Adv. Appl. Prob. **30**, 757–776 (1998)
Printed in Northern Ireland

PARAMETRIC ESTIMATION AND SPECTRAL ANALYSIS OF PIECEWISE LINEAR MAPS OF THE INTERVAL

ARTUR LOPES* AND
SÍLVIA LOPES,* *Universidade Federal do Rio Grande do Sul*

Abstract

We present an estimation procedure and analyse spectral properties of stochastic processes of the kind $Z_t = X_t + \xi_t = \phi(T^t(\psi)) + \xi_t$, for $t \in \mathbb{Z}$, where T is a deterministic map, ϕ is a given function and ξ_t is a noise process. The examples considered in this paper generalize the classical harmonic model $Z_t = A\cos(\omega_0 t + \psi) + \xi_t$, for $t \in \mathbb{Z}$. Two examples are developed at length. In the first one, the spectral measure is discrete and in the second it is continuous. In the second example, the time series is obtained from a chaotic map. These two examples exhibit the extremal cases of different possibilities for the spectral measure of time series and they are both associated with ergodic deterministic transformations with noise. We present a method for obtaining explicitly the spectral density function (second example) and the autocorrelation coefficients (first example). In the first example the rotation number plays an important role. We also consider large deviation properties of the estimated parameters of the model.

Keywords: Spectral analysis; chaotic time series; rotation number

AMS 1991 Subject Classification: Primary 60G10
Secondary 58F11

1. Introduction

Consider the stationary process

$$Z_t = A\cos(\omega_0 t + \psi) + \xi_t, \quad \text{for } t \in \mathbb{Z}, \tag{1.1}$$

where $A > 0$ and $\omega_0 \in (-\pi, \pi]$ are constants, $\{\xi_t\}_{t\in\mathbb{Z}}$ is a Gaussian white noise with mean zero and variance σ_ξ^2 and ψ is a uniformly distributed random variable in $(-\pi, \pi]$ independent of the noise process.

The process in (1.1) is the classical harmonic model (see Bloomfield (1976)) for time series analysis. Several different procedures to estimate the frequency ω_0 are known. The spectral distribution function of the model (1.1) is

$$dF_Z(\lambda) = \tfrac{1}{2}(\delta_{\omega_0} + \delta_{-\omega_0}) + \frac{1}{2\pi} \quad \text{for } \lambda \in [-\pi, \pi],$$

where δ_{ω_0} denotes the probability such that for any Borel set A, $\delta_{\omega_0}(A) = 1$, if $\omega_0 \in A$ and $\delta_{\omega_0}(A) = 0$, otherwise.

Received 18 April 1995; revision received 7 May 1997.
* Postal address: Instituto de Matemática, Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves 9500, Porto Alegre, RS–91540-000, Brasil.

Consider the map $T : (-\pi, \pi) \to (-\pi, \pi)$ given by $T(x) = \omega_0 + x \pmod{2\pi}$ and its iterates

$$T^t = \underbrace{T \circ T \circ \cdots \circ T}_{t \text{ times}} \tag{1.2}$$

which satisfy $T^t(x) = \omega_0 t + x \pmod{2\pi}$. We remind the reader that for any given number c, the value $c \pmod{2\pi}$ is the value d where $c = 2\pi n + d$, $0 \leq d < 2\pi$ and $n \in \mathbb{Z}$.

The process (1.1) can be rewritten as

$$Z_t = \phi(T^t(x)) + \xi_t, \quad \text{for } t \in \mathbb{Z}, \tag{1.3}$$

where ψ is a uniformly distributed random variable in $(-\pi, \pi]$ and $\phi(x) = \cos x$.

Our purpose is to analyse stochastic processes of the type (1.3) where the transformation T is a general bijective map from a set $K \subset \mathbb{R}$ (or, more generally, $K \subset \mathbb{R}^n$) to itself preserving an absolutely continuous invariant measure and ϕ is a general continuous function. Formal definitions will be given in the next section.

In this paper we consider two examples where the map T is a piecewise linear transformation. The map T of the second example defines a dynamical system with chaotic behavior. The map T of the first example is not chaotic since it does not involve sensitive dependence on the initial conditions. We will use techniques from ergodic theory (see Cornfeld *et al.* (1982) and Walters (1981)) and large deviations (see Dembo and Zeitouni (1993) and Ellis (1989)) in order to analyse the process (1.3). After the general definitions and properties of processes as in (1.3) given in Sections 2 and 3, we present Example 1 and Example 2, respectively, in Sections 4 and 5.

The spectral measure of the first example is singular with respect to the Lebesgue measure and for the second example there exists a spectral density function.

The stochastic process (1.1) is a particular case of Example 1 which will be analysed in Section 4. We are able to present an estimation procedure to find the parameters (they play the role of the frequency ω_0 in model (1.1)) and also to explicitly exhibit the spectral density function of Example 2 and all the Fourier coefficients of the spectral distribution function of Example 1. A remarkable fact in Example 1 is the appearance of a strong peak in the spectral distribution function at the value corresponding to the rotation number of the map T. The rotation number of a bijective map T is an important invariant previously analysed in dynamical systems (see Devaney (1989)) and it seems also to play an important role in the spectral analysis properties of certain time series obtained from bijective one-dimensional maps.

It is well known in the theory of time series analysis that different models require different estimation procedures. We do not know a general procedure that works for all models of the type (1.3). We propose to use here the sample autocovariance functions at low order to estimate the parameters, but each particular model will require a different approach to estimate the parameters involved.

We also carry out a complete analysis of the deviations of the mean estimated values in the case with no noise. In fact, Example 1 and Example 2 (in the case where $\sigma_\xi^2 = 0$) satisfy a large deviations principle, as will be shown below. The large deviations principle for the case $\sigma_\xi^2 \neq 0$ will be presented in a forthcoming paper (see Carmona *et al.* (1998)).

Only in the large deviations section of this paper have we to assume the existence of no noise. The large deviations properties of the model (see Sections 4 and 5) assure that the estimation procedure is, in some sense, robust.

The general techniques presented in Sections 2, 3 and 5 can also be applied to a wide range of different examples of the type (1.3) (see Lopes *et al.* (1997a), Lopes and Lopes (1996)).

In the examples we consider here, the attractor is the whole set K. Therefore, any initial point $x \in K$ will be in the basin of attraction of the attractor. Any initial point x (outside a set of Lebesgue measure zero on K) will define a sample path $w = (x, T(x), T^2(x), \ldots, T^n(x), \ldots)$ in a set of probability 1 for the associated stochastic stationary process (see Section 2). This phenomenon occurs for transformations T that leave invariant a measure absolutely continuous with respect to the Lebesgue measure. This fact is essential in the theory we consider here. The transformation of Example 2 is expanding and for such a class of maps there are always invariant probabilities absolutely continuous with respect to the Lebesgue measure (see Lasota and Yorke (1973)). The existence of such ergodic measures for Example 1 is shown in Coelho *et al.* (1994). From the ergodicity of the measure, we will be able to estimate integrals by using the method of moments.

As usual, we call $\{\phi(T^t(x))\}_{t\in\mathbb{Z}}$ the *signal process* and $\{\xi_t\}_{t\in\mathbb{Z}}$ the *noise process*. The value σ_ξ determines the strength of the noise. The *signal to noise ratio* is defined by

$$\mathrm{SNR} = 20\log_{10}\left(\frac{\text{standard signal}}{\text{standard noise}}\right). \tag{1.4}$$

As for other kinds of time series models, if the noise is much stronger than the signal, that is, if the signal to noise ratio is strongly negative, the estimation procedure works badly. We present, at the end of Sections 4 and 5, tables showing simulations that confirm the good performance of the method for estimation purposes when the signal to noise ratio has reasonable values.

There is a basic difference between the model (1.3) considered here and the previous work of Tong (1990) and others. In Tong (1990), the model is

$$X_{t+1} = \phi(X_t, \xi_t), \quad \text{for } t \in \mathbb{Z}, \tag{1.5}$$

where ϕ and X_t are deterministic and ξ_t is the noise. In this case, for $\phi(x) = x$, for instance, the influence of the noise as time goes on propagates like

$$T(T(x) + \xi_1) + \xi_2. \tag{1.6}$$

In the present situation, the noise propagation is like

$$T(T(x)) + \xi_2. \tag{1.7}$$

We refer the reader to Takens (1994), Kostelich and Yorke (1990), Ding *et al.* (1993) and Tong (1990) for general properties of time series with chaotic behavior.

2. Stationary stochastic processes

The general setting of chaotic time series we want to analyse is the following. Consider K a compact subset of $\mathbb{R}^n$ with a given Borel σ-algebra $\mathcal{F}$, an invertible continuous transformation $T : K \to K$, an invariant probability $\mathcal{P}$ on K (that is, $\mathcal{P}(T^{-1}(A)) = \mathcal{P}(A)$, for any set $A \in \mathcal{F}$) and $\phi : K \to \mathbb{R}$ a continuous function. We will analyse the stationary stochastic process $\{Z_t\}_{t\in\mathbb{Z}}$ given by

$$Z_t = X_t + \xi_t = (\phi \circ T)(X_{t-1}) + \xi_t, \quad \text{for } t \in \mathbb{Z}. \tag{2.1}$$

The natural measure on $K^{\mathbb{Z}}$ is the product measure on $K^{\mathbb{Z}}$ and it is invariant for the stationary process $\{X_t\}_{t\in\mathbb{Z}}$ or $\{Z_t\}_{t\in\mathbb{Z}}$. The process $\{\xi_t\}_{t\in\mathbb{Z}}$ is considered to be a Gaussian white noise process (see Brockwell and Davis (1987)) independent of $\{(\phi \circ T)(X_t)\}_{t\in\mathbb{Z}}$, with zero mean and variance σ_ξ^2. Observe that in the model (2.1) the random variables Z_t and Z_{t+1} are generally not independent.

We denote the above system by $(K, T, \mathcal{P}, \phi, \mathcal{F}, \sigma_\xi^2)$. Following the terminology of Tong (1990) we may call the system (2.1), when $\sigma_\xi^2 = 0$, the *skeleton* of the system.

For the following definitions we do not consider the noise process $\{\xi_t\}_{t\in\mathbb{Z}}$ in the model (2.1) and we denote the system by $(K, T, \mathcal{P})$. We say that two systems $(K_1, T_1, \mathcal{P}_1)$ and $(K_2, T_2, \mathcal{P}_2)$ (where, for the moment, we do not consider any continuous function ϕ) are *equivalent* in the ergodic theory sense if there exists a map $v : K_1 \to K_2$ invertible (that is, there exists $u : K_2 \to K_1$ such that $v \circ u = id$, $\mathcal{P}_1$-almost everywhere and $u \circ v = id$, $\mathcal{P}_2$-almost everywhere) such that

$$\begin{aligned} &\text{(i)} \quad v^*(\mathcal{P}_2) = \mathcal{P}_1, \text{ where } v^*(\mathcal{P}_2)(A) = \mathcal{P}_2(v^{-1}(A)), \quad \text{for any set } A \in \mathcal{F}. \\ &\text{(ii)} \quad T_2 \circ v = v \circ T_1, \quad \mathcal{P}_1\text{-almost everywhere.} \end{aligned} \tag{2.2}$$

One observes that v plays the role of a change of variables. When v satisfies property (2.2) we say that v is a *conjugacy* between the systems $(K_1, T_1, \mathcal{P}_1)$ and $(K_2, T_2, \mathcal{P}_2)$. We refer the reader to Walters (1981) for precise definitions and general results about equivalence in ergodic theory. It is a simple consequence of (ii) in (2.2) that

$$T_2^t \circ v = v \circ T_1^t, \quad \text{for any } t \in \mathbb{Z}.$$

Given a certain measurable function $\phi : K \to \mathbb{R}$ *the autocovariance function at lag h* (see Brockwell and Davis (1987)) *of the process* $\{X_t\}_{t\in\mathbb{Z}}$ as in (2.1) is given by

$$R_{XX}(h) = E(X_t X_{t+h}) - [E(X_t)]^2 = \int \phi(x)\phi(T^h(x))\,\mathrm{d}\mathcal{P}(x) - \left[\int \phi(x)\,\mathrm{d}\mathcal{P}(x)\right]^2. \tag{2.3}$$

The *autocorrelation function at lag h of the process* $\{X_t\}_{t\in\mathbb{Z}}$ (see Brockwell and Davis (1987)) is given by

$$\rho_X(h) = \frac{R_{XX}(h)}{R_{XX}(0)}, \quad \text{for } h \in \mathbb{N}, \tag{2.4}$$

where $R_{XX}(0) = E[(X_t - E(X_t))^2] = \mathrm{Var}(X_t)$ is the variance of the process.

Proposition 2.1. *If $(K_1, T_1, \mathcal{P}_1)$ and $(K_2, T_2, \mathcal{P}_2)$ are equivalent as in (2.2) then, for any ϕ, the autocovariance functions at lag h of the processes $X_t = \phi \circ v \circ T_1^t$ and $Y_t = \phi \circ T_2^t$ are the same, that is,*

$$R_{XX}(h) = R_{YY}(h), \quad \textit{for any } h \in \mathbb{Z}.$$

The proof is left for the reader.

When F_X satisfies

$$\rho_X(h) = \int_{-\pi}^{\pi} \mathrm{e}^{ih\lambda}\,\mathrm{d}F_X(\lambda), \quad \text{for any } h \in \mathbb{Z}, \tag{2.5}$$

F_X is called *the spectral distribution function*. When

$$\sum_{h=-\infty}^{\infty} |\rho_X(h)| < \infty, \tag{2.6}$$

then

$$f_X(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \mathrm{e}^{-ih\lambda} \rho_X(h) \tag{2.7}$$

is called *the spectral density function* of the process $\{X_t\}_{t\in\mathbb{Z}}$.

Remark. From Proposition 2.1 we conclude that the spectral distribution functions (see (2.5)) of the two stochastic processes, $X_t = (\phi \circ v)(T_1^t)$ and $Y_t = \phi(T_2^t)$, are the same. In conclusion, if we are able to analyse the spectral properties of the system $(K_1, T_1, \mathcal{P}_1, \phi)$ then we are also able to analyse the spectral properties of any equivalent system $(K_2, T_2, \mathcal{P}_2, \phi \circ v)$.

Remark. Expanding maps (see Lopes (1994) for the definition) always have an exponential decay of autocorrelations, for any ϕ Holder continuous function (see Parry and Pollicott (1990)). Therefore, in this case, the spectral density function always exists and the spectrum is of continuous type. The function T of Example 2 in Section 5 is an expanding map.

3. The estimation of the parameters

We shall consider ϕ as a fixed known continuous function, $\mathcal{P}$ is also fixed and T is an unknown transformation indexed by the parameters $(\alpha_1, \alpha_2, \ldots, \alpha_k)$. One of our main purposes in this paper is to estimate the map T, or equivalently, to estimate the parameters $(\alpha_1, \alpha_2, \ldots, \alpha_k)$ from a time series $\{Z_t\}_{t=1}^N$ of size N derived from the stationary stochastic process $\{Z_t\}_{t\in\mathbb{Z}}$ given by (2.1). We also would like to estimate the noise parameter σ_ξ^2.

In the example given before where $T(x) = \omega_0 + x \pmod{2\pi}$ and $\phi(x) = A\cos(x)$, one wants to estimate the frequency ω_0.

The ergodicity of the system justifies our use of the sample autocovariance function.

3.1. Birkhoff's ergodic theorem

In what follows we assume that the system $(K, T, \mathcal{P})$ is *ergodic* (that is, if $T^{-1}(A) = A$ then $\mathcal{P}(A) = 0$ or $\mathcal{P}(A) = 1$, for any $A \in \mathcal{F}$). Birkhoff's ergodic theorem claims that if $\mathcal{P}$ is ergodic and if $\phi : K \to \mathbb{R}$ is $\mathcal{P}$-integrable then for any y $\mathcal{P}$-almost everywhere

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} \phi(T^j(y)) = \int \phi(x)\, \mathrm{d}\mathcal{P}(x). \tag{3.1}$$

Our main tool for estimating the parameters $(\alpha_1, \alpha_2, \ldots, \alpha_k)$ is the ergodic theorem.

Each particular system $(K, T, \mathcal{P}, \phi, \mathcal{F}, \sigma_\xi^2)$ requires a particular method for estimating the parameters $(\alpha_1, \alpha_2, \ldots, \alpha_k)$. It is natural to try to estimate these parameters from the sample autocovariance function at lag h based on the time series $\{Z_t\}_{t=1}^N$, for small values of h.

3.2. Large deviations

The deviations from $(\hat{\alpha}_1, \hat{\alpha}_2, \ldots, \hat{\alpha}_k)$ to $(\alpha_1, \alpha_2, \ldots, \alpha_k)$ in the case where $\sigma_\xi^2 = 0$, that is, when the model (2.1) has only the signal process, is the content of the theory of large deviations as presented by Ellis (1989). The most important property for the large deviations estimates to be robust is the exponential convergence property (see Definition 3.1 below). This property means that the deviation rate is exponentially decreasing. This is true for Example 2 treated in Section 5 since, in that case, the map T is expanding (see Lopes (1994) for the definition and general properties). Example 1, presented in Section 4, does not fit in the context of Lopes

(1994) since the map T is not an expanding one. It is also true that Example 1 has exponentially decreasing deviation rate and this will be proved in Section 4.2. The case $\sigma_\xi^2 \neq 0$ requires a different analysis and is the subject of a forthcoming paper (see Carmona *et al.* (1998)). In this section we consider $\sigma_\xi^2 = 0$.

We now consider a general dynamical system $(K, T, \mathcal{P})$ and a continuous function $f : K \to \mathbb{R}$. We assume that $\mathcal{P}$ is an ergodic probability measure.

In general, there may exist points y such that the equality (3.1) does not hold. Given $\epsilon > 0$, consider the set

$$Q_n(\epsilon) = \left\{ y \in K; \left| n^{-1} \sum_{j=1}^{n} f(T^j(y)) - \int f(x)\,\mathrm{d}\mathcal{P}(x) \right| > \epsilon \right\}.$$

From expression (3.1) it follows that, for any $\epsilon > 0$,

$$\lim_{n\to\infty} \mathcal{P}(Q_n(\epsilon)) = 0. \tag{3.2}$$

If the convergence in (3.2) is very slow, even for large n, we have a certain reasonably large chance of choosing a bad y such that the mean

$$\frac{1}{n} \sum_{j=1}^{n} f(T^j(y))$$

is distant from $\int f(x)\,\mathrm{d}\mathcal{P}(x)$ by more than ϵ. This would be a very bad situation for the estimation purposes presented in Section 4.1. In Section 4.2 we show that the probability of choosing such a bad y is exponentially small with n.

Definition 3.1. The system $(K, T, \mathcal{P}, f)$ has the *exponential convergence property* if for any $\epsilon > 0$, there exists $M > 0$ such that, for any $n > 0$,

$$\mathcal{P}(Q_n(\epsilon)) < \mathrm{e}^{-nM}.$$

For the estimation procedure in Section 4.1 to work properly one should prove that, for any f, the system $(K, T, \mathcal{P}, f)$ satisfies Definition 3.1, where $T = T_{\alpha,\beta}$ (see definition in expression (4.7)). First, we prove that this property is true when the transformation T is given by $T(x) = \omega_0 + x$ and then we derive, by the contraction principle (see Theorem 3.2), the exponential convergence property for $(K, T_{\alpha,\beta}, \mathcal{P}, f)$. This will be the subject of Section 4.2.

Remark. When one needs to estimate

$$E(Z_t Z_{t+1}) = \int \phi(x)\phi(T(x))\,\mathrm{d}\mathcal{P}(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N-1} Z_t Z_{t+1}$$

one should consider large deviation properties for the function $f(x) = \phi(x)\phi(T(x))$ (in the notation of this section).

Definition 3.2. For each $n \in \mathbb{N}$ and $t \in \mathbb{R}$, consider the function

$$c_n(t) = \int \exp\left\{ t \sum_{j=1}^{n} f(T^j(x)) \right\} \mathrm{d}\mathcal{P}(x)$$

and the limit

$$c(t) = \lim_{n\to\infty} \frac{1}{n} \log c_n(t).$$

When such a limit exists, for all t, we call $c\,(t)$ *the free energy of f* or *moment generating function of f*.

When the process is independent the limit in the definition above is superfluous (see Ellis (1989)). However, not all examples we consider here are independent.

Note that, in Definition 3.2, $c\,(0) = 0$.

Definition 3.3. Given the free energy $c\,(t)$, $t \in \mathbb{R}$, of f we define $I(z)$, the Legendre transform of $c(t)$, by

$$I(z) = \sup_{t\in\mathbb{R}}\{tz - c(t)\}.$$

We call $I(z)$ *the deviation function of* f.

Remark. When $c\,(t)$ is differentiable and convex, the deviation function of f is

$$I(z) = t_0 z - c(t_0), \quad \text{where } c'(t_0) = z.$$

Note that if $c\,(t)$ is linear with inclination α, then $I(z) = \infty$ for $z \neq \alpha$, and $I(\alpha) = 0$.

Theorem 3.1. *If $c\,(t)$, the free energy of f, is differentiable on t, then*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathcal{P}(Q_n(\epsilon)) = - \inf_{|z-\int f(x)\,\mathrm{d}\mathcal{P}(x)|>\epsilon} I(z).$$

According to Theorem 3.1 (see Orey (1986)), one concludes that if $I(z)$ is such that $I(z) = 0 \Leftrightarrow z = \int f(x)\,\mathrm{d}\mathcal{P}(x)$, otherwise it is greater than zero, then the system $(K, T, \mathcal{P}, f)$ has the *exponential convergence property*. In particular, any system with a linear free energy (as presented in the above example) has the *exponential convergence property*. Systems which have a linear free energy present the best possible convergence rate.

We shall prove in Section 4.2 that for the transformation T given by $T(x) = \omega_0 + x$ and for any continuous function f, the free energy is linear.

Given a continuous function f, the deviation function I_f of f can be obtained in the following way (see Ellis (1989)):

$$I_f(z) = - \lim_{\epsilon\to 0} \lim_{n\to\infty} \frac{1}{n} \log \mathcal{P}\left\{x \in K;\ \frac{1}{n}\sum_{j=1}^{n} (f \circ T^j)(x) \in [z-\epsilon, z+\epsilon]\right\}. \tag{3.3}$$

The value $I_f(z)$ is the local rate of deviation of the Birkhoff mean around the value z (see Ellis (1989)).

We now explain the contraction principle for two equivalent systems. Given $(K_1, T_1, \mathcal{P}_1)$ and $(K_2, T_2, \mathcal{P}_2)$, suppose that v is a change of coordinates between two systems in the sense of (2.2). Given a function $f : K_2 \to \mathbb{R}$ one considers its deviation function I_f associated to T_2. Consider the random variable $f \circ v$ defined on K_1. We shall obtain similar properties for the deviation function $I_{f\circ v}(r)$ associated to T_1.

Theorem 3.2 (contraction principle for equivalent systems) *Let $f : K_2 \to \mathbb{R}$ be a function and let v be a conjugacy between the systems $(K_1, T_1, \mathcal{P}_1)$ and $(K_2, T_2, \mathcal{P}_2)$. Then $I_{f\circ v} = I_f$.*

We refer the reader to Orey (1986) for the proof of this theorem.

Let us consider now the generalization of the above considerations to a system of equations. When one considers a system $g_1(\alpha, \beta) = k_1 = \int f_1(x)\,\mathrm{d}\mathcal{P}(x)$ and $g_2(\alpha, \beta) = k_2 = \int f_2(x)\,\mathrm{d}\mathcal{P}(x)$, where $g_1, g_2 : \mathbb{R}^2 \to \mathbb{R}$, the analogous property is true (see Orey (1986)). Therefore, the exponential convergence property of I_{f_1} and I_{f_2} implies that $\hat{\alpha}$ and $\hat{\beta}$ have corresponding deviation functions $\tilde{I}_\alpha$ and $\tilde{I}_\beta$ with the exponential convergence property, that is, given $\epsilon > 0$, there exists $M > 0$ such that, for all $n > 0$,

$$\mathcal{P}\Bigg\{x \in K;\, |\hat{\alpha} - \alpha| > \epsilon,\, |\hat{\beta} - \beta| > \epsilon,\ \text{where } g_1(\hat{\alpha}, \hat{\beta}) = \frac{1}{n}\sum_{j=1}^{n} f_1(T^j(x)) = \hat{k}_1$$

$$\text{and } g_2(\hat{\alpha}, \hat{\beta}) = \frac{1}{n}\sum_{j=1}^{n} f_2(T^j(x)) = \hat{k}_2\Bigg\} \le \mathrm{e}^{-Mn}.$$

In conclusion, if I_{f_1} and I_{f_2} have the exponential convergence property and $\hat{\alpha}$ and $\hat{\beta}$ can be obtained by solving the equations $g_1(\hat{\alpha}, \hat{\beta}) = \hat{k}_1$ and $g_2(\hat{\alpha}, \hat{\beta}) = \hat{k}_2$, then one can use the contraction principle to conclude that $\hat{\alpha}$ and $\hat{\beta}$ satisfy the exponential convergence property.

4. Example 1

Consider the two-parameter mapping family $\{T_{a,b} : [0, 1] \to [0, 1];\ a, b \in \mathbb{R}\}$ where $T_{a,b}$ is given by

$$T_{a,b}(x) = \begin{cases} a + \dfrac{1-a}{b}x, & \text{if } 0 \le x < b \\ \dfrac{a}{1-b}(x - b), & \text{if } b \le x \le 1, \end{cases} \tag{4.1}$$

with a and b constants ($b \ne 0, 1$). Let α be the derivative of $T = T_{a,b}$ on $[0, b)$ and β its derivative on $[b, 1]$. Then,

$$\alpha = T'_{a,b}(x) = \frac{1-a}{b}, \quad \text{if } 0 \le x < b \quad \text{and} \quad \beta = T'_{a,b}(x) = \frac{a}{1-b}, \quad \text{if } b \le x \le 1. \tag{4.2}$$

The ergodic properties of the family $\{T_{a,b} : [0, 1] \to [0, 1];\ a, b \in \mathbb{R}\}$ are analysed in Coelho *et al.* (1994). In particular this transformation is uniquely ergodic, that is, there exists only one invariant probability measure for $T_{a,b}$. However, the transformation in Example 2 is not uniquely ergodic.

In Example 1 we want to analyse the estimation of the parameters a and b and the spectral analysis of the process $\{X_t\}_{t\in\mathbb{Z}}$ defined in (4.3) below. Notice that when $b = 1 - a$, the transformation $T_{a,b}$ of Example 1 is $T(x) = a + x \pmod 1$, which corresponds to the model (1.1).

First we examine the system with no noise. The case with noise can be subsequently analysed in a simple way.

4.1. Estimation

By using the notation introduced in Section 2, for a given transformation $T_{a,b}$ and $\phi(x) = x$ one considers the signal process $\{X_t\}_{t\in\mathbb{Z}}$ given by

$$X_t = T_{a,b}(X_{t-1}), \quad \text{for } t \in \mathbb{Z}. \tag{4.3}$$

To estimate the unknown constants a and b is the same as to estimate α and β, since one has the following identities:

$$\alpha = \frac{1-a}{b} \quad \text{and} \quad \beta = \frac{a}{1-b} \Leftrightarrow a = \frac{\beta(\alpha-1)}{\alpha-\beta} \quad \text{and} \quad b = \frac{1-\beta}{\alpha-\beta}. \tag{4.4}$$

Therefore, for the sake of simplicity in our analysis we shall estimate the parameters α and β.

The invariant measure $\mathcal{P}_{\alpha,\beta} = \mathcal{P}$ (see Coelho *et al.* (1994)) for the process $\{X_t\}_{t\in\mathbb{Z}}$, in terms of α and β, is given by the density

$$\varphi_{\alpha,\beta}(x) = \varphi(x) = \frac{1}{c}\frac{1}{x + (\beta/\alpha)(1-x)} = \frac{1}{c}\frac{1}{(\alpha-\beta)x+\beta}, \tag{4.5}$$

where

$$c = \frac{1}{\beta-\alpha}\log\left(\frac{\beta}{\alpha}\right) = \frac{1}{(\beta/\alpha)-1}\log\left(\frac{\beta}{\alpha}\right). \tag{4.6}$$

For a set $A \subset [0,1]\times[0,1]$, with Lebesgue measure equal to 1, for all $(\alpha,\beta) \in A$, the map $T_{\alpha,\beta}$ is ergodic for $\mathcal{P}_{\alpha,\beta} = \mathcal{P}$. We will assume $(\alpha,\beta) \in A$ in what follows.

From the expressions (4.1) and (4.4) the transformation $T_{\alpha,\beta}$ is given by

$$T_{\alpha,\beta}(x) = \begin{cases} \dfrac{\beta(\alpha-1)}{\alpha-\beta} + \alpha x, & \text{if } 0 \le x < \dfrac{1-\beta}{\alpha-\beta} \\ \beta\left(x - \dfrac{1-\beta}{\alpha-\beta}\right), & \text{if } \dfrac{1-\beta}{\alpha-\beta} \le x \le 1. \end{cases} \tag{4.7}$$

The following integrals are useful in understanding the estimation and the spectral analysis presented below.

$$\int_0^y \varphi(x)\,\mathrm{d}x = \frac{\log(((\alpha-\beta)y+\beta)/\beta)}{\log(\alpha/\beta)}. \tag{4.8a}$$

$$E(Z_t) = E(X_t) = \int_0^1 x\varphi(x)\,\mathrm{d}x = \frac{1}{\log(\alpha/\beta)} - \frac{\beta}{\alpha-\beta} = k_1. \tag{4.8b}$$

$$E(Z_t^2) = E(X_t^2) + \sigma_\xi^2 = \int_0^1 x^2\varphi(x)\,\mathrm{d}x + \sigma_\xi^2 = \left(\frac{\beta}{\alpha-\beta}\right)^2 + \frac{\alpha-3\beta}{2(\alpha-\beta)\log(\alpha/\beta)} + \sigma_\xi^2. \tag{4.8c}$$

$$E(Z_tZ_{t+1}) = E(X_tX_{t+1}) = \int_0^1 xT(x)\varphi(x)\,\mathrm{d}x = \left(\frac{\beta}{\alpha-\beta}\right)^2 + \frac{1+\alpha\beta-4\beta}{2(\alpha-\beta)\log(\alpha/\beta)} = k_2. \tag{4.8d}$$

Some of these integrals are obtained only after long calculations.

TABLE 1: Parameters of Example 1 and their respective estimates.

α	β	σ_ξ	SNR[a]	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}_\xi$
0.63049	3.31683	0.000	∞	0.63383	3.26947	0.00031
0.99566	1.03169	0.100	9.208	1.00230	0.98326	0.10487
1.21035	0.44972	0.100	9.138	1.21481	0.44430	0.08482
1.21035	0.44972	0.295	−0.258	1.21481	0.44298	0.08601
1.19998	0.80002	0.000	∞	1.19983	0.79988	0.00000
2.32675	0.19141	0.100	8.796	2.40095	0.19690	0.09674
2.32675	0.19141	0.430	−3.874	6.47962	0.03473	0.30758

[a] SNR, signal-to-noise ratio.

For estimation purposes, in the case where $\phi(x) = x$, one needs integrals (4.8b) and (4.8d). We are able to estimate k_1 and k_2 using

$$\frac{1}{N}\sum_{t=1}^{N} Z_t = \hat{k}_1 \approx E(Z_t) = k_1$$

$$\frac{1}{N}\sum_{t=1}^{N-1} Z_t Z_{t+1} = \hat{k}_2 \approx E(Z_t Z_{t+1}) = k_2. \tag{4.9}$$

Then, the estimates of α and β are obtained as the solutions of the system of equations given by (4.8b) and (4.8d).

After we find $\hat{\alpha}$ and $\hat{\beta}$, the value $\hat{\sigma}_\xi^2$ can be easily estimated from integral (4.8c) and from the value $N^{-1}\sum_{t=1}^{N} Z_t^2 \approx E(Z_t^2)$ obtained from a time series derived from the process $\{Z_t\}_{t\in\mathbb{Z}}$. In this way we find all the parameters of our model.

Remark. Notice that from the above equations one obtain two pairs of solutions. One pair is the value (α, β). The other pair is $(\tilde{\alpha}, \tilde{\beta})$ such that

$$T_{\alpha,\beta}^{-1} = T_{\tilde{\alpha},\tilde{\beta}}.$$

The stationary processes as in (1.3) generated by $T_{\alpha,\beta}$ and $T_{\tilde{\alpha},\tilde{\beta}}$, respectively, have the same spectral distribution. This indeterminacy is analogous to the one observed in the harmonic model (1.1) where ω_0 and $-\omega_0$ determine the same spectral measure $\frac{1}{2}(\delta_{\omega_0} + \delta_{-\omega_0})$.

Using the large deviations results of Section 4.2 the conclusion is that with very high probability (with infinite exponential velocity) the values $\hat{k}_1$ and $\hat{k}_2$ will be respectively close to k_1 and k_2.

In the simulations, where the sample size is $N = 5000$ whenever σ_ξ^2 is equal to zero and $N = 2000$ otherwise, we obtained the results shown in Table 1.

4.2. Large deviations

In this section we analyse the large deviations associated with the mean

$$\lim_{N\to\infty} \frac{1}{N}\sum_{j=1}^{N} (f \circ T^j)(x) \tag{4.10}$$

where f is a continuous function on [0, 1] and $T = T_{\alpha,\beta}$ is the map defined by (4.7).

In Coelho *et al.* (1994) it is shown that the function

$$v(y) = \int_0^y \varphi(x)\,\mathrm{d}x = \frac{\log([(\alpha-\beta)y+\beta]/\beta)}{\log(\alpha/\beta)} \tag{4.11}$$

is a change of coordinates in the sense of (2.2) between the systems ([0, 1], $T_{\alpha,\beta}$, dx) and ([0, 1], T_{ω_0}, dx), where $T_{\omega_0}(x) = \omega_0 + x$ with $\omega_0 = \log\alpha/\log(\alpha/\beta)$. The value ω_0 is called the *rotation number of* $T_{\alpha,\beta}$ (see Devaney (1989) for definitions). We first analyse the large deviations properties of $T_{\omega_0}(x) = \omega_0 + x$ and then, after that, we derive by a contraction principle argument (see Theorem 3.2) the exponential decreasing property for the system ([0, 1], $T_{\alpha,\beta}$, dx) by using the change of coordinates v.

Consider a rotation $T(x) = \omega_0 + x$, where ω_0 is an irrational number. It is well known that, in this case, the Lebesgue measure dx is ergodic for T (see Devaney (1989)) where $\mathcal{P}(A)$ is the length of A, for any interval A. We first analyse the deviation properties for the transformation $T(x) = \omega_0 + x$ and a continuous function $f : [0, 1] \to \mathbb{R}$.

We now present several results for a continuous function $f : [-1, 1] \to \mathbb{R}$ such that $f(-1) = f(1)$. The result is also true for any continuous function f, by using an approximation argument in $L^1(\mathrm{d}x)$.

Proposition 4.1. *Given $\epsilon > 0$, there exists $M > 0$ such that, for all $x \in S_1$ and all $N > M$,*

$$\frac{1}{N}\sum_{j=1}^{N} f(T^j(x)) \in \left[\int f(z)\,\mathrm{d}z - \epsilon, \int f(z)\,\mathrm{d}z + \epsilon\right].$$

We refer the reader to Lopes and Lopes (1996) for the proof.

Remark. We consider the function $f(x) = xT(x)$ to estimate large deviations of the autocovariance at lag 1 of the process $\{X_t\}_{t\in\mathbb{Z}} = \{T(X_{t-1})\}_{t\in\mathbb{Z}}$.

Now we show that the free energy $c\,(t)$ is linear.

Theorem 4.1. *The free energy $c\,(t)$ is linear and, therefore, the deviation function I satisfies $I(z) = \infty$ for $z \neq \int f(x)\,\mathrm{d}\mathcal{P}(x)$, otherwise it is zero.*

Proof. One needs to show that

$$c(t) = t\int f(x)\,\mathrm{d}\mathcal{P}(x).$$

One observes that

$$\begin{aligned}
&\left|n^{-1}\log\int \exp\left\{t\sum_{j=1}^{n} f(T^j(x))\right\}\mathrm{d}\mathcal{P}(x) - t\int f(x)\,\mathrm{d}\mathcal{P}(x)\right| \\
&\quad = \left|n^{-1}\log\left(\int\left(\exp\left\{t\sum_{j=1}^{n} f(T^j(x))\right\} - \exp\left\{nt\int f(x)\,\mathrm{d}\mathcal{P}(x)\right\}\right)\mathrm{d}\mathcal{P}(x)\right)\right| \\
&\quad = \left|n^{-1}\log\left[\int \exp\left\{nt\left(n^{-1}\sum_{j=1}^{n} f(T^j(x)) - \int f(x)\,\mathrm{d}\mathcal{P}(x)\right)\right\}\mathrm{d}\mathcal{P}(x)\right]\right|.
\end{aligned}$$

From Proposition 4.1, given $\epsilon > 0$, there exists $M > 0$ such that, for any $x \in S_1$ and all $n > M$,

$$\left| n^{-1} \sum_{j=1}^{n} f(T^j(x)) - \int f(x)\,\mathrm{d}\mathcal{P}(x) \right| < \epsilon.$$

Therefore, for all $n > M$,

$$\begin{aligned} &\left| n^{-1} \log \int \exp\left\{ t \sum_{j=1}^{n} f(T^j(x)) \right\} \mathrm{d}\mathcal{P}(x) - t \int f(x)\,\mathrm{d}\mathcal{P}(x) \right| \\ &\quad \le \left| n^{-1} \log \left[\int \exp\left\{ nt \left(n^{-1} \sum_{j=1}^{n} f(T^j(x)) - \int f(x)\,\mathrm{d}\mathcal{P}(x) \right) \right\} \mathrm{d}\mathcal{P}(x) \right] \right| \\ &\quad \le \left| n^{-1} \log \int \exp\{nt(\pm\epsilon)\}\,\mathrm{d}\mathcal{P}(x) \right| = n^{-1} \log \exp\{\pm \epsilon n t\} = \pm \epsilon t. \end{aligned}$$

As t is fixed, by taking $\epsilon \to 0$ one concludes that $c(t) = t \int f(x)\,\mathrm{d}\mathcal{P}(x)$.

Since for $T(x) = \omega_0 + x$ and for any continuous function f the deviation function I_f has the exponential convergence property and since $v(y) = \int_0^y \varphi_{\alpha,\beta}(x)\,\mathrm{d}x$ defines an equivalence between the systems $(K, T, \mathrm{d}x)$ and $(K, T_{\alpha,\beta}, \varphi_{\alpha,\beta})$ then, from Theorem 3.2, one concludes that, for a given continuous function g, the deviation function I_g associated with the system $(K, T_{\alpha,\beta}, \varphi_{\alpha,\beta}, g)$ also has the exponential convergence property. This follows from the fact that v is a continuous function and by considering, in Theorem 3.2, $g = f \circ v$, with $f = g \circ v^{-1}$.

4.3. Spectral analysis

For a given $T = T_{\alpha,\beta}$ and the corresponding invariant density $\varphi = \varphi_{\alpha,\beta}$ we consider the signal process $\{X_t\}_{t\in\mathbb{Z}} = \{(\phi \circ T_{\alpha,\beta})(X_{t-1})\}_{t\in\mathbb{Z}}$.

From the expressions (4.5) and (4.6) one observes that the density function $\varphi_{\alpha,\beta}(x)$ depends only on the quotient $\Delta = \alpha/\beta$. Consider now the transformation T^h, for any $h \in \mathbb{Z}$, where $T = T_{\alpha,\beta}$ is given by the expression (4.7). From Coelho *et al.* (1994) it is known that

$$T^h(x) = T_{\alpha_h,\beta_h}(x) \qquad \text{where } \alpha_h = \frac{b_h}{1-a_h} \quad \text{and} \quad \beta_h = \frac{a_h}{1-b_h} \tag{4.12}$$

with $a_h = T^h(0)$ and $b_h = T^{-h}(0)$ and also that

$$\frac{\alpha_h}{\beta_h} = \frac{\alpha}{\beta} \quad \text{for any } h \in \mathbb{N},$$

and hence

$$\varphi_{\alpha_h,\beta_h} = \varphi_{\alpha,\beta} \quad \text{for any } h \in \mathbb{N}.$$

Remark. The above property (4.12) is essential for the rest of our argument. This property makes the computation of the autocorrelation coefficients possible. The conclusion is that, for

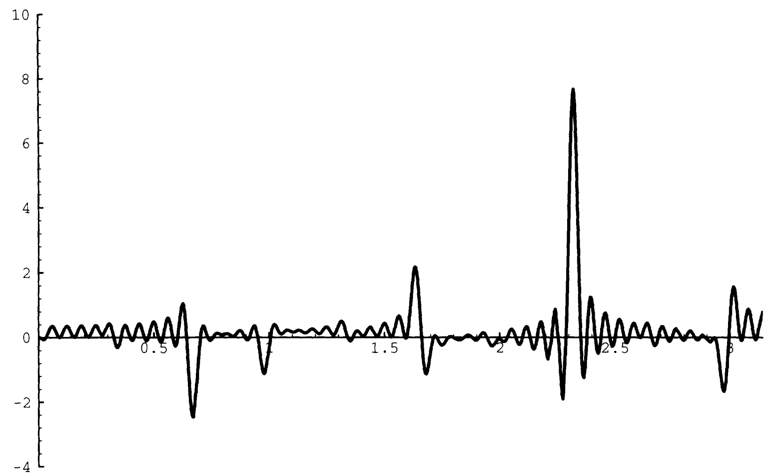

FIGURE 1: The spectral distribution function $f_X(\lambda)$ for Example 1, as in (4.17) when $\sigma_\xi^2 = 0$, $\alpha = 2.41809$ and $\beta = 0.22052$.

any continuous function ϕ and $h \in \mathbb{N}$,

$$\begin{aligned} E(X_t X_{t+h}) &= \int \phi(x)\phi(T^h(x))\varphi(x)\,\mathrm{d}x \\ &= \int \phi(x)\phi(T_{\alpha_h,\beta_h}(x))\varphi_{\alpha,\beta}(x)\,\mathrm{d}x \\ &= \int \phi(x)\phi(T_{\alpha_h,\beta_h}(x))\varphi_{\alpha_h,\beta_h}(x)\,\mathrm{d}x. \end{aligned}$$

As we know $\int \phi(x)\phi(T_{\alpha,\beta}(x))\varphi_{\alpha,\beta}(x)\,\mathrm{d}x$ (see integral (4.8d)), for any α and β, one can calculate $\int \phi(x)\phi(T_{\alpha_h,\beta_h}(x))\varphi_{\alpha_h,\beta_h}(x)\,\mathrm{d}x$, for any $h \in \mathbb{N}$.

Notice that $E(X_t X_{t+h}) = E(X_t X_{t-h})$, for all $h \in \mathbb{N}$. Therefore, we are able to obtain the exact values of $R_{XX}(h)$, for all $h \in \mathbb{Z}$, from the positive and negative orbit of zero by T (since α_h and β_h depend only on a_h and b_h).

We now consider $\phi(x) = x$. In this case the spectral measure is not a function as one can see from numerical experiments (see Figure 1).

First one observes that the process $\{X_t\}_{t\in\mathbb{Z}} = \{T_{\alpha,\beta}(X_{t-1})\}_{t\in\mathbb{Z}}$ has mathematical expectation given by integral (4.8b). That is,

$$E(X_t) = \frac{1}{\log(\alpha/\beta)} - \frac{\beta}{\alpha - \beta}, \quad \text{for all } t \in \mathbb{Z}.$$

We want to derive the spectral distribution function of the process $\{Z_t\}_{t\in\mathbb{Z}}$. We first consider the autocorrelation $\rho_X(h)$ at lag h of the process $\{X_t\}_{t\in\mathbb{Z}} = \{T_{\alpha,\beta}(X_{t-1})\}_{t\in\mathbb{Z}}$ and then use Herglotz's theorem for the process $\{X_t\}_{t\in\mathbb{Z}}$.

Remark. The Fourier coefficients of the spectral distribution function in the case where $T(x) = \omega_0 + x$ are given by $\rho_X(h) = \cos(h\omega_0) = \cos(T^h(0))$, for $h \in \mathbb{Z}$, that is, they are determined by the iterates T^h of zero. The next theorem obtains a similar property for the transformation $T_{\alpha,\beta}$ and $\phi(x) = x$.

Theorem 4.2. *The spectral distribution function of the process*

$$Z_t = T^t_{\alpha,\beta}(\cdot) + \xi_t = T_{\alpha,\beta}(X_{t-1}) + \xi_t, \quad \textit{for } t \in \mathbb{Z},$$

is given by

$$\mathrm{d}F_Z(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \mathrm{e}^{-i\lambda h} \rho_X(h) + \frac{1}{2\pi}, \quad \textit{for } \lambda \in [-\pi, \pi] \tag{4.13}$$

where $\rho_X(h)$ is given by $R_{XX}(h)/R_{XX}(0)$ with

$$R_{XX}(h) = \frac{1 + \alpha_h \beta_h}{2(\alpha_h - \beta_h)\log(\alpha_h/\beta_h)} - \frac{1}{[\log(\alpha_h/\beta_h)]^2} \tag{4.14}$$

and

$$R_{XX}(0) = \frac{\alpha + \beta}{2(\alpha - \beta)\log(\alpha/\beta)} - \frac{1}{[\log(\alpha/\beta)]^2}, \tag{4.15}$$

α and β being given by the expression (4.2) and

$$\alpha_h = \frac{1 - a_h}{b_h}, \quad \beta_h = \frac{a_h}{1 - b_h}, \quad a_h = T^h(0), \quad b_h = T^{-h}(0).$$

Now we consider $\phi(x) = \cos(2\pi x)$. One wants to calculate the spectral distribution of the process

$$Z_t = X_t + \xi_t = \cos(2\pi T_{\alpha,\beta}(X_{t-1})) + \xi_t, \quad \text{for } t \in \mathbb{Z}.$$

For this purpose we need the following integral:

$$E(X_t X_{t+1}) = \int_0^1 \cos(2\pi x)\cos(2\pi T(x))\varphi(x)\,\mathrm{d}x = \frac{1}{2\log(\alpha/\beta)} \times k \tag{4.16}$$

where, after long calculations,

$$\begin{aligned} k = {} & \cos(2d\beta)[\mathrm{ci}(d(\alpha+1)) + \mathrm{ci}(d\alpha(\beta+1)) - \mathrm{ci}(d\beta(\alpha+1)) - \mathrm{ci}(d(\beta+1))] \\ & + \sin(2d\beta)[\mathrm{si}(d(\alpha+1)) + \mathrm{si}(d\alpha(\beta+1)) - \mathrm{si}(d\beta(\alpha+1)) - \mathrm{si}(d(\beta+1))] \\ & + \mathrm{ci}(d(\alpha-1)) + \mathrm{ci}(d\alpha(\beta-1)) - \mathrm{ci}(d(\beta-1)) - \mathrm{ci}(d\beta(\alpha-1)), \end{aligned}$$

with $d = 2\pi/(\alpha - \beta)$, $\mathrm{ci}(x)$ is the cosine integral and $\mathrm{si}(x)$ is the sine integral (see Gradshteyn and Ryzhik (1965), p. 928).

In order to calculate the spectral distribution function, one should obtain the autocorrelation coefficients of order k, $E(X_t X_{t+k})$, of such distribution (the Fourier coefficients of the spectral measure) by substituting in (4.16) the values of α and β by α_h and β_h (see expression (4.12)).

Another example is considered in the next theorem.

Theorem 4.3. *The spectral distribution function of the process*

$$Z_t = \phi(T^t_{\alpha,\beta}(\cdot)) + \xi_t = \cos(2\pi\, T_{\alpha,\beta}(X_{t-1})) + \xi_t, \quad \textit{for } t \in \mathbb{Z},$$

is given by

$$\mathrm{d}F_Z(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \mathrm{e}^{-i\lambda h} \rho_X(h) + \frac{1}{2\pi}, \quad \textit{for } \lambda \in [-\pi, \pi], \tag{4.17}$$

where $\rho_X(h)$ *is given by* $R_{XX}(h)/R_{XX}(0)$ *with*

$$R_{XX}(h) = \frac{1}{2\log(\alpha_h/\beta_h)} \times k_h - \frac{1}{[\log(\alpha_h/\beta_h)]^2} \times l_h$$

where

$$\begin{aligned} k_h = {} & \cos(2d_h\beta_h)[\mathrm{ci}(d_h(\alpha_h+1)) + \mathrm{ci}(d_h\alpha_h(\beta_h+1)) - \mathrm{ci}(d_h\beta_h(\alpha_h+1)) - \mathrm{ci}(d_h(\beta_h+1))] \\ & + \sin(2d_h\beta_h)[\mathrm{si}(d_h(\alpha_h+1)) + \mathrm{si}(d_h\alpha_h(\beta_h+1)) - \mathrm{si}(d_h\beta_h(\alpha_h+1)) - \mathrm{si}(d_h(\beta_h+1))] \\ & + \mathrm{ci}(d_h(\alpha_h-1)) + \mathrm{ci}(d_h\alpha_h(\beta_h-1)) - \mathrm{ci}(d_h(\beta_h-1)) - \mathrm{ci}(d_h\beta_h(\alpha_h-1)), \end{aligned}$$

and

$$l_h = \{\cos(d_h\beta_h)[\mathrm{ci}(d_h\alpha_h) - \mathrm{ci}(d_h\beta_h)] + \sin(d_h\beta_h)[\mathrm{si}(d_h\alpha_h) - \mathrm{si}(d_h\beta_h)]\}^2$$

with

$$d_h = \frac{2\pi}{\alpha_h - \beta_h}, \quad \alpha_h = \frac{1-a_h}{b_h}, \quad \beta_h = \frac{a_h}{1-b_h},$$

$a_h = T^h(0)$ *and* $b_h = T^{-h}(0)$. *The variance* $\mathrm{Var}(X_t)$ *is given by*

$$\begin{aligned} R_{XX}(0) = {} & \frac{1}{2\log(\alpha/\beta)}\{\cos(2d\beta)[\mathrm{ci}(2d\alpha) - \mathrm{ci}(2d\beta)] + \sin(2d\beta)[\mathrm{si}(2d\alpha) - \mathrm{si}(2d\beta)]\} \\ & + \frac{1}{2} - \frac{1}{[\log(\alpha/\beta)]^2} \times l, \end{aligned}$$

where

$$l = \{\cos(d\beta)[\mathrm{ci}(d\alpha) - \mathrm{ci}(d\beta)] + \sin(d\beta)[\mathrm{si}(d\alpha) - \mathrm{si}(d\beta)]\}^2$$

with

$$d = \frac{2\pi}{\alpha - \beta}, \quad \alpha = \frac{1-a}{b}, \quad \beta = \frac{a}{1-b}.$$

In Figure 1 we plot the graph of the Fourier series $(1/2\pi)\sum_{h=-100}^{100} \mathrm{e}^{-i\lambda h}\rho_X(h)$ when $\alpha = 2.41809$ and $\beta = 0.22052$. We consider here an approximation of the generalized spectral density function $f_X(\lambda)$ up to an order of 100.

Remark. The rotation number of $T_{\alpha,\beta}$ is

$$\theta_1 = \frac{\log(\alpha)}{\log(\alpha/\beta)}$$

and the rotation number of $T_{\tilde{\alpha},\tilde{\beta}} = T_{\alpha,\beta}^{-1}$ is

$$\theta_2 = \frac{\log(\beta)}{\log(\beta/\alpha)}.$$

One observes that $\theta_1 + \theta_2 = 1$. We denote by ζ the smaller of θ_1 and θ_2. Therefore, $\zeta \leq 0.5$. We call ζ the *rotation number of the stochastic process*.

It is extremely interesting that, for any α and β, the spectral measure is not a Dirac delta function concentrated on the rotation number of $T_{\alpha,\beta}$ (we checked the coefficients $\rho_X(h)$) but it has a very strong peak on the value $2\pi\zeta$ where ζ is the rotation number of the process. In other words, the spectral distribution is very close to

$$\tfrac{1}{2}(\delta_{2\pi\,\zeta} + \delta_{-2\pi\,\zeta}) = \tfrac{1}{2}(\delta_{2\pi\,\theta_1} + \delta_{-2\pi\,\theta_1}),$$

where $\theta_1 \leq 0.5 \leq \theta_2$ were defined above.

In conclusion, if one applies the Fourier transform to the data a strong peak appears at the rotation number of the process. This property requires a deeper analysis in order to understand the spectral distribution function given by (4.17). Notice in Figure 1 the strong peak at the value $2\pi\zeta = 2.31671$, where ζ is the rotation number of the process when $\alpha = 2.41809$ and $\beta = 0.22052$ (corresponding to the values $a = 0.1423$ and $b = 0.3547$).

We remind the reader that if $a = 1 - b$ then the rotation number of $T_{\alpha,\beta}$ is equal to a and, in fact, in this case, the spectral distribution function is a Dirac delta function $\frac{1}{2}(\delta_{\pi a} + \delta_{-\pi a})$, when $\phi(x) = \cos(2\pi x)$.

Notice that for $T_{\alpha,\beta}(x) = a+x \pmod 1$, the inverse map $T_{\alpha,\beta}^{-1} = T_{\tilde{\alpha},\tilde{\beta}}$ is such that $T_{\tilde{\alpha},\tilde{\beta}}(x) = x - a \pmod 1$. In this case, $\zeta = \pi|a|$.

5. Example 2

Sakai and Tokumaru (1980) introduce the following model of chaotic time series. For a given constant $a \in (0, 1)$ consider the transformation $T_a : [0, 1] \to [0, 1]$ given by

$$T_a(x) = \begin{cases} \dfrac{x}{a}, & \text{if } 0 \leq x < a \\ \dfrac{1-x}{1-a}, & \text{if } a \leq x \leq 1. \end{cases} \tag{5.1}$$

The Lebesgue measure $\mathrm{d}x$ is invariant and ergodic for the transformation T_a (see Lasota and Yorke (1973)). In the notation of section 2, $\mathcal{P}(A)$ is the length of A, for any interval A.

We now consider the stochastic process

$$Z_t = X_t + \xi_t = T_a(X_{t-1}) + \xi_t, \quad \text{for } t \in \mathbb{Z}, \tag{5.2}$$

where $\phi(x) = x$.

The autocovariance function at lag h of the process $\{X_t\}_{t\in\mathbb{Z}}$ in (5.2) (see Sakai and Tokumaru (1980)) is given by

$$R_{XX}(h) = \int_0^1 xT^h(x)\,\mathrm{d}x - [E(X_t)]^2 = \tfrac{1}{12}(2a-1)^h, \quad \text{for } h > 0, \tag{5.3}$$

where $E(X_t) = \frac{1}{2}$ and $R_{XX}(0) = \mathrm{Var}(X_t) = \frac{1}{12}$.

One can use the above integral to estimate a from the system $(K, T, \mathcal{P}, \phi, \mathcal{F}, \sigma_\xi^2)$ where $\phi(x) = x$. This will be done in Section 5.1. After that, we shall analyse the spectral properties of the process in (5.2). The result presented here is generalized in Lopes *et al.* (1997a,b).

TABLE 2: Parameters of Example 2 and their estimates.

a	σ_ξ	SNR[a]	$\hat{a}$	$\hat{\sigma}_\xi$
0.273001011	0.100	9.208	0.27249	0.11406
0.273001011	0.295	−0.188	0.25564	0.29990
0.273001011	3.000	−20.334	2.33035	3.08462
0.273001011	0.000	∞	0.27551	0.01262
0.400010101	0.100	9.208	0.38870	0.08430
0.400010101	0.000	∞	0.39978	0.00852
0.400010101	0.000	∞	0.40707	0.02615
0.500010111	0.000	∞	0.49147	0.01856
0.783000101	0.000	∞	0.77875	0.06136

[a] SNR, signal-to-noise ratio.

5.1. Estimation

Let us consider now the case with noise. We want to estimate the parameters $a \in (0, 1)$ and σ_ξ^2. From the ergodic theorem, one observes that

$$\lim_{N\to\infty} N^{-1} \sum_{t=1}^{N-1} Z_t Z_{t+1} = \int_0^1 x T_a(x)\,\mathrm{d}x = \frac{a+1}{6} \tag{5.4}$$

since $\{X_t\}_{t\in\mathbb{Z}}$ and $\{\xi_t\}_{t\in\mathbb{Z}}$ are independent processes and since

$$E(\xi_t) = 0 \quad \text{and} \quad E(\xi_t \xi_{t+h}) = \begin{cases} \sigma_\xi^2, & \text{if } h = 0 \\ 0, & \text{if } h \neq 0, \end{cases}$$

for all $t \in \mathbb{Z}$. The last equality in expression (5.4) comes from (5.3) when $h = 1$ and from the fact that $E(X_t) = \frac{1}{2}$. From the ergodic theorem and the independence, one has

$$\lim_{N\to\infty} N^{-1} \sum_{t=1}^{N} Z_t^2 = \lim_{N\to\infty} N^{-1} \sum_{t=1}^{N} X_t^2 + \lim_{N\to\infty} N^{-1} \sum_{t=1}^{N} \xi_t^2 = \int_0^1 x^2\,\mathrm{d}x + \sigma_\xi^2 = \tfrac{1}{3} + \sigma_\xi^2. \tag{5.5}$$

Therefore, by using a time series $\{Z_t\}_{t=1}^N$ of size N derived from the stochastic process $\{Z_t\}_{t\in\mathbb{Z}}$ given by (5.2), the estimators $\hat{a}$ and $\hat{\sigma}_\xi^2$ of a and σ_ξ^2 can be obtained implicitly from expressions (5.4) and (5.5) and, thus, are given by

$$\hat{a} = 6\left(N^{-1} \sum_{t=1}^{N-1} Z_t Z_{t+1}\right) - 1 \approx 6 \int_0^1 x T_a(x)\,\mathrm{d}x - 1$$

$$\hat{\sigma}_\xi^2 = N^{-1} \sum_{t=1}^{N} Z_t^2 - \tfrac{1}{3} \approx \int_0^1 x^2\,\mathrm{d}x - \tfrac{1}{3}.$$

In the simulations, where the sample size is $N = 5000$ whenever σ_ξ^2 is equal to zero and $N = 2000$ otherwise, we obtained the results shown in Table 2.

5.2. Large deviations

The map T is an expanding one and the function $\phi(x) = x$ is Holder continuous. Therefore, from the differentiability of the free energy (see Lopes (1994)), the exponential convergence property is true. The conclusion is that, with very high probability the sample autocovariances at lags 1 and 0 estimate with high accuracy the autocovariance of lag 1 and the variance of the process, respectively.

Finally, by the contraction principle (see the end of Section 3.2) the estimates $\hat{a}$ and $\hat{\sigma}_{\xi}^2$ also satisfy the exponential convergence property.

The central limit theorem holds for expanding systems (see Parry and Pollicott (1990)) and therefore, for Example 2.

5.3. Spectral analysis

The main obstacle to the spectral analysis of Example 2 is that the map T_a is not invertible. Therefore, the autocovariance function $R_{XX}(h)$ of the process $\{X_t\}_{t\in\mathbb{Z}}$, given by expression (5.2), for negative lag h does not have a precise meaning. For the estimation of the parameters there is no problem, since we just need the positive lag h. In fact, $h = 0$ and $h = 1$ are enough.

We propose analysing the natural extension F of T_a, instead of T_a itself. The natural extension is a canonical way of embedding a non-invertible dynamical system in an invertible one. We refer the reader to Pollicott (1986) and Bogomolny and Carioli (1993) (see Section 3) for general considerations on the natural extension map.

In Example 2, the natural extension of T_a is the map $F : [0, 1] \times [0, 1] \to [0, 1] \times [0, 1]$ such that

$$F(x, y) = (T(x), G(x, y)), \quad \text{for any } (x, y) \in [0, 1] \times [0, 1],$$

where

$$G(x, y) = \begin{cases} ya, & \text{if } 0 \le x < a \\ (a-1)y + 1, & \text{if } a \le x \le 1. \end{cases}$$

The map F is invertible and it is easy to see that the Lebesgue measure $\mathrm{d}x\,\mathrm{d}y$ is invariant and ergodic for F.

As a particular example, we mention that the Baker map is the natural extension of the tent map (with inclination 2). Therefore, we consider the dynamical system $(K, F, \mathcal{P})$ where $K = [0, 1]\times[0, 1]$ and $\mathcal{P}$ is the Lebesgue measure $\mathrm{d}x\,\mathrm{d}y$ on $[0, 1]\times[0, 1]$. Instead of $\phi(x) = x$, one can consider $\phi(x, y) = \Pi(x, y) = x$ for any $(x, y) \in [0, 1] \times [0, 1]$ as a random variable. In the setting of Section 2, we analyse in this section the system $(K, F, \mathcal{P}, \Pi, \mathcal{F}, \sigma_{\xi}^2)$. Now, if $h \ge 0$ then

$$\int_0^1 xT^h(x)\,\mathrm{d}x = \int_0^1\int_0^1 x\Pi(F^h(x, y))\,\mathrm{d}x\,\mathrm{d}y = \int_0^1\int_0^1 \Pi(x, y)\Pi(F^h(x, y))\,\mathrm{d}x\,\mathrm{d}y$$

and we obtain, from the expression (5.3), $R_{XX}(h)$ for positive h when $X_t = \Pi \circ F^t$. As the map F is invertible, it makes sense to estimate, for $h > 0$, the integral

$$\int_0^1\int_0^1 \Pi(x, y)\Pi(F^{-h}(x, y))\,\mathrm{d}x\,\mathrm{d}y.$$

As the measure $\mathrm{d}x\,\mathrm{d}y$ is invariant for F, then

$$\int_0^1\int_0^1 \Pi(x,y)\Pi(F^{-h}(x,y))\,\mathrm{d}x\,\mathrm{d}y = \int_0^1\int_0^1 \Pi(F^h(x,y))\Pi(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_0^1 xT^h(x)\,\mathrm{d}x.$$

After these results one can obtain the spectral density function associated to the stochastic process $\{X_t\}_{t\in\mathbb{Z}}$. The last term in the above equalities has already been calculated (see (5.3)).

Theorem 5.1. *The spectral density function of the stochastic process*

$$Z_t = X_t + \xi_t = (\Pi \circ F)(X_{t-1}) + \xi_t, \quad \textit{for } t \in \mathbb{Z},$$

is given by

$$f_Z(\lambda) = \frac{2a(1-a)}{\pi[1-2(2a-1)\cos(\lambda)+(2a-1)^2]} + \frac{1}{2\pi}, \quad \textit{for } \lambda \in [-\pi,\pi]. \tag{5.6}$$

Proof. Since $R_{XX}(h)$ is given by the expression (5.3) and goes to zero exponentially when $h \to +\infty$, the spectral density function (see (2.7)) does exist and it is given after some algebra by

$$f_X(\lambda) = \frac{2a(1-a)}{\pi[1-2(2a-1)\cos(\lambda)+(2a-1)^2]},$$

for all $\lambda \in [-\pi,\pi]$. The spectral density function of the process $\{Z_t\}_{t\in\mathbb{Z}}$ follows from this.

In Lopes *et al.* (1997a,b) we also analyse the spectral density function of other chaotic time series.

The spectrum of the signal process $\{X_t\}_{t\in\mathbb{Z}}$ is continuous. Notice that if a is small then the spectral density is larger around π and if a is larger the spectral density is larger around zero.

References

BOGOMOLNY, E. AND CARIOLI M. (1993). Quantum maps from transfer operators. *Physica D* **67**, 88–112.

BLOOMFIELD, P. (1976). *Fourier Analysis of Time Series: An Introduction.* Wiley, New York.

BROCKWELL, P. J. AND DAVIS, R.A. (1987). *Time Series: Theory and Methods.* Springer, New York.

CARMONA, S., LANDIM, C., LOPES, A. AND LOPES, S. (1998). A level-2 large deviations principle for the autocorrelation function of uniquely ergodic transformations with noise. *J. Statist. Phys.* **91**, 395–421.

COELHO, Z., LOPES, A. AND ROCHA, L. F. (1994). Absolutely continuous invariant measures for a class of affine interval exchange maps. *Proc. Amer. Math. Soc.* **123**, 3533–3542.

CORNFELD, I. P., FOMIN, S. V. AND SINAI, YA. G. (1982). *Ergodic Theory.* Springer, New York.

DEMBO, A. AND ZEITOUNI O. (1993). *Large Deviations Techniques and Applications.* Jones and Bartlett, Boston.

DEVANEY R. L. (1989). *An Introduction to Chaotic Dynamical System.* Addison-Wesley, Redwood City, CA.

DING, M. Z., GREBOGI, C., OTT, E., SAUER, T. AND YORKE, J. A. (1993). Estimating correlation dimension from a chaotic time series: when does a plateau onset occur? *Physica D* **69**, 404–424.

ELLIS, R. (1989). *Entropy, Large Deviations and Statistical Mechanics.* Springer, New York.

GRADSHTEYN, I. S. AND RYZHIK, I. M. (1965). *Table of Integrals, Series and Products.* Academic Press, New York.

KOSTELICH, E. AND YORKE, J. A. (1990). Noise reduction: finding the simplest dynamical system consistent with data. *Physica D* **41**, 183–196.

LASOTA, A. AND YORKE, J. A. (1973). On the existence of invariant measures for piecewise monotonic transformations. *Trans. Amer. Math. Soc.* **186**, 481–488.

LOPES, A. (1994). Entropy, pressure and large deviation. *Cellular Automata, Dynamical Systems and Neural Networks.* Ed. E. Goles and S. Martinez. Kluwer, Massachusetts, pp. 79–146.

LOPES, A. AND LOPES, S. (1995). Parametric estimation and spectral analysis of chaotic time series. *Technical Report*, UFRGS, **42,** Series A.

LOPES, A. AND LOPES, S. (1996). Unique ergodicity, large deviations and parametric estimation. Preprint.

LOPES, A., LOPES, S. AND SOUZA, R. (1997a). On the spectral density of a class of chaotic time series. *J. Time Series Anal.* **18,** 465–474.

LOPES, A., LOPES, S. AND SOUZA, R. (1997b). Spectral analysis of expanding one-dimensional chaotic transformations. *Random and Computational Dynamics* **15,** 65–80.

OREY, S. (1986). Large deviations in ergodic theory. In *Seminar on Stochastic Processes*, ed. K. L. Chung *et al.* Birkhäuser, Boston, pp 195–249.

PARRY, W. AND POLLICOTT, M. (1990). Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Asterisque* **187–188.**

POLLICOTT, M. (1986). Distribution of closed geodesics on the modular surface and quadratic irrationals. *Bull. Soc. Math. France.* **114,** 431–446.

SAKAI, H. AND TOKUMARU, H. (1980). Autocorrelations of a certain chaos. *IEEE Trans. Acoust., Speech and Signal Processing* **28,** 588–590.

TAKENS, F. (1994). Analysis of non-linear time series, a survey. Preprint.

TONG, H. (1990). *Non-linear Time Series: A Dynamical System Approach.* Clarendon Press, Oxford.

VARADHAN, S. R. S. (1988). *Large deviations and applications.* Lecture Notes in Mathematics **1362**, ed. P. Diaconis *et al.* Springer, Berlin, pp. 1–49.

WALTERS, P. (1981). *An Introduction to Ergodic Theory.* Springer, New York.