On the weak invariance principle for ortho-martingale in Banach spaces. Application to stationary random fields

Han-Mai Lin, Florence Merlevède

On the weak invariance principle for ortho-martingale in Banach spaces. Application to stationary random fields

Han-Mai Lin[*] and Florence Merlevède[†]

December 31, 2021

Abstract

In this paper, we study the weak invariance principle for stationary ortho-martingales with values in 2-smooth or cotype 2 Banach spaces. Then, with the help of a suitable maximal ortho-martingale approximation, we derive the weak invariance principle for stationary random fields in L^1 under a condition in the spirit of Hannan. As an application, we get an asymptotic result for the L^1-Wasserstein distance between the common distribution function and the corresponding empirical distribution function of stationary random fields.

2020 Mathematics Subject Classification. *60F17, 60G60, 60F25.*
Key words and phrases. *Invariance principle, Brownian sheet, random field, ortho-martingale approximation, Empirical distribution function, Wasserstein distance*

1 Introduction

Let $(X_i)_{i\in\mathbb{Z}}$ be a stationary random sequence. If one may represent $(X_i)_{i\in\mathbb{Z}}$ in the form

$$X_i = d_i + \zeta_i, \tag{1.1}$$

where $(d_i)_{i\in\mathbb{Z}}$ is a stationary sequence of martingale differences, and $(\zeta_i)_{\in\mathbb{Z}}$ is a coboundary, which means that it can be written as $\zeta_i = \theta_i - \theta_{i-1}$ with $(\theta_i)_{i\in\mathbb{Z}}$ a stationary random sequence, then one may derive the CLT as well as other limit theorems from the corresponding ones for the martingale differences $(d_i)_{i\in\mathbb{Z}}$. This method is usually called the martingale approximation, but also known as the Gordin's method (see [11]). Note that as a variation of this method, the sequence $(\zeta_i)_{\in\mathbb{Z}}$ needs not to be coboundary, but suitably normalized still negligible, for instance in probability. This kind of approach has been developed in many papers. We refer to Chapter 4 in Merlevède et al. [17] for a survey concerning Gaussian approximation via martingale methods.

In this paper, we will use an adaptation of the martingale approximation method to prove the CLT and its functional form for the partial sums associated with multi-indexed sequences

[*]Université Gustave Eiffel, LAMA and CNRS UMR 8050. Email: han-mai.lin@univ-eiffel.fr
[†]Université Gustave Eiffel, LAMA and CNRS UMR 8050. Email: florence.merlevede@univ-eiffel.fr

(also called random fields) with values in some Banach spaces $\mathcal{X}$. Let us first recall some recent results concerning the limiting behavior of the partial sums associated with random sequences in Banach spaces.

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, T be an invertible bi-measurable measure preserving transformation on Ω. Let $\mathcal{F} \subset \mathcal{A}$ be a σ-algebra such that $T^{-1}\mathcal{F} \subset \mathcal{F}$, and define a non-decreasing filtration $\mathcal{F}_i = T^{-i}(\mathcal{F})$. We also define the stationary sequence of random elements $X := (X_i)_{i\in\mathbb{Z}}$ by $X_i = X_0 \circ T^i$. Let $\mathcal{X} = L^1(\mathbb{R}, \mu)$, where μ is a σ-finite measure on the real line, and $L^\infty(\mathbb{R}, \mu)$ be its dual space. When $(X_i)_{i\in\mathbb{Z}}$ is a L^1-valued stationary ergodic sequence of martingale differences (m.d.s.), Dede [4] proved the following CLT: If

$$\int_{\mathbb{R}} \|X_0(t)\|_2 \, \mu(dt) < \infty, \tag{1.2}$$

then

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X \circ T^i \Rightarrow_{n\to\infty} G \text{ in } L^1. \tag{1.3}$$

Here and along the paper, $\Rightarrow$ stands for the convergence in distribution and G is a centered Gaussian random variable with covariance operator $\mathcal{K}_X$ (see definition 3.1). Then, with the help of a martingale approximation, she extended the CLT to stationary and ergodic sequences of regular random elements with values in $L^1(\mathbb{R}, \mu)$ provided that (1.2) holds and

$$\sum_{i\in\mathbb{Z}} \int_{\mathbb{R}} \|P_0(X_i(t))\|_2 \, \mu(dt) < \infty, \tag{1.4}$$

where $P_0(\cdot) = \mathbb{E}(\cdot|\mathcal{F}_0) - \mathbb{E}(\cdot|\mathcal{F}_{-1})$. Her condition is in the spirit of Hannan's condition [13].

Later, Cuny [2] has proved the CLT and its functional form for stationary and ergodic sequences of m.d.s. with values in more general Banach spaces than L^1, namely cotype 2 or 2-smooth Banach spaces. Next, again with the help of a martingale approximation, he derived sufficient conditions in the spirit of the Maxwell-Woodroofe's condition [16] ensuring the weak invariance principle for a stationary and ergodic sequence with values in either 2-smooth or cotype 2 Banach spaces. Note that when applied to the L^1-space, his condition and (1.4) have a different area of applications.

The first aim of this paper is to generalize the functional CLT for martingale differences sequences with values in some Banach spaces as stated in [2, Prop 3.2] to higher dimension in the sense of multi-indexed sequences. Since there is no natural ordering in higher dimension, we choose to use the notion of completely commuting filtration (see Section 2 for a definition of this notion). It follows that the generalization of [2, Prop 3.2] will mean extending the functional CLT to ortho-martingale sequences with values in a 2-smooth or cotype 2 Banach spaces (see Section 3). Then, in Section 4, with the help of a suitable ortho-martingale approximation, we derive a functional CLT for L^1-valued stationary random fields. As in [4], our conditions are in the spirit of Hannan's condition. Hence, our theorem 4.1 can be viewed as a generalization of the results stated in [4] in two directions. First, it proves the functional form of [4, Th 2.3]. Secondly, it extends it to multi-indexed random sequences. As an application, in Section 5,

we give asymptotic results for the L^1-Wasserstein distance between the common distribution function and the corresponding empirical distribution function for stationary random fields. Section 6 is devoted to a suitable maximal ortho-martingale approximation in 2-smooth Banach spaces and has interest in itself. All the proofs are postponed to Section 7.

2 Notations and definitions

We will use the same notations as in [3] and [19]. We shall consider Banach-valued random fields. Hence, in all the paper, $(\mathcal{X}, |\cdot|_{\mathcal{X}})$ will be a real separable Banach space. We denote by $\mathcal{X}^*$ the topological dual of $\mathcal{X}$. We shall also denote by $L^0(\mathcal{X})$ the space of functions from $\Omega \to \mathcal{X}$ that are limits $\mathbb{P}$-a.s. of simple functions. In addition, for every $p \geq 1$, we define the usual Bochner spaces L^p as follows

$$L^p(\Omega, \mathcal{A}, \mathbb{P}, \mathcal{X}) = L^p(\mathcal{X}) = \left\{ Z \in L^0(\mathcal{X}) : \mathbb{E}\left(|Z|_{\mathcal{X}}^p\right) < \infty \right\}.$$

For every Z in $L^p(\mathcal{X})$, write $\|Z\|_{p,\mathcal{X}} = \left(\mathbb{E}\left(|Z|_{\mathcal{X}}^p\right)\right)^{1/p} = \left\||Z|_{\mathcal{X}}\right\|_p$. To define now random fields, we start by introducing $\mathbb{Z}^d$ actions. With this aim, we denote elements of $\mathbb{Z}^d$ by $\underline{i} := (i_1 \ldots, i_d)$. Let $(T_{\underline{i}})_{\underline{i} \in \mathbb{Z}^d}$ be a $\mathbb{Z}^d$ action on $(\Omega, \mathcal{A}, \mathbb{P})$ generated by commuting invertible and measure-preserving transformations $\{T_{e_1}, \ldots, T_{e_d}\}$, where, for $1 \leq i \leq d$, e_i is a $\mathbb{Z}^d$ vector with 1 at i-th place and 0 elsewhere. We denote those transformations by $\left\{T_{(1)}, \ldots, T_{(d)}\right\}$. By $U_{\underline{i}}$ we denote the operator in $L^p(\mathcal{X})$ $(1 \leq p < \infty)$ defined by $U_{\underline{i}} f = f \circ T_{\underline{i}}$. We assume that the filtration $(\mathcal{F}_{\underline{k}})_{\underline{k} \in \mathbb{Z}^d}$ is completely commuting, i.e. there exists a σ-algebra $\mathcal{F}$ such that $\mathcal{F}_{\underline{i}} = T_{-\underline{i}}\mathcal{F}$, for $\underline{i} \leq \underline{j}$ we have $\mathcal{F}_{\underline{i}} \subset \mathcal{F}_{\underline{j}}$ ($\underline{i} \leq \underline{j}$ means $i_q \leq j_q$ for all $1 \leq q \leq d$) and for an integrable X, we have

$$\mathbb{E}\left(\mathbb{E}\left(X|\mathcal{F}_{i_1,\ldots,i_d}\right)|\mathcal{F}_{j_1,\ldots,j_d}\right) = \mathbb{E}\left(X|\mathcal{F}_{i_1 \wedge j_1, \ldots i_d \wedge j_d}\right),$$

where $i \wedge j = \min\{i, j\}$. Note that a filtration defined by an independent and identically distributed (i.i.d.) random field is completely commuting. This kind of filtrations can also be constructed using stationary random fields with independent rows or columns (see [10]).

For a fixed $q \in \{1, \ldots, d\}$ and a fixed $\ell \in \mathbb{Z}$, we denote by $\mathcal{F}_\ell^{(q)}$ the σ-algebra generated by all $\mathcal{F}_{\underline{i}}$ with $\underline{i}$ such that $i_q \leq \ell$. For every $1 \leq q \leq d$, define also $\mathcal{F}_{-\infty}^{(q)} = \bigwedge_{i \in \mathbb{Z}} \mathcal{F}_i^{(q)}$, and $\mathcal{F}_{\underline{\infty}} = \bigvee_{\underline{i} \in \mathbb{Z}^d} \mathcal{F}_{\underline{i}}$.

For σ-algebras $\mathcal{G} \subset \mathcal{F} \subset \mathcal{A}$ and $1 \leq p < \infty$, we denote by $L^p(\mathcal{F}, \mathcal{X}) \ominus L^p(\mathcal{G}, \mathcal{X})$ the space of $X \in L^p(\mathcal{F}, \mathcal{X})$ for which $\mathbb{E}(X|\mathcal{G}) = 0$. We can define projection operators $P_\ell^{(q)}$ onto $L^p(\mathcal{F}_\ell^{(q)}, \mathcal{X}) \ominus L^p(\mathcal{F}_{\ell-1}^{(q)}, \mathcal{X})$ by $P_\ell^{(q)}(X) = \mathbb{E}(X|\mathcal{F}_\ell^{(q)}) - \mathbb{E}(X|\mathcal{F}_{\ell-1}^{(q)})$. Those operators commute and for $\ell \neq k$, $P_\ell^{(q)} P_k^{(q)} = 0$ (see [20, Lemma 2.4]). We now define the projections $P_{\underline{i}} = P_{i_1}^{(1)} \cdots P_{i_d}^{(d)}$ onto $\bigcap_{1 \leq q \leq d} L^p(\mathcal{F}_{i_q}^{(q)}, \mathcal{X}) \ominus L^p(\mathcal{F}_{i_q - 1}^{(q)}, \mathcal{X})$.

Definition 2.1. *Let $X \in L^1(\mathcal{X})$. The stationary random fields $(U_{\underline{i}} X)_{\underline{i} \in \mathbb{Z}^d}$ is said to be regular if X is $\mathcal{F}_{\underline{\infty}}$-measurable, and for every $1 \leq q \leq d$, $\mathbb{E}(X|\mathcal{F}_{-\infty}^{(q)}) = 0$ a.s.*

Next we give the definition of ortho-martingales.

Definition 2.2. *Let $D \in L^1(\mathcal{X})$. We say that $(U_{\underline{i}}D)_{\underline{i}\in\mathbb{Z}^d}$ is a field of ortho-martingale differences with respect to a completely commuting filtration $(\mathcal{F}_{\underline{i}})_{\underline{i}\in\mathbb{Z}^d}$, if D is $\mathcal{F}_{\underline{0}}$-measurable and for all $\underline{i},\underline{j} \in \mathbb{Z}^d$, such that $j_q < i_q$ for some $q \in \{1,\dots,d\}$, we have*

$$\mathbb{E}\left(D \circ T_{\underline{i}} | \mathcal{F}_{\underline{j}}\right) = 0\,.$$

In addition, $M_{\underline{n}} := \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} D \circ T_{\underline{i}}$ is said to be an ortho-martingale.

Since our results are stated for random fields in Banach spaces that are 2-smooth or of cotype 2, let us recall their definitions from [2]. We say that $\mathcal{X}$ is 2-smooth, if there exists $L \geq 1$, such that

$$|x+y|_{\mathcal{X}}^2 + |x-y|_{\mathcal{X}}^2 = 2\left(|x|_{\mathcal{X}}^2 + L\,|y|_{\mathcal{X}}^2\right) \qquad \forall x, y \in \mathcal{X}. \tag{2.1}$$

For a such L, we say that $\mathcal{X}$ is $(2,L)$-smooth. We shall also recall the following inequality for m.d.s. in 2-smooth Banach space ([1, Prop 1]): Assume that $\mathcal{X}$ is $(2,L)$-smooth, then for every m.d.s. $(D_i)_{i\in\mathbb{Z}}$, we have

$$\mathbb{E}\left(|D_1 + \cdots + D_N|_{\mathcal{X}}^2\right) \leq 2L^2 \sum_{i=1}^{N} \mathbb{E}\left(|D_i|_{\mathcal{X}}^2\right) \qquad \text{for all } N \in \mathbb{N}\,. \tag{2.2}$$

For example, for $p \geq 2$ the spaces $L^p(\mathbb{R})$ are $(2, \sqrt{p-1})$-smooth. As a counterpart, we recall that a separable Banach space $\mathcal{X}$ is said of cotype 2 if there exists $L > 0$ such that for every independent random variables $d_1, \dots, d_N \in L^2(\mathcal{X})$, with $\mathbb{E}(d_1) = \ldots = \mathbb{E}(d_N) = 0$, (2.2) holds in the reverse direction. As an example, note that for $1 \leq p \leq 2$, $L^p(\mathbb{R})$ is of cotype 2.

3 Functional CLT for ortho-martingales in Banach spaces

Let X be a random element from Ω to $\mathcal{X}$. Define its associated partial sum $S_{\underline{n}}(X)$ by

$$S_{\underline{n}}(X) = \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} U_{\underline{i}} X\,.$$

In addition, for $\underline{t} = (t_1, \dots, t_d) \in [0,1]^d$, we set

$$T_{\underline{nt}}(X) := \frac{1}{\sqrt{n_1 \cdots n_d}} S_{[\underline{nt}]}(X)\,, \tag{3.1}$$

where $[\underline{nt}] = ([n_1 t_1], \dots, [n_d t_d])$. Before stating the main result of this section, as in [2], we need to recall the definitions of Gaussian and pregaussian random elements.

Definition 3.1. *Let $\mathcal{K}$ be a bounded symmetric bilinear operator from $\mathcal{X}^* \times \mathcal{X}^*$ to $\mathbb{R}$. We say that $\mathcal{K} = \mathcal{K}_X$ is the covariance operator associated with X if*

$$\forall x^*, y^* \in \mathcal{X}^* \qquad \mathcal{K}(x^*, y^*) = \mathbb{E}\left[x^*(X) y^*(X)\right].$$

Definition 3.2. *The random variable $W \in L^0(\mathcal{X})$ is said to be Gaussian if, for every $x^* \in \mathcal{X}^*$, $x^*(W)$ has a normal distribution. We say that a random variable $X \in L^0(\mathcal{X})$ such that for every $x^* \in \mathcal{X}^*$, $\mathbb{E}\,(x^*(X)^2) < \infty$ and $\mathbb{E}\,(x^*(X)) = 0$, is pregaussian, if there exists a Gaussian variable $W \in L^0(\mathcal{X})$ with the same covariance operator. We denote by $G(X)$ the Gaussian variable having the same covariance operator as X.*

Definition 3.3. *Let $\mathbb{G}(\Omega, \mathcal{A}, \mathbb{P}, \mathcal{X}) = \mathbb{G}(\mathcal{X})$ be the set of pregaussian random elements that are in $L^2(\mathcal{X})$. For every $X \in \mathbb{G}(\mathcal{X})$, denote $\|X\|_{\mathbb{G}(\mathcal{X})} = \|X\|_{2,\mathcal{X}} + \|G(X)\|_{2,\mathcal{X}}$.*

We are now in position to state the functional form of the CLT for sequences of ortho-martingale differences with values in some Banach spaces as well as a L^2-maximal inequality for the corresponding partial sums. Below and all along the paper, $\underline{n} \to \infty$ means $\min(n_1, \ldots, n_d) \to \infty$.

Theorem 3.4. *Let $\mathcal{X}$ be a real separable Banach space that is either 2-smooth or of cotype 2. Let $D \in L^2(\mathcal{X})$ be such that $(U_{\underline{i}} D)_{\underline{i} \in \mathbb{Z}^d}$ is a field of ortho-martingale differences w.r.t. a completely commuting filtration $(\mathcal{F}_{\underline{i}})_{\underline{i} \in \mathbb{Z}^d}$. Assume in addition that $D \in \mathbb{G}\,(\mathcal{F}_{\underline{0}}, \mathcal{X})$ and that at least one of the $T_{(i)}$ for $1 \le i \le d$ is ergodic. Then, as $\underline{n} \to \infty$, $\{T_{\underline{nt}}(D)\}_{\underline{t} \in [0,1]^d}$ converges in distribution in $D([0,1]\,, \mathcal{X})$ (equipped with the uniform topology) to a Brownian motion $(W_{\underline{t}})_{\underline{t} \in [0,1]^d}$ with covariance $\mathcal{K}_D$ associated with D. In addition, there exists $C \ge 0$, such that*

$$\left\| \max_{\underline{1} \le \underline{k} \le \underline{n}} |S_{\underline{k}}(D)|_{\mathcal{X}} \right\|_2 \le C (n_1 \cdots n_d)^{1/2} \|D\|_{\mathbb{G}(\mathcal{X})}\,. \tag{3.2}$$

4 Application to stationary random fields in L^1

In this section we consider $(S, \mathcal{S}, \mu)$ be a σ-finite measure space such that $\mathcal{X} := L^1(S, \mathcal{S}, \mu)$ is separable. Recall that $L^1(S, \mu)$ is a Banach space of cotype 2. Let $L^\infty(S, \mu)$ be its dual space.

In all this section, X is a random variable in $L^0(\Omega, \mathcal{A}, \mathbb{P}, L^1(S, \mu))$ and $(\mathcal{F}_{\underline{i}})_{\underline{i} \in \mathbb{Z}^d}$ with $\mathcal{F}_{\underline{i}} = T_{-\underline{i}} \mathcal{F}$ is a completely commuting filtration. Let also $X_{\underline{i}} = U_{\underline{i}} X$.

Theorem 3.4 combined with a suitable ortho-martingale approximation (see Proposition 6.1 stated in Section 6) leads to the following result.

Theorem 4.1. *Assume that X is regular,*

$$\int_S \|X(s)\|_2 \, \mu(ds) < \infty\,, \tag{4.1}$$

and

$$\sum_{\underline{i} \in \mathbb{Z}^d} \int_S \|P_{\underline{0}}(X_{\underline{i}}(s))\|_2 \, \mu(ds) < \infty\,. \tag{4.2}$$

Then there exists a stationary random field of ortho-martingale differences $(U_{\underline{i}}D)_{\underline{i}\in\mathbb{Z}^d}$ *w.r.t* $(\mathcal{F}_{\underline{i}})_{\underline{i}\in\mathbb{Z}^d}$ *such that* $D \in \mathbb{G}(\mathcal{F}_{\underline{0}}, L^1(S,\mu))$ *and*

$$\lim_{\underline{n}\to\infty} \int_S \left\| \max_{\underline{1}\le\underline{k}\le\underline{n}} \left| \frac{S_{\underline{k}}(X(s))}{\sqrt{n_1\cdots n_d}} - \frac{S_{\underline{k}}(D(s))}{\sqrt{n_1\cdots n_d}} \right| \right\|_2 \mu(ds) = 0. \tag{4.3}$$

Suppose, in addition, that at least one of the transformations $T_{(i)}$ *for* $1 \le i \le d$ *is ergodic, then*

$$\{T_{\underline{nt}}(X), \underline{t} \in [0,1]^d\} \Longrightarrow_{\underline{n}\to\infty} W \text{ in } D([0,1], \mathcal{X}), \tag{4.4}$$

where W *is a* $\mathcal{X}$*-valued centered Brownian motion with covariance operator* $\mathcal{K}_D$ *that can be defined as follows: for all* f *and* g *in* $L^\infty(S,\mu)$*,*

$$\mathcal{K}_D(f,g) = \mathbb{E}(f(D)g(D)) = \sum_{\underline{i}\in\mathbb{Z}^d} \sum_{\underline{k}\in\mathbb{Z}^d} Cov(f(X_{\underline{i}}), g(X_{\underline{k}}))\,. \tag{4.5}$$

Remark 4.2. *Note that by [2, Lemma 2.4], condition* (4.1) *implies that* $X \in \mathbb{G}(L^1(S,\mu))$*.*

Next result gives sufficient conditions for (4.2) to hold.

Corollary 4.3. *Let* $X := (X_{\underline{k}})_{\underline{k}\in\mathbb{Z}^d}$ *be a regular centered random variable in* $L^0(\Omega, \mathcal{A}, \mathbb{P}, L^1(S,\mu))$*, such that* $\int_S \|X_{\underline{0}}(s)\|_2 \,\mu(ds) < \infty$*. If the following conditions hold:*

$$\sum_{\underline{i}=\underline{1}}^{\infty} \int_S \frac{1}{\sqrt{i_1\cdots i_d}} \left\|\mathbb{E}(X_{\underline{i}}(s)|\mathcal{F}_{\underline{0}})\right\|_2 \mu(ds) < \infty, \tag{4.6}$$

and

$$\sum_{\underline{i}=\underline{1}}^{\infty} \int_S \frac{1}{\sqrt{i_1\cdots i_d}} \left\|X_{-\underline{i}}(s) - \mathbb{E}(X_{-\underline{i}}(s)|\mathcal{F}_{\underline{0}})\right\|_2 \mu(ds) < \infty, \tag{4.7}$$

then (4.2) *is satisfied.*

Comment 4.4. *Using Proposition 6.1, a similar result as Theorem 4.1 can be obtained when* $X \in L^0(\Omega, \mathcal{A}, \mathbb{P}, \mathcal{X})$ *where* $\mathcal{X} = L^p(S,\mu)$ *with* $p \ge 2$*. More precisely if* $\|X_{\underline{0}}\|_{2,\mathcal{X}} < \infty$ *and*

$$\sum_{\underline{i}\in\mathbb{Z}^d} \left\| \left(\int_S |P_{\underline{0}}(X_{\underline{i}}(s))|^p \,\mu(ds) \right)^{1/p} \right\|_2 < \infty, \tag{4.8}$$

then

$$\lim_{\underline{n}\to\infty} \prod_{i=1}^d n_i^{-1/2} \left\| \sup_{\underline{1}\le\underline{k}\le\underline{n}} |S_{\underline{k}} - M_{\underline{k}}|_{\mathcal{X}} \right\|_2 = 0\,,$$

and the convergence (4.4) *holds in* $\mathcal{X}$*.*

When $p = 2$*, using Lemma 7.2 adapted to dimension* d *with* $u_{\underline{i}} = \|\|P_{-\underline{i}}(X_{\underline{0}})\|_{L^2}\|_2$ *and the*

fact that L^2 is 2-convex, we infer that (4.8) *holds provided*

$$\sum_{i_1=1}^{\infty}\cdots\sum_{i_d=1}^{\infty}\frac{1}{\sqrt{i_1\cdots i_d}}\left\|\left\|\mathbb{E}(X_{\underline{i}}(s)|\mathcal{F}_{\underline{0}})\right\|_{L^2}\right\|_2<\infty,$$

and

$$\sum_{i_1=1}^{\infty}\cdots\sum_{i_d=1}^{\infty}\frac{1}{\sqrt{i_1\cdots i_d}}\left\|\left\|X_{-\underline{i}}(s)-\mathbb{E}(X_{-\underline{i}}(s)|\mathcal{F}_{\underline{0}})\right\|_{L^2}\right\|_2<\infty.$$

Note that if $X_{\underline{0}}(s)=\sum_{\underline{i}\in\mathbb{Z}^d}a_{\underline{i}}Y_{-\underline{i}}(s)$ where $(Y_{\underline{i}})_{\underline{i}\in\mathbb{Z}^d}$ is an ortho-martingale differences, (4.8) *holds as soon as*

$$\|Y_{\underline{0}}\|_{2,\mathcal{X}}<\infty \text{ and } \sum_{\underline{i}\in\mathbb{Z}^d}|a_{\underline{i}}|<\infty. \tag{4.9}$$

Our condition (4.8) *and condition* (3) *in [2] have a different range of applications. For instance* (4.9) *is not enough for [2, Condition (3)] to hold.*

5 Application to empirical processes

In this section, $S:=\mathbb{R}$ and μ is the Lebesgue measure on the real line denoted by λ. Let $Y\in L^0(\Omega,\mathcal{F}_{\underline{0}},\mathbb{P},L^1(\mathbb{R},\lambda))$. For every $\underline{i}\in\mathbb{Z}^d$, we define $Y_{\underline{i}}:=U_{\underline{i}}Y$ and for every $s\in\mathbb{R}$, $X_{\underline{i}}(s):=\mathbb{1}_{Y_{\underline{i}}\le s}-F(s)$, where $F(s)=\mathbb{P}(Y\le s)$. Let $S_{\underline{n}}(X)(s):=\sum_{i_1=1}^{n_1}\cdots\sum_{i_d=1}^{n_d}X_{\underline{i}}(s)$. We also denote by $F_{\underline{n}}$ the empirical distribution function of $(Y_{\underline{i}})_{\underline{i}\in\mathbb{Z}^d}$:

$$\forall s\in\mathbb{R}\quad F_{\underline{n}}(s):=\frac{1}{n_1\cdots n_d}\sum_{i_1=1}^{n_1}\cdots\sum_{i_d=1}^{n_d}\mathbb{1}_{Y_{\underline{i}}\le s}.$$

We are interested in deriving the asymptotic behavior of

$$\int_{\mathbb{R}}\left|\frac{1}{n_1\cdots n_d}S_{\underline{n}}(X)(s)\right|ds=\int_{\mathbb{R}}|F_{\underline{n}}(s)-F(s)|\,ds. \tag{5.1}$$

Note that (5.1) is the L^1-Wasserstein distance between F the distribution function of Y, and $F_{\underline{n}}$ the corresponding empirical distribution function.

The following result is a direct application of Theorem 4.1 combined with Corollary 4.3.

Corollary 5.1. *Assume that Y defined as above is regular and that*

$$\int_{\mathbb{R}}\sqrt{F(s)\,(1-F(s))}ds<\infty \tag{5.2}$$

and

$$\sum_{i_1=1}^{\infty}\cdots\sum_{i_d=1}^{\infty}\frac{1}{\sqrt{i_1\cdots i_d}}\int_{\mathbb{R}}\left\|\mathbb{P}\left(Y_{\underline{i}}\le s|\mathcal{F}_{\underline{0}}\right)-F(s)\right\|_2 ds<\infty. \tag{5.3}$$

Suppose, in addition, that at least one of the transformations $T_{(i)}$ for $1 \leq i \leq d$ is ergodic. Then Then $\left\{ s \mapsto \sqrt{n_1 \cdots n_d} \left(F_{\underline{n}}(s) - F(s) \right), s \in \mathbb{R} \right\}$ converges in L^1 to a centered Gaussian random variable G, with covariance operator $\mathcal{K}_\mu$ defined by: for every $f, g \in L^\infty(\mathbb{R}, \lambda)$:

$$\mathcal{K}_\mu(f,g) = \sum_{\underline{i} \in \mathbb{Z}^d} \mathbb{E}\left(\int_{\mathbb{R}} \int_{\mathbb{R}} f(s) g(t) \left(\mathbb{1}_{Y_{\underline{0}} \leq s} - F(s) \right) \left(\mathbb{1}_{Y_{\underline{i}} \leq t} - F(t) \right) ds dt \right).$$

In particular, as $\underline{n} \to \infty$,

$$\sqrt{n_1 \cdots n_d} \int_{\mathbb{R}} \left| F_{\underline{n}}(s) - F(s) \right| ds \Rightarrow \int_{\mathbb{R}} |G(s)| \, ds.$$

Note that by Remark 4.2, condition (5.2) implies that $Y \in \mathbb{G}(\mathcal{F}_{\underline{0}}, L^1(\mathbb{R}, \lambda))$. Next we give a sufficient condition for (5.3) to hold, in terms of dependence coefficients. With this aim, in the spirit of [9], define

$$\widetilde{\phi}(\underline{i}) = \sup_{s \in S} \left\| \mathbb{P}\left(Y_{\underline{i}} \leq s | \mathcal{F}_{\underline{0}} \right) - \mathbb{P}\left(Y_{\underline{i}} \leq s \right) \right\|_\infty . \tag{5.4}$$

As a direct application of [5, Prop 2.1], we get the following proposition:

Proposition 5.2. *Condition* (5.3) *holds as soon as* (5.2) *holds and*

$$\sum_{i_1=1}^{\infty} \cdots \sum_{i_d=1}^{\infty} \frac{1}{\sqrt{n_1 \cdots n_d}} \sqrt{\widetilde{\phi}(\underline{i})} < \infty. \tag{5.5}$$

Comment 5.3. *Note that by Item 3 in [9, Lemma 2], we get that if the distribution function F of $Y_{\underline{0}}$ is continuous then*

$$\widetilde{\phi}(\underline{i}) \leq K \left\| F(Y_{\underline{i}}) - F(Y_{\underline{i}}^*) \right\|_\infty , \tag{5.6}$$

where Y^* is a random element distributed as Y and independent of $\mathcal{F}_{\underline{0}}$, and $Y_{\underline{i}}^* := Y^* \circ T_{\underline{i}}$.

6 Ortho-martingale approximation in Banach spaces

The following result allows to derive a useful ortho-martingale approximation in 2-smooth Banach spaces. It is an extension to multidimensional index sets and to 2-smooth Banach spaces of the estimate (3.3) in [7]. It can also be viewed as an extension in several directions of [21, Th. 1 (ii)] and gives a more precise estimate than [7, Theorem 8]. Below $(\mathcal{F}_{\underline{i}})_{\underline{i} \in \mathbb{Z}^d}$ defined by $\mathcal{F}_{\underline{i}} = U_{\underline{i}} \mathcal{F}$ is assumed to be a completely commuting filtration and denote by $X_{\underline{i}} = U_{\underline{i}} X$. By $\min_i(\underline{n})$ we mean the infimum of all possible different products of size i constructed from $\{n_1, \ldots, n_d\}$. So, for instance, $\min_1(\underline{n}) = n_1 \wedge \cdots \wedge n_d$, $\min_2(\underline{n}) = n_1 n_2 \wedge \cdots \wedge n_1 n_d \wedge \cdots \wedge n_{d-1} n_d$, and so on.

Proposition 6.1. *Let $\mathcal{X}$ be a $(2, L)$-smooth Banach spaces. Let $p \geq 2$ and $X_{\underline{0}}$ be a regular r.v. with values in $\mathcal{X}$ such that $\| |X_{\underline{0}}|_{\mathcal{X}} \|_p < \infty$. Assume that*

$$\sum_{\underline{i} \in \mathbb{Z}^d} \| |P_{\underline{0}}(X_{\underline{i}})|_{\mathcal{X}} \|_p < \infty . \tag{6.1}$$

Then setting $D_{\underline{k}} = \sum_{\underline{i}\in\mathbb{Z}^d} P_{\underline{k}}(X_{\underline{i}})$, $(D_{\underline{k}})_{\underline{k}\in\mathbb{Z}^d}$ *are stationary* $\mathbb{L}^p(\mathcal{X})$ *ortho-martingale differences with respect to* $(\mathcal{F}_{\underline{k}})_{\underline{k}\in\mathbb{Z}^d}$ *and the corresponding ortho-martingale* $M_{\underline{n}} = \sum_{k_1=1}^{n_1}\cdots\sum_{k_d=1}^{n_d} D_{\underline{k}}$ *satisfies the following maximal inequality: there exists a positive constant* C *only depending on* (L,p) *such that for any positive integers* $n_1,\dots,n_d,m$ *and for any positive real* $\lambda \geq m$,

$$\prod_{i=1}^{d} n_i^{-1/2} \left\| \sup_{1\leq \underline{k}\leq \underline{n}} |S_{\underline{k}} - M_{\underline{k}}|_{\mathcal{X}} \right\|_p \leq C \sum_{i=1}^{d} \sum_{|k_i|>m} \sum_{\underline{k}\setminus\{k_i\}\in\mathbb{Z}^{d-1}} \left\| |P_{\underline{k}}(X_{\underline{0}})|_{\mathcal{X}} \right\|_p$$
$$+ C\lambda^2 \sum_{i=1}^{d-1} m^{2d-i} \Big(\frac{1}{\min_{\mathrm{i}}(\underline{\mathrm{n}})^{1/2}} + \frac{\lambda^{-3}}{\min_{\mathrm{i}}(\underline{\mathrm{n}})^{(\mathrm{p}-2)/(2\mathrm{p})}} \Big) \left(\| |X_{\underline{0}}|_{\mathcal{X}} \|_p + \| |D_{\underline{0}}|_{\mathcal{X}} \|_p \right)$$
$$+ C \sum_{i=1}^{d-1} m^{2d-i} \frac{1}{\min_{\mathrm{i}}(\underline{\mathrm{n}})^{(\mathrm{p}-2)/(2\mathrm{p})}} \Big(\| |X_{\underline{0}}|_{\mathcal{X}} \mathbf{1}_{\{|X_{\underline{0}}|_{\mathcal{X}} > \lambda \| |X_{\underline{0}}|_{\mathcal{X}} \|_p\}} \|_p + \| |D_{\underline{0}}|_{\mathcal{X}} \mathbf{1}_{\{|D_{\underline{0}}|_{\mathcal{X}} > \lambda \| |D_{\underline{0}}|_{\mathcal{X}} \|_p\}} \|_p \Big) . \quad (6.2)$$

The proof of this result is posponed to Section 7.4

7 Proofs

7.1 Proof of Theorem 3.4

The idea of this proof is essentially the same as the proof of [2, Prop 2.3], except for the convergence of finite dimensional laws.

We first prove (3.2). By the Caroli's strong (p,p) inequality for ortho-submartingales [14, Th. 2.3.1], we have

$$\left\| \max_{\underline{1}\leq \underline{k}\leq \underline{n}} |S_{\underline{k}}(D)|_{\mathcal{X}} \right\|_2^2 \leq 2^{2d} \left\| |S_{\underline{n}}(D)|_{\mathcal{X}} \right\|_2^2 .$$

Assume first that $\mathcal{X}$ is $(2,L)$-smooth, then applying (2.2) to each index, we derive

$$\left\| |S_{\underline{n}}(D)|_{\mathcal{X}} \right\|_2^2 \leq 2^d L^{2d} \sum_{i_1=1}^{n_1} \cdots \sum_{i_d=1}^{n_d} \mathbb{E}\left(|D_{\underline{i}}|_{\mathcal{X}}^2 \right) .$$

Recall that any 2-smooth Banach space is of type 2. Therefore (3.2) holds by stationarity and the fact that, in Banach spaces of type 2, the norms $\|\cdot\|_{\mathbb{G}(\mathcal{X})}$ and $\|\cdot\|_{2,\mathcal{X}}$ are equivalent (see [15, Prop 9.24]).

Suppose now that $\mathcal{X}$ is of cotype 2. Since D is assumed to be pregaussian, so is $S_{\underline{n}}(D)$. In addition, by the orthogonality of ortho-martingale increments, we get $G(S_{\underline{n}}(D)/\sqrt{n_1\cdots n_d}) = G(D)$. Hence by [15, Prop 9.25], we deduce that

$$\|S_{\underline{n}}(D)\|_{2,\mathcal{X}} \leq C \, \|G(S_{\underline{n}}(D))\|_{2,\mathcal{X}} = C\sqrt{n_1\cdots n_d} \, \|G(D)\|_{2,\mathcal{X}} \leq C\sqrt{n_1\cdots n_d} \, \|D\|_{\mathbb{G}(\mathcal{X})} .$$

This ends the proof of (3.2).

Now we prove the functional central limit theorem by first proving the tightness and then the convergence of finite dimensional laws. For the tightness, the idea in [2, Proof of Prop 3.2] also applies in higher dimension. For reader convenience, let us give the details.

Since $\mathcal{X}$ is separable, then $\sigma(D)$ is countably generated. Thus there exists an increasing filtration $(\mathcal{G}_m)_{m\in\mathbb{N}^*}$, such that $\mathcal{G}_m$ is finite for every $m \geq 1$ and $\sigma(D) = \bigcup_{m\geq 1} \mathcal{G}_m$. We set $D_m := \mathbb{E}\,(D|\mathcal{G}_m)$. Since $\mathcal{G}_m$ is finite, there exist $A_1^{(m)}, \cdots A_{k_m}^{(m)} \in \mathcal{G}_m$ and $x_1^{(m)}, \cdots x_{k_m}^{(m)} \in \mathcal{X}$ such that $D_m = \sum_{k=1}^{k_m} x_k^{(m)} \mathbb{1}_{A_k^{(m)}}$. Using [2, Lemma 2.3], we have

$$\|D_m - D\|_{\mathbb{G}(\mathcal{X})} \longrightarrow 0 \text{ as } m \to \infty.$$

Now, setting

$$\tilde{D}_m := P_{\underline{0}}(D_m)\,.$$

One can see that $(U_{\underline{i}}\tilde{D}_m)_{\underline{i}\in\mathbb{Z}^d}$ is a sequence of ortho-martingale differences with respect to the completely commuting filtration $(\mathcal{F}_{\underline{i}})_{\underline{i}\in\mathbb{Z}}$. Then using [2, Lemma 2.2] and the fact that $\mathbb{E}(D|\mathcal{F}_{-e_1}) = \cdots = \mathbb{E}(D|\mathcal{F}_{-e_d}) = 0$, we deduce that $(U_{\underline{i}}\tilde{D}_m)_{\underline{i}\geq\underline{1}}$ converges in $\mathbb{G}(\mathcal{X})$ to D.

By [3, Theorem 1], for every $m \geq 1$, $\left\{(n_1\cdots n_d)^{-1/2} S_{[\underline{nt}]}(\tilde{D}_m)\right\}_{\underline{t}\in[0,1]^d}$ is tight in $D([0,1],\mathcal{X})$. Indeed $\tilde{D}_m$ takes only a finite number of values and therefore we can work on the finite dimensional vector space generated by these values. Note that Theorem 1 in [3] is stated for reversed ortho-martingale differences, but it is also obviously true in case of ortho-martingales in the usual sense. Now, by (3.2), we deduce that

$$\begin{aligned}
&\left\| \sup_{\underline{0}\leq\underline{t}\leq\underline{1}} \left| \frac{1}{\sqrt{n_1\cdots n_d}} \left(S_{[\underline{nt}]}(\tilde{D}_m) - S_{[\underline{nt}]}(D) \right) \right|_{\mathcal{X}} \right\|_2 \\
&\qquad = \frac{1}{\sqrt{n_1\cdots n_d}} \left\| \max_{\underline{1}\leq\underline{k}\leq\underline{n}} \left| \left(S_{\underline{k}}(\tilde{D}_m) - S_{\underline{k}}(D) \right) \right|_{\mathcal{X}} \right\|_2 \\
&\qquad \leq C \left\| \tilde{D}_m - D \right\|_{\mathbb{G}(\mathcal{X})} \longrightarrow 0 \text{ as } m \to \infty\,.
\end{aligned}$$

Hence the tightness of $\left\{(n_1\cdots n_d)^{-1/2} S_{[\underline{nt}]}(D)\right\}_{\underline{t}\in[0,1]^d}$ follows (see [2, Lemma B.1]).

We turn now to the proof of the convergence of the finite dimensional laws. Recall the notation (3.1) for $T_{\underline{nt}}(D)$. When no confusion is possible, we will denote $T_{\underline{nt}}(D)$ by $T_{\underline{nt}}$. Let $\underline{t}^{(k)} = (t_1^{(k)}, \ldots, t_d^{(k)})$. Our aim is to prove that for any $m \geq 1$ and any $(0,0,\ldots,0) < \underline{t}^{(1)} < \ldots < \underline{t}^{(m)} \leq (1,1,\ldots,1)$,

$$\left(T_{\underline{nt}^{(1)}}, \ldots, T_{\underline{nt}^{(m)}}\right) \Rightarrow \left(W_{\underline{t}^{(1)}}, \ldots, W_{\underline{t}^{(m)}}\right), \tag{7.1}$$

where $(W_{\underline{t}})_{\underline{t}\in[0,1]^d}$ is a Brownian sheet with covariance operator $\mathcal{K}_D$. For reader convenience, we shall give the complete proof in case $d = 2$, noticing that the general case can be proved by induction. Since either $T_{(1)}$ or $T_{(2)}$ is ergodic, let us assume from now that $T_{(2)}$ is. Using the Cramer-Wold device, it is sufficient to prove that for any $m \geq 1$, any $(0,0) < \underline{t}^{(1)} < \cdots < \underline{t}^{(m)} \leq$

$(1,1)$ and any $x_1^*, \cdots x_m^* \in \mathcal{X}^*$,

$$\sum_{i=1}^{m} x_i^* \left(T_{\underline{nt}^{(i)}}\right) \Rightarrow \sum_{i=1}^{m} x_i^* \left(W_{\underline{t}^{(i)}}\right). \tag{7.2}$$

With this aim we shall use similar arguments as those developed in [3, Section 3.2] and, in a sake of clarity, we shall give most of the details. Notice first that $\sum_{i=1}^{m} x_i^* \left(T_{\underline{nt}^{(i)}}\right)$ can be written as a weighted sum over disjoint and adjacent rectangles. Hence proving (7.2) is equivalent to show that for any positive integer m, any $0 = t_0 < t_1 < \cdots < t_m \leq 1$, $0 = s_0 < s_1 < \cdots < s_m \leq 1$ and $(x_{i,j}^*)_{1 \leq i,j \leq m} \in \mathcal{X}^*$

$$\begin{aligned} V_{n_1 n_2} :=& \frac{1}{\sqrt{n_1 n_2}} \sum_{i=1}^{m} \sum_{j=1}^{m} x_{i,j}^* \left(\sum_{k=[n_1 t_{i-1}]+1}^{[n_1 t_i]} \sum_{\ell=[n_2 s_{j-1}]+1}^{[n_2 s_j]} D \circ T_{k,\ell} \right) \\ \Rightarrow & \sum_{i=1}^{m} \sum_{j=1}^{m} x_{i,j}^* \left(W_{t_i,s_j} + W_{t_{i-1},s_{j-1}} - W_{t_i,s_{j-1}} - W_{t_{i-1},s_j} \right), \end{aligned} \tag{7.3}$$

as $n_1 \wedge n_2 \to \infty$. Notice that the random variable on the right-hand side is distributed according to $\mathcal{N}(0, \sigma^2)$ with

$$\sigma^2 = \sum_{i=1}^{m} \sum_{j=1}^{m} (t_i - t_{i-1}) (s_j - s_{j-1}) \, \mathbb{E}\left(x_{i,j}^*(D)^2\right).$$

Clearly, it suffices to prove the convergence (7.3) when $n_1, n_2 \to \infty$ along any sequence $(n_r, N_r)_{r \geq 1}$. Hence, let us fix a sequence $(n_r, N_r)_{r \geq 1}$ such that $n_r, N_r \to \infty$ as $r \to \infty$. It remains to prove that

$$V_r = \frac{1}{\sqrt{n_r N_r}} \sum_{i=1}^{m} \sum_{j=1}^{m} x_{i,j}^* \left(\sum_{k=[n_r t_{i-1}]+1}^{[n_r t_i]} \sum_{\ell=[N - r s_{j-1}]+1}^{[N_r s_j]} D \circ T_{k,\ell} \right) \Rightarrow \mathcal{N}(0, \sigma^2). \tag{7.4}$$

As in [3], the proof of (7.4) is based on the usual central limit theorem for triangular arrays of martingale differences due McLeish (see [12, Theorem 3.2] for an easy reference). The rest of the proof then follows the lines of [3, Section 3.2] by noticing that $\sup_{1 \leq i,j \leq m} \left\| x_{i,j}^*(D) \right\|_2 < \infty$ and the following modification of [3, Lemma 4].

Lemma 7.1. *Let $\Delta_i = \sum_{j=1}^{m} (s_j - s_{j-1}) \mathbb{E}\left(x_{i,j}(D)^2\right)$ and $\varepsilon > 0$. If $T_{0,1}$ is ergodic, there exist integers $v > 0$ (large enough) and $p(v)$ (large enough), such that for every $n \geq p(v)$*

$$\left\| \frac{1}{v} \sum_{k=1}^{v} \left(\sum_{j=1}^{m} \frac{1}{\sqrt{n}} \sum_{\ell=[n s_{j-1}]+1}^{[n s_j]} x_{i,j}^*(D) \circ T_{k,\ell} \right)^2 - \Delta_i \right\|_1 < \varepsilon. \tag{7.5}$$

Proof of Lemma 7.1. Note first that for any fixed $i, j \in \mathbb{N}$ and $k \in \mathbb{Z}$, we have that $\left(x_{i,j}^*(D) \circ T_{k,\ell}\right)_{\ell \in \mathbb{Z}}$ is a sequence of martingale differences with respect to the filtration $\left(\mathcal{F}_\ell^{(2)}\right)_{\ell \in \mathbb{Z}}$. Let $d_{k,\ell}(i,j) =$

$x^*_{i,j}(D) \circ T_{k,\ell}$ and define

$$D_\ell(i,j) = (d_{1,\ell}(i,j), \ldots, d_{v,\ell}(i,j))^t .$$

Note that $(D_\ell(i,j))_{\ell \geq 1}$ is a stationary and ergodic (since $T_{(2)}$ is ergodic) sequence of $\mathbb{R}^v$-valued martingale differences. Therefore from the functional form of the central limit theorem for $\mathbb{R}^v$-valued stationary and ergodic $\mathbb{L}^2$ martingale differences, we get that for any positive integer m, and any $0 \leq s_1 < \ldots < s_m \leq 1$

$$\left(\frac{1}{\sqrt{n}} \sum_{\ell=1}^{[ns_1]} D_\ell(i,j), \ldots, \frac{1}{\sqrt{n}} \sum_{\ell=[ns_{m-1}]+1}^{[ns_m]} D_\ell(i,j) \right)$$

converges in distribution to $(G_1, \ldots, G_m)$, where $(G_u)_{1 \leq u \leq m}$ are independent and centered Gaussian random variables with respective covariance matrix

$$\mathbb{E}\left(G_\ell G_\ell^t\right) = (s_\ell - s_{\ell-1}) \mathbb{E}\left(D_1(i,j) D_1(i,j)^t\right) = (s_\ell - s_{\ell-1}) \left\| x^*_{i,j}(d) \right\|_2^2 Id,$$

and for any i, j, k, $\left(n^{-1} \max_{u \leq n} \left(\sum_{\ell=1}^u d_{k,\ell}(i,j) \right)^2 \right)_{n \geq 1}$ is uniformly integrable. In particular, we have

$$\left(\frac{1}{\sqrt{n}} \sum_{j=1}^m \sum_{\ell=[ns_{j-1}]+1}^{[ns_j]} x^*_{i,j}(d) \circ T_{k,\ell} \right)_{1 \leq k \leq v} \tag{7.6}$$

converges in distribution to $(N_k)_{1 \leq k \leq v}$ where N_k are i.i.d random variables with common distribution $\mathcal{N}\left(0, \sum_{j=1}^m (s_j - s_{j-1}) \mathbb{E}\left(x^*_{i,j}(d)^2\right)\right)$. Now using the notations

$$F_{i,k,n} = \sum_{j=1}^m \frac{1}{\sqrt{n}} \sum_{\ell=[ns_{j-1}]+1}^{[ns_j]} x^*_{i,j}(d) \circ T_{k,\ell} \text{ and } V_{i,k,n} = F^2_{i,k,n} - \Delta_i ,$$

note that $\Delta_i = \mathbb{E}\left(F^2_{i,k,n}\right) = \mathbb{E}\left(N_k^2\right)$. To soothe the notation, we will drop the index i in the rest of the proof. Let $\varepsilon > 0$, $M > 0$ and define

$$A = \mathbb{E}\left| \frac{1}{v} \sum_{k=1}^v V_{k,n} \mathbb{1}\left(\left| \frac{1}{v} \sum_{k=1}^v V_{k,n} \right| \leq \varepsilon \right) \right| , \; B = \mathbb{E}\left| \frac{1}{v} \sum_{k=1}^v V_{k,n} \mathbb{1}\left(\varepsilon < \left| \frac{1}{v} \sum_{k=1}^v V_{k,n} \right| \leq M \right) \right|$$

and

$$C = \mathbb{E}\left| \frac{1}{v} \sum_{k=1}^v V_{k,n} \mathbb{1}\left(\left| \frac{1}{v} \sum_{k=1}^v V_{k,n} \right| > M \right) \right| .$$

We have

$$\mathbb{E}\left| \frac{1}{v} \sum_{k=1}^v V_{k,n} \right| \leq A + B + C .$$

Clearly $A \leq \varepsilon$. Next,

$$B \leq M\mathbb{P}\left(\left|\frac{1}{v}\sum_{k=1}^{v} V_{k,n}\right| > \varepsilon\right)$$

Since $(F_{i,k,n})_{1\leq k\leq v} \Rightarrow (N_k)_{1\leq k\leq v}$, we get that, for any $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\frac{1}{v}\sum_{k=1}^{v} V_{k,n}\right| > \varepsilon\right) \to_{n\to\infty} \mathbb{P}\left(\left|\frac{1}{v}\sum_{k=1}^{v} N_k^2 - \Delta_i\right| > \varepsilon\right)$$

which converge to zero as $v \to \infty$ by the law of large numbers.

Let us now deal with C. Letting $h_M(x) = \left(|x| - \frac{M}{2}\right)_+$, where $x_+ = x\mathbb{1}_{x>0}$, and noticing that $|x|\,\mathbb{1}_{|x|>M} \leq 2h_M(x)$, we get

$$C \leq 2\mathbb{E}\left(h_M\left(\frac{1}{v}\sum_{k=1}^{v} V_{k,n}\right)\right).$$

Since h_M is a convex function, we deduce that

$$C \leq 2\frac{1}{v}\sum_{k=1}^{v} \mathbb{E}\left(h_M\left(V_{k,n}\right)\right).$$

But since for each i and k, $\left(F_{i,k,n}^2\right)_{n\geq 1}$ is an uniformly integrable family,

$$\lim_{M\to\infty} \limsup_{n\to\infty} \mathbb{E}\left(h_M(V_{k,n})\right) = 0.$$

So overall the lemma follows by letting $\varepsilon \to 0$. □

7.2 Proof of Theorem 4.1

By applying Proposition 6.1 to $X_0(s)$ and $\mathcal{X} = \mathbb{R}$, it follows that there exists a positive constant C only depending on (L, d) such that for any positive integers $n_1, \ldots, n_d, m$ and for any positive real $\lambda \geq m$,

$$
\begin{aligned}
&(n_1 \cdots n_d)^{-1/2} \int_S \| \max_{\underline{1} \le \underline{k} \le \underline{n}} S_{\underline{n}}(X(s)) - S_{\underline{n}}(D(s))\|_2 \mu(ds) \\
&\le C \sum_{i=1}^{d} \sum_{|k_i|>m} \sum_{\underline{k} \setminus \{k_i\} \in \mathbb{Z}^{d-1}} \int_S \|P_{\underline{k}}(X_{\underline{0}}(s))\|_2 \, \mu(ds) \\
&+ C\lambda^2 m^{2d} \Big(\sum_{i=1}^{d} \frac{1}{n_i^{1/2}} + \frac{1}{\lambda^3} \Big) \int_S \big(\|X_{\underline{0}}(s)\|_2 + \|D_{\underline{0}}(s)\|_2 \big) \, \mu(ds) \\
&+ C m^{2d} \int_S \Big(\|X_{\underline{0}} \mathbf{1}_{\{|X_{\underline{0}}(s)| > \lambda \|X_{\underline{0}}(s)\|_2\}}\|_2 + D_{\underline{0}}(s) \mathbf{1}_{\{|D_{\underline{0}}(s)| > \lambda \|D_{\underline{0}}(s)\|_2\}}\|_2 \Big) \mu(ds) \\
&:= I_1 + I_2 + I_3 \, .
\end{aligned}
$$

Letting first $\underline{n} \to \infty$, then $\lambda \to \infty$ and after $m \to \infty$, the R.H.S. is tending to zero. Indeed, I_1 is tending to zero as $m \to \infty$ by taking into account condition (4.2). I_2 is tending to zero by letting first $\underline{n} \to \infty$, then $\lambda \to \infty$ and by taking into account that (4.2) implies that $\int_S \|D_{\underline{0}}(s)\|_2 \mu(ds) < \infty$. Finally, I_3 is tending to zero by letting $\lambda \to \infty$ and using dominated convergence theorem and conditions (4.1) and (4.2). Hence (4.3) holds. Then the convergence in distribution (4.4) follows by using (4.3) together with Theorem 3.4. Next to prove (4.5), as in the proof of [3, Th 8] we use [3, Lemma 7]. Hence it suffices to prove that for any $f \in L^\infty(S, \mu)$ we have

$$
\sum_{\underline{i} \in \mathbb{Z}^d} \|P_{\underline{0}}(f(X_{\underline{i}}))\|_2 < \infty \, ,
$$

which is implied by (4.2). Note that [3, Lemma 7] is stated for reversed ortho-martingales, but it holds also for ortho-martingales in the usual sense. □

7.3 Proof of Corollary 4.3

For the sake of clarity, we shall give the complete proof in the case $d = 2$ (the case $d > 2$ can be proved with similar arguments). We shall need the following lemma, whose proof follows by applying two times [6, Lemma 6.1].

Lemma 7.2. *For any double indexed sequence $(u_{i,j})_{i,j \in \mathbb{N}}$ of non-negative numbers, there exists a positive constant C such that*

$$
\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} u_{i,j} \le C \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \left(\frac{1}{ij} \sum_{k=i}^{\infty} \sum_{\ell=j}^{\infty} u_{k,\ell}^2 \right)^{1/2} .
$$

From Lemma 7.2 with $u_{i,j} = \|P_{-i,-j}(X_{0,0}(s))\|_2$, we get:

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\int_S \|P_{-i,-j}(X_{0,0}(s))\|_2\,\mu(ds) \le C\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\int_S \left(\frac{1}{ij}\sum_{k=i}^{\infty}\sum_{\ell=j}^{\infty}\|P_{-k,-\ell}(X_{0,0}(s))\|_2^2\right)^{1/2}\mu(ds)\,.$$

But by orthogonality and regularity

$$\sum_{k=i}^{\infty}\sum_{\ell=j}^{\infty}\|P_{-k,-\ell}(X_{0,0}(s))\|_2^2 = \left\|\sum_{k=i}^{\infty}\sum_{\ell=j}^{\infty}P_{-k,-\ell}(X_{0,0}(s))\right\|_2^2 = \|\mathbb{E}\left(X_{i,j}(s)|\mathcal{F}_{0,0}\right)\|\,.$$

Hence

$$\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\int_S \|P_{-i,-j}(X_{0,0}(s))\|_2\,\mu(ds) \le C\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\int_S \frac{1}{\sqrt{ij}}\,\|\mathbb{E}\left(X_{i,j}(s)|\mathcal{F}_{0,0}\right)\|\,\mu(ds)\,.$$

which is finite by condition (4.6) in case $d = 2$. It remains to prove that:

$$\sum_{i=1}^{\infty}\sum_{j=0}^{\infty}\int_S \|P_{0,0}(X_{i,-j}(s))\|_2\,\mu(ds) < \infty\,,\ \sum_{i=0}^{\infty}\sum_{j=1}^{\infty}\int_S \|P_{0,0}(X_{-i,j}(s))\|_2\,\mu(ds) < \infty\,, \tag{7.7}$$

and

$$\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\int_S \|P_{0,0}(X_{-i,-j}(s))\|_2\,\mu(ds) < \infty\,. \tag{7.8}$$

By Lemma 7.2 again,

$$\sum_{i=1}^{\infty}\sum_{j=0}^{\infty}\int_S \|P_{0,0}(X_{i,-j}(s))\|_2\,\mu(ds) \le C\sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\int_S \left(\frac{1}{ij}\sum_{k=i+1}^{\infty}\sum_{\ell=j}^{\infty}\|P_{k,-\ell}(X_{0,0}(s))\|_2^2\right)^{1/2}\mu(ds)\,. \tag{7.9}$$

But, for any $j \ge 0$, by stationarity,

$$\sum_{k=i+1}^{\infty}\sum_{\ell=j}^{\infty}\|P_{k,-\ell}(X_{0,0}(s))\|_2^2 \le \sum_{k=i+1}^{\infty}\sum_{\ell=-j}^{\infty}\|P_{k,-\ell}(X_{0,0}(s))\|_2^2 \le \sum_{k=1}^{\infty}\sum_{\ell=-\infty}^{0}\|P_{k,\ell}(X_{-i,-j}(s))\|_2^2\,.$$

Next, by orthogonality,

$$\sum_{k=1}^{\infty}\sum_{\ell=-\infty}^{0}\|P_{k,\ell}(X_{-i,-j}(s))\|_2^2 = \left\|\sum_{k=1}^{\infty}\sum_{\ell=-\infty}^{0}P_{k,\ell}(X_{-i,-j}(s))\right\|_2^2\,.$$

But, note that,

$$\begin{aligned} X_{-i,-j}(s) - \mathbb{E}(X_{-i,-j}(s)|\mathcal{F}_{0,0}) &= \sum_{k=1}^{\infty}\sum_{\ell=1}^{\infty} P_{k,\ell}(X_{-i,-j}(s)) \\ &\quad + \sum_{k=-\infty}^{0}\sum_{\ell=1}^{\infty} P_{k,\ell}(X_{-i,-j}(s)) + \sum_{k=1}^{\infty}\sum_{\ell=-\infty}^{0} P_{k,\ell}(X_{-i,-j}(s)) \\ &= I_1 + I_2 + I_3\,. \end{aligned}$$

Hence, by orthogonality, it follows that

$$\|X_{-i,-j}(s) - \mathbb{E}(X_{-i,-j}(s)|\mathcal{F}_{0,0})\|_2^2 = \|I_1\|_2^2 + \|I_2\|_2^2 + \|I_3\|_2^2\,.$$

So, overall,

$$\sum_{k=i+1}^{\infty}\sum_{\ell=j}^{\infty} \|P_{k,-\ell}(X_{0,0}(s))\|_2^2 \le \|I_3\|_2^2 \le \|X_{-i,-j}(s) - \mathbb{E}(X_{-i,-j}(s)|\mathcal{F}_{0,0})\|_2^2\,. \tag{7.10}$$

Therefore, starting from (7.9) and taking into account (7.10), it follows that the first part of (7.7) holds provided condition (4.7) (in case $d = 2$) is assumed. Using similar arguments, we derive that the second part of (7.7) as well as (7.8) are satisfied provided (4.7) holds. This ends the proof of the corollary in case $d = 2$. □

7.4 Proof of Proposition 6.1

In the proof the notation $a \ll b$ means that there exists a universal constant C (here only depending on (L, p)) such that $a \le Cb$.

For reader's convenience, let us consider the case $d = 2$. In this situation (6.1) reads as

$$\sum_{i,j\in\mathbb{Z}} \||P_{0,0}(X_{i,j})|_{\mathcal{X}}\|_p < \infty\,. \tag{7.11}$$

Then $D_{k,\ell} = \sum_{i,j\in\mathbb{Z}} P_{k,\ell}(X_{i,j})$. $(D_{k,\ell})_{k,\ell}$ is a stationary $\mathbb{L}^p(\mathcal{X})$ ortho-martingale differences with respect to $(\mathcal{F}_{k,\ell})_{(k,\ell)\in\mathbb{Z}^2}$ and the corresponding ortho-martingale $M_{n_1,n_2} = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} D_{i,j}$. Next (6.2) reads as: there exists a positive constant C only depending on (L, p) such that for any positive integers n_1, n_2, m and for any positive real $\lambda \ge m$,

$$
\begin{aligned}
&(n_1 n_2)^{-1/2}\Big\| \sup_{\substack{1\le k\le n_1\\ 1\le \ell\le n_2}} \big|S_{k,\ell}-M_{k,\ell}\big|_{\mathcal{X}}\Big\|_p \\
&\qquad \le C\sum_{|u|>m, v\in\mathbb{Z}} \||P_{u,v}(X_{0,0})|_{\mathcal{X}}\|_p + C\sum_{u\in\mathbb{Z},|v|>m} \||P_{u,v}(X_{0,0})|_{\mathcal{X}}\|_p \\
&\qquad + Cm^3\lambda^2\Big(\frac{1}{\sqrt{n_1\wedge n_2}} + \frac{\lambda^{-3}}{(n_1\wedge n_2)^{(p-2)/(2p)}}\Big)\big(\||X_{0,0}|_{\mathcal{X}}\|_p + \||D_{0,0}|_{\mathcal{X}}\|_p\big) \\
&\quad + C\frac{m^3}{(n_1\wedge n_2)^{(p-2)/(2p)}}\Big(\big\||X_{0,0}|_{\mathcal{X}}\mathbf{1}_{\{|X_{0,0}|_{\mathcal{X}}>\lambda\||X_{0,0}|_{\mathcal{X}}\|_p\}}\big\|_p + \big\||D_{0,0}|_{\mathcal{X}}\mathbf{1}_{\{|D_{0,0}|_{\mathcal{X}}>\lambda\||D_{0,0}|_{\mathcal{X}}\|_p\}}\big\|_p\Big)\Big\}.
\end{aligned}
\tag{7.12}
$$

To prove (7.12), we proceed as follows. Since $X_{0,0}$ is regular, we can write $X_{0,0}=\sum_{i,j\in\mathbb{Z}} P_{i,j}(X_{0,0})$. Let m be a fixed positive integer, and define

$$
\theta^{(m)}_{0,0} = \sum_{k=0}^{2m-2}\sum_{\ell=0}^{2m-2}\sum_{i=k-m+1}^{m-1}\sum_{j=\ell-m+1}^{m-1} P_{i,j}(X_{k,\ell}) \ \text{ and } \ \theta^{(m)}_{u,v} = \theta^{(m)}_{0,0}\circ T_{u,v}\,.
$$

By simple algebra we have

$$
\begin{aligned}
I &:= \theta^{(m)}_{0,0} - \theta^{(m)}_{1,0} - \theta^{(m)}_{0,1} + \theta^{(m)}_{1,1} \\
&= \sum_{k=-m+1}^{m-1}\sum_{\ell=-m+1}^{m-1} P_{k,\ell}(X_{0,0}) - \sum_{i=-m+1}^{m-1}\sum_{v=1}^{2m-1} P_{i,m}(X_{0,v}) \\
&\quad - \sum_{u=1}^{2m-1}\sum_{j=-m+1}^{m-1} P_{m,j}(X_{u,0}) + \sum_{u=1}^{2m-1}\sum_{v=1}^{2m-1} P_{m,m}(X_{u,v})\,.
\end{aligned}
$$

On another hand setting

$$
g^{(m)}_{0,0} = \sum_{u=1}^{2m-1}\sum_{b=0}^{2m-2}\sum_{\ell=b-m+1}^{m-1} P_{m,\ell}(X_{u,b}) \ \text{ and } \ g^{(m)}_{i,j} = g^{(m)}_{0,0}\circ T_{i,j}\,,
$$

and

$$
h^{(m)}_{0,0} = \sum_{v=1}^{2m-1}\sum_{a=0}^{2m-2}\sum_{k=a-m+1}^{m-1} P_{k,m}(X_{a,v}) \ \text{ and } \ h^{(m)}_{i,j} = h^{(m)}_{0,0}\circ T_{i,j}\,,
$$

we have

$$
\sum_{u=1}^{2m-1}\left(\sum_{j=-m+1}^{m-1} P_{m,j}(X_{u,0}) - \sum_{v=1}^{2m-1} P_{m,m}(X_{u,v})\right) = g^{(m)}_{0,0} - g^{(m)}_{0,1}
$$

and

$$\sum_{v=1}^{2m-1}\left(\sum_{i=-m+1}^{m-1}P_{i,m}(X_{0,v})-\sum_{k=1}^{2m-1}P_{m,m}(X_{k,v})\right)=h_{0,0}^{(m)}-h_{1,0}^{(m)}\,.$$

So, overall, for any positive integer m, the following decomposition is valid:

$$\begin{aligned}X_{0,0}-D_{0,0}\circ T_{m,m}:=I+(g_{0,0}^{(m)}-g_{0,1}^{(m)}(u))+(h_{0,0}^{(m)}-h_{1,0}^{(m)})\\+Y_{0,0}^{(m)}-Z_{0,0}^{(m)}\,,\end{aligned}\tag{7.13}$$

where we used the following notation:

$$Y_{0,0}^{(m)}:=\sum_{(k,\ell)\in\mathbb{Z}^2\setminus[-m+1,m-1]^2}P_{k,\ell}(X_{0,0})\text{ and }Z_{0,0}^{(m)}:=\sum_{(u,v)\in\mathbb{Z}^2\setminus[1,2m-1]^2}P_{m,m}(X_{u,v})\,.$$

Let $R_{k,\ell}=S_{k,\ell}-M_{k,\ell}$. We then derive that

$$\begin{aligned}R_{k,\ell}=-\sum_{i=1}^{k}\sum_{j=1}^{\ell}\big(D_{i,j}-D_{i,j}\circ T_{m,m}\big)+(\theta_{1,1}^{(m)}-\theta_{k+1,1}^{(m)}-\theta_{1,\ell+1}^{(m)}+\theta_{k+1,\ell+1}^{(m)})\\+\sum_{i=1}^{k}(g_{i,1}^{(m)}-g_{i,\ell+1}^{(m)})+\sum_{j=1}^{\ell}(h_{1,j}^{(m)}-h_{k+1,j}^{(m)})\\+\sum_{i=1}^{k}\sum_{j=1}^{\ell}Y_{0,0}^{(m)}\circ T_{i,j}-\sum_{i=1}^{k}\sum_{j=1}^{\ell}Z_{0,0}^{(m)}\circ T_{i,j}\,.\end{aligned}$$

Therefore

$$\begin{aligned}\Big\|\max_{\substack{1\le k\le n_1\\1\le\ell\le n_2}}\big|R_{k,\ell}\big|_{\mathcal{X}}\Big\|_p\le\Big\|\max_{\substack{1\le k\le n_1\\1\le\ell\le n_2}}\Big|\sum_{i=1}^{k}\sum_{j=1}^{\ell}\big(D_{i,j}-D_{i,j}\circ T_{m,m}\big)\Big|_{\mathcal{X}}\Big\|_p+4\Big\|\max_{\substack{1\le k\le n_1+1\\1\le\ell\le n_2+1}}\big|\theta_{k,\ell}^{(m)}\big|_{\mathcal{X}}\Big\|_p\\+2\Big\|\max_{\substack{1\le k\le n_1\\1\le\ell\le n_2+1}}\Big|\sum_{i=1}^{k}g_{i,\ell}^{(m)}\Big|_{\mathcal{X}}\Big\|_p+2\Big\|\max_{\substack{1\le k\le n_1+1\\1\le\ell\le n_2}}\Big|\sum_{j=1}^{\ell}h_{k,j}^{(m)}\Big|_{\mathcal{X}}\Big\|_p\\+\Big\|\max_{\substack{1\le k\le n_1\\1\le\ell\le n_2}}\Big|\sum_{i=1}^{k}\sum_{j=1}^{\ell}Y_{0,0}^{(m)}\circ T_{i,j}\Big|_{\mathcal{X}}\Big\|_p+\Big\|\max_{\substack{1\le k\le n_1\\1\le\ell\le n_2}}\Big|\sum_{i=1}^{k}\sum_{j=1}^{\ell}Z_{0,0}^{(m)}\circ T_{i,j}\Big|_{\mathcal{X}}\Big\|_p\,.\end{aligned}\tag{7.14}$$

Let $Y_{i,j}^{(m)}=Y_{0,0}^{(m)}\circ T_{i,j}$. Since $X_{0,0}$ is regular,

$$Y_{i,j}^{(m)}=\sum_{u,v\in\mathbb{Z}}P_{u-i,v-j}(Y_{i,j}^{(m)})\,.$$

Hence

$$\Big\| \max_{\substack{1\le k\le n_1\\ 1\le \ell\le n_2}} \Big| \sum_{i=1}^{k}\sum_{j=1}^{\ell} Y_{i,j}^{(m)}\Big|_{\mathcal{X}}\Big\|_p \le \sum_{u,v\in\mathbb{Z}} \Big\| \max_{\substack{1\le k\le n_1\\ 1\le \ell\le n_2}} \Big| \sum_{i=1}^{k}\sum_{j=1}^{\ell} P_{u-i,v-j}(Y_{i,j}^{(m)})\Big|_{\mathcal{X}}\Big\|_p .$$

Denoting by $U_{k,\ell} = \Big|\sum_{i=1}^{k}\sum_{j=1}^{\ell} P_{u-i,v-j}(Y_{i,j}^{(m)})\Big|_{\mathcal{X}}$, note that $(U_{k,\ell})_{k,\ell}$ is a ortho-submartinagle w.r.t. the completely commuting filtration $(\mathcal{F}_{k,\ell})_{k,\ell}$. By [14, Prop. 2.2.1], it follows that $(\max_{1\le k\le n_1} U_{k,\ell})_{\ell\ge 1}$ and $(\max_{1\le \ell\le n_2} U_{k,\ell})_{k\ge 1}$ are both one parameter submartingales. Therefore, for any fixed pair (n_1,n_2) of natural numbers, applying twice Doob's maximal inequality, we get

$$\Big\| \max_{\substack{1\le k\le n_1\\ 1\le \ell\le n_2}} \Big| \sum_{i=1}^{k}\sum_{j=1}^{\ell} Y_{i,j}^{(m)}\Big|_{\mathcal{X}}\Big\|_p \le \Big(\frac{p}{p-1}\Big)^2 \sum_{u,v\in\mathbb{Z}} \Big\|\Big| \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} P_{u-i,v-j}(Y_{i,j}^{(m)})\Big|_{\mathcal{X}}\Big\|_p .$$

Next, since $\mathcal{X}$ is $(2,L)$-smooth, using twice [18, Th. 2.6] (see also [8, Th. 2.2]), we get

$$\begin{aligned}\Big\|\Big| \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} P_{u-i,v-j}(Y_{i,j}^{(m)})\Big|_{\mathcal{X}}\Big\|_p^2 &\ll \sum_{i=1}^{n_1}\Big\|\Big| \sum_{j=1}^{n_2} P_{u-i,v-j}(Y_{i,j}^{(m)})\Big|_{\mathcal{X}}\Big\|_p^2\\ &\ll \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \big\||P_{u-i,v-j}(Y_{i,j}^{(m)})|_{\mathcal{X}}\big\|_p^2 .\end{aligned}$$

Hence, by stationarity,

$$\Big\| \max_{\substack{1\le k\le n_1\\ 1\le \ell\le n_2}} \Big| \sum_{i=1}^{k}\sum_{j=1}^{\ell} Y_{i,j}^{(m)}\Big|_{\mathcal{X}}\Big\|_p \ll \sqrt{n_1 n_2}\sum_{u,v\in\mathbb{Z}} \big\||P_{u,v}(Y_{0,0}^{(m)})|_{\mathcal{X}}\big\|_p .$$

Since $P_{u,v}P_{k,\ell}(\cdot)=0$ for $(u,v)\neq(i,j)$, it follows that

$$\begin{aligned}&\Big\| \max_{\substack{1\le k\le n_1\\ 1\le \ell\le n_2}} \Big| \sum_{i=1}^{k}\sum_{j=1}^{\ell} Y_{i,j}^{(m)}\Big|_{\mathcal{X}}\Big\|_p\\ &\qquad\ll \sqrt{n_1 n_2}\Big(\sum_{|u|>m}\sum_{v\in\mathbb{Z}} \big\||P_{u,v}(X_{0,0})|_{\mathcal{X}}\big\|_p + \sum_{u\in\mathbb{Z}}\sum_{|v|>m} \big\||P_{u,v}(X_{0,0})|_{\mathcal{X}}\big\|_p\Big) . \quad (7.15)\end{aligned}$$

Similarly, we get

$$\Big\| \max_{\substack{1\le k\le n_1\\ 1\le \ell\le n_2}} \Big| \sum_{i=1}^{k}\sum_{j=1}^{\ell} Z_{i,j}^{(m)}\Big|_{\mathcal{X}}\Big\|_p \ll \sqrt{n_1 n_2}\sum_{(u,v)\in\mathbb{Z}^2\setminus[1,2m-1]^2} \big\||P_{m,m}(X_{u,v})|_{\mathcal{X}}\big\|_p ,$$

implying that

$$\Big\| \max_{\substack{1\le k\le n_1\\ 1\le \ell\le n_2}} \Big| \sum_{i=1}^{k}\sum_{j=1}^{\ell} Z_{i,j}^{(m)}\Big|_{\mathcal{X}}\Big\|_p \ll \sqrt{n_1 n_2}\Big(\sum_{|u|>m}\sum_{v\in\mathbb{Z}} \||P_{u,v}(X_{0,0})|_{\mathcal{X}}\|_p + \sum_{u\in\mathbb{Z}}\sum_{|v|>m} \||P_{u,v}(X_{0,0})|_{\mathcal{X}}\|_p\Big)\,. \quad (7.16)$$

We handle now the first term in the R.H.S. of decomposition (7.14). Note first that

$$\sum_{i=1}^{k}\sum_{j=1}^{\ell}\big(D_{i,j}-D_{i,j}\circ T_{m,m}\big) = \sum_{i=1}^{m}\sum_{j=1}^{\ell} D_{i,j} - \sum_{i=k+1}^{k+m}\sum_{j=m+1}^{\ell+m} D_{i,j} + \sum_{i=m+1}^{k}\sum_{j=1}^{m} D_{i,j} - \sum_{i=m+1}^{k}\sum_{j=\ell+1}^{\ell+m} D_{i,j}\,. \quad (7.17)$$

Since for j fixed $(D_{i,j})_j$ is a martingale differences sequence, $\big(\big|\sum_{j=1}^{\ell} D_{i,j}\big|_{\mathcal{X}}\big)_\ell$ is a submartingale. Then, by stationarity and Doob's maximal inequality,

$$\Big\| \sum_{i=1}^{m} \max_{1\le \ell\le n_2} \Big| \sum_{j=1}^{\ell} D_{i,j}\Big|_{\mathcal{X}}\Big\|_p \le m\Big(\frac{p}{p-1}\Big)\Big\| \sum_{j=1}^{n_2} D_{0,j}\Big|_{\mathcal{X}}\Big\|_p\,.$$

Next, using that $\mathcal{X}$ is $(2,L)$-smooth and [18, Th. 2.6] , we derive

$$\Big\| \sum_{i=1}^{m} \max_{1\le \ell\le n_2} \Big| \sum_{j=1}^{\ell} D_{i,j}\Big|_{\mathcal{X}}\Big\|_p^2 \ll m^2 \sum_{j=1}^{n_2} \||D_{0,j}|_{\mathcal{X}}\|_p^2 \ll m^2 n_2 \||D_{0,0}|_{\mathcal{X}}\|_p^2\,. \quad (7.18)$$

Similarly, we obtain

$$\Big\| \sum_{j=1}^{m} \max_{m+1\le k\le n_1} \Big| \sum_{i=m+1}^{k} D_{i,j}\Big|_{\mathcal{X}}\Big\|_p^2 \ll m^2 n_1 \||D_{0,0}|_{\mathcal{X}}\|_p^2\,. \quad (7.19)$$

To deal with the quantity coming from the second term in the R.H.S. of (7.17), we first note

that for any positive real A,

$$\Big\| \max_{\substack{1\le k\le n_1\\ 1\le \ell\le n_2}} \Big| \sum_{i=k+1}^{k+m}\sum_{j=m+1}^{\ell+m} D_{i,j}\Big|_{\mathcal{X}} \Big\|_p$$
$$\le \Big\| \max_{\substack{1\le k\le n_1\\ 1\le \ell\le n_2}} \sum_{i=k+1}^{k+m} \Big| \sum_{j=m+1}^{\ell+m} D_{i,j}\Big|_{\mathcal{X}} \mathbf{1}_{\{\max_{1\le \ell\le n_2} |\sum_{j=m+1}^{\ell+m} D_{i,j}|_{\mathcal{X}}\le 4A\sqrt{n_2}\}} \Big\|_p$$
$$+ \Big\| \max_{\substack{1\le k\le n_1\\ 1\le \ell\le n_2}} \sum_{i=k+1}^{k+m} \Big| \sum_{j=m+1}^{\ell+m} D_{i,j}\Big|_{\mathcal{X}} \mathbf{1}_{\{\max_{1\le \ell\le n_2} |\sum_{j=m+1}^{\ell+m} D_{i,j}|_{\mathcal{X}}> 4A\sqrt{n_2}\}} \Big\|_p .$$

The first term in the R.H.S. is less than $4mA\sqrt{n_2}$ whereas to deal with the second one we note that, by stationarity,

$$\Big\| \max_{\substack{1\le k\le n_1\\ 1\le \ell\le n_2}} \sum_{i=k+1}^{k+m} \Big| \sum_{j=m+1}^{\ell+m} D_{i,j}\Big|_{\mathcal{X}} \mathbf{1}_{\{\max_{1\le \ell\le n_2} |\sum_{j=m+1}^{\ell+m} D_{i,j}|_{\mathcal{X}}> 4A\sqrt{n_2}\}} \Big\|_p^p$$
$$\le n_1 m^p \Big\| \max_{1\le \ell\le n_2} \Big| \sum_{j=m+1}^{\ell+m} D_{0,j}\Big|_{\mathcal{X}} \mathbf{1}_{\{\max_{1\le \ell\le n_2} |\sum_{j=m+1}^{\ell+m} D_{0,j}|_{\mathcal{X}}> 4A\sqrt{n_2}\}} \Big\|_p^p .$$

But, by [17, Corollary 2.10] and stationarity,

$$\Big\| \max_{1\le \ell\le n_2} \Big| \sum_{j=m+1}^{\ell+m} D_{0,j}\Big|_{\mathcal{X}} \mathbf{1}_{\{\max_{1\le \ell\le n_2} |\sum_{j=m+1}^{\ell+m} D_{0,j}|_{\mathcal{X}}> 4A\sqrt{n_2}\}} \Big\|_p^p$$
$$\le 2^p \Big(\frac{p}{p-1}\Big) \Big\| \Big| \sum_{j=1}^{n_2} D_{0,j}\Big|_{\mathcal{X}} \mathbf{1}_{\{|\sum_{j=1}^{n_2} D_{0,j}|_{\mathcal{X}}> 2A\sqrt{n_2}\}} \Big\|_p^p .$$

Let B be a positive real and define

$$d'_j = D_{0,j}\mathbf{1}_{\{|D_{0,j}|_{\mathcal{X}}\le B\}} - \mathbb{E}(D_{0,j}\mathbf{1}_{\{|D_{0,j}|_{\mathcal{X}}\le B\}}|\mathcal{F}_{0,j-1}) \,,\; M'_{n_2} = \sum_{j=1}^{n_2} d'_j$$

and

$$d''_j = D_{0,j}\mathbf{1}_{\{|D_{0,j}|_{\mathcal{X}}> B\}} - \mathbb{E}(D_{0,j}\mathbf{1}_{\{|D_{0,j}|_{\mathcal{X}}> B\}}|\mathcal{F}_{0,j-1}) \,,\; M''_{n_2} = \sum_{j=1}^{n_2} d''_j \,.$$

Observe that $M_{n_2} = \sum_{j=1}^{n_2} D_{0,j} = M'_{n_2} + M''_{n_2}$ and that for any nonnegative reals a, b and ε,

$(a+b)^p \mathbf{1}_{\{a+b\geq 2\varepsilon\}} \leq 2^p a^p \mathbf{1}_{\{a\geq\varepsilon\}} + 2^p b^p \mathbf{1}_{\{b\geq\varepsilon\}}$. Therefore

$$\Big\| \Big| \sum_{j=1}^{n_2} D_{0,j}\Big|_{\mathcal{X}} \mathbf{1}_{\{|\sum_{j=1}^{n_2} D_{0,j}|_{\mathcal{X}} > 2A\sqrt{n_2}\}} \Big\|_p^p \leq 2^p \big\| |M'_{n_2}|_{\mathcal{X}} \mathbf{1}_{\{|M'_{n_2}|_{\mathcal{X}} > A\sqrt{n_2}\}} \big\|_p^p + 2^p \big\| |M''_{n_2}|_{\mathcal{X}} \mathbf{1}_{\{|M''_{n_2}|_{\mathcal{X}} > A\sqrt{n_2}\}} \big\|_p^p .$$

Using that $\mathcal{X}$ is $(2,L)$-smooth, [18, Th. 2.6] and stationarity, we get

$$\big\| |M''_{n_2}|_{\mathcal{X}} \big\|_p^2 \ll n_2 \| |D_{0,0}|_{\mathcal{X}} \mathbf{1}_{\{|D_{0,0}|_{\mathcal{X}} > B\}} \|_p^2 .$$

Next, using again that $\mathcal{X}$ is $(2,L)$-smooth and [18, Th. 2.6], we derive

$$\big\| |M'_{n_2}|_{\mathcal{X}} \mathbf{1}_{\{|M'_{n_2}|_{\mathcal{X}} > A\sqrt{n_2}\}} \big\|_p^p \leq \frac{1}{A^p (n_2)^{p/2}} \mathbb{E}\big(|M'_{n_2}|_{\mathcal{X}}^{2p} \big) \ll \frac{1}{A^p (n_2)^{p/2}} \Big(\sum_{j=1}^{n_2} \| |d'_j|_{\mathcal{X}} \|_{2p}^2 \Big)^p$$
$$\ll \frac{(n_2)^{p/2}}{A^p} \| |D_{0,0} \mathbf{1}_{\{|D_{0,0}|_{\mathcal{X}} \leq B\}}|_{\mathcal{X}} \|_{2p}^{2p} \ll \frac{(n_2)^{p/2} B^p}{A^p} \| |D_{0,0}|_{\mathcal{X}} \|_p^p .$$

So, overall, for any positive reals A and B,

$$\Big\| \max_{\substack{1\leq k\leq n_1 \\ 1\leq \ell\leq n_2}} \Big| \sum_{i=k+1}^{k+m} \sum_{j=m+1}^{\ell+m} D_{i,j}\Big|_{\mathcal{X}} \Big\|_p \ll mA\sqrt{n_2} + mn_1^{1/p}\sqrt{n_2} \| |D_{0,0}|_{\mathcal{X}} \mathbf{1}_{\{|D_{0,0}|_{\mathcal{X}} > B\}} \|_p + m\frac{n_1^{1/p}\sqrt{n_2} B}{A} \| |D_{0,0}|_{\mathcal{X}} \|_p . \quad (7.20)$$

Similarly, we get

$$\Big\| \max_{\substack{1\leq k\leq n_1 \\ 1\leq \ell\leq n_2}} \Big| \sum_{i=m+1}^{k} \sum_{j=\ell+1}^{\ell+m} D_{i,j}\Big|_{\mathcal{X}} \Big\|_p \ll mA\sqrt{n_1} + mn_2^{1/p}\sqrt{n_1} \| |D_{0,0}|_{\mathcal{X}} \mathbf{1}_{\{|D_{0,0}|_{\mathcal{X}} > B\}} \|_p + m\frac{n_2^{1/p}\sqrt{n_1} B}{A} \| |D_{0,0}|_{\mathcal{X}} \|_p . \quad (7.21)$$

Let $\lambda > 0$. Starting from (7.17), taking into account (7.18)-(7.21) and selecting $A = \lambda^2 \| |D_{0,0}|_{\mathcal{X}} \|_p$ and $B = A/\lambda$, we derive

$$\Big\| \max_{\substack{1\leq k\leq n_1 \\ 1\leq \ell\leq n_2}} \Big| \sum_{i=1}^{k} \sum_{j=1}^{\ell} \big(D_{i,j} - D_{i,j} \circ T_{m,m} \big) \Big|_{\mathcal{X}} \Big\|_p \ll m(\lambda^2+1)(\sqrt{n}_1 + \sqrt{n}_2) \| |D_{0,0}|_{\mathcal{X}} \|_p$$
$$+ m(n_1^{1/p} n_2^{1/2} + n_1^{1/2} n_2^{1/p}) \| |D_{0,0}|_{\mathcal{X}} \mathbf{1}_{\{|D_{0,0}|_{\mathcal{X}} > \lambda \| |D_{0,0}|_{\mathcal{X}} \|_p\}} \|_p$$
$$+ \frac{m(n_1^{1/p} n_2^{1/2} + n_1^{1/2} n_2^{1/p})}{\lambda} \| |D_{0,0}|_{\mathcal{X}} \|_p . \quad (7.22)$$

Next note that, for any positive real M,

$$\Big\| \max_{\substack{1\le u\le n_1\\ 1\le v\le n_2}} \big|\theta^{(m)}_{u,v}\big|_{\mathcal{X}} \Big\|_p \le 4(2m-1)^4 M + (n_1n_2)^{1/p} \sum_{k=0}^{2m-2}\sum_{\ell=0}^{2m-2}\sum_{i=-m+1}^{m-1}\sum_{j=-m+1}^{m-1} \||P_{i,j}(X_{k,\ell}\mathbf{1}_{\{|X_{k,\ell}|_{\mathcal{X}}>M\}})|_{\mathcal{X}}\|_p \,.$$

Hence, by stationarity, setting $M=\lambda\||X_{0,0}|_{\mathcal{X}}\|_p$, we get

$$\Big\| \max_{\substack{1\le u\le n_1\\ 1\le v\le n_2}} \big|\theta^{(m)}_{u,v}\big|_{\mathcal{X}} \Big\|_p \ll \lambda m^4 \||X_{0,0}|_{\mathcal{X}}\|_p + m^4(n_1n_2)^{1/p}\||X_{0,0}|_{\mathcal{X}}\mathbf{1}_{\{|X_{0,0}|_{\mathcal{X}}>\lambda\||X_{0,0}|_{\mathcal{X}}\|_p\}}\|_p \,. \tag{7.23}$$

We deal now with the third and fourth term in the R.H.S of decomposition (7.14). For any positive real A, using stationarity, we infer that

$$\begin{aligned}\Big\| \max_{\substack{0\le k\le n_1\\ 0\le \ell\le n_2}} \Big|\sum_{i=1}^{k} g^{(m)}_{i,\ell+1}\Big|_{\mathcal{X}} \Big\|_p &\le \sum_{u=1}^{2m-1}\sum_{b=0}^{2m-2}\sum_{v=b-m+1}^{m-1} \Big\| \max_{\substack{0\le k\le n_1\\ 0\le \ell\le n_2}} \Big|\sum_{i=1}^{k} P_{m+i,v}(X_{u+i,b})\circ T_{0,\ell+1}\Big|_{\mathcal{X}} \Big\|_p \\ &\le 4(2m-1)^3 A\sqrt{n_1} + (2m-1)(n_2+1)^{1/p}\sum_{u=1-m}^{m-1}\sum_{v=-m+1}^{m-1} A(n_1,u,v)\,,\end{aligned}$$

where

$$A(n_1,u,v) := \Big\| \max_{1\le k\le n_1} \Big|\sum_{i=1}^{k} P_{i,v}(X_{u+i,0})\Big|_{\mathcal{X}} \mathbf{1}_{\{|\sum_{i=1}^{k} P_{i,v}(X_{u+i,0})|_{\mathcal{X}}>4A\sqrt{n_1}\}} \Big\|_p \,.$$

Note that $\big(\big|\sum_{i=1}^{k} P_{i,v}(X_{u+i,0})\big|_{\mathcal{X}}\big)_{k\ge1}$ is a submartingale. Therefore, by [17, Corollary 2.10] and stationarity,

$$A(n_1,u,v) \le 2\Big(\frac{p}{p-1}\Big)^{1/p} \Big\| \Big|\sum_{i=1}^{n_1} P_{i,v}(X_{u+i,0})\Big|_{\mathcal{X}} \mathbf{1}_{\{|\sum_{i=1}^{n_1} P_{i,v}(X_{u+i,0})|_{\mathcal{X}}>2A\sqrt{n_1}\}} \Big\|_p \,.$$

Next noticing that for any real B,

$$P_{i,v}(X_{u+i,0}) = P_{i,v}(X_{u+i,0}\mathbf{1}_{\{|X_{u+i,0}|_{\mathcal{X}}\le B\}}) + P_{i,v}(X_{u+i,0}\mathbf{1}_{\{|X_{u+i,0}|_{\mathcal{X}}>B\}})$$

and, proceeding as to get (7.20) and selecting $A=\lambda^2\||X_{0,0}|_{\mathcal{X}}\|_p$ and $B=A/\lambda$, we infer that, for any $\lambda>0$,

$$\begin{aligned}\Big\| \max_{\substack{1\le k\le n_1\\ 1\le \ell\le n_2}} \Big|\sum_{i=1}^{k} g^{(m)}_{i,\ell+1}\Big|_{\mathcal{X}} \Big\|_p &\ll m^3\lambda^2\sqrt{n_1}\||X_{0,0}|_{\mathcal{X}}\|_p \\ &+ m^3(n_1)^{1/2}(n_2)^{1/p}\||X_{0,0}|_{\mathcal{X}}\mathbf{1}_{\{|X_{0,0}|_{\mathcal{X}}>\lambda\||X_{0,0}|_{\mathcal{X}}\|_p\}}\|_p + \frac{m^3(n_1)^{1/2}(n_2)^{1/p}}{\lambda}\||X_{0,0}|_{\mathcal{X}}\|_p \,.\end{aligned} \tag{7.24}$$

With similar arguments, we obtain

$$\Big\| \max_{\substack{1\le k\le n_1\\ 1\le \ell\le n_2}} \Big| \sum_{j=1}^{\ell} h_{k+1,j}^{(m)}\Big|_{\mathcal{X}} \Big\|_p \ll m^3\lambda^2\sqrt{n_2}\||X_{0,0}|_{\mathcal{X}}\|_p$$
$$+ m^3(n_1)^{1/p}(n_2)^{1/2}\||X_{0,0}|_{\mathcal{X}}\mathbf{1}_{\{|X_{0,0}|_{\mathcal{X}}>\lambda\||X_{0,0}|_{\mathcal{X}}\|_p\}}\|_p + \frac{m^3(n_1)^{1/p}(n_2)^{1/2}}{\lambda}\||X_{0,0}|_{\mathcal{X}}\|_p\,. \quad (7.25)$$

Starting from decomposition (7.14) and considering the upper bounds (7.15), (7.16), (7.22), (7.23), (7.24) and (7.25), the inequality (7.11) follows.

Let us now indicate the main argument to extend (7.12) to dimension $d > 2$ and then get (6.2). By simple induction we infer that the following extension of (7.13) holds:

$$X_{\underline{0}} - D_{\underline{0}} \circ T_{\underline{m}} = \sum_{\emptyset\subsetneq J\subseteq\langle d\rangle} \prod_{s\in J}(I - U_s)g_J^{(m)} + Y_{\underline{0}}^{(m)} - Z_{\underline{0}}^{(m)}\,, \quad (7.26)$$

where

$$Y_{\underline{0}}^{(m)} = \sum_{\underline{k}\in\mathbb{Z}^d\setminus[-m+1,m-1]^d} P_{\underline{k}}(X_{\underline{0}}) \text{ and } Z_{\underline{0}}^{(m)} := \sum_{\underline{u}\in\mathbb{Z}^d\setminus[1,2m-1]^d} P_{\underline{m}}(X_{\underline{u}})\,,$$

and $g_J^{(m)} \in L^p(\mathcal{X})$. Even if it is not an easy task to give a rigorous formulation for g_J^m let us describe them for some given sets J: for $J = J_1 = \{1\}$ we have

$$g_{J_1}^{(m)} = \sum_{u_1=1}^{2m-1}\sum_{a_2=0}^{2m-2}\cdots\sum_{a_d=0}^{2m-2}\sum_{b_2=a_2-m+1}^{m-1}\cdots\sum_{b_d=a_d-m+1}^{m-1} P_{m,b_2,\dots,b_d}(X_{u_1,a_2,\dots,a_d})\,,$$

for $J = J_2 = \{1,2\}$ we have

$$g_{J_2}^{(m)} = \sum_{u_1=1}^{2m-1}\sum_{u_2=1}^{2m-1}\sum_{a_3=0}^{2m-2}\cdots\sum_{a_d=0}^{2m-2}\sum_{b_3=a_3-m+1}^{m-1}\cdots\sum_{b_d=a_d-m+1}^{m-1} P_{m,m,b_3,\dots,b_d}(X_{u_1,u_2,a_3,\dots,a_d})\,,$$

for $J = J_3 = \{1,3\}$ we have

$$g_{J_2}^{(m)} = \sum_{u_1=1}^{2m-1}\sum_{u_3=1}^{2m-1}\sum_{a_2=0}^{2m-2}\sum_{a_4=0}^{2m-2}\cdots\sum_{a_d=0}^{2m-2}\sum_{b_2=a_2-m+1}^{m-1}\sum_{b_4=a_4-m+1}^{m-1}\cdots\sum_{b_d=a_d-m+1}^{m-1} P_{m,b_2,m,b_4,\dots,b_d}(X_{u_1,a_2,u_3,a_4,\dots,a_d})\,,$$

and so on. Inequality (6.2) then follows using the decomposition (7.26) and the arguments used to prove (7.12). □

References

[1] P. Assouad. Espaces p-lisses et q-convexes, inégalités de Burkholder. *Séminaire Maurey-Schwartz*, pages 1–7, 1974-1975.

[2] C. Cuny. Invariance principles under the Maxwell–Woodroofe condition in Banach spaces. *Ann. proba*, 45(3):1578–1611, 2017.

[3] C. Cuny, J. Dedecker, and D. Volný. A functional clt for fields of commuting transformations via martingale approximation. *J. Math. Sci*, 219(5):765–781, 2016.

[4] S. Dede. An empirical central limit theorem in L1 for stationary sequences. *Stochastic Process. Appl.*, 119(10):3494–3515, 2009.

[5] J. Dedecker. Inégalités de covariance. *C. R. Math. Acad. Sci. Paris*, 339(7):503–506, 2004.

[6] J. Dedecker, F Merlevède, and M. Peligrad. Invariance principles for linear processes with application to isotonic regression. *Bernoulli*, 17(1):88–113, 2011.

[7] J. Dedecker, F. Merlevède, and F. Pène. Rates in the strong invariance principle for ergodic automorphisms of the torus. *Stoch. Dyn.*, 14(2):1350021, 30, 2014.

[8] J. Dedecker and F. Merlevède. Moment bounds for dependent sequences in smooth Banach spaces. *Stochastic Process. Appl.*, 125(9):3401–3429, 2015.

[9] J. Dedecker and C. Prieur. New dependence coefficients. examples and applications to statistics. *Probab. Theory. Relat. Fields*, 123(2):203–236, 2005.

[10] M. El Machkouri, D. Volný, and W. Wu. A central limit theorem for stationary random fields. *Stochastic Process. Appl.*, 123(1):1–14, 2013.

[11] M I. Gordin. The central limit theorem for stationary processes. In *Dokl. Akad. Nauk*, volume 188, pages 739–741. SSSR, 1969.

[12] P. Hall and CC. Heyde. Martingale limit theory and its application. 1980.

[13] E. J. Hannan. Central limit theorems for time series regression. *Z. Wahrscheinlichkeitstheor und verw. Gebiete*, 26(2):157–170, 1973.

[14] D. Khoshnevisan. *Multiparameter processes: An introduction to random fields.* Springer, 2002.

[15] M. Ledoux and M. Talagrand. *Probability in Banach Spaces: isoperimetry and processes.* Springer Science & Business Media, 2013.

[16] M. Maxwell and M. Woodroofe. Central limit theorems for additive functionals of Markov chains. *Ann. proba*, pages 713–724, 2000.

[17] F. Merlevède, M. Peligrad, and S. Utev. *Functional Gaussian Approximation for Dependent Structures.* Oxford Studies in Probability, 2019.

[18] I. Pinelis. Optimum bounds for the distributions of martingales in Banach spaces. *Ann. Probab.*, 22(4):1679–1706, 1994.

[19] D. Volný. Martingale-coboundary decomposition for stationary random fields. *Stoch. Dyn.*, 2018.

[20] D. Volný and Y. Wang. An invariance principle for stationary random fields under Hannan's condition. *Stochastic Process. Appl.*, 124(12):4012–4029, 2014.

[21] W.B. Wu. Strong invariance principles for dependent random variables. *Ann. Probab.*, 35(6):2294–2320, 2007.