ESAIM: Probability and Statistics September 2001, Vol. 5, 51–76
URL: `http://www.emath.fr/ps/`

DENSITY ESTIMATION FOR ONE-DIMENSIONAL DYNAMICAL SYSTEMS

Clémentine Prieur[1]

Abstract. In this paper we prove a Central Limit Theorem for standard kernel estimates of the invariant density of one-dimensional dynamical systems. The two main steps of the proof of this theorem are the following: the study of rate of convergence for the variance of the estimator and a variation on the Lindeberg–Rio method. We also give an extension in the case of weakly dependent sequences in a sense introduced by Doukhan and Louhichi.

Mathematics Subject Classification. 37D20, 37M10, 37A50, 60G07, 60G10.

Received January 19, 2001. Revised March 9, June 12 and July 3, 2001.

1. Introduction

This paper considers estimation of the marginal density f of a stationary sequence $(X_n)_{n\in\mathbb{N}}$ of dependent random variables. If $(X_n)_{n\in\mathbb{N}}$ satisfies mixing conditions, Robinson [24] obtains the following result:

$$\sqrt{nb_n}\left[\hat{f}_n(x) - \mathbf{E}\hat{f}_n(x)\right] \xrightarrow[n\to+\infty]{\mathcal{D}} \mathcal{N}\left(0, f(x)\int_{-\infty}^{+\infty} K^2(t)\mathrm{d}t\right), \tag{1.1}$$

where $\hat{f}_n(x)$ is a standard kernel density estimate defined as follows (see Rosenblatt [25]):

$$\hat{f}(x) = \hat{f}_n(x) = \frac{1}{nb_n}\sum_{k=0}^{n-1} K\left(\frac{x - X_k}{b_n}\right), \tag{1.2}$$

with a sequence $(b_n)_{n\in\mathbb{N}} \in (\mathbb{R}_+)^{\mathbb{N}}$ and a compact supported kernel $K : \mathbb{R} \to \mathbb{R}$ (we note D its support) satisfying:

$$b_n \xrightarrow[n\to+\infty]{} 0, \text{ and } \int_D K(t)\mathrm{d}t = 1,\ 0 < \int_D K^2(t)\mathrm{d}t < \infty. \tag{1.3}$$

Quote that the last assumption, $\int_D K^2(t)\mathrm{d}t < \infty$, holds for K measurable and bounded.

Keywords and phrases: Dynamical systems, decay of correlations, invariant probability, stationary sequences, Lindeberg theorem, Central Limit Theorem, bias, nonparametric estimation, s-weakly and a-weakly dependent.

[1] Université de Cergy-Pontoise, Laboratoire de Mathématiques, bâtiment A4, Site Saint-Martin, 95011 Cergy-Pontoise Cedex, France; e-mail: `prieur@math.u-cergy.fr`

The purpose of this paper is to prove such a result for a certain class of one-dimensional dynamical systems. Let us introduce the class $(\mathcal{T})$ of dynamical systems:

$$\forall n \geq 1,\ X_n = T^n X_0 \tag{1.4}$$

such that:

- $T : I \to \mathbb{R}$ is a function defined on a closed interval $I \subset \mathbb{R}$;
- T admits an invariant probability measure μ_0 absolutely continuous with respect to the Lebesgue measure;
- the random variable X_0 has the distribution μ_0.

Therefore $(X_n)_{n\in\mathbb{N}}$ is a stationary sequence. Let f denote the density of μ_0 with respect to Lebesgue.

We also assume a control of the correlations. Before stating this control, let us define the set of functions $\mathcal{BV}$. We first define the variation of a function $\varphi : I = [L, R] \to \mathbb{R}$ (L and R can be respectively set equal to $-\infty$ and $+\infty$) by

$$\mathcal{V}(\varphi) = \sup_J \sup \sum_{i=1}^{n} |\varphi(x_{i-1}) - \varphi(x_i)|$$

where the first supremum is taken over all compact subsets $J = [L_J, R_J]$ of $I = [L, R]$, and where the second supremum is taken over all finite partitions $L_J = x_0 < x_1 < \cdots < x_n = R_J$, $n \geq 1$, of J (see [26]). Now $\mathcal{BV}$ denotes the set of functions $h : I \to \mathbb{R}$ with bounded variation and whose $\mathbb{L}^1-$norm is finite (see *e.g.* [27]). If $||h||_{\mathcal{BV}} := \mathcal{V}(h) + ||h||_1$ where $\mathcal{V}(h)$ is the variation of h and $|| \cdot ||_1$ is the standard norm on $\mathbb{L}^1$, then $|| \cdot ||_{\mathcal{BV}}$ is a norm and $\mathcal{BV}$ endowed with this norm is a Banach space.

We can now state the assumption on the correlations.

There exist a positive constant κ and a sequence of non-negative numbers $(d_n)_{n\in\mathbb{N}}$ satisfying $\sum_{k=0}^{+\infty} d_k < \infty$ such that

$$\forall n \geq 0,\ \forall\ h,\ k \in \mathcal{BV}, \quad |\mathrm{Cov}\,(h(X_0), k(X_n))\,| \leq \kappa \|k\|_1 \|h\|_{\mathcal{BV}}\ d_n, \tag{1.5}$$

where Cov denotes the covariance with respect to the invariant measure μ_0. This control yields the Theorem 1.1 (see Sect. 4).

The paper is organized as follows. In Section 2, we precise the assumptions on the class $\mathcal{T}$ of dynamical systems. In Section 3 we study the convergence in mean squares of our density estimates. This is the purpose of Lemma 3.1 which constitutes the tool step of this paper and improves existing results (see Bosq and Guégan [5], Maës [19]). Section 4 is devoted to the statement and the proof of the main result: the Central Limit Theorem (CLT) of type (1.1). Lemma 3.1, together with a variation on the Lindeberg–Rio method [22, 23], yields the central limit Theorem 4.3. Hence the rate of convergence obtained in Lemma 3.1 is "the good one". Theorem 4.3 does not involve the bias term. However the analysis of the bias is standard (see [2, 25]) and is not linked with the dependent structure of the subjacent sequence $(X_n)_{n\in\mathbb{N}}$ but only on its marginal distribution; this yields Theorems 4.1 and 4.2. Some examples follow in Section 5.

In Section 6 we extend the results of Sections 3 and 4 to the case where X_0 does not follow the invariant law μ_0. Properties of the Perron–Frobenius operator allow to conclude for Lasota–Yorke transformations T (Lem. 6.1 and Ths. 6.2, 6.3 and 6.4). Thanks to the results of this section, we can choose a density function p and construct a (non-stationary) sequence $(X_n^{'})_{n\in\mathbb{N}}$ as follows:

$$X_0^{'} \text{ has the distribution } p(t)\mathrm{d}t,\ \ X_n^{'} = T^n X_0^{'},\ n \geq 1, \tag{1.6}$$

and we then estimate the invariant density f with

$$\hat{p}(x) = \hat{p}_n(x) = \frac{1}{nb_n} \sum_{k=0}^{n-1} K\left(\frac{x - X_k^{'}}{b_n}\right) \tag{1.7}$$

where b_n and K are defined as in (1.3).

Finally, the appendix is devoted to an extension to the case of weakly dependent sequences in a sense introduced by Doukhan and Louhichi [13] (Th. A.1).

2. Definition of the class $\mathcal{T}$ of transformations

In this part we detail the technical assumptions required for the Central Limit Theorems given in Section 4. We first assume that the kernel K is in the set $\mathcal{BV}$. We consider a closed interval $I := [L, R] \subset \mathbb{R}$, and T a function from I into itself. We denote λ the Lebesgue measure on I and $\mathrm{int}(I)$ the interior of I. We assume:

- for all k in $\mathbb{N}$, for all x in $\mathrm{int}(I)$, $\lim_{t\to 0^+} T^k(x+t) =: T^k(x^+)$ and $\lim_{t\to 0^-} T^k(x+t) =: T^k(x^-)$ exist;
- for all k in $\mathbb{N}^*$, denote $D_-^k := \{x \in \mathrm{int}(I),\ T^k(x^-) = x\}$ and $D_+^k := \{x \in \mathrm{int}(I),\ T^k(x^+) = x\}$. Let $I_1 := \mathrm{int}(I) \setminus B$, where $B := \bigcup_{k\in\mathbb{N}^*}(D_-^k \cup D_+^k)$. We assume $\lambda(B) = 0$;
- T admits at least an invariant probability measure μ_0 which is absolutely continuous with respect to Lebesgue measure: $d\mu_0 = f d\lambda$;
- let $S := \mathrm{supp}(\mu_0)$ be the support of μ_0. If $I_2 := S \cap C(f)$ where $C(f)$ denotes the continuity set of f, then $\lambda(S \setminus I_2) = 0$.

Recall that μ_0 is the distribution of X_0 and that $(X_n)_{n\in\mathbb{N}}$ is defined by (1.4). Finally we assume the control of correlations described by inequality (1.5) in the introduction (note that the absolutely continuous invariant probability measure μ_0 is unique because of the control of correlations (1.5)).

The following "tent-map" (see Fig. 1 in Sect. 5) belongs to $\mathcal{T}$ (hence $\mathcal{T} \neq \varnothing$):

$$T(x) = \begin{cases} \frac{3}{2}x & \text{if } 0 \le x \le \frac{1}{2} \\ \frac{-3}{2}x + \frac{3}{2} & \text{if } \frac{1}{2} < x \le 1. \end{cases}$$

In the next sections, $\lambda(D)$ denotes the Lebesgue measure of the compact set D.

3. Convergence in mean squares in the stationary case

In this section, we provide the mean squares convergence of the invariant density estimates for dynamical systems in the class $\mathcal{T}$. Let T be in $\mathcal{T}$, then the series of correlations are summable (1.5).

Remark 3.1. In many examples (see Sect. 5) we have a stronger property: the exponential decay of correlations. A classical way to prove it is by using the theory of transfer operators (see Collet [8]).

Recall that μ_0 is the absolute continuous invariant probability measure for T, that f is its density with respect to Lebesgue measure on I, and that $X_0 \sim \mu_0$. $(X_k)_{k\in\mathbb{N}}$ is then a stationary process with marginal density f. We estimate f by standard kernel estimates of the invariant density. We get the following mean squares convergence result:

Lemma 3.1. *Let T be in the class $\mathcal{T}$ and $\hat{f}(x)$ be defined by (1.2) and (1.3). Assume that $b_n \xrightarrow[n\to\infty]{} 0$. Then if $x \in I_1 \cap I_2$, we have*

$$\mathrm{Var}\left(\hat{f}(x)\right) = \frac{1}{nb_n}\left(f(x)\int_D K^2(s)\mathrm{d}s + o(1)\right).$$

Remark 3.2. A first evaluation, $\mathcal{O}\left(\frac{1}{nb_n^2}\right)$, of the rate of convergence of $\hat{f}_n(x)$ is given in [5] (or more recently in [19]). Our result provides $\lim_{n\to+\infty}(nb_n)\mathrm{Var}\left(\hat{f}(x)\right) = C,\ C \geq 0$. This accurate rate of convergence is necessary in order to obtain the forthcoming Central Limit Theorem (Th. 4.3) from Lindeberg–Rio technique.

To prove such a result in a mixing frame, the authors (*e.g.* [13]) usually assume that the couples $\{(X_0, X_k)\}_{k\in\mathbb{N}^*}$ have regular joint densities. Here these distributions are singular and therefore our study is quite different. We first need the two following lemmas:

Lemma 3.2. *Assume that T is in the class $\mathcal{T}$ and let $x \in I_1 \cap I_2$. Assume that $b_n \xrightarrow[n\to\infty]{} 0$. Then for each $k \in \mathbb{N}^*$, there exists a sequence $\varepsilon(n,k) \xrightarrow[n\to\infty]{} 0$ such that*

$$\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right) = b_n\,\varepsilon(n,k).$$

Lemma 3.3. *Assume that T is in the class $\mathcal{T}$ and let $x \in I_1 \cap I_2$. Assume that $b_n \xrightarrow[n\to\infty]{} 0$. Then*

$$\frac{1}{nb_n}\sum_{k=1}^{n-1}(n-k)\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right) \xrightarrow[n\to\infty]{} 0.$$

Proof of Lemma 3.2. Recall that D denotes the support of K. Let k fixed in $\mathbb{N}^*$, then

$$\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right) = \int_L^R K\left(\frac{x-T^k s}{b_n}\right)K\left(\frac{x-s}{b_n}\right)f(s)\mathrm{d}s - \left(\int_L^R K\left(\frac{x-s}{b_n}\right)f(s)\mathrm{d}s\right)^2.$$

- <u>Study of $Q_n := \left(\int_L^R K\left(\frac{x-s}{b_n}\right)f(s)\mathrm{d}s\right)^2$</u>

$$0 \leq Q_n = b_n^2\left(\int_{\frac{x-R}{b_n}}^{\frac{x-L}{b_n}} K(t)f(x-tb_n)\mathrm{d}t\right)^2 \leq b_n^2\,\left(\int_D |K(t)|\,f(x-tb_n)\mathrm{d}t\right)^2.$$

Let $l_n(t) := |K(t)|\,f(x-tb_n)$. As f is continuous at point x ($x \in I_2$ implies $x \in C(f)$), we get for all $t \in D$: $l_n(t) \xrightarrow[n\to\infty]{} |K(t)|\,f(x)$. Let $\varepsilon > 0$. As D is compact, the convergence of $f(x-tb_n)$ to $f(x)$ is uniform in $t \in D$. Hence there exists a positive integer N such that for $n \geq N$, $|l_n(t)| \leq |K(t)|\,(f(x)+\varepsilon)\ \forall t \in D$.

Then, as $\int_D |K(t)|\,(f(x)+\varepsilon)\,\mathrm{d}t \leq (f(x)+\varepsilon)\,\sqrt{\lambda(D)}\,\sqrt{\int_D K^2(t)\mathrm{d}t} < \infty$, the dominated convergence theorem yields ([26], p. 27)

$$\int_D |K(t)|\,f(x-tb_n)\mathrm{d}t \xrightarrow[n\to\infty]{} f(x)\int_D |K(t)|\,\mathrm{d}t < \infty.$$

Hence $b_n\left(\int_D |K(t)|\,f(x-tb_n)\mathrm{d}t\right)^2 \xrightarrow[n\to\infty]{} 0$ and there exists $\varepsilon_1(n) \xrightarrow[n\to\infty]{} 0$ such that $Q_n = b_n\,\varepsilon_1(n)$.

- <u>Study of $A_{n,k} := \int_L^R K\left(\frac{x-T^k s}{b_n}\right) K\left(\frac{x-s}{b_n}\right) f(s)\mathrm{d}s$</u>

$$
\begin{aligned}
0 \le |A_{n,k}| &= b_n \left| \int_{\frac{x-R}{b_n}}^{\frac{x-L}{b_n}} K\left(\frac{x-T^k(x-tb_n)}{b_n}\right) K(t) f(x-tb_n)\mathrm{d}t \right| \\
&\le b_n \int_D \left| K\left(\frac{x-T^k(x-tb_n)}{b_n}\right) K(t) \right| f(x-tb_n)\mathrm{d}t.
\end{aligned}
$$

As D is compact, $tb_n \to 0$ uniformly in $t \in D$ as n tends to infinity. Hence the existence of onesided limits at point x implies
$x - T^k(x-tb_n) \longrightarrow x - T^k(x^-)$, uniformly with respect to $t \in D_+^* := D \cap \mathbb{R}_+^*$;
$x - T^k(x-tb_n) \longrightarrow x - T^k(x^+)$, uniformly with respect to $t \in D_-^* := D \cap \mathbb{R}_-^*$.
Moreover, by assumption, $x - T^k(x^-) \neq 0$ and $x - T^k(x^+) \neq 0$. Let $D^* := D \setminus \{0\}$.

From the compactness of D we exhibit some $n_0 \in \mathbb{N}^*$ such that
$n \ge n_0 \Longrightarrow \left[\forall\, t \in D^* :\ K\left(\frac{x-T^k(x-tb_n)}{b_n}\right) = 0\right]$.
Then $n \ge n_0 \Longrightarrow \frac{A_{n,k}}{b_n} = 0$ and we can write $A_{n,k} = b_n \varepsilon_2(n,k)$, with $\varepsilon_2(n,k) = 0$ for n large enough.

Then $\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right) = b_n[\varepsilon_2(n,k) - \varepsilon_1(n)]$. Hence for each $k \in \mathbb{N}^*$, there exists $\varepsilon(n,k) \xrightarrow[n\to\infty]{} 0$ such that $\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right) = b_n\, \varepsilon(n,k)$, which concludes the proof. □

Proof of Lemma 3.3. We have

$$
\frac{1}{nb_n}\sum_{k=1}^{n-1}(n-k)\left|\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right)\right| \le \frac{1}{b_n}\sum_{k=1}^{n-1}\left|\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right)\right| \cdot \quad (3.1)
$$

By the control (1.5) of correlations and from Lemma 3.2 and inequality (3.1), there exists a constant $M > 0$ such that

$$
\frac{1}{nb_n}\sum_{k=1}^{n-1}(n-k)\left|\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right)\right| \le M\sum_{k=1}^{n-1}\min\left(\varepsilon(n,k), d_k\right).
$$

Let $\varepsilon > 0$. Since $\sum_{k=0}^{\infty} d_k < \infty$, there exists $k(\varepsilon) \in \mathbb{N}^*$ such that $\sum_{k=k(\varepsilon)}^{\infty} d_k < \frac{\varepsilon}{2}$. So for $n \ge k(\varepsilon)$,

$$
\sum_{k=1}^{n-1}\min\left(\varepsilon(n,k), d_k\right) < \sum_{k=1}^{k(\varepsilon)-1}\varepsilon(n,k) + \frac{\varepsilon}{2}\cdot
$$

Hence there exists some $n_0 \ge k(\varepsilon)$ such that $n \ge n_0$ implies

$$
\sum_{k=1}^{k(\varepsilon)-1}\varepsilon(n,k) + \frac{\varepsilon}{2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
$$

This entails $\sum_{k=1}^{n-1}\min\left(\varepsilon(n,k), d_k\right) \xrightarrow[n\to\infty]{} 0$ and concludes the proof. □

Remark 3.3. Using similar arguments, we can prove that for all $0 \leq i < j \leq n-1$, and for any bounded function φ,

$$\operatorname{Cov}\left(\varphi(X_0)\left[K\left(\frac{x-T^jX_0}{b_n}\right)-\mathbf{E}K(\frac{x-T^jX_0}{b_n})\right], K\left(\frac{x-T^iX_0}{b_n}\right)-\mathbf{E}K\left(\frac{x-T^iX_0}{b_n}\right)\right) \leq b_n \times \varepsilon(n,j-i),$$

with $\varepsilon(n,j-i) \xrightarrow[n\to\infty]{} 0$. We make use of this remark in the next section to prove the central limit Theorem 4.3.

We are now in position to prove Lemma 3.1.

Proof Lemma 3.1. Let $\operatorname{Var}_{\text{ind}}\hat{f}(x)$ denote the variance of $\frac{1}{nb_n}\sum_{k=0}^{n-1} K\left(\frac{x-\widetilde{X_k}}{b_n}\right)$ where the $\widetilde{X_k}$'s are independent copies of the X_k's. We have:

$$\operatorname{Var}(\hat{f}(x)) = \operatorname{Var}_{\text{ind}}\hat{f}(x) + \frac{2}{n^2b_n^2}\sum_{0\leq i<j\leq n-1}\operatorname{Cov}\left(K\left(\frac{x-X_i}{b_n}\right), K\left(\frac{x-X_j}{b_n}\right)\right). \tag{3.2}$$

As $x \in I_1 \subset \operatorname{int}(I)$, we have for n large enough

$$\operatorname{Var}_{\text{ind}}\hat{f}(x) = \frac{1}{nb_n}\int_D K^2(s)\, f(x-sb_n)\mathrm{d}s - \frac{1}{n}\left(\int_D K(s)\, f(x-sb_n)\mathrm{d}s\right)^2.$$

As f is continuous at point x, and as K is compactly supported and satisfies (1.3), apply twice the dominated convergence theorem (see *e.g.* [26], p. 27) to obtain

$$\begin{aligned}
\operatorname{Var}_{\text{ind}}\hat{f}(x) &= \frac{1}{nb_n}\textstyle\int_D K^2(s)\, f(x-sb_n)\mathrm{d}s - \frac{1}{n}\left(\int_D K(s)\, f(x-sb_n)\mathrm{d}s\right)^2 \\
&= \frac{f(x)}{nb_n}\textstyle\int_D K^2(s)\mathrm{d}s + o\left(\frac{1}{nb_n}\right) + \frac{1}{nb_n}\left\{b_n\left(\left(\int_D K(s)f(x)\mathrm{d}s\right)^2 + o(1)\right)\right\} \\
&= \frac{f(x)}{nb_n}\textstyle\int_D K^2(s)\mathrm{d}s + o\left(\frac{1}{nb_n}\right).
\end{aligned}$$

Remark 3.4. Quote that Bosq and Lecoutre ([6], p. 76) ask an additional differentiability assumption on f in order to obtain an equivalent of the bias together with this result.

Hence

$$\operatorname{Var}\left(\hat{f}(x)\right) = \frac{f(x)}{nb_n}\int_D K^2(s)\mathrm{d}s + o\left(\frac{1}{nb_n}\right) + \frac{2}{n^2b_n^2}\sum_{k=1}^{n-1}(n-k)\operatorname{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right).$$

Then

$$\operatorname{Var}\left(\hat{f}(x)\right) = \frac{f(x)}{nb_n}\int_D K^2(s)\mathrm{d}s + o\left(\frac{1}{nb_n}\right) + \frac{2}{nb_n}\left(\frac{1}{nb_n}\sum_{k=1}^{n-1}(n-k)\operatorname{Cov}\left(K\left(\frac{x-X_0}{b_n}\right), K\left(\frac{x-X_k}{b_n}\right)\right)\right).$$

Hence by Lemma 3.3,

$$\operatorname{Var}\left(\hat{f}(x)\right) = \frac{1}{nb_n}\left(f(x)\int_D K^2(s)\mathrm{d}s + o(1)\right),$$

which concludes the proof. □

Remark 3.5. We may also study the MISE defined as follows

$$\text{MISE}\,(\hat{f}, f) := \int \mathbf{E}(\hat{f} - f)^2 \mathrm{d}x.$$

As usual (see [2]) the rate of convergence of the MISE to 0 depends on the regularity of f. If T is Lasota–Yorke and Markov, we can deduce the regularity of the density f from the one of T (see [1]).

4. A Central Limit Theorem in the stationary case

Many versions of a Central Limit Theorem for the partial sums of dynamical systems $\frac{1}{\sqrt{n}} \sum_{j=1}^{[ns]} \phi(T^j(x))$ have been proved in the literature. For example Liverani [18], Viana [27] prove a CLT in the case where $s = 1$ for some piecewise expanding dynamical systems. Barbour *et al.* [4] prove a functional CLT with respect to s in the case where T is some expanding map of the unit interval into itself. They use first a coupling method: they prove that the iterates of T can be closely tied to an $m-$dependent process. Then they use techniques which are derived using Stein's method, so they obtain bounds on the rate of convergence. Here we prove a CLT for the density estimates.

We study the following process

$$U_n(x) := \sqrt{nb_n}(\hat{f}(x) - f(x)). \tag{4.1}$$

We do not use a decomposition in Bernstein blocks. Here the idea is to adapt the Lindeberg method after Rio [23]. To be in position to use such a method, we need the mean squares convergence result stated in Section 3 (Lem. 3.1). We then study the bias term by a Taylor's decomposition. Let us now state the central limit results for the invariant density estimates in the case of dynamical systems in the class $\mathcal{T}$.

We first precise the following notations. If $l \in \mathbb{N}^*$ and if $a,\ y_1, \dots, y_l \in \mathbb{R}$, $a * (y_i)_{1 \le i \le l}$ denotes the vector of $\mathbb{R}^l$ whose coordinates are $a * y_1, \dots, a * y_l$, and $a * \text{diag}\,(y_1, \dots, y_l)$ denotes the diagonal matrix whose diagonal terms are equal to $a * y_1, \dots, a * y_l$. We also define

$$\Sigma_l := \left(\int_D K^2(s)\mathrm{d}s \right) * \text{diag}\,(f(x_1), \dots, f(x_l))\,. \tag{4.2}$$

Theorem 4.1. *Let T be in the class $\mathcal{T}$ and $\hat{f}(x)$ be defined by (1.2) and (1.3). Assume that $b_n \xrightarrow[n \to +\infty]{} 0$, $nb_n \xrightarrow[n \to +\infty]{} +\infty$. Let l be a positive integer. For all $1 \le i \le l$, let $x_i \in I_1 \cap I_2$. Let m be a positive integer. Assume that for each i, $1 \le i \le l$, there exists a neighbourhood V_i of x_i such that the invariant density f is m-times continuously differentiable on V_i. Also assume that $\int_D s^j K(s)\mathrm{d}s = 0$ for all integer j such that $1 \le j \le m-1$ and that nb_n^{2m+1} converges to some non-negative constant ρ_m as n tends to infinity. Then*

$$(U_n(x_1),\ \dots,\ U_n(x_l)) \xrightarrow[n \to \infty]{\mathcal{D}} \mathcal{N}\left(\frac{(-1)^m\ \sqrt{\rho_m}}{m!}\ \int_D s^m\ K(s)\mathrm{d}s * \left(f^{(m)}(x_i)\right)_{1 \le i \le l},\ \Sigma_l \right),$$

where Σ_l is defined by (4.2).

Remark 4.1. For example, when $m = 1$ Theorem 4.1 yields

$$(U_n(x_1),\ \dots,\ U_n(x_l)) \xrightarrow[n \to \infty]{\mathcal{D}} \mathcal{N}\left(-\sqrt{\rho_1} \int_D s\ K(s)\mathrm{d}s * (f'(x_i))_{1 \le i \le l}\,,\ \Sigma_l \right).$$

We can also write such a theorem if the regularity of the invariant density f in terms of Hölder spaces is not necessarily an integer. Let ν denote the regularity of the function f, this means that setting $\nu = \alpha + \beta$ with $\alpha \in \mathbb{N}$ and $0 \le \beta < 1$ there exists a constant $A > 0$ such that f is α-times continuously differentiable with $|f^{(\alpha)}(x) - f^{(\alpha)}(y)| \le A|x - y|^\beta$ for $x,\ y$ belonging to an arbitrary compact interval. We get the following result:

Theorem 4.2. *Let T be in the class $\mathcal{T}$ and $\hat{f}(x)$ be defined by (1.2) and (1.3). Assume that $b_n \xrightarrow[n\to+\infty]{} 0$, $nb_n \xrightarrow[n\to+\infty]{} +\infty$. Let l be a positive integer. For all $1 \le i \le l$, let $x_i \in I_1 \cap I_2$. Let $\nu = \alpha + \beta$ with $\alpha \in \mathbb{N}$ and $0 \le \beta < 1$. Assume that for each i, $1 \le i \le l$, there exists a neighbourhood V_i of x_i such that the invariant density f has the regularity ν on V_i. Also assume that $\int_D s^j K(s)\mathrm{d}s = 0$ for all integer j such that $1 \le j \le \alpha$ and that $nb_n^{2\nu+1} \xrightarrow[n\to+\infty]{} 0$. Then*

$$(U_n(x_1),\ \ldots\ ,\ U_n(x_l)) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0,\ \Sigma_l\right),$$

where Σ_l is defined by (4.2).

Example 4.1. *In many cases, the regularity of the invariant density f can be deduced from the one of the dynamic T. Let us give some examples and references.*

We consider a transformation T of the interval $I = [0,1]$. We assume that T is Lasota-Yorke in a sense defined in Section 5. Then there exists a countable partition (finite or not) of I, $\{a_j\}_{j\in J}$, such that for all $j \in J$, on $]a_j, a_{j+1}[$, T is expansive (see Assumpt. 5.2). For all $j \in J$, we assume that T admits a continuous extension to $I_j := [a_j, a_{j+1}]$. Let us denote $\overline{T}\,(I_j)$ the image of I_j by the extension of T on I_j. We assume moreover that T is of Markov type, i.e.

1. *if the partition $\{a_j\}_{j\in J}$ is finite, we assume that for all $j \in J$, there exists $K_j \subset J$ such that $\overline{T}\,(I_j) = \bigcup_{k\in K_j} I_k$. Now if A is a subset of I, let clos(A) denotes the closure of A. We assume in that case that there exists a positive integer p such that for all $j \in J$, clos$\left(T^p\left(]a_j, a_{j+1}[\right)\right) = I$;*
2. *if the partition $\{a_j\}_{j\in J}$ is infinitely countable, we assume that for all $j \in J$, $\overline{T}\,(I_j) = I$.*

We now give easy examples of such transformations T.

1. *the r-adic maps*
$$T(x) = rx\ [1],\ \text{where } r \in \mathbb{N},\ r \ge 2.$$
We have $0 = a_0 < \frac{1}{r} < \cdots < \frac{r-1}{r} < a_r = 1$;
2. *generalization of the r-adic maps*
$$T(x) = rx + c\ [1],\ \text{where } r > 1,\ 0 \le c < 1,\ \text{and } r + c \in \mathbb{N},\ c(r+1) \in \mathbb{N}^*.$$
For these maps we can take $0 = a_0 < \frac{1-c}{r} < \cdots < \frac{(r+c-1)-c}{r} < a_{r+c} = 1$;
3. *some piecewise linear transformations.*
There exists a countable partition (finite or not), $\{a_j\}_{j\in J}$, of I such that for all $j \in J$, $\overline{T}\,(I_j) = I$, T is linear and continuously differentiable on $]a_j, a_{j+1}[$ and $|T'(x)| \ge 1 + \varepsilon$ where $\varepsilon > 0$;
4. *the Gauss map*
$$T(x) = \frac{1}{x} - \left[\frac{1}{x}\right] \text{ for } x \neq 0 \text{ and } T(0) = 0.$$
Then for all $j \in J = \mathbb{N}^*$, $I_j = [a_j, a_{j+1}] = \left[\frac{1}{j+1}, \frac{1}{j}\right]$.
For this map we know the exact form of the invariant density, $f(x) = \frac{1}{\log(2)}(1+x)^{-1}$, which is infinitely continuously differentiable.

For these transformations T, we know (e.g. [1, 7, 15]) that the regularity of the invariant density f depends on the one of T. If the partition is finite and if T is piecewise two-times continuously differentiable, then f is piecewise continuously differentiable. If moreover T is onto on each interval of the partition, then f is continuously differentiable. In the case where the partition is infinitely countable, a further assumption on the Schwarzian derivative of T is needed (see e.g. [1, 10]) to conclude that f is Hölder continuous.

If the regularity of f is not an integer, no equivalent of the bias seems to be known. The optimal rate is reached but we do not get an explicit equivalent of the bias. Hence we sometimes prefer not to consider the bias term but rather to restrict our attention to the centered estimation process

$$Y_n(x) := \sqrt{nb_n}(\hat{f}(x) - \mathbf{E}\hat{f}(x)). \tag{4.3}$$

We get the following result (see also Applications 6.1 and 6.2 in Sect. 6 for some use of this result):

Theorem 4.3. *Let T be in the class $\mathcal{T}$, and $\hat{f}(x)$ be defined by (1.2) and (1.3). Assume that $b_n \xrightarrow[n\to+\infty]{} 0$, $nb_n \xrightarrow[n\to+\infty]{} +\infty$. Then if for all $1 \le i \le l$, $x_i \in I_1 \cap I_2$,*

$$(Y_n(x_1),\ \dots,\ Y_n(x_l)) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0,\ \Sigma_l)\,,$$

where Σ_l is defined by (4.2).

Remark 4.2. For $l = 1$ Theorem 4.3 is a CLT. Let $J \subset I_1 \cap I_2$ be a compact subinterval of $I_1 \cap I_2$. Working with arbitrary l and with some $f > 0$ implies that the sequence of estimation processes $\left(\frac{Y_n(x)}{\sqrt{f(x)}},\ x \in J\right)_{n\in\mathbb{N}^*}$ is not tight in $C(J)$; its limit is indeed $\sqrt{\int_D K^2(s)\mathrm{d}s}\ \dot{W}$, where $\dot{W}$ is the Gaussian white noise. Now for sake of simplicity we develop the proof for $l = 1$. The general case is similar. If one wants to know the asymptotic behaviour in distribution of the vector $(Y_n(x_1),\ \dots,\ Y_n(x_l))$, it is sufficient to use the following proof of Theorem 4.3 (in the case $l = 1$) with $\frac{1}{\sqrt{nb_n}}\sum_{j=1}^{l} s_j K(\frac{x_j - X_k}{b_n})$, for arbitrary numbers $s_1,\ \dots,\ s_l \in \mathbb{R}$, instead of $\frac{1}{\sqrt{nb_n}}K(\frac{x-X_k}{b_n})$.

We first prove Theorem 4.3 and then deduce Theorems 4.1 and 4.2 by studying the bias term defined by $\mathcal{BIAS}_n(x) = \mathbf{E}\hat{f}_n(x) - f(x)$.

Proof of Theorem 4.3 with $l = 1$. Let us first notice that if $f(x) = 0$, then $Y_n(x)$ tends to zero in mean squares (Lem. 3.1), so it also converges to zero in law.

From now on, we suppose that $f(x) > 0$. Let $g_n(t) = \frac{1}{\sqrt{nb_n}}K(\frac{x-t}{b_n})$, $M_n = \|g_n\|_\infty$, $l_n = \|g_n\|_{\mathcal{BV}}$ and $\delta_n = \|g_n\|_1$.

In the following c will denote some constant independent of k and n, which may vary from line to line. We have $M_n \le \frac{c}{\sqrt{nb_n}}$, $l_n \le \frac{c}{\sqrt{nb_n}}$ and $\delta_n \le \frac{c\,b_n}{\sqrt{nb_n}}$, where c is positive. We recall that for all $h,\ k$ in $\mathcal{BV}$

$$||h \cdot k||_{\mathcal{BV}} \le \|h\|_\infty ||k||_{\mathcal{BV}} + \|k\|_\infty ||h||_{\mathcal{BV}}. \tag{4.4}$$

We set, for $k = 0,\ \dots,\ n-1$ and $n = 1,\ 2,\ \dots,$

$$Z_{n,k} = g_n(X_{n-k-1}) - \mathbf{E}\left(g_n(X_{n-k-1})\right),\ \text{and}\ S_n = Z_{n,0} + \dots + Z_{n,n-1}.$$

Now let $S_{k,n} = Z_{n,0} + \dots + Z_{n,k}$ for $0 \le k \le n-1$. Empty sums are, as usual, set equal to 0.

Hence

$$\lim_{n\to\infty} \operatorname{Var} S_n = f(x)\int_D K^2(s)\mathrm{d}s > 0. \tag{4.5}$$

Consider now a bounded thrice differentiable function $h : \mathbb{R} \to \mathbb{R}$ with continuous and bounded derivatives. Set $C_j = \|h^{(j)}\|_\infty$, for $j = 0,\ 1,\ 2,\ 3$. Set $\sigma_n^2 = \operatorname{Var} S_n$. For some standard Gaussian random variable η, write $\Delta_n(h) = \mathbf{E}(h(S_n) - h(\sigma_n\eta))$. The theorem will follow from (4.5), if we prove that $\lim_{n\to\infty}\Delta_n(h) = 0$. Let $v_{k,n} := \operatorname{Var} S_{k,n} - \operatorname{Var} S_{k-1,n}$, for $1 \le k \le n-1$.

$$v_{k,n} = 2\sum_{l=0}^{k-1} \operatorname{Cov}\left(Z_{n,k}, Z_{n,l}\right) + \mathbf{E}Z_{n,k}^2 =: a_{n,k}^1 + a_n^2,$$

with

$$a_n^2 \sim \frac{f(x)}{n}\int_D K^2(t)\mathrm{d}t,$$

a_n^2 independent of k, and

$$0 \le a_{n,k}^1 \le \frac{2}{n}\sum_{l=1}^{k}\min\left(\varepsilon(n,l), d_l\right).$$

So, as in Lemma 3.3, we show that $\sup_{0\le k\le n-1}(na_{n,k}^1) \longrightarrow 0$ as $n \longrightarrow \infty$. Then there exists $n_o \in \mathbb{N}^*$ such that for all $n \ge n_o$ and for all $k \in \mathbb{N}^*$ such that $0 \le k \le n-1$ we have $v_{k,n} > 0$. Therefore we can consider $Y_{n,k} \sim \mathcal{N}(0, v_{k,n})$ for sufficiently large n and for $0 \le k \le n-1$.

Let us assume that the array $\{Y_{n,k};\ 0 \le k \le n-1,\ n \ge n_0\}$ is independent and is independent of the sequence $(X_k)_{k\in\mathbb{N}}$. If $0 \le k \le n-1$, set $T_{n,k} = \sum_{j=k+1}^{n-1} Y_{n,j}$, still with empty sums set equal to 0. We can now write Rio's decomposition

$$\Delta_n(h) = \sum_{k=0}^{n-1}\Delta_{k,n}(h),$$

with $\Delta_{k,n}(h) = \mathbf{E}\left[h(S_{k-1,n} + Z_{n,k} + T_{n,k}) - h(S_{k-1,n} + Y_{n,k} + T_{n,k})\right]$.

The function $x \mapsto h_{k,n}(x) = \mathbf{E}h(x + T_{n,k})$ has the same derivability properties as h, *e.g.* for $0 \le j \le 3$, $\|h_{k,n}^{(j)}\| \le C_j$; now we write $\Delta_{k,n}(h) = \Delta_{k,n}^{(1)}(h) - \Delta_{k,n}^{(2)}(h)$, with

$$\begin{aligned}
\Delta_{k,n}^{(1)}(h) &= \mathbf{E}h_{k,n}(S_{k-1,n} + Z_{n,k}) - \mathbf{E}h_{k,n}(S_{k-1,n}) - \frac{v_{k,n}}{2}\mathbf{E}h_{k,n}^{''}(S_{k-1,n}),\\
\Delta_{k,n}^{(2)}(h) &= \mathbf{E}h_{k,n}(S_{k-1,n} + Y_{n,k}) - \mathbf{E}h_{k,n}(S_{k-1,n}) - \frac{v_{k,n}}{2}\mathbf{E}h_{k,n}^{''}(S_{k-1,n}).
\end{aligned}$$

- Bound of $\Delta_{k,n}^{(2)}(h)$.

  Using Taylor expansion yields for some (random) $\rho_{n,k} \in (0,\ 1)$:

$$\Delta_{k,n}^{(2)}(h) = \mathbf{E}h_{k,n}^{'}(S_{k-1,n})Y_{n,k} + \frac{1}{2}\mathbf{E}h_{k,n}^{''}(S_{k-1,n})(Y_{n,k}^2 - v_{k,n}) + \frac{1}{6}\mathbf{E}h_{k,n}^{(3)}(S_{k-1,n} + \rho_{n,k}Y_{n,k})Y_{n,k}^3.$$

From the independence of the Gaussian sequence $(Y_{n,k})_{n\in\mathbb{N},\ 0\le k\le n-1}$ and the process $(X_n)_{n\in\mathbb{N}}$,

$$|\Delta_{k,n}^{(2)}(h)| \le \frac{C_3}{6}\mathbf{E}|Y_{n,k}|^3\ ,$$

hence

$$|\Delta_{k,n}^{(2)}(h)| \le \frac{2C_3 v_{k,n}^{\frac{3}{2}}}{3\sqrt{2\pi}}\cdot$$

Now

$$v_{k,n} = \mathrm{Var}\, Z_{n,k} + 2\sum_{j=0}^{k-1}\mathrm{Cov}(Z_{n,j},\ Z_{n,k}),$$

hence

$$v_{k,n} \le 4M_n\delta_n + 2\sum_{j=1}^{k}2\delta_n l_n d_j. \tag{4.6}$$

We thus need

$$n^{\frac{2}{3}}\left(\frac{1}{n}+\sum_{j=1}^{k}\frac{d_j}{n}\right)\xrightarrow[n\to\infty]{}0.$$

As $\sum_{k=0}^{\infty} d_k<\infty$, the last assertion is always true.

- Bound of $\Delta^{(1)}_{k,n}(h)$.

Set $\Delta^{(1)}_{k,n}(h)=\mathbf{E}\delta^{(1)}_{k,n}(h)$. We write, again with some random $\tau_{n,k}\in(0,\ 1)$,

$$\delta^{(1)}_{k,n}(h)=h'_{k,n}(S_{k-1,n})Z_{n,k}+\frac{1}{2}h''_{k,n}(S_{k-1,n})(Z^2_{n,k}-v_{k,n})+\frac{1}{6}\left(h^{(3)}_{k,n}(S_{k-1,n}+\tau_{n,k}Z_{n,k})Z^3_{n,k}\right).$$

We analyze separately the terms in the previous expression. We have

$$\frac{1}{6}\left|\mathbf{E}h^{(3)}_{k,n}(S_{k-1,n}+\tau_{n,k}Z_{n,k})Z^3_{n,k}\right|\le\frac{C_3}{6}(4M_n^2)(2\delta_n).\tag{4.7}$$

To estimate the middle term, write (with Rio)

$$\mathrm{Cov}\left(h''_{k,n}(S_{k-1,n}),\ Z^2_{n,k}\right)=\sum_{j=0}^{k-1}\mathrm{Cov}\left(h''_{k,n}(S_{j,n})-h''_{k,n}(S_{j-1,n}),\ Z^2_{n,k}\right).$$

By Taylor,

$$\mathrm{Cov}\left(h''_{k,n}(S_{j,n})-h''_{k,n}(S_{j-1,n}),\ Z^2_{n,k}\right)=\mathrm{Cov}\left(Z_{n,j}\ h^{(3)}_{k,n}(S_{j-1,n}+uZ_{n,j}),Z^2_{n,k}\right),$$

for some $0<u<1$.

We can also write

$$Z^2_{n,k}=g_n^2(X_{n-k-1})-2\mathbf{E}\left(g_n(X_{n-k-1})\right)g_n(X_{n-k-1})+\left[\mathbf{E}\left(g_n(X_{n-k-1})\right)\right]^2,$$

and

$$Z_{n,j}=g_n(X_{n-j-1})-\mathbf{E}\left(g_n(X_{n-j-1})\right).$$

From those decompositions, the dominant term of

$$\mathrm{Cov}\left(Z_{n,j}\ h^{(3)}_{k,n}(S_{j-1,n}+uZ_{n,j}),Z^2_{n,k}\right)$$

is

$$\mathrm{Cov}\left(g_n(X_{n-j-1})\ h^{(3)}_{k,n}(S_{j-1,n}+uZ_{n,j}),g_n^2(X_{n-k-1})\right),$$

by replacing twice $Z_{n,l}$ by $g(X_{n-l-1})$. Hence using (4.4) and Remark 3.3 we obtain:

$$\mathrm{Cov}\left(g_n(X_{n-j-1})\ h^{(3)}_{k,n}(S_{j-1,n}+uZ_{n,j}),g_n^2(X_{n-k-1})\right)\le C_3\ \delta_n\ 2M_nl_nd_{k-j}.\tag{4.8}$$

So, as it is the dominant term, we get the same bound for the other terms. Summing up yields:

$$\left|\mathrm{Cov}\left(h''_{k,n}(S_{k-1,n}),\ Z^2_{n,k}\right)\right|\le 4\sum_{j=0}^{k-1}C_3\ \delta_n\ 2M_nl_nd_j.\tag{4.9}$$

Proceeding as for (4.9) implies:

$$\left|\mathrm{Cov}\left(h'_{k,n}(S_{i,n})-h'_{k,n}(S_{i-1,n}),\ Z_{n,k}\right)\right|\le 2C_2\delta_nl_nd_{k-i}.\tag{4.10}$$

Hence from (4.10) and from Remark 3.3:

$$\left|\mathrm{Cov}\left(h'_{k,n}(S_{i,n}) - h'_{k,n}(S_{i-1,n}),\ Z_{n,k}\right)\right| \le c\min\left(\delta_n l_n d_{k-i},\ \frac{\varepsilon(n,k-i)}{n}\right)\cdot \tag{4.11}$$

We also have

$$\left|\mathbf{E}h''_{k,n}(S_{k-1,n})\mathbf{E}Z_{n,i}Z_{n,k}\right| \le c\min\left(\delta_n l_n d_{k-i},\ \frac{\varepsilon(n,k-i)}{n}\right)\cdot \tag{4.12}$$

Adding (4.11) and (4.12) and summing up the expression for all i yields:

$$\left|\mathbf{E}h'_{k,n}(S_{k-1,n})Z_{n,k} - \mathbf{E}h''_{k,n}(S_{k-1,n})\sum_{i=0}^{k-1}\mathbf{E}Z_{n,i}Z_{n,k}\right| \le c\sum_{p=1}^{k}\min\left(\delta_n l_n d_p,\ \frac{\varepsilon(n,p)}{n}\right)\cdot \tag{4.13}$$

We add equations (4.7), $\frac{1}{2}$(4.9) and (4.13) to obtain:

$$\left|\Delta^{(1)}_{k,n}(h)\right| \le c\left(M_n^2\delta_n + \delta_n M_n l_n\sum_{p=0}^{k-1}d_p + \sum_{p=1}^{k}\min\left(\delta_n l_n d_p,\ \frac{\varepsilon(n,p)}{n}\right)\right)\cdot \tag{4.14}$$

We sum (4.14) for all k to conclude:

$$\left|\sum_{k=0}^{n-1}\Delta^{(1)}_{k,n}(h)\right| \le c\times n\left(M_n^2\delta_n + \delta_n M_n l_n\sum_{p=0}^{\infty}d_p + \sum_{p=1}^{\infty}\min\left(\delta_n l_n d_p,\ \frac{\varepsilon(n,p)}{n}\right)\right)\cdot \tag{4.15}$$

With the techniques used to prove Lemma 3.3 we can prove

$$\lim_{n\to\infty} n\left(\sum_{p=1}^{\infty}\min\left(\delta_n l_n d_p,\ \frac{\varepsilon(n,p)}{n}\right)\right) = 0.$$

Replacing $M_n,\ \delta_n,\ l_n$ by their upper bounds, we easily see on (4.15) that

$$\sum_{k=0}^{n-1}\Delta^{(1)}_{k,n}(h) \xrightarrow[n\to\infty]{} 0.$$

This concludes the proof of Theorem 4.3. □

Remark 4.3. The proof of Theorem 4.3 extends immediately to the case $g_n(t) = \frac{1}{\sqrt{nb_n}}K(\frac{x-\psi(t)}{b_n})$, where $\psi : I \to \mathbb{R}$ is some monotone function.

Instead of $\mathcal{BV}$ we can also take a Banach space $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ whose norm satisfies:

- there exists $M > 0$ such that for all n in $\mathbb{N}$, $\|\ K\left(\frac{x-\cdot}{b_n}\right)\|_{\mathcal{B}} \le M$;
- $\sqrt{nb_n}\ \|\ K^2\left(\frac{x-\cdot}{b_n}\right)\|_{\mathcal{B}} \xrightarrow[n\to\infty]{} 0.$

Equation (4.9) and Lemma 3.2 still hold, thus we can prove Theorem 4.3.

Lipschitz norm does not yield the second point above. Hence we cannot replace the norm $||.||_{\mathcal{BV}}$ by the Lipschitz norm. Therefore the class of examples is not really large. It is a real problem which is due to the kernel density estimates whose Lipschitz norm has order $\mathcal{O}\left(\frac{1}{b_n}\right) \xrightarrow[n\to+\infty]{} +\infty.$

Proof of Theorems 4.1 and 4.2. For sake of simplicity, we write the proof for $l = 1$. We first write

$$\hat{f}_n(x) - f(x) = \left[\hat{f}_n(x) - \mathbf{E}\hat{f}_n(x)\right] + \left[\mathbf{E}\hat{f}_n(x) - f(x)\right].$$

The behaviour of $\hat{f}_n(x) - \mathbf{E}\hat{f}_n(x)$ is given by Theorem 4.3. Hence we restrict our attention to $\mathcal{BIAS}_n(x) = \mathbf{E}\hat{f}_n(x) - f(x)$.

The sequence $(X_n)_{n\in\mathbb{N}}$ is stationary because X_0 follows the invariant law μ_0.

Hence as $\int_D K(s)\mathrm{d}s = 1$,

$$\mathcal{BIAS}_n(x) = \int_D K(s)[f(x - sb_n) - f(x)]\mathrm{d}s. \tag{4.16}$$

Let ν denote the regularity of f on V.

Case $\nu \in \mathbb{N}^*$ (Th. 4.1):

As D is compact and as $b_n \xrightarrow[n\to\infty]{} 0$, there exists some $n_0 \in \mathbb{N}^*$ such that for $n \geq n_0$ the interval $J_{s,n} := [\min(x - sb_n, x),\ \sup(x - sb_n, x)]$ is included in V for all $s \in D$. Hence using (4.16) and Taylor's decomposition on each $J_{s,n}$ we get for $n \geq n_0$

$$\mathcal{BIAS}_n(x) = \int_D K(s)\left\{\sum_{j=1}^{m-1}\frac{(-sb_n)^j}{j\,!}\,f^{(j)}(x) + \frac{(-sb_n)^m}{m!}f^{(m)}(x - st_{s,n}b_n)\right\}\mathrm{d}s,$$

where for all $n \geq n_0$ and for all $s \in D$, $t_{s,n}$ is some real satisfying $0 < t_{s,n} < 1$.

Hence, as $\int_D s^j K(s)\mathrm{d}s = 0$ for all $1 \leq j \leq m-1$,

$$\mathcal{BIAS}_n(x) = \int_D K(s)\frac{(-sb_n)^m}{m!}\,f^{(m)}(x - st_{s,n}b_n)\mathrm{d}s.$$

Then as $f^{(m)}$ is continuous in x and as $0 < t_{s,n} < 1$ for all $s \in D$ and for all $n \geq n_0$, we have for each fixed $s \in D : t_{s,n}sb_n \xrightarrow[n\to\infty]{} 0$. Hence, proceeding as in the proof of Lemma 3.2 (study of Q_n), we get, by the dominated convergence theorem,

$$\int_D \frac{(-s)^m}{m!}K(s)f^{(m)}(x - st_{s,n}b_n)\mathrm{d}s \xrightarrow[n\to\infty]{} f^{(m)}(x)\int_D \frac{(-s)^m}{m!}K(s)\mathrm{d}s.$$

Hence, as soon as there exists $\rho_m \in \mathbb{R}^+$ such that $nb_n^{2m+1} \xrightarrow[n\to\infty]{} \rho_m$,

$$\sqrt{nb_n}\,\mathcal{BIAS}_n(x) \xrightarrow[n\to\infty]{} \frac{(-1)^m\sqrt{\rho_m}}{m!}\,f^{(m)}(x)\int_D s^m K(s)\mathrm{d}s.$$

This concludes the study of Theorem 4.1.

General case (Th. 4.2):

Here we use the integral form of Taylor's decomposition. Let $n_0 \in \mathbb{N}^*$ be such that for $n \geq n_0$ the interval $J_s := [\min(x - sb_n, x),\ \sup(x - sb_n, x)]$ is included in V for all $s \in D$. Recall that $\nu = \alpha + \beta$ where $\alpha \in \mathbb{N}$ and $0 \leq \beta < 1$.

Empty sums are set equal to 0.

For $n \geq n_0$,

$$\mathcal{BIAS}_n(x) = \int_D \left\{\sum_{j=1}^{\alpha-1}\frac{(-sb_n)^j}{j\,!}f^{(j)}(x) + \int_0^1 \frac{(-sb_n)^\alpha}{(\alpha-1)!}(1-t)^{\alpha-1}f^{(\alpha)}(x - stb_n)\mathrm{d}t\right\}K(s)\mathrm{d}s.$$

As $\int_D s^j K(s)\mathrm{d}s = 0$ for all $1 \le j \le \alpha$ we deduce:

$$\mathcal{BIAS}_n(x) = \int_D \int_0^1 \frac{(1-t)^{\alpha-1}}{(\alpha-1)!}(-b_n)^\alpha s^\alpha [f^{(\alpha)}(x - stb_n) - f^{(\alpha)}(x)]\mathrm{d}tK(s)\mathrm{d}s.$$

As f has the regularity $\nu = \alpha + \beta$ (in terms of Hölder spaces) on V, there exists some constant A independent of n such that:

$$|\mathcal{BIAS}_n(x)| \le \int_D \frac{1}{(\alpha-1)!}\, b_n^\alpha s^\alpha \; A\, b_n^\beta s^\beta \; |K(s)|\,\mathrm{d}s = \frac{A\, b_n^\nu}{(\alpha-1)!}\int_D s^\nu\,|K(s)|\,\mathrm{d}s.$$

As D is compact and as $\int_D K^2(s)\mathrm{d}s < \infty$, we have $\int_D s^\nu\,|K(s)|\,\mathrm{d}s < \infty$. Hence there exists some non-negative constant C independent of n such that:

$$|\mathcal{BIAS}_n(x)| \le C\, b_n^\nu.$$

So $\sqrt{nb_n}\;\mathcal{BIAS}_n(x) \xrightarrow[n\to\infty]{} 0$ as soon as $nb_n^{2\nu+1} \xrightarrow[n\to\infty]{} 0$. This concludes the proof of Theorem 4.2. □

5. Examples of dynamical systems in the class $\mathcal{T}$

Without being exhaustive we will now give some examples of dynamics T which satisfy all the previous assumptions.

- Lasota–Yorke functions

Let T be some piecewise smooth expanding map of the interval $[0,1]$. Following Viana ([27], Chap. 3), we introduce the following set of assumptions for T.

Assumption 5.1. *(regularity). There exists $0 = a_0 < a_1 < \cdots < a_l = 1$ such that the restriction of T to each $\eta_i = (a_{i-1}, a_i)$ is of class C^1, with $|T'(x)| > 0$ for all $x \in \eta_i$ and $i = 1, \dots, l$.*
Moreover, the function $g_{\eta_i} = \frac{1}{|T'_{\eta_i}|}$ has bounded variation for $i = 1, \dots, l$.

If h is some function on I and if $J \subset I$, h_J denotes the restriction of h to J.
Using this notation: (T_{η_i}) and g_{η_i} admit continuous extensions to $\bar{\eta}_i = [a_{i-1}, a_i]$ for each $i = 1, \dots, l$. Since modifying the values of a map over a finite set of points does not change its statistical properties, we may assume that T is either left-continuous or right-continuous (or both) at a_i, for each $i = 1, \dots, l$.
Then let $P^{(1)}$ be some partition of I into intervals η such that $\eta_i \subset \eta \subset \bar{\eta}_i$ for some i and such that (T_η) is continuous.
For $n \ge 1$, $P^{(n)}$ is the Markov partition of I: $P^{(n)}(x) = P^{(n)}(y)$ if and only if $P^{(1)}(T^j(x)) = P^{(1)}(T^j(y))$ for all $0 \le j < n$. ($P^{(n)}$ is the largest partition on which T^n is monotone.)
Given $\eta \in P^{(n)}$, denote $g_\eta^{(n)} = \frac{1}{|(T_\eta^n)'|}$.

Assumption 5.2. *(expansivity). There exist $C_1 > 0$ and $\lambda_1 < 1$ such that $\sup_{t\in\eta} g_\eta^{(n)}(t) \le C_1\lambda_1^n$ for all $\eta \in P^{(n)}$ and all $n \ge 1$.*

Assumption 5.3. *(topological mixing). There is an interval $I_* \subset I = [0,1]$ such that $T(I_*) = I_*$, every orbit $T^n(x)$, $x \in (0,1)$, eventually enters I_*, and T_{I_*} is topologically mixing: for each interval $J \subset I_*$ there is $n \ge 1$ such that $T^n(J) = I_*$.*

Lasota and Yorke [16], Liverani [17], Viana [27] and others study such functions. It may be shown (Viana [27]) that T admits a unique absolutely continuous invariant probability measure μ_0 ($\mathrm{d}\mu_0 = f\ \mathrm{d}t$ where $\mathrm{d}t$ is the Lebesgue measure on $[0,1]$). In addition, μ_0 is ergodic and its support coincides with I_*.

We also have that f has bounded variation on $[0,1]$. This implies that f is continuous on $[0,1]$ except at most for countably many points. So our assumptions in Section 1 are satisfied.

The "tent-maps" having a large enough slope and the r-adic transformations, for $r > 1$, of the interval are two examples of Lasota–Yorke functions.

- **"tent-map"**
  It has $T^{'}$ constant and strictly larger than 1 in absolute value, in each of the monotonicity intervals $[0,\ c)$ and $(c,\ 1]$.
  Moreover if $c = \frac{1}{2}$ and $|T'(x)| = \sigma > \sqrt{2}$ for all $x \neq c$, we have $I_* = [T^2(c),\ T(c)]$.

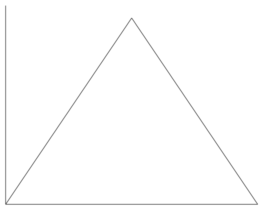

FIGURE 1. The "tent map" for $c = 0.5$ and $T(c) = 0.75$.

- **r-adic transformations** $(r > 1)$ of the interval $[0,\ 1]$. We have $I_* = I$.

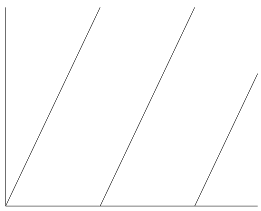

FIGURE 2. The r-adic map with $r = 8/3$.

- Functions with infinitely many monotonicity intervals. The precedent case extends, under appropriate conditions, to piecewise expanding maps with countably many domains of smoothness and monotonicity (see *e.g.* Broise [7], Viana [27]). For example if we consider the Gauss-map, that is the map T defined by $T(x) = \frac{1}{x} - [\frac{1}{x}]$ for $x \neq 0$ and $T(0) = 0$, we have summable decay of correlations and an invariant probability measure μ_0 absolutely continuous with respect to Lebesgue on $[0,1]$ and whose density has bounded variation. We have, keeping the former notation, $I_* = I$. Furthermore the Gauss-map satisfies the assumptions in Section 1. Therefore we have the result for the invariant density estimates in that case.

Remark 5.1. Let $T : [0,1] \to [0,1]$ defined by $T(x) = 4x(1-x)$, then Theorem 4.3 still holds because this map is obtained from a "tent map" by conjugation.

6. A non-stationary case

Lasota–Yorke functions T, introduced in Section 5, have additional properties which allow us to extend the previous results to the non-stationary case. We use the same definitions as before (*e.g.* (X_n) is the stationary dynamical system). Now let p be any density function on $I = [0,1]$ with bounded variation. We set $p(t) = 0$

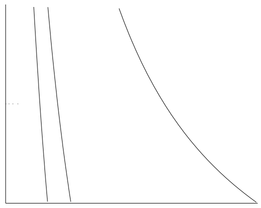

Figure 3. The Gauss-map.

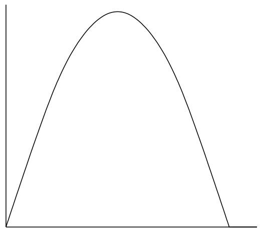

Figure 4. $T(x) = 4x(1-x)$.

for $t \notin I$. We define a random variable $X_0^{'}$ with distribution $p(t)dt$ and the (non-stationary) dynamical system $X_n^{'} = T^n X_0^{'}$, $n \geq 1$ (as in (1.6)).

6.1. Classical results

In the case of Lasota–Yorke function T, the invariant density f of T has bounded variation. We define the Perron–Frobenius operator $\mathcal{L}$ (for sake of simplicity we write $\mathcal{L}$ for $\mathcal{L}_T$) as follows:

$$\mathcal{L} : \begin{cases} \mathcal{BV} & \rightarrow & \mathcal{BV} \\ \omega & \mapsto & \mathcal{L}\omega(x) = \sum_{Ty=x} \dfrac{\omega(y)}{|T'(y)|} \end{cases}.$$

The following theorem collects properties of both the Perron–Frobenius operator $\mathcal{L}$ and the associated invariant density f of the Lasota–Yorke function T.

Theorem 6.1. (Liverani [17], Collet [8], Viana [27])

- $$h \in \mathcal{BV} \Longrightarrow \mathcal{L}^n h \in \mathcal{BV}\ \forall n \in \mathbb{N}. \textit{ Moreover } \sup_{n\in\mathbb{N}} ||\mathcal{L}^n h||_{\mathcal{BV}} < \infty. \tag{6.1}$$
- *If f denotes the invariant density, then there exists some $\gamma > 0$ such that:*
$$\frac{1}{\gamma} \leq f(t) \leq \gamma \textit{ for all } t \in I. \tag{6.2}$$
Then $\frac{1}{f}$ has also bounded variation.
- $$\exists\, R > 0,\ \exists\, 0 \leq \lambda < 1,\ \forall j \in \mathbb{N} : ||\mathcal{L}^j p - f||_\infty \leq R\, \lambda^j. \tag{6.3}$$
- *The correlations decrease exponentially fast. Hence there exists $\kappa > 0$ such that for any $h,\ k \in \mathcal{BV}$,*
$$|\mathrm{Cov}\,(h(X_0), k(X_n))\,| \leq \kappa\ \|k\|_1\ \|h\|_{\mathcal{BV}}\ \lambda^n \quad \forall n \geq 0, \tag{6.4}$$

where Cov *denotes the covariance with respect to the invariant probability measure* μ_0, *and* λ *is the same as in* (6.3).

Remark 6.1. The second assertion of Theorem 6.1 yields that if T is Lasota–Yorke, we can not have $f(x) = 0$.

6.2. Convergence in mean squares

Let us first recall the definition (1.7) of $\hat{p}(x)$:

$$\hat{p}(x) = \hat{p}_n(x) = \frac{1}{nb_n}\sum_{k=0}^{n-1} K\left(\frac{x - X_k'}{b_n}\right),$$

where b_n and K are defined in the introduction by (1.3). The following result extends Lemma 3.1 to non-stationary dynamical systems, I_1 and I_2 being defined as in Section 2.

$\lambda(D) < \infty$ still denotes the Lebesgue measure of the compact D.

Lemma 6.1. *Let T be a Lasota–Yorke function and $\hat{p}_n(x)$ be defined by (1.3, 1.6) and (1.7). Assume that $b_n \xrightarrow[n\to\infty]{} 0$. Then for $x \in I_1 \cap I_2$ we have*

$$\mathrm{Var}\,(\hat{p}(x)) = \frac{1}{nb_n}\left(f(x)\int_D K^2(s)\mathrm{d}s + o(1)\right).$$

Proof of Lemma 6.1. Write

$$(nb_n)\mathrm{Var}\,(\hat{p}(x)) = V_n + \frac{2}{nb_n}\sum_{1\le i<j\le n} C_{i,j} \tag{6.5}$$

with

$$V_n := \frac{1}{nb_n}\left(\sum_{k=0}^{n-1}\mathbf{E}K^2\left(\frac{x - X_k'}{b_n}\right) - \sum_{k=0}^{n-1}\left(\mathbf{E}K\left(\frac{x - X_k'}{b_n}\right)\right)^2\right) \tag{6.6}$$

and

$$C_{i,j} := \mathrm{Cov}\left(K\left(\frac{x - X_i'}{b_n}\right), K\left(\frac{x - X_j'}{b_n}\right)\right). \tag{6.7}$$

Note that $C_{i,j}$ depends on n.

- **Study of $\frac{2}{nb_n}\sum_{0\le i<j\le n-1} C_{i,j}$:**

$$\begin{aligned}
|C_{i,j}| &= \left|\int_0^1 K\left(\frac{x-t}{b_n}\right)K\left(\frac{x - T^{j-i}t}{b_n}\right)\mathcal{L}^i p(t)\mathrm{d}t - \int_0^1 K\left(\frac{x-t}{b_n}\right)\mathcal{L}^i p(t)\mathrm{d}t\int_0^1 K\left(\frac{x-t}{b_n}\right)\mathcal{L}^j p(t)\mathrm{d}t\right| \\
&\le \left|\int_0^1 K\left(\frac{x-t}{b_n}\right)K\left(\frac{x - T^{j-i}t}{b_n}\right)\mathcal{L}^i p(t)\mathrm{d}t - \int_0^1 K\left(\frac{x-t}{b_n}\right)\mathcal{L}^i p(t)\mathrm{d}t\int_0^1 K\left(\frac{x-t}{b_n}\right)f(t)\mathrm{d}t\right| \\
&\quad + \left|\int_0^1 K\left(\frac{x-t}{b_n}\right)\mathcal{L}^i p(t)\mathrm{d}t\int_0^1 K\left(\frac{x-t}{b_n}\right)\left(\mathcal{L}^j p(t) - f(t)\right)\mathrm{d}t\right| =: A_{n,i,j} + B_{n,i,j}.
\end{aligned}$$

By inequality (6.2) of Theorem 6.1, f is bounded below by $\frac{1}{\gamma} > 0$, so we can write:

$$A_{n,i,j} = \left| \int_0^1 K\left(\frac{x-t}{b_n}\right) K\left(\frac{x-T^{j-i}t}{b_n}\right) \frac{\mathcal{L}^i p(t)}{f(t)} f(t)\mathrm{d}t - \int_0^1 K\left(\frac{x-t}{b_n}\right) \frac{\mathcal{L}^i p(t)}{f(t)} f(t)\mathrm{d}t \int_0^1 K\left(\frac{x-t}{b_n}\right) f(t)\mathrm{d}t \right|.$$

We note that $A_{n,i,j} = |\mathrm{Cov}\left(h(X_0), k(X_{j-i})\right)|$, where $h : t \mapsto K(\frac{x-t}{b_n})\frac{\mathcal{L}^i p(t)}{f(t)}$ and $k : t \mapsto K(\frac{x-t}{b_n})$.

In the following c will denote some constant independent of n, i, and j which may vary from line to line.

Quote that there exists a constant c independent of n such that $||\mathcal{L}^n p||_\infty \le c$. Indeed, $||\mathcal{L}^n p||_\infty \le ||\mathcal{L}^n p||_{\mathcal{BV}}$ and by assertion (6.1) in Theorem 6.1, $\sup_{n\in\mathbb{N}} ||\mathcal{L}^n p||_{\mathcal{BV}} < \infty$.

We obtain $A_{n,i,j} \le c\, b_n \lambda^{j-i}$ by using Theorem 6.1.

Analogously, for $B_{n,i,j}$, we have

$$B_{n,i,j} \le \left| \int_0^1 K\left(\frac{x-t}{b_n}\right) \mathcal{L}^i p(t)\mathrm{d}t \right| \int_0^1 \left| K\left(\frac{x-t}{b_n}\right) \left(\mathcal{L}^j p(t) - f(t)\right) \right| \mathrm{d}t \le b_n^2\, c\, \lambda^j.$$

Using the bound $\lambda < 1$ now yields $A_{n,i,j} + B_{n,i,j} \le c\, b_n \lambda^{j-i}$.

Remark 6.2. Using similar arguments, we can prove that given any $h \in \mathcal{BV}$ and $k \in \mathbb{L}^1(dt)$ we have for all $0 \le i < j \le n-1$

$$|\mathrm{Cov}\left(h(X_i'), k(X_j')\right)| \le c\, ||h||_{\mathcal{BV}} ||k||_1 \lambda^{j-i}. \tag{6.8}$$

Using Theorem 6.1 and the proof of Lemma 3.2, we get

$$\begin{aligned} |C_{i,j}| &= \left| \int_0^1 K\left(\frac{x-t}{b_n}\right) K\left(\frac{x-T^{j-i}t}{b_n}\right) \mathcal{L}^i p(t)\mathrm{d}t - \int_0^1 K\left(\frac{x-t}{b_n}\right) \mathcal{L}^i p(t)\mathrm{d}t \int_0^1 K\left(\frac{x-t}{b_n}\right) \mathcal{L}^j p(t)\mathrm{d}t \right| \\ &\le b_n\, \varepsilon(n, j-i), \end{aligned}$$

where for $j-i$ fixed in $\mathbb{N}^*$, $\varepsilon(n, j-i) \xrightarrow[n\longrightarrow\infty]{} 0$.

Therefore

$$\left| \frac{2}{nb_n} \sum_{0\le i<j\le n-1} C_{i,j} \right| \le \frac{2\,c}{n} \sum_{0\le i<j\le n-1} \min\left(\lambda^{j-i}, \varepsilon(n, j-i)\right) \le c \sum_{k=1}^{n-1} \min\left(\lambda^k, \varepsilon(n,k)\right) \cdot \tag{6.9}$$

The right hand side of this inequality tends to 0 as n tends to infinity.

Remark 6.3. Using similar arguments, we can prove that for all $0 \le i < j \le n-1$, and for any bounded function φ,

$$\mathrm{Cov}\left(\varphi(X_0) \left(K\left(\frac{x - T^j X_0'}{b_n}\right) - \mathbf{E}K\left(\frac{x - T^j X_0'}{b_n}\right) \right), K\left(\frac{x - T^i X_0'}{b_n}\right) - \mathbf{E}K\left(\frac{x - T^i X_0'}{b_n}\right) \right) \le b_n\, \varepsilon(n, j-i),$$

with $\varepsilon(n, j-i) \xrightarrow[n\to\infty]{} 0$.

• **Study of V_n:**

Using $\mathrm{Var}_{\mathrm{ind}}\hat{f}(x)$ introduced in formula (3.2), it is worth decomposing V_n as follows:

$$V_n = (nb_n)\mathrm{Var}_{\mathrm{ind}}\hat{f}(x) + (s_n + s_n^{'}) := (nb_n)\mathrm{Var}_{\mathrm{ind}}\hat{f}(x) + \Big(V_n - (nb_n)\mathrm{Var}_{\mathrm{ind}}\hat{f}(x)\Big), \tag{6.10}$$

where

$$s_n := \frac{1}{nb_n}\left(\sum_{k=0}^{n-1}\mathbf{E}K^2\left(\frac{x-X_k^{'}}{b_n}\right) - \mathbf{E}K^2\left(\frac{x-X_k}{b_n}\right)\right)$$

and

$$s_n^{'} := \frac{1}{nb_n}\left(\sum_{k=0}^{n-1}\left(\mathbf{E}K\left(\frac{x-X_k^{'}}{b_n}\right)\right)^2 - \left(\mathbf{E}K\left(\frac{x-X_k}{b_n}\right)\right)^2\right).$$

We have

$$|s_n| \;=\; \left|\tfrac{1}{nb_n}\left(\sum_{k=0}^{n-1}\int_0^1 K^2\left(\frac{x-t}{b_n}\right)\left(\mathcal{L}^k p(t) - f(t)\right)\mathrm{d}t\right)\right|.$$

Hence, using (6.3) in Theorem 6.1 we get

$$|s_n| \le \frac{1}{n}\int_D K^2(s)\sum_{k=0}^{n-1}\left|\left(\mathcal{L}^k p(x-sb_n) - f(x-sb_n)\right)\right|\mathrm{d}s \le \frac{R\int_D K^2(s)\mathrm{d}s}{n}\sum_{k=0}^{n-1}\lambda^k. \tag{6.11}$$

The right hand side of this inequality tends to 0 as n tends to infinity.

Let $K_n' = K\left(\frac{x-X_n^{'}}{b_n}\right)$, $K_n = K\left(\frac{x-X_n}{b_n}\right)$. Then using the following identity

$$(\mathbf{E}K_n')^2 - (\mathbf{E}K_n)^2 = \left(\mathbf{E}\left(K_n' - K_n\right)\right)\left(\mathbf{E}\left(K_n' + K_n\right)\right),$$

we have by (6.1, 6.2) and (6.3) in Theorem 6.1:

$$\begin{aligned}|s_n^{'}| \;&=\; \left|\frac{1}{nb_n}\sum\nolimits_{k=0}^{n-1}\left(\int_0^1 K\left(\frac{x-t}{b_n}\right)\left(\mathcal{L}^k p(t) - f(t)\right)\mathrm{d}t\right)\left(\int_0^1 K\left(\frac{x-t}{b_n}\right)\left(\mathcal{L}^k p(t) + f(t)\right)\mathrm{d}t\right)\right| \\ &\le\; \frac{R}{n}\left(\sum\nolimits_{k=0}^{n-1}\lambda^k\right)\; b_n\;\left(\sup_n ||\mathcal{L}_n p||_\infty + \gamma\right)\;\lambda(D)\;\int_D K^2(s)\mathrm{d}s.\end{aligned}$$

As $0 \le \lambda < 1$ and as $\int_D K^2(s)\mathrm{d}s < \infty$, the right hand side of this inequality tends also to 0 as n tends to infinity.

Hence by (6.10, 6.11) and (6.2),

$$V_n \sim (nb_n)\;\mathrm{Var}_{\mathrm{ind}}\hat{f}(x) \sim f(x)\int_D K^2(t)\mathrm{d}t > 0 \;\text{ as } n \text{ tends to infinity.} \tag{6.12}$$

Collecting (6.5) and bounds (6.9) and (6.12) yields the result:

$$(nb_n)\mathrm{Var}\left(\hat{p}(x)\right) \xrightarrow[n\to+\infty]{} f(x)\int_D K^2(t)\mathrm{d}t.$$

□

6.3. **Central Limit Theorem**

Theorems 6.2, 6.3 and 6.4 below deal with the non-stationary case and are analogous to Theorems 4.1, 4.2 and 4.3 of Section 4. The setting is not really different from the stationary one. We first study the following process

$$U_n^{'}(x) := \sqrt{nb_n}(\hat{p}(x) - f(x)). \tag{6.13}$$

Theorem 6.2. *Let T be a Lasota–Yorke function and $\hat{p}_n(x)$ be defined by (1.3, 1.6) and (1.7). Assume that $b_n \xrightarrow[n\to+\infty]{} 0$, $nb_n \xrightarrow[n\to+\infty]{} +\infty$. Let l be a positive integer. For all $1 \le i \le l$, let $x_i \in I_1 \cap I_2$. Let m be a positive integer. Assume that for each i, $1 \le i \le l$, there exists a neighbourhood V_i of x_i such that the invariant density f is $m-$times continuously differentiable on V_i. Also assume that $\int_D s^j K(s)\mathrm{d}s = 0$ for all $1 \le j \le m-1$ and that nb_n^{2m+1} converges to some non-negative constant ρ_m as n tends to infinity. Then*

$$(U_n'(x_1),\ \dots,\ U_n'(x_l)) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(\frac{(-1)^m\ \sqrt{\rho_m}}{m!}\ \int_D s^m\ K(s)\mathrm{d}s * \left(f^{(m)}(x_i)\right)_{1\le i\le l},\ \Sigma_l\right),$$

where Σ_l is defined by (4.2).

Let us now consider the case where the invariant density f has a regularity $\nu = \alpha + \beta$ with $0 \le \beta < 1$ and $\alpha \in \mathbb{N}$.

Theorem 6.3. *Let T be a Lasota–Yorke function and $\hat{p}(x)$ be defined by (1.3, 1.6) and (1.7). Assume that $b_n \xrightarrow[n\to+\infty]{} 0$, $nb_n \xrightarrow[n\to+\infty]{} +\infty$. Let l be a positive integer. For all $1 \le i \le l$, let $x_i \in I_1 \cap I_2$. Let $\nu = \alpha + \beta$ with $0 \le \beta < 1$ and $\alpha \in \mathbb{N}$. Assume that for each i, $1 \le i \le l$, there exists a neighbourhood V_i of x_i such that the invariant density f has the regularity ν on V_i. Also assume that $\int_D s^j K(s)\mathrm{d}s = 0$ for all integer j such that $1 \le j \le \alpha$ and that $nb_n^{2\nu+1} \xrightarrow[n\to+\infty]{} 0$. Then*

$$(U_n'(x_1),\ \dots,\ U_n'(x_l)) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0,\ \Sigma_l\right),$$

where Σ_l is defined by (4.2).

As in the stationary case, we sometimes prefer to study the centered estimation process

$$Y_n^{'}(x) := \sqrt{nb_n}(\hat{p}(x) - \mathbf{E}\hat{p}(x)). \tag{6.14}$$

We then get:

Theorem 6.4. *Let T be a Lasota–Yorke function and $\hat{p}(x)$ be defined by (1.3, 1.6) and (1.7). For all $1 \le i \le l$, let $x_i \in I_1 \cap I_2$. Assume that $b_n \xrightarrow[n\to\infty]{} 0$, $nb_n \xrightarrow[n\to\infty]{} \infty$. Then the finite dimensional marginals $(\overline{Y_n'}(x_1),\ \dots,\ \overline{Y_n'}(x_l))$ of the process*

$$\overline{Y_n'}(x) \equiv \frac{Y_n^{'}(x)}{\sqrt{f(x)\int_{-\infty}^{\infty} K^2(t)\mathrm{d}t}}$$

converge in distribution to a standard $\mathcal{N}(0,\ I_l)$ random variable.

Remark 6.4. Here we have normalized the process $Y_n^{'}(x)$. It is possible as $f(x) > 0$ for Lasota–Yorke functions T (see Th. 6.1).

As in the stationary case, we first prove Theorem 6.4 and then deduce Theorems 6.2 and 6.3 by studying the bias term. For the proof of Theorem 6.4 we develop the proof for $l = 1$ for sake of simplicity. If one wants to know the asymptotic behaviour in distribution of the vector $(Y_n^{'}(x_1),\ \dots,\ Y_n^{'}(x_l))$, it is sufficient to use the following proof (in the case $l=1$) with $\frac{1}{\sqrt{nb_n}}\sum_{j=1}^{l} s_j K(\frac{x_j-X_k^{'}}{b_n})$, for arbitrary numbers $s_1, \dots, s_l \in \mathbb{R}$, instead of $\frac{1}{\sqrt{nb_n}} K(\frac{x-X_k^{'}}{b_n})$.

Let us first give two applications of Theorem 6.4.

Application 6.1 Let r be a positive integer. For all $1 \le i \le r$ we define

$$\hat{p}_n^{(i)}(x) = \frac{1}{nb_n}\sum_{k=0}^{n-1} K\left(\frac{x - {X_k^{'}}^{(i)}}{b_n}\right)$$

where for all n, $\left({X_0^{'}}^{(i)}, \dots, {X_{n-1}^{'}}^{(i)}\right) \sim \left(X_0^{'}, \dots, X_{n-1}^{'}\right)$. We assume that the sequences $X^{(i)} := \left({X_k^{'}}^{(i)}\right)_{k\in\mathbb{N}}$, for $i \in \mathbb{N}$, are independent of each other. Hence $\hat{p}_n^{(0)}(x), \hat{p}_n^{(1)}(x), \dots, \hat{p}_n^{(r)}(x)$ are $r+1$ independent copies of $\hat{p}_n(x)$, and we can consider

$$\hat{Y}_n^{'}(x) := \sqrt{nb_n}\left\{\hat{p}_n^{(0)}(x) - \frac{\hat{p}_n^{(1)}(x) + \dots + \hat{p}_n^{(r)}(x)}{r}\right\}.$$

Let assume that $b_n \xrightarrow[n\to\infty]{} 0$, $nb_n \xrightarrow[n\to\infty]{} \infty$. We also assume that r depends on n, $r = r(n)$, with $\frac{r(n)}{n} \le C$ where C is some positive constant and $r(n)b_n \xrightarrow[n\to\infty]{} \infty$. For example take $r(n) = n$. For all i and for large n,

$$\left|\hat{p}_n^{(i)}(x) - \mathbf{E}\hat{p}_n(x)\right| \le \frac{2\,||K||_\infty}{b_n}$$

and

$$\begin{aligned} \textstyle\sum_{i=1}^{r} \mathrm{Var}\left(\hat{p}_n^{(i)}(x)\right) &= r\ \mathrm{Var}\left(\hat{p}_n(x)\right) = \frac{r(n)}{nb_n}\left(f(x)\int_D K(s)\mathrm{d}s + o_n(1)\right) \\ &\le \tfrac{C}{b_n}\left(f(x)\int_D K(s)\mathrm{d}s + 1\right). \end{aligned}$$

Hence using Bernstein's inequality in Pollard ([20], pp. 192, 193) we get for all $\eta > 0$

$$P\left(\left|\tfrac{\hat{p}_n^{(1)}(x)+\dots+\hat{p}_n^{(r)}(x)}{r} - \mathbf{E}\hat{p}_n(x)\right| > \eta\right) \quad \le \quad 2\exp\left(\frac{-\eta^2\, r^2\, b_n}{2\,C\,\left(f(x)\,\int_D K^2(s)\mathrm{d}s+1\right)+\frac{4}{3}\,\eta\, r\,||K||_\infty}\right).$$

The exponential term above tends to 0 as n tends to infinity. Hence, we can approach $Y_n^{'}(x)$ by the empirical quantity $\hat{Y'}_n(x)$. The advantage of $\hat{Y'}_n(x)$ is that it can be simulated. Indeed it does not involve the knowledge of $f(x)$.

Application 6.2 Now let $\hat{p}_n^{(1)}(x)$ and $\hat{p}_n^{(2)}(x)$ be two independent copies of $\hat{p}_n(x)$. The difference

$$\Phi_n(x) = \sqrt{nb_n}\left(\{\hat{p}_n^{(1)}(x) - \mathbf{E}\hat{p}_n^{(1)}(x)\} - \{\hat{p}_n^{(2)}(x) - \mathbf{E}\hat{p}_n^{(2)}(x)\}\right)$$

does not depend on $\mathbf{E}\hat{p}_n(x)$. Indeed $\Phi_n(x) = \sqrt{nb_n}\left(\hat{p}_n^{(1)}(x) - \hat{p}_n^{(2)}(x)\right)$.

Moreover, $\Phi_n(x)$ converges in distribution to a $\mathcal{N}(0, 2\,f(x)\int_D K^2(s)\mathrm{d}s)$ random variable as n tends to infinity as soon as $b_n \xrightarrow[n\to\infty]{} 0$, $nb_n \xrightarrow[n\to\infty]{} \infty$. To estimate $f(x)$ it can be useful to work with $\Phi_n(x)$ instead of $U_n^{'}(x)$

as $f(x)$ only appears in the variance term of the limit of $\Phi_n(x)$ and not in the quantity $\Phi_n(x)$ itself. Hence to simulate $\Phi_n(x)$ we have neither to approach $\mathbf{E}\hat{p}_n(x)$ using exponential inequalities as in Application 6.1 nor to know $f(x)$ (as $X_0^{'}$ has the distribution $p(t)\,\mathrm{d}t$ where p is known).

Proof of Theorem 6.4 with $l = 1$. We use notations g_n, M_n, l_n, and δ_n of Theorem 4.3. As in the proof of Theorem 4.3 we obtain $M_n \le \frac{c}{\sqrt{nb_n}}$, $l_n \le \frac{c}{\sqrt{nb_n}}$ and $\delta_n \le \frac{c\, b_n}{\sqrt{nb_n}}$, for some positive constant c.

We set for $k = 0, \dots, n-1$ and $n = 1, 2, \dots,$

$$Z_{n,k}^{'} = g_n(X_{n-k-1}^{'}) - \mathbf{E}\left(g_n(X_{n-k-1}^{'})\right), \text{ and } S_n^{'} = Z_{n,0}^{'} + \cdots + Z_{n,n-1}^{'}.$$

Now let $S_{k,n}^{'} = Z_{n,0}^{'} + \cdots + Z_{n,k}^{'}$ for $0 \le k \le n-1$. Empty sums are, as usual, set equal to 0.

Recall that

$$\lim_{n\to\infty} \operatorname{Var} S_n = f(x) \int_D K^2(s)\mathrm{d}s > 0. \tag{6.15}$$

We still consider a bounded thrice differentiable function $h : \mathbb{R} \to \mathbb{R}$ with continuous and bounded derivatives, with $C_j = \|h^{(j)}\|_\infty$, for $j = 0, 1, 2, 3$. As in the proof of Theorem 4.3, also consider $\sigma_n^2 = \operatorname{Var} S_n$, and set in that case for some standard Gaussian r.v. η, $\Delta_n(h) = \mathbf{E}(h(S_n^{'}) - h(\sigma_n \eta))$. As in the proof of Theorem 4.3, the theorem will follow if we prove that $\lim_{n\to\infty} \Delta_n(h) = 0$.

Let us assume $\{Y_{n,k};\ 0 \le k \le n-1,\ n \ge n_0\}$ to be defined as in Section 3. $T_{n,k}$ is also defined as before. We are now in position to use Rio's decomposition

$$\Delta_n(h) = \sum_{k=0}^{n-1} \Delta_{k,n}(h),$$

with $\Delta_{k,n}(h) = \mathbf{E}(h(S_{k-1,n}^{'} + Z_{n,k}^{'} + T_{n,k}) - h(S_{k-1,n}^{'} + Y_{n,k} + T_{n,k}))$.

We still use the function $x \to h_{k,n}(x) = \mathbf{E}h(x + T_{n,k})$, which has the same derivability properties as the function h.

We proceed as in Section 4.

Inequality (6.8) replaces inequality (1.5).

For example inequality (4.12) in the proof of Theorem 4.3 is replaced by

$$|\mathbf{E}h_{k,n}^{''}(S_{k-1,n}^{'})\mathbf{E}Z_{n,i}^{'}Z_{n,k}^{'}| \le c \min\left(\delta_n l_n \lambda^{k-i},\ \frac{\varepsilon(n, k-i)}{n}\right).$$

Hence doing this with each inequality of the proof of Theorem 4.3 we conclude the proof of Theorem 6.4. □

Let us now prove Theorems 6.2 and 6.3.

Proof of Theorems 6.2 and 6.3. We have the following decomposition:

$$\hat{p}_n(x) - f(x) = \hat{p}_n(x) - \mathbf{E}\hat{p}_n(x) + \mathbf{E}\hat{p}_n(x) - \mathbf{E}\hat{f}_n(x) + \mathbf{E}\hat{f}_n(x) - f(x).$$

The term $\hat{p}_n(x) - \mathbf{E}\hat{p}_n(x)$ is studied in Theorem 6.4. The term $\mathbf{E}\hat{f}_n(x) - f(x) = \mathcal{BIAS}_n(x)$ is studied in the proof of Theorems 4.1 and 4.2. It does not depend on the density p. So we just have to study the term $\mathbf{E}\hat{p}_n(x) - \mathbf{E}\hat{f}_n(x)$. We have

$$\mathbf{E}\hat{p}_n(x) - \mathbf{E}\hat{f}_n(x) = \frac{1}{n} \int_{\frac{x-R}{b_n}}^{\frac{x-L}{b_n}} K(s) \sum_{k=0}^{n-1} \left[\mathcal{L}^k p(x - sb_n) - f(x - sb_n)\right] \mathrm{d}s. \tag{6.16}$$

Using inequality (6.3) in Theorem 6.1, equality (6.16) and $\int_D K^2(s)\mathrm{d}s < \infty$ we get

$$\left|\mathbf{E}\hat{p}_n(x) - \mathbf{E}\hat{f}_n(x)\right| \leq \frac{R}{n}\sum_{k=0}^{n-1}\lambda^k\sqrt{\int_D K^2(s)\mathrm{d}s}\,\sqrt{\lambda(D)} \leq \frac{R}{n}\,\frac{1}{1-\lambda}\sqrt{\int_D K^2(s)\mathrm{d}s}\,\sqrt{\lambda(D)} \tag{6.17}$$

where $0 \leq \lambda < 1$. As $b_n \xrightarrow[n\to\infty]{} 0$ and as $\int_D K^2(s)\mathrm{d}s < \infty$, inequality (6.17) yields $\sqrt{nb_n}\,\left|\mathbf{E}\hat{p}_n(x) - \mathbf{E}\hat{f}_n(x)\right| \xrightarrow[n\to\infty]{} 0$. It concludes the proof of Theorems 6.2 and 6.3. □

Remark 6.5. Notice that $\sup_{t\in I}|\mathcal{L}^n 1(t) - f(t)|$ tends to 0 exponentially fast, hence $\mathcal{L}^n 1$ also appears to be a good evaluation of f. Unfortunately explicit computations of $\mathcal{L}^n 1$ involve a complete knowledge of iterated preimages with respect to T. Example given for $r \in \mathbb{N}^*$ the $r-$adic transformation involves r^n such preimages.

Appendix A. Appendix

Extension to weak dependent sequences

We extend here the results of this paper to weak dependent sequences. Let us first introduce our dependence frame which is a variation on the definition in Doukhan and Louhichi [13]. Assume that, for convenient functions h and k,

$$\mathrm{Cov}\,(h(\text{“past”}), k(\text{“future”}))$$

converges to 0 as the distance between the “past” and the “future” converges to infinity. Here “past” and “future” refer to the values of some time series of interest. Asymptotically, this means that independence holds if we use a *determining* function class.

More precisely, E being some Euclidean space $\mathbb{R}^d$ endowed with its Euclidean norm $\|.\|$, we shall consider a sequence of E-valued random variables $(\xi_n)_{n\in\mathbb{N}}$. We define $\mathbb{L}^\infty$ as the set of measurable and bounded numerical functions on some space $\mathbb{R}^k$, $k \in \mathbb{N}^*$ and its norm is classically written $\|\bullet\|_\infty$.

Moreover, let $u \in \mathbb{N}^*$ be a positive integer. We endow the set $F = E^u$ with the norm

$$\|(x_1, \ldots, x_u)\|_F = \|x_1\| + \cdots + \|x_u\|.$$

Let now $h : F = E^u \to \mathbb{R}$ be a numerical function on F, we set

$$\mathrm{Lip}(h) = \sup_{x\neq y}\frac{|h(x) - h(y)|}{\|x-y\|_F}$$

the Lipschitz modulus of h. Define

$$\mathcal{L} = \bigcup_{u=1}^{\infty}\{h \in \mathbb{L}^\infty(E^u, \mathbb{R});\ \|h\|_\infty \leq 1, \mathrm{Lip}(h) < \infty\}\cdot \tag{A.1}$$

Definition A.1. The sequence $(\xi_n)_{n\in\mathbb{N}}$ is s-weakly (resp. a-weakly) dependent, if for some sequence $\theta = (\theta_r)_{r\in\mathbb{N}}$ decreasing to zero at infinity and any (u+1)-tuple $(i_1, \ldots, i_u, j_1)$ with $i_1 \leq \ldots \leq i_u < i_u + r \leq j_1$, for $h \in \mathbb{L}^\infty$ satisfying $\|h\|_\infty \leq 1$ and for $k \in \mathcal{L}$,

$$|\mathrm{Cov}\,(h(\xi_{i_1}, \ldots, \xi_{i_u}), k(\xi_{j_1}))| \leq \mathrm{Lip}(k)\theta_r, \tag{A.2}$$

and respectively for $h, k \in \mathcal{L}$

$$|\mathrm{Cov}\,(h(\xi_{i_1}, \ldots, \xi_{i_u}), k(\xi_{j_1}))| \leq \mathrm{Lip}(h)\mathrm{Lip}(k)\theta_r. \tag{A.3}$$

The results presented in this appendix improve CLTs stated by Doukhan and Louhichi in a more general non-causal frame (see [13]). We work here indeed under a fundamental causality assumption. Contrarly to Doukhan and Louhichi [13], we do not use Bernstein blocks but a variation on the Lindeberg–Rio method. We also relax assumptions in Coulon–Prieur and Doukhan [9]. Indeed, in [9], the authors need two points in the future. Here we just consider one point in the future ξ_{j_1}.

Note that the notions of weak dependence and dynamical systems are not that much different. For example let us define the autoregressive model by:

$$\xi_n = T(\xi_{n-1}) + \eta_n, \tag{A.4}$$

with $T : \mathbb{R} \to \mathbb{R}$ such that $|T(u) - T(u')| \leq c|u - u'|$ for some $0 \leq c < 1$ and for all $u,\ u' \in \mathbb{R}$, and with $(\eta_n)_{n\in\mathbb{Z}}$ some real valued i.i.d inovation process satisfying $\mathbf{E}|\eta_0| < \infty$. This model is s-weakly dependent. A generalization of this model is given by:

$$X_{n+1} = F(X_n, \varepsilon_{n+1}),$$

with $(\varepsilon_i)_{i\in\mathbb{N}}$ a sequence of independent random variables (r.v.s) and with F a measurable function. Such Markov chains are actually noisy dynamical systems (see Baladi *et al.* [3]).

We refer to [9] for further examples of weak dependent sequences.

Density estimation in the case of weak dependence

We are now going to extend Theorem 4.3.

$Y_n(x)$ is defined as in Section 4, that is

$Y_n(x) := \sqrt{nb_n}\left(\hat{f}(x) - \mathbf{E}\hat{f}(x)\right)$, where $\hat{f}_n(x) = \frac{1}{nb_n}\sum_{k=1}^n K\left(\frac{x-\xi_k}{b_n}\right)$, and K is supposed to be Lipschitz.

Theorem A.1. *Assume that the previous $s-$weak dependence (resp. $a-$) condition holds for the stationary real valued sequence $(\xi_n)_{n\in\mathbb{N}}$ with for some positive $a < \frac{1}{3}$ (resp. $a < \frac{1}{4}$) $\sum_{p=1}^\infty \theta_p^a < \infty$, then the finite dimensional marginals $(\overline{Y_n}(x_1), \ldots, \overline{Y_n}(x_l))$, of the process $\overline{Y_n}(x) \equiv Y_n(x)/\sqrt{f(x)\int_{-\infty}^\infty u^2(t)\mathrm{d}t}$ converge in distribution to an $\mathcal{N}(0, I_l)$ random variable if we assume moreover that $f(x_1) \neq 0, \ldots, f(x_l) \neq 0$, that ξ_0's marginal admits a continuous marginal density f and the marginal densities $f_k(x,y)$ of the bivariate random variables (ξ_0, ξ_k) exist for any $k > 0$ and satisfy $\sup_{k>0}\sup_{(x,y)\in\mathbb{R}^2} f_k(x,y) < \infty$.*

Remarks.

- Here we need the existence of marginal densities $f_k(x,y)$ of the bivariate random variables (ξ_0, ξ_k). It is a classical assumption in that frame. We recall that in the case of dynamical systems, such densities are singular. The strong estimate

$$\mathrm{Cov}\left(K\left(\frac{x-\xi_j}{b_n}\right),\ K\left(\frac{x-\xi_i}{b_n}\right)\right) \leq C\ b_n^2$$

is standard under this condition while in the dynamical case, we just use that

$$\mathrm{Cov}\left(K\left(\frac{x-T^jX_0}{b_n}\right),\ K\left(\frac{x-T^iX_0}{b_n}\right)\right) \leq C\ b_n\ \varepsilon(n, j-i)$$

for a sequence $\varepsilon(n,k) \xrightarrow[n\to\infty]{} 0$ for any k. We also replace the summable decay of correlations (see (1.5)) of dynamical systems in the class $\mathcal{T}$ by a weak dependence condition.

Furthermore, in the case of stationary dynamical systems, we do not have any reason to suppose that

$$\sup_{1\leq k\leq n}\frac{1}{b_n}\mathrm{Cov}\left(K\left(\frac{x-X_0}{b_n}\right),\ K\left(\frac{x-X_k}{b_n}\right)\right)$$

tends to 0 as n tends to infinity, so using our estimates it appears to be hopeless to consider the Lipschitz or an Hölder norms as in the case of weak dependence instead of the norm $||\bullet||_{\mathcal{BV}}$ without adding assumptions on the sequence $(b_n)_{n\in\mathbb{N}}$.

- The conditions hold respectively if $\theta_r = \mathcal{O}(r^{-a})$ for some $a > 3$ (resp. $a > 4$).
- This result improves on a previous result in Doukhan and Louhichi [14], *e.g.* under association we need $\mathrm{Cov}(\xi_0, \xi_r) = \mathcal{O}(r^{-a})$ for $a > 5$ while the previous result was obtained assuming $a > 12$ and for causal shifts it was needed that $\theta_r = \mathcal{O}(r^{-a})$ for some $a > \max\{9, \frac{3}{2}(1+\delta^{-1})\}$ if $b_n \sim n^{-\delta}$.
- For strongly mixing sequences, the condition $\alpha_n = \mathcal{O}(n^{-a})$ for $a > 1$ ensures this CLT as proved by Robinson [24] (and also Ango–Nze and Doukhan [2]); this asumption is of a different nature, *e.g.* linear processes satisfy mixing conditions (under additional regularity conditions, see Doukhan ([11], Chap. 2.3). The decay rate of the coefficients are there more restrictive.

The proof of Theorem A.1 is a variation on the proof of Theorem 4.3. We refer to Coulon–Prieur and Doukhan [9], where the proof is written under stronger assumptions (in terms of dimension of the "future"). The same techniques provide some general limit theorem for triangular arrays under weak dependence (see [9]).

The author is grateful to the anonymous referees for constructive comments.

References

[1] A. Amroun, *Systèmes dynamiques perturbés. Sur une classe de fonctions zéta dynamiques*, Thèse de Doctorat de l'Université Paris 6, Spécialité Mathématique (1995).

[2] P. Ango Nze and P. Doukhan, Non-parametric Minimax estimation in a weakly dependent framework I: Quadratic properties. *Math. Methods Statist.* **5-4** (1996) 404-423.

[3] V. Baladi, M. Benedicks and V. Maume–Deschamps, Almost sure rates of mixing for i.i.d. unimodal maps. *Ann. E.N.S.* (to appear).

[4] A.D. Barbour, R.M. Gerrard and G. Reinert, Iterates of expanding maps. *Probab. Theory Related Fields* **116** (2000) 151-180.

[5] D. Bosq and D. Guégan, Nonparametric estimation of the chaotic function and the invariant measure of a dynamical system. *Statist. Probab. Lett.* **25** (1995) 201-212.

[6] D. Bosq and J.P. Lecoutre, *Théorie de l'estimation fonctionnelle.* Collection "Économie et statistiques avancées". Série : École Nationale de la Statistique et de l'Administration Économique et Centre d'Études des Programmes Economiques". Economica (1987).

[7] A. Broise, F. Dal'bo and M. Peigné, Études spectrales d'opérateurs de transfert et applications. *Astérisque* **238** (1996) Société Math. de France.

[8] P. Collet, Some ergodic properties of maps of the interval, in *dynamical systems*, edited by R. Bamon, J.M. Gambaudo and S. Martinez. Hermann, Paris (1996).

[9] C. Coulon–Prieur and P. Doukhan, A triangular central limit Theorem under a new weak dependence condition. *Statist. Probab. Lett.* **47** (2000) 61-68.

[10] W. De Melo and S. Van Strien, *One-Dimensional Dynamics.* Springer-Verlag (1993).

[11] P. Doukhan, *Mixing: Properties and Examples.* Springer Verlag, *Lecture Notes in Statist.* **85** (1994).

[12] P. Doukhan, *Models, Inequalities and Limit Theorems for Stationary Sequences*, edited by P. Doukhan, G. Oppenheim and M. Taqqu. Birkhaüser (to appear).

[13] P. Doukhan and S. Louhichi, A new weak dependence condition and applications to moment inequalities. *Stochastic Process. Appl.* **84** (1999) 313-342.

[14] P. Doukhan and S. Louhichi, Functional estimation of a density under a new weak dependence condition. *Scand. J. Statist.* **28** (2001) 325-342.

[15] A. Lasota and M. Mackey, *Probabilistic properties of deterministic systems.* Cambridge University Press (1985).

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

[17] C. Liverani, Decay of correlations for piecewise expanding maps. *J. Statist. Phys.* **78** (1995) 1111-1129.

[18] C. Liverani, Central limit Theorem for deterministic systems, in *Proc. of the international Congress on Dynamical Systems*, Montevideo 95. Pittman, *Res. Notes Math.* (1997).

[19] J. Maës, *Statistique non paramétrique des processus dynamiques réels en temps discret.* Thèse de l'Université Paris 6 (1999).

[20] D. Pollard, *Convergence of Stochastic Processes.* Springer Verlag, *Springer Ser. Statist.* (1984).

[21] R. Prakasa, *Nonparametric functional estimation.* Academic Press, New York (1983).

[22] E. Rio, About the Lindeberg method for strongly mixing sequences. *ESAIM: PS* **1** (1995) 35-61. www.emath.fr/ps

[23] E. Rio, Sur le théorème de Berry–Esseen pour les suites faiblement dépendantes. *Probab. Theory Related Fields* **104** (1996) 255-282.
[24] P.M. Robinson, Non parametric estimators for time series. *J. Time Ser. Anal.* **4-3** (1983) 185-207.
[25] M. Rosenblatt, Stochastic curve estimation, in *NSF-CBMS Regional Conference Series in Probability and Statistics*, Vol. 3 (1991).
[26] W. Rudin, *Real and complex analysis*. McGraw-Hill Series in Higher Mathematics, Second Edition (1974).
[27] M. Viana, *Stochastic dynamics of deterministic systems, Instituto de Matematica Pura e Aplicada.* IMPA, Vol. 21 (1997).