Estimation of the Entropy and Other Dynamical Invariants for Piecewise Affine Chaotic Maps

Tommaso Addabbo, Ada Fort, D. Papini, S.Rocchi, Valerio Vignoli

Department of Information Engineering
University of Siena, 53100 Italy
(e-mail: addabbo@dii.unisi.it)

⋆ RESEARCH MANUSCRIPT ⋆

– PLEASE REFER TO THE PUBLISHED PAPER[1] –

Abstract

In this paper we discuss an efficient iterative method for the estimation of the chief dynamical invariants of chaotic systems based on stochastically stable piecewise affine maps (e.g., the invariant measure, the Lyapunov exponent as well as the Kolmogorov-Sinai entropy). Referring to the theory developed by C. Liverani, we provide a theoretical tool for determining *a priori* a sufficient number of iterations that guarantees a given estimation accuracy. The proposed method represents an alternative to the Monte-Carlo methods and to other methods based on the discretization of the Frobenius-Perron operator, such as the well known Ulam's method. Our estimation method converges not slower than exponentially and it requires a computation complexity that grows linearly with the iterations. The proposed approach can be used to efficiently estimate any order statistics of a symbolic source based on a piecewise affine mixing map.

1 Introduction

Chaotic dynamical systems have been widely investigated for the generation of sequences with predetermined statistical properties [3, 20, 21, 28]. Such kind of dynamical systems can be suitably used, e.g., for the design of truly random number generators [3], for issuing spreading sequences in DS-CDMA telecommunication systems [26], or for enhancing the electromagnetic compatibility of electronic devices [10].

Among the wide class of nonlinear dynamical systems considered in literature, several types of Piecewise Affine Expanding (PWAE) maps have been proposed for the circuit implementation of stochastic sources with *tunable statistical characteristics*: in such cases, depending on the application, by changing the system parameters one or more target statistical features of the source can

[1] International Journal of Bifurcation and Chaos, 2009, vol. 19, n. 12, p. 4175 - 4195.

be theoretically modulated at will (e.g., the autocorrelation, or any higher order expectations) [3, 21, 28, 30]. Unfortunately, when passing from theory to practice, in chaos-based circuits the involved non-idealities (e.g., due to the fabrication process tolerances) cause a drift of the system parameters with respect to the nominal design values, resulting into a possible critical deterioration of the statistical features of the generated sequences [3, 5].

Since most of the circuit non-idealities are random quantities, the sensitivity of the chaotic circuits to the non-idealities must be framed within the theory of stochastic perturbation of chaotic systems [2, 4, 7, 9]. An important condition for studying the effects of the perturbations on the statistical characterization of a dynamical system is to know the *information generation mechanism* which is associated to the perturbed chaotic system, i.e., to know how its ergodic behavior is affected by the circuit non-idealities. About this issue, in [1, 2] the authors provided an efficient and accurate method for estimating the invariant measure of PWAE chaotic maps.

Within the above discussed theoretical framework, as itemized in the following, the aim of this paper is to complete and extend the work presented in [2]. In detail, the chief contributions can be summarized in three points:

1) The method proposed in [2] is extended and applied for estimating both the Lyapunov exponent and the Kolmogorov-Sinai entropy associated to a generic PWAE map. In detail, the estimation of the Kolmogorov-Sinai entropy is obtained by the direct computation of high order statistics characterizing the chaotic process. As shown in this paper, the proposed approach results to be much more efficient than those algorithms based, e.g., on Monte-Carlo methods, or based on the discretization of the Frobenius-Perron operator (such as the well known Ulam's method) [6, 11, 14, 17, 18, 23, 27, 28].

2) The method for estimating the invariant measure proposed in [2] is based on an iterative procedure that converges not slower than exponentially. In this paper we provide a theoretical tool for determining *a priori* a sufficient number of iterations that guarantees a given estimation accuracy. The result is based on the theoretical work of C. Liverani, who studied the decay of the correlation in piecewise affine maps [25]. The theory of Liverani, that was proposed within a purely mathematical background, is reused and applied in this paper to set a method for determining the accuracy of our estimation procedure.

3) Finally, we carry out an exhaustive comparison between our estimation procedure and the well known Ulam's universal method. The comparison is discussed from a theoretical point of view, from an operative point of view and from an algorithmic complexity point of view.

Paper organization: The original contribution of this work is condensed in Sections 4, 5, 6 and 7, whereas the aim of Sections 2 and 3 is to make the paper self-contained: in Section 2 we introduce the notation and some basic definitions from the Ergodic Theory, whereas in Section 3 we complete the presentation of the iterative estimation method, firstly introduced in [2]. Most of the results given in Section 2 are very well known to those fellows working on chaotic symbolic dynamics. Nevertheless, in this paper we spend this brief section to recall such theoretical elements, since they represent the fundamentals for the rest of the paper. The Section 3 is devoted to a complete detailed description of the iterative estimation method. The proofs of theorems are not reported, since they can be found in [2].

In Section 4 we analyze the speed of convergence of the method, and we provide theoretical results about its assessment. In Section 5 we discuss the estimation of the Lyapunov Exponent, whereas we devote Section 6 for discussing the problem of estimating the Kolmogorov-Sinai entropy of a given PWAE map. More in general, in this section we face the problem of estimating any order statistics for a symbolic source based on a PWAE map. With the Section 7 we complete the comparison between our method and the well know Ulam's method. Conclusions and references close the paper.

2 Notation and Basic Setup

Let $I \subset \mathbb{R}$ be a bounded interval with endpoints $a, b \in \mathbb{R}$, $a < b$. We denote with $\chi_I : \mathbb{R} \to \{0, 1\}$ the usual characteristic function of I, with $\overline{I}$ the closure of I, with $\mathcal{B}_I$ the usual Borel σ-algebra of subsets of I, and with $\lambda : \mathcal{B}_I \to \mathbb{R}^+$ the Lebesgue measure. The triplet $(I, \mathcal{B}_I, \lambda)$ is the Lebesgue space measure over I. We say that a property holds almost everywhere (a.e.) on I if the subset $F \subset I$ on which the property fails has zero measure, i.e., $\exists H \in \mathcal{B}_I$ such that $F \subseteq H$ and $\lambda(H) = 0$. With reference to the Lebesgue integration theory, the notation $L^p(I)$ denotes the set of functions $f : I \to \mathbb{R}$ such that $\int_I |f(x)|^p dx < \infty$, with $0 < p \in \mathbb{N}$, whereas $L^\infty(I)$ is the set of a.e. bounded measurable functions. We recall that $L^p(I)$ and $L^\infty(I)$ can be made Banach spaces with respect to the norms $\|f\|_p = (\int_I |f(x)|^p dx)^{\frac{1}{p}}$ and $\|f\|_\infty = \inf\{M \in \mathbb{R}^+ \text{ such that } \{x \in I : |f(x)| > M\} \text{ has zero measure}\}$, respectively.

We define $\mathcal{P}_I$ as the set of all finite partitions of intervals[2] of I, and if $Q \in \mathcal{P}_I$, we denote the endpoints of Q with $\mathcal{E}_Q = \{p_0, \ldots, p_q\}$, with $p_i < p_j$ if $i < j$. If there exists a positive number M such that $\forall Q \in \mathcal{P}_I$

$$V_I(f, Q) = \sum_{k=1}^{\#Q} |f(p_k) - f(p_{k-1})| \leq M, \tag{1}$$

then f is said to be of *bounded variation* on I and the quantity $V_I(f) = \sup_{\mathcal{P}_I}\{V_I(f, Q)\}$ is the *total variation* of f on I. Accordingly, we can define the following subset of $L^1(I)$

$$BV(I) = \{f \in L^1(I) : \\ \exists g \in L^1(I) \text{ such that } V_I(g) < \infty \text{ and } g = f \text{ a.e.}\}. \tag{2}$$

Note that $BV(I)$ also contains functions with infinite variation that are a.e. equal to a bounded variation function. In such case we define the *infimum total variation* of $f \in BV(I)$ as $\tilde{V}_I(f) = \inf\{V_I(g) : g \in BV(I), g = f \text{ a.e.}\}$. The set $BV(I)$ is a vector space of functions, which can be made a Banach space with the norm $\|f\|_{BV} = \|f\|_1 + \tilde{V}_I(f)$, and it can be proved that that $BV(I)$ is dense in $L^1(I)$ [9]. Since in this paper we deal with probability density functions (pdfs), we introduce the subset $D_1(I) = \{f \in L^1(I) : \|f\|_1 = 1, f \geq 0\}$ and the subset $D_{BV}(I) = BV(I) \cap D_1(I)$.

In this paper we consider one dimensional dynamical systems $x_{n+1} = \tau(x_n)$, where $\tau : I \to I$ is a transformation satisfying the following definition.

[2] If $P \in \mathcal{P}_I$ then $P = \{I_i, i = 1, \ldots, p, \text{with } p > 0\}$, with $I_i \cap I_j = \oslash$ for $i \neq j$ and $\cup_{i=1}^{p} I_i = I$.

Definition 1 *The map $\tau : I \to I$ is piecewise affine expanding (PWAE) if and only if it is onto and there exists a partition $Q \in \mathcal{P}_I$, with $Q = \{I_1, \ldots, I_q\}$, such that $\tau|_{I_i}$ is a linear function of the form $\tau|_{I_i}(x) = \alpha_i x + \beta_i$, with $\alpha_i, \beta_i \in \mathbb{R}$ and $|\alpha_i| > 1$ $(i = 1, \ldots, q)$.*

We denote with $\mathcal{T}_{aex}(I)$ the class of PWAE transformations over I. For a given map in $\mathcal{T}_{aex}(I)$ there is an infinite number of partitions satisfying the above definition. We call a partition Q the *defining partition* of τ if it has the smallest number q of elements such that the restriction to each of them of τ is affine.

If either one or both of the endpoints a, b of I do not belong to I, we can extend any map $\tau \in \mathcal{T}_{aex}(I)$ to $\overline{I}$, and we have $\tau(a), \tau(b) \in \overline{I}$. Moreover, we denote with $\tau_+(x)$ and $\tau_-(x)$ the limits $\lim_{y \to x^+} \tau(y)$ and $\lim_{y \to x^-} \tau(y)$, respectively, and for $x = a, b$ we assume for notation convenience that $\tau_+(x) = \tau_-(x) = \tau(x)$.

2.1 The Frobenius-Perron operator

For any non-singular transformation $\tau : I \to I$ we define the Frobenius-Perron operator $\Theta_\tau : L^1(I) \to L^1(I)$ as follows: for any $f \in L^1(I)$, the function $\Theta_\tau f$ is the unique (up to λ-a.e. equivalence) element in $L^1(I)$ such that

$$\int_A \Theta_\tau f(x) dx = \int_{\tau^{-1}(A)} f(x) dx, \qquad A \in \mathcal{B}_{\mathbb{R}}. \tag{3}$$

The validity of the above definition derives from the Radon-Nikodym theorem. The mechanism ruled by the Frobenius-Perron operator is intuitively summarized by the following remark: assuming the initial condition x_0 of the dynamical system as a random variable[3] distributed over I according to the pdf $f_0 \in D_1(I)$, the pdf $f_1 = \Theta_\tau f_0$ describes the distribution of the random variable $x_1 = \tau(x_0)$, i.e., the system state after one iteration. Accordingly, the operator Θ_τ rules the evolution of the initial pdf f_0 induced by the deterministic transformation τ. If $\tau \in \mathcal{T}_{aex}(I)$ then the Frobenius-Perron operator Θ_τ satisfies the following properties [9].

Proposition 1 ([9]) *Let $\tau \in \mathcal{T}_{aex}(I)$. The Frobenius-Perron operator $\Theta_\tau : L^1(I) \to L^1(I)$ has the form*

$$\Theta_\tau f(x) = \sum_{\tau(y)=x} \frac{f(y)}{\left| \frac{d\tau}{dy}(y) \right|}. \tag{4}$$

The linear functional operator Θ_τ preserves integrals and $\Theta_\tau : D_1(I) \to D_1(I)$. Moreover, if $\phi \in D_{BV}(I)$ then $\Theta_\tau f \in D_{BV}(I)$.

2.2 Symbolic Dynamics

Symbolic dynamics describes the mechanism which turns a chaotic systems into a stochastic source of symbols. Formally, for a given chaotic map τ, we refer

[3] Once assuming the initial condition x_0 as a random variable, the obtained sequence is stochastic: for a chaotic map this assumption is founded on the sensitivity of the dynamics to the initial condition, which physically manifests itself, as the time passes, as an exponential growth of the uncertainty about the prediction of the dynamical evolution.

to a finite collection of nonempty subsets of I, i.e., a *symbolic partition* $R = \{S_1, \ldots, S_m\}$, such that $i \neq j \Leftrightarrow S_i \cap S_j = \oslash$, $\cup_{i=1}^{m} S_i = I$ and $\lambda(S_i) > 0$ for $i = 1, \ldots, m$. By calling the set $\mathcal{M} = \{1, \ldots, m\}$ the *alphabet of the source*, a sequence of symbols $\{s_n\}$ is determined by a function $\sigma_R : I \to \mathcal{M}$, defined as

$$\sigma_R(x_n) = i \Leftrightarrow x_n \in S_i, \qquad i \in \mathcal{M}. \tag{5}$$

Since we are dealing with deterministic chaos, each symbolic sequence is completely specified by the map τ, the initial condition x_0 and the symbolic partition R. Accordingly, if we denote with $\mathcal{M}^\infty$ the set of infinite symbolic sequences $\mathbf{s} = \{s_n\}$ with symbols in $\mathcal{M}$, the induced symbolic dynamics is described by a function $W_{\tau,R} : I \to \mathcal{M}^\infty$ that assigns to each initial condition $x_0 \in I$ a symbolic sequence $\mathbf{s}_{x_0}$ according to

$$\mathbf{s}_{x_0} = W_{\tau,R}(x_0) = \{\sigma_R(\tau^n(x_0)), n \in \mathbb{N}\}. \tag{6}$$

If we assume the initial condition x_0 as a random variable, the function $W_{\tau,R}$ turns a dynamical system into a symbolic source that can be analyzed by the classical Shannon Information Theory.

Definition 2 *A partition R is called a* generating partition *for the map τ iff the function $W_{\tau,R}$ is injective, i.e., for $x, y \in I$, $x \neq y \Rightarrow W_{\tau,R}(x) \neq W_{\tau,R}(y)$.*

2.3 Entropy and Other Dynamical Invariants

The ergodicity of chaotic systems reveals itself in many forms handled by Ergodic Theory. In this subsection we recall the definition of those theoretical quantities, named *dynamical invariants*, that are dealt with in this paper.

Invariant probability measures: Let $(I, \mathcal{B}_I, \mu)$ be a probability measure space. The measure $\mu : \mathcal{B}_I \to \mathbb{R}^+$ is *invariant with respect to* $\tau : I \to I$ if and only if for any $A \in \mathcal{B}_I$ we have that $\mu(A) = \mu(\tau^{-1}(A))$, and we say that τ preserves μ. In the following we consider measures that are absolutely continuous (a.c.) with respect to the Lebesgue measure, i.e., that are defined as

$$\mu(A) = \int_A f(x)dx, \qquad A \in \mathcal{B}_I, \tag{7}$$

where the pdf $f \in D^1(I)$ is called the density of μ, and if τ preserves μ the pdf f is said invariant. We abbreviate eq. (7) using the compact notation $\mu = f \cdot \lambda$, and if $f \in D_1(I)$ then μ is a probabilistic measure. Not all nonsingular transformations admit an a.c. invariant measure [22], whereas the existence of a.c. invariant measures with density of bounded variation is assured for maps belonging to $\mathcal{T}_{aex}(I)$ [9]. As a final remark, we note that if f is an invariant pdf, then $\Theta_\tau f = f$ a.e., and the pdf f is an eigenfunction of the functional linear operator Θ_τ, associated to the eigenvalue 1. If we assume x_0 as a random variable distributed over I according to an invariant pdf f, then $\forall n \in \mathbb{N}$ also the random variable $x_n = \tau^n(x_0)$ is distributed over I according to f.

Lyapunov Exponent: In this paper we adopt the ergodic definition of the Lyapunov Exponent[4] for a chaotic map τ in one dimension. Accordingly, let

[4]The Birkhoff Ergodic Theorem provides the link between this definition and the classical one used by experimentalists when analyzing chaotic systems [9].

$(I, \mathcal{B}_I, \mu)$ be a probability space and let the measure $\mu = f \cdot \lambda$ be invariant with respect to $\tau : I \to I$. The Lyapunov Exponent associated to τ with respect to μ is defined as

$$L_\tau(\mu) = \int_I \ln |\tau'(x)| \, f(x) dx, \tag{8}$$

where $\tau' = \frac{d\tau}{dx}$ and it is assumed that $\ln \circ \tau' \in L^1(I)$. The eq. (8) characterizes the almost sure rate of separation of trajectories starting from infinitesimally close initial conditions belonging to the support of μ.

Entropy: In the classical Information Theory the Shannon entropy is a quantity that can be used for quantifying the unpredictability of an information source. In this paragraph we recall some definitions about entropy developed within the Ergodic Theory. The link between the Shannon entropy and the following definitions is made clearer in the next sections.

Definition 3 *Let $(I, \mathcal{B}_I, \mu)$ be a probability space, and let $R = \{I_1, \ldots, I_m\}$ be a finite symbolic partition of I. The* entropy of the partition R *is the quantity*

$$H(R) = -\sum_{i=1}^{m} \mu(I_i) \log_2 \mu(I_i), \tag{9}$$

where we assume the quantity $0 \cdot \log_2 0 = 0$.

Since we are using logarithms with base 2, the result is expressed in bits. Given the partition R, we denote with $\tau^{-1}(R)$ the collection of sets

$$\tau^{-1}(R) = \{\tau^{-1}(I_1), \tau^{-1}(I_2), \ldots, \tau^{-1}(I_m)\}, \tag{10}$$

and we can define a new partition, called a *refinement of one pullback iteration*, by mutually intersecting the elements of R with those of $\tau^{-1}(R)$, i.e., referring to the notation introduced in Sec. 2.2,

$$R \vee \tau^{-1}(R) = \{I_i \cap \tau^{-1}(I_j) : (i, j) \in \mathcal{M}^2\}. \tag{11}$$

In general we write $\tau^0(R) = R$ and

$$\bigvee_{k=0}^{n-1} \tau^{-k}(R) = \left\{ \bigcap_{k=0}^{n-1} \tau^{-k}(I_{i_k}) : (i_0, \ldots, i_{n-1}) \in \mathcal{M}^n \right\}. \tag{12}$$

Definition 4 *The* measure-theoretic entropy *for the dynamical system $\tau : I \to I$ with respect to the partition R is defined as*

$$H_R(\tau, \mu) = \lim_{n \to \infty} \frac{1}{n} H\left(\bigvee_{k=0}^{n-1} \tau^{-k}(R) \right) \tag{13}$$

Definition 5 *The* Kolmogorov-Sinai entropy *for the dynamical system $\tau : I \to I$ is defined as*

$$H_{KS}(\tau, \mu) = \sup_{R \in \mathcal{P}_I} H_R(\tau, \mu), \tag{14}$$

where the supremum is taken over all the possible finite partitions.

3 An Efficient Method for Estimating the Invariant Measure in Mixing PWAE Maps

Invariant measures play a key role in the analysis of the stochastic aspects of chaotic dynamics, since they provide the necessary information for retrieving any order statistics of the chaotic process [28]. The main problem when dealing with PWAE maps with nontrivial parameter values is to determine their invariant measures. The issue may appear simplified for PWAE transformations that satisfy the Markovian property [2], but in practical cases the problem of determining if a given map is Markovian typically results unfeasible. Accordingly, the estimation of invariant measures is achieved resorting to numerical calculations, and in literature several methods have been proposed (see, e.g., [6, 11, 14, 17, 18, 23, 27, 28] and the references cited therein).

The efficient method for the estimation of the invariant measure proposed in [2] is valid for stochastically stable PWAE maps. For such kind of maps there exists an unique invariant pdf $f^* \in D_{BV}(I)$ such that[5]

$$\lim_{n\to\infty} \|\Theta_\tau^n(f) - f^*\|_1 = 0, \qquad \forall f \in D_{BV}(I). \tag{15}$$

Among the stochastically stable systems there are those defined by *mixing* maps [9]. The mixing assumption represents a minor practical restriction, since most of the chaotic systems used in ICT applications are exact, and exact systems always satisfy the mixing property [9]. For mixing PWAE maps, the stochastic stability derives from the inequality

$$\|\Theta_\tau^n(f) - f^*\|_\star \leq K r^n \|f\|_{BV}\,, \tag{16}$$

where K and r are positive quantities that depend on τ, with $0 < r < 1$ [9] and, since the above inequality holds regardless of the norm chosen among $\|\cdot\|_{BV}$ and $\|\cdot\|_p$, with $1 \leq p \leq \infty$, we adopt the 'any-norm' notation $\|\cdot\|_\star$. In a stochastically stable PWAE maps, regardless of the distribution of the initial condition, any order statistics of the stochastic process associated to the dynamics settle on the statistics of the ergodic process ruled by τ and its unique invariant measure $\mu^* = f^* \cdot \lambda$. According to (16) the transient statistical behavior vanishes at least exponentially.

3.1 Evolution of Piecewise Constant pdfs

The proofs of the following propositions were presented in [2]. We start this section stating a fundamental property of PWAE maps.

Proposition 2 *Let $\tau \in \mathcal{T}_{aex}(I)$, and let Q be its defining partition. If $f \in L^1(I)$ is piecewise constant a.e.[6], then $\Theta_\tau f$ is piecewise constant a.e..*

A piecewise constant pdf $f \in D_1(I)$ can be described with a matrix

$$M = \begin{pmatrix} x_0 & l_0 \\ \vdots & \\ x_N & l_N \\ x_{N+1} & 0 \end{pmatrix}, \tag{17}$$

[5] Formally, the uniqueness of f^* is understood up to an a.e. equivalence in $D_{BV}(I)$.

[6] We refer to functions a.e. equal to simple functions with a finite number of discontinuities and constant over positive measure intervals partitioning I.

such that $x_0 = a$, $x_{N+1} = b$ and if $x_i \leq x < x_{i+1}$ then $f(x) = l_i$. Furthermore, due to the normalization property of pdfs, $\sum_{k=0}^{N} l_k(x_{k+1} - x_k) = 1$. The matrix M has 2 columns and $N + 2$ rows, and the entries of its first column are the N discontinuity points of f, plus the two endpoints of I (a and b).

Due to the deterministic behavior of the Frobenius-Perron operator, the following proposition can be proven.

Proposition 3 *Let f be piecewise constant, and let N be its number of discontinuities in I. The pdf $\Theta_\tau f$ has at most $N + 2(q-1)$ discontinuities, where q is the number of elements of the defining partition Q.*

Using the Frobenius-Perron operator, one can obtain a sequence of pdfs using the rule $f_{n+1} = \Theta_\tau f_n = \Theta_\tau^{n+1} f_0$. The proposition 2 assures that if the starting pdf f_0 is piecewise constant, such sequence can be unambiguously described by a sequence of matrices $\{M_n, n \in \mathbb{N}\}$. In general, if $N(n)$ is the number of discontinuities of f_n, the number of rows of the matrix M_n is equal to $N(n)+2$. As a direct consequence of the previous proposition, we have the following

Corollary 1 *Let $x_0(n), \ldots, x_{N(n)+1}(n)$ be the entries of the first column of M_n, and let $\{p_i, 0 \leq i \leq q\}$ be the endpoints of Q. The entries of the first column of M_{n+1} are the elements of the set $\tau(\{x_0(n), \ldots, x_{M(n)+1}(n)\}) \cup \tau_-(\{p_i, 0 \leq i \leq q\}) \cup \tau_+(\{p_i, 0 \leq i \leq q\})$.*

We conclude this subsection describing the explicit method for computing the entries of the second column of the matrix M_{n+1}, once we assume to know its first column and the matrix M_n. Accordingly, being $Q = \{I_1, \ldots, I_q\}$ the defining partition of τ, we define $\tau(Q)$ to be the collection of intervals $\{\tau(I_1), \ldots, \tau(I_q)\}$. If we denote with $x_j = M_{n+1}(j, 1)$ and $x_{j+1} = M_{n+1}(j+1, 1)$ the j-th and $j + 1$-th entries of the first column of M_{n+1}, we have to compute the level $l_j = M_{n+1}(j, 2)$, i.e., the j-th entry of the second column of M_{n+1} (we recall that by construction the last entry of the second column of the matrix M_n is zero, $\forall n \in \mathbb{N}$). Since l_j represents the level of the piecewise constant pdf described by M_{n+1} for any $x_j < x < x_{j+1}$, it suffices to use (4) referring to the point $x = \frac{x_j + x_{j+1}}{2}$, as described by the algorithm in Fig. 1.

In this algorithm, for each interval of $\tau(Q)$ the instruction A.3 checks whether the point x belongs to it or not. If yes, instruction A.4 computes the point $y \in I_i$ such that $\tau(y) = x$. Referring to the partition of I having the first column of M_n as endpoints, the instructions A.5-A.8 find the unique interval of such partition containing y, increasing the value of l_j according to (4).

REMARK: It is worth noting that due to the analytical structure of both the PWAE map and the piecewise constant pdf f_n, the computation of $f_{n+1} = \Theta_\tau f_n$ can be obtained performing extremely simple calculation on the entries of matrices with just two columns.

3.2 The iterative procedure

The idea is to study the evolution of the uniform distribution for computing an estimation of the invariant density f^*. Accordingly, we note that if we consider f_0 as the uniform distribution over I, described by the matrix

$$H_0 = \begin{pmatrix} x_0(0) = a & l_0(0) = \frac{1}{b-a} \\ x_1(0) = b & l_1(0) = 0 \end{pmatrix}, \tag{18}$$

ALGORITHM A – Computation of $M_{n+1}(j,2)$

```
1.  x = (x_j + x_{j+1})/2; l_j = 0;
2.  for i = 1 → q
3.     if x ∈ τ(I_i) then
4.        y = (x − β_i)/α_i;
5.        for k = 0 → N(n) − 1
6.           if M_n(k,1) < y < M_n(k+1,1) then
7.              l_j = l_j/|α_i| + M_n(k,2);
8.              break loop;
9.           end if;
10.       end for;
11.    end if;
12. end for;
```

Figure 1: The basic algorithm for computing the j-th entry of the second column of the matrix M_{n+1}.

PROCEDURE 1 – Invariant Measure Estimation

1. set $n = 1$ and M_0 as in (18) (the uniform pdf);
2. compute the entries of the first column of M_1 by sorting the elements of the set $\{x_0(1), \ldots, x_{N(1)+1}(1)\} = \tau_-(\{p_i, 0 \le i \le q\}) \cup \tau_+(\{p_i, 0 \le i \le q\})$;
3. compute the levels $l_0(n), \ldots, l_{N(n)}(n)$ according to the algorithm A;
4. compute E_n and if $E_n \le \varepsilon$ then `STOP`;
5. compute the entries of the first columns of M_{n+1} by sorting the elements of the set $\tau(\{x_0(n), \ldots, x_{N(n)+1}(n)\}) \cup \tau_-(\{p_i, 0 \le i \le q\}) \cup \tau_+(\{p_i, 0 \le i \le q\})$;
6. set $n = n + 1$ and `REPEAT` from step 3);

Figure 2: The procedure for the estimation of the invariant pdf f^*.

directly from Proposition 3 we have the following

Corollary 2 *Let f_0 be uniform. The matrix M_n has at most $2+2n(q-1)$ rows.*

The stochastic stability (16) of the system assures that the evolution of the uniform distribution $\{\Theta_\tau^n f_0\}$ is a converging sequence of densities in $D_{BV}(I)$. Since a converging sequence is always a Cauchy sequence, if we define

$$E_n = \operatorname*{ess\,sup}_{x \in I} \left| \Theta_\tau^n f - \Theta_\tau^{n-1} f \right| = \left\| f_n - f_{n-1} \right\|_\infty , \tag{19}$$

we have that $\lim_{n\to\infty} E_n = 0$[7].

Since the invariant density f^* is unknown a priori, regardless of the considered norm $\|\cdot\|_\star$ we can not directly estimate the quantity $\|\Theta_\tau^n f - f^*\|_\star$. Accordingly, we refer to the Cauchy quantity E_n for determining a criterion to stop

[7] Referring to the inequality (16), it is easy to show that $E_n \le K r^n \|f_0\|_{BV} = K r^n$.

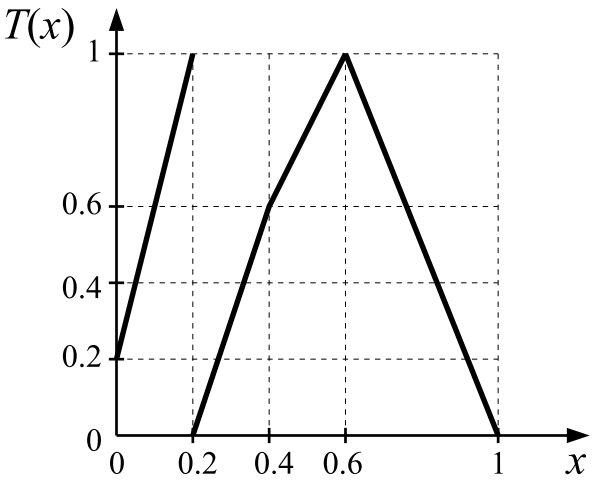


Figure 3: The Markovian chaotic map (20).

the procedure described in Fig. 2[8].

Each step of this algorithm requires basic low-complexity computations, and in step 4) the quantity ε is the accepted threshold for E_n, introduced for stopping the algorithm. The number of rows of the involved matrices grows linearly with n, whereas the convergence rate is assured from (16) to be at least exponential.

Exploiting the proposed algorithm a piecewise constant pdf $\tilde{f}^*$ can be computed for approximating f^*, with an accuracy that can be theoretically arbitrary. As a last remark, we stress that the first column of the final matrix $\tilde{M}^*$ describing $\tilde{f}^*$ are the endpoints of a refinement partition of Q.

3.3 Example

In this subsection we refer to an example PWAE map for evaluating the operation of Procedure 1, whereas in Section 7 the same system is taken into account for comparing the iterative procedure with the Ulam's method.

Accordingly, let us consider the Markovian chaotic transformation $T : [0, 1) \to [0, 1)$ defined as

$$T(x) = \begin{cases} 4x + \frac{1}{5}, & \text{if } x \in [0, \frac{1}{5}), \\ 3x - \frac{3}{5}, & \text{if } x \in [\frac{1}{5}, \frac{2}{5}), \\ 2x - \frac{1}{5}, & \text{if } x \in [\frac{2}{5}, \frac{3}{5}), \\ \frac{5}{2} - \frac{5}{2}x, & \text{if } x \in [\frac{3}{5}, 1). \end{cases} \tag{20}$$

The plot of this map is depicted in Fig. 3. For this special kind of map we can compute analytically its unique invariant pdf,

[8]The same criterion is used in the other numerical methods proposed in literature, as discussed in Section 7.

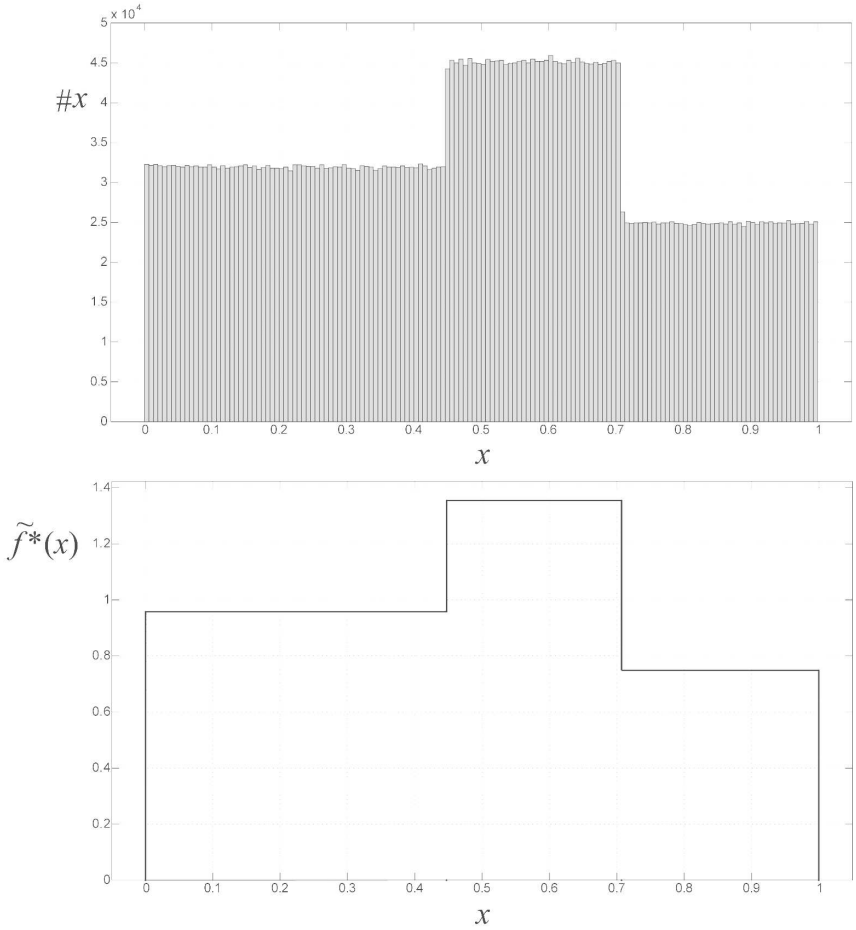


Figure 4: The unique invariant pdf f^* for the system (20).

$$f^*(x) = \begin{cases} \frac{60}{77}, & \text{if } x \in [0, \frac{1}{5}), \\ \frac{75}{77}, & \text{if } x \in [\frac{1}{5}, \frac{2}{5}) \cup [\frac{2}{5}, \frac{3}{5}), \\ \frac{175}{154}, & \text{if } x \in [\frac{3}{5}, 1), \end{cases} \tag{21}$$

which is the piecewise constant function[9] depicted in Fig. 4. It is worth noting that, since the invariant measure is known a priori, the choice of this map allows us for evaluating the accuracy of the estimation.

By setting in the procedure of Fig. 2 a value for ε equal to 10^{-4}, the invariant pdf f^* was estimated in only 10 steps with an estimation error lower than 10^{-8} in the L^1 norm, as depicted in Fig. 5. According to the logarithmic scale used for representing the amplitudes in the plot, straight lines denote exponential curves. When dealing with PWAE mixing maps, the procedure operation shown

[9]The levels of the function f^* are related to the invariant probability vector of the 4-states Markov chain associated to the system [9].

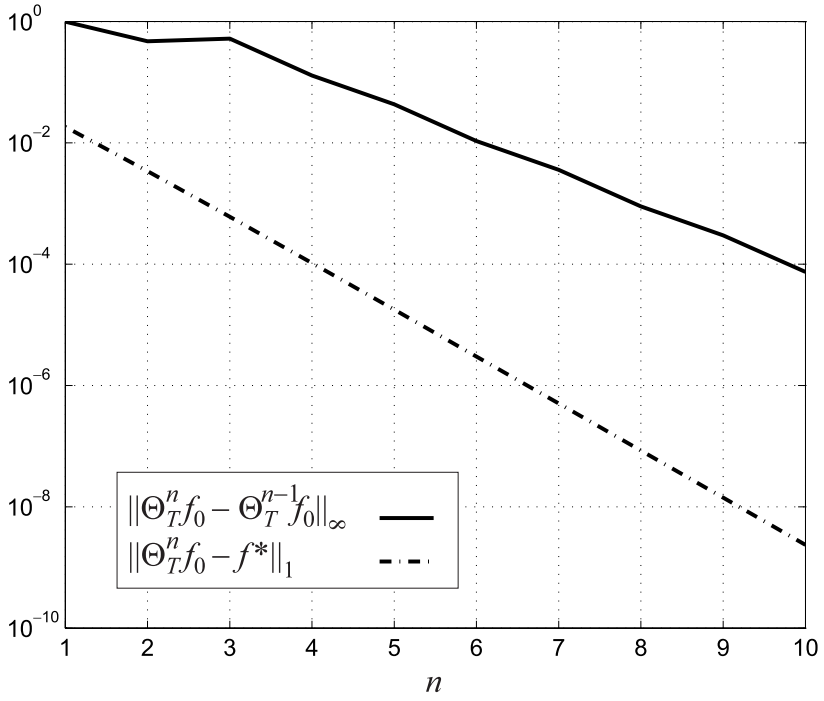


Figure 5: The Cauchy sequence $E_n = \|f_n - f_{n-1}\|_\infty$ and the absolute estimation error $\|\Theta_T^n f_0 - f^*\|_1$ obtained in the example of subsection 3.3. In the logarithmic scale used for representing the amplitudes, straight lines denote exponential curves. The estimated pdf $\tilde{f}^*$ is graphically indistinguishable form that one of Fig. 4, since the resulting error $\left\|\tilde{f}^* - f^*\right\|_1$ is lower than 10^{-8}.

in this example is typical. Obviously, when considering PWAE mixing maps with nontrivial parameter values the invariant density f^* is unknown, and the absolute estimation error can only be bounded according to the inequality (16). As discussed in the next Section, two conservative values for the constant K and r in (16) can be expressed as a function of the analytical structure of the PWAE map.

4 Speed of Convergence and Accuracy of the Estimation

In this section we discuss the accuracy and the speed of convergence of the proposed method. In particular, we discuss how the halting threshold ϵ determines the accuracy of the estimation of f^*, and we provide a theoretical method for determining *a priori* a sufficient number of iterations that guarantees a given estimation accuracy.

The problem of estimating the invariant density f^* is equivalent to finding

an approximated solution for the functional equation

$$\Theta_\tau f^* = f^*, \qquad f^* \in D_{BV}(I). \tag{22}$$

In general, if $f \in D_{BV}(I)$ we can write $\Theta_\tau f = f + e$, where $e \in BV(I)$, and $E = \|\Theta_\tau f - f\|_\infty = \|e\|_\infty$. Consequently, the halting threshold ϵ in the algorithm represents a direct measure of the accuracy according to which the approximated solution of the functional equation (22) is calculated. If n is the smallest integer such that $E_n < \epsilon$ then, by denoting $g = \Theta_\tau^{n-1} f_0$, we have

$$\Theta_\tau g = g + e, \qquad \text{with } \|e\|_\infty < \epsilon. \tag{23}$$

Depending on the accuracy of the required application, if ϵ is chosen 'sufficiently small' one can assume $\Theta_\tau g \approx g$. This fundamental aspect is analyzed more in detail in the next subsection.

We conclude this part with a further remark on the theoretical meaning of the sequence E_n. Indeed, we can write $f_n = \Theta_\tau^n(f_0) = f^* + e_n$, where $e_n \in BV(I)$ for all $n \in \mathbb{N}$ and the sequence $\{e_n\}$ is such that

$$\lim_{n\to\infty} \|e_n\|_\star = 0. \tag{24}$$

Accordingly, if $E_n = \|f_n - f_{n-1}\|_\infty$, due to the linearity of the Frobenius-Perron operator, we have $E_n = \|\Theta_\tau e_{n-1} - e_{n-1}\|_\infty$.

4.1 The constants K and r in (16)

The inequality (16) that holds for PWAE mixing maps is related to the general spectral properties of the Frobenius-Perron operator for mixing piecewise expanding maps [9]. Indeed, for these systems the Frobenius-Perron operator has an unique eigenvalue with modulus 1 (to which correspond the eigenfunction $f^* = \Theta f^*$), whereas the remaining eigenvalues have modulus strictly lower than 1.

There is an infinite number of (K, r) values that make the inequality (16) true [4, 19, 25]. In detail, the constant r is related to the *rate of mixing* (or to the *rate of decay of correlations*) that characterizes the dynamical system, or rather it is related to the gap between the eigenvalue 1 and the remaining part of the spectrum of the Frobenius-Perron operator Θ_τ (see, e.g., [4, 19]).

In particular, by denoting with $\sigma(\Theta_\tau)$ the spectrum of Θ_τ, the inequality (16) holds for any $r > r_0$ such that

$$r_0 = \sup\big\{|\lambda| : \lambda \in \sigma(\Theta_\tau) \setminus \{1\}\big\}, \tag{25}$$

and for a suitable constant K depending on r.

Unfortunately, in most cases an accurate estimation of the constant r_0 can only be obtained resorting to numerical calculations based on the Ulam's method [13].

Liverani in [24, 25] investigates the decay of correlations for a large class of non-Markov one-dimensional expanding maps, by directly studying the properties of the Frobenius-Perron operator. As a result, Liverani provides a method that can be used for calculating two sub-optimal constants K and r that work in (16). Even if in some cases the Liverani's approach can provide for K and

r rather conservative values [13, 24, 25], nonetheless these values are significant, since their expressions depend *directly on the analytical structure of the map* τ. In the following theorem, in which the result of Liverani is fit on our case, we report the main contribution of this section. The proof of the theorem is not reported, since a huge amount of notation and background would be indispensable to frame the theory of Liverani, whereas on the other hand the proof is obtained by rearranging the work of the same author presented in [25].

Theorem 1 *Let us consider* $\tau \in \mathcal{T}_{aex}(I)$ *with defining partition* $Q = \{I_j : j = 1, \dots, q\}$, *and define* $\alpha_{\min} = \inf_{x \in I} |\tau'(x)|$ *and* $\alpha_{\max} = \sup_{x \in I} |\tau'(x)|$. *Let us assume moreover that*[10] $\alpha_{\min} > 2$. *We define*

- $A = \max\{\frac{2}{|\alpha_j|\lambda(I_j)}, j = 1, \dots, q\}$;
- $a > a_0 = \dfrac{A\lambda(I)}{1 - \frac{2}{\alpha_{\min}}}$;
- $k_0 \in \mathbb{N}$ *as the minimum integer* k *such that all the intervals of the partition* $Q_k = \bigvee_{j=0}^{k} \tau^{-j}(Q)$ *have length smaller than* $\lambda(I)/2a$;
- $P_0 \in \mathbb{N}$ *as the minimum integer* n *such that* $\tau^n(J) = I$ *for all* $J \in Q_{k_0}$, *i.e., the map* τ^n *is* covering *on* Q_{k_0};
- $\delta = (\frac{2}{\alpha_{\min}})^{P_0} + \left(1 - (\frac{2}{\alpha_{\min}})^{P_0}\right)\frac{a_0}{a}$;
- $\Delta = 2 \ln \dfrac{\max\{1 + \delta; 1 + a\delta\}}{\min\{1 - \delta; \frac{1}{2}\alpha_{\max}^{-P_0}\}}$;
- $\Lambda = \tanh \frac{\Delta}{4}$.

Then, if we denote with $f_0 \in D_{BV}(I)$ *the uniform pdf over* I *it results* $\|\Theta_\tau^n f_0 - f^*\|_\infty \leq K r^n$, *where*

$$K = \frac{(1 + a_0)\Delta e^{\frac{\Delta}{\Lambda}}}{\lambda(I)\Lambda}, \qquad \textit{and} \qquad r = \Lambda^{\frac{1}{P_0}}. \tag{26}$$

Among the above quantities, the integers k_0, P_0 can result the most complex to calculate, since they depend on the arbitrary shape of the map τ. As an example, in [25] the author discusses in detail how to compute the constants defined above for a specific PWAE mixing map.

Even if for a given map τ one can use Theor. 1 for determining *a priori* a definite (and conservative) number n_0 of iterations in the Procedure 1 such to satisfy a certain absolute error associated to the approximating pdf $\tilde{f}^*$, we stress that n_0 can exhibit an high sensitivity to parameter perturbations. Accordingly, from a practical point of view, when using the Procedure 1 in huge Monte Carlo simulations (e.g., for determining the dependence of some statistical features of a chaotic circuit with respect to certain implementation tolerances), it is more convenient to use the simple criterion based on the Cauchy error sequence E_n given by (19), that assures an approximated solution for the functional equation $\Theta_\tau g = g$, as discussed in the previous subsection[11].

[10] If $1 < \alpha_{\min} \leq 2$ then for an expanding map we can focus the theorem on a properly chosen iterate of τ such that $2 < \inf_{x \in I} |\frac{d\tau^r}{dx}(x)|$.

[11] See, for an application example, [1].

5 Lyapunov Exponent Estimation

In this section we assume $\tilde{f}^* = \Theta_\tau^{n_0} f_0$, where f_0 is the uniform pdf over I, n_0 is a suitable positive integer and $\tau \in \mathcal{T}_{aex}(I)$ is a mixing map. If we consider $\tilde{\mu} = \tilde{f}^* \cdot \lambda$ as an approximation for the invariant measure $\mu^* = f^* \cdot \lambda$, we expect that the value of the Lyapunov Exponent associated to τ with respect to μ^* can be approximated as

$$L_\tau(\mu^*) \approx \int_I \ln|\tau'(x)|\, \tilde{f}^*(x)dx. \tag{27}$$

Indeed, by defining $|\alpha_{\max}| = \sup_{x\in I} |\tau'(x)|$, from (16) we have

$$\begin{aligned} |L_\tau(\mu^*) - L_\tau(\tilde{\mu}^*)| \leq \\ \leq \int_I \left|\ln|\tau'(x)|\,(f^*(x) - \tilde{f}^*(x))\right| dx \leq \\ \leq \left\|f^* - \tilde{f}^*\right\|_1 \ln|\alpha_{\max}| \leq K r^{n_0} \ln|\alpha_{\max}|. \end{aligned}$$

Recalling that the pdf $\tilde{f}^*$ is piecewise constant and that its analytical structure can be described by a matrix $\tilde{M} = M_{n_0}$ with $\tilde{N}+1 = N(n_0)+1$ rows and two columns, we simply have

$$\begin{aligned} L_\tau(\mu^*) \approx \\ \approx \sum_{i=1}^{q} \ln|\alpha_i| \sum_{j=j_1(i)}^{j_2(i)-1} \tilde{M}(j,2)(\tilde{M}(j+1,1) - \tilde{M}(j,1)), \end{aligned} \tag{28}$$

where $j_1(i) < j_2(i)$ are the indexes such that $\tilde{M}(j_1,1) = p_{i-1}$ and $\tilde{M}(j_2,1) = p_i$, being $\mathcal{E}_Q = \{p_0,\ldots,p_q\}$ the endpoints of the defining partition Q of τ[12].

6 Shannon and Kolmogorov-Sinai Entropy Estimation

Let us consider a mixing map $\tau \in \mathcal{T}_{aex}(I)$ with defining partition $Q = \{I_1,\ldots,I_q\}$. We assume that the partition Q is used for defining a symbolic dynamics $\mathbf{S} = \{\sigma_Q \circ \tau^n(x_0), n \in \mathbb{N}\}$, where $\sigma_Q : I \to \mathcal{M} = \{1,\ldots,q\}$ and we assume the initial condition x_0 is a random variable distributed according to the unique invariant pdf f^* of τ[13]. In this section we discuss the link that there exists between the classical Shannon entropy for such a symbolic source and the Kolmogorov-Sinai entropy defined in Sec. 2.2, and we provide an efficient method for their estimation.

6.1 The Symbolic Source

For a random symbolic source we write $P(s_n = i)$ to denote the probability of the source $\mathbf{S}$ to issue the symbol $i \in \mathcal{M}$ at the time-step n. When considering

[12] We recall that the entries in the first column of $\tilde{M}$ are the endpoints of a refinement partition of Q.

[13] Accordingly, the obtained symbolic source is ruled by an ergodic stochastic process.

the jointly probability $P(s_{n_1} = i_1, s_{n_2} = i_2, \ldots, s_{n_k} = i_k)$ it is convenient to adopt the notation $P\left(\bigwedge_{j=1}^{n} s_{n_j} = i_j\right)$. It results that

$$s_n = \sigma(x_n = \tau^n(x_0)) = i \Leftrightarrow x_0 \in \tau^{-n}(I_i), \tag{29}$$

and in general, given a n-tuple of integers $(k_0, k_1, \ldots, k_{n-1})$ and a finite n-tuple of symbols $(i_0, i_1, \ldots, i_{n-1}) \in \mathcal{M}^n$

$$\left(\bigwedge_{j=0}^{n-1} s_{k_j} = i_j\right) \Leftrightarrow x_0 \in \bigcap_{j=0}^{n-1} \tau^{-k_j}(I_{i_j}). \tag{30}$$

Since we are assuming that the mixing system is stabilized on its invariant probability measure $\mu^* = f^* \cdot \lambda$, if $A \in \mathcal{B}_I$, the quantity $\mu^*(A)$ represents the probability for the state x to belong to the set A at any time-step[14]. To completely specify the obtained ergodic information source it is sufficient to consider ordered n-tuples of integers $(k_0, k_2, \ldots, k_{n-1})$ as sequence time-step references, assuming $k_0 = 0$ and $k_i < k_j$ if $i < j$.

Accordingly, the probability for the information source to generate a symbolic sequence with $(\wedge_{j=0}^{n-1} s_{k_j} = i_j)$ is equal to

$$P\left(x_0 \in \bigcap_{j=0}^{n-1} \tau^{-k_j}(I_{i_j})\right) = \mu^*\left(\bigcap_{j=0}^{n-1} \tau^{-k_j}(I_{i_j})\right). \tag{31}$$

Expression (31) allows for the computation of any order statistics and specifies completely the information source defined by partitioning the phase state of the ergodic dynamical system τ with any symbolic partition.

6.2 Efficient Estimation of Any Order Statistics of the Symbolic Source

In the following, if $w_k = (i_0, i_2, \ldots, i_{k-1}) \in \mathcal{M}^k$, we write $P(w_k) = P\left(\bigwedge_{j=0}^{k-1} s_j = i_j\right)$. We recall from (31) that the probability $P(w_k)$ is equal to

$$P(w_k) = \int_{\cap_{j=0}^{k-1} \tau^{-j}(I_{i_j})} f^*(x)dx. \tag{32}$$

Therefore, the computation of the above quantity requires the calculation of the set $\cap_{j=0}^{k-1} \tau^{-j}(I_{i_j})$, which can be a nontrivial task even when k is small. Alternatively, when considering PWAE mixing maps, we propose a method which is based on both the properties of the Frobenius-Perron operator and the analytical structure of the piecewise constant pdf $\tilde{f}^*$ that we assume as a good approximation of the invariant pdf f^*.

We first prove a general result which is valid for any mixing map $\tau : I \to I$. In detail, given an arbitrary symbolic partition $R = \{J_1, \ldots, J_m\} \in \mathcal{P}_I$, we focus on the symbolic dynamics $W_{\tau,R} : I \to \mathcal{M}^\infty$.

[14] It is worth recalling that the resulting process is ergodic and hence it is strongly stationary.

Proposition 4 *Let us consider a symbolic sequence* $\{i_n, n \in \mathbb{N}\} \in \mathcal{M}^\infty$ *and let* $f^* \in D_1(I)$ *be the invariant pdf for* τ*. We define the sequence of functions* $\{g_n \in L^1(I), n \in \mathbb{N}\}$ *as*

$$\begin{aligned} g_0 &= f^* \chi_{J_{s_0}}, \\ g_{n+1} &= \Theta_\tau g_n \cdot \chi_{J_{i_{n+1}}}, \qquad n \in \mathbb{N}. \end{aligned} \tag{33}$$

It results that

$$P(w_k) = \int_I g_{k-1}(x)dx, \qquad k \in \mathbb{N}. \tag{34}$$

Proof. For notation convenience we write $J_{i_n} = E_n$. For any set $A_k \in \mathcal{B}_I$ we define the finite sequence $\{A_0, A_1, \ldots, A_k\}$ such that for $n = 1 \ldots, k$ it results $A_{n-1} = \tau^{-1}(A_n) \cap \tau^{-1}(E_n)$. We want to show that

$$A_0 = \cap_{j=1}^k \tau^{-j}(E_j) \cap \tau^{-k}(A_k). \tag{35}$$

Accordingly, we first note that by construction $A_{k-1} = \tau^{-1}(A_k) \cap \tau^{-1}(E_k)$. Proceeding by induction on $p = 1, 2, \ldots$ we assume that

$$A_{k-p} = \cap_{j=k-p+1}^k \tau^{-p-j+k}(E_j) \cap \tau^{-p}(A_k). \tag{36}$$

Hence, it results $A_{k-p-1} = \tau^{-1}(A_{k-p}) \cap \tau^{-1}(E_{k-p}) = \cap_{j=k-p}^k \tau^{(-p-1-j+k)}(E_j) \cap \tau^{-p-1}(A_k)$. Therefore, if we set $p = k$ in (36) we obtain (35).

Let us now note that for any set $A_k \in \mathcal{B}_I$ it results

$$\begin{aligned} \int_{A_k} g_k(x)dx &= \int_{A_k \cap E_k} \Theta_\tau g_{k-1}(x)dx = \\ = \int_{\tau^{-1}(A_k \cap E_k)} g_{k-1}(x)dx &= \int_{A_{k-1}} g_{k-1}(x)dx = \ldots \\ = \int_{A_0} g_0(x)dx &= \int_{\cap_{j=1}^k \tau^{-j}(E_j) \cap \tau^{-k}(A_k)} f^*(x)\chi_{E_0}dx = \\ &= \int_{\cap_{j=0}^k \tau^{-j}(E_j) \cap \tau^{-k}(A_k)} f^*(x)dx. \end{aligned}$$

Since $\tau^{-k}(I) = I$, if in the above equations we set $A_k = I$ we obtain

$$\int_I g_k(x)dx = \int_{\cap_{j=0}^k \tau^{-j}(E_j)} f^*(x)dx = P(w_{k+1}), \tag{37}$$

and the proof is completed. □

On the basis of Proposition 4, the quantity $P(w_k)$ can be easily computed for any k once g_{k-1} is known. Nevertheless, when considering PWAE maps the estimation of g_{k-1} can be iteratively computed in $k-1$ steps applying a modified version of the Procedure 1 to a suitable estimation $\tilde{f}^*$ of f^*. Indeed, the principles that we used for the estimation of the invariant pdf f^* are still valid, noting that if g_n is piecewise constant then $\Theta_\tau g_n \chi_{I_j}$ is still piecewise constant for any element I_j of the partition Q.

Accordingly, the estimation of $P(w_k)$ can be efficiently performed using the procedure of Fig. 6, where we denote with $\tilde{M}^*$ the matrix describing $\tilde{f}^*$ obtained by means of the Procedure 1.

PROCEDURE 2 – Estimation of $P(w_k)$

1. set $n = 0$ and $M_0 = \tilde{M}^*$;
2. for $i \in \{0, \ldots, N(n) - 1\}$ if $(M_n(i, 1), M_n(i + 1, 1)) \cap I_{s_n} = \oslash$ then set $M_n(i, 2) = 0$;
3. if $n = k - 1$ then `SKIP` to step 7);
4. compute the entries of the first columns of M_{n+1} by sorting the elements of the set $\tau(\{x_0(n), \ldots, x_{N(n)+1}(n)\}) \cup \tau_-(\{p_i, 0 \leq i \leq q\}) \cup \tau_+(\{p_i, 0 \leq i \leq q\})$;
5. compute the levels $l_0(n), \ldots, l_{N(n)}(n)$ according to the algorithm A;
6. set $n = n + 1$ and `REPEAT` from step 2);
7. set $P(w_k) = \sum_{i=0}^{N(k-1)-1} [M_n(i + 1, 1) - M_n(i, 1)] M_n(i, 2)$;

Figure 6: The procedure for the estimation of the generation probability $P(w_k) = (s_0, s_1, \ldots, s_{k-1}) \in \mathcal{M}^k$.

The above approach can be easily generalized for the computation of the generation probability of any k-tuple with form $\{s_0 = i_0, s_{j_1} = i_1, \ldots, s_{j_{k-1}} = i_{k-1}\}$ in j_{k-1} iteration of the Procedure 2, assuming to properly modify the step 2). Indeed, e.g., the 3-plet $\{s_0 = i_0, s_3 = i_1, s_4 = i_2\}$ can be though of as the special sequence

$$\{s_0 = i_0, s_1 = \sharp, s_2 = \sharp, s_3 = i_1, s_4 = i_2\}, \tag{38}$$

where the notation $\sharp$ means 'any symbol'. In such case, when encountering in the sequence the special symbol $\sharp$ the step 2) in the Procedure 2 can be skipped, since the subset of the domain associated to $\sharp$ is the whole domain I itself.

As a last remark, it can be noticed that if in the operation of the procedure 2 a time step n exists such that after the instruction 2) the matrix M_n has all the entries in the second column equal to zero, then any matrix M_j, with $j > n$, will have the entries in the second column equal to zero. Accordingly, when such event occurs, the procedure can be aborted setting $P(w_k) = 0$ and avoiding useless further calculations.

6.3 The Average Shannon Entropy

The Shannon entropy quantifies the degree of unpredictability of an information source.

Definition 6 *The* Average Shannon Entropy *(ASE) of the source $\boldsymbol{S}$ is equal to*

$$ASE(\boldsymbol{S}) = \lim_{k \to \infty} -\frac{1}{k} \sum_{\mathcal{M}^k} P(w_k) \cdot \log_2 P(w_k), \tag{39}$$

where the summation extends to the finite set collecting all k-tuples $(i_0, \ldots, i_{k-1}) \in \mathcal{M}^k$, and if $P(w_k) = 0$ then $P(w_k) \log_2 P(w_k)$ is assumed zero. Since we are using logarithms with base 2, the result is expressed in bits.

The ASE indicates, for a given source, the average amount of information issued at each time-step. For a chaotic symbolic source, the ASE limit is approached from the above [12]. All the quantities in the expression can be computed by using the Procedures 1 and 2. In the next subsection we will prove that for symbolic sources based on PWAE maps, the estimation of the Kolmogorov-Sinai entropy (14) can be directly obtained by estimating (39).

6.4 Estimation of the Kolmogorov-Sinai Entropy

We first highlight an inportant result about the defining partition of a PWAE map.

Proposition 5 *Let $Q = \{I_1, \ldots, I_q\}$ be the defining partition of a PWAE map τ, and let assume that*

$$x \in I_i \Leftrightarrow \sigma_Q(x) = i, \qquad \text{for } i = 1, \ldots, q. \tag{40}$$

Then, the partition Q is a generating partition.

Proof. We have to show that the function $W_{\tau,Q} : I \to \mathcal{M}^\infty$ is injective. Accordingly, let us consider $x_1 \neq x_2 \in I$. If they belong to different intervals $I_i \neq I_j$ then $\sigma_Q(x_1) \neq \sigma_Q(x_2)$ and the two sequences $W_{\tau,Q}(x_1)$ and $W_{\tau,Q}(x_2)$ are different. On the other hand, let $x_1, x_2 \in I_i$, for a certain $I_i \in Q$, and let us consider the open interval $A_0 = (x_1, x_2)$. Since τ is linearly expanding over I_i then the Lebesgue measure of the open interval $A_1 = \tau(A_0)$ is bounded from below by $\lambda(A_1) \geq \alpha_{\min}\lambda(A_0) > \lambda(A_0)$, where $\alpha_{\min} = \inf_{x\in I} |\tau'(x)|$. If ab-absurdo $W_{\tau,Q}(x_1) = W_{\tau,Q}(x_2)$ then for any $n \in \mathbb{N}$ both $\tau^n(x_1)$ and $\tau^n(x_2)$ belong to the same element of the partition Q, and the set $A_n = \tau^n(A_0) \subseteq I$ is an open interval whose Lebesgue measure is bounded from below by $\lambda(A_n) \geq \alpha^n_{\min}\lambda(A_0)$. This result is impossible since it implies that exists $m \in \mathbb{N}$, with $m > \frac{\ln \lambda(I) - \ln \lambda(A_0)}{\alpha_{\min}}$ such that $\lambda(\tau^m(A_0)) > \lambda(I)$. □

Generating partitions are important since in 1959 Y. Sinai proved the following theorem [29].

Theorem 2 *Let $\tau \in \mathcal{T}_{aex}(I)$, and let the a.c. measure μ be invariant for τ. If R is a generating partition then $H_{KS}(\tau, \mu) = H_R(\tau, \mu)$.*

On these bases, we can now prove the following

Proposition 6 *The ASE for the ergodic information source obtained by partitioning with the defining partition Q the phase space I of a mixing PWAE map is equal to the Kolmogorov-Sinai entropy $H_{KS}(\tau, \mu^*)$.*

Proof. We know from the proposition 5 that the partition Q is generating, and from Theor. 2 it results that the Kolmogorov-Sinai entropy is equal to the measure-theoretic entropy $H_Q(\tau, \mu^*)$. Accordingly, we have to show that the expression (39) and $H_Q(\tau, \mu^*)$ given by (13) represent the same quantity. Indeed, the elements of the partition

$$R_n = \bigvee_{k=0}^{n-1} \tau^{-k}(Q) \tag{41}$$

are sets of the form $I_{i_0} \cap \tau^{-1}(I_{i_1}) \cap \cdots \cap \tau^{-n+1}(I_{i_{n-1}})$, where $(i_0, \ldots, i_{n-1}) \in \mathcal{M}^n$. From (31) we know that the probability measure of each of these sets is equal to the probability $P(w_k)$ of the event $\left(\bigwedge_{k=0}^{n-1} s_k = i_k\right)$. Accordingly, referring to the probability space $(I, \mathcal{B}_I, \mu^*)$, the entropy of the partition R_n is equal to

$$H(R_n) =$$
$$= -\sum_{\mathcal{M}^n} \mu^* \left(\cap_{k=0}^{n-1} \tau^{-k}(I_{i_k})\right) \log_2 \mu^* \left(\cap_{k=0}^{n-1} \tau^{-k}(I_{i_k})\right) =$$
$$= -\sum_{\mathcal{M}^n} P(w_k) \log_2 P(w_k).$$

Form Def. (13) the quantity $H_Q(\tau, \mu^*)$ is equal to

$$H_Q(\tau, \mu^*) = \lim_{n\to\infty} \frac{1}{n} H(R_n) =$$
$$= \lim_{n\to\infty} -\frac{1}{n} \sum_{\mathcal{M}^n} P(w_k) \log_2 P(w_k),$$

which is the ASE defined in (39). □

Concluding, according to the Proposition 6, the Kolmogorov-Sinai entropy of a PWAE mixing map can be estimated as the ASE of the information source obtained by partitioning the phase space I with the defining partition Q.

REMARK: The estimation of the ASE typically involve almost insuperable numerical problems due to the exponential growth of the number of terms w_k in the summation in (39) as far as $n \to \infty$. For this reason, in practical applications the limiting sequence is truncated to a given maximum number of terms. Nevertheless, since in chaotic systems the asymptotic value is approached from above, if the partial ASE determined at the last term of the truncated sequence reveals a given level h of entropy, then for sure $ASE \leq h$. This fact can be useful for evaluating the effects of the non-idealities of chaos-based circuits on the statistical behavior of the information source. In the following we write ASE(n) for denoting the n-th term of the sequence in (39).

7 Comparison with the Ulam's method

In 1960 Ulam conjectured a method for estimating the invariant measure of a chaotic system based on the discrete approximation of the Frobenius-Perron operator [31]. The method, validated several years later by Li [23], is universally acknowledged as an advanced theoretical tool for the numerical analysis of mixing dynamical systems [8, 13–16, 27]. As stated in the introduction, in this section we compare our estimation method with the Ulam's method from a theoretical point of view, from an operative point of view and from an algorithmic complexity point of view.

7.1 Brief review of the Ulam's method

Let us consider a mixing map $\tau : I \to I$, where $I \subset \mathbb{R}$ is an interval with endpoints $a < b$. We define the sequence of partitions $R_n = \{J_{n,1}, \ldots, J_{n,n}\}$,

where for $k = 1, \ldots, n$ the set $J_{n,k}$ is a sub-interval of I with endpoints

$$a_{n,k} = a + \frac{b-a}{n}(k-1),$$

and

$$b_{n,k} = a + \frac{b-a}{n}k.$$

Accordingly, the partition R_n divides I in n equal subintervals. Let us define the n-states ergodic Markov chain $\mathbf{C}_n$ whose transition probability matrix $\mathbf{P}(n) \in \mathbb{R}^{n\times n}$ is given by

$$P_{i,j}(n) = \frac{\lambda(J_{n,i}) \cap \lambda(\tau^{-1}(J_{n,j}))}{\lambda(J_{n,i})}. \tag{42}$$

The quantity $P_{i,j}(n)$ agrees with the probability that the state of $\mathbf{C}_n$ at a given time-step is j, provided that at the previous time-step the state was i.

The invariant probability mass density for the Markov chain $\mathbf{C}_n$ is described by an *invariant row vector* $v_n^* \in \mathbb{R}^{+n}$, with $|v_n^*| = 1$, such that $v_n^*\mathbf{P}(n) = v_n^*$. On the basis of v^* one can define the piecewise constant pdf

$$\tilde{f}_n^*(x) = \sum_{1\leq i\leq n} \frac{l_{n,i}}{\lambda(J_{n,i})}\, \chi_{J_{n,i}}(x), \tag{43}$$

where $\{l_{n,1}, \ldots, l_{n,n}\}$ are the components of v_n^*. Li in [23] proved that

$$\lim_{n\to\infty} \left\| \tilde{f}_n^* - f^* \right\|_1 = 0, \tag{44}$$

where f^* is the invariant pdf for the dynamical system ruled by τ. The rate of convergence to the above limit was proved by Murray in [8] to be $O(\frac{\log n}{n})$. In general, since the invariant pdf f^* in unknown, the Ulam's method is iterated until $\left\| \tilde{f}_n^* - \tilde{f}_{n-1}^* \right\|_1 < \varepsilon$, for a suitable $\varepsilon > 0$ assumed 'sufficiently small' for the considered analysis [16]. It is worth underlining that the Ulam's method can be applied to the wide class of mixing maps, whereas the approach discussed in this paper is specific for the family of PWAE mixing maps.

7.2 Theoretical comparison

As stated in the work of Li [23], the Ulam's method is framed within the *theory of Markovian approximations* of mixing maps. According to this approach, a sequence of Markovian maps is taken into account such that the limit of their associated invariant pdfs converges in the norm $\|\cdot\|_1$ to the invariant pdf of the original transformation. As a result, the estimated invariant pdf obtained with the Ulam's method is a numerically computed fixed point for an approximated version of the original Frobenius-Perron operator. The higher is the resolution of the Markovian approximation, and the more accurate is the estimation.

On the other hand, our estimation method works with the exact Frobenius-Perron operator, and the estimated invariant pdf is obtained as a numerically computed fixed point for the same operator. The convergence of the method is theoretically assured by the theory presented in Sections 2 and 3. It is interesting noting that in our case the convergence is assured in any norm $\|\cdot\|_\star$.

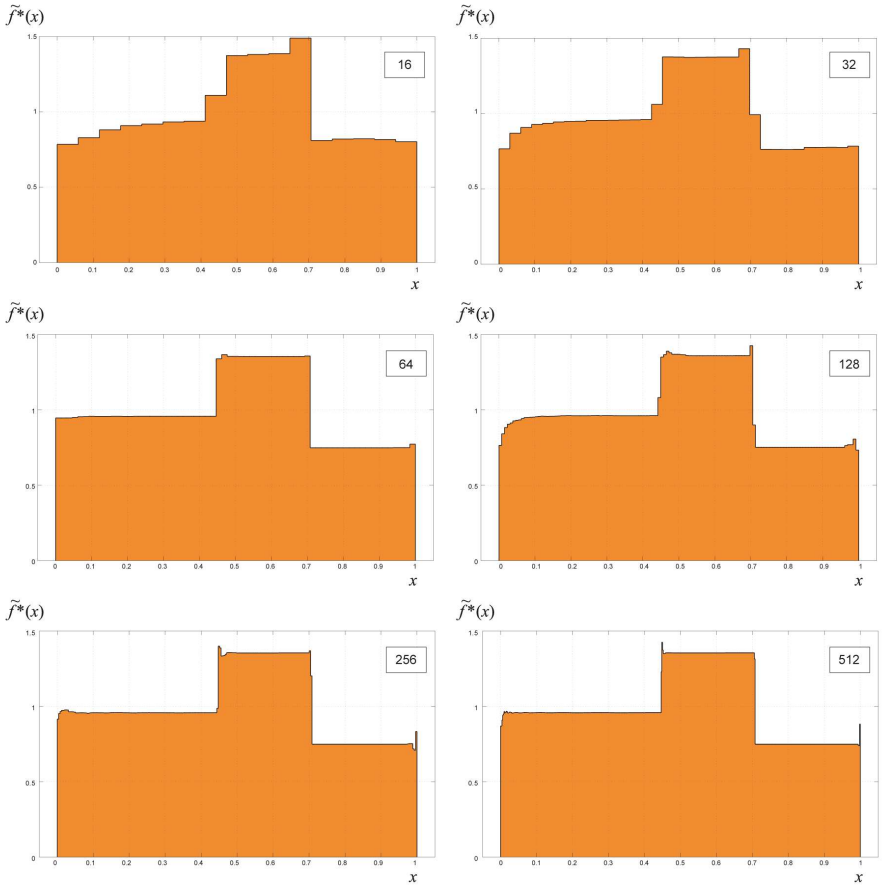


Figure 7: The estimated invariant pdf $\tilde{f}^*$ after 512 iterations of the Ulam's method. The circles highlight the main discontinuities of the pdf around which the estimation error is more relevant.

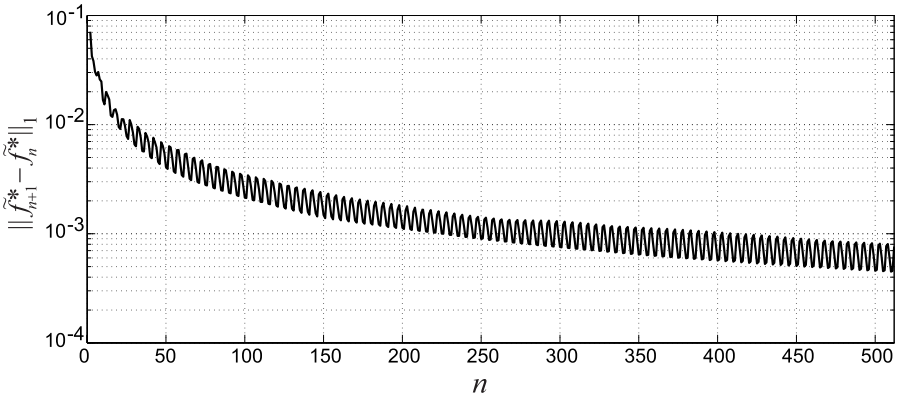


Figure 8: The Chaucy sequence $\left\|\tilde{f}^*_{n+1} - \tilde{f}^*_n\right\|_1$ for the first 512 steps of the Ulam's method.

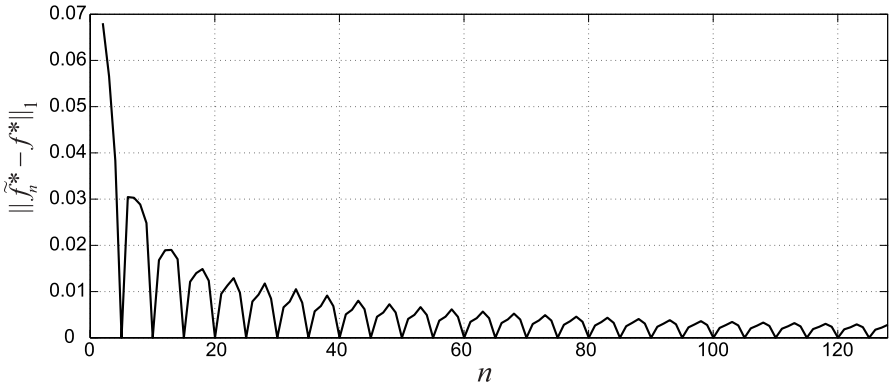


Figure 9: The absolute estimation error $\left\|\tilde{f}_n^* - f^*\right\|_1$ for the first 128 steps of the Ulam's method.

7.3 Operating comparison

For making an operating comparison between our method and the Ulam's method, we applied the latter one for estimating the invariant pdf f^* (21) of the Markovian map (20). Results are reported in Figures 7, 8 and 9.

The estimated invariant pdf depicted in Fig. 7 was obtained after 512 iterations of the procedure, and it can be compared with the exact invariant pdf reported in Fig. 4. Four circles were drawn to highlight the main discontinuities of the pdf around which the estimation error is more relevant. It is interesting noting in Fig. 8 that even if the method was iterated for more than 500 steps, the Chauchy quantity $\left\|\tilde{f}_n^* - \tilde{f}_{n-1}^*\right\|_1$ did not get lower than 10^{-4}, and the procedure was halted when the transition probability matrix was larger than a 512 by 512 square matrix. This latter figure has to be compared with Fig. 5, considering that $\|\cdot\|_\infty \geq \|\cdot\|_1$.

Since we knew a priori the unique invariant pdf f^* (21) for the map (20), we could calculate the absolute estimation error in the L^1 norm, as reported in Fig. 9. By inspecting this figure, it can be noticed that if n is an integer multiple of 5 then the absolute estimation error drops to zero: indeed, for the specific map considered in this example, for such values of n the Ulam's partition corresponds to a refinement of the defining Markov partition, and the quantity computed with (43) results to be the exact computation of the Markovian invariant pdf f^*. Unfortunately, a PWAE map with nontrivial parameters is either a non-Markovian map or it is a map with a high order Markovian defining partition [2], and the Ulam's method, according to a general approach, must be iterated until either the Chauchy quantity $\left\|\tilde{f}_n^* - \tilde{f}_{n-1}^*\right\|_1$ gets lower than a proper threshold, or the number of iterations exceed a predefined maximum value. Even if we knew the map (20) to be Markovian, we did not used this information since we were interested in showing the general operation of the Ulam's method for a PWAE map: indeed, the obtained overall behavior of the method is typical, and it is not strictly related to the special map considered in this example.

As a last remark, we want to stress that when a high accuracy in the estimation of the invariant pdf is required, the Ulam's method can result to be

susceptible to the numerical errors associated to the computation of the eigenvector v_n^* of the transition probability matrix $\mathbf{P}(n)$. This can happen when the matrix $\mathbf{P}(n)$ is close to be singular or badly scaled, causing the computation of v_n^* to be inaccurate, and according to our experience this event is not unfrequent.

7.4 Algorithmic complexity comparison

From an algorithmic complexity point of view, we can carry out a rough comparison between the two methods by taking into account the growing rate of the computing data size to be processed as a function of the estimation accuracy. According to what discussed in Section 3, at each iteration of our estimation procedure the size of the computing data grows linearly (i.e., the size of the matrix M_n), whereas the converging rate of the method is not slower than exponential of the form r^n, with $0 < r < 1$. Accordingly, for a given estimation error ρ the required number of iterations for our method is in the order of $O(\frac{\log \rho}{\log r})$, that implies, due to the linear growth of the computing data, an algorithmic complexity in the order of $O(\frac{\log \rho}{\log r})$.

On the other hand, for the Ulam's method the converging rate is $O(\frac{\log n}{n})$, that is even slower than $O(\frac{1}{n})$, whereas the size of the transition probability matrix of the Markovian approximation grows as $O(n^2)$. As a result, for a given estimation error ρ the required number of iterations of the Ulam's method is much greater than $O(\frac{1}{\rho})$, that implies, due to the quadratic growth of the computing data, an algorithmic complexity much greater than $O(\frac{1}{\rho^2})$. Since ρ is reasonably small, the algorithmic complexity involved by the Ulam's method is much worse than our estimation method proposed in [2].

8 Conclusions

In this paper we have discussed an efficient iterative method for the estimation of the chief dynamical invariants for the special class of stochastically stable PWAE maps (e.g., the invariant measure, the Lyapunov exponent as well as the Kolmogorov-Sinai entropy). In detail, the estimation of the Kolmogorov-Sinai entropy is obtained by the direct computation of high order statistics characterizing the chaotic process. As shown in this paper, the proposed approach results to be much more efficient than those algorithms based, e.g., on Monte-Carlo methods, or based on the discretization of the Frobenius-Perron operator (such as the well known Ulam's method) [6, 11, 14, 17, 18, 23, 27, 28].

Moreover, we have provided a theoretical method for determining *a priori* a sufficient number of iterations that guarantees a given estimation accuracy and we have carried out an exhaustive comparison between our estimation procedure and the well known Ulam's universal method. As discussed in the last Section, the comparison is carried out from a theoretical point of view, from an operative point of view and from an algorithmic complexity point of view, and results confirm the strong benefits provided by the proposed solution.

References

[1] Addabbo, T., A. Fort, S. Rocchi and V. Vignoli (2008*a*). An efficient and accurate method for computing the invariant measure of piecewise affine chaotic maps. In: *ISCAS 2008, Proceedings of the 2008 IEEE International Symposium on Circuits and Systems*. pp. 760–763.

[2] Addabbo, T., A. Fort, S. Rocchi, D. Papini and V. Vignoli (2008*b*). Invariant measures of tunable chaotic sources: robustness analysis and efficient estimation. *IEEE Transactions on Circuits and Systems - I.* Accepted, available online.

[3] Addabbo, T., M. Alioto, A. Fort, S. Rocchi and V. Vignoli (2006). A feedback strategy to improve the entropy of a chaos-based random bit generator. *IEEE Transaction on Circuits and Systems – part I* **53**(2), 326–337.

[4] Baladi, V. and L.S. Young (1993). On the spectra of randomly perturbed expanding maps. *Communication in Mathematical Physics* **156**, 355–385.

[5] Baranovski, A.I. and A.J. Lawrance (2007). Sensitive parameter dependence of autocorrelation function in piecewise linear maps. *International Journal of Bifurcation and Chaos* **17**, 1185–1197.

[6] Blank, M. (2001). Perron-Frobenius spectrum for random maps and its approximation. *Moscow Mathematical Journal*.

[7] Blank, M. and G. Keller (1998). Random perturbations of chaotic dynamical systems: stability of the spectrum. *Nonlinearity* **11**, 1351–1364.

[8] Bose, C. and R. Murray (2001). The exact rate of approximation in Ulam's mehtod. *Discrete and continuous dynamical systems* **7**(1), 219–235.

[9] Boyarsky, A. and P. Góra (1997). *Laws of Chaos*. Birkhäuser.

[10] Callegari, S., R. Rovatti and G. Setti (2003). Generation of constant-envelope spread-spectrum signals via chaos-based FM: theory and simulation results. *IEEE Trans. Circ. Systems I* **50**, 3–15.

[11] Dellnitz, M. and O. Junge (1999). On the approximation of complicated dynamical behavior. *SIAM Journal on Numerical Analysis* **36**(2), 491–515.

[12] Fraser, A.M. (1989). Information and entropy in strange attractors. *IEEE Transaction on Information Theory* **35**(2), 245–261.

[13] Froyland, G. (1997). Computer-assisted bounds for the rate of decay of correlations. *Communications in Mathematical Physics* **189**, 237–257.

[14] Froyland, G. (1998). Approximating physical invariant measures of mixing dynamical systems in higher dimensions. *Nonlinear Analysis, Theory, Methods & Applications* **32**(7), 831–860.

[15] Froyland, G. (1999*a*). Ulam's method for random interval maps. *Nonlinearity* **12**, 1029–1052.

[16] Froyland, G. (1999*b*). Ulam's method to calculate entropy and other dynamical invariants. *Nonlinearity* **12**, 79–101.

[17] Froyland, G., K. Judd and A.I. Mess (n.d.). Estimation of Lyapunov exponents in dynamical systems using spatial average.

[18] Guder, R., M. Dellnitz and E. Kreuzer (1997). An adaptive method for the approximation of the generalized cell mapping. *Chaos, Solitons and Fractals* **8**(4), 525–534.

[19] Keller, G. (1984). On the rate of convergence to equilibrium in one-dimensional systems. *Communications in Mathematical Physisc* **96**, 181–193.

[20] Kennedy, M. P., Rovatti, R. and Setti, G., Eds.) (2000). *Chaotic Electronics in Telecomunications*. Boca Raton, FL: CRC Press.

[21] Kohda, T. (2002). Information sources using chaotic dynamics. *Proc. of the IEEE* **90**(5), 641–661.

[22] Lasota, A. and J. A. Yorke (1973). On the existence of invariant measures for piecewise monotonic transformations. *Trans. Amer. Math. Soc.* **186**, 481–488.

[23] Li, T. (1976). Finite approximation for the Frobenius-Perron operator. a solution to Ulam's conjecture. *Journal of Approximation Theory* **17**, 177–186.

[24] Liverani, C. (1995*a*). Decay of correlations. *Annals of Mathematics* **142**, 239–301.

[25] Liverani, C. (1995*b*). Decay of correlations in piecewise expanding maps. *Journal of Statistical Physics* **78**, 1111–1129.

[26] Mazzini, G., G. Setti and R. Rovatti (1997). Chaotic complex spreading sequences for asynchronous CDMA– Part I: system modeling and results. *IEEE Trans. Circuits Syst. I* **45**, 937–947.

[27] Murray, R. (2004). Optimal partition choice for invariant measure approximation for one-dimensional maps. *Nonlinearity* **17**, 1623–1644.

[28] Setti, G., G. Mazzini, R. Rovatti and S. Callegari (2002). Statistical modeling of discrete-time chaotic processes: basic finite-dimensional tools and applications. *Proc. of the IEEE* **90**(5), 662–690.

[29] Sinai, Ya (1959). On the concept of entropy for a dynamic system. *Doklady Akademii Nauk SSSR* **124**, 768–771.

[30] Tsuneda, A. (2005). Design of binary sequences with tunable exponential autocorrelations and run statistics based on one-dimensional chaotic maps. *IEEE Transaction on Circuits and Systems I* **52**(3), 454–462.

[31] Ulam, S. (1960). *Problems in modern mathematics*. Interscience Publishers. New York.