Elect. Comm. in Probab. **2** (1997) 27–41

ELECTRONIC COMMUNICATIONS in PROBABILITY

STRONG LAWS AND SUMMABILITY FOR SEQUENCES OF φ-MIXING RANDOM VARIABLES IN BANACH SPACES

RÜDIGER KIESEL
Department of Statistics, Birkbeck College
University of London
Malet Street
London WC1E 7HX, England.
email: r.kiesel@stat.bbk.ac.uk

submitted September 15, 1996; *revised* March 19, 1997

AMS 1991 Subject classification: 60F15, (40G05, 40G10)
Keywords and phrases: Strong Laws, φ-mixing, Summability.

Abstract

In this note the almost sure convergence of stationary, φ-mixing sequences of random variables with values in real, separable Banach spaces according to summability methods is linked to the fulfillment of a certain integrability condition generalizing and extending the results for i.i.d. sequences. Furthermore we give via Baum-Katz type results an estimate for the rate of convergence in these laws.

1 Introduction and main result

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space rich enough so that all random variables used in the sequel can be defined on this space.
If $X_0, X_1, \ldots$ is a sequence of independent, identically distributed (i.i.d.) real valued random variables, then the almost sure (a.s.) convergence of such a sequence according to certain summability methods is equivalent to the fulfillment of certain integrability conditions on X_0, see e.g.[5, 6, 12, 19, 24].
Some of the above results have been extended to sequences of stationary, φ-mixing sequences of real-valued random variables [3, 5, 22, 28] and to i.i.d. Banach space-valued random variables [4, 9, 14, 15, 18, 25].
The aim of this paper is to prove a general result linking convergence according to summability methods with integrability conditions for stationary, φ-mixing random variables taking values in a real separable Banach space $(\mathbb{B}, \|.\|)$ equipped with its Borel σ-algebra.
A result of this type has useful statistical applications, e.g. to the empirical distribution function, to the likelihood functions, to strong convergence of density estimators and to least square regression with fixed design [25, 27, 31].

Before we state our main result we give the minimal amount of necessary definitions. We start out with the relevant summability methods (for general information on summability see [13, 29, 32]).
Let (p_n) be a sequence of real numbers such that

$$(1.1) \qquad \begin{cases} p_0 > 0, p_n \geq 0,\ n = 1, 2, \ldots, \quad \text{and the power series} \\ p(t) := \sum_{n=0}^{\infty} p_n t^n \quad \text{has radius of convergence } R \in (0, \infty]. \end{cases}$$

We say that a sequence (s_n) is summable to s by the **power series method** (P), briefly $s_n \to s\ (P)$, if

$$\begin{cases} p_s(t) = \sum_{n=0}^{\infty} s_n p_n t^n \quad \text{converges for } |t| < R \\ \text{and} \quad \sigma_p(t) = \dfrac{p_s(t)}{p(t)} \to s \quad \text{for} \quad t \to R- . \end{cases}$$

Observe that the classical Abel ($p_n \equiv 1$) and Borel methods ($p_n = 1/n!$) are in the class of power series methods. We assume throughout the following regularity condition

$$(1.2) \qquad p_n \sim \exp\{-g(n)\} \quad (n \to \infty)$$

with a real-valued function $g(.)$, which has the following properties, see [8],

$$(\mathrm{C}) \qquad \begin{cases} g \in C_2[t_0, \infty) \text{ with } t_0 \in \mathbb{N}; \\ g''(t) \text{ is positive and non-increasing with } \lim_{t\to\infty} g''(t) = 0; \\ G(t) := t^2 g''(t) \text{ is non-decreasing on } [t_0, \infty). \end{cases}$$

As the corresponding family of matrix methods we use the **generalized Nörlund methods**, see [20] for a general discussion. We say a sequence (s_n) is summable to s by the generalized Nörlund method $(N, p^{*\kappa}, p)$, briefly $s_n \to s\ (N, p^{*\kappa}, p)$, if

$$\sigma_n^\kappa := \frac{1}{p_n^{*(\kappa+1)}} \sum_{\nu=0}^{n} p_{n-\nu}^{*\kappa} p_\nu s_\nu \to s \quad (n \to \infty),$$

where we define the convolution of a sequence (p_n) by

$$p_n^{*1} := p_n \quad \text{and} \quad p_n^{*\kappa} := \sum_{\nu=0}^{n} p_{n-\nu}^{*(\kappa-1)} p_\nu, \quad \text{for} \quad \kappa = 2, 3, \ldots .$$

In order to define two more especially in probability theory widely used summability methods, we need a few function classes.
We call a measurable function $f : (0, \infty) \to (0, \infty)$

(i) **self-neglecting**, if f is continuous, $\boldsymbol{o}(t)$ at ∞, and

$$f(t + uf(t))/f(t) \to 1 \ \ (t \to \infty),\ \forall u \in \mathbb{R}.$$

We write briefly: $f \in SN$.

(ii) of **bounded increase**, if

$$f(\lambda t)/f(t) \leq C\lambda^{\alpha_0} \quad (1 \leq \lambda \leq \Lambda, t \geq t_0)$$

with suitable constants $\Lambda > 1, C, t_0, \alpha_0$. We write briefly: $f \in BI$.

(iii) **bounded decrease**, if

$$f(\lambda t)/f(t) \geq C\lambda^{\beta_0} \quad (1 \leq \lambda \leq \Lambda, t \geq t_0)$$

with suitable constants $\Lambda > 1, C, t_0, \beta_0$. We write briefly: $f \in BD$.

(iv) **regular varying with index** ρ, if

$$f(\lambda t)/f(t) \to \lambda^{\rho} \quad (t \to \infty) \quad \forall \lambda > 0.$$

We write briefly: $f \in R_\rho$.

For properties and relations of these classes see N.H.Bingham et al. [7], §§1.5, 1.8, 2.1, 2.2, 2.3.
We call a sequence (s_n) summable by the **moving average** (M_ϕ), briefly $s_n \to s$ (M_ϕ), if

$$\frac{1}{u\phi(t)} \sum_{t<n\leq t+u\phi(t)} s_n \to s \quad (t \to \infty) \quad \forall u > 0,$$

with $\phi \in SN$. The above convergence is locally uniform in u (N.H.Bingham et al., [7], §2.11).
Finally we call a sequence (s_n) summable by the **generalized Valiron method** (V_ϕ), briefly $s_n \to s$ (V_ϕ), if

$$\frac{1}{\sqrt{2\pi}\phi(t)} \sum_{n=0}^{\infty} s_n \exp\left\{-\frac{(t-n)^2}{2\phi(t)^2}\right\} \to s \quad (t \to \infty),$$

with $\phi \in SN$.
For a general discussion of these methods and a general equivalence Theorem see [20, 30]. As a measure of dependence we use a strong mixing condition. We write $\mathcal{F}_n^m := \sigma(X_k : n \leq k \leq m)$ for the canonical σ-algebra generated by $X_n, \ldots, X_m$ and define the $\varphi-$ mixing coefficient by

$$\varphi_n := \sup_{k\geq 1} \sup_{\substack{A\in\mathcal{F}_1^k, \mathbb{P}(A)\neq 0 \\ B\in\mathcal{F}_{k+n}^\infty}} |\mathbb{P}(B|A) - \mathbb{P}(B)|\,.$$

We say, that a sequence $X_0, X_1, \ldots$ is φ-mixing, if $\varphi_n \to 0$ for $n \to \infty$.
We can now state our main result

Theorem. *Let $\{X_n\}$ be a stationary φ-mixing sequence of random variables taking values in the Banach space $\mathbb{B}$ and assume that $\varphi_1 < 1/4$.*
Moreover let (p_n) be a sequence of real numbers satisfying condition (1.2) *and p_n/p_{n+1} is non-decreasing.*
If $\phi(.) = 1/\sqrt{g''(.)}$ has an absolutely continuous inverse $\psi(.) = \phi^{\leftarrow}(.)$ with positive derivative $\psi' \in BI \cap BD$ with $\beta_0 > 0$, then the following are equivalent

(M) $\mathbb{E}(\psi(\|X\|)) < \infty,\ E(X) = \mu$ *(in the Bochner sense);*

$(A1)$ $\sum_{n=1}^{\infty} n^{-1}(\psi(n+1)-\psi(n))I\!P\left(\max_{1\le k\le n}\|S_k-k\mu\|>\varepsilon n\right)<\infty \;\; \forall\varepsilon>0;$

$(A2)$ $\sum_{n=1}^{\infty} n^{-1}(\psi(n+1)-\psi(n))I\!P\left(\|S_n-n\mu\|>\varepsilon n\right)<\infty \;\; \forall\varepsilon>0;$

$(S1)$ $X_n \to \mu \; (M_\phi) \; a.s.;$

$(S2)$ $X_n \to \mu \; (V_\phi) \; a.s.;$

$(S3)$ $X_n \to \mu \; (N, p^{*\kappa}, p) \; a.s. \quad \forall\kappa\in I\!N;$

$(S4)$ $X_n \to \mu \; (P) \; a.s..$

Remark 1.

(i) As an example consider the Borel case, i.e. $p_n = 1/n!$. These weights satisfy (1.2) with function $g(t) = t\log t - t + \frac{1}{2}\log(2\pi)$ (use Stirlings formula). We then get $\phi(t) = \sqrt{t}$ and $\psi(t) = t^2$. Hence we can use the Theorem with moment condition $I\!E(\|X\|^2) < \infty$. For the i.i.d. case we therefore obtain the results proved in [25], Theorem 4. Observe that the matrix methods in $(S3)$ are the Euler methods.

(ii) Using $g(t) = -t^\alpha, 0 < \alpha < 1$ resp. $g(t) = t^\beta, 1 < \beta < 2$ in (1.2) we find $\phi(t) = c(\alpha)t^{1-\frac{\alpha}{2}}$ and $\psi(t) = \tilde{c}(\alpha)t^{\frac{2}{2-\alpha}}$ resp. $\phi(t) = c(\beta)t^{1-\frac{\beta}{2}}$ and $\psi(t) = \tilde{c}(\beta)t^{\frac{2}{2-\beta}}$ and therefore get via $(M) \Leftrightarrow (A1) \Leftrightarrow (A2)$ Theorem 1 in [25] and via $(M) \Leftrightarrow (S1) \Leftrightarrow (S2)$ Theorem 3.2 in [14] resp. Theorem 1.2 in [15] in the i.i.d. case.

(iii) For the case of real-valued random variables the Theorem was proved in [22] using techniques from [3, 5, 28].

(iv) The case of Abel's method, $p_n \equiv 1$, is not directly included, but it can be viewed as a limiting case, compare [8]. For the real-valued mixing case the equivalence of $(M) \Leftrightarrow (S3) \Leftrightarrow (S4)$ has basically been proved in Theorem 6 in [3] and in the i.i.d. Banach-valued case in [9]. Observe that the matrix method in $(S3)$ is Cesàro's method.

(v) Let $Y_1, Y_2, \ldots$ be φ-mixing and uniformly distributed on $(0,1)$ and $F_n(t) = n^{-1}\sum_{i=1}^n \mathbf{1}_{[Y_i\le t]}$ be the empirical distribution function based on $Y_1, Y_2, \ldots Y_n$. As in Lai [25] the above theorem might be extended to discuss the specific behaviour of $F_n(t) - t$ for t near 0 and 1. Likewise one can use the theorem to discuss certain likelihood functions (see also [25]).

(vi) The above theorem can also be used to obtain results on strong convergence of kernel estimators in non-parametric statistics in the spirit of Liebscher [27] (see also [23] for results of Erdős-Rényi-Shepp type related to kernel estimators).

(vii) Consider the problem of least square regression with fixed design (for a precise formulation of the problem and background see [31], §3.4.3.1). In that context stochastic processes of the form $\{n^{-\frac{1}{2}}\sum_{i=1}^n \theta(x_i)e_i : \theta\in\Theta\}$ play a central role (typically $x_i \in I\!R^d$, Θ is a set of functions with $\theta : I\!R^d \to I\!R$ and the error terms e_i are i.i.d.). Imposing regularity conditions on Θ the theorem can be used to discuss the speed of convergence of the above process (even if independence is replaced by the appropriate φ-mixing condition).

2 Auxiliary results

We start with an application of the Feller-Chung lemma for events generalizing a result for i.i.d. real-valued random variables in [12], Lemma 2.

Lemma 1. *Let $\{Y_n\}$ and $\{Z_n\}$ be sequences of $\mathbb{B}$-valued random variables. Set $\mathcal{K}_1^n := \sigma(Y_i; 1 \le i \le n)$ and assume*

$$\varphi := \varphi(\mathcal{K}_1^n, \sigma(Z_n)) = \sup_n \sup_{\substack{A\in\mathcal{K}_1^n, \mathbb{P}(A)\neq 0,\\ B\in\sigma(Z_n)}} |\mathbb{P}(B|A) - \mathbb{P}(B)| < 1.$$

Then $\|Z_n\| \xrightarrow{p} 0$ and $Y_n + Z_n \xrightarrow{a.s.} 0$ imply $Y_n \xrightarrow{a.s.} 0$.
The proof follows the lines of Lemma 2 in [12] using only standard properties of φ-mixing sequences, e.g. Lemma 1.1.1 in [17].
As a key result we now prove a Lévy-type inequality using techniques from [25], Lemma 1, [28] Lemma 3.2.

Lemma 2. *Let $\{X_n\}$ be a sequence of $\mathbb{B}$-valued random variables such that $\varphi_1 < 1/4$. Let $\{X_n'\}$ be an independent copy of $\{X_n\}$ and consider the symmetrized sequence $\{X_n^s\}$, such that for every n $X_n^s = X_n - X_n'$. Denote by $S_n^s = \sum_{k=1}^n X_k^s$. Then we have for every $\varepsilon > 0$*

$$\mathbb{P}\left(\max_{1\le k\le n} \|S_k^s\| > \varepsilon\right) \le \left(\frac{1}{2} - 2\varphi_1\right)^{-1} \mathbb{P}(\|S_n^s\| > \varepsilon).$$

Proof:
According to Bradley [10] Theorem 3.2 the φ-mixing coefficients for $\{X_n^s\}$, which we denote by $\{\varphi_n^s\}$, cannot exceed twice the size of the φ-mixing coefficients for $\{X_n\}$. So we have $\varphi_n^s \le 2\varphi_n$ $\forall n$. Now we can basically follow the proof of Lemma 1 in [25] with the only modification being that we use the definition of $\{\varphi_n^s\}$ instead of independence. We outline the main steps. For notational convenience we assume, that $\{X_n\}$ itself is the symmetrized sequence.
First assume that $X_k = (X_k^{(1)}, \ldots, X_k^{(d)})$ is a finite dimensional random vector. Let $S_k^{(j)} := X_1^{(j)} + \ldots + X_k^{(j)}$. Then we claim:

$$\mathbb{P}\left(\max_{1\le j\le d} \max_{1\le k\le n} \left|S_k^{(j)}\right| > \varepsilon\right) \le \left(\frac{1}{2} - 2\varphi_1\right)^{-1} \mathbb{P}\left(\max_{1\le j\le d} \left|S_n^{(j)}\right| > \varepsilon\right).$$

Define the stopping times

$$\tau_j := \inf\left\{1 \le k \le n : S_k^{(j)} > \varepsilon\right\} \quad \text{and} \quad \sigma_j := \inf\left\{1 \le k \le n : S_k^{(j)} < -\varepsilon\right\},$$

with $\inf\emptyset = n+1$. For $k = 1, \ldots, n$ consider the following sets

$$A_k^{(j)} := \{\tau_j = k \le \min\{\min\{\tau_\nu, \sigma_\nu\} : \nu \neq j\}, k < \sigma_j\};$$
$$B_k^{(j)} := \left\{\sigma_j = k \le \min\{\sigma_\nu : \nu \neq j\}, k < \min_{1\le l\le d} \tau_l\right\}.$$

Observe that for $1 \le k \le n$ we have $A_k^{(j)}, B_k^{(j)} \in \sigma(X_1, \dots X_k)\ \forall j$. Now it follows that

$$\begin{aligned} I\!P\left(\max_{1\le j\le d}\max_{1\le k\le n}\left|S_k^{(j)}\right| > \varepsilon\right) &\le \sum_{k=1}^{n} I\!P\left(\bigcup_{j=1}^{d} A_k^{(j)}\right) + \sum_{k=1}^{n} I\!P\left(\bigcup_{j=1}^{d} B_k^{(j)}\right) \\ &\le \sum_{k=1}^{n}\sum_{j=1}^{d} I\!P\left(C_k^{(j)}\right) + \sum_{k=1}^{n}\sum_{j=1}^{d} I\!P\left(D_k^{(j)}\right), \end{aligned}$$

where we define

$$C_k^{(1)} := A_k^{(1)}, \quad \dots \quad C_k^{(j)} := A_k^{(j)} \cap \left(A_k^{(1)} \cup \dots \cup A_k^{(j-1)}\right)^c, \ \ 2 \le j \le d;$$

$$D_k^{(1)} := B_k^{(1)}, \quad \dots \quad D_k^{(j)} := B_k^{(j)} \cap \left(B_k^{(1)} \cup \dots \cup B_k^{(j-1)}\right)^c, \ \ 2 \le j \le d.$$

Again $C_k^{(j)}, D_k^{(j)} \in \sigma(X_1, \dots X_k)$. Using the mixing condition (instead of independence as in Lai's Lemma) we get for $k < n, 1 \le j \le d$

$$\begin{aligned} I\!P(C_k^{(j)}) &\le \left(\frac{1}{2} - 2\varphi_1\right)^{-1} I\!P\left(C_k^{(j)} \cap \left\{(X_{k+1}^{(j)} + \dots + X_n^{(j)}) \ge 0\right\}\right) \\ &\le \left(\frac{1}{2} - 2\varphi_1\right)^{-1} I\!P\left(C_k^{(j)} \cap \left\{S_n^{(j)} > \varepsilon\right\}\right). \end{aligned}$$

Obviously

$$I\!P(C_n^{(j)}) \le \left(\frac{1}{2} - 2\varphi_1\right)^{-1} I\!P\left(C_n^{(j)} \cap \left\{S_n^{(j)} > \varepsilon\right\}\right).$$

With the same arguments we get for $1 \le j \le d, 1 \le k \le n$

$$I\!P(D_k^{(j)}) \le \left(\frac{1}{2} - 2\varphi_1\right)^{-1} I\!P\left(D_k^{(j)} \cap \left\{S_n^{(j)} < -\varepsilon\right\}\right).$$

Hence

$$\begin{aligned} & I\!P\left(\max_{1\le j\le d}\max_{1\le k\le n}\left|S_k^{(j)}\right| > \varepsilon\right) \\ \le\ & \left(\frac{1}{2} - 2\varphi\right)^{-1} \sum_{k=1}^{n}\sum_{j=1}^{d} I\!P\left(C_k^{(j)} \cap \left\{\max_{1\le\nu\le d}\left|S_n^{(\nu)}\right| > \varepsilon\right\}\right) \\ +\ & \left(\frac{1}{2} - 2\varphi\right)^{-1} \sum_{k=1}^{n}\sum_{j=1}^{d} I\!P\left(D_k^{(j)} \cap \left\{\max_{1\le\nu\le d}\left|S_n^{(\nu)}\right| > \varepsilon\right\}\right) \\ \le\ & \left(\frac{1}{2} - 2\varphi\right)^{-1} I\!P\left(\max_{1\le j\le d}\left|S_n^{(j)}\right| > \varepsilon\right). \end{aligned}$$

Turning to the general case we find, since $I\!B$ is separable, a countable, dense subset $D := \{f_1, f_2, \dots\}$ of $I\!B'$ with

$$\|b\| = \sup\left\{|f_n(b)| : f_n \in D\right\}$$

for all $b \in \mathbb{B}$, see [16] p.34. Since $\{X_n\}$ is a symmetrized sequence, so is $\{f(X_n)\}$ for all $f \in \mathbb{B}'$ with φ-mixing coefficients not exceeding the mixing coefficients of $\{X_n\}$ (see [10], p.170). Therefore

$$\begin{aligned} \mathbb{P}\left(\max_{1\le k\le n} \|S_k\| > \varepsilon\right) &= \mathbb{P}\left(\max_{1\le k\le n}\sup_{d\ge 1} |f_d(S_k)| > \varepsilon\right) \\ &= \lim_{N\to\infty} \mathbb{P}\left(\max_{1\le k\le n}\max_{1\le d\le N} |f_d(S_k)| > \varepsilon\right) \\ &\le \left(\frac{1}{2} - 2\varphi_1\right)^{-1} \lim_{N\to\infty} \mathbb{P}\left(\max_{1\le d\le N} |f_d(S_n)| > \varepsilon\right) \\ &= \left(\frac{1}{2} - 2\varphi_1\right)^{-1} \mathbb{P}(\|S_n\| > \varepsilon). \end{aligned}$$ □

3 A reduction principle

We now state and prove a general reduction principle which allows us to deduce results for $\mathbb{B}$-valued random variables from the corresponding results for real-valued random variables. Let (V) be a summability method with weights $c_n(\lambda) \ge 0,\ n = 0, 1, \ldots; \lambda > 0$ a discrete or continuous parameter $\sum_{n=0}^{\infty} c_n(\lambda) = 1\ \ \forall\lambda$. Denote the (V)-transform of a sequence (s_n) by

$$V_s(\lambda) := \sum_{n=0}^{\infty} c_n(\lambda) s_n.$$

We say $s_n \to s\ (V)$, if $V_s(\lambda) \to s\ (\lambda \to \infty)$.
Assume that if $\{X_n\}$ is a stationary φ-mixing sequence of real-valued random variables with mixing coefficient $\varphi_1 < 1/4$ and if ψ is a function as in our Theorem, then

$$\mathbb{E}(\psi(|X|)) < \infty, \mathbb{E}(X) = \mu \quad \Rightarrow V_X(\lambda) \to \mu\ a.s..$$

Under these assumptions we have

Proposition. *If $\{X_n\}$ is a stationary φ-mixing sequence of $\mathbb{B}$-valued random variables with $\varphi_1 < 1/4$, then*

$$\mathbb{E}(\psi(\|X\|)) < \infty, \mathbb{E}(X) = \mu \text{ (Bochner) } \quad \textit{implies} \quad X_n \to \mu\ (V)\ a.s..$$

Proof:
Since $\mathbb{B}$ is separable, we can find a dense sequence $(b_n), n \ge 1$. For each $n \ge 1$ define

$$\begin{aligned} A_{1n} &:= \left\{b \in \mathbb{B} : \|b - b_1\| < \tfrac{1}{n}\right\}; \\ A_{in} &:= \left\{b \in \mathbb{B} : \|b - b_i\| < \tfrac{1}{n}\right\} \cap A_{1n}^c \cap \ldots \cap A_{i-1,n}^c;\ i = 2, 3, \ldots. \end{aligned}$$

Hence for each n $(A_{ni})_{i=1}^{\infty}$ is a partition of $\mathbb{B}$.

Define for $m, n \geq 1$ the mappings $\tau_n, \sigma_{mn}^{(1)}, \sigma_{mn}^{(2)}$ by:

$$\tau_n(b) = b_i \quad \text{if } b \in A_{in},\ i \geq 1;$$

$$\sigma_{mn}^{(1)}(b) = \begin{cases} b_i & \text{if } b \in A_{in},\ i \geq m+1, \\ 0 & b \in \bigcup_{i=1}^{m} A_{in}; \end{cases}$$

$$\sigma_{mn}^{(2)}(b) = \begin{cases} b_i & \text{if } b \in A_{in},\ 1 \leq i \leq m, \\ 0 & b \notin \bigcup_{i=1}^{m} A_{in}. \end{cases}$$

Hence $\tau_n(b) = \sigma_{mn}^{(1)}(b) + \sigma_{mn}^{(2)}(b)$ for each $m, n \geq 1$ and

$$\|b - \tau_n(b)\| \leq \frac{1}{n} \quad \forall b \in \mathbb{B}.$$

For the real-valued random variable $\|\tau_n(X_k)\|$ with fixed k and n we have

$$\mathbb{E}(\psi(\|\tau_n(X_k)\|)) < \infty$$

and therefore likewise

$$\mathbb{E}(\psi(\|\sigma_{mn}^{(1)}(X_k)\|)) < \infty \quad \text{and} \quad \mathbb{E}(\psi(\|\sigma_{mn}^{(2)}(X_k)\|)) < \infty.$$

Since $\|\sigma_{mn}(.)\| : \mathbb{B} \to [0, \infty)$ is measurable, it follows by [10], p.170 and [11], Proposition 6.6, p.105, that the sequence $\left\{\|\sigma_{mn}^{(1)}(X_k)\|\right\}, k = 0, 1, \ldots$ is stationary and satisfies the same mixing condition as $\{X_n\}$. The assumed strong law for real-valued random variables therefore implies the almost sure (V)-convergence of $\left\{\|\sigma_{mn}^{(1)}(X_k)\|\right\}$ to

$$\xi_{mn}^{(1)} = \mathbb{E}(\|\sigma_{mn}^{(1)}(X_k)\|).$$

Since $\mathbb{E}(\|X\|) < \infty$ we furthermore get

$$\xi_{mn}^{(1)} = \mathbb{E}(\|\sigma_{mn}^{(1)}(X_k)\|) = \sum_{i \geq m+1} \|b_i\| \mathbb{P}(\omega \in \Omega : X(\omega) \in A_{in}) \to 0 \ (m \to \infty).$$

Consider now the sequence $\left\{\mathbf{1}_{\{b_i\}}(\sigma_{mn}^{(2)}(X_k))\right\}$ of $0-1$-valued random variables. Using again [10], p.170 and [11], Proposition 6.6, p.105, we see that this sequence is also stationary and satisfies the mixing condition. This sequence converges in the (V)-sense almost sure to

$$\xi_{mni} = \mathbb{P}(\omega \in \Omega, X(\omega) \in A_{in}), \text{ if } i \leq m, \quad \text{or to} \quad \xi_{mni} = 0 \text{ if } i \geq m+1.$$

Since we have

$$\sigma_{mn}^{(2)}(X_k) = \sum_{i=1}^{m} b_i \mathbf{1}_{\{b_i\}}(\sigma_{mn}^{(2)}(X_k))$$

it furthermore follows, that $\sigma^{(2)}_{mn}(X_k)$ is almost sure (V) summable for each pair $m, n \geq 1$ to

$$\xi^{(2)}_{mn} = \sum_{i=1}^{m} b_i \xi_{mni}.$$

Define

$$E_{mn} := \left\{ \omega \in \Omega : \|\sigma^{(1)}_{mn}(X_k(\omega))\|, k \geq 1 \text{ is (V)-summable to } \xi^{(1)}_{mn} \right\},$$

$$F_{mn} := \left\{ \omega \in \Omega : \sigma^{(2)}_{mn}(X_k(\omega)), k \geq 1 \text{ is (V)-summable to } \xi^{(2)}_{mn} \right\}$$

and $D := \bigcap_m \bigcap_n \bigcap_p \bigcap_q (F_{mn} \cap E_{pq})$. Observe that $I\!P(D) = 1$. For $\omega \in D$ we show that $X_k(\omega)$ is a Cauchy-sequence. For $\varepsilon > 0$ choose $N \geq 1$ such that $N \geq 8/\varepsilon$. Since $\xi^{(1)}_{mN}$ converges to 0, we can find a $M \geq 1$ such that $\xi^{(1)}_{MN} < \varepsilon/8$. Finally using the (V)-convergence of $(\|\sigma^{(1)}_{MN}(X_k)\|)$ and $(\sigma^{(2)}_{MN}(X_k))$ we can choose a $\lambda_0 > 0$, such that

$$\left| \sum_{k=0}^{\infty} c_k(\lambda) \|\sigma^{(1)}_{MN}(X_k(\omega))\| - \xi^{(1)}_{MN} \right| < \frac{\varepsilon}{8} \quad \text{if } \lambda \geq \lambda_0 \text{ and}$$

$$\left\| \sum_{k=0}^{\infty} c_k(\lambda) \sigma^{(2)}_{MN}(X_k(\omega)) - \xi^{(2)}_{MN} \right\| < \frac{\varepsilon}{8} \quad \text{if } \lambda \geq \lambda_0.$$

Using the triangle inequality we get

$$\begin{aligned} & \left\| \sum_{k=0}^{\infty} c_k(\lambda) X_k(\omega) - \xi^{(2)}_{MN} \right\| \\ \leq \quad & \left\| \sum_{k=0}^{\infty} c_k(\lambda) \left(X_k(\omega) - \tau_N(X_k(\omega)) \right) \right\| \\ + \quad & \left\| \sum_{k=0}^{\infty} c_k(\lambda) \left(\tau_N(X_k(\omega)) - \sigma^{(2)}_{MN}(X_k(\omega)) \right) \right\| - \xi^{(1)}_{MN} \\ + \quad & \xi^{(1)}_{MN} + \left\| \sum_{k=0}^{\infty} c_k(\lambda) \left(\sigma^{(2)}_{MN}(X_k(\omega)) - \xi^{(2)}_{MN} \right) \right\| \\ \leq \quad & \frac{1}{N} + \left\{ \sum_{k=0}^{\infty} c_k(\lambda) \left\| \sigma^{(1)}_{MN}(X_k(\omega)) \right\| - \xi^{(1)}_{MN} \right\} + \frac{\varepsilon}{8} + \frac{\varepsilon}{8} \leq \frac{\varepsilon}{2}. \end{aligned}$$

Hence if $\lambda_1, \lambda_2 \geq \lambda_0$ we get

$$\left\| \sum_{k=0}^{\infty} c_k(\lambda_1) X_k(\omega) - \sum_{k=0}^{\infty} c_k(\lambda_2) X_k(\omega) \right\| < \varepsilon$$

and $V_{X(\omega)}(\lambda)$ is a Cauchy-sequence in $I\!B$. Since $I\!B$ is complete (V)-summability of $\{X_n\}$ follows.

For simple random variables we immediately see from our proof, that the (V)-limit is the expected value. Using the approximating property of simple functions, see [16], pp.76-82, the same is true in the general case i.e. $X_n \to \mu = I\!E(X)$ (V) *a.s.*. □

Remark 2. For the special case of Borel summability of i.i.d. random variables the result has been proved in [9] Theorem 2.3.

4 Proof of the main results

Proof of $(M) \Rightarrow (A1)$:
W.l.o.g. we can assume that the random variables are centered at expectation. First assume that $\mathbb{P}(X \in \{b_1, b_2, \dots b_d\}) = 1$ with $d \in \mathbb{N}$. For $i = 1, 2, \dots d$ consider the real-valued random variables

$$Z_k^{(i)} := \mathbf{1}_{\{X_k = b_i\}} - \mathbb{P}(X_k = b_i), \quad S_n^{(i)} := \sum_{k=1}^{n} Z_k^{(i)}.$$

By [10], p.170 and [11], Proposition 6.6. p. 105, the sequence $\{Z_n^{(i)}\}$ is stationary and satisfies the above mixing condition. Furthermore, since $\mathbb{E}(\psi(|Z_k^{(i)}|)) < \infty$ and $\mathbb{E}(Z_k^{(i)}) = 0$ we can use the Baum-Katz-type law (see [22], Theorem) and get:

$$\sum_{n=1}^{\infty} n^{-1}(\psi(n+1) - \psi(n))\mathbb{P}\left(\max_{k \le n} |S_k^{(i)}| > \varepsilon n\right) < \infty \ \forall \varepsilon > 0.$$

Now since $\mathbb{E}(X_k) = 0$ and

$$X_k = \sum_{i=1}^{d} b_i \mathbf{1}_{\{X_k = b_i\}} = \sum_{i=1}^{d} b_i \mathbf{1}_{\{X_k = b_i\}} - \sum_{i=1}^{d} b_i \mathbb{P}(X_k = b_i) = \sum_{i=1}^{d} Z_k^{(i)} b_i,$$

it follows that

$$\|S_n\| = \left\| \sum_{k=1}^{n} X_k \right\| = \left\| \sum_{j=1}^{n} \sum_{k=1}^{d} Z_k^{(i)} b_i \right\| = \left\| \sum_{i=1}^{d} S_n^{(i)} b_i \right\| \le d \max_{1 \le i \le d} \|b_i\| \, |S_n^{(i)}|.$$

Therefore

$$\|S_n\| \le C_d \max_{1 \le i \le d} |S_n^{(i)}|.$$

This implies the following upper bound

$$\mathbb{P}\left(\max_{k \le n} \|S_k\| > \varepsilon n\right) \le \mathbb{P}\left(\max_{k \le n} \max_{1 \le i \le d} |S_k^{(i)}| > \frac{\varepsilon}{C_d} n\right) \le \sum_{i=1}^{d} \mathbb{P}\left(\max_{k \le n} |S_k^{(i)}| > \frac{\varepsilon}{C_d} n\right)$$

and the claim is proved using [22] in the finite dimensional case.
Now assume that $\mathbb{P}(X \in \{b_1, b_2, \dots\}) = 1$.
Since $\mathbb{E}(\psi(\|X\|)) < \infty$ and hence $\mathbb{E}(\|X\|) < \infty$, we find for $\varepsilon > 0$ a $d \in \mathbb{N}$ with

$$\sum_{\nu=d+1}^{\infty} \|b_\nu\| \mathbb{P}(X = b_\nu) < \varepsilon.$$

Consider the set $A := \{b_1, b_2, \dots b_d\}$ and define random variables

$$X_k' := X_k \mathbf{1}_A(X_k) - \mathbb{E}(X_k \mathbf{1}_A(X_k)), \ X_k'' := X_k - X_k'$$

with partial sums S_n' and S_n''. Now the X_k' assume only finitely many values and the sequence $\{X_n'\}$ is stationary and satisfies the mixing condition. Furthermore $\mathbb{E}(X') = 0$ and $\mathbb{E}(\psi(\|X'\|)) < \infty$. Using the first part of the proof it follows that

$$\sum_{n=1}^{\infty} n^{-1}(\psi(n+1) - \psi(n))\mathbb{P}\left(\max_{j \le n} \|S_k'\| > \varepsilon n\right) < \infty \ \forall \varepsilon > 0.$$

Furthermore $\{\|X_n''\|\}$ is a stationary φ-mixing sequence of real-valued random variables with $\mathbb{E}(\|X_n''\|) < \varepsilon$ and $\mathbb{E}(\psi(\|X_n''\|)) < \infty$. Hence

$$\begin{aligned} & \sum_{n=1}^{\infty} n^{-1}(\psi(n+1)-\psi(n))\mathbb{P}\left(\max_{k\le n}\|S_k''\| > 2\varepsilon n\right) \\ \le\; & \sum_{n=1}^{\infty} n^{-1}(\psi(n+1)-\psi(n))\mathbb{P}\left(\sum_{k=1}^{n}\|X_k''\| > 2\varepsilon n\right) \\ \le\; & \sum_{n=1}^{\infty} n^{-1}(\psi(n+1)-\psi(n))\mathbb{P}\left(\sum_{k=1}^{n}(\|X_k''\| - \mathbb{E}(\|X_k''\|) > \varepsilon n\right) < \infty, \end{aligned}$$

using again the Baum-Katz-type law in [22] for the partial sums $\tilde{S}_n = \sum_{k=1}^{n}(\|X_k''\| - \mathbb{E}(\|X_k''\|))$.
Since $X_k = X_k' + X_k''$ we have $S_n = S_n' + S_n''$ and the claim follows.
In the general case there exists a countable, dense subset $\{b_1, b_2, \ldots\}$ of $\mathbb{B}$. Given $\varepsilon > 0$ we define for $\nu = 1, 2, \ldots$

$$A_\nu := \{b \in \mathbb{B} : \|b - b_\nu\| < \varepsilon\}$$

and $B_1 := A_1, B_\nu := A_\nu \cap A_1^c \cap \ldots \cap A_{\nu-1}^c$. Hence $\mathbb{B}$ is the countable union of the disjunct sets B_ν. Let

$$X_k' := \sum_{\nu=1}^{\infty} b_\nu \mathbf{1}_{B_\nu}(X_k) - \sum_{\nu=1}^{\infty} b_\nu \mathbb{P}(X_k \in B_\nu), \quad X_k'' := X_k - X_k'.$$

Now the X_k' assume countably many values and $\mathbb{E}(X_k') = 0$ and $\mathbb{E}(\psi(\|X_k'\|)) < \infty$, using $\psi \in BI$. Since $\{X_k'\}$ is also stationary and mixing it follows that

$$\sum_{n=1}^{\infty} n^{-1}(\psi(n+1)-\psi(n))\mathbb{P}\left(\max_{k\le n}\|S_k'\| > \varepsilon n\right) < \infty \;\forall \varepsilon > 0.$$

Now consider

$$\begin{aligned} \|X_k''\| &= \|X_k - X_k'\| \le \varepsilon + \left\|\sum_{\nu=1}^{\infty} b_\nu \mathbb{P}(X_k \in B_\nu)\right\| \\ &\le \varepsilon + \sum_{\nu=1}^{\infty}\int_{B_\nu}\|X_k - b_\nu\| d\mathbb{P} < 2\varepsilon, \end{aligned}$$

using $\mathbb{E}(X) = 0$ (B). Hence

$$\max_{k\le n}\|S_k''\| \le \sum_{k=1}^{n}\|X_k''\| \le 2n\varepsilon,$$

and the claim follows.
The implication $(A1) \Rightarrow (A2)$ is trivial.
Proof of $(A2) \Rightarrow (S1)$:
This follows by repeating line by line the corresponding real-valued case proof in [22], see also the proof of Theorem 3, p.449 in [6], with the only modification is using $\|.\|$ for $|.|$.

The implications $(M) \Rightarrow (S2)$, $(M) \Rightarrow (S3)$, $(M) \Rightarrow (S4)$ follow directly from our Proposition and the corresponding real-valued result in [22]
To prove the reversed implication $(Si) \Rightarrow (M), i = 1, 2, 3, 4$ we use the following

Lemma 3. *Assume that (V) is a summability method with weights which have the following localization property*

$$(L) \qquad \sup_{\lambda \in [0,\infty)} c_n(\lambda) = c_n(\lambda_n) \ \ \forall n = 0, 1, 2, \ldots,$$

with a non-decreasing sequence $\lambda_n \to \infty$ $(n \to \infty)$. Let $\{X_n\}$ be a stationary, φ-mixing $\mathbb{B}$-valued random variables and $\varphi_1 < 1/4$ and $\{X_n^s\}$ the symmetrized sequence. Then

$$X_n^s \to 0 \ (V) \ a.s. \quad \Rightarrow \quad \sum_{n=1}^{\infty} \mathbb{P}\left(c_n(\lambda_n)\|X_n^s\| > \varepsilon\right) < \infty \ \forall \varepsilon > 0.$$

Proof:
Define

$$Y_m := \sum_{n=0}^{m} c_n(\lambda_m) X_n^s \quad \text{and} \quad Z_m := \sum_{n=m+1}^{\infty} c_n(\lambda_m) X_n^s.$$

We have $Y_m + Z_m \to 0 \quad a.s.$. Using Lemma 2 we get for $\varepsilon > 0$

$$\mathbb{P}(\|Z_m\| > \varepsilon) \leq (\frac{1}{2} - 2\varphi_1)^{-1} \mathbb{P}(\|Y_m + Z_m\| > \varepsilon) \to 0 \quad (m \to \infty).$$

Hence by Lemma 1

$$Y_m \to 0 \ a.s. \quad (m \to \infty).$$

We repeat the argument for Y_{m-1} and $c_m(\lambda_m)X_m^s$ and get

$$c_m(\lambda_m)X_m^s \to 0 \ a.s. \ (m \to \infty).$$

By the Borel-Cantelli Lemma for φ-mixing sequences of random variables, see [17], Proposition 1.1.3 we get

$$\sum_{n=1}^{\infty} \mathbb{P}\left(c_n(\lambda_n)\|X_n^s\| > \varepsilon\right) < \infty \ \ \forall \varepsilon > 0.$$

□

Returning to the proof of $(Si) \Rightarrow (M), i = 1, 2, 3, 4$ we observe, that for the summability methods $(M_\phi), (V_\phi), (N, p^{*\kappa}, p), (P)$ generated by a weight sequence (p_n), satisfying (1.2), the condition (L) holds true with $c_n(\lambda_n) \sim 1/\phi(\lambda_n)$ and $\phi(t) = 1/\sqrt{g''(t)}$ [23, 30].
Let $\{X_n'\}$ be an independent copy of $\{X_n\}$ and $\{X^s\}$ be the symmetrized sequence with $X_n^s = X_n - X_n'$, $\forall n$ Then $X_n^s \to 0 \ (V) \ a.s.$.
Hence using Lemma 3 and the above asymptotic we get

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\|X_n^s\| > \varepsilon\phi(n)\right) < \infty, \ \forall \varepsilon > 0$$

and therefore

$$\mathbb{E}(\psi(\|X^s\|)) < \infty.$$

We denote by $m = med(\|X\|)$ the median. Using the inequality $I\!P(\|X\| > t+m) \leq 2I\!P(\|X^s\| > t)$ $\forall t > 0$ from [26], p.150. and Tonelli's Theorem we get

$$I\!E(\psi(\|X\|)) = \int_0^\infty \psi'(t) I\!P(\|X\| > t)dt \leq m + C I\!E(\psi(\|X^s\|)) < \infty,$$

where also $\psi' \in BI$ is used in the inequality. That $I\!E(X) = 0$ (B) follows from the first part of the Theorem. □

Remark 3. We outline a direct proof of $(A2) \Rightarrow (M)$, see [21].
Consider again the symmetrized sequence $\{X_n^s\}$ with $X_n^s = X_n - X_n'$ and partial sums $S_n^s = \sum_{k=1}^n X_k^s$. Since $\{\|S_n^s\| > \varepsilon n\} \subseteq \left\{\|S_n\| > \frac{\varepsilon}{2} n\right\} \cup \left\{\|S_n'\| > \frac{\varepsilon}{2} n\right\}$ we get the inequality $I\!P\left(\|S_n^s\| > \varepsilon n\right) \leq 2 I\!P\left(\|S_n\| > \frac{\varepsilon}{2} n\right)$. Hence $(A2)$ is also true for the symmetrized sequence. By Lemma 2 $(A1)$ follows, that is

$$\sum_{n=1}^\infty n^{-1}(\psi(n+1) - \psi(n)) I\!P\left(\max_{k \leq n} \|S_k^s\| > \varepsilon n\right) < \infty \ \forall \varepsilon > 0.$$

From $\{\max_{1 \leq k \leq n} \|X_k^s\| > 2\varepsilon n\} \subseteq \{\max_{1 \leq k \leq n} \|S_k^s\| > \varepsilon n\}$ this implies

$$\sum_{n=1}^\infty n^{-1}(\psi(n+1) - \psi(n)) I\!P\left(\max_{k \leq n} \|X_k^s\| > \varepsilon n\right) < \infty \ \forall \varepsilon > 0.$$

Using the method of associated random variable (see [21] Lemma 3.2.2 or [28], proof of Theorem 1) we can assume, that the X_n^s are mutually independent and

$$I\!E(\psi(\|X^s\|)) < \infty$$

follows from [25], Theorem 1. Now (M) follows as in the first proof. □

Remark 4. For i.i.d. real-valued random variables our Theorem above is complemented by an Erdős-Rényi-Shepp type law, see [23], which is proved by using a result on large deviations of the above convergence. In the case of i.i.d. random variables taking values in a Banach space such a result is only known for the (C_1)-method resp. (M_ϕ)-method with $\phi(t) = t$, see [1, 2]. Hence in the case of our general summability methods this interesting question remains subject to further research.

Acknowledgement:
I would like to express my gratitude to two referees for pointing out some gaps in the previous form of this paper.

References

[1] de Acosta, A.: On large deviations of sums of independent random vectors, *In: Beck A. et al: Probability in Banach Spaces 5,* Lecture Notes in Mathematics **1153**, (1984) 1–14, Springer Verlag.

[2] de Acosta, A.; Kuelbs, J.: Limit Theorems for Moving Averages of independent Random Vectors, *Z.W-theorie verw.Geb.* **64** (1983) 67-123.

[3] Bingham, N.H.: Summability methods and dependent strong laws, *In: Eberlein, E.; Taqqu, M.S.: Dependence in Probability and Statistics*, (1986) 291-300, Birkhäuser.

[4] Bingham, N.H.: Extensions of the strong law, in Analytic and Geometric Stochastics: *Adv. Appl. Probab. Supplement* (1986) 27-36.

[5] Bingham, N.H.: Moving averages, *In: Edgar, G.A.; Sucheston, L.: Almost Everywhere Convergence* Proc. Conf. Columbus OH, (1988) 131–145, New York: Academic Press.

[6] Bingham, N.H.; Goldie, C.M.: Riesz Means and Self-Neglecting Functions, *Math. Z.* **199** (1988) 443-454.

[7] Bingham, N.H.; Goldie, C.M.; Teugels, J.L.: *Regular Variation,* (1987 Cambridge University Press.

[8] Borwein, D.; Kratz, W.: An O–Tauberian theorem and a High Indices theorem for power series methods of summability, *Camb. Phil. Math. Soc.* **115** (1994) 365–375.

[9] Bozorgnia, A.; Bhaskara, R.: On summability methods and limit theorems for Banach space random variables, *Bull.Inst.Math.Acad. Sinica* **7** (1979) 1–6.

[10] Bradley, R.: Basic properties of strong mixing conditions, *In: Eberlein, E.; Taqqu, M.S.: Dependence in Probability and Statistics*, (1986) 165–192, Birkhäuser.

[11] Breiman, L.: *Probability,* Classics in applied mathematics **7** (1992) SIAM.

[12] Chow, Y.S.: Delayed sums and Borel summability of independent, identically distributed random variables. *Bull.Inst.Math.Acad. Sinica* **1** (1973) 207–220.

[13] Hardy, G.H.: *Divergent Series,* (1949) Oxford Press.

[14] Heinkel, B.: On the almost sure summability of $\mathbb{B}$-valued random variables, *In: Dudley R.M. et al: Probability in Banach Spaces* **8**, (1991) 325–338, Birkhäuser.

[15] Heinkel, B.: On Valiron means of $\mathbb{B}$-valued random variables, *Lith. Math. Journal* **32** (2) (1992) 162-173.

[16] Hille, E.; Phillips, R.S.: *Functional Analysis and Semi-Groups* (1968) AMS.

[17] Iosifescu, M; Theodorescu, R: *Random Processes and Learning* (1969) Springer Verlag.

[18] Jain, N.C.: Tail probabilities for sums of independent Banach space random variables, *Z.Wahrscheinlichkeitstheorie verw. Gebiete* **33** (1975) 155-166.

[19] Kiesel, R.: Power series methods and almost sure convergence *Math.Proc.Camb.Phil.-Soc.* **113** (1993) 195-204.

[20] Kiesel, R.: General Nörlund transforms and power series methods,*Math.Z.* **214** (1993) 273-286.

[21] Kiesel, R.: *Starke Gesetze für gewichtete Summen von Zufallsvariablen*, Habilitationsschrift, Universität Ulm (1995).

[22] Kiesel, R.: Summability and strong laws for φ-mixing random variables, *J. Theo. Prob.* (to appear)

[23] Kiesel, R.; Stadtmüller, U.: Erdős-Rényi-Shepp laws and weighted sums of independent identically distributed random variables, *J. Theo. Prob.*, **9**(4) (1996) 961-982.

[24] Lai, T.L.: Summability methods for independent identically distributed random variables, *Proc.Amer.Math.Soc.* **45** (1974) 253–261.

[25] Lai, T.L.: Convergence rates in the strong law of large numbers for random variables taking values in Banach spaces, *Bull.Inst.Math.Acad. Sinica* **2** (1) (1974) 67–85.

[26] Ledoux, N.; Talagrand, M.: *Probability in Banach spaces,* Ergebnisse der Mathematik und ihrer Grenzgebiete, 3.Folge, Band **23** (1991) Springer Verlag.

[27] Liebscher, E.: Strong convergence of sums of α-mixing random variables with applications to density estimation, *Stoch. Proc. Appl.* **65** (1996) 69–80.

[28] Peligrad, M.: The r-quick Version of the Strong Law for Stationary φ-mixing sequences, *In: Edgar, G.A.; Sucheston, L.: Almost Everywhere Convergence*, 335–348, Proc.Conf. Columbus OH, (1988). New York: Academic Press.

[29] Peyerimhoff, A.: *Lectures on Summability*, Lecture Notes in Mathematics, No.107, (1969) Springer Verlag.

[30] Stadtmüller, U.: On a family of summability methods and one-sided Tauberian conditions, *J. Math. Analysis Appl.* (to appear)

[31] van der Vaart, A.W.; Wellner J.A.: *Weak Convergence and Empirical Processes*, (1996) Springer Verlag.

[32] Zeller, K.;Beekmann, W.: *Theorie der Limitierungsverfahren*, (1970) Springer Verlag.