Weak convergence of the empirical process of intermittent maps in L2 under long-range dependence

Jérôme Dedecker, Herold G. Dehling, Murad Taqqu

Weak convergence of the empirical process of intermittent maps in $\mathbb{L}^2$ under long-range dependence.

Jérôme Dedecker*, Herold Dehling† and Murad S. Taqqu‡

November 22, 2013

Abstract

We study the behavior of the empirical distribution function of iterates of intermittent maps in the Hilbert space of square inegrable functions with respect to Lebesgue measure. In the long-range dependent case, we prove that the empirical distribution function, suitably normalized, converges to a degenerate stable process, and we give the corresponding almost sure result. We apply the results to the convergence of the Wasserstein distance between the empirical measure and the invariant measure. We also apply it to obtain the asymptotic distribution of the corresponding Cramér-von-Mises statistic.

Keywords. Long-range dependence, intermittency, empirical process.

Mathematics Subject Classification (2010). 60F17, 60E07, 37E05.

1 Introduction and main results.

For γ in $]0,1[$, we consider the intermittent map T_γ (or simply T) from $[0,1]$ to $[0,1]$, introduced by Liverani, Saussol and Vaienti (1999):

$$T_\gamma(x) = \begin{cases} x(1+2^\gamma x^\gamma) & \text{if } x \in [0,1/2[\\ 2x-1 & \text{if } x \in [1/2,1]; \end{cases}$$

see Figure 1 for the graph of T_γ.

This kind of maps are known to exhibit a transition from a stable periodic behavior to a chaotic one, as described in Pomeau and Manneville (1980). Concerning the existence of T_γ-invariant probability measures which are absolutely continuous with respect to the Lebesgue measure, it follows from Thaler (1980) that:

- if $\gamma \in]0,1[$, there exists a unique absolutely continuous T_γ-invariant probability measure ν_γ (or simply ν) on $[0,1]$;
- if $\gamma \geq 1$, there is no absolutely continuous invariant probability measure.

For x near the neutral fixed point 0, the sequence $T^k(x), k \geq 0$, spends a lot of time around 0, since $T(x) = x(1+(2x)^\gamma) \approx x$ for $x \approx 0$; see Figure 2 for the time series of 500 iterations of $T_{0.5}$ (left) and $T_{0.9}$ (right). Note that the length of the periods spent in the neighborhood of zero increases as γ gets larger. Hence the density h_γ (or simply h) of the invariant distribution ν_γ explodes in the neighborood of 0. Even though no explicit formula is known for h, we can give a precise description thanks to the works by Thaler (1980) and Liverani, Saussol and Vaienti

*Université Paris Descartes, Sorbonne Paris Cité, Laboratoire MAP5 and CNRS UMR 8145, 75016 Paris, France.
†Faculty of Mathematics, Ruhr-University Bochum, 44780 Bochum, Germany.
‡Department of Mathematics, Boston University, Boston MA 02215, USA

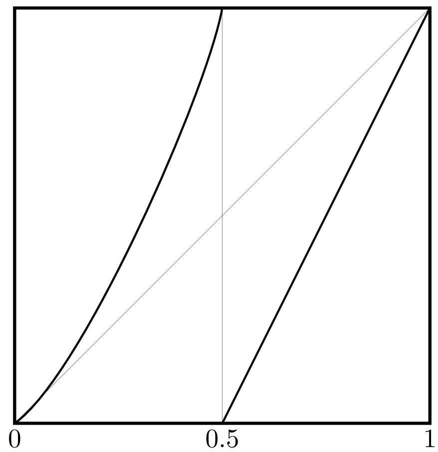

Figure 1: Graph of intermittent map $T_\gamma : [0,1] \to [0,1]$

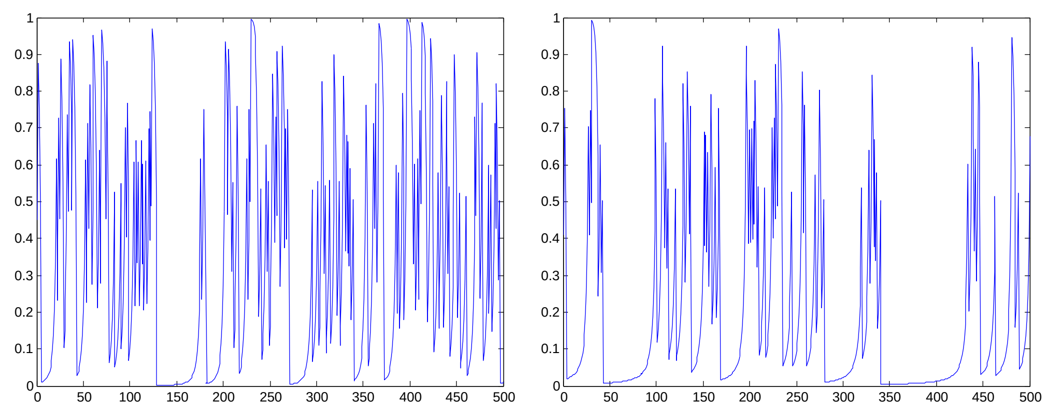

Figure 2: Time series of 500 iterations of the intermittent map $T_{0.5}$ (left) and $T_{0.9}$ (right).

(1999). From the first paper, we infer that the function $x \mapsto x^\gamma h(x)$ is bounded from above and below. From the second paper, we know that h is non increasing with $h(1) > 0$, and that it is Lipshitz on any interval $[a, 1]$ with $a > 0$. Since h is strictly positive on $[0, 1]$, the probability ν is equivalent to the Lebesgue measure on $[0, 1]$.

For $x \in [\epsilon, 1]$, the map T is expanding, i.e. $|T'(x)| \geq \alpha$, for some $\alpha = \alpha(\epsilon) > 1$, and thus the sequence $T^k(x)$ exhibits random behavior. As $T(1/2 + \epsilon) = 2\epsilon$, the sequence occasionally returns to a neighborhood of $x = 0$. Then, as $T(x) \approx x$ for $x \approx 0$, the sequence $T^k(x)$ hovers around the neighborhood of zero for a long time, which may explain the long range dependence of the process T^k, $k \geq 1$. Eventually, since $T(x) = x + 2^\gamma x^{\gamma+1} > x$, and $T'(x) = 1 + 2^\gamma(\gamma + 1)\gamma x^\gamma > 1$, the process escapes the neighborhood of zero and reenters the zone of chaotic behaviour.

From now on, we shall use the notation

$$\nu(f) = \int f(x)\nu(dx)\,,$$

which is valid for any $f \in \mathbb{L}^1([0, 1], \nu)$. For $\gamma \in]0, 1[$, we view T^k as a random variable from the probability space $([0, 1], \nu)$ to $[0, 1]$. The fact that ν is invariant by T implies that $\nu(f) = \nu(f \circ T^k)$ for any $f \in \mathbb{L}^1([0, 1], \nu)$, and more generally, it also implies that the process $(T^k)_{k\geq 0}$ is strictly stationary. Let us briefly recall some known results about the iterates of T:

1. *Decay of correlations.* For any bounded function f and any Hölder functions g, Young (1999) proved the following decay of the covariances

$$\nu\Big((g-\nu(g))\cdot(f-\nu(f))\circ T^n\Big)=\nu\Big((g-\nu(g))\cdot f\circ T^n\Big)=O(n^{(\gamma-1)/\gamma}), \qquad (1.1)$$

as $n\to\infty$. Some lower bounds for the covariance in (1.1) can be found in the paper by Sarig (2002), proving that the rate $n^{(\gamma-1)/\gamma}$ is optimal. Dedecker, Gouëzel and Merlevède (2010) have shown that (1.1) remains true if g is any bounded variation function.

2. *Behaviour of Birkhoff sums.* Liverani, Saussol and Vaienti (1999) have proved that the map T is mixing in the ergodic theoretic sense. Let then

$$s_n(f)=\sum_{k=1}^{n} f\circ T^k\,.$$

For any $f\in\mathbb{L}^1([0,1],\nu)$, it follows from Birkhoff's ergodic theorem that $n^{-1}s_n(f)$ converges to $\nu(f)$ almost everywhere.

Concerning the convergence in distribution of the sequence $s_n(f)-n\nu(f)$ (suitably normalized) on the probability space $([0,1],\nu)$, we must distinguish three cases. If $\gamma\in]0,1/2[$ and f is any Hölder function, Young (1999) proved that the sequence $n^{-1/2}(s_n(f)-n\nu(f))$ converges in distribution to a normal law. Next, Gouëzel (2004) has given a complete picture of the convergence in distribution of $s_n(f)-n\nu(f)$ when $\gamma\in[1/2,1[$. More precisely, if $\gamma=1/2$ and f is any Hölder function, he proved that the sequence $(n\log(n))^{-1/2}(s_n(f)-n\nu(f))$ converges in distribution to a normal law. If $\gamma\in]1/2,1[$ and f is any Hölder function, he proved that $n^{-\gamma}(s_n(f)-n\nu(f))$ converges in distribution to a stable law of index $1/\gamma$.

The power decay of the covariance (1.1) suggests that there may be long-range dependence for some values of γ. A finite variance stationary process with covariance $r(n)$ is said to be *short-range dependent* if $\sum_{n=0}^{\infty}|r(n)|<\infty$ and *long-range dependent* if $\sum_{n=0}^{\infty}|r(n)|=\infty$. In view of the optimality of (1.1), we see that the process $(T^k)_{k\geq 0}$ is short-range dependent if $\gamma\in]0,1/2[$ and long-range dependent if $\gamma\in]1/2,1[$. The case $\gamma=1/2$ is a boundary case. Moreover, the asymptotic behaviour of the normalized sums $s_n(f)$ is normal in the short-range dependent case (including $\gamma=1/2$), and is stable in the long-range dependent case.

Our aim is to study the limit in distribution of the empirical process

$$G_n(t)=\frac{1}{n}\sum_{k=1}^{n}\Big(\mathbf{1}_{T^k\leq t}-F(t)\Big),\quad t\in[0,1], \qquad (1.2)$$

with $F(t)=\nu([0,t])$, in the case where $\gamma\in[1/2,1[$.

Let us introduce another stationary process with the same law. Let first K be the Perron-Frobenius operator of T with respect to ν, defined as follows: for any functions f,g in $\mathbb{L}^2([0,1],\nu)$

$$\int f(T(x))g(x)\nu(dx)=\int f(x)(Kg)(x)\nu(dx)\,. \qquad (1.3)$$

The relation (1.3) states that K is the adjoint operator of the isometry $U:f\mapsto f\circ T$ acting on $\mathbb{L}^2([0,1],\nu)$. It is easy to see that the operator K is a transition kernel.[1] Let now $(X_i)_{i\in\mathbb{Z}}$ be a stationary Markov chain with invariant measure ν and transition kernel K. It is well known (see

[1]Indeed, by stationarity, the relation (1.3) can be written as follows: for any functions f,g in $\mathbb{L}^2([0,1],\nu)$,

$$\nu((f\circ T)\cdot g)=\nu((f\circ T)\cdot(Kg)\circ T)\,.$$

On the probability space $([0,1],\nu)$, this means precisely that $(Kg)\circ T=\mathbb{E}(g|T)$. Hence

$$(Kg)(x)=\mathbb{E}(g|T=x),$$

so that K is a transition kernel.

for instance Lemma XI.3 in Hennion and Hervé (2001)) that on the probability space $([0,1],\nu)$, the random vector $(T,T^2,\dots,T^n)$ is distributed as $(X_n,X_{n-1},\dots,X_1)$.[2] Hence, the process

$$L_n(t)=\frac{1}{n}\sum_{k=1}^{n}\Big(\mathbf{1}_{X_k\le t}-F(t)\Big),\quad t\in[0,1] \tag{1.4}$$

has the same distribution as $\{G_n(t),t\in[0,1]\}$.

In the short-range dependent case $\gamma\in]0,1/2[$, Dedecker (2010) proved that, on the probability space $([0,1],\nu_\gamma)$ the process $\{\sqrt{n}G_n(t),t\in[0,1]\}$ converges in distribution in the space $D([0,1])$ of cadlag functions equipped with the uniform metric to a centered Gaussian process G, whose sample paths are almost surely uniformly continuous. Moreover the covariance function of G is given by

$$\mathrm{Cov}(G(s),G(t))=\nu(f_t^{(0)}\cdot f_s^{(0)})+\sum_{k>0}\nu(f_t^{(0)}\cdot f_s^{(0)}\circ T^k)+\sum_{k>0}\nu(f_s^{(0)}\cdot f_t^{(0)}\circ T^k)\,, \tag{1.5}$$

where the function $f_t^{(0)}$ is defined by

$$f_t^{(0)}(x)=\mathbf{1}_{x\le t}-\nu([0,t])\,.$$

For $s=t$, the series (1.5) is the asymptotic variance of $\sqrt{n}G_n(t)$. This variance has the same structure as the asymptotic variance of the normalized partial sums of a stationary sequence $(Y_i)_{i\ge0}$ in the case where the covariance series converge, that is

$$\mathrm{Var}(Y_0)+2\sum_{k=1}^{\infty}\mathrm{Cov}(Y_0,Y_k)\,.$$

Observe that in this case, the limit process $\{G(t),t\in[0,1]\}$ is not degenerate.

In the long range-dependent case $\gamma\in[1/2,1[$, the series in (1.5) may not converge. The long-range dependent case has been studied by Dehling and Taqqu (1989) for the empirical process of a stationary Gaussian sequence. In that paper, the authors show that the empirical process, suitably normalized, converges in distribution in $D(\mathbb{R})$ to a degenerate Gaussian process. Following the approach of Dehling and Taqqu, Surgailis (2002) proved that the empirical process of a linear process whose innovations belong to the domain of normal attraction of a stable distribution, converges in distribution in $D(\mathbb{R})$ to a degenerate stable process. In the two papers cited above, the main idea is to approximate the empirical process by a sum of independent random variables whose asymptotic distribution is easy to derive. Such an approximation is not available in our context, and we shall use a completely different approach. For $\gamma\in[1/2,1[$, we shall obtain the same limit behavior as in Surgailis (2002), but in a space whose topology is much coarser than that of $D([0,1])$.

We shall investigate here the behavior of the empirical process

$$\{G_n(t),t\in[0,1]\}\quad\text{for } \gamma\in[1/2,1[,$$

in the Hilbert space

$$H=\mathbb{L}^2\big([0,1],dt\big)$$

with norm $\|\cdot\|_H$. The reason is that we can use very precise deviation inequalities for H-valued random variables to prove the tightness of the empirical process in H. An interesting question is whether our results remain true in $D([0,1])$, as in Dehling and Taqqu (1989) and Surgailis (2002).

Note that the empirical process, viewed as element of the Hilbert space H, is a centered and normalized partial sum of the random variables ξ_i, defined by

$$\xi_i(t)=1_{[0,t]}(X_i).$$

[2]For instance, by the Perron-Frobenius relation (1.3) and stationarity

$$\mathbb{E}(f(X_1)g(X_2))=\nu(f\cdot Kg)=\nu((f\circ T)\cdot g)=\nu(f(T^2)g(T)).$$

By setting $f(x)=\exp(itx)$ and $g(y)=\exp(iuy)$, we obtain that (X_1,X_2) is distributed as (T^2,T).

If the underlying random variables $(X_i)_{i\geq 1}$ are independent and indentically distributed (i.i.d.) or mixing, weak convergence of the empirical process is an immediate corollary of an appropriate central limit theorem for H-valued random variables. Such CLTs have been established in the i.i.d. case by Mourier (1953), for ϕ-mixing processes by Kuelbs and Philipp (1980) and for strongly mixing processes by Dehling (1983). The problem in our situation is that the process $(X_k)_{k\geq 1}$ is not mixing, so that none of the known CLTs for Hilbert space valued random variables can be applied. In this paper, we will present a proof that is tailor-made for the empirical process of intermittent maps.

Concerning the weak convergence of the empirical process $\{G_n(t), t\in[0,1]\}$ with $\gamma\in[1/2,1[$ in the space H, we shall prove the following theorem.

Theorem 1.1. *On the probability space $([0,1],\nu)$, the following results hold:*

1. *If $\gamma=1/2$, then the process*

$$\Big\{\frac{\sqrt{n}}{\sqrt{\log(n)}}G_n(t), t\in[0,1]\Big\}$$

converges in distribution in H to a degenerate Gaussian process $\{g(t)Z, t\in[0,1]\}$ where Z is a standard normal and $g(t)=\sqrt{h(1/2)}(1-F(t))$.

2. *If $\gamma\in]1/2,1[$, then the process*

$$\{n^{1-\gamma}G_n(t), t\in[0,1]\}$$

converges in distribution in H to a degenerate stable process $\{g(t)Z, t\in[0,1]\}$ where $g(t)=C_\gamma(h(1/2))^\gamma(1-F(t))$ with

$$C_\gamma=\frac{1}{4^\gamma\gamma}\Big(\Gamma(1-1/\gamma)\cos\Big(\frac{\pi}{2\gamma}\Big)\Big)^\gamma,$$

and Z is an $1/\gamma$-stable random variable totally skewed to the right, that is with characteristic function

$$\mathbb{E}(\exp(itZ))=\exp\Big(-|t|^{1/\gamma}\Big(1-\mathrm{sign}(t)\tan\Big(\frac{\pi}{2\gamma}\Big)\Big)\Big). \tag{1.6}$$

Remark 1.2. Recall that, on the probability space $([0,1],\nu)$ the process $\{G_n(t), t\in[0,1]\}$ is distributed as $\{L_n(t), t\in[0,1]\}$, defined in (1.4). Hence Theorem 1.1 is also valid for the empirical process $\{L_n(t), t\in[0,1]\}$.

Remark 1.3. Recall also that the distribution $S_\alpha(\sigma,\beta,\mu)$ of a stable random variable is characterized by the stability parameter $\alpha\in]0,2]$, the scale parameter $\sigma>0$, the skewness parameter $\beta\in[-1,1]$ and the shift parameter $\mu\in\mathbb{R}$ (see e.g. Samorodnitsky and Taqqu (1994)). When $\gamma\in]1/2,1[$, the random variable Z in (1.6) has a stability parameter $\alpha=1/\gamma\in]1,2[$ and hence has infinite variance and finite mean. Moreover $\sigma=1,\beta=1$ and $\mu=0$. Since $\beta=1$ it is said to be "totally skewed to the right".

Remark 1.4. The limits in the short-range and long-range dependent case are quite different. As noted after the relation (1.4), in the short-range dependent case, the limit is a non-degenerate Gaussian process, whereas in the long-range dependent case considered in Theorem 1.1 the limit is a degenerate process $\{g(t)Z, t\in[0,1]\}$, where Z can be Gaussian or not depending on the value of γ.

Concerning the almost sure behavior of $\|G_n\|_H$, we shall prove the following theorem.

Theorem 1.5. *The following results hold:*

1. *Let $\gamma=1/2$. Let a_n be any sequence of numbers such that $a_n\geq a$ for some $a>0$, and $\sum_{n>0}n^{-1}a_n^{-2}<\infty$. Then for any $\varepsilon>0$,*

$$\sum_{n=1}^{\infty}\frac{1}{n}\nu\Big(\max_{1\leq k\leq n}\frac{k\|G_k\|_H}{\sqrt{n\log(n+1)}a_n}\geq\varepsilon\Big)<\infty. \tag{1.7}$$

Assume moreover that a_n is non decreasing and such that $a_n \le c a_{n/2}$ for some $c \ge 1$. Then

$$\frac{\sqrt{n}}{a_n\sqrt{\log(n)}}\|G_n\|_H \quad \textit{converges almost everywhere to 0}. \tag{1.8}$$

2. *Let $\gamma \in]1/2,1[$. Let a_n be any sequence of positive numbers such that $\sum_{n>0} n^{-1}a_n^{-1/\gamma} < \infty$. Then for any $\varepsilon > 0$,*

$$\sum_{n=1}^{\infty}\frac{1}{n}\nu\Big(\max_{1\le k\le n}\frac{k\|G_k\|_H}{n^\gamma a_n} \ge \varepsilon\Big) < \infty. \tag{1.9}$$

Assume moreover that a_n is non decreasing and such that $a_n \le c a_{n/2}$ for some $c \ge 1$. Then

$$\frac{n^{1-\gamma}}{a_n}\|G_n\|_H \quad \textit{converges almost everywhere to 0}. \tag{1.10}$$

Remark 1.6. The corresponding almost sure result is true also for $\|L_n\|_H$, see the proof of Theorem 1.5.

Remark 1.7. For instance, all the conditions on a_n are satisfied if $a_n = (\log(n+1))^\delta$ for $\delta > \gamma$. For $\gamma \in]1/2,1[$ this is in accordance with the i.i.d. situation, which we now describe.

Let $(X_i)_{i\ge 1}$ be a sequence of i.i.d. centered random variables such that $n^{-\gamma}(X_1+\cdots+X_n)$ converges in distribution to the $1/\gamma$-stable distribution with characteristic function (1.6). It is well known (see for instance Feller (1966), page 547) that this implies that $x^{1/\gamma}\mathbb{P}(X_1 < -x) \to 0$ and $x^{1/\gamma}\mathbb{P}(X_1 > x) \to c > 0$ as $x \to \infty$. For any nondecreasing sequence $(b_n)_{n\ge1}$ of positive numbers, either $(X_1+\cdots+X_n)/b_n$ converges to zero almost surely or $\limsup_{n\to\infty}|X_1+\cdots+X_n|/b_n = \infty$ almost surely, according as $\sum_{n=1}^\infty \mathbb{P}(|X_1| > b_n) < \infty$ or $\sum_{n=1}^\infty \mathbb{P}(|X_1| > b_n) = \infty$ – this follows from the proof of Theorem 3 in Heyde (1969). If one takes $b_n = n^\gamma(\ln(n+1))^\delta$ we obtain the constraint $\delta > \gamma$ for the almost sure convergence of $n^{-\gamma}(\ln(n+1))^{-\delta}(X_1+\cdots+X_n)$ to zero. This is exactly the same constraint as in our dynamical situation.

This situation is similar to the one described in Theorem 1.7 of Dedecker, Gouëzel ans Merlevède (2010). Note that there is a mistake in Theorem 1.7 of this paper, in the case where $p = 1/2$ (weak moment of order 2): the exponent of the logarithm in (1.8) should satisfy $b > 1$ instead of $b > 1/2$.

Remark 1.8. In the short-range dependent case $\gamma \in]0,1/2[$, Dedecker, Merlevède and Rio (2013) have proved a strong approximation result for the empirical process $\{G_n(t), t \in [0,1]\}$. As a consequence, it follows that: almost everywhere, the sequence

$$\Big\{\frac{\sqrt{n}}{\sqrt{2\log\log(n)}}G_n(t), t\in[0,1]\Big\}$$

is relatively compact for the supremum norm, and the set of limit points is the unit ball of the reproducing kernel Hilbert space associated with the covariance function (1.5) of the limit Gaussian proces G. In particular, it follows that

$$\limsup_{n\to\infty}\frac{\sqrt{n}}{\sqrt{2\log\log(n)}}\|G_n\|_H = \sigma \quad \text{almost everywhere, where } \sigma = \sup_{\|g\|_H\le 1}\sqrt{\mathbb{E}(< g, G >^2)}\,.$$

To conclude this section, let us note that the conclusions of Theorems 1.1 and 1.5 also hold when T^k is replaced by $g(T^k)$ where g is a monotonic and Hölder continuous function from $[0,1]$ to $\mathbb{R}$. This fact will be proved in Theorem 5.1 of Section 5, and used in Section 3.

2 Application to the weak convergence of the Wasserstein distance W_1.

Let μ_n be the empirical measure of the iterates of T, that is

$$\mu_n = \sum_{k=1}^n \delta_{T^k}\,.$$

Consider the Wasserstein distance between the empirical measure μ_n and ν:

$$W_1(\mu_n,\nu) = \inf_{\pi \in M(\mu_n,\nu)} \int |x-y| \pi(dx,dy)\,,$$

where $M(\mu_n,\nu)$ is the set of probability measures on $[0,1]^2$ with marginals μ_n and ν. Since μ_n and ν are probability measures on the real line, it is well known that (see for instance Fréchet (1957))

$$W_1(\mu_n,\nu) = \int_0^1 |G_n(t)|dt\,, \tag{2.1}$$

where G_n is defined in (1.2). Since the functional $\psi(f) = \int_0^1 |f(t)|dt$ is continuous on the Hilbert space H, we can apply Theorem 1.1. Since

$$\int_0^1 (1-F(t))dt = \int_0^1 x\nu(dx) = \int_0^1 xh(x)dx\,,$$

we obtain

Corollary 2.1. *On the probability space $([0,1],\nu)$, the following results hold:*

1. *If $\gamma = 1/2$,*

$$\frac{\sqrt{n}}{\sqrt{\log(n)}} W_1(\mu_n,\nu) \quad \textit{converges in distribution to} \quad \sqrt{h(1/2)}|Z| \int_0^1 xh(x)dx\,.$$

   where Z is a standard normal.

2. *If $\gamma \in]1/2,1[$,*

$$n^{1-\gamma} W_1(\mu_n,\nu) \quad \textit{converges in distribution to} \quad C_\gamma (h(1/2))^\gamma |Z| \int_0^1 xh(x)dx\,.$$

   where Z is an $1/\gamma$-stable random variable with characteristic function (1.6).

Since by (2.1), $W_1(\mu_n,\nu) \le \|G_n\|_H$, we can apply Theorem 1.5.

Corollary 2.2. *The following results hold:*

1. *Let $\gamma = 1/2$ and let a_n be as in Item 1 of Theorem 1.5. Then*

$$\frac{\sqrt{n}}{a_n\sqrt{\log(n)}} W_1(\mu_n,\nu) \quad \textit{converges to zero almost everywhere.}$$

2. *Let $\gamma \in]1/2,1[$, and let a_n be as in Item 2 of Theorem 1.5. Then*

$$\frac{n^{1-\gamma}}{a_n} W_1(\mu_n,\nu) \quad \textit{converges to zero almost everywhere.}$$

3 Application to the Cramér-von-Mises statistic

Recall that F is the cumulative distribution function (cdf) of the absolutely continuous invariant measure ν. In order to test whether the cdf of ν is equal to the cdf G, we use the test statistic

$$\Psi_n = \int_0^1 (F_n(t) - G(t))^2 dG(t), \quad \text{where } F_n(t) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}_{T^i \le t}\,.$$

Note that, by Birkoff's ergodic Theorem combined with the Glivenko-Cantelli arguments,

$$\sup_{t\in[0,1]} |F_n(t) - F(t)| \quad \text{converges almost everywhere to 0.}$$

It follows that Ψ_n converges almost everywhere to 0 if $G = F$ and to a strictly positive number if G is absolutely continuous and $G \neq F$.

We can apply Proposition 5.1 of the Appendix (which is a generalization of Theorem 1.1) to derive the asymptotic distribution of Ψ_n on the probability space $([0,1],\nu)$, under the null hypothesis $H : F = G$.

Corollary 3.1. *On the probability space $([0,1],\nu)$, under the null hypothesis $H : F = G$, the following results hold:*

1. *If $\gamma = 1/2$,*
$$\frac{n}{\log n}\Psi_n \quad \textit{converges in distribution to} \quad \frac{1}{3}h(1/2)Z^2,$$
where Z is a standard normal random variable.
2. *If $\gamma \in]1/2,1[$,*
$$n^{2-2\gamma}\Psi_n \quad \textit{converges in distribution to} \quad \frac{C_\gamma^2(h(1/2))^{2\gamma}}{3}Z^2,$$
where Z is a $1/\gamma$-stable random variable with characteristic function (1.6).

Proof. Assume that $F = G$, and note that F is continuous and strictly increasing. By a change of variables, we obtain

$$\Psi_n = \int_0^1 (G_n(t))^2 \nu(dt) = \int_0^1 (G_n(F^{-1}(t)))^2 dt.$$

Since $F^{-1}(F(t)) = t$,

$$G_n(F^{-1}(t)) = \frac{1}{n}\sum_{i=1}^n \Big(\mathbf{1}_{T^i \le F^{-1}(t)} - t\Big) = \frac{1}{n}\sum_{i=1}^n \Big(\mathbf{1}_{F(T^i)\le t} - t\Big),$$

i.e. $G_n(F^{-1}(t))$ is the empirical process of the sequence $(F(T), F(T^2), \dots, F(T^n))$.

In addition of being increasing, the function F is also Hölder continuous. Indeed, let $\delta \in]0, 1-\gamma[$ and $p = 1/(1-\delta)$. By Hölder's inequality, for any $x, y \in [0,1]$ with $x \le y$,

$$F(y) - F(x) \le \Big(\int_x^y h(x)^p dx\Big)^{1/p} (y-x)^\delta \, .$$

Note that $\int_0^1 h(x)^p dx$ is finite, because $x^\gamma h(x)$ is bounded from above and $p\gamma < 1$. It follows that there exists a positive constant K such that

$$F(y) - F(x) \le K(y-x)^\delta \, ,$$

showing that F is Hölder continuous of index δ.

Hence one can apply Theorem 5.1 of Section 5: on the probability space $([0,1],\nu)$,

1. If $\gamma = 1/2$, then the process
$$\Big\{\frac{\sqrt{n}}{\sqrt{\log(n)}} G_n(F^{-1}(t)), t \in [0,1]\Big\}$$
converges in distribution in H to a degenerate Gaussian process $\{g(t)Z, t \in [0,1]\}$ where Z is a standard normal and $g(t) = \sqrt{h(1/2)}(1-t)$.
2. If $\gamma \in]1/2, 1[$, then the process
$$\{n^{1-\gamma} G_n(F^{-1}(t)), t \in [0,1]\}$$
converges in distribution in H to a degenerate stable process $\{g(t)Z, t \in [0,1]\}$ where $g(t) = C_\gamma (h(1/2))^\gamma (1-t)$ and Z is a $1/\gamma$-stable random variable with characteristic function (1.6).

Since the map $f \mapsto \|f\|_H^2 = \int_0^1 f^2(t)dt$ is continuous on H, we obtain in case $\gamma = 1/2$ that

$$\frac{n}{\log n}\Psi_n = \left\| \frac{\sqrt{n}}{\sqrt{\log n}} G_n \circ F^{-1} \right\|_H^2 \text{ converges in distribution to } h(1/2)Z^2 \int_0^1 (1-t)^2 dt = \frac{1}{3}h(1/2)Z^2.$$

The case $1/2 < \gamma < 1$ follows in the same way.

4 Proof of Theorems 1.1 and 1.5

In this section, C is a positive constant which may vary from line to line.

4.1 Some general facts

Let

$$Y_i(t) = \mathbf{1}_{X_i \le t} - F(t)\,, \tag{4.1}$$

and $S_n = \sum_{i=1}^n Y_i$. With these notations, by (1.4),

$$S_n(t) = \sum_{i=1}^n (\mathbf{1}_{X_i \le t} - F(t)) = nL_n(t)\,. \tag{4.2}$$

Let also $V_i(t) = \mathbf{1}_{T^i \le t} - F(t)$, and $\Sigma_n = \sum_{i=1}^n V_i$. With these notations, by (1.2),

$$\Sigma_n(t) = \sum_{i=1}^n (\mathbf{1}_{T^i \le t} - F(t)) = nG_n(t)\,. \tag{4.3}$$

Recall that, on the probability space $([0,1],\nu)$ the sequence $(V_1, V_2, \dots, V_n)$ is distributed as the sequence $(Y_n, Y_{n-1}, \dots, Y_1)$. It follows that L_n is distributed as G_n, and it is equivalent to prove Theorem 1.1 for L_n or for G_n.

1. Let us first prove the following inequality: for any $x \ge 0$,

$$\nu\Big(\max_{1\le k\le n} \|\Sigma_k\|_H \ge x\Big) \le \mathbb{P}\Big(2 \max_{1\le k\le n} \|S_k\|_H \ge x\Big)\,. \tag{4.4}$$

   Indeed,

$$\max_{1\le k\le n} \Big\|\sum_{i=1}^k V_i\Big\|_H \quad \text{is distributed as} \quad \max_{1\le k\le n} \Big\|\sum_{i=k}^n X_i\Big\|_H\,. \tag{4.5}$$

   Notice now that for any $k \in \{1, \dots, n\}$,

$$\sum_{i=k}^n X_i = \sum_{i=1}^n X_i - \sum_{i=1}^{k-1} X_i\,.$$

   Consequently

$$\max_{1\le k\le n} \Big\|\sum_{i=k}^n X_i\Big\|_H \le \max_{1\le k\le n-1} \Big\|\sum_{i=1}^k X_i\Big\|_H + \Big\|\sum_{i=1}^n X_i\Big\|_H \le 2 \max_{1\le k\le n} \Big\|\sum_{i=1}^k X_i\Big\|_H$$

   which together with (4.5) entails (4.4).

2. Let H be a separable Hilbert space with inner product $\langle x, y\rangle$ and norm $\|x\|_H^2 = \langle x, x\rangle$, and let $(u_i)_{i\ge1}$ be a complete orthonormal system in H. Thus, any vector $x \in H$ can be expanded into a series

$$x = \sum_{i=1}^\infty \langle x, u_i\rangle u_i.$$

   For any integer m, we define the finite-dimensional projection $P_m : H \to H$ by

$$P_m(x) = \sum_{i=1}^m \langle x, u_i\rangle u_i.$$

   Let now $(Z_n)_{n\ge1}$ be a sequence of H-valued random variables, and let Z be another H-valued random variable, satisfying

(i) For all integers m, as $n \to \infty$,

$$P_m(Z_n) \text{ converges in distribution in } H \text{ to } P_m(Z) \tag{4.6}$$

(ii) For any $\epsilon > 0$,

$$\lim_{m\to\infty} \limsup_{n\to\infty} P(\|Z_n - P_m(Z_n)\|_H \geq \epsilon) = 0. \tag{4.7}$$

Then, by Theorem 4.2 of Billingsley (1968), Z_n converges in distribution to Z, as $n \to \infty$. Observe that weak convergence in (4.6) is really weak convergence in a finite dimensional Euclidean space, since the map $a = (a_1, \ldots, a_m) \mapsto \sum_{i=1}^m a_i u_i$ defines an isometry between $\mathbb{R}^m$ and the subspace $P_m(H) \subset H$. Thus (4.6) holds if and only if

$$(\langle Z_n, u_1\rangle, \ldots, \langle Z_n, u_m\rangle) \text{ converges in distribution to } (\langle Z, u_1\rangle, \ldots, \langle Z, u_m\rangle)\,,$$

as $n \to \infty$. Hence, using the Cramér-Wold device, (4.6) is equivalent to

$$\langle Z_n, \sum_{i=1}^m \lambda_i u_i\rangle = \sum_{i=1}^m \lambda_i \langle Z_n, u_i\rangle \text{ converges in distribution to } \sum_{i=1}^m \lambda_i\langle Z, u_i\rangle = \langle Z, \sum_{i=1}^m \lambda_i u_i\rangle, \tag{4.8}$$

as $n \to \infty$, for all $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$.

3. On the Hilbert space $H = \mathbb{L}^2([0,1], dt)$, consider the usual orthonormal Fourier basis

$$(\mathbf{1}, (\sqrt{2}\cos(2\pi k\cdot), \sqrt{2}\sin(2\pi k\cdot))_{k\in\mathbb{N}^*})\,,$$

and let P_m be the projection operator on the space spanned by

$$(\mathbf{1}, (\sqrt{2}\cos(2\pi k\cdot), \sqrt{2}\sin(2\pi k\cdot))_{k\in\{1,\ldots,m\}})\,.$$

Let $\|\cdot\|_\infty$ denotes the essential supremum norm, namely

$$\|X\|_\infty = \inf\{M > 0 : \mathbb{P}(|X| > M) = 0\}\,.$$

Thus while $\|Y_0 - P_m(Y_0)\|_H$ is random,

$$M(m) = \|\|Y_0 - P_m(Y_0)\|_H\|_\infty \tag{4.9}$$

is not random. In fact, we have

$$M(m) = O(m^{-1/2})\,. \tag{4.10}$$

Indeed, the Fourier coefficients

$$a_k(x) = \sqrt{2}\int_x^1 \cos(2\pi kt)dt \quad \text{and} \quad b_k(x) = \sqrt{2}\int_x^1 \sin(2\pi kt)dt$$

of the functions $f_x(t) = \mathbf{1}_{t\geq x}$ are such that

$$a_k(x) + ib_k(x) = \sqrt{2}\int_x^1 \exp(i2\pi kt)dt = \frac{1}{i\sqrt{2}\pi k}(1 - \exp(i2\pi kx))\,,$$

in such a way that $a_k^2(x) + b_k^2(x) \leq 2/(\pi k)^2$. Hence,

$$\|Y_0 - P_m(Y_0)\|_H^2 = \sum_{k=m+1}^\infty (a_k^2(X_0) + b_k^2(X_0)) \leq \sum_{k=m+1}^\infty \frac{2}{\pi^2 k^2} \leq \frac{C}{m}\,, \tag{4.11}$$

proving (4.10).

4. Let Y_k be defined as in (4.1). Then, if $\mathbb{E}_0$ is the conditional expectation with respect to X_0 and $F_{X_k|X_0}$ is the conditional distribution function of X_k given X_0,

$$\begin{aligned}\mathbb{E}(\|\mathbb{E}_0(Y_k)\|_H) &= \mathbb{E}\Big(\Big(\int_0^1 (F_{X_k|X_0}(t)-F(t))^2 dt\Big)^{1/2}\Big)\\ &\leq \mathbb{E}\Big(\sup_{t\in[0,1]} |F_{X_k|X_0}(t)-F(t)|\Big) := \beta(k). \qquad (4.12)\end{aligned}$$

Here $\beta(k)$ is the weak β-mixing coefficient of the chain $(X_i)_{i\geq 0}$. Starting from the computations of the paper by Dedecker, Gouëzel and Merlevède (2010), we shall prove in the appendix that

$$\text{for any } \gamma \in]0,1[, \quad \beta(k) \leq \frac{C}{(k+1)^{(1-\gamma)/\gamma}}\,. \qquad (4.13)$$

4.2 Proof of Theorem 1.1 for $\gamma = 1/2$.

Let S_n be defined as in (4.2). We shall prove that $Z_n = S_n/\sqrt{n\log(n)}$ satisfies the points (i) and (ii) of Item 2 of Section 4.1.

We first prove (i): for any positive integer m, $P_m(S_n/\sqrt{n\log(n)})$ converges in distribution in H to $P_m(V)$, where $V=\{g(t)Z, t\in[0,1]\}$ is the process described in Item 1 of Theorem 1.1. For any $\mathbf{a}=(a_0,a_1,\ldots,a_m)$ in $\mathbb{R}^{m+1}$ and any $\mathbf{b}=(b_1,\ldots,b_m)$ in $\mathbb{R}^m$, let

$$f_{\mathbf{a},\mathbf{b}}(t) = a_0 + \sum_{k=1}^m a_k\cos(2\pi k t) + \sum_{k=1}^m b_k \sin(2\pi k t)\,. \qquad (4.14)$$

As noted in Section 4.1, this is equivalent to prove that

$$\frac{1}{\sqrt{n\log(n)}} < f_{\mathbf{a},\mathbf{b}}, S_n > \text{ converges in distribution to } \sqrt{h(1/2)}\Big(\int_0^1 f_{\mathbf{a},\mathbf{b}}(t)(1-F(t))dt\Big)Z\,. \qquad (4.15)$$

Defining the function $u_{\mathbf{a},\mathbf{b}}$ by

$$u_{\mathbf{a},\mathbf{b}}(x) = \int_x^1 f_{\mathbf{a},\mathbf{b}}(t)dt - \int_0^1 f_{\mathbf{a},\mathbf{b}}(t)F(t)dt\,, \qquad (4.16)$$

we obtain that

$$< f_{\mathbf{a},\mathbf{b}}, S_n > = \sum_{k=1}^n u_{\mathbf{a},\mathbf{b}}(X_k)\,.$$

Note that the function $u_{\mathbf{a},\mathbf{b}}$ is Lipschitz and that $\mathbb{E}(u_{\mathbf{a},\mathbf{b}}(X_k))=0$. Hence, it follows from Gouëzel (2004) that

$$\frac{1}{\sqrt{n\log(n)}}\sum_{k=1}^n u_{\mathbf{a},\mathbf{b}}(X_k) \text{ converges in distribution to } \sqrt{h(1/2)}u_{\mathbf{a},\mathbf{b}}(0)Z\,.$$

Since $u_{\mathbf{a},\mathbf{b}}(0)=\int_0^1 f_{\mathbf{a},\mathbf{b}}(t)(1-F(t))dt$, (4.15) holds and hence point (i) is proved.

We now prove (ii): for any $\varepsilon>0$,

$$\lim_{m\to\infty}\limsup_{n\to\infty}\mathbb{P}\Big(\frac{\|S_n-P_m(S_n)\|_H^2}{\sqrt{n\log(n)}} > \varepsilon\Big) = 0\,.$$

Hence, (ii) follows from

$$\lim_{m\to\infty}\limsup_{n\to\infty}\frac{\mathbb{E}(\|S_n-P_m(S_n)\|_H^2)}{n\log(n)} = 0\,.$$

By stationarity

$$\mathbb{E}(\|S_n - P_m(S_n)\|_H^2) = n\mathbb{E}(\|Y_0 - P_m(Y_0)\|_H^2) + 2\sum_{k=1}^{n-1}(n-k)\mathbb{E}(< Y_0 - P_m(Y_0), Y_k - P_m(Y_k) >)$$
$$\leq n\Big(\mathbb{E}(< Y_0 - P_m(Y_0), Y_0 >) + 2\sum_{k=1}^{n-1}|\mathbb{E}(< Y_0 - P_m(Y_0), Y_k >)|\Big)\,, \quad (4.17)$$

since $Y_0 - P_m(Y_0)$ is orthogonal to $P_m(Y_k)$ for any $k = 0, \ldots, n$.

Taking the conditional expectation with respect to X_0, it follows that

$$|\mathbb{E}(< Y_0 - P_m(Y_0), Y_k >)| \leq |\mathbb{E}(< Y_0 - P_m(Y_0), \mathbb{E}_0(Y_k) >)| \leq \|\|Y_0 - P_m(Y_0)\|_H\|_\infty \mathbb{E}(\|\mathbb{E}_0(Y_k)\|_H)\,.$$

Therefore (4.17) yields

$$\mathbb{E}(\|S_n - P_m(S_n)\|_H^2) \leq n\|\|Y_0 - P_m(Y_0)\|_H\|_\infty\Big(\mathbb{E}(\|Y_0\|_H) + 2\sum_{k=1}^{n-1}\mathbb{E}(\|\mathbb{E}_0(Y_k)\|_H)\Big)$$
$$\leq 2n\|\|Y_0 - P_m(Y_0)\|_H\|_\infty \sum_{k=0}^{n-1}\mathbb{E}(\|\mathbb{E}_0(Y_k)\|_H)\,. \quad (4.18)$$

By (4.10), $\|\|Y_0 - P_m(Y_0)\|_H\|_\infty = O(m^{-1/2})$. To evaluate the sum, we use the inequalities (4.12) and (4.13). Since $\gamma = 1/2$, it follows that $\beta(k) = O(k^{-1})$, so that

$$\sum_{k=1}^{n-1}\mathbb{E}(\|\mathbb{E}_0(Y_k)\|_H) \leq \sum_{k=1}^{n-1}\beta(k) \leq C\log(n)\,. \quad (4.19)$$

Consequently

$$\lim_{m\to\infty}\limsup_{n\to\infty}\frac{\mathbb{E}(\|S_n - P_m(S_n)\|_H^2)}{n\log(n)} \leq \lim_{m\to\infty}\limsup_{n\to\infty}\frac{Cn\log(n)}{\sqrt{m}n\log(n)} = 0\,,$$

and (ii) follows.

4.3 Proof of Theorem 1.1 for $\gamma \in]1/2, 1[$.

Let S_n be defined as in (4.2). We shall prove that $Z_n = S_n/n^\gamma$ satisfies the points (i) and (ii) of Item 2 of Section 4.1.

We first prove (i): for any positive integer m, $P_m(S_n/n^\gamma)$ converges in distribution in H to $P_m(V)$, where $V = \{g(t)Z, t \in [0,1]\}$ is the process described in Item 2 of Theorem 1.1. For any $\mathbf{a} = (a_0, a_1, \ldots, a_m)$ in $\mathbb{R}^{m+1}$ and any $\mathbf{b} = (b_1, \ldots, b_m)$ in $\mathbb{R}^m$, define the functions $f_{\mathbf{a},\mathbf{b}}$ and $u_{\mathbf{a},\mathbf{b}}$ as in (4.14) and (4.16) respectively. As in Section 4.2, it suffices to prove that

$$\frac{1}{n^\gamma}\sum_{k=1}^{n}u_{\mathbf{a},\mathbf{b}}(X_k) \text{ converges in distribution to } C_\gamma(h(1/2))^\gamma\Big(\int_0^1 f_{\mathbf{a},\mathbf{b}}(t)(1-F(t))dt\Big)Z\,.$$

where C_γ and Z are described in Item 2 of Theorem 1.1. Note that the function $u_{\mathbf{a},\mathbf{b}}$ is Lipschitz and that $\mathbb{E}(u_{\mathbf{a},\mathbf{b}}(X_k)) = 0$. Hence, it follows from Theorem 1.3 in Gouëzel (2004) that

$$\frac{1}{n^\gamma}\sum_{k=1}^{n}u_{\mathbf{a},\mathbf{b}}(X_k) \text{ converges in distribution to } C_\gamma(h(1/2))^\gamma u_{\mathbf{a},\mathbf{b}}(0)Z\,.$$

Since $u_{\mathbf{a},\mathbf{b}}(0) = \int_0^1 f_{\mathbf{a},\mathbf{b}}(t)(1-F(t))dt$, the point (i) follows.

We now prove (ii): for any $\varepsilon > 0$,

$$\lim_{m\to\infty}\limsup_{n\to\infty}\mathbb{P}\Big(\frac{\|S_n - P_m(S_n)\|_H^2}{n^\gamma} > \varepsilon\Big) = 0\,.$$

We shall apply Proposition 6.1 of the appendix to the random variables $Y_i - P_m(Y_i)$ and the σ-algebras $\mathcal{F}_i = \sigma(X_1, \ldots, X_i)$. For any $j \geq i + k$, since $\mathcal{F}_i \subset \mathcal{F}_{j-k}$, one has

$$\mathbb{E}(\|\mathbb{E}(Y_j - P_m(Y_j)|\mathcal{F}_i)\|_H) \leq \mathbb{E}(\|\mathbb{E}(Y_j - P_m(Y_j)|\mathcal{F}_{j-k})\|_H).$$

Combined with the Markov property, this implies that the coefficient $\theta(k)$ defined in Proposition 6.1 of the appendix is such that: for $k \in \{0, \ldots, n-1\}$,

$$\begin{aligned}\theta(k) &= \max\Big\{\mathbb{E}(\|\mathbb{E}(Y_j - P_m(Y_j)|\mathcal{F}_i)\|_H), (i,j) \in \{1, \ldots, n\}^2 \text{ such that } j \geq i + k\Big\} \\ &= \max\Big\{\mathbb{E}(\|\mathbb{E}(Y_j - P_m(Y_j)|\mathcal{F}_{j-k})\|_H), j \in \{k+1, \ldots, n\}\Big\} \\ &= \max\Big\{\mathbb{E}(\|\mathbb{E}(Y_j - P_m(Y_j)|X_{j-k})\|_H), j \in \{k+1, \ldots, n\}\Big\}.\end{aligned}$$

Let $\mathbb{E}_0$ be the conditional expectation with respect to X_0. By stationarity it follows that

$$\theta(k) = \mathbb{E}(\|\mathbb{E}_0(Y_k - P_m(Y_k))\|_H) \leq \mathbb{E}(\|\mathbb{E}_0(Y_k)\|_H). \tag{4.20}$$

the last inequality being satisfied because $\|\mathbb{E}_0(Y_k)\|_H^2 = \|\mathbb{E}_0(Y_k - P_m(Y_k))\|_H^2 + \|\mathbb{E}_0(P_m(Y_k))\|_H^2$ by orthogonality. By (4.13) it follows that

$$\theta(k) \leq \|\mathbb{E}_0(Y_k)\|_H \leq \beta(k).$$

Let $M(m) = \|\|Y_0 - P_m(Y_0)\|_H\|_\infty$ as in (4.9). Applying Proposition 6.1 of the appendix, for any positive integer q and $x \geq qM(m)$, one has

$$\mathbb{P}\Big(\max_{1 \leq k \leq n} \|S_k - P_m(S_k)\|_H \geq 4x\Big) \leq \frac{n\beta(q)}{x} + \frac{2nM(m)}{x^2} \sum_{k=0}^{q-1} \beta(q). \tag{4.21}$$

By (4.13), we know that $\beta(k) \leq C(k+1)^{(\gamma-1)/\gamma}$. Hence, it follows from (4.21) that, for $x \geq qM(m)$ and $\gamma \in]1/2, 1[$,

$$\mathbb{P}\Big(\max_{1 \leq k \leq n} \|S_k - P_m(S_k)\|_H \geq 4x\Big) \leq C\Big(\frac{n}{xq^{(1-\gamma)/\gamma}} + \frac{q^{(2\gamma-1)/\gamma} nM(m)}{x^2}\Big).$$

Taking $q = [x/M(m)]$ when $x \geq M(m)$, we finally obtain that

$$\mathbb{P}\Big(\max_{1 \leq k \leq n} \|S_k - P_m(S_k)\|_H \geq 4x\Big) \leq \frac{CnM(m)^{(1-\gamma)/\gamma}}{x^{1/\gamma}} \mathbf{1}_{x \geq M(m)} + \mathbf{1}_{x < M(m)} \tag{4.22}$$

(we bound this probability by 1 when $x < M(m)$). We now apply (4.22) with $x = n^\gamma \varepsilon/4$. In view of the definition (4.9) of $M(m)$, it follows that, for n large enough,

$$\mathbb{P}\Big(\frac{\|S_n - P_m(S_n)\|_H^2)}{n^\gamma} > \varepsilon\Big) \leq \frac{C}{\varepsilon^{1/\gamma}} \Big(\|\|Y_0 - P_m(Y_0)\|_H\|_\infty\Big)^{(1-\gamma)/\gamma}.$$

In view of (4.11),

$$\lim_{m \to \infty} \limsup_{n \to \infty} \mathbb{P}\Big(\frac{\|S_n - P_m(S_n)\|_H^2)}{n^\gamma} > \varepsilon\Big) \leq \frac{C}{\varepsilon^{1/\gamma}} \lim_{m \to \infty} \Big(\frac{1}{m}\Big)^{(1-\gamma)/\gamma} = 0,$$

since $\gamma < 1$. The point (ii) follows.

4.4 Proof of Theorem 1.5.

Applying Inequality (4.4), we get that, for any $x > 0$,

$$\nu\Big(\max_{1\le k\le n} k\|G_k\|_H \ge x\Big) \le \mathbb{P}\Big(2\max_{1\le k\le n}\|S_k\|_H \ge x\Big).$$

Hence, the inequalities (1.7) and (1.9) hold provided the same inequalities hold for $\|S_k\|_H$ instead of $k\|G_k\|_H$.

Let $\gamma = 1/2$. Replacing $Y_0 - P_m(Y_0)$ by Y_0, we note that the inequality (4.21) is valid for S_n with $M = \|\|Y_0\|_H\|_\infty$. By (4.13), we know that $\beta(q) \le C(q+1)^{-1}$. Hence, it follows from (4.21) that, for $x \ge qM$,

$$\mathbb{P}\Big(\max_{1\le k\le n}\|S_k\|_H \ge 4x\Big) \le C\Big(\frac{n}{xq} + \frac{\log(q)nM}{x^2}\Big).$$

Taking $q = [\sqrt{n}(\log(n))^\alpha]$ for some $\alpha \in]0,1/2[$, we finally obtain that

$$\mathbb{P}\Big(\max_{1\le k\le n}\|S_k\|_H \ge 4x\Big) \le C\Big(\frac{\sqrt{n}}{x(\log(n))^\alpha} + \frac{\log(n)nM}{x^2}\Big)\mathbf{1}_{x\ge M[\sqrt{n}(\log(n))^\alpha]} + \mathbf{1}_{x<M[\sqrt{n}(\log(n))^\alpha]}. \tag{4.23}$$

Let a_n be any sequence of numbers such that $a_n \ge a$ for some $a > 0$, and $\sum_{n>0} n^{-1}a_n^{-2} < \infty$. Taking $4x = \varepsilon\sqrt{n\log(n)}a_n$ in (4.23), we get that, for n large enough

$$\mathbb{P}\Big(\max_{1\le k\le n}\frac{\|S_k\|_H}{\sqrt{n\log(n)}a_n} \ge \varepsilon\Big) \le C\Big(\frac{1}{\varepsilon(\log(n))^{(1+2\alpha)/2}a_n} + \frac{M}{\varepsilon^2 a_n^2}\Big).$$

Since, by Cauchy-Schwarz,

$$\sum_{n=2}^\infty \frac{1}{n}\frac{1}{(\log(n))^{(1+2\alpha)/2}a_n} \le \Big(\sum_{n=2}^\infty \frac{1}{n(\log(n))^{(1+2\alpha)}}\Big)^{1/2}\Big(\sum_{n=2}^\infty \frac{1}{na_n^2}\Big)^{1/2} < \infty,$$

we infer that

$$\sum_{n=2}^\infty \frac{1}{n}\mathbb{P}\Big(\max_{1\le k\le n}\frac{\|S_k\|_H}{\sqrt{n\log(n)}a_n} \ge \varepsilon\Big) < \infty,$$

and (1.7) follows. Assume moreover that a_n is non decreasing and such that $a_n \le ca_{n/2}$ for some $c \ge 1$. Let $N \ge 2$ be a positive integer, and let $n \in \{2^N+1, \dots, 2^{N+1}\}$. Clearly

$$\max_{1\le k\le 2^N}\frac{k\|G_k\|_H}{\sqrt{2^N\log(2^N)}a_{2^N}} \le 2c\max_{1\le k\le n}\frac{k\|G_k\|_H}{\sqrt{n\log(n)}a_n} \le 4c^2\max_{1\le k\le 2^{N+1}}\frac{k\|G_k\|_H}{\sqrt{2^{(N+1)}\log(2^{N+1})}a_{2^{N+1}}}. \tag{4.24}$$

Using the first inequality of (4.24), it follows from (1.7) that

$$\begin{aligned}\sum_{N=2}^\infty \nu\Big(\max_{1\le k\le 2^N}\frac{k\|G_k\|_H}{\sqrt{2^N\log(2^N)}a_{2^N}} \ge \varepsilon\Big) &\le \sum_{N=2}^\infty\sum_{n=2^N+1}^{2^{N+1}}\frac{2}{n}\nu\Big(\max_{1\le k\le n}\frac{k\|G_k\|_H}{\sqrt{n\log(n)}a_n} \ge \frac{\varepsilon}{2c}\Big)\\ &\le \sum_{n=5}^\infty \frac{2}{n}\nu\Big(\max_{1\le k\le n}\frac{k\|G_k\|_H}{\sqrt{n\log(n)}a_n} \ge \frac{\varepsilon}{2c}\Big) < \infty.\end{aligned}$$

By the direct part of the Borel-Cantelli Lemma, and bearing in mind that ν is equivalent to the Lebesgue measure, we infer that

$$\lim_{N\to\infty}\max_{1\le k\le 2^N}\frac{k\|G_k\|_H}{\sqrt{2^N\log(2^N)}a_{2^N}} = 0 \quad \text{almost everywhere}. \tag{4.25}$$

Using (4.25) and the second inequality of (4.24), we conclude that

$$\frac{\sqrt{n}}{\sqrt{\log(n)}a_n}\|G_n\|_H \quad \text{converges almost everywhere to } 0,$$

proving (1.8).

Let $\gamma \in]1/2, 1[$. The inequality (4.22) is valid for S_n with $M = \| \|Y_0\|_H \|_\infty$ and gives

$$\mathbb{P}\Big(\max_{1\le k\le n} \|S_k\|_H \ge 4x \Big) \le \frac{CnM^{(1-\gamma)/\gamma}}{x^{1/\gamma}} \mathbf{1}_{x\ge M} + \mathbf{1}_{x<M} \,. \tag{4.26}$$

Let a_n be any sequence of positive numbers such that $\sum_{n>0} n^{-1} a_n^{-1/\gamma} < \infty$. Taking $x = 4\varepsilon n^\gamma a_n$ in (4.26), we get that

$$\sum_{n=1}^{\infty} \frac{1}{n} \mathbb{P}\Big(\max_{1\le k\le n} \frac{\|S_k\|_H}{n^\gamma a_n} \ge \varepsilon \Big) < \infty \,,$$

and (1.9) follows. Assume moreover that a_n is non decreasing and such that $a_n \le c a_{n/2}$ for some $c \ge 1$. Using the same arguments as for the case $\gamma = 1/2$, we infer from (1.9) that

$$\sum_{N=1}^{\infty} \nu\Big(\max_{1\le k\le 2^N} \frac{k\|G_k\|_H}{2^{N\gamma} a_{2^N}} \ge \varepsilon \Big) < \infty \,.$$

and we conclude that

$$\frac{n^{1-\gamma}}{a_n} \|G_n\|_H \quad \text{converges almost everywhere to } 0 \,,$$

proving (1.10).

5 Extensions of the main results to some functions of T^k.

Let g be a function from $[0,1]$ to a compact interval $[a,b]$. In this subsection, we modify the notations of Section 1 as follows: we denote now by G_n the empirical process of the sequence $(g(T), g(T^2), \ldots, g(T^n))$, that is

$$G_n(t) = \frac{1}{n} \sum_{k=1}^{n} \Big(\mathbf{1}_{g(T^k)\le t} - F(t) \Big), \quad t \in \mathbb{R} \tag{5.1}$$

where now $F(t) = \nu(\mathbf{1}_{g\le t})$. Let also $H = \mathbb{L}^2([a,b], dt)$.

Theorem 5.1. *Let g be a monotonic and Hölder continuous function from $[0,1]$ to $\mathbb{R}$. Then the conclusions of Theorems 1.1 and 1.5 apply to the process defined by (5.1).*

Proof of Theorem 5.1. Without loss of generality, assume that g is a function from $[0,1]$ to $[0,1]$, so that $H = \mathbb{L}^2([0,1], dt)$. The proof of this proposition is almost the same as that of Theorems 1.1 and 1.5. Let us check the main points.

Let S_n be the H-valued random variable defined by

$$S_n(t) = \sum_{i=1}^{n} (\mathbf{1}_{g(X_i)\le t} - F(t)) \,,$$

where $(X_k)_{k\ge 0}$ is a stationary Markov chain with invariant measure ν and transition kernel K defined in (1.3).

Let us start with the convergence in distribution.

The finite dimensional convergence (Point (i) of Item 2 of Section 4.1) can be proved as at the beginning of Sections 4.2 and 4.3. If $f_{\mathbf{a},\mathbf{b}}$ and $u_{\mathbf{a},\mathbf{b}}$ are defined by (4.14) and (4.16) respectively, we obtain that

$$< f_{\mathbf{a},\mathbf{b}}, S_n >= \sum_{k=1}^{n} u_{\mathbf{a},\mathbf{b}}(g(X_k)) \,.$$

As already noticed, the function $u_{\mathbf{a},\mathbf{b}}$ is Lipschitz, and consequently the function $u_{\mathbf{a},\mathbf{b}} \circ g$ is Hölder continuous. The finite dimensional convergence follows as in Sections 4.2 and 4.3, since Gouëzel's results (2004) apply to any Hölder function.

The tightness (Point (ii) of Item 2 of Section 4.1) can be proved exactly as in Sections 4.2 and 4.3 provided that (4.10) holds for $Y_0 = \mathbf{1}_{g(X_0)\leq t} - F(t)$, and provided that the new coefficient

$$\beta(k) = \mathbb{E}\Big(\sup_{t\in[0,1]} |F_{g(X_k)|X_0}(t) - F(t)|\Big) \tag{5.2}$$

satisfies (4.13). The first point can be proved as in (4.11): for some positive constant C,

$$\|Y_0 - P_m(Y_0)\|_H^2 = \sum_{k=m+1}^{\infty} (a_k^2(g(X_0)) + b_k^2(g(X_0))) \leq \sum_{k=m+1}^{\infty} \frac{2}{\pi^2 k^2} \leq \frac{C}{m},$$

the Fourier coefficients a_k and b_k being defined in Section 4.1. To prove the second point note that, since g is monotonic, the set $\{g(X_k) \leq t\}$ is of the form $\{X_k \leq u\}$, or $\{X_k < u\}$, or $\{X_k \geq u\}$, or $\{X_k > u\}$, for some $u \in [0,1]$. Hence

$$|F_{g(X_k)|X_0}(t) - F(t)| \leq \sup_{u\in[0,1]} |F_{X_k|X_0}(u) - \mathbb{P}(X_k \leq u)|,$$

and consequently

$$\beta(k) = \mathbb{E}\Big(\sup_{t\in[0,1]} |F_{g(X_k)|X_0}(t) - F(t)|\Big) \leq \mathbb{E}\Big(\sup_{u\in[0,1]} |F_{X_k|X_0}(u) - \mathbb{P}(X_k \leq u)|\Big). \tag{5.3}$$

By Proposition 6.2 of the appendix

$$\mathbb{E}\Big(\sup_{u\in[0,1]} |F_{X_k|X_0}(u) - \mathbb{P}(X_k \leq u)|\Big) \leq \frac{C}{(k+1)^{(1-\gamma)/\gamma}}, \tag{5.4}$$

for some positive constant C. From (5.3) and (5.4), it follows that the coefficient $\beta(k)$ defined in (5.2) satisfies (4.13).

For the almost sure behavior of G_n, the proof is exactly the same as that of Theorem 1.5, since the coefficient $\beta(k)$ defined in (5.2) satisfies (4.13).

6 Appendix

In this section, C is a positive constant which may vary from line to line.

6.1 A maximal inequality in Hilbert spaces

The following proposition is used in the proof of Theorem 1.1. It is adapted from Proposition 4 in Dedecker and Merlevède (2007).

Proposition 6.1. *Let $Y_1, Y_2, \ldots, Y_n$ be n random variables with values in a separable Hilbert space H with norm $\|\cdot\|_H$, such that $\mathbb{P}(\|Y_k\|_H \leq M) = 1$ and $\mathbb{E}(Y_k) = 0$ for any $k \in \{1, \ldots, n\}$. Let $\mathcal{F}_1, \ldots, \mathcal{F}_n$ be an increasing filtration such that Y_k is $\mathcal{F}_k$ mesurable for any $k \in \{1, \ldots, n\}$. Let $S_n = \sum_{k=1}^n Y_k$, and for $k \in \{0, \ldots, n-1\}$, let*

$$\theta(k) = \max\Big\{\mathbb{E}(\|\mathbb{E}(Y_j|\mathcal{F}_i)\|_H), (i,j) \in \{1, \ldots, n\}^2 \text{ such that } j \geq i + k\Big\}.$$

Then, for any $q \in \{1, \ldots, n\}$, and any $x \geq qM$, the following inequality holds

$$\mathbb{P}\Big(\max_{1\leq k\leq n} \|S_k\|_H \geq 4x\Big) \leq \frac{n\theta(q)}{x}\mathbf{1}_{q<n} + \frac{2nM}{x^2}\sum_{k=0}^{q-1}\theta(k).$$

Proof of Proposition 6.1. Let $S_0 = 0$ and define the random variables U_i by: $U_i = S_{iq} - S_{(i-1)q}$ for $i \in \{1, \dots, [n/q]\}$ and $U_{[n/q]+1} = S_n - S_{q[n/q]}$. By Proposition 4 in Dedecker and Merlevède (2007), for any $x \geq Mq$,

$$\begin{aligned}
\mathbb{P}\Big(\max_{1\leq k\leq n} \|S_k\|_H \geq 4x\Big) &\leq \frac{1}{x} \sum_{i=3}^{[n/q]+1} \mathbb{E}(\|\mathbb{E}(U_i|\mathcal{F}_{(i-2)q})\|_H) + \frac{1}{x^2} \sum_{i=1}^{[n/q]+1} \mathbb{E}(\|U_i - \mathbb{E}(U_i|\mathcal{F}_{(i-2)q})\|_H^2) \\
&\leq \frac{1}{x} \sum_{i=3}^{[n/q]+1} \mathbb{E}(\|\mathbb{E}(U_i|\mathcal{F}_{(i-2)q})\|_H) + \frac{1}{x^2} \sum_{i=1}^{[n/q]+1} \mathbb{E}(\|U_i\|_H^2)\,, \qquad (6.1)
\end{aligned}$$

the second inequality being satisfied because $\|U_i\|_H^2 = \|U_i - \mathbb{E}(U_i|\mathcal{F}_{(i-2)q})\|_H^2 + \|\mathbb{E}(U_i|\mathcal{F}_{(i-2)q})\|_H^2$ by orthogonality.

To handle the first term in (6.1), note that $\theta(k)$ decreases with k and, according to the definition of $\theta(k)$: for $i \in \{1, \dots, [n/q]\}$,

$$\mathbb{E}(\|\mathbb{E}(U_i|\mathcal{F}_{(i-2)q})\|_H) \leq \sum_{j=(i-1)q+1}^{iq} \|\mathbb{E}(Y_j|\mathcal{F}_{(i-2)q})\|_H \leq \sum_{j=(i-1)q+1}^{iq} \theta(j-(i-2)q) \leq q\theta(q)\,, \quad (6.2)$$

and

$$\mathbb{E}(\|\mathbb{E}(U_{[n/q]+1}|\mathcal{F}_{([n/q]-1)q})\|_H) \leq \sum_{j=q[n/q]+1}^{n} \|\mathbb{E}(Y_j|\mathcal{F}_{([n/q]-1)q})\|_H \leq (n-q[n/q])\theta(q)\,. \qquad (6.3)$$

From (6.2) and (6.3), and taking into account that the sum from $i = 3$ to $[n/q] + 1$ is 0 if $q = n$, we infer that

$$\sum_{i=3}^{[n/q]+1} \mathbb{E}(\|\mathbb{E}(U_i|\mathcal{F}_{(i-2)q})\|_H) \leq n\theta(q)\mathbf{1}_{q<n}\,. \qquad (6.4)$$

To handle the second term in (6.1), we start from the basic equalities: for $i \in \{1, \dots, [n/q]\}$,

$$\mathbb{E}(\|U_i\|_H^2) = \sum_{j=(i-1)q+1}^{iq} \mathbb{E}(\|Y_j\|_H^2) + 2 \sum_{j=(i-1)q+1}^{iq} \sum_{\ell=(i-1)q+1}^{j-1} \mathbb{E}(< Y_j, Y_\ell >)\,, \qquad (6.5)$$

and

$$\mathbb{E}(\|U_{[n/q]+1}\|_H^2) = \sum_{q[n/q]+1}^{n} \mathbb{E}(\|Y_j\|_H^2) + 2 \sum_{j=q[n/q]+1}^{n} \sum_{\ell=q[n/q]+1}^{j-1} \mathbb{E}(< Y_j, Y_\ell >)\,. \qquad (6.6)$$

Taking the conditional expectation of Y_j with respect to $\mathcal{F}_\ell$ and proceeding exactly as to prove (4.18), we obtain from (6.5) and (6.6) that: for $i \in \{1, \dots, [n/q]\}$,

$$\begin{aligned}
\mathbb{E}(\|U_i\|_H^2) &\leq 2 \sum_{j=(i-1)q+1}^{iq} \sum_{\ell=(i-1)q+1}^{j} \|Y_\ell\|_\infty \|\mathbb{E}(Y_j|\mathcal{F}_\ell)\|_H \\
&\leq 2M \sum_{j=(i-1)q+1}^{iq} \sum_{\ell=(i-1)q+1}^{j} \theta(j-\ell) \leq 2Mq \sum_{k=0}^{q-1} \theta(k)\,, \qquad (6.7)
\end{aligned}$$

and

$$\mathbb{E}(\|U_{[n/q]+1}\|_H^2) \leq 2M(n-q[n/q]) \sum_{k=0}^{n-q[n/q]-1} \theta(k) \leq 2M(n-q[n/q]) \sum_{k=0}^{q-1} \theta(k)\,. \qquad (6.8)$$

From (6.7) and (6.8), we infer that

$$\sum_{i=1}^{[n/q]+1} \mathbb{E}(\|U_i\|_H^2) = \mathbb{E}(\|U_{[n/q]+1}\|_H^2) + \sum_{i=1}^{[n/q]} \mathbb{E}(\|U_i\|_H^2) \leq 2nM \sum_{k=0}^{q-1} \theta(k)\,. \qquad (6.9)$$

Starting from (6.1) and using the upper bounds (6.4) and (6.9), Proposition 6.1 follows.

6.2 Proof of the upper bound (4.13) on the coefficient $\beta(k)$.

Proposition 6.2. *Let $\gamma \in]0,1[$, and let $(X_i)_{i\geq 0}$ be a stationary Markov chain with transition kernel K defined in (1.3) and invariant measure ν. Then the coefficient $\beta(k)$ defined in (4.12) satisfies the upper bound (4.13).*

Remark 6.1. In fact, the upper bound (4.13) holds for Markov chains associated with the class of generalized Pomeau-Manneville (GPM) maps introduced in Definition 1.1 of Dedecker, Gouëzel and Merlevède (2010).

Proof of Proposition 6.2. The proof is a slight modification of the proof of Proposition 1.16 in Dedecker, Gouëzel and Merlevède (2010) and is included here for the sake of completeness.

If f is supported in $[0,1]$, let $V(f)$ be the variation of the function f, given by

$$\mathrm{V}(f) = \sup_{x_0 < \cdots < x_N} \sum_{i=1}^{N} |f(x_{i+1}) - f(x_i)|\,, \tag{6.10}$$

where the x_i's are real numbers (not necessarily in $[0,1]$). Let also $\|df\|$ denotes the variation norm of the signed measure df on $[0,1]$, which is defined as in (6.10) with all the x_i's in $[0,1]$. Note that, since f is supported on $[0,1]$,

$$\|f\|_\infty \leq \mathrm{V}(f) = \|df\| + |f(0)| + |f(1)|\,, \tag{6.11}$$

where $\|f\|_\infty = \sup\{|f(x)|, x \in [0,1]\}$. Note also that, for any $x \in [0,1]$,

$$|f(x) - \nu(f)| = \Big|\int_0^1 (F(t) - \mathbf{1}_{x\leq t}) df(t)\Big| \leq \|df\|\,. \tag{6.12}$$

In particular, it follows from (6.11) and (6.12) that

$$\mathrm{V}(f - \nu(f)) = \|df\| + |f(0) - \nu(f)| + |f(1) - \nu(f)| \leq 3\|df\|\,. \tag{6.13}$$

Let K be the transition kernel defined in (1.3). Recall that an equivalent definition of the coefficient $\beta(k)$ defined in (4.12) is

$$\beta(k) = \nu\Big(\sup_{f:\|df\|\leq 1} |K^k f - \nu(f)|\Big) \tag{6.14}$$

(cf. Lemma 1 in Dedecker and Prieur (2005)). Recall also that one has the decomposition

$$K^n f = \sum_{a+k+b=n} A_a(\mathbf{1}_{(z_1,1]}) \cdot \nu(B_b f) + \sum_{a+k+b=n} A_a E_k B_b f + C_n f\,, \tag{6.15}$$

where the operators A_n, B_n, C_n and E_n and the sequence $(z_n)_{n\geq 0}$ are defined in Section 3 of the paper by Dedecker, Gouëzel and Merlevède (2010). In particular, it is proved in this paper that

$$\mathrm{V}(E_k f) \leq \frac{C}{k^{(1-\gamma)/\gamma}}\,\mathrm{V}(f) \quad \text{and} \quad \mathrm{V}(B_n f) \leq \frac{C\,\mathrm{V}(f)}{(n+1)^{1/\gamma}}\,. \tag{6.16}$$

Following the proof of Proposition 1.16 in Dedecker, Gouëzel and Merlevède (2010), one has that

$$|C_n(f)| \leq C\|f\|_\infty K^n \mathbf{1}_{[0,z_{n+1}]}\,, \tag{6.17}$$

and

$$\nu(K^n \mathbf{1}_{[0,z_{n+1}]}) = \nu([0,z_{n+1}]) \leq \frac{C}{(n+1)^{(1-\gamma)/\gamma}}\,. \tag{6.18}$$

We now turn to the term $\sum_{a+k+b=n} A_a E_k B_b f$ in (6.15). Following the proof of Proposition 1.16 in Dedecker, Gouëzel and Merlevède (2010), for any bounded function g,

$$|A_n(g)| \leq C\|g\|_\infty K^n \mathbf{1}_{(z_1,1]\cap T^{-1}[0,z_n]}\,, \tag{6.19}$$

and

$$\nu(K^n \mathbf{1}_{(z_1,1]\cap T^{-1}[0,z_n]}) = \nu((z_1,1]\cap T^{-1}[0,z_n]) \leq \frac{C}{(n+1)^{1/\gamma}}\,. \tag{6.20}$$

Using successively (6.19), (6.11) and (6.16), we obtain that

$$\begin{aligned}\Big|\sum_{a+k+b=n} A_a E_k B_b f\Big| &\leq C \sum_{a+k+b=n} \|E_k B_b f\|_\infty K^a \mathbf{1}_{(z_1,1]\cap T^{-1}[0,z_a]} \\ &\leq C \sum_{a+k+b=n} \mathrm{V}(B_b f) \frac{K^a \mathbf{1}_{(z_1,1]\cap T^{-1}[0,z_a]}}{(k+1)^{(1-\gamma)/\gamma}} \\ &\leq C\,\mathrm{V}(f) \sum_{a+k+b=n} \frac{K^a \mathbf{1}_{(z_1,1]\cap T^{-1}[0,z_a]}}{(k+1)^{(1-\gamma)/\gamma}(b+1)^{1/\gamma}}\,. \end{aligned} \tag{6.21}$$

We now turn to the term $\sum_{a+k+b=n} A_a(\mathbf{1}_{(z_1,1]})\cdot\nu(B_b f)$ in (6.15). From the displayed inequality right before (3.13) in Dedecker, Gouëzel and Merlevède (2010), one has

$$\Big|\sum_{b=0}^{n-a} \nu(B_b f)\Big| = \Big|\sum_{b>n-a} \nu(B_b f)\Big| \leq \sum_{b>n-a} \mathrm{V}(B_b f) \leq \sum_{b>n-a} \frac{C\,\mathrm{V}(f)}{(b+1)^{1/\gamma}} \leq \frac{D\,\mathrm{V}(f)}{(n+1-a)^{(1-\gamma)/\gamma}}\,. \tag{6.22}$$

From (6.22) and (6.19), we obtain

$$\Big|\sum_{a=0}^{n} A_a(\mathbf{1}_{(z_1,1]}) \cdot \Big(\sum_{b=0}^{n-a} \nu(B_b f)\Big)\Big| \leq C\,\mathrm{V}(f) \sum_{a=0}^{n} \frac{K^a \mathbf{1}_{(z_1,1]\cap T^{-1}[0,z_a]}}{(n+1-a)^{(1-\gamma)/\gamma}}\,. \tag{6.23}$$

From (6.13), $\mathrm{V}(f-\nu(f)) \leq 3\|df\|$. Hence, it follows from (6.15), (6.17), (6.21) and (6.23) that

$$\begin{aligned} &|K^n(f-\nu(f))| \\ &\leq C\|df\|\Big(K^n \mathbf{1}_{[0,z_{n+1}]} + \sum_{a=0}^{n} \frac{K^a \mathbf{1}_{(z_1,1]\cap T^{-1}[0,z_a]}}{(n+1-a)^{(1-\gamma)/\gamma}} + \sum_{a+k+b=n} \frac{K^a \mathbf{1}_{(z_1,1]\cap T^{-1}[0,z_a]}}{(k+1)^{(1-\gamma)/\gamma}(b+1)^{1/\gamma}}\Big)\,. \end{aligned} \tag{6.24}$$

From (6.14), (6.24), (6.18) and (6.20), it follows that

$$\begin{aligned}\beta(n) \leq C\Big(&\frac{1}{(n+1)^{(1-\gamma)/\gamma}} + \sum_{a=0}^{n} \frac{1}{(a+1)^{1/\gamma}(n+1-a)^{(1-\gamma)/\gamma}} \\ &+ \sum_{a+k+b=n} \frac{1}{(a+1)^{1/\gamma}(k+1)^{(1-\gamma)/\gamma}(b+1)^{1/\gamma}}\Big)\,. \end{aligned} \tag{6.25}$$

All the sums on right hand being of the same order (see Lemma 3.2 of Dedecker, Gouëzel and Merlevède (2010), and its application at the beginning of the proof of their Proposition 1.15), it follows that

$$\beta(n) \leq \frac{C}{(n+1)^{(1-\gamma)/\gamma}}\,,$$

and the proof is complete.

Acknowledgements. Herold Dehling was partially supported by the Collaborative Research Project *Statistical Modeling of Nonlinear Dynamic Processes* (SFB 823) of the German Research Foundation DFG. Murad S. Taqqu was partially supported by the NSF grants DMS-1007616 and DMS-1309009 at Boston University.

References

[1] J. Dedecker (2010), An empirical central limit theorem for intermittent maps. *Probab. Theory Relat. Fields* **148**, 177-195.

[2] J. Dedecker, S. Gouëzel, and F. Merlevède (2010), Some almost sure results for unbounded functions of intermittent maps and their associated Markov chains. *Ann. Inst. Henri Poincaré Probab. Stat.* **46**, 796-821.

[3] J. Dedecker and F. Merlevède (2007), Convergence rates in the law of large numbers for Banach-valued dependent random variables, *Teor. Veroyatnost. i Primenen* **52**, 562-587.

[4] J. Dedecker, F. Merlevède and E. Rio (2013), On strong approximation for the empirical process of stationary sequences. *Ann. Probab.* **41**, 3658–3696.

[5] J. Dedecker and C. Prieur (2005). New dependence coefficients. Examples and applications to statistics. *Probability Theory and Related Fields* **132**, 203–236.

[6] H. Dehling (1983), Limit theorems for sums of weakly dependent Banach space valued random variables. *Z. für Wahrsch. Verw. Gebiete* **63**, 393–432.

[7] H. Dehling and M. S. Taqqu (1989), The empirical process of some long-range dependent sequences with an application to U-statistics. *Ann. Statist.* **17**, 1767-1783.

[8] W. Feller (1966), An introduction to probability theory and its applications. Vol. II. John Wiley & Sons, Inc., New York-London-Sydney, 636 pp.

[9] M. Fréchet (1957), Sur la distance de deux lois de probabilité. (French) *Publ. Inst. Statist. Univ. Paris* **6**, 183–198.

[10] S. Gouëzel (2004), Central limit theorem and stable laws for intermittent maps. *Probab. Theory Relat. Fields* **128**, 82-122.

[11] H. Hennion and L. Hervé (2001), Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness. *Lecture Notes in Mathematics* **1766**, Springer.

[12] C. C. Heyde (1969), A note concerning behaviour of iterated logarithm type. *Proc. Amer. Math. Soc.* **23**, 85–90.

[13] J. Kuelbs and W. Philipp (1980), Almost sure invariance principles for partial sums of mixing B-values random variables. *Ann. Probab.* **8**, 1003–1036.

[14] C. Liverani, B. Saussol and S. Vaienti (1999), A probabilistic approach to intermittency. *Ergodic Theory Dynam. Systems* **19**, 671–685.

[15] E. Mourier (1953), Éléments aléatoires dans un espace de Banach. *Ann. Inst. Henri Poincaré* **13**, 161–244.

[16] Y. Pomeau and P. Manneville (1980), Intermittent transition to turbulence in dissipative dynamical systems. *Comm. Math. Phys.* **74**, 189–197.

[17] G. Samorodnitsky and M. S. Taqqu (1994), Stable non-Gaussian random processes : Stochastic models with infinite variance. *Stochastic Modeling.* Chapman & Hall, New York.

[18] O. Sarig (2002), Subexponential decay of correlations. *Inv. Math.* **150**, 629–653.

[19] D. Surgailis (2002), Stable limits of empirical processes of moving averages with infinite variance. *Stochastic Process. Appl.* **100**, 255–274.

[20] M. Thaler (1980), Estimates of the invariant densities of endomorphisms with indifferent fixed points. *Israel J. Math.* **37**, 303–314.

[21] L.-S. Young (1999), Recurrence times and rates of mixing. *Israel J. Math.* **110**, 153–188.