Yokohama Mathematical
Journal Vol. 62, 2016

DIFFERENCE EQUATIONS DEFINED BY SOME DEPENDENT RANDOM VECTORS

By

Kenichi Yoshihara

(Received April 22, 2015; Revised November 20, 2015)

Summary. Corresponding to the Black-Scholes equation and other SDE's Yoshihara (2012,2013) considered difference equations based on strong mixing random variables and proved that their solutions converge almost surely to the analogous solutions to those of the SDE's. In Takahashi-Kanagawa-Yoshihara (2015) their multidimensional versions are considered. In this paper, we show that the same results hold for other weakly dependent random vectors for which strong invariance principles hold. By the way we prove the strong invariance principle for weakly $\mathcal{M}$-dependent random vectors.

1. Introduction

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space. Let $\{\xi_i\}$ be a stationary sequence, with $E\xi_1 = 0$ and $E\xi_1^2 = 1$, satisfying the strong mixing condition, a kind of dependence conditions. Let $\{X(t); t \geq 0\}$ be a one-dimensional continuous process. Corresponding to the Black-Scholes equation

$$dX(t) = X(t)(\nu dt + \sigma dW(t)), \quad X(0) = z > 0, \tag{1}$$

where ν and $\sigma > 0$ are constants and $\{W(t); t \geq 0\}$ is a standard Wiener process, Yoshihara (2012) considered difference equation

$$\begin{aligned} \Delta X(s_i) = X(s_i) - X(s_{i-1}) &= X(s_{i-1})\left\{\nu\frac{T}{n} + \sigma\sqrt{\frac{T}{n}}\xi_i\right\}, \\ X(0) &= z > 0 \end{aligned} \tag{2}$$

with $s_i = (iT)/n$ $(1 \leq i \leq n)$ and showed that the solution $X^{(n)}(T)$ of (2) converges almost surely to

$$X(t) = z\exp\left\{\left(\nu - \frac{\sigma^2}{2}\right)T + \sigma\gamma_1 W(T)\right\}. \tag{3}$$

2010 Mathematics Subject Classification:
Key words and phrases:

as $n \to \infty$ where $\gamma_1 > 0$ is a constant which comes from the property of the sequence $\{\xi_i\}$. We note here that if the $\{\xi_i\}$ is a sequence of i.i.d zero-mean random variables, with $E\xi_1^2 = 1$ and $E|\xi_1|^p < \infty$ $(p > 2)$, then $\gamma_1 = 1$ and (3) coinsides with the solution of (1).

Further, corresponding to the one-dimensional Itô formula with respect to the SDE

$$dX(t) = h(X(t), t)dt + v(X(t), t)dW(t), \tag{4}$$

Yoshihara (2013) obtained asymptotics of functions of solutions of difference equation

$$\Delta X(s_k) = h(X(s_{k-1}), s_{k-1})\frac{T}{n} + v(X(s_{k-1}), s_{k-1})\xi_k\sqrt{\frac{T}{n}} \tag{5}$$

and proved that the functions of solutions of (5) behave analogously to the functions of the solution of (4). Further, Takahasi-Kanagawa-Yosihara (2015) considered the asymptoics of multidimendional difference equations based on the strong mixing random vectors.

The crucial tools of the proofs are moment inequalities for maxima of partial sums and the stong invariance approximations of sums, as will be seen in the proof of Theorem 1 in Section 2. Hence, many other dependent sequences may be used instead of the strong mixing sequences.

In this paper, firstly, we prove the general results (Section 2), and then the strong invariance for weakly $\mathcal{M}$-dependent random vector sequence which is the multidimensional version of Berkes *et al* (Section 3). In Sections 4 and 5 the same problems of another types of weak dependency are considered.

In the following sections, we use "c" to denote some absolute constant which does not depend on i, j, k, n and may differ from line to line. Also, we use ϵ's to denote arbitrarily small positive numbers. For any vector $\mathbf{a} \in \mathsf{R}^d$, $|\mathbf{a}|$ denotes the Euclidean norm, $\mathbf{a}^\mathsf{T}$ denotes the transpose of the vector $\mathbf{a}$. For any random vector $\mathbf{X}$ we write $\|\mathbf{X}\|_p = \{E|\mathbf{X}|^p\}^{1/p}$ if the right-side exists and by $\mathbf{Cov}(\mathbf{X})$ denote the covariance matrix of $\mathbf{X}$. Further, we define the norm of any $d \times d$ matrix M by

$$|M| = \sup_{|\mathbf{x}| \leq 1} |M\mathbf{x}|, \quad \mathbf{x} \in \mathsf{R}^d.$$

2. The general case

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space. Let $d \geq 1$. Let $\{\xi_i\} = \{(\xi_{i,1}, \cdots, \xi_{i,d})\}$ be a stationary d-dimensional sequence of centered vectors. Let

$$\mathbf{S}_m = \sum_{j=1}^{m} \xi_j.$$

We consider the following assumption.

Assumption A. *Suppose* $E|\xi_1|^{p+\delta} < \infty$ *for* $p \geq 6$ *and* $0 < \delta < 1$.
(I) *For any* r $(2 \leq r < p)$ *and any sequence* $\{a_k\}$ *of real numbers such that* $|a_k| \leq K < \infty$ $(k \geq 1)$

$$E\left(\max_{1\leq m\leq n}\left|\sum_{k=}^{m} a_k \xi_k\right|\right)^r \leq CK^r\left(n\|\xi_i\|_{2+\delta}^{\frac{2}{2+\delta}}\right)^{\frac{r}{2}}, \tag{6}$$

where K *is some positive constant.*
(II) *Put*

$$\lim_{n\to\infty} \frac{1}{n}\mathbf{S}_n^{\mathsf{T}}\mathbf{S}_n = \mathbf{\Gamma} \quad a.s., \tag{7}$$

where $\mathbf{\Gamma} = (\gamma_{q,q'})$ *is the* $d \times d$ *matrix with*

$$\gamma_q = \gamma_{q,q} = E\xi_{1,q}^2 + 2\sum_{i=2}^{\infty} E\xi_{1,q}\xi_{i,q},$$

$$\gamma_{q,q'} = E\xi_{1,q}\xi_{i,q'} + \sum_{i=2}^{\infty} E\xi_{1,q}\xi_{i,q'} + \sum_{i=2}^{\infty} E\xi_{i,q}\xi_{1,q'}.$$

$\mathbf{\Gamma}$ *is positive definite.*
(III) *there exists an Wiener process with covariance matrix* $\mathbf{\Gamma}$ *such that*

$$\left|\sum_{j\leq t} \xi_j - \mathbf{W}(t)\right| = o(n^{\frac{1}{4}}). \tag{8}$$

If $E|\xi_1|^{6+\delta} < \infty$, then there exists a κ with $1/3 < \kappa < (4+\delta)/(2(6+\delta))$ such that

$$\frac{|\xi_1|}{n^{\frac{1}{2}}} = O(n^{-\kappa}) \quad a.s. \tag{9}$$

Define the $d \times d$ matrix $\mathbf{R} = (r_{q,q'})$ with $r_{q,q'} = E\xi_{1,q}\xi_{1,q'}$ $(q, q' = 1, \cdots, d)$. By the stationarity and the moment condition of $\{\xi_i\}$

$$\left|\frac{1}{n}\sum_{j=1}^{n} \xi_{j,q}\xi_{j,q'} - r_{q,q'}\right| = O\big(n^{-\frac{1}{6}+\epsilon_1}\big) \quad a.s. \quad (q, q' = 1, \cdots, d). \tag{10}$$

Here, $0 < \epsilon_1 < \frac{1}{6}$ is arbitrary.

Let $\mathcal{F}_a^b$ be the σ-algebra generated by $\xi_a, \cdots, \xi_b$ $(a \le b)$. Let $\{\mathcal{F}_t\}$ be a family of sub-σ-algebras defined by $\mathcal{F}_t = \mathcal{F}_{-\infty}^t = \cup_{a\le t}\mathcal{F}_{-\infty}^a$. Let $\{X(t); t \ge 0\}$ be a time-continuous process with filteration $\{\mathcal{F}_t\}$.

Let $T > 0$ be fixed. For any integer n put

$$s_k = \frac{kT}{n} \quad (1 \le k \le n), \quad s_0 = 0.$$

Firstly, we prove the following theorem.

Theorem 1. *Let $d \ge 1$. Let $T > 0$ be arbitrarily given. Let $h(t)$ and $v_i(t) > 0$ $(i = 1, \cdots, d)$ be continuous functions with bounded derivatives on $[0, T]$ and consider the d-dimensional vector function $\mathbf{v}(t) = (v_1(t), \cdots, v_d(t))$ on $[0, T]$. Suppose Assumption A is satisfied. Let $\{X(t); 0 \le t \le T\}$ be a time-continuous process. Suppose the difference equation*

$$\Delta X(s_k) = X(s_{k-1})\left\{h(s_{k-1})\frac{T}{n} + \sqrt{\frac{T}{n}}\mathbf{v}(s_{k-1})\xi_k^{\mathsf{T}}\right\}, \quad (1 \le k \le n) \qquad (11)$$
$$X(0) = z > 0$$

holds for all n and denote by $X^{(n)}(T)$ with $X^{(n)}(0) = z > 0$ the solution of the difference equation (11), i.e.,

$$\frac{X^{(n)}(T)}{z} = \exp\left\{\sum_{k=1}^{n}\log\left(1 + h(s_{k-1})\frac{T}{n} + \sqrt{\frac{T}{n}}\mathbf{v}(s_{k-1})\xi_k^{\mathsf{T}}\right)\right\}. \qquad (12)$$

Then, as $n \to \infty$, $X^{(n)}(T)/z$ converges almost surely to

$$\frac{X(T)}{z} = \exp\left\{\int_0^T h(t)dt - \frac{1}{2}\int_0^T \mathbf{v}(t)\mathbf{R}\mathbf{v}(t)^{\mathsf{T}}dt + \sum_{i=1}^{d}\int_0^T v_i(t)dW_i(t)\right\} \qquad (13)$$

where $\{\mathbf{W}(t); 0 \le t \le T\} = \{(W_1(t), \cdots, W_d(t)); 0 \le t \le T\}$ is a Wiener process with covariance matrix $\mathbf{\Gamma}$.

Proof. Theorem 1 may be proved by the same method used in the strong mixing case. (See, Yoshihara (2013) and Takahasi-Kanagawa-Yoshihara (2015).) But, we repeat here the method in the case $d = 1$

Let $1 \le k \le n$. Rewriting (11) and using the Taylor theorem

$$\log \frac{X(s_k)}{X(s_{k-1})} = \log\left(1 + h(s_{k-1})\frac{T}{n} + \sqrt{\frac{T}{n}}\mathbf{v}(s_{k-1})\xi_k\right)$$
$$= \left(h(s_{k-1})\frac{T}{n} + \sqrt{\frac{T}{n}}\mathbf{v}(s_{k-1})\xi_k\right)$$
$$-\frac{1}{2}\left(h(s_{k-1})\frac{T}{n} + \sqrt{\frac{T}{n}}\mathbf{v}(s_{k-1})\xi_k\right)^2 + R_k^{(n)}.$$

Here, $R_k^{(n)}$ is the remainder term, that is,

$$R_k^{(n)} := \frac{1}{3}\frac{\left(h(s_{k-1})\frac{T}{n} + \sqrt{\frac{T}{n}}\mathbf{v}(s_{k-1})\xi_k\right)^3}{\left(1 + \theta_{k-1}\left(h(s_{k-1})\frac{T}{n} + \sqrt{\frac{T}{n}}\mathbf{v}(s_{k-1})\xi_k\right)\right)^3}$$

and θ_{k-1} is a random variable such that $|\theta_{k-1}| < 1$. Since h and v are continuous functions with bounded derivatives, by (9) we have

$$|R_k^{(n)}| = O(n^{-3\kappa}) \quad a.s. \tag{14}$$

Now, we have

$$\log \frac{X(T)}{X(0)} = \sum_{k=1}^{n} \log \frac{X(s_k)}{X(s_{k-1})} \tag{15}$$
$$= \sum_{k=1}^{n} \log\left(1 + h(s_{k-1})\frac{T}{n} + \sqrt{\frac{T}{n}}v(s_{k-1})\xi_k\right)$$
$$= \sum_{k=1}^{n}\left\{\left(h(s_{k-1})\frac{T}{n} + \sqrt{\frac{T}{n}}v(s_{k-1})\xi_k\right)\right.$$
$$\left. -\frac{1}{2}\left(h(s_{k-1})\frac{T}{n} + \sqrt{\frac{T}{n}}v(s_{k-1})\xi_k\right)^2 + R_k^{(n)}\right\}$$

for all n.

Let $l_q = l_q(n)$ and $m_q = m_q(n)$ $(q = 1, 2)$ be integer-valued functions of n defined by

$$m_1 = [n^{\frac{2}{11}}], l_1 = [n/m_1] \quad \text{and} \quad m_2 = [n^{\frac{3}{16}}], l_2 = [n/m_2],$$

where $[a]$ denotes the integer part of a, and put

$$s_{i,j}^{(q)} = \frac{T}{l_q m_q}\{l_q(i-1) + j\} \quad (1 \le i \le m_q, 1 \le j \le l_q),$$
$$s_{0,0}^{(q)} = 0 \quad (q = 1, 2)$$

Then, we can rewrite (15) as follows:

$$\begin{aligned}
(16)\qquad \log\frac{X(T)}{X(0)} &= \sum_{k=1}^{n} h(s_{k-1})\frac{T}{n} \\
&+\Big\{\sqrt{\frac{T}{n}}\sum_{i=1}^{m_1}\sum_{j=1}^{l_1} v(s^{(1)}_{i-1,j-1})\xi_{l_1(i-1)+j} + \sum_{k=l_1m_1+1}^{n}\sqrt{\frac{T}{n}}v(s_{k-1})\xi_k\Big\} \\
&-\frac{1}{2}\Big\{\sum_{i=1}^{m_2}\sum_{j=1}^{l_2}\Big(h(s^{(2)}_{i-1,j-1})\frac{T}{n} + \sqrt{\frac{T}{n}}v(s^{(2)}_{i-1,j-1})\xi_{l_2(i-1)+j}\Big)^2 \\
&+\sum_{k=l_2m_2+1}^{n}\Big(h(s_{k-1})\frac{T}{n} + \sqrt{\frac{T}{n}}v(s_{k-1})\xi_k\Big)^2\Big\} + \sum_{k=1}^{n} R^{(n)}_k \\
&= U^{(n)}_0 + (U^{(n)}_1 + V^{(n)}_1) - \frac{1}{2}(U^{(n)}_2 + V^{(n)}_2) + U^{(n)}_3 \quad \text{(say)},
\end{aligned}$$

for all $n \geq 1$.

Since $h(t)$ is a continuous function with a bounded derivative, it is obvious that

$$(17)\qquad \left|U^{(n)}_0 - \int_0^T h(t)dt\right| \leq cn^{-1}.$$

By (14) we have

$$(18)\qquad |U^{(n)}_3| \leq c\sum_{k=1}^{n} n^{-3\kappa} = O(n^{-3\kappa+1}) = o\big(n^{-\frac{\delta}{2(6+\delta)}}\big) \quad a.s.$$

Next, we consider $U^{(n)}_1$ and $V^{(n)}_1$. Since $h(t)$ and $v(t)$ are continuous functions with bounded derivatives on $[0, T]$, we put

$$M = \max\Big\{\sup_{0\leq t\leq T}|h(t)|, \sup_{0\leq t\leq T}|v(t)|\Big\}.$$

Since $n - l_1m_1 \sim m_1 = [n^{\frac{2}{11}}]$ and $\kappa > (1/3)$, by (9)

$$(19)\qquad |V^{(n)}_1| \leq \sqrt{\frac{T}{n}}M\sum_{k=l_1m_1+1}^{n}|\xi_i| = m_1O\big(n^{-\kappa}\big) = O\big(n^{-(\kappa-\frac{2}{11})}\big)$$

Now, we consider

$$\begin{aligned}
(20)\qquad U^{(n)}_1 &= \sum_{i=1}^{m_1}\sqrt{\frac{T}{n}}v(s^{(1)}_{i-1,0})\sum_{j=1}^{l_1}\xi_{l_1(i-1)+j} \\
&+\sum_{i=1}^{m_1}\sum_{j=1}^{l_1}\sqrt{\frac{T}{n}}(v(s^{(1)}_{i-1,j}) - v(s^{(1)}_{i-1,0}))\xi_{l_1(i-1)+j} \\
&= U^{(n)}_{1,1} + U^{(n)}_{1,2} \quad \text{(say.)}
\end{aligned}$$

By Assumption A (III) there is a Wiener process W with $EW^2(t) = \gamma_1 t$ such that

$$\Big|\sum_{j=1}^{l_1} \xi_{l_1(i-1)+j} - \{W(il_1) - W((i-1)l_1)\}\Big| = O\big(l_1^{\frac{1}{4}}\big) = O\big(n^{\frac{9}{44}}\big) \tag{21}$$

Hence, noting $l_1 m_1 \sim n$, we have

$$\begin{aligned}
&\sqrt{\frac{1}{l_1 m_1}} \sum_{i=1}^{m_1} \sum_{j=1}^{l_1} v(s_{i-1,0}^{(1)}) \xi_{l_1(i-1)+j} \\
&= \sqrt{\frac{1}{l_1 m_1}} \sum_{i=1}^{m_1} v(s_{i-1,0}^{(1)}) \sum_{j=1}^{l_1} \xi_{l_1(i-1)+j} \\
&= \sqrt{\frac{1}{m_1 l_1}} \sum_{i=1}^{m_1} v(s_{i-1,0}^{(1)}) \big\{\{W(il_1) - W((i-1)l_1)\} + O\big(l_1^{\frac{1}{4}}\big)\big\} \\
&= \sum_{i=1}^{m_1} v(s_{i-1,0}^{(1)}) \Big(W\Big(\frac{i}{m_1}\Big) - W\Big(\frac{i-1}{m_1}\Big)\Big) + O\big(n^{-\frac{1}{44}}\big) \quad a.s.
\end{aligned}$$

Since $m_1 = [n^{\frac{2}{11}}]$ and $v(t)$ is a continuous function with bounded derivative, we have

$$U_{1,1}^{(n)} \to \int_0^T v(t) dW(t) \quad a.s. \tag{22}$$

Denote

$$M(m_1, l_1) := \max_{1 \le i \le l_1} \max_{1 \le j \le m} |v(s_{i-1,j-1}^{(1)}) - v(s_{i-1,0}^{(1)})| \le \frac{c}{m_1}, \tag{23}$$

We proceed to estimate $U_{1,2}^{(n)}$. By the Markov inequality and Assumption A (I)

$$\begin{aligned}
&P\big(|U_{1,2}^{(n)}| > n^{-\frac{1}{132}}\big) \\
&\le cn^{\frac{1}{22}} E\Big(\sum_{i=1}^{m_1} \sum_{j=1}^{l_1} \sqrt{\frac{T}{n}} (v(s_{i-1,j-1}^{(1)}) - v(s_{i-1,0}^{(1)}) \xi_{l_1(i-1)+j}\Big)^6 \\
&\le cn^{\frac{1}{22}} M^6(m_1, l_1) \le cn^{\frac{1}{22}} m_1^{-6} \le cn^{-\frac{12}{11} + \frac{1}{22}} \le cn^{-\frac{23}{22}}.
\end{aligned} \tag{24}$$

Hence,

$$|U_{1,2}^{(n)}| = O\big(n^{-\frac{1}{132}}\big) \quad a.s. \tag{25}$$

Now, we consider $U_2^{(n)}$ and $V_2^{(n)}$. By (9)

$$\begin{aligned}
V_2\ (n) = &\sum_{k=l_2 m_2 + 1}^{n} \Big(h(s_k)\frac{T}{n} + \sqrt{\frac{T}{n}} v(s_{k-1})\xi_k\Big)^2 \\
&\le cm_2 n^{-2\kappa} = o\big(n^{-\frac{23}{48}}\big) \le cn^{-\frac{1}{3}}
\end{aligned} \tag{26}$$

To estimate $U_2^{(n)}$, we rewrite $U_2^{(n)}$ as

$$
\begin{aligned}
(27)\qquad U_2^{(n)} &= \sum_{i=1}^{m_2}\sum_{j=1}^{l_2}\left(h(s_{i-1,j-1})\frac{T}{n}\right)^2 \\
&\quad +2\sum_{i=1}^{m_2}\sum_{j=1}^{l_2}\frac{T}{n}\sqrt{\frac{T}{n}}h(s_{i-1,j-1}^{(2)})v(s_{i,j-1}^{(2)})\xi_{l_2(i-1)+j} \\
&\quad +\sum_{i=1}^{m_2}\sum_{j=1}^{l_2}\left(\sqrt{\frac{T}{n}}v(s_{i-1,j-1}^{(2)})\xi_{l_2(i-1)+j}\right)^2 \\
&= U_{2,1}^{(n)}+2U_{2,2}^{(n)}+U_{2,3}^{(n)} \quad \text{(say)}.
\end{aligned}
$$

It is obvious that

$$
(28)\qquad |U_{2,1}^{(n)}| \le M^2\sum_{i=1}^{m_2}\sum_{j=1}^{l_2}\left(\frac{T}{n}\right)^2 \le cM^2 n\left(\frac{T^2}{n}\right)^2 \le cn^{-1}
$$

and by (9)

$$
(29)\qquad |U_{2,2}^{(n)}| \le cM\sqrt{\frac{T}{n}}\left(\frac{T}{n}\sum_{i=1}^{m_2}\sum_{j=1}^{l_2}|\xi_{l_2(i-1)+j}|\right) \le cn^{-\kappa} \quad a.s.
$$

It remains to consider $U_{2,3}^{(n)}$. Firstly, we consider

$$
\frac{T}{l_2m_2}\sum_{i=1}^{m_2}\sum_{j=1}^{l_2}v^2(s_{i-1,0}^{(2)})\xi_{l_2(i-1)+j}^2 = \frac{T}{m_2}\sum_{i=1}^{m_2}v^2\left(\frac{(i-1)T}{m_2}\right)\frac{1}{l_2}\sum_{j=1}^{l_2}\xi_{l_2(i-1)+j}^2.
$$

Since $E\xi_1^2=1$, by (10)

$$
\left|\frac{1}{l_2}\sum_{j=1}^{l_2}\xi_{l_2(i-1)+j}^2-1\right| \le \frac{c}{l_2^{\frac{1}{6}-\epsilon}} \le n^{-\frac{1}{8}} \quad a.s..
$$

for each $1\le i\le m_2$. Noting

$$
(30)\qquad \frac{T}{m_2}\sum_{i=1}^{m_2}v^2(s_{i,0}^{(2)}) = \frac{T}{m_2}\sum_{i=1}^{m_2}v^2\left(\frac{(i-1)T}{m_2}\right) \le M^2T,
$$

we have

$$
\begin{aligned}
(31)\qquad &\left|\frac{T}{l_2m_2}\sum_{i=1}^{m_2}\sum_{j=1}^{l_2}v^2(s_{i-1,0}^{(2)})\xi_{l_2(i-1)+j}^2-\frac{T}{m_2}\sum_{i=1}^{m_2}v^2\left(\frac{(i-1)T}{m_2}\right)\right| \\
&\le \frac{T}{m_2}\sum_{i=1}^{m_2}v^2\left(\frac{(i-1)T}{m_2}\right)\left|\frac{1}{l_2}\sum_{j=1}^{l_2}\xi_{l_2(i-1)+j}^2-1\right| \le cn^{-\frac{1}{8}} \quad a.s.
\end{aligned}
$$

Further, since v is a continuous function with bounded derivative, we have

$$\left|\frac{T}{m_2}\sum_{i=1}^{m_2} v^2\left(\frac{(i-1)T}{m_2}\right) - \int_0^T v^2(t)dt\right| \le \frac{cT}{m_2} \le cn^{-\frac{1}{8}}. \tag{32}$$

Hence from (31) and (32) we obtain

$$\left|\frac{T}{l_2 m_2}\sum_{i=1}^{m_2}\sum_{j=1}^{l_2} v^2(s^{(2)}_{i-1,0})\xi^2_{l_2(i-1)+j} - \int_0^T v^2(t)dt\right| \le cn^{-\frac{1}{8}} \quad a.s., \tag{33}$$

which, in turn, implies

$$\left|U^{(n)}_{2,3} - \int_0^T v^2(t)dt\right| \le cn^{-\frac{1}{8}} \quad a.s. \tag{34}$$

Combining (16)-(20), (22), (25)-(29) and (34), we have

$$\begin{aligned}
\log\frac{X(T)}{X(0)} &= \lim_{n\to\infty}\sum_{k=1}^{n}\log\left(1 + h(s_{k-1})\frac{T}{n} + \sqrt{\frac{T}{n}}v(s_{k-1})\xi_k\right) \\
&= \lim_{n\to\infty}\{U^{(n)}_0 + (U^{(n)}_1 + V^{(n)}_1) - \frac{1}{2}(U^{(n)}_2 + V^{(n)}_2) + U^{(n)}_3\} \\
&= \int_0^T h(t)dt - \frac{1}{2}\int_0^T v^2(t)dt + \gamma_1\int_0^T v(t)dW(t) \quad a.s.
\end{aligned}$$

and the proof is completed. □

Remark. Instead of Assumption (III), we can use the bound $O_{a.s.}(n^{\frac{1}{2}-\epsilon_0})$ ($\epsilon_0 > 0$ being sufficiently small) in (8). But, the proof is slightly complex, because we must choose m_1, and $l_1 = [n/m_1]$ so that for some positive constants ϵ_1 and ϵ_2

$$\sqrt{\frac{m_1}{l_1}}\, l_1^{(\frac{1}{2}-\epsilon_0)} = \sqrt{m_1}\, l_1^{-\epsilon_0} = o(n^{-\epsilon_1}),$$

and $|U_{1,2}| = o_{a.s.}(n^{-\epsilon_2})$.

Denote by $\mathcal{C}^a_*(A^b)$ the set of functions $A^b \to \mathbf{R}$ which possess continuous bounded partil derivatives up to order a. For $F(x_1,\cdots,x_r) \in \mathcal{C}^3_*(\mathsf{R}^d)$ write

$$\begin{aligned}
F_{x_q}(x_1,\cdots,x_r) &= \frac{\partial F(x_1,\cdots,x_r)}{\partial x_q}, \\
F_{x_q,x_{q'}}(x_1,\cdots,x_r) &= \frac{\partial^2 F(x_1,\cdots,x_r)}{\partial x_q \partial x_{q'}}, \quad \text{etc.}
\end{aligned}$$

Let $\mathbf{h} = (h_1, \cdots h_d) : \mathsf{R}^d \times [0,\infty) \to \mathsf{R}^d$ and $\mathbf{v} = (v_1, \cdots, v_d) : \mathsf{R}^d \times [0,\infty) \to \mathsf{R}^p \times [0,\infty)$ be component-wise $\mathcal{C}_*^1(\mathsf{R}^r \times [0,\infty)$ functions such that

$$\begin{aligned}
&v_q(\mathbf{x}, t) > 0 \quad (1 \le q \le d), \\
&\|\mathbf{h}(\mathbf{x}, t) - \mathbf{h}(\mathbf{y}, t)\| + \|\mathbf{v}(\mathbf{x}, t) - \mathbf{v}(\mathbf{y}, t)\| \le K\|\mathbf{x} - \mathbf{y}\|, \\
&\|\mathbf{h}(\mathbf{x}, t)\| + \|\mathbf{v}(\mathbf{x}, t)\| \le K(1 + \mathbf{x}\|.
\end{aligned}$$

Referring the proof of Theorem 1 in Yoshihara (2013) and those of theorems in Takahasi-Kanagawa-Yoshihara (2015) and using the above method of proof of Theorem 1 we can also prove the following theorem which corresponds to the Itô formula.

Theorem 2. *Let $T > 0$. Let $\{\xi_i\}$ be a sequence of p-dimentional centered random vectors. Suppose Assumption holds. Let $\mathbf{h}$ and $\mathbf{v}$ be functions defined above. Further, let $\{\mathbf{X}(t); 0 \le t \le T\}$ be a continuous process satisfying the difference equation*

$$\begin{aligned}
\Delta X_q(s_k) &= h_q(\mathbf{X}(s_{k-1}), s_{k-1})\frac{kT}{n} + v_q(\mathbf{X}(s_{k-1}), s_{k-1})\xi_k\sqrt{\frac{kT}{n}} \\
\mathbf{X}(0) &= \mathbf{x} \qquad\qquad (1 \le k \le n, 1 \le q \le d).
\end{aligned} \tag{35}$$

Let $F(\mathbf{x}, t) = F(x_1, \cdots, x_d, t) : \mathsf{R}^d \times [0,\infty) \to \mathsf{R}$ be an element of $\mathcal{C}_*^3(\mathsf{R}^d \times [0,\infty))$ and consider the sum of difference

$$Z^{(n)}(T) = \sum_{k=1}^{n} \Delta Z(s_k) = \sum_{k=1}^{n} F(\mathbf{X}(s_k), s_k) - F(\mathbf{X}(s_{k-1}), s_{k-1}) \tag{36}$$

with $z = F(\mathbf{X}(0), 0)$.

Then, $Z^{(n)}(T)$ coverges almost surely to

$$\begin{aligned}
Z(T) = z &+ \sum_{q=1}^{d} \int_0^T F_{x_q}(\mathbf{X}(t), t) h_q(\mathbf{X}(t), t) + \int_0^T F(\mathbf{X}(t), t)dt \\
&+ \frac{1}{2}\sum_{q,q'=1}^{d} \int_0^T r_{q,q'} F_{x_q, x_{q'}}(\mathbf{X}(t), t) v_q(\mathbf{X}(t), t) v_{q'}^{\mathsf{T}}(\mathbf{X}(t), t)dt \\
&+ \sum_{q=1}^{d} \int_0^T F_{x_q}(\mathbf{X}(t), t) v_q(\mathbf{X}(t), t) d\mathbf{W}^{\mathsf{T}}(t)
\end{aligned} \tag{37}$$

as $n \to \infty$, where $\{\mathbf{W}(t); 0 \le t \le T\}$ is a p-dimensional Wiener process with covariance matrix $\mathbf{\Gamma}$.

The proof is omitted.

3. Weakly $\mathcal{M}$-dependent sequence and the strong invariance principle

Let $d \geq 1$ and $p \geq 1$. Let $\{\mathbf{Y}_k\} = \{(Y_{k,1} \cdots, Y_{k,d})\}$ be a stationary sequence of d-dimensional centered random vectors and $\delta(m) \to 0$ as $m \to \infty$. We say that $\{\mathbf{Y}_k\}$ is weakly $\mathcal{M}$-dependent in L^p with rate function $\delta(\cdot)$ if the following Condition is satisfied;

Condition. (A) *for any $k \in \mathsf{Z}$ and $m \geq 0$ one can find a d-dimensional random vector $\mathbf{Y}_k^{(m)} = (Y_{k,1}^{(m)}, \cdots Y_{k,d}^{(m)})$ with finite p-th moment such that*

$$E\{\max_{1\leq i\leq d} |Y_{k,i} - Y_{k,i}^{(m)}|^p\} \leq \delta^p(m); \tag{38}$$

(B) *For any disjoint intervals $I_1, \cdots, I_r$ $(r \geq 1)$ of integers and any positive integers $m_1, \cdots, m_r$ the vectors $\{\mathbf{Y}_j^{(m)}, j \in I_1\}, \cdots, \{\mathbf{Y}_j^{(m)}, j \in I_r\}$ are independent provided*

$$d(I_k, I_l) > \max(m_k, m_l) \quad (1 \leq k < l \leq r).$$

Here,

$$d(A, B) = \inf\{|a - b| : a \in A, b \in B\}$$

if A and B are subsets of Z.

Remark. Suppose $\{\mathbf{Y}_k\}$ is weakly $\mathcal{M}$-dependent in L^p with rate function $\delta(m)$. Then, $E|\mathbf{Y}_k|^p$ is finite, since for each $1 \leq i \leq d$

$$\|Y_{k,i}\|_p \leq \|Y_{k,i}^{(m)}\|_p + \|Y_{k,i} - Y_{k,i}^{(m)}\|_p \leq \|Y_{k,i}^{(m)}\|_p + \delta(m). \tag{39}$$

Further, if h is a Lipschitz function with Lipschitz constant K, then

$$\|h(\mathbf{Y}_{k,i}) - h(\mathbf{Y}_k^{(m)})\|_p \leq K\|\mathbf{Y}_k - \mathbf{Y}_k^{(m)}\|_p \leq K\delta(m)$$

and thus $\{h(\mathbf{Y}_k)\}$ is also weakly $\mathcal{M}$-dependent in L^p with rate function $K\delta(m)$.

In Berkes *et al* (2011) the following moment inequality is shown.

Lemma A. *Let $\{Y_k\}$ be a centered stationary sequence, weakly $\mathcal{M}$-dependent in L^p with rate function $\delta(m)$ satisfying*

$$D_p = \sum_{m=0}^{\infty} \delta(m) < \infty.$$

Then, the following inequalities hold:
(I) *If* $p \geq 2$*, for any* $n \in \mathsf{N}$*,* $b \in \mathsf{Z}$ *we have*

$$E\Big|\sum_{k=b+1}^{b+n} Y_k\Big|^p \leq C_p n^{\frac{p}{2}} \tag{40}$$

where C_p *is a constant depending on* p *and the sequence* $\{Y_k\}$*.*
(II) *If* $p > 2$*, for any* $2 < q \leq p$ *,* $n \in \mathsf{N}$*,* $b \in \mathsf{Z}$ *we have*

$$E\Big\{\max_{1\leq l\leq n}\Big|\sum_{k=b+1}^{b+l} Y_k\Big|^q\Big\} \leq C'_{p,q} n^{\frac{q}{2}} \tag{41}$$

where $C'_{p,q}$ *is a constant depending on* p, q *and the sequence* $\{Y_k\}$*.*

Lemma A may be generalized as follows:

Lemma 1. *Suppose* $\{\mathbf{Y}_k\}$ *is a stationary sequence of* d*-dimensional centered random vectors which is weakly* $\mathcal{M}$*-dependent in* L^p *with rate function* $\delta(m)$ *satisfying*

$$D_p = \sum_{m=0}^{\infty} \delta(m) < \infty. \tag{42}$$

Then, the following hold:
(I)

$$\|\mathbf{Y}_k - \mathbf{Y}_k^{(m)}\|_p \leq C d^{\frac{1}{2}} \delta(m). \tag{43}$$

(II) *If* $p \geq 2$*, for any any* $n \in \mathsf{N}$*,* $b \in \mathsf{Z}$ *we have*

$$E\Big|\sum_{k=b+1}^{b+n} \mathbf{Y}_k\Big|^p \leq C_p d^{\frac{1}{2}} n^{\frac{p}{2}} \tag{44}$$

and

$$E\Big|\sum_{k=b+1}^{b+n} \mathbf{Y}_k^{(m)}\Big|^p \leq C_{1,p} d^{\frac{p}{2}} n^{\frac{p}{2}} + C_{2,p} n^p \delta^p(m) \tag{45}$$

where s are positive constants independent of m *and* n*.*
(III) *If* $p > 2$*, for any* $2 < q \leq p$ *,* $n \in \mathsf{N}$*,* $b \in \mathsf{Z}$ *we have*

$$E\Big\{\max_{1\leq l\leq n}\Big|\sum_{k=b+1}^{b+l} \mathbf{Y}_k\Big|\Big\}^q \leq C'_{p,q} n^{\frac{q}{2}} \tag{46}$$

Here, $C_p, C_{1,p}, C_{2,p}, C'_{p,q}$ *are positive constant dependining only on* p, q, d *and the sequence* $\{\mathbf{Y}_k\}$*.*

Proof. From Condition (A) and the Minkowski inequality

$$
\begin{aligned}
\|\mathbf{Y}_k - \mathbf{Y}_k^{(m)}\|_p &= \Big\{ E\Big(\sqrt{\sum_{i=1}^{d} (Y_{k,i} - Y_{k,i}^{(m)})^2} \Big)^p \Big\}^{\frac{1}{p}} \\
&\leq \Big[\Big\{ E\Big(\sum_{i=1}^{d} (Y_{k,i} - Y_{k,i}^{(m)})^2 \Big)^{\frac{p}{2}} \Big\}^{\frac{2}{p}} \Big]^{\frac{1}{2}} \leq \Big[\sum_{i=1}^{d} \big(E\big|Y_{k,i} - Y_{k,i}^{(m)}\big|^p \big)^{\frac{2}{p}} \Big]^{\frac{1}{2}} \\
&= \Big[\sum_{i=1}^{d} \|Y_{k,i} - Y_{k,i}^{(m)}\|_p^2 \Big]^{\frac{1}{2}} \leq c d^{\frac{1}{2}} \delta(m).
\end{aligned}
$$

Hence, (43) is obtained.

Next, without loss of generality we assume $b = 0$. By the above method and (40) we have

$$
\begin{aligned}
\Big\| \sum_{k=1}^{n} \mathbf{Y}_k \Big\|_p &\leq \Big\{ E\Big(\sum_{i=1}^{d} \Big(\sum_{k=1}^{n} Y_{k,i} \Big)^2 \Big)^{\frac{p}{2}} \Big\}^{\frac{1}{p}} \\
&\leq \Big\{ \sum_{i=1}^{d} \Big(E\Big| \sum_{k=1}^{n} Y_{k,i} \Big|^p \Big)^{\frac{2}{p}} \Big\}^{\frac{1}{2}} \leq c\big\{ \big(d n^{\frac{p}{2}} \big)^{|frac2p} \big\}^{\frac{1}{2}} \leq c d^{\frac{1}{2}} n^{\frac{1}{2}},
\end{aligned}
$$

which, via (40), implies (45).

By the Minkowski inequality and (44)

$$
\begin{aligned}
E\Big| \sum_{k=b+1}^{b+n} \mathbf{Y}_k^{(m)} \Big|^p &\leq c\Big\{ E\Big| \sum_{k=b+1}^{b+n} \mathbf{Y}_k \Big|^p + E\Big| \sum_{k=b+1}^{b+n} \mathbf{Y}_k - \sum_{k=b+1}^{b+n} \mathbf{Y}_k^{(m)} \Big|^p \Big\} \\
&\leq c\Big\{ E\Big| \sum_{k=b+1}^{b+n} \mathbf{Y}_k \Big|^p + \Big(\sum_{k=b+1}^{b+n} \|\mathbf{Y}_k - \mathbf{Y}_k^{(m)}\|_p \Big)^p \Big\} \\
&\leq c d^{\frac{p}{2}} n^{\frac{p}{2}} + c n^p \delta^p(m).
\end{aligned}
$$

and (45) follows.

(46) is obtained from Lemma A (II) by the method of proof of (41). □

Remark. From the proof of Lemma 1 it easily follows that if $\{a_i\}$ is a sequence of real numbers such that $\sup_{i\geq 1} |a_i| \leq K < \infty$ and the condition of Lemma 1 holds, then

$$
E\Big| \sum_{i=b+1}^{b+n} a_i \mathbf{Y}_i \Big|^p \leq c_p d^{\frac{p}{2}} K^p n^{\frac{p}{2}} \tag{47}
$$

for all b.

Suppose $\{\mathbf{Y}_k\}$ is a stationary sequence of d-dimensional centered random vectors which is weakly $\mathcal{M}$-dependent in L^p withe rate function $\delta(\cdot)$. Put

$$\mathbf{S}_n = \sum_{k=1}^{n} \mathbf{Y}_k \quad \text{and} \quad \mathbf{Cov}(\mathbf{S}_n) = E\mathbf{S}_n\mathbf{S}_n^{\mathsf{T}}$$

and let $\mathbf{\Gamma} = (\gamma_{q,q'})$ be the $d \times d$ matrix such that for q, q' $(1 \le q, q' \le d)$

$$\gamma_q^2 = \gamma_{q,q'} = EY_{1,q}^2 + 2\sum_{i=2}^{\infty} EY_{1,q}Y_{i,q}, \tag{48}$$

$$\gamma_{q,q'} = EY_{1,q}Y_{1,q'} + \sum_{i=2}^{\infty}(EY_{1,q}Y_{i,q} + EY_{i,q}Y_{1,q'}). \tag{49}$$

The existence of $\mathbf{\Gamma}$ is guranteed by the following lemma.

Lemma 2. *Let $p > 2$. Let $\{\mathbf{Y}_k\}$ be a stationary sequence of d-dimensional centered random vectors which is weakly $\mathcal{M}$-dependent in L^p with rate function $\delta(\cdot)$ satisfying* (42). *Then, the series in* (48) *and* (49) *are absolutely convergent and hence $\mathbf{\Gamma}$ exists.*

Further,

$$\left|\mathbf{\Gamma} - \frac{1}{n}\mathbf{Cov}(\mathbf{S}_n)\right| \le c\left\{\frac{1}{n}\sum_{j=1}^{n-1} j\delta(j) + \sum_{j=n}^{\infty}\delta(j)\right\} \tag{50}$$

and consequently

$$\lim_{n\to\infty} \frac{1}{n}\mathbf{Cov}(\mathbf{S}_n) = \mathbf{\Gamma}. \tag{51}$$

Proof. Without loss of generality we assume that $E\mathbf{Y}_k^{(m)} = 0$ for all $k \in \mathsf{Z}$ and $m \ge 0$. Since we can use the same method, we consider only the case $\gamma_{1,2}$. We write

$$\begin{aligned} Y_{k,1}Y_{k+j,2} = &(Y_{k,1} - Y_{k,1}^{(j-1)})Y_{k+j,2} + Y_{k,1}^{(j-1)}(Y_{k+j,2} - Y_{k+j,2}^{(j-1)}) \\ &+ Y_{k,1}^{(j-1)}Y_{k+j,2}^{(j-1)}. \end{aligned}$$

Since by Condition (B) $EY_{k,1}^{(j-1)}Y_{k+j,2}^{(j-1)} = 0$, from Conditions (A) and (B) we have that for $j \ge 1$

$$\begin{aligned} |EY_{k,1}Y_{k+j,2}| &\le |E(Y_{k,1} - Y_{k,1}^{(j-1)})Y_{k+j,2}| + |EY_{k,1}^{(j-1)}(Y_{k+j,2} - Y_{k+j,2}^{(j-1)})| \\ &\le \|Y_{k,1} - Y_{k,1}^{(j-1)}\|_2\|Y_{k+j,2}\|_2 + \|Y_{k,1}^{(j-1)}\|_2\|Y_{k+j,2} - Y_{k+j,2}^{(j-1)}\|_2 \\ &\le (\|Y_{k+j,2}\|_2 + \|Y_{k,1}^{(j-1)}\|_2)\delta(j-1) \\ &\le (2\|Y_{1,1}^{(j-1)}\|_2 + D_2)\delta(j-1), \end{aligned}$$

which implies

$$\sum_{j=n}^{\infty} |Y_{k,1}Y_{k+j,2}| \le c \sum_{j=n}^{\infty} \delta(j) \to 0 \quad (n \to \infty). \tag{52}$$

Thus, the first part of Lemma 2 is obtained.

Next, since the (q, q')-component of $\mathbf{Cov}(\mathbf{S}_n)/n$ is

$$\frac{1}{n} E\Big(\sum_{i=1}^{n} Y_{i,q}\Big)\Big(\sum_{j=1}^{n} Y_{j,q'}\Big),$$

using the stationarity and the above method we can easily show (51). □

In the sequel, we alway assume that the matrix Γ is positive definite and denote by $\mathbf{I}$ the $d \times d$ identity matrix.

To prove Theorem 3 we need the following Theorem due to Götze and Zaitsev (2009).

Theorem A. *Suppose that $\xi_1, \cdots, \xi_n$ are independent R^d-valued random vectors with $E\xi_j = \mathbf{0}, j = 1, \cdots, n$. Let $p \ge 2$ and put*

$$M_p = \sum_{j=1}^{n} E|\xi_j|^p < \infty. \tag{53}$$

Let σ^2 be the maximal eigen value of $\mathbf{Cov}(\sum_{j=1}^n \xi_j)$. Assume that $\sigma \le C_1 M_p^{\frac{1}{p}}$ with some positive constant C_1. Then, for any construction on a probability space of a sequence of independent random vectors $\mathbf{X}_1, \cdots, \mathbf{X}_n$ and a corresponding sequence of independent Gaussian random vectors $\mathbf{Y}_1, \cdots, \mathbf{Y}_n$ such that $\mathcal{L}(\mathbf{X}_j) = \mathcal{L}(\xi_j)$, $E\mathbf{Y}_j = 0, \mathbf{Cov}(\mathbf{Y}_j) = \mathbf{Cov}(\mathbf{X}_j)$ $(j = 1, \cdots, n)$. For all $z > 0$

$$P\Big(\max_{1\le k\le n}\Big|\sum_{i=1}^{k}\mathbf{X}_i - \sum_{i=1}^{k}\mathbf{Y}_i\Big| > z\Big) \le C_2 d^{1+(p/2)} M_p z^{-p} \tag{54}$$

where $C_2 > 0$ is a constant depending only on p and C_1.

Remark. From Theorem A, we see that if $p > 6$, then

$$\max_{1\le k\le n}\Big|\sum_{i=1}^{k}\mathbf{X}_i - \sum_{i=1}^{k}\mathbf{Y}_i\Big| = o\big(n^{\frac{1}{6}}\big) \quad a.s. \tag{55}$$

The following theorem is a multi-dimensional version of the result due to Berkes *et al.* (2011)

Theorem 3. *Let $p > 2$. Let $\{\mathbf{Y}_k\}$ be a stationary sequence of d-dimensional centered random vectors which is weakly $\mathcal{M}$-dependent in L^p with rate function*

$$\delta(m) \le Cm^{-A} \tag{56}$$

where $\kappa > 0$ and

$$A > \frac{p-2}{2\kappa}\left(1 - \frac{1+\kappa}{p}\right) \vee 1, \quad \frac{1+\kappa}{p} < \frac{1}{2}. \tag{57}$$

Suppose $\mathbf{\Gamma}$ is positive definite.

Then, $\{\mathbf{Y}_k\}$ can be defined on a new probability space together with two d-dimensional Wiener processes with covariance $\mathbf{\Gamma}$, $\{\mathbf{W}_1(t); t \ge 0\}$ and $\{\mathbf{W}_2(t); t \ge 0\}$ such that

$$\sum_{k=1}^{n} \mathbf{Y}_k = \mathbf{W}_1(s_n) + C_2\mathbf{W}_2(t_n) + O\big(n^{\frac{1+\kappa}{p}}\big) \quad a.s. \tag{58}$$

where $\{s_n\}$ and $\{t_n\}$ are nondecreasing numerical sequences with

$$s_n \sim n, \quad t_n \sim C_1 n^{\kappa'}, \quad 0 < \kappa' < 1 \tag{59}$$

and C_1 and C_2 are positive constants.

Remark. $\mathbf{W}_1$ and $\mathbf{W}_2$ are not independent. But, as in Berkes *et al*, we can show that

$$\mathbf{Cov}(\mathbf{W}_1(s_n), \mathbf{W}_2(t_n)) \to 0 \quad (m, n \to \infty). \tag{60}$$

Proof. Since by assumption $\mathbf{\Gamma}$ is positive definite we can assume $\mathbf{\Gamma} = \mathbf{I}$.

Let us specify some constants that will be used for the proof. By assump tion on A it is possible to find a constant $0 < \epsilon_0 < \frac{1}{2}$ such that

$$A > \frac{p-2}{2\kappa(1-\epsilon_0)^2}\left(1 - \frac{1+\kappa}{p}\right) > 1. \tag{61}$$

Then, we set

$$\alpha = \frac{2\kappa(1-\epsilon_0)}{p-2(1+\kappa)}, \quad \beta = (1-\epsilon_0)\alpha, \quad \rho = \frac{\beta}{1+\alpha}. \tag{62}$$

For some $\epsilon_1 > 0$ (which will be specified later) we now define $m_k = [\epsilon_1 k^\rho]$.

The first step is to show that it is sufficient to provide the strong approximation for the perturbed sequence $\mathbf{Y}_k^{(m_k)}$. By Lemma 1 (I)

$$\|\mathbf{Y}_k - \mathbf{Y}_k^{(m_k)}\|_p \le ck^{-A\rho}. \tag{63}$$

If $A\rho < 1$, then

$$P\Big(\max_{2^n\le l\le 2^{n+1}}\Big|\sum_{j=1}^{l}(\mathbf{Y}_j-\mathbf{Y}_j^{(m_j)})\Big| > \frac{1}{n}\,2^{\frac{n}{p}(1+\kappa)}\Big)$$
$$\le P\Big(\sum_{j=1}^{2^{n+1}}|\mathbf{Y}_j-\mathbf{Y}_j^{(m_j)}| > \frac{1}{n}\,2^{\frac{n}{p}(1+\kappa)}\Big)$$
$$\le 2^{-n(1+\kappa)}n^p\Big(\sum_{j=1}^{2^{n+1}}\|\mathbf{Y}_j-\mathbf{Y}_j^{(m_j)}\|_p\Big)^p \le c2^{-C_1 n}n^p$$

where $C_1=(1+\kappa)-(1-A\rho)p>0$. Thus, by the Borel-Cantelli lemma we have

$$\sum_{j=1}^{l}\mathbf{Y}_j=\sum_{j=1}^{l}\mathbf{Y}_j^{(m_j)}+o\big(l^{\frac{1+\kappa}{p}}\big)\quad a.s.$$

If $A\rho\ge 1$ we get an (even better) error term of order $o(l^{(1/p)})$.

Next, we partition N into disjoint blocks as

$$\mathsf{N}=J_1\cup I_1\cup J_2\cup I_2\cup\cdots$$

where $|I_l|=[l^\alpha]$ and $|J_l|=[l^\beta|$ with α and β as in (62). Set $I_l=\{\underline{i}_l,\cdots,\bar{i}_l\}$ and $J_l=\{\underline{j}_l,\cdots,\bar{j}_l\}$ and put

$$\eta_l^{(1)}=\sum_{k\in I_l}\mathbf{Y}_k^{(m_k)}\quad\text{and}\quad \eta_l^{(2)}=\sum_{k\in J_l}\mathbf{Y}_k^{(m_k)} \tag{64}$$

Note that $\bar{i}_l=O(l^{1+\alpha})$ and if ϵ_1 in the definition of m_l is chosen small enough, then

$$|J_l|=[l^\beta]>[\epsilon_1\underline{i}_l^\rho]=m_{\underline{i}_1}.$$

Hence, by Condition (B) we have that $\{\eta_l\}$ and $\{\zeta_l\}$ each define a sequence of independent cetered random vectors.

By (43)

$$E\Big|\sum_{k\in I_l}\mathbf{Y}_k-\eta_l^{(1)}\Big|^p\le\Big\|\sum_{k\in I_l}(\mathbf{Y}_k-\mathbf{Y}_k^{(m_k)})\Big\|_p^p=O\big((|I_l|\cdot\underline{i}_l^{-A\rho(1+\alpha)})^p\big).$$

Further, by the restriction on the parameters A,ρ,α, and ϵ_0

$$|I_l|\cdot\underline{i}_l^{-A\rho(1+\alpha)}\le cl^\alpha l^{-A\rho(1+\alpha)}\le c\,l^{\frac{\alpha}{2}}\le c|I_l|^{\frac{1}{2}}. \tag{65}$$

Thus, we can find a constant C (which does not depend on l)

$$E\Big|\sum_{k\in I_l}\mathbf{Y}_k-\eta_l^{(1)}\Big|^p\le C|I_l|^{\frac{p}{2}}. \tag{66}$$

Similarly, we have

$$E\Big|\sum_{k\in J_l}\mathbf{Y}_k-\eta_l^{(2)}\Big|^p\le C|J_l|^{\frac{p}{2}}. \tag{67}$$

It is obvious that

$$L_l:=\sum_{k=1}^{l}|I_k|=O\big(l^{1+\alpha}\big).$$

Put

$$a_l=L_{2^n}^{\frac{1+\kappa}{p}}\quad(2^n\le l<2^{n+1};n=0,1,2,\cdots).$$

Then,

$$\begin{aligned}&\sum_{l=1}^{\infty}\frac{1}{a_l^2}E\Big|\sum_{k\in I_l}\mathbf{Y}_k-\eta_l^{(1)}\Big|^p\\&=\sum_{n=0}^{\infty}\sum_{l=2^n}^{2^{n+1}-1}\frac{1}{a_l^p}E\Big|\sum_{k\in I_l}\mathbf{Y}_k-\eta_l^{(1)}\Big|^p\\&\le\sum_{n=0}^{\infty}\frac{1}{L_{2^n}^{1+\kappa}}\sum_{l=1}^{2^{n+1}-1}\frac{1}{a_l^p}E\Big|\sum_{k\in I_l}\mathbf{Y}_k-\eta_l^{(1)}\Big|^p\\&\le c\sum_{n=0}^{\infty}2^{[-(1+\alpha)(1+\kappa)+(\alpha p/2)+1]n}.\end{aligned} \tag{68}$$

The exponent in the last line of (68) will be negative if $(1+\alpha)(1+\kappa)>(\alpha p/2)+1$. This is equivalent to $\alpha<2\kappa/(p-2(1+\kappa))$. Hence, by (66) we have

$$\sum_{l=1}^{\infty}P\bigg(\Big|\sum_{k\in I_l}\mathbf{Y}_k-\eta_l^{(1)}\Big|>a_l\bigg)\le\sum_{l=1}^{\infty}\frac{1}{a_l^p}E\Big|\sum_{k\in I_l}\mathbf{Y}_k-\eta_l^{(1)}\Big|^p<\infty. \tag{69}$$

which, via the Borel-Cantelli lemma, implies

$$\sum_{k\in I_l}\mathbf{Y}_k-\eta_l^{(1)}=O(a_l)\quad a.s. \tag{70}$$

By the same method, we can prove that

$$\sum_{k\in J_l}\mathbf{Y}_k - \eta_l^{(2)} = O(a_l) \quad a.s. \tag{71}$$

Next, put

$$\Sigma_{1,l} = \mathbf{Cov}(\eta_l^{(1)}) \quad \text{and} \quad \Sigma_{2,l} = \mathbf{Cov}(\eta_l^{(2)}) \quad (l = 1, 2, \cdots).$$

Since by Lemma 2

$$\frac{\Sigma_{1,l}}{|I_l|} \to \mathbf{\Gamma} = \mathbf{I} \quad (l \to \infty), \tag{72}$$

$\Sigma_{1,l}$ is positive definite and thus $\Sigma_{1,l}^{\frac{1}{2}}$ exists. Put

$$\zeta_{1,l}^{(1)} = \Sigma_{1,l}^{-\frac{1}{2}}\eta_l^{(1)} \quad (l = 1, 2, \cdots)$$

Then, by Theorem A and (70), we can construct a new probability space $(\Omega_0, \mathcal{F}_0, P_0)$ and two sequences of independent d-dimensional random vectors $\{\mathbf{Z}_{1,k}\}$, $\{\mathbf{Z}_{1,k}^*\}$,with $P_0 \circ \mathbf{Z}_{1,k} = P_0 \circ \zeta_k^{(1)}$, $P_0 \circ \mathbf{Z}_{1,k}^* = N(\mathbf{0}, |I_l|\mathbf{I})$, $(k \in \mathsf{N})$ such that

$$\sum_{k=1}^{2^n}\mathbf{Z}_{1,k} - \sum_{k=1}^{2^n}\mathbf{Z}_{1,k}^* = O(a_n) \quad a.s. \tag{73}$$

Similarly, Put

$$\zeta_{1,l}^{(2)} = \Sigma_{1,l}^{-\frac{1}{2}}\eta_l^{(2)} \quad (l = 1, 2, \cdots),$$

Then, by Theorem A and (71), we can construct a new probability space $(\Omega_0, \mathcal{F}_0, P_0)$ and two sequences of independent d-dimensional random vectors $\{\mathbf{Z}_{2,k}\}$, $\{\mathbf{Z}_{2,k}^*\}$ with $P_0 \circ \mathbf{Z}_{2,k} = P_0 \circ \zeta_k^{(2)}$, $P_0 \circ \mathbf{Z}_{2,k}^* = N(\mathbf{0}, |J_l|\mathbf{I})$, $(k \in \mathsf{N})$ such that

$$\sum_{k=1}^{2^n}\mathbf{Z}_{2,k} - \sum_{k=1}^{2^n}\mathbf{Z}_{2,k}^* = O(a_n) \quad a.s. \tag{74}$$

Next, we show that

$$|I_l|^{\frac{1}{2}}\zeta_l^{(1)} - \eta_l^{(1)} = O(a_l) \quad a.s, \tag{75}$$

Note that for any $l \geq 1$

$$|I_l|^{\frac{1}{2}}\zeta_l^{(1)} - \eta_l^{(1)} = (|I_l|^{\frac{1}{2}}\mathbf{I} - \Sigma_{1,l}^{\frac{1}{2}})\zeta_l^{(1)}. \tag{76}$$

Since $\|(I-M)^2\| \le \|I-M^2\|^2$ for any semi-positive definite matrix M, we have

$$\big\|\big(|I_l|^{\frac{1}{2}}\mathbf{I}-\Sigma_{1,l}^{\frac{1}{2}}\big)^2\big\| \le |I_l|^{-1}\||I_l|\mathbf{I}-\Sigma_{1,l}\|^2. \tag{77}$$

Hence, by (77)

$$\begin{aligned} E\||I_l|^{\frac{1}{2}}\zeta_l^{(1)}-\eta_l^{(1)}\|^p &\le E\big|(|I_l|^{\frac{1}{2}}\mathbf{I}-\Sigma_{1,l}^{\frac{1}{2}})\zeta_l^{(1)}\big|^p \\ &\le c\big\||I_l|^{\frac{1}{2}}\mathbf{I}-\Sigma_{1,l}^{\frac{1}{2}}\big\|^p \le c|I_l|^{-\frac{p}{2}}\big\||I_l|\mathbf{I}-\Sigma_{1,l}\big\|^p. \\ &\le |I_l|^{\frac{p}{2}}\left(\frac{\||I_l|\mathbf{I}-\Sigma_{1,l}\|}{|I_l|}\right)^p \le c|I_l|^{\frac{p}{2}} \end{aligned}$$

and

$$\begin{aligned} \|\Sigma_{1,l}\|^{\frac{1}{2}} &\ge |I_l|^{\frac{1}{2}}-|I_l|^{-\frac{1}{2}}\||I_l|^{\frac{1}{2}}\mathbf{I}-\Sigma_{1,l}^{\frac{1}{2}}\| \\ &\ge |I_l|^{\frac{1}{2}}-|I_l|^{\frac{1}{2}}\frac{\||I_l|\mathbf{I}-\Sigma_{1,l}\|}{|I_l|} \ge |I_l|^{\frac{1}{2}}-|I_l|^{\frac{1}{2}}\cdot o(1) \ge c|I_l|^{\frac{1}{2}}. \end{aligned}$$

Thus, by (45) we have

$$E|\zeta_l^{(1)}|^p = \frac{1}{\|\Sigma_{1,l}\|^{\frac{p}{2}}}\left|\sum_{k\in I_l}\mathbf{Y}_k^{|I_l|}\right|^p \le \frac{c}{|I_l|^{\frac{p}{2}}}|I_l|^{\frac{p}{2}}|I_l|^{\frac{p}{2}} \le c|I_l|^{\frac{p}{2}}. \tag{78}$$

As (68) we have

$$\sum_{l=1}^{\infty}\frac{1}{a_l^p}E\||I_l|^{\frac{1}{2}}\zeta_l^{(1)}-\eta_l^{(1)}\|^p \le c\sum_{n=0}^{\infty}2^{[-(1+\alpha)(1+\kappa)+(\alpha p/2)+1]n} < \infty,$$

which, via the Borel-Cantelli lemma, implies (75).

Now, for each k let

$$\sigma_l = 1(l\in I_k),\ = 0\ (l\in J_k); \quad \tau_l = 0(l\in I_k),\ = 1\ (l\in J_k)$$

and put

$$s_n = \sum_{l=1}^n \sigma_l, \quad t_n = \sum_{l=1}^n \tau_l.$$

Then, it is obvious that $s_n \sim n$ and $t_n \sim n^{\kappa'}$ ($\kappa' > 0$). By a Wiener process $\{\mathbf{W}_1(t)\}$ with covariance matrix $\mathbf{I}$ we can write

$$\sum_{k=1}^n \mathbf{Z}_{1,k}^n = \mathbf{W}_1(s_n) \quad a.s. \tag{79}$$

and similarly by a Wiener process $\{\mathbf{W}_2(t)\}$ with covariance matrix $\mathbf{I}$ we can write

$$\sum_{k=1}^{n} \mathbf{Z}_{2,k}^{n} = \mathbf{W}_2(s_n) \quad a.s. \tag{80}$$

Combining (68), (70), (73)-(75) we have (58) on an enlarged probability space.

To finish the proof, we have to show that the fluctuations of the partial sums and the Wiener processes $\{\mathbf{W}_1(t)\}$ and $\{\mathbf{W}_2(t)\}$ within the blocks are small enough. Since fluctuation properties of Wiener processes are easy to handle using standard deviation inequality, we only investigate the partial sums. The fact is shown, since by Lemma 2 (III) we have

$$\begin{aligned} P\Big(\max_{\underline{i}_k \le l \le \bar{i}_k} \Big| \sum_{j=\underline{i}_k}^{l} \mathbf{Y}_j > \underline{i}_k^{\frac{1+\kappa}{p}}\Big) &\le \underline{i}_k^{-(1+\kappa)} E\Big|\sum_{j=\underline{i}_k}^{\bar{i}_k} \mathbf{Y}_j\Big|^p \\ \le \underline{i}_k^{-(1+\kappa)} |I_k|^{\frac{p}{2}} \le ck^{-(1+\kappa)(1+\alpha)+(\alpha p/2)} &= O\big(k^{-(1+\epsilon_1)}\big) \end{aligned}$$

where $\epsilon_1 > 0$ is some number. □

Thus, if the conditions of Theorem 3 are satisfied, then the conclusion holds for any stationary sequence of weakly $\mathcal{M}$-dependent multi-dimensional random vectors. This implies that we can apply Theorems 1 and 2 to multi-dimensional weakly $\mathcal{M}$-dependent sequences.

4. Martingale generalizations

Eberline (1986) proved the strong invariance principles under the assumptions include various generalizations of martingales such as asymptotic martingales and mixingales.

Let $\{\xi_k\}$ be a stationary sequence of R^d-valued random vectors. Denote

$$\mathbf{S}_n(m) = \sum_{k=m+1}^{m+n} \xi_k \quad \text{and} \quad \mathbf{S}_n = \mathbf{S}_n(0).$$

Let $(\Gamma_n(m))_{q,q'}$ $(1 \le q, q' \le d)$ be the covariances defined by

$$\Gamma_n(m)_{q,q'} = n^{-1} E\{\mathbf{S}_n(m)_q \mathbf{S}_n(m)_{q'}\}$$

and write $(\Gamma)_{q,q'} = (\Gamma_n(0))_{q,q'}$. Denote the limit by $\mathbf{\Gamma} = (\Gamma_{q,q'})$ if exists.

We consider the following:

ASSUMPTION B. Let $\{\xi_k\}$ be a stationary sequence of R^d-valued random vectors with $E|\xi_1|^{6+\delta}$ $(0 \le d < 1)$.

(I) For all $m, n \geq 1$ there is a constant $C > 0$ such that

$$\|E\{\mathbf{S}_n(m)|\mathcal{F}_m\}\|_2 \leq C \tag{81}$$

(II) For each $e \in \mathsf{R}^d$ of length 1

$$\text{Var}(\langle e, \mathbf{S}_n\rangle) \geq r(n), \quad r(n) \to \infty \text{ as } n \to \infty, \tag{82}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product of vectors.
(III) There exists $\theta > 0$ such that for $1 \leq q, q' \leq d$.

$$\left\| E\{\mathbf{S}_n(m)_q \mathbf{S}_n(m)_{q'}|\mathcal{F}_m\} - E\{\mathbf{S}_n(m)_q \mathbf{S}_n(m)_{q'}\right\| \leq Cn^{1-\theta} \tag{83}$$

uniformly in m.

Remark. If Assumption B (II) holds, then $\mathbf{\Gamma}$ exists. Moreover, Assumption B is more strict conditions than those of Theorem 2 in Eberline (1986) and the strong invariance hold with bound $O_{a.s.}(n^{\frac{1}{2}-\rho})$ $(0 < \rho < \frac{1}{2})$. Hence, by Remark to Theorem 1, Theorems 1 and 2 may be applied to this case. Since Assumption B implies Assumption A, Theorems 1 and 2 may be applied.

5. Weak dependence defined by Berkes-Liu-Wu

By the new approximation method, Berkes-Liu-Wu (2014) proved the strong invariance principle for all $p > 2$ and for a large class of dependent sequences. In the sequel of this section, we use the result in the special case only.

For $k \in \mathsf{Z}$ define the shift process $\mathcal{E}_k = (\varepsilon_{l+k}, l \in \mathsf{Z})$. The central element of $\mathcal{E}_k$ (belonging to $l = 0$) is ε_k.

Let $\{\xi_k\}$ be a stationary sequence defined as

$$\xi_k = G(\mathcal{E}_k) = G(\cdots, \varepsilon_{k-1}, \varepsilon_k, \varepsilon_{k+1}, \cdots)$$

where $\{\varepsilon_k\}$ is an i.i.d.sequence and $G : \mathsf{R}^{\mathsf{Z}} \to \mathsf{R}$ is a measurable function. Let $\{\varepsilon'_j; j \in \mathsf{Z}\}$ be an i.i.d.copy of $\{\varepsilon_j; j \in \mathsf{Z}\}$ and for $i, j \in \mathsf{Z}$ let $\mathcal{E}_{k,\{j\}}$ denote the process obtained from $\mathcal{E}_k$ by replacing the coordinate ε_j by ε'_j. We assume $E|\xi_1|^6 < \infty$.

Put

$$\delta_{k,6} = \|\xi_k - \xi_{k,\{0\}}\|_6 \quad \text{where} \quad \xi_{k,\{0\}} = G(\mathcal{E}_{k,\{0\}}). \tag{84}$$

We assume $E|\xi_1|^6 < \infty$. The dependence condition is expresses by

$$\Theta_{k,p} = \sum_{|j| \geq k} \delta_{j,p} \quad (k \geq 0). \tag{85}$$

If

$$\Theta_{m,6} = m^{-2} \quad (m \geq 1)$$

then by Corollary 2.1 to Theorem 1 in Berkes-Liu-Wu (2014) for the stationary sequence $\{\xi_k\}$ of random variables the strong invariance principle holds with bound order $o_{a.s.}(n^{1/6})$. Hence, Theorems 1 and 2 may be applied to this sequence $\{\xi_k\}$.

Acknowledgement. We thank the anonymous referees for their careful reading.

References

[1] Berkes, I., Hörmann, S., Shauer, J. (2011). Split invariance principles for stationary processes. *Ann. Probab.* **39** 2441-2473.

[2] Berkes, I., Liu, W., Wu, W. B. (2014). Komlós-Major-Tusnády approximation under dependence. *Ann. Probab.* **42** 794-817.

[3] Eberlein, E. (1986). On strong invariance principles under dependence assumptions. *Ann. Probab.* **14** 260-270.

[4] Götze, F., Zaitsev, A. Yu. (2009). Bounds for the rate of strong approximation in the multidimensional invariance principle. *Theory Probab. Appl.* **53** 59-80.

[5] Lin, W., Lin, Z. (2009). Strong approximation for a class of stationary processes. *Stoch. Process. Appl.* **119** 249-280.

[6] Takahasi, H., Kanagawa, S., Yoshihara, K. (2015) Asymptotic behavior of solutions of some difference equations defined by weakly dependent random vectors. Stoch. Analy. Appli. **33**, 740-755.

[7] Takahasi, H., Saigo, T., Kanagawa, S. Yoshihara, K. (2015) . Optimal portfolio based on weakly dependent data. Dynamical Systems and Differential Equations, Proceedings of the 10th International Conference. (To appear).

[8] Yoshihara, K. (1992). *Weakly dependent stochastic sequences and their applications, Vol.I, Summation theorey for weakly dependent random variables.* Sanseido, Tokyo.

[9] Yoshihara, K. (2012). Asymptotic behavior of solutions of Blach-Sholes type equations based on weakly dependent random variables. *Yokohama Math. J.* **58** 1-16.

[10] Yoshihara, K. (2013). Difference equations of functions of processes based on weakly dependent data. *Yokohama Math. J.* **59** 99-119

[11] Yoshihara, K. (2014). U-statistics and U-processes based on weakly $\mathcal{M}$-dependent data. *Yokohama Math. J.* **60** 91-112.

[12] Yoshihara,K., Kanagawa, S. (2013). Approximation to optimal prices when basic data are weakly dependent. (Preprint.)