle x_i x_j \delta_{ij}(t,\tau) \rangle. \tag{43}$$

The autoinformation of Eq. (5) in nats is also

$$I_{xx}(t,\tau) = \sum_{i,j} P_{ij}(t,\tau) \ln \frac{P_{ij}(t,\tau)}{P_i(t)P_j(t+\tau)}, \tag{44}$$

now expressable as

$$I_{xx}(t,\tau) = \langle [1 + \delta_{ij}(t,\tau)] \ln[1 + \delta_{ij}(t,\tau)] \rangle. \tag{45}$$

The forms of Eqs. (43) and 45 show that, in general, there is no direct relation connecting alone the autocovariance $C_{xx}(t,\tau)$ and the autoinformation $I_{xx}(t,\tau)$, even for stationary signals when the dependence in t of $\delta_{ij}(t,\tau)$ vanishes.

In the situation of large delay $\tau \to \pm\infty$, an asymptotic regime where $x(t)$ and $x(t+\tau)$ return to independence produces $\delta_{ij}(t,\tau) \to 0$, and Eq. (45) leads to

$$I_{xx}(t,\tau) \approx \tfrac{1}{2} \langle \delta_{ij}^2(t,\tau) \rangle. \tag{46}$$

In such a regime, there is no general relation either connecting alone $C_{xx}(t,\tau)$ of Eq. (43) and $I_{xx}(t,\tau)$ of Eq. (46).

4.2. Stationary signal

If we add the assumption of $x(t)$ stationary, one has the probabilities $P_i(t) = P_i$ independent of t, and $P_{ij}(t, \tau) = P_{ij}(\tau)$ with τ-dependence only and $P_{ij}(-\tau) = P_{ji}(\tau)$, for all $(i,j) \in \{1, \ldots, N\}^2$. Also, one can write for any $i \in \{1, \ldots, N\}$ and any τ,

$$P_i = \sum_{j=1}^{N} P_{ij}(\tau) = \sum_{j=1}^{N} P_{ji}(\tau), \tag{47}$$

expressing $2N$ constraints imposed by stationarity for the N^2 probabilities $P_{ij}(\tau)$. In terms of the $\delta_{ij}(\tau)$'s these constraints can be written

$$0 = \sum_{j=1}^{N} P_i P_j \delta_{ij}(\tau) = \sum_{j=1}^{N} P_i P_j \delta_{ji}(\tau). \tag{48}$$

In the general case where Eqs. (43) and (45) hold, these constraints of Eq. (48) due to stationarity, allow no direct relation either connecting $C_{xx}(t, \tau)$ and $I_{xx}(t, \tau)$ alone. Yet, if under the assumption of $x(t)$ stationary, we further impose the sufficient condition for the $\delta_{ij}(\tau)$'s that

$$\delta_{ij}(\tau) = \delta'_{ij} f(\tau) \quad \forall (i,j) \in \{1, \ldots, N\}^2, \tag{49}$$

with δ'_{ij} independent of τ. This amounts to requiring that, for every $(i,j) \in \{1, \ldots, N\}^2$, each $\delta_{ij}(\tau)$ bears a dependence in τ expressable by a common multiplicative factor $f(\tau)$ accompanied by a complementary term δ'_{ij} possibly function of i and j but independent of τ. This relation of Eq. (49) is verified by the stationary Gaussian noise of Section 3.1 and by its binarized version of Section 3.3. Asymptotically, at large t and τ, the condition of Eq. (49) is also verified by the random telegraph signal of Section 3.4.

Based on Eq. (49), one obtains for instance for the autocorrelation coefficient, from Eq. (43),

$$r_{xx}(\tau) = \frac{C_{xx}(\tau)}{\sigma^2} = \frac{f(\tau)}{\sigma^2} \langle x_i x_j \delta'_{ij} \rangle, \tag{50}$$

with the stationary standard deviation $\sigma = \mathrm{std}[x(t)]$, and for the autoinformation in the asymptotic regime of Eq. (46),

$$I_{xx}(\tau) \approx \tfrac{1}{2} \langle \delta'^2_{ij} \rangle f^2(\tau). \tag{51}$$

Combining Eqs. (50) and (51), one gets at large τ for stationary signals verifying Eq. (49),

$$I_{xx}(\tau) \approx A \tfrac{1}{2} r_{xx}^2(\tau), \tag{52}$$

with the prefactor

$$A = \frac{\langle \delta'^2_{ij} \rangle}{\left\langle \frac{x_i}{\sigma} \frac{x_j}{\sigma} \delta'_{ij} \right\rangle^2}. \tag{53}$$

The prefactor A of Eq. (53) is a nondimensional constant, independent of τ, since under stationarity the linear operator $\langle \cdot \cdot \rangle$ introduces no dependence in τ (since P_i and P_j are independent of τ). It turns out that $A = 1$ for the stationary signals of Sections 3.1 and 3.3, and also for the nonstationary signal of Section 3.4 in the asymptotic regime of large t and τ, yielding the uniform behavior of Eqs. (12), (29) and (40) which is a form of Eq. (52). Also, Eq. (18) is another form of Eq. (52), with $A = (\frac{15}{9})^2$.

Therefore, for stationary signals verifying Eq. (49), at large τ, Eq. (52) expresses that the autoinformation $I_{xx}(\tau)$ is proportional to the squared autocorrelation $r_{xx}^2(\tau)$. Both functions return to zero at large τ, but the autoinformation $I_{xx}(\tau)$, which measures the complete statistical dependence between $x(t)$ and $x(t + \tau)$, always returns to zero faster than the autocorrelation $r_{xx}(\tau)$ which measures only partial dependence. Nevertheless, the short-range versus long-range character of the dependence in $x(t)$ is assessed qualitatively in the same way, i.e., exponential decay of $r_{xx}(\tau)$ is associated to exponential decay of $I_{xx}(\tau)$, and power-law decay of $r_{xx}(\tau)$ is associated to power-law decay of $I_{xx}(\tau)$.

We found no alternative way for a more concrete characterization of the stationary signals verifying Eq. (49). Eq. (49) is equivalent to $P_{ij}(\tau) = P_i P_j + a_{ij} f(\tau)$ with τ-independent a_{ij}, for the departure of the τ-dependent joint probability $P_{ij}(\tau)$ from the product of the τ-independent marginals P_i, P_j. Yet, as indicated, simple useful models of random signals, as those reported in Section 3, verify this property.

Ref. [14] restricted to stationary symbolic sequences, also arrived at results similar to those of this Section 4.2. From relations similar to $C_{xx}(\tau) = \langle x_i x_j \delta_{ij}(\tau)\rangle$ for the stationary autocorrelation and to $I_{xx}(\tau) \approx \langle \delta_{ij}^2(\tau)\rangle/2$ for the asymptotic stationary autoinformation, [14] directly jumps to an asymptotic proportionality relation $I_{xx}(\tau) \propto C_{xx}^2(\tau)$. Here in addition we have a form for a proportionality coefficient A from Eq. (53).

4.3. Stationary binary signal

In the special case of a stationary signal $x(t)$ with only $N = 2$ states, the $2N = 4$ (unindependent) constraints of Eqs. (47)–(48) impose severe limitations on the degrees of freedom of the random signal contained in the $N^2 = 4$ probabilities $P_{ij}(\tau)$. In this circumstance, it necessarily results the following forms:

$$P_{11}(\tau) = P_1^2 + f(\tau), \tag{54}$$

$$P_{12}(\tau) = P_{21}(\tau) = (1 - P_1)P_1 - f(\tau), \tag{55}$$

$$P_{22}(\tau) = (1 - P_1)^2 + f(\tau) \tag{56}$$

Any such stationary binary signal will in fact be defined by (in addition to the states x_1 and x_2) one scalar P_1 and one even function $f(\tau)$. The binary signal of Section 3.3 is one member of this class. Eqs. (54)–(56) are equivalent to

$$\delta_{11}(\tau) = \frac{1}{P_1^2} f(\tau), \tag{57}$$

$$\delta_{12}(\tau) = \delta_{21}(\tau) = -\frac{1}{P_1 P_2} f(\tau), \tag{58}$$

$$\delta_{22}(\tau) = \frac{1}{P_2^2} f(\tau), \tag{59}$$

with $P_2 = 1 - P_1$, which obviously satisfy Eq. (49). The sufficient condition Eq. (49) is thus always satisfied by any stationary binary signal. Therefore, Eq. (52) is expected to hold for any stationary binary signal. Moreover, as we shall show next, the prefactor A in Eq. (52) is always 1 for a stationary binary signal.

From Eqs. (54) to (56) or Eqs. (57) to (59), one obtains from their definitions exact expressions for the mean $\mathrm{E}[x(t)] = x_1 P_1 + x_2 P_2$, the variance $\mathrm{std}^2[x(t)] = (x_1 - x_2)^2 P_1 P_2$, the autocovariance function

$$C_{xx}(\tau) = (x_1 - x_2)^2 f(\tau), \tag{60}$$

the autocorrelation coefficient

$$r_{xx}(\tau) = \frac{1}{P_1 P_2} f(\tau), \tag{61}$$

and the autoinformation function

$$\begin{aligned} I_{xx}(\tau) = & -P_1 \log(P_1) - P_2 \log(P_2) + 2[P_1 P_2 - f(\tau)] \log[P_1 P_2 - f(\tau)] \\ & + [P_1^2 + f(\tau)] \log[P_1^2 + f(\tau)] + [P_2^2 + f(\tau)] \log[P_2^2 + f(\tau)]. \end{aligned} \tag{62}$$

Since from Eq. (61) one has $f(\tau) = P_1 P_2 r_{xx}(\tau)$, it results from Eq. (62) that for a stationary binary signal the autocorrelation $r_{xx}(\tau)$ uniquely determines the autoinformation $I_{xx}(\tau)$. Moreover, when $P_1 = P_2 = \frac{1}{2}$, Eq. (62) reduces to the form of Eq. (28), showing that, not only the binarized stationary Gaussian noise of Section 3.3, but any stationary signal binarized at its median (to get $P_1 = P_2 = \frac{1}{2}$), will verify Eq. (28).

In the asymptotic regime where Eq. (51) holds, one obtains

$$I_{xx}(\tau) \approx \tfrac{1}{2}\left(\frac{1}{P_1 P_2}\right)^2 f^2(\tau), \tag{63}$$

Finally, as announced, for the stationary binary signal, one always observes the relation

$$I_{xx}(\tau) \approx \tfrac{1}{2} r_{xx}^2(\tau) \tag{64}$$

between the autocorrelation and autoinformation functions in the asymptotic regime of τ large.

For the special case of a stationary binary sequence with states $x_1 = 1$ and $x_2 = 0$, [13] arrived at results similar to those of this Section 4.3.

4.4. Summary of asymptotic behavior

In the general situation, because of Eqs. (43) and (45) no direct relation can be expected to hold between the autocorrelation and autoinformation functions, even for stationary signals. In the asymptotic situation of large delay $\tau \to \pm\infty$ when $x(t)$ and $x(t+\tau)$ return to independence, it sufficies that the signal verify the condition of Eq. (49), to allow relation (52) to hold between the autocorrelation and autoinformation functions.

A stationary binary signal always satisfies sufficient condition (49). Therefore in the asymptotic situation of large delay $\tau \to \pm\infty$, it verifies relation (52), and moreover with a prefactor A which is always 1.

We shall now examine another model, for a nonstationary random signal, which especially displays long-range dependence, but which does not satisfy sufficient condition (49).

5. A model of long-range dependent signal

We now consider a discrete-time model [16,17], with time step Δt, and $t = k\Delta t$ with k integer. For the discrete time $t > 0$, we introduce the system

$$U(t) = U(t - \Delta t) + u(t), \tag{65}$$

$$X(t) = \max[X(t - \Delta t), U(t)], \tag{66}$$

$$x(t) = X(t) - X(t - \Delta t), \tag{67}$$

with the initial condition $U(0) = X(0) = 0$. For all times $t = k\Delta t > 0$ in Eq. (65), the input quantities $u(t)$ are independent and identically distributed random variables. For simplicity of notation, we set for the sequel $\Delta t = 1$ making t an integer time.

The system of Eqs. (65)–(67) produces a random signal $x(t)$ which represents the increments of the running maximum $X(t)$ of the random walk $U(t)$ having independent increments $u(t)$. Concretely, $x(t)$ is formed by successions of intervals where $x(t) = 0$ interrupted by bursts where $x(t) > 0$, with these successions occurring in a self-similar fashion over all time scales. The resulting self-similar structure for $x(t)$ is visible in Fig. 1 on a realization; it will also be expressed by power-law forms for the autocorrelation and autoinformation functions of $x(t)$ as we shall see.

According to Eq. (66), $x(t) > 0$ at each time t where the walk $U(t)$ realizes a first passage. Also $x(t)$ possesses a renewal property, since according to Eq. (66), at each time t where $x(t) > 0$, one has $X(t) = U(t)$, and for the subsequent evolution of the increment x it is just as if the system had been reset to its initial condition $X = U = 0$.

We now consider the simple case of the binary input $u(t) = \pm 1$ equiprobably, for any integer $t > 0$. In this case, the values accessible to the increment $x(t)$ reduce to 0 or 1. For any integers $t > 0$ and $\tau > 0$, we introduce the marginal probabilities $P_i(t) = \Pr\{x(t) = i\}$ for $i \in \{0, 1\}$, and the joint probabilities $P_{ij}(t, \tau) = \Pr\{x(t) = i, x(t+\tau) = j\}$ for $(i,j) \in \{0, 1\}^2$. One has $P_{11}(t, \tau) = \Pr\{x(t+\tau) = 1 \mid x(t) = 1\} \times P_1(t)$. Because of the renewal property, $\Pr\{x(t+\tau) = 1 \mid x(t) = 1\} = \Phi(\tau)$, a function which only depends on τ, and verifies $\Phi(0) = 1$ and

$$\Phi(t) = P_1(t) = \mathrm{E}[x(t)] = \mathrm{E}[x^2(t)] \tag{68}$$

for all $t > 0$. One then deduces

$$P_0(t) = 1 - P_1(t) = 1 - \Phi(t), \tag{69}$$

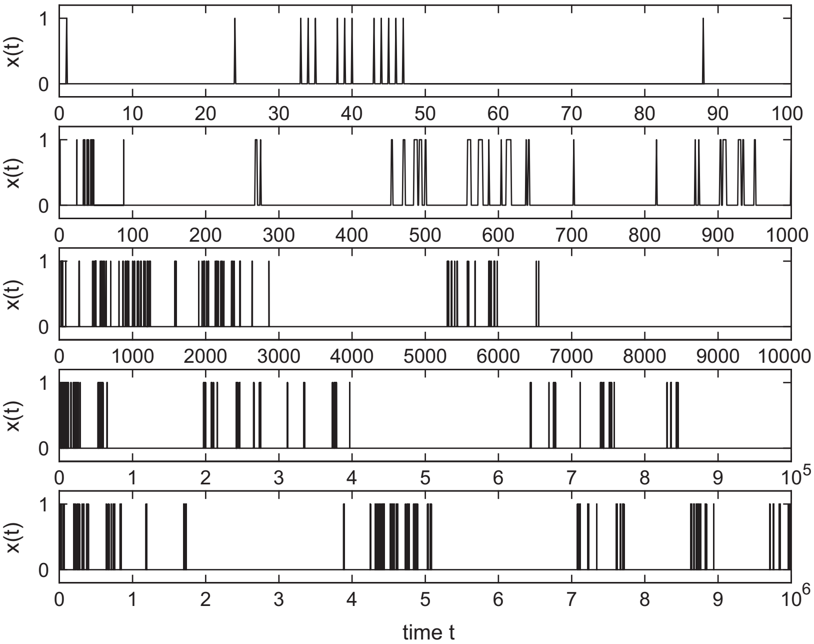

Fig. 1. One realization of the random signal $x(t)$ from Eqs. (65) to (67) when $u(t) = \pm 1$ equiprobably, represented over time intervals of increasing durations showing the self-similarity of $x(t)$.

and

$$P_{11}(t, \tau) = \Phi(\tau)\Phi(t). \tag{70}$$

One can right away deduce the autocorrelation function

$$R_{xx}(t, \tau) = \mathrm{E}[x(t)x(t + \tau)] = 1 \times 1 \times P_{11}(t, \tau) = \Phi(t)\Phi(\tau). \tag{71}$$

One also has $P_{10}(t, \tau) = \Pr\{x(t + \tau) = 0 | x(t) = 1\} \times P_1(t)$, yielding

$$P_{10}(t, \tau) = [1 - \Phi(\tau)]\Phi(t). \tag{72}$$

Since $P_1(t + \tau) = P_{11}(t, \tau) + P_{01}(t, \tau)$, it comes

$$P_{01}(t, \tau) = \Phi(t + \tau) - \Phi(t)\Phi(\tau), \tag{73}$$

and since $P_0(t + \tau) = P_{00}(t, \tau) + P_{10}(t, \tau)$, it comes

$$P_{00}(t, \tau) = 1 - \Phi(t + \tau) - [1 - \Phi(\tau)]\Phi(t). \tag{74}$$

The marginals $P_i(t)$ and joint probabilities $P_{ij}(t, \tau)$ for any $(i, j) \in \{0, 1\}^2$, are therefore all known by Eqs. (68)–(74) through function $\Phi(\cdot)$. This function $\Phi(\cdot)$ can be expressed, according to the properties of the random walk $U(t)$ with increments $u(t) = \pm 1$, as

$$\Phi(t) = \sum_{n=1}^{t} \varphi(n, t) \tag{75}$$

for any integer $t \geqslant 1$, where $\varphi(n, t)$ is the probability of a first passage in $U = n$ at step t of the random walk U started from $U = 0$ at step 0. This is because $x(t) = 1$ is equivalent, according to Eq. (66), to a first passage of the walk U at step t.

According to Ref. [18, p. 89, Eq. (7.5)], one has a $\varphi(n, t) = 0$ for n and t with opposite parity, and for n and t with same parity

$$\varphi(n, t) = \frac{2^{-t}}{t} n \,\mathrm{bino}[t, \ (t + n)/2], \tag{76}$$

where $\mathrm{bino}(\cdot, \cdot)$ is the standard binomial coefficient.

The sum of Eq. (75) can be explicitly evaluated: for $t \geqslant 2$ even, one has

$$\Phi(t) = \frac{2^{-t}}{t} t! \frac{1}{\left(\frac{t-2}{2}\right)! \left(\frac{t}{2}\right)!}, \tag{77}$$

and for $t > 2$ odd, one has

$$\Phi(t) = \frac{2^{-t}}{t} t! \frac{1}{\left(\frac{t-1}{2}\right)! \left(\frac{t-1}{2}\right)!}. \tag{78}$$

Eqs. (77)–(78) allow exact evaluations of both the autocorrelation and autoinformation functions of $x(t)$ for any $t > 0$ and $\tau > 0$, which however come as rather complicated expressions. We shall be interested now in the special case of the asymptotic regime where $1 \ll t \ll \tau$. In this regime, from Eqs. (77)–(78) one finds at leading orders in t and τ,

$$\Phi(t) = \mathrm{E}[x(t)] \approx at^{-1/2}, \tag{79}$$

with the constant $a = 1/\sqrt{2\pi}$, and the autocorrelation function

$$R_{xx}(t,\tau) = \Phi(t)\Phi(\tau) \approx a^2 t^{-1/2} \tau^{-1/2}. \tag{80}$$

Eqs. (79)–(80) establish $x(t)$ as a nonstationary long-range dependent random signal with power-law autocorrelation. The autoinformation $I_{xx}(t,\tau)$ will depend separately on $\Phi(t)$, $\Phi(\tau)$ and $\Phi(t+\tau)$, just as the joint probabilities $P_{ij}(t,\tau)$ of Eqs. (70)–(74) do. Since $R_{xx}(t,\tau)$ from Eq. (71) only depends on the product $\Phi(t)\Phi(\tau)$, no general relation is expected to hold between $I_{xx}(t,\tau)$ and $R_{xx}(t,\tau)$ alone.

Next, from the definition of Eq. (41), one finds at leading orders in t and τ,

$$\delta_{11}(t,\tau) \approx \frac{1}{2} t\tau^{-1}, \tag{81}$$

$$\delta_{10}(t,\tau) \approx -\frac{a}{2} t\tau^{-3/2}, \tag{82}$$

$$\delta_{01}(t,\tau) \approx -\frac{a}{2} t^{1/2}\tau^{-1}, \tag{83}$$

$$\delta_{00}(t,\tau) \approx \frac{a^2}{2} t^{1/2}\tau^{-3/2}. \tag{84}$$

Clearly, the above coefficients $\delta_{ij}(t,\tau)$ do not verify the sufficient condition of Eq. (49) which would guarantee the asymptotic relation of Eq. (52) to hold between autocorrelation $r_{xx}(\tau)$ and autoinformation $I_{xx}(\tau)$.

However, in the asymptotic regime, one also finds at leading orders in t and τ, from their definitions,

$$C_{xx}(t,\tau) \approx \frac{a^2}{2} t^{1/2}\tau^{-3/2}, \tag{85}$$

$$r_{xx}(t,\tau) \approx \frac{a}{2} t^{3/4}\tau^{-5/4}, \tag{86}$$

and

$$I_{xx}(t,\tau) \approx \frac{a^2}{8} t^{3/2}\tau^{-5/2}. \tag{87}$$

Finally, in this asymptotic regime, one observes again the same relation

$$I_{xx}(\tau) \approx \tfrac{1}{2} r_{xx}^2(\tau) \tag{88}$$

between the autocorrelation and autoinformation functions.

The theoretical expressions of Eqs. (86)–(88) are verified in Fig. 2 through numerical simulation of the self-similar random signal $x(t)$ of Fig. 1.

The results of Fig. 2 confirm the long-range dependence in the signal $x(t)$ with power-law evolutions for both the autocorrelation and autoinformation functions. Moreover, the simple asymptotic relation of Eq. (88) between the autocorrelation and autoinformation functions is obtained for this type of nonstationary long-range dependent random signal.

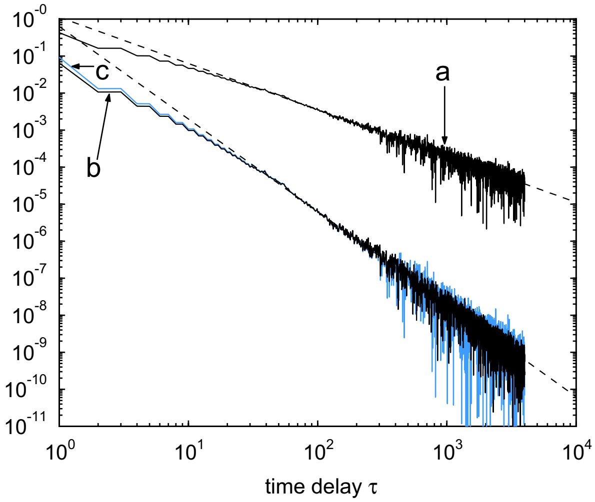

Fig. 2. For the self-similar random signal $x(t)$ of Fig. 1, as a function of the time delay τ at $t = 10$: (a) autocorrelation function $r_{xx}(t, \tau)$ numerically evaluated, and its asymptotic model of Eq. (86) (dashed line); (b) autoinformation function $I_{xx}(t, \tau)$ numerically evaluated, and its asymptotic model of Eq. (87) (dashed line); (c) function $\frac{1}{2}r_{xx}^2(t, \tau)$ matching $I_{xx}(t, \tau)$ as predicted by Eq. (88).

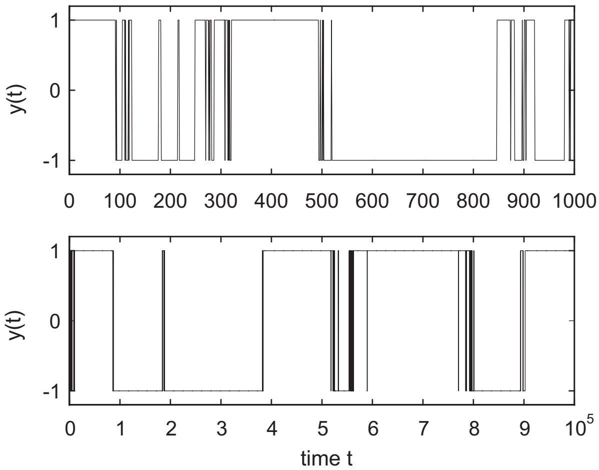

Fig. 3. One realization of the two-state random signal $y(t)$ represented over two time intervals of increasing durations showing the self-similarity of $y(t)$.

The signal $x(t)$ of Figs. 1 and 2 results from a very simple generation algorithm, based on the first-order recurrence of Eqs. (65)–(67). Its numerical implementation is straightforward and provides a very simple approach for producing a random signal with long-range dependence persisting over time duration of arbitrary length, as long as the process is left to run. To obtain more variety in the appearance of the random signal while preserving the long-range dependence and simplicity of generation, its is possible to use the process of Eqs. (65)–(67) to drive a secondary random signal, in various ways.

For instance, we used the random signal $x(t)$ of Fig. 1 to drive a binary random signal $y(t) = \pm 1$, with $y(t)$ flipping its state between -1 and $+1$ at each time t where $x(t) > 0$. A realization of this $y(t)$ is plotted in Fig. 3, over two time intervals of increasing durations, which provides visual appreciation of the self-similar character of $y(t)$, inherited from the self-similarity of the underlying process $x(t)$ from Eqs. (65) to (67).

Fig. 4 presents for $y(t)$ the numerical evaluation of autocorrelation function $r_{yy}(t, \tau)$ and autoinformation function $I_{yy}(t, \tau)$. The theoretical statistical properties of $y(t)$ of Fig. 3 are more intricate than those of $x(t)$ of

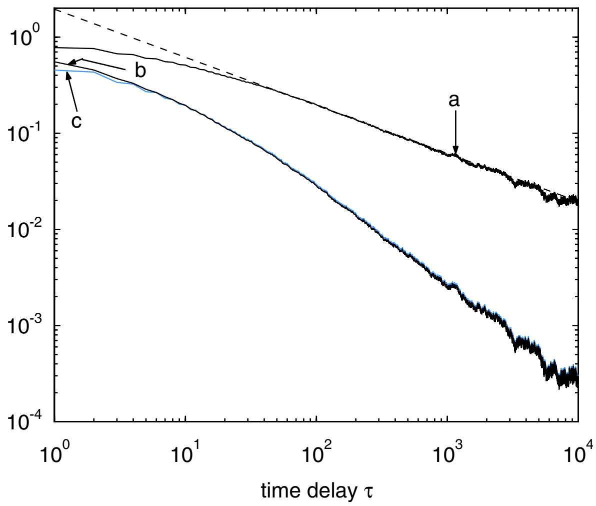

Fig. 4. For the self-similar random signal $y(t)$ of Fig. 3, as a function of the time delay τ at $t = 10$: (a) autocorrelation function $r_{yy}(t, \tau)$, (b) autoinformation function $I_{yy}(t, \tau)$, (c) function $1.5 \times \frac{1}{2} r_{yy}^2(t, \tau)$ matching $I_{yy}(t, \tau)$. The dashed line has slope $-\frac{1}{2}$.

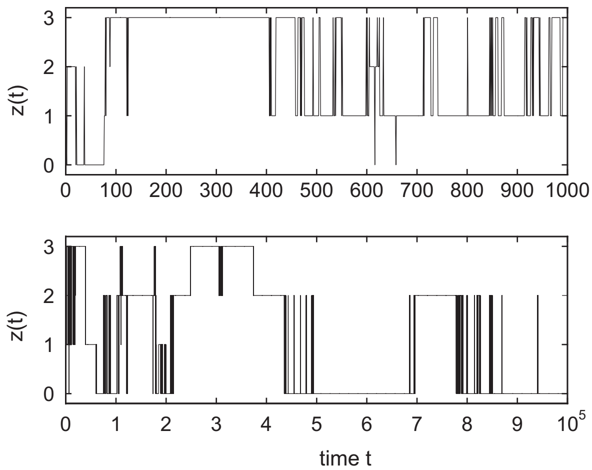

Fig. 5. One realization of the four-state random signal $z(t)$ represented over two time intervals of increasing durations showing the self-similarity of $z(t)$.

Fig. 1, and we did not obtain analytical expressions for $r_{yy}(t, \tau)$ and $I_{yy}(t, \tau)$. Yet, the numerical evaluations of Fig. 4 clearly manifest the long-range dependence of $y(t)$ with power-law evolutions of the autocorrelation $r_{yy}(t, \tau) \sim \tau^{-1/2}$ and of the autoinformation $I_{yy}(t, \tau) \sim \tau^{-1}$ at large delay τ. Also, although $y(t)$ is a nonstationary binary signal and therefore does not belong to the conditions of Section 4.3, Fig. 4 shows that $y(t)$ nevertheless verifies the simple asymptotic relation $I_{yy}(t, \tau) \approx 1.5 \times \frac{1}{2} r_{yy}^2(t, \tau)$ in place of Eq. (64).

Next, we used a random signal $x_1(t)$ similar to $x(t)$ of Fig. 1 to drive a binary random signal $z_1(t) = 0/1$, with $z_1(t)$ flipping its state between 0 and 1 at each time t where $x_1(t) > 0$. Another similar but independent $x_2(t)$ was used to drive another binary signal $z_2(t) = 0/1$ in the same way. At each time t, the binary values of $z_1(t)$ and $z_2(t)$ are collected to form the two bits of the binary representation for a signal $z(t)$; this $z(t)$ therefore results as a four-state random signal. A realization of $z(t)$ is plotted in Fig. 5, over two time intervals of increasing durations, also to provide visual appreciation of self-similarity in $z(t)$.

The self-similarity of $z(t)$ is quantitatively confirmed by its autocorrelation coefficient $r_{zz}(t, \tau)$ which was numerically evaluated and is shown in Fig. 6, revealing a power-law evolution $r_{zz}(t, \tau) \sim \tau^{-1/2}$ at large delay τ. In the same way, numerical evaluation was performed of the autoinformation function $I_{zz}(t, \tau)$ shown in

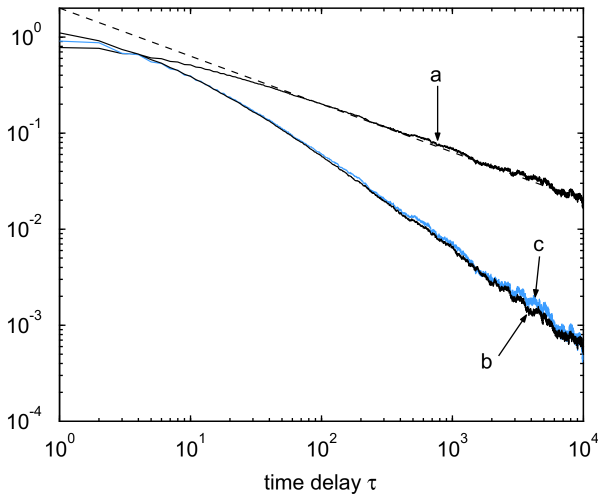

Fig. 6. For the self-similar random signal $z(t)$ of Fig. 5, as a function of the time delay τ at $t = 10$: (a) autocorrelation function $r_{zz}(t, \tau)$, (b) autoinformation function $I_{zz}(t, \tau)$, (c) function $3 \times \frac{1}{2} r_{zz}^2(t, \tau)$ matching $I_{zz}(t, \tau)$. The dashed line has slope $-\frac{1}{2}$.

Fig. 6, which is also found to follow a power-law $I_{zz}(t, \tau) \sim \tau^{-1}$ at large delay τ. Moreover, an asymptotic relation is observed in Fig. 6 as $I_{zz}(t, \tau) \approx 3 \times \frac{1}{2} r_{zz}^2(t, \tau)$.

6. Summary and discussion

In this paper we have presented different models of random signal for which both the autocorrelation function and the autoinformation function are given explicitly in analytical form. These models can especially serve as reference models to test various estimation methods for these functions on experimental signals, although this issue of estimation was not explicitly addressed here [19–22]. In general the autocorrelation and autoinformation functions are not connected in a one-to-one way. However, in the asymptotic regime of large separation times when both functions return to zero, we have investigated conditions (summarized in the sufficient condition of Eq. (49)) where a relation can exist between them. This sufficient condition is fulfilled by common models of stationary random signal, but may not be fulfilled by more complex nonstationary signals. In this direction, we have presented a nonstationary model for several random signals with long-range dependence, that do not a priori fulfill this sufficient condition, but that nevertheless exhibit also a simple asymptotic relation between the autocorrelation and autoinformation functions. The various behaviors observed for the autocorrelation and autoinformation functions are summarized in Table 1.

The results of Table 1 clearly show that the behaviors of the autocorrelation and autoinformation functions do not generally coincide, so that both remain complementary tools for investigating random signals. From Table 1 however, there seems to exist a uniform asymptotic behavior where the autoinformation is proportional to the square of the autocorrelation. Only in specific cases reported in Table 1 has this proportionality been strictly proved to hold, for stationary signals verifying sufficient condition (49). The more complex nonstationary signals also tested in Table 1, although they do not satisfy (49), were also proved theoretically (signal of Fig. 2) or observed numerically (signals of Figs. 4 and 6) to display this proportionality. This asymptotic proportionality between the autoinformation and the squared autocorrelation thus appears to be a very common behavior for random signals, stationary or nonstationary. However, based on Eqs. (43) and (46), this proportionality cannot a priori be expected to hold in general, as long as the square of a linear average (Eq. (43)) is not the average of the squares (Eq. (46)). These possibilities of behaviors exhibited here may be useful for the interpretation of the properties of complex real signals, and the evolution at large separation of the statistical dependence they may contain.

Also, we want to emphasize that the model of Section 5 for long-range dependent random signals, is a recent model, for which the autocorrelation and autoinformation functions are calculated and compared here for the first time. The main advantage of the model of Section 5 is that it is able to generate long-range dependence

Table 1
Summary of the results concerning the behaviors of the autocorrelation and autoinformation functions

	General relation between autocorrelation and autoinformation	Asymptotic relation $I_{xx} \approx A \times \frac{1}{2} r_{xx}^2$
Nonstationary signal	No, based on Eqs. (43) and (45) even for Gaussian or binary signal.	Yes, example of Fig. 6 ($A \approx 3$).
Nonstationary binary signal	No, counterexample of Eq. (71).	Yes, examples of Eq. (40) ($A = 1$), Fig. 2 ($A = 1$), Fig. 4 ($A \approx 1.5$).
Stationary signal	No, based on Eqs. (43) and (45), except for Gaussian or binary signals.	Yes, sufficient condition Eq. (49), example Eq. (18) ($A = (15/9)^2$).
Stationary binary signal	Yes, proof of Eq. (62).	Yes, always with $A = 1$, proof of Eq. (64).
Stationary Gaussian signal	Yes, proof of Eq. (11).	Yes, always with $A = 1$, proof of Eq. (12).

over time horizons of arbitrary durations. This is made possible by the recurrent form of the model, expressed by the first-order recurrence of Eqs. (65)–(67). This recurrence is straightforward to numerically implement, and as long as it is left to run, it generates a random output with long-range dependence which keeps developing in time. This has to be contrasted with other methods for generating long-range dependent random signals, for instance wavelet synthesis, inverse Fourier transform, Cholesky decomposition [23,24]. These methods do not implement a recurrent synthesis over an a priori unlimited time horizon, but a block synthesis, where the duration of the long-range dependent signal has to be specified and limited before the method is run. Especially, with these methods, if one a posteriori wants to add to the signal only one more point with long-range dependence, there is no other way than to restart anew a complete synthesis with a block size longer by one unity. Other methods based on fractional integration can also be used to generate long-range dependent random signals [23,24]. These methods can lead to recurrent algorithms for synthesis. But fractional systems are infinite-order systems, and implementing them in a recurrent form requires a truncation to finite-order, which translates necessarily into a long dependence restricted to a limited time horizon. The method of Section 5, based on a first-order recurrence, does not have these limitations. Another possible approach allowing to generate random sequences with long-range dependence is based on an iterative transformation which randomly alternates expansion and substitution, and which is repeatedly applied to evolve a string of symbols. An interesting example is given by the expansion-modification system presented in Ref. [25], with an extension in Ref. [26]. Both processes of Refs. [25,26] reach power-law behavior of the autocorrelation, in common with our model of Section 5, yet with a generating mechanism which has more a spatiotemporal nature, and which differs from the simple temporal recurrence of Eqs. (65)–(67). Our recurrence model of Section 5 is specially convenient for studying the asymptotic behavior of random signals, there where usually the properties of long-range dependence emerge. This served to us here to compare the asymptotic behaviors of the autocorrelation and autoinformation functions. Also the method for producing $z(t)$ in Section 5, where Eqs. (65)–(67) are used to generate the random bits of the binary representation of a multi-state long-range dependent signal, is new here. This approach, or other approaches using Eqs. (65)–(67) to drive other secondary random signals, can ground various techniques for generation of long-range dependence over arbitrary time horizons. The present theoretical results and models may be useful to contribute to the study of structures and informational contents in complex processes and signals.

References

[1] R.F. Voss, Evolution of long-range fractal correlations and $1/f$ noise in DNA base sequences, Phys. Rev. Lett. 68 (1992) 3805–3808.
[2] H.E. Stanley, S.V. Buldyrev, A.L. Goldberger, S. Havlin, C.-K. Peng, M. Simons, Long-range power-law correlations in condensed matter physics and biophysics, Physica A 200 (1993) 4–24.

[3] W. Ebeling, T. Pöschel, Entropy and long-range correlations in literary English, Europhys. Lett. 26 (1994) 241–245.
[4] W. Li, The study of correlation structures of DNA sequences: a critical review, Comput. Chem. 21 (1997) 257–271.
[5] C. Kayser, W. Einhäuser, P. König, Temporal correlations of orientations in natural scenes, Neurocomputing 52–54 (2003) 117–123.
[6] A.M. Fraser, H.L. Swinney, Independent coordinates for strange attractors from mutual information, Phys. Rev. A. 33 (1986) 1134–1140.
[7] H. Herzel, W. Ebeling, A.O. Schmitt, Entropies of biosequences: the role of repeats, Phys. Rev. E 50 (1994) 5061–5071.
[8] G.A. Darbellay, D. Wuertz, The entropy as a tool for analysing statistical dependences in financial time series, Physica A 287 (2000) 429–439.
[9] M.J. Berryman, A. Allison, D. Abbott, Mutual information for examining correlations in DNA, Fluctuation Noise Lett. 4 (2004) L237–L246.
[10] L. Feng, K. Ju, K.H. Chon, A method for segmentation of switching dynamic modes in time series, IEEE Trans. Systems Man Cybernet. Part B 35 (2005) 1058–1064.
[11] D. Hoyer, B. Pompe, K.H. Chon, H. Hardraht, C. Wicher, U. Zwiener, Mutual information function assesses autonomic information flow of heart rate dynamics at different time scales, IEEE Trans. Biomed. Eng. 52 (2005) 584–592.
[12] M. Dehnert, W.E. Helm, M.T. Hütt, Information theory reveals large-scale synchronization of statistical correlations in eukaryote genome, Gene 345 (2005) 81–90.
[13] W. Li, Mutual information functions versus correlation functions, J. Stat. Phys. 60 (1990) 823–837.
[14] H. Herzel, I. Große, Measuring correlations in symbol sequences, Physica A 216 (1995) 518–542.
[15] A. Papoulis, Probability, Random Variables, and Stochastic Processes, third ed., McGraw-Hill, New York, 1991.
[16] F. Chapeau-Blondeau, (max, +) dynamic systems for modeling traffic with long-range dependence, Fractals 6 (1998) 305–311.
[17] F. Chapeau-Blondeau, A. Monir, Generation of signals with long-range correlation, Electron. Lett. 37 (2001) 599–600.
[18] W. Feller, An Introduction to Probability Theory and its Applications, vol. I, third ed., Wiley, New York, 1968.
[19] R. Moddemeijer, On estimation of entropy and mutual information of continuous distributions, Signal Process. 16 (1989) 233–248.
[20] Y.-I. Moon, B. Rajagopalan, U. Lall, Estimation of mutual information using kernel density estimators, Phys. Rev. E 52 (1995) 2318–2321.
[21] G.A. Darbellay, I. Vajda, Estimation of the information by an adaptive partitioning of the observation space, IEEE Trans. Inf. Theory 45 (1999) 1315–1321.
[22] A. Kraskov, H. Stögbauer, P. Grassberger, Estimating mutual information, Phys. Rev. E 69 (2004) 066138,1–15.
[23] J. Beran, Statistical methods for data with long-range dependence, Stat. Sci. 7 (1992) 404–427.
[24] G. Wornell, Signal Processing with Fractals, Prentice-Hall, New York, 1996.
[25] W. Li, Expansion-modification systems: a model for spatial $1/f$ spectra, Phys. Rev. A 43 (1991) 5240–5260.
[26] M.A. Zaks, Multifractal Fourier spectra and power-law decay of correlations in random substitution sequences, Phys. Rev. E 65 (2001) 011111,1–5.