WEAK INVARIANCE PRINCIPLES FOR SUMS OF DEPENDENT RANDOM FUNCTIONS

ISTVÁN BERKES, LAJOS HORVÁTH, AND GREGORY RICE

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.

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

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

$$\epsilon_j(t) = \epsilon_j(t, \omega) \text{ is jointly measurable in } (t, \omega) \;\; (j = 1, 2, \ldots), \tag{1.2}$$

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

and

(1.4) 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 $X_{n,\ell} = g(\epsilon_n, \epsilon_{n-1}, ..., \epsilon_{n-\ell+1}, \boldsymbol{\epsilon}^*_{n,\ell})$, $\boldsymbol{\epsilon}^*_{n,\ell} = (\epsilon^*_{n,\ell,n-\ell}, \epsilon^*_{n,\ell,n-\ell-1}, \ldots)$, where the $\epsilon^*_{n,\ell,k}$'s are independent copies of ϵ_0, independent of $\{\epsilon_i, -\infty < i < \infty\}$.

1991 *Mathematics Subject Classification.* 60F17, 60F25, 62G10.
Key words and phrases. variables in Hilbert spaces, m–approximability, weak convergence.
Research supported by FWF grant P24302-N18 (Berkes) and NSF grant DMS 0905400 (Horváth).

We note that assumption (1.1) implies that X_n is a stationary and ergodic sequence. Hörmann and Kokoszka (2010) 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 Ibragimov (1962) and is used frequently in the literature (cf. Billingsley (1968)). Aue et al (2012) provide several examples when assumptions (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. Hörmann et al (2010+)) and bilinear models (cf. Hörmann and Kokoszka (2010)) satisfy (1.4).

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

$$C(t,s) = EX_0(t)X_0(s) + \sum_{\ell=1}^{\infty} EX_0(t)X_\ell(s) + \sum_{\ell=1}^{\infty} EX_0(s)X_\ell(t) \tag{1.5}$$

are convergent in L^2. The function $C(t,s)$ is positive definite, and therefore there exist $\lambda_1 \geq \lambda_2 \geq \ldots \geq 0$ 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

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

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

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

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 Bosq (2000) and Horváth and Kokoszka (2012). In the real valued case, the martingale approach to weak dependence was developed by Gordin (1969) and Philipp and Stout (1975) and using such techniques, Merlevède (1996) and Dedecker and Merlevède (2003) 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 Kuelbs and Philipp (1980), Dehling and Philipp (1982) and Dehling (1983). 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. Bradley (2007)). The weak dependence concept of Doukhan and Louhichi (1999) (cf. also Dedecker et al. (2007)) solves this difficulty, but so far this concept has not been extended to variables in Hilbert spaces. Wu (2005, 2007) 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 Wu (2005, 2007). For extensions to the Hilbert space case we refer to Hörmann and Kokoszka (2010).

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*

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

Proof. The result is an immediate consequence of Markov's inequality and Theorem 3.2. □

Define

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

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

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

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

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

$$\iint C_m^2(t,s)dtds < \infty \quad \textit{for all} \;\; m \ge 1, \tag{2.4}$$

$$\lim_{m\to\infty}\iint (C(t,s)-C_m(t,s))^2 dtds = 0. \tag{2.5}$$

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

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

and

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

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

$$\iint (EX_0(t)X_0(s))^2 dtds \le \iint ((EX_0^2(t))^{1/2}(EX_0^2(s))^{1/2})^2 dtds = (E||X_0||^2)^2 < \infty.$$

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

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

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

$$\iint\left\{\sum_{i=1}^{m} EX_{0,m}(t)X_{i,m}(s)\right\}^2 dtds < \infty.$$

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 $1 \le i \le m$. Also, $(X_{i,m}, X_{i,i})$ has the same distribution as $(X_0, X_{0,i}), 1 \le i \le m$. Hence following the arguments in (2.9) we get

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

The proof of (2.4) is now complete. The arguments used above also prove (2.5). Repeating the previous arguments we have

$$\begin{aligned}
\int C(t,t)dt &\le \int EX_0^2(t)dt + 2\sum_{i=1}^{\infty} \int |EX_0(t)X_i(t)|dt \\
&= \int EX_0^2(t)dt + 2\sum_{i=1}^{\infty} \int |EX_0(t)(X_i(t) - X_{i,i}(t))|dt \\
&= \int EX_0^2(t)dt + 2\sum_{i=1}^{\infty} \int (EX_0^2(t))^{1/2}(E(X_i(t) - X_{i,i}(t))^2)^{1/2}dt \\
&\le E||X_0||^2 + 2\sum_{i=1}^{\infty} \left(\int EX_0^2(t)dt \right)^{1/2} \left(\int E(X_i(t) - X_{i,i}(t))^2 dt \right)^{1/2} \\
&= E||X_0||^2 + 2(E||X_0||^2)^{1/2} \sum_{i=1}^{\infty} (E||X_0 - X_{0,i}||^2)^{1/2} \\
&< \infty.
\end{aligned}$$

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). □

By the Karhunen–Loéve expansion, we have that

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

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*

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

then for all $x > 0$ *we have that*

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

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

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

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

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

then it follows from Doob's maximal inequality (Chow and Teicher, 1988 p. 247) that

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

which completes the proof. □

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

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

Proof. Define $Q_k(j) = \{i : 1 \le i \le k, i = j(mod\ \ m)\}$ for $j = 0, 1, ..., m-1$, 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

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

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

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

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 the assumption (1.3), we have that

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

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

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*

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

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

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

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 (1968), it is tight. This implies that the K dimensional process

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

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. DasGupta (2008) 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.

□

In light of the Skorkohod–Dudley–Wichura theorem (cf. Shorack and Wellner (1986), 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,*

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

as $N\to\infty$.

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*

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

as $N\to\infty$.

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

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

as $N\to\infty$, by Corollary 2.1. □

Lemma 2.7. *If (1.1)–(1.4) hold,*

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

as $K\to\infty$, 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

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

as $K\to\infty$. Therefore (2.20) follows from the Markov inequality. □

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*

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

as $N\to\infty$.

Proof. It follows from Lemmas 2.4-2.7. □

Since the distribution of $W_{\ell,N}, 1\le\ell<\infty$ 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 $\bar W_{\ell,m}(x), 1\le\ell<\infty$ such that*

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

as $m\to\infty$.

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 $x_i=i/M, 0\le i\le M$. It is easy to see that

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

$$= E\left\{\max_{0\le i<M}\sup_{0\le h\le 1/M}(W_1(x_i+h)-W_1(x_i))^2\right\}\sum_{\ell=1}^\infty \lambda_{\ell,m}.$$

Using Lemma 2.2 we get that

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

So by the modulus of continuity of the Wiener process (cf. Garsia (1970)) we get that

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

By the Karhunen-Loéve 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

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

Also, for any positive integer ℓ,

$$E(\langle\Delta_m(x,\cdot),\phi_\ell\rangle)^2=\iint C_m(t,s)\phi_\ell(t)\phi_\ell(s)dtds\to\iint C(t,s)\phi_\ell(t)\phi_\ell(s)dtds=\lambda_\ell,$$

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

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

The joint distribution of $\langle\Delta(x_i,\cdot),\phi_\ell\rangle, 1\le i\le M, 1\le\ell\le K$ 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

$$E\langle\Delta(x_i,\cdot),\phi_\ell\rangle\langle\Delta(x_j,\cdot),\phi_k\rangle=\min(x_i,x_j)\iint C_m(t,s)\phi_\ell(t)\phi_k(s)dtds$$
$$\to\min(x_i,x_j)\iint C(t,s)\phi_\ell(t)\phi_k(s)dtds=\min(x_i,x_j)\lambda_\ell I\{k=\ell\}.$$

Due to this covariance structure and the Skorkohod–Dudley–Wichura theorem (cf. Shorack and Wellner (1986), p. 47) we can find independent Wiener processes $\bar W_{\ell,m}(x), 1\le\ell<\infty$ such that

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

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

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

and therefore similarly to (2.23)

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

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

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

where W is a Wiener process. This also completes the proof of Lemma 2.9. □

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).

3. **A Maximal Inequality**

In this section we prove a maximal inequality for the sums of $Y_i = X_i - X_{i,m}$, where $X_{i,m}$ is defined in (1.4).
Our first lemma is a Hilbert space version of Doob's (1953 p. 226) inequality.

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

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

Proof. Since $0 < \delta \le 1$, for any $A, B \ge 0$ we have that $(A+B)^\delta \le A^\delta + B^\delta$ (cf. Hardy et al (1969, p. 32)). An application of this inequality along with Minkowski's inequality gives that

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

We also have by Hölders inequality that

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

This yields that

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

$$
\begin{aligned}
&= E||Z_1||^{2+\delta} + E||Z_2||^{2+\delta} + E||Z_1||^2 E||Z_2||^\delta + E||Z_2||^2 E||Z_1||^\delta \\
&\le E||Z_1||^{2+\delta} + E||Z_2||^{2+\delta} + E||Z_1||^2 (E||Z_2||^2)^{\delta/2} + E||Z_2||^2 (E||Z_1||^2)^{\delta/2},
\end{aligned}
$$

which proves the lemma.

□

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

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

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 $2 \le r \le 2 + \delta$.

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

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

where

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

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

$$
\begin{aligned}
E\left|\left|\sum_{i=1}^{n} Y_i\right|\right|^2 &= \int E\left(\sum_{i=1}^{n} Y_i(t)\right)^2 dt \qquad (3.3)\\
&= n\int EY_0^2(t)dt + 2\int \sum_{i=1}^{n-1}(n-i)EY_0(t)Y_i(t)dt \\
&\le n\int EY_0^2(t)dt + 2n\sum_{i=1}^{n-1}\int |EY_0(t)Y_i(t)|dt \\
&\le n\int EY_0^2(t)dt + 2n\sum_{i=1}^{\infty}\int |EY_0(t)Y_i(t)|dt.
\end{aligned}
$$

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

$$EY_0(t)Y_i(t) = EY_0(t)(Y_i(t) - Y_{i,i}(t)).$$

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

$$\int |EY_0(t)(Y_i(t)-Y_{i,i}(t))|dt \le \int \left(EY_0^2(t)\right)^{1/2}\left(E\left[Y_i(t)-Y_{i,i}(t)\right]^2\right)^{1/2} dt$$
$$\le \left(\int EY_0^2(t)dt\right)^{1/2}\left(\int E\left[Y_i(t)-Y_{i,i}(t)\right]^2 dt\right)^{1/2}.$$

Also,

$$\int E\left[Y_i(t)-Y_{i,i}(t)\right]^2 dt \le 2\left(\int E\left[X_i(t)-X_{i,i}(t)\right]^2 dt + \int E\left[X_{i,m}(t)-Z_i(t)\right]^2 dt\right)$$

The substitution of this expression into (3.3) gives that

$$E\left|\left|\sum_{i=1}^n Y_i\right|\right|^2 \le n\int EY_0^2(t)dt + 2^{3/2}n\sum_{i=1}^{\infty}\left(\int EY_0^2(t)dt\right)^{1/2}\left\{\left(\int E\left[X_i(t)-X_{i,i}(t)\right]^2 dt\right)^{1/2}\right.$$
$$\left. + \left(\int E\left[X_{i,m}(t)-Z_i(t)\right]^2 dt\right)^{1/2}\right\}$$
$$\le n\left[\int EY_0^2(t)dt + 2^{5/2}\left(\int EY_0^2(t)dt\right)^{1/2} I(2)\right],$$

which completes the proof.

□

Lemma 3.3. *If $a, b \ge 0$, then for $r > 2$ we have that*

$$(a+b)^r \le a^r + ra^{r-1}b + \frac{r(r-1)}{2}(a+b)^{r-2}b^2.$$

Proof. By Taylor expansion we have that

$$(a+b)^r = a^r + ra^{r-1}b + \frac{r(r-1)}{2}\xi^{r-2}b^2,$$

where $a \le \xi \le a+b$. Therefore, since $r > 2$, the function $g(\xi) = \xi^{r-2}$ is strictly increasing, and we have that

$$a^r + ra^{r-1}b + \frac{r(r-1)}{2}\xi^{r-2}b^2 \le a^r + ra^{r-1}b + \frac{r(r-1)}{2}(a+b)^{r-2}b^2.$$

We may then conclude that

$$(a+b)^r \le a^r + ra^{r-1}b + \frac{r(r-1)}{2}(a+b)^{r-2}b^2,$$

as needed. This completes the proof.

□

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

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

where

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

$$+ (c_\delta J_m^2)^{1/(1-\delta)}$$

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 $n+1 \le i \le 2n$, we define

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

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

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

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

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

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

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

Since both of the expected values in the last line of the inequality above are positive, we obtain by Lemma 3.3 that

$$\begin{aligned} E\left|\left|\sum_{i=1}^{2n} Y_i\right|\right|^{2+\delta} \le & E\Theta^{2+\delta} + (2+\delta)(E\Theta^{2+\delta})^{(1+\delta)/(2+\delta)}(E\Psi^{2+\delta})^{1/(2+\delta)} \\ & + (2+\delta)(1+\delta)\big[(E\Theta^{2+\delta})^{1/(2+\delta)} \\ & + (E\Psi^{2+\delta})^{1/(2+\delta)}\big]^{\delta}(E\Psi^{2+\delta})^{2/(2+\delta)}. \end{aligned} \tag{3.6}$$

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

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

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

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

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

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

and

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

Thus we get

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

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

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

The substitution of these two bounds into (3.6) give that

$$\begin{aligned}
E\left|\left|\sum_{i=1}^{2n} Y_i\right|\right|^{2+\delta} &\le 2n^{1+\delta/2}B + 2(nA)^{1+\delta/2} \\
&\quad + (2+\delta)[2n^{1+\delta/2}B + 2(nA)^{1+\delta/2}]^{(1+\delta)/(2+\delta)}J_m \\
&\quad + (2+\delta)(1+\delta)\left[2n^{1+\delta/2}B + 2(nA)^{1+\delta/2} + J_m\right]^{\delta}J_m^2.
\end{aligned} \tag{3.7}$$

Furthermore, by the definition of B, we may further bound each summand on the right hand side of (3.7). We obtain for the first two terms that

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

A similar factoring procedure applied to the expression in the second line of (3.7) yields that

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

Since $0 < \delta < 1$, the expression in the third line of (3.7) may be broken into three separate terms:

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

Furthermore by again applying the definition of B we have that

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

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

and

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

The application of these bounds to the right hand side of (3.7) give that

$$E\left|\left|\sum_{i=1}^{2n} Y_i\right|\right|^{2+\delta} \le (2n)^{1+\delta/2}B\left[2^{-\delta/2} + 36c_\delta^{-1}\right]$$

$$= (2n)^{1+\delta/2}B,$$

which concludes the induction step and thus the proof. □

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

$$E\left(\max_{1\le k\le N}\left|\left|\sum_{i=1}^{k}(X_i - X_{i,m})\right|\right|\right)^{2+\delta} \le a_m N^{1+\delta/2} \tag{3.8}$$

with some sequence a_m satisfying $a_m \to 0$ as $\to \infty$.

Proof. By examining the proofs, it is evident that Theorem 3.1 in Móricz et al (1982) 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

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

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

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 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 then for all $N \ge 1$

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

where

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

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

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 Berkes et al (2011) to random variables in Hilbert spaces; we have computed how B_* depends on the distribution of $\mathbf{X}$ explicitly.

Acknowledgement We are grateful to Professor Harold Dehling to pointing out several important references.

References

[1] Aue, A., Hörmann, S., Horváth, L., Hušková, M. and J. Steinebach, J.: Sequential testing for the stability of portfolio betas. *Econometric Theory*, **28**(2012), 804–837.

[2] Berkes, I. Hörmann, S. and Schauer, J.: Split invariance principles for stationary processes, *Annals of Probability*, to appear.

[3] Billingsley, P.:*Convergence of Probability Measures*, Wiley, New York, 1968.

[4] Bosq, D.: *Linear Processes in Function Spaces. Theory and Applications.* Lecture Notes in Statistics, 149. Springer-Verlag, New York, 2000.

[5] Bradley, R.C.: *Introduction to Strong Mixing Conditions I–III.* Kendrick Press, Heber City UT, 2007.

[6] Chow, Y. and Teicher, H.: *Probability Theory: Independence, Interchangeability, Martingales*, Springer-Verlag, Heidelberg, 1978.

[7] DasGupta, A.:*Asymptotic Theory of Statistics and Probability*, Springer, New York, 2008.

[8] Dedecker, J. Doukhan, P. Lang, G. León, R.J.R. Louhichi, S. and Prieur, C.:*Weak Dependence: with Examples and Applications.* Lecture Notes in Statistics, **190** Springer, New York, 2007. xiv+318 pp.

[9] Dedecker, J. and Merlevède, F.: The conditional central limit theorem in Hilbert spaces. *Stochastic Processes and Their Applications* **108** (2003), 229262.

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

[11] Dehling, H. and Philip, W.: Almost sure invariance principles for weakly dependent vector-valued random variables.*Annals of Probability* **10** (1982), 689-701.

[12] Doob, J.: *Stochastic Processes* , Wiley, New York, 1953.

[13] Doukhan, P. and Louhichi, S.: A new weak dependence condition and applications to moment inequalities.*Stochastic Processes and Their Applications* **84** (1999), 313-342.

[14] Eberlein, E;: An invariance principle for lattices of dependent random variables. *Z. Wahrsch. Verw. Gebiete* **50**(1979), 119-133.

[15] Garsia, A.M.: Continuity properties of Gaussian processes with multidimensional time parameter. In: *Proceedings of the* 6^{th}*Berkeley Symposium Math. Stat. Probab.*, University of California, Berkeley, Vol. 2 (1970), pp. 369–374.

[16] Gordin, M.I.: The central limit theorem for stationary processes. (Russian) *Dokl. Akad. Nauk SSSR* **188** (1969), 739-741.

[17] Hardy, G.H., Littlewood, J.E. and Pólya, G.: *Inequalities.* (Second Edition) Cambridge University Press, 1959.

[18] Hörmann, S. and Kokoszka, P.: Weakly dependent functional data. *Annals of Statistics* **38** (2010), 1845–1884.

[19] Hörmann, S., Horváth, L. and Reeder, R.: *A functional version of the ARCH model.* Preprint.

[20] Horváth, L. and Kokoszka, P.: *Inference for Functional Data with Applications.* Springer, New York, 2012.

[21] Ibragimov, I. A.: Some limit theorems for stationary processes.*Theory of Probabability and Its Applications* **7** (1962), 349–382.

[22] Kuelbs, J. and Philipp, W.: Almost sure invariance principles for partial sums of mixing B-valued random variables. *Annals of Probabability* **8** (1980), 1003-1036.

[23] Merlevède, F.: Central limit theorem for linear processes with values in a Hilbert space. *Stochastic Processes and Their Applications* **65** (1996), 103-114.

[24] Móricz, F. Serfling, R. and Stout, W.: Moment and probability bounds with quasi-superadditive structure for the maximal partial sum, *Annals of Probability*, **10**(1982), 1032–1040.

[25] Philipp, W. and Stout, W.: Almost sure invariance principles for partial sums of weakly dependent random variables.*Memoirs of the Amererican Mathematical Society* **161**(1975), iv+140 pp.

[26] Shorack, G. and Wellner, J.: *Empirical Processes With Applications To Statistics*, Wiley, New York, 1986.

[27] Wu, W.: Nonlinear system theory: another look at dependence. *Proc. Natl. Acad. Sci. USA* **102** (2005), 14150-14154.

[28] Wu, W.: Strong invariance principles for dependent random variables. *Annals of Probability* **35** (2007), 2294-2320.

István Berkes, Institute of Statistics, Graz University of Technology, Kopernikus Gasse 24/3, A–8010 Graz, Austria.

Lajos Horváth, Department of Mathematics, University of Utah, Salt Lake City, UT 84112–0090 USA

Gregory Rice, Department of Mathematics, University of Utah, Salt Lake City, UT 84112–0090 USA