



**HAL**  
open science

# Exponential inequalities and functional estimations for weak dependent datas; applications to dynamical systems.

Véronique Maume-Deschamps

► **To cite this version:**

Véronique Maume-Deschamps. Exponential inequalities and functional estimations for weak dependent datas; applications to dynamical systems.. Stoch. Dyn., 2006, pp.535. 10.1142/S0219493706001876 . hal-00022453

**HAL Id: hal-00022453**

**<https://hal.science/hal-00022453v1>**

Submitted on 9 Apr 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**EXPONENTIAL INEQUALITIES AND FUNCTIONAL  
ESTIMATIONS FOR WEAK DEPENDENT DATAS ;  
APPLICATIONS TO DYNAMICAL SYSTEMS.**

V. MAUME-DESCHAMPS

ABSTRACT. We estimate density and regression functions for weak dependant datas. Using an exponential inequality obtained in [DeP] and in [Mau2], we control the deviation between the estimator and the function itself. These results are applied to a large class of dynamical systems and lead to estimations of invariant densities and on the mapping itself.

Dynamical systems are widely used by scientists to modalize complex systems ([ABST]). Therefore, estimating functions related to dynamical systems is crucial. Of particular interest are : the invariant density, the mapping itself, the pressure function. We shall see that many dynamical systems have the same behavior as weak dependant processes (as defined in [DoLo]). We obtain results of deviation for regression functions and densities for weak dependant processes and apply these results to dynamical systems.

In [Mau2] we gave an estimation (with control of the deviation) of the pressure function for some expanding maps of the interval. Results on the estimation of the invariant density for the same kind of maps were obtained in [P] and stated in [Mae]. In this later article, results on the estimation of the mapping were also stated. These last two papers dealt with convergence in quadratic mean. Our goal here is, on one hand, to consider more general dynamical systems : non uniformly hyperbolic maps on the interval, dynamics in higher dimension ... On the second hand, we obtain bounds on the deviation between the estimator and the regression function as well as almost sure convergence. Related results on the estimation of the regression function may also be found in [FV] and [Mas] where strongly mixing processes are considered and almost everywhere convergence and asymptotic normality are proved. Our aim is to provide an estimation of the deviation between the estimator and the regression function for a larger class of mixing processes and for regression functions that may have singularities.

Before giving the precise definitions and results, let us state our main results informally. We consider a weak dependant stationary process  $X_0, \dots, X_i, \dots$  taking values in  $\Sigma \subset \mathbb{R}^d$ . Our condition of weak dependence is with respect to a Banach space  $\mathcal{C}$  of bounded functions on  $\Sigma$  (see Definition 1 below). Let  $(Y_i)_{i \in \mathbb{N}}$  be a stationary process taking values in  $\mathbb{R}$  and satisfying a condition of weak dependence according with the process  $X_0, \dots, X_i, \dots$ . Consider the

---

2000 *Mathematics Subject Classification.* 37A50, 60E15, 37D20.

*Key words and phrases.* Exponential inequalities, functional estimation, dynamical systems, weak dependance.

regression function  $r(x) = \mathbb{E}(Y_i | X_i = x)$ , we shall assume some regularity on  $r$  (see Assumption 2 below). We shall also consider that the process  $(X_i)_{i \in \mathbb{N}}$  has a density  $f$  with respect to the Lebesgue measure  $m$ .  $f$  is assumed to have some regularity properties that allow localised singularities. Consider a non negative kernel  $K \in \mathcal{C}$  satisfying Assumption 1 and the estimators (introduced in [R] and used for example in [P], [Mae], [FV], [Mas]) :

$$(0.1) \quad \begin{aligned} \hat{f}_n(x) &= \frac{1}{nh} \sum_{i=0}^{n-1} K\left(\frac{x - X_i}{h}\right) & \hat{g}_n(x) &= \frac{1}{nh} \sum_{i=0}^{n-1} Y_i K\left(\frac{x - X_i}{h}\right) \\ \hat{r}_n(x) &= \frac{\hat{g}_n(x)}{\hat{f}_n(x)}. \end{aligned}$$

Remark that  $\hat{f}_n(x) = 0$  implies  $\hat{g}_n(x) = 0$ , in that case, we define  $\hat{r}_n(x) = 0$ .

**Theorem.** *Assume some weak dependence condition on  $X_i$  and  $Y_i$  (see Definition 1), assume that  $\inf f > 0$ . There exists  $M > 0$ ,  $L > 0$ ,  $R > 0$ ,  $0 \leq \beta < 1$ ,  $D > 0$ ,  $\gamma' > 0$  such that, outside a set of measure less than  $Dh^{\gamma'}$ , for all  $n \in \mathbb{N}^*$ , for all  $t \in \mathbb{R}^+$ , for all  $u \geq Cth$ ,*

$$\mathbb{P}(|\hat{f}_n(x) - f(x)| > t - u^\alpha) \leq 2e^{\frac{1}{e}} \exp[-t^2 M h^{\beta+2} n] \text{ if } f \text{ is } \alpha \text{ regular}$$

if  $r$  is bounded and  $\alpha$ -regular,

$$\mathbb{P}(|\hat{r}_n(x) - r(x)| > t - u^\alpha) \leq e^{\frac{1}{e}} \left( 2 \exp[-t^2 L n h^{\beta+2}] + \exp[-L' n h^{\beta+2}] \right),$$

if  $Y_i \in L^\infty$  and  $r$  is  $\alpha$ -regular,

$$\mathbb{P}(|\hat{r}_n(x) - r(x)| > t - u^\alpha) \leq 2e^{\frac{1}{e}} \exp[-t^2 L n h^{\beta+2}].$$

As a consequence, provided  $h = h_n$  goes to zero and  $nh_n^{\beta+2} = O(n^\varepsilon)$ ,  $\varepsilon > 0$ , we obtain the following convergences provided  $r$  (and  $f$ ) are  $\alpha$ -regular :

- for  $m$  almost all  $x \in \Sigma$ ,  $\hat{f}_n(x)$  converges to  $f(x)$  and  $\hat{r}_n(x)$  converges to  $r(x)$  almost surely and in  $L^p$  for any  $1 \leq p$ ,
- $\mathbb{E} \left( \int_{\Sigma} |\hat{f}_n(x) - f(x)| dx \right)$  and  $\mathbb{E} \left( \int_{\Sigma} |\hat{r}_n(x) - r(x)| \right)$  go to zero.
- for almost all  $x \in \Sigma$ , for  $a < \frac{1}{2}$ ,  $|\hat{g}_n(x) - g(x)| = O_{\mathbb{P}}(n^{-a})$  where  $\hat{g}_n$  is either  $\hat{f}_n$  or  $\hat{r}_n$  and  $g$  is either  $f$  or  $r$ .

Using a “time inverted process” (as in [DeP, Mau2]), we deduce an estimation of the invariant density and of the mapping itself for dynamical systems. Consider a discrete dynamical system from a “good class” (see Definition 1.4 for precise settings), denote by  $T$  the mapping and  $f$  the invariant density of interest. Pick  $X_0$  at random with stationary law  $\mu = fm$  and let  $X_i = T^i(X_0)$ . Consider a kernel  $K$  and  $\hat{f}_n$  as above, let  $\hat{r}_n^j$  be the estimator for  $Y_i^j = X_{i+1}^j + \varepsilon_i^j$ ,  $j = 1, \dots, d$ ,  $X_i^j$  is the  $j$ th coordinate of  $X_i$  and  $(\varepsilon_i^j)_{i \in \mathbb{N}}$  are independent stationary process with 0 mean, independent of  $(X_i)_{i \in \mathbb{N}}$ . Let  $\hat{T}_n$  be an estimator of the map  $T$  which has coordinates  $\hat{r}_n^j$ . Let  $\|\cdot\|_d$  be the sup norm in  $\mathbb{R}^d$ , we have the following results.

**Corollary.** *Assume that the dynamical is in the good class. There exists  $M > 0$ ,  $L > 0$ ,  $R > 0$ ,  $0 \leq \beta \leq 1$ ,  $\gamma' > 0$  such that outside a set of measure*

less than  $Rh^{\gamma'}$ , for all  $t \in \mathbb{R}^+$ , for all  $u \geq Cth$ ,

$$(0.2) \quad \text{is } f \text{ is regular } \mathbb{P}(|\widehat{f}_n(x) - f(x)| > t - u^\alpha) \leq 2e^{\frac{1}{\varepsilon}} \exp[-t^2 Mh^{\beta+2}n]$$

$$(0.3) \quad \mathbb{P}(\|\widehat{T}_n(x) - T(x)\|_d > t - u^\alpha) \leq 2e^{\frac{1}{\varepsilon}} \exp[-t^2 Lh^{\beta+2}n].$$

As a consequence, provided  $h = h_n$  goes to zero and  $nh_n^{\beta+2} = O(n^\varepsilon)$ ,  $\varepsilon > 0$ , we obtain the following convergences :

- for  $m$  almost all  $x \in \Sigma$ ,  $\widehat{T}_n(x)$  converges to  $T(x)$  almost surely and in  $L^p$  for any  $1 \leq p$ , the same holds for  $\widehat{f}_n(x)$  provided that  $f$  is  $\alpha$ -regular.
- $\mathbb{E} \left( \int_{\Sigma} \|\widehat{T}_n(x) - T(x)\|_d \right)$  go to zero, the same holds for  $\widehat{f}_n(x)$  provided that  $f$  is  $\alpha$ -regular.
- for almost all  $x \in \Sigma$ , for  $a < \frac{1}{2}$ ,  $|\widehat{g}_n(x) - g(x)| = O_{\mathbb{P}}(n^{-a})$  where  $\widehat{g}_n$  is either  $\widehat{f}_n$  or  $\widehat{T}_n$  and  $g$  is either  $f$  (provided it is  $\alpha$ -regular) or  $T$ .

In a first section, we state our hypothesis on the process and the dynamical systems. We also give the precise results.

The second section is devoted to the proofs.

In the last section we provide examples of dynamical systems satisfying our hypothesis. We also provide some simulations.

## CONTENTS

1. Hypothesis and statement of the results	3
1.1. Weak dependence	3
1.2. Regularity conditions	4
1.3. Main result	5
1.4. Dynamical systems	6
2. Proofs	7
2.1. Main ingredient : exponential inequality	7
2.2. Proof of Theorem 1.1	8
2.3. Dynamical systems and time reversed process	11
3. Examples and simulations	12
3.1. In dimension one : piecewise expanding maps	12
3.2. In dimension one : unimodal maps	19
3.3. Piecewise expanding maps in higher dimension	20
References	22

## 1. HYPOTHESIS AND STATEMENT OF THE RESULTS

**1.1. Weak dependence.** As mentioned quickly in the introduction, we shall consider a class of weak mixing process with respect to a Banach space of bounded functions  $\mathcal{C}$ . This functional definition of dependence is more general than strong mixing used in [Mas] and [FV]. We shall see that it encounters a large class of dynamical systems (see also [DeP] for other examples than dynamical systems). This kind of functional definition has been introduced in [DoLo] and [DeP] for Lipschitz or BV functions (see also

[Mau2]).

For simplicity, let  $\Sigma \subset \mathbb{R}^d$  and  $(X_i)_{i \in \mathbb{N}}$  a process taking values in  $\Sigma$ . Our proofs could probably be extended to sets  $\Sigma$  included in more general Banach spaces. In the following,  $\|\cdot\|_d$  denotes the sup norm on  $\mathbb{R}^d$ :  $\|(x_1, \dots, x_d)\|_d = \max(|x_i|, i = 1, \dots, d)$ .

**Definition 1.** Let  $\mathcal{C}$  be a Banach space of bounded functions on  $\Sigma$ . We consider a norm on  $\mathcal{C}$  of the form :

$$\|g\|_{\mathcal{C}} = \|g\|_{\infty} + C(g)$$

where  $C(\cdot)$  is a semi-norm on  $\mathcal{C}$  and  $\|\cdot\|_{\infty}$  is the sup. norm on  $\mathcal{C}$ . Let  $\mathcal{C}_1$  be the ‘‘semi-ball’’ of functions  $g \in \mathcal{C}$  such that  $C(g) \leq 1$ . Let  $\mathcal{M}_i$  be the  $\sigma$ -algebra generated by  $X_0, \dots, X_i$ . The  $\mathcal{C}$ -mixing coefficients are :

$$\begin{aligned} \Phi_{\mathcal{C}}(n) &= \sup\{|\mathbb{E}(Zg(X_{i+n})) - \mathbb{E}(Z)\mathbb{E}(g(X_{i+n}))| \mid i \in \mathbb{N}, Z \text{ is} \\ (*) \quad &\mathcal{M}_i - \text{measurable and } \|Z\|_1 \leq 1, g \in \mathcal{C}_1\}. \end{aligned}$$

The process  $(X_i)_{i \in \mathbb{N}}$  is  $\mathcal{C}$ -weakly dependant if  $(\Phi_{\mathcal{C}}(n))_{n \in \mathbb{N}}$  is summable.

Let  $Y_i$  be a  $L^1$  stationary process, taking values in  $\mathbb{R}$ . Let  $\widetilde{\mathcal{M}}_i$  be the  $\sigma$ -algebra generated by  $X_0, Y_0, \dots, X_i, Y_i$ . We shall say that  $(X_i)_{i \in \mathbb{N}}$  is  $\mathcal{C}$ -weakly dependant with respect to  $(Y_i)_{i \in \mathbb{N}}$  if the coefficients :

$$\begin{aligned} \Phi_{\mathcal{C}, (Y_i)_{i \in \mathbb{N}}}(n) &= \sup\{|\mathbb{E}(ZY_{i+n}g(X_{i+n})) - \mathbb{E}(Z)\mathbb{E}(Y_{i+n}g(X_{i+n}))| \mid i \in \mathbb{N}, Z \text{ is} \\ (**) \quad &\widetilde{\mathcal{M}}_i - \text{measurable and } \|Z\|_1 \leq 1, g \in \mathcal{C}_1\}, \end{aligned}$$

are summable.

*Remark.* Examples of spaces  $\mathcal{C}$  that we shall consider are :

- the space of function with bounded variations, in that case,  $C(g)$  is the total variation of  $g$  ;
- the space of Lipschitz (resp. Hölder) functions, in that case,  $C(g)$  is the Lipschitz (resp. Hölder) constant ;
- the space of  $C^1$  function could also be considered, in that case,  $C(g)$  is the sup norm of  $g'$ .

**1.2. Regularity conditions.** We now state the conditions on the kernel  $K$ , the regression function  $r$  and the density function  $f$ .

**Assumption 1.** The kernel  $K$  is a nonnegative function in  $\mathcal{C}$  (so it is bounded), with compact support  $D$  and with integral 1 :

$$\int_{\Sigma} K dm = 1.$$

For  $h > 0$  and  $x \in \Sigma$ , let  $K_{h,x}(t) = K\left(\frac{x-t}{h}\right)$ . We assume that there exists  $0 \leq \beta < 1$  such that  $C(K_{h,x}) \leq \frac{C(K)}{h^{\beta}}$ .

**Assumption 2.** Consider a function  $g$  on  $\Sigma$  and  $0 < \alpha \leq 1$ . Let  $B_g(u, h)$  be the set of points  $x$  such that

$$\sup_{d(x,y) < h} |g(x) - g(y)| > u^{\alpha}.$$

If  $g$  is a map from  $\Sigma$  to  $\Sigma$ ,  $B_g(u, h)$  is the set of points  $x$  such that

$$\sup_{d(x,y) < h} \|g(x) - g(y)\|_d > u^{\alpha}.$$

For any decreasing to zero sequences  $(u_n)_{n \in \mathbb{N}}$  and  $(h_n)_{n \in \mathbb{N}}$ , let  $A_n = B_g(u_n, h_n)$  and  $B_N = \bigcap_{n \geq N} \bigcup_{p \geq n} A_p$ .

We shall say that  $g$  is  $\alpha$ -regular if for any sequence  $(u_n)_{n \in \mathbb{N}}$  decreasing to 0, for any sequence  $(h_n)_{n \in \mathbb{N}}$  with  $h_n = o(u_n)$ , there exist constants  $D(g), D'(g) > 0$  and  $\gamma, \gamma' > 0$  such that  $m(B_N) \leq D(g)h_N^\gamma$  and  $m(A_N) \leq D'(g)h_N^{\gamma'}$ .

All our results are proved under the assumption that  $r$  and/or  $f$  are  $\alpha$ -regular for some  $0 < \alpha \leq 1$ .

*Remarks.* If  $g$  is  $\alpha$ -Hölder then it is  $\alpha$ -regular with  $m(B_g(u, h)) = 0$  for all  $u \geq H(g)h$  with  $H(g)$  the Hölder constant of  $g$ .

If  $g$  is Lipschitz on  $\Sigma \setminus \{x\}$  and discontinuous at  $x$  then  $m(B_g(u, h)) \leq Cth^d$  for all  $u \geq L(g)h$  where  $L(g)$  is the Lipschitz constant of  $g$  on  $\Sigma \setminus \{x\}$ .

Functions of bounded variations satisfy this condition (see section 3).

An  $\alpha$ -regular function need not to be Hölder everywhere but we control the points where this is not the case.

**1.3. Main result.** Let us state our main result for  $\mathcal{C}$ -weak dependant process.

**Theorem 1.1.** *Assume that  $(X_i)_{i \in \mathbb{N}}$  is a stationary process absolutely continuous and is  $\mathcal{C}$ -weakly dependant with respect to  $(Y_i)_{i \in \mathbb{N}}$ ,  $(Y_i)_{i \in \mathbb{N}}$  is a stationary process taking values in  $\mathbb{R}$ . Let  $r$  be the regression function ( $r = \mathbb{E}(Y_i | X_i)$ ) and  $f$ , the density function, we assume that  $\inf f > 0$ . Then there exists  $M > 0$ ,  $L > 0$ ,  $L' > 0$ ,  $R > 0$ , a set  $A_{u,h}$  such that for all  $n \in \mathbb{N}^*$ , for all  $t \in \mathbb{R}^+$ , for  $x \notin A_{u,h}$ , for all  $u \geq h \cdot \text{Diam}(D)$ ,  $m(A_{u,h}) \leq Rh^{\gamma'}$ , if  $f$  is  $\alpha$ -regular,*

$$(1.1) \quad \mathbb{P}(|\hat{f}_n(x) - f(x)| > t - u^\alpha) \leq 2e^{\frac{1}{e}} \exp[-t^2 Mnh^{\beta+2}],$$

if  $r$  is  $\alpha$ -regular and bounded,

$$(1.2) \quad \mathbb{P}(|\hat{r}_n(x) - r(x)| > t - u^\alpha) \leq e^{\frac{1}{e}} \left( 2 \exp[-t^2 Lnh^{\beta+2}] + \exp[-L'nh^{\beta+2}] \right),$$

if  $r$  is  $\alpha$ -regular and  $Y_i \in L^\infty$ ,

$$(1.3) \quad \mathbb{P}(|\hat{r}_n(x) - r(x)| > t - u^\alpha) \leq 2e^{\frac{1}{e}} \exp[-t^2 Lnh^{\beta+2}].$$

As a consequence, provided  $h = h_n$  goes to zero and  $nh_n^{\beta+2} = O(n^\varepsilon)$ ,  $\varepsilon > 0$ , we obtain the following convergences :

- for  $m$  almost all  $x \in \Sigma$ ,  $\hat{r}_n(x)$  converges to  $r(x)$  almost surely and in  $L^p$  for any  $1 \leq p$ , the same holds for  $\hat{f}_n(x)$  provided  $f$  is  $\alpha$ -regular.
- $\mathbb{E} \left( \int_\Sigma |\hat{r}_n(x) - r(x)| \right)$  goes to zero, the same holds for  $\hat{f}_n(x)$  provided  $f$  is  $\alpha$ -regular.
- for almost all  $x \in \Sigma$ , for  $a < \frac{1}{2}$ ,  $|\hat{r}_n(x) - r(x)| = O_{\mathbb{P}}(n^{-a})$ , the same holds for  $\hat{f}_n(x)$  provided  $f$  is  $\alpha$ -regular.

**1.4. Dynamical systems.** We turn now to our main motivation : dynamical systems. Consider a dynamical system  $(\Sigma, T, \mu)$ . That is,  $T$  maps  $\Sigma$  into itself,  $\mu$  is a  $T$ -invariant probability measure on  $\Sigma$ , absolutely continuous with respect to the Lebesgue measure  $m$ . We assume that the dynamical system satisfy the following mixing property : for all  $\varphi \in L^1(\mu)$ ,  $\psi \in \mathcal{C}$ ,

$$(1.4) \quad \left| \int_{\Sigma} \psi \cdot \varphi \circ T^n d\mu - \int_{\Sigma} \psi d\mu \int_{\Sigma} \varphi d\mu \right| \leq \Phi(n) \|\varphi\|_1 \|\psi\|_{\mathcal{C}},$$

with  $\Phi(n)$  summable.

**Definition 2.** A dynamical system  $(\Sigma, T, \mu)$  is of the good class if

- (1)  $\Sigma$  is a bounded subset of  $\mathbb{R}^d$ ,
- (2) the map  $T : \Sigma \rightarrow \Sigma$  is  $\alpha$ -regular,
- (3) the invariant density  $f$  ( $\mu = fm$ ) verify  $\inf f > 0$ ,
- (4) it satisfy the mixing property (1.4).

We shall also assume that the Banach space  $\mathcal{C}$  is *regular* in the sense that for any  $\varphi \in \mathcal{C}$ , there exists  $R(\varphi) \in \mathbb{R}$  such that

$$(1.5) \quad \|\varphi + R(\varphi)\|_{\infty} \leq C(\varphi).$$

We consider the stationary process taking values in  $\Sigma : X_i = T^i$  with law  $\mu$  and  $Y_i = X_{i+1} + \varepsilon_i$  where the  $\varepsilon_i$  are  $L^1$  independent random vectors, with independent coordinates, identically distributed random variables with 0 mean and independent on  $(X_i)_{i \in \mathbb{N}}$ . Then  $T$  is the regression function  $T(x) = \mathbb{E}(Y_i | X_i = x)$ . Let  $\widehat{r}_n^j$  is the estimator for  $Y_i^j = X_{i+1}^j + \varepsilon_i^j$ , defined by (0.1),  $j = 1, \dots, d$ ,  $X_i^j$  is the  $j$ th coordinate of  $X_i$  and  $\varepsilon_i^j$  is the  $j$ th coordiante of  $\varepsilon_i$ . We shall denote  $\widehat{T}_n$  the estimator of  $T$  which has coordinates  $\widehat{r}_n^j$ . Theorem 1.1 applied in the context of dynamical systems gives the following result.

**Corollary 1.2.** *Let  $T$  be a dynamical system of the good class, assume that the Banach space  $\mathcal{C}$  is regular. There exists  $M > 0$ ,  $L > 0$ ,  $R > 0$ ,  $0 \leq \beta \leq 1$ ,  $\gamma' > 0$  such that outside a set of measure less than  $Rh^{\gamma'}$ , for all  $t \in \mathbb{R}^+$ , for all  $u \geq Cth$ ,*

$$\mathbb{P}(\|\widehat{T}_n(x) - T(x)\|_d > t - u^\alpha) \leq 2de^{\frac{1}{\varepsilon}} \exp[-t^2 Lh^{\beta+2}n]$$

$$\mathbb{P}(|\widehat{f}_n(x) - f(x)| > t - u^\alpha) \leq 2e^{\frac{1}{\varepsilon}} \exp[-t^2 Mh^{\beta+2}n] \text{ if } f \text{ is } \alpha\text{-regular.}$$

As a consequence, provided  $h = h_n$  goes to zero and  $nh_n^{\beta+2} = O(n^\varepsilon)$ ,  $\varepsilon > 0$ , we obtain the following convergences :

- for  $m$  almost all  $\widehat{T}_n(x)$  converges to  $T(x)$   $x \in \Sigma$ ,  $\widehat{f}_n(x)$  converges to  $f(x)$  almost surely and in  $L^p$  for any  $1 \leq p$ , if  $f$  is  $\alpha$ -regular,  $\widehat{f}_n(x)$  converges to  $f(x)$  almost surely and in  $L^p$  for any  $1 \leq p$ ,
- $\mathbb{E} \left( \int_{\Sigma} \|\widehat{T}_n(x) - T(x)\|_d \right)$  go to zero, if  $f$  is  $\alpha$ -regular,
- $\mathbb{E} \left( \int_{\Sigma} |\widehat{f}_n(x) - f(x)| dx \right)$  go to zero.
- for almost all  $x \in \Sigma$ , for  $a < \frac{1}{2}$ ,  $|\widehat{g}_n(x) - g(x)| = O_{\mathbb{P}}(n^{-a})$  where  $\widehat{g}_n$  is either  $\widehat{f}_n$  or  $\widehat{T}_n$  and  $g$  is either  $f$  (provided  $f$  is  $\alpha$ -regular) or  $T$ .

## 2. PROOFS

Let us now prove the results stated in the above section. The main ingredient is an exponential inequality that has been obtained in [DeP] in the case  $\mathcal{C} = BV$ , and in [Mau2] for more general spaces  $\mathcal{C}$ .

## 2.1. Main ingredient : exponential inequality.

**Proposition 2.1.** *Let  $(X_i)_{i \in \mathbb{N}}$  be a  $\mathcal{C}$ -weakly mixing process. Let the coefficients  $\Phi_{\mathcal{C}}(k)$  be defined by (\*). For  $\varphi \in \mathcal{C}$ ,  $p \geq 2$ , define*

$$S_n(\varphi) = \sum_{i=1}^n \varphi(X_i)$$

and

$$b_{i,n} = \Phi_{\mathcal{C}}(0) \left( \sum_{k=0}^{n-i} \Phi_{\mathcal{C}}(k) \right) \|\varphi(X_i) - \mathbb{E}(\varphi(X_i))\|_{\frac{p}{2}} C(\varphi).$$

For any  $p \geq 2$ , we have the inequality :

$$\begin{aligned} \|S_n(\varphi) - \mathbb{E}(S_n(\varphi))\|_p &\leq \left( 2p \Phi_{\mathcal{C}}(0) \sum_{i=1}^n b_{i,n} \right)^{\frac{1}{2}} \\ (2.1) \quad &\leq C(\varphi) \left( 2p \sum_{k=0}^{n-1} (n-k) \Phi_{\mathcal{C}}(k) \right)^{\frac{1}{2}}. \end{aligned}$$

As a consequence, we obtain

$$\begin{aligned} (2.2) \quad &\mathbb{P}(|S_n(\varphi) - \mathbb{E}(S_n(\varphi))| > t) \\ &\leq e^{\frac{1}{e}} \exp \left( \frac{-t^2}{2e(C(\varphi))^2 \Phi_{\mathcal{C}}(0) \sum_{k=0}^{n-1} (n-k) \Phi_{\mathcal{C}}(k)} \right). \end{aligned}$$

Applying Proposition 2.1 to the function  $\phi = K$  will provide an exponential inequality for  $\widehat{f}_n$ . The proof of Proposition 2.1 may be found in [Mau2] (Proposition 1.1), see also [DeP] for a version of this proposition with  $\mathcal{C} = BV$ . In order to have an exponential inequality for  $\widehat{g}_n$ , we need the following result.

**Proposition 2.2.** *Let  $(X_i)_{i \in \mathbb{N}}$  be a  $\mathcal{C}$ -weakly mixing process with respect to  $(Y_i)_{i \in \mathbb{N}}$ . Let the coefficients  $\Phi_{\mathcal{C},(Y_i)_{i \in \mathbb{N}}}(k)$  be defined by (\*\*). For  $\varphi \in \mathcal{C}$ ,  $p \geq 2$ , define*

$$\widetilde{S}_n(\varphi) = \sum_{i=1}^n Y_i \varphi(X_i)$$

and

$$\widetilde{b}_{i,n} = \Phi_{\mathcal{C}(0),(Y_i)_{i \in \mathbb{N}}} \left( \sum_{k=0}^{n-i} \Phi_{\mathcal{C},(Y_i)_{i \in \mathbb{N}}}(k) \right) \|Y_i \varphi(X_i) - \mathbb{E}(Y_i \varphi(X_i))\|_{\frac{p}{2}} C(\varphi).$$



For any  $p \geq 2$ , we have the inequality :

$$(2.3) \quad \begin{aligned} \|\tilde{S}_n(\varphi) - \mathbb{E}(\tilde{S}_n(\varphi))\|_p &\leq \left( 2p \sum_{i=1}^n \tilde{b}_{i,n} \right)^{\frac{1}{2}} \\ &\leq C(\varphi) \left( 2p \Phi_{\mathcal{C},(Y_i)_{i \in \mathbb{N}}}(0) \sum_{k=0}^{n-1} (n-k) \Phi_{\mathcal{C},(Y_i)_{i \in \mathbb{N}}}(k) \right)^{\frac{1}{2}}. \end{aligned}$$

As a consequence, we obtain

$$(2.4) \quad \begin{aligned} &\mathbb{P}(|S_n(\varphi) - \mathbb{E}(S_n(\varphi))| > t) \\ &\leq e^{\frac{1}{e}} \exp \left( \frac{-t^2}{2e(C(\varphi))^2 \Phi_{\mathcal{C}(0),(Y_i)_{i \in \mathbb{N}}}(0) \sum_{k=0}^{n-1} (n-k) \Phi_{\mathcal{C},(Y_i)_{i \in \mathbb{N}}}(k)} \right). \end{aligned}$$

*Proof.* Proposition 2.2 is proved as Proposition 1.1 in [Mau2] using the variable  $Z_i = Y_i \varphi(X_i) - \mathbb{E}(Y_i \varphi(X_i))$  for which we control  $\|\mathbb{E}(Z_k | \tilde{\mathcal{M}}_i)\|_\infty$  with the coefficients  $\Phi_{\mathcal{C},(Y_i)_{i \in \mathbb{N}}}(k-i)$ , for  $k \geq i$ .  $\square$

**2.2. Proof of Theorem 1.1.** As already mentioned, the proof of Theorem 1.1 uses Propositions 2.1 and 2.2 applied to  $\varphi = K$ . This gives a control on the deviation between  $\hat{f}_n(x)$  (resp.  $\hat{g}_n(x)$ ) and  $\mathbb{E}(\hat{f}_n(x))$  (resp.  $\mathbb{E}(\hat{g}_n(x))$ ). Then it remains to see that  $\mathbb{E}(\hat{f}_n(x))$  is close to  $f(x)$ . In order to obtain the result for the regression function, we deduce from the previous results a control on the deviation between  $\hat{r}_n(x)$  and  $\frac{\mathbb{E}(\hat{g}_n(x))}{\mathbb{E}(\hat{f}_n(x))}$ , then we prove that

$\frac{\mathbb{E}(\hat{g}_n(x))}{\mathbb{E}(\hat{f}_n(x))}$  is close to  $r(x)$ . To this aim, we begin with two lemmas.

**Lemma 2.3.** *We assume that the density function  $f$  is  $\alpha$ -regular. For all  $u \geq h \text{diam}(D) =: k$ , if  $x \notin B_f(u, k)$  we have :*

$$|\mathbb{E}(\hat{f}_n(x)) - f(x)| \leq u^\alpha.$$

*Proof.* We have :

$$\mathbb{E}(\hat{f}_n(x)) = \int_D K(y) f(x - hy) dm(y).$$

If  $x \notin B_f(u, k)$  then for  $y \in D$ ,

$$|f(x) - f(x - hy)| \leq u^\alpha.$$

The result follows from the fact that the integral of  $K$  is 1.  $\square$

**Lemma 2.4.** *We assume that the regression function  $r$  is  $\alpha$ -regular. For all  $u \geq h \text{diam}(D) =: k$ , if  $x \notin B_r(u, k)$  we have :*

$$\left| \frac{\mathbb{E}(\hat{g}_n(x))}{\mathbb{E}(\hat{f}_n(x))} - r(x) \right| \leq u^\alpha.$$

*Proof.* We have :

$$\frac{\mathbb{E}(\hat{g}_n(x))}{\mathbb{E}(\hat{f}_n(x))} = \frac{\int_D r(x - yh) K(x) f(x - yh) dm(y)}{\int_D K(x) f(x - yh) dm(y)}.$$

If  $x \notin B_r(u, k)$  then for  $y \in D$ ,

$$|r(x) - r(x - hy)| \leq u^\alpha$$

and the result follows.  $\square$

*Proof of Theorem 1.1.* Proposition 2.1 applied to  $\varphi(t) = K \left( \frac{x-t}{h} \right)$  gives : for all  $t > 0$ , for all  $n \in \mathbb{N}$ , for all  $x \in \Sigma$ ,

$$\mathbb{P}(|\widehat{f}_n(x) - \mathbb{E}(\widehat{f}_n(x))| > t) \leq e^{\frac{1}{e}} \exp \left( -\frac{t^2 n h^{\beta+2}}{2eR\Phi_C(0)C(K)} \right),$$

where  $R$  is the smallest positive number such that

$$\sum_{k=0}^{n-1} (n-k)\Phi_C(k) \leq Rn.$$

This together with Lemma 2.3 gives (1.1) for  $x \notin B_f(u, k)$  ( $k = h \text{diam}(D)$ ). In order to obtain the estimation for the regression function  $r$ , we apply Proposition 2.2 to  $\varphi(t) = K \left( \frac{x-t}{h} \right)$ . We obtain :

$$\mathbb{P}(|\widehat{g}_n(x) - \mathbb{E}(\widehat{g}_n(x))| > t) \leq e^{\frac{1}{e}} \exp \left( -\frac{t^2 n h^{\beta+2}}{2eR'\Phi_{C,(Y_i)_{i \in \mathbb{N}}}(0)C(K)} \right),$$

where  $R'$  is the smallest positive number such that

$$\sum_{k=0}^{n-1} (n-k)\Phi_{C,(Y_i)_{i \in \mathbb{N}}}(k) \leq R'n.$$

Now, for any  $t > 0$ ,

$$\begin{aligned} \mathbb{P} \left( \left| \widehat{r}_n(x) - \frac{\mathbb{E}(\widehat{g}_n(x))}{\mathbb{E}(\widehat{f}_n(x))} \right| > t \right) &\leq \mathbb{P} \left( |\widehat{g}_n(x) - \mathbb{E}(\widehat{g}_n(x))| > \frac{t}{2} \mathbb{E}(\widehat{f}_n(x)) \right) \\ &\quad + \mathbb{P} \left( |\widehat{f}_n(x) - \mathbb{E}(\widehat{f}_n(x))| > \frac{t}{2} \mathbb{E}(\widehat{f}_n(x)) \widehat{r}_n(x)^{-1} \right). \end{aligned}$$

Now, assume that  $Y_i \in L^\infty$ , we remark that  $|\widehat{r}_n(x)| \leq \|Y_i\|_\infty =: y_{\max}$  thus :

$$\begin{aligned} \mathbb{P} \left( \left| \widehat{r}_n(x) - \frac{\mathbb{E}(\widehat{g}_n(x))}{\mathbb{E}(\widehat{f}_n(x))} \right| > t \right) &\leq \mathbb{P} \left( |\widehat{g}_n(x) - \mathbb{E}(\widehat{g}_n(x))| > \frac{t}{2} \inf f \right) \\ &\quad + \mathbb{P} \left( |\widehat{f}_n(x) - \mathbb{E}(\widehat{f}_n(x))| > \frac{t \inf f}{2 y_{\max}} \right). \end{aligned}$$

Propositions 2.1 and 2.2 give :

$$\mathbb{P} \left( \left| \widehat{r}_n(x) - \frac{\mathbb{E}(\widehat{g}_n(x))}{\mathbb{E}(\widehat{f}_n(x))} \right| > t \right) \leq 2e^{\frac{1}{e}} \exp[-Ct^2 h^{\beta+2} n]$$

where

$$C = \max \left( \frac{(\inf f)^2}{8eR'\Phi_{C,(Y_i)_{i \in \mathbb{N}}}(0)C(K)}, \frac{(\inf f)^2}{8eR\Phi_C(0)C(K)y_{\max}^2} \right).$$

Lemma 2.4 gives (1.3) for  $x \notin B_r(u, k)$ . Let  $A_h = B_f(u, k) \cup B_r(u, k)$ , the first part of Theorem 1.1 is now proved for  $x \notin A_h$  if  $Y_i \in L^\infty$ .

If we don't assume  $Y_i \in L^\infty$ , but  $r$  is bounded by  $r_{\max}$ , we write :

$$\begin{aligned} \mathbb{P} \left( \left| \widehat{r}_n(x) - \frac{\mathbb{E}(\widehat{g}_n(x))}{\mathbb{E}(\widehat{f}_n(x))} \right| > t \right) &\leq \mathbb{P} \left( |\widehat{g}_n(x) - \mathbb{E}(\widehat{g}_n(x))| > \frac{t}{2} \widehat{f}_n(x) \right) \\ &\quad + \mathbb{P} \left( \left| \widehat{f}_n(x) - \mathbb{E}(\widehat{f}_n(x)) \right| > \frac{t}{2} \widehat{f}_n(x) \left[ \frac{\mathbb{E}(\widehat{g}_n(x))}{\mathbb{E}(\widehat{f}_n(x))} \right]^{-1} \right) \\ &\leq \mathbb{P} \left( |\widehat{g}_n(x) - \mathbb{E}(\widehat{g}_n(x))| > \frac{t}{4} \inf f \right) + \mathbb{P} \left( \left| \widehat{f}_n(x) - \mathbb{E}(\widehat{f}_n(x)) \right| > \frac{t \inf f}{4 r_{\max}} \right) \\ &\quad + \mathbb{P} \left( \widehat{f}_n(x) < \frac{\inf f}{2} \right). \end{aligned}$$

Propositions 2.1 and 2.2 give :

$$\mathbb{P} \left( \left| \widehat{r}_n(x) - \frac{\mathbb{E}(\widehat{g}_n(x))}{\mathbb{E}(\widehat{f}_n(x))} \right| > t \right) \leq e^{\frac{1}{e}} \left[ 2 \exp[-C' t^2 h^{\beta+2} n] + \exp[-C'' n h^{\beta+2}] \right],$$

where  $C'$  and  $C''$  may be expressed in terms of  $\inf f$ ,  $r_{\max}$ ,  $R$ ,  $R'$ ,  $\Phi_{\mathcal{C}}(0)$ ,  $\Phi_{\mathcal{C}, (Y_i)_{i \in \mathbb{N}}}(0)$ ,  $C(K)$ . Then (1.2) follows as above for  $x \notin A_h$ .

Let us conclude with the almost everywhere and  $L^p$  convergence, for all  $1 \leq p$ . Let us begin with the convergence in  $L^p(m)$ . We fix  $1 \leq p$ . We remark that since the kernel  $K$  is bounded, so is  $\widehat{f}_n$  :  $\sup(\widehat{f}_n) \leq \frac{\sup K}{h}$ . Also, if  $x \notin B_f(u, \varepsilon)$  then  $f(x) \leq \frac{1+u^\alpha}{\varepsilon^d}$ . Let  $E_{t,u}(x)$  be the event

$$E_{t,u}(x) = \{ |\widehat{f}_n(x) - f(x)| > t - u^\alpha \},$$

we have :

$$\begin{aligned} \int |\widehat{f}_n(x) - f(x)|^p d\mathbb{P} &= \int_{E_{t,u}(x)} |\widehat{f}_n(x) - f(x)|^p d\mathbb{P} \\ &\quad + \int_{[E_{t,u}(x)]^c} |\widehat{f}_n(x) - f(x)|^p d\mathbb{P} \\ &\leq \mathbb{P}(E_{t,u}(x)) \left( \frac{1 + u^\alpha}{\text{diam}(D)h} + \frac{\sup K}{h} \right)^p + (t - u^\alpha)^p \\ &\leq \frac{Ct}{h^p} \exp[-Ct^2 h^\beta n] + (t - u^\alpha)^p. \end{aligned}$$

Take  $u$  such that  $u^\alpha = \frac{1}{\ln n}$ ,  $t = 2u^\alpha$  and  $h = h_n = O(\frac{1}{n^\xi})$  with  $0 < \xi < \frac{1}{\beta+2}$ . Then for  $x \notin A_n = B_f(u_n, h_n)$ , there exists  $0 < \kappa < 1$  and constants  $C_1, C_2 > 0$  such that :

$$\|\widehat{f}_n(x) - f(x)\|_p^p \leq C_1 e^{-C_2 n^\kappa} + \left( \frac{1}{\ln n} \right)^p.$$

Thus,  $\|\widehat{f}_n(x) - f(x)\|_p$  goes to zero for  $x \notin B_N = \bigcap_{n \geq N} \bigcup_{p \geq n} A_p$  and  $m(B_N) = O(h_N^\gamma)$ . We conclude that for almost all  $x \in \Sigma$ ,  $\widehat{f}_n(x)$  goes to  $f(x)$  in  $L^p(m)$ .

The proof of the almost everywhere convergence is in the same spirit. Consider a sequence  $t_m$  decreasing to 0, for  $x \in \Sigma$ ,

$$\mathbb{P}(\{\widehat{f}_n(x) \not\rightarrow f(x)\}) \leq \lim_{m \rightarrow \infty} \lim_{N \rightarrow \infty} \sum_{n \geq N} \mathbb{P}(\{|\widehat{f}_n(x) - f(x)| > t_m\}).$$

Take  $h = h_n = O(\frac{1}{n^\xi})$  with  $0 < \xi < \frac{1}{\beta+2}$  then

$$\lim_{N \rightarrow \infty} \sum_{n \geq N} \mathbb{P}(\{|\widehat{f}_n(x) - f(x)| > t_m\}) = 0$$

provided  $t_m > u^\alpha$  with  $u^\alpha > Cth_n$  and  $x \notin B_{u, h_n}$ , we take  $u = u_n = \frac{1}{\ln n}$  and  $A_n = B_{u_n, h_n}$ , for  $x \notin \cap_N \cup_{n \geq N} A_n =: B$ ,  $\widehat{f}_n(x)$  goes to  $f(x)$  and  $m(B) = 0$ . The proofs of  $L^p(m)$  and almost everywhere convergence for  $\widehat{r}_n$  follow the same lines.

Finally, to prove the bound for the convergence in probability, recall that  $u_n = O_{\mathbb{P}}(n^{-a})$  if and only if

$$\lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(n^a |u_n| > M) = 0.$$

Following the same lines as above, we prove that for  $g$  either  $f$  or  $r$  and  $\widehat{g}_n$  either  $\widehat{f}_n$  or  $\widehat{r}_n$ ,

$$\lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(n^a |\widehat{g}_n(x) - g(x)| > M) = 0$$

for  $a < \frac{1}{2}$ ,  $x \notin B = \cap_{n \geq 0} \cup_{p \geq n} A_p$  and  $m(B) = 0$ .  $\square$

Let us show how we can apply the above result to dynamical systems.

**2.3. Dynamical systems and time reversed process.** It turns out that, in general, the process  $(X_i)_{i \in \mathbb{N}}$  associated to a dynamical system  $(\Sigma, T, \mu)$  is not weakly dependent. Nevertheless the condition of weak dependence is satisfied for a “time reversed process” whose law is the same as  $(X_i)_{i \in \mathbb{N}}$ . Indeed, Condition 1.4 together with the fact that the space  $\mathcal{C}$  is regular (recall (1.5)) gives (see [DeP] or [Mau2] for the details) : for all  $i \in \mathbb{N}$ , for  $\psi \in \mathcal{C}_1$ ,  $\varphi \in L^1$ ,  $\|\varphi\|_1 \leq 1$ ,

$$(2.5) \quad |\text{Cov}(\psi(X_i), \varphi(X_{n+i}))| \leq 2\Phi(n).$$

So that, if we consider the process  $(\widetilde{X}_i)_{i \in \mathbb{N}}$  defined by

$$(\widetilde{X}_0, \dots, \widetilde{X}_n) \stackrel{\text{Law}}{=} (X_n, \dots, X_0) \quad \forall n \in \mathbb{N},$$

it is  $\mathcal{C}$ -weakly dependent.

Recall that  $Y_i = X_{i+1} + \varepsilon_i$  with  $(\varepsilon_i)_{i \in \mathbb{N}}$  independent random vectors, with independent coordinates, of zero mean, which is also independent of  $(X_i)_{i \in \mathbb{N}}$ . In order to estimate the regression function  $\mathbb{E}(Y_i | X_i) = T$ , we shall estimate separately each coordinates of  $X_{i+1}$ . Let

$$\begin{aligned} \widehat{g}_n^j(x) &= \frac{1}{nh} \sum_{i=0}^{n-1} Y_i^j K\left(\frac{x - X_i}{h}\right) \\ &= \frac{1}{nh} \sum_{i=0}^{n-1} X_{i+1}^j K\left(\frac{x - X_i}{h}\right) + \frac{1}{nh} \sum_{i=0}^{n-1} \varepsilon_i^j K\left(\frac{x - X_i}{h}\right). \end{aligned}$$

Equation (2.5) implies also that  $(\tilde{X}_i)_{i \in \mathbb{N}}$  is  $\mathcal{C}$ -weakly dependant with respect to each  $(\tilde{X}_{i+1}^j)_{i \in \mathbb{N}}$ , and also to each  $(\varepsilon_i^j)$ ,  $j = 1, \dots, d$ . We obtain inequalities (1.3) separately for the two parts of  $\hat{g}_n^j(x)$  and then the two parts of  $\hat{r}_n^j(x)$ . Then we deduce Corollary 1.2.

### 3. EXAMPLES AND SIMULATIONS

We shall now give some examples of discrete dynamical systems satisfying (2.5), such that  $T$  is  $\alpha$ -regular, admits a unique invariant density (with respect to the Lebesgue measure),  $f$  which may be also  $\alpha$ -regular.

There is a large class of dynamical systems satisfying (2.5), we refer to the literature on dynamical systems : [Y1, Y2, Mau1, Li, Col, BuMau1, BuMau2, LiSV, Bro] and many other, see [Ba] for a review on these topics. Below we consider Lasota-Yorke maps, unimodal maps, piecewise expanding maps in higher dimension. Our results should also apply for hyperbolic maps but a more intricate study on the invariant density is required.

**3.1. In dimension one : piecewise expanding maps.** We shall consider piecewise expanding maps or ‘‘Lasota-Yorke’’ maps. Consider  $I = [0, 1]$ , partitioned into subintervals  $I_j$ . On the interior of the  $I_j$ 's, the map  $T$  is  $C^2$ , uniformly expanding  $|T'| \geq \lambda > 1$ , and continuous on the closure of the  $I_j$ 's. Under additional conditions of mixing or covering (see below), it is well known ([Bro, LiSV, Li, Col]) that  $T$  admits a unique absolutely continuous invariant measure whose density  $f$  belongs to the space  $BV$  of functions of bounded variations.

Before we give precisely the hypothesis on the map, we prove that functions of bounded variations are regular. We shall denote  $m$  the Lebesgue measure on  $I$ .

Recall that a function  $g$  on  $I$  belongs to  $BV$  if

$$\bigvee g = \sup \sum_{i=0}^n |g(a_i) - g(a_{i+1})| < \infty$$

where the supremum is taken over all the finite subdivisions of  $I$  ;  $a_0 = 0 < a_1 < \dots < a_{n+1} = 1$ , then  $\bigvee g$  is the total variation of  $g$ .

**Lemma 3.1.** *Let  $g$  belongs to  $BV$ . Then for any sequences  $(u_n)_{n \in \mathbb{N}}$  and  $(h_n)_{n \in \mathbb{N}}$  decreasing to zero with  $h_n^{2-\alpha} \leq u_n^2$ ,  $0 < \alpha < 2$ , let  $B_g(u, h)$  be the set of points  $x$  such that*

$$\sup_{d(x,y) < h} |g(x) - g(y)| > u,$$

$$A_n = B_g(u_n, h_n) \text{ and } B_N = \bigcap_{n \geq N} \bigcup_{p \geq n} A_p.$$

Then

$$m(B_N) \leq 3(\bigvee g) h_N^{\frac{\alpha}{2}} \text{ and } m(A_n) \leq 2(\bigvee g) \frac{h_n}{u_n}$$

*Proof.* Let  $x_0 = \inf B_N$  and  $x_1$  such that  $0 \leq x_1 - x_0 \leq h_N$  and  $x_1 \in B_N$ . Then  $x_1$  belongs to  $A_p$  for infinitely many  $p \geq N$ . In particular, there exists  $p_1 \geq N$  and  $x_2$  with  $|x_1 - x_2| \leq h_{p_1}$  and  $|g(x_1) - g(x_2)| \geq u_{p_1}$ . Set

$x_0 = x_{1,0}$ ,  $x_1 = x_{1,1}$  and  $x_2 = x_{1,2}$ . We construct sequences (maybe finite)  $(x_{i,0}, x_{i,1}, x_{i,2})$  and  $(p_i)$  such that :

- (1)  $p_i \geq N$  for all  $i \geq 1$ ,
- (2)  $x_{i,0} \leq \min(x_{i,1}, x_{i,2})$  ;  $x_{i+1,0} \geq \max(x_{i,1}, x_{i,2})$ ,  $x_{i,1} \in A_{p_i}$ ,
- (3)  $|x_{i,0} - x_{i,1}| \leq h_{p_{i-1}}$  for all  $i \geq 1$ , with the convention that  $h_{p_0} = h_N$ ,
- (4)  $|x_{i,1} - x_{i,2}| \leq h_{p_i}$  for all  $i \geq 1$ ,
- (5)  $|g(x_{i,1}) - g(x_{i,2})| \geq u_{p_i}$  for all  $i \geq 1$ ,
- (6)  $B_N \subset \bigcup_{i \geq 1} J_i$  with  $J_i = [a_i, b_i + h_{p_{i-1}}]$ ,  $[a_i, b_i] = [x_{i,0}, x_{i,1}] \cup [x_{i,1}, x_{i,2}]$ .

If  $x_{i,0}, x_{i,1}, x_{i,2}$  are already constructed, let  $x_{i+1,0} = \inf[x \in B_N, x \geq \max(x_{i,1}, x_{i,2}) + h_{p_i}]$ . Now, there exists  $x_{i+1,1} \in B_N$  with  $0 \leq x_{i+1,1} - x_{i+1,0} \leq h_{p_i}$ . Since  $x_{i+1,1} \in B_N$ , there exists  $p_{i+1} \geq p_i$  and  $x_{i+1,2}$  with  $|x_{i+1,1} - x_{i+1,2}| \leq h_{p_{i+1}}$  and  $|g(x_{i+1,1}) - g(x_{i+1,2})| \geq u_{p_{i+1}}$ . In that way, all

the points of  $B_N$  smaller than  $\max(x_{i+1,2}, x_{i+1,1}) + h_{p_{i+1}}$  are in  $\bigcup_{j=1}^{i+1} J_j$ . The

construction stops when  $\{x \in B_N, x \geq \max(x_{i,1}, x_{i,2}) + h_{p_i}\}$  is empty.

As a consequence of this construction, we get :

$$\sum_{i \geq 1} u_{p_i} \leq \bigvee g \text{ remark that the } J_i \text{'s are disjoint,}$$

$$m(B_N) \leq 3 \sum_{i \geq 0} h_{p_i}.$$

Now, using Cauchy-Schwartz inequality,

$$\begin{aligned} \sum_{i \geq 0} h_{p_i} &\leq \left( \sum_{i \geq 1} \frac{h_{p_i}^2}{u_{p_i}} \right)^{\frac{1}{2}} \cdot \left( \sum_{i \geq 1} u_{p_i} \right)^{\frac{1}{2}} \\ &\leq \left( \bigvee g \right)^{\frac{1}{2}} \cdot h_N^{\frac{\alpha}{2}} \left( \sum_{i \geq 1} \frac{h_{p_i}^{2-\alpha}}{u_{p_i}} \right)^{\frac{1}{2}} \\ &\leq \bigvee g \cdot h_N^{\frac{\alpha}{2}} \text{ recall that } h_n^{2-\alpha} \leq u_n^2. \end{aligned}$$

Finally, we have proven :  $m(B_N) \leq 3(\bigvee g)h_N^{\frac{\alpha}{2}}$ .

The estimation on the measure of  $A_n$  is much simpler. We may cover  $A_n$  by balls of radius  $2h_n$ , centered on points  $x$  such that  $\sup_{d(x,y) < h_n} |g(x) - g(y)| > u_n$ . Then,  $m(A_n) \leq 2 \cdot k \cdot h_n$  where  $k$  is the number of such balls. Now,  $\bigvee g \geq k \cdot u_n$ , thus  $k \leq \frac{\bigvee g}{u_n}$ .  $\square$

**3.1.1. Finite number of pieces.** This subsection is devoted to Lasota-Yorke maps with a finite number of intervals of monotonicity. The main differences between infinite and finite pieces are some additional technical assumptions in the infinite case.

We consider the following conditions that may be found in [Col, Li].

**Assumption 3.**

- (1) The partition  $(I_i)_{i=1,\dots,r}$  of  $I$  is finite, denote by  $\mathcal{P}_k$  the partition of invertibility of  $T^k$ .
- (2)  $T$  satisfies the covering property : for all  $k \in \mathbb{N}$ , there exists  $N(k)$  such that for all  $P \in \mathcal{P}_k$ ,

$$T^{N(k)}P = [0, 1].$$

- (3) On each  $\overline{I_i}$ , the map  $T$  is  $C^2$ .

The following result is standard and may be found in ([Col, Li, Ba]).

**Theorem 3.2.** *Under Assumptions 3, the map  $T$  admits a unique absolutely invariant probability measure, its density  $f$  belongs to  $BV$  and  $\inf f > 0$ . Moreover, if  $\mu = fm$  is this invariant probability measure, then (1.4) is satisfied with  $\mathcal{C} = BV$  and  $\Phi(n) = \gamma^n$ ,  $0 < \gamma < 1$  : there exists  $0 < \gamma < 1$ ,  $C > 0$ , such that for any  $\psi \in BV$  and  $\varphi \in L^1(\mu)$ , for any  $n \in \mathbb{N}$ ,*

$$\left| \int_I \psi \cdot \varphi \circ T^n d\mu - \int_I \psi d\mu \int_I \varphi d\mu \right| \leq C\gamma^n \|\varphi\|_1 \|\psi\|_{BV}.$$

We already know that  $f$  is regular by Lemma 3.1. In order to have the conditions of Corollary 1.2, it remains to prove that  $T$  is regular. This is indeed clear because  $T$  is piecewise  $C^2$ , with finitely many points of discontinuity. As a consequence, for any sequences  $(u_n)_{n \in \mathbb{N}}$  and  $(h_n)_{n \in \mathbb{N}}$  decreasing to zero with  $h_N \leq \frac{u_N}{\sup |T'|}$ ,

$$m(A_N) \leq \sum_{i=1}^p |g_i| \cdot h_N$$

where we have denoted by  $x_i$ ,  $i = 1, \dots, p$  the points of discontinuity, and  $g_i$  the gap at  $x_i$  :  $g_i = |T(x_i^-) - T(x_i^+)|$ . Moreover, as soon as  $u_N \leq \inf g_i$ , we have that  $A_n \supset A_{n+1}$ , so  $B_N = A_N$  and we have the required control on the measure of  $A_N$ .

Thus, Corollary 1.2 applies and we get the following result.

**Corollary 3.3.** *Let  $T$  satisfies Assumption 3, let  $K$  be a Kernel belonging to  $BV$ , let  $\widehat{f}_N$  and  $\widehat{T}_N$  be the estimators of  $f$  and  $T$ . For all  $0 < \alpha < 1$ , here exists  $M > 0$ ,  $L > 0$ ,  $R > 0$ , such that outside a set of measure less than  $Rh^\alpha$ , for all  $t \in \mathbb{R}^+$ , for all  $u \geq Cth$ ,*

$$(3.1) \quad \mathbb{P}(|\widehat{f}_n(x) - f(x)| > t - u^\alpha) \leq 2e^{\frac{1}{e}} \exp[-t^2 Mnh^2]$$

$$(3.2) \quad \mathbb{P}(|\widehat{T}_n(x) - T(x)| > t - u^\alpha) \leq 2e^{\frac{1}{e}} \exp[-t^2 Lnh^2].$$

As a consequence, provided  $h = h_n$  goes to zero and  $nh_n^2 = O(n^\varepsilon)$ ,  $\varepsilon > 0$ , we obtain the following convergences :

- for  $m$  almost all  $x \in \Sigma$ ,  $\widehat{f}_n(x)$  converges to  $f(x)$  and  $\widehat{T}_n(x)$  converges to  $T(x)$  almost surely and in  $L^p$  for any  $1 \leq p$ ,
- $\mathbb{E} \left( \int_I |\widehat{f}_n(x) - f(x)| dx \right)$  and  $\mathbb{E} \left( \int_I |\widehat{T}_n(x) - T(x)| \right)$  go to zero.
- for almost all  $x \in \Sigma$ , for  $a < \frac{1}{2}$ ,  $|\widehat{g}_n(x) - g(x)| = O_{\mathbb{P}}(n^{-a})$  where  $\widehat{g}_n$  is either  $\widehat{f}_n$  or  $\widehat{T}_n$  and  $g$  is either  $f$  or  $T$ .

*Proof.* We apply Corollary 1.2. Since the Kernel  $K$  belongs to  $BV$ , we may take  $\beta = 0$ . Also, with our hypothesis,  $T$  and  $f$  are in  $BV$  so they are regular according to Lemma 3.1.  $\square$

Simulations for  $\beta$ -maps. For  $\beta > 1$ , the  $\beta$ -map is  $T_\beta(x) = \beta x \bmod 1$ . This map has been widely studied. It is well known that it satisfies Assumptions 3.3. It is the simplest example of a non-markov map on the interval. There is an explicit formula for the invariant density (see [Ta]) :

$$f_\beta(x) = C \sum_{i \geq 0} \beta^{-(i+1)} \mathbf{1}_{[0, T^i 1](x)},$$

where  $C$  is a constant chosen so that  $f_\beta$  has integral 1.

We have performed some simulations for several values of  $\beta$ , different Kernels and for several noises  $\varepsilon_i$ . These simulations are summarized in the following table. In this table, AMEf means absolute mean error for the density i.e.

$$\text{AMEf} = \frac{1}{p} \sum_{k=1}^p |\hat{f}_n(x_k) - f(x_k)|,$$

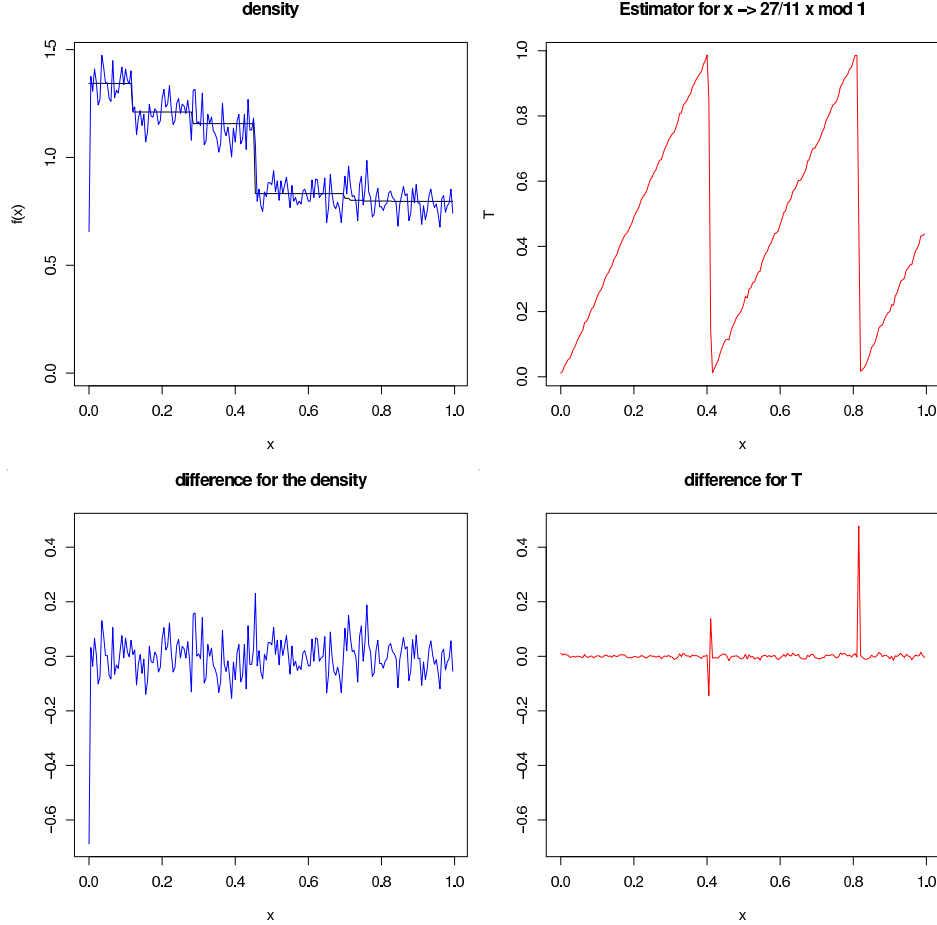
where the  $x_k$  are  $p$  points regularly espaced on  $I$  on which we have calculate  $\hat{f}_n$ . Below,  $p = 200$ . AMET means absolute mean error for the map  $T$  (same formulae). For the Kernel,  $P_2$  is the degree 2 polynomial :  $\frac{3}{4}(1 - x^2)$ .

n	h	AMEf	AMET	$\beta$	kernel	noise, $\varepsilon_i$
$10^4$	0.01	0.08234419	0.008309136	$\frac{27}{11}$	$P_2$	no, $\varepsilon_i = 0$
$10^4$	0.005	0.09906515	0.004301326	$\frac{27}{11}$	$P_2$	no, $\varepsilon_i = 0$
$5 \cdot 10^4$	0.007	0.04428149	0.005530895	$\frac{27}{11}$	$P_2$	no, $\varepsilon_i = 0$
$2 \cdot 10^5$	0.001	0.05107575	0.001799785	$\frac{27}{11}$	$P_2$	no, $\varepsilon_i = 0$
$5 \cdot 10^4$	0.007	0.05492035	0.003809815	$\frac{27}{11}$	$\mathbf{1}_{[-\frac{1}{2}, \frac{1}{2}]}$	no, $\varepsilon_i = 0$
$5 \cdot 10^4$	0.007	0.04728425	0.008303824	$\frac{27}{11}$	$P_2$	$\mathcal{U}[-0.3, 0.3]$
$2 \cdot 10^5$	0.0005	0.07806642	0.011328519	$\frac{27}{11}$	$P_2$	$\mathcal{U}[-0.3, 0.3]$
$10^4$	0.01	0.07473928	0.020744986	$\frac{27}{11}$	$P_2$	$\mathcal{N}(0, 0.3)$
$5 \cdot 10^4$	0.007	0.04269281	0.011423138	$\frac{27}{11}$	$P_2$	$\mathcal{N}(0, 0.3)$
$2 \cdot 10^5$	0.001	0.05107575	0.001799785	$\frac{27}{11}$	$P_2$	$\mathcal{N}(0, 0.3)$
$5 \cdot 10^4$	0.007	0.05329131	0.007722570	$\frac{27}{11}$	$\mathbf{1}_{[-\frac{1}{2}, \frac{1}{2}]}$	$\mathcal{U}[-0.3, 0.3]$
$10^4$	0.01	0.08165332	$1.648713e - 02$	$\frac{46}{11}$	$P_2$	no, $\varepsilon_i = 0$
$5 \cdot 10^4$	0.007	0.04259507	$1.071092e - 02$	$\frac{46}{11}$	$P_2$	no, $\varepsilon_i = 0$
$2 \cdot 10^5$	0.001	0.05249840	$1.536396e - 04$	$\frac{46}{11}$	$P_2$	no, $\varepsilon_i = 0$
$5 \cdot 10^4$	0.007	0.03810643	$1.482175e - 02$	$\frac{46}{11}$	$P_2$	$\mathcal{U}[-0.3, 0.3]$
$5 \cdot 10^4$	0.007	0.03961733	$1.763502e - 02$	$\frac{46}{11}$	$P_2$	$\mathcal{N}(0, 0.3)$
$5 \cdot 10^4$	0.007	0.05467709	$7.079913e - 03$	$\frac{46}{11}$	$\mathbf{1}_{[-\frac{1}{2}, \frac{1}{2}]}$	no
$5 \cdot 10^4$	0.007	0.05109682	$1.036748e - 02$	$\frac{46}{11}$	$\mathbf{1}_{[-\frac{1}{2}, \frac{1}{2}]}$	$\mathcal{U}[-0.3, 0.3]$

*Remark.* It seems that there no real influence of the kind of Kernel nor on the noise. What seems more surprising is the fact that the estimators for  $T$  are much better than for  $f$ . Also, remark that it seems that the best  $h_n$  is not the same for  $f$  and for  $T$ .



The graphics below correspond to  $\beta = \frac{27}{11}$ ,  $K = \mathbf{1}_{[-\frac{1}{2}, \frac{1}{2}]}$ ,  $\varepsilon_i \sim \mathcal{U}[-0.2, 0.2]$ ,  $n = 50000$  and  $h = 0.007$ .



**3.1.2. Infinite number of pieces.** If the number of pieces is infinite, there are many different settings leading to existence and uniqueness of an absolutely continuous invariant probability measure which is exponentially mixing. Let us cite [Bro, LiSV]. We will consider the conditions of Liverani-Saussol-Vaianti [LiSV] in a restricted case (our potential is  $|T'|^{-1}$ , they consider more general potential). These conditions are the following.

- Assumption 4.**
- (1) There is a countable partition  $(I_j)_{j \in \mathbb{N}}$  of  $I$  into intervals. On each  $I_j$ , the map  $T$  is monotonic and  $C^2$ , it is continuous on each  $\overline{I_j}$ . Denote by  $\mathcal{P}$  the partition  $(I_j)_{j \in \mathbb{N}}$  and by  $\mathcal{P}_k$  the partition of monotonicity of  $T^k$ .
  - (2)  $\frac{1}{|T'|} \in BV$  and  $\sum_{P \in \mathcal{P}} \sup \frac{1}{|T'|} < \infty$ .
  - (3) The partition  $\mathcal{P}$  is generating.
  - (4) The map  $T$  is covering, which means, in the infinite case, that for any  $n \in \mathbb{N}$ , for any  $P \in \mathcal{P}_n$ ,  $I$  may be covered by a finite number of

subintervals of  $T^N P$  :

$$\forall n \in \mathbb{N}, \forall P \in \mathcal{P}_n, \exists N, \exists \text{ finite } \mathcal{J} \subset \mathcal{P}_N \bigvee \{P\} / \bigcup_{Q \in \mathcal{J}} T^N Q = X.$$

*Remark.* The condition  $\frac{1}{|T'|} \in BV$  is satisfied provided that  $T$  has the bounded distortion property :

$$\sup_{P \in \mathcal{P}} \sup_{x \in P} \frac{|T''(x)|}{|T'(x)|^2} < \infty$$

and that  $\sum_{P \in \mathcal{P}} \sup \frac{1}{|T'|} < \infty$ .

The following result has been proved in a general setting in [LiSV] and in [Bro] under the condition of bounded distortion.

**Theorem 3.4.** *Under Assumptions 4, the map  $T$  admits a unique absolutely invariant probability measure, its density  $f$  belongs to  $BV$  and  $\inf f > 0$ . Moreover, if  $\mu = fm$  is this invariant probability measure, then (1.4) is satisfied with  $\mathcal{C} = BV$  and  $\Phi(n) = \gamma^n$ ,  $0 < \gamma < 1$  : there exists  $0 < \gamma < 1$ ,  $C > 0$ , such that for any  $\psi \in BV$  and  $\varphi \in L^1(\mu)$ , for any  $n \in \mathbb{N}$ ,*

$$\left| \int_I \psi \cdot \varphi \circ T^n d\mu - \int_I \psi d\mu \int_I \varphi d\mu \right| \leq C \gamma^n \|\varphi\|_1 \|\psi\|_{BV}.$$

**Lemma 3.5.** *The map  $T$  satisfying Assumptions 4 is regular.*

*Idea of the proof.* This is a simple consequence of the fact that  $T$  is piecewise  $C^2$  (thus piecewise  $C^1$ ) and that  $\sum_{P \in \mathcal{P}} \sup \frac{1}{|T'|} < \infty$ .  $\square$

There are several natural examples of dynamical systems satisfying Assumptions 4. Let us cite the Gauss map :

$$T(x) = \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor$$

which appears in the continuous fractions decomposition ([Bro]) and in analysis a gcd algorithms ([V]). An natural extension of these systems are ‘‘Japanese systems’’ or  $\alpha$ -Gauss maps : for  $0 < \alpha \leq 1$ ,

$$T_\alpha(x) = \left\lfloor \frac{1}{x} \right\rfloor - \left[ \left\lfloor \frac{1}{x} \right\rfloor + 1 - \alpha \right],$$

$T_\alpha$  maps the interval  $[\alpha - 1, \alpha]$  into itself. See [BoDaV] for a description of these systems as well as an application to analysis of generalized Euclidian algorithms. The maps  $T_\alpha$  satisfy Assumption 4 for  $0 < \alpha \leq 1$ .

**Corollary 3.6.** *Let  $T$  satisfies Assumption 4, let  $K$  be a Kernel belonging to  $BV$ , let  $\hat{f}_N$  and  $\hat{T}_N$  be the estimators of  $f$  and  $T$ . For all  $0 < \alpha < 1$ , here*

exists  $M > 0$ ,  $L > 0$ ,  $R > 0$ , such that outside a set of measure less than  $Rh^\alpha$ , for all  $t \in \mathbb{R}^+$ , for all  $u \geq Cth$ ,

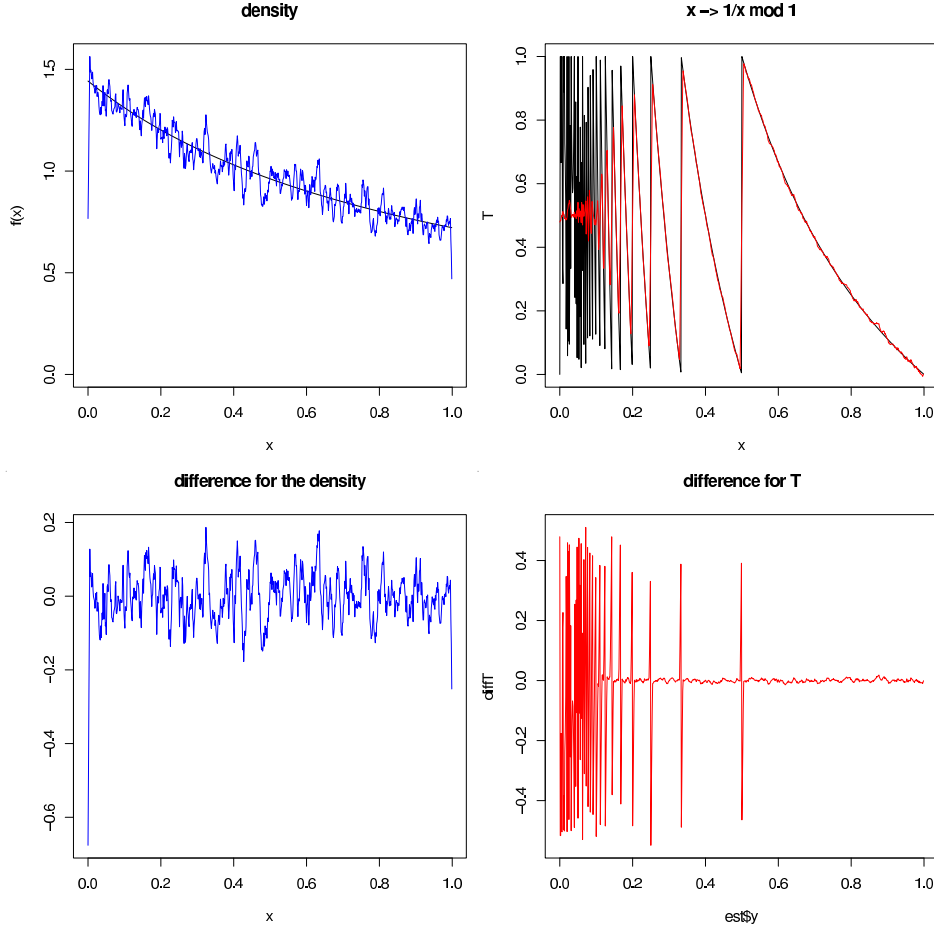
$$(3.3) \quad \mathbb{P}(|\widehat{f}_n(x) - f(x)| > t - u^\alpha) \leq 2e^{\frac{1}{\varepsilon}} \exp[-t^2 Mnh^2]$$

$$(3.4) \quad \mathbb{P}(|\widehat{T}_n(x) - T(x)| > t - u^\alpha) \leq 2e^{\frac{1}{\varepsilon}} \exp[-t^2 Lnh^2].$$

As a consequence, provided  $h = h_n$  goes to zero and  $nh_n^2 = O(n^\varepsilon)$ ,  $\varepsilon > 0$ , we obtain the following convergences :

- for  $m$  almost all  $x \in I$ ,  $\widehat{f}_n(x)$  converges to  $f(x)$  and  $\widehat{T}_n(x)$  converges to  $T(x)$  almost surely and in  $L^p$  for any  $1 \leq p$ ,
- $\mathbb{E} \left( \int_I |\widehat{f}_n(x) - f(x)| dx \right)$  and  $\mathbb{E} \left( \int_I |\widehat{T}_n(x) - T(x)| \right)$  go to zero.
- for almost all  $x \in I$ , for  $a < \frac{1}{2}$ ,  $|\widehat{g}_n(x) - g(x)| = O_{\mathbb{P}}(n^{-a})$  where  $\widehat{g}_n$  is either  $\widehat{f}_n$  or  $\widehat{T}_n$  and  $g$  is either  $f$  or  $T$ .

The following graphics are for the Gauss map  $T(x) = \frac{1}{x} \pmod{1}$  for which it is known that the unique invariant probability density is  $f(x) = \frac{1}{\log 2} \frac{1}{1+x}$ . We have performed simulations with  $n = 50000$ ,  $h = 0.009$  and the noise  $\varepsilon_i \sim \mathcal{U}[-0.2, 0.2]$ . Over 800 points, we get  $EMEf = 0.05046933$ ,  $AMET = 0.05141938$ , if we restrict to points in  $[0.2, 1]$  we get  $AMET = 0.01439787$ .



**3.2. In dimension one : unimodal maps.** Let  $I = [-1, 1]$  and  $f : I \rightarrow I$  be a  $C^2$  unimodal map (i.e.,  $f$  is increasing on  $[-1, 0]$ , decreasing on  $[0, 1]$ ) satisfying  $f''(0) \neq 0$ , and,

**(H1)** There are  $0 < \alpha < 1$ ,  $K > 1$ , and  $\tilde{\lambda} \leq \lambda \leq 4$  with  $e^{2\alpha} < \tilde{\lambda}$ , and  $\sup_I |f'| \leq \tilde{\lambda}^K < 8$  so that

(i)  $|(f^n)'(f(0))| \geq \tilde{\lambda}^n$  for all  $n \in \mathbb{N}$  and  $\lambda = \lim_{n \rightarrow \infty} |(f^n)'(f(0))|^{1/n}$ .

(ii)  $|f^n(0)| \geq e^{-\alpha n}$ , for all  $n \geq 1$ .

**(H2)** For each small enough  $\delta > 0$ , there is  $M = M(\delta) \in \mathbb{N}_+$  for which

(i) If  $x, \dots, f^{M-1}(x) \notin (-\delta, \delta)$  then  $|(f^M)'(x)| \geq \tilde{\lambda}^M$ ; (ii) For each  $n$ , if  $x, \dots, f^{n-1}(x) \notin (-\delta, \delta)$  and  $f^n(x) \in (-\delta, \delta)$ , then  $|(f^n)'(x)| \geq \tilde{\lambda}^n$ .

**(H3)**  $f$  is topologically mixing on  $[f^2(0), f(0)]$ , that is for any two open sets  $U, V \subset I$ , there exists  $N \in \mathbb{N}$  such that  $\forall n \geq N, T^{-n}U \cap V \neq \emptyset$ .

Examples of unimodal maps satisfying (H1)–(H3) are quadratic maps  $1 - a \cdot x^2$  for a positive measure set of parameters  $a$ . (See e.g. [BeCa]).

The following theorem is obtained in two steps. First, it is proven that unimodal maps satisfying (H1)–(H3) are conjugated to a kind of hyperbolic markov maps called Young towers (see [Y1, KeNo]). Then the estimation on the speed of mixing is obtained on the tower ([BuMau1]). Let us emphasize the fact that the kind a mixing we need (namely 1.4) and especially the fact that  $\|\varphi\|_1$  appears, is not obtained in [Y1, KeNo] (a form with the  $\|\varphi\|_\infty$  was obtained there).

**Theorem 3.7.** *Let  $T$  be a unimodal map satisfying (H1)–(H3). The map  $T$  admits a unique absolutely invariant probability measure with density  $f$  satisfying  $\inf_{[f^2(0), f(0)]} f > 0$ . Moreover, if  $\mu = f\mu$  is this invariant probability measure, then (1.4) is satisfied with  $\mathcal{C}$  the space  $\text{Lip}$  of Lipschitz functions on  $I$  and  $\Phi(n) = \gamma^n$ ,  $0 < \gamma < 1$  : there exists  $0 < \gamma < 1$ ,  $C > 0$ , such that for any  $\psi \in \text{Lip}$  and  $\varphi \in L^1(\mu)$ , for any  $n \in \mathbb{N}$ ,*

$$\left| \int_I \psi \cdot \varphi \circ T^n d\mu - \int_I \psi d\mu \int_I \varphi d\mu \right| \leq C \gamma^n \|\varphi\|_1 \|\psi\|_{\text{Lip}}.$$

This is clear that  $T$  is 1-regular : for well chosen  $\varepsilon_n$  and  $u_n$ , the sets  $A_N$  and  $B_N$  are empty because  $T$  is  $C^2$ . Nevertheless, it is known that the invariant density is very irregular (see [Y2, KeNo]). With a more intricate study, we should probably prove that  $f$  is also regular. Here, we restrict ourselves to the estimation of  $T$ .

Since the invariant measure has its support in  $S = [f^2(0), f(0)]$ , our estimates are valid only for  $x \in S$ .

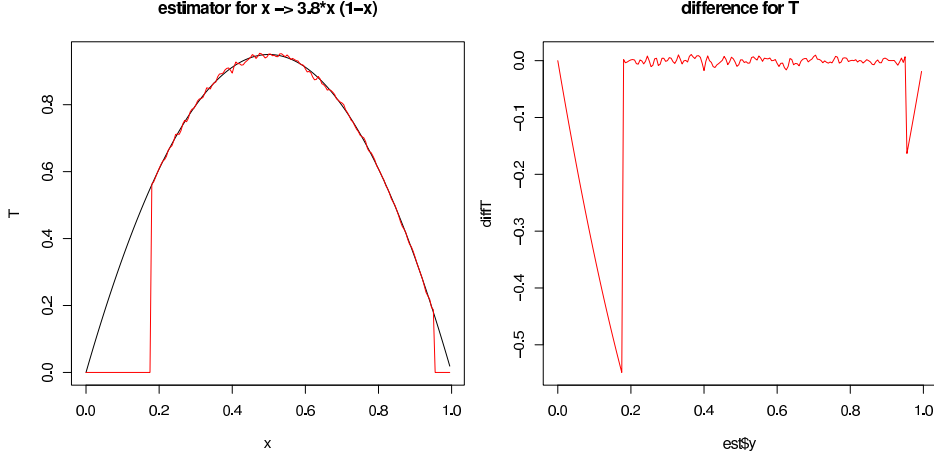
**Corollary 3.8.** *Let  $T$  be a unimodal map satisfying (H1)–(H3), let  $K$  be a Kernel belonging to  $\text{Lip}$ , let  $\hat{T}_N$  be the estimator of  $T$ . There exists  $M > 0$ ,  $L > 0$ ,  $R > 0$ , such that outside a subset of  $S$ , of measure less than  $Rh$ , for all  $t \in \mathbb{R}^+$ , for all  $u \geq Cth$ ,*

$$\mathbb{P}(|\hat{T}_n(x) - T(x)| > t - u) \leq 2e^{\frac{1}{\varepsilon}} \exp[-t^2 Lh^2 n].$$

As a consequence, provided  $h = h_n$  goes to zero and  $nh_n^2 = O(n^\varepsilon)$ ,  $\varepsilon > 0$ , we obtain the following convergences :

- for  $m$  almost all  $x \in S$ ,  $\widehat{T}_n(x)$  converges to  $T(x)$  almost surely and in  $L^p$  for any  $1 \leq p$ ,
- $\mathbb{E} \left( \int_S |\widehat{T}_n(x) - T(x)| \right)$  go to zero.
- for almost all  $x \in S$ , for  $a < \frac{1}{2}$ ,  $|\widehat{T}_n(x) - T(x)| = O_{\mathbb{P}}(n^{-a})$ .

The following graphics are for the map  $T(x) = 3.8 * x * (1 - x)$ . We have performed simulations with  $n = 50000$ ,  $h = 0.01$  and the noise  $\varepsilon_i \rightsquigarrow \mathcal{U}[-0.2, 0.2]$ . Over 154 points in  $S$ , we get  $AMET = 0.004114143$ .



**3.3. Piecewise expanding maps in higher dimension.** These are generalisations of Lasota-Yorke maps in higher dimension. These maps have been studied in [Bu4, BuPaS, Bu1, Bu2, Bu5, Cow, GoBo, Sau] from various point of view. The control of the speed of mixing may be found in [BuMau1, BuMau2], the strategy is as for unimodal maps : the map is conjugated to a “Young tower”.

The setting is the following.  $(X, \mathcal{Z}, T)$  will be a *piecewise invertible map*, i.e.:

- $X = \overline{\bigcup_{Z \in \mathcal{Z}} Z}$  is a locally connected compact subset of  $\mathbb{R}^d$ .
- $\mathcal{Z}$  is a finite collection of pairwise disjoint, bounded and open subsets of  $X$ , each with a non-empty boundary. Let  $Y = \bigcup_{Z \in \mathcal{Z}} Z$ .
- $T : Y \rightarrow X$  is a map such that each restriction  $T|_Z$ ,  $Z \in \mathcal{Z}$ , coincides with the restriction of a homeomorphism  $T_Z : U \rightarrow V$  with  $U, V$  open sets such that  $U \supset \bar{Z}$ ,  $V \supset \overline{T(Z)}$ .

$T$  will be assumed to be *non-contracting*, i.e., such that for all  $x, y$  in the same element  $Z \in \mathcal{Z}$ ,  $d(Tx, Ty) \geq d(x, y)$ . Also  $\mathcal{Z}$  will be assumed to be *generating*, i.e.,  $\lim_{n \rightarrow \infty} \text{diam}(\mathcal{Z}^n) = 0$  where  $\mathcal{Z}^n$  denotes the set of  $n$ -cylinders, i.e., the non-empty sets of the form:

$$[A_0 \dots A_{n-1}] := A_0 \cap \dots \cap T^{-n+1} A_{n-1}$$

for  $A_0, \dots, A_{n-1} \in \mathcal{Z}$ . Finally the *boundary of the partition*,  $\partial \mathcal{Z} = \bigcup_{Z \in \mathcal{Z}} \partial Z$ , will play an important role in our analysis. In particular, we shall assume “small boundary pressure” (see below), a fundamental condition which already appeared in [Bu4, Bu5, BuPaS]. To formulate the crucial “small

boundary pressure" condition, we need first some definitions. The *topological pressure* [DGS] of a subset  $S$  of  $X$  is:

$$P(S, T) = \limsup_{n \rightarrow \infty} \frac{1}{n} \log \sum_{\substack{A \in \mathcal{Z}^n \\ A \cap S \neq \emptyset}} g^{(n)}(A)$$

where  $g^{(n)}(A) = \sup_{x \in A} g(x)g(Tx) \dots g(T^{n-1}x)$ ,  $g = |\det T'|^{-1}$ . The *small boundary pressure* condition is:

$$P(\partial \mathcal{Z}, T) < P(X, T).$$

This inequality is satisfied in many cases. In particular, if  $T$  is expanding and  $X$  is a Riemannian manifold and the weight is  $|\det T'(x)|^{-1}$  or close to it, then it is satisfied: (i) in dimension 1, in all cases; (ii) in dimension 2, if  $T$  is piecewise real analytic [Bu5, Ts1]; (iii) in arbitrary dimension, for all piecewise affine  $T$  [Ts2] or for generic  $T$  [Bu3, Cow].

A basic example is given by the *multidimensional  $\beta$ -transformations* [Bu1], i.e., maps  $T : [0, 1]^d \rightarrow [0, 1]^d$ ,  $T(x) = B(x) \bmod \mathbb{Z}^d$  with  $B$  an expanding affine map on  $\mathbb{R}^d$ . Let us summarize our hypothesis on  $T$ .

**Assumption 5.** Let  $(X, \mathcal{Z}, T, g)$  be a weighted piecewise invertible dynamical system. Assume that:

- $T$  is expanding, i.e., there is some  $\lambda > 1$  such that for all  $x, y$  in the same element of  $\mathcal{Z}$ ,  $d(Tx, Ty) \geq \lambda \cdot d(x, y)$ ;
- $g = |\det T'|^{-1}$  is Hölder continuous with exponent  $\gamma$  and is positively lower bounded.
- the boundary pressure is small:  $P(\partial \mathcal{Z}, T) < P(X, T)$ .
- $T$  is topologically mixing.

Let  $K(f) = \max_{Z \in \mathcal{Z}} \sup_{x \neq y \in Z} \frac{|f(x) - f(y)|}{d(x, y)^\gamma}$  where  $\gamma$  is some Hölder exponent of  $g$ .

**Theorem 3.9.** [BuMau1] *Let  $T$  satisfy Assumption 5. Then,  $T$  admits a unique invariant measure  $\mu$ , absolutely continuous w.r.t. the Lebesgue measure  $m$ . This measure is exponentially mixing :*

$$\left| \int_X \varphi \circ T^n \cdot \psi d\mu - \int_X \varphi d\mu \int_X \psi d\mu \right| \leq C \cdot \|\varphi\|_{C^\gamma(X)} \cdot \|\psi\|_{L^1} \kappa^n.$$

with constants  $C < \infty$  and  $\kappa < 1$  depending only on  $(X, \mathcal{Z}, T, g)$ , for any measurable functions  $\varphi, \psi : X \rightarrow \mathbb{R}$  such that  $\psi$  is bounded and  $\varphi$  is  $\gamma$ -Hölder continuous.

This is clear that  $T$  is 1-regular because it is piecewise  $C^1$ . Nevertheless, it is known that the invariant density has discontinuities on  $\partial[T^n(\mathcal{Z}^n)]$ . With a more intricate study, we should probably prove that  $f$  is also regular. Here, we restrict ourselves to the estimate of  $T$ .

**Corollary 3.10.** *Let  $T$  satisfy Assumption 5, let  $K$  be a  $\gamma$ -Hölder Kernel, let  $\widehat{T}_N$  be the estimator of  $T$ . There exists  $M > 0$ ,  $L > 0$ ,  $R > 0$ , such that outside a set of measure less than  $Rh$ , for all  $t \in \mathbb{R}^+$ , for all  $u \geq Cth$ ,*

$$\mathbb{P}(|\widehat{T}_n(x) - T(x)| > t - u) \leq 2e^{\frac{1}{e}} \exp[-t^2 Lh^{\gamma+2}n].$$

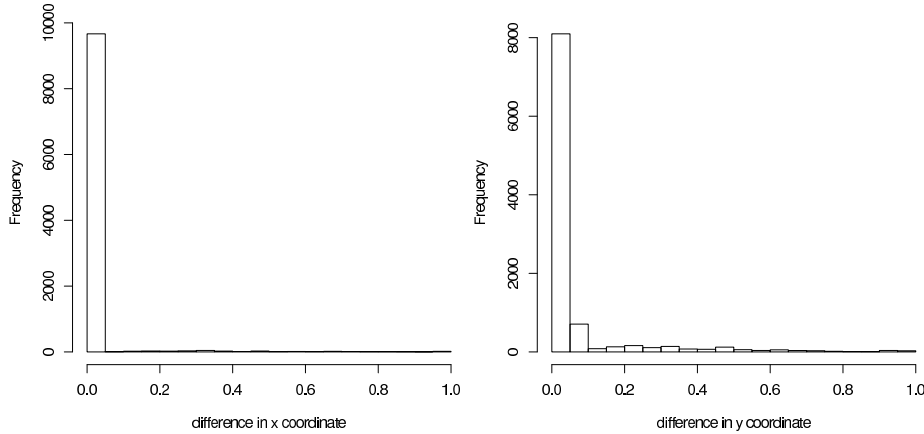
As a consequence, provided  $h = h_n$  goes to zero and  $nh_n^{\gamma+2} = O(n^\varepsilon)$ ,  $\varepsilon > 0$ , we obtain the following convergences :

- for  $m$  almost all  $x \in X$ ,  $\widehat{T}_n(x)$  converges to  $T(x)$  almost surely and in  $L^p$  for any  $1 \leq p$ ,
- $\mathbb{E} \left( \int_X |\widehat{T}_n(x) - T(x)| \right)$  go to zero.
- for almost all  $x \in X$ , for  $a < \frac{1}{2}$ ,  $|\widehat{T}_n(x) - T(x)| = O_{\mathbb{P}}(n^{-a})$ .

We have performed some simulations for  $T(x) = Bx \bmod \mathbb{Z}^2$ , with  $B$  the matrix

$$\begin{pmatrix} 2.5 & 3.4 \\ 4.6 & 3.2 \end{pmatrix}$$

Below are the histograms for  $diff_x$  (resp.  $diff_y$ ), the difference between the  $x$  (resp.  $y$ ) coordinate of  $T$  and the  $x$  (resp.  $y$ ) coordinate of  $\widehat{T}_n$ , over a grid of 100 times 100 points in  $[0, 1]^2$ . The Kernel is  $K = \frac{1}{4} \mathbf{1}_{[-1,1] \times [-1,1]}$ , there is no noise,  $\varepsilon_i = 0$ ,  $n = 66668$  and  $h = 0.004$ . The AME for the coordinate  $x$  is 0.01882885, for the coordinate  $y$ , the AME is 0.06723186.



*Remark* (Anosov maps). Our technics should also apply to estimate the invariant density and the application  $T$  for Anosov maps for which there exists an invariant measure absolutely continuous with respect to the Lebesgue measure. A more intricate study could also lead results of the same kind for Axiom A diffeomorphisms.

## REFERENCES

- [ABST] H. D Abarbanel, R. Brown, J. J. Sidorowich, L. S. Tsimring, The analysis of observed chaotic data in physical systems. *Rev. Modern Phys.* 65 (1993), no. 4, 1331-1392.
- [Ba] V. Baladi *Positive Transfer Operators and Decay of Correlations*, Book, Advanced Series in Nonlinear Dynamics, Vol 16, World Scientific, Singapore (2000).
- [BeCa] M. Benedicks, L. Carleson *On iterations of  $1 - ax^2$  on  $(-1,1)$*  *Ann. of Math.* (2), 122, (1985), 1-25
- [Bro] A. Broise, *Transformations dilatantes de l'intervalle et théorèmes limites*. Études spectrales d'opérateurs de transfert et applications. *Astérisque* 1996, no. 238, 1-109.

- [BoDaV] J. Bourdon, B. Daireaux, B. Vallée *Dynamical Analysis of  $\alpha$ -Euclidean Algorithms*, Journal of Algorithms, **44**, (1), (2002), pp. 246-285.
- [Bu1] J. Buzzi Intrinsic ergodicity of affine maps in  $[0, 1]^d$ . *Monatsh. Math.* 124 (1997), 97–118.
- [Bu2] J. Buzzi *Markov extensions for multi-dimensional dynamical systems*. Israel J. Math. 112 (1999), 357–380.
- [Bu3] J. Buzzi *Absolutely continuous invariant measures for generic piecewise affine and expanding maps* Int. J. Chaos & Bif. (1999), **9** (9), 1743-1750.
- [Bu4] J. Buzzi *Thermodynamical formalism for piecewise invertible maps : absolutely continuous invariant measures as equilibrium states* in Smooth Ergodic Theory and its Applications (Seattle, WA 1999), Proceedings of Symposia in Pure Mathematics **69**, AMS, RI, (2001), 749–783.
- [Bu5] J. Buzzi *Absolutely continuous invariant probability measures for arbitrary expanding piecewise  $\mathbf{R}$ -analytic mappings of the plane*. Ergodic Theory Dynam. Systems **20** (2000), 697–708.
- [BuMau1] J. Buzzi, V. Maume-Deschamps *Decay of correlations on towers for potentials with summable variation*, Discrete and Continuous Dynamical Systems, 12, (2005), no 4, 639-656.
- [BuMau2] J. Buzzi, V. Maume-Deschamps *Decay of correlations for piecewise invertible maps in higher dimensions*, Israel J. Math. 131 (2002), 203-220.
- [BuPaS] J. Buzzi, F. Paccaut, B. Schmitt, *Conformal measures for multidimensional piecewise invertible maps*. Ergodic Theory Dynam. Systems 21 (2001), no. 4, 1035–1049.
- [Col] P. Collet, *Some ergodic properties of maps of the interval*. Dynamical systems (Temuco, 1991/1992), 55–91, Travaux en Cours, **52**, Hermann, Paris, 1996.
- [CMarS] P. Collet, S. Martinez, B. Schmitt, *Exponential inequalities for dynamical measures of expanding maps of the interval*. Probab. Theory Related Fields 123 (2002), no. 3, 301–322.
- [Cow] W. Cowieson, *Piecewise smooth expanding maps in  $\mathbb{R}^d$* , Ph.D. Thesis, U.C.L.A., Berkeley (see <http://math.usc.edu/~cowieson>).
- [DeDo] J. Dedecker, P. Doukhan, *A new covariance inequality and applications*. Stochastic Process. Appl. 106 (2003), no. 1, 63–80.
- [DGS] M. Denker, C. Grillenberger, K. Sigmund *Ergodic theory on compact spaces*. Lecture notes in mathematics **527**, Springer, Berlin, 1976.
- [DeP] J. Dedecker, C. Priour, *New dependence coefficients. Examples and applications to statistics*. To appear in Probab. Theory and Relat. Fields.
- [DoLo] P. Doukhan, S. Louhichi, *A new weak dependence condition and applications to moment inequalities*. Stochastic Process. Appl. 84 (1999), no. 2, 313–342.
- [FV] F. Ferraty, P. Vieu, *Nonparametric models for functional data, with applications in regression, time series prediction and curves discrimination*. J. of Nonparametric Statistics, 16, 111-125, (2004).
- [GoBo] P. Góra, A. Boyarsky *Absolutely continuous invariant measures for piecewise expanding  $C^2$  transformation in  $R^N$* . *Israel J. Math.* **67** (1989), 272–286.
- [KeNo] G. Keller, T. Nowicki *Spectral theory, zeta functions and the distribution of periodic points for Collet-Eckmann maps*. Comm. Math. Phys. (1992), **149**, 1, 31-69.
- [Li] C. Liverani, *Decay of correlations for piecewise expanding maps*. J. Statist. Phys. 78 (1995), no. 3-4, 1111–1129.
- [LiSV] C. Liverani, B. Saussol, S. Vaienti, *Conformal measure and decay of correlation for covering weighted systems*. Ergodic Theory Dynam. Systems 18 (1998), no. 6, 1399–1420.
- [Mae] J. Maes *Estimation non paramétrique pour des processus dynamiques dilatants*. C. R. Acad. Sci. Paris Sér. I Math. 330 (2000), no. 9, 831–834.
- [Mas] E. Masry *Nonparametric regression estimation for dependent functional data: asymptotic normality*. Stochastic Processes and their Applications, 115, (2005), 155–177.



- [Mau1] V. Maume-Deschamps *Projective metrics and mixing properties on towers*, Trans. A.M.S. **353**, (2001), 8, 3371-3389.
- [Mau2] V. Maume-Deschamps *Exponential inequalities and estimation of conditional probabilities*. dans *Dependence in probability and statistics*, **Lect. notes in Stat., Springer**, Vol. 187 Bertail, Patrice; Doukhan, Paul; Soulier, Philippe (Eds.), (2006). <http://www.springer.com/sgw/cda/frontpage/0,11855,1-0-22-122342957-0,00.html?referer=www.springer.com>
- [P] C. Prieur *Density estimation for one-dimensional dynamical systems*. ESAIM Probab. Statist. 5 (2001), 51–76.
- [R] M. Rosenblatt, *Conditional probability density and regression estimators*. Multivariate Analysis, II (Proc. Second Internat. Sympos., Dayton, Ohio, 1968), (1969), pp. 25–31 Academic Press, New York.
- [Sau] B. Saussol Absolutely continuous invariant measures for multidimensional expanding maps. *Israel J. Math.* **116** (2000), 223–248.
- [Ta] Y. Takahashi,  *$\beta$ -transformations and symbolic dynamics*. Proceedings of the Second Japan-USSR Symposium on Probability Theory (Kyoto, 1972), pp. 455–464. Lecture Notes in Math., Vol. 330, Springer, Berlin, (1973).
- [Ts1] M. Tsujii, Absolutely continuous invariant measures for piecewise real-analytic expanding maps on the plane. *Comm. Math. Phys.* **208** (2000), 605–622.
- [Ts2] M. Tsujii, Absolutely continuous invariant measures for expanding piecewise linear maps *Invent. Math.* (to appear).
- [Ts3] M. Tsujii, Piecewise Expanding maps on the plane with singular ergodic properties. *Ergod. th. and Dynam. Syst.* (to appear).
- [V] B. Vallée *Digits and continuants in Euclidean algorithms. Ergodic versus Tauberian theorems*. J. Théor. Nombres Bordeaux 12 (2000), no. 2, 531–570.
- [Y1] L.-S. Young *Statistical properties of dynamical systems with some hyperbolicity*. Ann. of Math. (2) (1998), **147**, 3, 585-650.
- [Y2] L.-S. Young *Recurrence times and rates of mixing*. Israel J. Math. 110 (1999), 153–188.

UNIVERSITÉ DE BOURGOGNE B.P. 47870 21078 DIJON CEDEX FRANCE  
*E-mail address:* vmaume@u-bourgogne.fr