



Weak invariance principles for sums of dependent random functions[☆]

István Berkes^a, Lajos Horváth^{b,*}, Gregory Rice^b

^a *Institute of Statistics, Graz University of Technology, Kopernikus Gasse 24/3, A-8010 Graz, Austria*

^b *Department of Mathematics, University of Utah, Salt Lake City, UT 84112-0090, USA*

Received 4 February 2012; received in revised form 28 September 2012; accepted 1 October 2012

Available online 6 October 2012

Abstract

Motivated by problems in functional data analysis, in this paper we prove the weak convergence of normalized partial sums of dependent random functions exhibiting a Bernoulli shift structure.

© 2012 Elsevier B.V. All rights reserved.

MSC: 60F17; 60F25; 62G10

Keywords: Variables in Hilbert spaces; m -approximability; Weak convergence

1. Introduction

Functional data analysis in many cases requires central limit theorems and invariance principles for partial sums of random functions. The case of independent summands is much studied and well understood but the theory for the dependent case is less complete. In this paper we study the important class of Bernoulli shift processes which are often used to model econometric and financial data. Let $\mathbf{X} = \{X_i(t)\}_{i=-\infty}^{\infty}$ be a sequence of random functions, square integrable on $[0, 1]$, and let $\|\cdot\|$ denote the $L^2[0, 1]$ norm. To lighten the notation we use f for $f(t)$ when it does not cause confusion. Throughout this paper we assume that

\mathbf{X} forms a sequence of Bernoulli shifts, i.e. $X_j(t) = g(\epsilon_j(t), \epsilon_{j-1}(t), \dots)$ for some nonrandom measurable function $g : S^\infty \mapsto L^2$ and iid random functions $\epsilon_j(t)$,

[☆] Research supported by NSF grant DMS 0905400.

* Corresponding author.

E-mail address: horvath@math.utah.edu (L. Horváth).

$$-\infty < j < \infty, \text{ with values in a measurable space } S, \tag{1.1}$$

$$\epsilon_j(t) = \epsilon_j(t, \omega) \text{ is jointly measurable in } (t, \omega) \text{ } (-\infty < j < \infty), \tag{1.2}$$

$$EX_0(t) = 0 \text{ for all } t, \text{ and } E\|X_0\|^{2+\delta} < \infty \text{ for some } 0 < \delta < 1, \tag{1.3}$$

and

the sequence $\{X_n\}_{n=-\infty}^\infty$ can be approximated by ℓ -dependent sequences

$\{X_{n,\ell}\}_{n=-\infty}^\infty$ in the sense that

$$\sum_{\ell=1}^\infty (E\|X_n - X_{n,\ell}\|^{2+\delta})^{1/\kappa} < \infty \text{ for some } \kappa > 2 + \delta,$$

where $X_{n,\ell}$ is defined by $X_{n,\ell} = g(\epsilon_n, \epsilon_{n-1}, \dots, \epsilon_{n-\ell+1}, \epsilon_{n,\ell}^*)$,

$\epsilon_{n,\ell}^* = (\epsilon_{n,\ell,n-\ell}^*, \epsilon_{n,\ell,n-\ell-1}^*, \dots)$, where the $\epsilon_{n,\ell,k}^*$'s are independent copies of ϵ_0 ,

independent of $\{\epsilon_i, -\infty < i < \infty\}$. (1.4)

We note that assumption (1.1) implies that X_n is a stationary and ergodic sequence. Hörmann and Kokoszka [19] call the processes satisfying (1.1)–(1.4) L^2 m -decomposable processes. The idea of approximating a stationary sequence with random variables which exhibit finite dependence first appeared in [21] and is used frequently in the literature (cf. [3]). Aue et al. [1] provide several examples when assumption (1.1)–(1.4) hold which include autoregressive, moving average and linear processes in Hilbert spaces. Also, the non-linear functional ARCH(1) model (cf. [18]) and bilinear models (cf. [19]) satisfy (1.4).

We show in Section 2 (cf. Lemma 2.2) that the series in

$$C(t, s) = E[X_0(t)X_0(s)] + \sum_{\ell=1}^\infty E[X_0(t)X_\ell(s)] + \sum_{\ell=1}^\infty E[X_0(s)X_\ell(t)] \tag{1.5}$$

are convergent in L^2 . The function $C(t, s)$ is positive definite, and therefore there exist $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$ and orthonormal functions $\phi_i(t), 0 \leq t \leq 1$ satisfying

$$\lambda_i \phi_i(t) = \int C(t, s) \phi_i(s) ds, \quad 1 \leq i < \infty, \tag{1.6}$$

where \int means \int_0^1 . We define

$$\Gamma(x, t) = \sum_{i=1}^\infty \lambda_i^{1/2} W_i(x) \phi_i(t),$$

where W_i are independent and identically distributed Wiener processes (standard Brownian motions). Clearly, $\Gamma(x, t)$ is Gaussian. We show in Lemma 2.2 that $\sum_{\ell=1}^\infty \lambda_\ell < \infty$, and therefore

$$\sup_{0 \leq x \leq 1} \int \Gamma^2(x, t) dt < \infty \text{ a.s.}$$

Theorem 1.1. *If assumption (1.1)–(1.4) hold, then for every N we can define a Gaussian process $\Gamma_N(x, t)$ such that*

$$\{\Gamma_N(x, t), 0 \leq x, t \leq 1\} \stackrel{D}{=} \{\Gamma(x, t), 0 \leq x, t \leq 1\}$$

and

$$\sup_{0 \leq x \leq 1} \int (S_N(x, t) - \Gamma_N(x, t))^2 dt = o_P(1),$$

where

$$S_N(x, t) = \frac{1}{N^{1/2}} \sum_{i=1}^{\lfloor Nx \rfloor} X_i(t).$$

The proof of [Theorem 1.1](#) is given in [Section 2](#). The proof is based on a maximal inequality which is given in [Section 3](#) and is of interest in its own right.

There is a wide literature on the central limit theorem for sums of random processes in abstract spaces. For limit theorems for sums of independent Banach space valued random variables we refer to [Ledoux and Talagrand \(1991\)](#). For the central limit theory in the context of functional data analysis we refer to the books of [\[4,20\]](#). In the real valued case, the martingale approach to weak dependence was developed by [Gordin \[16\]](#), [Philipp and Stout \[25\]](#) and [Eberlein \[14\]](#), and by using such techniques [\[23,9\]](#) obtained central limit theorems for a large class of dependent variables in Hilbert spaces. For some early influential results on invariance for sums of mixing variables in Banach spaces we refer to [\[22,11,10\]](#). These papers provide very sharp results, but verifying mixing conditions is generally not easy and without additional continuity conditions, even autoregressive (1) processes may fail to be strong mixing (cf. [\[5\]](#)). The weak dependence concept of [\[13\]](#) (cf. also [\[8\]](#)) solves this difficulty, but so far this concept has not been extended to variables in Hilbert spaces. [Wu \[27,28\]](#) proved several limit theorems for one-dimensional stationary processes having a Bernoulli shift representation. Compared to classical mixing conditions, Wu’s physical dependence conditions are easier to verify in concrete cases. [Condition \(1.3\)](#) cannot be directly compared to the approximating martingale conditions of [\[27,28\]](#). For extensions to the Hilbert space case we refer to [\[19\]](#).

2. Proof of [Theorem 1.1](#)

The proof is based on three steps. We recall the definition of $X_{i,m}$ from [\(1.4\)](#). For every fixed m , the sequence $\{X_{i,m}\}$ is m -dependent. According to our first lemma, the sums of the X_i ’s can be approximated with the sums of m -dependent variables. The second step is the approximation of the infinite dimensional $X_{i,m}$ ’s with finite dimensional variables ([Lemma 2.4](#)). Then the result in [Theorem 1.1](#) is established for finite dimensional m -dependent random functions ([Lemma 2.6](#)).

Lemma 2.1. *If [\(1.1\)–\(1.4\)](#) hold, then for all $x > 0$ we have*

$$\lim_{m \rightarrow \infty} \limsup_{N \rightarrow \infty} P \left\{ \max_{1 \leq k \leq N} \frac{1}{\sqrt{N}} \left\| \sum_{i=1}^k (X_i - X_{i,m}) \right\| > x \right\} = 0. \tag{2.1}$$

Proof. The proof of this lemma requires the maximal inequality of [Theorem 3.2](#). [Section 3](#) is devoted to the proof of this result. Using [Theorem 3.2](#), [\(2.1\)](#) is an immediate consequence of Markov’s inequality. \square

Define

$$C_m(t, s) = E[X_{0,m}(t)X_{0,m}(s)] + \sum_{i=1}^m E[X_{0,m}(t)X_{i,m}(s)] + \sum_{i=1}^m E[X_{0,m}(s)X_{i,m}(t)]. \tag{2.2}$$

We show in the following lemma that for every m the function C_m is square-integrable. Hence there are $\lambda_{1,m} \geq \lambda_{2,m} \geq \dots \geq 0$ and corresponding orthonormal functions $\phi_{i,m}, i = 1, 2, \dots$ satisfying

$$\lambda_{i,m}\phi_{i,m}(t) = \int C_m(t, s)\phi_{i,m}(s)ds, \quad i = 1, 2, \dots$$

Lemma 2.2. *If (1.1)–(1.4) hold, then we have*

$$\iint C^2(t, s)dt ds < \infty, \tag{2.3}$$

$$\iint C_m^2(t, s)dt ds < \infty \quad \text{for all } m \geq 1, \tag{2.4}$$

$$\lim_{m \rightarrow \infty} \iint (C(t, s) - C_m(t, s))^2 dt ds = 0, \tag{2.5}$$

$$\int C(t, t)dt = \sum_{k=1}^{\infty} \lambda_k < \infty, \tag{2.6}$$

$$\int C_m(t, t)dt = \sum_{k=1}^{\infty} \lambda_{k,m} < \infty \tag{2.7}$$

and

$$\lim_{m \rightarrow \infty} \int C_m(t, t)dt = \int C(t, t)dt. \tag{2.8}$$

Proof. Using the Cauchy–Schwarz inequality for expected values we get

$$\begin{aligned} \iint (E[X_0(t)X_0(s)])^2 dt ds &\leq \iint ((EX_0^2(t))^{1/2}(EX_0^2(s))^{1/2})^2 dt ds \\ &= (E\|X_0\|^2)^2 < \infty. \end{aligned}$$

Recalling that X_0 and $X_{i,i}$ are independent and both have 0 mean, we conclude first using the triangle inequality and then the Cauchy–Schwarz inequality for expected values that

$$\begin{aligned} &\left\{ \iint \left(\sum_{i=1}^{\infty} E[X_0(t)X_i(s)] \right)^2 dt ds \right\}^{1/2} \\ &= \left\{ \iint \left(\sum_{i=1}^{\infty} E[X_0(t)(X_i(s) - X_{i,i}(s))] \right)^2 dt ds \right\}^{1/2} \\ &\leq \left(\iint \left(\sum_{i=1}^{\infty} E|X_0(t)(X_i(s) - X_{i,i}(s))| \right)^2 dt ds \right)^{1/2} \\ &\leq \sum_{i=1}^{\infty} \left(\int \int \{E|X_0(t)(X_i(s) - X_{i,i}(s))|\}^2 dt ds \right)^{1/2} \end{aligned}$$

$$\begin{aligned}
 &\leq \sum_{i=1}^{\infty} \iint \left\{ (EX_0^2(t))^{1/2} (E(X_i(s) - X_{i,i}(s))^2)^{1/2} \right\}^2 dt ds \\
 &= \int EX_0^2(t) dt \sum_{i=1}^{\infty} \int E(X_i(s) - X_{i,i}(s))^2 ds \\
 &= E\|X_0\|^2 \sum_{i=1}^{\infty} E\|X_0 - X_{0,i}\|^2 \\
 &< \infty
 \end{aligned} \tag{2.9}$$

on account of (1.4). This completes the proof of (2.3).

Since $EX_{0,m}(t)X_{0,m}(s) = EX_0(t)X_0(s)$, in order to establish (2.4), it is enough to show that

$$\iint \left\{ \sum_{i=1}^m E[X_{0,m}(t)X_{i,m}(s)] \right\}^2 dt ds < \infty.$$

It follows from the definition of $X_{i,m}$ that the vectors $(X_{0,m}, X_{i,m})$ and $(X_0, X_{i,m})$ have the same distribution for all $1 \leq i \leq m$. Also, $(X_{i,m}, X_{i,i})$ has the same distribution as $(X_0, X_{0,i})$, $1 \leq i \leq m$. Hence following the arguments in (2.9) we get

$$\begin{aligned}
 &\left\{ \iint \left\{ \sum_{i=1}^m |EX_{0,m}(t)X_{i,m}(s)| \right\}^2 dt ds \right\}^{1/2} \\
 &= \left\{ \iint \left\{ \sum_{i=1}^m |EX_0(t)X_{i,m}(s)| \right\}^2 dt ds \right\}^{1/2} \\
 &\leq E\|X_0\|^2 \sum_{i=1}^m \int E(X_{i,m}(s) - X_{i,i}(s))^2 ds \\
 &\leq E\|X_0\|^2 \sum_{i=1}^{\infty} E\|X_0 - X_{0,i}\|^2 \\
 &< \infty.
 \end{aligned}$$

The proof of (2.4) is now complete. The arguments used above also prove (2.5).

Repeating the previous arguments we have

$$\begin{aligned}
 \int C(t, t) dt &\leq \int EX_0^2(t) dt + 2 \sum_{i=1}^{\infty} \int |E[X_0(t)X_i(t)]| dt \\
 &= \int EX_0^2(t) dt + 2 \sum_{i=1}^{\infty} \int |E[X_0(t)(X_i(t) - X_{i,i}(t))]| dt \\
 &= \int EX_0^2(t) dt + 2 \sum_{i=1}^{\infty} \int (EX_0^2(t))^{1/2} (E[X_i(t) - X_{i,i}(t)]^2)^{1/2} dt \\
 &\leq E\|X_0\|^2 + 2 \sum_{i=1}^{\infty} \left(\int EX_0^2(t) dt \right)^{1/2} \left(\int E[X_i(t) - X_{i,i}(t)]^2 dt \right)^{1/2}
 \end{aligned}$$

$$\begin{aligned}
 &= E\|X_0\|^2 + 2(E\|X_0\|^2)^{1/2} \sum_{i=1}^{\infty} (E\|X_0 - X_{0,i}\|^2)^{1/2} \\
 &< \infty.
 \end{aligned}$$

Observing that

$$\int C(t, t)dt = \sum_{i=1}^{\infty} \lambda_i \int \phi_i^2(t)dt = \sum_{i=1}^{\infty} \lambda_i,$$

the proof of (2.6) is complete. The same arguments can be used to establish (2.7). The relation in (2.8) can be established along the lines of the proof of (2.5). \square

By the Karhunen–Loève expansion, we have that

$$\bar{X}_{i,m}(t) = \sum_{\ell=1}^{\infty} \langle X_{i,m}, \phi_{\ell,m} \rangle \phi_{\ell,m}(t). \tag{2.10}$$

Define

$$X_{i,m,K}(t) = \sum_{\ell=1}^K \langle X_{i,m}, \phi_{\ell,m} \rangle \phi_{\ell,m}(t) \tag{2.11}$$

to be the partial sums of the series in (2.10), and

$$\bar{X}_{i,m,K}(t) = X_{i,m}(t) - X_{i,m,K}(t) = \sum_{\ell=K+1}^{\infty} \langle X_{i,m}, \phi_{\ell,m} \rangle \phi_{\ell,m}(t). \tag{2.12}$$

Lemma 2.3. *If $\{Z_i\}_{i=1}^N$ are independent L^2 valued random variables such that*

$$EZ_1(t) = 0 \quad \text{and} \quad E\|Z_1\|^2 < \infty, \tag{2.13}$$

then for all $x > 0$ we have that

$$P \left\{ \max_{1 \leq k \leq N} \left\| \sum_{i=1}^k Z_i \right\|^2 > x \right\} \leq \frac{1}{x} E \left\| \sum_{i=1}^N Z_i \right\|^2. \tag{2.14}$$

Proof. Let \mathcal{F}_k be the sigma algebra generated by the random variables $\{Z_j\}_{j=1}^k$. By assumption (2.13) and the independence of the Z_i 's we have that

$$E \left(\left\| \sum_{i=1}^{k+1} Z_i \right\|^2 \middle| \mathcal{F}_k \right) = \left\| \sum_{i=1}^k Z_i \right\|^2 + E\|Z_{k+1}\|^2 \geq \left\| \sum_{i=1}^k Z_i \right\|^2.$$

Therefore $\left\{ \left\| \sum_{i=1}^k Z_i \right\|^2 \right\}_{k=1}^{\infty}$ is a non-negative submartingale with respect to the filtration $\{\mathcal{F}_k\}_{k=1}^{\infty}$. If we define

$$A = \left\{ \omega : \max_{1 \leq k \leq N} \left\| \sum_{i=1}^k Z_i \right\|^2 > x \right\},$$

then it follows from Doob’s maximal inequality [6, p. 247] that

$$\begin{aligned}
 x P \left\{ \max_{1 \leq k \leq N} \left\| \sum_{i=1}^k Z_i \right\|^2 > x \right\} &\leq E \left(\left\| \sum_{i=1}^N Z_i \right\|^2 I_A \right) \\
 &\leq E \left\| \sum_{i=1}^N Z_i \right\|^2,
 \end{aligned}$$

which completes the proof. \square

Lemma 2.4. *If (1.1)–(1.4) hold, then for all $x > 0$,*

$$\lim_{K \rightarrow \infty} \limsup_{N \rightarrow \infty} P \left\{ \max_{1 \leq k \leq N} \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^k \bar{X}_{i,m,K} \right\| > x \right\} = 0. \tag{2.15}$$

Proof. Define $Q_k(j) = \{i : 1 \leq i \leq k, i = j \pmod{m}\}$ for $j = 0, 1, \dots, m - 1$, and all positive integers k . It is then clear that

$$\sum_{i=1}^k \bar{X}_{i,m,K} = \sum_{j=0}^{m-1} \sum_{i \in Q_k(j)} \bar{X}_{i,m,K}.$$

We thus obtain by the triangle inequality that

$$P \left\{ \max_{1 \leq k \leq N} \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^k \bar{X}_{i,m,K} \right\| > x \right\} \leq P \left\{ \sum_{j=0}^{m-1} \max_{1 \leq k \leq N} \left\| \frac{1}{\sqrt{N}} \sum_{i \in Q_k(j)} \bar{X}_{i,m,K} \right\| > x \right\}.$$

It is therefore sufficient to show that for each fixed j ,

$$\lim_{K \rightarrow \infty} \limsup_{N \rightarrow \infty} P \left\{ \max_{1 \leq k \leq N} \left\| \frac{1}{\sqrt{N}} \sum_{i \in Q_k(j)} \bar{X}_{i,m,K} \right\| > x \right\} = 0.$$

By the definition of $Q_k(j)$, $\{\bar{X}_{i,m,K}\}_{i \in Q_k(j)}$ is an iid sequence of random variables. So, by applications of Lemma 2.3 and assumption (1.3), we have that

$$\begin{aligned}
 P \left\{ \max_{1 \leq k \leq N} \left\| \frac{1}{\sqrt{N}} \sum_{i \in Q_k(j)} \bar{X}_{i,m,K} \right\|^2 > x \right\} &\leq \frac{1}{x} E \left\| \frac{1}{\sqrt{N}} \sum_{i \in Q_N(j)} \bar{X}_{i,m,K} \right\|^2 \\
 &\leq \frac{1}{x} E \|\bar{X}_{0,m,K}^2\| \\
 &= \frac{1}{x} \sum_{\ell=K+1}^{\infty} \lambda_{\ell,m}.
 \end{aligned} \tag{2.16}$$

Since the right hand side of (2.16) tends to zero as K tends to infinity independently of N , (2.15) follows. \square

‘Clearly, with $k = \lfloor Nx \rfloor$ we have

$$\frac{1}{\sqrt{N}} \sum_{i=1}^k X_{i,m,K}(t) = \sum_{j=1}^K \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{\lfloor Nx \rfloor} \langle X_{i,m}, \phi_{j,m} \rangle \right) \phi_{j,m}(t). \tag{2.17}$$

Lemma 2.5. *If (1.1)–(1.4) hold, then the K dimensional random process*

$$\left(\frac{1}{\sqrt{N}} \sum_{i=1}^{\lfloor Nx \rfloor} \langle X_{i,m}, \phi_{1,m} \rangle, \dots, \frac{1}{\sqrt{N}} \sum_{i=1}^{\lfloor Nx \rfloor} \langle X_{i,m}, \phi_{K,m} \rangle \right)$$

converges, as $N \rightarrow \infty$, in $\mathcal{D}^K[0, 1]$ to

$$\left(\lambda_{1,m}^{1/2} W_1(x), \dots, \lambda_{K,m}^{1/2} W_K(x) \right), \tag{2.18}$$

where $\{W_i\}_{i=1}^K$ are independent, identically distributed Wiener processes.

Proof. A similar procedure as in Lemma 2.4 shows that for each j , $\frac{1}{\sqrt{N}} \sum_{i=1}^{\lfloor Nx \rfloor} \langle X_{i,m}, \phi_{j,m} \rangle$ can be written as a sum of sums of independent and identically distributed random variables, and thus, by Billingsley [3], it is tight. This implies that the K dimensional process

$$\left(\frac{1}{\sqrt{N}} \sum_{i=1}^{\lfloor Nx \rfloor} \langle X_{i,m}, \phi_{1,m} \rangle, \dots, \frac{1}{\sqrt{N}} \sum_{i=1}^{\lfloor Nx \rfloor} \langle X_{i,m}, \phi_{K,m} \rangle \right)$$

is tight, since it is tight in each coordinate. Furthermore, the Cramér–Wold device and the central limit theorem for m -dependent random variables (cf. [7, p. 119]) shows that the finite dimensional distributions of the vector process converge to the finite dimensional distributions of the process in (2.18). The lemma follows. \square

In light of the Skorokhod–Dudley–Wichura theorem (cf. [26, p. 47]), we may reformulate Lemma 2.5 as follows.

Corollary 2.1. *If (1.1)–(1.4) hold, then for each positive integer N , there exists K independent, identically distributed Wiener processes $\{W_{i,N}\}_{i=1}^K$ such that for each j ,*

$$\sup_{0 \leq x \leq 1} \left| \frac{1}{\sqrt{N}} \sum_{i=1}^{\lfloor Nx \rfloor} \langle X_{i,m}, \phi_{j,m} \rangle - \lambda_{j,m}^{1/2} W_{j,N}(x) \right| \xrightarrow{P} 0,$$

as $N \rightarrow \infty$.

Lemma 2.6. *If (1.1)–(1.4) hold, then for $\{W_{i,N}\}_{i=1}^K$ defined in Corollary 2.1, we have that*

$$\sup_{0 \leq x \leq 1} \int \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{\lfloor Nx \rfloor} X_{i,m,K}(t) - \sum_{\ell=1}^K \lambda_{\ell,m}^{1/2} W_{\ell,N}(x) \phi_{\ell,m}(t) \right)^2 dt \xrightarrow{P} 0, \tag{2.19}$$

as $N \rightarrow \infty$.

Proof. By using (2.17), we get that

$$\begin{aligned} & \frac{1}{\sqrt{N}} \sum_{i=1}^{\lfloor Nx \rfloor} X_{i,m,K}(t) - \sum_{\ell=1}^K \lambda_{\ell,m}^{1/2} W_{\ell,N}(x) \phi_{\ell,m}(t) \\ &= \sum_{\ell=1}^K \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{\lfloor Nx \rfloor} \langle X_{i,m}, \phi_{\ell,m} \rangle - \lambda_{\ell,m}^{1/2} W_{\ell,N}(x) \right) \phi_{\ell,m}(t). \end{aligned}$$

The substitution of this into the expression in (2.19) along with a simple calculation shows that

$$\begin{aligned} & \sup_{0 \leq x \leq 1} \int \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{\lfloor Nx \rfloor} X_{i,m,K}(t) - \sum_{\ell=1}^K \lambda_{\ell,m}^{1/2} W_{\ell,N}(x) \phi_{\ell,m}(t) \right)^2 dt \\ &= \sup_{0 \leq x \leq 1} \sum_{\ell=1}^K \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{\lfloor Nx \rfloor} \langle X_{i,m}, \phi_{\ell,m} \rangle - \lambda_{\ell,m}^{1/2} W_{\ell,N}(x) \right)^2 \\ &\leq \sum_{\ell=1}^K \sup_{0 \leq x \leq 1} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{\lfloor Nx \rfloor} \langle X_{i,m}, \phi_{\ell,m} \rangle - \lambda_{\ell,m}^{1/2} W_{\ell,N}(x) \right)^2 \xrightarrow{P} 0, \end{aligned}$$

as $N \rightarrow \infty$, by Corollary 2.1. \square

Lemma 2.7. *If (1.1)–(1.4) hold,*

$$\sup_{0 \leq x \leq 1} \int \left(\sum_{\ell=K+1}^{\infty} \lambda_{\ell,m}^{1/2} W_{\ell}(x) \phi_{\ell,m}(t) \right)^2 dt \xrightarrow{P} 0, \tag{2.20}$$

as $K \rightarrow \infty$, where W_1, W_2, \dots are independent and identically distributed Wiener processes.

Proof. Since the functions $\{\phi_{\ell,m}\}_{\ell=1}^{\infty}$ are orthonormal, we have that

$$\begin{aligned} E \sup_{0 \leq x \leq 1} \int \left(\sum_{\ell=K+1}^{\infty} \lambda_{\ell,m}^{1/2} W_{\ell}(x) \phi_{\ell,m}(t) \right)^2 dt &= E \sup_{0 \leq x \leq 1} \sum_{\ell=K+1}^{\infty} \lambda_{\ell,m} W_{\ell}^2(x) \\ &\leq \sum_{\ell=K+1}^{\infty} \lambda_{\ell,m} E \sup_{0 \leq x \leq 1} W_{\ell}^2(x) \longrightarrow 0, \end{aligned}$$

as $K \rightarrow \infty$. Therefore (2.20) follows from Markov’s inequality. \square

Lemma 2.8. *If (1.1)–(1.4) hold, then for each N we can define independent identically distributed Wiener processes $\{W_{i,N}\}_{i=1}^K$ such that*

$$\sup_{0 \leq x \leq 1} \int \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{\lfloor Nx \rfloor} X_{i,m}(t) - \sum_{\ell=1}^{\infty} \lambda_{\ell,m}^{1/2} W_{\ell,N}(x) \phi_{\ell,m}(t) \right)^2 dt \xrightarrow{P} 0,$$

as $N \rightarrow \infty$.

Proof. It follows from Lemmas 2.4–2.7. \square

Since the distribution of $W_{\ell,N}$, $1 \leq \ell < \infty$ does not depend on N , it is enough to consider the asymptotics for $\sum_{\ell=1}^{\infty} \lambda_{\ell,m}^{1/2} W_{\ell}(x) \phi_{\ell,m}(t)$, where W_{ℓ} are independent Wiener processes.

Lemma 2.9. *If (1.1)–(1.4) hold, then for each m we can define independent and identically distributed Wiener processes $\bar{W}_{\ell,m}(x)$, $1 \leq \ell < \infty$ such that*

$$\sup_{0 \leq x \leq 1} \int \left(\sum_{\ell=1}^{\infty} \lambda_{\ell,m}^{1/2} W_{\ell}(x) \phi_{\ell,m}(t) - \sum_{\ell=1}^{\infty} \lambda_{\ell}^{1/2} \bar{W}_{\ell,m}(x) \phi_{\ell}(t) \right)^2 dt \xrightarrow{P} 0, \tag{2.21}$$

as $m \rightarrow \infty$.

Proof. Let

$$\Delta_m(x, t) = \sum_{\ell=1}^{\infty} \lambda_{\ell,m}^{1/2} W_{\ell}(x) \phi_{\ell,m}(t).$$

Let M be a positive integer and define $x_i = i/M, 0 \leq i \leq M$. It is easy to see that

$$\begin{aligned} E \max_{0 \leq i < M} \sup_{0 \leq h \leq 1/M} \int (\Delta_m(x_i + h, t) - \Delta_m(x_i, t))^2 dt \\ \leq \sum_{\ell=1}^{\infty} \lambda_{\ell,m} E \left\{ \max_{0 \leq i < M} \sup_{0 \leq h \leq 1/M} (W_{\ell}(x_i + h) - W_{\ell}(x_i))^2 \right\} \\ = E \left\{ \max_{0 \leq i < M} \sup_{0 \leq h \leq 1/M} (W_1(x_i + h) - W_1(x_i))^2 \right\} \sum_{\ell=1}^{\infty} \lambda_{\ell,m}. \end{aligned}$$

Using Lemma 2.2 we get that

$$\sum_{\ell=1}^{\infty} \lambda_{\ell,m} = \int E \Delta_m^2(1, t) dt = \int C_m(t, t) dt \rightarrow \int C(t, t) dt = \sum_{\ell=1}^{\infty} \lambda_{\ell}.$$

So by the modulus of continuity of the Wiener process (cf. [15]) we get that

$$\lim_{M \rightarrow \infty} \limsup_{m \rightarrow \infty} E \max_{0 \leq i < M} \sup_{0 \leq h \leq 1/M} \int (\Delta_m(x_i + h, t) - \Delta_m(x_i, t))^2 dt = 0. \tag{2.22}$$

By the Karhunen–Loève expansion we can also write Δ_m as

$$\Delta_m(x, t) = \sum_{\ell=1}^{\infty} \langle \Delta_m(x, \cdot), \phi_{\ell} \rangle \phi_{\ell}(t)$$

and

$$E \int \Delta_m^2(x, t) dt = \sum_{\ell=1}^{\infty} E(\langle \Delta_m(x, \cdot), \phi_{\ell} \rangle)^2.$$

So by Lemma 2.2 we have

$$\sum_{\ell=1}^{\infty} E(\langle \Delta_m(x, \cdot), \phi_{\ell} \rangle)^2 \rightarrow x \sum_{\ell=1}^{\infty} \lambda_{\ell}.$$

Also, for any positive integer ℓ ,

$$E(\langle \Delta_m(x, \cdot), \phi_{\ell} \rangle)^2 = \iint C_m(t, s) \phi_{\ell}(t) \phi_{\ell}(s) dt ds \rightarrow \iint C(t, s) \phi_{\ell}(t) \phi_{\ell}(s) dt ds = \lambda_{\ell},$$

as $m \rightarrow \infty$. Hence for every $z > 0$ we have

$$\limsup_{K \rightarrow \infty} \limsup_{m \rightarrow \infty} P \left\{ \int \left(\sum_{\ell=K+1}^{\infty} \langle \Delta_m(x, \cdot), \phi_{\ell} \rangle \phi_{\ell}(t) \right)^2 dt > z \right\} = 0. \tag{2.23}$$

The joint distribution of $\langle \Delta(x_i, \cdot), \phi_{\ell} \rangle, 1 \leq i \leq M, 1 \leq \ell \leq K$ is multivariate normal with zero mean. Hence they converge jointly to a multivariate normal distribution. To show their joint

convergence in distribution, we need to show the convergence of the covariance matrix. Using again [Lemma 2.2](#) we get that

$$\begin{aligned} E\langle \Delta(x_i, \cdot), \phi_\ell \rangle \langle \Delta(x_j, \cdot), \phi_k \rangle &= \min(x_i, x_j) \iint C_m(t, s) \phi_\ell(t) \phi_k(s) dt ds \\ &\rightarrow \min(x_i, x_j) \iint C(t, s) \phi_\ell(t) \phi_k(s) dt ds \\ &= \min(x_i, x_j) \lambda_\ell I\{k = \ell\}. \end{aligned}$$

Due to this covariance structure and the Skorokhod–Dudley–Wichura theorem (cf. [26, p. 47]) we can find independent Wiener processes $\bar{W}_{\ell,m}(x)$, $1 \leq \ell < \infty$ such that

$$\max_{1 \leq i \leq M} \max_{1 \leq \ell \leq K} |\langle \Delta(x_i, \cdot), \phi_\ell \rangle - \lambda_\ell^{1/2} \bar{W}_{\ell,m}(x_i)| = o_P(1), \quad \text{as } m \rightarrow \infty.$$

Clearly, for all $0 \leq x \leq 1$

$$E \int \left(\sum_{\ell=K+1}^{\infty} \lambda_\ell^{1/2} \bar{W}_{\ell,m}(x) \phi_\ell(t) \right)^2 dt = x \sum_{\ell=K+1}^{\infty} \lambda_\ell \rightarrow 0, \quad \text{as } m \rightarrow \infty,$$

and therefore similarly to (2.23)

$$\limsup_{K \rightarrow \infty} \limsup_{m \rightarrow \infty} P \left\{ \int \left(\sum_{\ell=K+1}^{\infty} \lambda_\ell^{1/2} \bar{W}_{\ell,m}(x) \phi_\ell(t) \right)^2 dt > z \right\} = 0$$

for all $z > 0$. Similarly to (2.22) one can show that

$$\begin{aligned} E \max_{0 \leq i < M} \sup_{0 \leq h \leq 1/M} \int \left(\sum_{\ell=1}^{\infty} (\bar{W}_{\ell,m}(x_i + h) - \bar{W}_{\ell,m}(x_i)) \phi_\ell(t) \right)^2 dt \\ \leq E \left\{ \max_{0 \leq i < M} \sup_{0 \leq h \leq 1/M} (W(x_i + h) - W(x_i))^2 \right\} \sum_{\ell=1}^{\infty} \lambda_\ell \rightarrow 0, \quad \text{as } M \rightarrow \infty, \end{aligned}$$

where W is a Wiener process. This also completes the proof of [Lemma 2.9](#). \square

Proof of Theorem 1.1. First we approximate $S_N(x, t)$ with m -dependent processes ([Lemma 2.1](#)). The second step of the proof is the approximation of the sums of m -dependent processes with a Gaussian process with covariance function $\min(x, x')C_m(t, s)$, where C_m is defined in (2.2) ([Lemma 2.8](#)). The last step of the proof is the convergence of Gaussian processes with covariance functions $\min(x, x')C_m(t, s)$ to a Gaussian process with covariance function $\min(x, x')C(t, s)$ ([Lemma 2.9](#)). \square

3. Some moment and maximal inequalities

In this section we give the proof of the maximal inequality used in [Lemma 2.1](#) which is a crucial ingredient of the proof of [Theorem 1.1](#). Actually, we will prove below some moment and maximal inequalities for partial sums of function valued Bernoulli shift sequences which have their own interest and can be used in various related problems.

Our first lemma is a Hilbert space version of Doob’s [12, p. 226] inequality.

Lemma 3.1. *If Z_1 and Z_2 are independent mean zero Hilbert space valued random variables, and if $0 < \delta \leq 1$, then*

$$E\|Z_1 + Z_2\|^{2+\delta} \leq E\|Z_1\|^{2+\delta} + E\|Z_2\|^{2+\delta} + E\|Z_1\|^2(E\|Z_2\|^2)^{\delta/2} + E\|Z_2\|^2(E\|Z_1\|^2)^{\delta/2}.$$

Proof. Since $0 < \delta \leq 1$, for any $A, B \geq 0$ we have that $(A + B)^\delta \leq A^\delta + B^\delta$ (cf. [17, p. 32]). An application of this inequality along with Minkowski’s inequality gives that

$$\|Z_1 + Z_2\|^\delta \leq (\|Z_1\| + \|Z_2\|)^\delta \leq \|Z_1\|^\delta + \|Z_2\|^\delta.$$

We also have by Hölder’s inequality that

$$E\|Z_1\|^\delta \leq (E\|Z_1\|^2)^{\delta/2}.$$

This yields that

$$\begin{aligned} E\|Z_1 + Z_2\|^{2+\delta} &= E\|Z_1 + Z_2\|^2\|Z_1 + Z_2\|^\delta \\ &\leq E\|Z_1 + Z_2\|^2(\|Z_1\|^\delta + \|Z_2\|^\delta) \\ &= E[\|Z_1\|^2 + \|Z_2\|^2 + 2\langle Z_1, Z_2 \rangle](\|Z_1\|^\delta + \|Z_2\|^\delta) \\ &= E\|Z_1\|^{2+\delta} + E\|Z_2\|^{2+\delta} + E\|Z_1\|^2E\|Z_2\|^\delta + E\|Z_2\|^2E\|Z_1\|^\delta \\ &\leq E\|Z_1\|^{2+\delta} + E\|Z_2\|^{2+\delta} + E\|Z_1\|^2(E\|Z_2\|^2)^{\delta/2} \\ &\quad + E\|Z_2\|^2(E\|Z_1\|^2)^{\delta/2}, \end{aligned}$$

which proves the lemma. \square

Remark 3.1. If Z_1 and Z_2 are independent and identically distributed, then the result of Lemma 3.1 can be written as

$$E\|Z_1 + Z_2\|^{2+\delta} \leq 2E\|Z_1\|^{2+\delta} + 2(E\|Z_1\|^2)^{1+\delta/2}.$$

Let

$$I(r) = \sum_{\ell=1}^{\infty} (E\|X_0 - X_{0,\ell}\|^r)^{1/r}. \tag{3.1}$$

We note that by (1.4), $I(r) < \infty$ for all $2 \leq r \leq 2 + \delta$.

Lemma 3.2. *If (1.1)–(1.4) hold, then we have*

$$E \left\| \sum_{i=1}^n (X_i - X_{i,m}) \right\|^2 \leq nA,$$

where

$$A = \int E(X_0(t) - X_{0,m}(t))^2 dt + 2^{5/2} \left(\int E(X_0(t) - X_{0,m}(t))^2 dt \right)^{1/2} I(2). \tag{3.2}$$

Proof. Let $Y_i = X_i - X_{i,m}$. By Fubini’s theorem and the fact that the random variables are identically distributed, we conclude

$$\begin{aligned}
 E \left\| \sum_{i=1}^n Y_i \right\|^2 &= \int E \left(\sum_{i=1}^n Y_i(t) \right)^2 dt \\
 &= n \int E Y_0^2(t) dt + 2 \int \sum_{i=1}^{n-1} (n-i) E[Y_0(t) Y_i(t)] dt \\
 &\leq n \int E Y_0^2(t) dt + 2n \sum_{i=1}^{n-1} \int |E[Y_0(t) Y_i(t)]| dt \\
 &\leq n \int E Y_0^2(t) dt + 2n \sum_{i=1}^{\infty} \int |E[Y_0(t) Y_i(t)]| dt.
 \end{aligned}
 \tag{3.3}$$

We recall $X_{i,i}$ from (1.4). Under this definition, the random variables Y_0 and $X_{i,i}$ are independent for all $i \geq 1$. Let $Z_i = X_{i,m}$, if $i > m$ and $Z_i = g(\epsilon_i, \dots, \epsilon_1, \delta_i)$, if $1 \leq i \leq m$, where $\delta_i = (\delta_{i,0}, \delta_{i,-1}, \dots)$ and $\delta_{i,j}$ are iid copies of ϵ_0 , independent of the ϵ_ℓ 's and $\epsilon_{k,\ell}$'s. Clearly, Z_i and Y_0 are independent and thus with $Y_{i,i} = X_{i,i} - Z_i$ we have

$$E[Y_0(t) Y_i(t)] = E[Y_0(t) (Y_i(t) - Y_{i,i}(t))].$$

Furthermore, by first applying the Cauchy–Schwarz inequality for expected values and then by the Cauchy–Schwarz inequality for functions in L^2 , we get that

$$\begin{aligned}
 \int |E[Y_0(t) (Y_i(t) - Y_{i,i}(t))]| dt &\leq \int (E Y_0^2(t))^{1/2} (E [Y_i(t) - Y_{i,i}(t)]^2)^{1/2} dt \\
 &\leq \left(\int E Y_0^2(t) dt \right)^{1/2} \left(\int E [Y_i(t) - Y_{i,i}(t)]^2 dt \right)^{1/2}.
 \end{aligned}$$

Also,

$$\begin{aligned}
 &\int E [Y_i(t) - Y_{i,i}(t)]^2 dt \\
 &\leq 2 \left(\int E [X_i(t) - X_{i,i}(t)]^2 dt + \int E [X_{i,m}(t) - Z_i(t)]^2 dt \right).
 \end{aligned}$$

The substitution of this expression into (3.3) gives that

$$\begin{aligned}
 E \left\| \sum_{i=1}^n Y_i \right\|^2 &\leq n \int E Y_0^2(t) dt + 2^{3/2} n \sum_{i=1}^{\infty} \left(\int E Y_0^2(t) dt \right)^{1/2} \\
 &\quad \times \left\{ \left(\int E [X_i(t) - X_{i,i}(t)]^2 dt \right)^{1/2} \right. \\
 &\quad \left. + \left(\int E [X_{i,m}(t) - Z_i(t)]^2 dt \right)^{1/2} \right\} \\
 &\leq n \left[\int E Y_0^2(t) dt + 2^{5/2} \left(\int E Y_0^2(t) dt \right)^{1/2} I(2) \right],
 \end{aligned}$$

which completes the proof. \square

Theorem 3.1. *If (1.1)–(1.4) hold, then for all $N \geq 1$*

$$E \left\| \sum_{i=1}^N (X_i - X_{i,m}) \right\|^{2+\delta} \leq N^{1+\delta/2} B,$$

where

$$B = E \|X_0 - X_{0,m}\|^{2+\delta} + c_\delta^{2+\delta} [A^{1+\delta/2} + J_m^{2+\delta} + J_m A^{(1+\delta)/2} + A^{(1+\delta/2)\delta} J_m^2] + (c_\delta J_m^2)^{1/(1-\delta)} \tag{3.4}$$

with A defined in (3.2),

$$c_\delta = 36 \left(1 - \frac{1}{2^{\delta/2}} \right)^{-1} \tag{3.5}$$

and

$$J_m = 2(E \|X_0 - X_{0,m}\|^{2+\delta})^{(\kappa-2-\delta)/(\kappa(2+\delta))} \sum_{\ell=1}^\infty (E \|X_0 - X_{0,\ell}\|^{2+\delta})^{1/\kappa}.$$

Proof. We prove Theorem 3.1 using mathematical induction. By the definition of B , the inequality is obvious when $N = 1$. Assume that it holds for all k which are less than or equal to $N - 1$. We assume that N is even, i.e. $N = 2n$. The case when N is odd can be done in the same way with minor modifications. Let $Y_i = X_i - X_{i,m}$. For all i satisfying $n + 1 \leq i \leq 2n$, we define

$$X_{i,n}^* = g(\epsilon_i, \epsilon_{i-1}, \dots, \epsilon_{n+1}, \epsilon_n^*, \epsilon_{n-1}^*, \dots)$$

where the ϵ_j^* 's denote iid copies of ϵ_0 , independent of $\{\epsilon_i, -\infty < i < \infty\}$ and $\{\epsilon_{k,\ell}^*, -\infty < k, \ell < \infty\}$. We define $Z_{i,n} = X_{i,m}$, if $m + n + 1 \leq i \leq 2n$ and

$$Z_{i,n} = g(\epsilon_i, \dots, \epsilon_{n+1}, \epsilon_n^*, \dots, \epsilon_{i-m+1}^*, \delta_i) \quad \text{with } \delta_i = (\delta_{i,n}, \delta_{i,n-1}, \dots),$$

if $n + 1 \leq i \leq n + m$, where the $\delta_{k,\ell}$'s are iid copies of ϵ_0 , independent of the ϵ_k 's and $\epsilon_{k,\ell}^*$'s. Let $Y_{i,n}^* = X_{i,n}^* - Z_{i,n}$, if $n + 1 \leq i \leq 2n$. Under this definition, the sequences $\{Y_i, 1 \leq i \leq n\}$ and $\{Y_{i,n}^*, n + 1 \leq i \leq 2n\}$ are independent and have the same distribution. Let

$$\Theta = \left\| \sum_{i=1}^n Y_i + \sum_{i=n+1}^{2n} Y_{i,n}^* \right\| \quad \text{and} \quad \Psi = \left\| \sum_{i=n+1}^{2n} (Y_i - Y_{i,n}^*) \right\|.$$

By applying the triangle inequality for L^2 and expected values, we get

$$\begin{aligned} E \left\| \sum_{i=1}^{2n} Y_i \right\|^{2+\delta} &= E \left\| \sum_{i=1}^n Y_i + \sum_{i=n+1}^{2n} Y_{i,n}^* + \sum_{i=n+1}^{2n} (Y_i - Y_{i,n}^*) \right\|^{2+\delta} \\ &\leq E (\Theta + \Psi)^{2+\delta} \\ &\leq \left((E \Theta^{2+\delta})^{1/(2+\delta)} + (E \Psi^{2+\delta})^{1/(2+\delta)} \right)^{2+\delta}. \end{aligned} \tag{3.6}$$

The two-term Taylor expansion gives for all $a, b \geq 0$ and $r > 2$ that

$$(a + b)^r \leq a^r + r a^{r-1} b + \frac{r(r-1)}{2} (a + b)^{r-2} b^2. \tag{3.7}$$

Since both of the expected values in the last line of the inequality in (3.6) are positive, we obtain by (3.7) that

$$\begin{aligned}
 E \left\| \sum_{i=1}^{2n} Y_i \right\|^{2+\delta} &\leq E \Theta^{2+\delta} + (2 + \delta)(E \Theta^{2+\delta})^{(1+\delta)/(2+\delta)} (E \Psi^{2+\delta})^{1/(2+\delta)} \\
 &\quad + (2 + \delta)(1 + \delta) \left[(E \Theta^{2+\delta})^{1/(2+\delta)} + (E \Psi^{2+\delta})^{1/(2+\delta)} \right]^\delta (E \Psi^{2+\delta})^{2/(2+\delta)}. \tag{3.8}
 \end{aligned}$$

We proceed by bounding the terms $(E \Psi^{2+\delta})^{1/(2+\delta)}$, and $E \Theta^{2+\delta}$ individually. Applications of both the triangle inequality for L^2 and for expected values yield that

$$\begin{aligned}
 (E \Psi^{2+\delta})^{1/(2+\delta)} &= \left(E \left\| \sum_{i=n+1}^{2n} (Y_i - Y_{i,n}^*) \right\|^{2+\delta} \right)^{1/(2+\delta)} \\
 &\leq \left(E \left(\sum_{i=n+1}^{2n} \|Y_i - Y_{i,n}^*\| \right)^{2+\delta} \right)^{1/(2+\delta)} \\
 &\leq \sum_{i=n+1}^{2n} (E \|Y_i - Y_{i,n}^*\|^{2+\delta})^{1/(2+\delta)}.
 \end{aligned}$$

By Hölder’s inequality we have, with κ in (1.4),

$$\begin{aligned}
 (E \|Y_i - Y_{i,n}^*\|^{2+\delta})^{1/(2+\delta)} &= (E [\|Y_i - Y_{i,n}^*\|^{(2+\delta)^2/\kappa} \|Y_i - Y_{i,n}^*\|^{(2+\delta)-(2+\delta)^2/\kappa}])^{1/(2+\delta)} \\
 &\leq (E \|Y_i - Y_{i,n}^*\|^{2+\delta})^{1/\kappa} (E \|Y_i - Y_{i,n}^*\|^{2+\delta})^{(\kappa-2-\delta)/(\kappa(2+\delta))}.
 \end{aligned}$$

It follows from the definition of $Y_i, Y_{i,n}^*$ and the convexity of $x^{2+\delta}$ that

$$\begin{aligned}
 E \|Y_i - Y_{i,n}^*\|^{2+\delta} &\leq 2^{1+\delta} (E \|X_i - X_{i,n}^*\|^{2+\delta} + E \|X_{i,m} - Z_{i,n}\|^{2+\delta}) \\
 &\leq 2^{2+\delta} E \|X_0 - X_{0,i-n}\|^{2+\delta}
 \end{aligned}$$

and

$$\begin{aligned}
 E \|Y_i - Y_{i,n}^*\|^{2+\delta} &\leq 2^{1+\delta} (E \|X_i - X_{i,m}\|^{2+\delta} + E \|X_{i,n}^* - Z_{i,n}\|^{2+\delta}) \\
 &\leq 2^{2+\delta} E \|X_0 - X_{0,m}\|^{2+\delta}.
 \end{aligned}$$

Thus we get

$$\begin{aligned}
 (E \Psi^{2+\delta})^{1/(2+\delta)} &\leq 2(E \|X_0 - X_{0,m}\|^{2+\delta})^{(\kappa-2-\delta)/(\kappa(2+\delta))} \sum_{\ell=1}^{\infty} (E \|X_0 - X_{0,\ell}\|^{2+\delta})^{1/\kappa} \\
 &= J_m.
 \end{aligned}$$

To bound $E \Theta^{2+\delta}$, since $\sum_{i=1}^n Y_i$ and $\sum_{i=n+1}^{2n} Y_{i,n}^*$ are independent and have the same distribution, we have by Lemma 3.2, Remark 3.1 and the inductive assumption that

$$E \Theta^{2+\delta} = E \left\| \sum_{i=1}^n Y_i + \sum_{i=n+1}^{2n} Y_{i,n}^* \right\|^{2+\delta}$$

$$\begin{aligned} &\leq 2E \left\| \sum_{i=1}^n Y_i \right\|^{2+\delta} + 2 \left(E \left\| \sum_{i=1}^n Y_i \right\|^2 \right)^{1+\delta/2} \\ &\leq 2n^{1+\delta/2} B + 2(nA)^{1+\delta/2}. \end{aligned}$$

The substitution of these two bounds into (3.8) give that

$$\begin{aligned} E \left\| \sum_{i=1}^{2n} Y_i \right\|^{2+\delta} &\leq 2n^{1+\delta/2} B + 2(nA)^{1+\delta/2} \\ &\quad + (2 + \delta)[2n^{1+\delta/2} B + 2(nA)^{1+\delta/2}]^{(1+\delta)/(2+\delta)} J_m \\ &\quad + (2 + \delta)(1 + \delta) \left[2n^{1+\delta/2} B + 2(nA)^{1+\delta/2} + J_m \right]^\delta J_m^2. \end{aligned} \tag{3.9}$$

Furthermore, by the definition of B , we may further bound each summand on the right hand side of (3.9). We obtain for the first two terms that

$$\begin{aligned} 2n^{1+\delta/2} B + 2(nA)^{1+\delta/2} &\leq (2n)^{1+\delta/2} B \left[2^{-\delta/2} + \frac{A^{1+\delta/2}}{B} \right] \\ &\leq (2n)^{1+\delta/2} B \left[2^{-\delta/2} + 6c_\delta^{-1} \right]. \end{aligned}$$

A similar factoring procedure applied to the expression in the second line of (3.9) yields that

$$\begin{aligned} &(2 + \delta) \left[2n^{1+\delta/2} B + 2(nA)^{1+\delta/2} \right]^{(1+\delta)/(2+\delta)} J_m \\ &\leq 6 \left[(n^{1+\delta/2} B)^{(1+\delta)/(2+\delta)} + (nA)^{(1+\delta/2)[(1+\delta)/(2+\delta)]} \right] J_m \\ &\leq (2n)^{1+\delta/2} B \left[\frac{6J_m}{B^{1/(2+\delta)}} + \frac{6J_m A^{(1+\delta/2)[(1+\delta)/(2+\delta)]}}{B} \right] \\ &\leq (2n)^{1+\delta/2} B \left[12c_\delta^{-1} \right]. \end{aligned}$$

Since $0 < \delta < 1$, the expression in the third line of (3.9) may be broken into three separate terms:

$$\begin{aligned} &(2 + \delta)(1 + \delta) \left[2n^{1+\delta/2} B + 2(nA)^{1+\delta/2} + J_m \right]^\delta J_m^2 \\ &\leq 6(2n^{1+\delta/2} B)^\delta J_m^2 + 6(2(nA)^{(1+\delta/2)})^\delta J_m^2 + 6J_m^{2+\delta}. \end{aligned}$$

Furthermore by again applying the definition of B we have that

$$\begin{aligned} 6(2n^{1+\delta/2} B)^\delta J_m^2 &= (2n)^{1+\delta/2} B \left[\frac{6(2n^{1+\delta/2} B)^\delta J_m^2}{(2n)^{1+\delta/2} B} \right] \\ &\leq (2n)^{1+\delta/2} B \left[\frac{6J_m^2}{B^{1-\delta}} \right] \\ &\leq (2n)^{1+\delta/2} B [6c_\delta^{-1}], \\ 6(2(nA)^{(1+\delta/2)})^\delta J_m^2 &= (2n)^{1+\delta/2} B \left[\frac{6(2(nA)^{(1+\delta/2)})^\delta J_m^2}{(2n)^{1+\delta/2} B} \right] \end{aligned}$$

$$\begin{aligned} &\leq (2n)^{1+\delta/2} B \left[\frac{6A^{(1+\delta/2)\delta} J_m^2}{B} \right] \\ &\leq (2n)^{1+\delta/2} B [6c_\delta^{-1}], \end{aligned}$$

and

$$6J_m^{2+\delta} = (2n)^{1+\delta/2} B \left[\frac{6J_m^{2+\delta}}{(2n)^{1+\delta/2} B} \right] \leq (2n)^{1+\delta/2} B \left[\frac{6J_m^{2+\delta}}{B} \right] \leq (2n)^{1+\delta/2} B [6c_\delta^{-1}].$$

The application of these bounds to the right hand side of (3.9) give that

$$\begin{aligned} E \left\| \sum_{i=1}^{2n} Y_i \right\|^{2+\delta} &\leq (2n)^{1+\delta/2} B \left[2^{-\delta/2} + 36c_\delta^{-1} \right] \\ &= (2n)^{1+\delta/2} B, \end{aligned}$$

which concludes the induction step and thus the proof. \square

Theorem 3.2. *If (1.1)–(1.4) hold, then we have*

$$E \left(\max_{1 \leq k \leq N} \left\| \sum_{i=1}^k (X_i - X_{i,m}) \right\| \right)^{2+\delta} \leq a_m N^{1+\delta/2} \tag{3.10}$$

with some sequence a_m satisfying $a_m \rightarrow 0$ as $m \rightarrow \infty$.

Proof. By examining the proofs, it is evident that Theorem 3.1 in [24] holds for L^2 valued random variables. Furthermore, by the stationarity of the sequence $\{X_i - X_{i,m}\}_{i=1}^\infty$ and Theorem 3.1, the conditions of Theorem 3.1 in Móricz are satisfied and therefore

$$E \left(\max_{1 \leq k \leq N} \left\| \sum_{i=1}^k (X_i - X_{i,m}) \right\| \right)^{2+\delta} \leq c_\delta^* N^{1+\delta/2} B,$$

with some constant c_δ^* , depending only on δ and B is defined in (3.4). Observing that $B = B_m \rightarrow 0$ as $m \rightarrow \infty$, the result is proven. \square

Theorem 3.1 provides inequality for the moments of the norm of partial sums of $X_i - X_{i,m}$ which are not Bernoulli shifts. However, checking the proof of Theorem 3.1, we get the following result for Bernoulli shifts.

Theorem 3.3. *If (1.1), (1.3) are satisfied and \mathbf{X} is a Bernoulli shift satisfying*

$$I(2 + \delta) = \sum_{\ell=1}^\infty (E \|X_0 - X_{0,\ell}\|^{2+\delta})^{1/(2+\delta)} < \infty \quad \text{with some } 0 < \delta < 1,$$

where $X_{0,\ell}$ is defined by (1.4), then for all $N \geq 1$

$$E \left\| \sum_{i=1}^N X_i \right\|^{2+\delta} \leq N^{1+\delta/2} B_*,$$

where

$$\begin{aligned} B_* &= E \|X_0\|^{2+\delta} + c_\delta^{2+\delta} [A_*^{1+\delta/2} + I^{2+\delta}(2 + \delta) \\ &\quad + I(2 + \delta) A_*^{(1+\delta)/2} + A_*^{(1+\delta/2)\delta} I^2(2 + \delta)] + (c_\delta I^2(2))^{1/(1-\delta)}, \end{aligned}$$

$$A_* = \int EX_0^2(t)dt + 2 \left(\int EX_0^2(t)dt \right)^{1/2} I(2)$$

and c_δ is defined in (3.5) and $I(2)$ in (3.1).

Remark 3.2. The inequality in Theorem 3.1 is an extension of Proposition 4 in [2] to random variables in Hilbert spaces; we have computed how B_* depends on the distribution of \mathbf{X} explicitly.

Acknowledgment

We are grateful to Professor Herold Dehling for pointing out several important references.

References

- [1] A. Aue, S. Hörmann, L. Horváth, M. Hušková, J. Steinebach, Sequential testing for the stability of portfolio betas, *Econometric Theory* 28 (2012) 804–837.
- [2] I. Berkes, S. Hörmann, J. Schauer, Split invariance principles for stationary processes, *Annals of Probability* 39 (2011) 2441–2473.
- [3] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968.
- [4] D. Bosq, *Linear Processes in Function Spaces, Theory and Applications*, in: *Lecture Notes in Statistics*, vol. 149, Springer-Verlag, New York, 2000.
- [5] R.C. Bradley, *Introduction to Strong Mixing Conditions I–III*, Kendrick Press, Heber City UT, 2007.
- [6] Y. Chow, H. Teicher, *Probability Theory: Independence, Interchangeability, Martingales*, Springer-Verlag, Heidelberg, 1978.
- [7] A. DasGupta, *Asymptotic Theory of Statistics and Probability*, Springer, New York, 2008.
- [8] J. Dedecker, P. Doukhan, G. Lang, R.J.R. León, S. Louhichi, C. Prieur, Weak Dependence: With Examples and Applications, in: *Lecture Notes in Statistics*, vol. 190, Springer, New York, 2007, p. xiv+318.
- [9] J. Dedecker, F. Merlevède, The conditional central limit theorem in Hilbert spaces, *Stochastic Processes and Their Applications* 108 (2003) 229–262.
- [10] H. Dehling, Limit theorems for sums of weakly dependent Banach space valued random variables, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 63 (1983) 393–432.
- [11] H. Dehling, W. Philip, Almost sure invariance principles for weakly dependent vector-valued random variables, *Annals of Probability* 10 (1982) 689–701.
- [12] J. Doob, *Stochastic Processes*, Wiley, New York, 1953.
- [13] P. Doukhan, S. Louhichi, A new weak dependence condition and applications to moment inequalities, *Stochastic Processes and Their Applications* 84 (1999) 313–342.
- [14] E. Eberlein, An invariance principle for lattices of dependent random variables, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 50 (1979) 119–133.
- [15] A.M. Garsia, Continuity properties of Gaussian processes with multidimensional time parameter, in: *Proceedings of the 6th Berkeley Symposium Mathematical Statistics and Probability*, vol. 2, University of California, Berkeley, 1970, pp. 369–374.
- [16] M.I. Gordin, The central limit theorem for stationary processes, *Doklady Akademii Nauk SSSR* 188 (1969) 739–741 (in Russian).
- [17] G.H. Hardy, J.E. Littlewood, G. Pólya, *Inequalities*, second ed., Cambridge University Press, 1959.
- [18] S. Hörmann, L. Horváth, R. Reeder, A functional version of the ARCH model, *Econometric Theory* 2012 (in press).
- [19] S. Hörmann, P. Kokoszka, Weakly dependent functional data, *Annals of Statistics* 38 (2010) 1845–1884.
- [20] L. Horváth, P. Kokoszka, *Inference for Functional Data with Applications*, Springer, New York, 2012.
- [21] I.A. Ibragimov, Some limit theorems for stationary processes, *Theory of Probability and Its Applications* 7 (1962) 349–382.
- [22] J. Kuelbs, W. Philipp, Almost sure invariance principles for partial sums of mixing B -valued random variables, *Annals of Probability* 8 (1980) 1003–1036.
- [23] F. Merlevède, Central limit theorem for linear processes with values in a Hilbert space, *Stochastic Processes and Their Applications* 65 (1996) 103–114.
- [24] F. Móricz, R. Serfling, W. Stout, Moment and probability bounds with quasi-superadditive structure for the maximal partial sum, *Annals of Probability* 10 (1982) 1032–1040.

- [25] W. Philipp, W. Stout, Almost sure invariance principles for partial sums of weakly dependent random variables, *Memoirs of the American Mathematical Society* 161 (1975) iv+140.
- [26] G. Shorack, J. Wellner, *Empirical Processes With Applications To Statistics*, Wiley, New York, 1986.
- [27] W. Wu, Nonlinear system theory: another look at dependence, *Proceedings of the National Academy of Sciences of the USA* 102 (2005) 14150–14154.
- [28] W. Wu, Strong invariance principles for dependent random variables, *Annals of Probability* 35 (2007) 2294–2320.