Bernoulli **23**(3), 2017, 2083–2127
DOI: 10.3150/16-BEJ805

Behavior of the Wasserstein distance between the empirical and the marginal distributions of stationary α-dependent sequences

JÉRÔME DEDECKER[1] and FLORENCE MERLEVÈDE[2]

[1]*Université Paris Descartes, Sorbonne Paris Cité, Laboratoire MAP5 and CNRS UMR 8145. E-mail: jerome.dedecker@parisdescartes.fr*
[2]*Université Paris Est, UPEM, UPEC, LAMA and CNRS UMR 8050*

We study the Wasserstein distance of order 1 between the empirical distribution and the marginal distribution of stationary α-dependent sequences. We prove some moments inequalities of order p for any $p \geq 1$, and we give some conditions under which the central limit theorem holds. We apply our results to unbounded functions of expanding maps of the interval with a neutral fixed point at zero. The moment inequalities for the Wasserstein distance are similar to the well-known von Bahr–Esseen or Rosenthal bounds for partial sums, and seem to be new even in the case of independent and identically distributed random variables.

Keywords: central limit theorem; empirical process; intermittency; moments inequalities; stationary sequences; Wasserstein distance

1. Introduction

Let $(X_i)_{i\in\mathbb{Z}}$ be a stationary sequence of integrable real-valued random variables, with common marginal distribution μ. Let μ_n be the empirical measure of $\{X_1, \dots, X_n\}$, that is

$$\mu_n = \frac{1}{n}\sum_{k=1}^{n} \delta_{X_k}.$$

In this paper, we study the behavior of the quantity $W_1(\mu_n, \mu)$ for a large class of stationary sequences, where $W_1(\mu_1, \mu_2)$ is the *Wasserstein distance* of order 1 between two probability measures μ_1, μ_2 having finite first moments. The precise definition is as follows:

$$W_1(\mu_1, \mu_2) = \inf_{\pi \in M(\mu_1,\mu_2)} \int |x - y| \pi(dx, dy), \tag{1.1}$$

where $M(\mu_1, \mu_2)$ is the set of probability measures on $\mathbb{R}^2$ with marginal distributions μ_1 and μ_2. The distance W_1 belongs to the general class of minimal distances, as the total variation distance. Since the cost function $c_1(x, y) = |x - y|$ is regular, W_1 can be used to compare two singular measures, which is not possible with the total variation distance, whose cost function is given by the discrete metric $c_0(x, y) = \mathbf{1}_{x\neq y}$.

1350-7265

The quantity $W_1(\mu_n, \mu)$ appears very frequently in statistics, and can be understood from many points of view:

- The well-known dual representation of W_1 implies that

$$W_1(\mu_n, \mu) = \sup_{f \in \Lambda_1} \left| \frac{1}{n} \sum_{k=1}^{n} \big(f(X_k) - \mu(f) \big) \right|, \tag{1.2}$$

  where Λ_1 is the set of Lipschitz functions f from $\mathbb{R}$ to $\mathbb{R}$ such that $|f(x) - f(y)| \le |x - y|$. Hence, $W_1(\mu_n, \mu)$ is a measure of the concentration of μ_n around μ through the class Λ_1.
- In the one dimensional setting the minimization problem (1.1) can be explicitly solved, and leads to the expression

$$W_1(\mu_n, \mu) = \int_0^1 \left| F_n^{-1}(t) - F^{-1}(t) \right| dt, \tag{1.3}$$

  where F_n and F are the distribution functions of μ_n and μ, and F_n^{-1} and F^{-1} are their usual generalized inverses. Hence, $W_1(\mu_n, \mu)$ is the $\mathbb{L}^1$-distance between the empirical quantile function F_n^{-1} and the quantile function of μ.
- Starting from (1.3), it follows immediately that

$$W_1(\mu_n, \mu) = \int_{\mathbb{R}} \left| F_n(t) - F(t) \right| dt. \tag{1.4}$$

  Hence, $W_1(\mu_n, \mu)$ is the $\mathbb{L}^1$-distance between the empirical distribution function F_n and the distribution function of μ.

At this point, it should be clearly quoted that, if (1.3) and (1.4) have no analogue in higher dimension, the dual expression (1.2) is very general and holds if the X_i's take their values in a Polish space $\mathcal{X}$, as soon as the cost function c is a lower semi-continuous metric (the class Λ_1 being the class of 1-Lipschitz functions from $\mathcal{X}$ to $\mathbb{R}$ with respect to c).

Assume now that the sequence $(X_i)_{i \in \mathbb{Z}}$ is ergodic. Since μ has a finite first moment, it is well known that $W_1(\mu_n, \mu)$ converges to zero almost surely, and that $\mathbb{E}(W_1(\mu_n, \mu))$ converges to zero (this is a uniform version of Birkhoff's ergodic theorem, which can be easily deduced from the Glivenko–Cantelli theorem for ergodic sequences, see, for instance, [25]). However, without additional assumptions on μ the rate of convergence can be arbitrarily slow.

The purpose of this paper is to give some conditions under which the central limit theorem (CLT) holds (meaning that $\sqrt{n} W_1(\mu_n, \mu)$ converges in distribution to a certain law), and to prove some inequalities for $\|W_1(\mu_n, \mu)\|_p$ when $p \ge 1$ (von Bahr–Esseen type inequalities for $p \in (1, 2)$ and Rosenthal type inequalities for $p > 2$). We will do this for the class of α-dependent sequences, which is quite natural in this context, since the related dependency coefficients are defined through indicator of half lines. Hence, our results apply to mixing sequences in the sense of Rosenblatt [24], but also to many other dependent sequences including a large class of one dimensional dynamical systems. We shall illustrate our results through the examples of Generalized Pomeau–Manneville maps, as defined in [7].

The central limit question for $\sqrt{n} W_1(\mu_n, \mu)$ has been already investigated for dependent sequences in the papers by Dédé [5] and Cuny [3] (see Sections 6 and 7 for more details). This is not the case of the upper bounds for $\|W_1(\mu_n, \mu)\|_p$, even for sequences of independent and identically distributed (i.i.d.) random variables (except for $p = 1$, see for instance [1]). Hence, for $p > 1$, our moment bounds seem to be new even in the i.i.d. context.

Thanks to the relation (1.4), the central limit question for $\sqrt{n} W_1(\mu_n, \mu)$ is closely related to the empirical central limit theorem in $\mathbb{L}^1(dt)$, as first quoted by del Barrio, Giné and Matrán [10]. We shall deal with the more general central limit question for $\mathbb{L}^1(m)$-valued random variables in the separate Section 6. In Section 7, we shall express some of our conditions in terms of the quantile function of X_0, in the spirit of Doukhan, Massart and Rio [11]. It will then be easier to compare our conditions for the CLT to previous ones in the literature.

For $r > 1$, the quantity $W_r^r(\mu_n, \mu)$ may be defined as in (1.1), with the cost function $c_r(x, y) = |x - y|^r$ instead of c_1 (W_r is the Wasserstein distance of order r). In the i.i.d. case, some sharp upper bounds on $\mathbb{E}(W_r^r(\mu_n, \mu))$ are given in the recent paper [1]. In particular, if μ has an absolutely component with respect to the Lebesgue measure which does not vanish on the support of μ, then the optimal rate $n^{-r/2}$ can be reached. But in general, the rate can be much slower. Note that for $W_r^r(\mu_n, \mu)$ there is no such nice dual expression as (1.2). However the minimization problem can still be explicitly solved and implies that $W_r(\mu_n, \mu)$ is the $\mathbb{L}^r$-distance between F_n^{-1} and F^{-1}. There is no simple way to express $W_r^r(\mu_n, \mu)$ in terms of F_n and F (as in (1.4)), but the following upper bound due to Èbralidze [12] holds:

$$W_r^r(\mu_n, \mu) \le \kappa_r \int_{\mathbb{R}} |x|^{r-1} \big| F_n(x) - F(x) \big| \, dx, \tag{1.5}$$

where $\kappa_r = 2^{r-1} r$. Starting from this inequality, we shall also give some upper bounds on $\|W_r^r(\mu_n, \mu)\|_p$ for $p \ge 1$, but it is very likely that these bounds can be improved by assuming the existence of an absolutely regular component for μ, as in [1].

To be complete, let us mention the recent paper by Fournier and Guillin [14], who give some upper bounds for $\mathbb{E}(W_r^r(\mu_n, \mu))$ in any dimension, starting from an inequality which can be viewed as a d-dimensional analogue of (1.5). Note that the case of ρ-mixing sequences is also considered in this paper.

The paper is organized as follows. In Section 2, we give the notations and definitions which will be used all along the paper. In Section 3, we state the main results of the paper: a central limit theorem in Section 3.1, some upper bounds for the moments of order 1 and 2 in Section 3.2, for the moments of order $p \in (1, 2)$ in Section 3.3, and for the moments of order $p \ge 2$ in Section 3.4. In Section 4, we apply our main results to the empirical measure of unbounded observables of the iterates of intermittent maps. The proofs of the moment bounds are given in Section 5. In Section 6, we give some general results for the weak convergence of partial sums in $\mathbb{L}^1(dt)$. Finally, in Section 7, we give another expression of our condition for the central limit theorem, and we compare this condition to previous ones in the literature.

2. Definitions and notations

Let us start with the notation $a_n(x) \ll b_n(x)$, which means that there exists a positive constant C not depending on n nor x such that $a_n(x) \le C b_n(x)$, for all positive integer n and all real x.

We now introduce the probabilistic framework, and give our definition of a stationary sequence. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, and $T : \Omega \mapsto \Omega$ be a bijective bi-measurable transformation preserving the probability $\mathbb{P}$. Let $\mathcal{F}_0$ be a sub-σ-algebra of $\mathcal{A}$ satisfying $\mathcal{F}_0 \subseteq T^{-1}(\mathcal{F}_0)$. We say that the couple $(T, \mathbb{P})$ is ergodic if any $A \in \mathcal{A}$ satisfying $T(A) = A$ has probability 0 or 1.

Let X_0 be an $\mathcal{F}_0$-measurable and integrable real-valued random variable with distribution μ. Define the stationary sequence $\mathbf{X} = (X_i)_{i\in\mathbb{Z}}$ by $X_i = X_0 \circ T^i$.

Note that every stationary sequence $\mathbf{X} = (X_i)_{i\in\mathbb{Z}}$ in the usual sense can be represented this way by considering the state space $\Omega = (\mathbb{R}^{\mathbb{Z}}, \mathcal{B}(\mathbb{R})^{\mathbb{Z}}, \mathbb{P}_{\mathbf{X}})$, where $\mathcal{B}(\mathbb{R})$ is the Borel σ-field and $\mathbb{P}_{\mathbf{X}}$ is the distribution of $\mathbf{X}$, and by taking the shift operator $T : \mathbb{R}^{\mathbb{Z}} \mapsto \mathbb{R}^{\mathbb{Z}}$ defined by $(T(\omega))_i = \omega_{i+1}$. Clearly $\mathbb{P}_{\mathbf{X}}$ is invariant by T and, if $\pi_0 : \mathbb{R}^{\mathbb{Z}} \mapsto \mathbb{R}$ is the projection defined by $\pi_0(\omega) = \omega_0$, the sequence $\pi = (\pi_0 \circ T^i)_{i\in\mathbb{Z}}$ is stationary and distributed as $\mathbf{X}$. Hence, this formalism is very general, and enables to define the notion of ergodicity in a precise manner.

Let us now define the dependency coefficients of the sequence $(X_i)_{i\in\mathbb{Z}}$. These coefficients are less restrictive than the usual mixing coefficients of Rosenblatt [24].

Definition 2.1. *For any integrable random variable Z, let $Z^{(0)} = Z - \mathbb{E}(Z)$. For any random variable $Y = (Y_1, \dots, Y_k)$ with values in $\mathbb{R}^k$ and any σ-algebra $\mathcal{F}$, let*

$$\alpha(\mathcal{F}, Y) = \sup_{(x_1,\dots,x_k)\in\mathbb{R}^k} \left\| \mathbb{E}\left(\prod_{j=1}^{k} (\mathbf{1}_{Y_j \le x_j})^{(0)} \Big| \mathcal{F} \right) - \mathbb{E}\left(\prod_{j=1}^{k} (\mathbf{1}_{Y_j \le x_j})^{(0)} \right) \right\|_1 .$$

For the stationary sequence $\mathbf{X} = (X_i)_{i\in\mathbb{Z}}$, let

$$\alpha_{k,\mathbf{X}}(n) = \max_{1\le l\le k} \sup_{n\le i_1\le\cdots\le i_l} \alpha\big(\mathcal{F}_0, (X_{i_1}, \dots, X_{i_l})\big). \tag{2.1}$$

Note that $\alpha_{1,\mathbf{X}}(n)$ is then simply given by

$$\alpha_{1,\mathbf{X}}(n) = \sup_{x\in\mathbb{R}} \big\| \mathbb{E}(\mathbf{1}_{X_n\le x} | \mathcal{F}_0) - F(x) \big\|_1, \tag{2.2}$$

where F is the distribution function of μ.

All the results of Section 3 below involve only the coefficients $\alpha_{1,\mathbf{X}}(n)$, except for the Rosenthal bounds (Section 3.4) for which the coefficient $\alpha_{2,\mathbf{X}}(n)$ is needed.

To conclude this section, we introduce the tail and quantile functions of the random variable X_0.

Definition 2.2. *The tail function $H : \mathbb{R}^+ \to [0, 1]$ of X_0 is defined by $H(t) = \mathbb{P}(|X_0| > t)$. The quantile function $Q : [0, 1] \to \mathbb{R}^+$ of X_0 is the generalized inverse of H, that is*

$$Q(u) = \inf\big\{t \ge 0 : H(t) \le u\big\}.$$

3. Main results

In all this section, we use the notations of Sections 1 and 2.

3.1. Central limit theorem

Our first result is a central limit theorem for $W_1(\mu_n, \mu)$. It is a straightforward consequence of a CLT in $\mathbb{L}^1(m)$ for the empirical distribution function given in Proposition 6.2 of Section 6.4 (it suffices to consider the case where m is the Lebesgue measure on $\mathbb{R}$ and to use the continuous mapping theorem).

Proposition 3.1. *Assume that the couple* $(T, \mathbb{P})$ *is ergodic, and that*

$$\int_0^\infty \sqrt{\sum_{k=0}^\infty \min\{\alpha_{1,\mathbf{X}}(k), H(t)\}}\, dt < \infty. \tag{3.1}$$

Then $\sqrt{n} W_1(\mu_n, \mu)$ *converges in distribution to the random variable* $\int |G(t)|\, dt$, *where* G *is a Gaussian random variable in* $\mathbb{L}^1(dt)$ *whose covariance function may be described as follows: for any* f, g *in* $\mathbb{L}_\infty(\mu)$,

$$\begin{aligned} &\operatorname{Cov}\biggl(\int f(t)G(t)\,dt, \int g(t)G(t)\,dt\biggr) \\ &\quad = \sum_{k\in\mathbb{Z}} \mathbb{E}\biggl(\iint f(t)g(s)\bigl(\mathbf{1}_{X_0\le t} - F(t)\bigr)\bigl(\mathbf{1}_{X_k\le s} - F(s)\bigr)\,dt\,ds\biggr). \end{aligned} \tag{3.2}$$

Remark 3.1. Let m be a nonnegative integer. As usual, the stationary sequence $\mathbf{X}$ is m-dependent if $\sigma(X_i, i \le 0)$ is independent of $\sigma(X_i, i \ge m+1)$, and $m = 0$ corresponds to the i.i.d. case. In the m-dependent case, the condition (3.1) becomes simply

$$\int_0^\infty \sqrt{H(t)}\, dt < \infty, \tag{3.3}$$

which is exactly the condition given by del Barrio, Giné and Matrán [10] in the i.i.d. case. Note that these authors also proved that, in the i.i.d. case, the condition (3.3) is necessary and sufficient for the stochastic boundedness of $\sqrt{n} W_1(\mu_n, \mu)$.

In the dependent context, other general criteria have been proposed by Dédé [5] and Cuny [3]. We shall discuss these conditions in Sections 6 and 7, and show that, in the α-dependent case, the condition (3.1) is weaker than the corresponding condition obtained by applying the criteria by Dédé or Cuny.

3.2. Upper bounds for moments of order 1 and 2

In this section, we give some upper bounds for the quantities $\mathbb{E}(W_1(\mu_n,\mu))$ and $\|W_1(\mu_n,\mu)\|_2$ in terms of the coefficients $\alpha_{1,\mathbf{X}}(k)$ and of the tail function H. The proof of Proposition 3.2 below will be done in Section 5.1.

For any $t \geq 0$, let

$$S_{\alpha,n}(t) = \sum_{k=0}^{n} \min\{\alpha_{1,\mathbf{X}}(k), H(t)\}. \tag{3.4}$$

Proposition 3.2. *The following upper bounds hold*:

$$\mathbb{E}\big(W_1(\mu_n,\mu)\big) \leq 4\int_0^\infty \sqrt{\min\left\{\big(H(t)\big)^2, \frac{S_{\alpha,n}(t)}{n}\right\}}\, dt, \tag{3.5}$$

and

$$\big\|W_1(\mu_n,\mu)\big\|_2 \leq \frac{2\sqrt{2}}{\sqrt{n}}\int_0^\infty \sqrt{S_{\alpha,n}(t)}\, dt. \tag{3.6}$$

Remark 3.2. As will be clear from the proof, one can also get some upper bounds involving the quantity $B(t) = F(t)(1 - F(t))$ instead of $H(t)$. For instance, we can obtain an extension of the upper bound given in Theorem 3.5 of [1] to α-dependent sequences. We have chosen to express the upper bounds in terms of the function H, because they are easier to compute in the α-dependent case (see Remark 3.4 below). Note also that the upper bound (3.6) can be viewed as an extension of Theorem 3.2 of [1]: in the i.i.d. case, it gives a similar upper bound (up to constants) bur for the $\mathbb{L}^2$-norm instead of the expectation.

The proof of Proposition 3.2 is based on Minkowski's integral inequality applied to $p = 1$ and $p = 2$:

$$\text{For any } p \geq 1, \qquad \left\|\int \big|F_n(t) - F(t)\big|\, dt\right\|_p \leq \int \big\|F_n(t) - F(t)\big\|_p\, dt.$$

One could also start from this inequality in the case where $p \in (1,2)$ (resp. $p > 2$) by applying a von Bahr–Esseen bound (resp. a Rosenthal bound) to $\|F_n(t) - F(t)\|_p$. However, this would give less satisfactory bounds than in Sections 3.3 and 3.4, even in the i.i.d. case. For instance, in the i.i.d. case and $p \in (1,2)$, this would give

$$\big\|W_1(\mu_n,\mu)\big\|_p^p \ll \frac{1}{n^{p-1}}\left(\int \big(H(t)\big)^{1/p}\, dt\right)^p. \tag{3.7}$$

Note that the condition $\int (H(t))^{1/p}\, dt < \infty$ is more restrictive than $\|X_0\|_p < \infty$. Hence, the upper bound (3.18) of Section 3.3 is always better than (3.7).

Remark 3.3. Starting from Inequality (1.5) and following the proof of Proposition 3.2, we obtain the upper bounds

$$\mathbb{E}\big(W_r^r(\mu_n,\mu)\big) \le 4\int_0^\infty t^{r-1}\sqrt{\min\Big\{\big(H(t)\big)^2, \frac{S_{\alpha,n}(t)}{n}\Big\}}\,dt, \tag{3.8}$$

and

$$\big\|W_r^r(\mu_n,\mu)\big\|_2 \le \frac{2\sqrt{2}}{\sqrt{n}}\int_0^\infty t^{r-1}\sqrt{S_{\alpha,n}(t)}\,dt. \tag{3.9}$$

Remark 3.4. As a consequence of Proposition 3.2, the following upper bounds hold:

1. If (3.1) holds, then $\|W_1(\mu_n,\mu)\|_2 \ll n^{-1/2}$.
2. If $\alpha(k) = O(k^{-a})$ for some $a > 1$, then

$$\mathbb{E}\big(W_1(\mu_n,\mu)\big) \ll \bigg(\int_0^{n^{-a/(a+1)}} Q(u)\,du + \frac{1}{\sqrt{n}}\int_{n^{-a/(a+1)}}^1 \frac{Q(u)}{u^{(a+1)/(2a)}}\,du\bigg), \tag{3.10}$$

and

$$\big\|W_1(\mu_n,\mu)\big\|_2 \ll \bigg(\int_0^{n^{-a}} \frac{Q(u)}{\sqrt{u}}\,du + \frac{1}{\sqrt{n}}\int_{n^{-a}}^1 \frac{Q(u)}{u^{(a+1)/(2a)}}\,du\bigg). \tag{3.11}$$

Let us briefly explain how the bound (3.10) can be obtained. This will indicate to the reader how to handle the computations for the others items of this remark, and for all the examples of Section 4 (see also Section 7 for other computational tricks). First, since $\alpha(k) = O(k^{-a})$ for some $a > 1$, it follows that

$$S_{\alpha,n}(t) \ll \big(H(t)\big)^{(a-1)/a}.$$

Plugging this upper bound in (3.5) and using that Q is the generalized inverse of H, we obtain that

$$\mathbb{E}\big(W_1(\mu_n,\mu)\big) \ll \bigg(\int_{Q(n^{-a/(a+1)})}^\infty H(t)\,dt + \frac{1}{\sqrt{n}}\int_0^{Q(n^{-a/(a+1)})} \big(H(t)\big)^{(a-1)/(2a)}\,dt\bigg). \tag{3.12}$$

To handle the first term on right-hand side in (3.12), we write that

$$\begin{aligned}\int_{Q(n^{-a/(a+1)})}^\infty H(t)\,dt &= \int_{Q(n^{-a/(a+1)})}^\infty \bigg(\int_0^1 \mathbf{1}_{u<H(t)}\,du\bigg)\,dt = \int_0^{n^{-a/(a+1)}} \bigg(\int_0^\infty \mathbf{1}_{t<Q(u)}\,dt\bigg)\,du \\ &= \int_0^{n^{-a/(a+1)}} Q(u)\,du,\end{aligned}$$

which gives the first term on right-hand side in (3.10). For the second term on right-hand side in (3.12), the same kind of computations yields

$$\frac{1}{\sqrt{n}}\int_0^{Q(n^{-a/(a+1)})}\big(H(t)\big)^{(a-1)/(2a)}\,dt=\int_0^1 Q\big(y^{(2a)/(a-1)}\big)\mathbf{1}_{y^{2a/(a-1)}>n^{-a/(a+1)}}\,dy,$$

and the change of variable $u=y^{2a/(a-1)}$ gives the second term on right-hand side in (3.10).

3. If $\alpha(k)=O(a^k)$ for some $a<1$, then

$$\mathbb{E}\big(W_1(\mu_n,\mu)\big)\ll\left(\int_0^{\ln(n)/n}Q(u)\,du+\frac{1}{\sqrt{n}}\int_{\ln(n)/n}^1\frac{Q(u)|\ln(u)|}{\sqrt{u}}\,du\right),$$

and

$$\big\|W_1(\mu_n,\mu)\big\|_2\ll\left(\int_0^{e^{-n}}\frac{Q(u)}{\sqrt{u}}\,du+\frac{1}{\sqrt{n}}\int_{e^{-n}}^1\frac{Q(u)|\ln(u)|}{\sqrt{u}}\,du\right).$$

4. Assume that the α_k's converge to zero, but are not summable, and let

$$u_n=\frac{1}{n}\sum_{k=1}^n\alpha_k.$$

Then

$$\mathbb{E}\big(W_1(\mu_n,\mu)\big)\ll\int_0^{\sqrt{u_n}}Q(u)\,du,\tag{3.13}$$

and

$$\big\|W_1(\mu_n,\mu)\big\|_2\ll\int_0^{u_n}\frac{Q(u)}{\sqrt{u}}\,du.\tag{3.14}$$

Remark 3.5. In the m-dependent case, the inequality (3.10) holds with $a=\infty$, that is

$$\mathbb{E}\big(W_1(\mu_n,\mu)\big)\ll\left(\int_0^{n^{-1}}Q(u)\,du+\frac{1}{\sqrt{n}}\int_{n^{-1}}^1\frac{Q(u)}{\sqrt{u}}\,du\right).$$

In particular, if $H(t)=O(t^{-1}(\ln(t))^{-a})$ for some $a>1$ (which implies that $\mathbb{E}(|X_0|)<\infty$), then $Q(u)=O(u^{-1}|\ln(u)|^{-a})$, and consequently

$$\mathbb{E}\big(W_1(\mu_n,\mu)\big)\ll\frac{1}{(\ln(n))^{a-1}}.$$

3.3. A von Bahr–Esseen type inequality

In this section, we give some upper bounds for the quantity $\|W_1(\mu_n,\mu)\|_p$ when $p\in(1,2)$ in terms of the coefficients $\alpha_{1,\mathbf{X}}(k)$ and of the quantile function Q. The proof of Proposition 3.3 below will be done in Section 5.2.

For $u \in (0,1)$, let

$$\alpha_{1,\mathbf{X}}^{-1}(u) = \sum_{k=0}^{\infty} \mathbf{1}_{u<\alpha_{1,\mathbf{X}}(k)}. \tag{3.15}$$

Proposition 3.3. *For $p \in (1,2)$, the following inequality holds*

$$\left\| W_1(\mu_n, \mu) \right\|_p^p \ll \frac{1}{n^{p-1}} \int_0^1 \left(\alpha_{1,\mathbf{X}}^{-1}(u) \wedge n\right)^{p-1} Q^p(u)\, du. \tag{3.16}$$

Note that Inequality (3.16) writes also

$$\left\| W_1(\mu_n, \mu) \right\|_p^p \ll \frac{1}{n^{p-1}} \sum_{k=0}^{n} \frac{1}{(k+1)^{2-p}} \int_0^{\alpha_{1,\mathbf{X}}(k)} Q^p(u)\, du.$$

Remark 3.6. Let $r \geq 1$ and $p \in (1,2)$. Starting again from (1.5) and following the proof of Proposition 3.3, we obtain the upper bound

$$\left\| W_r^r(\mu_n, \mu) \right\|_p^p \ll \frac{1}{n^{p-1}} \int_0^1 \left(\alpha_{1,\mathbf{X}}^{-1}(u) \wedge n\right)^{p-1} Q^{pr}(u)\, du. \tag{3.17}$$

Remark 3.7. In the m-dependent case, Inequality (3.17) becomes

$$\left\| W_1(\mu_n, \mu) \right\|_p^p \ll \frac{1}{n^{p-1}} \|X_0\|_p^p. \tag{3.18}$$

This inequality seems to be new even in the i.i.d. case. It is noteworthy that the upper bound (3.18) is the same as the moment bound of order p for partial sums of i.i.d. random variables, which can be deduced from the classical inequality of von Bahr and Esseen [26].

3.4. A Rosenthal type inequality

In this section, we give some upper bounds for the quantity $\|W_1(\mu_n, \mu)\|_p$ when $p > 2$ in terms of the coefficients $\alpha_{2,\mathbf{X}}(k)$ and of the quantile function Q. The proof of Proposition 3.4 below will be done in Section 5.3.

The function $\alpha_{2,\mathbf{X}}^{-1}$ is defined as in (3.15) by replacing the coefficient $\alpha_{1,\mathbf{X}}(k)$ by $\alpha_{2,\mathbf{X}}(k)$.

Proposition 3.4. *For $p > 2$, the following inequality holds*:

$$\left\| W_1(\mu_n, \mu) \right\|_p^p \ll \frac{s_{\alpha,n}^p}{n^{p/2}} + \frac{1}{n^{p-1}} \int_0^1 \left(\alpha_{2,\mathbf{X}}^{-1}(u) \wedge n\right)^{p-1} Q^p(u)\, du, \tag{3.19}$$

where

$$s_{\alpha,n} = \int_0^{\infty} \sqrt{S_{\alpha,n}(t)}\, dt$$

and $S_{\alpha,n}$ is the function defined in (3.4).

Note that Inequality (3.19) writes also

$$\left\|W_1(\mu_n,\mu)\right\|_p^p \ll \frac{s_{\alpha,n}^p}{n^{p/2}} + \frac{1}{n^{p-1}}\sum_{k=0}^{n}(k+1)^{p-2}\int_0^{\alpha_{2,\mathbf{X}}(k)} Q^p(u)\,du.$$

Remark 3.8. Inequality (3.19) is similar to the Rosenthal inequality for partial sums given in Theorem 6.3 of Rio [23], with however two main differences:

- First, the variance term is not the same, but this is because we consider the quantity $W_1(\mu_n,\mu)$ and not only the partial sums, in accordance with the upper bounds for $\|W_1(\mu_n,\mu)\|_2$ given in Section 3.2.
- Second, Rio's inequality is stated for α-mixing sequences in the sense of Rosenblatt [24], and its proof relies on the coupling properties of these coefficients. Our result is valid for the larger class of α-dependent sequences as defined in (2.1) (with $k=2$ for the index of the dependency), and the proof is based on a version of the Rosenthal inequality for martingales given in [22]. Note that Rio's inequality cannot be applied to GPM maps, because the associated Markov chain is not α-mixing in the sense of Rosenblatt.

Remark 3.9. Let $r\geq 1$ and $p>2$. Starting again from (1.5) and following the proof of Proposition 3.4, we obtain the upper bound

$$\begin{aligned}&\left\|W_r^r(\mu_n,\mu)\right\|_p^p\\ &\quad\ll \frac{1}{n^{p/2}}\left(\int_0^\infty t^{r-1}\sqrt{S_{\alpha,n}(t)}\,dt\right)^p + \frac{1}{n^{p-1}}\int_0^1\left(\alpha_{2,\mathbf{X}}^{-1}(u)\wedge n\right)^{p-1}Q^{rp}(u)\,du.\end{aligned}\tag{3.20}$$

Remark 3.10. Inequality (3.19) implies in particular that if $p>2$ and

$$\int_0^1\left(\alpha_{2,\mathbf{X}}^{-1}(u)\right)^{p/2}Q^p(u)\,du<\infty,\tag{3.21}$$

then

$$\left\|W_1(\mu_n,\mu)\right\|_p \ll \frac{1}{\sqrt{n}}.$$

Remark 3.11. In the m-dependent case, Inequality (3.19) becomes

$$\left\|W_1(\mu_n,\mu)\right\|_p^p \ll \frac{1}{n^{p/2}}\left(\int_0^\infty\sqrt{H(t)}\,dt\right)^p + \frac{1}{n^{p-1}}\|X_0\|_p^p.$$

This inequality seems to be new even in the i.i.d. case. Compared to the usual Rosenthal bound for sums of i.i.d. random variables, the variance term is replaced by the integral involving H, in accordance with the upper bound (3.6).

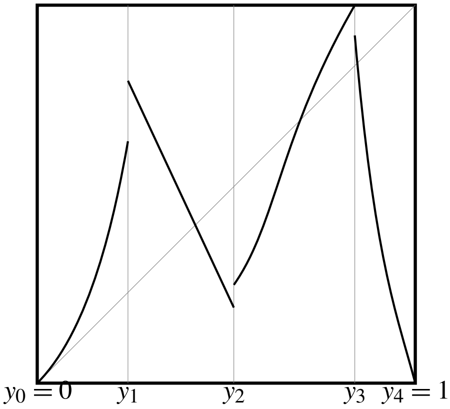

Figure 1. The graph of a GPM map, with $d = 4$.

4. Application to intermittent maps

4.1. Intermittent maps and Markov chains

Let us first recall the definition of the generalized Pomeau–Manneville maps introduced in [7] (see Figure 1 for the graph of such a map).

Definition 4.1. *A map* $\theta : [0, 1] \to [0, 1]$ *is a generalized Pomeau–Manneville map* (*or GPM map*) *of parameter* $\gamma \in (0, 1)$ *if there exist* $0 = y_0 < y_1 < \cdots < y_d = 1$ *such that*, *writing* $I_k = (y_k, y_{k+1})$:

1. *The restriction of* θ *to* I_k *admits a* C^1 *extension* $\theta_{(k)}$ *to* $\overline{I_k}$.
2. *For* $k \geq 1$, $\theta_{(k)}$ *is* C^2 *on* $\overline{I_k}$, *and* $|\theta'_{(k)}| > 1$.
3. $\theta_{(0)}$ *is* C^2 *on* $(0, y_1]$, *with* $\theta'_{(0)}(x) > 1$ *for* $x \in (0, y_1]$, $\theta'_{(0)}(0) = 1$ *and* $\theta''_{(0)}(x) \sim cx^{\gamma-1}$ *when* $x \to 0$, *for some* $c > 0$.
4. θ *is topologically transitive*.

The third condition ensures that 0 is a neutral fixed point of θ, with $\theta(x) = x + c'x^{1+\gamma}(1 + o(1))$ when $x \to 0$. The fourth condition is necessary to avoid situations where there are several absolutely continuous invariant measures, or where the neutral fixed point does not belong to the support of the absolutely continuous invariant measure.

The following well-known example of GPM map with only two branches has been introduced by Liverani, Saussol and Vaienti [20]:

$$\theta(x) = \begin{cases} x\big(1 + 2^\gamma x^\gamma\big), & \text{if } x \in [0, 1/2[, \\ 2x - 1, & \text{if } x \in [1/2, 1]. \end{cases} \tag{4.1}$$

As quoted in [7], a GPM map θ admits a unique invariant absolutely continuous (with respect to the Lebesgue measure) probability ν with density h. Moreover, it is ergodic, has full support, and $x^\gamma h(x)$ is bounded from above and below.

We shall illustrate each result of Section 3 by controlling, on the probability space $([0,1],\nu)$, the quantity $W_1(\tilde{\mu}_n,\mu)$, where

$$\tilde{\mu}_n = \frac{1}{n}\sum_{k=1}^{n}\delta_{g\circ\theta^k}, \tag{4.2}$$

θ is a GPM map, g is a monotonic function from $(0,1)$ to $\mathbb{R}$ (which can blow up near 0 or 1), and μ is the distribution of g.

To do this, we go back to the Markov chain associated with θ, as we describe now. Let first K be the Perron–Frobenius operator of θ with respect to ν, defined as follows: for any functions u, v in $\mathbb{L}^2([0,1],\nu)$

$$\nu(u\cdot v\circ\theta) = \nu\big(K(u)\cdot v\big). \tag{4.3}$$

The relation (4.3) states that K is the adjoint operator of the isometry $U: u\mapsto u\circ\theta$ acting on $\mathbb{L}^2([0,1],\nu)$. It is easy to see that the operator K is a transition kernel, and that ν is invariant by K. Let now $\mathbf{Y}=(Y_i)_{i\geq 0}$ be a stationary Markov chain with invariant measure ν and transition kernel K. It is well known (see, for instance, Lemma XI.3 in [18]) that on the probability space $([0,1],\nu)$, the random vector $(\theta,\theta^2,\dots,\theta^n)$ is distributed as $(Y_n, Y_{n-1},\dots,Y_1)$.

Let T be the shift operator from $[0,1]^{\mathbb{Z}}$ to $[0,1]^{\mathbb{Z}}$ defined by $(T(x))_i = x_{i+1}$, and let π_i be the projection from $[0,1]^{\mathbb{Z}}$ to $[0,1]$ defined by $\pi_i(x)=x_i$. By Kolmogorov's extension theorem, there exists a shift-invariant probability $\mathbb{P}$ on $([0,1]^{\mathbb{Z}}, (\mathcal{B}([0,1]))^{\mathbb{Z}})$, such that $\pi=(\pi_i)_{i\geq 0}$ is distributed as $\mathbf{Y}$.

Let then $X_0 = g\circ\pi_0$ and $X_i = X_0\circ T^i = g\circ\pi_i$, and define $\mathcal{F}_0=\sigma(\pi_i, i\leq 0)$. From the above considerations, we infer that the two random variables $W_1(\mu_n,\mu)$ (defined on the probability space $([0,1]^{\mathbb{Z}},\mathbb{P})$) and $W_1(\tilde{\mu}_n,\mu)$ (defined on the probability space $([0,1],\nu)$) have the same distribution. Hence, any information on the distribution of $W_1(\tilde{\mu}_n,\mu)$ can be derived from the distribution of $W_1(\mu_n,\mu)$.

From Proposition 1.17 (and the comments right after) in [7], we know that for any positive integer k, there exist two positive constants C and D such that, for any $n>0$,

$$\frac{D}{n^{(1-\gamma)/\gamma}} \leq \alpha_{k,\pi}(n) \leq \frac{C}{n^{(1-\gamma)/\gamma}}.$$

Since $X_i = g\circ\pi_i$, and since g is monotonic, it follows immediately that

$$\alpha_{k,\mathbf{X}}(n) \leq \alpha_{k,\pi}(n) \leq \frac{C}{n^{(1-\gamma)/\gamma}}. \tag{4.4}$$

This control of the coefficients $\alpha_{k,\mathbf{X}}(n)$ (for $k=1$ or $k=2$) and a control of the tail $\nu(|g|>t)$ are all we need to apply the results of Section 3 to the random variable $W_1(\tilde{\mu}_n,\mu)$.

4.2. Central limit theorem

Let θ be a GPM map of parameter $\gamma \in (0, 1/2)$, with absolutely continuous invariant probability ν. Let $\tilde{\mu}_n$ be defined as in (4.2), where g is a monotonic function from $(0, 1)$ to $\mathbb{R}$. Let then $(X_i)_{i\in\mathbb{Z}}$ be the stationary sequence constructed in Section 4.1, whose dependency coefficients $\alpha_{k,\mathbf{X}}(n)$ satisfy (4.4). Note that $H(t) = \mathbb{P}(|X_0| > t) = \nu(|g| > t)$. From Section 4.1, Proposition 3.1 and Item 3 of Proposition 7.3, we infer that $\sqrt{n}W_1(\tilde{\mu}_n, \mu)$ converges in distribution to the random variable $\int |G(t)|\, dt$, where G is a Gaussian random variable in $\mathbb{L}^1(dt)$ as soon as

$$\int_0^\infty \big(H(t)\big)^{(1-2\gamma)/(2(1-\gamma))}\, dt < \infty. \tag{4.5}$$

As a consequence:

1. If g is positive and nonincreasing on $(0, 1)$, with

$$g(x) \le \frac{C}{x^{(1-2\gamma)/2}|\ln(x)|^b} \qquad \text{near 0, for some } C > 0 \text{ and } b > 1,$$

then (4.5) holds.

2. If g is positive and nondecreasing on $(0, 1)$, with

$$g(x) \le \frac{C}{(1-x)^{(1-2\gamma)/(2-2\gamma)}|\ln(1-x)|^b} \qquad \text{near 1, for some } C > 0 \text{ and } b > 1,$$

then (4.5) holds.

Recall from (1.2) that $W_1(\tilde{\mu}_n, \mu) = \sup_{f\in\Lambda_1} |\tilde{\mu}_n(f) - \mu(f)|$, so that the condition (4.5) allows to control the supremum of $\sqrt{n}(\mu_n(f) - \mu(f))$ over the class Λ_1. Now if we only want a central limit theorem for $\sqrt{n}(\mu_n(f) - \mu(f))$ where f is an element of Λ_1, then it follows from [7] that the condition

$$\int_0^\infty t\big(H(t)\big)^{(1-2\gamma)/(1-\gamma)}\, dt < \infty \tag{4.6}$$

is sufficient. For the two simple examples above, this would give the constraint $b > 1/2$ instead of $b > 1$.

4.3. Moments of order 1 and 2

We continue the example of Section 4.2. Here γ can be any number in $(0, 1)$.

1. If g is positive and nonincreasing on $(0, 1)$, with

$$g(x) \le \frac{C}{x^b} \qquad \text{near 0, for some } C > 0 \text{ and } b \in [0, 1-\gamma),$$

then $Q(u) \le Du^{-b/(1-\gamma)}$ for some $D > 0$. Applying (3.10)–(3.11) and (3.13)–(3.14), the following upper bounds hold.

For $\gamma \in (0, 1/2)$,

$$\mathbb{E}\big(W_1(\tilde{\mu}_n, \mu)\big) \ll \begin{cases} n^{-1/2}, & \text{if } b < (1-2\gamma)/2, \\ n^{-1/2}\ln(n), & \text{if } b = (1-2\gamma)/2, \\ n^{b+\gamma-1}, & \text{if } b > (1-2\gamma)/2, \end{cases}$$

and

$$\big\| W_1(\tilde{\mu}_n, \mu)\big\|_2 \ll \begin{cases} n^{-1/2}, & \text{if } b < (1-2\gamma)/2, \\ n^{-1/2}\ln(n), & \text{if } b = (1-2\gamma)/2, \\ n^{(2b+\gamma-1)/2\gamma}, & \text{if } (1-2\gamma)/2 < b < (1-\gamma)/2. \end{cases}$$

For $\gamma = 1/2$,

$$\mathbb{E}\big(W_1(\tilde{\mu}_n, \mu)\big) \ll \left(\frac{\ln(n)}{n}\right)^{(1-2b)/2} \quad \text{and} \quad \big\| W_1(\tilde{\mu}_n, \mu)\big\|_2 \ll \left(\frac{\ln(n)}{n}\right)^{(1-4b)/2} \quad \text{if } b < 1/4.$$

For $\gamma \in (1/2, 1)$,

$$\mathbb{E}\big(W_1(\tilde{\mu}_n, \mu)\big) \ll n^{(b+\gamma-1)/(2\gamma)} \quad \text{and} \quad \big\| W_1(\tilde{\mu}_n, \mu)\big\|_2 \le C n^{(2b+\gamma-1)/(2\gamma)} \quad \text{if } b < (1-\gamma)/2.$$

2. If g is positive and nondecreasing on $(0, 1)$, with

$$g(x) \le \frac{C}{(1-x)^b} \qquad \text{near 1, for some } C > 0 \text{ and } b \in [0, 1),$$

then $Q(u) \le Du^{-b}$ for some $D > 0$. Applying (3.10)–(3.11) and (3.13)–(3.14), the following upper bounds hold.

For $\gamma \in (0, 1/2)$,

$$\mathbb{E}\big(W_1(\tilde{\mu}_n, \mu)\big) \ll \begin{cases} n^{-1/2}, & \text{if } b < (1-2\gamma)/2(1-\gamma), \\ n^{-1/2}\ln(n), & \text{if } b = (1-2\gamma)/2(1-\gamma), \\ n^{(\gamma-1)(1-b)}, & \text{if } b > (1-2\gamma)/2(1-\gamma), \end{cases}$$

and

$$\big\| W_1(\tilde{\mu}_n, \mu)\big\|_2 \ll \begin{cases} n^{-1/2}, & \text{if } b < (1-2\gamma)/2(1-\gamma), \\ n^{-1/2}\ln(n), & \text{if } b = (1-2\gamma)/2(1-\gamma), \\ n^{(\gamma-1)(1-2b)/2\gamma}, & \text{if } (1-2\gamma)/2(1-\gamma) < b < 1/2. \end{cases}$$

For $\gamma = 1/2$,

$$\mathbb{E}\big(W_1(\tilde{\mu}_n, \mu)\big) \ll \left(\frac{\ln(n)}{n}\right)^{(1-b)/2} \quad \text{and} \quad \big\| W_1(\tilde{\mu}_n, \mu)\big\|_2 \ll \left(\frac{\ln(n)}{n}\right)^{(1-2b)/2} \quad \text{if } b < 1/2.$$

For $\gamma \in (1/2, 1)$,

$$\mathbb{E}\big(W_1(\tilde{\mu}_n, \mu)\big) \ll n^{(\gamma-1)(1-b)/(2\gamma)} \quad \text{and} \quad \big\| W_1(\tilde{\mu}_n, \mu)\big\|_2 \ll n^{(\gamma-1)(1-2b)/(2\gamma)} \quad \text{if } b < 1/2.$$

4.4. Moments of order $p \in (1, 2)$

We continue the example of Section 4.2. Here γ can be any number in $(0, 1)$.

1. Let $p \in (1, 2)$, and let g be positive and non increasing on $(0, 1)$, with

$$g(x) \le \frac{C}{x^b} \qquad \text{near 0, for some } C > 0 \text{ and } b \in \big[0, (1-\gamma)/p\big).$$

Applying Proposition 3.3, the following upper bounds hold.
For $\gamma \in (0, 1/p]$,

$$\big\| W_1(\tilde{\mu}_n, \mu)\big\|_p \ll \begin{cases} n^{(1-p)/p}, & \text{if } b < (1-p\gamma)/p, \\ \big(n^{(1-p)} \ln(n)\big)^{1/p}, & \text{if } b = (1-p\gamma)/p, \\ n^{(pb+\gamma-1)/p\gamma}, & \text{if } b > (1-p\gamma)/p. \end{cases}$$

Moreover, if $b = (1 - p\gamma)/p$, Proposition 5.1 of Section 5.2 gives the upper bound

$$\mathbb{P}\big(W_1(\mu_n, \mu) \ge x\big) \ll \frac{1}{n^{p-1}x^p}. \tag{4.7}$$

For $\gamma \in (1/p, 1)$, $\|W_1(\tilde{\mu}_n, \mu)\|_p \ll n^{(pb+\gamma-1)/p\gamma}$.

2. Let $p \in (1, 2)$, and let g be positive and nondecreasing on $(0, 1)$, with

$$g(x) \le \frac{C}{(1-x)^b} \qquad \text{near 1, for some } C > 0 \text{ and } b \in [0, 1/p).$$

Applying Proposition 3.3, the following upper bounds hold.
For $\gamma \in (0, 1/p]$,

$$\big\| W_1(\tilde{\mu}_n, \mu)\big\|_p \ll \begin{cases} n^{(1-p)/p}, & \text{if } b < (1-p\gamma)/\big(p(1-\gamma)\big), \\ \big(n^{(1-p)} \ln(n)\big)^{1/p}, & \text{if } b = (1-p\gamma)/\big(p(1-\gamma)\big), \\ n^{(\gamma-1)(1-pb)/p\gamma}, & \text{if } b > (1-p\gamma)/\big(p(1-\gamma)\big). \end{cases}$$

Moreover, if $b = (1 - p\gamma)/(p(1 - \gamma))$, Proposition 5.1 of Section 5.2 gives the upper bound (4.7).

For $\gamma \in (1/p, 1)$, $\|W_1(\tilde{\mu}_n, \mu)\|_p \ll n^{(\gamma-1)(1-pb)/p\gamma}$.

Remark 4.2. The upper bound (4.7) is in accordance with a result by Gouëzel [16]. He proved that, if g is exactly of the form $g(x) = x^{-(1-p\gamma)/p}$ and θ is the LSV map defined by (4.1), then for any positive real x,

$$\lim_{n\to\infty} \nu\left(\frac{1}{n^{1/p}}\left|\sum_{k=1}^{n}\big(g \circ \theta^k - \nu(g)\big)\right| > x\right) = \mathbb{P}\big(|Z_p| > x\big),$$

where Z_p is a p-stable random variable such that $\lim_{x\to\infty} x^p \mathbb{P}(|Z_p| > x) = c > 0$.

4.5. Moments of order $p > 2$

We continue the example of Section 4.2. Here, γ can be any number in $(0, 1)$.

1. Let $p > 2$, and let g be positive and nonincreasing on $(0, 1)$, with

$$g(x) \le \frac{C}{x^b} \qquad \text{near } 0\text{, for some } C > 0 \text{ and } b \in \big[0, (1-\gamma)/p\big).$$

Applying Proposition 3.4, the following upper bounds hold.
For $\gamma \in (0, 1/2)$

$$\|W_1(\tilde{\mu}_n, \mu)\|_p \ll \begin{cases} n^{-1/2}, & \text{if } b \le \big(2-\gamma(p+2)\big)/2p, \\ n^{(pb+\gamma-1)/p\gamma}, & \text{if } b > \big(2-\gamma(p+2)\big)/2p. \end{cases}$$

For $\gamma \in [1/2, 1)$, $\|W_1(\tilde{\mu}_n, \mu)\|_p \ll n^{(pb+\gamma-1)/p\gamma}$.

2. Let $p > 2$, and let g be positive and nondecreasing on $(0, 1)$, with

$$g(x) \le \frac{C}{(1-x)^b} \qquad \text{near } 1\text{, for some } C > 0 \text{ and } b \in [0, 1/p).$$

Applying Proposition 3.4, the following upper bounds hold.
For $\gamma \in (0, 1/2)$

$$\|W_1(\tilde{\mu}_n, \mu)\|_p \ll \begin{cases} n^{-1/2}, & \text{if } b \le \big(2-\gamma(p+2)\big)/2p(1-\gamma), \\ n^{(\gamma-1)(1-pb)/p\gamma}, & \text{if } b > \big(2-\gamma(p+2)\big)/2p(1-\gamma). \end{cases}$$

For $\gamma \in [1/2, 1)$, $\|W_1(\tilde{\mu}_n, \mu)\|_p \ll n^{(\gamma-1)(1-pb)/p\gamma}$.

Remark 4.3. In the case where θ is the LSV map defined by (4.1) and g is the identity (which is a particular case of Item 2, $b = 0$, of the example above) all the rates for $\|W_1(\tilde{\mu}_n, \mu)\|_p$ given in Sections 4.3, 4.4 and 4.5 have been obtained in Corollary 4.1 of [8] by using a different approach. Moreover, all the bounds are optimal in that case (see the discussion in Section 4.2 of [8]).

5. Proofs of the moment bounds

5.1. Proof of Proposition 3.2

Starting from (1.4), we immediately see that

$$\mathbb{E}\big(W_1(\mu_n, \mu)\big) \le \int \|F_n(t) - F(t)\|_1 \, dt \quad \text{and} \quad \|W_1(\mu_n, \mu)\|_2 \le \int \|F_n(t) - F(t)\|_2 \, dt. \tag{5.1}$$

Let $B(t) = F(t)(1 - F(t))$, and note first that

$$\|F_n(t) - F(t)\|_1 \le \|\mathbf{1}_{X_0 \le t} - F(t)\|_1 = 2B(t). \tag{5.2}$$

On another hand

$$\big\|F_n(t)-F(t)\big\|_1^2 \le \big\|F_n(t)-F(t)\big\|_2^2 \le \frac{1}{n}\mathrm{Var}(\mathbf{1}_{X_0\le t}) + \frac{2}{n}\sum_{k=1}^{n}\big|\mathrm{Cov}(\mathbf{1}_{X_0\le t},\mathbf{1}_{X_k\le t})\big|. \tag{5.3}$$

Now, the two following upper bounds hold:

$$\big|\mathrm{Cov}(\mathbf{1}_{X_0\le t},\mathbf{1}_{X_k\le t})\big| \le \big\|\mathbb{E}(\mathbf{1}_{X_k\le t}|\mathcal{F}_0)-F(t)\big\|_1 \le \alpha_{1,\mathbf{X}}(k), \tag{5.4}$$

$$\big|\mathrm{Cov}(\mathbf{1}_{X_0\le t},\mathbf{1}_{X_k\le t})\big| \le \mathrm{Var}(\mathbf{1}_{X_0\le t}) = B(t). \tag{5.5}$$

From (5.2), (5.3), (5.4) and (5.5) it follows that

$$\big\|F_n(t)-F(t)\big\|_1 \le 2\sqrt{\min\Big\{\big(B(t)\big)^2, \frac{1}{n}\sum_{k=0}^{n}\min\big\{\alpha_{1,\mathbf{X}}(k), B(t)\big\}\Big\}}$$

and

$$\big\|F_n(t)-F(t)\big\|_2 \le \sqrt{\frac{2}{n}\sum_{k=0}^{n}\min\big\{\alpha_{1,\mathbf{X}}(k), B(t)\big\}}.$$

These two upper bounds combined with (5.1) imply that

$$\begin{aligned}\mathbb{E}\big(W_1(\mu_n,\mu)\big) &\le 2\int\sqrt{\min\Big\{\big(B(t)\big)^2, \frac{1}{n}\sum_{k=0}^{n}\min\big\{\alpha_{1,\mathbf{X}}(k), B(t)\big\}\Big\}}\,dt\\ &\le 4\int_0^\infty\sqrt{\min\Big\{\big(H(t)\big)^2, \frac{S_{\alpha,n}(t)}{n}\Big\}}\,dt\end{aligned}$$

and

$$\big\|W_1(\mu_n,\mu)\big\|_2 \le \sqrt{\frac{2}{n}}\int\sqrt{\sum_{k=0}^{n}\min\big\{\alpha_{1,\mathbf{X}}(k), B(t)\big\}}\,dt \le \frac{2\sqrt{2}}{\sqrt{n}}\int_0^\infty\sqrt{S_{\alpha,n}(t)}\,dt,$$

which are the desired inequalities.

5.2. Proof of Proposition 3.3

For any $n\in\mathbb{N}$, let us introduce the following notations:

$$R_n(u) = \big(\min\big\{q\in\mathbb{N}^* : \alpha_{1,\mathbf{X}}(q)\le u\big\}\wedge n\big)Q(u) \quad\text{and}\quad R_n^{-1}(x) = \inf\big\{u\in[0,1] : R_n(u)\le x\big\}.$$

The proof is based on the following proposition:

Proposition 5.1. *For any positive integer* n, *any* $x > 0$, *and any* $\eta \in [1, 2[$, *the following inequality holds*:

$$\mathbb{P}\big(nW_1(\mu_n, \mu) \geq 6x\big) \leq c_1 \frac{n}{x} \int_0^{R_n^{-1}(x)} Q(u)\,du + c_2 \frac{n}{x^\eta} \int_{R_n^{-1}(x)}^1 R_n^{\eta-1}(u) Q(u)\,du, \tag{5.6}$$

where $c_1 = 36$ *and* $c_2 = 128(2-\eta)^{-2}$.

Before proving the proposition above, let us see how it entails Proposition 3.3. We have

$$\big\|nW_1(\mu_n, \mu)\big\|_p^p = 6^p p \int_0^\infty x^{p-1} \mathbb{P}\big(nW_1(\mu_n, \mu) \geq 6x\big)\,dx.$$

Therefore applying Inequality (5.6) with $\eta \in (p, 2)$ and using the fact that

$$u < R_n^{-1}(x) \quad \Longleftrightarrow \quad x < R_n(u),$$

we get

$$\begin{aligned} \big\|nW_1(\mu_n, \mu)\big\|_p^p &\leq 6^p pnc_1 \int_0^1 Q(u) \int_0^\infty x^{p-2} \mathbf{1}_{x<R_n(u)}\,dx\,du \\ &\quad + 6^p pnc_2 \int_0^1 R_n^{\eta-1}(u) Q(u) \int_0^\infty x^{p-1-\eta} \mathbf{1}_{x \geq R_n(u)}\,dx\,du, \end{aligned}$$

which gives the desired result since $1 < p < \eta < 2$. Hence, it remains to prove Proposition 5.1.

Proof of Proposition 5.1. Let

$$v = R_n^{-1}(x), \qquad M = Q(v). \tag{5.7}$$

Since $\alpha_{2,\mathbf{X}}(0) = 1/2$, it follows that $v \in [0, 1/2]$. Note first that if $v = 1/2$, then by using (1.2) and Markov's inequality, we have

$$\begin{aligned} \mathbb{P}\big(nW_1(\mu_n, \mu) \geq x\big) &\leq \frac{2}{x} \sum_{k=1}^n \mathbb{E}\big(|X_k|\big) \leq \frac{2n}{x} \int_0^1 Q(u)\,du \\ &\leq \frac{4n}{x} \int_0^{1/2} Q(u)\,du = \frac{4n}{x} \int_0^{R_n^{-1}(x)} Q(u)\,du, \end{aligned} \tag{5.8}$$

which then proves the proposition in case where $v = 1/2$.

Therefore, we can assume in the rest of the proof that $v < 1/2$. From now on, we set $g_M(y) = (y \wedge M) \vee (-M)$. For any integer i, let

$$X_i' = g_M(X_i) \quad \text{and} \quad X_i'' = X_i - X_i'. \tag{5.9}$$

Starting from (1.2), we first notice that

$$nW_1(\mu_n,\mu) \le \sup_{f\in\Lambda_1}\sum_{i=1}^n \big(f(X_i')-\mathbb{E}\big(f(X_i')\big)\big)+\sup_{f\in\Lambda_1}\sum_{i=1}^n \big(f(X_i)-f(X_i')-\mathbb{E}\big(f(X_i)-f(X_i')\big)\big).$$

Therefore,

$$nW_1(\mu_n,\mu) \le \sup_{f\in\Lambda_1}\sum_{i=1}^n \big(f(X_i')-\mathbb{E}\big(f(X_i')\big)\big)+\sum_{i=1}^n \big(|X_i''|+\mathbb{E}\big(|X_i''|\big)\big). \tag{5.10}$$

Let now

$$q=\min\big\{k\in\mathbb{N}:\alpha_{1,\mathbf{X}}(k)\le v\big\}\wedge n. \tag{5.11}$$

Since R_n is right continuous, we have $R_n(R_n^{-1}(w))\le w$ for any w, hence

$$qM=R_n(v)=R_n\big(R_n^{-1}(x)\big)\le x. \tag{5.12}$$

Assume first that $q=n$. Bounding $f(X_i')-\mathbb{E}(f(X_i'))$ by $2M$ in (5.10), we obtain

$$nW_1(\mu_n,\mu)\le 2qM+\sum_{k=1}^n\big(|X_k''|+\mathbb{E}\big(|X_k''|\big)\big). \tag{5.13}$$

Taking into account (5.12) this gives

$$\mathbb{P}\big(nW_1(\mu_n,\mu)\ge 6x\big)\le \frac{1}{2x}\sum_{k=1}^n\mathbb{E}\big(|X_k''|\big).$$

Writing $\varphi_M(x)=(|x|-M)_+$, we have

$$\sum_{k=1}^n\mathbb{E}\big(|X_k''|\big)\le\sum_{k=1}^n\mathbb{E}\big(\varphi_M(X_k)\big).$$

But $Q_{\varphi_M(X_k)}\le Q_{|X_k|}\mathbf{1}_{[0,v]}\le Q\mathbf{1}_{[0,v]}$. Consequently,

$$\sum_{k=1}^n\mathbb{E}\big(|X_k''|\big)\le n\int_0^{R_n^{-1}(x)}Q(u)\,du. \tag{5.14}$$

From (5.13) and (5.14), we infer that

$$\mathbb{P}\big(nW_1(\mu_n,\mu)\ge 6x\big)\le \frac{n}{2x}\int_0^{R_n^{-1}(x)}Q(u)\,du, \tag{5.15}$$

which then proves the proposition in case where $q=n$.

From now on, we assume that $q < n$. Therefore $q = \min\{k \in \mathbb{N} : \alpha_{1,\mathbf{X}}(k) \leq v\}$ and then $\alpha_{1,\mathbf{X}}(q) \leq v$. Recall also that since v is assumed to be strictly less than $1/2$ then $q \geq 1$. Starting from (5.10), we notice that

$$\mathbb{P}\big(nW_1(\mu_n,\mu) \geq 6x\big) \leq \mathbb{P}\left(\sup_{f\in\Lambda_1}\sum_{i=1}^{n}\big(f(X_i') - \mathbb{E}\big(f(X_i')\big)\big) \geq 5x\right) + \frac{2}{x}\sum_{k=1}^{n}\mathbb{E}\big(|X_k''|\big).$$

Therefore taking into account (5.14),

$$\begin{aligned}&\mathbb{P}\big(nW_1(\mu_n,\mu) \geq 6x\big)\\ &\quad\leq \mathbb{P}\left(\sup_{f\in\Lambda_1}\sum_{i=1}^{n}\big(f(X_i') - \mathbb{E}\big(f(X_i')\big)\big) \geq 5x\right) + \frac{2n}{x}\int_0^{R_n^{-1}(x)} Q(u)\,du.\end{aligned} \tag{5.16}$$

To control the first term on the right-hand side, we first notice that

$$\begin{aligned}\sup_{f\in\Lambda_1}\sum_{i=1}^{n}\big(f(X_i') - \mathbb{E}\big(f(X_i')\big)\big) &\leq \sup_{f\in\Lambda_1}\sum_{i=1}^{[n/q]q}\big(f(X_i') - \mathbb{E}\big(f(X_i')\big)\big) + 2\big(n - [n/q]q\big)M\\ &\leq \int_{\mathbb{R}}\left|\sum_{i=1}^{[n/q]q}\big(\mathbf{1}_{X_i'\leq t} - \mathbb{E}(\mathbf{1}_{X_i'\leq t})\big)\right|dt + 2qM.\end{aligned}$$

Using (5.12), it follows that

$$\mathbb{P}\left(\sup_{f\in\Lambda_1}\sum_{i=1}^{n} f(X_i') - \mathbb{E}\big(f(X_i')\big) \geq 5x\right) \leq \mathbb{P}\left(\int_{\mathbb{R}}\left|\sum_{i=1}^{[n/q]q}\big(\mathbf{1}_{X_i'\leq t} - \mathbb{E}(\mathbf{1}_{X_i'\leq t})\big)\right|dt \geq 3x\right).$$

For any integer i, define

$$U_i(t) = \sum_{k=(i-1)q+1}^{iq}\big(\mathbf{1}_{X_k'\leq t} - \mathbb{E}(\mathbf{1}_{X_k'\leq t})\big).$$

Consider now the σ-algebras $\mathcal{G}_i = \mathcal{F}_{iq}$ and define the variables $\tilde{U}_i(t)$ as follows: $\tilde{U}_{2i-1}(t) = U_{2i-1}(t) - \mathbb{E}(U_{2i-1}(t)|\mathcal{G}_{2(i-1)-1})$ and $\tilde{U}_{2i}(t) = U_{2i}(t) - \mathbb{E}(U_{2i}(t)|\mathcal{G}_{2(i-1)})$. Substituting $\tilde{U}_i(t)$ to $U_i(t)$, we obtain the inequality

$$\begin{aligned}&\left|\sum_{i=1}^{[n/q]q}\big(\mathbf{1}_{X_i'\leq t} - \mathbb{E}(\mathbf{1}_{X_i'\leq t})\big)\right|\\ &\quad= \left|\sum_{i=1}^{[n/q]} U_i(t)\right|\\ &\quad\leq \max_{2\leq 2j\leq [n/q]}\left|\sum_{i=1}^{j}\tilde{U}_{2i}(t)\right| + \max_{1\leq 2j-1\leq [n/q]}\left|\sum_{i=1}^{j}\tilde{U}_{2i-1}(t)\right| + \sum_{i=1}^{[n/q]}\big|U_i(t) - \tilde{U}_i(t)\big|.\end{aligned} \tag{5.17}$$

Therefore

$$\mathbb{P}\left(\sup_{f\in\Lambda_1}\sum_{i=1}^{n}\big(f\big(X_i'\big)-\mathbb{E}\big(f\big(X_i'\big)\big)\big)\geq 5x\right)\leq I_1(n)+I_2(n)+I_3(n), \tag{5.18}$$

where

$$I_1(n)=\mathbb{P}\left(\int_{\mathbb{R}}\sum_{i=1}^{[n/q]}\big|U_i(t)-\tilde{U}_i(t)\big|\,dt\geq x\right),$$

$$I_2(n)=\mathbb{P}\left(\int_{\mathbb{R}}\max_{2\leq 2j\leq[n/q]}\left|\sum_{i=1}^{j}\tilde{U}_{2i}(t)\right|dt\geq x\right),$$

$$I_3(n)=\mathbb{P}\left(\int_{\mathbb{R}}\max_{1\leq 2j-1\leq[n/q]}\left|\sum_{i=1}^{j}\tilde{U}_{2i-1}(t)\right|dt\geq x\right).$$

Using Markov's inequality and stationarity, we get

$$I_1(n)\leq\frac{n}{x}\int_{\mathbb{R}}\mathbb{E}\big|\mathbb{E}(\mathbf{1}_{X_1'\leq t}|\mathcal{F}_{-q})-\mathbb{E}(\mathbf{1}_{X_1'\leq t})\big|\,dt=\frac{n}{x}\int_{-M}^{M}\mathbb{E}\big|\mathbb{E}(\mathbf{1}_{X_1'\leq t}|\mathcal{F}_{-q})-\mathbb{E}(\mathbf{1}_{X_1'\leq t})\big|\,dt.$$

But,

$$\sup_{t\in\mathbb{R}}\big\|\mathbb{E}(\mathbf{1}_{X_1'\leq t}|\mathcal{F}_{-q})-\mathbb{E}(\mathbf{1}_{X_1'\leq t})\big\|_1=\sup_{t\in\mathbb{R}}\big\|\mathbb{E}(\mathbf{1}_{g_M(X_1)\leq t}|\mathcal{F}_{-q})-\mathbb{E}(\mathbf{1}_{g(X_1)\leq t})\big\|_1\leq\alpha_{1,\mathbf{X}}(q+1),$$

where the inequality comes from the fact that g_M is a nondecreasing function. Therefore,

$$I_1(n)\leq\frac{2n}{x}Q(v)\alpha_{1,\mathbf{X}}(q+1)\leq\frac{2n}{x}vQ(v)\leq\frac{2n}{x}\int_0^v Q(u)\,du. \tag{5.19}$$

We handle now the term $I_2(n)$ in the decomposition (5.18). Using again Markov's inequality, we get

$$I_2(n)\leq\frac{1}{x^2}\left(\int_{-Q(v)}^{Q(v)}\left\|\max_{2\leq 2j\leq[n/q]}\left|\sum_{i=1}^{j}\tilde{U}_{2i}(t)\right|\right\|_2 dt\right)^2.$$

By Doob's maximal inequality,

$$\left\|\max_{2\leq 2j\leq[n/q]}\left|\sum_{i=1}^{j}\tilde{U}_{2i}(t)\right|\right\|_2^2\leq 2\sum_{i=1}^{1/2[n/q]}\big\|\tilde{U}_{2i}(t)\big\|_2^2\leq 2\sum_{i=1}^{1/2[n/q]}\big\|U_{2i}(t)\big\|_2^2.$$

Now

$$\begin{aligned}2\sum_{i=1}^{1/2[n/q]}\big\|U_{2i}(t)\big\|_2^2&\leq\frac{n}{q}\left\|\sum_{k=1}^{q}\big(\mathbf{1}_{X_k'\leq t}-\mathbb{E}(\mathbf{1}_{X_k'\leq t})\big)\right\|_2^2\\&\leq 2n\sum_{k=0}^{q-1}\big|\mathbb{E}\big(\big(\mathbf{1}_{g_M(X_0)\leq t}-\mathbb{E}(\mathbf{1}_{g_M(X_0)\leq t})\big)\big(\mathbf{1}_{g_M(X_k)\leq t}-\mathbb{E}(\mathbf{1}_{g_M(X_k)\leq t})\big)\big)\big|.\end{aligned}$$

Note that since g_M is a nondecreasing function,

$$\begin{aligned}&\sup_{t\in\mathbb{R}}\bigl|\mathbb{E}\bigl(\bigl(\mathbf{1}_{g_M(X_0)\le t}-\mathbb{E}(\mathbf{1}_{g_M(X_0)\le t})\bigr)\bigl(\mathbf{1}_{g_M(X_k)\le t}-\mathbb{E}(\mathbf{1}_{g_M(X_k)\le t})\bigr)\bigr)\bigr|\\&\quad\le\sup_{t\in\mathbb{R}}\bigl|\mathbb{E}\bigl(\bigl(\mathbf{1}_{X_0\le t}-\mathbb{E}(\mathbf{1}_{X_0\le t})\bigr)\bigl(\mathbf{1}_{X_k\le t}-\mathbb{E}(\mathbf{1}_{X_k\le t})\bigr)\bigr)\bigr|.\end{aligned}$$

Moreover

$$\sup_{t\in\mathbb{R}}\bigl|\mathbb{E}\bigl(\bigl(\mathbf{1}_{X_0>t}-\mathbb{E}(\mathbf{1}_{X_0>t})\bigr)\bigl(\mathbf{1}_{X_k>t}-\mathbb{E}(\mathbf{1}_{X_k>t})\bigr)\bigr)\bigr|\le\sup_{t\in\mathbb{R}}\bigl\|\mathbb{E}(\mathbf{1}_{X_k>t}|\mathcal{F}_0)-\mathbb{E}(\mathbf{1}_{X_k>t})\bigr\|_1=\alpha_{1,\mathbf{X}}(k).$$

On an other hand, the following bound is also valid

$$\begin{aligned}&\bigl|\mathbb{E}\bigl(\bigl(\mathbf{1}_{g_M(X_0)\le t}-\mathbb{E}(\mathbf{1}_{g_M(X_0)\le t})\bigr)\bigl(\mathbf{1}_{g_M(X_k)\le t}-\mathbb{E}(\mathbf{1}_{g_M(X_k)\le t})\bigr)\bigr)\bigr|\\&\quad\le\operatorname{Var}(\mathbf{1}_{g_M(X_0)\le t})\le\min\bigl\{\mathbb{E}(\mathbf{1}_{g_M(X_0)\le t}),\mathbb{E}(\mathbf{1}_{g_M(X_0)>t})\bigr\}.\end{aligned}$$

So, overall, we get

$$\begin{aligned}I_2(n)&\le\frac{2n}{x^2}\Biggl(\int_0^{Q(v)}\Biggl(\sum_{k=0}^{q-1}\alpha_{1,\mathbf{X}}(k)\wedge\mathbb{P}\bigl(g_M(X_0)>t\bigr)\Biggr)^{1/2}\\&\qquad+\Biggl(\sum_{k=0}^{q-1}\alpha_{1,\mathbf{X}}(k)\wedge\mathbb{P}\bigl(-g_M(X_0)\ge t\bigr)\Biggr)^{1/2}dt\Biggr)^2\\&\le\frac{2n}{x^2}\Biggl(\int_0^{Q(v)}\Biggl(\sum_{k=0}^{q-1}\alpha_{1,\mathbf{X}}(k)\wedge\mathbb{P}\bigl(|X_0|>t\bigr)\Biggr)^{1/2}+\Biggl(\sum_{k=0}^{q-1}\alpha_{1,\mathbf{X}}(k)\wedge\mathbb{P}\bigl(|X_0|\ge t\bigr)\Biggr)^{1/2}dt\Biggr)^2.\end{aligned}$$

We then derive that

$$\begin{aligned}I_2(n)&\le\frac{8n}{x^2}\Biggl(\int_0^{Q(v)}\Biggl(\sum_{k=0}^{q-1}\alpha_{1,\mathbf{X}}(k)\wedge H(t)\Biggr)^{1/2}dt\Biggr)^2\\&=\frac{8n}{x^2}\Biggl(\int_0^{Q(v)}\Biggl(\sum_{k=0}^{q-1}\int_0^{H(t)}\mathbf{1}_{u\le\alpha_{1,\mathbf{X}}(k)}\,du\Biggr)^{1/2}dt\Biggr)^2.\end{aligned}$$

Using the fact that $\sum_{k=0}^{q-1}\mathbf{1}_{u\le\alpha_{1,\mathbf{X}}(k)}=\alpha_{1,\mathbf{X}}^{-1}(u)\wedge q$, we then get

$$\begin{aligned}I_2(n)&\le\frac{8n}{x^2}\Biggl(\int_0^{Q(v)}\Biggl(vq+\int_v^{H(t)}\bigl(\alpha_{1,\mathbf{X}}^{-1}(u)\wedge n\bigr)\,du\Biggr)^{1/2}dt\Biggr)^2\\&\le\frac{16n}{x^2}vq\bigl(Q(v)\bigr)^2+\frac{16n}{x^2}\Biggl(\int_0^{Q(v)}\Biggl(\int_v^{H(t)}\bigl(\alpha_{1,\mathbf{X}}^{-1}(u)\wedge n\bigr)\,du\Biggr)^{1/2}dt\Biggr)^2 \qquad (5.20)\\&\le\frac{16n}{x}\int_0^v Q(u)\,du+\frac{16n}{x^2}\Biggl(\int_0^{Q(v)}\Biggl(\int_v^{H(t)}\bigl(\alpha_{1,\mathbf{X}}^{-1}(u)\wedge n\bigr)\,du\Biggr)^{1/2}dt\Biggr)^2,\end{aligned}$$

where for the last inequality we have used (5.12) and the fact that $vQ(v) \le \int_0^v Q(u)\,du$, since Q is non increasing. To handle the last term on the right-hand side, we proceed as follows. For any η in $[1,2)$, we first note that

$$\begin{aligned}\int_v^{H(t)} \big(\alpha_{1,\mathbf{X}}^{-1}(u)\wedge n\big)\,du &= \int_v^{H(t)} \big(\alpha_{1,\mathbf{X}}^{-1}(u)\wedge n\big) Q^{\eta}(u) Q^{-\eta}(u)\,du \\ &\le \frac{1}{t^{\eta}} \int_v^{H(t)} \big(\alpha_{1,\mathbf{X}}^{-1}(u)\wedge n\big) Q^{\eta}(u)\,du,\end{aligned}$$

where the inequality comes from the fact that $u < H(t) \iff t < Q(u)$, and then $u < H(t)$ implies that $Q^{-\eta}(u) < t^{-\eta}$. Now, since $u > v$ implies that $\alpha_{1,\mathbf{X}}^{-1}(u) \le \alpha_{1,\mathbf{X}}^{-1}(v)$, we get

$$\int_v^{H(t)} \big(\alpha_{1,\mathbf{X}}^{-1}(u)\wedge n\big)\,du \le \frac{1}{t^{\eta}} \big(\alpha_{1,\mathbf{X}}^{-1}(v)\wedge n\big)^{2-\eta} \int_v^1 \big(\alpha_{1,\mathbf{X}}^{-1}(u)\wedge n\big)^{\eta-1} Q^{\eta}(u)\,du.$$

Therefore, since $\eta/2 < 1$,

$$\begin{aligned}&\bigg(\int_0^{Q(v)} \bigg(\int_v^{H(t)} \big(\alpha_{1,\mathbf{X}}^{-1}(u)\wedge n\big)\,du\bigg)^{1/2} dt\bigg)^2 \\ &\quad\le \big(\alpha_{1,\mathbf{X}}^{-1}(v)\wedge n\big)^{2-\eta} \bigg(\int_0^{Q(v)} t^{-\eta/2}\,dt\bigg)^2 \int_v^1 \big(R_n(u)\big)^{\eta-1} Q(u)\,du \\ &\quad\le \frac{4}{(2-\eta)^2} \big(\alpha_{1,\mathbf{X}}^{-1}(v)\wedge n\big)^{2-\eta} Q^{2-\eta}(v) \int_v^1 \big(R_n(u)\big)^{\eta-1} Q(u)\,du.\end{aligned}$$

But, by (5.12), $(\alpha_{1,\mathbf{X}}^{-1}(v)\wedge n)^{2-\gamma} Q^{2-\gamma}(v) = R_n^{2-\gamma}(v) \le x^{2-\gamma}$. Therefore,

$$\begin{aligned}&\frac{1}{x^2}\bigg(\int_0^{Q(v)} \bigg(\int_v^{H(t)} \big(\alpha_{1,\mathbf{X}}^{-1}(u)\wedge n\big)\,du\bigg)^{1/2} dt\bigg)^2 \\ &\quad\le \frac{4}{x^{\eta}(2-\eta)^2} \int_v^1 \big(R_n(u)\big)^{\eta-1} Q(u)\,du,\end{aligned}$$

which combined with (5.20) gives

$$I_2(n) \le \frac{16n}{x} \int_0^v Q(u)\,du + \frac{64n}{x^{\eta}(2-\eta)^2} \int_v^1 \big(R_n(u)\big)^{\eta-1} Q(u)\,du. \tag{5.21}$$

With similar arguments, we get

$$I_3(n) \le \frac{16n}{x} \int_0^v Q(u)\,du + \frac{64n}{x^{\eta}(2-\eta)^2} \int_v^1 \big(R_n(u)\big)^{\eta-1} Q(u)\,du. \tag{5.22}$$

Starting from (5.18) and using the upper bounds (5.19), (5.21) and (5.22), we derive that

$$\mathbb{P}\biggl(\sup_{f\in\Lambda_1}\sum_{i=1}^{n}\bigl(f\bigl(X_i'\bigr)-\mathbb{E}\bigl(f\bigl(X_i'\bigr)\bigr)\bigr)\geq 5x\biggr)$$
$$\leq \frac{34n}{x}\int_0^v Q(u)\,du+\frac{128n}{x^\eta(2-\eta)^2}\int_v^1\bigl(R_n(u)\bigr)^{\eta-1}Q(u)\,du,$$

which combined with (5.16) ends the proof of the proposition. □

5.3. Proof of Proposition 3.4

Inequality (3.19) follows from Proposition 5.2 below.

Proposition 5.2. *For any $\eta>2$ and any $\beta\in(\eta-2,\eta)$, there exists a positive constant c such that, for any positive integer n and any $x>0$, the following inequality holds*:

$$\begin{aligned}\mathbb{P}\bigl(nW_1(\mu_n,\mu)\geq x\bigr)\leq c\frac{n^{\eta/2}}{x^\eta}s_{\alpha,n}^\eta+\frac{n}{x^{1+\beta/2}}\int_0^{R_n^{-1}(x)}R_n^{\beta/2}(u)Q(u)\,du\\ +c\frac{n}{x^{1+\eta/2}}\int_{R_n^{-1}(x)}^1 R_n^{\eta/2}(u)Q(u)\,du,\end{aligned}\tag{5.23}$$

where

$$R_n(u)=\bigl(\min\bigl\{q\in\mathbb{N}^*:\alpha_{2,\mathbf{X}}(q)\leq u\bigr\}\wedge n\bigr)Q(u)\quad\textit{and}\quad R_n^{-1}(x)=\inf\bigl\{u\in[0,1]:R_n(u)\leq x\bigr\}.$$

Indeed,

$$\begin{aligned}\bigl\|nW_1(\mu_n,\mu)\bigr\|_p^p&=p\int_0^\infty x^{p-1}\mathbb{P}\bigl(nW_1(\mu_n,\mu)\geq x\bigr)\,dx\\ &\ll n^{p/2}s_{\alpha,n}^p+\int_{n^{1/2}s_{\alpha,n}}^\infty x^{p-1}\mathbb{P}\bigl(nW_1(\mu_n,\mu)\geq x\bigr)\,dx.\end{aligned}\tag{5.24}$$

To handle the second term on the right-hand side, we apply (5.23) with $\eta\in(2p-2,2p)$ and $\beta\in(\eta-2,2p-2)$. This gives

$$\begin{aligned}&\int_{n^{1/2}s_{\alpha,n}}^\infty x^{p-1}\mathbb{P}\bigl(nW_1(\mu_n,\mu)\geq x\bigr)\,dx\\ &\quad\ll n^{\eta/2}s_{\alpha,n}^\eta\int_{n^{1/2}s_{\alpha,n}}^\infty x^{p-\eta-1}\,dx\\ &\qquad+n\int_0^1 R_n^{\beta/2}(u)Q(u)\biggl(\int_0^\infty x^{p-\beta/2-2}\mathbf{1}_{u<R_n^{-1}(x)}\,dx\biggr)du\\ &\qquad+n\int_0^1 R_n^{\eta/2}(u)Q(u)\biggl(\int_0^\infty x^{p-\eta/2-2}\mathbf{1}_{u\geq R_n^{-1}(x)}\,dx\biggr)du.\end{aligned}$$

Since $u < R_n^{-1}(x) \iff x < R_n(u)$, the choice of η and β implies that, for any $p > 2$,

$$\int_{n^{1/2}s_{\alpha,n}}^{\infty} x^{p-1}\mathbb{P}\big(nW_1(\mu_n,\mu) \geq x\big)\,dx \\ \ll n^{p/2}s_{\alpha,n}^p + n\int_0^1 R_n^{p-1}(u)Q(u)\,du,$$

which together with (5.24) give (3.19). Hence, it remains to prove Proposition 5.2.

Proof of Proposition 5.2. We proceed as for the proof of Proposition 5.1 with the following modifications: in the definition of R_n (and then also of v defined in (5.7)), $\alpha_{1,\mathbf{X}}$ is replaced by $\alpha_{2,\mathbf{X}}$, and in the definition of q given in (5.11), $\alpha_{1,\mathbf{X}}$ is also replaced by $\alpha_{2,\mathbf{X}}$. Assuming first that either $v = 1/2$ or $q = n$, we note by following the proof of Proposition 5.1, that the bounds (5.8) and (5.15) are still valid. In addition since $u < R_n^{-1}(x) \iff x < R_n(u)$,

$$\int_0^{R_n^{-1}(x)} Q(u)\,du \leq x^{-\beta/2}\int_0^{R_n^{-1}(x)} R_n^{\beta/2}(u)Q(u)\,du, \tag{5.25}$$

which combined with (5.8) or (5.15) proves the proposition in case where $v = 1/2$ or $q = n$.

From now on, we assume that $v < 1/2$ (so that $q \geq 1$) and $q < n$ (therefore $\alpha_{2,\mathbf{X}}(q) \leq v$). The bound (5.16) is still valid and combined with (5.25) gives

$$\mathbb{P}\big(nW_1(\mu_n,\mu) \geq 6x\big) \leq \mathbb{P}\Bigg(\sup_{f\in\Lambda_1}\sum_{i=1}^n\big(f\big(X_i'\big) - \mathbb{E}\big(f\big(X_i'\big)\big)\big) \geq 5x\Bigg) \\ + \frac{2n}{x^{1+\beta/2}}\int_0^{R_n^{-1}(x)} R_n^{\beta/2}(u)Q(u)\,du. \tag{5.26}$$

As in the proof of Proposition 5.1, the first term on the right-hand side can be handled with the help of the decomposition (5.18). Clearly since $\alpha_{1,\mathbf{X}}(q) \leq \alpha_{2,\mathbf{X}}(q) \leq v$, the term $I_1(n)$ in (5.18) satisfies the inequality (5.19). Therefore taking into account (5.25), it follows that

$$I_1(n) \leq \frac{2n}{x^{1+\beta/2}}\int_0^{R_n^{-1}(x)} R_n^{\beta/2}(u)Q(u)\,du. \tag{5.27}$$

We handle now the term $I_2(n)$ in the decomposition (5.18). Using again Markov's inequality, we get that for any $\eta > 2$,

$$I_2(n) \leq \frac{1}{x^\eta}\Bigg(\int_{-Q(v)}^{Q(v)}\Bigg\|\max_{2\leq 2j\leq[n/q]}\Bigg|\sum_{i=1}^j \tilde{U}_{2i}(t)\Bigg|\Bigg\|_\eta\,dt\Bigg)^\eta.$$

Note that $(\tilde{U}_{2i}(t))_{i\in\mathbb{Z}}$ (resp. $(\tilde{U}_{2i-1}(t))_{i\in\mathbb{Z}}$) is a stationary sequence of martingale differences with respect to the filtration $(\mathcal{G}_{2i})_{i\in\mathbb{Z}}$ (resp. $(\mathcal{G}_{2i-1})_{i\in\mathbb{Z}}$). By using the Rosenthal inequality of

Merlevède and Peligrad [22] for martingales (see their Theorem 6), we get

$$\begin{aligned}&\Big\| \max_{2\le 2j\le [n/q]} \Big|\sum_{i=1}^{j} \tilde U_{2i}(t)\Big| \Big\|_\eta \\ &\quad \ll (n/q)^{1/\eta} \|\tilde U_2(t)\|_\eta + (n/q)^{1/\eta} \left(\sum_{k=1}^{[n/q]} \frac{1}{k^{1+2\delta/\eta}} \left\| \mathbb{E}_0\left(\left(\sum_{i=1}^{k} \tilde U_{2i}(t) \right)^2 \right) \right\|_{\eta/2}^{\delta} \right)^{1/(2\delta)},\end{aligned}$$

where $\delta = \min\{1, (\eta-2)^{-1}\}$. Since $(\tilde U_{2i}(t))_{i\in\mathbb{Z}}$ is a stationary sequence of martingale differences with respect to the filtration $(\mathcal{G}_{2i})_{i\in\mathbb{Z}}$,

$$\mathbb{E}_0\left(\left(\sum_{i=1}^{k} \tilde U_{2i}(t) \right)^2 \right) = \sum_{i=1}^{k} \mathbb{E}_0\big(\tilde U_{2i}^2(t)\big).$$

Moreover $\mathbb{E}_0(\tilde U_{2i}^2(t)) \le \mathbb{E}_0(U_{2i}^2(t))$. Therefore

$$\left\| \mathbb{E}_0\left(\left(\sum_{i=1}^{k} \tilde U_{2i}(t) \right)^2 \right) \right\|_{\eta/2} \le \sum_{i=1}^{k} \big\| \mathbb{E}_0\big(U_{2i}^2(t)\big) - \mathbb{E}\big(U_{2i}^2(t)\big) \big\|_{\eta/2} + \sum_{i=1}^{k} \mathbb{E}\big(U_{2i}^2(t)\big).$$

By stationarity

$$\sum_{i=1}^{k} \mathbb{E}\big(U_{2i}^2(t)\big) = k \|S_q'(t)\|_2^2,$$

where

$$S_q'(t) = \sum_{i=1}^{q} \big(\mathbf{1}_{X_i' \le t} - \mathbb{E}(\mathbf{1}_{X_i' \le t})\big).$$

It follows that

$$\begin{aligned}&\Big\| \max_{2\le 2j\le [n/q]} \Big|\sum_{i=1}^{j} \tilde U_{2i}(t)\Big| \Big\|_\eta \\ &\quad \ll (n/q)^{1/\eta} \|S_q'(t)\|_\eta + (n/q)^{1/2} \|S_q'(t)\|_2 + (n/q)^{1/\eta} \left(\sum_{k=1}^{[n/q]} \frac{1}{k^{1+2\delta/r}} D_{k,q}^{\delta}(t) \right)^{1/(2\delta)},\end{aligned}$$

where

$$D_{k,q}(t) = \sum_{i=1}^{k} \big\| \mathbb{E}_0\big(U_{2i}^2(t)\big) - \mathbb{E}\big(U_{2i}^2(t)\big) \big\|_{\eta/2}.$$

We have

$$
\begin{aligned}
D_{k,q}(t) &\le q^2 \sum_{i=1}^{k} \sup_{j\ge \ell\ge (2i-1)q+1} \sup_{t\in\mathbb{R}} \big\| \mathbb{E}_0\big(\big(\mathbf{1}_{X'_\ell\le t} - \mathbb{E}(\mathbf{1}_{X'_\ell\le t})\big)\big(\mathbf{1}_{X_j\le t} - \mathbb{E}(\mathbf{1}_{X'_j\le t})\big)\big) \\
&\qquad - \mathbb{E}\big(\big(\mathbf{1}_{X'_\ell\le t} - \mathbb{E}(\mathbf{1}_{X'_\ell\le t})\big)\big(\mathbf{1}_{X'_j\le t} - \mathbb{E}(\mathbf{1}_{X'_j\le t})\big)\big)\big\|_{\eta/2} \\
&\le q^2 \sum_{i=1}^{k} \alpha_{2,\mathbf{X}}^{2/\eta}(iq+1),
\end{aligned}
$$

where we have used the fact that g_M is nondecreasing for the second inequality. Since $\beta < \eta$, Hölder's inequality gives

$$
D_{k,q}(t) \ll q^2 k^{(\eta-\beta)/\eta} \left(\sum_{i=1}^{k} i^{\beta/2-1} \alpha_{2,\mathbf{X}}(iq+1) \right)^{2/\eta}.
$$

Therefore, since $\beta > \eta - 2$,

$$
\begin{aligned}
&\left(\int_{-Q(v)}^{Q(v)} \frac{n^{1/\eta}}{q^{1/\eta}} \left(\sum_{k=1}^{[n/q]} \frac{1}{k^{1+2\delta/r}} D_{k,q}(t)^{\delta} \right)^{1/(2\delta)} dt \right)^{\eta} \\
&\quad \ll n q^{\eta-1} Q^{\eta}(v) \left(\sum_{k=1}^{[n/q]} \frac{k^{\delta(\eta-\beta)/\eta}}{k^{1+2\delta/\eta}} \right)^{\eta/(2\delta)} \sum_{i=1}^{[n/q]} i^{\beta/2-1} \alpha_{2,\mathbf{X}}(iq+1) \\
&\quad \ll n q^{\eta-1} Q^{\eta}(v) \sum_{i=1}^{[n/q]} i^{\beta/2-1} \alpha_{2,\mathbf{X}}(iq+1).
\end{aligned}
$$

Note that since $y < \alpha_{2,\mathbf{X}}^{-1}(u) \iff \alpha_{2,\mathbf{X}}(y) > u$ and $\alpha_{2,\mathbf{X}}(q) \le v$,

$$
\begin{aligned}
\sum_{i=1}^{[n/q]} i^{\beta/2-1} \alpha_{2,\mathbf{X}}(iq+1) &= \sum_{i=1}^{[n/q]} i^{\beta/2-1} \int_0^1 \mathbf{1}_{u<\alpha_{2,\mathbf{X}}(iq+1)}\, du \\
&\le \int_0^v \sum_{i=1}^{[n/q]} i^{\beta/2-1} \mathbf{1}_{i\le q^{-1}\alpha_{2,\mathbf{X}}^{-1}(u)}\, du \le q^{-\beta/2} \int_0^v \big(\alpha_{2,\mathbf{X}}^{-1}(u) \wedge n\big)^{\beta/2}\, du.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&\left(\int_{-Q(v)}^{Q(v)} \frac{n^{1/\eta}}{q^{1/\eta}} \left(\sum_{k=1}^{[n/q]} \frac{1}{k^{1+2\delta/r}} D_{k,q}(t)^{\delta} \right)^{1/(2\delta)} dt \right)^{\eta} \\
&\quad \ll n q^{\eta-1-\beta/2} Q^{\eta}(v) \int_0^v \big(\alpha_{2,\mathbf{X}}^{-1}(u) \wedge n\big)^{\beta/2}\, du.
\end{aligned}
$$

Using (5.12) and the fact that $u < v \iff Q(v) < Q(u)$, we infer that

$$\left(\int_{-Q(v)}^{Q(v)} \frac{n^{1/\eta}}{q^{1/\eta}} \left(\sum_{k=1}^{[n/q]} \frac{1}{k^{1+2\delta/r}} D_{k,q}(t)^{\delta}\right)^{1/(2\delta)} dt\right)^{\eta} \ll n x^{\eta-\beta/2-1} \int_0^v R_n^{\beta/2}(u) Q(u)\, du. \quad (5.28)$$

On another hand, since

$$\|S_q'(t)\|_2^2 \le 2q \sum_{k=0}^{q-1} \left|\mathbb{E}\left(\left(\mathbf{1}_{g_M(X_0)\le t} - \mathbb{E}(\mathbf{1}_{g_M(X_0)\le t})\right)\left(\mathbf{1}_{g_M(X_k)\le t} - \mathbb{E}(\mathbf{1}_{g_M(X_k)\le t})\right)\right)\right|,$$

proceeding as to bound $I_2(n)$ in the proof of Proposition 5.1, we infer that

$$\left(\int_{-Q(v)}^{Q(v)} \frac{n^{1/2}}{q^{1/2}} \|S_q'(t)\|_2\, dt\right)^{\eta} \ll n^{\eta/2} s_{\alpha,n}^{\eta}. \quad (5.29)$$

We prove now that

$$\begin{aligned}&\left(\int_{-Q(v)}^{Q(v)} \frac{n^{1/\eta}}{q^{1/\eta}} \|S_q'(t)\|_{\eta}\, dt\right)^{\eta} \\ &\quad \ll n x^{\eta-\beta/2-1} \int_0^v R_n^{\beta/2}(u) Q(u)\, du + n x^{\eta/2-1} \int_v^1 R_n^{\eta/2}(u) Q(u)\, du.\end{aligned} \quad (5.30)$$

With this aim, assume first that we can prove that

$$\int_{-Q(v)}^{Q(v)} \|S_q'(t)\|_{\eta}\, dt \ll q^{1/2} \int_0^{Q(v)} \left(\int_0^{H(t)} \left(\alpha_{2,\mathbf{X}}^{-1}(u) \wedge q\right)^{\eta/2} du\right)^{1/\eta} dt, \quad (5.31)$$

then

$$\left(\int_{-Q(v)}^{Q(v)} \frac{n^{1/\eta}}{q^{1/\eta}} \|S_q'(t)\|_{\eta}\, dt\right)^{\eta} \ll A(n) + B(n),$$

where

$$A(n) = n v q^{\eta-1} Q^{\eta}(v) \quad \text{and}$$
$$B(n) = n q^{\eta/2-1} \left(\int_0^{Q(v)} \left(\int_v^{H(t)} \left(\alpha_{2,\mathbf{X}}^{-1}(u) \wedge q\right)^{\eta/2} du\right)^{1/\eta} dt\right)^{\eta} dx.$$

Using (5.12), the fact that $u < v \iff Q(v) < Q(u)$ and that $u < R_n^{-1}(x) = v \iff x < R_n(u)$, we successively derive

$$A(n) \ll n x^{\eta-1} v Q(v) \ll n x^{\eta-1} \int_0^v Q(u)\, dx \ll n x^{\eta-\beta/2-1} \int_0^v R_n^{\beta/2}(u) Q(u)\, du. \quad (5.32)$$

On the other hand, since $u < H(t) \iff t < Q(u)$, we have

$$B(n) \le nq^{\eta/2-1}\left(\int_0^{Q(v)} \frac{1}{t^{1/2+1/\eta}}\left(\int_v^{H(t)} \left(\alpha_{2,\mathbf{X}}^{-1}(u)\wedge q\right)^{r/2} Q^{\eta/2+1}(u)\,du\right)^{1/\eta} dt\right)^{\eta}$$
$$\ll n\left(qQ(v)\right)^{\eta/2-1}\int_v^1 \left(\alpha_{2,\mathbf{X}}^{-1}(u)\wedge n\right)^{\eta/2} Q^{\eta/2+1}(u)\,du.$$

Using (5.12), it follows that

$$B(n) \ll nx^{\eta/2-1}\int_v^1 R_n^{\eta/2}(u)Q(u)\,du.$$

This last upper bound together with (5.32) show that to prove (5.30) it suffices to prove (5.31). To prove this moment inequality, we use Corollary 2 in [6]. Since, for any $t\in\mathbb{R}$, $|\mathbf{1}_{X_0'\le t} - \mathbb{E}(\mathbf{1}_{X_0'\le t})| \le 1$, this gives

$$\left\|S_q'(t)\right\|_\eta \le \sqrt{2q\eta}\left(\int_0^{\|Y(t)\|_1} \left(\gamma^{-1}(u)\wedge q\right)^{\eta/2} du\right)^{1/\eta},$$

where $Y(t) = \mathbf{1}_{g_M(X_0)\le t} - \mathbb{E}(\mathbf{1}_{g_M(X_0)\le t})$ and

$$\gamma^{-1}(u) = \sum_{k=0}^{\infty} \mathbf{1}_{u\le\gamma(k)} \qquad \text{with } \gamma(k) = \left\|\mathbb{E}_0\left(\mathbf{1}_{g_M(X_k)\le t} - \mathbb{E}(\mathbf{1}_{g_M(X_k)\le t})\right)\right\|_1.$$

Since g_M is nondecreasing $\gamma(k)\le\alpha_{1,\mathbf{X}}(k)\le\alpha_{2,\mathbf{X}}(k)$ in such a way that $\gamma^{-1}(u)\le\alpha_{2,\mathbf{X}}^{-1}(u)$. Moreover, for any $t\in\mathbb{R}$,

$$\left\|Y(t)\right\|_1 = 2\mathbb{P}\left(g_M(X_0)\le t\right)\mathbb{P}\left(g_M(X_0)>t\right) \le 2F(t)\left(1-F(t)\right),$$

where F is the distribution function of X_0. All these considerations end the proof of (5.31).

So, overall, we get

$$I_2(n) \ll x^{-\eta}n^{\eta/2}s_{\alpha,n}^{\eta} + nx^{-\beta/2-1}\int_0^v R_n^{\beta/2}(u)Q(u)\,du + nx^{-\eta/2-1}\int_v^1 R_n^{\eta/2}(u)Q(u)\,du.$$

With similar arguments, we can prove that

$$I_3(n) \ll x^{-\eta}n^{\eta/2}s_{\alpha,n}^{\eta} + nx^{-\beta/2-1}\int_0^v R_n^{\beta/2}(u)Q(u)\,du + nx^{-\eta/2-1}\int_v^1 R_n^{\eta/2}(u)Q(u)\,du.$$

Therefore starting from (5.18) and taking into account (5.27), (5.33) and (5.33), it follows that

$$\mathbb{P}\left(\sup_{f\in\Lambda_1}\sum_{i=1}^n\left(f\left(X_i'\right)-\mathbb{E}\left(f\left(X_i'\right)\right)\right)\ge 5x\right)$$
$$\ll x^{-\eta}n^{\eta/2}s_{\alpha,n}^{\eta} + nx^{-\beta/2-1}\int_0^v R_n^{\beta/2}(u)Q(u)\,du + nx^{-\eta/2-1}\int_v^1 R_n^{\eta/2}(u)Q(u)\,du,$$

which combined with (5.26) ends the proof of Proposition 5.2. □

6. Weak convergence of partial sums in $\mathbb{L}^1(m)$

Let $(S, \mathcal{S}, m)$ be a σ-finite measure space such that $\mathbb{L}^1(S, \mathcal{S}, m)$ is separable. In what follows, we shall denote by $\mathbb{L}^1(m)$ the space $\mathbb{L}^1(S, \mathcal{S}, m)$.

We use the notations of Section 2. Let $Y_0 = \{Y_0(t), t \in S\}$ be a random variable with values in $\mathbb{L}^1(m)$, such that

$$\int \|Y_0(t)\|_1 m(dt) < \infty \quad \text{and} \quad \int Y_0(t) m(dt) = 0.$$

Define the stationary sequence $\mathbf{Y} = (Y_i)_{i \in \mathbb{Z}}$ by $Y_i = Y_0 \circ T^i$, and let

$$S_n = \sum_{k=1}^{n} Y_k.$$

6.1. Previous results

If $\mathbf{Y}$ is a sequence of i.i.d. random variables, Jain [19] proved that $n^{-1/2} S_n$ satisfies the CLT (i.e., converges in distribution to an $\mathbb{L}^1(m)$-valued Gaussian random variable) if and only if

$$\int \|Y_0(t)\|_2 m(dt) < \infty. \tag{6.1}$$

Using a general result by de Acosta, Araujo and Giné [4], Dédé [5] proved that the CLT remains valid under (6.1) for stationary and ergodic martingale differences (meaning that Y_0 is $\mathcal{F}_0$-measurable and $\mathbb{E}(Y_1|\mathcal{F}_0) = 0$ almost surely). Starting from a martingale approximation, she proved then that, if $\mathbf{Y}$ is ergodic, the CLT holds as soon as (6.1) holds and

$$\sum_{k \in \mathbb{Z}} \int \|P_0(Y_k(t))\|_2 m(dt) < \infty, \tag{6.2}$$

where $P_0(Y_k(t)) = \mathbb{E}(Y_k(t)|\mathcal{F}_0) - \mathbb{E}(Y_k(t)|\mathcal{F}_{-1})$.

In a recent paper, Cuny [3] has given many new results concerning the behavior of partial sums of dependent sequences in Banach spaces of cotype 2. Among these results, he showed that, if $\mathbf{Y}$ is ergodic, Y_0 is $\mathcal{F}_0$-measurable, (6.1) holds and

$$\sum_{n>0} \int \frac{\|\mathbb{E}(S_n|\mathcal{F}_0)\|_2}{n^{3/2}} m(dt) < \infty, \tag{6.3}$$

then the CLT and the weak invariance principle (WIP) hold. By WIP, we mean that the partial sum process $\{n^{-1/2} S_{[nt]}, t \in [0, 1]\}$ converges in distribution to an $\mathbb{L}^1(m)$-valued Wiener process in the space $D_{\mathbb{L}^1(m)}([0, 1])$ of $\mathbb{L}^1(m)$-valued càdlàg functions equipped with the uniform metric. As usual, an $\mathbb{L}^1(m)$-valued Wiener process with covariance operator Λ is a centered Gaussian process $W = \{W_t, t \in [0,1]\}$ such that $\mathbb{E}(\|W_t\|^2_{\mathbb{L}^1(m)}) < \infty$ for all $t \in [0,1]$ and, for all f, g in $\mathbb{L}^\infty(m)$,

$$\mathrm{Cov}\left(\int f(u)W_t(u)m(du), \int g(u)W_s(u)m(du)\right) = \min\{s,t\}\Lambda(f,g)$$

(as usual, we identify a function f in $\mathbb{L}^\infty(m)$ with an element of the dual of $\mathbb{L}^1(m)$, and Λ is then a bounded symmetric bilinear operator from $\mathbb{L}^\infty(m) \times \mathbb{L}^\infty(m)$ to $\mathbb{R}$).

Note that Cuny [3] also proved that the WIP holds under (6.2), and that the almost sure invariance principle with rate $o(\sqrt{n \ln\ln n})$ is true if either (6.2) or (6.3) holds.

The condition (6.2) is the $\mathbb{L}^1(m)$ version of Hannan's criterion [17], and the condition (6.3) is the $\mathbb{L}^1(m)$ version of Maxwell–Woodroofe's criterion [21]. If Y_0 is $\mathcal{F}_0$-measurable, both criteria hold as soon as

$$\sum_{k=0}^{\infty} \frac{1}{k+1} \int \left\|\mathbb{E}\big(Y_k(t)|\mathcal{F}_0\big)\right\|_2 m(dt) < \infty. \tag{6.4}$$

As shown in [3], if either (6.2) or (6.3) holds, there exists a stationary and ergodic sequence of martingale differences $(D_i)_{i\in\mathbb{Z}}$ with values in $\mathbb{L}^1(m)$, such that, setting $M_n = \sum_{k=1}^n D_k$,

$$\left\| \max_{1\le k\le n} \int \left|S_k(t) - M_k(t)\right| m(dt) \right\|_2 = o(\sqrt{n}).$$

In the next subsections, we shall rather look for a martingale approximation in $\mathbb{L}^1$, in the spirit of Gordin [15]. Our criterion will not be directly comparable to either (6.2) or (6.3), but its application to the empirical distribution function of α-dependent sequences will lead to weaker conditions (see Section 7 for a deeper discussion).

6.2. A central limit theorem in $\mathbb{L}^1(m)$ for non-adapted sequences

In this section, we give an extension of Gordin's criterion [15] for the central limit theorem to $\mathbb{L}^1(m)$-valued random variables. For the sake of readability, we denote by $\mathbb{E}_i(\cdot)$ the conditional expectation with respect to $\mathcal{F}_i = T^{-i}(\mathcal{F}_0)$.

Theorem 6.1. *Assume that, for m almost every t, the series*

$$U(t) = \sum_{k=1}^{\infty} \mathbb{E}_0\big(Y_k(t)\big) \quad \textit{and} \quad V(t) = -\sum_{k=-\infty}^{0} \big(Y_k(t) - \mathbb{E}_0\big(Y_k(t)\big)\big) \tag{6.5}$$

converge in probability, and let

$$D_0(t) = \sum_{k\in\mathbb{Z}} \big(\mathbb{E}_0\big(Y_k(t)\big) - \mathbb{E}_{-1}\big(Y_k(t)\big)\big) \quad \textit{and} \quad M_n(t) = \sum_{k=1}^{n} D_0(t) \circ T^k.$$

If

$$\int \|U(t)+V(t)\|_1 m(dt) < \infty, \tag{6.6}$$

then

$$\lim_{n\to\infty} \int \left\| \frac{S_n(t)}{\sqrt{n}} - \frac{M_n(t)}{\sqrt{n}} \right\|_1 m(dt) = 0. \tag{6.7}$$

If moreover, for m almost every t,

$$C(t) = \liminf_{n\to\infty} \frac{1}{\sqrt{n}} \mathbb{E}\big(|S_n(t)|\big) < \infty \quad \text{and} \quad \int C(t) m(dt) < \infty, \tag{6.8}$$

then

$$\int \|D_0(t)\|_2 m(dt) < \infty, \tag{6.9}$$

and for any $(s_1, \dots, s_d)$ in $[0,1]^d$, the random vector $n^{-1/2}(S_{[ns_1]}, \dots, S_{[ns_d]})^t$ converges in distribution in $(\mathbb{L}^1(m))^d$ to the Gaussian random vector $(W_{s_1}, \dots, W_{s_d})$, where W is the $\mathbb{L}_1(m)$-valued Wiener process W with covariance operator Λ defined by: for any f, g in $\mathbb{L}_\infty(m)$,

$$\Lambda(f,g) = \mathbb{E}\left(\int\int f(t) g(s) D_0(t) D_0(s) m(dt) m(ds) \right). \tag{6.10}$$

Proof of Theorem 6.1. We first state the following intermediate result:

Proposition 6.1. *Assume that, for m almost every t,*

$$Y_0(t) = D_0(t) + Z(t) - Z(t) \circ T, \tag{6.11}$$

where $D_0(t)$ is an $\mathcal{F}_0$-measurable integrable random variable such that $\mathbb{E}(D_0(t)|\mathcal{F}_{-1}) = 0$ almost surely. Let then $M_n(t) = \sum_{k=1}^n D_0(t) \circ T^k$. If

$$\int \|Z(t)\|_1 m(dt) < \infty, \tag{6.12}$$

then (6.7) holds. If moreover (6.9) holds, then the conclusion of Theorem 6.1 holds.

Before proving Proposition 6.1, let us continue the proof of Theorem 6.1. Note first that, if (6.5) is satisfied, then (6.11) holds, with

$$D_0(t) = \sum_{k\in\mathbb{Z}} \big(\mathbb{E}_0\big(Y_k(t)\big) - \mathbb{E}_{-1}\big(Y_k(t)\big)\big) \quad \text{and}$$

$$Z(t) = \sum_{k=0}^{\infty} \mathbb{E}_{-1}\big(Y_k(t)\big) - \sum_{k=-\infty}^{-1} \big(Y_k(t) - \mathbb{E}_{-1}\big(Y_k(t)\big)\big).$$

Now, if $Z(t)$ is defined as above, the conditions (6.6) and (6.12) are the same. Hence, it follows from Proposition 6.1 that (6.7) holds as soon as (6.12) is satisfied. The second part of Theorem 6.1 will follow from Proposition 6.1 if we prove that (6.8) implies (6.9). By (6.6) it follows that

$$\lim_{n\to\infty} \frac{\|Z(t)\|_1}{\sqrt{n}} = 0 \qquad \text{for } m\text{-almost every } t. \tag{6.13}$$

Since $S_n(t) = M_n(t) + Z(t) - Z(t) \circ T^n$, we infer from (6.13) that, for m almost every t,

$$\liminf_{n\to\infty} \frac{\|M_n(t)\|_1}{\sqrt{n}} = \liminf_{n\to\infty} \frac{\|S_n(t)\|_1}{\sqrt{n}}. \tag{6.14}$$

From (6.14) and (6.8), it follows that, for m almost every t,

$$C(t) = \liminf_{n\to\infty} \frac{\|M_n(t)\|_1}{\sqrt{n}} < \infty.$$

Now, applying Theorem 1 and Remark 1.1 in Esseen and Janson [13], we deduce that, for m almost every t,

$$\|D_0(t)\|_2 = \sqrt{\frac{\pi}{2}}C(t),$$

so that (6.8) implies (6.9). This completes the proof of Theorem 6.1. □

Proof of Proposition 6.1. Since $S_n(t) = M_n(t) + Z(t) - Z(t) \circ T^n$, it follows that

$$\int \left\| \frac{S_n(t)}{\sqrt{n}} - \frac{M_n(t)}{\sqrt{n}} \right\|_1 m(dt) \le \frac{2}{\sqrt{n}} \int \|Z(t)\|_1 m(dt),$$

and (6.7) follows from (6.12).

Now, let d be a positive integer, and let f be a separately Lipschitz function from $(\mathbb{L}^1(m))^d$ to $\mathbb{R}$. This means that there exists non-negative constants $c_1, \dots, c_d$ such that

$$\left| f(x_1, \dots, x_d) - f(y_1, \dots, y_d) \right| \le \sum_{i=1}^d c_i \int \left| x_i(t) - y_i(t) \right| m(dt).$$

For such a f and any $(s_1, \dots, s_d)$ in $[0, 1]^d$, we get that

$$\left| \mathbb{E}\left(f\left(\frac{S_{[ns_1]}}{\sqrt{n}}, \dots, \frac{S_{[ns_d]}}{\sqrt{n}} \right) \right) - \mathbb{E}\left(f\left(\frac{M_{[ns_1]}}{\sqrt{n}}, \dots, \frac{M_{[ns_d]}}{\sqrt{n}} \right) \right) \right|$$

$$\le \sum_{i=1}^d c_i \int \left\| \frac{S_{[ns_i]}(t)}{\sqrt{n}} - \frac{M_{[ns_i]}(t)}{\sqrt{n}} \right\|_1 m(dt),$$

and it follows from (6.7) that

$$\lim_{n\to\infty} \left| \mathbb{E}\left(f\left(\frac{S_{[ns_1]}}{\sqrt{n}}, \dots, \frac{S_{[ns_d]}}{\sqrt{n}} \right) \right) - \mathbb{E}\left(f\left(\frac{M_{[ns_1]}}{\sqrt{n}}, \dots, \frac{M_{[ns_d]}}{\sqrt{n}} \right) \right) \right| = 0. \tag{6.15}$$

Now, when (6.9) holds, Cuny [3] proved that the process $\{n^{-1/2}M_{[nt]}, t \in [0,1]\}$ converges in distribution in the space $D_{\mathbb{L}_1(m)}([0,1])$ to an $\mathbb{L}_1(m)$-valued Wiener process W, with covariance operator Λ given by (6.10). Together with (6.15), this completes the proof of Proposition 6.1. □

6.3. An invariance principle in $\mathbb{L}^1(m)$ for adapted sequences

In this subsection, we assume that the random variable Y_0 is $\mathcal{F}_0$-measurable.

Theorem 6.2. *Assume that, for m-almost every t, the series $U(t)$ defined in* (6.5) *converges in probability. Assume also that, for m-almost every t, the series*

$$\sum_{k=0}^{n} Y_0(t)\mathbb{E}_0\big(Y_k(t)\big) \tag{6.16}$$

converge in $\mathbb{L}^1$, and let

$$L(t) = \sup_{n\geq 0}\left\| \sum_{k=0}^{n} Y_0(t)\mathbb{E}_0\big(Y_k(t)\big)\right\|_1. \tag{6.17}$$

If moreover $\int \|U(t)\|_1 m(dt) < \infty$ and

$$\int \sqrt{L(t)}m(dt) < \infty, \tag{6.18}$$

then $\{n^{-1/2}S_{[nt]}, t \in [0,1]\}$ converges in distribution in the space $D_{\mathbb{L}_1(m)}([0,1])$ to an $\mathbb{L}_1(m)$-valued Wiener process W, with covariance operator Λ defined by (6.10).

As an immediate consequence of Theorem 6.2, the following corollary holds.

Corollary 6.1. *Assume that*

$$\int \sqrt{\sum_{k\geq 0}\big\|\max\{1, |Y_0(t)|\}\big|\mathbb{E}_0\big(Y_k(t)\big)\big|\big\|_1} m(dt) < \infty. \tag{6.19}$$

Then the conclusion of Theorem 6.2 holds.

Remark 6.3. Under the assumptions of Theorem 6.2, we shall prove that the sequence

$$T_n = \frac{1}{n}\left(\max_{1\leq k\leq n}\int \big|S_k(t)\big|m(dt)\right)^2 \tag{6.20}$$

is uniformly integrable (see Lemma 6.1 below). By standard arguments, this implies the following extension of Theorem 6.2: let ψ be any continuous function from $(D_{\mathbb{L}_1(m)}([0,1]), \|\cdot\|_\infty)$ to

$\mathbb{R}$ such that $|\psi(x)| \leq C(1 + \|x\|_\infty^2)$ for some positive constant C. Then

$$\lim_{n\to\infty} \mathbb{E}\left(\psi\left(\frac{S_n}{\sqrt{n}}\right)\right) = \mathbb{E}\big(\psi(W)\big).$$

In particular

$$\lim_{n\to\infty} \mathbb{E}(T_n) = \mathbb{E}\left(\left(\max_{t\in[0,1]} \int \big|W_t(s)\big| m(ds)\right)^2\right).$$

Proof of Theorem 6.2. Note first that, in this adapted case, all the conditions of Theorem 6.1 are satisfied. Indeed, since for m almost every t the series (6.16) converge in $\mathbb{L}^1$, it follows that the series $\sum_{k=0}^\infty \mathrm{Cov}(Y_0(t), Y_k(t))$ converge, and then

$$\lim_{n\to\infty} \frac{\|S_n(t)\|_2^2}{n} = \mathrm{Var}\big(Y_0(t)\big) + 2\sum_{k=1}^\infty \mathrm{Cov}\big(Y_0(t), Y_k(t)\big).$$

Now, by definition of $L(t)$,

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

Hence, the condition (6.8) follows from (6.18) and the fact that

$$C(t) \leq \sqrt{\mathrm{Var}\big(Y_0(t)\big) + 2\sum_{k=1}^\infty \mathrm{Cov}\big(Y_0(t), Y_k(t)\big)} \leq \sqrt{2L(t)}.$$

So, the conclusion of Theorem 6.1 holds with the covariance function defined by (6.10).

As usual it remains to prove the tightness, which reduces through Ascoli's theorem to: for any $\varepsilon > 0$,

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \frac{1}{\delta} \mathbb{P}\left(\max_{1\leq k\leq[n\delta]} \int \big|S_k(t)\big| m(dt) > \sqrt{n}\varepsilon\right) = 0.$$

But this follows straightforwardly from Lemma 6.1 below by applying Markov inequality at order 2. The proof of Theorem 6.2 is complete. □

Lemma 6.1. *Assume that, for m-almost every t, the series defined in* (6.16) *converges in* $\mathbb{L}^1$. *Assume moreover that the function L defined in* (6.17) *satisfies* (6.18). *Then the sequence $(T_n)_{n\geq 1}$ defined in* (6.20) *is uniformly integrable.*

Proof. We first note that, for any positive random variable V,

$$\begin{aligned}\mathbb{E}\left(\left(\max_{1\leq k\leq n} \int \big|S_k(t)\big| m(dt)\right)^2 V\right) &\leq \mathbb{E}\left(\left(\int \sqrt{V} \max_{1\leq k\leq n} \big|S_k(t)\big| m(dt)\right)^2\right) \\ &\leq \left(\int \left\|\sqrt{V} \max_{1\leq k\leq n} \big|S_k(t)\big|\right\|_2 m(dt)\right)^2.\end{aligned}$$

Taking $V = \mathbf{1}_{T_n > M}$, we obtain that

$$\mathbb{E}(T_n \mathbf{1}_{T_n > M}) \le \frac{1}{n}\Big(\int \Big\| \Big(\max_{1\le k\le n} \big|S_k(t)\big|\Big)\mathbf{1}_{T_n>M}\Big\|_2 m(dt)\Big)^2. \tag{6.21}$$

Applying Inequality (3.12) in [9] with $\lambda = 0$, we get that

$$\Big\| \max_{1\le k\le n} \big|S_k(t)\big|\Big\|_2^2 \le 16 \sum_{k=1}^{n} \Big\| Y_k(t) \sum_{i=k}^{n} \mathbb{E}_k\big(Y_i(t)\big)\Big\|_1 \le 16nL(t). \tag{6.22}$$

Using (6.18), (6.21), (6.22) and the reverse Fatou lemma, we infer that

$$\lim_{M\to\infty} \limsup_{n\to\infty} \mathbb{E}(T_n \mathbf{1}_{T_n>M}) = 0$$

as soon as, for m-almost every t,

$$\lim_{M\to\infty} \limsup_{n\to\infty} \Big\| \frac{1}{\sqrt{n}}\Big(\max_{1\le k\le n}\big|S_k(t)\big|\Big)\mathbf{1}_{T_n>M}\Big\|_2 = 0. \tag{6.23}$$

It remains to prove (6.23). In fact this follows quite easily from Proposition 1 in [9]. Indeed, since for m-almost every t, the series defined in (6.16) converges in $\mathbb{L}^1$, it follows from this proposition that the sequence

$$\frac{1}{n}\Big(\max_{1\le k\le n}\big|S_k(t)\big|\Big)^2$$

is uniformly integrable for m-almost every t. Hence, (6.23) holds as soon as

$$\lim_{M\to\infty} \limsup_{n\to\infty} \mathbb{P}(T_n > M) = 0. \tag{6.24}$$

Now, applying (6.21) and (6.22),

$$\mathbb{P}(T_n > M) \le \frac{\mathbb{E}(T_n)}{M} \le \frac{16}{M}\Big(\int \sqrt{L(t)} m(dt)\Big)^2,$$

and (6.24) follows. This completes the proof of Lemma 6.1. □

6.4. An invariance principle in $\mathbb{L}_1(m)$ for the empirical distribution function

In this subsection, $S = \mathbb{R}$, and m is a σ-finite measure on $\mathbb{R}$ equipped with the Borel σ-field. As in Section 2, let X_0 be an $\mathcal{F}_0$-measurable and integrable real-valued random variable with distribution function F. Define the stationary sequence $\mathbf{X} = (X_i)_{i\in\mathbb{Z}}$ by $X_i = X_0 \circ T^i$, and denote by $F_{X_k|\mathcal{F}_0}$ the conditional distribution function of X_k given $\mathcal{F}_0$.

The random variable Y_k is then defined by $Y_k(t) = \mathbf{1}_{X_k \le t} - F(t)$, in such a way that

$$S_n = \sum_{k=1}^{n} Y_k = n(F_n - F),$$

where F_n is the empirical distribution function of $\{X_1, \dots, X_n\}$. Note that Y_0 is a $\mathbb{L}^1(m)$-valued random variable as soon as $\mathbb{E}(|X_0|) < \infty$.

Theorem 6.4. *Assume that*

$$\int \sqrt{\sum_{k=0}^{\infty} \left\| F_{X_k|\mathcal{F}_0}(t) - F(t) \right\|_1} m(dt) < \infty. \tag{6.25}$$

Then $\{n^{-1/2} S_{[ns]}, s \in [0,1]\}$ *converges in distribution in the space* $D_{\mathbb{L}_1(m)}([0,1])$ *to an* $\mathbb{L}_1(m)$-*valued Wiener process* W. *Moreover the explicit form of the covariance operator of* W *is obtained via equation* (6.10) *of Theorem* 6.1 *by taking* $Y_k(t) = \mathbf{1}_{X_k \le t} - F(t)$.

When applied to α-dependent sequences as defined in Section 2, Theorem 6.4 yields the following result.

Proposition 6.2. *Let* $B(t) = F(t)(1 - F(t))$. *The condition*

$$\int \sqrt{\sum_{k=0}^{\infty} \min\{\alpha_{1,\mathbf{X}}(k), B(t)\}} m(dt) < \infty \tag{6.26}$$

implies the condition (6.25), *and hence the conclusion of Theorem* 6.4. *Moreover, the covariance operator* Λ *of* W *can be expressed as follows: for any* f, g *in* $\mathbb{L}_\infty(m)$,

$$\Lambda(f,g) = \sum_{k \in \mathbb{Z}} \mathbb{E}\left(\int\int f(t) g(s) \left(\mathbf{1}_{X_0 \le t} - F(t)\right)\left(\mathbf{1}_{X_k \le s} - F(s)\right) m(dt) m(ds) \right). \tag{6.27}$$

Proofs of Theorem 6.4 and Proposition 6.2. Theorem 6.4 is a direct consequence of Corollary 6.1 applied to the random variables $Y_k(t) = \mathbf{1}_{X_k \le t} - F(t)$. More precisely, since $|Y_0(t)| \le 1$, the criterion (6.19) is exactly the criterion (6.25).

It remains to prove Proposition 6.2. We first quote that condition (6.26) implies (6.25): this follows easily from (5.2) and (5.4). It remains to prove that the covariance operator Λ given in (6.10) can be expressed as in (6.27). As usual, we identify a function f in $\mathbb{L}^\infty(m)$ with an element of the dual of $\mathbb{L}^1(m)$, and we write

$$f(Y_k) = \int f(t) Y_k(t) m(dt).$$

By Remark 6.3, we know that, for any f in $\mathbb{L}^\infty(m)$,

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{E}\big(\big(f(S_n)\big)^2\big)=\mathbb{E}\big(\big(f(W_1)\big)^2\big)=\Lambda(f,f). \tag{6.28}$$

Now, if we can prove that, for any f, g in $\mathbb{L}^\infty(m)$,

$$\sum_{k\in\mathbb{Z}}\big|\mathrm{Cov}\big(f(Y_0),g(Y_k)\big)\big|<\infty, \tag{6.29}$$

then the series

$$\bar\Lambda(f,g)=\sum_{k\in\mathbb{Z}}\mathrm{Cov}\big(f(Y_0),g(Y_k)\big)$$

is well defined, and

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{E}\big(\big(f(S_n)\big)^2\big)=\bar\Lambda(f,f). \tag{6.30}$$

From (6.28) and (6.30), we infer that, for any f in $\mathbb{L}^\infty(m)$, $\Lambda(f,f)=\bar\Lambda(f,f)$. Applying this equality to f, g, and $f+g$ it follows that, for any f, g in $\mathbb{L}^\infty(m)$,

$$\Lambda(f,g)=\bar\Lambda(f,g),$$

which is the desired result. To prove (6.29), we first note that

$$\big|\mathrm{Cov}\big(f(Y_0),g(Y_k)\big)\big|\le\|f\|_\infty\|g\|_\infty\int\!\!\int\big\|\big(\mathbf{1}_{X_0\le t}-F(t)\big)\mathbb{E}_0\big(\mathbf{1}_{X_k\le s}-F(s)\big)\big\|_1 m(dt)m(ds).$$

Now

$$\big\|\big(\mathbf{1}_{X_0\le t}-F(t)\big)\mathbb{E}_0\big(\mathbf{1}_{X_k\le s}-F(s)\big)\big\|_1\le\min\big\{\alpha_{1,\mathbf{X}}(k),2B(t),2B(s)\big\}.$$

Hence

$$\begin{aligned}&\sum_{k=0}\big\|\big(\mathbf{1}_{X_0\le t}-F(t)\big)\mathbb{E}_0\big(\mathbf{1}_{X_k\le s}-F(s)\big)\big\|_1\\&\quad\le 2\sqrt{\sum_{k=0}^\infty\min\big\{\alpha_{1,\mathbf{X}}(k),B(t)\big\}}\sqrt{\sum_{k=0}^\infty\min\big\{\alpha_{1,\mathbf{X}}(k),B(s)\big\}}.\end{aligned}$$

This implies that

$$\sum_{k\in\mathbb{Z}}\big|\mathrm{Cov}\big(f(Y_0),g(Y_k)\big)\big|\le 2\|f\|_\infty\|g\|_\infty\left(\int\sqrt{\sum_{k=0}^\infty\min\big\{\alpha_{1,\mathbf{X}}(k),B(t)\big\}}m(dt)\right)^2,$$

and (6.29) follows from (6.26). This completes the proof of Proposition 6.2. □

7. Quantile conditions

As a consequence of the results by Dédé [5] or Cuny [3] (see the condition (6.4) of Section 6.1) we know that the conclusion of Theorem 6.4 holds as soon as

$$\sum_{k=0}^{\infty} \frac{1}{\sqrt{k+1}} \int \left\| F_{X_k|\mathcal{F}_0}(t) - F(t) \right\|_2 m(dt) < \infty. \tag{7.1}$$

Moreover, it follows from [3] that the condition (7.1) also implies the strong invariance principle.

Let $B(t) = F(t)(1 - F(t))$. As quoted by Dédé (2009), the condition (7.1) is implied by

$$\sum_{k=0}^{\infty} \frac{1}{\sqrt{k+1}} \int \sqrt{\min\{\alpha_{1,\mathbf{X}}(k), B(t)\}} m(dt). \tag{7.2}$$

The conditions (6.26) of Proposition 6.2 and the condition (7.2) are not easy to compare. However, if either m has finite mass or X_0 is bounded, then (6.26) is equivalent to

$$\sum_{k=1}^{\infty} \alpha_{1,\mathbf{X}}(k) < \infty \tag{7.3}$$

and (7.2) is equivalent to

$$\sum_{k=1}^{\infty} \sqrt{\frac{\alpha_{1,\mathbf{X}}(k)}{k}} < \infty. \tag{7.4}$$

Hence, in that case, the condition (6.26) is weaker than the condition (7.2), and is in fact equivalent to the minimal condition to get the central limit theorem for partial sums of stationary α-dependent sequences of bounded random variables.

We shall now focus on the case where $m = \lambda$ is the Lebesgue measure on $\mathbb{R}$. In that case, the condition (7.2) is equivalent to

$$\sum_{k=0}^{\infty} \frac{1}{\sqrt{k+1}} \int_0^{\infty} \sqrt{\min\{\alpha_{1,\mathbf{X}}(k), H(t)\}}\, dt, \tag{7.5}$$

and the condition (6.26) is equivalent to (3.1). We shall see that the condition (3.1) is always weaker than the condition (7.5). The first step is to express (3.1) and (7.5) in terms of the quantile function of X_0, as done in [11] for the invariance principle of stationary α-mixing sequences. More precisely, we shall compare the three following conditions:

$$\int_0^1 \alpha^{-1}(u) Q^2(u)\, du < \infty, \tag{7.6}$$

$$\int_0^1 \frac{\alpha^{-1}(u) Q(u)}{\sqrt{\int_0^u \alpha^{-1}(x)\, dx}} du < \infty, \tag{7.7}$$

$$\int_0^1 \frac{\sqrt{\alpha^{-1}(u)}Q(u)}{\sqrt{u}}\,du < \infty, \tag{7.8}$$

where for simplicity we denote by α^{-1} the function $\alpha_{1,\mathbf{X}}^{-1}$ defined in (3.15). The condition (7.6) has been introduced by Doukhan, Massart and Rio [11], but in that paper the function α^{-1} is defined with the α-mixing coefficients of Rosenblatt [24]. These authors showed that (7.6) implies the functional central limit theorem for the Donsker line

$$\left\{\frac{1}{\sqrt{n}}\sum_{k=1}^{[nt]}\bigl(X_k - \mathbb{E}(X_k)\bigr), t \in [0,1]\right\},$$

and that it is optimal in a precise sense. The optimality of this condition has been further discussed in a paper by Bradley [2]. The fact that, for ergodic sequences, this functional central limit theorem remains true with the much weaker coefficients $\alpha_{1,\mathbf{X}}(k)$ is a consequence of a result by Dedecker and Rio [9].

Concerning these three quantile conditions, our first result is Proposition 7.1 below.

Proposition 7.1. *The following equivalences hold*:

1. *The condition* (7.6) *is equivalent to*

$$\int_0^\infty t\left(\sum_{k=0}^\infty \min\bigl\{\alpha_{1,\mathbf{X}}(k), H(t)\bigr\}\right)dt < \infty. \tag{7.9}$$

2. *The condition* (7.7) *is equivalent to* (3.1).
3. *The condition* (7.8) *is equivalent to* (7.5).

The hierarchy of these quantile conditions is given in Proposition 7.2 below.

Proposition 7.2. *The following implications hold*: (7.8) $\Rightarrow$ (7.7) $\Rightarrow$ (7.6).

Remark 7.1. At this point, it should be noticed that these three conditions are in fact very close. Indeed, by a simple application of Cauchy–Schwarz inequality, for any $b > 1/2$,

$$(7.6) \quad \Rightarrow \quad \int_0^1 \frac{\sqrt{\alpha^{-1}(u)}Q(u)}{\sqrt{u}|1+\ln(u)|^b}\,du < \infty$$

and the condition on right-hand side is a slight reinforcement of (7.8).

Proof of Proposition 7.1. Assume that $\sum_{k\geq 0}\alpha_{1,\mathbf{X}}(k) < \infty$. Then the function S defined on $\mathbb{R}^+$ by

$$S(t) = \sum_{k=0}^\infty \min\bigl\{\alpha_{1,\mathbf{X}}(k), H(t)\bigr\} \tag{7.10}$$

is finite and non-increasing.

Proof of Item 1. By a simple change of variables, we see that the condition (7.9) is equivalent to

$$\int_0^\infty S(\sqrt{t})\,dt < \infty.$$

Since

$$S(t) = \sum_{k=0}^\infty \int_0^1 \mathbf{1}_{u \le \min\{\alpha_{1,\mathbf{X}}(k), H(t)\}}\,du = \int_0^{H(t)} \alpha^{-1}(u)\,du, \tag{7.11}$$

it follows that

$$\begin{aligned}\int_0^\infty S(\sqrt{t})\,dt &= \int_0^\infty \left(\int_0^1 \alpha^{-1}(u)\mathbf{1}_{u \le H(\sqrt{t})}\,du\right)dt \\ &= \int_0^1 \alpha^{-1}(u)\left(\int_0^\infty \mathbf{1}_{t \le Q^2(u)}\,dt\right)du \\ &= \int_0^1 \alpha^{-1}(u)Q^2(u)\,du,\end{aligned}$$

which concludes the proof of Item 1.

Proof of Item 2. Starting from (7.11), it follows that

$$\int_0^\infty \sqrt{\sum_{k=0}^\infty \min\{\alpha_{1,\mathbf{X}}(k), H(t)\}}\,dt = \int_0^\infty \sqrt{\int_0^{H(t)} \alpha^{-1}(u)\,du}\,dt. \tag{7.12}$$

Let

$$G_\alpha(x) = \sqrt{\int_0^x \alpha^{-1}(u)\,du}.$$

From (7.12), we infer that

$$\begin{aligned}\int_0^\infty \sqrt{\sum_{k=0}^\infty \min\{\alpha_{1,\mathbf{X}}(k), H(t)\}}\,dt &= \int_0^\infty \int_0^1 \mathbf{1}_{v \le G_\alpha(H(t))}\,dv\,dt \\ &= \int_0^1 \int_0^\infty \mathbf{1}_{t \le Q(G_\alpha^{-1}(v))}\,dt\,dv \\ &= \int_0^1 Q \circ G_\alpha^{-1}(v)\,dv.\end{aligned}$$

Making the change of variables $u = G_\alpha^{-1}(v)$, the result follows.

Proof of Item 3. Note first that

$$\int_0^\infty \sqrt{\min\{\alpha_{1,\mathbf{X}}(k), H(t)\}}\,dt = \int_0^\infty \left(\int_0^1 \mathbf{1}_{u^2\le\alpha_{1,\mathbf{X}}(k)}\mathbf{1}_{u^2\le H(t)}\,du\right)dt \tag{7.13}$$

$$= \int_0^1 Q(u^2)\mathbf{1}_{u^2\le\alpha_{1,\mathbf{X}}(k)}\,du. \tag{7.14}$$

Now $u^2 \le \alpha_{1,\mathbf{X}}(k)$ if and only if $k \le \alpha^{-1}(u^2)$. Hence, there exists two positive constants A and B such that

$$A\sqrt{\alpha^{-1}(u^2)} \le \sum_{k=1}^\infty \frac{1}{\sqrt{k}}\mathbf{1}_{u^2\le\alpha_{1,\mathbf{X}}(k)} \le B\sqrt{\alpha^{-1}(u^2)}.$$

Finally

$$\sum_{k=1}^\infty \frac{1}{\sqrt{k}}\int_0^\infty \sqrt{\min\{\alpha_{1,\mathbf{X}}(k), H(t)\}}\,dt < \infty \quad \text{iff} \quad \int_0^1 \sqrt{\alpha^{-1}(u^2)}Q(u^2)\,du < \infty.$$

Making the change of variables $v = u^2$, the result follows. □

Proof of Proposition 7.2. Since the function α^{-1} is non-increasing, one has

$$\int_0^u \alpha^{-1}(x)\,dx \ge u\alpha^{-1}(u),$$

which proves that (7.8) implies (7.7).

It remains to prove that (7.7) implies (7.6). By Proposition 7.1, it is equivalent to prove that (3.1) implies (7.9). If (3.1) holds, then the function S defined on $\mathbb{R}^+$ by (7.10) is finite and non-increasing. Hence, using again (3.1),

$$t\sqrt{S(t)} \le 2\int_{t/2}^t \sqrt{S(s)}\,ds \le C, \qquad \text{with } C = 2\int_0^\infty \sqrt{S(s)}\,ds.$$

Consequently $tS(t) \le C\sqrt{S(t)}$, proving that (3.1) implies (7.9). □

7.1. Sufficient conditions

In this subsection, we give some simple conditions on $\alpha_{1,\mathbf{X}}(k)$ and H under which (7.7) (and hence (3.1)) is satisfied.

Proposition 7.3. *The following conditions imply* (7.7):

1.
$$\mathbb{E}\big(|X_0|^p\big) < \infty \qquad \textit{for some } p > 2, \quad \textit{and} \quad \sum_{k>0} \frac{(\alpha_{1,\mathbf{X}}(k))^{(p-2)/(2(p-1))}}{k^{(p-2)/(2(p-1))}} < \infty.$$

2.
$$H(t) = O\big(t^{-p}\big) \qquad \textit{for some } p > 2, \quad \textit{and} \quad \sum_{k>0} \frac{(\alpha_{1,\mathbf{X}}(k))^{(p-2)/(2p)}}{\sqrt{k}} < \infty.$$

3.
$$\int_0^\infty \big(H(t)\big)^{(a-1)/(2a)}\, dt < \infty \quad \textit{and} \quad \alpha_{1,\mathbf{X}}(k) = O\left(\frac{1}{k^a}\right) \qquad \textit{for some } a > 1.$$

4.
$$\int_0^\infty \left(\ln\left(1 + \frac{1}{H(t)}\right)\right)^{-(a-1)/2} dt < \infty \quad \textit{and} \quad \alpha_{1,\mathbf{X}}(k) = O\left(\frac{1}{k(\ln(k))^a}\right) \qquad \textit{for some } a > 1.$$

5.
$$\int_0^\infty \sqrt{H(t)\big|\ln\big(H(t)\big)\big|}\, dt < \infty \quad \textit{and} \quad \alpha_{1,\mathbf{X}}(k) = O\big(a^k\big) \qquad \textit{for some } a < 1.$$

Proof of Proposition 7.3. *Proof of Item* 1. Since (7.8) implies (7.7), it suffices to prove that Item 1 implies (7.8). Applying Cauchy–Schwarz, we obtain that

$$\int_0^1 \frac{\sqrt{\alpha^{-1}(u)}\,Q(u)}{\sqrt{u}}\, du \le \left(\int_0^1 Q(u)^p\, du\right)^{1/p} \left(\int_0^1 \left(\frac{\alpha^{-1}(u)}{u}\right)^{p/(2(p-1))} du\right)^{(p-1)/p}.$$

Since $\mathbb{E}(|X_0|^p) < \infty$, the first integral on right hand is finite. It remains to prove that

$$\int_0^1 \left(\frac{\alpha^{-1}(u)}{u}\right)^{p/(2(p-1))} du < \infty.$$

By definition of α^{-1}, this is equivalent to

$$\sum_{k>0} k^{p/(2(p-1))} \int_{\alpha_{1,\mathbf{X}}(k+1)}^{\alpha_{1,\mathbf{X}}(k)} u^{-p/(2(p-1))}\, du < \infty.$$

The last condition means exactly that

$$\sum_{k>0} k^{p/(2(p-1))} \Big(\big(\alpha_{1,\mathbf{X}}(k)\big)^{(p-2)/(2(p-1))} - \big(\alpha_{1,\mathbf{X}}(k+1)\big)^{(p-2)/(2(p-1))}\Big) < \infty,$$

which is equivalent to the condition of Item 1.

Proof of Item 2. Again, it suffices to prove that Item 2 implies (7.8). Now, the condition $H(t) = O(t^{-p})$ is equivalent to $Q(u) = O(u^{-1/p})$. Hence, the condition (7.8) holds as soon as

$$\int_0^1 \frac{\sqrt{\alpha^{-1}(u)}}{u^{1/p+1/2}}\, du < \infty.$$

By definition of α^{-1}, the last condition means exactly that

$$\sum_{k>0} \sqrt{k}\Big(\big(\alpha_{1,\mathbf{X}}(k)\big)^{(p-2)/(2p)} - \big(\alpha_{1,\mathbf{X}}(k+1)\big)^{(p-2)/(2p)}\Big) < \infty,$$

which is equivalent to the condition of Item 2.

Proofs of Item 3, 4 *and* 5. For the proof of these points, we start from condition (3.1) which is equivalent to (7.7). Since we can control the behavior of $\alpha_{1,\mathbf{X}}(k)$, we can give upper bounds for the function S defined by (7.10).

If $\alpha_{1,\mathbf{X}}(k) = O(\frac{1}{k^a})$ for some $a > 1$, then $S(t) = O((H(t))^{(a-1)/a})$.

If $\alpha_{1,\mathbf{X}}(k) = O(\frac{1}{k(\ln(k))^a})$ for some $a > 1$, then

$$S(t) = O\bigg(\bigg(\ln\bigg(1 + \frac{1}{H(t)}\bigg)\bigg)^{-(a-1)/2}\bigg).$$

If $\alpha_{1,\mathbf{X}}(k) = O(a^k)$ for some $a < 1$, then $S(t) = O(H(t)|\ln(H(t))|)$.

Item 3, 4 and 5 follow from these upper bounds and condition (3.1). □

Acknowledgments

We thank the two anonymous referees and the Editor-in-Chief for their careful reading of the paper, and for helpful suggestions.

References

[1] Bobkov, S. and Ledoux, M. (2014). One-dimensional empirical measures, order statistics and Kantorovich transport distances. Preprint.

[2] Bradley, R.C. (1997). On quantiles and the central limit question for strongly mixing sequences. *J. Theoret. Probab.* **10** 507–555. MR1455156

[3] Cuny, C. (2016). Limit theorems under the Maxwell–Woodroofe condition in Banach spaces. *Ann. Probab.* To appear. Available at arXiv:1403.0772.

[4] de Acosta, A., Araujo, A. and Giné, E. (1978). On Poisson measures, Gaussian measures and the central limit theorem in Banach spaces. In *Probability on Banach Spaces*. *Adv. Probab. Related Topics* **4** 1–68. New York: Dekker. MR0515429

[5] Dede, S. (2009). An empirical central limit theorem in $\mathbf{L}^1$ for stationary sequences. *Stochastic Process. Appl.* **119** 3494–3515. MR2568284

[6] Dedecker, J. and Doukhan, P. (2003). A new covariance inequality and applications. *Stochastic Process. Appl.* **106** 63–80. MR1983043

[7] Dedecker, J., Gouëzel, S. and Merlevède, F. (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. MR2682267

[8] Dedecker, J. and Merlevède, F. (2015). Moment bounds for dependent sequences in smooth Banach spaces. *Stochastic Process. Appl.* **125** 3401–3429. MR3357614

[9] Dedecker, J. and Rio, E. (2000). On the functional central limit theorem for stationary processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **36** 1–34. MR1743095

[10] del Barrio, E., Giné, E. and Matrán, C. (1999). Central limit theorems for the Wasserstein distance between the empirical and the true distributions. *Ann. Probab.* **27** 1009–1071. MR1698999

[11] Doukhan, P., Massart, P. and Rio, E. (1994). The functional central limit theorem for strongly mixing processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **30** 63–82. MR1262892

[12] Èbralidze, Š.S. (1971). Inequalities for the probabilities of large deviations in terms of pseudomoments. *Teor. Verojatnost. i Primenen.* **16** 760–765. MR0295417

[13] Esseen, C. and Janson, S. (1985). On moment conditions for normed sums of independent variables and martingale differences. *Stochastic Process. Appl.* **19** 173–182. MR0780729

[14] Fournier, N. and Guillin, A. (2015). On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields* **162** 707–738. MR3383341

[15] Gordin, M.I. (1973). Abstracts of communication, T.1:A-K. In *International Conference on Probability Theory*. Vilnius.

[16] Gouëzel, S. (2004). Central limit theorem and stable laws for intermittent maps. *Probab. Theory Related Fields* **128** 82–122. MR2027296

[17] Hannan, E.J. (1973). Central limit theorems for time series regression. *Z. Wahrsch. Verw. Gebiete* **26** 157–170. MR0331683

[18] Hennion, H. and Hervé, L. (2001). *Limit Theorems for Markov Chains and Stochastic Properties of Dynamical Systems by Quasi-Compactness. Lecture Notes in Math.* **1766**. Berlin: Springer. MR1862393

[19] Jain, N.C. (1977). Central limit theorem and related questions in Banach space. In *Probability* (*Proc. Sympos. Pure Math., Vol. XXXI, Univ. Illinois, Urbana, Ill.*, 1976) 55–65. Providence, RI: Amer. Math. Soc. MR0451328

[20] Liverani, C., Saussol, B. and Vaienti, S. (1999). A probabilistic approach to intermittency. *Ergodic Theory Dynam. Systems* **19** 671–685. MR1695915

[21] Maxwell, M. and Woodroofe, M. (2000). Central limit theorems for additive functionals of Markov chains. *Ann. Probab.* **28** 713–724. MR1782272

[22] Merlevède, F. and Peligrad, M. (2013). Rosenthal-type inequalities for the maximum of partial sums of stationary processes and examples. *Ann. Probab.* **41** 914–960. MR3077530

[23] Rio, E. (2000). *Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants. Mathématiques & Applications* (*Berlin*) [*Mathematics & Applications*] **31**. Berlin: Springer. MR2117923

[24] Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Natl. Acad. Sci. USA* **42** 43–47. MR0074711

[25] Tucker, H.G. (1959). A generalization of the Glivenko–Cantelli theorem. *Ann. Math. Statist.* **30** 828–830. MR0107891

[26] von Bahr, B. and Esseen, C. (1965). Inequalities for the rth absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Ann. Math. Statist* **36** 299–303. MR0170407

Received February 2015 and revised October 2015